Epigenetic reprogramming of hepatocyte-like cells

DISSERTATION

zur Erlangung des Grades

des Doktors der Naturwissenschaften

der Naturwissenschaftlich-Technischen Fakultäten

der Universität des Saarlandes

von

Wachiraporn Wanichnopparat

Saarbrücken

2020

Tag des Kolloquiums: 01.07.2021

Dekan: Prof. Dr. Jörn E. Walter

Berichterstatter: Prof. Dr. Jörn E. Walter

                         Prof. Dr. Sandra Iden

Vorsitz: Prof. Dr. Uli Müller

Akad. Mitarbeiter: Dr. Britta Diesel

ACKNOWLEDGEMENTS

This thesis cannot be successful without the support of all wonderful people around me, to whom I would like to express my sincere gratitude.

Foremost, I would like to express my sincere gratitude to my Doktorvater - Prof. Dr. Jörn Walter for the opportunity to join his group and for his kind guidance and valuable advice thoughout my Ph.D. study and thesis writing. His tremendous mental support always makes me brave and not giving up when I faced difficult times during the study. I could not have imagined having a better supervisor and mentor for my Ph.D. study.

My sincere thanks goes to Dr. Sascha Tierling and Nicole Jundel, who always support and believe in me, even when I am loss my confidence and all I believed in. Especially to Sascha, it was an absolute pleasure having him as my advisor. I am so grateful for all his support. He is always patient, dedicated and gives me a lot of wonderful suggestions. And especially to Nicole, I feel like home being with her, as my painfulness from homesick became better. All her physical and mental support really means so much to me.

My sincere thanks also goes to Prof. Dr. Fransiska Lautenschläger and Dr. Luiza Stankevicins for offering me the opportunity to learn and perform this exciting biophysical experiment. Special thanks to Dr. Luiza Stankevicins, for her collaboration and giving me very nice comments.

I am also thankful to the DPST scholarship – The Royal Thai Government for financial support throughout my study, also all staffs at Office of Educational Affairs – The Royal Thai Embassy for their assistance throughout the time I studied in Germany. My life could not be easy without all of them.

Many thanks to my colleagues in Epigenetic group for the precious times during my study. I always memorized our lab outing, Christmas party and Birthday brunch we shared the moments together.

I am very lucky being in our group. Many thanks to Beate, Christina, Jasmine, Katharina, Tabia, Karl, Abdul, Nicole, Pavlo, Sarah, Jie, Kathrin, Anna, Annemaria and Michael for all support during my study. Many thanks to Narges, although we spent very short time together, I was very happy talking and spending times with her. Special thanks to Nina, Jil and Kostya for their warm hug and cheering me up when I was frustrated. Also, special thanks to Pascal and Judith for their very good taking care during my difficult times. I was very happy and enjoy spending precious moments with them. Thank you very much for being my best friend.

My sincere thanks goes to my parents – Mr. Suttiporn and Mrs. Watana Wanichnopparat – my beloved Dad and Mom, and other family members, for their unconditional love and unlimited encouragement. Many thanks to Thai friends for all wonderful journeys, especially to Mr. Anan Schütt for his assistance when I am not in Germany. Finally, I am also very thankful to Mr. Wasakorn Laesanklang and Mr. Watcharapong Naraballobh for their brotherhood and very good taking care throughout my study.

ABSTRACT

The impact of physical microenvironment on cells was recently reported to have major influence on multipotency, self-renewal, cells plasticity and cell fate determination. In this study, we used physical-based methods, physical constraint (PC) and controlled confinement (CC), to induce multipotency of HepaRG, a bipotent progenitor of liver cells. RRBS and mRNA sequencing were performed to investigate changes in DNA methylation and mRNA levels during reprogramming and PC-induced differentiation. The results showed that PC and CC yielded similar results during reprogramming, e.g. locus-specific methylation changes towards hypomethylation and activation of immunomodulation process were observed. Particularly, CC also facilitated cell differentiation via IL-17 signaling. PC-induced differentiation of multipotent-like cells to hepatocyte-like cells was also combined with hydrocortisone and DMSO. Global hypomethylation was observed during PC-induced differentiation with the tendency towards a focal regain of methylation. In line with this, genes related to cytochrome P450 pathways were rebooted, suggesting that hepatic differentiation was partially induced. Furthermore, integrated analysis between differential methylation changes and chromatin states (ChromHMM) of differentiated HepaRG indicated that chromatin states were important in building the macromolecular structure that determine DNA methylation and gene expression during reprogramming and PC-induced differentiation of HepaRG.

ZUSAMMENFASSUNG

Physische Mikroumgebung von Zellen hat einen großen Einfluss auf die Multipotenz, Selbsterneuerung, Zellplastizität und Bestimmung des Zellschicksals. In der vorliegenden Arbeit wurden die physikalisch basierten Methoden, „*Physical Constraint*" (*PC*) und „*Controlled Confinement*" (*CC*) verwendet, um die Multipotenz von HepaRG- Zellen, einem bipotenten Vorläufer von Leberzellen, zu induzieren. Die genomweite Analyse molekularer Signaturen (DNA-Methylierung und Transkriptom) ergaben während der Reprogrammierung ähnliche Ergebnisse zwischen *PC* und *CC*, z.B. wurden lokus-spezifische Methylierungsänderungen in Richtung Hypomethylierung, sowie die Aktivierung des Immunmodulationsprozesses, beobachtet. Insbesondere *CC* förderte die Zelldifferenzierung mittels IL-17 Signaltransduktion. Die *PC*-induzierte Differenzierung (*PC*i) multipotenter Zellen zu Hepatozyten-ähnlichen Zellen unter Zugabe von Hydrocortison und DMSO resultierte in globaler Hypomethylierung, mit der Tendenz zur erneuten Etablierung fokaler Methylierung. Parallel hierzu wurden Gene, die mit Cytochrom P450 Signalwegen verbunden sind, wieder aktiviert, was darauf hindeutet, dass die *PC*i-Differenzierung teilweise erreicht wurde. Darüber hinaus zeigte die integrative Analyse von differentiellen Methylierungsänderungen und Chromatinzuständen (ChromHMM) differenzierter HepaRG die Wichtigkeit des Chromatinzustands beim Aufbau makromolekularer Strukturen, die die DNA-Methylierung und Genexpression während der Reprogrammierung und *PC*i-Differenzierung dieser Zellen bestimmen.

Table of Contents

Chapter 1 Introduction

1.1. Epigenetics

1.1.1. Epigenetic mechanisms

Epigenetics is the study of the modifications of DNA without changing nucleotide sequences. Epigenetic states are stable and mitotically inherited leading to cellular identity maintenance during environmental changes (Meissner, 2010). Most of epigenetic mechanisms are reversible enzymatic modifications, including DNA methylation, histone, and non-histone modifications (Kouzarides, 2007; Weber and Schübeler, 2007). Besides, micro RNAs, non-coding transcribed RNAs, are included in the epigenetic mechanisms as RNA-based mechanisms as well (Mattick *et al.,* 2009) (Figure 1.1). Epigenetic mechanisms are also fine-tuning tools to provide an alteration of phenotypes without changing genotypes. Many relevant biological processes are highly influenced by epigenetics such as embryogenesis, cell-differentiation, and cancer etc. (Herranz and Esteller, 2007; Reik, 2007; Ballestar and Esteller, 2008; Meissner, 2010). Defects in some parts of epigenetic mechanisms can cause malformation and lethality (Okano *et al.,*1999; Cortazar *et al.,* 2011).

Regarding histone proteins and non-histone proteins, the modifications are set on a post-translational level. Histone is a protein that structurally supports the wrapping of DNA to form more complex structure, known as chromatin, inside the cell nuclei (Figure 1.1). Several types of modifications such as acetylation, methylation, and ubiquitination etc. occur at N-terminal tails of histone core subunits, frequently on lysine residue or arginine residue of histones H3 and H4 (Kouzarides, 2007). However, in non-histone proteins, the only modification found is lysine phosphorylation (Zhang and Pradhan, 2014).
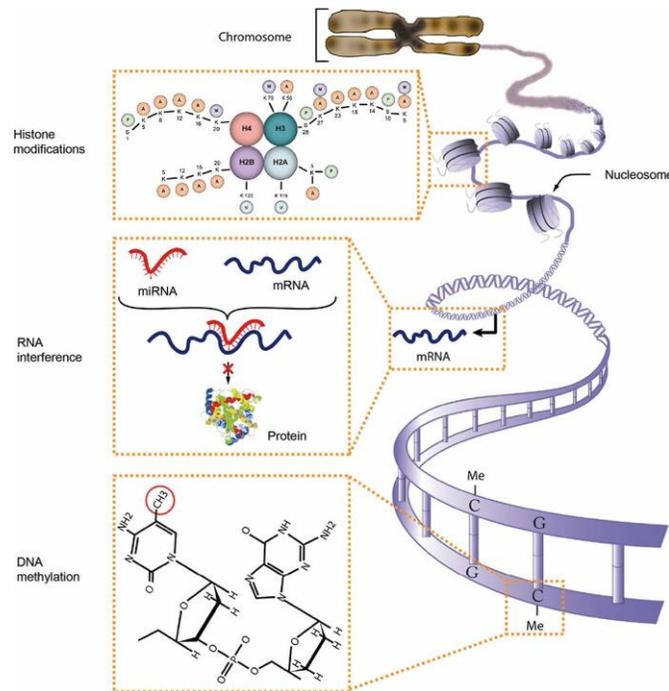
Figure 1.1 Epigenetic mechanisms and genome organization (Puumala and Hoyme, 2015). This figure displays that DNA methylation and RNA interference are the epigenetic mechanisms found on DNA, regulating in transcriptional level. DNA can be packed with histone proteins forming the nucleosome, which is further folded to chromatin and chromosome, respectively. Histone modifications belong to the modifications that organize chromatin condensation and influence local transcription activity.

DNA methylation is another enzymatic modification found mostly in CpG dinucleotides throughout the genome context (Lister *et al.,* 2009; Jones, 2012). Methylation and demethylation of DNA are closely associated with open or closed chromatin (Tamaru, 2010). DNA methylation also influences the binding activity of numerous transcription factors to the DNA, and regulation of transcription process of the genes (Bell and Felsenfeld, 2000; Lister *et al.,* 2009).

MicroRNAs are small non-coding RNAs that regulate gene activity on the post-transcriptional level. This single strand RNA interacts with a messenger RNA of target genes and induces a degradation of those gene transcripts. Moreover, microRNA can act on the DNA level by hypermethylation of its host gene and its own promoters, which affects their transcription activity as well (Mattick *et al.,* 2009; Augoff *et al.*, 2012; Gulyaeva and Kushlinskiy, 2016).

13

1.1.2. <u>Genome organization</u>

As eukaryotic genomes are complex in their structure, in general, the genome consists of a combination of DNA and histone proteins called "histone octamers" which is densely packed in the nucleus. Histone octamer is a protein that composes of two units of histones H2A, H2B, H3 and H4 (Luger *et al.,* 1997). This core protein is wrapped by 147 base pairs of double stranded DNA approximately to form a nucleosome, a basic packaging unit that is important for determining DNA accessibility (Luger *et al.,* 1997; Richmond and Davey, 2003). Each nucleosome is separated from its neighbor nucleosome by 20-50 base pairs linker DNA (Routh *et al.,* 2008). The folding of nucleosomes can form a higher order as a chromatin unit, and further a chromosome (Bednar *et al.,* 1998). Thus, the eukaryotic genomes take major advantage of folding their structure for preventing genetic codes from any harm, such as nucleases and restriction enzymes. DNA methylation and histone modifications are responsible for genome organization, as the chromatin states can be determined through the reversibility of modifications (Tamaru, 2010). Euchromatin, an active state of chromatin, is loose and accessible, and frequently present in gene-rich regions. Moreover, this active chromatin state is marked by acetylated histones H3 and H4. On the other hand, heterochromatin, a densely packed chromatin is found in gene-poor regions, as well as repetitive elements. Methylated DNA and trimethylation of H3K9 and H3K27 can be found in heterochromatin regions (Fischle *et al.,* 2003; Lehnertz *et al.,* 2003; Gilbert *et al.,* 2004). However, during cell replication, epigenetic marks like DNA methylation and histone modifications are maintained and inherited to the daughter cells (Martin and Zhang, 2007).

1.1.3. <u>DNA methylation and demethylation processes in mammals</u>

DNA methylation is a repressive epigenetic mark playing crucial roles in several biological processes. For instance, the regulation of monoallelic gene expression such as x-inactivation and imprinted genes, the silencing of transposable elements, embryogenesis and differentiation, and development and maintenance of cancer states are controlled by DNA methylation (Yoder *et al.,*

1997; Bell and Felsenfeld, 2000; Hellman and Chess, 2007; Herranz and Esteller, 2007; Laurent *et al.,* 2010). Generally, the methylation of DNA sequences can be found throughout the genome, predominantly at CpG dinucleotides (Holliday and Pugh, 1975). Although DNA methylation of CpGs is established widespread, non-CpG methylation (CA, CT, or CC) is also recognized as it is important for development of embryonic stem cells and brain development (Ramsahoye *et al.,* 2000; Lister *et al.,* 2009; Dyachenko *et al.,* 2010; Jang *et al.,* 2017).

Methylation patterns differ among genomic regions. CpG-rich promoters in CpG islands are mainly unmethylated since they regulate the expression of downstream regions, while the other regions show partially methylation (Lister *et al.,* 2009; Jones, 2012). Because of the variable methylation levels in specific compartments of the genome, the pattern of established methylation leads to specification of tissue differentiation (Song *et al.,*2005; Lister *et al.,* 2009; Laurent *et al.,* 2010).

To methylate DNA, the methyl group of s-adenosylmethionine (SAM) is transferred to the 5th carbon of cytosine (C) in palindromic CpG sequences by DNA methyltransferases (DNMTs). Thus, the cytosine base becomes 5-methylcytosine (5mC) (Kohli and Zhang, 2013) (Figure 1.2)
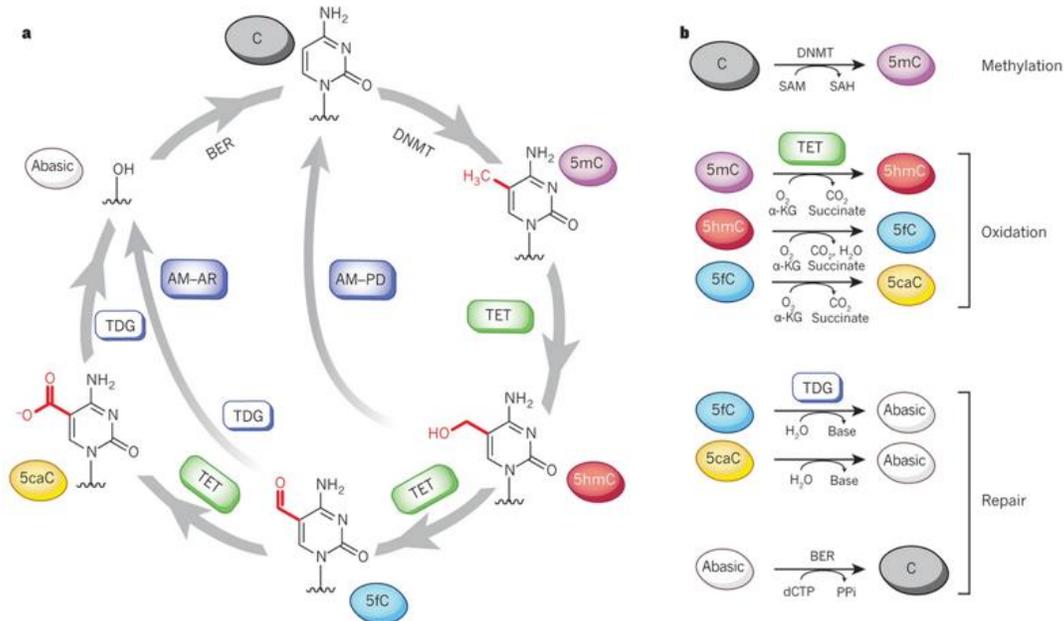
Figure 1.2 Enzymatic reactions leading to DNA methylation and demethylation (Kohli and Zhang, 2013). a) showing a circle diagram of cytosine and other derivatives in methylation and demethylation processes, while b) showing details and co-factors corresponding to each step of the circle diagram.

As described above, the establishment and maintenance of DNA methylation in the genome are regulated by active enzymes of the DNA methyltransferase family (DNMTs) coordinated with other co-factors. There are two groups of DNMTs related to this process. First, DNMT1 is an enzyme that functions predominantly in maintenance of DNA methylation patterns of the daughter strand after DNA replication. Working together with UHRF1, DNMT1 preferentially adds a methyl group to hemimethylated CpGs (Hermann *et al.,* 2004; Bostick *et al.,* 2007). The other enzymes of this family are DNMT3A and DNMT3B, which function in the de novo methylation process and collaborat with their co-factor, DNMT3L, in the establishment of parental DNA methylation in germ line cells.(Okano *et al.,* 1999; Ooi *et al.*, 2007). However, establishment and maintenance functions can be compensated partially by DNMT1 and DNMT3A/B, respectively (Fatemi *et al.,* 2002; Chen *et al.,* 2003).

To remove the methylation from DNA, there are two processes of demethylation in cells, passive and active demethylation. The passive demethylation occurs during DNA replication, while the active demethylation is linked to the activities of enzymes from the Ten-Eleven Translocation

(TET; TET1, TET2 and TET3) family and to the DNA repair pathway (Wu and Zhang, 2010; Kagiwada *et al.,* 2013). For active demethylation, TETs oxidize methylated cytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxycytosine (5caC), which can be detected and removed by enzymes from the base-excision repair (BER) pathway (Kohli and Zhang, 2013; Wu and Zhang, 2014) (Figure 1.2). On the other hand, in *In vitro* experiments, demethylation can be induced via the addition of demethylating agents such as 5-Azacytidine (5-Aza), 2'-deoxy-5-Azacytidine (DAC) and other derivatives. Previously, 5-Aza was frequently used in cancer therapy as it can result in activation of tumor suppressor genes. 5-Aza can be phosphorylated to 5-AzaCTP, which was used for the incorporation into newly synthesized DNA strand, instead of CTP, leading to demethylation of the whole genome (Christman, 2002; Sigalotti *et al.,* 2007). However, 5-Aza was also found to incorporate into tRNA and influence gene expression (Schaefer *et al.,* 2009; Qiu *et al.,* 2010).

1.1.4. DNA methylation and demethylation during embryogenesis

DNA methylation and demethylation are relevant mechanisms to establish a successful embryo development. Those processes are conserved in mouse and human (Fulka *et al.,* 2004; Xu *et al.,* 2005; Guo *et al.,* 2014; Li *et al.,* 2018). Several studies found that the methylation level of the paternal genome is higher than the maternal genome, and rapidly decrease after fertilization (Guo *et al.,* 2014; Li *et al.,* 2018). Demethylation of the paternal genome is a result from active demethylation processes, as the level of 5mC decrease while the level of 5hmC is detected increasing. Furthermore, TET3 was stronger expressed than the other TET family members during early embryogenesis, supporting their role in the increase of 5hmC in the paternal genome (Oswald *et al,* 2000; Iqbal *et al.,* 2011). Contrary to the paternal genome, the methylation level of the maternal genome is lower (Guo H *et al.,* 2014; Li *et al.,* 2018). In the maternal genome, methylated was protected from active demethylation process leading to a slow demethylation process. Stella, a protein encoded by *PGC7* and predominantly found in maternal genome, played

roles in the protection of demethylation in associated with H3K9me2 (Iqbal *et al.,* 2011; Nakamura *et al.,* 2012; Guo H *et al.,* 2014). Moreover, methylation of the maternal genome seemed to dependently decrease after DNA-replication. Passive demethylation was proposed to be a mechanism regulating the methylation level in the maternal genome (Santos *et al.,* 2002; Wu and Zhang, 2010; Messerschmidt *et al.,* 2014). However, there were some studies revealing that demethylation processes during embryogenesis were performed passively and actively in the paternal and maternal genomes (Guo F *et al.,* 2014; Shen *et al.,* 2014). The resetting of methylation in paternal and maternal genomes is maintained steadily from the morula to the blastocyst stage. Then the methylation is newly established by DNMTs in the implantation stage (Guo H *et al.,* 2014; Li *et al.,* 2018).

1.2. Pluripotency and multipotency of the cells

1.2.1. Molecular features of induced pluripotent stem cells

Cellular reprogramming is a process that converts somatic cells from differentiated state to pluripotent state. Several approaches are used in cellular reprogramming, such as somatic cell nuclear transfer and stem cell fusion, but the well-known approach is direct reprogramming or induction of pluripotency by defined transcription factors (Takahashi and Yamanaka, 2006; Takahashi *et al.,* 2007; Yamanaka and Blau, 2010; Buganim *et al.,* 2013).

Direct reprogramming applies an ectopic expression of OSKM (*POU5F1* (OCT3/4), *SOX2*, *KLF4* and *MYC*) to establish a pluripotent state of mature cells (Takahashi and Yamanaka, 2006; Takahashi *et al.,* 2007). Cells which are successfully reprogrammed are known as "induced pluripotent stem cells (iPSC)". Although iPSCs are reported to be comparable to embryonic stem cells (ESC), some molecular features, e.g. gene expression and epigenetic states are still different from ESC.

During reprogramming, the alteration of transcription and epigenetic modifications are prominently observed. On the level of transcription, two waves of changes were found to mainly drive throughout the whole process. Primary wave starts in the early reprogramming process, and this stochastic process is conducted by *MYC* and *KLF4* (Hanna *et al.,* 2009; Buganim *et al.,* 2012; Polo *et al.,* 2012; Buganim *et al.,* 2013)*. MYC* plays important roles in the introduction of early genes, and the expression of *MYC* is maintained steadily until the end of reprogramming (Sridharan *et al.,* 2009; Polo *et al.,* 2012, Rand *et al.,* 2018). Meanwhile, *KLF4* suppresses somatic gene expression, leading to pluripotent gene activation (Zhang *et al.,* 2010; Polo *et al.,* 2012). Furthermore, a surface marker gene, *SSEA1*, is gradually expressed and maintained in cells undergoing pluripotent stage (Buganim *et al.,* 2012; Polo *et al.,* 2012). During primary wave, mesenchymal to epithelium transition (MET) is also initiated through BMP signaling as mesenchymal genes are potentially switched off and prepared for the activation of epithelial genes (Samavarchi-Tehrani *et al.,* 2010). Genes related to cell proliferation are also upregulated leading to high proliferation. Moreover, a metabolic switch was also observed (Hanna *et al.,* 2009; Hansson *et al.,* 2012; Polo *et al.,* 2012). Oxidative phosphorylation seemed to be suppressed, while the glycolysis was predominantly activated (Folmes *et al.,* 2011; Hansson *et al.,* 2012; Panopoulos *et al.,* 2012; Zhang *et al.,* 2012). Secondary wave was observed in the late period of reprogramming. This period was a hierarchical process driven by *POU5F1* (OCT3/4)*, SOX2* and *KLF4.* Since primary transcription wave was successfully established, the consequently expressed genes e.g. *ESRRB* or *UTF1* became a key regulator to activate *SOX2* (Buganim *et al.,* 2012; Polo *et al.,* 2012) in late reprogramming. *POU5F1* (OCT3/4) and *SOX2* played similar roles in activation of endogenous pluripotent circuity and downstream networks, which were a target of OS factors. Nevertheless, expression of *POU5F1* seems to be less necessary than the expression of *SOX2*. Therefore, *SOX2* serves as a main regulator of late reprogramming. By the end of the secondary transcriptional wave, successful iPSCs are stably established and genes related to embryonic development and stem cell maintenance were upregulated (Polo *et al.,* 2012). The low

efficiency of direct reprogramming was found to be involved in the heterogeneity of the cells undergoing iPSC, which correlated with the origin and epigenetic memory of somatic cells (Kim *et al.,* 2010; Ohi *et al.,* 2011; Hansson *et al.,* 2012).

1.2.2. Epigenetic regulation of iPSCs

Passing through epigenetic barriers is a necessary process iPSCs need to complete since epigenetics provides the underlying mechanisms for the regulation of pluripotency. Concerning the histone modifications, H3K4me3 and H3K27me3 could be found in regions known as bivalent domains. These bivalent domains contain both repressive and active histone marks and found predominantly in genes which need rapid activation or repression; number of these domains are increased gradually during iPSC induction (Mikkelsen *et al.,* 2007; Polo *et al.*, 2012). Furthermore, H3K36me3 was also found in active genes, whereas H3K9me3 and H4K20me3 were found in silenced domains, e.g. in transposable elements or in centromeric regions etc. (Mikkelsen *et al.,* 2007). However, H3K36 and H3K20 methylation seem to be not important for the regulation of iPSC fate (Chen *et al.,* 2013). Several genes encoding histone modification enzymes were also found to have a strong correlation with OSKM factors. For instance, *KDM2B* encoding the H3K36me2 demethylase was upregulated during early reprogramming by OSK. *KDM2B* facilitated iPSC induction by demethylation of epithelial gene promoters, leading to the activation of downstream pluripotent cascades (Liang *et al.,* 2012). Another example is *SETDB1* encoding the H3K9 tri-methyltransferase. Knockdown of this gene in the appearance or absence of vitamin C in the culture medium resulted in rescue reprogramming mediated by OSKM. Similarly, knockdown of *SUV39H1,* a BMP target H3K9 methyltransferase, enabled pre-iPSC cells to establish and maintain an epigenetic barrier during reprogramming (Chen *et al.,* 2013; Sridharan *et al.,* 2013).

DNA methylation was found to be reset, particularly in the late process of reprogramming (Polo *et al.*, 2012). Since Takahashi *et al.* established a direct reprogramming method by the induction

of defined transcription factors, DNA methylation was also analyzed to figure out the epigenetic status of iPSC (Takahashi and Yamanaka, 2006; Takahashi *et al.,* 2007). Bisulfite genome sequencing revealed promotor demethylation of pluripotency-associated genes e.g. *POU5F1, NANOG* and *REX1* (Takahashi *et al.,* 2007; Huangfu *et al.,* 2008). Nishino *et al.* reported that the methylation patterns of human iPSCs was similar to human ESCs and distinct from the parental cells although a small number of CpG sites of iPSCs were slightly higher methylated than in ESCs and the parental cells. Differentially methylated regions (DMRs) were also analyzed in this study and the results were consistent with the results extracted from single CpG sites. Moreover, aberrant methylation of iPSCs was also reported and preferentially occurred in CpG islands. Aberrant methylation was inherited from the parental cells as an epigenetic memory and varied among iPSCs lines. Furthermore, incomplete establishment and aberrant DNA methylation during early passages were still detected influencing iPSCs induction efficiency and some somatic gene expressions. However, after long-term cultivation of iPSCs, aberrantly methylated regions were decreased gradually along passaging, suggesting that methylation profiles of iPSCs became closer to ESCs (Deng *et al.,* 2009; Lister *et al.,* 2011; Nishino *et al.,* 2011; Ohi *et al.,* 2011; Nishino and Umezawa, 2016). Genes related to DNA demethylation were also reported to influence the reprogramming process, particularly genes belonging to the TET family. As described previously, members of TET family played important roles in the oxidation of 5mC to become 5hmC and other derivatives, TET enzymes became relevant factors in reprogramming process, since 5hmC was shown to accumulate upon human iPSCs induction. *TET1* was reported to promote reprogramming efficiency via hydroxymethylation of CpGs (Wang *et al.,* 2013). Although the reprogramming efficiency was lower, other studies revealed that *TET1* was used to replace OCT4 protein in OSKM cocktails, leading to the initiation of reprogramming in mice. The studies also observed the enrichment of 5hmC at the promoter and enhancer of *POU5F1* as well as at OCT4 and KLF4 binding sites, which was a result of TET1 activity (Gao *et al.,* 2013; Wang *et al.,* 2013). In mice, during an early stage of reprogramming, TET2 and Parp1 were observed to be recruited

21

to *Nanog* and *Esrrb* loci. Not only playing a role in oxidizing 5mC to 5hmC, *Tet2* is also found to associate with H3K27 methylation level at those pluripotent loci. Therefore, *Tet2,* also *Parp1,* seem to be relevant in establishing epigenetic marks during the early reprogramming (Doege *et al.,* 2012). A triple knockdown of genes encoding TET and TDG enzymes resulted in unsuccessful reprogramming as MET process was blocked in Tet/Tdg deficient cells (Hu *et al.,* 2014).

1.2.3. Multipotent stem cells

Multipotent stem cells are another kind of stem cells found in mammals. This kind of stem cells was defined by their capability of self-renewal and their plasticity in differentiation to several cells (Tanabe, 2014). Generally, multipotent stem cells were isolated from adult samples although they were fully differentiated. Nevertheless, limitations in isolation and prolonged cultivation are still critical for using them in regenerative medicine and drug testing, as the function of daughter cells was incomplete and incomparable to primary cells, depending on their microenvironment (Tanabe, 2014; Lv *et al.,* 2014; Almalki *et al.,* 2016; Wang *et al.,* 2016). Well-known multipotent stem cells include hematopoietic stem cells (HSCs), bone-narrow derived mesenchymal stem cells (MSCs), neuronal stem cells (NSCs) etc. (Dan *et al.,* 2006; Mehanni *et al.,* 2013; Tanabe, 2014). Particularly, multipotent cells are also known and can be isolated from human fetal liver. For the maintenance in a specific microenvironment, human fetal liver cells are kept on feeder layers in an undifferentiated stage and differentiated to various cell types, such as hepatocytes, biliary cells, bone, endothelial cells etc. by culturing in defined conditions (Dan *et al.,* 2006).

DNA methylation is known as an essential mechanism in the regulation of MSCs differentiation. Bock *et al.* reported that DNA methylation profiles of HSCs and stem cell differentiation to skin fibroblasts shed light on the initiation and maintenance of their lineages. In HSCs, DNA methylation levels acted as a regulator of transcription factors binding, particularly between myeloerythroid and lymphoid lineages. DNMT1 depletion in mouse HSCs also resulted in alteration of HSCs homeostasis suggesting the dynamics of methylation in stem cell function and

differentiation (Bröske *et al.,* 2009; Bock *et al.,* 2012). Furthermore, genes of the MBD family, which encode methyl-CpG-binding proteins, influenced multipotency and differentiation of NSCs as well (Lax and Sapozhnikov, 2017).

1.3. Physical constraint and cell geometric confinement

1.3.1. <u>Mechanotransduction, Physical constaint and cell geometric confinement</u>

Controlling of cellular properties, particularly in MSCs maintenance and differentiation, affects the coordination of microenvironment and intracellular signals. The differences of microenvironments have an impact on cell shapes and physical structure since cells are directly attached to the material surface and surrounded by cultured medium (Engler *et al.,* 2006; Wang *et al.,* 2016). Morphogens such as growth factors or supplements in cultured medium also contribute to the control of cell morphology, as well as physical characteristics rise from the surface of the materials, to which the cells are attached (Ruiz and Chen, 2008; Murphy *et al.,* 2014). Thus, microenvironment is an important factor to generate bio-mechanic stress and influence signal switching, from extracellular to intracellular levels (Figure 1.3). Thereafter, cells transduce mechanical signals to biochemical signaling networks, leading to changes in cell behavior and

functionality called "mechanotransduction" (Wang *et al.,* 2009; Wozniak and Chen, 2009; Humphrey *et al.,* 2014).
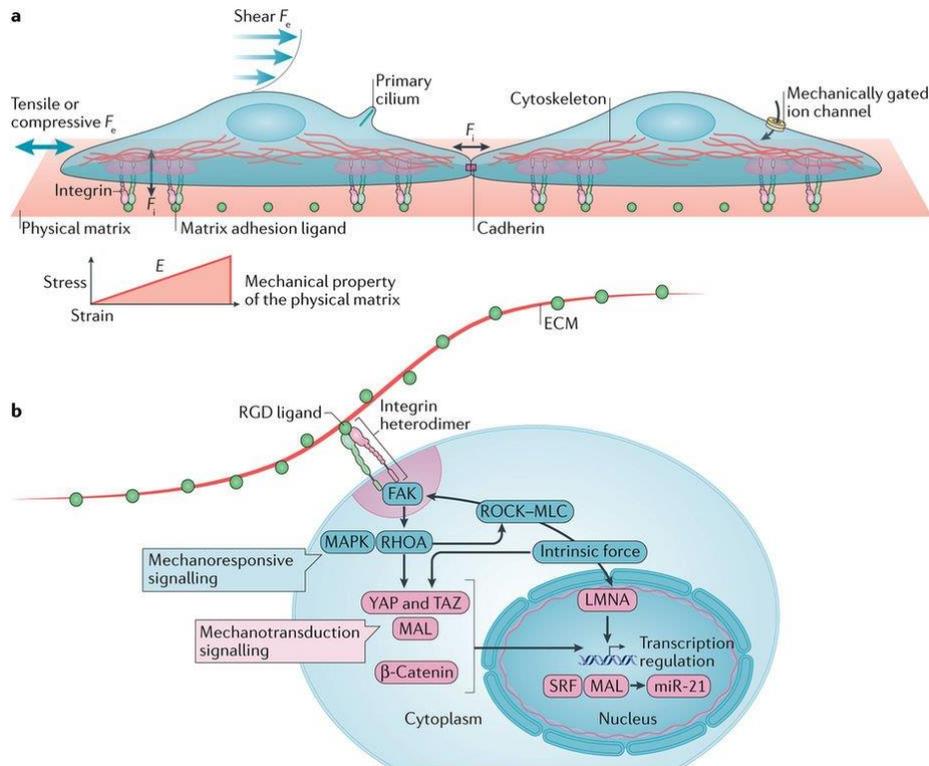

Figure 1.3 Mechanotransduction and mechanoresponsive signaling (Vining and Mooney, 2017). a) shows several kinds of physical forces applying to the cells, such as shear stresses, tensile or compressive and focal adhesion etc. While strains (mechanical property of the physical matrix) change, cellular stresses also become increasing. b) shows external forces are transduced to the mechanoresponsive signaling, leading to transcriptional changes in the cell nuclei.

Physical constraint (PC) is a study method used to restrict the microenvironments of cells physically, e.g. by limiting cell shapes/areas or by altering substrate elasticity (stiffness) of surface materials to which the cells are attached. PC was closely associated to mechanotransduction processes. When the physical microenvironment changes, cell shape, cell adhesiveness and cell softness are relevant for the responsiveness to stimuli. As ESCs are softer than adult cells, ESCs were found to better response to stress than adult cells, leading to changes in cell function and behavior (Holle and Engler, 2010).

Several studies revealed that PC is important in biological processes during embryonic stem cell differentiation. Chowdhury *et al.* revealed that when cultured on the same substrate/surface, mouse ESCs were found to be more responsive to stress than differentiated cells. By applying stress to undifferentiated ESCs, the expression of *Pou5f1* was decreased, even in the presence of LIF in the culture medium, suggesting that a stress-induced internal force was generated and facilitated differentiation of those undifferentiated cells (Chowdhury *et al.*, 2010). In mesenchymal stem cells (MSC) cultured in a limited circular shape, with medium containing adipocyte and osteocyte growth factors, it was found that the patterning of differentiation was determined by a gradient of mechanical forces. At the edge of the circle, which comprises of high stress, MSC preferentially differentiated to osteocytes. In contrast to the edge of the circle, the middle of the circle was a low stress area. MSCs in the low stress area preferentially differentiated to adipocytes (Ruiz and Chen, 2009).

Previous studies showed that PC facilitates maintenance and differentiation in stem cells. Regarding the study of cell migration or cell division, PC seems difficult to precisely manipulate, either in single cells or in cell populations. Consequently, geometric confinement (GC) or two-dimensional cell confinement (referred to "controlled confinement (CC)" in this study) is a method designed to provide homogeneous experiment conditions to study roles of geometry and cell microenvironment (Le Berre *et al.,* 2012; Le Berre *et al.,* 2014).

The impact of GC in cell migration and cell division have been reported recently. Generally, fibroblast cells migrated along substrates in mesenchymal fashion via focal adhesions, which were observed to be elongated in a spindle-like shape. During physical confinement and low adhesion, the mesenchymal-amoeboid transition was induced as well as fast amoeboid-like migration was observed. These findings were frequently found in the migration of metastatic cancer cells and dendritic cells etc. (Heuze *et al.,* 2013; Lui *et al.,* 2015). Effects of physical confinement on cell division appear to be more complex than on cell migration, since the effects

influencing cell division can have impact on phases of mitotic progression. Under strong confinement, cell division was halted during G0/G1 phase of mitosis and prometaphase was delayed and increased with time (Desmaison *et al.*, 2013; Desmaison *et al.,* 2018). Furthermore, application of high tension on the cell membrane led to a delayed cutting of the cytokinetic bridge during anaphase (Lafaurie-Janvore *et al.,* 2013). Limiting cell height also had an impact on rounding of the cells during mitosis. Lancester *et al.* showed that mitotic spindle assembly and poles splitting were defect when cells failed to round up (Lancester *et al.,* 2013). Additionally, F-actin was observed to be involved in mitotic spindle formation. Increasing confinement forces caused loss of cortical F-actin homogeneity and promoted impaired microtubule spindle and persistent blebbing during mitotic progression. Therefore, actin proved to be essential for cells to pass through mitosis under physical confinement (Lancester *et al.,* 2013; Cattin *et al.,* 2015).

1.3.2. <u>Cell geometric confinement in reprogramming and differentiation</u>

Any alterations directly affecting the nuclei of the cells also have the effects on chromatin leading to chromatin deformation and changes in gene expression influencing cellular properties (Chalut *et al.,* 2012; Le Berre *et al.,* 2012). Several studies showed that alteration of physical microenvironment is closely associated with maintenance of cellular properties or lineage-specific determination during differentiation through regulation of epigenetic factors. Chalut *et al.* generated cells harbouring high (HN) and low Nanog (LN) expression, which is closely connected to high and low cellular pluripotency, respectively. During optical stretching, a significant deformation was observed in nuclei of LN cells, rather than HN cells. Similarly, lower distribution of heterochromatin foci markers such as HP1α or H3K9me3 were found in LN cells, suggesting that decondensation of chromatin occurs genome-wide, particularly in cells primed to differentiation (Chalut *et al.,* 2012). Volume confinement with microfluidics also leads to higher efficiency in the delivery of defined transcription factors during reprogramming, as well as cell compression-based cell stiffness promotes the reprogramming efficiency of cells. (Pagliara *et al.,*

2014; Luni *et al.,* 2016). Epigenetic factors were also changed when the physical microenvironment becomes limited. For instance, histone H3 acetylation and methylation were increased and enriched, particularly at promoters of pluripotent genes e.g. *Sox2*. Moreover, mesenchymal-epithelial transition (MET), which is known as a hallmark of pluripotency, was also activated during compression (Downing *et al.,* 2013; Caiazzo *et al.,* 2016).

1.4. Biology of HepaRG and HepaRG differentiation

1.4.1. Biology of HepaRG

Previously, toxicological studies were performed in animal models touching the ethics of animal usage for drug discovery and testing. Therefore, cell lines that were comparable to primary cells were introduced providing better ethical solutions compared to animal-based experiments.

HepaRG is a cell line that is frequently used in toxicological studies. Originally, HepaRG was isolated from a female patient suffering from an Edmonson grade I differentiated hepatocellular carcinoma with chronic hepatitis C virus (HCV) infection. The genome of HepaRG was pseudodiploid with partially remodeled trisomy of chr7 and chr12:22 translocation. Fortunately, although the HCV genome was not detected after the isolation, functions supporting the viral infection were still maintained. (Gripon *et al.,* 2002; Troadec *et al.,* 2006). Therefore, HepaRG becomes an ideal model for drug testing and mechanisms of HBV infection studies since the cells provide comparable functions to adult hepatocytes.

Differentiation of HepaRG can be promoted using culture supplements, corticoids and DMSO, and an increased cell density in the culture. Undifferentiated HepaRG were maintained when seeding cells in the culture at low density. To induce differentiation, high density HepaRG culture was treated with 2% DMSO in the presence of hydrocortisone for 4 weeks. HepaRG cells rapidly proliferated for 2-7 days with a subsequent decrease in proliferation activity after 7 days. In this phase, cells displayed an epithelial morphology. Thereafter, hepatocyte-like and biliary-like

morphology were observed between days 7-14 as a mixed culture. Distinct population of cells could be observed after 14 days. While hepatocyte-like cells displayed small polarized cells with refractile borders and dark cytosol, biliary-like cells displayed small regions of bile canaliculi-like structures (Gripon *et al.,* 2002; Parent *et al.,* 2004; Aninat *et al.,* 2006; Troadec *et al.,* 2006).

1.4.2. <u>Gene expression and epigenetic regulation during HepaRG differentiation</u>

Gene expression profiles of undifferentiated and differentiated HepaRG were previously analyzed using Affymetrix U133 plus 2.0 array. In comparison to primary human hepatocytes (PHH), liver tissues and HepG2, gene expression profiles of both stages of HepaRG, undifferentiated and differentiated, were more closely associated to PHH than liver tissues and HepG2 profiles. Genes encoded for phase I and II of drug metabolizing enzymes were also expressed at a comparable level to PHH, such as cytochrome P450 and UDP family (Hart *et al.,* 2010). Moreover, in differentiated HepaRG, the expression of enzymes and their activities were similar, even when cell density and passages were changed. The level of major cytochrome P450 was still preserved in different conditions, it even became stronger in the presence of DMSO (Antherieu *et al.,* 2010; Hart *et al.,* 2010;). However, Bell *et al.* argued that PHH spheroids revealed transcriptomic differences between fresh isolated PHH, HepaRG and hepatocyte-derived iPSCs, leading to differences in the regulation of cellular functions. Gene ontology analysis of upregulated genes in HepaRG and hepatocyte-derived iPSCs were shown to be involved in biological processes related to cell cycle, ribosome and RNA transport, while upregulated genes in PHH spheroids were involved in metabolic pathways, retinol metabolism and xenobiotic metabolism (Bell *et al.,* 2017). Although many studies generated and worked with transcriptomic data during HepaRG differentiation, DNA methylation data were not frequently found. Whole genome bisulfite sequencing and partially methylated domain (PMD) analysis revealed that the HepaRG methylome seem to have similarities to HepG2. In contrast, PHH and normal liver tissues became more similar to primary liver cancer than HepaRG and HepG2, suggesting that cultivation has

major effects on epigenetic marks (Salhab *et al.,* 2018). Furthermore, in HepaRG containing HVB infection and lipid accumulations, hypermethylation was observed at various imprinted loci (Lambert *et al.,* 2015).

1.4.3. Roles of HepaRG in the induction of pluripotency

As undifferentiated HepaRG cells harbour bipotent properties in differentiation to two cells types, HepaRG became a candidate cell line used in iPSC induction. Many studies revealed that the pluripotency and lineage-specific differentiation of HepaRG were switched by physical microenvironment such as cultivating them in three-dimensional form or in medium containing growth factors and various oxygen concentrations. van Wenum *et al.* reported that cultivating HepaRG in hypoxia condition (5% O_2) led to the maintenance of stem cell functions via upregulation of *SOX2* and *HIF1α* as well as inhibition of hepatic differentiation (van Wenum *et al.,* 2018). Cultivated on synthetic polymeric hydrogels or in 3D spheroid form, HepaRG preferentially differentiated to hepatocyte-like cells. Those results were confirmed by immunostaining and expression of hepatocyte markers. Meanwhile, inhibition of differentiation to biliary cells were observed as biliary marker genes were suppressed (Higuchi *et al.,* 2016; Kanninen *et al.,* 2016). To induce cholangiocytes, HepaRG cells were cultivated in the medium containing interleukin-6, sodium taurocholate hydrate, followed by the combination of sodium taurocholate hydrate and sodium butyrate to prevent hepatic differentiation. At the end, HepaRG-derived cholangiocytes were successfully differentiated, *GGT1, CK19, TGR5* were up-regulated, whereas *ALB* and *HNF4α*, hepatotic markers, were down-regulated (Dianat *et al.,* 2014).

1.5. Aims of the study

Limitation of HepaRG usages in toxicological and virological studies was arisen from restricted use of the cells after differentiation and prolonged cultivation. As the capacity of HepaRG as bipotent progenitor cells was known, highlighted an opportunity to generate cells containing multipotent properties by changing the conditions of the microenvironment. Using PC-based methods in previous differentiation studies revealed promising preliminary results e.g. hepatic differentiation was increased and speeded-up significantly. Analysis by Affymetrix data showed that fewer transcripts related to cell plasticity or reprogramming signals were induced after PC (unpublished data, biopredic). Besides, although methylation of some hepatocyte-specific genes such as *CYP1A2, CYP3A4, CYP2D6* etc. was not changed after PC, hypomethylation of those genes were observed after Aza treatment, suggesting that Aza enabled cells to stabilize cellular properties. Gradual restoration of methylation was also observed further after PC-induced differentiation (PCi-differentiation or redirected differentiation) to hepatocyte-like cells. Thus, preliminary data indicated the PC method to be partially successful inducing reprogramming and PCi-differentiation. However, epigenetic regulation and other underlying mechanisms were not clearly understood yet. Therefore, this study aimed to investigate:

- using a PC-based method, HepaRG cells were subjected to reprogramming into multipotent cells, then multipotent cells were subjected to PCi-differentiation to hepatocyte-liked cells

- how DNA methylation and genome-wide expression were altered during reprogramming and PCi-differentiation

- whether any epigenetic mechanisms contributed to reprogramming and PCi-differentiation

Furthermore, this study also attempted to establish a controlled confinement (CC) experimental procedure to mimic PC-based approaches in a more reproducible and precisely controllable manner. Based on preliminary studies and comparable results between PC and CC-based

approaches, alterations in DNA methylation and genome-wide expression of CC vs. PC experiments were examined and compared to shed light on the pattern of molecular signatures during reprogramming and PCi-differentiation.

Chapter 2 Materials and methods

2.1. Materials

2.1.1. Chemicals and reagents

Table 2.1 Chemicals, reagents, and sources of suppliers

Chemical	Source of suppliers
Adenosine triphosphate (ATP)	NEB, Frankfurt a.M., Germany
Agarose	Biozym Scientific GmbH, Oldendorf, Germany
Agencourt AMPure XP beads	Beckman Coulter GmbH, Krefeld, Germany
Aqua double distilled (milli-Q)	In house Millipore system
Boric acid	Carl Roth GmbH und Co. KG, Karlsruhe, Germany
Bromophenol blue	Sigma-Aldrich, St. Louis, USA
Chloroform	VWR International GmbH, Darmstadt, Germany
Ethanol (absolute)	Sigma-Aldrich, St. Louis, USA
Ethidium bromide	Carl Roth GmbH und Co. KG, Karlsruhe, Germany
Ethylenediaminetetraacetic acid (EDTA)	Carl Roth GmbH und Co. KG, Karlsruhe, Germany
Magnesium chloride ($MgCl_2$)	Carl Roth GmbH und Co. KG, Karlsruhe, Germany
Nonident P-40 (NP40)	Carl Roth GmbH und Co. KG, Karlsruhe, Germany
Pefabloc	Sigma-Aldrich, Taufkirchen, Germany
Potassium chloride (KCl)	Merck, Darmstadt, Germany
Sodium acetate (NaAc)	Carl Roth GmbH und Co. KG, Karlsruhe, Germany
Sodium chloride (NaCl)	VWR International GmbH, Darmstadt, Germany
Sodium dodecyl sulfate (SDS)	AppliChem, Darmstadt, Germany
Tris hydrochloride (Tris-HCl)	Carl Roth GmbH und Co. KG, Karlsruhe, Germany
Xylen-Cyanol	Sigma-Aldrich, Taufkirchen, Germany
100 bp DNA Ladder Ready to Load	Solis BioDyne, Tartu, Estonia

<2.1.2. <u>Media, buffers, and solutions</u>

<u>Media and cell culture solutions</u>

Table 2.2 Media, cell culture solutions, and sources of suppliers

Chemical	Source of suppliers
William's E Medium, GlutaMAX™	Life Technologies, Darmstadt, Germany
Hydrocortisone (50µM)	Sigma-Aldrich, Taufkirchen, Germany
Fetal calf serum (FCS)	PAA Laboratories, Pasching, Austria
Dimethyl Sulfoxide (DMSO)	Sigma-Aldrich, Taufkirchen, Germany
Trypsin-EDTA (0.05%)	Life Technologies, Darmstadt, Germany

<u>Buffer and solution formulars</u>

<u>10X TAE</u>

400 mM Tris base pH8

200 mM Acetic acid

0.5 mM EDTA pH 8

<u>10 x TBE</u>

890 mM Tris base

890 mM Boric acid

20 mM EDTA pH 8

<u>6 x Loading buffer</u>

0.2 % (w/v) Bromophenol blue

0.2 % (v/v) Xylen-Cyanol

40 % (w/v) Sucrose

<u>10 x TE</u>

100 mM Tris base pH 8

10 mM EDTA pH 8

<u>Ethidium bromide bath solution</u>

5 µg Ethidium bromide / ml TBE

Enzyme buffers and PCR reagents

Table 2.3 Enzyme buffers, PCR reagents, and sources of suppliers

Buffer	Source of suppliers
Cutsmart (10x)	NEB, Frankfurt a.M., Germany
HotStar PCR Buffer (10x)	Qiagen, Hilden, Germany
Deoxyribonucleotides (dA/T/C/GTP)	Solis BioDyne, Tartu, Estonia
Hot Start-IT Binding Protein	Affymetrix, Cleveland, USA
PCR Buffer BD (10x)	Solis BioDyne, Tartu, Estonia
5x RT Buffer	Solis BioDyne, Tartu, Estonia
5x HOT FIREPol® EvaGreen® qPCR Mix	Solis BioDyne, Tartu, Estonia

2.1.3. Enzymes

Table 2.4 Enzyme and sources of suppliers

Enzyme	Source of suppliers
HOT FIREPol®DNA Polymerase (5 U/µl)	Solis BioDyne, Tartu, Estonia
HotStar Taq™DNA Polymerase (5 U/µl)	Qiagen, Hilden, Germany
MspI (20,000 U/ml)	NEB, Frankfurt a.M., Germany
Klenow Fragment (3' to 5' exo-; 5 U/µl)	NEB, Frankfurt a.M., Germany
Proteinase K (20 mg/ml)	Sigma-Aldrich, Taufkirchen, Germany
T4 DNA Ligase (400 U/µl)	NEB, Frankfurt a.M., Germany
Reverse transcriptase (200 U/µl)	Sigma-Aldrich, Taufkirchen, Germany
RNasin® Ribonuclease Inhibitors (40 U/µl)	Promega, Walldorf, Germany
Random hexaprimer	Sigma-Aldrich, Taufkirchen, Germany

2.1.4. Oligonucleotides (all adapters and primers)

Miseq adapters for local-deep sequencing (Ampicon sequencing)

Universal adapters were fused 5´ to amplicon primer sequences, as shown in Table 2.5

(adaptors) and Table 2.8 (amplicon primers).

Table 2.5 Sequences of universal adaptors for amplicon sequencing

Primer	Sequence (5' to 3')
PCR-F Fusion primer	TCTTTCCCTACACGACGCTCTTCCGATCT_AmpliconPrimer
PCR-R Fusion primer	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT_AmpliconPrimer

Truseq adapters for Hiseq (Library sequencing)

Truseq primer sequences with adaptor index (*NNNNNN*) were shown in Table 2.6, whereas index sequences were shown in

Table 2.7.

Table 2.6 Sequences of Truseq primer sequences for library sequencing

Primer	Sequence (5' to 3')
PCR-F with TruSeq Adaptor	CAAGCAAGACGGCATACGAGAT*NNNNNN*GTGACTGGAGTTCAGACGTGTGCTGCTCTTCCGATCT
PCR-R Truseq DNA Methyl Reverse	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT

Table 2.7 Truseq adaptor index sequences for library sequencing

PrimerID	Index sequence (5' to 3')
TruSeq_AD002	ACATCG
TruSeq_AD004	TGGTCA
TruSeq_AD005	CACTGT
TruSeq_AD006	ATTGGC
TruSeq_AD007	GATCTG

PrimerID	Index sequence (5' to 3')
TruSeq_AD012	TGGTCA
TruSeq_AD013	CACTGT
TruSeq_AD014	ATTGGC
TruSeq_AD015	GATCTG
TruSeq_AD018	GCGGAC
TruSeq_AD019	TTTCAC

Amplicon primers

Table 2.8 Amplicon primer sequences. Annealing temperature (T_a), number of CpG site and product length are also provided.

Gene Location (hg19)	Strand	(Sequence 5' to 3')	T_a (ºC)	CpG (site)	Length (bp)
AMIGO2 Chr12:47474651-47474830	F	GTTATTTTTTCGTTTTTTGTAAG	54	16	179
	R	ATAATCTCCACCTCATTTCTAAA			
CSMD3 Chr8:114421443-114421740	F	TTGGGATTATAGGTGTATATTAG	54	7	297
	R	AACATACCATTATAAAATTTAC			
DTD1 Chr20:18569004-18569163	F	GGTAGAGGGTGTAGGATTAG	58	10	159
	R	CAACAATAAAAAAAAAAAAATACCTAC			
DUS3L Chr19:5789973-5790340	F	GGAGTTTTTTAATAATTATTCGTTG	58	17	357
	R	AAACAATTTCACCAATTCCT			
FA2H Chr16:74808013-74808501	F	ATAAGTGAAGAGGAAGGAAAGT	58	30	488
	R	AACCAAAACRAACCAAAACATCAA			
FAM5C Chr1:190447233-190447582	F	GTAGTAAGGAAGTTTAATATAAT	54	21	295
	R	CTTAAATCACAAACAACAAAAAC			
OCRL ChrX:128674242-128674486	F	TGGGTTAGATTTTTAGTTTTTAGTT	54	29	244
	R	ACCTCCCCTCTCCCTTCTCT			
SPARC Chr5:151043575-151043843	F	GGGGGATGATAATTTAGTATTTT	54	7	268
	R	TAAAACTTTTCCATACCTCA			

Gene / Location (hg19)	Strand	(Sequence 5' to 3')	T_a (°C)	CpG (site)	Length (bp)
TF / Chr3:133464928-133465037	F	TGGAGTTAGGAGTAGAGTTTT	54	8	110
	R	TATTAATAAAAAAAAACATAAAACAC			
TOM1L2 / Chr17:17806802-17806960	F	TTGTAGTTTTAGTTATTTAGGAGGTT	58	7	158
	R	TCCATCTATTTTACCCAACATTTTT			
TUBA1A / Chr12:49582325-49582693	F	GTTTTTGTTTTTTTTGGGTGTTT	54	32	368
	R	AAACCTTTAAACTCTTTTCTTTC			
ZNF814 / Chr19:58399924-58400016	F	AGTTTATTTTTAGGTTTTTGT	54	15	192
	R	ATATCCTCAAATCACCTCATCATC			

qPCR primers

Table 2.9 qPCR primer sequences. NCBI accession number and product length are also provided.

Gene	NCBI Accession No.	Strand	(Sequence 5' to 3')	Length (bp)
ALB	NM_000477.7	F	CCTGTTGCCAAAGCTCGATG	140
		R	GAAATCTCTGGCTCAGGCGA	
CDK1	NM_001786.5	F	GGATTTTCAGAGCTTTGGGCA	192
		R	TGCCAGAAATTCGTTTGGCTG	
ETNK2	NM_018208.4	F	ACCGGGAGAATGAGGTCAGA	161
		R	GGCGATTAACCTGAAAAGCCG	
LGALS1	NM_002305.4	F	CTGGAAGTGTTGCAGAGGTGT	130
		R	CCGTCAGCTGCCATGTAGTT	
SPARC	NM_003118.4	F	CATTGACGGGTACCTCTCCC	102
		R	TCCAGGTCACAGGTCTCGAA	
TF	NM_001063.4	F	CTGCTTTGCCTGGACAACAC	149
		R	AAATGTTCCTGGGCCTGGTT	
TOP2A	NM_001067.4	F	GAATCGGTACTGGGTGGTCC	115
		R	TTGTAACTTGGAAGCATTGGCA	

Gene	NCBI Accession No.	Strand	(Sequence 5' to 3')	Length (bp)
ZNF330	NM_014487.6	F	CATGCAATCGGCTTGGTCAG	172
		R	TTCAGGGAGCGTGTTGACAT	

2.1.5. Molecular weight standard

100 bp DNA Ladder Ready to Load (Solis BioDyne) was used for gel electrophoresis as a size standard.

2.1.6. Reaction kits

Table 2.10 Reaction kits and sources of suppliers

Reaction kit	Source of suppliers
Bioanalyzer High Sensitivity DNA Kit	Invitrogen™, Karlsruhe, Germany
EZ DNA Methylation Kit Gold	Zymo Research, Irvine, USA
Qubit® ds DNA HS Assay Kit	Invitrogen™, Karlsruhe, Germany
MiSeq Reagent Kit v3	Illumina, Inc., San Diego, USA
TruSeq DNA Methylation Kit	Illumina, Inc., San Diego, USA
TruSeq PE Cluster Kit v3-cBot-HS	Illumina, Inc., San Diego, USA
TruSeq SBS Kit v3-HS	Illumina, Inc., San Diego, USA
Nextera XT DNA Library Preparation Kit	Illumina, Inc., San Diego, USA

2.1.7. Machines and devices

Table 2.11 Machines, devices, and sources of suppliers

Machine and device	Source of suppliers
MiSeq™	Illumina, Inc., San Diego, USA
HiSeq™ 2500	Illumina, Inc., San Diego, USA
Agilent 2100 Bioanalyzer	Agilent Technologies, Santa Clara, USA
cBot System	Illumina, Inc., San Diego, USA
Qubit ® 2.0	Invitrogen™, Karlsruhe, Germany

2.1.8. Databases and software

Databases

NCBI   http://www.ncbi.nlm.nih.gov

UCSC Genome Browser   http://genome.ucsc.edu

Software

bamUtil   http://genome.sph.umich.edu/wiki/BamUtil

BiQ Analyzer HT   http://biq-analyzer-ht.bioinf.mpi-inf.mpg.de

BisSNP   http://sourceforge.net/projects/bissnp/files

ChIPseek   http://chipseek.cgu.edu.tw

Fast QC (Version 0.11.2)   http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc

Galaxy   http://internal.genetik.uni-sb.de/galaxy

IGV Browser   http://broadinstitute.org/igv

ImageJ   https://imagej.nih.gov/ij

MACS2 (2.1.0.20140616)   https://pypi.python.org/pypi/MACS2

Pathview   https://pathview.uncc.edu

Picard tools   http://broadinstitute.github.io/picard

Primer-BLAST   https://www.ncbi.nlm.nih.gov/tools/primer-blast

Primer Design and Search Tool   http://bisearch.enzim.hu

The R Project for Statistical Computing   http://www.r-project.org

Trim Galore   http://www.bioinformatics.babraham.ac.uk/projects/trim_galore

2.2. Methods

2.2.1. Overall experimental procedure

Physical constraint experiment

Regarding the precursor of reprogramming experiment, HepaRG was cultured in a basic medium with 0.5 µM hydrocortisone (passage 0). After reaching a confluent level, cells were sub-cultured to two groups and hydrocortisone was removed. At this point, one group was in the culture as a control of cells without physical constraint (PC), while another group was subjected to physical constraint for 20 hours to introduce cell reprogramming. The PC cells were again sub-cultured to two subgroups, one supplemented with 10 µM 5'-Azacytidine treatment and one without Aza treatment. All groups of cells were prolonged to passage 10 in the same condition and harvested.

In parallel, cells without physical constraint were treated with Aza and used as a control of the PCi-differentiation experiment. After removal of Aza, differentiation of cells was induced by physical constraint, followed by cultivation in basic medium, together with 50 µM hydrocortisone and 2% DMSO treatment. The cultivation was prolonged and harvested the same way as for the induced reprogramming cells.

All experiments were performed under the supervision of Dr. Christiane Guillouzo at the Biopredic company. The cell pellets and extracted RNA were delivered to Saarland University for further analysis.

<u>Controlled confinement experiment</u>

Precursor cells of differentiated HepaRG were cultivated in William's E medium. After reaching confluence, cells were sub-cultured to multiwell confiner, which consists of fibronectin coated, and uncoated 6-welled plates. Approximately, $2x10^5$ cells were seeded in each well. Cells in those plates were cultivated in William's E medium without hydrocortisone for further use.

For controlled confinement experiment, cells were then incubated for 6 hours. To mimic the conditions in physical constraint experiments, confined roofs (see Section 2.2.9) were applied to multiwell confiner for 20 hours. In addition, an uncoated plate without confined roof was used as a control of this experiment. After 20 hours, cell pellets were collected followed by nucleic acid extraction (DNA and RNA) to prepare sequencing libraries. This experiment was performed in biological duplicate.

2.2.2. <u>Sequencing library preparation</u>

<u>Reduced representative bisulfite sequencing (RRBS) library</u>

Regarding the reprogramming and PCi-differentiation experiments, 12 samples covering all processes were selected to analyze the DNA methylation status. DNA of selected samples were extracted, then dissolved in 1xTE, followed by quality and quantity measurement using Nanodrop photometer.

The Reduced representative bisulfite sequencing (RRBS) technique captures CpG dinucleotides throughout the genome at single-base resolution, according to restriction enzyme digestion. In

this study, *MspI*, which restrictively cuts to 5' C^CGG 3', was used to generate the RRBS library.

Each step of the RRBS library preparation procedure was shown in Table 2.12.

Table 2.12 RRBS library preparation. Reaction mixtures and incubation/cycle conditions for each step is provided.

Reaction mixture (1x)	Vol. (μl)	Incubation/cycle condition
• **Digestion (adjusted reaction vol. by H$_2$O to 30 μl.)**		
Genomic DNA	100 ng	37 ^{o}C 16-18 hours
10x Cutsmart buffer	3	After 2 hours, add 0.5 μl *MspI* to boost up the reaction.
MspI	1	
• **End-repair and A-tailing**		
10 mM dATP and 1mM dCTP+dGTP	1	30 ^{o}C 30 min, 37 ^{o}C 20 min, and
Klenow exo- (5U/μl)	1	75 ^{o}C 20 min
• **Adaptor ligation**		
(1:10) Truseq adaptor (see Table 2.7)	2	16 ^{o}C 22-24 hours and 65 ^{o}C 20 min
10 mM ATP	4	
T4-Ligease (2000 U/μl)	1	
10x Cutsmart buffer	1	
• **Bisulfite conversion (see Section 2.2.3)**		
• **Library amplification**		
Bisulfite samples	19	95 ^{o}C 15 min
10x Hotstar PCR buffer	2.5	15 cycles of (95 ^{o}C 30 sec, 60 ^{o}C 30 sec, and 72 ^{o}C 1 min)
2.5 mM each dNTPs	2.5	
10 μM forward primer (see Table 2.6)	0.25	72 ^{o}C 7 min
10 μM reverse primer (see Table 2.6)	0.25	
HotStart Taq polymerase (5 U/μl)	0.5	

After library amplification, size selection and purification by AMPure XP beads were performed (see Section 2.2.5). Finally, the libraries were eluted in 11 µl of 0.1xTE. The RRBS libraries were quantified by Qubit Fluorometer (DNA High sensitivity kit) and fragment size was checked using the Agilent Bioanalyzer DNA High Sensitivity chip on the Agilent 2100 Bioanalyzer.

