Comprehensive Profiling and Functional Dynamics of microRNAs: A Developmental Perspective

Thesis submitted to University of Hyderabad

for the award of Ph.D. degree in

Department of Animal Biology

By

Rakhee Nayak

14LAPH03

Department of Animal Biology

School of Life Sciences

University of Hyderabad

Hyderabad - 500 046

India

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 **"Comprehensive Profiling and Functional Dynamics of microRNAs: A Developmental Perspective"** has been carried out by me under the supervision of **Prof. Sreenivasulu Kurukuti** and this work has not been submitted for any degree or diploma of any other university earlier.

Dated: 7th November 2021

Rakhee Nayak

Rakhee Nayak

14LAPH03

University of Hyderabad
(A Central University by an act of Parliament)
Department of Animal Biology
School of Life Sciences
P.O. Central University, Gachibowli, Hyderabad-500046

CERTIFICATE

This is to certify that the thesis entitled **"Comprehensive Profiling and Functional Dynamics of microRNAs: A Developmental Perspective"** submitted by **Rakhee Nayak** bearing registration number **14LAPH03** in partial fulfilment of the requirements for award of Doctor of philosophy in the **School of Life Sciences** is a bona fide work carried out by her 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. Parts of the thesis had been **Published in following publications:**

1. Sornapudi TR, Nayak R, Guthikonda PK, Kethavath S, Yellaboina S and Kurukuti S* **(2018)**. RNA sequencing of murine mammary epithelial stem-like cells (HC11) undergoing lactogenic differentiation and its comparision with embryonic stem cells. ***BMC Research Notes***. 11-241.
2. Sornapudi TR, Nayak R, Guthikonda PK, Pasupuleti AK, Kethavath S, Uppada V, Mondal S, Yellaboina S and Kurukuti S* **(2018)**. Comprehensive profiling of transcriptional networks specific for lactogenic differentiation of HC11 mammary epithelial cells. ***Scientific Reports***. 8:11777.

Presented in the following conference:

1. NextGen Genomics, Biology, Bioinformatics and Technologies (NGBT), "miRNA-mRNA integrative expression mapping during mouse embryonic stem cell to Neuron progenitor differentiation". Poster Presentation at Bhubaneswar, 2017.

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

Course Code	Title of the Course	Credits	Pass/Fail
1. AS 801	Analytical Techniques	4	PASS
2. AS 802	Research Ethics, Data Analysis and Biostatics	3	PASS
3. AS 803	Lab Work and seminar	5	PASS

K. Sreenivasulu
Supervisor
Dr. K. Sreenivasulu
Professor
Department of Animal Biology
School of Life Sciences
University of Hyderabad
Gachibowli, Hyderabad-500 046. T.S.

K. Sreenivasulu
Head of the Department
अध्यक्ष / HEAD
जंतु जैविकी विभाग
Department of Animal Biology

Dean of the School
DEAN
School of Life Sciences
University of Hyderabad
Hyderabad - 500 046.

ACKNOWLWDGEMENTS

- *I am very grateful to my PhD supervisor* ***Prof. Sreenivasulu Kurukuti*** *for his immense support and valuable guidance throughout my PhD work.*
- *I am grateful to my Doctoral Committee Members* ***'Late' Prof. Aparna Dutta-Gupta, Dr. Bindu Madhava Reddy and Prof. Anita Jagota*** *for their semester wise assessment of my work and valuable inputs.*
- *I am thankful to* ***Dr. Anil k Pasupulati and Prof. G Ravi Kumar*** *for their valuable inputs.*
- *I Sincerely thank former Heads of the Dept. of Animal Biology* ***Prof. B Senthilkumaran, Prof. Jagan Pongubala, Prof. Anita Jagota*** *and present HOD* ***Prof. Sreenivasulu Kurukuti*** *for allowing me to use departmental facilities.*
- *I heartily thank former Deans of the school* ***Prof. Reddanna, Prof. KVA Ramaiah*** *and present Dean* ***Prof. S Dayananda*** *for allowing me to use school facilities.*
- *I am thankful to all* ***non-teaching staff*** *of Dept. of animal Biology and School of Life Sciences for their administrative supports.*
- *I greatly acknowledge* ***UGC-RGNF-ST*** *for providing me fellowship during my work.*
- *I would like to thank my dear Lab Mates* ***Dr. EV Trinadharao Sornapudi, Prashanth Guthikonda, Sukalpa Mondal, Srinivas Kethavath, Netrika Tiwari, Yuva Sri Golivi, Satyanarayana Nadiminti, Dr Dhammapal Bharne*** *and* ***Sharmistha Chaitali Ghosh*** *for valuable inputs and moral support.*
- *I am grateful to my friends* ***Priyanka Kritinarayan, Dipti Singh, Jahnabi Ramchary, Minurani Dalai, Sangeeta Kumari, Nisha Chouhan, Sandeep Day, Dr Kavyashree Puranik, Rutuparna Jena, Bindia Chawla, Joytisha Basantia, Shivani Navalakha*** *and* ***Akash Patel****.*
- *I express my priceless gratitude towards my Parents: Mr.* ***Somanath Nayak*** *and Mrs.* ***Sarojini Nayak****, my brother Mr.* ***Bandhan Kumar Nayak****, my husband* ***Dr. EV Trinadharao Sornapudi****, my beloved daughter Miss.* ***Shivanshika Sornapudi*** *and my doggie 'Late'* ***Bhallu*** *for constant hope, encouragement and mental support.*

Table of Contents

	Pages
Abbreviations	**0**
Introductions	**1-37**

1. **One Genome Regulates Formation of Multiple Cell-Types**
2. **Epigenetic Modes of Gene Regulation during Cellular Differentiation and development of mammalian cells**
   a. **Gene regulation during pre-transcription at the level of genome packing**
      - **3D- Chromatin Organization**
      - **Chromatin accessibility**
      - **Chromatin condensation**
      - **Replication Timing**
      - **DNA methylation**
   b. **Gene regulation at transcription level**
      - **Transcription factors**
      - **Histone modifications**
      - **Alternative Splicing**
   c. **Gene regulation at post-transcription level**
      - **RNA binding proteins**
      - **RNA methylation**
      - **Alternative splicing**
      - **Long non-coding RNA**
   d. **Gene regulation at Translational level**
      - **Specialized ribosomes**
      - **microProteins**
   e. **MicroRNA**
      - **Historical perspective;**
      - **miRNA structural organization;**
      - **MicroRNA mechanism of action**
      - **MicroRNA developmental perspective (development, differentiation and disease)**
3. **MiRNA in the context of Pluripotency**
4. **MiRNA in the context of Neurogenesis**
5. **MiRNA in the context of Lactogenesis**
6. **MiRNA what is not known?**
7. **Logical hypothesis of miRNA functional dynamics in the context of cellular differentiation**

Materials and Methods	**38-57**
Results and Discussions	**58-151**
Summary	**152-153**
Tables	**154-223**
References	**224-240**
Publications	**241**
Anti-plagiarism certificate	

Abbreviations

ESC: Embryonic Stem Cells

ESC (G): Ground state of ESC

ESC (N+2i): Naïve+2i state of ESC

ESC (N): Naïve state of ESC

LIF: Leukemia Inhibitory Factor

2i: PD+CH

PD: MAP2K inhibitor

CH: GSK3ß inhibitor

TF: Transcription Factors

ER: Epigenetic Regulators

NC: Normalized Count

FPKM: Fragments Per Kilobase per Million

FACS: Fluorescence Activated Cell Sorting

NPC: Neuron Progenitor Cells

CN: Cortical Neurons

CN-Si: CN with negative inhibitor condition

CN+301bi: CN with *mmu-miR-301b-3p* knockdown

HITS-CLIP: High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation

CLASH: Cross-linking, Ligation and Sequencing of Hybrids

OPC: Oligodendrocytes Precursor Cells

MEC: Mammary Epithelial Cells

MEC (N): Normal state with Insulin & EGF

MEC (P): Primed state with Insulin & HC

MEC (Prl): Prolactin state with Insulin, HC & Prl

Prl/PRL: Prolactin

GR: Glucocorticoid Receptor

HC: Hydrocortisone

EGF: Epidermal Growth Factor

MEC (N-122i): MEC (N) state with *mmu-miR-122-5p* knockdown

MEC (P-122i): MEC (P) state with *mmu-miR-122-5p* knockdown

MEC (Prl-122i): MEC (Prl) state with *mmu-miR-122-5p* knockdown

WT: Wild Type

miR-122i: *miR-122-5p* knockdown condition

MEC (Prl-Si): MEC (N) state with negative inhibitor condition

cLAD: conserved lamina associated domains

ciLAD: conserved inter lamina associated domains

cfLAD: conserved facultative lamina associated domains

GE: Gene Expression

UCSC: University of California Santa Cruz

DGE: Differential Gene Expression

Introduction

One Genome Regulates Formation of Multiple Cell-Types

A journey from a haploid cell to a diploid organism is a wonder of nature. From the time of fertilization when sperm fertilizes oocyte, the chromosomes intermingle and the first diploid zygote develops. Zygote transcribes its RNA when it enters the two-cell stage. Before this, it depends upon maternally generated and stored RNAs (Stoeckius et al., 2014). These two cells divide mitotically until it reaches 8 cell blastomere, where it gains polarization properties. Due to polarization, cells undergo compaction and develop morula (Tsai et al., 2019). In morula, the 8-cell stage undergoes asymmetric division to become the 16-cell stage, known as the blastocyst. Among them, polarized cells occupy an exterior position, called trophectoderm and non-polarized cells migrate to the interior of an embryo to form inner cell mass (ICM) (Wigger et al., 2017). Trophectoderm gives rise to the placenta that connects with the mother's uterus but does not involve in the development of the embryo proper. Whereas inner cell mass develops into three primary germ layers, named ectoderm, mesoderm, and endoderm (Silvia Pellegrini et al., 2020). Ectoderm develops into neurons and epidermis; mesoderm gives rise to blood and muscle cells and endoderm turns into the gastrointestinal tract. This inner cell mass is isolated from the embryo and cultured on plates *in vitro* termed embryonic stem cells (Thomson et al., 1995, Bongso et al., 1994).

Embryonic stem cells (ESC) are pluripotent and procure the ability of self-renewal (Gerrard et al., 2005). The first ESC line was isolated in 1994 from the rhesus macaque embryo. At that time the pluripotent ESCs were maintained on fibroblast feeder cells for more than one year. Later the same group Thomson et al. isolated ESC from human patients. Gradually, it didn't take much time to discover the ability of ESC to differentiate to any kind of germ layers (Thomson et al., 1998). Differentiation in a controlled fashion could lead to the formation of multiple cell types including mature neurons, cardiac cells (Dvash et al., 2006), and insulin-producing cells (Kroon E et al., 2008). The undifferentiated state of ESCs is maintained by core transcription factors and other epigenetic features such as DNA and histone methylation. ESC state is controlled by core transcriptional factors such as OCT4, SOX2, and NANOG. OCT4 and NANOG are key regulators. OCT4 heterodimerizes with SOX2 in ESCs and regulates target genes. NANOG may not be required but provides stability (Zakrzewski et al., 2019). Lack of any one of these factors promotes the differentiation of ESCs. These core transcription factors have three ways to maintain ESC state that is through positive regulation of their promoters, activation of ESCs specific genes, and downregulation of lineage determining factors. These studies on the regulation of gene

expression in ESCs development and differentiation are useful for understanding embryonic development, differentiation, and diseases progression which should pave the way for efficient management in regenerative medicine.

Promoters of core transcription factors are regulated by Histone modifying enzyme (TRXG, DOT1, SET2), chromatin regulators (PcG, SetDB1), Mediator (cohesion, and the cohesion-loading factor NIPBL), and Co-repressors (DAX1, CNOT3, TRIM28). TRXG is a negative regulator of transcription, which trimethylates at H3K4 to maintain an active state during development (Boland et al., 2015). Chromatin regulator, PcG is antagonistic to TRXG (Klymenko et al., 2004). PcG does ubiquitylation of H2AK119U and trimethylate H3K27me3 to remain undifferentiated (Boland et al., 2015). Cohesion is a chromatin regulator that mediates chromatin loops to bring enhancers and promoters together, which initiates transcription. NIPBL helps in loading cohesion to DNA (Hnisz et al., 2013). Co-repressor, DAX1 directly acts upon OCT4. The increasing amount of DAX1 leads to cell differentiation (Zhang et al., 2014). c-Myc is another transcriptional regulator required for the maintenance of ESCs state. C-MYC binds to P-TEFb that phosphorylates the C-terminal domain of RNA polymerase II and negatively phosphorylates two pause factors (DSIF and NELF) (Rahl et al., 2010). DNA and histone methylation are under the control of DNA and Histone methyltransferases respectively. Methylation at H3K27 and H3K9 and ubiquitylation of H2A lead to repression of gene expression. Tip60-P400 complex does acetylation of histones and loss of this complex affect ESCs state (Chen PB et al., 2013). Methyltransferases such as DNMT1 1, 2, 3a, 3b & 3l is also involved in ESC differentiation (Liao et al., 2015). Hence, understanding the epigenetic mechanisms in the maintenance of embryonic stem cell state and differentiation would pave the way for a fundamental understanding of embryonic development and disease.

Epigenetic Modes of Gene Regulation during Cellular Differentiation and development of mammalian cells

Gene regulation during pre-transcription at the level of genome packing:

3D- Chromatin Organization

Differentiation of ESCs depends upon epigenetic modifiers, chromatin regulators and many distal enhancers to the promoter of lineage-specific gene to modulate pluripotency towards multipotency and terminally differentiated lineage. The interplay between promoter, enhancer and insulator

region of chromatin is vital. These enhancers can modulate the promoter which need not be the immediate neighbour but located far in the genome. That's why, the interaction of these two regions at linear fashion may not be feasible, which increase the probability of genome folding (Yu Miao et al., 2017). Early studies had shown the importance of chromatin looping is necessary to regulate lineage-specific marker during development and genome function is greatly influenced by chromatin architecture that arranged inside nuclear space with defined boundaries (Cremer et al., 2008). To understand the changing gene expression concerning changing developmental lineage, firstly it's necessary to track the spatial organization of chromatin fibres.

In recent years many tools have been reported to dissect the 3D arrangement of chromatin looping. Previously Fluorescence in situ hybridization (FISH) was used massively to measure the distance between two genomic loci (Langer-Safer et al., 1982). Here, 200kb probes have been used from BAC or YAC, but now fosmid is used most frequently which gives coverage of around 40kb (Cremer et al., 2008). Nowadays to increase the resolution between loci super-resolution microscopy was implemented. Live-cell imaging by direct labelling through the implementation of dCas9 fusion proteins provides better coverage to encounter interaction frequency between a promoter and its distal elements. It has provided evidence for the existence of chromosome territories, for a decade (Boveri et al., 1909). But, it has its limitation towards genome coverage and resolution. Also, the variation between cell to cell can't be ignored because of which at least 100 cells are need to analysed with statistical significance.

To overcome the limitation in genome coverage, molecular biology-based technique chromosome conformation capture (3C) (Dekker et al., 2002) was introduced. Here, we cross-link the cells with formaldehyde which preserves all the chromatin-protein interaction that makes it easy to track chromatin loops from distal genomic loci. Afterward, the genome was digested by using either four based or six based restriction enzymes which provide around 256bp and 4096bp resolutions respectively (Dekker et al., 2002). Spatially close genomic loci can be identified through the ligated product from a digested genome. This will provide proximity ligation between different gene loci that are present far away in the genome. Though it is a recent technique, in these courses of the period time it emerged so efficiently with other 3C based techniques called 4C, 5C, and Hi-C. Everything is fundamentally similar to 3C, only differ in their efficiency and resolution as per their requirements. 3C provides information regarding the one-to-one genomic region. Likewise, 4C is for one to many, 5C for many to many and Hi-C is about all to all. Ligated 3C library further proceeds with PCR extension by using a biotin tagged primer. Generally, multiple targeted primers

are being used in a single PCR pool. This provides information regarding targeted loci and it's all possible genome-wide interaction (Simonis et al., 2009). 5C, on the other hand, is used to predict interaction frequency within a megabase scale. Like 4C, completely targeted regionwith in a loop is covered with 5C primer pools which are used to amplify the 3C library (Philips-cremins et al., 2013, Rao et al., 2014). Unlike, these above mentioned C-based techniques, Hi-C offers all possible genome-wide interactions. It only differs from 3C at the biotin end filling step after ligation to enrich more distally ligated products. Later on, in nucleus Hi-C was developed where interactions are restricted to one particular nucleus. This reduces noise from a background in comparison with 3C and diluted Hi-C (Hsieh et al., 2015, Rao et al., 2014). To cover cell to cell variability, even single-cell Hi-C was developed (Nagano et al., 2013). To capture specific protein-based chromatin interactions, scientists ligated DNA fragments from pulldown products of ChIP. This specifies a genomic interaction map due to a particular binding of a protein and this technique is named ChIA-PET (Fullwood et al., 2009). Further, to capture many to all interactions capture Hi-C has been used. Same as 4C, more than 100 biotins tagged primers are used in a primer pool to target genome-wide interactions. These C-based techniques will only provide a probable contact frequency regarding proximity ligation in 3D genome architecture, that needs to be further validated with microscopy techniques.

All these tools support the non-randomness of chromatin in the nucleus. Both C-based techniques along with imaging through chromosome painting proved the existence of discrete territory for individual chromosomes. It disproved the concept of randomly organized chromatin. These defined territories of each chromosome termed chromosome territory (CT) (Bolzer et al., 2005, Branco & Pombo, 2006). The intermingling of chromosomes is restricted to chromosome territory boundaries. Chromosome territories are specific to cell types which also proved in Hi-C datasets that they maintain chromosome neighbourhoods (Liberman et al., 2009). After mitosis, positions of CTs are partially conserved. It had been also observed in FISH and intra-chromosomal Hi-C studies the genes that belong to CT occupy a specific position in the nucleus. Generally, gene-rich early replicating and active genes tend to localize towards the nuclear interior. Similarly, gene-poor, late replicating, and inactive genes are placed towards the nuclear periphery (Miao et al., 2017). Further, principle component analysis of observed vs expected matrix from Hi-C datasets revealed the existence of compartments inside CTs. They were named Compartments A and B (Liberman et al). Compartment A contains early replicating regions and active genes, whereas compartment B contains late replicating and inactive genes (Ryba et al., 2010). The proportion of compartments A and B and switching between them during differentiation is cell-type specific. These

compartments are framed based upon both structural and functional significance (Pope et al., 2014). Later, it was introduced as Topologically Associated Domains (TAD), having hundreds of kilobases to million in lengths (Dixon et al., 2012). Concerning the functional site of the genome, TADs are not only physically defined boundaries in chromatin but also having the most active interaction between genes within it, rather than other TADs (Dixon et al., 2012).

In the context of Embryonic stem cells, long-range contacts establish around promoter regions of Oct4, Sox2, and Nanog. Long-range contact between distal enhancer and promoter of Oct4 is mediated by binding of Cohesin and KLF4 proteins. OCT4, SOX2, and NANOG by itself associated with many ESC-specific distal interactions (Wei et al., 2013, Zhang et al., 2014). In pluripotency state, TAD boundaries are maintained by both CTCF and Cohesin. Also, developmentally associated genes are surrounded by polycomb group proteins. During differentiation, both inter and intra TAD interactions change based on epigenetic modifiers and gene activation. Many compartments change B to A with active histone marks and increase in gene expression parallelly with increase intra TAD interactions. Similarly, compartment changing from A to B shows reduced interaction along with minimal gene expression (Miao et al., 2017).

Chromatin accessibility

Chromatin fibres are packed inside the nucleus with the help of nucleosomes. Nucleosomes are the first level of packaging composed of an octamer of histone proteins which is rapped by 147bp of DNA fragment (Olins et al., 2003). Chromatin accessibility is the degree of availability of DNA fragments to be bounded by macromolecules. Generally, more nucleosomes are found in compact regions of the genome like heterochromatin compared to decompacted euchromatin. The openness of chromatin is associated with the binding of either active or repressive histone markers. Further, these decondensed DNA fragments are bounded by transcription factors to initiate gene transcription. Chromatin accessibility can be measured by quantification of DNase hypersensitivity or availability of chromatin for restriction enzyme digestion (Kornberg et al., 1974, Wu. C et al., 1979). In 2006, using DNase I, chromatin accessibility was measured for the first time. Later on, many different protocols were adopted which infers 80% accessible regions in distal enhancers of genes in comparison to promoters and transcription start sites (Thurman et al., 2012). By targeting ATAC sequence open chromatin is accessed by using Tn5 Transposase (Jin at al., 2015). These transposases insert illumine adapters after cutting the DNA fragments. ATAC-seq is more sensitive and accurate than DNase-seq and libraries of ATAC-seq can be generated from a minimal number of cells within two hours (Klemm et al., 2019). To measure accessibility through

micrococcal nuclease sequencing, is also adopted widely due to its both exo and endonuclease activities (Jin W et al., 2015). MNase-seq varies from ATAC-seq and DNase-seq in efficiency to cleave nucleosomal DNA (Klemm et al., 2019). MNase cleaves nucleosomal DNA through its endonuclease activity and again these cleaved DNA products are degraded due to exonuclease (Allan et al., 2012). Further, the use of methyltransferase in NOMe-seq could able to provide information regarding both chromatin accessibility and methylation status. Methyl Transferase from M. CviPI generated ectopic methylation at GcP sites throughout the genome (Krebs et al., 2017). Though all are very unique in their aspects but provide similar genome-accessiblility information.

Chromatin accessibility and transcription factors binding are broadly dependent upon nucleosome density and turnover rates. Low nucleosome occupancy regions are generally observed with active transcription start sites and CTCF-bound insulators (John et al., 2011). Generally, larger nucleosome depleted regions have been observed with active gene promoters compared to the inactive ones. It has been seen factors like BRG1, RNA polymerase II, CTCF and ATP-dependent chromatin remodelers are largely associated with nucleosomal exclusion (Gilchrist et al., 2010). Linker histones also have a critical role in maintaining heterochromatin by neutralizing the charge and changing the exit angle between histone and DNA. Massive decondensation inside the genome was observed upon depletion of H1 Which is also seen in transcriptionally active chromatin (Nalabothula et al., 2014). Decompaction status even depends upon the acetylation of histone and shorter DNA linkers. Due to the high specificity of TFs in DNA binding, it majorly modulates chromatin. There is always a competition between histone and TFs and it depends upon the availability of concentration of TFs (Swinstead et al., 2016). TFs can bind to nucleosome bound DNA during nucleosome turnover. TFs like, C/EBP involve chromatin accessibility at distal regulatory regions and also found to be involved in the reprogramming of induced pluripotency stem cells by proving access to pluripotent induced genes (Di Stefano et al., 2016). A hormonally regulated TF, like glucocorticoid receptor (GR) massively modulates the chromatin. GR has extensive binding sites throughout the genome including inaccessible heterochromatin region. This decompacted chromatin becomes accessible upon hormonal treatment and available for GR binding (John S et al., 2011, Klemm et al., 2019).

Chromatin condensation

In general., the heterochromatin region is classified into facultative and constitutive. Facultative heterochromatin depends upon developmental cues and upon signal can be transformed into

euchromatin. On the other hand, constitutive heterochromatin is composed of repetitive elements and transposons and maintains its stability throughout differentiation (Saksouk et al., 2015). Sequence-specific DNA binding protein-like, HP1 and long non-coding RNAs recruit deacetylases (HDACs) and H3K9 methyltransferases (HMTs) at the nucleation centre where initiation of heterochromatin happens (Bulut-Karslioglu et al., 2012). These HDACs and HMTs involve in hypo-acetylation and hyper-methylation of H3K9 (Zhang et al., 2008). Once H3K9 is methylated, it further recruits HP1, HMTs, and HDACs and spreads heterochromatin (Wang et al., 2007). Recruited H3K9 heterochromatin domain in parents cell can able to retrieve epigenetic memory in daughter cells and maintain heterochromatin domains (Audergon et al., 2015). During the S phase of cell division, DNA polymerase opens up the heterochromatin when heavy transcription of repetitive DNA occurs (Chen ES et al., 2008). RNA-dependent RNA polymerase complex converts these single stranded RNA to double and later these are cleaved by DICER to produce siRNA (Sugiyama et al., 2005). These siRNAs are loaded into Argonaute protein 1 to form RNA inducing transcriptional silencing complex that targets the CLRC complex and initiates H3K9 methylation (Bayne EH et al., 2010). Monomethylated H3K9 was found in the pericentromeric region which is maintained due to Prdm3 and Prdm16 (Pinheiro I et al., 2012). Similarly, di and tri-methylation of H3K9 is mediated by SUV39H1 and SUV39H2 respectively in this region. At pericentromeric heterochromatin, the availability of H3K20me3 is quite high that is initiated by H3K9me3 through SUV420H (Lachner et al., 2001). Except for H3K9me and H3K20me, pericentromeric heterochromatin is also enriched with H3K27me1 and H3K64me3 (Peters AH et al., 2003, Lange UC et al., 2013). The Pericentromeric region is also enriched with histone variants like ATRX, DAXX, and SSRP1 (McDowell TL et al., 1999, Ishov et al., 2004, Lewis et al., 2010).

Majorly transcription is controlled by nucleosomes. Active gene promoters are devoid of nucleosome or demodulated by ATP-dependent chromatin remodellers (Helbling et al., 2009). Centromere and telomere regions are the only constitutive heterochromatin in lower eukaryotes. It's necessary to maintain genome stability. But, in higher eukaryotes except for centromere and telomere many repetitive and non-coding regions are also involved. Some chromatin regions are interchanged between heterochromatin and euchromatin state, known as facultative heterochromatin (Trojer et al., 2007). Generally, decondensation of facultative heterochromatin depends upon factors that change cell fate developmentally, chromatin reorganization in the nucleus, and monoallelic gene expression (Trojer et al., 2007). The facultative heterochromatin state is maintained by histone H1. It can be localized with only a few nucleosomes near inactive gene promoter regions (Albright et al., 1979). H1 is loaded onto chromatin by trans-acting –factors

or chromatin-modifying factors like histone deacetylase (HDAC) and SIRT1. L3MBTL1, a trans-acting chromatin-binding factor interacts with methylated H1 in order to compact the chromatin (Trojer et al., 2007). The inactive X chromosome is an example of facultative heterochromatin that gets reactivated during the blastocyst stage of an embryo and undergoes inactivation before implantation (Boroviak et al., 2017). The selection between both the X chromosome for inactivation is majorly relied on the concentration of chromatin modifiers and trans-acting elements that leads to hypoacetylation and hypermethylation in histone markers like, H4K20me1 and H3K27me3 (Heard et al., 2005). These changes in histone marks initiate transcription of long non-coding RNA Xist which coats the entire inactive X chromosome (Brown et al., 1991). Inactivated X chromosome maintains a ratio of canonical H2A and macroH2A. MacroH2A increases the contact frequency between internucleosomal regions and stabilizes facultative heterochromatin (fHC) (Changolkar et al., 2002). The establishment of fHC also depends upon polycomb group proteins, a component of the PRC1 complex. RING1B and H2AK119ub1 recruited along with other PRC1 members to facilitate fHC formation (Fang et al., 2004, Dorigo et al., 2003). Methylated H3K27 is a marker for facultative heterochromatin. H3K27me3 is localized with H3K4me3 in a poised gene when the gene activates or inactivates upon differentiation (Guenther et al., 2007). Facultative heterochromatin regions are carried out through the cell cycle and maintained in daughter cells with the help of PcG proteins (Trojer et al., 2007).

Replication Timing

S phase of the cell cycle is very crucial during the replication of the entire chromosome. After cell division chromatin maintains its 3D organization and epigenetic modification in daughter cells. These are established in the G1 phase and maintained throughout interphase. FISH studies had shown that chromosomes are arranged to a confined territory even after the cell cycle (Cremer et al., 1993). It has been observed that chromatins near to nuclear periphery and around the nucleolus are late-replicating and chromatin at the nuclear centre are early replicating (Solovei I et al., 2009). Variation in replication timing also had been observed from principal component analysis of a Hi-C datasets that highlighted early replication in compartment A (active regions) and late replication in compartment B (inactive regions). Lamin Associated domains (LADs) came under compartment B and are also associated with late replication (Ryba et al., 2010). Though LADs do not have an origin of replications but LADs regions are overlap with early replicating domains from where replication forks spread rapidly (Kind et al., 2013). TAD regions are associated with early replication control elements (ERCEs), cis-regulatory elements for early replication. ERCE regions are binding sites for master transcriptional regulatory factors and deletion of it affects the

TAD architecture (Rao SSP et al., 2017). The involvement of non-coding RNAs (ncRNA) can't be denied concerning replication timing. Delay in replication timing of inactive X chromosome was observed due to Xist but the deletion of Xist resulted in more delay that is even after later in the S phase (Diaz-perez et al., 2006). Deletion of ncRNAs like ASAR6 and ASARA15 delayed replication timing in their respective binding regions (Donlet N et al., 2015).

The timing decision point (TDP) appears in the early G1 phase and established a replication program. TDP initiates the selection program of replication origin sites, called the origin decision point. In the case of force replication before TDP or between TDP, origin decision point lead to random replication in the genome (Lu J et al., 2010). During TDP, chromatin interaction and TAD boundaries are re-established and this is confirmed with single-cell Hi-C and 4C technologies (Dileep V et al., 2015, Nagano T et al., 2017). It has been observed in Hi-C that chromatin interactions are still preserved in the G2 phase (Dileep V et al., 2015). Replication timing and compartmentalization are indirectly related to each other. Boundaries of TADs and compartments are weaker in the S phase and strongly established during early G1 upon removal of the replication timing program (Nagano et al., 2017). Further, the replication timing program will be established once TADs and compartments are established. Replication timing also varies between alleles of a gene. Even replication can delay due to chromatin silencing factors and DNA methylation. H3K9me2 is strongly correlated with late replication. Instead, H3K4me1, H3K4me2, H3K4me3, H3K20me1, H3K36me3, H3K9ac, and H3K27ac are associated with early replication (Claire Marchal et al., 2019). Replication domains are temporally regulated and show variation during cell fate transitions. But, in embryonic stem cells replication timing is less correlated with compartments. Even TADs structures are not correlated with replication timing during development. Formation of TADs can be observed during the four-cell stage of the embryo but spatiotemporal patterning of replication was observed before that (Dileep V et al., 2019).

DNA methylation

One of the pre-transcriptional modifications in the genome that regulates transcription is different types of DNA methylation. There are two major DNA methyltransferases, one is DNMT3A and the other one DNMT3B (Okano et al., 1999). In germline cells, the presence of DNMT3L stimulates the activity of DNMT3A and DNMT3B (Ooi SK et al., 2007). During replication, symmetrical CpG methylation has been observed. DNA methylation is directly correlated with CpG density in the gene promoter but, the binding of a transcription factor is indirectly correlated. Specific types of DNA methylation also contribute to the maintenance of heterochromatin. Five

methyl-CpG binding proteins exhibit in mammals, those are MBD1, MBD2, MBD3, MBD4, and MeCP2. Expect MBD3, the other four bindings to CpG depend upon CpG methylation (Baubec T et al., 2013). They cause gene silencing by interacting with histone deacetylase and nucleosome remodelling complex. CGI promotors occupied two-third of gene promoters in the mammalian genome, all house-keeping genes are included among them. Generally, CGI promoters are unmethylated or maintain gene silencing by H3K27 methylation through polycomb recessive complex 2 which is a more flexible mode of gene silencing (Marasca et al., 2018).

Majorly stable DNA methylation had been observed under germ-line specific genes, inactive X chromosome, and imprinted genes. DNMT3B and SMCHD1 are required for X-linked CGI promoter silencing in mice (Gendrel AV et al., 2012, Gdula et al., 2019). DNA methylation in parental germlines establishes differentially during early embryogenesis. 20 imprinting control regions (ICRs) had been identified in mice and the human genome. These ICRs are CpG rich CGIs and being methylated during oocyte. ICRs force mono-allelic expression for their neighbouring genes (Proudhon et al., 2012). DNMT3L methylates expressed genes in oocytes during their growth phase through DNMT3A. Parental ICRs are methylated at the TGCCGC sequence motif during gametogenesis and maintained in postfertilization embryos. This specific sequence motif is identified by KRAB-ZFP57 recruits silencing factors like KAP1, ZFP445, and DNMTs. DNMT3B and PRC1.6 are involved in DNA methylation in germline-specific genes (Li X et al., 2008, Takahashi et al., 2019). L3MBTL2 also interacts with H3K9 methyltransferase to pursue germline-specific methylation (Greenberg et al., 2019).

Gene regulation at transcription level:

Transcription factors

Though every cell of the body restrains exactly similar genomic content but not all the cells show similar gene expression. It is being tightly regulated at the chromatin level but still, regulations are established at the transcription level. Regulation of gene transcription is majorly taken care of by transcription factors, a specific DNA sequence binding proteins. TFs can control gene expression through direct binding as a monomeric form to cis-regulatory elements like a promoter, enhancer, and silencer or indirectly through dimerization with other TFs (Mitsis et al., 2019). Post-transcriptional modifications like phosphorylation, acetylation, methylation and glycosylation occur due to external stimuli that modulate TFs' stability, localization, and interaction with the cofactor. This activity leads to a change in gene expression by directly affecting enhancer-promoter

interactions. Also, TFs binding is affected by post-transcriptional modification of histones. TFs with MAD orthologues, like SMAD and pMAD, involve in cell fate transition due to binding with temporal-specific enhancers (Spitz et al., 2012).

Temporal regulation is generally associated with several binding sites and the concentration of TFs. It's not always the availability of TFs but also depends upon chromatin accessibility due to spatiotemporal arrangement. Cooperative binding between TFs is often associated with enhancer activation during development and first observed with phase lambda cI repressor (Spitz et al., 2012). Assisted binding of cooperative TFs through the collaborative competition to the same enhancer can modulate nucleosomal repositioning. Cooperativity between two TFs also means the formation of a DNA loop by one TF, such that a DNA binding site can be available to another TF. The binding specificity of TFs can also vary with the availability of co factors. TFs binding depends upon chromosome accessibility which is modulated by nucleosome repositioning and chromatin remodelling. For example, GR binding in mammals depends upon chromatin remodellers like SWI/SNF and BRG1 (Vicent GP et al., 2009). Upon enhancer activation through TF during development can influence a post-translational modification in histone tails like H3K4me1, H3K27ac, and H3K79me3, within nearby nucleosomes (Creyghton et al., 2010, Bonn S et al., 2012). In a developmental context, the existence of pioneer TF has been reported. Pioneer TFs like, PAX5, FOXA1, MYOD1, and PU.1 can bind to inaccessible DNA and recruit chromatin remodellers that repositioned nucleosome (Lupien et al., 2008). Recent studies on embryonic stem cells revealed several methylations protected enhancer which has subsequent developmental specific roles. For example, a core transcription factor of ESC, SOX2 binds to many enhancers to maintain pluripotency but, later during development replaced with SOX3 and SOX11 to differentiate to neuron and SOX4 during B cell development. The multiplicity of enhancers also observed under embryonic development. Generally, the primary enhancer refers to the nearest enhancer to a gene promoter and the secondary one is the distal enhancer. Both the enhancers show similar spatiotemporal activities. A secondary enhancer provides robustness to the deterministic gene expression irrespective of any environmental fluctuations (Spitz et al., 2012).

Histone modifications

A nucleosome is a functional unit of chromatin that is composed of four core histones protein as, H2A, H2B, H3, and H4. Modifications in these histone proteins impact chromosome compaction, nucleosome stability, and gene transcription (Zhao et al., 2019). Higher-order chromatin architecture depends upon histone modifications in H3K9 and its functional association with

lamins. Gene activation depends upon the accessibility of respective promoters that can be formulated by H3K9 acetylation and H3K4me2. These two can decompact chromatin such that actively transcribed genes can come out from the condensed region (Bartova et al., 2008).

N terminal end of histone proteins which are protruded from nucleosome complex is loaded with post-translational modifications like acetylation, methylation, phosphorylation, ubiquitinylation, sumoylation, ADP ribosylation, deamination, propionylation, and butyrylation (Kouzarides et al., 2007). Histone modifications like H2AS1P, H3T3P, H3S10P, H3T11, and H4S1P are involved in chromatin assembly during mitosis (Barber CM et al., 2004, Rea S et al., 2000). Among this H3S10P is also associated with meiosis. H2AK119P and H2BK120uq are shown involvement during spermatogenesis (Baarends WM et al., 2007). Ubiquitinylated and methylated histones like H2AK119uq, H3K9me3, H3K27me3, H4K20me1, and H4K20me3 are responsible for transcriptional silencing (Rea S et al., 2000, Zhang K et al., 2002, Wang et al., 2004, Nishioka et al., 2002, Schotta G et al., 2004). Among these H4K20me3 mediates heterochromatin formation, whereas H3K27me3 has a wide role in X chromosome inactivation and gene poising (Zhang K et al., 2002). Many acetylation and mono-methylation marks have been observed with transcriptional activation as, H2K4/5ac, H2AK7ac, H2BS33P, H2BK5ac, H2BK11/12ac, H2BK15/16ac, H2BK20ac, H2BK123uq, H3R17me, H3K4ac, H3K9ac, H3K14ac, H1K18ac, H3K23ac, H3K27ac, H4R3me, H4K5ac, H4K8ac, H4K12ac, and H4K16ac. Some of them are involved in the DNA repair mechanism as, H3K14ac, H1K18ac, H3K23ac, H4K5ac, H4K8ac, H4K12ac, and H4K16ac. H4K12ac is also known as a telomeric silencing factor (Lawrence et al., 2015). Transcriptional elongation regions are enriched with H3K4me3, H3K36me3, and H4K8ac. Similarly, H2BS14P modification is reported in apoptosis (Cheung WL et al., 2003).

RNA Methylation

In early 1970s, methylation at the 6th N position of adenylate of RNA was found in mRNAs and long non-coding RNAs which is well known as m6A (Desrosiers et al., 1974). Later more than 150 different kinds of RNA modifications were discovered. RNA methylation proteins can be divided into three different kinds, such as methyltransferase, demethylase, and RNA methylation recognition protein (Zhou et al., 2020). m6Am is a special modification found in first and second nucleotides behind 5' cap m7G (N7 methylguanosine) of mRNAs. These nucleotides are methylated at both N6 and 2' the hydroxyl group of adenylate (Keith et al., 1978). PC1F1 and FTO are special methyltransferases and demethylase of m6Am respectively (Sun et al., 2019, Liu

et al., 2020). m6Am modification provides more translation efficiency and stability by weakening DCP2 mediated decapitation of mRNA (Mauer et al., 2017).

m6A methyltransferase consist of METTL3, METTL14, WTAP, VIRMA, ZC3H13, RBM15/15B, HAKAI, etc. METTL3 has a catalytic role (Liu et al., 2014). METTL14 is responsible for RNA substrate binding, m6A methylation at 3'UTR of mRNA and cell localization. Others provide stability to the complex (Zhou et al., 2020). METTL16 is an independent methyltransferase that is responsible for the methylation of U6 snRNA (Pendleton et al., 2017). Likewise, METTL5 and ZCCHC4 methyltransferases are responsible for m6A modification in 18srRNA and AAC sequence of A4220 on 28srRNA respectively (Ma et al., 2019).

Coming to m6A demethylase, FTO/ALKBH9 is the first identified m6A RNA demethylase from alpha-KG dependent ALKB family of dioxygenases in 2017 (Mauer et al., 2017, Jia et al., 2011). The function of FTO is largely dependent upon its localization whether in the nucleus or cytoplasm. In the nucleus, FTO involves in the demethylation of m6A of poly-A RNA, m6A and m6Am of snRNA, and m1A of tRNA. FTO acts as a demethylase of m6A and m6Am of poly-A RNA and m1A of tRNA in the cytoplasm. FTO is also associated with the demethylation of m6Am snRNA that affects alternative splicing (Mauer et al., 2019). Another demethylase, ALKBH5 demethylates m6A in the form of Fe II and alpha-ketoglutarate (Aik et al., 2014). ALKBH5 has a major role in brain development and mostly found in the nucleus of adult neurons, especially from the cerebellum and olfactory bulb region (Du et al., 2020).

RNA methylation recognition proteins are basically of three different kinds. Proteins with YTH domains YTHDF1, YTHDF2, YTHDF3, YTHDC1, and YTHDC2 are the first type that binds to m6A through the YTH domain (Du et al., 2020). YTHDF2 recognizes the m6A site through its C-terminal domain. YTHDF2 facilitates m6A modified RNA degradation by recruiting CCR deadenylase complex through its N-terminal binding to SH domain of CCR4 complex subunit 1. YTHDC1 helps in the binding of m6A modified mRNA to RNA nucleoprotein complex and also responsible for mRNA nucleation. YTHDC2 seems to be involved in RNA degradation and translation (Zhang et al., 2018). The second type is heteronuclear ribonucleoproteins that include HNRNPC, HNRNPG, and HNRNPA2B1. Heteronuclear ribonucleoproteins regulate RNA substrates maturation in the nucleus. Splicing events in secondary RNA structure due to m6A are recognized by HNRNPC. HNRNPA2B1 interacts with DGCR8 by recognizing m6A marker in pri-miRNA and facilitates miRNA biogenesis (Zhao et al., 2017). The third types are insulin-like

growth factor 2 consists of IGF2BPs, YTHDF1, YTHDF2, YTHDF3, YTHDC1, and YTHDC2. IGF2BPs increases translational ability and mRNA stability of m6A modified mRNA (Huang et al., 2018).

Gene regulation at post-transcription level:

Alternative Splicing

The concept of alternative splicing was introduced 7 years before the discovery of exon and introns, which is in the 1970s by Walter Gilbert. 5% of genes in eukaryotes were predicted previously to be involved in splicing events. 2000 onwards through high-throughput sequencing, complete mRNA information had captured via expressed sequence tags (Sharp PA et al., 1994). Only 10% exonic sequence of mRNAs are processed otherwise, 90% intronic sequence is removed through splicing (Stamm S et al., 2005). Alternative splicing is mediated by spliceosome which is an association of many proteins along with five smaller nuclear RNAs. Spliceosome cleaves the phosphodiester bond in between exons and introns but formulates a phosphodiester bond among exons. Transcription, including 5' capping, 3' polyadenylation, and nuclear expert processes are being executed while splicing regulatory proteins are attached to pre-mRNA (S Stamm et al., 2008, MJ Moore et al., 2005). The co-transcriptional the event was explained by genome-wide high-throughput analysis of chromatin-associated RNA fraction. There is two major class of splicing regulatory proteins as, heterogeneous nuclear ribonucleoproteins (hnRNPs) and SR proteins. But, seven different modes of splicing events were described until now (M Hiller et al., 2007). Those are cassette alternative exon, alternative 5' splice sites, alternative 3' splice sites, intron retention, mutually exclusive alternative exons, alternative promoter, and first exon and alternative poly-A site and terminal exon. Spliceosome assembly each time formulates removal of one intron. splicing regulatory proteins bind to single-stranded pre-mRNA with low specificity and mediate RNA-RNA interaction at 5' splicing site (Roy et al., 2012).

Alternative splicing also influenced by chromatin architecture, as we know changes in nucleosomes can change the chromatin accessibility which leads to a change in transcription state. Generally, introns with low GC content compared to exons with high GC contents and intronic sequence flanking to splicing sites are less occupied with nucleosomes. Also, more nucleosome occupancy had been observed with constitutively spliced exons and exons with weaker splice sites (Amit M et al., 2012, Tilgner H et al., 2012). Nucleosomes in exons affect transcriptional elongation that defines exons and also it alters splicing factors in speckle compartments (Schor IF et al., 2012).

Chromatin remodellers like, SWI/SNF, ISWI, CHD and INO80 are directly involved in spliceosome regulation (Clapier CR et al., 2009). SWI/SNF influences exon recognition by interacting with spliceosome proteins like U5 snRNP, Sam68, and hnRNP. CHD1 associates with U2 snRNP and SF3A and also regulate RNA polII elongation by interacting with FACT and PAFc (Batsche E et al., 2006). But, U1 snRNP binding at 5' end and overexpression of SRSF2 and SRSF1 deplete nucleosome positioning (Keren-Shaul H et al., 2013). Histone modifications as, H3K36me3, H3K79me1, H2BK5me1, H3K27me1, K3K27me2, and H3K27me3 are more correlated with exons (Spies N et al., 2009, Schwartz S et al., 2009). H3K4me3 directly involves in mRNA maturation by bridging CHD1 with spliceosome components (Sims RJ et al., 2007). Hyperacetylation of histones and H2B monoubiquitylation increases RNApolII elongation and decreases exonic insertion. Not only chromatin modulators but also alternative splicing is directly influencing epigenetic markers and changing chromatin architectures (Naftelberg et al., 2015).

RNA binding proteins

RNA binding proteins (RBP) bind to RNA in a sequence-specific manner to form ribonucleoprotein complex which is involved in RNA stability, biogenesis, and cellular localization. Ribonucleoprotein complexes play a major role during pre-mRNA splicing, polyadenylation, and transportation. During spliceosome, small nucleolar RNA is folded three-dimensionally so that protein binding sequence will be available at the surface for RBPs. Recently, lncRNA and nascent transcripts binding to chromatin-associated factors and DNA binding proteins have been identified. During the formation of paraspeckles lncRNA, NEAT1 sequestrates required proteins (Clemson CM et al., 2009). RNA binding proteins are sometimes metabolic enzymes. A good example is thymidylate synthase, who converts dUMP to dTMP, and during deficiency of dUMP, it binds to its mRNA to inhibit translation. Massive regulation of RBPs had been observed during mouse embryonic stem cell differentiation (de Breyne et al., 2009). In mESC, RBPs are involved in MYC-dependent cell fate transition. Induction of induced pluripotency stem cells from mouse embryonic fibroblasts upregulates many RBPs (Kwon SC et al., 2013).

Recently, many techniques have been evolved to interrogate RBPs and RNP complexes. Through RNA interactome capture (RIP), 860 and 791 RBP had identified in human HeLa and HEK293 cells (Baltz et al., 2012, Castello A et al., 2012). RIP includes UV cross-linking followed by RNP complex capture by targeting polyadenylated tail and then mass spectroscopy. Picked up RBPs from RIP was further validated with cross-linking immunoprecipitation (CLIP),

immunoprecipitation of a GFP-RBP fusion protein, reverse transcriptase followed by real-time PCR, high-throughput sequencing, and detection with a fluorescent oligo-dT probe (Strein C et al., 2014, Matia-Gonzalez et al., 2015). A comparison between different RNA binding and whole proteomes suggests three different classes of RBPs. 1015 RBPs come under class 1, whose RNA binding efficiency doesn't change during maternal to zygotic transitions. Similarly underclass 2, 78 RBPs are reported, those showed changes in RNA binding due to variation at the differential gene expression level. Class 3 are having 38 RBPs and showed variation during RNA binding but, there was no change observed at the gene expression level. Further to dissect novel RBPs, high throughput methods had emerged like RNA Immunoprecipitations (RIP) or its derivatives like cross-linking immunoprecipitation (CLIP), photoactivatable-ribonucleoside-CLIP (PAR-CLIP), individual nucleotide CLIP (iCLIP) and enhanced CLIP (eCLIP) (Matthias W et al., 2018).

Long non-coding RNA

The regulatory mechanism of the eukaryotic transcriptome is much more complex than the massive genome in the nucleus from where it gets transcribed. Evolutionary complexity did not evolve based on some genes but the mechanism of gene regulation. Transcription is the initial step of the central dogma, and it decides the stage of the cell. That's why it is conducted in a highly controlled manner during pre and post-transcriptional reactions. It involves three major phases, which are started with the recognition of promoter to elongation of transcripts followed by termination. In eukaryotes, in between these two steps another phase is present, that allow the binding of transcriptional machinery to the promoter, but keeps it in a pause state. It helps the cell to express the desired gene at instant stimuli. Binding of transcription machinery to open promoter triggers the process of transcription and as soon as the nascent primary transcript emerges from RNA Polymerase II, 5' end of the transcript is protected by 5'cap. Another stretch of protein cuts the nascent RNA from moving RNA PolII and starts adding adenine to 3' end, called polyA tail. 5' capping initiates after transcription of 20nts fragment. It acts as a recognition signal for eIF-4E, which is crucial to start the translational process by the ribosome. A hexanucleotide, as AAUAAA is reported at 3' UTR region of mRNA that is around 10-30nts upstream of the cleavage site of pre-mRNAs, is identified for 3' end processing. PolyA tail is added by template-independent poly(A) polymerases (PAPs) as a post-translational modification. Comparatively long transcript stables for 30 minutes to 1 hour. Short tails are easy to be degraded with the help of exosomes by exonucleolytic degradation.

3.2gb length of the mouse genome only encodes 30% coding regions that can give rise to protein and follow the above procedure. Remaining, thought to be junk DNA, in the previous era that only increase the chromatin length and space in the nucleus without any functional aspects. But, the ignorance of them, showed us lethality in a complex eukaryotic cell, without proper subsistence. In the 1970s, interest emerged for the importance of noncoding/Junk DNAs regions at the post-transcription level and rRNA and tRNA were discovered (Holmes et al., 1972). Gradually during the 1980s, the main player of post-transcriptional modifications was discovered which are small nuclear RNA and small nucleolar RNA and that followed by long non-coding RNA (lncRNA) in the 1990s (Brannan et al., 1990).

LncRNAs have many similarities with mRNAs as it also transcribes by RNA PolII and headed for poly-A tail addition followed by splicing (Brockdorff et al., 1992). The length is around 1kb, found in poorly conserved regions. H3K4me3 and H3K36me3 are identified histone marks for lncRNAs that had been observed with 5000 lncRNA from the human genome (Brown et al., 1992). Some well-known examples are Xist, NEAT1, MALAT1, HOTAIR, and H19. The extensive role of lncRNA can be figured out with X chromosome inactivation that has been observed with therian XX mammalian female. The high expression of Xist has marked with an inactive X chromosome rather than an active one (Lee et al., 2011). Xist coats inactive X chromosome entirely and form Xist cloud on to which silencing factors like polycomb repressive complex 2 are recruited (Brown et al., 1992). Tsix, another lncRNA is antagonistic to Xist and prevents Xist from binding to the active X chromosome (Xu et al., 2006). HOTAIR, a lncRNA guides polycomb repressor complex 2 to the required location during development (Rinn et al., 2007). Base on its structure, function, localization, metabolism, and interaction pattern, lncRNA is classified into sense, antisense, bidirectional, intronic, and intergenic. It can compete with DNA binding proteins to control particular gene transcription and can recruit epigenetic modifiers into a genome. LncRNA also serves as a miRNA precursor (Kung et al., 2013).

Gene regulation at translation level:

Specialized ribosomes

The eukaryotic ribosome consists of four ribosomal RNA and 79 ribosomal proteins. Ribosome biogenesis needs significant energy from cells and a highly coordinated process. It takes only 60ms to elongate a polypeptide chain. Around 2000 genes are encoded as ribosomal protein genes in mammals. Among them most of the genes are pseudogenes. Single genes are coding for important

ribosomal proteins, unlike yeast and plants where one ribosomal protein has two active genomic copies (Balasubramanian et al., 2009, Kellis et al., 2004). Deletion of one of them results in differences in cell phenotypes (Ni et al., 2001). In yeast, splicing events in ribosomal proteins can regulate both paralogue genes. Removal of intron during splicing of S29A/Rps29a ribosomal protein can reduce the expression of both of its paralogue genes. Around 70% of paralogue genes in yeast are regulated asymmetrically (Parenteau et al., 2011). They are not functional substitutes for each other. Two to seven paralogue genes have been observed in plants to single ribosomal proteins.

But, ribosomal proteins in mammals are represented through a single copy of genes. An exception can be seen with RPS4, having three paralogues such as, RPS4X, RPS4Y1, and RPS4Y2. These are present in the X and Y chromosomes. RPS4Y2 is expressed in the testis and prostate gland in human males. Whereas, RPS4X and RPS4Y1 are ubiquitously expressed in males. RPS4Y2 has a distinct carboxyl terminus that showed unique interaction with ribosomal factors of testes (Lopes et al., 2010). Similarly, in mice RPL10, RPL22 and RPL39 have paralogue genes. RPL10 and RPL39 are specific to testes and RPL22 is specific to the liver and mammary gland. RPL39 also showed localization to nucleoli with the 80S and polysome fractions (Sugihara et al., 2010). RPL38 has increased expression levels in developing somites and motor neurons specific to the spinal cord (Kondrashov et al., 2011).

Ribosomal proteins expressions were believed to be at an equimolar concentration in previous days, which is later got disproved with evidence of varied expression patterns in different state of life cycles (Xue et al., 2012). Many ribosomal proteins are post-translationally modified in a highly regulated manner. post-translationally modification may include phosphorylation, methylation, acetylation, and ubiquitylation. 11 large subunits and most of the small subunits are post-translationally modified in human (Odintsova et al., 2003). post-translationally modification can be done at Ser and Thr residue of the ribosome by adding O-linked ß-D-N-acetylglucosamine (Zeidan et al., 2010). Ubiquitylation of Rpl28 in *S. cerevisiae* was observed in the S phase of the cell cycle which was later reduced in the G1 phase. Polyubiquitylated Rpl28 can able to translate reporter gene faster than monoubiquitylated Rpl28 (Spence et al., 2000).

Along with ribosomal proteins many ribosomal-associated factors are also involved in ribosome activity. Glycogen synthase 1 is one of the ribosomal protein factors, associated with actively translating ribosomes and the polysomes may be compromised with its absence (Fuchs et al.,

2011). A scaffold protein RACK1 is associated with the 40S subunit and acts as a receptor for PKC (Protein kinase C) (Adams et al., 2011). Ribosome bound RACK1 facilitates phosphorylation of eIF6, a translation initiation factor with PKCßII (Ceci et al., 2003). RACK1 also facilitates recruitment of the miRISC complex to the ribosome (Jannot et al., 2011). Localization of ribosomes to the cell membrane was carried out by RACK1 include integrin receptor (Nilsson et al., 2004). mTORC2 interacts with ribosomes upon insulin signalling and facilitates mTORC2 independent translation and AKT substrate independent translation (Zinzalla et al., 2011, Oh et al., 2010).

Micro Proteins

Like microRNAs. Micro Proteins are small proteins with a single protein-protein interaction domain and can prevent their targeted proteins by forming homodimeric, heterodimeric, or multimeric complexes (Staudt et al., 2011). It inactivates its targets by forming heterodimeric complexes and can also deactivate its biological function by engaging it to different protein complexes (Graeff et al., 2012, Staudt et al., 2011). Micro Proteins target larger proteins by titrating interacting interaction partners of the targeted protein. The protein-protein interaction is either with an identical domain, known as homotypic miP inhibition, or with a nonidentical compatible domain, known as heterotypic miP inhibition (Eguen et al., 2015). The 16kDa, inhibitory of DNA binding micro Protein was discovered first. It has a helix-loop-helix a domain that interacts with proteins having a basic helix-loop-helix domain (Benezra et al., 1990). Transcription factors are mainly targeted by micro Proteins. Even in human micro Protein Vpu, sequestrates non-transcriptional factors like TASK1 with a non-functional protein complex and regulates K+ ion channel (Hsu et al., 2004).

Depending upon origin, micro Proteins are classified as *trans* and *cis*-micro Proteins. *Cis*-micro Proteins are products of alternative translation, splicing, and post-translational processing by proteolytic cleavage. Some of the small open reading frame (sORF) also can encode micro Proteins. On the other hand, *trans*-micro Proteins are evolutionary evolved through genome amplification and subsequent domain loss (Floyd et al., 2014). Micro Protein named, LITTLE ZIPPER negatively regulates HD-ZIPIII, a transcription factor by forming a heterodimer. LITTLE ZIPPER is also transcriptionally controlled by HD-ZIPIII (Bhati et al., 2017). In some cases, transcription factors are suppressed through the formation of homodimer by micro Protein till its requirement during certain stages of development.

Micro RNA

Historical Perspective:

The first time in 1993, the developmental role of lin-4 on the larval stage of C. elegans was discovered by Ambros et al. Instead of coding for a protein, lin-4 locus produces two small RNAs, of which, one is 22nts and another 61nts (Lee et al., 1993). The larger transcript was shown to have a step loop structure and a precursor for the smaller one. Later, multiple complementarities of lin-4 were noticed with the 3' UTR region of the lin-14 gene (Wightman et al., 1993). Further, it was shown that binding of complementary region of lin-4 to lin-14 caused its translational arrest. Because of its shorter length (22nts), it was named as microRNA (Lagos-Quintana et al., 2001, Lau et al., 2001, Lee and Ambros et al., 2001). Both RNA Polymerase II and III are involved in the transcription of miRNA. The majority of miRNAs fallen within the gene body, mainly from introns (mintrons). The remaining were intergenic and had their promoters. Few were transcribed in a cluster due to co-localization in the genome and having similar seed sequences, called a miRNA family.

Biogenesis:

MiRNA biogenesis largely follows canonical pathways. Transcribed primary miRNAs (pri-miRNA) were shown to be processed into precursor-miRNA (pre-miRNA) by cleavage through ribonuclease III enzyme, Drosha and leaves a 2nt overhang at the 3' site (Denli et al., 2004). This cleavage region was shown to be recognized by DiGeorge Syndrome Critical Region 8 (DGCR8) through the N6-methyladenylated GGAC sequence in pri-miRNA (Alarcon et al., 2015). All these above events were shown to be carried out within the cell nucleus and later processed pre-miRNAs are transported into the cytoplasm through exportin-5 (XPO5). In the cytoplasm, pre-miRNA is cleaved by RNase III endonuclease, Dicer, and produced two mature miRNA strands (Okada et al., 2009). Among them, 5p named after mature miRNA comes from 5' end of pre-miRNA and 3p for 3' end. In general, in most cases, 5p serves as a guide strand and 3p as passenger strand, that later cleaved by AGO2. Further, it was noted that the strands with lower 5' stability and the presence of uracil at 5' end are preferentially bounded by AGO (Brien et al., 2018). Non-canonical miRNA biogenesis was shown to be two different ways, some were Drosha/DGCR8-independent and others DICER-independent. In Drosha/DGCR8-independent pathways, pre-miRNA does not require to be processed through Drosha/DGCR8 and were directly exported to the cytoplasm by exportin 1. In this regard, mirtrons were spliced products of genes and came from intronic regions (Ruby et al., 2007). Nascent RNA, 7-methylguanosine capped pre-miRNAs were also

transported to the cytoplasm directly (Xie et al., 2013). Dicer independent mechanism was shown to be followed by endogenous short hairpin RNAs (shRNA), which were with insufficient length to be processed by Dicer. the entire double-stranded shRNA was shown to be loaded onto AGO2 and continued to maturation (Yang et al., 2010).

Structural Organization:

The functional complex of miRNA i.e miRNA induced silencing complex (miRISC), consists of mature miRNA strand and AGO protein (Kawamata et al., 2010). miRISC binding to 3' UTR region of mRNAs is based on sequence complementarity with miRNA, called miRNA response elements (MRE) (Jo et al., 2015). The degree of complementarity was shown to define, whether it undergoes mRNA degradation or translational arrest. Complete complementarity between RISC and MRE promotes mRNA decay. However, the majority of the miRNA-mRNA pairs showed central mismatches, except to their seed region, thus preventing AGO2 endonuclease activity. In the case of mRNA degradation, miRISC recruits GW182 which facilitate recruitment of PAN2-PAN3 and CCR4-NOT (Jonas et al., 2015, Bhem-Ansmant et al., 2006, Christie et al., 2013). Firstly, poly(A)-deadenylation is initiated by PAN2-PAN3 which is further efficiently executed by the tryptophan repeats of GW182 and poly (A) binding protein C (PABPC). Later, decapping protein 2 (DCP2), decap mRNA at 5' end followed by 5'-3' degradation by exoribonuclease 1 (XRN1) (Braun et al., 2012).

Mechanism of Action:

In order to understand the miRNA targets, many computational methods had been developed. But, because of few perfect complementary between miRNA and its targets in the case of animals make noisy and false-positive interpretations (Bartel et al., 2004). Because of which experimental validations are essential. On the other hand, complementarity is not only a factor that determines specificity, but also 3-dimensional structure of protein and mRNA may restrict miRISC binding. Despite the above constraints, complementarity with the seed region (6-8 nts) is most important. This short sequence complementarity implies multiple mRNAs being targeted by a single miRNA (Lewis et al., 2003). Though single miRNA can target many mRNAs, still miRNA circuitry is purely cell-type and developmentally regulated. Mode of action of miRNAs can be implemented as per the requirements i.e. if mRNA expression diminishes in a particular cell type that can be controlled immediately through miRNAs by target switching. For more customized expression of mRNAs, miRNAs can fine-tune the targets and maintain a uniform level. Some mRNAs were targeted naturally and the action of miRNAs can be suppressed by cells through a feedback mechanism.

Cells also evolve with anitargets to titrate miRNA actions, which shows complementarity to particular miRNA (Bartel et al., 2004). Considering these possibilities, we exploited the dynamic cell differentiation system at three developmental stages of animal development.

MiRNA in the context of Pluripotency

Pluripotent mouse embryonic stem cells (mESCs) were established first time in 1981 by Evans et al. and Martin et al. from inner cell mass (ICM) of the late blastocyst. The purity of pluripotency through the period of culture conditions was only validated through the generation of chimeric mice by injecting cultured ESC into the blastocyst. When human embryonic stem cells (hESCs) were isolated from ICM, they did not follow the same culture condition as mESCs. Also unlike mESCs, female hESCs cell lines showed inactive X chromosomes. But later in 2007, mESCs from ICM of post-implanted blastocyst of embryonic day 5.5 showed similarities with hESCs, which was rarely developed a chimeric mouse. This later stage of mESCs was termed as mouse epiblast-derived stem cells (mEpiSCs) and mESCs were designated as Naïve state of pluripotency. mEpiSCs exhibited inactive X chromosome and stated as a later stage of mESCs. Naïve state was isolated from ICM of pre-implated blastocyst of embryonic day 3.5 (Kevin et al., 2014). Later days, the Naïve state was started culturing in serum-free medium by adding 2i inhibitors along with LIF. 2i inhibitors are designated for two pharmacological agents GSK3ß and MAP2K inhibitors, which inhibit GSK3ß in Wnt pathway and MAP2K in ERK signaling respectively. Naïve ESCs cultured in this condition is termed as the Ground state of pluripotency. The ground state showed more resemblance with ICM of the pre-implated blastocyst in terms of transcriptome, epigenome, DNA hypomethylation, and genome-wide redistribution of H3K27me3 (Hackett et al., 2014).

Noticeable differences had been marked between mESCs and hESCs in terms of gene expression and colonization. Generally, hESC appear flattened instead of a dome, shaped, unlike mESCs. Also unlike mESCs, hESCs expresses SSEA-3 and SSEA-4 rather than SSEA-1. hESCs requires FGF/TGFß signaling to maintain their self-renewal but, mESCs need LIF/BMP4 (Wu et al., 2015). Naïve mESC show more resemblance with ICM from blastocyst of embryonic day 4.5 and mEpiSCs are closely related to mature E5.5 to E8.25. mEpiSCs depend upon FGF/TGFß signaling for its self-renewal (Kojima et al., 2014). Naïve mESCs could be incorporated into a pre-implantation embryo and showed colonization with ICM of blastocyst and chimera formation but, mEpiSCs failed to integrate, proliferate and differentiate (Wu et al., 2015). Post-implanted epiblasts in *in vivo* condition are intended to form primordial germ cells (PGCs) in presence of extra-embryonic tissues but, cultured mEpiSCs lose PGC competency (Hayashi et al., 2009). A transient

cellular state developed from the Naïve state which is more like epiblast (EpiLCs) and showed similarities with pre-gastrulation epiblast. EpiLCs are efficient to induce PGC (Aramaki et al.,2013) and to develop germ cells in vitro (Hayashi et al., 2011). The previous study has shown EpiSCs exhibit two distinguished cell population which was demarcated by GFP signals from 18kb the regulatory region of *the Oct4* gene (GOF18) (Han et al., 2010). Surprisingly, *Oct4*-GFP+ cells could integrate for chimera formation. Later, *Oct4*-GFP+ cells were stabilized in culture condition in the presence of FGF4 (Joo et al., 2014). Some group isolated intermediate epiblast like cells (IESCs) which responded to both LIF-STAT3 and Activin-SMAD2/3 signalling. IESCs could be efficiently incorporated into ICM but showed defects later in normal embryonic development (Chang and Li, 2013). Activation Wnt signalling in the Naïve state contributes to self-renewal and stabilization but, its inhibition rapidly shifts the Naïve state of ESC to Primed (Berge et al., 2011). IWP2, a porcupine inhibitor that blocks WNT signalling was used to stabilize EpiSCs. Interestingly, IWP2-EpiSCs could convert to its Naïve state and efficiently integrated to chimeric embryo formation.

Several studies also described culture condition for 2C-like (2 cell stage of zygote) cells which retains totipotency and could contribute to both embryonic and extra-embryonic lineages. These 2C-like cells did not show expression of pluripotent gene markers like *Oct4, Sox2*, and *Nanog* (Macfarian et al., 2012). Downregulation of CAF-1, which is responsible for chromatin assembly, facilitates chromatin reprogramming and induced 2C-like cells. But single cell RNA sequencing data of 2C-like cells showed more similarities with blastocysts rather than a *in vivo* two-cell stage of the embryo (Kolodziejczyk et al., 2015). In other hand, fraction of population in the Ground state may be functionally totipotent (Morgani et al., 2013).

Global DNA hypomethylation in 2i/LIF ESCs is a signature characterization of the Ground state, like ICM. Transcriptional silencing is mediated by methylation at CpG dinucleotides which is vital for the maintenance of genome integrity (Smith et al., 2013). During early embryonic development to establish cellular identity, DNA methylation (5mC) is stabilized throughout the genome and retained through cell divisions (Wu et al., 2015). Global DNA hypomethylation is crucial to remove epigenetic barriers against pluripotency (Hackett et al., 2013) but global DNA methylation was observed during lineage-restricted differentiation (Meissner et al., 2008). Naïve state of DNA hypermethylation with 5mC is 3 fold higher as compared to the Ground state of ESCs (Ficz et al., 2013) which showed more correlation with EpiSCs from E6.5. But, Ground state hypomethylation is more comparable with pre-implanted blastocyst from E3.5-E4.5. The methylation state in ESC

is very unstable (Shipony et al., 2014). XX ESCs can able to transit towards global DNA hypomethylation even in serum conditions as compared to XY ESCs (Schulz et al., 2014). Hypomethylation in XX ESCs was observed 3 fold more than XY ESCs in serum. It may be due to the presence of two active X chromosome that represses de novo methyltransferases and pERK activity (Hackett et al., 2013). Bivalency is a prominent feature in the ground state of ESC that includes activating histone marker H3K4me3 and repressive polycomb marker H3K27me3 on the promoter of developmentally important genes. H3K27me3 is globally reduced in the Ground state that may be due to inhibition of EED activity due to suppression of ERK signaling through PD inhibitor (Tee et al., 2014). Several epigenomic markers linked to chromatin decondensation are get activated in Ground ESCs while multiple repressive markers such as 5mC, H3K27me3, H3K9me2, and H3K9me3 were depleted. But, H3K4me3 showed collateral increment towards Ground transition. Surprisingly, decondensed chromatin does not promote precocious transcription. No transcriptional hyperactivity was observed neither in the Ground nor the Naïve state of ESCs. One possible mechanism to describe this phenomenon was stated as RNA polymerase II (Pol II) pausing at the proximal region of the promoter which was majorly observed under lineage commitment genes. RNA Pol II pausing at the proximal region of the promoter was observed at a greater extent in Ground ESCs in comparison to Naïve ESCs (Marks et al., 2012). RNA Pol II pausing in Naïve ESCs in presence of serum, is mediated by ERK1/2 by causing phosphorylation at CTD of developmental genes (Tee et al., 2014). But, ERK activity is blocked in the Ground state which is why unclear that which mechanism is responsible for increased transcriptional pausing (Wu et al., 2015).

Defects in embryonic development have been observed as a downstream result of disruption of miRNA processing enzymes (Murchison et al., 2005, Bernstein et al., 2003). Dicer and DGCR8 deficient mice showed abnormality in cell cycle and stem cell proliferation (Bernstein et al., 2003, Wang et al., 2007). Defects in the differentiation of stem cells also had been seen with the continuous expression of *Oct4, Nanog, Sox2*, and *Rex1* and reduced differentiation markers (Bodnar et al., 2004, Menendez et al., 2006). Exogenous miR-290 family can rescue self-renewal with the expression of *Oct4* in Dicer-null mouse embryonic stem cells (Sinkkonen et al., 2008). miR-290 cluster which comprises miR-290, miR-291, miR-292, miR-294 and miR-295 (Houbaviy et al., 2003) and its homolog miR-371 cluster which includes miR-371, miR-372, miR-373, and miR-373* (Suh et al., 2004) showed high expression in mouse and human embryonic stem cells respectively. Core regulatory factor of ESC such as *Oct4, Nanog* and *Sox2* have a binding site at the promoter region of the miR-290 cluster (Marson et al., 2008). That's how by targeting the inhibitor of *Oct4*, it can able to stabilize its expression along with a state of pluripotency. miR-290 cluster

also regulates differentiation in ESC by targeting epigenetic repressor RBL2, a DNMTs which suppresses OCT4 (Hayashi et al., 2008). *Oct4, Nanog, Sox2*, and *Rex1* are also the upstream regulator of the promoter of miR-302-367 cluster (Barroso-delJesus et al., 2008). miR-302-367 cluster comprises miR-302a, miR-302a*, miR-302b, miR-302b*, miR-302c, miR-302c*, miR-302d, miR-367, miR-367* (Landgraf et al., 2007). miR-302-367 cluster regulates cell cycle progression by targeting *Cyclin D1* and *Cdk4* (Card et al., 2008) and inhibits intermediate negative regulators of TGFß/Nodal/Activin pathway to maintain pluripotency (Barroso-delJesus et al., 2009). Another cluster miR-17-92 consists of miR-17, miR-18a, miR-19a, miR-20a, miR-19b-1, and miR-92a-1, which is activated by c-Myc, is highly expressed in an undifferentiated state of ESC (Houbaviy et al., 2003, Gu et al., 2008, Chang et al., 2008). C-Myc also directly represses let-7 in the undifferentiated state of ESC which facilitates LIN28B expression. Binding of Oct4, Nanog, Sox2, and Tcf3 in ESC facilitate a regulatory loop that activates LIN28 and suppresses let-7g to maintain self-renewal and pluripotency (Marson et al., 2008, Lakshmipathy et al., 2010)

MiRNA in the context of Neurogenesis

The carboniferous era marked its end with a magnificent developmental alteration within the amniotes regarding their brain size. Evolutionally, expansion of the dorsal telencephalon induces generation of the neocortex (Cardenas et al., 2018, De Juan-Romero C et al., 2015). The closure of the neural tube at embryonic day 10.5 of mouse results in its development into the forebrain, midbrain, and hindbrain. The forebrain gives rise to the telencephalon and the hindbrain to the myelencephalon. The cerebral cortex emerges from the telencephalic neuroepithelial cells. The cortical neurons and glutamatergic neurons arise from the dorsal telencephalon while GABAergic interneurons develop from the ventral telencephalon (Agirman et al., 2017). The asymmetric division of neuroepithelial cells produces self-renewing progenitor cells called radial glial cells in the ventricular zone. Thereafter, it generates one mother-daughter cell while the other daughter cell's fate is to commit for differentiation. The mouse neocortex is composed of 90% neural cells that comprise excitatory neurons, interneurons, astrocytes, and oligodendrocytes: and the remaining 10% are non-neural cells that consist of microglia and endothelial cells (Varrault et al., 2019).

In the developing cerebral cortex of the mouse, symmetric division of neural stem cells (NSC) maintains the stem cell pool in the subventricular zone. However, NSC undergoes asymmetric division on attaining stimulus for differentiation and follows a series of transit-amplifying

progenitors, interneurons, and neuroblasts to generate post-mitotic mature neurons. Intermediate progenitors from radial glial cells promote migration into dentate gyrus outward of neurogenic niche for further differentiation and maturation. However, there are short neuronal precursor cells in the ventricular zone that give rise to neurons via symmetric differentiative divisions (Tyler et al., 2015). Neuroepithelial cells, radial glial cells, and short neural precursor cells are collectively called apical progenitors (AP) (Agirman et al., 2017). The apical progenitors on encountering lateral ventricles assist in corticogenesis as they contain cerebrospinal fluid and diffusible morphogens. Radial glial cells serve as a common precursor for both neuronal and microglial lineages. The gliogenic switching of radial glial cells produces oligodendroglial progenitor cells (OPCs) that are the origin source for astrocytes and oligodendrocytes (Bergles et al., 2015).

Sonic-hedgehog (Shh) is a diffusible morphogen secreted from the ventral telencephalon, is involved in the development of forebrain (Agirman et al., 2017). Shh delocalizes Patched 1 (PTCH1), a 12 transmembrane receptor that activates a seven-transmembrane G-protein coupled receptor SMO (Smoothened homolog), thus transduces Hedgehog signaling proteins by downstream activation of transcription factors GLI2 and GLI3 to initiate transcription of GLI1. GLI1 controls the ventral telencephalon specification by enhancing Nkx2.1 expression (Niewiadomski et al., 2014, Agirman et al., 2017). It promotes the formation of oligodendrocytes, and GABAergic interneurons (Baudoin et al., 2012, Xu et al., 2020). At embryonic day 8.5, BMP (bone morphogenetic protein) and Wnt (Wingless-related proteins such as 2b, 3a, 5a, 7b, and 8b) are secreted from the dorsomedial telencephalon of the cortical hem. It aids in dorso-ventral and mediolateral telencephalon specification (Caronia-Brown G et al. 2014). Wnt ligands are glycoproteins that bind to lipoprotein receptors of apical progenitors and enhances the expression of the TCF/LEF transcription factor family and its downstream targets through ß-catenin (Cadigan KM et al., 2012). Wnt regulates adherens junction complexes and oligodendrogliogenesis of the radial glial population (Agirman et al., 2017). Repression of the Wnt pathway leads to downregulation of PAX6 expression, which as a result, depletes the AP population (Gan Q et al., 2014). Canonical Wnt signaling plays a prominent role in the self-renewal of apical progenitors and differentiation into cortical neurons during the early phases of corticogenesis. However, it promotes differentiation into oligodendrocytes in the later stages of development. The binding of BMP2 and BMP4 (of TGF-ß superfamily) to hetero-tetrameric complex receptors on apical progenitors leads to the phosphorylation of cytoplasmic R-SMADs (1, 5, or 8). It binds to co-SMAD and recruits other factors essential for the transcription of genes that are key for cortical development (Bond AM et al., 2012) but later produces oligodendrocytes. Fibroblast growth

factors (FGF) are secreted from the anterior neural ridge of the telencephalon at embryonic day 9.5, which supports the self-renewal capacity of the radial glial cells. A remarkable reduction in the cortical surface area can arise due to depletion in FGF receptors in the dorsal telencephalon (Agirman et al., 2017). Notch signalling is predominant in the radial glial cells (RG) as they express receptors for Notch 1 and Notch 5. NICD translocation and binding to CBF1 in the nucleus initiates the transcription of Hes, which is a crucial player in the maintenance of stemness of RG and suppressor of proneural gene expressions (Kageyama et al., 2008). Therefore, the synchrony of these signalling pathways is paramount for diverting neural progenitors towards corticogenesis.

The intricate process of neurogenesis is recapitulated in vitro by using pluripotent stem cells like embryonic stem cells (ESC) derived from inner cell mass (ICM) of the developing blastocyst as well as induced pluripotent stem cells (iPSCs) by differentiating into neurons. ESCs tend to proliferate and differentiate in the absence of extrinsic factors and start expressing neuronal markers, which recapitulate forebrain identity (Gaspard et al., 2009, Juliandi et al., 2010). The addition of inhibitors of BMP and Wnt pathways to the culture medium has the potential to direct the cell lineage to progress that has a telencephalic cortical cell identity (Kirkeby et al., 2012). The inhibition of Shh signalling by the addition of cyclopamine leads to the differentiation of cells towards the dorsal-telencephalic progenitors that mostly give rise to glutamatergic pyramidal neurons (Ameele et al., 2014). However, differentiation into ventral-telencephalic progenitors i.e. GABAergic interneurons requires both Shh and Wnt signalling. FGF8, FGF15, and activin collectively serve as external cues for the conversion of ventral telencephalic progenitors into anterior and posterior regional patterning interneurons (Danjo et al., 2011, Cambray S et al., 2012). A fully developed neocortex displays six layers of different neurons; during differentiation and maturation, the deep layer emerges early and the upper layer in later stages (Gaspard et al., 2008, Greig et al., 2013). In vitro differentiation of ESC can give rise to all the layers of neocortical neurons on receiving external signalling factors such as retinoic acid, which generates a high proportion of upper-layer neurons.

MicroRNAs are endogenous non-coding small RNAs and cleaves their target mRNAs. They alter the overall protein-coding machinery of the cell through modulating the epigenetic control and post-translational modulation. Cellular differentiation is a complex process with multiple regulatory players acting at different scales and regulatory nodes. Corticogenesis is accompanied with regulatory dynamics of miRNAs. The spatiotemporal expression patterns of miRNAs during cortical development have been mostly categorized into four groups (Olga Barca-Mayo et al.,

2013). The first group comprises of miRNAs whose expression is detectable throughout the development (similar to that of the housekeeping ones) like the let-7 family and miR-9, which are necessary for the maintenance and proliferation of neural stem cells (NSC) by targeting *Tlx* (Roush et al., 2008) and CyclinD1 (Zhao et al., 2009).

MicroRNAs, which depict enhanced expression at early stages of development with a gradual decrease towards the perinatal stages, are put under the second category like miR-125b, miR-181 family, and miR-17-92 clusters. During this particular stage of development, miRNAs tend to have a specific and significant role in lineage determination and cellular differentiation, just like miR-17-92 cluster is essential for the maintenance of apical progenitor population (Olga Barca-Mayo et al., 2014). Embryonic day 10 marks the upregulation of miR-34c, miR-152, miR-219-5p, miR-301b, miR-449a, miR-451, and miR-532-5p that signify their roles in NSC viability and proliferation (Yao MJ et al., 2012). A conglomeration of miRNAs is recruited during this tenure that aids in NSC self-renewal capability in developing cortex like miR-181d and miR-30c, which targets *HtrA1* (Nigro et al., 2012) and miR-34a and miR-29a which controls the p53 pathway that is a checkpoint between NSC proliferation and cell death.

MicroRNAs with accelerated expression at the mid-temporal stage and retains the expression levels till postnatal stage of brain development are under the third category, which comprises miR-124a, miR-99a, miR-266, and miR-128 with function to decide the fate of the neural stem cell commitment. MiR-124 and miR-9 support the neuronal commitment of NSCs by targeting the REST gene, and miR-124 specifically binds and represses SCP1, a phosphate component of REST (Laneve et al., 2010). Notch signaling maintains the NSC proliferation; miRNAs target Notch1 (Guruharsha KG et al., 2012) and its downstream target genes like *Blbp* (Kuang et al., 2012) that suppress the self-renewal and proliferation of NSC and pushes the cell towards neuronal commitment, which is by the action of miR-34a, miR-23b, miR-24, miR-27b, and miR-9. STAT3 is essential for the glial cell commitment: miR-9 and miR-124 repress the amount of phosphorylated STAT3 in the cell, thus enhances the probability of cell fidelity towards neuronal fate (Krichevsky et al., 2006). However, miR-124 also targets Sox9 as Sox9 promotes gliogenesis. MiR-9 maintains the cell identity in oligodendrocytes by suppressing non-oligodendrocyte lineage genes (Cheng et al., 2009).

MicroRNAs with abundant expression during cortical neuron formation and maturation are under the fourth category. The dynamics of microRNAs at embryonic day 21 reveals a decrease in

expression of miR-19b and extremely high levels in miR-137, miR-128, let-7b, and miR-185 during the formation of glutamatergic cortical neurons (Olga Barca-Mayo et al., 2014). MiR-132 and miR-212 are engaged in the synchronization of synaptogenesis, synaptic plasticity, and radial migration of glutamatergic cortical neurons in the adult brain. Mir-132 and miR-134 induce neurite outgrowth and neuronal migration.

The majority of studies performed to date are somehow focused on determining the role of miRNA in cellular development and lacks relevant insights for lineage specification at various stages of neurogenesis. The *invivo* studies are very important but due to the complex regulatory mechanism of tissue development, it becomes very difficult to the dissect the mechanistic cues associated with corticognenesis. To address this concern, we have used an *invitro* model system with well-defined and established characteristics i.e. mESCs to NSCs to Cortical neurons to map out the regulatory dynamics of miRNAs as one state transitions to another.

MiRNA in the context of Lactogenesis

The mammary gland is a unique organ that distinguishes mammals from other animals and the main function of this gland is to secrete milk to nourish offspring. The mammary gland is the only gland whose most of the development starts after birth. Gland development takes place in different cycles starts from the embryonic stage, virgin stage till attains puberty, pregnancy stage, and lactation stage. After the lactation cycle completes the gland undergoes a process called involution and comes back to its original virgin state where the gland doesn't secrete milk. During these courses of events, cells receive proliferate, differentiate, and apoptosis signals by various hormones that are secreted by the pituitary gland and adrenal glands. Misregulation of these cycles of events leads to the development of breast cancer which mimics these developmental processes.

Mouse mammary gland development starts during embryonic days 10.5 to 18.5 and ceases its growth till birth. After birth, it attains the growth signals and continues during the pregnancy cycle (Hens and Wysolmerski, 2005, Sakakura, 1987, Veltmaat et al., 2003). From day 10.5, cells from ectoderm start enlarging and extend from anterior to posterior limb bud to form five pairs of mammary fat buds (Hens and Wysolmerski, 2005, Propper, 1978; Robinson, 2007). Embryonic days 11.5 to 13.5 epithelial cells proliferate to mammary placodes (Sakakura, 1987; Watson and Khaled, 2008). In male embryos mesenchyme at day between 13.5 to 15.5 androgen receptor activation signals for mammary bud degradation (Sakakura, 1987). Gland development continues

in the female embryo at day 16.5 to form the nipple by the rapid proliferation of epithelial cells overlying the bud and lumen formation in the sprout (Hogg et al., 1983).

After birth, the mammary epithelium at the nipple remains in the quiescent state till puberty achieves. During puberty, mammary epithelium starts invading into the mammary fat pad by the process called branching morphogenesis and form the terminal end buds (Lyons, 1958; Nandi, 1958). These buds contain highly proliferative terminal cap epithelial cells that surround the multi-layered body epithelial cells and myoepithelial cells (Silberstein and Daniel, 1982; Williams and Daniel, 1983). The invading epithelial cells show some characteristics of epithelial-to-mesenchymal transition (EMT) (Kouros-Mehr and Werb, 2006). The EMT process in mammary gland development is tightly regulated by a transcription factor Ovol2 negative regulator of EMT which is required for tight regulation unlike the EMT process in cancer (Watanabe et al., 2014). The branching process continues to fill the mammary fat pad and stops and again during pregnancy gland starts differentiating in lactation then remodels through involution.

Various hormones were shown to be involved in the regulation of mammary gland development from puberty to lactation stage. Estrogen is the golden hormone in the female which allows gland proliferation, branching, terminal bud (TEB) formation in the female mammary gland. ERα is the receptor for estrogens and expressed in both mammary epithelium and stroma of the gland (Silberstein and Daniel, 1987). To examine estrogen role in epithelium and stroma in wild type (wt) female mice, in 3-week-old female mice endogenous mammary epithelium and stroma was cleared off and engrafted with ER α -/- female epithelium and stroma (Mallepell et al., 2006). Wild type mammary epithelium engrafted into cleared fat pads of mice epithelium proliferate and grows to fill the entire mammary fat pad (DeOme et al., 1959). Whereas ER α -/- mammary epithelium fails to grow. During pregnancy, wt epithelium showed side branching, in contrast ER α -/- remains rudiment. Stromal ERα signaling was shown to be not required for mammary gland development. when wt abdominal muscle wall and epithelium was grafted into ER α -/- female mammary gland grows normally even the ER α -/- epithelium coexists (Mallepell et al., 2006). Similar experiments were carried out to assess the role of progesterone in female mammary gland development due to its expression in both epithelium and stroma. Data showed PR-/- mammary epithelium grows normally when grafted into wt female showing growth of clear mammary fat pad in the absence of PR signalling (Haslam and Shyamala 1981; Haslam 1989). But side branching and alveoli formation are affected indicates epithelial PR signalling is required for ductal side branching during pregnancy. When experiments conducted with PR-/- stroma showed normal

outgrowth and side branching and showed no effect in gland development (Lydon et al., 1995; Brisken et al., 1998). Similarly, prolactin hormone deficient (Prl-/-) epithelium graft, showed normal outgrowth of ductal branching but showed defect alveogenesis and differentiation during late pregnancy. This has led to defects in the production of milk proteins such as ß-CASEIN and WAP. These data showed that Prl signalling was critical during late pregnancy for milk production. It also showed undetectable levels of phosphorylated STAT5A required for milk protein synthesis (Brisken et al., 1999; Gallego et al., 2001). When stroma Prl-/- grafted, experiments showed normal development (Ormandy et al., 2003). These experiments showed the importance of stage-specific hormonal signals for the proper development of the mammary gland after puberty.

The cellular composition of the mammary gland is distinct compared to other glands in mammals. The mammary gland uniquely possesses bipotent stem cells which are CD29hi/CD49fhi/ CD24+/mod/Sca-1- positive (Shackleton et al., 2006; Stingl et al., 2006). These stem cells can give rise to a common progenitor which gives rise to two lineages a myoepithelial progenitor and luminal progenitor. Myoepithelial progenitors differentiate to myoepithelial cells. Luminal progenitors produce luminal and alveolar epithelial cells which can further differentiate for the capable of milk production during lactation (Asselin-Labat et al., 2008). Stromal cells also contribute a major portion of the mammary gland that includes fibroblast cells and immune cells. Various transcription factor networks interplay during mammary gland development and differentiation. Among them, lobuloalveolar development and growth were under the control of CCAAT/enhancer-binding protein (C/EBPbeta) (Grimm and Rosen, 2003; LaMarca et al., 2010) during mid-pregnancy. STAT5a/5b are the downstream regulators of Jak-stat signaling which is a critical pathway during mid-pregnancy and lactation required for milk synthesis. Experiments conduct by conditional deletion of these transcription factors STAT5a/5b showed no effect on mammary stem cell population of ductal lineage CD69+ but there is a drastic reduction of another CD69+ luminal progenitor. Results in loss of differentiation capacity that ultimately inhibits milk production during lactation (Yamaji et al., 2009). Further, transcription factor GATA3 was shown to be very important from an early stage to a late stage of mammary gland development. Conditional deletion of GATA3 at early-stage results in impairment of placode formation. During mid-stage, GATA3 restricted to the luminal epithelial portion which is resulted in impairment of ductal elongation that indicates GATA3 requirement in maintaining mammary progenitor cell population (Kouros-Mehr et al., 2006; Asselin-Labat et al., 2007). In later stages of gland development, GATA3 was shown to be restricted to the alveolar compartment. Conditional deletion GATA2 or Gata3 results in loss of lobuloalveolar development with lactation deficient

mammary gland (Kouros-Mehr et al., 2006; Asselin-Labat et al., 2007). Elf5 is an ETS family-related transcription factor that functions along with GATA3 during mammary gland development. Loss Elf5 did not impact branching and ductal elongation but severely affected lobuloalveolar development leading to deficient in lactation (Oakes et al., 2008). STAT3 is another transcription factor that expresses at the terminal stage of mammary gland differentiation that controls the remodelling of the gland by inducing apoptosis and gland involution (Chapman et al., 1999; Humphreys et al., 2002).

The various epigenetic mechanism has shown to be influential in development and differentiation of mammary gland. Specifically, DNMTs are essential for the maintenance and proliferation of mammary stem and progenitor cells (Santos et al., 2015, Ivanova et al., 2021). Among DNMTs DNMT1 is critical for the development of ductal and terminal end bud formation (Pathania et al., 2015). A DNA methylation modulator, TET2 directs mammary stem cell differentiation. Chromatin complex which is made up of TET2 and FOXP1 involves in mammary luminal lineage specification by demethylating *Esr1, Gata3*, and *Foxa1* (Asselin-Labat et al., 2006). Mammary stem cells have a low level of H3K27me3 which appears to be increased during differentiation (Pal et al., 2013). A demethylase, JARIDIB/PLU1/KDM5B recruits GATA3 to its targeted genes and also exhibits mammary gland development, maintenance of estrogen level, and fertility rate (Zou et al., 2014). Another demethylase, UTX/KDM6A activates many luminal transcriptional factors by demethylating H3K27me2/3 (Agger et al., 2007). JHDM1B, histone demethylase acts as a tumour suppressor in mammary epithelial cells by controlling the cell cycle through demethylation of H3K4me3 and H3K36me2. JMJD2B which demethylates H3K9me3, responsible for mammary gland development and morphogenesis (Kawazu et al., 2011). SUZ12, an essential component of PRC2 complexes maintain progenitor cell activity and normal mammary gland development (Michalak et al., 2018). Another methyltransferase, EZH2 is essential for the maintenance of the luminal cell population and postnatal mammary gland development (Michalak et al., 2013, Ivanova et al., 2021).

