# Investigating the role of Lysine Specific Demethylase 1 (LSD1) in embryonic development using a gastruloid model system

Thesis submitted for the degree of Doctor of Philosophy at the University of Leicester

By

Megan Jennifer Broderick MRes (University of Portsmouth)

College of Life Sciences

Department of Molecular and Cell Biology

University of Leicester

December 2022

# Abstract

Investigating the role of Lysine Specific Demethylase 1 (LSD1) in embryonic development using a gastruloid model system

Megan J. Broderick

Lysine specific demethylase 1 (LSD1) is a chromatin modifying protein which specifically demethylates the permissive histone marks H3K4me1/2 and thereby acts as a transcriptional repressor. It performs this activity as part of the CoREST (co-repressor of Repressor Element1 Silencing Transcription Factor) complex, which also encompasses histone deacetylase 1/2 (HDAC1/2) activity and is required for stability of the complex. LSD1 is essential in embryonic development, with loss of LSD1 resulting in embryonic lethality at ~E6.5, a developmental stage which correlates with the onset of gastrulation. To further investigate the role of LSD1 in embryonic development, we employed a model system, gastruloids, which closely mimic early embryogenesis and aspects of gastrulation. We generated gastruloids from both induced and control *Lsd1* conditional knockout (KO) mouse embryonic stem cells (ESCs) and utilised RNA-sequencing (RNA-seq) analysis to identify differentially expressed genes at timepoints representative of ESCs, early gastrulation and late gastrulation. We identified dysregulated expression of genes associated with mesodermal lineages, epithelial-to-mesenchymal transition (EMT) and the bone morphogenic protein (BMP) pathway. To investigate whether gene expression changes were a result of the loss of direct LSD1 demethylase activity or of broader CoREST complex activity, we performed similar RNA-seq analysis on gastruloids generated from cells with wildtype (WT) or catalytically inactive LSD1 rescue constructs. We identified genes whose expression was not rescued, suggesting dependence on LSD1 demethylase activity, and genes whose expression was fully or partially rescued, including genes associated with EMT and the BMP pathway. The work in this thesis has enabled us to gain further insight into the role of LSD1 during embryonic development.

## Covid Impact Statement

Due to the restrictions imposed by the COVID-19 pandemic, my ability to conduct all of the planned lab work of my PhD project was heavily impacted. Between March and July 2020, our lab was shut entirely and upon return, due to building occupancy limits, I was only able to conduct lab work during 25-50% of pre-pandemic hours until July 2021. This impacted the work I was able to complete. For example, I was unable to conduct planned HCR experiments which would have provided spatial gene expression information to complement the work in Chapter 4. Despite receiving a three-month extension, for which I am very grateful, I still lost a significant amount of lab time overall.

## Acknowledgements

Firstly, I would like to thank my supervisor, Professor Shaun Cowley, for giving me the opportunity to join his research group and for his invaluable support and guidance throughout my PhD. I would also like to thank my second supervisor, Professor John Schwabe, and my committee members, Professor Ian Eperon, Dr Pietro Roversi and Dr Catherine Vial, for their helpful comments and discussions. I would like to thank Dr David Turner at the University of Liverpool for his helpful advice and troubleshooting of the gastruloid system. Thank you to all the members of the Cowley lab for their help and support, in particular to Dr Grace Adams and Dr David English for showing me the ropes, and to India-May Baker and Sam Lee for their help in all things R. I am eternally grateful for the support of my friends and family. My thanks must especially go to my parents, who have graciously put up with me for the past few months and whose endless care and encouragement has been a defining factor in getting me to where I am today. I am very grateful to Holly whose constant friendship has made this process immeasurably easier. I would also like to thank Dan Leggat for the help and support over the years. Finally, thank you to Dan, for helping me maintain my sanity and without whom the writing of this thesis would have been far more stressful.

# Table of Contents

## List of Figures and Tables

# Abbreviations

| | |
|---|---|
| ac | acetylation |
| ADD | ATRX-DNMT3-DNMT3L |
| AML | acute myeloid leukemia |
| AOD | amine oxidase-like domain |
| AR | androgen receptor |
| ATPase | adenosine triphosphatase |
| AVE | anterior visceral endoderm |
| BioID | biotin identification |
| BMP | bone morphogenic protein |
| BPTF | bromodomain and PHD domain transcription factor |
| CD | catalytic domain |
| CGI | CpG island |
| CHD | chromodomain helicase DNA-binding |
| CoREST | co-repressor of Repressor Element1 Silencing Transcription Factor |
| DEG | differentially expressed gene |
| DNMT | DNA methyltransferase |
| DVE | distal visceral endoderm |
| EB | embryoid body |
| EHT | endothelial-to-haemopoietic transition |
| EMT | epithelial-to-mesenchymal transition |
| EpiLSC | epiblast-like stem cells |
| EpiSC | epiblast stem cells |
| ESC | embryonic stem cell |
| GO | gene ontology |
| H2A | histone 2A |
| H2B | histone 2B |
| H3 | histone 3 |
| H4 | histone 4 |
| HAT | histone acetyltransferase |
| HCR | hybridisation chain reaction |
| HDAC | histone deacetylase |
| HID | HDAC interaction domain |
| HP1 | heterochromatin protein 1 |
| ICM | inner cell mass |
| ING2 | inhibitor of growth 2 |
| ISWI | imitation switch |
| JAK | Janus-associated kinase |
| JmjC | Jumonji-C |
| KDM | lysine demethylase |
| KMT | lysine methyltransferase |
| KO | knockout |

| | |
|---|---|
| LIF | leukemia inhibitory factor |
| LIF-R | LIF receptor |
| LSD1 | lysine specific demethylase 1 |
| MBD2 | methyl-CpG-binding domain protein 2 |
| me1 | monomethylation |
| me2 | dimethylation |
| me3 | trimethylation |
| MECP2 | methyl-CpG-binding protein |
| MEF | mouse embryonic fibroblast |
| MEK | MAPK/ERK kinase |
| MH1 | mad homology 1 |
| MiDAC | mitotic deacetylase complex |
| NCoR | nuclear receptor corepressor |
| NuRD | nucleosome remodelling and deacetylase complex |
| NURF | nucleosome remodelling factor |
| OGT | O-GlcNAc transferase |
| PAH | paired amphipathic helix |
| PCA | principal component analysis |
| PRC1 | polycomb repressive complex 1 |
| PRC2 | polycomb repressive complex 2 |
| PrE | primitive endoderm |
| REST/NRSF | RE1 Silencing Transcription Factor/neural-restrictive silencing factor |
| RNA-seq | RNA-sequencing |
| scRNA-seq | single cell RNA-sequencing |
| SERCA | sarcoendoplasmic reticulum calcium ATPase |
| SET | Su(var)3–9, Enhancer of Zeste, and Trithorax |
| SMRT | silencing mediator of retinoid and thyroid hormone receptor |
| STAT3 | signal transducer and activator of transcription 3 |
| SUMO | small ubiquitin-related modifier protein |
| SWI/SNF | switch/sucrose non-fermentable |
| TE | trophectoderm |
| TSS | transcription start site |
| ub | ubiquitination |
| WT | wildtype |

# 1 Introduction

## 1.1 Chromatin structure

For DNA to be packaged into the nuclei of cells, it requires a large level of compaction. To achieve this, DNA is wrapped around histone proteins to form chromatin, with 146bp of DNA circling each histone octamer to produce a nucleosome (Kornberg, 1997, Luger et al., 1997). The histone octamer comprises a tetramer formed of two units each of histone 3 (H3) and histone 4 (H4), as well as two dimers formed of histone 2A (H2A) and histone 2B (H2B) (Eickbush and Moudrianakis, 1978). The classical view of the organisation of chromatin suggested nucleosomes along DNA created a 10 nm diameter fibre resembling 'beads on a string', with the linker histone H1 binding to DNA between nucleosomes and promoting further condensation into the 30nm chromatin fibre (Figure 1.1A) (Paranjape et al., 1994). However, it is now largely accepted that these regular chromatin structures observed *in vitro* are not representative of the conformation of chromatin *in vivo*, with this instead presenting a highly random organisation in which chromatin fibres intertwine (Figure 1.1B) (Moraru and Schalch, 2019). The presence of positively charged residues on the histone octamer surface enables binding of DNA around the proteins (Arents and Moudrianakis, 1993). These electrostatic interactions between DNA and the histone octamer inhibit transcription of the associated DNA (Wasylyk and Chambon, 1979).

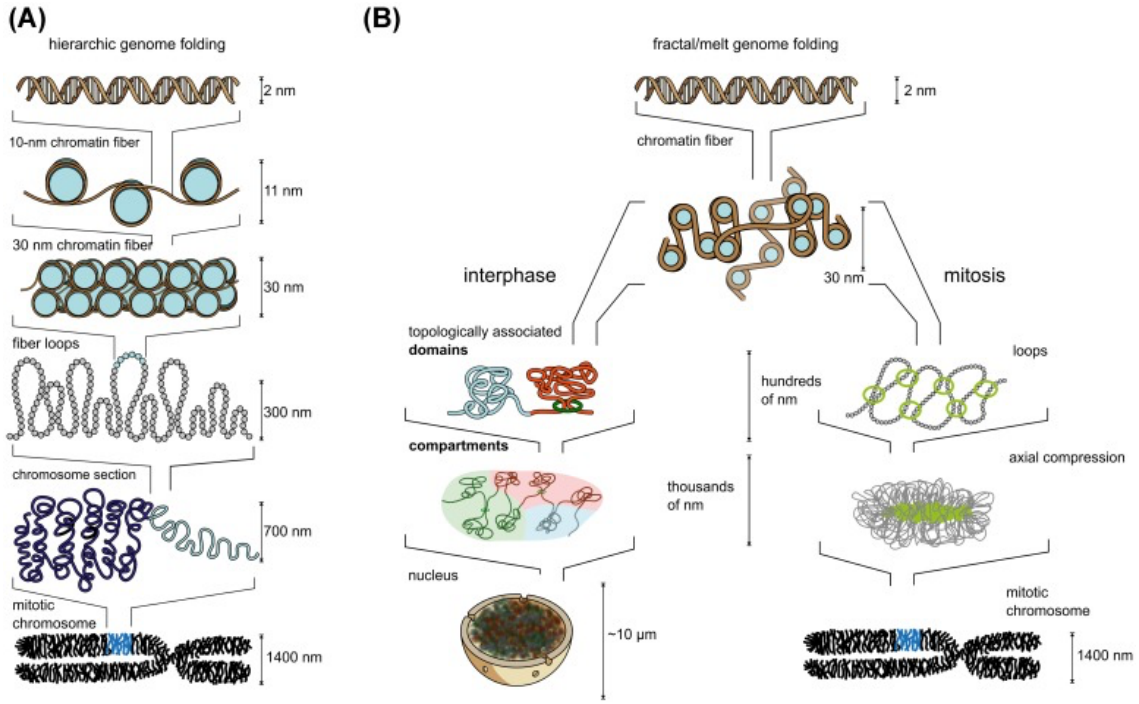


***Figure 1.1 Classical versus modern model of chromatin.*** *(**A**) Diagram of the classically accepted hierarchical model of chromatin. (**B**) Diagram of the more recently established fractal/melt model of chromatin folding. Taken from (Moraru and Schalch, 2019)*

Not only does the compaction of DNA into chromatin allow effective packaging into cells, but it also regulates the accessibility of the underlying DNA. Condensation of chromatin into heterochromatin restricts access to the DNA, resulting in inaccessibility and repression of associated genes. In contrast, euchromatin is less condensed regions of chromatin which allow access of transcription machinery to underlying genes and is permissive for gene expression (Dillon and Festenstein, 2002). Due to the requirement for dynamic gene expression depending on different spatial and temporal contexts, it is necessary for the compaction and decompaction of chromatin to be an equally dynamic process. This allows for differential accessibility of gene promoters and enhancers to transcription factors and the transcriptional machinery. This is achieved in a number of ways, including DNA methylation, modifications to histone tails (Section 1.2) and chromatin remodelling.

Chromatin remodelling involves the action of adenosine triphosphatase (ATPase) dependent complexes, which translocate DNA and break contacts between histones and DNA. These complexes can be separated into four subfamilies: imitation switch

(ISWI), chromodomain helicase DNA-binding (CHD), switch/sucrose non-fermentable (SWI/SNF) and INO80 (Clapier et al., 2017). These remodelling complexes mediate nucleosomal spacing, exchanging of histone variants, and movement or ejection of nucleosomes to enable binding of transcription factors. The nucleosome remodelling and deacetylase (NuRD) complex combines nucleosome remodelling activities with histone deacetylase activities, through its association with CHD3/4 and HDAC1/2, respectively (Millard et al., 2016). The NuRD complex is therefore an example of a complex possessing concerted chromatin modifying activities.

DNA methylation at the C-5 position of cytosine (5mC) is associated with transcriptional repression. *De novo* DNA methylation is catalysed by DNMT3A and DNMT3B (Okano et al., 1999), with DNMT3L stimulating this activity specifically in the germline (Greenberg and Bourc'his, 2019). The maintenance of DNA methylation following DNA replication is performed by DNMT1 in concert with UHRF1 (Bostick et al., 2007). The majority of mammalian gene promoters are GC rich, and some promoters contain CpG rich regions termed CpG islands (CGIs), which are rarely methylated. There are, however, three circumstances in which DNA methylation-based silencing at CGIs are important: in X-chromosome inactivation, genomic imprinting and germline-specific genes (Greenberg and Bourc'his, 2019). DNA methylation is excluded from the promoters of actively transcribed genes through the presence of H3K4 methylation. This occurs through disruption of the autoinhibition of the catalytic domain (CD) of DNMT3A by its ATRX–DNMT3–DNMT3L (ADD) domain by unmethylated H3K4, but not by H3K4me3 (Guo et al., 2015). In most cases, methylation of DNA leads to loss of or weakened binding of associated transcription factors, causing gene repression. However, some transcriptional factors show binding preference for methylated DNA, so in some cases DNA methylation can promote gene activation (Yin et al., 2017). DNA methylation has also been shown to influence transcriptional silencing through recruitment of both histone modifying and chromatin remodelling complexes. For example, MECP2 (methyl-CpG-binding protein) has been shown to recruit the NCoR/SMRT histone deacetylase complex to methylated DNA (Lyst et al., 2013). Additionally, the NuRD complex is directed to methylated DNA through interaction with MBD2 (methyl-CpG-binding domain protein 2), resulting in chromatin remodelling (Zhang et al., 1999).

## 1.2 Histone modifications

Histone structure consists of a core globular domain with N-terminal tail protrusions which contain the sites of post-translational modifications (Luger et al, 1997). The primary histone modifications include acetylation, methylation, phosphorylation, ubiquitination and sumoylation (Kouzarides, 2007). The modifications are typically representative of either an active or repressive chromatin state to a point which their presence can be predictive of these states (Heintzman et al., 2007). For example, H3K4me3, H4K16ac and H3K27ac are transcriptionally permissive gene marks, whereas H3K27me3 and H3K9me3 are repressive (Schubeler et al., 2004, Wang et al., 2008). The presence of histone modifications affects epigenetic regulation in two main ways: firstly, through alteration of higher-order chromatin structure and secondly, through recruitment of regulatory proteins which bind via specific domains (Kouzarides, 2007). For example, acetylation is recognised by bromodomains and tandem PHD fingers, methylation is recognised by the Royal family domains (Tudor, PWWP, MBT and chromodomains) as well as PHD fingers, and 14-3-3 proteins recognise phosphorylation (Yap and Zhou, 2011). The binding of proteins containing these domains cause direct or indirect changes, often through tethering of proteins with enzymatic activity to chromatin (Kouzarides, 2007). Modifications are limited to specific residues, for example, with acetylation occurring on particular lysines, methylation restricted to certain lysines and arginines, and phosphorylation on serine/threonine residues (Figure 1.2) (Kouzarides, 2007). A number of these residues can harbour multiple different modifications, for example H3K9 can be acetylated or methylated, with the mutually exclusive nature of these modifications resulting in antagonism (Rea et al., 2000). Adjacent modifications can also influence the binding of enzymes, via disrupted or increased binding, or with a requirement for the presence of an adjacent modification for the placement of secondary modifications. For example, phosphorylation of H3S10 inhibits methylation of the adjacent lysine 9 and additionally prevents recruitment of heterochromatin protein 1 (HP1) to H3K9me3 (Rea et al., 2000, Fischle et al., 2005). The PHD finger of RAG2 shows increased binding to H3K4me3 and symmetrically dimethylated H3R2 than to H3K4me3 alone (Ramón-Maiques et al., 2007).

In yeast, ubiquitination of H2B is required prior to methylation of H3K4 and H3K79 (Chandrasekharan et al., 2010). The ability of particular residues to exhibit multiple levels of modification, for example, unmodified, mono-, di- or trimethylated H3K4, adds another level of complexity (Bannister and Kouzarides, 2005).

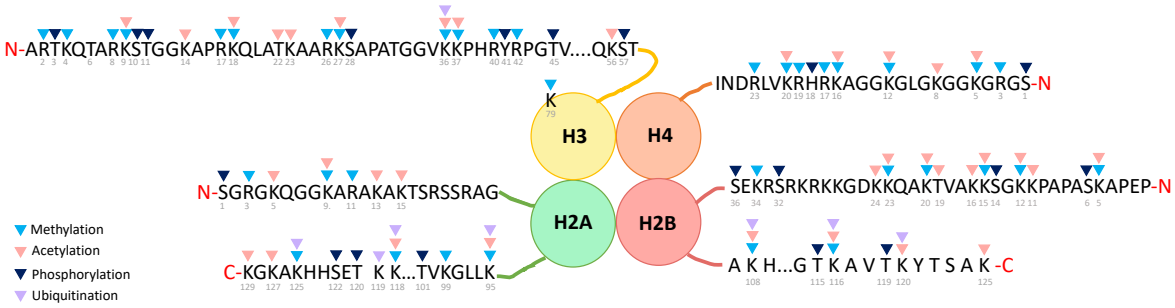


***Figure 1.2 Known histone modification sites on the core histone proteins.*** *The known sites of histone methylations, acetylations, phosphorylations and ubiquitinations are shown. Adapted from (Huang et al., 2014)*

## 1.2.1 Acetylation

A major player in the level of compaction of chromatin structure is histone acetylation. Generally associated with euchromatin, acetylation of histones acts to neutralise the positive charges of lysines, thereby 'unfolding' chromatin and increasing accessibility of genes to transcriptional machinery (Hong et al., 1993). In contrast, deacetylation of histones promotes condensed chromatin and contributes to a repressed state of neighbouring genes. Addition of acetyl groups to lysine residues occurs through the enzymatic action of histone acetyltransferases (HATs) and deacetylation is performed by histone deacetylases (HDACs) (Kouzarides, 2007).

HATs can be classified into Type A, nuclear HATs and Type B, cytoplasmic HATs. Type A HATs can further be classified into GCN5/PCAF, MYST and p300/CBP families, based on sequence similarity (Hodawadekar and Marmorstein, 2007). HATs are recruited into multi-protein complexes, which contributes to histone substrate specificity and thereby directs acetyltransferase activity (Sterner and Berger, 2000). For example, in yeast, GCN5 can be recruited into either SAGA or ADA complexes, which confers specificity for H3K-9, -14, -18, and -23 or H3K-14 and -18, respectively (Grant and Berger, 1999).

HDACs belong to either the zinc-dependent classical family, split into 4 classes (class I, IIa, IIb and IV), or the NAD+ dependent deacetylases, sirtuins (class III). Class I HDACs which include HDAC1, HDAC2, HDAC3, and HDAC8 are localised primarily in the nucleus. Of the Class I HDACs, HDAC1, HDAC2, and HDAC3 form complexes, showing reduced activity when not in these complexes (Kelly and Cowley, 2013), whereas HDAC8 deacetylates its substrates in isolation (Hu et al., 2000). Class II HDACs are divided into class IIa, which contains HDAC4, HDAC5, HDAC7 and HDAC9, and class IIb, which contains HDAC6 and HDAC10. HDAC11 is the sole member of class IV (Delcuve et al., 2012).

Mapping of individual histone acetylation sites has revealed functional significance, with H3K9ac, H3K18ac, H3K27ac located primarily around transcription start sites (TSSs) and with H3K4, H4K5, H4K8, H4K12 and H4K16 acetylations prominent at promoters and transcribed regions of active genes (Wang et al., 2008). H3K27ac is enriched at active enhancers, which are distal regulatory regions that enhance transcription (Creyghton et al., 2010). Acetylation of H4K16 specifically impedes the compaction of chromatin into the 30nm fibre (Shogren-Knaak et al., 2006).

Acetylated lysines are recognised and bound by bromodomains. The presence of bromodomains within chromatin-related and transcription-related protein complexes allows direction of these complexes to acetylated residues. For example, HATs such as PCAF, GCN5, and p300/CBP possess bromodomains, directing additional acetylation to previously acetylated areas (Smith and Zhou, 2016). The multi-protein complex TFIID, which initiates the assembly of the transcription machinery, comprises two tandem bromodomains in its subunit $TAF_{II}250$ (Jacobson et al., 2000). More recently, the YEATS and tandem PHD finger domains have also been identified as domains that recognise acetylated lysines (Zeng et al., 2010, Zhao et al., 2017).

## 1.2.2 Methylation

Histone methylation occurs on arginine and lysine residues (Bannister and Kouzarides, 2005). Unlike acetylation, the addition of methyl groups to these residues does not neutralise their charge and therefore doesn't directly alter the conformation of chromatin. Also, in contrast to acetylation, methylation can occur in multiple states, with lysine residues being mono-, di- or tri-methylated and arginine residues being mono- or (symmetrically/asymmetrically) di-methylated (Figure 1.3) (Bannister and Kouzarides, 2005). Interestingly, lysine methylation can be associated with both permissive and repressive chromatin states. Typically, methylation of H3K4 and H3K79 mark transcriptionally permissive chromatin regions, whereas H3K9 and H3K27 mark silenced regions of chromatin (Hyun et al., 2017).

### 1.2.2.1 Lysine Methylation

The sites for lysine methylation include H1K26, H4K20 and six residues on H3: K4, K9, K26, K27, K36 and K79 (Musselman et al., 2012). These sites are recognised by a number of different protein domains, including chromodomains, PHD finger domains, PWWP domains, MBT domains and Tudor domains (Musselman et al., 2012). The recognition of these sites occurs primarily through a conserved mechanism, in which aromatic cages within the domains form around the methylated lysine residues (Musselman et al., 2014). The composition and size of the aromatic cages confers substrate specificity, with smaller cages excluding higher methylation states due to steric hindrance and the presence of negatively charged residues contributing preference to mono- and di-methylated over tri-methylated lysines. For example, the PHD finger of BHC80, which constitutes part of the CoREST complex, binds unmethylated H3K4 and this binding is negated by methylation of this residue (Lan et al., 2007). In contrast, the PHD finger domains of tumour suppressor ING2 (inhibitor of growth 2) and of BPTF (bromodomain and PHD transcription factor), a subunit of NURF (nucleosome remodelling factor), bind H3K4me3 preferentially (Pena et al., 2006, Li et al., 2006). Employing these different

domains within multiprotein complexes allows for downstream functional effects for gene transcription which are dependent on the marks present.

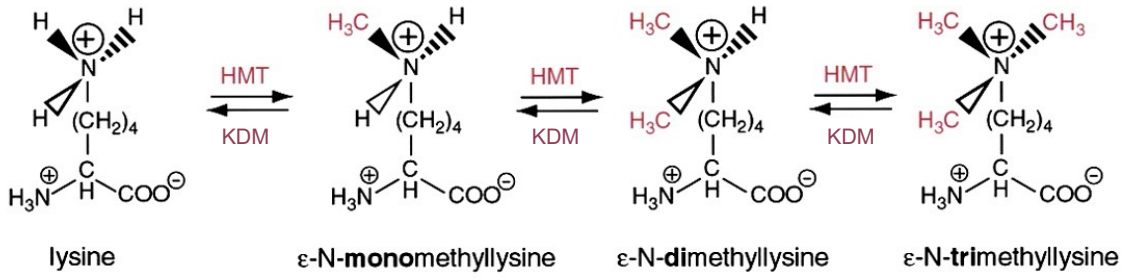


***Figure 1.3 Structures of methylated lysines in histone tails.*** *Chemical structures of lysine and its methylated forms are shown. Actions of HMTs and KDMs are indicated by arrows. Adapted from (Bannister et al., 2002)*

H3K4 methylations are generally associated with permissive chromatin, with different levels associated with different regions (Bernstein et al., 2005, Barski et al., 2007). All three methylation states are associated with the transcription start sites (TSSs) of known genes, with H3K4me3 correlating with gene expression. Mono-methylation of H3K4 marks enhancers and is enriched across gene bodies. H3K4me2 similarly correlates with gene bodies of actively transcribed genes and has shown to be present at enhancers (Barski et al., 2007, Heintzman et al., 2007). H3K4me3 also colocalises with H3K27me3, a mark associated with gene silencing, on regions termed “bivalent domains”, which poises developmental genes in embryonic stem cells for activation (Bernstein et al., 2006). H3K36me3 is present in the transcribed regions of active genes and peaks at the 3’ end of the genes (Bernstein et al., 2005, Barski et al., 2007).

Methylation of H3K9 and H3K27 is associated with gene silencing, in part because these marks are antagonistic to acetylation at these sites. H3K9 di- and tri-methylation is recognised by heterochromatin protein 1 (HP1), leading to silencing of genes and heterochromatin assembly (Bannister et al., 2001). H3K9 methylation and HP1 binding is essential for the correct formation of heterochromatin at the inactive X chromosome and pericentric heterochromatin (Peters et al., 2001, Boggs et al., 2002). H3K9me2 and H3K9me3 are enriched at the TSSs of silenced genes and, in contrast, H3K9me can be associated with active promotors (Barski et al., 2007). H3K27me3 is enriched at silenced

promoters and reduced at active promoters, with H3K27me2 showing similar patterns (Barski et al., 2007). Tri-methylation of H3K27 is associated with PRC2 (polycomb repressive complex 2) mediated repression of developmental genes (Boyer et al., 2006).

### 1.2.2.3 Histone Methyltransferases

Histone lysine methyl transferases (KMTs) are responsible for laying down methyl groups at competent sites. All but one of the numerous KMTs that have been identified possess a SET (Su(var)3–9, Enhancer of Zeste, and Trithorax) domain (Dillon et al., 2005). The exception to this is the DOT1 protein and its homologs, which specifically methylate H3K79 and do not contain SET domains (Feng et al., 2002). KMTs are highly specific to particular residues and methylation states and can be grouped accordingly (Figure 1.4). In mammals, H3K4 methylations are predominantly carried out by six homologs of the yeast SET1 protein: MLL1-4 (KMT2A-D), SET1A (KMT2F) and SET1B (KMT2G) (Hyun et al., 2017). These KMTs form multi-subunit protein complexes, of which RBBP5, ASH2L and WDR5 are common core components (Dou et al., 2006). In addition to methylation by these SET1 homologs, H3K4 mono-, di- and tri-methylation has also been shown to be catalysed by PRDM9 (Wu et al., 2013). Mono-methylation of H3K4 can be catalysed by SET7 (Keating and El-Osta, 2013) and by SMYD2, which usually catalyses H3K36 methylation (Abu-Farha et al., 2008).

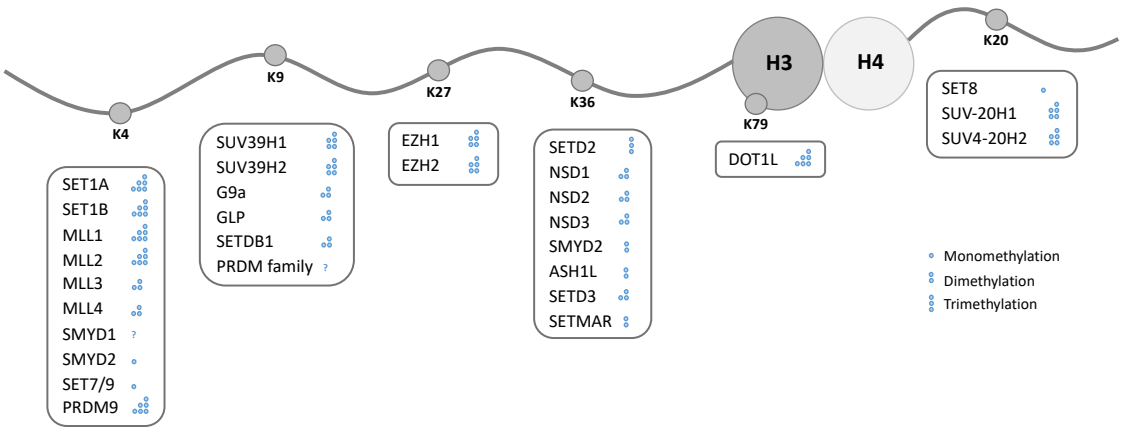


***Figure 1.4 Sites specific to individual methyltransferases.*** *Diagram showing the sites of lysine methylation on histones H3 and H4 and their corresponding methyltransferases. Methylation state specificities for each methyltransferase are indicated by dots: single dot, me; double dot, me2; and triple dot, me3. Adapted from (Hyun et al., 2017)*

### 1.2.2.4 Histone Demethylases

Until the discovery of the first histone demethylase, LSD1, histone methylation was thought to be a stable modification (Shi et al., 2004). Histone lysine demethylases (KDMs) can be split into two main classes. Firstly, the monoamine oxidase-like KDMs, LSD1 and LSD2, and secondly the KDMs which contain a Jumonji-C (JmjC) domain (Black et al., 2012). These demethylases show high specificity for both site and degree of methylation (Figure 1.5).

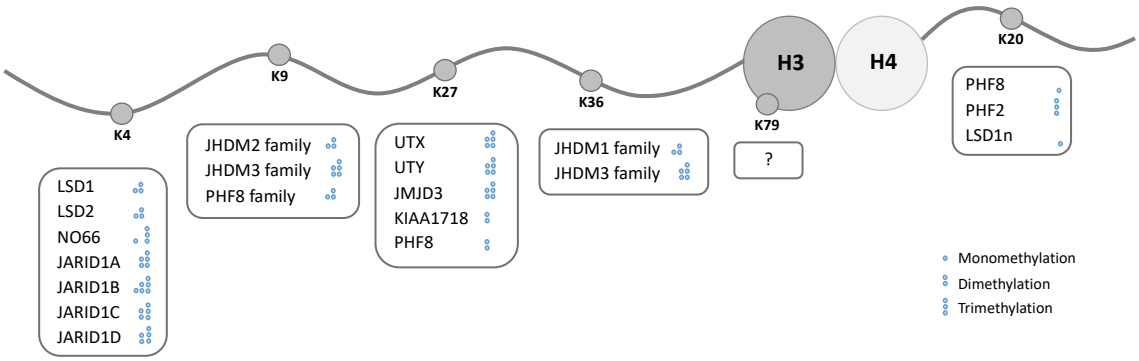


***Figure 1.5 Sites specific to individual demethylases.*** *Diagram showing the sites of lysine methylation on histones H3 and H4 and their corresponding demethylases. Methylation state specificities for each demethylase are indicated by dots: single dot, me; double dot, me2; and triple dot, me3. Adapted from (Hyun et al., 2017)*

KDMs which contain a JmjC domain catalyse the removal of methyl groups via a Fe(II) and $\alpha$-ketoglutarate dependent oxidative reaction (Tsukada et al., 2006). The JmjC family contains 30 members, which can be split into sub-families based on homology and the domains present (Franci et al., 2014). The JmjC KDMs are highly specific for individual sites and methylation levels and, due to their mode of catalysis, are able to demethylate tri-methylated residues (Tsukada et al., 2006, Klose et al., 2006).

LSD1 specifically demethylates H3K4me1/H3K4me2, which are marks associated with active transcription (Shi et al., 2004). This, combined with the fact it has been found to be associated with a corepressor complex, implicates it in transcriptional repression (Shi et al., 2005). LSD1 demethylase activity is dependent on its cofactor, flavin adenine dinucleotide (FAD). The first step in the demethylation reaction catalysed by LSD1 is flavin-mediated oxidation, which produces an imine intermediate that is subsequently

hydrolysed to form an amine and an aldehyde (Forneris et al., 2005b, Shi et al., 2004). The reduced cofactor, $FADH_2$, is then deoxidised, forming $H_2O_2$ and allowing for subsequent reactions (Figure 1.6). The activity of LSD1 is limited to mono- and di-methylated H3K4, due to the requirement of a protonated nitrogen in substrates of flavin-dependent oxidases (Shi et al., 2004). Binding of LSD1 to the H3 histone tail is greatly influenced by the modification state of the neighbouring residues. For example, Ser10 is involved in interactions which stabilise the conformation of the peptide within the enzyme pocket, hence the disruption that phosphorylation of this residue causes (Forneris et al., 2007). Arginine methylation at R2 and R8 also greatly reduce binding affinity of LSD1, as these residues are involved in forming intramolecular hydrogen bonds (Forneris et al., 2007). In contrast, substrate binding of LSD1 was greatly increased by the presence of H3K9 acetylation (Forneris et al., 2005a). Substrate recognition by LSD1 requires, at a minimum, 20 amino acids of the H3 tail, suggesting significant levels of specificity (Forneris et al., 2005a). The first seven residues of the H3 tail fit into the active site cavity and the number of residues on the N-terminal side of the methyllysine is limited to three, due to the size and conformation of this cavity (Yang et al., 2007). LSD1 has been implicated in the activation of androgen-receptor (AR) target genes, through demethylation of H3K9 when in association with AR (Metzger et al., 2005). However, the specificity of binding within the catalytic pocket of LSD1 outlined above brings into question this secondary AR related activity, and this suggested activity has yet to be structurally solved.

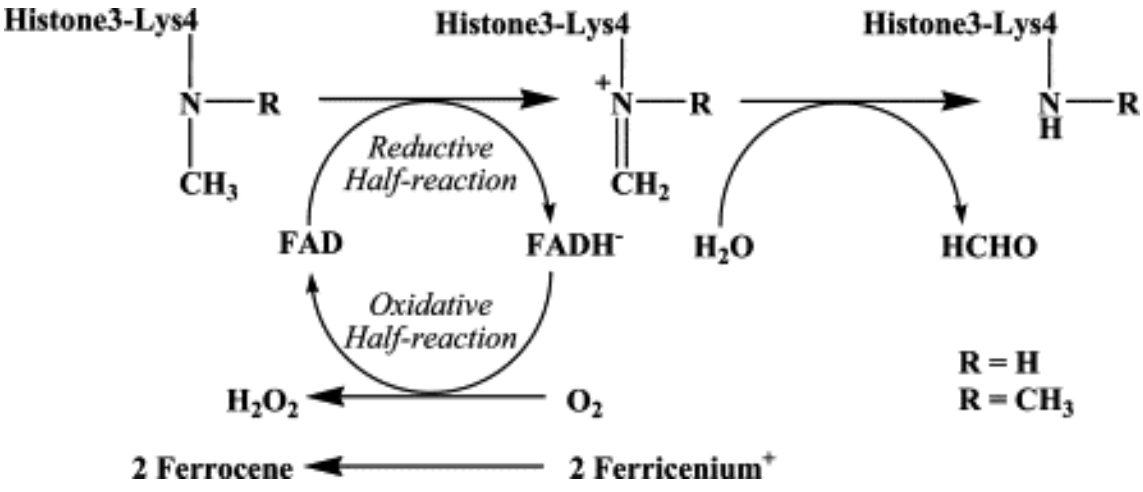


***Figure 1.6 Schematic of the demethylation reaction catalysed by LSD1.*** *Taken from (Forneris et al., 2005b)*

Homology analysis led to the discovery of a second flavin-dependent histone demethylase, LSD2 (Karytinos et al., 2009). Demethylation by LSD2 is, similarly to LSD1, specific for H3K4 mono- and di-methylation and occurs using the same catalytic reaction. However, LSD2 lacks a tower domain and so does not have the same association with CoREST as LSD1 (Karytinos et al., 2009). Instead, it possesses a zinc finger domain within its N-terminal, which consists of a C4H2C2-type zinc finger and a CW-type zinc finger, and this domain is required for LSD2 activity (Zhang et al., 2013a). Genome-wide mapping of LSD2 shows that it occupies gene bodies of active genes and, along with its complex constituents, appears to contribute to gene activation (Fang et al., 2010). LSD2 has also recently been shown to possess E3 ubiquitin ligase activity, presenting both autoubiquitination activity and ubiquitination of O-GlcNAc transferase (OGT), promoting degradation (Yang et al., 2015).

### 1.2.3 Phosphorylation

Phosphorylation of serine, threonine and tyrosine residues operates similarly to lysine acetylation, in that it involves addition of a negatively charged modification which affects the electrostatic properties and therefore conformation of chromatin (Musselman et al., 2012). One of the key roles of phosphorylation is in the DNA damage response, where phosphorylation of S139 on the H2AX variant histone, also referred to as γH2AX, is recognised by DNA damage repair factors (Rossetto et al., 2012). Phosphorylation of H3S10, T11 and S28 residues have been implicated in transcriptional activation, as they have been shown to be linked with acetylation of H3 (Rossetto et al., 2012). H3S10 phosphorylation has also been implicated in chromatin condensation during mitosis, which conflicts with its role in transcriptional activation (Nowak and Corces, 2004).

### 1.2.4 Ubiquitination and Sumoylation

In contrast to other histone modifications, ubiquitination involves the addition of a much larger 76 amino acid peptide to histones through the successive action of E1, E2 and E3 enzymes (Bannister and Kouzarides, 2011). The most common histone ubiquitinations are mono-ubiquitinations of H2AK119, which is associated with gene silencing, and H2BK120, which plays a role in transcriptional activation and elongation (Cao and Yan, 2012). H2AK119 is ubiquitinated by RING1B, an E3 ubiquitin ligase, as part of the polycomb repressive complex 1 (PRC1) thereby linking this modification with polycomb-related gene silencing (Wang et al., 2004). H2BK120ub enhances H3K4 and H3K79 methylation, influencing transcriptional activation, and is involved in the stabilisation-destabilisation cycle of nucleosomes required for RNA polymerase II elongation (Chandrasekharan et al., 2010). As H2BK120ub stabilises nucleosomes, deubiquitination of this residue during transcription allows RNA polymerase II progression.

The addition of small ubiquitin-related modifier protein (SUMO) to residues, termed sumoylation, is involved with transcriptional repression through recruitment of histone deacetylase complexes (Shiio and Eisenman, 2003). Recognition of sumoylated H4 by a SUMO interaction motif in CoREST increases occupancy of the LSD1-CoREST complex and mediates the demethylation of H3K4 (Dhall et al., 2017).

