Spatiotemporal Dynamics of 3-D Genome Architecture and Gene Expression during Lactogenic Differentiation of Murine Mammary Epithelial Cells

Thesis submitted to University of Hyderabad
for the award of Ph.D. degree in
Department of Animal Biology

By

E V Trinadha Rao Sornapudi

11LAPH04

Department of Animal Biology
School of Life Sciences
University of Hyderabad
Hyderabad - 500 046
India

July 2018

CERTIFICATE

This is to certify that the thesis entitled **"Spatiotemporal Dynamics of 3-D Genome Architecture and Gene Expression during Lactogenic Differentiation of Murine Mammary Epithelial Cells"** submitted by **E V Trinadha Rao Sornapudi** bearing registration number **11LAPH04** in partial fulfilment of the requirements for award of Doctor of philosophy in the **School of Life Sciences** is a bonafide work carried out by him under my supervision and guidance.

This thesis is free from plagiarism and has not been submitted previously in part or in full to this or any other University or Institution for award of any degree or diploma.

A. Published in the following publications:

1. Sornapudi et al., 2018 BMC Res Notes. 2018; 11: 241 (PMID: 29642945)
2. Scientific Reports. SREP-17-54387A (Under Communication)

B. Presented in the following conferences:

1. NextGen Genomics, Biology, Bioinformatics and Technologies (NGBT) conference, 2015 (International).
2. Genomic Architecture and cell fate regulation, 2014 (International).

Further, the student has passed the following courses towards fulfillment of coursework requirement for Ph.D.

Course Code	Name	Credits	Pass/Fail
1. AS 801	Seminar 1	1	PASS
2. AS 802	Research Ethics & Management	2	PASS
3. AS 803	Biostatistics	2	PASS
4. AS 804	Analytical Techniques	3	PASS
5. AS 805	Lab Work	4	PASS

Supervisor **Head of Department** **Dean of School**

UNIVERSITY OF HYDERABAD

Central University (P.O.), Hyderabad-500046, INDIA

DECLARATION

I hereby declare that the results of the study incorporated in the thesis entitled **"Spatiotemporal Dynamics of 3-D Genome Architecture and Gene Expression during Lactogenic Differentiation of Murine Mammary Epithelial Cells"** has been carried out by me under the supervision of ***Dr.Sreenivasulu Kurukuti*** and this work has not been submitted for any degree or diploma of any other university earlier.

Dated

E V Trinadha Rao Sornapudi

(Research Scholar)

Acknowledgements

- I would like to express my deepest sense of gratitude to my supervisor ***Dr. Sreenivasulu Kurukuti***, for giving me an opportunity to work under his guidance, constant cooperation, and encouragement throughout my work. I am highly grateful to him for all the efforts he has put in for the successful completion of this thesis.
- I thank Head, Dept of Animal Biology, Prof. Anita Jagota, and previous HOD's Prof. Jagan Pongubala, Prof. Senthilkumaran, Prof. Manjula Sritharan and Prof. S Dayananda for the departmental facilities
- I thank the Dean, School of Life Sciences, Prof. Ramaiah and former Dean Prof. P Reddanna, Prof. A.S. Raghavendra, Prof Aparna Dutta Gupta, Prof. R.P. Sharma and Prof. Ramanadham for providing the central facilities at the School of Life Sciences
- I thank the doctoral committee members **Prof. Aparna Dutta Gupta** and **Dr.Durga Bhavani** for constructive criticism and guidance.
- I thank **Prof. Aparna Dutta Gupta** for constant support and encouragement during my PhD work.
- I thank **Dr. Anil kumar pasupuleti** for helping in scientific writing skills during manuscript preparation.
- I thank all my lab mates G. Prashanth Kumar, Srinivas Kethavath, Rakhee Nayak, Sukalpa Mondal, Netrika Tiwari, Yuvasri and previous lab members for their help and joyful environment during my stay in lab
- I thank Rakhee Nayak and G. Prashanth Kumar for helping in my PhD work.
- I thank all my friends Dr. Sravan, Dr. Ravinder, Dr.Suman Asalla, Dr.Rameswar Reddy, Dr.Suresh Kumar Challapareddy,Dr.Shanmuk, Dr.Maruthi, Dr.Suresh Chava.
- I thank my uncle Akula Satyanarayna and Dr. Venkat Rao Vantaku for their financial and moral support throughout my education.
- I thank Prashanth and Sravan for assistance in the laboratory.
- The financial support from DBT, NIAB, CSIR, ICMR, DST, UGC, PURSE and DBT-CREBB is highly acknowledged
- Financial support through CSIR-UGC JRF and SRF is highly acknowledged
- I am extremely grateful to my **parents** and my **sisters** for their unwavering love, support and patience throughout my life.
- Above the all, I owe my thanks to **Almighty,** for everything.

Trinadha Rao…

Table of Contents

Abbreviations Page. No

1. **Review of Literature**

1.1 Introduction **1 - 30**

1.1.1 Nuclear structure and functions

1.1.1.1 Nuclear pore complex

1.1.1.2 Functional Domains in the Cell Nucleus

1.1.1.2.1 Promyelocytic Leukemia (PML) Nuclear Bodies

1.1.1.2.2 Histone locus bodies

1.1.1.2.3 Nuclear speckles

1.1.1.2.4 Peri nucleolar Compartment

1.1.1.2.5 Nuclear Stress Bodies

1.1.1.2.6 Orphan nuclear bodies

1.1.1.2.7 Polycomb bodies

1.1.1.2.8 Nucleolus

1.1.1.2.9 Cajal bodies

1.1.1.2.10 Paraspeckles

1.1.1.2.11 Transcription Factories

1.1.2 Chromatin organization in 2-Dimension and in 3-Dimensions

1.1.2.1 Chromosome Territories

1.1.2.2 Chromosome Compartment

1.1.2.3 Topologically Associated Domains

1.1.2.4 Sub-Topologically Associated Domains

1.1.2.5 Architectural proteins and their role in 3D genome organization

1.1.2.6 LAD organization and role in 3D genome organization

1.1.3 Spatial organization of 3D genome during cellular division (Time and Space)

1.1.4 3D genome studies in context of differentiation (Time and Space) Model system and objectives

1.1.5 HC11 mammary epithelial lactogenic differentiation model system to study interplay between signalling and chromatin dynamics

1.1.6 Mammary Epithelial differentiation and the role of signaling hormones:

2: Materials and Methods **31 - 46**

3: Results **47 - 141**

4: Discussion **142 - 149**

5: Summary **150 - 154**

6. Tables **155 - 174**

References **175 -194**

Publications

Anti-plagiarism certificate

Abbreviations:

MEC: mammary epithelial cells

ESC: embryonic stem cells

GC: Glucocorticoids

PRL: Prolactin

GRE: Glucocorticoid response elements

GR: Glucocorticoid Receptor

cLAD: conserved lamina associated domains

ciLAD: conserved inter lamina associated domains

cfLAD: conserved facultative lamina associated domains

LRI: Long Range Interaction

PI: Propidium Iodide

HC: Hydrocortisone

FACS: Fluorescence Activated Cell Sorting

CTN: Chromosome Territory Neighbourhood

GE: Gene Expression

UCSC: University of California Santa Cruz

FPKM: Fragments Per Kilobase per Million

VC: Vennela Coverage normalization

PCA: Principle Component Analysis

dGE: Differential gene expression

d-PC1: subtraction of principle component 1

chr: chromosomes

LvS: Long-range vs Short-range ratio

FISH: Fluorescence *In Situ* Hybridization

FBS: Foetal Bovine Serum

Mb: Megabase

Kb: Kilobase

AU: arbitrary units

LRI: Long Range Interactions

DAPI: 4',6-Diamidino-2-Phenylindole, Dihydrochloride

T: Translocation

NT: Non-Translocation

PCA: Principle Component Analysis

Csn2: Casein gene

EGF: Epidermal Growth Factor

TAD: Topologically Associated Domain

Review of Literature

1.1 INTRODUCTION

1.1.1 Nuclear Structure and Function

Eukaryotic cell is composed of well-defined membrane bound nucleus and cytoplasm surrounded by plasma membrane. Nucleus is a double membrane bound organelle and is highly complex in its organization. It carries genetic material and is involved in numerous functions related to replication, transcription and ribosome biosynthesis etc. Nucleus is a major storage organelle for genetic material which is passed to the offspring during cell division. Non-membrane compartmentalization is the one of the key aspect of nucleus. It is now postulated that compartmentalization of the eukaryotic genome inside the nuclei is important for its proper packing, for facilitating local increase of important regulatory factors and reducing mutational events. Packing the genome inside cell nucleus requires a complex process of chromatin compaction. At the primary level, double standard DNA wraps around histones to form "beads on a string" like structures called nucleosomes. Each nucleosome is composed of histone octamer in duplicate of H2A, H2B, H3 and H4 and 146bp of DNA that is wrapped around histone octamer. Chain of nucleosomes further arrange themselves to form 30 nm fiber having packing ratio of 50. Then the 30nm fiber loops around central scaffolding proteins to give rise to euchromatin and further compaction leads to heterochromatin. There are various models postulated to describe secondary structure of chromatin in interphase nuclei. The most favourable models are solenoid model and two-start helical model (Woodcock et al., 1984, Dorigo et al., 2004). Solenoid model explains arrangement of nucleosomes in a helical manner of six to eight nucleosomes per turn and the linker H1 histone arranged inside the fibre. In this model, linker H1 bent to connect with each nucleosome in its surrounding in order to form helical fibre whereas in two start helical model, nucleosomes are arranged in zig-zag fashion and the linker H1 is straight or parallel to the fibre axis. This packing level is maintained by formation of loops, scaffolds and domains in the interphase nucleus and chromatin compaction in mitotic stage is maintained by important architectural proteins like condensins. In mitosis, condensed chromatin forms rod shaped structures and are highly compacted in nature. After segregation of sister chromatids to the daughter cells chromatin in interphase nuclei is again reconstituted. Earlier electron microscopic images of cell nucleus have shown that electron dense highly condensed heterochromatic regions are spatially segregated near to the nuclear periphery whereas dencondensed euchromatin is present at the interior of the cell nucleus (Peric-Hupkes et al., 2010). Epigenetic and transcriptomic studies have revealed that

highly condense regions are transcriptionally more silent and marked with H3K9me2 and dencondensed regions are transcriptionally more active and marked with H3K4me2.

1.1.1.1 Nuclear pore complex

Nucleus is a double membrane organelle which is composed of outer and inner layers. Outer layer is continuum with rough endoplasmic reticulum whereas inner layer is localized towards to the nucleoplasm and binds to various meshwork of proteins including lamins and nuclear pore complex proteins (NPC) (de Las Heras et al., 2013). NPC proteins play important role in export and import business between nucleoplasm and cytoplasm. In addition to this traditional role, NPCs are also involved in genome organization, maintenance of the genome and regulation of gene expression (Bukata et al., 2013, Ibarra and Hetzer, 2015). In genome maintenance process NPCs facilitate DNA replication, DNA damage and telomere maintenance. It was observed that chromatin tethering and transcriptional activity that take place at nuclear pore complex act as a barrier to DNA replication which results in increased topological stress at these sites during replication (Bermejo et al., 2011). In order to relieve topological stress. S-phase check point machinery Mec1 and Rad53 in yeast and ATR/Chk1 in vertebrates reduces the stress by detaching tethered chromatin from NPC proteins (Bermejo et al., 2011, Matsuoka et al., 2007, Blasius et al., 2011). After replication fork passage, the regions re-anchor back to the NPC and resume their transcriptional activity in G1 stage of cell cycle. In S-phase, regions that are transcriptionally active at periphery get detached and move to the interior of the cell nuclei. In G2 stage, the regions re-anchor to NPC and is mediated by phosphorylated Nurp1 (Brickner and Brickner, 2011). In general, the function of NPC, prevents the formation of DNA lesions during DNA damage. A series of studies on yeast and metazoans have shown that the mutation in Nups results in accumulation of DNA damage and cells become hypersensitive to the DNA damage agents, thus confirming the potential role of NCPs in DNA damage response. Another class of proteins Nups, are highly dynamic in NCPs and there is constant mobility of Nups into the core complex. Nup153, Nup98, or Nup50 are highly mobile and found to be connected with RNA pol-II mediated transcription in NCPs (Griffis et al., 2002, Buchwalter et al., 2014). Further ChIP, DamID and immune histochemical studies have confirmed that these mobile Nups are highly interactive with transcriptionally active chromatin regions in the nucleus (Capelson et al., 2010, Kalverda et al., 2010). This study also showed the importance of the Nups as depletion of the Nup98 results in dysregulation of hundreds of genes in nuclear interior and at NPC. Nups are also found to be involved in transcriptional silencing by binding to the chromatin regions. For example, Nup88 is

associated with transcriptionally inactive regions in drosophila salivary gland (Capelson et al., 2010), whereas Chip-on-ChIP studies on Hela cells found that Nup93 is enriched in repressive chromatin marked with H3K27me3 or H3K9me3. Nups are evolutionary conserved and participates in global gene regulation, independent of sub nuclear organization of transcriptional activity.

1.1.1.2 Functional Domains in the Cell Nucleus

Compartmentalization is one of the key properties associated with structural and functional aspects of nuclei. It is now believed that functional compartmentalization of the cell nucleus is interconnected with spatial genome organization. Although nucleus is a membrane bound organelle, however within the nucleus various functions are carried out by multiple sub-nuclear dynamic domains. It has recently been proposed that macromolecular crowding might be involved in the formation of intra nuclear compartments due to phase separation. Some known nuclear bodies that are present in nucleus are nucleoli, nuclear speckles, Cajal bodies (CBs), para speckles, Polycomb bodies (Mao et al., 2011) etc. Several nuclear bodies are found to have active transcription sites which consist of transcriptional machinery components including key transcription factors, RNA polymerase etc.,

1.1.1.2.1 Promyelocytic Leukemia (PML) Nuclear Bodies

PMLs are non-membrane nuclear bodies found in nucleus and are wide spread in mammals. PML nuclear bodies are composed majorly of PML-proteins which act like tumor suppressor proteins (Koken et al., 1994). PML-Nuclear bodies are dynamic structures which vary in size from 0.3 to 1.0 μm in size and numbers about 5-30 per nucleus. PML bodies response to stress signaling like DNA damages and subsequently their numbers increase. (Zhong et al., 2000, Van Damme et al., 2010) PML bodies are associated with diverse repertoire of protein molecules that carry various functions including transcriptional regulation, protein modification, DNA repair, DNA damage response, cell proliferation, senescence and apoptosis (Bernardi et al., 2008, Bernardi and Pandolfi, 2007, Salomoni and Pandolfi, 2002). Any disruption in PML structure or function leads to a diseased state like acute promyelocytic leukemia (APL) (Vitoux et al., 2007). PML protein is a member of tripartite motif protein family (TRIM) and because of alternating splicing, PML bodies exist in cell with 11 different isoforms and all isoforms share a common N-terminal end but varying C terminal end which results in each isoform carrying a well-defined function. This division of labor allow PML bodies to perform multiple functions. All isoforms contain nuclear localization signal (NLS) except VIIb which is the shortest and present in cytosol. Another isoform PML-I which contains nuclear export signal (NES) and thus it exist in both

nucleoplasm and cytoplasm (Nisole et al., 2013). All PML bodies plays an important role in dynamically storing nuclear proteins like RAD51 and Mre11 which are mainly involved in DNA damage and repair mechanisms respectively. Absence of DNA/RNA molecules within PML bodies suggests that these phase separated macro molecular complexes have DNA-Protein, and protein-protein interactions. (Lombard and Guarente, 2000). PML bodies are synthesized/expressed primarily during G1 phase and in S-phase the number of PML bodies decrease due to degradation of PML bodies. During DNA damage and repair, PML bodies recruits proteins likeNBS1, ATM, CHK2, and ATR.

1.1.1.2.2 Histone locus bodies

Histone locus bodies were first identified in *Drosophila melanogaster*. Most of the proteins present in HLB play important role in processing histone pre-mRNAs. HLBs are generally assembled at replication dependent histone genes and they help in concentrating biological factors, required for biosynthesis of histone mRNA. Major histones coding genes do not have introns in them and their pre-mRNAs are not spliced. Histone pre-mRNAs were found to carry a 3' extension which gets removed before being exported to the cytoplasm (Marzluff et al., 2008). Cleavage specific sites is determined by the U7 snRNP acting in cohort with other factors (Dominski et al., 2005, Kolev and Steitz, 2005). In addition to U7 snRNA, a few more proteins have been detected in HLBs, U7 specific snRNP proteins; Lsm10 and Lsm11, stem-loop binding protein (SLBP), FLASH, negative elongation factor (NELF), symplekin etc. Functional aspects of all these proteins have not been worked out yet.

1.1.1.2.3 Nuclear speckles

Nuclear speckles are nuclear domains enriched with pre-mRNA splicing factors and are located in the inter chromatin regions. The punctate nuclear localization pattern of pre-mRNA splicing machinery involving diverse components like small nuclear ribonucleoprotein particles (snRNPs), spliceosome subunits and non-snRNP protein splicing factors is termed as "a speckled pattern". Immunofluorescence microscopy studies have revealed that these nuclear bodies appear as 20–50 irregularly shaped structures and that variation in their sizes is common. size. Experimental studies have recently identified speckle targeting signals in some compartments. For example, arginine/serine-rich domain (RS domain) of SR pre-mRNA splicing factors, threonine-proline repeats of SF3b1, and forkhead-associated domain present in NIPP1 were found to be important in speckles targeting. Nuclear speckles are associated with transcription factors (Larsson et al., 1995), 3'-end RNA processing factors (Krause et al., 1994, Schul et al., 1998), eukaryotic translation initiation factor eIF4E (Dostie et al., 2000), eif4AII, which is involved

in translation inhibition (Li et al., 1999) and A-type lamins (Jagatheesan et al., 1999). Association of such diverse regulatory factors suggest possible mechanistic role in genome architecture and function. Immunofluorescence studies have revealed serine-2-phosphorylated form of the RNA polymerase II (RNAPII) large subunit (LS) also localizes to nuclear speckles, suggesting the functional role of transcription in maintaining this non-membrane bound nuclear body. Recently long non-coding RNA, MALAT1 was found to recruit SR splicing factors from nuclear speckles to sites of active transcription suggesting that these speckles act like biological non membrane signaling hubs. Speckles are dynamic structures and exchange of constitutive components with active transcription sites is a common feature. Recent paradigm changing studies have shown that highly transcribing genomic regions organize three dimensionally around nuclear speckles (Quinodoz et al., 2018). Studies on structural and functional aspects of nuclear speckles have provided us with an important paradigm for understanding the functional organization of the nucleus.

1.1.1.2.4 Peri nucleolar Compartment

The peri nucleolar compartment (PNC) is a non-membrane bound nuclear body generally present at the periphery of the nucleolus. Microscopic studies of HeLa cells have revealed that PNCS have multiple electron dense strands of approximately 80-180 nm. The PNCs remain physically associated with nucleolus but remain structurally distinct (Huang da et al., 2009). PNCs are also enriched with RNA-binding proteins and pol III RNA. Many proteins present in PNCs are involved in alternative splicing, de-adenylation and RNA decay. RNA component of PNC has also been implicated in RNA maturation, replication of mitochondrial DNA and pre-ribosomal RNA processing. It was later established that continual existence of PNC is dependent on continuous pol III transcription and the presence of key proteins. Presence of PNC in nucleoplasm has been strongly implicated in malignancy both in vivo and in vitro and is a potential cancer marker.

1.1.1.2.5 Nuclear Stress Bodies

In nucleoplasm, Nuclear stress bodies (NSBs) are formed in response to heat shock. Initiation of these stress bodies are seeded by coordinated interaction between heat shock transcription factor 1 (HSF1) and pericentric tandem repeats of satellite III sequences. NSBs functionally correspond to active transcription sites for noncoding satellite III transcripts. These nuclear bodies differ in structure and composition when compared to other known transcription sites and in response to heat shock, they orchestrate rapid

amplification transcriptome thus paving the way for chromatin remodeling and trapping of transcription and splicing factors.

1.1.1.2.6 Orphan nuclear bodies

Orphan nuclear bodies are less characterized and studied when compared to other nuclear bodies in cell nucleus. These bodies do not have traditional "markers" as is expected of most nuclear bodies. Orphan nuclear bodies are not detected in all cell types and are assembled and disassembled upon environmental signals at specific stage of the cell cycle. Experimental evidences in Caenorhabditis elegans suggest that sumolytation of the PcG protein, SOP-2 is needed to localize into these nuclear bodies *in vivo* and leading to physiological repression of *Hox* genes (Zhang et al., 2004). Recent studies have also shown that these bodies act as protein modification hubs mediated by SUMO and/or proteasome degradation of ubiquitin-tagged protein. Earmarking constrained 3-D space in nuclear volume for protein modification and degradation, lead to increased survival capacity of cells under changing environmental conditions.

1.1.1.2.7 Polycomb bodies

Polycomb bodies consist of Polycomb Group of proteins which are present as discrete foci in the nucleus at various intensities. It was observed that the number of polycomb bodies are far fewer when compared to target sites. In Drosophila melanogaster, Hox gene clusters are present in two distinct clusters termed Antennapedia and bithorax complex. These clusters are linearly separated by 10 Mb in the same chromosomal arm. During the course of development, in stage dependent manner, these linearly separated clusters co-localizes in 3D space and are co-repressed by Polycomb group of proteins (Bantignies et al., 2011) thus giving rise to the concept of polycomb bodies.

1.1.1.2.8 Nucleolus

Nucleolus is the one of the prominent nuclear domains existing in eukaryotic cells. Nucleolus is a membrane less organelle and carries functions related to ribosomal RNA biogenesis and assembly of ribosomal proteins with RNA. Hence, nucleolus is also termed as ribosome factory of the cell. It has been hypothesized that presence of clustered multiple copies of rDNA potentiate the RNA polymerase I dependent transcription in a cell cycle dependent manner. Recent studies have s shown that nucleolus is involved in multiple functions including telomere processing, DNA repair, pre-mRNA processing and cell cycle regulation etc., In higher organisms, nucleolus is tend to be assembled and disassembled during cell cycle. Onset of mitosis increases cyclin-dependent kinase 1 (CDK1)/cyclin B activity which triggers

repression of rRNA transcription that leads to breakdown of nucleolus. Similarly, decrease in cyclin-dependent kinase 1 (CDK1)/cyclin B activity decreases on mitosis stage exits rRNA transcription starts within nucleolar organizing regions (NORs) regions still remains by recruiting all downstream factors and followed by re-assembly of nucleolus (Hernandez-Verdun et al., 2010). The key transcription factor UBF component of Pol I initiation complex is still binds with NORs during mitosis even when transcription halts and nucleoli degradation. Few studies showed UBF helps in recruiting Pol I pre-initiation complex by invite studies (Prieto and McStay, 2007, Mais et al., 2005). Interestingly Cohesin a known player in chromatin organization, also plays a vital role in rDNA production and nucleolus organization. Studies have shown that depletion of Cohesin results in rDNA looping and disorganization of nucleolus in yeast (Gard et al., 2009, Harris et al., 2014). Recently it has been shown that transcriptionally inactive and centromeric regions organize three dimensionally around nucleus thus playing important role in establishment of higher order inter chromosomal networks (Quinodoz et al., 2018).

1.1.1.2.9 Cajal bodies

Cajal bodies was discovered by Santiago Ramón y Cajal in 1903 and they are typically 5μm in diameter. Cajal bodies are the sites of initial modification of U snRNPs, which are imported from cytoplasm by the SMN (Spinal Muscular Atrophy) complexes. Some of the U snRNPs are mainly involved in spliceosome or RNA spicing machinery (Morris, 2008). Cajal bodies are highly dynamic in nature and often they don't exist as a visible structure in the nucleus of highly proliferative cells like Hela cells. Only a small proportion 5-10% of cells might have Cajal bodies in any time (Hao le et al., 2007). Moreover, some cells might lack these Cajal bodies including smooth, cardiac muscle, skin cells and spleen parenchymal cells (Young et al., 2000). Coilin is known to play important role in maintaining structural integrity of spliceosomes. Mouse, lacking both coilin and SMN complexes shows embryonic lethality (Schrank et al., 1997). Recent studies have shown that Cajal bodies are linked mechanistically to transcriptional output and depletion of Cajal bodies lead to altered gene positions leading to the hypothesis that CBs have a genome organization role. Specific assembly events that take place in Cajal bodies are, assembly of U2 snRNP and the assembly of U5 snRNP with the U4/U6 snRNPs after each round of the splicing to reform tri-snRNP (Stanek and Neugebauer, 2006). Cajal bodies are also sites for the modification of 2'-*O*-ribose-methylation, pseudo uridylation (Verheggen et al., 2002) and replication dependent histone mRNA processing (Pillai et al., 2003).

1.1.1.2.10 Paraspeckles

Paraspeckles are a new class of sub nuclear organelle found within the nucleoplasm (Andersen et al., 2002). Para speckles are generally diffused and present at a concentration of 5-20 sub nuclear foci within the nucleoplasm (Fox et al., 2002). Co-localization studies have showed that these speckles did not coincide or overlap with any other previously identified sub nuclear organelle. These speckles were observed in interchromatin region near to nuclear speckles and thus named as para speckles (Fox and Lamond, 2010). A novel protein which is a marker of paraspeckles was identified and named as paraspeckle Protein 1" (PSPC1). Paraspeckles are enriched with small number of proteins and RNA. Most of the paraspeckle proteins are involved in Pol-II transcription and RNA processing. These proteins include DBHS (Drosophila behavior/human splicing) protein family, many transcription factors and splicing factors. The RNA component of paraspeckles includes ctn RNA, first discovered by Spector group (Prasanth et al., 2005) is mouse specific and majorly retained in paraspeckles. Ctn is majorly implicated in the control of gene expression which contains inverted and repeated Alu elements by nuclear retention (Levanon et al., 2005, Chen and Carmichael, 2008). It was later observed that paraspeckles remain unaffected after knocking down Ctn RNA (Prasanth et al., 2005). Various groups have later reported that NEAt1 (Nuclear Enriched Autosomal non-coding Transcript), a long nuclear RNA is essential for the formation structural and functional integrity of paraspeckles (Clemson et al., 2009, Sasaki et al., 2009, Sunwoo et al., 2009). In 2007, two ncRNA associated with paraspeckles were identified as NEAT1 and NEAT2 (MALAT-1) (Hutchinson et al., 2007). Both of these ncRNA were transcribed by RNA pol-II independently of protein coding genes and are well conserved in mammalian system (Fox and Lamond, 2010) and were shown to have diverse biological roles.

1.1.1.2.11 Transcription Factories

In classical view of Transcription, RNA polymerases diffuse freely through out nucleoplasm to search, recognize and bind to their respective promoters to initiate transcription a. It is difficult to comprehend that with such a large motley of transcription process involved and with large amount of DNA acting as template, transcription occurs with randomly diffusing components. It is now evident that most nuclear processes benefit from high local concentration of progressive factors and relevant template inside the nuclear volume. Presence of nuclear bodies like PML, splicing speckles, Cajal bodies suggest that compartmentalization acts to cluster relevant factors with common goal and function in nuclear volume to facilitate efficiency.

Several experimental and imaging studies suggest that RNA polymerases are not diffused throughout the nucleoplasm but are present in distinct foci leading to concept of Transcription factories. These factories

contain RNA-polymerase and protein machinery required for making RNA (Wachsmuth et al., 2008). The term Transcription factory was first coined in 1993 by Jackson and colleagues. They conducted an experiment in which mRNAs of permeabilized Hela cells, were labelled with Br-UTP during transcription and encapsulated in agarose microbeads. Observing these cells under Confocal microscope, they later observed that transcription process occurs at 300-500 discrete foci within the nucleus (Hozak et al., 1993, Iborra et al., 1996). These results were further corroborated by work done in different groups. Br-UTP labeling of nascent RNA and visualization by immunofluorescence (Wei et al., 1999) and (Iborra et al., 1996), Br-UTP or biotin-14 CTP labeling of nascent RNA and visualization by electron microscopy using immune gold particles revealed that signals are clustered in distinct foci and increase in signals is related to increase in labelling time. De-novo formation of new factories was not observed and total number of foci remained same during increase in labelling time, suggesting in principle the existence of Transcription factories.

Imaging studies further revealed that two different population of RNA-polymerase exist in nuclear volume, one that is scattered and other which is clustered in multiple foci. Scattered polymerases are not co-localized with transcripts whereas clustered polymerase showed co-localization with nascent transcripts, suggesting that clustered foci are engaged with transcriptionally active RNA-polymerases. Formation of nucleolus by phase separation is a good model to study Transcription factories. Nucleolus is a specialized site for ribosomal biogenesis and has mainly three primary regions. First region is the Fibrillar component, where clusters of RNA polymerase I is mainly concentrated. Second region is the dense fibrillary centers (DFC), which are rich in nascent RNA transcripts. The third region is the Granular component (GC) which contain ribosomal units and processed soluble RNAs. Regions of active transcription in nucleolus were visualized by labelling of mRNA with BrU followed by immune-labelling and fluorescent microscopy. Clustering of ribosomal genes in active transcriptional foci demonstrates the key postulates of transcriptional factory model. Transcriptional machinery and polymerases recruit genes to their location and are clustered in presence of a transcription factor, with transcription of specific subset of genes. Imaging studies have shown that initiating/elongating RNAPII also form transcription factories and they vary in numbers in different cell types. Transcription factory specialization is also observed in cell nuclei with RNAP II and RNAP III being present in different factories but never together (Pombo et al., 1999) thus suggesting that division of labor is associated with differences in spatial distribution of the factories. Further specialization was seen in RNAPII transcription factories where specific subsets of genes cluster together suggesting that these factories are important for co-regulated

gene expression (Berry et al., 2015) showed that RNA transcription in nucleoli is modulated by phase transition driven nuclear assembly. It is now evident that phase separation is important for organizing the intracellular environment inside the cell nuclei. A phase separation model has recently been postulated for transcriptional control and formation of transcription factories.

1.1.2 Chromatin organization in 2-Dimension and in 3-Dimensions

Early studies showed that chromosomes stained with Giemsa stain, show light and dark bands also called R and G bands respectively (Holmquist, 1992). G-band corresponds to gene poor regions and AT-rich isochores whereas R-band corresponds to gene rich regions and GC-rich isochores. It was then realised that, chromatin was compartmentalized into condensed and open state. Gene expression studies later clearly demonstrated that open chromatin is associated with higher gene expression and are clustered in specific chromosomal regions. Closed chromatin has relatively lower gene expression and are in gene deserts. (Caron et al., 2001, Lercher et al., 2002, Kosak et al., 2007, Versteeg et al., 2003). Further epigenome studies revealed that open or gene rich or early replicating chromatin of the genome is enriched for active histone marks such as histone acetylation (Barski et al., 2007, Martens et al., 2005). Compartmentalization of genome is functionally and structurally relevant and several hypotheses were put forward over the years to explain how compartmentalization helps in establishment of physical gene networks in interphase nuclei. Based on transcriptome analysis, it was postulated that genes located near to each other tend to be co-expressed or co-regulated, even though they are functionally unrelated (Hurst et al., 2004). The mechanism that drives this co-regulated and co-ordinated gene expression is not known but looping interactions by a few selective enhancers with multiple promoters are postulated to play an important role. This model was further supported by an experiment in which GFP reporter gene was inserted randomly in the mouse genome and was traced for GFP expression. It was observed that GFP expression was related to gene expression status of the chromosomal domain into which it was inserted (Gierman et al., 2007). This observation also correlated with expression pattern of related genes situated next to each other in close proximity in linear genome and were supported by gene clusters of α and β-globin and HOX genes. Clustering can also be supported by coordinated regulation of gene expression by sharing a common set of nearby regulatory elements (de Laat and Grosveld, 2003, Sproul et al., 2005). Genes can also be regulated by multiple regulatory elements that are located within chromosome nearby (cis) or might be in another chromosome (trans).

For a long time, we had a gene centric view to understand biological flow of information. Many breakthroughs were made by applying this reductionist approach to understand important biological processes. However, the role of regulatory regions and long-range looping interactions for establishing stage specific gene expression program was under appreciated for a long time. The discovery of non-coding RNAs opened a new vista in our understanding about the complex circuitry that regulate both transient and stable gene expression program. With the recent advances in imaging studies and development of Chromosome conformation capture(3C) based techniques, it has finally become possible to dissect the fundamental organizing principle of the interphase nuclei from 3D genome perspectives. 3C-based methods are based on frequency of physical interactions or proximity among any pair of genomic loci. Experimental data suggest that genome is not randomly folded and chromatin have preferential neighbourhood in interphase nuclei. From evolutionary perspective, several common features in metazoans are evolutionary conserved. This led to the hypothesis that there are a few key mechanistic players which play important role in genome folding. Moreover, the evolutionary conserved features point to the fact that 3D genome is intricately linked to spatio-temporal gene expression pattern in diverse species. Recent findings suggest that mammalian genome is hierarchically organized into different scales. From larger to finer scale, Chromosome Territory (CT), compartment A/B, Topologically Associating Domains (TADs), sub-TADs.

1.1.2.1 Chromosome Territories

Studies in past 20 years by using florescent in-situ studies have shown that chromosomes in interphase nucleus have preferential positions known as chromosome territories (CTs). Study of the nuclear organization gives a brief idea of it influences on gene expression. The concept of CT was first invented by Carl Rabl in 1985 by studies on salamandra maculate. Combination of various fluorophores and chromosome paints in single cells confirmed the existence of CTs and that they occupy preferential places in a nucleus (Zorn et al., 1979, Stack et al., 1977). FISH experiments further showed that CTs have distinct dark and light bands (Dietzel et al., 1998), which account for the position of active and inactive genes.(Volpi et al., 2000, Dietzel et al., 1999). CT positions were found to be relatively stable within a given cell type and are invariant between cell types (Walter et al., 2003, Gibcus and Dekker, 2013), indicating that trans interactions have little role in exacting nuclear position (Finlan et al., 2008). Recent data from Hi-C experiments have suggested that chromosomes are segregated into non-overlapping territories in interphase nuclei. Hi-C is a modified method of Chromosome Conformation Capture (3C) that gives preferential contacts of interacting loci genome wide in an unbiased matter.

This was evidenced by the fact that much greater intra-chromosomal contact probability was observed when compared to inter chromosomal interactions. Development and differentiation allows re-localisation of selective genes in and out from nuclear periphery which causes significant invariance between the cell-types (Ruault et al., 2008, Finlan et al., 2008). Experimental evidences of CTs existence by laser-UV-micro irradiation in isotonic cells by Giemsa stain clearly showed that chromosomes remain in non-random fashion and have preferential domains in interphase nucleus (Zorn et al., 1976, Zorn et al., 1979). In earlier times, researchers studied non-random radial distribution of chromatin in human lymphocyte nucleus by painting chr19, which is having highest gene density and also chr18, which is a gene poor chromosome. These results clearly showed that gene rich chr19 CT is consistently located interior of lymphocyte nuclei whereas gene poor chr18 CT is preferentially located at the periphery of the nucleus. Nuclear CTs are partially conserved after mitosis, yielding non-random but altered patterns in daughter cells when comparted to mother cells. The exact mechanistic details of epigenomic memory have not yet been clearly elucidated. Experimental studies on radial arrangements of chromatin within CTs further revealed that this non-random radial arrangement of CTs within nuclei is depends on their gene density. Similar studies when conducted in rodents (Mayer et al., 2005, Neusser et al., 2007), cattle (Koehler et al., 2009) as well as in birds (Habermann et al., 2001) showed similar results. Further studies showed that there is a causal relationship between nuclear position of chromatin elements and their properties like gene density, transcriptional activity and its radial position in the nucleus (Croft et al., 1999, Boyle et al., 2001, Cremer et al., 2001, Tanabe et al., 2002, Kozubek et al., 2002). Genes from CT which are highly expressed are positioned in the interior of cell e nucleus, whereas inactive regions are preferentially positioned at nuclear periphery. Interphase chromatin is more dynamic, mobile and has sites for active transcription and these sites are compartmentalized within CT (Fakan, 1994, Verschure et al., 2002, Cmarko et al., 1999). FISH studies have shown that chromosomal regions within compacted CTs loop out from the compacted CT (Boyle et al., 2011). There are only handful of studies till date that investigated changes in chromosomal organization in different cell types or tissues in developmental and disease progression context (Mayer et al., 2005, Parada et al., 2004, Neusser et al., 2007, Kuroda et al., 2004), further in cancer cells (Parada et al., 2002).

1.1.2.2 Sub-Genome Compartments

Pearson correlation analysis of dilution Hi-C data at 1Mb resolution revealed a plaid pattern of interactions in each chromosome. Principle Component Analysis revealed that two chromosomal domains with positive and negative values exist which were termed as A and B compartments

(Lieberman-Aiden et al., 2009, Rao et al., 2014). Compartment A is nearly active chromatin and correlates with early replicating of euchromatin of the genome. Whereas compartment B is highly inactive chromatin and is majorly correlates with late replicating or heterochromatic DNA. HiC analysis also reveals that compartment A is highly interacting with A whereas compartment B interacts only with compartment B. Upon close observations of these compartments by overlaying with epigenome and various DNase hypersensitivity data sets revealed that compartment A is high with gene density, transcriptionally active, has active histone modification H3K36me3, DNase hypersensitivity region and high GC content. Similarly compartment B contains densely packed DNA which is transcriptionally inactive, has inactive histone mark H3K9me3 (Hawkins et al., 2010, Bernstein et al., 2010, Fortin and Hansen, 2015) high AT rich chromatin and is localized at nuclear periphery (Ryba et al., 2010). Later higher resolution *in situ* Hi-C datasets at 5kb resolution confirmed that the organising principle and folding pattern observed with low resolution Hi-C data exist at a much finer scale also. At 5Kb resolution, compartments A and B were further subdivided to A1-A2 and B1-B3 based on histone modifications. A1 and A2 regions are associated with H3K36me3, H3K79me2, H3K27ac and H3K4me1 and are depleted at the nuclear envelope and at nucleolus associated domains (NADs). A1 and A2, both exhibit early replication times. Replication of A1 is done at the beginning of S-phase, whereas A2 continues replication till the middle of S-phase. A2 s strongly associated with H3K9me3 compared to A1 and has lower GC content, and contains longer genes. B1 correlates positively with H3K27me3 and negatively with H3K36me3, suggestive of facultative heterochromatin and its replication peaks during middle of the S-phase. B2 is positioned at peri-centromeric heterochromatin and is mostly associated with nuclear lamina and at NADs. B3 is ordinary heterochromatin and is associated with nuclear lamina but is strongly depleted at NADs. Sub-compartments B2 and B3 do not replicate till the end of S-phase. This distribution was similar in all autosomes in a given cell type. Further to understand changes in compartment distribution in senescent cells, recent studies found that there is high correlation between compartment A and B with active and repressed chromatin in senescent cells (Criscione et al., 2016) Further Hi-C studies on senescent cells revealed that there is a decrease in long-range (>2Mb) and concomitant increase in short range (<2Mb) interactions. At much higher resolution TADs are well conserved between senescent and proliferating cells but only few TADs markedly switch between B to A compartment and shows transcriptional activity (Criscione et al., 2016). It raises questions regarding at dynamics of compartmental reshuffling during embryonic stem cell differentiation has found that differentiation is accompanied dynamics

switches between A and B compartments. Only a subset of genes gets effected by compartment changes and remaining were found to be unaffected genes show nuclear LAD tethering (Dixon et al., 2015). These changes are highly predicted by H3K4me1 and enhancer elements indicating that enhancer elements are more likely key players in shaping the higher order chromatin changes (Dixon et al., 2015). Genome is organized in a hierarchical manner and fractal globule model was put forward to understand the polymeric nature of chromatin folding in interphase nuclei. Hi-C studies were based on population of cells, so naturally questions were raised whether individual cells follow similar folding principles. Multiplexed FISH studies by (Wang et al., 2016) confirmed the existence of Compartments A and B and depleted contacts were observed between A and B. Single cell Hi-C studies by multiple groups (Stevens et al., 2017, Nagano et al., 2017) also confirmed the existence of compartments. Now it is widely accepted fact that compartments are separated in space and are organized in spatially polarised manner in individual cells.

1.1.2.3 Topologically Associated Domains (TADs)

Analysis of Hi-C data at low resolution revealed that interactions between loci of same compartment on an average 800 kb in size are higher than interactions of different loci between compartments (Lieberman-Aiden et al., 2009). With advancement in technology and presence of high resolution Hi-C and 5C (derivative of 3C) datasets allowed for the systematic studies of self-associating chromosomal domains. Analysis of high resolution datasets at 40 kb resolution or higher revealed the presence of self-interacting chromosomal regions termed as Topological Associated Domains (TAD) (Dixon et al., 2012, Nora et al., 2012). In mouse embryonic stem cells more than 90 percent of the genome was found to be organized into 2200 distinct TADs with median size of 880kb. TADS were also observed in other mammalian cell lines (Nora et al., 2012). Similar domain organization albeit with smaller size was also observed in other non-mammalian species like Drosophila (Sexton et al., 2012), zebrafish (Gomez-Marin et al., 2015). In Arabidopsis thaliana, TADs were not observed suggesting that plants might have different folding mechanisms. TADs are majorly invariant after many cell divisions and are evolutionary conserved among different cell types and thus TADS are considered as a basic unit of genome organization (Dekker and Heard, 2015). TADs are also highly conserved among related species and are mostly independent of tissue and cell type specific gene expression pattern. Notably, more than 50% of TAD boundaries are conserved in human and mouse ESCs. Surprisingly in mouse ESC differentiation, it was observed that even in presence of large scale compartment switching between A and B, TADs positioning remained stable. High correlation in TAD positioning between sperm nuclei and fibroblast also suggest that TADs

are indeed a basic unit of genome organisation. When compared with the interactions of enhancers and promoters present in the same TAD with that of neighbouring TADs, weak correlation was observed, suggesting the interactions of enhancers and promoters between the TADs are weakly co-ordinated (Dixon et al., 2015, Nora et al., 2012, Zhan et al., 2017). To further dissect whether TADs are indeed basic unit of 3D genome organisation or not, a GFP reporter gene was randomly inserted in moue genome. Expression pattern of the reporter gene was found to be directly co-related with expression pattern within the same TAD. Moreover recent experimental evidences also suggest that early and late replicating DNA domains co-relate with TAD domains and are majorly involved in regulation of genes (Pope et al., 2014). Replication domain of 400-800kb whose positions are invariant across different cell types and species exist. All this relevant information has led to the conclusion that TADs are indeed a basic functional units of 3-D genome organisation.

Contact frequency of genomic loci in the same TADs is several folds (at least 2 folds) higher compared to contact foci of loci present in neighbouring TADs. Intra-TAD interactions are relatively more compared to Inter-TAD interactions. The relative stable nature of TAD boundaries has opened up intriguing questions about how TAD borders are maintained. Epigenome data suggest that TAD borders are enriched with transcription start sites, housekeeping genes, H3K4me3, H3K36me3, short interspersed nuclear elements and CTCF binding sites. Particularly striking was the enrichment of CTCF binding sites and housekeeping genes at TAD boundaries. 33% of all TAD boundaries are enriched with CTCF binding sites and 75 % of all TAD boundaries are enriched with housekeeping genes in mouse ESCs. This led to the hypothesis that high transcriptional activity might contribute to TAD formation. TADs are invariant between the cell types but TADs are not constitutive feature of the chromosome for example TADs are absent in condensed chromatin structure at mitotic stage (Naumova et al., 2013). TAD structures are established in early phase of G1 stage and are present during interphase and G2 phase but they disappear during mitosis. This concludes that TADs are of basic units of chromosome whereas in mitosis high condensation of chromosome favours loss of TADs and compartments and allowing for the chromosome to fold into a condensed and rod shape like structure. Role of architectural proteins like CTCF in genome organisation is now widely investigated. Down regulation of CTCF by RNAi technology resulted in increased inter-TAD interactions and reduction in intra-TAD interactions, suggesting weakening of TAD boundaries. This observation was in line with the hypothesis that CTCFs play a major role in delineating TAD boundaries. Acute depletion of endogenous CTCF in ESCs with an auxin-inducible degradation

resulted in TADs structure disruption. It was also observed that loss of CTCF results in disruption of local insulation and cause transcriptional mis regulation (Nora et al., 2017). Surprisingly TAD disruption did not affect higher order chromatin structure like compartments, suggesting different mechanistic approaches in folding principles of the genome at different scales. Overall chromosome compaction and organization remain unaffected. It is now hypothesized that other factors might play major role in packaging and compartmentalization of the genome (Nora et al., 2017). Recently loop extrusion model has been put forward to explain how CTCF mediated long range interaction can lead to de-novo formation of TADs. In loop extrusion model, a chromatin motor complex (like cohesion complex) loads onto DNA and then extrudes to form a progressively larger loop unless it is stalled by CTCF, binding at convergent sites and thus forming TADs. This model is in line with various polymer simulations that try to model observed TAD structures from various Hi-C datasets. This model also makes accurate prediction regarding de-novo loop formation when anchored CTCF sites are deleted (Fudenberg et al., 2016, Sanborn et al., 2015). One drawback with loop extrusion model is that it cannot explain the systematic formation of long-range *cis* and *trans* chromosomal interactions to form domains of the scale reported by Hi-C datasets. It is also known that not all TAD boundaries have convergent CTCF binding sites giving rise to hypothesis that alternative CTCF independent mechanisms also exist in determining TAD boundaries. One attractive alternate hypothesis regarding formation of TADs is mediated by DNA supercoiling (Benedetti-2014). Boundaries of supercoiling domains partially overlap with TAD boundaries. Topoisomerase II beta (TOP2B) is also known to localize with CTCF and cohesion at TAD borders. It is important to note that till data conclusive evidence regarding how TADs are formed and how loops are established at different scales is missing. Further experiments are needed in this direction.

Alteration in TAD boundaries lead to gene dysregulation due to ectopic interactions between regulatory elements. Naturally occurring mutations affecting TADs boundaries have been implicated in cancer. Hi-C data has recently been used to study balanced and unbalanced chromosomal rearrangements in primary human tumors (Harewood et al., 2017) and cancer samples. These studies have revealed new layers of complexity present from disease onset to disease progression. In *IDH* mutant gliomas, activities of TET proteins which are involved in DNA demethylation is significantly reduced leading to low CTCF occupancy at some key CTCF sites because of hyper methylation. This results in loss of insulation and disruption in TAD boundaries. Increased ectopic interactions between a constitutive enhancer and an established glioma oncogene *PDGFRA* trigger upregulation of the gene by many folds (Flavahan et al.,

2016). Oncogene activation because of disruption in TAD boundaries is also observed in T-cell acute lymphoblastic leukemia (T-ALL). TAL1 and LMO2 are two oncogenes located near impaired TAD boundaries in HEK-293T cells and deletion of this specific boundary results in ectopic interactions with distant enhancers leading to their activation (Hnisz et al., 2016). CTCF sites are found to be frequently mutated in multiple cancer types and as CTCF is one of the key players in TAD formation, it is now postulated disruptions in 3D genome architecture is one of the hall marks of cancer cells (Katainen et al., 2015). Higher order structure around EPHA4 is very important for proper functioning and any deletions of the TAD encompassing EPHA4 at the telomeric side results in brachydactyly disease, whereas in other case inversion and duplication of TAD involving EPHA4 on the centromeric side results in syndactlyly indicating the functional role of higher order structure in onset of disease and progression. (Franke et all 2016).

1.1.2.4 Sub-Topologically associated domains (Sub-TADs):

TADs are further divided into smaller units known as sub-TADs. They were first observed with high resolution datasets from Hi-C and 5C experiments (Rao et al., 2014, Phillips-Cremins et al., 2013). Sub TADs, similar to TADs and are self-associating domains and it was observed that there was marked reduction in frequency of contacts between genomic loci present across sub-TAD boundaries. This observation mirrors the findings observed in TADs. Similar to TADs, boundaries of Sub-TAD are enriched with CTCF and cohesin and are less conserved across different cell types. In response to external and developmental signalling cues, sub-TAD dynamics in space and time helps in establishing local insulating chromatin loops, thus facilitating in establishment of physical gene networks in stage dependent manner. It is well postulated that the sub-TAD partitioning from TADs might be due to cell type specific promoter and its regulatory element interactions mediated by CTCF, cohesion, mediators and lineage specific proteins. The role of lineage specific proteins has not yet been established though but is postulated. Further experimental evidences obtained from experiments that looked into protein centric view of 3D genome conformation suggested smaller structural and functional units with in TADs. Analysis of ChIA-PET data for CTCF and cohesion revealed that chromatin loops mediated by these local factors are units of gene regulation. Mutation at local factor binding sites lead to gene dysregulation (Dowen et al., 2014).

1.1.2.5 Architectural proteins and their role in 3-D genome organisation:

At a much higher level, chromosome occupies preferential positions within interphase nucleus and is further decomposed into to so called topologically associated domains (TADs). TADs are highly self-interactive entities and are invariant between cell types. Interactions within TAD are highly enriched when compared to the interactions between the TADs, which suggest TADs limits the spread of interactions in genome and helps in establishing insulator neighbourhoods. In mammalian cells, TAD boundaries are enriched for Short Interspersed nuclear element (SINEs), CTCF and Cohesin proteins (Dixon et al., 2012).

CTCF is the major insulator protein in mammals, and its function is required for early mouse development particularly in oocyte development (Fedoriw et al., 2004) Conditional knockdown of CTCF in somatic cells suggest its role in differentiation, apoptosis and cell cycle progression (Heath et al., 2008, Soshnikova et al., 2010, Splinter et al., 2006), CTCF is a zinc finger protein that regulate various aspects of gene regulation, one such example is insulating gene promoters from distant enhancers. It was later found that enhancer blocking activity is necessary for gene expression in H19-IGF2 locus (Bell and Felsenfeld, 2000, Hark et al., 2000) and further experimental studies revealed that gene expression in H19-IGF2 locus is mediated by CTCF by forming allele specific chromatin loops (Kurukuti et al., 2006). Remarkably cohesion was also found to be necessary for enhancer blocking activity. Most importantly these requirements were observed in G1 phase of the cell cycle suggesting important functional role for Cohesin. Several hypothesis regarding formation of chromatin loops were put forward where CTCF work in conjunction with cohesion for establishing insulated neighbourhoods. CTCF is ubiquitous in nature recent study indicates that nearly ~14000 CTCF sites are present in human fibroblast cells, However the distribution is 48% in intergenic, 22% intronic, 12% exonic and relative to this 20% sites are present within 2.5kb upstream of transcription start site (TSS) (Kim et al., 2007). ENCODE project revealed that ~68000 sites CTCF sites are present in 38 different cell types(Wang et al., 2012). Mass spectroscopy analysis of purified CTCF from Hela cells revealed that it can form homodimers and multimers *in vivo*. This property of CTCF was confirmed by yeast two hybrid system (Yusufzai et al., 2004). In vitro study showed that glutathione S-transferase (GST) purified CTCF binds to 11-zinc-finger domain of CTCF (Pant et al., 2004).

It is now known that CTCF binds to chromatin and establishes loops between the promoters with its regulatory elements and Cohesin binds with C- terminus of CTCF and stabilizes these interactions (Zuin et al., 2014). Loop extrusion model is now considered to play important role in formation of loops inside

the nuclei. Inactivation of CTCF gene in mouse embryonic fibroblast cells induces miss-regulation of nearly 698 genes (Busslinger et al., 2017), Similar kind of studies performed in post mitotic embryonic and postnatal neurons revealed changes in 400 and 800 genes respectively (Hirayama et al., 2012, Sams et al., 2016). Recent studies in ESCs suggest that in absence of CTCF, compartments retain their structure but at TAD level, 3D genome architecture is disrupted leading to changes in gene expression (Nora et al., 2017). Further experimental evidences suggest that CTCF is not the sole factor that influences 3D genome organisation. Regulators like YY1 factor along with Cohesin and CTCF establishes long range loop interactions between promoters and its enhancers (Weintraub et al., 2017, Zuin et al., 2014). Cohesin orchestrates two types of interactions, which includes Cohesin independent organization involving local transcriptional activity defined by epigenetic landscape and secondly, Cohesin dependent orchestration through loop extrusion model, allowing for preferential contacts in between regulatory elements and their target genes there by formation of TADs (Schwarzer et al., 2017). In general, CTCF mediated loops requires two CTCF molecules and each one binding at the end of the chromatin loop in a convergent fashion in both primary and secondary loops (de Wit et al., 2015, Guo et al., 2015, Rao et al., 2014). Apart from transcription and boundary role, CTCF also participates in imprinting and in X-chromosome inactivation and these functional roles could be explained by the ability of CTCF to make longa range DNA looping interactions (Ling et al., 2006, Zhao et al., 2006). Deletion of CTCF results in decrease in shorter loops, and consequent increase in inter TAD interactions and loss of intra TAD interactions within TADs (Baranello et al., 2014). Disruption of cohesin complex leads to decrease or loss of short range interaction but is not accompanied by breakdown of TADs (Zuin et al., 2014).

Polycomb group (PcG) proteins identified in *Drosophila* and vertebrates which regulates vast number of developmentally regulated genes which play an important role in chromatin organisation. PcG proteins form multimeric complexes and acts directly by binding to specific chromatin region via histone modifications. Two proteins of the PcG complex, PRC1 and PRC2 were first identified in drosophila. Few studies on drosophila showed that formation of H3K27me3 marked repressive chromatin is mediated by recruitment of PRC1 and PRC2 to the polycomb responsive elements (PREs). PRC1 protein inhibits transcription elongation by recognizing H3K27me3 through H2A ubiquitinoylation (Eskeland et al., 2010, Zhou et al., 2008). PRC1 also maintain compaction state of HOX gene clusters in ESC during development. Methylation of H3K27 by PRC2 results in chromatin silencing mediated by PcG proteins. In mammals, YY1 factor which is a homolog to PRC gets recruited and methylate H3K27 (Brown et al., 1998). In-vitro and in-vivo studies also showed that PcG proteins are involved in chromatin compaction

in both drosophila and mammals through recruitment of methyltransferases like EZH1 and EZH2 (Margueron et al., 2008, Kuzmichev et al., 2004). Significant number of stalled RNA-pol II are found at many promoters repressed by PcG (Stock et al., 2007, Enderle et al., 2011) suggesting that PcG is involved in transcription modulation and probably has mechanistic role in chromatin folding and its spatial arrangement. Recent study of tethered Dam methyltransferase in drosophila demonstrated that distantly located PREs can interact with its regulatory elements in bithorax complex (Cleard et al., 2006). FISH studies combined with polycomb immunostaining with 3C have shown that PcG form multi looped structures and is conserved in mammals (Lanzuolo et al., 2007). Depending upon the external signalling cue, genes get displaced from the repressive hub by looping out, suggesting that transcriptional output is dependent on its higher order chromatin organization.

1.1.2.6 LAD organization and its role in 3D genome organization

Heitz in 1928, identified chromatin segments that failed to dencondense at the telophase stage of cell cycle. These compacted segments termed heterochromatin remain condensed throughout the interphase nucleus. In general heterochromatin is highly condensed, mostly localized at nuclear periphery, and is primarily associated with late replicating domains of cell cycle. Heterochromatin is also highly concentrated at centromere and telomere regions of the chromosome. Electron and light microscopy observations have shown that there is close connection of heterochromatin with nuclear lamina (Fawcett, 1966). Nuclear lamina is the major structural support for the nuclear envelope and carries out various functions within the nucleus. Laminas are composed of complex proteins which connect inner nuclear membrane with chromatin and nuclear pore complex. Lamins are type V intermediate filament (IF) proteins and are divided into two types; Type A and type B. In mammalian cells there are two major A-type Lamins (A and C) that are transcribed from a single gene, LMNA and B-type lamins (B1 and B2) are transcribed from two different genes LMNB1 and LMNB2 (Lin and Worman, 1993, Peter et al., 1989, Vorburger et al., 1989). Lamins regulate various functions within nucleus including maintenance of nuclear shape and mechanical support (Ostlund et al., 2001, Favreau et al., 2003, Vaughan et al., 2001, Goldman et al., 2004). Close connection of lamins with heterochromatin suggest that lamins may have structural and functional role in 3-D genome organization by orchestrating heterochromatin at nuclear periphery which is transcriptionally silent (Fawcett, 1966, Galiova et al., 2008).

Nuclear localization of heterochromatin at nuclear periphery by Lamins regulate transcription by frequently positioning gene desert regions at nuclear periphery. Lamins and chromatin interactions were

identified through DamID technique. In this technique, Lamin proteins are fused with DNA adenine methyl transferase (Dam) from *E. coli*. Dam fused Lamins interact with chromatin present at nuclear periphery and methylate it. Methylated DNA is then sequenced and analyzed (Pickersgill et al., 2006). LADs which are mapped to human and mouse numbers about ~1000-1500 and spans about 10kb-10Mb in size and covers about one third of the genome. LADs are gene desert regions and are enriched with repressive histone modifications H3K9me2 and H3K9me3 (Guelen et al., 2008, Wen et al., 2009). In mammals, LADs that are invariant between the cell types are known as "constitute LADs (cLADs)", while LADs that are interacting with lamina in cell specific context are termed "Facultative LADs (fLADs)" (Meuleman et al., 2013, Peric-Hupkes et al., 2010).

In general, cLADs mostly coincide with gene desert regions of the genome and AT-rich isochores (Meuleman et al., 2013). It is now believed that cLADs act as a scaffold which tether chromatin to nuclear periphery. fLADs are highly mapped to GC-rich isochores and have relatively high gene density regions. fLADs are less conserved in mouse and human cells than cLADs. During the course of cellular development and differentiation, LADS play an important role in activation and repression of stage specific genes which are fate determinants (Loh et al., 2016). It was also observed that during cellular development and lineage restriction large number of genes change their position relative to their nuclear lamina (Meuleman et al., 2013). Differentiation signalling induces detachment of genes from nuclear lamina to interior of cell nuclei and vice versa and is associated with establishment of physical gene network pertaining to lineage restriction (Peric-Hupkes et al., 2010, Robson et al., 2016). In mESC, regions of key cardiac genes are spatially positioned within H3K9me2 marked heterochromatin at nuclear periphery in an Hdac3 dependent manner and differentiation induces repositioning of genes essential for cardiac lineages to the interior of cell nuclei (Poleshko et al., 2017). Several proteins are involved in driving LADs to NL like H3K9 methyltransferases. In *C. elegans*, it was observed that transgenes present in heterochromatic regions are localized at nuclear periphery and MET-2 and SET-25 are the two H3K9 methyltransferase that are required for peripheral localization of heterochromatin in nucleus (Towbin et al., 2012). Depletion of these methyltransferase results in re-localization of transgene to nuclear center by partial loss of NL interactions. Weakening of LAD and NL interactions were observed by inhibition or disruption of G9A methyltransferase in H3K9me2 (Bian et al., 2013, Chen et al., 2014). Studies on *C. elegans* embryos showed that small protein CEC-4 which is anchored at inner membrane of nuclear periphery binds to mono, di and tri methylated H3K9 chromatin, and loss of CEC-4 results in detachment of heterochromatin from NL thus forming a functional association between H3K9

methylation and NL (Gonzalez-Sandoval et al., 2015). Distribution of Lamins in nucleus is not uniform for example B-type Lamins are more confined to NL present at nuclear periphery whereas lamin A-type is spread throughout interior of the nucleus (Gesson et al., 2014). Deletion of Lamin B type in drosophila results in detaching of genes from NL (Kohwi et al., 2013, Lund et al., 2015). DamID profiles of Lamin A and B showed that both are similar in structure. It is now hypothesized that Lamin A bind with LADs when they are placed at the interior of cell nuclei. Recent ChIP experiments have shown that Lamin A and C binding sites are also present in inter LAD regions (Kind and van Steensel, 2014, Gesson et al., 2016, Lund et al., 2015). Interestingly deletion of Lamins from mouse ESCs results in marginal effect on NL contacts, thus suggesting that Lamins might have redundant roles in mESCs. It is now known from Hi-C data that metazoan genome can be partitioned into two discrete Compartments (A and B). Compartment A is transcriptionally active and compartment B is transcriptionally inactive. Analysis of Dam ID and Hi-C studies have shown that LADs mostly correlate with compartment B i.e. repressive compartment suggesting both structural and functional role. LADs also correspond to heterochromatin compartments present in nucleoli and centromere (Stevens et al., 2017). Collectively lamina tethered LADs are divided into cLADs and fLADs in which cLADs are invariant between cell types whereas fLADs are dynamically shuffled from NL to the interior of the nucleus during cellular differentiation. It is now known that LADs are important for coordinating cell specific gene expression program but little is known about LAD dynamics and 3D genome in the context of differentiation. Time and space studies are important to understand how dynamic switch in LADs facilitates stage specific physical gene networks in the context of differentiation.

1.1.3 Spatial organization of genome during cell division (Time and Space)

How 3-D genome architecture is restored after cell division in interphase nuclei of daughter cells is an open question in the field. Chromosome dynamics during cell cycle was difficult to study due to several technical limitations but with the availability of different synchronization techniques and refinement in high throughput methodologies, it is now possible to study 3D genome organization in 4D (space and time). Cell cycle involves stages of interphase, metaphase and telophase; however, in interphase, chromosomes are mostly static at global level wherein metaphase, chromosomes are linearly condensed and are ready to transfer to daughter cells. Mistakes in mitotic condensation and decondensation, afterwards in daughter cells can lead to life-threatening diseases like cancer, neurodegenerative diseases etc. Analysis of single cells and population based Hi-C have revealed stage dependent chromatin interactions in different stages of cell cycle. G1 stage chromatin is transcription-compatible and is

characterized by a higher probability of short range intra-chromosomal interactions compared to long range intra and inter-chromosomal interactions. It was also observed that the extent and nature of short range contacts differs between organisms depending on their chromosomes and genome size. S phase or replication phase is characterized by increased long range intra-chromosomal interactions compared to short range intra-chromosomal contacts. Cells in their G2 and mitotic phase further showed increase in short range interactions which are indicative of gradual axial compaction and suggesting individualization of chromosomes as observed in cell cycles. (Nagano et al., 2017) showed that CTCF mediated chromatin loops, TADS insulation, compartmentalization are present during the cell cycle. TAD insulation was found to be reduced in S-phase i.e when coupled to replication process. Increase in compartmentalization was observed throughout G1 and S-phase. Mechanistic details about epigenomic memory has not yet been elucidated, but several mechanisms are hypothesized. It was observed CTCF can remain bound to chromatin throughout the cell cycle and thus probably acting as a scaffold in reconstituting 3D genome architecture in daughter nuclei. Role of CTCFs and its role in epigenomic memory for faithful reconstitution of chromosome organization are now being widely explored. To understand the folding principles at mitotic phase, several models were proposed such as "Chromosome stress cycle hypothesis". This model is based on dynamic participation of three key players in chromosome compaction, Condensin subunits (SMCs), TopoII and KIF4. Condensin participates in loop extrusion, TopoII catalyzes DNA passing reaction and sister chromatid segregation requires TopoII decatenation (Nagasaka et al., 2016). KIF4 is a microtubule associated motor chemo kinesin and is majorly associated with chromatin remodeling complex and histone chaperons. Loss of KIF4 results in de-condensation of chromosomes in interphase whereas hyper condensation is observed in mitosis (Samejima et al., 2012). Loop extrusion model is now envisioned to provide the structural framework to understand the mitotic chromosome dynamics. It explains that chromosome first compacts in prophase due to cohesin and the accumulating stress is released by TopoII which leads the structure to a lower energy state. Further compaction occurs due to Condensin in metaphase (Goloborodko et al., 2016). Recent study by (Gibcus et al., 2018) modelled polymer simulations using imaging and HiC data of synchronous DT40 cells to understand the mechanistic principles of chromosome folding in mitosis when cells enter from G2 stage to metaphase. In this study they also observed that the loss of interphase chromatin organization in Prophase, is a Condensin-dependent process and arrays of consecutive 60kb loops are consequently formed. Polymer simulation modelling showed that loop extrusion principles in synchronized cells is dependent on the dynamics of Condensin 1 and Condensin 2 during prometaphase, centrally located axis

acquires helical twisting giving rise to loops emanating as spiral staircase. ~400 kb outer loops are further nested within 80kb inner chromatin loops. With the progression of prometaphase, chromosomes shorten themselves through progressive helical winding, by increasing the number of loops per turn. During prometaphase helical turn size progressively increases to ~12 Mb (150 loops) from 3 Mb (40 loops). Differential action of Condensins were demonstrated by acute depletion of both I and II Condensins. Which reveals condensin II is required for helical winding and condensin I modulate the size and arrangement of nested inner loops.

1.1.4 3D genome studies in context of differentiation (Time and Space)

Lineage restriction of pluripotent stem cells is accompanied by changes in gene repositioning within the nuclei, major compartmental switches and changes in long range chromatin interactions. One of the early studies concerning mechanistic details of differentiation, revealed that β-globin gene locus is repositioned from nuclear periphery to the interior of cell nuclei in later stages of erythroid differentiation (Palstra et al., 2003). This movement was found to correlate with the establishment of long range interactions mediated by chromatin loop between β-globin gene locus and locus control regions (Krivega and Dean, 2012, Krivega et al., 2014). Further studies in differentiating B-cells have shown that immunoglobulin heavy chain gene drastically undergo intra-nuclear spatial repositioning during differentiation. *De novo* long-range physical interactions get established in differentiation of B-cells due to re-localization between enhancer and VDG gene cluster (Guo et al., 2011). Application of Hi-C methodology during stem cell differentiation has revealed that TADS boundary positioning remains largely unchanged in hESC and in hESC-derived restricted lineages. However, marked changes in both intra-TAD and inter-TAD interactions correspond to alternations in epigenetic states and transcriptional levels (Dixon et al., 2015). During Neuronal differentiation, it was surprisingly observed that TADs tend to interact with neighboring TADs forming mega domains or meta-TADs like structure (Fraser et al., 2015b). Meta-TADs were found to be relatively more stable during differentiation but TAD-TAD interactions were subjected to changes during differentiation and accompanied by changes in gene expression (Fraser et al., 2015b). HOX genes provide an important paradigm in dissecting 3-D genome dynamics in context of differentiation. HOXD genes are upregulated sequentially during limb development and are silenced in embryonic stem cells. It was observed that silenced HOXD clusters are present in a single repressed chromatin compartment containing all genes. Onset of Transcription results in physical reorganization of the repressed chromatin with the formation of active and repressed compartments. Developmental signaling induced changes in HoxD clusters is accompanied with

switching of gene expression status and long -range interaction pattern of the active and repressed compartments (Andrey et al., 2013). Studies on HOXA clusters have also shown that embryonic stem cell differentiation is associated with establishment of insulating neighborhood which separate borders between active and repressed genes (Narendra et al., 2015).

Differentiation of terminal B-lymphocytes into plasma cells provides a new perspective to our understanding of genome dynamics in space and time. G0 B-cells (a naive, resting, or quiescent) undergo a global change in their chromatin state and subsequent amplification of their transcriptome as they enter the cell cycle. Molecular mechanisms that drive epigenetic accessibility is mostly unknown. Recent studies have shown that chromatin decondensation in B-cells, which facilitate transcriptome amplification is orchestrated by at least three mechanisms. First, the chromosomes spread and reposition themselves from periphery to the nuclear interior to prevent transcriptional silencing (Therizols et al., 2014). Second, decompaction of chromatin nano-domains results in increased accessibility to DNA binding regulatory proteins. Third, changes in chromatin dynamics, with long-range interactions becoming predominant compared to short range interactions, mediated by doubling of CTCF-mediated loops and contact domains. Chromatin architectural changes helps in establishing physical cognate promoter-enhancer contacts and requires the activity of Myc and continual ATP production (Kieffer-Kwon et al., 2017).