In an effort to understand the significance of spatiotemporal expression of various hormones during lactogenic differentiation of the mammary glands, in vitro lactogenic differentiation of HC11 mammary epithelial cells would serve as an excellent model system to dissect the roles played by these hormones. Towards this goal, mRNA-sequencing at various stages of lactogenic differentiation had shown the gene regulatory networks that were orchestrated by these hormones. In conjunction with these studies, regulatory roles of lactogenesis mediated through microRNAs

have not to be elucidated comprehensively. To understand mechanistically, the role of miRNA mediated transcriptional networks in the differentiation process, we have chosen a well-established in vitro mouse mammary epithelial lactogenic differentiation system (HC11 cell lines) (Aydogdu et al., 2012). HC11 cells are prolactin responsive clones of the COMMA-1D cell line. These cells have epithelial stem-like properties (Normal) and can differentiate into luminal and myoepithelial progenitors (Primed). Further, these progenitors based on developmental cues can differentiate into ductal, alveolar and myoepithelial lineages during lactogenesis (Prolactin). This stage is characterized by mammospheres formation and by the secretion of the ß-CASEIN protein. This model system is well-established to study the role of miRNAs in mammalian cell differentiation. Micro-RNA transcriptome can be effectively characterized by high throughput sequencing. Here, we used Illumina platform-based next-generation deep sequencing of miRNAs in HC11 (Normal), Primed (P) and Prolactin (PRL) states. The role of miRNAs in mammary differentiation was reported by Aydogdu et al., 2012 using microarray, a probe hybridization-based technology, which included 700 probes. The list of known microRNAs have been grown in recent years and now more than 2000 miRNAs were updated in miRbase-V22 (Griffiths et al., 2008). This is one of the caveats from hybridization-based technologies that can be limiting which concern in certain cases. Moreover, RNAseq can also predict novel miRNA which is a major restrain in microarray studies. Also, hybridization-based technology shows major limitations in low signal vs background and high signal vs saturation point. RNAseq on the other hand provides more accurate quantitative profiles that clearly illustrate differentially expressed genes.

Previous studies have shown that initiation of lactogenesis was accompanied by significantly increased expression of miR-200a, miR-200b, miR-148a, miR-152, and miR-30b/d family. miR-146b and miR-21 expression were upregulated in primed conditions (Aydogdu et al., 2012). miR-200 family with miR-205 was shown to regulate epithelial to mesenchymal transition and mammospheres formation during prolactin treatment (Wyatt et al., 2007). miR-200a and miR-200b were documented to inhibit the expression of *Zeb1, Zeb2*, and polycomb complex *Suz12*, leading to increased expression of the downstream targets of these genes such as *E-cadherin* (Iliopoulos et al., 2010). Increased *E-cadherin* disturbs epithelial to mesenchymal transition, which is essential for development and metastasis. Also, miR-200a downregulates the expression of the ephrin receptor of the protein-tyrosine kinase family (*EphA2*), which promotes tumorigenesis and metastasis in the mammary gland by regulating mammary gland branching (Vaught et al., 2009). But, in the Normal stage miR-200a and miR-200b were shown to be downregulated. Their overexpression in both Normal and breast cancer stem cells suppresses stem cell factors and

mammospheres formation. Many tumour suppresser miRNAs are activated during lactogenesis like the let-7 family, miR-148a, miR-200b, miR-27b, miR-205, and miR-146a/b (Bhaumik et al., 2008). Downstream product of Wnt pathway, *SerpineI* is highly expressed in Normal condition and breast cancer. *Serpine1* is directly targeted by miR-148a and miR-181a. Further, many genes that are affecting *Serpine1* are also targeted by miR-148a, miR-27b, miR-200b, miR-205, miR-26a, and miR-181a (Krzyzanowski et al., 2007). OncomiR like miR-17 and miR-206 shows high expression in Normal and downregulated upon differentiation (He L et al., 2005). In a Normal state, miR-17 suppresses tumour suppresser genes *Dab2* and *Celsr2* (Williams C et al., 2009). MiR-17 also targets the expression of *Stat3* and its downstream target genes *Pik3r1, Igfbp5*, and *Cdkn1b* in Normal conditions. Mir-17 and miR-206 together target *Stat3* pathway by targeting its downstream gene, *Mxd4* (Aydogdu et al., 2012). *Stat3* and *Stat1* are also suppressed by *Dnmtl* in the undifferentiated stage, but upon differentiation high expression of miR-148a and miR-152 suppress *Dnmtl* (Aydogdu et al., 2012). All the above studies provide evidence on the role of miRNAs during lactogenesis which was further dissected deeply with the help of high-resolution microRNA-seq analysis.

MiRNA what is not known?

Importance of microRNA cannot be ignored starting from zygote to complete individual development. During development, we have seen above the impact of miRNA on controlled gene regulation. Successful cellular differentiation is required highly controlled gene expression patterns which are tightly undertaken by miRNA. Many important miRNAs are reported during zygote development through blastocyst to gastrula. The formation of inner cell mass during the blastocyst stage is regulated by miR-302 cluster, miR-371 cluster, miR-520 cluster, miR-17-92 cluster and miR 93 in human. In mice, miR-106a, miR-93, miR-20, miR17-5p, and miR-290 are responsible for the same. Heading towards gastrulation, the Early gastrula stage is regulated by miR-124a in human and by miR-200a, miR-200b, miR-200c, miR-141 and miR429 in mice. Coming to the late gastrula stage, Mesoderm is controlled by miR-145 and miR-302 in humans and in mice miR-290, miR-93, and miR-17-5p are responsible. In human, miR-145, miR-10a, miR-24, miR-375, miR-122, miR-192, miR-196a and miR-196b are involved in differentiation into endoderm and in mice, miR-93, miR-338-5p, and miR-340-3p are engaged. Differentiation towards ectoderm in human is facilitated by miR-125, miR-30b, and miR-30c and in mice by miR-29c, miR-125a, miR-376a, miR-297, miR-96, miR-21, let-7, miR-424 and miR-214. Not only during embryo development but trophectoderm specification also controlled by miRNAs. miR-125, miR-30b, and miR-30c are in human and miR-297, miR-214, miR-96, miR-125a, miR-21, miR-93, miR-424, miR-29c, let-7 and

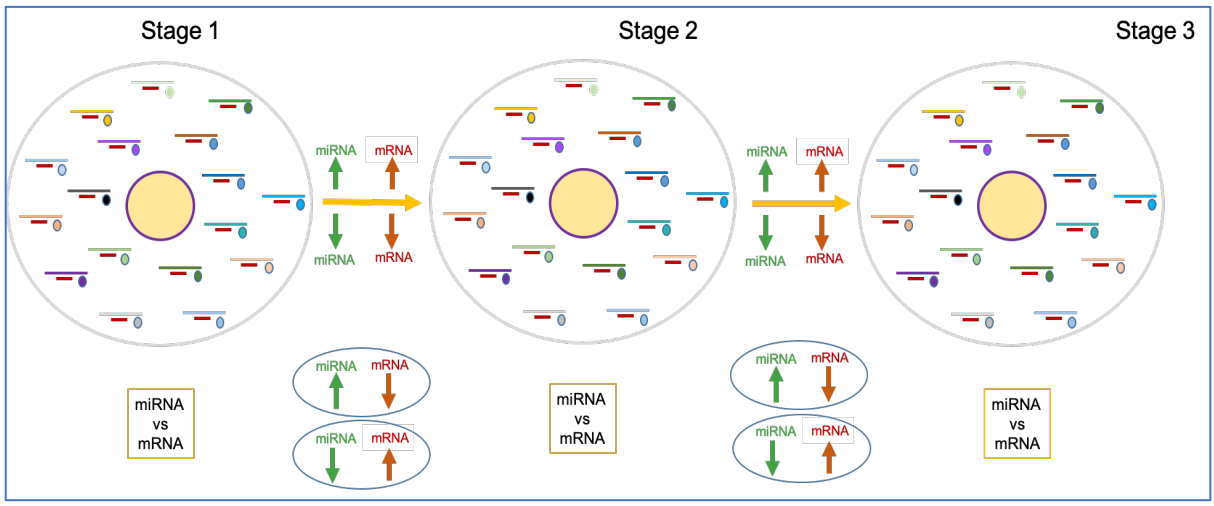
miR-376a in mice are engaged in trophectoderm specification. Even pluripotent embryonic stem cells can differentiate into various other kinds of cells with controlled miRNA regulations. Like, cardiomyocytes differentiation is regulated by miR-1, miR-133, miR-206, miR-208 and miR-499, smooth muscle cells differentiation by miR-1, miR-123, miR-141, miR-145 and miR-21, endothelial cells differentiation by miR-17-92 cluster, miR-126, miR-130a, miR-133, miR196, miR-210 and let-7, osteocyte differentiation by miR-140, miR-143, miR-21, miR-23, miR-27, miR-29, miR-2861, miR-3960 and miR-676, erythrocytes differentiation by miR-144, miR-150, miR-155, miR-181, miR-24, miR223 and miR-451 and neurons differentiation by miR-9 and miR-124. Though many miRNAs has been reported with lineage specification, still many turning points yet to be disclosed that is affected by the expression of miRNA.

Not only in the field of differentiation but also in the disease field significant miRNAs were characterized. Neurological diseases like Alzheimer's disease which is the result of variation in the expression of let7b, miR-34, miR-107, miR125b. Mis regulation of miR-9 and miR-132 can result in Parkinson's disease. Hutchinson's disease and Amyotrophic Lateral Sclerosis are appeared due to abnormalities in the regulation of miR-132, miR-184, and miR-146a. In the disease field, miRNAs are well studied in the field of cancer. The suppression of tumour suppression miRNA and elevation of oncogenic miRNAs lead to carcinogenesis. Cells have a tight regulation system even to control expressed miRNAs instantly by antimiRNAs. They have sequence complementarity with miRNAs which prevents the binding of miRNAs with mRNAs. All these control systems need to be further revised through many experiments within different tissue and differentiation system.

Logical hypothesis of miRNA's functional dynamics in the context of cellular differentiation

Always the best way to estimate the role of a particular miRNA is to study its role in a cell differentiation system that has two or more differentiation stages. Having a transcriptome profile of both mRNA and miRNA throughout different stages of differentiation is always useful to look into the differential patterns of gene expression. Critical evaluation of upregulated miRNAs and downregulated mRNAs between two stages defines the importance of mRNAs in the first stage which is needed to be downregulated during the second stage. It also provides information about miRNAs that are responsible for the downregulation of these mRNAs for the differentiation. Similarly, analysing downregulated miRNAs and upregulated mRNAs between stages helps in

determining mRNAs that are not required in the first stage and suppressed by downregulated miRNAs. Now those miRNAs need to be downregulated for the expression of essential mRNAs for the second stage. Once after detecting a related miRNA, an efficient knockdown approach can be used to further dismantle the role of that particular miRNA in any desired differentiation system. That's how a differentiation system provides a useful platform for the proper characterization of many miRNAs. Here, I am trying to replicate a similar approach to dissect the role of miR-301b-3p in neurons and miR-122-5p in mammary epithelial cell differentiation.



Materials and Methods

Materials

10% complete medium: DMEM supplemented with 10% FBS, 1X Antibiotic Antimitotic, Insulin (5 µg/ml; Sigma # 16634) and EGF (20ng/ml; Sigma # E4127).

5% complete medium: DMEM supplemented with 5% FBS, 1X Antibiotic Antimitotic, Hydrocortisone (1ug/ml; Sigma # H4001) and Insulin (5ug/ml).

10% ESC medium: DMEM (Gibco#10569-010) supplemented with 10% FBS, 1X Antibiotic Antimitotic, 1000 U/ml leukaemia inhibitory factor (Merck#ESG1106), 1X non-essential Amino acids (NEAA, Gibco#11140-050) and 0.11 mM ß-mercaptoethanol.

N2 medium: It composed of 1:1 of DMEM and F12, 1X N2 supplement (Gibco#17502048), 1X NEAA, 1X Glutamax, 1X Sodium Pyruvate, BSA (50µg/ml) and 0.11 mM ß-mercaptoethanol.

B27-A medium: B27-A medium was prepared by adding 1X B27-minusVitA supplement (Gibco#12587010; without Vitamin A) and 1X L-Glutamine into neurobasal medium (Gibco#21103049).

N2B27-A medium: N2B27-A medium contains 1:1 of N2 and B27-A medium.

2i medium: 2i medium composed of knockout DMEM (Gibco#10829-018) supplemented with 1X N2 supplement (Gibco#17502048), 1X B27 supplement (Gibco#17504044: with Vitamin A), 1µM MEK inhibitor (Mitogen-activated protein Kinase inhibitor; Selleckchem#PD0325901), 3µM GSK inhibitor (Glycogen Synthase Kinase inhibitor; Selleckchem#CHIR-99021), 1000 U/ml leukaemia inhibitory factor (LIF) and 0.11 mM ß-mercaptoethanol.

ESC freezing medium: DMEM (Gibco#10569-010) supplemented with 15% FBS, 1X Antibiotic Antimitotic, 1X non-essential Amino acids (NEAA, Gibco#11140-050), 0.11 mM ß-mercaptoethanol and 5%DMSO.

Gelatin coated plate: 0.1% gelatin was poured into cultured flask in such a manner that it should cover the entire surface. Flask was placed at 37^0C for 10min or half an hour at room temperature. Then gelatin was removed by pipetting and allowed to air dry.

PORN-Laminin coated plate: Poly-Ornithine (Sigma#P4957) was poured into cultured plate and incubated at 37^0C for overnight. Next day, PORN was removed and plate was washed two times thoroughly with mili-Q water. Then Laminin (Merck#11243217001) was added and incubated at 37^0C for 3hrs to overnight. Cells with media were plated directly after laminin removal; caution was taken for not to dry the laminin coated plates before the addition of cells.

50X TAE buffer: 2M Tris base (Sigma#T1503), 1M Glacial acetic acid (Sigma#1005706), and 50mM EDTA were added to miliQ water and made the volume up to 500ml.

1X PXL buffer: 0.2% Igepal/NP40 substitute (Merck# 492016), 0.1% sodium deoxycholate (D7650), and 0.02% SDS (L3771) were dissolved in 1X PBS (P3813).

Bead wash buffer: 0.02% Tween-20 (P9416) dissolved in 1X PBS.

High-stringency wash buffer: 1% Igepal/NP40 substitute, 1% sodium deoxycholate, 0.1% SDS, 15mM Tris-HCl (Thermo Fisher# 15567027) (pH 7.5), 5mM EDTA (Sigma#E9884) (pH 8.0), 2.5 mM EGTA (Sigma#E3889) (pH 8.0), 120 mM NaCl (S7653), and 25 mM KCl (P9541) were dissolved in nuclease free water.

High-salt wash buffer: 1% Igepal/NP40 substitute, 0.5% sodium deoxycholate and 0.1% SDS, and 860mM NaCl were dissolved in 1X PBS.

Low-salt wash buffer: 15 mM Tris-HCl (pH 7.5) and 5 mM EDTA were dissolved in nuclease free water.

1X PNK buffer: 0.5% Igepal/NP40 substitute, 50 mM Tris-HCl (pH 7.5), and 10 mM MgCl2 (Sigma#M8266) were added to nuclease free water.

Beads Preparation: For 10 million cells 1.5mg (50µl) protein G (Dynabeads™ Protein G; 10004D) beads were taken in a 2ml safe lock tubes and kept on a magnetic stand to remove the supernatant. Beads were washed 3 times with bead wash buffer and resuspended in 50µl bead wash buffer. 5µg anti AGO antibody (Anti-pan Ago Antibody, clone 2A8; MABE56) was added to it and incubated for 30min at room temperature. Beads were washed 3 times with 1X PXL buffer followed by resuspension in 50µl 1X PXL and stored at 4^0C till use.

Bioinformatics packages

1. **Cutadapt**: https://cutadapt.readthedocs.io/en/stable/
2. **FastQC:** http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
3. **Bowtie:** http://bowtie-bio.sourceforge.net/manual.shtml
4. **Bowtie2:** http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
5. **Tophat2:** https://ccb.jhu.edu/software/tophat/index.shtml
6. **Cufflinks:** http://cole-trapnell-lab.github.io/cufflinks/
7. **DESeq2:** http://bioconductor.org/packages/release/bioc/html/DESeq2.html
8. **miRDeep2:** https://github.com/rajewsky-lab/mirdeep2
9. **Clust:** https://github.com/BaselAbujamous/clust
10. **miRNet:** https://www.mirnet.ca/
11. **Novoalign:** http://www.novocraft.com/products/novoalign/
12. **Hyb:** https://github.com/gkudla/hyb
13. **CIMS:** http://zhanglab.c2b2.columbia.edu/index.php/CTK
14. **miRNA-seq-codes:**
    https://github.com/Rakheenayak/HC11_miRNA_analysis_pipeline

Experimental models

Culture condition for Ground, Naïve+2i and Naïve state:

Ground state: E14TG2a, Embryonic stem cell lines (purchased from Dr. Smith) were thawed freshly from liquid nitrogen by incubating tube at 37°C water bath for 5min. Then cells were collected by centrifugation with 10ml 10% ESC medium at 800rpm for 5min. These cells were plated on a 0.1% gelatin-coated plate, supplemented with 2% FBS in 2i medium. After 80% confluency medium was removed and cells were treated with 0.05% trypsin for 5min. Detached cells were collected through a pipette and transferred to a centrifuge tube with 10ml 10% ESC medium. The cell pellet was collected after centrifugation at 800rpm for 5min. the cell pellet was dissolved with 1ml 2i medium and plated on a 0.1% gelatin-coated plate, supplemented with 2% FBS in 2i medium. These cells were allowed to grow up to 3rd passages at 37°C and 7% CO2. 4th passage onwards, cells were grown under 1% FBS in 2i medium and at 5th passage cells were harvested, as Ground state ESCs.

Naïve+2i state: R1, Embryonic stem cell lines (kind gift from Dr. Andras Nagi) were thawed from liquid nitrogen by placing in a 37°C water bath for 5min. Then cells were collected by centrifugation with 10ml 10% ESC medium at 800rpm for 5min. These cells were plated on to 0.1% gelatin-coated plate having a monolayer mitomycin-C treated inactive mouse embryonic fibroblast feeder cells (Bibel et al. 2007). These cells were supplemented with 10% FBS in 2i medium and were grown up to 80% confluency at 37°C with 7% CO2. ESCs were selectively enriched over fibroblasts and harvested as Naïve+2i for further experiments. For selective enrichment of ESC from fibroblast, after trypsinization, the 10ml mixture of ESC and fibroblast were poured onto a 100mm petri dish. As fibroblasts are attached first to the surface than ESC, it was allowed to settle down for 30min at 37°C and 7% CO2. Afterward, a complete 10 ml medium was collected slowly and transferred to another gelatin-coated petri dish with 10% ESC medium. The above selective enrichment procedure was repeated for three passages that will provide a 90% pure population of ESC, devoid of fibroblasts.

Naïve state: E14TG2a, Embryonic stem cell lines were thawed from liquid nitrogen by placing in a 37°C water bath for 5min. Then cells were collected by centrifugation with 10ml 10% ESC medium at 800rpm for 5min and then plated on a 0.1% gelatin-coated plate. ESCs were supplemented with 10% ESC medium and allowed to grow up to 80% confluency in 37°C and 7% CO2 then harvested as a Naïve state.

Freezing of embryonic stem cells: Embryonic stem cells were allowed to grow up to 80% confluency. Then the medium was removed and 0.05% trypsin was added to it for 5min. Cells were collected on to a 15ml conical flask with 10ml 10% ESC medium and proceeded for centrifugation for 5 min at 1000rpm. The supernatant was removed and the pellet was resuspended in ESC freezing medium. Then 1ml freezing medium with 1million cells were distributed into 1.5ml cryo-vials and transferred to cryo box. The cryo box was kept at -20°C for 30min followed by -80°C for 2hrs then liquid nitrogen.

Cortical neuron differentiation:

E14TG2a Embryonic stem cell lines, Oct4-GFP, 46C (Sox1-GFP), TK23 (Tau-GFP), were allowed to grow in 15% ESC medium till it has reached 50% confluency on 0.1% gelatin-coated plates. Afterward, Medium was replaced with N2 medium and considered as day 0. After 2 days,

the medium was exchanged with fresh N2 medium along with 1X Cyclopamine (Merck#C4116). Every after 2 days, fresh N2 medium along with 1X Cyclopamine was added till day 10. After day 10 Cyclopamine was removed and fresh N2 medium was added and continued to culture up to the 12th day. On day 12, cells were incubated with 0.05% trypsin (diluted freshly with PBS and pre-heated at 37^0C) and proceeded for centrifugation with 10% PBS (PBS with 10% FBS). The cell pellet was dismantled properly and plated onto Poly-Laminin coated plate in the presence of N2B27-A medium. Neuron Progenitor cells (NPC) were harvested after 2 days of plating (day 14). For Cortical Neuron differentiation, cells were continued to culture in N2B27-A culture medium up to the 21st day. During this period, fresh medium was added every 2 days to the old medium instead of exchanging it (Gaspard et al. 2009).

Mammary Epithelial cells differentiation:

Mouse mammary epithelial stem cell (HC11) were allowed to grow simultaneously in three T25 flasks (Corning#430639) at 37^0C and 5% CO_2, in 10% complete medium with EGF, till they were grown to confluency. For all the experiments cells were grown up to confluent. One of the T25 flasks was harvested for Normal (HC11-N) state. Remaining two flasks, the medium was changed to 5% complete medium (Hydrocortisone) and cultured for 48 hours (two days). After two days, one of the flasks, cells were harvested in all the experiments and were termed as Primed (Primed condition is considered after 2days of GC treatment.). In the third flask, the medium was replaced with a 5% complete medium supplemented with Prolactin hormone (5 µg/ml); NIH # NIDDK-oPRL-21 and incubated for up to 72hours. These cells were termed as Prolactin state (HC11-PRL). The above-mentioned protocol was repeated for biological replicate samples and all stages of cells were harvested on its 9^{th} passage. To freeze, HC11 cells were resuspended in 7% DMSO in 10% complete medium with additional 10% FBS and then transferred into -20^0C for 1hr followed by -80^0C 2hrs to overnight and followed by liquid nitrogen.

Methods

Isolation of total RNA: Cells of all the stages of HC11 differentiation (HC11-N, P and PRL), ESC Ground, Naïve, NPC, and CN were harvested for RNA isolation by directly adding TRIzol (Invitrogen#15596018). (how much per what kind of flask) into the respective flasks. Total RNA was isolated from cells by using TRIzol reagent as per the recommendation. In brief, cells were lysed in presence of 1ml TRIzol by gently shaking the flasks and pipetting in and out with a

microtip. The suspension was transferred to a 1.5ml Eppendorf tube. Up and down pipetting was done vigorously to make proper lysis of cells. 200µl Chloroform was added, vortexed thoroughly, and centrifuged at 12,000rpm for 15min. The upper aqueous phase was carefully collected into a fresh to 2ml Eppendorf tube. 500µl Isopropanol was added to it, vortexed, and centrifuged at 13,000rpm for 30min. The supernatant was discarded and the pellet was washed in 1ml 75% ethanol then centrifuged at 13,000rpm for 30min at 4^{O}C. The supernatant was discarded and the pellet was allowed to air dry. Pellet was resuspended in nuclease-free water and stored at -20^{O}C. Total RNA was quantified by using a spectrophotometer (Nanodrop)

DNase treatment of RNA sample: 20µg RNA was incubated with 10units of RNase-free-DNaseI (Merck#4716728001) and 1X DNaseI buffer at 37^{0}C for half an hour followed by RNA isolation by using TRIzol for the second time. Isolated total RNA was further quantified by using a nanodrop spectrophotometer.

Isolation of microRNAs: microRNAs from HC11 cells undergoing lactogenic differentiation (HC11-N, P and PRL) and ESC Ground, Naïve, NPC, and CN were isolated by using miRVana miRNA isolation Kit (Invitrogen# AM1560) as per the recommendations with minor changes. In brief, 300µl of lysis buffer was added directly to 3-5 million cells (cells from one T25/30-60mm Petri plate) at room temperature. 30µL miRNA homogenate additive was added and incubated further 10min on ice. Then 300µL acid-phenol:chloroform in 1:1 ratio was added, mixed or vortexed and centrifuged at 10,000rpm 5min in room temperature. The aqueous phase was separated and 100µL of 100% ethanol (1/3th of the total volume of the aqueous phase) was added. It was mixed properly and passed through the spin column1 through centrifugation at 10,000rpm for 1min at room temperature. Column1 was kept on a 2ml tube for mRNA extraction. 266µL 100% ethanol (2/3th of the total flow through) was added to flow through, mixed properly, and passed through column2, which contains miRNA, through centrifugation at 10,000rpm 1min in room temperature. Both column1 and column2 was washed with 700 µL miRNA Wash Solution1 followed by 2 times 500µL Wash Solution 2/3 through centrifugation at 10,000rpm 1min at room temperature. Columns were dried by centrifugation at 10,000rpm 2min at room temperature. 50µL and 30µL nuclease-free water were added into column1 and column2 respectively. Columns were incubated for 1min and centrifuged at 10,000rpm 1min at room temperature. Isolated RNA's concentration was measured by a spectrophotometer (nanodrop).

Quantitation of Samples:

DNA: DNA concentration was measured by using the Qubit dsDNA HS assay kit (Thermo Fisher#Q32854) in Qubit Fluorometer 4.0. Standard samples 1 and 2 were prepared by adding 190µl dsDNA HS buffer and 10µl of Standard #1 and Standard #2 into two different tubes separately. Samples were incubated for 2min in dark at room temperature. First, Standard #1 was set followed by standard #2. Unknown samples were prepared by adding 1µl of DNA sample and 199µl dsDNA HS working solution. It was incubated for 2min in dark at room temperature and read on a fluorometer.

RNA: RNA concentration was measured by using the Qubit RNA HS assay kit (Thermo Fisher#Q32855) in Qubit Fluorometer 4.0. Standard samples were prepared by adding 1µl RNA HS reagent, 189µl RNA HS buffer, and 10µl of both Standard #1and Standard #2 to two different tubes. Samples were incubated for 2min in dark at room temperature. First, Standard #1 was read followed by Standard #2, and the standard was set. Unknown samples were prepared by adding 1µl of RNA sample, 1µl RNA HS reagent, and 198µl RNA HS buffer. It was incubated for 2min in dark at room temperature and read on a fluorometer.

Protein: Protein concentration was measured by using the Qubit Protein assay kit (Thermo Fisher#Q33211) in Qubit Fluorometer 4.0. Standard samples were prepared by adding 1µl Qubit protein reagent, 189µl Qubit protein buffer, and 10µl of Standard #1, Standard #2, and Standard #3 to three different tubes. Samples were incubated for 15min in dark at room temperature. First, Standard #1 was read on fluorometer followed by Standard #2, and Standard #3 and standard was set. Unknown samples were prepared by adding 1µl of protein sample, 1µl Qubit protein reagent, and 198µl Qubit protein buffer. It was incubated for 30min in dark at room temperature and read on a fluorometer.

Preparation of complementary DNA (cDNA):

mRNA: 1µg total RNA was taken as input for cDNA synthesis by using iScript cDNA synthesis kit (BioRad#1708891). To 1µg RNA, nuclease-free water was added up to 15µl in a PCR tube and incubated at 70^0C for 5min followed by snap chilling. 4µl cDNA synthesis buffer and 1µl Reverse transcriptase enzyme were added into it and the reaction was set up at 25^0C for 5min, 42^0C for 30min, and 85^0C for 5min on a thermocycler.

miRNA: miRNA cDNA was prepared by using miScript II RT kit (Qiagen#218161). 1µg miRNA, 4µl HiSpec buffer, 2µl nucleotide, 1µl reverse transcriptase enzyme, and nuclease-free water up to 20µl were added and incubated at 37^0C for 60min followed by 95^0C for 5min on a thermocycler.

Designing of primers:

mRNA: Primers used in this experiment were spanning exon-exon junction having annealing temperature of 57^0C and GC content >40. Gene sequences were retrieved from UCSC genome browser (https://genome.ucsc.edu/) and primer was designed on Primer3 tool (http://bioinfo.ut.ee/primer3-0.4.0/).

miRNA: Exact mature miRNA sequence was considered as miRNA PCR forward primers. Only the nucleotide U was changed to T for cDNA primers. Universal reverse primer was used from miScript SYBR Green PCR kit (Qiagen#218073).

Real time PCR:

mRNA: cDNA was diluted with NF water up to 20ng/µl concentration. The total reaction volume of each well was 10µl (5µl 2X KAPA SYBR FAST universal mix, 1µl cDNA template, 1µl of 5µm primer mix, and 3µl NF water). RT-PCR was kept in CFX96 Touch Real-Time PCR (Bio-Rad) and conditions were set as per KAPA#KK4618. *Gapdh* was used as the house-keeping gene for ESC, NPC, and CN samples, and *ß-Actin* was used for HC11 (N), (P) and (PRL) samples. All the data were analyzed by using the $-2^{\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). In neurogenesis, the *Gapdh* gene and ESC sample were considered as control. In lactogenesis, the *ß-Actin* gene and HC11 (N) sample were considered as control. Bar graphs were generated with one-way ANOVA by using Graph-pad PRISM software.

miRNA: miRNA cDNA was diluted with NF water up to 20ng/µl concentration. The total reaction volume of each well was 10µl (5µl 2X miScript SYBR Green PCR kit, 1µl miRNA cDNA template, 1µl of 5µm primer mix, and 3µl NF water). RT-PCR was kept in CFX96 Touch Real-Time PCR (Bio-Rad) and conditions were set as per miScript SYBR Green PCR kit (Qiagen#218073). The *Rnu6* gene was used as a housekeeping gene. All the data were analysed by using the $-2^{\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). In neurogenesis, *Rnu6* gene and ESC sample were considered as control. In lactogenesis, the *Rnu6* gene and HC11 (N) sample were

considered as control. Bar graphs were generated with one-way ANOVA by using Graphpad PRISM software.

Agarose-gel electrophoresis

mRNA & miRNA: 500ng miRNA and mRNA samples of HC11(N), HC11(P), HC11(Prl) ESC, NPC, and CN were loaded on to 0.8% agarose (Lonza#SeaKem LE Agarose; 50004) gel (0.8g of agarose in 100ml of 1X TAE) with 1µl 100% glycerol and run in 1X TAE buffer (20ml of 50X TAE in 980ml of Mili Q water) with 80V constant power. After the completion of the run, the gel was visualized under Gel doc system (BioRad#Gel Doc XR+ system). Images were captured and processed with Image Lab (version-V6.0) software from BioRad.

PCR products: 20µl PCR products were loaded on to 1.5% agarose gel (1.5g of agarose in 100ml of 1X TAE; Wt/Vol) with 6X purple loading dye (NEB#B7024S) and run in 1X TAE buffer (20ml of 50X TAE in 980ml of Mili Q water) with 100V constant power. After the run completion, the gel was visualized under Gel doc system (BioRad#Gel Doc XR+ system). Images were captured and processed with Image Lab (version-V6.0) software from BioRad.

Cross-linking Ligation and Sequencing of Hybrids (CLASH-seq):

Day 1: CLASH-seq was performed in HC11(N), HC11(P), HC11(Prl) state with 10 million cells as input from T75 flask. Cells were thoroughly washed with ice-cold PBS (pH-7.5) and submerged with fresh 20 ml of ice-cold PBS(pH-7.5). Cells were exposed to UV radiation while samples were in ice with Stratalinker at 400 mJ cm2/4000. After UV exposure, cells were removed by scraper and were collected by centrifugation (1500rpm for 5min at 4^0C). Cell pallet could be stored at -80^0C for future experiments or processed immediately by the addition of 1X PXL (lysis buffer) which is 3 times the volume of the pellet. 40Unit of SUPERase.In (2µl) (Invitrogen#AM2696) was added to it, mixed gently, and incubated on ice for 10min. 30µl DNase-1 (RQ1, Promega#M6101), was added (30U) and incubated on a thermomixer at 37^0C, 1000rpm for 10min. 2% of the total volume of the sample was saved as an input sample and stored at -20^0C. The rest of the sample was transferred to an Eppendorf tube containing anti-AGO tagged Protein-G beads and kept at 4^0C for 2hr on rotation. After rotation, the supernatant was removed and beads were washed 3 times with 1X PXL, 2 times with high salt wash buffer, 2 times with high stringent wash buffer, 2 times with low salt wash buffer, and 2 times with 1X PNK buffer with the help of a magnetic stand. Then beads were suspended in 500µl 1X PNK buffer with 1 µl (0.5Unit) RNase A

(Agilent#RNace-IT#400720) and was incubated at 20^0C for 7min followed by immediate chilling on ice. Beads were washed 2 times with 1X PNK and kept for a 5' Phosphorylation* reaction with 80µl of the total volume containing 4µl T4 PNK enzyme (NEB, M0201S), 8µl ATP (10mM), 0.5µl SUPERaseIn (10Unit), 8µl 10X PNK buffer, and 59.5µl NF water at 20^0C for 2.5hrs. Beads were washed 2 times with 1X PNK. Then ligation reaction was kept at 16^0C for overnight in 300rpm with 80µl of the total volume containing 2µl T4 RNA Ligase I (NEB; M0437M), 8µl ATP (10mM), 0.5µl SUPERase.In (10Unit), 8µl 10X PNK buffer, and 61.5µl NF water.

Day 2: Beads were washed 2 times with 1X PNK buffer and incubated with 80µl dephosphorylation reaction mix, that contains 8 µl TSAP (Thermo Stable Alkaline Phosphatase; Promega; M9910), 0.5 µl SUPERase.In (10Unit), 8 µl 10X PNK buffer, and 63.5 µl NF water at 20^0C for 45min. Further beads were washed twice with 1X PNK buffer and again* kept for 5' Phosphorylation reaction (80µl volume) at 37^0C for 30min. Then beads were washed 3 times with the 1X PNK. After that RNA-Protein complexes were separated from beads by adding the mixture of 1X LDS sample buffer (ThermoFisher#NP0007) and 1X sample reducing buffer (ThermoFisher#NP0009)in a 1X PNK solution and incubated at 70^0C for 10min in 1000rpm. The supernatant was collected and loaded into 8% SDS-Polyacrylamide gel and a run was performed with 200V in 4^0C for 4hrs. RNA-Protein complexes were transferred into nitrocellulose membrane () at 100V in 4^0C for 1hr by using Biorad Mini Trans-Blot Electrophoretic Transfer Cell (M170-3930). Nitrocellulose membrane at 90 kDa to 250 kDa size was separated by cutting with a sterile blade. Further, the membrane was cut into small pieces and transferred to one 1.5ml Eppendorf tube for extraction of protein-RNA complexes. Proteinase K treatment was done by adding Proteinase K (10 µg; 8 µl from 10mg/ml stock solution) (Sigma#P2308) in to 200µl NF water at 37^0C for 20min on a rotator (1000rpm). To the above sample equal volume of acid phenol (Invitrogen#AM9720) and chloroform in 1:1 ratio was added and mixed vigorously. After the centrifugation at 10000RPM at room temperature for 10min, the aqueous phase was collected into a fresh Eppendorf tube. To the aqueous phase $1/10^{th}$ of 3M sodium acetate (Sigma#S2889) 1 µl glycoblue (Thermo Fisher#AM9515), and 1ml 100% ethanol was added and incubated at -20^0C for overnight precipitation.

Day 3: Then sample was proceeded for centrifugation in 15000rpm at 4^0C for 1hr followed by 2 times washing with 750µl 70% ice cold ethanol in 15000rpm at 4^0C for 20min. RNA pallet was resuspended in 20µl nuclease-free water.

RNA Library preparation and sequencing:

RNA concentration was taken by using Qubit RNA HS Kit (Invitrogen# Q32852) in Fluorimeter 4. Then samples were loaded onto the Bioanalyzer (Agilent 2100) for the RNA integrity. miRNA and CLASH-seq libraries were generated with NEBNext® Small RNA Library Prep Set for Illumina (NEB# E7330L). mRNA library was generated by using NEBNext® UltraTM RNA Library Prep Kit (NEB # E7530S). Further, a Library quality check was performed with Bioanalyzer and proceeded for illumine single end sequencing for miRNA-seq (1X50bp) and CLASH-seq (1X150bp) and paired-end (2X150bp) sequencing for mRNA-seq with Illumina Hiseq 2500 platform.

knockdown of miRNA:

knockdown of microRNA mmu-miR-301b-3p in neuron: Hairpin inhibitors for mmu-miR-301b-3p (IH-310775-04-0002) was ordered from Dharmacon miRIDIAN. For negative control, miRIDIAN microRNA hairpin inhibitor negative control #2; IN-002005-01-05 (based on cel-miR-239b) was used. Neuron progenitor cells were plated on to poly-laminin-coated plate and cultured for 12days. On the day 13th, 8nM microRNA inhibitors were transfected by using Xfect RNA transfection reagent (Takara#631450). The Xfect RNA Transfection polymer was mixed thoroughly by vortexing. RNA Transfection mix for 1ml culture was prepared by adding 8nm microRNA inhibitor mmu-miR-301b-3p (IH-310775-04-0002)/miRIDIAN microRNA hairpin inhibitor negative control (#2; IN-002005-01-05), 0.8µl Xfect RNA Transfection polymer and 100µl Xfect reaction buffer. This reaction was mixed well for 5sec by vortexing and incubated for 10min at room temperature. 100µl reaction mixture was added drop-wise to the cell culture medium and incubated at 37°C in a CO2 incubator for 4 hrs. After 4 hrs cell culture transfection medium was replaced with a fresh medium. Cells were continuously transfected as mentioned above transfection procedure after every 2 days till 21st day of differentiation.

knockdown of microRNA mmu-miR-122-5p in HC11 cells: Hairpin inhibitors for mmu-miR-122-5p (miRIDIAN microRNA; IH-310775-04-0002) was ordered from Dharmacon miRIDIAN. For negative control miRIDIAN microRNA hairpin inhibitor negative control #2; IN-002005-01-05 (based on cel-miR-239b) was used. 8nM MicroRNA hairpin inhibitor for mmu-miR-122-5p (miRIDIAN microRNA; IH-310775-04-0002) and negative inhibitor control was

transfected to three separate T25 flasks with Xfect RNA transfection reagent when HC11 (N) cells are at 30% confluency. The Xfect RNA Transfection polymer was mixed thoroughly by using a vortex. RNA Transfection mix for 1ml culture was prepared by adding 8nm microRNA inhibitor mmu-miR-122-5p (miRIDIAN microRNA; IH-310775-04-0002)/miRIDIAN microRNA hairpin inhibitor negative control (#2; IN-002005-01-05), 0.8µl Xfect RNA Transfection polymer, and 100µl Xfect reaction buffer. This reaction was mixed well for 5sec by vortexing and incubated for 10min at room temperature. 100µl reaction mixture was added drop-wise to the cell culture medium and incubated at 37°C in a CO2 incubator for 4 hrs. After 4 hrs cell culture transfection medium was replaced with a fresh medium. Cells were transfected as mentioned above transfection procedure after 2 days again then harvested one of the T25 flasks as HC11(N) when it reached 100% confluency. Differentiation protocol was continued with the other two flasks to harvest HC11(P) and HC11(Prl) stage.

Western Blot analysis of Proteins:

Sample Preparation: Single-cell suspension from a total of 10million cells was made from HC11 (N), (P), (PRL), ESC, NPC, and CN. Cells were washed with cold PBS and 1ml RIPA buffer (150mM NaCl, 1% IGEPAL, 0.5% sodium deoxycholate, 0.1% SDS and 50mM Tris pH 8) (Sigma#R0278) was added along with 3X Protease Inhibitor cocktail (Merck#5056489001). Then the sample was incubated in ice for 30min for cell lysis. The cell lysate was collected in a tube and a portion of it was saved for RNA isolation (same as mentioned above). The rest of the lysate was centrifuged at 1000 RPM at 4^0C for 5 min to remove the debris. The supernatant was collected into a fresh Eppendorf tube and its concentration was measured by using a Qubit Protein assay kit (ThermoFisher#Q33211) in Fluorimeter 4.

Casting of stacking and running gel: The SDS-PAGE gel was cast by using BioRad Mini-Protein vertical electrophoresis cell #1658000FC. 5ml 8% running gel was prepared by using 1.3ml 30% acrylamide solution (29.22g acrylamide and 0.78g bisacrylamide in 100ml NF water), 1.25ml 1.5M Tris-HCl pH-8.8, 50µl 10% SDS, 25µl 10% APS, 5µl TEMED, and remaining Mili Q water. Then it was poured into a casting tray and allowed to polymerize for 15min. After that 2ml 4%, stacking gel was poured above it. 2ml 4% stacking gel was prepared by using 266µl 30% acrylamide solution, 504µl 0.5M Tris-HCl pH-6.8, 20µl 10% SDS, 10µl 10% APS, 2µl TEMED,

and remaining Mili Q water. Immediately after pouring comb was placed and allowed to polymerize for 30min. After that comb was removed and placed in a running tray.

Sample loading: 20µg protein with 1X LDS (Lithium Dodecyl Sulfate) sample buffer (Invitrogen#NP0007) and 1X sample reducing buffer (Invitrogen#NP0004) was heated at 70^0C for 5min. Samples were loaded into 8-15% SDS-PAGE under 200V constant volt at 4^0C incubator. Size fractionated protein samples from PAGE were transferred to nitrocellulose membrane (Bio-Rad#1620115) by wet transfer method with 100V constant voltage for 1 hour at 4^0C. The membrane was blocked with 5%BSA for 1hr at room temperature followed by incubation with primary antibody (1:1000) in 3% BSA overnight at 4^0C. The next day, a specific secondary antibody (1:5000) was added with 5% BSA for 1hr at room temperature. Blot was developed by using ECL western-blot detection reagents and images were processed in Chemidoc (BioRad).

Immunofluorescence:

Around half a million cells were cultured on a coverslip inside a 30mm Petri dish. The medium was removed from the Petri dish and cells were fixed with 100% ice-cold methanol at room temperature for 5min. Then cells were kept for blocking with 1% BSA in PBST (PBS+ 0.1% Tween 20) at 4^0C for 1hr. Primary antibody was added to blocking solution in 1:500 dilutions and incubated overnight inside a humidified chamber at 4^0C. The next day, cells were thoroughly washed with PBST for 3 times, and respective secondary antibody was applied for 1hrs in dark at 1:1000 dilutions. After that, 0.1µg/ml DAPI solution was added and incubated for 5min in dark followed by PBST wash 3 times. Coverslip was transferred to a slide with mounting medium and sealed with nail polish. Images were processed under a fluorescence microscope.

Mammosphere counting:

Number of mammospheres: The 40X bright field images of MEC(N), MEC(N+si), MEC(P), MEC(P+si), MEC(Prl) and MEC(Prl+si) were taken by using a bright field microscope. For each of the sample three different images from three different random spots on flasks were taken. This was repeated with three biological replicates. The round and spherical mammospheres were counted manually from every image. Then the number of mammospheres were represented by using bar graph. Bar graphs were generated by using Graph-pad PRISM software with one-way ANOVA significance test.

Size of mammospheres: The 40X bright field images of MEC(N), MEC(N+si), MEC(P), MEC(P+si), MEC(Prl) and MEC(Prl+si) were taken by using a bright field microscope. For each of the sample three different images from three different random spots on flasks were taken. This was repeated with three biological replicates. The diameter of mammospheres were measured by using image J software. Then the size of mammospheres were represented by using bar graph. Bar graphs were generated by using Graph-pad PRISM software with one-way ANOVA significance test.

Bioinformatics data analysis

Bioinformatics analysis of mRNA-seq data:

Generation of FPKM values: Approximately 20million paired end reads were received for each sample. First, quality of sequence data was checked with FastQC program. Adapters and low-quality bases were trimmed using Cutadapt program. Correlation between biological replicate was predicted by using Spearman's correlation method. Filtered reads were mapped to the *Mus musculus* reference genome GRCm38 by using Tophat2 program. Output Bam files were given as input to Cufflinks to predict FPKM values. FPKM value above 1 were considered as expressed for all the cell stages.

Analysis of differentially expressed genes: For differential gene expression analysis (log2 fold change) bam files were processed with DESeq2 algorithm. Log2 fold change above 1 were considered as upregulated genes and below -1 were downregulated between stages of differentiation.

Analysis of TF and ER: Differentially regulated and expressed transcription factors were filtered out from log2 fold change and FPKM value of mRNA respectively from updated list of mouse TFs (Riken Transcription Factor Database (http://genome.gsc.riken.jp/TFdb/)) and also using the sequence specific DNA binding (Ashburner et al., 2000). Similar way, differentially regulated and expressed epigenetic regulators from log2 fold change and FPKM value of mRNA respectively were filtered out from published literatures (Shipra et al., 2006, Fazzio et al., 2008, Gendler et al., 2008).

Bioinformatics analysis of miRNA-seq data:

Generation of Normalized count: 20million single end reads were obtained for each sample. Sequenced reads were processed first for quality check with FastQC. Then adapter was removed with Cutadapt and correlation was predicted between replicates. High quality sequencing reads were mapped to mouse reference genome (mm10) by using Bowtie. Mapped reads were processed through miRDeep2 to predict normalized count for both known and novel miRNAs.

Analysis of differentially expressed miRNAs: Output file of miRDeep2 was proceeded further with DESeq2 to generate differentially expressed miRNAs. From differentially expressed miRNAs list, Log2 fold change above 1 were considered as upregulated genes and below -1 were downregulated during differentiation. Likewise, Normalized Count value above 10 were considered as expressed in that particular state.

Analysis of miRNAs associated with LADs: Differentially regulated and expressed miRNAs were filtered out from published Lamin-associated Domain datasets (Dann Peric-Hupkes et al., 2010) and further miRNAs belong to either iLADs, fLADs or cLADs were segregated.

Generation of Venn diagram and heatmap:

Venn diagram: All venn diagrams were generated by using online tool Venny 2.1 (https://bioinfogp.cnb.csic.es/tools/venny/). Venn diagram for expressed mRNA, TF, and ER in neurogenesis (Naïve, NPC, and CN) and ESCs were generated by considering FPKM values above 1 in all the stages. Venn diagram for expressed miRNA in ESCs, lactogenesis (Normal, Primed, and Prolactin) and neurogenesis (Naïve, NPC, and CN) were generated by considering normalized count above 10 in all the stages. Upregulated venn diagram of mRNA, TF, ER and miRNA was generated for ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i) in ESCs, Primed vs Normal, and Prolactin vs Primed in lactogenesis and NPC vs Naïve, and CN vs NPC in neurogenesis by considering log2 fold change above 1 with Pvalue below 0.005 for mRNA and 0.01 for miRNA. Downregulated venn diagram of mRNA and miRNA was generated for ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i) in ESCs, Primed vs Normal, and Prolactin vs Primed in lactogenesis and NPC vs Naïve, and CN vs NPC in neurogenesis by considering log2 fold change below -1 with Pvalue below 0.005 for mRNA and 0.01 for miRNA.

Heatmap for mRNA, TF, and ER: Heatmaps were generated by using Graphpad PRISM software version V 7.0. Heatmap for uniquely expressed top 20 mRNA, TF, and ER in ESCs (ESC(G), ESC(N+2i) and ESC(N)) and neurogenesis (Naïve, NPC, and CN) were generated by considering unique list from expressed venn diagram. Heatmap for highly expressed top 20 mRNA, TF, and ER in ESCs (ESC(G), ESC(N+2i) and ESC(N)) and neurogenesis (Naïve, NPC, and CN) were generated by filtering top 20 mRNA with highest FPKM value. Heatmap for upregulated top 20 mRNA, TF, and ER in ESCs (ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i)) and neurogenesis (NPC vs Naïve, and CN vs NPC) were generated by filtering top 20 mRNA having log2 fold change above 1 and Pvalue below 0.005. Heatmap for downregulated top 20 mRNA, TF, and ER in ESCs (ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i)) and neurogenesis (NPC vs Naïve, and CN vs NPC) were generated by filtering down 20 mRNA having log2 fold change below -1 and Pvalue below 0.005.

Heatmap for miRNA: Heatmap for uniquely expressed top 20 miRNA in ESCs (ESC(G), ESC(N+2i) and ESC(N)), lactogenesis (Normal, Primed, and Prolactin), and neurogenesis (Naïve, NPC, and CN) were generated by considering unique list from expressed venn diagram. Heatmap for highly expressed top 20 miRNA in ESCs (ESC(G), ESC(N+2i) and ESC(N)), lactogenesis (Normal, Primed, and Prolactin), and neurogenesis (Naïve, NPC, and CN) were generated by filtering top 20 miRNA with highest normalized count. Heatmap for upregulated top 20 miRNA in ESCs (ESC(G), ESC(N+2i) and ESC(N)), lactogenesis (Primed vs Normal, and Prolactin vs Primed), and neurogenesis (NPC vs Naïve, and CN vs NPC) were generated by filtering top 20 miRNA having log2 fold change above 1 and Pvalue below 0.01. Heatmap for downregulated top 20 miRNA in ESCs (ESC(G), ESC(N+2i) and ESC(N)), lactogenesis (Primed vs Normal, and Prolactin vs Primed), and neurogenesis (Naïve vs Ground, NPC vs Naïve, and CN vs NPC) were generated by filtering down 20 log2 value having log2 fold change below -1 and Pvalue below 0.01.

Co-expressed gene cluster analysis:

***Clust* analysis:** mRNA FPKM value and normalized count of miRNA were taken as an input for co-expressed cluster analysis by using *Clust* algorithm. Separate group of *Clust* map was generated for mRNA and miRNA. For each state *Clust* evaluated list of upregulated mRNA with downregulated miRNA and vice versa were given as an input to generate cluster map in miRNet.

miRNet analysis: mRNA-miRNA cluster analysis was performed in miRNet version 2.0, 2019. Expressed miRNA and expressed mRNA interaction map was generated by considering

expressed miRNA of ESCs (ESC(G), ESC(N+2i) and ESC(N)), lactogenesis (Normal, Primed, and Prolactin) and neurogenesis (Naïve, NPC, and CN) as input which one later processed through manual batch filter with expressed mRNA of ESCs (ESC(G), ESC(N+2i) and ESC(N)), lactogenesis (Normal, Primed, and Prolactin) and neurogenesis (Naïve, NPC, and CN) respectively with their respective stages.

Upregulated miRNA and expressed mRNA interaction map was generated by considering upregulated miRNA having log2 fold change above 1 and Pvalue below 0.01 of ESCs (ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i)), lactogenesis (Primed vs Normal, and Prolactin vs Primed) and neurogenesis (NPC vs Naïve, and CN vs NPC) as input and by considering expressed mRNA of ESCs (ESC(G), ESC(N+2i) and ESC(N)), lactogenesis (Normal, Primed, and Prolactin) and neurogenesis (Naïve, NPC, and CN) for their respective stages through manual batch filter.

Expressed miRNA and upregulated mRNA interaction map was generated by considering expressed miRNA of ESCs (ESC(G), ESC(N+2i) and ESC(N)), lactogenesis (Normal, Primed, and Prolactin) and neurogenesis (Naïve, NPC, and CN) as input and by considering upregulated mRNA having log2 fold change above 1 and Pvalue below 0.005 of ESCs (ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i)), lactogenesis (Primed vs Normal, and Prolactin vs Primed) and neurogenesis (NPC vs Naïve, and CN vs NPC) for their respective stages through manual batch filter.

Upregulated miRNA and downregulated mRNA interaction map was generated by considering upregulated miRNA having log2 fold change above 1 and Pvalue below 0.01 of ESCs (ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i)), lactogenesis (Primed vs Normal, and Prolactin vs Primed) and neurogenesis (NPC vs Naïve, and CN vs NPC) as input and by considering downregulated mRNA having log2 fold change below -1 and Pvalue below 0.005 of ESCs (ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i)), lactogenesis (Primed vs Normal, and Prolactin vs Primed) and neurogenesis (NPC vs Naïve, and CN vs NPC) for their respective stages through manual batch filter.

Downregulated miRNA and upregulated mRNA interaction map was generated by considering downregulated miRNA having log2 fold change below -1 and Pvalue below 0.01 of ESCs (ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i)), lactogenesis (Primed vs Normal, and Prolactin vs Primed) and neurogenesis (NPC vs Naïve, and CN vs NPC) as input and by considering upregulated mRNA having log2 fold change above 1 and Pvalue below 0.005 of ESCs (ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i)), lactogenesis (Primed vs Normal, and Prolactin vs Primed)

and neurogenesis (NPC vs Naïve, and CN vs NPC) for their respective stages through manual batch filter.

KEGG pathway and GO analysis:

KEGG pathway analysis and Gene Ontology study were conducted in miRNet by considering P value above 1 for both KEGG pathway and GO analysis**.** Upregulated KEGG pathways and GO studies were extracted from downregulated miRNA and upregulated mRNA interaction map which was generated by considering downregulated miRNA having log2 fold change below -1 and Pvalue below 0.01 of ESCs (ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i)), lactogenesis (Primed vs Normal, and Prolactin vs Primed) and neurogenesis (NPC vs Naïve, and CN vs NPC) as input and by considering upregulated mRNA having log2 fold change above 1 and Pvalue below 0.005 of ESCs (ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i)), lactogenesis (Primed vs Normal, and Prolactin vs Primed) and neurogenesis (NPC vs Naïve, and CN vs NPC) for their respective stages through manual batch filter.

Downregulated KEGG pathways and GO studies were extracted from upregulated miRNA and downregulated mRNA interaction map which was generated by considering upregulated miRNA having log2 fold change above 1 and Pvalue below 0.01 of ESCs (ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i)), lactogenesis (Primed vs Normal, and Prolactin vs Primed) and neurogenesis (NPC vs Naïve, and CN vs NPC) as input and by considering downregulated mRNA having log2 fold change below -1 and Pvalue below 0.005 of ESCs (ESC(N+2i) vs ESC(G), ESC(N) vs ESC(N+2i)), lactogenesis (Primed vs Normal, and Prolactin vs Primed) and neurogenesis (NPC vs Naïve, and CN vs NPC) for their respective stages through manual batch filter.

QIAGEN ingenuity pathway analysis:

Associated pathways and diseases were extracted by considering together upregulated and downregulated mRNA having log2 fold change above 1 and below -1 respectively with Pvalue below 0.005 of lactogenesis (Primed vs Normal, and Prolactin vs Primed) and neurogenesis (NPC vs Naïve, and CN vs NPC) as input.

Bioinformatics analysis of CLASH-seq analysis:

***Hyb* analysis:** For each sample 10million single end reads were generated and sequence quality check was performed with FastQC program followed by trimming adapters using Cutadapt.

Correlation between biological replicate was predicted by Spearman's correlation after normalization of data. Chimeras (RNA-RNA ligation, that came from different region in the genome) were estimated by using hybrid analysis as described in *hyb* algorithm.

Generation of FPKM and log2 fold change: Also, separately mRNA expression was predicted by mapping filtered reads to mouse reference genome by Tophat2 followed by FPKM prediction with Cufflinks and differential expression analysis with DESeq2.

Generation of Normalized count and log2 fold change: Same for miRNA, filtered reads were normalized with miRDeep2 and differential expression was analysed with DESeq2 by considering normalized value from miRDeep2.

CIMS analysis: Cross-link Induced Mutation Sites (CIMS) analysis for AGO was performed as described in CLIP Tool Kit (CTK). Filtered reads were mapped to reference genome mm10 by using Novoalign. Novoalign also provides information about mutations caused by UV-crosslinking. After mapping, it detects mutations based on substitution, deletion and insertion separately. Uniquely CLIP-tagged mapped reads were collapsed to avoid PCR duplicates based on coordinates. Here, we considered only deletion profile because AGO binding sites show deletion rather than substitution upon UV-crosslinking. Total expressed mRNAs were filtered out with mRNAs, that showed deletion. Then, -20 upstream and +20 downstream to deletion sites were analysed for exact sequence match with seed region (GTGAGG) of mmu-miR-122-5p.

Results and Discussions

Objective I

MicroRNA dynamics during embryonic stem cell maintenance and differentiation

Morphologically ESC (N +2i) showed more similarities with ESC (G):

The state of pluripotency is the fundamental state of embryonic development. Pluripotent cells have capacity to undergo all somatic and germline lineages. This stage is acquired during early development of zygote. Zygote is a totipotent state which develops into blastocyst and trophectoderm on embryonic day 3.5. Blastocyst contains inner cell mass (ICM) which is covered by trophectoderm, an extraembryonic epithelial layer. The inner cell mass cells at this stage are pluripotent in nature and harbour the potential to develop into any kind of germ layers. This state of cells on embryonic day 3.5 is considered as Ground state of pluripotency ESC (G). This state is well characterized by uniform expression of core pluripotent factors, reactivation of X chromosome and global DNA hypomethylation. This is a kind of blank state with unbiased developmental potential. As development proceeds on embryonic day 5, intrinsic stimuli facilitate the transition towards Primed state of pluripotency which is a epigenetically restricted stage and initiate lineage specification. The transition state of cells in between embryonic day 3.5 to 5 is considered as Naive state of pluripotency ESC (N). Though the gene expression of core regulatory factors are very much similar in between transient states of pluripotency, still much variations are there in global gene expression pattern. Even changes in the expression of miRNAs during this transitions are very critical. MiRNA modulates changes in pluripotent specific genes by targeting them that leads to lineage commitment. These state of pluripotency can be achieved *in vitro* through various external factors like Leukemia Inhibitory Factor (LIF), MAP2K inhibitor (PD) and GSK3ß inhibitor (CH). The above extent of pluripotency depends upon various culture condition and source of origin. From 1981 to till date, there has been many standard protocol described to maintain a pure population of embryonic stem cells (ESCs). (Fig. A1. A1-A2). ESC (G) stage was cultured with standard 2i/LIF serum free medium in N2B27 medium whereas ESC (N+2i) was grown under 10% of serum medium with 2i/LIF. ESC (N) state was cultured with normal serum ESC medium. There has been a detailed study of mRNA expression profile of Ground and Naive

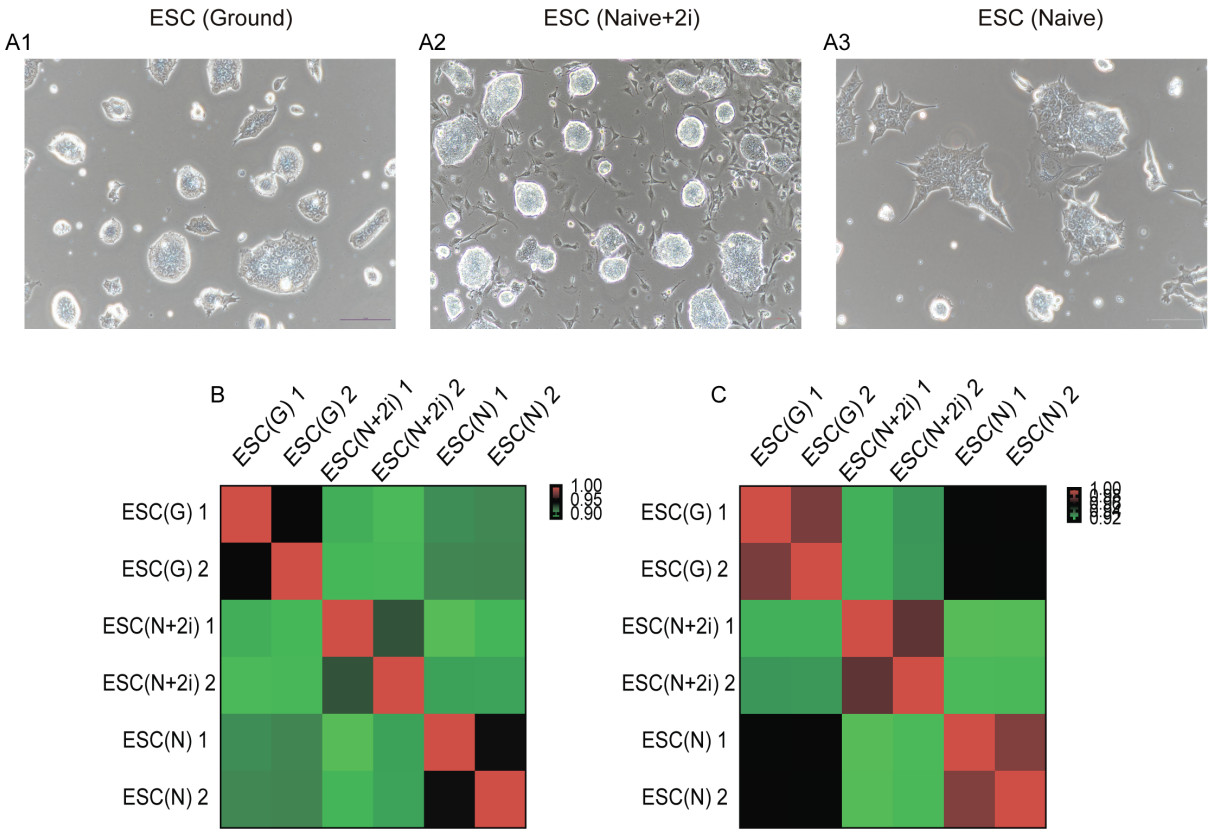


Fig. A1. Evaluation of ESC colonies and correlation matrix of biological replicates: 20X bright field images of Embryonic stem cells cultured in different culture conditions; **A1.** ESC (ESC (G)) stage was cultured with standard 2i/LIF serum free medium in N2B27 medium., **A2.** ESC (Naive+2i) stage was cultured with 10% of serum medium with 2i/LIF., A3. ESC (Naive) stage was cultured with 10% serum in ESC medium. Heatmap representing Spearman's correlation map of ESC (G), ESC (N+2i) and ESC (N) **B**. miRNA–seq and **C.** mRNA-seq in between biological replicates and different samples with Pvalue<2.2e-16. Spearman's correlation was 0.90 between the replicate data sets which suggested the more similarities between biological replicates than between different samples. Details of correlation values are available in Table 1-2.

state by Ghimire et al., 2017 and miRNA expression profile by Moradi et al., 2017 but, how the extent of pluripotency will be compromised by the addition of 2i in presence of serum will be interesting to look forward because both follows different pathways to maintain pluripotency. To obtain a detail view of transcription, both mRNAs' along with miRNAs' transcriptome profile were analysed.

Previous study has shown that compared to Naive ESCs, Ground ESCs appears more as dome shaped colony. In this study also ESC (N) are grown as flattened like colony as shown in Fig. A1. During transition from ESC (N) to ESC (G) state ESC colonies are more round like dome shaped and brightened. But ESCs cultured in Naive+2i which is in serum along with 2i, PD (MAP2K inhibitor), CH (GSK3ß inhibitor) inhibitors and also maintained on inactivated feeder layer, showed more like domed shape colony. The morphology of ESC (N+2i) is closely resemblance with ESC (G) state.

Comprehensive microRNA and mRNA transcriptome profiles during early embryonic development

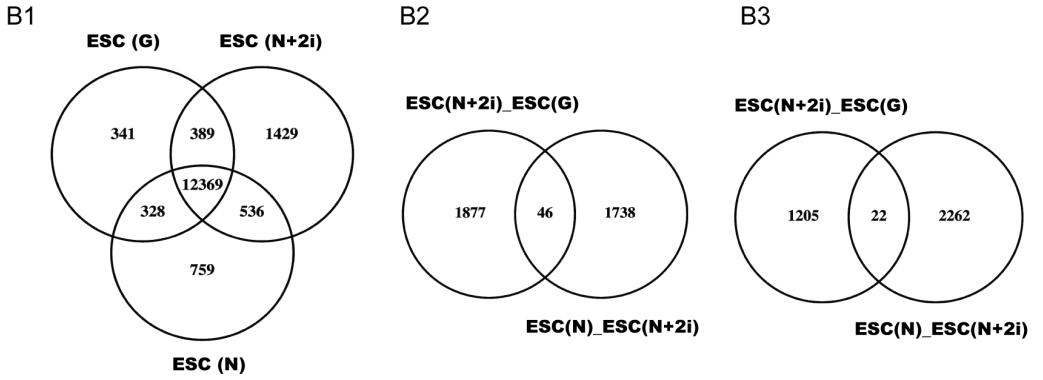
Two biological replicate of ESC (G), ESC (N+2i) and ESC (N) condition were processed for miRNAs and mRNAs sequencing using next generation illumina sequencing platform at a depth of 20 million reads per sample. Quality checking of sequenced reads were performed using FastQC which showed ~50% GC content and above 95% accuracy of Q20 and Q30 (Table 3, Table 4). Spearman's correlation was performed between the replicate data sets which showed the value of 0.90 with pvalue<2.2e-16 (Fig. A1. B-C) (Table 1, Table 2). Firstly, complete mRNA transcriptome profiles were analysed for all three stages of ESCs. In ESC (G) stage a total of 13427 genes were found to be expressed and in ESC (N+2i) and ESC (N) stages a total of 14723 and 14798 genes were found to be expressed respectively. Among them 12369 genes were found to be expressed commonly in all three stages and 341, 1429 and 759 number of genes were uniquely expressed in ESC (G), ESC (N+2i) and ESC (N) respectively (Fig. A2. B1). A comparison of differentially expressed miRNAs between stages of ESC development provided information of total of 1923 genes that were upregulated and 1227 genes were downregulated between ESC (N+2i) vs ESC (G) and 1784 genes were upregulated and 2284 gene were downregulated in ESC (N) vs ESC (N+2i). Further, it was found that 46 genes were upregulated both in ESC (N+2i) vs ESC (G) as well as ESC (N) vs ESC (N+2i) stages. Also, 22 common genes were found to be downregulated (Fig. A2. B2-B3). Among these uniquely expressed (Table 5), highly expressed (Table 6), top 30 upregulated (Table 7) and top 30 downregulated (Table 8) list of mRNAs lists were listed out.

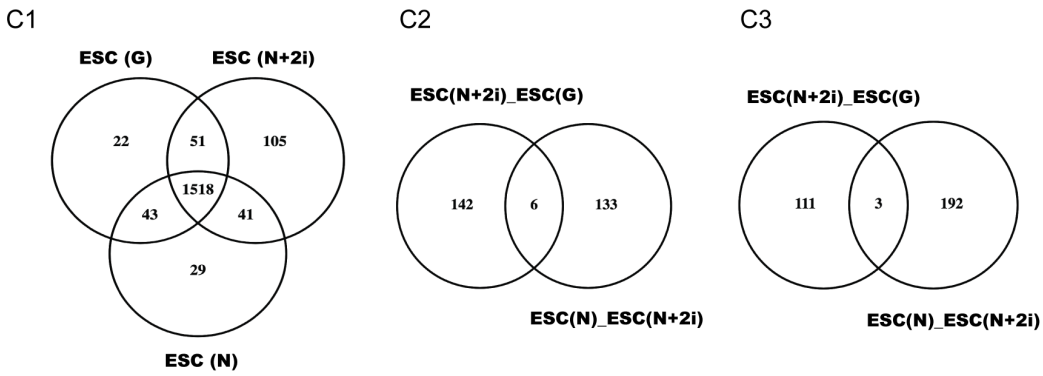
Further, to have a deeper understanding about differentially expressed transcription factors (TF) and epigenetic regulators (ER), their profiles have been derived from the mRNA list. It was found that there were 1634, 1715 and 1631 TF genes expressed in ESC (G), ESC (N+2i) and ESC (N) condition respectively and 1518 were found to express in all three conditions/stages of ESCs. 22,

105 and 29 TFs were uniquely found in ESC (G), ESC (N+2i) and ESC (N) respectively (Fig. A2. C1). Genes like *Asz1, Hesx1, Foxr1, Klf1* and *Lhx2* are uniquely expressed TFs for ESC (G). *Nkx6-3, Cspg4, Foxo6, Foxg1* and *Sox1* are showing unique expression in ESC (N+2i). Similarly, *Egr, Rem2, Pax6, Nlrp9b* and *Dll1* are uniquely enriched for ESC (N). From ESC (G) to ESC (N+2i) 148 TFs were found to be upregulated and 139 were found to be down regulated. Some developmentally important TFs like *Sox9, Irf2, Dlx2, Zeb2* and *Ebf3* are getting upregulated and *Arntl2, Dppa3, fgfbp1, Asz1* and *Rhox6* are getting downregulated from ESC (G) to ESC (N+2i) state. During ESC transition, 6 genes were found to be commonly upregulated and 142 were specific to ESC (N+2i) vs ESC (G). 133 genes were specific to ESC (N) vs ESC (N+2i). Similarly, 114 genes were found to be downregulated during ESC (G) to ESC (N+2i) and 195 genes during ESC (N+2i) to ESC (N) transition. Three genes were found to be commonly downregulated in both whereas, 111 genes were found to be unique to ESC (G) and ESC (N) and 192 genes were in ESC (N) vs ESC (N+2i) (Fig. A2. C2-C3). Among these above-mentioned TFs uniquely expressed (Table 9), highly expressed (Table 10), top 30 upregulated (Table 11) and top 30 downregulated (Table 12) list of TFs lists were listed out. Like TFs, there are 747, 747 and 741 total ERs are expressing among ESC (G), ESC (N+2i) and ESC (N) condition respectively. 8, 28 and 4 ERs are unique to ESC (G), ESC (N+2i) and ESC (N) condition respectively but 695 are common in all three (Fig. A2. D1). Epigenetic regulators like *Grid2* and *9130023H24Rik* are unique to ESC (G) stage. *Gata6, Runx2* and *Gata1* are unique ER for ESC (N+2i). *Smc1b* and *Gata3* are uniquely expressing in ESC (N). Majority of ERs seems to be stabilized during different state of pluripotency transition. Minimal changes were observed in differentially regulated ERs. Total 9 (Ciita, Bloc1s1, Gata2, Gata6, Phldb1, Gata4, H1f0, D1Pas1 and Ptpn23) are upregulating from ESC (G) to ESC (N+2i) stage and 8 (Ddx4, Dnmt3b, Hat1, etc.) are from ESC (N+2i) to ESC (N). In downregulation, total 8 (Grid2, Ddx4, Cited1, Bicd1, etc.) are appeared in ESC (N+2i) vs ESC (G) and 12 (Bloc1s1, Ciita, Runx2, Cbx4, Gata4, etc.) are in ESC (N) vs ESC (N+2i) (Fig. A2. D2-D3). Among these ERs highly expressed (Table 13), uniquely expressed (Table 14), top 30 upregulated (Table 15) and top 30 downregulated (Table 16) list of ERs lists were listed out.

To understand the fine tuning of various mRNA transcripts including TF and ERs by miRNAs at three stages of ESC development, microRNAs were comprehensively profiled by miRNA-seq using illumine NGS platform. miRNA seq data analysis revealed expression of 328, 480 and 388 number of miRNA genes in ESC (G), ESC (N+2i) and ESC (N) stages respectively (Fig. A5. B1). Among them 289 miRNAs were found to be common in three stages of ESCs and 14, 79 and 28 miRNAs were found to exclusively expressed in ESC (G), ESC (N+2i) and ESC (N) respectively.

miRNA like mmu-miR-743b-5p, mmu-miR-871-5p, mmu-miR-743b-3p, mmu-miR-743a-3p, etc. are unique to ESC (G). mmu-miR-690, mmu-miR-199a-3p, mmu-miR-199b-3p, mmu-miR-199a-





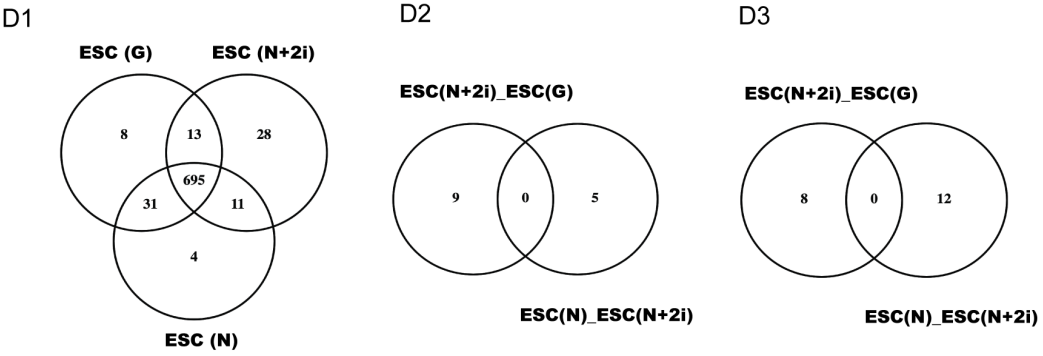


Fig. A2. Schematic representation of statistically analysed mRNAseq dataset: A. Flow chart representing expressed (FPKM≥1) and differentially regulated (log2 fold change) mRNAs, TFs and ERs in ESC (G), ESC (N+2i), ESC (N). Venn diagram representing **B1.** expressed (FPKM≥1) mRNAs, **B2.** differentially upregulated (log2 fold change≥1) mRNAs and **B3.** differentially downregulated (log2 fold change≤1) mRNAs in between ESC (G), ESC (N+2i), ESC (N). Venn diagram representing **C1.** expressed (FPKM≥1) TFs, **C2.** differentially upregulated (log2 fold change≥1) TFs and **C3.** differentially downregulated (log2 fold change≤1) TFs in between ESC (G), ESC (N+2i), ESC (N). Venn diagram representing **D1.** expressed (FPKM≥1) ERs, **D2.** differentially upregulated (log2 fold change≥1) ERs and **D3.** differentially downregulated (log2 fold change≤1) ERs in between ESC (G), ESC (N+2i), ESC (N).

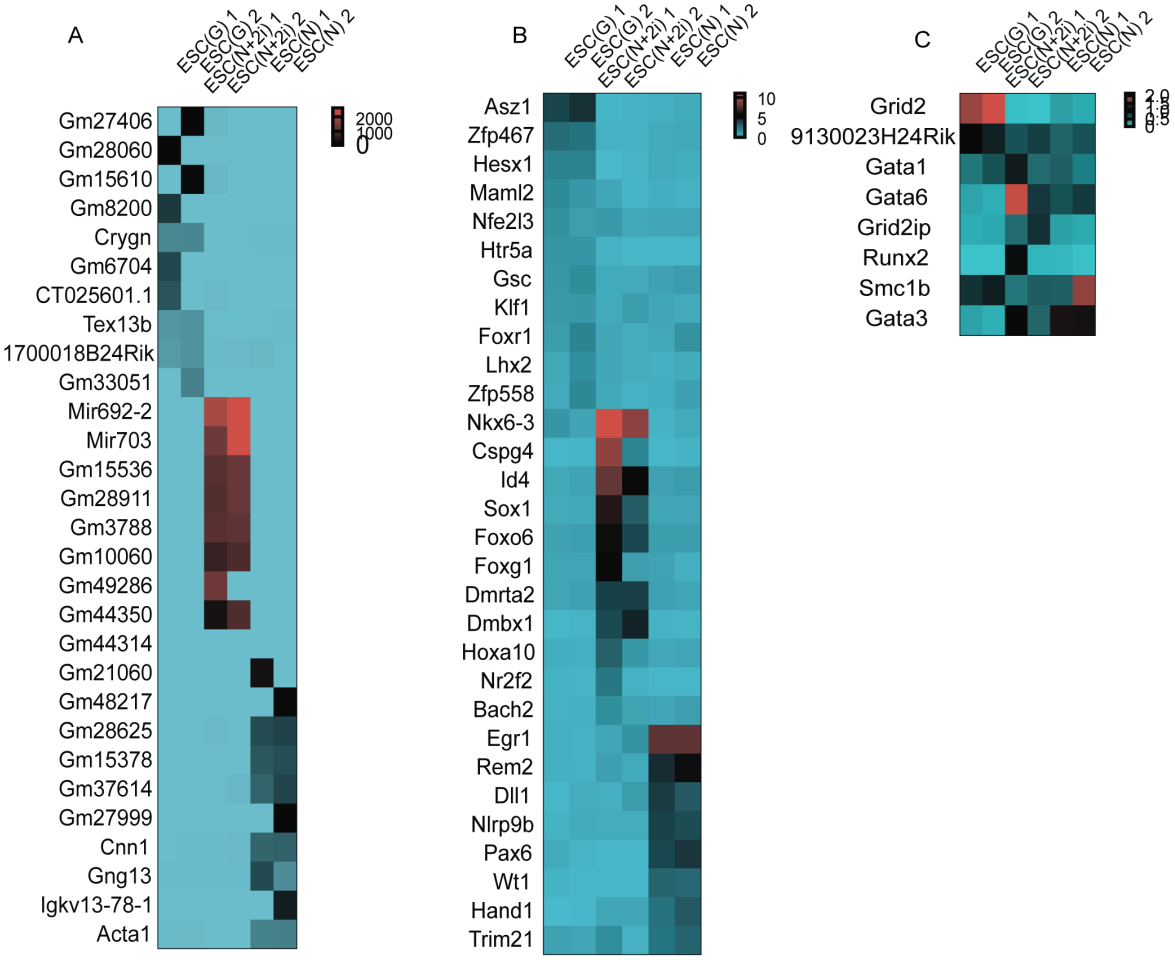


Fig. A3. Comparative analysis of uniquely expressed genes in different stages of ESC: Heatmap representing Exclusively expressed **A.** mRNAs, **B.** TFs and **C.** ERs in ESC (G), ESC (N+2i), ESC (N). Lists of genes uniquely expressed mRNAs, TFs, and ERs with FPKM values are available in Table 5, Table 9 and Table 14.

5p, mmu-miR-677-5p, etc. are exclusively expressing in ESC (N+2i). mmu-miR-1a-3p, mmu-miR-669m-5p, mmu-miR-466m-5p, mmu-miR-466d-5p, mmu-miR-466n-5p, etc. are uniquely expressing in ESC (N). Total 123 (mmu-miR-6238, mmu-miR-466i-5p, mmu-miR-3963, mmu-miR-3968, etc.) are showing upregulation from ESC (G) to ESC (N+2i) and 147 (mmu-miR-451a, mmu-miR-302c-5p, mmu-miR-592-5p, mmu-miR-1298-5p, etc.) are from ESC (N+2i) to ESC (N).

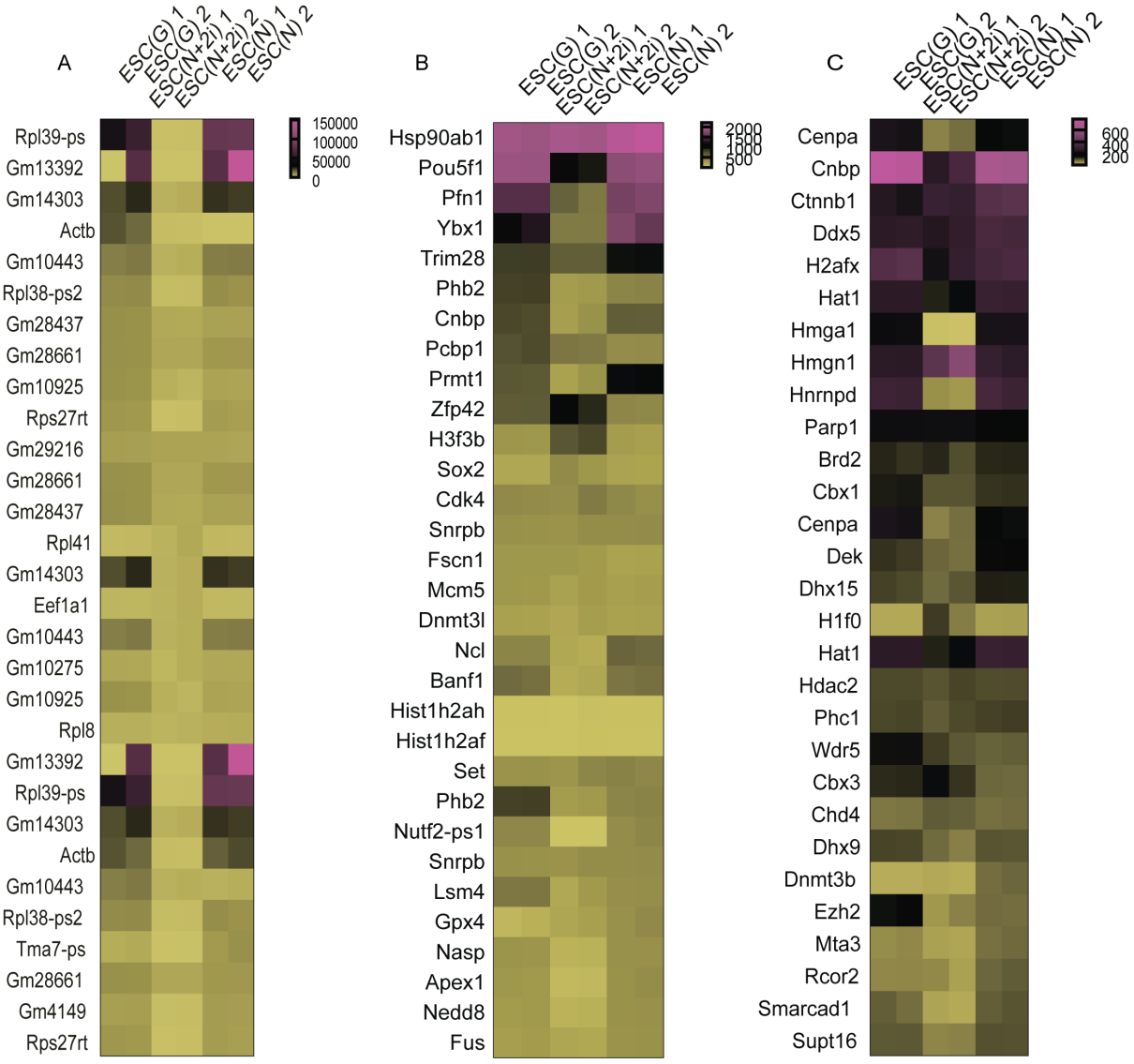


Fig. A4. Comparative analysis of highly expressed genes in different stages of ESC: Heatmap representing Exclusively expressed **A.** mRNAs, **B.** TFs and **C.** ERs in ESC (G), ESC (N+2i), ESC (N). Lists of highly expressed mRNAs, TFs, and ERs with FPKM values are available in Table 6, Table 10 and Table 13.

Among which only 2 are common and 121 are unique to ESC (N+2i) vs ESC (G) and 115 are in ESC (N) vs ESC (N+2i). Similarly, 117 (mmu-miR-451a, mmu-miR-743b-5p, 881-3p, 463-5p, etc.) was downregulating from ESC (G) to ESC (N+2i) and 257 (mmu-miR-5106, mmu-miR-3963, mmu-miR-690, mmu-miR-6238, etc.) are from ESC (N+2i) to ESC (N) condition among them 15 are common in both the transition (Fig. A5. B2-B3). Among these above-mentioned miRNAs highly expressed (Table 17), top 25 upregulated (Table 18) and top 25 downregulated (Table 19) list of miRNAs lists were listed out (Fig. A6).

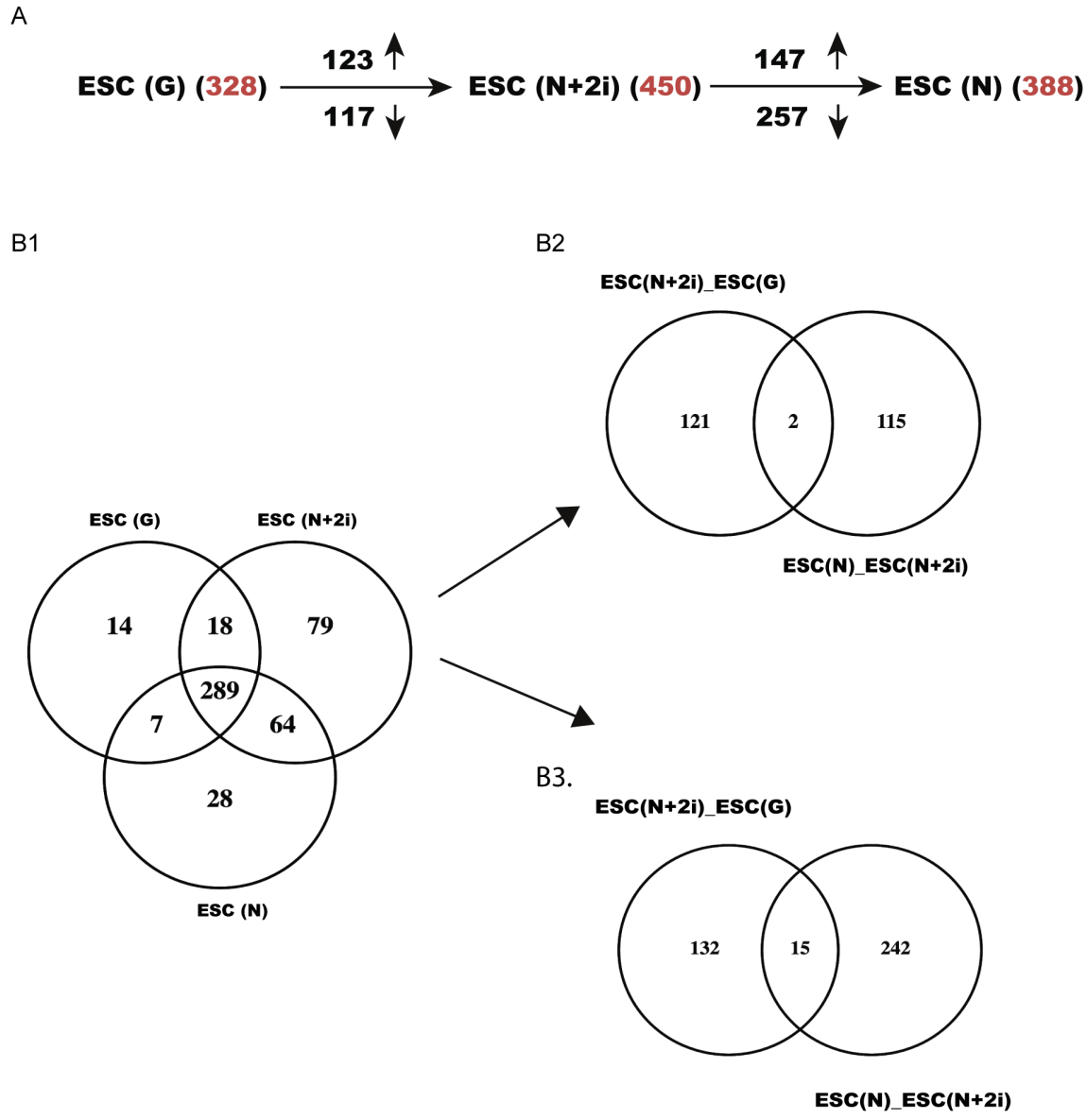


Fig. A5. Schematic representation of statistically analysed miRNAseq dataset: A. Flow chart representing expressed (NC≥10) and differentially regulated miRNAs in ESC (G), ESC (N+2i), ESC (N). Venn diagram representing **B1.** expressed miRNAs (NC≥10), **B2.** differentially upregulated (log2 fold change≥1) miRNAs, **B3.** differentially downregulated (log2 fold change≤1) miRNAs in between ESC (G), ESC (N+2i), ESC (N).