## 1.3 HDAC1/2 complexes

Class I HDACs, with the exception of HDAC8, are recruited into multi-subunit complexes which impart chromatin binding activity (Yang and Seto, 2008). The CoREST, SIN3, NuRD and MiDAC complexes associate with both HDAC1 and HDAC2, whereas the SMRT/NCoR complex contains HDAC3 (Figure 1.7) (Wang et al., 2020). HDAC1 and HDAC2 are highly related, sharing around 80% sequence similarity, and exhibit partial redundancy, compensating for each other when the other is lost. The deacetylase activity of each complex is site specific, for example, H3K14ac is resistant to the deacetylase activity of

CoREST but not of other HDAC complexes, suggesting that each complex has unique cellular functions (Wang et al., 2020). These complexes often possess a cohort of chromatin modifying activity, for example, the NuRD complex possesses both histone deacetylase and chromatin remodeling activity, and the CoREST complex has both lysine deacetylase and demethylase subunits (Hayakawa and Nakayama, 2011). Apart from the SIN3 complex, the complexes interact with HDAC1/2 via a ELM2-SANT domain within one of their constituent components (Millard et al., 2017).

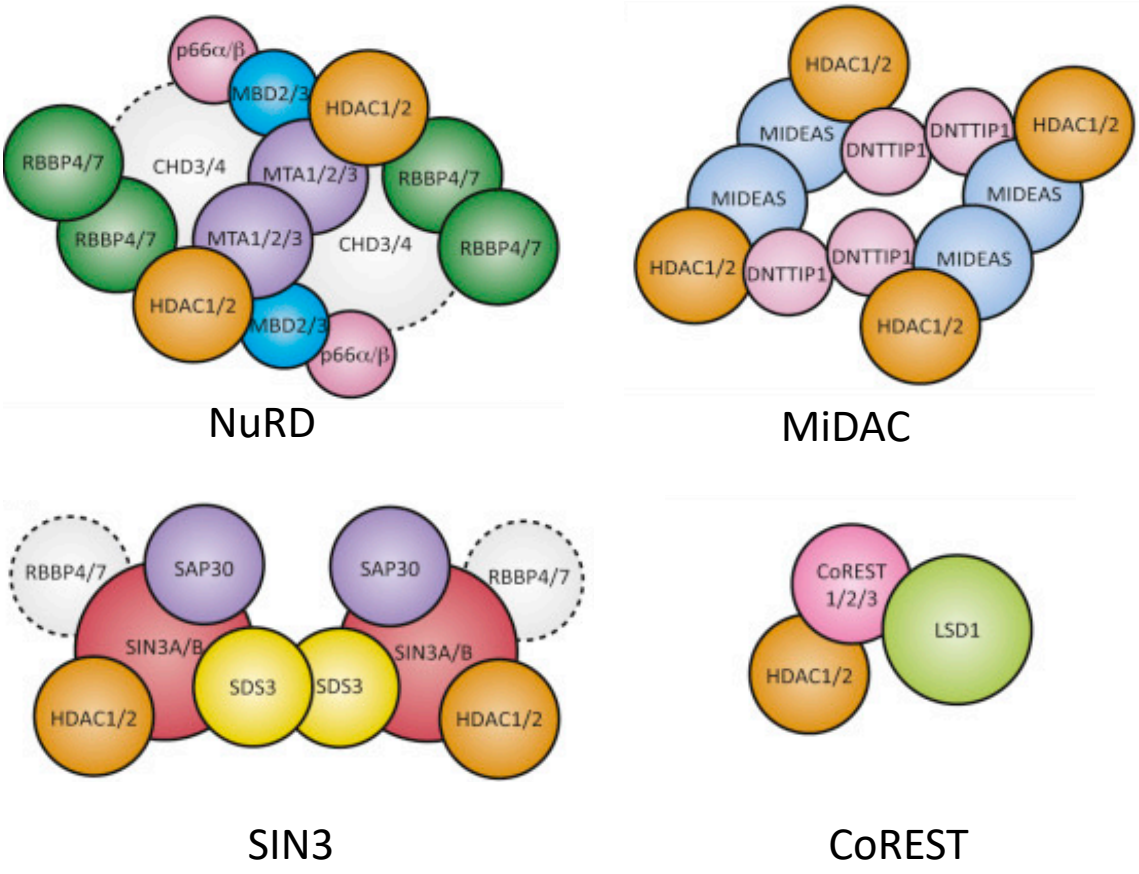


***Figure 1.7 Structure of HDAC1/2 containing complexes.*** *Schematics of each HDAC1/2 containing complex, and their constituent components are shown. Adapted from (Millard et al., 2017)*

### 1.3.1 SIN3

Mammals have two SIN3 homologues, SIN3A and SIN3B. The core SIN3 complex consists of SIN3A/B, HDAC1, and HDAC2, the retinoblastoma-binding proteins RBBP4 and RBBP7, SAP30, SAP18, and SDS3 (Grzenda et al., 2009). In addition to the core complex, SIN3 also interacts with additional subunits that can target the complex to specific areas of

DNA. For example, the ING2 subunit interacts with H3K4me3 via its PHD domain, targeting the repressive activity of the SIN3 complex to gene promoters (Shi et al., 2006). SIN3A interacts with HDAC1 and HDAC2 via its HID (HDAC interaction domain) (Laherty et al., 1997), and with other protein interactors through its paired amphipathic helix (PAH) domains (Le Guezennec et al., 2006). Despite exhibiting 57% shared identity, SIN3B and SIN3B have been shown to have individual, non-redundant roles, as knock out of the SIN3A isoform results in embryonic lethality (Cowley et al., 2005). In ESCs, SIN3A and Nanog co-occupy pluripotency genes and induce their transcription (Saunders et al., 2017).

### 1.3.2 NuRD

The NuRD complex consists of CHD3/4, HDAC1/2, MBD2/3, RBBP7/4, MTA1/2/3 and p66α/β (Allen et al., 2013). CHD3/4 ATPase confers ATP-dependent nucleosomal sliding and repositioning activities to the complex. The metastasis associated proteins, MTA1, MTA2 and MTA3, act as scaffolding proteins, with their ELM2-SANT domains interacting with HDAC1 (Millard et al., 2013) and a C-terminal motif interacting with RBBP7/4 (Alqarni et al., 2014). The complex is formed from a HDAC1:MTA1 dimer which in turn recruits four RBBP4/7 subunits (Millard et al., 2016). The WD40 repeat found within the RBBP7 and RBBP4 subunits interacts with histones H3 and H4, potentially recruiting the complex to chromatin (Murzina et al., 2008, Zhang et al., 2013b). MBD2 and MBD3 are present interchangeably in the NuRD complex and are required for its stabilisation. While MBD2 binds and directs the complex to methylated DNA, MBD3 lacks this activity (Allen et al., 2013). p66α/β interact with MBD2 and enhance its transcriptional repression (Brackertz et al., 2006). These proteins have also shown direct binding activities to all histone tails, though their binding is affected by post-translational modifications.

Existence of an LSD1/NuRD complex has previously been suggested, as LSD1 was shown to coimmunoprecipitate with components of the NuRD complex, and this association has been implicated in enhancer decommissioning (Wang et al., 2009b, Whyte et al., 2012).

However, more recently, biotin identification (BioID) analysis of LSD1 proximal proteins only detected the CHD3/4 and MTA1 components of the NuRD complex, and association with the other components was absent (Barnes et al., 2022).

### 1.3.3 MiDAC

In 2011, a novel HDAC containing complex, termed the mitotic deacetylase complex (MiDAC), due to its increased activity in mitotic cells, was discovered via proteomic screening (Bantscheff et al., 2011). The MiDAC complex consists of a mitotic deacetylase-associated SANT domain protein (MIDEAS), HDAC1/2 and DNTTIP1. MIDEAS contains an ELM2-SANT domain which facilitates its association with HDAC1/2 (Itoh et al., 2015). DNTTIP1 contains an N-terminal domain which confers association with the complex and a C-terminal domain, structurally related to the SKI/SNO/DAC domain, which has both DNA and nucleosomal binding activity. The complex is tetrameric, containing four copies of each of its constituent proteins (Itoh et al., 2015).

### 1.3.5 LSD1-CoREST

The core CoREST complex consists of a single copy of LSD1, CoREST1/2/3, and HDAC1/2 (Song et al., 2020). CoREST contains two 50 amino acid SANT domains and associates with HDAC1/2 via the N-terminal SANT domain (You et al., 2001). The C-terminal region of CoREST contains the second SANT domain, which was previously thought to facilitate binding to chromatin, as deletions of this SANT domain leads to loss of the ability of LSD1 to demethylate nucleosomal substrates (Shi et al., 2005). However, structural studies of the complex when both LSD1 and HDAC1 are bound have revealed that this is unlikely to be the case, as the SANT2 domain is positioned in close proximity to HDAC1 (Song et al., 2020). CoREST binds LSD1 through its linker-SANT2 region, via the Tower domain of LSD1 (Figure 1.8) (Yang et al., 2006). Differences in the particular CoREST paralogue associated with the complex can serve to alter its activity. In contrast to CoREST1 and CoREST2, association of CoREST3 with the complex inhibits demethylation of nucleosomes by LSD1 (Upadhyay et al., 2014). CoREST2 has a proposed reduced

interaction with HDAC1/2, compared to CoREST1 and CoREST3, due to a singular residue difference in the first SANT domain, and it has therefore been suggested that HDAC1/2 is found in the LSD1-CoREST2 complex less often (Barrios et al., 2014). However, BioID experiments investigating HDAC1 proximal proteins have determined CoREST1, CoREST2, and CoREST3 all to be in close proximity to HDAC1 (Barnes et al., 2022). LSD1 is required for the stability of the CoREST complex, therefore loss of LSD1 not only leads to loss of demethylase activity but some deacetylase activity, with observed increases in H3K9ac and H3K56ac upon loss of LSD1 (Foster et al., 2010).

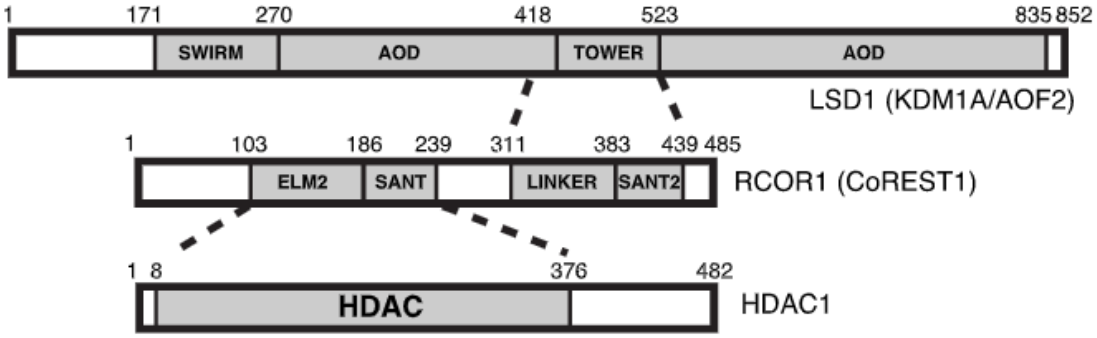


***Figure 1.8 Domains of CoREST complex components.*** *Schematic showing the domains of LSD1, CoREST1 and HDAC1. Dashed lines indiate domain interactions. Taken from (Song et al., 2020).*

#### 1.3.5.1 The diverse functions of LSD1-CoREST are dependent on its interactors

The CoREST complex serves as a tool which can be recruited by a number of proteins and complexes to chromatin to mediate a repressive effect on target genes. The complex was initially identified as a cofactor of the REST/NRSF (RE1 silencing transcription factor/neural-restrictive silencing factor), which plays a key role in repressing neuronal genes in non-neuronal cells (Andrés et al., 1999). The core CoREST complex can be recruited as a module of the CtBP corepressor complex (Hayakawa and Nakayama, 2011, Shi et al., 2003). ZNF516 then recruits the CtBP complex to *EGFR*, which plays a role in cell proliferation and epithelial-to-mesenchymal transition (EMT), to mediate repression of this gene (Li et al., 2017). Another zinc finger protein, ZNF217, can bind at promoters, for example at the promoter of *BRCA1*, and recruit repressive proteins, including the CoREST complex (Banck et al., 2009). SNAI1, a key player in EMT, directs the repressive activity of the CoREST complex to the *E-cadherin* promoter. SNAI1

interacts with the amine oxidase-like domain (AOD) of LSD1 via its SNAG domain, which mimics the structure of histone H3 tail (Lin et al., 2010b). The CoREST complex is also recruited by the SNAG domains of the zinc finger proteins GFI1 and GFI1B to their target genes (Saleque et al., 2007). This process is critical for repressing genes in the haemogenic endothelium during the endothelial-to-haemopoietic transition (EHT), which is required for the generation of haemopoietic stem cells (Thambyrajah et al., 2016). For example, GFI1B recruits the repressive activity of CoREST to *meis1* in erythroid cell lineages and loss of the complex, or of GFI1B, leads to upregulation of *meis1* (Chowdhury et al., 2013). Treatment of AML (acute myeloid leukemia) cells with an LSD1 inhibitor showed that the resultant facilitation of differentiation was not due to the loss of catalytic activity of the enzyme, but rather the disruption of the interaction of LSD1/CoREST with GFI1 (Maiques-Diaz et al., 2018). This was reflected in the lack of accumulation of H3K4me1/2 marks at LSD1 bound promoters and enhancers, the ability to rescue the effects of LSD1 depletion with a catalytically dead mutant, K661A, as well as the prevention of the inhibitor activity with a LSD1-GFI1 fusion protein. Another SNAG domain containing transcription factor is INSM1, which recruits the CoREST complex to repress specific genes in the pituitary, which allows differentiation of endocrine cells (Welcker et al., 2013). BHC80, a PHD finger containing protein, in complex with CoREST preferentially binds unmethylated H3K4 and the demethylation of H3K4me2 by LSD1 is required for this interaction (Lan et al., 2007). BHC80 is required for gene repression by LSD1, through sustaining its presence at target genes and preventing re-methylation of H3K4. RACK7 is present at superenhancers and associates with LSD1 and KDM5C, preventing overactivation of enhancers through both H3K4me3 and H3K4me1/2 demethylase activity (Shen et al., 2016).

#### 1.3.5.2 Substrates of LSD1-CoREST

Hypoacetylated histones are the preferred substrate for LSD1, suggesting the deacetylase activity by the HDAC component of the complex precedes the demethylase activity of LSD1 (Shi et al., 2005). This is supported by the increased demethylase activity of the LSD1-CoREST complex when also complexed with HDAC1, and the counteraction

of this increase by the addition of HDAC inhibitors (Lee et al., 2006). Specifically, CoREST mediated demethylation of H3K4me2 on substrates acetylated at H3K9, H3K14 or H3K18 was markedly reduced by inhibition of HDAC1, particularly H3K14ac (Wu et al., 2018). Interestingly, acetylation of H3K14 is also the slowest of these marks to be deacetylated by CoREST, suggesting that histone tails with this mark present are resistant to CoREST modifying activities. Though this data suggests that HDAC activity likely precedes LSD1 activity, the full picture may be more complicated. The activities of LSD1 and HDAC1 within CoREST are closely linked, with K4 methylation affecting deacetylase activity and K9 acetylation affecting demethylase activity, though binding of substrate by HDAC1 or LSD1 is mutually exclusive (Song et al., 2020).

## 1.4 LSD1

### 1.4.1 LSD1 structure

LSD1 (BHC110/KIAA0601/p110b/AOF2/KDM1) is a 110 kDa protein which consists of an N-Terminal unstructured region, a SWIRM domain and a C-terminal amine oxidase-like domain (AOD), which itself contains a Tower domain, an antiparallel coiled-coil which extends out from the main body of the protein (Figure 1.9) (Chen et al., 2006). The SWIRM and AOL domains interact with each other to form a globular structure, which the Tower domain protrudes from. Mutations that weaken the interaction between these two domains reduce the catalytic activity of LSD1 (Stavropoulos et al., 2006). This is potentially due to interruption of a cleft formed between these domains which may provide further interactions with the H3 peptide (Chen et al., 2006, Stavropoulos et al., 2006, Tochio et al., 2006). The conformation of the complex changes upon substrate binding, with the AOD rotating in respect to the Tower domain in a clamping motion (Baron and Vellore, 2012). Despite SWIRM domains present in other histone-modifying complexes possessing DNA binding activity, the LSD1 SWIRM domain lacks the positively charged region which confers this activity, suggesting it is not directly involved in DNA binding (Stavropoulos et al., 2006).

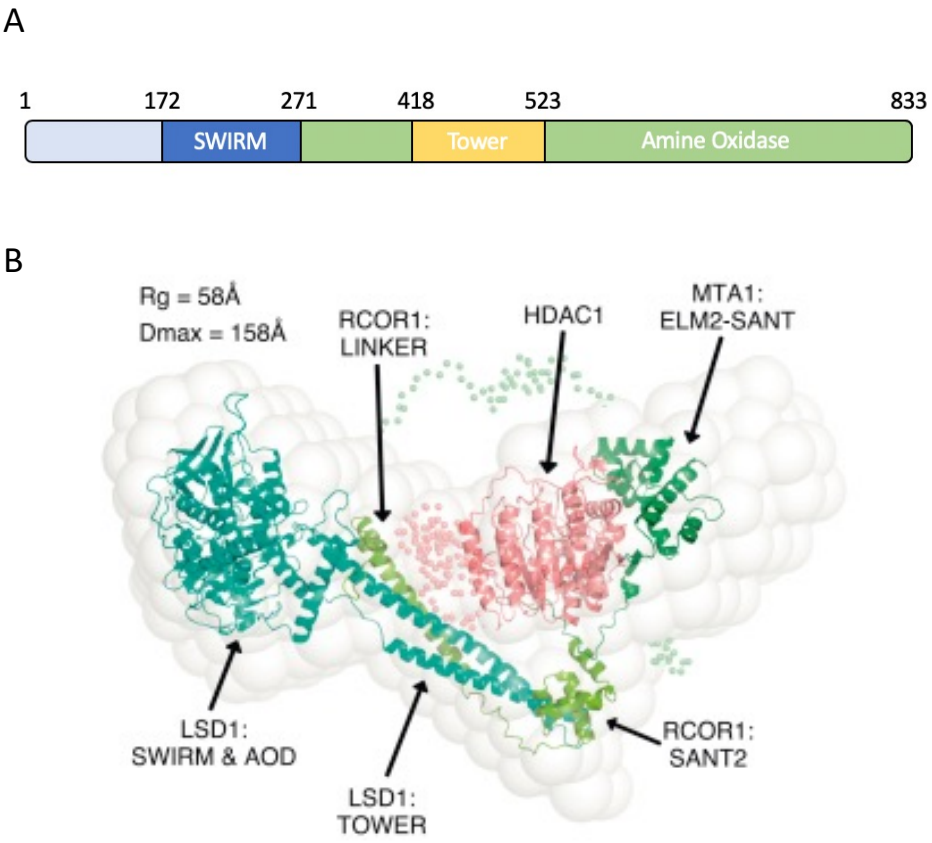


***Figure 1.9 Structure of LSD1within the CoREST ternary complex.*** *(**A**) Schematic of LSD1 domains. (**B**) CoREST ternary structure fitted using crystal structures of MTA1:HDAC1 and LSD1:RCOR1 complexes, taken from (Song et al., 2020).*

The AOD is split into two subdomains: the FAD-binding subdomain and the substrate binding subdomain, and the active site is present in the large cavity formed between these subdomains (Chen et al., 2006, Stavropoulos et al., 2006). The AOD shows structural similarity to the polyamine oxidase domain, though they differ in the size and shape of the catalytic cavity, reflecting the bulkier methylated histone substrate of LSD1 (Chen et al., 2006). The residues present on the surface of the catalytic cavity are important for activity, as these are conserved across species and multiple individual point mutations affect activity (Chen et al., 2006, Stavropoulos et al., 2006). In particular, a point mutation of K661A was previously shown to completely abolish LSD1 demethylase activity, as it plays an essential role in flavin reduction, and LSD1 with this mutation has been widely employed as a catalytically inactive version (Lee et al., 2005, Chen et al., 2006, Lee et al., 2006, Huang et al., 2007, Adamo et al., 2011, Maiques-Diaz et al., 2018, Sehrawat et al., 2018). Interactions between the active site residues and the H3 substrate are extensive, with the H3 tail adopting a compressed conformation

within the site (Yang et al., 2007). The AOD interacts with extranucleosomal DNA positioned 100Å away from the nucleosome (Kim et al., 2020).

### 1.4.2 Role of LSD1 in embryonic development

LSD1 has been shown to have an important role in development. Knockout of LSD1 has been shown to be embryonic lethal at approximately embryonic day (E) 6.5, showing it to be essential during development of the early embryo (Wang et al., 2007, Wang et al., 2009a, Foster et al., 2010). Knockout of LSD1 in embryonic stem cells (ESCs) in one study has revealed that these cells are stable in a proliferative state, however, upon differentiation, elevated levels of cell death are observed (Foster et al., 2010). This is demonstrated by the reduced number and size of embryoid bodies (EBs) lacking LSD1 compared to controls and increased cell death following retinoic acid stimulated differentiation (Foster et al., 2010). In contrast to this, another study found that knockout of LSD1 in ESCs led to cell death and defects in cell cycle progression when in a proliferative state (Wang et al., 2009a), however, both studies observed clear defects in differentiation.

LSD1 occupies enhancers of active genes in ESCs, together with master pluripotency transcription factors, OCT4, SOX2 and NANOG. Differentiation of ESCs in the absence of LSD1 activity showed that genes neighbouring LSD1 occupied enhancers failed to be repressed (Whyte et al., 2012). Loss of LSD1 upregulates a number of enhancers in ESCs, reflected in an increase in H3K4 methylation, H3K27 acetylation and enhancer RNA transcription (Agarwal et al., 2021). The interaction of OCT4 with LSD1 inhibits LSD1 activity at these enhancers, and overexpression of OCT4 has been shown to lead to retention of H3K4 methylation (AlAbdi et al., 2020)

Circumventing the embryonic lethality of LSD1 loss using conditional knockouts has revealed a number of roles of LSD1 throughout development. Conditional knockout of LSD1 in the developing pituitary gland showed that, although LSD1 was not required for its early morphology and cell fate determination, it was required for late and terminal

cell-lineage differentiation (Wang et al., 2007). Conditional deletion in spermatogonia showed LSD1 to be essential for spermatogenesis, evidenced by a complete loss of germ cells in adult males (Lambrot et al., 2015). Similarly, conditional knockouts of LSD1 in adult mice revealed an ongoing requirement for LSD1 in adult spermatogenesis (Myrick et al., 2017). Conditional knockouts in blood lineage specific cells revealed that LSD1 is required for both immature hematopoietic stem cell and mature hematopoietic cell differentiation (Kerenyi et al., 2013).

LSD1 has been implicated in haematopoiesis through association with hematopoietic transcriptional regulators including TAL1, GFI1/GFI1B, SALL4 (Saleque et al., 2007, Hu et al., 2009, Liu et al., 2013). LSD1 is recruited by GFI1/GFI1B to its target genes, resulting in transcriptional repression. Inhibition of LSD1 leads to derepression of these genes and disrupts differentiation of several hematopoietic lineages (Saleque et al., 2007). LSD1 associates with TAL1 at genes involved with erythroid differentiation, repressing this differentiation programme (Hu et al., 2009). SALL4 recruits LSD1 to genes in hematopoietic precursor cells, resulting in their repression, and knockdown of LSD1 results in altered expression of SALL4 target genes (Liu et al., 2013)

Differentiation of mesodermal lineages such as adipogenesis, myogenesis and osteogenesis appears to be intimately linked with LSD1 activity (Musri et al., 2010, Choi et al., 2010, Ge et al., 2014). LSD1 interacts with MYOD and MEF2 transcription factors on myogenic promoters and has been suggested to activate these genes by removing methylation of H3K9 (Choi et al., 2010). In muscle stem cells, LSD1 promotes differentiation into myocytes during muscle regeneration and loss of LSD1 not only slows this process but also redirects cell fate to brown adipogenesis (Tosic et al., 2018). LSD1 is required for adipogenesis, demethylating H3K9 at the promoter of the adipogenic transcription factor *cebpa*, resulting in its activation (Musri et al., 2010). LSD1 also promotes brown adipogenesis by demethylating H3K4 at Wnt pathway genes, leading to repression (Chen et al., 2016). LSD1 represses the osteogenic differentiation programme and inhibition of LSD1 in adipose-derived stem cells leads to promotion of osteogenic cell fate (Ge et al., 2014).

The LSD1-CoREST complex interacts with the transcriptional repressor REST to repress neuronal genes in non-neuronal cells (Hakimi et al., 2002). Proteosomal degradation of REST is required for the transition of ESCs to neural progenitors. However, throughout differentiation into cortical neurons the CoREST complex persists at RE1 sites, resulting in low levels of expression of related genes (Ballas et al., 2005). It has been shown that, in addition to genes targeted by both REST and CoREST, there are a selection of genes targeted solely by CoREST in neuronal subtypes, suggesting a role for CoREST in neurogenesis independent of REST activity (Abrajano et al., 2009). In neuronal stem cells, LSD1 is recruited to target genes by TLX where it represses the expression of cell proliferation regulators. Inhibition of LSD1 thereby leads to reduced proliferation of neural stem cells (Sun et al., 2010). During cerebral cortex development, LSD1-CoREST negatively regulates the Notch pathway, allowing expression of proneural genes and differentiation of neurons from neuronal stem cells (Lopez et al., 2016).

## 1.4 Mouse embryonic stem cells

Stem cells are defined by their ability to self-renew indefinitely, as well as their pluripotency, and they exhibit the ability to differentiate into cells from any of the three germ layers: endoderm, mesoderm, and ectoderm. Mouse ESCs are derived from the inner cell mass (ICM) of preimplantation blastocysts (Evans and Kaufman, 1981, Martin, 1981). These cells possess the ability to differentiate *in vitro* and to form teratocarcinomas when injected into mice. ESCs isolated in this way were shown to form chimeras when reintroduced into embryos, with the resulting embryos able to continue development and later produce offspring (Bradley and Robertson, 1986). The contribution of ESCs to chimeric germ cells allows introduction of genetic modifications into subsequent offspring, which can be exploited to produce mutants, including knockouts (Robertson et al., 1986, Thompson et al., 1989).

### 1.4.1 Maintenance of ESCs in culture

The original culture of ESCs relied on co-culture with mouse embryonic fibroblast (MEF) feeder cells in media supplemented with serum (Evans and Kaufman, 1981, Martin, 1981). Identification of a secreted protein responsible for the suppression of differentiation by feeder cells, Leukemia Inhibitory Factor (LIF), enabled culture of ESCs independently from feeder cells (Williams et al., 1988, Smith et al., 1988). The effects of LIF are mediated through the heterodimerisation of LIF receptor (LIF-R) and GP130, which leads to phosphorylation and activation of Janus-associated Kinases (JAKs). This results in activation of signal transducer and activator of transcription 3 (STAT3), which is an essential factor in LIF signal transduction, as inhibition of STAT3 blocks cell self-renewal and promotes differentiation (Niwa et al., 1998). In serum free cultures, LIF is not sufficient to maintain pluripotency alone, due to the absence of bone morphogenetic proteins (BMPs). BMPs act through activation of SMAD transcription factors, resulting in induction of Id genes, and the combination of BMP4 and LIF enhances self-renewal in the absence of serum (Ying et al., 2003). Ideally, culture of ESCs would give rise to a homogenous cell population, however, ESCs cultured in serum and LIF have been shown to be heterogenous, with subpopulations resembling epiblast and primitive ectoderm in marker expression (Toyooka et al., 2008, Hayashi et al., 2008).

More recently, culture of ESCs with chemical suppression of the MEK/ERK pathway and inhibition of glycogen synthase kinase 3 (GSK3) in combination with LIF, referred to as 2i+LIF, has presented a serum free technique (Ying et al., 2008). FGF4 activates MAPK/ERK kinase (MEK) which subsequently activates mitogen-activated protein kinases ERK1/2. Deletion or inhibition of components of this pathway restricts exit from pluripotency (Kunath et al., 2007). Wnt signalling promotes accumulation of β-catenin in the nucleus, where it acts as a transcription factor, activating transcription of downstream genes. GSK3 phosphorylates β-catenin, promoting its degradation and thereby negatively regulating Wnt signalling. Inhibition of GSK3 prevents this phosphorylation, allowing nuclear accumulation of β-catenin, which then interacts with TCF3 and relieves TCF3-mediated repression (Wray et al., 2011). 2i+LIF therefore

presents a more defined culture method that produces a homogenous population of self-renewing cells.

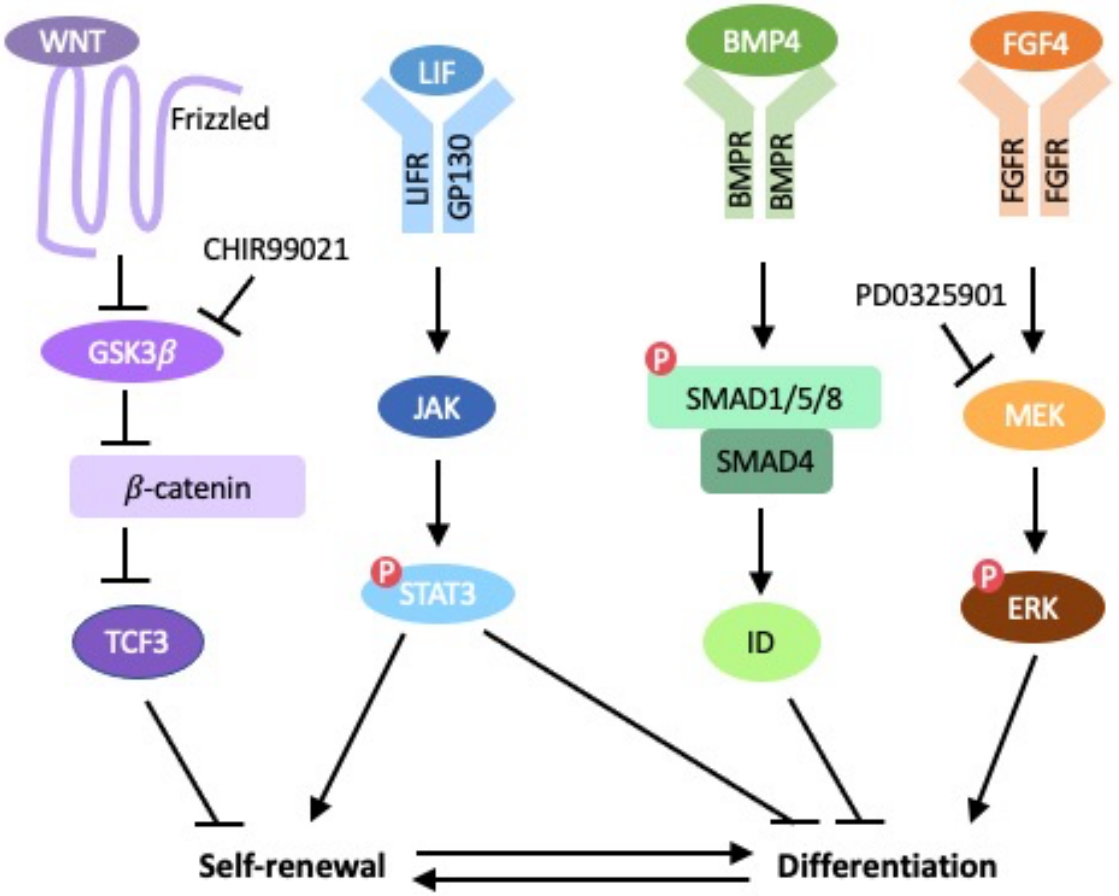


***Figure 1.10 Signalling pathways regulating ESC pluripotency.*** *Diagram of four key signalling pathways involved in the regulation of pluripotency in ESCs: LIF, Wnt, BMP and FGF, through promotion of self-renewal and inhibition of differentiation.The two inhibitors used in 2i+Lif media (CHIR99021 and PD0325901) are also shown. Arrows indicate activation and lines depict inhibition. Adapted from (Hassani et al., 2019, Weinberger et al., 2016).*

## 1.4.2 Pluripotency transcriptional network

ESCs are maintained in a pluripotent state by a network of transcription factors, at the core of which are OCT4, SOX2 and NANOG (Morey et al., 2015). OCT4, SOX2 and NANOG co-occupy a number of promoters of developmentally important genes, as well as their own promoters, resulting in autoregulatory loops (Boyer et al., 2005). The core factors regulate many of the same genes, activating transcription of pluripotency associated factors and repressing transcription of lineage-specific factors (Figure 1.11)

(Loh and Lim, 2011). A precise balance of OCT4 is required for pluripotency, as removal of OCT4 leads to loss of pluripotency and dedifferentiation into trophectoderm, while overexpression promotes differentiation into primitive endoderm and mesoderm (Niwa et al., 2000). Similar dedifferentiation into trophectoderm is seen upon deletion of SOX2, but this effect can be rescued by overexpression of OCT4 (Masui et al., 2007). This is in part because SOX2 maintains expression of OCT4 through regulating multiple transcription factors. ESCs deficient in NANOG lose pluripotency and differentiate into extraembryonic endoderm and exogenous expression of NANOG drives self-renewal of ESCs (Mitsui et al., 2003, Chambers et al., 2003). However, it has been shown that NANOG is not essential for the pluripotent state but instead helps to stabilise it, as ESCs are able to self-renew in the absence of NANOG, albeit with an impaired capacity and a predisposition for differentiation (Chambers et al., 2007). Aside from the core pluripotency factors, a number of other factors important for ESC pluripotency have been described, including KLF2/4/5 (Bourillot and Savatier, 2010), ESRRB (Festuccia et al., 2012), TFCP2l1 (Martello et al., 2013), TBX3 and DPPA4 (Ivanova et al., 2006).

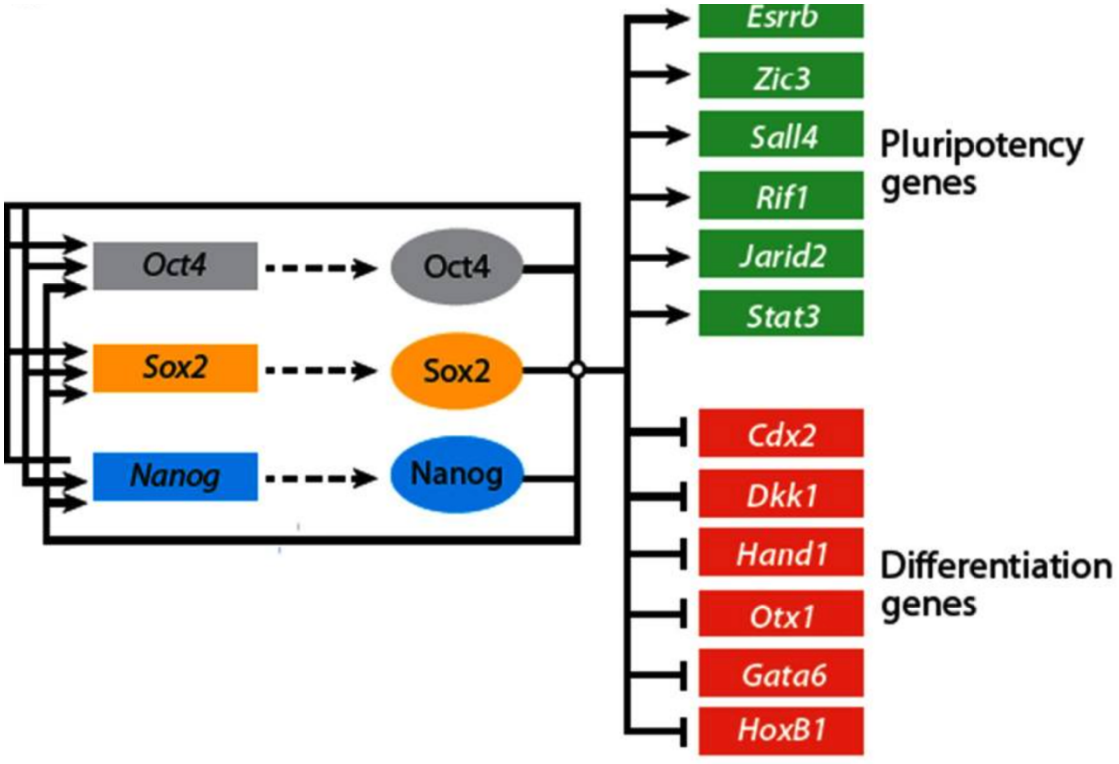


***Figure 1.11 Regulation of ESC state by key pluripotency factors.*** *Pluripotency factors OCT4, SOX2 and NANOG positively regulate their own expression and expression of other pluripotency genes, whilst inhibiting expression of differentiation genes. Taken from (Loh et al., 2011).*

## 1.5 Differentiation of mouse embryonic stem cells

### 1.5.1 Mouse embryogenesis

Mouse embryogenesis follows a highly coordinated series of events progressing from the single cell zygote (Figure 1.12). The single cell undergoes several divisions to form the preimplantation blastocyst, which consists of three cell types: the trophectoderm (TE), epiblast and primitive endoderm (PrE). Following implantation, extraembryonic and post-implantation epiblast tissues are formed from the epiblast while the primitive endoderm forms the visceral and parietal endoderm (Loebel et al., 2003). Gastrulation then proceeds with the formation of the primitive streak, where epiblast cells undergo epithelial-to-mesenchymal transition (EMT) and ingress to form mesoderm and definitive endoderm (Tam and Behringer, 1997). The posterior-to-anterior positioning of the primitive streak determines the mesodermal and endodermal lineages, with the posterior giving rise to the lateral plate mesoderm and subsequently the intermediate mesoderm, anterior paraxial mesoderm, axial mesoderm and then definitive endoderm forming from the anterior primitive streak (Bardot and Hadjantonakis, 2020). Cells that do not move through the primitive streak form the definitive ectoderm (Tam and Behringer, 1997).