Ultra-deep Hi-C sequencing of mESCs, neuronal progenitors and cortical neurons in G1 phase of cell cycle have recently shed some light on some key genome folding principles in regard to neuronal differentiation. It was observed that neural differentiation is associated with a progressive increase in compartment size, reduced interactions within the A compartment and increased interactions between B-type domains. It was also observed that genes located near to the neural-specific TAD boundaries get up-regulated during neuronal differentiation whereas genes present near to ES-specific borders are down regulated in differentiated cells. Promoters of active genes are strongly insulated, which co-relates with transcriptional activity, and are mostly CTCF independent. Active Transcription start sites have enriched contacts with the gene body and interacts with other transcribed transcription start sites (TSS) present across TAD boundaries. Transcribing genes with large number of exons regularly engage in establishing long-range chromatin loops in clusters both in *cis*- and in *trans*. Transcriptional activation was found to be insufficient for maintaining chromatin insulation, as other regulatory factors such as enhancer-promoter interactions and specific TFs play important role in creating insulating neighborhood for spatiotemporal gene expression pattern (Bonev et al., 2017).This study also shed light upon the role of

pioneer transcription factors in establishing the epigenomic landscape for transitions from one state to another , as it was observed that cell type specific chromatin folding in the context of neuronal differentiation is organized around a few important pioneer transcription factors like, NeuroD2,Pax6 and Tbr1 in-vivo. Analysis of Polycomb-mediated interactions in 3-point neuronal differentiation system revealed that these were among the strongest in mouse ESCs but get disrupted progressively during neuronal differentiation. However, Polycomb-mediated interactions were found to be present or even enhanced at a specific gene, in a manner which correlates with the levels of Ring1B bound to its target loci.

1.1.5 HC11 mammary epithelial lactogenic differentiation model system to study interplay between signalling and chromatin dynamics

Lactogenic differentiation of mouse mammary epithelial cells is an important paradigm in understanding signaling induced genome dynamics. This model is robust and reconstitutes in vivo mammary gland development. For our studies, we have chosen a 3point differentiation system using HC11 cells as a representative of various stages of mammary glands development *in vivo*. HC11 cells are MECs derived from pregnant mice, COMMA-1D clone, which are responsive to GC and PRL hormones, thus widely used in lactogenic cellular differentiation system *in vitro* (Ball et al., 1988) . To understand 3-D genome organization in time and space in an unbiased manner, in-situ Hi-C is the method of choice. But the major conundrum in the field is how to account for the effect of the cell division on chromatin architecture and most studies ignore them. Different studies, using different methods highlights different views of chromosome folding (Nagano et al., 2017, Naumova et al., 2013). We now know that TADs are established at the G1 phase and to really understand the mechanistic folding pattern of interphase nuclei, most of the Hi-C studies need to be done at G1 phase. We reasoned that as most cell synchronization methods are physiologically irrelevant, they do not always reconstitute folding principles as expected in-vivo. HC-11 Cells differentiate in response to external signaling in time dependent manner and after reaching confluency, they donot divide with most of the cells remaining at G1 phase (Merlo et al., 1996). Thus, we reasoned that this growth arrested cells in G1 phase under normal cell culture conditions can be a good model to study genome dynamics. Moreover, there is another side of the story. Glucorticoids signaling is necessary for various physiological processes like maintaining homeostasis. Clinically, synthetic variants of this hormone are widely used in treating multiple diseases including autoimmune disorders, hematological cancer and inflammatory diseases. Till date, very few studies are done on the overall effect of glucorticoid signaling genome- wide and possible role of 3-D genome dynamics in

establishing stage-specific gene networks in response to GR signalling. In this thesis, I have tried to address the mechanistic folding principle, its relation with evolving gene networks mediated by GR in mammary gland developmental context.

1.1.6 Mammary Epithelial differentiation and the role of signaling hormones

Mammary gland is an apocrine gland and its key function is to produce and secrete milk. The basic components of adult mammary gland are alveoli, which are lined with milk-secreting mammary epithelial cells (MECs). Myoepithelial cells and stromal compartment that are derived from embryonic ectoderm and mesoderm respectively surround these epithelial cells. Several alveoli join to form lobules that have a lactiferous duct, which drains milk into opening of the nipple (Macias and Hinck, 2012). Milk is considered as a complete diet as it contains all essential components such as proteins, fats, carbohydrates and other micronutrients required for proper growth and maintenance of the infant body. The process of lactation is highly conserved throughout the mammalian species. Much of our current understanding of morphological, cellular and molecular changes that accompany mammary gland development and differentiation comes from the studies on mouse system.

In mice, development of mammary gland initiates during E10-11.5 of embryonic state, with the formation of bilateral stripes on either side of abdomen of developing embryo, which further organizes into placodes. During puberty, these placodes differentiate into multiple mammary trees and terminal end buds under the influence of estrogen and insulin like growth factor (IGF) (Macias and Hinck, 2012). Further development of mammary gland triggers at virgin stage and continues through early, mid and late pregnancy till the end of parturition under the influence of various lactogenic hormones including epidermal growth factor (EGF) (Hynes and Watson, 2010), glucocorticoids (GC) (Wintermantel et al., 2005) and prolactin (PRL) (Kelly et al., 2002). There are a few studies that define the cellular and molecular mechanisms responsible for protein production and lactose content of milk. The known major cue that defines the protein content of milk is signal transduction by Jak-Stat pathway, with Stat5 playing the central role (Yang et al., 2000b). The levels or transcriptional activity of Stat5 in non-pregnant and non-lactating mammary gland is very less compared to pregnant or lactating mammary gland (Yang et al., 2000a). Prolactin has been shown to activate Jak-Stat5 pathway. In a few instances, it has been shown that Jak-Stat5 regulates primary milk protein, *Casein* gene transcription and therefore, mutation of Stat5 results in loss of *Casein* gene transcription (Schmitt-Ney et al., 1991). However, comprehensive picture of how PRL induced signaling pathways execute their effects on epigenome of mammary epithelial cells

for efficient milk production is yet incomplete and further studies are required for complete understanding of the core processes involved.

Interestingly, mouse *Casein* gene cluster is located at the nuclear interior or outside its Chromosome territory (CT) in a Lactogenic hormone stimulated mouse epithelial cells, while in non-stimulated or non-mammary cells residing mostly at nuclear periphery or inside of the CT (Ballester et al., 2008). Moreover, it has been documented that milk protein coding genes are in open chromatin conformation during lactation, but in earlier stages of development and as well in other non-mammary tissues, they are in closed conformation (Kress et al., 2010). It is now known that gene repositioning within the cell nuclei in response to external signaling cues is essential for establishing stage specific gene networks.

The epithelial cells of mammary gland undergo profound morphological and functional changes in accordance with the developmental window of puberty, pregnancy and lactation (Watson and Khaled, 2008). These changes occur in conjunction with stage specific signalling events and are accompanied by differential gene expression. Several epigenetic phenomena such as DNA methylation, histone modifications occur upon extra cellular signaling and act in concert in regulating cell-type specific chromatin organization (Mattout et al., 2015). Recent studies indicate that cell-type specific chromatin organization within the 3-dimensional space of the cell nucleus play a fundamental role in cell-type specific gene expression (Fraser et al., 2015b). Further, it has been shown that chromatin regions encompassing genes relocate to either silent or active compartment to establish signaling induced differential gene expression patterns. Further, it has been demonstrated that chromatin is spatially segregated to nuclear lamina enriched at nuclear periphery (van Steensel and Belmont, 2017). It has also been shown that during interphase, chromatin is constrained to its territory, named as chromosome territories (CTs) and CTs are non-randomly arranged and their positioning and neighborhood are shown to be cell-type specific (Cremer et al., 1982). Further, interphase chromosomal chromatin is partitioned into hierarchical, self-interacting multi-mega-base length, sub chromosomal domains at micro, meso and macro scale referred to as Topological Associated Domains (TADs), meta and super domains respectively (Rocha et al., 2015, Dixon et al., 2016). TADs were shown to be cell-type invariant and evolutionarily conserved but a few studies suggest that TADs are dynamic and undergo an extensive rewiring to enable differential expression of genes as per given environmental cues (Nora et al., 2012, Dixon et al., 2012). Based on these compelling evidences, we hypothesize that different developmental signaling cues such as GC, PRL hormonal signaling lead to profound structural and functional changes at the level of

chromatin in mammary epithelial cells, which leads to differential clustering of group of active/silent genes to enable expression of genes that are involved in milk biosynthesis pathways. Based on this hypothesis, we set out to study interplay between spatial organization of genome and gene expression in normal mammary epithelial cells under EGF signalling and its dynamics under GC primed signalling alone and in combination with PRL hormone signalling.

Materials & Methods

1. **Cell culture and lactogenic differentiation of HC11 mammary epithelial cells**:

HC11 MECs cells were cultured and differentiated essentially as described earlier (Schwartz, 1988) Briefly, murine MECs stem-like cell line, HC11 with passage number 6 (obtained from Dr. Nancy E Hynes laboratory), were cultured with RPMI1640-GlutaMAX™ (Gibco) medium supplemented with 10% FBS (US Origin GIBCO), Insulin (5µg/ml, Sigma Catno#I6634), 1X Anti-biotic/Anti-mycotic (Gibco), and Epidermal growth factor (20ng/ml, Sigma Cat.no#E4127) until they reach confluence under 5% $C0_2$. Lactogenic differentiation of HC11 cells was initiated by withdrawal of EGF containing medium and replacement with fresh priming medium consisting of RPMI1640-GlutaMAX™complete medium with 10% FBS, Insulin (5µg/ml), Hydrocortisone (1µg/ml, Sigma #H4001) and were grown at 37°C for 48hrs under 5% $C0_2$. Samples for primed condition were harvested at this time point. Later, HC primed cells after 48hrs, were subjected to PRL treatment by supplementing with fresh RPMI1640-GlutaMAX™ complete medium with 5% FBS, Insulin (5µg/ml), hydrocortisone (1µg/ml) and prolactin (5µg/ml, NIH, #NIDDK-oPRL-21) and cells were then incubated at 37°C for 72hrs under 5% $C0_2$ to complete lactogenic differentiation. Cells were harvested after 72 hours for PRL stage.

2. **Culture of mouse embryonic stem cells (ESCs):**

Feeder dependent R1 mouse embryonic stem cells cultured according to published protocol (Bibel et al., 2007) with minor modifications. Initially, mEF were revived from cryo-preservation and cultured in 100mm 0.1% gelatin coated Petri dish, containing 10ml DMEM medium supplemented with 10% FBS and 1X Antibiotic/Antimitotic (Invitrogen) and were incubated under 5% CO_2 at 37°C till they were grown up to 60% confluence. At this stage, mEF were inactivated by supplementation with fresh DMEM medium containing 5% FBS and 2.5% Mitomycin-C and incubated at 37°C, under 5% CO_2 for 3hrs. Inactivated mEF were washed twice with 1X PBS. Mouse ESCs were revived from cryopreservation and were culture over inactivated feeder cells in presence of DMEM medium containing 10% FBS, 1X Antibiotic/Antimitotic, LIF (10µg/ml) and 2i inhibitors (MEK 0.5µg/ml: GSK: 1.5µg/ml), 1X Non-essential amino acids, 2mM L-glutamine, 50µM 2-Marcaptoethanol (Gibco#21985) incubated at 37°C under 7% CO_2 for 24-48 hours. To enrich ESCs for experimental studies ESCs were trypsin treated for 30 seconds and plated immediately after removal of trypsin from medium allows MEF attached faster than ESC allows to separate relatively pure population of ESC by repeating this step for 2-3 times. Collected ESCs were used for all the experiments used in this study.

3. Flow cytometry:

ESCs that were grown at 70-80% confluence, fully confluent normal, GC primed and PRL treated HC11 MEC cells were subjected to cell cycle analysis using Flow cytometry. All the samples were washed twice with PBS pH (7.8) to remove media and cell were collected by centrifugation at 3000 RPM for 2min. Cell pellet was re-suspended in ice cold 70% methanol and were stored at -20^{0}C for few hours. Later, cells were centrifuged and the cell pellet was re-suspended in 0.5ml of PBS (pH:7.8) solution containing DNase free RNase-A (500µg/ml) (Roche#11119915001) and incubated at 37^{0}C for 1hr. Later these cells were supplemented with 5µg (1mg/ml) of Propidium iodide (Sigma#P4170) and injected into BD Fortessa™ for cell cycle analysis.

4. RNA extraction, Library construction and sequencing:

Total RNA from ESCs, Normal HC11 cells, GC primed and PRL treated HC11 cells isolated by using TRIzol™ (Invitrogen # 15596026) according to manufacturer instruction. RNA was further purified using a commercially available kit (GCC Biotech # GR1003). 20µg of purified RNA from each sample was treated with 10 Units of DNAse1 (Roche # 04716728001) and were further purified by using G Sure cell culture RNA isolation kit. From each RNA sample, Ribosomal RNA was removed using Ribo-Zero kit (NEB#E6310L) and further mRNAs were enriched using Oligo (dT) beads. Illumina paired end library was prepared as per the NEBNext® Ultra™ RNA Library Prep Kit (NEB # E7530S). All the libraries were paired end sequenced using illumina HiSeq 2500 sequencing platform.

5. RNA-seq reads quality control and mapping:

Raw sequence reads in FASTQ format were further processed to remove Illumina adaptor sequences by using Trimmomatic. The resultant raw reads were compressed to .gz format and were deposited in GEO repository [GEO-Accession No-GSE107419**]**. After filtering**,** approximately 55-65 million paired end reads were processed for each sample. The reads were mapped to mouse reference genome (GRCm38/mm10) using TopHat2.1.0(Kim et al., 2013).The mismatch parameter was set to 2 and all other parameters to default.

6. Quantification and Identification of differentially expressed genes:

Quantification of the expression of different genes was done with the tool Sub-read version1.5.0 using the features count mode method. Briefly, Total read count for each gene was obtained and were then normalized to obtain FPKM (Fragments per Kilo base exon per million reds sequenced) values. An FPKM ≥ 1 was considered as threshold for determining whether a gene is expressed or not(Love et

al., 2014). For differential gene expression analysis, DESeq2 version 1.24 was used. Differential gene expression analysis involved steps like model dependent p-value estimation using negative binomial distribution and Wald test, and adjusted p-value estimation based on multiple hypothesis testing. A log2Fold change value of ≥1 is taken as cutoff to define upregulated and a value of ≤-1 to define down-regulated genes. The p-adjusted threshold value was set to 0.01 to maintain statistical significance. gg-plots package (R programming) was used to generate the heat maps representing the above data. Venn diagrams' representing the data were generated using Ugent web tool (http://bioinformatics.psb.ugent.be/webtools/Venn/).

7. Identification of differentially expressed TFs and ERs:

TFs in mouse were identified using the sequence specific DNA binding or transcription regulation terms in GO functional annotation (Ashburner et al., 2000). List of ERs and their coding genes were obtained from previous literature (Shipra et al., 2006, Fazzio et al., 2008, Gendler et al., 2008). Corresponding expression and log2 fold change values of TFs and ERs were extracted from the RNA-Seq data. Here, it should be however, note that some of the genes are represented both under TFs and ERs. Heat maps and Venn diagrams were generated as described in previous section.

8. Pathway analysis of differentially expressed genes:

Top up-regulated and down-regulated genes in ESC vs normal, normal vs Primed and Primed vs PRL were uploaded to DAVID database (Huang da et al., 2009) to derive the KEGG pathway enrichment and GO annotation data. All differentially expressed gene lists were uploaded in DAVID functional annotation from which KEGG pathway enrichment scores list was derived. Biocarta and Wiki pathways were obtained from existing databases (Liberzon et al., 2011, Pico et al., 2008). Gene set enrichment analysis was calculated similarly to the method published previously(Arun et al., 2016, Ideker et al., 2002) and log2fold change of each gene was calculated by comparing different samples. Mean log2 fold change of a pathway was calculated by summing the scores of individual genes. Gene-set enrichment score was calibrated against the background distribution, by using randomly sampled n (number of genes in a pathway) scores and calculating mean log2 score. This process was repeated over 10000 times. The mean and standard deviation of the sampling distribution thus obtained was used for correction of the original score.

9. Real-time RT-PCR analysis:

Differentially expressed genes including TFs and ERs from ESC, normal HC11 cells, GC primed and PRL treated HC11 cells were quantified by quantitative real-time PCR with the CFX96 Touch Real-Time PCR (Bio-Rad). DNase treated 1μg RNA was taken as input from each sample to synthesize cDNA using iScript cDNA synthesis kit (BioRad). cDNA template was diluted and PCR conditions were followed according to manufacturer protocol kappa syber mix (Kappa bio systems). *β-actin* was used as housekeeping internal control gene and normalized all genes with it. No template controls (NTC) were also included in PCR plate. All primers used in this experiment were designed to span exon-exon junctions in order to minimize genomic DNA contamination. Relative gene expression was calculated using the method $-2^{\Delta\Delta CT}$(Livak and Schmittgen, 2001). All the graphs were generated using Graph pad prism software. Significance was calculated using two-way Anova by using graph pad prism and significance was measured and represented as [*] on top of error bars. Gene specific primers used for this study were listed in Table 13.

10. Immunoblotting:

Cells representing ESC, normal, GC primed and PRL treated HC11 cells are scrapped from the plates and cell lysate was prepared by using RIPA buffer (Sigma; CatNo#R0278) after incubating them for 30min. Cell lysates were centrifuged to remove debris and clear supernatant was collected for subsequent analysis. Protein concentration was measured by BCA by taking BSA as a standard protein. A total of 30ug of protein form each sample was electrophoresed on 8% SDS-PAGE. After electrophoresis, the proteins were transferred to nitrocellulose membrane by electro blotting. The nitrocellulose membrane was then blocked in 5% non-fat milk for 1h and subsequently probed with the respective primary antibody dilution and secondary antibodies. Blots were then developed using the ECL Western blotting substrate. All antibodies used for this study were listed in Table 14.

11. In-nucleus Chromosome Conformation Capture (3C):

3C was performed essentially as described (Dekker et al., 2002, Tolhuis et al., 2002) with a few critical modifications. Briefly $1X10^7$ cells from each stage were harvested and were cross-linked with 2% Formaldehyde for 5min at room temperature. Glycine was added to a final concentration of 0.125M to quench the reaction. Cross linked cell pellet was re-suspended in freshly prepared cell lysis buffer (10mM Tris pH-8.0, 0.2% Nonidet P-40,10mM NaCl, 1 tablet of Protease inhibitor cocktail (Roche#) for 12 min on ice. Resulting intact nuclear pellet was washed once with 10mM Tris-Cl (pH: 8.0) and

finally re-suspended in 1.2X restriction enzyme buffer (NEB buffer 2). Un-crosslinked proteins were cleared by the treatment of 0.3% of SDS at final concentration. Reaction mix was neutralized by the addition of Triton-X-100 to a final concentration of 2% and 400U of CSP61 restriction enzyme (Fermatas# ER0211) was added and incubated at 37^{O}C overnight on rotation. Following centrifugation, Intra nuclear restriction enzyme digested nuclear pellet was further re-suspended in 100µl of 1.1X Ligase buffer containing 800U of T4 DNA Ligase enzyme (NEB) and incubated for 4 hrs at 16^{O}C following 30min incubation at room temperature. Nuclei were de-crosslinked by the addition of Proteinase K (10mg/ml) and incubation at 65^{O}C for overnight. Later, the nuclei were treated with DNase free RNAse A (500µg/ml) (Roche#) for 1hr at 37^{0}C. The mixture was treated with phenol: chloroform and the DNA was precipitated by the addition of $1/10^{th}$ volume of 3M Sodium acetate and double the volume of ethanol. Finally, the DNA pellet was air dried and re-suspended in appropriate amount of nuclease free H_2O. Concentration of the DNA was measure by using Qubit fluorometer.

12. in-Nucleus Hi-C assay:

in-nucleus Hi-C, developed by kurukuti and Vetrie was used as descried (Sofueva et al., 2013) with few minor modification. ~ 8-10 million adherent HC11 cells in T75 flask were cross liked directly by the addition of 2% Formaldehyde (Sigma#) final concentration and incubated for 5min at room temperature. During incubation period of fixation, cells were scraped and passed through 70µM cell strainer (Genexy#93070). 1M glycine was added at a final concentration of 0.125M to quench the formaldehyde reaction. Cells were centrifuged, washed once with 1XPBS (8.0) and the resulting cells were re-suspended in ice cold, freshly prepared lysis buffer (10mM Tris (8.0), 10 mM NaCl, 0.2% Nonidet P-40) supplemented with 1 tablet of protease inhibitor cocktail (Roche#) was added, and incubated on ice for 15 minutes. After incubation cells were centrifuged at 2000rpm for 10 minutes at 4° C. Supernatant was discarded and nuclei were collected in the 500µl of 1.2X NEB buffer 2 restriction buffer. SDS was added to a final concentration of 0.3%, and the nuclei were incubated at 37 °C for 1hrs. Triton X-100 was then added to the final concentration of 2% to sequester the SDS. Digestion was performed with restriction enzyme at 37 °C overnight. End filling was performed with the enzyme Klenow fragment (NEB) and out of four dNTP's one of the dNTP, dCTP was modified to biotin dCTP followed by ligation by using 1600 units of ligase (NEB) enzyme and incubated at 16°C for 4hrs. 30µl of Proteinase K (10mg/ml) for + ligase samples and 5µl for minus ligase samples

were added, and samples were incubated overnight at 65°C to reverse the cross-links. 10µl (500µg/ml) for + ligase samples and 3µl (500µg/ml) for – ligase samples RNaseA (Roche) was added to remove RNA contamination followed by incubation at 37°C for 1 hrs. DNA was extracted from all samples by using phenol/ chloroform extraction method. DNA was precipitated by Ethanol precipitation and finally pellet was dissolved in TE buffer. Hi-C DNA concentration was measured by Qubit and Integrity of DNA, restriction enzyme digestion, Ligation was assessed by agarose gel electrophoresis. Authenticity of biotin nucleotide incorporation followed by in-nucleus ligation was assessed by the restriction enzyme digestion with HindIII as well as Hha1 of a PCR product of generated by a pair of primers from two far distant locations of a actively transcribed *Ercc3* gene.

13. Hi-C Library Preparation:

Removal of biotin from un-ligated junction by T4 DNA polymerase by taking 5µg of DNA, 5U of T4DNA polymerase, 200µM of dATP and dGTP, 0.2mg of BSA at final concentration in 100 µl reaction volume for 2 hours at 12°C in thermomixer. DNA was purified after T4 DNA polymerase by phenol-chloroform separation followed by ethanol precipitation. 10µg of HiC DNA was taken for Covaris sonication and settings were setup to achieve 300-500bp range. Enrichment of fragments in the range of 300-500bp by Ampure bead separation then DNA was eluted from beads and small portion was run on 0.8% agarose gel and Q-bit quantification to observe fragment range. Biotin pull down was done by taking 50 µl of streptavidin beads M280 from Invitrogen was added to the DNA which was purified by Ampure beads and incubated at 37°C for 1 hour and beads were magnetized and supernatant was discarded. For further End repair we proceeded the beads with end repair mix from Trueseq kit (FC-121-2001). followed by A-tailing and Adapter ligation mix at 30°C for 30 minutes. Finally, beads were washed and mixed with PCR master mix and PE primers from illumina for 6 cycles. After PCR beads were magnetized and supernatant was collected and purified with AMPure-XP beads finally DNA was eluted with nuclease free water. Library was analysed by Bio analyser.

14. Preparation of metaphase chromosomes from HC11 cells and chromosome painting:

Highly proliferating HC11 cells were treated with Cholchicine (10µM) and continued to culture for 2hrs. Cells were harvested by trypsinization and collected cells in 0.56% hypotonic solution for 15min at 37^{O}C. Later, cells were fixe by methanol:Gacial acetic acid (1:3). Cell in fixative were dropped from six-inch distance on to glass slides to get a well spread mitotic chromosomes. Slides were air drier and

subsequently processed for chromosome painting. These metaphase chromosome slides were incubated with chr4 paint (XMP4-green (D-1404-050FI) (Metasystems) and chr8 paint (XMP8-Orange (D-1408-050-OR) combined and covered with coverslip. Slides were heated to 74^{O}C for 3min to denature the chromosomal DNA and allowed them to hybridize at 37^{O}C in a humidified chamber for overnight. Following day, slide was washed with pre-warmed 0.4X SSC at 72^{0}C for 2min. followed by 2X SSC with 0.05% Tween-20 wash for 5min. Slide was allowed to air dry and incubated with DAPI (1μg/ml) for 5min at room temperature. Later, the slides were overlaid with Gold antifade mounting medium (Invitrogen) and observed under fluorescence microscope (Zeiss).

15. DNA FISH:

DNA FISH was performed essentially as described. Isolated ESCs, normal, GC primed and PRL treated HC11 cells were harvested from T25 flasks by trypsinization. Cells were crosslinked with freshly prepared glacial acetic acid: methanol (1:3) for 10min at room temperature. Cross-linked cells were evenly spread on to glass slides that were initially treated with fixative and air dried. These slides were subjected to pre-treatment, hybridization and post-hybridization essential as described (Bolland et al., 2013). Briefly, crosslinked cells were initially permeabilized with buffer consisting of 0.1% Saponin + 0.1% Triton X100 in PBS for 10min, following incubation them with 20% glycerol for 20min, snap freeze and thawed three times in liquid nitrogen followed by 5min of PBS wash. Further, cells were treated with 0.1M HCl for 30min and washed in PBS for 5min followed by second round of permeabilization with 0.1% Saponin + 0.1% Triton X100 in PBS for 30min. Cells were then equilibrated with 50% formamide and 50% 2X SSC for 10min at room temperature.

DNA probes were representing cLADs from Chr6: 102,777,358 -102,985,022 (**L0**; Clone RP23-77B7, labelled with Alexa Fluor-555: Red) and Chr6: 45,804,687- 45,966,460 (**L2**; RP23-77B7, labelled with Alexa Fluor-488: Green) and ciLADs from Chr6: 100,781,363-100,975,441 (**iL0a**; RP23-278J23, labelled with Alexa Fluor-647: Magenta) and Chr6: 112,359,511-112,583,069 (**iL0b**; RP23-40E10, labelled with Alexa Fluor-594: Cyan) were labelled by nick translation followed by chemical conjugation of the amino-modified nuclotides with activated fluorophores using FISH Tag DNA Multicolor Kit- (Invitrogen# F32951). Hybridization and post hybridization were performed with respective probes following protocol by Daniel et al., (2013) The purified probe was further used for hybridization onto the ESC, normal, GC primed and PRL treated HC11 cells at specific stages of lactogenic differentiation based on paper published in Jove by Daniel et al. 2013 with minor

modifications. Hybridization was performed with respective probe in hybridization buffer (% Formamide, % Dextron sulfate in water) were added to the cells on slide and covered with coverslip and sealed with rubber cement. At this stage slides were kept at hot plate at 78^{O}C for 3min to denature the DNA and were incubated at 37^{O}C for overnight in a humid chamber. Later, cells were washed with 0.4X SSC at 72^{0}C for 5min followed by 2X SSC and 0.05% Tween-20 at room temperature for 1min. Slides were allowed to air dry and stained with DAPI for 5min. Later slides were washed with water, slightly air dried and then mounted with ProLong Gold antifade reagent (Thermo Scientific). Cells were covered with coverslip, sealed with nail paint and were observed under confocal laser microscope (LSM710) at 63X (oil immersion). Z-stack images of thickness were taken and images were processed by using Zen 2.3 Lite software. To assess the interaction between LADs, BAC probes L0 labelled Alexa Fluor 555: red, L2-Alexa Fluor 488: green, we counted total 226 (ESC), 145 (Normal HC11 cels), 228 in (GC primed) and 270 (PRL treated HC11 cells). To assess the interaction between iLADs, BACs probes iL0a-Alexa fluor 647: magenta, iL0b-Alexa fluor-594: cyan was counted a total of 410 cell in ESC, 319 in Normal, 231 in GC primed and 300 in PRL treated HC11 cells. Likewise, to assess the interaction between cLAD-interLAD, BA probes labelled with L2-AlexaFluor 488: green, iL0a-Alexa Fluor 647: magenta was counted total 224 number in ESC, 296 in Normal, 110 in Primed and 270 in Prolactin. Significance of interaction among different states and graphs were calculated by using one-way ANOVA with the help of Prism 7 software. All the images were compiled by using Adobe Illustrator and Adobe Photoshop.

16. Scoring gene expression patterns in iLADs and cLADs:

List of mRNA genes were retrieved from mouse GRCm38/mm10 assembly from UCSC genome browser. Coordinates of Conserved LADs, Conserved inter-LADs and facultative-regions were retrieved from available data sets (Peric-Hupkes et al., 2010). Based on the coordinates of genes, we assigned genes to either conserved ci/cLADs. All the genes that have a minimal FPKM value were considered as list of genes expressed in ESC, normal, GC primed and PRL treated HC11 cells.

17. Hi-C data analysis:

I) Mapping to the reference genome:

Two biological replicates each of ESCs, normal, GC primed and PRL treated HC11 cells Hi-C libraries were sequenced using illumine paired end sequencing. Data with one mismatch was taken for the further analysis. All of the raw reads of replicates were processed separately and later replicates were

merged. The raw reads were inspected for quality check using "FastQC". Overrepresented adapter sequences were trimmed using "cutadapt" (http://dx.doi.org/10.14806/ej.17.1.200). "HICUP" (Wingett et al., 2015) was used to truncate, map, filer and perform deduplication of raw reads. Briefly, HICUP uses PERL scripts to truncate reads at the putative Hi-C ligation junctions, aligns read pairs independently to a reference genome using Bowtie2, filters out commonly encountered Hi-C artefacts and removes (retaining one copy) putative PCR duplicates (deduplication).

II) Derivation of contact frequency maps:

The mapped SAM/BAM files were used to generate contact frequency matrices using HOMER tools (http://homer.salk.edu/homer/interactions/index.html) (Heinz et al., 2010). Using HOMER, the genome is divided into bins of equal sizes and contact frequency between each bin was calculated and represented as a contact matrix. We have generated matrices at different resolutions (1MB, 100kb, 40kb and 10kb) as per requirement in raw and SimpleNorm formats. SimpleNorm outputs the ratio of observed to expected interactions to normalise them for sequence depth between the regions. All the contact matrices were visualized using Java TreeView software (http://jtreeview.sourceforge.net/) (Heinz et al., 2010). Further, to observe the loss of gain of contacts between the samples, the contact matrices were subtracted from one another in R, which were visualized in JavaTreeview.

III) Identification and substitution for chromosomal translocations:

To visualize chromosomal translocations, we normalized the genome-wide raw contact matrix to the number of restriction enzyme (HindIII) sites present in each bin as previously described (Harewood et al., 2017)). Briefly, each element in the raw contact matrix was divided by square root of the product of number of restriction enzyme sites in the 1Mb region of corresponding row and column of the element. The normalized matrix shows enriched regions of inter-chromosomal contacts and were visually identified. The coordinates of the identified translocations are in the Table 15B. The values corresponding to the translocation regions in the contact matrix were replaced by the mean inter-chromosomal contact frequency for each matrix. Subtraction matrices were generated by directly subtracting one with the other matrix in R and are visualized using JavaTreeview.

IV) Generation of whole genome chromosome neighbourhood maps:

We used Vanilla coverage normalization to normalize within sample inter-chromosomal contact matrices as previously described (Rao et al., 2014). Briefly, intra-chromosomal contacts were removed

in raw genome wide matrices, in which the translocated regions were replaced by mean inter-chromosomal contact frequencies as mentioned earlier. Then all the values of the matrix corresponding to an inter-chromosomal pair were summed up and divided by the product of lengths of the corresponding chromosomes of that pair to account for contact frequency biases due to different lengths of the chromosomes. This normalized matrix (20 X 20) is termed chromosome neighbourhood matrix. Spearman correlation between the neighbourhood maps of cell-types was calculated using R. Paired Wilcoxon sign rank test was performed to find the statistical difference between the neighbourhood maps of between cell-types. Subtraction matrices were generated all these matrices were visualized using GI Tools (Perez-Llamas and Lopez-Bigas, 2011) where mean is taken as middle value during heat map generation. The clustering analysis was done, as previously described (Barutcu et al., 2015), with some modifications. Briefly, ratio matrices of VC normalized contacts were calculated between cell types (N/ES, P/N and PRL/P). For each ratio matrix, mean interactions between the set of chromosomes (Chr 8, 9, 10 and 11) and their mean interactions with rest of the chromosomes were calculated and were plotted as Bar plot in R. The same was repeated for the smaller chromosome set (Chr 15, 16, 17, 18 and 19). Unpaired Wilcoxon test was used to calculate the statistical significance. Hierarchical clustering of CT on both row and columns was done using the package 'pheatmap' in R using complete clustering method and using default parameters (Sarnataro et al., 2017).

V) Evaluation of long vs Short-range interactions:

The chromatin interactions between any two 100kb bins separated by 2Mb was denoted as short-range interaction and above 2Mb as long-range in *Cis* interaction matrix. LvS was calculated as the ratio of sum of all short-range interactions matrix to the sum of all the long-range interactions in a *Cis* interaction as described previously (Criscione et al., 2016)with few modification. These values were calculated for all chromosome *Cis* matrices and were represented as bar charts generated in Microsoft Excel. All individual chromosomal values were averaged to derive whole genome LvS ratio for the four cell-types.

VI) Generation of distance decay plots:

Each element in the SimpleNorm contact frequency matrix was antilog2 transformed and then log10 transformed, since HOMER outputs log2 transformed values. The mean interaction frequency of all

the contacts at a given distance (in multiples of 100kb) were calculated and were plotted as line plots to get the distance decay plots. 'Geom' smooth function in 'ggplot2' package in R (Wickham, 2009) was used to smoothen the lines and the smoothening method was put to default. To avoid severe outliers in every matrix, the contact frequencies between bins separated only above 8Mb were considered for this decay analysis.

VII) Principle Component Analysis (PCA1) at 100kb resolution:

Correlation matrices of each intra-chromosomal contact matrix were generated using HOMER. Briefly, for each element in the matrix, HOMER calculates the Pearson correlation between all the elements in its row and all the elements in its column. From the correlation matrices at 100kb resolution, HOMER calculates the first Eigen vector and generates PCA (Principal component analysis) of each intra-chromosomal contact matrix. In the first PCA, the positive PCA components represents active/A compartments and negative PCA components represent repressed/B compartment.

VIII) Relationship between Compartment changes A-B / B-A 100kb and gene expression through Box Plot:

Genomic bins corresponding to A and B were separated in each cell-type and expression (FPKM values) of the genes present physically in those genomic bins were derived. The gene expression distribution is visualized as a boxplot, using R. Similarly, the bins which shifted their compartments were separated and their corresponding gene expression was derived. The gene expression distribution was plotted as a boxplot with no colour.

IX) LAD Analysis:

The boundaries of cLADs, ciLADs and cfLADs were downloaded from GSE17051 (Peric-Hupkes et al., 2010). These mm9 coordinates were converted to mm10 using LiftOver tool in UCSC genome browser. Then the bed files were uploaded in UCSC genome browser to visualize the c/ciLADs along with other datasets bed files such as PCA1, Gene expression (FPKM ratio), GC/AT content, Gene density and chromosomal banding patterns from four cell-types studies.

X) Evaluation of intra-chromosomal c/ciLAD interactions:

To measure the contact frequencies, intra-chromosomal c/ciLADs larger than 400kb were considered for analysis and all the c/ciLADs were rounded off to nearest 100kb value. For each *Cis* matrix

(100kb), three categories of sub-matrices were derived from the upper triangle. The submatrices corresponding to inter-LAD –LAD, inter iLAD-iLAD and inter LAD-iLAD interactions were retrieved using R. All the values are concatenated to a single vector. Student t-test was performed on values of all the submatrices of a chromosome corresponding to a category between the cell-types to measure the statistical differences. P-value less than 0.01 was considered significant. For each chromosome, mean of the contact frequencies of a category was calculated for all four cell-types and were represented as Bar plots.

XI) Evaluation of inter-chromosomal c/ciLAD interactions:

To derive correlation between inter-chromosomal cLAD-cLAD, cLAD-iLAD and iLAD-iLAD interaction, values corresponding to the translocated regions were excluded from the analysis. Genome-wide *Trans* matrices were generated at 250kb resolution. c/iLADs larger than 250kb were taken for the analysis and all the c/ciLAD boundaries were rounded off to nearest 250kb bin. Then for each genome-wide *Trans* matrix, three categories of sub-matrices were derived from the upper triangle as mentioned above. Paired student's t-test was performed on the values of all the sub matrices of a category between the cell-types to measure the statistical difference. Mean of the contact frequencies of each category was calculated for four cell-types and were represented as Bar pots. To analyze the interactions in translocated regions, the three categories of submatrices were generated inside of these regions and were compared to Trans interactions of each chromosome of the pair with rest of the genome for each category. Student t-test was used measure the statistical difference. All the above analysis was performed using 'R'.

XII) Evaluation of translocated chromosomal c/ciLAD interactions:

Inter-chromosomal interaction matrix that was generated by vanilla coverage normalization (250kb) was utilized to analyze interactions of c/ciLAD between translocated chromosomes 1:6 and 4:8 for each cell-type including ESC. For chr1-6 translocation, two submatrices were derived. A matrix having translocated segment of Chr1 on one axis and whole chr6 on another axis was derived. A second matrix was derived with non–translocated region of Chr1 (similar length as that of translocated part) and whole Chr6. Then four possible combinations of the interactions (ci-ciLAD,ci-cLAD,c-cLAD and c-ciLAD) between LADs are derived from both the matrices and were represented as Bar plots. Similar analysis was done on chr4-8 pair, where translocated and non-translocated regions were taken from chr4, and translocated region of Chr8 to derive the matrices. Student t test was used to calculate

the statistical significance.

XIII) Generation of hybrid-heatmaps:

The cis matrices of each chromosome of a translocated pair (Chr4:8) are taken at 1st and 4th quadrants and the corresponding trans matrix was taken at 2nd and third quadrant to derive the hybrid matrix of a translocated pair. The interaction values (au) are plotted as heatmap in Java Treeview.

XIV) Gene expression analysis in the LADs:

The genes and their expression present in cLADs and iLADs were derived using In-house PERL script and FPKM from RNA-Seq analysis. Wilcoxon rank sum test was used to determine the statistical significance between two samples and R was used to draw the boxplots. DAVID tool was used to derive the pathways associated with list of genes.

XV) Statistical correlation between cLADs, ciLADs and PCA derived genomic compartment A and B: To evaluate the relation between the cLADs, ciLADs and A, B genomic compartments, PCA at 20kb resolution was generated for four cell-types using HOMER. For each segment of cLAD region, we counted number of A and B compartments that are physically located within the coordinates separately. If the end or start of the c/ciLAD spanned over more than half of the 20kb bin, then the bin was considered to be part of the c/ciLAD. The number of A/B compartments were generated separately from randomized PCA and were calculated from the same sample for the same set of c/ciLADS. Paired Wilcoxon rank sum test was used to see the statistical difference. The same was repeated for ciLADs and the boxplots were drawn using R.

XVI) Derivation and analysis of Topologically Associated Domains (TAD):

Topologically Associated Domains (TADs) were generated using two tools. To identify smaller and nested TADs, we used TADtree on simpleNorm Cis matrices (100kb resolution) by setting the parameters in Table 20. TAD is defined as conserved between two cell-types if the boundaries of the TAD in the two cell-types overlap each other with some relaxation criteria mentioned in (Table 20). To generate general TADs, HOMER was used. HOMER uses directionality Index (DI) concept to generate the domains boundaries. To derive the conserved TADs, the same criteria as described above was used. Venn diagrams representing the data were generated using Ugent web tool (http://bioinformatics.psb.ugent.be/webtools/Venn/).

XVII) Whole genome analysis of TAD dynamics:

To deduce such TAD to sub-TAD interplay, we derived the sub-TADs with in a TAD boundary and vice versa between cell-types. If the boundaries of a TAD from one cell-type were falling under boundaries of a TAD in second sample, the first TAD was considered as sub-TAD of second TAD and vice versa. The relaxation criteria taken to derive conserved TADs were used here as well. The Conserved TADs are deleted from the list before this analysis Note that we are not deriving sub-TADs within a TAD boundary in the same cell type. TADs derived using HOMER didn't show any TAD to sub-TAD transitions and vice versa.

XVIII) Hierarchical clustering of TADs Hierarchical clustering of TADs:

To compare the hierarchical clustering of TADs between the cell-types was performed using algorithms described previously (Weinreb and Raphael, 2016) with few modifications. TADs generated using HOMER, were taken for the analysis. Since only conserved TADs could be taken for analysis, we relaxed the parameters for defining conserved TADs as in Table 40. Conserved TADs were derived between the pairs for all the cell-types (i.e., ESC-N, N-P and P-PRL). For a pair of TADs T1 (start:x1, end:x2) and T2 (start: y1, end: y2), a score was calculated by summing up the inter-TAD and intra-TAD interaction frequencies in the submatrices matrix [x1:x2, x1:x2], matrix [y1:y2, y1:y2], matrix [x1:x2, y1:y2] and matrix [y1:y2, x1:x2], and dividing it by the length of both TADs and linear distance (with respect to the resolution of the Hi-C matrix) between those TADs.

$$\text{Score} = \frac{\text{sum(mat[x1:x2,x1:x2])} + \text{sum(mat[y1:y2,y1:y2])} + \text{sum(mat[x1:x2,y1:y2])} + \text{sum(mat[y1:y2,x1:x2])}}{\text{i. (x2-x1) * (y2-y1) * (linear distance in terms of HiC resolution)}}$$

A symmetrical score matrix was built by calculating scores among all possible combinations of TAD pairs. Then, a distance matrix was generated by the function 'dist' with default parameters. Function 'hclust' was used to do hierarchical clustering to generate dendrogram and cophenetic correlation was calculated by function 'cophenetic' from the object created by 'hclust'. All the functions were derived by using R (version 3.4.2)

XIX) Generation of sub-chromosomal domain contact matrices:

Genes, each of that are shown to be significantly up-regulated between ESC vs N (*Etv4*, *Krt42*), N vs P (*Krt23*, *Wfdc18*) and P vs PRL states (*Csn2* and *Wap*) (Table 19), were chosen. For Circos plots, cis simpleNorm matrices of the respective genes were antilog2 transformed, and a virtual 4C kind of contacts were derived by taking the 100kb bin, with the gene, as bait and the interactions of the bin with all the bins of the chromosome. The interaction values only greater than 1 were filtered out and colour scale was applied to the ranges in the interactions, as shown in the fig. Circos version 0.69-6 was used for the analysis. For Arc plots, 1Mb regions of respective genes were derived from 10kb cis matrices, such that the length of the total 1Mb region includes the gene length and gene lies at the centre. Interaction value which is greater than 0 were represented as a single line in Arc plot so that the plot represents the total no of interacting bins rather than intensity of interactions. Diagonally sliced subtraction simpleNorm heatmaps were overlaid with UCSC tracks representing, PCA, TADs, GE, GC%, Gene density and chromosomal bands.

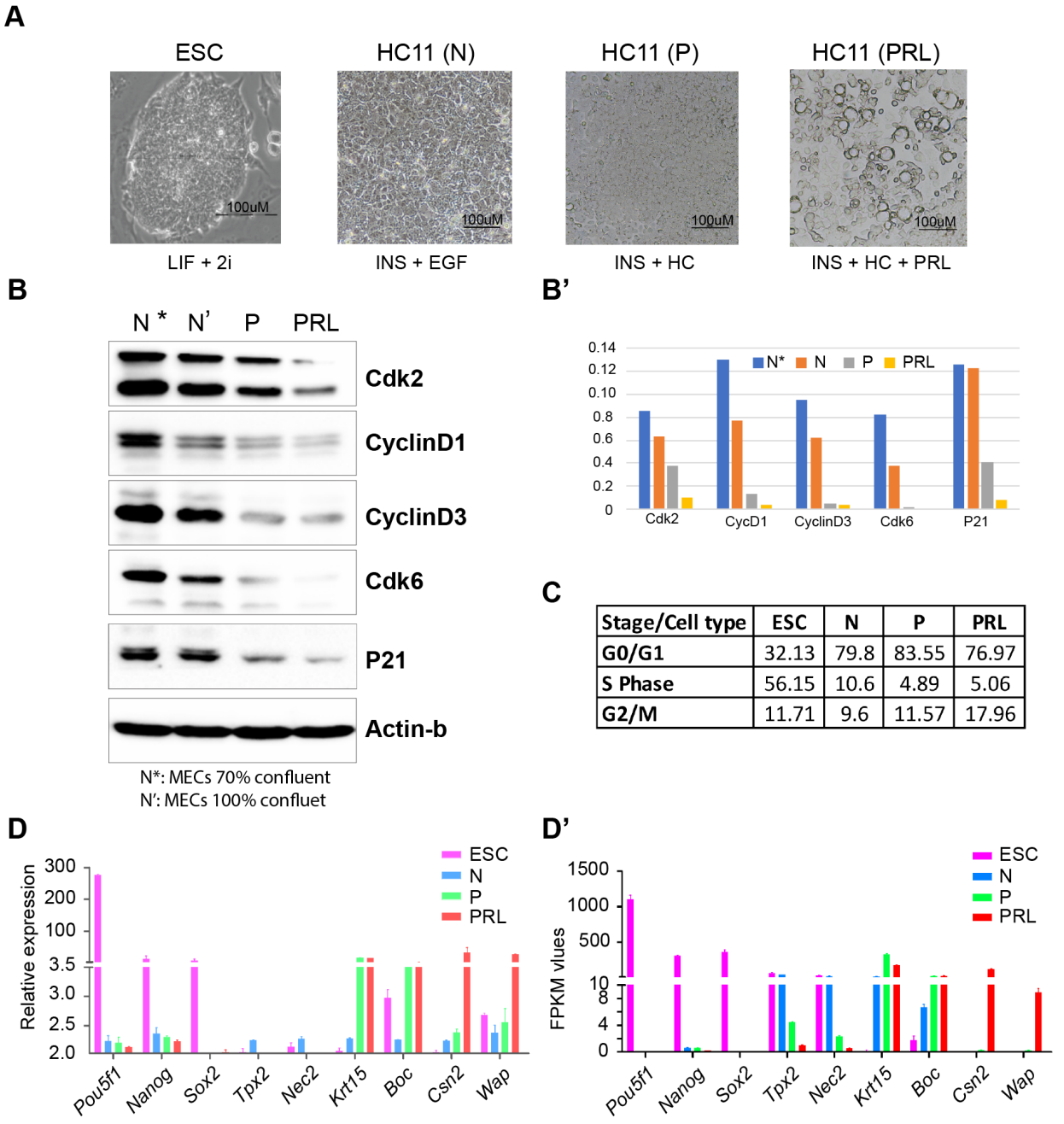
Results

1. **HC11 MEC undergo G0/G1 arrest during lactogenic differentiation:**

It is known that a regulated withdrawal from the cell cycle allows cells of any kind to differentiate. Therefore, to assess whether cell cycle arrest is accompanied by lactogenic differentiation of MEC, we performed lactogenic differentiation of MEC as described earlier(Morrison and Cutler, 2009). We confirmed the formation of mammospheres, which were more prominent after the treatment with PRL (Fig.1A). Cellular differentiation is associated with cell cycle withdrawal; therefore, we analysed the expression of cell cycle regulators by immunoblot analysis. Expression of cyclinD1 and D3, CDK2, CDK6 and p21 decreased in differentiated MEC compared to normal MEC and ESC (Fig.1B&B'). Cell cycle analyses revealed that majority of cells (77-84%) in MEC (normal & differentiated) were accumulated in G**0**/G1 stage, while most of the cells in ESC are in S-phase (Fig. 1C). Accumulation of MEC in G**0**/G1 stage and expression of cell cycle regulators (Fig.1B&C) suggests the progression to cell cycle arrest, which may facilitate differentiation. Further, we assessed the expression of specific markers pertinent to stage specific lactogenic differentiation. ESC expressed the putative markers such as *Pou5f1, Nanog and Sox2*. HC11 MECs expressed*Tpx2* and *Nek2*, whereas, HC11 primed cells expressed *Krt15* and *Boc* and PRL treated HC11 cells expressed *Csn2* and *Wap*. (Fig.1D). We assessed the expression of specific markers pertinent to lactogenic differentiation based on FPKM values obtained from RNA-seq analysis (See below) and found that identical sets of genes are induced during lactogenic differentiation of HC11 MEC (Fig.1D'

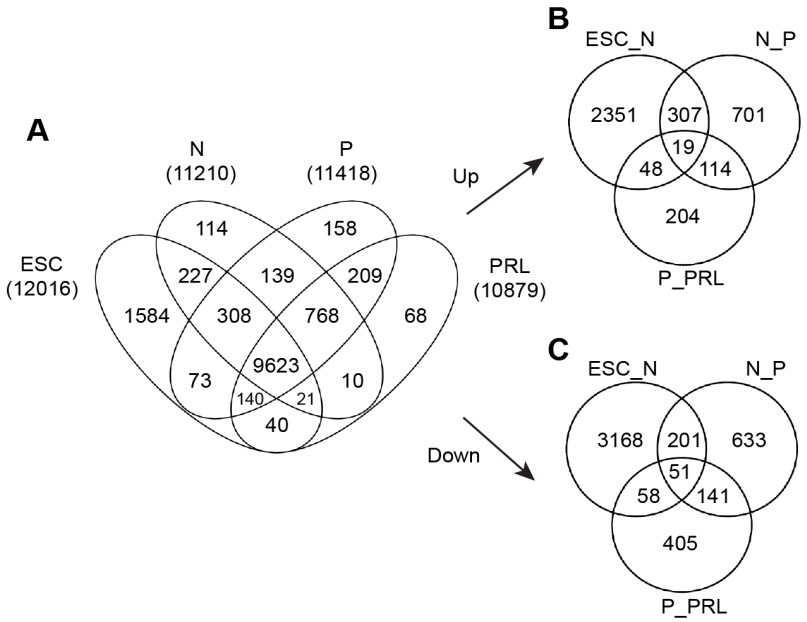
2. **RNA-seq analysis of ESC and differentiated HC11 MEC:**

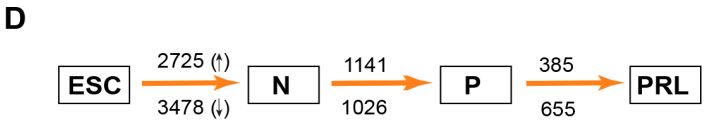
In order to quantify the changes in the expression levels of each transcript during lactogenic differentiation and to comprehensively understand the profile of all species of transcripts, we performed RNA-seq and analysed the data in ESC, normal MEC and MEC treated with GC and PRL. Qualitative analysis of RNA-seq data is provided in the (Table 1) .We compared the transcriptome profile of HC11 MEC with ESC after aligning them with GRCm38/mm10 mouse reference genome assembly. In case of ESC, 78% of total reads were uniquely mapped to the reference mouse genome, whereas ~88% of the reads were uniquely mapped in case of both normal and treated HC11 MECs employed (Table 1). Mixed datasets of highly correlated (>0.9) duplicate samples (Table 2) from ESC, normal, primed and PRL treated HC11 MEC were used to derive respective 'fragments per kilo base per million reads' (FPKM) values using Sub-read version 1.5.0.



Stage/Cell type	ESC	N	P	PRL
G0/G1	32.13	79.8	83.55	76.97
S Phase	56.15	10.6	4.89	5.06
G2/M	11.71	9.6	11.57	17.96

Fig.1. Characterization of HC11 MEC undergoing lactogenic differentiation. (**A)** Bright field microscopic images of actively growing ESC, undifferentiated HC11 cells at confluent stage in presence of EGF and INS and HC primed (P) alone and in combination with PRL are shown. Note the formation of clear dome shaped mammospheres under PRL condition. Scale bar represents 100μM. **(B)** Immunoblot analysis of cell cycle regulators in actively growing (N*), confluent stage undifferentiated normal (N') HC11 cells along with HC primed (P) and PRL treated cells showing gradual reduction in their levels in comparison with Actin-B. Full length blot ECL images are provided in supplementary Fig.S2. **(B')** Quantitative analysis of cell cycle regulator protein levels normalized against Actin-B showing gradual reduction in their levels during lactogenic differentiation. **(C)** Table showing percentage of ESC, N, P and PRL treated HC11 cells at G0/G1, S and G2/M phase of cell cycle showing Predominantly in S phase for ESCs and G0/G1 phage for rest of HC11 cell types. **(D)** Real time PCR analysis of cell-type specific gene expression analysis representing ESC, N, P and PRL treated HC11 cells. **(E)** RNA-seq data presentative FPKM values for the respective cell-type specific genes.


A
N
(11210)
P
(11418)
ESC
(12016)
114
158
227
139
209
1584
308
768
68
9623
73
10
140
21
40
Up
PRL
(10879)
Down
B
ESC_N
N_P
2351
307
701
19
48
114
204
P_PRL
C
ESC_N
N_P
3168
201
633
51
58
141
405
P_PRL


D
ESC
2725 (↑)
3478 (↓)
N
1141
1026
P
385
655
PRL

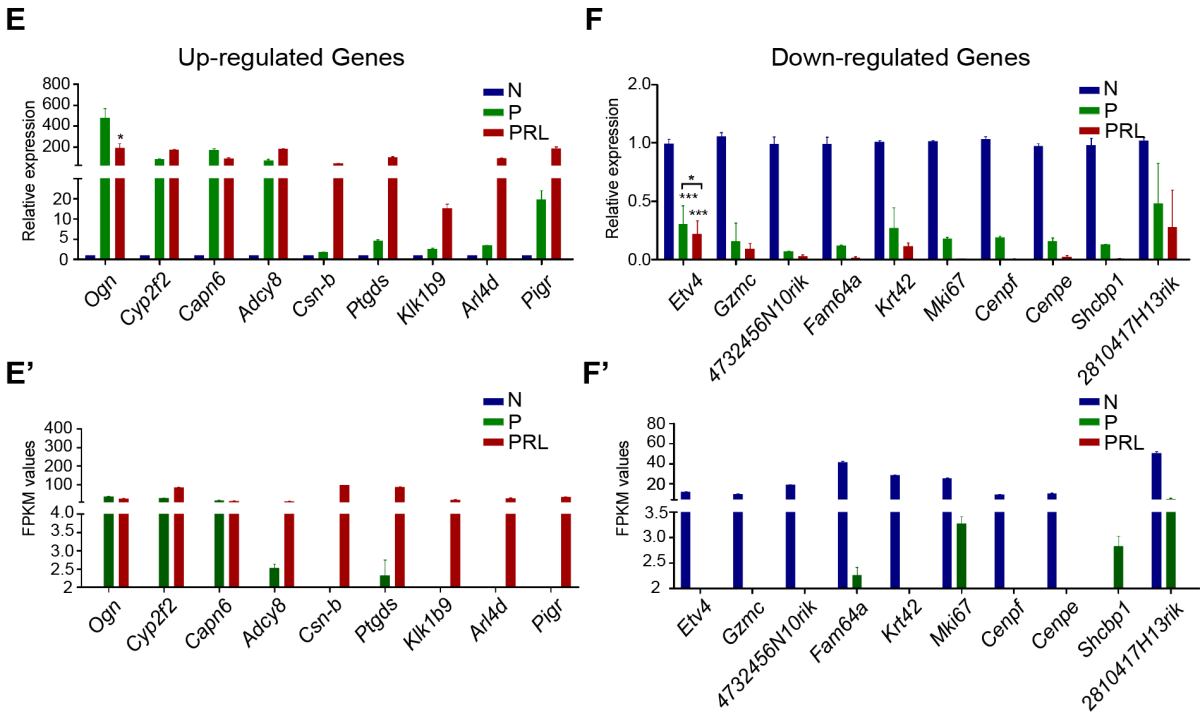

E
Up-regulated Genes
Relative expression
N
P
PRL
Ogn
Cyp2f2
Capn6
Adcy8
Csn-b
Ptgds
Klk1b9
Arl4d
Pigr
F
Down-regulated Genes
Relative expression
Etv4
Gzmc
4732456N10rik
Fam64a
Krt42
Mki67
Cenpf
Cenpe
Shcbp1
2810417H13rik
E'
FPKM values
F'
FPKM values

Fig.2. RNA-seq expression analysis of ESC and HC11 cells undergoing lactogenic differentiation and its validation by real time PCR. **(A)** Venn diagram showing expression of total, unique and overlapping genes (>1 FPKM value) in ESC, Normal (N), GC primed (P) and PRL treated HC11 cells. **(B)** Venn diagram showing unique and overlapping, upregulated genes between ESC vs N, N vs P and P vs PRL treated HC11 cells (Log2 >=1). **(C)** Venn diagram showing unique and overlapping, down regulated genes between ESC vs N, N vs P and P vs PRL treated HC11 cells (Log2 <=-1). **(D)** Schematic diagram showing total number of genes which are differentially up regulated (Up arrow) and down regulated (Down arrow) between ESC vs Normal (N), N vs GC primed (P) and PRL treated HC11 cells. **(E)** Real time PCR analysis of the top up-regulated and **(F**) downregulated in Normal (N), GC primed (P) and PRL treated HC11 cells and their respective RNA-seq FPKM values of up regulated genes and **(G)** down regulated genes **(H)**

We considered genes with a FPKM value of minimum 1, as expressed and found a total of 12016 genes expressed in ESC, 11210 in normal MEC, 11418 in GC treated MEC and 10879 with PRL treated HC11 MEC (Fig.2A). While a large number of genes (9623) express constitutively in ESC and MEC, relatively less number of genes are differentially expressed during lactogenic differentiation (Fig.2A). 1584 genes were selectively expressed in ESC, 114 in HC11 normal MEC, 158 in primed MEC and 68 in PRL treated MEC (Fig.2A). It is noteworthy that ESC express large number of unique genes compared to normal MEC and differentiated MECs (Fig.2A&3A). We further categorized highly expressed genes from ESC, normal, primed and PRL states based on FPKM values. Among them, genes with higher FPKM values were predominantly housekeeping and ribosomal genes. However, some of these genes were cell-specific, indicating that these might be induced selectively during GC and PRL treatment (Fig.2A&3B).

3. **Lactogenic differentiation induces differential expression of large number of transcripts:**

We next sought to understand the differentially expressed genes during various stages of lactogenic differentiation. We analyzed differentially expressed genes using DEseq2(Anders and Huber, 2010)A Log2 fold change of >=+1 and <=-1 values were considered as cutoff to define up and downregulation respectively, with p-adjusted value of <0.001, and. This analysis showed that 2725, 1141 and 385 genes were up-regulated and 3478, 1026 and 655 genes were downregulated when compared ESCs vs. normal, normal vs. GC and GC vs. PRL treated HC11 cells, respectively (Fig.2B, C&D). It should be noted that the total number of differentially expressed genes (both up and downregulated) were decreased from ESC vs. normal MEC to normal vs. GC treated MEC to GC vs. PRL treated MEC. Although, both GC and PRL are key promoters of lactogenic differentiation, a large number of genes were differentially expressed (upregulated: 1141; downregulated: 1026) upon GC priming alone when compared to both GC and prolactin treatment (PRL; upregulated: 385 and downregulated: 655). We selected a list of genes such as *Ogn, Cyp2f2, Capn6, Adcy8, Csn2, Ptgds, Klk1b9, Arl4d* and *Pigr* (upregulated), and *Etv4, Gzmc, 4732456N10rik, Fam64a, Krt42, Mki67, Cenpf, Cenpe, Shcbp1* and *2810417H13Rik* (downregulated) based on RNA-Seq data from all three types of MEC and validated their expression by RT-PCR (Fig.2E&F). Expression of each transcript by RT-PCR is in concurrence with respective FPKM values (Fig.2E'&F'). Top 20 uniquely expressed genes in ESC, N: Normal P: Primed and Prolactin treated MEC are represented as heatmap (Fig. 3A). Similarly, top 20 highly expressed genes (Fig.3B, Table 4), top 20 highly upregulated genes (Fig.3C, Table 5) and top 20 highly downregulated (Fig.3D, Table 6) genes between ESC, N, P and PRL treated MEC were represented as heat maps.

4. Lactogenic differentiation induces differential expression of large number of Transcription factors and epigenetic regulators:

It is generally considered that the expression of cell-type specific genes is under the control of TFs(Sharmin et al., 2016, Arvey et al., 2012) & ERs(Choukrallah and Matthias, 2014). Therefore, we analyzed TFs and ERs that were differentially transcribed during lactogenic differentiation of MEC. We observed that 2196, 1990, 2002 and 1998 TFs are expressed in ESC, normal, primed and PRL states of MEC, respectively (Fig.4A). On the other hand, 837, 797, 800 and 797 ERs were expressed in ESC, normal, primed and PRL states of MEC, respectively (Fig.4D, Table 7&10). It is noteworthy that 1894 TFs and 775 ERs were constitutively expressed in both ESC and MEC (Fig. 4A&D). Next, differential expression of these regulatory factors during various stages of lactogenic differentiation was analyzed. We observed that 315, 74 and 47 TFs (Fig.4B), and 91, 16 and 7 ERs (Fig.4E) are upregulated in ESC vs. normal MEC, normal vs. GC primed, and primed vs. PRL treated MEC, respectively. Similarly, we found that 546, 186 and 114 TFs (Fig.4C), and 229, 107 and 55 ERs (Fig.4F) are downregulated in ESC vs. normal, normal vs. GC primed and primed vs. PRL treated MEC, respectively. Although a large number of TFs and ERs are differentially expressed between ESC and MEC, priming of MEC with GC resulted in more differentially expressed TFs and ERs than that of treatment with PRL. Total number of up and downregulated TFs and ERs during lactogenic differentiation of MEC is summarized in Fig. 4G. Heatmaps for highly expressed, up-and downregulated TFs is provided in Fig. 5A-C. Heatmaps for highly expressed, up and downregulated ERs is provided in Fig. 5D-F.

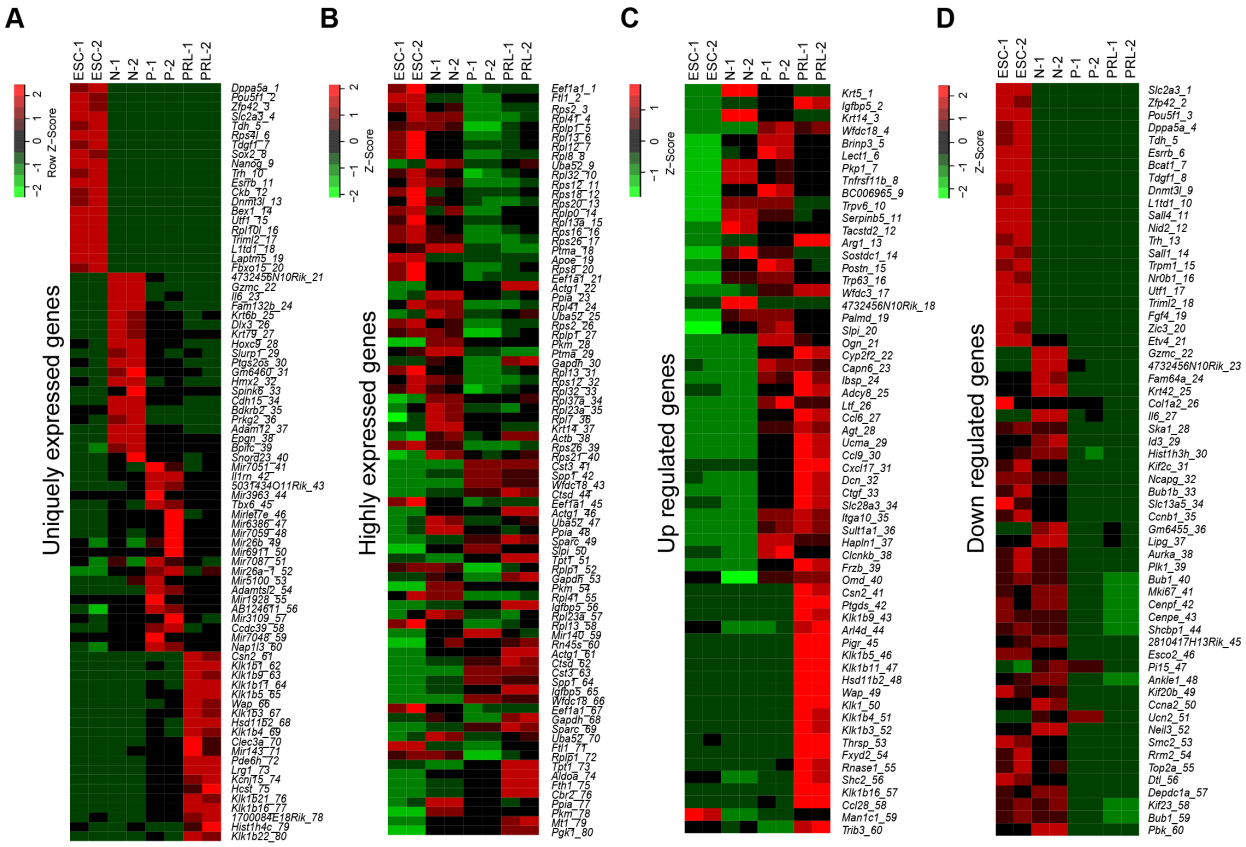

A
B
C
D
ESC-1
ESC-2
N-1
N-2
P-1
P-2
PRL-1
PRL-2
Row Z-Score
Z-Score
Uniquely expressed genes
Highly expressed genes
Up regulated genes
Down regulated genes

Fig.3. Comparative analysis of unique and differentially expressed genes in ESC and HC11 cells undergoing lactogenic differentiation. (A) Heat map showing top 20 uniquely expressed genes, **(B)** highly expressed, **(C)** Up-regulated and **(D)** Downregulated genes in replicates of ESC (ESC 1 & 2), Normal HC11 cells (N1 & 2), GC primed (P1 & P2) and PRL treated HC11 cells (PRL1 & PRL 2).