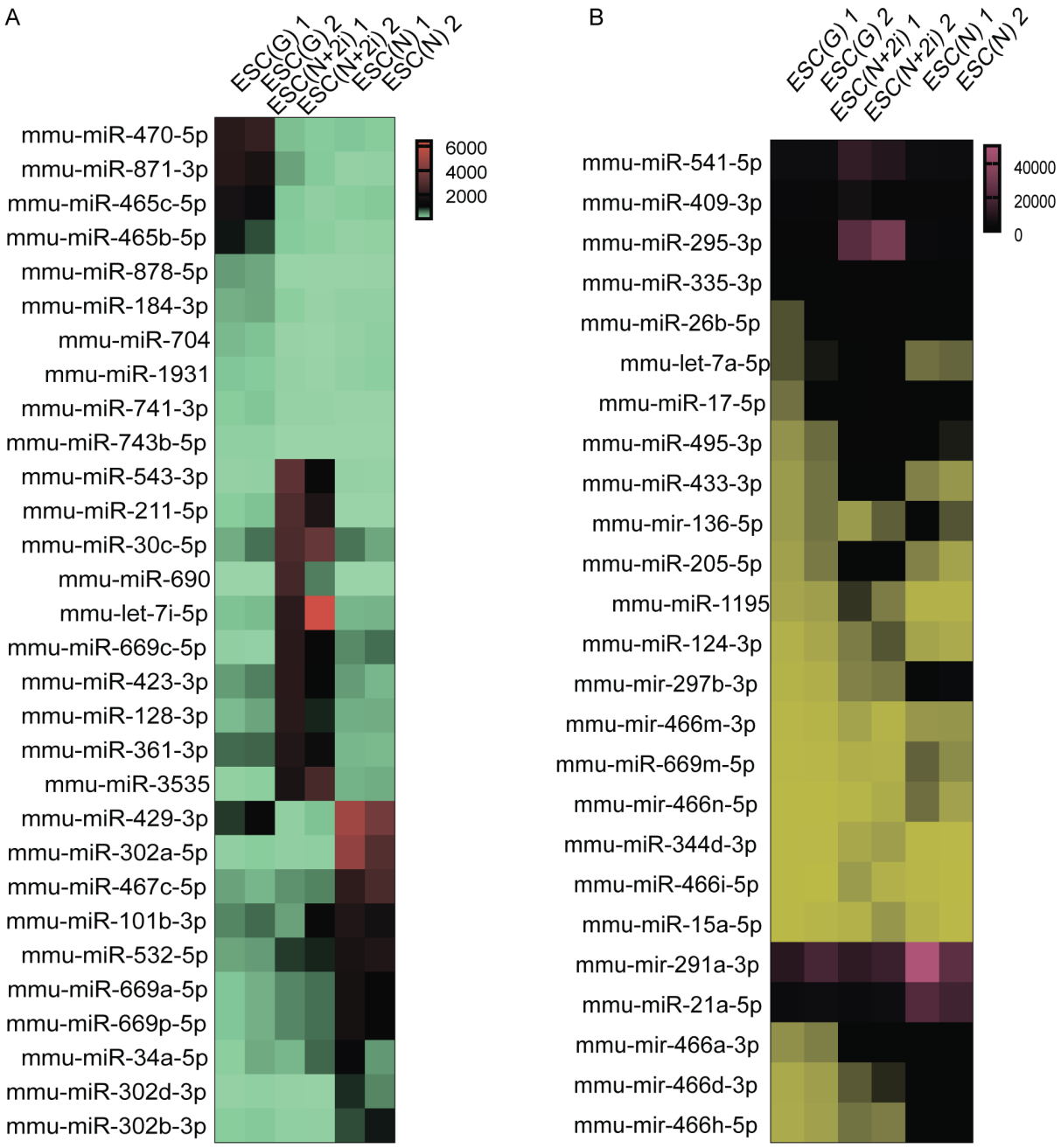


Fig. A6. Comparative analysis of uniquely and highly expressed miRNA genes in between different stages of ESC: Heatmap representing **A.** Uniquely expressed and **B.** Highly expressed miRNAs in ESC (G), ESC (N+2i), ESC (N). Lists of uniquely and highly expressed miRNAs with Normalized counts are available in Table 17.

ESC (G) stage specific miRNA-mRNA interactome analysis:

In order to understand the role of miRNAs in maintenance of ESC (G) pluripotency network, expressed mRNAs (>1FPKM) and expressed miRNA (>10 Normalized count) from ESC (G) stage were considered for miRNA-mRNA interactome analysis (Fig. A7. A-A'). Interactome map has some major hubs of miRNAs those are controlling more than 100 coding genes as, *Wnt11, Gpc4, Fzd7, Fzd8, Pkm, Camk2b, Ccnd1, Myc, Jun, Tcf7*, etc. in that particular state. Among these coding genes many important TFs like *Pcgf2, Id1, Bmi1, Nanog, Stat3, Myc, Zfhx3, Mapk14, Pou5f1, Klf4, Sox2, Smad,* etc. are regulated by mmu-miR-9-5p, mmu-miR-24-3p, mmu-miR-1195, mmu-miR-7b-5p, etc. and controlling Signaling pathways regulating pluripotency. Cell cycle in ESC (G) is regulated by targeting TFs as, *Cdk7, Mcm5, Tfdp2, Atm, Tgfb1, Hdac2, E2f1, Smad3*, etc. through mmu-miR-1195, mmu-miR-24-3p, mmu-miR-466a-3p, mmu-miR-122-5p, etc. These above-mentioned miRNAs are also regulating expression of *Tgfb1, Tgif1, Sp1, Smad4, Smad7, Smad1,* etc. to control TGF-beta signalling pathway. Some highly expressed miRNAs' in this interactome hub are mmu-mir-340-5p, mmu-mir-9-5p, mmu-mir-329, mmu-mir-7b-5p, mmu-mir-149-5p, mmu-mir-26a-5p, mmu-mir-425-5p, mmu-mir-17-5p, mmu-mir-30e-5p, and others are listed out in the Table 17. Pathways regulated by these networks are Hippo signaling pathway, Wnt signaling pathway, P53 signaling pathway, cell cycle, FoxO signaling pathway, Signaling pathways regulating pluripotency, etc.

ESC(N+2i) specific miRNA-mRNA interactome analysis:

ESC(N+2i) specific interactome map was extracted by considering expressed miRNAs (>10 Normalized count) and expressed mRNAs (>1FPKM) in ESC (N+2i) condition (Fig. A7. B-B'). Interactome map has some major hubs of miRNAs who are controlling many developmentally important genes like, *Wnt11, Twist2, Fzd7, Fzd8, Fzd5, Lmna, Mdm2, Pkm, Pau, Tgfb1, Actb, Akt1, Akt3*, etc. in that particular state. Among these coding genes many important TFs like *Pcgf2, Meis1, Id1, Bmi1, Nanog, Stat3, Myc, Zfhx3, Mapk14, Pou5f1, Klf4, Sox2, Smad,* etc. are regulated by mmu-miR-9-5p, mmu-miR-124-3p, mmu-miR-15a-5p, mmu-miR-7b-5p, etc. and controlling Signaling pathways regulating pluripotency. Cell cycle in ESC (G) is regulated by targeting TFs as, *Cdk7, Mcm5, Tfdp2, Atm, Ccna2, Tgfb1, Hdac2, E2f1, Smad3, Myc, Mcm7*, etc. through mmu-miR-1195, mmu-miR-181a-5p, mmu-miR-17-5p, mmu-miR-30e-5p, etc. These above-mentioned miRNAs are also regulating expression of *Tgfb1, Id1, Id2, Tgif1, Sp1, Smad4, Smad7, Smad1,* etc. to control TGF-beta signalling pathway. Due to addition of Gsk3ß inhibitor, elevation of Wnt signalling

pathway was observed which is regulated through mmu-mir-24-3p, mmu-mir-181a-5p, mmu-mir-329-3p, mmu-mir-466f-3p, etc. by targeting TFs like, *Nfatc2, Ctbp2, Nfatc3, Dkk-1, Ctbp1, Ctnnb1, Myc, Smad4, Nik*, etc. ESC (N+2i) specific miRNA-mRNA interactome analysis displayed major miRNAs hubs as mmu-mir-340-5p, mmu-mir-15a-5p, mmu-mir-9-5p, mmu-mir-329-3p, mmu-mir-124-3p, mmu-mir-181a-5p, mmu-mir-7b-5p, mmu-mir-149-5p, mmu-mir-26a-5p, mmu-mir-17-5p, mmu-mir-425-5p, mmu-mir-30e-5p, mmu-mir-24-3p, mmu-mir-324-3p, mmu-mir-34b-5p, mmu-mir-1195, and others are listed out in the Table 17. Among these many are common as described previously in ESC (G) ESCs. But mmu-mir-290 family which is a land mark miRNA marker in pluripotency is maintaining major central hubs in ESC (N+2i) as compared to ESC (G) state. ESC (N+2i) interactome map was enriched with pathways like, MAPK signaling pathway, FoxO signaling pathway, Oxyticin signaling pathway, Neurotrophin signaling pathway, Hippo signaling pathway, signaling pathway regulating pluripotency, ErbB signaling pathway, etc.

ESC (N) specific miRNA-mRNA interactome analysis:

Expressed miRNAs (>10 Normalized count) and mRNAs (>1FPKM) from ESC (N) condition were used to predict miRNA-mRNA interactome map of Naive ESC (Fig. A7. C-C'). Interactome map has some major hubs of miRNAs who are controlling many developmentally important genes like, *Wnt11, Gng2, Nfkbia, Vegfb, Fzd3, Ctbp2, Birc5, Gsk3b, mTOR, Cdh1, Cycs, Fgf5, Grb2, Mdm2, Pdgfb, Pkm, Rac2, Wnt3, Bmp4, Fgf13, Fos, Ccnd1, Fzd1*, etc. in that particular state. Among these coding genes many important TFs like *Pcgf2, Pax6, Id2, Id1, Bmi1, Fzd1, Nanog, Stat3, Myc, Zfhx3, Mapk14, Pou5f1, Klf4, Sox2, Smad,* etc. are regulated by mmu-miR-329-3p, mmu-miR-466f-3p, mmu-miR-24-3p, mmu-miR-9-5p, etc. and controlling Signaling pathways regulating pluripotency. Cell cycle in ESC (G) is regulated by targeting TFs as, *Cdk7, Mcm5, Tfdp2, Atm, Ccna2, Tgfb1, Hdac2, E2f1, Smad3, Myc, Mcm7*, etc. through mmu-miR-495-3p, mmu-miR-26a-5p, mmu-miR-122-5p, mmu-miR-24-3p, etc. Due to removal of MAP2K inhibitor, elevation of MAPK signalling pathway was observed which is regulated through mmu-mir-340-5p, mmu-mir-30e-5p, mmu-mir-129-5p, mmu-mir-466f-3p, etc. by targeting TFs like, *Relb, Asb3, Tgfb1, Nfatc3, Atf2, Srf, Nf1, Elk1, Myc, Sos1, Max, Jun, Mapk14*, etc. ESC (N) specific miRNA-mRNA interactome analysis provided many central miRNAs hubs like, mmu-mir-340-5p, mmu-mir-9-5p, mmu-mir-329-3p, mmu-mir-124-3p, mmu-mir-7b-5p, mmu-mir-149-5p, mmu-mir-17-5p, mmu-mir-301b-3p, and others are listed out in the Table 17. This interactome map is enriched with pathways like, Hippo signaling pathway, MAPK signaling pathway, signaling pathway regulating pluripotency, Neurotrophin signaling pathway, FoxO signaling pathway, Oxytocin signaling pathway, ErbB signaling pathway, etc.

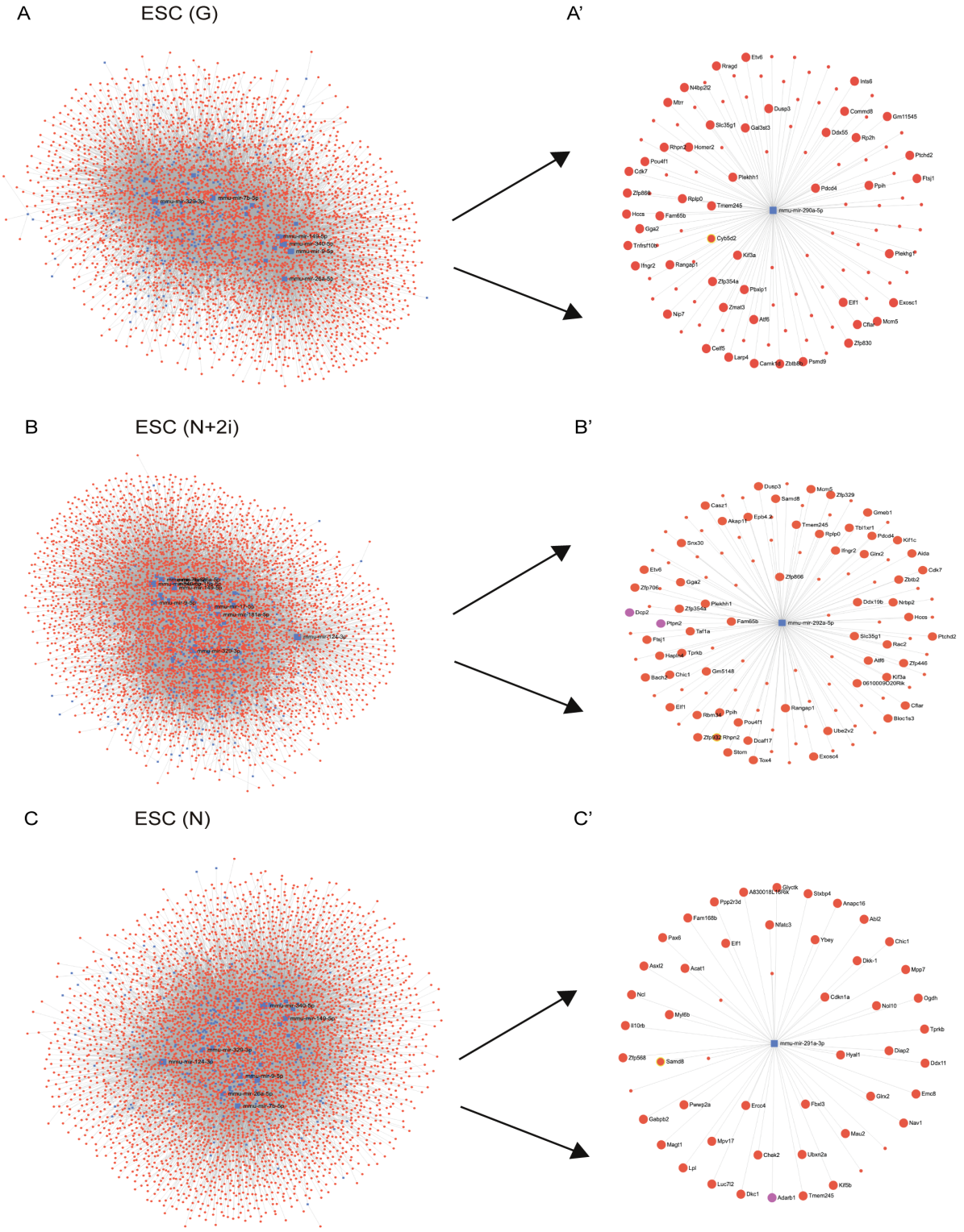


Fig. A7. miRNA-mRNA network map of expressed miRNAs and its experimentally validated mRNAs targets in different stages of ESC: Computationally analysed interactome map of expressed miRNA and expressed mRNA as a whole and zoomed in view of one of the highly expressed miRNA. **A.** expressed mRNA

(FPKM≥1) with expressed miRNA (NC≥10) of ESC(G), **A'.** Zoom in view of mmu-mir-290a-5p, one of the important miRNA's hub from A, **B.** expressed mRNA (FPKM≥1) with expressed miRNA (NC≥10) of ESC(N+2i), **B'.** Zoom in view of mmu-mir-292a-5p, one of the important miRNA's hub from B, **C.** expressed mRNA (FPKM≥1) with expressed miRNA (NC≥10) of ESC(N), **C'**. Zoom in view of mmu-mir-291a-3p, one of the important miRNA's hub from C.

Interactome analysis of upregulated mRNAs vs down regulated miRNAs during ESC (G) to ESC (N+2i) transition

To analyse the difference between two culture condition under 2i/LIF and 2i/LIF/serum, interaction map was extracted by considering mRNAs which are upregulating during ESC (G) to ESC (N+2i) transition and miRNA which are downregulating (Fig. A8. B-B'). This will give an idea about miRNAs which are now no longer required to maintain pluripotency in 2i/LIF/serum medium. But, they had vital roles in suppressing mRNAs in ESC (G) state and now those mRNAs are important for ESC (N+2i) condition. Some developmentally important upregulated genes which are targeted by downregulated miRNAs are *mTOR, IL-6, Csf1, Itgb4, Gys1, Egfr, Gnb4, Rps6, Gng12, Slc1a1, Nos3, Prkca, Jun, Adcy6, Ptgs2, Tlr4*, etc. Among these above-mentioned genes, some transcription factors like *Dkk, Nfatc4, Jun, Gucy1a3, Whsc111, Per1, Bmal1*, etc. are also controlled by downregulated miRNAs during ESC (G) to ESC (N+2i) transition. This interaction map has important central miRNAs hub with target genes 10 or above are, mmu-mir-129-5p, mmu-mir-149-5p, mmu-let-7b-5p, mmu-mir-425-5p, mmu-mir-301b-3p, mmu-mir-136-5p, mmu-mir-298-5p, mmu-mir-10b-5p, mmu-mir-132-3p, mmu-mir-342-3p, mmu-mir-743a-3p and 743b-3p. Top upregulated pathways from ESC (G) to ESC (N+2i) through this interaction map are PI3K-Akt signaling pathway, ECM receptor interaction, Glutamatergic synapse, Oxytocin signaling pathway, HIF-1 signaling pathway, GABAergic synapse, NF-kappa B signaling, Wnt signaling pathway, etc. Here, presence of 2i in serum is showing upregulation of Wnt signaling which is due to GSK3ß inhibitor. Here upregulation of Wnt is mediated by downregulation of miRNAs such as, mmu-mir-1a-3p, mmu-mir-342-3p, mmu-mir-149-5p, mmu-mir-3470a, mmu-mir-320-3p, mmu-mir-122-5p, mmu-mir-741-3p and mmu-mir-291a-3p These miRNAs are targeting *Fzd7, Dkk-1, Nfatc4, Prickle2, Jun* and *Prkca*. Also, highly upregulated Akt pathway is controlled by mmu-mir-1a-3p, mmu-mir-342-3p, mmu-mir-149-5p, mmu-mir-3470a, mmu-mir-320-3p, mmu-mir-122-5p,

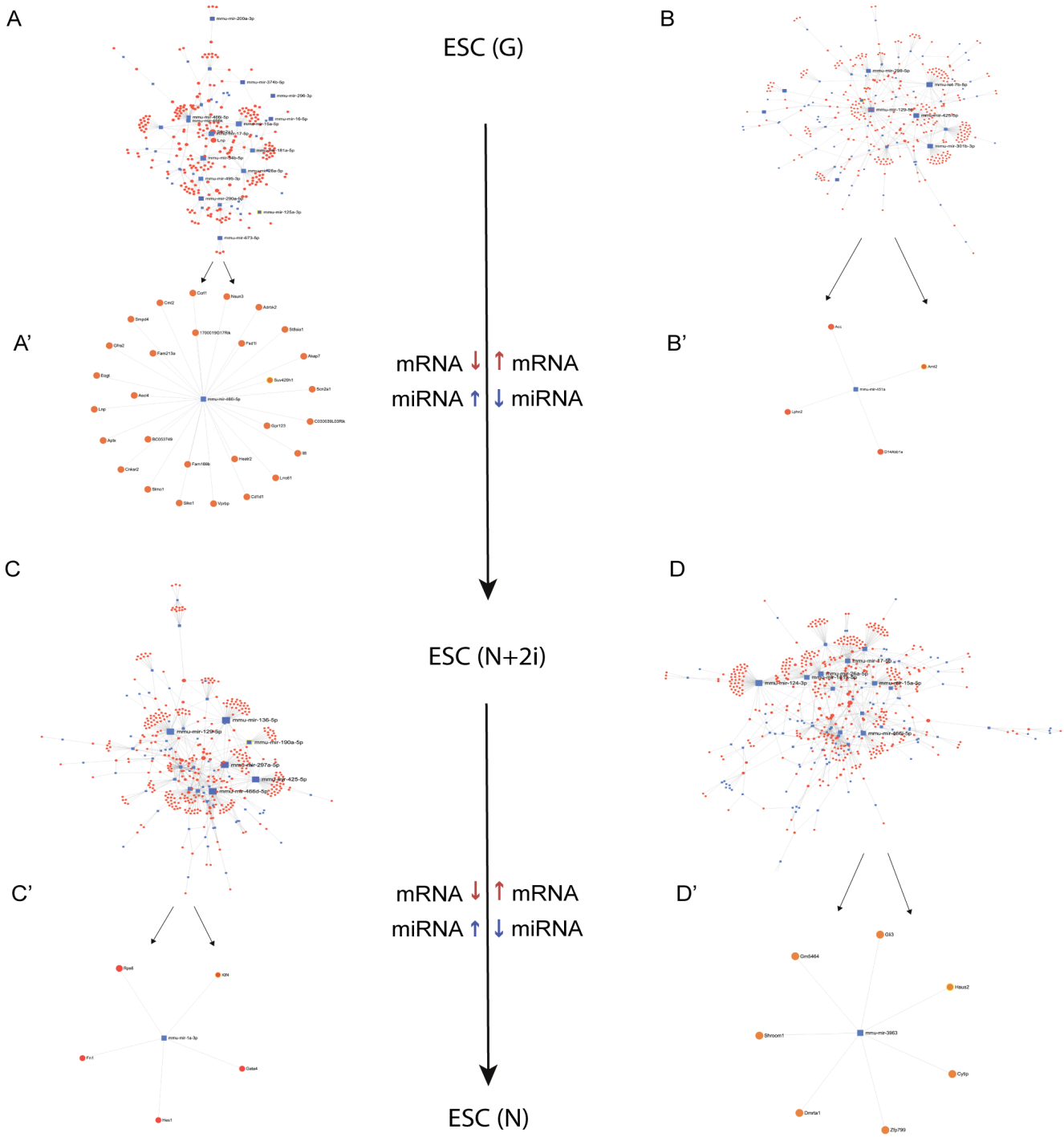


Fig. A8. miRNA-mRNA network map of differentially expressed miRNAs and its experimentally validated mRNAs targets between different stages of ESC: Computationally analysed interactome map of differentially expressed mRNA and differentially expressed miRNA as a whole and zoomed in view of one of the highly expressed miRNA. **A.** differentially downregulated mRNA (log2 fold change≤-1) with differentially upregulated miRNA (log2 fold change≥1) from ESC(N+2i) vs ESC(G), **A'.** Zoom in view of mmu-mir-466i-

5p, one of the important miRNA's hub from A, **B**. differentially upregulated mRNA (log2 fold change≥1) with differentially downregulated miRNA (log2 fold change≤-1) from ESC(N+2i) vs ESC(G), **B'.** Zoom in view of mmu-mir-451a, one of the important miRNA's hub from B, **C.** differentially downregulated mRNA (log2 fold change≤-1) with differentially upregulated miRNA (log2 fold change≥1) from ESC(N) vs ESC(N+2i), **C'.** Zoom in view of mmu-mir-1a-3p, one of the important miRNA's hub from C, **D.** differentially upregulated mRNA (log2 fold change≥1) with differentially downregulated miRNA (log2 fold change≤-1) from ESC(N) vs ESC(N+2i), **D'.** Zoom in view of mmu-mir-3963, one of the important miRNA's hub from D.

mmu-mir-741-3p, mmu-mir-142a-3p, mmu-mir-298-5p, mmu-let-7b-5p, mmu-mir-425-5p, mmu-mir-136-5p, mmu-mir-129-5p, mmu-mir-3057-5p, mmu-mir-1981-5p, mmu-mir-301b-3p, mmu-mir-542-3p, mmu-mir-132-3p and mmu-mir-291a-3p which are targeting *Gng2, Itga9, Csf1, mTOR, IL-6, Itgb4, Nos3, Srgn, Col4a1, Col4a2, Gys1, Itga7, Egfr, F2r, Fn1, Gnb4, Gng12, Tcn, Igf1r, Rps6, Thbs1, Tlr4* and *Prkca.*

Interactome analysis of down-regulated mRNAs vs upregulated miRNAs during ESC (G) to ESC (N+2i) transition

Further to investigate important pathways and genes which are vital in 2i/LIF but not required in 2i/LIF/serum, interactome map was analysed by considering downregulating mRNAs and upregulating miRNAs during ESC (G) to ESC (N+2i) transition (Fig. A8. A-A'). This also provide information about miRNAs necessary during 2i/LIF/serum to maintain pluripotency state. Some developmentally important downregulated genes which are targeted by upregulated miRNAs are *Ntrk3, Kras, Kit, Pik3r3, Raf1, Gls, Acsl3, Acsl4, Scd1, Ptplad1, Cyp51, Dhcr7, Bcat1, Polr1b, Alg6, B3gnt1, Ak4, polr3k, Mat2a, Fut9,* etc. Among these above-mentioned genes, some transcription factors like *Whsc111, Suv420h1, Agpat6, Ppap2b,* etc. are also controlled by upregulated miRNAs during ESC (G) to ESC (N+2i) transition. Among them some major central miRNAs hubs are mmu-mir-15a-5p, mmu-mir-34b-5p, mmu-mir-466k, mmu-mir-466i-5p, mmu-mir-17-5p, mmu-mir-181a-5p, mmu-mir-26a-5p, mmu-mir-362-5p, mmu-mir-665-3p, mmu-mir-495-3p, mmu-mir-344d-3p, mmu-mir-466p-5p, mmu-mir-1187 and mmu-mir-290a-5p. These miRNAs are downregulating pathways like regulation of actin cytoskeleton, fatty acid metabolism, VEGF signaling pathway, etc.

Interactome analysis of upregulated mRNAs vs down regulated miRNAs during ESC (N+2i) to ESC (N) transition

Maintainance of ESC in serum follows different pathways to exhibit pluripotency. Here, comparison of 2i/LIF/serum and LIF/serum was conducted by considering upregulated mRNAs and downregulated miRNAs from ESC (N) vs ESC (N+2i) (Fig. A8. D-D'). This interactome map gave information about depleted miRNAs during ESC (N) vs ESC (N+2i) which are important for 2i/LIF/serum but now no longer required in LIF/serum. Also, it will provide information about the mRNAs which are necessary for LIF/serum but was suppressed in 2i/LIF/serum. Some developmentally important upregulated genes which are targeted by downregulated miRNAs are *Orc4, Chek2, Tgfb2, Cul1, CCnd1, Ccne2, Cdc27, Rb1, Skp2, Myc, Egfr, Mapkapk3, Fgf13, Cacna2d1, Srf, Soga2,* etc. Among these above-mentioned genes, some transcription factors like *Tead3, Id1, Id2, Neurod1, Ptplad1, Agpat6, Ept1, Ppap2b*, etc. are also controlled by downregulated miRNAs during ESC (N+2i) to ESC (N) transition. Among them some major miRNAs hubs in this network are mmu-mir-124-3p, mmu-mir-17-5p, mmu-mir-26a-5p, mmu-mir-181a-5p, mmu-mir-15a-5p, mmu-mir-324-3p, mmu-mir-1195, mmu-mir-466i-5p, mmu-mir-541-5p, mmu-mir-665-3p, mmu-mir-298-5p, mmu-mir-362-5p, mmu-mir-495-3p, mmu-mir-10a-5p. These miRNAs are regulating pathways like, cell cycle, Axon guidance, Hippo signalling pathway, regulation of actin cytoskeleton, FoxO signaling pathway, MAPK signaling pathway, etc. These pathways are upregulating from ESC (N+2i) to ESC (N). Due to absence of PD inhibitor, here we can observe upregulation of MAPK signaling. MAPK signaling Pathway is regulated by mmu-mir-205-5p, mmu-let-7c-5p, mmu-mir-3473a, mmu-mir-1195, mmu-mir-335-3p, mmu-mir-495-3p, mmu-mir-344d-3p, mmu-mir-541-5p, mmu-mir-466i-5p, mmu-mir-295-3p, mmu-mir-15a-5p, mmu-mir-17-5p, mmu-mir-433-3p, mmu-mir-26b-5p, mmu-mir-124-3p and mmu-mir-125a-1-3p which are targeting *Soga2, Fgf5, Tgfb2, B230120H23Rik, Srf, Cacna2d1, Fgf13, Egfr, Mapkapk3* and *Myc.*

Interactome analysis of down-regulated mRNAs vs upregulated miRNAs during ESC (N+2i) to ESC (N) transition

Similarly, here downregulated mRNA and upregulated miRNAs were considered from ESC (N+2i) to ESC (N) (Fig. A8. C-C'). This will provide information about miRNAs which are necessary in LIF/serum but no longer required in ESC (N+2i) state. Also, it will display the genes

targeted by these miRNAs that need to be downregulated in ESC (N) condition. Some developmentally important downregulated genes which are targeted by upregulated miRNAs are *mTOR, Nos3, Serpine1, Arnt, Enos2, Camk2b, Cdkn1a, Tlr4, Prkca, Stat1, Prkca, Thrb, Atp1b1, Ncoa2, Casp9, Ryr2, Gucy1a3,* etc. Among these above-mentioned genes, some transcription factors like *Mmp9, Sfpi1, Aff1, Bcl6, Runx2, Whsc1, Stat1, Thrb, Ncoa2, Nfatc2, Dkk-1,* etc. are also controlled by upregulated miRNAs during ESC (G) to ESC (N+2i) transition. Among them some vital miRNAs hubs in this interactome map are mmu-mir-129-5p, mmu-mir-425-5p, mmu-mir-466d-5p, mmu-mir-301b-3p, mmu-mir-297a-5p. Downregulated pathways from ESC (N+2i) to ESC (N) due to these miRNAs are HIF-1 signaling pathway, Axon guidance, Thyroid hormone signaling, Oxytocin signaling pathway, Wnt signaling pathway, etc. Here, Wnt pathway is getting downregulated due to removal of CH inhibitor and the miRNAs that downregulating Wnt signaling pathway are mmu-mir-136-5p, mmu-mir-291a-3p, mmu-mir-466m-3p, mmu-mir-466d-3p, mmu-mir-466a-3p, mmu-mir-297b-3p, mmu-mir-466n-5p, mmu-mir-466h-5p, mmu-mir-669m-5p and mmu-mir-425-5p by targeting *Nfatc2, Btrc, Camk2b, Dkk-1, Csnk1e, Sfrp1* and *Prkca.*

LIF induced genes are highly upregulated in ESC (N+2i):

Leukemia Inhibitory Factor (LIF) is provided to ESC culture medium as extrinsic signal to maintain self-renewal. LIF bind to gp130/LIF-R cell surface receptor which phosphorylates JAK kinases and activate STAT3. Further STAT3, as a transcription factor upregulates many genes responsible for self-renewal of ESC (Yosidha et al., 1994). Over expression of STAT3 is sufficient to drive self-renewal in ESCs (Niwa et al., 2007). LIF also activates PI3K/Akt signaling pathway and MEK/ERK signaling pathway (Hackett et al.,2014). Phosphorylated STAT3 is translocated into nucleus and modulates expression of *Klf4, Gbx2, c-Myc* and *Tfcp2l1* (Ying et al., 2013). Though LIF is added to all three state but, *Stat3* expression is observed more in ESC (G) and ESC (N+2i) as compared to ESC (N). Expression of *Klf4* and *Tfcp2l1* are also more in ESC (G) and ESC (N+2i) (Fig. A10. B). But, expression of *Gbx2* and *c-Myc* is upregulated in case of ESC (N). Overexpression of *Tcfp2l1* can reprogram EpiSCs to ESC (N) state (Grabole et al., 2013).

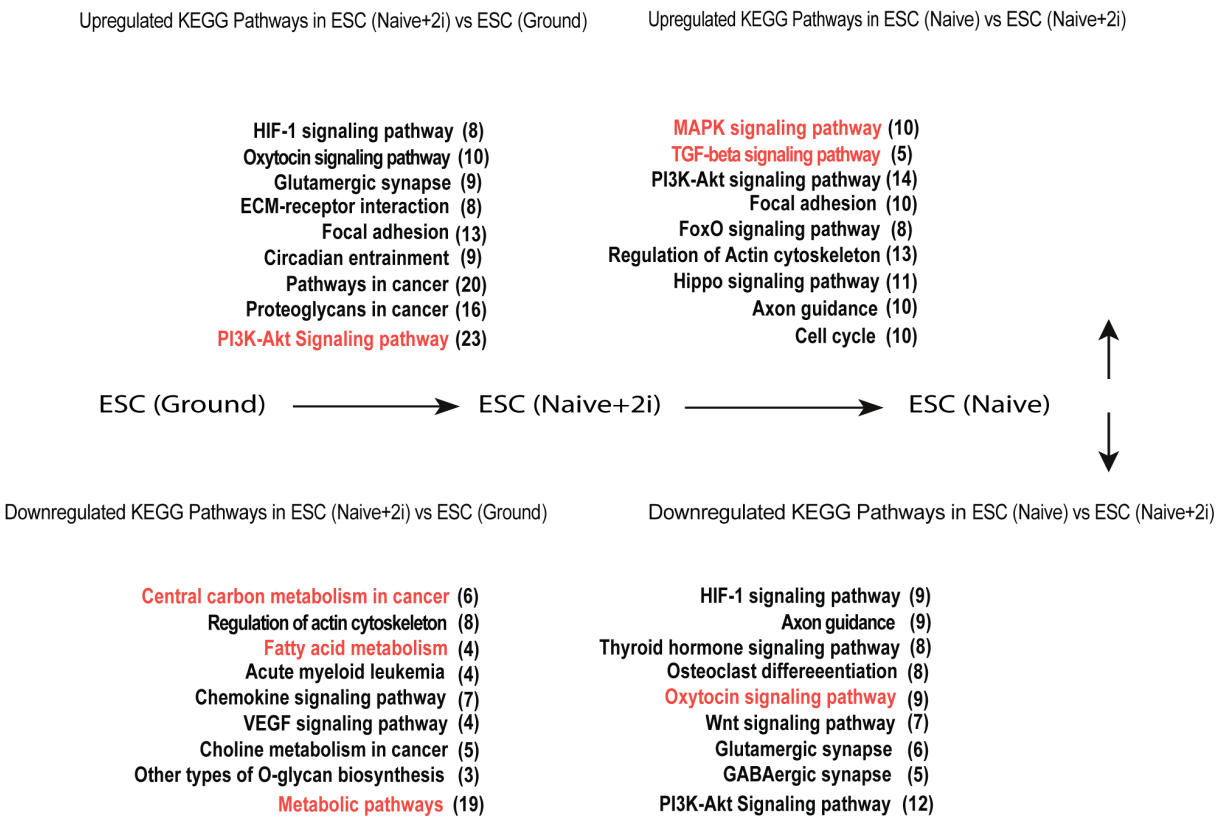


Fig. A9. KEGG Pathways prediction during neuron differentiation of mouse Embryonic stem cell: Flow chart representing KEGG pathways in between stages by considering differentially regulated (log2 fold change values) miRNAs and its filtered targets from differentially regulated (log2 fold change values) mRNAs. Full list of KEGG pathways are available in Table 20-23.

SALL4 stabilizes OCT4 expression by repressing trophectoderm genes (Zhang et al., 2006). *Sall4* expression is comparatively low in ESC (G) than ESC (N+2i) and ESC (N). *Tbx3* and *Tcl1,* pluripotency gene markers, are also getting upregulated by addition of LIF. *Tbx3* is showing high expression in ESC (G) and ESC (N+2i) but its expression is significantly much higher in ESC (N+2i). *Tcl1* expression is significant high in both ESC (G) and ESC (N+2i).

2i induced genes are showing elevated expression in ESC (G) and ESC (N+2i):

FGF/ERK pathway shifts Naive status of ESC to Primed which is susceptible to lineage specific differentiation. KLF2, a pluripotent gene marker, is phosphorylated by ERK2 and cause

proteosomal degradation. *Klf2* expression is elevated in case of ESC (N+2i) (Fig. A10. E). Also, ERK1/2 directly promotes many poised genes that are required for differentiation. *Fgf4*$^{-/-}$ and *Erk2*$^{-/-}$ ESCs are difficult to proceed for differentiation into neural and meso-endoderm lineage (Kunath et al., 2007). Therefore, use of MEK inhibitor PD03 which blocks phosphorylation of ERK1/2, promotes ESC self-renewal (Ying et al., 2008). When the expression of *Mapk3* and *Mapk1* were analysed, it showed downregulation in ESC (G) and ESC (N+2i) as compared to ESC (N) which is without MEK inhibitor inhibitor (Fig. A10. E).

However, only MEK inhibitor PD03 is not sufficient to support ESC self-renewal in absence of LIF. Along with PD03, there is an another GSK3 inhibitor called CHIRON enhances self-renewal and ESC propagation (Ogawa et al., 2006). Both *Gsk3a*$^{-/-}$ and *Gsk3b*$^{-/-}$ ESCs can able to propagate self-renewal in ESC without LIF and serum (Ying et al., 2008). CHIRON stabilizes ß-catenin inside cell which effectively stimulates canonical Wnt signaling (Berge et al., 2011). ß-catenin translocated into nucleus and enhances the expression of Oct4 and Tcf1 (Kelly et al., 2011) whose expression are much higher in ESC (G) but in ESC (N+2i) condition both of their expression are comparatively lower. TCF3 is a transcription factor that represses OCT4/SOX2 by directly binding to it. But TCF3 mediated suppression is disturbed by binding of ß-catenin which stabilizes ESCs (Faunes et al., 2013). Thus, *Tcf3*$^{-/-}$ ESCs was reported with enhanced self-renewal (Guo et al., 2011). Another gene, *Esrrb* which is important to maintain self-renewal, also has capacity to substitute CHIRON. *Esrrb*$^{-/-}$ ESCs did not response to 2i condition (Martello et al., 2012). ESC (N+2i) condition in our data showed elevated expression of *Esrrb* (Fig. A10. E). *Esrrb* expression is extremely vital for self-renewal in 2i condition but not necessarily in presence of serum.

2i/LIF ESC populations exhibit *Rex1/Zfp42* positive cells whose expression gradually diminishes towards Primed EpiSCs state. Though *Rex1* expression is detected in LIF/serum condition but, Rex-negetive Naive population are developmentally more matured than *Rex1*-positive Ground ESCs (Marks et al., 2012). ESC population with more *Rex1*-positive cells are comparatively more hypomethylated (Singer et al., 2014). *Rex1* expression is high in ESC (G) and ESC (N+2i) than ESC (N) (Fig. A10. F). Even *Rex1* expression is much higher in ESC (N+2i) in our datasets. 2i condition also enhances the transcription of *Prdm14*. *Prdm14* promotes homogeneous *Rex1*-positive population. *Prdm14* also represses FGF signaling and DNA methylation (Grabole et al., 2013). The expression of Prdm14 is comparatively high in ESC (G) and ESC (N+2i). *Prdm14* represses de novo methylases Dnmt3a and Dnmt3b which recruit 5mC in genome whose expressions are elevated in ESC (N). *Prdm14*$^{-/-}$ ESCs did not show hypomethylation in 2i/LIF

(Yamaji et al., 2013). Activity of TET enzyme also directly regulated by PRDM14 (Okashita et al., 2014). Some genomic region of ESC (G) state contains oxidized state of 5mC to 5-hydroxymethicytosinne (5hmC) to stabilize hypomethylation state which is mediated by TET enzymes and even erasure of 5mC is compromised in the absence of *Tet1* and *Tet2* (Ficz et al., 2013). Tet1 expression is elevated in ESC (G) but Tet2 is showing higher expression in ESC (N+2i) (Fig. A10. E). These results implies persistence of 2i activity in presence of serum.

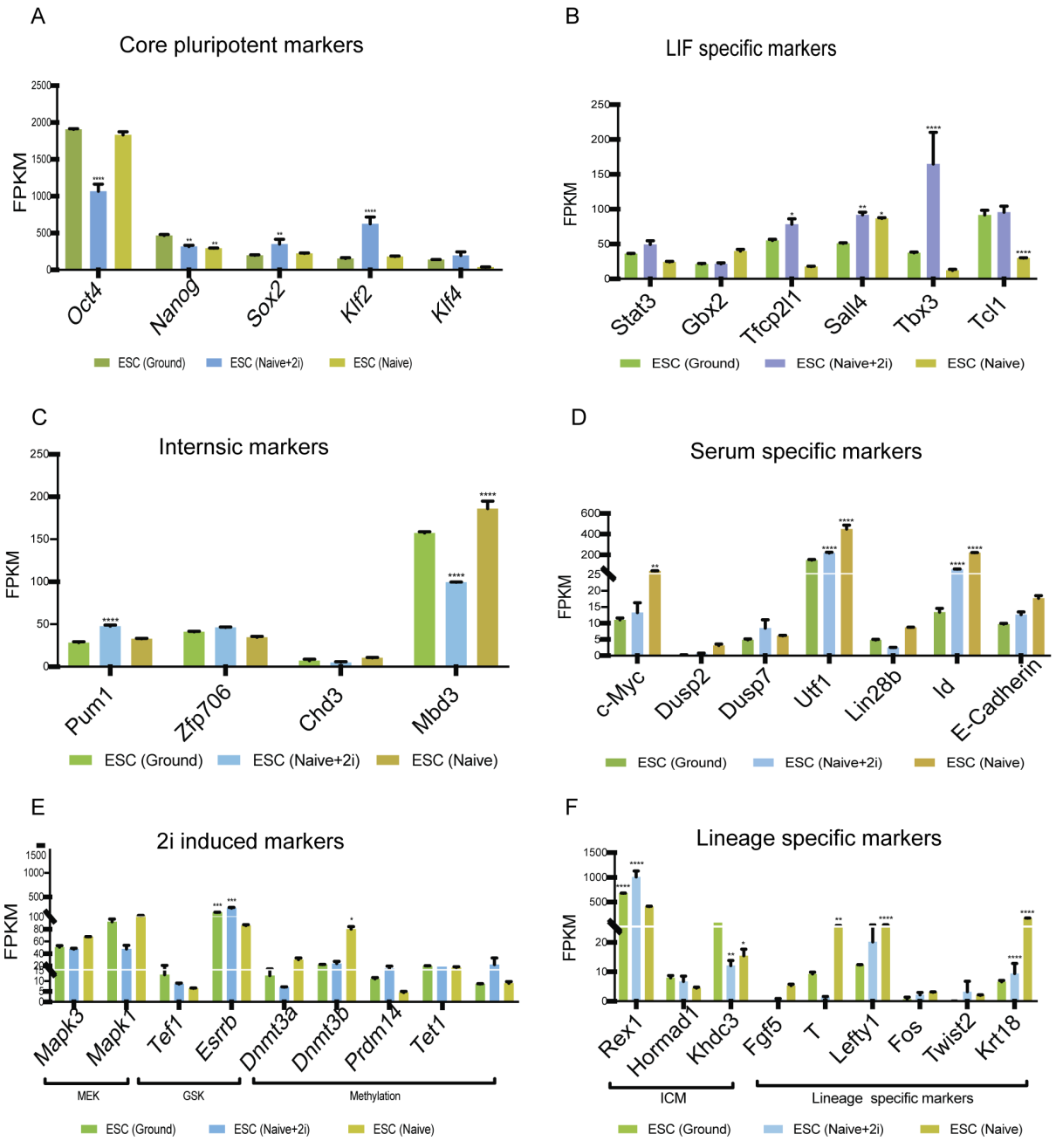


Fig. A10. Characterization of different developmental stages of ESC: Bar chart representing comparision of mRNAs FPKM values in between ESC (G), ESC (N+2i), ESC (N) by considering **A.** Core pluripotent markers, **B.** LIF specific markers, **C.** Internsic markers, **D.** Serum specific markers, **E.** 2i induced markers, **F.** Lineage specific markers.

Serum mediated genes are highly upregulated in ESC (N):

Specifically, presence of serum in ESC medium elevates different module to maintain plutipotency, one of them is *c-Myc*. *Myc* rapidly promotes transition through G1 cell cycle that suppresses differentiation. *Myc* also phosphorylates Dusp2 and Dusp7 that suppress FGF/ERK signaling. *Myc* and *Dusp2* transcription levels are high in ESC (N) condition but, Dusp7 expression is much higher in case of ESC (N+2i) (Fig. A10. D). Utf1, Lin28b and Id are specifically elevated in serum but showed downregulation in 2i/LIF medium which is even correlating in our datasets. *Id* is activated in presence of serum by BMP4 which targets downstream SMAD signaling pathway. Addition of BMP4 extrinsically or overexpression of *Id* can able to maintain self-renewal in absence of serum due to elevated expression of *E-Cadherin*. *E-Cadherin* prevents cell fate commitment (Malaguti et al., 2013). The expression of *E-Cadherin*/*Chd1* is gradually increases from ESC (G) to ESC (N+2i) and to ESC (N). Absence of LIF from culture medium differentiate ESC to non-neuronal lineage and absence of BMP4 redirects differentiation towards neuroectoderm derivatives (Ying et al., 2013). Therefore, presence of both LIF and BMP4 in ESC medium is necessary to maintain self-renewal (Hackett et al., 2014) and that's how ESC(N) can stabilize the pluripotency state.

Intrinsic differentiation factors increase in case of ESC (N+2i) and Naive:

The core transcription factors *Oct4, Sox2* and *Nanog* are essential to obtain Naive pluripotency but maintenance depends upon many other crucial factors (Silva et al., 2009). Together *Oct4*/*Sox2* and *Nanog* separately play distinct functional role to regulate pluripotency gene markers. These are cis-regulatory elements and can directly recruit coactivators and transcription factors onto promoters of specified gene targets. *Oct4* and *Sox2* are uniformly express irrespective of difference in ESC culture medium. The role of *Sox2* may be to activate Oct4 (Avilion et al., 2003). Overexpression of both *Oct4* and *Sox2* could promote lineage specification but, in absence of them

differentiation redirected towards tophectoderm lineage. Indeed, balance of *Oct4* is crucial to establish homogeneous population (Karwacki et al., 2013). The expression of *Oct4* was observed more in ESC (G) and ESC (N) compared to ESC (N+2i) but Sox2 showed more expression in ESC (N+2i) (Fig. A10. C).

Similarly, *Nanog* expression is high in ESC (G) and gradually reduces to ESC (N+2i) and ESC (N). ESC (G) pluripotency was confirmed with minimal expression of *Tsix* in real-time PCR, where X chromosome activation was yet to be initiated (Ghimire et al. 2018). Elevated expression of *Klf4* and *Nanog* (Fig 10. A) were marked distinctly in ESC (G) state compared to ESC (N) (Dvash et al. 2006). Also early pluripotency marker like *Hormad1* and *Khdc3* are clearly upregulated in ESC (G) condition (Guo et al. 2016). There are many intrinsic factors that promote to exit self-renewal state in ESC. Transcriptional regulators like PUM1 and ZFP706 are such factors. *Pum1* expression was observed to be more in ESC (N+2i) and *Zfp706* is high both in ESC (G) and ESC (N+2i). NURD corepressor complex is also encourage differentiation. Component of NURD like, *Chd3* and *Mbd3* are showing high expression in ESC (N) condition. ESC derived autocrine factor is FGF4 which instructs differentiation signal through MEK/ERK signaling. Surprisingly, *Fgf4* is activated by OCT4 and SOX2 (Kunath et al., 2007). *Fgf4* is expressing more in ESC (N+2i) in our datasets. Lineage specific gene markers *Fgf5* and *Lefty1* expression are also elevated slightly in ESC (N+2i) compared to ESC (G) and express more in ESC (N). Another differentiation commitment gene, *T* is showing significant high expression in ESC (N) but much lower in case of ESC (N+2i) (Fig. A10. C). These lineage commitment genes are showing comparatively elevated expression in ESC (N).

2i induction modulates miRNAs expression that leads to change in pathways:

Major effects of 2i was observed in ESC (N) condition where 2i was removed and ESC (N) condition was maintained only with LIF/serum. As in interactome map we have seen removal of 2i in ESC (N) leads to downregulation of Wnt signaling and upregulation of MAPK signaling pathway due to absence of CHIRON and PD03 inhibitors respectively. Once CHIRON was removed from ESC medium Wnt signalling related genes such as, *Nfatc2, Camk2b, Dkk-1, Csnk1e, Sfrp1* and *Prkca* had shown down regulation from ESC (N+2i) to ESC (N). These genes are mainly targeted by mmu-mir-136-5p, mmu-mir-291a-3p, mmu-mir-466m-3p, mmu-mir-466d-3p, mmu-mir-466a-3p, mmu-mir-297b-3p, mmu-mir-466n-5p, mmu-mir-466h-5p, mmu-mir-669m-5p and mmu-mir-425-5p whose expressions are elevated in ESC (N) (Fig. A11). These miRNAs showed

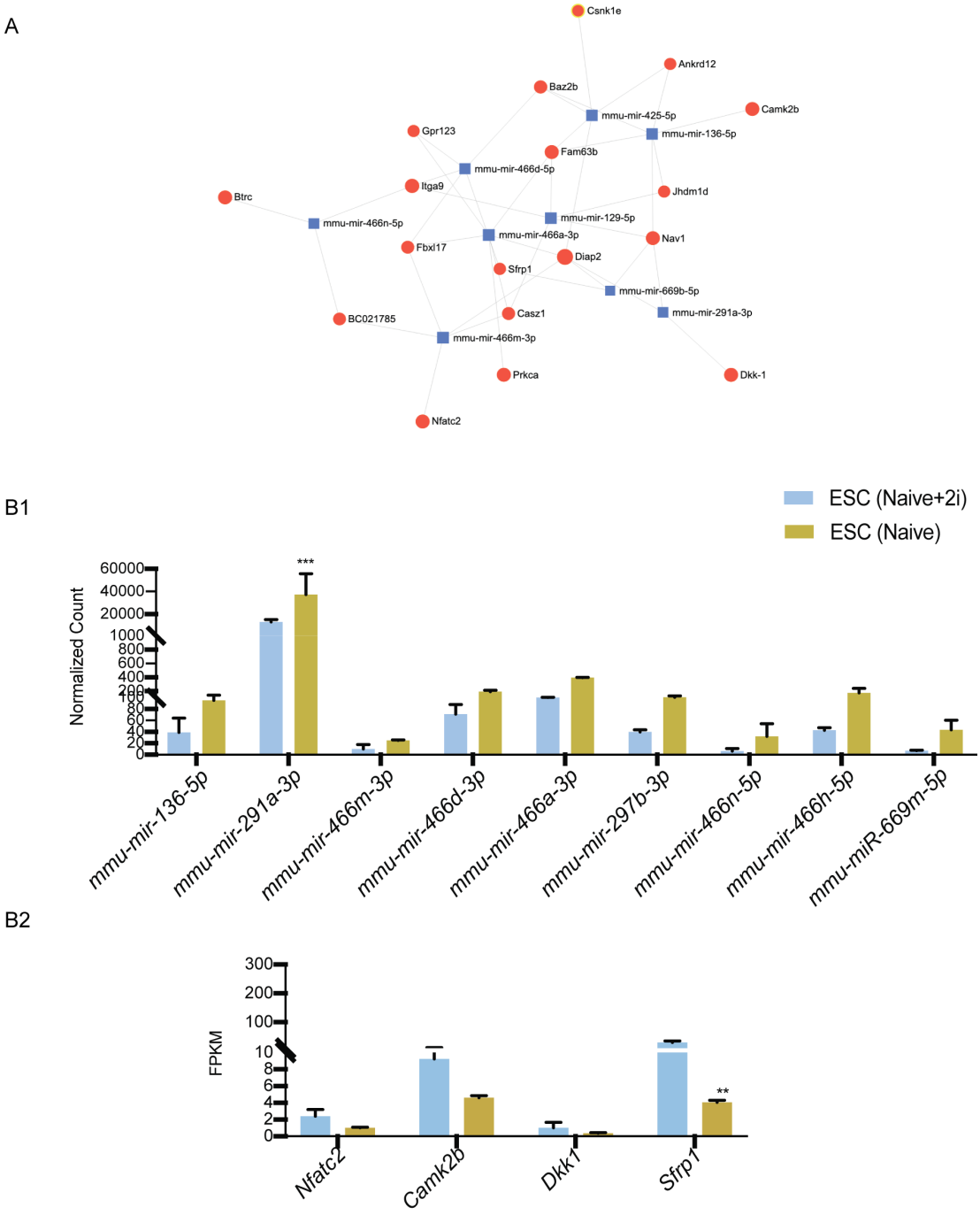


Fig. A11. Downregulation of Wnt signaling pathway upon removal of 2i in ESC(N) condition: A. miRNA-mRNA interactome map, **B1.** Normalized count of miRNAs and **B2.** FPKM value of mRNAs of Wnt signaling pathway by considering upregulated (log2 fold change≥1) miRNAs and downregulated (log2 fold change<-1) mRNAs from ESC (Naive) vs ESC (Naive+2i). Lists of mRNAs and miRNAs with FPKMs and Normalized counts are available in Table 24.

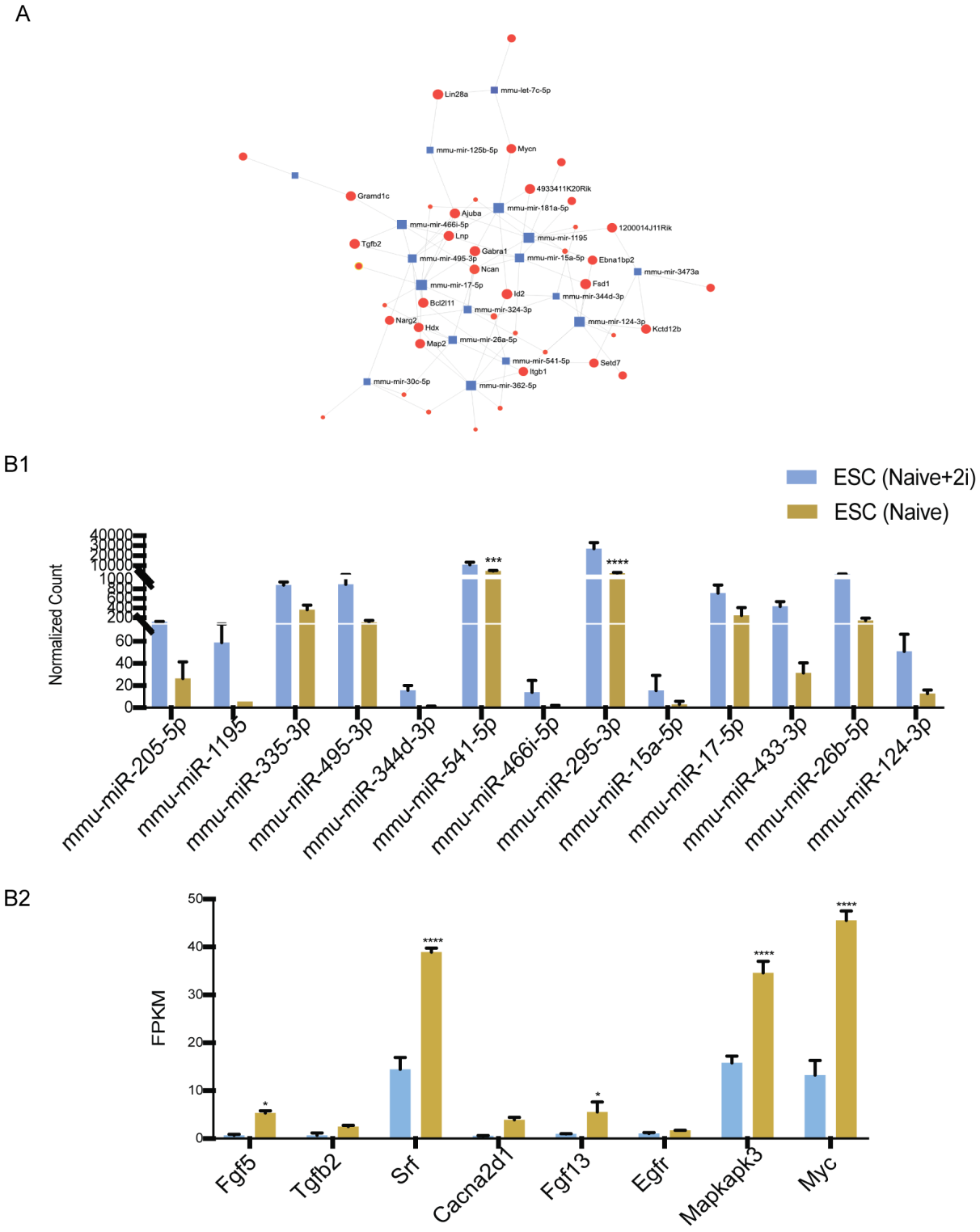


Fig. A12. Upregulation of MAPK signaling pathway upon removal of 2i in ESC(N) condition: A. miRNA-mRNA interactome map, **B1.** Normalized count of miRNAs, and **B2.** FPKM value of mRNAs of MAPK signaling pathway by considering downregulated (log2 fold change≤ß-1) miRNAs and upregulated (log2 fold change≥1) mRNAs from ESC (Naive) vs ESC (Naive+2i). Lists of mRNAs and miRNAs with FPKMs and Normalized counts are available in Table 25.

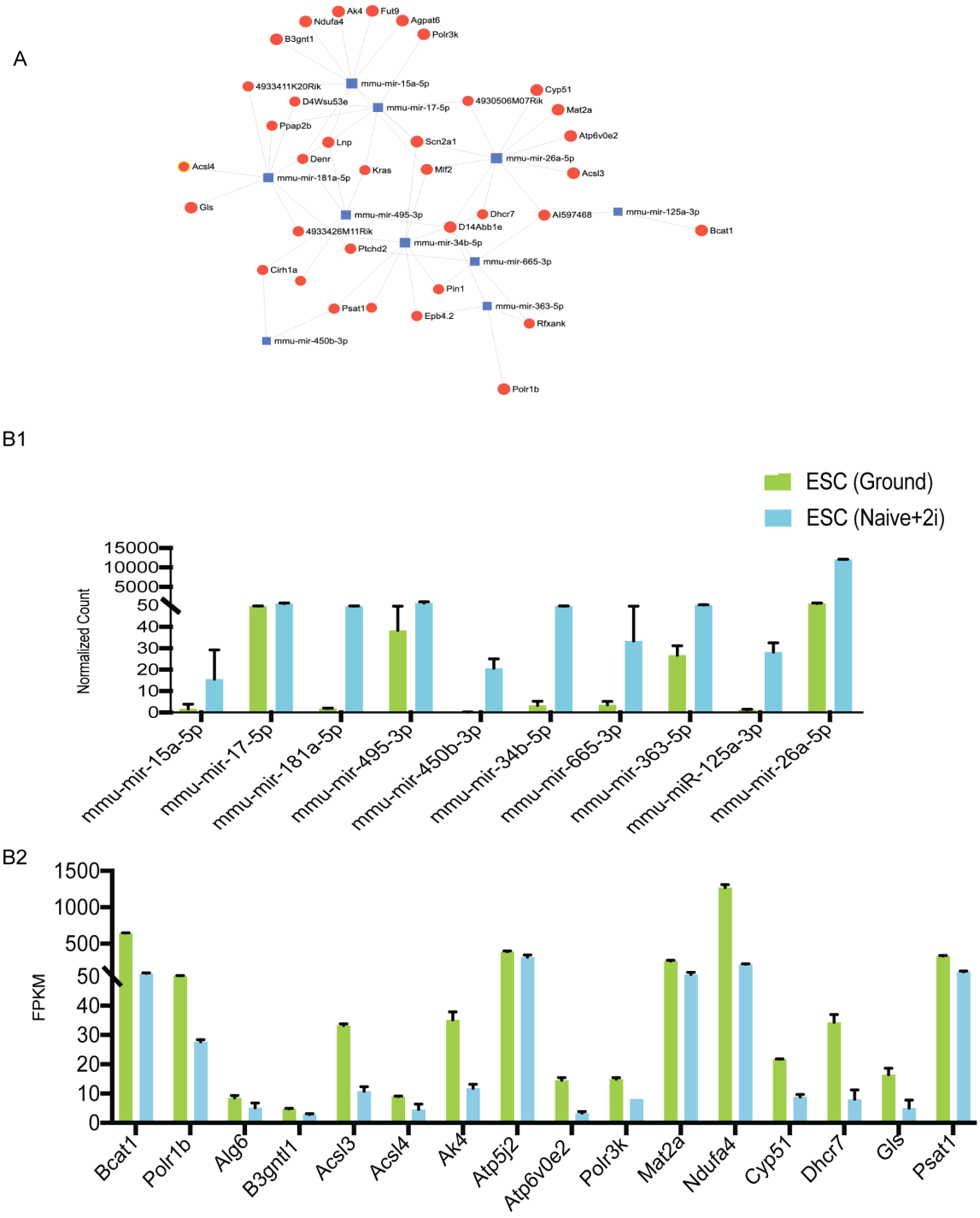


Fig. A13. Upregulation of metabolic pathway in ESC(G) condition: A. miRNA-mRNA interactome map, **B1.** Normalized count of miRNAs and **B2.** FPKM value of mRNAs of Metabolic pathway by considering downregulated (log2 fold change≤-1) miRNAs and upregulated (log2 fold change≥1) mRNAs from ESC (ESC (N+2i)) vs ESC (G). Lists of mRNAs and miRNAs with FPKMs and Normalized counts are available in Table 26.

gradual upregulation from ESC (G) to ESC (N+2i) and then highly expressing in ESC (N). Also, removal of PD03 from culture medium cause upregulation of *Fgf5, Tgfb2, Srf, Cacna2d1, Fgf13, Egfr, Mapkapk3* and *Myc*. These genes are related to MAPK signaling that was suppressed in 2i/LIF condition. The expression of these genes is highly upregulated in Naive ESCs. Above mentioned genes were targeted by mmu-mir-205-5p, mmu-let-7c-5p, mmu-mir-3473a, mmu-mir-1195, mmu-mir-335-3p, mmu-mir-495-3p, mmu-mir-344d-3p, mmu-mir-541-5p, mmu-mir-466i-5p, mmu-mir-295-3p, mmu-mir-15a-5p, mmu-mir-17-5p, mmu-mir-433-3p, mmu-mir-26b-5p and mmu-mir-124-3p miRNAs whose expressions are downregulated upon removal of MAK inhibitor (Fig. A12). If we analyse the expression pattern of these miRNAs in all three state, it is significantly much higher in ESC (N+2i) and drastically reduced in ESC (N) state. Thus, with these observations we could conclude that effects of 2i did not modulate much in presence of serum.

Ground state of ESCs are metabolically more active than ESC (N+2i):

Ground state of ESCs showed higher expression of metabolic genes and down regulation of cell cycle regulated genes similar to early embryonic development (Zhou et al., 2012). suggesting that ESCs (G) were metabolically more active than ESC (N+2i) and ESC (N) state. Metabolic pathway specific genes were observed to be downregulated from ESC (G) to ESC (N+2i) and its related genes *Bcat1, Polr1b, Alg6, B3gntl1, Acsl3, Acsl4, Ak4, Atp5j2, Atp6v0e2, Polr3k, Mat2a, Ndufa4, Cyp51, Dhcr7, Gls and Psat1* were also getting downregulated in ESC (N+2i) condition (Fig. A13. B). The miRNAs that targets these metabolic genes were getting upregulated from ESC (G) to ESC (N+2i), are mmu-mir-15a-5p, mmu-mir-17-5p, mmu-mir-181a-5p, mmu-mir-495-3p, mmu-mir-450b-3p, mmu-mir-34b-5p, mmu-mir-665-3p, mmu-mir-363-5p, mmu-mir-125a-3p and mmu-mir-26a-5p (Fig. A13. C). Along with metabolic pathway, pathways like central carbon metabolism in cancer, fatty acid metabolism, Choline metabolism in cancer and other types of O-glycan biosynthesis were showing downregulation from Ground state of ESC to ESC (N+2i) (Fig. A9). This implies serum free culture of ESC in 2i medium is metabolically more active.

Objective II

MicroRNA dynamics during directed cellular differentiation in mammals.

Comprehensive mRNA transcriptome profiles during Neurogenesis:

Neurogenesis is difficult to recapitulate *in vitro*. Three distinct feeder-independent embryonic stem cell lines were chosen that are modified from parental cell line E14TG2a for in vitro investigation. To capture the status quo of miRNAs at three different stages of development i.e., naive state (neuroectodermal stem cells), neural progenitors, and cortical neurons, thus adds an advantage over the classical method of procuring heterogeneous cell population obtained from the mouse brain dissection and cell sorting by FACS. For Naive state, Oct4-GiP E14TG2a cell line with GFP tagged to Oct4 (Fig. B1. A-A'): for neuronal progenitor state, 46C cell line with GFP tagged to Sox1 (Fig. B1. B-B'): and for cortical neurons, TK-23 cell line with GFP tagged to Tau (Fig. B1. C-C') were selected respectively. The rationale for adopting these cell lines is achieving the stage-specific expression of specific markers, which will aid in monitoring the efficiency of neuronal differentiation from stem cells. Sox1 is one of the earliest markers for neuroepithelial lineage and Tau for neurons (Gaspard et al. 2009). The cell identity was tracked for neural progenitors and cortical neurons by fluorescence of cells expressing GFP in 46C and TK23 cell lines.

The whole process of differentiation was confirmed with immunofluorescence assays using PAX6, TUJ1, and ßTUBB3 antibodies. PAX6 is a marker for progenitors and early neurons in NPCs; intense signals were detected for PAX6 on day 14 (Fig. B2. A), which gradually subsided in cortical neurons on day 21 of differentiation. On the other hand, TUJ1 is a marker of neurons with minimalistic expression patterns in NPCs and extensive signal output in CNs (Fig. B2. B); high *Vglut* expression in CN state reinforced our reliance on the differentiation system (Fig. B1. D). Astrocytes were rare in the CN population during immunostaining and real-time PCR experimental data. Negative expression of Dlx1 and Mag gene transcripts in q-PCR data suggested the presence of very few no. of cells that belong to neuro-ventral and oligodendrocyte lineage in the CN population. Neuronal progenitors and neurons were harvested on their respective days of differentiation, followed by mRNA and microRNA sequencing.

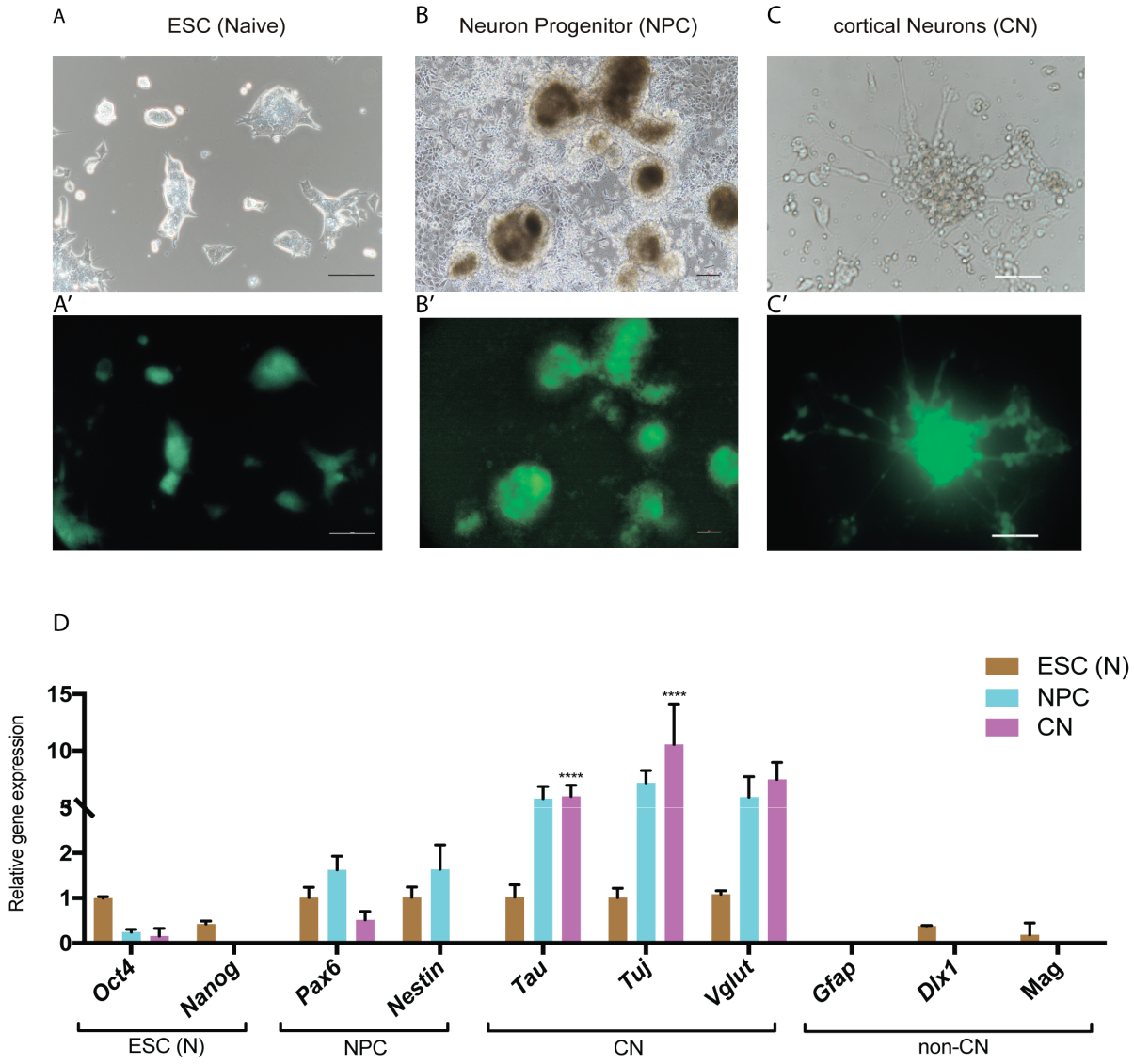


Fig. B1. Characterization of ESC to Cortical Neuron differentiation: 20X bright field and fluorescence images of different stages of cortical neuron differentiation; A-A'. ESC (Naive); For Naive state, Oct4-GiP E14TG2a cell line was used with GFP tagged to Oct4. ESC (Naive+2i) stage was cultured with 10% of serum medium with LIF, B-B'. Neuron Progenitor (NPC); for neuronal progenitor state, 46C cell line with GFP tagged to Sox1 was used. 50% confluent ESCs were grown with N2 medium for 2 days, then Cyclopamine was added till day 10. After day 10 Cyclopamine was removed and fresh N2 medium was added and continued to culture up to day 12. On day 12, cells were plated onto Poly-Laminin coated plate in presence of N2B27-A medium. Neuron Progenitor cells (NPC) were harvested on day 14., C-C'. Cortical Neuron (CN) state; TK-23 cell line was used with GFP tagged to Tau. Cortical Neuron were harvested on day 21 of neuron differentiation. D. Bar chart representing Real-time PCR analysis of stage specific markers of ESC (Naive), NPC, CN and non-CN genes. CNs are enriched with neuronal marker whereas no signal was captured from non-neuronal gene markers. Data was analysed by using $-2^{\Delta\Delta CT}$ method. *Gapdh* and ESC(N) were considered as gene control and sample control respectively. Bar graphs were generated by using Graph-pad PRISM software with one-way ANOVA significance test.

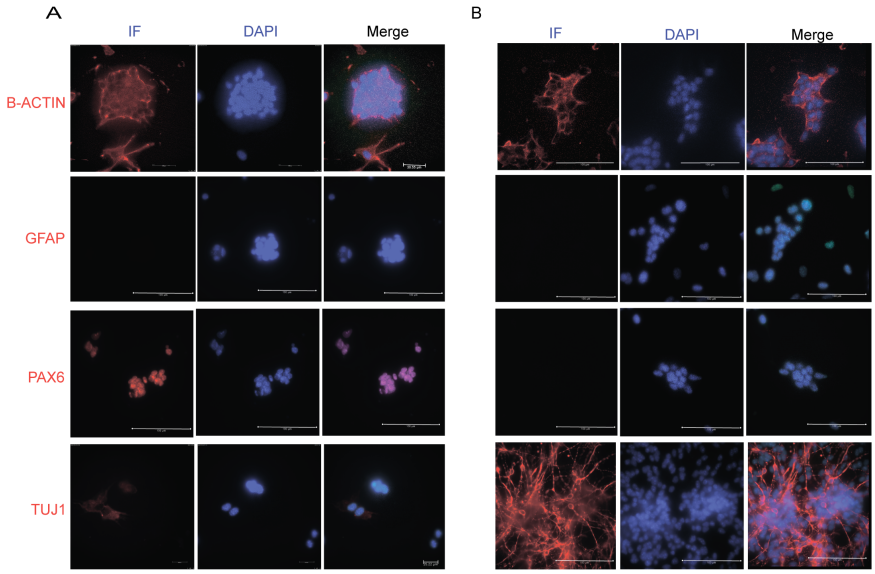


Fig. B2. Characterization of NPC and CN stage by Immunofluorescence: 63X fluorescence images were taken of A. Neuron Progenitors (NPC) and B. Cortical Neurons (CN) by using B-ACTIN, PAX6, GFAP and TUJ1 as primary antibodies. Here, PAX6 antibody is showing enriched signal in NPC where as TUJ1 is enriched in CN.

Differentiation of pluripotent cells into ectodermal dorsal neural lineage is tightly regulated and orchestrated by massive and dynamic modifications in the transcription, followed by the translation of genes to generate neurons. These changes at the transcriptome levels were apprehended for ESC (N), NPC, and CN states by using RNA-sequencing depth was set to 20 million reads for samples in biological duplicates. Qualitative analysis of obtained sequenced reads displayed ~50% GC content with 95% accuracy of Q-20 and Q-30 (Table 29). A positive correlation of above 0.90 was observed between the two biological duplicates with P-value <2.2e-16 (Table 27). A total of 14798 mRNA transcripts were of neural progenitor and 14060 of cortical neuron states, out of which 639 were unique to NPC and 941 to CN stages. The transition of ESC(N) to NPC state was associated with up-regulation of 2883 and downregulation of 2041 mRNA transcripts. The upregulation of 3270 mRNAs and downregulation of 3632 mRNA transcripts were found to be associated with NPCs differentiation to CN. Moreover, among the 11373 mRNAs, that were found to be commonly expressed, 90 mRNAs showed differential upregulation and followed by 40 mRNAs which showed differential downregulation in all the 3 states of differentiation (Fig. B4. A). Relevant data regarding transcriptomic expression of each

states have been tabulated. The mRNAs with unique expression patterns displayed as a list in Table 31, highly expressing mRNAs in Table 32, top 30 upregulated in Table 33, and 30 highly downregulated transcripts in Table 34.

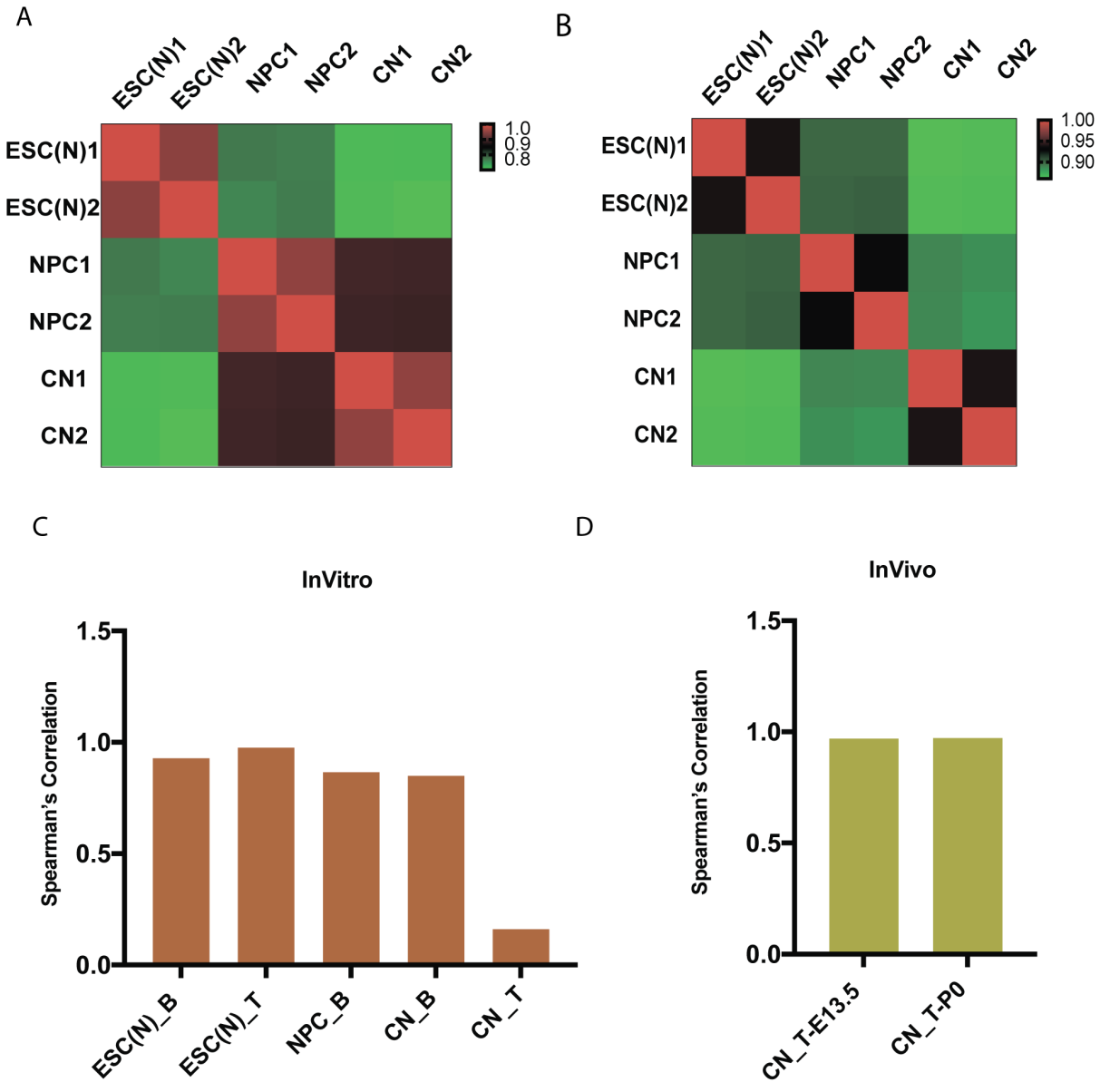


Fig. B3. Comparison of mRNA-seq with publically available datasets: Correlation between replicates and publicly available in vivo (GSE58523), in vitro (GSE96107) datasets: Heatmap representing Spearman's correlation map of A. miRNA-seq, B. mRNA-seq in between biological replicates and different samples of ESC (N), NPC and CN with Pvalue<2.2e-16. Bar graph representing more than 0.9 Spearman's correlation of NPC and CN samples with C. in vivo and D. in vitro datasets. Details of correlation values are available in Table 27-28.

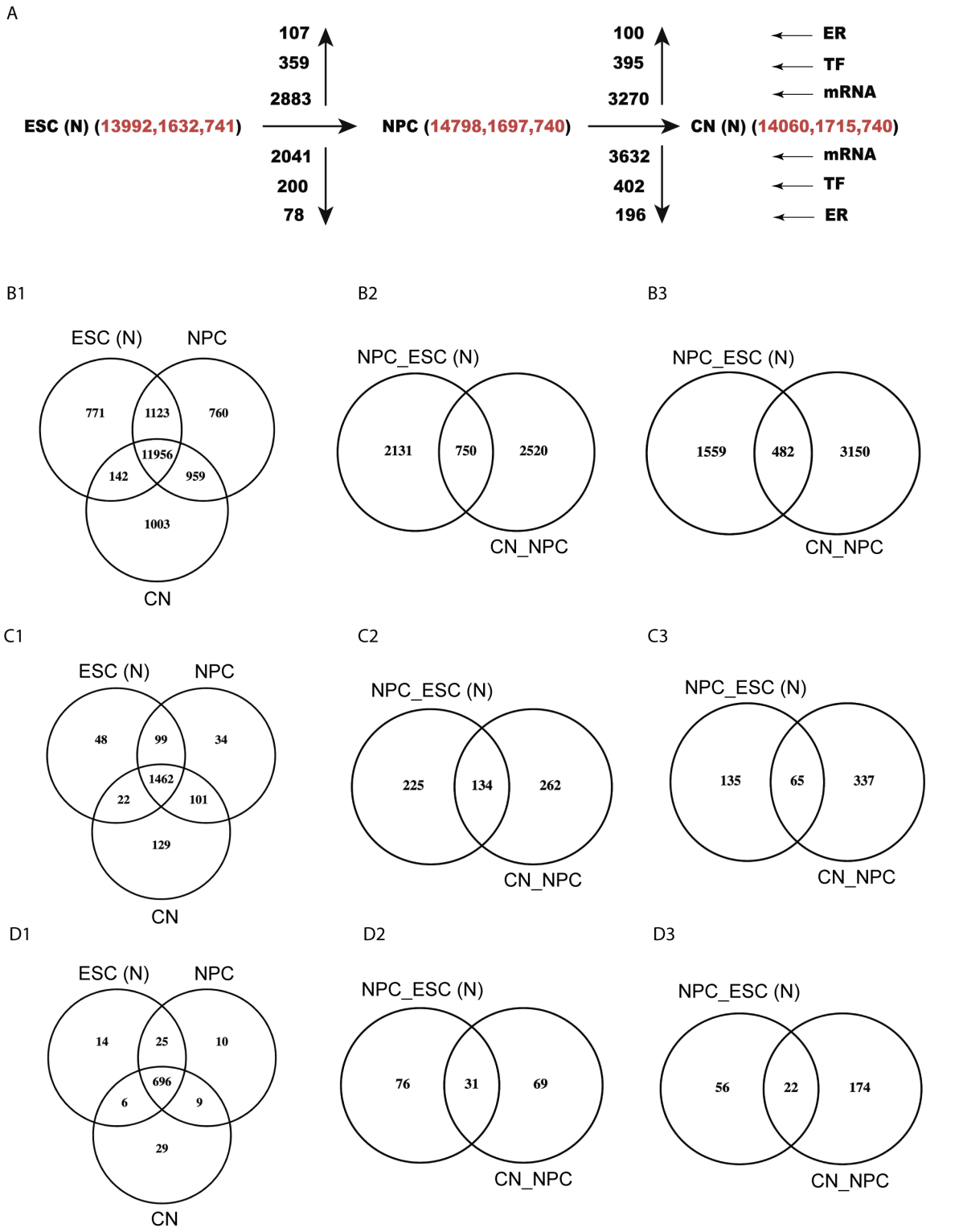


Fig. B4. Schematic representation of statistically analysed mRNAseq dataset: A. Flow chart representing expressed (FPKM≥1) and differentially regulated (log2 fold change) mRNAs, TFs and ERs in ESC (N), NPC, and CN. Venn diagram representing expressed, differentially upregulated (log2 fold change≥1) and differentially downregulated (log2 fold change≤1) B1-B3. mRNAs, C1-C3. TFs and D1-D3. ERs in between ESC (N), NPC, and CN.

To benchmark our mRNA-seq datasets, we compared our datasets with existing published datasets from Bonev *et al.* (Bonev et al., 2017) (designated as; B). The analysis was carried out considering that all three-cell lines were of the same E14TG2a ancestry. Correlation in both the datasets were observed i.e., for cortical neurons 0.84, for NPCs 0.86, and for ESC(N) 0.92 with P-value < 2.2e-16 was estimated (see Fig. B3. C). Comparison between acquired *in vitro* cortical neurons mRNA datasets with published *in vivo* (designated as; T-E13.5 and T-P0) for cortical neurons from Bouschet *et al.* datasets (Bouschet et al., 2017) was also performed, and a very high correlation of 0.97 was observed (Fig. B3. D). The correlation obtained for the mRNA datasets were found to be satisfactory enough to proceed further for microRNA sequencing; miRNA isolated from the same stage specific samples that were used for mRNA sequencing.

Upregulation of neuronal specific Transcription Factors (TFs) and Epigenetic Regulators (ERs):

The developmental regulatory networks involved in the differentiation processes are under the tight regulation of numerous transcription factors and epigenetic regulators. Neurogenesis also encompasses a massive variety of TFs with 1632, 1697, and 1715 TFs expressed in ESC(N), NPC, and CN states (Fig.B4. A). The sequencing output revealed 1423 TFs commonly expressed in all states, some stage-specific TFs numbering up to 12, 25, and 120 were found to be unique to ESC(N), NPC, and CN states (Fig. B4. B). TFs with unique expression patterns are tabulated in a listed format in Table 35, highly expressing TFs in Table 36, top 30 upregulated in Table 37, and 30 top downregulated TFs in Table 38. Differentially expressing TFs were analysed, and significant changes were observed during NPC differentiation. Amongst all 359 upregulated TFs, *Sox1* and *Gli3* displayed high upregulation during neuron transition. *Sox1* is the earliest marker for neuronal lineage (Elkouris et al., 2011), and *Gli3* is crucial for the self-renewal of neural progenitors (Wang et al., 2011). TFs like *Foxg1* (Kawauchi et al., 2009) and *Isl1* (Liang et al., 2011) are vital for neural development; *Pbx1* (Thiaville et al., 2012) and *Dlx1* (Petryniak et al., 2007) help in redirecting neural differentiation apart from oligodendrogliogenesis by upregulating specifically at NPC versus ESC(N) state. During ESC(N) to NPC transition, 200 TFs displayed downregulation. *Emx2* upregulation with other 395 highly expressed genes during cortical neuron development signified the beginning of corticogenesis (Chiara Cecchi, 2001). Co-upregulation in *Neurog1* and *Neurog2* established the redirection of neural lineage switch towards dorsal telencephalic progenitors, which give rise to glutamatergic neurons (Shekar et al., 2012). Traces of *NeuroD6* (Ulttenbogaard et al., 2010) and *NeuroD4* in CN state (Masserdotti et al., 2015) supported the survival and development

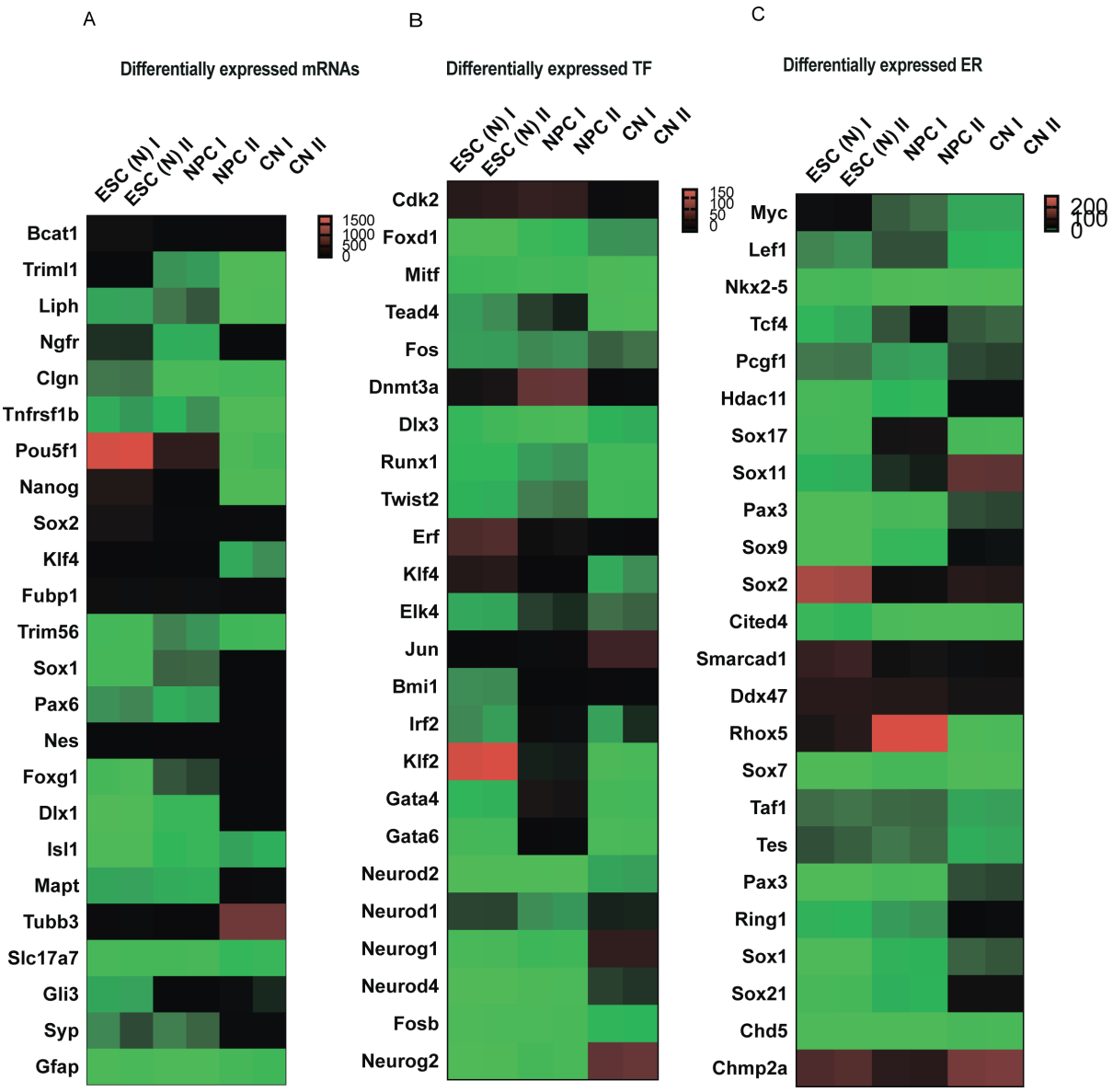


Fig. B5. Comparative analysis of uniquely expressed genes in Cortical Neuron differentiation from ESC: Heatmap representing Exclusively expressed A. mRNAs, B. TFs and C. ERs in ESC (N), NPC, and CN. Lists of genes uniquely expressed mRNAs, TFs, and ERs with FPKM values are available in Table 31, Table 35 and Table 39.

of neurons: high expression of *NeuroD1* signified the terminal differentiation of neurons (Boutin et al., 2012) and *NeuroD2* expression marked synaptic neurotransmission of cortical neurons (Chen et al., 2016). These findings suggest the presence of enriched cell populations in all stages of development.