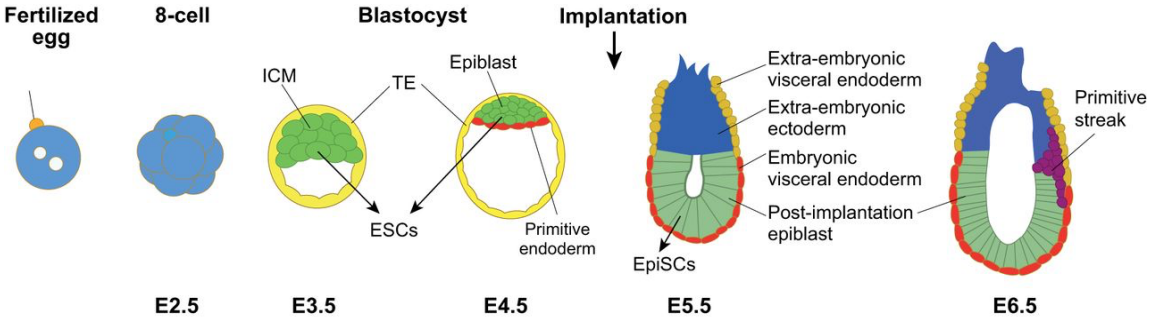


***Figure 1.12 Stages of early mouse embryogenesis.*** *Following fertilisation and cell division, the preimplantation blastocyst is formed at E3.5, consisting of the trophoectoderm (TE) and inner cell mass (ICM). Subsequently the ICM develops into the epiblast and primitive endoderm. Following implantation, the epiblast differentiates into multiple lineages before the primitive streak is formed at E6.5. Taken from (Davidson et al., 2015)*

### 1.5.2 Directed differentiation of ESCs

The pluripotent state of ESCs can be utilised in culture to model cell differentiation processes. ESCs are able to spontaneously form embryoid bodies (EBs) in culture in non-adherent conditions without the addition of extrinsic factors. Extended culture of EBs shows them to recapitulate some aspects of the developing embryo, with differentiation towards all three germ layers (Doetschman et al., 1985). More defined methods which include the addition of signalling factors enable directed differentiation of ESCs into different lineages.

Among these methods is the generation of epiblast-like stem cells (EpiLSCs) from ESCs cultured with a combination of Activin A and FGF2 (Hayashi et al., 2011, Buecker et al., 2014). Stem cells derived from the post implantation epiblast, termed epiblast stem cells (EpiSCs), can be isolated and maintained in culture. These EpiSCs represent a primed pluripotent state, expressing *Oct4* and *Nanog* alongside downregulation of *Rex1*, a marker present in ESCs, and upregulation of *Fgf5* and *Nodal*, which are expressed in late epiblast following implantation (Brons et al., 2007, Tesar et al., 2007). EpiSCs show a narrower capacity for pluripotency than ESCs, as chimera formation was impaired compared to ESCs upon injection into blastocysts (Brons et al., 2007).

More recently, gastruloids, which closely mimic the multiple germ layer organisation and morphology of gastrulating embryos, have been generated. This technique involves aggregation of ~300 ESCs in non-adherent conditions with a pulse of a GSK3 inhibitor, CHIR99021, between 48 hours and 72 hours post plating (Beccari et al., 2018b). The gastruloids establish the three major body axes and express *Hox* genes along the anterior-posterior axis. Gene expression analysis of gastruloids show expression of gastrulation associated genes at 48 hours and lineage specific markers associated with mesendoderm and neuroectoderm at 72 hours, reproducing spatial and temporal gene expression of the developing embryo (Beccari et al., 2018b). It is worth noting that despite abundant similarities between gastruloids and the developing embryo,

gastruloids do not express genes associated with extraembryonic structures and anterior neural plate lineages (Beccari et al., 2018b).

## 1.6 Project aims

LSD1 plays a role in numerous developmental pathways, including myogenesis, adipogenesis, osteogenesis, haematopoiesis, and spermatogenesis, which indicates its importance during development (Saleque et al., 2007, Musri et al., 2010, Choi et al., 2010, Ge et al., 2014, Lambrot et al., 2015). The embryonic lethality of LSD1 KO mice at ~E6.5 and the correspondence of this embryonic stage with the onset of early gastrulation indicates a role for LSD1 in this key developmental process (Wang et al., 2007, Wang et al., 2009a, Foster et al., 2010, Bardot and Hadjantonakis, 2020). Gastrulation is a highly coordinated and dynamic period which incorporates multiple processes, including EMT and axis formation, and the embryonic lethality of LSD1 KO during this period suggests disruption of one or more of these processes. This project aimed to elucidate the role of LSD1 during this developmental period by utilising gastruloids as a model system. We employed previously characterised *Lsd1*$^{lox/\Delta 3}$ ESCs (Foster et al., 2010), with a tamoxifen inducible *Lsd1* conditional knockout, to generate gastruloids and assess differential gene expression.

The requirement of LSD1 for the stability of the CoREST complex and loss of HDAC1/2 deacetylase activity upon loss of LSD1 (Foster et al., 2010), prompted us to distinguish between differential gene expression as a direct result of LSD1 demethylase activity and as an indirect result of loss of more generalised CoREST complex activity. To achieve this, we employed a catalytic inactive mutant of LSD1, K661A, which abrogates LSD1 demethylase activity whilst retaining CoREST complex stability (Lee et al., 2005, Chen et al., 2006, Lee et al., 2006, Huang et al., 2007, Adamo et al., 2011, Maiques-Diaz et al., 2018, Sehrawat et al., 2018), and produced gastruloids to assess differential gene expression.

# 2 Methods

## 2.1 Cell culture

Previously characterised *Lsd1*$^{lox/\Delta3}$ ESCs (Foster et al., 2010), with a tamoxifen inducible LSD1 conditional knockout, were used for all experiments. These cells were generated using E14 ESCs expressing a Cre/estrogen receptor fusion protein expressed from the *ROSA26* locus. A full schematic of the generation of these cells is given in Figure 2.1.

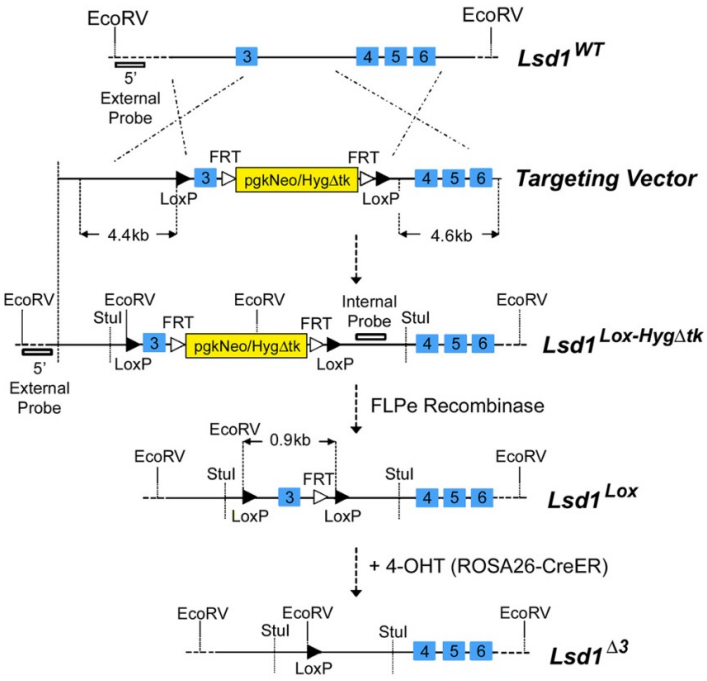


***Figure 2.1 Generation of conditional KO Lsd1$^{lox/\Delta3}$ ESCs.*** *A "Lsd1cKO-Hyg" gene targeting vector was used to flank Exon 3 of the Lsd1 gene with LoxP sites to produce Lsd1 $^{+/Lox\text{-}Hyg\Delta TK}$ cells. These were then treated with tamoxifen for 6h to induce LoxP recombination, generating Lsd1 $^{+/\Delta3}$ cells. Subsequently, the second allele of Lsd1 $^{+/\Delta3}$ cells was targeted with the Lsd1cKO-Hyg vector to produce Lsd1 $^{Lox\text{-}Hyg\Delta TK/\Delta3}$ cells and the selection cassette was removed using transient transfection with FLPe recombinase, generating the final Lsd1$^{lox/\Delta3}$ cells. Black triangles – LoxP positions, White triangles – FRT sites. Taken from (Foster et al., 2010).*

Previously generated *Lsd1*$^{lox/\Delta3}$ ESCs with *piggyBac* WT-*Lsd1* and *Lsd1*-K661A inserts were used for experiments in Results Chapter 5 (Barnes, 2018). Generation and screening of these cells for similar protein expression of inserts and endogenous LSD1 is outlined in Figure 2.2.

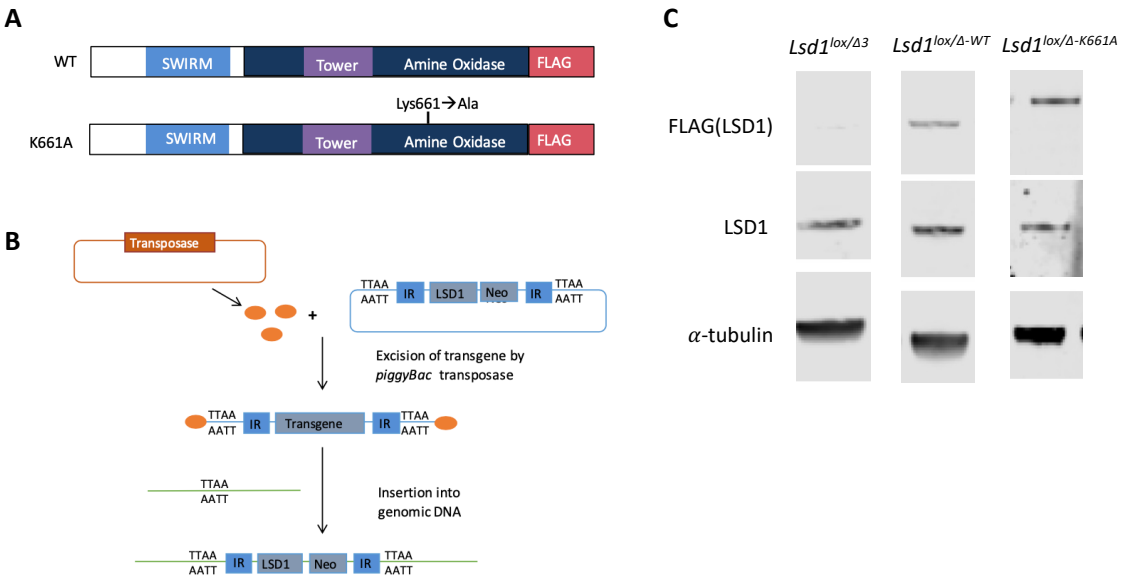


***Figure 2.2 Generation and characterisation of rescue construct cell lines.*** *A) Schematic of WT-Lsd1 and Lsd1-K661A inserts. B) Diagram of transposase insertion. DNA regions flanked by inverted repeats (IR) are excised by piggyBac transposase and then inserted into genomic DNA at any TTAA site. C) Screening of rescued Lsd1$^{lox/\Delta 3}$ cells protein expression levels. Western blots for FLAG-tagged mutant and WT LSD1 inserts in the clones used in this thesis using LSD1 and FLAG specific antibodies. α-tubulin was used as a loading control.*

Cells were either cultured in M15 + LIF or in 2i media, on 0.1% gelatin-coated 6-well/6cm plates. When switching cells from M15 + LIF to 2i media, or vice versa, cells were acclimatised to the new media for 2 weeks before being used in experiments.

## 2.1.1 Cell culture media and reagents

**M15 +LIF media**

500mL Knockout DMEM (GIBCO; 10829-018)

90mL Fetal Bovine Serum (LSP; S-001A-BR)

6mL Penicillin/Streptomycin/Glutamine-100x (GIBCO; 10378-016)

600µL 50mM b-mercaptoethanol (Fisher Scientific; 125472500)

25µL Leukaemia inhibitory factor (made in house)

**2i Stock**

500mL Knockout DMEM (GIBCO; 10829-018)

6mL Penicillin/Streptomycin/Glutamine-100x (GIBCO; 10378-016)

600µL 50mM b-mercaptoethanol (Fisher Scientific; 125472500)

25µL Leukaemia inhibitory factor (made in house)

**2i media**

50mL 2i Stock

1mL B27 (GIBCO; 17504044)

500μL N2 (GIBCO; 17502048)

6μL (3μM) CHIR99021- GSKi (Merck-Millipore; 25917-06-9)

5μL (1μM) PD032591- MEKi (Sigma; A3734)

**0.1% gelatin**

475mL Dulbecco's Phosphate Buffered Saline (GIBCO,14190144)

25mL 2% Gelatin from Porcine Skin (Sigma, G1890)

**TrypLE Express** (GIBCO; 12604021)

**Dulbecco's Phosphate Buffered Saline (PBS)** (Sigma; D8662-500ML)

**Freezing media**

4mL KO DMEM (GIBCO; 10829-018)

5mL Foetal Bovine Serum (LSP; S-001A-BR)

1mL DMSO (Sigma; D2650-100ML)

**Gastruloid Differentiation Medium (N2B27):**

25mL Neurobasal (Thermo Fisher Scientific, 21103-049)

25mL DMEM-F12 (Sigma-Aldrich, D6421)

250μL N2 (GIBCO; 17502048)

250μL B27 (GIBCO; 17504044)

500μL glutamine (GIBCO; 25030081)

50μL B-mercaptoethanol (Fisher Scientific; 125472500)

### 2.1.2 Thawing ESCs

10cm plates (Corning; 403167) were gelatinised with PBS containing 0.1% gelatin. Cryovials (Corning) containing cells were thawed rapidly and added to 9mL of prewarmed ESC media (M15+LIF/2i media). Cells were centrifuged at 1100rpm for 3

minutes and resuspended in ESC media (M15+LIF/2i media) before plating onto gelatinised plates. Plates were incubated in a 5% $CO_2$ incubator at 37°C.

### 2.1.3 Passaging ESCs

After plating, cells were passaged onto new plates every 2-3 days, once 75-90% confluency was reached. Passaging included aspirating media and washing cells in 1xPBS before dissociation from the plate using TrypLE. TrypLE was then neutralised using M15 + LIF media before centrifugation at 1100rpm for 3 minutes. The supernatant was aspirated from the resulting pellet, which was resuspended in fresh, prewarmed media before being split onto pre-gelatinised plates. Plates were maintained in a 5% $CO_2$ incubator at 37°C.

### 2.1.4 Freezing ESCs

ESCs were washed with PBS and disassociated from the plate using TrypLE, as described in 2.1.2. After centrifugation, cells were resuspended in 1mL freezing media and transferred to a cryovial. Cryovials were placed in freezing containers and placed in an -80°C freezer. After 24-48 hours, cryovials were transferred to liquid nitrogen.

### 2.1.5 Generation of Gastruloids

Gastruloids were generated following the protocol previously established by Beccari et al. (2018a). *Lsd1*$^{lox/\Delta 3}$ ESCs previously cultured in M15 + LIF media, and treated with OHT 6 days previous, were plated down onto ultra-low attachment 96 well plates (Corning 7007) in N2B27 media at a concentration of 7.5 ESCs/μL, alongside non-treated controls. At this stage, approximately 1.67 x $10^6$ cells from both conditions were taken for the 0-hour timepoint. Following incubation for 48 hours, 150μL of secondary media (N2B27 with 3μM CHIR99021) was added to each well. Following this, media was removed from each well and 150μL fresh media was added at 24h intervals until 120h. Images were taken at 48h, 72h, 96h and 120h timepoints.

For each replicate, there were a total of four plates of gastruloids each for *Lsd1* KO and controls. From these, two full plates and the outer perimeter of the remaining two plates were harvested at 72h, for a total of 216 gastruloids. This was due to the increased likelihood for gastruloids in the outer wells of the plates to become "ragged" after 72h. For the 120h timepoint 90 gastruloids, which were determined by eye to have elongated, were harvested from both *Lsd1* KO and control conditions.

For the rescue experiments in Results Chapter 5, $Lsd1^{lox/\Delta 3\text{-}WT}$ and $Lsd1^{lox/\Delta 3\text{-}K661A}$ ESCs were previously cultured in M15 + LIF media, and treated with OHT 6 days previous, to knock out endogenous expression of *Lsd1.* Gastruloids generated from WT-*Lsd1* and Lsd1-K661A rescue construct cell lines were produced as outlined for *Lsd1* KO and control gastruloids, with the same number of plates, number of gastruloids harvested, and following the same protocol.

#### 2.1.5.1 Optimisation and troubleshooting

Optimisation of media quality, initial cell number, and initial culture conditions was performed. Solutions to common issues are given in Table 2.1.

| Issue | Possible cause | Solution |
|---|---|---|
| Aggregates fail to form.<br>OR<br>Aggregates form but later begin to disaggregate. | Cell number is too low. | 300 cells per well is optimum. Check cell counting calculations and that cell pellet is properly resuspended before plating. |
| | Quality of media/pH of media. | Ensure N2B27 media is fresh, warm, and well-mixed.<br>Ensure the pH of media is appropriate. Overly pink/orange media could indicate incorrect incubator $CO_2$ concentration. |

| | | |
|---|---|---|
| | Quality of initial cell culture. | Ensure cell culture is not overgrown and has undergone as few passages as possible. Try thawing another cell aliquot and see if this improves it. |
| | Increased evaporation in outer wells. | When culturing gastruloids past 72h, avoid using wells in the outer perimeter of the plates, as cumulative evaporation can affect their growth. |
| Aggregates are irregularly shaped.<br>OR<br>Multiple aggregates form. | Cell number is too high. | 300 cells per well is optimum. Check cell counting calculations and that cell pellet is properly resuspended before plating. |
| | Brand of plate is not appropriate. | Ensure Corning 7007 ultra-low attachment plates are used, as other brands led to formation of multiple aggregates. |
| Aggregates fail to elongate. | Quality of media/pH of media. | Ensure N2B27 media is fresh, warm, and well-mixed.<br>Ensure the pH of media is appropriate. Overly pink/orange media could indicate incorrect incubator $CO_2$ concentration. |
| | Quality of initial cell culture. | Ensure cell culture is not overgrown and has undergone as few passages as possible. Try thawing another cell aliquot and see if this improves it. |

***Table 2.1 Solutions to common gastruloid issues***

### 2.1.6 Image acquisition

For both *Lsd1* KO versus controls and WT-*Lsd1* versus *Lsd1*-K661A experiments, five wells from each of the four plates for both conditions were randomly selected and images were taken with 20x magnification at 48h and 72h. Following harvesting at 72h, the same five wells in the remaining two plates imaged previously and five additional randomly selected wells for each plate were imaged at 10x magnification at 96h and 120h. Pixel measurements were taken using the measuring tool of GNU Image Manipulation Program (GIMP, version 2.10.34) and converted to μm using a scaling factor of 430 for 20x magnification and 215 for 10x magnification, calculated using images of a 1mm scale bar. For 48h, 72h, and 96h images, measurements were taken of the perpendicular x- and y-diameters of gastruloids, and area was calculated using $\pi \times x/2 \times y/2$. For 120h images, measurements were taken of the longest gastruloid axis.

## 2.2 RNA analysis

### 2.2.1 RNA extraction

For each replicate, from each condition, approximately 1.67 x $10^6$ cells, 216 gastruloids, and 90 gastruloids were taken for the 0-, 72-, and 120-hour timepoints, respectively. Gastruloids were harvested, pelleted and kept frozen at -80°C until simultaneous RNA extraction of samples from each timepoint, separately for each replicate. RNA extraction was performed using a Direct-zol RNA mini prep kit (ZYMO RESEARCH; R2053). When extracting RNA from >96-hour gastruloids, samples were first centrifuged after addition of TRI reagent for 2 mins in QIAshredder columns (QIAGEN; Cat No./ID: 79654). RNA was eluted in 20μL DNA/RNAse free water. RNA was stored at -80°C.

### 2.2.2 cDNA

cDNA was generated using 500ng of RNA in a 20μL reaction with 4μL q-Script (Qantabio; 95048) and using the below conditions. cDNA was diluted 5-fold with RNase/DNase free water and stored at -20°C.

**Reaction conditions:**

Lid temperature - 105°C

Step 1- 25°C 5 mins

Step 2- 42°C 30 mins

Step 3- 85°C 5 mins

Step 4- 12°C Infinite hold

### 2.2.3 RT-qPCR

RT-qPCR was performed in triplicate for each cDNA sample and primer combination. RT-qPCR reactions were set up in 96-well plates with 10μL volume consisting of: 0.5μL of 20mM primers, 5μL of SensiFAST SYBR, 2.5μL of DNA/RNAse free water and 2μL of cDNA. RT-qPCR was performed on a BioRad CFx Connect qPCR machine using the programme below.

**qPCR programme:**

95°C 15 mins

95°C 10 secs
55°C/60°C 30 secs } x 40 cycles
72°C 1 min

Plate read

65°C-95°C in 1°C increments (melt curve)

| Target gene | Primer | Sequence | Tm (°C) |
|---|---|---|---|
| *β-actin* | β-actin F | GGC TCC TAG CAC CAT GAA GA | 60 |
| | β-actin R | AGC TCA GTA ACA GTC CGC CT | |
| *Nanog* | Nanog F | AGG GTC TGC TAC TGA GAT GCT CTG | 60 |
| | Nanog R | CAA CCA CTG CTT TTT CTG CCA CCG | |
| *Sox2* | Sox2 F | TAG AGC TAG ACT CCG GGC GAT GA | 60 |
| | Sox2 R | TTG CCT TAA ACA AGA CCA CGA AA | |
| *Pou5f1 (Oct4)* | Oct4 F | AGT ATG AGG CTA CAG GGA CA | 60 |
| | Oct4 R | CAGT ATG AGG CTA CAG GGA CA | |
| *Brachyury* | Bra F | GCT TCA AGG AGC TAA CTA ACG AG | 55 |
| | Bra R | CCA GCA AGA AAG AGT ACA TGG C | |
| *Foxa2* | Foxa2 F | CCC TAC GCC AAC ATG AAC TCG | 55 |
| | Foxa2 R | GTT CTG CCG GTA G(g)**A**A AGG GA | |
| *Gsc* | Gsc F | CAG ATG CTG CCC TAC ATG AAC | 60 |
| | Gsc R | TCT GGG TAC TTC TGG | |
| *Gata6* | Gata6 F | TTG CTC CGG TAA CAG CAG TG | 60 |
| | Gata6 R | GTG GTC GCT TGT GTA GAA GGA | |
| *Zscan4a* | Zscan4a F | ACA GAT GCC AGT AGA CAC C | 60 |
| | Zscan4a R | ACA AAC AAC ATT CCC CTC TTC | |
| *Zscan4c* | Zscan4c F | CAG TAG ACA CCA CAC AAG ATA G | 60 |
| | Zscan4c R | ACA AAC AAC ATT CCC CTC TTC | |
| *Zscan4d* | Zscan4d F | AAA GAG GTG AGG TGG AGG AG | 60 |
| | Zscan4d R | GTG GTG ACA ATG GTG TGA AAG | |
| *Mixl1* | Mixl1 F | ACG CAG TGC TTT CCA AAC C | 60 |
| | Mixl1 R | CCC GCA AGT GGA TGT CTG G | |
| *Nkx1-2* | Nkx1-2 F | CGC TCT GCC CTA TCA GAC TTT | 60 |
| | Nkx1-2 R | GGC CCA AGG AAT GGA GTG A | |

***Table 2.2 RT-qPCR primers***

## 2.2.4 RNA-seq sample preparation

Quality of extracted RNA was checked using an Agilent RNA 6000 Nano Chip (Agilent; 5067-1511) on a 2100 Bioanalyzer (Agilent) to ensure that samples met the requirements for Novogene's sequencing service.

## 2.2.5 RNA-seq library preparation and sequencing

RNA samples were sent to Novogene for library preparation and sequencing. Libraries were prepared using the NEB Next Ultra RNA Library Prep kit (NEB; 37530). The Illumina NovaSeq 6000 sequencing system was used to perform 150bp paired-end sequencing

to a 20 million read depth per sample. 4 biological replicates for each condition were sequenced. Data was quality control checked before being released for analysis.

### 2.2.6 RNA-seq data analysis

The University of Leicester's SPECTRE High Performance Computing Facility was used to perform the initial data analysis. FastQC was used to check the quality of the FASTQ files and MultiQC was used to compile FastQC reports (Ewels et al., 2016). HISAT2 (version 2.1.0) was used to map reads to the mm10 genome, using default parameters (Kim et al., 2019a). Resultant SAM files were sorted and converted to BAM files using SAMtools (version 1.9) with default parameters, and SAMtools was subsequently used to index the BAM files (Li et al., 2009). BAM files were converted to bigWig files using deeptools bamCoverage and the resulting bigWig files were merged using the python package bigwigmerge. LiBiNorm (version 2.4) was used to generate read counts in htseq-count compatible mode using the Mus-musculusGRCm38.92 reference genome from Ensembl (Dyer et al., 2019). Count versions were loaded into RStudio running R version 4.2.2. DESeq2 was used for modelling and normalisation of the data, using 'normal' for LFC shrinkage (Love et al., 2014). Principal component analysis plots were used to ensure clustering between replicates within conditions and variance between conditions. Cook's distance plots, which use the distance of individual data points from the line of best fit to identify outliers, were then used to ensure that there were no outlying samples. Differential gene expression analysis was performed to compare samples and differentially expressed genes were determined using adjusted p-value (padj) of <0.01 and a Log2 fold change of > +1/< -1 as parameters.

Analysis and plot generation in R involved the following packages: DESeq2, basicPlotteR, ggplot2, RColorBrewer, ggrepel, ggpubr, PCAtools, dplyr, VennDiagram, topGO and tidyverse. The Bioconductor package topGO was used to perform Gene Ontology (GO) analysis on lists of differentially expressed genes (Alexa and Rahnenfuhrer, 2022). Fisher's exact test was used to determine significance (Fisher's <0.01)

# 3 Generation of gastruloids

## 3.1 Chapter Aims

The work in this chapter involves the generation of gastruloids and determination of their effectiveness as a model system to study the effects of *Lsd1* knockout in early development. The aim was to establish a model system which closely mirrors the developmental timepoint at which loss of LSD1 leads to embryonic lethality, ~E6.5, which correlates with early gastrulation (Wang et al., 2007, Wang et al., 2009a, Foster et al., 2010, Bardot and Hadjantonakis, 2020). To do this we produced gastruloids from previously characterised conditional knock out *Lsd1*$^{lox/\Delta 3}$ mouse ESCs (Foster et al, 2010) and analysed the levels of gene expression at timepoints over the course of 5 days (Figure 3.1A). We assessed gene expression in the control condition to confirm the functional relevance of the system. The timepoints chosen; 0 hours, 72 hours, and 120 hours, reflect ESC, early gastrulation, and late gastrulation stages, respectively, as has been previously demonstrated in gastruloids (Beccari et al., 2018b). Although it could be expected that differential gene expression which results in the embryonic lethality of *Lsd1* KO would be observed at the 72 hour timepoint, 120 hours was chosen as the endpoint to enable elucidation of any downstream gene expression changes in affected developmental pathways.

## 3.2 Optimisation of gastruloid generation

Following the previously established protocol for the generation of gastruloids by Beccari et al. (2018a), we were able to robustly and reliably produce gastruloids from the cell lines used. However, initial attempts to produce gastruloids highlighted a number of issues which required optimisation before this was achieved (Figure 3.1). Firstly, on multiple occasions, aggregates were either not forming or were disaggregating after the 72h timepoint, resulting in small, ragged aggregates as seen in Figure 3.1A. This issue occurred in multiple circumstances, including when cells had

undergone many passages (>20), when the number of cells plated was too low, or when media quality was impaired. The use of freshly thawed cells which allowed for fewer passages (<10) before plating and ensuring cell counts were accurate to allow for plating of approximately 300 cells per well rectified this issue in most cases. On one occasion, there was a sustained period of time for which aggregates were failing to form or later disaggregating whenever cells were plated, despite employing the countermeasures which had previously remedied this issue. Discolouration of the media indicated unsuitable pH levels, which were found to be caused by inaccurate $CO_2$ levels in the incubator. This was corrected by using an alternative incubator which accurately maintained $CO_2$ levels at 5%.

As indicated in Figure 3.1B, another issue in the generation of gastruloids which had to be overcome was the incidence of irregularly shaped, “amorphous” gastruloids. The formation of these gastruloids tended to occur when the plated cell number was too high. This was remedied by counting cells with a haemocytometer, as opposed to the previously used automated cell counter (Bio-Rad TC20), to ensure accuracy, and by ensuring cell pellets were sufficiently resuspended before plating. Occasionally, the plating of too many cells also resulted in the formation of multiple aggregates, as demonstrated in Figure 3.1C, and could be rectified by similar methods. In addition to this, multiple aggregates also formed when using a particular brand of low attachment 96 well plate (Greiner 650979) and this issue could be avoided by using a brand of plate which didn’t exhibit this issue (Corning 7007).

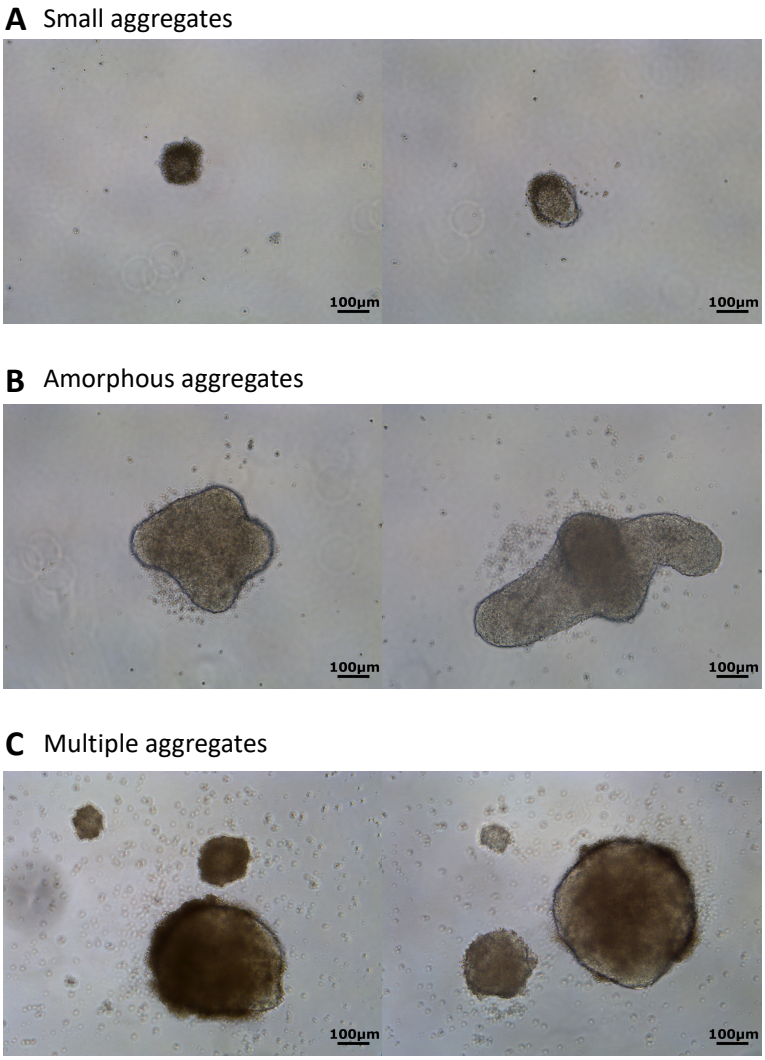


***Figure 3.1 Issues encountered during gastruloid generation.*** *Representative images at 96h of common issues. A) Small aggregates formed when the quality of initial cell culture or media was poor, or when too few cells were plated. B) Amorphous aggregates formed when too many cells were plated. C) Multiple aggregates formed when too many cells were plated, or a certain brand of plate was used.*

## 3.3 Gastruloid morphology mimics embryo elongation

Previous descriptions of gastruloid morphology showed that cells aggregate into spherical structures at early stages and then extend from 72 hours onwards to form elongated gastruloids at 120 hours (Beccari et al., 2018b) (Figure 3.2C). At 120 hours, the gastruloids display a cell dense rostral region and an extended caudal region. The results from our experiments exhibit similar morphology, though where elongation begins at 72 hours in previous experiments (Beccari et al., 2018b), the beginning of this process seemed to be delayed until 96 hours (Figure 3.2B).

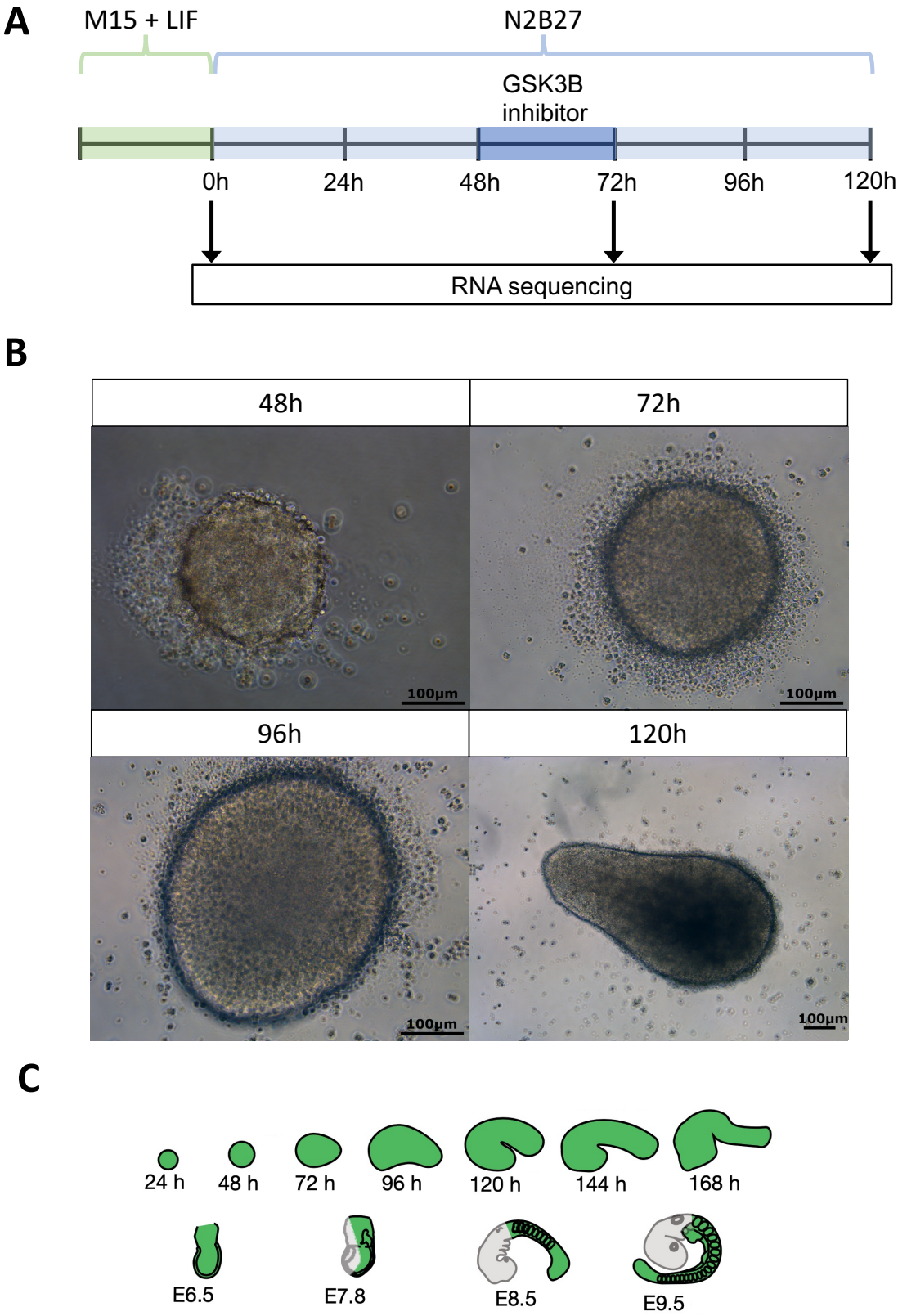


***Figure 3.2 Morphology of gastruloids mimics embryo elongation*** *(**A**) Diagram showing the timeline of gastruloid generation and harvest for RNA sequencing. (**B**) Representative images of gastruloids at 48-, 72-, 96- and 120-hour timepoints. (**C**) Diagram showing how gastruloid timepoints align with embryonic day stages in the embryo (Adapted from Beccari et al., 2018b)*

The delayed elongation prompted investigation into whether this corresponded with delayed induction of gastrulation associated markers. Using RT-qPCR analysis, it was determined that gastrulation markers were induced at 72 hours, as seen in previous studies (Figure 3.3) (Beccari et al., 2018b). Specifically, expression levels of the early differentiation marker *Brachyury* and gastrulation associated genes, *Mixl1*, *Gsc*, and *Nkx1-2* increased dramatically from 48 to 72 hours. This showed 72 hours was representative of early gastrulation and informed the choice of intervals used for RNA-seq. Despite the observed delay in elongation, these results indicate that gastruloids could be reliably produced using the previously established method.

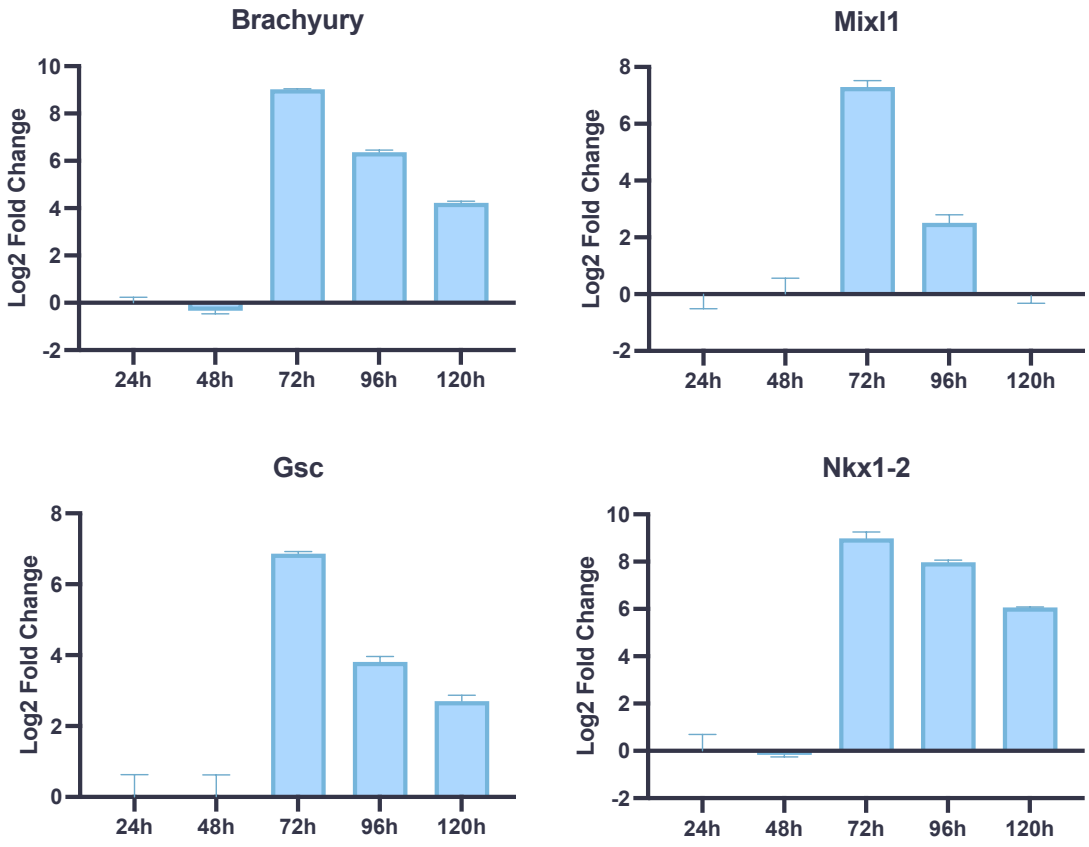


***Figure 3.3 Gastrulation associated genes were induced in gastruloids at 72 hours.*** *RT-qPCR data showing the fold-change in expression levels of Brachyury, Mixl1, Gsc, and Nkx1-2. Expression levels were normalised to the control gene b-actin (n=3 +/-SD).*

## 3.4 Gastruloids exit pluripotency and express markers of gastrulation

Gastruloids were harvested at 0, 72, and 120 hour timepoints for RNA-seq. Principal component analysis (PCA) was used to assess the variance of samples and samples clustered well by timepoint (Figure 3.4A). A plot of Cook's distance showed there were no outliers among the samples (Figure 3.4B).