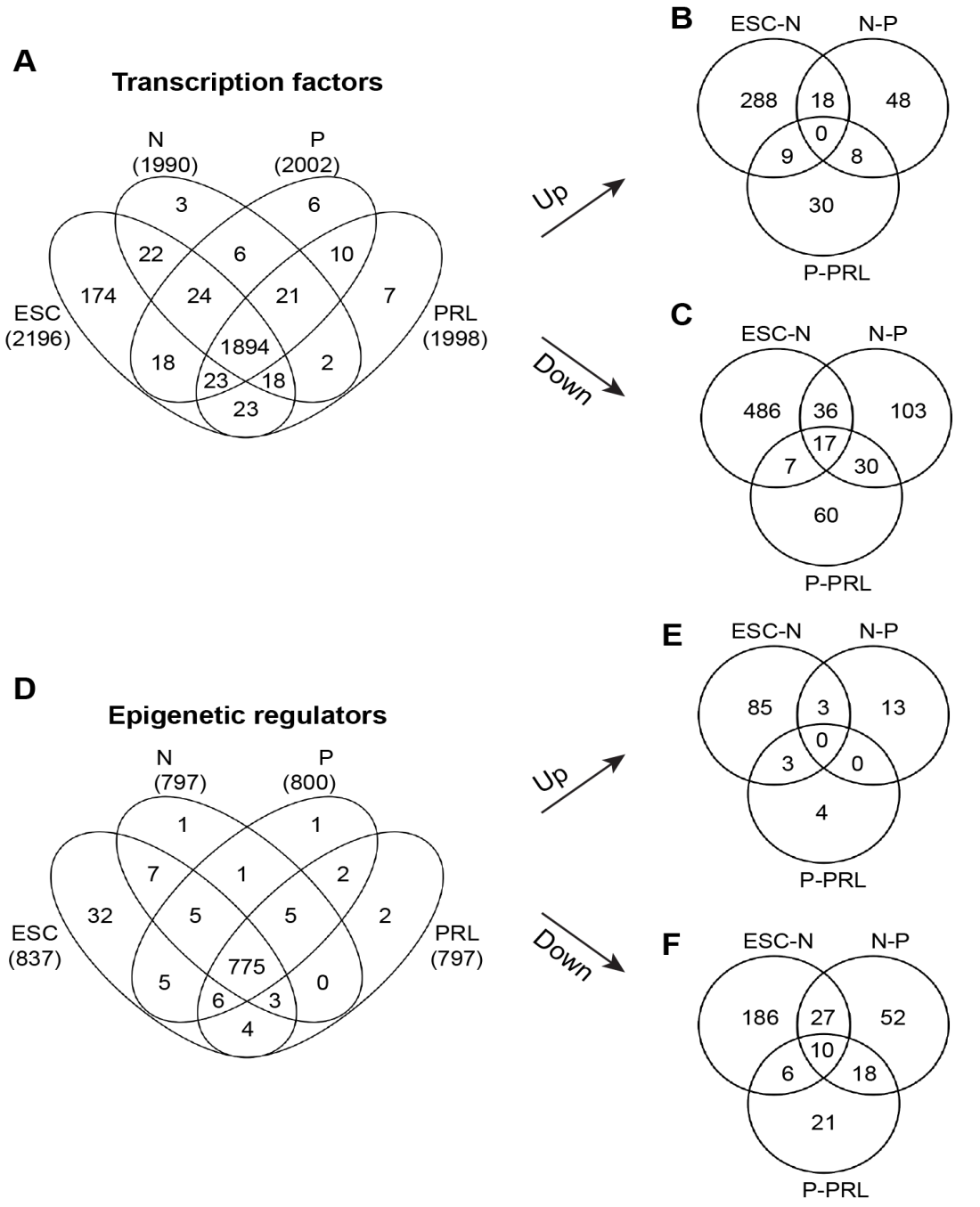

A
Transcription factors
N
(1990)
P
(2002)
ESC
(2196)
PRL
(1998)
3
6
22
6
10
174
24
21
7
1894
18
2
23
18
23
Up
Down
B
ESC-N
N-P
288
18
48
0
9
8
30
P-PRL
C
ESC-N
N-P
486
36
103
17
7
30
60
P-PRL
D
Epigenetic regulators
N
(797)
P
(800)
ESC
(837)
PRL
(797)
1
1
7
1
2
32
5
5
2
775
5
6
3
0
4
Up
Down
E
ESC-N
N-P
85
3
13
0
3
0
4
P-PRL
F
ESC-N
N-P
186
27
52
10
6
18
21
P-PRL
G
ESC
91
315 (↑)
546
229 (↓)
N
16
74
186
107
P
7
47
114
55
PRL
Epigenetic regulators
Transcription factors
Transcription factors
Epigenetic regulators

Fig.4. RNA-seq expression analysis of transcription factors (TFs) and epigenetic regulators (ERs) in ESC and HC11 cells undergoing lactogenic differentiation. **(A)** Venn diagram showing total, unique and overlapping expression of TFs between ESC vs N, N vs P and P vs PRL treated HC11 cells (>0 FPKM). **(B)** Venn diagram showing unique and overlapping expression of up-regulated (log2 >=1) and **(C)** down regulated (<=-1 FPKM) TFs between ESC vs N, N vs P and P vs PRL treated HC11 cells. **(D)** Venn diagram showing unique and overlapping expression of ERs between ESC vs N, N vs P and P vs PRL treated HC11 cells (>0 FPKM). **(E)** Venn diagram showing unique and over lapping expression of up-regulated (Log2 >=1) and **(F)** down regulated (Log2 <=-1) ERs between ESC vs N, N vs P and P vs PRL treated HC11 cells. **(G)** Schematic diagram showing total number of TFs/ERs which are differentially up regulated (Up arrow) and down regulated (Down arrow) between ESC vs Normal (N), N vs GC primed (P) and PRL treated HC11 cells.

5. Analysis of gene expression between ESC vs. normal HC11 MEC:

Upregulated genes: We analyzed the expression of genes, including TFs and ERs, which are differentially expressed in normal MEC, compared to ESC and performed pathway analysis. When compared to ESC, a total of 2725 genes were significantly up-regulated in normal MEC (Fig.2D). Among them, *krt5, Igfbp5, Krt14* and *Brinp3* were significantly upregulated (Fig.3C, Table 5). These genes encode for keratins, that are involved in extracellular matrix support, and Igf binding protein, which helps in the proliferation and survival of MECs(Allar and Wood, 2004). We also found that 315 TFs and 91 ERs are significantly upregulated in normal MEC compared to ESC (Fig.4G;). *Trpv6, Wfdc3, Barx2, Elf5* and *Ehf* (TFs), and *Dmp1*, *Ifih1, Sox9, Aim1* and *Sox8* (ERs) were significantly upregulated in normal MEC (Fig.5B&E, Table 8&11). We confirmed the expression of top upregulated TFs (Fig.6A) and ERs (Fig.6C) by RT-PCR. FPKM data of RNA-Seq for these TFs and ERs is comparable to that of RT-PCR data (Fig.6A vs.6A' & 6C vs. 6C'). Among TFs which are upregulated, *Elf5* is very specific to normal MEC and shown to regulate mammary stem or progenitor cell population(Choi et al., 2009, Chakrabarti et al., 2012). Based on upregulated genes, we performed KEGG pathway analysis and found several pathways activated in MECs including focal adhesion, proteoglycans and PI3k-Akt signaling pathway (Fig.7A). We also performed different pathway analyses like BioCarta and Wiki pathways which are shown in Fig.8.**Downregulated genes:** Similarly, we observed that about 3478 genes are downregulated in normal MEC compared to ESC (Fig.2D). The downregulated genes include *Nanog, Apoe, Akap12, Mybl2* and *Gm2373* (Fig.3D, Table 6). Most of the genes that were downregulated were found to be associated with the maintenance of stem cell state. We identified that 546 and 229 TFs and ERs are downregulated in MEC, respectively (Fig.4G). *Zfp42, Pou5f1, Esrrb* and *Dnmt3l* are among the top downregulated TFs (Fig.5C, Table 9), whereas *Sox2, Mtcl1, Ddx3y* and *Cdyl2* are top downregulated ERs (Fig.5F, Table 12). Among them, *Pou5f1 and Sox2* are very important for the maintenance of ESC pluripotency(Zhang and Cui, 2014, Hailesellasse Sene et al., 2007). The decreased expression of these factors in MEC suggests that these cells are deprived of pluripotency, which is required for differentiation. We confirmed the expression of top downregulated TFs (Fig.6B) and ERs (Fig.6D) by RT-PCR. FPKM data obtained from RNA-Seq for these TFs and ERs is comparable to that of RT-PCR data (Fig.6B vs.6B' & 6D vs. 6D'). KEGG pathway analysis showed that genes involved in insulin secretion, dilated cardiomyopathy, arrhythmogenic right ventricular cardiomyopathy (ARVC), cell adhesion molecules and hypertrophic cardiomyopathy (HCM) were downregulated in MEC compared to ESC (Fig.7B).

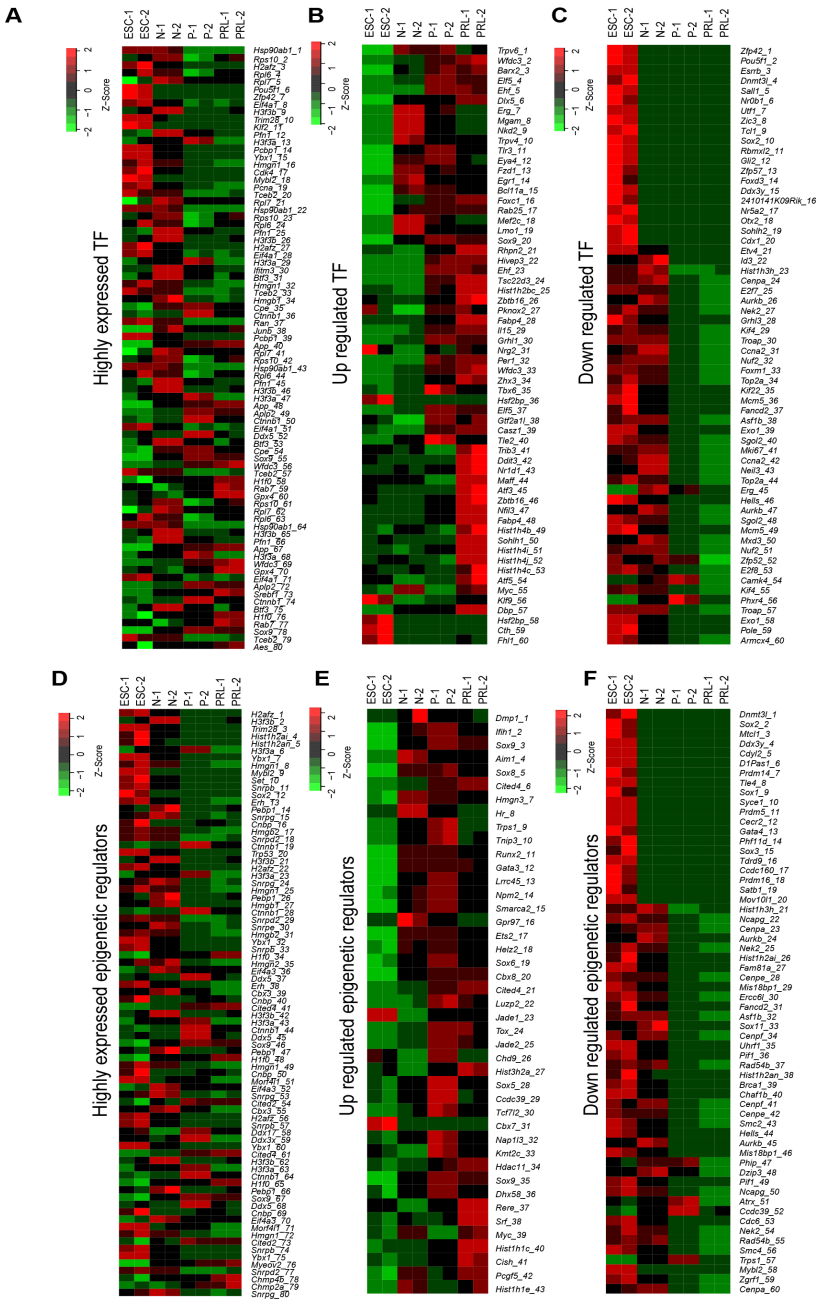


Fig.5. Comparative analysis of unique and differentially expressed TFs and ERs in ESC and HC11 cells undergoing lactogenic differentiation. (A, B & C) Heat map showing top 20 highly expressed, up-regulated and downregulated TFs and **(D, E & F)** ERs respectively in replicates of ESC (ESC 1 & 2), Normal HC11 cells (N1 & N2), GC primed (P1 & P2) and PRL treated HC11 cells (PRL1 & PRL 2).

6. **Analysis of gene expression in HC11 MEC during GC priming:**

   Upregulated genes: A total of 1141 genes were differentially expressed in MEC treated with GC compared to normal MEC (Fig.2D). Some of the upregulated genes include *Ogn, Cyp2f2, Capn6, Ibsp* and *Adcy8* (Fig.3C, Table 5). Interestingly, we found only 74 TFs and 16 ERs upregulated in GC treated primed vs normal MEC (Fig.4G). Upregulated TFs in GC primed HC11 MECs include *Rhpn2, Hivep3, Ehf* and *Tsc22d3* etc., whereas upregulated ERs include *Cited4, Luzp2, Jade1* and *Tox* (Fig.5B&E, Table 8&11). We have validated the expression of significantly upregulated TFs and ERs by RT-PCR (Fig.6A&C). The expression of upregulated TFs and ERs is in concurrence with FPKM data of RNA-Seq (Fig.6A'&C'). The epigenetic regulator and a transcriptional coactivator, *Cited4*, which is induced in MEC during GC priming, is known to be involved in milk secretion during pregnant and lactating mammary gland(Yahata et al., 2002). Further, genes shown to play role in chromatin remodeling and decompaction such as *Chd9(Ooga et al., 2018)* and *Myc(Kieffer-Kwon et al., 2017)* are significantly upregulated during GC priming (Fig5E, Fig6E). Most of the upregulated genes are known to be involved in pathways corresponding to protein digestion and absorption, gastric acid secretion, salivary secretion, glutamatergic synapse, melanogenesis and circadian entrainment (Fig.7C). **Downregulated genes:** We found a total of about 1026 genes that were significantly downregulated in HC11 MEC during priming with GC (Fig.2D), which include *Etv4, Gzmc, 4732456N10rik, Fam64a* and *Krt42 etc.,* (Fig.3D, Table 6). Moreover, we found that 186 TFs and 107 ERs were downregulated in GC primed HC11 MEC (Fig.4G). *Etv4, Id3, Hist1h3h, Cenpa* and *E2f7* are some of TFs that were downregulated during priming (Fig.5C, Table 9), whereas *Ncapg, Aurkb, Nek2, Hist1h2ai* and *Cenpe* are some of the downregulated ERs (Fig.5F, Table 12). The TFs *Etv4, E2f7* and *Aurkb* are known to be involved in cellular proliferation(Kurpios et al., 2009, Westendorp et al., 2012, Regan et al., 2013). Further, we validated the expression of top downregulated TFs (Fig.6B) and ERs (Fig.6D) by RT-PCR and decreased expression of these TFs and ERs during priming is comparable with FPKM data obtained from RNA-Seq (Fig.6B vs.6B' & 6D vs.6D'). The significantly downregulated genes were mostly involved in the cellular proliferation, DNA binding and cellular growth-related functions. Pathway analysis of the downregulated genes showed that pathways related to cell cycle, DNA replication, Fanconi anemia pathway, homologous recombination as well as pyrimidine metabolism are all downregulated (Fig.7D).

7. **Analysis of gene expression in PRL treated HC11 MEC: Upregulated genes:** To gain further insight into the changes in HC11 MEC during lactogenic differentiation, gene expression profile, pathway analysis and expression of TFs and ERs were analyzed in GC primed vs. prolactin+GC (PRL) treated

HC11 MEC. Our gene expression analysis revealed a total of 385 genes to be significantly upregulated in MEC treated with PRL vs. MEC treated with GC (Fig.2D). *Csn2, Wap, Pip, Ptgds, Klk1b9, Arl4d* and *Pigr* are some of the highly expressed genes in MEC treated with PRL (Fig.3A&C, Table 3&5). It is interesting to note that PRL triggers the expression of major milk proteins such as *Csn2*, *Wap* and *Pip*, which constitute upto 80% of total milk proteins. We noticed that 47 TFs and 7 ERs were upregulated with PRL treatment (Fig.4G). *Trib3, Ddit3, Nr1d1, Maff* and *Atf3* are some of the TFs upregulated with PRL treatment and *Rere, Srf* and *Myc* are some of the ERs highly expressed in PRL treated HC11 MECs (Fig.5B&E, Table 8&11). Upon validation of top upregulated TFs (Fig.6A) and ERs (Fig.6C) by RT-PCR, it was revealed that the RT-PCR confirms the FPKM data of RNA-seq (Fig.6A'&C'). *Trib3,* a TF that is associated with various ER stress responses, is upregulated in lactating period, which requires high demand for protein synthesis(Szegezdi et al., 2006). *Xbp1 is another TFs* that was upregulated during PRL treatment and is required for proliferation and differentiation of MEC(Davis et al., 2016). KEGG analysis revealed that PRL activates several other pathways that include endocrine and other factor-regulated calcium reabsorption, mineral absorption, renin-angiotensin system and biosynthesis of amino acids (Fig.7E).**Down regulated genes:** In PRL treated HC11 MEC, 655 genes were downregulated compared with primed MECs (Fig.2D). Genes such as *Mki67, Cenpf, Cenpe,* and *Shcbp1* were downregulated upon PRL treatment (Fig.3D, Table 6). We found that 114 TFs and 46 ERs, respectively are downregulated in PRL treated MEC (Fig.4G). Highly downregulated TFs are *Mki67, Ccna2, Neil3, Top2a* and *Erg,* and ERs are *Cenpf, Cenpe, Smc2, Hells* and *Aurkb* (Fig.5C&F, Table 9&12). We validated the expression of down regulated TFs and ERs by RT-PCR (Fig.6B&D) and is comparable with FPKM data of RNA-Seq (Fig.6B'&D'). KEGG analysis for downregulated genes revealed several pathways that are affected such as cell cycle, p53 signaling, fanconi anemia pathway, and focal adhesion (Fig.7F).

8. **Confirmation of specific factors expression during lactogenic differentiation:** In addition to mRNA expression analysis, immunoblotting for selected markers were performed to determine their expression during lactogenic differentiation of HC11 MECs (Fig.9A&B). Expression of Stat5a and 5b were increased during differentiation and particularly, the expression of Stat saturated during priming. Interestingly, expression of Dnmt3l protein level is increased during the course of differentiation of HC11 MECs in contrast to RT-PCR data and FRKM data (Fig.6). Furthermore, our data reveals that Hdac11 expression increased with PRL treatment (Fig.9).

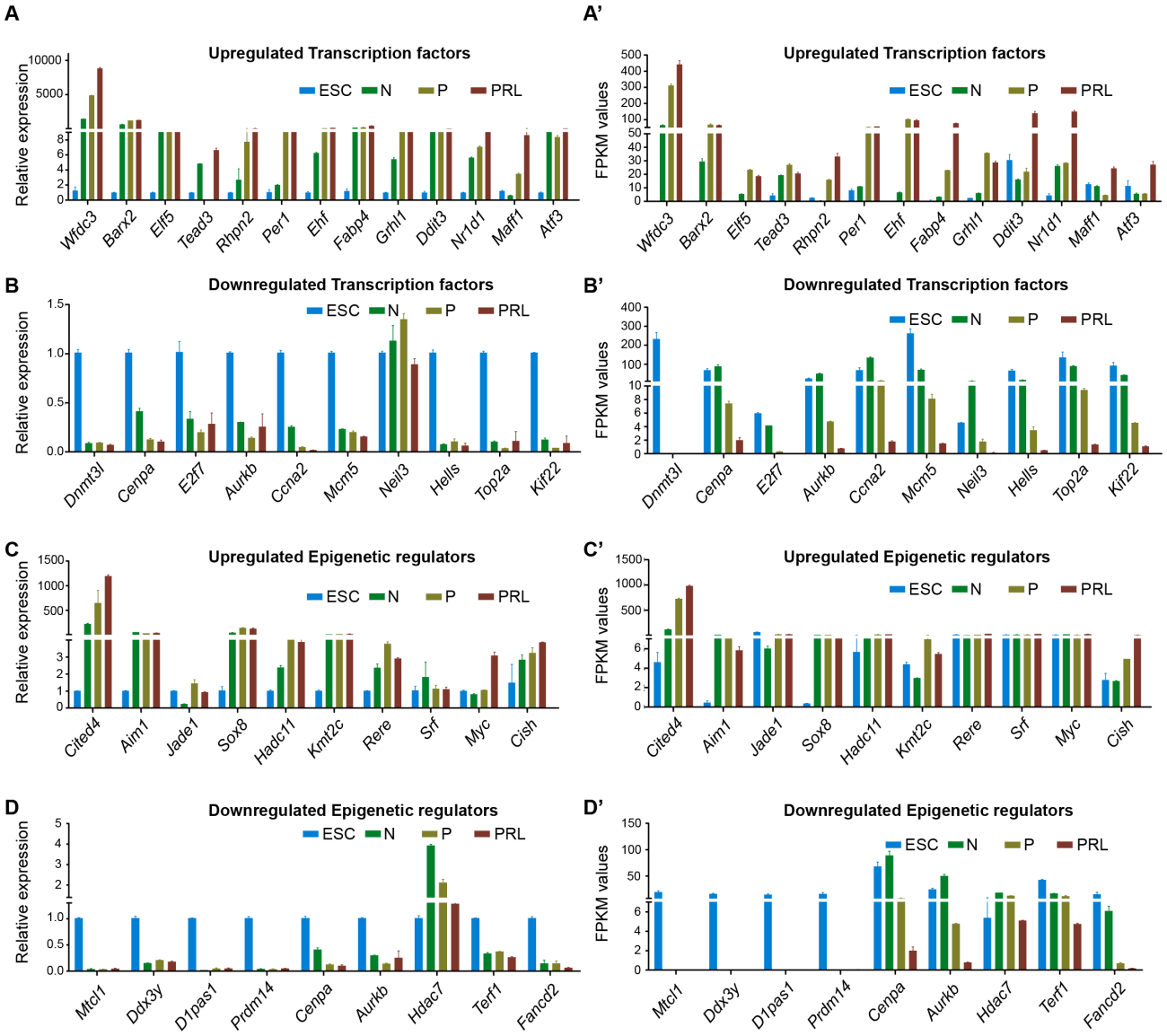


Fig.6. Real time PCR validation of top up or down regulated TFs and ERs in ESC and HC11 cells undergoing lactogenic differentiation. **(A)** Bar chart showing real time PCR measured relative expression of up-regulated and **(B)** downregulated TFs along with its respective RNA-Seq FPKM values **(A' and B')** in ESC, N, P and PRL treated HC11 cells. **(C)** Bar chart showing real time PCR measured relative expression of up-regulated and **(D)** down- regulated ERs along with its respective RNA-Seq FPKM values **(C' and D')** in ESC, N, P and PRL treated HC11 cells.

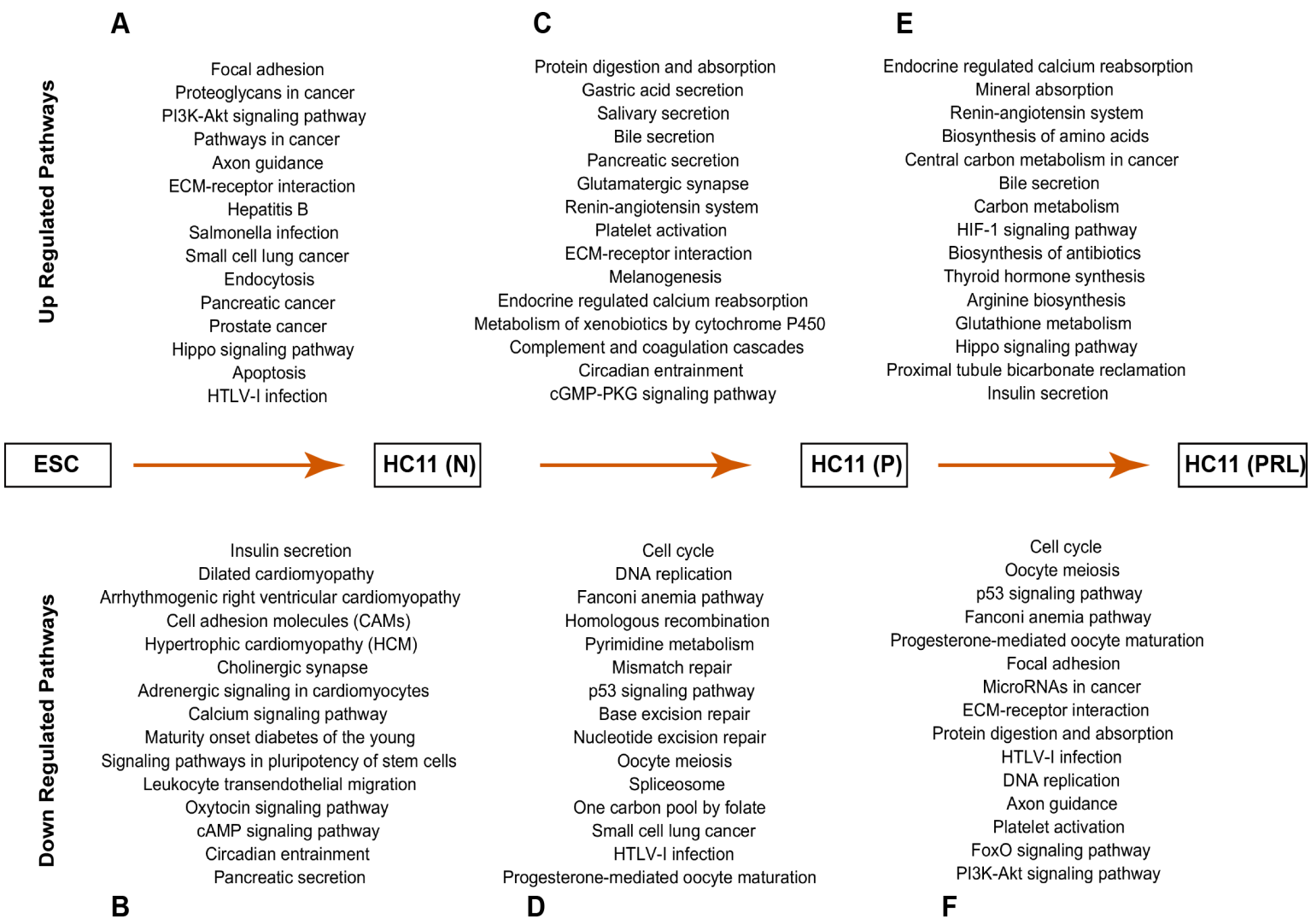


Fig.7. Pathway analysis of genes that are differentially expressed between ESC and HC11 cells undergoing lactogenic differentiation. (A, C & E) KEGG pathways analysis of genes that are either upregulated and **(B, D & F)** downregulated between ESC vs normal HC11 cells (N), N vs GC primed (P) and P vs PRL treated HC11 cells respectively.

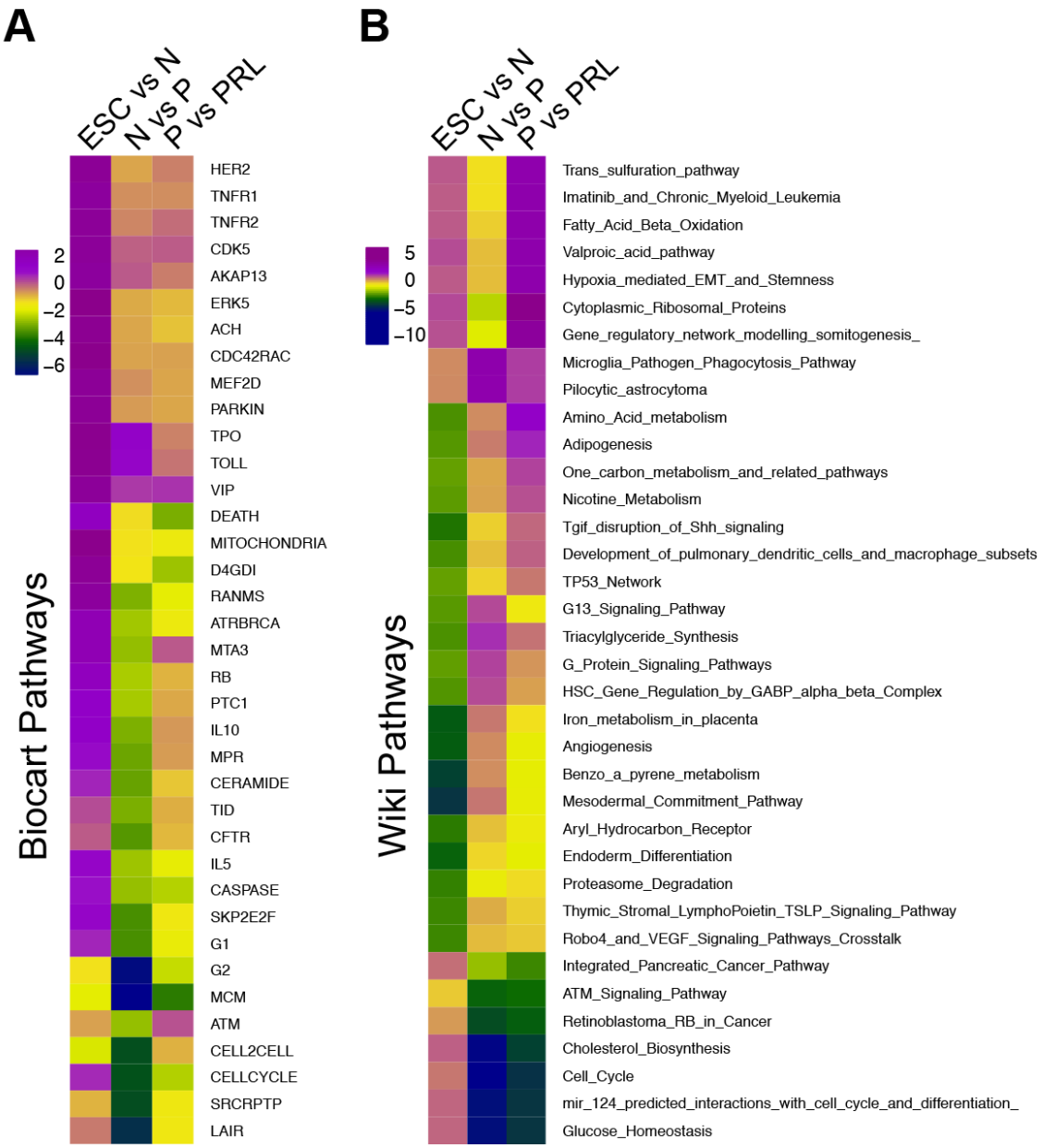


Fig.8. Comparative pathway analysis of differentially expressed genes between ESC and HC11 cell-types. Heatmap showing enrichment score (Z-score) of up or down-regulated genes using **(A)** BioCarta and **(B)** Wiki pathway tools between ESC vs N, N vs P and P v PRL treated HC11 cells. Z-score greater than 2 denote significantly upregulated and less than 2 denotes significantly downregulated.

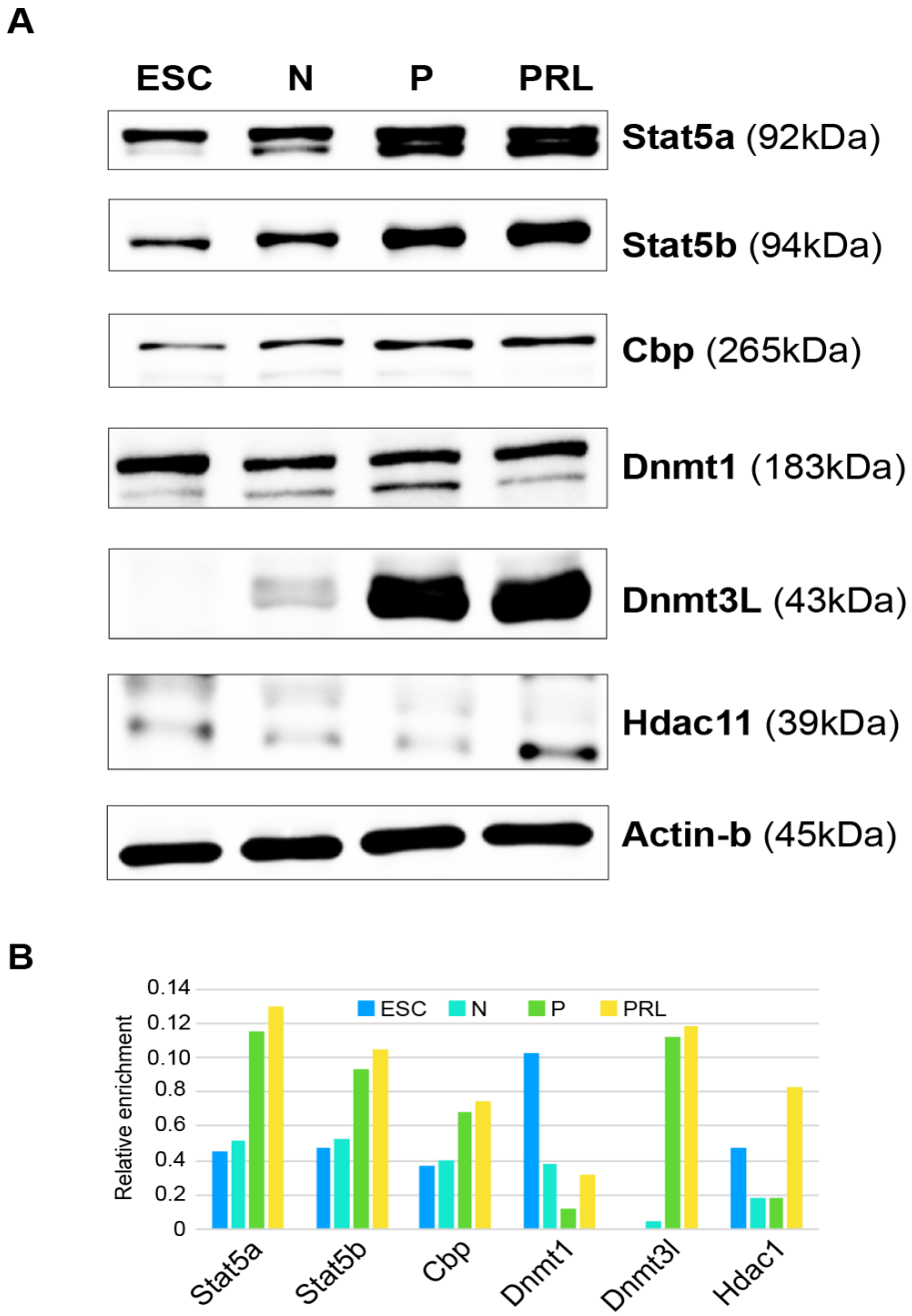


Fig.9. Immunoblot validation of selective markers in ESC and HC11 cells undergoing lactogenic differentiation. **(A)** Immunoblots showing ECL signal representing steady-state levels of selective protein markers in ESC, N, P and PRL treated HC11 cells. **(B)** Bar chart showing relative abundance of respective protein normalized against β-Actin.

Objective 2 & 3

9. **HC11 MEC lactogenic differentiation accompanied by changes in gene expression patterns but occurs in nonproliferative cells and predominantly at G0/G1 phase of cell cycle:**
HC11 MECs were able to maintain its epithelial phenotype in presence of EGF at confluence *in vitro*. Withdrawal of EGF and priming with GC followed by PRL hormonal treatment would initiate lactogenic differentiation, characterized by growth arrest, mammospheres formation (Fig. 10A) and expression of milk protein coding genes such as *β-Casein* (*Csn2*) and *Way acidic protein* (*Wap*). GC is known to initiate glucocorticoid signalling upon binding to cytoplasmic localized GC receptor (GR) and elicit its action in both genomic and non-genomic way (Kadmiel and Cidlowski, 2013). Prolactin alone could not induce lactogenic differentiation specific genes such as Csn2 suggesting that GR signalling is a prerequisite in mediating the effects of PRL signalling (Kabotyanski et al., 2006, Pittius et al., 1988). Previously, we have comprehensively mapped transcriptional networks that orchestrate normal HC11 cells, GC alone primed and in combination with PRL hormone treated HC11 and compared it with ESCs using RNA sequencing (Sornapudi et al., 2018). We have shown differential expression of thousands of genes, including transcription factors and epigenetic regulators in between ESC, HC11 cells and HC11 cells exposed to GC and PRL hormones (Fig. 10C). It is an established fact that lactogenic differentiation of HC11 cells occur at confluence stage under nonproliferative condition. It is of our interest to know if these chromatin changes, if any, occur in the same cell generation. Three lines of evidence support cell being arrested at G0/G1 phase of cell cycle via (i) visually observation of cells being at confluence (ii) FACS analysis show that most of the cells are being in Go/G1 phase of cell cycle (Fig. 10B) (iii) Depletion of cell cycle regulated proteins such as PCNA, CD4 (iv) RNA-seq gene expression analysis suggest shutting down of group of genes involved in cell cycle regulation. These results suggest that GC and PRL mediated differential expression of thousands of genes occur in cells that are predominantly at G0/G1 phase of cell cycle. It is therefore expected that the gene expression changes might be preceded by chromatin higher order structural changes, as is emerging as one of key epigenetic phenomena in establishing cell-type specific gene expression patterns.

10. **In-nucleus Hi-C of undifferentiated HC11 MECs and its comparison with ESCs and GC and PRL treated MECs:** To probe chromatin higher order structural changes in nonproliferative HC11 MECs during lactogenic differentiation, in-nucleus Hi-C developed by kurukuti and Vetrie [45] (see methods) was performed on two biological replicates of undifferentiated normal HC11 cells (N), GC primed (P) and GC and PRL hormone treated HC11 cells, along with mouse ESC and were sequenced

over illumina sequencing platform. The workflow of the analysis done was represented in Fig. 10D. All the sample sequences are mapped to mouse reference genome (mm10) and filtered by using HICUP [46]. We have generated nearly 16, 18, 10 & 10 million uniquely mapped paired-end reads from ESC, normal HC11 (N), GC primed (P) and PRL treated HC11 cells, respectively (Table 15). Genome-wide contact matrices were generated using HOMER at 1Mb resolution (see materials and methods for details). Correlation of raw contact frequency among replicates is high (>0.97) (Table 16A) as well as among the replicates normalized by SimpleNorm (0.74) (Table 16B) and hence the replicates were merged and processed further. Pearson correlation between genome-wide contact matrix of merged samples of ESC, Normal, GC primed and PRL treated HC11 cells were performed and found that ESCs were correlated with normal HC11 cells (p-value 0.946), when compared with Primed (p-value 0.956) and PRL treated HC11 cells (p-value 0.49) suggesting differences in chromosomal contacts between these four cell conditions (Table 16C). Genome-wide contact maps of all the cell conditions showed typical enrichment of intra-chromosomal interactions, suggestive of territorial nature of interphase chromosomes in the cell nucleus (Fig. 10E). However, certain highly enriched than expected contacts between chromosomes were observed in HC11 cell-types, when compared to relatively uniform distribution of inter-chromosomal interactions in ESCs. To rule out the possibility of restriction enzyme sites being over represented in those locations, raw contact matrices were normalized to the number of restriction sites and found to have significantly enriched interactions at five different locations representing chr1, 2, 5, 8 and 11 regions (Fig. 11A). Based on the suggestions made by Harewood et al., [47] ,these regions were identified as reminiscent of naturally occurring chromosomal translocations in HC11 cells, the observation which was not reported so far. To enable us to compare different conditions of HC11 cells with ESCs, we removed the translocation contacts (Fig. 11B, Table 15B) and substituted them with the mean inter-chromosomal contact frequency for each cell-condition (Fig. 11C). These results show that HC11 cell-types not only differ from their gene expression patterns but also differ in higher order arrangement of chromatin under nonproliferative conditions of HC11 MECs lactogenic differentiation.

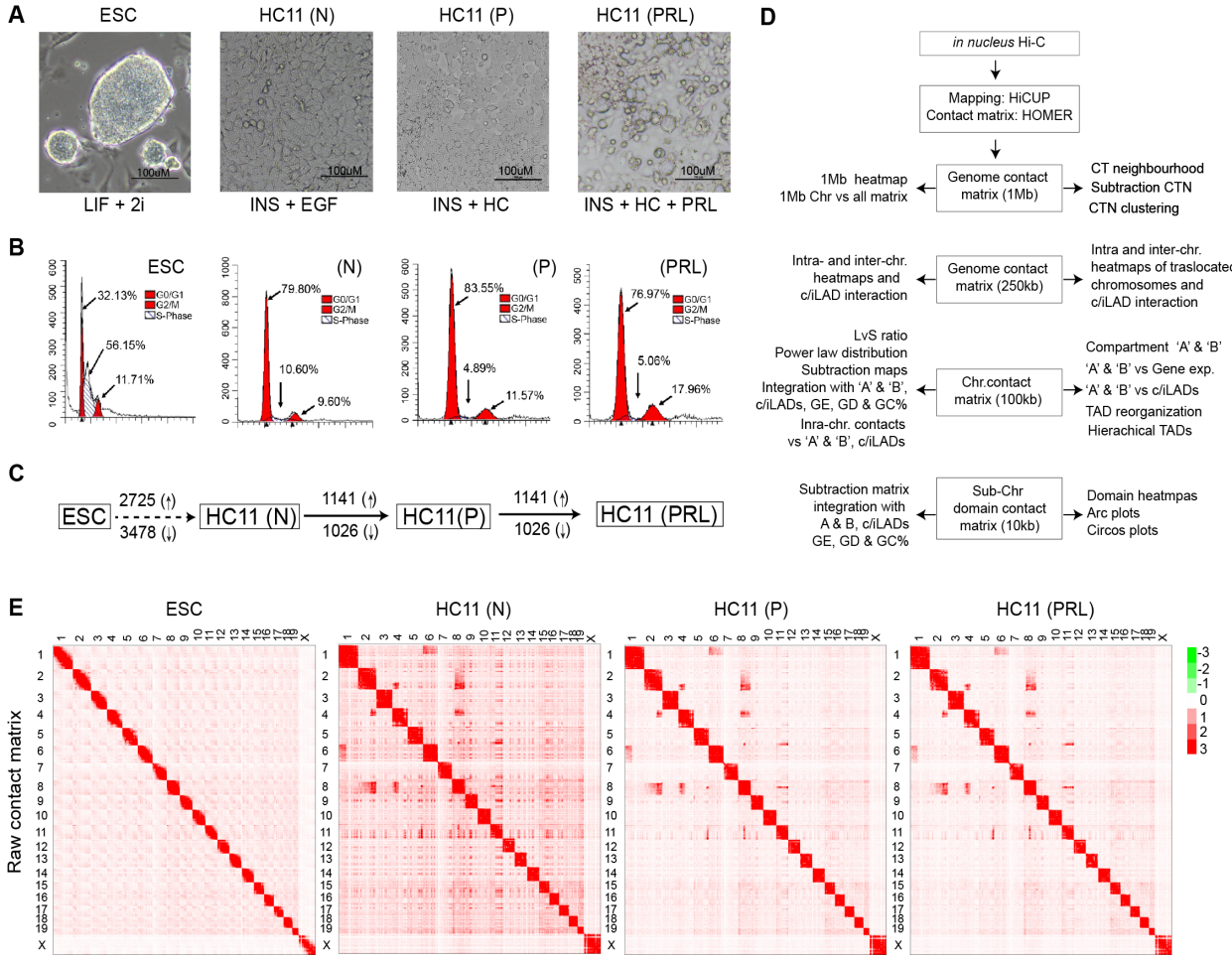


Fig.10. HC11 MECs lactogenic differentiation cell system, interplay between gene expression and genome architecture: **(A)** Schematic representation of ESCs and HC11 MECs lactogenic differentiation system and conditions used. Note the appearance of mammospheres in HC11 cells treated with HC and PRL. Scale bar represents 100μM. **(B)** FACS analysis of PI stained cell suspension from ESC and Undifferentiated normal HC11 cells (N), HC11 cells primed with HC alone (P) and in combination with prolactin (PRL), showing cells predominantly in 'S' phase in ESCs and G0/G1 phase in different stages of HC11 cells. **(C)** Schematic representation of number of differentially expressed genes between ESC- HC11 (N), HC11 (N) vs (P) and, (P) vs PRL treated cells. Upper arrow represents up-regulation and down-arrow represents down-regulation. **(D)** Schematic representation of experiments and bioinformatics analysis performed on the ESCs and HC11 MECs undergoing lactogenic differentiation. **(E)** Heatmaps representing genome-wide contact frequency at 1Mb.

11. **Comparison of probabilistic chromosome territory neighbourhood (CTN) maps reveal only subtle reorganization of CTs during lactogenic differentiation of HC11 MECs:** We generated chromosome territory neighbourhood maps (CTN) from raw inter-chromosomal contacts in ESC and translocations deleted and mean substituted raw matrices of HC11 cells-types. We further normalized these datasets by vanilla coverage (VC) normalization (Fig. 11D) [48]. VC normalized contact matrices show relatively less correlation between ESC and normal HC11 cells (0.181) compared to normal and GC primed (0.219) and between GC primed and PRL treated HC11 cells (0.185) (Table 17A). VC matrices were further corrected for the length of each chromosome and the resultant inter-chromosomal interactions were represented as heatmaps, thus deriving probabilistic CTN maps of ESC, normal HC11 cells, GC primed and PRL treated HC11 cells (Fig. 12A). From these maps, we observed that the CTN of ESC looks different from that of the HC11 cell-types and CTN look similar among HC11 cell-types. It is apparent from these heatmaps that smaller chromosomes 8-19,X in ESCs and chr15-19,X in HC11 cell-types tend to interact within themselves compared to rest, which appeared to be more non-random. The spearman correlation between CTNs of ESC and Normal HC11 cells is 0.835, but there is relatively high correlation between CTNs of normal vs GC primed HC11 cells (0.975) and even higher between primed vs PRL treated HC11 cells (0.988). Paired Wilcoxon test showed that, HC11 cell-types CTN are not significantly different among themselves but are different from that of ESCs (Table 17D). Then, we generated subtraction matrices of CTNs between the samples, which clearly showed the changes in CTNs between ESC-N, N-P and P-PRL (Fig. 12B). We observed a dramatic reorganization of CTs between ESC and normal HC11 cells (N), however, there seemed to be very subtle reorganization between normal and GC primed and further less between primed and PRL treated HC11 cells, in agreement with the spearman correlation and Wilcoxon tests. Further, the normal HC11 cells vs GC primed CTN subtraction matrix showed that smaller chromosomes are clustered and clustering was further increased a little in GC primed vs PRL treated HC11 cells subtraction CTNs maps We also measured differences in clustering, as described in Barutcu et al., [49] among smaller chromosomes such as chr 15,16,17,18 and 19, and with the rest of the genome, between the cell conditions, which clearly showed significantly higher increase in clustering among the smaller chromosomes than the rest of the genome between ESC-N, N-P and P-PRL treated HC11 cells (Fig. 12C). However, similar analysis with larger chromosomes such as 8,9,10 and 11, showed that they tend to interact less within themselves than with the rest (Fig. 12D). Further, hierarchical clustering of the CTN maps showed cell-type specific patterns (Fig. 12E), reflective of the observations made above. In

our previous work [44], we have shown dramatic changes in gene expression networks and pathways upon GC and PRL treatment. These results suggest that despite extensive changes in gene expression patterns, HC11 cells differentiate to Primed state with subtle changes in CTNs and further differentiate under PRL hormone with even lesser changes CTN. These results suggest that under nonproliferative, growth arrest conditions, GC and PRL treated HC11 cells undergo lactogenic differentiation with dramatic changes in gene expression patterns with only subtle changes in CTN but with few smaller chromosomes significantly gaining interactions within and between them.

12. **Lactogenic differentiation of HC11 MECs accompanies an extensive reorganization of short/long-range intra- and inter-chromosomal contacts**: Lack of dramatic changes in CTN, but with differential expression of thousands of genes upon GC and PRL treated HC11 cells prompted us to investigate changes if any at the level of sub-chromosomal domain. Individual chromosome contact matrices normalized by SimpleNorm at 100kb resolution were generated using HOMER (Heinz et al., 2010) (see methods). A close look at the genome-wide raw as well as simpleNorm 1Mb intra-chromosomal heatmaps of ESC and HC11 cell types, showed restricted interactions which appeared as diagonally rhomboid beads, which is indicative of lack of long-range intra-chromosomal contacts in highly proliferating ESCs. Whereas in HC11 cell types, it is more sparse and appears as square beads, which is indicative of presence of long-range intra-chromosomal contacts (Fig. 10E, Fig. 13A). During the mapping of the reads with HICUP, by looking at percentage of short-range (<10kb), long-range (>10kb) intra as well as inter-chromosomal contacts, we found that ESC have less percentage of inter chromosomal contacts (7%) compared to HC11 cell stages (16%, 13% and 13% respectively at this depth of data for normal HC11, GC primed and PRL treated cells) (Fig. 13B, Table 15A). To quantitate the ratio of short/long range chromatin interactions, we followed the method described by Chandra et al., (Chandra et al., 2015), where, we measured genome-wide ratio of intra-chromosomal long-range (>2Mb) vs short-range (<2Mb) interactions (LvS) and found that ESC genome harbour majorly short-range interactions when compared with HC11 cell-types, which showed significantly higher long-range contacts (Fig. 13C). Among MECs, HC11 cells treated with GC hormone showed significantly higher LvS ratio than normal cells, suggesting dramatic changes in long-range contacts upon GC treatment. Interestingly, in PRL treated cells however, the ratio slightly reduces when compared with GC treated HC11 cells.


ESC
HC11 (N)
HC11 (P)
HC11 (PRL)
A
Raw + RE normalized
B
Raw - Translocations (TD)
C
Raw - TD + Medien filled
D
VC inter-chr contact matrix
-3
-2
-1
0
1
2
3

Fig.11. Generation of genome-wide contact matrix heatmaps using various normalization procedures: (A) Raw interaction matrix of 1MB domains across the genome normalized for the restriction enzyme occurrence biases showing significant presence of naturally occurring chromosomal translocations involving chr1:6/ 2:4/ 4:8/ 5:8/ 5:11/ 8:11 (Arrow heads) in all cell-types of HC11 cells. Note the lack of such translocation in ESCs. **(B)** Genome-wide raw interaction matrix of 1MB domains across the genome, where translocated regions were deleted and **(C)** replaced with mean inter-chromosomal interaction frequency (arrow heads). **(D)** Genome-wide 1MB contact matrix heatmap of vanilla-coverage normalized inter-chromosomal data (excluding intra-chromosomal interactions) was used to construct probabilistic chromosome neighbourhood maps in ESC, N, P and PRL treated HC11 cells.

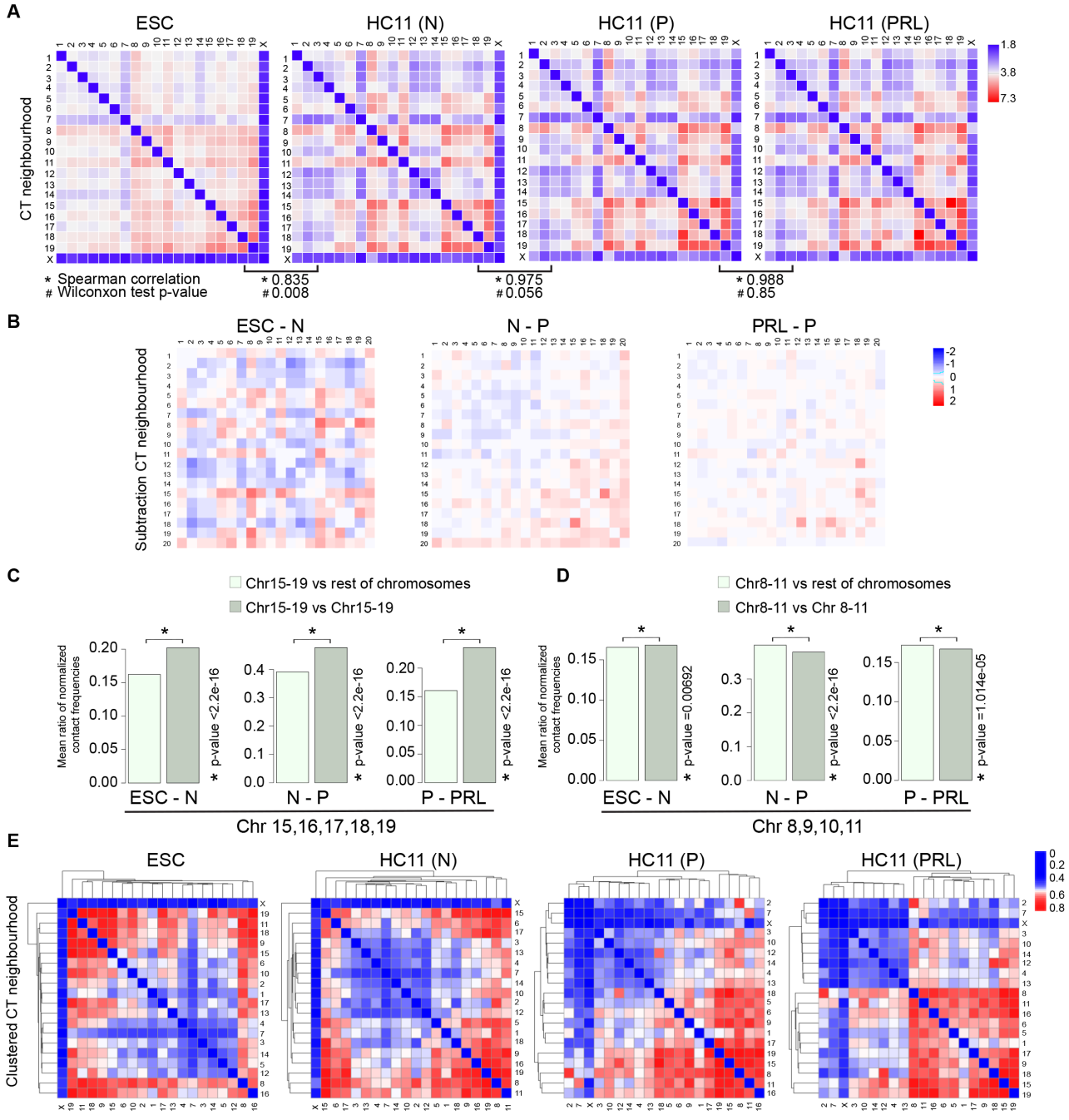

A
ESC
HC11 (N)
HC11 (P)
HC11 (PRL)
CT neighbourhood
1.8
3.8
7.3
* Spearman correlation
# Wilconxon test p-value
* 0.835
# 0.008
* 0.975
# 0.056
* 0.988
# 0.85
B
ESC - N
N - P
PRL - P
Subtraction CT neighbourhood
-2
-1
0
1
2
C
Chr15-19 vs rest of chromosomes
Chr15-19 vs Chr15-19
Mean ratio of normalized contact frequencies
* p-value <2.2e-16
* p-value <2.2e-16
* p-value <2.2e-16
ESC - N
N - P
P - PRL
Chr 15,16,17,18,19
D
Chr8-11 vs rest of chromosomes
Chr8-11 vs Chr 8-11
Mean ratio of normalized contact frequencies
* p-value =0.00692
* p-value <2.2e-16
* p-value =1.014e-05
ESC - N
N - P
P - PRL
Chr 8,9,10,11
E
ESC
HC11 (N)
HC11 (P)
HC11 (PRL)
Clustered CT neighbourhood
0
0.2
0.4
0.6
0.8

Fig.12. Probabilistic Chromosome Territory Neighbourhood (CTN) maps reveal similarities and differences in ESCs and HC11 MECs undergoing lactogenic differentiation: **(A)** CTN maps derived from respective cell-type, using vanilla coverage normalized inter-chromosomal contacts only, where ESCs show more random arrangement of CTs of smaller chromosomes (8-19) compared to larger chromosomes (1-7 and X), whereas HC11 cell-types show non-random organization of CTN and seen to be similar among themselves (as evident from their spearman correlation and Wilcoxon Rank sum test p-values). Scale bar represents intensity of inter-chromosomal contacts normalized to the length of chromosomes. **(B)** Subtraction CTN maps between ESC&N, N&P and P&PRL cell-types, showing more pronounced differences between ESC&N compared to N&P and very few in P&PRL cellular states. Scale bar represents intensity of gain or loss of interactions between two states in arbitrary units. **(C & D)** Bar plot representing mean ratios of normalized contacts which shows difference in the clustering of smaller chromosomes 8, 9, 10 and 11 (C) between cell types . The same is shown for larger chromosomes Chr15-19 (D). p-values show that larger chromosomes cluster less than smaller chromosomes. **(E)** Hierarchical clustering of CTN shows similarities and differences in CTN among four cell-types.

We next asked if these changes that we see globally among the four cell condition are represented at individual chromosomes. Surprisingly, we found that every chromosome behaves in a similar way as was seen on genome-wide scale (Fig. 13D), suggesting the general nature of action of respective GC/PRL signalling on all chromosomes. These changes in Short Vs Long range interactions were also evident These results suggest that both GC and PRL hormones execute differential expression of thousands of genes in non-proliferating HC11 cells by dramatic alterations in long-range intra-chromosomal contacts within the context of minimally altered CTNs.

13. **Dynamic reorganization of chromatin interactions in alternative sub-chromosomal domains of A and B in succession accompanies lactogenic differentiation of HC11 cells:**

In order to assess whether the dramatic changes in gene expression patters, with subtle changes in CTN during lactogenic differentiation of HC11 cells would alter changes within chromosomes, we generated simpleNorm matrices at 100kb resolution genome-wide (Fig. 18). To understand the nature of chromatin regions undergoing reorganization among ESCs and HC11 cells during lactogenic differentiation, we performed Principle Component Analysis (PCA) to derive cell-type specific compartments A and B regions at 100kb resolution and integrated them with steady state levels of gene expression (RNA-seq), chromosome G-bands, with chr6 Hi-C interaction matrix heatmaps (Fig. 14A1, 2, 3, & 4), and the same for chr1 (Fig. 15A), chr15 (Fig. 16A) and chrX (Fig. 17A). We focussed our attention on Chr1, 6, 15 and X for deeper analysis. Unless and otherwise stated, throughout the text, we referred to chr6 as an example, to distinguish the differences in intra-chromosomal interactions among four cells-types studied. This analysis clearly showed that regions that gained Long-Range Interactions (LRI) between ESC and normal HC11 cells, span both compartment A and B (Fig. 14A2) but they were predominantly seen enhanced in compartment B in GC primed (Fig. 14A3) and to compartment A in PRL treated cells (Fig. 14A4). However, we found no correlation between the alternative chromatin regions with chromosomal G-banding patterns (Fig. 14A1, 2, 3, & 4). Distance decay plot of interactions in four cells-types show that in ESCs sudden drop in LRI after 40Mb compared to HC11 cell-types, where they had LRI, and interestingly GC primed cells seems to harbour more LRI than rest of the cell-types (Fig. 14B), including chr1 (Fig. 15B), chr15 (Fig. 16B), chrX (Fig. 17B) and rest of the chromosomes (Fig. 19). Further, we noted that compartment A correlates with higher gene expression patterns when compared with compartment B and the compartments are dynamically switching between the cell-types (ESC-N, N-P and P-PRL states). To elucidate the relation between various chromatin structural features of changing chromatin interactions, we overlaid the

compartments, gene expression, conserved Lamina (cLADs) and inter lamina associated domains (ciLADs) along with GC%, gene density and chromosomal bands and visualized in UCSC genomic browser for chromosome 6 (Fig. 14C), chr1 (Fig. 15C), chr15 (Fig. 16C) and chrX (Fig. 17C). This integrative approach of visualization clearly showed similarities, differences and inter-relationship of Hi-C heatmap with respective cell-type specific features. A close look at these integrated maps show striking visual correlation (which we later tested statistically for their significance), such as (i) conservation of compartment A and B across the four cell-types with few switching between A-B or B-A compartment, (ii) Compartment A with higher gene expression than B, (iii) similarities of compartment A with conserved inter lamina associated domain (ciLADs) regions (denoted as Y), which are correlated with GC rich isochores of genome that harbours higher gene density, and compartment B with conserved lamina associated domains (cLADs) regions (denoted as X), which are correlated with AT rich isochores of genome that harbours lower gene density, (iv) lack of correlation of X or Y domains with chromosomal bands. We calculated conserved and dynamically switching chromatin bins between A and B compartment in all four cell-types studied within chr6 (Fig. 14D1, 2), chr1 (Fig. 15D), chr15 (Fig. 16D) and chrX (Fig. 17D) and on genomic scale (Fig. 20C), which clearly showed that most of the chromatin maintain in their respective compartments, are conserved but some are overlapping between any two cell-types and some are exclusive. Then we analyzed percentage of bins that are maintained (A-A and B-B) and that switched (A-B and B-A) between ESC-N, N-P and P-PRL states (Fig. 14E), which showed more or less equal proportion of compartment A and B in all the states. However, we noted that GC treatment leads to predominant switching of chromatin regions from compartment B to A, whereas it is from compartment B to A in PRL treated HC11 cells, in chromosome 6. Interestingly, we found that GC treatment leads to less of A to B switching but more of B-A switching, which is strikingly inverse to that of PRL treatment (Fig. 14E). Similar trends are seen in chr1 (Fig. 15E), chr15 (Fig. 16E) but not in chrX (Fig. 17E). We extended our analysis to see this behaviour on genomic scale and found reciprocal behaviour of A-B compartments under GC and PRL treatment (Fig. 20C). Next, we wanted to see whether the switch in the compartments from one cell-type to another is accompanied by changes in the gene expression. When the expression of genes presents in A and B compartments were plotted separately, it was obviously seen that the expression of the genes in A compartment are much higher than in B in chr6 (Fig. 14F), chr1 (Fig. 15F), chr15 (Fig. 15F), chrX (Fig. 17F) and is true genome-wide as well (Fig. 20B).

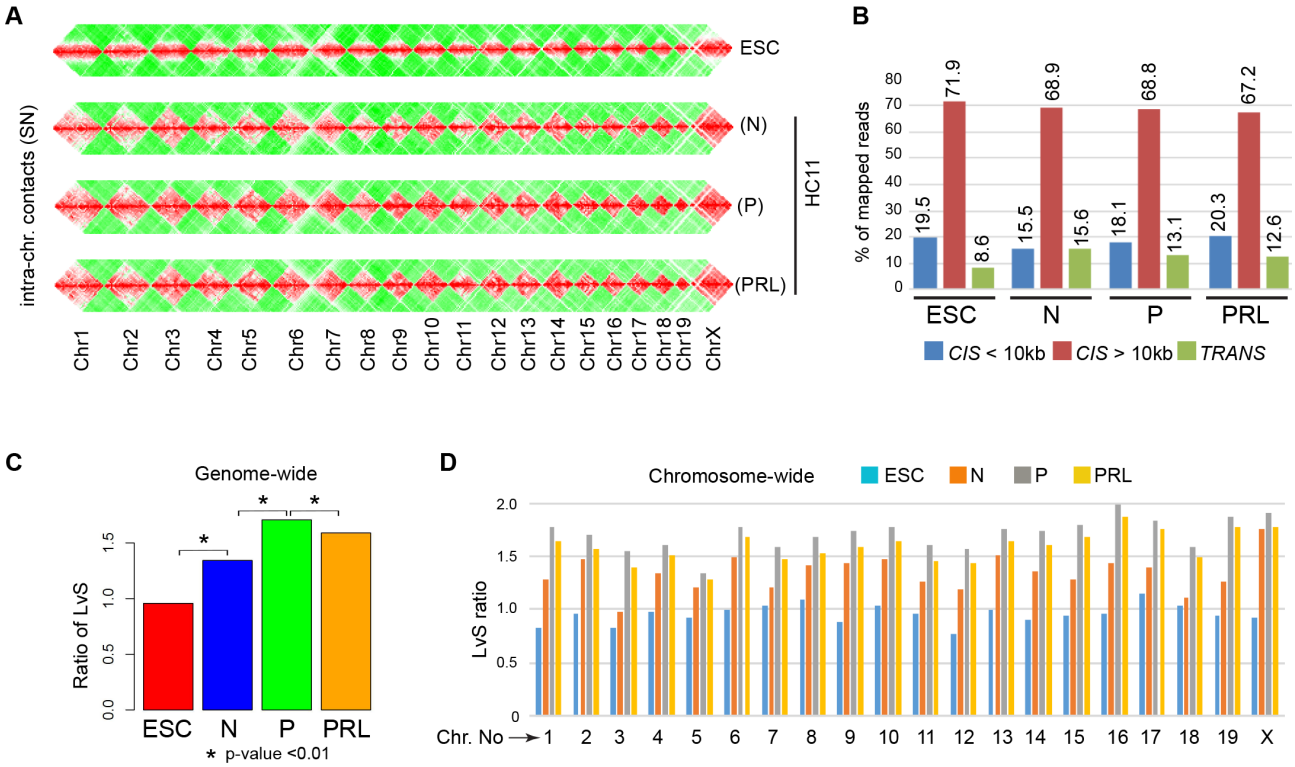

A
ESC
(N)
(P)
(PRL)
HC11
intra-chr. contacts (SN)
Chr1 Chr2 Chr3 Chr4 Chr5 Chr6 Chr7 Chr8 Chr9 Chr10 Chr11 Chr12 Chr13 Chr14 Chr15 Chr16 Chr17 Chr18 Chr19 ChrX
B
% of mapped reads
19.5 71.9 8.6
15.5 68.9 15.6
18.1 68.8 13.1
20.3 67.2 12.6
ESC N P PRL
CIS < 10kb CIS > 10kb TRANS
C
Genome-wide
Ratio of LvS
ESC N P PRL
* p-value <0.01
D
Chromosome-wide
ESC N P PRL
LvS ratio
Chr. No → 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X

Fig. 13. Global distribution of long-range vs short-range chromatin interactions in ESC and HC11 MECs undergoing lactogenic differentiation: **(A)** Heatmap showing SimpleNorm normalized intra-chromosomal contact matrix in all mouse chromosomes (1-19+X) in ESC, Normal (N), Primed (P) and PRL treated HC11 cells. Note the presence of significantly higher short-range intra-chromosomal contacts in ESCs compared to rest of the HC11 cell-types. **(B)** Bar plot showing percentages of short-range (*Cis* <10kb), long-range intra-chromosomal (*Cis* >10kb) and inter-chromosomal interactions (*Trans*), showing lower percentage of inter-chromosomal contacts and higher percentage of intra-chromosomal contacts in ESC when compared to rest of the HC11 cell-types. Note the gradual increase in short-range intra-chromosomal contacts among (N), (P) and (PRL) treated HC11 cell-types. **(C)** Genome-wide ratio of long-range (L: >2Mb) vs Short-range (S: <2Mb) (LvS) *Cis*-chromosomal contacts in ESC, N, P and PRL treated HC11 cells. Note the dramatic increase in long-range intra-chromosomal contacts between ESC and N and between N and P states of HC11 cells, with slight reduction under PRL state of HC11 cells. Paired willcoxon rank sum test was used for statistical testing. **(D)** Chromosome-wide distribution of LvS *Cis* chromatin interactions. Note the striking similarity in ratios of LvS contacts across all chromosomes in ESC, N, P and PRL treated cells, especially in primed state to have higher long-range *Cis* contacts.