Epigenetic regulators (ERs) are modulators, which provide regulatory instructions to genes without altering the primary nucleotide sequence but have a profound impact on the overall developmental process of neurons. A wide variety of ERs expression during Naive (741), NPC (740), and CN (740) were prominent (Fig. B4. A) from the generated dataset. ERs expressing in all stages of development were 688 in total, while very few unique stage-specific ERs came up with three for Naive, 9 for NPC, and 24 for CN state (Fig. B4. B). Differentially upregulated and downregulated ERs during NPC-Naive transition were 107 and 78 respectively. Similarly, 100 differentially upregulated and 196 downregulated ERs were observed during the CN-NPC transitions (Fig. B4. C-D). Only two ERs displayed upregulation and downregulation in all transit stages. During cortical neuron differentiation, 168 ERs showed downregulation. ERs with unique expression listed in Table 39, highly expressed in Table 40, top 30 upregulated in Table 41, and top 30 downregulated in Table 42.

MicroRNA dynamics during neurogenesis enriched neuronal specific pathways:

Like mRNAs, 20 million depth sequencing reads were obtained from both biological duplicates of ESC(N), NPC and CN. Quality check of sequenced reads were analysed which gave ~50% GC content and above 95% accuracy of Q20 and Q30 (Table 30). A good correlation of above 0.90 was observed in between two biological duplicates with Pvalue<2.2e-16 (Table 28). MicroRNAs showed huge transition in expression during cortical neuron differentiation. It has been observed that a greater number of miRNAs being expressed in NPC state (Fig. B6. A). Among them, 242 were marked as expressed in all the conditions and 22, 47 and 32 were unique to ESC(N), NPC and CN respectively (Fig. B6. B). During differentiation, 288 miRNAs have shown upregulation and 151 downregulation in ESC(N) cell transition to neuron progenitors. But from NPC to cortical neurons, 247miRNAs have shown downregulation and 155 miRNAs have shown upregulation (Fig. B6. B2-B3). The datasets are tabulated with, miRNAs uniquely expressed (Table 43), highly expressed (Table 44), top 30 upregulated (Table 45) and top 30 downregulated (Table 46) list of TFs lists were listed out.

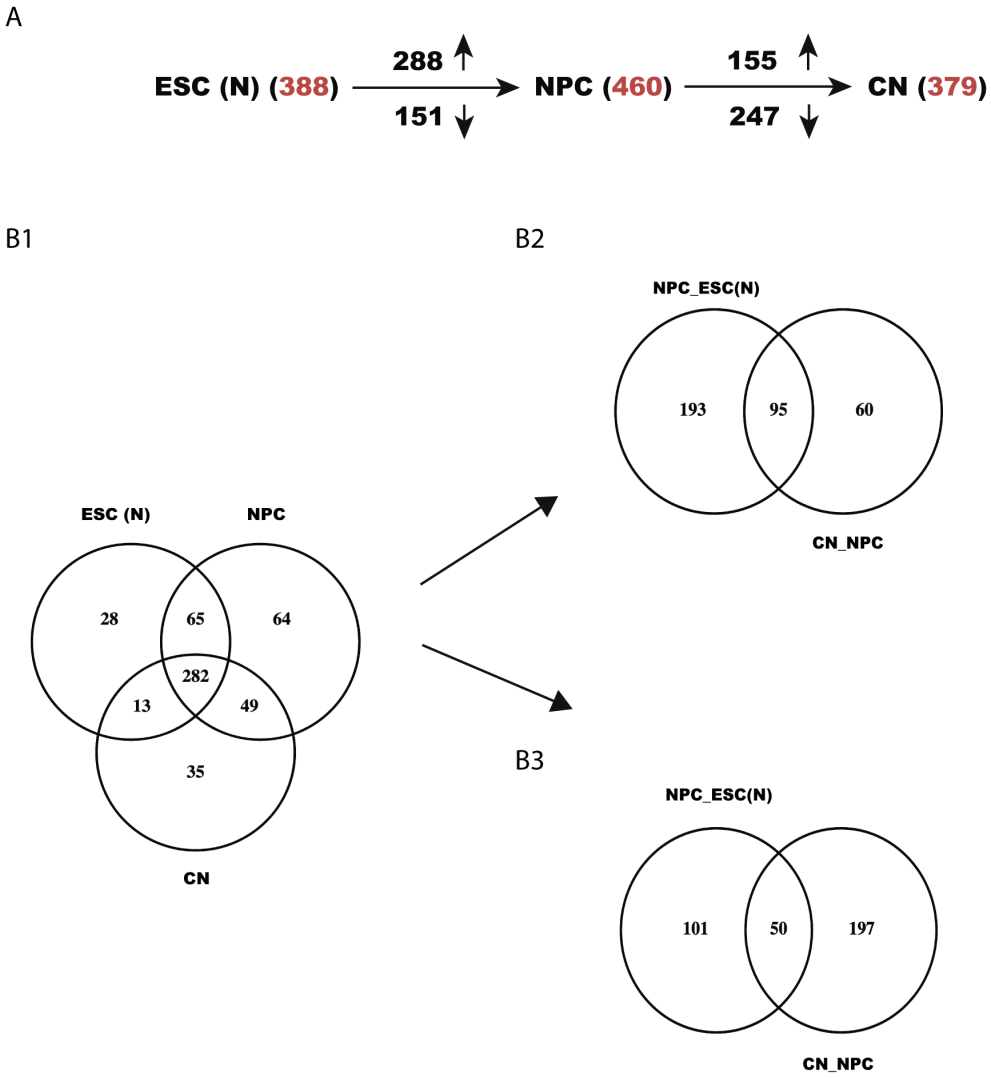


Fig. B6. Schematic representation of statistically analysed miRNAseq dataset: A. Flow chart representing expressed (NC≥10) and differentially regulated miRNAs in ESC (N), NPC, and CN. Venn diagram representing B1. expressed miRNAs (NC≥10), B2. differentially upregulated (log2 fold change≥1) miRNAs, B3. differentially downregulated (log2 fold change≤1) miRNAs in between ESC (N), NPC, and CN.

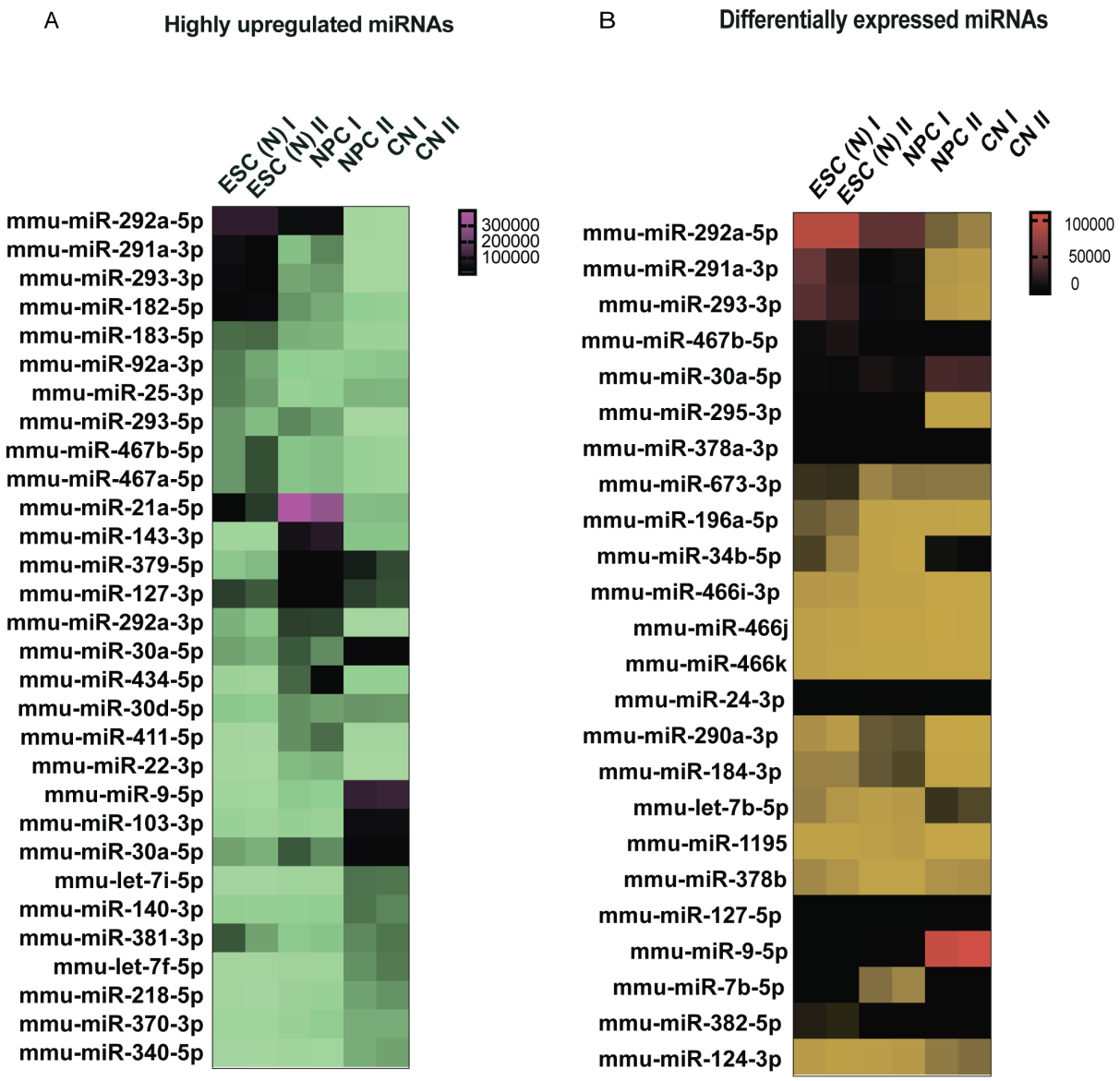


Fig. B7. Comparative analysis of uniquely and highly expressed miRNA genes in Cortical Neuron differentiation from ESC: Heatmap representing A. Uniquely expressed and B. Highly expressed miRNAs in ESC (N), NPC, and CN. Lists of uniquely and highly expressed miRNAs with Normalized counts are available in Table 43-44.

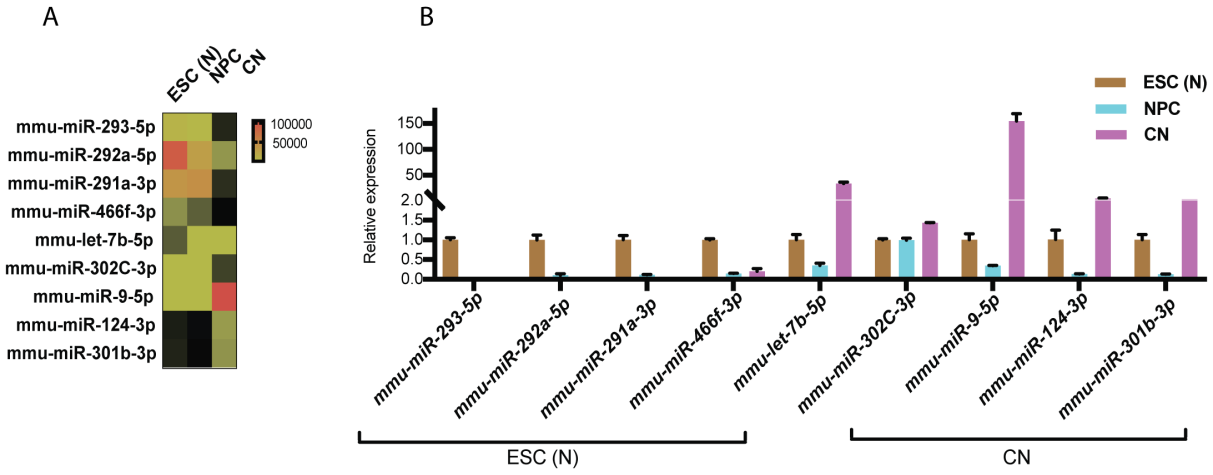


Fig. B8. Validation of miRNA-seq by using stage specific miRNA markers in real-time PCR: A. Heatmap representing expression (Normalized Count) status of stage specific miRNAs, B. Validation of few important miRNAs by using Real-time PCR in ESC (N), NPC and CN. Here, Real-time data showed similarities with miRNA-seq. Real-time data was analysed by using $-2^{\Delta\Delta CT}$ method. *Rnu6* and ESC(N) were considered as gene control and sample control respectively. Bar graphs were generated by using Graph-pad PRISM software with one-way ANOVA significance test.

Let-7 and miR-125b have lineage specific role in neuron development (Cho et al., 2019) that are highly upregulated in CN stage. Similarly, miR-122 expresses in NPC, supports neuronal commitment (Meza-Sosa et al., 2012). A well-studied miR-9 plays key role in neuronal development and is essential for neuronal differentiation (Rajman et al., 2017). MicroRNAs like miR-219 and miR-338, whose presence can redirect neuronal differentiation to oligodendrocytes (Meza-Sosa et al., 2012), are found to be lowly expressed in cortical neurons. Highly expressed microRNAs in astrocytes, such as miR-23 and miR-29 (Meza-Sosa et al., 2012), also showed low expression in CN condition

Integration analysis of miRNA datasets with mRNA showed enriched neuronal lineage specific pathways towards corticogenesis (Fig. B9). Gene ontology study of biological processes showed upregulation of synaptic transmission, transmission of nerve impulse, axonogenesis and neuron differentiation and development pathways, etc. (Fig. B10. A). Further cellular component analysis also showed involvement in formation of neuron projection, synapse, axon and dendrite (Fig. B10. C). Elevation of Notch signalling towards terminal differentiation has been observed with downregulation of mmu-miR-150-5p, mmu-miR-3095-5p, mmu-miR-324-3p and mmu-miR-1195 (Fig. B11. F). Similarly, upregulation of Nrg3 along with other ErbB pathway specific genes are mandatory for proper establishment of neuron circuitry and neuron transmission (Mei et al., 2014).

Cholinergic signal in cortical neuron influences release of acetylcholine in cortical neurons that establish neuronal connections (Bruel-Jungerman et al., 2011). Cholinergic signalling genes are controlled by mmu-miR-24-3p, mmu-miR-324-3p, mmu-miR-1195, mmu-miR-466f-3p and mmu-

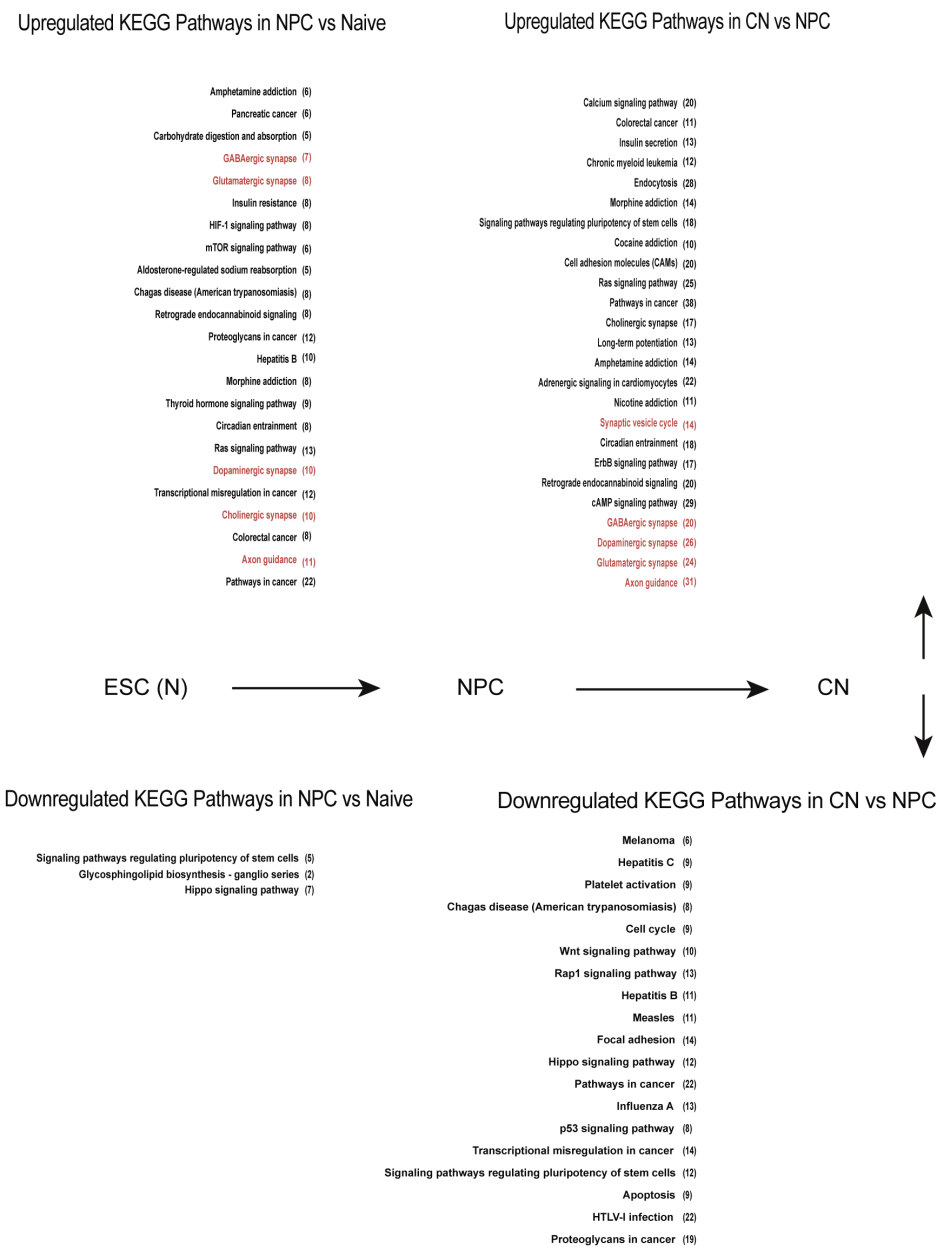


Fig. B9. Pathways prediction during ESC to Cortical neuron differentiation: Flow chart representing KEGG pathways in between stages by considering differentially regulated (log2 fold change values) miRNAs and its filtered targets from differentially regulated (log2 fold change values) mRNAs during CN differentiation. Number in bracket of right side of every pathway represents number of genes which are associated with that particular pathway during this phase of differentiation. Full list of KEGG pathways are available in Table 47-50.

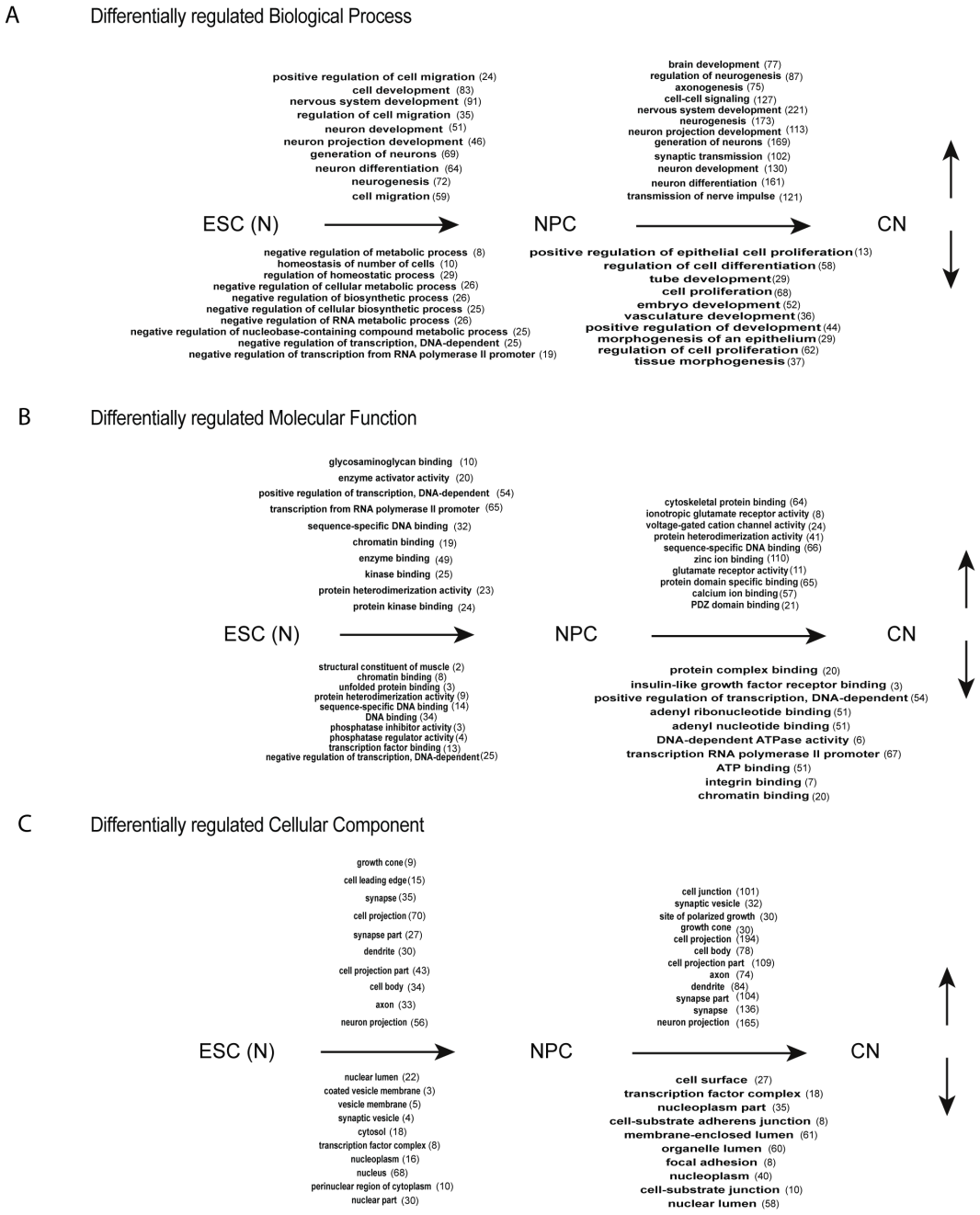


Fig. B10. Gene Ontology prediction during ESC to Cortical neuron differentiation: Flow chart representing Gene Ontology studies A. Biological Processes, b. Molecular Function, c. Cellular Component in between stages by considering differentially regulated (log2 fold change values) miRNAs and its filtered targets from differentially regulated (log2 fold change values) mRNAs during CN differentiation. Number in bracket of right side of every pathway represents number of genes which are associated with that particular pathway during this phase of differentiation. Full list of KEGG pathways are available in Table 51-62.

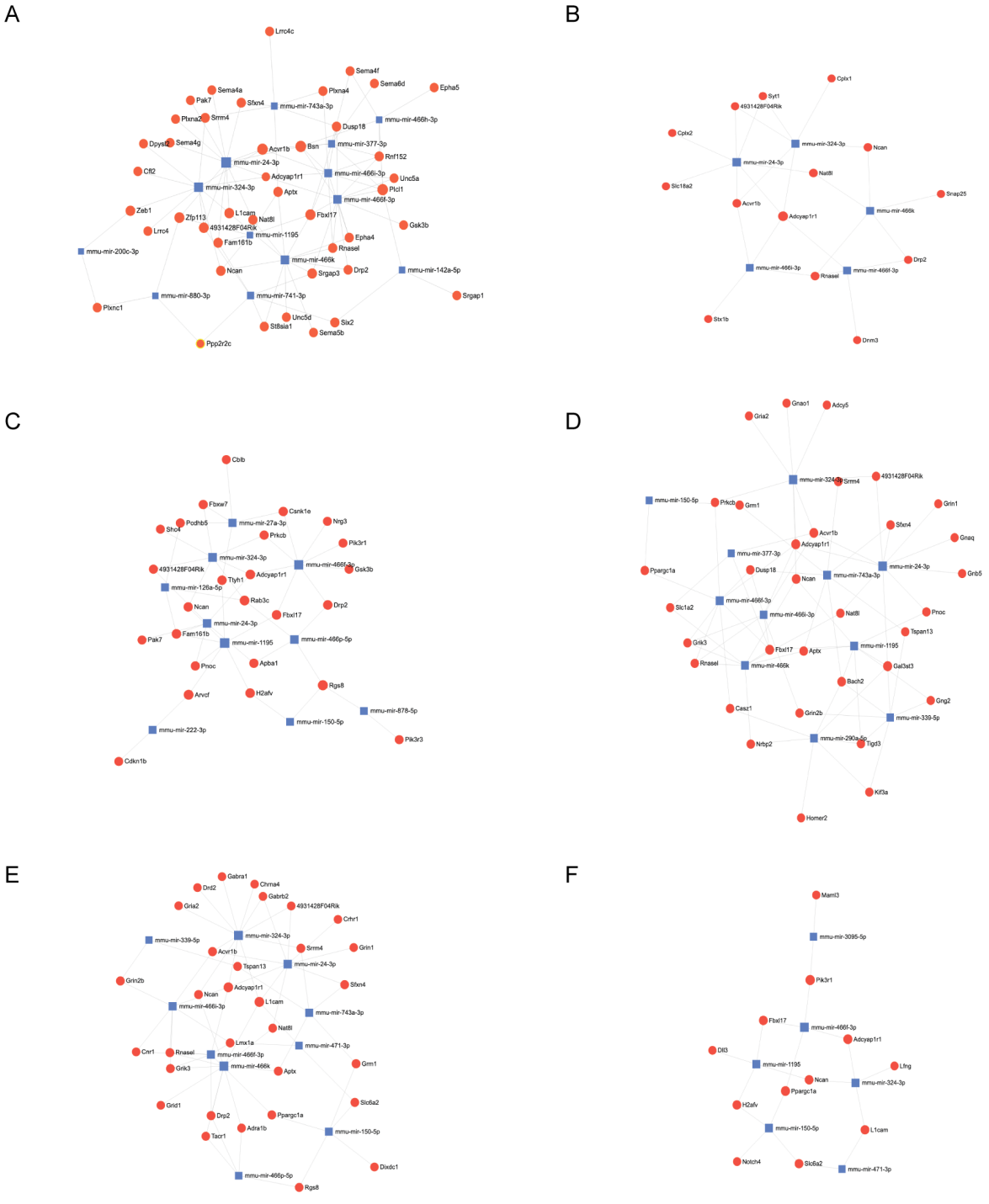


Fig.B11. MicroRNA-mRNA interactome map of highly enriched KEGG pathways from integration of upregulated miRNA and downregulated mRNAs from NPC to CN differentiation: A. Axon Guidance, B. Synaptic vesicle cycle, C. ErbB Signalling, D. Glutamatergic Synapse, E. Neuroactive-ligand-receptor binding and F. Notch Signalling.

miR-466i-3p (Fig. B11. C). Along with these miRNAs', downregulation of mmu-miR-466k, mmu-miR-466i-5p and miR-290 family promoted Glutamatergic synapse formation (Fig. B11. D). All of these above pathways with Synaptic vesicle cycle (Fig. B11. B) and Neuroactive ligand-receptor interaction (Fig. B11. E) are majorly controlled by mmu-miR-24-3p, mmu-miR-324-3p, mmu-miR-1195, mmu-miR-743 and mmu-miR-466 family. Downregulation of these miRNAs during cortical neuron development seems to be essential for proper neuronal function.

Pathways prediction with Qiagen Ingenuity pathway analysis:

Pathway analysis during neurogenesis through Qiagen ingenuity pathway analysis (IPA) provides upregulation of neuron specific pathways. During ESC(N) ESC to NPC differentiation pathways like Axonal Guidance Signalling, Wnt/ß-catenin Signalling, GP6 signalling pathway and Regulation of the Epithelial-Mesenchymal Transition pathway showed highly upregulated (Table T40). Similarly, during cortical neuron generation from NPC showed upregulation of pathways like Synaptogenesis Signalling Pathway, Axonal Guidance Signalling, Synaptic long term Depression, Wnt/ß-catenin Signalling, CREB Signalling in neurons, Neuropathic pain signalling in Dorsal Horn Neurons, Endocannabinoid Neuronal Synapse Pathway, Reelin Signalling in Neurons, GABA Receptor Signalling and Glutamate Receptor Signalling (Fig. B12) (Table 63-64).

Neural Progenitor cell specific miRNA-mRNA interactome analysis:

To dissect regulatory networks and related pathways, it became important to examine the relationships among differentially regulated and expressed mRNAs and miRNAs during the course of neurogenesis (Fig. B13. A). Differentially expressed datasets provide many useful insights regarding the role of miRNAs during corticogenesis. In our differentially expressed datasets, we looked into the mRNA and miRNAs that have significant P-value with –log2 fold change. There is also a possibility that with high cut-off we might overlook many lowly expressed or having insignificant differential changes during the course of differentiation. To overcome this issue, we considered interaction map of expressed mRNA and expressed miRNAs. Considering NPC state, there were many miRNAs that showed upregulation and act as major hubs with more than 100 expressed miRNAs such as, mmu-mir-340-5p, mmu-mir-9-5p, mmu-mir-362-3p, mmu-mir-329-3p, mmu-mir-124-3p, mmu-mir-181a-5p, mmu-mir-7b-5p, mmu-mir-149-5p, mmu-mir-26a-5p, mmu-mir-17-5p, mmu-mir-425-5p, mmu-mir-30e-5p, mmu-mir-24-3p, mmu-mir-324-3p, mmu-let-7b-5p, etc. This interaction map was found to be enriched with KEGG pathways like Hippo

Fig. B12. Pathways prediction during ESC to Cortical neuron differentiation by using Qiagen Ingenuity Pathway Analysis: Flow chart representing Qiagen-IPA pathways in between stages by considering differentially regulated (log2 fold change values) miRNAs and differentially regulated (log2 fold change values) mRNAs during CN differentiation. Full list of Qiagen Ingenuity pathways are available in Table 63-64.

signalling pathway (Total 91 expressed genes are regulated), Signalling pathway regulating pluripotency (81), Axon guidance (73), Neurotrophin signalling pathway (70), FoxO signalling pathway (72), ErbB signalling pathway (53), Wnt signalling pathway (74), AMPK signaling pathway (67), mTOR signalling pathway (38), Dopaminergic synapse (57), etc. Though the interaction map of differentially miRNAs provide neuron specific miRNA-mRNA interaction maps and specified neuronal developmental pathways. There are many expressed mRNAs in existing state those have minimal expression but have high functional significance. But due to marginal cut-off could be omitting those interactions. Reconsidering these drawbacks, we tried to emphasize here interactions of differentially upregulated and downregulated miRNAs with expressed mRNAs of previous state. Firstly, we proceeded with interaction map of differentially upregulated miRNAs between NPC to CN with expressed genes in Neuron progenitors to rule out the targets of differentially upregulated miRNAs which did not show significant differences in expression during NPC to CN transition (Fig. B13. B). This interaction network also provided major miRNAs' hub as we have seen in previous interaction map of upregulated miRNA and downregulated mRNAs during NPC to CN transition. But here it captured miRNAs' hub with a greater number of targeted genes. Like, mmu-mir-9-5p, mmu-mir-124-3p, mmu-mir-301b-3p are targeting total 458, 423 and 337 expressed mRNAs. Top regulated pathways came up in this interaction are Hippo signalling pathway (77), Thyroid hormone signalling (62), Neurotrophin signalling pathway (62), Axon guidance (61), Dopaminergic synapse (58), mTOR signalling pathway (33), Ras signalling pathway (83), etc.

Further we proceeded with interaction maps of differentially downregulated miRNAs between NPC to CN with expressed genes in Neuron progenitors to decipher the miRNAs whose presence in NPC is essential to control the expressed genes and have minimal role during CN differentiation (Fig. B13. C). The major miRNAs' hub in these interactions are mmu-mir-24-3p, mmu-mir-324-3p, mmu-mir-466i-3p, mmu-mir-1195, mmu-mir-466k, mmu-mir-466i-5p, mmu-mir-466f-3p, mmu-mir-339-5p, mmu-mir-377-3p, mmu-mir-290a-5p, mmu-mir-292a-5p, mmu-mir-466p-5p, etc. These miRNAs' hubs are involved in regulation of significant pathways like pathways in cancer (92), axon guidance (41), Hippo signalling pathway (42), Signalling pathway regulating pluripotency (39), FoxO signalling pathway (37), mTOR signalling pathway (22), PI3K-Akt signalling pathway (70), cell cycle (33), Ras signalling pathway (48), Wnt signalling pathway (35), ErbB signalling pathway (24), p53 signalling pathway (20) etc.

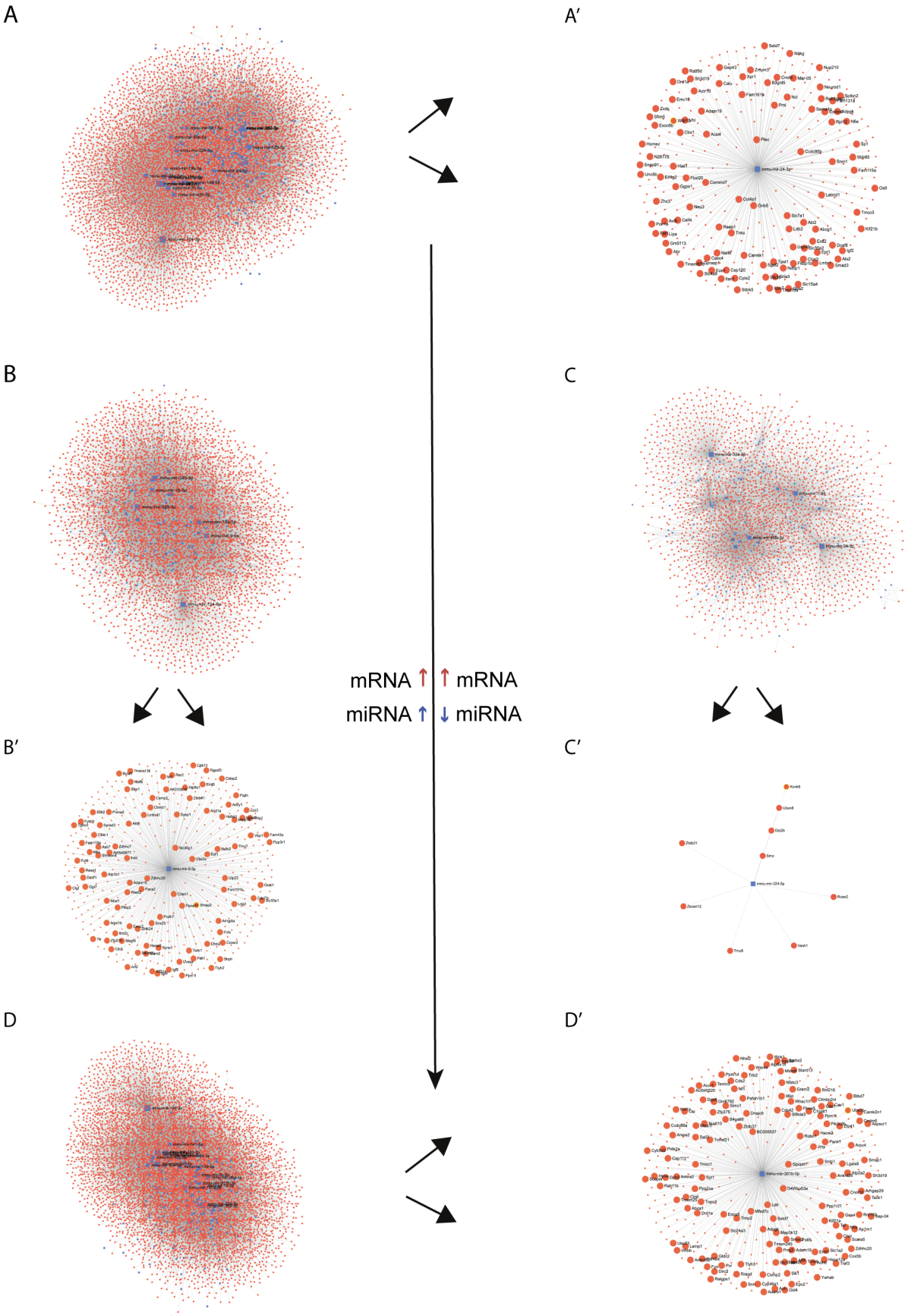

A
A′
B
C
mRNA ↑ ↑ mRNA
miRNA ↑ ↓ miRNA
B′
C′
D
D′

Fig. B13. miRNA-mRNA network map of miRNAs and its experimentally validated mRNAs targets in neuronal cells during neurogenesis: Computationally analysed interactome map of miRNA and mRNA as a whole and zoomed in view of one of the highly expressed miRNA. A. expressed mRNA (FPKM≥1) with expressed miRNA (NC≥10) of NPC, A'. Zoom in view of mmu-mir-24-3p, one of the important miRNA's hub from A, B. expressed mRNA (FPKM≥1) of NPC with differentially upregulated miRNA (log2 fold change≥1) from CN vs NPC, B'. Zoom in view of mmu-mir-9-5p, one of the important miRNA's hub from B, C. expressed mRNA (FPKM≥1) of NPC with differentially downregulated miRNA (log2 fold change≤-1) from CN vs NPC. C'. Zoom in view of mmu-mir-324-5p, one of the important miRNA's hub from C, D. expressed mRNA (FPKM≥1) with expressed miRNA (NC≥10) of CN, D'. Zoom in view of mmu-mir-301b-3p, one of the important miRNA's hub from D.

Cortical neurons specific miRNA-mRNA interactome analysis:

Similarly, interaction map of expressed mRNAs in CN and expressed miRNAs in CN captured many essential neuronal specific miRNAs' hubs. These interactions would clarify the list of expressed miRNAs whose presence might be essential to maintain CN state by controlling neuron specific pathway genes (Fig. B13. d). Cortical neuron miRNA-mRNA interactome map is enriched with major miRNAs' hubs such as, mmu-mir-340-5p, mmu-mir-9-5p, mmu-mir-362-3p, mmu-mir-329-3p, mmu-mir-7b-5p, mmu-mir-124-3p, mmu-mir-149-5p, mmu-mir-26a-5p, mmu-mir-17-5p, mmu-mir-425-5p, mmu-mir-30e-5p, mmu-mir-301b-3p, etc. These miRNAs are involved in regulation of pathways like Hippo signalling pathway (86), Axon guidance (74), Signalling pathway regulating pluripotency (78), MAPK signalling pathway (115), Neurotrophin signalling pathway (66), FoxO signalling pathway (68), ErbB signalling pathway (53), Wnt signalling pathway (69), AMPK signalling pathway (65), mTOR signalling pathway (38), Dopaminergic synapse (64), Adherens junction (39), Regulation of Actin cytoskeleton (85), Ras signalling pathway (89), Cell cycle (55), TGFß signalling pathway (40), Glutamatergic synapse (50), Hedgehog signalling pathway (26), Cholinergic synapse (47), Notch signalling pathway (23), Synaptic vessicle cycle (27), Tight junction (51), etc.

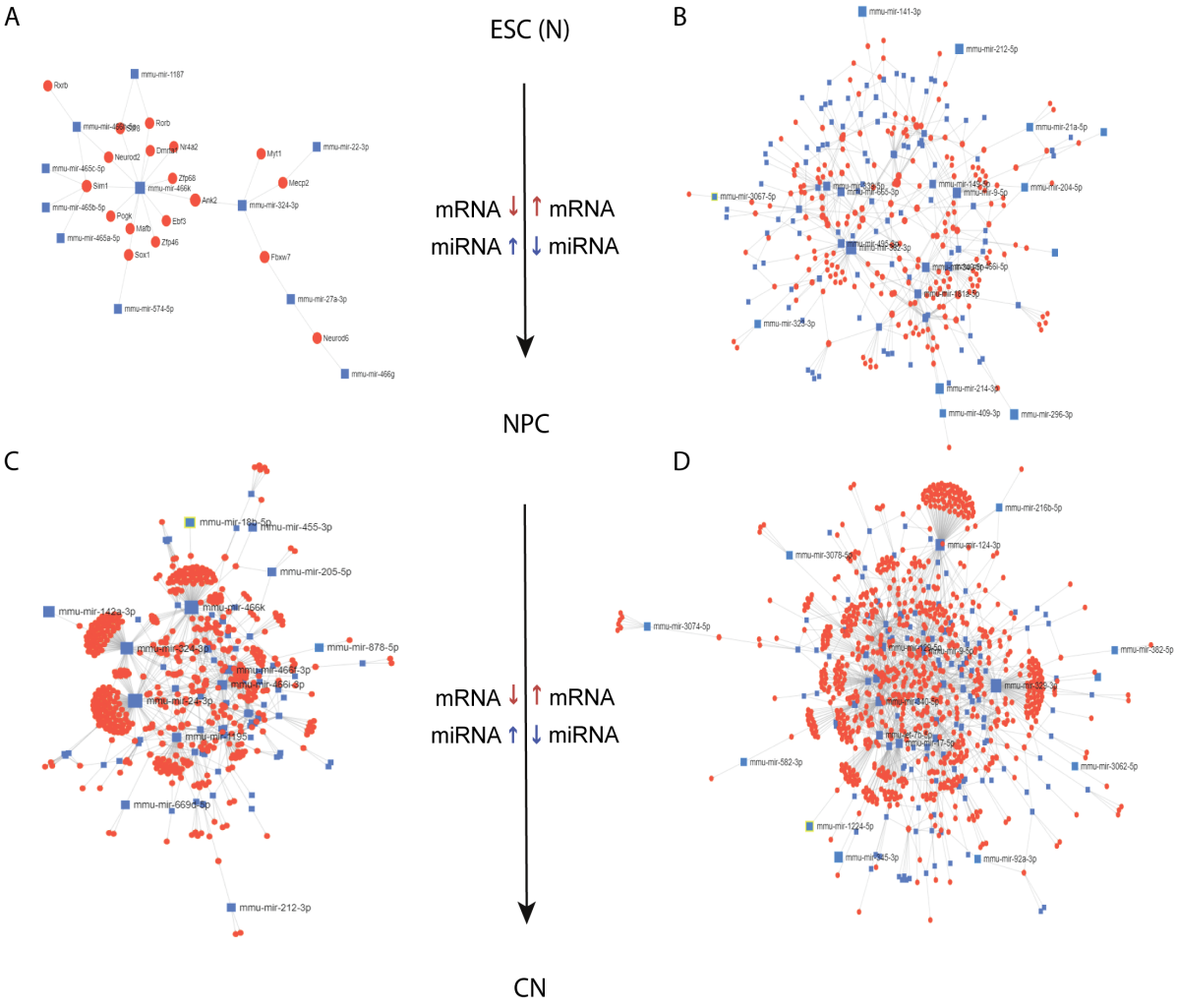


Fig. B14. miRNA-mRNA network map of differentially expressed miRNAs and its experimentally validated differentially expressed mRNAs in neuronal cells during neurogenesis: Computationally analysed interactome map of differentially expressed mRNA and differentially expressed miRNA as a whole. A. differentially downregulated mRNA (log2 fold change≤-1) with differentially upregulated miRNA (log2 fold change≥1) from NPC vs ESC(N), B. differentially upregulated mRNA (log2 fold change≥1) with differentially downregulated miRNA (log2 fold change≤-1) from NPC vs ESC(N), C. differentially downregulated mRNA (log2 fold change≤-1) with differentially upregulated miRNA (log2 fold change≥1) from CN vs NPC, D. differentially upregulated mRNA (log2 fold change≥1) with differentially downregulated miRNA (log2 fold change≤-1) from CN vs NPC.

ESC(N) to NPC upregulated mRNA vs downregulated miRNA interactome analysis:

Here, we considered differentially downregulated miRNAs and differentially upregulated mRNAs during ESC(N) to Neuron progenitor differentiation to know about the miRNAs that were involved in suppression of neuron specific genes in ESC(N) (ESC condition) but plays minimal role in NPC to enrich neuron specific genes in NPC (Fig. B14. B). Some miRNAs contribute major hubs in these networks and are targeting more than 25 number of genes, those are mmu-mir-7b-5p, mmu-mir-466l-5p, mmu-mir-425-5p, mmu-mir-466k, mmu-mir-466d-5p, mmu-mir-301b-3p, mmu-mir-297a-5p, etc. These miRNAs are also involved in regulation of major neurogenesis pathways like, Axon Guidance, Colinergic synapse, Glutamatergic synapse, FoxO signalling pathway, mTOR signalling pathway, etc. Axon Guidance pathway is controlled by these miRNAs in ESC(N) state by targeting genes, *Sema5b, Plxna1, Gsk3ß, plxna4, Ablim1, Arhgef12, Efna5, Tbc1d24, CxcI12, Sema6d, DpysI2, Efnb2, Epha5, Epha7, L1cam, Srgap3, Sema3c, Sema4g, Sema5a, Unc5c* and *Robo2*. Similarly, Colinergic synapse pathway is regulating by targeting *Pik3r3, Kcnq2, Prkca, Creb5, Adcy1, Adcy6, Akt1, Pik3rl, Bcl2, Camk2d, Camk4, Gnao1, Kcn2, Gnaq, Gnb5* and *GngI2*. These miRNAs are also involved in gene ontology related functions. These networks were enriched with neurogenesis related biological processes such as, nervous system development, regulation of cell migration, neurogenesis, neuron differentiation, generation of neurons, axonogenesis, brain development, neuron development, central nervous system, neuron projection development, etc. Also, its showed high enrichment in cellular component towards neuron development like neuron projection, cell protection part, dendrite, axon, synapse part, synapse, cell protection, etc.

ESC(N) to NPC upregulated TF vs downregulated miRNA interactome analysis:

Differentially downregulated miRNAs and differentially upregulated TFs during ESC(N) to Neuron progenitor differentiation were considered to dissect about miRNAs that were involved in suppression of neuron specific TFs in ESC(N) condition but were no longer required in NPC to enrich neuron specific genes in NPC (Fig. B15. B). Some miRNAs contribute major hubs in this network are mmu-miR-466i-5p, mmu-miR-466k, mmu-miR-466d-5p, mmu-miR-7b-5p, mmu-let-7b-5p, mmu-miR-425-5p, mmu-miR-297a-5p, mmu-miR-34b-5p, mmu-miR-466f-3p,

mmu-miR-466i-3p, mmu-miR-301b-3p, etc. Downregulation of these miRNAs in NPC were found to be related to upregulation of Notch signalling pathway, Adherens Junction, Chemokine signalling pathway, Oxytocin signalling pathway, cAMP signalling pathway, Wnt signalling pathway, Hippo signalling pathway, Glutamatergic synapse, etc. Here, Glutamatergic synapse pathway was found to be getting upregulated due to upregulation of *Adcy1, Adrbk1* and *Adrbk2*. These transcription factors were silent in ESC(N) condition due to probable activities of mmu-miR-34b-5p, mmu-miR-136-5p, mmu-miR-7b-5p, mmu-miR-187-3p, mmu-miR-425-5p, mmu-miR-466k, mmu-miR-466f-3p, mmu-miR-466i-3p and mmu-miR-466d-5p.

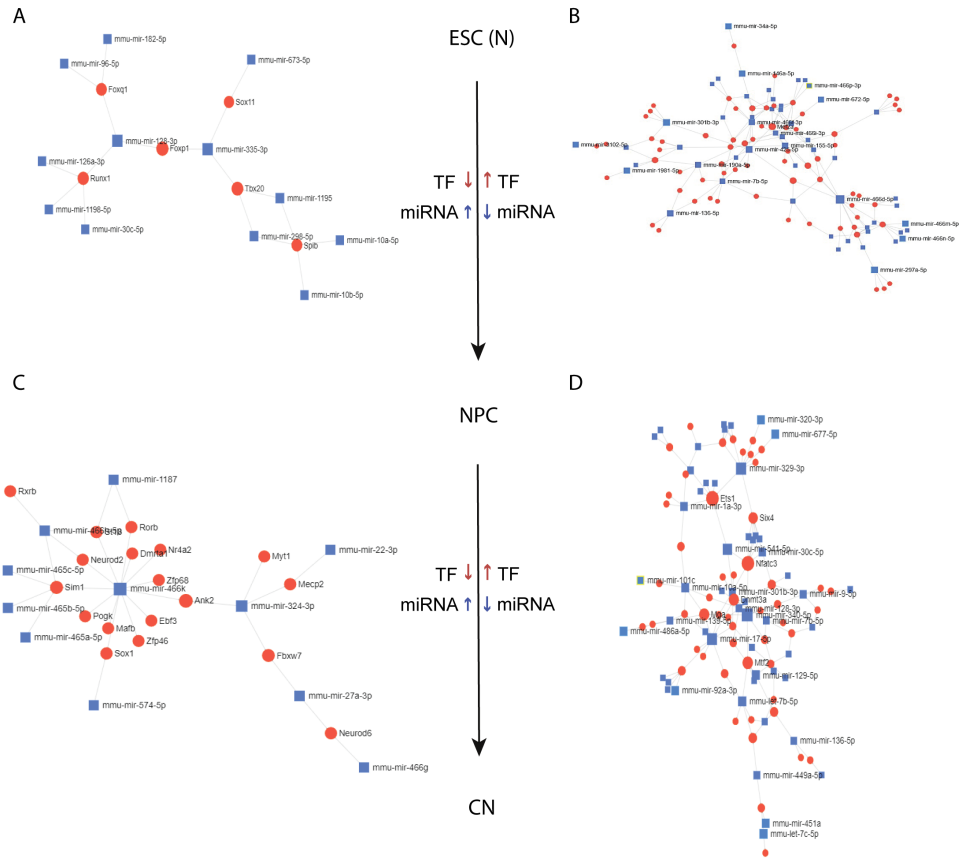


Fig. B15. miRNA-mRNA network map of differentially expressed miRNAs and its experimentally validated differentially expressed TFs in neuronal cells during neurogenesis: Computationally analysed interactome map of differentially expressed TF and differentially expressed miRNA as a whole. A. differentially downregulated TF (log2 fold change≤-1) with differentially upregulated miRNA (log2 fold change≥1) from NPC vs ESC(N), B. differentially upregulated TF (log2 fold change≥1) with differentially downregulated miRNA (log2 fold change≤-1) from NPC vs ESC(N), C. differentially downregulated TF (log2 fold change≤-1) with differentially upregulated miRNA (log2 fold change≥1) from CN vs NPC, D. differentially upregulated TF (log2 fold change≥1) with differentially downregulated miRNA (log2 fold change≤-1) from CN vs NPC.

ESC(N) to NPC downregulated mRNA vs upregulated miRNA interactome analysis:

We then further analysed the networks of differentially upregulated miRNAs and differentially downregulated mRNAs to dissect the important miRNAs required for neuronal differentiation which suppress ESC specific pluripotency genes (Fig. B14. A). Following are the list of miRNAs involved in major central hubs by targeting many relevant mRNAs; mmu-mir-362-3p, mmu-mir-340-5p, mmu-mir-1195, mmu-mir-466i-5p, mmu-mir-9-5p, mmu-mir-181a-5p, 339-5p, 30e-5p, mmu-mir-495-3p, mmu-mir-298-5p, mmu-mir-324-3p, mmu-let-7b-5p, mmu-mir-665-3p, mmu-mir-10b-5p, mmu-mir-149-5p. These above miRNAs were downregulating pathways like Hippo signalling pathway, Signalling Pathway regulating Pluripotency, Tight junction, TGFß signalling pathway, etc. Hippo signalling pathway was found to be majorly downregulated during NPC differentiation. A few miRNAs such as mmu-let-7c-5p, mmu-mir-410-3p, mmu-mir-200c-3p, mmu-mir-134-5p, mmu-mir-93-5p, mmu-mir-668-3p, mmu-mir-344d-3p were predicted to target Hippo signalling pathways by probably targeting expression of *Fzd5, Wnt3, Id1, Id2, Bhc3, Ppp1cb, Wtip, Ppp2r2c, Sox2, Smad7* and *Myc*. Signalling Pathway regulating Pluripotency network involves miRNAs like mmu-mir-9-5p, mmu-mir-340-5p, mmu-mir-134-5p, mmu-mir-673-3p, mmu-mir-410-3p, mmu-mir-344d-3p, mmu-let-7c-5p, mmu-mir-137-3p, mmu-mir-206-3p, mmu-mir-93-5p, mmu-mir-145a-5p, mmu-mir-10b-5p, mmu-mir-204-5p, mmu-mir-26a-5p which supress the expression of *Fzd5, Wnt3, Id1, Id2, Jard2, Klf4, Sox2, Zic3, Tbx3* and *Myc*.

ESC(N) to NPC downregulated TF vs upregulated miRNA interactome analysis:

Differentially upregulated miRNAs and differentially downregulated TFs were visualized to dissect the important miRNAs required for neuron differentiation which suppress ESC specific pluripotent TFs in NPC (Fig. B15. A). Some of the major miRNAs were found to be mmu-miR-340-5p, mmu-miR-362-3p, mmu-miR-181a-5p, mmu-let-7b-5p, mmu-miR-324-3p, mmu-miR-30e-5p, mmu-miR-1195, mmu-miR-466i-5p, mmu-miR-9-5p, mmu-miR-495-3p. These miRNAs were found to be involved in downregulation of Signalling pathway regulating pluripotency, TGF-beta signalling pathway, Hippo signalling pathway, Tight junction, etc. Signalling pathway regulating pluripotency regulating TFs like *Id1, Id2, Jarid2, Klf4, Sox2, Zic3, Tbx3* and *Myc* were also being predicted to be targeted by these above miRNAs in NPC stage.

NPC to CN upregulated mRNA vs downregulated miRNA interactome analysis:

Further analysis was proceeded towards differentiation of cortical neuron from neuron progenitor with differentially downregulated miRNAs with differentially upregulated mRNAs to infer the essential miRNAs need to be suppressed for terminal differentiation of neurons and to enrich terminal neuron specific genes (Fig. B14. D). It provided top significant miRNAs which were occupying major hubs in networks, such as mmu-mir-24-3p, mmu-mir-466k, mmu-mir-466i-5p, mmu-mir-324-3p, mmu-mir-466i-3p, mmu-mir-466f-3p, mmu-mir-1195, mmu-mir-377-3p, mmu-mir-339-5p, mmu-mir-466p-5p, mmu-mir-1187, mmu-mir-743a-3p, mmu-mir-743b-3p, mmu-mir-150-5p, mmu-mir-290a-5p, etc. These above lists of miRNAs were getting downregulated from NPC to CN and resulted in upregulation of mRNAs that were involved in major neuron differentiation pathways like, Axon guidance, Glutamatergic synapse, Dopaminergic synapse, Synaptic vesicle cycle, Neuroactive ligand-receptor interaction. Glutamatergic synapse pathway related genes *Ging2, Slc17a6, Adrbk2, Grik3, Plcb1, Adcy5, Gnao1, Grm1, Homer2, Gnaq, Gnb5, Slc1a2, Gria2, Grin1, Grin2b* and *Prkcb* showed high upregulation in CN due to downregulation of the above-mentioned miRNAs. Also, suppression of these miRNAs resulted in possible upregulation of *Gng2, Bmal1, Plcb1, Drd1a, Adcy5, Gnao1, Drd2, Kif5c, Gna1, Gnaq, Gnb5, Gria2, Grin2b, Prkcb, Slc18a2, Ppp2r2c* genes which were involved in Dopaminergic synapse pathway. Synaptic vesicle cycle pathway, that we observe in active neurons during release of neurotransmitters, also showed upregulation with increase expression of *Fam203a, Stx1b, Slc17a6, Dnm3, Cplx1, Cplx2, Snap25, Syt1, Slc18a2*. Similarly, Neuroactive ligand-receptor interactions, a pathway activated in mature neurons showed upregulation in CN state with increase expression of *Adra1b, Dixdc1, Grid1, Grik3, Tacr1, Drd1a, Adcyap1r1, Chrna4, Cnr1, Crhr1, Grm1, Drd2, Gabra1, Gabrb2, Gria2, Grin1, Grin2b.*More genes related to neuronal development were found to be expressed at terminal stage. Important neurogenesis related biological processes like, transmission of nerve impulse, synaptic transmission, cell-cell signalling, locomotory behaviour, neuron development, neuron projection development, neuron differentiation, axonogenesis, brain development, generation of neurons, nervous system development, regulation of neurotransmitter, central nervous system development, axon guidance, learning or memory, neurotransmitter secretion, synapse organization, regulation of neuron apoptotic process, glutamate receptor signalling pathway, regulation of neurogenesis, regulation of synapse structure and activity, regulation of action potential, synapse assembly, regulation of axonogenesis, etc were found to be present during CN development. Further, significant cellular component which were

essential for neuron development such as, neuron projection, synapse, synapse part, axon, dendrite, synaptic vesicle, cell protection part, cell junction, etc. were enriched during the course of terminal differentiation.

NPC to CN upregulated TF vs downregulated miRNA interactome analysis:

Differentially downregulated miRNAs with differentially upregulated TFs interactome maps were analysed to infer the essential miRNAs needed to be suppressed for terminal differentiation of neurons and to enrich terminal neuron specific TFs (Fig. B15. D). These major miRNAs' hubs were found to be mmu-miR-24-3p, mmu-miR-466k, mmu-miR-466i-5p, mmu-miR-324-3p, mmu-miR-1195, mmu-miR-466f-3p, mmu-miR-466i-3p, mmu-miR-339-5p, etc. and these miRNAs were targeting pathways like Glutamatergic synapse, Notch signalling pathway, AMPK signalling pathway, Insulin signalling pathway, Gap Junction, Thyroid hormone signalling pathway, Dopaminergic synapse, etc. Transcription factors *Bmal1* and *Drd1a* which were regulator for Dopaminergic synapse were kept silent in NPC, probably due to the activities of mmu-miR-142a-3p and mmu-miR-24-3p. But, in CN the downregulation of these miRNAs resulted in elevation of Dopaminergic synapse pathway.

NPC to CN downregulated mRNA vs upregulated miRNA interactome analysis:

To visualize the interaction map and pathways that needed to be downregulated during terminal differentiation of neuron and to segregate miRNAs which were targeting these genes, we considered, differentially upregulated miRNAs with differentially downregulated mRNAs interaction map (Fig. B14. C). These interactome maps were analysed and which provided some significant miRNAs with major central hubs such as, mmu-mir-329-3p, mmu-mir-124-3p, mmu-mir-340-5p, mmu-mir-9-5p, mmu-mir-17-5p, mmu-mir-129-5p, mmu-mir-15a-5p, mmu-mir-301b-3p, mmu-let-7b-5p, mmu-mir-7b-5p, mmu-mir-425-5p, mmu-mir-34b-5p, mmu-mir-362-5p, mmu-mir-10a-5p, etc. These miRNAs were involved in downregulation of pathways like Hippo signalling pathway, Cell cycle, p53 signalling pathway, etc. Genes related to Hippo signalling pathway were *Tead1, Yap1, Tead4, Fzd7, Fzd5, Tcf7l1, Wnt2b, Zim1, Pard6b, Serpine1, Tgfb1, Tgfb2, Afp, Ajuba, Amot, Bmp4, Bmp6, Tgfbr2, Lef1, Myc, Fzd6 and Wtip* were downregulated during differentiation upon upregulation these above-mentioned miRNAs. Downregulation of cell cycle was observed during terminal differentiation of cells. Here these miRNAs were involved in

downregulation of cell cycle related genes like, *Cdk7, Mcm5, Atm, Bub1, Ttk, Chek2, Sfn, Sltm, Mdm2, Tgfb1, Tgfb2, Ccne2, Cdc14b, Cdc27, Cdc7, Cdk6, Chek1, E2f3* and *Myc*. These counter expressions of mRNAs and miRNAs throughout the neurogenesis gave us a cleared picture of tight regulations during neuronal development through miRNAs.

NPC to CN downregulated TF vs upregulated miRNA interactome analysis:

To visualize the TFs related pathways and those that were needed to be supressed during terminal differentiation of neuron, we analysed interaction maps of downregulated TFs and upregulated miRNAs during NPC to CN transition (Fig. B15. C). These interactome maps have major miRNA hubs like, mmu-miR-340-5p, mmu-miR-329-3p, mmu-miR-17-5p, mmu-miR-129-5p, mmu-miR-301b-3p, mmu-miR-9-5p, mmu-let-7b-5p, mmu-miR-124-3p, etc. These miRNAs were involved in downregulation of pathways like, cell cycle, NF-kappa B signalling pathway, Apoptosis, signalling pathway regulating pluripotency, Hippo signalling Pathway, etc. in terminally differentiated neurons.

Interplay between Pbx1 and PbxIP1 along with mmu-miR-301b-3p determines neuronal lineage:

Pre B-cell Leukaemia homeodomain (PBX) is a transcription factor that contributes towards regional identity during embryonic development and controls organ development (Manavathi et al., 2012). Previous studies have highlighted the association of PBX1 and MEIS during neuronal fate commitment (Golonzka et al., 2015, Grebbin et al., 2016). Pbx1 also has a critical role in decision making in early neurogenesis and found in *Tuj1* and *Nestin* positive cells, but found to be lost in *Gfap* positive astrocytes and *O4* positive oligodendrocytes (Grebbin et al., 2016). Upon knockout condition of same, neuronal fate shifted towards oligodendroglial lineage. In our dataset Pbx1 expression increases from ESC(N) to NPC and keep on increasing towards cortical neuron (Fig. B16. C). PBX1 has a well-studied corepressor, PBX1 interacting protein (PBXIP1). The interaction of these two proteins were reported previously in Erythroid differentiation (Manavathi et al., 2012) and recently in osteoarthritis (Ji Q et al., 2019). This information made us to reanalyse interaction status of PBX1 and PBXIP in neurogenesis. It has been observed, that the expression of *Pbxip1* is quite high in NPC but get downregulated from NPC to CN condition, which was contrary to Pbx1 expression (Fig. B16. C).

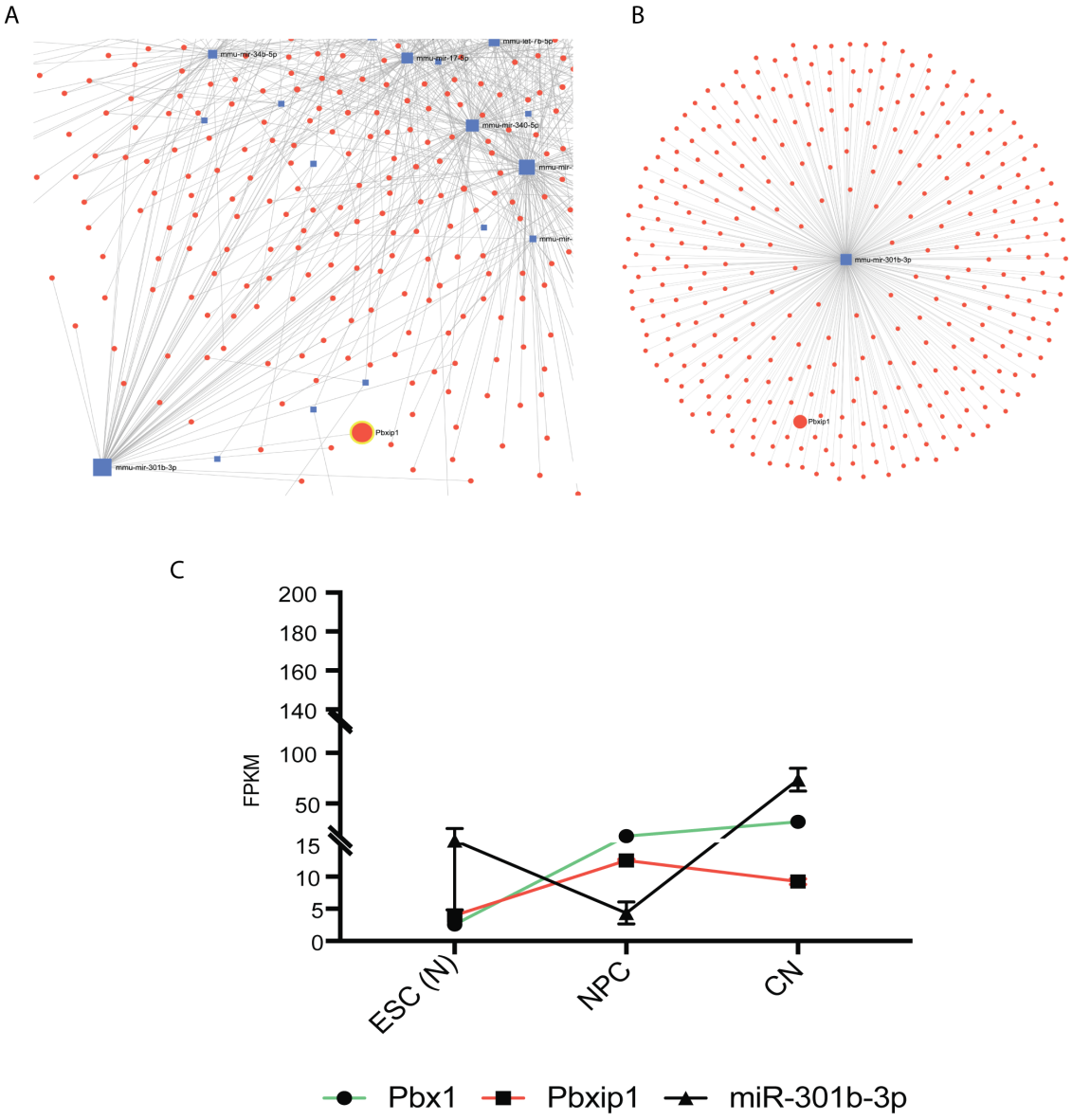


Fig. B16. Gene expression and interaction study of mmu-miR-301b-3p: A. Zoomed in view of mmu-miR-301b-3p hub from expressed mRNA (FPKM≥1) of NPC with differentially upregulated miRNA (log2 fold change≥1) from CN vs NPC, B. miRNA-mRNA interactome map representing total 335 expressed mRNAs are targeted by *mmu-miR-301b-3p* in CN, C. Line graph representing FPKM status of *Pbx1, Pbxip1* and *miR-301b-3p* in ESC (N), NPC and CN which showed elevated expression of *Pbx1* and *mmu-miR-301b-3p* but depletion of *Pbxip1* in CN.

Here, Pbx1 expression induced with downregulation of *Pbxip1*. This downregulation of *Pbxip1* might be regulated by *mmu-miR-301b-3p* (Fig. B16. A). Role of *mmu-miR-301b-3p* was reported first time in prostate cancer pathogenesis (Fort et al., 2018). This miRNA was involved in tumour initiation, migration and invasion properties. The interaction of *mmu-miR-301b-3p* and Pbxip1 was

reported first time in HITS-CLIP study of mouse brain (Chi et al., 2009). But, significance of the interaction remained elusive. *Mmu-miR-301b-3p* expression was found to be downregulated in NPC vs ESC(N) but was differentially upregulated from NPC to CN. Expression of *mmu-miR-301b-3p* and *Pbxip1* were found to be following a reciprocal expression pattern during neurogenesis. *mmu-miR-301b-3p* has total 335 expressed targets in cortical neuron state (Fig. B16. B). It maintains a large core hub of differentially downregulated genes, among them Pbxip1 was found to be solely regulated by *mmu-miR-301b-3p*. We then focussed our attention on the role of *mmu-miR-301b-3p* in the context of cortical neuron differentiation.

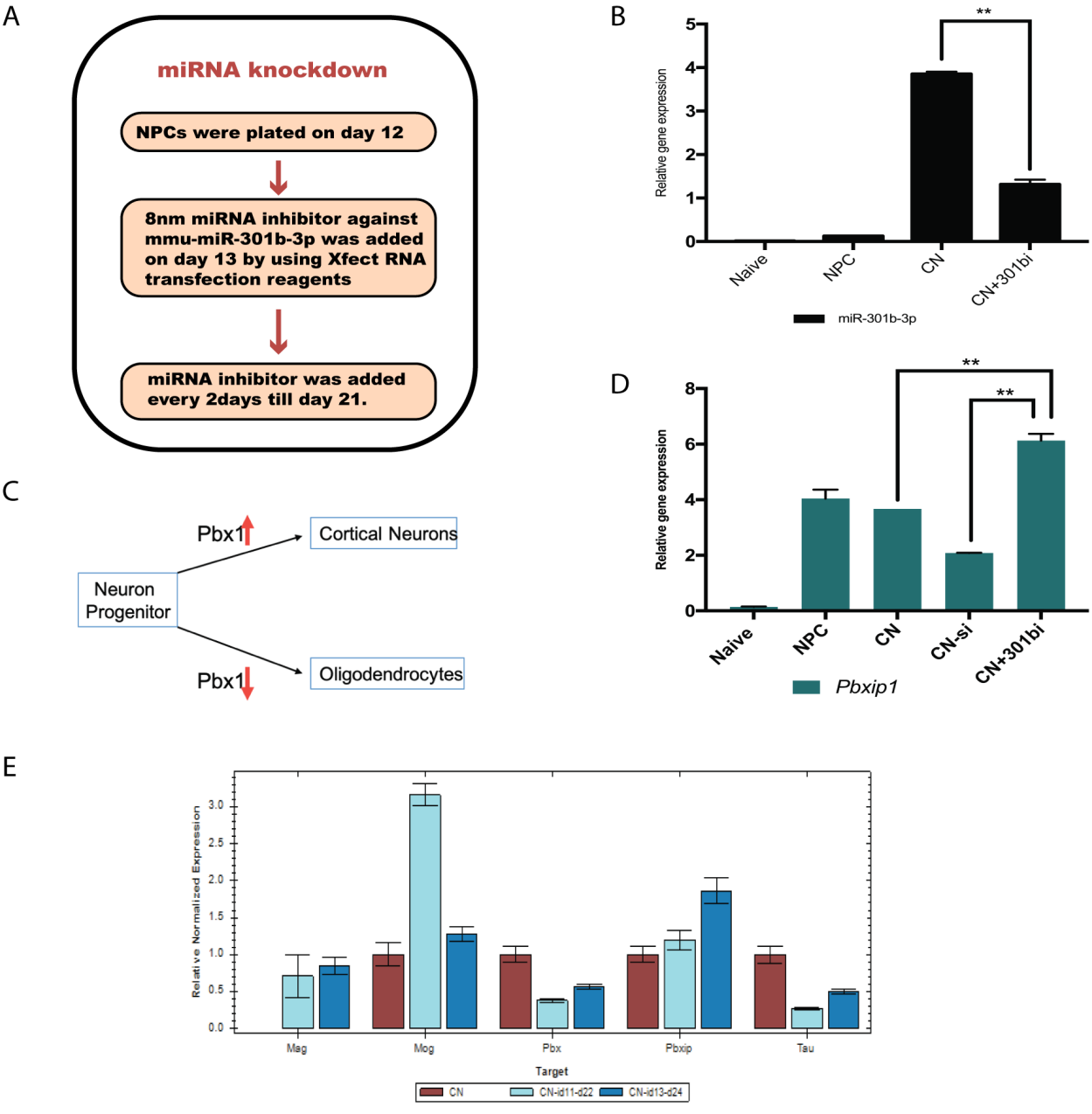


Fig. B17. Knockdown study of mmu-miR-301b-3p on Neurogenesis: A. Flow chart representing mmu-miR-301b-3p knockdown study in Corticogenesis (Detail protocol of mmu-miR-301b-3p knockdown in cortical neurons has been provided in method section), B. Relative gene expression status of *miR-301b-3p* through Real-time PCR in between ESC (N), NPC, CN and knockdown sample (CN+301bi). Data was analysed by using $-2^{\Delta\Delta CT}$ method. *Rnu6* and ESC(N) were considered as gene control and sample control respectively. Bar graphs were generated by using Graph-pad PRISM software with one-way ANOVA significance test. The expression of *miR-301b-3p* was significantly dropped down in knockdown sample (CN+301bi). C. Relative gene expression status of *Pbxip1* through Real-time PCR in between ESC (N), NPC, CN and knockdown sample (CN+301bi). Data was analysed by using $-2^{\Delta\Delta CT}$ method. *Gapdh* and ESC(N) were considered as gene control and sample control respectively. Bar graphs were generated by using Graph-pad PRISM software with one-way ANOVA significance test. The expression of *Pbxip1* was significantly increased in knockdown sample (CN+301bi). D. Graphical representation showing importance of PBX1 during Corticogenesis. E. Transfection efficiency comparison by real-time PCR between transfection day 11 and allowed to grown up to day 22 and transfected at day 13 allowed to grown up to day 24. Here, day13 of transfection was used for further experimentation.

Knockdown of *mmu-miR-301b-3p* changed cell fate towards oligodendrogliogenesis:

To investigate the potential role of *mmu-miR-301b-3p* during neurogenesis, we knocked down mmu-miR-301b-3p by transfecting hairpin inhibitors against *mmu-miR-301b-3p* on day 13 of neuronal differentiation (Fig. B17. A). It was convenient to transfect cells after plating of neuron progenitors on day 12 on a poly-laminin coated plate, when expression of *mmu-miR-301b-3p* levels were low. It was observed that addition of transfection reagents before plating of NPC resulted in massive cell death. Hence, we have provided a comparison in between day 11 transfected and allowed cells to grow up to day 22 along with day 13 transfected and allowed the cells to grow up to day 24. As these hairpin inhibitors works best after 96 hours of post- transfection and the expression of *mmu-miR-301b-3p* also start getting increased after plating of NPC, the functional aspects of *mmu-miR-301b-3p* was analysed by comparing transfection in NPCs on day 11 and on day 13. In both the cases, reduced expression of *Pbx1* and increased expression of *Pbxip1* were observed. Even more reduced expression of *Tau* and significant increase in *Mog* were observed on day11 of transfection (Fig. B17. E). But Cell death was marked more in case of day11 of transfection. Therefore, day13 of transfection was used for further experimentation.

Experiment was monitored parallelly during cell differentiation with negative inhibitor condition (CN-si condition). Upon differentiation till day 21, reduced expression of *Tuj1* has been observed in knockdown condition but no significant differences were found at translation level (Fig. B18. B). Reduced *Pbx1* expression was marked compared to CN and CN-si condition (Fig. B18. A). But *Pbxip1* expression in both transcriptional and translational level increased significantly in knockdown condition (Fig. B18. B). As we hypothesized that with *mmu-miR-301b-3p* knockdown (CN+301bi) elevates *Pbxip1* expression and resulted in decrease in *Pbx1* transcripts. This made us to check the status of *Tuj1* transcripts in neuron cell populations. In knocked down cells, increased oligodendrocytes population was observed in neuronal population, with increased expression of oligodendrocyte markers *Mog* and *Mag* (Fig. B18. A). Increased Mag expression at protein level was confirmed in CN+301bi compared to CN and CN-si (Fig. B18. B). Also, this was confirmed with immunofluorescence studies where knockdown sample showed decreased TAU and enhanced MAG signal as compared to CN and CN-si (Fig. B19).

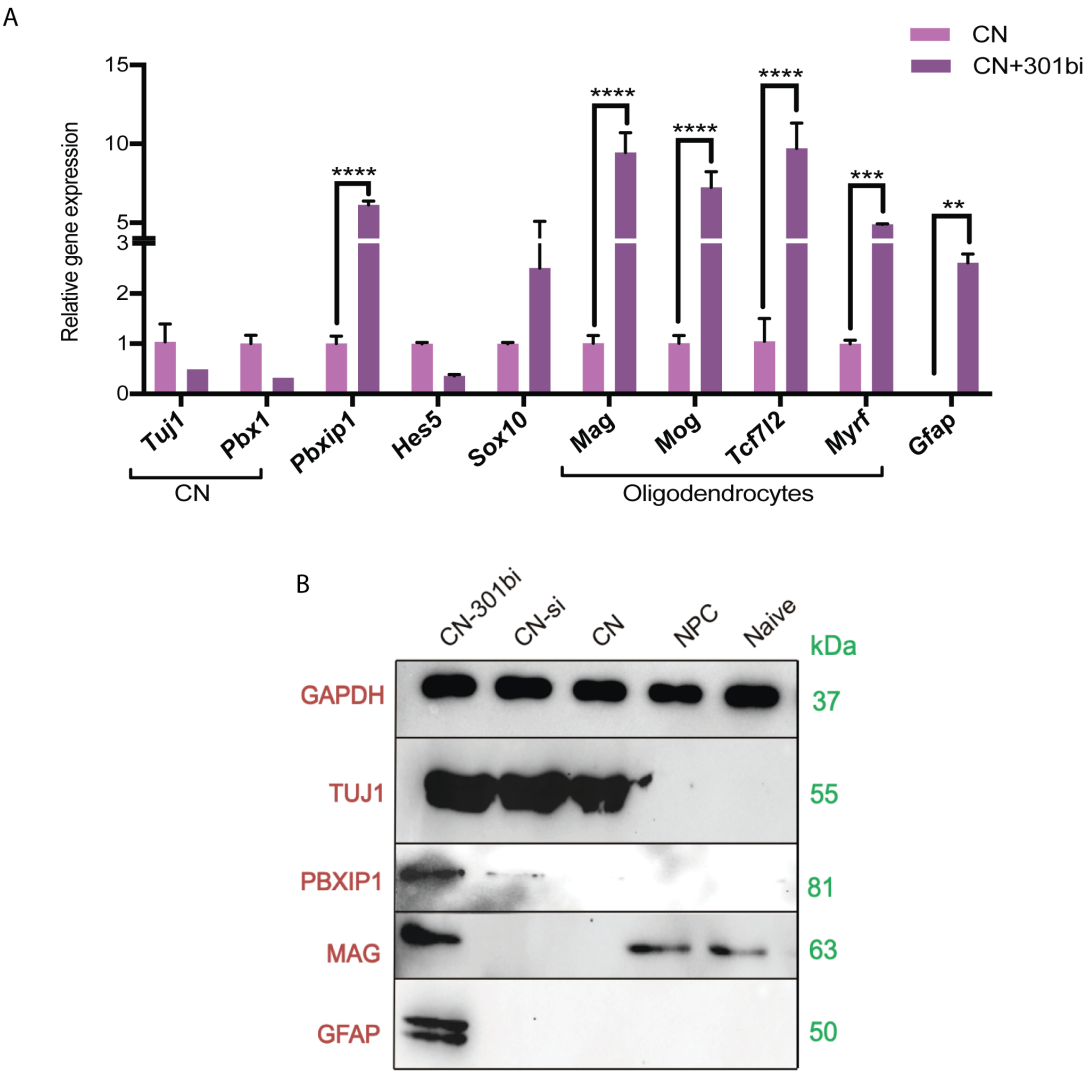


Fig. B18. Study of mmu-miR-301b-3p knockdown effect on cortical neuron differentiation: A. Relative gene expression pattern of gene markers from Neurons, Neuron stem cells, Oligodendrocytes and Astrocytes through Real-time PCR in CN+301bi knockdown condition. Data was analysed by using $-2^{\Delta\Delta CT}$ method. *Gapdh* and CN were considered as gene control and sample control respectively. Bar graphs were generated by using Graph-pad PRISM software with one-way ANOVA significance test. B. Comparison of expression of TUJ1, PBXIP1, MAG and GFAP in between ESC(Naive), NPC, CN, CN+301bi knockdown condition by using western blot.

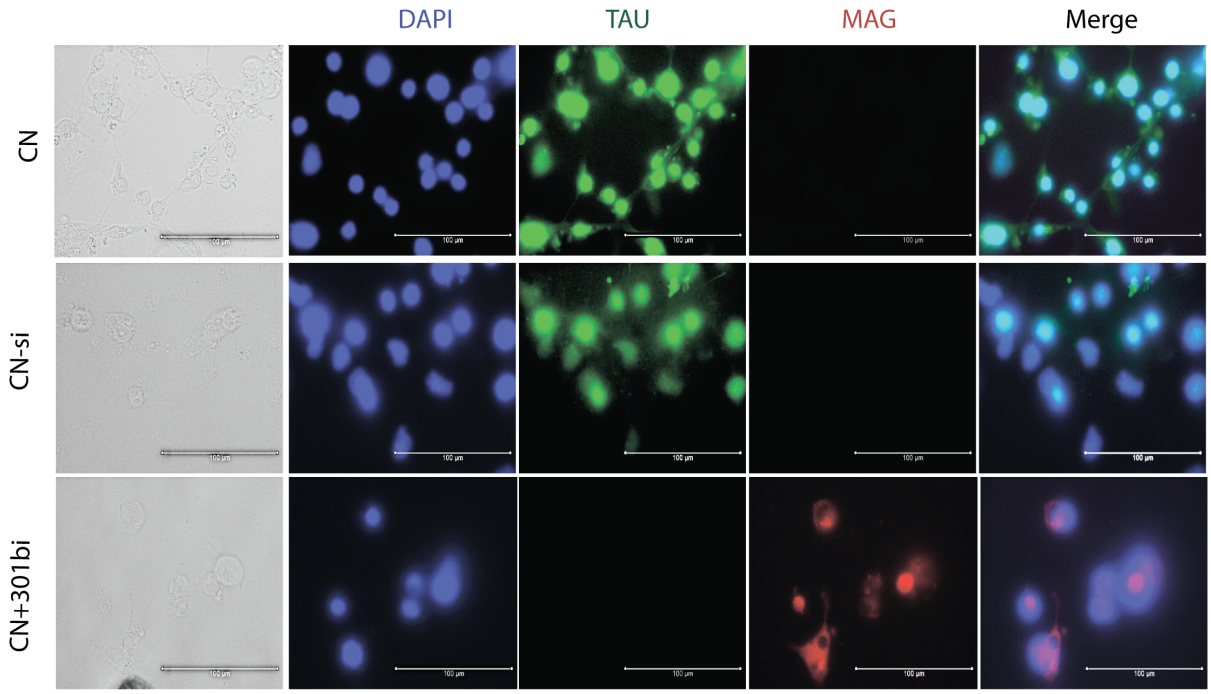


Fig. B19. Immunofluorescence study of mmu-miR-301b-3p knockdown effect on cortical neuron differentiation: Comparison of 63X fluorescence images from CN, CN-si and CN+301bi by using Oligodendrocyte marker MAG and Neuron marker TAU. Here, knockdown sample, CN+301bi is clearly showing decreased TAU and enhanced MAG signal as compared to CN and CN-si

Precursor cells of neurons, astrocytes and oligodendrocytes, altogether originate from multipotent neuroepithelial progenitor cells (Emery et al.,2015). Oligodendrocytes precursor cells (OPC) are highly proliferative and motile in nature. Once they migrate to the specified position adjacent to neurons, they undergo terminal differentiation for myelinating axons. Many transcription factors regulate this migration and specification of OPC towards differentiated myelinating oligodendrocytes. Oligodendrocytes are seen in three different stages during differentiation, as

OPCs, premyelinating and then myelinating. To understand the nature of oligodentrocytes generated *invitro* in our knock down and time course studies, we used some stage specific transcription factor markers. Hes5 is targeted by Notch to limit OPC differentiation by competing with promyelinating factor Sox10 (Liu et al., 2019). Hes5 is OPC specific markers and show gradual decrease in expression towards myelinating oligodendrocytes. *Hes5* came negative in knockdown population, which implies absent of OPCs after 9days of *mmu-miR-301b-3p* knockdown culture (Fig. B18. A). While increase in Sox10 expression upon knockdown confirms that a population of NSCs have commit towards oligodendrocyte lineage. But it showed significant increased expression for *Tcf7l2*, which is a marker of premyelinating oligodendrocytes and *Myrf,* which is a marker of both premyelinating and myelinating oligodendrocytes (Emery et al., 2015). It indicates majority of oligodendrocyte population were in premyelinating and myelinating stages.

Few percentage of oligoderoglial progenitor cells also differentiate in to astrocytes (Tao et al., 2016). That's why when astrocyte marker Gfap was checked in CN+301bi population, it gave significantly increased expression upon *mmu-miR-301b-3p* knockdown in both transcriptional and translational level (Fig. B18. B). We thus inferred divergence of cell lineage towards Oligoderoglial progenitor cells upon knockdown of *mmu-miR-301b-3p* (Fig. B20).

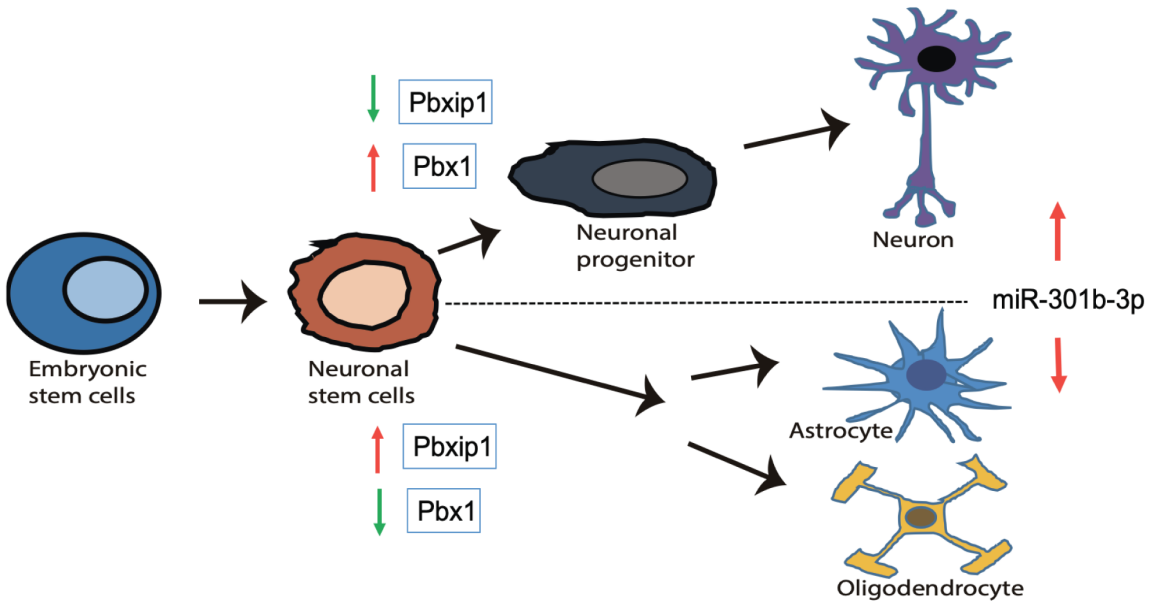


Fig. B20. Graphical representation showing role of mmu-miR-301b-3p during neuron differentiation.

Objective III

MicroRNA dynamics during signal induced cellular differentiation in mammals

High throughput miRNA sequencing of mouse mammary epithelial stem like cells during lactogenic differentiation:

Undifferentiated state of proliferative HC11 MECs (N) are maintained under the influence of EGF. Their lactogenic (secretory) differentiation is induced by growing them to confluency, then priming them with glucocorticoids (P) followed by induction with MEC(Prl) hormone treatment (Prl). Comprehensive profiling of mRNA transcriptional networks during HC11 MECs lactogenic differentiation have shown the dynamic nature in expression of thousands of mRNAs in stage specific context (Sornapudi et al., 2018, Wang et al., 2009, Williams et al., 2009, Perotti et al., 2009). In conjunction with this study, a potential role of miRNAs in post-transcriptional control of gene expression was comprehensively studied by high throughout next generation sequencing using MEC (N), (P) and (Prl) stages (Fig. C1. A-C). Though, a previous microarray based study described expression of miRNAs but is suffered from low complexity (Aydogdu et al. 2012). The above-mentioned study also highlighted the role of miR-200a and miR-200b in inhibiting EMT. However, considering the recent updates at miRbase-V22 (Griffiths-jones et al., 2006), the previous studies were not comprehensive and failed to capture quantitative profiles. Our study address the lacunae present in previous studies and is more comprehensive and provides more accurate quantitative profiles.