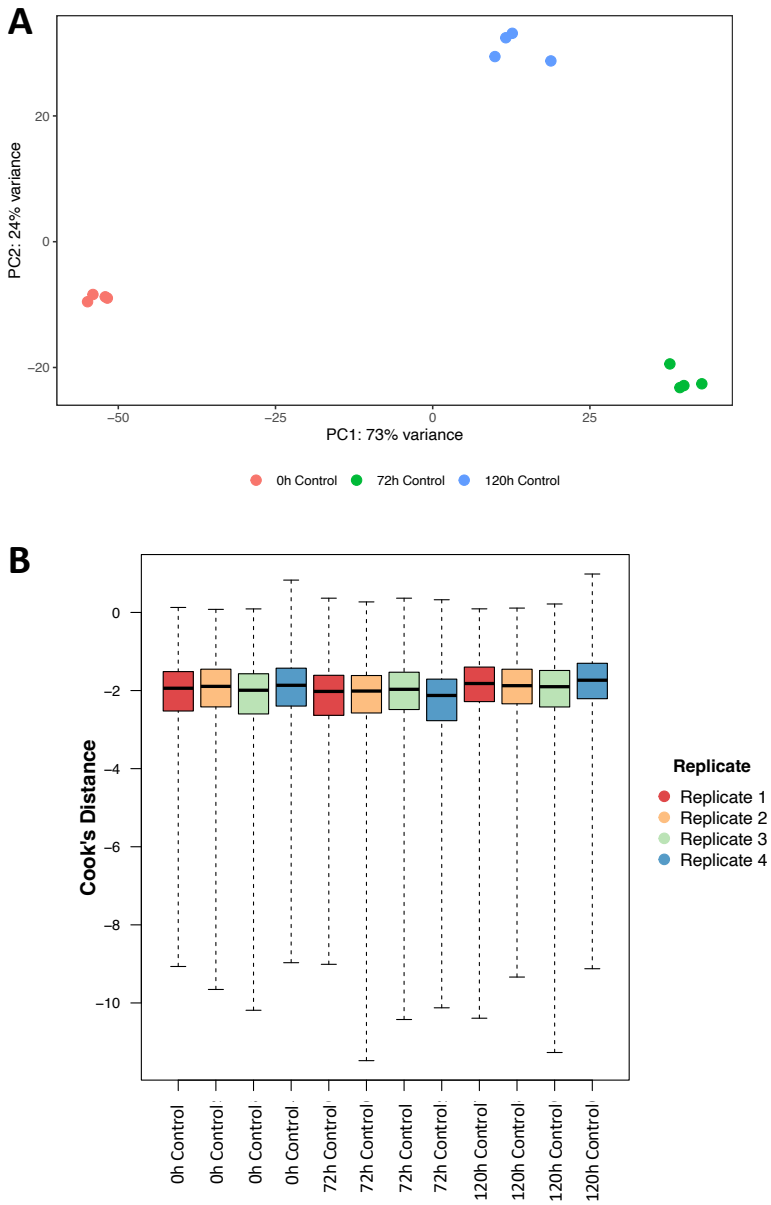


***Figure 3.4 Gastruloid samples were suitable for RNA-seq analysis.*** *(**A**) Plot of principal component analysis (PCA) showing the variance of control samples (**B**) Plot of Cook's distance showing there were no outliers among the samples.*

Using parameters of a fold change cut off of ≥2 and an adjusted p-value (padj) of ≤0.01, numerous differentially expressed genes (DEGs) were identified between timepoints, with 5554 differentially expressed genes identified between 0 hours and 72 hours, and 3668 genes between 72 hours and 120 hours (Figure 3.5). This is consistent with an alteration of gene expression as cells exit pluripotency and begin to differentiate.

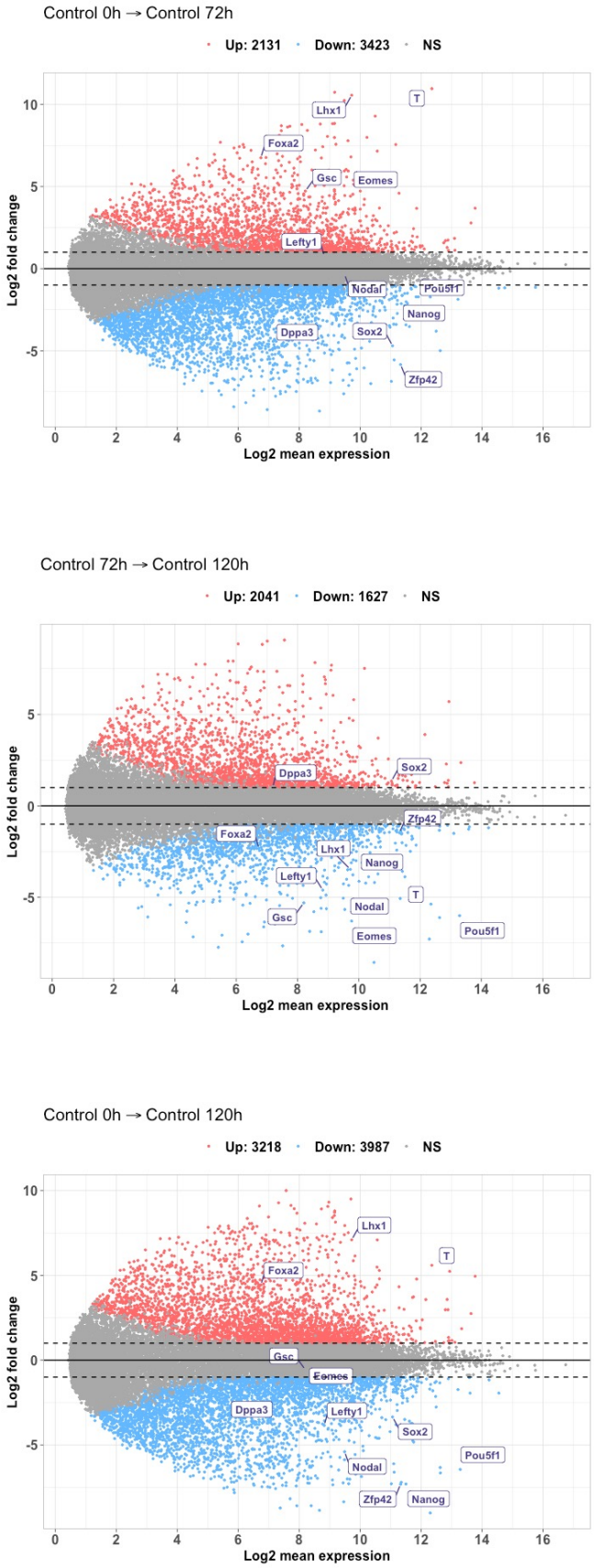


***Figure 3.5 Differential gene expression showed differentiation between timepoints.*** *MA plots showing genes that are differentially expressed between timepoints, using an adjusted p-value of <0.01 and a log2 fold change of >1. Upregulated genes are highlighted in red and downregulated genes in blue. Black dashed lines indicate the cut offs for log2 fold change.*

Gastruloids showed a reduction in pluripotency markers: *Nanog*, *Oct4 (Pou5f1)*, *Sox2*, *Rex1* (*Zfp42)*, and *Dppa3,* from 0 hours to 72 hours and these remained low at 120 hours (Figure 3.5). Reduction in the levels of the core pluripotency factors *Nanog, Oct4* and *Sox2* are required for the exit from pluripotency (Loh and Lim, 2011). *Rex1* expression is high in ESCs and reduces when cells are induced to differentiate (Rogers et al., 1991). *Dppa3* expression is associated with pluripotency, as cells expressing *Dppa3* resemble ICM, whereas *Dppa3* negative cells resemble epiblast (Hayashi et al., 2008). *Dppa3* shares a superenhancer with *Nanog*, which regulates the expression of both genes (Blinka et al., 2016).

Overall, a reduction in expression of genes associated with pluripotency and an increase in expression of genes associated with early differentiation was seen as gastruloids progress from 0 hours to 72 hours and 120 hours, as shown in Figure 3.6B. As pluripotency factors decrease in expression, levels of key early differentiation genes, *T*, *Foxa2*, and *Fgf5* increase, with a peak at 72 hours, representing a state primed for differentiation (Figure 3.6A).

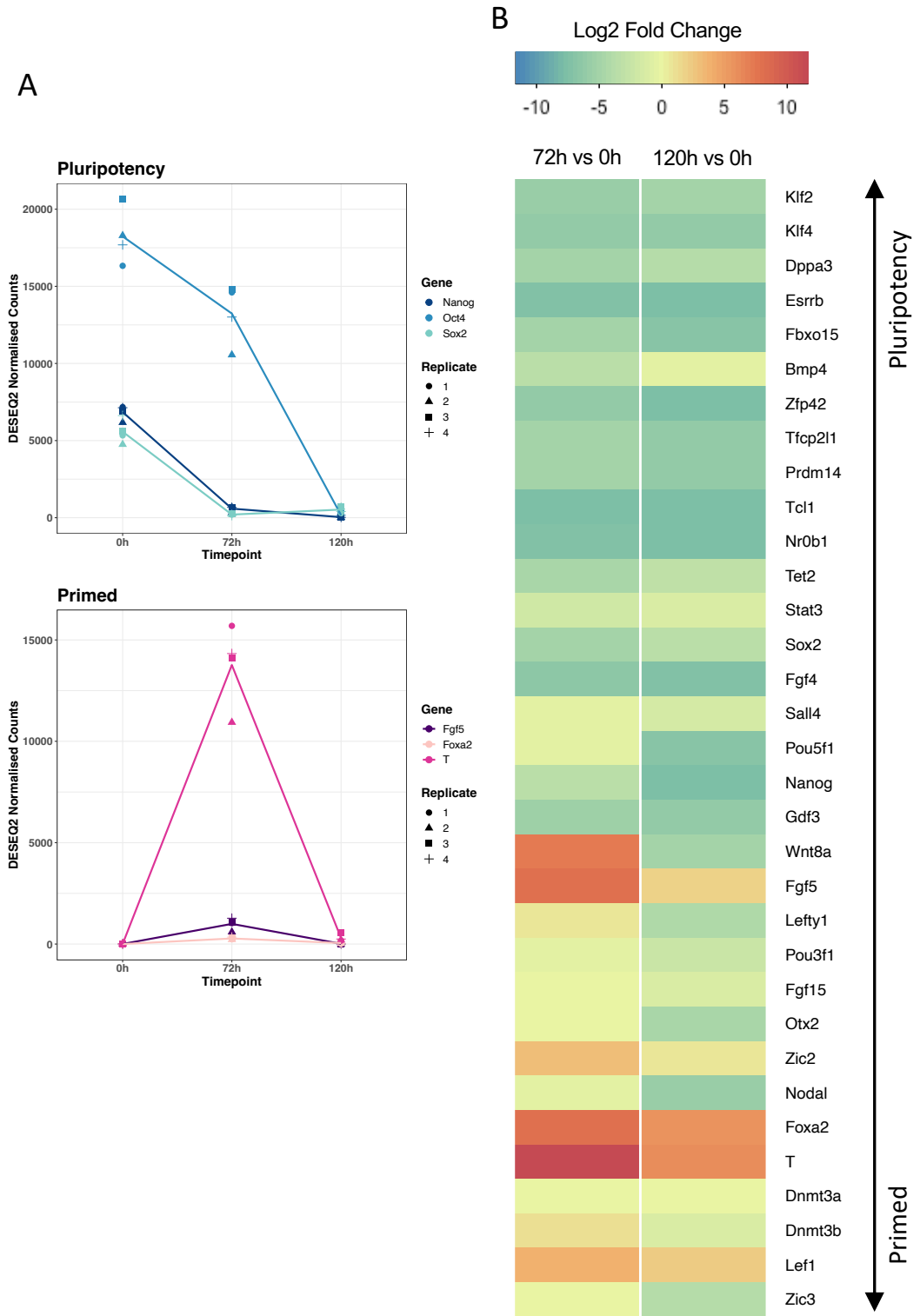


***Figure 3.6 Analysis of pluripotency markers shows an exit from pluripotency in control gastruloids*** *(**A**) Line graphs showing expression of genes associated with pluripotency and a state primed for differentiation at different timepoints. Points represent values for individual replicates for each gene, lines show mean values of replicates. (**B**) Heatmap showing log2 fold change values between 0 hours and 72 hours and between 0 hours and 120 hours for genes associated with pluripotency and a state primed for differentiation.*

Genes associated with gastrulation, *Brachyury (T)*, *Foxa2*, *Lhx1*, *Gsc* (*Goosecoid*), *Eomes* (*Eomesodermin*), and *Lefty1*, show a peak in upregulation at 72 hours (Figure 3.5). *Brachyury* and *Foxa2* are expressed in the posterior and anterior primitive streak, respectively, and their expression domains partially overlap (Bardot and Hadjantonakis, 2020). Cells expressing *Brachyury* form posterior mesoderm and *Foxa2* expressing cells form endoderm, while populations coexpressing these genes form axial mesoderm (Burtscher and Lickert, 2009). *Brachyury* expression in the primitive streak partially overlaps with expression of *Eomes*, and there is a requirement for either or both of these two factors for the determination of all mesoderm and definitive endodermal lineages. Cells deficient in both factors fail to form mesoderm and definitive endoderm, instead retaining primed pluripotency markers (Tosic et al., 2019). *Lhx1*, alongside *Mixl1*, has an important role in the movement of mesoderm and endoderm during gastrulation (Tam and Loebel, 2007). *Lefty1* is expressed in ESCs, with KLF4 cooperating with OCT4 and SOX2 at its promoter to activate its expression (Nakatake et al., 2006). However, it is also expressed in the epiblast, where its antagonism of *Nodal*, alongside *Cer1*, is essential for the correct positioning of the primitive streak (Tam and Loebel, 2007). *Goosecoid* is expressed at E6.5 in the developing primitive streak and localises towards its anterior end before contributing to an area of mesoderm which forms the head process. *Goosecoid* expression peaks over a very short window in embryos, with expression starting at 6 days, 9 hours and ending by 6 days, 21 hours, which corresponds with the expression peak seen in gastruloids at 72 hours (Figure 3.7A) (Blum et al., 1992).

This peak of expression in gastrulation markers at 72 hours is further reflected in Figure 3.7A, which highlights the sharp increase between 0 and 72 hours followed by a decrease between 72 and 120 hours of markers *Eomes*, *Fgf8*, *Gsc*, *Mixl1* and *Nkx1-2*. BMP and Wnt signalling are essential for the initiation of the primitive streak through activation of genes, including *Nkx1-2* (Sharma et al., 2017, Tamashiro et al., 2012). *Nkx1-2* is a transcriptional repressor and is essential for expression of *Brachyury* in the primitive streak, through repression of targets such as *Tcf3* (Tamashiro et al., 2012). *Fgf8* is expressed in the primitive streak, where it upregulates expression of *Fgf4*. Lack of *Fgf8* and/or *Fgf4* causes a failure of cells to migrate from the primitive streak and leads to complete loss of mesodermal and endodermal lineages (Sun et al., 1999). Upregulation

of *Fgf8* is partly mediated by *Brachyury*, as this upregulation is reduced upon knockdown of *Brachyury* (Tamashiro et al., 2012).

The expression of epiblast markers, *Nodal*, *Lefty1*, *Otx2* and *Fgf5*, in gastruloids aligns with their pre-gastrulation roles in the embryo (Figure 3.7A). *Nodal* is expressed in the epiblast pre-gastrulation and is required for the formation of the anterior visceral endoderm (AVE), one of the first steps in defining the anterior-posterior axis in the embryo. The AVE expresses the NODAL antagonists LEFTY1 and CER1, which restrict NODAL signalling to the proximal and posterior side of the embryo. This results in localisation of NODAL at the onset of gastrulation to the site which gives rise to the primitive streak, and this is essential for its formation (Whitman, 2001, Bardot and Hadjantonakis, 2020). OTX2 is required for the transition from the ES cell state to EpiSCs and prevents epiblast cells from regaining pluripotency by directly repressing *Oct4* and *Nanog* (Acampora et al., 2013, Di Giovannantonio et al., 2021). OTX2 is required for the anterior migration of the visceral endoderm to form the AVE and is essential for the downstream expression of LEFTY1 (Perea-Gomez et al., 2001). The expression of *Nodal*, *Lefty1* and *Otx2* seen at 0 and 72 hour correlates with their function in the pre-gastrulation epiblast (Figure 3.7A). The minimal expression of these epiblast related factors at 120 hours reflects an exit from a state representative of epiblast at the onset of gastrulation and entry into lineage specification.

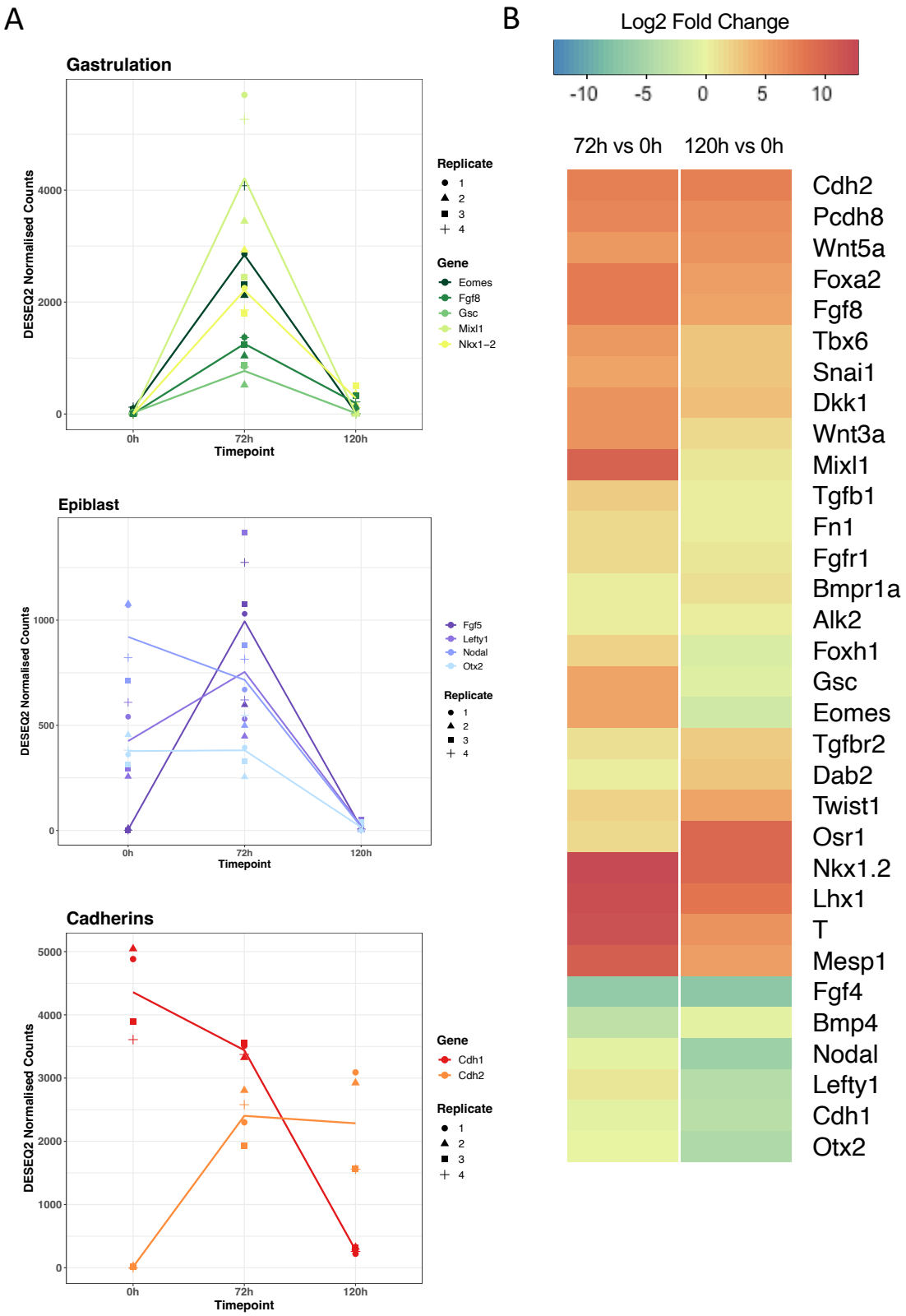


***Figure 3.7 Analysis of gastrulation associated genes shows induction of gastrulation in control gastruloids*** *(**A**) Line graphs showing expression of genes associated with gastrulation, epiblast and cadherins at different timepoints. Points represent values for individual replicates for each gene, lines show mean values of replicates. (**B**) Heatmap showing log2 fold change values between 0 hours and 72 hours and between 0 hours and 120 hours for genes associated with gastrulation.*

The epithelial-to-mesenchymal transition that occurs at the primitive streak is essential for the migration of cells that contribute to mesodermal and endodermal lineages (Tam and Loebel, 2007). *E-cadherin* (*Cdh1*) confers an epithelial state to cells through maintenance of cell-cell adhesions and its downregulation, through *Snai1* mediated repression, is essential for EMT (Cano et al., 2000). As cells undergo EMT, they encounter a switch from *E-cadherin* to *N-cadherin* (*Cdh2*) expression, and initial *N-cadherin* expression occurs at E7.5 in mesoderm emerging from the primitive streak (Radice et al., 1997). Corresponding with this, we observed downregulation of *E-cadherin* from 0 hours through to 120 hours and an upregulation of *N-cadherin* at 72 hours which is maintained at 120 hours (Figure 3.7A). This is consistent with EMT that occurs at the primitive streak in early gastrulation in the embryo.

## 3.5 Lineage specification in late stage gastruloids

Gastruloids have previously been shown to express lineage specific genes representing all three germ layers in spatial domains from 72 hours, though this gene expression does not lead to formation of defined structures such as the notochord, as seen in the embryo (Beccari et al., 2018b). This is consistent with results shown in Figure 3.8A, as induction of endodermal and early mesodermal markers can be seen at 72 hours. Expression of the endodermal determinant *Foxa2*, which acts upstream of multiple endodermal markers, peaks at 72 hours. Expression of endodermal markers, *Gata4, Mmp14* and *Sox17*, increases through 72 hours and 120 hours. *Gata4* is a direct target of *Foxa2* and plays a role in the specification of early endoderm (Rojas et al., 2010). The metalloproteinase gene *Mmp14*, involved in basement membrane remodelling, is expressed in *Foxa2* positive definitive endoderm (Scheibner et al., 2021). *Sox17* is a determinant of definitive endoderm and its depletion leads to loss of gut endoderm (Kanai-Azuma et al., 2002).

Upregulation of genes involved in the determination of early mesoderm, *Fgf8, Wnt3a, Tbx6, Brachyury (T)* and *Msgn1*, is observed at 72 hours in gastruloids (Figure 3.8A). *Wnt3a*, *Tbx6* and *Msgn1* are specifically expressed in the primitive streak where they are critical for the formation of paraxial mesoderm (Nowotschin et al., 2012b). *Brachyury* is a direct target of *Wnt3a* and acts upstream of *Tbx6* and so is essential in this pathway (Yamaguchi et al., 1999, Hofmann et al., 2004). *Fgf8* establishes the position of the somite determination front in presomitic mesoderm, where segmentation will occur (Dubrulle et al., 2001). Expression of these genes peak at 72 hours, reflecting their role in establishment of early mesoderm, and reduce at 120 hours, suggesting further differentiation into more specific lineages, including somites.

Genes related to somite formation are upregulated at 120 hours, suggesting formation of later, more defined mesodermal linages (Figure 3.8A). *Meox1* is expressed in presomitic mesoderm and in differentiating somites (Candia et al., 1992), which is reflected in its expression at 72 hours which is then upregulated at 120 hours. *Tcf15* (*paraxis*) is first expressed in paraxial mesoderm and then at higher levels in newly formed somites (Burgess et al., 1995), aligning with results in gastruloids. *Tcf15* is also expressed in a subpopulation of ESCs, priming them for differentiation (Davies et al., 2013), which may explain its additional expression at 0 hours. *Pax3* is expressed in newly formed somites and its expression is later restricted to the dermomyotome (Goulding et al., 1994). *Foxc1* is essential for the formation of somites, as loss of *Foxc1*, and the closely related *Foxc2*, leads to complete lack of somite formation (Kume et al., 2001).

Ectodermal gene expression is also established at 120 hours, but expression levels are not as high as mesodermal and endodermal genes (Figure 3.8A). This observed reduced induction of neuroectodermal markers coincides with the specification of neuroectoderm during late gastrulation (Bardot and Hadjantonakis, 2020). *Sox1* and *Pax6* are markers of neuroectoderm and their overexpression can induce neuronal lineages in embryoid bodies (Suter et al., 2009). *Pax6* expression was previously shown to be lower at 120 hours in gastruloids than at the equivalent stage in embryos, and genes associated with anterior neural plate were poorly expressed, suggesting ectoderm lineages are underrepresented in gastruloids (Beccari et al., 2018b).

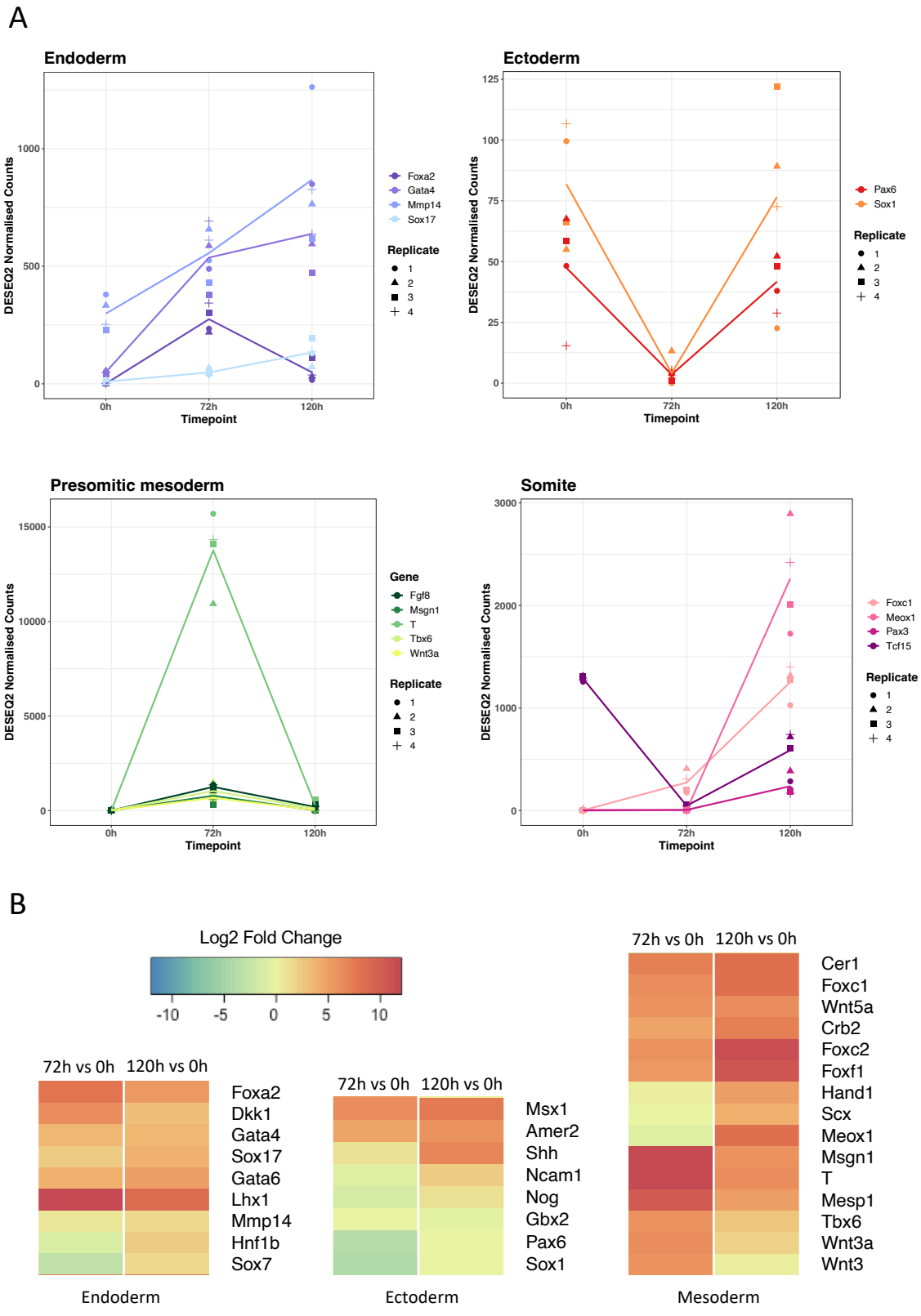


***Figure 3.8 Analysis of lineage specific genes indicates differentiation into all three germ layers in control gastruloids*** *(**A**) Line graphs showing expression of genes associated with endodermal, ectodermal, presomitic mesoderm and somites at different timepoints. Points represent values for individual replicates for each gene, lines show mean values of replicates. (**B**) Heatmap showing log2 fold change values between 0 hours and 72 hours and between 0 hours and 120 hours for genes associated with each germ layer.*

Transcription of Hox gene clusters were previously observed in gastruloids from 72 hours (Beccari et al., 2018b). Homeobox-containing *Hox* genes play a key role in establishing anterior-posterior patterning during embryogenesis. *Hox* genes consist of four paralogue groups which exist in chromosomal clusters: *Hoxa*, *Hoxb*, *Hoxc* and *Hoxd*, and genes are expressed in a spatio-temporal order according to their 3' to 5' positioning within these clusters (Figure 3.9A) (Deschamps and van Nes, 2005). For example, 3' *Hox2* genes are expressed earlier in the posterior primitive streak, whereas *Hox4* and *Hox9* genes, positioned more towards the 5' end of the cluster, are expressed later (Figure 3.9A). This sequential activation of *Hox* transcripts was previously observed in gastruloids. Specifically, expression of *Hoxa1* and *Hoxa3* was seen at 72 hours, followed by transcription of *Hoxa5/7/9* at 96 and 120 hours (Beccari et al., 2018b). Similarly, we see induction of 3' located *Hox* gene transcription at 72 hours, which are then upregulated further at 120 hours, followed by induction of more 5' positioned genes at 120 hours (Figure 3.8B). In previous studies, expression of *Hoxa10* and *Hoxa11* did not occur until 144 hours, when a downregulation of H*oxa1/2/3* was also seen (Beccari et al., 2018b). Expression of these transcripts were minimal at 120 hours in the current work, and downregulation of early *Hox* genes was not observed (Figure 3.8B), owing to the fact gastruloids were not cultured to the 144 hour timepoint. These results are indicative of progressive transcription of *Hox* gene clusters in an order corresponding to the 3' to 5' positioning of the genes.

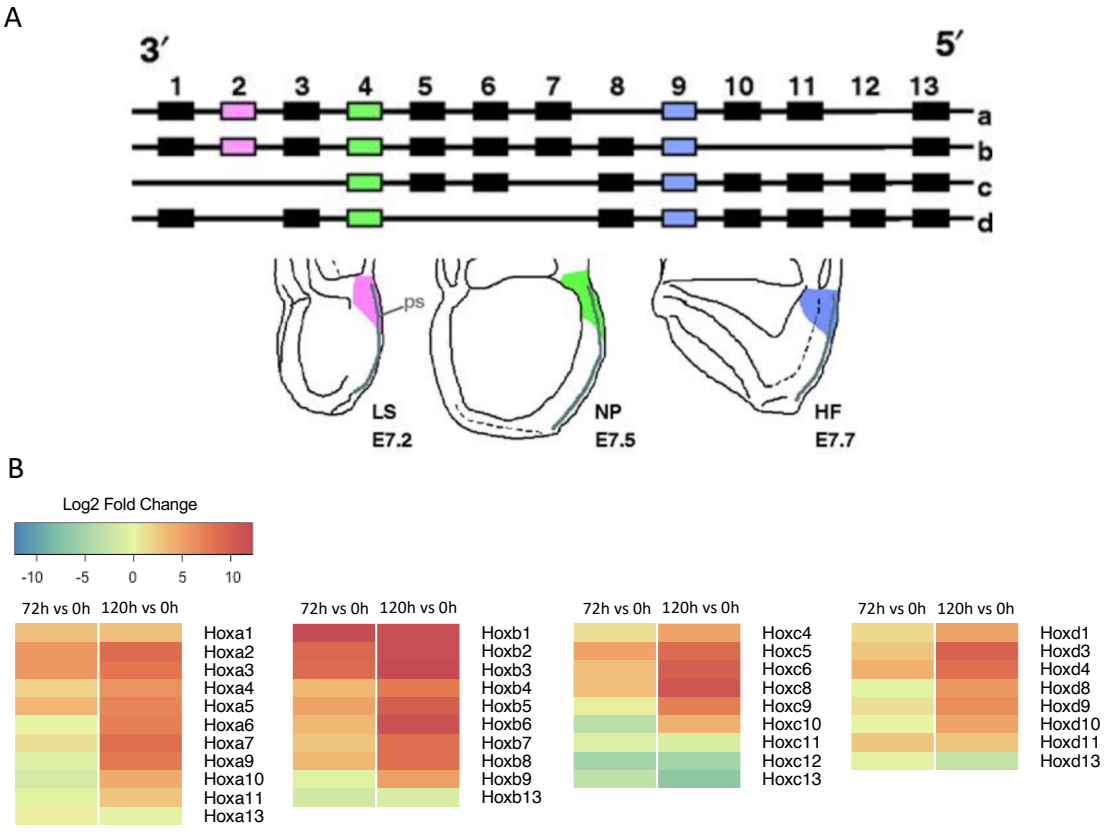


***Figure 3.9 Analysis of Hox genes reflects their progressive temporal expression in control gastruloids*** ***(A)*** *Schematic showing the arrangement of the four Hox gene clusters (a to d) on DNA and the temporal expression patterns of three paralog groups; 2, 4 and 9 at the late streak (LS), neural plate (NP) and head fold (HF) stages of mouse embryos (Adapted from Deschamps and van Nes, 2005)* ***(B)*** *Heatmap showing log2 fold change values between 0 hours and 72 hours and between 0 hours and 120 hours for Hox genes, separated into each Hox gene cluster.*

## 3.6 Summary

The results in this chapter have shown the successful generation of gastruloids that closely resembled gastruloids from previous studies and that recapitulated aspects of early embryogenesis. Gastruloids showed a decrease in pluripotency markers and an increase in genes associated with gastrulation, indicating differentiation over the timepoints. The expression of gastrulation associated genes and the subsequent expression of markers of all three germ layers suggests this model system emulates embryonic gastrulation processes, including EMT, as evidenced by the switch in cadherin expression. Overall, these results highlight the appropriateness of gastruloids as a model system to investigate the effects of *Lsd1* KO during gastrulation.

# 4 Investigating the effects of *Lsd1* knockout in gastruloids

## 4.1 Chapter Aims

Following the establishment of gastruloids as an appropriate model for recapitulation of the embryonic stage at which *Lsd1* KO is embryonic lethal in Chapter 3, the work in this chapter aimed to investigate the effect of *Lsd1* knockout on the differentiation of ESCs into gastruloids. This involved comparison of gastruloids lacking *Lsd1* to control gastruloids at 0 hour, 72 hour and 120 hour timepoints, including analysis of morphology and differential gene expression. LSD1 has been shown to act as a transcriptional repressor through its demethylation of H3K4me1/2, marks which are permissive for gene expression (Hyun et al., 2017, Heintzman et al., 2007). The aim of this work was therefore to investigate dysregulated genes in gastruloids following depletion of *Lsd1* and to determine how these changes might functionally contribute to the lethality of *Lsd1* KO in the gastrulating embryo.

## 4.2 Gastruloid growth and morphology is impaired in *Lsd1* KOs

Assessment of *Lsd1* KO gastruloids compared to controls revealed reduced size and changes in morphology (Figure 4.1A). These gastruloids exhibited slightly reduced sphericity and rougher edges at 72 hours compared to controls. At 96 hours, elongation was less apparent than in controls, though this appeared to be compensated for at 120 hours. This was also evidenced by the proportion of gastruloids that had elongated at 96 and 120 hours (Figure 4.1B), as there were 15-60% non-elongated *Lsd1* KO gastruloids at 96 hours, depending on the replicate, compared to 0-4% non-elongated control gastruloids at this timepoint. At 120 hours, the number of non-elongated gastruloids in *Lsd1* KO and control gastruloids converged, with 0-6% and 0-3% non-elongated, respectively. Where control gastruloids display a darker, cell-dense rostral region, gastruloids with *Lsd1* depletion appear to lack this definition and exhibit reduced delamination of cells at the gastruloid border.