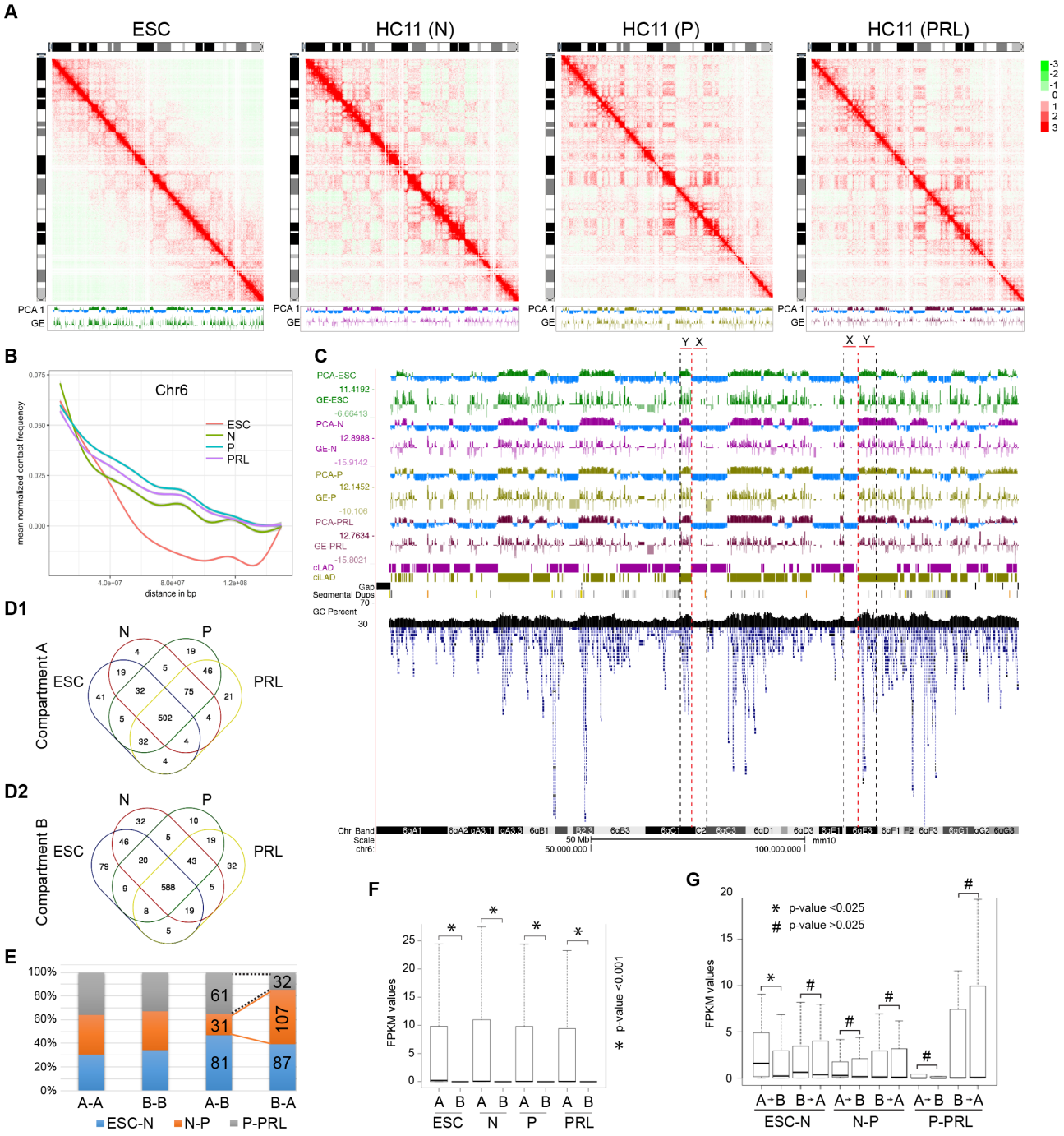

A
ESC
HC11 (N)
HC11 (P)
HC11 (PRL)
PCA 1
GE
B
Chr6
ESC
N
P
PRL
mean normalized contact frequency
distance in bp
C
PCA-ESC
GE-ESC
PCA-N
GE-N
PCA-P
GE-P
PCA-PRL
GE-PRL
cLAD
ciLAD
Gap
Segmental Dups
GC Percent
Chr Band
Scale
chr6:
50 Mb
mm10
D1
Compartment A
N
P
ESC
PRL
D2
Compartment B
E
A-A
B-B
A-B
B-A
ESC-N
N-P
P-PRL
F
FPKM values
p-value <0.001
G
p-value <0.025
p-value >0.025
FPKM values
ESC-N
N-P
P-PRL

Fig.14. **Dynamics of reorganization of intra-chromosomal (Chr-6) interactions and gene expression in ESC, N, P and PRL treated HC11 cells**: **(A)** Chr6 contact frequency heatmap at 100kb resolution in ESC, N, P and PRL treated HC11 cells showing fewer long-range interactions in ESCs when compared to other HC11 cell-types. Visually distinguishable and enriched long-range interactions of certain domains are observed in primed state. Each heat map is overlaid on PCA1 derived A/B compartments and gene expression (GE) tracks from UCSC, along with chromosome G-banding pattern in both X and Y axis, showing existence two kinds of compartments in each state and accompanied gene expression patterns. Scale bar represents intensity of chromatin interactions at 100kb resolution in arbitrary units. **(B)** Decay plot showing mean normalized contact frequency in chr6 in ESC (Red), N (green), P (blue) and PRL (pink) along the distance of the chromosome, showing differences in the distributions. **(C)** UCSC snapshot view of Chr6 with respective cell-type specific PCA1, gene expression, cLADs, ciLADs, GC/AT %, gene density and chromosome bands showing dynamic switching of genes between compartments and accompanied differential gene expression patterns. Note the striking relationship between compartment A with ciLADs, GC rich isochores, gene rich domains of the chromosomes (Highlighted lines showing 'X' region), whereas compartment B with cLADs, AT rich isochores and gene poor regions of chromosomes (Highlighted dotted lanes showing 'Y' region). Note the lack of values in the regions of GAP and segmental duplicated regions due to mapping issues. Highlighted lines show relationship of cLADs, ciLADs with gene density, gene expression and GC/AT %. **(D1,2)** Venn diagram showing number of conserved, overlapping and unique A & B compartments in Chr6 of ESC, N, P and PRL treated HC11 cells. Note that a large number of compartments are conserved between samples. **(E)** Bar plot showing number of compartments switching between A and B from one cell to another in chr6. Note dramatic switching of compartments B to A under GC primed (P) condition of HC11 cells (Shown by dotted lanes). **(F)** Relation between A/B compartments and gene expression was shown as boxplots (outliers excluded). The gene expression is generally seen high in compartment A than in B. Wilcoxon rank sum test was performed for statistical testing. **(G)** Boxplot representing the changes in the gene expression (values in FPKM) of the genes, which are shifting from (A-B, B-A), with outliers excluded. Note that switching from A-B or B-A compartments are not accompanied by decrease or increase in gene expression patterns respectively. significant statistical difference is depicted in p-values.

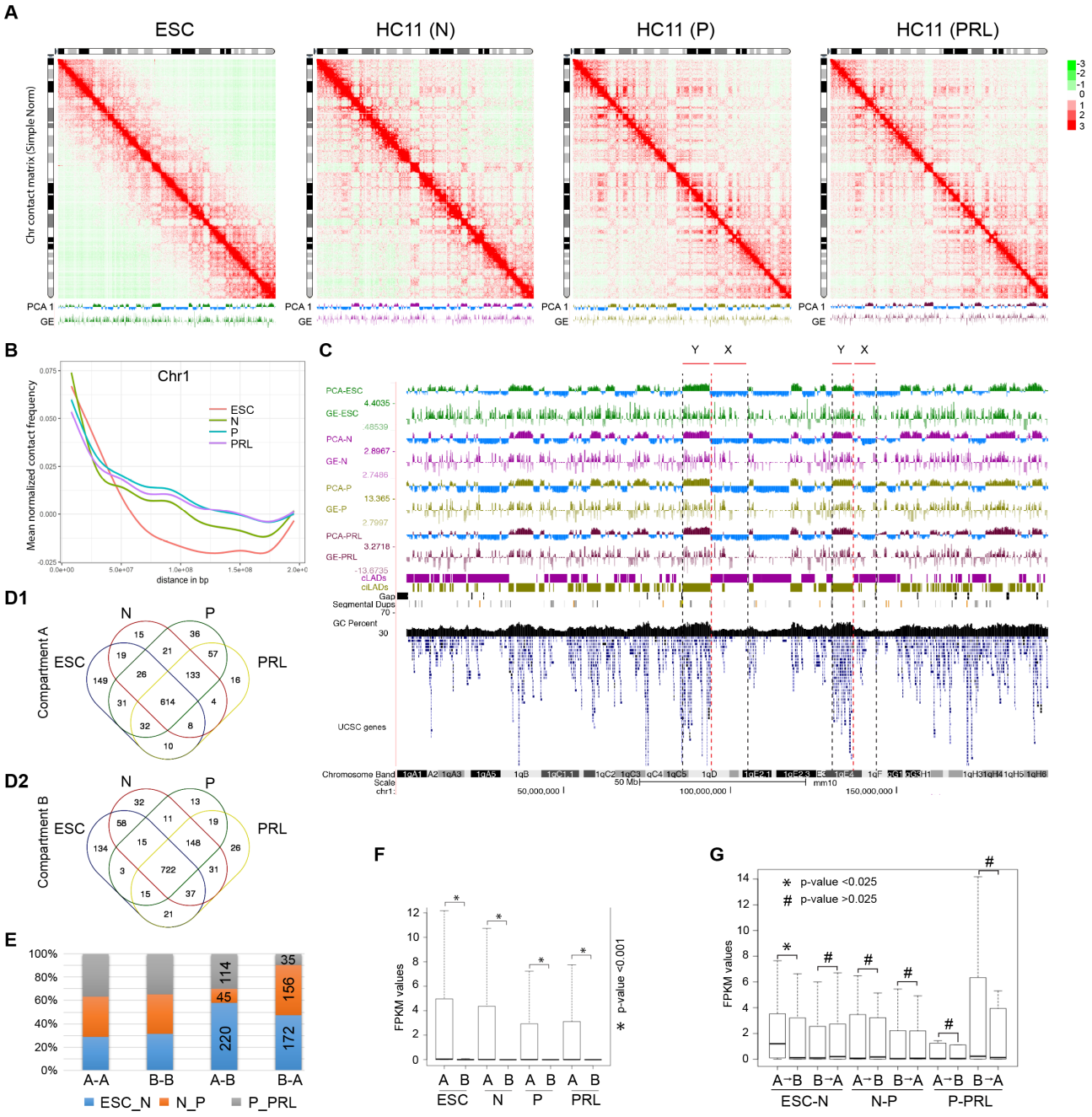

A
ESC
HC11 (N)
HC11 (P)
HC11 (PRL)
Chr contact matrix (Simple Norm)
PCA 1
GE
B
Chr1
ESC
N
P
PRL
Mean normalized contact frequency
distance in bp
C
PCA-ESC
GE-ESC
PCA-N
GE-N
PCA-P
GE-P
PCA-PRL
GE-PRL
cLADs
ciLADs
Gap
Segmental Dups
GC Percent
UCSC genes
Chromosome Band
D1
Compartment A
D2
Compartment B
E
A-A
B-B
A-B
B-A
ESC_N
N_P
P_PRL
F
FPKM values
p-value <0.001
ESC
N
P
PRL
G
p-value <0.025
p-value >0.025
ESC-N
N-P
P-PRL

Fig.15. Dynamics of reorganization of intra-chromosomal (Chr-1) interactions and gene expression in ESC, N, P and PRL treated HC11 cells: **(A)** Chr1 contact frequency heatmap at 100kb resolution in ESC, N, P and PRL treated HC11 cells showing fewer long-range interactions in ESCs when compared to other HC11 cell-types. Visually distinguishable and enriched long-range interactions of certain domains are observed in primed state. Each heat map is overlaid on PCA 1 derived A/B compartments and gene expression (GE) tracks from UCSC, along with chromosome G-banding pattern in X and Y axis respectively, showing existence two kinds of compartments in each state and accompanied gene expression patterns. **(B)** Decay plot showing mean normalized contact frequency in Chr1 in ESC (Red), N (green), P (blue) and PRL (pink) along the distance between genomic regions showing differences in the distributions. **(C)** UCSC snapshot view of Chr1 with respective cell-type specific PCA1, gene expression, cLADs, ciLADs, GC/AT %, gene density and chromosome bands showing dynamic switching of genes between compartments and accompanied differential gene expression patterns. Note the striking relationship between compartment A with ciLADs, GC rich isochores, gene rich domains of the chromosomes (Highlighted lines showing 'X' region), whereas compartment B with cLADs, AT rich isochores and gene poor regions of chromosomes (Highlighted dotted lanes showing 'Y' region). Note the lack of values in the regions of GAP and segmental duplicated regions due to mapping issues. Highlighted lines show relationship of cLADs, ciLADs with gene density, gene expression and GC/AT %. **(D1,2)** Venn diagram showing number of conserved, overlapping and unique A & B compartments in Chr1 of ESC, N, P and PRL treated HC11 cells. Note that a large number of compartments are conserved between samples. **(E)** Bar plot showing number of compartments switching between A and B from one cell to another in Chr1. Note dramatic switching of compartments B to A under GC primed (P) condition of HC11 cells (Shown by dotted lanes). **(F)** Relation between A/B compartments and gene expression was shown as boxplots (outliers excluded). The gene expression is generally seen high in compartment A than in B. Wilcoxon rank sum test was performed for statistical testing. **(G)** Boxplot representing the changes in the gene expression (FPKM) of the genes, which are shifting from (A-B, B-A), with outliers excluded. Note that switching from A-B or B-A compartments are not accompanied by decrease or increase in gene expression patterns respectively in between N-P and P-PRL states of HC11 cells. Significance of difference is depicted in p-values.

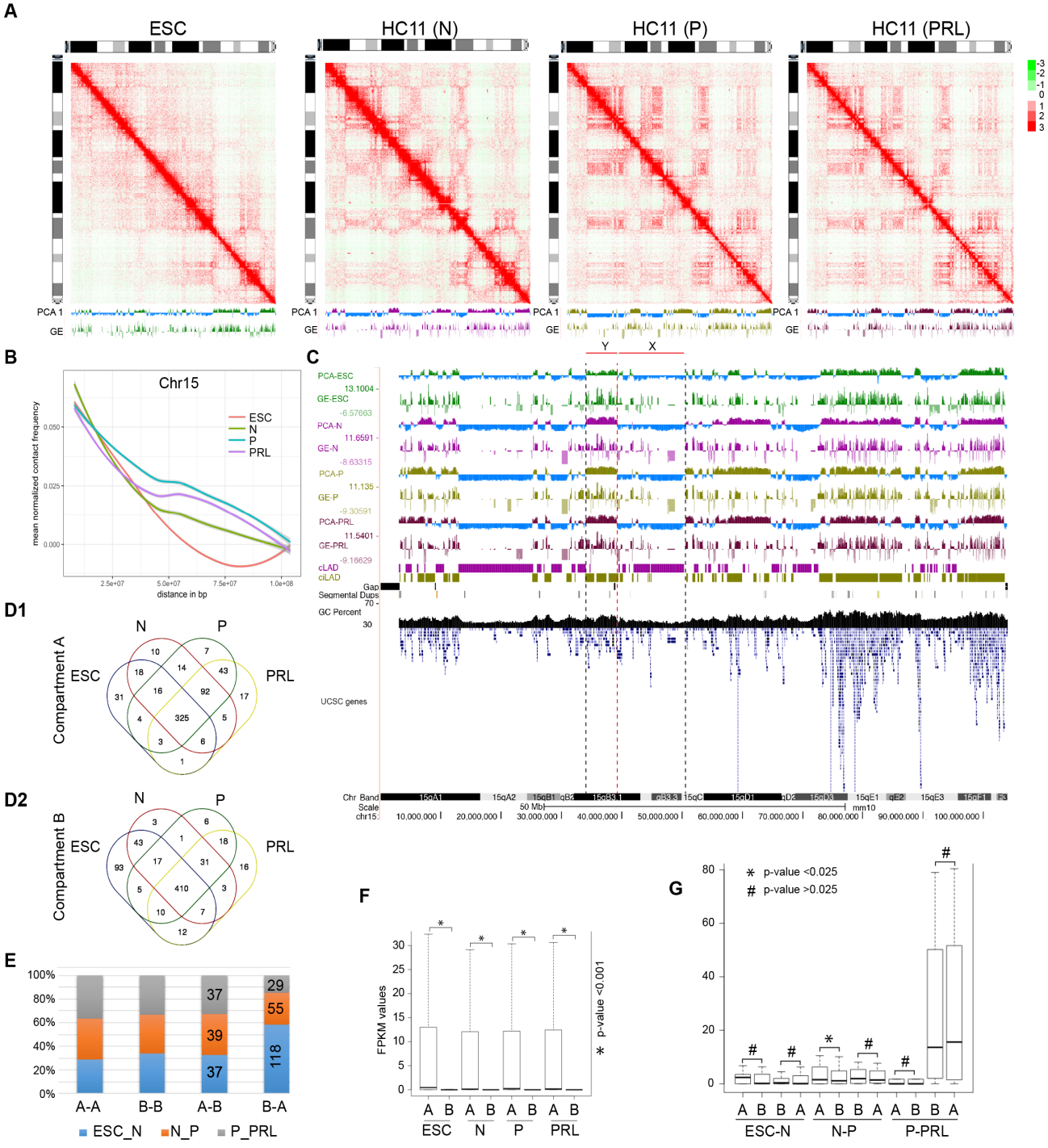

A
ESC
HC11 (N)
HC11 (P)
HC11 (PRL)
PCA 1
GE
B
Chr15
ESC
N
P
PRL
mean normalized contact frequency
distance in bp
C
PCA-ESC
GE-ESC
PCA-N
GE-N
PCA-P
GE-P
PCA-PRL
GE-PRL
cLAD
ciLAD
Gap
Segmental Dups
GC Percent
UCSC genes
Chr Band
Scale
chr15:
D1
Compartment A
N
P
ESC
PRL
D2
Compartment B
E
ESC_N
N_P
P_PRL
A-A
B-B
A-B
B-A
F
FPKM values
p-value <0.001
ESC
N
P
PRL
G
p-value <0.025
p-value >0.025
ESC-N
N-P
P-PRL

Fig.16. Dynamics of reorganization of intra-chromosomal (Chr-15) interactions and gene expression in ESC, N, P and PRL treated HC11 cells: **(A)** Chr15 contact frequency heatmap at 100kb resolution in ESC, N, P and PRL treated HC11 cells showing fewer long-range interactions in ESCs when compared to other HC11 cell-types. Visually distinguishable and enriched long-range interactions of certain domains are observed in primed state. Each heat map is overlaid on PCA 1 derived A/B compartments and gene expression (GE) tracks from UCSC, along with chromosome G-banding pattern in X and Y axis respectively, showing existence two kinds of compartments in each state and accompanied gene expression patterns. **(B)** Decay plot showing mean normalized contact frequency in chr15 in ESC (Red), N (green), P (blue) and PRL (pink) along the distance, between genomic regions showing differences in the distributions. **(C)** UCSC snapshot view of Chr15 with respective cell-type specific PCA1, gene expression, cLADs, ciLADs, GC/AT %, gene density and chromosome bands showing dynamic switching of genes between compartments and accompanied differential gene expression patterns. Note the striking relationship between compartment A with ciLADs, GC rich isochores, gene rich domains of the chromosomes (Highlighted lines showing 'X' region), whereas compartment B with cLADs, AT rich isochores and gene poor regions of chromosomes (Highlighted dotted lanes showing 'Y' region). Note the lack of values in the regions of GAP and segmental duplicated regions due to mapping issues. Highlighted lines show relationship of cLADs, ciLADs with gene density, gene expression and GC/AT %. **(D1,2)** Venn diagram showing number of conserved, overlapping and unique A & B compartments in Chr15 of ESC, N, P and PRL treated HC11 cells. Note that a large number of compartments are conserved between samples. **(E)** Bar plot showing number of compartments switching between A and B from one cell to another in chr15. Note dramatic switching of compartments B to A under GC primed (P) condition of HC11 cells (Shown by dotted lanes). **(F)** Relation between A/B compartments and gene expression was shown as boxplots (outliers excluded). The gene expression is generally seen high in compartment A than in B. Wilcoxon rank sum test was performed for statistical testing. **(G)** Boxplot representing the changes in the gene expression (FPKM) of the genes, which are shifting from (A-B, B-A), with outliers excluded. Note that switching from A-B or B-A compartments are not accompanied by decrease or increase in gene expression patterns respectively between the cell-types except in for A-B compartment shift in between N-P stages of HC11 cells. Significance of difference is depicted in p-values.

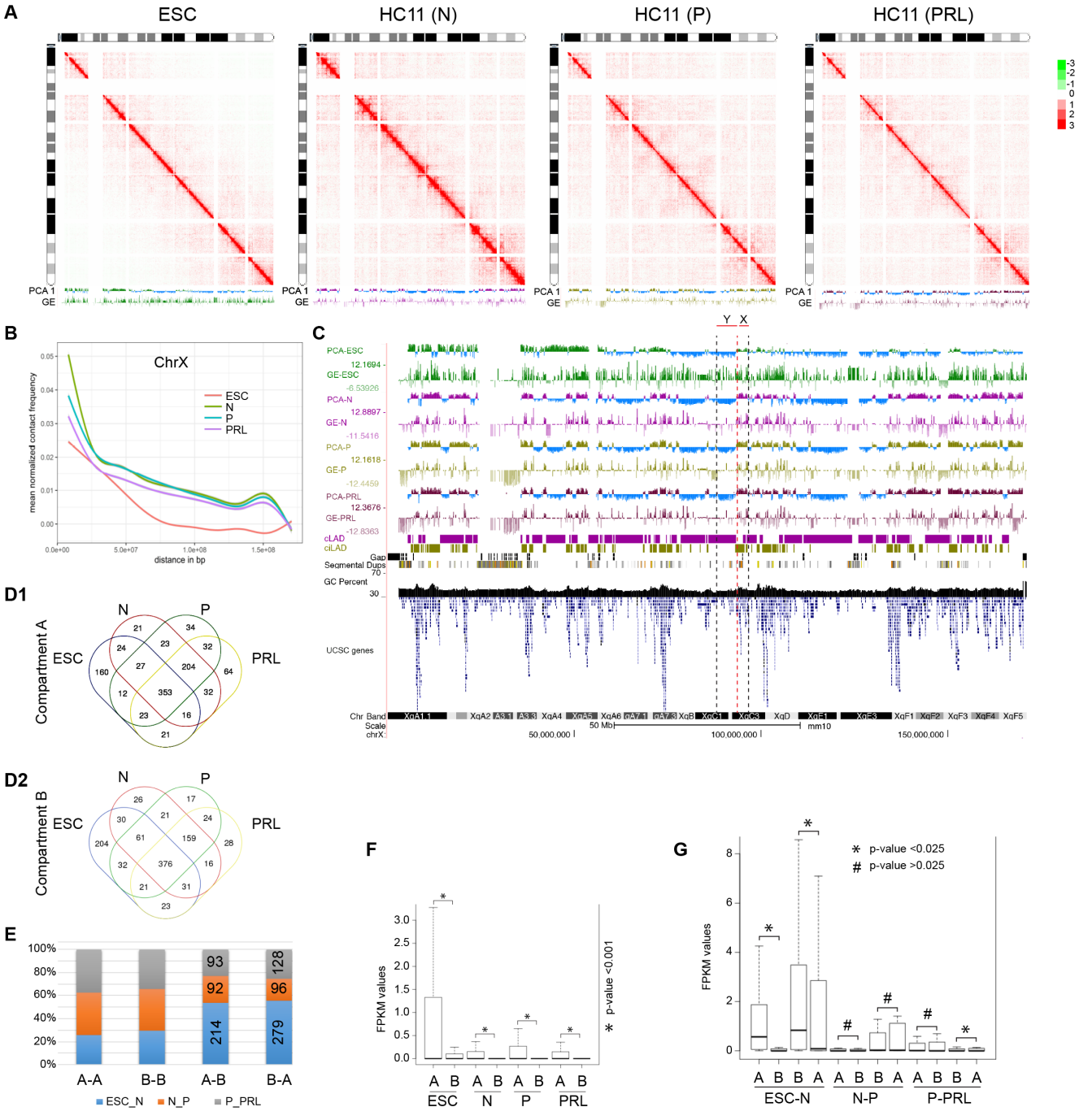

A
ESC
HC11 (N)
HC11 (P)
HC11 (PRL)
PCA 1
GE
B
ChrX
ESC
N
P
PRL
mean normalized contact frequency
distance in bp
C
PCA-ESC
GE-ESC
PCA-N
GE-N
PCA-P
GE-P
PCA-PRL
GE-PRL
cLAD
ciLAD
Gap
Segmental Dups
GC Percent
UCSC genes
Chr Band
Scale
chrX
D1
Compartment A
N
P
ESC
PRL
D2
Compartment B
E
A-A
B-B
A-B
B-A
ESC_N
N_P
P_PRL
F
FPKM values
p-value <0.001
ESC
N
P
PRL
G
p-value <0.025
p-value >0.025
FPKM values
ESC-N
N-P
P-PRL

Fig.17. Dynamics of reorganization of intra-chromosomal (Chr-X) interactions and gene expression in ESC, N, P and PRL treated HC11 cells: **(A)** ChrX contact frequency heatmap at 100kb resolution in ESC, N, P and PRL treated HC11 cells showing fewer long-range interactions in ESCs when compared to other HC11 cell-types. Visually distinguishable and enriched long-range interactions of certain domains are observed in primed state. Each heat map is overlaid on PCA 1 derived A/B compartments and gene expression (GE) tracks from UCSC, along with chromosome G-banding pattern in X and Y axis respectively, showing existence two kinds of compartments in each state and accompanied gene expression patterns. **(B)** Decay plot showing mean normalized contact frequency in ChrX in ESC (Red), N (green), P (blue) and PRL (pink) along the distance, between genomic regions showing differences in the distributions. **(C)** UCSC snapshot view of Chr6 with respective cell-type specific PCA1, gene expression, cLADs, ciLADs, GC/AT %, gene density and chromosome bands showing dynamic switching of genes between compartments and accompanied differential gene expression patterns. Note the striking relationship between compartment A with ciLADs, GC rich isochores, gene rich domains of the chromosomes (Highlighted lines showing 'X' region), whereas compartment B with cLADs, AT rich isochores and gene poor regions of chromosomes (Highlighted dotted lanes showing 'Y' region). Note the lack of values in the regions of GAP and segmental duplicated regions due to mapping issues. Highlighted lines show relationship of cLADs, ciLADs with gene density, gene expression and GC/AT %. **(D1,2)** Venn diagram showing number of conserved, overlapping and unique A & B compartments in ChrX of ESC, N, P and PRL treated HC11 cells. Note that a large number of compartments are conserved between samples. **(E)** Bar plot showing number of compartments switching between A and B from one cell to another in ChrX. Note dramatic switching of compartments B to A under GC primed (P) condition of HC11 cells (Shown by dotted lanes). **(F)** Relation between A/B compartments and gene expression was shown as boxplots (outliers excluded). The gene expression is generally seen high in compartment A than in B. Wilcoxon rank sum test was performed for statistical testing. **(G)** Boxplot representing the changes in the gene expression (FPKM) of the genes, which are shifting from (A-B, B-A), with outliers excluded. Note that switching from A-B or B-A compartments are not accompanied by decrease or increase in gene expression patterns respectively, between N-P HC11 cells types except A-B switching in P-PRL states of HC11 cells. Significance of difference is depicted in p-values.

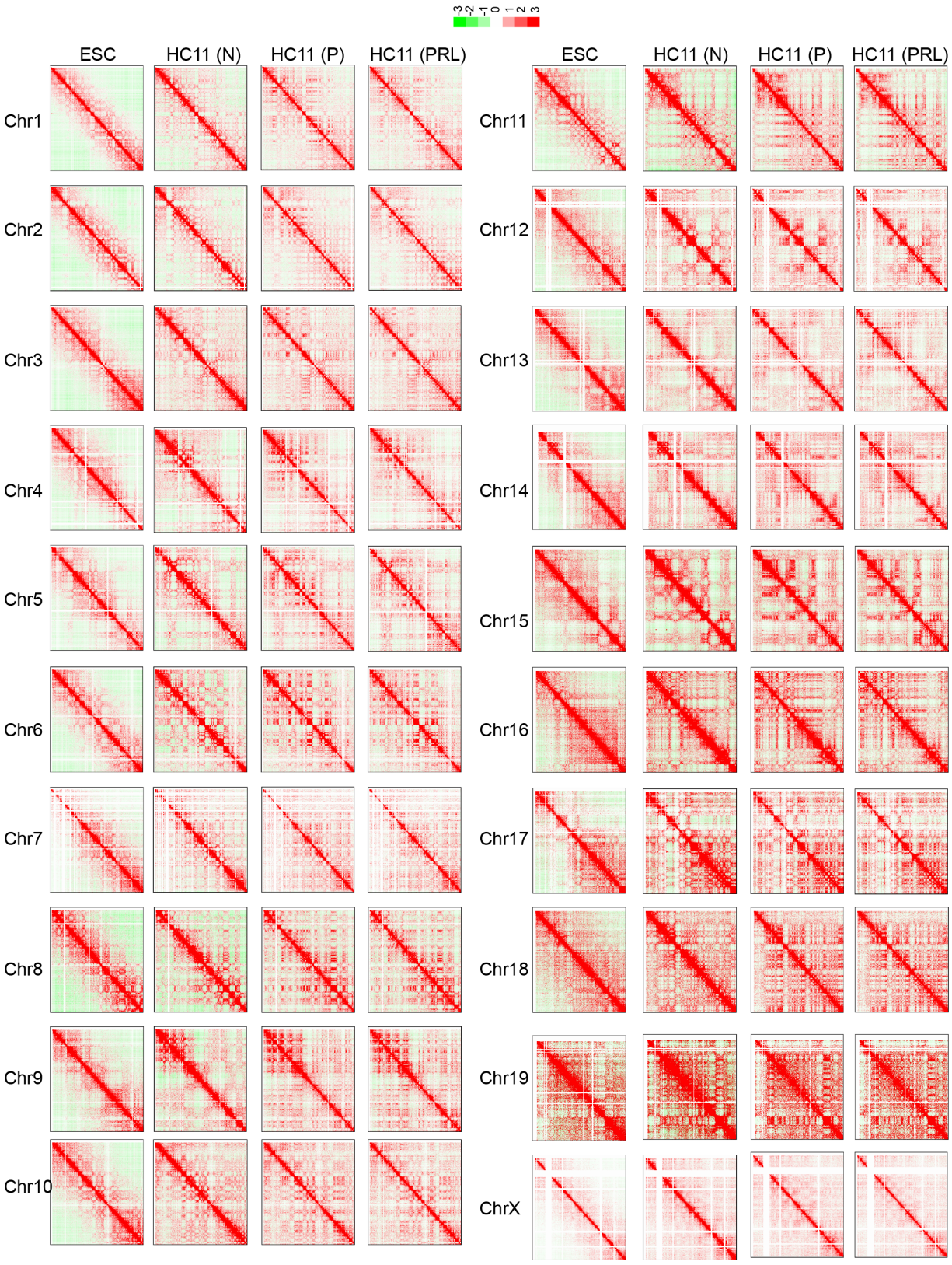

-3 -2 -1 0 1 2 3
ESC
HC11 (N)
HC11 (P)
HC11 (PRL)
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
Chr12
Chr13
Chr14
Chr15
Chr16
Chr17
Chr18
Chr19
ChrX

Fig.18. Dynamic reorganization of intra-chromosomal domain in all mouse chromosomes (1-19+X) across ESC, N, P and PRL treated HC11 cells: Heatmaps of intra-chromosomal interaction matrices of mouse chromosomes 1-19+X in ESCs showing presence of restricted short-range interactions along the diagonal. Notice the presence of fewer long-range intra- chromosomal interactions among all the chromosomes. However, all other HC11 cell-types, show more long-range intra-chromosomal interactions across all the chromosomes. Note the presence of enriched interactions among intra-chromosomal domains in GC primed (P) cells in all the chromosomes and slight reduction in the intensity of those interactions were seen in all the chromosomes of PRL treated HC11 cells. Scale bar represents intensity of chromatin interaction frequency at 1Mb resolution in arbitrary units.

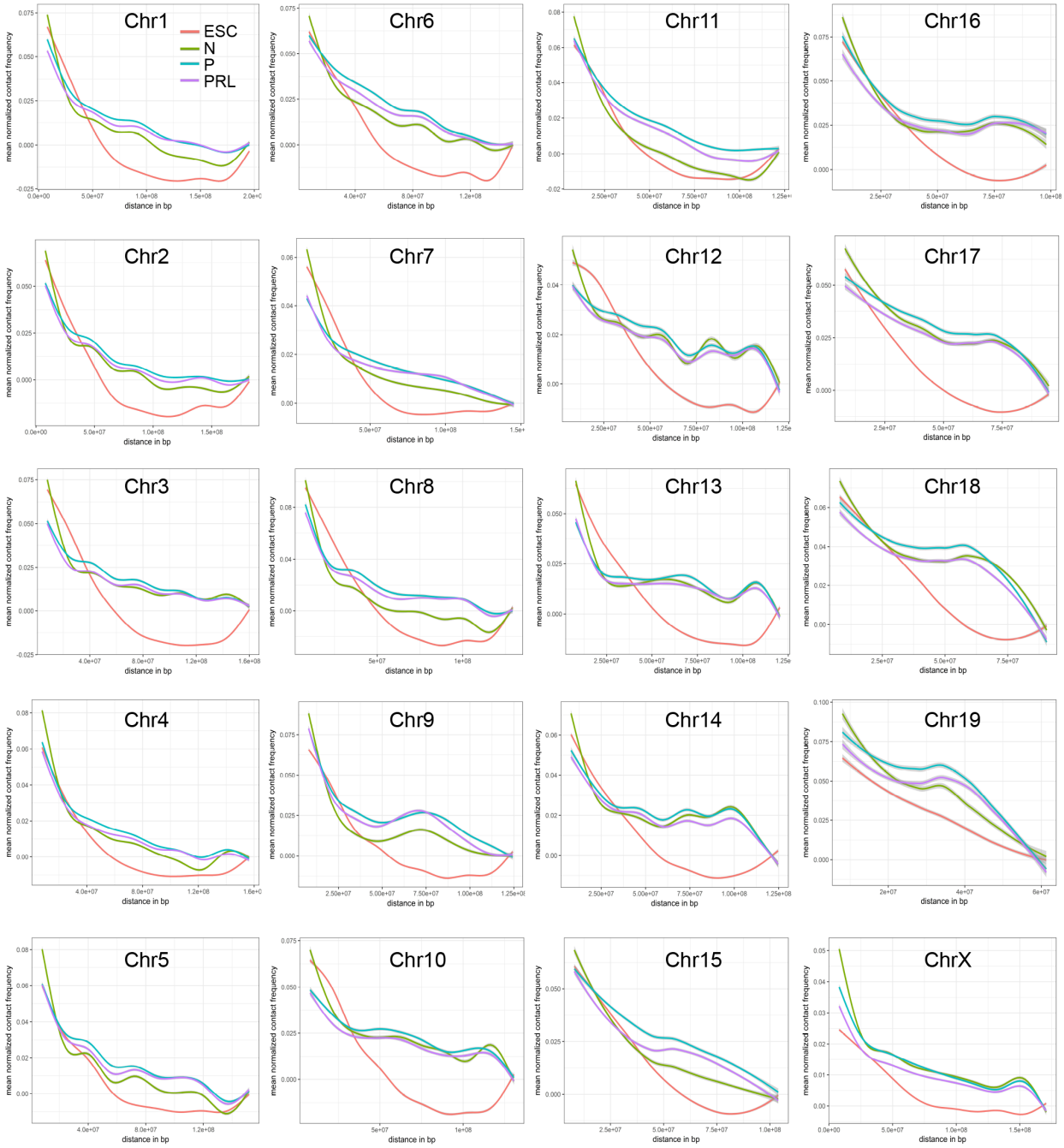

Chr1
ESC
N
P
PRL
Chr6
Chr11
Chr16
Chr2
Chr7
Chr12
Chr17
Chr3
Chr8
Chr13
Chr18
Chr4
Chr9
Chr14
Chr19
Chr5
Chr10
Chr15
ChrX
mean normalized contact frequency
distance in bp

Fig.19. Dynamic switching of short-range and long-range chromatin interactions in ESC and in HC11 cells undergoing lactogenic differentiation revealed by distance decay plots. Decay plot showing mean normalized contact frequency in all the chromosomes in ESC (Red), N (green), P (blue) and PRL (pink) along the distance between genomic regions showing differences in the distributions of short and long-range interactions. Interestingly, note the occurrence of more long-range interactions in GC primed HC11 cells in most of the length in all the chromosomes followed by PRL treated cells and them normal HC11 cells.

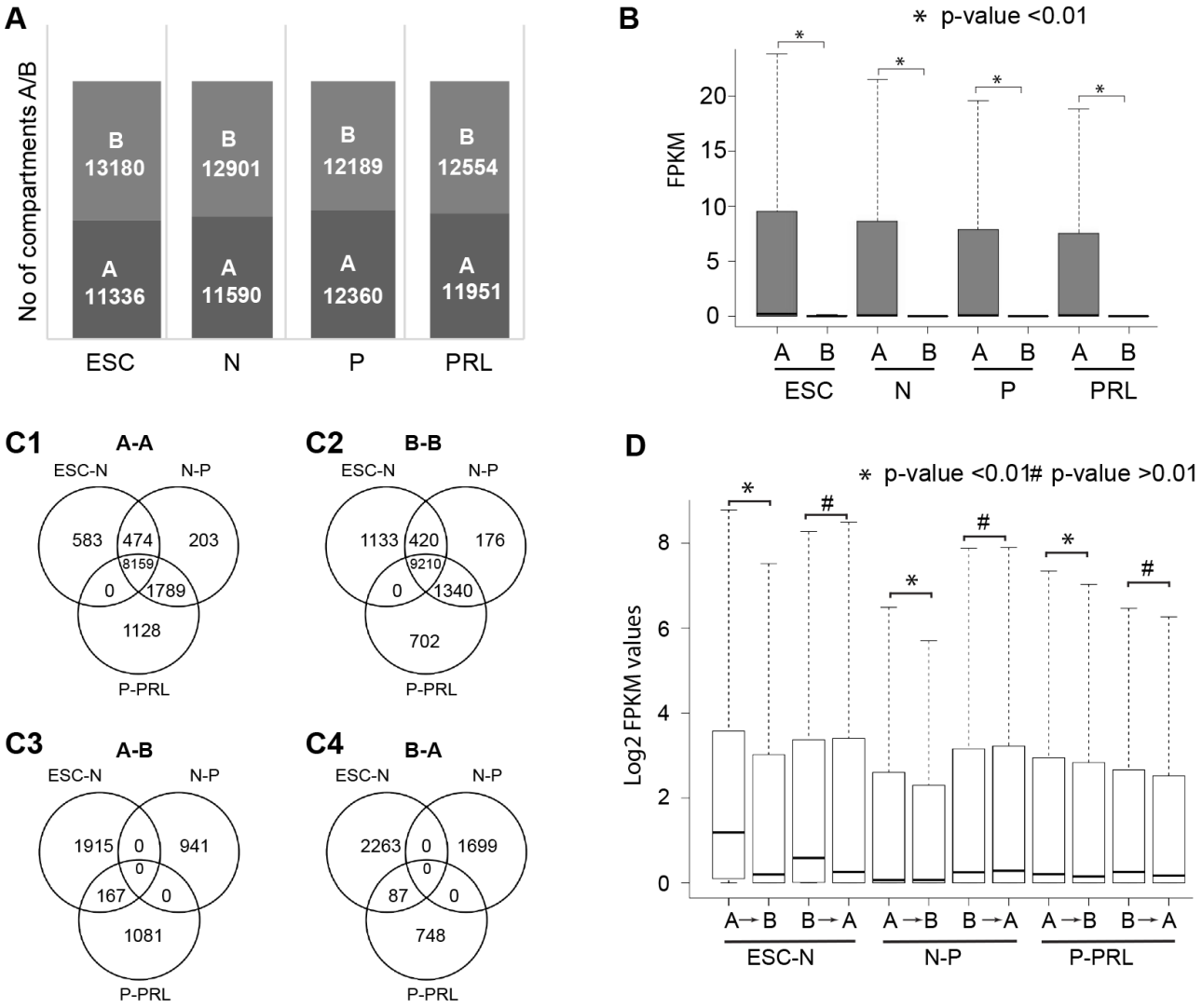

A
No of compartments A/B
B 13180
A 11336
B 12901
A 11590
B 12189
A 12360
B 12554
A 11951
ESC
N
P
PRL
B
* p-value <0.01
FPKM
A B A B A B A B
ESC N P PRL
C1 A-A
ESC-N N-P
583 474 203
8159
0 1789
1128
P-PRL
C2 B-B
ESC-N N-P
1133 420 176
9210
0 1340
702
P-PRL
C3 A-B
ESC-N N-P
1915 0 941
0
167 0
1081
P-PRL
C4 B-A
ESC-N N-P
2263 0 1699
0
87 0
748
P-PRL
D
* p-value <0.01# p-value >0.01
Log2 FPKM values
A→B B→A A→B B→A A→B B→A
ESC-N N-P P-PRL

Fig.20. Genome-wide analysis of relation between A/B compartments and gene expression in ESCs, N, P and PRL treated HC11 cells: **(A)** Bar plot showing total number of A & B compartment in ESCs, N, P & PRL treated HC11 cells. **(B)** Boxplots showing distribution of expression of genes (FPKM values) present in A and B genomic compartments across ESC, N, P and PRL treated HC11 cells (outliers excluded), showing global nature of higher gene expression in A compartment, compared to compartment B. Venn diagram showing conservation of A-A **(C1)**, B-B **(C2)** compartments and compartment A-B **(C3)** and B-A **(C4)** switching between ESC-N, N-P and P-PRL treated HC11 cells. Note the dramatic switching of B-A and A-B compartments under primed and PRL states of HC11 cells respectively. **(D)** Distribution of gene expression (FPKM values) of genes switching from compartments A-B or B-A between ESC-N, N-P and P-PRL treated HC11 cells (outliers excluded). The plot shows significant difference when genes are switching from A-B compartment from ESC to N, N to P and P to PRL states, however, expression in genes switching from B-A compartments does not show significant difference. Although, there is statistically significant change in gene expression of genes when they switch from A-B, but they do not correlate with the nature of A/B compartments. There is no statistical significance of gene expression difference in B-A.

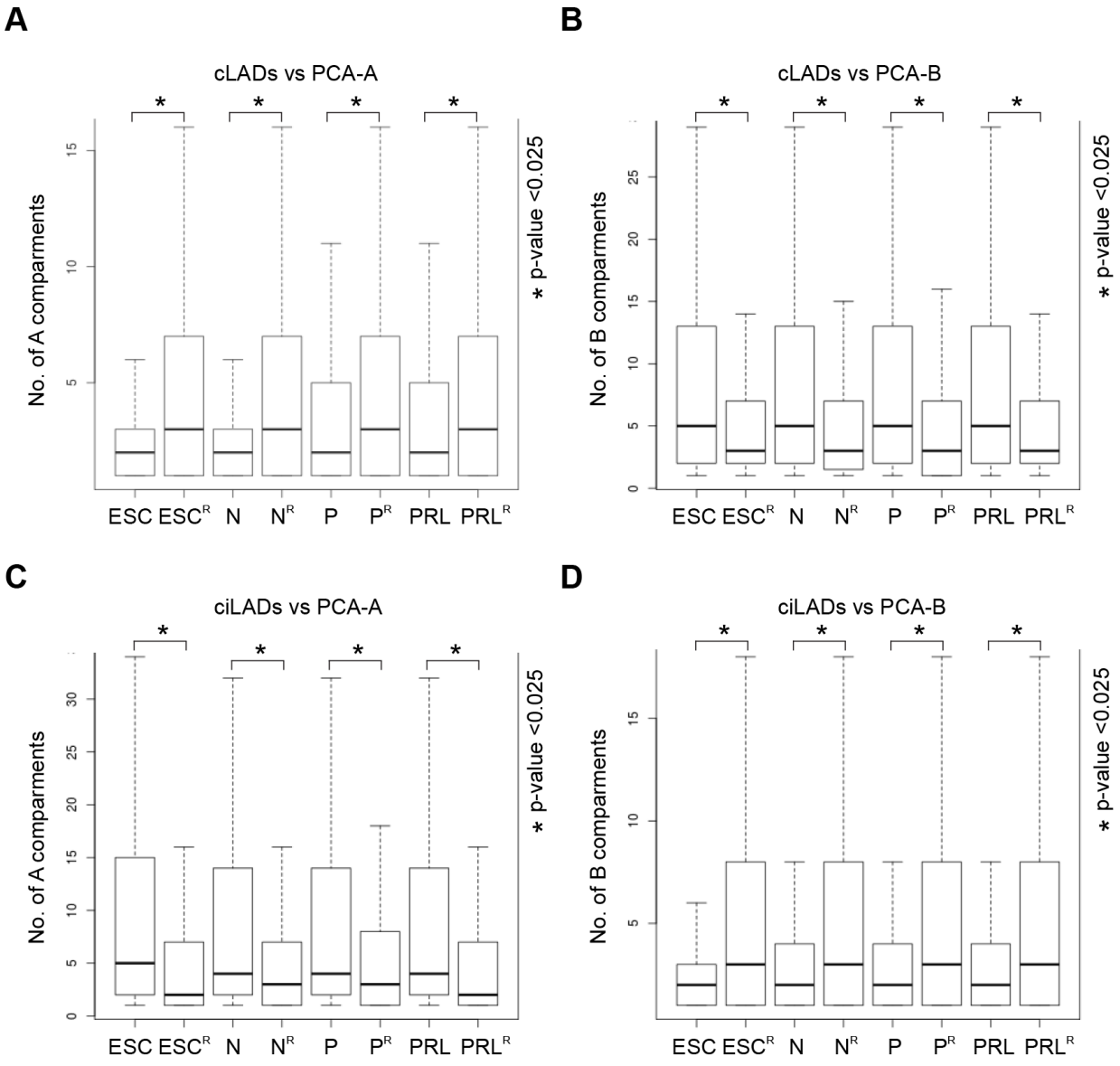

A
cLADs vs PCA-A
No. of A comparments
* p-value <0.025
ESC ESC^R N N^R P P^R PRL PRL^R
B
cLADs vs PCA-B
No. of B comparments
* p-value <0.025
ESC ESC^R N N^R P P^R PRL PRL^R
C
ciLADs vs PCA-A
No. of A comparments
* p-value <0.025
ESC ESC^R N N^R P P^R PRL PRL^R
D
ciLADs vs PCA-B
No. of B comparments
* p-value <0.025
ESC ESC^R N N^R P P^R PRL PRL^R

Fig.21.Relationship between Genomic A & B compartments with c/ciLADs in ESC, N, P and PRL treated HC11 cells: Boxplots showing number of A compartments within the coordinates of cLAD regions in each sample of ESC, N, P and PRL treated HC11 cells, in comparison with number of randomized PCA generated A compartments specific cLADs of the same sample (ESC^R, N^R, P^R and PRL^R) (outliers excluded) **(A)**, and compartment B **(B),** and conserved inter LADs in compartment A **(C),** and B **(D).** Paired Wilcox test was used for statistical testing and significance of each are represented above each pair.

But when we plotted expression of the genes which switch between the cell-types, there was no causal correlation between gene expression and switch in A and B compartments in chr6 (Fig. 14G), chr1 (Fig. 15G), chr15 (Fig. 16G), chrX (Fig. 17G) and is also true genome-wide (Fig. 20D). These results suggest that genes present in A compartment in general show higher expression that that of in B compartment, but there is no significant change in the gene expression of the genes present in the switching compartmets (A-B or B-A) between any two cell stages. the switching compartments (A-B or B-A) between any two cell stages. However, it is important to note that resolution that we performed PCA is of 100kb, hence it is possible that effect of few genes within 100kb cannot be attributed to the entire 100kb region. Nevertheless, these results demonstrate that GC alone and in combination with PRL promotes differential expression of thousands of genes in nonproliferative HC11 cells by selectively switching chromatin predominantly from compartment B-A and A-B but were not accompanied by significant increase and decrease in gene expression respectively.

14. **Dynamic reorganization of alternative intra and inter-chromosomal cLADs and ciLADs interactions, in succession accompanies lactogenic differentiation of HC11 cells:** Gain of long-range interactions upon GC and PRL and selective switching of chromatin regions within compartment B and A respectively for GC and PRL treatments without changing gene expression patterns prompted us to investigate nature of these interactions under GC and PRL signalling. It is evident from the above results that changes in the long-range interactions are not uniform but were concentrated at certain regions. When the heatmaps were overlaid on PCA, we could clearly see that the regions corresponded to intra-chromosomal A and B compartments. In line with previous observations (Jabbari and Bernardi, 2017), we find a concordance of compartment B with conserved lamina associated domains (cLADs) and compartment A with conserved inter lamina associated domains (ciLADs) in all four cell-types (Fig. 21), In the light of above results, we wanted to further study the dynamics of interactions corresponding to cLADs and ciLADs during lactogenic differentiation of HC11 cells. To this end, we generated chromosome subtraction matrices of interaction frequencies between cell-types (ESC-N, N-P and P-PRL treated HC11 cells) of all chromosomes to visualize the differences in chromatin contact frequencies (Fig. 26). As the off-diagonal long-range interactions are not visible at regular scale, we adjusted the scale to visualize the interactions. We noted that increase in long-range contacts between ESC and normal HC11 cells, span across the regions of chromosomes (Fig. 22A1), whereas they were restricted to alternative domains (X or Y) in GC primed (Fig. 22A2) and PRL treated HC11 cells (Fig.

22A3). By overlaying the subtraction matrices over the PCA and c/ciLADs, we found that the increase in the long-range interactions between ESC and normal HC11 cells span both cLADs (X domain) and ciLADs (Y domain) (Fig. 22B1). Whereas they were predominantly from cLADs (X domain) in subtraction matrices of normal and GC primed HC11 cells (Fig. 22B2) and were predominantly from ciLADs between GC primed and PRL treated HC11 cells (Fig. 22B3). Similar trends were observed even in chr1 (Fig. 23), chr15 (Fig. 24) and chrX (Fig. 25). Next, we wanted to assess the nature of inter-chromosomal interactions and generated chr6 vs genome interactions with vanilla coverage normalized contacts. Concordant with the previous results, we observed that ciLADs of chr6 are interacting with certain regions of all the chromosomes (Fig. 27A), interestingly, the pattern of interactions were quite different between ESC and HC11 cell types. In ESCs, we noticed that, apart from interactions at telomeric regions, broadly four different locations of chr6, which coincide with ciLADs, are preferentially making contacts with some regions of all other chromosomes. Further, we noted that these inter-chromosomal contacts are enriched in HC11 (N) cells and gradually seen to diminish under primed and PRL conditions. In order to understand the importance of c/ciLAD interactions, we quantified both intra and inter-chromosomal interactions between cLAD-cLAD (L-L), ciLAD-ciLADs (i-i) and cLAD-ciLAD (L-i) interaction in ESC, Normal, GC primed and PRL treated HC11 cells. Intra-chromosomally, we found L-L interactions were significantly increased from ESC to normal HC11 cells and is further enhanced under GC primed and later slightly decreased in PRL treated HC11 cells (Fig. 28A1). Same trend was seen in 16 out of 20 other mouse chromosomes as well (Fig. 29). PRL treatment leads to dramatic reduction in L-L interaction in all the chromosomes. Further, our genome-wide analysis of intra-chromosomal interactions (cis) between L-L does show similar trend (Fig. 28B1). Inter-chromosomally, genome-wide, normal HC11 cells showed higher L-L than that of ESCs. Interestingly, these interactions were decreased in GC primed state and then moderately increased in PRL state (Fig. 28C1). Similarly, we quantitated intra-chromosomal interactions among (i-i) domains in ESC, normal, GC primed and PRL treated HC11 cells. We observed that i-i interactions were significantly increased from ESC to normal HC11 in chr6 (Fig. 28A2) as well as genome-wide (Fig. 28B2, 30). However, they were dramatically reduced in all the chromosomes in GC primed cells and they seemed to be maintained even under PRL condition genome-wide. Inter-chromosomally, interactions in HC11 normal cells were drastically increased compared to ESC, but they are further decreased in GC primed and then to PRL state (Fig. 28C2).

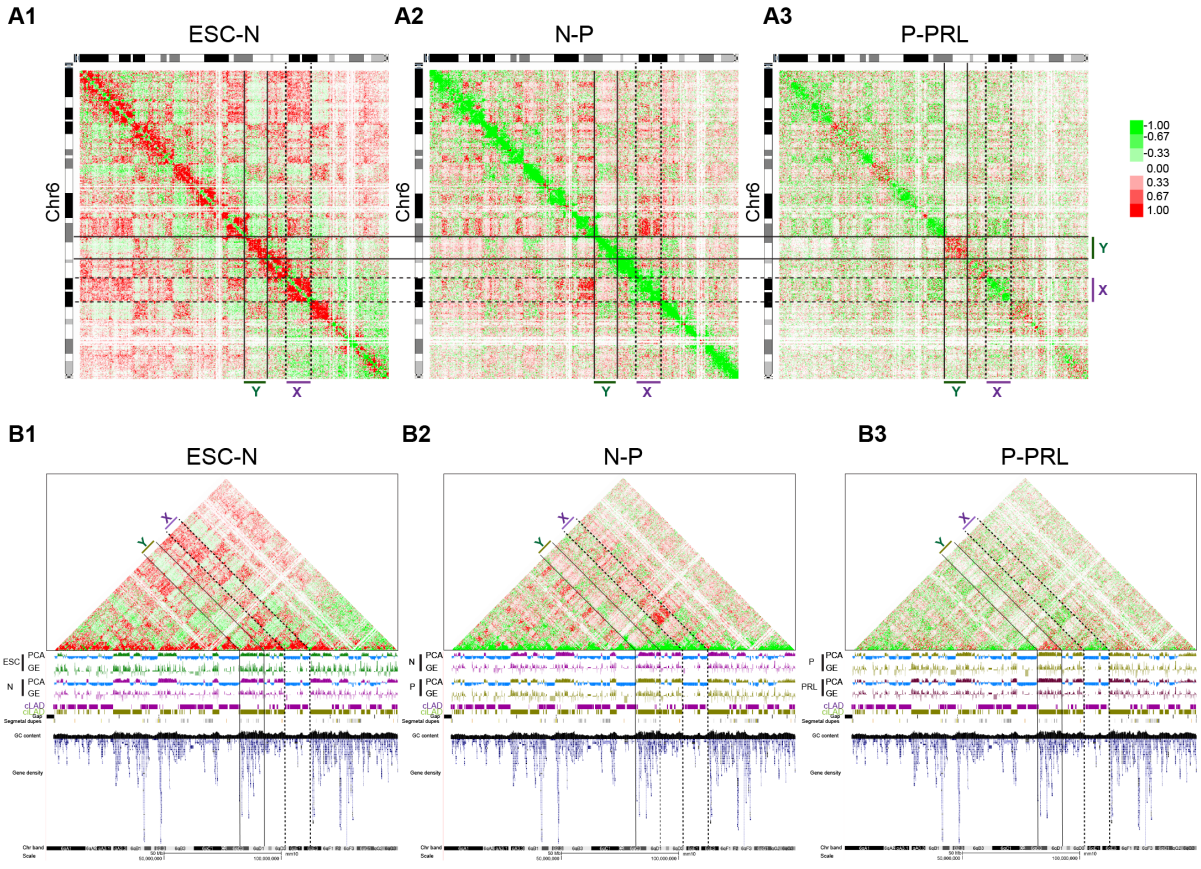


Fig.22. Dynamics and nature of reorganized intra-chromosomal sub-domains revealed by chromosome specific subtraction matrices between ESC and different states of HC11 cells: **(A1)** Chr6 specific subtraction matrix heatmaps between ESCs and normal HC11 cells (N), showing gain of long-range intra-chromosomal interactions (Red colour) across the chromosome (includes both 'X' and 'Y' regions highlighted in dotted and solid lanes respectively), whereas they seem to be predominantly restricted to alternating domains of 'X' or 'Y' in GC primed and PRL treated HC11 cells respectively, as revealed in subtraction matrix between N-P **(A2)** and P-PRL states **(A3)** respectively. Scale bar showing the intensity of interaction frequency between 100kb segments arbitrary units. **(B1)** Integration of subtraction matrices with PCA1 derived A and B compartments, gene expression (GE), conserved LADs (cLADs), conserved inter LADs (ciLADs), GC/AT %, gene density and chromosomal G-bands reveals the nature of alternative domains of 'X' (dotted lanes) as predominantly derived from compartment B with less gene expression, correlated with gene poor, AT rich cLAD, whereas 'Y' domains (solid lanes) as predominantly derived from compartment A with higher gene expression, correlated with gene rich, GC rich ciLADs. Subtraction matrix between ESCs and HC11 cells (N) show gain of interactions in both 'X' and 'Y' domains, whereas gain in interactions are predominantly at 'X' domain in N-P **(B2)** and 'Y' domain in P-PRL states **(B3)** of HC11 cells**.**

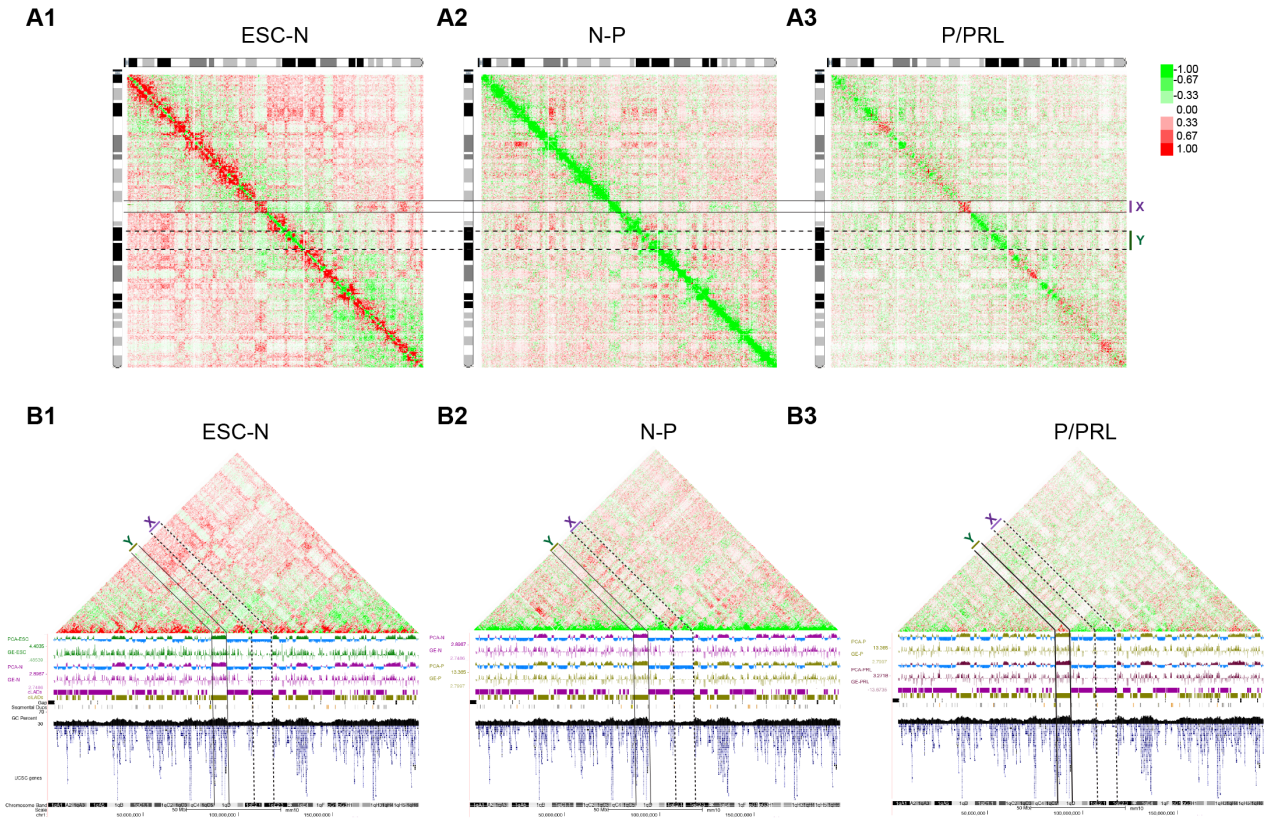


Fig.23. Dynamics and nature of reorganized Chr1 intra-chromosomal sub-domains revealed by chromosome specific subtraction matrices between ESC and different states of HC11 cells: **(A1)** Chr1 specific subtraction matrix heatmaps between ESCs and normal HC11 cells (N), showing gain of long-range intra-chromosomal interactions (Red colour) across the chromosome (includes both 'X' and 'Y' regions highlighted in dotted and solid lanes respectively), whereas they seems to be predominantly restricted to alternating domains of 'X' or 'Y' in GC primed and PRL treated HC11 cells respectively, as revealed in subtraction matrix between N-P **(A2)** and P-PRL states **(A3)** respectively. Scale bar showing the intensity of interaction frequency between 100kb segments. **(B1)** Integration of subtraction matrices with PCA1 derived A and B compartments, gene expression (GE), conserved LADs (cLADs), conserved inter LADs (ciLADs), GC/AT %, gene density and chromosomal G-bands reveals the nature of alternative domains of 'X' (dotted lanes) as predominantly derived from compartment B with less gene expression, correlated with gene poor, AT rich cLAD, whereas 'Y' domains (solid lanes) as predominantly derived from compartment A with more gene expression, correlated with gene rich, GC rich ciLADs. Subtraction matrix between ESCs and HC11 cells (N) show gain of interactions in both 'X' and 'Y' domains, whereas gain in interactions are predominantly at 'X' domain in N-P **(B2)** and 'Y' domain in P-PRL states **(B3)** of HC11 cells**.**

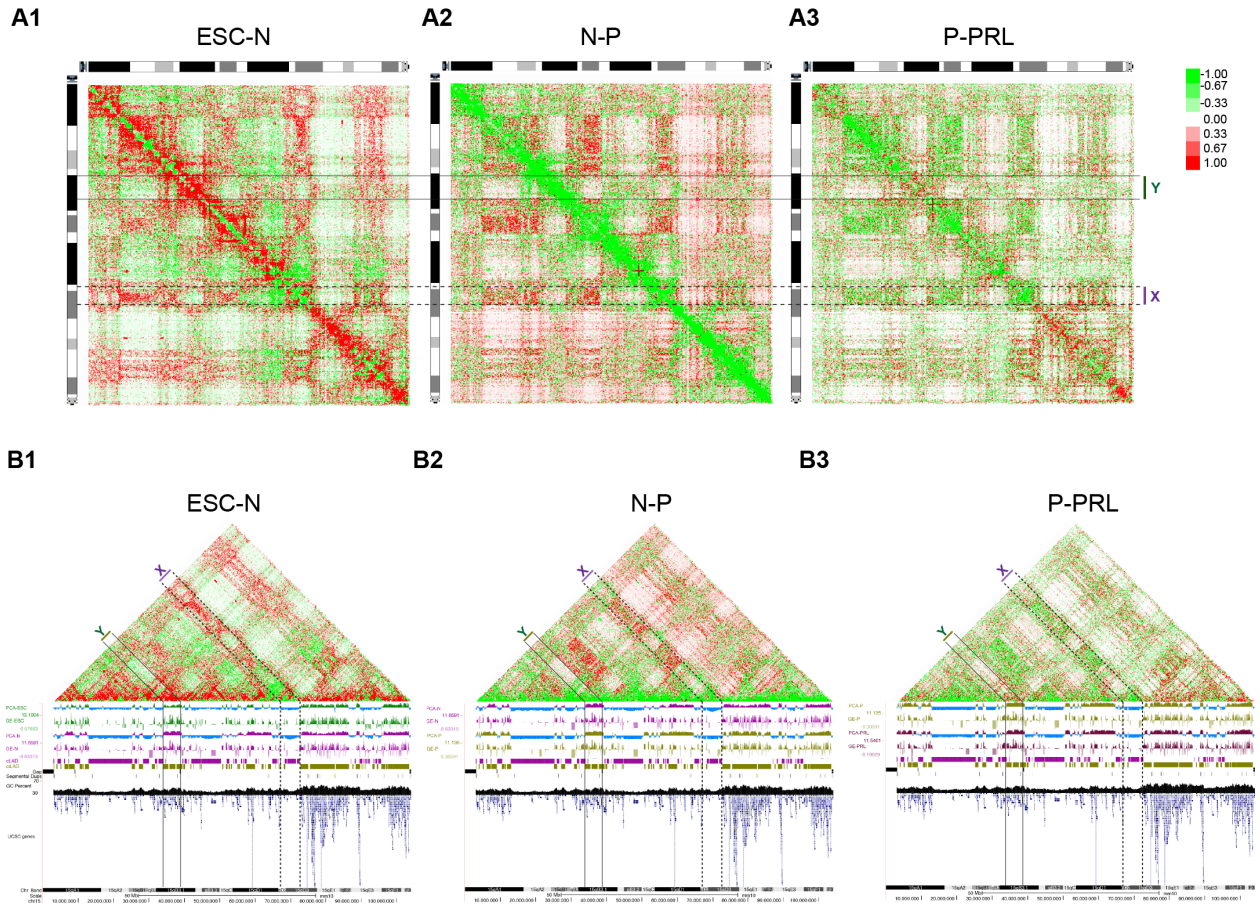


Fig.24. Dynamics and nature of reorganized Chr15 intra-chromosomal sub-domains revealed by chromosome specific subtraction matrices between ESC and different states of HC11 cells: **(A1)** Chr15 specific subtraction matrix heatmaps between ESCs and normal HC11 cells (N), showing gain of long-range intra-chromosomal interactions (Red colour) across the chromosome (includes both 'X' and 'Y' regions highlighted in dotted and solid lanes respectively), whereas they seems to be predominantly restricted to alternating domains of 'X' or 'Y' in GC primed and PRL treated HC11 cells respectively, as revealed in subtraction matrix between N-P **(A2)** and P-PRL states **(A3)** respectively. Scale bar showing the intensity of interaction frequency between 100kb segments. **(B1)** Integration of subtraction matrices with PCA1 derived A and B compartments, gene expression (GE), conserved LADs (cLADs), conserved inter LADs (ciLADs), GC/AT%, gene density and chromosomal G-bands reveals the nature of alternative domains of 'X' (dotted lanes) as predominantly derived from compartment B with less gene expression, correlated with gene poor, AT rich cLAD, whereas 'Y' domains (solid lanes) as predominantly derived from compartment A with more gene expression, correlated with gene rich, GC rich ciLADs. Subtraction matrix between ESCs and HC11 cells (N) show gain of interactions in both 'X' and 'Y' domains, whereas gain in interactions are predominantly at 'X' domain in N-P **(B2)** and 'Y' domain in P-PRL states **(B3)** of HC11 cells.