Differential expression of miRNAs during lactogenic differentiation of HC11 MECs:

Analysis of miRNA-seq data showed dramatic alterations in miRNA gene expression profiles during the course of lactogenesis. Our previous study (Sornapudi et al., 2018) showed that lactogenesis is mostly accompanied with cell cycle arrest, so the dramatic alterations in miRNA expression during different stages of lactogenic differentiation of HC11 MECs occurs in absence

of cell cycle progression. Expression of 122, 142, 114 number of miRNAs were observed in MEC (N), MEC (P) and MEC (Prl) stages respectively (Fig. C2. A). Among them 6 (mmu-miR-3535, mmu-miR-501-3p, mmu-miR-200c-3p, mmu-miR-210-3p, mmu-let-7a-1-3p and mmu-let-7c-2-3p), 19 (mmu-miR-1a-3p, mmu-miR-155-5p, mmu-miR-328-3p, mmu-miR-196a-5p, mmu-miR-215-5p,

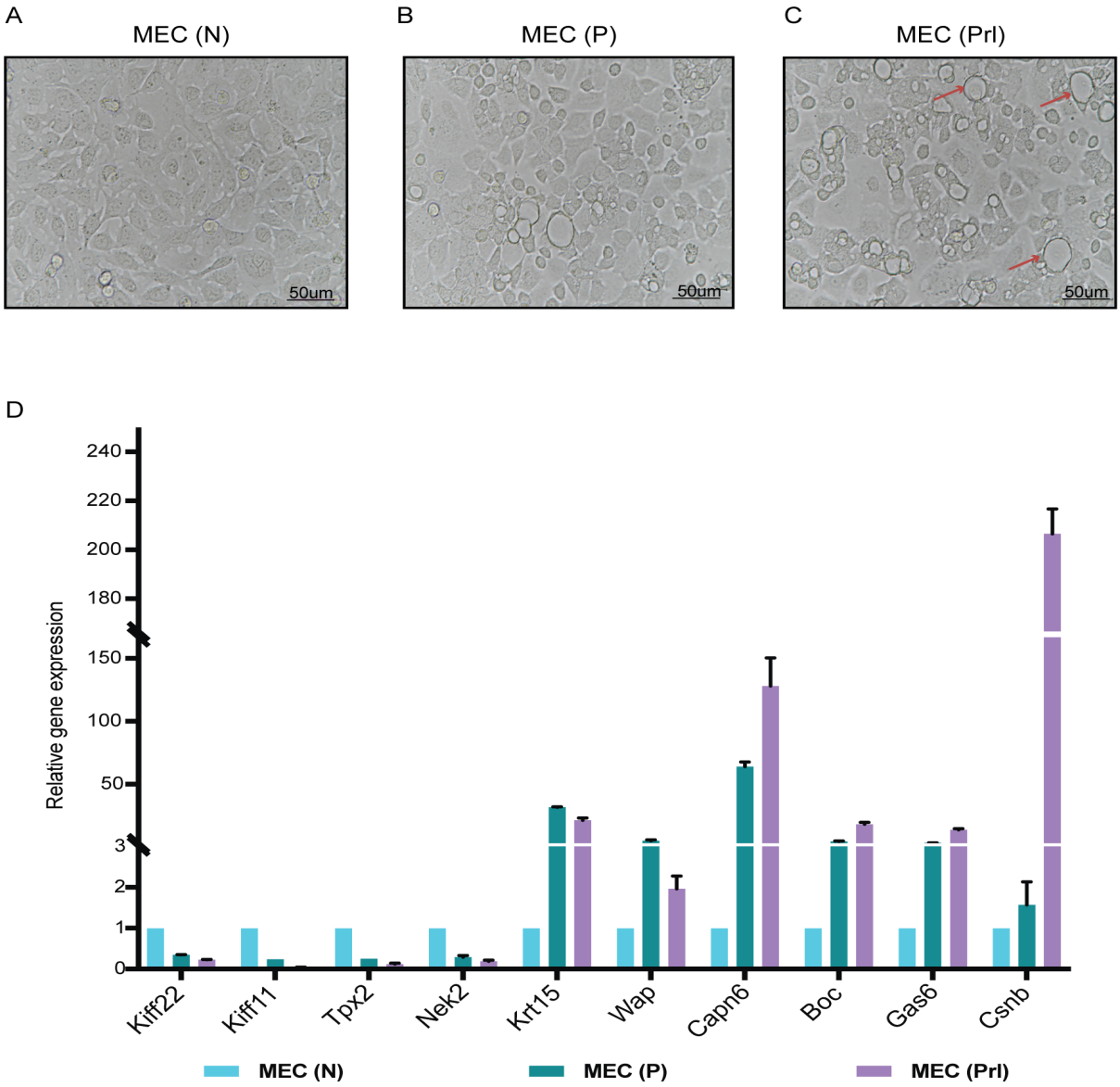


Fig. C1. Characterization of lactogenic differentiation of Mammary Epithelial Cells (MECs): Bright field 40X microscopic images of growing. **A.** undifferentiated HC11 MEC(N) cells at confluent stage in presence of EGF and Insulin, **B.** differentiated MEC(P) cells in presence of Insulin and HC, and **C.** differentiated MEC(Prl) cells in presence of Insulin, HC, and Prolactin. Red arrow marks represent the formation of mammospheres under Prolactin condition. Scale bar represents 50 μ M. **D.** Bar chart representing Real time PCR analysis for stage specific markers of MEC (N), MEC (P), MEC (Prl). Data was analysed by using $-2^{\Delta\Delta CT}$ method. *Beta-Actin* and MEC(N) were considered as gene control and sample control respectively. Bar graphs were generated by using Graph-pad PRISM software with one-way ANOVA significance test.

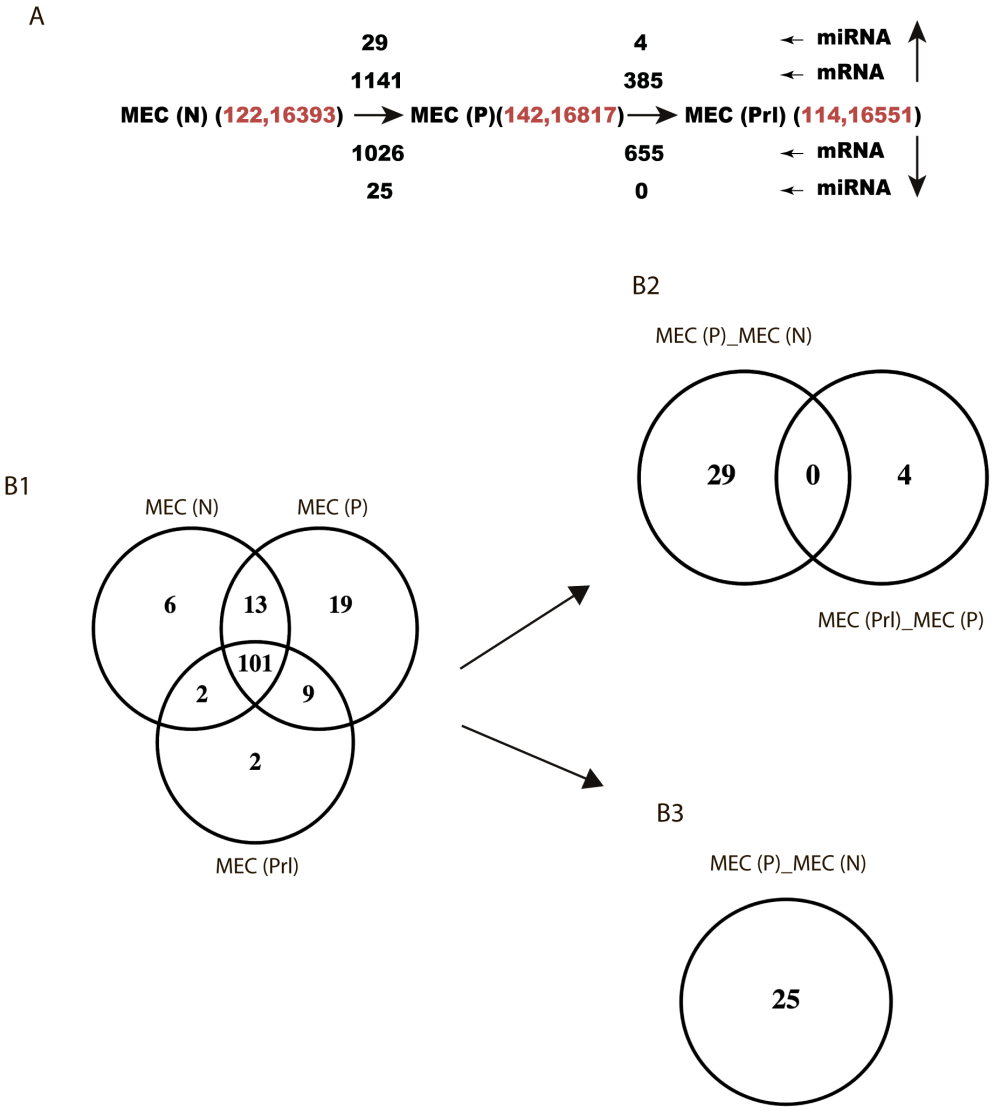


Fig.C2. Schematic representation of statistically analysed miRNAseq dataset of HC11 MECs differentiation: A. Flow chart representing expressed and differentially regulated miRNAs in MEC (N), MEC (P), MEC (Prl) stages of lactogenesis. Venn diagram representing **B1.** unique and overlapping expressed miRNAs (NC≥10) in MEC (N), MEC (P), MEC (Prl), **B2.** unique and overlapping differentially upregulated miRNAs (Log2 Fold change ≥ 1) between MEC(P) vs MES(N) and MEC(Prl) vs MES(P), **B3.** unique and overlapping differentially downregulated (Log2 Fold change ≤ -1) miRNAs between MEC(P) vs MEC(N) and MEC(Prl) vs MEC(P).

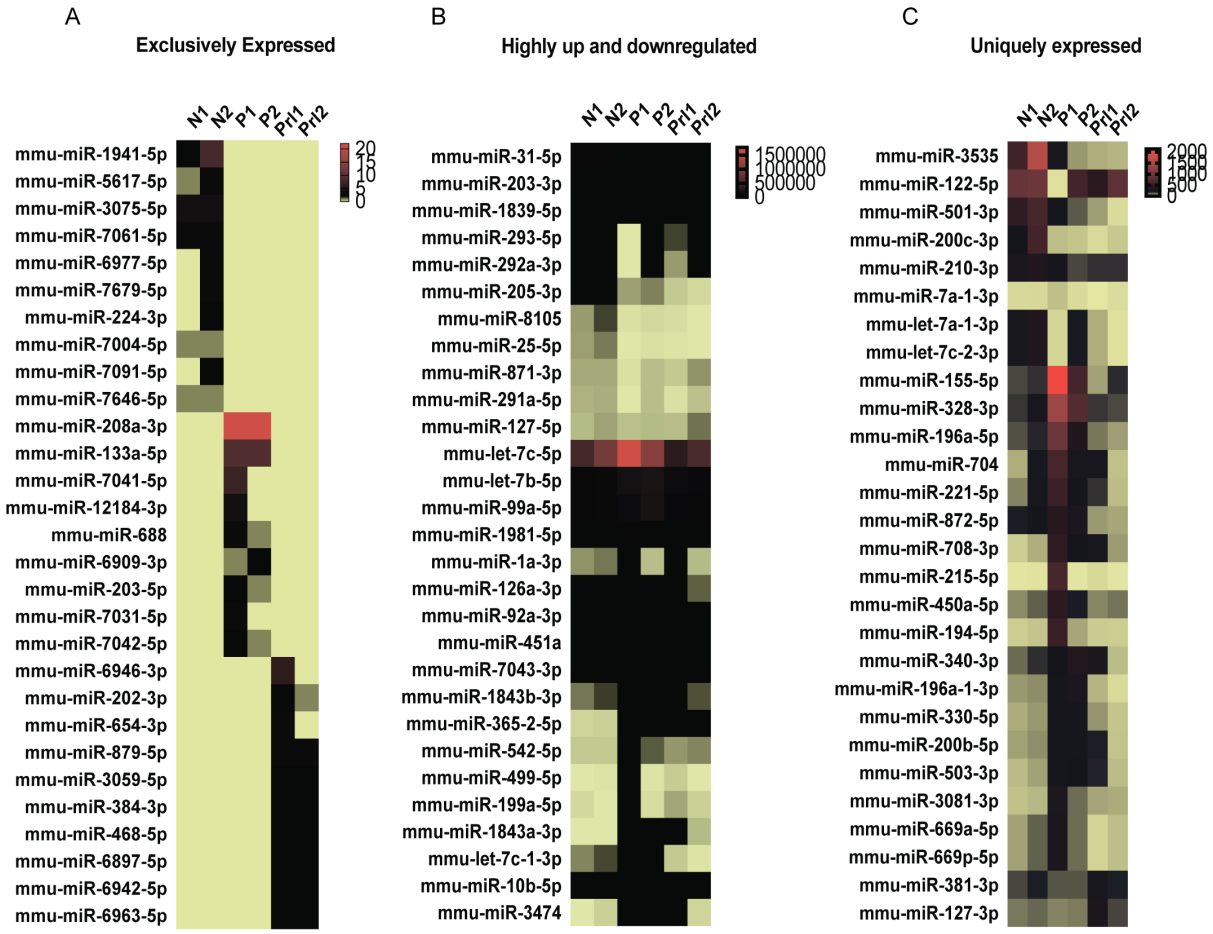


Fig.C3. Comparative analysis of uniquely and differentially expressed genes in HC11 MECs cells undergoing lactogenic differentiation: Heatmap representing **A.** exclusively expressed, **B.** Highly up and downregulated, **C.** Uniquely expressed miRNAs in MEC (N), MEC (P), and MEC(Prl). The list of miRNAs based on, uniquely expressed (Table 68), highly expressed (Table 69), top 30 upregulated (Table 70) and top 30 downregulated (Table 71) were available in Table section.

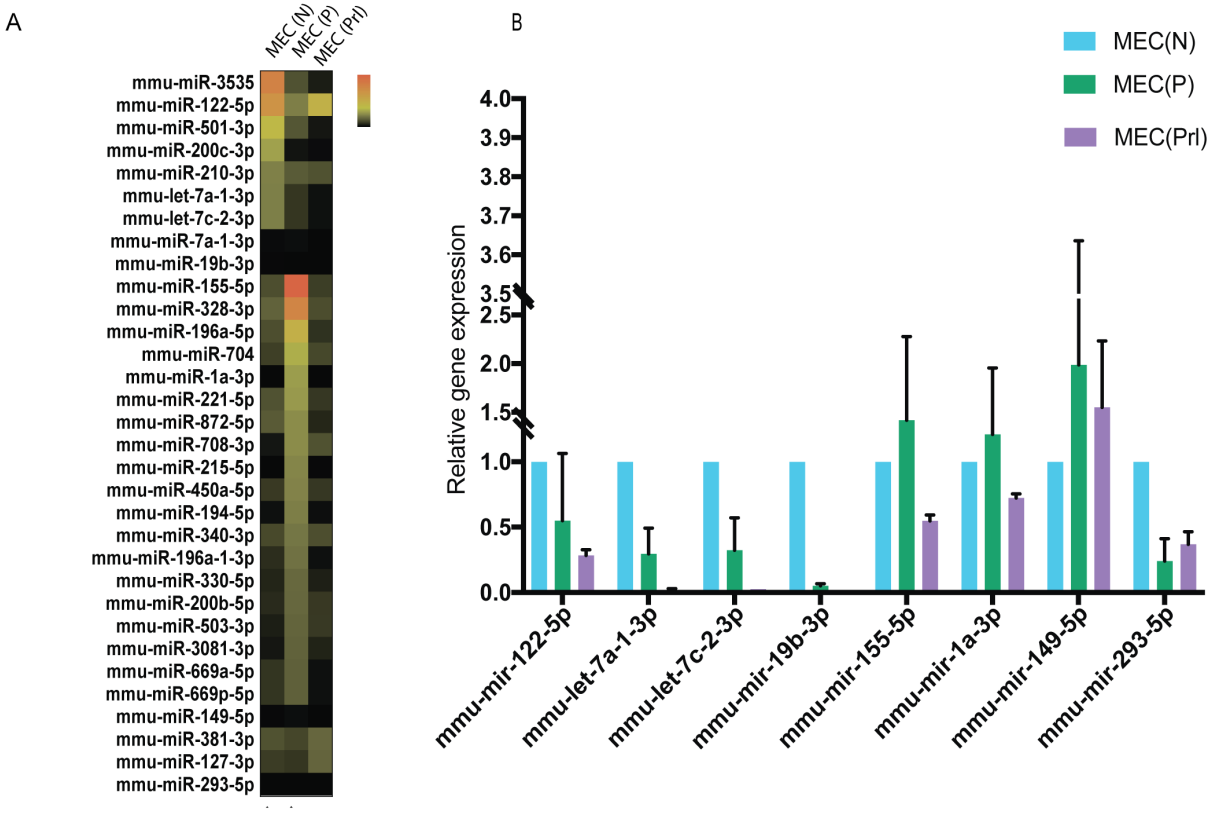


Fig.C4. Validation of stage specific miRNAs during HC11 MECs lactogenic differentiation; A. Heatmap representing expression status of MEC (N), MEC (P), and MEC (Prl) stage specific miRNAs, **B.** Validation of few miRNAs by Real-time PCR for MEC (N), MEC (P), and MEC (Prl) stages. Data was analysed by using -$2^{\Delta\Delta CT}$ method. *Rnu6* and MEC(N) were considered as gene control and sample control respectively. Bar graphs were generated by using Graph-pad PRISM software with one-way ANOVA significance test. The miRNA genes were found to mimic respective miRNA-seq datasets.

mmu-miR-503-3p, etc.), and 2 miRNAs (mmu-miR-381-3p and mmu-miR-127-3p) are uniquely expressed in MEC (N), MEC (P) and MEC (Prl) stages (Fig. C2. B1). We found differential upregulation of 29 and 4 miRNAs during developmental transition from MEC (N) to (P) and MEC (P) to (Prl) respectively (Fig. C2. B2). Transition from MEC(P) states to Prl state is associated with slight upregulation of only 4 microRNA. Significant changes were observed at mRNA level during transition from MEC(P) to (Prl) but had little co-relation at microRNA level. We also observed differential downregulation of 25 miRNAs between MEC (N) vs MEC (P). When analysed for novel miRNAs, four unique novel miRNAs corresponding to respective MEC(N) and (P) state and one in MEC (Prl) state were observed. To identify and characterize stage specific

miRNAs, top 20 highly expressed, uniquely expressed and differentially upregulated miRNAs were short listed (Fig. C3). Some of these miRNAs include mmu-miR-122-5p, mmu-let-7a-1-3p, mmu-let-7c-2-3p and mmu-miR-19b-3p which were specific for MEC (N), mmu-miR-155-5p, mmu-miR-1a-3p and mmu-miR-149-5p which are specific for MEC (P) and mmu-miR-293-5p which is specific for MEC (Prl) were validated through real-time PCR (Fig. C4. B) and found to mimic respective miRNA-seq datasets. Further listing of miRNAs based on, uniquely expressed (Table 68), highly expressed (Table 69), top 30 upregulated (Table 70) and top 30 downregulated (Table 71) were then done. Subsequently, we have used these data to dissect complex gene regulatory networks governing lactogenic differentiation of HC11 MECs.

Interestingly, it has been observed that no miRNAs were downregulated during MEC (P) to MEC (Prl) transition though most of the miRNAs maintain their expression during MEC(Prl) stages. These observations, indicate that major changes in miRNA transcript levels were mediated under the condition of ECM and GR signalling and were maintained even after the addition of MEC(Prl) hormone.

Inter-LAD region contains most of the miRNAs:

To understand the spatial distribution of microRNA in mouse genome, localized information of miRNAs in genome was extracted and was analysed for their location and expression in published DamID data sets of mouse that have previously characterized chromatin into Constitutive LADs(cLADs), constitutive inter LADs(ciLADs) and facultative LADs(fLADs). On an average 79 miRNAs from constitutive LADs (cLAD) showed expression in MEC(N), 78 on MEC(P) and 79 in MEC(Prl). Among the miRNAs only one was found to be differentially upregulated from ESC to MEC(N); 3 miRNAs were upregulated from MEC(N) to MEC(P) and no change from MEC(P) to MEC(Prl). We also observed that no miRNAs were down regulated from MEC(N) to MEC(P) state (Fig. C5). We repeated our analysis by now focusing on miRNAs from constitutive inter LADs(ciLAD). It was observed that 519 miRNAs from constitutive inter LADs (ciLAD) showed expression in MEC(N), 537 in MEC(P) and 453 in MEC(Prl). Among these miRNAs, 29 are from MEC(N) to MEC(P) and 4 from MEC(P) to MEC(Prl). Further analysis also revealed that 20 were downregulated from MEC(N) to MEC(P) condition (Fig. C5). A set of miRNAs from ciLADs were well characterized based upon the expression in basal or luminal cells of mammary tumours. Mir-155, miR-135b and miR-505 were reported in basal type specific tumour expression.

Similarly, miR-100, miR-130, miR-29b, miR-152, Let-7a, and Let-7f were reported in luminal type specific tumour expression. The expression of miR-486, miR-148a,

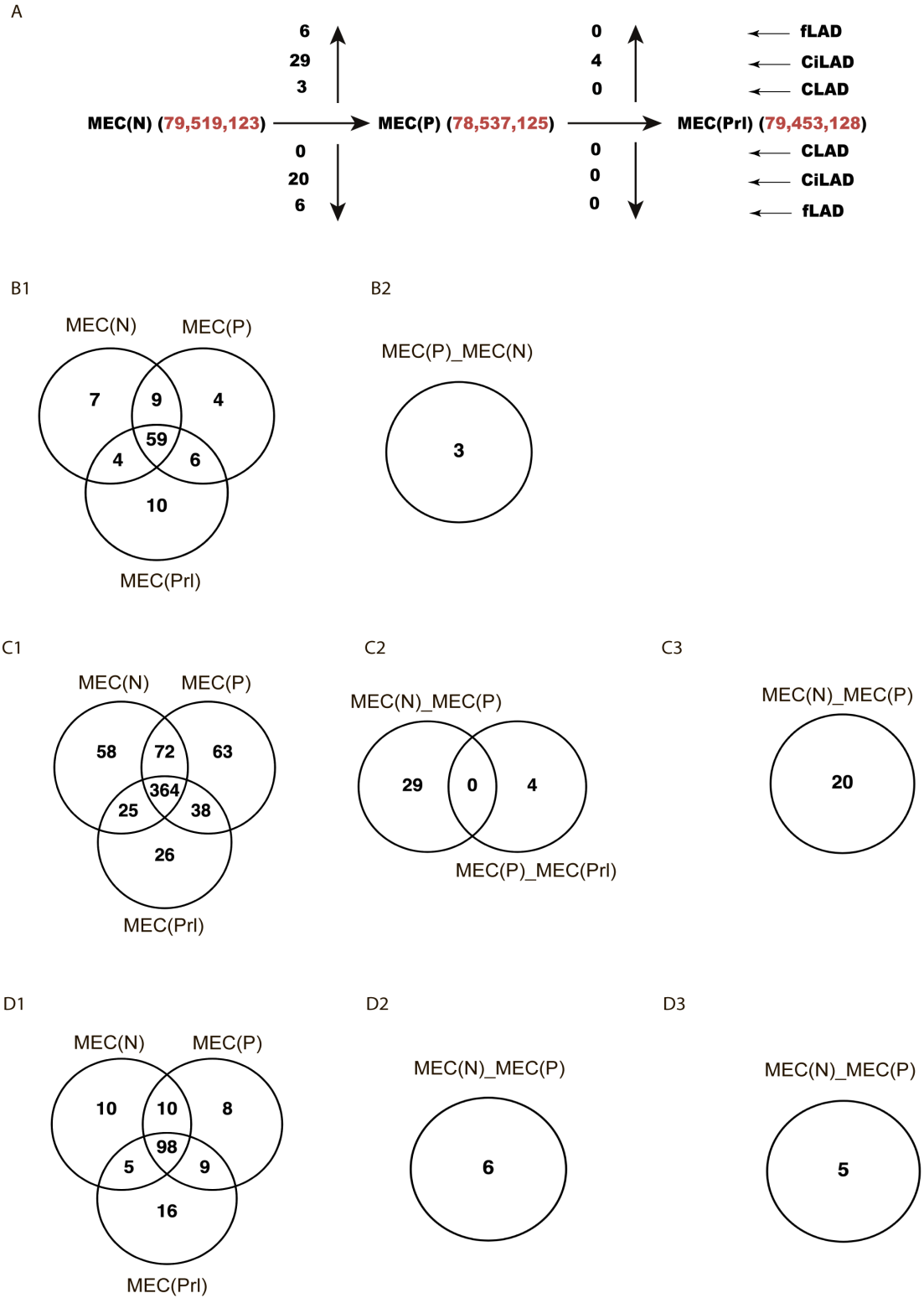


Fig.C5. Schematic representation of statistically analysed miRNAseq dataset of HC11 MECs differentiation to figure out miRNA's genes present in the LAD region of chromatin: A. Flow chart

representing expressed and differentially regulated miRNAs' genes present in CLAD, CiLAD, and fLAD in MEC (N), MEC (P), MEC (Prl) stages of lactogenesis. Venn diagram representing **B1.** Expressed unique and overlapped miRNA's (NC≥10) genes present in CLAD in MEC (N), MEC (P), MEC (Prl), **B2.** unique and overlapped differentially upregulated miRNA's (Log2 Fold change ≥ 1) genes present in CLAD in between MEC(P) vs MES(N) and MEC(Prl) vs MES(P). **C1.** Expressed unique and overlapped miRNA's (NC≥10) genes present in CiLAD in MEC (N), MEC (P), MEC (Prl), C**2.** unique and overlapped differentially upregulated miRNA's (Log2 Fold change ≥ 1) genes present in CiLAD in between MEC(P) vs MES(N) and MEC(Prl) vs MES(P), **C3.** unique and overlapped differentially downregulated (Log2 Fold change ≤ -1) miRNA's genes present in CiLAD in between MEC(P) vs MES(N) and MEC(Prl) vs MEC(P). **D1.** Expressed unique and overlapped miRNA's (NC≥10) genes present in fLAD in MEC (N), MEC (P), MEC (Prl), **D2.** unique and overlapped differentially upregulated miRNA's (Log2 Fold change ≥ 1) genes present in fLAD in between MEC(P) vs MES(N) and MEC(Prl) vs MES(P), **D3.** unique and overlapped differentially downregulated (Log2 Fold change ≤ -1) miRNA's genes present in fLAD in between MEC(P) vs MES(N) and MEC(Prl) vs MEC(P).

miR-10b, miR-199a, and miR-150 are expressed only in normal mammary epithelium cells (Zhu et al., 2011). In case of facultative LADs (fLAD), 123 in MEC(N), 125 on MEC(P) and 128 in MEC(Prl). Among them 6 were upregulated from MEC(N) to MEC(P). Further analysis revealed that 5 were downregulated from MEC(N) to MEC(P) (Fig. C5). Mmu-mir-206 is present in fLAD region of the chromatin which showed high expression in MEC(N) but, gradual reduction upon differentiation. Mmu-mir-206 is mainly responsible for the arrest of epithelial to mesenchymal transition and progression of G1-S cell cycle. Because of which, this miRNA can able to reduce colony formation during breast tumour progression (Wang et al., 2019). Another mirNA miR-184 belongs to fLAD, is having anti-tumerogenic properties (Phua et al., 2015). Among these above-mentioned miRNAs, the expression status of few of the miRNAs those are involved in cLADs (Table 72), ciLADs (Table 73) and fLADs (Table 74) were listed out.

Stage specific miRNA-mRNA interactome analaysis:

MEC (N) stage after reaching confluency are mostly arrested at GO/G1 phase of cell cycle (sornapudi et et al., 2018). Dissecting regulatory principles of gene expression at this state is very important as dysregulation at this point leads to failure of lactogenesis or might skew developmental program towards carcinogenesis (Wang et al., 2019). To dissect microRNA mediated gene regulation, interactome map of expressed mRNAs and miRNAs in MEC (N) state were analysed by using miRNet tool which considers experimentally validated miRNAs' targets

for predicting targets (Chang et al., 2020) (Fig. C6. A). Genes in MEC (N) stage were found to be regulated by miRNAs such as, mmu-mir-340-5p, mmu-mir-9-5p, mmu-mir-181a-5p, mmu-mir-26a-5p, mmu-mir-149-5p, mmu-mir-24-3p, mmu-let-7-5p etc. Many of these miRNAs are predicated to be involved in the regulation of pathways such as, FoxO signalling pathway, Focal adhesion, Hippo signalling pathway, HIF-1 signalling pathway, TNF signalling pathway, ErbB signalling pathway, PI3K-Akt signalling pathway, AMPK signalling pathway, mTOR signalling pathway, Wnt signaling pathway, etc. These pathways are known to play important role in cellular homeostasis and play important role in metabolic activity, growth and development.

miRNA-mRNA interactome analysis for MEC(P) cells by using miRNet tool resulted in some interesting predictions. (Fig. C6. B). Major miRNAs' hub towards MEC(P) transition were found to be centred around microRNAs such as mmu-mir-340-5p, mmu-mir-9-5p, mmu-mir-181a-5p, mmu-mir-26a-5p, mmu-mir-149-5p, mmu-mir-425-5p etc. These miRNAs in MEC (P) stage are predicted to regulate expression of genes involved in pathways such as, FoxO signalling pathway, Focal adhesion, HIF-1 signalling pathway, MAPK signalling pathway, Hippo signalling pathway, ErbB signalling pathway, , PI3K-Akt signalling pathway, mTOR signalling pathway, TNF signalling pathway, AMPK signalling pathway, Insulin signalling pathway etc.

To understand the role of MEC(Prl) signalling in driving lactogenic differentiation of MECs through miRNA mediated gene regulation, expressed miRNA-mRNA interactome map have been generated and analysed (Fig. C6. C). Major miRNAs' hubs that were identified in MEC(Prl) stages were found to be derived from mmu-mir-9-5p, mmu-mir-181a-5p, mmu-mir-26a-5p. It has already been established that mmu-mir-26a-5p is highly upregulated during lactogenesis and play key role in lipidogenesis in rodent liver which is essential for lactogenesis (Wang et al., 2016) (Fig. C6. C'). These miRNAs in MEC (Prl) stages were also found to be involved in regulation of pathways such as, FoxO signalling pathway, Focal adhesion, HIF-1 signalling pathway, MAPK signalling pathway, , Hippo signalling pathway, mTOR signalling pathway, AMPK signalling pathway, PI3K-Akt signalling pathway, Insulin signalling pathway, , MEC(Prl) signalling pathway, cell cycle, etc.

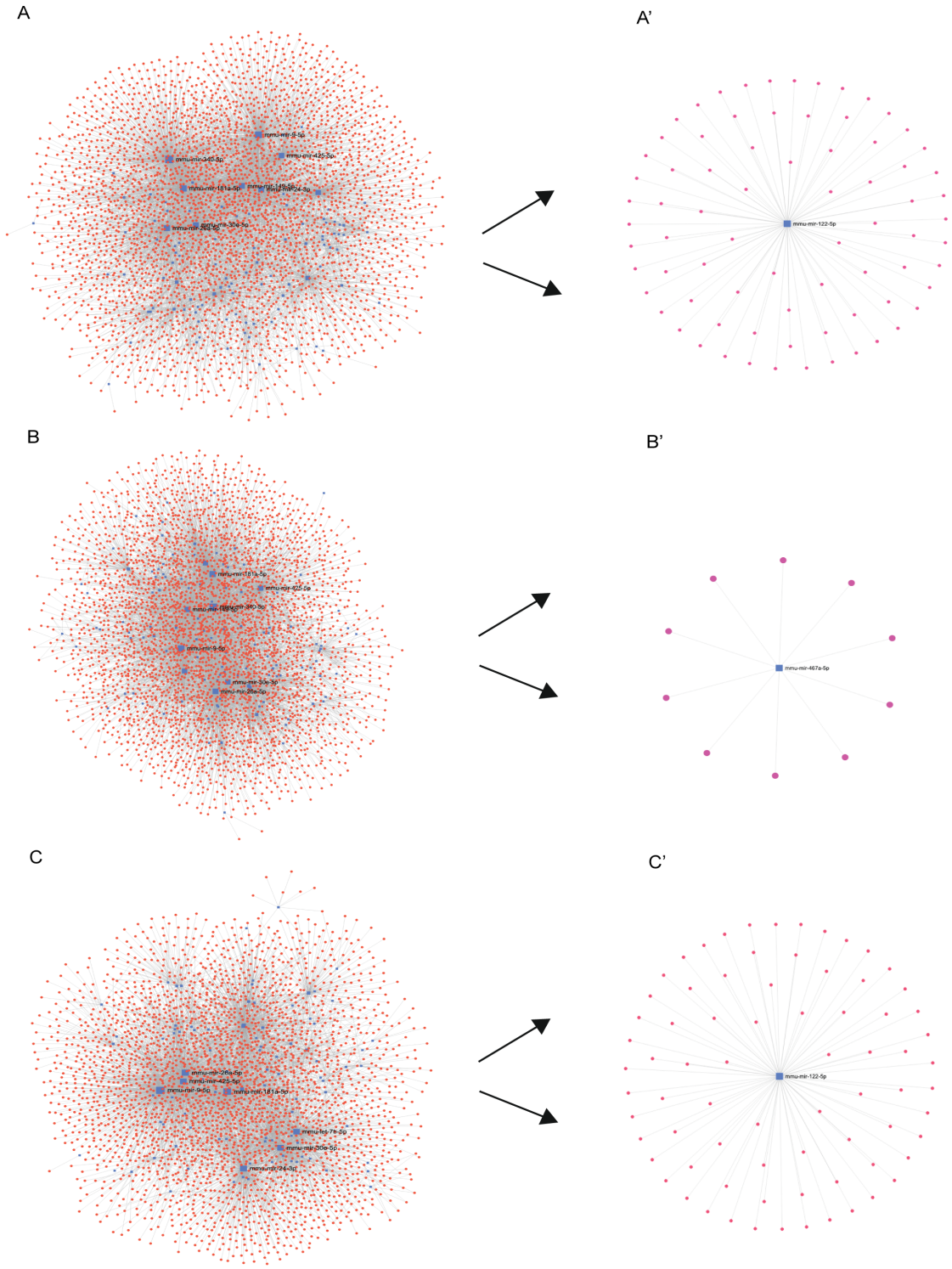


Fig. C6. miRNA-mRNA network map of expressed miRNAs and its experimentally validated mRNAs targets in mammary epithelial cells during lactogenesis: A. expressed mRNA (FPKM≥1) with expressed miRNA (NC≥10) of MEC(N), **A'.** Zoom in view of mmu-mir-122-5p, one of the important miRNA's hub from A, **B.** expressed mRNA (FPKM≥1) with expressed miRNA (NC≥10) of MEC(P), **B'.** Zoom in view of mmu-mir-467a-5p, one of the important miRNA's hub from B, **C.** expressed mRNA (FPKM≥1) with expressed miRNA (NC≥10) of MEC(Prl), **C'.** Zoom in view of mmu-mir-122-5p, one of the important miRNA's hub from C.

Further to understand the regulation of mRNAs by miRNAs during developmental transition of MEC(N) to (P) stages, dynamic expression profiles of miRNAs and mRNAs between these stages were analysed, keeping in mind the inverse functional co-relation miRNA and mRNA data sets have. Interactome maps were generated for down regulated miRNAs and upregulated mRNAs between MEC(N) to (P) transition (Fig. C7. B). Downregulated microRNAs which might be controlling MEC(P) specific mRNAs in MEC(N) stage were found to be mmu-mir-340-5p, mmu-mir-19b-3p, mmu-mir-30e-5p, mmu-mir-292a-5p etc. Down regulation of these microRNAs was found to be inversely corelated with upregulation of genes involved in the pathways such as

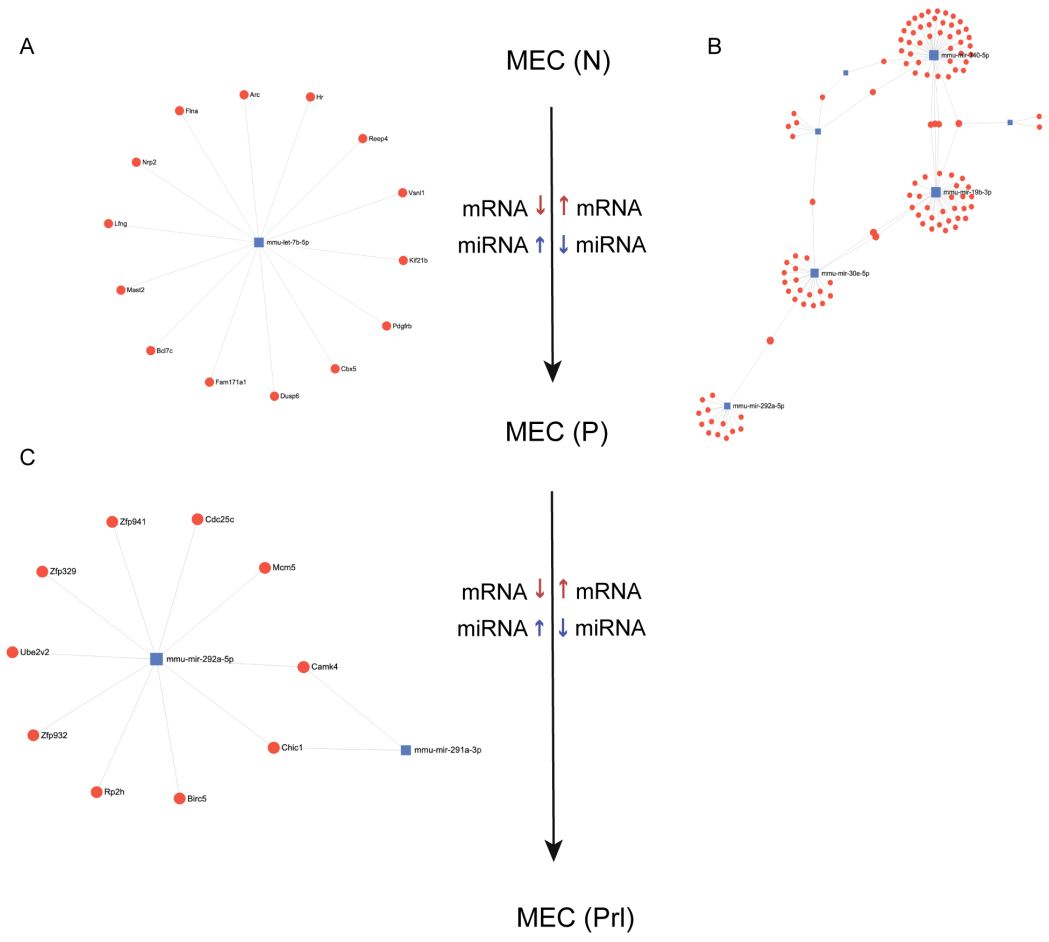


Fig. C7. Study of differentially regulated miRNAs and its experimentally validated mRNAs targets through miRNA-mRNA network map of mammary epithelial cells during lactogenesis: A. downregulated mRNA (Log2 Fold change ≤ -1) with upregulated miRNA (Log2 Fold change ≥ 1) from MEC(P) vs MEC(N), **B.** upregulated mRNA (Log2 Fold change ≥ 1) with downregulated miRNA (Log2 Fold change ≤ -1) from MEC(P) vs MEC(N), **C.** downregulated mRNA (Log2 Fold change ≤ -1) with upregulated miRNA (Log2 Fold change ≥ 1) from MEC(P) vs MEC(N).

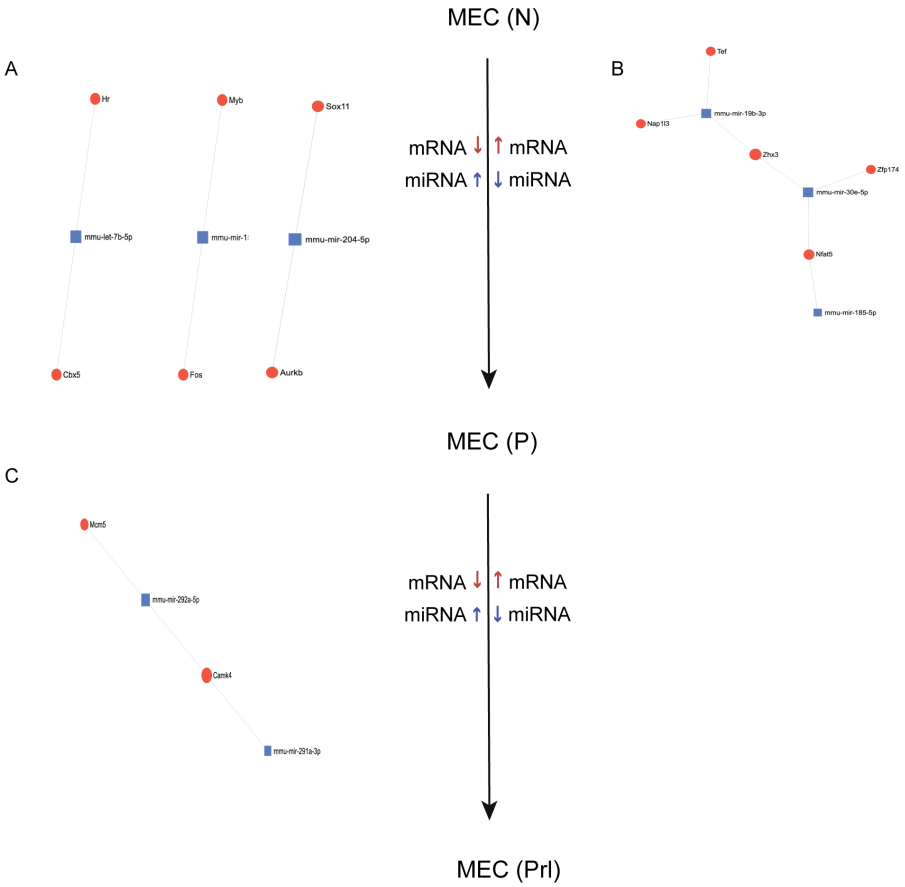


Fig. C8. Study of differentially regulated miRNAs and its experimentally validated TFs targets through miRNA-mRNA network map of mammary epithelial cells during lactogenesis: A. downregulated TF (Log2 Fold change ≤ -1) with upregulated miRNA (Log2 Fold change ≥ 1) from MEC(P) vs MEC(N), **B.** upregulated TF (Log2 Fold change ≥ 1) with downregulated miRNA (Log2 Fold change ≤ -1) from MEC(P) vs MEC(N), **C.** downregulated TF (Log2 Fold change ≤ -1) with upregulated miRNA (Log2 Fold change ≥ 1) from MEC(P) vs MEC(N).

MEC(Prl) signalling pathway, Glycerophospholipid metabolism, FoxO signalling pathway etc. Here, Prolactin signalling pathway started showing upregulation from MEC (P) state probably due to addition to glucocorticoids. When downregulated miRNAs during MEC(N) to (P) transitions were specifically analysed with filtered upregulated TFs (Fig. C8. B); revealed upregulation of very specific TF targets of miRNAs instead of a complex transcriptome network. Downregulation of mmu-miR-19b-3p caused upregulation of *Tef* which is involved in Hippo signalling pathway and is found to be essential during pregnancy (Chen et al., 2014). *Nap1l3* and *Zhx3*. *Zhx3* is predicted to be targeted by mmu-miR-30e-5p that even targets *Zfp174*. *Nfat5* is predicted to be targeted by both mmu-miR-30e-5p and mmu-miR-185-5p.

Interaction maps of downregulated mRNAs and upregulated miRNAs between MEC (N) to MEC (P) stages (Fig. C7. A) revealed major hubs of miRNAs, like mmu-let-7b-5p, mmu-mir-155-5p, mmu-mir-126a-3p, etc. Further analysis of miRNA-TF interaction map of upregulated miRNAs and downregulated TFs from MEC (N) to MEC (P) transition also revealed very specific targets of miRNAs (Fig. C8. A). Mmu-let-7b-5p is predicted to regulate *Hr* and *Cbx5*. Mmu-miR-1 is predicted to target transcription factors *Myb* and *Fos* and causing its downregulation. Regulation of *Myb* is essential to control tumorigenesis during mammary gland development (Miao et al., 2011) and *Fos* which is required during the process of involution seems to be suppressed during pregnancy (Jaggi et al., 1995). Further, mmu-miR-204-5p is predicted to be involved in regulation of *Sox11* and *Aurkb*. Sox11 is an embryonic mammary marker and remains silent in postnatal development of mammary gland (Umeh-Garcia et al., 2020) which is controlled by mmu-miR-204-5p.

In order to understand the significance of upregulated miRNAs in MEC (Prl), miRNA-mRNA interaction map of downregulated mRNAs and upregulated miRNAs from MEC (P) to MEC (Prl) stages were analysed (Fig. C7. C). This analysis showed that thee major miRNAs' hubs centred around multiple mRNAs such as mmu-let-7b-5p, mmu-mir-499-5p, mmu-mir-291a-5p, mmu-mir-126a-3p, mmu-mir-1a-3p, etc. Similarly, analysis of upregulated miRNA and downregulated TFs from MEC (P) to MEC (Prl) revealed possible controlling of *Mcm5* and *Camk4* due to upregulation of mmu-miR-292a-5p and mmu-miR-291a-3p respectively (Fig. C8. C).

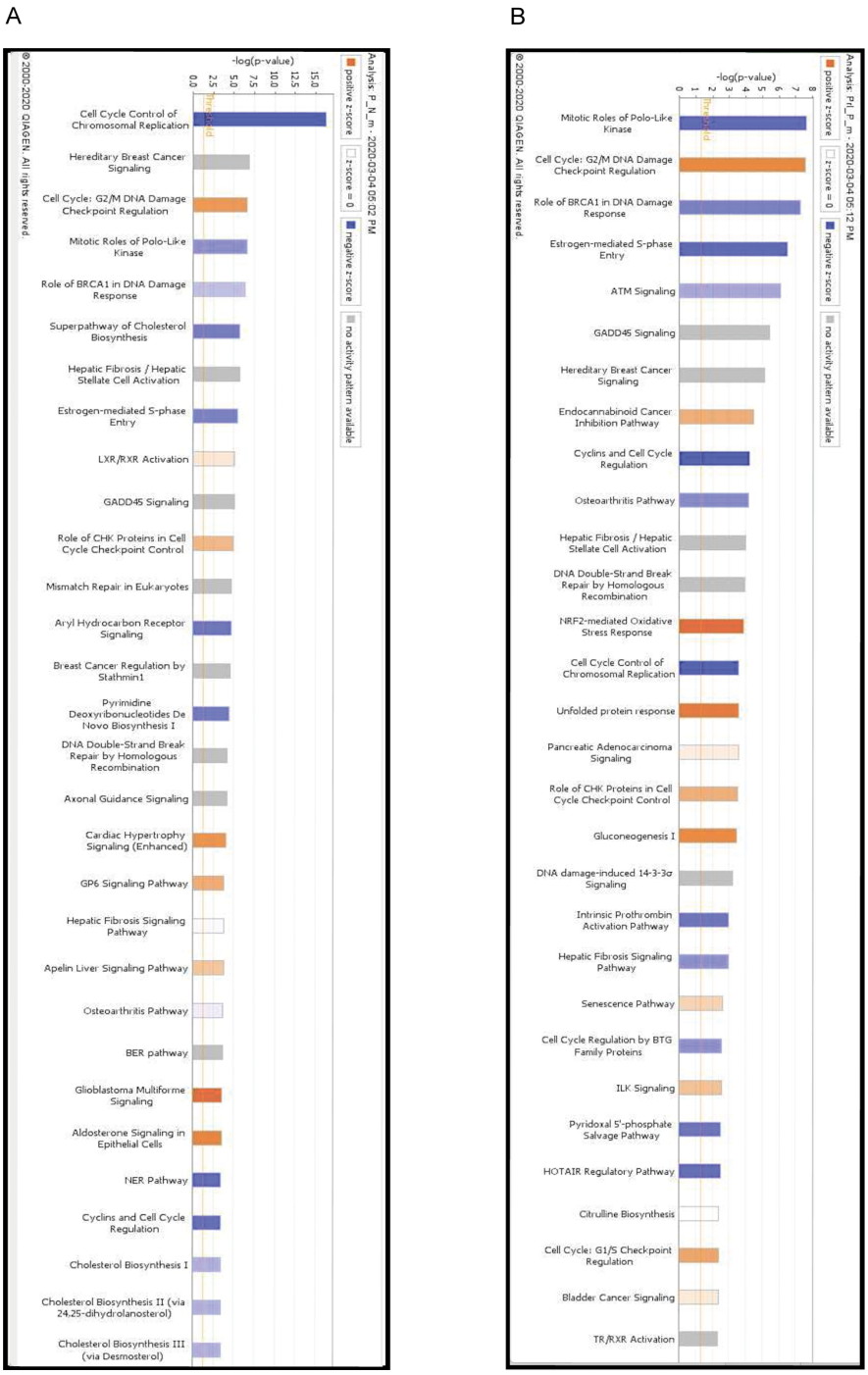


Fig.C9. Pathways prediction during Lactogenic differentiation of mouse mammary epithelial cells by using Qiagen-IPA analysis: A. Qiagen-IPA pathways in between MEC(P) vs MEC(N) stages by providing differentially regulated miRNAs (both downregulated; Log2 fold change≤-1 and upregulated; Log2 fold change ≥ 1) and mRNAs (both downregulated; Log2 fold change≤-1 and upregulated; Log2 fold change ≥ 1) together as an input. B. Qiagen-IPA pathways in between MEC(Prl) vs MEC(P) stages by providing differentially regulated miRNAs (both downregulated; Log2 fold change≤-1 and upregulated; Log2 fold change ≥ 1) and mRNAs (both downregulated; Log2 fold change≤-1 and upregulated; Log2 fold change ≥ 1) together as an input. Full list of Qiagen Ingenuity pathways are available in Table 75-76.

miRNA-mRNA interactome analysis reveals functionally linked biological pathways:

Differentially regulated mRNA and miRNA gene lists were given as an input to Qiagen IPA software (Fig. C9). It provided list of pathways and associated diseases during each stage of differentiated mammary epithelial cells. MEC (N) condition is enriched with pathways like Epithelial adherens junction signalling and Wnt/ß-catenin Signalling which are essential for rapid cell proliferation and differentiation. MEC (P) condition upon Glucocorticoid signalling is enriched with pathways related to cell cycle, DNA damage, Cell cycle: G2/M DNA damage checkpoint regulation, Mitotic roles of polo-like Kinase, Role of BRCA1 in DNA damage response, Estrogen-mediated-S-phase entry, Role of CHK proteins in cell cycle checkpoint control, Mismatch repair in eukaryotes, Breast cancer regulation by Stathmin1, DNA double-strand break repair by homologous recombination, and Cyclins and cell cycle regulation (Table 75). Similarly, Prolactin addition to MEC (P) cells showed enrichment of majorly cell cycle regulation pathways. Among top upregulated pathways are Cell cycle: G2/M DNA damage checkpoint regulation, Estrogen-mediated-S-phase entry, Cyclins and cell cycle regulation, Cell cycle regulation by BTG family proteins, Cell cycle: G1/S checkpoint regulation (Table 76). These above results imply major regulation in terms of cell cycle control during mouse mammary epithelial stem like cells, HC11 differentiation.

miRNA-mRNA and KEGG pathway integrative analysis reveals functional significance of miRNA mediated regulation of mRNAs during lactogenic differentiation of MECs:

To understand the functional state of MECs during their lactogenic differentiation, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were derived by integrative mapping of miRNAs and its targeted mRNAs (Fig. C10). It was noted that a few miRNAs such as mmu-mir-9-5p, mmu-mir-181a-5p, mmu-mir-26a-5p, mmu-mir-24-3p, mmu-mir-292a-5p, etc were commonly involved in majority of the regulatory pathways. Of which, upregulated upon glucocorticoid treatment were mRNAs encoding adherents junctions, which were shown to supports pubertal development of mammary gland and helps in survival of epithelial cells during lactation (Shamir et al., 2015). Upregulation of MAPK signalling pathway was shown to be vital for cell survival as its inhibition was shown to promote apoptosis (Healy et al. 2000). MAPK

signalling pathway specific mRNAs in MEC(P) and MEC(Prl) stages were predicted to be regulated by downregulation of miRNAs such as mmu-let-7c-5p, mmu-let-7d-5p, mmu-mir-101b-3p, mmu-

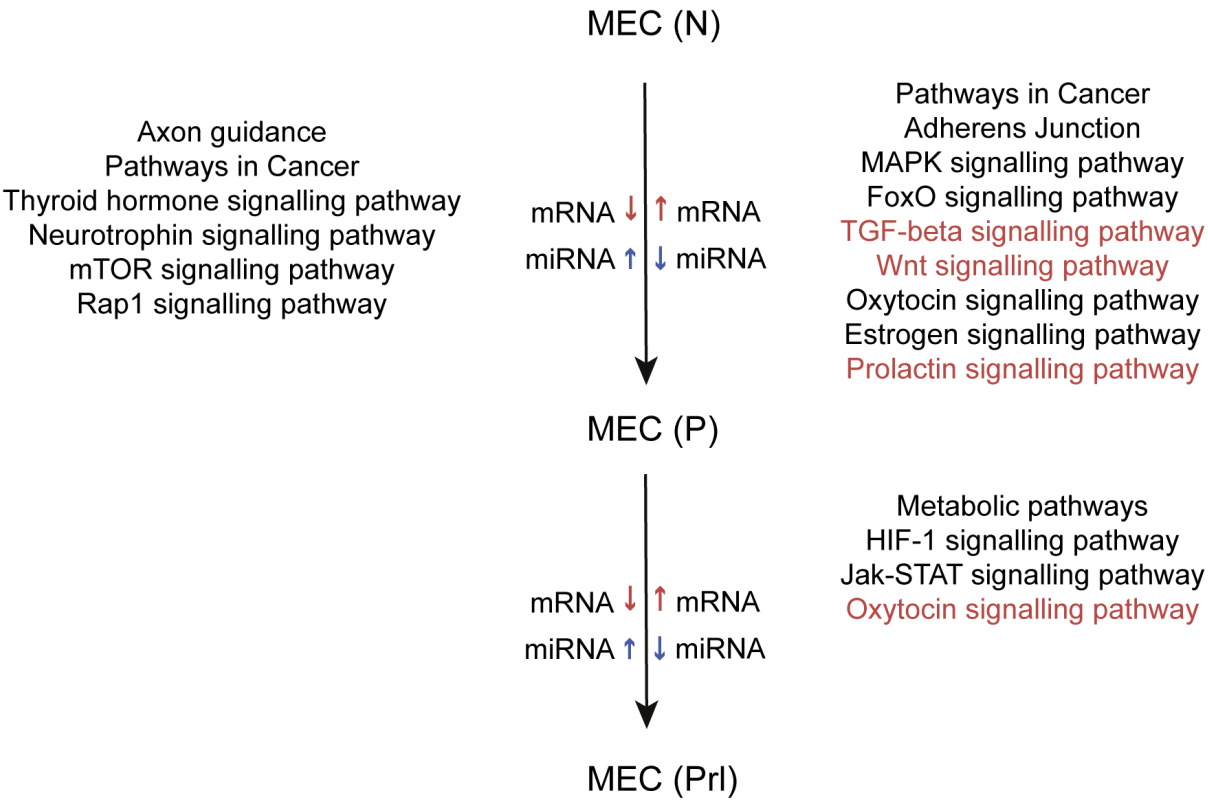


Fig.C10. Pathways prediction during Lactogenic differentiation of mouse mammary epithelial cells by using KEGG pathway analysis: Flow chart representing KEGG pathways in between stages by considering differentially regulated miRNAs and its filtered targets from differentially regulated mRNAs during MEC (Prl) differentiation.

Fig.C11. Study of pathway specific microRNA-mRNA interactome map of highly enriched KEGG pathways during MEC(Prl) differentiation: Pathway specific miRNA-mRNA interaction networks during integration of expressed miRNA (NC ≥ 10) and expressed mRNAs (FPKM ≥ 1) in MEC (Prl) stage. provides detail of **A.** MAPK signalling pathway, **B.** Oxytocin signalling pathway and **C.** Wnt signalling pathway during Lactogenesis.

mir-200b-3p, miR-30 family, etc. and those that promote upregulation of *Dusp7* and *Ntrk2* (Fig. C11. A). Upregulation of mRNAs involved in FoxO signalling also play critical roles in maintenance of mammary stem cell homeostasis (Sreekumar et al., 2017). Our integrative analysis has shown that upregulation of mRNAs related to FoxO signalling were correlated with downregulation of miRNAs such as mmu-mir-19b-3p, mmu-mir-340-5p, mmu-mir-292a-5p and mmu-mir-30e-5p upon glucocorticoid treatment which are predicted to control Foxo3, Homer2 and Pik3r1 in MEC(N) condition. TGF-ß signalling pathway control mammary epithelial cells proliferation which was noted to be upregulated during developmental transition from MEC(N) to (P) stages owing to the initiation of differentiation process upon cell cycle arrest (Wakefield et al., 2001). Wnt signalling pathway is critical for mammary gland growth, differentiation and involution (Turashvill et al. 2006) and was found to upregulated in MEC (P) and MEC (Prl) RNA-seq datasets. In MEC (Prl) stage, mmu-mir-27a-3p, mmu-mir-27b-3p, mmu-mir-10a-5p, mmu-mir-122-5p, etc. were predicted/identified to be involved in regulation of Wnt signalling specific mRNAs. Estrogen signalling play an important role in development of mammary gland at pubertal stage by facilitating ductal morphogenesis (LaMarca et al. 2007) as predicted to be regulated by mmu-miR-340-5p at MEC (P) stage (Fig. C11. C). mTOR signalling showed upregulation upon Prolactin treatment and is essential for epithelial cell proliferation and differentiation (Jankiewicz et al., 2006). Upon inhibition of mTOR, reduction in expression of milk protein in cultured cells had been observed. On the other hand, Prolactin signalling pathway gets upregulated upon glucocorticoid and Prolactin treatment which activates Jak-STAT pathway. Jak-STAT pathway modulates temporal expression of many transcripts required for lactogenic differentiation (Lavnilovitch et al., 2002). Prolactin signalling pathway upregulation, was correlated with downregulation of mmu-mir-19b-3p, mmu-mir-340-5p and mmu-let-7a-5p during MEC (Prl) condition. At the end, a fully functional mammary gland requires Oxytocin signalling pathway to eject milk by contraction of myoepithelial cells (Lollivier et al., 2006) which was potentially modulated by expression of mmu-mir-1198-5p, mmu-mir-381-3p, mmu-mir-146a-5p, mmu-mir-291a-3p, mmu-mir-143-3p, mmu-mir-103-3p, mmu-mir-122-5p, etc (Fig. C11. B).

Elucidating authentic miRNA-mRNA interacting partners during lactogenesis by CLASH-seq method:

Computational prediction of miRNA targeted mRNAs alone does not guarantee their function *in vivo*. To authenticate such interactions, we performed CLASH-seq method (Helwak et al., 2014).

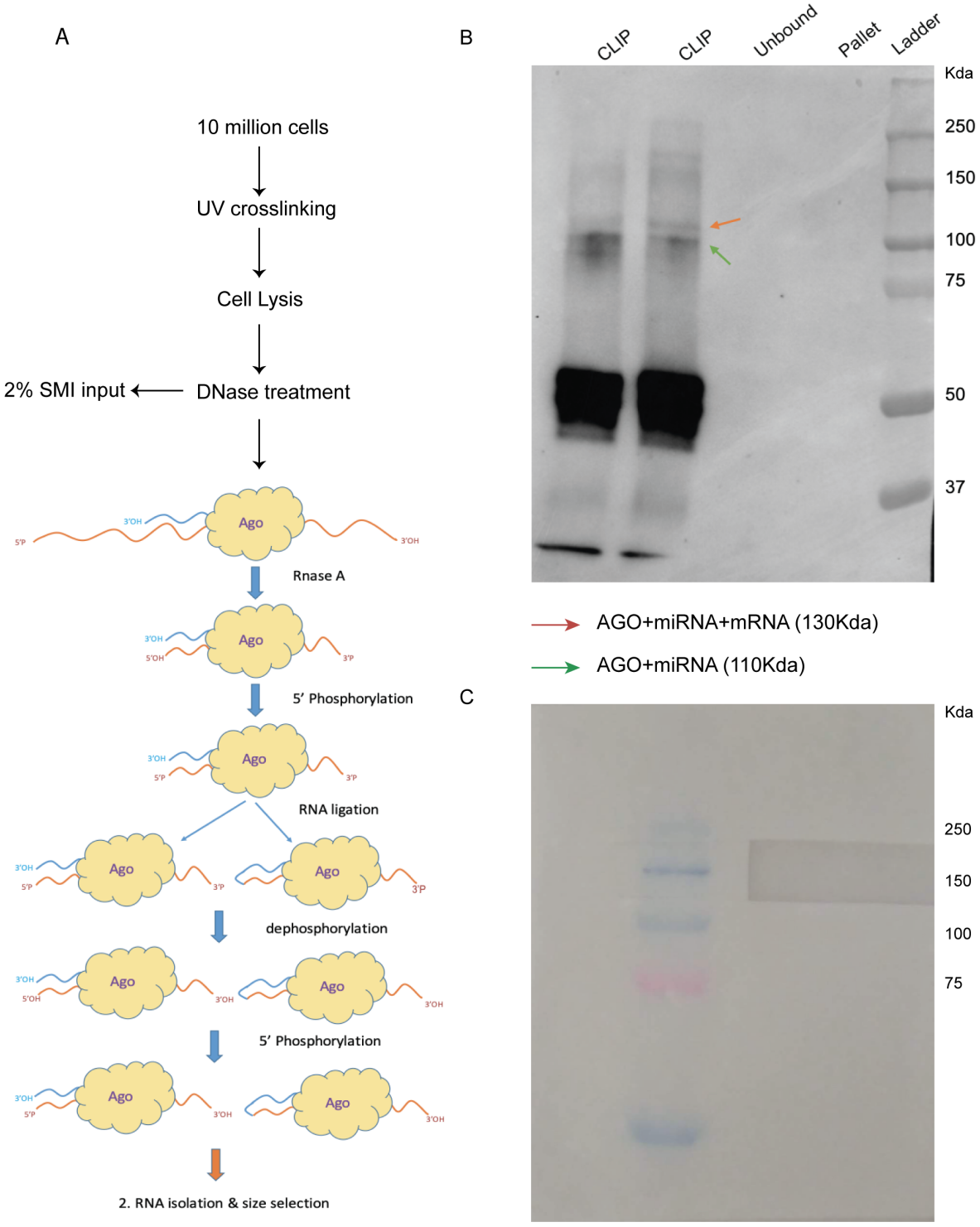


Fig.C12. Cross linking, ligation and sequencing of Hybrids: A. Work flow of CLASH-seq. protocol by considering 10 million cells as input. **B.** Confirmation of AGO complex at 130 Kda (AGO+miRNA+mRNA) and 110Kda (AGO+miRNA) through western blot by using anti-AGO antibody at 1;1000 dilution and mouse secondary antibody at 1:5000 dilutions. **C.** Representation of size selection of AGO complex from a nitrocellulose membrane.

This method relies upon the fact that all the physiologically engaged miRNA-mRNA along with RNA inducing silencing complex (RISC) can be immune-precipitated with the antibodies that specifically recognizes one of the components of RISC complex i.e. Argonaut protein. Immunoprecipitation with Ago antibody followed by ligation allow ligation of physically proximal mRNA and miRNAs resulting in chimeric ligation product which is later characterized by high throughput sequencing. CLASH sequencing of HC11 MECs undergoing lactogenic differentiation was performed. 10 million cells from each biological replicates of MEC (N), MEC (P) and MEC (Prl) were subjected to UV crosslinking, followed by lysis and were then immunoprecipitated with anti-AGO antibody followed by RNA-RNA ligation by T4 RNA ligase I. Ligated chimeric RNAs with RISC complex were subjected to isolation of RNA, reverse transcription and high throughput sequencing using sequencing platform, Illumina Hiseq 2500 (Fig. C12). For each sample around 10 million reads were generated. To validate the immunoprecipitation, one of the highly expressed miRNA in MEC (N) stages such as mmu-mir-122-5p presence was taken in to consideration. The microRNA show high expression during MEC (N) condition but is downregulated upon differentiation towards MEC(P) and MEC(Prl) stages of differentiation. Both biological replicates of MEC (N) showed a presence of mmu-mir-122-5p but it was absent in MEC (P), MEC (Prl) stages (Fig. C13. B). CLASH sequencing datasets were analysed by using hyb and CIMS algorithm (Zhang et al., 2012) which relies upon the fact of UV mediated deletions. Hyb algorithm provides miRNA-mRNA chimeric reads from the sequencing reads and were found to represent below 1% of total reads (Fig. C14. A).

Due to mapped reads containing hybrids of miRNA-mRNAs, these chimeras were identified by using *Hyb* algorithm. *Hyb* recognizes chimeras based on non-contiguous mapped reads to genome and contiguous reads had been discarded. Finally, it was used to predict folding patterns computationally to infer true RNA-RNA interaction molecules. *Hyb* extracted 13 chimeras in MEC (N), 34 in MEC (P) and 8 in MEC (Prl) state, among which one chimera exhibits in all three state and two are common in MEC (N) and MEC (P) condition (Fig. C14. A, D). Due to lower RNA-RNA ligation efficiency, CLASH-seq is limited to provide enough chimeras. It has been reported that total number of chimeras recovered through this is below 2% of total sequenced reads. Because of this limitation, reanalysis of discarded contiguous reads by *Hyb* through Cross-linked Induced Mutation (CIMS) algorithm was performed. CIMS detects mutation near to protein binding sites that were caused by UV cross-linking. In case of AGO, deletion has been observed majorly on RNA strands.

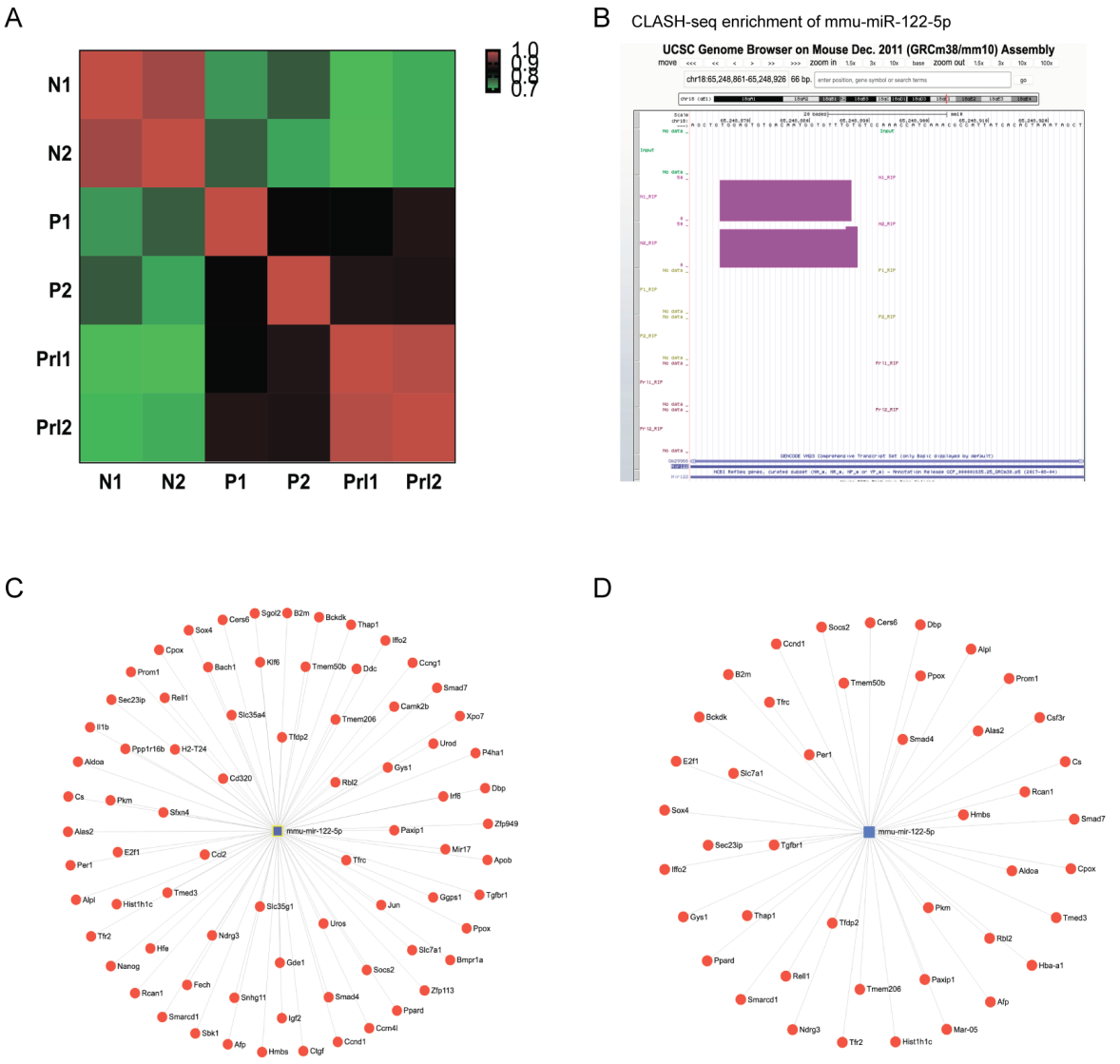


Fig. C13. Quality control check and validation of CLASH-seq sample's enrichment: A. Heatmap representing CLASH-seq Spearman's correlation map between biological replicates and different stages of MEC (N), MEC (P), MEC (Prl). **B.** Screenshot window from UCSC genome browser showing confirmation of immunoprecipitation of mmu-miR-122-5p in MEC (N). **C.** Total 76 expressed mRNAs targets of mmu-miR-122-5p in MEC (N). **D.** Total 45 expressed mRNAs targets of mmu-miR-122-5p in MEC (N) from CLASH-seq.

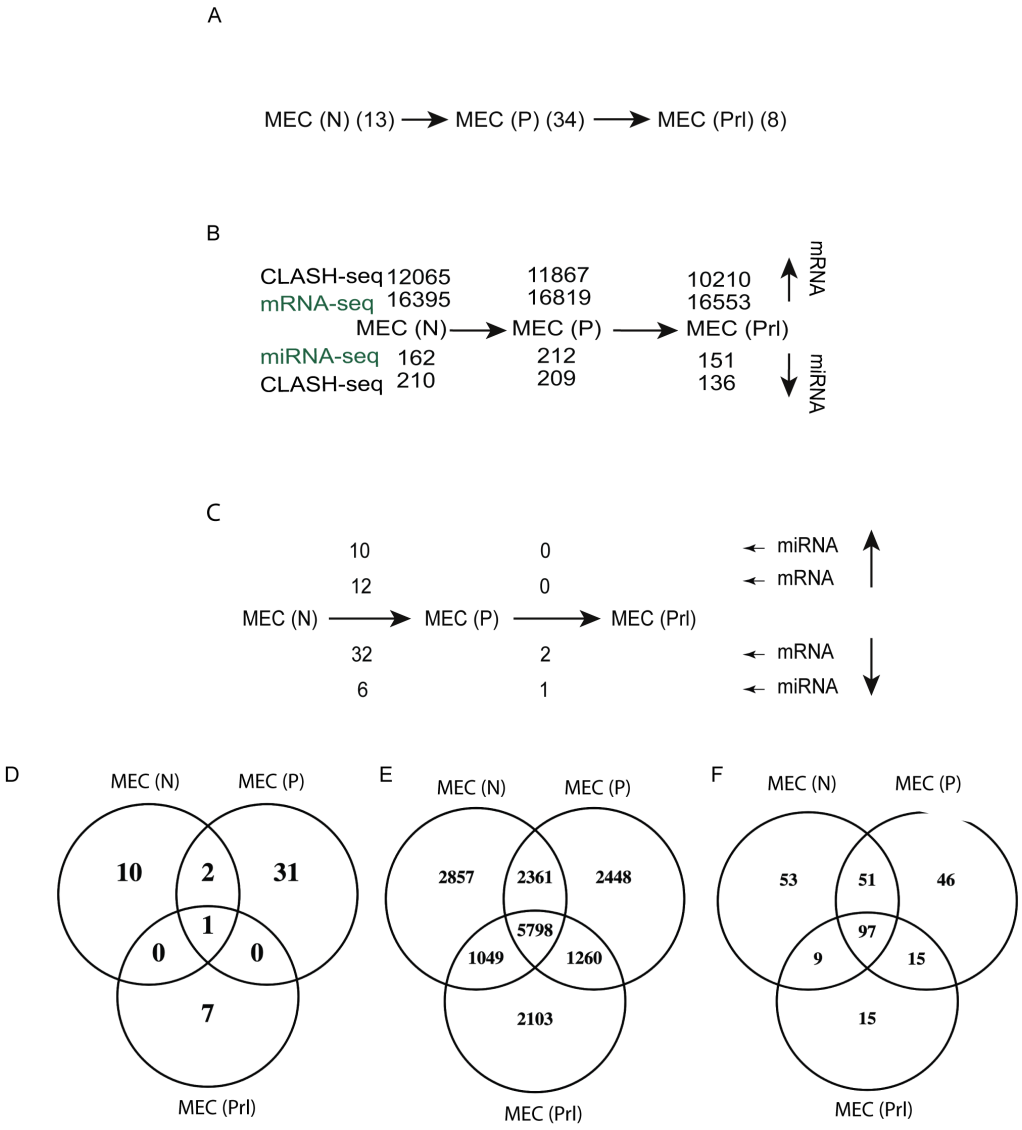


Fig. C14. Schematic representation of statistically analysed CLASHseq dataset of HC11 MECs differentiation to figure out miRNA's genes and its putative targets: A. Flow chart representing miRNA-mRNA chimeras in MEC (N), MEC (P), MEC (Prl). **B.** Flow chart representing expressed mRNAs and miRNAs from miRNA-seq, mRNA-seq and CLASH-seq in MEC (N), MEC (P), MEC (Prl). **C**. Flow chart representing differentially regulated miRNAs and mRNAs from CLASH-seq experiment. Venn diagram representing **D.** mRNA-miRNA chimeras through hyb analysis of CLASH-seq, **E.** expressed (FPKM≥1) mRNAs and **F.** expressed (NC≥1) miRNAs in MEC (N), MEC (P), MEC (Prl) from CLASH-seq.

Further, CLASH -seq procedure can be used to decipher the authentic binding of Ago complexed cognate mRNA by the fact that UV crosslinking promotes nucleotide deletions, near to AGO binding sites. These sites were detected from CLASH seq data by using CIMS algorithm (Zhang et al., 2012) which relies on the UV mediated mutations. Further, deletion of nucleotides on mRNAs should correspond to the seed region of miRNAs which would confirm the authentic target of a given miRNA. Some genes showed multiple AGO binding sites as evident from multiple deletions in a given mRNA and some showed single deletion sites. It should be however, noted that some genes might have enriched for multiple or single AGO binding sites but might not show any deletions due to the fact that during reverse transcription, reverse transcriptase enzyme might have added the wrong or correct nucleotides. However, occurrence of AGO binding sites would confirm the presence of the seed region of miRNAs. But genes without deletion and without any AGO binding sites were considered negative. The seed region of miRNA near to deletion or AGO binding sites were considered positive for one nucleotide mismatch, which was either a wobbled or bulged base pair.

CLASH-seq can determine RNA-RNA ligation which exhibits either coding-coding, coding-noncoding or noncoding-noncoding interaction in higher eukaryotes. It can be also being used to predict miRNA-mRNA interactions that are bounded by AGO. Anti-AGO antibody was used for pulldown and immunoprecipitation which was confirmed with western blot. In western blot, clearly it has given two bands at 110 and 130kDa. 110kDa refers to AGO-miRNA complex and 130kDa refers to AGO-miRNA-mRNA complex (Fig. C20. c). Extracted RNA from these two regions with input sample were sequenced up to 10million reads in average. Sequenced reads showed 99% Q20 with 50% GC content. Correlation between biological replicates were 0.9 with p-value < 2.2e-16 (Fig. C13. A). Target enrichment was confirmed by analysing mapped bam files with UCSC genome browser (Fig. C13. B).

CLASH-seq analysis by using Cufflinks and miRDeep2 algorithm. MEC (N), showed up 12,065 mRNAs had been pulled down through immunoprecipitation compared to 16,396 expressed mRNA from mRNA-seq. Similarly, in MEC (P) and MEC (Prl), out of 16,819 and 16,553 expressed mRNAs 11,867 and 10,210 had been captured by antibody pulldown respectively. This indicates approx. 60% expressed mRNAs had precipitated through anti-AGO pulldown. But, in case of miRNAs, the average number of expressed miRNAs are almost equal to the number of precipitated miRNAs. Out of 162, 212 and 151 expressed miRNAs in MEC (N), MEC (P) and MEC (Prl), 210, 209 and 136 miRNAs were getting pulled down (Fig. C14. B). 5,798 expressed

mRNAs and 97 expressed miRNAs are common in all three conditions (Fig. C14. E-F). Even CLASH-seq data was further analysed by using DESeq2 algorithm which provided information about differentially expressed mRNAs and miRNAs those were pulled down in CLASH. Among 1141 upregulated and 1026 downregulated mRNAs from MEC (N) to MEC (P) transition only 12 and 32 came up in CLASH experiment respectively. Similarly, among 385 upregulated and 655 downregulated mRNAs from MEC (P) to MEC (Prl) transition 0 and only 2 came up in CLASH experiment respectively. Coming to comparison of miRNAs list of miRNA-seq and CLASH-seq, among 51 upregulated and 29 downregulated mRNAs from MEC (N) to MEC (P) transition only 10 and 6 came up in CLASH experiment respectively. Similarly, among 385 upregulated and 655 downregulated mRNAs from MEC (P) to MEC (Prl) transition 0 and only 2 came up in CLASH experiment respectively. Similarly, among 4 upregulated and 0 downregulated mRNAs from MEC (P) to MEC (Prl) transition 0 and only 1 came up in CLASH experiment respectively (Fig. C14. C).

CLASH-seq derived miRNA-mRNA interactome analysis:

CLASH-seq derived Interactome map of upregulated mRNAs and downregulated miRNAs from MEC (N) to MEC (P) were analysed to understand the relevance of miRNA in maintenance of MEC (P) stage specific mRNAs upon Glucocorticoid treatment (Fig. C15. A). Majorly mmu-miR-712-5p was predicted to involve in controlling *Bcom1, Gpr116, Gpr64, Gm4944* and *B630005N14Rik* genes in MEC (N). These genes are involved in metabolic pathways and Retinol metabolism. These genes are getting upregulated in MEC (P) differentiation while mmu-miR-712-5p was downregulated.

In order to understand downregulated miRNA-mRNA network towards MEC (P) differentiation; downregulated mRNAs and upregulated miRNAs from MEC (N) to MEC (P) were analysed (Fig. C15. B). Mmu-miR-1195, mmu-miR-709 and mmu-miR-3473b were major central hubs and were downregulating tight junctions, protein processing in Endoplasmic reticulum and MAPK signalling pathway. Along with these above miRNAs, mmu-miR-466f-3p and mmu-miR-3470b upregulation in MEC (P) was observed.

Functional significance of mmu-miR-122-5p during MECs lactogenic differentiation:

A careful analysis of miRNA-mRNA interactions based on CLASH-seq data from HC11 MECs undergoing lactogenic differentiation and its associated pathways have been carried out. In

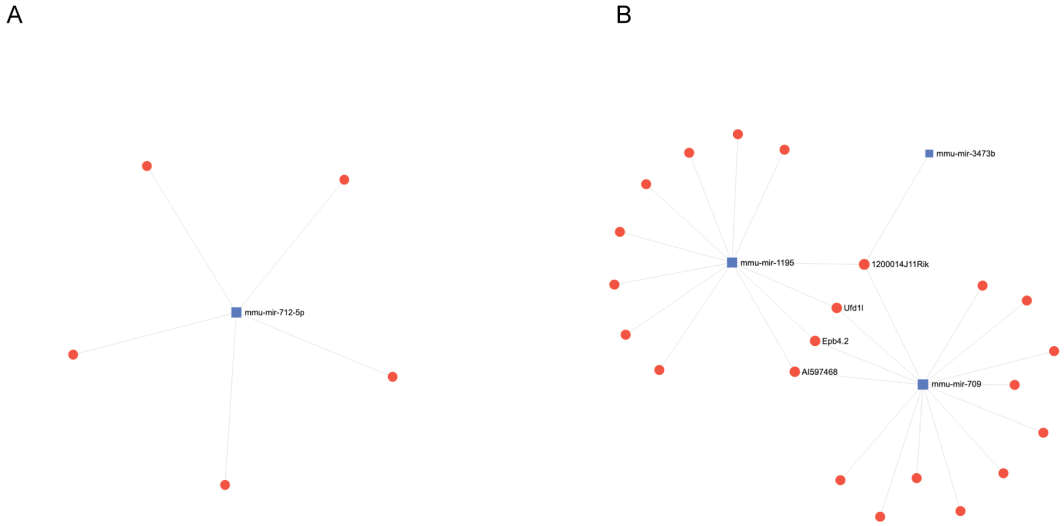


Fig.C15. miRNA-mRNA interactome map of A. Upregulated mRNA with downregulated miRNA and **B.** Downregulated mRNA with upregulated miRNA during MEC (N) to MEC (P) transition.

MEC(P) conditions, a total of six miRNA-mRNA networks were derived from down regulated miRNAs and upregulated mRNAs between MEC(N) and (P) stages. One of the network contains mmu-miR-122-5p node with six different mRNA targets such as *as Per1, Ndrg3, Ctgf, Igf2, Alpl and Alas2* (Fig. C16. A). These genes were reported to be important for mammary gland development. For example, *Per1* was known to be upregulated during transition of MEC(N) to (P) stage and reported to be involved in mammary gland development (McQueen et al., 2018). *Ndrg3* was shown to be required for MECs cell proliferation, but higher expression of it has been seen in breast cancer conditions. *Igf2* is upregulated upon MEC(Prl) treatment and induces cell differentiation in mammary gland (Brisken et al., 2002). *Ctgf* has been seen to be involved in mamospheres formation (Morrison et al., 2010). However, there was no report about *Alpl* and *Alas2* and their potential involvement in mammary gland development. This has attracted us to investigate the potential role of miR-122-5p in regulating these sets of genes that were known to play important roles in mammary gland development and differentiation. Towards this end, we derived all the CLASH

seq derived interacting partners for the miR-122-5p in MEC(N) condition and found that it targets 76 expressed mRNAs (Fig. C13. C). Along with above mentioned mRNAs, other mRNAs such as *Gys1*, *CyclinD1* which play important roles in lactogenic differentiation were found in the list of miR-122-5p interacting mRNAs. Considering the importance of these mRNA genes and their potential regulation by miR-122-5p, functional studies have been carried out.

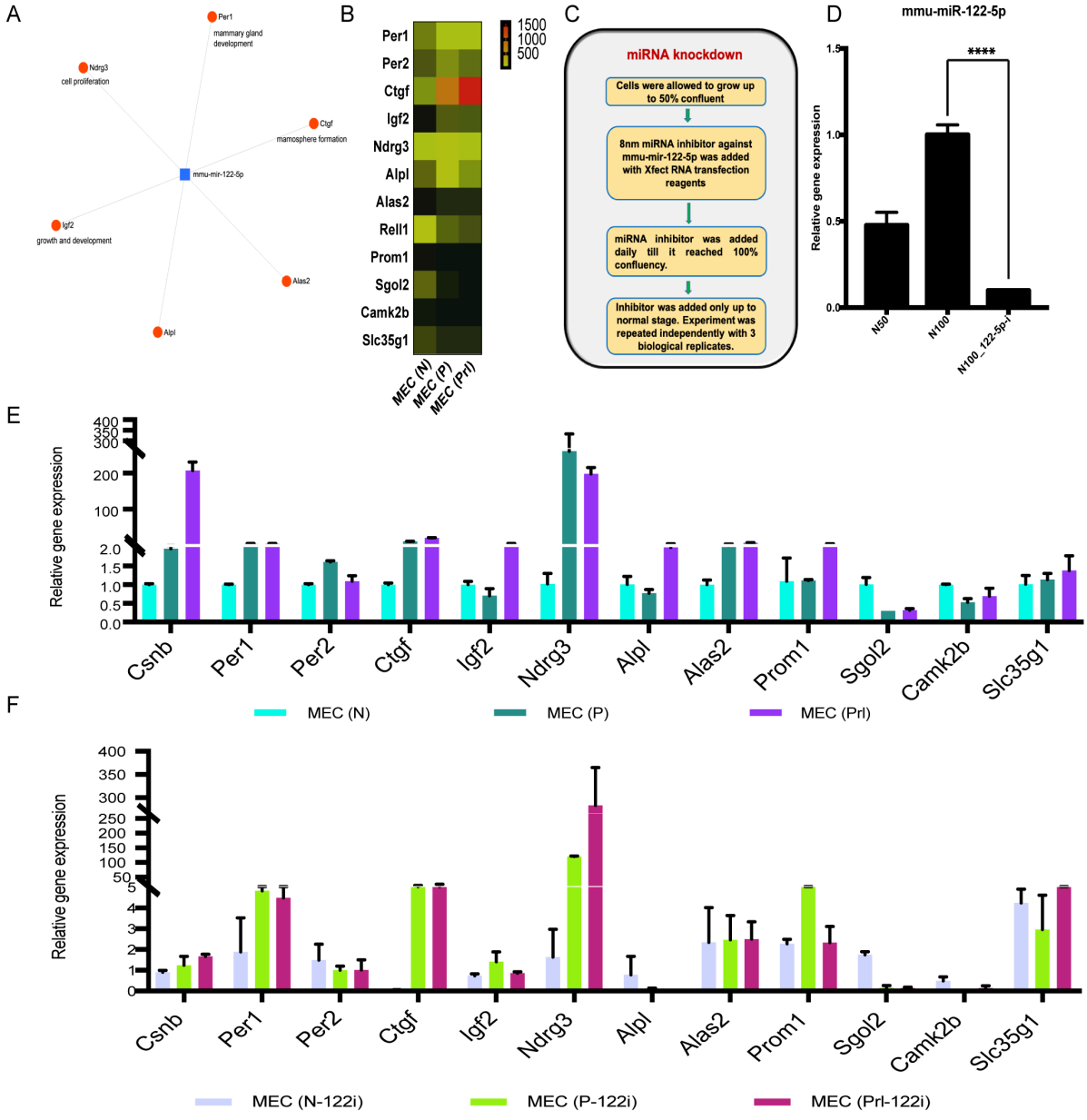


Fig. C16. Study of mmu-miR-122-5p knockdown in MEC(N): A. miRNA-mRNA interactome map of downregulated mmu-miR-122-5p and upregulated mRNAs from MEC (P) vs MEC (N), **B.** Heatmap representing expression status of mmu-miR-122-5p targets, **C.** Flow chart representing mmu-miR-122-5p knockdown study in Lactogenesis. **D.** Relative gene expression status of miR-122-5p upon knockdown by using real-time PCR. Data was analysed by using $-2^{\Delta\Delta CT}$ method. *Rnu6* and MEC(N) were considered as gene control

and sample control respectively. Bar graphs were generated by using Graph-pad PRISM software with one-way ANOVA significance test. Relative gene expression status of mRNAs that are targeted by miR-122-5p in **E.** Normal and **F.** mmu-miR-122-5p knockdown condition by using Real-time PCR. Data was analysed by using -$2^{\Delta\Delta CT}$ method. *Beta-Actin* and MEC(N) were considered as gene control and sample control respectively. Bar graphs were generated by using Graph-pad PRISM software with one-way ANOVA significance test.