Smart-Seq2 messenger RNA sequencing library

While the DNA methylation data of selected samples were obtained by RRBS, the transcriptome data of some selected samples were obtained by mRNA-sequencing using smart-seq2 protocol. The smart-seq2 protocol is a method that can generate quantitative transcriptome data beyond the limitation of small amount of RNA material. In this study, the RNA samples were provided by Dr. Christiane Guillouzo. To prepare mRNA library, each step was shown following Table 2.13.

In the beginning of library preparation, 1 ng RNA (in 2 µl) was mixed with 1µl of 10 mM oligo-dT primer and 1 µl of 10 mM dNTP and denatured at 72 °C for 3 min. After denaturation, premixed RNA sample was placed on ice immediately. RT reaction and library preparation were further performed after premixed RNA was prepared.

Table 2.13 mRNA-seq library preparation. Reaction mixtures and incubation/cycle conditions for each step is provided.

Reaction mixture (1x)	Vol. (µl)	Incubation/cycle condition
• **Reverse transcription (RT)**		
Premixed RNA	4	42 °C 90 min
SuperScript II RT (200 U/µl)	0.5	10 cycles of (50 °C 2 min and 42 °C 2 min)
RNA Inhibitor (40 U/µl)	0.25	70 °C 15 min and hold at 4 °C
5x SuperScript II first-strand buffer	2	
100 mM DTT	0.5	
5M Betaine	2	
0.5M $MgCl_2$	0.12	
100 µM TSO	0.1	
Nuclease-free water	0.23	
• **PCR pre-amplification**		
2x KAPA HiFi HotStart Ready mix	12.5	98 °C 3 min
10mM IS PCR primers	0.25	18 cycles of (98 °C 20 sec, 67 °C 15 sec, and 72 °C 6 min)
Nuclease-free water	2.25	72 °C 5 min and hold at 4 °C
RT sample	9.7	
• **PCR purification and quality control (see Section 2.2.5)**		
• **Tagmentation (Scaled volume to 4 µl)**		
2x TD buffer	2	55 °C 5 min and hold at 4 °C, following by an addition of NT buffer
Amplicon tagment mix	1	
cDNA (500 pg)	1	
NT buffer	1	5 min at room temperature
• **Amplification of adaptor-ligated fragments (Scaled volume to 10 µl)**		
Tagmented DNA samples	5	72 °C 3 min and 95 °C 30 sec
NPM	3	10 cycles of (95 °C 10 sec, 55 °C 30 sec, and 72 °C 30 sec)
5 µM AD1 primer	1	72 °C 5 min and hold at 4 °C
5 µM AD2 primer	1	

Similar to RRBS library, size selection and purification by AMPure XP bead were performed after amplification of adaptor-ligated fragments (see Section 2.2.5). The libraries were then eluted in 10 μl of elution buffer (or nuclease-free water). The mRNA-sequencing libraries were quantified by Qubit Fluorometer (DNA High sensitivity kit) and fragment size was checked using the Agilent Bioanalyzer DNA High Sensitivity chip on the Agilent 2100 Bioanalyzer.

2.2.3. Bisulfite treatment of DNA

Bisulfite treatment is a method to distinguish methylated cytosines from unmethylated cytosines. First, double stranded DNAs are denatured to single stranded DNAs. Those single stranded DNAs are subjected to sodium bisulfite treatment. The reaction turns unmethylated cytosine to be sulphonated and deaminated intermediate product. After desulphonation by the addition of NaOH into the reaction, intermediate product is converted to uracil. Bisulfite-treated DNAs are then subject of bisulfite-specific PCRs. In the PCR reaction, uracil will be translated to thymines, while methylated cytosines will still be amplified as cytosines.

In this study, bisulfite treatment was performed using the EZ-DNA methylation Kit/ Gold Kit (Zymo Research) following the manufacturer's instructions. For local-deep sequencing (Miseq for PCR amplicons), bisulfite DNA was eluted in 30 μl elution buffers. For library preparation with the TruSeq DNA Methylation Kit (see Section 2.2.2), bisulfite DNA was eluted in 2x12 μl MilliQ water.

2.2.4. Polymerase Chain Reaction (PCR)

Bisulfite-specific PCR

Polymerase chain reaction (PCR) is a reaction that mimic DNA replication processes, but it is performed *in vitro* to amplify target sequences. PCR steps include denaturation of DNA templates, annealing of primers and elongation by DNA polymerase enzymes. In this study, reaction mixtures were prepared as described in Table 2.14 (left). PCR reaction was performed in the AB 2720 or

Veriti thermocycling machines (Applied Biosystems). Cycling conditions were provided in Table 2.14 (right).

Table 2.14 Reaction mixture for 1x PCR (left) and PCR cycle condition (right).

Reaction mixture (1x)	Vol.(µl)
Bisulfite DNA	3
10x buffer BD	3
25 mM MgCl$_2$	3
10mM dNTPs (2.5mM each)	2.4
10 µM forward primer	0.5
10 µM reverse primer	0.5
HotFire polymerase (5 U/µl)	0.5
Milli Q water	Adjust to 30

Segment	Condition	
Initial denaturation	95 ^{o}C, 15 mins	1 cycle
Denaturation	95 ^{o}C, 1 min	Repeat for 40 cycles
Annealing*	X ^{o}C, 2 mins	
Elongation	72 ^{o}C, 1 min	
Final elongation	72 ^{o}C, 7 mins	1 cycle
Cooling	4-10 ^{o}C, ∞	

*Annealing temperatures (T$_a$) for each amplicon were provided in Table 2.8.

Quantitative Reverse Transcription (RT-PCR)

In this study, 1 µg RNA, 1 µl random hexamers, and nuclease-free water were adjusted to a total volume of 12 µl. Then, the mixture was incubated at 65° C for 5 mins and chilled on ice. After incubation, RT reaction mix (Table 2.15) was added and incubated in a cycler at 25 ^{o}C for 5 mins, followed by 42 ^{o}C for 60 mins. Finally, the reaction was terminated at 70 ^{o}C for 5 mins.

Table 2.15 Reaction mixture for 1x RT reaction mixture.

Reaction mixture (1x)	Vol. (20 µl)
RNA mixture	12
5x RT buffer	4
10 mM dNTPs mix	2
RNase inhibitor (40 U/µl)	1
Reverse transcriptase (200 U/µl)	1

Prepared cDNA was used as a template in quantitative PCR (real-time qPCR). This step includes fluorescent dye (EvaGreen) to be incorporated into the nascent PCR product., Gene expression level was detected by Ct value during the annealing step. qPCR reaction was prepared as listed in Table 2.16 (left), qPCR conditions are shown in Table 2.16 (right). qPCR reaction was performed in Multiplex Quantitative PCR Systems (Mx3000P machine, Stratagene).

Table 2.16 Reaction mixture for 1x qPCR (left) and PCR cycle conditions (right).

Reaction mixture (1x)	Vol.(µl)
cDNA	3
EvaGreen® qPCR Mix	4
10 µM forward primer	0.3
10 µM reverse primer	0.3
Milli Q water	Adjust to 20

Segment	Condition	
Initial Denaturation	95° C, 15 mins	1 cycle
Denaturation	95° C, 30 sec	Repeat for 40 cycles
Annealing	60° C, 20 sec	
Elongation	72° C, 30 sec	
Final detection	95° C, 1 min	1 cycle
	60° C, 30 sec	
	95° C, 30 sec	

2.2.5. DNA Purification, quantification, and quality control

Amplicon purification

Amplicons were purified using AMPure XP beads to get rid of other PCR components e.g. buffer, dNTPs and primers. PCR product was mixed with 1.1x magnetic beads and placed on magnetic rack. In case of RRBS library and mRNA library purification, the mixtures were mixed with 0.9x and 0.8x magnetic beads, respectively. Supernatant was discarded after 10 minutes, whereas PCR fragments remain bound to magnetic beads. The beads were washed with 80% ethanol twice and dried. Dried pellets were then dissolved and eluted using 0.1x TE buffer. After placing back on the magnetic rack to separate magnetic beads, the solution was transferred to a new tube. Quality control was performed by gel electrophoresis (1.2% agarose gel in 0.5x TBE).

Amplicons were quantified using Qubit dsDNA HS assay kit on a Qubit fluorometer. Amplicons can be stored at 4°C or -20°C until sequencing.

2.2.6. Next Generation Sequencing

Local-deep sequencing on the Illumina Miseq platform

Sequencing of bisulfite amplicons was performed using Illumina Miseq platform. Amplicons were generated with primers containing adapter sequences as shown in Table 2.5 and Table 2.8. On Illumina Miseq platform, sequencing by synthesis technology was used. The amplicon samples were bound to fixed oligos on a flow cell, where bridge amplification was performed. Reversible terminator nucleotides (dA/T/C/GTPs) labelled with different fluorescent dyes were used in the elongation step, leading to fluorescent detection of each base during every round of DNA synthesis. In this study, the samples were sequenced 2 x 300 bp paired-end, aiming at 10,000 reads per amplicon. This work was carried out by Jasmin Kirch, Christina Lo Porto and Dr. Sascha Tierling.

RRBS and mRNA libraries sequencing on the Illumina Hiseq2500 platform

The qPCR was performed to measure amplifiable library concentration and adjusted to 2 nM. For the cluster generation, the cBot system was used following the manufacturer´s instructions. After cluster generation, all libraries were sequenced 1 x 100 bp single read on the Illumina HiSeq 2500 system. This work was carried out by Christina Lo Porto, Dr. Nina Gasparoni and Dr. Gilles Gasparoni.

2.2.7. Analysis of sequencing data

Bisulfite amplicons

Initial quality control and adaptor trimming were computationally processed by Dr. Karl Nordström and Dr. Abdulrahman Salhab. Amplicon reads were further extracted by primer sequence and

sample index using the Galaxy Organizer tool for Illumina data. Finally, amplicon reads were analyzed for methylation level using BiQAnalyzer HT (Lutsik *et al.*, 2011). Parameters on BiQAnalyzer HT were set as minimal sequence identity = 0.9, minimal bisulfite conversion = 0.95 and maximal fraction of unrecognized CpG sites = 0.1, respectively. The methylation of each amplicon and pattern maps were provided as results from BiQAnalyzer HT.

<u>RRBS and mRNA sequencing</u>

Comparable to local-deep sequencing, raw sequenced reads were computationally processed including initial quality control (FastQC), adapter trimming (Trim Galore! wrapper of cutadapt), mapping to GRCh37/hg19 reference genome (GSNAP), DNA methylation calling (BisSNP), removing duplicates (Picard tools) and elimination of overlapping read pairs (bamUtil). Computational processes were carried out by Dr. Karl Nordström and Dr. Abdulrahman Salhab.

Regarding RRBS sequencing, processed reads were provided as bed and bam files. Those file types were loaded into the IGV browser for visualization. Further analysis was performed using MethylKit. 10x coverage was used as a filtering threshold. Tiling regions were determined as 500 bp windows including at least 3 CpG sites. In case of duplicates, only tiling regions that are found in duplicated samples were considered. Differential methylation was detected using f-test (no replicate) or student´s t-test (duplicate) to obtain statistical significance.

mRNA sequencing results were further analyzed using RNA-SeQC. Bigwig files were provided for visualization as well as FPKM of each transcript was also provided. Differential expression was performed using Cuffdiff (no replicate) or DeSeq2 (duplicate) to obtain statistical significance. This work was carried out by Dr. Karl Nordström.

2.2.8. Gene selection for validation experiments and Primer design

Bisulfite amplicons

Amplicons were selected for validation experiments from pairwise comparison of each state. Selected DMRs should show a differential methylation of at least 15% and an FDR-adjusted P-value less than 0.01. After region selection, nucleotide sequences were obtained from UCSC genome database (hg19). Bisulfite primer design was performed using the BiSearch online tool (Tusnády *et al.,* 2005; Arányi *et al.,* 2006) following these parameters:

- Primer should be 20 - 35 bp in length
- Amplicon should be 150 - 500 bp in length
- Primer melting temperature (Tm) should be at 48 – 62 $^{\circ}$C
- Primer concentration should be 0.167 pM
- Magnesium chloride concentration should be 2.5 mM
- Maximum Tm difference between the primers should be 2.0° C

To conclude primer design, ePCR was performed using BiSearch as a control for the specificity of those primers. Primers were manufactured by biomers.net GmbH, Ulm, Germany or Microsynth, Lindau, Germany. Primers were dissolved in 1xTE obtaining 100µM, as stock solution, and diluted in 0.1xTE to 10µM to obtain the working solution. Stock and working solution can be stored at -20° C until they are used in further steps.

Gene expression validation

Differential expression of genes was validated based on pairwise comparison of each state. Log2 fold change of selected DEGs should be at least 2 and be significant (q-value < 0.01). Accession number of genes were obtained from NCBI database (see Table 2.9). Primers for qPCR were designed using Primer-BLAST tools following these parameters:

- PCR product should be 70 - 200 bp in length

- Primer melting temperature (Tm) should be at 60 +/- 3° C

- At least one of each primer should be placed on an exon-exon junction

- Database should be Refseq mRNA, excluding predicted Refseq transcripts and uncultured/environmental sample sequences.

Similar to bisulfite primers, these primers were manufactured by biomers.net GmbH, Ulm, Germany or Microsynth, Lindau, Germany. Stock and working solution of primers were also prepared in the same way as described for bisulfite primers.

2.2.9. Glass slide fabrication for controlled confinement

Glass slide fabrication is a preparation step of multiwell confiner. In this study, glass slide fabrication had been modified (Le Berre *et al.,* 2014) and shown in Figure 2.1.
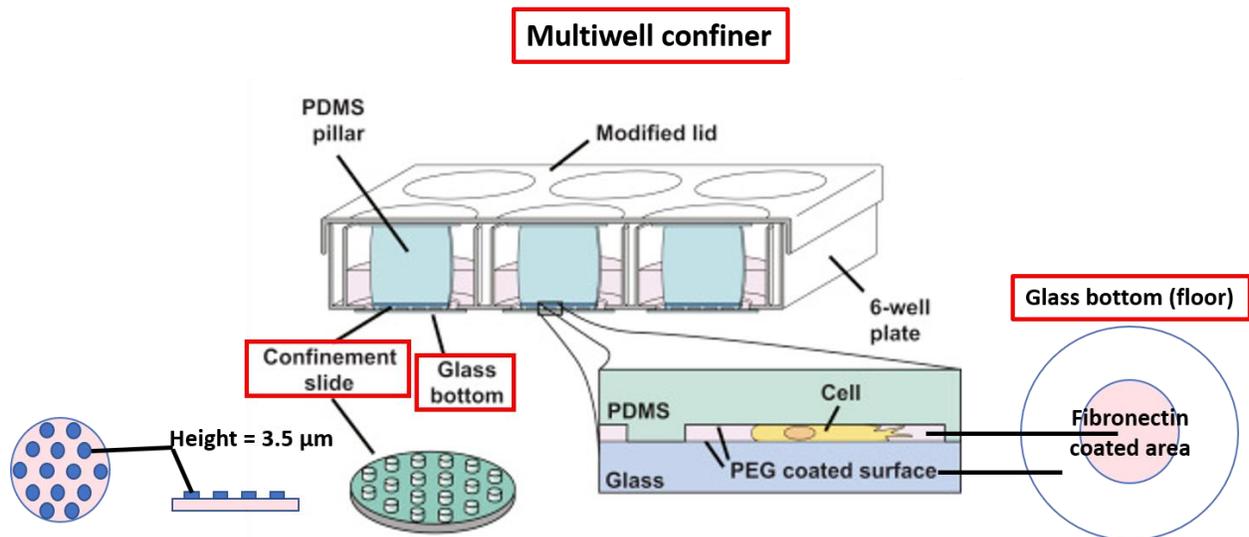
Figure 2.1 Assembled multiwell confiner (Modified, Liu et al., 2015). After floor glass treatment and slide fabrication, glass bottom slides are attached to the floor of 6-well plate by PDMS sheet, while confinement slides are attached to PDMS pillars, which are are attached directly to 6-well plate lid.

Floor glass treatment

A floor glass slide was placed in microplasma vacuum for 1 minute to activate surface. After that, parafilm sheet was used to cover a glass surface, except in the middle of the floor. Meanwhile, soft silicone was attached to the middle of the floor to prevent the middle area from coating. PEG

solution was then coated around the middle area under covered parafilm sheet and let it dry at room temperature overnight. After PEG coating, the middle area was marked on the uncoated side. Soft silicone was then removed from the glass, while parafilm sheet was not. In the middle area, fibronectin solution was used to coat in this area, since cells were attached in the middle of the floor during the experiment. After fibronectin coating, the floor glass was dried at room temperature for 6 hours. Then, the parafilm sheet was removed after drying step. Coated glass was washed using distilled water to remove excess of coating solution and removed excess water using an air jet. Coated floor glass was then stored at 4 °C until being assembled to a floor of mutiwell confiner.

Slide fabrication

A 10mm glass slide was placed in microplasma vacuum for 1 minute to activate surface, similar to the first step of the floor glass treatment. Meanwhile, 10 g of PDMS/cross-linker mix (8:1, w/w) was performed and dropped on an imprinted mold. To fabricate a confinement slide, an activated glass was placed on the mold by pressing activated surface onto PDMS drop. The mold was then incubated on a hot plate at 95° C for 15-20 minutes. After that, the PDMS rims were removed, retaining only the fabricated confinement slide. Detaching the fabricated confinement slide was performed using a razor blade and Isopropanol, since the slide needed to be kept wet during detachment step. The detached confinement slide was cleaned with Isopropanol and air dried. This fabricated confinement slide was stored at 4° C until assembled to a cover of mutiwell confiner as the confinement roof.

Multiwell confiner assembly

To assemble the multiwell confiner, sticky PDMS sheets were used to attach floor glasses and PDMS pillars to a 6-well plate and a lid, respectively. PDMS pillars were soft and flexible and had a similar height to the height of 6-well plate. Floor glasses were attached to each well of the 6-

well plate whereas confinement slides were attached to PDMS pillars. Then, the height of the pillars was slightly increased, providing a controlled pressure when this modified lid was applied in the confinement experiment.

2.2.10. Immunofluorescence for 5mC/5hmC

After 20 hours of confinement, cells were harvested by trypsinization and washed twice in 1x PBS. Cells were then fixed on the slide, including sucking out small amount of PBS, which are replaced with 3.7% paraformaldehyde in PBS, instead. After that, fixed cells were permeabilized, followed by staining with antibodies against DAPI, 5mC and 5hmC. For this immunofluorescent staining, all works were carried out by Dr. Konstantin Lepikhov. Fluorescent signal was measured using ImageJ and relative signals of 5mC and 5hmC were obtained by normalization a raw signal of 5mC or 5hmC by a raw signal of DAPI e.g. 5mC/DAPI or 5hmC/DAPI, respectively.

Chapter 3 Results

3.1. Physical constraint and PC-induced differentiation experiment

HepaRG cells were used for the following reprogramming experiment since they assemble molecular features of (pre-) differentiated hepatocytes on the one side and improved proliferation efficiency comparable to cancer cells on the other. To achieve the switch from a (pre-) differentiated cell to a multipotent cell, HepaRGs were physically restricted in their microenvironments (Physical constraint; PC). It was described previously that physical constraint affects chromosome structure, and this alteration was shown to have an impact on the epigenome and transcriptome of the cells (Huang *et al.*, 2015; Uhler and Shivashankar, 2016). Therefore, in this study, DNA methylation and expression profiles were studied through the reprogramming process, including physical constraint and maintenance state, in which cells were treated with 5-Azacytidine to keep multipotency. Moreover, the molecular signatures after PC-induced differentiation (PCi-differentiation or redirected differentiation), which was induced again by the physical constraint, followed by the application of hydrocortisone and DMSO to redifferentiate multipotent cells to hepatocyte-like cells, was also examined.

At the beginning of the experimental procedure, the experiment was conducted by Biopredic staffs under the supervision of Dr.Guillouzo. Preparing of HepaRG cells for the PC path, the precursor of differentiated HepaRG cells were cultured with hydrocortisone to push proliferation of the cells. After reaching confluency, the cells were divided into two sub-cultures, and hydrocortisone was removed from the medium. Cells growing in high density were subjected to physical constraint for 20 hours (RP_P0PC), while cells growing in low density were cultivated further without physical stress (RP_P0nonPC) (Figure 3.1).
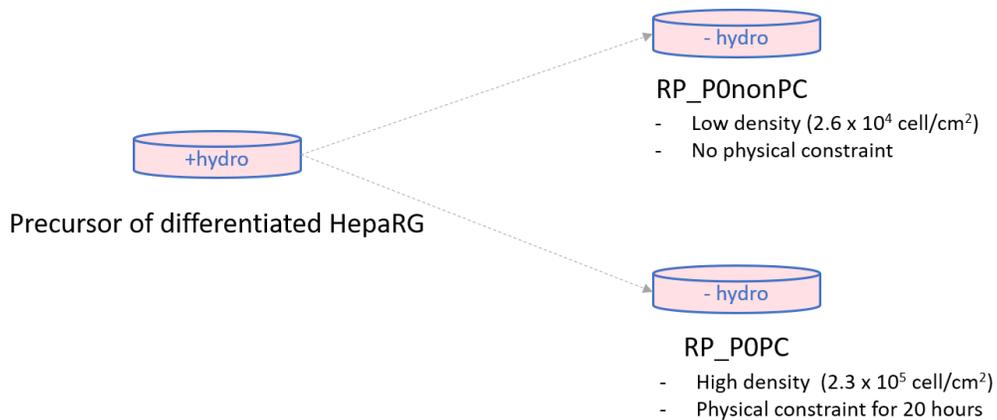
Figure 3.1 Experimental scheme of reprogramming HepaRG cells by physical constraint. Differentiated HepaRG precursor cells were divided into 2 sub-cultures, physical constraint was applied to the high-density culture while cells growing in low density were cultured without physical stress.

RP_P0PC/nonPC cells were further cultivated and prolonged to passage 10 (RP_P10nonPC).

The RP_P0PC were divided into two sub-cultures, i) treated with 10 μM 5-Azacytidine and ii) without Aza-treatment, and prolonged to passage 10 as well (RP_P10PC/PC+A) (Figure 3.2).
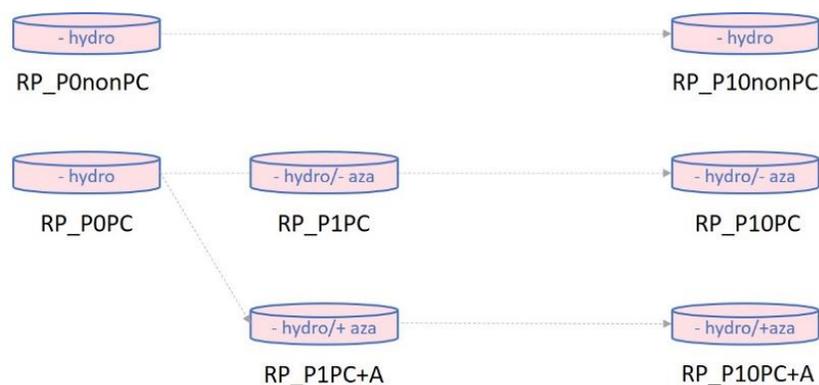
Figure 3.2 Experimental scheme of maintaining PC-HepaRG cells. RP_P0PC/nonPC, from physical constraint experiment, were further prolonged to passage 10. Cells were treated with 5-Azacytidine to stabilize PC-based reprogramming through the experiment.

For the PCi-differentiation path, the experiment was performed independently. Non-PC-HepaRG cells were cultivated with 5-Azacytidine (RD_P0nonPC+A). RD_P0nonPC+A were then subjected to physical constraint for 20 hours the same way as described for the RP_P0PC cells. After that, 5-Azacytidine was removed from the culture and cells were treated with 50μM hydrocortisone

(RD_P0+Hydro) and DMSO (RD_P0+DMSO) and cultivated to passage 20 (RD_P20), respectively (Figure 3.3).
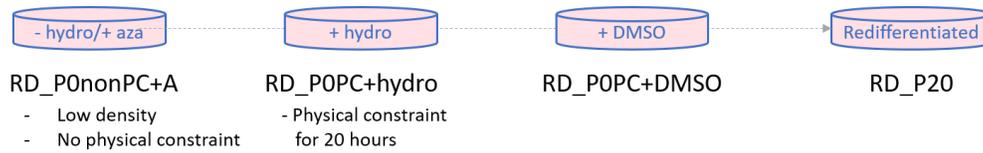
Figure 3.3 Schematic workflow showing PCi-differentiation strategy of reprogrammed HepaRG. RD_P0nonPC+A were subjected to physical constraint for 20 hours, followed by hydrocortisone/DMSO induction and prolonged cultivation to passage 20 (RD_P20).

During reprogramming and PCi-differentiation processes, cells were harvested at each time point and subjected to methylome and transcriptome analysis. To sum up the experimental process as a whole, Figure 3.4 illustrates the experimental scheme by timing.
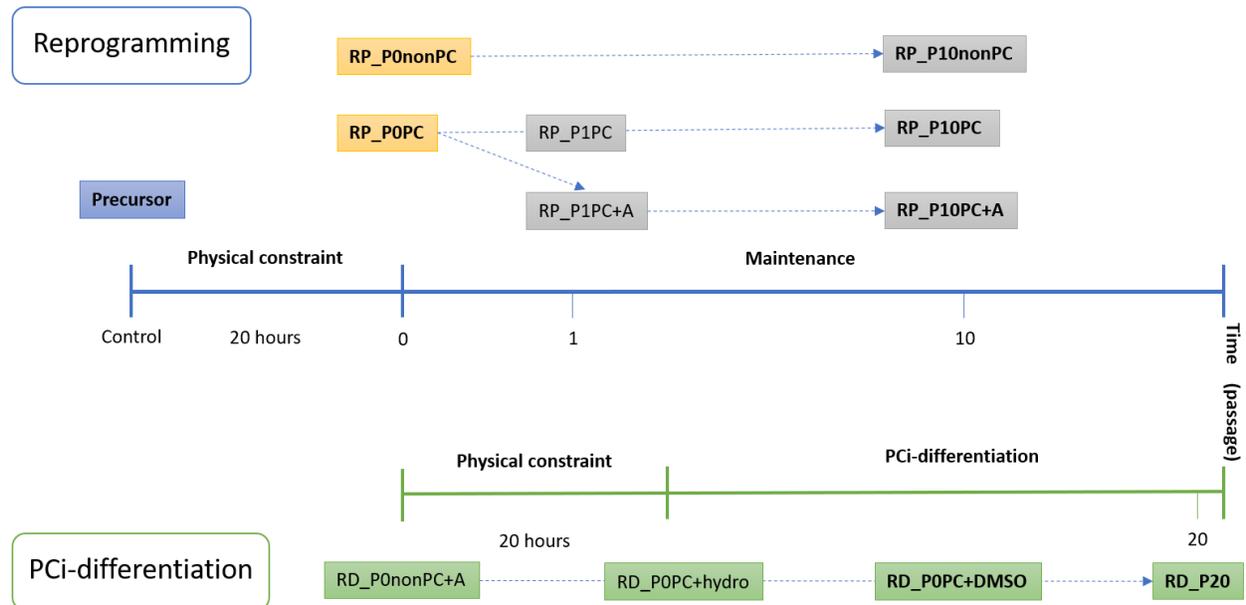
Figure 3.4 Schematic overview of the experimental workflow. Blue-dashed arrow represents time points of cultivation and prolonged culture. All samples are subjected to methylome sequencing, whereas samples with bold letters are subjected for transcriptome sequencing.; blue = precursor, orange = physical constraint, gray = maintenance, and green = PCi-differentiation.

3.1.1. DNA methylation changes during reprogramming and PCi-differentiation processes

The methylome data of this study were obtained by reduced-representation bisulfite sequencing (RRBS) on a HiSeq2500 machine (Illumina). In this study, the restriction enzyme *Msp*I was used to study promoters and other regulatory CpG-rich regions.
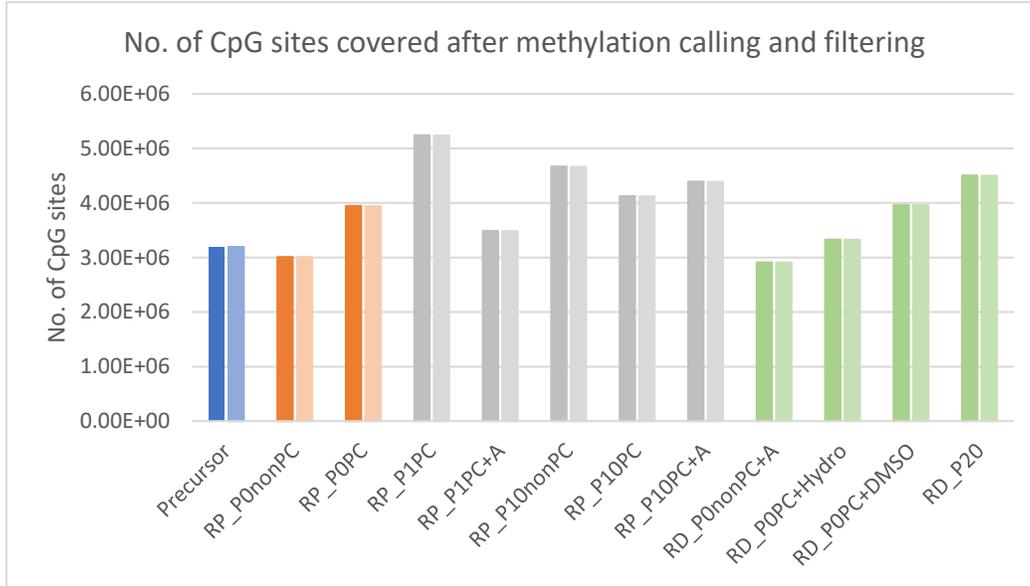
Figure 3.5 Number of CpG sites covered after methylation calling and filtering. Number of CpG sites after methylation calling are in dark colors, while number of CpG sites after filtering are in light color; blue = precursor, orange = physical constraint, gray = maintenance, and green = PCi-differentiation.

The sequencing yielded approx. 40-120 million reads, from each of the 12 samples (see Supplementary Table 7.1). Sequenced methylome data of all samples were analyzed by MethylKit (Akalin *et.al.*, 2012) setting a 10x coverage threshold. After methylation calling, for each sample 2-5 million CpG sites remained, depending on the number of sequenced reads. Even after 10x coverage filtering, more than 90% of reads were taken into account for analysis (Figure 3.5).

Overall methylation of CpGs of samples
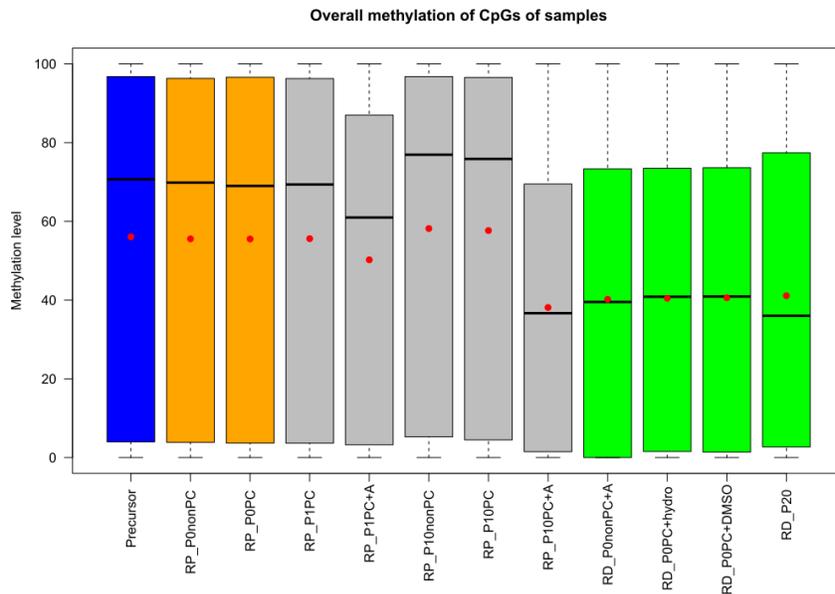
Sample	Avg.Met
Precursor	56.1
RP_P0nonPC	55.5
RP_P0PC	55.5
RP_P1PC	55.6
RP_P1PC+A	50.2
RP_P10nonPC	58.2
RP_P10PC	57.7
RP_P10PC+A	38.1
RD_P0nonPC+A	40.2
RD_P0+Hydro	40.5
RD_P0+DMSO	40.6
RD_P20	41.1

Figure 3.6 Overall methylation level of all samples. The black line in the box represents the methylation median, while the red dot indicates the average methylation level of the respective sample. Average methylation level (%) of each sample is listed in a table next to the box plot.

First, the average methylation level of all samples was assessed and found to be slightly different between the states (Figure 3.6). Between precursor cells and reprogrammed cells passage 0 (RP_P0 vs. PC/nonPC), the average methylation level was almost similar (56.1% and 55.5%, respectively). Notably, the average methylation level of samples treated with Aza (RP_P1 and RP_P10) was slightly decreased in passage 1 (from 55.6% to 50.2%), then considerably decreased in passage 10 (from 58.2% to 38.1%). For the PCi-differentiation path, the average methylation level of starter cells (RD_P0nonPC+A) and PCi-differentiation passage 0 and 20 (RD_P0 and RD_P20) were 15% lower than samples (40% – 41%) in the reprogramming and maintenance states. Although the samples in this group were treated with hydrocortisone and DMSO, the average methylation of those remained constant. However, the average methylation level tended towards a slight increase throughout the PCi-differentiation process.

58
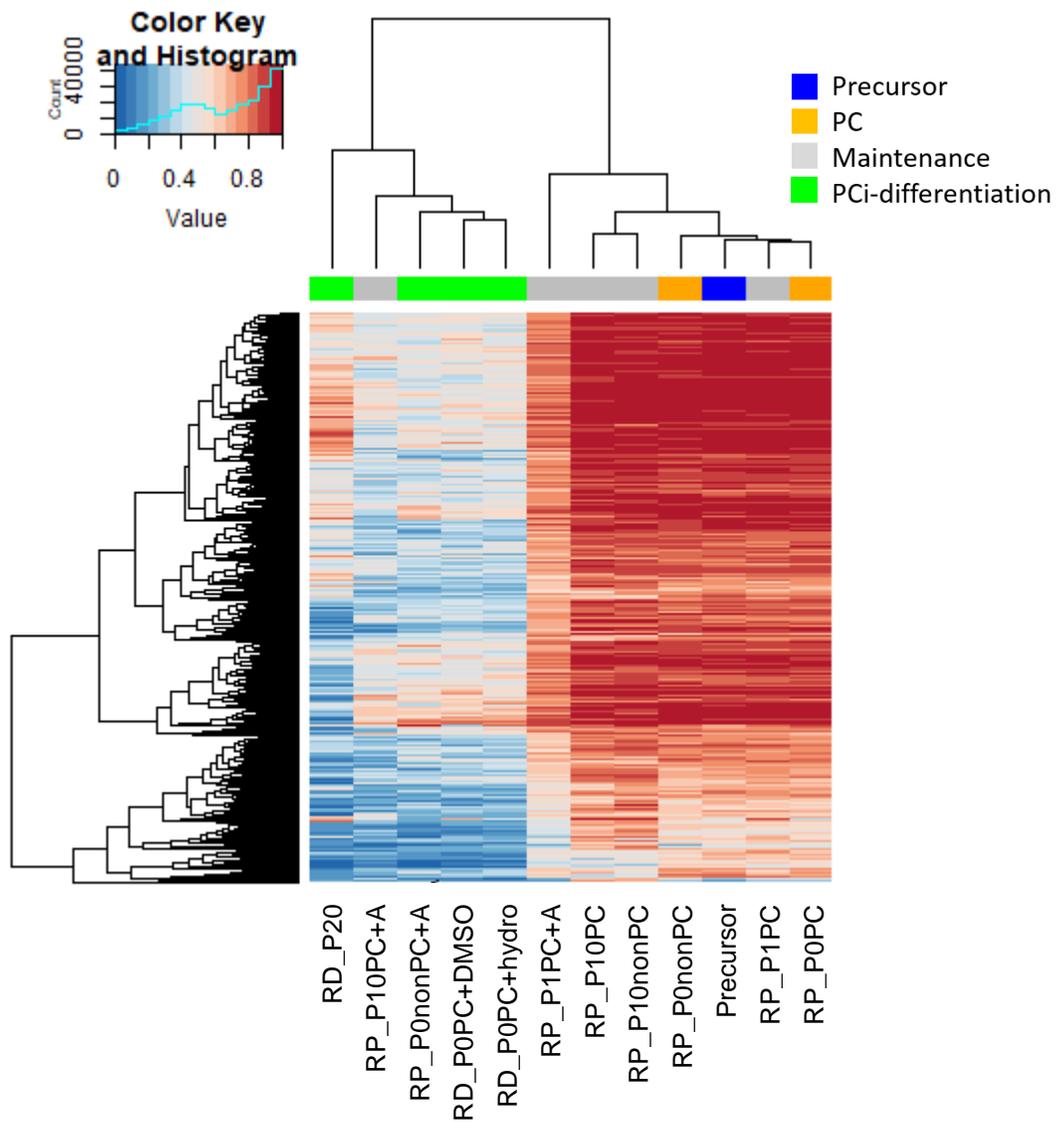
Figure 3.7 Heatmap of 20K most variable 500 bp tiles. The colors, blue and red, represent the methylation level from 0 to 1. The states of samples are defined by colors presented on the top right.

Using the MethylKit package, methylome data of all samples were assessed using a 500 bp tiling window containing at least 3 CpGs filtering all reads with less than 10-fold coverage. The heatmap of 20K most variable tiles showed significant differences between reprogramming and PCi-differentiation states (Figure 3.7). The precursor cells (blue) and samples belonging to reprogramming state (orange and grey) were in the same cluster showing highly methylated tiles with increasing methylation from early passage (P0) to late passage (P10), which were consistent

with the average methylation level of all CpG sites. During the reprogramming process, the 5-Azacytidine was introduced to the culture to stabilize the hypomethylated status of the reprogrammed cells, particularly in the RP_P10PC+A. Therefore, conversely to the other reprogramming samples, RP_P10PC+A clustered to the PCi-differentiation samples (green) with more low methylated tiles. Moreover, the methylation profile of the RD_P20 confirms the tendency towards an increase of methylation during the PCi-differentiation process.

3.1.2. <u>Comparison of differential methylation during reprogramming and PCi-differentiation processes</u>

In this study, the reprogramming process was divided into two parts, i) physical-constraint and ii) maintenance. For i) hydrocortisone was removed when the precursor cells reached confluency and these cells were sub-cultured for using in the subsequent PC experiment. Sub-cultured cells were seeded with different density and cultured in medium without hydrocortisone. The bottle of high-density cells was subjected to physical force (PC), while the bottle of low-density cells was used as a control (nonPC). By comparing the methylome and transcriptome of PC and nonPC physical constraint effects on molecular signatures can be studied. Studying cells before and after removal of hydrocortisone (+/-hydro), the effect of hydrocortisone on HepaRG molecular signatures as well as the effect of PC after 20 hours will become visible. For ii) PC and nonPC were prolonged passage. PC samples were also stabilized for cellular properties using 5-Azacytidine (Aza). Studying cells in maintenance state will also provide molecular signatures influenced by prolonged cultivation and Aza treatment, respectively.

For pairwise comparisons, the differentially methylated regions (DMRs) were analyzed using F-test in MethylKit package. In this study, the DMRs were defined as tiles with more than 10% DNA methylation difference and an FDR adjusted p-value less than 0.01.
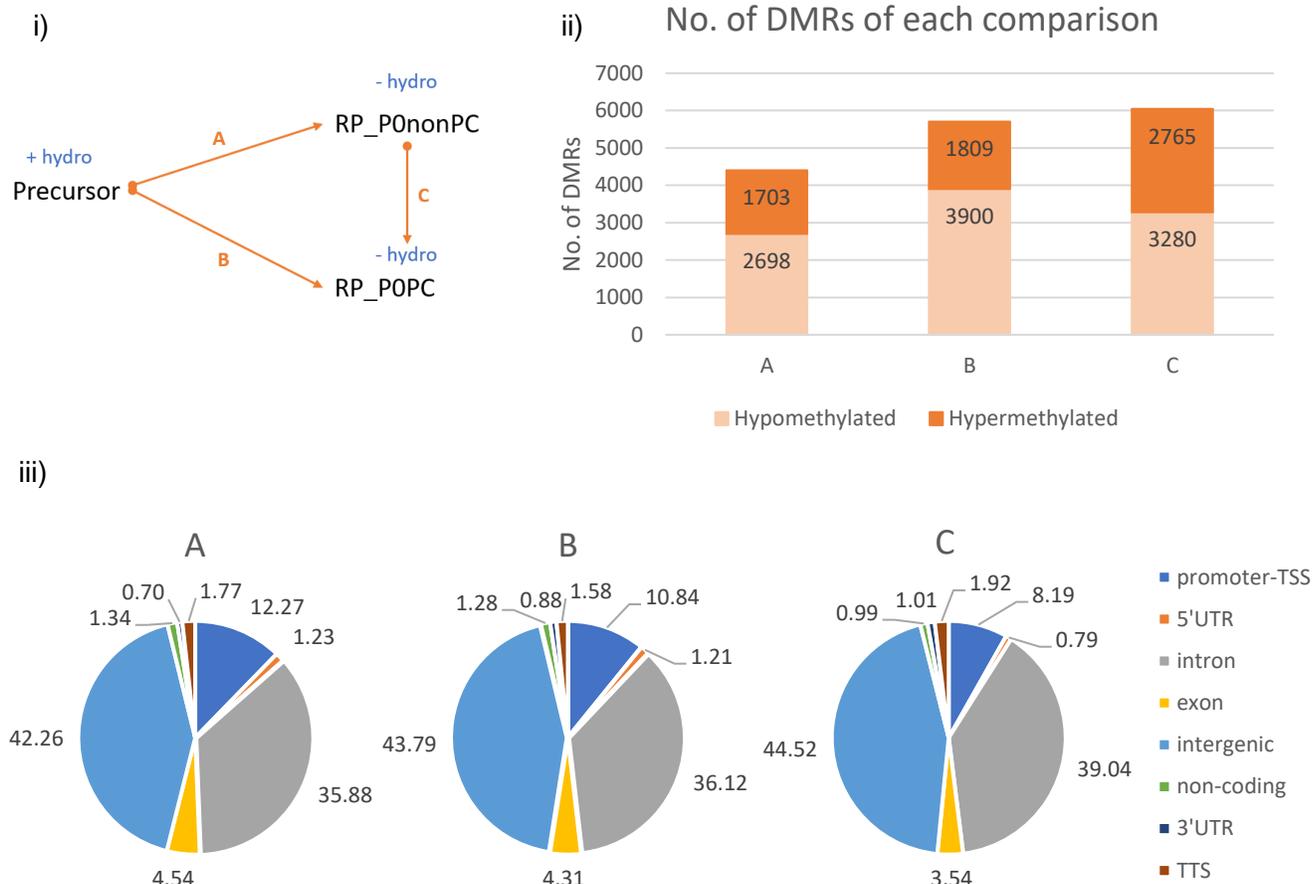
Figure 3.8 Differentially methylated regions (DMRs) after physical constraint. i) shows each pairwise comparison (A-C). Dot in each comparison represents the reference sample, while the sample at the end of the arrow is the observed sample. ii) Bar graph shows hyper- and hypomethylated DMRs for each comparison. The number of DMRs were also shown in each box. and iii) Pie charts show distribution of DMRs according to genome annotation (UTR – untranslated region, TSS – transcriptional start site, TTS – transcriptional termination site).

As a result of physical constraint, the overall number of DMRs in A-C was slightly different but tended to decreased methylation levels in all comparisons (Figure 3.8 ii). While hypomethylated DMRs appeared more prominent in the comparison of +/-hydro (A and B), the comparison of PC effect (C) showed almost equal numbers of hypermethylated DMRs and hypomethylated DMRs. Functional annotation of each comparison was obtained from ChIPSeek website (Chen *et.al.,* 2014), and all DMRs were included in the analysis. The result revealed that DMRs of all comparisons were mostly located in intronic, intergenic and promoter-TSS regions, respectively. DMRs obtained from comparisons A and B were similarly distributed among different transcription-related features, whereas minor differences in distribution were found in DMRs from

C, i.e., the promoter-TSS decreased approx. 2-4%, but the intron and intergenic increased 3%, when compared to A and B (Figure 3.8 iii).

Because comparison B was a combination between cultivation effect and PC effect, A-B DMRs and B-C DMRs were compared to figure out if DMRs evolved from PC or culturing. Interestingly, B-specific DMRs overlapped with A- and C-specific DMRs in 11.43% and 9.38%, resp., of all tiles (see Supplementary Figure 7.1). To extract the PC-specific DMRs, DMRs obtained A and C was performed (Figure 3.9 i). The result showed that 13.9% of DMRs are common between A and C splitting up into 3 categories;

(1) Hypomethylated DMRs in A → Hypomethylated DMRs in C (0.11%)

(2) Hypomethylated DMRs in A → Hypermethylated DMRs in C (6.39%)

(3) Hypermethylated DMRs in A → Hypomethylated DMRs in C (7.4%)

DMRs in the categories (2) and (3) appeared to a similar extent (6.39% and 7.40%, respectively) while DMRs grouping into category (1) were underrepresented (0.11%) (Figure 3.9 ii).
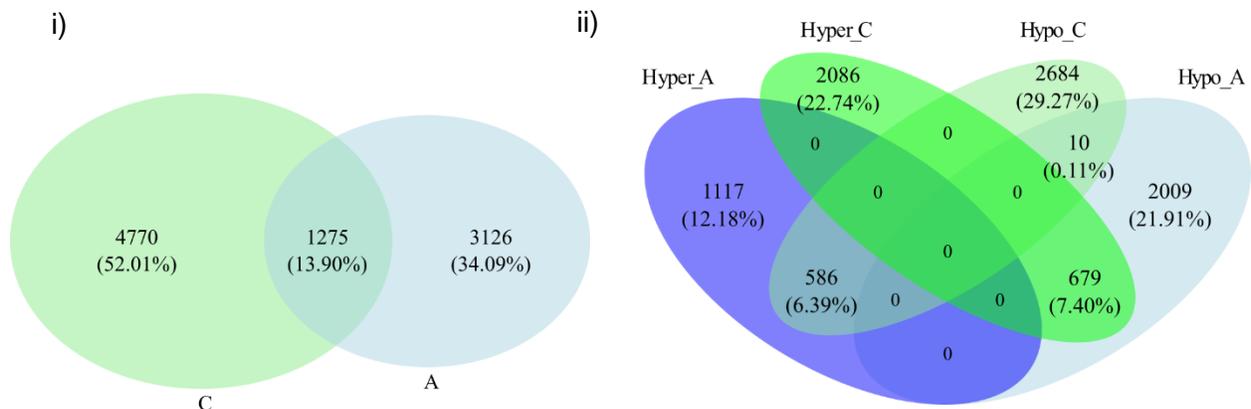
Figure 3.9 Specific DMRs related to physical constraint. i) Overlapping of A and C extracts DMRs evolved specifically by physical constraint. ii) The overlapping is then performed separately by methylation status, yielding 3 groups of overlapping DMRs.

Prolonged cultivation of the cells derived from the RP_P0PC/nonPC represented a maintenance procedure that could reveal effects based on cultivation only. The cells derived from the RP_P0PC were treated with a stabilizer, 5-Azacytidine (5-Aza/Aza), to maintain molecular and morphological features of the cells. Because the 5-Azacytidine treatment directly affects global DNA methylation, the effect of prolonged cultivation and the effect induced by 5-Aza only is of major interest.

For the effect of prolonged cultivation, the PC/nonPC-derived samples were observed independently (Figure 3.10 i). The results showed that number of significant DMRs influenced by long-term cultivation in nonPC-derived comparison (A: 30,393 DMRs) was higher than in PC-derived comparisons (B-D: 26,596, 8,687 and 28,366 DMRs, respectively). Notably, it was found that prolonged cultivation contributed to hypermethylation of the cells. A higher number of DMRs was observed more clearly when comparing early and late passages of cells (B and C). This effect was found sharply increased from early passage (8,687 DMRs in C) to late passage (26,596 DMRs in B) as shown by the comparisons of PC-derived samples (Figure 3.10 ii). Genome annotation analysis was performed for the comparisons A, B and C. While nonPC-derived (A) and PC-derived comparisons (B) revealed similar proportions, remarkable change was observed approx. to 5% in early (C) and late passages (B), particularly within introns (Figure 3.10 iii).
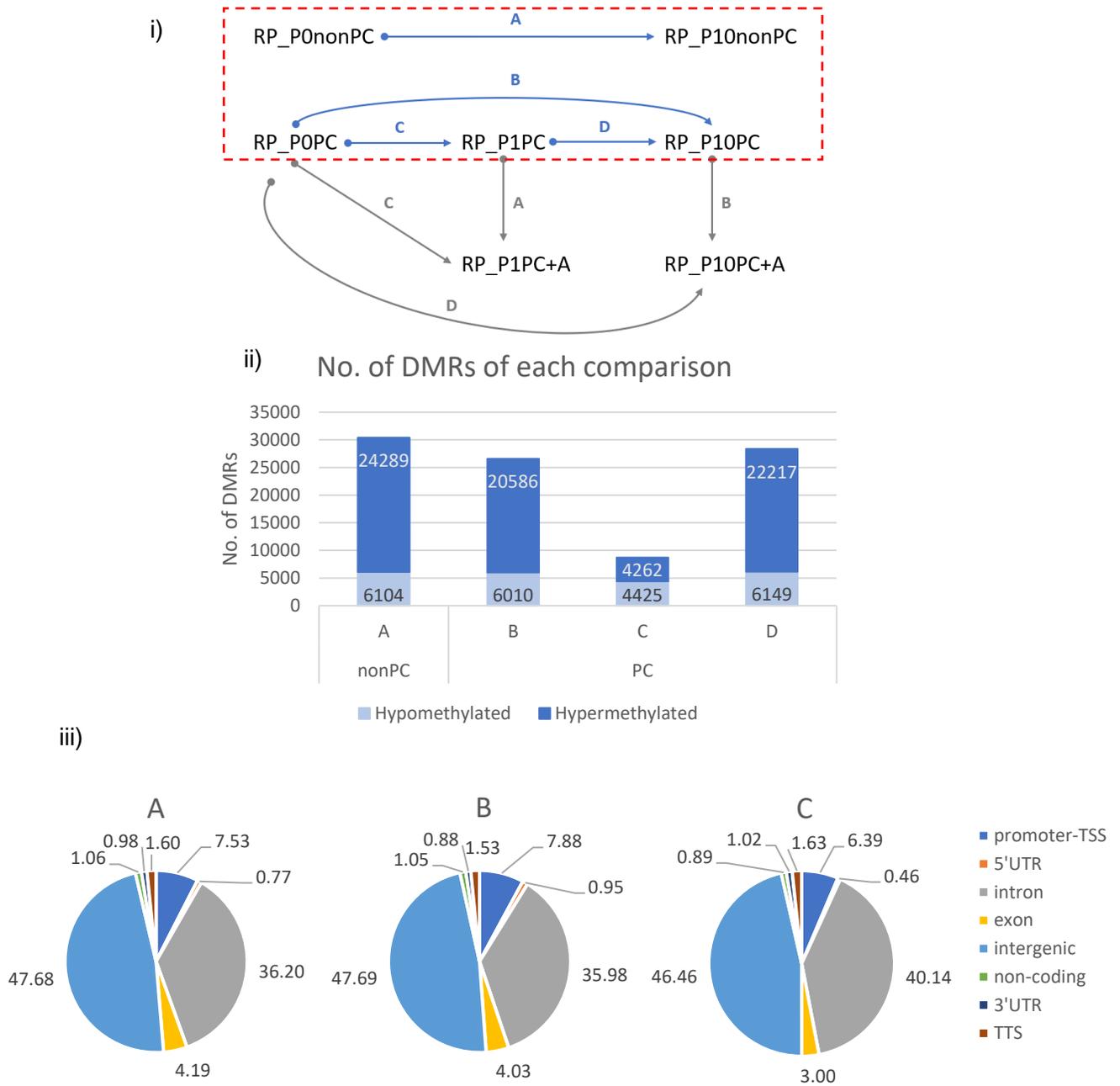
Figure 3.10 DMRs emerging by prolonged cultivation. i) shows each pairwise comparison (A-D). Dot in each comparison represents the reference sample, while the sample at the end of the arrow is the observed sample. ii) Bar graph shows hyper- and hypomethylated DMRs for each comparison. The number of DMRs were also shown in each box. and iii) Pie charts show distribution of DMRs according to genome annotation (UTR – untranslated region, TSS – transcriptional start site, TTS – transcriptional termination site).

To figure out if the differences observed in the early passage remained detectable also in the late passage, comparisons B and C were compared. Shared DMRs (sDMRs) were found only in 9.85% of tiles (Figure 3.11 i), splitting up into 4 categories:

(1) Hypomethylated DMRs in B → Hypomethylated DMRs in C (3.70%)

(2) Hypomethylated DMRs in B → Hypermethylated DMRs in C (0.10%)

(3) Hypermethylated DMRs in B → Hypermethylated DMRs in C (5.63%)

(4) Hypermethylated DMRs in B → Hypomethylated DMRs in C (0.42%)

While the most sDMRs grouped into category (3), followed by category (1), not many sDMRs grouped into (4) and (2) (Figure 3.11 ii).

PC/nonPC-derived DMRs (pcDMRs) were identified by comparing A vs. B. pcDMRs were shown for 29%, which were grouped into 4 categories as already shown for sDMRs (see Supplementary Figure 7.2).
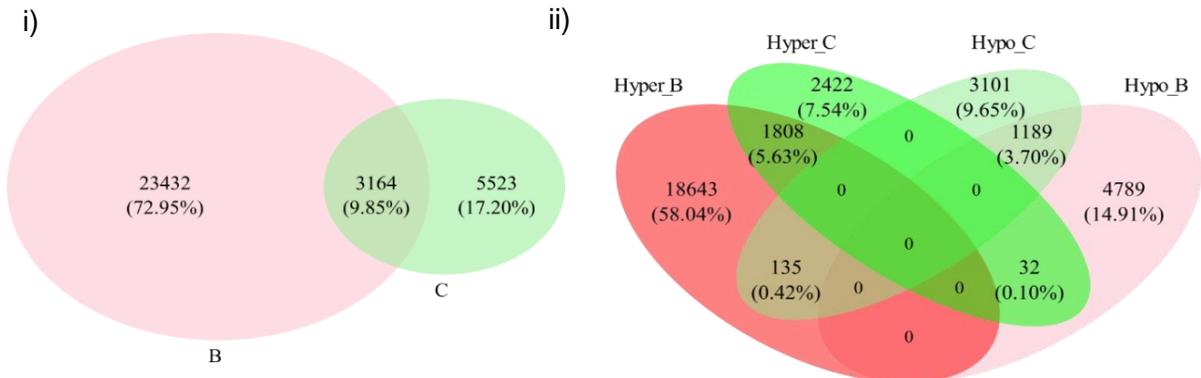
Figure 3.11 Shared DMRs after prolonged cultivation i) Overlapping of B and C extracts shared DMRs (sDMRs) in the comparison between early and late passages. ii) The overlapping is then performed separately by methylation status, yielding 4 groups of sDMRs.

The 5-Aza-treatment was applied solely in PC-derived samples. After treatment, the whole genome hypomethylation was clearly observed in all comparisons. Number of hypomethylated DMRs was increased significantly from passage 1 to passage 10 (A to B and C to D, respectively). The number of significant DMRs was approximately 3-folds higher, compared to the hypomethylated DMRs from passage 1 (A) to passage 10 (B). Meanwhile, hypermethylated DMRs barely occurred (less than 1%, 1,556 DMRs in passage 1(A) and 209 DMRs in passage 10 (B)). To identify DMRs between Aza-treated samples and PC sample passage 0, pairwise comparison of C and D was performed. Nevertheless, the resulting numbers seemed similar to the comparison A and B (early to late passage), but in opposite directions. A slight increase of DMRs was observed from A to C, whereas a slight decrease of DMRs was observed from B to D (Figure 3.12 ii). Genome annotation analysis showed that DMRs were mostly located in intergenic and intronic regions. Interestingly, passage number contributed to the distribution of DMRs into genome annotation categories, for example the distribution of DMRs in A was comparable to that in C and the distribution of DMRs in B was comparable to that in D (Figure 3.12 iii).

i)

RP_P0nonPC ●━━━━━━━━ **A** ━━━━━━━━▶ RP_P10nonPC

B

RP_P0PC ●━━ **C** ━━▶ RP_P1PC ●━━ **D** ━━▶ RP_P10PC

C ↓   **A** ↓   **B** ↓

RP_P1PC+A   RP_P10PC+A

D

ii)

No. of DMRs of each comparison

Comparison	Hypermethylated	Hypomethylated
A	1556	79182
B	209	249902
C	1432	85615
D	577	191234

■ Hypomethylated   ■ Hypermethylated

iii)

A

1.58   2.32   3.38
0.91   0.26
34.17
53.03
4.36

B

1.43   2.19   4.37
0.99   0.39
37.55
48.28
4.81

C

1.55   2.29   3.27
0.87   0.23
34.05
53.62
4.13

D

1.49   2.26   4.57
1.02   0.38
36.42
48.52
5.33

■ promoter-TSS
■ 5'UTR
■ intron
■ exon
■ intergenic
■ non-coding
■ 3'UTR
■ TTS

Figure 3.12 DMRs emerging by the application of 5-Aza. i) shows each pairwise comparison (A-D). Dot in each comparison represents the reference sample, while the sample at the end of the arrow is the observed sample. ii) Bar graph shows hyper- and hypomethylated DMRs for each comparison. The number of DMRs

were also shown in each box. and iii) Pie charts show distribution of DMRs according to genome annotation (UTR – untranslated region, TSS – transcriptional start site, TTS – transcriptional termination site).

3.1.3. Changes in DNA methylation during PCi-differentiation process

PCi-differentiation process starts from RD_P0nonPC+A, which was subjected to physical constraint for 20 hours. After that, the cells were treated with hydrocortisone to reboot some hepatic functions, followed by DMSO to induce cell differentiation. In this state, medium containing hydrocortisone and DMSO was used continuously in the culture until cells were harvested at passage 20 (RD_P20).
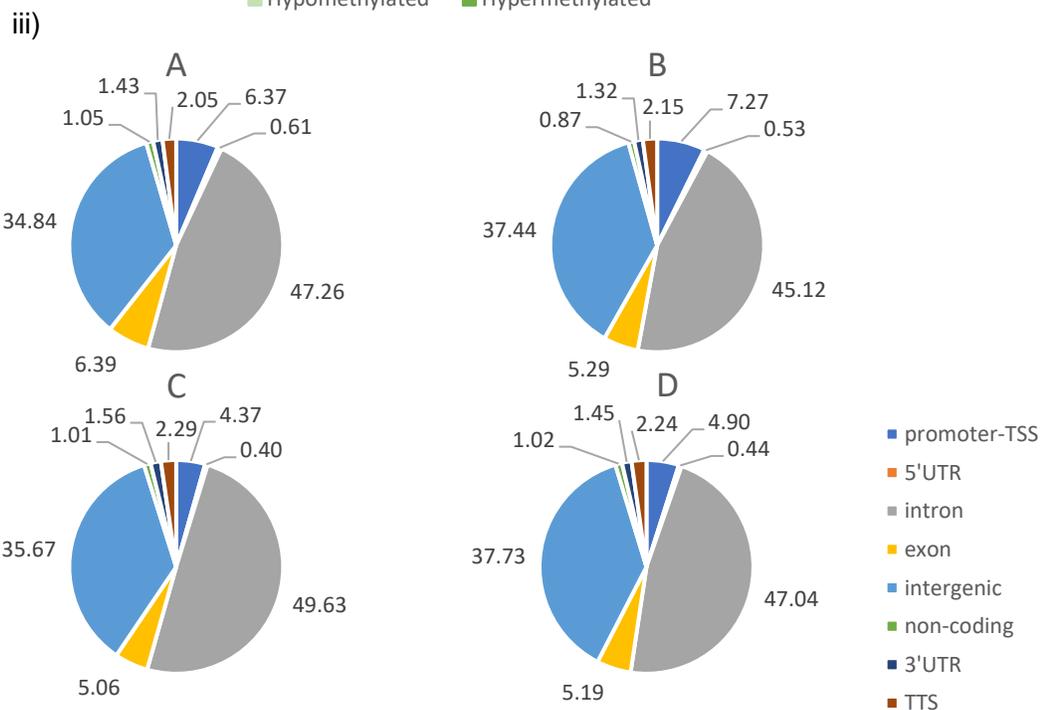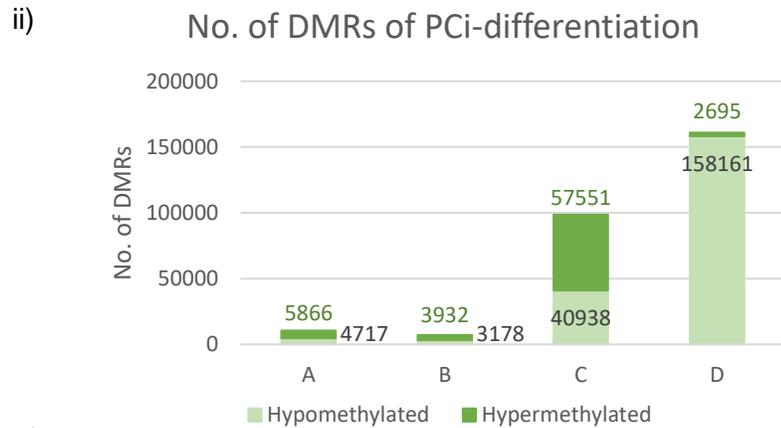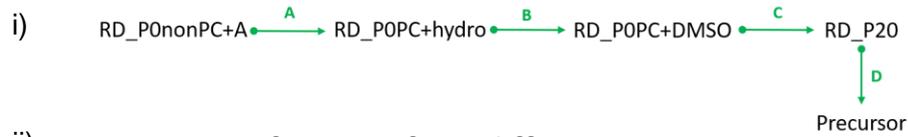
Figure 3.13 DMRs emerging during PCi-differentiation process. i) shows each pairwise comparison (A-D). Dot in each comparison represents the reference sample, while the sample at the end of the arrow is the observed sample. ii) Bar graph shows hyper- and hypomethylated DMRs for each comparison. The number of DMRs were also shown in each box. and iii) Pie charts show distribution of DMRs according to genome annotation (UTR – untranslated region, TSS – transcriptional start site, TTS – transcriptional termination site).