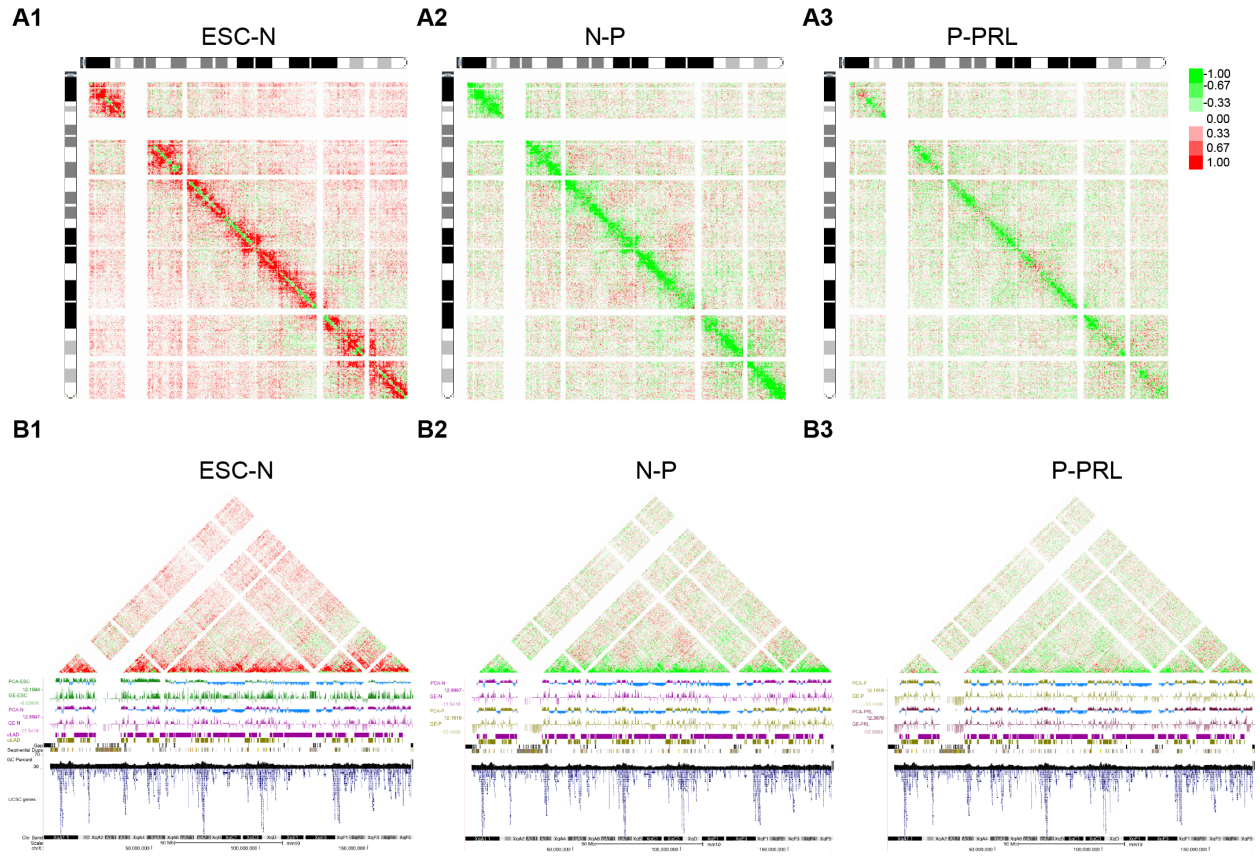


Fig.25. Dynamics and nature of reorganized ChrX intra-chromosomal sub-domains revealed by chromosome specific subtraction matrices between ESC and different states of HC11 cells: **(A1)** ChrX specific subtraction matrix heatmaps between ESCs and normal HC11 cells (N), showing gain of long-range intra-chromosomal interactions (Red colour) across the chromosome (includes both 'X' and 'Y' regions highlighted in dotted and solid lanes respectively), whereas they seems to be predominantly restricted to alternating domains of 'X' or 'Y' in GC primed and PRL treated HC11 cells respectively, as revealed in subtraction matrix between N-P **(A2)** and P-PRL states **(A3)** respectively. Scale bar showing the intensity of interaction frequency between 100kb segments. **(B1)** Integration of subtraction matrices with PCA1 derived A and B compartments, gene expression (GE), conserved LADs (cLADs), conserved inter LADs (ciLADs), GC/AT%, gene density and chromosomal G-bands reveals the nature of alternative domains of 'X' (dotted lanes) as predominantly derived from compartment B with less gene expression, correlated with gene poor, AT rich cLAD, whereas 'Y' domains (solid lanes) as predominantly derived from compartment A with more gene expression, correlated with gene rich, GC rich ciLADs. Subtraction matrix between ESCs and HC11 cells (N) show gain of interactions in both 'X' and 'Y' domains, whereas gain in interactions are predominantly at 'X' domain in N-P **(B2)** and 'Y' domain in P-PRL states **(B3)** of HC11 cells

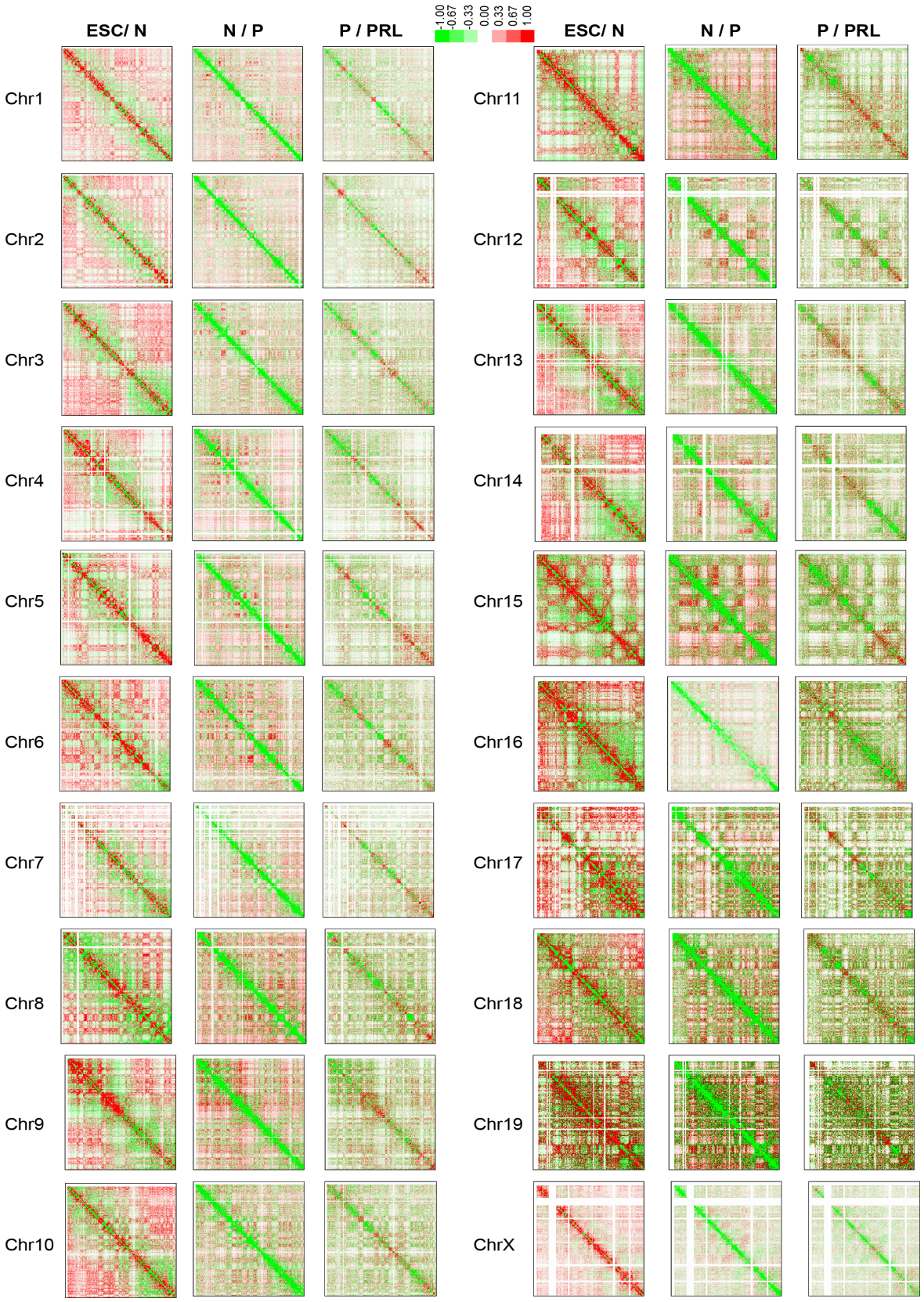

ESC/ N
N / P
P / PRL
-1.00
-0.67
-0.33
0.00
0.33
0.67
1.00
ESC/ N
N / P
P / PRL
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
Chr12
Chr13
Chr14
Chr15
Chr16
Chr17
Chr18
Chr19
ChrX

Fig.26. Dynamic reorganization of intra-chromosomal sub-domains in ESC and HC11 cells during their lactogenic differentiation revealed by subtraction matrices between cell-types. 100kb resolution subtraction heatmaps of simpleNorm normalized interaction matrices between ESC-N, N-P and P-PRL states of HC11 cells, showing gain of long-range interactions across the length of chromosomes in between ESC-HC11 (N) cells in all of the mouse chromosomes. However, these interactions seen to be restricted predominantly, to certain intra-chromosomal sub-domains (Compartment B) in all the chromosomes in between HC11 (N) and GC primed HC11 cells. Interestingly PRL treated HC11 show, enrichment in long-range interaction in adjacent domains (compartment A) of GC induced sub-chromosomal domains in all the chromosomes. Scale bar represents intensity of chromatin interactions in arbitrary units at 1Mb resolution.

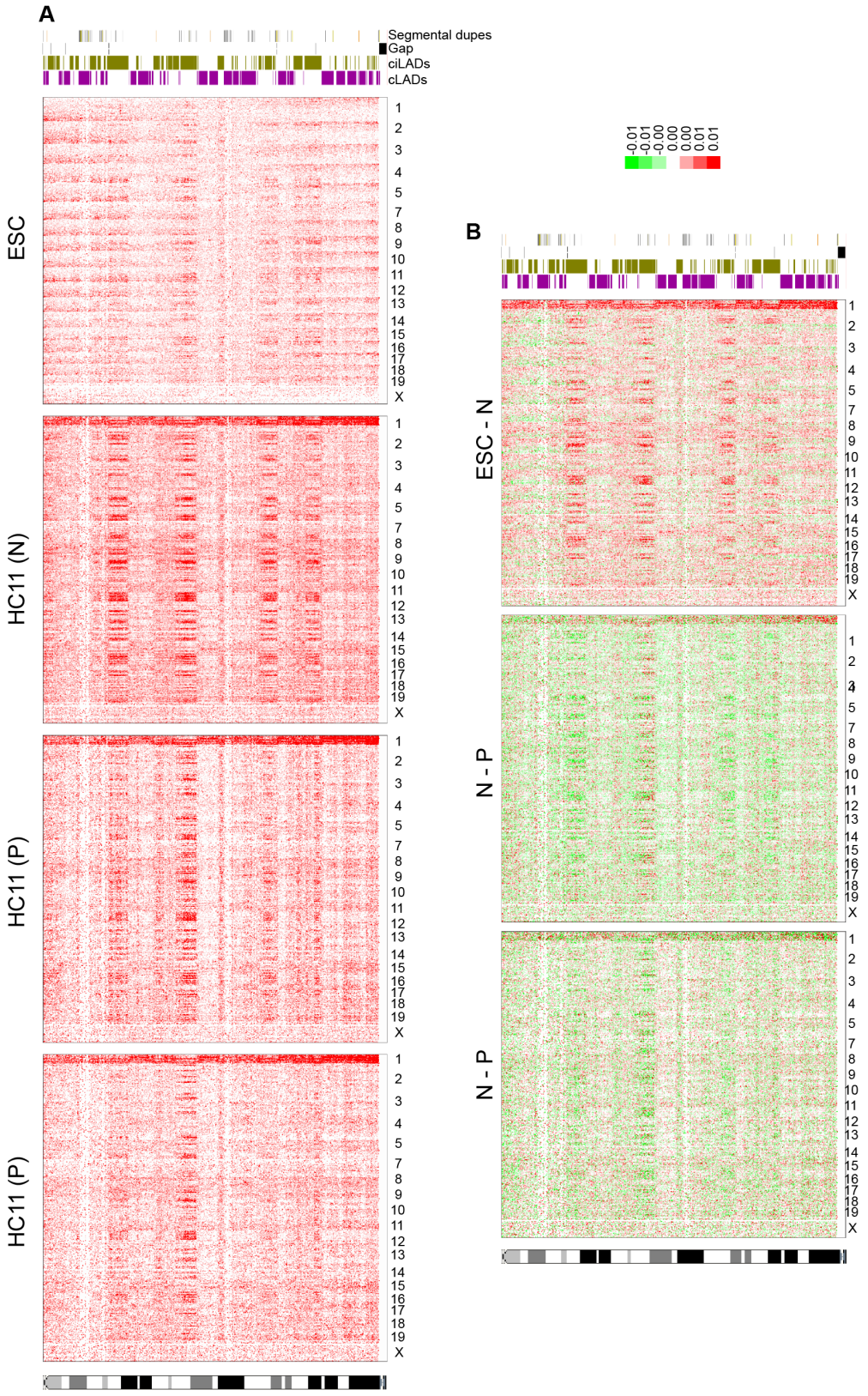

A
Segmental dupes
Gap
ciLADs
cLADs
ESC
HC11 (N)
HC11 (P)
HC11 (P)
-0.01
-0.01
-0.00
0.00
0.00
0.01
0.01
B
ESC - N
N - P
N - P

Fig.27.Dynamics and nature of reorganized inter-chromosomal sub-domains revealed by Chr6 specific genome-wide inter-chromosomal interaction analysis in ESC and different states of HC11 cells: (A). VC normalized 250kb resolution, whole genome inter-chromosomal contact matrix heatmaps of Chr6 vs entire genome representing all other Nineteen mouse (1-5, 7-19 and X) chromosomes in ESC, HC11 (N), (P) and (PRL) cell conditions, showing preferential interactions of multiple regions of Chr6 encompassing both c/ciLADs, making contacts at different locations of rest of the chromosomes genome-wide. Heat maps are overlaid with c/ciLADs (Top) and Chromosome G-banding pattern (Bottom). **(B).** Subtraction heatmaps of VC normalized 250kb resolution Chr6 inter-chromosomal matrix ore nineteen other mouse chromosomes between ESC-N, N-P and P-PRL states of HC11 cells showing preferential enrichment of sub-domains encompassing either c/ciLADs. Heat maps are overlaid with c/ciLADs (Top) and Chromosome G-banding pattern (Bottom). Scale bar represent frequency of interactions. All contact frequencies are in arbitrary units.

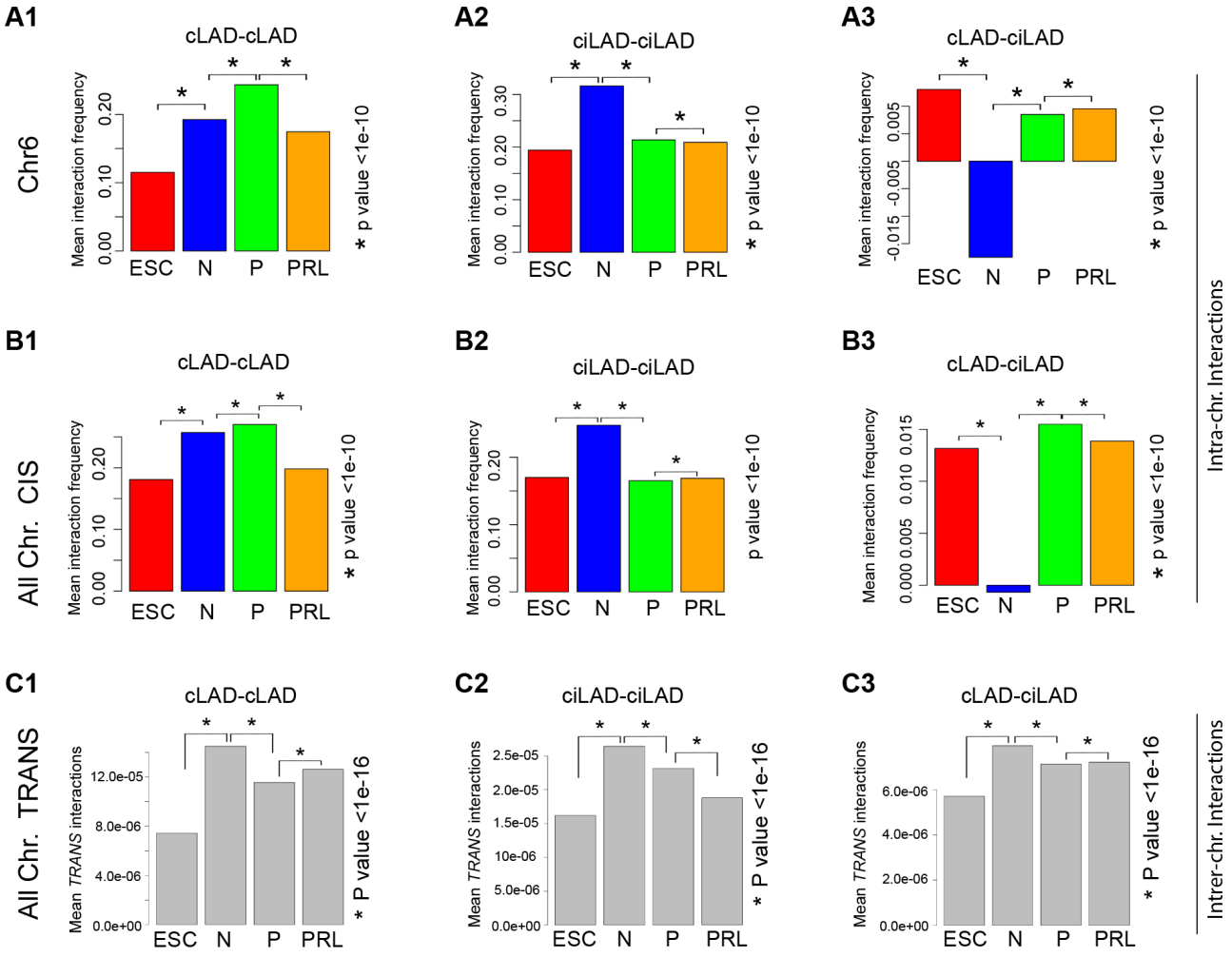

A1
A2
A3
Chr6
cLAD-cLAD
ciLAD-ciLAD
cLAD-ciLAD
Mean interaction frequency
ESC N P PRL
* p value <1e-10
B1
B2
B3
All Chr. CIS
p value <1e-10
Intra-chr. Interactions
C1
C2
C3
All Chr. TRANS
Mean TRANS interactions
* P value <1e-16
Intrer-chr. Interactions

Fig.28. **(A1) Dynamics reorganization of intra as well as inter-chromosomal cLADs and ciLADs in ESC and HC11 cells undergoing lactogenic differentiation: (A1)** Bar plot showing mean interaction frequency among regions of cLADs of Chr6 in ESC, N, P and PRL treated HC11 cells, showing significant gain of interaction in N, P and PRL treated HC11 cells in comparison with ESCs. Note significant gain and loss of interactions between cLADs under (P) and (PRL) conditions respectively. **(A2)** Bar chart showing mean interaction frequency between regions of ciLADs of Chr6 in ESC, N, P and PRL treated HC11 cells showing significantly higher interaction in HC11 (N) cells, whereas they decrease significantly in HC11 (P) and PRL states. **(A3)** Bar chart showing mean interaction frequency between regions of cLAD vs ciLADs of Chr6, showing significant drop from ESC to N but seen to gain in P and PRL states of HC11 cells. **(B1)** Bar diagram showing mean interaction frequency of genome-wide Intra-chromosomal cLADs showing significant rise in interactions between ESC to N state and further in (P) state. However, significant reduction in interactions are seen between P and PRL states. **(B2)** Bar diagram showing significant gain of intra-chromosomal interactions among ciLADs between ESC to HC11 (N) states. There is significant reduction in interactions seen further in primed (P) condition. However, interactions increase from P to PRL states of HC11 cells. **(B3)** Bar graphs showing significant presence of cross-domain interactions between cLADs and ciLADs in ESCs. No such significant interactions are seen in HC11 (N) cells. However, dramatic rise in cLAD-ciLAD interaction are seen in P and PRL states HC11 cells. **(C1)** Bar plot showing genome-wide mean inter-chromosomal interaction frequency of cLADs are significantly increased from ESC to HC11 (N) cells following significant reduction under (P) condition and were seen to rise significantly under PRL state. **(C2)** Bar plot showing significant rise inter-chromosomal contact of ciLADs between ESC and HC11 (N) cells and they were seen to reduce upon GC and further under PRL states. **(C3)** Bar plot showing significant rise in inter-chromosomal contacts between cLAD-ciLADs between ESC and HC11 (N) cells and reduction, between N-P and were significantly increased between P and PRL states of HC11 cells. Paired student's t-test was used for statistical testing.

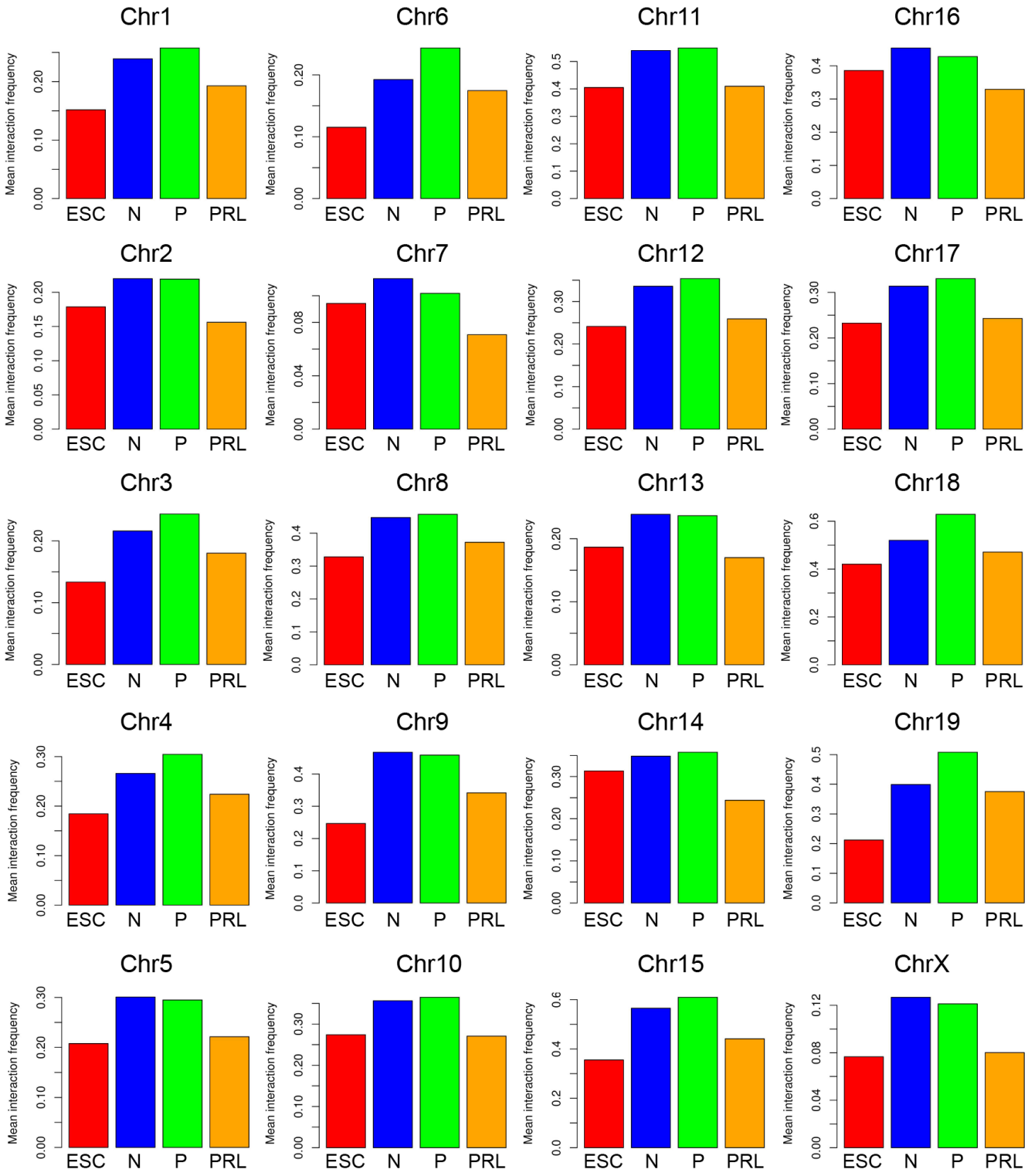

Chr1
Chr6
Chr11
Chr16
Chr2
Chr7
Chr12
Chr17
Chr3
Chr8
Chr13
Chr18
Chr4
Chr9
Chr14
Chr19
Chr5
Chr10
Chr15
ChrX
Mean interaction frequency
ESC
N
P
PRL

Fig.29. Dynamics of intra-chromosomal cLAD-cLAD interactions across all the mouse chromosomes in ESC, N, P and PRL treated HC11 cells, reveal differences: Bar plot showing mean intra-chromosomal interaction frequency between cLADs and cLADs in ESCs (Red colour), normal HC11 cells (N) (Blue colour), GC primed (P) (Green colour) and PRL treated (Orange colour) HC11 cells. Note, significant increase in interaction between cLAD regions from ESC to HC11 (N) cells and further dramatic increase interaction in GC primed HC11 cells in all chromosomes except in chr7, 16 and ChrX. Under PRL condition, there is a dramatic reduction in interactions among cLADs in all the chromosomes.

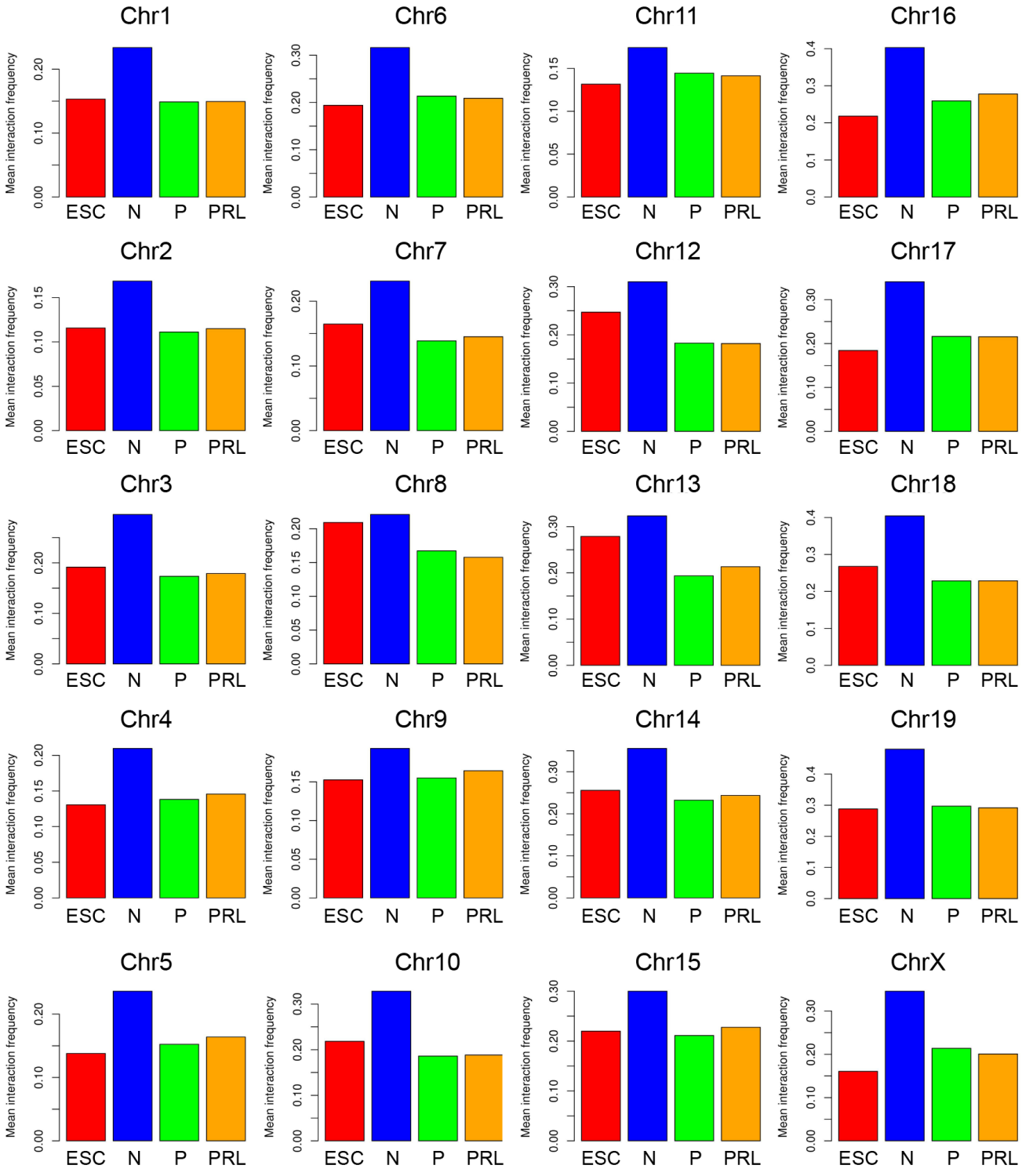

Chr1
Chr6
Chr11
Chr16
Chr2
Chr7
Chr12
Chr17
Chr3
Chr8
Chr13
Chr18
Chr4
Chr9
Chr14
Chr19
Chr5
Chr10
Chr15
ChrX
Mean interaction frequency
ESC
N
P
PRL

Fig.30. Dynamics of intra-chromosomal ciLAD-ciLAD interactions across all the mouse chromosomes in ESC, N, P and PRL treated HC11 cells reveal differences: Bar plot showing mean intra-chromosomal ciLADs and ciLAD interactions in ESCs (Red colour), normal HC11 cells (N) (Blue colour), GC primed (P) (Green colour) and PRL treated (Orange colour) HC11 cells. Note significant increase in interaction between ciLAD regions from ESC to HC11 (N) cells in all of the mouse chromosomes. However, GC and PRL treatment leads to dramatic reduction in ciLAD-ciLAD interactions in all mouse chromosomes.

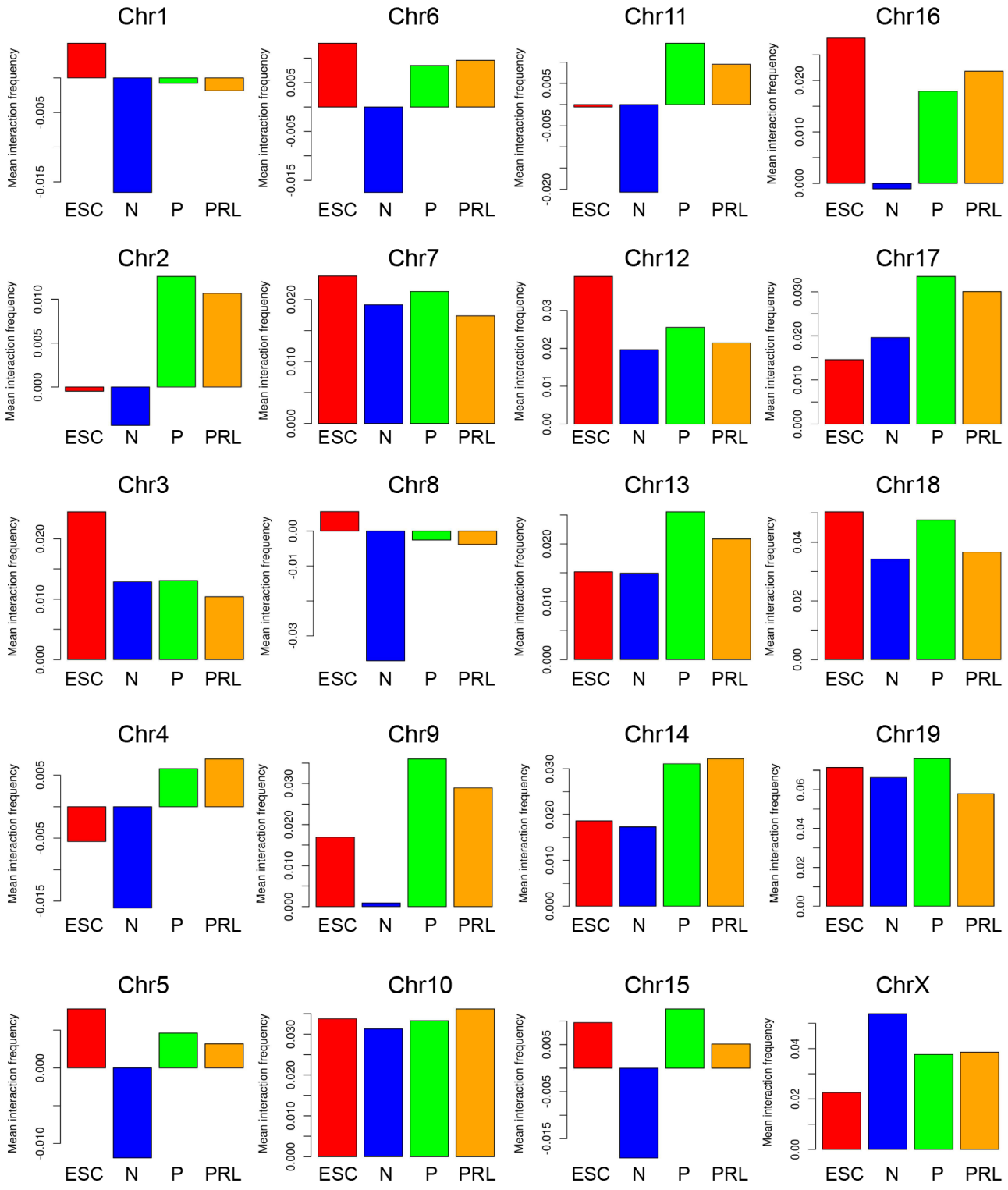

Chr1
Chr6
Chr11
Chr16
Chr2
Chr7
Chr12
Chr17
Chr3
Chr8
Chr13
Chr18
Chr4
Chr9
Chr14
Chr19
Chr5
Chr10
Chr15
ChrX
Mean interaction frequency
ESC
N
P
PRL

Fig.31. Dynamics of intra-chromosomal cLAD-ciLAD interactions across all the mouse chromosomes in ESC, N, P and PRL treated HC11 cells reveal differences. Bar plot showing mean intra-chromosomal cLADs and ciLAD cross-interaction frequencies in ESCs (Red colour), normal HC11 cells (N) (Blue colour), GC primed (P) (Green colour) and PRL treated (Orange colour) HC11 cells. Note significant reduction in cross interaction in normal HC11 cells, though, a few chromosomes (Chr3, 7,10, 13, 14, 18, 19 and X) showed presence of significant interactions when compared to ESCs where there seems to be significantly more interactions in many chromosomes. However, GC priming leads to dramatic raise in interactions between cLAD-ciLADs and it seem to be more or less maintained under PR treated HC11 cells.

Further, our analysis of cross-domain interactions (L-i) intra-chromosomally showed, ESCs to harbour significantly higher interactions compared to normal HC11 cells. Interestingly, we found that in normal HC11 cells there is dramatic reduction i-L interaction are seen in Chr6 (Fig. 28A3), genome-wide (Fig. 28B3, 31) However, closer examination of individual chromosomes such as chr3, 7, 10, 12, 13, 14, 18, 19 and X has shown to harbour interactions between i-L domains (Fig. 31). Interestingly, we observed significant raise in cross domain i-L interaction in GC primed cells in chr6 and also on genomic scale except chr1 and 8. Whereas, i-L interactions among PRL treated HC11 cells significantly gained interactions in chr6 (Fig. 28A3) and also on genomic scale (Fig. 28B3, 31) except for chr1 and 8. It is interesting to note that chr1 and 8 behave very similar to each other compared to rest intra-chromosomally. Surprisingly, Inter-chromosomally, i-L interactions increase in normal HC11 cells, compared to ESC. Further, they were decreased in GC primed and almost maintained in PRL state (Fig. 28C3). These results on intra-chromosomal c/ciLAD interactions suggest that (i) HC11 normal cells acquire long-range contacts encompassing both c/ciLAD regions, (ii) GC priming leads to dramatic enhancement in long-range interaction encompassing cLAD domains, (iii) PRL treatment leads to increase in short-range interaction among ciLAD regions, (iv) HC11 cells dramatically reduce interaction between cLAD-ciLADs, but are dramatically enhanced under GC primed and further reduced in PRL state. Further, our results on inter-chromosomal interaction c/ciLAD interactions suggest existence of dynamic and preferential interaction between and among c/ciLAD regions.

15. **Preferential clustering of cLADs and ciLADs in linked (cis) and unlinked (trans) chromosome are revealed by DNA FISH and statistical analysis on translocated chromosomes respectively:** In order to validate the observed intra-chromosomal preferential clustering of c/ciLADs in ESC and HC11 cells undergoing lactogenic differentiation at single cell level, we have performed DNA FISH with the probes representing cLAD and iLAD regions on chr6 in ESC, normal, GC primed and PRL treated HC11 cells (Fig. 32A). We noticed that significant number of cells show co-association of distant cLAD regions under GC primed conditions (Fig. 32B, B') and co-association of ciLAD regions under PRL condition (Fig. 32C, C'). Similarly, significant number of co-associations between ciLAD and cLAD regions was seen in GC primed cells. These results showed that there is an overall increase in interactions between cLADs in a chromosome upon GC signalling and increase in interactions between ciLAD regions upon PRL signalling. To demonstrate, the preferential clustering of c/ciLAD regions in cis and trans, we took advantage of naturally occurring HC11 cells specific chromosomal translocations

of chr1:6 and chr4:8, where we reasoned that translocated part of other chromosome should behave like a part of *cis* chromosome and interactions of c/ciLAD therein. To evaluate Hi-C genome-wide contact matrix predicted chromosomal translocation between chr4 and 8, we performed chromosome painting on the metaphase spread of proliferating HC11 cells. Under fluorescent microscope, we observed occurrence of full length chromosome 4 (Fig. 33A1) and 8 (Fig. 33A2), along with a chromosome where we observed part of chr4 and 8 fused along with some part of other chromosome (mostly part from chr2 based on Hi-C genome contact matrix) (Fig. 33A3).The Presence of parts of three different chromosomes together in a single region has allowed us to assess the interactions of c/ciLAD regions of translocated chr4 (Chr4-T)vs translocated part of chr8(Chr8-T). As a control, we also assessed interactions of c/ciLAD regions of non-translocated part of chr4(Chr-4NT) vs translocated part of chr8. Subtraction heatmaps between chr4 and chr8 between ESC-N (Fig. 33B1), N-P (Fig. 33B2) and P-PRL (Fig. 33B3) clearly shows enhanced interactions specifically in the translocated portion of chr4 and were seen to follow the trends that were observed under GC primed and PRL cellular states intra-chromosomally. We assessed these observations computationally by performing analysis of c/ciLAD regions of translocated & non-translocated parts of Chr4 vs translocated portion of chr8 in ESCs (Fig. 33C1), normal HC11 cells (Fig. 33C2), GC primed (Fig. 33C3) and PRL treated HC 11 cells (Fig. 33C4). This clearly shows the preferential enhanced interaction between cLAD-cLAD regions of chr8-T with chr-4T only but not with chr-4NT, they also were seen much higher under GC primed condition, reinforcing the idea that clADs and ciLAD regions cluster more preferentially within the cis chromosome in GC primed and PRL states respectively. Further to substantiate observed changes in c/ciLAD interaction among translocated chromosomes, we took advantage of another naturally existing chr1:6 translocations in HC11 cells (Fig. 10E, Fig. 11A, B and C; Table15B) and analyzed interactions of c/ciLAD in translocated and non-translocated portion of chr1 vs c/ciLADs of entire chr6 (Fig. 34A1-2; B1-3). We found similar trends in c/ciLAD interaction among translocated and non-translocated portion of chr1 vs entire chr6 (Fig. 34C1-4) as that of chr4:8 translocations. These results unequivocally demonstrate that GC priming leads to preferential clustering of cLAD regions intra-chromosomally and PRL treatment leads to preferential clustering of ciLAD regions. Reciprocal behavior of cLADs under GC and ciLADs under PRL signalling led us to investigate consequences of such interaction on the status of transcription of genes in those domains.

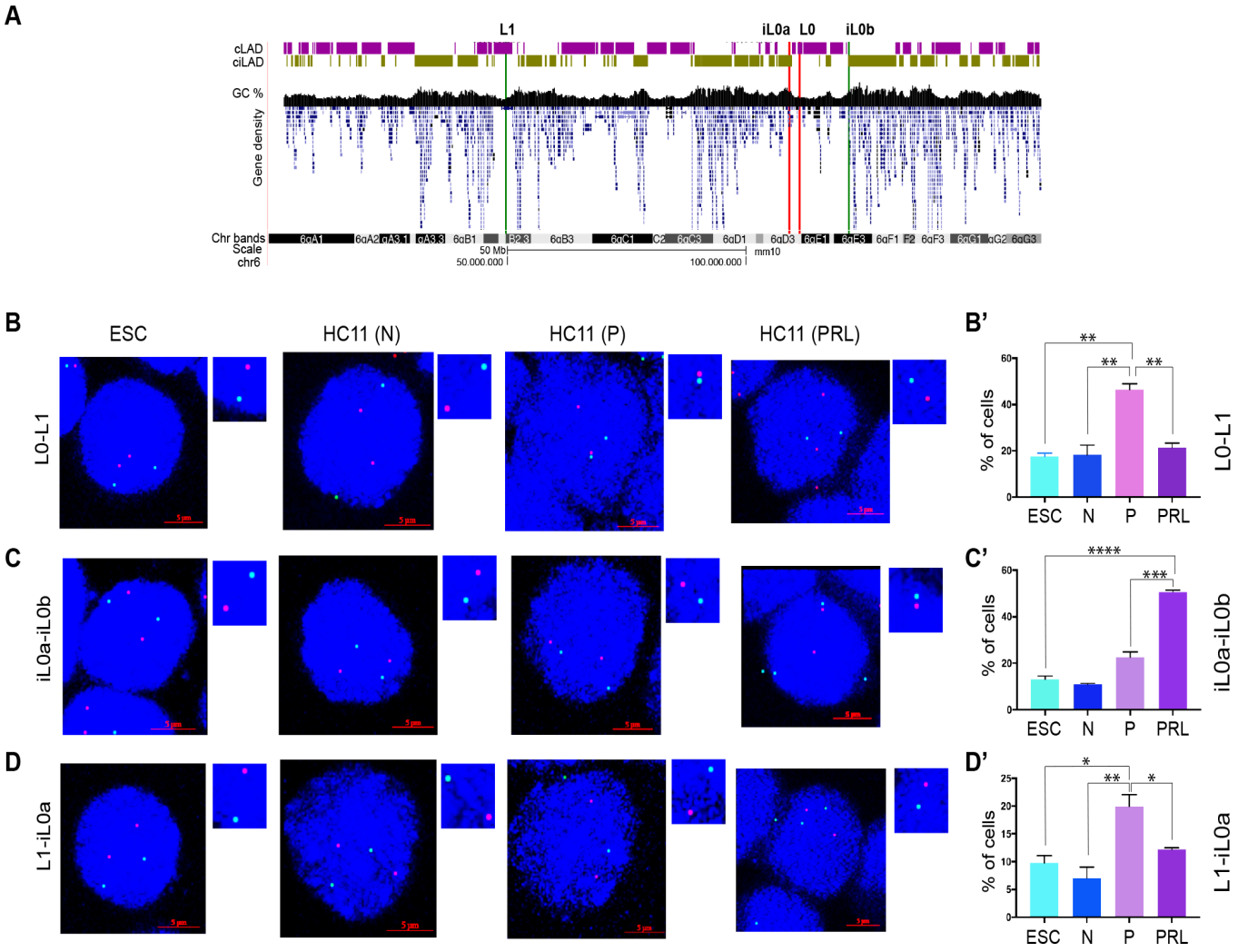

A
L1
iL0a
L0
iL0b
cLAD
ciLAD
GC %
Gene density
Chr bands
Scale
chr6
50 Mb
50.000.000
100.000.000
mm10
B
ESC
HC11 (N)
HC11 (P)
HC11 (PRL)
L0-L1
B'
% of cells
ESC
N
P
PRL
C
iL0a-iL0b
C'
D
L1-iL0a
D'

Fig.32. FISH analysis of long-range chromatin interaction between and among c/ciLADs of Chr6 in ESC, N, P and PRL treated HC11 cells: **(A)** Schematic view of UCSC genome browser Chr6 specific tracks of c/ciLADs, GC%, gene density, chromosome band with scale upon which representative FISH probes from cLAD (L1 (green lane) and L0 (Red lane) and ciLAD (iL0a (red lane) and iL0b (green lane) are shown. **(B)** 3D DNA FISH images of nuclei from ESC, N, P and PRL treated HC11 cells using the probes L0 (Red) and L1 (Green) from cLAD regions. Inset shows the most representative magnified image of a portion of image from each cell-type. Co-association of signals are seen only in HC11 (P) condition. **(B')** Bar plot showing the percentage of cells showing representative co-association signals. Note significant increase in number of cells showing co-associations of distant cLADs only in HC11 (P) condition. **(C)** 3D DNA FISH images of nuclei from ESC, N, P and PRL treated HC11 cells using probes iL0a (Red) and iL0b (Green) from ciLAD regions. Inset shows the most representative magnified image of a portion of image from each cell-type. Co-association of signals are seen only under HC11 (PRL) condition. **(C')** Bar plot showing the percentage of cells showing representative co-association signals. Note significant increase in number of cells showing co-associations of distant ciLAD only in HC11 (PRL) condition. (**D)** 3D DNA FISH images of nuclei from ESC, N, P and PRL treated HC11 cells using probes L1 (Green) from cLAD and iL0a (Red) from ciLAD regions. Inset shows the most representative magnified image of a portion of image from each cell-type. Co-association of signals for this pair of probes are seen only under HC11 (P) condition and to some extent in PRL condition. **(D')** Bar plot showing the percentage of cells showing representative co-association signals. Note significant increase in number of cells showing co-associations of distant c/ciLAD only in HC11 (P) condition and to some extent in PRL condition. Image scale is 5μM.

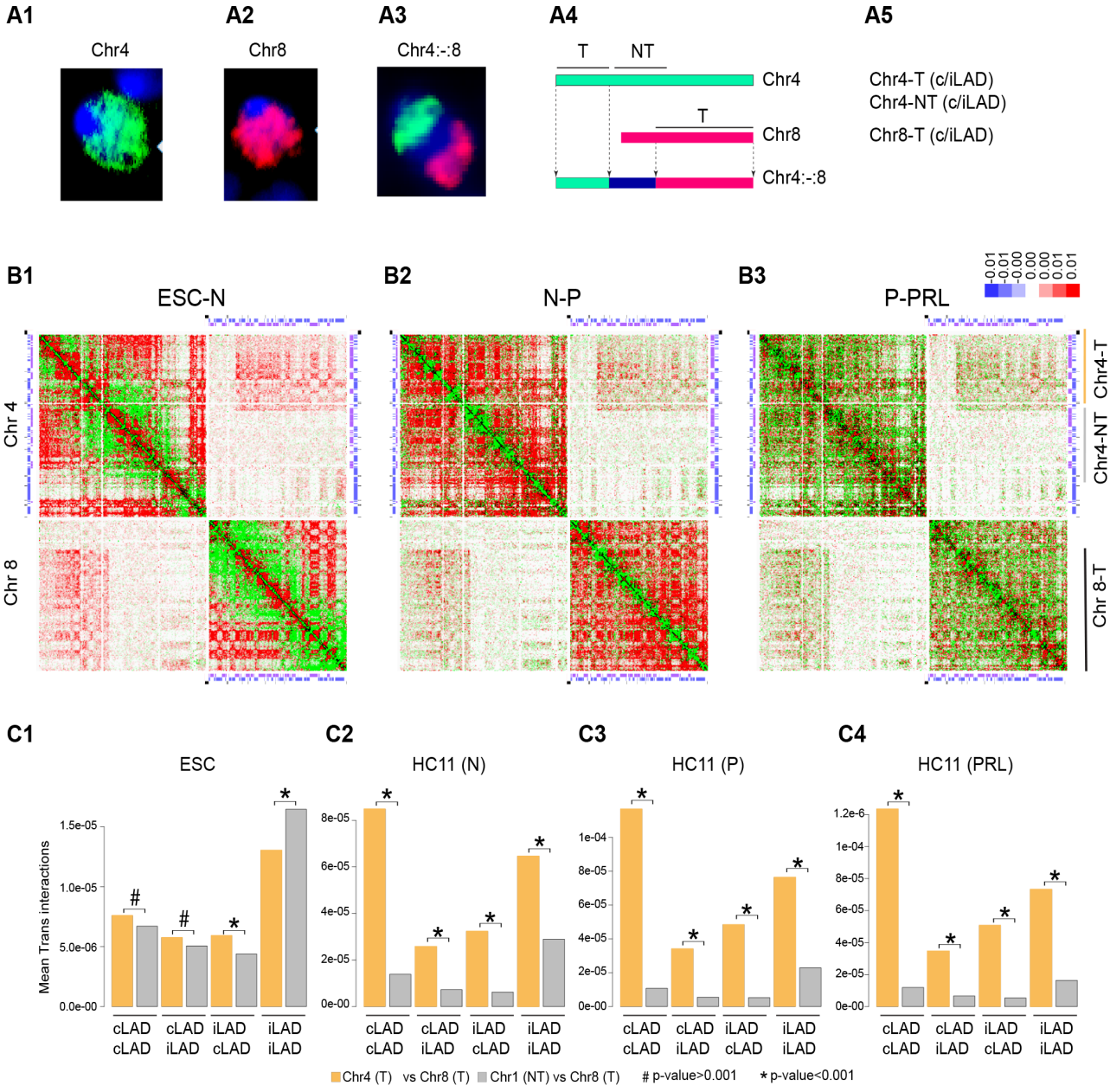

A1
Chr4
A2
Chr8
A3
Chr4:-:8
A4
T
NT
Chr4
T
Chr8
Chr4:-:8
A5
Chr4-T (c/iLAD)
Chr4-NT (c/iLAD)
Chr8-T (c/iLAD)
B1
ESC-N
B2
N-P
B3
P-PRL
-0.01
-0.01
-0.00
0.00
0.00
0.01
0.01
Chr 4
Chr 8
Chr4-T
Chr4-NT
Chr 8-T
C1
ESC
C2
HC11 (N)
C3
HC11 (P)
C4
HC11 (PRL)
Mean Trans interactions
cLAD cLAD
cLAD iLAD
iLAD cLAD
iLAD iLAD
Chr4 (T) vs Chr8 (T)
Chr1 (NT) vs Chr8 (T)
p-value>0.001
* p-value<0.001

Fig.33. Evaluation of chromatin interactions encompassing c/ciLAD regions between translocated chromosome 4 and 8: DAPI stained metaphase chromosome with Chr4 specific paint (Green) and Chr8 specific paint (Red) showing normal chr4 **(A1),** Chr8 **(A2)** and Chr4:8 translocated chromosome **(A3). (A4)** Schematic representation of parts of Chr4 and Chr8 that are translocated (T) and non-translocated (NT). **(A5)** Schematic representation of regions of translocated (T) and non-translocated region of Chr4 specific c/ciLADs interactions with translocated portion of Chr8. **(B1)** Heatmaps representing subtraction matrices of Chr4 and Chr8 intra and inter- chromosomal interaction matrix between ESC & N, **(B2)** N & P and **(B3)** P & PRL states, showing relatively enriched interactions at inter-chromosomal area of translocated region. Heatmaps are overlaid with UCSC track of Chr4 specific cLAD (Pink) and ciLADs (Purple). Note preferentially enriched interaction of cLADs in translocated region (T) in N-P states compared to P-PRL state of HC11 cells. Scale bar represent frequency of interactions among 250kb bins in arbitrary units. **(C1)** Bar plot showing interaction of c/ciLADs of translocated region (T) vs non-translocated (NT) regions of Chr4 vs c/ciLADs of a portion of translocated Chr8 (T) in ESCs, normal **(C2)**, GC primed **(C3)** and PRL treated HC11 cells **(C4)**, showing significant rise in c/ciLAD regions of Chr4-T with Chr8 (T), compared to CHr4-NT, except in ESCs where cLAD Chr4 (T) and Chr4 (NT) vs c/ciLADs of Chr8(T) were found to be insignificant. Student's t test was used to calculate the statistical significance.

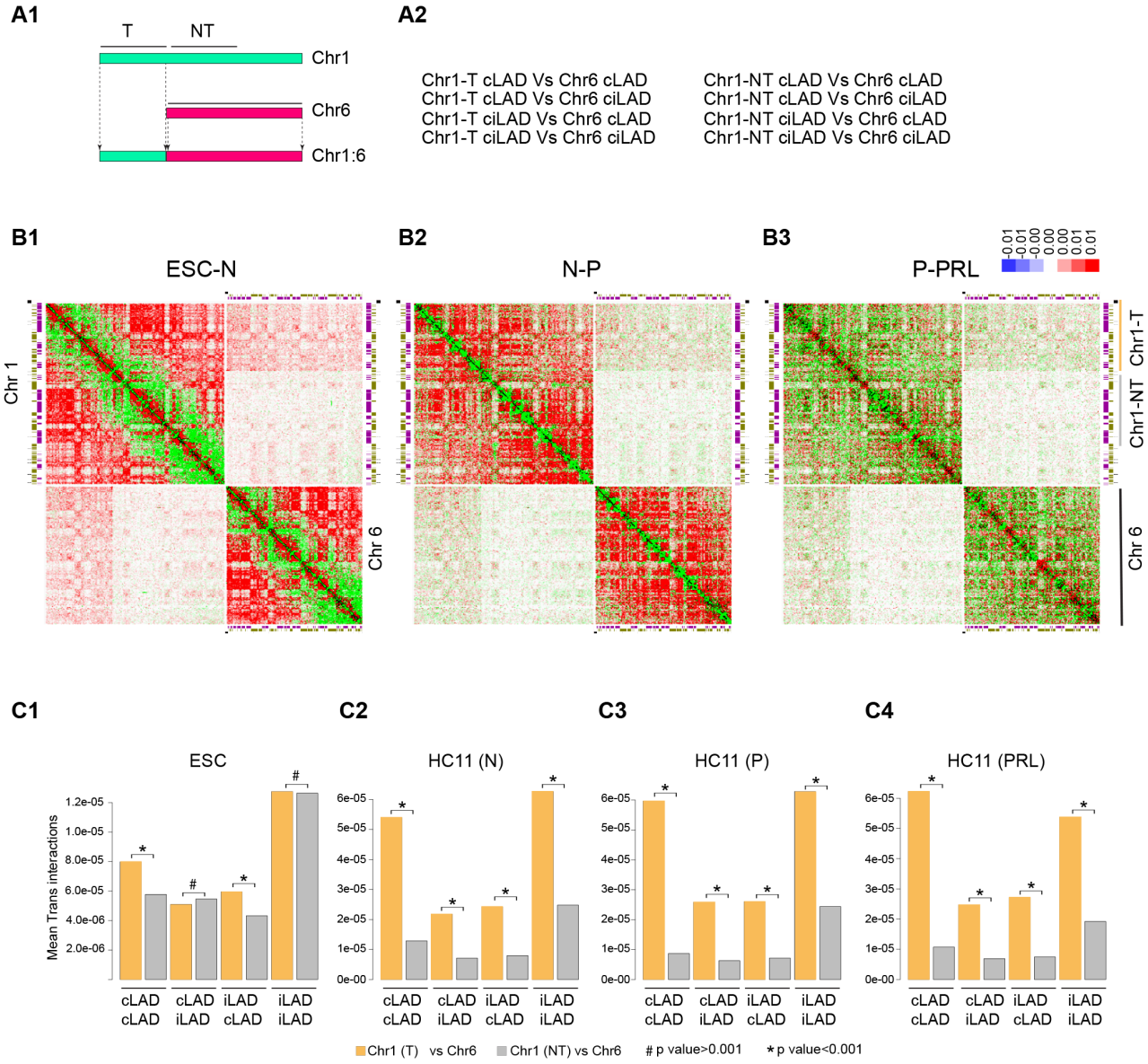

A1
T
NT
Chr1
Chr6
Chr1:6
A2
Chr1-T cLAD Vs Chr6 cLAD
Chr1-T cLAD Vs Chr6 ciLAD
Chr1-T ciLAD Vs Chr6 cLAD
Chr1-T ciLAD Vs Chr6 ciLAD
Chr1-NT cLAD Vs Chr6 cLAD
Chr1-NT cLAD Vs Chr6 ciLAD
Chr1-NT ciLAD Vs Chr6 cLAD
Chr1-NT ciLAD Vs Chr6 ciLAD
B1
ESC-N
B2
N-P
B3
P-PRL
-0.01
-0.01
-0.00
0.00
0.00
0.01
0.01
Chr 1
Chr 6
Chr1-T
Chr1-NT
Chr 6
C1
ESC
C2
HC11 (N)
C3
HC11 (P)
C4
HC11 (PRL)
Mean Trans interactions
1.2e-05
1.0e-05
8.0e-06
6.0e-06
4.0e-06
2.0e-06
6e-05
5e-05
4e-05
3e-05
2e-05
1e-05
0e-00
cLAD cLAD
cLAD iLAD
iLAD cLAD
iLAD iLAD
Chr1 (T) vs Chr6
Chr1 (NT) vs Chr6
p value>0.001
* p value<0.001

Fig.34. Evaluation of chromatin interactions encompassing c/ciLAD regions between translocated chromosome 1 and 6: **(A1)** Schematic representation of parts of Chr1 and Chr6 that are translocated (T) and non-translocated (NT). **(A2)** Schematic representation of regions of translocated (T) and non-translocated region of Chr1 specific c/ciLADs interactions with entire Chr6 that were considered for the analysis. **(B1)** Heatmaps representing subtraction matrices of Chr1 and Chr6 intra and inter- chromosomal interaction matrix between ESC and N, **(B2)** N and P and **(B3)** P and PRL states, showing relatively enriched interactions at inter-chromosomal area of translocated region. Heat maps are overlaid with respective chromosome specific cLADs (Pink) and ciLADs (olive green). Note preferential enriched interaction of cLADs in translocated region (T) in N-P states compared to P-PRL state. **(C1).** Bar plot showing interaction of c/ciLADs of translocated (T) and non-translocated (NT) regions of Chr1 vs c/ciLADs of entire Chr6 showing significant rise in c/ciLAD interactions between translocated portion of Chr1-T with Chr6 compared to Chr1-NT in all HC11 cell types when compared with ESC. Student's t test was used to calculate the statistical significance.

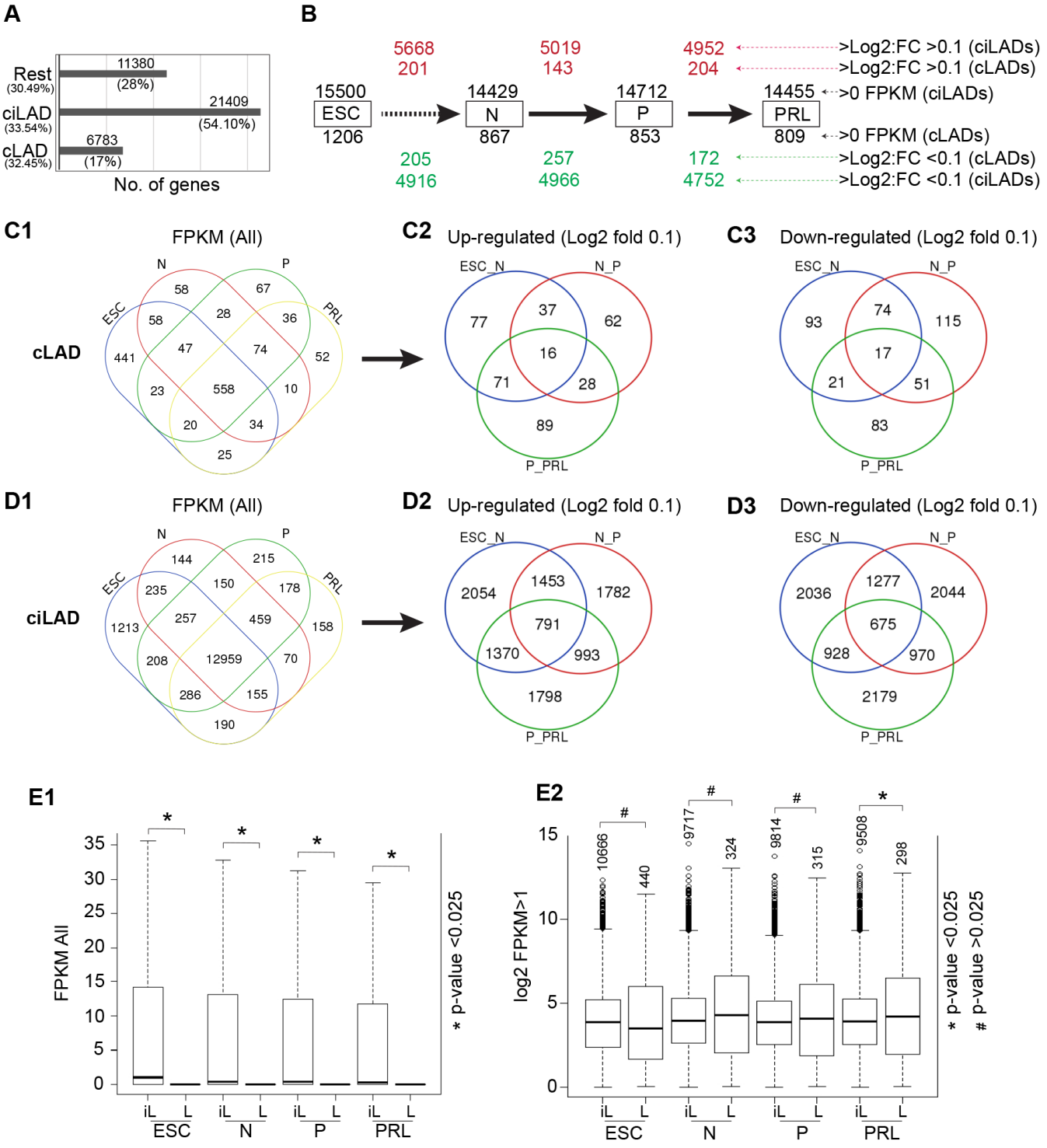

A
Rest
(30.49%)
11380
(28%)
ciLAD
(33.54%)
21409
(54.10%)
cLAD
(32.45%)
6783
(17%)
No. of genes
B
15500
ESC
1206
14429
N
867
14712
P
853
14455
PRL
809
5668
201
5019
143
4952
204
>Log2:FC >0.1 (ciLADs)
>Log2:FC >0.1 (cLADs)
>0 FPKM (ciLADs)
>0 FPKM (cLADs)
205
4916
257
4966
172
4752
>Log2:FC <0.1 (cLADs)
>Log2:FC <0.1 (ciLADs)
C1
FPKM (All)
cLAD
C2
Up-regulated (Log2 fold 0.1)
C3
Down-regulated (Log2 fold 0.1)
D1
FPKM (All)
ciLAD
D2
Up-regulated (Log2 fold 0.1)
D3
Down-regulated (Log2 fold 0.1)
E1
FPKM All
* p-value <0.025
E2
log2 FPKM>1
* p-value <0.025
# p-value >0.025
iL
L
ESC
N
P
PRL

Fig.35. Overview of gene expression differences between cLADs and ciLADs in ESC, N, P and PRL treated HC11 cells: **(A)** Total number and percentage of genes physically located within the boundaries of cLADs, ciLADs and rest of the genome. **(B)** Schematic representation of total number of transcribed mRNA genes from cLADs (below the square box) and ciLADs (above the square box) in all the four cell-types studied, along with the number of differentially expressed genes between ESC-N, N-P and P-PRL treated HC11 cells (red coloured numbers above the arrowhead with Log2FC >0.1 of ciLADs or cLADs). **(C1)** Venn diagram showing number of genes in cLADs and ciLADs **(D1)** of ESC, N, P and PRL treated HC11 cells showing conserved, overlapping and uniquely transcribed genes along with number of up-regulated genes **(C2)** and down-regulated genes **(C3)** between ESC-N, N-P and P-PRL states in cLADs and in cILADs respectively **(D2&D3). (E1)** Boxplot representing the distribution of gene expression values (FPKM) in ciLADs (iL) and cLADs (L) in ESC, N, P and PRL states (outliers excluded). **(E2)** Boxplot showing distribution of gene expression values of FPKM only >1 in ciLADs and cLADs in ESC, N, P and PRL states. Wilcoxon rank sum test was used for statistical testing.

16. **Reorganization of cLAD and ciLADs has inverse effect on transcriptional status under GC and PRL treated HC11 cells:**

Mouse genome is constituted by nearly 32.45% cLADs, 33.5% ciLADs and rest (facultative regions: f-reg) 30% of chromatin, but interestingly, these regions harbour17%, 54.1% and 28% of total 39,572 genes respectively (Fig. 35A). Pathway analysis of genes that are physically present in cLADs showed genes that are involved in Olfactory transduction, Linoleic acid metabolism, Graft-versus-host disease, Type I diabetes mellitus, Allograft rejection, Steroid hormone biosynthesis, Retinol metabolism, Arachidonic acid metabolism, Chemical carcinogenesis, Taste transduction, Drug metabolism - other enzymes, Ascorbate and Aldarate metabolism, Metabolism of xenobiotics by cytochrome P450, Pentose and glucuronate interconversions, Porphyrin and chlorophyll metabolism, Drug metabolism - cytochrome P450, PPAR signalling pathway, Fatty acid degradation, Starch and sucrose metabolism, Ovarian steroidogenesis, Glucagon signaling pathway, Systemic lupus erythematosus, Alcoholism etc. When we wanted to see if there are any significantly differentially expressed genes between the cell-stages in cLADs, we observed that most of the genes, but not all, in the cLADs were very lowly expressed and didn't show much differential expression when we considered >1 or <-1 log2fold change for differential gene expression, when compared to iLADs (data note shown). However, when we considered >0.1 (up-regulation) or < -0.1 log2fold change (down-regulation) there are some genes which are differentially expressed and were present in cLADs (Fig. 35B, C1, C2, C3) when compared with ciLADs (Fig. 35D1, D2, D3). The general gene expression distribution in ciLADs vs cLADs showed, as expected, higher gene expression in ciLADs (Fig. 35E1). But surprisingly, when we considered only genes which have at least FPKM value of 1, we observed that the distribution of the expression of genes in cLADs are not so different from ciLADs (Fig. 35E2) which suggests that although most of the genes present in the cLADs have very low expression, cLADs do harbour some genes which are expressed as high as the genes in ciLAD

17. **Preferential reorganization of c/iLAD interactions in succession are recapitulated at individual gene loci during lactogenic differentiation of HC11 cells:**

In order to see the observed changes in long-range contacts over c/ciLAD regions and gene expression at sub-chromosomal gene specific domains (100kb), we extracted all the pair of interactions of a *Csn2* (highly transcribed genes in PRL treated HC11 cells, resides in cLAD region) gene locus 100kb bin vs rest of the bins in chr5 and integrated with c/ciLAD and gene expression and visualized by using Circos

plots of ESC, N, P and PRL treated HC11 cells (Fig. 36A1-4). These maps clearly recapitulate global trends of (i) lack of long-range interactions in ESC, (ii) presence of long-range contacts both in c/ciLADs in HC11 (N) cells, (iii) preferential enhanced interactions of c/iLAD regions in GC primed and PRL treated HC11 cells. Further, to assess the interactions more locally, we generated 10kb resolution Simplenorm normalized contact matrix over 1MB domain encompassing *Csn2* gene locus among ESC, N, P and PRL treated cells (Fig. 36B1-4). All pair wise interactions encompassing 10kb bins within 1MB domain were extracted and visualized as an Arc plot over the 1MB domain heat map and further integrated with PC1, gene expression. This integrative visualization clearly showed (i) presence of more local short-range interactions in ESC (Fig. 36B1), (ii) Gradual decline in local short-range interactions through HC11 (N), P and PRL stages (Fig. 36B2-4) and (iii) that under transcriptionally active condition, (PRL state) Csn2 locus seems to make very few but specific contacts within domain. Further, to observe differences among cell-types, we generated subtraction matrices between ESC-N, N-P and P-PRL states of HC11 cells (Fig. 36C1-3) and integrated with subtraction PCA & differential gene expression maps. This integrative visualization clearly showed Csn2 1MB domain progressive shift from B to A compartment and is accompanied by loss of short-range interactions under GC primed condition (N-P) and gain of local but specific interactions under PRL condition (P-PRL). Similarly, we also analysed *Wap* gene locus (PRL stage specific highly expressed gene locus which falls within ciLAD region) specific 100kb domain in mouse chr11 through Circos plots in ESC, N, P and PRL states (Fig. 37A1-4) which showed rhythmic mode of interactions through ESC (few and short range interactions), HC11 (N) (few but more long-range interactions, HC11 (P) (more long-range contacts and HC11 (PRL) cells (less but specific long-range contacts encompassing both c/ciLADs). Similar trends were observed in 10kb resolution *Wap* locus specific 1MB domain in ESC, N, P and PRL treated HC11 cells (Fig. 37B1-4) and were further evident in subtraction matrices between ESC-N, N-P and P-PRL treated HC11 cells (Fig. 37C1-3). Similar analysis was performed along with other highly expressed cell-type specific genes such as *Etv4* (Fig. 38), *Krt42* (Fig. 39) (highly expressed highly expressed in normal HC11 cells, *Krt23* (Fig. 40) and *Wfdc18* (Fig. 41) (highly expressed in GC primed cells) and found to be in accordance with the global trends. The coordinates of all the above 1Mb gene loci are mentioned in Table 19. These integrative visualizations at 10kb resolution clearly demonstrates that gain and loss of interactions at very micro scale (1MB domain) are reciprocally promoted as loss or gain respectively of interactions at much larger distances (cis or trans chromosomes) apart from the status of a stage-specific gene expression.