In an effort to understand the functional significance of miR-122-5p in HC11 lactogenic differentiation, ShRNA mediated knockdown strategy was employed. Relative levels of miR-122-5p in MEC (N) stages studied by quantitative PCR, showed significant decreased expression of

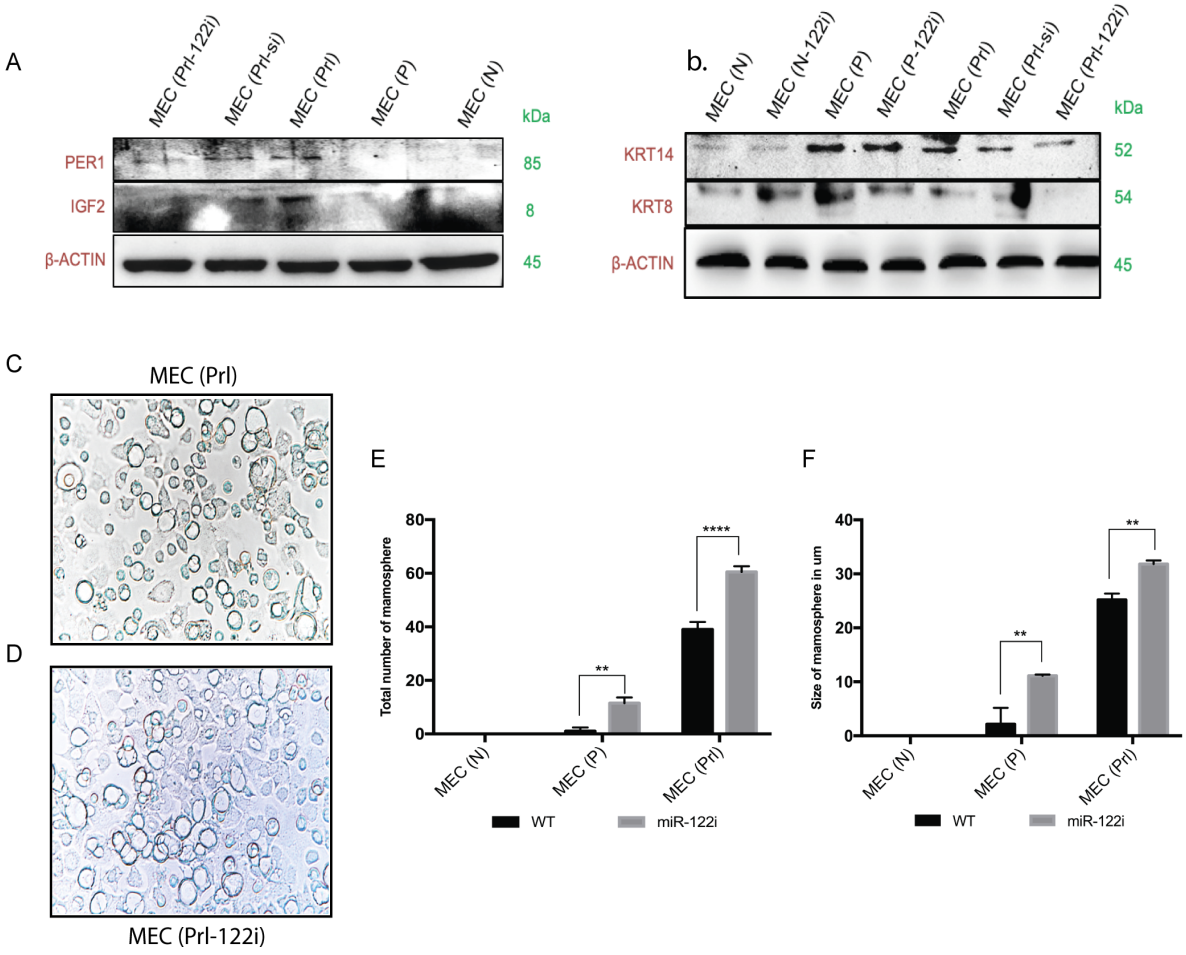


Fig. C17. Increase in mamospheres number and size upon mmu-miR-122-5p knockdown: A-B. Comparison of expression of PER1, IGF2, KRT14 and KRT8 in mmu-miR-122-5p knockdown sample by using Western bolt. **C-D.** 40X bright field images representing effect of mmu-miR-122-5p knockdown in MEC (Prl) cells. Bar graph representing comparison of **E.** total number of mamospheres and **F.** size of mamospheres after mmu-miR-122-5p knockdown. Increase in the mamospheres number and size were observed upon mmu-miR-122-5p knockdown:

miR-122-5p upon knockdown (Fig. C16. C). Further cognate mRNA partners as analysed through CLASH-seq showed increased expression of its targeted genes. These experiments thus suggest that miR-122-5p could potentially pair with its cognate mRNAs in HC11 N stages, and its subsequent down-regulation might have upregulated corresponding mRNAs. To functionally test this possibility, miR-122-5p was knocked down by ShRNA. To knock down miR-122-5p, HC11 cells which were grown to confluence were then incubated with ShRNA specific to miR-122-5p for 4hours (Fig. C16. C). Knockdown of miR-122-5p was validated by real time PCR analysis and significant reductions in its levels were observed (Fig. C16. D). Mmu-miR-122-5p knockdown samples were harvested from MEC (N), MEC (P) and MEC (Prl) stages and then proceeded for miRNA and mRNA isolation, cDNA preparation and then subsequent quantitative PCR experiments. Further, the expression of its targeted mRNAs such as *Per1, Ctgf, Igf2* and Ndrg3 was evaluated by RT-PCR in MEC(N) cells and were found to be significantly upregulated (Fig. C16. F). Significant increase of Ctgf and Ndrg3 have been observed in MEC (P-122i) condition. Also, the expression of Igf2 was increased in MEC (P-122i) condition but showed reduced expression upon MEC(Prl) treatment. Further, reduction in the level of cognate mRNAs were validated using western blot analysis (Fig. C17. A). These experiments were performed in triplicate and its subsequent analysis showed that, Per1 protein was noted to be elevated in MEC (Prl-122i) compared to MEC (Prl) stages. In line with these observations, we also noted that reduced expression of *Csn-b* was both at mRNA and at proteins levels (Fig. C20. A). These changes were also apparent in terms of number and size of mamospheres formation, as their number were observed to be enhanced *in vitro* differentiation conditions (Fig. C17. E-F). We attribute this possibly due to enhanced levels of expression *Ctgf* (as assessed by real-time PCR). In parallel with these, we so noted reductions in the level of Csn-b, Krt14a and Krt18 mRNA and protein levels under MEC(Prl) stages of differentiation (Fig. C17. A-B).

The elevated expression of CTGF/CCN2 is depend upon the level of glucocorticoids during lactogenic differentiation, not on the Prolactin or TGFβ (Wang et al. 2008). In the HC11 mouse mammary epithelial cell background, CTGF/CCN2 expression enhanced the early transcription of β-casein in response to lactogenic hormone. Exogenous addition of CTGF/CCN2 contributed to the formation of mammospheres and MCF10A acini, hallmarks of terminal differentiation (Wang et al. 2008; Debnath and Brugge 2005; Morrison et al. 2010). CTGF/CCN2 enhanced the expression of laminin in mammary epithelial cells resulting in a decreased requirement for exogenous laminin for the activation of β-casein transcription (Morrison et al. 2010). CTGF/CCN2 increased expression of fibronectin and stabilized the surface expression of the α6

and β1 integrins; PINCH1 and Rsu1, proteins found in an integrin- ILK-linked protein complex were also elevated (Wang et al. 2008; Morrison et al. 2010). Collectively, these results suggest that the mechanism by which CTGF/CCN2 enhances lactogenic differentiation is through stabilization of the interaction between laminin and α6β1 integrin that is required early in the lactogenic differentiation process. Lactogenic differentiation may also depend on expression and stabilization of CTGF/CCN2 -integrin complexes that promote MEC(Prl)-induced Stat5 activity (Xu et al. 2007, 2009)

Gradually, when this bipotent cells are started differentiating into luminal and myoepithethial cells during MEC (P) transition from MEC (N), they develop distinct luminal and basal layers. That can be demarcated by the expression of Krt14 in myoepithelial cells and Krt8 in luminal cells. Because major developmental genes such as *Per1, Igf2* and *Csn-b* were getting affected due to mmu-miR-122-5p knockdown, the expression of Krt14 and Krt8 showed reduction in their protein levels (Fig. C17. B). However, reduction in the levels under MEC (P) condition might be an indirect effect of loss of function of miR-122-5p in MEC (N) condition. Keeping this possibility in mind, expression status of downregulated genes like *Rell1, Prom1, Sgol2, Camk2b* and *Slc35g1* during MEC (P) condition and which are potential targets of mmu-miR-122-5p (Fig. C16. E-F) were analysed. Among these *Camk2b* and *Slc35g1* were responsible for calcium ion intake into MECs (Chi et al., 2016). *Rell1* was involved in apoptosis and *Prom1* induces ductal branching in mammary gland during pubertal development (Anderson et al., 2011) suggesting that knocking down the function of miR-122-5p during N stage would impair differentiation of MEC(P) stage suggesting a pivotal role of this microRNA in Lactogenic differentiation of HC11 MECs and its potential role during mammary developmental.

CLASH-seq reveals putative targets for *mmu-miR-122-5p* in HC11:

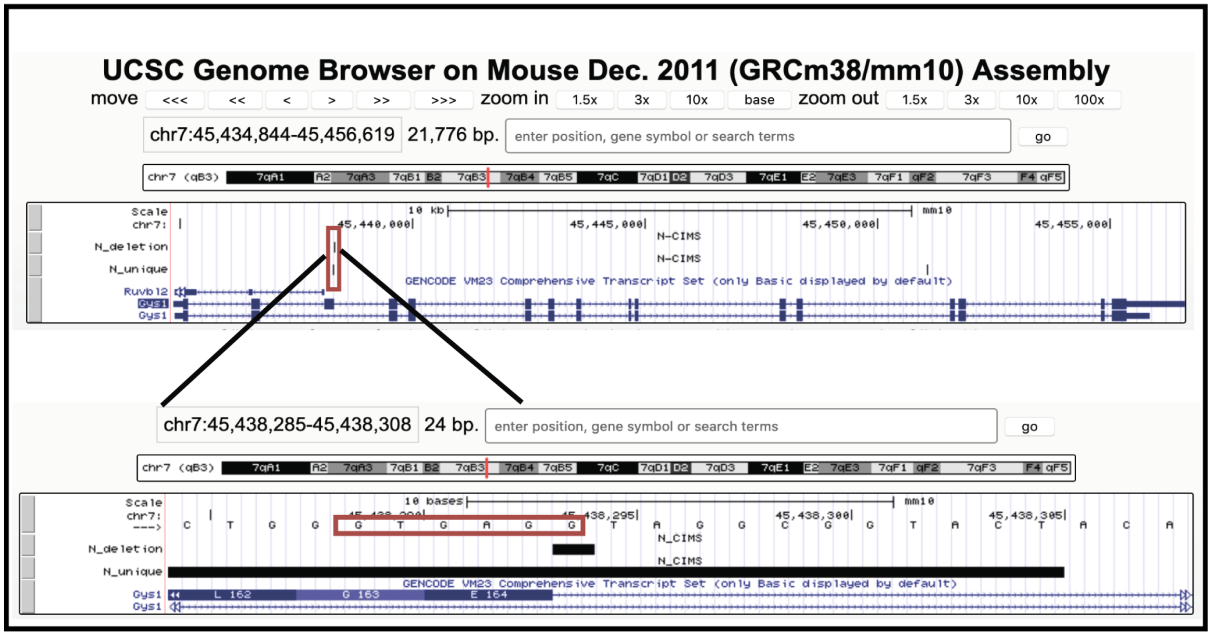
Highly expressed MEC (N) state specific mmu-miR-122-5p showed enrichment only in both the biological replicates of MEC (N) condition but not in either MEC (P), MEC (Prl) or input samples (Fig. C13. B). Filtration of mapped reads by CIMS analysis, have shown that in many of the miR-122-5p interacting mRNAs under MEC(N) stage that showed deletions near AGO binding sites or deletions sites (-20 upstream and +20 downstream) mmu-miR-122-5p seed region were identified. Here, expressed targets for mmu-miR-122-5p was also filtered out that showed in total 45 target mRNAs out of 76 from RNA-seq. Likewise, this analysis ended up with seven mRNA targets for mmu-miR-122-5p. Among them *Gys1, Sec23ip* and *Paxip1 mRNAs* showed exact match

to seed sequence and others viz. *Pkm, Tmed3, Zfp113* and *Fech* showed one base mismatch. It is to be mentioned miR-122-5p and its interactions with these seven mRNAs in HC11 system that were also known to target in other cellular systems as shown by HITS-CLIP experiment. A careful analysis of expression of *Gys1, Sec23ip* and *Paxip1 by RT-PCR* showed its high expression in MEC (N) and gradual reduction upon lactogenic differentiation (Fig. C19. C). This might be due to activation of AMPK signalling pathway upon glucocorticoid treatment, that might lead to downregulation of *Gys1* and regulates glycogen synthesis (Zhang et al., 2010) during lactogenesis. But, there is no experimental evidences about the possible role *Sec23ip* and *Paxip1* genes in relation to mammary gland development.

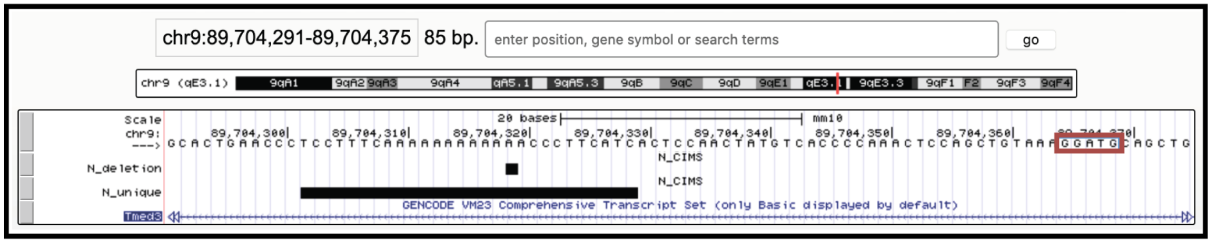
mmu-miR-122-5p **controlling cell proliferation in 100% confluent MEC (N) stage:**

*mmu-mi*R-*122-5p* was well studied in liver physiology, lipid metabolism and stress response (Esau C et al., 2006). Its overexpression was shown to be linked to cell cycle arrest and apoptosis (Ma et al., 2010). miR-122-5p has also been studied in human breast cancer cells (MCF-7), where it's expression was found to be low. Moreover, overexpression of *mmu-mi*R-*122-5p* inhibited the rate of cell proliferation and decreased colony formation (Wang et al., 2012). Further, IGF1R, a gene involved in with Akt pathway was shown to interact with mmu-*mi*R-*122-5p through* luciferase assay (Wang et al., 2012). Expression of Akt, mTOR and P70S6K were effectively downregulated upon *mmu-mi*R-*122-5p* overexpression which suggested inhibition of cell proliferation by targeting PI3K/Akt signaling pathway. *mmu-mi*R-*122-5p* was known to be a negative regulator of breast cancer by suppressing cell growth, colony formation and tumorigenesis (Wang et al., 2012). Keeping these studies in mind and its potential role in down regulation of Akt signalling pathway specific mRNAs in HC11 lactogenic differentiation, downstream target of Akt signalling pathway specific genes such as Cyclin D1 (*Ccnd1*) and *Gys1* and their regulation by *mmu-mi*R-*122-5p* has been investigated. We found that *mmu-mi*R-*122-5p* complementarity in *Gys1* mRNA in mammary epithelial cell lines through CLASH seq analysis. Gys1 was highly upregulated in MEC(N) stage but upon differentiation towards MEC(Prl) state, its levels were gradually reduced (Fig. C19. C). It is to be noted that the rate of glycogenesis was found to be very high in pregnancy mice mammary gland, compared to virgin but found to be lowest in case of lactating mice (Emerman et al. 1980). Glycogen synthesis before parturition was essential to prevent lactose accumulation by converting UDP-4-glucose to glycogen. During parturition glycogen breakdowns to supply free glucose for

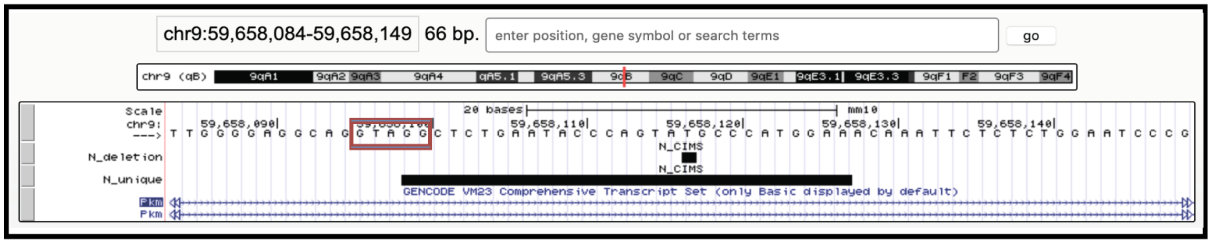
A


UCSC Genome Browser on Mouse Dec. 2011 (GRCm38/mm10) Assembly
move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x 100x
chr7:45,434,844-45,456,619 21,776 bp.
enter position, gene symbol or search terms
go
chr7:45,438,285-45,438,308 24 bp.
GENCODE VM23 Comprehensive Transcript Set (only Basic displayed by default)
N-CIMS
N_deletion
N_unique
Ruvbl2
Gys1

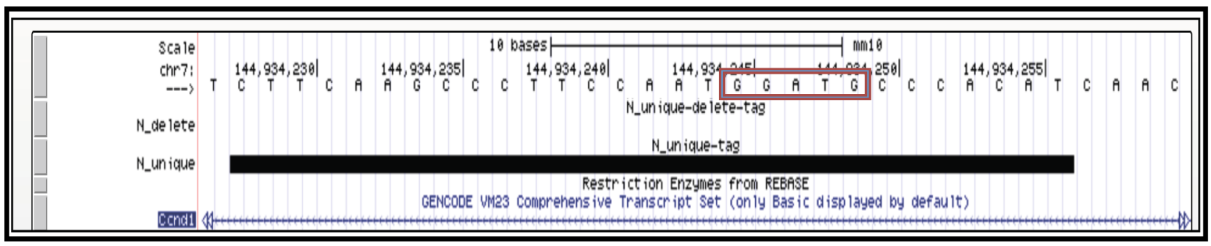
B


chr9:89,704,291-89,704,375 85 bp.
enter position, gene symbol or search terms
go
GENCODE VM23 Comprehensive Transcript Set (only Basic displayed by default)
N-CIMS
N_deletion
N_unique
Tmed3

C


chr9:59,658,084-59,658,149 66 bp.
enter position, gene symbol or search terms
go
GENCODE VM23 Comprehensive Transcript Set (only Basic displayed by default)
N-CIMS
N_deletion
N_unique
Pkm

D


N_unique-delete-tag
N_delete
N_unique-tag
N_unique
Restriction Enzymes from REBASE
GENCODE VM23 Comprehensive Transcript Set (only Basic displayed by default)
Ccnd1

E

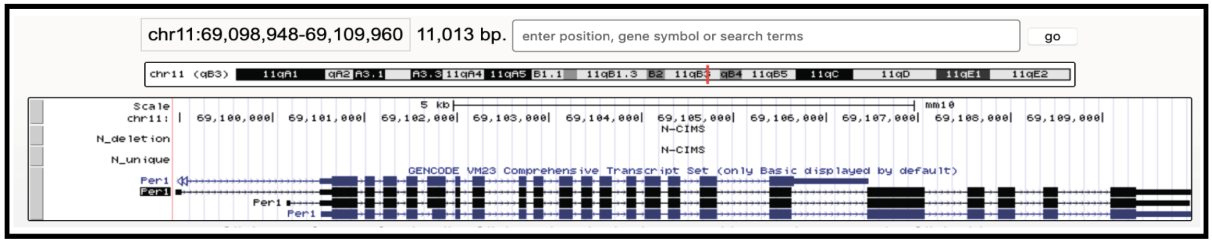

chr11:69,098,948-69,109,960 11,013 bp.
enter position, gene symbol or search terms
go
GENCODE VM23 Comprehensive Transcript Set (only Basic displayed by default)
N-CIMS
N_deletion
N_unique
Per1

Fig. C18. Revealing putative targets of mmu-mir-122-5p by locating the seed region, 'GTGAGG' within its targeted mRNAs in MEC (N) by detecting UV-deletion sites through CIMS analysis of CLASH-seq: A. Screenshot window from UCSC genome browser which is focusing on the seed region of mmu-mir-122-5p detected in *Gys1*. **B.** Screenshot window from UCSC genome browser which is focusing on the seed region of mmu-mir-122-5p detected in *Tmed3*. **C.** Screenshot window from UCSC genome browser which is focusing on the seed region of mmu-mir-122-5p detected in *Pkm*. **D.** Screenshot window from UCSC genome browser which is focusing on the seed region of mmu-mir-122-5p detected in *Ccnd1*. **E.** Screenshot window from UCSC genome browser which is focusing on the seed region of mmu-mir-122-5p not detected in *Per1*.

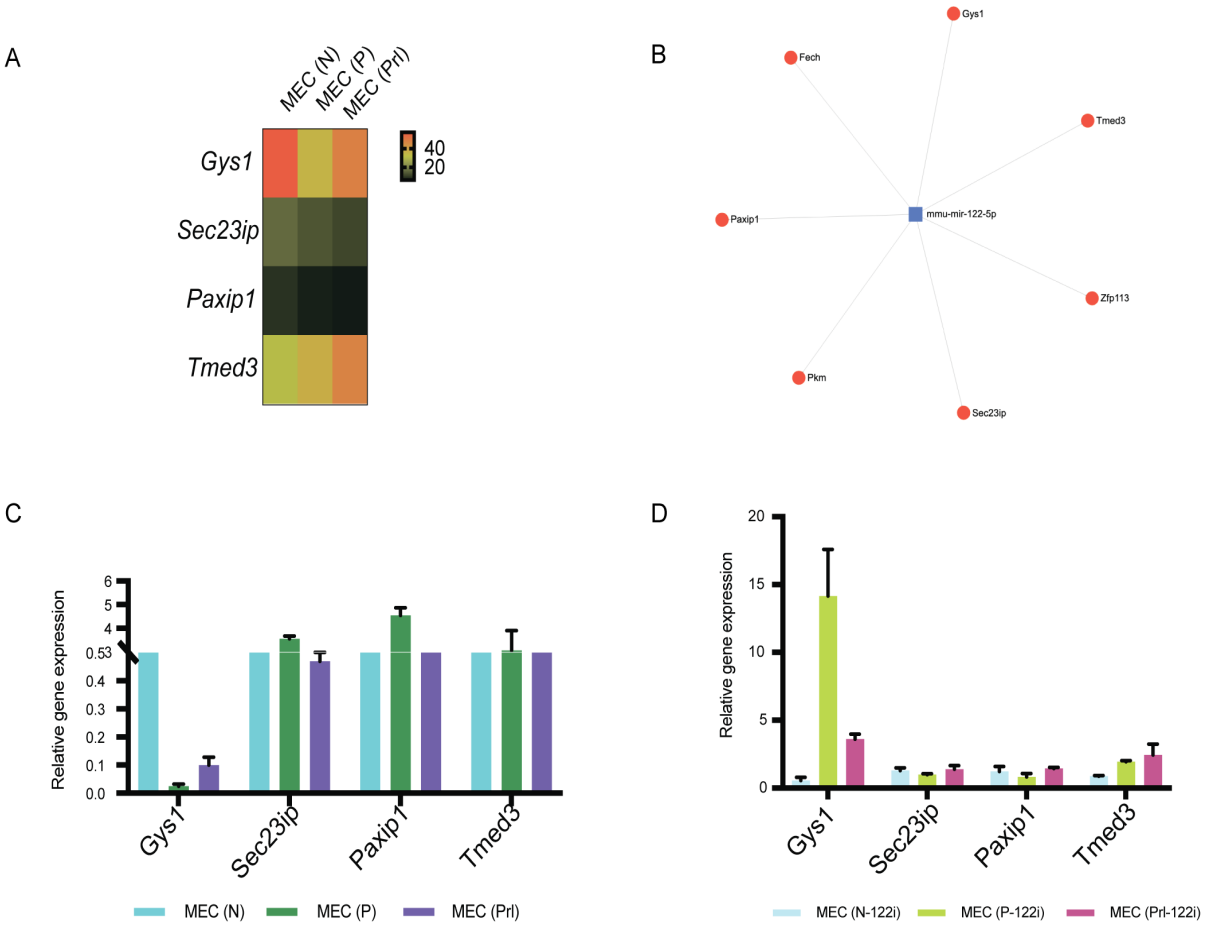


Fig. C19 Validation of putative targets of mmu-miR-122-5p from CLASH-seq by using Real-time PCR: A. Heatmap representing expression status of mmu-miR-122-5p targets, **B.** miRNA-mRNA interactome map of downregulated mmu-miR-122-5p and expressed mRNAs from MEC (N) which were filtered from CLASH-seq having deletion profile in CIMS analysis. Relative gene expression status of mRNAs that are targeted by miR-122-5p (confirmed by CLASH-seq) in **C**. Normal and **D.** mmu-miR-122-5p knockdown condition by using Real-time PCR. Data was analysed by using $-2^{\Delta\Delta CT}$ method. *Beta-Actin* and MEC(N) were considered as gene control and sample control respectively. Bar graphs were generated by using Graph-pad PRISM software with one-way ANOVA significance test. The expression of *Gys1* among all is effectively downregulating upon mmu-miR-122-5p knockdown in MEC(N).

lactose formation. Limiting Gys1, during pregnancy mammary cells, might make availability of free UDP-glucose to convert into UDP-galactose which is essential for lactose formation (Fig. C20. c). Apart from Gys1, Pkm was found to be a potential target for *mmu-miR-122-5p through* CLASH seq data. PKM is a pleiotropic protein that acts like a transcriptional coactivator along with ß-catenin, that might facilitate transcriptional upregulation of *c-Myc, Glut-1* and *Ldh*. Pkm was also known to be a downstream target of Wnt signaling pathway that was promoted by Akt mediated inhibition of GSK3ß (Sherwood, 2015). The transcriptional activation and upregulation of c-Myc was mediated by canonical Wnt pathway through ß-catenin-TCF dependent manner. The upregulation of c-Myc was shown to facilitates the expression of many glycolytic genes such as as glucose transporter 1 (GLUT-1), LDH and PKM. PKM is a catalytic enzyme that generates ATP and pyruvate at the final step of glycolysis. Upregulation of c-Myc controls the expression of genes involved in cell cycle regulation, including cyclins and cyclin dependent kinases (CDKs) (Sherwood, 2015).

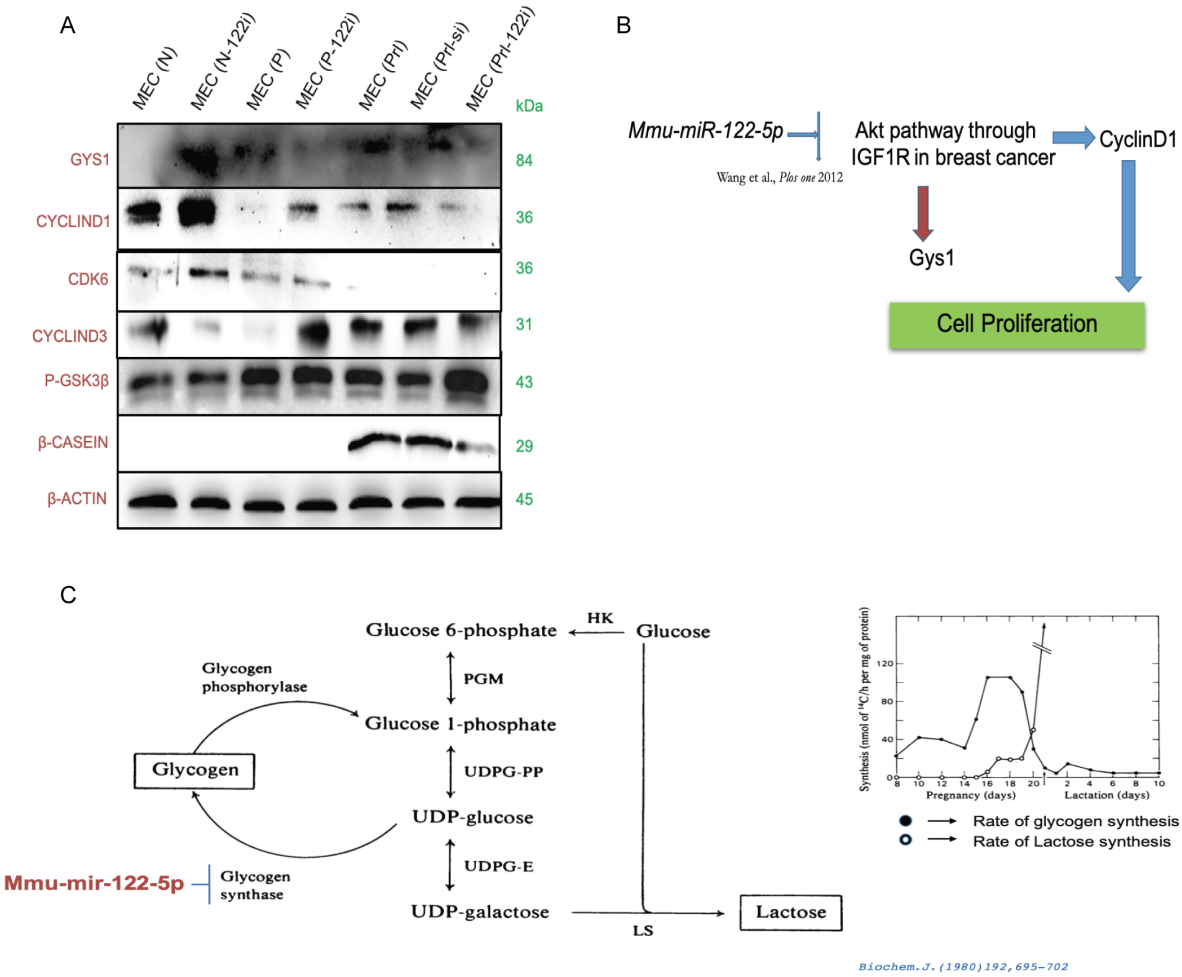


Fig. C20. *mmu-miR-122-5p* knockdown upregulates cell cycle regulators and enhances rate of glycogenesis: A. Comparision of expression of GYS1, CYCLIND1, CDK6, CYCLIND3, P-GSK3B and B-CASEIN in mmu-miR-122-5p knockdown samples through western blot. **B.** Graphical representation showing role of mmu-miR-122-5p in controlling cell proliferation through Akt signalling pathway. **C.** Graphical representation showing role of mmu-miR-122-5p in controlling rate of glycogenesis by targeting Gys1.

mmu-miR-122-5p knockdown upregulates cell cycle regulators during differentiation:

Expression of *mmu-miR-122-5p* is higher in fully confluent MEC (N) state and its target of 45 expressed mRNAs which showed association with RISC complex. KEGG pathway analysis of all interacting 45 mRNAs targets were performed. This analysis revealed that mmu-miR-122-5p potentially can target essential pathways required for mammary gland development. We identified that JAK/STAT pathway, AKT pathway, Wnt pathway, Oxytocin signalling pathway, FoxO signalling pathway, MAPK signalling pathway, ErbB signalling pathway, AMPK signalling pathway, MEC(Prl) signalling pathway and Insulin signalling pathways were also found to be targeted by miR-122-5p. *mmu-miR-122-5p* seems to act on JAK/STAT and MEC(Prl) signalling pathway by directly controlling *Socs2* and *CyclinD1* (*Ccnd1*) gene mRNAs. Binding of MEC(Prl) to MEC(Prl) receptor leads to dimerization of MEC(Prl) receptor, that activates JAK2 followed by phosphorylation of MEC(Prl) receptor and STAT5 (Watson et al., 1996). SOCS members SOCS1, SOCS2, and SOCS3 were noted to be suppressor of cytokine signalling (Pezet et al., 1999) and SOCS2 which was a target of *mmu-miR-122-5p,* associates with MEC(Prl) receptor (PRLR). SOCS1 and SOCS3 are early expressed SOCS genes those inhibit PRL signalling but, SOCS2 which expresses later, suppresses the inhibitory activity of SOCS1 and increases the sensitivity towards PRL signalling (Pezet et al., 1999). *Cyclin D1* is a downstream target of both JAK/STAT and MEC(Prl) signalling pathway. Cytokines like, IL3 and IL6 can stimulate the activity of *Cyclin D1* through STAT3 and STAT5 (Klein et al., 2008). CyclinD1 regulates cell cycle during G1-S phase transition and also activates CDK4 and CDK6 (Landis et al., 2006). *Cyclin D1* is a direct target of *mmu-miR-122-5p* and are also controlled by inhibiting JAK/STAT pathway by limiting Socs2 which itself is targeted by *mmu-miR-122-5p*. *Igf-2* is also a downstream target of MEC(Prl) signalling which has STAT binding sites at its promoter and induces *Igf-2* transcription through STAT5 (Brisken et al., 2002). Igf-2 acts as an upstream activator of *Cyclin D1* and also a direct target of *mmu-miR-122-5p*.

Mitogenic growth factors like Cyclin D1 activators that through MAPK signaling pathway are well studied. The promoter region of Cyclin D1, at 930bp upstream to transcription start site, contains a AP-1 binding site (Shen et al., 2007). AP-1 positively regulates the expression of Cyclin D1 and this AP-1 binding site is regulated by Fos and Jun dimers. Jun also form heterodimer with activating transcription factor 2 (ATF-2) and regulate Cyclin D1 by binding to cAMP response element (CRE) at the promoter of Cyclin D1. Here, *mmu-miR-122-5p* is controlling Cyclin D1 expression via MAPK signalling by targeting Jun. Also, Wnt signalling stimulates CyclinD1 by translocating ß-catenin to nucleus which forms a complex with ternary complex factor (TCF) and/or lymphoid enhancer-binding factor (LEF) transcription factors and enhances Cyclin D1 transcription (Klein et al., 2008). CyclinD1 showed significant increased expression upon *mmu-miR-122-5p* knockdown in MEC (N-122i) condition. When checked for other cell cycle regulators like CyclinD3 and Cdk6 by western blot; CyclinD3 did not show any change of expression in MEC (N-122i) but was significantly increased in MEC (P-122i). Similar to CyclinD1 Cdk6 showed increased expression in MEC (N-122i) condition (Fig. C20. A). Thus, increased in cell proliferation take place upon knockdown of *mmu-miR-122-5p* (Fig. C20. B).

Summary

One haploid genome turns into a multicellular organism that gave us insight how a single genome regulates formation of multiple cell types. Inner cell mass from blastocyst is pluripotent in nature. If we isolate inner cell mass and allow them to culture on petri dishes, it is called Embryonic stem cells (ESC). ESC, like inner cell mass gives rise to three primary germ layers, named ectoderm, mesoderm, and endoderm. ESC pluripotency is regulated by the core transcription factors like, Oct4, Sox2 and Nanog. They positively regulates promotors of many ESC specific genes and downregulates lineage specific markers. During lineage specific differentiation majorly epigenetic modes of gene regulation determines the cell fate. We can classify three levels of epigenetic modes of gene regulation such as, during pre-transcript level, transcript level and post-transcript level. pre-transcript epigenetic modifications generally depends upon three dimensional chromatin organization, chromatin accessibility, chromatin condensation, replication timing of genome, etc. Similarly, transcript level epigenetic modifications depends upon DNA methylation patterns, transcription factors binding sides, histone modifications, alternative splicing, etc. Post-transcript epigenetic modifications involves RNA binding proteins, long non-coding RNAs, microRNAs, etc. Here I am focused on miRNA mode of gene regulation during lineage directed cell differentiation and signal induced cellular differentiation in mammals.

Evidences from these above results and discussions clearly elucidated the importance of gene regulation post-transcriptionally in developmental contexts. Considering miRNA-mRNA interaction analysis during early embryonic development, neurogenesis and lactogenesis demonstrated role of miRNAs in cell fate commitment. Study of spectrum of pluripotency in early embryonic development showed dramatic changes in miRNAs in different stages of ESC though all are in pluripotency state. Depending upon culture medium, ESCs showed different mRNAs and miRNAs expression along with change in pathways for maintenance of pluripotency. Effect of 2i is not suppressed due to presence of serum in ESC (Naïve+2i) state. Downregulation of Metabolic pathway specific mRNAs from ESC (G) to ESC (N+2i) that are targeted by upregulated 2i/Serum/LIF specific miRNAs, were identified. Removal of 2i mediates downregulation of Wnt signalling specific genes during ESC transitions were correlated with upregulation of Serum/LIF specific miRNAs and upregulation of MEK signalling specific genes with downregulation of 2i/LIF specific miRNAs.

During neurogenesis, ESCs to CN differentiation is accompanied by differential expression of many mRNAs and miRNAs. Dynamic reciprocal expression of miRNAs are correlated with expression of neuronal specific genes and pathways involved in neuronal differentiation. *mmu-mi*R-*301b-3p* is highly upregulated in CN which can potentially target 335 mRNA genes including Pbx1 indirectly which is involved in corticogenesis. Knockdown of *mmu-mi*R-*301b-3p* increases the *Pbxip1* mRNA and Protein levels and it's correlated with reduction in the level of *Pbx1* mRNA and enhancement in the level of oligodendrocyte markers such as MAG, *Mog, Tcf712* along with an astrocyte specific marker GFAP. A potential role of *mmu-mi*R-*301b-3p* in the cell fate decisions in between cortical neuron and oligodendrocyte has been elucidated.

Glucocorticoid signalling majorly altered many developmentally important miRNAs. But upon MEC(Prl) signalling during MEC(P) to MEC(Prl) transition no miRNAs downregulation and a few miRNAs have shown upregulation. miRNA and TF analysis inferred upregulation of key developmental TF upon down-regulation of *mmu-mir-122-5p*. *mmu-mir-122-5p* is having 76 expressed mRNA targets in MEC(N) stage among them *Gys1, Tmed3, CyclinD1*, *Per1, Ndrg3, Ctgf* and *Igf2* are developmentally important and showed variation upon knockdown with reduced β-CASEIN level. CIMS analysis confirmed *Gys1, CyclinD1 and Tmed3* are potentially targeted by *mmu-mir-122-5p*. Upregulation of GYS1 and CYCLIND1 with increased p-GSK3β infers upregulation of Akt signalling pathway upon knockdown of *mmu-mir-122-5p* that mediates increase in cell proliferation markers CYCLIND1, CYCLIND3 and CDK6. *mmu-mir-122-5p* downregulates *Gys1* for availability of UDP-glucose during lactose synthesis. A potential role of *mmu-mi*R-*122-5p* in cell cycle arrest and process of lactose synthesis during lactogenesis has been elucidated.

Tables

	ESC(G) 1	ESC(G) 2	ESC(N+2i) 1	ESC(N+2i) 2	ESC(N) 1	ESC(N) 2
ESC(G) 1	1	0.9799726	0.9187531	0.9256711	0.9557035	0.9569179
ESC(G) 2	0.9799726	1	0.9187091	0.9253029	0.9548657	0.9558106
ESC(N+2i) 1	0.9187531	0.9187091	1	0.9744941	0.9116571	0.9120085
ESC(N+2i) 2	0.9256711	0.9253029	0.9744941	1	0.9146821	0.9149469
ESC(N) 1	0.9557035	0.9548657	0.9116571	0.9146821	1	0.9817977
ESC(N) 2	0.9569179	0.9558106	0.9120085	0.9149469	0.9817977	1

Table: 1 ESCs mRNAseq Spearman's correlation statistics between samples and replicates with p-value < 2.2e-16.

	ESC(G) 1	ESC(G) 2	ESC(N+2i) 1	ESC(N+2i) 2	ESC(N) 1	ESC(N) 2
ESC(G) 1	1	0.9404575	0.8930997	0.8875898	0.9051593	0.9066946
ESC(G) 2	0.9404575	1	0.889291	0.8885801	0.9073724	0.9081039
ESC(N+2i) 1	0.8930997	0.889291	1	0.9225088	0.8833857	0.8899646
ESC(N+2i) 2	0.8875898	0.8885801	0.9225088	1	0.8975764	0.8972952
ESC(N) 1	0.9051593	0.9073724	0.8833857	0.8975764	1	0.9464444
ESC(N) 2	0.9066946	0.9081039	0.8899646	0.8972952	0.9464444	1

Table: 2 ESCs miRNAseq Spearman's correlation statistics between samples and replicates with p-value < 2.2e-16.

mRNA Sample name	Raw reads	GC Content (%)	Q20 (%)	Q30 (%)	Mapped reads (%)
ESC(G)_1	1,98,59,926	49.98%	97.42	97.42	89.10%
ESC(G)_2	2,02,36,536	50.13%	97.72	93.66	89.50%
ESC(N+2i)_1	19416948	48	99.81	99.52	98.07
ESC(N+2i)_2	20272824	48	99.68	99.15	98.62
ESC(N)_1	2,25,23,225	49.73%	97.7	93.59	89.70%
ESC(N)_2	2,36,69,902	50.00%	97.75	93.69	89.70%

Table: 3 ESCs mRNA sequencing raw reads and their quality reports.

Sample name	Raw reads	GC Content (%)	Q20 (%)	Q30 (%)	Mapped reads (%)
ESC(G)_1	2,61,78,461	49.23	99.62	98.68	94.85
ESC(G)_2	1,75,85,551	49.31	99.56	98.8	93.92
ESC(N+2i)_1	1,95,58,854	44	99.99	95.72	88.08
ESC(N+2i)_2	2,37,04,914	48	99.99	95.9	80.96
ESC(N)_1	1,90,73,329	49.39	99.54	98.73	93.62
ESC(N)_2	2,04,56,334	49.76	99.52	98.74	92.87

Table: 4 ESCs miRNA sequencing raw reads and their quality reports.

ESC(G)	FPKM	ESC(N+2i)	FPKM	ESC(N)	FPKM
Gm27406	49.7843	Mir692-2	2327.915	Gm21060	164.9035
Gm28060	37.1096	Mir703	1961.4	Gm48217	63.6475
Gm15610	24.34085	Gm15536	1132.725	Gm28625	34.49195
Gm8200	19.0517	Gm28911	1101.9025	Gm15378	31.8422
Crygn	17.3519	Gm3788	1080.885	Gm37614	31.30995
Gm6704	17.2937	Gm10060	772.9105	Gm27999	30.30815
CT025601.1	15.7449	Gm49286	658.37	Cnn1	27.5042
Tex13b	13.18165	Gm44350	634.486	Gng13	24.732
1700018B24Rik	12.4638	Bc1	456.681	Igkv13-78-1	22.0071
Gm33051	9.54285	Gm44314	359.202	Acta1	19.07665
Il4i1	9.17162	Gm20430	346.6695	Gm11534	18.624
CT010429.1	8.99705	Gm44454	325.377	Gm29284	15.814155
Pkd2l1	8.35054	Gm47441	298.181	Ccno	15.1976
H3f3aos	7.753559	Gm44394	297.999	Gm44954	15.0757
1700066B19Rik	7.36463	Gm42845	294.2415	Gbp2	11.9798
Gm47135	6.9953395	Gm44357	291.2635	Car4	11.706
Gm38804	6.7432	Gm44486	236.8695	Aass	11.3922
Gm29630	6.47443	Gm7266	207.1065	Gm43107	11.3339
Osm	6.33876	Gm19810	130.994	Ptprt	10.883645
Gm43810	6.15415	Gm14287	127.4255	Vgf	10.7241
1700102P08Rik	5.72588	Gm11539	127.3455	Wnt3	10.130035
Tnni1	5.268875	Gm12013	125.439	Gm7942	9.770645
Acsm4	5.172875	Gm2174	113.1845	Pdlim3	9.492725
AI427809	4.895605	Gm25100	100.6355	Arg2	9.33906
Srarp	4.89417	Gm7434	97.7403	Gbp2b	9.17845
G430095P16Rik	4.89361	Rps15a-ps8	91.42635	Tacstd2	9.14987
Cd180	4.84976	AC165157.1	89.01025	Fndc3c1	9.028585
Gm44877	4.765745	Gm6394	88.9093	Tmem54	8.938295
Casq2	4.688165	Gm44484	87.50365	Sema3e	8.87543

Table: 5 Top uniquely expressed mRNAs among ESC(G), ESC(N+2i) and ESC(N) state.

ESC(G)	FPKM	ESC(N+2i)	FPKM	ESC(N)	FPKM
Rpl39-ps	50844.15	Gm29216	6214.54	Gm13392	113648.35
Gm13392	36956.95	Gm28661	5470.925	Rpl39-ps	84643.9
Gm14303	22630.9	Gm28437	4900.01	Gm14303	23255.5
Actb	17810	Rpl41	4200.43	Actb	18990.35
Gm10443	13272.2	Gm14303	3955.16	Gm10443	13083.05
Rpl38-ps2	10633.7	Eef1a1	3900.595	Rpl38-ps2	9628.975
Gm28437	9586.6	Gm10443	3801.92	Tma7-ps	8705.72
Gm28661	9412.535	Gm10275	3313.935	Gm28661	8295.975
Gm10925	9172.935	Gm10925	3264.44	Gm4149	7405.285
Rps27rt	8326.935	Rpl8	3261.255	Rps27rt	7324.655
Gm28438	7590.67	Apoe	3150.825	Gm7536	7058.68
Rpl37rt	7552.31	Gm4332	2795.295	Rpl36-ps12	6860.18
Gm11808	7289.65	Ftl2-ps	2692.37	Gm43110	6765.33
Gm4149	7127.465	Rps16-ps2	2670.165	Gm28437	6573.49
Gm29216	7089.76	Gm9794	2589.145	Rpl37rt	6153.36
Gm7536	6718.66	Gm15427	2525.15	Gm10925	6144.94
Rpl36-ps12	6243.02	Rpl37rt	2426.47	Rps26-ps1	6135.48
Gm14539	6194.755	Mir692-2	2327.915	Rps12-ps3	5946.36
Gm43110	6019.275	Gm11808	2321.615	Gm29216	5919.135
Rps12-ps3	5983.42	Rps5	2190.12	Rpl13a	5882.22
Rpl32	5973.76	Tpt1-ps3	2142.34	Gm11808	5787.855
Rpl13a	5790.695	Rack1	2101.565	Gm2000	5631.26
Gm9794	5715.17	Rps23-ps1	2085.515	Gm12338	5539.65
Rps26-ps1	5625.51	Rpl13a	2001.305	Rpl36a-ps2	5514.72
Gm16418	5563.975	Mir703	1961.4	Gm9794	5358.345
Gm10275	5403.06	Ftl1-ps1	1953.99	Gm10123	5349.05
Rpl31-ps8	5349.865	Gm8203	1886.305	Gm10275	5284.85
Rpl36a-ps2	5323.795	Gm13456	1878.2	Gm14539	5201.095
Gm12338	5190.8	Gm11478	1850.81	Gm28438	5184.05
Rps23-ps1	5054.79	Gm12918	1831.41	Rpl34-ps1	5045.69

Table: 6 Top highly expressed mRNAs among ESC(G), ESC(N+2i) and ESC(N) state.

ESC(N+2i) vs ESC(G)	log2FoldChange	ESC(N) vs ESC(N+2i)	log2FoldChange
Gm7266	14.18077769	Mroh3	9.90954219
Gm13498	13.52063789	Gbp2b	9.634737898
Rps11-ps2	13.12703617	Mir6381	8.626658178
Rps2-ps4	12.80703613	Gm44393	8.505810957
Gm9061	12.77770622	Gm15247	8.474145812
Gm10313	12.63542466	Mif-ps3	8.167136961
Gm4754	12.61667258	Gpm6a	7.957240381
Gm15464	12.52631174	Grm6	7.721642227
Gm7985	12.49334316	Mid1-ps1	7.688184764
Gm5558	12.36786185	2310007B03Rik	7.680179332
Gm5452	12.32941448	Kynu	7.677281276
Gm7434	12.00050671	Fndc3c1	7.642353565
Rpl22-ps1	11.95852418	Ccl20	7.640347598
Rps15a-ps8	11.86282213	St6gal2	7.610870124
Gm11362	11.80979455	Scin	7.571285317
Gm5160	11.80574352	Gbp2	7.370369481
Gm6218	11.76887608	Rasef	7.151097512
Rpsa-ps11	11.76162129	Mid1	7.135900219
Rpl31-ps4	11.74839398	Olfr539	7.033619235
Gm18078	11.69757667	Rdh1	6.802493621
Gm8692	11.53851164	Gm48529	6.680098463
Rpl21-ps6	11.514942	Gm21060	6.647240362
Gm7507	11.33838726	Tagln	6.590309724
Gm37348	11.33815941	Gm15726	6.566032704
Gm7857	11.32434908	Syt4	6.546181982
Gm7293	11.32097866	Gm12660	6.524555607
Gm7565	11.3137715	Atoh1	6.510195904
Gm32899	11.29455831	Dppa3	6.507033378
Hspa9-ps1	11.27086399	Macc1	6.499405563
Rpl13a-ps1	11.15971689	Jakmip2	6.482416018

Table: 7 Highly upregulated mRNAs during ESC(N+2i) vs ESC(G) and ESC(N) vs ESC(N+2i).

ESC(N+2i) vs ESC(G)	log2FoldChange	ESC(N) vs ESC(N+2i)	log2FoldChange
Mroh3	-9.52386098	Gm7266	-14.42453437
Gm15247	-8.908652499	Gm5160	-13.44829204
Rhox2c	-7.793621981	Rpsa-ps11	-13.42919045
Mid1-ps1	-7.635717697	Rps2-ps4	-13.04944162
Rhox2a	-7.273106336	Gm6394	-13.01495511
Mid1	-7.191367424	Gm10313	-12.8776278
9430069I07Rik	-7.134828652	Gm7985	-12.73584152
Mif-ps3	-6.94050949	Gm5558	-12.61104977
Gm44460	-6.798472114	Gm7434	-12.24478435
Gm44323	-6.762387936	Gm12017	-12.23687125
Gm15726	-6.679047526	Gm3809	-12.15301267
Gm21742	-6.665138839	Rps15a-ps8	-12.10640013
Tex13b	-6.40762604	Rpl31-ps4	-11.99045753
Grid2	-6.245621712	Gm18078	-11.94142328
Tslrn1	-6.043657045	Gm8692	-11.781518
Acsm4	-5.926120234	Gm9061	-11.58550044
Akr1cl	-5.894100714	Gm37348	-11.58085472
Gm47730	-5.748992507	Gm7857	-11.56679718
Fndc7	-5.711211491	Gm7565	-11.55684905
Dnajc5g	-5.700981649	Gm32899	-11.5383296
Gm30717	-5.631176529	Hspa9-ps1	-11.51472623
B230311B06Rik	-5.622891662	Gm4754	-11.43688734
Gm13655	-5.537722586	Hnrnpa1l2-ps2	-11.41719689
Rhox2e	-5.525991642	Gm11449	-11.38736375
Nlrp4c	-5.426974268	Gm8318	-11.37900122
Ptprtos	-5.265938671	Gm3699	-11.37281643
Pkd2l1	-5.107838733	Gm15464	-11.32699112
Cela2a	-5.10206889	Gm13498	-11.32291779
Arntl2	-5.100892672	Gm19810	-11.31395573
Arhgef38	-5.054437457	Rps19-ps5	-11.30123194

Table: 8 Highly downregulated mRNAs during ESC(N+2i) vs ESC(G) and ESC(N) vs ESC(N+2i).

ESC(G)	FPKM	ESC(N+2i)	FPKM	ESC(N)	FPKM
Asz1	4.121735	Nkx6-3	9.769115	Mar-05	90.15945
Zfp467	2.54128	Cspg4	7.481415	Sep-09	50.6649
Hesx1	1.95907	Foxo6	4.4554	Sep-10	11.0739
Foxr1	1.425332	Foxg1	4.10601	Egr1	7.82888
Maml2	1.41257	Sox1	3.959	Rem2	4.85684
Gsc	1.371695	Hoxa10	3.905265	Pax6	4.04487
Htr5a	1.238995	Dmbx1	3.899005	Nlrp9b	3.775465
Nfe2l3	1.15322	Dmrta2	3.739635	Dll1	3.691505
Klf1	1.124895	Id4	3.482855	Wt1	2.88635
Zfp558	1.121655	Bach2	3.21395	Hand1	2.879795
Lhx2	1.0253885	Nr2f2	3.1769685	Trim21	2.641835
		Hoxa11	3.01503	Ankrd22	2.276485
		Dlx2	2.992771	Sox18	1.97379
		Irx1	2.8172445	Lmx1a	1.670095
		Zbtb7c	2.59006	1700020N01Rik	1.61594
		Pou2f2	2.532305	Rxrg	1.570455
		Aff3	2.38463	Hist1h1b	1.503645
		Prdm16	2.30452	Creb5	1.491184
		Sox9	2.2678725	Stard13	1.4064665
		Gsx2	2.25021	Nkx2-2	1.401029
		Hsf4	2.228495	Zfp595	1.37283
		Zbtb4	2.17025	En1	1.357778
		Dlx1	2.112925	Nfam1	1.344045
		Foxq1	2.062085	Magel2	1.317415
		Zfp316	2.05822	Nkd2	1.28118
		Hoxa5	1.979181	Tbx4	1.27658
		Dpf3	1.83125	Lmo2	1.219965
		Six3	1.83088	Smc1b	1.21247
		Prrx1	1.818432	Gata3	1.20035
		Irx5	1.754514	Ascl4	1.1760085

Table: 9 Top uniquely expressed TFs among ESC(G), ESC(N+2i) and ESC(N) state.

ESC(G)	FPKM	ESC(N+2i)	FPKM	ESC(N)	FPKM
Hsp90ab1	1936.19	Hsp90ab1	1692.89	Hsp90ab1	2171.12
Pou5f1	1906.675	Pou5f1	1108.33	Pou5f1	1831.89
Pfn1	1498.82	Zfp42	807.159	Pfn1	1705.8
Ybx1	1163.905	Pcbp2	788.5575	Ybx1	1658.38
Trim28	836.3875	Trim28	705.8905	Prmt1	1085.74
Phb2	813.55	Cdk4	675.9905	Trim28	1037.11
Cnbp	767.469	Pfn1	673.389	Pcbp2	776.488
Pcbp2	743.199	Ybx1	647.894	Hnrnpu	766.742
Pcbp1	739.719	Klf2	625.8445	Ranbp1	714.3675
Prmt1	694.3325	Prmt1	548.6925	Cnbp	663.7435
Zfp42	679.2895	H3f3b	529.99	Ncl	607.561
Banf1	569.903	Pcbp1	478.153	Banf1	536.4215
Hnrnpu	531.5235	Hist1h2ah	453.437	Ifitm3	510.8625
Ranbp1	522.132	Cnbp	422.5985	Utf1	449.119
Lsm4	515.7115	Ctnnb1	389.1285	Tardbp	436.56
Nanog	466.765	Fscn1	379.7335	Phb2	429.6285
Tardbp	423.5825	Snrpb	375.992	Set	428.8835
Ncl	422.411	Dnmt3l	374.6665	C230062I16Rik	421.464
Nutf2-ps1	413.3275	Ncl	365.6835	Zfp42	410.845
Cdk4	395.3725	Sox2	358.8225	Nutf2-ps1	400.818
Snd1	385.2185	Banf1	355.3815	Pcbp1	387.219
H2afx	377.195	Mcm5	326.815	Ctnnb1	383.8835
Snrpb	351.162	Hist1h2af	315.0375	Cdk4	383.469
Set	338.2885	Esrrb	312.5625	Snrpb	382.1945
Nasp	326.742	Ash2l	312.4055	Gpx4	378.146
H3f3b	315.3415	Nanog	312.027	Lsm4	367.77
Fscn1	311.6785	Cct4	304.6115	Apex1	362.201
Mcm5	305.0875	Hnrnpu	302.072	Nasp	354.953
Apex1	303.4945	Ran	299.54	Nedd8	346.887
Aes	298.842	Apex1	294.3085	Fus	343.2725

Table: 10 Top highly expressed TFs among ESC(G), ESC(N+2i) and ESC(N) state.

ESC(N+2i) vs ESC(G)	log2FoldChange	ESC(N) vs ESC(N+2i)	log2FoldChange
E130201H02Rik	7.769204653	Atoh1	6.510195904
Hist1h2ab	7.132621042	Dppa3	6.507033378
Sox9	6.894421199	Wt1	6.005988711
Irf2	6.715822152	Fgfbp1	5.658409039
Hist1h2af	6.672314073	Egr3	5.218253641
Foxa2	6.565614448	Pax6	4.942869123
Cspg4	6.499368432	Arntl2	4.687329574
Twist2	6.261001425	Bcl11a	3.874633353
Hist1h2ah	6.146879608	Snai3	3.862431886
Dlx2	6.111483004	Cdx2	3.711483111
Hoxa11	5.914901115	Lef1	3.506404098
Aebp1	5.718507541	Rhox9	3.437686381
Zeb2	5.662535586	Magel2	3.388756938
Ppargc1a	5.439590348	Egr1	3.365628119
Ebf3	4.944746672	Etv2	3.302447822
Dmbx1	4.940479652	Pou3f1	3.270199323
Otx2	4.789358402	Tbx20	3.135242297
Zfhx4	4.682687635	Rxrg	2.966502723
Ciita	4.55658061	Rem2	2.801966046
Hopx	4.54084136	Id2	2.781800325
Hist1h2ak	4.46467524	Pitx2	2.754095313
Irx1	4.459023388	Ifitm3	2.72202978
Fosl1	4.301130627	Mycn	2.713010033
Nfatc1	4.257821085	Neurod1	2.695426654
Dlx1	4.0977136	Mlf1	2.695370496
Foxc2	4.078515216	Lmx1a	2.693570436
Foxq1	3.989139877	Hdx	2.678802692
Hoxa10	3.92980988	Nlrp9b	2.669440076
Ebf1	3.690503881	Ankrd22	2.626187818
Hoxb4	3.56147788	Rhox6	2.622119874

Table: 11 Highly upregulated TFs during ESC(N+2i) vs ESC(G) and ESC(N) vs ESC(N+2i).

ESC(N+2i) vs ESC(G)	log2FoldChange	ESC(N) vs ESC(N+2i)	log2FoldChange
Arntl2	-5.100892672	Il1rl1	-6.738216326
Dppa3	-4.7457798	Ifi204	-6.471318674
Fgfbp1	-4.707068185	Hist1h2af	-6.300869165
Asz1	-4.501692953	Hist1h2ab	-6.282548293
Rhox6	-4.305043451	Six3	-6.056244019
Pax8	-4.215245081	Hist1h2ah	-5.942055856
Tlx2	-3.74283819	Lyl1	-5.820564772
Rhox9	-3.641117742	Foxd1	-5.785853676
Htr5a	-3.518594892	Cspg4	-5.756603348
Scml2	-3.210734702	Ciita	-5.737201546
Ssbp1	-2.807099248	E130201H02Rik	-5.570886787
Kcnh7	-2.581825597	Vax2	-5.429411181
Bola3	-2.402327684	Gsx2	-5.289906656
Nkx1-2	-2.158026734	Ppargc1a	-5.229722301
Maml2	-2.121521003	Prrx1	-5.138474839
Pcgf1	-2.063151027	Il2	-5.097405455
Spic	-2.020580285	Runx2	-5.02356484
Polr2k	-1.993942279	Nkx6-3	-4.892503849
Cdca7l	-1.989385405	Thrb	-4.856756093
Pole4	-1.984680027	Dmbx1	-4.752500097
Dazl	-1.826152402	Hist1h2ak	-4.729167173
Srebf1	-1.791585151	Hnf4a	-4.570382112
Msc	-1.74520948	Nfix	-4.417723426
Strap	-1.743722272	Foxc2	-4.405066445
Cited1	-1.712527114	Dlx1	-4.352995691
N6amt1	-1.684528649	Hoxc12	-4.311991378
Mycbp	-1.680866442	Pou2f3	-4.262816561
Setmar	-1.664967344	Zfp9	-4.020805903
Nthl1	-1.654606411	Sox9	-3.973740498
Zfp775	-1.65327122	Hoxc10	-3.93187411

Table: 12 Highly downregulated TFs during ESC(N+2i) vs ESC(G) and ESC(N) vs ESC(N+2i).

ESC(G)	FPKM	ESC(N+2i)	FPKM	ESC(N)	FPKM
Cnbp	767.469	Cnbp	422.5985	Cnbp	663.7435
H2afx	377.195	Ctnnb1	389.1285	Ctnnb1	383.8835
Hnrnpd	300.6955	H2afx	293.2985	H2afx	337.62
Ddx5	264.8385	H3f3a	242.753	Ddx5	334.0375
Hat1	260.399	Eed	234.774	Hnrnpd	316.041
Hmgn1	259.0845	Ddx5	222.539	Hat1	290.406
Ctnnb1	231.481	Hnrnpd	208.4575	Hmgn1	274.807
Cenpa	219.393	Parp1	191.4705	Hmga1	220.0755
Parp1	194.3935	Hmgn1	174.734	Ddx3x	188.3205
Hmga1	190.9275	Hmga1	174.5755	H3f3a	180.679
Ddx3x	145.8115	Brd2	165.4745	Eed	168.248
Eed	145.3335	Hat1	135.6885	Dek	165.4475
Ezh2	143.679	Dhx15	110.3165	Parp1	150.414
Wdr5	140.007	Phc1	109.673	Cenpa	146.0795
H3f3a	135.2135	Cbx1	109.136	Dhx15	131.3065
Cbx1	134.3955	Cenpa	108.5966	Brd2	127.3925
Cbx3	125.6145	Dek	106.264	Cbx1	120.192
Brd2	124.2615	Wdr5	102.09705	Phc1	111.3345
Dek	117.3825	H1f0	101.13485	Hdac2	102.6915
Dhx9	107.5505	Hdac2	100.1783	Tpr	101.3984
Dhx15	107.244	Ddx3x	95.15355	Supt16	98.6628
Phc1	106.8815	Chd4	86.8334	Dhx9	96.42325
Hdac2	104.2665	Kdm3a	78.25105	Smarcad1	93.1715
Supt16	94.53655	Tpr	75.34935	Wdr5	88.79705
Tpr	88.6878	Dnmt1	74.95955	Rcor2	87.56305
Kdm3a	86.7472	Supt16	70.08475	Cbx3	81.66
Bloc1s1	85.30475	Ezh2	68.6189	Dnmt3b	80.3429
Smarcad1	84.06675	Dhx9	68.4901	Ezh2	78.29315
Dnmt1	72.97425	Ehmt2	63.2971	Mta3	77.8363
Chd4	71.6596	Ctcf	62.8254	Chd4	77.38215

Table: 13 Top highly expressed ERs among ESC(G), ESC(N+2i) and ESC(N) state.

ESC(G)	FPKM	ESC(N+2i)	FPKM	ESC(N)	FPKM
Grid2	1.9872	Gata6	1.484538	Smc1b	1.21247
9130023H24Rik	1.019427	Runx2	1.1677869	Gata3	1.20035
		Gata1	1.049108		

Table: 14 Top uniquely expressed ERs among ESC(G), ESC(N+2i) and ESC(N) state.

ESC(N+2i) vs ESC(G)	log2FoldChange	ESC(N) vs ESC(N+2i)	log2FoldChange
Ciita	4.55658061	Ddx4	3.323297178
Bloc1s1	4.089982377	Dnmt3b	2.044160322
Gata2	3.337909266	Hat1	1.07121518
Gata6	3.188232751	Ddx3x	1.064938388
Phldb1	2.055669772	Dffb	1.025225407
Gata4	1.98358673		
H1f0	1.755412725		
D1Pas1	1.440579796		
Ptpn23	1.044246625		

Table: 15 Highly upregulated ERs during ESC(N+2i) vs ESC(G) and ESC(N) vs ESC(N+2i).

ESC(N+2i) vs ESC(G)	log2FoldChange	ESC(N) vs ESC(N+2i)	log2FoldChange
Grid2	-6.245621712	Bloc1s1	-6.313991178
Ddx4	-1.800070356	Ciita	-5.737201546
Cited1	-1.712527114	Runx2	-5.02356484
Bicd1	-1.543008122	Cbx4	-3.257167753
Cbx3	-1.34203263	Gata4	-2.084141079
Ezh2	-1.230886989	Aire	-1.852649715
Cnbp	-1.113233252	H1f0	-1.422857288
Hmgb3	-1.055697079	Cux1	-1.373511808
		Ptpn23	-1.169983268
		Bmi1	-1.117380254
		Mbd6	-1.028625382
		Hmgn2	-1.008161754

Table: 16 Highly downregulated ERs during ESC(N+2i) vs ESC(G) and ESC(N) vs ESC(N+2i).

ESC(G)	Normalized Count	ESC(N+2i)	Normalized Count	ESC(N)	Normalized Count
mmu-miR-148a-3p	532424.905	mmu-miR-292a-5p	245334.31	mmu-miR-148a-3p	150124.92
mmu-miR-292a-5p	80492.68	mmu-miR-148a-3p	59389.775	mmu-miR-292a-5p	93399.02
mmu-miR-182-5p	70547.34	mmu-miR-293-3p	58748.715	mmu-miR-7a-5p	43197.535
mmu-miR-183-5p	32450.66	mmu-miR-7a-5p	42954.175	mmu-miR-7a-5p	42837.76
mmu-miR-7a-5p	26017.16	mmu-miR-7a-5p	42492.92	mmu-miR-291a-3p	37274.68
mmu-miR-7a-5p	25819.91	mmu-miR-182-5p	35832.585	mmu-miR-182-5p	36679.97
mmu-miR-293-3p	15949.02	mmu-miR-183-5p	32567.44	mmu-miR-293-3p	34245.395
mmu-miR-291a-3p	13950.65	mmu-miR-292a-3p	30953.425	mmu-miR-21a-5p	18752.455
mmu-miR-127-3p	12303.795	mmu-miR-21a-5p	30431.17	mmu-miR-127-3p	14701.655
mmu-miR-25-3p	9771.39	mmu-miR-295-3p	27124.455	mmu-miR-183-5p	12372.86
mmu-miR-881-3p	5703.03	mmu-miR-293-5p	21351.575	mmu-miR-467b-5p	11399.67
mmu-miR-151-3p	5662.715	mmu-miR-294-3p	18319.955	mmu-miR-467b-5p	11399.67
mmu-miR-541-5p	4750.155	mmu-miR-25-3p	14434.95	mmu-miR-467b-5p	11399.67
mmu-miR-10b-5p	4438.665	mmu-miR-291a-3p	13174.79	mmu-miR-467b-5p	11399.67
mmu-miR-92a-3p	4379.115	mmu-miR-92a-3p	12974.465	mmu-miR-467b-5p	11399.67
mmu-miR-293-5p	4150.225	mmu-miR-26a-5p	12047.98	mmu-miR-467b-5p	11399.67
mmu-miR-21a-5p	4070.87	mmu-miR-26a-5p	12047.49	mmu-miR-467b-5p	11399.67

mmu-miR-1981-5p	3993.23	mmu-miR-541-5p	11082.76	mmu-miR-467b-5p	11399.67
mmu-miR-381-3p	3397.385	mmu-miR-381-3p	7812.16	mmu-miR-467b-5p	11399.67
mmu-miR-378a-3p	3051.98	mmu-miR-27b-3p	6853.47	mmu-miR-467b-5p	11399.67
mmu-miR-30d-5p	2701.405	mmu-miR-30d-5p	6114.04	mmu-miR-467b-5p	11399.67
mmu-miR-292a-3p	2635.27	mmu-miR-378a-3p	5778.545	mmu-miR-467a-5p	11399.445
mmu-miR-106b-3p	2585.075	mmu-miR-5099	5306.76	mmu-miR-467a-5p	11399.445
mmu-miR-409-3p	2525.415	mmu-miR-92a-3p	4741.99	mmu-miR-467a-5p	11399.445
mmu-miR-470-5p	2082.195	mmu-miR-409-3p	4335.28	mmu-miR-467a-5p	11399.445
mmu-miR-181d-5p	2081.05	mmu-let-7i-5p	4213.685	mmu-miR-467a-5p	11399.445
mmu-miR-425-5p	2004.3	mmu-miR-379-5p	3791.205	mmu-miR-467a-5p	11399.445
mmu-miR-140-3p	2001.215	mmu-miR-30a-5p	3787.86	mmu-miR-467a-5p	11399.445
mmu-miR-148b-3p	1977.16	mmu-miR-99b-5p	3771.51	mmu-miR-467a-5p	11399.445
mmu-miR-871-3p	1795.13	mmu-let-7f-5p	3664.55	mmu-miR-467a-5p	11399.445

Table: 17 Top highly expressed miRNAs among ESC(G), ESC(N+2i) and ESC(N) state.

ESC(N)-2i vs ESC(G)	log2FoldChange	ESC(N) vs ESC(N+2i)	log2FoldChange
mmu-miR-6238	10.00062215	mmu-miR-451a	7.811896977
mmu-miR-466i-5p	9.593229383	mmu-miR-302c-5p	7.103611396
mmu-miR-466i-5p	9.276701332	mmu-miR-592-5p	6.345628915
mmu-miR-3963	8.006930301	mmu-miR-1298-5p	6.25764339
mmu-miR-3968	7.285613245	mmu-miR-302d-5p	5.950829845
mmu-miR-1969	6.561838118	mmu-miR-302a-5p	5.849209671
mmu-miR-214-3p	6.478392933	mmu-miR-693-3p	5.680793048
mmu-miR-466i-5p	6.295657942	mmu-miR-702-5p	5.505331023
mmu-miR-450b-3p	6.264639137	mmu-miR-429-3p	5.499789632
mmu-miR-181a-1-3p	6.225236544	mmu-miR-1a-3p	5.356323128
mmu-miR-6239	6.171640826	mmu-miR-1a-3p	5.356323128
mmu-miR-21b	6.107402457	mmu-miR-3102-5p	5.353261488
mmu-miR-450a-1-3p	6.107402457	mmu-miR-302c-3p	5.250774983
mmu-miR-690	6.030865755	mmu-miR-483-5p	4.606602241
mmu-miR-199a-3p	6.015839058	mmu-miR-302d-3p	4.294936195
mmu-miR-199b-3p	6.015839058	mmu-miR-653-5p	4.174602744
mmu-miR-199a-3p	6.011540415	mmu-miR-367-3p	4.128552467
mmu-miR-199b-3p	6.011540415	mmu-miR-203b-3p	4.12748017
mmu-miR-199a-3p	6.011540415	mmu-miR-425-5p	3.964325211
mmu-miR-199b-3p	6.011540415	mmu-miR-3473f	3.904519129
mmu-miR-21c	5.977965422	mmu-miR-669b-3p	3.902536888
mmu-miR-214-5p	5.876796784	mmu-miR-302b-3p	3.8289948
mmu-miR-466i-5p	5.710714718	mmu-miR-466f-5p	3.76686253
mmu-miR-299b-3p	5.636626723	mmu-miR-466f-5p	3.76686253

Table: 18 Highly upregulated miRNAs during ESC(N+2i) vs ESC(G) and ESC(N) vs ESC(N+2i).

ESC(N)-2i vs ESC(G)	log2FoldChange	ESC(N) vs ESC(N+2i)	log2FoldChange
mmu-miR-451a	-7.232084113	mmu-miR-5106	-10.64952759
mmu-miR-743b-5p	-6.614514227	mmu-miR-3963	-9.736306307
mmu-miR-693-3p	-6.514160157	mmu-miR-690	-9.560107699
mmu-miR-743a-3p	-6.395485719	mmu-miR-6238	-9.403526419
mmu-miR-881-3p	-6.336594621	mmu-miR-211-5p	-9.183861793
mmu-miR-463-5p	-6.319065012	mmu-miR-214-5p	-8.064339767
mmu-miR-1981-5p	-6.30631603	mmu-miR-3968	-8.050793662
mmu-miR-878-5p	-6.18250108	mmu-miR-147-5p	-7.584734957
mmu-miR-6965-3p	-6.059342647	mmu-miR-196a-2-3p	-7.217619928
mmu-miR-3473f	-5.846204679	mmu-miR-6952-5p	-6.886055115
mmu-miR-871-5p	-5.743667099	mmu-miR-505-5p	-6.853514265
mmu-miR-465c-5p	-5.733798774	mmu-miR-450a-1-3p	-6.608389876
mmu-miR-465c-5p	-5.733798774	mmu-miR-211-3p	-6.54526476
mmu-miR-704	-5.682163898	mmu-miR-199a-3p	-6.520866016
mmu-miR-702-5p	-5.660386725	mmu-miR-199b-3p	-6.520866016
mmu-miR-470-5p	-5.465638044	mmu-miR-199a-3p	-6.516558597
mmu-miR-743b-3p	-5.373748098	mmu-miR-199b-3p	-6.516558597
mmu-miR-425-5p	-5.052806482	mmu-miR-199a-3p	-6.516558597
mmu-miR-7059-5p	-4.956832379	mmu-miR-199b-3p	-6.516558597
mmu-miR-429-3p	-4.948420314	mmu-miR-12199-3p	-6.495552998
mmu-miR-148a-3p	-4.940021503	mmu-miR-8114	-6.371836011
mmu-miR-542-5p	-4.934685375	mmu-miR-12184-5p	-6.346959811
mmu-miR-871-3p	-4.813142838	mmu-miR-450a-2-3p	-6.346959811
mmu-miR-465b-5p	-4.705162202	mmu-miR-6919-5p	-6.346959811

Table: 19 Highly downregulated miRNAs during ESC(N+2i) vs ESC(G) and ESC(N) vs ESC(N+2i).

Name	Hits	Pval	adj.Pval
PI3K-Akt signaling pathway	44	3.61E-09	1.44E-07
Focal adhesion	32	4.05E-09	1.44E-07
ECM-receptor interaction	20	4.33E-09	1.44E-07
Proteoglycans in cancer	29	1.48E-07	0.0000037
Ribosome	20	0.0000195	0.00039
Protein digestion and absorption	14	0.0000934	0.001556667
Rap1 signaling pathway	23	0.000251	0.003585714
Lysosome	16	0.00029	0.003625
Glutamatergic synapse	14	0.00123	0.01366667
MAPK signaling pathway	23	0.00252	0.0252
Wnt signaling pathway	15	0.00395	0.03142857
Glycolysis / Gluconeogenesis	9	0.00419	0.03142857
Platelet activation	14	0.0042	0.03142857
Amoebiasis	13	0.0044	0.03142857
VEGF signaling pathway	8	0.00779	0.04363636
Malaria	7	0.00863	0.04363636
Arrhythmogenic right ventricular cardiomyopathy (ARVC)	9	0.00893	0.04363636
Oxytocin signaling pathway	15	0.00924	0.04363636
Axon guidance	13	0.00926	0.04363636
Pathways in cancer	30	0.0095	0.04363636
GnRH signaling pathway	10	0.0096	0.04363636
Dilated cardiomyopathy	10	0.0096	0.04363636
Salmonella infection	9	0.0124	0.05391304
TNF signaling pathway	11	0.0169	0.0704
Galactose metabolism	5	0.0176	0.0704
Hypertrophic cardiomyopathy (HCM)	9	0.0195	0.075
Ovarian steroidogenesis	7	0.0209	0.07740741
Fructose and mannose metabolism	5	0.0225	0.08
NF-kappa B signaling pathway	10	0.0236	0.08
ErbB signaling pathway	9	0.024	0.08
Serotonergic synapse	12	0.0269	0.08677419
alpha-Linolenic acid metabolism	4	0.0303	0.0946875
Complement and coagulation cascades	8	0.0317	0.09485714
Glycosphingolipid biosynthesis - globo series	3	0.033	0.09485714
Regulation of actin cytoskeleton	17	0.0332	0.09485714
Leishmaniasis	7	0.0364	0.1010811
HIF-1 signaling pathway	10	0.0374	0.1010811
Maturity onset diabetes of the young	4	0.039	0.1026316

Table: 20 Upregulated KEGG Pathways in ESC(N+2i) vs ESC(G).

Name	Hits	Pval	adj.Pval
Metabolic pathways	104	6.51E-10	6.51E-08
Parkinson's disease	26	2.68E-09	1.34E-07
Non-alcoholic fatty liver disease (NAFLD)	25	3.72E-08	0.00000124
Oxidative phosphorylation	23	6.39E-08	1.5975E-06
Huntington's disease	28	8.72E-08	0.000001744
Spliceosome	20	0.00000253	4.21667E-05
Proteasome	11	0.00000392	0.000056
Alzheimer's disease	23	0.00000515	0.000064375
Peroxisome	13	0.0000889	0.000987778
Ribosome	18	0.000101	0.00101
Cysteine and methionine metabolism	9	0.00023	0.002090909
Terpenoid backbone biosynthesis	6	0.000586	0.004883333
Carbon metabolism	14	0.000858	0.0066
Steroid biosynthesis	5	0.00133	0.0095
Pentose phosphate pathway	6	0.00205	0.01366667
Central carbon metabolism in cancer	9	0.00291	0.01788235
Biosynthesis of amino acids	10	0.00304	0.01788235
Cardiac muscle contraction	10	0.00334	0.01855556
Primary immunodeficiency	6	0.00462	0.02431579
Alanine, aspartate and glutamate metabolism	6	0.00701	0.035
Purine metabolism	16	0.00774	0.035
Homologous recombination	5	0.00802	0.035
Fatty acid metabolism	7	0.00805	0.035
Synthesis and degradation of ketone bodies	3	0.0119	0.0476
Ubiquinone and other terpenoid-quinone biosynthesis	3	0.0119	0.0476
Propanoate metabolism	5	0.0124	0.04769231
Citrate cycle (TCA cycle)	5	0.0142	0.05259259
Proximal tubule bicarbonate reclamation	4	0.0164	0.05857143
Pyrimidine metabolism	10	0.0202	0.06965517
RNA transport	14	0.0236	0.07866667
Colorectal cancer	7	0.0261	0.08419355
Glycosphingolipid biosynthesis - globo series	3	0.0287	0.0896875
Pyruvate metabolism	5	0.0312	0.09454545
Biosynthesis of unsaturated fatty acids	4	0.0329	0.09676471
Aldosterone-regulated sodium reabsorption	5	0.0344	0.09828571
Ribosome biogenesis in eukaryotes	8	0.0354	0.09833333
Regulation of actin cytoskeleton	16	0.04	0.1081081
Protein export	4	0.0414	0.1089474

Table: 21 Downregulated KEGG Pathways in ESC(N+2i) vs ESC(G).

Name	Hits	Pval	adj.Pval
Proteasome	15	3.84E-08	0.00000384
Cell cycle	23	0.00000202	0.000101
Regulation of actin cytoskeleton	30	0.0000263	0.000876667
Arrhythmogenic right ventricular cardiomyopathy (ARVC)	15	0.0000381	0.0009525
Focal adhesion	26	0.000466	0.00932
Spliceosome	19	0.000573	0.00955
p53 signaling pathway	12	0.000743	0.00995
Pathways in cancer	41	0.000796	0.00995
Dilated cardiomyopathy	14	0.000956	0.01062222
Hypertrophic cardiomyopathy (HCM)	13	0.00186	0.0186
Nitrogen metabolism	5	0.00299	0.02564286
Bladder cancer	8	0.00322	0.02564286
Wnt signalling pathway	18	0.00336	0.02564286
Cardiac muscle contraction	12	0.00359	0.02564286
MAPK signalling pathway	27	0.00389	0.02593333
Ribosome biogenesis in eukaryotes	12	0.00488	0.0305
PI3K-Akt signalling pathway	34	0.00609	0.03582353
Hippo signalling pathway	18	0.00684	0.038
Hedgehog signalling pathway	8	0.00869	0.0445
RNA transport	19	0.0089	0.0445
Oocyte meiosis	14	0.0106	0.04714286
Parkinson's disease	17	0.0107	0.04714286
Melanoma	10	0.0117	0.04714286
Tight junction	16	0.0128	0.04714286
Signaling pathways regulating pluripotency of stem cells	16	0.0128	0.04714286
Peroxisome	11	0.0131	0.04714286
Rap1 signaling pathway	22	0.0132	0.04714286
Thyroid hormone signalling pathway	14	0.0132	0.04714286
Purine metabolism	19	0.015	0.0484375
HTLV-I infection	27	0.0151	0.0484375
TGF-beta signalling pathway	11	0.0155	0.0484375
Small cell lung cancer	11	0.0155	0.0484375
Glutathione metabolism	8	0.0192	0.05685714
ECM-receptor interaction	11	0.0197	0.05685714
Central carbon metabolism in cancer	9	0.0199	0.05685714
Axon guidance	14	0.0286	0.07944444
Thyroid cancer	5	0.0311	0.08405405

Table: 22 Upregulated KEGG Pathways in ESC(N) vs ESC(N+2i).

Name	Hits	Pval	adj.Pval
Lysosome	29	2.68E-10	2.68E-08
Ribosome	23	0.0000254	0.00127
PI3K-Akt signaling pathway	41	0.0000539	0.0014625
Other glycan degradation	7	0.0000585	0.0014625
Axon guidance	19	0.000342	0.00684
PPAR signaling pathway	14	0.00045	0.0075
Signaling pathways regulating pluripotency of stem cells	19	0.000964	0.01377143
VEGF signaling pathway	10	0.00346	0.03692308
Oxytocin signaling pathway	19	0.00397	0.03692308
Adipocytokine signaling pathway	11	0.00444	0.03692308
Pathways in cancer	38	0.00453	0.03692308
Osteoclast differentiation	16	0.00465	0.03692308
Peroxisome	12	0.0048	0.03692308
Alzheimer's disease	20	0.00641	0.04045
Rap1 signaling pathway	23	0.00666	0.04045
HIF-1 signaling pathway	14	0.00711	0.04045
TNF signaling pathway	14	0.0077	0.04045
Calcium signaling pathway	20	0.0077	0.04045
Glycosaminoglycan degradation	5	0.00793	0.04045
Ras signaling pathway	24	0.00809	0.04045
Sphingolipid metabolism	8	0.00859	0.04090476
Amino sugar and nucleotide sugar metabolism	8	0.00972	0.04418182
Glycosphingolipid biosynthesis - globo series	4	0.0115	0.04791667
Glycosphingolipid biosynthesis - ganglio series	4	0.0115	0.04791667
alpha-Linolenic acid metabolism	5	0.0169	0.0676
Focal adhesion	21	0.018	0.06923077
ECM-receptor interaction	11	0.0194	0.07185185
Fc epsilon RI signaling pathway	9	0.0215	0.07482759
NF-kappa B signaling pathway	12	0.0217	0.07482759
Thyroid hormone signaling pathway	13	0.0284	0.09466667
Platelet activation	14	0.0316	0.09578947
Phospholipase D signaling pathway	15	0.0326	0.09578947
Long-term depression	8	0.0334	0.09578947
B cell receptor signaling pathway	9	0.0354	0.09578947
Influenza A	17	0.0355	0.09578947
Hepatitis B	15	0.0363	0.09578947
Fatty acid metabolism	7	0.0364	0.09578947

Table: 23 Downregulated KEGG Pathways in ESC(N) vs ESC(N+2i).

Gene name	ESC(N+2i) I FPKM	ESC(N+2i) II FPKM	ESC(N) I FPKM	ESC(N) II FPKM
mmu-mir-136-5p	21.49	56.96	128.31	62.45
mmu-mir-291a-3p	11749.35	14600.23	50334	24215.36
mmu-mir-466m-3p	15.63	4.33	24.29	25.74
mmu-mir-466d-3p	59.59	82.92	207.94	180.12
mmu-mir-466a-3p	104.53	111.76	396.94	397.84
mmu-mir-297b-3p	37.12	42.54	124.9	96.55
mmu-mir-466n-5p	2.93	9.37	47.89	16.54
mmu-mir-466h-5p	45.91	39.66	222.73	131
mmu-miR-669m-5p	7.82	6.49	55.51	31.47
Nfatc2	2.972774675	1.821160454	0.984047	1.08364
Camk2b	11.29392964	7.21820726	4.44593	4.79621
Dkk1	1.489987174	0.561918881	0.331644	0.432652
Sfrp1	33.94937876	26.40117339	4.24395	3.88107

Table: 24 List of miRNAs and mRNAs involved in Wnt Signaling Pathways in ESC(N) vs ESC(N+2i).

Gene name	ESC(N+2i) I FPKM	ESC(N+2i) II FPKM	ESC(N) I FPKM	ESC(N) II FPKM
mmu-miR-205-5p	122.11	115.37	37.03	15.96
mmu-miR-1195	78.15	39.66	5.46	5.47
mmu-miR-335-3p	927.09	838.56	436.98	306.53
mmu-miR-495-3p	1088.28	705.89	133.03	87.84
mmu-miR-344d-3p	12.7	18.75	1.37	1.02
mmu-miR-541-5p	12889.41	9276.11	5057.25	4299.26
mmu-miR-466i-5p	21.49	6.49	1.76	0.31
mmu-miR-295-3p	22924.28	31324.63	3023.57	1764.22
mmu-miR-15a-5p	5.86	25.24	5.01	0.98
mmu-miR-17-5p	594.94	834.24	363.1	137.8
mmu-miR-433-3p	508.97	370.61	37.99	25.03
mmu-miR-26b-5p	1063.86	1666.31	178.02	109.48
mmu-miR-124-3p	40.05	62.01	15.02	10.58
Fgf5	0.841754833	0.53522923	5.01565	5.67011
Tgfb2	1.036069604	0.336826268	2.69073	2.36129
Srf	16.21310974	12.72924968	38.325	39.5353
Cacna2d1	0.526296534	0.637196363	3.61759	4.28694
Fgf13	0.996379458	0.960650361	4.1236	7.0347
Egfr	1.214521248	0.923765081	1.71651	1.70055
Mapkapk3	16.80485766	14.84301274	32.8754	36.3146
Myc	15.41795951	11.13818944	46.9333	44.1913

Table: 25 List of miRNAs and mRNAs involved in MAPK Signaling Pathways in ESC(N) vs ESC(N+2i).

Gene name	ESC(N+2i) I FPKM	ESC(N+2i) II FPKM	ESC(N) I FPKM	ESC(N) II FPKM
mmu-mir-15a-5p	0.09	3.22	5.86	25.24
mmu-mir-17-5p	46.66	119.16	594.94	834.24
mmu-mir-181a-5p	1.69	2	37.12	122.58
mmu-mir-495-3p	27.07	49.61	1088.28	705.89
mmu-mir-450b-3p	0.18	0	17.58	23.79
mmu-mir-34b-5p	1.96	4.72	21.49	103.83
mmu-mir-665-3p	2.55	4.72	21.49	45.43
mmu-mir-363-5p	29.85	23.66	415.19	310.77
mmu-miR-125a-3p	1	1.43	31.26	25.24
mmu-mir-26a-5p	591.41	831.31	12011.16	12083.82
Bcat1	647.521	643.226	87.86434826	98.71299675
Polr1b	56.9409	61.9301	27.17641929	28.20047576
Alg6	8.03206	9.11362	6.310118043	3.927248772
B3gntl1	4.6313	4.92805	3.076456241	2.732372441
Acsl3	33.6263	32.8862	11.89430856	9.733516176
Acsl4	9.11232	8.85944	5.923551359	3.347362125
Ak4	37.1081	33.2903	12.81601742	10.87203815
Atp5j2	384.179	396.742	294.3472116	336.4583345
Atp6v0e2	15.1815	13.8456	3.707310129	2.877367826
Polr3k	15.2731	14.5608	8.146944102	8.144059691
Mat2a	267.074	251.513	98.76308561	54.30877847
Ndufa4	1298.08	1243.41	212.2679189	223.3689987
Cyp51	21.7994	21.6578	9.465470496	8.351321969
Dhcr7	32.5167	36.2078	10.33402166	5.859866753
Gls	17.9906	14.9491	7.001720883	3.043093077
Psat1	334.588	325.317	106.4781439	121.2904423

Table: 26 List of miRNAs and mRNAs involved in Metabolic Pathways in ESC(N+2i) vs ESC(G).

	ESC(N)1	ESC(N)2	NPC1	NPC2	CN1	CN2
ESC(N)1	1	0.9386419	0.8996591	0.899727	0.8604449	0.861289
ESC(N)2	0.9386419	1	0.9009399	0.9020787	0.8615175	0.8627291
NPC1	0.8996591	0.9009399	1	0.9337942	0.8886294	0.8854112
NPC2	0.899727	0.9020787	0.9337942	1	0.8877946	0.882937
CN1	0.8604449	0.8615175	0.8886294	0.8877946	1	0.93955
CN2	0.861289	0.8627291	0.8854112	0.882937	0.93955	1

Table: 27 Neurogenesis mRNAseq Spearman's correlation statistics between samples and replicates with p-value < 2.2e-16.

	ESC(N)1	ESC(N)2	NPC1	NPC2	CN1	CN2
ESC(N)1	1	0.9465168	0.7711927	0.7674678	0.707778	0.7084993
ESC(N)2	0.9465168	1	0.760938	0.7690785	0.705418	0.7005638
NPC1	0.7711927	0.760938	1	0.9497777	0.8920891	0.8911706
NPC2	0.7674678	0.7690785	0.9497777	1	0.8892873	0.8876061
CN1	0.707778	0.705418	0.8920891	0.8892873	1	0.9491964
CN2	0.7084993	0.7005638	0.8911706	0.8876061	0.9491964	1

Table: 28 Neurogenesis miRNAseq Spearman's correlation statistics between samples and replicates p-value < 2.2e-16.

Sample name	Raw reads	GC content (%)	Q20 (%)	Q30 (%)	Mapped
ESC(N)I	2,25,23,225	49.73%	97.7	93.59	89.70%
ESC(N)II	2,36,69,902	50.00%	97.75	93.69	89.70%
NPCI	2,31,39,901	49.80%	97.53	93.25	89.60%
NPCII	1,95,23,776	56.29%	99.50	98.61	89.40%
CNI	2,89,88,882	51.42%	97.71	93.66	87.70%
CNII	2,04,08,746	51.31%	97.62	93.53	88.40%

Table: 29 Neurogenesis mRNAseq raw read sequences quality and mapping percentages.

Sample name	Raw reads	GC content (%)	Q20 (%)	Q30 (%)	Mapped
ESC(N)I	1,90,73,329	49.39%	99.54	98.73	69.39%
ESC(N)II	2,04,56,334	49.76%	99.52	98.74	73.17%
NPCI	2,02,59,269	54.97%	99.45	98.47	89.60%
NPCII	1,95,23,776	56.29%	99.50	98.61	89.40%
CNI	2,31,72,410	49.22%	99.50	98.62	94.36%
CNII	1,89,34,399	49.35%	99.76	99.38	93.56%

Table: 30 Neurogenesis miRNAseq raw read sequences quality and mapping percentages.