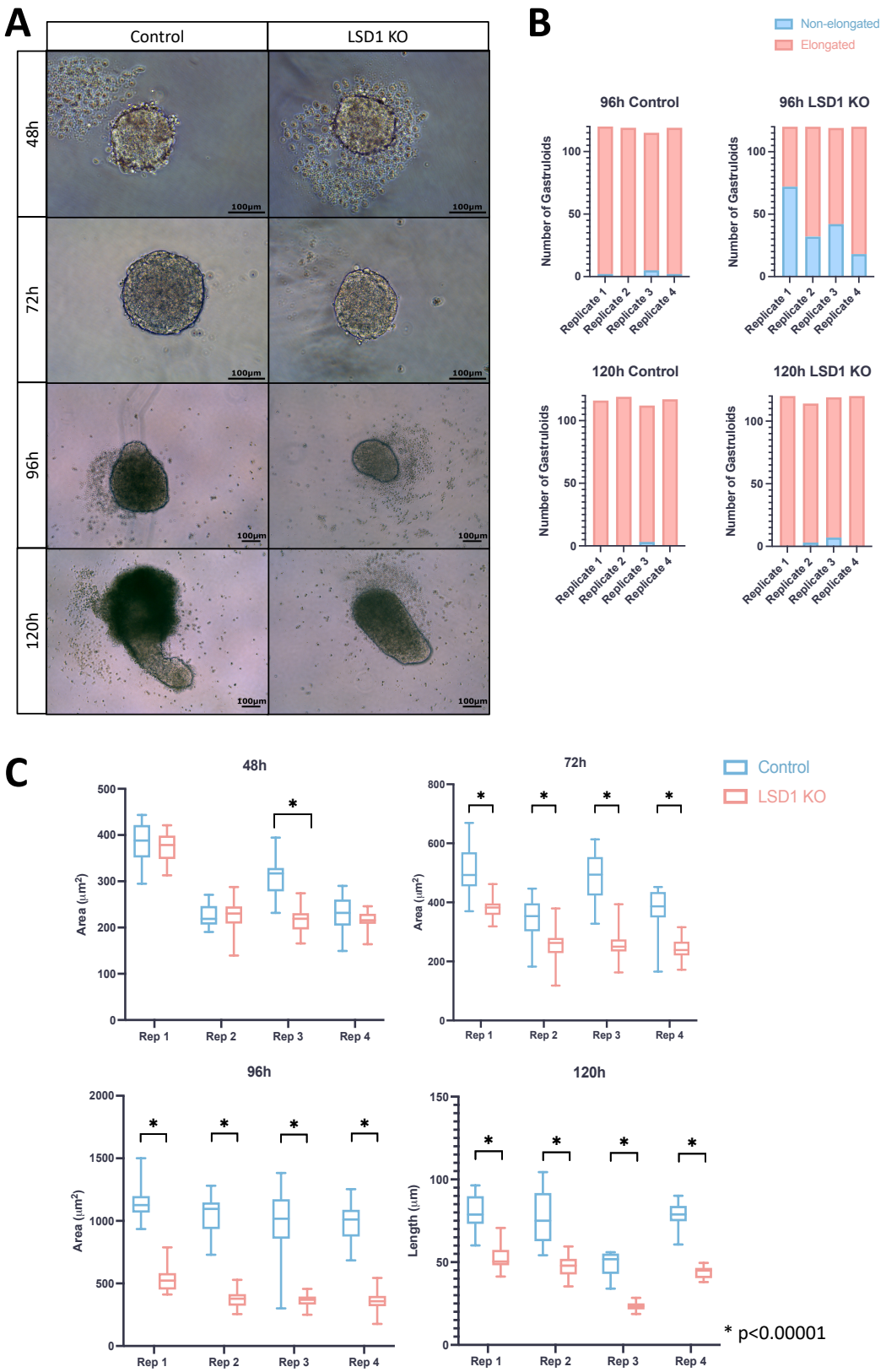


***Figure 4.1 Morphology of Lsd1 KO gastruloids differs from controls*** *(**A**) Representative images of Lsd1 KO versus control gastruloids at 48-, 72-, 96-, and 120-hour timepoints. (**B**) Bar graphs showing the proportion of gastruloids that exhibited elongation in Lsd1 KO and control gastruloids at 96- and 120-hour timepoints. (**C**) Boxplots showing measurements of area (µm²) of 20 per replicate Lsd1 KO versus control gastruloids at 48-, 72- and 96-hours, and length (µm) of 10 per replicate Lsd1 KO versus control gastruloids at 120 hours. Unpaired T-tests were used to test for significant differences and a p-value of <0.00001 is represented by *.*

Measurements of gastruloids over the timepoints illustrate the reduced size of *Lsd1* KO gastruloids versus controls (Figure 4.1C). Gastruloids without *Lsd1* were significantly smaller than controls in all replicates from 72 hours onwards, as assessed by unpaired t-tests. The similarity in size of gastruloids in both conditions at 48 hours suggests this was due to reduced proliferation at later stages rather than increased cell death at earlier stages.

## 4.3 Gastruloids lacking *Lsd1* exhibit differential expression of gastrulation associated genes

RNA-seq was performed on *Lsd1* KO and control gastruloids concurrently for samples harvested at 0, 72 and 120 hours. The variance of samples was assessed using a PCA plot, which showed clustering of samples by timepoint, as well as by treatment condition (Figure 4.2A). This indicates differences between samples at different timepoints and between control and *Lsd1* KO gastruloids. Analysis of Cook's distance revealed no outliers among the samples (Figure 4.2B)

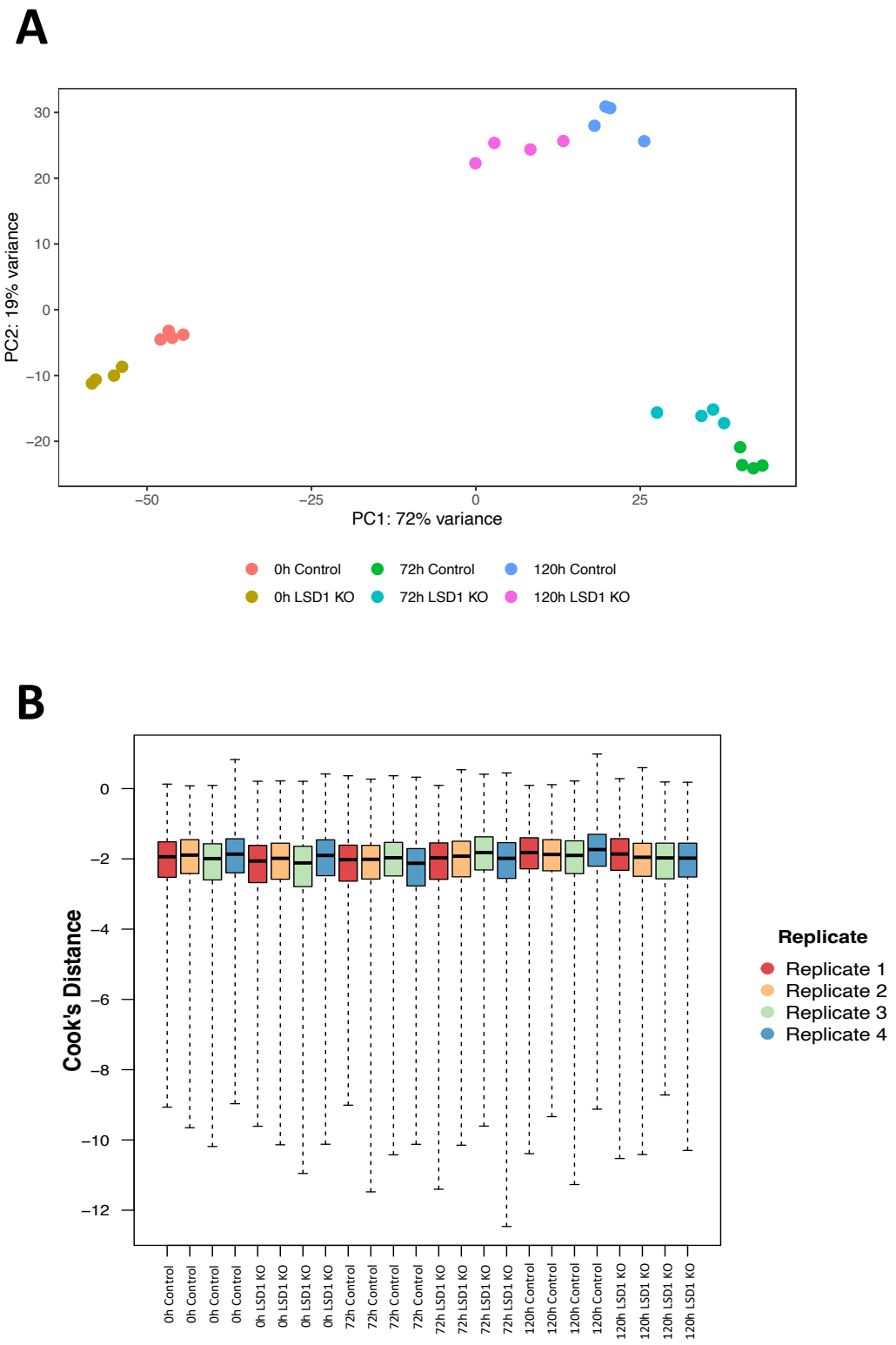


***Figure 4.2 Gastruloid samples were suitable for RNA-seq analysis.*** *(**A**) Plot of principal component analysis (PCA) showing the variance of control and Lsd1 KO samples (**B**) Plot of Cook's distance showing there were no outliers among the samples.*

Numerous differentially expressed genes (DEGs) were identified between control and *Lsd1* KO gastruloids at each timepoint, using parameters of a fold change of >+2/<-2 and an adjusted p-value (padj) of <0.01 (Figure 4.3). Specifically, totals of 881, 1251 and 1220 differentially expressed genes were identified in the 0 hour, 72 hour and 120 hour timepoints, respectively (Figure 4.3B). The role of LSD1 in decommissioning active genes through its transcriptional repressor activity is particularly important during differentiation, where previously active genes must be repressed and vice versa. This is highlighted by the higher number of differentially expressed genes in 72 and 120 hour gastruloids compared to ESCs, as cells undergo differentiation during these timepoints. It would be expected, due to the transcriptionally repressive nature of LSD1, that the majority of differentially expressed genes in gastruloids lacking *Lsd1* would be upregulated. This appears to be the case, with 705, 1084 and 777 of differentially expressed genes being upregulated in the 0 hours, 72 hours and 120 hours, respectively. However, interestingly, there is also a proportion of downregulated genes in each timepoint, with 176, 167 and 443 downregulated genes at 0 hours, 72 hours, and 120 hours, respectively. The higher proportion of downregulated genes at 120 hours could reflect the dysregulation of differentiation processes, potentially as upregulation of transcriptional repressors would lead to downregulation of downstream genes.

Interestingly, there is little overlap in the number of genes differentially expressed at each timepoint (Figure 4.3A) The highest overlap is between 72 hour and 120 hour timepoints and the majority of these are upregulated genes (Figure 4.3A). As is to be expected, *Lsd1* (*Kdm1a*) is among the 7 genes downregulated in all timepoints, as is further evidenced in Figure 4.4.

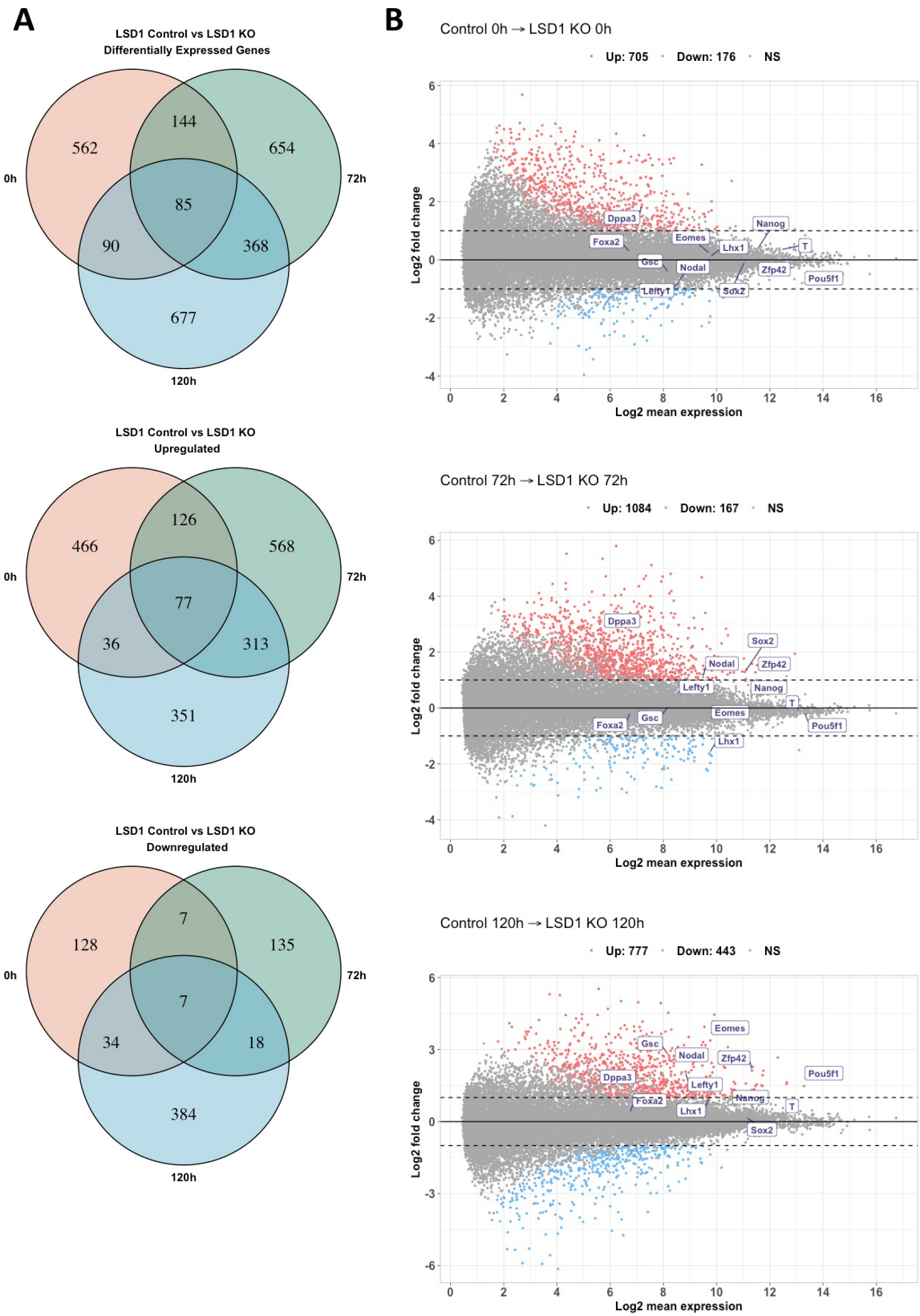


***Figure 4.3 Genes were differentially expressed between Lsd1 KO and control gastruloids at each timepoint.*** *(**A**) Venn diagrams showing the overlapping differentially expressed, upregulated, and downregulated genes between each timepoint. (**B**) MA plots showing genes that are differentially expressed between Lsd1 KO and control gastruloids at each timepoint, using an adjusted p-value of <0.01 and a log2 fold change of >1/<1. Upregulated genes are highlighted in red and downregulated genes in blue. Black dashed lines indicate the cut offs for log2 fold change.*

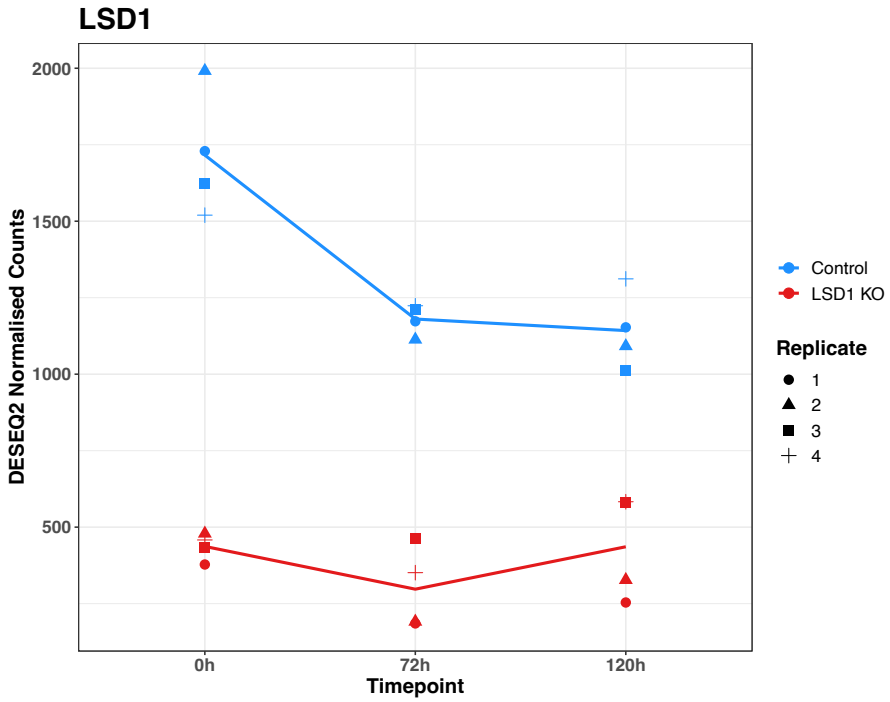


***Figure 4.4 Expression levels of Lsd1 in control and knockout gastruloid samples.*** *Line graphs showing expression of Lsd1 at different timepoints. Points represent values for individual replicates, lines show mean values of replicates.*

Similarity in expression of pluripotency associated genes was seen between control and *Lsd1* KO gastruloids in the ESC state, with only *Dppa3* showing differential expression (Figure 4.3B). This upregulation of *Dppa3*, alongside slight upregulation of *Sox2* and *Zfp42*, continued at 72 hours. Despite these differences, pluripotency markers *Oct4* and *Nanog* were not significantly differentially expressed in 0 hour and 72 hour gastruloids lacking *Lsd1*. At 120 hours, *Dppa3, Zfp42* and *Oct4* are overexpressed in following *Lsd1* depletion whereas *Nanog* and *Sox2* are not.

Overall, expression of pluripotency factors in *Lsd1* KO is similar to controls at 0 hours, with only minor differences at 72 and 120 hours (Figure 4.5B). The suggested role of LSD1 in enhancer decommissioning would suggest that pluripotency factors may fail to switch off during the switch from pluripotency to a differentiated state (Whyte et al., 2012). However, the present data suggests this isn't the case, with the reduction in pluripotency factors in gastruloids with depleted *Lsd1* largely mimicking controls. In particular, analysis of the key pluripotency markers *Oct4*, *Nanog* and *Sox2* show similar levels in *Lsd1* KOs and controls (Figure 4.5A). Interestingly, *Dppa3* and *Nanog* share a

super enhancer positioned 45 kb upstream of *Nanog* which regulates both genes (Blinka et al., 2016). Consequently, it might be expected that *Dppa3* and *Nanog* would show similar expression levels, however, *Dppa3* is upregulated at all timepoints in gastruloids lacking *Lsd1*, whereas *Nanog* is not. Expression levels of early differentiation markers *T*, *Foxa2* and *Fgf5* are similar in *Lsd1* KO and control gastruloids (Figure 4.5A). Overall, these results suggest that the exit from pluripotency and induction of early differentiation are not impaired in gastruloids following *Lsd1* depletion.

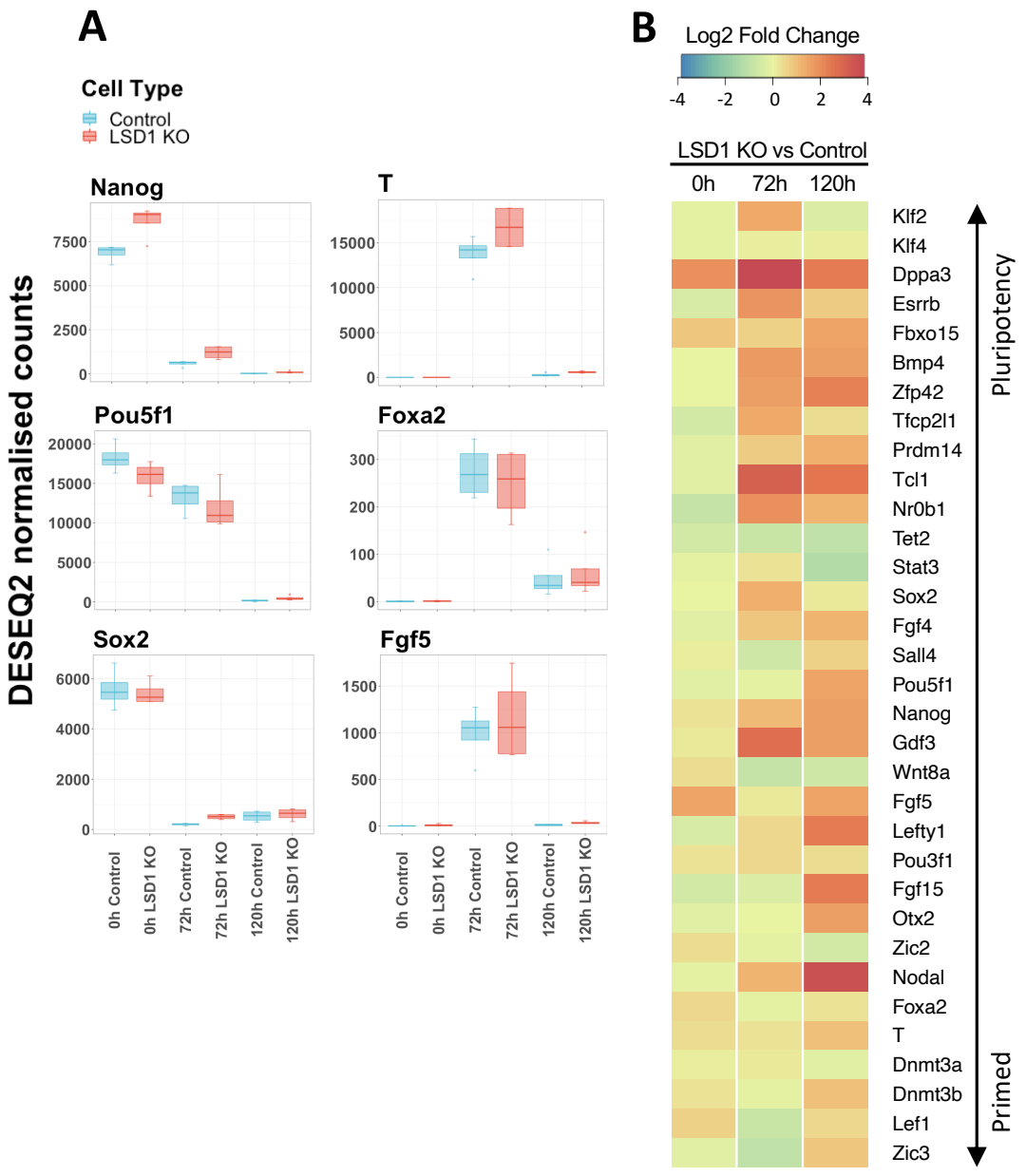


***Figure 4.5 Analysis of pluripotency markers shows exit from pluripotency is unaffected in Lsd1 KO gastruloids compared to controls.*** *(**A**) Boxplots showing expression of genes associated with pluripotency and early differentiation in Lsd1 KO and control gastruloids at different timepoints. (**B**) Heatmap showing log2 fold change values of Lsd1 KO versus control gastruloids at 0-, 72-, and 120-hours for genes associated with pluripotency and early differentiation.*

As is the case with pluripotency markers, genes associated with gastrulation, *Brachyury (T)*, *Foxa2*, *Lhx1*, *Gsc*, *Eomes*, and *Lefty1* are expressed at similar levels in *Lsd1* KO and control ESCs and at 72 hours, with the exception of a downregulation of *Lhx1* in knockouts at 72 hours (Figure 4.3B). However, a number of gastrulation associated genes appear to be upregulated at 120 hours in gastruloids lacking *Lsd1* (Figure 4.6B). In particular, *Gsc*, *Eomes*, *Fgf8*, *Foxh1* and *Mixl1* are significantly upregulated at 120 hours, despite being expressed at similar levels as in controls when expression peaks at 72 hours (Figure 4.6A). This indicates that loss of *Lsd1* does not prevent these genes switching on, but perhaps might impair the switching off of these genes, which correlates with the transcriptional repressive role of the CoREST complex.

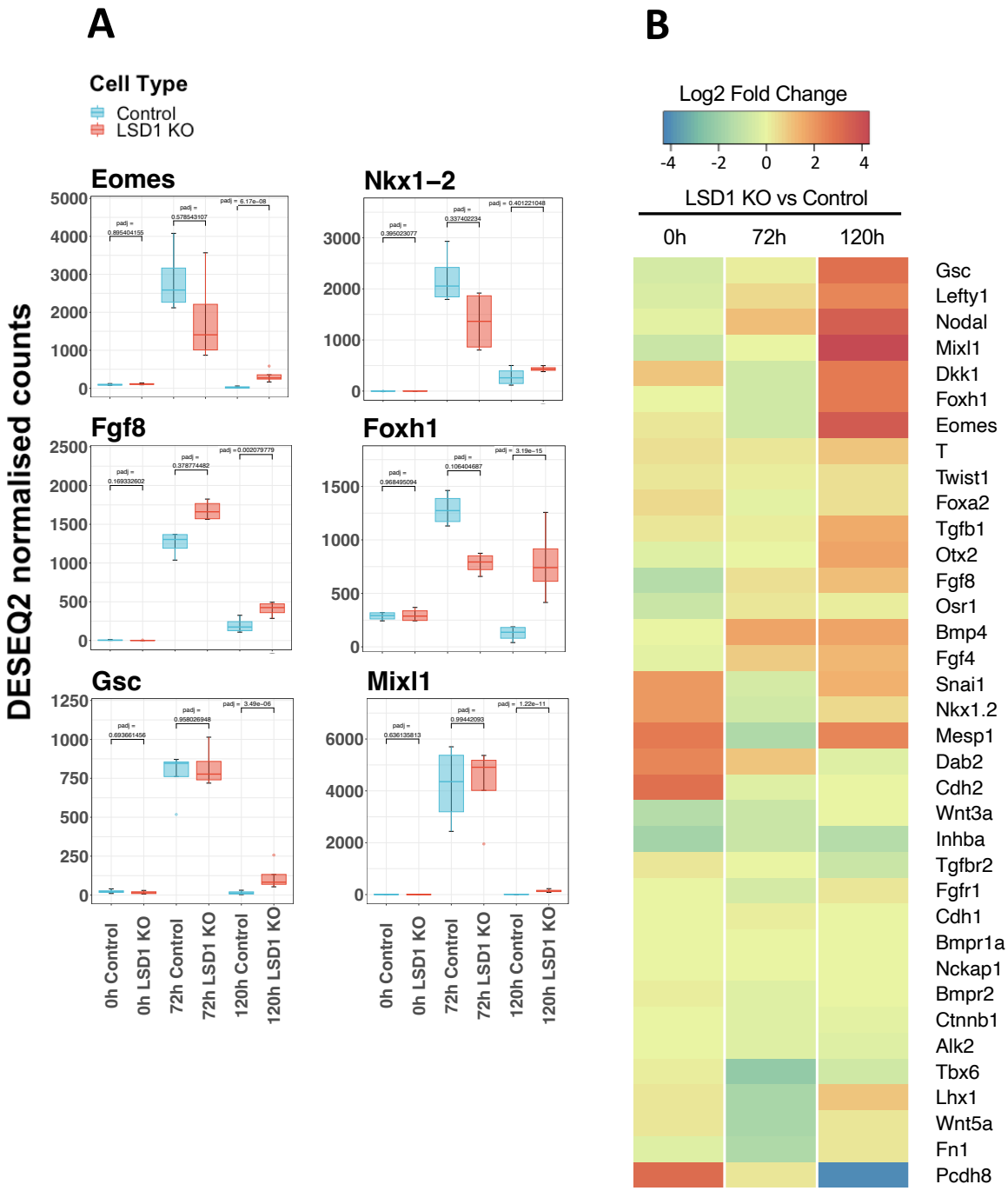


***Figure 4.6 Analysis of gastrulation associated genes shows differential expression in Lsd1 KO versus control gastruloids at 120 hours.*** *(**A**) Boxplots showing expression of genes associated with gastrulation in Lsd1 KO and control gastruloids at different timepoints. Adjusted p-values (padj) for Lsd1 KO versus control gastruloids are displayed. (**B**) Heatmap showing log2 fold change values of Lsd1 KO versus control gastruloids at 0-, 72-, and 120-hours for genes associated with gastrulation.*

A similar effect is seen in the expression of epiblast markers, *Nodal* and *Lefty1*, with significant upregulation of these genes at 120 hours in *Lsd1* KOs (Figure 4.7). *Nodal* is also significantly upregulated at 72 hours. This could also be reflective of reduced ability to switch these genes off from an active state, as in control gastruloids expression of the transcripts decreases from 72 hours onwards.

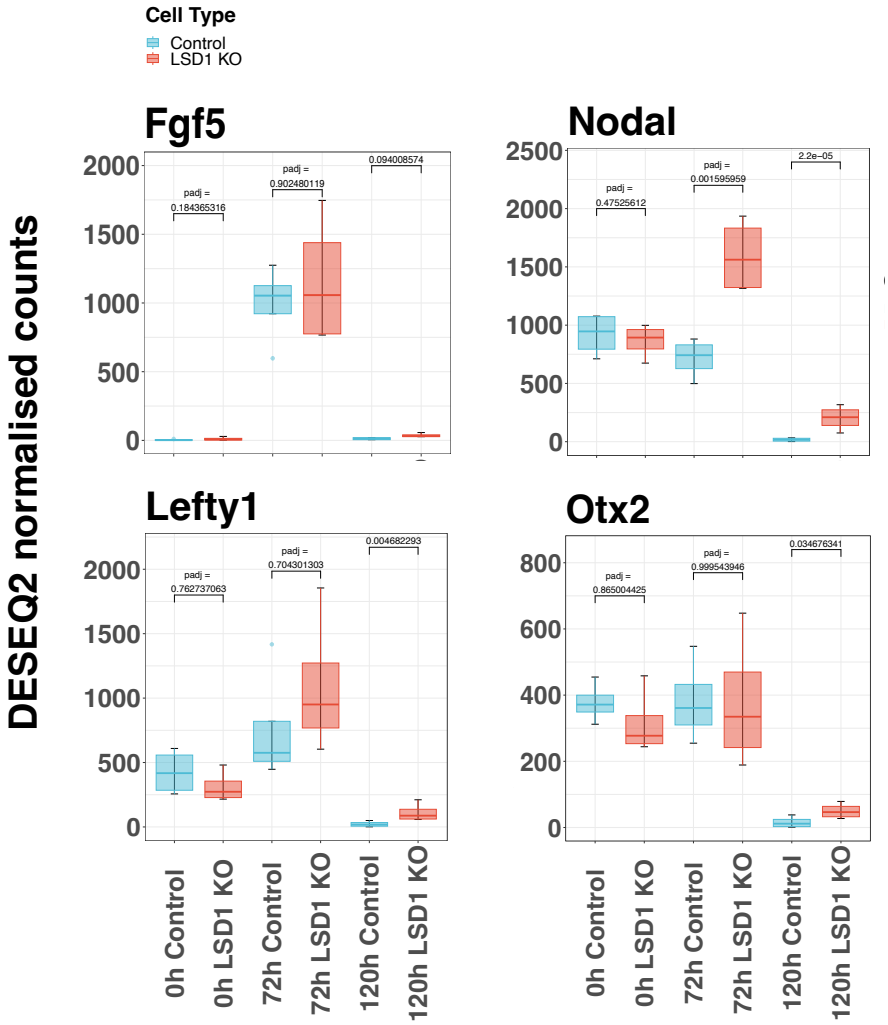


***Figure 4.7 Expression of epiblast markers are differentially expressed in Lsd1 KO versus control gastruloids.*** *Boxplots showing expression of epiblast markers in Lsd1 KO and control gastruloids at different timepoints. Adjusted p-values (padj) for Lsd1 KO versus control gastruloids are displayed.*

## 4.4 Markers of mesodermal lineages are differentially expressed in gastruloids following *Lsd1* depletion

Analysis of lineage specific markers revealed that, while expression of endodermal and ectodermal markers were largely normal, a subset of mesodermal markers were dysregulated in gastruloids lacking *Lsd1* (Figure 4.8A). In particular, *Wnt3*, *Hand1* and *Mesp1*, along with the endodermal marker *Dkk1*, were significantly upregulated at 120 hours (Figure 4.8B). *Hand1* is expressed in lateral mesoderm, the developing heart and a subset of neural crest cells, from around E7.0 (Firulli, 2003). Expression of *Wnt3* is required for formation of the primitive streak and induction of mesoderm, as well as for

establishment of the anterior-posterior axis (Barrow et al., 2007). *Mesp1* determines multiple mesodermal lineages, including cardiac, haemopoietic or skeletal myogenic lineages (Chan et al., 2013). *Eomes* directly activates *Mesp1* expression (Costello et al., 2011), so it would follow that overexpression of *Eomes* at 120 hours (Figure 4.6A) would lead to overexpression of *Mesp1*. Inhibition of Wnt signalling leads to failure to induce many downstream mesodermal genes, including *Brachyury, Mixl1* and *Mesp1*, as well as genes associated with EMT, *Snai1, Fn1* and *Cdh2* (Lindsley et al., 2006), and so overexpression of *Wnt3* may lead to dysregulated expression of these genes. *Dkk1* (*Dickkopf*) is a secreted inhibitor of Wnt signalling which is expressed in the anterior visceral endoderm (AVE) and is required for head induction (Hoshino et al., 2015, Mukhopadhyay et al., 2001). Interestingly, as well as *Mesp1* activation occurring downstream of Wnt signalling, *Dickkopf* is also a direct target of *Mesp1*, indicating that the upregulation of *Wnt3*, *Mesp1* and *Dickkopf* at 120 hours may be due to dysregulation of this gene regulatory cascade (David et al., 2008).

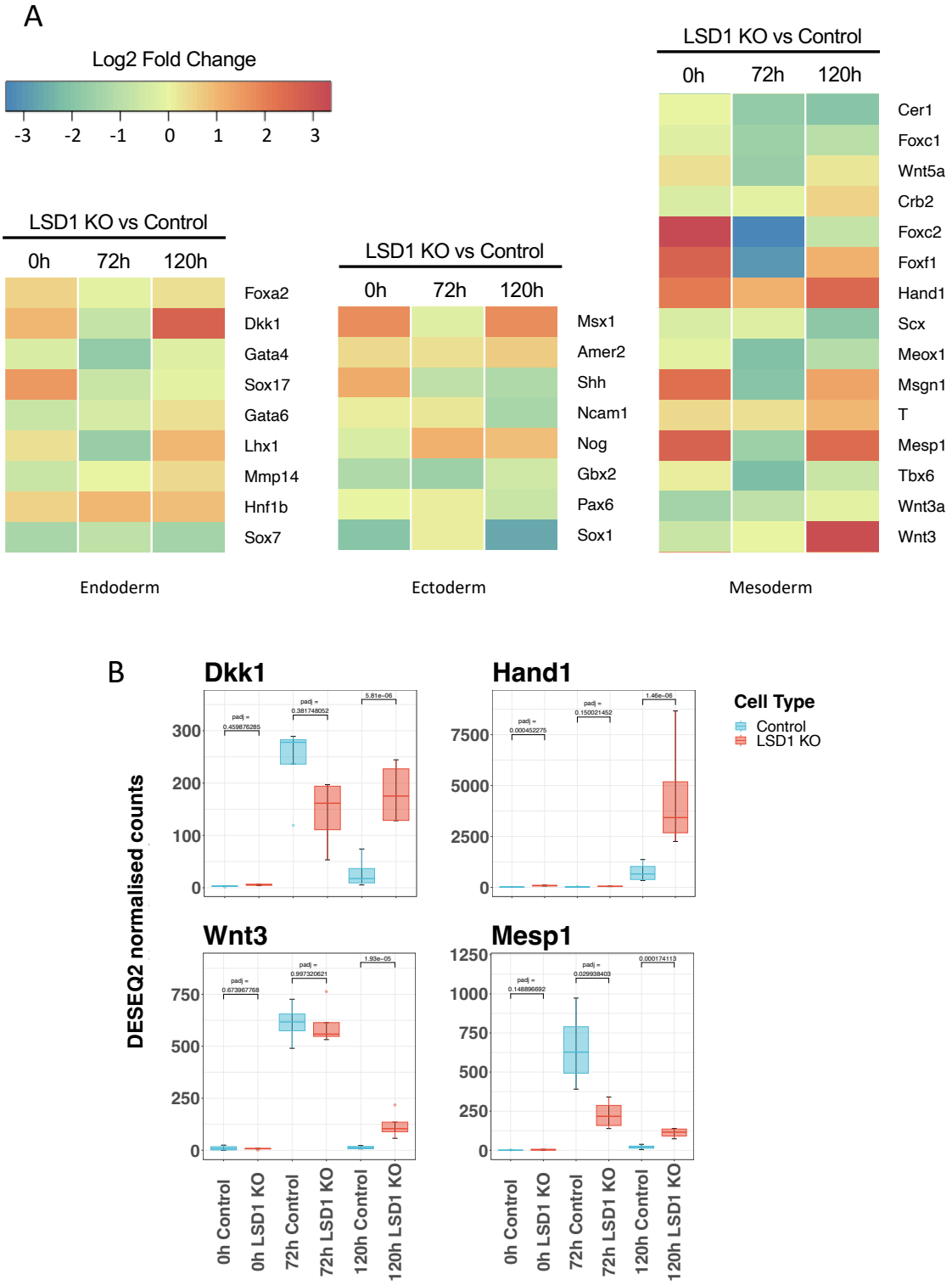


***Figure 4.8 Analysis of lineage specific genes shows differential expression of mesodermal genes in Lsd1 KO versus control gastruloids.*** ***(A)*** *Boxplots showing expression of genes most upregulated in Lsd1 KO versus control gastruloids at 120 hours at different timepoints. Adjusted p-values (padj) for Lsd1 KO versus control gastruloids are displayed.* ***(B)*** *Heatmap showing log2 fold change values of Lsd1 KO versus control gastruloids at 0-, 72-, and 120-hours for genes associated with endoderm, ectoderm and mesoderm.*

Markers of presomitic mesoderm show similar levels of expression in both conditions (Figure 4.9), with the exception of a downregulation of *Tbx6* at 72 hours and

upregulation of *Fgf8* at 120 hours in *Lsd1* KO gastruloids. Genes associated with somite formation, *Tcf15* and *Foxc1* are significantly downregulated at 120 hours in gastruloids lacking *Lsd1*, suggesting the gene transcription programme for somite formation is impaired without LSD1.