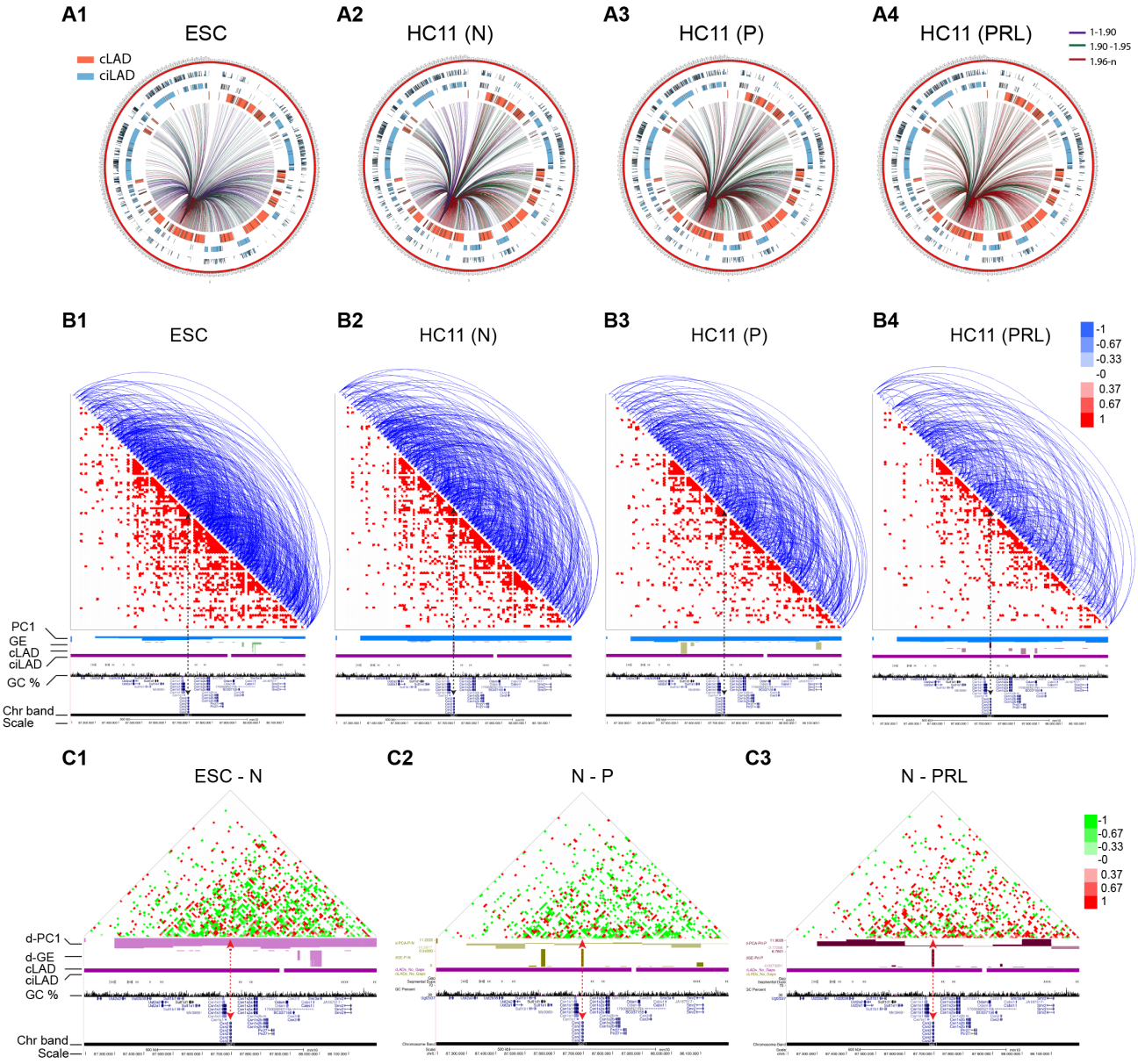

A1
ESC
cLAD
ciLAD
A2
HC11 (N)
A3
HC11 (P)
A4
HC11 (PRL)
1-1.90
1.90 -1.95
1.96-n
B1
ESC
B2
HC11 (N)
B3
HC11 (P)
B4
HC11 (PRL)
-1
-0.67
-0.33
-0
0.37
0.67
1
PC1
GE
cLAD
ciLAD
GC %
Chr band
Scale
C1
ESC - N
C2
N - P
C3
N - PRL
d-PC1
d-GE
cLAD
ciLAD
GC %
Chr band
Scale

Fig.36. Dynamics reorganization of *Csn-2* locus-specific intra-chromosomal networks in ESCs, N, P and PRL treated HC11 cells: **(A1)** Circos plots representing 100kb bin encompassing Csn2 gene locus specific chromatin interaction across mouse Chr5, among ESC, **(A2)** N , (**A3)** P and **(A4)** PRL treated HC11 cells, showing preferential enriched interactions of cLADs and ciLADs under P and PRL states. Intensity of interaction values are depicted with different colours; 1-1.90 (Pink), 1.90-1.95 (Green), 1.96-N (any number) (Red). 1MB domain of *Csn-2* gene specific heatmap (10kb resolution) overlaid with Arc plot showing all pairs of interactions within 1MB domain, which is further integrated with PCA, gene expression (GE), c/ciLADs, gene density, GC/AT % and chromosomal band in ESCs **(B1),** normal HC11 cells **(B2),** GC primed HC11 cells **(B3)** and PRL treated HC11 cells **(B4)** showing differences in long-range contacts. To visualize reorganization differences, subtraction matrices were generated between ESC-N HC11 cells **(C1),** N-P **(C2)** and between P-PRL treated HC11 cells **(C3).** Scale bar represents frequency of interaction between two loci in arbitrary units. d-PCA: subtraction PCA between samples; d-GE: Differential gene expression between samples.

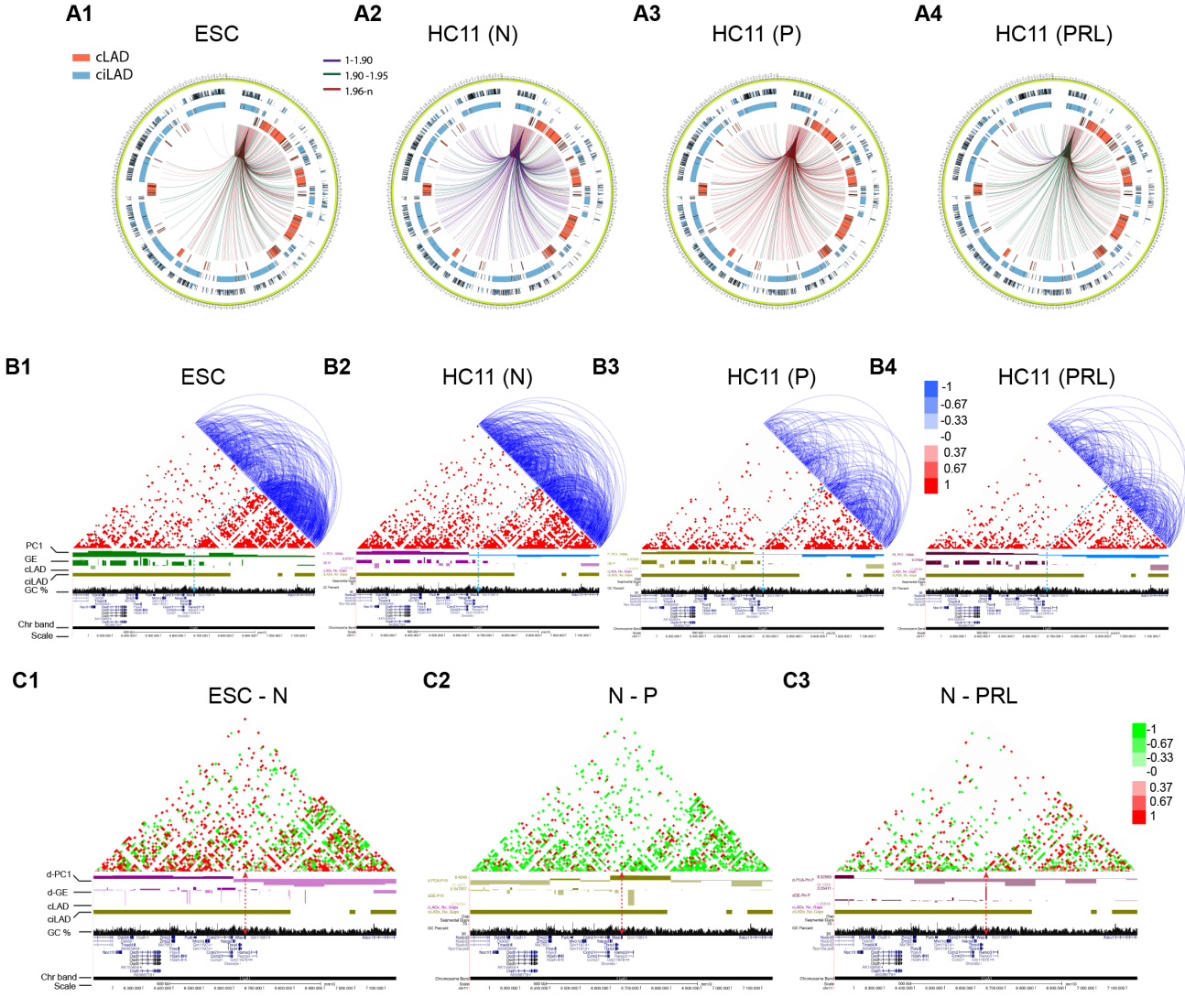

A1
ESC
cLAD
ciLAD
A2
HC11 (N)
1-1.90
1.90 -1.95
1.96-n
A3
HC11 (P)
A4
HC11 (PRL)
B1
ESC
B2
HC11 (N)
B3
HC11 (P)
B4
HC11 (PRL)
-1
-0.67
-0.33
-0
0.37
0.67
1
PC1
GE
cLAD
ciLAD
GC %
Chr band
Scale
C1
ESC - N
C2
N - P
C3
N - PRL
d-PC1
d-GE
cLAD
ciLAD
GC %
Chr band
Scale
-1
-0.67
-0.33
-0
0.37
0.67
1

Fig.37. Dynamics of chromatin interactions within 1MB domain of *Wap* gene locus and its integration with gene expression and c/ciLADs: **(A1)** Circos plots of 10kb bin representing mouse *Wap* gene specific intra-chromosomal (Chr11) interaction among ESC, N, P and PRL treated HC11 cells showing preferential enriched interactions of cLADs and ciLADs under P and PRL states respectively. 1MB domain of *Wap* gene specific heatmap overlaid with 'Arc' plot showing all pairs of interactions within 1MB domain, which is further integrated with PCA1, gene expression (GE), c/ciLADs, gene density, GC/AT % and chromosomal G-band in ESCs **(B1),** normal HC11 cells **(B2),** GC primed HC11 cells **(B3)** and PRL treated HC11 cells **(B4)** showing differences in long-range contacts. To visualize reorganization differences, subtraction matrices were made between ESC-N HC11 cells **(C1),** N-P **(C2)** and between P-PRL treated HC11 cells **(C3).** Location of *Wap* gene is highlighted with dotted lanes across panel. Scale bar represents frequency of interaction between two loci in arbitrary units. d-PCA: subtraction PCA between samples; d-GE: Differential gene expression between samples.

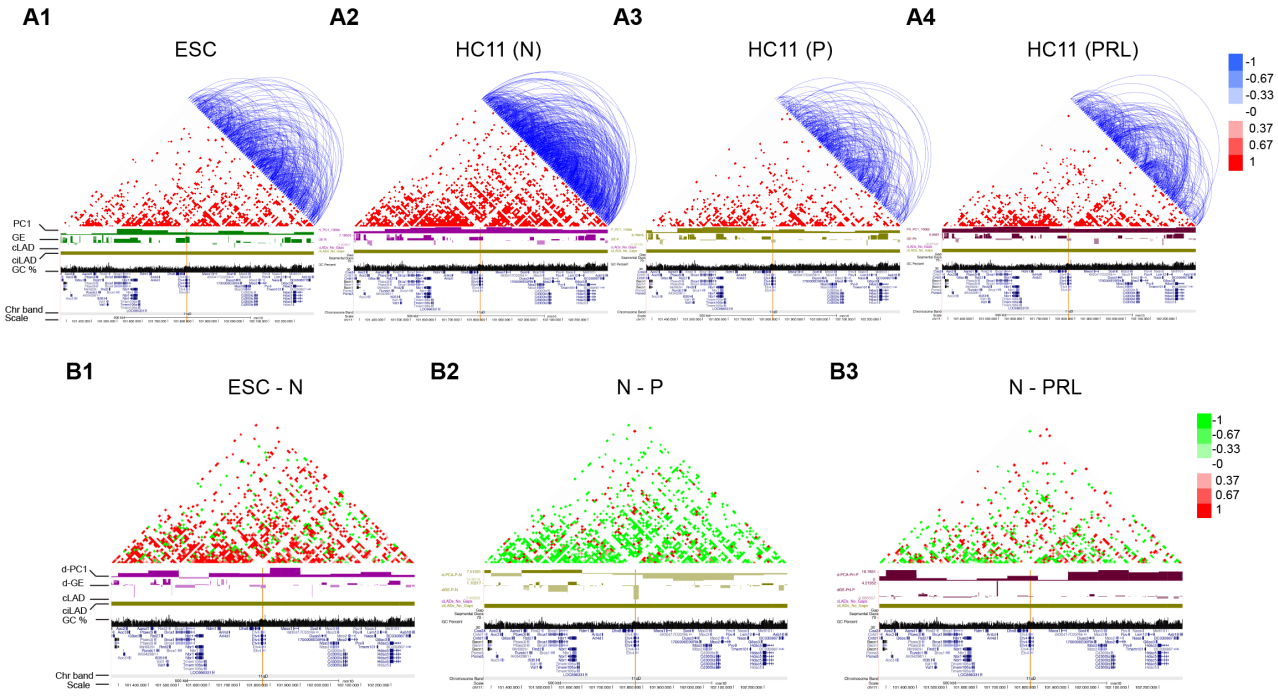


Fig.38. Dynamics of chromatin interactions within 1MB domain of *Etv4* gene locus and its integration with gene expression and c/ciLAD regions: 1MB domain of *Etv4* gene specific heat map (10kb resolution) overlaid with Arc plot showing all pair of interaction within 1MB domain, is further integrated with levels of PCA1 component, gene expression (GE), c/ciLADs, gene density, GC/AT % and chromosomal G-bands in ESCs **(A1),** normal (N) HC11 cells **(A2),** GC primed (P) HC11 cells **(A3)** and PRL treated HC11 cells **(A4)** showing differences in long-range contacts within 1MB domain. To visualize differences in reorganization, subtraction matrices were made between ESC-N HC11 cells **(B1),** N-P **(B2)** and between P-PRL treated HC11 cells **(B3).** Location of *Etv4* gene is highlighted with solid orange coloured lane in each heatmap. Scale bar represent frequency of interaction between two loci in arbitrary units. d-PCA: subtraction PCA between samples; d-GE: Differential gene expression between samples.

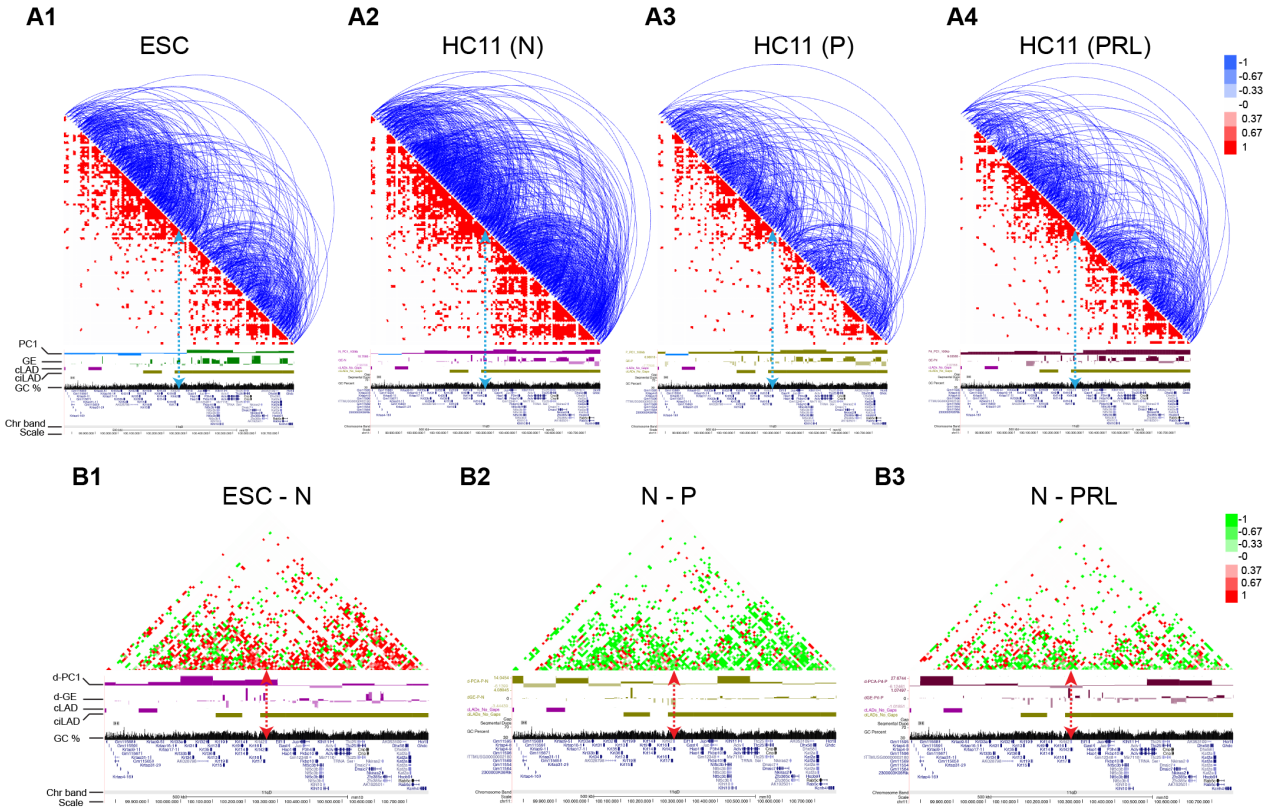


Fig.39. Dynamics of chromatin interactions within 1MB domain of *Krt42* gene locus and its integration with gene expression and c/ciLAD regions: 1MB domain of *Krt42* gene specific heatmap (10kb resolution) overlaid with Arc plot showing all pair of interaction within 1MB domain is further integrated with levels of PCA, gene expression (GE), c/ciLADs, gene density, GC/AT % and chromosomal band in ESCs **(A1),** normal HC11 cells **(A2),** GC primed HC11 cells **(A3)** and PRL treated HC11 cells **(A4)** showing differences in long-range contacts within 1MB domain. To visualize reorganization differences subtraction matrices were made between ESC-N HC11 cells **(B1),** N-P **(B2)** and between P-PRL treated HC11 cells **(B3).** Location of *Krt42* gene is highlighted with dotted lanes in each heatmap**.** Scale bar represent frequency of interaction between two loci in arbitrary units. d-PCA: subtraction PCA between samples; d-GE: Differential gene expression between samples.

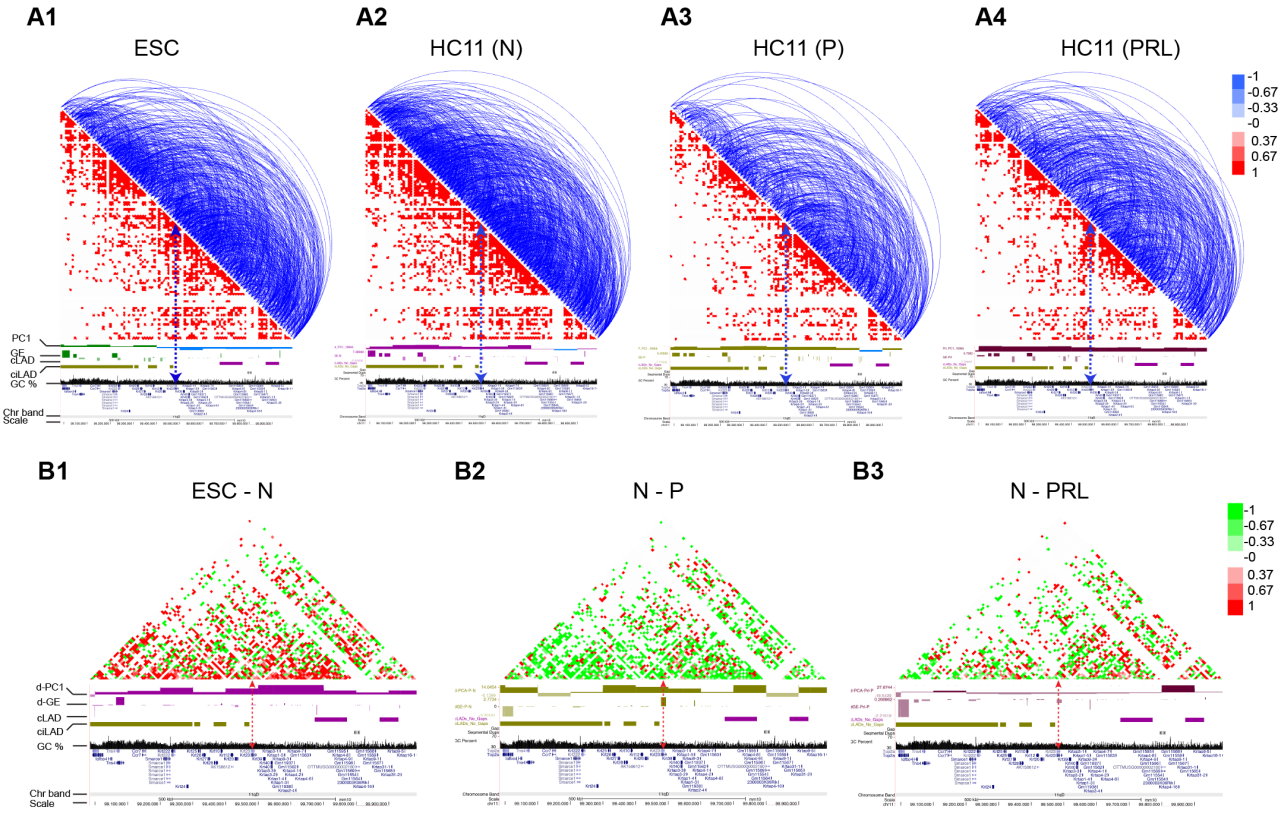


Fig.40. Dynamics of chromatin interactions within 1MB domain of *Krt23* gene locus and its integration with gene expression and c/ciLAD regions: 1MB domain of *Krt23* gene specific heatmap (10kb resolution) overlaid with Arc plot showing all pair of interaction within 1MB domain is further integrated with levels of PCA, gene expression (GE), c/ciLADs, gene density, GC/AT % and chromosomal band in ESCs **(A1),** normal HC11 cells **(A2),** GC primed HC11 cells **(A3)** and PRL treated HC11 cells **(A4)** showing differences in long-range contacts within 1MB domain. To visualize reorganization differences subtraction matrices were made between ESC-N HC11 cells **(B1),** N-P **(B2)** and between P-PRL treated HC11 cells **(B3).** Location of *Krt23* gene is highlighted with dotted lanes in each heatmap. Scale bar represent frequency of interaction between two loci in arbitrary units. d-PCA: subtraction PCA between samples; d-GE: Differential gene expression between samples.

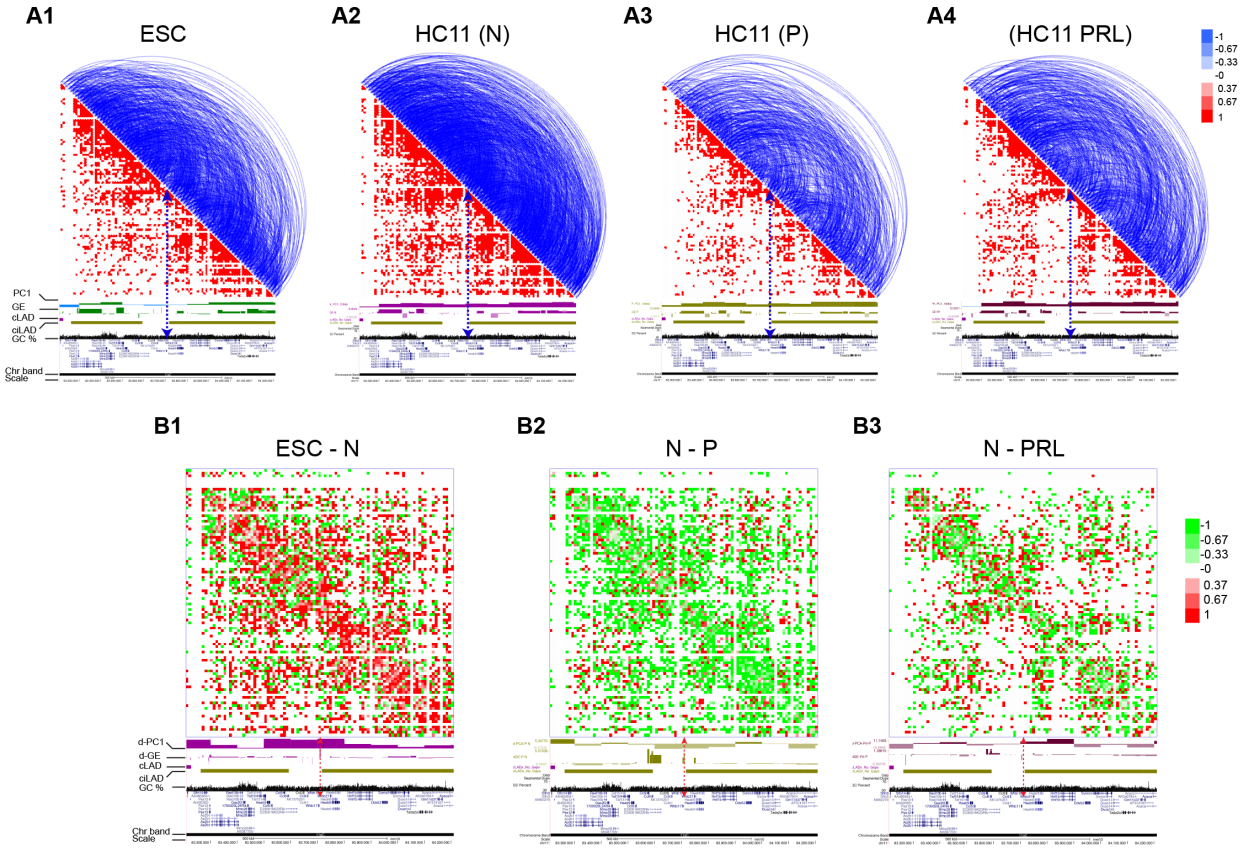


Fig.41. Dynamics of chromatin interactions within 1MB domain of *Wfdc18* gene locus and its integration with gene expression and c/ciLAD regions: 1MB domain of *Wfdc18* gene specific heatmap (10kb resolution) overlaid with Arc plot showing all pair of interaction within 1MB domain is further integrated with levels of PCA, gene expression (GE), c/ciLADs, gene density, GC/AT % and chromosomal band in ESCs **(A1),** normal HC11 cells **(A2),** GC primed HC11 cells **(A3)** and PRL treated HC11 cells **(A4)** showing differences in long-range contacts within 1MB domain. To visualize reorganization differences subtraction matrices were made between ESC-N HC11 cells **(B1),** N-P **(B2)** and between P-PRL treated HC11 cells **(B3).** Location of *Wfdc18* gene is highlighted with dotted lanes in each heat map. Scale bar represent frequency of interaction between two loci in arbitrary units. d-PCA: subtraction PCA between samples; d-GE: Differential gene expression between samples.

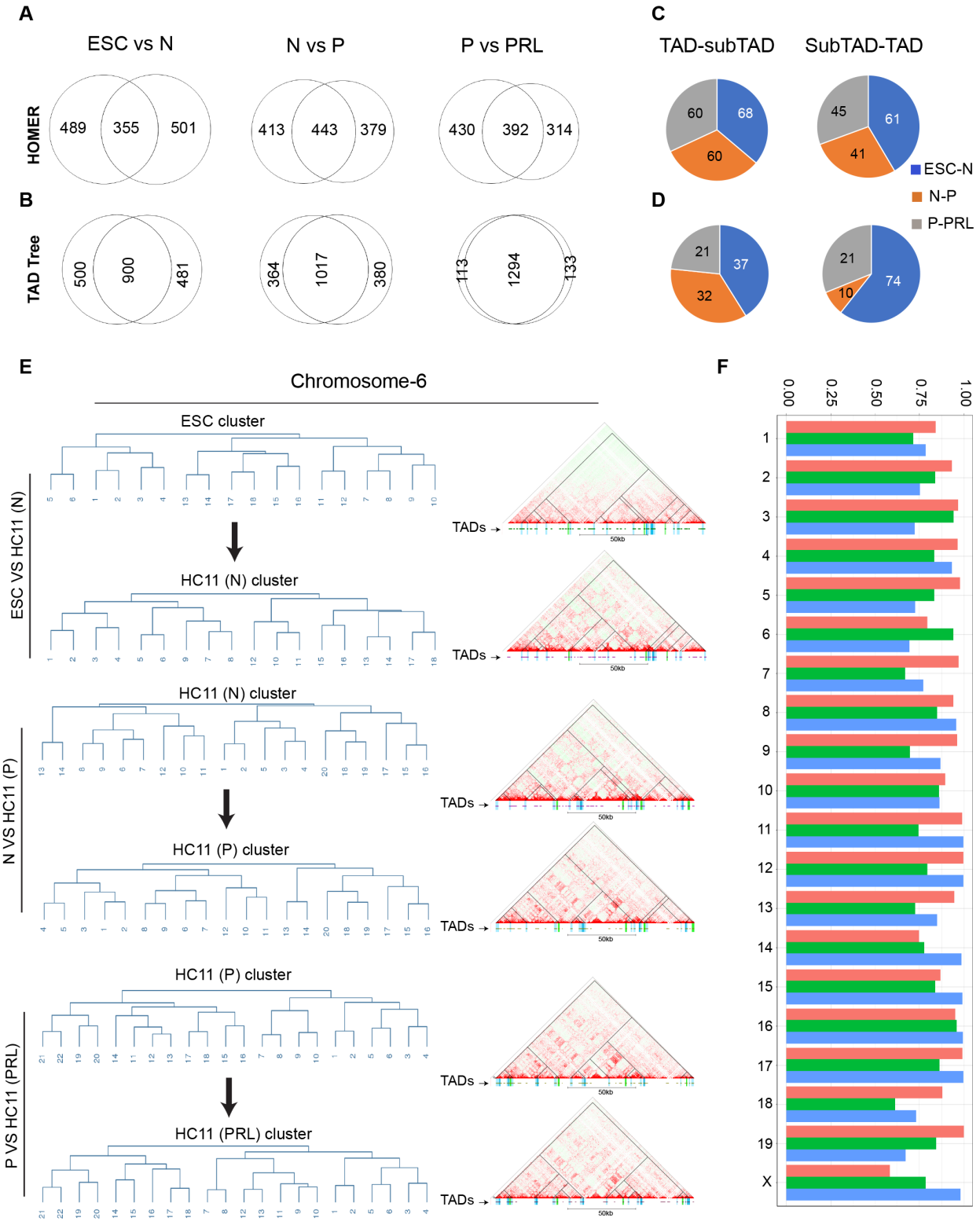

A
ESC vs N
N vs P
P vs PRL
HOMER
489
355
501
413
443
379
430
392
314
B
TAD Tree
500
900
481
364
1017
380
113
1294
133
C
TAD-subTAD
SubTAD-TAD
60
68
60
45
61
41
ESC-N
N-P
P-PRL
D
21
37
32
21
74
10
E
Chromosome-6
ESC cluster
ESC VS HC11 (N)
HC11 (N) cluster
TADs
50kb
HC11 (N) cluster
N VS HC11 (P)
HC11 (P) cluster
HC11 (P) cluster
P VS HC11 (PRL)
HC11 (PRL) cluster
F
0.00
0.25
0.50
0.75
1.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
X
ESC - N correlation
N - P correlation
P - PRL correlation

Fig.42. **Dynamic, hierarchical reorganization of TADs in ESCs and during lactogenic differentiation of HC11 cells. (A)** Venn diagram showing extent of conservation of TADs between ESC-N, N-P and P-PRL states of HC11 cells, derived using HOMER tool showing relatively less conservation when compared with TADs derived from TAD Tree **(B)**, which showed higher conservation. **(C)** Pie chart showing number of transitions from TADs to sub-TAD and sub-TAD to TAD between ESC-N, N-P and P-PRL states of HC11 cells derived using HOMER and **(D)** TADTree in between N, P and PRL treated HC11 cells. **(E)** Dendrogram representing pair-wise analysis of hierarchical organization of conserved TADs between ESC&HC11 (N), N&P, P&PRL states of HC11 cells. TADs are overlaid onto Chr6 heatmap of ESC, N, P and PRL treated HC11 cells showing stage-specific, hierarchical reorganization of conserved TADs. **(F)** Hierarchical clustering of conserved TADs of all the mouse chromosomes (Chr1-19+X), in ESC, N, P and PRL treated HC11 cells, showing clustering correlation analysis between distances and in between cophenetic correlations showing variable patterns of clustering among chromosomes between ESC-N, N-P and P-PRL conditions.

18. **Dynamic reorganization of topologically associated domains (TADs) accompanies GC and PRL signalling:** We derived TADs in ESC and at different stages of lactogenic differentiation of HC11 cells by using two different tools; HOMER which derives TADs based on directional Index concept, whereas TADtree tool generates nested TADs (Weinreb and Raphael, 2016) Upon comparison between ESC-N, N-P and P-PRL states, we observed that majority of TADs show conservation between cell-types, however, there are significant number of TADs did not overlap between samples suggesting an extensive reorganization. But TADs derived from HOMER show slightly higher conservation between N -P than between ESC- N, and P- PRL treated cells (Fig. 42A). Interestingly, TADs generated with TADtree provides higher conservation of TAD boundaries between GC primed and PRL treated stages followed by normal to GC primed, and then ESC to normal HC11 cells. (Fig. 42B). We also observed and derived the set of TADs which are losing the boundaries and become smaller (TAD to sub-TAD) and the opposite (sub-TADs to TADs) from one cell state to another using HOMER (Fig. 42C) and Tadtree (Fig. 42D). We observe that most of the TAD-subTAD interplay between ESC and normal HC11 cells are prominent but to a lesser extent between N-P and P-PRL treated cells which could be attributed to the fact that MECs are much closer in their lineage hierarchy than with ESC, which is also evident from correlation values between cis matrices (Table 18). Overall this study concluded that TADs are highly dynamic and highly reorganized during GC and PRL signalling, which was consistent with a previous study (Cubenas-Potts and Corces, 2015, Boya et al., 2017, Fraser et al., 2015a)
19. **Hierarchical folding /clustering of TADs reflects changes in long-range interactions:** We theorized that, the hierarchical folding /clustering between the TADs, which involves the long-range contacts would reflect the trends observed in cLADs. With a slight modification from a previously described method to see the clustering (Fraser et al., 2015a), We performed hierarchical folding/clustering in samples' pair wise (ES-N, N-P, P-PRL) comparison by taking the corresponding conserved TADs. Shown in the figure (Fig. 42E) is the reorganization of chr6, with the remaining chromosomes in Figs 43-47. In chr6, the correlation between N-P is higher than that of ESC-N and P-PRL (Fig. 42F). But, in most of the chromosomes (except chr2, 3, 5, 14, 16, 19 and X), the correlation between N-P is lower than that of the other two comparisons (Fig. 35A). This suggests that, the changes in hierarchical clustering are more pronounced after GC signalling but not between ESC and normal HC11 cells. Interestingly, the larger chromosomes have a higher correlation between ESC-N and

smaller chromosomes show the opposite. This suggests that, at hierarchical folding level, larger chromosomes are not much reorganized from ESC to N but smaller chromosomes do.

Fig.43. Hierarchical organization of Chr1-4 TADs accompanies ESC and HC11 lactogenic differentiation: Dendrogram representing hierarchical clustering analysis between distances and in between cophenetic correlations showing variable patterns of clustering among chromosomes (1-4) between ESC-N, N-P and P-PRL conditions.

Fig.44. Hierarchical organization of Chr5-8 TADs accompanies ESC and HC11 lactogenic differentiation: Dendrogram representing hierarchical clustering analysis between distances and in between cophenetic correlations showing variable patterns of clustering among chromosomes (5-8) between ESC-N, N-P and P-PRL conditions.

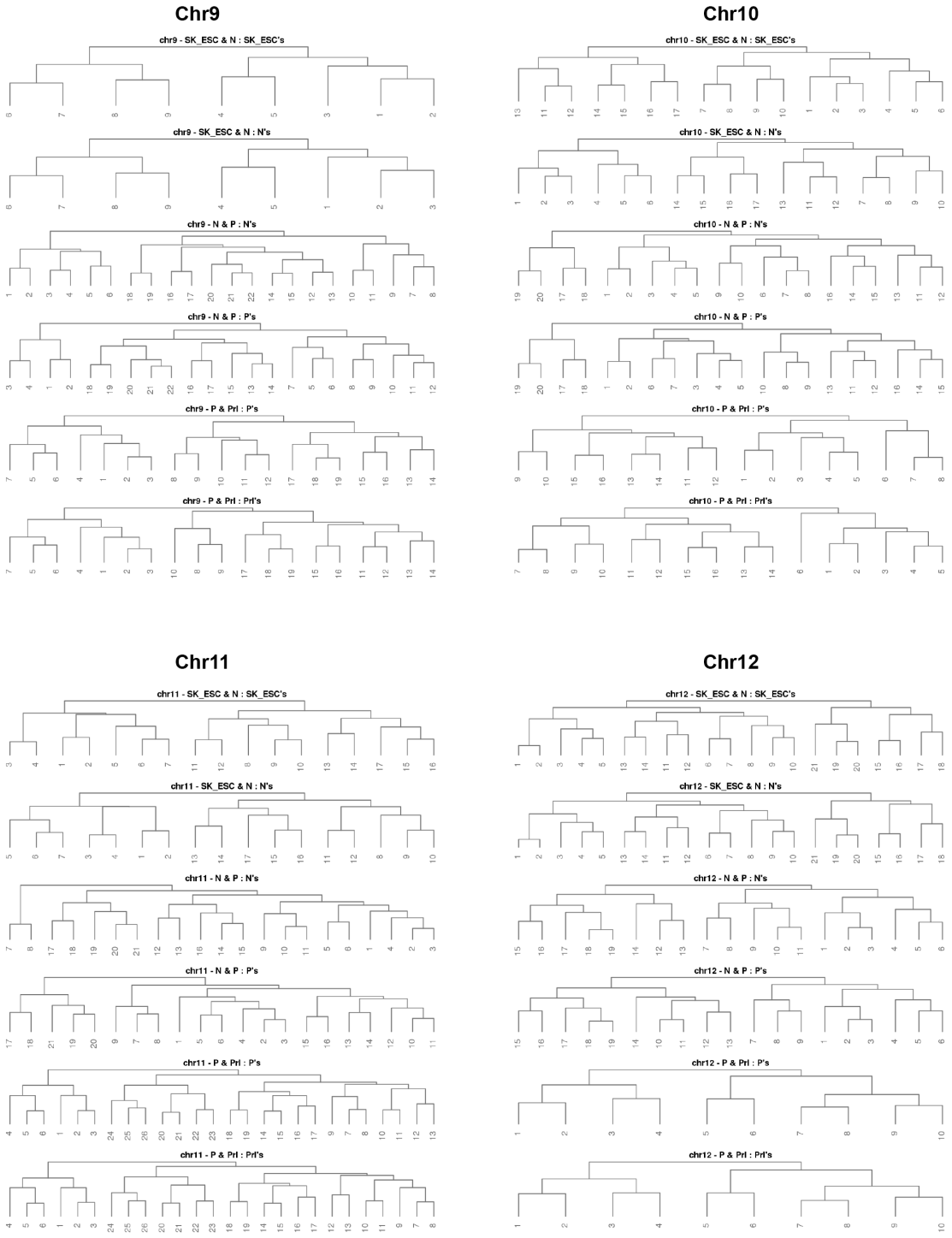


Fig.45. Hierarchical organization of Chr9-12 TADs accompanies ESC and HC11 lactogenic differentiation: Dendrogram representing hierarchical clustering analysis between distances and in between cophenetic correlations showing variable patterns of clustering among chromosomes (9-12) between ESC-N, N-P and P-PRL conditions.

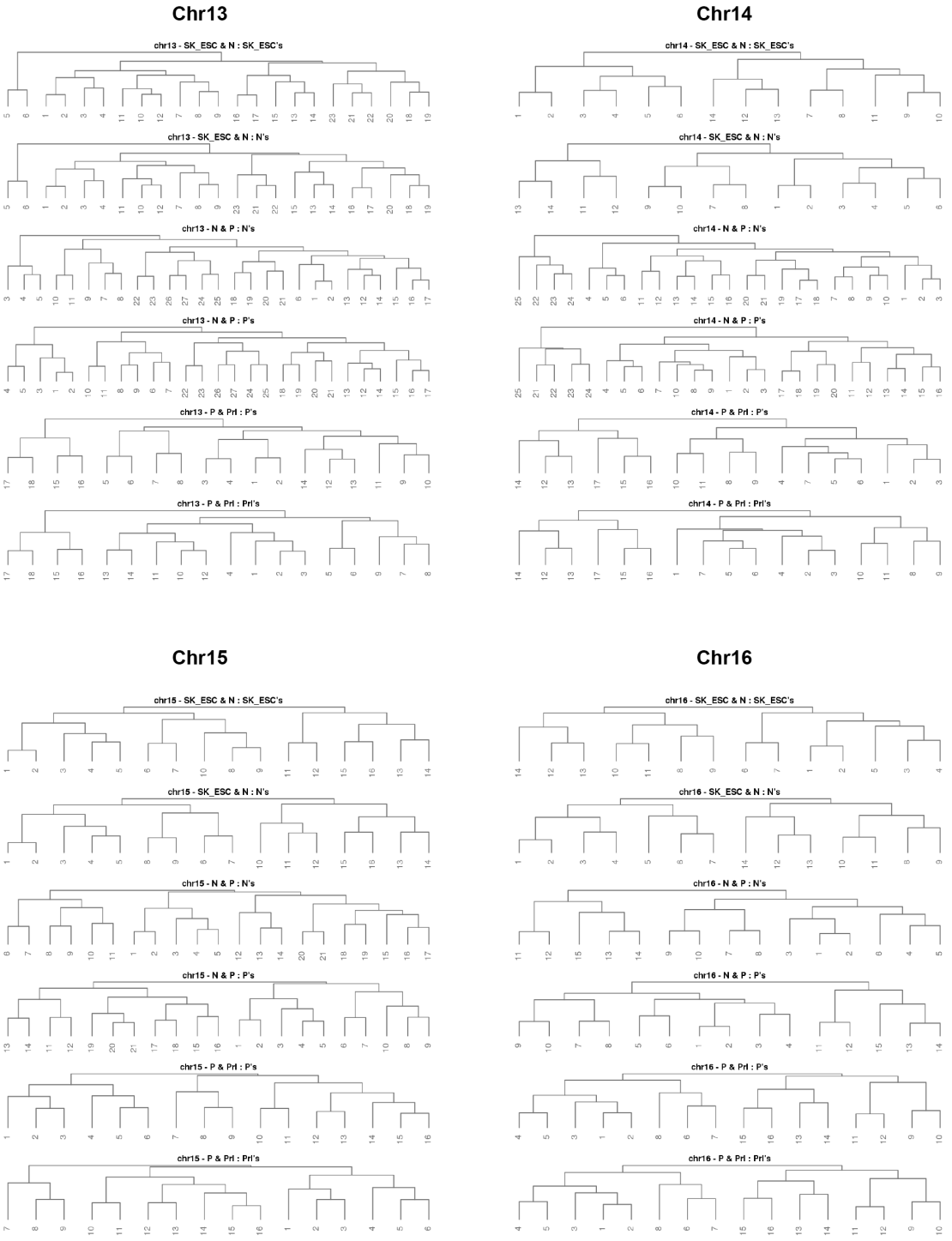


Fig.46. Hierarchical organization of Chr13-16 TADs accompanies ESC and HC11 lactogenic differentiation: Dendrogram representing hierarchical clustering analysis between distances and in between cophenetic correlations showing variable patterns of clustering among chromosomes (13-16) between ESC-N, N-P and P-PRL conditions.

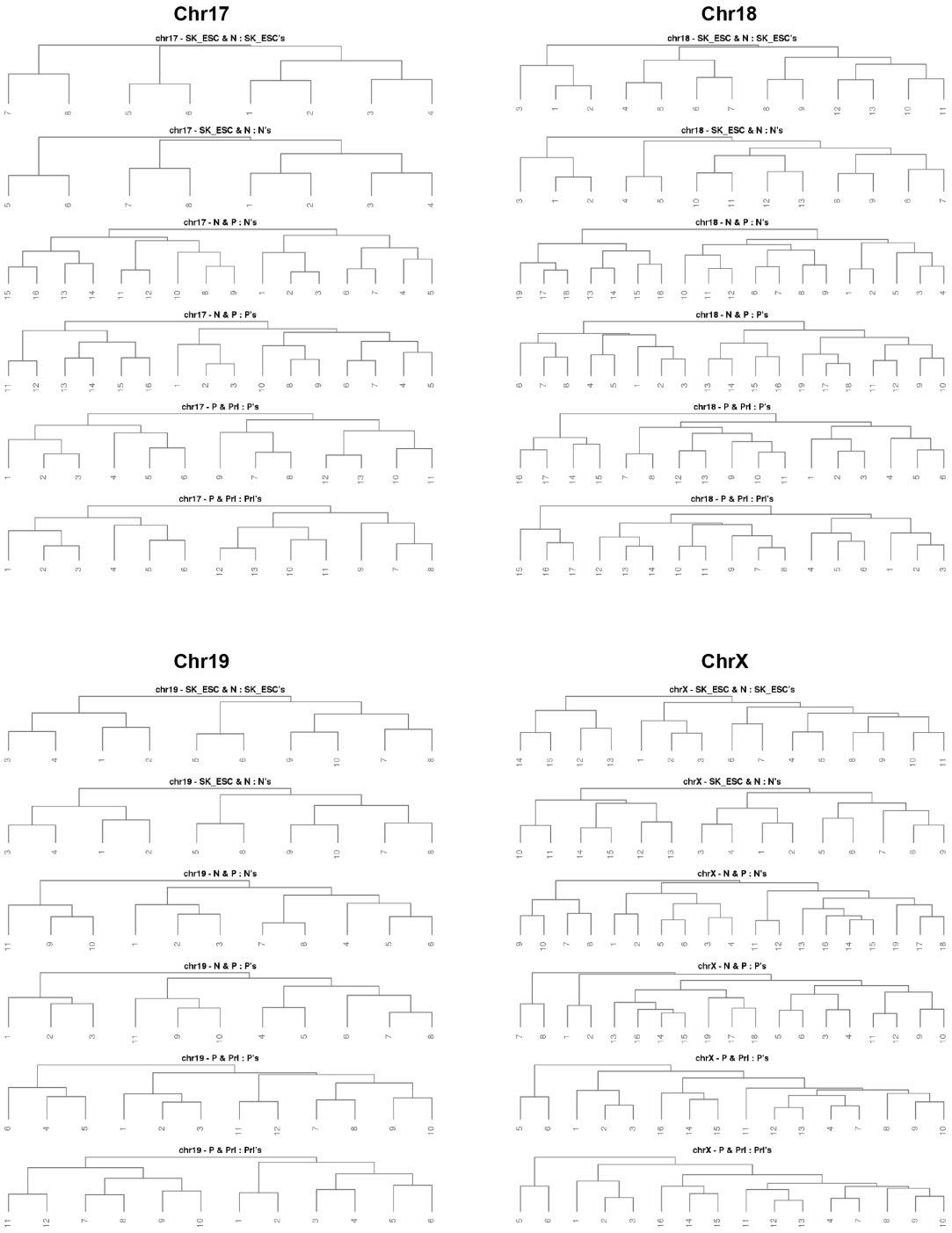


Fig.47. Hierarchical organization of Chr17-X TADs accompanies ESC and HC11 lactogenic differentiation: Dendrogram representing hierarchical clustering analysis between distances and in between cophenetic correlations showing variable patterns of clustering among chromosomes (17-X) between ESC-N, N-P and P-PRL condition.

Discussion

Objective 1:

The development of mammary gland and lactogenic differentiation of MECs are highly complex biological processes and mediated by several endocrine factors including glucocorticoids, prolactin, insulin and EGF. A precise and coordinated control of gene expression is vital for lactogenesis. Recent microarray based studies in HC11 MECs investigated the role of PRL and its relevance to breast cancer(Williams et al., 2009). However, the mechanism of multistage lactogenic differentiation of MECs remains unclear. To gain insights into the mechanism of lactogenic differentiation, we employed HC11 MECs and compared the profile of transcripts from normal, GC primed and PRL treated HC11 cells. Further, we garnered the clues about key TFs and ERs that drive the lineage specific differentiation of ESC to MECs. It is interesting to note that priming induces PRL receptor in MECs, which helps in eliciting PRL signalling following treatment with PRL. Our data suggests that lactogenic differentiation of MECs occurs in a hierarchical manner, where priming with GC might alter chromatin accessibility and sets a base required for the PRL to elicit its action during lactation.

HC11 MEC are widely used in lactogenic differentiation system *in vitro* and form mammospheres during lactogenic differentiation(Morrison and Cutler, 2009). The transcriptome profile revealed that only a subset (22%) of genes was differentially expressed in MEC, when compared to ESCs. *Sox9, Runx2, Gata3, Ehf, Elf5 a*nd *Bcl11a* are TFs and ERs that are critical for the mammary gland development and are reported to maintain luminal progenitor population in the mammary gland(Malhotra et al., 2014, Ferrari et al., 2013, Kouros-Mehr et al., 2006, Zhou et al., 2005). We found that these factors are highly expressed in normal HC11, which resembles luminal progenitor cells of the mammary gland. We found up-regulation of Smad3 in MEC when compared to ESC. Smad3 is a downstream molecule of TGF-β and antagonizes STAT5/PRL signalling and promotes cellular proliferation of normal mammary epithelial cells(Cocolakis et al., 2008). We also observed that major regulators of stemness such as *Pou5f1, Sox2* and *Nanog* are downregulated in MECs indicating that cell type specific gene expression is prerequisite for specialized functions.

We observed that some of the genes (*Brca1, Krt6a, Krt6b, Krt5* and *Melk),* which are indicative of stem cell-like characteristics are highly expressed in normal MEC, and are dow regulated during differentiation, are majorly involved in cell proliferation. We speculate that normal HC11 cells become proliferative upon exposure to EGF. *Ncapg, Cenpa, Cenpe, Arukb, Kif4* and *Kif22* are some of the genes

involved in mitotic chromosome condensation and segregation and are majorly required for cellular proliferation. These genes are upregulated in normal MEC (treated with EGF) whereas in response to GC treatment, expressions of these genes are down \regulated. Another class of Ets family of TFs are *Etv4* and *Etv5* which are majorly involved in cell motility and cell growth, are highly down regulated since they are not required in normal MECs(Hollenhorst et al., 2011). During priming with GCs, HC11 MEC ceases cell proliferation concomitant with downregulation of above-mentioned genes.

Further, our data suggests GC alone brings about differential expression of vast number of genes (Fig.2). Priming of MECs with GC induces various ERs; *Cited4, Luzp2, Jade1, Jade2, Chd9, Myc* and *Tox*. Earlier studies showed that Cited4 is strongly involved in priming of MECs during lactogenic differentiation(Yahata et al., 2002). *Cited4* binds to CBP/p300, a transcriptional regulator as well as isoforms of *TFAP2* transcription factor. Cited4 thereby modulates *TFAP2* dependent genes, thus might be helpful to access the chromatin by signal transducers and transcriptional activators during PRL treatment(Yahata et al., 2000, Yahata et al., 2002, Braganca et al., 2002, Shioda et al., 1997). Jade family proteins serve as transcriptional activators via their histone acetyltransferase activity(Panchenko et al., 2004) and *Id* family TFs are diversely expressed during mammary gland development. *Id2*, which gets induced during GC treatment, is crucial for survival and differentiation of MEC. This suggests that GC is critical for differentiation of MEC during the course of lactogenesis(Kim et al., 2013, Roschger and Cabrele, 2017). GC activated GCR was shown to remodel chromatin in Brg1 dependent and independent manner(John et al., 2008). In this regard, GC induction of Chd9, a remodeling enzyme shown to play role in loosening chromatin in growing oocyte(Ooga et al., 2018) and Myc, a protein shown to play an important role in decompaction of chromatin(Kieffer-Kwon et al., 2017) suggest that there might be a link between these genes in global remodeling of chromatin, however, its relevance in the context of GC signalling remains to be explored further.

Transcriptome profile of HC11 cells treated with PRL revealed that only 3.5% genes were differentially expressed compared to GC primed HC11 cells. Out of the highly expressed TFs and ERs, *c-Myc* is expressed during lactation. Knockdown studies of *c-Myc* in mammary gland showed defect in alveolar cell proliferation and differentiation resulting in decreased milk production(Stoelzle et al., 2009). High expression of E-Cadherin (*Cdh1*) and downregulation of *Klf4* in HC11 cell-types suggest that HC11 cells are partial primary cells(Yori et al., 2010) as it was known that Klf4 suppresses epithelial to mesenchymal transition in breast cancer cells through induction of E-cadherin. PRL treatment induces

milk fatty acid biosynthesis by inducing expression of various TFs such as *Fabp4, Lipe* and *Srebf1*, which are major key regulators of fatty acid homeostasis in lactating mammary gland(Zhou et al., 2015, Zidi et al., 2010, Rudolph et al., 2010). In this context, we observed PRL mediated actions are preferably localized to milk production and survival of lactating epithelial cells, as GC sets a base for PRL by modulating accessibility of chromatin and survival of epithelial cells.

Previously reported transcriptome data of HC11 was obtained using microarray and are of 15 and 20K gene probe resolution(Perotti et al., 2009, Wang et al., 2009). In this study we provided RNA-seq of higher resolution with nearly 50-60 million reads, whereas the recommended coverage is 15-20 million read(Liu et al., 2014). The data in our study provides a greater insight about gene expression in response to GC and PRL treatment during lactogenic differentiation of HC11 mammary epithelial cells.

Objective 2 & 3

i) **Lactogenic differentiation of HC11 cells accompany subtle changes in chromosome territory neighbourhood:**

Experiments performed in human cells to monitor CT changes throughout the interphase and during mitosis have demonstrated that in absence of gross changes in CT proximities, CT undergo rotational changes, thus potentially pairing up with different chromatin region in linked and un-linked chromosomes. In this regard, HC11 cells grown to confluence make extensive cell-cell contacts and were known to cease to proliferate, perhaps by growth arrest, and enter lactogenic differentiation process by serial additional of GC and PRL hormones. This was confirmed by FACs analysis in our study. This model system allowed us to monitor changes in gene expression patterns in states, where there is a cessation of growth, and chromosomal structural changes that precede gene expression in the absence of cell cycle progression. This system provides the advantage of studying multiple signalling events on differentiated cells. Also, the comparison of the data with mESC, which is the default ground state of a cell, provided us with lots of insights about the 3D reorganization of chromatin during differentiation process, in general.

We could clearly see in the subtraction CTNs that HC11 Normal is very different from that of ESC. It can be attributed to the fact that ESCs are proliferative and have to go through many intermediate states in the lineage to reach the HC11 N state. However, since the HC11 MECs were lineage committed, do

not divide and are closely related to each other, there are very subtle differences among HC11 MECs. Interestingly, the difference is less pronounced between P and Prl states, which means that there is little to no reorganization in the CTN of PRL state from primed one. These subtle changes in CTN among GC and PRL treated HC11 cells could be due to local interactions between genes or repetitive elements. It was known that CTs do not change their position throughout interphase stage of cell cycle. Moreover, long-range contacts in cell nuclei may involve CT rotational movements as well as constrained chromatin movements locally(Strickfaden et al., 2010). In line with these propositions, HC11 cells upon confluence are non-proliferative but undergo differentiation upon GC and PRL treatment, which was seen as subtle changes in chromosome territory positioning but dramatic reorganization happened within and in between chromosome's long and short-range contacts.

ii) Dramatic reorganization of sub-chromosomal domains accompany lactogenic differentiation of HC11 cells

Intra-chromosomally, there is lot of switching between A and B compartments between cell types. Genes present in the switched Compartments are generally assumed to change their gene expression, accordingly. But in our analysis, it was shown that, although the genes present in A show higher expression that that of in B, there is no significant change in the gene expression of the genes present in the switching compartments between two cell types. This led us to study the subtraction matrices of cis matrices between the cell-types, where, we observed that changes in the interactions between two cell types are not uniform along the chromosome but are confined to specific regions or blocks in the heatmaps. Upon studying, these regions turned out to be hetero- and Euchromatin domains, which in turn were in concordant with the boundaries of cLADs and iLADs, respectively. The interactions between the cLADS intra-chromosomally increased from N to P and then decreased from P to PRL state. On the contrary, interactions between the ciLADs decreased from N to P and then are maintained from P to PRL. On the other hand, it is interesting to note that ESCs have fewer intra-chromosomal long-range interactions than that of HC11 MECs, which was evident from LvS ratios. Increase in LvS ratio from normal to GC primed HC11 cells suggests that decrease in local interactions within cLADs is accompanied by increase in long-range interactions between cLADs. Alternatively, decrease in long-range contacts from GC primed to PRL treated HC11 cells suggests a gain in local interactions, as was seen during senescence associated heterochromatin loci (Chandra et al., 2015). Perhaps, PRL receptor activated stat5 proteins may mediate local short-range contacts to facilitate enhancer-promotor

interactions in PRL state. Infact, studies suggest that GR mediates long-range interactions (Kuznetsova et al., 2015) and large-scale chromatin decompaction of GR regulated genes (Jubb et al., 2017).

It is also interesting to note that cilAD-ciLAD interactions within chromosome in PRL state are similar to that GC primed cells. On the other hand, through gene expression analysis suggest that thousands of genes are differentially expressed. based on these observations, we speculate that PRL induced enhanced interaction between iLADs may lead to clustering of these loci for transcriptional activation resulting in interactions similar to that of GC treated cells. Long-range Interactions among ciLADs within chromosomes under primed and PRL state are significantly lower than that of normal HC11 cells and is in line with increase in short-range interactions (<10kb) in GC primed (18.1%), PRL (20.3%) when compared to normal HC11 cells (15.5%).

The intra-chromosomal interactions between cLAD and ciLAD provide new insights into dynamics of LADs. Our genome-wide analysis of cLAD and ciLAD interactions show that ESCs have interactions among themselves but surprisingly these interactions are dramatically lower in normal HC11 cells which is in line with the assumption that differentiated cells gain a clear LAD structure at nuclear periphery which are relatively dense, and hence reduction in long-range interaction among them. However, Upon GC treatment, there seems to be significant gain in long-range interactions suggesting GC activated rise in long-range contacts in cLADs might have participated in regulation of genes in ciLAD regions and continue to do so under PRL state. It will be interesting to see if this is linked to remodeling of LAD regions so that those regions move away from nuclear lamina and interact internally with other active genes. Parallelly, we also analyzed the inter-chromosomal interactions and their dynamics during differentiation. The cLAD-cLAD interactions decreased from N to P and then increased from P to PRL state. This could be due to the fact that cLADs, which were interacting with trans ciLADs, are not available for interacting with cLADs of other chromosomes in P state.

Our genome-wide analysis of interactions among cLADs intra-chromosomally and inter chromosomally showed significant increase, which is contradictory to the observations made, perhaps due to the fact that we analyzed all the chromosomes in trans but not any two different chromosomes. Indeed, our analysis of interactions among cLADs shows that they interact significantly within chromosome rather than between chromosomes. LADs are known to harbour two SINE and SINE

repeat independent regions. It is also know that SINE containing elements from different chromosome cluster to assume chromocenters (Kuznetsova et al., 2016). Perhaps the inter chromosomal interactions among cLADs is a reminiscence of its possible involvement in formation of chromocenters. However, possibility of such things remains to be explored further.

iii) GC promotes large scale clustering of cLADs but accompanied by subtle changes in CTN

It is intriguing to see that GC treatment lead to long-range interactions among cLADs and is not accompanied by change in gene expression patters other that few key genes which found to be expressed at high levels such as Csn2. This phenomenon of increase in long-range contacts between cLADs globally without having dramatic effects on gene expression indicates other essential role for cLADs in the genome, however, physiological significance of such phenomenon requires further investigation. Though, it is evident that GC treatment is pre-requisite for Csn2 to be transcribed, based on our observations, it is suggesting that GC treatment could lead to relocation of genes from LAD regions to as to expose them to subsequent PRL signalling. It is known that GR receptor upon binding to the enhancer elements later recruits Stat5 for augmented transcription of Csn2 suggesting pivotal role for GR in in making accessible to Stat5 binding and gene expression. GC and PRL treatment leads to dramatic reduction in inter-chromosomal ciLAD-ciLAD interactions. However, gene expression analysis suggests that there is dramatic increase in expression of many genes. It may be because the Stat5 binding leads to smaller loops and thus the interactions among them decrease. It may be also because GC and PRL treatment may lead to clustering of genes three dimensionally leading to dramatic decrease in interaction among themselves or there is overall reduction in gene expression among ciLAD regions when compared to normal cells. However, it remains to be determined. Based on gene expression data suggest that there is overall reduction in gene expression profiles under GC and PRL conditions. Inter-chromosomal cross domain interactions between ciLADs and cLADs in normal HC11 cells some showed dramatic increase in interactions from ESC. These interactions are then decreased in P state and increased in PRL state. This is just the opposite of these interactions intra-chromosomally. The fact that N has lot of inter-chromosomal ci–clAD interactions suggests that there are more than one type of cLADs, which is evident from the fact that pericentromeric satellite DNA from different chromosomes cluster to hold all the chromosomes together during nuclear membrane break down (Jagannathan et al., 2018). GC mediated expression of genes through GR is shown to be cell and tissue specific owing to the fact that pre-existing chromatin structure (DHS) would determine

the functional output of transcriptome and seems to sets a stage for other signalling molecules to act upon. Hence, synergistic action of multiple signalling events in a hierarchical fashion by GC followed by PRL might facilitate lactogenic specific gene expression program. Our studies on integrative visualization of gene specific microdomains (1Mb) of stage specifically expressed genes such as Etv4 (expressed highly in HC11 normal cells), Krt23 (expressed highly in GC primed cells) and Csn2 (Expressed highly in PRL state) suggest that irrespective of transcription status, 1Mb microdomains undergo rhythmic loss or gain of short vs long-range contacts in a stage specific manner. It remains to be determined how does this fits into formation of active or silent domains within the 3-dimensonal space of the nucleus in space and time. Further, the analysis on TADs showed that undergo dramatic reorganization during lactogenic differentiation. These results suggest that TADs are dynamic and undergo an extensive reorganization upon GC and PRL signalling. Perhaps future studies should aim at dissecting interactions pertaining to genes that fall within these reorganize TADs which require much deeper sequencing of Hi-C libraries from these samples.

iv) TADs undergo dramatic reorganization during lactogenic differentiation of HC11 cells

Hierarchical reorganization of TADS under three states (ESC-N, N-P and P-PRL) could be attributed to combinatorial usage of various regulatory elements for execution of stage and signal specific gene expression patterns, however, to draw meaningful evidence for these studies require more in-depth sequencing of Hi-C libraries as well as other epigenomic features association such as active or silent histone modifications, DNAse, GRO-seq. In the hierarchical TAD clustering analysis, when only MECs were analysed, it was clear that most of the reorganization happens upon GC signalling and that the N and PRL states have higher correlation, which reflects the changes we have observed in long-range interactions between the sub-chromosomal domains. When we analysed samples' pairwise, unlike in the TAD-subTAD interplay, most of the reorganization happens upon GC signalling, as opposed to in the expected ES-N comparison. Also it was observed that larger chromosomes didn't show much reorganization from ESC-N, compared to the other two comparisons, but smaller chromosomes did. This may possibly suggest that, larger chromosomes do not change much from ESC to N in terms of TAD clustering, which requires a lot of reorganization at folding level, which requires lot of energy input, and hence only smaller chromosomes reorganized. But upon GC signalling, the TAD clustering of all the chromosomes is extensively reorganized, making lot of architectural changes in Primed state. This provides another support for the importance of role of GC signalling in this system.

Summary

Summary:

It is now well established that 3-dimensional architecture of the mammalian genome plays an important role in cell-type specific gene expression and regulation during development and disease. Mammary gland is an apocrine gland and its key function is to produce and secrete milk. There are a few studies that define the cellular and molecular mechanisms responsible for production of protein and lactose content of milk. However, comprehensive picture of how lactogenic hormones, induced signaling pathways execute their effects on epigenome of mammary epithelial cells and the reorganization of 3D genome architecture for efficient milk production is yet incomplete and further studies are required for complete understanding of the core processes involved. Here in this study, I sought to study the 3D genome architecture and reorganization of mouse mammary epithelial cells undergoing lactogenic differentiation under the influence of glucocorticoid (GC) and Prolactin (PRL) signaling at the level of chromosome territories, sub-chromosomal domains, Topologically Associated Domains and micro-domains level. We achieved this by employing cellular and molecular methods in combination with extensive bioinformatics and computational tools. This helped us to understand the mechanisms, dynamics and trends associated with genome architecture during signal induced lactogenic differentiation and development, in general. These are the following objectives of our study.