ESC(N)	FPKM	NPC	FPKM	CN	FPKM
Gm21060	164.9035	Gm6272	99.17	Neurog2	86.1985
Gm15181	75.1095	Gm11382	68.814	Nhlh1	75.2854
Gm48217	63.6475	Gm43078	66.33	Scrt2	68.56845
Gm26448	56.8486	Gm20305	39.5992	Actl6b	61.6403
Gm28625	34.49195	Gm15079	30.15435	Lypd1	56.9286
Igkv13-78-1	22.0071	Gm49146	20.296909	Gsx1	52.713
Gm48323	21.08705	Gm14200	12.33695	Nxph4	50.43485
Gm29284	15.814155	Gm48017	12.21955	Resp18	47.91025
Gm43655	9.7353	AC122821.1	10.22828	C130021I20Rik	44.23255
Gm13193	8.84895	Gm43013	8.276125	Ralyl	43.5731
Gm47126	7.47424	Gm14298	8.23495	Scrt1	42.09575
Gm13195	6.6364	Gm14915	7.16805	Otp	39.5467
Gm14282	5.68165	AC119957.1	6.0033	Caly	35.7351
Trav5-2	5.67625	Gm32460	5.44505	Cbln2	32.9731
AC154762.1	3.7021	Gm32950	5.44505	Th	29.7692
Gm37934	3.65622	Gm3054	5.273755	Syt4	29.3593
Gm10772	3.41961	Gm49258	4.054045	Myt1	29.12875
Gm21863	3.25542	Gm13396	3.978705	Lbx1	27.42565
Gm13165	2.751195	Gm25363	3.778625	Fbll1	27.16485
Ighv8-13	2.541165	Gm48494	3.5272025	Msx3	25.0947

Table: 31 Top uniquely expressed mRNAs among ESC(N), NPC and CN state.

ESC(N)	FPKM	NPC	FPKM	CN	FPKM
Gm13392	113648.35	Gm13392	39242.3	Gm13392	179209.45
Rpl39-ps	84643.9	Rpl39-ps	19594.7	CT010467.1	18946.5
Gm14303	23255.5	Gm14303	11542.8	Gm14303	18364.45
Actb	18990.35	Actb	9747.275	Tpi-rs11	13525.75
Gm10443	13083.05	Rpl38-ps2	8839.29	Actb	12008.45
Rpl38-ps2	9628.975	Gm10443	7328.625	Rpl38-ps2	9836.27
Tma7-ps	8705.72	Gm28661	7272.445	Gm10443	9336.54
Gm28661	8295.975	Gm29216	6205.2	Gm29216	7922.72
Gm4149	7405.285	Tma7-ps	6156.905	Gm28437	6562.105
Rps27rt	7324.655	AC123659.1	5642.535	Gm4149	5954.95
Gm7536	7058.68	Gm10925	5641.965	Rpl37rt	5900.64
Rpl36-ps12	6860.18	Gm28437	4784.68	Rpl13a	5811.31
Gm43110	6765.33	Gm9794	4533.335	Rps26-ps1	5463.415
Gm28437	6573.49	Gm10222	4507.495	Gm28661	5274.855
Rpl37rt	6153.36	Rps26-ps1	4316.055	Gm14539	5274.305
Gm10925	6144.94	Gm28438	4221.45	Gm10925	5157.095
Rps26-ps1	6135.48	Gm7536	4169.225	Rps27rt	5123.405
Rps12-ps3	5946.36	Gm43110	4164.725	AC123659.1	5007.31
Gm29216	5919.135	Tpi-rs11	4000.295	Rpl36-ps12	4809.13
Rpl13a	5882.22	Rpl37rt	3976.795	Rpl31-ps8	4654.625

Table: 32 Top highly expressed mRNAs among ESC(N), NPC and CN state.

NPC vs ESC(N)	log2 Fold change	CN vs NPC	log2 Fold change
Nepn	11.82930484	Gm28902	10.40896568
Fat4	11.62769285	Prdm13	9.634767526
Svep1	11.37348149	Lbx1	8.883461742
Dnm3os	10.73616816	Neurod2	8.207540195
Ttr	10.1280164	Gm38604	7.968790924
Postn	9.586714188	Helt	7.886765529
Angpt1	9.545090311	Resp18	7.731691851
Abca9	9.340367194	1110015O18Rik	7.481724378
Dbx1	8.995238475	Ralyl	7.459548148
Col6a3	8.745760489	1700010K23Rik	7.457629704
Zic4	8.742186567	Neurod6	7.44440337
Prrx1	8.677562765	2610028E06Rik	7.432725539
Cavin2	8.530244161	Pcdha1	7.417857735
Myh8	8.516279373	Gm13872	7.388321183
Tenm2	8.201283048	Drd1	7.374407942
Cartpt	8.172957012	AW047730	7.373248388
Slitrk6	8.023519024	Gm30177	7.319833284
Cdh11	8.013695839	Neurog2	7.261092646
Dkk2	8.009488661	Rgs13	7.196286002
Fibin	7.930512672	Nms	7.16070191
Col3a1	7.881948821	Gm20647	7.135321828
Rtl3	7.880048697	Cdr1	7.082748061
Nhlh2	7.823804524	Actl6b	7.035505957
Fezf2	7.743196024	Tlx1	6.985431642
Adamts12	7.70216742	AA387200	6.958931844
Fbn2	7.690547474		6.905364883
Gria2	7.688464526	Lamp5	6.896936018
Enpep	7.668281297	AI849053	6.880469473
Aqp1	7.62762595	Lingo2	6.844246973
Bace2	7.592281557	Th	6.77059888
Slc28a2	7.561161422	Gsx1	6.745633549
Dcn	7.528969994	Gjb6	6.725948719
Lum	7.506888834	Fibcd1	6.716778212
Adamtsl3	7.491830543	Gm3764	6.713770234

Table: 33 Highly upregulated mRNAs during NPC vs ESC(N), CN vs NPC.

NPC vs ESC(N)	log2 Fold change	CN vs NPC	log2 Fold change
Gm11238	-12.04617347	Spink1	-12.01803375
Gm428	-11.95040757	Dppa4	-11.98105285
Gm6346	-11.16780078	Zfp988	-11.71391649
Gm4301	-10.96198239	Zscan10	-11.51160047
Gm4312	-10.94010239	Zfp987	-11.48312396
Gm4308	-10.93436305	Dppa2	-11.32106355
Gm11237	-10.80322356	Gm8935	-11.29156326
Gm11236	-10.49814179	AC158554.1	-11.24007565
Gm3176	-10.42571462	L1td1	-11.12531155
Gm11239	-10.30201633	Morc1	-11.10123094
Gm5039	-10.25346835	Zfp978	-11.02166771
Gm4307	-9.75056019	Grb7	-11.01928006
Gm4303	-9.748409028	2410141K09Rik	-11.01589351
Gm4305	-9.728540673	4930447C04Rik	-10.87857189
Gm15280	-9.582276172	Pramef12	-10.74739136
Gm6351	-9.551489168	Gm17067	-10.68821899
Gm7982	-9.534652248	Zfp989	-10.67820937
Zscan4a	-9.446144354	Slfn9	-10.5897446
Gm8723	-9.389293581	Fthl17b	-10.53130604
Gm8701	-9.345712149	Gm10324	-10.48959776
AC166159.1	-9.301811229	Gm6798	-10.44933934
Gm21542	-9.285631543	Tdh	-10.42297812
Tdpoz4	-9.165174652	Apob	-10.41945302
Gm7942	-9.136214637	Platr15	-10.40584284
Zscan4f	-9.13430896	Gm17232	-10.40274982
Gm21033	-9.073088224	AC154506.1	-10.36489986
Gm3147	-9.008211327	Gm9788	-10.3595559
Gm8711	-8.992432144	Fthl17c	-10.19171437
Gm9343	-8.991606058	Utf1	-9.906335301
Gm6763	-8.989803232	Sohlh2	-9.854949192
Gm21034	-8.867523942	Tuba3a	-9.775755786
Gm11543	-8.863106325	Zfp981	-9.741792624
Gm42662	-8.759769871	Fgf4	-9.734561682
Zscan4e	-8.698488659	Pla2g10	-9.666242281

Table: 34 Highly downregulated mRNAs during NPC vs ESC(N), CN vs NPC.

ESC(N)	FPKM	NPC	FPKM	CN	FPKM
Zfp595	1.37283	A630089N07Rik	3.30745	Neurog2	86.1985
Prdm1	2.275075	Ankrd1	2.31751	Nhlh1	75.2854
Cebpa	2.75912	Eya4	2.752865	Scrt2	68.56845
Hist2h3c1	1.72654	Foxq1	7.509295	Gsx1	52.713
Hr	2.562445	Gata6	13.85395	Neurog1	47.55015
Mesp2	1.351855	Glis3	3.827565	Scrt1	42.09575
Stat4	3.68407	Hif3a	1.83706	Otp	39.5467
Nlrp9b	3.775465	Hnf1b	4.213705	Myt1	29.12875
Ankrd22	2.276485	Ikzf2	4.253625	Lbx1	27.42565
Zfp109	1.40674	Myb	9.865115	Msx3	25.0947
1700020N01Rik	1.61594	Myog	2.29046	Rfx4	24.91385
Nfam1	1.344045	Nfatc2	1.872055	Lmo3	22.4105
		Nostrin	4.92425	Phox2b	22.3244
		Nr3c2	2.317687	Pax3	21.54965
		Pitx1	1.96876	Npas3	20.25564
		Prdm16	2.366305	Shh	17.6658
		Ryr2	1.91937	Tlx1	17.4849
		Scx	3.15553	Dlx2	12.6688
		Smc1b	3.260695	Dlx1	12.21825
		Snai2	5.395395	Prdm8	11.8969
		Sox7	2.125725	Tox	10.81415
		Svep1	6.680175	Cxxc4	10.64935
		Tbx20	2.749465	Eya2	10.42225
		Vgll2	2.02779	Barhl2	9.919805
		Zfp677	2.388985	Heyl	9.79803
				Pou2f2	9.58205

Table: 35 Top uniquely expressed TFs among ESC(N), NPC and CN state.

ESC(N)	FPKM	NPC	FPKM	CN	FPKM
Hsp90ab1	2171.12	Hsp90ab1	1414.995	Gm13392	179209.45
Pou5f1	1831.89	Pfn1	1401.815	CT010467.1	18946.5
Pfn1	1705.8	Ybx1	808.494	Gm14303	18364.45
Ybx1	1658.38	Prmt1	720.709	Tpi-rs11	13525.75
Prmt1	1085.74	Hnrnpu	514.3395	Actb	12008.45
Trim28	1037.11	Ddx5	507.425	Gm10443	9336.54
Pcbp2	776.488	H3f3b	460.976	Gm29216	7922.72
Hnrnpu	766.742	Phb2	450.167	Gm28437	6562.105
Ranbp1	714.3675	Pcbp2	432.474	Gm4149	5954.95
Cnbp	663.7435	Ncl	426.74	Rpl37rt	5900.64
Ncl	607.561	Hmga1b	403.0545	Rpl13a	5811.31
Hmga1b	584.9885	Pou5f1	393.825	Gm28661	5274.855
Banf1	536.4215	Tardbp	383.8755	Gm14539	5274.305
Ifitm3	510.8625	Cnbp	375.047	Gm10925	5157.095
Marcksl1	472.0435	Ranbp1	372.014	Rps27rt	5123.405
Utf1	449.119	Trim28	358.203	AC123659.1	5007.31
Tardbp	436.56	Banf1	349.5375	Gm7536	4469.695
Phb2	429.6285	Marcksl1	339.0135	Rpl32	4357.775
Set	428.8835	Fus	320.445	Gm11808	4334.8
C230062I16Rik	421.464	Tcp1	313.226	Gm9794	4287.725
Zfp42	410.845	Cdk4	304.9545	Gm28438	3979.885
Pcbp1	387.219	Ctnnb1	301.2115	Gm43110	3819.845
Ctnnb1	383.8835	Nasp	293.312	Gm16418	3788.965
Cdk4	383.469	Sap18b	291.263	Gm27544	3618.075
Snrpb	382.1945	Cct4	283.634	Gm9844	3438.47
Gpx4	378.146	Dnmt3b	272.9205	Tuba1a	3360.88

Table: 36 Top highly expressed TFs among ESC(N), NPC and CN state.

NPC vs ESC(N)	log2 Fold Change	CN vs NPC	log2 Fold Change
Svep1	11.37348149	Lbx1	8.883461742
Dbx1	8.995238475	Neurod2	8.207540195
Zic4	8.742186567	Helt	7.886765529
Prrx1	8.677562765	Neurod6	7.44440337
Nhlh2	7.823804524	Neurog2	7.261092646
Ifi204	7.275708417	Tlx1	6.985431642
Zic1	6.999865543	AI849053	6.880469473
Otp	6.891521403	Gsx1	6.745633549
Vax1	6.765900193	Scrt2	6.639951438
Six3	6.595442425	Nhlh1	6.587701171
Zmiz1os1	6.432897024	Ferd3l	6.440091321
Nr2f2	6.321438466	Scrt1	6.239232753
Six3os1	6.225067878	Grm1	6.212065674
Pcdhb16	6.125929009	Onecut3	6.148927589
Mybph	6.090024839	Otp	6.103364033
Zbtb16	6.086927122	Irx4	6.071208313
Lmo3	6.079905187	Hmx3	5.946871889
Zfp9	5.98239386	Phf24	5.852552474
Shh	5.847798375	Fev	5.756300141
Myt1	5.770218942	Ptf1a	5.731325414
Otx1	5.76827118	Neurog1	5.620514862
Zeb2os	5.764114629	Msx3	5.575069487
Pax5	5.721051402	Prdm8	5.554567338
Ntf3	5.612011164	Prrxl1	5.542450029
Pou3f2	5.564887168	Phox2b	5.538216476
Foxa2	5.375526748	Myt1	5.534709345
Nr4a3	5.200178548	Gsx2	5.341862282
Pcdh10	5.192461144	Lmo3	5.309060911
Pou3f3	5.142706601	Lhx8	5.245706659
Lbh	5.066655557	Lhx1os	5.219659544
Phox2b	4.906371598	Barhl2	5.176508817
Hand2	4.890546235	Lmo1	5.016164688
Sox1ot	4.819714005	Lrfn5	4.913207512
Ebf3	4.810266918	Hist3h2ba	4.752429221

Table: 37 Highly upregulated TFs during NPC vs ESC(N), CN vs NPC.

NPC vs ESC(N)	log2 Fold Change	CN vs NPC	log2 Fold Change
Cdx1	-8.343412189	Zscan10	-11.51160047
Zfp352	-7.206038148	Dppa2	-11.32106355
Nkx6-3	-6.250695039	Zfp978	-11.02166771
Ankrd22	-5.41174276	2410141K09Rik	-11.01589351
Bcl3	-5.027703216	Utf1	-9.906335301
Tcl1	-4.542384068	Sohlh2	-9.854949192
Nr0b1	-4.228349881	Pou5f1	-9.530934837
Fgfbp1	-4.220798599	Foxh1	-9.381053385
Cdx4	-4.170906282	Smc1b	-8.682917107
Sp110	-3.96446517	Rbmxl2	-8.641322803
Gbx2	-3.802456418	Dnmt3l	-8.50668901
Nr5a2	-3.793093952	Mybpc3	-8.472410298
Msc	-3.713199117	Zfy1	-8.463000984
Snai3	-3.581732149	Rab25	-8.325171031
Prdm14	-3.418404071	Sox7	-7.95961184
Zfp977	-3.414851978	Dazl	-7.947066722
Irx4	-3.263643684	Tbx4	-7.939453692
Spz1	-3.256661941	Aire	-7.8466728
Pou4f2	-3.238922652	Rex2	-7.821389678
Sox18	-3.160119496	Syce1	-7.810168143
Egr4	-3.114173832	Hnf1b	-7.426864279
Rbmxl2	-3.044041919	Hnf4a	-7.367641686
Tcf15	-3.018412253	Nanog	-7.292060519
Hoxa1	-3.00425546	Tcl1	-7.289929632
Spic	-2.989795712	Hand1	-7.214335803
Hr	-2.905750038	Elf3	-7.125626507
Id1	-2.824423724	Tcf23	-7.068654489
Dnmt3l	-2.799522126	Zfp42	-7.060890114
D1Pas1	-2.793938356	Ripk4	-6.759871133
Mesp2	-2.79103921	Pycard	-6.722303216
Junb	-2.768494153	Nostrin	-6.682673449
Dppa2	-2.715061835	Hsf2bp	-6.648898662
Syce1	-2.683297147	Rhox9	-6.645860296
Ifitm3	-2.652863613	Grhl2	-6.545825143

Table: 38 Highly downregulated TFs during NPC vs ESC(N), CN vs NPC.

ESC(N)	FPKM	NPC	FPKM	CN	FPKM
Prdm1	2.562445	Sox7	2.125725	Actl6b	61.6403
Hr	2.275075	Smc1b	3.260695	Syt4	29.3593
Nkx2-5	1.845045	Gata6	13.85395	Pax3	21.54965
		Myb	9.865115	Prdm13	14.7582
		Nfatc2	1.872055	Vit	12.3673
		Shroom4	4.19035	Tox	10.81415
		Helz2	1.90437	Sox5	9.279275
		Svep1	6.680175	Prdm12	6.58821
		Prdm16	2.366305	Gata3	6.311375
				Nap1l3	6.255495
				Tal1	5.99049
				Sox14	5.396465
				Pygo1	5.08857
				Dpf3	4.820705
				Npcd	4.130705
				Hdac9	4.028555
				Wdr5b	3.26856
				Ccdc57	2.883775
				Tspyl5	2.79227
				Asxl3	2.246635
				Sox8	2.104395
				Brdt	1.903295
				Hist1h1e	1.69947
				Ccdc96	1.653935

Table: 39 Top uniquely expressed ERs among ESC(N), NPC and CN state.

ESC(N)	FPKM	NPC	FPKM	CN	FPKM
Ybx1	1658.38	Ybx1	808.494	H3f3b	691.7375
Trim28	1037.11	Snrpert	515.9825	Snrpert	659.4315
Snrpert	997.312	Ddx5	507.425	Ybx1	631.7025
Cnbp	663.7435	H3f3b	460.976	Cops9	543.814
Hmga1b	584.9885	Hmga1b	403.0545	Ctnnb1	458.3785
Banf1	536.4215	Cnbp	375.047	Ddx5	444.4205
Mkrn1	531.4405	Trim28	358.203	Supt4b	430.02
Snrpe	512.9755	Banf1	349.5375	H2afv	428.025
Supt4b	430.0575	Tcp1	313.226	Snrpb	344.2175
Set	428.8835	Supt4b	301.3175	Banf1	310.781
Ctnnb1	383.8835	Ctnnb1	301.2115	Pebp1	309.5575
Snrpb	382.1945	Snrpe	292.9895	H2afj	295.276
Ddx39	372.719	Rhox5	291.339	Cnbp	293.7925
Sap18b	339.2525	Sap18b	291.263	Snrpe	284.867
H2afx	337.62	Dnmt3b	272.9205	Tcp1	284.4175
Ddx5	334.0375	Snrpb	267.9995	H3f3a	281.635
Tcp1	316.828	Psmc5	257.5015	Trim28	245.5535
Hnrnpd	316.041	H2afj	230.522	H1f0	232.5415
Hat1	290.406	Hnrnpd	229.767	Ddx39b	232.1175
H2afj	283.232	Ddx39b	227.2215	H2afx	226.619
Trp53	279.6695	Set	221.1285	Sap18b	226.4905
Hmgn1	274.807	Cops9	219.1315	Sin3b	211.4745
Psmc5	266.252	Morf4l2	213.5335	Psmc5	204.6425
Morf4l2	265.3735	Hmgn1	206.475	Trp53i11	204.616
Cops9	261.5905	Mybbp1a	199.8165	Ube2b	199.7365
H3f3b	260.3725	Mkrn1	194.756	H2afy	195.119
Ddx39b	258.4085	Rbbp7	192.4775	Snrpn	193.932
Rbbp7	247.3275	Ddx3x	176.724	Cadm4	193.738
Dnmt3l	236.9515	Cbx1	170.596	Mbd3	187.1805
H2afz	230.788	H2afx	158.6905	Morf4l2	173.798
Sox2	223.259	Hat1	158.227	Hnrnpd	169.992
Hmga1	220.0755	Ddx39	157.9165	H1fx	169.218
Ruvbl2	215.639	Phc1	151.7955	Chmp2a	167.7965
Dhx16	212.736	Acin1	150.9	Mllt11	166.571
Pebp1	207.641	Trp53	149.3945	Hmga1b	163.3925

Table: 40 Top highly expressed ERs among ESC(N), NPC and CN state.

NPC vs ESC(N)	log2 Fold Change	CN vs NPC	log2 Fold Change
Svep1	11.37348149	Prdm13	9.634767526
Mybph	6.090024839	Actl6b	7.035505957
Pax5	5.721051402	Phf24	5.852552474
Kdm5d	5.487209034	Syt4	5.166599545
Vit	5.483012052	Grid2	4.960942626
Sox9	5.215794711	Prdm12	4.695010485
Sox1ot	4.819714005	Tal1	4.45228279
Pax3	4.304795778	Nap1l5	4.278244674
Sox17	4.303153074	Gata2	3.964462367
Sox14	4.261568957	Pax3	3.844496935
Ddx60	4.199970906	Hist3h2a	3.841166581
Gata4	4.151756164	Tspyl4	3.740508323
Gata6	4.103767093	Mllt11	3.586645882
Prdm6	3.971740778	Sox10	3.546467371
Smarca2	3.791883873	Hdac11	3.384156213
Runx2	3.685182649	Dpf3	3.368559756
Helz2	3.548530602	Sox21	3.368403247
Sox6	3.35138213	Sox18	3.338687602
Mybpc3	3.325094521	Hist1h1e	3.255949932
Cbx4	3.227958914	Tox3	3.114040346
Tox2	3.133148344	Phf21b	3.083959623
Sox1	2.957045099	Vit	3.062297344
Myb	2.88463166	Tox	3.060053507
Tox3	2.768486652	Soga3	2.891638277
Ifih1	2.71810631	Sox9	2.844496876
Ing4	2.650456803	H3f3aos	2.843446029
Phf21b	2.513784669	Nova1	2.785101082
Phf2	2.508908066	Nap1l3	2.652374801
Pygo1	2.464684383	Gata3	2.604924358
Cited1	2.438657813	Sox8	2.510587244
Sox7	2.367178864	Sox1ot	2.450207575
Syne2	2.334451677	Hdac9	2.412227407
Chd3	2.266586264	H2afv	2.358524903
Cbx8	2.259106711	Pygo1	2.33976821

Table: 41 Highly upregulated ERs during NPC vs ESC(N), CN vs NPC.

NPC vs ESC(N)	log2 Fold Change	CN vs NPC	log2 Fold Change
Nkx2-5	-3.491635855	Smc1b	-8.682917107
Prdm14	-3.418404071	Dnmt3l	-8.50668901
Sox18	-3.160119496	Mybpc3	-8.472410298
Hr	-2.905750038	Rhox5	-8.058608642
Dnmt3l	-2.799522126	Sox7	-7.95961184
D1Pas1	-2.793938356	Aire	-7.8466728
Syce1	-2.683297147	Syce1	-7.810168143
Ddx4	-2.637236264	4930548H24Rik	-7.775621874
Cited4	-2.63329489	Ddx4	-7.530577124
Sox15	-2.629223479	Prdm14	-6.79206635
Zbtb32	-2.407563725	Ddx60	-6.2183714
Sox2	-2.290775717	Phf11d	-5.805661497
Gata2	-2.254655486	Piwil1	-5.313988765
Hspbap1	-2.146770177	Piwil2	-5.269326821
Piwil2	-2.132257099	Gata6	-5.175668467
Mybl2	-2.103884002	Gata4	-5.106423798
Phf11d	-2.056145635	Kdm5d	-5.080719858
Prdm1	-1.958871565	Sox17	-5.080131495
Hist1h3g	-1.920042506	Jade2	-4.812116198
Eed	-1.891323018	Helz2	-4.659935254
Mycn	-1.849600188	Cbx7	-4.139517419
Mov10l1	-1.784127802	Crocc2	-3.744323304
Trmt11	-1.722737142	Hells	-3.653693241
Ube2a	-1.610323505	Svep1	-3.600625858
Zfp961	-1.605573146	Ccdc18	-3.49378257
Ddx25	-1.590648261	Phf11c	-3.455531464
Hist1h2ae	-1.513693277	Myb	-3.388439025
Myc	-1.508177867	Dnajc22	-3.385999076
Mtf2	-1.437040062	Hnf1a	-3.322790523
Terf1	-1.432152152	Ddx58	-3.241604747
Cenpt	-1.422619003	Mov10	-3.226736796
Ash2l	-1.42171788	Pax5	-3.221800115
Mgmt	-1.400256156	Rad50	-3.11840574
Wdr5b	-1.396558212	Zbtb32	-3.031892734

Table: 42 Highly downregulated ERs during NPC vs ESC(N), CN vs NPC.

ESC(N)	Normalize count	NPC	Normalize count	CN	Normalize count
mmu-miR-297c-5p	0.39	mmu-miR-7069-3p	0.375	mmu-miR-873a-3p	2.975
mmu-miR-466l-5p	0.365	mmu-miR-1971	0.29	mmu-miR-3093-5p	2.895
mmu-miR-6912-5p	0.22	mmu-miR-7076-5p	0.185	mmu-miR-344g-5p	2.29
mmu-miR-6935-3p	0.2			mmu-miR-6977-5p	1.63
mmu-miR-7025-3p	0.165			mmu-miR-181b-2-3p	1.545
mmu-miR-7080-5p	0.16			mmu-miR-9b-5p	1.03
mmu-miR-23a-5p	0.155			mmu-miR-128-2-5p	0.81
mmu-miR-1247-5p	0.115			mmu-miR-9b-5p	0.755
mmu-miR-5108	0.09			mmu-miR-9b-5p	0.755
mmu-miR-669i	0.09			mmu-miR-7093-5p	0.635
mmu-miR-682	0.09			mmu-miR-544-5p	0.585
mmu-miR-7020-5p	0.09			mmu-miR-879-5p	0.585
mmu-miR-3109-5p	0.085			mmu-miR-133a-5p	0.565
mmu-miR-5619-3p	0.085			mmu-miR-133a-5p	0.565
mmu-miR-669k-3p	0.085			mmu-miR-6928-5p	0.545
mmu-miR-7037-5p	0.085			mmu-miR-344f-5p	0.525

Table: 43 Top uniquely expressed miRNAs among ESC(N), NPC and CN state.

ESC(N)	Normalized count	NPC	Normalized count	CN	Normalized count
mmu-miR-148a-3p	150124.92	mmu-miR-148a-3p	289269.355	mmu-miR-148a-3p	124827.07
mmu-miR-292a-5p	93399.02	mmu-miR-291a-3p	44219.785	mmu-miR-9-5p	106556.83
mmu-miR-7a-5p	43197.535	mmu-miR-182-5p	38926.45	mmu-miR-9-5p	106554.92
mmu-miR-7a-5p	42837.76	mmu-miR-292a-5p	28842.6	mmu-miR-9-5p	106554.795
mmu-miR-291a-3p	37274.68	mmu-miR-467b-5p	15781.275	mmu-miR-103-3p	46931.51
mmu-miR-182-5p	36679.97	mmu-miR-467b-5p	15781.275	mmu-miR-103-3p	46762.78
mmu-miR-293-3p	34245.395	mmu-miR-467b-5p	15781.275	mmu-miR-7a-5p	39779.185
mmu-miR-21a-5p	18752.455	mmu-miR-467b-5p	15781.275	mmu-miR-7a-5p	39612.92
mmu-miR-127-3p	14701.655	mmu-miR-467b-5p	15781.275	mmu-miR-30a-5p	33330.33
mmu-miR-183-5p	12372.86	mmu-miR-467b-5p	15781.275	mmu-miR-379-5p	16453.565
mmu-miR-467b-5p	11399.67	mmu-miR-467b-5p	15781.275	mmu-miR-127-3p	15190.025
mmu-miR-467b-5p	11399.67	mmu-miR-467b-5p	15781.275	mmu-let-7i-5p	11449.53
mmu-miR-467b-5p	11399.67	mmu-miR-467b-5p	15781.275	mmu-miR-140-3p	10723.88
mmu-miR-467b-5p	11399.67	mmu-miR-467b-5p	15781.275	mmu-miR-381-3p	10561.735
mmu-miR-467b-5p	11399.67	mmu-miR-467b-5p	15781.275	mmu-let-7f-5p	9942.205
mmu-miR-467b-5p	11399.67	mmu-miR-467a-5p	15781.275	mmu-let-7f-5p	9835.195

Table: 44 Top highly expressed miRNAs among ESC(N), NPC and CN state.

NPC vs ESC(N)	log2 fold change	CN vs NPC	log2 fold change
mmu-miR-214-3p	10.45883694	mmu-miR-295-5p	12.83085704
mmu-miR-214-5p	10.14334117	mmu-miR-290a-3p	11.10946883
mmu-miR-758-5p	8.982324414	mmu-miR-292a-3p	10.40606801
mmu-miR-153-5p	8.727139659	mmu-miR-294-3p	9.016742627
mmu-miR-351-3p	8.316050705	mmu-miR-291b-5p	8.37606363
mmu-miR-100-5p	8.091868618	mmu-miR-292a-5p	8.192871235
mmu-miR-181a-1-3p	7.79425726	mmu-miR-471-3p	8.080277315
mmu-miR-3095-5p	7.744242711	mmu-miR-291a-5p	8.063554171
mmu-miR-133a-5p	7.361166538	mmu-miR-463-3p	7.868009398
mmu-miR-133a-5p	7.361166538	mmu-miR-293-5p	7.732188831
mmu-miR-344f-3p	7.320352638	mmu-miR-293-3p	7.676905197
mmu-miR-877-3p	7.141333673	mmu-miR-878-3p	7.628113873
mmu-miR-675-3p	7.122945501	mmu-miR-880-3p	7.590687031
mmu-miR-675-5p	7.1052269	mmu-miR-294-5p	7.45722623
mmu-miR-463-3p	7.003295374	mmu-miR-200c-5p	7.353129554
mmu-miR-216b-5p	6.990362788	mmu-miR-465a-3p	7.289731076
mmu-miR-8114	6.990234623	mmu-miR-465b-3p	7.289731076
mmu-miR-199a-5p	6.902216307	mmu-miR-465c-3p	7.289731076
mmu-miR-199a-5p	6.902213442	mmu-miR-465a-3p	7.289731076
mmu-miR-6538	6.834099503	mmu-miR-465b-3p	7.289731076

Table: 45 Highly upregulated miRNAs during NPC vs ESC(N), CN vs NPC.

NPC vs ESC(N)	log2 fold change	CN vs NPC	log2 fold change
mmu-miR-181d-5p	-4.509167667	mmu-miR-153-3p	-6.432922443
mmu-miR-466f	-4.525565758	mmu-miR-3970	-6.52777587
mmu-miR-466f	-4.525565758	mmu-miR-9b-5p	-6.686767352
mmu-miR-466f	-4.525565758	mmu-miR-34b-5p	-6.779449219
mmu-miR-34b-5p	-4.537142738	mmu-miR-129-1-3p	-6.839068381
mmu-miR-466f-5p	-4.568597989	mmu-miR-488-3p	-6.916804388
mmu-miR-466f-5p	-4.568718022	mmu-miR-9-3p	-7.160329134
mmu-miR-466f-5p	-4.568718022	mmu-miR-9-3p	-7.163458178
mmu-miR-466f-5p	-4.568718022	mmu-miR-9-3p	-7.163458178
mmu-miR-466f	-4.569553079	mmu-miR-216a-5p	-7.199013171
mmu-miR-467e-3p	-4.605920432	mmu-miR-3093-3p	-7.236144604
mmu-miR-467b-3p	-4.660595436	mmu-miR-217-5p	-7.254206229
mmu-miR-704	-4.688820029	mmu-miR-490-5p	-7.263970495
mmu-miR-1931	-4.775490953	mmu-miR-217-3p	-7.348959104
mmu-miR-12182-3p	-4.822521114	mmu-miR-3081-3p	-7.561368255
mmu-miR-3473b	-5.071498795	mmu-miR-9b-5p	-7.585425796
mmu-miR-467f	-5.097313657	mmu-miR-9b-5p	-7.585425796
mmu-miR-3473f	-5.192389654	mmu-miR-704	-7.827407103
mmu-miR-451a	-5.596310725	mmu-miR-544-5p	-8.466910954
mmu-miR-669b-3p	-5.655070226	mmu-miR-873a-3p	-9.572707686

Table: 46 Highly downregulated miRNAs during NPC vs ESC(N), CN vs NPC.

Name	Hits	Pval
Pathways in cancer	22	0.00256
Axon guidance	11	0.00407
Colorectal cancer	8	0.00407
Cholinergic synapse	10	0.00502
Transcriptional misregulation in cancer	12	0.0102
Dopaminergic synapse	10	0.0147
Ras signaling pathway	13	0.0163
Circadian entrainment	8	0.0163
Thyroid hormone signaling pathway	9	0.0163
Morphine addiction	8	0.0163
Hepatitis B	10	0.0163
Proteoglycans in cancer	12	0.0163
Retrograde endocannabinoid signaling	8	0.0185
Chagas disease (American trypanosomiasis)	8	0.0185
Aldosterone-regulated sodium reabsorption	5	0.0197
mTOR signaling pathway	6	0.0212
HIF-1 signaling pathway	8	0.0221
Insulin resistance	8	0.0235
Glutamatergic synapse	8	0.0241
GABAergic synapse	7	0.0241
Carbohydrate digestion and absorption	5	0.0241
Pancreatic cancer	6	0.0241
Amphetamine addiction	6	0.026
Sphingolipid signaling pathway	8	0.0318
Neurotrophin signaling pathway	8	0.0318
Gastric acid secretion	6	0.0318
Chronic myeloid leukemia	6	0.0318
Rap1 signaling pathway	11	0.0327
Basal cell carcinoma	5	0.0426
Non-small cell lung cancer	5	0.0445
Apoptosis	6	0.0464
cAMP signaling pathway	10	0.0494
cGMP-PKG signaling pathway	9	0.0507
VEGF signaling pathway	5	0.0507
Signaling pathways regulating pluripotency of stem cells	8	0.0507
Long-term depression	5	0.0507
Insulin secretion	6	0.0507

Table: 47 Upregulated KEGG Pathways during NPC vs ESC(N).

Name	Hits	Pval
Hippo signaling pathway	7	0.0786
Glycosphingolipid biosynthesis - ganglio series	2	0.639
Signaling pathways regulating pluripotency of stem cells	5	0.639
Glycolysis / Gluconeogenesis	1	1
Fructose and mannose metabolism	1	1
Galactose metabolism	2	1
Purine metabolism	2	1
Pyrimidine metabolism	2	1
Alanine, aspartate and glutamate metabolism	1	1
Cysteine and methionine metabolism	1	1
Valine, leucine and isoleucine degradation	1	1
beta-Alanine metabolism	1	1
Taurine and hypotaurine metabolism	1	1
Starch and sucrose metabolism	1	1
N-Glycan biosynthesis	1	1
Other glycan degradation	1	1
Amino sugar and nucleotide sugar metabolism	1	1
Glycosaminoglycan degradation	1	1
Inositol phosphate metabolism	2	1
Glycerophospholipid metabolism	3	1
Ether lipid metabolism	1	1
Arachidonic acid metabolism	1	1
Linoleic acid metabolism	1	1
alpha-Linolenic acid metabolism	1	1
Sphingolipid metabolism	1	1
Glycosphingolipid biosynthesis - lacto and neolacto series	1	1
Glycosphingolipid biosynthesis - globo series	1	1
Butanoate metabolism	1	1
Nicotinate and nicotinamide metabolism	2	1
Pantothenate and CoA biosynthesis	1	1
Aminoacyl-tRNA biosynthesis	1	1
Metabolic pathways	13	1
Carbon metabolism	1	1
2-Oxocarboxylic acid metabolism	1	1
Biosynthesis of amino acids	1	1
Ribosome	2	1
RNA transport	2	1
mRNA surveillance pathway	2	1

Table: 48 Downregulated KEGG Pathways during NPC vs ESC(N).

Name	Hits	Pval
Axon guidance	31	6.32E-12
Glutamatergic synapse	24	5.62E-08
Dopaminergic synapse	26	5.62E-08
GABAergic synapse	20	2.63E-07
cAMP signaling pathway	29	0.00000243
Retrograde endocannabinoid signaling	20	0.00000247
ErbB signaling pathway	17	0.0000197
Circadian entrainment	18	0.0000209
Synaptic vesicle cycle	14	0.0000253
Nicotine addiction	11	0.0000434
Adrenergic signaling in cardiomyocytes	22	0.0000608
Amphetamine addiction	14	0.000062
Long-term potentiation	13	0.000219
Cholinergic synapse	17	0.000407
Pathways in cancer	38	0.000418
Ras signaling pathway	25	0.00127
Cell adhesion molecules (CAMs)	20	0.00127
Cocaine addiction	10	0.00127
Signaling pathways regulating pluripotency of stem cells	18	0.00147
Morphine addiction	14	0.00152
Endocytosis	28	0.00158
Chronic myeloid leukemia	12	0.00182
Insulin secretion	13	0.00214
Colorectal cancer	11	0.00214
Calcium signaling pathway	20	0.00352
Wnt signaling pathway	17	0.00468
Gastric acid secretion	11	0.00608
Estrogen signaling pathway	13	0.0066
Aldosterone synthesis and secretion	12	0.0066
Melanogenesis	13	0.00696
Glioma	10	0.00792
Salivary secretion	11	0.00814
Insulin signaling pathway	16	0.00816
Pancreatic cancer	10	0.00816
Oxytocin signaling pathway	17	0.0101
Proteoglycans in cancer	20	0.0128
Neurotrophin signaling pathway	14	0.0137
MAPK signaling pathway	23	0.0143

Table: 49 Upregulated KEGG Pathways during CN vs NPC.

Name	Hits	Pval
Proteoglycans in cancer	19	0.0000756
HTLV-I infection	22	0.0000794
Apoptosis	9	0.00599
Signaling pathways regulating pluripotency of stem cells	12	0.00599
Transcriptional misregulation in cancer	14	0.00599
p53 signaling pathway	8	0.00622
Influenza A	13	0.00622
Pathways in cancer	22	0.00622
Hippo signaling pathway	12	0.00782
Focal adhesion	14	0.00911
Measles	11	0.00911
Hepatitis B	11	0.0142
Rap1 signaling pathway	13	0.0324
Wnt signaling pathway	10	0.0398
Cell cycle	9	0.0485
Chagas disease (American trypanosomiasis)	8	0.049
Platelet activation	9	0.0588
Hepatitis C	9	0.0712
Melanoma	6	0.0919
NF-kappa B signaling pathway	7	0.11
PI3K-Akt signaling pathway	16	0.11
Melanogenesis	7	0.11
African trypanosomiasis	4	0.11
Basal cell carcinoma	5	0.11
Herpes simplex infection	11	0.111
Acute myeloid leukemia	5	0.114
Regulation of actin cytoskeleton	11	0.134
mTOR signaling pathway	5	0.134
Hypertrophic cardiomyopathy (HCM)	6	0.134
TNF signaling pathway	7	0.137
MAPK signaling pathway	12	0.141
ECM-receptor interaction	6	0.145
Toxoplasmosis	7	0.145
Ras signaling pathway	11	0.158
Thyroid hormone signaling pathway	7	0.159
Amoebiasis	7	0.162
RIG-I-like receptor signaling pathway	5	0.162
Arginine and proline metabolism	4	0.195

Table: 50 Downregulated KEGG Pathways during CN vs NPC.

Name	Hits	Pval
cell migration	59	2.34E-09
neurogenesis	72	1.17E-08
neuron differentiation	64	1.17E-08
generation of neurons	69	1.17E-08
neuron projection development	46	6.51E-08
neuron development	51	6.51E-08
regulation of cell migration	35	1.21E-07
nervous system development	91	1.24E-07
cell development	83	1.82E-07
positive regulation of cell migration	24	0.00000287
regulation of cell adhesion	23	0.0000241
cell morphogenesis involved in differentiation	37	0.0000453
muscle cell differentiation	27	0.0000683
cell-substrate adhesion	18	0.000139
central nervous system development	37	0.000142
axonogenesis	26	0.000184
regulation of small GTPase mediated signal transduction	23	0.000323
brain development	29	0.000359
anatomical structure formation involved in morphogenesis	78	0.000381
behavior	34	0.000478
tube development	29	0.000597
cell-cell signaling	43	0.000658
regulation of heart contraction	13	0.001
response to hypoxia	13	0.00103
striated muscle tissue development	25	0.00104
axon guidance	15	0.00131
intracellular signal transduction	74	0.00138
extracellular structure organization	14	0.00138
transmission of nerve impulse	34	0.00173
positive regulation of cell adhesion	12	0.00173
regulation of signal transduction	76	0.00181
anatomical structure morphogenesis	92	0.00184
organ morphogenesis	43	0.00208
positive regulation of cell differentiation	31	0.00238
regulation of Ras protein signal transduction	19	0.00239
regulation of transport	51	0.00322
positive regulation of developmental process	39	0.00375

Table: 51 Upregulated Gene Ontology: Biological Processes during NPC vs ESC(N).

Name	Hits	Pval
Name	Hits	Pvalue
negative regulation of transcription from RNA polymerase II promoter	19	0.00696
negative regulation of transcription, DNA-dependent	25	0.00696
negative regulation of nucleobase-containing compound metabolic process	25	0.00696
negative regulation of RNA metabolic process	26	0.00713
negative regulation of cellular biosynthetic process	25	0.00713
negative regulation of biosynthetic process	26	0.0158
negative regulation of cellular metabolic process	26	0.0171
regulation of homeostatic process	29	0.0488
homeostasis of number of cells	10	0.0918
negative regulation of metabolic process	8	0.103
cell cycle arrest	30	0.126
skeletal muscle tissue development	8	0.126
negative regulation of cell cycle	8	0.134
DNA catabolic process	9	0.173
immune system development	3	0.202
regulation of cell adhesion	15	0.213
negative regulation of sequence-specific DNA binding transcription factor activity	8	0.213
regulation of DNA binding	5	0.213
embryo development	4	0.213
tissue development	20	0.213
myoblast differentiation	26	0.246
protein folding	3	0.246
hematopoietic or lymphoid organ development	5	0.262
myeloid cell differentiation	14	0.262
hemopoiesis	8	0.265
regulation of binding	13	0.272
carbohydrate transport	6	0.272
muscle cell differentiation	5	0.32
negative regulation of cellular process	9	0.32
response to steroid hormone stimulus	48	0.32
response to hormone stimulus	6	0.32
actin filament organization	10	0.372
female pregnancy	6	0.375
sexual reproduction	4	0.375
negative regulation of cell differentiation	11	0.375
embryonic morphogenesis	11	0.375
striated muscle tissue development	12	0.375
nucleus organization	9	0.384

Table: 52 Downregulated Gene Ontology: Biological Processes during NPC vs ESC(N).

Name	Hits	Pval
protein kinase binding	24	0.013
protein heterodimerization activity	23	0.013
kinase binding	25	0.0135
enzyme binding	49	0.0144
chromatin binding	19	0.0348
sequence-specific DNA binding	32	0.0348
transcription from RNA polymerase II promoter	65	0.0438
positive regulation of transcription, DNA-dependent	54	0.0438
enzyme activator activity	20	0.0631
glycosaminoglycan binding	10	0.102
identical protein binding	35	0.102
insulin-like growth factor receptor binding	3	0.102
GTPase activator activity	14	0.131
zinc ion binding	45	0.131
transcription factor binding	21	0.138
negative regulation of transcription, DNA-dependent	39	0.141
calmodulin binding	9	0.152
protein homodimerization activity	24	0.152
protein dimerization activity	36	0.158
protein tyrosine phosphatase activity	7	0.19
cytoskeletal protein binding	27	0.19
SMAD binding	6	0.19
actin binding	16	0.201
chloride channel activity	6	0.201
PDZ domain binding	7	0.201
transition metal ion binding	49	0.201
double-stranded DNA binding	9	0.201
histone deacetylase binding	5	0.27
anion channel activity	6	0.286
transcription cofactor activity	16	0.322
transcription corepressor activity	8	0.326
protein binding transcription factor activity	17	0.398
growth factor binding	7	0.417
protein domain specific binding	22	0.417
GTPase regulator activity	19	0.417
heparin binding	6	0.518
ionotropic glutamate receptor activity	3	0.592
RNA polymerase II transcription cofactor activity	4	0.626
transcription coactivator activity	9	0.626

Table: 53 Upregulated Gene Ontology: Molecular Function during NPC vs ESC(N).

Name	Hits	Pval
negative regulation of transcription, DNA-dependent	25	0.0113
transcription factor binding	13	0.132
phosphatase regulator activity	4	0.312
phosphatase inhibitor activity	3	0.373
DNA binding	34	0.454
sequence-specific DNA binding	14	0.556
protein heterodimerization activity	9	0.638
unfolded protein binding	3	0.638
chromatin binding	8	0.652
structural constituent of muscle	2	0.652
histone deacetylase binding	3	0.735
thyroid hormone receptor binding	2	0.735
NF-kappaB binding	2	0.735
endodeoxyribonuclease activity	2	0.784
transcription from RNA polymerase II promoter	26	0.784
carbohydrate transmembrane transporter activity	2	0.917
nucleotide binding	17	1
magnesium ion binding	1	1
protein binding transcription factor activity	8	1
RNA polymerase II transcription cofactor activity	2	1
peptide receptor activity	2	1
DNA helicase activity	2	1
damaged DNA binding	1	1
steroid hormone receptor activity	2	1
transcription cofactor activity	5	1
transcription coactivator activity	2	1
transcription corepressor activity	3	1
RNA binding	8	1
mRNA binding	2	1
structural constituent of ribosome	2	1
translation initiation factor activity	1	1
motor activity	1	1
actin binding	6	1
GTPase activity	2	1
endopeptidase activity	1	1
helicase activity	2	1
nuclease activity	2	1
endonuclease activity	2	1
deoxyribonuclease activity	2	1
phospholipase activity	3	1

Table: 54 Downregulated Gene Ontology: Molecular Function during NPC vs ESC(N).

Name	Hits	Pval
neuron projection	56	2.15E-09
axon	32	7.61E-09
cell body	34	1.71E-07
cell projection part	43	1.71E-07
dendrite	30	0.00000552
synapse part	27	0.000482
cell projection	70	0.000686
synapse	35	0.000994
cell leading edge	15	0.0103
growth cone	9	0.0127
site of polarized growth	9	0.0141
intercalated disc	6	0.0181
membrane raft	14	0.0259
perinuclear region of cytoplasm	18	0.0833
basement membrane	8	0.0842
nuclear membrane	6	0.123
nuclear chromatin	9	0.139
anchored to membrane	9	0.139
plasma membrane part	55	0.151
nuclear matrix	5	0.152
cell junction	31	0.152
intrinsic to plasma membrane	27	0.169
transcription factor complex	15	0.198
extracellular matrix part	11	0.205
sarcomere	8	0.248
microtubule associated complex	7	0.266
integral to plasma membrane	24	0.282
receptor complex	9	0.282
lamellipodium	6	0.326
extracellular matrix	18	0.372
contractile fiber part	8	0.372
ruffle	5	0.42
cytoskeletal part	39	0.426
myofibril	8	0.464
microtubule	12	0.475
trans-Golgi network transport vesicle	2	0.475
histone deacetylase complex	3	0.576
microtubule cytoskeleton	24	0.576
basolateral plasma membrane	7	0.576
contractile fiber	8	0.576

Table: 55 Upregulated Gene Ontology: Cellular Component during NPC vs ESC(N).

Name	Hits	Pval
nuclear part	30	0.315
perinuclear region of cytoplasm	10	0.315
nucleus	68	0.322
nucleoplasm	16	0.322
transcription factor complex	8	0.322
cytosol	18	0.322
synaptic vesicle	4	0.322
vesicle membrane	5	0.322
coated vesicle membrane	3	0.322
nuclear lumen	22	0.322
organelle lumen	23	0.322
protein complex	45	0.322
chromosomal part	11	0.322
nucleoplasm part	15	0.322
synapse part	9	0.322
membrane-enclosed lumen	25	0.322
clathrin-coated vesicle	5	0.322
nuclear chromosome part	6	0.403
protein serine/threonine phosphatase complex	2	0.403
chromosome	12	0.437
cell junction	13	0.437
coated vesicle	5	0.437
cytoplasmic vesicle membrane	4	0.437
macromolecular complex	50	0.437
cytoplasmic vesicle part	4	0.437
membrane-bounded vesicle	11	0.437
nuclear chromosome	6	0.439
vesicle	15	0.449
organelle part	56	0.449
synapse	11	0.449
cell cortex	4	0.55
actin cytoskeleton	6	0.55
nuclear body	4	0.55
cytosolic part	4	0.557
intracellular organelle part	53	0.599
DNA-directed RNA polymerase II, core complex	1	0.599
neuron projection	12	0.619
endomembrane system	11	0.687
ubiquitin ligase complex	3	0.702
chromatin	5	0.721

Table: 56 Downregulated Gene Ontology: Cellular Component during NPC vs ESC(N).

Name	Hits	Pval
transmission of nerve impulse	121	6.99E-24
neuron differentiation	161	2.32E-23
neuron development	130	6.80E-22
synaptic transmission	102	1.63E-21
generation of neurons	169	1.67E-21
neuron projection development	113	4.07E-20
neurogenesis	173	4.15E-20
nervous system development	221	3.15E-17
cell-cell signaling	127	4.00E-17
axonogenesis	75	5.62E-17
regulation of neurogenesis	87	1.81E-14
brain development	77	9.01E-13
cell morphogenesis involved in differentiation	91	1.12E-12
central nervous system development	91	3.18E-11
regulation of neurotransmitter levels	34	1.98E-10
cell development	179	3.99E-10
synapse organization	33	4.36E-09
homophilic cell adhesion	22	4.71E-09
regulation of membrane potential	54	7.07E-09
behavior	81	8.37E-09
neurotransmitter secretion	25	8.99E-09
glutamate receptor signaling pathway	23	3.64E-08
regulation of axonogenesis	28	3.91E-08
locomotory behavior	34	5.99E-08
axon guidance	35	6.15E-08
generation of a signal involved in cell-cell signaling	52	1.35E-07
synapse assembly	20	1.77E-07
learning or memory	34	1.77E-07
regulation of synapse structure and activity	19	2.34E-07
regulation of cellular component organization	138	3.18E-07
positive regulation of cellular component organization	72	7.73E-07
regulation of transport	124	9.94E-07
regulation of cell morphogenesis	46	0.00000355
cell-cell adhesion	47	0.00000623
respiratory gaseous exchange	13	0.0000225
regulation of cell differentiation	125	0.0000591
developmental maturation	29	0.000108
regulation of action potential	28	0.000115
cellular membrane organization	46	0.000136
cellular homeostasis	87	0.000196

Table: 57 Upregulated Gene Ontology: Biological Processes during CN vs NPC.

Name	Hits	Pval
tissue morphogenesis	37	0.000152
regulation of cell proliferation	62	0.000204
morphogenesis of an epithelium	29	0.00057
positive regulation of developmental process	44	0.000659
vasculature development	36	0.000772
embryo development	52	0.000772
cell proliferation	68	0.000772
tube development	29	0.00138
regulation of cell differentiation	58	0.00138
positive regulation of epithelial cell proliferation	13	0.00138
regulation of developmental process	75	0.002
regulation of chromosome organization	12	0.00207
tube morphogenesis	22	0.00268
positive regulation of cell proliferation	37	0.00335
negative regulation of cell cycle	20	0.00452
positive regulation of cell differentiation	31	0.00505
tissue development	64	0.00525
regulation of cellular component organization	56	0.00562
regulation of organelle organization	27	0.0079
angiogenesis	23	0.00845
regulation of transcription from RNA polymerase II promoter	68	0.00852
negative regulation of cell proliferation	27	0.00852
muscle cell differentiation	22	0.00852
regulation of cell cycle	31	0.0114
apoptotic signaling pathway	15	0.0118
regulation of gene expression, epigenetic	13	0.0128
wound healing	16	0.0135
transcription from RNA polymerase II promoter	67	0.0151
positive regulation of metabolic process	86	0.0178
positive regulation of cellular metabolic process	81	0.0199
skeletal muscle tissue development	15	0.0225
negative regulation of response to stimulus	37	0.025
response to endogenous stimulus	40	0.0251
negative regulation of signal transduction	31	0.0251
regulation of protein modification process	40	0.0251
negative regulation of cell differentiation	26	0.0251
homeostasis of number of cells	14	0.0251
positive regulation of RNA metabolic process	55	0.0255
regulation of anatomical structure morphogenesis	31	0.0259
interphase of mitotic cell cycle	13	0.0259

Table: 58 Downregulated Gene Ontology: Biological Processes during CN vs NPC.

Name	Hits	Pval
PDZ domain binding	21	0.00106
calcium ion binding	57	0.00405
protein domain specific binding	65	0.00405
glutamate receptor activity	11	0.0179
zinc ion binding	110	0.0284
sequence-specific DNA binding	66	0.0539
protein heterodimerization activity	41	0.0539
voltage-gated cation channel activity	24	0.0539
ionotropic glutamate receptor activity	8	0.077
cytoskeletal protein binding	64	0.0838
phosphatase inhibitor activity	7	0.143
phospholipid binding	41	0.242
calmodulin binding	17	0.248
GABA receptor activity	5	0.415
protein complex binding	38	0.415
tubulin binding	19	0.439
transition metal ion binding	114	0.439
transmembrane receptor protein kinase activity	14	0.469
voltage-gated potassium channel activity	13	0.488
amine transmembrane transporter activity	3	0.488
channel regulator activity	9	0.488
phosphatase regulator activity	8	0.488
protein kinase binding	36	0.488
phosphatase binding	13	0.488
inorganic cation transmembrane transporter activity	50	0.488
metal ion transmembrane transporter activity	45	0.488
excitatory extracellular ligand-gated ion channel activity	50	0.488
protein N-terminus binding	9	0.488
microtubule binding	11	0.488
monovalent inorganic cation transmembrane transporter activity	13	0.517
neurotransmitter binding	35	0.569
lipid kinase activity	4	0.569
microtubule motor activity	2	0.607
Rho GTPase activator activity	7	0.607
voltage-gated sodium channel activity	5	0.607
protein phosphatase binding	4	0.607
sterol binding	9	0.607
protein dimerization activity	5	0.607
protein C-terminus binding	75	0.607

Table: 59 Upregulated Gene Ontology: Molecular Function during CN vs NPC.

Name	Hits	Pval
chromatin binding	20	0.163
integrin binding	7	0.163
ATP binding	51	0.163
transcription from RNA polymerase II promoter	67	0.163
DNA-dependent ATPase activity	6	0.163
adenyl nucleotide binding	51	0.163
adenyl ribonucleotide binding	51	0.163
positive regulation of transcription, DNA-dependent	54	0.163
insulin-like growth factor receptor binding	3	0.163
protein complex binding	20	0.209
enzyme activator activity	19	0.232
growth factor activity	9	0.256
zinc ion binding	46	0.256
transition metal ion binding	53	0.256
DNA helicase activity	4	0.283
transcription corepressor activity	9	0.283
protein tyrosine phosphatase activity	7	0.283
ribonucleoprotein complex binding	5	0.283
ATP-dependent DNA helicase activity	3	0.283
transcription factor binding	20	0.37
ATP-dependent helicase activity	6	0.37
negative regulation of transcription, DNA-dependent	38	0.431
aminopeptidase activity	3	0.632
ATPase activity	14	0.632
protein domain specific binding	23	0.632
protein homodimerization activity	22	0.632
double-stranded RNA binding	4	0.632
microtubule motor activity	4	0.652
carbonate dehydratase activity	2	0.652
protein serine/threonine kinase activity	21	0.652
phospholipid binding	17	0.652
transition metal ion transmembrane transporter activity	3	0.652
hydro-lyase activity	4	0.652
heparin binding	6	0.68
purine ribonucleotide binding	53	0.68
phosphate transmembrane transporter activity	2	0.68
ATPase activity, coupled	11	0.68
transcription cofactor activity	15	0.686
actin binding	14	0.686
calmodulin binding	7	0.686

Table: 60 Downregulated Gene Ontology: Molecular Function during CN vs NPC.

Name	Hits	Pval
neuron projection	165	4.56E-35
synapse	136	2.02E-29
synapse part	104	1.19E-28
dendrite	84	1.50E-18
axon	74	3.07E-18
cell projection part	109	8.22E-18
cell body	78	2.85E-14
cell projection	194	7.96E-12
growth cone	30	1.27E-11
site of polarized growth	30	2.50E-11
synaptic vesicle	32	5.14E-11
cell junction	101	7.36E-08
clathrin-coated vesicle	34	1.25E-07
integral to plasma membrane	82	6.01E-07
intrinsic to plasma membrane	84	0.00000373
coated vesicle	34	0.00000628
voltage-gated potassium channel complex	13	0.000237
membrane raft	32	0.000708
coated vesicle membrane	13	0.0033
anchored to membrane	22	0.00402
plasma membrane part	144	0.00692
perinuclear region of cytoplasm	41	0.0137
kinesin complex	6	0.0137
cytoplasmic membrane-bounded vesicle	58	0.0182
membrane-bounded vesicle	61	0.0755
trans-Golgi network	15	0.0784
cytoplasmic vesicle part	18	0.0784
cytoplasmic vesicle	80	0.0829
cytoplasmic vesicle membrane	17	0.0864
early endosome	15	0.134
endosome	53	0.137
nuclear membrane	10	0.16
integral to Golgi membrane	6	0.169
intrinsic to Golgi membrane	6	0.169
cell leading edge	23	0.169
intercalated disc	7	0.222
integral to organelle membrane	18	0.222
vesicle membrane	17	0.23
Golgi apparatus	75	0.298
vesicle	82	0.326

Table: 61 Upregulated Gene Ontology: Cellular Component during CN vs NPC.

Name	Hits	Pval
nuclear lumen	58	0.00557
cell-substrate junction	10	0.00598
nucleoplasm	40	0.0158
focal adhesion	8	0.0158
organelle lumen	60	0.0158
membrane-enclosed lumen	61	0.0158
cell-substrate adherens junction	8	0.0179
nucleoplasm part	35	0.0236
transcription factor complex	18	0.0456
cell surface	27	0.0456
basement membrane	8	0.117
spindle	9	0.144
adherens junction	9	0.144
nuclear body	10	0.144
integral to Golgi membrane	4	0.144
intrinsic to Golgi membrane	4	0.144
nuclear part	61	0.144
spindle pole	5	0.158
nucleolus	12	0.195
RNA polymerase complex	6	0.195
DNA-directed RNA polymerase complex	6	0.195
nuclear DNA-directed RNA polymerase complex	6	0.195
extracellular matrix part	11	0.209
cell leading edge	11	0.209
nuclear chromosome	13	0.215
nuclear chromosome part	12	0.215
cell cortex	9	0.216
integrin complex	3	0.216
intercalated disc	4	0.216
PML body	4	0.264
cell-cell junction	14	0.266
DNA-directed RNA polymerase II, holoenzyme	5	0.277
proteinaceous extracellular matrix	16	0.282
endomembrane system	27	0.282
vesicle membrane	8	0.282
basolateral plasma membrane	8	0.282
apical part of cell	13	0.282
late endosome	6	0.293
Golgi apparatus	31	0.293
extracellular matrix	18	0.293

Table: 62 Downregulated Gene Ontology: Cellular Component during CN vs NPC.

Molecules in Network	Score	Focus Molecules	Top Diseases and Functions
ADAMTSL1,ATN1,ATRN,ATXN7,AUTS2,B3GLCT,CRIM1,DPY19L3,FBN2,FOXF2,HIBADH,HSPBAP1,KBTBD6,KBTBD8,KLHL29,Marcks,MATN2,MEGF6,MEGF8,MFAP2,NCALD,NELL2,NPHP3,PDCD2,RBFA,RNF157,ROBO3,RSPO1,RSPO3,SRC (family),ZNF579,ZNF608,ZNF692,ZSWIM8,ZZEF1	30	34	[Embryonic Development, Nervous System Development and Function, Organ Development]
Aff2,BAHD1,CEP170,CUEDC1,CYTH3,DLGAP3,DPPA2,Dux,Eif4a3l1,GABRB2,Gm2016 (includes others),Gm4340 (includes others),GRASP,Gsk3,HEYL,KLHL13,KLHL14,LAPTM5,MAPKBP1,MIS18BP1,NEUROG2,NHS,NHSL2,PDZRN3,Pramel7,SNX29,SP8,SYCE1,Tdpoz4/Tdpoz8,TOM1L2,TOX3,UBE2T,ZBTB5,Zfp352,Zscan4c (includes others)	30	34	[Cell-To-Cell Signaling and Interaction, Dental Disease, Nervous System Development and Function]
ADAM11,ADAM19,ADAM33,ADAMTS1,ADAMTS3,ADAMTS5,ADAMTS7,ARHGAP21,ARHGAP32,CBLN4,COCH,CSGALNACT1,CTNNA2,DCHS1,FBXO2,FURIN,GRINA,IQSEC1,IQSEC3,KALRN,LGI4,MPP3,NECTIN1,NMDA Receptor,OAF,PAM,PAPSS1,PAPSS2,TMEM37,TMX4,TNIK,VASN,VIT,ZC3H8,ZCCHC12	30	34	[Cancer, Cell Morphology, Cellular Development]
ACAD11,AJUBA,AMOT,ANKRD52,C16orf54,CRYBG2,DCLK1,DND1,DOCK7,DYNLL2,FAM199X,IKK (complex),JCAD,LATS1,LATS2,LIMD1,MOB1B,MOB3B,OCIAD2,PELI3,PKP4,PLEKHG1,PPP6R2,PTPN14,RASAL2,RASSF2,RASSF4,RRAGD,SIPA1L1,SYDE1,TAOK1,TBKBP1,TESK2,VPS37C,ZNF618	30	34	[Cancer, Gastrointestinal Disease, Organismal Injury and Abnormalities]
APCDD1,CD248,CNR1,CTBP2,DACH1,DYRK2,EFEMP1,EFEMP2,FBLN5,FCGBP,FMN1,FOXB1,FOXC1,FOXG1,FOXQ1,GREB1,HCN4,HOXD13,MAL,Meis1,NOL4L,PDYN,Pias,PITX2,PLSCR3,RBMXL2,RERE,SATB1,SATB2,SHISA5,TCF15,TFAP2C,TXNDC5,ZNF316,ZNF780A	30	34	[Connective Tissue Disorders, Dermatological Diseases and Conditions, Developmental Disorder]
BCL11A,CDYL2,COL23A1,CRYL1,DPP8,EBF1,FOXK2,IFT122,KCTD12,KIAA1109,LDB2,LHX2,LOC728392,LRRK1,MSL3,NOVA1,NPY,PFKP,PHF20L1,PHKA1,Plekha6,PODN,Rag,RASIP1,SIX3,Six3os1,Snhg11,SRRM4,STAC2,TSHZ3,TTC21B,VGF,ZNF428,ZNF521,ZNF616	30	34	[Connective Tissue Disorders, Organismal Injury and Abnormalities, Skeletal and Muscular System Development and Function]
BCOR,CCDC6,CDX2,CEP55,GATA4,Gm428 (includes others),GPRC5A,Histone h4,IGFBPL1,IMPDH2,KMT2A,MED1,MSRB3,MTMR4,MTMR7,MYCN,NCOA2,NR5A2,PDE7B,PRDM1,PRRG1,RBBP4,Rhox6/Rhox9,RUVBL2,SET,Slc22a21,SMIM15,SOX2,SUMO1,TCF4,TCF7L2,TMEM121B,TMEM141,TMEM163,ZNF827	30	34	[Cellular Development, Digestive System Development and Function, Gene Expression]
ADAD2,AIFM3,ALKBH7,ANKRD45,APP,Aspartyl Protease,BACE2,Beta Secretase,C5orf15,C7orf31,CAPSL,CCDC149,CPED1,EBF4,EPB41L4A,Fnbp1l,IGSF10,INKA1,KIAA0513,LHFPL2,MCEE,MND1,MUCL3,NAPSA,R3HDML,RAPGEFL1,SERAC1,TCP11,TERB2,TMCC1,TMEM35A,TRIML1,TSPAN12,Zfp811,ZNF641	28	33	[Auditory Disease, Developmental Disorder, Hereditary Disorder]
ANK3,ANKRD16,C16orf70,CHL1,CNTN1,CRAMP1,DOK4,EID2,EXOC3L2,GPRIN1,KIDINS220,KIF21B,LIX1,MAGED2,MAGEE1,MAPRE3,MICALL2,OSBPL5,OXNAD1,PTPRB,PTPRZ1,RBP7,SCN1A,SCN3B,SCN5A,SLC6A8,SNTA1,SNTB2,sodium channel,SPZ1,ST5,STARD8,TTBK2,VEGFA,voltage-gated sodium channel	28	33	[Behavior, Molecular Transport, Neurological Disease]
ANO6,DIP2A,DLK1,EPHX1,FAM189A2,FAM214B,FBLIM1,FERMT1,IgG2c,KCNQ1OT1,MBLAC2,MKX,MYORG,PCDHGA5,PRICKLE1,RAB34,S1PR1,Sox,SOX1,SOX14,SOX15,SOX18,SOX21,SOX3,STEAP3,TMCC2,TMTC4,TRABD2B,TTYH3,WLS,WNT3,WNT4,WNT6,WNT8B,ZFP57	28	33	[Cellular Assembly and Organization, Cellular Compromise, Gene Expression]
ACTG2,AFAP1,ARHGEF17,CDC42EP5,DBN1,EFHD2,HBE1,HMCN1,LAP3,LIMD2,LRR1,MYH10,MYL6B,MYO18A,MYO6,NEB,OGN,PHLDB2,PPP1CB,PPP1R18,PRSS35,RAPGEF2,RAPGEF6,RERG,Septin,SEPTIN1,SEPTIN3,SEPTIN8,SMO,SORBS2,SVIL,SYNPO,TNFAIP1,TROPONIN,WFDC2	28	33	[Cellular Assembly and Organization, Cellular Function and Maintenance, Cellular Movement]
AMT,BAG2,CACYBP,CAPG,CARNMT1,CDKL3,CHCHD10,CHIP/Hsc70/Hsp70/Hsp90/Bag2,CNTLN,CORO2B,CREBRF,DDX59,EEF1B2,Eif4g,FAF1,FAT3,FIZ1,GET3,HERC6,HIKESHI,HSPA4L,HSPA8,LRRC2,MKKS,NADSYN1,PACS2,PHLPP2,PNMA5,PTGES3,RAD23A,RPL10L,SH3RF2,SUGT1,UBQLN2,ZBTB8A	28	33	[Connective Tissue Disorders, Developmental Disorder, Gastrointestinal Disease]

Table: 63 Top Network of molecules associated with diseases and function in NPC vs ESC(N) state.

Molecules in Network	Score	Focus Molecules	Top Diseases and Functions
AKT1,ANKRD26,CCDC14,CCDC18,CCDC66,CCDC77,CEP135,CEP162,CEP290,CEP295,CEP95,CGNL1,CSPP1,Ehbp1l1,FGFR1OP,FOPNL,FXYD5,H2BC10,HAUS3,KIAA0753,LRRC49,LRRCC1,MPHOSPH9,NCBP2AS2,NOL4,NOL4L,Ott (includes others),PIBF1,PLEKHG1,RAP1GAP2,SKA1,SKA3,SNED1,TBC1D31,TPGS1	26	35	[Developmental Disorder, Hereditary Disorder, Neurological Disease]
AGAP3,ARPIN/ARPIN-AP3S2,C9orf64,COQ10A,ELAVL3,FAM171A1,FMNL3,FOCAD,FSD1,GDAP1,MAF1,MEX3A,NAA11,NAA15,NAA16,NAA25,NAA50,NHLRC2,NTRK1,PCDHB3,PLEKHH3,POLR3A,POLR3B,POLR3G,POLR3GL,PRRT4,RTL1,SEZ6L2,SLC25A14,SUSD5,TMEM59,TRIM46,URGCP,ZNF316,ZNF629	26	35	[Hereditary Disorder, Neurological Disease, Organismal Injury and Abnormalities]
ABI3,GAMT,GM2A,INPP4A,ISL2,LDB1,LDB2,LHX1,LHX4,LHX8,LHX9,LMO1,Lmo3,LMO4,LMX1A,LMX1B,MAB21L1,MAB21L2,MEIS2,MEIS3,PAK1,PBX1,PBX2,PBX3,PBX4,PGBD5,PTAR1,PYGO1,RAB9B,RABGGTB,RLIM,SSBP2,SSBP4,TAL2,YKT6	26	35	[Cardiovascular Disease, Congenital Heart Anomaly, Developmental Disorder]
AARSD1,ANKRD27,ARRB2,C11orf74,C18orf25,CREB3L2,FBXL17,HCFC2,IFT140,KBTBD8,KIF1A,KLHL21,KLHL29,KLHL9,MTO1,MYO5C,MYO9B,NDUFA3,NPHP3,PEAK1,PJA2,PNMA8A,PSMG4,RFTN2,SFXN2,TCEAL1,TCEANC2,TTC21B,VPS26B,WASHC2A/WASHC2C,WASHC4,WDR35,ZBTB46,ZCCHC12,ZNF692	26	35	[Connective Tissue Disorders, Developmental Disorder, Hereditary Disorder]
AFDN,ALDH1L2,CBY1,CDC25B,CGN,DCLK1,DENND4C,DGCR2,FAM110A,FAM110B,FAS,FGD6,GAB2,INTS2,KIF13B,KIF15,KIF18B,KIF1C,KIF3C,LARP1,LIMA1,LRFN1,MDGA2,NUAK2,OSBPL6,PHLDB2,PLEKHA7,PPM1H,PTPN14,RAB11FIP2,STARD13,STARD9,SYDE1,TBC1D1,ZNF638	26	35	[Cancer, Gastrointestinal Disease, Organismal Injury and Abnormalities]
B9D2,CCDC68,CEP192,CHST10,CHST12,CLBA1,CNPPD1,COMMD2,COMMD9,EDRF1,EFR3A,ESS2,FAT3,FBXO31,GNB2,GNGT2,GXYLT1,HS6ST2,LOXL2,MIER1,MNS1,NDC80,NPTX1,OMA1,PCDHB15,PON2,QSER1,RAB28,SIL1,SMOC1,TINAGL1,TMEM25,TMEM266,TTC14,ZNF358	26	35	[Endocrine System Development and Function, Lipid Metabolism, Organismal Injury and Abnormalities]
ANKRD29,BRICD5,DLGAP4,ESCO2,FAM122B,FAM171A2,GGT7,IMPACT,KLF12,LIN54,LRRTM2,LSR,MCOLN3,MRPL14,NAALAD2,NCAPD3,NDRG3,NRN1,RNF130,SARAF,SFXN4,SHANK1,SHANK2,SSTR2,ST7,STXBP1,STXBP2,SYNE4,THEM6,TMEM140,TMEM219,TMEM65,TRIM9,TSPAN18,TUSC3	26	35	[Auditory Disease, Cell-To-Cell Signaling and Interaction, Endocrine System Development and Function]
BMS1,DDX10,DOK5,FGF8,GPATCH4,GRM1,KLHDC2,Ktn1,LAP3,MACROH2A2,MPHOSPH10,MYBBP1A,NLE1,NOL8,NOP2,NOP56,PDCD11,PLA2R1,PLEKHO1,PUM3,RAI2,RBM19,RBM28,REXO4,RIOX2,RPL30,RPL37A,RTN4RL1,SDAD1,SGTB,SPTBN2,SQOR,THAP7,URB1,URB2	26	35	[Cellular Function and Maintenance, Hereditary Disorder, RNA Post-Transcriptional Modification]
ALOX5AP,CARHSP1,CCDC85A,CLCN4,DACH1,DMRT3,DPPA5,DPYSL5,ESRRB,FBXO15,FOXN2,FXYD6,GNG2,HACD4,IER5L,KHDRBS3,MGAT5B,mir-290,MPZL2,NANOG,NR0B1,NR6A1,PCDHA2,PDLIM2,PIK3C2B,PREX1,SALL1,SALL4,STC2,Tcf7,TCF7L1,Tdh,VRTN,Zfp345 (includes others),ZFP42	26	35	[Cell Death and Survival, Cellular Development, Cellular Function and Maintenance]
ACKR3,AQR,C11orf68,CARMIL1,CAV2,CEP126,CP,DENND2D,ELAVL2,GCNT2,GNA15,GOLGB1,HSBP1,IGF2BP1,KMT2E,LARP4,LIN28B,ODF2L,PABPC1,PDLIM1,PIPOX,PNN,PRKRA,PRSS35,RALYL,RBMX,SAFB2,SH3D21,SREK1,SRSF11,STAU2,STEAP2,THSD7A,TRIM71,WDR12	26	35	[Cell Morphology, Molecular Transport, RNA Post-Transcriptional Modification]
ADCYAP1R1,ARHGDIB,ARID4A,ATXN7L3B,CRTC1,EMB,LOC102724788/PRODH,LRFN5,MACROH2A1,MAGEL2,MCPH1,MDFIC,Otub1,OTUD7A,PAX7,PEG3,PIP4K2C,PJA1,PLAGL1,PPA1,PPRC1,RBM24,RFFL,RNF128,S100A16,SPINK1,TM6SF1,UCHL5,USP20,USP22,USP28,USP48,USP8,VXN,ZNF428	26	35	[Amino Acid Metabolism, Gene Expression, Post-Translational Modification]
ACAT1,AGPAT1,AHCY,ALDH18A1,ALDH7A1,ATAD2,BAG4,CCDC125,CDC27,EIF1AX,EIF1B,ETFBKMT,GMNN,GPRC5A,IQCK,KIF14,KIF20B,KIF3A,LIG3,Macf1,MTDH,NCOA3,NUDT21,PRDX5,PYCR1,RBMXL2,RXRB,SIRT2,TARBP1,THOC2,TMEM132A,TRIM14,ZC3H11A,ZDHHC12,ZNF217	26	35	[Connective Tissue Disorders, Dermatological Diseases and Conditions, Developmental Disorder]

Table: 64 Top Network of molecules associated with diseases and function in CN vs NPC state.

	MEC(N) 1	MEC(N) 2	MEC(P) 1	MEC(P) 2	MEC(Prl) 1	MEC(Prl) 2
MEC(N) 1	1	0.999419403	0.988768558	0.997999983	0.999616365	0.999637651
MEC(N) 2	0.999419403	1	0.992742733	0.99899199	0.998337763	0.998953479
MEC(P) 1	0.988768558	0.992742733	1	0.995096153	0.986165454	0.988945042
MEC(P) 2	0.997999983	0.99899199	0.995096153	1	0.997298491	0.998513792
MEC(Prl) 1	0.999616365	0.998337763	0.986165454	0.997298491	1	0.99974388
MEC(Prl) 2	0.999637651	0.998953479	0.988945042	0.998513792	0.99974388	1

Table: 65 Lactogenesis miRNAseq Spearman's correlation statistics between samples and replicates p-value < 2.2e-16.

Sample name	Raw reads	GC content (%)	Q20 (%)	Q30 (%)
MEC(N) 1	2,20,52,831	48.00%	99.62	98.99
MEC(N) 2	2,33,17,021	48.00%	99.59	98.91
MEC(P) 1	1,94,16,948	48.00%	99.81	99.52
MEC(P) 2	2,02,72,824	48.00%	99.68	99.15
MEC(Prl) 1	2,21,36,347	48.00%	99.5	98.61
MEC(Prl) 2	1,97,25,994	48.00%	99.62	99.01

Table: 66 Lactogenesis miRNA sequencing raw reads and their quality reports.

Sample name	Processed reads	Mapped	Unmapped
MEC(N) 1	2,16,17,038	21490194(99.41%)	126844(0.59%)
MEC(N) 2	2,28,19,178	22565783(98.89%)	253395(1.11%)
MEC(P) 1	1,92,63,199	18890590(98.07%)	372609(1.93%)
MEC(P) 2	1,97,40,412	19468523(98.62%)	271889(1.38%)
MEC(Prl) 1	2,20,11,880	21779875(98.95%)	1232005(1.05%)
MEC(Prl) 2	1,92,63,199	19161270(99.45%)	106569(0.55%)

Table: 67 Lactogenesis processed miRNA sequencing reads and mapping percentages.

MEC (N)		**MEC (P)**		**MEC (PRL)**	
Gene	Normalize count	Gene	Normalize count	Gene	Normaliz count
mmu-miR-3535	47.155	mmu-miR-1a-3p	387.96	mmu-miR-381-3p	11.01
mmu-miR-501-3p	23.585	mmu-miR-155-5p	65.555	mmu-miR-127-3p	10.72
mmu-miR-200c-3p	17.58	mmu-miR-328-3p	50.64		
mmu-miR-210-3p	13.32	mmu-miR-196a-5p	30.57		
mmu-let-7a-1-3p	13.185	mmu-miR-704	21.445		
mmu-let-7c-2-3p	13.185	mmu-miR-221-5p	17.96		
		mmu-miR-872-5p	16.2		
		mmu-miR-708-3p	16.075		
		mmu-miR-215-5p	14.935		
		mmu-miR-450a-5p	14.82		
		mmu-miR-194-5p	14.135		
		mmu-miR-340-3p	13.345		
		mmu-miR-196a-1-3p	12.59		
		mmu-miR-330-5p	11.29		
		mmu-miR-200b-5p	11.18		
		mmu-miR-503-3p	10.755		

Table: 68 Top uniquely expressed miRNAs among MEC (N), MEC (P) and MEC (PRL) state.

MEC (N)		**MEC (P)**		**MEC (PRL)**	
Gene	Normalize Count	Gene	Normalize Count	Gene	Normalize Count
mmu-let-7c-5p	31224.7	mmu-let-7c-5p	69109.015	mmu-let-7c-5p	22036.6
mmu-miR-140-3p	17017.93	mmu-miR-140-3p	20477.615	mmu-miR-140-3p	5156.775
mmu-miR-30a-5p	14494.595	mmu-let-7b-5p	9791.93	mmu-let-7f-5p	4697.445
mmu-miR-30a-5p	14494.595	mmu-let-7i-5p	8555.065	mmu-let-7b-5p	4022.1
mmu-let-7f-5p	11869.255	mmu-miR-99a-5p	7631.505	mmu-let-7i-5p	3901.535
mmu-miR-183-5p	10128.13	mmu-miR-10a-5p	5706.365	mmu-miR-10a-5p	2424.67
mmu-miR-146b-5p	5548.135	mmu-miR-151-3p	4059.445	mmu-miR-148b-3p	2347.96
mmu-miR-30d-5p	5021.215	mmu-miR-200a-5p	2647.475	mmu-miR-99a-5p	2342.94
mmu-miR-21a-5p	4609.71	mmu-miR-24-3p	2482.235	mmu-miR-21a-5p	2236.45
mmu-let-7i-5p	3355.53	mmu-miR-429-3p	1390.84	mmu-miR-182-5p	2027.245
mmu-miR-99a-5p	2278.915	mmu-miR-30a-3p	649.05	mmu-miR-183-5p	1623.095
mmu-miR-182-5p	2197.815	mmu-miR-1a-3p	387.96	mmu-miR-151-3p	1319.18
mmu-miR-10a-5p	2194.59	mmu-miR-205-5p	355.13	mmu-miR-200a-5p	1057.955
mmu-let-7b-5p	1985.29	mmu-miR-125a-5p	304.575	mmu-let-7a-5p	958.315
mmu-miR-148b-3p	1678.84	mmu-miR-24-2-5p	282.63	mmu-miR-27b-3p	605.51
mmu-miR-200b-3p	1517.17	mmu-miR-451a	269.99	mmu-miR-30d-5p	578.33

Table: 69 Highly expressed miRNAs among MEC (N), MEC (P) and MEC (PRL) state.

MEC (P) vs MEC (N)		**MEC (PRL) vs MEC (P)**	
miRNA	log2FoldChange	miRNA	log2FoldChange
mmu-miR-1843a-3p	8.41635249	mmu-miR-292a-3p	4.730596851
mmu-miR-365-2-5p	6.652287666	mmu-miR-122-3p	4.505284654
mmu-miR-499-5p	5.2203524	mmu-miR-291a-5p	4.50208051
mmu-miR-3474	4.964323324	mmu-miR-127-5p	2.437933802
mmu-miR-199a-5p	4.155263006		
mmu-miR-1a-3p	3.831039868		
mmu-miR-1843b-3p	3.658140592		
mmu-miR-1981-5p	3.548591001		
mmu-miR-126a-3p	3.51492022		
mmu-miR-542-5p	3.485427114		
mmu-miR-671-3p	3.343794576		
mmu-miR-222-3p	3.241035414		
mmu-miR-125a-5p	2.973613806		
mmu-miR-151-3p	2.91000943		
mmu-miR-204-5p	2.862736512		
mmu-miR-7043-3p	2.820626307		
mmu-miR-429-3p	2.74889702		
mmu-miR-669c-5p	2.699855288		
mmu-miR-708-3p	2.588787678		
mmu-miR-155-5p	2.453792734		

Table: 70 Highly upregulated miRNAs during MEC (P) vs MEC (N) and MEC (PRL) vs MEC (P).

MEC (P)_vs_MEC (N)	
miRNA	log2FoldChange
mmu-miR-25-5p	-5.183707002
mmu-miR-185-5p	-4.174040998
mmu-miR-122-5p	-4.052650645
mmu-miR-292a-5p	-4.004922511
mmu-miR-340-5p	-3.996313329
mmu-miR-291a-3p	-3.874345495
mmu-miR-16-5p	-3.780642917
mmu-miR-16-5p	-3.770619355
mmu-miR-293-5p	-3.753622085
mmu-miR-31-5p	-3.53280134
mmu-miR-200c-3p	-3.461139584
mmu-miR-30a-5p	-3.418104361
mmu-miR-542-3p	-2.987661195
mmu-miR-3535	-2.98137172
mmu-miR-30e-5p	-2.717101077
mmu-miR-19b-3p	-2.626223569
mmu-miR-19b-3p	-2.626223569
mmu-let-7a-1-3p	-2.573792859
mmu-let-7c-2-3p	-2.573792859
mmu-let-7a-1-3p	-2.573792859
mmu-let-7c-2-3p	-2.573792859
mmu-miR-203-3p	-2.544796859
mmu-miR-378b	-2.271966962
mmu-miR-146a-5p	-2.220605932
mmu-miR-30d-5p	-2.208601647
mmu-miR-146b-5p	-2.121794232
mmu-miR-378d	-2.114879854
mmu-miR-183-5p	-2.107124831
mmu-let-7e-5p	-1.970257014

Table: 71 Highly downregulated miRNAs during MEC (P) vs MEC (N).

miRNAs	MEC (N)	MEC (P)	MEC (PRL)
mmu-miR-467b-5p	72.22	542.125	57.375
mmu-miR-467a-5p	72.22	542.125	57.375
mmu-let-7j	63.98	159.46	56.43
mmu-miR-5099	49.655	164.125	135.145
mmu-miR-467c-5p	22.35	61.285	13.31
mmu-miR-669p-5p	4.6	10.15	1.575
mmu-miR-669a-5p	4.6	10.15	1.575
mmu-miR-466b-3p	2.635	2.53	0.56
mmu-miR-466p-3p	2.635	2.53	0.56
mmu-miR-466c-3p	2.635	2.53	0.56
mmu-miR-466e-3p	2.635	2.53	0.56
mmu-miR-466a-3p	2.635	2.53	0.56
mmu-miR-669c-5p	2.475	20.5	31.825
mmu-miR-881-3p	2.36	1.96	1.505
mmu-miR-669a-3p	2.23	2.6	0.44
mmu-miR-669o-3p	2.23	2.6	0.44
mmu-miR-467e-5p	1.9	5.145	0.835
mmu-miR-467a-3p	1.075	0.385	0.295
mmu-miR-467d-3p	1.075	0.385	0.295
mmu-miR-466d-3p	1.055	1.505	0.44
mmu-miR-466c-5p	0.78	2.07	0.3
mmu-miR-470-5p	0.77	0.535	0.35
mmu-miR-465c-5p	0.745	0.195	0.3
mmu-miR-871-3p	0.705	0.315	0.745
mmu-miR-504-5p	0.7	3.59	0.145
mmu-miR-669o-5p	0.465	1.485	0.35
mmu-miR-297b-3p	0.455	0.235	0.025
mmu-miR-297c-3p	0.455	0.235	0.025
mmu-miR-297a-3p	0.455	0.235	0.025
mmu-miR-223-5p	0.44	0.17	0.685
mmu-miR-217-5p	0.4	0.48	0.46
mmu-miR-467d-5p	0.395	1.425	0.2
mmu-miR-466b-5p	0.28	0.425	0.05
mmu-miR-466o-5p	0.28	0.425	0.05
mmu-miR-466f-3p	0.265	0.355	0.125
mmu-miR-466f-5p	0.235	0.385	0.045
mmu-miR-669d-5p	0.22	0.215	0.025
mmu-miR-465b-5p	0.195	0.05	0.18
mmu-miR-669l-5p	0.16	0.99	0.53

Table: 72 Expression status of miRNAs present in cLADs.

miRNAs	MEC (N)	MEC (P)	MEC (PRL)
mmu-miR-148a-3p	820572.52	720332.385	905889.005
mmu-let-7c-5p	31224.7	69109.015	22036.6
mmu-miR-140-3p	17017.93	20477.615	5156.775
mmu-let-7f-5p	11869.255	6572.645	4697.445
mmu-miR-183-5p	10128.13	3279.475	1623.095
mmu-miR-146b-5p	5548.135	1710.305	376.1
mmu-miR-21a-5p	4609.71	7138.89	2236.45
mmu-let-7i-5p	3355.53	8555.065	3901.535
mmu-miR-182-5p	2197.815	9090.34	2027.245
mmu-miR-10a-5p	2194.59	5706.365	2424.67
mmu-let-7b-5p	1985.29	9791.93	4022.1
mmu-miR-148b-3p	1678.84	1891.29	2347.96
mmu-let-7g-5p	1201.315	1104.2	446.14
mmu-let-7a-5p	1175.23	1195.42	958.315
mmu-miR-27b-3p	964.805	1569.17	605.51
mmu-miR-24-3p	826.695	2482.235	254.32
mmu-miR-320-3p	616.98	526.115	118.25
mmu-miR-378a-3p	491.245	755.635	131
mmu-miR-26a-5p	464.25	456.965	239.3
mmu-miR-99b-5p	391.705	1157.025	239.44
mmu-miR-151-3p	385.225	4059.445	1319.18
mmu-miR-27a-3p	323.665	523.96	160.26
mmu-miR-30e-5p	289.24	60.335	17.815
mmu-miR-25-3p	282.6	1064.305	147.935
mmu-miR-191-5p	269.705	160.81	31.75
mmu-miR-185-5p	253.48	18.045	7.63
mmu-let-7e-5p	241.65	84.295	64.005
mmu-miR-152-3p	239.435	153.775	76.825
mmu-miR-203-3p	213.125	50.17	18.83
mmu-miR-103-3p	205.855	186.455	40.46
mmu-miR-9-5p	200.435	300.395	159.265
mmu-miR-542-3p	187.25	42.885	13.585
mmu-let-7d-5p	171.985	173.985	60.875
mmu-miR-7a-5p	169.34	198.645	106.32
mmu-miR-24-2-5p	166.495	282.63	84.55
mmu-miR-16-5p	135.02	13.295	4.06
mmu-miR-96-5p	134.91	94.275	36.725
mmu-miR-140-5p	131.94	186.305	72.58
mmu-miR-340-5p	118.775	10.62	3.535

Table: 73 Expression status of miRNAs present in iLADs.

miRNAs	MEC (N)	MEC (P)	MEC (PRL)
mmu-miR-30a-5p	14494.595	2120.375	545.99
mmu-miR-30d-5p	5021.215	1467.155	578.33
mmu-miR-99a-5p	2278.915	7631.505	2342.94
mmu-miR-31-5p	1230.255	143.61	34.33
mmu-miR-30a-3p	135.06	649.05	288.975
mmu-miR-30c-5p	114.93	118.06	21.93
mmu-miR-378d	93.035	29.375	8.62
mmu-miR-146a-5p	83.565	24.495	10.215
mmu-miR-125b-5p	78.13	259.575	50.615
mmu-miR-30c-2-3p	44.26	50.725	21.4
mmu-miR-181b-5p	33.365	91.6	22.64
mmu-miR-361-3p	25.67	72.01	15.845
mmu-miR-222-3p	16.69	213.275	135.815
mmu-miR-221-3p	15.445	65.83	19.91
mmu-miR-125b-2-3p	14.52	39.32	11.215
mmu-miR-206-3p	13.185	10.065	4.46
mmu-miR-181a-5p	12.135	46.26	17.39
mmu-miR-224-5p	12.08	75.365	16.22
mmu-miR-1843b-5p	9.445	15.69	7.44
mmu-miR-30b-5p	8.835	7.38	1.3
mmu-miR-872-5p	8.705	16.2	3.675
mmu-miR-221-5p	7.71	17.96	5.17
mmu-miR-381-3p	7.7	7.08	11.01
mmu-miR-744-5p	7.18	9.83	5.71
mmu-miR-30b-3p	6.43	9.48	3.095
mmu-miR-379-5p	6.14	5.405	6.63
mmu-miR-704	5.67	21.445	6.92
mmu-miR-107-3p	5.27	5.17	1.74
mmu-miR-186-5p	4.82	9.765	2.35
mmu-miR-361-5p	3.94	5.34	1.575
mmu-miR-582-3p	3.29	4.035	1.575
mmu-miR-370-3p	2.045	1.655	1.39
mmu-miR-30d-3p	2.025	3.79	1.58
mmu-miR-31-3p	1.77	3.465	0.855
mmu-miR-34c-5p	1.725	2.655	2.515
mmu-miR-1843b-3p	1.49	27.3	10.285
mmu-let-7c-1-3p	1.375	3.785	0.21
mmu-miR-134-5p	1.075	0.775	1.43
mmu-miR-204-5p	0.98	9.785	3.38

Table: 74 Expression status of miRNAs present in fLADs

Molecules in Network	Score	Focus Molecules	Top Diseases and Functions
AFF3,AIFM3,Anp32e,ASPM,C10orf90,CCT3,DUT,FGF11,FKBP5,FZD4,GABBR1,ILF2,ITM2C,KLHDC8B,KLHL33,KPNA2,MCC,MCM2,MCM3,MTR,NBEA,PIPOX,POC1A,RUVBL1,RUVBL2,SH2D5,SYT13,TMEM151A,TTC36,TUBA1A,TXNIP,UBC,UHRF1,XRCC5,Zim1	42	35	[Cancer, Organismal Injury and Abnormalities, Reproductive System Disease]
AEBP1,ARHGAP26,ARHGAP42,BIN3,BUB1,CCDC18,CDC42EP2,CENPK,CENPL,CENPM,CENPN,CENPT,CENPW,DSN1,FGF13,GMDS,HID1,LIME1,Mapk,MIA,NDC80,NSL1,NUF2,PKN3,PMF1/PMF1-BGLAP,PSRC1,SKA1,SKA2,SKA3,SPC24,SPC25,SPRED2,SYNM,TPD52L1,TTLL12	39	34	[Cell Cycle, Cellular Assembly and Organization, DNA Replication, Recombination, and Repair]
ACTR3B,ALG10,CDC42EP3,DPEP1,FAM131B,FAM214B,FOXL1,gelatinase,GULP1,H19,IL1,KIF20A,KIF4A,LMAN1L,LMNB2,LONRF1,LRRK1,MAP7D2,NCAPD2,NCAPD3,NCAPG,NCAPG2,NCAPH,NEK6,PANX1,PANX3,POGK,SMC2,SMC4,SYNGR3,TLCD4,TTF2,UCK2,UMPS,UTP20	37	33	[Cell Cycle, Cellular Assembly and Organization, DNA Replication, Recombination, and Repair]
CEP128,CEP135,CEP192,CEP55,CEP57L1,CEP72,FAM83D,GEN1,HAUS1,HAUS4,HAUS5,HAUS6,HAUS8,HYLS1,KIF23,MASTL,MEX3B,MTORC1,PLEKHG1,PLK4,Rab5,RACGAP1,RTTN,SHCBP1,SIPA1L2,SLC1A5,SLC26A8,SPATA24,SPATA2L,SYBU,TBC1D31,TUBG1,WDR62,WDR90,YWHAH	37	33	[Cell Cycle, Cellular Assembly and Organization, DNA Replication, Recombination, and Repair]
14-3-3,AIRE,AKAP12,ARHGAP25,ARHGEF39,AZGP1,C15orf39,CAMK2D,CD34,CDH19,CKAP5,COBL,Cr3,GM2A,HEXA,HP,LMO4,LTF,MAP4K1,MCM5,Nradd,PCMTD1,PIGR,RBPMS,RPA2,RPA3,SORT1,SPIRE2,SPN,SSRP1,SYK,THBS3,TOP2A,TUBB,ZWILCH	37	33	[Cancer, Cellular Development, Cellular Growth and Proliferation]
26s Proteasome,ACKR3,ADSS1,ANKRD35,C4orf47,CTSD,DNMT1,EZH2,FAM107A,FHOD3,FZD8,H2AX,HCAR2,Histone h4,IFRD2,KCTD14,KLF4,LYAR,mir-8,MKI67,MSI2,NPM1,NUDT21,ODC1,PARD3B,PARP1,PKD1,PLAU,PLEKHF1,PTN,RAMP3,TSPAN1,VIM,VIPR1,Wfdc3	37	33	[Cancer, Hematological Disease, Organismal Injury and Abnormalities]
ABCB9,ABHD14B,AMT,ATIC,CDK2-CyclinE,CEACAM,DNPH1,Gar1,GCSH,GGH,GLS2,Hmgn2 (includes others),IER5,ITGBL1,KNDC1,LRRC8B,LRWD1,ME2,MLANA,MYC,NHP2,NIBAN1,Npm,PDZD2,PERP,PPAT,SARDH,SCAMP5,SEC31B,SEPHS2,SLC38A1,Snrpc,Sprr1b,YY2,ZNF385B	35	32	[Amino Acid Metabolism, Post-Translational Modification, Small Molecule Biochemistry]
AGAP1,ARRB1,ATAD2,ATP1B1,ATPase,AURK,BLM,CDC25B,CGN,CHD1L,CMBL,DDX39A,DEPDC1B,DNA2,FAM110B,GAB2,H2BC17,KIF13B,KIF14,KIF18A,KIF18B,KIF20B,MELK,MICALL1,MXD3,MYH1,MYH6,PDRG1,plus-end-directed kinesin ATPase,PPM1H,PRC1,SESN1,SH3RF3,SLC38A2,SLFN13	35	32	[Cancer, Gastrointestinal Disease, Organismal Injury and Abnormalities]
AKAP8L,C9orf152,CBX5,CDC6,CDCA4,CLEC3A,DLX3,EFHD1,GCNT4,GMNN,GRIK5,HDAC11,HISTONE,Histone h3,HOXC9,Id,ID4,JADE1,LBR,LRCH1,OSGIN1,PROM2,PRR15L,PTGR1,RBM14,RNF43,SDSL,SERP2,SMAD6,TBC1D4,TM4SF1,TMPO,TRIM16,TUBE1,ZBTB7C	35	32	[Cell Cycle, Connective Tissue Disorders, Hair and Skin Development and Function]
alcohol group acceptor phosphotransferase,BANF1,BCL7C,CNTROB,CRYBB3,DAPK1,DMPK,DPF1,ESPL1,FAM126B,GPATCH4,HECTD2,KIF24,KIF2C,KLF5,LENG8,LPAR1,MYBBP1A,NEIL3,NEK2,PAPSS2,PLEKHO1,PLK1,POP1,PRKCH,Ribosomal 40s subunit,Rnr,RPS2,RPSA,SRPX,THOP1,TTK,UTP14C,YBX2,ZMYM5	35	32	[Cell Cycle, Cellular Assembly and Organization, DNA Replication, Recombination, and Repair]
ABLIM1,ANLN,ANTXR1,B4GALNT4,BASP1,Beta Arrestin,CAD,CaMKII,CFL1,CHAF1A,CHRM1,CORO1C,DNA-PK,EIF5A,FAM126A,FLNA,HMGB2,KIF11,LHFPL2,LRRC59,MYH10,MYO6,NUDCD2,NUPR1,P2RX5,PDLIM7,PLCB1,RBM3,RTN4RL2,SHPK,SPDL1,STIL,TMOD1,ZGRF1,ZSWIM6	35	32	[Developmental Disorder, Hereditary Disorder, Organismal Injury and Abnormalities]
ACE2,ADCY9,ANKRD23,AREG,ARHGAP11A,AURKB,BIK,C17orf53,CBX7,CD27,CDCA2,CDCA3,CDCA7,CDCA8,CPED1,DCUN1D4,FOXM1,GSTCD,Importin alpha/beta,Importin beta,KANK4,KIF22,let-7,MAD2L1,NAP1L1,Nucleoporin,OSCP1,PHF19,Polycomb,RFC3,RHBDL3,RRM1,SIVA1,TPX2,TTLL10	33	31	[Cell Cycle, Cellular Assembly and Organization, DNA Replication, Recombination, and Repair]

Table: 75 Top Network of molecules associated with diseases and function in MEC(P) vs MEC(N)state.