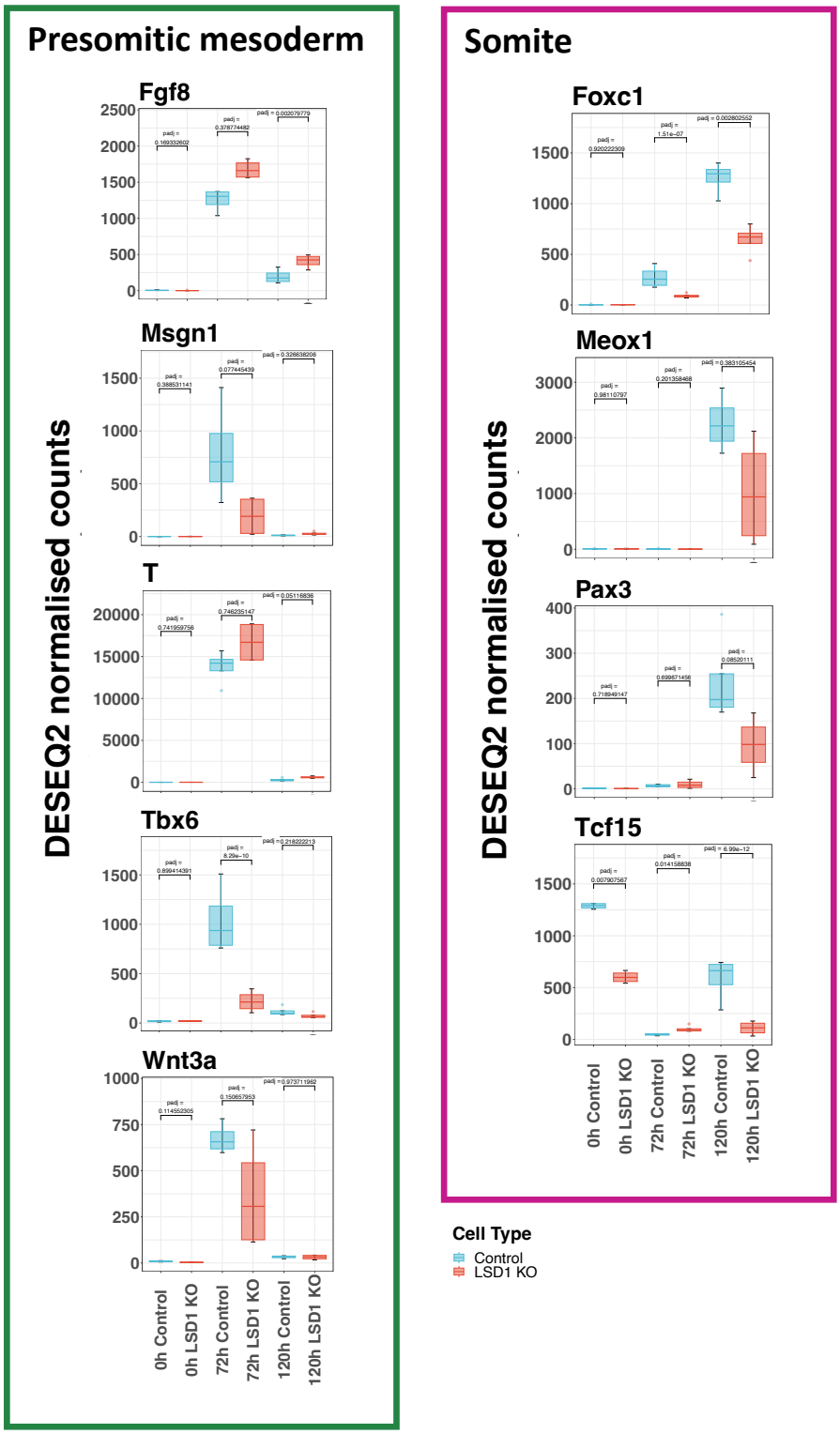


***Figure 4.9 Genes associated with formation of somites are dysregulated in Lsd1 KO versus control gastruloids.*** *Boxplots showing expression of genes associated with presomitic mesoderm and somites in Lsd1 KO versus control gastruloids at different timepoints. Adjusted p-values (padj) for Lsd1 KO versus control gastruloids are displayed.*

Expression of the 3' positioned *Hoxd1* gene was significantly upregulated in *Lsd1* KO gastruloids at all timepoints, reflecting an early and sustained overexpression of this gene (Figure 4.10). At 72 hours, *Hoxc4* was significantly downregulated, followed by downregulation of *Hoxc5* and *Hoxc6* at 120 hours, suggesting incorrect temporal expression of this cluster in gastruloids lacking *Lsd1*.

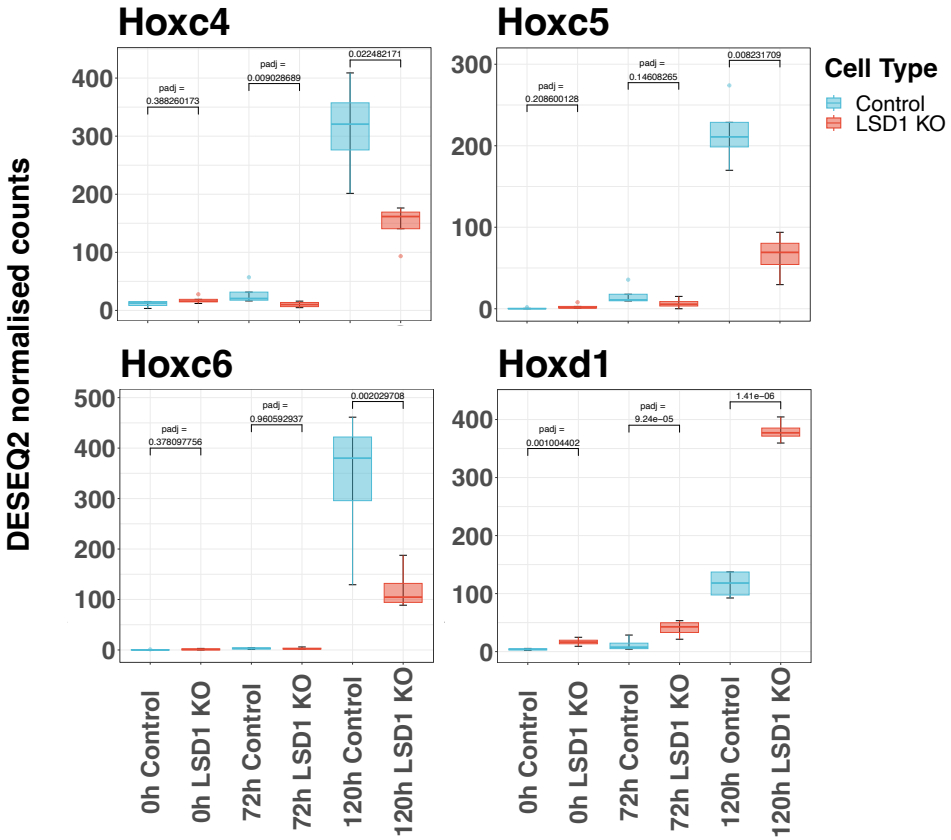


***Figure 4.10 Analysis of Hox genes shows differential expression of select genes in Lsd1 KO versus control gastruloids.*** *Boxplots showing expression of a subset of Hox genes in Lsd1 KO versus control gastruloids at different timepoints. Adjusted p-values (padj) for Lsd1 KO versus control gastruloids are displayed.*

## 4.5 Gene ontology analysis shows alteration of multiple pathways in *Lsd1* KO gastruloids

Gene ontology (GO) analysis of biological processes was performed to identify gene ontology terms associated with differentially expressed genes in each of the timepoints (Figure 4.11). The GO term with the most related differentially expressed genes at both 72 and 120 hours was cell adhesion, closely followed at 72 hours by the term regulation of cell migration. Because these biological processes are linked with EMT, an essential process during gastrulation, this prompted investigation into whether this process is dysregulated in gastruloids following depletion of *Lsd1*.

GO analysis of the 120 hour timepoint also revealed that a number of differentially expressed genes were associated with GO terms relating to the BMP pathway, including cellular response to BMP stimulus and pathway-restricted SMAD protein phosphorylation. Therefore, expression of genes associated with the BMP pathway was further investigated.

Also of note were GO terms associated with primordial germ cells, including male meiotic nuclear division, DNA methylation involved in gamete generation and piRNA metabolic process, as well as with development of the circulatory system, including vascular process in circulatory system, blood vessel morphogenesis and cardiac septum morphogenesis.

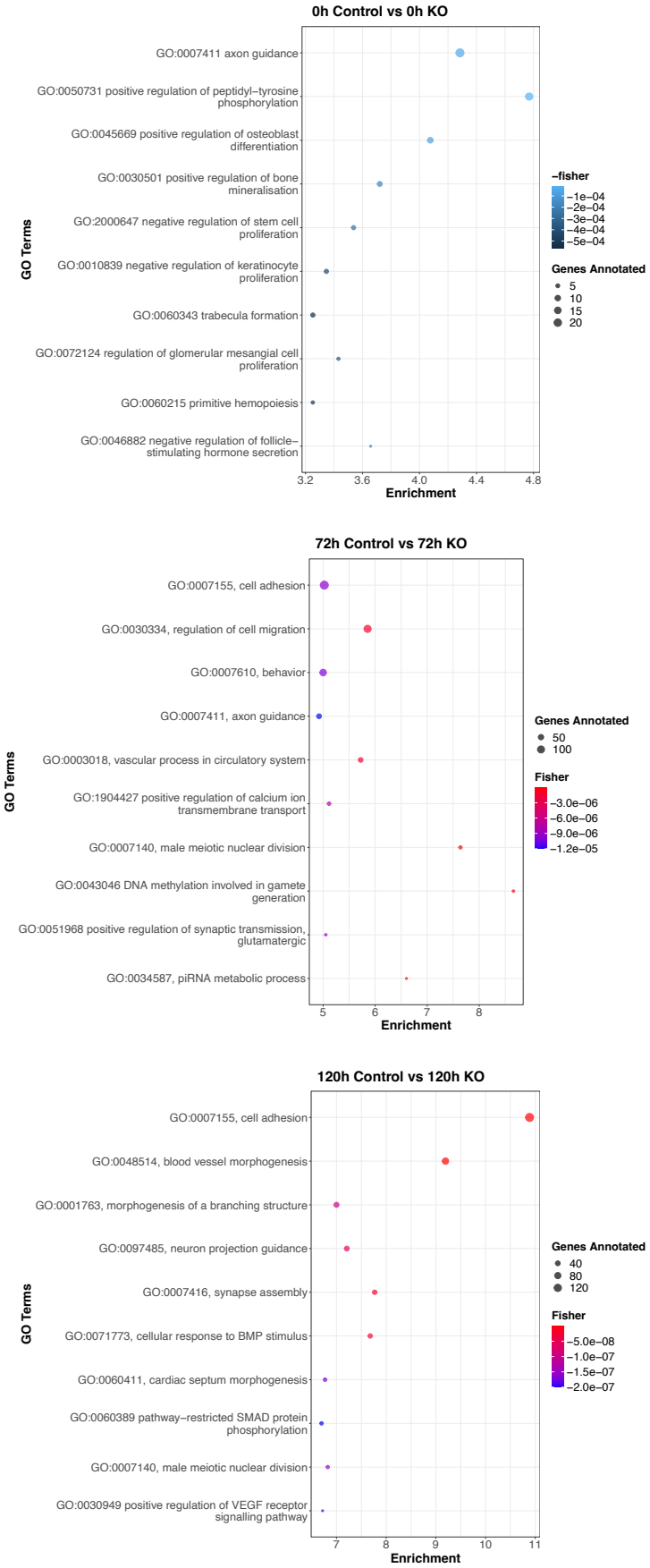


***Figure 4.11 Processes affected by Lsd1 KO were analysed through Biological Process Gene Ontology****. The top 10 biological process GO terms, that are enriched in Lsd1 KO gastruloids relative to controls for each timepoint are shown. GO terms were determined through TopGO analysis (Alexa and Rahnenfuhrer, 2022).*

### 4.5.1 Genes associated with EMT are differentially expressed in gastruloids lacking *Lsd1*

Further investigation into the genes found in the GO term cell adhesion show that both upregulation and downregulation of genes contribute to this term (Figure 4.12). Within this term, the top three upregulated genes in *Lsd1* KO gastruloids were *Rtn4rl1 (NgR2), L1cam* and *Miat* at 120 hours, and *Miat*, *Myo1f* and *Lama3* at 72 hours (Figure 4.13). The top three downregulated genes were *Cdh5*, *Cldn5* and *Pcdh8* at 120 hours and *Foxc2*, *Kdr* and *Pdgfra* at 72 hours (Figure 4.13).

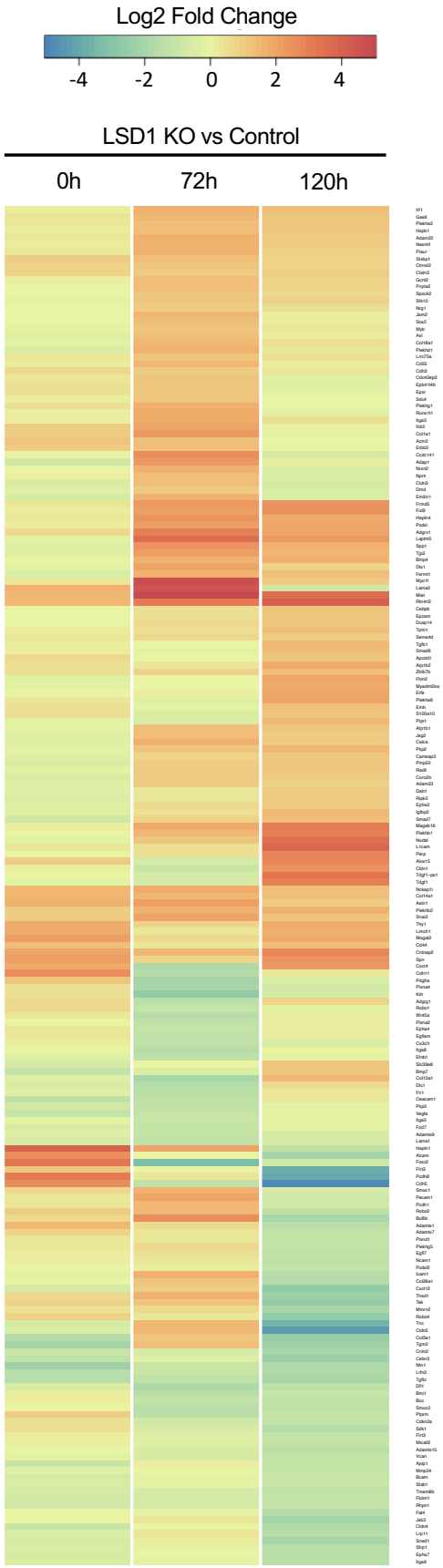


***Figure 4.12 Genes in the GO term cell adhesion were differentially expressed In Lsd1 KO gastruloids.*** *Heatmap showing log2 fold change values of LSD1 KO versus control gastruloids at 0-, 72-, and 120-hours for genes associated with the GO term cell adhesion.*

*Rtn4rl2 (NgR2)* is primarily expressed in the brain where it acts as a receptor for myelin inhibitory proteins which inhibit axonal growth (Venkatesh et al., 2005). *L1cam* encodes a cell-adhesion protein expressed in the nervous system which plays a role in neurite outgrowth and, notably, is regulated by REST (Samatov et al., 2016, Kallunki et al., 1997). *Cdh5* (*VE-cadherin*) is the main component in endothelial adherens junctions and plays a role in organisation of the vascular system during development (Giannotta et al., 2013). *Cdh5* upregulates the tight junction gene *Cldn5* (*Claudin5*) (Taddei et al., 2008), which may explain its parallel downregulation in 120 hour *Lsd1* KO gastruloids.

A number of the top up-/downregulated genes at 72 and 120 hours play a role in the process of EMT. The long non-coding RNA MIAT (myocardial infarction associated transcript), which is upregulated at both 72 and 120 hours in *Lsd1* KO gastruloids, induces EMT through suppressing *E-cadherin* and stimulating *N-cadherin, vimentin* and *Snai1* via the Wnt signalling pathway (Zhong et al., 2019). VEGF has an inhibitory effect on EMT, which is mediated by its receptor KDR (Yang et al., 2008), so downregulation of *Kdr* at 72 hours may reduce inhibition of EMT. Conversely, *Foxc2* is important for maintenance of mesenchymal cell identity following EMT, and loss of *Foxc2* leads to reduced expression of *N-cadherin*, *fibronectin* and *vimentin* (Hollier et al., 2013), meaning downregulation of *Foxc2* at 72 hours might impede EMT. *Pdgfra* promotes EMT and decreases expression of epithelial integrity genes, including TFF3 (Lopez-Campistrous et al., 2021), which suggests that downregulation of *Pdgfra* would also impair EMT. These results together suggest the process of EMT may be dysregulated in *Lsd1* KO gastruloids.

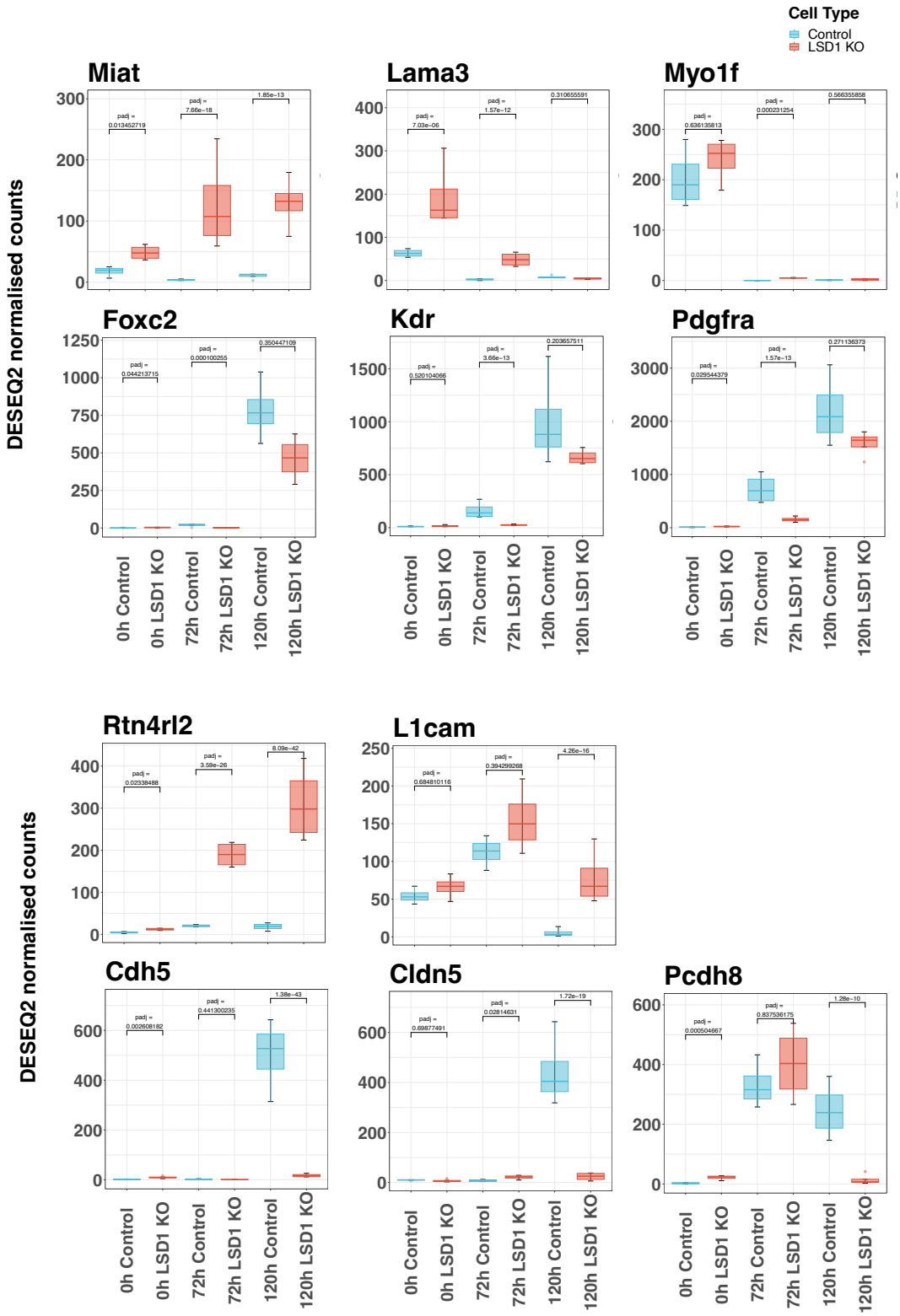


***Figure 4.13 Top up- and downregulated genes In the GO term cell adhesion.*** *Boxplots showing expression of the top three up- and downregulated genes associated with the GO term cell adhesion at 72- and 120-hours in Lsd1 KO versus control gastruloids at different timepoints. Adjusted p-values (padj) for Lsd1 KO versus control gastruloids are displayed.*

To further investigate whether EMT is dysregulated in gastruloids following *Lsd1* depletion, the expression levels of key genes that mediate EMT were investigated (Figure 4.14). Despite no differential expression of *E-cadherin* and *N-cadherin,* which might be expected if EMT is dysregulated, apart from a slight upregulation of *N-cadherin* at 0 hours, differential expression of *fibronectin* and *vimentin* is observed. *Fibronectin* is significantly downregulated at 72 hours, whereas *vimentin* is upregulated at all timepoints in *Lsd1* KO gastruloids. *Fibronectin* and *vimentin* are mesenchymal markers expressed in migrating mesodermal cells during gastrulation (Klinowska et al., 1994, Mendez et al., 2010), the upregulation of *vimentin* and downregulation of *fibronectin* is therefore unusual, as both genes contribute to mesenchymal cell identity. FGFR1 maintains expression of *Snai1* in the primitive streak to induce EMT (Ciruna and Rossant, 2001). Although *Fgfr1* is significantly downregulated at 72 hours in gastruloids lacking *Lsd1*, *Snai1* is not downregulated, but rather upregulated. *Snai1* (*Snail*) and *Snai2* (*Slug*) regulate EMT through repression of epithelial genes, such as *E-cadherin,* and activation of mesenchymal genes, such as *N-cadherin. Snai1* is upregulated at 120 hours and *Snai2* is upregulated at 72 and 120 hours in *Lsd1* KO gastruloids. The sustained overexpression of the genes at 120 hours could suggest increased EMT at this timepoint. *Eomes* also regulates EMT, and lack of *Eomes* leads to failure of mesoderm to delaminate and migrate from the primitive streak (Arnold et al., 2008). The upregulation of *Eomes* at 120 hours in gastruloids lacking *Lsd1* could similarly represent increased mesenchymal identity later than in controls. Together these results indicate dysregulation of EMT in gastruloids following *Lsd1* depletion.

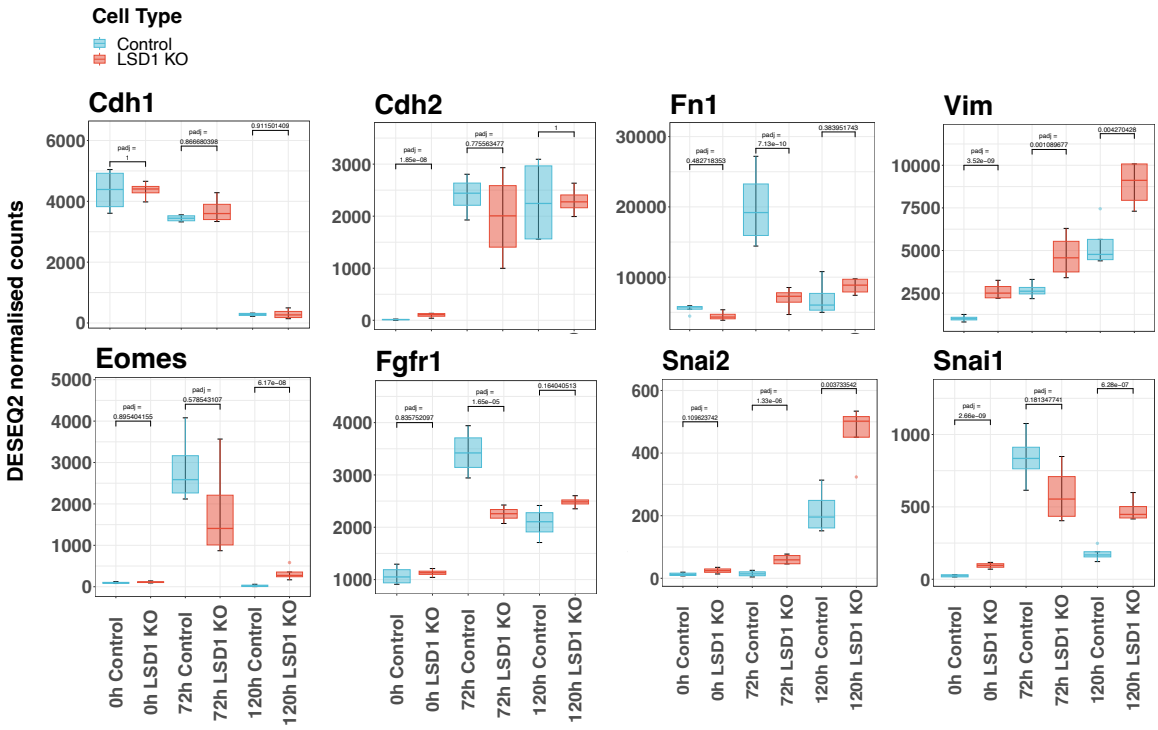


***Figure 4.14 Genes associated with the epithelial-to-mesenhymal transition are dysregulated in Lsd1 KO versus control gastruloids****. Boxplots showing expression of genes associated with EMT in Lsd1 KO versus control gastruloids at different timepoints. Adjusted p-values (padj) for Lsd1 KO versus control gastruloids are displayed.*

## 4.5.2 Genes involved in the BMP pathway are dysregulated in *Lsd1* KO gastruloids

Investigation into genes related to the GO terms cellular response to BMP stimulus and pathway-restricted SMAD protein phosphorylation revealed a number of these genes were upregulated at 120 hours in gastruloids following *Lsd1* depletion (Figure 4.15). These included initiators of the BMP pathway, *Bmp4* and *Bmp7*, which were significantly upregulated at 72 and 120 hours, and at 120 hours, respectively.

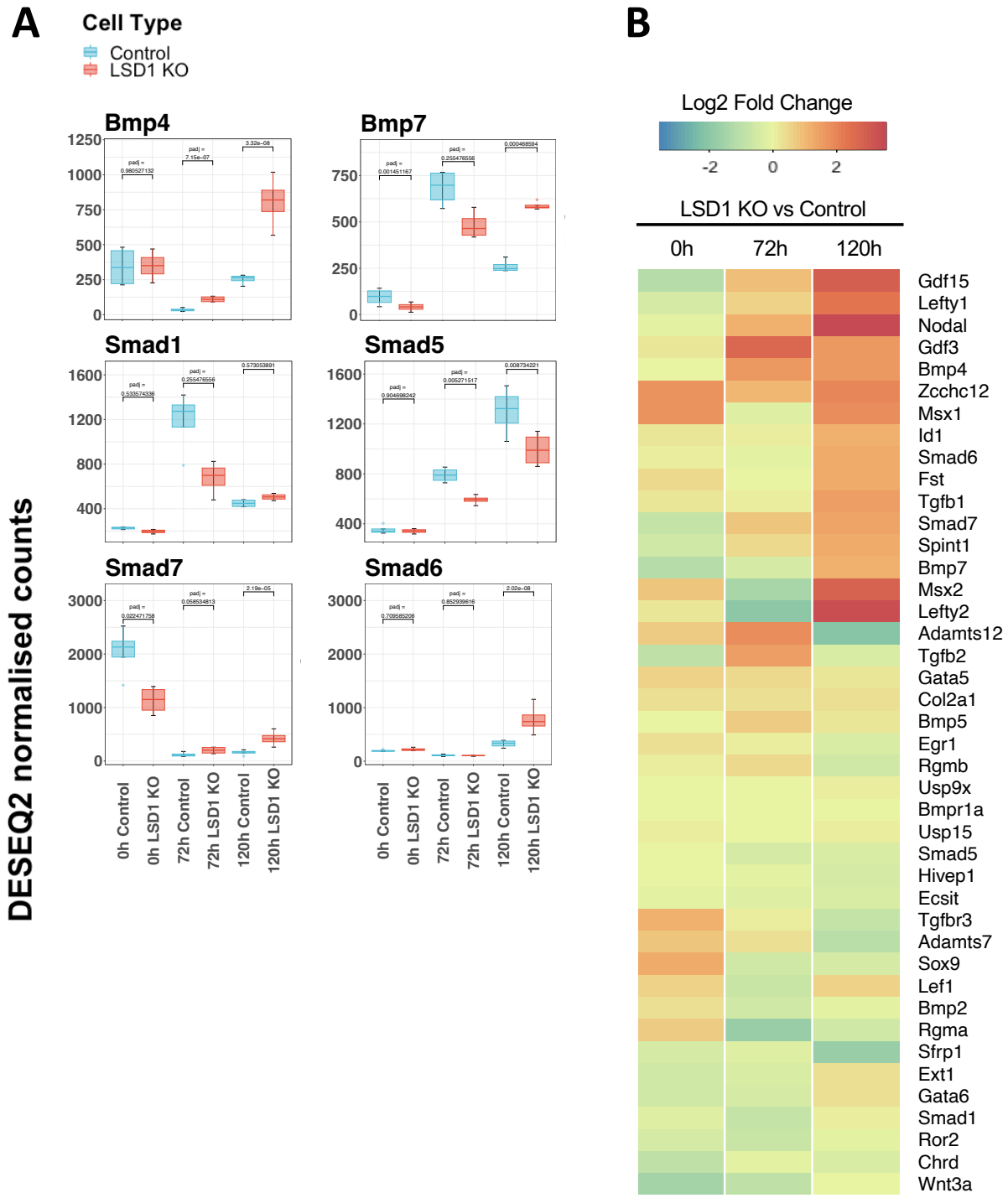


***Figure 4.15 Analysis of lineage specific genes shows differential expression of mesodermal genes in Lsd1 KO versus control gastruloids.*** *(**A**) Boxplots showing expression of genes associated with the GO terms cellular response to BMP stimulus and pathway-restricted SMAD protein phosphorylation in Lsd1 KO versus control gastruloids at different timepoints. Adjusted p-values (padj) for Lsd1 KO versus control gastruloids are displayed. (**B**) Heatmap showing log2 fold change values of Lsd1 KO versus control gastruloids at 0-, 72-, and 120-hours for genes associated with the GO terms cellular response to BMP stimulus and pathway-restricted SMAD protein phosphorylation.*

SMAD1, SMAD5 and SMAD8 are receptor-regulated SMADs (R-SMADs) which are activated by BMP type I receptors (BMPRI) through phosphorylation and go on to complex with SMAD4 in order to regulate transcription of downstream genes. R-SMADs possess both MH1 (Mad homology 1) and MH2 domains, which confer DNA binding activity and association with type I receptors and SMAD4, respectively. In contrast, SMAD6 and SMAD7 lack the DNA binding domain but retain the MH2 domain, allowing them to associate with type I receptors in a way that inhibits R-SMAD activity (Nishimura et al., 2003). As can be seen in Figure 4.15, expression of these inhibitory SMADs is significantly increased at 120 hours in *Lsd1* KO gastruloids, suggesting there may be reduced phosphorylation and activation of R-SMADs, including SMAD1 and SMAD5. In addition to this, expression levels of *Smad5* are significantly reduced at 72 and 120 hours. Therefore, not only are gene levels of BMP ligands aberrantly expressed at 120 hours, but also downstream effectors of the BMP pathway are differentially expressed in gastruloids lacking *Lsd1*.

BMP signalling has previously been associated with dorsal-ventral patterning through a gradient interpreted by downstream genes. Specifically, high expression of BMP ventrally and low expression dorsally leads to expression of *Sfrp1* ventrally and *Bambi* dorsally, which is mediated by pSMAD5 (Greenfeld et al., 2021). Changes in this expression is reflected in Figure 4.16, where, although *Bmp4* and *Bmp7* are overexpressed, reduced *Smad5* expression at 120 hours may lead to the increased expression of *Bambi* and reduced expression of *Sfrp1*. This could indicate that the gradient of BMP signalling responsible for dorsoventral patterning is not appropriately established in gastruloids following depletion of *Lsd1*, despite evidence of a dorsoventral axis in previous gastruloid studies (Beccari et al., 2018b). *Bambi* is a BMP inhibitor and has also been shown to share a spatio-temporal expression pattern with *Bmp4*, which induces its expression, so increased *Bmp4* could also directly lead to increased *Bambi* expression (Grotewold et al., 2001).

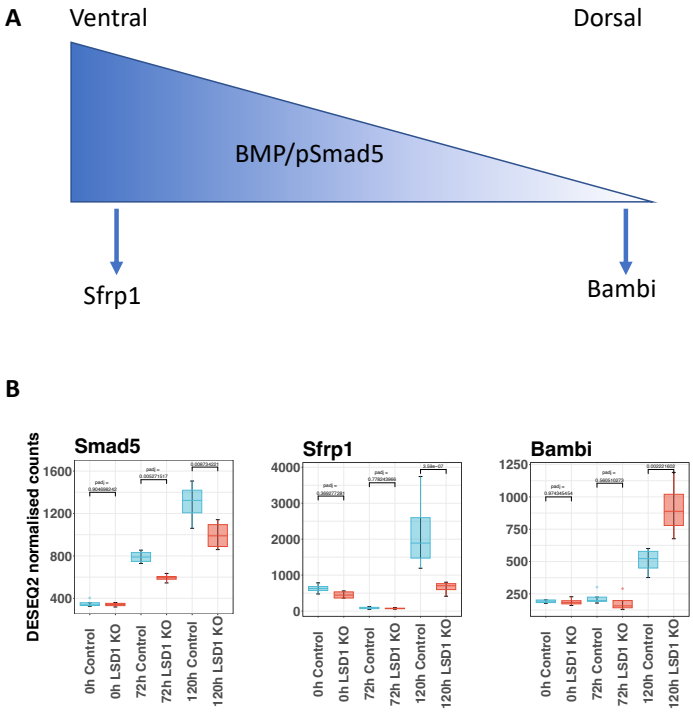


***Figure 4.16 BMP signalling affects dorsoventral patterning.*** *(**A**) Schematic showing the gradient of BMP signalling and its downstream effectors. (**B**) Boxplots showing expression of genes downstream of BMP dorsoventral patterning in Lsd1 KO versus control gastruloids at different timepoints. Adjusted p-values (padj) for Lsd1 KO versus control gastruloids are displayed.*

## 4.6 Summary

The results in this chapter revealed numerous differentially expressed genes at each timepoint in *Lsd1* KO gastruloids versus controls. Although gastruloids lacking *Lsd1* demonstrated exit from pluripotency, they also exhibited overexpression of a number of gastrulation associated genes at 120 hours, reflecting incorrect timing of this process. In particular, expression of mesodermal markers was affected in gastruloids following depletion of *Lsd1*, suggesting incorrect establishment of this lineage. Gene ontology analysis revealed several biological processes affected in *Lsd1* KO gastruloids, including genes in GO terms associated with cell adhesion and the BMP pathway. Further analysis suggested a dysregulation of EMT and BMP signalling in gastruloids lacking *Lsd1*, which may explain in part the lethal phenotype seen upon loss of *Lsd1* in embryos at E6.5.

# 5 Investigation into the requirement of LSD1 demethylase activity in gastruloids

## 5.1 Chapter Aims

The stability of the CoREST complex is dependent on LSD1 and therefore loss of LSD1 would be predicted to disrupt both demethylase activity and HDAC deacetylase activity (Foster et al., 2010). The work in this chapter sought to discriminate between differential gene expression resulting from depleted LSD1 demethylase activity specifically and that resulting from loss of the broader CoREST complex activity. To this end, we utilised previously generated *Lsd1*$^{lox/\Delta 3}$ ESCs with *piggyBac* wildtype *Lsd1* and *Lsd1* mutant K661A inserts to rescue the effects of *Lsd1* knockout. *Lsd1* K661A is a catalytically inactive mutant which has been employed extensively in previous studies (Lee et al., 2005, Chen et al., 2006, Lee et al., 2006, Huang et al., 2007, Adamo et al., 2011, Maiques-Diaz et al., 2018, Sehrawat et al., 2018). Gastruloids were generated using ESCs with wildtype and K661A mutant rescue constructs, hereafter referred to as WT-*Lsd1* and *Lsd1*-K661A, respectively, and RNA-seq was performed at 0, 72, and 120 hours to assess differentially expressed genes (DEGs) between these conditions.

## 5.2 Morphology of WT-*Lsd1* and *Lsd1*-K661A gastruloids

WT-*Lsd1* and *Lsd1*-K661A gastruloids were similar in their morphology, despite size differences between the two conditions (Figure 5.1A). Comparing the morphology to control and *Lsd1* KO gastruloids in Chapter 4, morphology of WT-*Lsd1* and *Lsd1*-K661A resemble *Lsd1* KO more than the control condition, as they exhibit lighter rostral regions with less cell delamination. Size differences between the two conditions varied between replicates, with replicate 1 and 2 *Lsd1*-K661A gastruloids starting out smaller than WT-*Lsd1* gastruloids at earlier timepoints before increasing to a similar size, and in replicate 3 and 4 *Lsd1*-K661A gastruloids were larger at all timepoints (Figure 5.1C).