Pairwise comparisons revealed that number of DMRs increased throughout the PCi-differentiation process. In the beginning of PCi-differentiation process, minor changes were found in the DMRs of comparison A and B. Apparently, when partially differentiation occurred in comparison C

(RD_P20), the number of DMRs dramatically increased (10-folds increase). In particular, the results from A to C revealed that the number of hypermethylated DMRs were raised by the supplements. Interestingly, the precursor cells used for the reprogramming experiment still showed clearly distinct methylation patterns compared to the RD_P20. Moreover, hypomethylated DMRs appeared to be more prominent in RD-P20 (D in Figure 3.13 ii). As shown in Figure 3.13 iii for genome annotation, most DMRs were found in intergenic and intronic regions. The relative amount of DMRs in promoter-TSS regions were comparable between the samples analyzed during the PCi-differentiation process (Figure 3.13 iii).

3.1.4. Summary of methylation analyses during reprogramming and PCi-differentiation of HepaRG

*Msp*I-RRBS shed light on major DNA methylation changes during reprogramming and PCi-differentiation processes. Overall, high methylation patterns were observed in most samples undergoing the reprogramming process, while overall low methylation patterns were observed in the samples analyzed during PCi-differentiation.

The reprogramming process was divided to 2 states, physical constraint and maintenance state. Both sample states showed overall high methylation levels. The pairwise comparison of PC vs. nonPC showed minor differences of methylation since the number of DMRs was the lowest among all comparisons. Prolonged cultivation and treatment with 5-Aza in the maintenance state, showed pronounced DNA methylation effects. Prolonged cultivation led to gain of methylation levels in late passage samples, whereas Aza-treatment led to whole-genome hypomethylation in early and late passages.

After induction of PCi-differentiation, DNA methylation tended to increase throughout the PCi-differentiation process, however, did not reach the levels observed for the starter HepaRG culture.

Genome annotation analysis revealed distinct patterns for samples in the reprogramming process and those in the PCi-differentiation process, respectively. While DNA methylation was found to be stable in promoters and 5´UTRs, most DMRs were found in intronic and intergenic regions.

3.1.5. The expression profiles of reprogrammed and PCi-differentiated HepaRG

In parallel with the methylation analysis, 8 of 12 samples were chosen as representatives of each state to perform mRNA sequencing. Quality control of mRNA-sequencing showed that all samples obtained different numbers of reads (from 3×10^7 to 8×10^7, see Supplementary Figure 7.3), but number of detected transcripts and detected genes were comparable, average at 1.2×10^5 transcripts and 2.1×10^4 genes (see Supplementary Figure 7.4). However, the transcript-associated relative distribution of reads was slightly different in each sample, e.g. the samples obtained during the redifferentiation process had higher read numbers in intronic and intergenic regions compared to other samples (Figure 3.14).
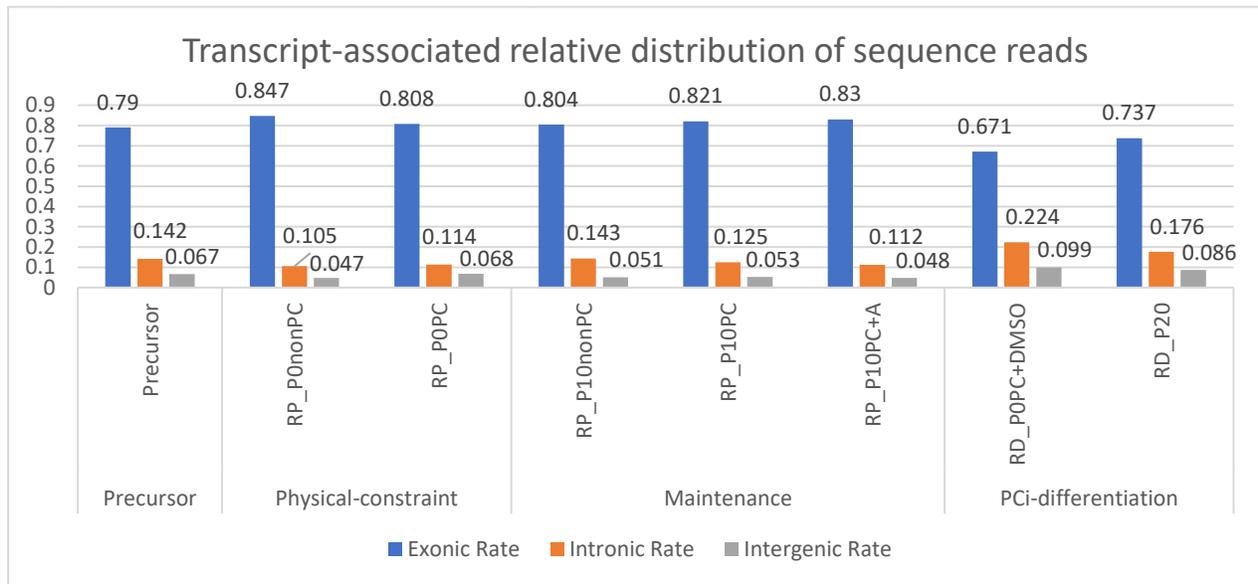
Figure 3.14 Transcript-associated relative distribution of reads of PC samples. The bar diagram shows the relative read distribution in exon, intron, and intergenic regions for all analyzed samples in different epigenetic states.

Similar to the methylation analysis, mRNA-Seq samples from each branch of the experiment were analyzed separately. The overall expression profiles revealed only minor differences between the samples. Although the sample clustering seemed comparable to the clustering defined by DNA methylation, the transcription profile of the HepaRG precursor cells clustered together with those of the samples obtained in the PCi-differentiation process. (Blue bar in Figure 3.7 and Figure 3.15). Particularly, samples from the same passage were close to each other, for instance RP_P0PC and RP_P0nonPC (Orange bar in Figure 3.15).
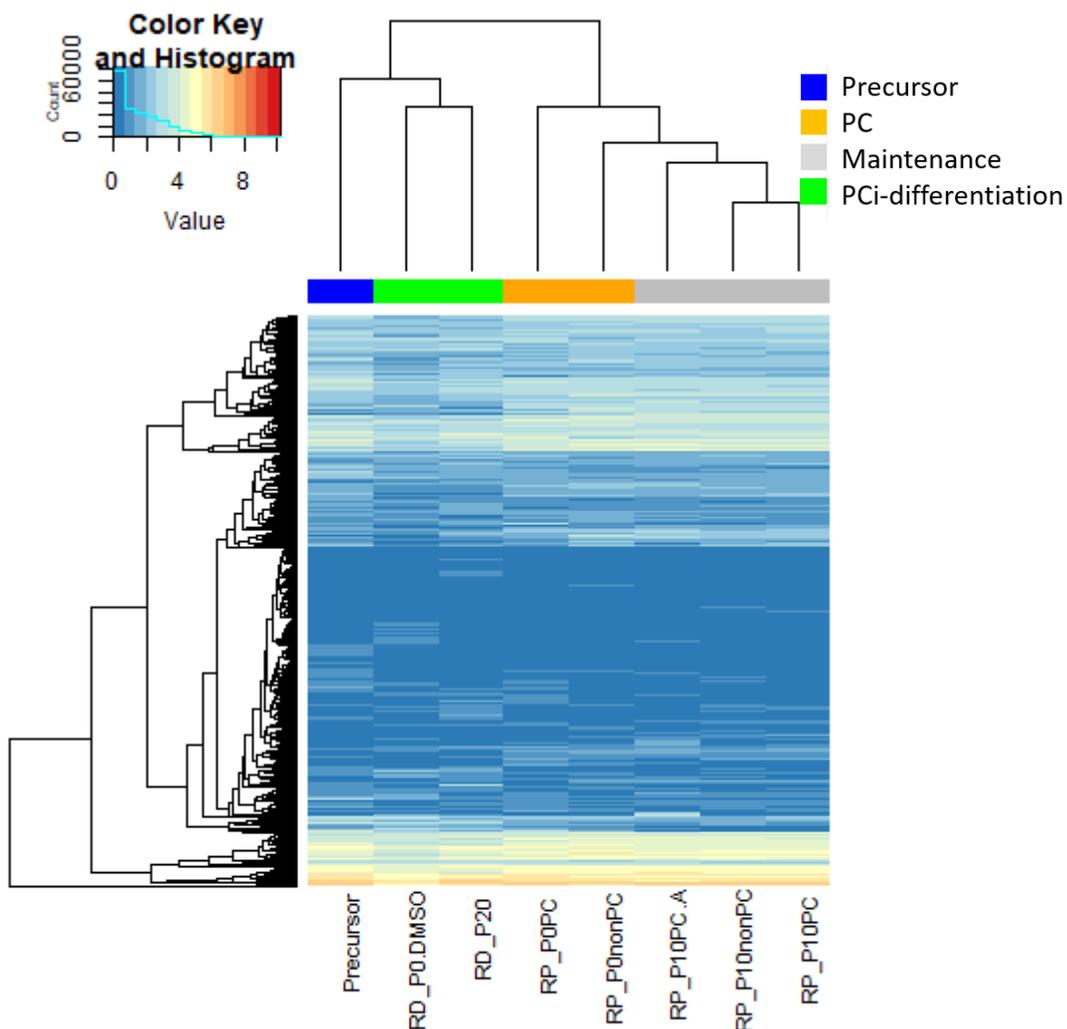
Figure 3.15 Heatmap of 20k variable expressed genes. A gradient of blue to red represents log (FPKM+1) value from low expression to high expression. The processes when samples were obtained are defined by different colors.

3.1.6. Alterations of gene expression during reprogramming process

The analysis of differential expression during reprogramming process were separated to physical constraint and maintenance state. Differentially expressed genes were computed using Cuffdiff in the Cufflink package (Trapnell *et al.,* 2010; Trapnell *et al.,* 2013). Genes with an FDR adjusted p-value less than 0.01 were classified as Differentially Expressed Genes (DEGs). Furthermore, Gene Ontology (GO) of biological process and KEGG pathways of DEGs were analyzed by String-DB database at FDR adjusted p-value less than 0.05 (von Mering *et al.*, 2003; Szklarczyk *et al.*, 2017).
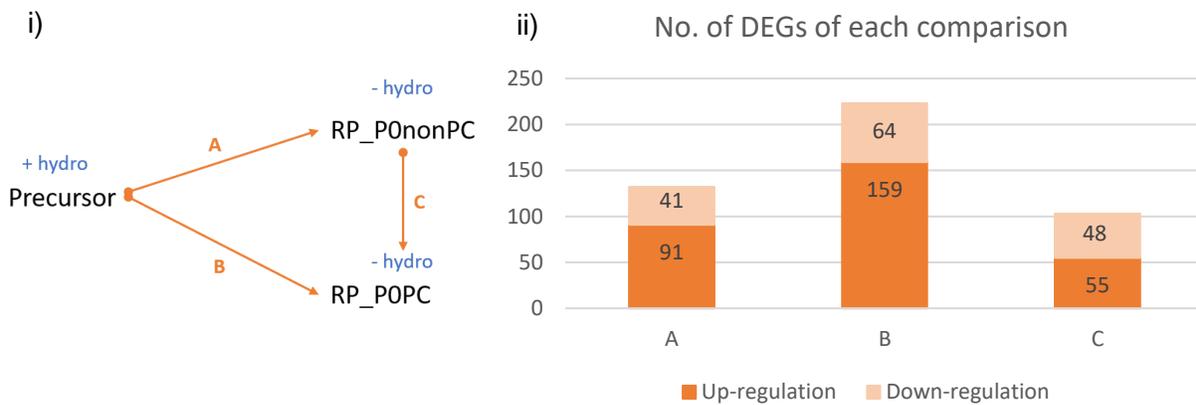
Figure 3.16 Differentially expressed genes (DEGs) in physical constraint state. i) shows each pairwise comparison (A-C). Dot in each comparison represents the reference sample, while the sample at the end of the arrow is the observed sample. ii) Bar graph shows up- (dark color) and down-regulation (light color) for each comparison. Number of DEGs in each direction are shown in boxes.

Pairwise comparisons of samples obtained in the physical constraint branch of the experiment showed that the highest number of DEGs was observed in B, followed by A and C. The up-regulated DEGs were observed predominantly in comparison A and B, while both, up-regulation and down-regulation, was observed almost equally in comparison C (Figure 3.16 i and ii).

Since comparison B represents a combination between hydrocortisone removal (A) and physical constraint (C), overlapping of A vs. B and B vs. C was performed. The results showed that A contained common DEGs with B more than with C (see Supplementary Figure 7.5).

To figure out if hydrocortisone removal (A) influenced the physical constraint (C), comparison of A vs. C was performed. GO and KEGG pathway were also examined for both unique and common DEGs.

The result of A vs. C comparison showed that most DEGs of A and C in both regulation directions were unique, but common DEGs were observed for 13 genes (Figure 3.17 i and iii). Unique DEGs of both regulation directions could be found in GO of biological process and KEGG pathways, except down-regulation in C. For the hydrocortisone removal effect (A), down-regulation mainly related to metabolic processes such as steroid and lipid metabolic processes, while up-regulation related to protein-containing complex binding process (Figure 3.17 ii). As mentioned previously that no GO and KEGG pathway was found in down-regulation in C, the up-regulated transcripts in C, on the other hand, was found to be involved in several biological processes such as immune response and metabolic processes (Figure 3.17 iv).
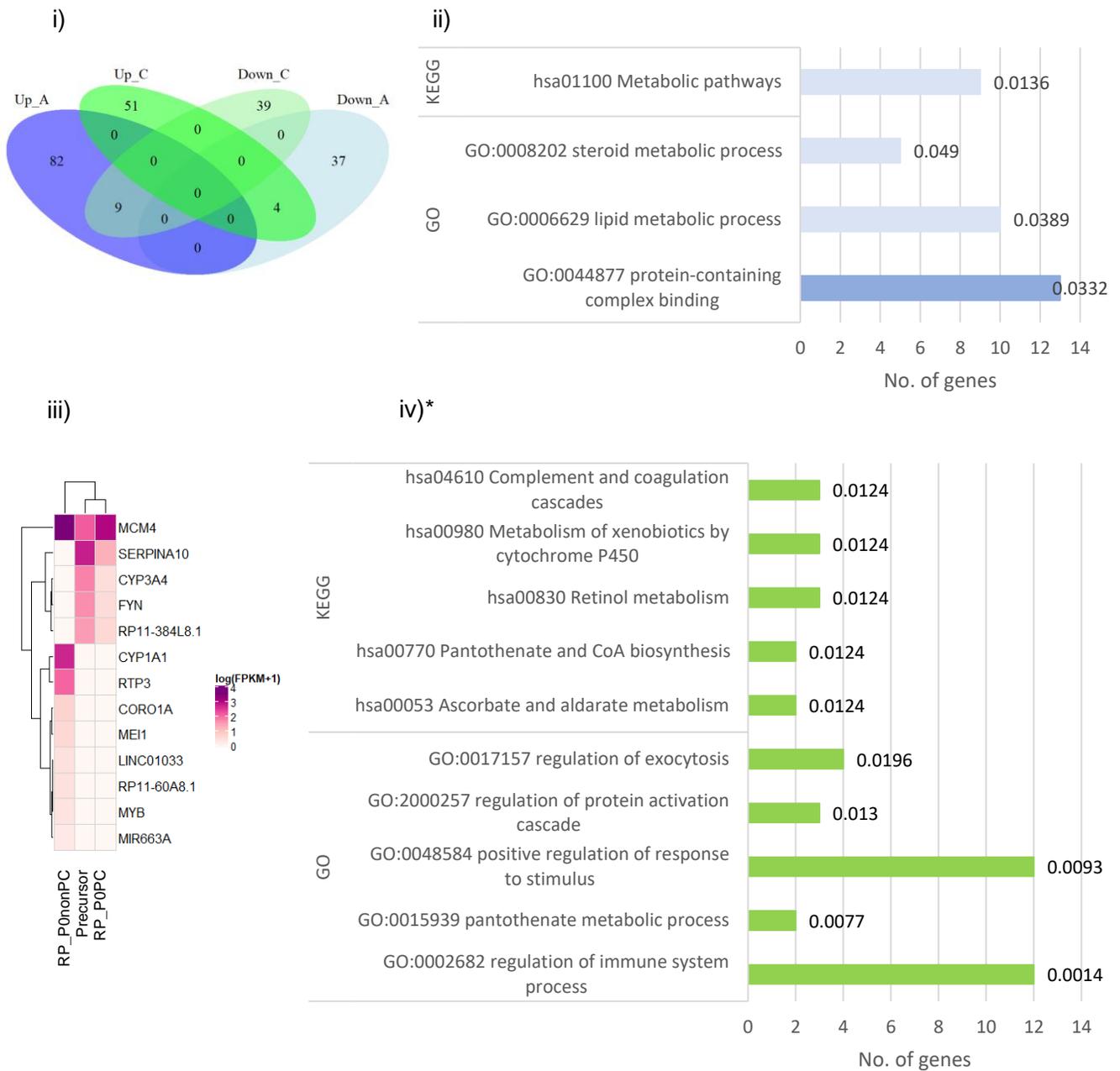
Figure 3.17 GO terms and KEGG pathways of DEGs in samples of the physical constraint branch of the experiment. i) Venn diagram showing number of DEGs the overlapping between hydrocortisone removal effect (A; Blue) and physical constraint (C; Green) in each regulation direction. Dark colors represent up-regulated genes, while light colors represent down-regulated genes. ii) and iv) Bar graph shows number of genes related to GO terms and KEGG pathways of hydrocortisone removal effect (ii) and physical constraint (iv). FDR-adjusted p-value are written at the end of bars. Colors are also defined following the Venn diagram. *No GO terms and KEGG pathways of down-regulated genes are presented (no light green). iii) Heatmap shows the expression value (log (FPKM+1)) of common DEGs. Gradient colors from white to purple represent low to high expression, respectively.

Three samples from passage 10 were selected to be subjects of mRNA sequencing representing the maintenance state focusing on long-term (prolonged) cultivation and Aza-treatment effects (Figure 3.18). Interestingly, the sample derived from PC (B) showed strong differences, compared to the sample derived from nonPC (A). However, at the late passage, there were not many differences in differential expression between PC and nonPC (C), compared to the early passage in the physical constraint state (Figure 3.16 ii, comparison C). Moreover, the long-term cultivation effect tended towards down-regulation in all comparisons (A-C).
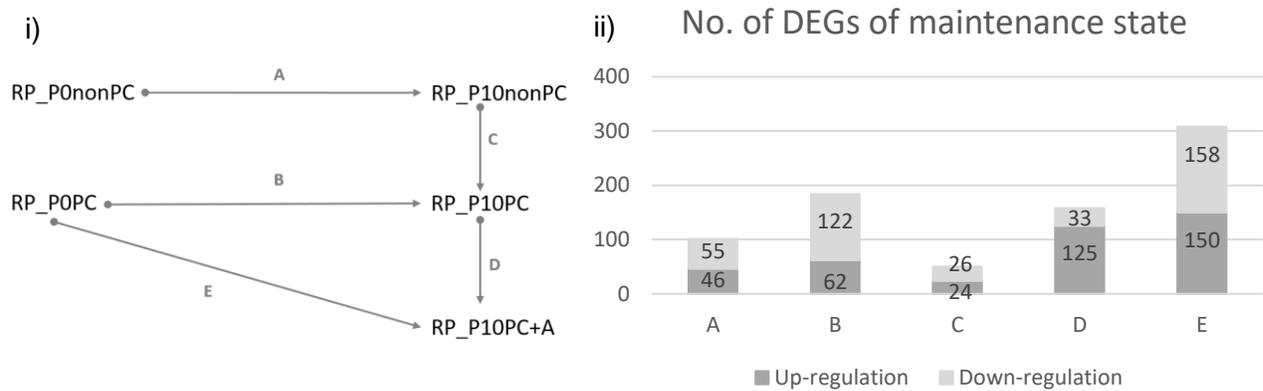
Figure 3.18 DEGs obtained during maintenance of PC and nonPC cells. i) shows each pairwise comparison (A-E). Dot in each comparison represents the reference sample, while the sample at the end of the arrow is the observed sample. ii) Bar graph shows up- (dark color) and down-regulation (light color) for each comparison. Number of DEGs in each direction are shown in boxes.

The results of PC vs. nonPC comparisons showed that most of DEGs (235 genes) are unique, while common DEGs were found only in 25 genes (Figure 3.19 i and iii).

Figure 3.19 GO terms and KEGG pathways of DEGs in samples of the maintenance state, particularly for prolonged cultivation branch. i) Venn diagram showing number of DEGs the overlapping between nonPC (A; Blue) and PC (C; Red) of passage 10 in each regulation direction. Dark colors represent up-regulated genes, while light colors represent down-regulated genes. ii) and iv) Bar graph shows number of genes related to GO terms and KEGG pathways of nonPC (ii) and PC (iv). FDR-adjusted p-value are written at the end of bars. Colors are also defined following the Venn diagram. iii) Heatmap shows the expression value (log (FPKM+1)) of common DEGs. Gradient colors from white to purple represent low to high expression, respectively.

Unique DEGs could be found in GO of biological process and KEGG pathways in all directions (up-up, up-down, down-up and down-down). Down-regulated DEGs of A related to stress responses and cellular matrix organization, while down-regulated DEGs of B mainly related to catabolic/metabolic processes. On the other hand, up-regulated DEGs of A mainly related to many tumorigenesis processes, while up-regulated DEGs of B related to regulation of tyrosine kinase activity (Figure 3.19 ii and iv).

To study the Aza-treatment effect, Aza-treated sample was compared to the same passage of the PC-derived sample and the original PC sample (Figure 3.18 i). The results showed that Aza-treatment led to remarkable increase of up-regulated DEGs. However, significant differences between D and E were observed, since down-regulated DEGs increased dramatically when compared to the PC sample (Figure 3.18 i and ii, comparison D-E).

When looking at GO categories and presence of DEGs in KEGG pathways, up-regulated DEGs were found related to hsa04512 ECM-receptor interaction (5 observed genes, FDR = 0.0075) in comparison D, while in comparison E, both up- and down-regulated DEGs were strongly related to several GO terms and KEGG pathways. Up-regulated DEGs of E related to GO terms concerning cell cycle and movement, and tissue morphogenesis, whereas down-regulated DEGs of E related to biological processes such as platelet degranulation, protein activation etc. (Figure 3.20).
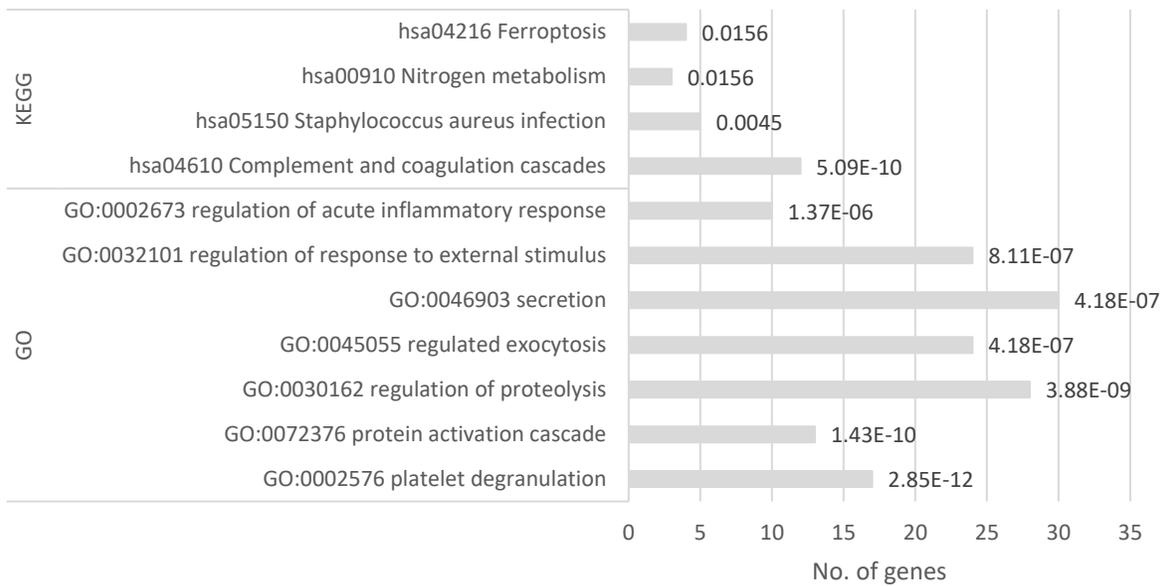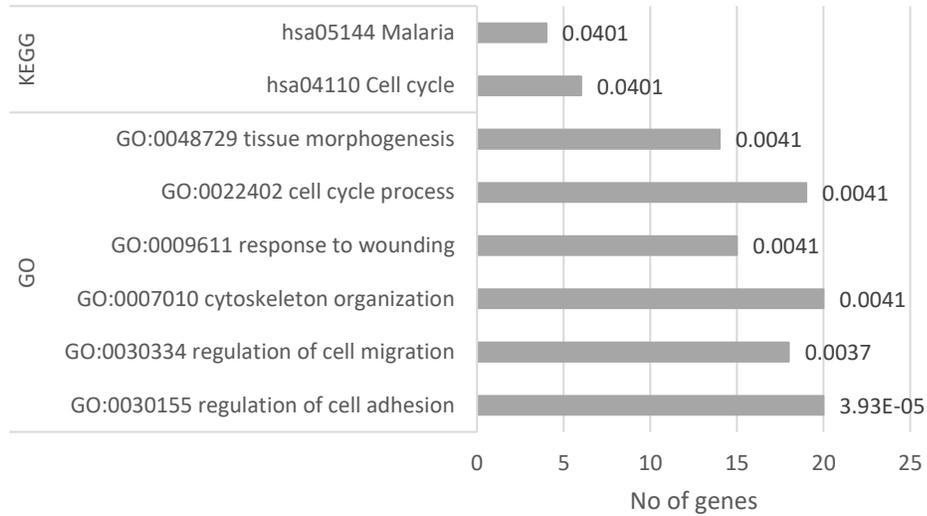
Figure 3.20 GO terms and KEGG pathways of DEGs related to Aza treatment effect (Comparison E). The upper part is a bar graph showing GO terms and KEGG pathways related to up-regulated DEGs, while the lower part belongs to down-regulated DEGs. FDR-adjusted p-values are indicated at the end of each bar.

3.1.7. Alterations of gene expression during PCi-differentiation process

i)

ii) No. of DEGs of PCi-differentiation process
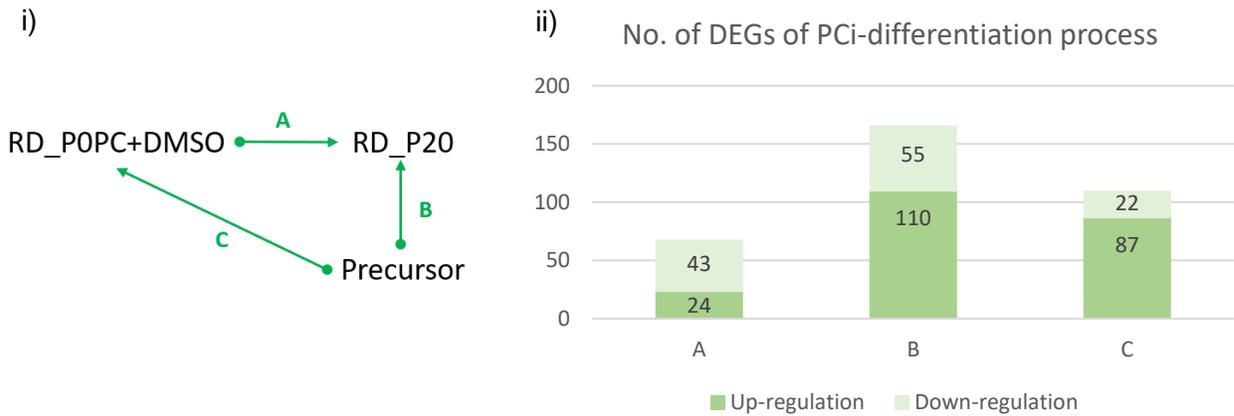
Figure 3.21 DEGs in PCi-differentiation process. i) shows each pairwise comparison (A-C). Dot in each comparison represents the reference sample, while the sample at the end of the arrow is the observed sample. ii) Bar graph shows up- (dark color) and down-regulation (light color) for each comparison. Number of DEGs in each direction are shown in boxes.

Two samples (RD_P0PC+DMSO and RD_P20) that were obtained during the PCi-differentiation branch of the experiment were subjected to perform mRNA-sequencing (Figure 3.21 i). The result showed that there were not many differences in expression (67 genes) between both RD samples, and 64% of those were down-regulated DEGs (A). On the other hand, the comparisons of both RD against precursor of differentiated cells (B and C) revealed higher numbers (165 and 99 genes, respectively), which were more up-regulated DEGs than down-regulated DEGs obviously (Figure 3.21 ii).

However, GO and KEGG pathway analysis revealed that hsa04977 Vitamin digestion and absorption related to up-regulated DEGs of comparison A (2 observed genes, FDR = 0.0181). Similar to comparison B and C, not many GO terms and KEGG pathways related to DEGs in all regulation directions (Figure 3.22).
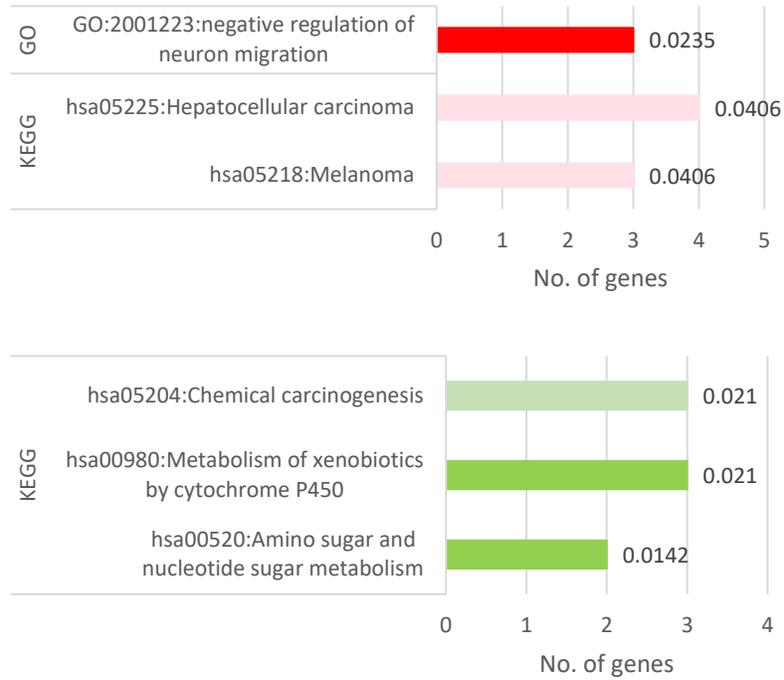
Figure 3.22 GO terms and KEGG pathways of DEGs in PCi-differentiation process. The red bar graph represents GO terms and KEGG pathways related to comparison B, while the green bar graph related to comparison C. Dark colors is up-regulated gene, whereas light color is down-regulated gene. FDR-adjusted p-value are written at the end of bars.

3.1.8. mRNA expression of cell type-specific genes, pluripotency markers and epigenetic modifiers during reprogramming and PCi-differentiation processes
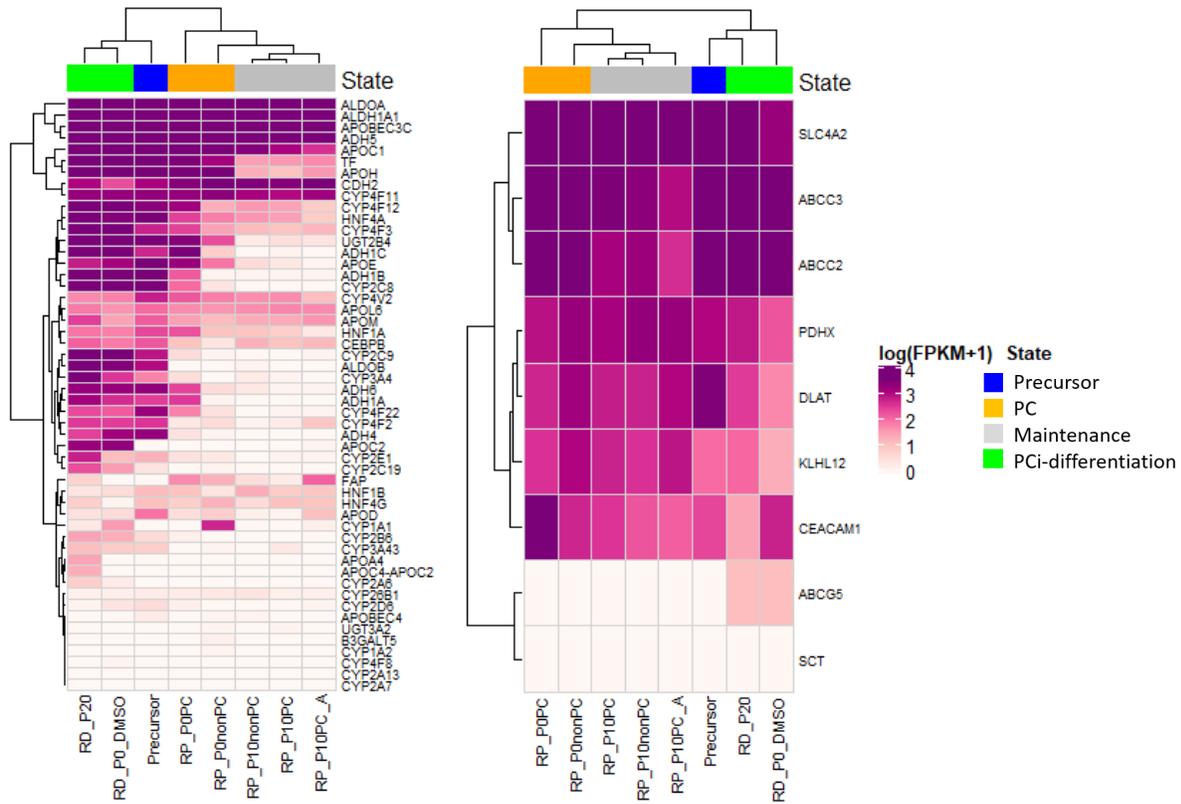
Figure 3.23 Expression of cell type-specific genes. Heatmap shows the expression value (log (FPKM+1)) of hepatocyte-specific genes (left) and biliary-specific genes (right). Gradient colors from white to purple represent low to high expression, respectively.

Hepatocyte- and biliary-specific genes were obtained from The Human Protein Atlas database (Uhren *et al.*, 2010; Uhren *et al.*, 2015). The heatmap of those gene sets showed consistency in clustering with the heatmap of most variable expressed genes (see Figure 3.15).

Most hepatocyte-specific genes showed high expression, particularly in precursor and PCi-differentiation samples, but became lower expressed in physical constraint and maintenance samples (Figure 3.23). In addition, although the clustering was consistent with hepatocyte-specific genes, minor differences in expression of biliary-specific genes were detected in all

states. For instance, *ABCC2* was lower expressed in maintenance samples and *ABCG5* was expressed exclusively in PCi-differentiated samples etc.
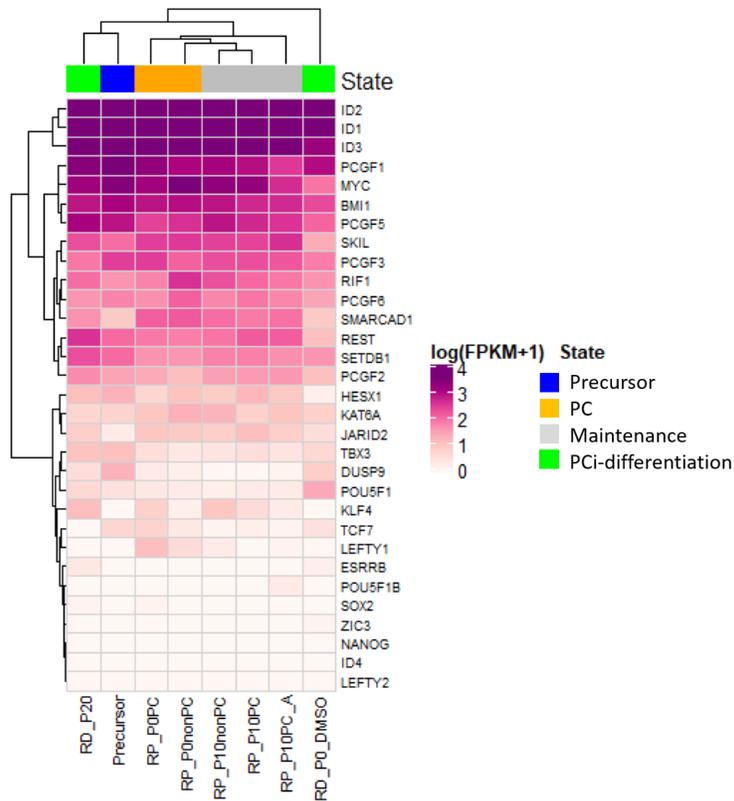
Figure 3.24 Expression of pluripotent marker genes. Heatmap shows the expression value (log (FPKM+1)) of genes related to pluripotency. Gradient colors from white to purple represent low to high expression, respectively.

Genes related to pluripotency were collected from the KEGG signaling pathway regulating pluripotency of stem cells (KEGG: hsa04550). The highest expression was shown in the DNA-binding protein inhibitor family (*ID*), followed by the polycomb group family (*PcG*). On the other hand, OSKM factors and *NANOG* were barely expressed, except *MYC* (Figure 3.24). Signal transduction pathways that related to pluripotency were also analyzed. Interestingly, TFGβ signaling pathway showed prominent expression in Physical constraint and maintenance state, while JAK-STAT-MAPK and WNT signaling pathways did not show any remarkable differences between reprogramming and PCi-differentiation processes (see Supplementary Figure 7.6).
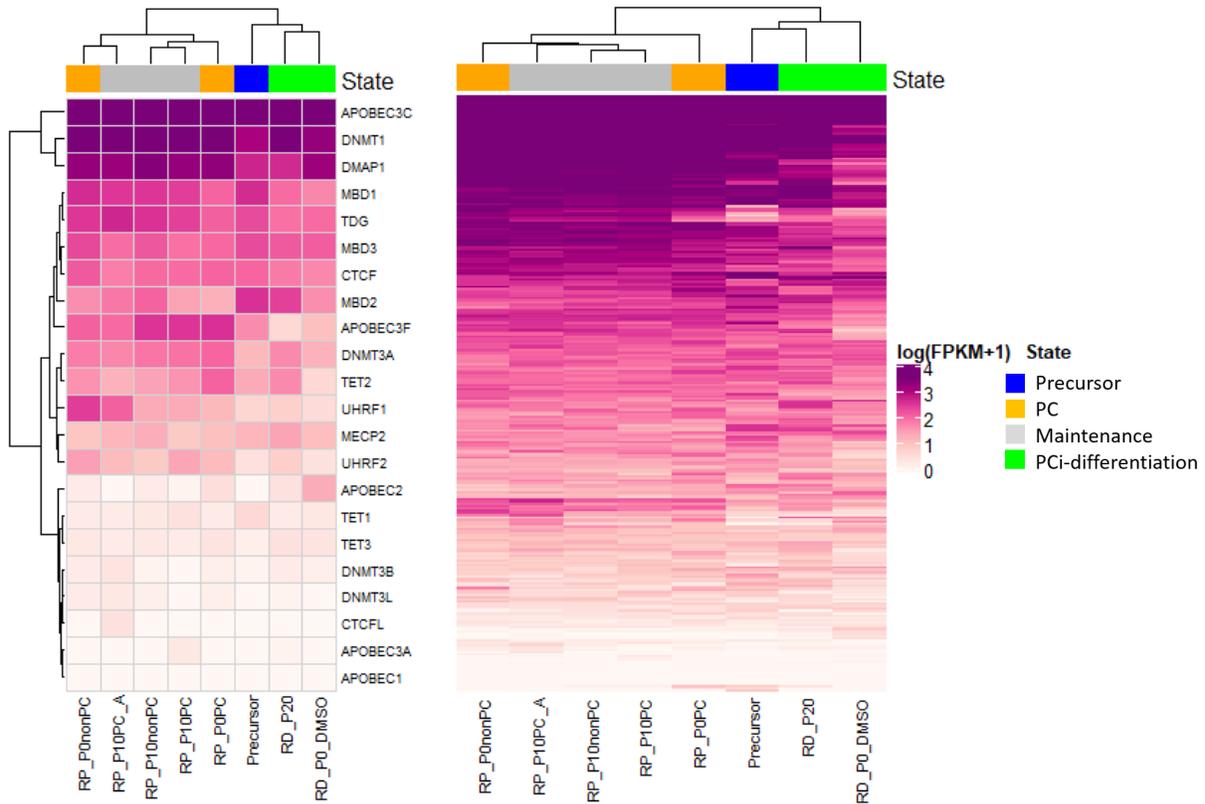
Figure 3.25 mRNA expression of genes related to epigenetic modifications. Heatmap shows the expression value (log (FPKM+1)) of genes involved in modification of DNA methylation (left) and genes related to the modification of histone modifications (right). Gradient colors from white to purple represent low to high expression, respectively.

For genes related to epigenetic modifications, genes involved in the process of DNA methylation (left) and histone modifications were observed independently (right). The clustering of genes from both mechanisms showed slight differences from the heatmap of 20k variable expressed genes, since cells in RP passage 0 were not found in the same cluster (see Figure 3.15). However, minor differences can still be observed, particularly between physical constraint and maintenance samples.

Focusing on DNA methylation mechanism, well-known genes related to DNA methylation and demethylation were selected and analyzed. The results showed that the expression of genes related to the establishment and maintenance DNA methylation, such as *DNMT1* and *DNMT3A,* seemed to be stronger expressed than genes related to DNA demethylation e.g. members of the TET family (*TET1/2/3*) (Figure 3.25).

Genes related to the modification of histones showed gradual expression changes, since there are a lot of genes that orchestrate the process of modification. Therefore, well-known genes related to histone acetylation and histone methylation were examined independently.
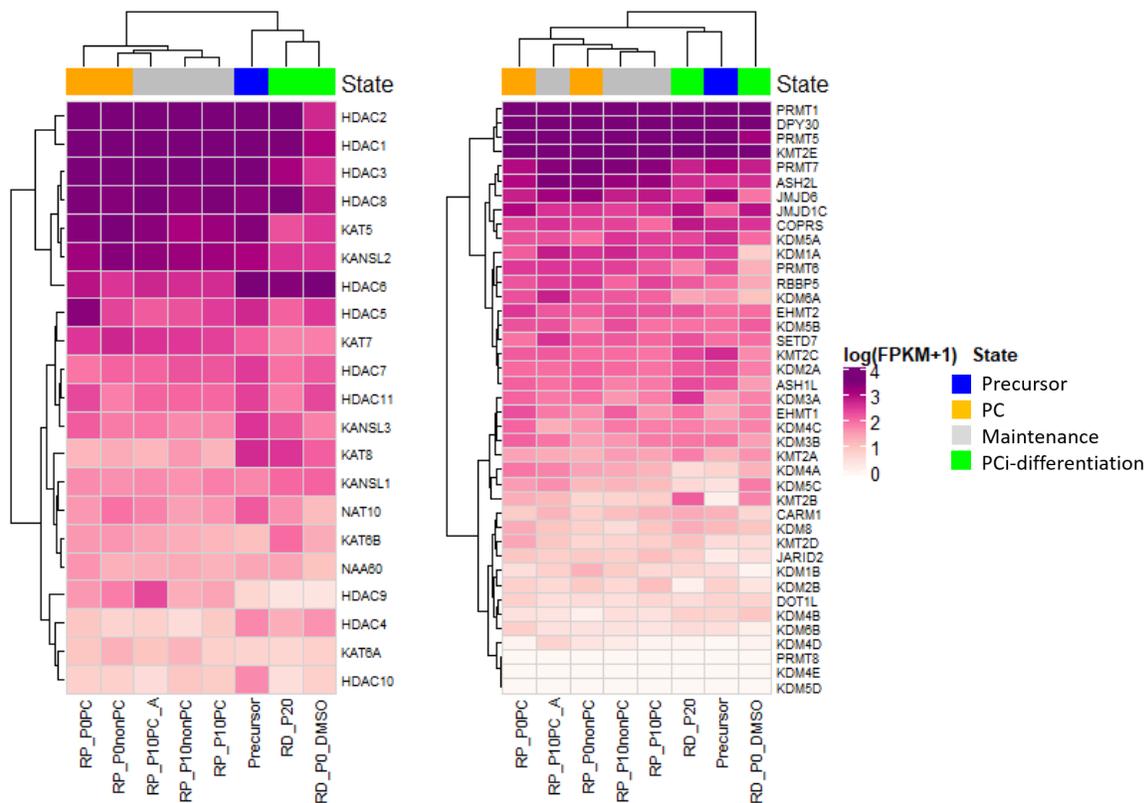
Figure 3.26 mRNA expression of histone tail modifiers. Heatmap shows the expression value (log (FPKM+1)) of genes related to histone acetylation (left) and genes related to histone methylation (right). Gradient colors from white to purple represent low to high expression, respectively.

The results showed that genes related to histone acetylation (left) and histone methylation(right) are highly expressed during reprogramming and PCi-differentiation processes. Differences in clustering of both modifications were observed. For histone acetylation genes, precursor and PCi-differentiated samples were found in the same cluster, whereas samples from the physical constraint branch clustered together with samples from the maintenance branch. Interestingly, members of HDAC gene family were strongly expressed in all states, whereas for histone methylation, the arginine methyltransferase (PRMTs) showed higher expression compared to the other genes (Figure 3.26).

3.1.9. <u>Summary of expression changes during reprogramming and PCi-differentiation processes</u>

Although transcriptomic profiles from mRNA sequencing showed similar patterns among reprogramming and PCi-differentiation processes, the clustering of those processes seemed distinguishable. Reprogrammed samples were grouped in the same node, but all can be separated between PC and maintenance states. On the other hand, samples belonging to the PCi-differentiation were grouped in the same node, but clearly separated from precursor. The GO terms and KEGG pathways were summarized in the table below (Table 3.1).

Expression of genes related to the following pathways were observed. i) Hepatocyte specific genes showed high expression in precursor and PCi-differentiated samples, while biliary specific genes were found to be expressed in all states. ii) Genes related to pluripotency, particularly the OSK factors and *NANOG*, showed low overall expression. iii) Genes related to signal transduction pathways were also examined. Only genes of the TGFβ signaling pathway showed remarkable differences between reprogramming and PCi-differentiation processes. iv) Genes related to DNA methylation processes were stronger expressed, compared to genes related to DNA demethylation in all states. v) Genes related to histone acetylation and methylation were found in all states to be expressed similarly to genes involved in DNA methylation.

Table 3.1 Summary of GO terms and KEGG pathways related to DEGs in each state of PC experiment.

State (case vs. control)	GO terms and KEGG pathways	
	Up-regulation	Down-regulation
Physical constraint		
- Hydrocortisone Removal (RP_P0nonPC vs Precursor)	Protein-containing complex binding process	Metabolic processes
- Physical constraint (RP_P0PC vs RP_P0nonPC)	Exocytosis, Protein activation cascades, Response to stimulus, Complement and coagulation, Metabolic processes	-
Maintenance		
- Prolonged culture of nonPC (RP_P10nonPC vs RP_P0nonPC)	Tumorigenesis processes	Stress responses, Cellular matrix organizations
- Prolonged culture of PC (RP_P10PC vs RP_P0PC)	Tyrosine kinase activity	Catabolic/metabolic processes
- Aza-treatment effect (RP_P10PC+A vs RP_P10PC)	Cell cycle and movement, Tissue morphogenesis	Exocytosis, Protein activation cascades, Response to stimulus etc.
PCi-differentiation		
- PCi-differentiation (RD_P20 vs RD_P0PC+DMSO)	Vitamin digestion and absorption	-

3.1.10. Correlation of DNA methylation and gene expression during reprogramming and PCi-differentiation processes

To figure out a potential correlation between epigenetic modifications and mRNA expression levels of the whole genome, DNA methylation and gene expression profiles were plotted in the heatmap next to each other. Although the overall DNA methylation showed distinctive patterns during reprogramming and PCi-differentiation processes, mRNA expression patterns of DMR-neighbouring genes seemed to be robust (Figure 3.27 and Supplementary Figure 7.7).
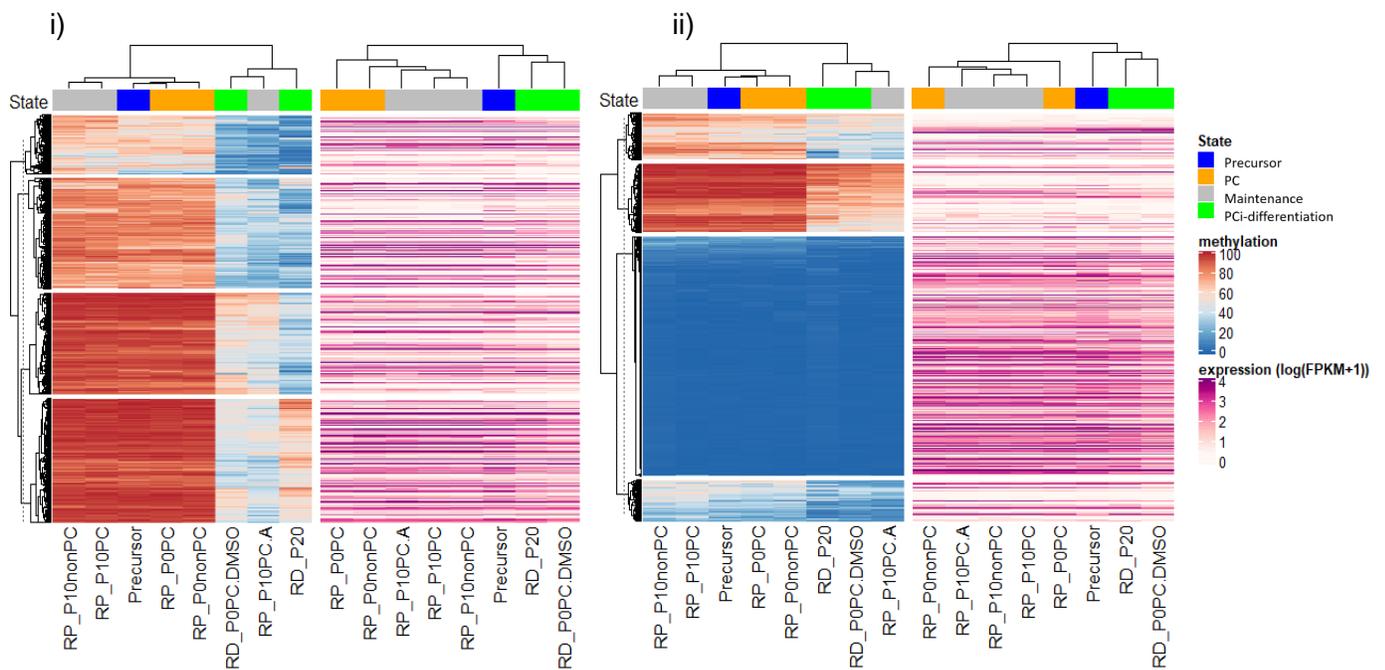
Figure 3.27 Correlation of DNA methylation and gene expression. i) showing heatmap of methylation and expression of 20k variable regions. ii) showing heatmap of methylation of tiling regions related to promoter-TSS and expression of genes associated to those tiling regions. Left panel of i) and ii) shows DNA methylation profile, while right panel shows gene expression profile. The colors, blue and red, represent the methylation level from 0 to 1, and gradient colors from white to purple represent low to high expression, respectively.

Functional annotation of DMRs revealed that anti-correlation of DNA methylation and gene expression was observed in the 5'UTR-promoter-related DMRs, whereas other functional annotation-related DMRs, such as intragenic, intergenic or TTS regions, had no clear tendencies (see Supplementary Figure 7.8).

3.1.11. <u>Correlation of DMRs and DEGs of samples undergoing reprogramming and PCi-differentiation</u>

To determine the correlation of DMRs and DEGs of each state during reprogramming and PCi-differentiation, genes related to DMRs and DEGs were overlapped related to their methylation and expression status, respectively. All DMRs were divided into 2 groups for further analysis: DMRs that show methylation change in opposite direction to transcriptional change (e.g. hyper-DMR/down-DEG or hypo-DMR/up-DEG) were classified as anti-correlated DMRs (aDMRs). Conversely, DMRs that show methylation change in the same direction to transcriptional change (e.g. hyper-DMR/up-DEG or hypo-DMR/down-DEG) were classified as positively correlated DMRs (cDMRs).
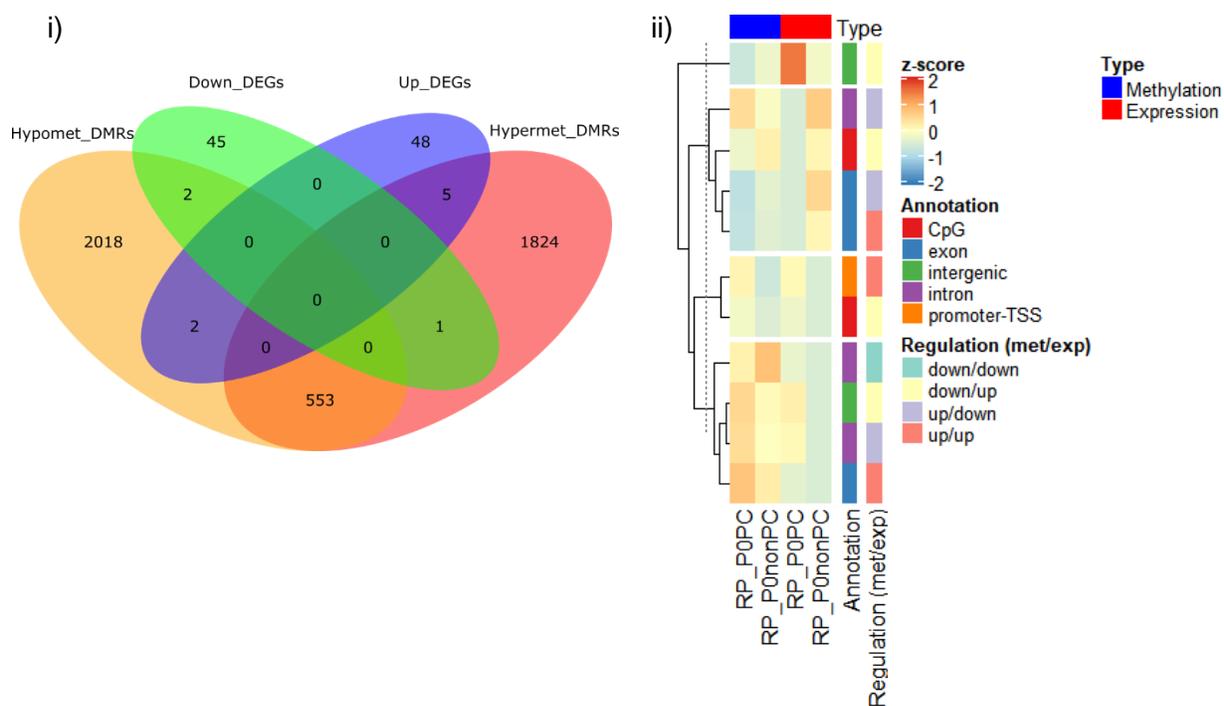
Figure 3.28 Correlation of DMRs and DEGs after physical constraint of HepaRG precursor cells i) venn diagram shows the overlapping genes between hypo/hypermethylatedDMRs and up/down regulated DEGs. ii) Heatmap of methylation (blue) and expression (red) was drawn based on row Z-score. Gradient colors from red to blue represent the value of z-score from high to low, respectively. Types of genome annotations related to DMRs and regulatory functions are also defined on the right side of the heatmap (CpG – CpG islands, TSS – transcription start site).

After physical constraint, 10 DEGs with 11 DMRs were found in both regulation directions (Figure 3.28 i). Remarkably, 7 anti-correlations (yellow and purple in regulation bar) and 4 positive correlations (green and red in regulation bar) were also found, but all of those were assigned to different functional annotation groups (Figure 3.28 ii). For instance, while 4 out of 7 anti-correlated DMRs corresponding to down/up regulation (yellow bar) were found in CpG islands and intergenic regions, 3 of 7 anti-correlated DMRs corresponding to up/down regulation (purple bar) were found in intron and exon regions.
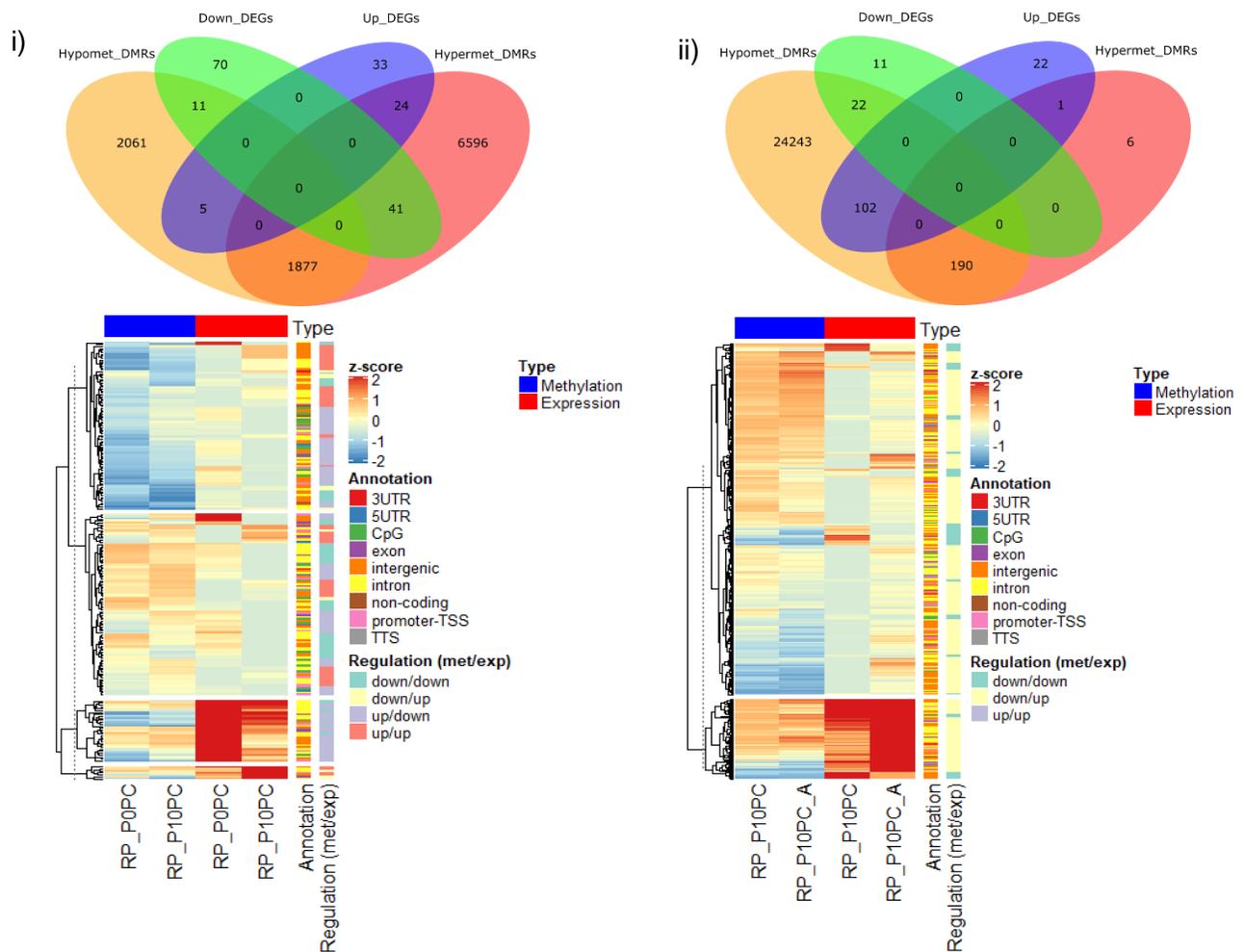
Figure 3.29 Correlation of DMRs and DEGs after maintenance of reprogrammed HepaRG cells. Venn diagrams (upper part) and heatmaps (lower part) show the overlapping genes between hypo/hypermethylatedDMRs and up/down regulated DEGs of i) long-term (prolonged) cultivation effect and ii) Aza-treatment effect. Heatmap of methylation (blue) and expression (red) was drawn based on row Z-score. Gradient colors from red to blue represent the value of z-score from high to low, respectively. Genome annotation categories related to DMRs and regulatory functions are also defined on the right side of the heatmap (CpG – CpG islands, UTR – untranslated region, TSS – transcriptional start site, TTS – transcription termination site).