<u>Objective 1</u>: Spatial dynamics of chromosome territory neighborhood and gene expression during Glucocorticoid and Prolactin signal induced lactogenic cellular differentiation of HC11 mammary epithelial cells

It's been shown that the interphase chromosomal chromatin is non-randomly organized to occupy discrete locations in nucleus, termed as chromosome territories (CT). Within and between the CTs, dynamic intermingling of chromatin fibersoccur in inter-chromosomal space, where important nuclear processes such as transcription, replication, splicing and repair take place. Initially, fluorescence and electron microscopy imaging techniques were used to probe the DNA and visualize the structural and functional organization of genomic elements, with very low resolution. With the advent of Chromosome Conformation Capture based technologies such as Hi-C (high throughput sequencing method based of 3C), it became possible to dwell into the 3D structures of chromosomes and genomic elements with unprecedented scale and resolution.

Signal induced lactogenic differentiation of HC11 mouse mammary epithelial stem-like cells system was established, according to the protocol, and morphological mammospheres formation was confirmed by light microscopic observations. Further, HC11 lactogenic differentiation was confirmed by performing real time PCR of selected cell-type specific gene markers such as *Csn2*, *Wap etc.* As per the literature, HC11 cells differentiate after cell cycle arrest and in order to assess the cell cycle progression during the course of HC11 cell differentiation, FACS analysis was performed which showed that ~80% all HC11 cells were at GO/G1 phase.

Next, whole transcriptome analysis was performed from three different stages of HC11 MECs differentiation process i.e normal HC11 cells (N), GC primed (GC signalling) (P) and PRL (PRL signalling) treated HC11 cell by RNA-sequencing and were compared with embryonic stem cells. Gene expression analysis revealed thousands of genes that are differentially expressed between the cell-types and the relevant pathways, which are in accordance with the functional profiles of the cell-types. To assess the dynamics of 3D genome architecture, we performed in-nucleus Hi-C on all the three HC11 states, along with mouse Embryonic Stem Cells (mESC). At the genome level, large blocks of inter-chromosomal enriched contacts were observed, which were pre-existing translocations in HC11 cells. Upon normalizing for this and other biases, we generated probabilistic CT neighborhood maps (CTN), which showed that CTN of ESC were different from HC11 cells. Also, there are very subtle changes between the CTN of HC11 cells, which was also evident from subtraction CTNs, showing that lactogenic differentiation of HC11 cells under GC and PRL signaling undergo very subtle changes at CTN level.

<u>Objective 2</u>: Spatial dynamics of intra- and inter-chromosomal interactions and gene expression during GC and PRL signal induced lactogenic cellular differentiation of HC11 cells

At chromosomal level, interaction matrices showed that ESCs have lesser long-range interactions compared to HC11 cells, which was evident from the rhomboid shape of intra-chromosomal Hi-C matrices of ESC. To quantitate the long-range interactions, we calculated LvS ratio for cis matrices of all the cell-types. The LvS values of ESC were lower than HC11 cells. Interestingly, the values hike in HC11 (N) and in GC primed state and then decrease moderately in PRL state. This trend was observed globally in all chromosomes and this was supported by the mapping percentage of Hi-C reads with HiCUP. Then, we overlaid the Hi-C heatmaps on gene expression, PCA and LAD data to

see the nature of long-range interactions. We observed that the increase in the long-range interactions from ESC to N span across both A and B compartments. Surprisingly, the increase in LRI from N to P span compartment B and the LRI from P to PRL span compartment A, which were respectively cLADs and ciLAds. To look at the gene expression perspective, the expression (RNA-Seq) of the genes physically located in A and B compartments were derived and found to be higher and lower respectively, as expected. But the expression of the genes which are located in switching compartments between cell-types are not correlated unlike what was expected.

Objective 3: Spatial dynamics of sub-chromosomal domains and gene expression during GC and PRL signal induced lactogenic differentiation of HC11 cells

Next, we sought to study the dynamics of these sub-chromosomal domains, which are cLADs and ciLADS during the HC11 differentiation. It was evident that compartment A highly correlates with ciLADs and compartment B with cLADs respectively. The *cis* subtraction matrices clearly showed visually that intra-chromosomal interactions between cLADs increase from N to P and then decrease from P to PRL state. Conversely, the interactions between ciLADs decrease from N to P and are then maintained from P to PRL. We quantitated these interactions globally and found that this is true for almost all the chromosomes statistically. These results were validated through 3C and 3D DNA FISH, by selecting probes from Chr6. To further confirm these results, we took advantage of naturally occurring chromosomal translocated data. We reasoned that the translocated part of a chromosome would show interactions similar to a chromosome in which it is translocated and was extensively proved by bioinformatics analysis. Further, inter-chromosomal interactions were studied and was found that cLAD interactions between the chromosomes increase from ESC to HC11 (N), decrease from N to P and then increase from P to PRL state. Surprisingly, cross domain interactions (cLAD-ciLAD) between chromosomes increased from ESC to HC11 (N), which was unexpected. Further, we studied 1Mb microdomains of important cell-type specific genes and showed their interaction dynamics through circos, Arc plots and subtraction matrices.

We then studied the dynamics of TADs during in ESCs and during HC11 cell lactogenic differentiation. We found that the TADs were not as conserved as previously thought, though the TADs of HC11 cells were more conserved among themselves than with that of ESC. Also, there is a lot of TAD-subTAD interplay between the cell-types, where TADs lose boundaries and vice versa. Further, hierarchical clustering of TADs showed that HC11 (N) TADs cluster similar to PRL TADs.

Moreover, it was also seen that reorganization of TAD clustering happens mostly between N and P but not from ESC to N.

Over all our results unravel the basic principles of chromatin organization and its dynamics in the context of Glucocorticoid and Prolactin signalling in orchestrating lactogenic differentiation specific gene expression pattern which has implications in understanding mechanisms of cell differentiation and it dysregulation in disease perspective in general.

Tables

Table 1

RNA-seq paired end sequencing data quality statistics				
Sample name	Raw Paired End reads	GC content (%)	Q20 (%)	Q30 (%)
NI	3,34,66,574	50.00%	97.44	93.53
NII	2,89,86,442	50.00%	97.19	93.03
PI	3,20,41,815	49.00%	97.28	93.22
PII	3,25,27,414	49.00%	97.18	93.03
PrlI	3,22,53,820	50.00%	97.12	92.85
PrlII	4,12,98,528	50.00%	97.05	92.91
ESC I	4,82,11,782	51.00%	97.13	93.32
ESC II	3,49,59,564	51.00%	97.03	93.06

RNA-seq paired end sequencing reads maping statistics								
Sample name	NI	NII	PI	PII	PRLI	PRLII	ESCI	ESCII
Total reads	66005456	56766030	62850938	63991192	63735374	81540128	41434474	30850216
Total mapped	58212828(88.19%)	50025174(88.13%)	56227898(89.46%)	56396196(88.13%)	56227898(88.22%)	71603961(87.81%)	32640360 (78.7759%)	26156498 (84.7855%)
Uniquely mapped	58212828(88.19%)	50025174(88.13%)	56227898(89.46%)	56396196(88.13%)	56227898(88.22%)	71603961(87.81%)	32640360 (78.7759%)	26156498 (84.7855%)
Read-1	29643572(44.91%)	29643572(52.22%)	28209994(44.88%)	28826381(45.05%)	28825154(45.23%)	37138361(45.55%)	16340858 (39.4378%)	13078513 (42.3936%)
Read-2	28569256(43.28%)	24446137(43.06%)	27040086(43.02%)	27569815(43.08%)	27402744(42.99%)	34465600(42.27%)	16299502 (39.338%)	13077985 (42.3919%)
Reads map to '+'	29153302(44.17%)	25051011(44.13%)	27653702(44.00%)	28232196(44.12%)	28177178(44.21%)	35887200(44.01%)	16333607 (39.4203%)	13089066 (42.4278%)
Reads map to '-'	28569256(43.28%)	20381602(35.90%)	28017904(44.58%)	27569815(43.08%)	27402744(42.99%)	34465600(42.27%)	16306753(39.3556%)	13067432(42.3577%)

Table: 1 RNA-seq read sequnces statistics and maped sequences statistics

Table 2

RNA-seq correlation statistics between replicates and samples								
	ESCI	**ESCII**	**NI**	**NII**	**PI**	**PII**	**PRLI**	**PRLII**
ESCI	1	0.979018	0.599605	0.60562	0.437505	0.44077	0.433995	0.451157
ESCII	0.979018	1	0.594121	0.598564	0.421669	0.425186	0.418062	0.436234
NI	0.599605	0.594121	1	0.996613	0.749792	0.757137	0.746801	0.776996
NII	0.60562	0.598564	0.996613	1	0.751264	0.758722	0.748788	0.778865
PI	0.437505	0.421669	0.749792	0.751264	1	0.998428	0.938284	0.938147
PII	0.44077	0.425186	0.757137	0.758722	0.998428	1	0.938891	0.940874
PRLI	0.433995	0.418062	0.746801	0.748788	0.938284	0.938891	1	0.996789
PRLII	0.451157	0.436234	0.776996	0.778865	0.938147	0.940874	0.996789	1

Table:2 RNA-seq correlation statistics between replicates and samples

Table 3								
	Esc		**N**		**P**		**Prl**	
S.No.	**Gene**	**FPKM**	**Gene**	**FPKM**	**Gene**	**FPKM**	**Gene**	**FPKM**
1	*Dppa5a*	1583.9	*4732456N10*	19.02312	*Mir7051*	3.960329	*Csn2*	119.5648
2	*Pou5f1*	1066.8	*Gzmc*	9.653638	*Il1rn*	3.630448	*Klk1b1*	27.94152
3	*Zfp42*	1010.4	*Il6*	8.674601	*5031434O11*	3.216255	*Klk1b9*	17.04681
4	*Slc2a3*	478.28	*Fam132b*	5.860947	*Mir3963*	2.975995	*Klk1b11*	14.40445
5	*Tdh*	448.26	*Krt6b*	4.918103	*Tbx6*	2.811827	*Klk1b5*	14.28051
6	*Rps4l*	413.23	*Dlx3*	4.723839	*Mirlet7e*	2.633297	*Wap*	8.950675
7	*Tdgf1*	360.16	*Krt79*	4.360654	*Mir6386*	2.550191	*Klk1b3*	8.72859
8	*Sox2*	351.9	*Hoxc9*	4.039718	*Mir7059*	2.419857	*Hsd11b2*	6.318191
9	*Nanog*	316.91	*Slurp1*	3.961538	*Mir26b*	2.169829	*Klk1b4*	4.829544
10	*Trh*	305.56	*Ptgs2os*	3.645179	*Mir6911*	2.061781	*Clec3a*	4.025162
11	*Esrrb*	272.38	*Gm6460*	3.590083	*Mir7087*	2.061395	*Mir143*	3.827188
12	*Ckb*	254.46	*Hmx2*	3.253492	*Mir26a-1*	2.054936	*Pde6h*	3.461363
13	*Dnmt3l*	233.08	*Spink6*	3.193648	*Mir5100*	1.9371	*Lrg1*	3.38325
14	*Bex1*	223.43	*Cdh15*	2.898545	*Adamtsl2*	1.912625	*Kcnj15*	2.944297
15	*Utf1*	218.82	*Bdkrb2*	2.510905	*Mir1928*	1.893815	*Hcst*	2.274105
16	*Rpl10l*	205.58	*Prkg2*	2.486092	*AB124611*	1.891634	*Klk1b21*	2.24748
17	*Triml2*	168.91	*Adam12*	2.482224	*Mir3109*	1.859131	*Klk1b16*	2.186739
18	*L1td1*	163.34	*Epgn*	2.462729	*Ccdc39*	1.842384	*1700084E1*	2.112219
19	*Laptm5*	160.98	*Bpifc*	2.45596	*Mir7048*	1.795859	*Hist1h4c*	1.958821
20	*Fbxo15*	146.66	*Snord23*	2.417601	*Nap1l3*	1.785778	*Klk1b22*	1.886563

Table: 3 Top 20 Uniquely expressed genes among ESC, N, P, PRL treated HC11 MECs

Table 4								
	ESC		N		P		Prl	
S.No.	Gene	FPKM	Gene	FPKM	Gene	FPKM	Gene	FPKM
1	*Eef1a1*	8144.65	*Eef1a1*	5088.9301	*Cst3*	6375.449	*Actg1*	8107.961
2	*Ftl1*	6259.55	*Actg1*	4790.0333	*Spp1*	6297.858	*Ctsd*	7593.725
3	*Rps2*	5236.57	*Ppia*	4755.2997	*Wfdc18*	5472.739	*Cst3*	6397.193
4	*Rpl41*	4063.9	*Rpl41*	4209.9603	*Ctsd*	5224.604	*Spp1*	5298.26
5	*Rplp1*	3723.04	*Uba52*	4029.4308	*Eef1a1*	3448.458	*Igfbp5*	4647.586
6	*Rpl13*	3342.33	*Rps2*	3887.8535	*Actg1*	3416.57	*Wfdc18*	4552.576
7	*Rpl12*	3332.51	*Rplp1*	3686.3781	*Uba52*	2881.579	*Eef1a1*	4141.19
8	*Rpl8*	2785.05	*Pkm*	3219.0967	*Ppia*	2780.603	*Gapdh*	3698.89
9	*Uba52*	2774.49	*Ptma*	3067.8125	*Sparc*	2416.252	*Sparc*	3591.025
10	*Rpl32*	2704.97	*Gapdh*	2992.8205	*Slpi*	2415.386	*Uba52*	3337.091
11	*Rps12*	2543.97	*Rpl13*	2590.0953	*Tpt1*	2328.907	*Ftl1*	3120.081
12	*Rps18*	2374.17	*Rps12*	2568.3407	*Rplp1*	2068.475	*Rplp1*	2956.715
13	*Rps20*	2370.24	*Rpl32*	2462.0911	*Gapdh*	2019.64	*Tpt1*	2923.565
14	*Rplp0*	2364.12	*Rpl37a*	2364.5684	*Pkm*	1972.405	*Aldoa*	2923.381
15	*Rpl13a*	2309.82	*Rpl23a*	2254.6483	*Rpl41*	1912.622	*Fth1*	2877.746
16	*Rps16*	2274.59	*Rpl7*	2129.5976	*Igfbp5*	1854.247	*Cbr2*	2659.615
17	*Rps26*	2250.49	*Krt14*	2077.0081	*Rpl23a*	1846.209	*Ppia*	2600.268
18	*Ptma*	2234.59	*Actb*	2045.8061	*Rpl13*	1817.512	*Pkm*	2395.564
19	*Apoe*	2218.59	*Rps26*	1966.2169	*Mir140*	1794.448	*Mt1*	2284.33
20	*Rps8*	2179.74	*Rps21*	1963.9867	*Rn45s*	1793.467	*Pgk1*	2271.16

Table:4 Top 20 Highly expressed genes in ESC, N, P, PRL

Table 5						
	Esc vs N		N vs P		P vs Prl	
S.No.	Gene	log2_FoldChan	Gene	log2_FoldChange	Gene	log2_FoldChange
1	*Krt5*	15.32823401	*Ogn*	7.75532407	*Csn2*	7.801687537
2	*Igfbp5*	14.08950686	*Cyp2f2*	7.142109765	*Ptgds*	5.206113922
3	*Krt14*	13.39420752	*Capn6*	6.861207298	*Klk1b9*	5.014153122
4	*Wfdc18*	11.52898508	*Ibsp*	6.609787835	*Arl4d*	5.010026293
5	*Brinp3*	11.44661622	*Adcy8*	6.369662606	*Pigr*	4.747801554
6	*Lect1*	11.27772993	*Ltf*	6.262762115	*Klk1b5*	4.148100209
7	*Pkp1*	11.24094278	*Ccl6*	6.170247887	*Klk1b11*	4.06871544
8	*Tnfrsf11b*	11.16635554	*Agt*	6.134832855	*Hsd11b2*	3.998158154
9	*BC006965*	11.02459587	*Ucma*	6.066072138	*Wap*	3.822632144
10	*Trpv6*	10.88496196	*Ccl9*	5.996742829	*Klk1*	3.804933922
11	*Serpinb5*	10.59677658	*Cxcl17*	5.990458828	*Klk1b4*	3.660879237
12	*Tacstd2*	10.56961486	*Dcn*	5.964853598	*Klk1b3*	3.437304572
13	*Arg1*	10.30600439	*Ctgf*	5.803830583	*Thrsp*	3.434664712
14	*Sostdc1*	10.26888283	*Slc28a3*	5.712901259	*Fxyd2*	3.35000972
15	*Postn*	10.1283288	*Itga10*	5.667809036	*Rnase1*	3.243550893
16	*Trp63*	10.08127265	*Sult1a1*	5.656325294	*Shc2*	2.909987656
17	*Wfdc3*	10.0230382	*Hapln1*	5.633495677	*Klk1b16*	2.901719417
18	*4732456N10Rik*	9.893710149	*Clcnkb*	5.603104893	*Ccl28*	2.844203135
19	*Palmd*	9.725131561	*Frzb*	5.590883692	*Man1c1*	2.751420208
20	*Slpi*	9.690368429	*Omd*	5.576741172	*Trib3*	2.723961494

Table: 5 Top 20 Up regulated genes between ESC vs N, N vs P, P vs PRL

Table 6						
	Esc vs N		N vs P		P vs PRL	
S.No.	**Gene**	**log2_FoldChange**	**Gene**	**log2_FoldChange**	**Gene**	**log2_FoldChange**
1	*Slc2a3*	-15.2181384	*Etv4*	-6.11085387	*Mki67*	-4.0517314
2	*Zfp42*	-15.16510123	*Gzmc*	-5.267010409	*Cenpf*	-3.897146995
3	*Pou5f1*	-14.94811449	*4732456N10Ri*	-4.52269063	*Cenpe*	-3.106607047
4	*Dppa5a*	-14.37677166	*Fam64a*	-4.183544007	*Shcbp1*	-3.078873505
5	*Tdh*	-14.14285591	*Krt42*	-4.106585421	*2810417H*	-2.958282585
6	*Esrrb*	-14.14203349	*Col1a2*	-4.063174476	*Esco2*	-2.78715017
7	*Bcat1*	-14.00489899	*Il6*	-3.925880661	*Pi15*	-2.687193274
8	*Tdgf1*	-13.8913459	*Ska1*	-3.916288942	*Ankle1*	-2.630049182
9	*Dnmt3l*	-13.72613328	*Id3*	-3.868965754	*Kif20b*	-2.604506806
10	*L1td1*	-13.42538795	*Hist1h3h*	-3.858399228	*Ccna2*	-2.599911318
11	*Sall4*	-13.38825634	*Kif2c*	-3.768115685	*Ucn2*	-2.523132605
12	*Nid2*	-13.2608396	*Ncapg*	-3.76527331	*Neil3*	-2.509479321
13	*Trh*	-12.65260982	*Bub1b*	-3.736363342	*Smc2*	-2.508097535
14	*Sall1*	-12.63766435	*Slc13a5*	-3.732063608	*Rrm2*	-2.495595717
15	*Trpm1*	-12.62208882	*Ccnb1*	-3.71246059	*Top2a*	-2.472868925
16	*Nr0b1*	-12.60219401	*Gm6455*	-3.685609436	*Dtl*	-2.448202926
17	*Utf1*	-12.59415214	*Lipg*	-3.649463807	*Depdc1a*	-2.436905448
18	*Triml2*	-12.5550186	*Aurka*	-3.638385884	*Kif23*	-2.424788606
19	*Fgf4*	-12.54561559	*Plk1*	-3.626224828	*Bub1*	-2.393794638
20	*Zic3*	-12.53728368	*Bub1*	-3.62113512	*Pbk*	-2.363063088

Table: 6 Top 20 Down regulated genes between ESC vs N, N vs P, P vs PRL

Table 7								
	Esc		N		P		PRL	
S.No.	Gene	FPKM	Gene	FPKM	Gene	FPKM	Gene	FPKM
1	*Hsp90ab1*	1976.098	*Rpl7*	2129.5976	*Rpl7*	1527.939	*Rps10*	1672.344
2	*Rps10*	1556.802	*Hsp90ab1*	1867.4739	*Rps10*	1345.569	*Rpl7*	1667.002
3	*H2afz*	1500.889	*Rps10*	1824.4116	*Hsp90ab1*	1094.349	*Rpl6*	1138.742
4	*Rpl6*	1305.209	*Rpl6*	1317.2611	*Rpl6*	935.3236	*Hsp90ab1*	1090.279
5	*Rpl7*	1097.269	*Pfn1*	1186.6408	*Pfn1*	831.9473	*H3f3b*	749.0038
6	*Pou5f1*	1066.83	*H3f3b*	1017.2563	*H3f3b*	720.9077	*Pfn1*	667.0828
7	*Zfp42*	1010.35	*H2afz*	665.58024	*H3f3a*	719.7943	*App*	599.849
8	*Eif4a1*	851.9483	*Eif4a1*	626.87744	*App*	574.1085	*H3f3a*	492.3438
9	*H3f3b*	736.1619	*H3f3a*	606.47896	*Aplp2*	489.7475	*Wfdc3*	442.4106
10	*Trim28*	664.6277	*Ifitm3*	575.44072	*Ctnnb1*	457.9231	*Gpx4*	422.1437
11	*Klf2*	626.9358	*Btf3*	470.11022	*Eif4a1*	379.4752	*Eif4a1*	384.0017
12	*Pfn1*	568.4511	*Hmgn1*	395.99021	*Ddx5*	366.4971	*Aplp2*	379.6544
13	*H3f3a*	536.6472	*Tceb2*	386.34822	*Btf3*	365.7076	*Srebf1*	366.1363
14	*Pcbp1*	510.8119	*Hmgb1*	372.49377	*Cpe*	360.3272	*Ctnnb1*	355.7742
15	*Ybx1*	494.86	*Cpe*	321.19102	*Sox9*	353.1805	*Btf3*	353.5031
16	*Hmgn1*	465.0941	*Ctnnb1*	303.50326	*Wfdc3*	311.6097	*H1f0*	341.9432
17	*Cdk4*	431.2959	*Ran*	291.07836	*Tceb2*	291.2541	*Rab7*	340.0984
18	*Mybl2*	417.6968	*Junb*	290.08049	*H1f0*	258.3652	*Sox9*	301.1385
19	*Pcna*	399.1924	*Pcbp1*	276.98778	*Rab7*	254.402	*Tceb2*	293.7541
20	*Tceb2*	396.6716	*App*	274.34078	*Gpx4*	233.0738	*Aes*	256.1077

Table: 7 Top 20 Highly expressed TFs in ESC, N, P, PRL

Table 8						
	Esc vs N		N vs P		P vs PRL	
S.No.	Gene	log2_FoldChange	Gene	log2_FoldChange	Gene	log2_FoldChange
1	*Trpv6*	11.01634506	*Rhpn2*	4.792380543	*Trib3*	2.723961494
2	*Wfdc3*	10.15577656	*Hivep3*	4.20745936	*Ddit3*	2.720577952
3	*Barx2*	9.376058743	*Ehf*	3.87275599	*Nr1d1*	2.615021016
4	*Elf5*	8.253576	*Tsc22d3*	3.382347247	*Maff*	2.57062252
5	*Ehf*	8.053754084	*Hist1h2bc*	3.040856345	*Atf3*	2.420964948
6	*Dlx5*	7.779149547	*Zbtb16*	2.94590288	*Zbtb16*	2.30423178
7	*Erg*	7.758035584	*Pknox2*	2.842219852	*Nfil3*	2.021541716
8	*Mgam*	7.71989528	*Fabp4*	2.610665183	*Fabp4*	1.941134778
9	*Nkd2*	7.074574561	*Il15*	2.507962962	*Hist1h4h*	1.866721592
10	*Trpv4*	6.209123375	*Grhl1*	2.460607602	*Sohlh1*	1.828916401
11	*Tlr3*	6.15455753	*Nrg2*	2.224168285	*Hist1h4i*	1.722094112
12	*Eya4*	6.081600167	*Per1*	2.199331579	*Hist1h4j*	1.633739979
13	*Fzd1*	5.981540284	*Wfdc3*	2.195140034	*Hist1h4c*	1.615628957
14	*Egr1*	5.867101772	*Zhx3*	2.168073032	*Atf5*	1.610315153
15	*Bcl11a*	5.823859545	*Tbx6*	2.126978174	*Myc*	1.456465341
16	*Foxc1*	5.810591903	*Hsf2bp*	2.097340632	*Klf9*	1.421452916
17	*Rab25*	5.720428103	*Elf5*	2.036550438	*Dbp*	1.404040652
18	*Mef2c*	5.609427415	*Gtf2a1l*	1.975079462	*Hsf2bp*	1.369430358
19	*Lmo1*	5.571011663	*Casz1*	1.910497958	*Cth*	1.355361089
20	*Sox9*	5.489814429	*Tle2*	1.890630369	*Fhl1*	1.35417149

Table:8 Top 20 Highly Up regulated TFs between ESC vs N, N vs P, P vs PRL

Table 9						
	Esc vs N		N vs P		P vs PRL	
S.No.	Gene	log2FoldChange	Gene	log2FoldChange	Gene	log2FoldChange
1	*Zfp42*	-15.26507018	*Etv4*	-6.11085387	*Mki67*	-4.0517314
2	*Pou5f1*	-15.04763559	*Id3*	-3.868965754	*Ccna2*	-2.599911318
3	*Esrrb*	-14.20371579	*Hist1h3h*	-3.858399228	*Neil3*	-2.509479321
4	*Dnmt3l*	-13.83011285	*Cenpa*	-3.602352579	*Top2a*	-2.472868925
5	*Sall1*	-12.73781141	*E2f7*	-3.561684068	*Erg*	-2.351562526
6	*Nr0b1*	-12.70401851	*Aurkb*	-3.414986342	*Hells*	-2.255545874
7	*Utf1*	-12.69374375	*Nek2*	-3.374209445	*Aurkb*	-2.13215254
8	*Zic3*	-12.63735501	*Grhl3*	-3.364614776	*Sgol2*	-2.122413389
9	*Tcl1*	-12.06323049	*Kif4*	-3.351110438	*Mcm5*	-2.077207001
10	*Sox2*	-12.05419632	*Troap*	-3.347689607	*Mxd3*	-2.075288576
11	*Rbmxl2*	-11.76323243	*Ccna2*	-3.333606758	*Nuf2*	-2.05321289
12	*Gli2*	-11.31095982	*Nuf2*	-3.317154441	*Zfp52*	-2.037175181
13	*Zfp57*	-11.15061339	*Foxm1*	-3.304696463	*E2f8*	-1.960844788
14	*Foxd3*	-11.10528126	*Top2a*	-3.302675033	*Camk4*	-1.92323071
15	*Ddx3y*	-11.07977661	*Kif22*	-3.260651475	*Kif4*	-1.906360162
16	*2410141K09Rik*	-10.86114487	*Mcm5*	-3.167783689	*Phxr4*	-1.895070251
17	*Nr5a2*	-10.75326755	*Fancd2*	-3.114306298	*Troap*	-1.870362152
18	*Otx2*	-10.74398862	*Asf1b*	-3.080042895	*Exo1*	-1.867012693
19	*Sohlh2*	-10.69000497	*Exo1*	-3.048947647	*Pole*	-1.858784935
20	*Cdx1*	-10.68702972	*Sgol2*	-3.045704784	*Armcx4*	-1.848435819

Table: 9 Top 20 Down regulated TFs in ESC vs N, N vs P, P vs PRL

Table 10								
	ESC		N		P		PRL	
S.No.	Gene	FPKM	Gene	FPKM	Gene	FPKM	Gene	FPKM
1	*H2afz*	1500.89	*H3f3b*	1017.2563	*Cited4*	727.465	Cited4	982.906
2	*H3f3b*	736.162	*H2afz*	665.58024	*H3f3b*	720.908	H3f3b	749.004
3	*Trim28*	664.628	*H3f3a*	606.47896	*H3f3a*	719.794	H3f3a	492.344
4	*Hist1h2ai*	644.743	*Snrpg*	487.2517	*Ctnnb1*	457.923	Ctnnb1	355.774
5	*Hist1h2an*	589.921	*Hmgn1*	395.99021	*Ddx5*	366.497	H1f0	341.943
6	*H3f3a*	536.647	*Pebp1*	392.92459	*Sox9*	353.181	Pebp1	324.916
7	*Ybx1*	494.86	*Hmgb1*	372.49377	*Pebp1*	316.648	Sox9	301.138
8	*Hmgn1*	465.094	*Ctnnb1*	303.50326	*H1f0*	258.365	Ddx5	250.094
9	*Mybl2*	417.697	*Snrpd2*	274.76863	*Hmgn1*	212.007	Cnbp	190.95
10	*Set*	371.116	*Snrpe*	270.06182	*Cnbp*	186.887	Eif4a3	179.452
11	*Snrpb*	356.39	*Hmgb2*	267.23078	*Morf4l1*	179.168	Morf4l1	165.393
12	*Sox2*	351.901	*Ybx1*	266.43246	*Eif4a3*	175.148	Hmgn1	148.391
13	*Erh*	324.924	*Snrpb*	257.43498	*Snrpg*	170.131	Cited2	145.305
14	*Pebp1*	319.574	*H1f0*	250.22416	*Cited2*	154.577	Snrpb	142.25
15	*Snrpg*	301.875	*Hmgn2*	246.63384	*Cbx3*	153.158	Ybx1	141.939
16	*Cnbp*	299.646	*Eif4a3*	238.21409	*H2afz*	143.998	Myeov2	135.994
17	*Hmgb2*	295.269	*Ddx5*	234.83801	*Snrpb*	139.866	Snrpd2	135.716
18	*Snrpd2*	289.501	*Erh*	233.99955	*Ddx17*	136.637	Chmp4b	132.257
19	*Ctnnb1*	286.983	*Cbx3*	228.35804	*Ddx3x*	133.889	Chmp2a	131.857
20	*Trp53*	270.804	*Cnbp*	188.83422	*Ybx1*	129.802	Snrpg	128.71

Table: 10 Highly Expressed Ers in ESC, N, P, PRL

Table 11

	ESC vs N		N vs P		P vs PRL	
S.No.	**Gene**	**log2FoldChange**	**Gene**	**log2FoldChange**	**Gene**	**log2FoldChange**
1	*Dmp1*	7.239662936	*Cited4*	2.5000789	*Rere*	1.245950496
2	*Ifih1*	6.232851479	*Luzp2*	2.226990088	*Srf*	1.201762539
3	*Sox9*	5.489814429	*Jade1*	1.933437169	*Myc*	1.456465341
4	*Aim1*	5.397658876	*Tox*	1.712963649	*Hist1h1c*	1.314487583
5	*Sox8*	5.35193064	*Jade2*	1.702924522	*Cish*	1.150745135
6	*Cited4*	4.810961562	*Chd9*	1.656489313	*Pcgf5*	1.115113197
7	*Hmgn3*	4.625788013	*Hist3h2a*	1.5238735	*Hist1h1e*	1.071895668
8	*Hr*	4.556241465	*Sox5*	1.44531088		
9	*Trps1*	4.459375255	*Ccdc39*	1.363617352		
10	*Tnip3*	4.287717713	*Tcf7l2*	1.342828784		
11	*Runx2*	4.083408183	*Cbx7*	1.270597383		
12	*Gata3*	3.988160156	*Nap1l3*	1.219217907		
13	*Lrrc45*	3.935627477	*Kmt2c*	1.162634011		
14	*Npm2*	3.615440481	*Hdac11*	1.12457349		
15	*Smarca2*	3.485086164	*Sox9*	1.033186751		
16	*Gpr97*	3.303220525	*Dhx58*	1.001269664		
17	*Ets2*	3.216794341				
18	*Helz2*	3.183900702				
19	*Sox6*	3.13678863				
20	*Cbx8*	2.979187249				

Table: 11 Top 20 highly Up regulated Ers in ESC vs N, N vs P, P vs PRL

Table 12						
	ESC vs N		N vs P		P vs PRL	
S.No	Gene	log2FoldChange	Gene	log2FoldChange	Gene	log2FoldChange
1	*Dnmt3l*	-13.83011285	*Hist1h3h*	-3.858399228	*Cenpf*	-3.897146995
2	*Sox2*	-12.05419632	*Ncapg*	-3.76527331	*Cenpe*	-3.106607047
3	*Mtcl1*	-11.23432766	*Cenpa*	-3.602352579	*Smc2*	-2.508097535
4	*Ddx3y*	-11.07977661	*Aurkb*	-3.414986342	*Hells*	-2.255545874
5	*Cdyl2*	-10.53235945	*Nek2*	-3.374209445	*Aurkb*	-2.13215254
6	*D1Pas1*	-10.49069554	*Hist1h2ai*	-3.308668624	*Mis18bp1*	-1.952022471
7	*Prdm14*	-10.3105828	*Fam81a*	-3.290018198	*Phip*	-1.837066055
8	*Tle4*	-10.30700995	*Cenpe*	-3.199387723	*Dzip3*	-1.818279308
9	*Sox1*	-9.28297948	*Mis18bp1*	-3.198726703	*Pif1*	-1.788941064
10	*Syce1*	-9.225132396	*Ercc6l*	-3.1300734	*Ncapg*	-1.752763551
11	*Prdm5*	-8.956398803	*Fancd2*	-3.114306298	*Atrx*	-1.752044457
12	*Cecr2*	-8.949381689	*Asf1b*	-3.080042895	*Ccdc39*	-1.636800798
13	*Gata4*	-8.777082066	*Sox11*	-3.024831135	*Cdc6*	-1.611750376
14	*Phf11d*	-8.622448031	*Cenpf*	-3.01493462	*Nek2*	-1.593489822
15	*Sox3*	-8.561866783	*Uhrf1*	-3.012965808	*Rad54b*	-1.576923643
16	*Tdrd9*	-8.514790252	*Pif1*	-2.99015814	*Smc4*	-1.567334735
17	*Ccdc160*	-8.403675998	*Rad54b*	-2.974559239	*Trps1*	-1.554976108
18	*Prdm16*	-8.286785361	*Hist1h2an*	-2.973361694	*Mybl2*	-1.521701065
19	*Satb1*	-8.25959285	*Brca1*	-2.915556691	*Zgrf1*	-1.50615673
20	*Mov10l1*	-8.043908361	*Chaf1b*	-2.851071709	*Cenpa*	-1.489307372

Table: 12 Top 20 Highly Down regulated ERs in ESC vs N, N vs P, P vs PRL

Table 13				
Complete list of gene specific primers used and its respective PCR prouct length				
S.No.	Gene Name	Forward Primer 5'----3'	Reverse Primer 5'-------3'	Size (bp)
1	*Beta-Actin*	TTACTGCTCTGGCTCCTAGCA	GACTCATCGTACTCCTGCTTGC	145
2	*Ogn+B24B30BB2:B24*	TACCATCATTACCAACCAAGAAAGA	TCACAGTAGACAGAGCCACTTA	85
3	*Cyp2f2*	TCCAGGACCCAAACCTCT	CAGGTACACCGTGAACACC	118
4	*Capn6*	TAACAACCGTGATACCTTCTTGC	GCGGTAAGTGCGTAGGTCC	109
5	*Adcy8*	CTTGGAGACAGAGAACCAAAGA	TTGGTCATGTCGTTGATCATTTC	90
6	*Etv4*	GCAGGAAGCCACCACTC	GGGACTTGATGGCGATTTGT	89
7	*Gzmc*	GCTTCCTGGTTCGAGACAAA	CTCCTCCTTAGCCTTGATGTTG	100
8	*4732456N10Rik*	CAACTTCCTGAGAGCTTTCTTTG	TTCAGGCTGCGGTTGTT	102
9	*Fam64a*	TTCCTCCACAGAACCTCTCT	CTCTGTCCAGCCTTTGAGAC	107
10	*Krt42*	AGCAGCCGGACAGAGAT	CAGGGCTGTTCTCCAGTGATG	102
11	*Csn-beta*	CCTCCTCTCTTGTCCTCCAC	TGTTCAACAGATTCCTCACTGG	123
12	*Ptgds*	TATTCAGCAGAGGCACCAAG	AGGTGGTGAATTTCTCCTTCAG	96
13	*Klk1b9*	GGCAGCACTACACCTTTCA	ACCTTCTCTATGTGGGCTTTG	105
14	*Arl4d*	GACTTCCCTTCTTTACCGTCTC	CACGCGGATCTTCTCAGTATT	85
15	*Pigr*	TGGCCTGGGTACCAGTAA	ATGGTCACATTTCTGCCTATGT	120
16	*Wap*	TGCCTCATCAGCCTTGTTCT	CACACTCCTCGTTGGTTTGG	157
17	*Mki67*	TCGTATCCAGCTGCCTGTA	TGAGTTGGATTGGTAGAACTGAAA	99
18	*Cenpf*	CCACAGAAACTCTTTGCAACTC	CCTCTAACCTCTTCCGTTCTTC	100
19	*Cenpe*	ATTGCGGTGCCCATCAT	CCCATCATCGTGTGTGTCTT	98
20	*Shcbp1*	CCACGAATTATCCAAACAAATGAAC	CCTTTACCTCAGAGGCCTTAC	97
21	*2810417H13Rik*	CCTCCACCTTTGTCACCAAT	TTGCCACTTGGGAGTTGG	101
22	*Tpx2*	CTTACTCTTTCGATGCCCCC	TCTCCAAGTTGGCCTTCTCA	108
23	*Nek2*	TCTGATGGCTTGAATGACCTC	TCCTTTGCTCTTCTGCAACC	121
24	*Krt15*	CAGATCGGGACTACAGCCAT	GTCAATCTCCAGGACAACGC	101
25	*Boc*	CATTCTCACACTCTCGCACC	ACAAGAGGACACACACCACG	132
26	*Pouf51*	GGCGTTCTCTTTGGAAAGGTGTTC	CTCGAACCACATCCTTCTCT	359
27	*Nanog*	CTCAAGTCCTGAGGCTGACA	TGAAACCTGTCCTTGAGTGC	120
28	*Sox2*	CCGCGTCAAGAGGCCCATGAA	CCCGCTTCTCGGTCTCGGACAA	149
29	Wfdc3	AGCTGTGTACCGGAGACGAA	CCAGGATTCTGGGACATTGA	120
30	Barx2	GAGAAGAAATTCCAGAAGCAGAAG	TTCCATTTCATCCTGCGATT	122
31	Elf5	GACGCAGGAGGAGTTCATTG	TGCCAGTCTTGGTCTCTTCAG	122
32	Tead3	CCGCTACATCAAGCTGAGGA	AATGCCAACCTGGTATTCCC	106
33	Rhpn2	CTGCAGATGCTGAAGGAGGA	CAGAGGAATGGTGAACGCTT	87
34	Per1	AGGCTTCGTGGACTTGACAC	GCTTTGCTTTAGATCGGCAG	106
35	Ehf	GTCTGCAGGAGTTCACGAGG	ATGACATTGTGTGCGGACTG	121
36	Fabp4	GTGGGAACCTGGAAGCTTGT	TGATGATCATGTTGGGCTTG	115
37	Grhl1	ACAGCTATAACAACCGCAGCA	TTTTCGTTCTTCATCCCGAA	107
38	Ddit3	CAGCACACATTCAGGCACTC	ACACCGTCTCCAAGGTGAAA	107
39	Nr1d1	GACATGACGACCCTGGACTC	CTGCCATTGGAGCTGTCACT	122
40	Maff	CGACAAGCACGCACTGAG	ACAAACACACACCCCGAAGT	92
41	Atf3	ACCCCTGGAGATGTCAGTCA	ACACTTGGCAGCAGCAATTT	100
42	Dnmt3l	GCATGGAAGAGACAGCCAGT	CCACGTACTTCAGCGTTCCT	121
43	Cenpa	ATTCCTCCAAGGATTCTGCC	TCTTCTGCGCAGTGTCTGAG	90
44	E2f7	GGAGCTGTTCAGACCCATTG	TCTTCTGTTTCCTGCTTGGC	122
45	Aurkb	GGTCTGCAGGGAGAACTGAA	AATCATCTCTGGGGGCAGAT	111
46	Kif22	CCTGTGTCCGAGCCATAGAC	CTGAGTGCTCTTCTCGCCAT	110
47	Ccna2	CTCTTTACCCGGAGCAAGAAA	TTCATTAACGTTCACTGGCTTG	124
48	Neil3	TACAGTGGCGTGGAAACCTT	AGCTCTGTTCTCACCCTCCC	123
49	Top2a	ACTGTGGAGATGAAAGCTGAGG	TTCTCTATTCGTTGCCGGAG	122
50	Mcm5	AGCTGCCACAATACCCTCAC	AAGTATGGATCCAGTGGGCA	119
51	Hells	CTGCGGGACTTGAGAAAGAG	CGGTACCGAATTTCAGTGGA	82
52	Cited4	CCTCCTGCCAGATGACAGTT	CAAGAACGTGGAATCCGAAG	96
53	Aim1	AAGACGTGTTCCATCAAAGTGC	CCTCGGTATCCCAGAAGGAT	124
54	Sox8	CTCAGCAAGACCCTAGGCAA	CTGGGTGGTCTTTCTTGTGC	109
55	Jade1	CCTGCTCTAGACCTGAAGATCG	GCACGTAGTAGTCATCCGGG	111
56	Hdac11	AGAGAAGCTGCTGTCCGATG	ATGGTAGCCACCACAAAGGA	123
57	Kmt2c	CTGGGAATATGCCAAATGGA	TTTGATAGTGGCTCCAAGGTG	111
58	Rere	CCAGCTCCTGCTTTCTGTGA	CGTTCATGAGTAGATGGTCCCT	124
59	Srf	ACAGACCTCACGCAGACCTC	CATACATCACTGCATGGGGA	124
60	Myc	GTGCTGCATGAGGAGACACC	AGGGGTTTGCCTCTTCTCC	105
61	Cish	CTAGGTGTGGGTGGCATTTG	AGTGGGCAGATAGGCAACAG	96
62	Mtcl1	AGGCCAGAGCTGTTTCAACC	AGGCTGTGCATCTCCAGTTC	81
63	Ddx3y	CAGGTCTACGGTGGCTGTTC	TGATTTCAGGTCAAGACCAACA	117
64	D1Pas1	TTGAGGCGACAGGGAATAAC	AGGAGTTGGGCGTGTGTAAC	113
65	Prdm14	AGATTTGGACCCTTTCGAGG	TGGCTCAGGTGACCATCTTC	104
66	Hdac7	CCACCTGTCAGACCCAAGTC	GGGATGCTTGCTGTTGTCTC	113
67	Terf1	TTTTGACAAGAGTTGCATCTGG	TCCAAGGGTGTAATACGCTCA	73
68	Fancd2	TCACGAGCCTTCCTGAAATC	TCAGGGAAGTATTCTGCACCA	90

Table: 13 All gene primers usend in this study

Table 14					
No.	**Antibody**	**Dilution**	**Catalog No.**	**Company**	Mol. wt
1	Stat5a	1:1000	SC-1081	Santa Cruz	92kDa
2	Stat5b	1:1000	SC-1656	Santa Cruz	94kDa
3	Cbp	1:1000	SC-369	Santa Cruz	265kDa
4	Dnmt1	3ug	ab92453	Abcam	183kDa
5	Dnmt3L	1:10000	SC-20705	Santa Cruz	43kDa
6	Hdac11	1:1000	SC-130776	Santa Cruz	39kDa
7	Cdk2	1:1000	78B2	Cell signaling	33kDa
8	Cyclin-D1	1:1000	2978	Cell signaling	36kDa
9	Cyclin-D3	1:2000	2936	Cell signaling	31kDa
10	Cdk6	1:2000	3136	Cell signaling	36kDa
11	P21	1:1000	2947	Cell signaling	21kDa
12	Beta-Actin	1:1000	3700	Cell signaling	45kDa
13	Anti-Mouse HRP linke	1:3000	7076S	Cell signaling	
14	Anti-Rabbit HRP linke	1:3000	7074S	Cell signaling	

Table: 14 Complte list of primary, secondary antibodies, its respective dilutions and moleular weight in kDa

Table 15

Hicup Read statistics

	ESC (Merged)		MEC (N) Merged		MEC (P) Merged		MEC (PRL) Merged	
	Read1	Read2	Read1	Read2	Read1	Read2	Read1	Read2
Total Reads	16,92,13,117	16,92,13,117	5,13,15,811	5,13,15,811	5,00,11,891	5,00,11,891	3,37,00,585	3,37,00,585
Not Truncated	15,37,62,943	15,38,91,774	2,93,94,822	2,97,31,498	3,60,65,699	3,61,33,944	1,95,73,342	1,97,79,191
Truncated	1,54,50,174	1,53,21,343	2,19,20,989	2,15,84,313	1,39,46,192	1,38,77,947	1,41,27,243	1,39,21,394
Too short to map	36,66,141	36,31,546	34,44,134	31,90,341	22,60,655	20,96,554	23,19,843	20,62,090
Average length of truncated	33.75	33.76	46.95	47.63	46.91	47.51	46.94	48.03
Unique Alignments	10,74,03,077	10,43,38,124	3,77,60,783	3,70,19,157	3,65,14,410	3,60,23,995	2,45,83,138	2,39,12,112
Multiple Alignments	3,45,89,905	3,36,39,825	87,23,009	84,58,907	91,84,502	90,18,480	57,81,165	55,04,446
Failed To Align	2,35,53,994	2,76,03,622	13,87,885	26,47,406	20,52,324	28,72,862	10,16,439	22,21,937
Paired	8,30,49,546	8,30,49,546	2,57,21,642	2,57,21,642	2,69,11,513	2,69,11,513	1,64,66,603	1,64,66,603
Valid Pairs	1,67,07,119		1,88,60,609		1,05,06,693		1,05,48,059	
Invalid Pairs	6,63,42,427		68,61,033		1,64,04,820		59,18,544	
Same Circularised	3,69,098		4,08,523		3,20,657		3,03,767	
Same Fragment Dangling End	67,89,468		1,18,263		11,58,922		1,41,730	
Same Fragment Internal	4,91,93,223		8,67,438		1,01,96,667		11,68,977	
Re-ligation	52,79,433		10,28,953		14,29,263		12,00,071	
Contiguous Sequence	2,84,454		86,543		80,298		83,375	
Wrong Size	44,26,751		43,51,313		32,19,013		30,20,624	
Total Pairs	8,30,49,546		2,57,21,642		2,69,11,513		1,64,66,603	
	All Di-Tags	**Unique Di-Tags**	**All Di-Tags**	**Unique Di-Tags**	**All Di-Tags**	**Unique Di-Tags**	**All Di-Tags**	**Unique Di-Tags**
Read Pairs	1,67,07,119	1,61,31,704	1,88,60,609	1,87,41,679	1,05,06,693	1,04,33,151	1,05,48,059	1,04,87,416
Cis-close (< 10Kbp)	32,76,149	31,53,345	29,23,930	29,04,852	18,99,571	18,86,101	21,37,326	21,24,371
Cis-far (> 10Kbp)	1,19,97,830	1,15,92,229	1,29,99,117	1,29,17,721	72,28,199	71,77,483	70,84,693	70,44,498
Trans	14,33,140	13,86,130	29,37,562	29,19,106	13,78,923	13,69,567	13,26,040	13,18,547

HC11 Translocaton Coordinates

No.	Traslocation Pair	1st Chr	Start	End	2nd Chr	Start	End
1	Chr1:6	chr1	100000	73000000	chr6	100000	148000000
2	Chr2:4	chr2	125000000	180000000	chr4	2000000	65000000
3	Chr2:8	chr2	52000000	180000000	chr8	26000000	127000000
4	Chr4:8	chr4	2000000	65000000	chr8	26000000	127000000
5	Chr5:11	chr5	131000000	150000000	chr11	42000000	120000000
6	Chr5:8	chr5	131000000	150000000	chr8	54000000	71000000
7	Chr8:11	chr8	55000000	71000000	chr11	2000000	120000000

Table:15 Hicup read statistics and HC11 chromosomal translocation co-ordinates

Table 16

Whole Genome raw 1Mb			
Sample 1	**Sample 2**	**Pearson Correlation**	**P_Value**
ESC1 Raw	ESC2 Raw	0.9946072	2.20E-16
N1 Raw	N2 Raw	0.9788713	2.20E-16
P1 Raw	P2 Raw	0.9926041	2.20E-16
Prl1 Raw	Prl2 Raw	0.9953005	2.20E-16

Whole Genome SimpleNorm (SN) 1Mb			
Sample 1	**Sample 2**	**Pearson Correlation**	**P_Value**
ESC1 SN	ESC2 SN	0.8160674	2.20E-16
N1 SN	N2 SN	0.742939	2.20E-16
P1 SN	P2 SN	0.788653	2.20E-16
Prl1 SN	Prl2 SN	0.7941851	2.20E-16

Whome Genome Raw 1Mb combined				
	ESC	**N**	**P**	**PRL**
ESC	1	0.946180251	0.96707818	0.94912153
N	0.946180251	1	0.9563313	0.98780221
P	0.967078178	0.956331296	1	0.97524151
PRL	0.949121532	0.987802214	0.97524151	1

Table: 16 Whole genome 1Mb correlations

Table 17		
Vanilla coverage WG 1MB		
Sample	**Spearmans correlation**	
ESC_N	0.181	
N_P	0.219	
P_PRL	0.185	
Chromosome Territory Correlations (p-values)		
Sample	**Person correlation**	**Spearmans correlation**
ESC_N	0.9204	0.835
N_P	0.9836	0.975
P_PRL	0.9936	0.988
Sample	**T-Test (p-values)**	
ESC_N	0.03	
N_P	0.004	
P_PRL	0.1116	
Sample	**Wilcoxon Test (p-values)**	
ESC_N	0.008	
N_P	0.05643	
P_PRL	0.85	

Table: 17 Correlations and P-values of Chromosome Negihbourhood Territories

Table 18

Chromosome 100kb matrix ESC-MEC (N)				
S No	**Variable1**	**Variable2**	**Correlation**	**P Value**
1	ESC_chr1	N_chr1	0.66322973	2.20E-16
2	ESC_chr2	N_chr2	0.68108148	2.20E-16
3	ESC_chr3	N_chr3	0.67091695	2.20E-16
4	ESC_chr4	N_chr4	0.68974983	2.20E-16
5	ESC_chr5	N_chr5	0.68303408	2.20E-16
6	ESC_chr6	N_chr6	0.67710787	2.20E-16
7	ESC_chr7	N_chr7	0.69200989	2.20E-16
8	ESC_chr8	N_chr8	0.69261428	2.20E-16
9	ESC_chr9	N_chr9	0.69535734	2.20E-16
10	ESC_chr10	N_chr10	0.68846025	2.20E-16
11	ESC_chr11	N_chr11	0.70758541	2.20E-16
12	ESC_chr12	N_chr12	0.7060029	2.20E-16
13	ESC_chr13	N_chr13	0.71537048	2.20E-16
14	ESC_chr14	N_chr14	0.69993925	2.20E-16
15	ESC_chr15	N_chr15	0.70683611	2.20E-16
16	ESC_chr16	N_chr16	0.69626768	2.20E-16
17	ESC_chr17	N_chr17	0.71365225	2.20E-16
18	ESC_chr18	N_chr18	0.69819366	2.20E-16
19	ESC_chr19	N_chr19	0.74676995	2.20E-16
20	ESC_chrX	N_chrX	0.65181552	2.20E-16

Chromosome 100kb matrix MEC (N) - (P)				
S No	**Variable1**	**Variable2**	**Correlation**	**P Value**
1	N_chr1	P_chr1	0.59546972	2.20E-16
2	N_chr2	P_chr2	0.62052709	2.20E-16
3	N_chr3	P_chr3	0.61773298	2.20E-16
4	N_chr4	P_chr4	0.63865358	2.20E-16
5	N_chr5	P_chr5	0.6235005	2.20E-16
6	N_chr6	P_chr6	0.61591382	2.20E-16
7	N_chr7	P_chr7	0.62221954	2.20E-16
8	N_chr8	P_chr8	0.65965209	2.20E-16
9	N_chr9	P_chr9	0.63523308	2.20E-16
10	N_chr10	P_chr10	0.62770166	2.20E-16
11	N_chr11	P_chr11	0.65697972	2.20E-16
12	N_chr12	P_chr12	0.65094954	2.20E-16
13	N_chr13	P_chr13	0.67239627	2.20E-16
14	N_chr14	P_chr14	0.64128406	2.20E-16
15	N_chr15	P_chr15	0.66306944	2.20E-16
16	N_chr16	P_chr16	0.66249034	2.20E-16
17	N_chr17	P_chr17	0.65912357	2.20E-16
18	N_chr18	P_chr18	0.63574754	2.20E-16
19	N_chr19	P_chr19	0.69245997	2.20E-16
20	N_chrX	P_chrX	0.58183262	2.20E-16

Chromosome 100kb MEC (P) - (PRL)				
S No	**Variable1**	**Variable2**	**Correlation**	**P Value**
1	P_chr1	Prl_chr1	0.56660031	2.20E-16
2	P_chr2	Prl_chr2	0.60050217	2.20E-16
3	P_chr3	Prl_chr3	0.59418246	2.20E-16
4	P_chr4	Prl_chr4	0.61464492	2.20E-16
5	P_chr5	Prl_chr5	0.6008954	2.20E-16
6	P_chr6	Prl_chr6	0.59090492	2.20E-16
7	P_chr7	Prl_chr7	0.59528959	2.20E-16
8	P_chr8	Prl_chr8	0.63358241	2.20E-16
9	P_chr9	Prl_chr9	0.61615021	2.20E-16
10	P_chr10	Prl_chr10	0.60894135	2.20E-16
11	P_chr11	Prl_chr11	0.63812747	2.20E-16
12	P_chr12	Prl_chr12	0.62687646	2.20E-16
13	P_chr13	Prl_chr13	0.66218929	2.20E-16
14	P_chr14	Prl_chr14	0.61231826	2.20E-16
15	P_chr15	Prl_chr15	0.64029848	2.20E-16
16	P_chr16	Prl_chr16	0.64766492	2.20E-16
17	P_chr17	Prl_chr17	0.64519093	2.20E-16
18	P_chr18	Prl_chr18	0.61639619	2.20E-16
19	P_chr19	Prl_chr19	0.67992125	2.20E-16
20	P_chrX	Prl_chrX	0.54659741	2.20E-16

Table: 18 Correlation and P-value stats of simple norm 100kb matrices of all chromosomes

Table 19						
HC11 MECs conditions	**Gene Names**	**Chromosome**	**Gene_Start**	**Gene_End**	**1Mb_Start**	**1Mb_End**
Normal	Etv4	chr11	101769748	101785310	101277529	102277529
Normal	Krt42	chr11	100262882	100269871	99766376.5	100766376.5
Primed	Wfdc18	chr11	83709004	83711360	83210182	84210182
Primed	Krt23	chr11	99477973	99493110	98985541.5	99985541.5
Prolactin	Csn2	chr5	87692619	87699425	87196022	88196022
Prolactin	Wap	chr11	6635483	6638649	6137066	7137066

Table: 19 Gene 1Mb domain co-ordinates used to generate gene locus interactions at 10Kb resolution

Table 20

TADTree parameters		
S	25	max. size of TAD (in bins)
M	10	max. number of TADs in each tad-tree
p	3	boundary index parameter
q	12	boundary index parameter
N	1 per size of the chr in Mb	Total TADs per chromosome

Relaxation parameters		
if TAD size <=500kb	relaxation is +/-100kb	
if TAD size > 500kb and <=1500kb	relaxation is +/-200kb	
if TAD size > 1500kb	relaxation is +/-300kb	

References

References:

ALLAR, M. A. & WOOD, T. L. 2004. Expression of the insulin-like growth factor binding proteins during postnatal development of the murine mammary gland. *Endocrinology,* 145**,** 2467-77.

ANDERS, S. & HUBER, W. 2010. Differential expression analysis for sequence count data. *Genome Biol,* 11**,** R106.

ANDERSEN, J. S., LYON, C. E., FOX, A. H., LEUNG, A. K., LAM, Y. W., STEEN, H., MANN, M. & LAMOND, A. I. 2002. Directed proteomic analysis of the human nucleolus. *Curr Biol,* 12**,** 1-11.

ANDREY, G., MONTAVON, T., MASCREZ, B., GONZALEZ, F., NOORDERMEER, D., LELEU, M., TRONO, D., SPITZ, F. & DUBOULE, D. 2013. A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science,* 340**,** 1234167.

ARUN, P. V., MIRYALA, S. K., CHATTOPADHYAY, S., THIYYAGURA, K., BAWA, P., BHATTACHARJEE, M. & YELLABOINA, S. 2016. Identification and functional analysis of essential, conserved, housekeeping and duplicated genes. *FEBS Lett,* 590**,** 1428-37.

ARVEY, A., AGIUS, P., NOBLE, W. S. & LESLIE, C. 2012. Sequence and chromatin determinants of cell-type-specific transcription factor binding. *Genome Res,* 22**,** 1723-34.

ASHBURNER, M., BALL, C. A., BLAKE, J. A., BOTSTEIN, D., BUTLER, H., CHERRY, J. M., DAVIS, A. P., DOLINSKI, K., DWIGHT, S. S., EPPIG, J. T., HARRIS, M. A., HILL, D. P., ISSEL-TARVER, L., KASARSKIS, A., LEWIS, S., MATESE, J. C., RICHARDSON, J. E., RINGWALD, M., RUBIN, G. M. & SHERLOCK, G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet,* 25**,** 25-9.

BALL, R. K., FRIIS, R. R., SCHOENENBERGER, C. A., DOPPLER, W. & GRONER, B. 1988. Prolactin regulation of beta-casein gene expression and of a cytosolic 120-kd protein in a cloned mouse mammary epithelial cell line. *EMBO J,* 7**,** 2089-95.

BALLESTER, M., KRESS, C., HUE-BEAUVAIS, C., KIEU, K., LEHMANN, G., ADENOT, P. & DEVINOY, E. 2008. The nuclear localization of WAP and CSN genes is modified by lactogenic hormones in HC11 cells. *J Cell Biochem,* 105**,** 262-70.

BANTIGNIES, F., ROURE, V., COMET, I., LEBLANC, B., SCHUETTENGRUBER, B., BONNET, J., TIXIER, V., MAS, A. & CAVALLI, G. 2011. Polycomb-dependent regulatory contacts between distant Hox loci in Drosophila. *Cell,* 144**,** 214-26.

BARANELLO, L., KOUZINE, F. & LEVENS, D. 2014. CTCF and cohesin cooperate to organize the 3D structure of the mammalian genome. *Proc Natl Acad Sci U S A,* 111**,** 889-90.

BARSKI, A., CUDDAPAH, S., CUI, K., ROH, T. Y., SCHONES, D. E., WANG, Z., WEI, G., CHEPELEV, I. & ZHAO, K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell,* 129**,** 823-37.

BARUTCU, A. R., LAJOIE, B. R., MCCORD, R. P., TYE, C. E., HONG, D., MESSIER, T. L., BROWNE, G., VAN WIJNEN, A. J., LIAN, J. B., STEIN, J. L., DEKKER, J., IMBALZANO, A. N. & STEIN, G. S. 2015. Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. *Genome Biol,* 16**,** 214.

BELL, A. C. & FELSENFELD, G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature,* 405**,** 482-5.

BERMEJO, R., CAPRA, T., JOSSEN, R., COLOSIO, A., FRATTINI, C., CAROTENUTO, W., COCITO, A., DOKSANI, Y., KLEIN, H., GOMEZ-GONZALEZ, B., AGUILERA, A., KATOU, Y., SHIRAHIGE, K. & FOIANI, M. 2011. The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores. *Cell,* 146**,** 233-46.

BERNARDI, R. & PANDOLFI, P. P. 2007. Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. *Nat Rev Mol Cell Biol,* 8**,** 1006-16.

BERNARDI, R., PAPA, A. & PANDOLFI, P. P. 2008. Regulation of apoptosis by PML and the PML-NBs. *Oncogene,* 27**,** 6299-312.

BERNSTEIN, B. E., STAMATOYANNOPOULOS, J. A., COSTELLO, J. F., REN, B., MILOSAVLJEVIC, A., MEISSNER, A., KELLIS, M., MARRA, M. A., BEAUDET, A. L., ECKER, J. R., FARNHAM, P. J., HIRST, M., LANDER, E. S., MIKKELSEN, T. S. & THOMSON, J. A. 2010. The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol,* 28**,** 1045-8.

BERRY, S., HARTLEY, M., OLSSON, T. S., DEAN, C. & HOWARD, M. 2015. Local chromatin environment of a Polycomb target gene instructs its own epigenetic inheritance. *Elife,* 4.

BIAN, Q., KHANNA, N., ALVIKAS, J. & BELMONT, A. S. 2013. beta-Globin cis-elements determine differential nuclear targeting through epigenetic modifications. *J Cell Biol,* 203**,** 767-83.

BIBEL, M., RICHTER, J., LACROIX, E. & BARDE, Y. A. 2007. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. *Nat Protoc,* 2**,** 1034-43.

BLASIUS, M., FORMENT, J. V., THAKKAR, N., WAGNER, S. A., CHOUDHARY, C. & JACKSON, S. P. 2011. A phospho-proteomic screen identifies substrates of the checkpoint kinase Chk1. *Genome Biol,* 12**,** R78.

BOLLAND, D. J., KING, M. R., REIK, W., CORCORAN, A. E. & KRUEGER, C. 2013. Robust 3D DNA FISH using directly labeled probes. *J Vis Exp*.

BONEV, B., MENDELSON COHEN, N., SZABO, Q., FRITSCH, L., PAPADOPOULOS, G. L., LUBLING, Y., XU, X., LV, X., HUGNOT, J. P., TANAY, A. & CAVALLI, G. 2017. Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell,* 171**,** 557-572 e24.

BOYA, R., YADAVALLI, A. D., NIKHAT, S., KURUKUTI, S., PALAKODETI, D. & PONGUBALA, J. M. R. 2017. Developmentally regulated higher-order chromatin interactions orchestrate B cell fate commitment. *Nucleic Acids Res,* 45**,** 11070-11087.

BOYLE, S., GILCHRIST, S., BRIDGER, J. M., MAHY, N. L., ELLIS, J. A. & BICKMORE, W. A. 2001. The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum Mol Genet,* 10**,** 211-9.

BOYLE, S., RODESCH, M. J., HALVENSLEBEN, H. A., JEDDELOH, J. A. & BICKMORE, W. A. 2011. Fluorescence in situ hybridization with high-complexity repeat-free oligonucleotide probes generated by massively parallel synthesis. *Chromosome Res,* 19**,** 901-9.

BRAGANCA, J., SWINGLER, T., MARQUES, F. I., JONES, T., ELORANTA, J. J., HURST, H. C., SHIODA, T. & BHATTACHARYA, S. 2002. Human CREB-binding protein/p300-interacting transactivator with ED-rich tail (CITED) 4, a new member of the CITED family, functions as a co-activator for transcription factor AP-2. *J Biol Chem,* 277**,** 8559-65.

BRICKNER, D. G. & BRICKNER, J. H. 2011. Gene positioning is regulated by phosphorylation of the nuclear pore complex by Cdk1. *Cell Cycle,* 10**,** 392-5.

BROWN, J. L., MUCCI, D., WHITELEY, M., DIRKSEN, M. L. & KASSIS, J. A. 1998. The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. *Mol Cell,* 1**,** 1057-64.

BUCHWALTER, A. L., LIANG, Y. & HETZER, M. W. 2014. Nup50 is required for cell differentiation and exhibits transcription-dependent dynamics. *Mol Biol Cell,* 25**,** 2472-84.

BUKATA, L., PARKER, S. L. & D'ANGELO, M. A. 2013. Nuclear pore complexes in the maintenance of genome integrity. *Curr Opin Cell Biol,* 25**,** 378-86.

BUSSLINGER, G. A., STOCSITS, R. R., VAN DER LELIJ, P., AXELSSON, E., TEDESCHI, A., GALJART, N. & PETERS, J. M. 2017. Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature,* 544**,** 503-507.

CAPELSON, M., LIANG, Y., SCHULTE, R., MAIR, W., WAGNER, U. & HETZER, M. W. 2010. Chromatin-bound nuclear pore components regulate gene expression in higher eukaryotes. *Cell,* 140**,** 372-83.

CARON, H., VAN SCHAIK, B., VAN DER MEE, M., BAAS, F., RIGGINS, G., VAN SLUIS, P., HERMUS, M. C., VAN ASPEREN, R., BOON, K., VOUTE, P. A., HEISTERKAMP, S., VAN KAMPEN, A. & VERSTEEG, R. 2001. The human transcriptome map: clustering of highly expressed genes in chromosomal domains. *Science,* 291**,** 1289-92.

CHAKRABARTI, R., HWANG, J., ANDRES BLANCO, M., WEI, Y., LUKACISIN, M., ROMANO, R. A., SMALLEY, K., LIU, S., YANG, Q., IBRAHIM, T., MERCATALI, L., AMADORI, D., HAFFTY, B. G., SINHA, S. & KANG, Y. 2012. Elf5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2. *Nat Cell Biol,* 14**,** 1212-22.

CHANDRA, T., EWELS, P. A., SCHOENFELDER, S., FURLAN-MAGARIL, M., WINGETT, S. W., KIRSCHNER, K., THURET, J. Y., ANDREWS, S., FRASER, P. & REIK, W. 2015. Global reorganization of the nuclear landscape in senescent cells. *Cell Rep,* 10**,** 471-83.

CHEN, L. L. & CARMICHAEL, G. G. 2008. Gene regulation by SINES and inosines: biological consequences of A-to-I editing of Alu element inverted repeats. *Cell Cycle,* 7**,** 3294-301.

CHEN, X., YAMMINE, S., SHI, C., TARK-DAME, M., GONDOR, A. & OHLSSON, R. 2014. The visualization of large organized chromatin domains enriched in the H3K9me2 mark within a single chromosome in a single cell. *Epigenetics,* 9**,** 1439-45.

CHOI, Y. S., CHAKRABARTI, R., ESCAMILLA-HERNANDEZ, R. & SINHA, S. 2009. Elf5 conditional knockout mice reveal its role as a master regulator in mammary alveolar development: failure of Stat5 activation and functional differentiation in the absence of Elf5. *Dev Biol,* 329**,** 227-41.

CHOUKRALLAH, M. A. & MATTHIAS, P. 2014. The Interplay between Chromatin and Transcription Factor Networks during B Cell Development: Who Pulls the Trigger First? *Front Immunol,* 5**,** 156.

CLEARD, F., MOSHKIN, Y., KARCH, F. & MAEDA, R. K. 2006. Probing long-distance regulatory interactions in the Drosophila melanogaster bithorax complex using Dam identification. *Nat Genet,* 38**,** 931-5.

CLEMSON, C. M., HUTCHINSON, J. N., SARA, S. A., ENSMINGER, A. W., FOX, A. H., CHESS, A. & LAWRENCE, J. B. 2009. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell,* 33**,** 717-26.

CMARKO, D., VERSCHURE, P. J., MARTIN, T. E., DAHMUS, M. E., KRAUSE, S., FU, X. D., VAN DRIEL, R. & FAKAN, S. 1999. Ultrastructural analysis of transcription and splicing in the cell nucleus after bromo-UTP microinjection. *Mol Biol Cell,* 10**,** 211-23.