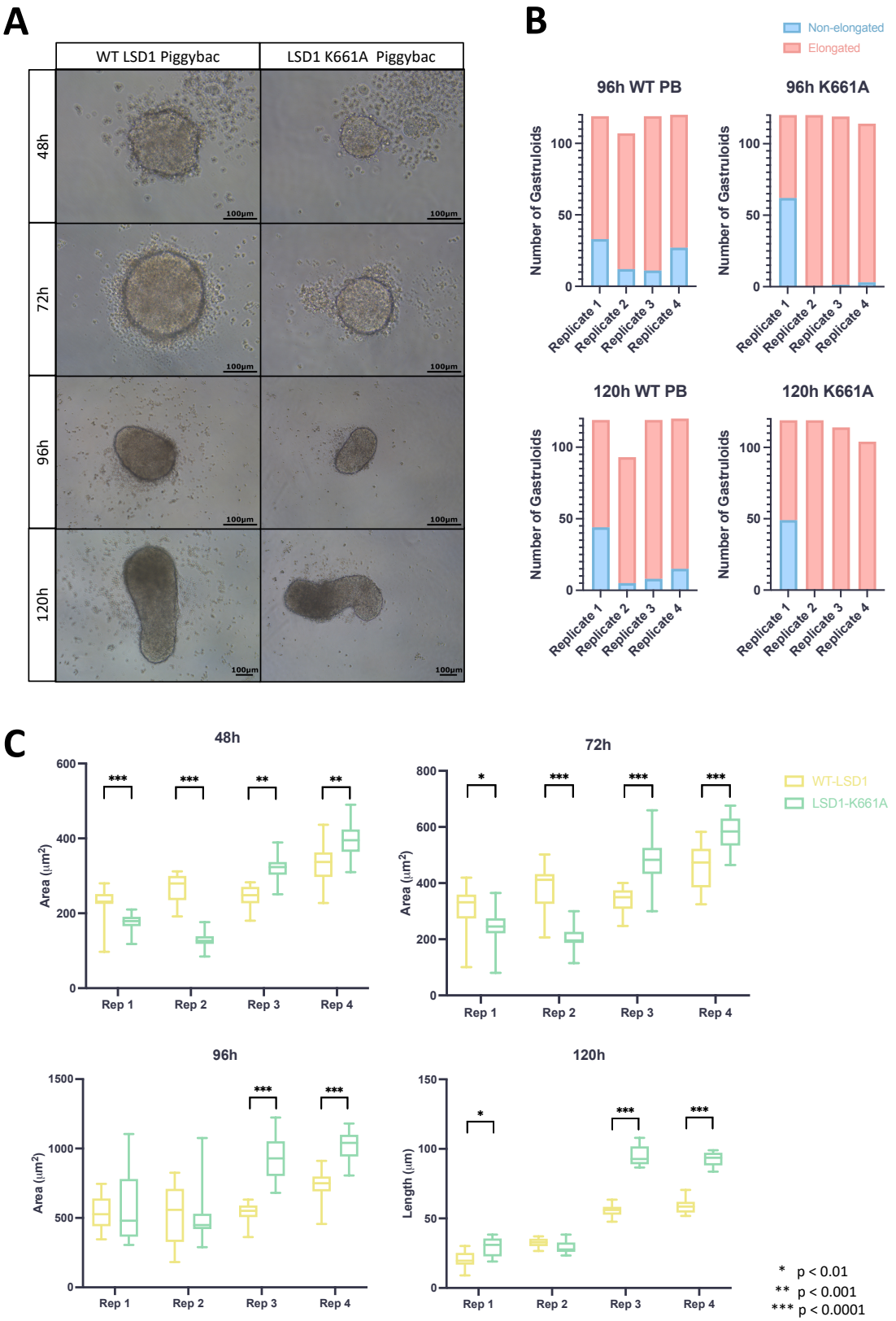


***Figure 5.1 Morphology of Lsd1-K661A and WT-Lsd1 gastruloids exhibit some differences*** *(**A**) Representative images of Lsd1-K661A versus WT-Lsd1 gastruloids at 48-, 72-, 96-, and 120-hours. (**B**) Bar graphs showing the proportion of gastruloids elongated in Lsd1-K661A and WT-Lsd1 gastruloids at 96- and 120-hour timepoints. (**C**) Boxplots showing measurements of area (µm²) of 20 per replicate Lsd1-K661A versus WT-Lsd1 gastruloids at 48-, 72- and 96-hours, and length (µm)of 10 per replicate Lsd1-K661A versus WT-Lsd1 gastruloids at 120 hours. Unpaired T-tests were used to test for significant differences and p-values of <0.01, <0.001 and <0.0001 are represented by *, ** and ***, respectively.*

Replicate 1 appeared to be outlying in its elongation data in both conditions compared to other replicates, with 28% and 37% non-elongated WT-*Lsd1* gastruloids at 96 and 120 hours, respectively, and with 52% and 41% non-elongated *Lsd1*-K661A gastruloids at 96 and 120 hours, respectively (Figure 5.1B). In all other replicates, elongation appeared to be less prevalent in WT-*Lsd1* gastruloids than in *Lsd1*-K661A gastruloids at 96 hours, with 11-23% non-elongated and 0-3% non-elongated, respectively. This was also the case at 120 hours, despite an increase in elongated gastruloids at this timepoint, with 5-13% non-elongated WT-*Lsd1* gastruloids and no non-elongated *Lsd1*-K661A gastruloids.

## 5.3 Differentially expressed genes between gastruloids with wildtype and mutant *Lsd1* rescue constructs

Samples of 0, 72 and 120 hour WT-*Lsd1* and *Lsd1*-K661A gastruloids were used for RNA-sequencing. Variance between samples was assessed using a PCA plot, which showed that samples clustered well by timepoint, but that there was less distinct clustering between conditions (Figure 5.2A). Cook's distance analysis showed no outliers among the samples (Figure 5.2B).

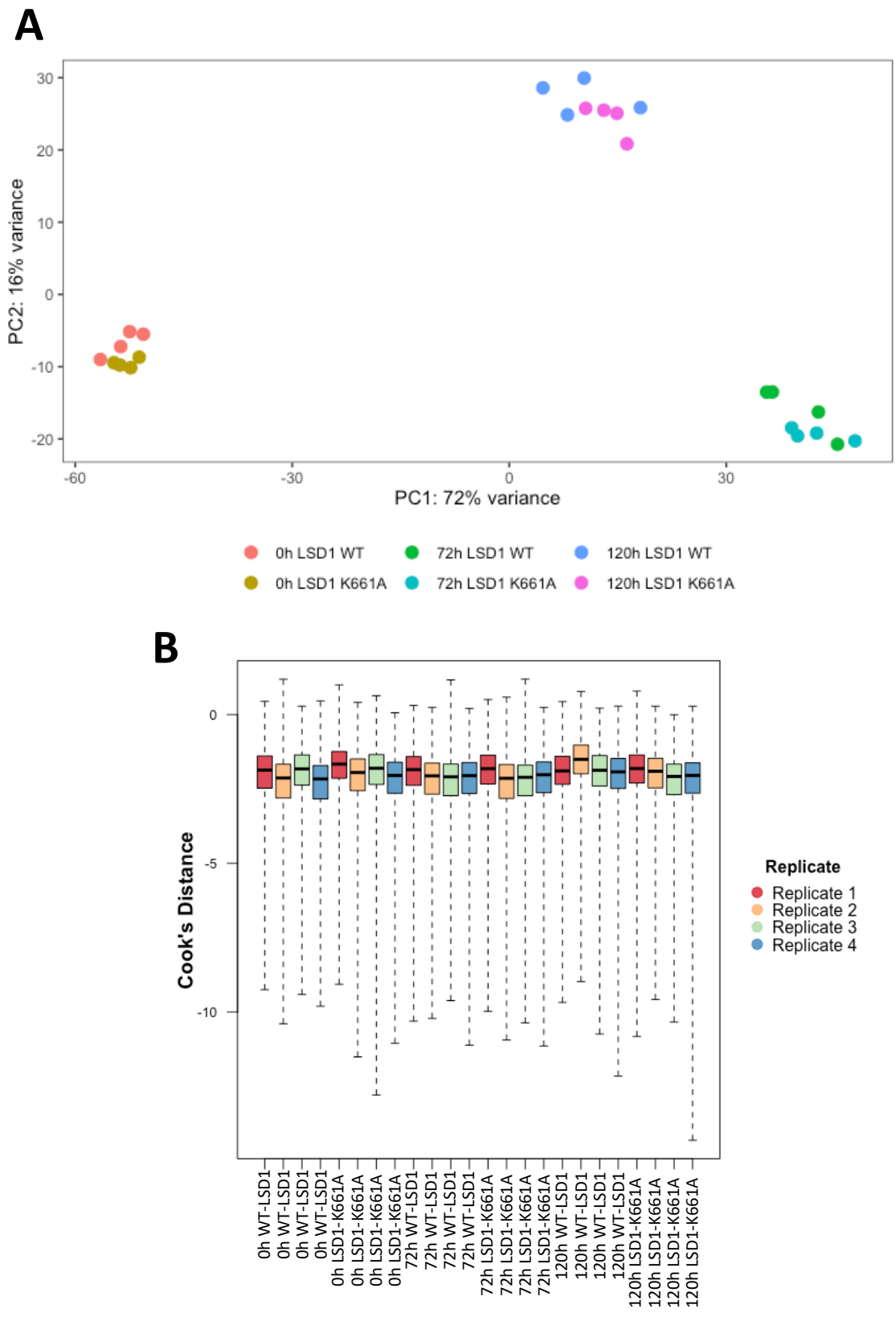


***Figure 5.2 Gastruloid samples were suitable for RNA-seq analysis.*** *(**A**) Plot of principal component analysis (PCA) showing the variance of WT-Lsd1 and Lsd1-K661A samples (**B**) Plot of Cook's distance showing there were no outliers among the samples.*

Differentially expressed genes (DEGs) were identified using parameters of a fold change greater or less than 2/-2 and an adjusted p-value (padj) of <0.01 (Figure 5.3). Overall, totals of 1135, 313, and 1032 DEGs were identified between WT-*Lsd1* and *Lsd1*-K661A gastruloids in the 0 hour, 72 hour and 120 hour timepoints, respectively. The comparatively small number of DEGs identified at 72 hours is indicative of more similar gene expression between the conditions at this time, which could reflect similar progression of differentiation. The higher proportion of upregulated genes at all timepoints in gastruloids with the *Lsd1* mutant is consistent with the transcriptional repressor activity of LSD1. Specifically, 867, 253 and 686 DEGs were upregulated at the 0 hour, 72 hour and 120 hour timepoints, respectively.

Expression of pluripotency and gastrulation markers between the conditions at 0 hours shows upregulation of *Brachyury*, *Foxa2* and *Eomes*, and downregulation of *Gsc* and *Zfp42* in *Lsd1*-K661A gastruloids (Figure 5.3). This might suggest a slightly altered ESC state in this condition. At 72 hours, all gastrulation markers and *Dppa3* are upregulated, while expression of pluripotency markers *Nanog*, *Oct4* and *Sox2* is similar in both conditions, perhaps reflecting increased differentiation in mutant LSD1 gastruloids. Upregulation of these genes, aside from *Lhx1*, was retained at 120 hours.

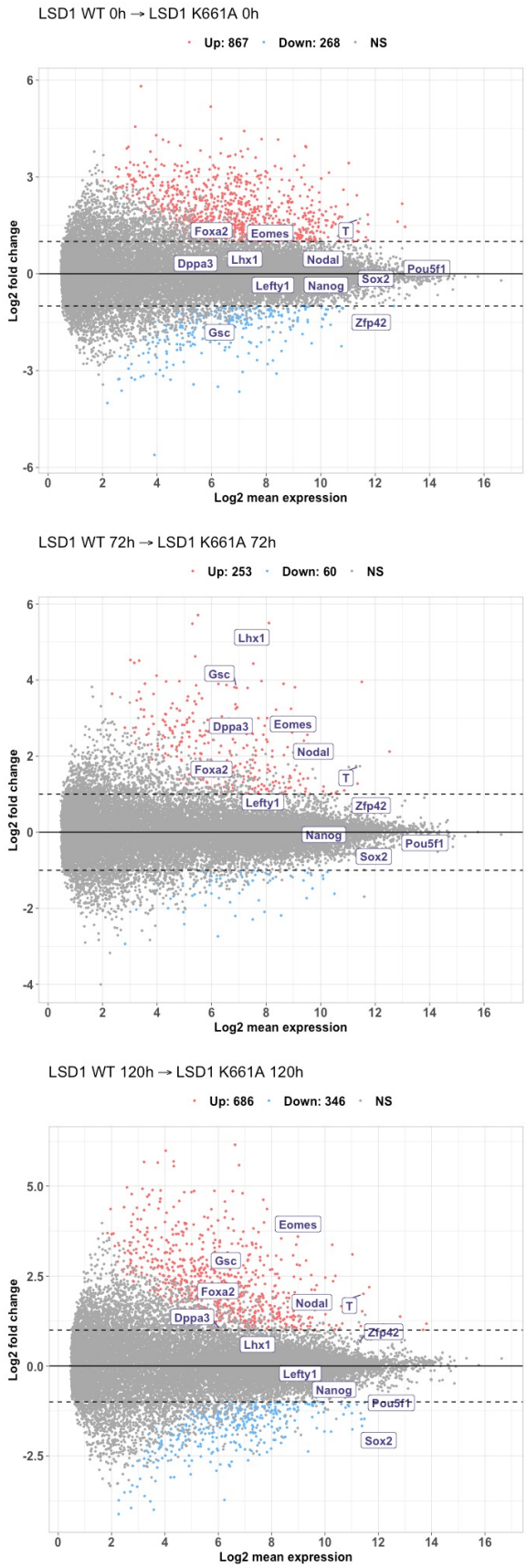


***Figure 5.3 Genes were differentially expressed between Lsd1-K661A and WT-Lsd1 gastruloids at each timepoint.*** *MA plots showing genes that are differentially expressed between Lsd1-K661A and WT-Lsd1 gastruloids at each timepoint, using an adjusted p-value of <0.01 and a log2 fold change of >1/<1. Upregulated genes are highlighted in red and downregulated genes in blue. Black dashed lines indicate the cut offs for log2 fold change.*

Gene ontology analysis was used to determine the terms associated with differentially expressed genes at each timepoint (Figure 5.4). Terms associated with the 0 hour timepoint appeared to primarily represent aspects of cell adhesion and migration, with the top term being cell adhesion, followed by actin cytoskeleton organisation and positive regulation of locomotion. Also enriched were terms: regulation of cell-substrate adhesion, cell-matrix adhesion, and tight junction organisation. Overall, these terms suggest that cells in one of the conditions are more adherent and less migratory than in the other condition, possibly representing a more epithelial cell identity.

At 72 hours, multiple terms relate to development of the circulatory system, including angiogenesis, branching morphogenesis of an epithelial tube, and vascular process in circulatory system. This may suggest this process is impaired or promoted in *Lsd1*-K661A gastruloids. This developmental process was also highlighted in terms at 120 hours, including blood vessel development, morphogenesis of a branching structure, and cardiac chamber development. As well as this, at 120 hours, terms associated with cell adhesion and migration resurfaced, including positive regulation of cell motility, cell-substrate adhesion, extracellular matrix organisation, and regulation of epithelial cell migration.

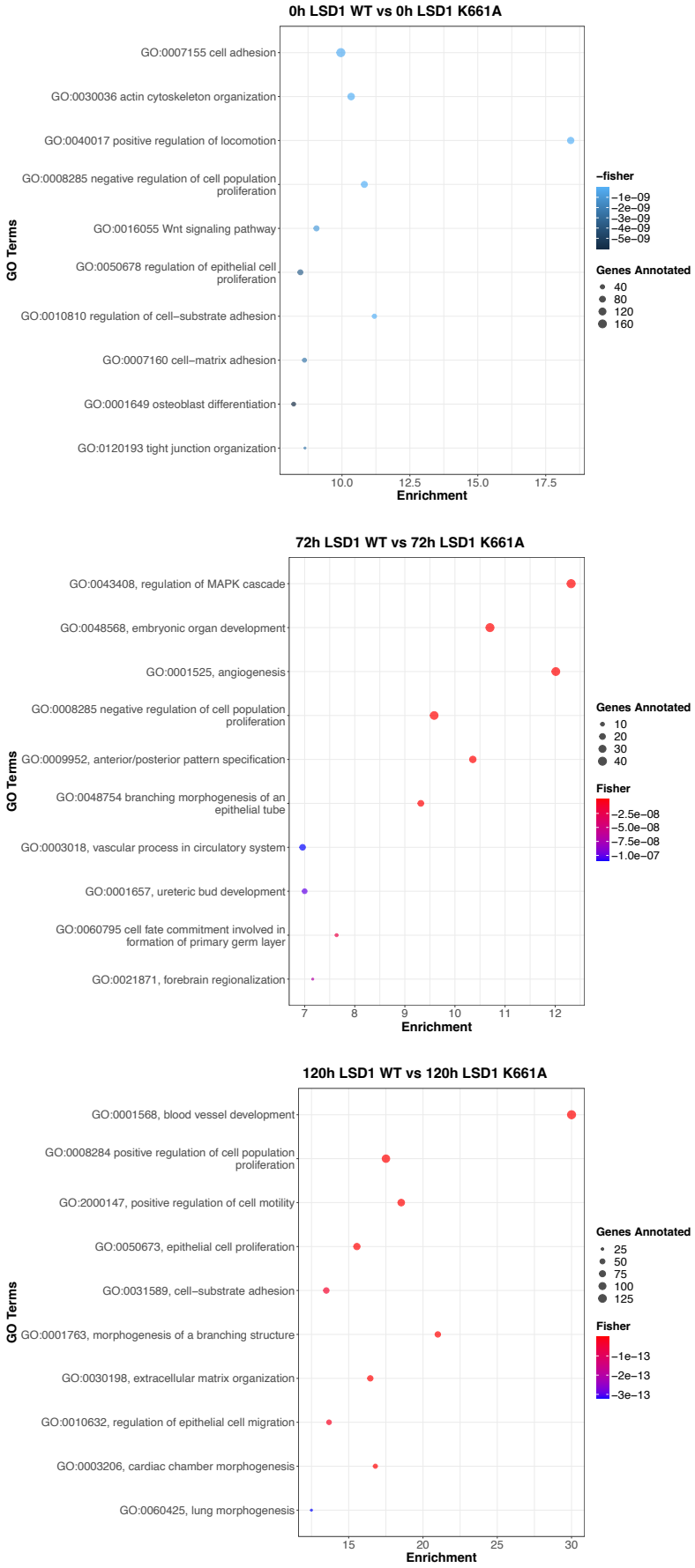


***Figure 5.4 Processes affected by the Lsd1 K661A mutant were analysed through Biological Process Gene Ontology.*** *The top 10 biological process GO terms, that are enriched in Lsd1-K661A relative to WT-Lsd1 gastruloids for each timepoint are shown. GO terms were determined through TopGO analysis (Alexa and Rahnenfuhrer, 2022).*

## 5.4 WT-*Lsd1* gastruloids display impaired differentiation potential

During the analysis, it became apparent that the WT-*Lsd1* gastruloids failed to induce genes associated with early differentiation and gastrulation at the 72 hour timepoint. Despite this, pluripotency markers *Nanog*, *Oct4* and *Sox2* were correctly switched off in a similar fashion to control gastruloids from the first dataset analysed in Chapter 4, suggesting that exit from pluripotency was not impaired. Interestingly, *Sox2* was re-expressed at high levels at 120 hours. *Sox2* is not only expressed in ESCs, but also in multipotent cell populations of endodermal, mesodermal, and ectodermal lineages, where it regulates stem cell characteristics, such as proliferation (Sarkar and Hochedlinger, 2013, Hagey et al., 2018). Therefore, whilst expression of *Sox2* at this timepoint is not unusual in a developmental context, its overexpression in these gastruloids may indicate increased proliferative potential of cells. Early differentiation markers *T*, *Foxa2* and *Fgf5* were expressed at much lower levels at 72h than in any other gastruloid condition (Figure 5.5). The reduced induction of *Brachyury* is particularly interesting, as it has previously been shown that LSD1 negatively regulates *Brachyury* expression (Foster et al., 2010). It is therefore possible that reduced expression of *Brachyury* could be due to overexpression of *Lsd1*. However, expression levels of *Lsd1* were similar at 0 and 72 hour timepoints in WT-*Lsd1* gastruloids when compared to control gastruloids from the first analysed dataset (Figure 5.6).

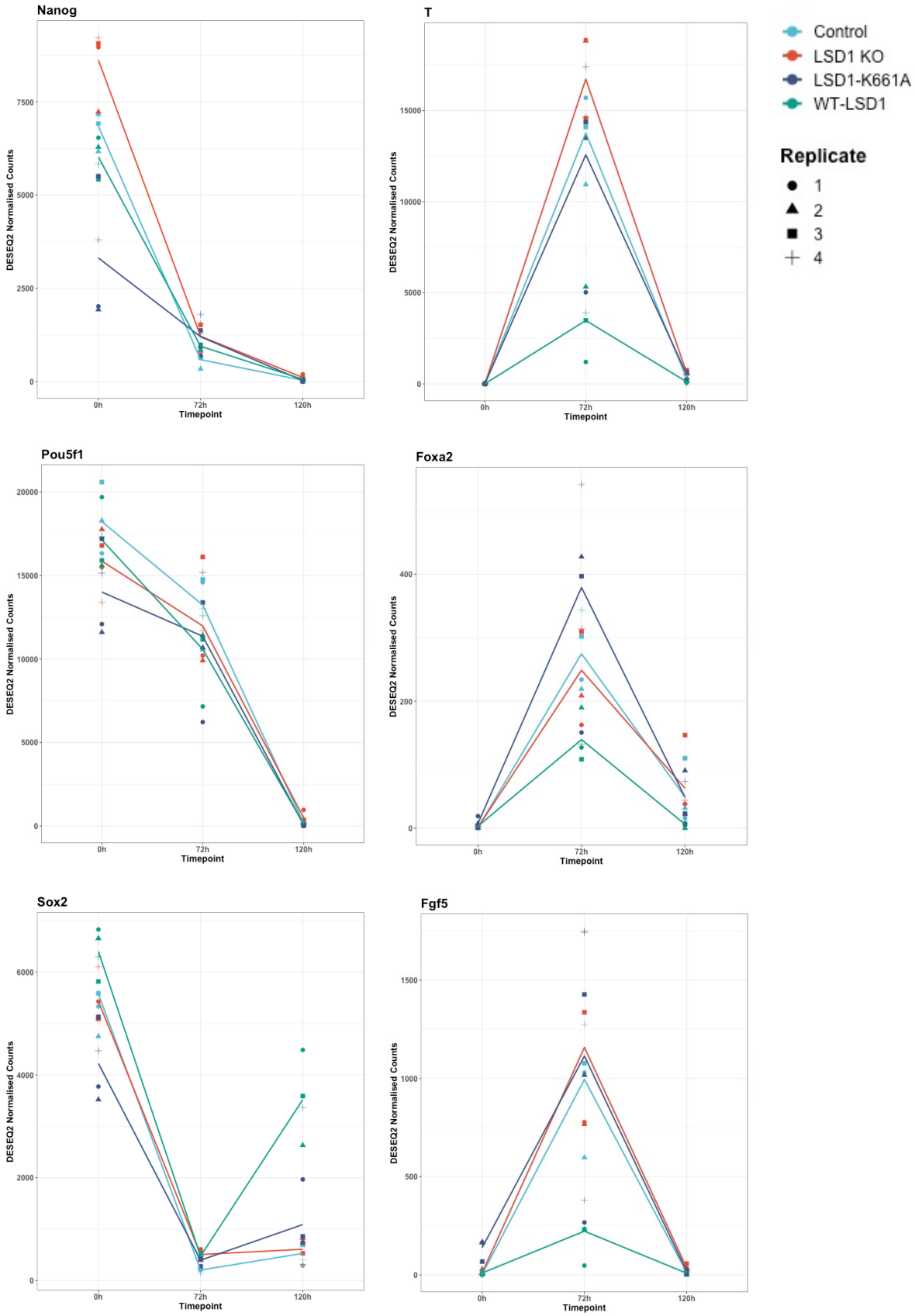


***Figure 5.5 Analysis of pluripotency and early differentiation markers shows limited differentiation potential of WT-Lsd1 gastruloids***. *Line graphs showing expression of genes associated with pluripotency and early differentiation at different timepoints. Points represent values for individual replicates for each gene, lines show mean values of replicates.*

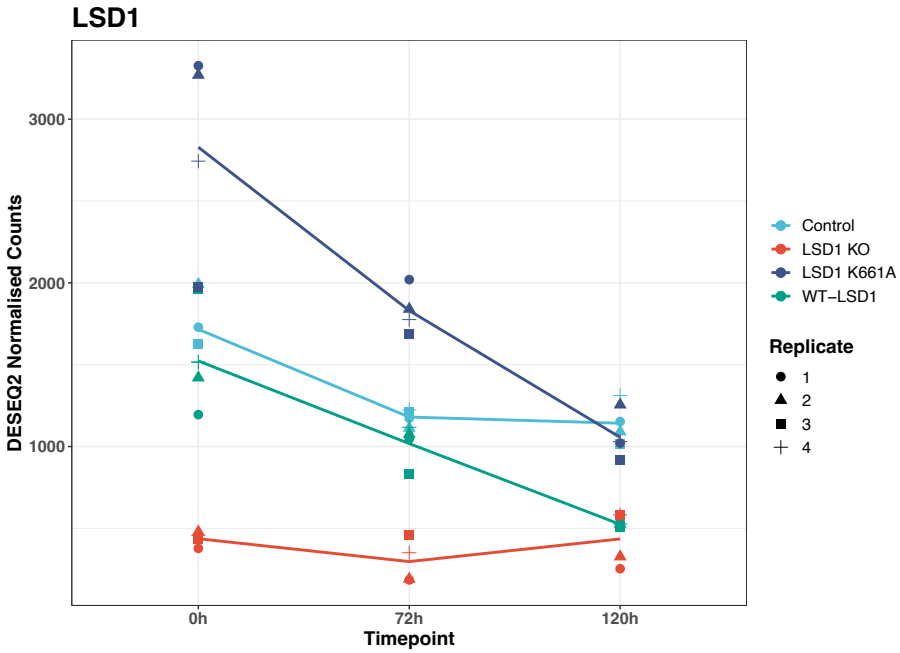


***Figure 5.6 Expression levels of Lsd1 in all gastruloid samples.*** *Line graphs showing expression of LSD1 at different timepoints. Points represent values for individual replicates, lines show mean values of replicates.*

Expression of gastrulation markers *Eomes*, *Gsc*, *Fgf8*, and *Mixl1* are not induced at 72 hours in WT-*Lsd1* gastruloids to the extent seen in other gastruloids (Figure 5.7). In contrast, expression of *Nkx1-2* is dramatically upregulated at this timepoint. Expression of *Foxh1* is lower than in control gastruloids but around the same levels as in gastruloids with *Lsd*1 KO. Together these results show that normal induction of gastrulation associated genes at 72 hours is impaired in WT-*Lsd1* gastruloids. This, in concert with reduced expression of early differentiation genes, suggests that this cell line has reduced differentiation potential.

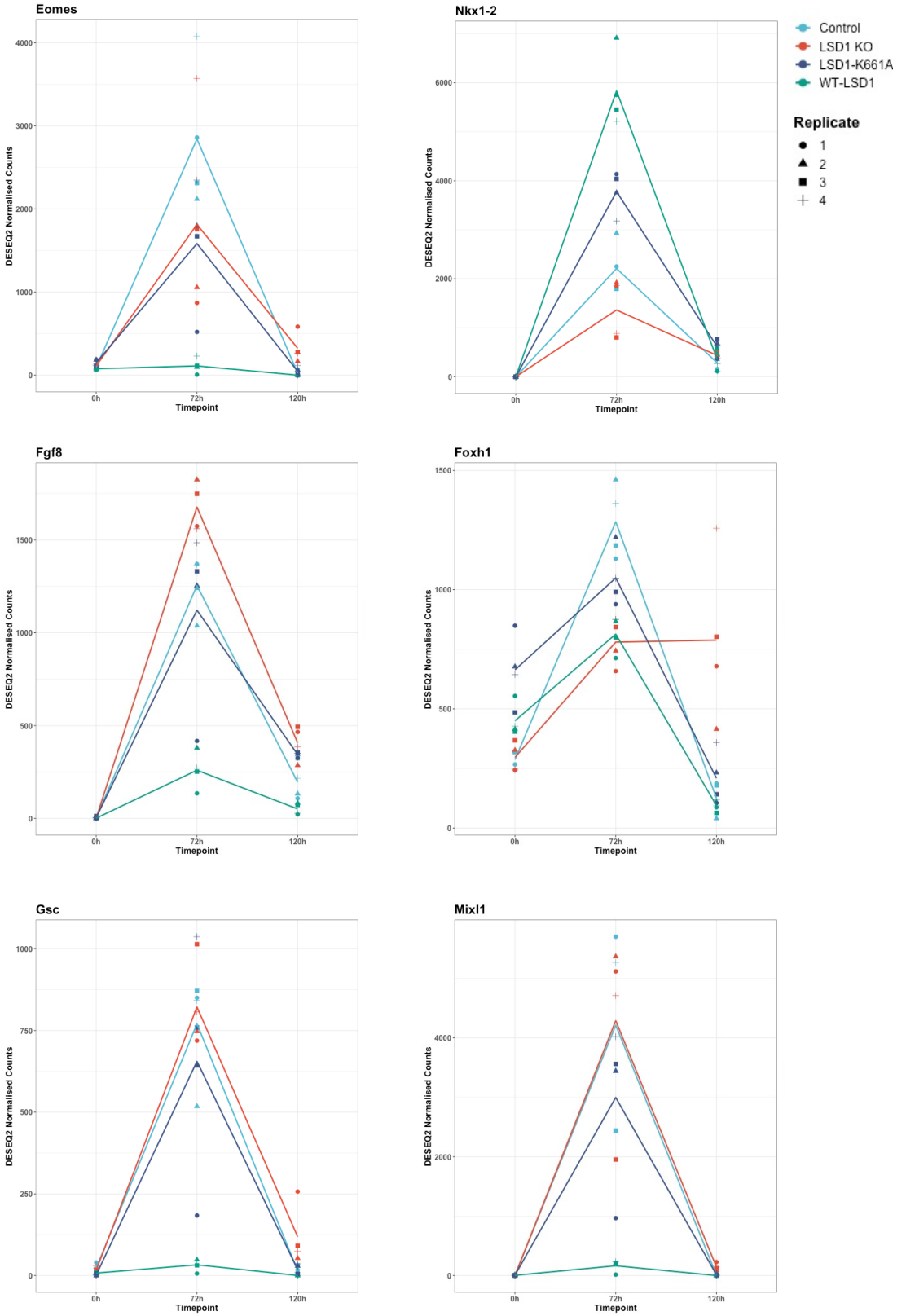


***Figure 5.7 Expression of gastrulation associated genes is impaired in WT-Lsd1 gastruloids.*** *Line graphs showing expression of genes associated with gastrulation at different timepoints. Points represent values for individual replicates for each gene, lines show mean values of replicates.*

## 5.4 Identification of genes whose repression is either dependent on or independent of LSD1 demethylase activity

We compared RNA-seq datasets from control versus *Lsd1* KO (Chapter 4) and our rescue experiments but found there was little overlap of DEGs between the datasets (Figure 5.8). Specifically, only 148, 102 and 254 DEGs were shared between the datasets at 0, 72 and 120 hours, respectively. Since early differentiation markers and markers of gastrulation were similarly expressed in both control and *Lsd1* KO gastruloids at 72 hours, the DEGs that result from WT-*Lsd1* impaired differentiation would not show up in the intersect of DEG lists of each dataset. Therefore, analysis was focused on these intersecting lists in order to identify genes which were dysregulated in both datasets.

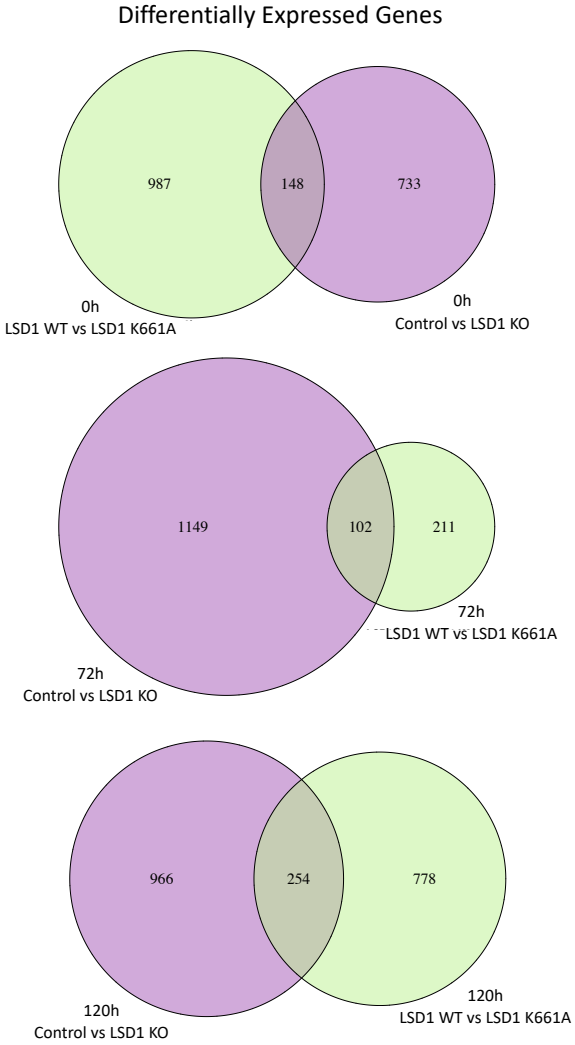


***Figure 5.8 Differentially Expressed Genes show little overlap between the two datasets.*** *Venn diagrams showing the overlapping differentially expressed genes between each timepoint between dataset 1 (Lsd1 KO versus Control gastruloids) and 2 (Lsd1-K661A versus WT-Lsd1 gastruloids).*

Investigating gene lists of the shared DEGs from each dataset allowed identification of genes in three categories: genes which are similarly downregulated/upregulated in both *Lsd1* KO and mutant gastruloids; genes which are upregulated in *Lsd1* mutant gastruloids compared to *Lsd1* KO; genes which are upregulated in gastruloids lacking *Lsd1* but are expressed at control or near control levels in gastruloids with mutant *Lsd1*.

Within the first category, *Klhl13* is upregulated in both *Lsd1* KO and *Lsd1*-K661A gastruloids at 0 hours, suggesting that LSD1 demethylase activity is required for its repression (Figure 5.9A). *Klhl13* encodes an adaptor protein which forms a complex with an Cul3-based E3 ligase to coordinate mitotic progression and cytokinesis (Sumara et al., 2007) and its upregulation may reflect dysregulation of this process. Genes which were downregulated in both knockout and mutant conditions at 0 hours included *Atp1a3*, *Atp2a3* and *Inhbb* (Figure 5.9B). *Atp1a3* encodes a subunit of the $Na^+$,$K^+$-ATPase transport pump, which uses ATP to maintain gradients of these ions across the plasma membrane, and expression of this particular isoform is restricted to neurons (Bøttger et al., 2011). *Atp2a3* similarly encodes an ATP ion pump, namely SERCA3, a member of the Sarcoendoplasmic Reticulum Calcium ATPase (SERCA) family which is expressed in non-muscle cells, such as neuron and epithelial cells (Xu and Van Remmen, 2021). Interestingly, multiple cancer cell lines exhibit decreased *Atp2a3* expression and treatment with HDAC inhibitors upregulates *Atp2a3* expression (Contreras-Leal et al., 2016). At 72 hours, *Fn1* and *Mesp1*, which were previously highlighted in Chapter 4, are upregulated in control gastruloids but not in gastruloids with the mutant or knockout of LSD1 (Figure 5.9B). This indicates that the failure to derepress these genes without LSD1 is not rescued by reintroducing CoREST complex activity, and therefore it is likely that this derepression relies on LSD1 demethylase activity. This is also the case for *Robo2*, a neural cell adhesion molecule that stimulates neurite outgrowth (Hivert et al., 2002), which is upregulated at 120 hours in controls.

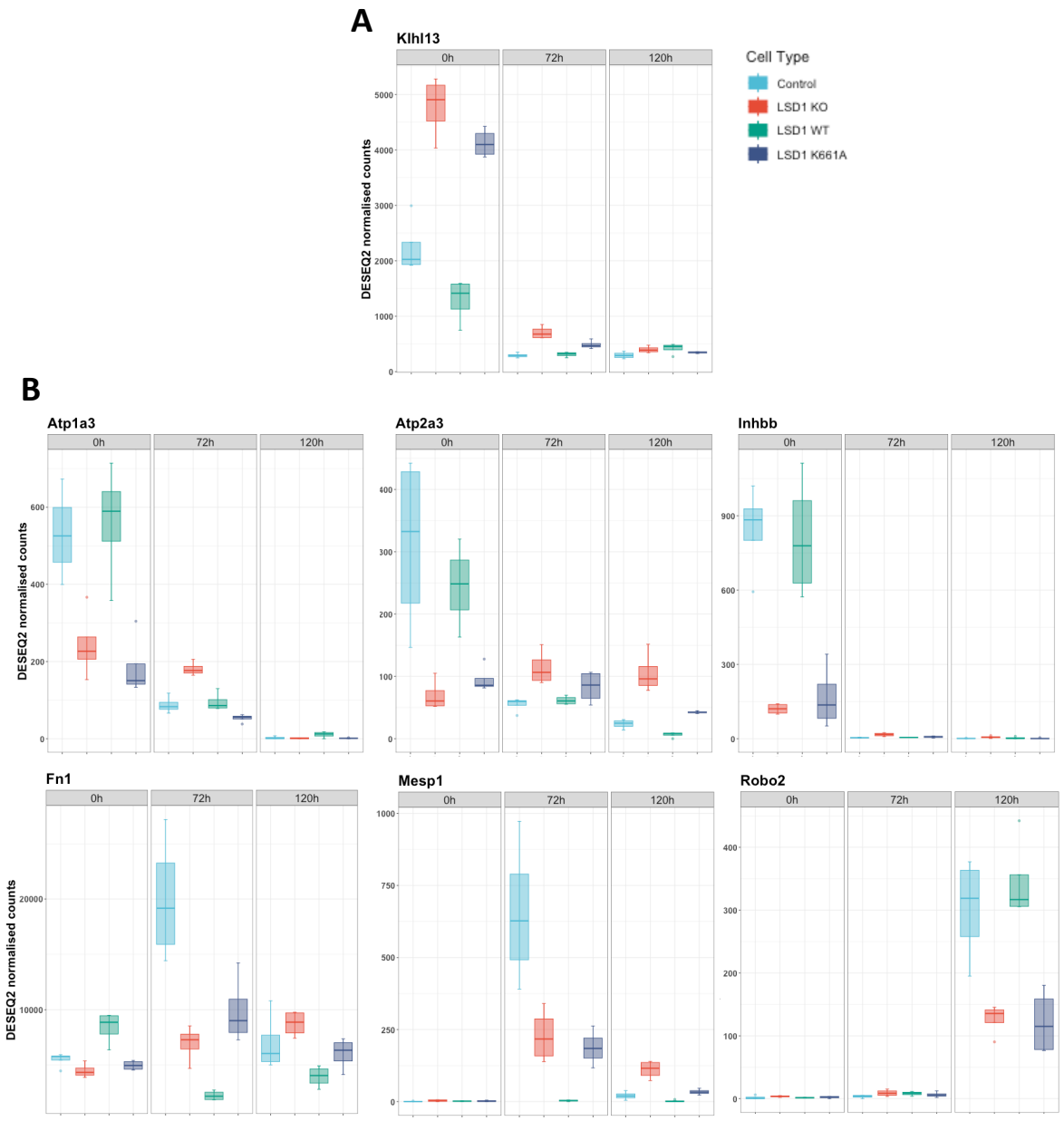


***Figure 5.9 Genes which are similarly downregulated/upregulated in both Lsd1 KO and Lsd1-K661A gastruloids.*** *Boxplots showing expression levels of selected genes in control, Lsd1 KO, WT-Lsd1 and Lsd1-K661A gastruloids, for 0-, 72- and 120-hour timepoints.*

Genes which are upregulated in gastruloids with the *Lsd1* mutant but not in gastruloids lacking *Lsd1* included *Foxd1*, *Foxd2* and *Tcf15*, which are all upregulated at 120 hours (Figure 5.10). Expression of *Tcf15*, as previously mentioned in Chapter 4, is downregulated in *Lsd1* KO gastruloids compared to controls, but upregulated in *Lsd1*-K661A gastruloids compared to both controls and knockouts. *Foxd1* and *Foxd2* show similar expression patterns.