Correlations and anti-correlations between DEGs and DMRs in the maintenance state were detected in both, long-term cultivation, and Aza-treatment effects. For the long-term (prolonged) cultivation effect, there were 81 genes that correspond to 218 DMRs. The results showed that 65 genes correlated to hypermethylation and 16 genes correlated to hypomethylation. Depending on transcriptomic data, there were 29 genes correlated to up-regulation and 52 down-regulation (Figure 3.29 i). Furthermore, the Aza-treatment effect showed 125 genes with corresponding 1230 DMRs. Interestingly, 124 genes, 102 up-regulated and 22 down-regulated genes, correlated to

hypomethylation (Figure 3.29 ii). *PRAME* is the solely up-regulated gene, which showed hypermethylation at an intergenic region next to the 3'UTR of the gene (10Kb from TSS).

i)

Anti-correlated DMRs
1.79 0.00
1.79 12.50
26.79 3.57
31.25
6.25

Correlated DMRs
0.00 1.85 3.70 4.63
0.93 0.93
37.96 45.37
4.63

Legend: promoter-TSS, 5'UTR, intron, exon, intergenic, non-coding, 3'UTR, TTS, CpG

ii)

Anti-correlated DMRs
1.07 1.65 4.66 8.63
0.87 1.07
37.25 39.96
4.85

Correlated DMRs
1.01 8.04 3.52 0.50
0.50
0.50 34.17
48.74 3.02

Legend: promoter-TSS, 5'UTR, intron, exon, intergenic, non-coding, 3'UTR, TTS, CpG

Figure 3.30 Distribution of functional genome annotation categories within the different DMRs to DEGs correlation groups i) shows distribution of functional genome annotation of long-term cultivation effect, whereas ii) shows distribution of functional genome annotations of Aza-treatment effect. Distribution of anti-correlating DMRs is shown on the left, while the distribution of correlated DMRs is shown on the right (CpG – CpG islands, UTR – untranslated region, TSS – transcriptional start site, TTS – transcription termination site).

In this study, long-term cultivation gave rise to 112 (51.38%) aDMRs and 106 (48.62%) cDMRs, whereas Aza-treatment revealed 1031 (84%) aDMRs and 199 (16%) cDMRs. Functional genome annotation of aDMRs and cDMRs showed that most prominent differences were found in introns, intergenic and promoter-TSS regions for long-term cultivation only. While aDMRs related to introns and intergenic regions markedly decreased to 14.12% and 11.17%, respectively, aDMRs related to promoter-TSS regions greatly increased to 7.87%, when compared to cDMRs (Figure 3.30 i). For Aza-treatment, aDMRs and cDMRs showed slight changes in CpG islands (-3.38%),

promoter-TSSs (+5.11%), introns (+5.79%) and exons (-1.83%), whereas more pronounced changes were found in aDMRs related to intergenic regions (-11.49%), in comparison to intergenic cDMRs (Figure 3.30 ii).

i)

ii)

Anti-correlated DMRs
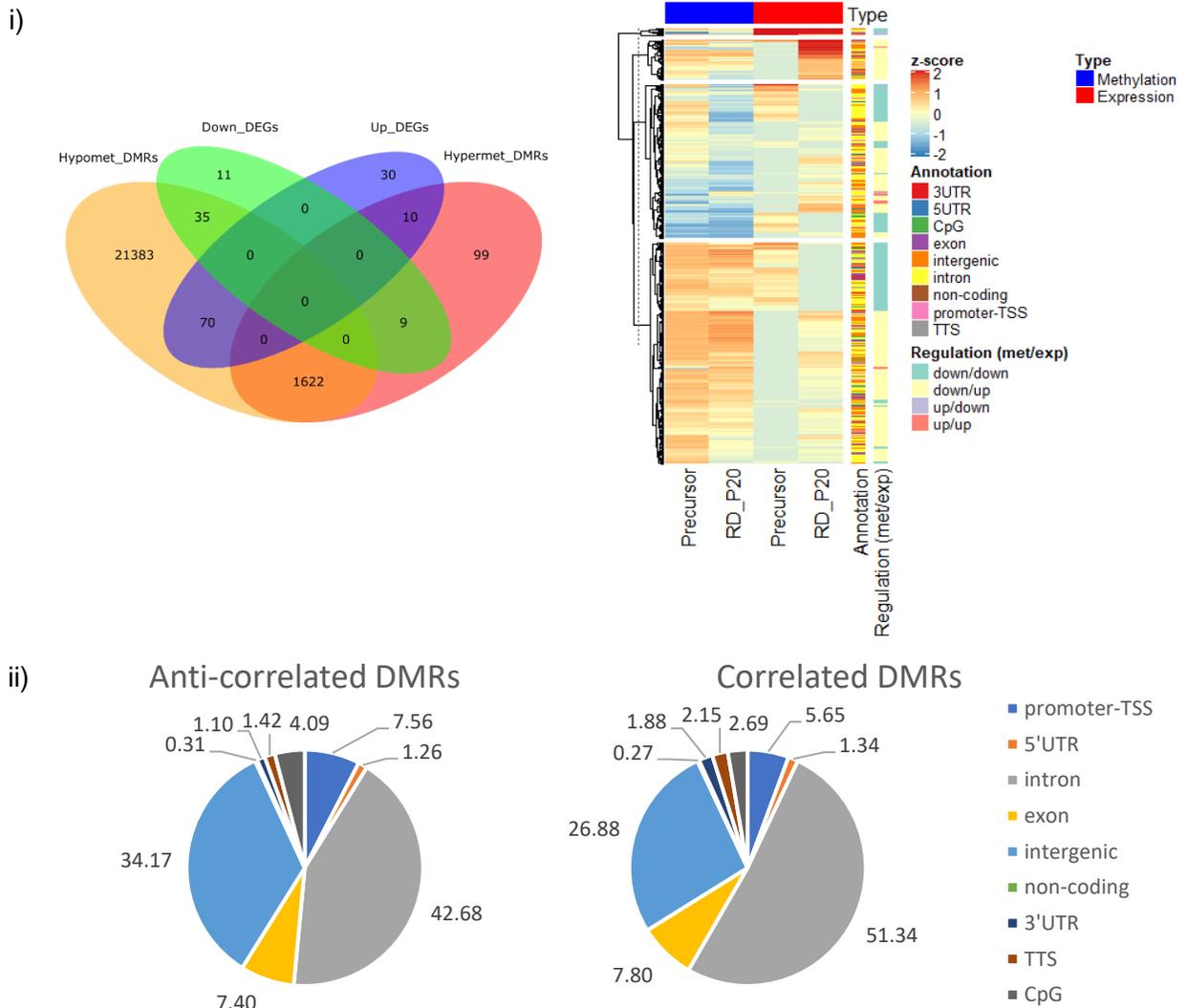
Correlated DMRs

Figure 3.31 Correlation of DMRs and DEGs after PCi-differentiation of reprogrammed HepaRG. i) Venn diagrams (left) and heatmaps (right) show the overlapped genes between hypo/hypermethylatedDMRs and up/down regulated DEGs. Heatmap of methylation (blue) and expression (red) was drawn based on row Z-score. Gradient colors from red to blue represent the value of z-score from high to low, respectively. Types of genome annotations related to DMRs and regulations are also defined on the right side of the heatmap. ii) Pie charts show distribution of genome annotation according to anti-correlation DMRs (left) and correlation DMRs (right) (CpG – CpG islands, UTR – untranslated region, TSS – transcriptional start site, TTS – transcription termination site).

During the PCi-differentiation process, overlapping of DEGs and DMRs between RD_P20 and precursor showed 124 genes that correspond to 1007 DMRs, 19 genes correlated to

hypermethylated DMRs and 105 genes correlated to hypomethylated DMRs. Looking at DEGs, there were 80 DEGs found to be up-regulated and 44 DEGs found to be down-regulated (Figure 3.31 i).

Out of the 1007 observed DMRs 635 (63.06%) aDMRs and 372 (36.94%) cDMRs were found. Genomic annotation analysis showed most prominent differences of aDMRs and cDMRs in intronic (8.66%) and intergenic (7.29%) regions, respectively (Figure 3.31 ii).

3.1.12. Technical validation of methylation and expression data

To validate the results obtained by RRBS and get further insight into the methylation profile of prominent DMRs, bisulfite local deep sequencing (LDS) on an Illumina MiSeq was performed (Gries *et al.,* 2013). DMRs related to 11 genes were selected from each branch of the PC experiment and primers were designed across DMRs (Chapter 2, Section 2.2.8). All amplicons were sequenced and analyzed using BiQ Analyzer HT software (Lutsik *et al.*, 2011).
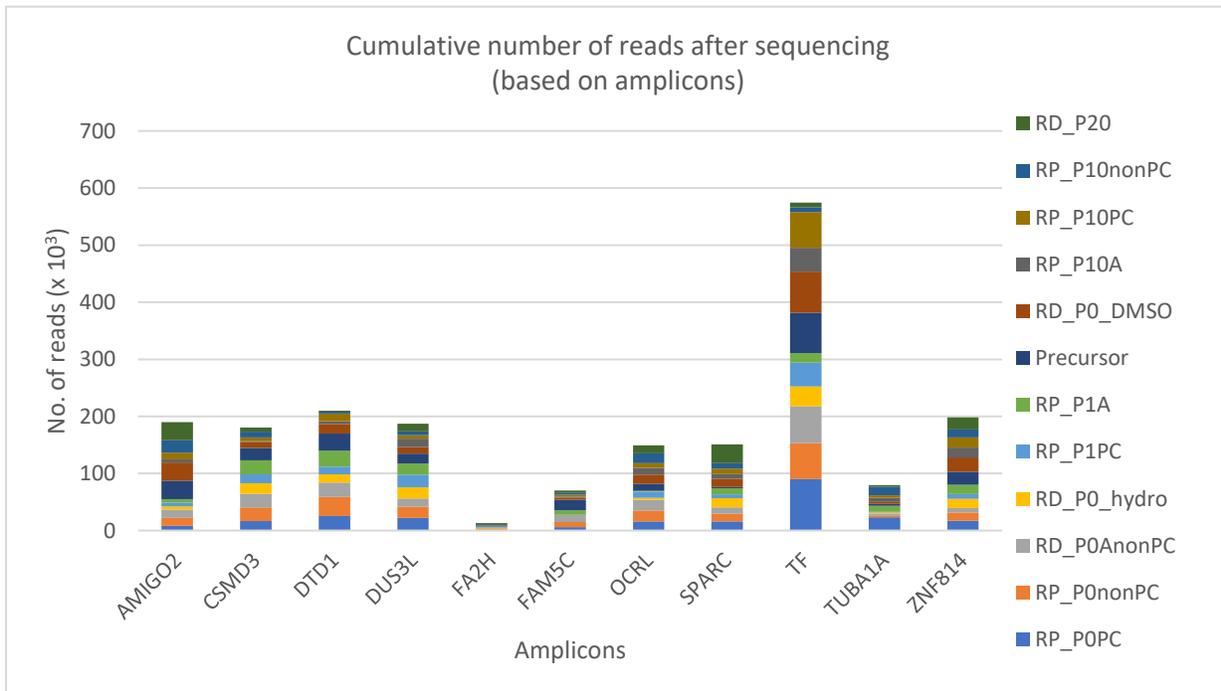
Figure 3.32 Cumulative number of reads after sequencing. Bar graph shows cumulative number of reads, based on amplicon. All samples are defined by colors on the right side of the graph.

The results showed that, for most amplicons, the cumulative number of reads obtained was higher than 100K, except for *FA2H, FAM5C* and *TUBA1A* (13K, 70K, 79K, respectively). With approx. 570K the number of reads was outstanding for the amplicon linked to *TF* (Figure 3.32).

The correlation between RRBS and LDS was performed at the CpG level. Therefore, the methylation of CpG sites present in both sequencing approaches (n = 1279) was considered. Plotting the overall results showed moderate correlation between both sequencing methods (R = 0.76). Overall, a tendency towards higher methylation in the local deep sequencing compared to RRBS was observed, probably due to some CpG sites being only sparsely covered in the RRBS data (cov. < 10x) (Figure 3.33 i).

Generally, a variation of correlation value between LDS and RRBS were observed among the selected amplicons (R = 0.14 - 0.90). The strength of correlation obviously was partially depending on the number of reads obtained from RRBS, e.g. CpGs within the OCRL amplicon revealed significant higher correlation when low coverage CpGs were removed (0.49 vs. 0.66) (Supplementary Table 7.2 and Supplementary Figure 7.25 - Figure 7.35). However, overall correlation was not improved even after low coverage CpGs (cov. < 10x) were excluded (R = 0.74, Supplementary Figure 7.9 i). Remarkably, the methylation levels of the CpGs within the *TUBA1A* amplicon were observed to be constant in RRBS but variable in the LDS data (Supplementary Figure 7.34 ii). After the CpGs within the *TUBA1A* amplicon were excluded from the data, correlation of overall CGs between LDS and RRBS was improved (R = 0.82, Figure 3.33 ii). Interestingly, the average of all CpGs within the LDS amplicons was highly correlated to RRBS regions (n = 130, R = 0.82), even when low coverage CpGs were still included in the analysis (Supplementary Figure 7.9 ii).

i)
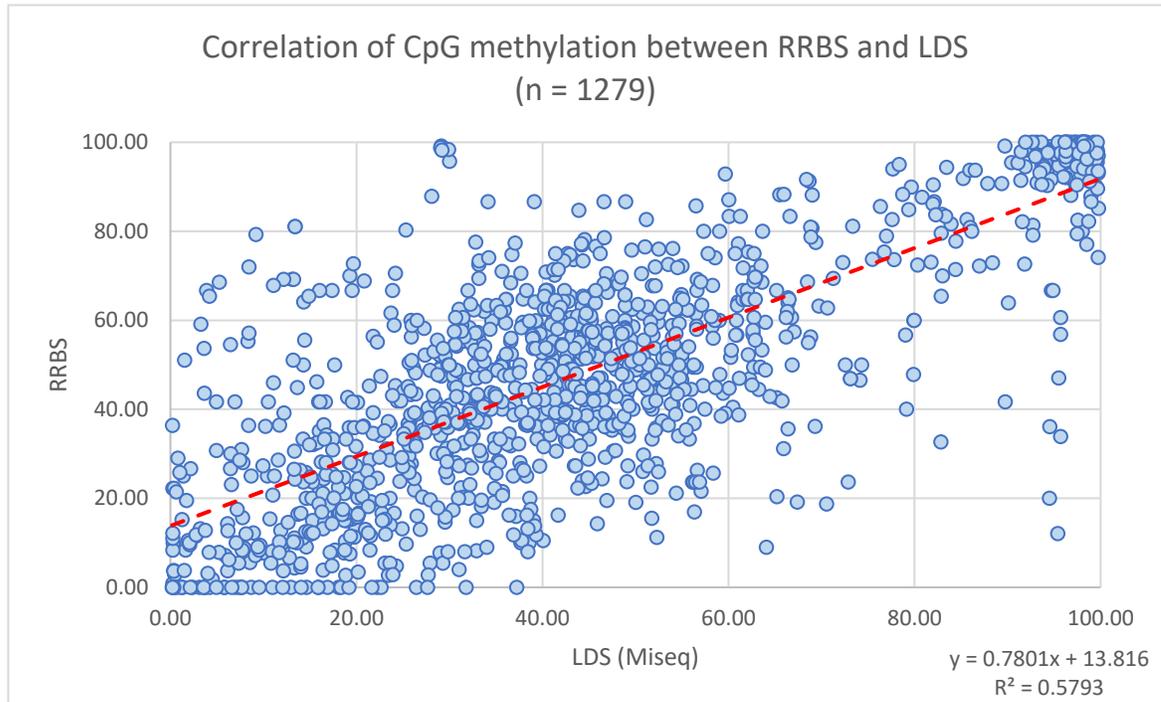
ii)
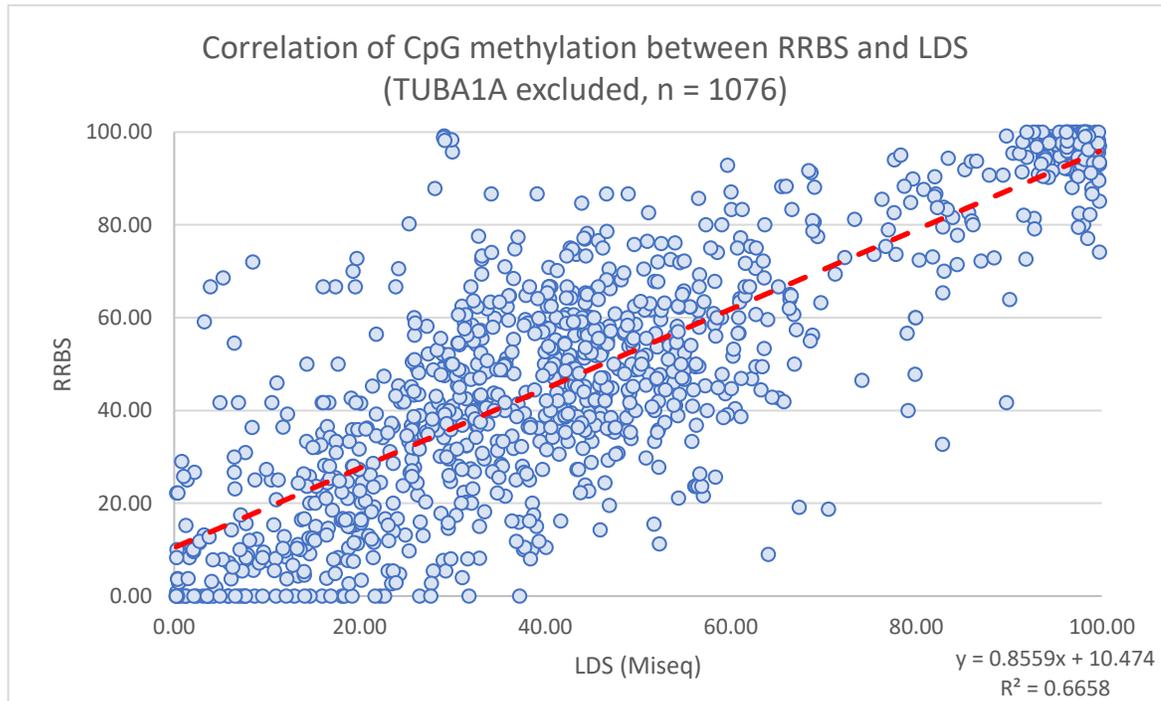
Figure 3.33 Correlation of methylation of all CpG sites present in the RRBS and LDS data sets. i) shows correlation of single CpG methylation between RRBS and LDS, whereas ii) shows correlation of single CpG methylation between RRBS and LDS after CpGs of *TUBA1A* amplicon were excluded. Each blue dot represents a CpG site of each amplicon, which is also present in the RRBS data. Red dash line presents a trendline according to equation and R^2 are also shown on the bottom right.

For the technical validation of expression data, 8 genes and 7 samples, except RD_P0DMSO, were selected to validate mRNA-seq results by reverse-transcription quantitative PCR (RT-qPCR). CT values were normalized to the house-keeping gene *UBE2D1*, then normalized CT values obtained from the precursor cells ($2^{-delta\ delta\ CT}$method, Livak and Schmidtgen, 2001). The results also showed 3 groups of expression changes, i) *ZNF330* showed slight expression changes in all samples, ii) *ETNK2, ALB and TF* were lower expressed in all samples, except in RD_P20, and iii) *CDK1, LGALS1, SPARC* and *TOP2A* showed higher expression, compared to the precursor (Figure 3.34 i). Moreover, overall results from RT-qPCR were comparable to mRNA-sequencing (R = 0.94), but log2 expression values from RT-qPCR were stronger than the values from mRNA-sequencing (Figure 3.34 ii).

i)

Log2 expression of selected genes in comparison with precursor

Legend: ETNK2, ZNF330, ALB, TF, CDK1, LGALS1, SPARC, TOP2A

Samples: RP_P0PC, RP_P0nonPC, RP_P10A, RP_P10PC, RP_P10nonPC, RD_P20

ii)

Correlation of log2 expression between mRNA-sequencing and RT-qPCR

y = 0.8973x + 0.6827
R^2 = 0.8769
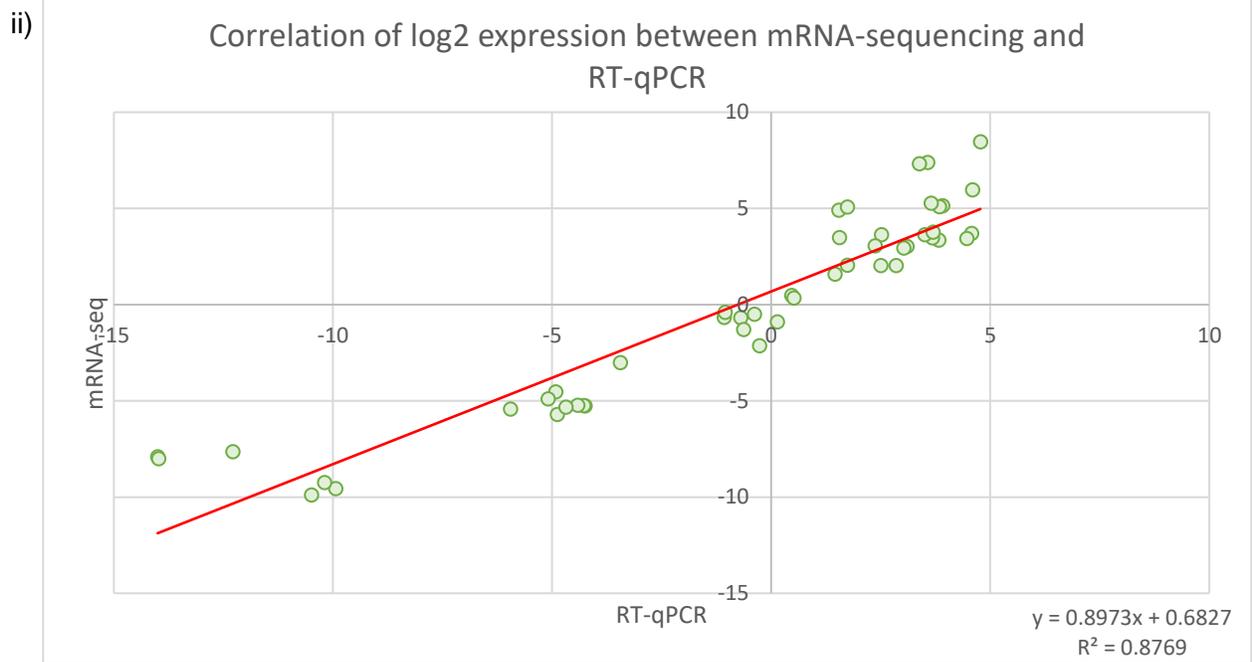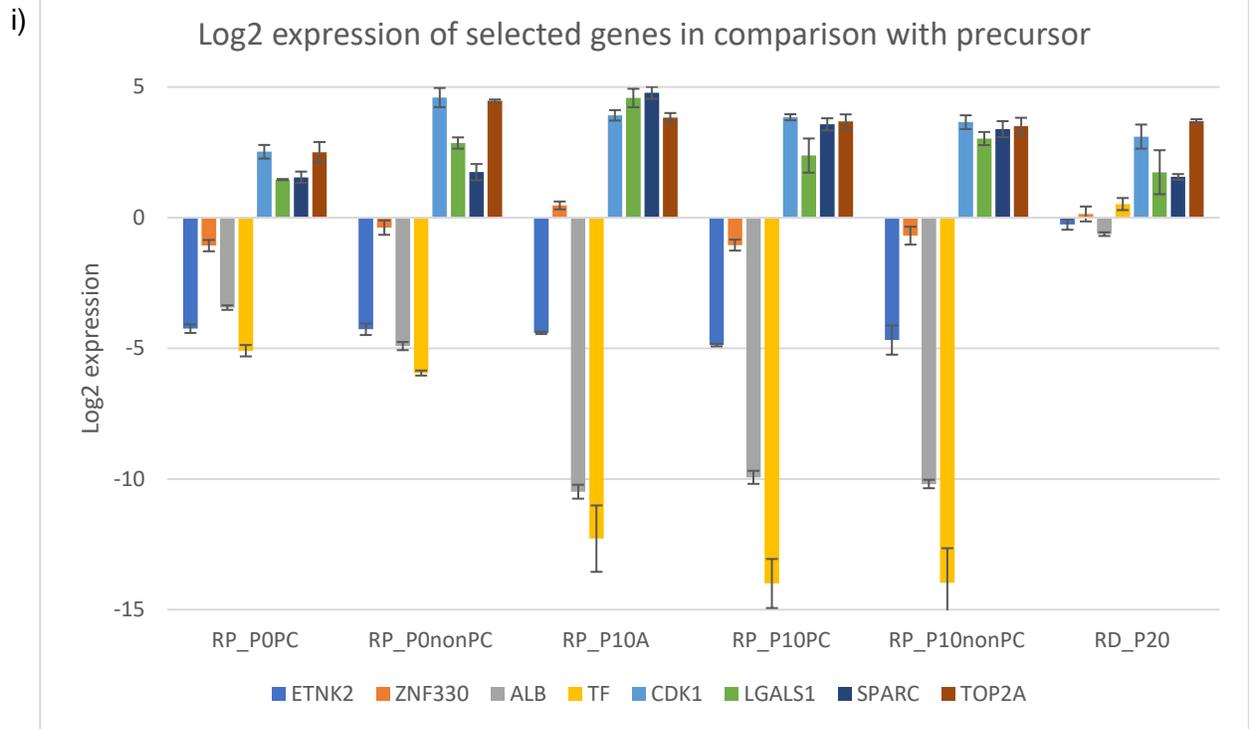
Figure 3.34 Correlation between mRNA-sequencing and RT-qPCR. i) Bar graph shows log2 expression of selected genes in comparison with precursor cells. Sample names are written below. ii) Correlation of log2 expression between mRNA-sequencing and RT-qPCR. Each dot represents log2 expression of selected genes of each sample, while red line represents a trendline. Equation and R^2 are also shown on the bottom right.

3.2. Controlled confinement

As the PC experiment was performed by Biopredic Co. (Dr. Christiane Guillouzo) with partially empiric experimental steps, a controlled confinement (CC) experiment was set up to mimic the PC experiment, particularly in reprogramming - PC state, in a well-documented and reproducible manner. The importance of CC lies in the controlled stress and stiffness of the cells by changing the microenvironment of culture conditions in a reproducible and controllable manner. Thus, the cells performing CC were squeezed through a gap between single pillars resulting in the restriction of movement and cell growth.

In the first experiment, confinement height was varied to select for the optimal CC condition. After squeezing the cells for 20h, immunostaining of 5mC and 5hmC found that the number of cells with high relative immunofluorescence signals (5mC/DAPI and 5hmC/DAPI) were increased in the squeezed cells of 2.7 and 3.5 μm height, compared to the non-squeezed cells. However, the number of cells of both heights were different, since the 2.7 μm condition contained lower number of cells than the 3.5 μm condition (Supplementary Figure 7.10 - Figure 7.12). Consequently, after checking the IF signals in relation to the cell number, a pillar height of 3.5 μm was chosen for all subsequent CC experiments.
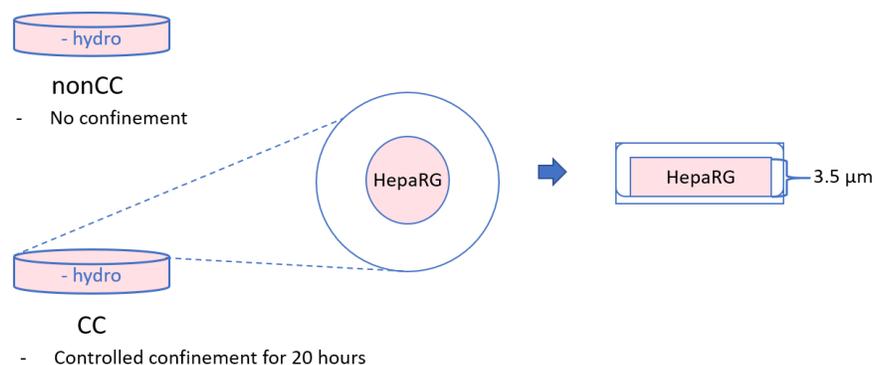
Figure 3.35 Experimental scheme of controlled confinement experiment. Differentiated HepaRG precursor cells were seeded in 6-well plates coated with fibronectin (pink area in the middle of circle). Cells in CC plates were squeezed for 20 hours, similar to the PC experiment, whereas nonCC plates were used as control samples, without squeezing.

To prepare for the CC procedure, HepaRG precursor cells were cultured without hydrocortisone until they reached confluency. HepaRG were sub-cultured into 6-welled plates for $2x10^5$ cells/well, but the cells were attached only in the middle of the well. For CC, cells were squeezed using pillars at 3.5 μm for 20 hours. As non-CC control, cells were grown in the very same way, without squeezing (Figure 3.35). The experiment was performed in technical duplicates and cells were harvested for nucleic acid preparations, DNA for RRBS and RNA for mRNA sequencing.

3.2.1. The alteration of DNA methylation after controlled confinement

After methylome sequencing, CC samples obtained higher number of reads than nonCC samples, but number of reads in all samples were between $1.5-2.0 \times 10^7$ reads (Supplementary Table 7.3). After methylation calling and 10x coverage filtering, both CC replicates yielded around 1.8 million CpG sites, while the nonCC replicates yielded 1-1.5 million CpGs (Figure 3.36).
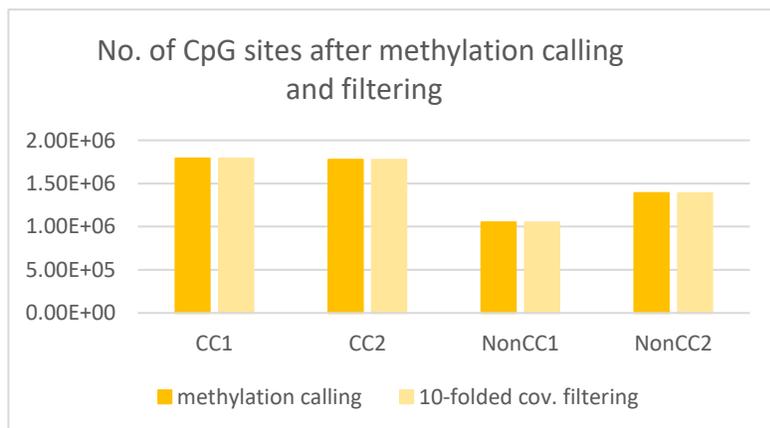
Figure 3.36 Number of CpG sites covered after methylation calling and filtering. Number of CpG sites after methylation calling is shown in dark color, while number of CpG sites after filtering is shown in light color.

The methylome data of CC and nonCC samples were analyzed by MethylKit with the same filtering parameters used for the analysis of PC experiment (see Page 59). Particularly, for CC experiment, standard T-test was used in DMRs analysis to extract DMRs that could be found in both technical replicates. CpG sites, which appeared in only one of the duplicates (CC/nonCC) and had a coverage less than 10-fold, were excluded from the analysis.
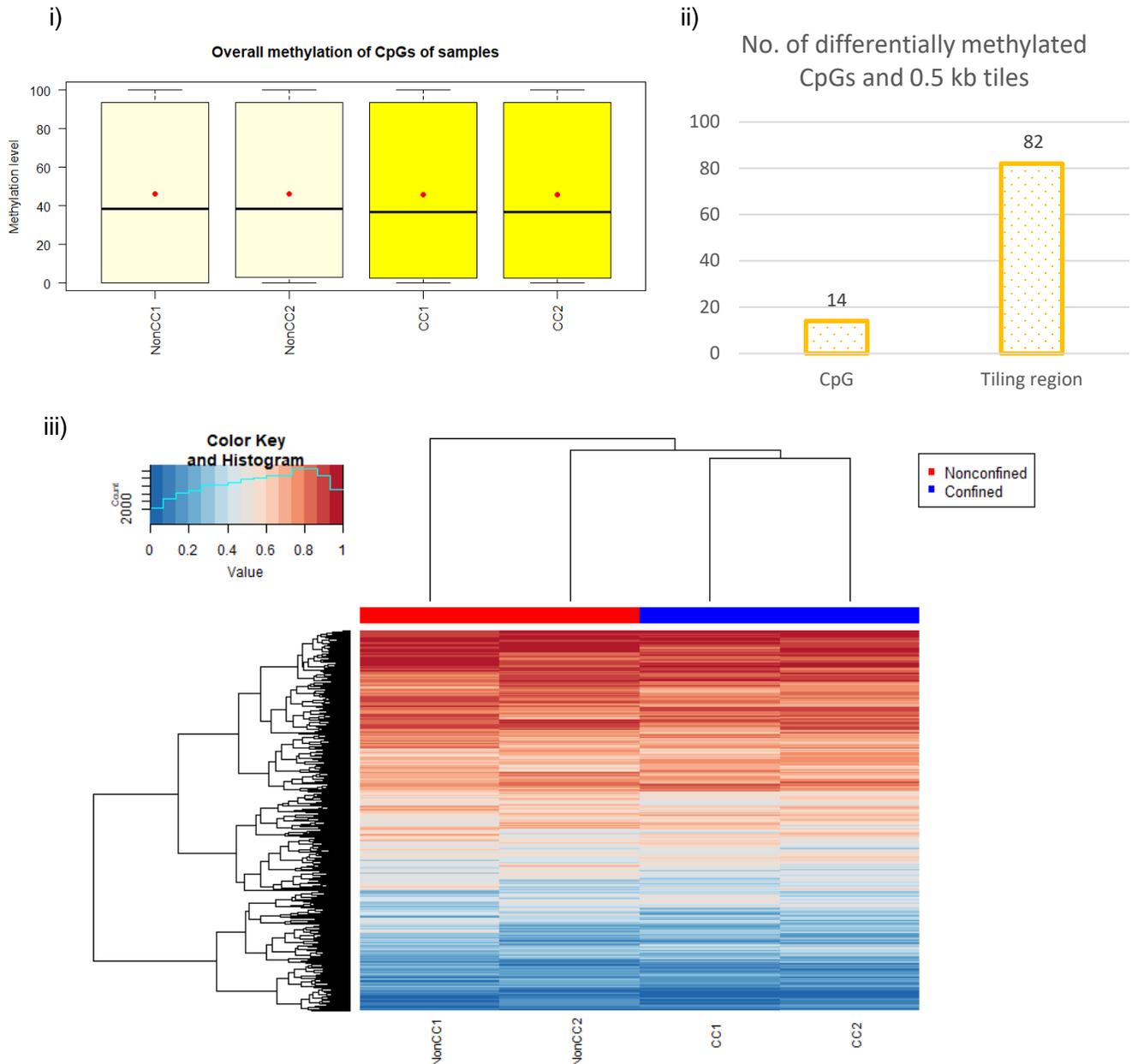
Figure 3.37 Overall methylation, DMCs and DMRs of CC experiment. i) Box plot shows overall methylation level of all samples. The black line and red dot are median and average of methylation levels, respectively. ii) Bar graph presents number of differentially methylated CpGs/regions. iii) Heatmap of 20K most variable 500 bp tiles. The colors, blue and red, represent the methylation level from 0 to 1. CC and nonCC samples are defined by colors presented on the top right.

In this study, the differential methylation analyses were computed at single site level and 500bp tiling region level. At single CpG methylation level, there were 638,220 sites covered throughout the genome. Average methylation of CC samples was found in high similarity to nonCC samples (Figure 3.37 i, CC: 45.8, nonCC: 46.4). The clustering distinguished CC from nonCC (Figure 3.37

iii). The heatmap of 20k most variable regions revealed highly methylated DMRs rather than low methylated DMRs in both CC and nonCC samples.

Out of 638,220 CpG sites, only 14 differentially methylated CpGs (DMC) were observed, consisting of 6 hypermethylated (40-58% mean difference) and 8 hypomethylated DMCs (39-87% mean difference). For 500 bp tiling regions, out of 72,860 tiles covering the whole genome only 82 DMRs were found. Among those, 32 DMRs were hypermethylated (10-33% difference), whereas 50 DMRs were hypomethylated (10-34% difference) (Figure 3.37 ii).
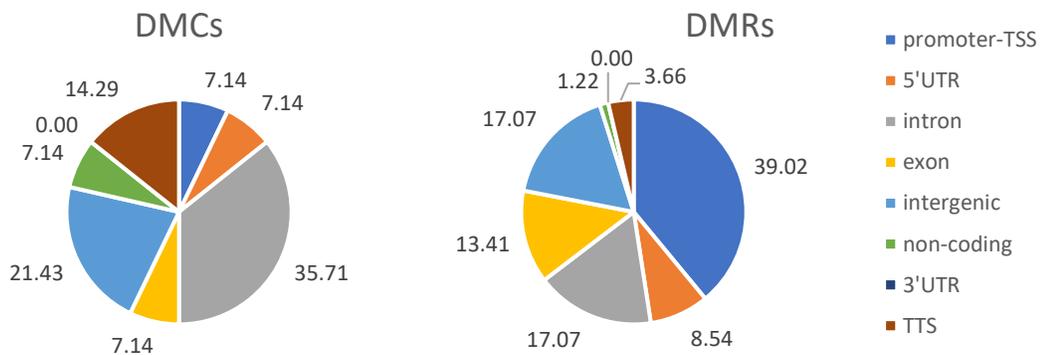
Figure 3.38 Functional annotation of DMCs and DMRs. Pie charts show distribution of DMCs (left) and DMRs (right) according to functional genome annotations (UTR – untranslated region, TSS – transcriptional start site, TTS – transcriptional termination site).

Genomic annotation analysis showed distinctive patterns among DMCs and DMRs. DMCs were found to be prominent in introns (35.71%) and intergenic regions (21.43%), whereas DMRs were mostly found in promoter-TSS regions (39.02%). However, no DMCs and DMRs were found in 3'UTR regions (Figure 3.38).

3.2.2. The alteration of gene expression after controlled confinement
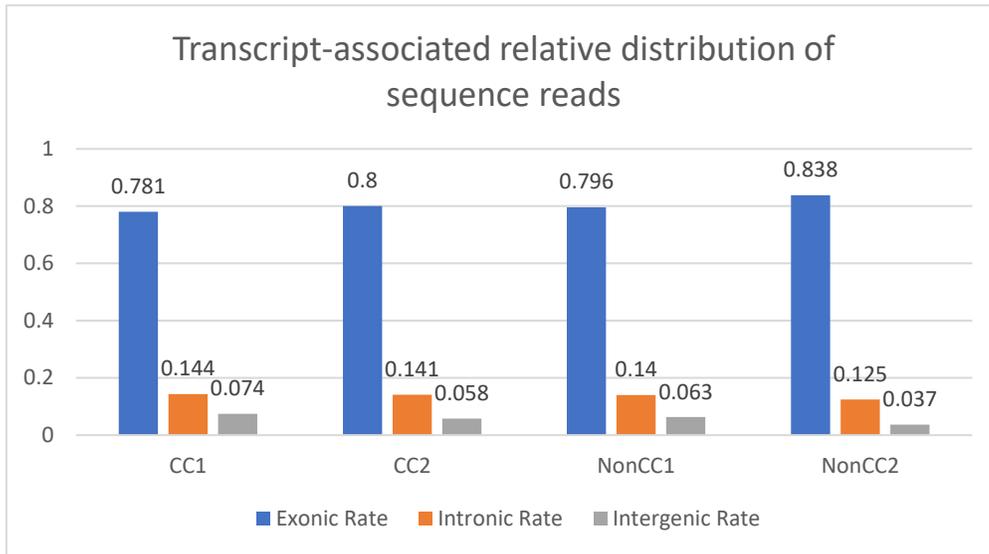
Figure 3.39 Transcript-associated relative distribution of reads of CC samples. The bar diagram shows the relative read distribution in exons, introns and intergenic regions for all samples in different conditions.

The overall obtained sequence reads of mRNA sequencing for all samples were slightly variable and found to be between 1.9-5.0 x10^7 reads (average 3.3 x 10^7 reads, see Supplementary Table 7.3). However, the number of genes mapped to the reads was similar (2 x 10^4 genes approx.) between the samples, as can be seen from the transcript-associated reads ratio (Figure 3.39 and Supplementary Figure 7.13).
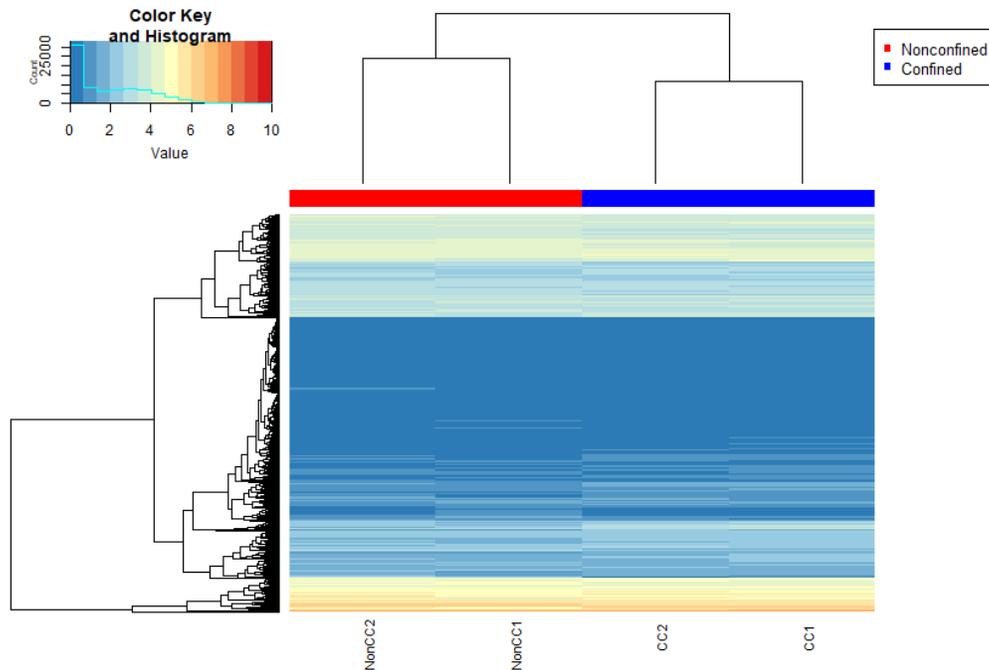
Figure 3.40 Heatmap of 20k variable expressed genes. A gradient of blue to red represents log (FPKM+1) value from low expression to high expression. CC and nonCC samples are defined by colors presented on the top right.

The heatmap of 20k most variable expressed genes revealed minor differences between CC and nonCC samples. However, the clustering was distinct between CC and nonCC samples (Figure 3.40).

Pairwise comparison between CC and nonCC samples were grouped and analyzed by DESeq2 (Love *et al.,* 2014) package. Genes with FDR adjusted p-value less than 0.01 were DEGs, the same as classified in the PC experiment. The result showed that there were 193 down-regulated genes and 381 up-regulated genes. GO terms and KEGG pathways were analyzed for the top 100 significantly up- and down-regulated DEGs. Remarkably, GO terms of upregulated DEGs frequently showed biological processes related to stresses responses, cell motility and signal transductions. For the DEGs grouped into signal transduction processes found in GO analyses, several KEGG pathways were identified which correlated to signal transduction as well, such as the IL-17 signaling pathway, TNF signaling pathway and NF-kappa B signaling pathway etc..

Unfortunately, no significantly enriched GO terms and KEGG pathways were observed for downregulated DEGs (Figure 3.41).

Relevant GO terms related to top 100 upregulated DEGs

GO term	p-value
GO:0030155 regulation of cell adhesion	6.50E-03
GO:0006355 regulation of transcription, DNA-templated	5.10E-03
GO:2000026 regulation of multicellular organismal development	2.00E-03
GO:0060548 negative regulation of cell death	9.00E-04
GO:0007165 signal transduction	1.50E-04
GO:0030335 positive regulation of cell migration	3.85E-05
GO:0006950 response to stress	6.17E-06
GO:0032101 regulation of response to external stimulus	3.97E-06
GO:0032496 response to lipopolysaccharide	2.78E-10
GO:0006954 inflammatory response	2.78E-10

No. of gene

Relevant KEGG pathways related to top 100 upregulated DEGs

KEGG pathway	p-value
hsa04380 Osteoclast differentiation	4.59E-02
hsa04620 Toll-like receptor signaling pathway	2.86E-02
hsa04640 Hematopoietic cell lineage	2.54E-02
hsa05204 Chemical carcinogenesis	1.59E-02
hsa00980 Metabolism of xenobiotics by cytochrome P450	1.44E-02
hsa05146 Amoebiasis	3.10E-03
hsa04064 NF-kappa B signaling pathway	1.68E-06
hsa04668 TNF signaling pathway	2.78E-07
hsa04060 Cytokine-cytokine receptor interaction	1.27E-08
hsa04657 IL-17 signaling pathway	2.30E-13

No. of gene

Figure 3.41 GO terms enrichment and KEGG pathways of DEGs related to top 100 upregulated DEGs of controlled confinement. The upper part is a bar graph of GO terms, while the lower part belongs to KEGG pathways. FDR-adjusted p-values are written at the end of bars.

3.2.3. Correlation between methylation and expression of CC experiment

	Methylation		Expression	
	CC	NonCC	CC	NonCC
GEM	77.11	60.63	3.16	4.35
CDCP1	32.28	46.4	6.14	5.06
SFMBT1	29.73	52.98	3.37	4.37
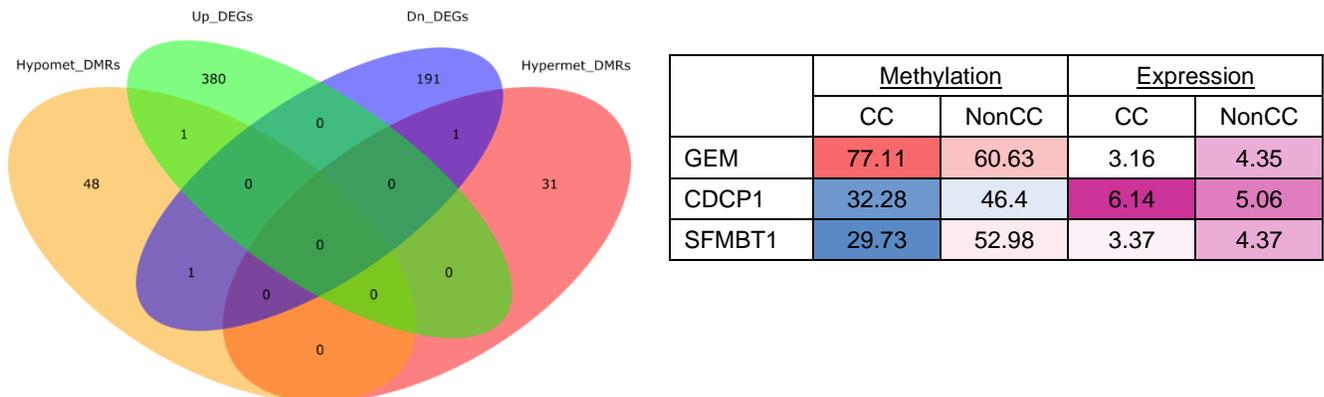
Figure 3.42 Correlation of DMRs and DEGs of controlled confinement. Venn diagram on the left shows two overlapping genes between hypo/hypermethylated DMRs and up/down regulated DEGs. Table on the right shows methylation and expression of *GEM*, *CDCP1* and *SFMBT1*. The colors, dark blue to dark red, represent the methylation level from low to high. A gradient of purple from light to dark represents the expression level from low to high as well.

A strong correlation between methylation and expression of genes after CC was not observed since there were only 3 DEGs corresponding to 3 DMRs. Those three DEGs included *GEM* – a protein receptor of plasma membrane, *CDCP1* – a transmembrane protein involved in tumor invasion and metastasis, and *SFMBT1* – a transcriptional corepressor involved in myogenesis. From the analysis, *GEM* and *CDCP1* were found to be aDMR, whereas *SFMBT1* was found as cDMRs. Notably, while the DMRs of *SFMBT1* located in promoter-TSS region, the DMRs of *GEM* and *CDCP1* located in the 1st intron and the 1st exon, respectively (Figure 3.42).

3.2.4. Summary of methylation and expression changes during controlled confinement

Regarding DNA methylation, only minor changes of DNA methylation were revealed between CC and nonCC. Low methylation and high methylation profiles were equally observed in both groups. Moreover, number of DMCs and DMRs was extremely low. Genome annotation revealed that most DMRs were in promoter-TSS regions, whereas most DMCs were in intronic regions.

Similar to DNA methylation, low differences in mRNA expression between CC and nonCC were observed. Number of DEGs grouped into higher and lower expressed genes were slightly different as number of downregulated DEGs were moderately lower than number of upregulated DEGs.

GO terms enrichment of DEGs and presence in KEGG pathways of upregulated DEGs were found in Table 3.2.

Table 3.2 Summary of GO terms and KEGG pathways related to DEGs of CC experiment.

GO terms and KEGG pathways	
Up-regulation	**Down-regulation**
Stresses responses, Cell motility, Signal transductions and Development	-

3.3. Comparison between PC and CC experiments

3.3.1. Comparison of methylation and transcription data between PC and CC experiments

To compare DNA methylation data between PC and CC, CpG sites and 0.5 kb tiling regions have to be overlapped between both experimental series. Although 77.92% of CpG sites and 96.23% of tiling regions of CC were found overlapping in PC (Figure 3.43), only 11 common DMRs (9.5%) between PC and CC were found (3 hypermethylated DMRs and 8 hypomethylated DMRs). Those DMRs were found in promoter-TSS (5 DMRs), introns (4 DMRs) and exons (2 DMRs).

Figure 3.43 No. of CpG sites and tiling regions of PC and CC experiment after 10x coverage filtering. The overlapping of CpG sites is shown on the left, while the overlapping of tiling regions is shown on the right. The colors of circle, pink and blue, represents PC and CC experiments, respectively.

The overall methylation data clearly separated CC and PC samples by PC1 and PC4 in the PCA plot. PC1, which was the highest proportion, separated PC from CC experiment, while PC4 seemed to separate PC/CC from controlled counterparts (nonPC/nonCC). However, the CC and nonCC samples remained in the large cluster (Figure 3.44 i). This result was confirmed by clustering the 20k most variable regions. The PC samples were all clustered in a different branch compared to CC and NonCC (Figure 3.44 ii).
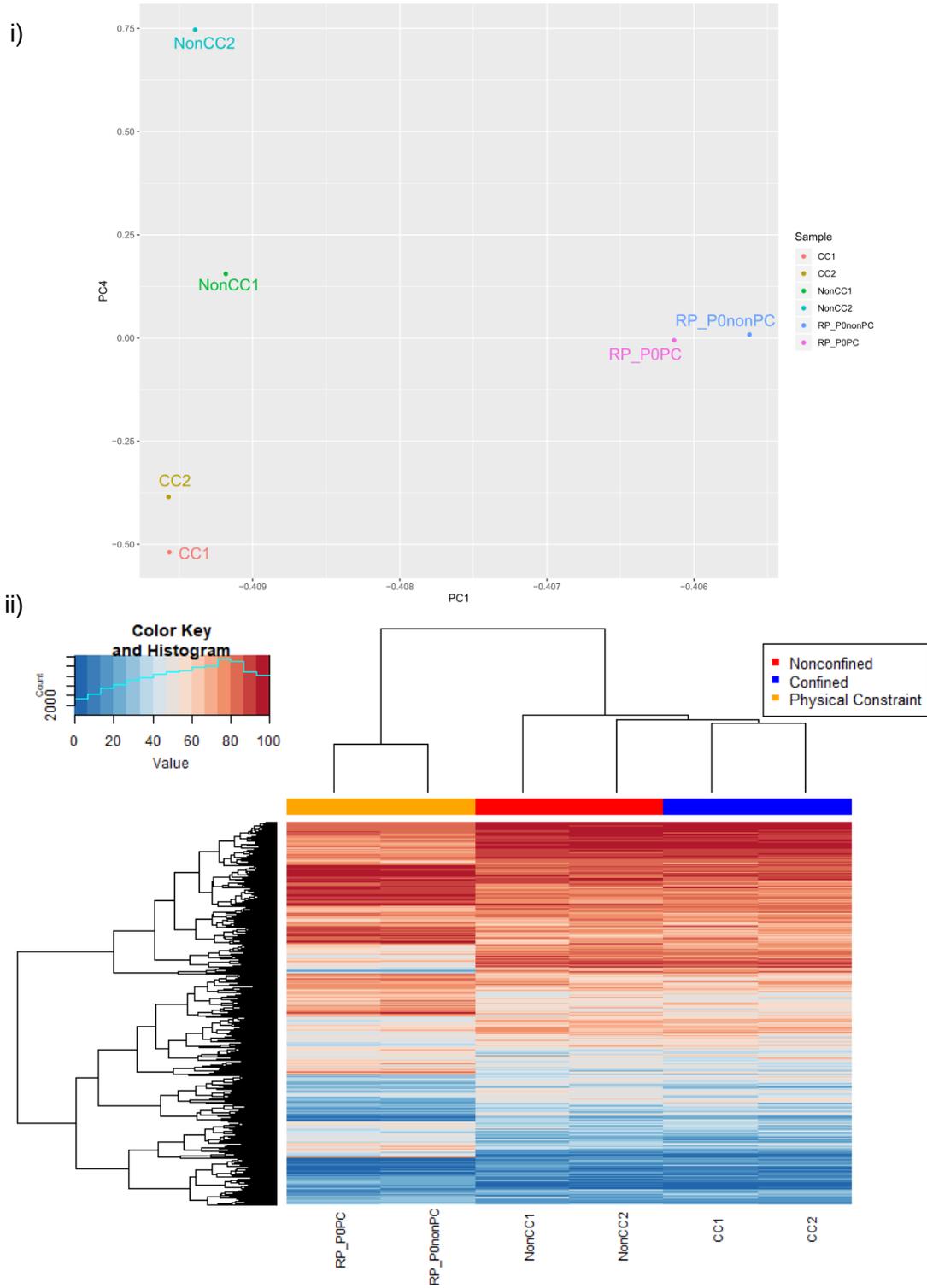
Figure 3.44 Principal component analysis (PCA) and heatmap of 20K DMRs in PC and CC. i) showing PCA of PC and CC samples according to experimental procedures. ii) showing heatmap of 20K most variable 500 bp tiles. The colors, blue and red, represent the methylation level from 0 to 1. PC, CC and nonCC samples are defined by colors presented on the top right.

Comparison of mRNA-seq data revealed that PC samples were separated from CC and nonCC by PC2, while both control samples (nonCC/nonPC) were separated from CC/PC samples by PC3 (Figure 3.45 i) However, PC1 did not separate neither experiments nor samples. Interestingly, the overall expression of HepaRG cells used in the CC experiment seemed comparable to the cells used for the PC experiment. Besides, the technical replicate of each group clustered together suggesting a high reproducibility of the CC experiment (Figure 3.45 ii).
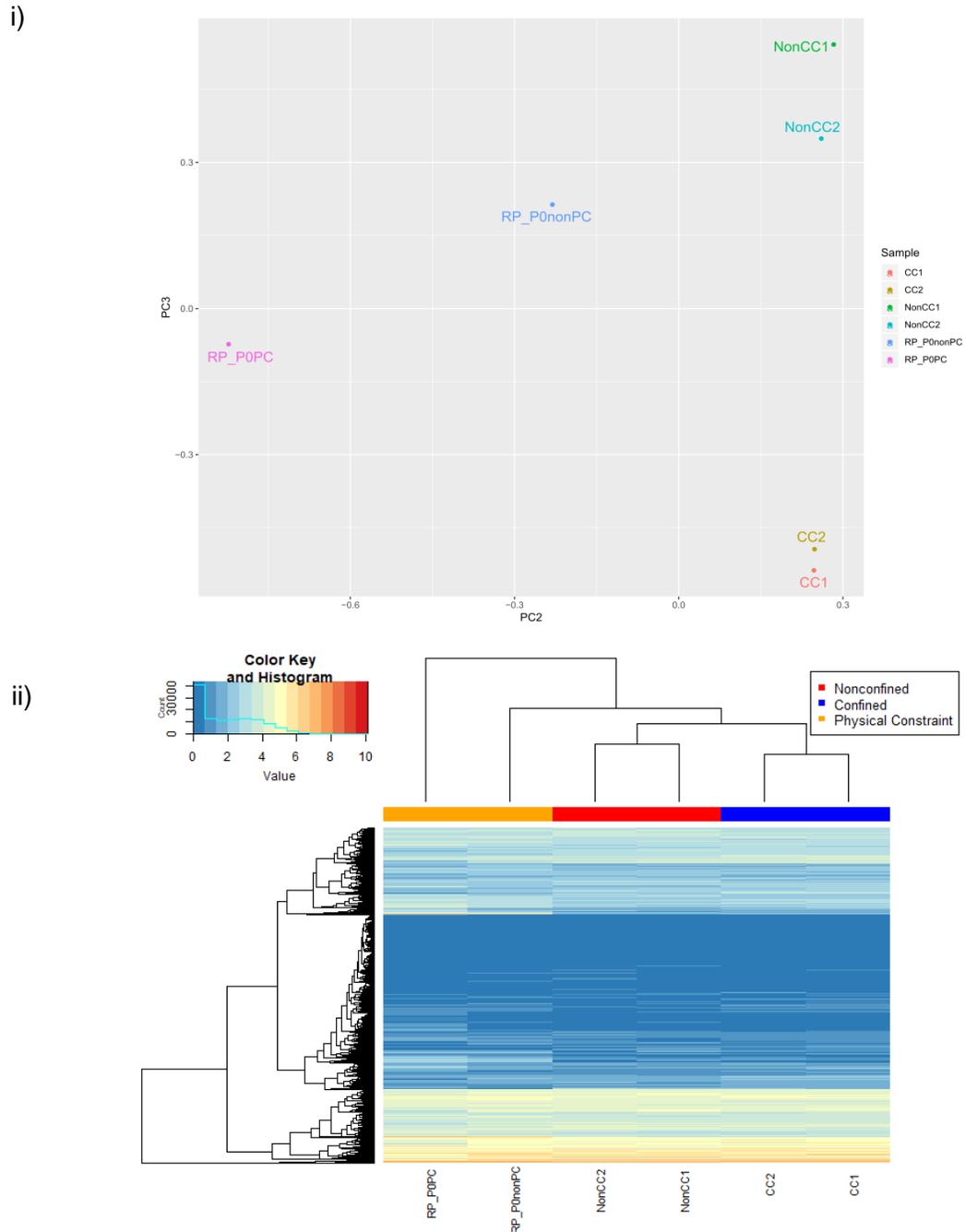
Figure 3.45 Principal component analysis (PCA) and heatmap of 20K DEGs in PC and CC experiment. i) showing PCA of PC and CC samples separately according to the experimental procedures and treatment (PC/nonPC or CC/nonCC). ii) showing heatmap of 20k variable expressed genes. A gradient of blue to red represents log (FPKM+1) value from low expression to high expression. CC and nonCC samples are defined by colors presented on the top right.

Pairwise comparison of PC and CC revealed DEGs for 103 genes and 574 genes, respectively. The CC experiment showed higher number of DEGs than the PC experiment. As shown in Figure 3.46, there is a slight tendency towards up-regulation of genes in both CC and PC experiments. Notably, three genes were found to be down-regulated in PC, but became up-regulated in CC, namely, *CYP1A1* encoding an enzyme in xenobiotic metabolism, *KRTAP2-3* encoding keratin associated protein 2-3 and *SERPIN2B* encoding a protein inhibitor of serine/cysteine proteases.
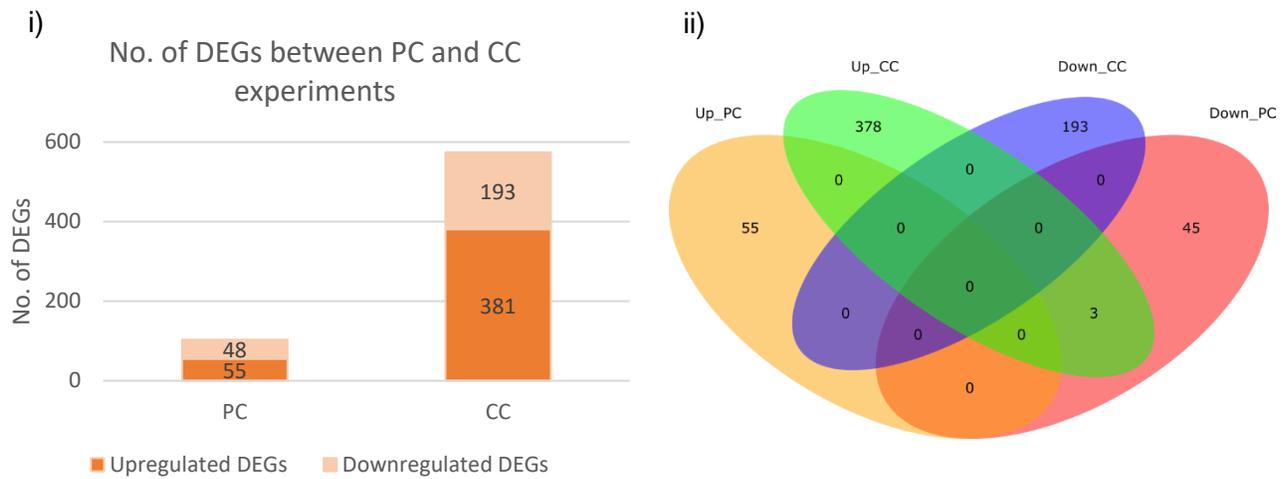
Figure 3.46 Comparison of DEGs obtained from PC and CC experiments. i) Bar graph shows number of up- (dark color) and down-regulated DEGs (light color). ii) Venn diagram shows two overlapping DEGs between up- and down-regulation of PC and CC experiment.