COCOLAKIS, E., DAI, M., DREVET, L., HO, J., HAINES, E., ALI, S. & LEBRUN, J. J. 2008. Smad signaling antagonizes STAT5-mediated gene transcription and mammary epithelial cell differentiation. *J Biol Chem,* 283**,** 1293-307.

CREMER, M., VON HASE, J., VOLM, T., BRERO, A., KRETH, G., WALTER, J., FISCHER, C., SOLOVEI, I., CREMER, C. & CREMER, T. 2001. Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosome Res,* 9**,** 541-67.

CREMER, T., CREMER, C., BAUMANN, H., LUEDTKE, E. K., SPERLING, K., TEUBER, V. & ZORN, C. 1982. Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments. *Hum Genet,* 60**,** 46-56.

CRISCIONE, S. W., DE CECCO, M., SIRANOSIAN, B., ZHANG, Y., KREILING, J. A., SEDIVY, J. M. & NERETTI, N. 2016. Reorganization of chromosome architecture in replicative cellular senescence. *Sci Adv,* 2**,** e1500882.

CROFT, J. A., BRIDGER, J. M., BOYLE, S., PERRY, P., TEAGUE, P. & BICKMORE, W. A. 1999. Differences in the localization and morphology of chromosomes in the human nucleus. *J Cell Biol,* 145**,** 1119-31.

CUBENAS-POTTS, C. & CORCES, V. G. 2015. Topologically Associating Domains: An invariant framework or a dynamic scaffold? *Nucleus,* 6**,** 430-4.

DAVIS, K. R., GIESY, S. L., LONG, Q., KRUMM, C. S., HARVATINE, K. J. & BOISCLAIR, Y. R. 2016. XBP1 Regulates the Biosynthetic Capacity of the Mammary Gland During Lactation by Controlling Epithelial Expansion and Endoplasmic Reticulum Formation. *Endocrinology,* 157**,** 417-28.

DE LAAT, W. & GROSVELD, F. 2003. Spatial organization of gene expression: the active chromatin hub. *Chromosome Res,* 11**,** 447-59.

DE LAS HERAS, J. I., MEINKE, P., BATRAKOU, D. G., SRSEN, V., ZULEGER, N., KERR, A. R. & SCHIRMER, E. C. 2013. Tissue specificity in the nuclear envelope supports its functional complexity. *Nucleus,* 4**,** 460-77.

DE WIT, E., VOS, E. S., HOLWERDA, S. J., VALDES-QUEZADA, C., VERSTEGEN, M. J., TEUNISSEN, H., SPLINTER, E., WIJCHERS, P. J., KRIJGER, P. H. & DE LAAT, W. 2015. CTCF Binding Polarity Determines Chromatin Looping. *Mol Cell,* 60**,** 676-84.

DEKKER, J. & HEARD, E. 2015. Structural and functional diversity of Topologically Associating Domains. *FEBS Lett,* 589**,** 2877-84.

DEKKER, J., RIPPE, K., DEKKER, M. & KLECKNER, N. 2002. Capturing chromosome conformation. *Science,* 295**,** 1306-11.

DIETZEL, S., JAUCH, A., KIENLE, D., QU, G., HOLTGREVE-GREZ, H., EILS, R., MUNKEL, C., BITTNER, M., MELTZER, P. S., TRENT, J. M. & CREMER, T. 1998. Separate and variably shaped chromosome arm domains are disclosed by chromosome arm painting in human cell nuclei. *Chromosome Res,* 6**,** 25-33.

DIETZEL, S., SCHIEBEL, K., LITTLE, G., EDELMANN, P., RAPPOLD, G. A., EILS, R., CREMER, C. & CREMER, T. 1999. The 3D positioning of ANT2 and ANT3 genes within female X chromosome territories correlates with gene activity. *Exp Cell Res,* 252**,** 363-75.

DIXON, J. R., GORKIN, D. U. & REN, B. 2016. Chromatin Domains: The Unit of Chromosome Organization. *Mol Cell,* 62**,** 668-80.

DIXON, J. R., JUNG, I., SELVARAJ, S., SHEN, Y., ANTOSIEWICZ-BOURGET, J. E., LEE, A. Y., YE, Z., KIM, A., RAJAGOPAL, N., XIE, W., DIAO, Y., LIANG, J., ZHAO, H., LOBANENKOV, V. V., ECKER, J. R., THOMSON, J. A. & REN, B. 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature,* 518**,** 331-6.

DIXON, J. R., SELVARAJ, S., YUE, F., KIM, A., LI, Y., SHEN, Y., HU, M., LIU, J. S. & REN, B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature,* 485**,** 376-80.

DOMINSKI, Z., YANG, X. C., PURDY, M., WAGNER, E. J. & MARZLUFF, W. F. 2005. A CPSF-73 homologue is required for cell cycle progression but not cell growth and interacts with a protein having features of CPSF-100. *Mol Cell Biol,* 25**,** 1489-500.

DORIGO, B., SCHALCH, T., KULANGARA, A., DUDA, S., SCHROEDER, R. R. & RICHMOND, T. J. 2004. Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science,* 306**,** 1571-3.

DOSTIE, J., FERRAIUOLO, M., PAUSE, A., ADAM, S. A. & SONENBERG, N. 2000. A novel shuttling protein, 4E-T, mediates the nuclear import of the mRNA 5' cap-binding protein, eIF4E. *EMBO J,* 19**,** 3142-56.

DOWEN, J. M., FAN, Z. P., HNISZ, D., REN, G., ABRAHAM, B. J., ZHANG, L. N., WEINTRAUB, A. S., SCHUJIERS, J., LEE, T. I., ZHAO, K. & YOUNG, R. A. 2014. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell,* 159**,** 374-387.

ENDERLE, D., BEISEL, C., STADLER, M. B., GERSTUNG, M., ATHRI, P. & PARO, R. 2011. Polycomb preferentially targets stalled promoters of coding and noncoding transcripts. *Genome Res,* 21**,** 216-26.

ESKELAND, R., LEEB, M., GRIMES, G. R., KRESS, C., BOYLE, S., SPROUL, D., GILBERT, N., FAN, Y., SKOULTCHI, A. I., WUTZ, A. & BICKMORE, W. A. 2010. Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Mol Cell,* 38**,** 452-64.

FAKAN, S. 1994. Perichromatin fibrils are in situ forms of nascent transcripts. *Trends Cell Biol,* 4**,** 86-90.

FAVREAU, C., DUBOSCLARD, E., OSTLUND, C., VIGOUROUX, C., CAPEAU, J., WEHNERT, M., HIGUET, D., WORMAN, H. J., COURVALIN, J. C. & BUENDIA, B. 2003. Expression of lamin A mutated in the carboxyl-terminal tail generates an aberrant nuclear phenotype similar to that observed in cells from patients with Dunnigan-type partial lipodystrophy and Emery-Dreifuss muscular dystrophy. *Exp Cell Res,* 282**,** 14-23.

FAWCETT, D. W. 1966. On the occurrence of a fibrous lamina on the inner aspect of the nuclear envelope in certain cells of vertebrates. *Am J Anat,* 119**,** 129-45.

FAZZIO, T. G., HUFF, J. T. & PANNING, B. 2008. Chromatin regulation Tip(60)s the balance in embryonic stem cell self-renewal. *Cell Cycle,* 7**,** 3302-6.

FEDORIW, A. M., STEIN, P., SVOBODA, P., SCHULTZ, R. M. & BARTOLOMEI, M. S. 2004. Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. *Science,* 303**,** 238-40.

FERRARI, N., MCDONALD, L., MORRIS, J. S., CAMERON, E. R. & BLYTH, K. 2013. RUNX2 in mammary gland development and breast cancer. *J Cell Physiol,* 228**,** 1137-42.

FINLAN, L. E., SPROUL, D., THOMSON, I., BOYLE, S., KERR, E., PERRY, P., YLSTRA, B., CHUBB, J. R. & BICKMORE, W. A. 2008. Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet,* 4**,** e1000039.

FLAVAHAN, W. A., DRIER, Y., LIAU, B. B., GILLESPIE, S. M., VENTEICHER, A. S., STEMMER-RACHAMIMOV, A. O., SUVA, M. L. & BERNSTEIN, B. E. 2016. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature,* 529**,** 110-4.

FORTIN, J. P. & HANSEN, K. D. 2015. Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome Biol,* 16**,** 180.

FOX, A. H., LAM, Y. W., LEUNG, A. K., LYON, C. E., ANDERSEN, J., MANN, M. & LAMOND, A. I. 2002. Paraspeckles: a novel nuclear domain. *Curr Biol,* 12**,** 13-25.

FOX, A. H. & LAMOND, A. I. 2010. Paraspeckles. *Cold Spring Harb Perspect Biol,* 2**,** a000687.

FRASER, J., FERRAI, C., CHIARIELLO, A. M., SCHUELER, M., RITO, T., LAUDANNO, G., BARBIERI, M., MOORE, B. L., KRAEMER, D. C., AITKEN, S., XIE, S. Q., MORRIS, K. J., ITOH, M., KAWAJI, H., JAEGER, I., HAYASHIZAKI, Y., CARNINCI, P., FORREST, A. R., CONSORTIUM, F., SEMPLE, C. A., DOSTIE, J., POMBO, A. & NICODEMI, M. 2015a. Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. *Mol Syst Biol,* 11**,** 852.

FRASER, J., WILLIAMSON, I., BICKMORE, W. A. & DOSTIE, J. 2015b. An Overview of Genome Organization and How We Got There: from FISH to Hi-C. *Microbiol Mol Biol Rev,* 79**,** 347-72.

FUDENBERG, G., IMAKAEV, M., LU, C., GOLOBORODKO, A., ABDENNUR, N. & MIRNY, L. A. 2016. Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep,* 15**,** 2038-49.

GALIOVA, G., BARTOVA, E., RASKA, I., KREJCI, J. & KOZUBEK, S. 2008. Chromatin changes induced by lamin A/C deficiency and the histone deacetylase inhibitor trichostatin A. *Eur J Cell Biol,* 87**,** 291-303.

GARD, S., LIGHT, W., XIONG, B., BOSE, T., MCNAIRN, A. J., HARRIS, B., FLEHARTY, B., SEIDEL, C., BRICKNER, J. H. & GERTON, J. L. 2009. Cohesinopathy mutations disrupt the subnuclear organization of chromatin. *J Cell Biol,* 187**,** 455-62.

GENDLER, K., PAULSEN, T. & NAPOLI, C. 2008. ChromDB: the chromatin database. *Nucleic Acids Res,* 36**,** D298-302.

GESSON, K., RESCHENEDER, P., SKORUPPA, M. P., VON HAESELER, A., DECHAT, T. & FOISNER, R. 2016. A-type lamins bind both hetero- and euchromatin, the latter being regulated by lamina-associated polypeptide 2 alpha. *Genome Res,* 26**,** 462-73.

GESSON, K., VIDAK, S. & FOISNER, R. 2014. Lamina-associated polypeptide (LAP)2alpha and nucleoplasmic lamins in adult stem cell regulation and disease. *Semin Cell Dev Biol,* 29**,** 116-24.

GIBCUS, J. H. & DEKKER, J. 2013. The hierarchy of the 3D genome. *Mol Cell,* 49**,** 773-82.

GIBCUS, J. H., SAMEJIMA, K., GOLOBORODKO, A., SAMEJIMA, I., NAUMOVA, N., NUEBLER, J., KANEMAKI, M. T., XIE, L., PAULSON, J. R., EARNSHAW, W. C., MIRNY, L. A. & DEKKER, J. 2018. A pathway for mitotic chromosome formation. *Science,* 359.

GIERMAN, H. J., INDEMANS, M. H., KOSTER, J., GOETZE, S., SEPPEN, J., GEERTS, D., VAN DRIEL, R. & VERSTEEG, R. 2007. Domain-wide regulation of gene expression in the human genome. *Genome Res,* 17**,** 1286-95.

GOLDMAN, R. D., SHUMAKER, D. K., ERDOS, M. R., ERIKSSON, M., GOLDMAN, A. E., GORDON, L. B., GRUENBAUM, Y., KHUON, S., MENDEZ, M., VARGA, R. & COLLINS, F. S. 2004. Accumulation of mutant lamin A causes progressive changes in nuclear architecture in Hutchinson-Gilford progeria syndrome. *Proc Natl Acad Sci U S A,* 101**,** 8963-8.

GOLOBORODKO, A., IMAKAEV, M. V., MARKO, J. F. & MIRNY, L. 2016. Compaction and segregation of sister chromatids via active loop extrusion. *Elife,* 5.

GOMEZ-MARIN, C., TENA, J. J., ACEMEL, R. D., LOPEZ-MAYORGA, M., NARANJO, S., DE LA CALLE-MUSTIENES, E., MAESO, I., BECCARI, L., ANEAS, I., VIELMAS, E., BOVOLENTA, P., NOBREGA, M. A., CARVAJAL, J. & GOMEZ-SKARMETA, J. L. 2015. Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. *Proc Natl Acad Sci U S A,* 112**,** 7542-7.

GONZALEZ-SANDOVAL, A., TOWBIN, B. D., KALCK, V., CABIANCA, D. S., GAIDATZIS, D., HAUER, M. H., GENG, L., WANG, L., YANG, T., WANG, X., ZHAO, K. & GASSER, S. M. 2015. Perinuclear Anchoring of H3K9-Methylated Chromatin Stabilizes Induced Cell Fate in C. elegans Embryos. *Cell,* 163**,** 1333-47.

GRIFFIS, E. R., ALTAN, N., LIPPINCOTT-SCHWARTZ, J. & POWERS, M. A. 2002. Nup98 is a mobile nucleoporin with transcription-dependent dynamics. *Mol Biol Cell,* 13**,** 1282-97.

GUELEN, L., PAGIE, L., BRASSET, E., MEULEMAN, W., FAZA, M. B., TALHOUT, W., EUSSEN, B. H., DE KLEIN, A., WESSELS, L., DE LAAT, W. & VAN STEENSEL, B. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature,* 453**,** 948-51.

GUO, C., YOON, H. S., FRANKLIN, A., JAIN, S., EBERT, A., CHENG, H. L., HANSEN, E., DESPO, O., BOSSEN, C., VETTERMANN, C., BATES, J. G., RICHARDS, N., MYERS, D., PATEL, H., GALLAGHER, M., SCHLISSEL, M. S., MURRE, C., BUSSLINGER, M., GIALLOURAKIS, C. C. & ALT, F. W. 2011. CTCF-binding elements mediate control of V(D)J recombination. *Nature,* 477**,** 424-30.

GUO, Y., XU, Q., CANZIO, D., SHOU, J., LI, J., GORKIN, D. U., JUNG, I., WU, H., ZHAI, Y., TANG, Y., LU, Y., WU, Y., JIA, Z., LI, W., ZHANG, M. Q., REN, B., KRAINER, A. R., MANIATIS, T. & WU, Q. 2015. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell,* 162**,** 900-10.

HABERMANN, F. A., CREMER, M., WALTER, J., KRETH, G., VON HASE, J., BAUER, K., WIENBERG, J., CREMER, C., CREMER, T. & SOLOVEI, I. 2001. Arrangements of macro- and microchromosomes in chicken cells. *Chromosome Res,* 9**,** 569-84.

HAILESELLASSE SENE, K., PORTER, C. J., PALIDWOR, G., PEREZ-IRATXETA, C., MURO, E. M., CAMPBELL, P. A., RUDNICKI, M. A. & ANDRADE-NAVARRO, M. A. 2007. Gene function in early mouse embryonic stem cell differentiation. *BMC Genomics,* 8**,** 85.

HAO LE, T., FULLER, H. R., LAM LE, T., LE, T. T., BURGHES, A. H. & MORRIS, G. E. 2007. Absence of gemin5 from SMN complexes in nuclear Cajal bodies. *BMC Cell Biol,* 8**,** 28.

HAREWOOD, L., KISHORE, K., ELDRIDGE, M. D., WINGETT, S., PEARSON, D., SCHOENFELDER, S., COLLINS, V. P. & FRASER, P. 2017. Hi-C as a tool for precise detection and characterisation of chromosomal rearrangements and copy number variation in human tumours. *Genome Biol,* 18**,** 125.

HARK, A. T., SCHOENHERR, C. J., KATZ, D. J., INGRAM, R. S., LEVORSE, J. M. & TILGHMAN, S. M. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature,* 405**,** 486-9.

HARRIS, B., BOSE, T., LEE, K. K., WANG, F., LU, S., ROSS, R. T., ZHANG, Y., FRENCH, S. L., BEYER, A. L., SLAUGHTER, B. D., UNRUH, J. R. & GERTON, J. L. 2014. Cohesion promotes nucleolar structure and function. *Mol Biol Cell,* 25**,** 337-46.

HAWKINS, R. D., HON, G. C., LEE, L. K., NGO, Q., LISTER, R., PELIZZOLA, M., EDSALL, L. E., KUAN, S., LUU, Y., KLUGMAN, S., ANTOSIEWICZ-BOURGET, J., YE, Z., ESPINOZA, C., AGARWAHL, S., SHEN, L., RUOTTI, V., WANG, W., STEWART, R., THOMSON, J. A., ECKER, J. R. & REN, B. 2010. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell,* 6**,** 479-91.

HEATH, H., RIBEIRO DE ALMEIDA, C., SLEUTELS, F., DINGJAN, G., VAN DE NOBELEN, S., JONKERS, I., LING, K. W., GRIBNAU, J., RENKAWITZ, R., GROSVELD, F., HENDRIKS, R. W. & GALJART, N. 2008. CTCF regulates cell cycle progression of alphabeta T cells in the thymus. *EMBO J,* 27**,** 2839-50.

HEINZ, S., BENNER, C., SPANN, N., BERTOLINO, E., LIN, Y. C., LASLO, P., CHENG, J. X., MURRE, C., SINGH, H. & GLASS, C. K. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell,* 38**,** 576-89.

HERNANDEZ-VERDUN, D., ROUSSEL, P., THIRY, M., SIRRI, V. & LAFONTAINE, D. L. 2010. The nucleolus: structure/function relationship in RNA metabolism. *Wiley Interdiscip Rev RNA,* 1**,** 415-31.

HIRAYAMA, T., TARUSAWA, E., YOSHIMURA, Y., GALJART, N. & YAGI, T. 2012. CTCF is required for neural development and stochastic expression of clustered Pcdh genes in neurons. *Cell Rep,* 2**,** 345-57.

HNISZ, D., WEINTRAUB, A. S., DAY, D. S., VALTON, A. L., BAK, R. O., LI, C. H., GOLDMANN, J., LAJOIE, B. R., FAN, Z. P., SIGOVA, A. A., REDDY, J., BORGES-RIVERA, D., LEE, T. I., JAENISCH, R., PORTEUS, M. H., DEKKER, J. & YOUNG, R. A. 2016. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science,* 351**,** 1454-1458.

HOLLENHORST, P. C., PAUL, L., FERRIS, M. W. & GRAVES, B. J. 2011. The ETS gene ETV4 is required for anchorage-independent growth and a cell proliferation gene expression program in PC3 prostate cells. *Genes Cancer,* 1**,** 1044-1052.

HOLMQUIST, G. P. 1992. Chromosome bands, their chromatin flavors, and their functional features. *Am J Hum Genet,* 51**,** 17-37.

HOZAK, P., HASSAN, A. B., JACKSON, D. A. & COOK, P. R. 1993. Visualization of replication factories attached to nucleoskeleton. *Cell,* 73**,** 361-73.

HUANG DA, W., SHERMAN, B. T. & LEMPICKI, R. A. 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res,* 37**,** 1-13.

HURST, L. D., PAL, C. & LERCHER, M. J. 2004. The evolutionary dynamics of eukaryotic gene order. *Nat Rev Genet,* 5**,** 299-310.

HUTCHINSON, J. N., ENSMINGER, A. W., CLEMSON, C. M., LYNCH, C. R., LAWRENCE, J. B. & CHESS, A. 2007. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics,* 8**,** 39.

HYNES, N. E. & WATSON, C. J. 2010. Mammary gland growth factors: roles in normal development and in cancer. *Cold Spring Harb Perspect Biol,* 2**,** a003186.

IBARRA, A. & HETZER, M. W. 2015. Nuclear pore proteins and the control of genome functions. *Genes Dev,* 29**,** 337-49.

IBORRA, F. J., POMBO, A., JACKSON, D. A. & COOK, P. R. 1996. Active RNA polymerases are localized within discrete transcription "factories' in human nuclei. *J Cell Sci,* 109 (Pt 6)**,** 1427-36.

IDEKER, T., OZIER, O., SCHWIKOWSKI, B. & SIEGEL, A. F. 2002. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics,* 18 Suppl 1**,** S233-40.

JABBARI, K. & BERNARDI, G. 2017. An Isochore Framework Underlies Chromatin Architecture. *PLoS One,* 12**,** e0168023.

JAGANNATHAN, M., CUMMINGS, R. & YAMASHITA, Y. M. 2018. A conserved function for pericentromeric satellite DNA. *Elife,* 7.

JAGATHEESAN, G., THANUMALAYAN, S., MURALIKRISHNA, B., RANGARAJ, N., KARANDE, A. A. & PARNAIK, V. K. 1999. Colocalization of intranuclear lamin foci with RNA splicing factors. *J Cell Sci,* 112 (Pt 24)**,** 4651-61.

JOHN, S., SABO, P. J., JOHNSON, T. A., SUNG, M. H., BIDDIE, S. C., LIGHTMAN, S. L., VOSS, T. C., DAVIS, S. R., MELTZER, P. S., STAMATOYANNOPOULOS, J. A. & HAGER, G. L. 2008. Interaction of the glucocorticoid receptor with the chromatin landscape. *Mol Cell,* 29**,** 611-24.

JUBB, A. W., BOYLE, S., HUME, D. A. & BICKMORE, W. A. 2017. Glucocorticoid Receptor Binding Induces Rapid and Prolonged Large-Scale Chromatin Decompaction at Multiple Target Loci. *Cell Rep,* 21**,** 3022-3031.

KABOTYANSKI, E. B., HUETTER, M., XIAN, W., RIJNKELS, M. & ROSEN, J. M. 2006. Integration of prolactin and glucocorticoid signaling at the beta-casein promoter and enhancer by ordered recruitment of specific transcription factors and chromatin modifiers. *Mol Endocrinol,* 20**,** 2355-68.

KADMIEL, M. & CIDLOWSKI, J. A. 2013. Glucocorticoid receptor signaling in health and disease. *Trends Pharmacol Sci,* 34**,** 518-30.

KALVERDA, B., PICKERSGILL, H., SHLOMA, V. V. & FORNEROD, M. 2010. Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm. *Cell,* 140**,** 360-71.

KATAINEN, R., DAVE, K., PITKANEN, E., PALIN, K., KIVIOJA, T., VALIMAKI, N., GYLFE, A. E., RISTOLAINEN, H., HANNINEN, U. A., CAJUSO, T., KONDELIN, J., TANSKANEN, T., MECKLIN, J. P., JARVINEN, H., RENKONEN-SINISALO, L., LEPISTO, A., KAASINEN, E., KILPIVAARA, O., TUUPANEN, S., ENGE, M., TAIPALE, J. & AALTONEN, L. A. 2015. CTCF/cohesin-binding sites are frequently mutated in cancer. *Nat Genet,* 47**,** 818-21.

KELLY, P. A., BACHELOT, A., KEDZIA, C., HENNIGHAUSEN, L., ORMANDY, C. J., KOPCHICK, J. J. & BINART, N. 2002. The role of prolactin and growth hormone in mammary gland development. *Mol Cell Endocrinol,* 197**,** 127-31.

KIEFFER-KWON, K. R., NIMURA, K., RAO, S. S. P., XU, J., JUNG, S., PEKOWSKA, A., DOSE, M., STEVENS, E., MATHE, E., DONG, P., HUANG, S. C., RICCI, M. A., BARANELLO, L., ZHENG, Y., TOMASSONI ARDORI, F., RESCH, W., STAVREVA, D., NELSON, S., MCANDREW, M., CASELLAS, A., FINN, E., GREGORY, C., ST HILAIRE, B. G., JOHNSON, S. M., DUBOIS, W., COSMA, M. P., BATCHELOR, E., LEVENS, D., PHAIR, R. D., MISTELI, T., TESSAROLLO, L., HAGER, G., LAKADAMYALI, M., LIU, Z., FLOER, M., SHROFF, H., AIDEN, E. L. & CASELLAS, R. 2017. Myc Regulates Chromatin Decompaction and Nuclear Architecture during B Cell Activation. *Mol Cell,* 67**,** 566-578 e10.

KIM, D., PERTEA, G., TRAPNELL, C., PIMENTEL, H., KELLEY, R. & SALZBERG, S. L. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol,* 14**,** R36.

KIM, T. H., ABDULLAEV, Z. K., SMITH, A. D., CHING, K. A., LOUKINOV, D. I., GREEN, R. D., ZHANG, M. Q., LOBANENKOV, V. V. & REN, B. 2007. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell,* 128**,** 1231-45.

KIND, J. & VAN STEENSEL, B. 2014. Stochastic genome-nuclear lamina interactions: modulating roles of Lamin A and BAF. *Nucleus,* 5**,** 124-30.

KOEHLER, D., ZAKHARTCHENKO, V., FROENICKE, L., STONE, G., STANYON, R., WOLF, E., CREMER, T. & BRERO, A. 2009. Changes of higher order chromatin arrangements during major genome activation in bovine preimplantation embryos. *Exp Cell Res,* 315**,** 2053-63.

KOHWI, M., LUPTON, J. R., LAI, S. L., MILLER, M. R. & DOE, C. Q. 2013. Developmentally regulated subnuclear genome reorganization restricts neural progenitor competence in Drosophila. *Cell,* 152**,** 97-108.

KOKEN, M. H., PUVION-DUTILLEUL, F., GUILLEMIN, M. C., VIRON, A., LINARES-CRUZ, G., STUURMAN, N., DE JONG, L., SZOSTECKI, C., CALVO, F., CHOMIENNE, C. & ET AL. 1994. The t(15;17) translocation alters a nuclear body in a retinoic acid-reversible fashion. *EMBO J,* 13**,** 1073-83.

KOLEV, N. G. & STEITZ, J. A. 2005. Symplekin and multiple other polyadenylation factors participate in 3'-end maturation of histone mRNAs. *Genes Dev,* 19**,** 2583-92.

KOSAK, S. T., SCALZO, D., ALWORTH, S. V., LI, F., PALMER, S., ENVER, T., LEE, J. S. & GROUDINE, M. 2007. Coordinate gene regulation during hematopoiesis is related to genomic organization. *PLoS Biol,* 5**,** e309.

KOUROS-MEHR, H., SLORACH, E. M., STERNLICHT, M. D. & WERB, Z. 2006. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. *Cell,* 127**,** 1041-55.

KOZUBEK, M., SKALNIKOVA, M., MATULA, P., BARTOVA, E., RAUCH, J., NEUHAUS, F., EIPEL, H. & HAUSMANN, M. 2002. Automated microaxial tomography of cell nuclei after specific labelling by fluorescence in situ hybridisation. *Micron,* 33**,** 655-65.

KRAUSE, S., FAKAN, S., WEIS, K. & WAHLE, E. 1994. Immunodetection of poly(A) binding protein II in the cell nucleus. *Exp Cell Res,* 214**,** 75-82.

KRESS, C., BALLESTER, M., DEVINOY, E. & RIJNKELS, M. 2010. Epigenetic modifications in 3D: nuclear organization of the differentiating mammary epithelial cell. *J Mammary Gland Biol Neoplasia,* 15**,** 73-83.

KRIVEGA, I., DALE, R. K. & DEAN, A. 2014. Role of LDB1 in the transition from chromatin looping to transcription activation. *Genes Dev,* 28**,** 1278-90.

KRIVEGA, I. & DEAN, A. 2012. Enhancer and promoter interactions-long distance calls. *Curr Opin Genet Dev,* 22**,** 79-85.

KURODA, M., TANABE, H., YOSHIDA, K., OIKAWA, K., SAITO, A., KIYUNA, T., MIZUSAWA, H. & MUKAI, K. 2004. Alteration of chromosome positioning during adipocyte differentiation. *J Cell Sci,* 117**,** 5897-903.

KURPIOS, N. A., MACNEIL, L., SHEPHERD, T. G., GLUDISH, D. W., GIACOMELLI, A. O. & HASSELL, J. A. 2009. The Pea3 Ets transcription factor regulates differentiation of multipotent progenitor cells during mammary gland development. *Dev Biol,* 325**,** 106-21.

KURUKUTI, S., TIWARI, V. K., TAVOOSIDANA, G., PUGACHEVA, E., MURRELL, A., ZHAO, Z., LOBANENKOV, V., REIK, W. & OHLSSON, R. 2006. CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc Natl Acad Sci U S A,* 103**,** 10684-9.

KUZMICHEV, A., JENUWEIN, T., TEMPST, P. & REINBERG, D. 2004. Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. *Mol Cell,* 14**,** 183-93.

KUZNETSOVA, I. S., OSTROMYSHENSKII, D. I., KOMISSAROV, A. S., PRUSOV, A. N., WAISERTREIGER, I. S., GORBUNOVA, A. V., TRIFONOV, V. A., FERGUSON-SMITH, M. A. & PODGORNAYA, O. I. 2016. LINE-related component of mouse heterochromatin and complex chromocenters' composition. *Chromosome Res,* 24**,** 309-23.

KUZNETSOVA, T., WANG, S. Y., RAO, N. A., MANDOLI, A., MARTENS, J. H., ROTHER, N., AARTSE, A., GROH, L., JANSSEN-MEGENS, E. M., LI, G., RUAN, Y., LOGIE, C. & STUNNENBERG, H. G. 2015. Glucocorticoid receptor and nuclear factor kappa-b affect three-dimensional chromatin organization. *Genome Biol,* 16**,** 264.

LANZUOLO, C., ROURE, V., DEKKER, J., BANTIGNIES, F. & ORLANDO, V. 2007. Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat Cell Biol,* 9**,** 1167-74.

LARSSON, S. H., CHARLIEU, J. P., MIYAGAWA, K., ENGELKAMP, D., RASSOULZADEGAN, M., ROSS, A., CUZIN, F., VAN HEYNINGEN, V. & HASTIE, N. D. 1995. Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing. *Cell,* 81**,** 391-401.

LERCHER, M. J., URRUTIA, A. O. & HURST, L. D. 2002. Clustering of housekeeping genes provides a unified model of gene order in the human genome. *Nat Genet,* 31**,** 180-3.

LEVANON, K., EISENBERG, E., RECHAVI, G. & LEVANON, E. Y. 2005. Letter from the editor: Adenosine-to-inosine RNA editing in Alu repeats in the human genome. *EMBO Rep,* 6**,** 831-5.

LI, Y., KANE, T., TIPPER, C., SPATRICK, P. & JENNESS, D. D. 1999. Yeast mutants affecting possible quality control of plasma membrane proteins. *Mol Cell Biol,* 19**,** 3588-99.

LIBERZON, A., SUBRAMANIAN, A., PINCHBACK, R., THORVALDSDOTTIR, H., TAMAYO, P. & MESIROV, J. P. 2011. Molecular signatures database (MSigDB) 3.0. *Bioinformatics,* 27**,** 1739-40.

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., SANDSTROM, R., BERNSTEIN, B., BENDER, M. A., GROUDINE, M., GNIRKE, A., STAMATOYANNOPOULOS, J., MIRNY, L. A., LANDER, E. S. & DEKKER, J. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science,* 326**,** 289-93.

LIN, F. & WORMAN, H. J. 1993. Structural organization of the human gene encoding nuclear lamin A and nuclear lamin C. *J Biol Chem,* 268**,** 16321-6.

LING, J. Q., LI, T., HU, J. F., VU, T. H., CHEN, H. L., QIU, X. W., CHERRY, A. M. & HOFFMAN, A. R. 2006. CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science,* 312**,** 269-72.

LIU, Y., ZHOU, J. & WHITE, K. P. 2014. RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics,* 30**,** 301-4.

LIVAK, K. J. & 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**,** 402-8.

LOH, K. M., CHEN, A., KOH, P. W., DENG, T. Z., SINHA, R., TSAI, J. M., BARKAL, A. A., SHEN, K. Y., JAIN, R., MORGANTI, R. M., SHYH-CHANG, N., FERNHOFF, N. B., GEORGE, B. M., WERNIG, G., SALOMON, R. E. A., CHEN, Z., VOGEL, H., EPSTEIN, J. A., KUNDAJE, A., TALBOT, W. S., BEACHY, P. A., ANG, L. T. & WEISSMAN, I. L. 2016. Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types. *Cell,* 166**,** 451-467.

LOMBARD, D. B. & GUARENTE, L. 2000. Nijmegen breakage syndrome disease protein and MRE11 at PML nuclear bodies and meiotic telomeres. *Cancer Res,* 60**,** 2331-4.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**,** 550.

LUND, E. G., DUBAND-GOULET, I., OLDENBURG, A., BUENDIA, B. & COLLAS, P. 2015. Distinct features of lamin A-interacting chromatin domains mapped by ChIP-sequencing from sonicated or micrococcal nuclease-digested chromatin. *Nucleus,* 6**,** 30-9.

MACIAS, H. & HINCK, L. 2012. Mammary gland development. *Wiley Interdiscip Rev Dev Biol,* 1**,** 533-57.

MAIS, C., WRIGHT, J. E., PRIETO, J. L., RAGGETT, S. L. & MCSTAY, B. 2005. UBF-binding site arrays form pseudo-NORs and sequester the RNA polymerase I transcription machinery. *Genes Dev,* 19**,** 50-64.

MALHOTRA, G. K., ZHAO, X., EDWARDS, E., KOPP, J. L., NARAMURA, M., SANDER, M., BAND, H. & BAND, V. 2014. The role of Sox9 in mouse mammary gland development and maintenance of mammary stem and luminal progenitor cells. *BMC Dev Biol,* 14**,** 47.

MAO, Y. S., ZHANG, B. & SPECTOR, D. L. 2011. Biogenesis and function of nuclear bodies. *Trends Genet,* 27**,** 295-306.

MARGUERON, R., LI, G., SARMA, K., BLAIS, A., ZAVADIL, J., WOODCOCK, C. L., DYNLACHT, B. D. & REINBERG, D. 2008. Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell,* 32**,** 503-18.

MARTENS, J. H., O'SULLIVAN, R. J., BRAUNSCHWEIG, U., OPRAVIL, S., RADOLF, M., STEINLEIN, P. & JENUWEIN, T. 2005. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J,* 24**,** 800-12.

MARZLUFF, W. F., WAGNER, E. J. & DURONIO, R. J. 2008. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet,* 9**,** 843-54.

MATSUOKA, S., BALLIF, B. A., SMOGORZEWSKA, A., MCDONALD, E. R., 3RD, HUROV, K. E., LUO, J., BAKALARSKI, C. E., ZHAO, Z., SOLIMINI, N., LERENTHAL, Y., SHILOH, Y., GYGI, S. P. & ELLEDGE, S. J. 2007. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science,* 316**,** 1160-6.

MATTOUT, A., CABIANCA, D. S. & GASSER, S. M. 2015. Chromatin states and nuclear organization in development--a view from the nuclear lamina. *Genome Biol,* 16**,** 174.

MAYER, R., BRERO, A., VON HASE, J., SCHROEDER, T., CREMER, T. & DIETZEL, S. 2005. Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC Cell Biol,* 6**,** 44.

MERLO, G. R., GRAUS-PORTA, D., CELLA, N., MARTE, B. M., TAVERNA, D. & HYNES, N. E. 1996. Growth, differentiation and survival of HC11 mammary epithelial cells: diverse effects of receptor tyrosine kinase-activating peptide growth factors. *Eur J Cell Biol,* 70**,** 97-105.

MEULEMAN, W., PERIC-HUPKES, D., KIND, J., BEAUDRY, J. B., PAGIE, L., KELLIS, M., REINDERS, M., WESSELS, L. & VAN STEENSEL, B. 2013. Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res,* 23**,** 270-80.

MORRIS, G. E. 2008. The Cajal body. *Biochim Biophys Acta,* 1783**,** 2108-15.

MORRISON, B. & CUTLER, M. L. 2009. Mouse Mammary Epithelial Cells form Mammospheres During Lactogenic Differentiation. *J Vis Exp*.

NAGANO, T., LUBLING, Y., VARNAI, C., DUDLEY, C., LEUNG, W., BARAN, Y., MENDELSON COHEN, N., WINGETT, S., FRASER, P. & TANAY, A. 2017. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature,* 547**,** 61-67.

NAGASAKA, K., HOSSAIN, M. J., ROBERTI, M. J., ELLENBERG, J. & HIROTA, T. 2016. Sister chromatid resolution is an intrinsic part of chromosome organization in prophase. *Nat Cell Biol,* 18**,** 692-9.

NARENDRA, V., ROCHA, P. P., AN, D., RAVIRAM, R., SKOK, J. A., MAZZONI, E. O. & REINBERG, D. 2015. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science,* 347**,** 1017-21.

NAUMOVA, N., IMAKAEV, M., FUDENBERG, G., ZHAN, Y., LAJOIE, B. R., MIRNY, L. A. & DEKKER, J. 2013. Organization of the mitotic chromosome. *Science,* 342**,** 948-53.

NEUSSER, M., SCHUBEL, V., KOCH, A., CREMER, T. & MULLER, S. 2007. Evolutionarily conserved, cell type and species-specific higher order chromatin arrangements in interphase nuclei of primates. *Chromosoma,* 116**,** 307-20.

NISOLE, S., MAROUI, M. A., MASCLE, X. H., AUBRY, M. & CHELBI-ALIX, M. K. 2013. Differential Roles of PML Isoforms. *Front Oncol,* 3**,** 125.

NORA, E. P., GOLOBORODKO, A., VALTON, A. L., GIBCUS, J. H., UEBERSOHN, A., ABDENNUR, N., DEKKER, J., MIRNY, L. A. & BRUNEAU, B. G. 2017. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell,* 169**,** 930-944 e22.

NORA, E. P., LAJOIE, B. R., SCHULZ, E. G., GIORGETTI, L., OKAMOTO, I., SERVANT, N., PIOLOT, T., VAN BERKUM, N. L., MEISIG, J., SEDAT, J., GRIBNAU, J., BARILLOT, E., BLUTHGEN, N., DEKKER, J. & HEARD, E. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature,* 485**,** 381-5.

OOGA, M., FUNAYA, S., HASHIOKA, Y., FUJII, W., NAITO, K., SUZUKI, M. G. & AOKI, F. 2018. Chd9 mediates highly loosened chromatin structure in growing mouse oocytes. *Biochem Biophys Res Commun,* 500**,** 583-588.

OSTLUND, C., BONNE, G., SCHWARTZ, K. & WORMAN, H. J. 2001. Properties of lamin A mutants found in Emery-Dreifuss muscular dystrophy, cardiomyopathy and Dunnigan-type partial lipodystrophy. *J Cell Sci,* 114**,** 4435-45.

PALSTRA, R. J., TOLHUIS, B., SPLINTER, E., NIJMEIJER, R., GROSVELD, F. & DE LAAT, W. 2003. The beta-globin nuclear compartment in development and erythroid differentiation. *Nat Genet,* 35**,** 190-4.

PANCHENKO, M. V., ZHOU, M. I. & COHEN, H. T. 2004. von Hippel-Lindau partner Jade-1 is a transcriptional co-activator associated with histone acetyltransferase activity. *J Biol Chem,* 279**,** 56032-41.

PANT, V., KURUKUTI, S., PUGACHEVA, E., SHAMSUDDIN, S., MARIANO, P., RENKAWITZ, R., KLENOVA, E., LOBANENKOV, V. & OHLSSON, R. 2004. Mutation of a single CTCF target site within the H19 imprinting control region leads to loss of Igf2 imprinting and complex patterns of de novo methylation upon maternal inheritance. *Mol Cell Biol,* 24**,** 3497-504.

PARADA, L. A., MCQUEEN, P. G. & MISTELI, T. 2004. Tissue-specific spatial organization of genomes. *Genome Biol,* 5**,** R44.

PARADA, L. A., MCQUEEN, P. G., MUNSON, P. J. & MISTELI, T. 2002. Conservation of relative chromosome positioning in normal and cancer cells. *Curr Biol,* 12**,** 1692-7.

PEREZ-LLAMAS, C. & LOPEZ-BIGAS, N. 2011. Gitools: analysis and visualisation of genomic data using interactive heat-maps. *PLoS One,* 6**,** e19541.

PERIC-HUPKES, D., MEULEMAN, W., PAGIE, L., BRUGGEMAN, S. W., SOLOVEI, I., BRUGMAN, W., GRAF, S., FLICEK, P., KERKHOVEN, R. M., VAN LOHUIZEN, M., REINDERS, M., WESSELS, L. & VAN STEENSEL, B. 2010. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell,* 38**,** 603-13.

PEROTTI, C., WIEDL, T., FLORIN, L., REUTER, H., MOFFAT, S., SILBERMANN, M., HAHN, M., ANGEL, P. & SHEMANKO, C. S. 2009. Characterization of mammary epithelial cell line HC11 using the NIA 15k gene array reveals potential regulators of the undifferentiated and differentiated phenotypes. *Differentiation,* 78**,** 269-82.

PETER, M., KITTEN, G. T., LEHNER, C. F., VORBURGER, K., BAILER, S. M., MARIDOR, G. & NIGG, E. A. 1989. Cloning and sequencing of cDNA clones encoding chicken lamins A and B1 and comparison of the primary structures of vertebrate A- and B-type lamins. *J Mol Biol,* 208**,** 393-404.

PHILLIPS-CREMINS, J. E., SAURIA, M. E., SANYAL, A., GERASIMOVA, T. I., LAJOIE, B. R., BELL, J. S., ONG, C. T., HOOKWAY, T. A., GUO, C., SUN, Y., BLAND, M. J., WAGSTAFF, W., DALTON, S., MCDEVITT, T. C., SEN, R., DEKKER, J., TAYLOR, J. & CORCES, V. G. 2013. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell,* 153**,** 1281-95.

PICKERSGILL, H., KALVERDA, B., DE WIT, E., TALHOUT, W., FORNEROD, M. & VAN STEENSEL, B. 2006. Characterization of the Drosophila melanogaster genome at the nuclear lamina. *Nat Genet,* 38**,** 1005-14.

PICO, A. R., KELDER, T., VAN IERSEL, M. P., HANSPERS, K., CONKLIN, B. R. & EVELO, C. 2008. WikiPathways: pathway editing for the people. *PLoS Biol,* 6**,** e184.

PILLAI, R. S., GRIMMLER, M., MEISTER, G., WILL, C. L., LUHRMANN, R., FISCHER, U. & SCHUMPERLI, D. 2003. Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev,* 17**,** 2321-33.

PITTIUS, C. W., SANKARAN, L., TOPPER, Y. J. & HENNIGHAUSEN, L. 1988. Comparison of the regulation of the whey acidic protein gene with that of a hybrid gene containing the whey acidic protein gene promoter in transgenic mice. *Mol Endocrinol,* 2**,** 1027-32.

POLESHKO, A., SHAH, P. P., GUPTA, M., BABU, A., MORLEY, M. P., MANDERFIELD, L. J., IFKOVITS, J. L., CALDERON, D., AGHAJANIAN, H., SIERRA-PAGAN, J. E., SUN, Z., WANG, Q., LI, L., DUBOIS, N. C., MORRISEY, E. E., LAZAR, M. A., SMITH, C. L., EPSTEIN, J. A. & JAIN, R. 2017. Genome-Nuclear Lamina Interactions Regulate Cardiac Stem Cell Lineage Restriction. *Cell,* 171**,** 573-587 e14.

POMBO, A., JACKSON, D. A., HOLLINSHEAD, M., WANG, Z., ROEDER, R. G. & COOK, P. R. 1999. Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. *EMBO J,* 18**,** 2241-53.

POPE, B. D., RYBA, T., DILEEP, V., YUE, F., WU, W., DENAS, O., VERA, D. L., WANG, Y., HANSEN, R. S., CANFIELD, T. K., THURMAN, R. E., CHENG, Y., GULSOY, G., DENNIS, J. H., SNYDER, M. P., STAMATOYANNOPOULOS, J. A., TAYLOR, J., HARDISON, R. C., KAHVECI, T., REN, B. & GILBERT, D. M. 2014. Topologically associating domains are stable units of replication-timing regulation. *Nature,* 515**,** 402-5.

PRASANTH, K. V., PRASANTH, S. G., XUAN, Z., HEARN, S., FREIER, S. M., BENNETT, C. F., ZHANG, M. Q. & SPECTOR, D. L. 2005. Regulating gene expression through RNA nuclear retention. *Cell,* 123**,** 249-63.

PRIETO, J. L. & MCSTAY, B. 2007. Recruitment of factors linking transcription and processing of pre-rRNA to NOR chromatin is UBF-dependent and occurs independent of transcription in human cells. *Genes Dev,* 21**,** 2041-54.

QUINODOZ, S. A., OLLIKAINEN, N., TABAK, B., PALLA, A., SCHMIDT, J. M., DETMAR, E., LAI, M. M., SHISHKIN, A. A., BHAT, P., TAKEI, Y., TRINH, V., AZNAURYAN, E., RUSSELL, P., CHENG, C., JOVANOVIC, M., CHOW, A., CAI, L., MCDONEL, P., GARBER, M. & GUTTMAN, M. 2018. Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. *Cell*.

RAO, S. S., 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. & AIDEN, E. L. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell,* 159**,** 1665-80.

REGAN, J. L., SOURISSEAU, T., SOADY, K., KENDRICK, H., MCCARTHY, A., TANG, C., BRENNAN, K., LINARDOPOULOS, S., WHITE, D. E. & SMALLEY, M. J. 2013. Aurora A kinase regulates mammary

epithelial cell fate by determining mitotic spindle orientation in a Notch-dependent manner. *Cell Rep,* 4**,** 110-23.

ROBSON, M. I., DE LAS HERAS, J. I., CZAPIEWSKI, R., LE THANH, P., BOOTH, D. G., KELLY, D. A., WEBB, S., KERR, A. R. W. & SCHIRMER, E. C. 2016. Tissue-Specific Gene Repositioning by Muscle Nuclear Membrane Proteins Enhances Repression of Critical Developmental Genes during Myogenesis. *Mol Cell,* 62**,** 834-847.

ROCHA, P. P., RAVIRAM, R., BONNEAU, R. & SKOK, J. A. 2015. Breaking TADs: insights into hierarchical genome organization. *Epigenomics,* 7**,** 523-6.

ROSCHGER, C. & CABRELE, C. 2017. The Id-protein family in developmental and cancer-associated pathways. *Cell Commun Signal,* 15**,** 7.

RUAULT, M., DUBARRY, M. & TADDEI, A. 2008. Re-positioning genes to the nuclear envelope in mammalian cells: impact on transcription. *Trends Genet,* 24**,** 574-81.

RUDOLPH, M. C., MONKS, J., BURNS, V., PHISTRY, M., MARIANS, R., FOOTE, M. R., BAUMAN, D. E., ANDERSON, S. M. & NEVILLE, M. C. 2010. Sterol regulatory element binding protein and dietary lipid regulation of fatty acid synthesis in the mammary epithelium. *Am J Physiol Endocrinol Metab,* 299**,** E918-27.

RYBA, T., HIRATANI, I., LU, J., ITOH, M., KULIK, M., ZHANG, J., SCHULZ, T. C., ROBINS, A. J., DALTON, S. & GILBERT, D. M. 2010. Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res,* 20**,** 761-70.

SALOMONI, P. & PANDOLFI, P. P. 2002. The role of PML in tumor suppression. *Cell,* 108**,** 165-70.

SAMEJIMA, K., SAMEJIMA, I., VAGNARELLI, P., OGAWA, H., VARGIU, G., KELLY, D. A., DE LIMA ALVES, F., KERR, A., GREEN, L. C., HUDSON, D. F., OHTA, S., COOKE, C. A., FARR, C. J., RAPPSILBER, J. & EARNSHAW, W. C. 2012. Mitotic chromosomes are compacted laterally by KIF4 and condensin and axially by topoisomerase IIalpha. *J Cell Biol,* 199**,** 755-70.

SAMS, D. S., NARDONE, S., GETSELTER, D., RAZ, D., TAL, M., RAYI, P. R., KAPHZAN, H., HAKIM, O. & ELLIOTT, E. 2016. Neuronal CTCF Is Necessary for Basal and Experience-Dependent Gene Regulation, Memory Formation, and Genomic Structure of BDNF and Arc. *Cell Rep,* 17**,** 2418-2430.

SANBORN, A. L., RAO, S. S., HUANG, S. C., DURAND, N. C., HUNTLEY, M. H., JEWETT, A. I., BOCHKOV, I. D., CHINNAPPAN, D., CUTKOSKY, A., LI, J., GEETING, K. P., GNIRKE, A., MELNIKOV, A., MCKENNA, D., STAMENOVA, E. K., LANDER, E. S. & AIDEN, E. L. 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci U S A,* 112**,** E6456-65.

SARNATARO, S., CHIARIELLO, A. M., ESPOSITO, A., PRISCO, A. & NICODEMI, M. 2017. Structure of the human chromosome interaction network. *PLoS One,* 12**,** e0188201.

SASAKI, Y. T., IDEUE, T., SANO, M., MITUYAMA, T. & HIROSE, T. 2009. MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc Natl Acad Sci U S A,* 106**,** 2525-30.

SCHMITT-NEY, M., DOPPLER, W., BALL, R. K. & GRONER, B. 1991. Beta-casein gene promoter activity is regulated by the hormone-mediated relief of transcriptional repression and a mammary-gland-specific nuclear factor. *Mol Cell Biol,* 11**,** 3745-55.

SCHRANK, B., GOTZ, R., GUNNERSEN, J. M., URE, J. M., TOYKA, K. V., SMITH, A. G. & SENDTNER, M. 1997. Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. *Proc Natl Acad Sci U S A,* 94**,** 9920-5.

SCHUL, W., VAN DRIEL, R. & DE JONG, L. 1998. A subset of poly(A) polymerase is concentrated at sites of RNA synthesis and is associated with domains enriched in splicing factors and poly(A) RNA. *Exp Cell Res,* 238**,** 1-12.

SCHWARTZ, H. C. 1988. Use of the hemicoronal flap to approach the coronoid process. *J Oral Maxillofac Surg,* 46**,** 519.

SCHWARZER, W., ABDENNUR, N., GOLOBORODKO, A., PEKOWSKA, A., FUDENBERG, G., LOE-MIE, Y., FONSECA, N. A., HUBER, W., C, H. H., MIRNY, L. & SPITZ, F. 2017. Two independent modes of chromatin organization revealed by cohesin removal. *Nature,* 551**,** 51-56.

SEXTON, T., YAFFE, E., KENIGSBERG, E., BANTIGNIES, F., LEBLANC, B., HOICHMAN, M., PARRINELLO, H., TANAY, A. & CAVALLI, G. 2012. Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell,* 148**,** 458-72.

SHARMIN, M., BRAVO, H. C. & HANNENHALLI, S. 2016. Heterogeneity of transcription factor binding specificity models within and across cell lines. *Genome Res,* 26**,** 1110-23.

SHIODA, T., FENNER, M. H. & ISSELBACHER, K. J. 1997. MSG1 and its related protein MRG1 share a transcription activating domain. *Gene,* 204**,** 235-41.

SHIPRA, A., CHETAN, K. & RAO, M. R. 2006. CREMOFAC--a database of chromatin remodeling factors. *Bioinformatics,* 22**,** 2940-4.

SOFUEVA, S., YAFFE, E., CHAN, W. C., GEORGOPOULOU, D., VIETRI RUDAN, M., MIRA-BONTENBAL, H., POLLARD, S. M., SCHROTH, G. P., TANAY, A. & HADJUR, S. 2013. Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J,* 32**,** 3119-29.

SORNAPUDI, T. R., NAYAK, R., GUTHIKONDA, P. K., KETHAVATH, S., YELLABOINA, S. & KURUKUTI, S. 2018. RNA sequencing of murine mammary epithelial stem-like cells (HC11) undergoing lactogenic differentiation and its comparison with embryonic stem cells. *BMC Res Notes,* 11**,** 241.

SOSHNIKOVA, N., MONTAVON, T., LELEU, M., GALJART, N. & DUBOULE, D. 2010. Functional analysis of CTCF during mammalian limb development. *Dev Cell,* 19**,** 819-30.

SPLINTER, E., HEATH, H., KOOREN, J., PALSTRA, R. J., KLOUS, P., GROSVELD, F., GALJART, N. & DE LAAT, W. 2006. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev,* 20**,** 2349-54.

SPROUL, D., GILBERT, N. & BICKMORE, W. A. 2005. The role of chromatin structure in regulating the expression of clustered genes. *Nat Rev Genet,* 6**,** 775-81.

STACK, S. M., BROWN, D. B. & DEWEY, W. C. 1977. Visualization of interphase chromosomes. *J Cell Sci,* 26**,** 281-99.

STANEK, D. & NEUGEBAUER, K. M. 2006. The Cajal body: a meeting place for spliceosomal snRNPs in the nuclear maze. *Chromosoma,* 115**,** 343-54.

STEVENS, T. J., LANDO, D., BASU, S., ATKINSON, L. P., CAO, Y., LEE, S. F., LEEB, M., WOHLFAHRT, K. J., BOUCHER, W., O'SHAUGHNESSY-KIRWAN, A., CRAMARD, J., FAURE, A. J., RALSER, M., BLANCO, E., MOREY, L., SANSO, M., PALAYRET, M. G. S., LEHNER, B., DI CROCE, L., WUTZ, A., HENDRICH, B., KLENERMAN, D. & LAUE, E. D. 2017. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature,* 544**,** 59-64.

STOCK, J. K., GIADROSSI, S., CASANOVA, M., BROOKES, E., VIDAL, M., KOSEKI, H., BROCKDORFF, N., FISHER, A. G. & POMBO, A. 2007. Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat Cell Biol,* 9**,** 1428-35.

STOELZLE, T., SCHWARB, P., TRUMPP, A. & HYNES, N. E. 2009. c-Myc affects mRNA translation, cell proliferation and progenitor cell function in the mammary gland. *BMC Biol,* 7**,** 63.

STRICKFADEN, H., ZUNHAMMER, A., VAN KONINGSBRUGGEN, S., KOHLER, D. & CREMER, T. 2010. 4D chromatin dynamics in cycling cells: Theodor Boveri's hypotheses revisited. *Nucleus,* 1**,** 284-97.

SUNWOO, H., DINGER, M. E., WILUSZ, J. E., AMARAL, P. P., MATTICK, J. S. & SPECTOR, D. L. 2009. MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res,* 19**,** 347-59.

SZEGEZDI, E., LOGUE, S. E., GORMAN, A. M. & SAMALI, A. 2006. Mediators of endoplasmic reticulum stress-induced apoptosis. *EMBO Rep,* 7**,** 880-5.

TANABE, H., HABERMANN, F. A., SOLOVEI, I., CREMER, M. & CREMER, T. 2002. Non-random radial arrangements of interphase chromosome territories: evolutionary considerations and functional implications. *Mutat Res,* 504**,** 37-45.

THERIZOLS, P., ILLINGWORTH, R. S., COURILLEAU, C., BOYLE, S., WOOD, A. J. & BICKMORE, W. A. 2014. Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science,* 346**,** 1238-42.

TOLHUIS, B., PALSTRA, R. J., SPLINTER, E., GROSVELD, F. & DE LAAT, W. 2002. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell,* 10**,** 1453-65.

TOWBIN, B. D., GONZALEZ-AGUILERA, C., SACK, R., GAIDATZIS, D., KALCK, V., MEISTER, P., ASKJAER, P. & GASSER, S. M. 2012. Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell,* 150**,** 934-47.

VAN DAMME, E., LAUKENS, K., DANG, T. H. & VAN OSTADE, X. 2010. A manually curated network of the PML nuclear body interactome reveals an important role for PML-NBs in SUMOylation dynamics. *Int J Biol Sci,* 6**,** 51-67.

VAN STEENSEL, B. & BELMONT, A. S. 2017. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell,* 169**,** 780-791.

VAUGHAN, A., ALVAREZ-REYES, M., BRIDGER, J. M., BROERS, J. L., RAMAEKERS, F. C., WEHNERT, M., MORRIS, G. E., WHITFIELD, W. G. F. & HUTCHISON, C. J. 2001. Both emerin and lamin C depend on lamin A for localization at the nuclear envelope. *J Cell Sci,* 114**,** 2577-90.

VERHEGGEN, C., LAFONTAINE, D. L., SAMARSKY, D., MOUAIKEL, J., BLANCHARD, J. M., BORDONNE, R. & BERTRAND, E. 2002. Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. *EMBO J,* 21**,** 2736-45.

VERSCHURE, P. J., VAN DER KRAAN, I., ENSERINK, J. M., MONE, M. J., MANDERS, E. M. & VAN DRIEL, R. 2002. Large-scale chromatin organization and the localization of proteins involved in gene expression in human cells. *J Histochem Cytochem,* 50**,** 1303-12.

VERSTEEG, R., VAN SCHAIK, B. D., VAN BATENBURG, M. F., ROOS, M., MONAJEMI, R., CARON, H., BUSSEMAKER, H. J. & VAN KAMPEN, A. H. 2003. The human transcriptome map reveals extremes in gene density, intron length, GC content, and repeat pattern for domains of highly and weakly expressed genes. *Genome Res,* 13**,** 1998-2004.

VITOUX, D., NASR, R. & DE THE, H. 2007. Acute promyelocytic leukemia: new issues on pathogenesis and treatment response. *Int J Biochem Cell Biol,* 39**,** 1063-70.

VOLPI, E. V., CHEVRET, E., JONES, T., VATCHEVA, R., WILLIAMSON, J., BECK, S., CAMPBELL, R. D., GOLDSWORTHY, M., POWIS, S. H., RAGOUSSIS, J., TROWSDALE, J. & SHEER, D. 2000. Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. *J Cell Sci,* 113 (Pt 9)**,** 1565-76.

VORBURGER, K., KITTEN, G. T. & NIGG, E. A. 1989. Modification of nuclear lamin proteins by a mevalonic acid derivative occurs in reticulocyte lysates and requires the cysteine residue of the C-terminal CXXM motif. *EMBO J,* 8**,** 4007-13.

WACHSMUTH, M., CAUDRON-HERGER, M. & RIPPE, K. 2008. Genome organization: balancing stability and plasticity. *Biochim Biophys Acta,* 1783**,** 2061-79.

WALTER, J., SCHERMELLEH, L., CREMER, M., TASHIRO, S. & CREMER, T. 2003. Chromosome order in HeLa cells changes during mitosis and early G1, but is stably maintained during subsequent interphase stages. *J Cell Biol,* 160**,** 685-97.

WANG, H., MAURANO, M. T., QU, H., VARLEY, K. E., GERTZ, J., PAULI, F., LEE, K., CANFIELD, T., WEAVER, M., SANDSTROM, R., THURMAN, R. E., KAUL, R., MYERS, R. M. & STAMATOYANNOPOULOS, J. A. 2012. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res,* 22**,** 1680-8.

WANG, W., JOSE, C., KENNEY, N., MORRISON, B. & CUTLER, M. L. 2009. Global expression profiling reveals regulation of CTGF/CCN2 during lactogenic differentiation. *J Cell Commun Signal,* 3**,** 43-55.

WATSON, C. J. & KHALED, W. T. 2008. Mammary development in the embryo and adult: a journey of morphogenesis and commitment. *Development,* 135**,** 995-1003.

WEI, X., SOMANATHAN, S., SAMARABANDU, J. & BEREZNEY, R. 1999. Three-dimensional visualization of transcription sites and their association with splicing factor-rich nuclear speckles. *J Cell Biol,* 146**,** 543-58.

WEINREB, C. & RAPHAEL, B. J. 2016. Identification of hierarchical chromatin domains. *Bioinformatics,* 32**,** 1601-9.

WEINTRAUB, A. S., LI, C. H., ZAMUDIO, A. V., SIGOVA, A. A., HANNETT, N. M., DAY, D. S., ABRAHAM, B. J., COHEN, M. A., NABET, B., BUCKLEY, D. L., GUO, Y. E., HNISZ, D., JAENISCH, R., BRADNER, J. E., GRAY, N. S. & YOUNG, R. A. 2017. YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell,* 171**,** 1573-1588 e28.

WEN, B., WU, H., SHINKAI, Y., IRIZARRY, R. A. & FEINBERG, A. P. 2009. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet,* 41**,** 246-50.

WESTENDORP, B., MOKRY, M., GROOT KOERKAMP, M. J., HOLSTEGE, F. C., CUPPEN, E. & DE BRUIN, A. 2012. E2F7 represses a network of oscillating cell cycle genes to control S-phase progression. *Nucleic Acids Res,* 40**,** 3511-23.

WICKHAM, H. 2009. *Ggplot2 : elegant graphics for data analysis,* New York, Springer.

WILLIAMS, C., HELGUERO, L., EDVARDSSON, K., HALDOSEN, L. A. & GUSTAFSSON, J. A. 2009. Gene expression in murine mammary epithelial stem cell-like cells shows similarities to human breast cancer gene expression. *Breast Cancer Res,* 11**,** R26.

WINGETT, S., EWELS, P., FURLAN-MAGARIL, M., NAGANO, T., SCHOENFELDER, S., FRASER, P. & ANDREWS, S. 2015. HiCUP: pipeline for mapping and processing Hi-C data. *F1000Res,* 4**,** 1310.

WINTERMANTEL, T. M., BOCK, D., FLEIG, V., GREINER, E. F. & SCHUTZ, G. 2005. The epithelial glucocorticoid receptor is required for the normal timing of cell proliferation during mammary lobuloalveolar development but is dispensable for milk production. *Mol Endocrinol,* 19**,** 340-9.

WOODCOCK, C. L., FRADO, L. L. & RATTNER, J. B. 1984. The higher-order structure of chromatin: evidence for a helical ribbon arrangement. *J Cell Biol,* 99**,** 42-52.

YAHATA, T., DE CAESTECKER, M. P., LECHLEIDER, R. J., ANDRIOLE, S., ROBERTS, A. B., ISSELBACHER, K. J. & SHIODA, T. 2000. The MSG1 non-DNA-binding transactivator binds to the p300/CBP coactivators, enhancing their functional link to the Smad transcription factors. *J Biol Chem,* 275**,** 8825-34.

YAHATA, T., TAKEDATSU, H., DUNWOODIE, S. L., BRAGANCA, J., SWINGLER, T., WITHINGTON, S. L., HUR, J., COSER, K. R., ISSELBACHER, K. J., BHATTACHARYA, S. & SHIODA, T. 2002. Cloning of mouse Cited4, a member of the CITED family p300/CBP-binding transcriptional coactivators: induced expression in mammary epithelial cells. *Genomics,* 80**,** 601-13.

YANG, J., KENNELLY, J. J. & BARACOS, V. E. 2000a. The activity of transcription factor Stat5 responds to prolactin, growth hormone, and IGF-I in rat and bovine mammary explant culture. *J Anim Sci,* 78**,** 3114-25.

YANG, J., KENNELLY, J. J. & BARACOS, V. E. 2000b. Physiological levels of Stat5 DNA binding activity and protein in bovine mammary gland. *J Anim Sci,* 78**,** 3126-34.

YORI, J. L., JOHNSON, E., ZHOU, G., JAIN, M. K. & KERI, R. A. 2010. Kruppel-like factor 4 inhibits epithelial-to-mesenchymal transition through regulation of E-cadherin gene expression. *J Biol Chem,* 285**,** 16854-63.

YOUNG, P. J., LE, T. T., THI MAN, N., BURGHES, A. H. & MORRIS, G. E. 2000. The relationship between SMN, the spinal muscular atrophy protein, and nuclear coiled bodies in differentiated tissues and cultured cells. *Exp Cell Res,* 256**,** 365-74.

YUSUFZAI, T. M., TAGAMI, H., NAKATANI, Y. & FELSENFELD, G. 2004. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell,* 13**,** 291-8.

ZHAN, Y., MARIANI, L., BAROZZI, I., SCHULZ, E. G., BLUTHGEN, N., STADLER, M., TIANA, G. & GIORGETTI, L. 2017. Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes. *Genome Res,* 27**,** 479-490.

ZHANG, H., SMOLEN, G. A., PALMER, R., CHRISTOFOROU, A., VAN DEN HEUVEL, S. & HABER, D. A. 2004. SUMO modification is required for in vivo Hox gene regulation by the Caenorhabditis elegans Polycomb group protein SOP-2. *Nat Genet,* 36**,** 507-11.

ZHANG, S. & CUI, W. 2014. Sox2, a key factor in the regulation of pluripotency and neural differentiation. *World J Stem Cells,* 6**,** 305-11.

ZHAO, Z., TAVOOSIDANA, G., SJOLINDER, M., GONDOR, A., MARIANO, P., WANG, S., KANDURI, C., LEZCANO, M., SANDHU, K. S., SINGH, U., PANT, V., TIWARI, V., KURUKUTI, S. & OHLSSON, R. 2006. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet,* 38**,** 1341-7.

ZHONG, S., SALOMONI, P. & PANDOLFI, P. P. 2000. The transcriptional role of PML and the nuclear body. *Nat Cell Biol,* 2**,** E85-90.

ZHOU, H., CHENG, L., AZIMU, W., HODGE, S., EDWARDS, G. R. & HICKFORD, J. G. 2015. Variation in the bovine FABP4 gene affects milk yield and milk protein content in dairy cows. *Sci Rep,* 5**,** 10023.

ZHOU, J., CHEHAB, R., TKALCEVIC, J., NAYLOR, M. J., HARRIS, J., WILSON, T. J., TSAO, S., TELLIS, I., ZAVARSEK, S., XU, D., LAPINSKAS, E. J., VISVADER, J., LINDEMAN, G. J., THOMAS, R., ORMANDY, C. J., HERTZOG, P. J., KOLA, I. & PRITCHARD, M. A. 2005. Elf5 is essential for early embryogenesis and mammary gland development during pregnancy and lactation. *EMBO J,* 24**,** 635-44.

ZHOU, W., ZHU, P., WANG, J., PASCUAL, G., OHGI, K. A., LOZACH, J., GLASS, C. K. & ROSENFELD, M. G. 2008. Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. *Mol Cell,* 29**,** 69-80.

ZIDI, A., FERNANDEZ-CABANAS, V. M., CARRIZOSA, J., JORDANA, J., URRUTIA, B., POLVILLO, O., GONZALEZ-REDONDO, P., GALLARDO, D., AMILLS, M. & SERRADILLA, J. M. 2010. Genetic variation at the goat hormone-sensitive lipase (LIPE) gene and its association with milk yield and composition. *J Dairy Res,* 77**,** 190-8.

ZORN, C., CREMER, C., CREMER, T. & ZIMMER, J. 1979. Unscheduled DNA synthesis after partial UV irradiation of the cell nucleus. Distribution in interphase and metaphase. *Exp Cell Res,* 124**,** 111-9.

ZORN, C., CREMER, T., CREMER, C. & 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**,** 83-9.

ZUIN, J., DIXON, J. R., VAN DER REIJDEN, M. I., YE, Z., KOLOVOS, P., BROUWER, R. W., VAN DE CORPUT, M. P., VAN DE WERKEN, H. J., KNOCH, T. A., VAN, I. W. F., GROSVELD, F. G., REN, B. & WENDT, K. S. 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc Natl Acad Sci U S A,* 111**,** 996-1001.