Molecules in Network	Score	Focus Molecules	Top Diseases and Functions
Aconitase,ATAD2,ATPase,CARNS1,CENPE,CIT,DACT1,DGKH,DQX1,ERCC6L,FAM199X,FAM83D,Fetal Hemoglobin,HMMR,Jnk,KCMF1,KIF14,KIF15,KIF18A,KIF18B,KIF20B,KIF23,KIF4A,KNSTRN,MAFF,MAFK,plus-end-directed kinesin ATPase,PRC1,RBM41,RFK,SH3BP5,SPAG5,TROAP,USP34,ZNF354C	41	30	[Cell Cycle, Cellular Assembly and Organization, Cellular Movement]
Angiotensin II receptor type 1,ANKRD26,CCDC18,CCDC66,CENPA,CENPH,CENPI,CENPK,CENPM,CENPT,CSPP1,EDAR,Fascin,FOLR1,Hif,IFT74,IFT80,IFT81,LRRCC1,NCOR-LXR-Oxysterol-RXR-9 cis RA,NDC80,NFkB (complex),NUF2,PCM1,peptidase,RUSC2,SKA3,SLC13A2,SLC15A2,SPC24,TEX9,Trim30a/Trim30d,TSPAN33,VSNL1,Wfdc17	39	29	[Connective Tissue Disorders, Developmental Disorder, Hereditary Disorder]
ADCY9,AKAP12,ANGPTL6,APOA4,ARNTL2,BRWD3,BTN1A1,CCL28,CKAP2,CSN3,EFNA3,EGLN,EGLN1,ERO1A,FAIM2,Ferritin,FTL,hemoglobin,IKBIP,LZTFL1,MSRA,OSR1,PDGFRL,Pdi,Pgk,RRAGD,SELENBP1,SLC20A1,SMTNL2,STC1,Sult1d1,TEAD4,TPX2,UBE2C,Vegf	39	29	[Cellular Function and Maintenance, Connective Tissue Development and Function, Reproductive System Development and Function]
AChR,ADRB,ANO1,ARRDC4,ASNS,ATF3,ATF4,ATF5,CACNA1G,Calmodulin,CEBPD,CKAP2L,CTH,CTXN1,DDIT3,DDIT4,ENTPD3,FAM107A,FAM111A,GUCY,Gucy1b2,JAK1/2,MTHFD2,NUPR1,PIGR,PPP1R15A,Proinsulin,SAMD4A,SGPP2,SLC7A11,TENM4,TRIB3,UBIQUITIN LIGASE,XBP1,ZNF488	37	28	[Cancer, Cell Morphology, Cellular Function and Maintenance]
ATAD5,AURKA,BRCA1,C10orf90,CCNC,CFP,CRYBB3,ECT2,FANCD2,FXYD6,GADD45A,GTPase,Ige,IgG1,Igg3,IKK (complex),Immunoglobulin,KCNA6,KLF4,MKI67,MSS51,MTORC1,PCMTD1,PLK1,PRPF39,SLC22A23,SLC5A6,SMC3,SRPX,SYNE4,TPR,UBA6,UBC,Zim1,ZNF280D	37	28	[Cell Cycle, Cellular Assembly and Organization, DNA Replication, Recombination, and Repair]
48s,ACSS1,ASF1B,CDCA2,CEACAM,CRIP1,Crnde,E2f,Enolase,ESCO1,ESCO2,ESF1,FAR1,FERMT1,GPX2,HIGD1A,IMPACT,ITGBL1,Mcm,MCM10,MCM5,MCM6,MIRLET7,MIS18BP1,MLANA,MYC,MYO9A,PEG3,POLA1,pyruvate dehydrogenase,Smad2/3-Smad4,SMOC2,SPEG,TONSL,WDHD1	35	27	[Cell Cycle, DNA Replication, Recombination, and Repair, Lipid Metabolism]
arginase,AURKB,BTLA,C/EBP,CCDC13,CCDC71L,CCP110,CDCA3,CDCA8,CEP162,CEP192,CEP290,CNTRL,Ctbp,cytokine,DEK,GPBP1,HAUS3,histone deacetylase,Hnrnpa3,KIF20A,KIF2C,NCAPG,NCAPH,NEK1,Pias,Psg18 (includes others),RBM12B,RPGRIP1L,RPS6KA,RRP12,SARS1,SMC2,TIA1,Top2	35	27	[Cell Morphology, Cellular Assembly and Organization, DNA Replication, Recombination, and Repair]
BCLAF1,CAPNS2,Casein,Casp12,CHM,CK1,CPM,CST6,DEPTOR,Diap,Ecm,FHL1,FTH1,HPN,IGFBP5,KLK10,KLK3,MTORC2,NUP210,NUSAP1,PBK,POF1B,PRSS23,PRSS53,PSRC1,RAB17,Ras homolog,RGPD4 (includes others),SDF2L1,Serine Protease,SPINK5,TRPV6,VitaminD3-VDR-RXR,ZBTB26,ZBTB6	35	27	[Post-Translational Modification, Protein Degradation, Tissue Morphology]
ACTA2,ACTBL2,ATP11C,CEACAM1,CG,Collagen type I (complex),CXCR4,EDN2,ERG,ESPL1,FABP4,FABP5,FDX1,FKBP5,FSH,G-Actin,GNRH1,let-7,Lh,Notch,PLAU,PLC,Rap1,RAPGEF5,RNPC3,SPDL1,SRF,STEAP1,TMEM44,TSPAN13,TTC7B,VCAN,ZNF433,ZNF519,ZNF644	35	27	[Cancer, Cellular Movement, Organismal Injury and Abnormalities]
ALDOA,ARL4D,C16orf89,CWF19L2,DCUN1D1,ENO2,GAPDH,GBP7,Hdac,HNRNPH1,Hsp70,Hsp90,HSPA1A/HSPA1B,HSPA9,IFI30,Interferon alpha,IQGAP3,MAST1,MGP,NDRG1,Ngf,p85 (pik3r),PLXNA3,RAD51AP1,RAD54L,RNA polymerase II,RNASE1,RNF207,SAMD9L,Secretase gamma,SLC7A5,TERF1,TRAPPC2,ZC3HAV1L,ZFAND2A	35	27	[Cancer, Cellular Assembly and Organization, Cellular Function and Maintenance]
14-3-3,Ap1,BUB1,Ccl9,CYB5R1,EIF4EBP1,FOXJ1,FOXRED2,FRZB,Hsp27,IGFBP3,Keratin,KRT17,KRT18,KRT5,MELK,mir-10,MRPS18B,NFAT (complex),NFAT5,p70 S6k,PDGF BB,PNISR,PNN,PPP2R2C,Rsk,RSPO3,S100A4,SERTAD4,SLC7A1,SRSF10,TIMP1,TRA2A,TSH,VEGFA	32	26	[Cancer, Cell Death and Survival, Cellular Movement]
20s proteasome,26s Proteasome,ATP1B1,BARD1,BHLHE40,BIRC5,Caspase 3/7,CHCHD10,cytochrome-c oxidase,DLGAP5,DZIP3,FKBP3,GPI,HLTF,HSPB8,IAP,IL1,MAP1LC3,P glycoprotein,Perm1,PLCH2,PLK4,RIF1,SAT1,SQSTM1,STIL,TFAP4,TK1,TKT,TRIM5,UBE2V2,Ubiquitin,UBN2,XIAP,ZFP36L2	32	26	[Cell Death and Survival, Hereditary Disorder, Neurological Disease]

Table: 76 Top Network of molecules associated with diseases and function in MEC(Prl) vs MEC(P) state.

Antibodies	Dilution	Catalog no	Company	Mol. Wt
ß-ACTIN	1:1500	3700	Cell Signaling	45 Kda
GAPDH	1:1500	SC-32233	Santa Cruz	37 Kda
ß-CASEIN	1:1500	SC-166530	Santa Cruz	29 Kda
KRT8	1:1500	SC-8020	Santa Cruz	40-55 Kda
KRT14	1:1500	SC-53253	Santa Cruz	50 Kda
IGFII	1:1500	SC-515805	Santa Cruz	8 Kda
PER2	1:1500	SC-377290	Santa Cruz	140 Kda
p-GSK3ß	1:1500	SC-373800	Santa Cruz	47 Kda
GYS	1:1500	3886	Cell Signaling	84 Kda
p-GYS	1:1500	3891	Cell Signaling	85-90 Kda
TUBB3/TUJ1	1:1500	MMS-435P	Bio Legend	50 Kda
PBXIP1	1:1500	A301-628A	Bethyl Lab.	110 Kda
GFAP	1:1500	12389	Cell Signaling	50 Kda
MAG	1:1500	SC-166849	Santa Cruz	100 Kda

Table: 77 List of Antibodies used in studies.

miRNA Gene names	Forward Primer
mmu-miR-293-5p	ACTCAAACTGTGTGACATTTTG
mmu-miR-19b-3p	TGTGCAAATCCATGCAAAACTGA
mmu-miR-122-5p	TGGAGTGTGACAATGGTGTTTG
mmu-miR-155-5p	TTAATGCTAATTGTGATAGGGGT
mmu-miR-1a-3p	TGGAATGTAAAGAAGTATGTAT
mmu-miR-149-5p	TCTGGCTCCGTGTCTTCACTCCC
mmu-let-7a-1-3p	CTATACAATCTACTGTCTTTCC
mmu-let-7c-2-3p	CTATACAATCTACTGTCTTTCC

Table: 78 List of miRNA primers used during Lactogenesis study.

miRNA Gene names	Forward Primer
mmu-miR-9-5p	UCTTTGGTTATCTAGCTGTATGA
mmu-miR-124-3p	TAAGGCACGCGGTGAATGCC
mmu-let-7b-5p	TGAGGTAGTAGGTTGTGTGGTT
mmu-mir-466f-3p	CATACACACACACATACACAC
mmu-miR-302c-3p	AAGTGCTTCCATGTTTCAGTGG
mmu-miR-301b-3p	CAGTGCAATGGTATGTCAAAGC
mmu-miR-293-5p	ACTCAAACTGTGTGACATTTTG
mmu-miR-292a-5p	ACUCAAACUGGGGGCUCUUUUG
mmu-miR-291a-3p	AAAGUGCUUCCACUUUGUGUGC

Table: 79 List of miRNA primers used in neurogenesis.

Gene	Forward Primer	Reverse Primer	Product size (bp)
ß-Actin	TTACTGCTCTGGCTCCTAGCA	GACTCATCGTACTCCTGCTTGC	145
ß-Casein	CCTCCTCTCTTGTCCTCCAC	TGTTCAACAGATTCCTCACTGG	123
Kif22	CCTGTGTCCGAGCCATAGAC	CTGAGTGCTCTTCTCGCCAT	110
Kif11	ATTAAGGATGGCAGTGCGAA	GTGCTGTCGTGGTAATGGTG	112
Tpx2	CTTACTCTTTCGATGCCCCC	TCTCCAAGTTGGCCTTCTCA	108
Nek2	TCTGATGGCTTGAATGACCTC	TCCTTTGCTCTTCTGCAACC	121
Krt15	CAGATCGGGACTACAGCCAT	GTCAATCTCCAGGACAACGC	101
Wap	TGCCTCATCAGCCTTGTTCT	CACACTCCTCGTTGGTTTGG	157
Capn6	TAACAACCGTGATACCTTCTTGC	GCGGTAAGTGCGTAGGTCC	109
Boc	CATTCTCACACTCTCGCACC	ACAAGAGGACACACACCACG	132
Gas6	ATACCTGCCACTGTGATGGG	GGCCCAGGTACAAGGACTTC	123
Per1	AACTTCGACTGCCACCAGAG	ACCCTCCTCCAGACTCCACT	134
Per2	CTCCAGGAAGACGTGGACAT	TGTGCTCTGCCTCTGTCATC	184
Igf2	GGAGGGGAGCTTGTTGACAC	GGGGTGGCACAGTATGTCTC	159
Ctgf	CATTCTAGCCAGACAGCTCCA	CTCCACCCGAGTTACCAATG	151
Ndrg3	CCACCCGAGTTACCAATGAC	ACACTGGTGCAGCCAGAAAG	175
Prom1	GGCCAAGTACTATCGCAGGA	GACCACTGATGCCATGTTCC	172
Gys1	ATCTGGTGGGACCATACACG	CCTACATCCAGGAGCACCAC	172
Sec23ip	GACCCCCTGTGCAGACATAC	GCACACTGAAAGGCATCCAT	224
Paxip1	TGCAGAATCAAGCAGCACAC	CGACATCTGCTCGGGATAGT	212
Tmed3	ACTTCCAAGTGGGTGACGAG	GGGCAATAGTCTCGCCTACA	219
Krt14	GCCAACACTGAACTGGAGGT	GTCGATCTGCAGGAGGACAT	159
Krt8	AAGTTCGTGCCCAGTACGAG	CGGAGATCTCTGTCTTTGTGC	141
Gsk3-ß	CCTCTGGCCACCATCCTTAT	CCACGGTCTCCAGCATTAGT	104

Table: 80 List of mRNA primers used during Lactogenesis study.

Gene	Forward Primer	Reverse Primer	Product size (bp)
Gapdh	TTACTGCTCTGGCTCCTAGCA	GACTCATCGTACTCCTGCTTGC	145
Pou5f1	GGCGTTCTCTTTGGAAAGGTGTTC	CTCGAACCACATCCTTCTCT	359
Nanog	CTCAAGTCCTGAGGCTGACA	TGAAACCTGTCCTTGAGTGC	120
Sox2	CCGCGTCAAGAGGCCCATGAA	CCCGCTTCTCGGTCTCGGACAA	149
Pax6	TAACGGAGAAGACTCGGATGAAGC	CGGGCAAACACATCTGGATAATGG	140
Nestin	CTCTTCCCCCTTGCCTAATACC	TTTAGGATAGGGAGCCTCAGACAT	133
Tuj1	AAGGTAGCCGTGTGTGACATC	ACCAGGTCATTCATGTTGCTC	201
Tau/Mapt	CTTTGAACCAGTATGGCTGACCCT	CGAGGTGTGGCGATCTTCG	157
Vglut	TAACAACCGTGATACCTTCTTGC	GCGGTAAGTGCGTAGGTCC	109
Gfap	CATTCTCACACTCTCGCACC	ACAAGAGGACACACACCACG	132
Dlx1	ATACCTGCCACTGTGATGGG	GGCCCAGGTACAAGGACTTC	123
Mag	CCTGGGCCTACGAAACTGTA	AACTGACCTCCACTTCCGTTC	194
Mog	GAGCAAGCACCTGAATACCG	GGGGTTGACCCAATAGAAGG	179
Pbx1	GGGTGCAGGTTCAGACAACT	GCTTTGCTCTCGAAGGAGGT	151
Pbxip1	TCCACAACTATGGCCTCCTG	CCATCCAAGGTCCCAGCTAA	182
Trk-ß	CAGTATTAACTCGCTTCTGGC	TTCATCCACGTCAAAGGCAG	281
Blbp	CAGTCAGGAAGGTGGCAAG	CACCGGATAAAGCTGCCTCT	172
NeuroD1	TTAAATTAAGGCGCATGAAGGCC	GGACTGGTAGGAGTAGGGATG	374
Sox1	GGAAAACCCCAAGATGCACAAC	CGCAGTCTCTTGGCCTCGTC	111
Foxg1	CAGCACTTTGAGTTACAACG	TGGTCTGCGAAGTCATTGAC	326
Neurog1	ATGCCTGCCCCTTTGGAGAC	TGCATGCGGTTGCGCTCGC	320
Neurog2	GCTGGCATCTGCTCTATTCC	ATGAAGCAATCCTCCCTCCT	342
Gli3	ACCGTTCAAAGCCCAGTACA	CAGACGTATGGCTTCTCTCCA	150
Sox3	CAGGCAACGGGGGCAGCGGG	CCGCATCGGTCAGCAGTTTC	210
Hes6	CTCCCTCGTGTTCACCTCTC	GAGGAGCAGCTTCAGTGACC	204
Otx2	ACAAGTGGCCAGTTCAGTCC	CTGGGTGGAAAGAGAAGCTG	345
Notch1	CGGTGAACAATGTGGATGCT	ACTTTGGCAGTCTCATAGCT	127
Olig2	CACAGGAGGGACTGTGTCCT	GGTGCTGGAGGAAGATGACT	145
Gata6	GAACGTACCACCACCACCAT	CCATGTAGGGCGAGTAGGTC	51
Gata4	CCCCAATCTCGTAGATATGTTTGAT	GTCCCATCTCGCCTCCAG	110
Tcf712	CGAGATAAATCCCGGGAAAG	GGGATCATGATGAAGGGGTAG	101
Id2	GCAAAGCTTCACGCTAAACC	GAATTGCCATTGGTGGAAGG	138
Id4	AGACTCACCCTGCTTTGCTG	ATGCTGTCACCCTGCTTGTT	148
Sox10	CACGGTTTTCCACTTCCTCA	GTCTTGTTCCTCGGCCATGT	151
Hes5	AGCAGCATAGAGCAGCTGAAG	TAGTCCTGGTGCAGGCTCTT	164
Zeb2	CATTCCCTCATACGGTCAGG	AGAGCGGATCAGATGGCAGT	151
Nkx6-2	ATGACCGAGAGCCAAGTGAA	CCGGTTGTATTCGTCATCGT	162
Myrf	CTGCAACGGGACCTCTACAT	TAGAGGGGTGTGGAGGGAGT	160

Table: 81 List of mRNA primers used in neurogenesis.

References

1. Agger, K., et al. (2007). "UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development." Nature **449**(7163): 731-734.
2. Agirman, G., et al. (2017). "Cerebral cortex development: an outside-in perspective." FEBS Lett **591**(24): 3978-3992.
3. Alarcon, C. R., et al. (2015). "N6-methyladenosine marks primary microRNAs for processing." Nature **519**(7544): 482-485.
4. Albright, S. C., et al. (1979). "Histone molar ratios among different electrophoretic forms of mono- and dinucleosomes." J Biol Chem **254**(4): 1065-1073.
5. Allan, J., et al. (2012). "Micrococcal nuclease does not substantially bias nucleosome mapping." J Mol Biol **417**(3): 152-164.
6. Allar, M. A. and T. L. Wood (2004). "Expression of the insulin-like growth factor binding proteins during postnatal development of the murine mammary gland." Endocrinology **145**(5): 2467-2477.
7. Anders, S. and W. Huber (2010). "Differential expression analysis for sequence count data." Genome Biol **11**(10): R106.
8. Arun, P. V., et al. (2016). "Identification and functional analysis of essential, conserved, housekeeping and duplicated genes." FEBS Lett **590**(10): 1428-1437.
9. Arvey, A., et al. (2012). "Sequence and chromatin determinants of cell-type-specific transcription factor binding." Genome Res **22**(9): 1723-1734.
10. Ashburner, M., et al. (2000). "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium." Nat Genet **25**(1): 25-29.
11. Asselin-Labat, M. L., et al. (2006). "Steroid hormone receptor status of mouse mammary stem cells." J Natl Cancer Inst **98**(14): 1011-1014.
12. Asselin-Labat, M. L., et al. (2007). "Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation." Nat Cell Biol **9**(2): 201-209.
13. Audergon, P. N., et al. (2015). "Epigenetics. Restricted epigenetic inheritance of H3K9 methylation." Science **348**(6230): 132-135.
14. Aydogdu, E., et al. (2012). "MicroRNA-regulated gene networks during mammary cell differentiation are associated with breast cancer." Carcinogenesis **33**(8): 1502-1511.
15. Baarends, W. M., et al. (2007). "Increased phosphorylation and dimethylation of XY body histones in the Hr6b-knockout mouse is associated with derepression of the X chromosome." J Cell Sci **120**(Pt 11): 1841-1851.
16. Bach, K., et al. (2017). "Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing." Nat Commun **8**(1): 2128.
17. Barber, C. M., et al. (2004). "The enhancement of histone H4 and H2A serine 1 phosphorylation during mitosis and S-phase is evolutionarily conserved." Chromosoma **112**(7): 360-371.
18. Barca-Mayo, O. and D. De Pietri Tonelli (2014). "Convergent microRNA actions coordinate neocortical development." Cell Mol Life Sci **71**(16): 2975-2995.
19. Barroso-delJesus, A., et al. (2008). "Embryonic stem cell-specific miR302-367 cluster: human gene structure and functional characterization of its core promoter." Mol Cell Biol **28**(21): 6609-6619.
20. Bartel, D. P. (2004). "MicroRNAs: genomics, biogenesis, mechanism, and function." Cell **116**(2): 281-297.
21. Bartova, E., et al. (2008). "Histone modifications and nuclear architecture: a review." J Histochem Cytochem **56**(8): 711-721.
22. Baubec, T., et al. (2013). "Methylation-dependent and -independent genomic targeting principles of the MBD protein family." Cell **153**(2): 480-492.
23. Baudoin, J. P., et al. (2012). "Tangentially migrating neurons assemble a primary cilium that promotes their reorientation to the cortical plate." Neuron **76**(6): 1108-1122.
24. Bayne, E. H., et al. (2010). "Stc1: a critical link between RNAi and chromatin modification required for heterochromatin integrity." Cell **140**(5): 666-677.
25. Behm-Ansmant, I., et al. (2006). "mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes." Genes Dev **20**(14): 1885-1898.

26. Berardi, E., et al. (2012). "miRNAs in ESC differentiation." Am J Physiol Heart Circ Physiol **303**(8): H931-939.
27. Bergles, D. E. and W. D. Richardson (2015). "Oligodendrocyte Development and Plasticity." Cold Spring Harb Perspect Biol **8**(2): a020453.
28. Bernstein, E., et al. (2003). "Dicer is essential for mouse development." Nat Genet **35**(3): 215-217.
29. Bhaumik, D., et al. (2008). "Expression of microRNA-146 suppresses NF-kappaB activity with reduction of metastatic potential in breast cancer cells." Oncogene **27**(42): 5643-5647.
30. Boland, M. J., et al. (2014). "Epigenetic regulation of pluripotency and differentiation." Circ Res **115**(2): 311-324.
31. Bolger, A. M., et al. (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data." Bioinformatics **30**(15): 2114-2120.
32. Bolzer, A., et al. (2005). "Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes." PLoS Biol **3**(5): e157.
33. Bommer, G. T., et al. (2007). "p53-mediated activation of miRNA34 candidate tumor-suppressor genes." Curr Biol **17**(15): 1298-1307.
34. Bond, A. M., et al. (2012). "The dynamic role of bone morphogenetic proteins in neural stem cell fate and maturation." Dev Neurobiol **72**(7): 1068-1084.
35. Bonev, B., et al. (2017). "Multiscale 3D Genome Rewiring during Mouse Neural Development." Cell **171**(3): 557-572 e524.
36. Bongso, A., et al. (1994). "Isolation and culture of inner cell mass cells from human blastocysts." Hum Reprod **9**(11): 2110-2117.
37. Bonn, S., et al. (2012). "Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development." Nat Genet **44**(2): 148-156.
38. Boroviak, T. and J. Nichols (2017). "Primate embryogenesis predicts the hallmarks of human naive pluripotency." Development **144**(2): 175-186.
39. Bouschet, T., et al. (2017). "In Vitro Corticogenesis from Embryonic Stem Cells Recapitulates the In Vivo Epigenetic Control of Imprinted Gene Expression." Cereb Cortex **27**(3): 2418-2433.
40. Boutin, C., et al. (2010). "NeuroD1 induces terminal neuronal differentiation in olfactory neurogenesis." Proc Natl Acad Sci U S A **107**(3): 1201-1206.
41. Braganca, J., et al. (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**(10): 8559-8565.
42. Branco, M. R. and A. Pombo (2006). "Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations." PLoS Biol **4**(5): e138.
43. Braun, J. E., et al. (2012). "A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5' exonucleolytic degradation." Nat Struct Mol Biol **19**(12): 1324-1331.
44. Brisken, C., et al. (1999). "Prolactin controls mammary gland development via direct and indirect mechanisms." Dev Biol **210**(1): 96-106.
45. Brown, C. J., et al. (1991). "A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome." Nature **349**(6304): 38-44.
46. Bruel-Jungerman, E., et al. (2011). "Cholinergic influences on cortical development and adult neurogenesis." Behav Brain Res **221**(2): 379-388.
47. Bulut-Karslioglu, A., et al. (2012). "A transcription factor-based mechanism for mouse heterochromatin formation." Nat Struct Mol Biol **19**(10): 1023-1030.
48. Buse, P., et al. (1995). "Glucocorticoid-induced functional polarity of growth factor responsiveness regulates tight junction dynamics in transformed mammary epithelial tumor cells." J Biol Chem **270**(47): 28223-28227.
49. Cadigan, K. M. (2012). "TCFs and Wnt/beta-catenin signaling: more than one way to throw the switch." Curr Top Dev Biol **98**: 1-34.
50. Cadigan, K. M. and M. L. Waterman (2012). "TCF/LEFs and Wnt signaling in the nucleus." Cold Spring Harb Perspect Biol **4**(11).

51. Cahoy, J. D., et al. (2008). "A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function." J Neurosci **28**(1): 264-278.
52. Cambray, S., et al. (2012). "Activin induces cortical interneuron identity and differentiation in embryonic stem cell-derived telencephalic neural precursors." Nat Commun **3**: 841.
53. Card, D. A., et al. (2008). "Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells." Mol Cell Biol **28**(20): 6426-6438.
54. Cardenas, A. and V. Borrell (2020). "Molecular and cellular evolution of corticogenesis in amniotes." Cell Mol Life Sci **77**(8): 1435-1460.
55. Cardenas, A., et al. (2018). "Evolution of Cortical Neurogenesis in Amniotes Controlled by Robo Signaling Levels." Cell **174**(3): 590-606 e521.
56. Caronia-Brown, G., et al. (2014). "The cortical hem regulates the size and patterning of neocortex." Development **141**(14): 2855-2865.
57. Cecchi, C. (2002). "Emx2: a gene responsible for cortical development, regionalization and area specification." Gene **291**(1-2): 1-9.
58. Chakrabarti, R., et al. (2012). "Elf5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2." Nat Cell Biol **14**(11): 1212-1222.
59. Chandradoss, K. R., et al. (2020). "CTCF-Mediated Genome Architecture Regulates the Dosage of Mitotically Stable Mono-allelic Expression of Autosomal Genes." Cell Rep **33**(4): 108302.
60. Chandradoss, K. R., et al. (2020). "Biased visibility in Hi-C datasets marks dynamically regulated condensed and decondensed chromatin states genome-wide." BMC Genomics **21**(1): 175.
61. Chang, T. C., et al. (2008). "Widespread microRNA repression by Myc contributes to tumorigenesis." Nat Genet **40**(1): 43-50.
62. Changolkar, L. N. and J. R. Pehrson (2002). "Reconstitution of nucleosomes with histone macroH2A1.2." Biochemistry **41**(1): 179-184.
63. Chapman, R. S., et al. (1999). "Suppression of epithelial apoptosis and delayed mammary gland involution in mice with a conditional knockout of Stat3." Genes Dev **13**(19): 2604-2616.
64. Chen, E. S., et al. (2008). "Cell cycle control of centromeric repeat transcription and heterochromatin assembly." Nature **451**(7179): 734-737.
65. Chen, F., et al. (2016). "The transcription factor NeuroD2 coordinates synaptic innervation and cell intrinsic properties to control excitability of cortical pyramidal neurons." J Physiol **594**(13): 3729-3744.
66. Chen, P. B., et al. (2013). "Hdac6 regulates Tip60-p400 function in stem cells." Elife **2**: e01557.
67. Chen, Q., et al. (2014). "A temporal requirement for Hippo signaling in mammary gland differentiation, growth, and tumorigenesis." Genes Dev **28**(5): 432-437.
68. Chen, Y. T., et al. (2005). "Identification of CT46/HORMAD1, an immunogenic cancer/testis antigen encoding a putative meiosis-related protein." Cancer Immun **5**: 9.
69. Cheng, L. C., et al. (2009). "miR-124 regulates adult neurogenesis in the subventricular zone stem cell niche." Nat Neurosci **12**(4): 399-408.
70. Chi, S. W., et al. (2009). "Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps." Nature **460**(7254): 479-486.
71. Cho, K. H. T., et al. (2019). "Emerging Roles of miRNAs in Brain Development and Perinatal Brain Injury." Front Physiol **10**: 227.
72. Choi, Y. S., et al. (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**(2): 227-241.
73. Choukrallah, M. A. and P. Matthias (2014). "The Interplay between Chromatin and Transcription Factor Networks during B Cell Development: Who Pulls the Trigger First?" Front Immunol **5**: 156.

74. Christie, M., et al. (2013). "Structure of the PAN3 pseudokinase reveals the basis for interactions with the PAN2 deadenylase and the GW182 proteins." Mol Cell **51**(3): 360-373.
75. Cocolakis, E., et al. (2008). "Smad signaling antagonizes STAT5-mediated gene transcription and mammary epithelial cell differentiation." J Biol Chem **283**(3): 1293-1307.
76. Coleman, S., et al. (1988). "Ductal morphogenesis in the mouse mammary gland: evidence supporting a role for epidermal growth factor." Dev Biol **127**(2): 304-315.
77. Cremer, M., et al. (2008). "Multicolor 3D fluorescence in situ hybridization for imaging interphase chromosomes." Methods Mol Biol **463**: 205-239.
78. Cremer, T., et al. (1993). "Role of chromosome territories in the functional compartmentalization of the cell nucleus." Cold Spring Harb Symp Quant Biol **58**: 777-792.
79. Creyghton, M. P., et al. (2010). "Histone H3K27ac separates active from poised enhancers and predicts developmental state." Proc Natl Acad Sci U S A **107**(50): 21931-21936.
80. Cui, Y., et al. (2004). "Inactivation of Stat5 in mouse mammary epithelium during pregnancy reveals distinct functions in cell proliferation, survival, and differentiation." Mol Cell Biol **24**(18): 8037-8047.
81. Danjo, T., et al. (2011). "Subregional specification of embryonic stem cell-derived ventral telencephalic tissues by timed and combinatory treatment with extrinsic signals." J Neurosci **31**(5): 1919-1933.
82. Davis, K. R., et al. (2016). "XBP1 Regulates the Biosynthetic Capacity of the Mammary Gland During Lactation by Controlling Epithelial Expansion and Endoplasmic Reticulum Formation." Endocrinology **157**(1): 417-428.
83. De Juan Romero, C. and V. Borrell (2015). "Coevolution of radial glial cells and the cerebral cortex." Glia **63**(8): 1303-1319.
84. Dekker, J., et al. (2002). "Capturing chromosome conformation." Science **295**(5558): 1306-1311.
85. Denli, A. M., et al. (2004). "Processing of primary microRNAs by the Microprocessor complex." Nature **432**(7014): 231-235.
86. Deome, K. B., et al. (1959). "Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice." Cancer Res **19**(5): 515-520.
87. Di Stefano, B., et al. (2016). "C/EBPalpha creates elite cells for iPSC reprogramming by upregulating Klf4 and increasing the levels of Lsd1 and Brd4." Nat Cell Biol **18**(4): 371-381.
88. Diaz-Perez, S. V., et al. (2006). "A deletion at the mouse Xist gene exposes trans-effects that alter the heterochromatin of the inactive X chromosome and the replication time and DNA stability of both X chromosomes." Genetics **174**(3): 1115-1133.
89. Dileep, V., et al. (2015). "Topologically associating domains and their long-range contacts are established during early G1 coincident with the establishment of the replication-timing program." Genome Res **25**(8): 1104-1113.
90. Dileep, V., et al. (2019). "Rapid Irreversible Transcriptional Reprogramming in Human Stem Cells Accompanied by Discordance between Replication Timing and Chromatin Compartment." Stem Cell Reports **13**(1): 193-206.
91. Donley, N., et al. (2015). "ASAR15, A cis-acting locus that controls chromosome-wide replication timing and stability of human chromosome 15." PLoS Genet **11**(1): e1004923.
92. Dorigo, B., et al. (2003). "Chromatin fiber folding: requirement for the histone H4 N-terminal tail." J Mol Biol **327**(1): 85-96.
93. Dos Santos, C. O., et al. (2015). "An epigenetic memory of pregnancy in the mouse mammary gland." Cell Rep **11**(7): 1102-1109.
94. Du, P., et al. (2018). "An Intermediate Pluripotent State Controlled by MicroRNAs Is Required for the Naive-to-Primed Stem Cell Transition." Cell Stem Cell **22**(6): 851-864 e855.
95. Dvash, T., et al. (2006). "Human embryonic stem cells as a powerful tool for studying human embryogenesis." Pediatr Res **60**(2): 111-117.
96. Eguizabal, C., et al. (2019). "Two decades of embryonic stem cells: a historical overview." Hum Reprod Open **2019**(1): hoy024.

97. Eiraku, M., et al. (2008). "Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals." Cell Stem Cell **3**(5): 519-532.
98. Elkouris, M., et al. (2011). "Sox1 maintains the undifferentiated state of cortical neural progenitor cells via the suppression of Prox1-mediated cell cycle exit and neurogenesis." Stem Cells **29**(1): 89-98.
99. Emerman, J. T., et al. (1980). "Interrelationship of glycogen metabolism and lactose synthesis in mammary epithelial cells of mice." Biochem J **192**(2): 695-702.
100. Emery, B. and Q. R. Lu (2015). "Transcriptional and Epigenetic Regulation of Oligodendrocyte Development and Myelination in the Central Nervous System." Cold Spring Harb Perspect Biol **7**(9): a020461.
101. Faigle, R. and H. Song (2013). "Signaling mechanisms regulating adult neural stem cells and neurogenesis." Biochim Biophys Acta **1830**(2): 2435-2448.
102. Fang, J., et al. (2004). "Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation." J Biol Chem **279**(51): 52812-52815.
103. Fazzio, T. G., et al. (2008). "Chromatin regulation Tip(60)s the balance in embryonic stem cell self-renewal." Cell Cycle **7**(21): 3302-3306.
104. Fenton, S. E. and L. G. Sheffield (1993). "Prolactin inhibits epidermal growth factor (EGF)-stimulated signaling events in mouse mammary epithelial cells by altering EGF receptor function." Mol Biol Cell **4**(8): 773-780.
105. Ferrari, N., et al. (2013). "RUNX2 in mammary gland development and breast cancer." J Cell Physiol **228**(6): 1137-1142.
106. Fort, R. S., et al. (2018). "An integrated view of the role of miR-130b/301b miRNA cluster in prostate cancer." Exp Hematol Oncol **7**: 10.
107. Friedman, R. C., et al. (2009). "Most mammalian mRNAs are conserved targets of microRNAs." Genome Res **19**(1): 92-105.
108. Fullwood, M. J., et al. (2009). "An oestrogen-receptor-alpha-bound human chromatin interactome." Nature **462**(7269): 58-64.
109. Gallego, M. I., et al. (2001). "Prolactin, growth hormone, and epidermal growth factor activate Stat5 in different compartments of mammary tissue and exert different and overlapping developmental effects." Dev Biol **229**(1): 163-175.
110. Gan, Q., et al. (2014). "Pax6 mediates ss-catenin signaling for self-renewal and neurogenesis by neocortical radial glial stem cells." Stem Cells **32**(1): 45-58.
111. Gao, C., et al. (2020). "Role of microRNA-33a in malignant cells." Oncol Lett **20**(3): 2537-2556.
112. Gaspard, N., et al. (2009). "Generation of cortical neurons from mouse embryonic stem cells." Nat Protoc **4**(10): 1454-1463.
113. Gaspard, N., et al. (2008). "An intrinsic mechanism of corticogenesis from embryonic stem cells." Nature **455**(7211): 351-357.
114. Gdula, M. R., et al. (2019). "The non-canonical SMC protein SmcHD1 antagonises TAD formation and compartmentalisation on the inactive X chromosome." Nat Commun **10**(1): 30.
115. Gendler, K., et al. (2008). "ChromDB: the chromatin database." Nucleic Acids Res **36**(Database issue): D298-302.
116. Gendrel, A. V., et al. (2012). "Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome." Dev Cell **23**(2): 265-279.
117. Gerrard, L., et al. (2005). "Stably transfected human embryonic stem cell clones express OCT4-specific green fluorescent protein and maintain self-renewal and pluripotency." Stem Cells **23**(1): 124-133.
118. Ghimire, S., et al. (2018). "Comparative analysis of naive, primed and ground state pluripotency in mouse embryonic stem cells originating from the same genetic background." Sci Rep **8**(1): 5884.
119. Gilchrist, D. A., et al. (2010). "Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation." Cell **143**(4): 540-551.

120. Golonzhka, O., et al. (2015). "Pbx Regulates Patterning of the Cerebral Cortex in Progenitors and Postmitotic Neurons." Neuron **88**(6): 1192-1207.
121. Grebbin, B. M., et al. (2016). "Pbx1 is required for adult subventricular zone neurogenesis." Development **143**(13): 2281-2291.
122. Greenberg, M. V. C. and D. Bourc'his (2019). "The diverse roles of DNA methylation in mammalian development and disease." Nat Rev Mol Cell Biol **20**(10): 590-607.
123. Gregor, M. F., et al. (2013). "The role of adipocyte XBP1 in metabolic regulation during lactation." Cell Rep **3**(5): 1430-1439.
124. Greig, L. C., et al. (2013). "Molecular logic of neocortical projection neuron specification, development and diversity." Nat Rev Neurosci **14**(11): 755-769.
125. Griffiths-Jones, S., et al. (2008). "miRBase: tools for microRNA genomics." Nucleic Acids Res **36**(Database issue): D154-158.
126. Grimm, S. L. and J. M. Rosen (2003). "The role of C/EBPbeta in mammary gland development and breast cancer." J Mammary Gland Biol Neoplasia **8**(2): 191-204.
127. Guenther, M. G., et al. (2007). "A chromatin landmark and transcription initiation at most promoters in human cells." Cell **130**(1): 77-88.
128. Guo, G., et al. (2016). "Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass." Stem Cell Reports **6**(4): 437-446.
129. Guruharsha, K. G., et al. (2012). "The Notch signalling system: recent insights into the complexity of a conserved pathway." Nat Rev Genet **13**(9): 654-666.
130. Hackett, J. A. and M. A. Surani (2014). "Regulatory principles of pluripotency: from the ground state up." Cell Stem Cell **15**(4): 416-430.
131. Hailesellasse Sene, K., et al. (2007). "Gene function in early mouse embryonic stem cell differentiation." BMC Genomics **8**: 85.
132. Haslam, S. Z. (1989). "The ontogeny of mouse mammary gland responsiveness to ovarian steroid hormones." Endocrinology **125**(5): 2766-2772.
133. Haslam, S. Z. and G. Shyamala (1981). "Relative distribution of estrogen and progesterone receptors among the epithelial, adipose, and connective tissue components of the normal mammary gland." Endocrinology **108**(3): 825-830.
134. Hayashi, K., et al. (2008). "MicroRNA biogenesis is required for mouse primordial germ cell development and spermatogenesis." PLoS One **3**(3): e1738.
135. He, L., et al. (2005). "A microRNA polycistron as a potential human oncogene." Nature **435**(7043): 828-833.
136. Heard, E. (2005). "Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome." Curr Opin Genet Dev **15**(5): 482-489.
137. Helbling Chadwick, L., et al. (2009). "The Mi-2/NuRD complex associates with pericentromeric heterochromatin during S phase in rapidly proliferating lymphoid cells." Chromosoma **118**(4): 445-457.
138. Helwak, A. and D. Tollervey (2014). "Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH)." Nat Protoc **9**(3): 711-728.
139. Hens, J. R. and J. J. Wysolmerski (2005). "Key stages of mammary gland development: molecular mechanisms involved in the formation of the embryonic mammary gland." Breast Cancer Res **7**(5): 220-224.
140. Hnisz, D., et al. (2013). "Super-enhancers in the control of cell identity and disease." Cell **155**(4): 934-947.
141. Hogg, N. A., et al. (1983). "Lumen formation in the developing mouse mammary gland." J Embryol Exp Morphol **73**: 39-57.
142. Hollenhorst, P. C., et al. (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**(10): 1044-1052.

143. Houbaviy, H. B., et al. (2003). "Embryonic stem cell-specific MicroRNAs." Dev Cell **5**(2): 351-358.
144. Hovey, R. C. and L. Aimo (2010). "Diverse and active roles for adipocytes during mammary gland growth and function." J Mammary Gland Biol Neoplasia **15**(3): 279-290.
145. Hsieh, T. H., et al. (2015). "Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C." Cell **162**(1): 108-119.
146. Hu, Z., et al. (2008). "Genetic variants of miRNA sequences and non-small cell lung cancer survival." J Clin Invest **118**(7): 2600-2608.
147. Huang da, W., et al. (2009). "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists." Nucleic Acids Res **37**(1): 1-13.
148. Huang da, W., et al. (2009). "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources." Nat Protoc **4**(1): 44-57.
149. Huang, K., et al. (2014). "The naive state of human pluripotent stem cells: a synthesis of stem cell and preimplantation embryo transcriptome analyses." Cell Stem Cell **15**(4): 410-415.
150. Humphreys, R. C., et al. (2002). "Deletion of Stat3 blocks mammary gland involution and extends functional competence of the secretory epithelium in the absence of lactogenic stimuli." Endocrinology **143**(9): 3641-3650.
151. Ideker, T., et al. (2002). "Discovering regulatory and signalling circuits in molecular interaction networks." Bioinformatics **18 Suppl 1**: S233-240.
152. Iliopoulos, D., et al. (2010). "Loss of miR-200 inhibition of Suz12 leads to polycomb-mediated repression required for the formation and maintenance of cancer stem cells." Mol Cell **39**(5): 761-772.
153. Iorio, M. V. and C. M. Croce (2012). "MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review." EMBO Mol Med **4**(3): 143-159.
154. Ishov, A. M., et al. (2004). "Heterochromatin and ND10 are cell-cycle regulated and phosphorylation-dependent alternate nuclear sites of the transcription repressor Daxx and SWI/SNF protein ATRX." J Cell Sci **117**(Pt 17): 3807-3820.
155. Ivanova, E., et al. (2021). "Epigenetics: New Insights into Mammary Gland Biology." Genes (Basel) **12**(2).
156. Jaggi, R., et al. (1996). "Regulation of a physiological apoptosis: mouse mammary involution." J Dairy Sci **79**(6): 1074-1084.
157. Ji, Q., et al. (2019). "Hematopoietic PBX-interacting protein mediates cartilage degeneration during the pathogenesis of osteoarthritis." Nat Commun **10**(1): 313.
158. Jin, W., et al. (2015). "Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples." Nature **528**(7580): 142-146.
159. Jo, M. H., et al. (2015). "Human Argonaute 2 Has Diverse Reaction Pathways on Target RNAs." Mol Cell **59**(1): 117-124.
160. John, S., et al. (2008). "Interaction of the glucocorticoid receptor with the chromatin landscape." Mol Cell **29**(5): 611-624.
161. John, S., et al. (2011). "Chromatin accessibility pre-determines glucocorticoid receptor binding patterns." Nat Genet **43**(3): 264-268.
162. Jonas, S. and E. Izaurralde (2015). "Towards a molecular understanding of microRNA-mediated gene silencing." Nat Rev Genet **16**(7): 421-433.
163. Juliandi, B., et al. (2010). "Epigenetic regulation in neural stem cell differentiation." Dev Growth Differ **52**(6): 493-504.
164. Juliandi, B., et al. (2012). "Induction of superficial cortical layer neurons from mouse embryonic stem cells by valproic acid." Neurosci Res **72**(1): 23-31.
165. Kabotyanski, E. B., et al. (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**(10): 2355-2368.

166. Kageyama, R., et al. (2008). "Roles of Hes genes in neural development." Dev Growth Differ **50 Suppl 1**: S97-103.
167. Kawamata, T. and Y. Tomari (2010). "Making RISC." Trends Biochem Sci **35**(7): 368-376.
168. Kawauchi, S., et al. (2009). "The role of foxg1 in the development of neural stem cells of the olfactory epithelium." Ann N Y Acad Sci **1170**: 21-27.
169. Kawazu, M., et al. (2011). "Histone demethylase JMJD2B functions as a co-factor of estrogen receptor in breast cancer proliferation and mammary gland development." PLoS One **6**(3): e17830
170. Khaled, W. T., et al. (2015). "BCL11A is a triple-negative breast cancer gene with critical functions in stem and progenitor cells." Nat Commun **6**: 5987.
171. Kieffer-Kwon, K. R., et al. (2017). "Myc Regulates Chromatin Decompaction and Nuclear Architecture during B Cell Activation." Mol Cell **67**(4): 566-578 e510.
172. Kim, D., et al. (2013). "TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions." Genome Biol **14**(4): R36.
173. Kind, J., et al. (2013). "Single-cell dynamics of genome-nuclear lamina interactions." Cell **153**(1): 178-192. Kirkeby, A., et al. (2012). "Generation of regionally specified neural progenitors and functional neurons from human embryonic stem cells under defined conditions." Cell Rep **1**(6): 703-714.
174. Klemm, S. L., et al. (2019). "Chromatin accessibility and the regulatory epigenome." Nat Rev Genet **20**(4): 207-220.
175. Klymenko, T. and J. Muller (2004). "The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins." EMBO Rep **5**(4): 373-377.
176. Kornberg, R. D. and J. O. Thomas (1974). "Chromatin structure; oligomers of the histones." Science **184**(4139): 865-868.
177. Kouros-Mehr, H., et al. (2006). "GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland." Cell **127**(5): 1041-1055.
178. Kouzarides, T. (2007). "Chromatin modifications and their function." Cell **128**(4): 693-705.
179. Krebs, A. R., et al. (2017). "Genome-wide Single-Molecule Footprinting Reveals High RNA Polymerase II Turnover at Paused Promoters." Mol Cell **67**(3): 411-422 e414.
180. Krichevsky, A. M., et al. (2006). "Specific microRNAs modulate embryonic stem cell-derived neurogenesis." Stem Cells **24**(4): 857-864.
181. Kroon, E., et al. (2008). "Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo." Nat Biotechnol **26**(4): 443-452.
182. Krzyzanowski, P. M. and M. A. Andrade-Navarro (2007). "Identification of novel stem cell markers using gap analysis of gene expression data." Genome Biol **8**(9): R193.
183. Kuang, Y., et al. (2012). "Dicer1 and MiR-9 are required for proper Notch1 signaling and the Bergmann glial phenotype in the developing mouse cerebellum." Glia **60**(11): 1734-1746.
184. Kurpios, N. A., et al. (2009). "The Pea3 Ets transcription factor regulates differentiation of multipotent progenitor cells during mammary gland development." Dev Biol **325**(1): 106-121.
185. Lachner, M., et al. (2001). "Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins." Nature **410**(6824): 116-120.
186. Lagos-Quintana, M., et al. (2001). "Identification of novel genes coding for small expressed RNAs." Science **294**(5543): 853-858.
187. Lakshmipathy, U., et al. (2010). "miRNA in pluripotent stem cells." Regen Med **5**(4): 545-555.
188. Lall, S., et al. (2006). "A genome-wide map of conserved microRNA targets in C. elegans." Curr Biol **16**(5): 460-471.
189. LaMarca, H. L., et al. (2010). "CCAAT/enhancer binding protein beta regulates stem cell activity and specifies luminal cell fate in the mammary gland." Stem Cells **28**(3): 535-544.
190. Landgraf, P., et al. (2007). "A mammalian microRNA expression atlas based on small RNA library sequencing." Cell **129**(7): 1401-1414.
191. Laneve, P., et al. (2010). "A minicircuitry involving REST and CREB controls miR-9-2 expression during human neuronal differentiation." Nucleic Acids Res **38**(20): 6895-6905.

192. Lange, U. C., et al. (2013). "Dissecting the role of H3K64me3 in mouse pericentromeric heterochromatin." Nat Commun **4**: 2233.
193. Langer-Safer, P. R., et al. (1982). "Immunological method for mapping genes on Drosophila polytene chromosomes." Proc Natl Acad Sci U S A **79**(14): 4381-4385.
194. Lau, N. C., et al. (2001). "An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans." Science **294**(5543): 858-862.
195. Lawrence, M., et al. (2016). "Lateral Thinking: How Histone Modifications Regulate Gene Expression." Trends Genet **32**(1): 42-56.
196. Lee, R. C. and V. Ambros (2001). "An extensive class of small RNAs in Caenorhabditis elegans." Science **294**(5543): 862-864.
197. Lee, R. C., et al. (1993). "The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14." Cell **75**(5): 843-854.
198. Lee, Y., et al. (2004). "MicroRNA genes are transcribed by RNA polymerase II." EMBO J **23**(20): 4051-4060.
199. Lehtinen, M. K., et al. (2011). "The cerebrospinal fluid provides a proliferative niche for neural progenitor cells." Neuron **69**(5): 893-905.
200. Lewis, B. P., et al. (2005). "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets." Cell **120**(1): 15-20.
201. Lewis, B. P., et al. (2003). "Prediction of mammalian microRNA targets." Cell **115**(7): 787-798.
202. Lewis, P. W., et al. (2010). "Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres." Proc Natl Acad Sci U S A **107**(32): 14075-14080.
203. Li, X., et al. (2008). "A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints." Dev Cell **15**(4): 547-557.
204. Liang, X., et al. (2011). "Isl1 is required for multiple aspects of motor neuron development." Mol Cell Neurosci **47**(3): 215-222.
205. Liao, J., et al. (2015). "Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells." Nat Genet **47**(5): 469-478.
206. Liberzon, A., et al. (2011). "Molecular signatures database (MSigDB) 3.0." Bioinformatics **27**(12): 1739-1740.
207. Lieberman-Aiden, E., et al. (2009). "Comprehensive mapping of long-range interactions reveals folding principles of the human genome." Science **326**(5950): 289-293.
208. Liu, Y., et al. (2014). "RNA-seq differential expression studies: more sequence or more replication?" Bioinformatics **30**(3): 301-304.
209. Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." Methods **25**(4): 402-408.
210. Love, M. I., et al. (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome Biol **15**(12): 550.
211. Lu, J., et al. (2010). "G2 phase chromatin lacks determinants of replication timing." J Cell Biol **189**(6): 967-980.
212. Lupien, M., et al. (2008). "FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription." Cell **132**(6): 958-970.
213. Lydon, J. P., et al. (1995). "Mice lacking progesterone receptor exhibit pleiotropic reproductive abnormalities." Genes Dev **9**(18): 2266-2278.
214. Lyons, W. R. (1958). "Hormonal synergism in mammary growth." Proc R Soc Lond B Biol Sci **149**(936): 303-325.
215. Macias, H. and L. Hinck (2012). "Mammary gland development." Wiley Interdiscip Rev Dev Biol **1**(4): 533-557.
216. Malhotra, G. K., et al. (2014). "The role of Sox9 in mouse mammary gland development and maintenance of mammary stem and luminal progenitor cells." BMC Dev Biol **14**: 47.

217. Mallepell, S., et al. (2006). "Paracrine signaling through the epithelial estrogen receptor alpha is required for proliferation and morphogenesis in the mammary gland." Proc Natl Acad Sci U S A **103**(7): 2196-2201.
218. Manavathi, B., et al. (2012). "Functional regulation of pre-B-cell leukemia homeobox interacting protein 1 (PBXIP1/HPIP) in erythroid differentiation." J Biol Chem **287**(8): 5600-5614.
219. Marasca, F., et al. (2018). "How Polycomb-Mediated Cell Memory Deals With a Changing Environment: Variations in PcG complexes and proteins assortment convey plasticity to epigenetic regulation as a response to environment." Bioessays **40**(4): e1700137.
220. Marchal, C., et al. (2019). "Control of DNA replication timing in the 3D genome." Nat Rev Mol Cell Biol **20**(12): 721-737.
221. Marson, A., et al. (2008). "Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells." Cell **134**(3): 521-533.
222. Maruyama, R., et al. (2011). "Epigenetic regulation of cell type-specific expression patterns in the human mammary epithelium." PLoS Genet **7**(4): e1001369.
223. Masserdotti, G., et al. (2015). "Transcriptional Mechanisms of Proneural Factors and REST in Regulating Neuronal Reprogramming of Astrocytes." Cell Stem Cell **17**(1): 74-88.
224. Matsuzaki, F. and A. Shitamukai (2015). "Cell Division Modes and Cleavage Planes of Neural Progenitors during Mammalian Cortical Development." Cold Spring Harb Perspect Biol **7**(9): a015719.
225. McDowell, T. L., et al. (1999). "Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes." Proc Natl Acad Sci U S A **96**(24): 13983-13988.
226. Mei, L. and K. A. Nave (2014). "Neuregulin-ERBB signaling in the nervous system and neuropsychiatric diseases." Neuron **83**(1): 27-49.
227. Meijer, H. A., et al. (2013). "Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation." Science **340**(6128): 82-85.
228. Menendez, P., et al. (2013). "Diagnostic and prognostic significance of serum microRNAs in colorectal cancer." J Surg Oncol **107**(2): 217-220.
229. Messmer, T., et al. (2019). "Transcriptional Heterogeneity in Naive and Primed Human Pluripotent Stem Cells at Single-Cell Resolution." Cell Rep **26**(4): 815-824 e814.
230. Meza-Sosa, K. F., et al. (2012). "Role of microRNAs in central nervous system development and pathology." J Neurosci Res **90**(1): 1-12.
231. Miao, R. Y., et al. (2011). "MYB is essential for mammary tumorigenesis." Cancer Res **71**(22): 7029-7037.
232. Michalak, E. M., et al. (2018). "Canonical PRC2 function is essential for mammary gland development and affects chromatin compaction in mammary organoids." PLoS Biol **16**(8): e2004986.
233. Michelsen, K. A., et al. (2015). "Area-specific reestablishment of damaged circuits in the adult cerebral cortex by cortical neurons derived from mouse embryonic stem cells." Neuron **85**(5): 982-997.
234. Moradi, S., et al. (2017). "Small RNA Sequencing Reveals Dlk1-Dio3 Locus-Embedded MicroRNAs as Major Drivers of Ground-State Pluripotency." Stem Cell Reports **9**(6): 2081-2096.
235. Morrison, B. and M. L. Cutler (2009). "Mouse Mammary Epithelial Cells form Mammospheres During Lactogenic Differentiation." J Vis Exp(32).
236. Muratore, C. R., et al. (2014). "Comparison and optimization of hiPSC forebrain cortical differentiation protocols." PLoS One **9**(8): e105807.
237. Murchison, E. P., et al. (2005). "Characterization of Dicer-deficient murine embryonic stem cells." Proc Natl Acad Sci U S A **102**(34): 12135-12140.
238. Murtagh, J., et al. (2004). "Organization of mammary epithelial cells into 3D acinar structures requires glucocorticoid and JNK signaling." J Cell Biol **166**(1): 133-143.

239. Nagano, T., et al. (2013). "Single-cell Hi-C reveals cell-to-cell variability in chromosome structure." Nature **502**(7469): 59-64.
240. Nagano, T., et al. (2017). "Cell-cycle dynamics of chromosomal organization at single-cell resolution." Nature **547**(7661): 61-67.
241. Nalabothula, N., et al. (2014). "The chromatin architectural proteins HMGD1 and H1 bind reciprocally and have opposite effects on chromatin structure and gene regulation." BMC Genomics **15**: 92.
242. Nandi, S. (1958). "Endocrine control of mammarygland development and function in the C3H/ He Crgl mouse." J Natl Cancer Inst **21**(6): 1039-1063.
243. Nichols, J. and A. Smith (2009). "Naive and primed pluripotent states." Cell Stem Cell **4**(6): 487-492.
244. Niewiadomski, P., et al. (2014). "Gli protein activity is controlled by multisite phosphorylation in vertebrate Hedgehog signaling." Cell Rep **6**(1): 168-181.
245. Nigro, A., et al. (2012). "MiR-30e and miR-181d control radial glia cell proliferation via HtrA1 modulation." Cell Death Dis **3**: e360.
246. Nishioka, K., et al. (2002). "PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin." Mol Cell **9**(6): 1201-1213.
247. Noctor, S. C., et al. (2001). "Neurons derived from radial glial cells establish radial units in neocortex." Nature **409**(6821): 714-720.
248. Nowak, J. S., et al. (2017). "Lin28a uses distinct mechanisms of binding to RNA and affects miRNA levels positively and negatively." RNA **23**(3): 317-332.
249. O'Brien, J., et al. (2018). "Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation." Front Endocrinol (Lausanne) **9**: 402.
250. Oakes, S. R., et al. (2008). "The Ets transcription factor Elf5 specifies mammary alveolar cell fate." Genes Dev **22**(5): 581-586.
251. Okada, C., et al. (2009). "A high-resolution structure of the pre-microRNA nuclear export machinery." Science **326**(5957): 1275-1279.
252. Okano, M., et al. (1999). "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." Cell **99**(3): 247-257.
253. Olena, A. F. and J. G. Patton (2010). "Genomic organization of microRNAs." J Cell Physiol **222**(3): 540-545.
254. Olins, D. E. and A. L. Olins (2003). "Chromatin history: our view from the bridge." Nat Rev Mol Cell Biol **4**(10): 809-814.
255. Ooga, M., et al. (2018). "Chd9 mediates highly loosened chromatin structure in growing mouse oocytes." Biochem Biophys Res Commun **500**(3): 583-588.
256. Ooi, S. K., et al. (2007). "DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA." Nature **448**(7154): 714-717.
257. Ormandy, C. J., et al. (2003). "Investigation of the transcriptional changes underlying functional defects in the mammary glands of prolactin receptor knockout mice." Recent Prog Horm Res **58**: 297-323.
258. Ortega, F., et al. (2013). "Oligodendrogliogenic and neurogenic adult subependymal zone neural stem cells constitute distinct lineages and exhibit differential responsiveness to Wnt signalling." Nat Cell Biol **15**(6): 602-613.
259. Pal, B., et al. (2013). "Global changes in the mammary epigenome are induced by hormonal cues and coordinated by Ezh2." Cell Rep **3**(2): 411-426.
260. Panchenko, M. V., et al. (2004). "von Hippel-Lindau partner Jade-1 is a transcriptional co-activator associated with histone acetyltransferase activity." J Biol Chem **279**(53): 56032-56041.
261. Pathania, R., et al. (2015). "DNMT1 is essential for mammary and cancer stem cell maintenance and tumorigenesis." Nat Commun **6**: 6910.

262. Pellegrini, S., et al. (2020). "Transcriptional dynamics of induced pluripotent stem cell differentiation into beta cells reveals full endodermal commitment and homology with human islets." Cytotherapy.
263. Perotti, C., et al. (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**(5): 269-282.
264. Peters, A. H., et al. (2003). "Partitioning and plasticity of repressive histone methylation states in mammalian chromatin." Mol Cell **12**(6): 1577-1589.
265. Petryniak, M. A., et al. (2007). "Dlx1 and Dlx2 control neuronal versus oligodendroglial cell fate acquisition in the developing forebrain." Neuron **55**(3): 417-433.
266. Phillips-Cremins, J. E., et al. (2013). "Architectural protein subclasses shape 3D organization of genomes during lineage commitment." Cell **153**(6): 1281-1295.
267. Phua, Y. W., et al. (2015). "MicroRNA profiling of the pubertal mouse mammary gland identifies miR-184 as a candidate breast tumour suppressor gene." Breast Cancer Res **17**: 83.
268. Pico, A. R., et al. (2008). "WikiPathways: pathway editing for the people." PLoS Biol **6**(7): e184.
269. Pinheiro, I., et al. (2012). "Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity." Cell **150**(5): 948-960.
270. Propper, A. Y. (1978). "Wandering epithelial cells in the rabbit embryo milk line. A preliminary scanning electron microscope study." Dev Biol **67**(1): 225-231.
271. Proudhon, C., et al. (2012). "Protection against de novo methylation is instrumental in maintaining parent-of-origin methylation inherited from the gametes." Mol Cell **47**(6): 909-920.
272. Rahl, P. B., et al. (2010). "c-Myc regulates transcriptional pause release." Cell **141**(3): 432-445.
273. Rajman, M. and G. Schratt (2017). "MicroRNAs in neural development: from master regulators to fine-tuners." Development **144**(13): 2310-2322.
274. Rao, S. S., et al. (2014). "A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping." Cell **159**(7): 1665-1680.
275. Rao, S. S. P., et al. (2017). "Cohesin Loss Eliminates All Loop Domains." Cell **171**(2): 305-320 e324.
276. Rea, S., et al. (2000). "Regulation of chromatin structure by site-specific histone H3 methyltransferases." Nature **406**(6796): 593-599.
277. Regan, J. L., et al. (2013). "Aurora A kinase regulates mammary epithelial cell fate by determining mitotic spindle orientation in a Notch-dependent manner." Cell Rep **4**(1): 110-123.
278. Robinson, G. W. (2007). "Cooperation of signalling pathways in embryonic mammary gland development." Nat Rev Genet **8**(12): 963-972.
279. Rohatgi, R., et al. (2007). "Patched1 regulates hedgehog signaling at the primary cilium." Science **317**(5836): 372-376.
280. Roschger, C. and C. Cabrele (2017). "The Id-protein family in developmental and cancer-associated pathways." Cell Commun Signal **15**(1): 7.
281. Roush, S. and F. J. Slack (2008). "The let-7 family of microRNAs." Trends Cell Biol **18**(10): 505-516.
282. Ruby, J. G., et al. (2007). "Intronic microRNA precursors that bypass Drosha processing." Nature **448**(7149): 83-86.
283. Rudolph, M. C., et al. (2010). "Sterol regulatory element binding protein and dietary lipid regulation of fatty acid synthesis in the mammary epithelium." Am J Physiol Endocrinol Metab **299**(6): E918-927.
284. Ryba, T., et al. (2010). "Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types." Genome Res **20**(6): 761-770.
285. Sakakura, T. (1991). "New aspects of stroma-parenchyma relations in mammary gland differentiation." Int Rev Cytol **125**: 165-202.
286. Saksouk, N., et al. (2015). "Constitutive heterochromatin formation and transcription in mammals." Epigenetics Chromatin **8**: 3.

287. Salasznyk, R. M., et al. (2005). "Focusing of gene expression as the basis of stem cell differentiation." Stem Cells Dev **14**(6): 608-620.
288. Schotta, G., et al. (2004). "A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin." Genes Dev **18**(11): 1251-1262.
289. Shackleton, M., et al. (2006). "Generation of a functional mammary gland from a single stem cell." Nature **439**(7072): 84-88.
290. Shaker, T., et al. (2012). "Neurog1 and Neurog2 coordinately regulate development of the olfactory system." Neural Dev **7**: 28.
291. Sharmin, M., et al. (2016). "Heterogeneity of transcription factor binding specificity models within and across cell lines." Genome Res **26**(8): 1110-1123.
292. Sharp, P. A. (1994). "Split genes and RNA splicing." Cell **77**(6): 805-815.
293. Shioda, T., et al. (1997). "MSG1 and its related protein MRG1 share a transcription activating domain." Gene **204**(1-2): 235-241.
294. Shipra, A., et al. (2006). "CREMOFAC--a database of chromatin remodeling factors." Bioinformatics **22**(23): 2940-2944.
295. Silberstein, G. B. and C. W. Daniel (1987). "Reversible inhibition of mammary gland growth by transforming growth factor-beta." Science **237**(4812): 291-293.
296. Simonis, M., et al. (2009). "High-resolution identification of balanced and complex chromosomal rearrangements by 4C technology." Nat Methods **6**(11): 837-842.
297. Solovei, I., et al. (2009). "Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution." Cell **137**(2): 356-368.
298. Sornapudi, T. R., et al. (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**(1): 241.
299. Sornapudi, T. R., et al. (2018). "Comprehensive profiling of transcriptional networks specific for lactogenic differentiation of HC11 mammary epithelial stem-like cells." Sci Rep **8**(1): 11777.
300. Spitz, F. and E. E. Furlong (2012). "Transcription factors: from enhancer binding to developmental control." Nat Rev Genet **13**(9): 613-626.
301. Sternlicht, M. D. (2006). "Key stages in mammary gland development: the cues that regulate ductal branching morphogenesis." Breast Cancer Res **8**(1): 201.
302. Stingl, J., et al. (2006). "Purification and unique properties of mammary epithelial stem cells." Nature **439**(7079): 993-997.
303. Stoeckius, M., et al. (2014). "Paternal RNA contributions in the Caenorhabditis elegans zygote." EMBO J **33**(16): 1740-1750.
304. Stoelzle, T., et al. (2009). "c-Myc affects mRNA translation, cell proliferation and progenitor cell function in the mammary gland." BMC Biol **7**: 63.
305. Sugiyama, T., et al. (2005). "RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production." Proc Natl Acad Sci U S A **102**(1): 152-157.
306. Suh, M. R., et al. (2004). "Human embryonic stem cells express a unique set of microRNAs." Dev Biol **270**(2): 488-498.
307. Swinstead, E. E., et al. (2016). "Pioneer factors and ATP-dependent chromatin remodeling factors interact dynamically: A new perspective: Multiple transcription factors can effect chromatin pioneer functions through dynamic interactions with ATP-dependent chromatin remodeling factors." Bioessays **38**(11): 1150-1157.
308. Szegezdi, E., et al. (2006). "Mediators of endoplasmic reticulum stress-induced apoptosis." EMBO Rep **7**(9): 880-885.
309. Takahashi, N., et al. (2019). "ZNF445 is a primary regulator of genomic imprinting." Genes Dev **33**(1-2): 49-54.

310. Takahashi, S., et al. (2018). "Epigenetic differences between naive and primed pluripotent stem cells." Cell Mol Life Sci **75**(7): 1191-1203. Tao, Y. and S. C. Zhang (2016). "Neural Subtype Specification from Human Pluripotent Stem Cells." Cell Stem Cell **19**(5): 573-586.
311. Thiaville, M. M., et al. (2012). "Identification of PBX1 target genes in cancer cells by global mapping of PBX1 binding sites." PLoS One **7**(5): e36054.
312. Thomson, J. A., et al. (1998). "Embryonic stem cell lines derived from human blastocysts." Science **282**(5391): 1145-1147.
313. Thomson, J. A., et al. (1995). "Isolation of a primate embryonic stem cell line." Proc Natl Acad Sci U S A **92**(17): 7844-7848.
314. Thurman, R. E., et al. (2012). "The accessible chromatin landscape of the human genome." Nature **489**(7414): 75-82.
315. Trapnell, C., et al. (2012). "Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks." Nat Protoc **7**(3): 562-578.
316. Trojer, P. and D. Reinberg (2007). "Facultative heterochromatin: is there a distinctive molecular signature?" Mol Cell **28**(1): 1-13.
317. Tsai, N. C., et al. (2019). "Developmental potential of surplus morulas with delayed and/or incomplete compaction after freezing-thawing procedures." Reprod Biol Endocrinol **17**(1): 87.
318. Tyler, W. A. and T. F. Haydar (2013). "Multiplex genetic fate mapping reveals a novel route of neocortical neurogenesis, which is altered in the Ts65Dn mouse model of Down syndrome." J Neurosci **33**(12): 5106-5119.
319. Tyler, W. A., et al. (2015). "Neural precursor lineages specify distinct neocortical pyramidal neuron types." J Neurosci **35**(15): 6142-6152.
320. Uittenbogaard, M., et al. (2010). "NeuroD6 genomic signature bridging neuronal differentiation to survival via the molecular chaperone network." J Neurosci Res **88**(1): 33-54.
321. van den Ameele, J., et al. (2014). "Thinking out of the dish: what to learn about cortical development using pluripotent stem cells." Trends Neurosci **37**(6): 334-342.
322. Van Keymeulen, A., et al. (2011). "Distinct stem cells contribute to mammary gland development and maintenance." Nature **479**(7372): 189-193.
323. Varrault, A., et al. (2019). "Cerebral Cortex Generated from Pluripotent Stem Cells to Model Corticogenesis and Rebuild Cortical Circuits: In Vitro Veritas?" Stem Cells Dev **28**(6): 361-369.
324. Vaught, D., et al. (2009). "Regulation of mammary gland branching morphogenesis by EphA2 receptor tyrosine kinase." Mol Biol Cell **20**(10): 2572-2581.
325. Veltmaat, J. M., et al. (2003). "Mouse embryonic mammogenesis as a model for the molecular regulation of pattern formation." Differentiation **71**(1): 1-17.
326. Vicent, G. P., et al. (2009). "Two chromatin remodeling activities cooperate during activation of hormone responsive promoters." PLoS Genet **5**(7): e1000567.
327. Visvader, J. E. and J. Stingl (2014). "Mammary stem cells and the differentiation hierarchy: current status and perspectives." Genes Dev **28**(11): 1143-1158.
328. Wang, G. G., et al. (2007). "Chromatin remodeling and cancer, Part II: ATP-dependent chromatin remodeling." Trends Mol Med **13**(9): 373-380.
329. Wang, H., et al. (2011). "Gli3 is required for maintenance and fate specification of cortical progenitors." J Neurosci **31**(17): 6440-6448.
330. Wang, H., et al. (2016). "MicroRNA-26a/b and their host genes synergistically regulate triacylglycerol synthesis by targeting the INSIG1 gene." RNA Biol **13**(5): 500-510.
331. Wang, H., et al. (2004). "Role of histone H2A ubiquitination in Polycomb silencing." Nature **431**(7010): 873-878.
332. Wang, J., et al. (2019). "A miR-206 regulated gene landscape enhances mammary epithelial differentiation." J Cell Physiol **234**(12): 22220-22233.
333. Wang, W., et al. (2009). "Global expression profiling reveals regulation of CTGF/CCN2 during lactogenic differentiation." J Cell Commun Signal **3**(1): 43-55.

334. Wang, Y., et al. (2007). "DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal." Nat Genet **39**(3): 380-385.
335. Watanabe, K., et al. (2014). "Mammary morphogenesis and regeneration require the inhibition of EMT at terminal end buds by Ovol2 transcriptional repressor." Dev Cell **29**(1): 59-74.
336. Watson, C. J. and W. T. Khaled (2008). "Mammary development in the embryo and adult: a journey of morphogenesis and commitment." Development **135**(6): 995-1003.
337. Wei, Z., et al. (2013). "Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency." Cell Stem Cell **13**(1): 36-47.
338. Westendorp, B., et al. (2012). "E2F7 represses a network of oscillating cell cycle genes to control S-phase progression." Nucleic Acids Res **40**(8): 3511-3523.
339. Wigger, M., et al. (2017). "Plasticity of the inner cell mass in mouse blastocyst is restricted by the activity of FGF/MAPK pathway." Sci Rep **7**(1): 15136.
340. Wightman, B., et al. (1993). "Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans." Cell **75**(5): 855-862.
341. Williams, C., et al. (2009). "Gene expression in murine mammary epithelial stem cell-like cells shows similarities to human breast cancer gene expression." Breast Cancer Res **11**(3): R26.
342. Woo, P. L., et al. (1999). "Requirement for Ras and phosphatidylinositol 3-kinase signaling uncouples the glucocorticoid-induced junctional organization and transepithelial electrical resistance in mammary tumor cells." J Biol Chem **274**(46): 32818-32828.
343. Wu, C., et al. (1979). "The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activity." Cell **16**(4): 807-814.
344. Wu, J. and J. C. Izpisua Belmonte (2015). "Dynamic Pluripotent Stem Cell States and Their Applications." Cell Stem Cell **17**(5): 509-525.
345. Wyatt, L., et al. (2007). "The protein tyrosine phosphatase Pez regulates TGFbeta, epithelial-mesenchymal transition, and organ development." J Cell Biol **178**(7): 1223-1235.
346. Xie, M., et al. (2013). "Mammalian 5'-capped microRNA precursors that generate a single microRNA." Cell **155**(7): 1568-1580.
347. Xu, Q., et al. (2010). "Sonic hedgehog signaling confers ventral telencephalic progenitors with distinct cortical interneuron fates." Neuron **65**(3): 328-340.
348. Xu, Y., et al. (2020). "GABAergic Inhibitory Interneuron Deficits in Alzheimer's Disease: Implications for Treatment." Front Neurosci **14**: 660.
349. Yahata, T., et al. (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**(12): 8825-8834.
350. Yahata, T., et al. (2002). "Cloning of mouse Cited4, a member of the CITED family p300/CBP-binding transcriptional coactivators: induced expression in mammary epithelial cells." Genomics **80**(6): 601-613.
351. Yamaji, D., et al. (2009). "Development of mammary luminal progenitor cells is controlled by the transcription factor STAT5A." Genes Dev **23**(20): 2382-2387.
352. Yang, J. S., et al. (2010). "Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis." Proc Natl Acad Sci U S A **107**(34): 15163-15168.
353. Yao, M. J., et al. (2012). "Transcriptome analysis of microRNAs in developing cerebral cortex of rat." BMC Genomics **13**: 232.
354. Yori, J. L., et al. (2010). "Kruppel-like factor 4 inhibits epithelial-to-mesenchymal transition through regulation of E-cadherin gene expression." J Biol Chem **285**(22): 16854-16863.
355. Yu, M. and B. Ren (2017). "The Three-Dimensional Organization of Mammalian Genomes." Annu Rev Cell Dev Biol **33**: 265-289.
356. Yuan, K., et al. (2017). "The miR-290-295 cluster as multi-faceted players in mouse embryonic stem cells." Cell Biosci **7**: 38.
357. Zakrzewski, W., et al. (2019). "Stem cells: past, present, and future." Stem Cell Res Ther **10**(1): 68.

358. Zampetaki, A., et al. (2010). "Plasma microRNA profiling reveals loss of endothelial miR-126 and other microRNAs in type 2 diabetes." Circ Res **107**(6): 810-817.
359. Zhang, J., et al. (2014). "Dax1 and Nanog act in parallel to stabilize mouse embryonic stem cells and induced pluripotency." Nat Commun **5**: 5042.
360. Zhang, K., et al. (2008). "Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin." Nat Struct Mol Biol **15**(4): 381-388.
361. Zhang, K., et al. (2002). "Histone acetylation and deacetylation: identification of acetylation and methylation sites of HeLa histone H4 by mass spectrometry." Mol Cell Proteomics **1**(7): 500-508.
362. Zhang, M., et al. (2018). "Highly efficient methods to obtain homogeneous dorsal neural progenitor cells from human and mouse embryonic stem cells and induced pluripotent stem cells." Stem Cell Res Ther **9**(1): 67.
363. Zhang, S. and W. Cui (2014). "Sox2, a key factor in the regulation of pluripotency and neural differentiation." World J Stem Cells **6**(3): 305-311.
364. Zhao, C., et al. (2010). "MicroRNA let-7b regulates neural stem cell proliferation and differentiation by targeting nuclear receptor TLX signaling." Proc Natl Acad Sci U S A **107**(5): 1876-1881.
365. Zhao, C., et al. (2009). "A feedback regulatory loop involving microRNA-9 and nuclear receptor TLX in neural stem cell fate determination." Nat Struct Mol Biol **16**(4): 365-371.
366. Zhao, Z. and A. Shilatifard (2019). "Epigenetic modifications of histones in cancer." Genome Biol **20**(1): 245.
367. Zhou, H., et al. (2015). "Variation in the bovine FABP4 gene affects milk yield and milk protein content in dairy cows." Sci Rep **5**: 10023.
368. Zhou, J., et al. (2005). "Elf5 is essential for early embryogenesis and mammary gland development during pregnancy and lactation." EMBO J **24**(3): 635-644.
369. Zhu, M., et al. (2011). "Integrated miRNA and mRNA expression profiling of mouse mammary tumor models identifies miRNA signatures associated with mammary tumor lineage." Genome Biol **12**(8): R77.
370. Zidi, A., et al. (2010). "Genetic variation at the goat hormone-sensitive lipase (LIPE) gene and its association with milk yield and composition." J Dairy Res **77**(2): 190-198.
371. Zou, M. R., et al. (2014). "Histone demethylase jumonji AT-rich interactive domain 1B (JARID1B) controls mammary gland development by regulating key developmental and lineage specification genes." J Biol Chem **289**(25): 17620-17633.

Publications

- Trinadha Rao Sornapudi, **Rakhee Nayak**, Prashanth Kumar Guthikonda, Anil Kumar Pasupulati, Srinivas Kethavath, VanitaUppada, SukalpaMondal, SailuYellaboina, and Sreenivasulu Kurukuti*. Comprehensive profiling of transcriptional networks specific for lactogenic differentiation of mammary epithelial stem-like cells. (**Scientific Reports 2018, PMID: 30082875**).

- Trinadha Rao Sornapudi, **Rakhee Nayak**, Prashanth Kumar Guthikonda, Srinivas Kethavath, Sailu Yellaboina, Sreenivasulu Kurukuti*.RNA sequencing of murine mammary epithelial stem-like cells (HC11) undergoing lactogenic differentiation and its comparison with embryonic stem cells. (**BMC Research Notes 2018, PMID:29642945**).

- Keerthivasan Raanin Chandradoss, Prashanth Kumar Guthikonda, Srinivas Kethavath, Monika Dass, Harpreet Singh, **Rakhee Nayak**, Sreenivasulu Kurukuti* and Kuljeet Singh Sandhu*. Biased visibility in Hi-C datasets marks dynamically regulated condensed and decondensed chromatin states genome- wide (**BMC Genomics 2020, PMID: 32087673**).

- Chandradoss KR, Chawla B, Dhuppar S, **Nayak R**, Ramachandran R, Kurukuti S, Mazumder A and Sandhu KS (2020). CTCF mediated genome architecture regulates the dosage of mitotically stable mono-allelic expression of autosomal genes. (**Cell Reports** 2020, PMID: 33113374).

Comprehensive Profiling and Functional Dynamics of microRNAs: A Developmental Perspective

by Rakhee Nayak

Submission date: 23-Aug-2021 03:30PM (UTC+0530)
Submission ID: 1634753494
File name: Rakhee_Nayak.docx (121.29K)
Word count: 34430
Character count: 202274

Comprehensive Profiling and Functional Dynamics of microRNAs: A Developmental Perspective

ORIGINALITY REPORT

7%	7%	4%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	genecopoeia.com.cn Internet Source	2%
2	downloads.hindawi.com Internet Source	1%
3	www.genecopoeia.com.cn Internet Source	<1%
4	www.biomedcentral.com Internet Source	<1%
5	academic.oup.com Internet Source	<1%
6	bmcgenomics.biomedcentral.com Internet Source	<1%
7	Azhwar Raghunath, Dhivyalakshmi Jeyabaskar, Kiruthika Sundarraj, Lakshmikanthan Panneerselvam, Ekambaram Perumal. "In silico prediction of microRNAs on fluoride induced sperm toxicity in mice", Food and Chemical Toxicology, 2016 Publication	<1%

8	Gilbert, Scott F.. "Developmental Biology", Oxford University Press Publication	<1 %
9	www.mdpi.com Internet Source	<1 %
10	Xiao Ma, Shuangshuang Cen, Luming Wang, Chao Zhang, Limin Wu, Xue Tian, Qisheng Wu, Xuejun Li, Xiaoqing Wang. "Genome-wide identification and comparison of differentially expressed profiles of miRNAs and lncRNAs with associated ceRNA networks in the gonads of Chinese soft-shelled turtle, Pelodiscus sinensis", Research Square, 2019 Publication	<1 %
11	link.springer.com Internet Source	<1 %
12	ajpheart.physiology.org Internet Source	<1 %
13	compbio.uthsc.edu Internet Source	<1 %
14	www.dharmacon.com Internet Source	<1 %
15	www.researchsquare.com Internet Source	<1 %
16	www.mirbase.org Internet Source	<1 %

17	hdl.handle.net Internet Source	<1 %
18	worldwidescience.org Internet Source	<1 %
19	www.embopress.org Internet Source	<1 %
20	Submitted to University of Nottingham Student Paper	<1 %
21	eprints.nottingham.ac.uk Internet Source	<1 %
22	oro.open.ac.uk Internet Source	<1 %
23	pericles.pericles-prod.literatumonline.com Internet Source	<1 %
24	www.repository.cam.ac.uk Internet Source	<1 %
25	etheses.whiterose.ac.uk Internet Source	<1 %
26	livrepository.liverpool.ac.uk Internet Source	<1 %
27	tdx.cat Internet Source	<1 %
28	mafiadoc.com Internet Source	<1 %

29	othes.univie.ac.at Internet Source	<1 %
30	Submitted to Middlesex University Student Paper	<1 %
31	www.nature.com Internet Source	<1 %
32	escholarship.umassmed.edu Internet Source	<1 %
33	wellcomeopenresearch.org Internet Source	<1 %
34	open.library.ubc.ca Internet Source	<1 %
35	www.promocell.com Internet Source	<1 %
36	archive.org Internet Source	<1 %
37	ir.uiowa.edu Internet Source	<1 %
38	www.freepatentsonline.com Internet Source	<1 %
39	www.jcancer.org Internet Source	<1 %
40	www.wjgnet.com Internet Source	<1 %

41	edocs.fu-berlin.de Internet Source	<1%
42	Submitted to University of Keele Student Paper	<1%
43	www.jneurosci.org Internet Source	<1%
44	Submitted to University of Leicester Student Paper	<1%
45	authors.library.caltech.edu Internet Source	<1%
46	www.diss.fu-berlin.de Internet Source	<1%
47	www.qiagen.com Internet Source	<1%
48	Cooper, Geoffrey. "The Cell XE", Oxford University Press Publication	<1%
49	arthritis-research.com Internet Source	<1%
50	dspace.alquds.edu Internet Source	<1%
51	www.hindawi.com Internet Source	<1%

Exclude quotes	On	Exclude matches	< 14 words
Exclude bibliography	On