3.3.2. Comparison of cell type-specific, pluripotency-associated genes and epigenetic modifiers between PC and CC experiment

PC and CC experiments were compared to each other choosing the same gene sets in the same level.
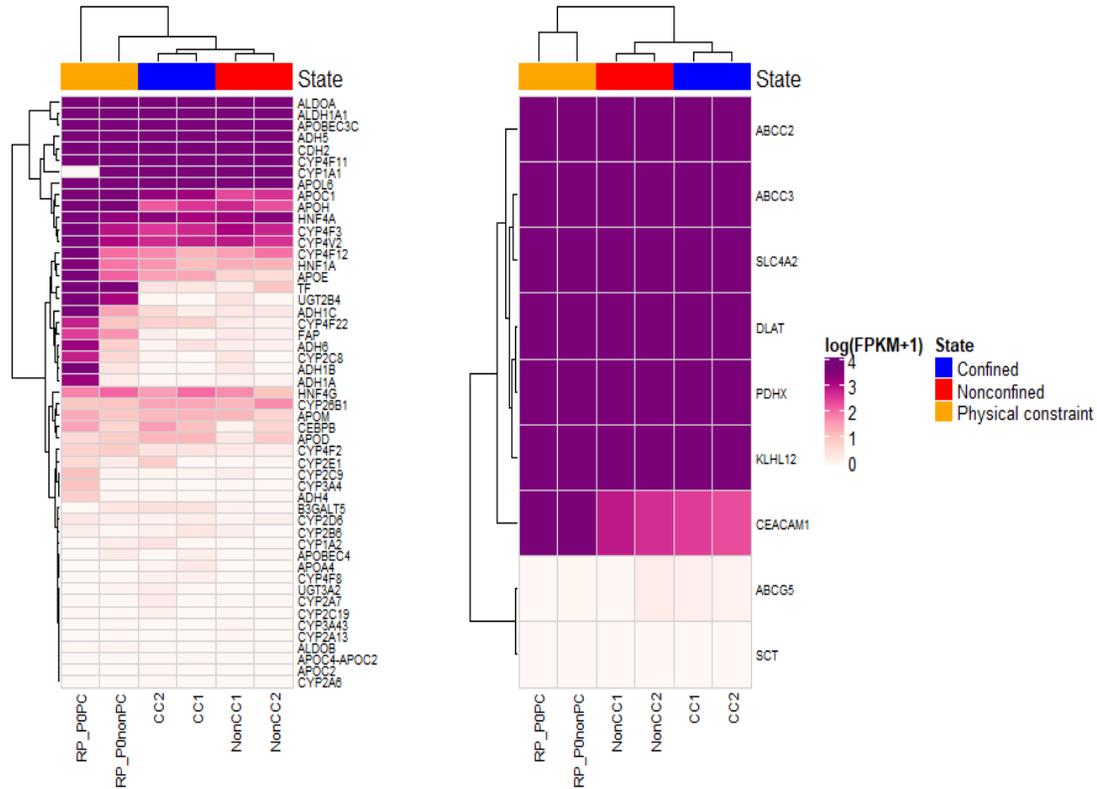
Figure 3.47 mRNA expression of cell type-specific genes. Heatmap shows the expression value (log(FPKM+1)) of hepatocyte-specific genes (left) and biliary-specific genes (right). Gradient colors from white to purple represent low to high expression, respectively.

According to the heatmaps, expression of hepatocyte-specific genes varied among samples, whereas expression of biliary-specific genes revealed to be more similar between PC and CC. Clustering of PC and CC experiment was observed solely in biliary-specific genes since nonPC sample clustered with all CC samples, in the large branch. Notably, PC samples showed stronger expression of tissue-specific genes compared to CC samples (Figure 3.47).
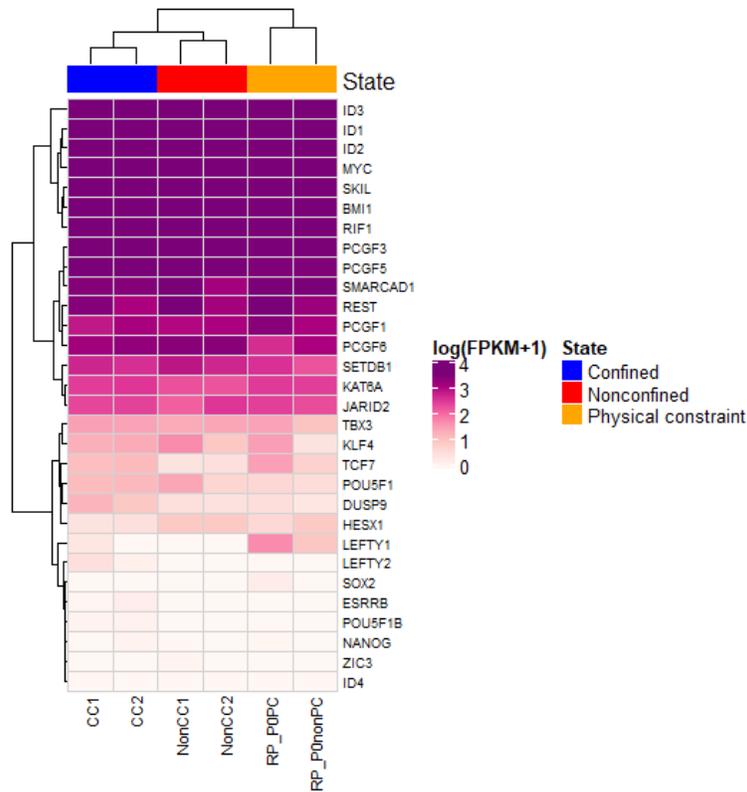
Figure 3.48 mRNA expression of pluripotent marker genes. Heatmap shows the expression value (log(FPKM+1)) of pluripotent genes. Gradient colors from white to purple represent low to high expression, respectively.

For pluripotency genes, PC and CC samples were clustered in different branches. Highly expressed genes were comparable among those samples and related to genes in IDs family (DNA-binding protein inhibitor), *MYC* and *SKIL* etc. Moreover, defined transcription factors, which were commonly used in the induction of epigenetic reprogramming of cells such as *KLF4, POU5F1, SOX2* and *NANOG,* were weakly expressed as well as early reprogramming markers, *DUSP9* and *ESRRB* (Figure 3.48). Furthermore, the signaling pathways regulating pluripotency such as JAK-STAT, WNT and TGFβ pathways were also found comparable among the samples. Clustering of those pathways revealed a clear separation between PC and CC experiment, but not between PC/CC and nonPC/nonCC (Supplementary Figure 7.14).

Figure 3.49 mRNA expression of epigenetic modifiers. Heatmap shows the expression value (log(FPKM+1)) of genes related to DNA methylation (left) and genes related to histone modifications (right). Gradient colors from white to purple represent low to high expression, respectively.

For DNA methylation modifiers, stronger expression was observed in CC experiment, although clustering of CC samples vs. nonCC was not detected. On the other hand, genes related to histone modifications showed minor differences between PC and CC. Notably, the cluster of NonPC samples in genes related to histone modifications was more similar to nonCC than PC (Figure 3.49).

Figure 3.50 mRNA expression of genes related to different kinds of histone modifications. Heatmap shows the expression value (log(FPKM+1)) of genes related to histone acetylation (left) and genes related to histone methylation (right). Gradient colors from white to purple represent low to high expression, respectively.

Genes related to histone modifications were grouped into histone acetylation modifiers and histone methylation modifiers to figure out if one of those groups had an impact on clustering. Heatmap of genes related to different kinds of histone modifications revealed partial clustering of PC against CC. Genes related to histone acetylation separated CC and PC samples from each other. Surprisingly, genes related to histone methylation separated PC/CC samples from nonPC/nonCC samples (Figure 3.50). This confirmed the results of the previous PCA analysis (PC3) that genes related to histone methylation separated the samples of the treatment vs. non-treatment (PC/CC against nonPC/nonCC), and not by the used experimental procedure (see Figure 3.45 i).

3.3.3. <u>Summary of comparison between PC and CC experiments</u>

Reads obtained from the RRBS experiments were different in number between the PC and CC experiment, concerning raw sequencing reads and filtered reads. Therefore, the overlap of CpGs and tiling regions were only around 20-30%. Only 11 DMRs were found to be common between PC and CC, the methylation profiles of PC and CC experiment were comparable. PC1 separated PC samples from all CC samples, while PC4 seemed to separate PC/CC samples from nonPC/nonCC samples.

Regarding transcriptome data, although raw sequencing reads of PC and CC were comparable, large differences of DEGs were still detected in PC and CC comparisons. *CYP1A1, KRTAP2-3* and *SERPINB2* were down-regulated genes that were common in PC and CC. Furthermore, the expression profiles showed similar patterns between PC and all CC samples. PC2 separated PC samples from all CC samples and PC3 separated PC/CC samples from nonPC/nonCC samples, respectively.

The expression of cell-type specific and pluripotency genes together with epigenetic modifiers showed strong or moderate expression in both PC and CC samples. Notably, genes linked to histone modifications, particularly histone methylation, showed distinct expression levels between PC/CC samples and nonPC/nonCC samples.

Chapter 4 Discussions and further directions

HepaRG is a non-cancerous cell line that take advantages in studies of hepatic functions, such as drug metabolisms, hepatotropic virus infection and stem cells etc., similar to primary hepatocyte cells (Gripon *et al.*, 2002; Parent *et al.*, 2004; Aninat *et al.*, 2006; Guillouzo *et al.*, 2007; Narayan *et al.*, 2009; Guguen-Guillouzo *et al.*, 2010). Several advantages and flexibility in use of HepaRG enabled us to study molecular mechanisms when this bipotent progenitor cell is induced to develop multipotent function and redirected differentiation. Therefore, this study focused on DNA methylation and gene expression during reprogramming and PC-induced differentiation of HepaRG. Using RRBS and mRNA sequencing allowed us to access genome-wide methylation and transcription data along the processes.

4.1. Quality control of RRBS and mRNA sequencing and technical validation

The detection of DNA methylation by RRBS provided representative methylome data throughout the experiment. Works by Meissner and others revealed that RRBS-*MspI* of mouse ES cells show a bimodal methylation pattern, as a huge number of reads was found either highly methylated (>80%) or lowly methylated (<20%). Moreover, *MspI* fragment reads contained at least one CpG site, and half of the reads were in CpG islands. In agreement with previous reports, a bimodal distribution of CpG sites regarding their methylation level in all samples of PC were also observed, except for Aza-treated samples. Instead of a bimodal distribution, a unimodal distribution, tending towards low methylation levels, was indeed observed in Aza-treated samples. Notably, reprogramming process did not increase the number of CpGs with mosaic methylation (20-80%), but mosaic methylation tended toward increasing when Azacytidine was applied to the culture (see Supplementary Figure 7.15).

Obtaining high read numbers for samples of CC experiment proved to be very difficult, since each multiwell confiner allowed only a very small cell number. Although cells from similar multiwell confiners were pooled collectively, the material for further sequencing library preparation was still

insufficient. Because of the limited concentration of those libraries, re-sequencing was not possible. However, a bimodal distribution of CpG sites against methylation levels in all CC samples could still be observed (see Supplementary Figure 7.16). Furthermore, CpGs with at least 10x coverage, both in PC and CC experiment, was discovered in CpG islands and other regulatory regions as well, similar to previous reports (Meissner *et al.,* 2008).

After mRNA-sequencing, the number of reads obtained for all PC and CC samples revealed various numbers of total read counts. Generally, a million to 20 million reads had been obtained from mRNA sequencing, with at least 60% of all exonic regions covered (Picelli *et. al,* 2013 and Picelli *et. al,* 2014). In both PC and CC studies, raw sequenced reads were obtained in expected range covering approx. 80% of exonic regions (see Figure 3.14 and Figure 3.39). Moreover, detected genes and transcripts were comparable among samples (see Supplementary Figure 7.4 and Figure 7.13). Other types of RNA, such as micro RNAs, non-coding RNAs and long non-coding RNAs etc., were also targets of library preparation and sequencing method. Although those RNAs play roles in promotion of pluripotentcy (Loewer *et al.,* 2010; Leonardo *et al.,* 2012; Flynn and Chang, 2014), the expression of those types of RNA was barely observed in our transcriptional profiles. Therefore, we decided to focus on protein coding genes only.

Technical validation of RRBS data using local deep sequencing (LDS) revealed a fairly strong correlation of CpG methylation levels between RRBS and LDS (R = 0.76), with the tendency to overall higher methylation levels obtained from LDS. Generally, low correlation coincided with low coverage in RRBS (coverage less than 10x). Unfortunately, after low coverage CpG sites were excluded, the overall correlation of CpGs between RRBS and LDS did not improve as expected (R = 0.74), even when considered for each amplicon separately. Most of them still showed a decrease of their correlation coefficient. This effect can be clearly observed in amplicons containing more than 20% of low coverage CpGs, e.g. within the amplicons linked to *TF* and *ZNF814* etc., suggesting that low coverage CpGs had a high impact on the correlation of CpG

119

methylation (see Supplementary Table 7.2). Furthermore, clonality of reads caused by PCR amplification of low template samples could be observed (*TUBA1A*, Supplementary Figure 7.34). After CpGs within the *TUBA1A* amplicon were excluded from the analysis, the correlation between RRBS and LSD results was stronger, as the correlation coefficient value increased from 0.76 to 0.82 (see Figure 3.33 ii). Remarkably, the correlation of average methylation across the amplicons to the CpGs present in the RRBS data was stronger than the correlation of single CpG methylation (R = 0.82), although the low coverage CpGs were still included(Supplementary Figure 7.9 ii). This points to neighbouring CpGs still belonging to the same regulatory domain as covered by MspI-RRBS which confirms RRBS to be a strong technique for the analysis of regulatory units like promoters or CpG-rich enhancers (Smith *et al.*, 2009; Bock *et al.*, 2010; Wang *et al.,* 2012).

Transcriptome data was validated by semi-quantitative reverse transcription PCR. RT-qPCR used specific primers to amplify those transcripts specifically, while in mRNA sequencing, universal poly-dT primers were used to amplify the library. Moreover, specific primers for RT-qPCR were designed across exon-exon boundary regions, which reduced false positive results from DNA contamination. Therefore, better sensitivity and specificity of RT-qPCR provided a better resolution leading to stronger log2 expression compared to mRNA sequencing. Anyway, strong correlation was obtained from RT-qPCR and mRNA sequencing results (R = 0.94).

Although technical validation was not performed in CC experiment because of the limited material, more strict filter criteria and evaluation parameters used in CC still give us an advantage in avoiding false positive results. For those reasons together with technical validation results, the sequencing data obtained from both PC and CC experiments were reliable and reasonable to use in this study.

4.2. The roles of PC and CC in the alterations of DNA methylation and gene expression during reprogramming process

PC and CC were experimental approaches used to study whether mechanical forces enabled cells to establish multipotent properties or not. In this study, we investigated DNA methylation and expression, molecular signatures of reprogrammed cells and compared them to other multipotent cells. Thus, the experiments, both PC and CC, were performed in the same manner.

The overall methylation level detected in PC and CC were almost similar, but the pattern of methylation profiles including 20k variable regions were slightly different. The methylation profiles observed in non-PC and non-CC provided distinctive patterns which might be caused by different batches of precursor cells used (see Figure 3.44 ii). The tendency of results between PC and CC were comparable; only minor changes on locus-specific DMRs were observed in force-applied samples, preferentially towards hypermethylation. However, as mentioned in Section 4.1 (see Page 118 on the 2nd paragraph), the number of DMRs of CC (n = 82) was extremely low, compared to PC (n = 6045), and solely 11 DMRs were found to be common. Low amount of DNA together with strict filter criteria could be the reason for this difference.

To investigate the observed difference in DMR detection further, integrated analysis between DMRs of PC and CC using ChromHMM was also performed and revealed extremely different results. DMRs related to PC were detected in inactive states and active states approx. 60% and 40%, respectively, while DMRs related to CC were detected prominently in active states, briefly more than 80% of DMRs (see Figure 4.1). Remarkably, our CC experiment was also performed in a different analysis procedure as described before. Although our study showed that DMRs related to chromatin states were preferentially found in the active states, histone modification changes were not observed directly. Additionally, a very low number of DMRs was also previously observed. Consequently, any implementation needed to be carefully considered.

Several studies about microenvironmental changes reported that mechanical forces were closely associated with the regulation of chromatin states and transcription (Chalut *et al.,* 2012; Pagliara *et al.,* 2014; Le *et al.,*2016; Tajik *et al.,* 2016; Miroshnikova *et al.,* 2017). During the irreversible transition state between reprogramming and differentiation, particularly under compression, the nuclei of embryonic stem cells were auxetic and stiffer, leading to nuclear deformation and more decondensed chromatin (Pagliara *et al.,* 2014). On the contrary, some studies argued that applying forces to mesenchymal stem cells resulted in an increase of cell contractility and chromatin condensation. Moreover, rate and degree of condensation were correlated with frequency and duration time of applying mechanical forces (Heo *et al.,* 2016). In mouse epidermal stem cells, mechanical strain also promoted chromatin compaction via alterations between H3K9me2/3 and H3K27me3 occupancy, leading to transcription repression (Le *et al.,*2016). Independent of experimental procedures, chromatin compaction seemed to react to mechanical forces in different directions.
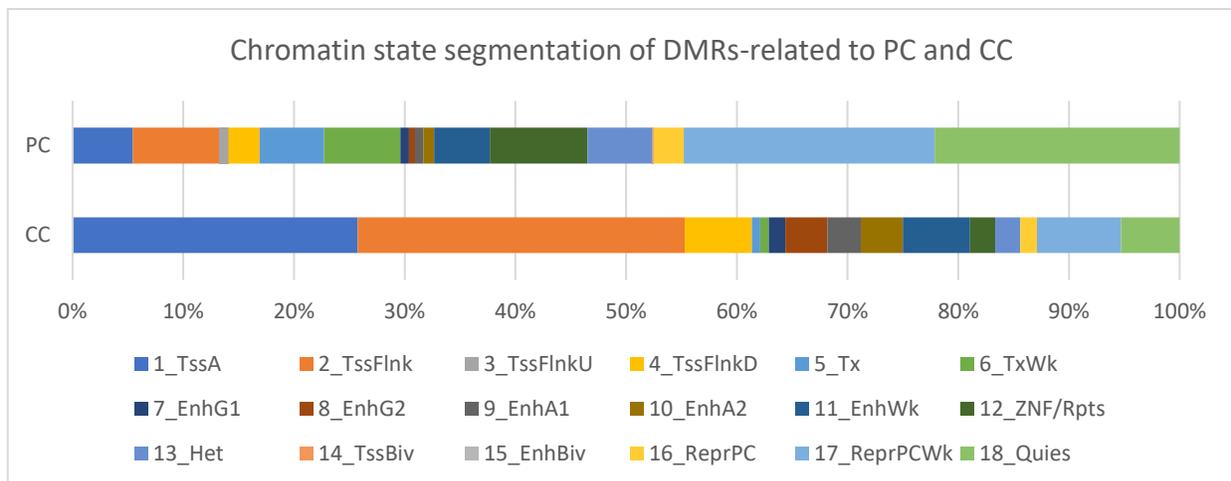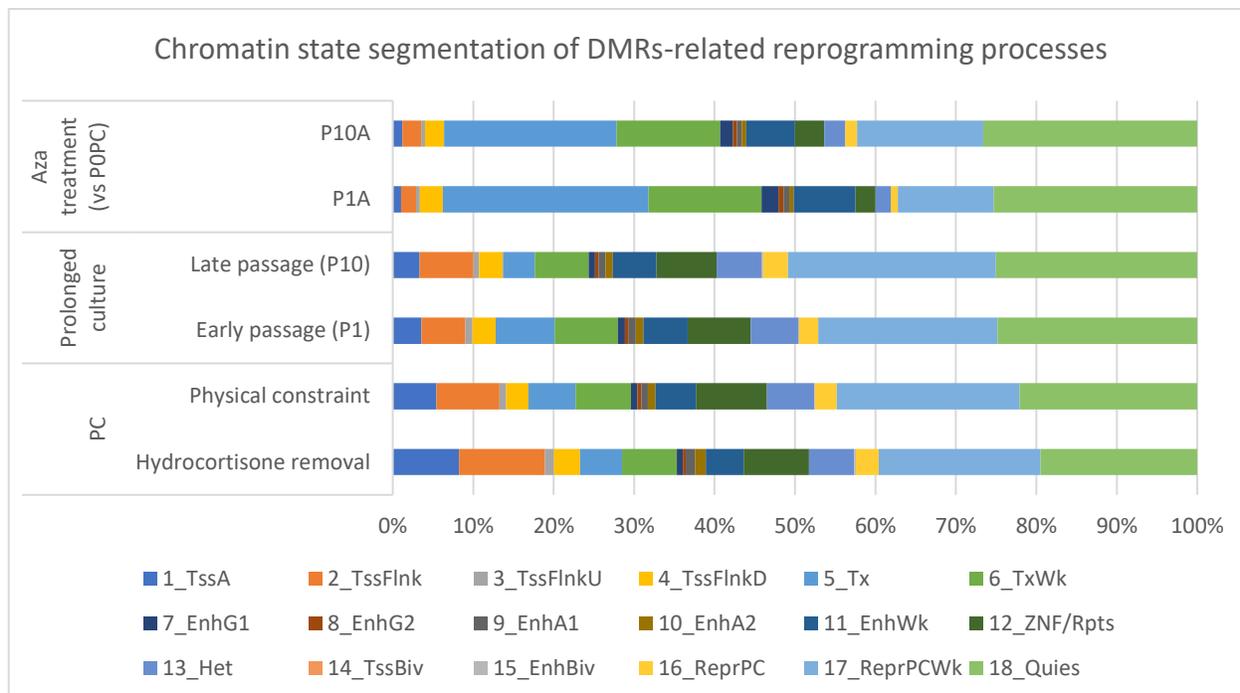

Figure 4.1 Chromatin state segmentation of DMRs-related to PC and CC. Stack bar graph shows percentages of DMRs overlapping to 18 states of segmentation (ChromHMM). PC and CC conditions are shown on the left side of each bar graph. Each state of chromatin segmentation is more provided in supplementary table 6.4.

Although the experiments were performed under controllable manners, several factors seemed undisciplined e.g. direction of forces applied to cells, and nutrients and oxygen level flown to the cells during experiment. Those factors probably influenced transcriptome profiles of PC and CC, rather than observed in methylome data. Overall mRNA profiles of PC and CC provided high

similarities. The number of DEGs observed in PC and CC showed slight differences, but similar in the tendency, as most DEGs were found to be upregulated, rather than downregulated (see Figure 3.46 i).

Generally, we observed that genes related to the regulation of response to stimuli and any immunomodulation pathways were upregulated commonly in PC and CC, but the regulation processes seemed to contribute to different pathways. Microenvironmental changes were reported for immune balance and inflammation of MSCs, switching between pro-inflammation and anti-inflammation (La Berre *et al.,* 2012; Jiang and Xu, 2020). Here, in PC, we still observed the activation of genes related to regulation of immune system process, particularly in complement and coagulation cascades. Moreover, some parts of the Hadden *et al* study revealed that cell proliferation were suppressed due to stiffness causing by differences in cell density (Hadden *et al.,* 2017). Cell proliferation also have a close relation to the regulation of the immune system in mesenchymal stem cells (MSCs), especially in a suppressive fashion. Altogether with activation of immune system process, we found that cell proliferation of passage 1 decreased (from 2×10^6 to 1.5×10^6) after PC. Therefore, effect of PC seemed to fit to those models, as the expression changes related to immunomodulation during the induction of multipotency.

Our CC experiment was performed in line with the works of Le Berre *et al.*, which the CC experiment was performed at 3.5 µm and they analyzed genome-wide expression changes, however, they used Hela cells, a cervical cancer cell line (Le Berre *et al.,* 2012). The difference in performing the experiment is expressed in the difference of the numbers of DEGs. Gene expression analysis in the works of Le Berre *et al.* was performed using Affymetrix genome wide exon array (U133 plus 2.0), while our study was performed using mRNA-sequencing, providing more accurate data. Still, upregulated DEGs of CC were found to be predominantly related to inflammation and stress responses, which was also observed in PC experiment and previous works (Le Berre *et al.,* 2012). We found that IL-17 signaling was the most significantly upregulated

pathway. Interestingly, IL-17 signaling is not only relevant in autoimmune diseases, but also plays crucial roles in cell differentiation (Mojsilović *et al,* 2015). Furthermore, NF-kappa B signaling pathways, a central inflammatory regulator and downstream target of IL-17 signaling, was upregulated as well. Genes on the upper right of the cascade were strongly expressed, which contributed to the canonical pathway (see Supplementary Figure 7.17) of NF-kappa B which was reported to be activated as response to various stimuli as well as various stresses corresponding to DNA damage and inflammation (Mojsilović *et al,* 2015; Lui *et al.,* 2017).

Remarkably, in contrast to the PC experiment, several pathways related to differentiation of mesenchymal stem cells (MSCs) were upregulated in CC experiment e.g. osteoclast differentiation and hematopoietic cell lineage. These differentiations observed by our transcriptomic data were frequently reported under the promotion of IL-17 signaling (Krstic *et al.,* 2012; Lee 2013). Although some studies provided inhibitory effects of IL-17 against neuronal and myogenic differentiation, the roles of IL-17 in cell differentiation have been suggested in relation to tissue-determination of MSCs, and microenvironmental factors (Kocić *et al.,* 2012; Li *et al.,* 2013; Mojsilović *et al,* 2015).

Our transcriptome data potentially pointed to a role of PC-induced reprogramming in the initiation of immunomodulation responding to microenvironmental changes. However, instead of immunomodulation-based adaption to microenvironmental changes, CC preferentially promoted some multipotency features, via cytokine signaling pathways, leading to cell differentiation.

Based on our results, we suggested that on epigenetic level, PC and CC induced a response in immunomodulation, and were more general than observed for mRNA levels. Namely, DNA hypomethylation and hypermethylation at specific loci were observed (see Figure 4.2, left panel). Notably, we found that those signals coincided with genes related to histone methylation/demethylation processes (see Figure 3.50, right panel), particularly with the up regulation of *KDM4B*, encoding a H3K9 demethylase (JMJD2B). This was in line with other studies describing that mechanical stretch was found to increase nuclear deformation, leading to alterations in H3K9me3 and H3K27me3 modifications. Moreover, transcriptional changes in stretched cells were reported, e.g. genes were found up-regulated that locate to facultative heterochromatin marks (Chalut *et al.,* 2012; Le *et al.,* 2016; Nava *et al.,* 2020).
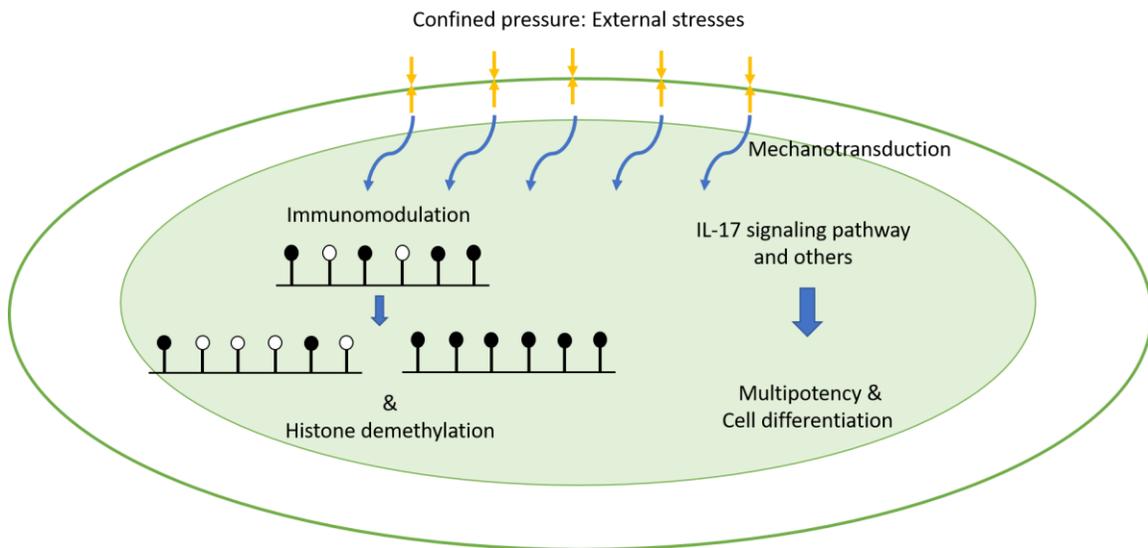
Figure 4.2 The Alteration in epigenetic and transcriptional levels of HepaRG under CC condition. The diagram illustrates that when cells (green oval) are under confined pressure, external stress signals are generated and transduced to the nucleus (green filled oval) via a process known as mechanotransduction. The effect on the left panel, particularly in epigenetic level, is generally found as immunomodulation, whereas the effect on the right panel is more specific and found in transcriptional level. Mechanotransduction induced by confined pressure activates IL-17 signaling pathways and other related pathways, leading to initiation of multipotency and differentiation.

Furthermore, responded to the confined pressure with the up-regulation of IL-17 signaling, which plays a role in both immunomodulation responses and induction of multipotency, the core pluripotency-linked networks were not activated under constant pressure condition, but some genes related to ESC signaling were activated (see Supplementary Figure 7.18, right panel).

Therefore, in this study, we suggested that mechanical forces play a role in the promotion of multipotency and facilitate differentiation observed by the activation of genes related to IL-17 signaling, osteoblast, and hematopoietic differentiation, respectively (see Figure 4.2, right panel).

4.3. Other effects observed during reprogramming and PCi-differentiation

4.3.1. Hydrocortisone-specific effects

Alteration of molecular signatures caused by hydrocortisone removal had been observed in PC state of reprogramming. In this experiment, hydrocortisone was removed after precursor cells reached confluence. Literally, the effect of hydrocortisone influencing on DNA methylation changes was barely reported, as hydrocortisone directly affects rather histone acetylation, and not DNA methylation and related cofactors (Russell *et al.,* 2010). Therefore, solely minor changes of DNA methylation, primarily focal methylation changes, in both directions - hypo- and hypermethylation - were observed. Although direct significant effects of hydrocortisone on DNA methylation were not observed , in stress-response studies cortisones derivates were frequently reported to be associated with gene-specific DNA methylation changes e.g.for *KITLG*, a gene encoding a ligand of the tyrosine-kinase receptor (Houtepen *et al.,* 2016; Wrigglesworth *et al.,* 2019), and for *CYP11B1*, a gene that encodes the 11β-hydroxylase enzyme which catalyzes cortisol biosynthesis. Particularly, significant hypomethylation of *CYP11B1* promoter was found in hypercortisolemia in cortisol-producing adenoma (Kometani *et al.,* 2017). Our study seemed to be not comparable to the previous studies. However, we observed that, after hydrocortisone was removed, DMRs in gene body of *CYP11B1* became slightly hypermethylated, oppositely to the finding in hypercortisolemia (see Supplementary Figure 7.19).

Additionally, the alteration of mRNA levels was minor, the number of DEGs was low. The strongest observation in this condition related to down-regulation of genes involved in metabolic processes. Hepatocytes indeed play important roles in metabolic pathways, particularly in glucose and lipid metabolism in liver (Bechmann *et al.,* 2012; Rui, 2014). As hydrocortisone, a growth

factor specific to hepatocytes, was withdrawn from the culture, downregulation of genes related to glucose and lipid metabolisms, such as *G6PC* – catalytic subunit of G6Pase, *ACOX2* – Acetyl-Co A oxidase 2, and *ETNK2* – Ethanolamine Kinase 2 etc., was commonly found. The alteration of DNA methylation and gene expression observed in this condition leads to the suggestion that adaptation to short-term stress induction and suppression of hepatocyte specificity were promoted when hydrocortisone was removed from the cultures.

Moving forward to the maintenance state, two well-known factors influencing DNA methylation patterns needed to be concerned – i) prolonged culture, and ii) Aza treatment effect, respectively.

4.3.2. Long-term cultivation effect

Focusing on prolonged or long-term cultivation, global hypomethylation and focal hypermethylation have been proposed as signatures of long-term *in vitro* cultivation found in several cell types (Bork *et al.*, 2010; Cruickshanks *et al.*, 2013 and Dmitrijeva *et al.*, 2018). In contrast, our observation was that prolonged cultivation of HepaRG to passage 10, either after applying PC or without PC, resulted in genome wide hypermethylation, also detected in methylation levels of LINE1 and HERVK elements (see Supplementary Figure 7.20). Consistent with the global level, hypermethylation at specific loci were preferentially observed, rather than hypomethylation (see Figure 3.7 and Figure 3.10 ii). Global hypomethylation in prolonged culture was reported in association with a proliferation arrest and diminished DNMT activity (Cruickshanks *et al.*, 2013). Since our RP samples could be cultivated for up to 20 passages (unpublished data, Biopredic) and *DNMT1* and *DNMT3A* were still strongly expressed at P10 (see Supplementary Figure 7.21), global hypomethylation could not be expected in this experiment. Additionally, HepaRG cells were obtained from hepatocellular carcinoma, some properties of cancerous cells were still maintained (Gripon *et al.,* 2002), such as DNMTs activities, which were reported previously in hepatocellular carcinoma and many cancers (Oh *et al.,* 2007; Subramaniam *et al.,* 2014).

Differences in gene expression between long-term cultivation of PC and nonPC cells could also be observed. Long-term cultivation of nonPC led to a down-regulation of genes related to the regulation of cell adhesion and extracellular matrix organization, whereas up-regulated genes were related to morphogenesis and vasculature development. Those findings indicate that DEGs of long-term culture of nonPC possibly promote inflammation response as well as angiogenesis and vascular development, as generally found in tumorigenesis (Ziyad and Iruela-Arispe, 2011; Jiang and Xu, 2020). In long-term cultivated PC cells, DEGs included down-regulated genes related to some metabolic pathways while up-regulated genes promoted protein tyrosine kinase activity (see Figure 3.19 iv). Consequently, overall results of long-term cultivation of PC revealed that HepaRG did not only exhibit reduced tissue specificity, but also partially maintain stress responses caused by microenvironmental changes via immunomodulation processes.

4.3.3. Azacytidine -specific effects

5-Azacytidine (Aza) was used to stabilize multipotent cell-like cellular properties after PC. We therefore proposed that Aza-treated samples biologically – on the molecular level - get closer to multipotent cells. As expected, an extremely high number of hypomethylated DMRs was discovered.

For cellular reprogramming, several transcription factors and methods were used to induce pluripotent properties. As shown in previous studies, hypermethylation was observed at pluripotent-related genes in partially reprogrammed cells. After Aza treatment reduction of methylation levels at those genes was found to provide rapid and full reprogramming with high efficiency, as pluripotency-related genes were further reactivated (Mikkelsen *et.al.,* 2008). Notably, the transcripts of pluripotency-related genes, such as *KLF4* and *DUSP9,* were barely observed in the Aza-treated samples, although the methylation of those genes was gradually decreased, from passage 0 of PC sample to passage 10 of PC+A (see Supplementary Figure 7.22). Along with this, the results from Aza treatment suggest that the epigenetic state of the cells

particularly seemed to be stabilized at lower methylation levels, but the effect was not sufficient for the activation of the pluripotency state.
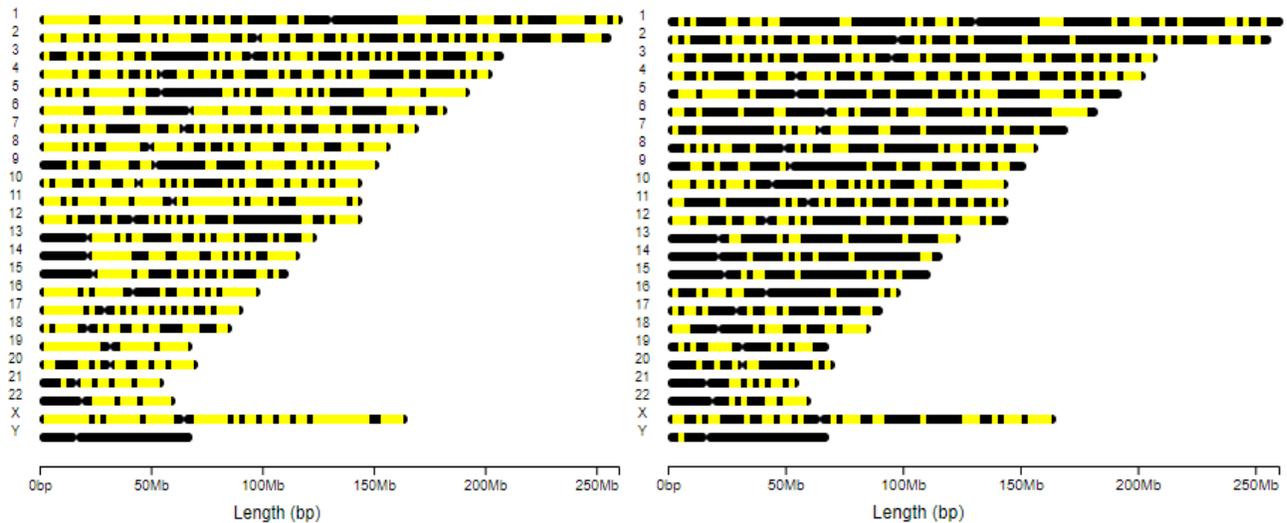
Figure 4.3 Hypermethylated DMRs found in Aza treated samples. The figures provide ideograms of each chromosome showing hypermethylated DMRs in yellow, whereas the rest of the genome is in blue. On the left side is the figure of Aza treated cells in passage 1, while the figure on the right displays Aza treated cells in passage 10.

Although Aza-treatment was highly effective in promoting genome-wide hypomethylation in both, passages 1 and 10, hypermethylated DMRs could still be observed (less than 1%, diff meth = 10-60%). Figure 4.3 shows that hypermethylated DMRs were dispersed throughout the genome, not clustering particularly at persistent heterochromatic regions such as centromeric and telomeric regions.

Several studies on iPSCs suggest that hypermethylation occurring during reprogramming probably underlines an epigenetic aberration of iPSCs as a consequence of epigenetic memory of iPSCs (Lister *et al.,* 2011; Ohi *et al.,* 2011; Nishizawa *et al.,* 2016). This aberrant hypermethylation was reported to have functional roles in reprogramming and differentiation. For instance, in blood-iPSCs, aberrant hypermethylation was associated to high maturation capacity of iPSCs (Nishizawa *et al.,* 2016). Unfortunately, our study was performed without biological replicates, differential methylation analyses comparing the cells with precursors or their passage counterparts could not be included. In addition, functional annotations of DMRs associated with

somatic cell origin influencing DNA methylation seemed to be barely observed in this study, but hypermethylated DMRs were found related to nervous system development (34 genes, FDR =0.0038). This effect was frequently found in the establishment of hepatocyte-like cells derived from hepatocellular carcinoma (Tao *et al.,* 2011).

During Aza treatment, global hypomethylation was a major effect since Azacytidine is a cytidine analog, which competes with cytidine during DNA replication, leading to global hypomethylation of the genome and transcriptional activation (Christman, 2002; Sigalotti *et al.,* 2007). Although Aza was reported to influence RNA transcription processes (Schaefer *et al.,* 2009; Qiu *et al.,* 2010), the effect of Aza treatment on RNA transcription was barely observed in this study. Our results from mRNA-Seq revealed that the number of up-regulated DEGs was dramatically increased in the Aza treated samples (see Figure 3.18 ii, comparison D and E). Up-regulated DEGs were mainly found to be associated with regulation of cell adhesion and cell cycle. Conversely, tissue specifically expressed genes, such as genes related to complement and coagulation cascades, were continuously downregulated (see Figure 3.20, lower part).

Notably, during reprogramming, GO terms and KEGG pathways obtained from DEGs of PC or maintenance state revealed some similarities, but the terms were allocated to different regulation directions. For instance, genes associated with hsa04610 complement and coagulation cascades were strongly expressed in PC. The expression of those genes became lower in the maintenance state, particularly in Aza-treated samples (see Supplementary Figure 7.23). Therefore, Aza treatment not only provided an epigenetic resetting of DNA methylation, but the treatment also facilitated multipotency by decreasing the specificity of the cells and suppressing genes related to immune responses during reprogramming.

4.3.4. <u>PC-induced differentiation to hepatocyte-liked cells</u>

On the other hand, during HepaRG differentiation, work by Oriana Genolet reported a minor change in DNA methylation during HepaRG differentiation, the methylation profiles of samples collected during the differentiation process did not show distinctive patterns. Methylation changes in the samples which were supplemented with DMSO from day 15 onwards were also not prominent, meaning that the methylation of HepaRG undergoing differentiation was not influenced by hydrocortisone and DMSO similar to the findings after reprogramming (Genolet, Master thesis 2012). However, in the mentioned Master thesis, the 450K BeadChip array (Illumina) was used which is not comparable for all CpGs detected by RRBS, so any conclusion from comparisons have to be taken with care.

In contrast to previous differentiation studies, we used hydrocortisone for re-establishing differentiated cells and DMSO to restore hepatic functions after PC. This also showed mildly different results. Methylation profiles of RD samples still clearly showed a global hypomethylation with focal hypermethylation (see Figure 3.7 – green bar), assuming that Aza treatment effects in the starter (RD_P0nonPC+A) were probably retained. However, a tendency towards increasing methylation levels during PCi-differentiation appeared throughout the process.

Concerning PCi-differentiation from multipotent cells to hepatocyte like cells, we pinpointed the steps in which PCi-differentiation was induced in the cells by hydrocortisone and DMSO treatment. Interestingly, PCi-differentiation with a combination of hydrocortisone and DMSO treatment of the culture showed promising results. As expected, several genes related to cytochrome P450 family became upregulated during PCi-differentiation, suggesting that hepatocyte-specific properties were partially rebooted (see Supplementary Figure 7.24).

4.4. Chromatin state segmentation in correlation to DNA methylation and gene expression

Previous analyses showed that the distribution of DMRs according to genome annotation provided highly similar patterns among each state. The DMRs were frequently located in introns and intergenic regions, leading to the question if any regulatory regions were influenced by the alteration of DNA methylation. Since histone modification enrichment and DNA accessibility profiles of HepaRG were available, segmentation of chromatin states segmentation (ChromHMM) of HepaRG was done and linked to functional sequence units in the genome (Ernst and Kellis, 2012; Salhab *et.al,* 2018). This study took advantage of ChromHMM to deliver predictive chromatin states of DMRs. DMRs that overlapped with chromatin states with at least 10 bp were defined to be DMRs located within those chromatin states. However, the analyses should be handled carefully, because DMRs potentially can overlap with more than one chromatin segment (active and inactive). Additionally, predictive chromatin states were also obtained from differentiated HepaRG.

Overall, DMRs related to reprogramming processes were found to be more prominent in inactive states (No.12-18) of chromatin, particularly during PC, as well as during prolonged cultivation. Dramatic changes were observed when Aza was applied to the culture, e.g. number of active states (No.1-11) increased significantly (Figure 4.4).

Figure 4.4 Chromatin state segmentation of DMR-related reprogramming processes. Stack bar graph shows percentages of DMRs overlapping to 18 states of segmentation (ChromHMM). Sample state and factors influencing methylation changes are shown on the left side of each bar graph. Detailed information about each state of chromatin segmentation is provided in supplementary table 6.4

Besides, as shown in   Figure 4.4, ChromHMM seemed to reveal a specific signature of methylation during reprogramming since DMRs related to each state of the process were predominantly altered within active chromatin states. Apparently, active TSS (1_TssA; 8.25%) and flanking TSS (2_TssFlnk; 10.66%) sites, prominently found in PC and prolonged cultivation, dramatically decreased in Aza-treated samples. Instead of active TSS and flanking TSS regions, regions indicating strong transcription (5_Tx) and weak transcription (6_TxWk) significantly increased either in passage 1 (39.61%) or in passage 10 (34.32%), suggesting that Aza treatment potentially influenced chromatin state to be more open. Another significant effect was observed in more closed chromatin, particularly in DMRs related to weakly repressed polycomb (17_ReprPCwk) in passage 10 (15.68%). Regions indicating weakly repressed polycomb state

probably seem to be sensitive to methylation changes, the same effect was found in long term Aza-treated samples.
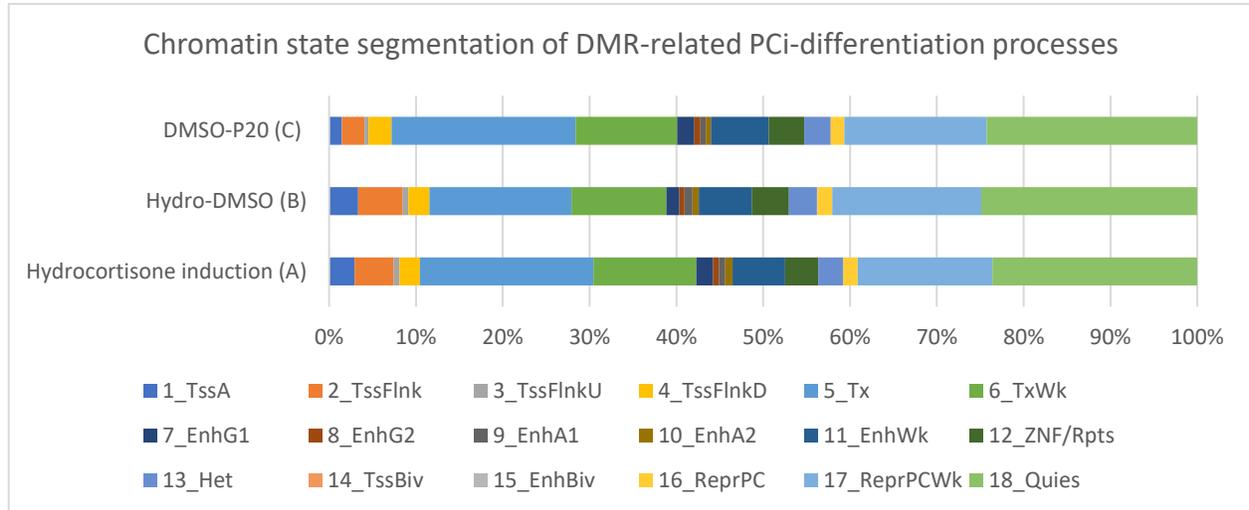
Figure 4.5 Chromatin state segmentation of DMR-related PCi-differentiation processes. Stack bar graphs show percentages of DMRs overlapping with 18 states of segmentation (ChromHMM). Sample states and factors influencing methylation changes are shown on the left side of each bar and correspond to figure 3.13-i. Detailed information of each state of chromatin segmentation is provided in supplementary table 6.4

During PCi-differentiation, the patterns of ChromHMM were comparable to those in the reprogramming process, particularly in the maintenance state, as the effect of Aza treatment remained in the cells (Figure 4.4 and Figure 4.5). High proportions of DMRs located in segments associated with strong transcription (5_Tx), weak transcription (6_TxWk), weak repressed polycomb (17_ReprPCwk) and quiescent regions (18_Quies) were found at similar levels. In this state, an impact of hydrocortisone on DNA methylation was also noticed. Figure 4.5 shows that, after hydrocortisone induction, the number of DMRs associated with active TSS (1_TssA) and TSS flanking regions (2_TSSFlnk) was increased, while the number of DMRs associated with strong transcription (5_Tx) and weak transcription (6_TxWk) segments declined. These findings may strongly support the role of hydrocortisone in influencing DNA methylation changes, particularly at the TSS and TSS flanking regions, although they were barely reported in previous studies. Therefore, the combination of methylation data and chromatin state segmentation confirmed that PCi-differentiation of HepaRG was partially induced after hydrocortisone and

DMSO administration. Besides, prolonged cultivation of PCi-differentiated samples revealed similar effects as detected for the Aza-treated samples.

In summary, according to ChromHMM, DMRs are predominantly located in closed chromatin rather than in more open segments and, furthermore, can be steadily found in the same proportions, approx. 60% of closed and 40% of open states, either in reprogramming or PCi-differentiation processes. However, during PC and maintenance, alteration of chromatin states was preferentially found in DMRs located in open segments, while alterations in PCi-differentiation processes were found in similar patterns compared to DMRs after Aza treatment.

Finally, exploring the correlation of DNA methylation and gene expression revealed a frequent association between DNA methylation and mRNA levels in promoter regions. As reported in many disease-related studies, promoter hypomethylation promoted activation of genes, whereas promoter hypermethylation promoted suppression of genes related to these promoters (Fan *et al.,* 2006; Jones 2012; Deng *et al.,* 2019). Overall, the results of this study were consistent with these previous reports, but the correlation of DMRs and DEGs could not be applied thoroughly with those models. For example, after prolonged cultivation, although a negative correlation had mostly been found, a positive correlation between DMRs and DEGs could still be observed (see Figure 3.29 – regulation bar). Notably, negative correlation, as well as positive correlation, were not detected solely in promoter regions, but also in other regulatory regions (see Figure 3.30). Gene body methylation correlated with mRNA levels, particularly when found in the first intron, where predicted enhancers and transcription factor binding motifs were shown to be frequently located (Yang *et al.,* 2014; Anastasiadi *et al.,* 2018). Moreover, a progressive loss of methylation in gene bodies also showed a positive correlation with H3K36me3, a histone modification often enriched in heterochromatic regions. Focusing on partially methylated domains (PMDs) of HepaRG, PMDs associated with high transcription were highly methylated and lowly enriched for heterochromatic markers, compared to the methylation average across ChromHMM of HepaRG

(Chantalat *et al.,* 2011; Salhab *et al.,* 2018). Although PMD data was not obtained in this study, ChromHMM analysis revealed that DMRs were associated with open segments during reprogramming and PCi-differentiation processes and thus possibly correlate with active genes. Prominent examples were discovered in reprogramming processes when DMRs, particularly in maintenance state, were increasingly found in open segments which may result in an increase of gene expression. Along with this, in addition to DNA methylation, the results from the integrative analysis with ChromHMM convinced us that chromatin states are one of the important underlying mechanisms regulating gene expression in HepaRG during reprogramming and PCi-differentiation.

4.5. Further directions

To investigate reprogramming events further, some parts of the experimental procedures still need to be improved, e.g. using more cell material for further analysis. Based on the observation that histone demethylation coincided with CC, together with the results that DMRs correlate to chromatin state segmentation, a broad analysis of histone modifications and nucleosome occupancies might be interesting, since these chromatin signatures are known to be influenced directly from mechanical forces. Additionally, our studies provided evidence of supporting roles of 5-hydroxymethylcytosine (5-hmC) in cell cycle progression during CC. Investigation of 5-hmC in association with cell cycle states during CC, particularly at single cell level, might provide a better understanding about DNA methylation and cell cycle dynamics during organogenesis, an *in vivo* mechanical process found during development.

Chapter 5 Summary

5.1. Overall summary

In this study, we used physical constraint and confinement to induce cellular reprogramming and redirected differentiation of hepatocyte-like cells. Analysing the molecular signature of HepaRG cells, we found that cellular reprogramming probably did not entirely occur, but molecular processes tending towards PCi-differentiation were partially found. Indeed, we observed minor changes in overall DNA methylation that tended towards hypomethylation in the PC and CC experiments. Obtained data from prolonged cultivation of HepaRG pointed to overall increased methylation levels. In the process of PCi-differentiation, genome-wide hypomethylation acquired by treatment of HepaRG with Aza was still maintained and constantly remained when hepatocyte differentiation was induced. However, overall methylation levels of PCi-differentiated hepatocyte-like cells tended towards an increase upon continued cultivation.

Looking at genome-wide mRNA profiles, genes related to immunomodulation processes were predominantly activated to respond to PC-induced stress and prolonged cultivation. Hepatocyte-specific genes were found with reduced mRNA levels. During PCi-differentiation process, members of the cytochrome P450 family were restored after DMSO was applied to the culture, confirming that hepatocyte specificity was partially activated during the process.

Distribution of DMRs related to functional genome elements provided distinctive patterns, particularly when integrated to chromatin segments using ChromHMM. We found that the proportion of active and inactive chromatin states seemed to remain constantly along the reprogramming and PCi-differentiation processes. Notably, DMRs integrated to ChromHMM were found to correlate to gene expression more significantly than DMRs only. Therefore, chromatin states potentially represent another underlying mechanism related to the alteration of DNA methylation and gene expression during our study.

Chapter 6 Bibliography

1.	Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biol. 2012;13(10):R87.

2.	Almalki SG, Agrawal DK. Key transcription factors in the differentiation of mesenchymal stem cells. Differentiation. 2016;92(1-2):41-51.

3.	Anastasiadi D, Esteve-Codina A, Piferrer F. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. Epigenetics Chromatin. 2018;11(1):37.

4.	Aninat C, Piton A, Glaise D, Le Charpentier T, Langouët S, Morel F, et al. Expression of cytochromes P450, conjugating enzymes and nuclear receptors in human hepatoma HepaRG cells. Drug Metab Dispos. 2006;34(1):75-83.

5.	Anthérieu S, Chesné C, Li R, Camus S, Lahoz A, Picazo L, et al. Stable expression, activity, and inducibility of cytochromes P450 in differentiated HepaRG cells. Drug Metab Dispos. 2010;38(3):516-25.

6.	Arányi T, Váradi A, Simon I, Tusnády GE. The BiSearch web server. BMC Bioinformatics. 2006;7:431.

7.	Augoff K, McCue B, Plow EF, Sossey-Alaoui K. miR-31 and its host gene lncRNA LOC554202 are regulated by promoter hypermethylation in triple-negative breast cancer. Mol Cancer. 2012;11:5.

8.	Ballestar E, Esteller M. Epigenetic gene regulation in cancer. Adv Genet. 2008;61:247-67.

9.	Bechmann LP, Hannivoort RA, Gerken G, Hotamisligil GS, Trauner M, Canbay A. The interaction of hepatic lipid and glucose metabolism in liver diseases. J Hepatol. 2012;56(4):952-64.

10.	Bednar J, Horowitz RA, Grigoryev SA, Carruthers LM, Hansen JC, Koster AJ, et al. Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. Proc Natl Acad Sci U S A. 1998;95(24):14173-8.

11.	Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature. 2000;405(6785):482-5.

12.	Bell CC, Lauschke VM, Vorrink SU, Palmgren H, Duffin R, Andersson TB, et al. Transcriptional, Functional, and Mechanistic Comparisons of Stem Cell-Derived Hepatocytes, HepaRG Cells, and Three-Dimensional Human Hepatocyte Spheroids as Predictive In Vitro Systems for Drug-Induced Liver Injury. Drug Metab Dispos. 2017;45(4):419-29.

13.   Bock C, Beerman I, Lien WH, Smith ZD, Gu H, Boyle P, et al. DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. Mol Cell. 2012;47(4):633-47.

14.   Bock C, Tomazou EM, Brinkman AB, Müller F, Simmer F, Gu H, et al. Quantitative comparison of genome-wide DNA methylation mapping technologies. Nat Biotechnol. 2010;28(10):1106-14.

15.   Bork S, Pfister S, Witt H, Horn P, Korn B, Ho AD, et al. DNA methylation pattern changes upon long-term culture and aging of human mesenchymal stromal cells. Aging Cell. 2010;9(1):54-63.

16.   Bostick M, Kim JK, Estève PO, Clark A, Pradhan S, Jacobsen SE. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science. 2007;317(5845):1760-4.

17.   Bröske AM, Vockentanz L, Kharazi S, Huska MR, Mancini E, Scheller M, et al. DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. Nat Genet. 2009;41(11):1207-15.

18.   Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, et al. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell. 2012;150(6):1209-22.

19.   Buganim Y, Faddah DA, Jaenisch R. Mechanisms and models of somatic cell reprogramming. Nat Rev Genet. 2013;14(6):427-39.

20.   Caiazzo M, Okawa Y, Ranga A, Piersigilli A, Tabata Y, Lutolf MP. Defined three-dimensional microenvironments boost induction of pluripotency. Nat Mater. 2016;15(3):344-52.

21.   Cattin CJ, Düggelin M, Martinez-Martin D, Gerber C, Müller DJ, Stewart MP. Mechanical control of mitotic progression in single animal cells. Proc Natl Acad Sci U S A. 2015;112(36):11258-63.

22.   Cerec V, Glaise D, Garnier D, Morosan S, Turlin B, Drenou B, et al. Transdifferentiation of hepatocyte-like cells from the human hepatoma HepaRG cell line through bipotent progenitor. Hepatology. 2007;45(4):957-67.

23.   Chalut KJ, Höpfler M, Lautenschläger F, Boyde L, Chan CJ, Ekpenyong A, et al. Chromatin decondensation and nuclear softening accompany Nanog downregulation in embryonic stem cells. Biophys J. 2012;103(10):2060-70.

24.   Chantalat S, Depaux A, Héry P, Barral S, Thuret JY, Dimitrov S, et al. Histone H3 trimethylation at lysine 36 is associated with constitutive and facultative heterochromatin. Genome Res. 2011;21(9):1426-37.

25.   Chen F, Zhang G, Yu L, Feng Y, Li X, Zhang Z, et al. High-efficiency generation of induced pluripotent mesenchymal stem cells from human dermal fibroblasts using recombinant proteins. Stem Cell Res Ther. 2016;7(1):99.

26.   Chen J, Liu H, Liu J, Qi J, Wei B, Yang J, et al. H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. Nat Genet. 2013;45(1):34-42.

27.   Chen T, Ueda Y, Dodge JE, Wang Z, Li E. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Mol Cell Biol. 2003;23(16):5594-605.

28.   Chowdhury F, Na S, Li D, Poh YC, Tanaka TS, Wang F, et al. Material properties of the cell dictate stress-induced spreading and differentiation in embryonic stem cells. Nat Mater. 2010;9(1):82-8.

29.   Christman JK. 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. Oncogene. 2002;21(35):5483-95.

30.   Cortázar D, Kunz C, Selfridge J, Lettieri T, Saito Y, MacDougall E, et al. Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. Nature. 2011;470(7334):419-23.

31.   Cruickshanks HA, McBryan T, Nelson DM, Vanderkraats ND, Shah PP, van Tuyn J, et al. Senescent cells harbour features of the cancer epigenome. Nat Cell Biol. 2013;15(12):1495-506.

32.   Dan YY, Riehle KJ, Lazaro C, Teoh N, Haque J, Campbell JS, et al. Isolation of multipotent progenitor cells from human fetal liver capable of differentiating into liver and mesenchymal lineages. Proc Natl Acad Sci U S A. 2006;103(26):9912-7.

33.   Deng GX, Xu N, Huang Q, Tan JY, Zhang Z, Li XF, et al. Association between promoter DNA methylation and gene expression in the pathogenesis of ischemic stroke. Aging (Albany NY). 2019;11(18):7663-77.

34.   Deng J, Shoemaker R, Xie B, Gore A, LeProust EM, Antosiewicz-Bourget J, et al. Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. Nat Biotechnol. 2009;27(4):353-60.

35.   Desmaison A, Frongia C, Grenier K, Ducommun B, Lobjois V. Mechanical stress impairs mitosis progression in multi-cellular tumor spheroids. PLoS One. 2013;8(12):e80447.

36.   Desmaison A, Guillaume L, Triclin S, Weiss P, Ducommun B, Lobjois V. Impact of physical confinement on nuclei geometry and cell division dynamics in 3D spheroids. Sci Rep. 2018;8(1):8785.

37.   Dianat N, Dubois-Pot-Schneider H, Steichen C, Desterke C, Leclerc P, Raveux A, et al. Generation of functional cholangiocyte-like cells from human pluripotent stem cells and HepaRG cells. Hepatology. 2014;60(2):700-14.

38.   Dmitrijeva M, Ossowski S, Serrano L, Schaefer MH. Tissue-specific DNA methylation loss during ageing and carcinogenesis is linked to chromosome structure, replication timing and cell division rates. Nucleic Acids Res. 2018;46(14):7022-39.

39.   Doege CA, Inoue K, Yamashita T, Rhee DB, Travis S, Fujita R, et al. Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. Nature. 2012;488(7413):652-5.

40.   Downing TL, Soto J, Morez C, Houssin T, Fritz A, Yuan F, et al. Biophysical regulation of epigenetic state and cell reprogramming. Nat Mater. 2013;12(12):1154-62.

41.   Dyachenko OV, Schevchuk TV, Kretzner L, Buryanov YI, Smith SS. Human non-CG methylation: are human stem cells plant-like? Epigenetics. 2010;5(7):569-72.

42.   Engler AJ, Sen S, Sweeney HL, Discher DE. Matrix elasticity directs stem cell lineage specification. Cell. 2006;126(4):677-89.

43.   Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nat Methods. 2012;9(3):215-6.

44.   Fan M, Yan PS, Hartman-Frey C, Chen L, Paik H, Oyer SL, et al. Diverse gene expression and DNA methylation profiles correlate with differential adaptation of breast cancer cells to the antiestrogens tamoxifen and fulvestrant. Cancer Res. 2006;66(24):11954-66.

45.   Fatemi M, Hermann A, Gowher H, Jeltsch A. Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. Eur J Biochem. 2002;269(20):4981-4.

46.   Fischle W, Wang Y, Jacobs SA, Kim Y, Allis CD, Khorasanizadeh S. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev. 2003;17(15):1870-81.

47.   Flynn RA, Chang HY. Long noncoding RNAs in cell-fate programming and reprogramming. Cell Stem Cell. 2014;14(6):752-61.

48.   Folmes CD, Nelson TJ, Martinez-Fernandez A, Arrell DK, Lindor JZ, Dzeja PP, et al. Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. Cell Metab. 2011;14(2):264-71.

49.   Fulka H, Mrazek M, Tepla O, Fulka J. DNA methylation pattern in human zygotes and developing embryos. Reproduction. 2004;128(6):703-8.

50.   Gao Y, Chen J, Li K, Wu T, Huang B, Liu W, et al. Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. Cell Stem Cell. 2013;12(4):453-69.

51.   Garcia-Perez JL, Widmann TJ, Adams IR. The impact of transposable elements on mammalian development. Development. 2016;143(22):4101-14.

52.   Genolet O. Analysis of epigenetic signatures involved in the differentiation process of the HepaRG cell line. Saarbruecken, Germany: Universitaet des Saarlandes; 2012.

53.   Gilbert N, Boyle S, Fiegler H, Woodfine K, Carter NP, Bickmore WA. Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. Cell. 2004;118(5):555-66.

54.   Gries J, Schumacher D, Arand J, Lutsik P, Markelova MR, Fichtner I, et al. Bi-PROF: bisulfite profiling of target regions using 454 GS FLX Titanium technology. Epigenetics. 2013;8(7):765-71.

55.   Gripon P, Rumin S, Urban S, Le Seyec J, Glaise D, Cannie I, et al. Infection of a human hepatoma cell line by hepatitis B virus. Proc Natl Acad Sci U S A. 2002;99(24):15655-60.

56.   Grossniklaus U, Paro R. Transcriptional silencing by polycomb-group proteins. Cold Spring Harb Perspect Biol. 2014;6(11):a019331.

57.   Gu H, Smith ZD, Bock C, Boyle P, Gnirke A, Meissner A. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nat Protoc. 2011;6(4):468-81.

58.   Guguen-Guillouzo C, Corlu A, Guillouzo A. Stem cell-derived hepatocytes and their use in toxicology. Toxicology. 2010;270(1):3-9.

59.   Guillouzo A, Corlu A, Aninat C, Glaise D, Morel F, Guguen-Guillouzo C. The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics. Chem Biol Interact. 2007;168(1):66-73.

60.   Gulyaeva LF, Kushlinskiy NE. Regulatory mechanisms of microRNA expression. J Transl Med. 2016;14(1):143.

61.   Guo F, Li X, Liang D, Li T, Zhu P, Guo H, et al. Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. Cell Stem Cell. 2014;15(4):447-59.

62.   Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, et al. The DNA methylation landscape of human early embryos. Nature. 2014;511(7511):606-10.

63.   Hadden WJ, Young JL, Holle AW, McFetridge ML, Kim DY, Wijesinghe P, et al. Stem cell migration and mechanotransduction on linear stiffness gradient hydrogels. Proc Natl Acad Sci U S A. 2017;114(22):5647-52.

64.   Hanna J, Saha K, Pando B, van Zon J, Lengner CJ, Creyghton MP, et al. Direct cell reprogramming is a stochastic process amenable to acceleration. Nature. 2009;462(7273):595-601.

65.   Hansson J, Rafiee MR, Reiland S, Polo JM, Gehring J, Okawa S, et al. Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. Cell Rep. 2012;2(6):1579-92.

66.   Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, et al. Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. Nat Biotechnol. 2010;28(10):1097-105.

67.   Hart SN, Li Y, Nakamoto K, Subileau EA, Steen D, Zhong XB. A comparison of whole genome gene expression profiles of HepaRG cells and HepG2 cells to primary human hepatocytes and human liver tissues. Drug Metab Dispos. 2010;38(6):988-94.

68.   Hellman A, Chess A. Gene body-specific methylation on the active X chromosome. Science. 2007;315(5815):1141-3.

69.   Heo SJ, Han WM, Szczesny SE, Cosgrove BD, Elliott DM, Lee DA, et al. Mechanically Induced Chromatin Condensation Requires Cellular Contractility in Mesenchymal Stem Cells. Biophys J. 2016;111(4):864-74.

70.   Hermann A, Goyal R, Jeltsch A. The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. J Biol Chem. 2004;279(46):48350-9.

71.   Herranz M, Esteller M. DNA methylation and histone modifications in patients with cancer: potential prognostic and therapeutic targets. Methods Mol Biol. 2007;361:25-62.

72.   Heuzé ML, Vargas P, Chabaud M, Le Berre M, Liu YJ, Collin O, et al. Migration of dendritic cells: physical principles, molecular mechanisms, and functional implications. Immunol Rev. 2013;256(1):240-54.

73.   Higuchi Y, Kawai K, Kanaki T, Yamazaki H, Chesné C, Guguen-Guillouzo C, et al. Functional polymer-dependent 3D culture accelerates the differentiation of HepaRG cells into mature hepatocytes. Hepatol Res. 2016;46(10):1045-57.

74.   Holle AW, Engler AJ. Cell rheology: Stressed-out stem cells. Nat Mater. 2010;9(1):4-6.

75.   Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. Science. 1975;187(4173):226-32.

76.   Houtepen LC, Vinkers CH, Carrillo-Roa T, Hiemstra M, van Lier PA, Meeus W, et al. Genome-wide DNA methylation levels and altered cortisol stress reactivity following childhood trauma in humans. Nat Commun. 2016;7:10967.

77.   Hu X, Zhang L, Mao SQ, Li Z, Chen J, Zhang RR, et al. Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. Cell Stem Cell. 2014;14(4):512-22.

78.   Huang K, Zhang X, Shi J, Yao M, Lin J, Li J, et al. Dynamically reorganized chromatin is the key for the reprogramming of somatic cells to pluripotent cells. Sci Rep. 2015;5:17691.

79.   Huangfu D, Osafune K, Maehr R, Guo W, Eijkelenboom A, Chen S, et al. Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol. 2008;26(11):1269-75.

80.   Humphrey JD, Dufresne ER, Schwartz MA. Mechanotransduction and extracellular matrix homeostasis. Nat Rev Mol Cell Biol. 2014;15(12):802-12.

81.   Iizaka M, Han HJ, Akashi H, Furukawa Y, Nakajima Y, Sugano S, et al. Isolation and chromosomal assignment of a novel human gene, CORO1C, homologous to coronin-like actin-binding proteins. Cytogenet Cell Genet. 2000;88(3-4):221-4.

82.   Iqbal K, Jin SG, Pfeifer GP, Szabó PE. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. Proc Natl Acad Sci U S A. 2011;108(9):3642-7.

83.   Jang HS, Shin WJ, Lee JE, Do JT. CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function. Genes (Basel). 2017;8(6).

84.   Jiang W, Xu J. Immune modulation by mesenchymal stem cells. Cell Prolif. 2020;53(1):e12712.

85.   Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13(7):484-92.

86.   Kagiwada S, Kurimoto K, Hirota T, Yamaji M, Saitou M. Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. EMBO J. 2013;32(3):340-53.

87.   Kanninen LK, Porola P, Niklander J, Malinen MM, Corlu A, Guguen-Guillouzo C, et al. Hepatic differentiation of human pluripotent stem cells on human liver progenitor HepaRG-derived acellular matrix. Exp Cell Res. 2016;341(2):207-17.

88.   Kar G, Kim JK, Kolodziejczyk AA, Natarajan KN, Torlai Triglia E, Mifsud B, et al. Flipping between Polycomb repressed and active transcriptional states introduces noise in gene expression. Nat Commun. 2017;8(1):36.

89.   Kim K, Doi A, Wen B, Ng K, Zhao R, Cahan P, et al. Epigenetic memory in induced pluripotent stem cells. Nature. 2010;467(7313):285-90.

90.   Kocić J, Santibañez JF, Krstić A, Mojsilović S, Dorđević IO, Trivanović D, et al. Interleukin 17 inhibits myogenic and promotes osteogenic differentiation of C2C12 myoblasts by activating ERK1,2. Biochim Biophys Acta. 2012;1823(4):838-49.

91.   Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. Nature. 2013;502(7472):472-9.

92.   Kometani M, Yoneda T, Demura M, Koide H, Nishimoto K, Mukai K, et al. Cortisol overproduction results from DNA methylation of CYP11B1 in hypercortisolemia. Sci Rep. 2017;7(1):11205.

93.   Kouzarides T. Chromatin modifications and their function. Cell. 2007;128(4):693-705.

94.   Krstic A, Mojsilovic S, Jovcic G, Bugarski D. The potential of interleukin-17 to mediate hematopoietic response. Immunol Res. 2012;52(1-2):34-41.

95.   Kuilman T, Michaloglou C, Mooi WJ, Peeper DS. The essence of senescence. Genes Dev. 2010;24(22):2463-79.

96.   Lafaurie-Janvore J, Maiuri P, Wang I, Pinot M, Manneville JB, Betz T, et al. ESCRT-III assembly and cytokinetic abscission are induced by tension release in the intercellular bridge. Science. 2013;339(6127):1625-9.

97.   Lambert MP, Ancey PB, Esposti DD, Cros MP, Sklias A, Scoazec JY, et al. Aberrant DNA methylation of imprinted loci in hepatocellular carcinoma and after in vitro exposure to common risk factors. Clin Epigenetics. 2015;7:15.

98.   Lancaster OM, Le Berre M, Dimitracopoulos A, Bonazzi D, Zlotek-Zlotkiewicz E, Picone R, et al. Mitotic rounding alters cell geometry to ensure efficient bipolar spindle formation. Dev Cell. 2013;25(3):270-83.

99.   Lanzoni G, Cardinale V, Carpino G. The hepatic, biliary, and pancreatic network of stem/progenitor cell niches in humans: A new reference frame for disease and regeneration. Hepatology. 2016;64(1):277-86.

100.   Laurent L, Wong E, Li G, Huynh T, Tsirigos A, Ong CT, et al. Dynamic changes in the human methylome during differentiation. Genome Res. 2010;20(3):320-31.

101.   Lax E, Sapozhnikov DM. Adult Neural Stem Cell Multipotency and Differentiation Are Directed by the Methyl-CpG-Binding Protein MBD1. J Neurosci. 2017;37(16):4228-30.

102.   Le Berre M, Aubertin J, Piel M. Fine control of nuclear confinement identifies a threshold deformation leading to lamina rupture and induction of specific genes. Integr Biol (Camb). 2012;4(11):1406-14.

103.   Le Berre M, Zlotek-Zlotkiewicz E, Bonazzi D, Lautenschlaeger F, Piel M. Methods for two-dimensional cell confinement. Methods Cell Biol. 2014;121:213-29.

104.   Le HQ, Ghatak S, Yeung CY, Tellkamp F, Günschmann C, Dieterich C, et al. Mechanical regulation of transcription controls Polycomb-mediated gene silencing during lineage commitment. Nat Cell Biol. 2016;18(8):864-75.

105.   Lee Y. The role of interleukin-17 in bone metabolism and inflammatory skeletal diseases. BMB Rep. 2013;46(10):479-83.

106.   Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, et al. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol. 2003;13(14):1192-200.

107.   Leonardo TR, Schultheisz HL, Loring JF, Laurent LC. The functions of microRNAs in pluripotency and reprogramming. Nat Cell Biol. 2012;14(11):1114-21.

108.   Li C, Fan Y, Li G, Xu X, Duan J, Li R, et al. DNA methylation reprogramming of functional elements during mammalian embryonic development. Cell Discov. 2018;4:41.

109.   Li W, Ren G, Huang Y, Su J, Han Y, Li J, et al. Mesenchymal stem cells: a double-edged sword in regulating immune responses. Cell Death Differ. 2012;19(9):1505-13.

110.   Li Z, Li K, Zhu L, Kan Q, Yan Y, Kumar P, et al. Inhibitory effect of IL-17 on neural stem cell proliferation and neural cell differentiation. BMC Immunol. 2013;14:20.

111.   Liang G, He J, Zhang Y. Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming. Nat Cell Biol. 2012;14(5):457-66.

112.   Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009;462(7271):315-22.

113.   Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G, et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature. 2011;471(7336):68-73.

114.   Liu T, Zhang L, Joo D, Sun S-C. NF-κB signaling in inflammation. Signal Transduction and Targeted Therapy. 2017;2(1):17023.

115.   Liu YJ, Le Berre M, Lautenschlaeger F, Maiuri P, Callan-Jones A, Heuzé M, et al. Confinement and low adhesion induce fast amoeboid migration of slow mesenchymal cells. Cell. 2015;160(4):659-72.

116.   Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402-8.

117.   Loewer S, Cabili MN, Guttman M, Loh YH, Thomas K, Park IH, et al. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. Nat Genet. 2010;42(12):1113-7.

118.   Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

119.   Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997;389(6648):251-60.

120.   Luni C, Giulitti S, Serena E, Ferrari L, Zambon A, Gagliano O, et al. High-efficiency cellular reprogramming with microfluidics. Nat Methods. 2016;13(5):446-52.

121.   Lutsik P, Feuerbach L, Arand J, Lengauer T, Walter J, Bock C. BiQ Analyzer HT: locus-specific analysis of DNA methylation by high-throughput bisulfite sequencing. Nucleic Acids Res. 2011;39(Web Server issue):W551-6.

122.   Lv FJ, Tuan RS, Cheung KM, Leung VY. Concise review: the surface markers and identity of human mesenchymal stem cells. Stem Cells. 2014;32(6):1408-19.

123.   Martin C, Zhang Y. Mechanisms of epigenetic inheritance. Curr Opin Cell Biol. 2007;19(3):266-72.

124.   Martinez-Arguelles DB, Lee S, Papadopoulos V. In silico analysis identifies novel restriction enzyme combinations that expand reduced representation bisulfite sequencing CpG coverage. BMC Res Notes. 2014;7:534.

125.   Mattick JS, Amaral PP, Dinger ME, Mercer TR, Mehler MF. RNA regulation of epigenetic processes. Bioessays. 2009;31(1):51-9.

126.   Mehanni SS, Ibrahim NF, Hassan AR, Rashed LA. New approach of bone marrow-derived mesenchymal stem cells and human amniotic epithelial cells applications in accelerating wound healing of irradiated albino rats. Int J Stem Cells. 2013;6(1):45-54.

127.   Meissner A. Epigenetic modifications in pluripotent and differentiated cells. Nat Biotechnol. 2010;28(10):1079-88.

128.   Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res. 2005;33(18):5868-77.

129.   Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008;454(7205):766-70.

130.   Messerschmidt DM, Knowles BB, Solter D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev. 2014;28(8):812-28.

131.   Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P, et al. Dissecting direct reprogramming through integrative genomic analysis. Nature. 2008;454(7200):49-55.

132.   Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature. 2007;448(7153):553-60.

133.   Miroshnikova YA, Nava MM, Wickström SA. Emerging roles of mechanical forces in chromatin regulation. J Cell Sci. 2017;130(14):2243-50.

134.   Mojsilović S, Jauković A, Santibañez JF, Bugarski D. Interleukin-17 and its implication in the regulation of differentiation and function of hematopoietic and mesenchymal stem cells. Mediators Inflamm. 2015;2015:470458.

135.   Murphy WL, McDevitt TC, Engler AJ. Materials as stem cell regulators. Nat Mater. 2014;13(6):547-57.

136.   Nakamura T, Liu YJ, Nakashima H, Umehara H, Inoue K, Matoba S, et al. PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. Nature. 2012;486(7403):415-9.

137.   Nakamura T, Takeuchi K, Muraoka S, Takezoe H, Takahashi N, Mori N. A neurally enriched coronin-like protein, ClipinC, is a novel candidate for an actin cytoskeleton-cortical membrane-linking protein. J Biol Chem. 1999;274(19):13322-7.

138.   Narayan R, Gangadharan B, Hantz O, Antrobus R, García A, Dwek RA, et al. Proteomic analysis of HepaRG cells: a novel cell line that supports hepatitis B virus infection. J Proteome Res. 2009;8(1):118-22.

139.   Nava MM, Miroshnikova YA, Biggs LC, Whitefield DB, Metge F, Boucas J, et al. Heterochromatin-Driven Nuclear Softening Protects the Genome against Mechanical Stress-Induced Damage. Cell. 2020;181(4):800-17.e22.

140.   Nishino K, Toyoda M, Yamazaki-Inoue M, Fukawatase Y, Chikazawa E, Sakaguchi H, et al. DNA methylation dynamics in human induced pluripotent stem cells over time. PLoS Genet. 2011;7(5):e1002085.

141.   Nishino K, Umezawa A. DNA methylation dynamics in human induced pluripotent stem cells. Hum Cell. 2016;29(3):97-100.

142.   Nishizawa M, Chonabayashi K, Nomura M, Tanaka A, Nakamura M, Inagaki A, et al. Epigenetic Variation between Human Induced Pluripotent Stem Cell Lines Is an Indicator of Differentiation Capacity. Cell Stem Cell. 2016;19(3):341-54.

143.   Oh BK, Kim H, Park HJ, Shim YH, Choi J, Park C, et al. DNA methyltransferase expression and DNA methylation in human hepatocellular carcinoma and their clinicopathological correlation. Int J Mol Med. 2007;20(1):65-73.

144.   Ohi Y, Qin H, Hong C, Blouin L, Polo JM, Guo T, et al. Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. Nat Cell Biol. 2011;13(5):541-9.

145.   Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999;99(3):247-57.

146.   Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature. 2007;448(7154):714-7.

147.   Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, et al. Active demethylation of the paternal genome in the mouse zygote. Curr Biol. 2000;10(8):475-8.

148.   Pagliara S, Franze K, McClain CR, Wylde G, Fisher CL, Franklin RJM, et al. Auxetic nuclei in embryonic stem cells exiting pluripotency. Nat Mater. 2014;13(6):638-44.

149.   Panopoulos AD, Yanes O, Ruiz S, Kida YS, Diep D, Tautenhahn R, et al. The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming. Cell Res. 2012;22(1):168-77.

150.   Parent R, Marion MJ, Furio L, Trépo C, Petit MA. Origin and characterization of a human bipotent liver progenitor cell line. Gastroenterology. 2004;126(4):1147-56.

151.   Picelli S, Björklund Å, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods. 2013;10(11):1096-8.

152.   Picelli S, Faridani OR, Björklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc. 2014;9(1):171-81.

153.   Polo JM, Anderssen E, Walsh RM, Schwarz BA, Nefzger CM, Lim SM, et al. A molecular roadmap of reprogramming somatic cells into iPS cells. Cell. 2012;151(7):1617-32.

154.   Puumala SE, Hoyme HE. Epigenetics in pediatrics. Pediatr Rev. 2015;36(1):14-21.

155.   Qiu X, Hother C, Ralfkiær UM, Søgaard A, Lu Q, Workman CT, et al. Equitoxic doses of 5-azacytidine and 5-aza-2'deoxycytidine induce diverse immediate and overlapping heritable changes in the transcriptome. PLoS One. 2010;5(9).

156.   Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc Natl Acad Sci U S A. 2000;97(10):5237-42.

157.   Rand TA, Sutou K, Tanabe K, Jeong D, Nomura M, Kitaoka F, et al. MYC Releases Early Reprogrammed Human Cells from Proliferation Pause via Retinoblastoma Protein Inhibition. Cell Rep. 2018;23(2):361-75.

158.   Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. Nature. 2007;447(7143):425-32.

159.   Richmond TJ, Davey CA. The structure of DNA in the nucleosome core. Nature. 2003;423(6936):145-50.

160.   Routh A, Sandin S, Rhodes D. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. Proc Natl Acad Sci U S A. 2008;105(26):8872-7.

161.   Rui L. Energy metabolism in the liver. Compr Physiol. 2014;4(1):177-97.

162.   Ruiz SA, Chen CS. Emergence of patterned stem cell differentiation within multicellular structures. Stem Cells. 2008;26(11):2921-7.

163.   Russell SB, Russell JD, Trupin KM, Gayden AE, Opalenik SR, Nanney LB, et al. Epigenetically altered wound healing in keloid fibroblasts. J Invest Dermatol. 2010;130(10):2489-96.

164.   Salhab A, Nordström K, Gasparoni G, Kattler K, Ebert P, Ramirez F, et al. A comprehensive analysis of 195 DNA methylomes reveals shared and cell-specific features of partially methylated domains. Genome Biol. 2018;19(1):150.

165.   Samavarchi-Tehrani P, Golipour A, David L, Sung HK, Beyer TA, Datti A, et al. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell. 2010;7(1):64-77.

166.   Santos F, Hendrich B, Reik W, Dean W. Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev Biol. 2002;241(1):172-82.

167.   Sarkar U, Rivera-Burgos D, Large EM, Hughes DJ, Ravindra KC, Dyer RL, et al. Metabolite profiling and pharmacokinetic evaluation of hydrocortisone in a perfused three-dimensional human liver bioreactor. Drug Metab Dispos. 2015;43(7):1091-9.

168.   Schaefer M, Hagemann S, Hanna K, Lyko F. Azacytidine inhibits RNA methylation at DNMT2 target sites in human cancer cell lines. Cancer Res. 2009;69(20):8127-32.

169.   Shashidhar S, Lorente G, Nagavarapu U, Nelson A, Kuo J, Cummins J, et al. GPR56 is a GPCR that is overexpressed in gliomas and functions in tumor cell adhesion. Oncogene. 2005;24(10):1673-82.

170.   Shen L, Inoue A, He J, Liu Y, Lu F, Zhang Y. Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. Cell Stem Cell. 2014;15(4):459-71.

171.   Sigalotti L, Fratta E, Coral S, Cortini E, Covre A, Nicolay HJ, et al. Epigenetic drugs as pleiotropic agents in cancer treatment: biomolecular aspects and clinical applications. J Cell Physiol. 2007;212(2):330-44.

172.   Simmons D. Epigenetic Influences and Disease2008; 1(1) :6.

173.   Smith JR, Venable S, Roberts TW, Metter EJ, Monticone R, Schneider EL. Relationship between in vivo age and in vitro aging: assessment of 669 cell cultures derived from members of the Baltimore Longitudinal Study of Aging. J Gerontol A Biol Sci Med Sci. 2002;57(6):B239-46.

174.   Smith ZD, Gu H, Bock C, Gnirke A, Meissner A. High-throughput bisulfite sequencing in mammalian genomes. Methods. 2009;48(3):226-32.

175.   Song F, Smith JF, Kimura MT, Morrow AD, Matsuyama T, Nagase H, et al. Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. Proc Natl Acad Sci U S A. 2005;102(9):3336-41.

176.   Sridharan R, Gonzales-Cope M, Chronis C, Bonora G, McKee R, Huang C, et al. Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1γ in reprogramming to pluripotency. Nat Cell Biol. 2013;15(7):872-82.

177.   Sridharan R, Tchieu J, Mason MJ, Yachechko R, Kuoy E, Horvath S, et al. Role of the murine reprogramming factors in the induction of pluripotency. Cell. 2009;136(2):364-77.

178.   Subramaniam D, Thombre R, Dhar A, Anant S. DNA methyltransferases: a novel target for prevention and therapy. Front Oncol. 2014;4:80.

179.   Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017;45(D1):D362-D8.

180.   Tajik A, Zhang Y, Wei F, Sun J, Jia Q, Zhou W, et al. Transcription upregulation via force-induced direct stretching of chromatin. Nat Mater. 2016;15(12):1287-96.

181.   Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 2007;131(5):861-72.

182.   Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126(4):663-76.

183.   Tamaru H. Confining euchromatin/heterochromatin territory: jumonji crosses the line. Genes Dev. 2010;24(14):1465-78.

184.   Tanabe S. Role of mesenchymal stem cells in cell life and their signaling. World J Stem Cells. 2014;6(1):24-32.

185.   Tang K, Wang L, Yang Z, Sui Y, Li L, Huang Y, et al. Comparison of hydrocortisone and prednisone in the glucocorticoid replacement therapy post-adrenalectomy of Cushing's Syndrome. Oncotarget. 2017;8(62):106113-20.

186.   Tao R, Li J, Xin J, Wu J, Guo J, Zhang L, et al. Methylation profile of single hepatocytes derived from hepatitis B virus-related hepatocellular carcinoma. PLoS One. 2011;6(5):e19862.

187.   Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol. 2013;31(1):46-53.

188.   Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28(5):511-5.

189.   Troadec MB, Glaise D, Lamirault G, Le Cunff M, Guérin E, Le Meur N, et al. Hepatocyte iron loading capacity is associated with differentiation and repression of motility in the HepaRG cell line. Genomics. 2006;87(1):93-103.

190.   Tusnády GE, Simon I, Váradi A, Arányi T. BiSearch: primer-design and search tool for PCR on bisulfite-treated genomes. Nucleic Acids Res. 2005;33(1):e9.

191.   Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419.

192.   Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, et al. Towards a knowledge-based Human Protein Atlas. Nat Biotechnol. 2010;28(12):1248-50.

193.   Uhler C, Shivashankar GV. Geometric control and modeling of genome reprogramming. Bioarchitecture. 2016;6(4):76-84.

194.   van der Wijst MG, Venkiteswaran M, Chen H, Xu GL, Plösch T, Rots MG. Local chromatin microenvironment determines DNMT activity: from DNA methyltransferase to DNA demethylase or DNA dehydroxymethylase. Epigenetics. 2015;10(8):671-6.

195.   van Wenum M, Adam AAA, van der Mark VA, Chang JC, Wildenberg ME, Hendriks EJ, et al. Oxygen drives hepatocyte differentiation and phenotype stability in liver cell lines. J Cell Commun Signal. 2018;12(3):575-88.

196.   Vining KH, Mooney DJ. Mechanical forces direct stem cell behaviour in development and regeneration. Nat Rev Mol Cell Biol. 2017;18(12):728-42.

197.   von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. STRING: a database of predicted functional associations between proteins. Nucleic Acids Res. 2003;31(1):258-61.

198.   Wang KY, Chen CC, Shen CK. Active DNA demethylation of the vertebrate genomes by DNA methyltransferases: deaminase, dehydroxymethylase or demethylase? Epigenomics. 2014;6(3):353-63.

199.   Wang L, Sun J, Wu H, Liu S, Wang J, Wu B, et al. Systematic assessment of reduced representation bisulfite sequencing to human blood samples: A promising method for large-sample-scale epigenomic studies. J Biotechnol. 2012;157(1):1-6.

200.   Wang N, Tytell JD, Ingber DE. Mechanotransduction at a distance: mechanically coupling the extracellular matrix with the nucleus. Nat Rev Mol Cell Biol. 2009;10(1):75-82.

201.   Wang PY, Thissen H, Kingshott P. Modulation of human multipotent and pluripotent stem cells using surface nanotopographies and surface-immobilised bioactive signals: A review. Acta Biomater. 2016;45:31-59.

202.   Wang T, Wu H, Li Y, Szulwach KE, Lin L, Li X, et al. Subtelomeric hotspots of aberrant 5-hydroxymethylcytosine-mediated epigenetic modifications during reprogramming to pluripotency. Nat Cell Biol. 2013;15(6):700-11.

203.   Wang Y, Chen X, Cao W, Shi Y. Plasticity of mesenchymal stem cells in immunomodulation: pathological and therapeutic implications. Nat Immunol. 2014;15(11):1009-16.

204.   Weber M, Schübeler D. Genomic patterns of DNA methylation: targets and function of an epigenetic mark. Curr Opin Cell Biol. 2007;19(3):273-80.

205.   Wozniak MA, Chen CS. Mechanotransduction in development: a growing role for contractility. Nat Rev Mol Cell Biol. 2009;10(1):34-43.

206.   Wrigglesworth J, Ancelin ML, Ritchie K, Ryan J. Association between DNA methylation of the KITLG gene and cortisol levels under stress: a replication study. Stress. 2019;22(1):162-8.

207.   Wu H, Zhang Y. Reversing DNA methylation: mechanisms, genomics, and biological functions. Cell. 2014;156(1-2):45-68.

208.   Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. Nat Rev Mol Cell Biol. 2010;11(9):607-20.

209.   Xu Y, Zhang JJ, Grifo JA, Krey LC. DNA methylation patterns in human tripronucleate zygotes. Mol Hum Reprod. 2005;11(3):167-71.

210.   Yamanaka S, Blau HM. Nuclear reprogramming to a pluripotent state by three approaches. Nature. 2010;465(7299):704-12.

211.   Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G. Gene body methylation can alter gene expression and is a therapeutic target in cancer. Cancer Cell. 2014;26(4):577-90.

212.   Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. 1997;13(8):335-40.

213.   Zhang G, Pradhan S. Mammalian epigenetic mechanisms. IUBMB Life. 2014;66(4):240-56.

214.   Zhang J, Nuebel E, Daley GQ, Koehler CM, Teitell MA. Metabolic regulation in pluripotent stem cells during reprogramming and self-renewal. Cell Stem Cell. 2012;11(5):589-95.

215.   Zhang P, Andrianakos R, Yang Y, Liu C, Lu W. Kruppel-like factor 4 (Klf4) prevents embryonic stem (ES) cell differentiation by regulating Nanog gene expression. J Biol Chem. 2010;285(12):9180-9.

216.   Ziller MJ, Gu H, Müller F, Donaghey J, Tsai LT, Kohlbacher O, et al. Charting a dynamic DNA methylation landscape of the human genome. Nature. 2013;500(7463):477-81.

217.   Ziyad S, Iruela-Arispe ML. Molecular mechanisms of tumor angiogenesis. Genes Cancer. 2011;2(12):1085-96.

Chapter 7 Appendices

7.1. Supplementary data 1

Table 7.1 Total read counts and number of CpG sites at 10-fold coverage of all samples. Samples with asterisk (*) were re-sequenced to obtain more comparable CpG sites after 10x coverage filtering.

Samples	Total read counts	Total CpG sites after meth. called	Total CpG sites after 10x cov.	Read count of each sequencing	
Precursor	5.18E+07	3.20E+06	3.20E+06		
RP_P0nonPC	4.04E+07	3.01E+06	3.01E+06		
RP_P0PC	1.10E+08	3.95E+06	3.94E+06		
RP_P1PC*	1.20E+08	5.24E+06	5.24E+06	3.30E+07	8.74E+07
RP_P1PC+A	6.66E+07	3.49E+06	3.49E+06		
RP_P10nonPC*	8.30E+07	4.67E+06	4.66E+06	4.11E+07	4.19E+07
RP_P10PC	4.61E+07	4.13E+06	4.12E+06		
RP_P10PC+A*	7.60E+07	4.39E+06	4.39E+06	2.74E+07	4.86E+07
RD_P0nonPC+A	4.24E+07	2.91E+06	2.91E+06		
RD_P0PC+Hydro	5.69E+07	3.33E+06	3.33E+06		
RD_P0PC+DMSO*	6.27E+07	3.96E+06	3.96E+06	2.61E+07	3.66E+07
RD_P20*	7.84E+07	4.51E+06	4.50E+06	2.35E+07	5.50E+07

Table 7.2 Total read numbers of each amplicons. The table also provides additional information including the number of CpG sites on Miseq amplicons, number of CpG found in both Miseq amplicons and RRBS, correlation coefficient between RRBS and LDS and percentage of low coverage CpGs (cov.<10x) included in each amplicon.

| Amplicons | Total read numbers | No. of CpG sites | | Correlation coefficient (R) | | % of CpGs with cov.<10x |
		LDS	Overlapped with RRBS	All CpGs	All CpGs excluding low coverage CpGs	
AMIGO2	1.90E+05	16	14	0.78	-	-
CSMD3	1.81E+05	6	4	0.61	-	-
DTD1	2.10E+05	9	8	0.38	-	-
DUS3L	1.87E+05	16	11	0.90	0.88	4.55
FA2H	1.34E+04	29	13	0.59	0.60	18.75
FAM5C	7.01E+04	20	13	0.66	0.64	28.20
OCRL	1.49E+05	29	12	0.49	0.66	43.75
SPARC	1.51E+05	7	4	0.54	0.14	25.00
TF	5.74E+05	8	7	0.50	0.10	53.57
TUBA1A	7.94E+04	30	17	0.21	0.04	12.02
ZNF814	1.98E+05	14	5	0.69	0.56	23.33

Table 7.3 Total read counts of all samples. Table shows read counts after sequencing and mapping separated by methods of sequencing.

| Samples | RRBS-sequencing | | | mRNA-sequencing |
	Total read counts	Total CpG sites after meth. called	Total CpG sites after 10x cov.	Total read counts
CC1	2.00E+07	1.79E+06	1.79E+06	2.69E+07
CC2	2.20E+07	1.78E+06	1.77E+06	5.03E+07
NonCC1	1.87E+07	1.05E+06	1.05E+06	1.93E+07
NonCC2	1.48E+07	1.39E+06	1.39E+06	3.66E+07

Table 7.4 Description of expanded 18-state of chromatin segmentation.

STATE NO.	MNEMONIC	DESCRIPTION
1	TssA	Active TSS
2	TssFlnk	Flanking TSS
3	TssFlnkU	Flanking TSS Upstream
4	TssFlnkD	Flanking TSS Downstream
5	Tx	Strong transcription
6	TxWk	Weak transcription
7	EnhG1	Genic enhancer1
8	EnhG2	Genic enhancer2
9	EnhA1	Active Enhancer 1
10	EnhA2	Active Enhancer 2
11	EnhWk	Weak Enhancer
12	ZNF/Rpts	ZNF genes & repeats
13	Het	Heterochromatin
14	TssBiv	Bivalent/Poised TSS
15	EnhBiv	Bivalent Enhancer
16	ReprPC	Repressed PolyComb
17	ReprPCWk	Weak Repressed PolyComb
18	Quies	Quiescent/Low
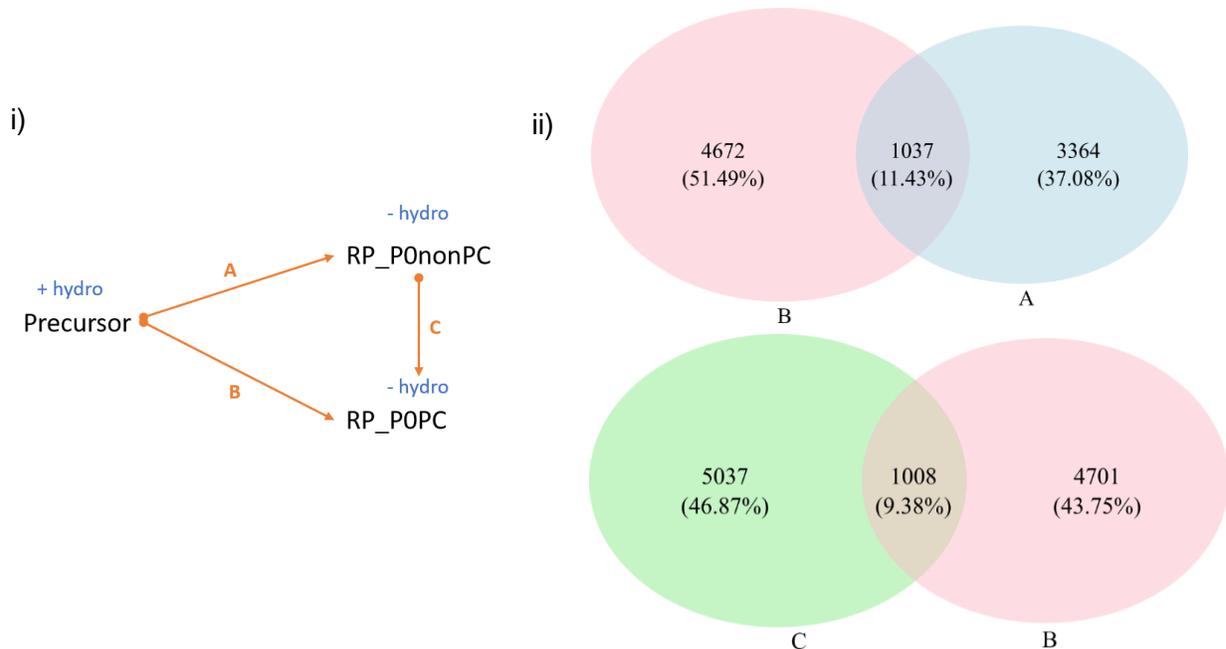
Figure 7.1 Common DMRs emerging from physical constraint. i) shows each pairwise comparison (A-C) of PC. ii) Overlapping of A–B and B-C were extracted to see whether common DMRs of B were greater in A or in C.
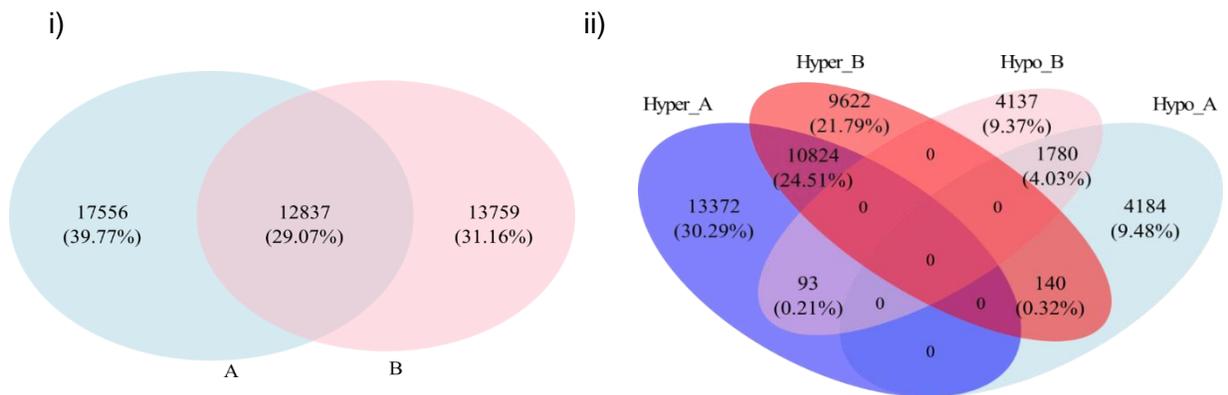
Figure 7.2 pcDMRs found in prolonged cultivation. i) Overlapping of prolonged nonPC (A) and PC (B). ii) The overlapping is then performed separately by methylation status, yielding 4 groups of pcDMRs. Notably, pcDMRs were highly conserved in hypermethylation direction (group 3).

(1) Hypomethylated DMRs in A → Hypomethylated DMRs in B (4.03%)

(2) Hypomethylated DMRs in A → Hypermethylated DMRs in B (0.32%)

(3) Hypermethylated DMRs in A → Hypermethylated DMRs in B (24.51%)

(4) Hypermethylated DMRs in A → Hypomethylated DMRs in B (0.21%)
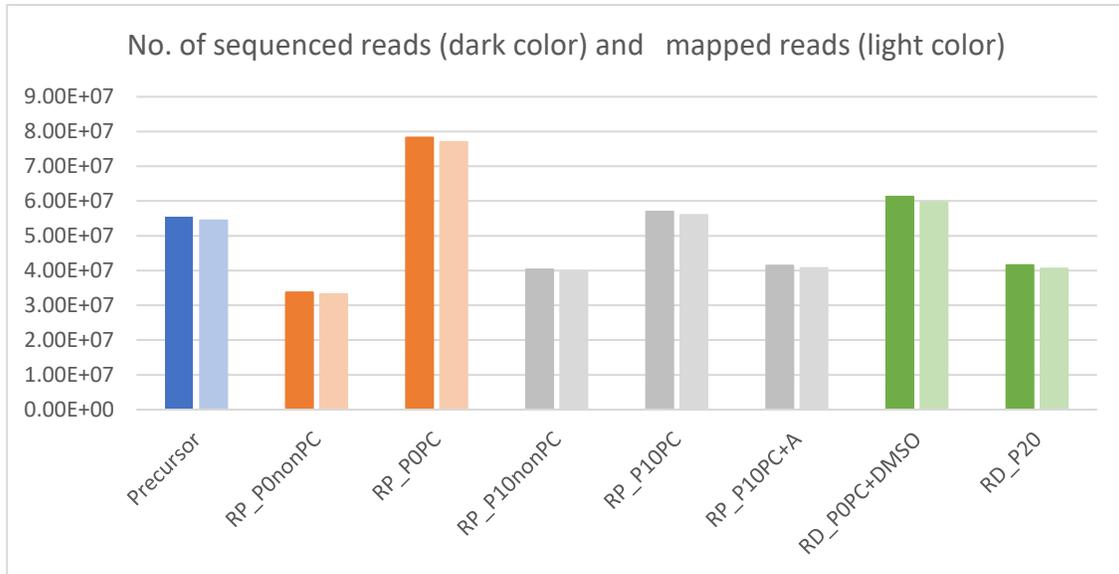
Figure 7.3 Total number of read obtained from all samples of PC experiment. Dark colored bars are total reads obtained after mRNA-sequencing, while light colored bars are reads that mapped to reference genome. Bar colors define states of each samples; blue = precursor, orange = PC, gray = maintenance and green = PCi-differentiation.
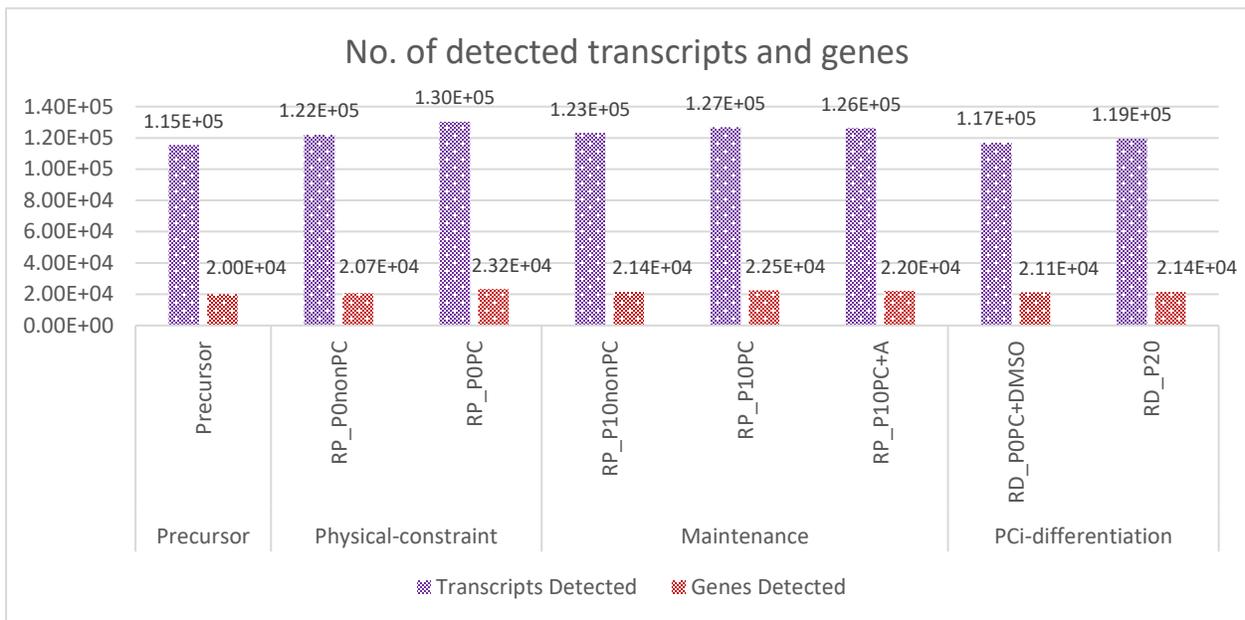
Figure 7.4 Number of detected transcripts and genes after mRNA-sequencing of PC experiment. Purple bars are detected transcripts, while red bars are detected genes.
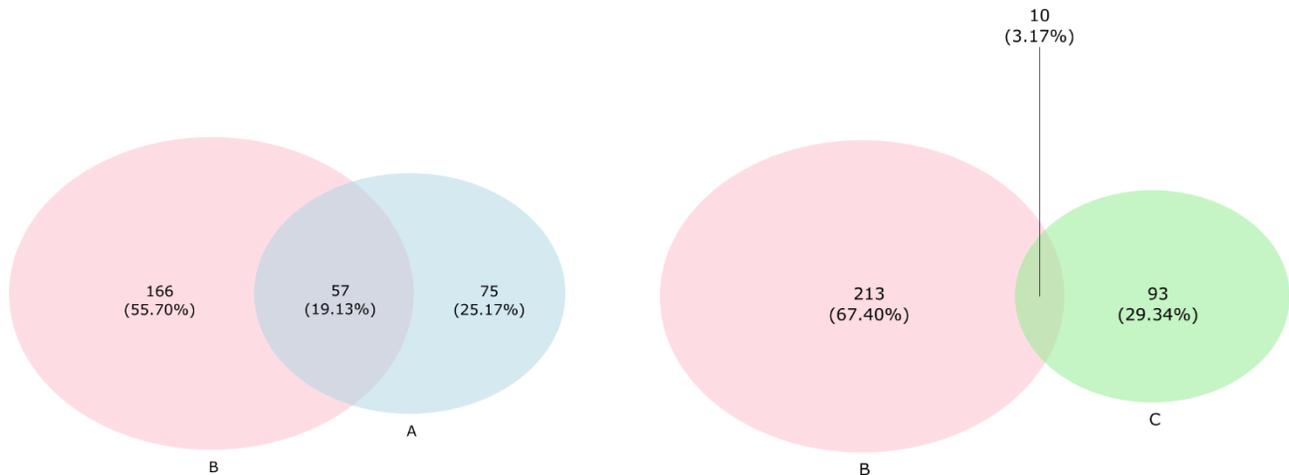
Figure 7.5 Common DEGs emerging from physical constraint. Overlapping of A–B and B-C were extracted to see whether common DEGs of B were greater in A or in C.
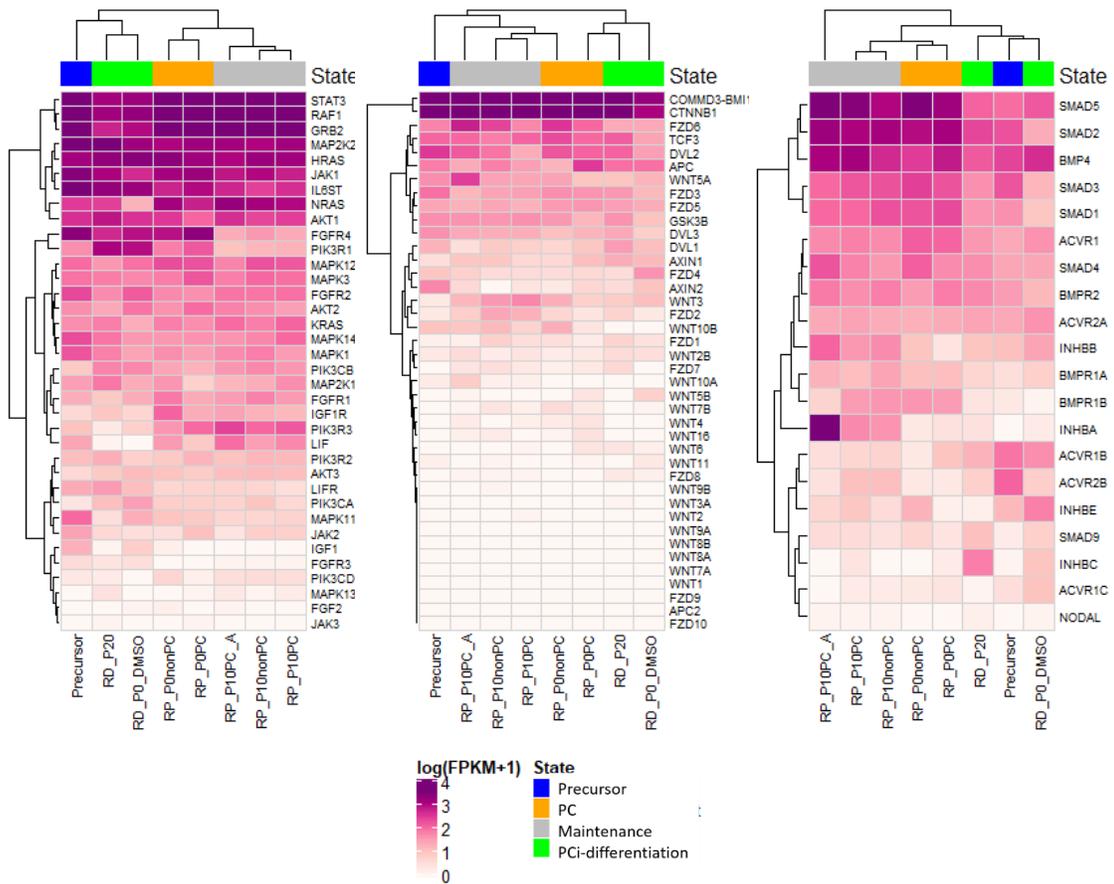
Figure 7.6 mRNA expression of genes in signaling pathways related to pluripotency. From left to right were JAK-MAPK-STAT, WNT and TGFβ signaling pathway, respectively. Heatmap shows the expression value (log (FPKM+1)) of genes. Gradient colors from white to purple represent low to high expression, respectively.
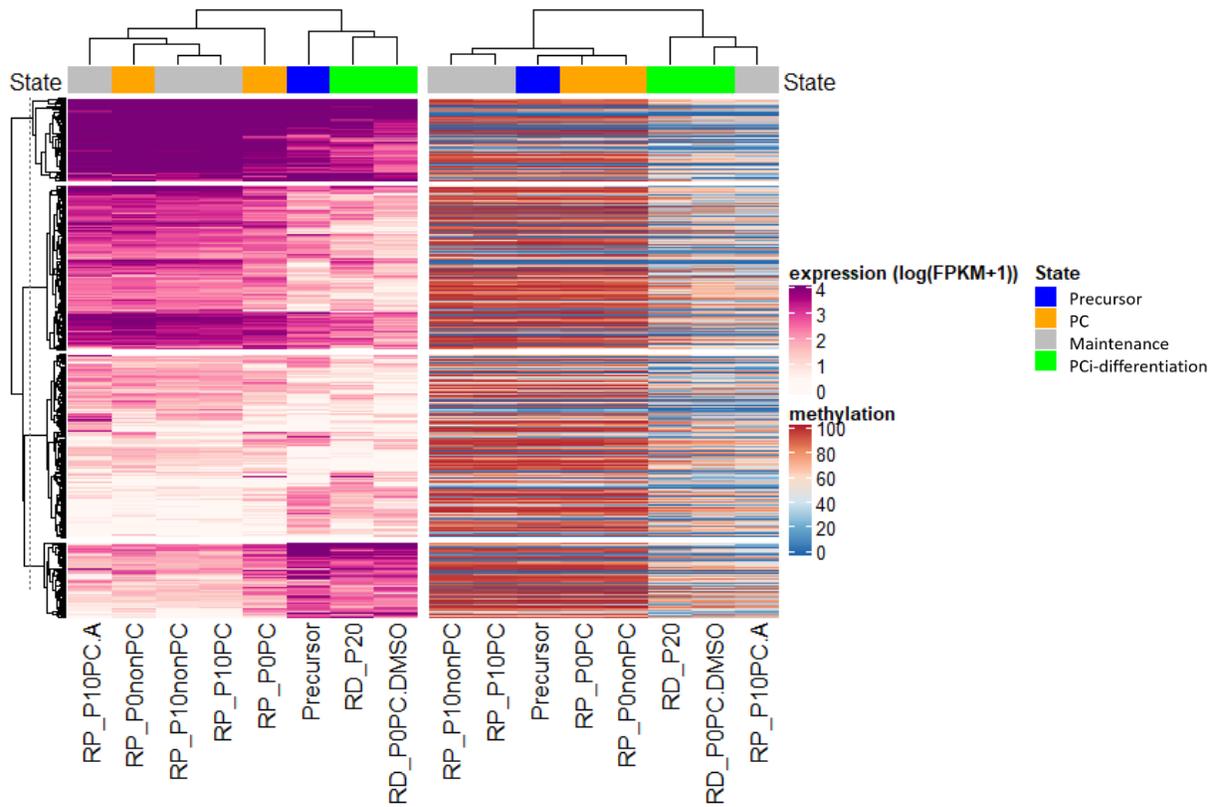
Figure 7.7 Correlation of gene expression and DNA methylation. Left panel of heatmap shows gene expression profile, while right panel shows DNA methylation profile. The colors, blue and red, represent the methylation level from 0 to 1, and gradient colors from white to purple represent low to high expression, respectively.
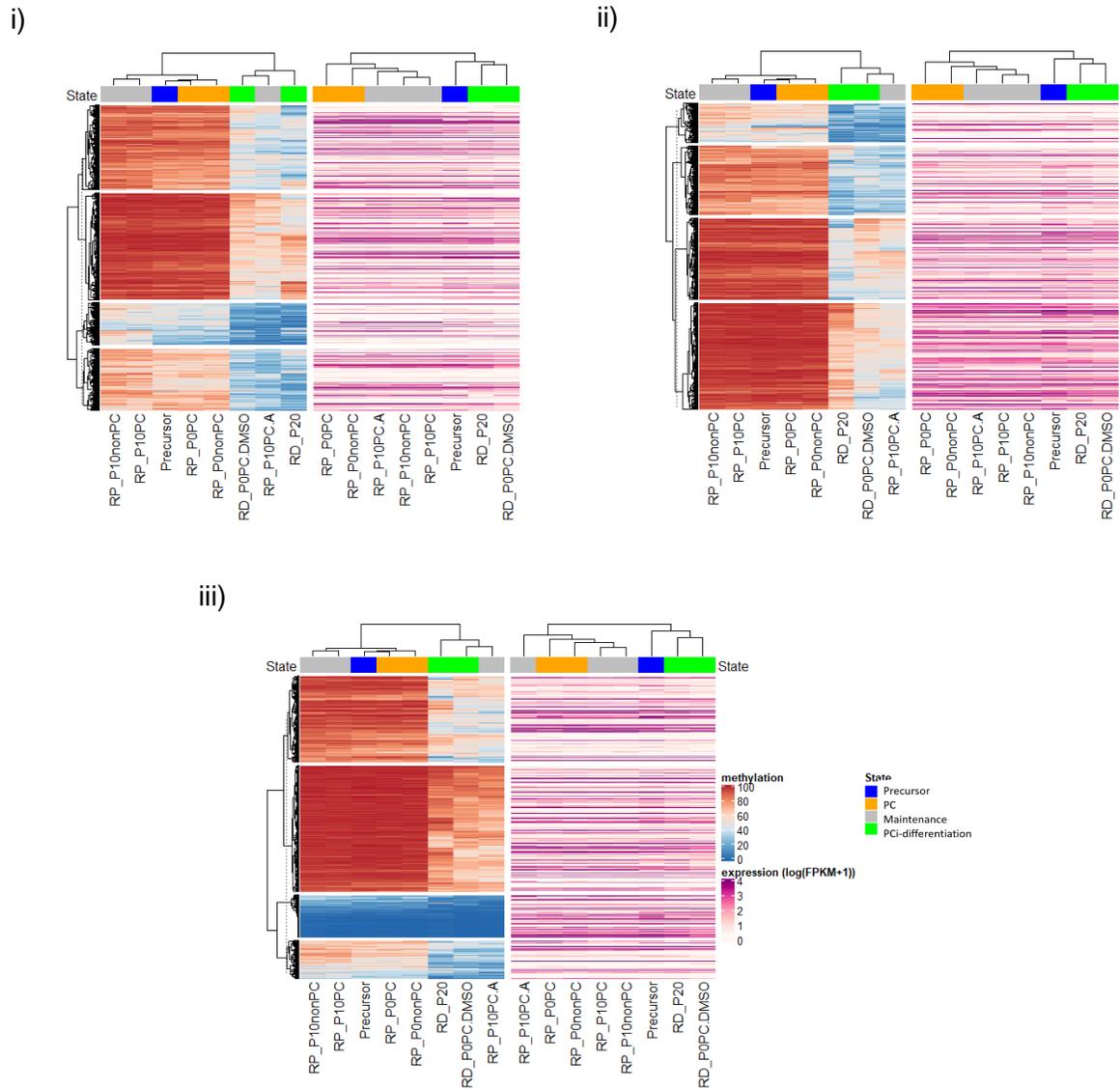
Figure 7.8 Correlation of gene expression and DNA methylation of specific regions. Figures shows heatmaps of methylation of tiling regions related to i) intergenic, ii) intragenic and iii) TTS regions, respectively, and expression of genes associated to those tiling regions. Left panel of those heatmaps shows DNA methylation profile, while right panel shows gene expression profile. The colors, blue and red, represent the methylation level from 0 to 1, and gradient colors from white to purple represent low to high expression, respectively.

i)
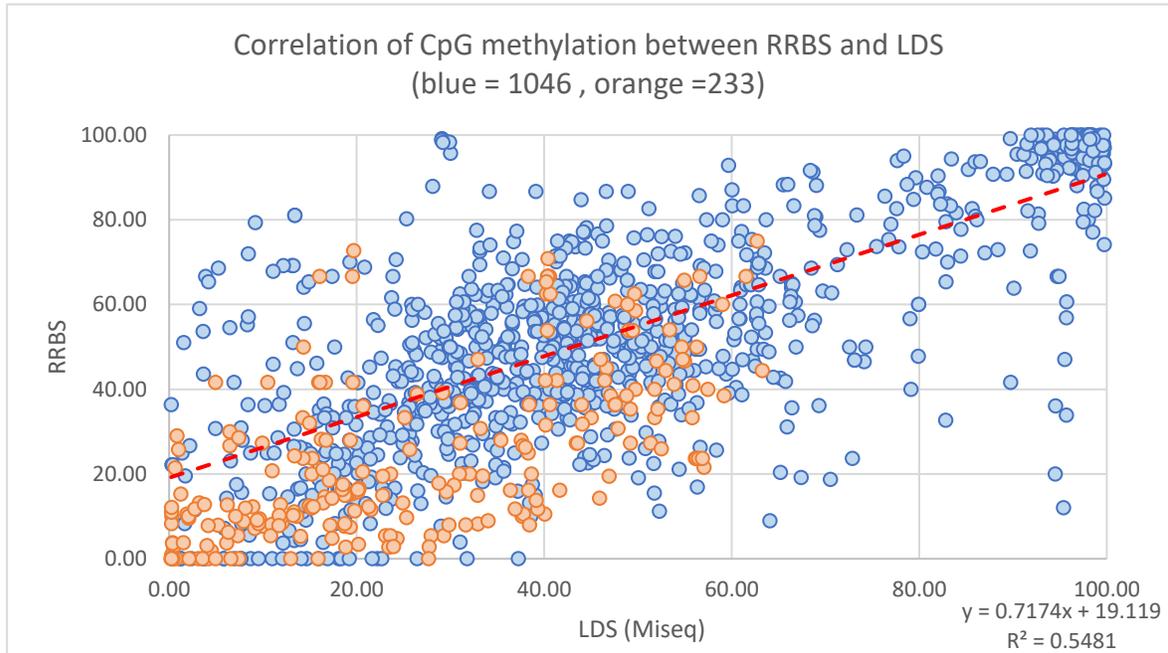
ii)
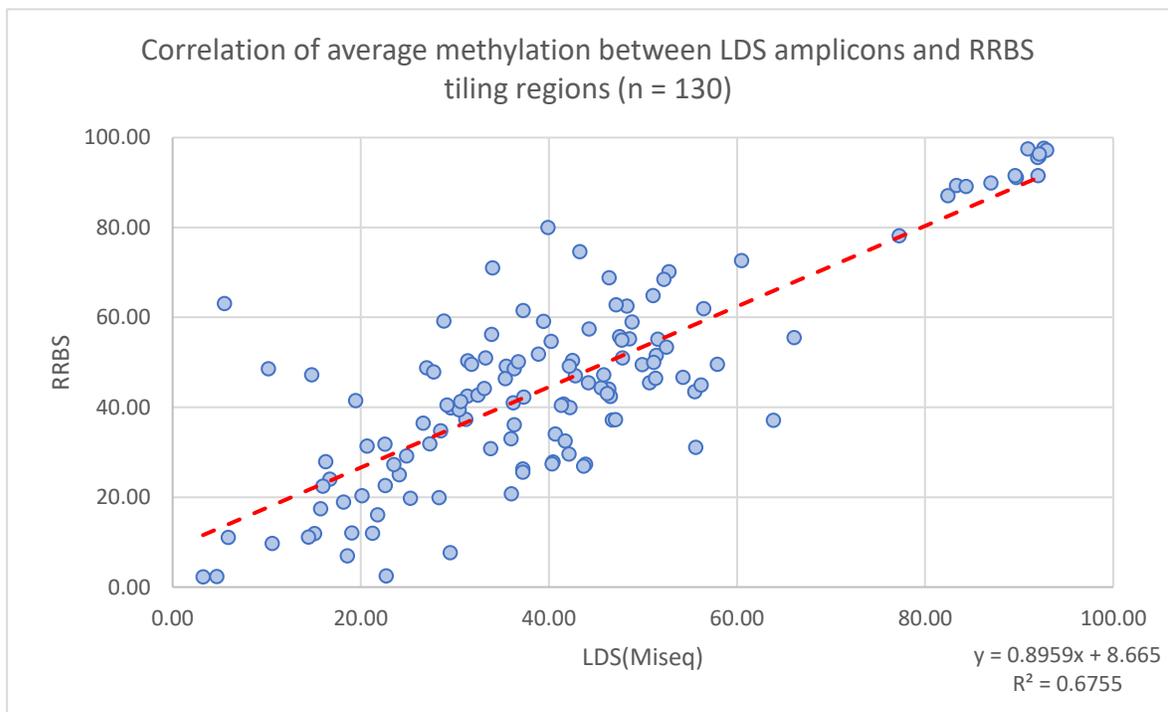
Figure 7.9 Correlation of methylation between RRBS and LDS. i) correlation of CpG methylation after low coverage CpGs are excluded. Blue dot is a CpG site that RRBS coverage are more than 10x, whereas orange dot is a CpG site that RRBS coverage are less than 10x. Red dash line presents a trendline according to blue dot. ii) correlation of average methylation between LDS amplicons and RRBS tiling regions. Equation and R^2 are also shown on the bottom right.
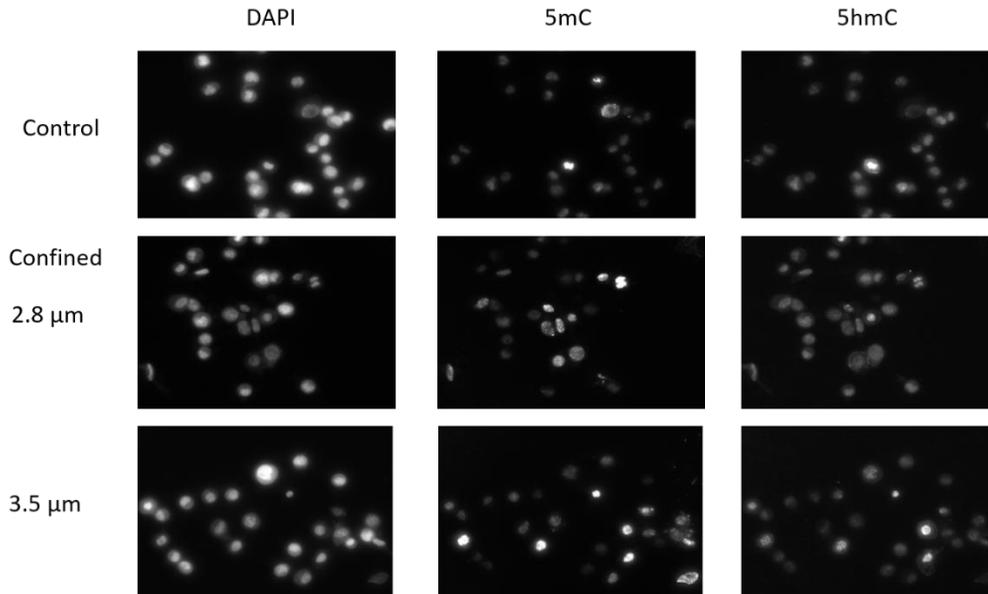
Figure 7.10 Immunofluorescent signal (IF) of DAPI, 5mC and 5hmC of CC conditions. Each condition is defined on the left side of figures, where IF signals are defined on the top of figures.
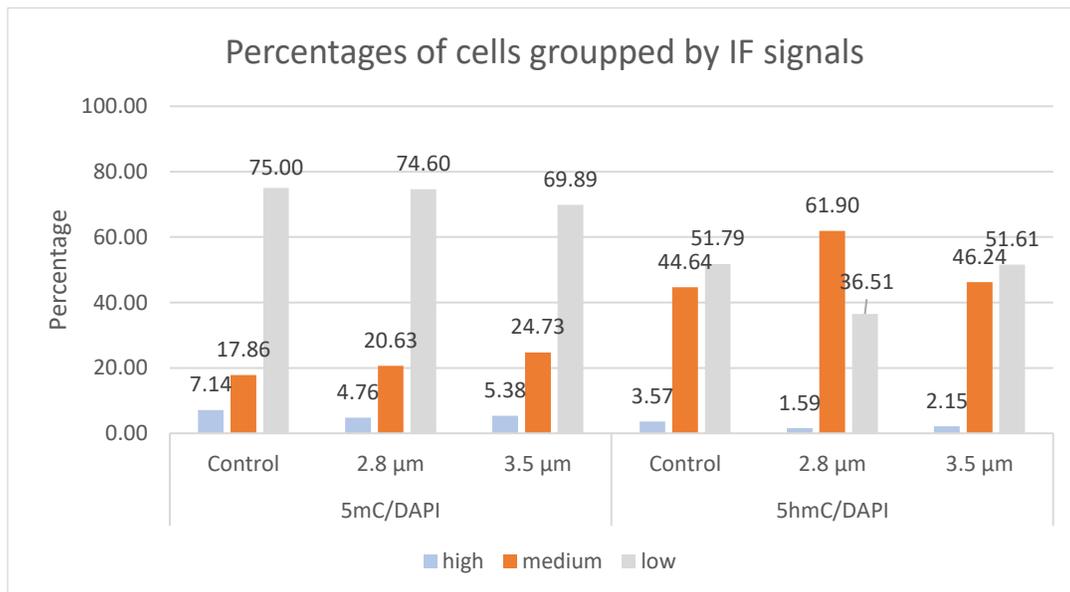
Figure 7.11 Relative IF signals and cell numbers of each condition. Bar graph shows percentage of cells of each Relative IF signals. 5mC/DAPI is on the left panel, while 5hmC/DAPI is on the right panel of the graph.
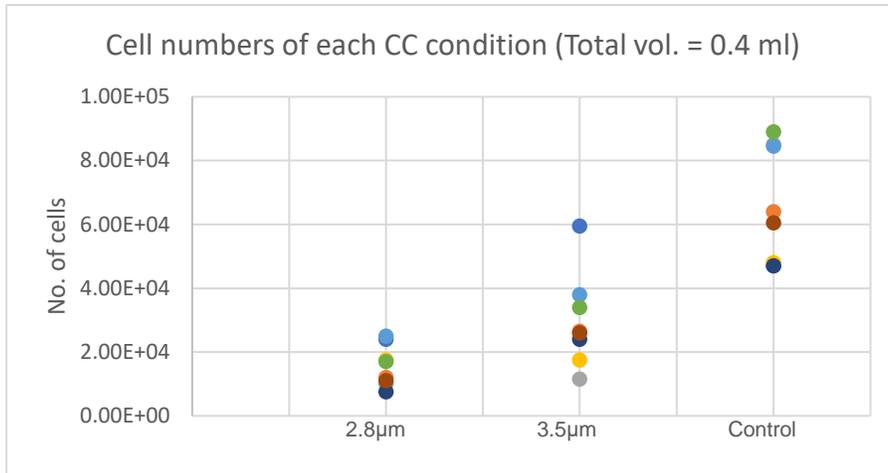
Figure 7.12 Number of cells in total volume (0.4 ml) of each CC condition. Each dot presents number of cells of each replicate. Each CC condition is written below the graphs.
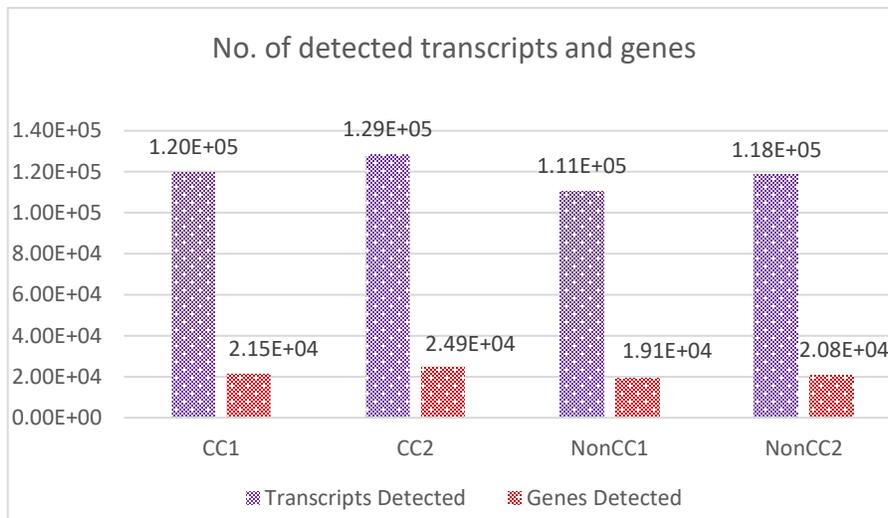
Figure 7.13 Number of detected transcripts and genes after mRNA-sequencing of CC experiment. Purple bars are detected transcripts, while red bars are detected genes.
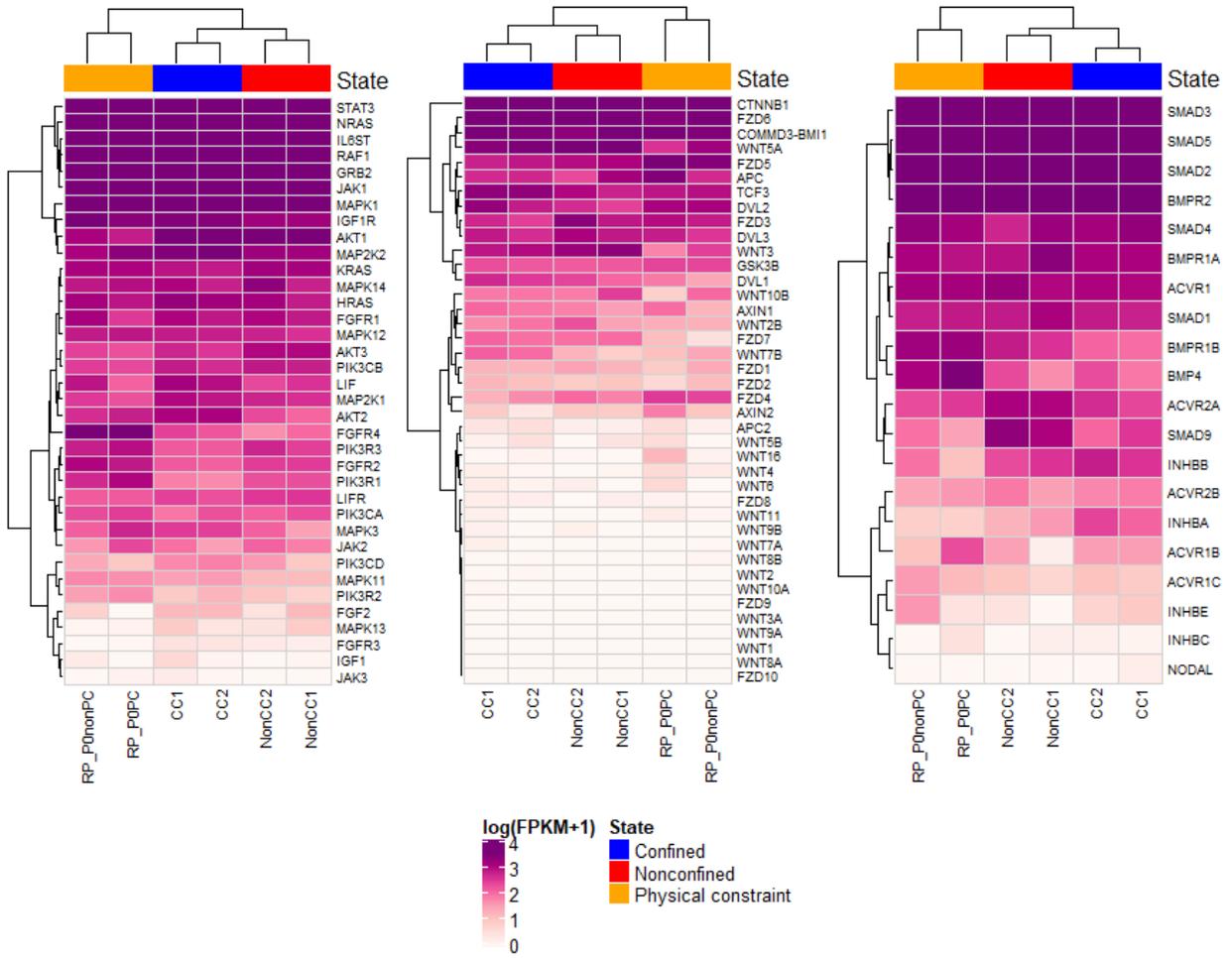
Figure 7.14 mRNA expression of genes in signaling pathways related to pluripotency of PC and CC experiment. From left to right were JAK-MAPK-STAT, WNT and TGFβ signaling pathway, respectively. Heatmap shows the expression value (log (FPKM+1)) of genes. Gradient colors from white to purple represent low to high expression, respectively.
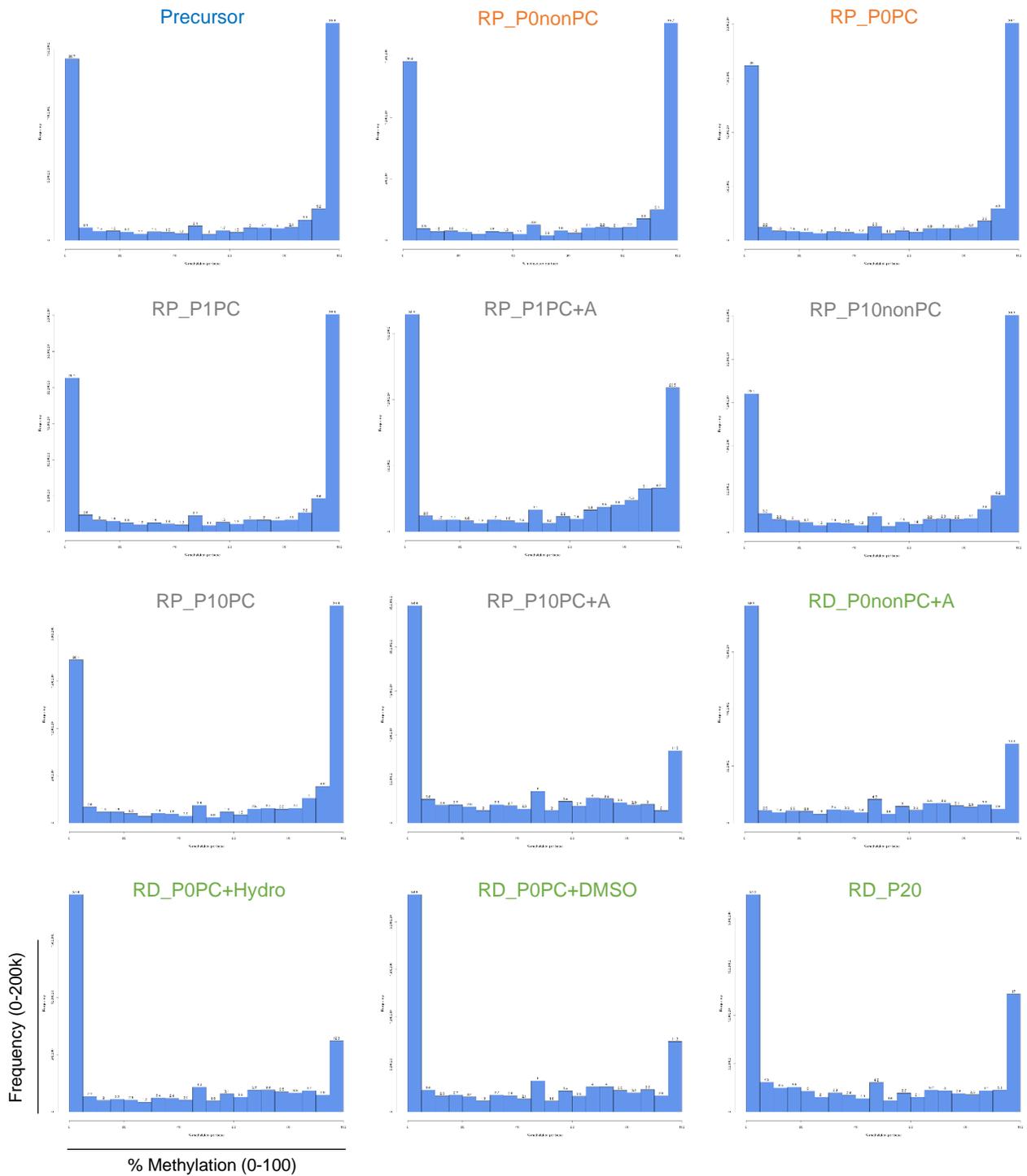
Figure 7.15 Histogram of methylation distribution of all sample. Each figure displays histogram of methylation distribution. Vertical axis represents a frequency (50k, 100k, 150k and 200k), while horizontal axis represents percentages of methylation (0, 20, 40, 60, 80 and 100). Name of each sample is written on the top by the colors representing their states; blue = precursor, orange = PC, gray = maintenance and green = PCi-differentiation.
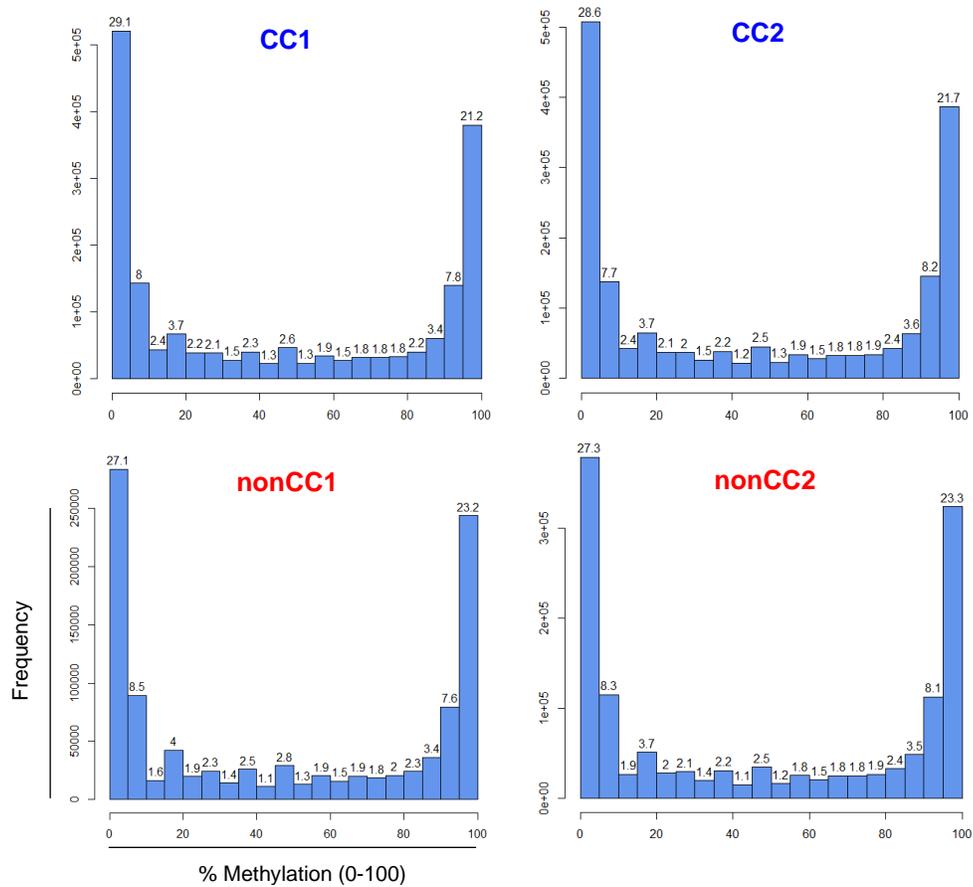
Figure 7.16 Histogram of methylation distribution of CC sample. Each figure displays histogram of methylation distribution. Vertical axis represents a frequency, while horizontal axis represents percentages of methylation (0, 20, 40, 60, 80 and 100). Name of each sample is written on the top by the colors representing their states; blue = CC samples, red = nonCC samples.
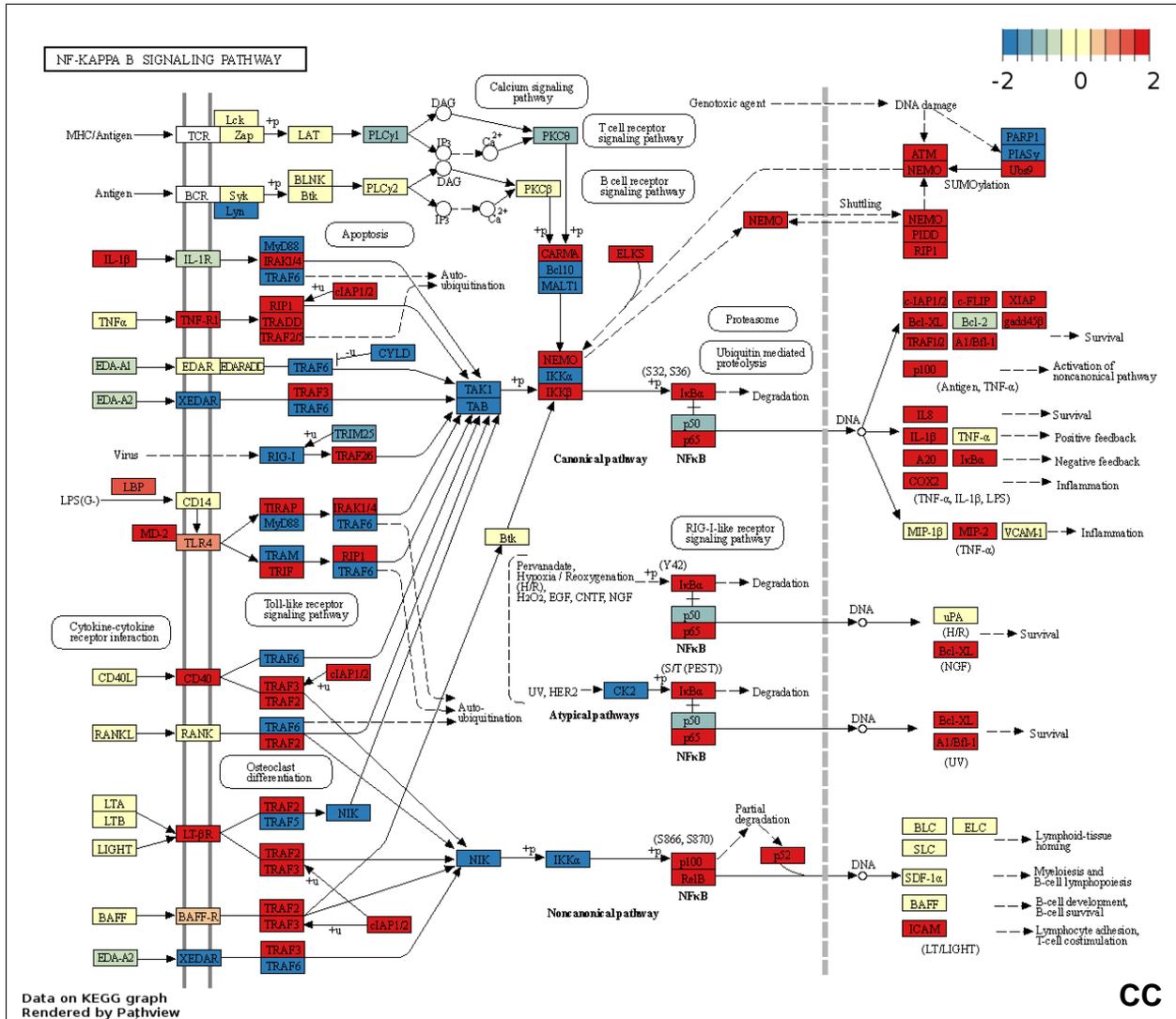
Figure 7.17 KEGG pathway of NF-kappa B (hsa04064). The figure shows the expression of genes in NF-kappa B signaling pathway during CC experiment. Gradient colors from blue, yellow and red represent log2 fold change of each gene from -2 to 2, respectively.
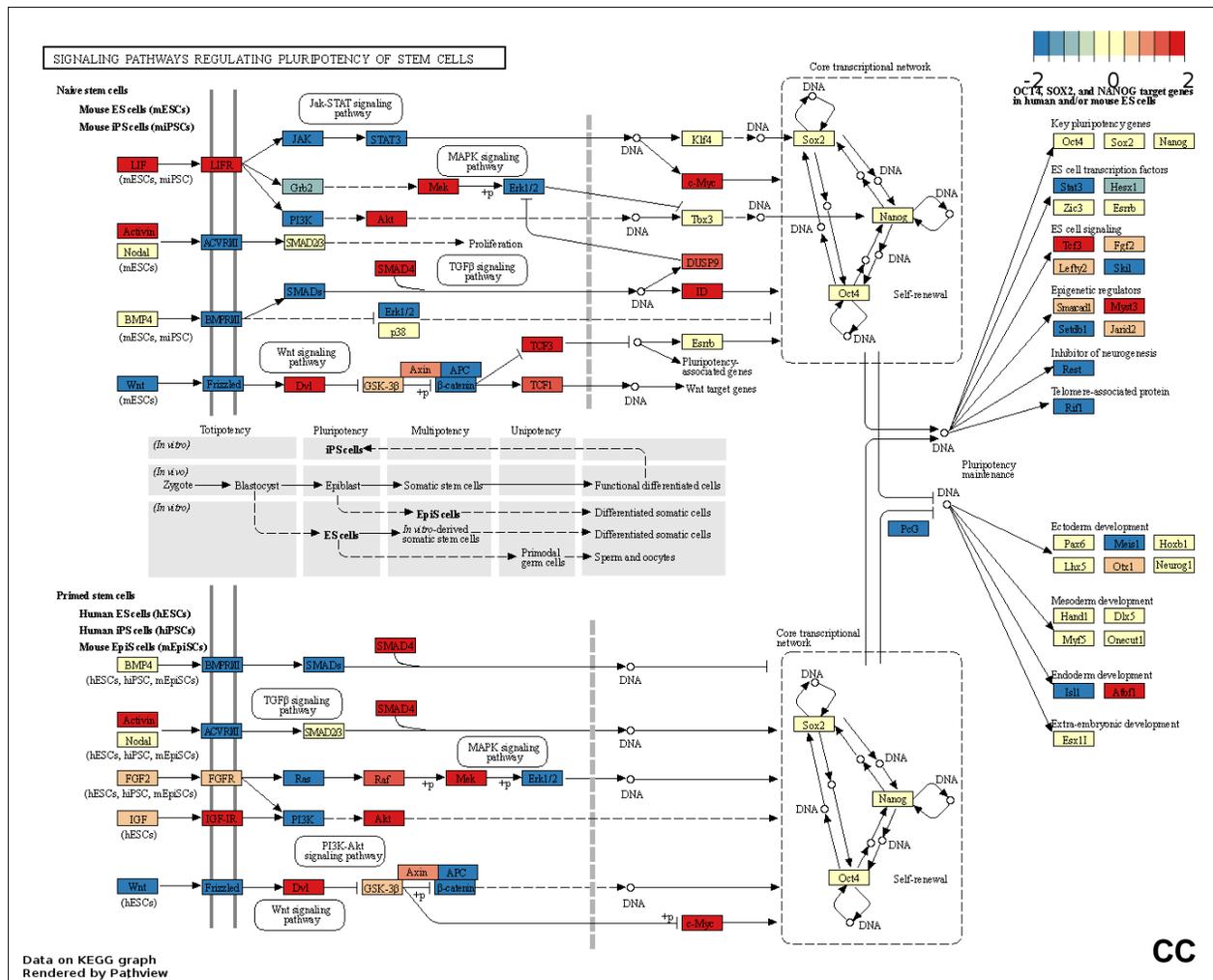
Figure 7.18 KEGG pathways of signaling pathways regulating pluripotency of stem cells (hsa04550). The figure shows the expression of genes in hsa04550 during CC experiment. Gradient colors from blue, yellow and red represent log2 fold change of each gene from -2 to 2, respectively.
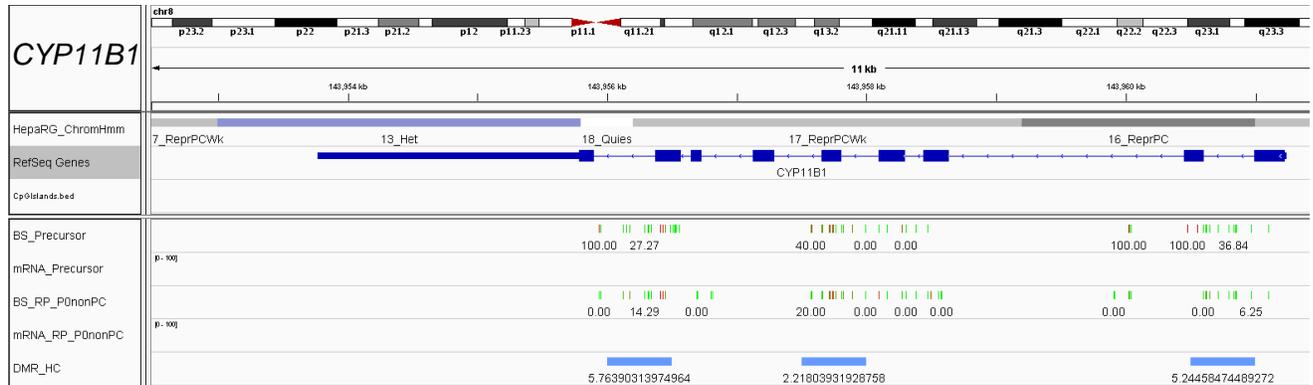
Figure 7.19 IGV snapshots of *CYP11B1*. The figure shows the RRBS data labeled as "BS_samples", following by expression data labeled as "mRNA_sample". The figure provides slightly hypermethylated DMRs of hydrocortisone removal sample (DMR_HC; blue bar).
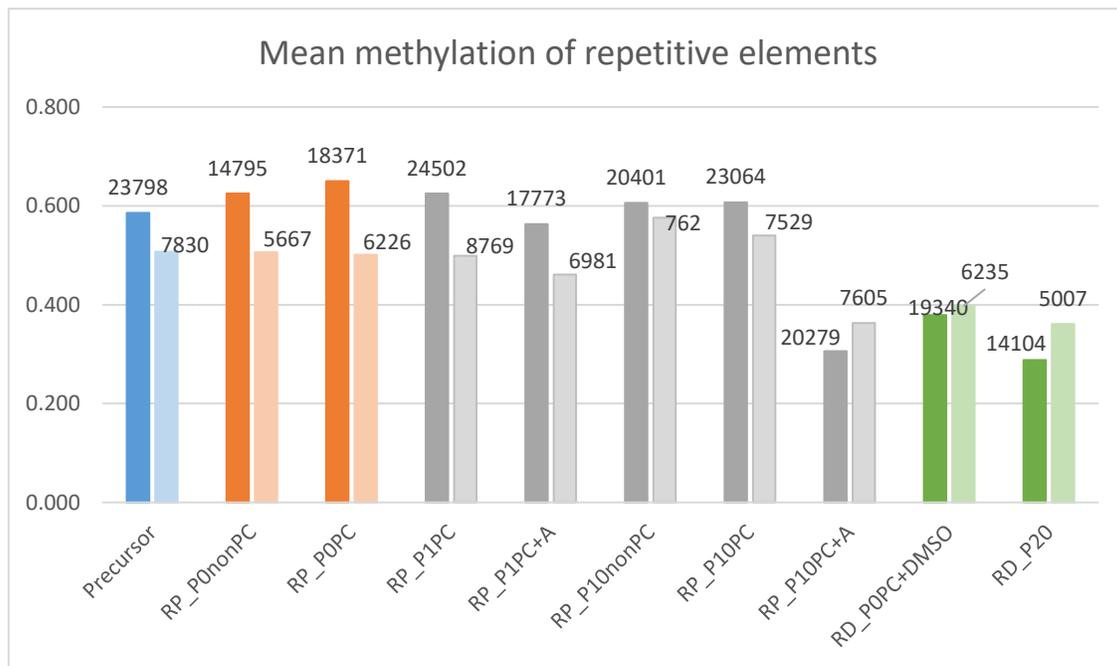
Figure 7.20 Mean methylation of repetitive elements. Bar graph shows mean methylation of HERVK (dark color bars) and LINE1 (light color bars) detected by local deep sequencing (Miseq). Coverage of each sample is displayed at the end of bar. Bar colors define states of each samples; blue = precursor, orange = PC, gray = maintenance and green = PCi-differentiation. Noted that RD_P0nonPC+A and RD_P0PC+hydro are excluded because of limited materials.
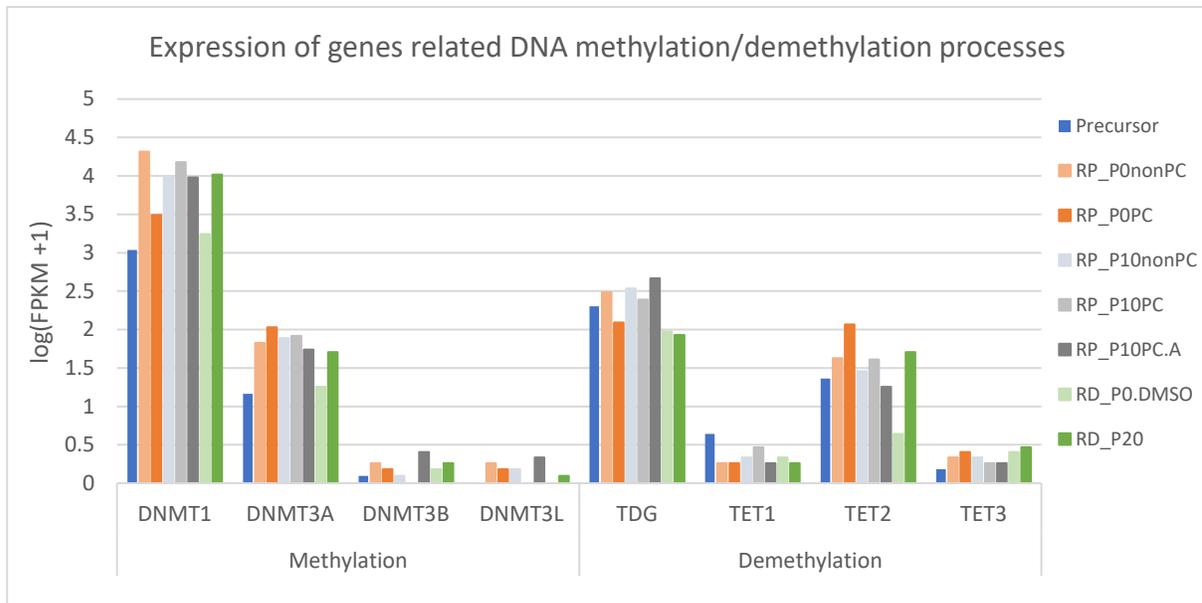
Figure 7.21 Expression of genes related to DNA methylation and demethylation processes. State of samples is defined by colors; blue = precursor, orange = PC, gray = maintenance and green = PCi-differentiation.
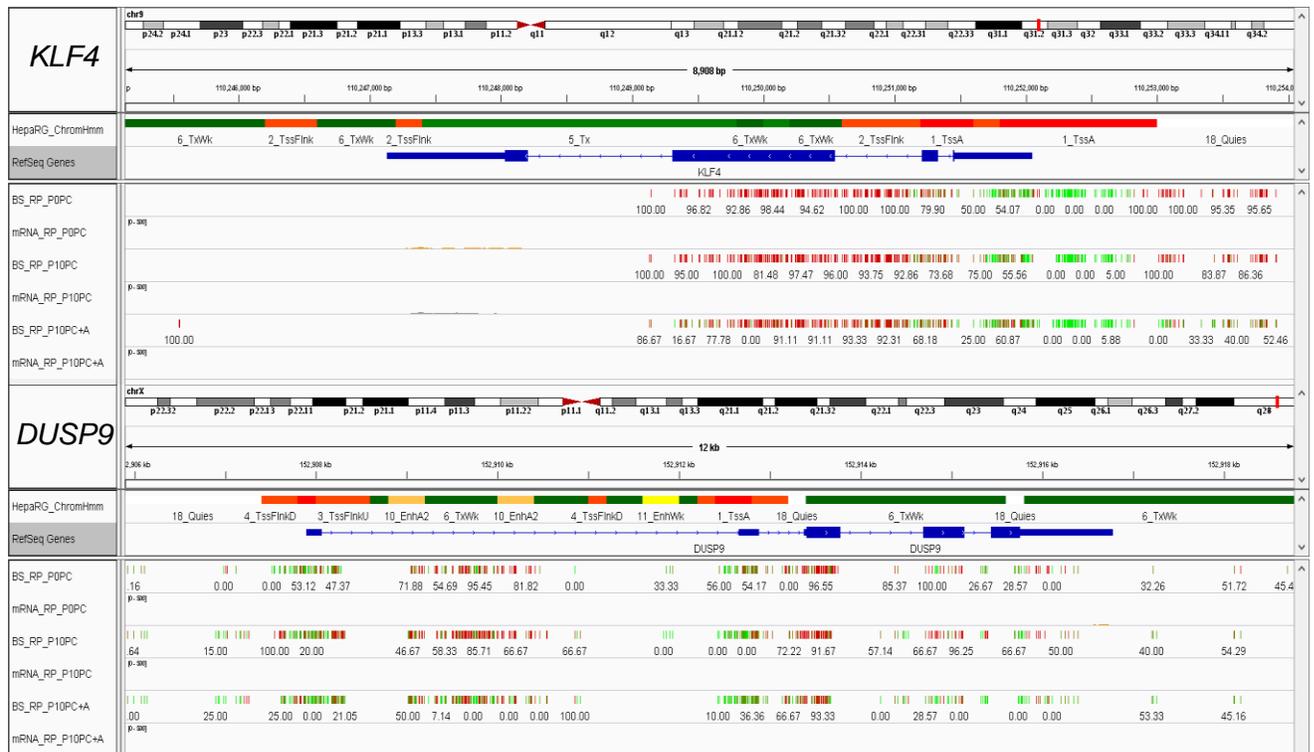
Figure 7.22 IGV snapshots of pluripotent related genes. The figure shows the RRBS data labeled as "BS_samples", following by expression data labeled as "mRNA_sample". Upper part is KLF4, whereas lower part is DUSP9. Samples include RP_P0PC, RP_P10PC and RP_P10PC+A, respectively. Colors from green to red represents the methylation level from low to high (0-100).

Figure 7.23 KEGG pathways of complement and coagulation cascades (hsa04610). Upper and lower parts show the expression of genes in hsa04610 during PC and Aza treatment. Gradient colors from blue, yellow and red represent log2 fold change of each gene from -2 to 2, respectively.
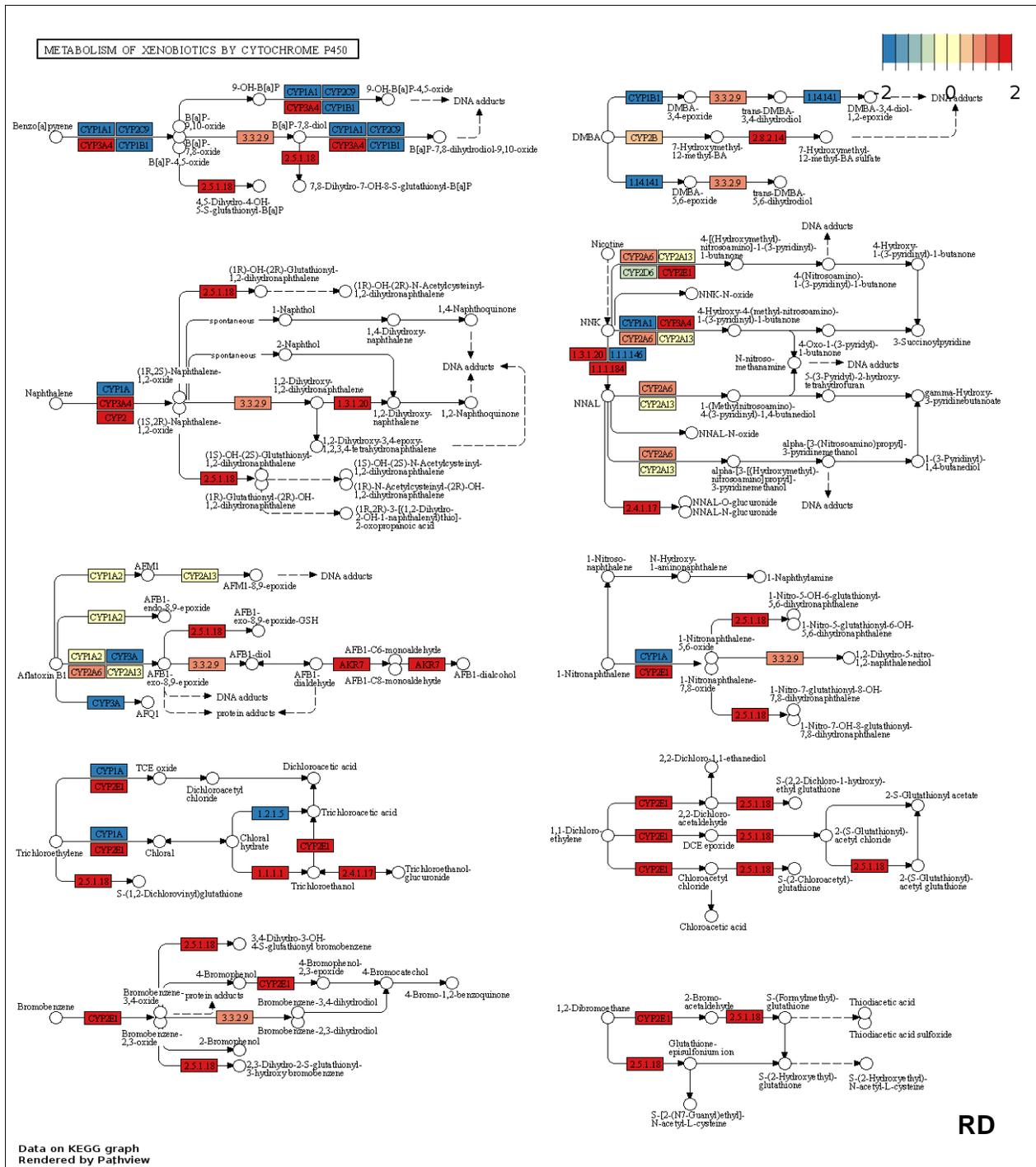
Figure 7.24 KEGG pathways of metabolism of xenobiotics by cytochrome P450 (hsa00980). Upper and lower figures show the expression of genes in hsa00980 during PCi-differentiation. Gradient colors from blue, yellow and red represent log2 fold change of each gene from -2 to 2, respectively.

7.2. Supplementary data 2

<u>Figure legends for technical validation by local deep sequencing</u>

Figure 7.25 - Figure 7.35 i) show methylation profiles of amplicons obtained from local deep sequencing (Miseq) and generated by BiQ Analyzer HT. Mean methylation (Mean met) and coverage (Cov.) of each sample are also provided. Sample without methylation profile is indicated as NA. ii) and iii) show scatter plot providing correlation of methylation between local deep sequencing (Miseq) and RRBS of each amplicon. In plot ii), blue dot represents all CpGs found in the amplicon. In plot iii), blue dot represents CpGs containing high RRBS coverage (≥10x), while orange dot represents CpGs containing low RRBS coverage (<10x). Equation, R^2 value and trendline (red dash) are calculated according to blue dot. Amplicon without CpGs containing low RRBS coverage will be marked by asterisk (*). Green line over the methylation profile indicates common CpGs found in RRBS and LDS (Miseq).

<u>List of amplicons</u>

- *AMIGO**
- *CSMD3**
- *DTD1**
- *DUS3L*
- *FA2H*
- *FAM5C*
- *OCRL*
- *SPARC*
- *TF*
- *TUBA1A*
- *ZNF814*

Figure 7.25 *AMIGO2* - i

Precursor

Mean met = 0.927

Cov. = 32448

RP_P0nonPC

Mean met = 0.922

Cov. = 14437

RP_P0PC

Mean met = 0.929

Cov. = 8002

RP_P1PC

Mean met = 0.911

Cov. = 6737
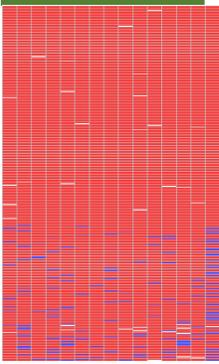
RP_P1PC+A

Mean met = 0.819

Cov. = 6345
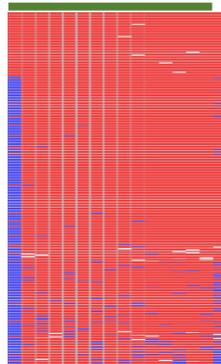
RP_P10nonPC

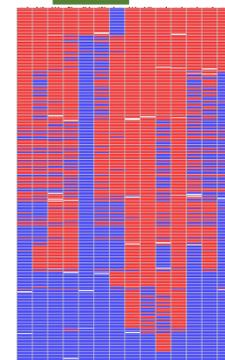Mean met = 0.915

Cov. = 21771
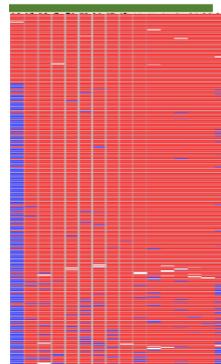
RP_P10PC

Mean met = 0.915

Cov. = 11953
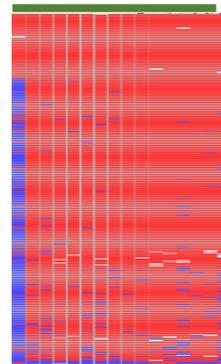
RP_P10PC+A

Mean met = 0.201

Cov. = 5548

RD_P0nonPC+A

Mean met = 0.49

Cov. = 14126

RD_P0PC+hydro

Mean met = 0.457

Cov. = 5509

RD_P0PC+DMSO

Mean met = 0.496

Cov. = 31370
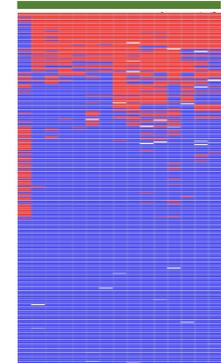
RD_P20

Mean met = 0.496

Cov. = 31955

AMIGO2 - ii

y = 0.7207x + 23.279
R² = 0.6035
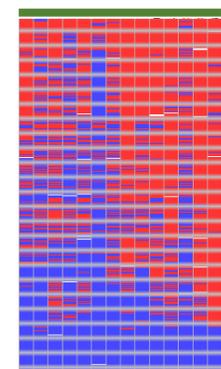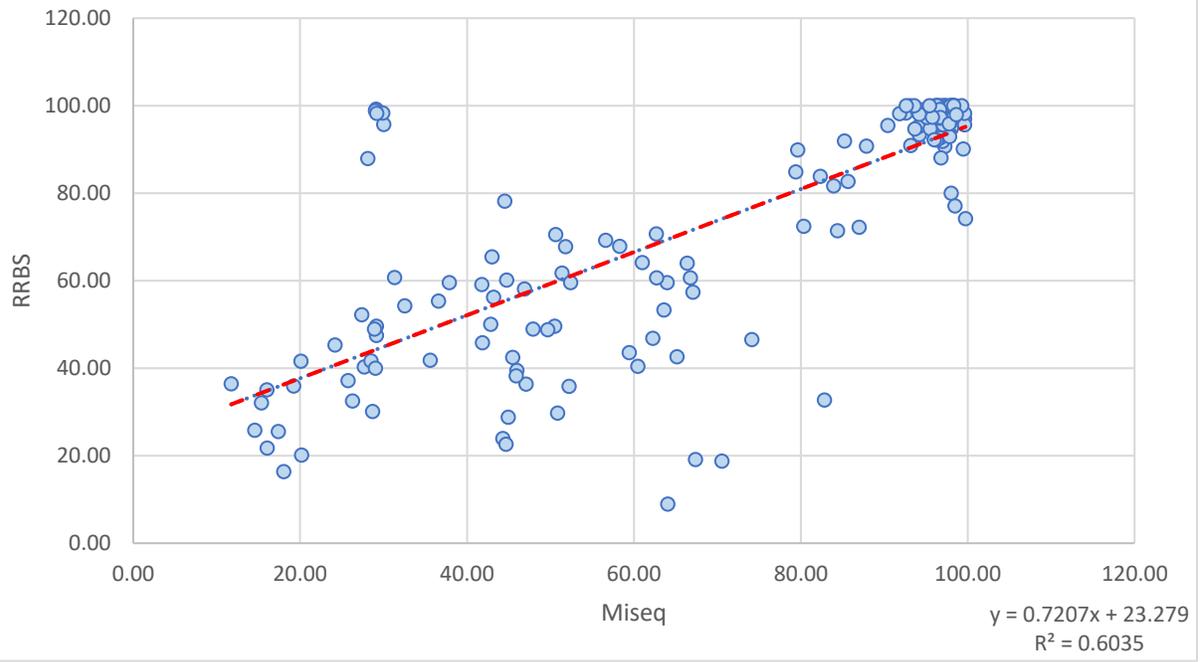
Figure 7.26 *CSMD3* - i

Precursor
Mean met = 0.226
Cov. = 21899
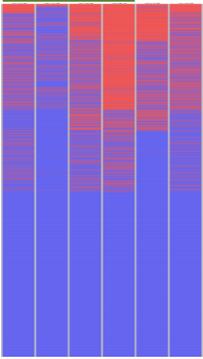
RP_P0nonPC
Mean met = 0.265
Cov. = 23688
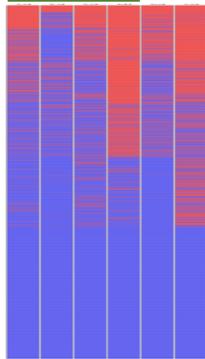
RP_P0PC
Mean met = 0.284
Cov. = 16875
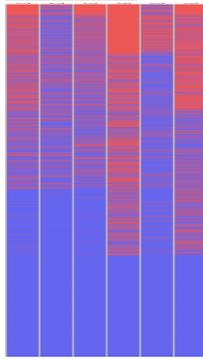
RP_P1PC
Mean met = 0.222
Cov. = 15975
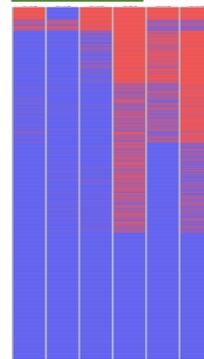
RP_P1PC+A
Mean met = 0.255
Cov. = 24047
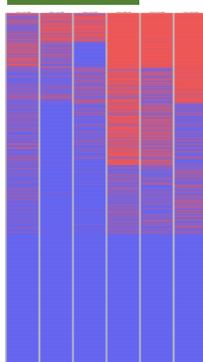
RP_P10nonPC
Mean met = 0.462
Cov. = 10078
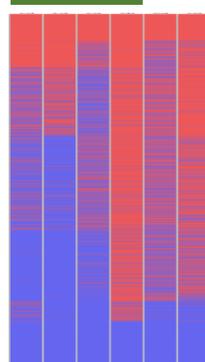
RP_P10PC
Mean met = 0.302
Cov. = 5842
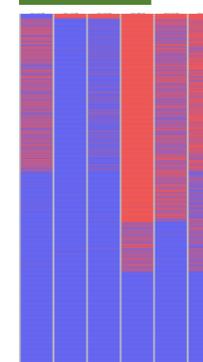
RP_P10PC+A
Mean met = 0.104
Cov. = 1726
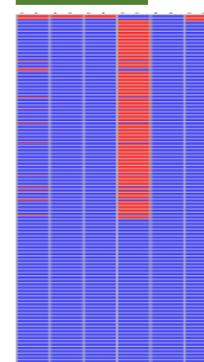
RD_P0nonPC+A
Mean met = 0.245
Cov. =23774
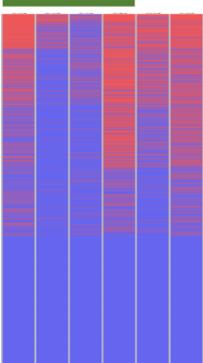
RD_P0PC+hydro
Mean met = 0.296
Cov. = 18422
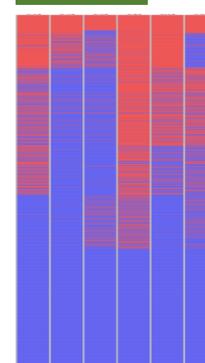
RD_P0PC+DMSO
Mean met = 0.324
Cov. = 10787
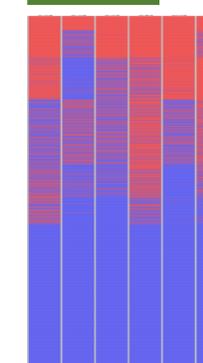
RD_P20
Mean met = 0.355
Cov. = 7417

CSMD3 - ii

y = 0.8039x + 7.0577
R² = 0.3751
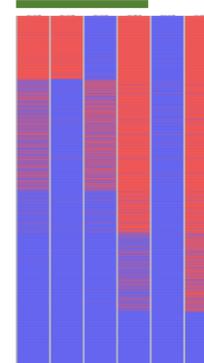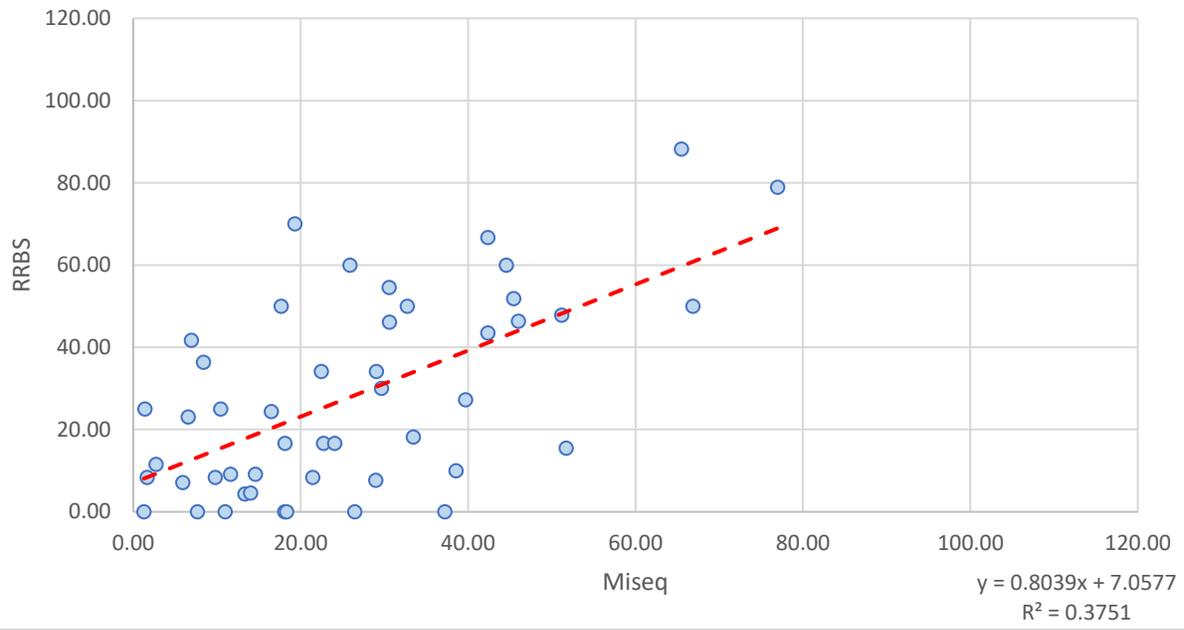
Figure 7.27 *DTD1* - i

Precursor
Mean met = 0.445
Cov. = 29866

RP_P0nonPC
Mean met = 0.434
Cov. = 33326

RP_P0PC
Mean met = 0.423
Cov. = 25602

RP_P1PC
Mean met = 0.398
Cov. = 13690
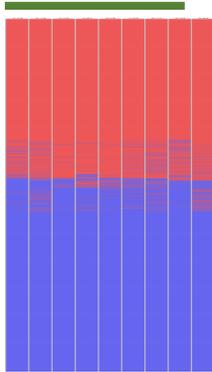
RP_P1PC+A
Mean met = 0.341
Cov. = 27824

RP_P10nonPC
Mean met = 0.423
Cov. = 3367

RP_P10PC
Mean met = 0.352
Cov. = 13349

RP_P10PC+A
Mean met = 0.268
Cov. = 4990
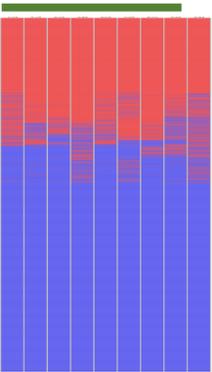
RD_P0nonPC+A
Mean met = 25357
Cov. = 0.383
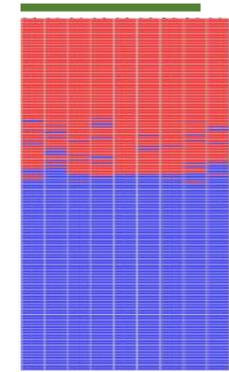
RD_P0PC+hydro
Mean met = 14098
Cov. = 0.312
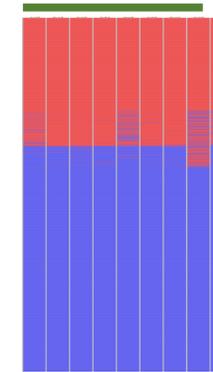
RD_P0PC+DMSO
Mean met = 0.393
Cov. = 17007
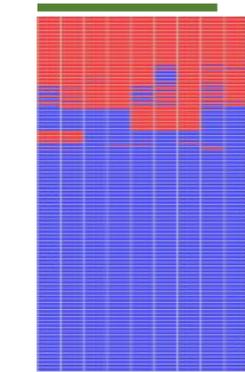
RD_P20
Mean met = 0.285
Cov. = 1479

DTD1 - ii

y = 0.6758x + 29.442
R² = 0.1419
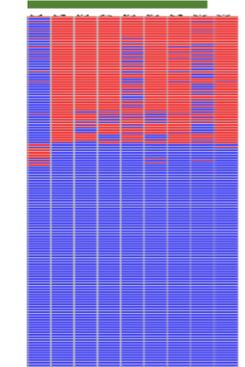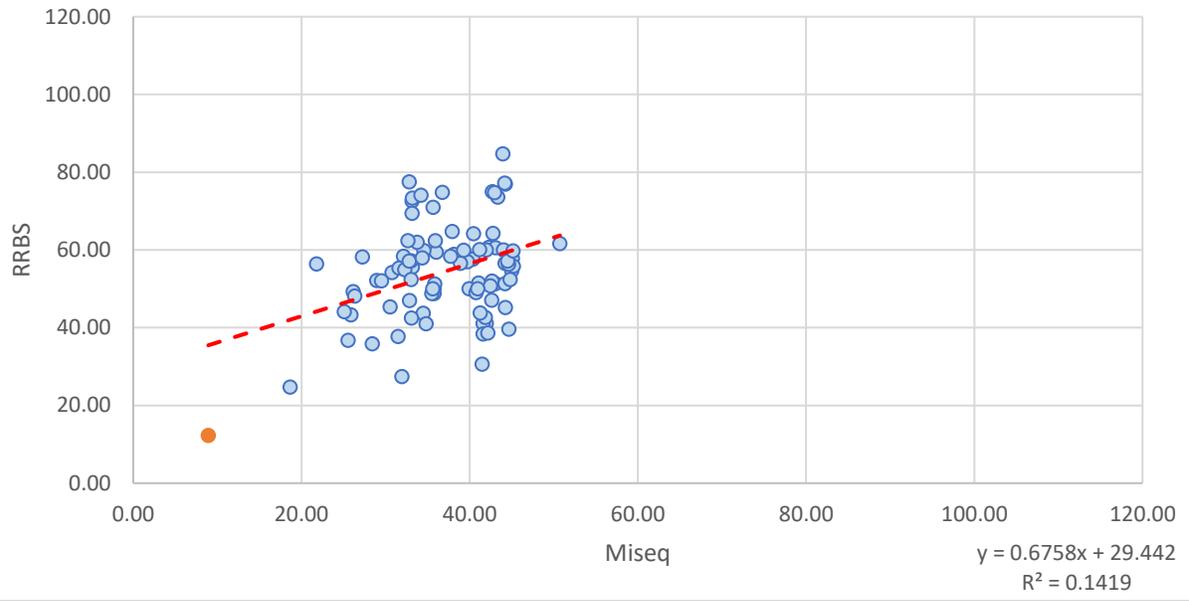
Figure 7.28 *DUS3L* - i

Precursor
Mean met = 0.918
Cov. = 17823

RP_P0nonPC
Mean met = 0.887
Cov. = 19808

RP_P0PC
Mean met = 0.844
Cov. = 22101

RP_P1PC
Mean met = 0.873
Cov. = 22673
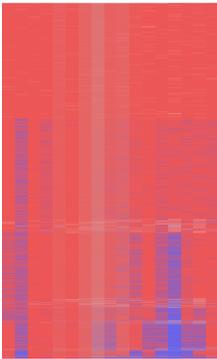
RP_P1PC+A
Mean met = 0.754
Cov. = 19060

RP_P10nonPC
Mean met = 0.939
Cov. = 6785

RP_P10PC
Mean met = 0.926
Cov. = 6065

RP_P10PC+A
Mean met = 0.491
Cov. = 14438
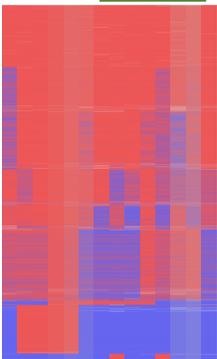
RD_P0nonPC+A
Mean met = 0.388
Cov. = 14264
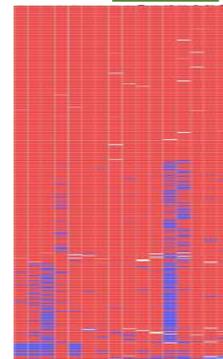
RD_P0PC+hydro
Mean met = 0.439
Cov. = 19296
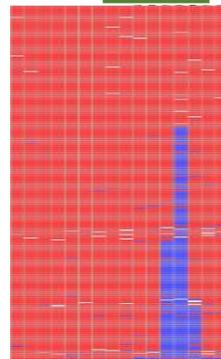
RD_P0PC+DMSO
Mean met = 0.364
Cov. = 11548
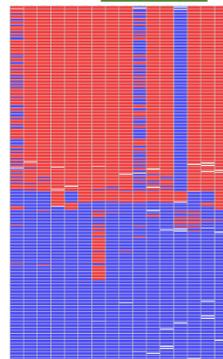
RD_P20
Mean met = 0.149
Cov. = 13162

DUS3L - ii

y = 0.8612x + 10.393
R² = 0.8151

DUS3L - iii

y = 0.8198x + 13.739
R² = 0.7714
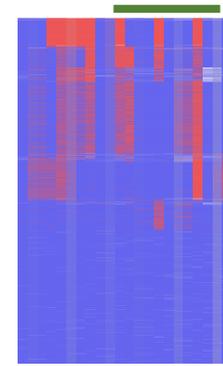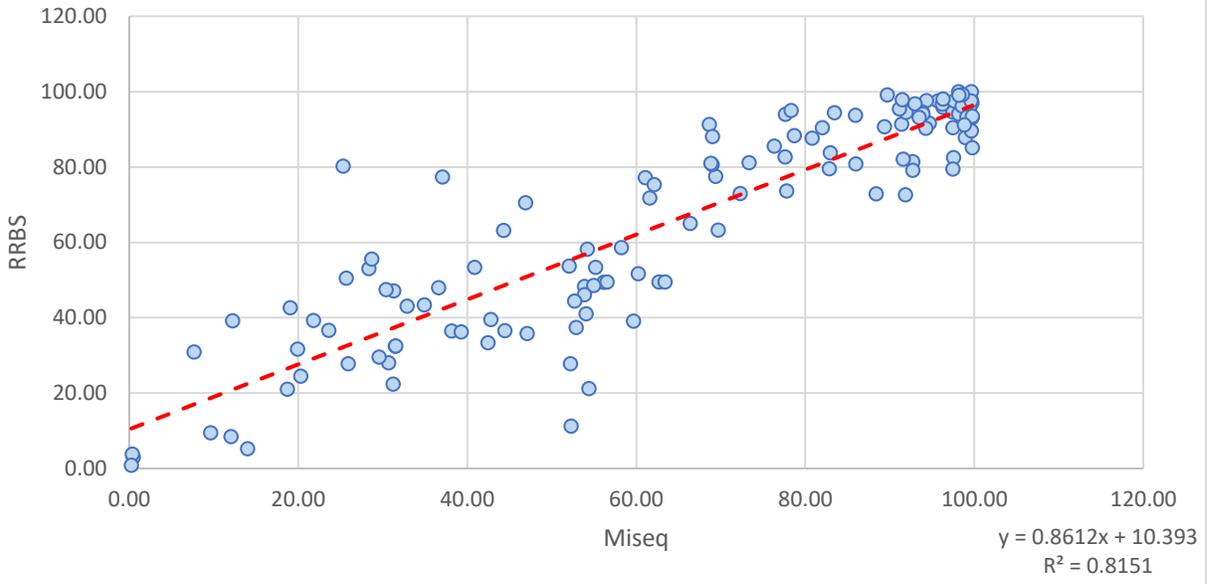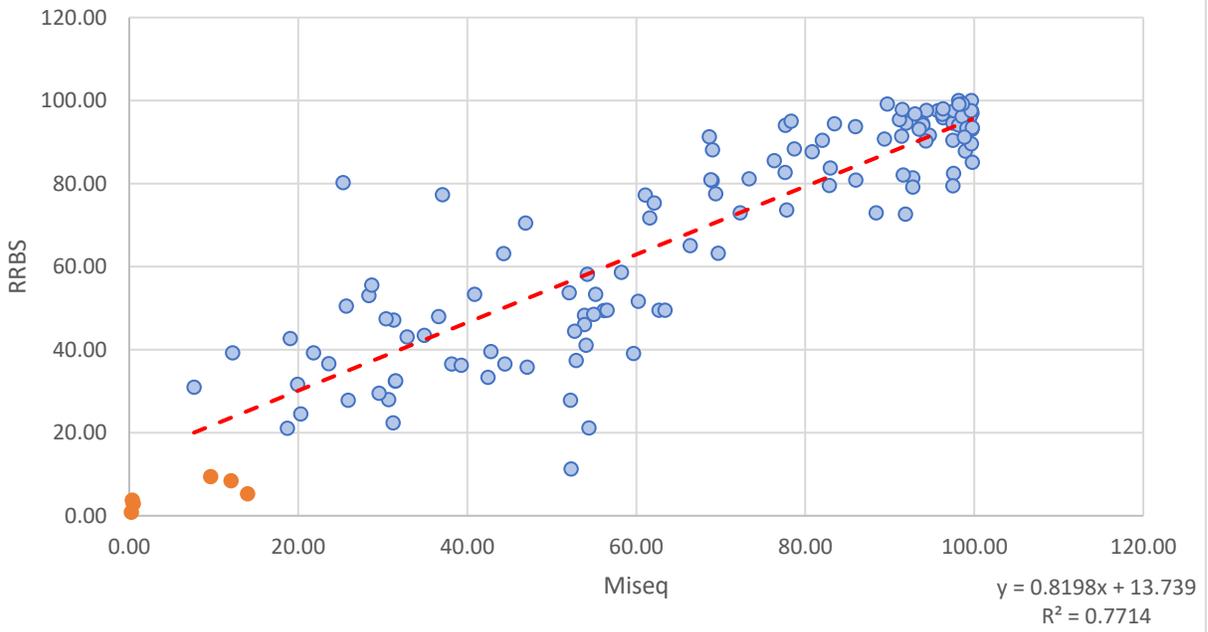
Figure 7.29 *FA2H* - i

Precursor	RP_P0nonPC	RP_P0PC	RP_P1PC
Mean met = 0.564	Mean met = 0.598	Mean met = 0.611	Mean met = 0.566
Cov. = 1002	Cov. = 2320	Cov. = 1566	Cov. = 813
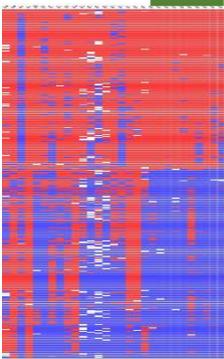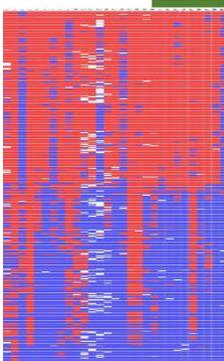
RP_P1PC+A	RP_P10nonPC	RP_P10PC	RP_P10PC+A
Mean met = 0.553	Mean met = 0.688	Mean met = 224	Mean met = 0.288
Cov. = 1351	Cov. =1645	Cov. = 0.456	Cov. = 248

RD_P0nonPC+A	RD_P0PC+hydro	RD_P0PC+DMSO	RD_P20
Mean met = 0.815	Mean met = 0.426	Mean met = 1060	Mean met = 0.327
Cov. = 1	Cov. = 1129	Cov. = 0.393	Cov. = 2050

NA

FA2H - ii

y = 0.5985x + 25.663
R² = 0.352

FA2H - iii

y = 0.5474x + 30.297
R² = 0.3608
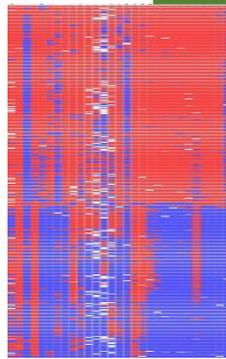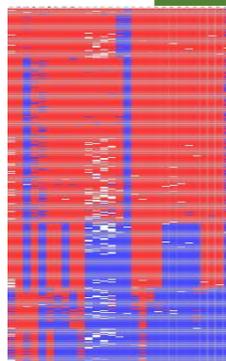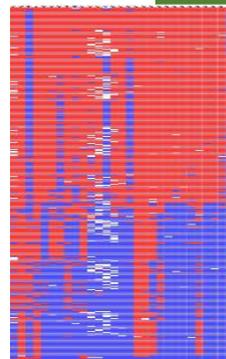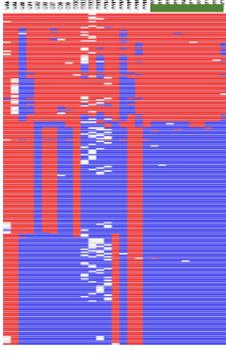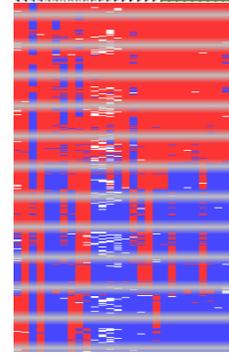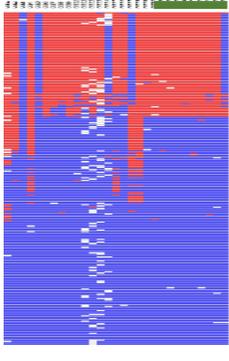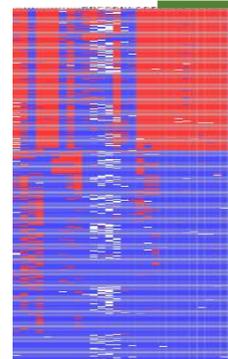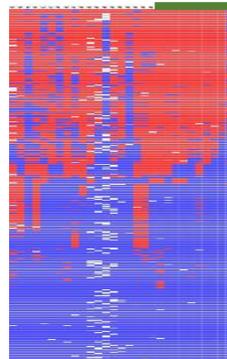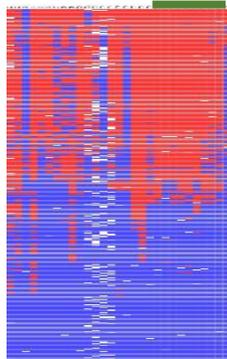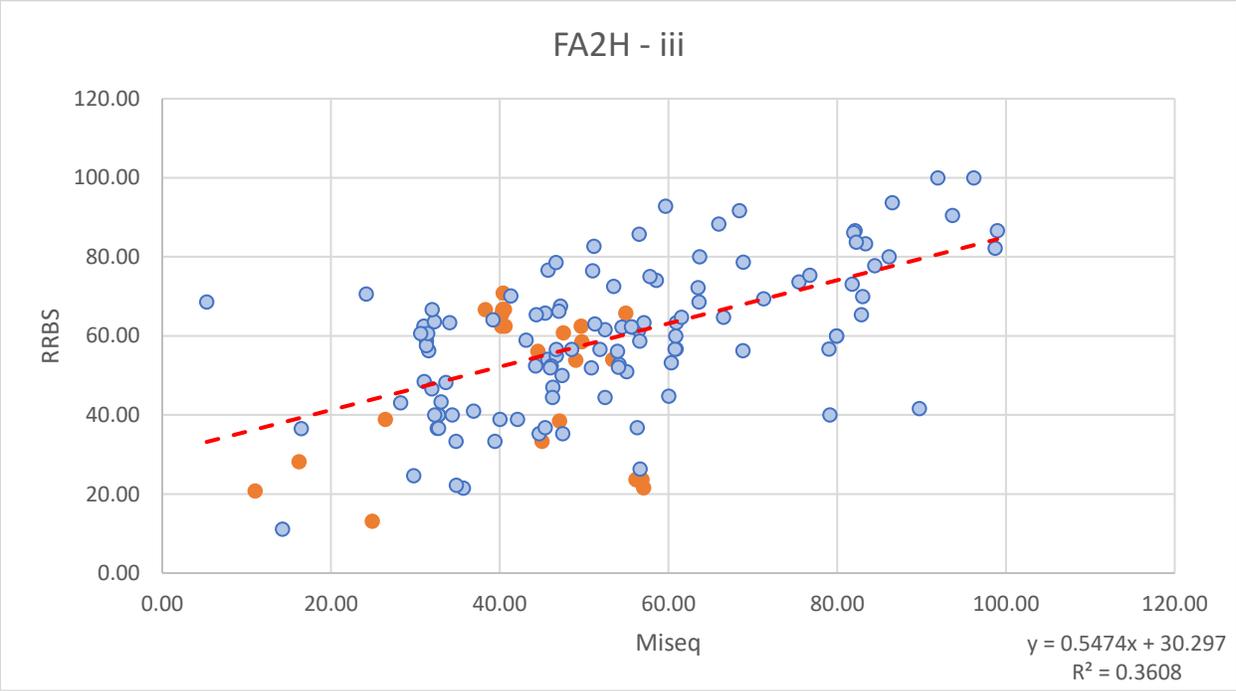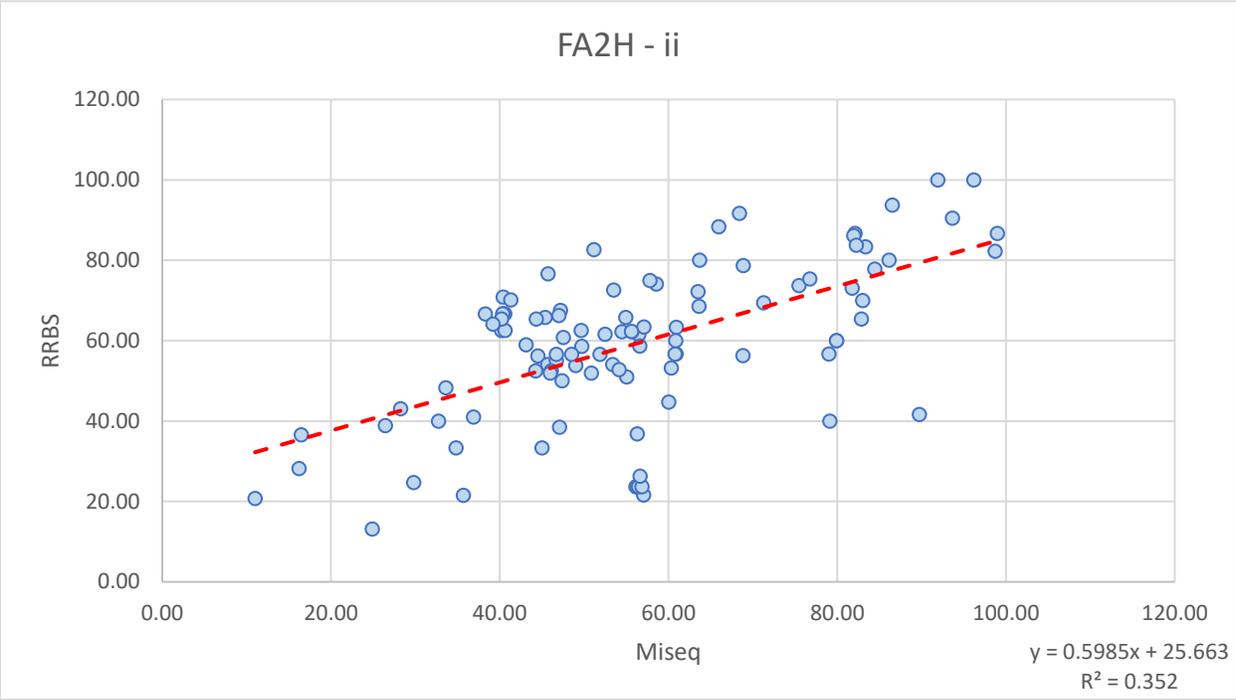
Figure 7.30 *FAM5C* - i

Precursor	RP_P0nonPC	RP_P0PC	RP_P1PC
Mean met = 0.267	Mean met = 0.338	Mean met = 0.340	Mean met = 0.260
Cov. = 18330	Cov. = 8767	Cov. = 6178	Cov. = 593
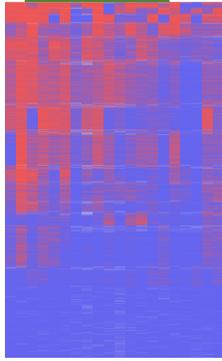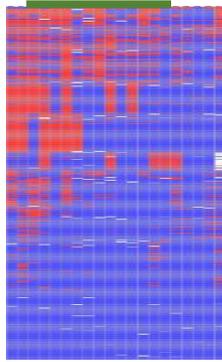
RP_P1PC+A	RP_P10nonPC	RP_P10PC	RP_P10PC+A
Mean met = 0.216	Mean met = 0.172	Mean met = 0.231	Mean met = 0.142
Cov. = 7470	Cov. = 3566	Cov. = 3491	Cov. = 1806

RD_P0nonPC+A	RD_P0PC+hydro	RD_P0PC+DMSO	RD_P20
Mean met = 0.141	Mean met = 0.170	Mean met = 0.104	Mean met = 0.08
Cov. = 12030	Cov. = 399	Cov. = 4004	Cov. = 3448

FAM5C - ii

y = 0.7626x + 3.3912
R² = 0.4329

FAM5C - iii

y = 0.7901x + 4.4666
R² = 0.4159
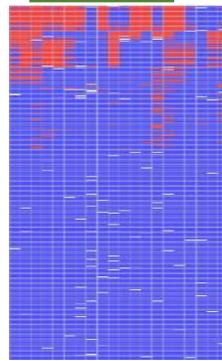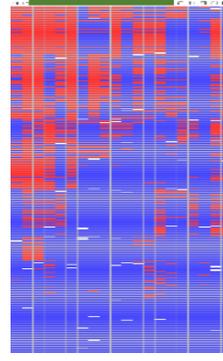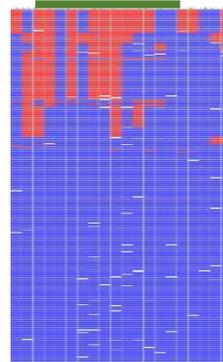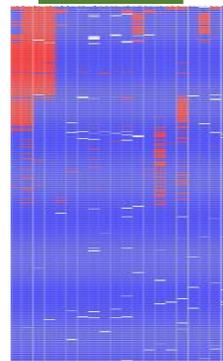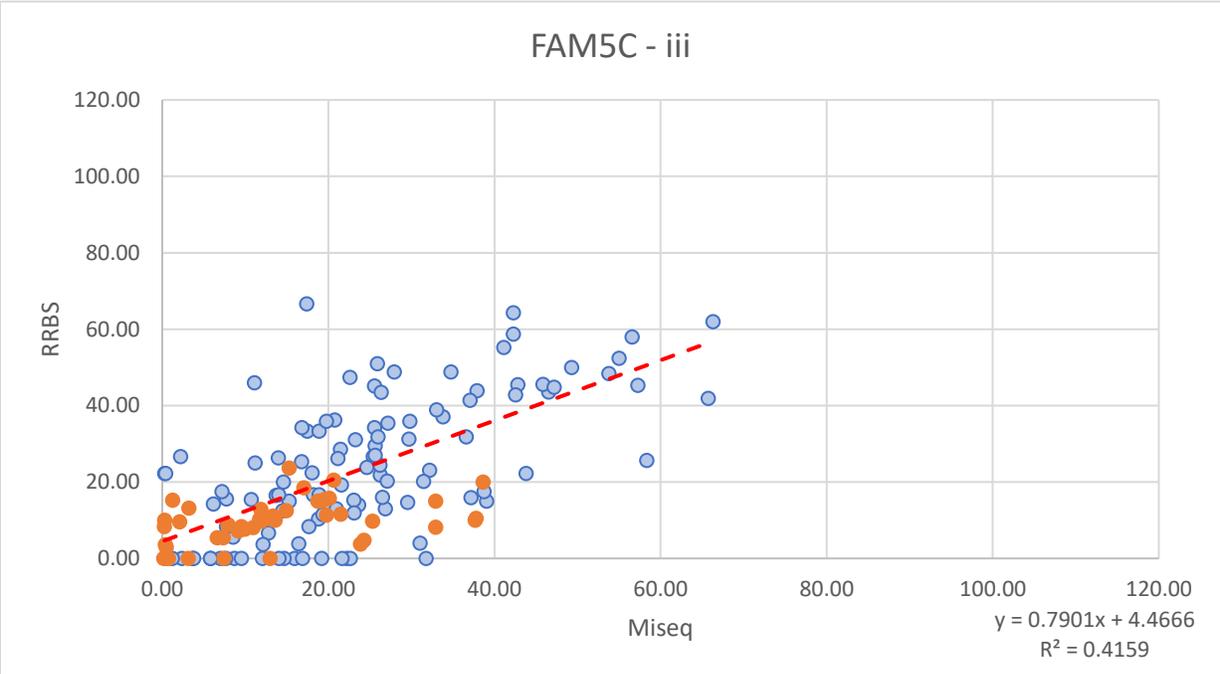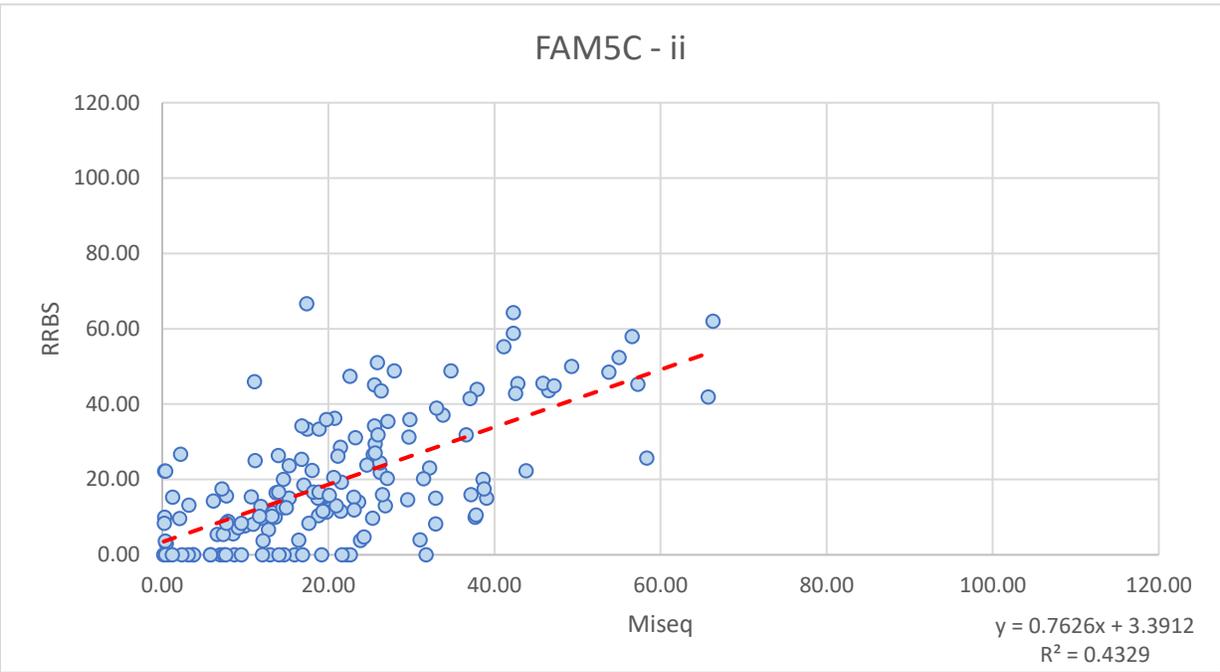
Figure 7.31 *OCRL* - i

Precursor
Mean met = 0.343
Cov. = 12597

RP_P0nonPC
Mean met = 0.293
Cov. = 18911

RP_P0PC
Mean met = 0.299
Cov. = 16304

RP_P1PC
Mean met = 0.361
Cov. = 11056
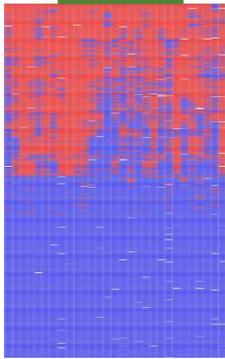
RP_P1PC+A
Mean met = 0.323
Cov. = 1276
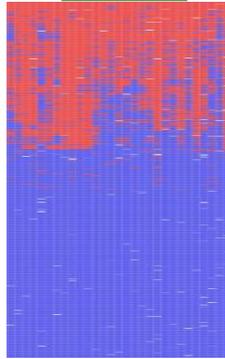
RP_P10nonPC
Mean met = 0.341
Cov. = 17557
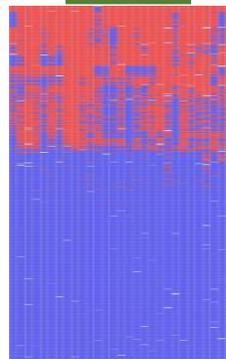
RP_P10PC
Mean met = 0.149
Cov. = 8218
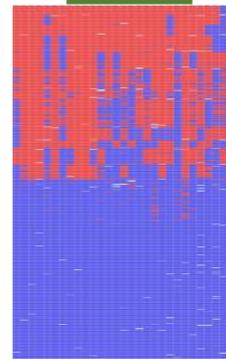
RP_P10PC+A
Mean met = 0.216
Cov. = 11867

RD_P0nonPC+A
Mean met = 0.212
Cov. =18539

RD_P0PC+hydro
Mean met = 0.244
Cov. = 3261

RD_P0PC+DMSO
Mean met = 0.197
Cov. = 15789
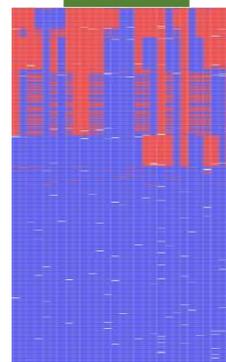
RD_P20
Mean met = 0.159
Cov. = 14006

OCRL - ii

y = 0.707x + 14.068
R² = 0.2405

OCRL - iii

y = 0.7646x + 16.724
R² = 0.3453
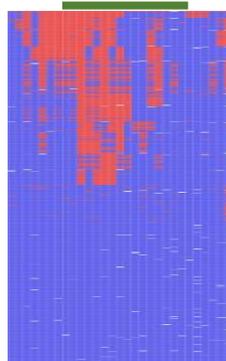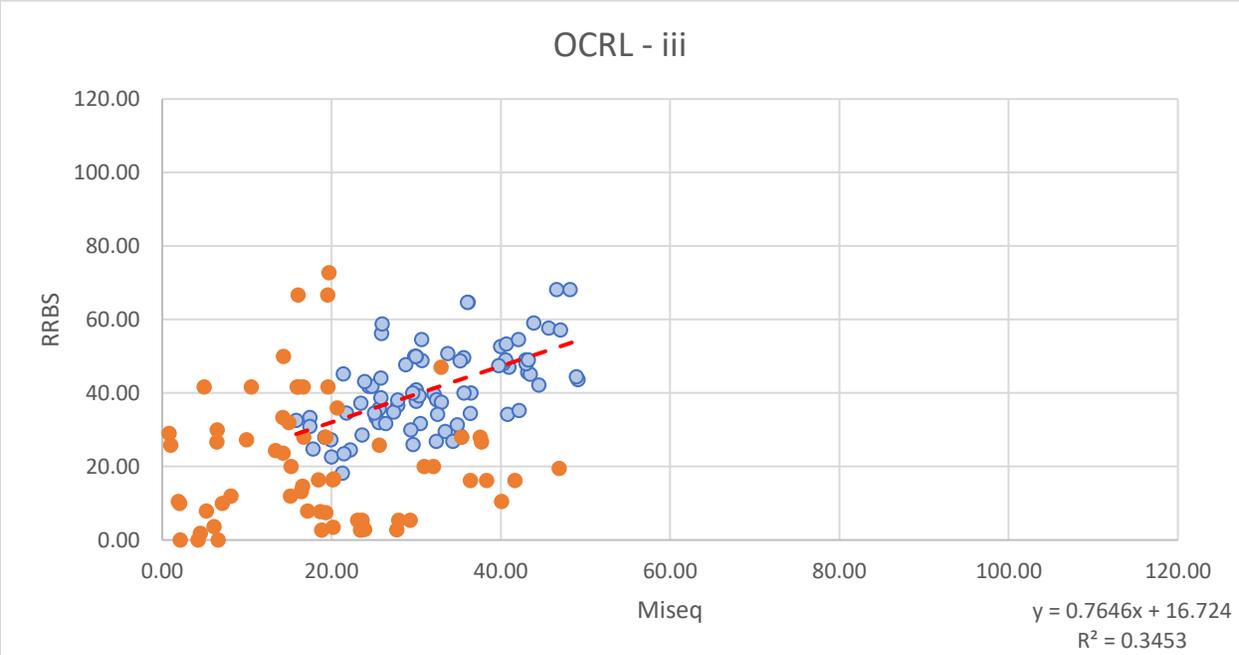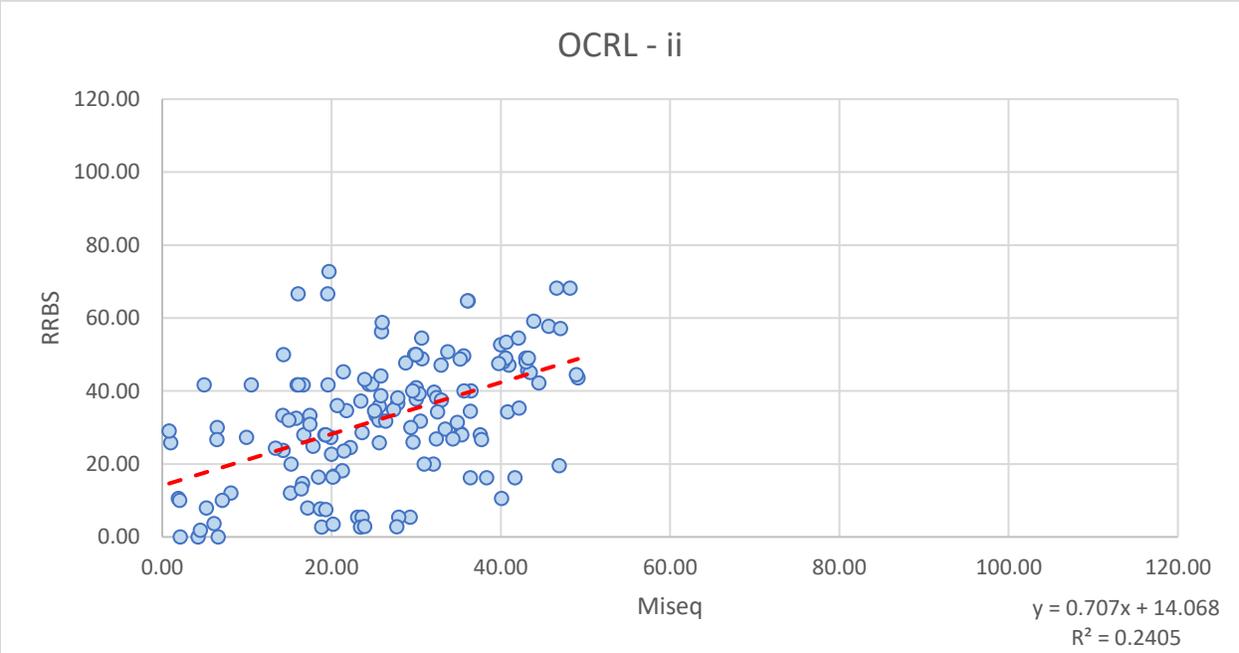
Figure 7.32 *SPARC* - i

Precursor	RP_P0nonPC	RP_P0PC	RP_P1PC
Mean met = 0.496	Mean met = 0.483	Mean met = 0.54	Mean met = 0.503
Cov. = 2002	Cov. = 13318	Cov. = 16102	Cov. = 6621
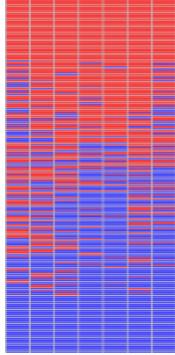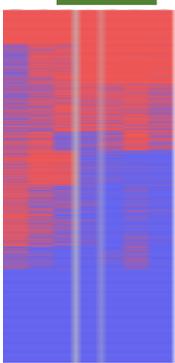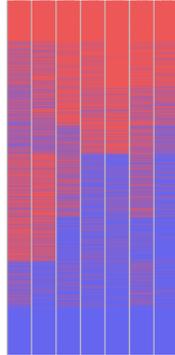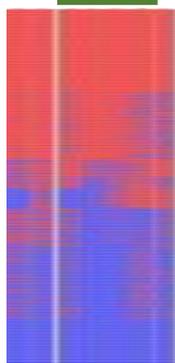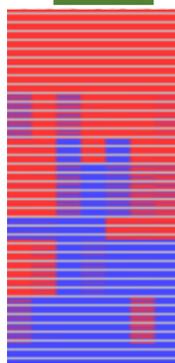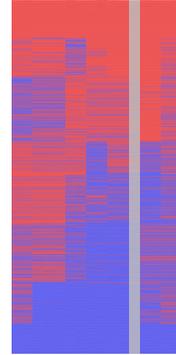
RP_P1PC+A	RP_P10nonPC	RP_P10PC	RP_P10PC+A
Mean met = 0.420	Mean met = 0.529	Mean met = 0.572	Mean met = 0.315
Cov. = 10947	Cov. = 9622	Cov. = 9685	Cov. = 8013

RD_P0nonPC+A	RD_P0PC+hydro	RD_P0PC+DMSO	RD_P20
Mean met = 0.086	Mean met = 0.064	Mean met = 0.068	Mean met = 0.350
Cov. = 11083	Cov. = 16011	Cov. = 14548	Cov. = 32829
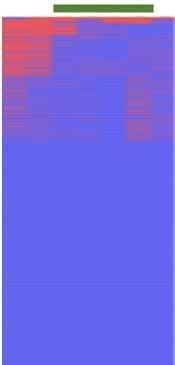
SPARC - ii

y = 0.6134x + 19.075
R² = 0.2954

SPARC - iii

y = -0.1211x + 56.039
R² = 0.0195
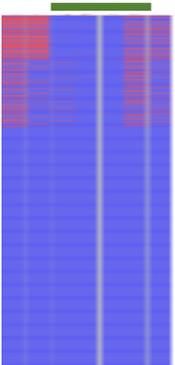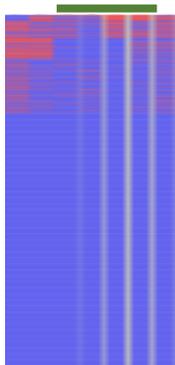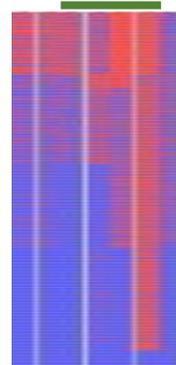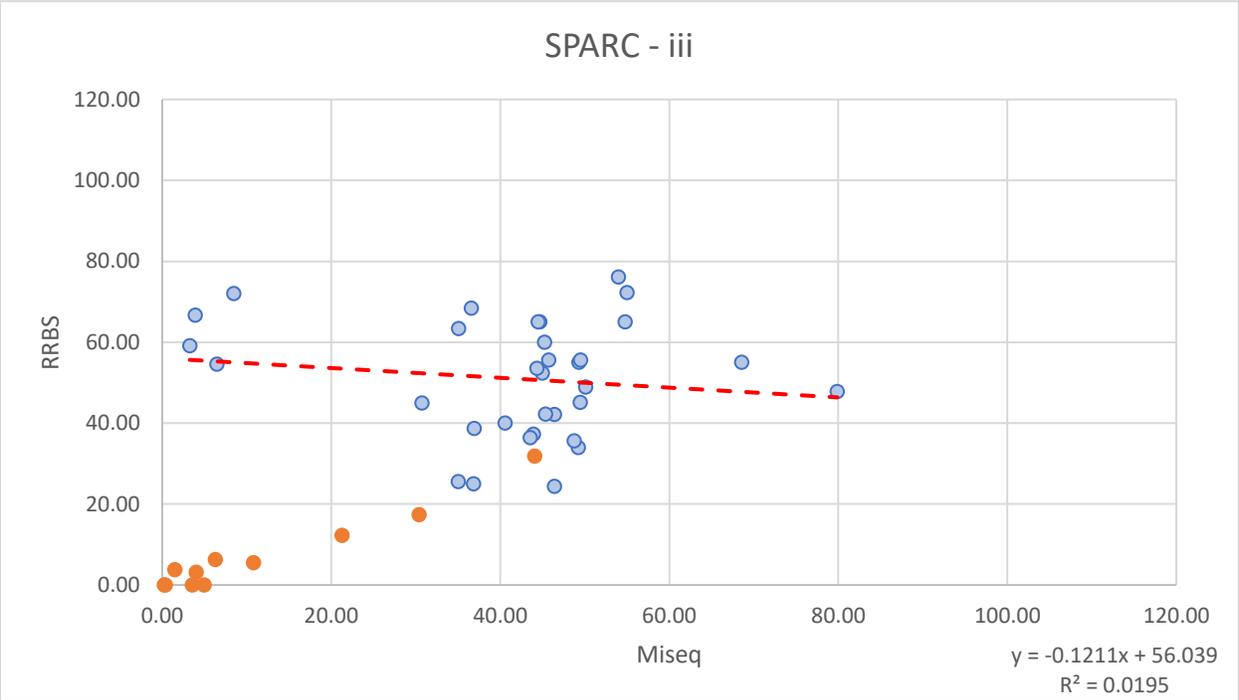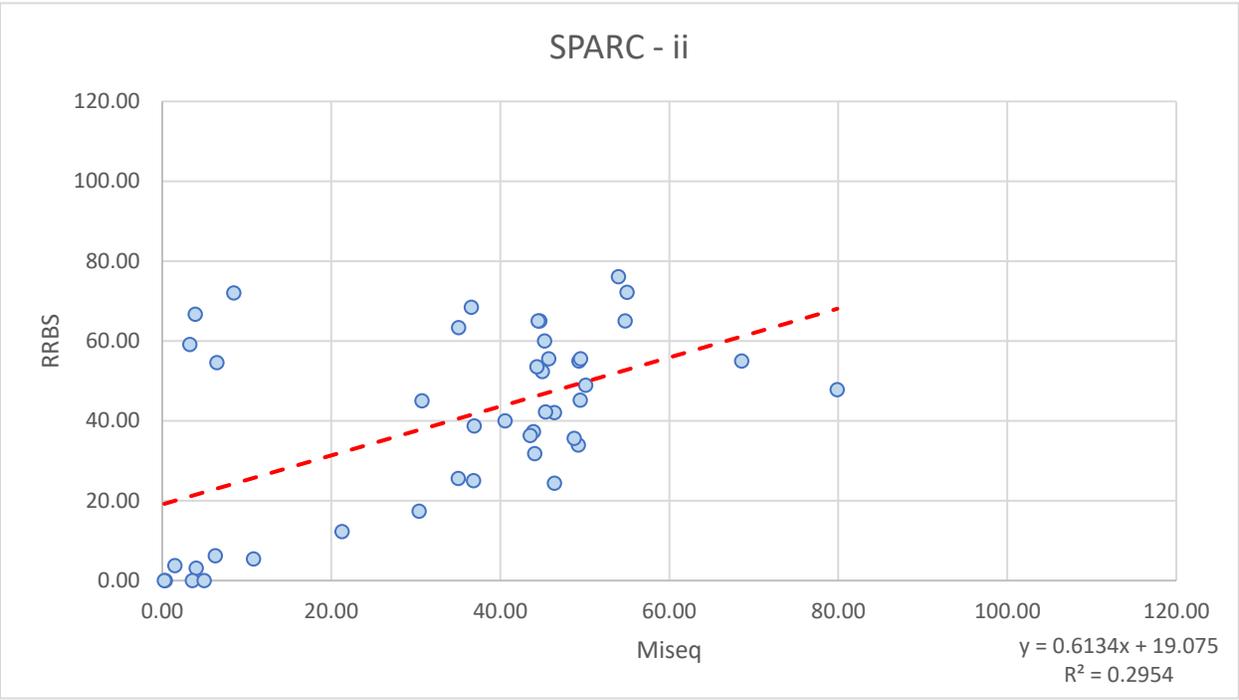
Figure 7.33 *TF* - i

Precursor
Mean met = 0.488
Cov. = 70795

RP_P0nonPC
Mean met = 0.549
Cov. = 63265

RP_P0PC
Mean met = 0.476
Cov. = 90011

RP_P1PC
Mean met = 0.507
Cov. = 42227
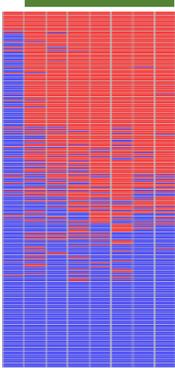
RP_P1PC+A
Mean met = 0.405
Cov. = 16512
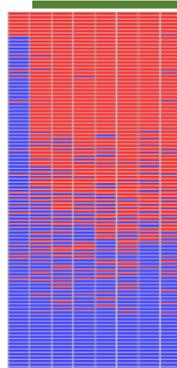
RP_P10nonPC
Mean met = 0.436
Cov. = 8389
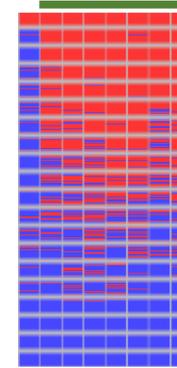
RP_P10PC
Mean met = 0.546
Cov. = 62694
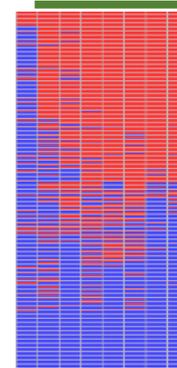
RP_P10PC+A
Mean met = 0.297
Cov. = 41328

RD_P0nonPC+A
Mean met = 0.356
Cov. = 64777

RD_P0PC+hydro
Mean met = 0.373
Cov. = 34034

RD_P0PC+DMSO
Mean met = 0.388
Cov. = 72151

RD_P20
Mean met = 0.341
Cov. = 8035

TF - ii

y = 0.9799x + 1.0972
R² = 0.2531

TF - iii

y = 0.1761x + 51.898
R² = 0.0101
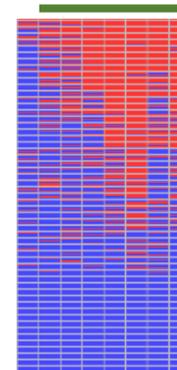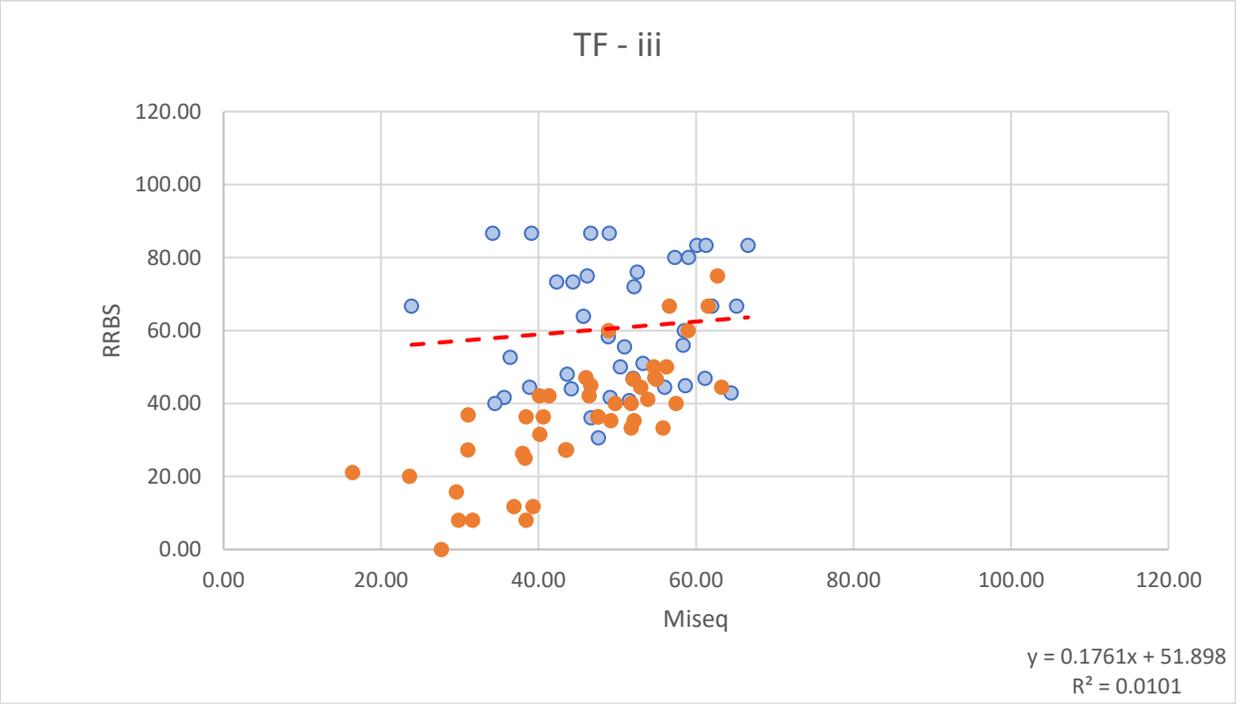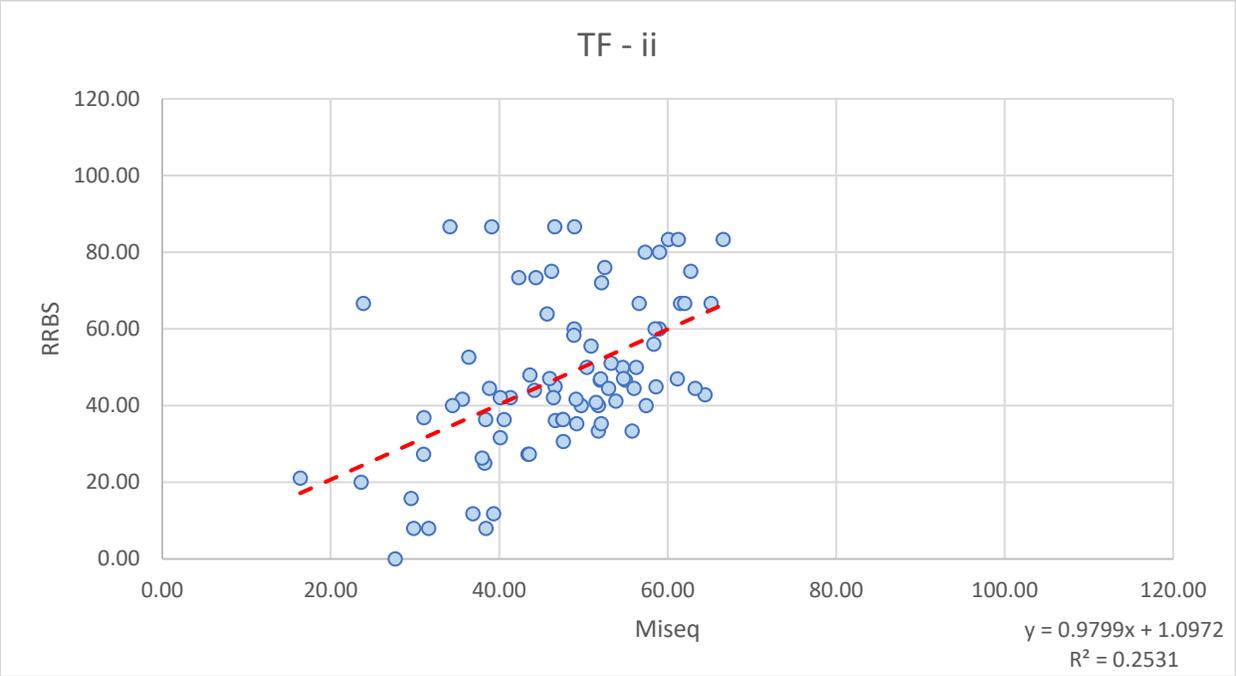
Figure 7.34 *TUBA1A* - i

Precursor	RP_P0nonPC	RP_P0PC	RP_P1PC
Mean met = 0.409	Mean met = 0.107	Mean met = 0.373	Mean met = 0.619
Cov. = 3592	Cov. = 3358	Cov. = 23162	Cov. = 1114
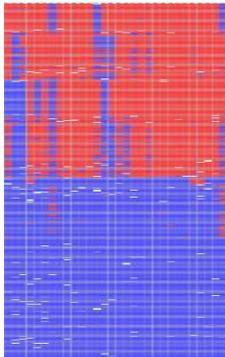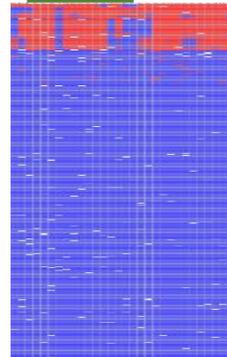
RP_P1PC+A	RP_P10nonPC	RP_P10PC	RP_P10PC+A
Mean met = 0.416	Mean met = 0.472	Mean met = 0.559	Mean met = 0.756
Cov. = 10475	Cov. = 15055	Cov. = 4275	Cov. = 7241

RD_P0nonPC+A	RD_P0PC+hydro	RD_P0PC+DMSO	RD_P20
Mean met = 0.406	Mean met = 0.313	Mean met = 0.296	Mean met = 0.342
Cov. = 3266	Cov. = 1927	Cov. = 3299	Cov. =2657

TUBA1A - ii

y = 0.1699x + 38.117
R² = 0.0435

TUBA1A - iii

y = 0.032x + 46.933
R² = 0.002
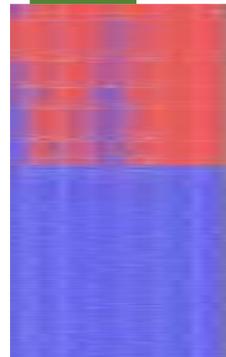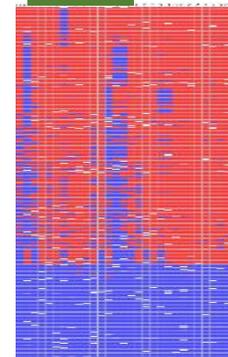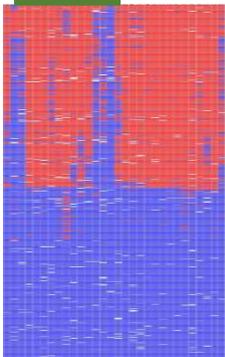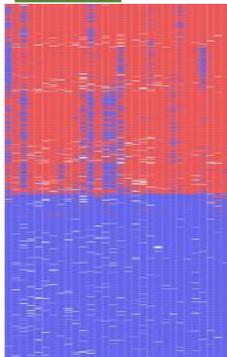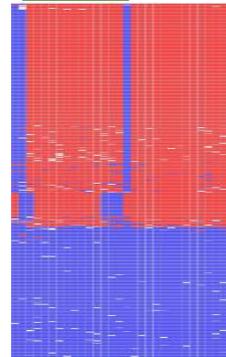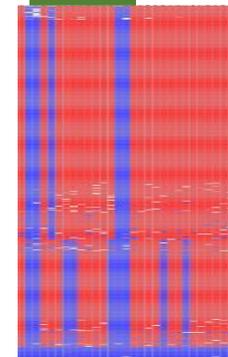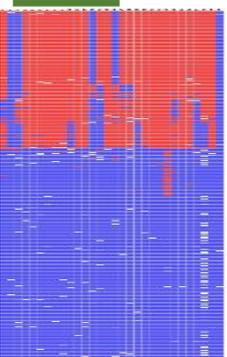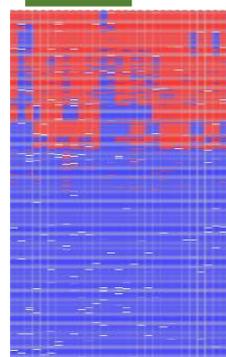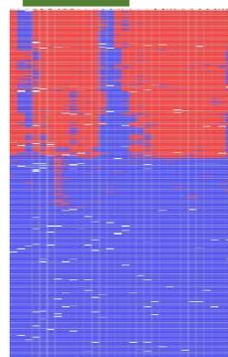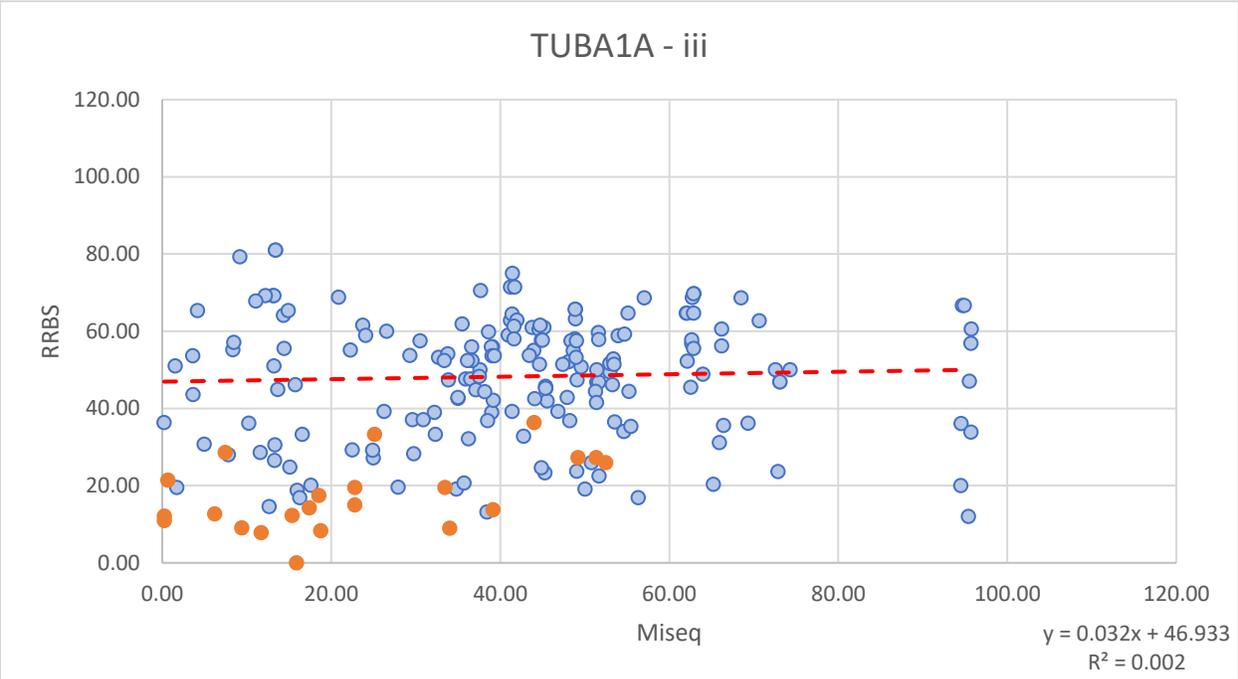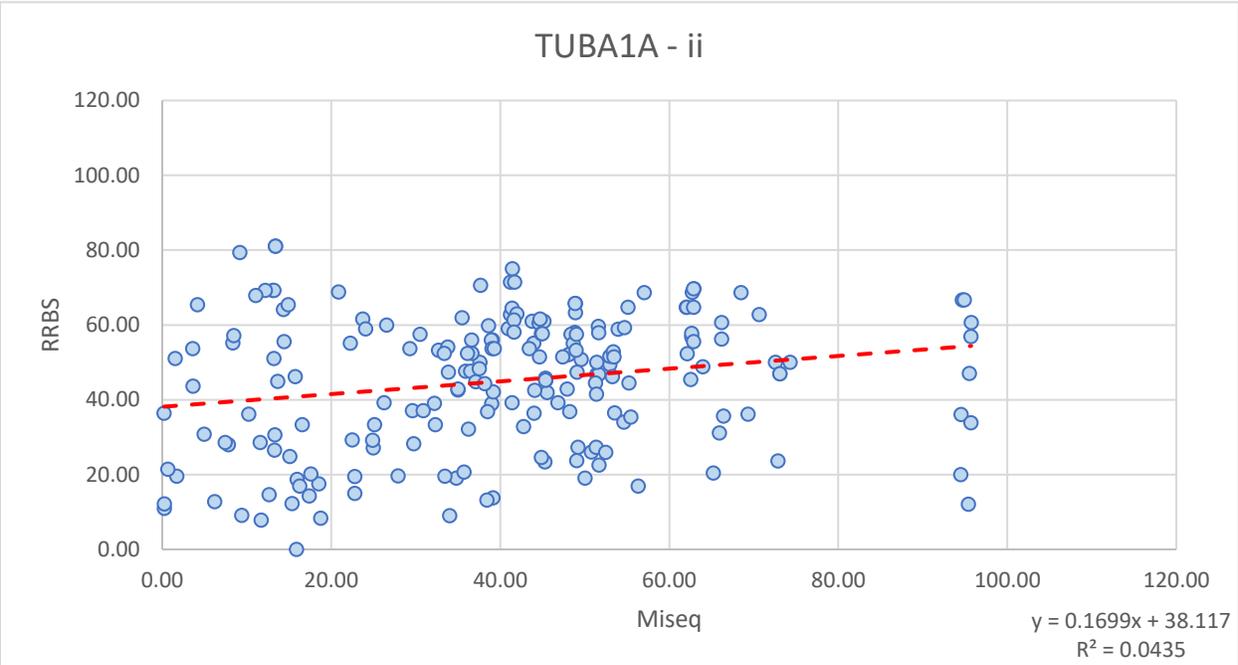
Figure 7.35 *ZNF814* - i

Precursor
Mean met = 0.563
Cov. = 22577
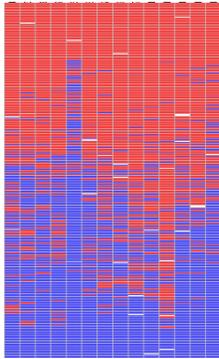
RP_P0nonPC
Mean met = 0.566
Cov. = 14185

RP_P0PC
Mean met = 0.515
Cov. = 17184
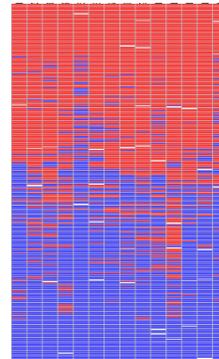
RP_P1PC
Mean met = 0.461
Cov. = 8642
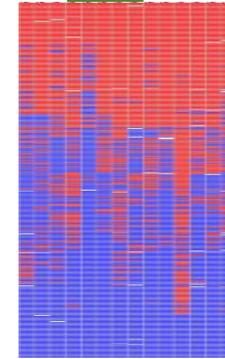
RP_P1PC+A
Mean met = 0.498
Cov. = 15956

RP_P10nonPC
Mean met = 0.477
Cov. = 14076
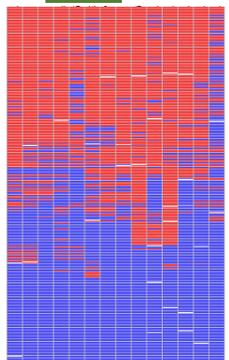
RP_P10PC
Mean met = 0.515
Cov. = 17891
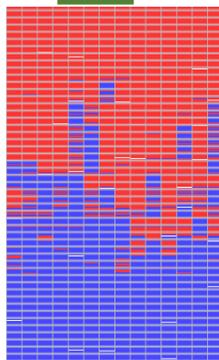
RP_P10PC+A
Mean met = 0.562
Cov. = 17167

RD_P0nonPC+A
Mean met = 0.502
Cov. = 7907
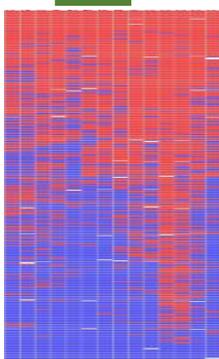
RD_P0PC+hydro
Mean met = 0.494
Cov. = 16289
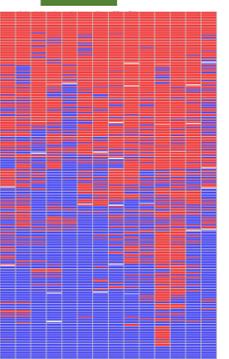
RD_P0PC+DMSO
Mean met = 0.469
Cov. = 25383
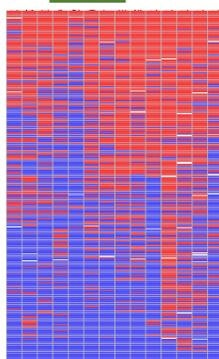
RD_P20
Mean met = 0.275
Cov. = 20960

ZNF814 - ii

y = 0.7734x + 11.195
R² = 0.4712

ZNF814 - iii

y = 0.6443x + 19.937
R² = 0.3195
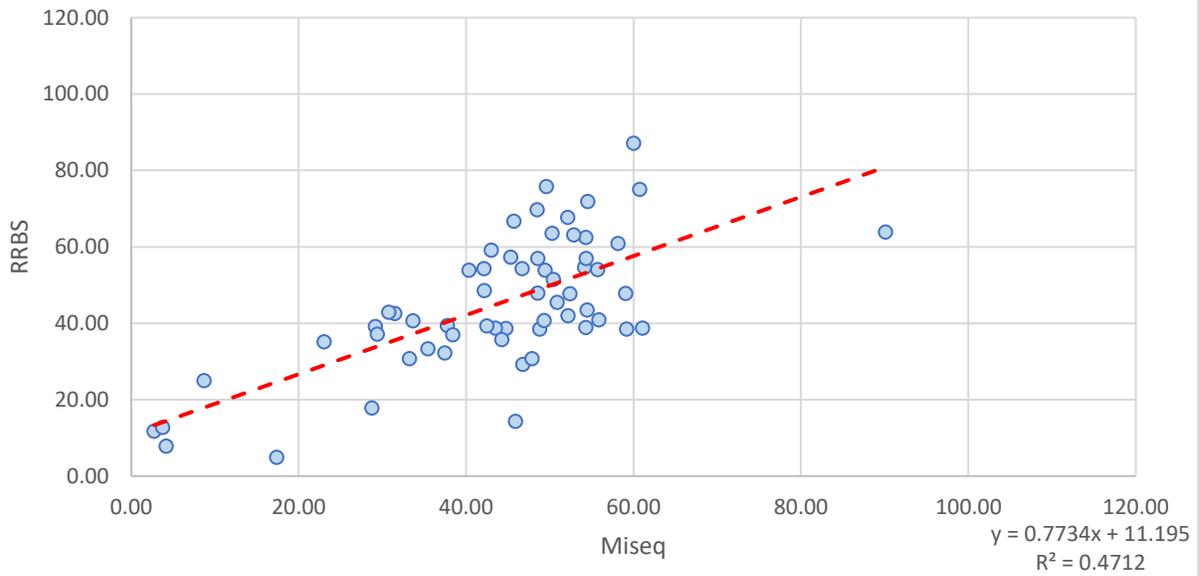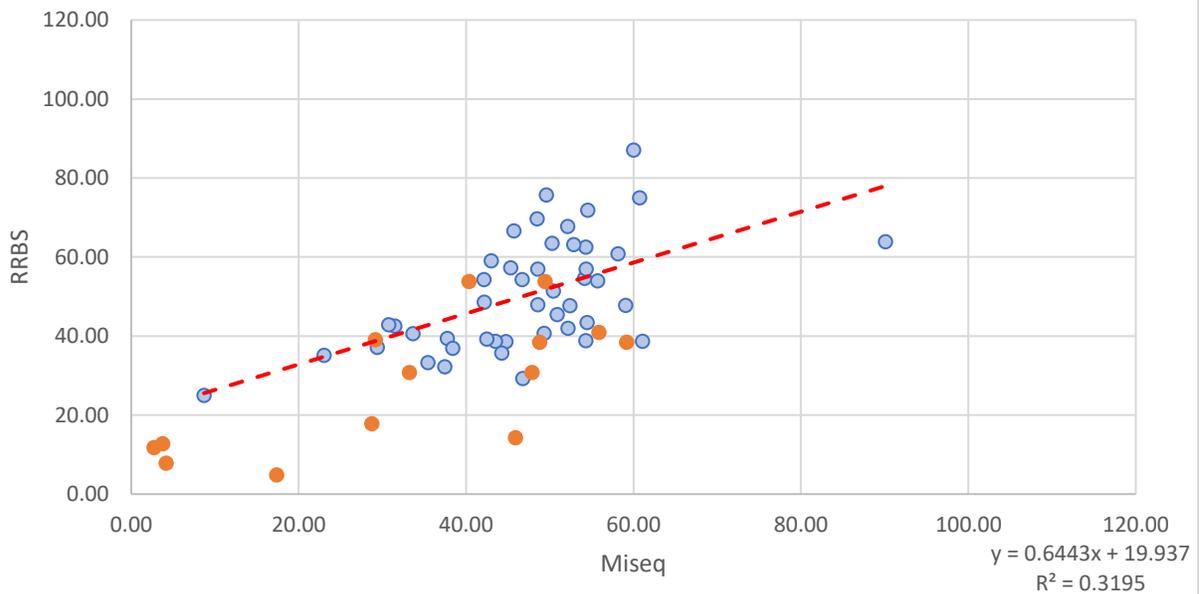
7.3. List of abbreviation

Abbreviation	Description	Abbreviation	Description
3'	Downstream	LIF	Leukemia inhibitory factor
5'	Upstream	me	Methylation
5' Aza/Aza	5-Azacytidine	MET	Mesenchymal-Epithelial transformation
5caC	5-carboxylcytosine	mRNA	Messenger RNA
5fC	5-formylcytosine	MSC	Mesenchymal stem cells
5hmC	5-hydroxymethylcytosine	NSC	Neuronal stem cells
5mC	5-methylcytosine	O_2	Oxigen
ac	Acetylation	^oC	Degree Celcius
aDMR	Anti-correlated DMR	PBS	Phosphate-buffered saline
ATP	Adenosinetriphosphate	pcDMR	PC-specific DMR
bp	Base pair	PC	Physical constraint
CC	Controlled confinement	PC1/2/3/4	Principal component no.1/2/3/4
cDMR	Correlated DMR	PCA	Principal component analysis
Chr	Chromosome	PCR	Polymerase chain reaction
ChromHMM	Chromatin state segmentation	PDMS	Polydimethylsiloxane
Cov	Coverage	PEG	Polyethylene glycol
CpG	Cytosine-Guanine dinucleotide	PHH	Primary human hepatocyte
CTP	Cytidinetriphosphate	qPCR	Quantitative polymerase chain reaction
DEG	Differential expressed gene	R	Correlation coefficient value
DMR	Differential methylated region	RD	PCi-differentiated sample
DMSO	Dimethyl sulfoxide	Refseq	Reference sequence
DNA	Deoxyribonucleic acid	RNA	Ribonucleic acid
ePCR	Electronic or in silico polymerase chain reaction	RP	Reprogramed sample
ESC	Embryonic stem cells	RRBS	Reduces representative bisulfite sequencing
FDR	False discovery rate	RT	Reverse transcription
FPKM	Fragments per kilobase of exon model per million reads mapped	SAM	S-adenosyl methionine
GC	Geometric confinement	sDMR	Stable DMR
GO	Gene Ontology	tRNA	Transfer RNA
HBV	Hepatitis B virus	TSS	Transcription start site
HCV	Hepatitis C virus	TTS	Transcription termination site
HSC	Haematopoietic stem cells	UTR	Untranslated region
IF	Immunofluorescence	vol	volume
iPSC	Induced pluripotent stem cells	w/w	weight by weight
LDS	Local deep sequencing	µ	Micro

7.4. List of figures

7.5. List of tables