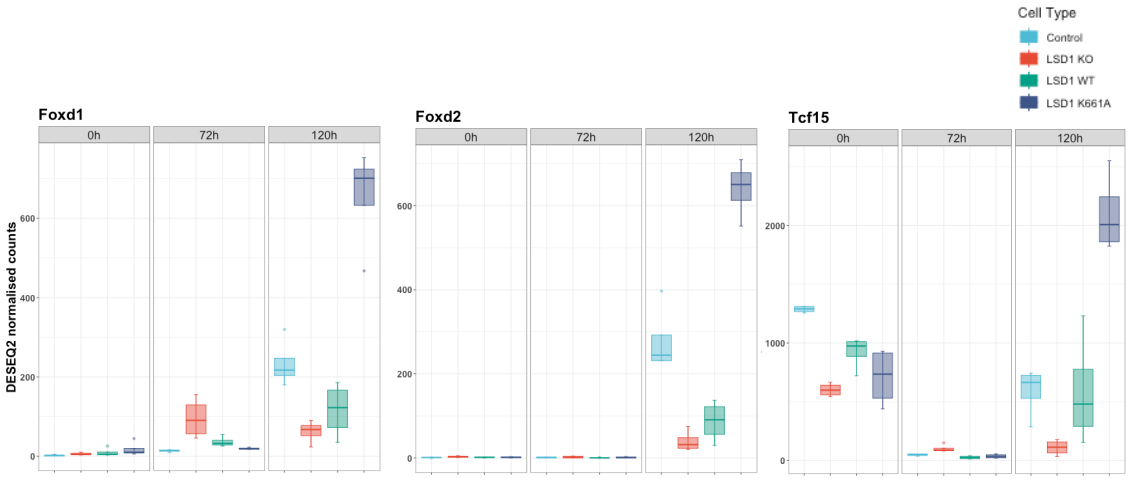


***Figure 5.10 Genes which are upregulated in Lsd1-K661A gastruloids compared to Lsd1 KO.*** *Boxplots showing expression levels of selected genes in control, Lsd1 KO, WT-Lsd1 and Lsd1-K661A gastruloids, for 0-, 72- and 120-hour timepoints.*

A number of genes are upregulated in gastruloids with mutant LSD1 but are expressed near to or below control levels in gastruloids lacking LSD1, suggesting that reintroduction of an intact CoREST complex was able to rescue some effects of *Lsd1* knockout (Figure 5.11). Previously identified genes involved with the BMP pathway and EMT, *Bmp4*, *Bmp7*, *Snai1*, and *Snai2*, were among these genes, suggesting LSD1 demethylase activity is not required for their repression. *Meioc* and *Usp26* were also upregulated at all timepoints following knockout of LSD1 but was expressed at the same or near to same levels as controls in gastruloids with mutant LSD1. Interestingly, expression of *Usp26* in *Lsd1*-K661A gastruloids sat somewhere between control and knockout levels at all timepoints, suggesting that while reinstatement of the CoREST complex was sufficient to partially rescue this effect, LSD1 demethylase activity is required for full repression of this gene.

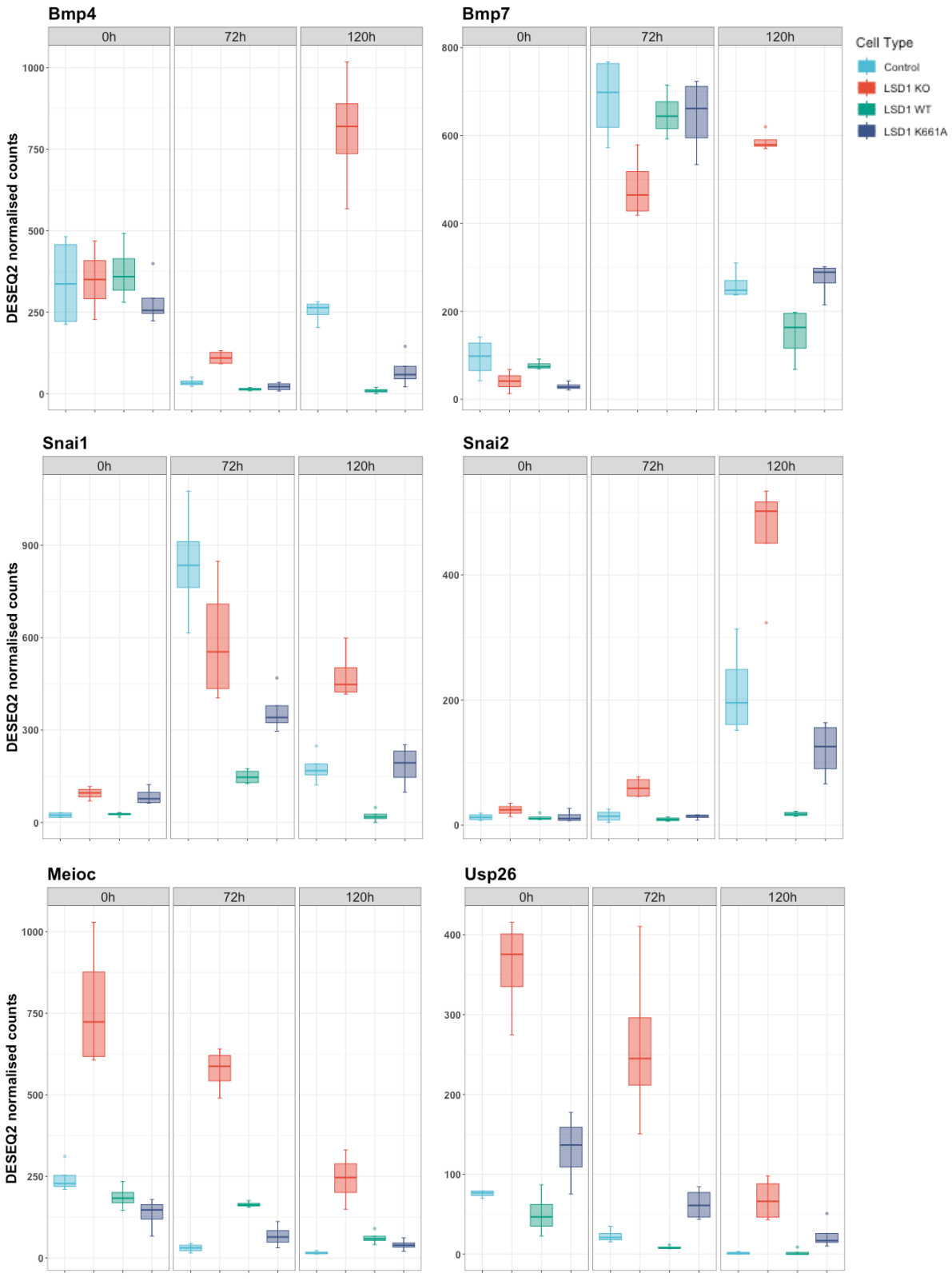


***Figure 5.11 Genes which are upregulated in Lsd1-KO compared to Lsd1-K661A gastruloids.*** *Boxplots showing expression levels of selected genes in control, Lsd1 KO, WT-Lsd1 and Lsd1-K661A gastruloids, for 0-, 72- and 120-hour timepoints*

Members of the *Zscan4* gene cluster, including *Zscan4a*, *Zscan4c*, and *Zsan4d*, also followed this pattern of expression at 0 hours (Figure 5.12A). *Zscan4* is expressed exclusively in late 2-cell embryos and in ESCs (Falco et al., 2007), which corresponds with the expression seen at 0 hours in controls, and this expression is highly upregulated following *Lsd1* knockout. In ESCs, ZSCAN4 is transiently expressed in short bursts, termed Z4 events, which are associated with telomere extension and contribute to genomic stability (Zalzman et al., 2010). Z4 events are accompanied by rapid and transient derepression of constitutive heterochromatin and select regions of facultative heterochromatin (Akiyama et al., 2015). Active histone marks, especially H3K27ac, are enriched at the derepressed heterochromatin during these events and ZSCAN4 forms complexes with chromatin remodellers, including LSD1. The overexpression of *Zscan4* as a result of *Lsd1* KO could therefore result in upregulation of a number of genes which are usually not expressed. ESCs expressing mutant *Lsd1* showed upregulation of *Zscan4*, but not to the same levels as in knockouts, showing that LSD1 demethylase activity contributes to full repression of these genes.

The transient expression of *Zscan4* is important, as some gene expression resulting from extended derepression of heterochromatin can be harmful to cells. The slight upregulation of *Zscan4* in *Lsd1* KO gastruloids seen at 72 hours suggested that repression of the genes is delayed without LSD1. To further investigate this, RT-qPCR was performed at 0, 24, 48 and 72 hour timepoints, which revealed a gradual reduction in gene expression over time, suggesting that correct timing of repression is delayed (Figure 5.12B). Interestingly, this delayed repression was also seen in gastruloids with mutant *Lsd1*, though not to the same extent, with expression seen at 24 hours but not at 48 hours. This further indicates that, although restoration of CoREST complex activity can rescue some repression of these genes, LSD1 demethylase activity is required for return to control levels.

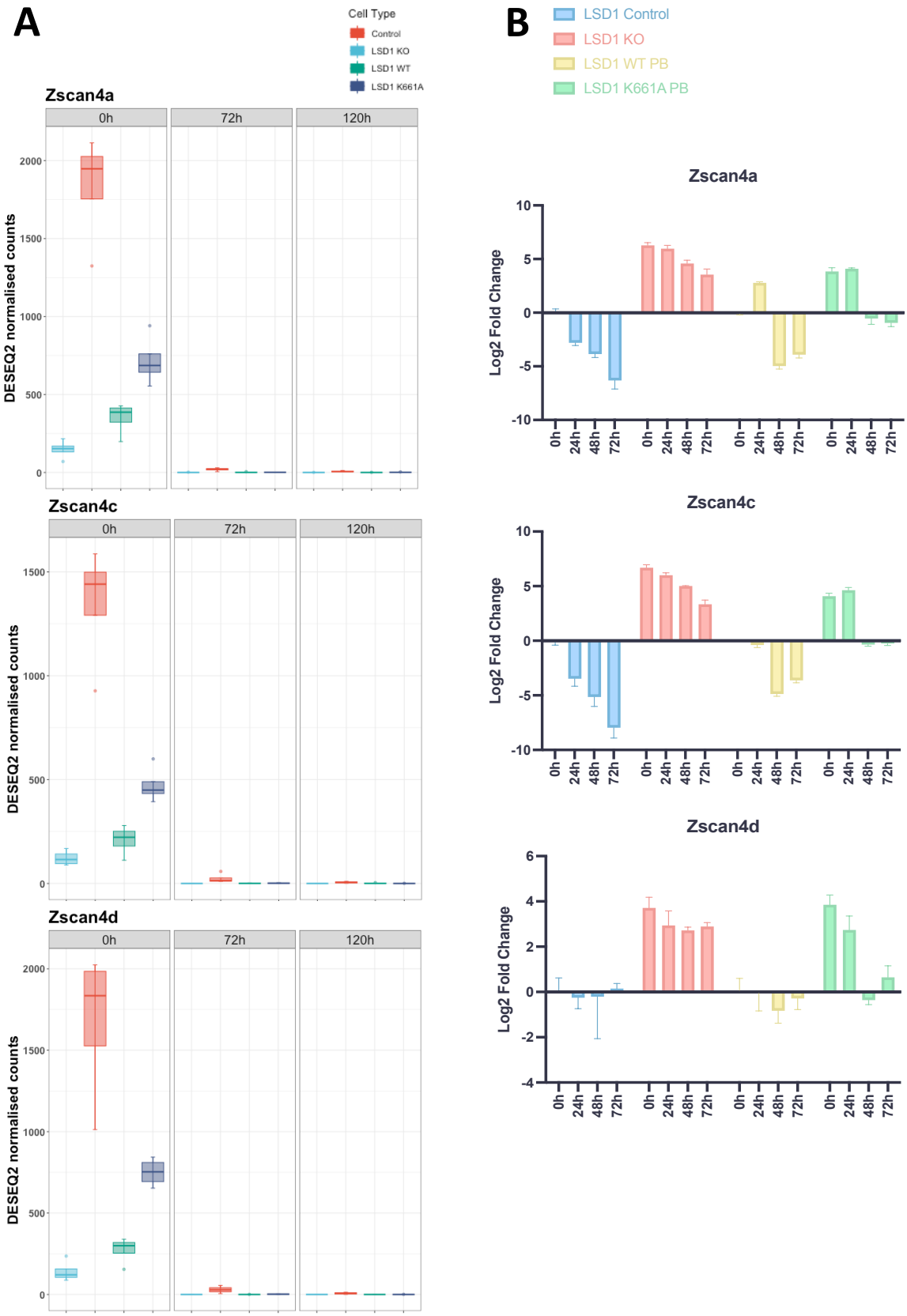


***Figure 5.12 Expression of Zscan4 is upregulated in Lsd1 knockout and mutant gastruloids.*** *(**A**) Boxplots showing expression levels of Zscan4 in control, Lsd1 KO, WT-Lsd1 and Lsd1-K661A gastruloids, for 0-, 72- and 120-hour timepoints. (**B**) RT-qPCR data showing the fold-change in expression levels of Zscan4. Expression levels were normalised to the control gene b-actin (n=3 +/-SD).*

## 5.5 Summary

The results in this chapter have highlighted the effects of the *Lsd1*-K661A mutation on gene expression and allowed some distinction between the requirement of CoREST complex activity versus LSD1 demethylase activity. WT-*Lsd1* cells were found to exhibit compromised differentiation, showing reduced induction of early differentiation and gastrulation associated genes at 72 hours, limiting downstream analysis. Despite the impaired differentiative potential of WT-*Lsd1* cells, a number of genes were shown to either require LSD1 demethylase activity for their correct expression or be rescued, either completely or partially, by reintroduction of the CoREST complex.

# 6 Discussion

## 6.1 Characterisation of the gastruloid model system

The stage at which *Lsd1* loss is lethal in embryos (~E6.5) aligns with the onset of gastrulation, a critical developmental process which involves concerted execution of key processes, including embryo patterning and formation of definitive germ layers. Therefore, we sought to determine the role of *Lsd1* KO in the dysregulation of the highly coordinated process of gastrulation. The *in utero* development of mouse embryos, and particularly the increased difficulty of recovering embryos post-implantation, makes the study of gastrulating embryos in mice challenging. The establishment of a model system which circumvents these limitations and yet closely mimics gastrulation was therefore highly desirable for this project.

The results in Chapter 3 have demonstrated the successful generation of gastruloids from *Lsd1*$^{lox/\Delta 3}$ control ESCs, following troubleshooting of issues encountered when following the protocol. The gastruloids were shown to morphologically replicate gastruloids produced previously in other laboratories (Beccari et al., 2018b), aside from an initiation of elongation at 96 hours rather than the 72 hours previously seen (Figure 3.1). Despite this delay to elongation, it was determined that gastrulation markers were induced at 72 hours, as seen in previous studies (Beccari et al., 2018b) (Figure 3.2). Analysis of RNA-seq data showed that gastruloids exit from pluripotency and express early differentiation markers (Figure 3.5) as well as induce gastrulation markers at 72 hours (Figure 3.6), increasing confidence in the robustness of the model system. The epithelial-to-mesenchymal transition (EMT) that occurs as cells migrate through the primitive streak is essential for the determination of mesendodermal lineages during gastrulation (Tam and Loebel, 2007). An observed decrease in *E-cadherin* and increase in *N-cadherin* as gastruloids progressed suggests that this model recapitulates the EMT seen in the gastrulating embryo (Figure 3.6). Gastruloids were shown to express markers of all three germ layers, although with underrepresentation of ectodermal markers, as

suggested in previous studies (Beccari et al., 2018b) (Figure 3.7). *Hox* genes were expressed in a temporally progressive manner corresponding to their placement in the *Hox* gene cluster, further confirming that gastruloids share similar patterning with embryos (Figure 3.8). Overall, these results suggest that gastruloids were an appropriate model system for investigating the effects of *Lsd1* knockout during gastrulation.

The gastruloid model system is not without its limitations. Gastruloids do not mimic the embryo exactly, they lack extraembryonic tissues which provide signalling to the epiblast in the embryo. For example, expression of BMP4 from the extraembryonic ectoderm restricts the formation of the DVE (distal visceral endoderm) to the distal region of the embryo, a key process for the establishment of the antero-posterior axis (Yamamoto et al., 2009). Despite this, gastruloids display axial patterning similar to the embryo (Beccari et al., 2018b), suggesting there may be some compensatory mechanism for the lack of signalling from extraembryonic tissues. Additionally, *Lsd1* expression has been shown to be restricted to the embryonic portion of the embryo (Foster et al., 2010), therefore the lack of extraembryonic tissue in gastruloids should not impede investigation into *Lsd1* function.

It is also important to consider the maternal contribution of mRNA during embryonic development. It is probable that maternal contribution of *Lsd1* expression compensates for loss of endogenous *Lsd1* in embryos, as depletion of maternal *Lsd1* leads to embryonic arrest at the 1-2 cell stage (Wasson et al., 2016). The half-life of LSD1 protein is relatively long and has been shown to remain in the cell for around 3 to 4 days post knockout (Foster et al., 2010) so it is possible that maternal LSD1 persists until the developmental stage at which loss of LSD1 is lethal. It is therefore likely that maternal contribution of *Lsd1* conceals defects that may occur earlier in embryos and these defects may become apparent in a model system without maternal mRNA. Consequently, investigating *Lsd1* knockout in gastruloids does not exactly mimic the situation in embryos, however, it allows for uncovering of the mechanisms that may contribute to the earlier lethality observed when maternally provided *Lsd1* is absent. For example, the observed increased and sustained expression of *Zscan4* upon *Lsd1* knockout, as evidenced in Figure 5.12, could contribute to this earlier lethality.

Expression of *Zscan4* has been shown to be required for the correct transition from the 2 cell to the 4 cell stage in embryos (Falco et al., 2007). *Zscan4* expression is also associated with derepression of heterochromatin (Akiyama et al., 2015) and, while this process appears to be essential for normal development, compromised heterochromatin organisation generally leads to genomic instability and often appears in cancer (Carone and Lawrence, 2013). Therefore, the prolonged upregulation of *Zscan4* in *Lsd1* knockouts could be a contributing factor to the early embryonic arrest of embryos lacking maternal *Lsd1* (Wasson et al., 2016).

Recent advances in next generation sequencing have generated new methods such as single cell RNA sequencing (scRNA-seq) which enables examination of the transcriptome of single cells (Tang et al., 2009). This technique has been used to investigate the effects of individual gene deletion on gastrulating embryos (Pijuan-Sala et al., 2019). The more sensitive nature of this technique makes it suitable for smaller amounts of input material. While the work in this thesis has employed RNA-seq on pooled gastruloids, interrogating the transcriptome of individual cells would provide both spatial and temporal information.

## 6.2 *Lsd1* knockout gastruloids exhibit dysregulation of developmental processes

The embryonic lethality of *Lsd1* knockout occurs at around E6.5 (Wang et al., 2009a, Wang et al., 2007, Foster et al., 2010), which corresponds approximately with the 72 hour timepoint in gastruloids. It may then be expected that gastruloids without *Lsd1* would not develop past this timepoint, however, this was not the case, as gastruloids continued to develop past this point, albeit with some differences in morphology (Figure 4.1). The abundance of differentially expressed genes at each timepoint, specifically 881, 1251 and 1220 genes at 0 hours, 72 hours, and 120 hours, respectively (Figure 4.3), suggests that the effects of *Lsd1* deletion are widespread, affecting multiple pathways and processes. The majority of these genes were upregulated, consistent with the role of *Lsd1* as a transcriptional repressor. A role of LSD1 in transcriptional activation has

been proposed, through its suggested H3K9 demethylase activity when in association with androgen receptor (AR) (Metzger et al., 2005). Though this activity has yet to be structurally verified, if LSD1 does behave in this way, it could explain the proportion of genes which are downregulated upon *Lsd1* KO. Alternatively, it is possible that LSD1 represses genes which encode other transcriptional repressors, resulting in reduced repression of selected genes in the absence of LSD1. Additionally, Zscan??

Gastruloids lacking *Lsd1* exhibited normal exit from pluripotency and induction of early differentiation markers (Figure 4.4) and gastrulation associated genes at 72 hours (Figure 4.5). However, the downregulation of gastrulation genes at 120 hours was impaired, suggesting a reduced ability to switch these genes off. Although *Lsd1* knockout gastruloids did not fail to develop in a way that might be reminiscent of the lethality in embryos, we identified a number of dysregulated processes which may cumulatively contribute to this phenotype.

### 6.2.1 Differential expression of mesodermal genes

Analysis of lineage specific markers in gastruloids lacking *Lsd1* showed that, although endodermal and ectodermal markers were generally expressed at similar levels as in controls, mesodermal markers were differentially expressed (Figure 4.7). In particular, genes associated with the formation of somites appeared to be downregulated at 120 hours. Somites, which are the originators of skeletal muscle, are formed from paraxial mesoderm in an anterior to posterior fashion (Chal and Pourquie, 2017). Presomitic mesoderm is a transient tissue which is specified though Wnt and FGF signalling pathways, and which can be divided into two regions: the nascent posterior region and the more committed anterior region, from which somites emerge. The presomitic mesoderm expresses Wnt/FGF targets, including *Wnt3a*, *T*, *Tbx6*, and *Msgn1*, each of which is essential for the formation of paraxial mesoderm (Chal and Pourquie, 2017, Nowotschin et al., 2012a). Embryos with Wnt3a$^{-/-}$, Tbx6$^{-/-}$ and Msgn1$^{-/-}$ mutants showed loss of posterior somites and, in the case of Tbx6$^{-/-}$, aberrant formation of anterior somites (Nowotschin et al., 2012a). Therefore, reduced expression of *Tbx6* at

72 hours could result in reduced induction of somite associated genes, *Foxc1* and *Tcf15*, at 120 hours (Figure 4.8). The reduced expression of somite related genes and therefore anticipated dysregulation of somite formation is unlikely to be the cause of *Lsd1* knockout embryonic lethality. This is partly because somitogenesis occurs after the point of lethality in embryos lacking *Lsd1* and because embryos with defective somite formation, in particular as a result of reduced levels of *Tbx6*, do not present a lethal phenotype until E12.5 (White et al., 2003).

### 6.2.2 Dysregulation of the epithelial-to-mesenchymal transition

Following gene ontology analysis of genes differentially expressed in gastruloids lacking *Lsd1,* the prevalence of terms associated with cell adhesion prompted investigation into this pathway (Figure 4.11). Genes associated with EMT were among the most upregulated and downregulated genes within this term at 72 and 120 hours (Figure 4.12). This encouraged investigation into expression levels of key EMT associated genes, many of which were dysregulated (Figure 4.13). The upregulation of *Snail*, *Slug*, *Vimentin*, *MIAT*, and *Eomesodermin*, all of which are factors that stimulate EMT, would suggest increased transition towards mesenchymal cell identity. Conversely, the downregulation of *Fgfr1*, *Fibronectin*, *Foxc2* and *Pdgfra*, which all similarly stimulate EMT, would suggest the opposite. It is therefore difficult to dissect whether *Lsd1* knockout leads to an overall promotion or suppression of the transition to mesenchymal cell identity, but it is clear that it contributes to dysregulation of the process. The switch from expression of *E-cadherin* to expression of *N-cadherin* is a hallmark of EMT (Loh et al., 2019). It has also been previously shown that chemical inhibition of LSD1 in colon cancer cells leads to upregulation of *E-cadherin* and downregulation of *N-cadherin*, and that the upregulation is a result of LSD1 demethylase activity at the promoter of *E-cadherin* (Ding et al., 2013). Consequently, it is unusual that there were no observed changes in the expression levels of *E-cadherin* and *N-cadherin* following *Lsd1* depletion despite changes in factors which regulate the cadherin switch, such as *Snai1*.

The process of EMT during gastrulation is essential for the ingression of cells through the primitive streak to form mesodermal and endodermal lineages. As well as this, EMT can also occur in the adult, either in tissue regeneration and organ fibrosis processes, or in cancer invasion and metastasis (Kalluri and Weinberg, 2009). Though each of these types of EMT occur in distinct biological contexts, they share underlying genetic mechanisms. *Lsd1* has previously been implicated in regulating the process of EMT, both through epigenetic modulation of EMT related genes and through association with EMT regulators within the CoREST complex (Ambrosio et al., 2017).

*Lsd1* overexpression is present in cancer cells, and this contributes to increased proliferation, cell motility and EMT (Hino et al., 2016, Hayami et al., 2011, Lv et al., 2012). Inhibition of *Lsd1* has been shown to suppress proliferation, migration and invasion in prostate, lung, and colon cancer cells (Wang et al., 2015, Lv et al., 2012, Ding et al., 2013). H3K4me3 levels increase and H3K9me2 decrease during TGF-β stimulated EMT in AML12 hepatocytes and these chromatin modifications were blocked upon *Lsd1* knockdown, suggesting LSD1 plays an important role in chromatin reprogramming during EMT (McDonald et al., 2011). Interestingly, in this study, the normal decrease in *E-cadherin* was observed following siRNA knockdown of *Lsd1*, similar to the results presented here.

LSD1 physically interacts with mediators of EMT, such as UTX, Snail, and Slug (Choi et al., 2015, Lin et al., 2010a, Ferrari-Amorotti et al., 2013). The association of the CoREST complex with UTX has been shown to repress EMT-inducing genes, *Snai1*, *Zeb1* and *Zeb2* in breast cancer stem cells (Choi et al., 2015). The upregulation of *Snai1* in our results could be due to loss of this activity. In contrast, the association of LSD1 with Snail recruits its repressive activity to epithelial related genes, including *E-cadherin (Lin et al., 2010a)*. Consequently, loss of transcriptional repressive activity of the CoREST complex at epithelial related genes results in increased epithelial cell identity and reduced EMT. While some of our results are in agreement with this, for example the downregulation of genes which stimulate EMT (*Fgfr1*, *Fibronectin*, *Foxc2* and *Pdgfra*), the upregulation of mesenchymal markers and inducers of EMT (*Snail*, *Slug*, *Vimentin*, *MIAT*, and *Eomesodermin*) is in stark contrast to this, perhaps relating to loss of UTX mediated gene

repression. This presents a more complicated situation of the role of LSD1 in EMT regulation. ChIP-seq analysis to interrogate the levels of H3K4 methylation levels at EMT related genes upon *Lsd1* knockout could aid in deciphering the role of LSD1 in this process, but this was beyond the scope of the work in this thesis.

### 6.2.3 Differential expression of BMP pathway related genes

Gene ontology analysis showed a number of differentially expressed genes were associated with terms relating to the BMP pathway, including cellular response to BMP stimulus and pathway-restricted SMAD protein phosphorylation (Figure 4.11). Investigation into genes related to these terms revealed dysregulation of a number of genes relating to the BMP pathway (Figure 4.14). Although upregulation of *Bmp4* and *Bmp7* may suggest increased BMP pathway activity, the downregulation of downstream effector *Smad5* and upregulation of inhibitory SMADs, *Smad6* and *Smad7*, suggests the opposite. This is apparent in the reduced expression of *Sfrp1*, which is usually expressed in regions with high BMP activity, and increased expression of *Bambi*, which is usually expressed in regions with low BMP expression. The expression of these genes, important in the specification of the dorsal-ventral axis, is mediated by pSMAD5 (Greenfeld et al., 2021), and reduced expression of *Smad5*, as well as increased expression of inhibitive SMADs which may prevent its phosphorylation, could be the cause of their differential expression.

The first established role of the BMP proteins was in bone and cartilage formation, and were hence given the nomenclature bone morphogenic proteins (BMPs) (Zhao, 2003). Aside from this developmental role, the BMP pathway regulates many processes in embryonic development, including cavitation of the ICM, axis formation and organogenesis (Mishina, 2003). Firstly, in the determination of the antero-posterior axis, expression of BMP4 from the extraembryonic ectoderm inhibits the formation of the DVE, restricting it to the distal pole of the embryo (Yamamoto et al., 2009). Although gastruloids do not include extraembryonic tissues, they do possess an antero-posterior axis (Beccari et al., 2018b), and therefore overexpression of *Bmp4* in *Lsd1* knockout

gastruloids could result in perturbation of the specification of this axis. *Bmp4* expression from the extraembryonic ectoderm is also required for the generation of primordial germ cells (PGCs) in the embryo (Lawson et al., 1999). Again, while extraembryonic tissues are not present in gastruloids, the upregulation of *Bmp4* could account for the differentially expressed genes associated with GO terms relating to germ cells, including male meiotic nuclear division, DNA methylation involved in gamete generation and piRNA metabolic process (Figure 4.11). During embryogenesis, BMP7 forms heterodimers with BMP2 and BMP4, the presence of which are required for the correct formation of the heart (Kim et al., 2019b). The upregulation of *Bmp7* and *Bmp4* following *Lsd1* depletion could therefore contribute to the differentially expressed genes in the identified GO terms relating to cardiovascular development, including vascular process in circulatory system, blood vessel morphogenesis and cardiac septum morphogenesis.

Due to the involvement of the BMP pathway in the establishment of the dorsoventral and antero-posterior axes in embryos, it would be useful to gain spatial expression information of genes including *Bmp4* during gastruloid development. This could be achieved through scRNA-seq or alternatively through direct visualisation methods, such as hybridisation chain reaction (HCR) experiments. HCR involves hybridisation of an mRNA of choice with multiple RNA probes which present an initiator sequence that can then be targeted by fluorescent DNA hairpins (Tsuneoka and Funato, 2020). Interrogation of spatial expression of, for example *Bmp4*, would give insight into whether the upregulation of *Bmp4* in *Lsd1* KO gastruloids is due to overexpression in cells which already express the gene in controls. Alternatively, the increased expression may be in cells which do not normally express *Bmp4* in control gastruloids, thereby potentially affecting gradients of the gene that enable establishment of axes.

## 6.3 LSD1 demethylase activity is required for differential expression of some identified genes, but dispensable for others

CoREST complex stability is dependent on LSD1 and knockout of *Lsd1* therefore results in loss of CoREST complex activity as well as LSD1 demethylase activity. Consequently, the DEGs identified in Chapter 4 could be a result of loss of either direct or indirect functionalities. To discriminate between these possibilities, gastruloids were generated from cells with the catalytically inactive mutant *Lsd1*-K661A and cells with WT-*Lsd1* as a control. Morphologically, these gastruloids were similar and resembled *Lsd1* KO gastruloids from Chapter 4, despite reduced elongation potential presented by WT-*Lsd1* gastruloids (Figure 5.1). This was likely due to the reduced differentiation capacity of these cells, as evidenced in Chapter 5.4. Totals of 1135, 313, and 1032 DEGs were identified between gastruloid conditions at 0 hours, 72 hours and 120 hours, respectively, with the majority of genes at each timepoint being upregulated, which aligns with the repressive nature of LSD1 demethylase activity (Figure 5.3).

The impaired capability of WT-*Lsd1* cells to differentiate was demonstrated through reduced induction of early differentiation (Figure 5.5) and gastrulation (Figure 5.7) associated genes. Examination of *Lsd1* expression levels suggested that overexpression of *Lsd1* was not the cause of this (Figure 5.6). The *piggyBac* transposon system, which was used to originally generate these cells, can carry and integrate genes into the genome via a 'cut and paste' mechanism (Ding et al., 2005). *piggyBac* mediated integration is biased towards transcriptional units and there is potential for insertional mutagenesis of genes using this system. It is therefore possible that insertional mutagenesis is responsible for the impaired differentiation capability of this cell line, and determination of integration sites by PCR could be used to ascertain this.

The overlap between DEGs in *Lsd1* KO versus control gastruloids and WT-*Lsd1* versus *Lsd1*-K661A gastruloid datasets was relatively small (Figure 5.8). The number of genes differentially expressed in the dataset with the mutant variant of *Lsd1* is surprising, as

the restoration of the CoREST complex would be expected to rescue expression of some DEGs from the first dataset, and not to dysregulate unrelated genes. This result could be due in part to the decreased differentiation phenotype of WT-*Lsd1* cells, or perhaps due to insertional mutagenesis of genes in *Lsd1*-K661A cells. Genes were also identified which were upregulated in gastruloids with mutant *Lsd1* but not with *Lsd1* KO (Figure 5.10), suggesting that genes with this expression pattern may also contribute to this subset of differentially expressed genes. Despite the reduced differentiation defect seen in gastruloids with exogenous WT *Lsd1*, a selection of genes which require, or conversely do not require, LSD1 demethylase activity for their repression or expression were identified by examining genes which were dysregulated in both datasets.

Reintroduction of CoREST complex functionality was unable to rescue the overexpression of *Klhl13* and the reduced expression of *Atp1a3*, *Atp2a3, Inhbb, Fn1, Mesp1,* and *Robo2* (Figure 5.9). LSD1 demethylase activity therefore appears to be required for repression of *Klhl13* and for the activation of the other genes. The requirement for LSD1 for this activation is presumably because of an indirect effect, potentially due to LSD1 mediated transcriptional repression of transcription factors or chromatin modifiers which would otherwise act to repress expression of these genes. There are likely more genes which have a direct requirement for LSD1 demethylase activity, which we have been unable to identify here due to limitations presented by the unanticipated WT-*Lsd1* cell phenotype.

A number of genes were identified whose expression was rescued or partially rescued in gastruloids with CoREST complex activity but not LSD1 demethylase activity (Figure 5.11, Figure 5.12). Among the genes whose expression levels were rescued were genes associated with the BMP pathway, *Bmp4* and *Bmp7*, as well as with EMT, *Snai1* and *Snai2*, suggesting regulation of these genes rely on CoREST complex repressive activity but not on LSD1 demethylase activity. This is not unexpected in the case of *Snai1* and *Snai2,* whose SNAG domains interact directly with the active domain of LSD1 (Lin et al., 2010b). The partial restoration of expression levels of *Usp26* and *Zscan4* genes showed that LSD1 activity is required for full repression of these genes to control levels. The dispensability of LSD1 catalysis for the repression of these genes could be due to a number of reasons. Firstly, and perhaps most simply, the histone modification state of

these genes when active may include histone acetylations specifically facilitated by the CoREST complex but not H3K4me1/2, making LSD1 activity redundant for their repression. Alternatively, there are a number of lysine demethylases that are able to demethylate H3K4me2, including LSD2, KDM5A, KDM5B, KDM5C, and KDM5D (Hyun et al., 2017), which may be able to compensate for the loss of LSD1 demethylase activity. Another reason may be explained by the results of a recent study which demonstrated that the observed catalytically inactive status of the K661A mutation in *in vitro* conditions may not reflect its status *in vivo* (Kim et al., 2020). Assessment of the demethylase activity of the K661A mutant on nucleosomal substrates including extranucleosomal DNA revealed that the mutant retains around 20% of the activity of WT *Lsd1*. This would explain the observed partial rescue of selected genes in gastruloids with the K661A mutant, as the retained 20% of WT activity may enable full repression of some genes but only partial repression of others.

The downstream analysis was severely limited by the unexpected phenotype of WT-*Lsd1* gastruloids and the potential retained demethylase activity in the *Lsd1*-K661A gastruloids. Repeating these experiments with another WT-*Lsd1* cell line and with the double LSD1 mutation A539E/K661A, which has undetectable demethylase activity on nucleosomal substrates (Kim et al., 2020), would circumvent these limitations.

## 6.4 Summary

The work in this thesis has characterised the role of LSD1 during gastrulation by employing gastruloids to model this developmental stage. Loss of LSD1 was shown to dysregulate expression of genes associated with mesodermal lineages, EMT and the BMP pathway; the cumulative effects of which may contribute to the embryonic lethal phenotype of *Lsd1* knockout. The increased and sustained expression of *Zscan4* observed upon loss of *Lsd1*, partially rescued by the reintroduction of the CoREST complex, could be particularly deleterious during embryonic development. Despite limitations in our analysis, we identified a number of genes whose repression appeared to be dependent on or independent of LSD1 demethylase activity which warrant further investigation.

# References


ABRAJANO, J. J., QURESHI, I. A., GOKHAN, S., ZHENG, D., BERGMAN, A. & MEHLER, M. F. 2009. REST and CoREST Modulate Neuronal Subtype Specification, Maturation and Maintenance. *PLOS ONE,* 4**,** e7936.

ABU-FARHA, M., LAMBERT, J.-P., AL-MADHOUN, A. S., ELISMA, F., SKERJANC, I. S. & FIGEYS, D. 2008. The Tale of Two Domains: Proteomics and Genomics Analysis of SMYD2, A New Histone Methyltransferase *. *Molecular & Cellular Proteomics,* 7**,** 560-572.

ACAMPORA, D., DI GIOVANNANTONIO, L. G. & SIMEONE, A. 2013. Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development,* 140**,** 43-55.

ADAMO, A., SESE, B., BOUE, S., CASTANO, J., PARAMONOV, I., BARRERO, M. J. & IZPISUA BELMONTE, J. C. 2011. LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. *Nat Cell Biol,* 13**,** 652-9.

AGARWAL, S., BONEFAS, K. M., GARAY, P. M., BROOKES, E., MURATA-NAKAMURA, Y., PORTER, R. S., MACFARLAN, T. S., REN, B. & IWASE, S. 2021. KDM1A maintains genome-wide homeostasis of transcriptional enhancers. *Genome Res,* 31**,** 186-97.

AKIYAMA, T., XIN, L., ODA, M., SHAROV, A. A., AMANO, M., PIAO, Y., CADET, J. S., DUDEKULA, D. B., QIAN, Y., WANG, W., KO, S. B. H. & KO, M. S. H. 2015. Transient bursts of Zscan4 expression are accompanied by the rapid derepression of heterochromatin in mouse embryonic stem cells. *DNA Research,* 22**,** 307-318.

ALABDI, L., SAHA, D., HE, M., DAR, M. S., UTTURKAR, S. M., SUDYANTI, P. A., MCCUNE, S., SPEARS, B. H., BREEDLOVE, J. A., LANMAN, N. A. & GOWHER, H. 2020. Oct4-Mediated Inhibition of Lsd1 Activity Promotes the Active and Primed State of Pluripotency Enhancers. *Cell Rep,* 30**,** 1478-1490 e6.

ALEXA, A. & RAHNENFUHRER, J. 2022. topGO: Enrichment Analysis for Gene Ontology. *R package version 2.50.0*.

ALLEN, H. F., WADE, P. A. & KUTATELADZE, T. G. 2013. The NuRD architecture. *Cell Mol Life Sci,* 70**,** 3513-24.

ALQARNI, S. S., MURTHY, A., ZHANG, W., PRZEWLOKA, M. R., SILVA, A. P., WATSON, A. A., LEJON, S., PEI, X. Y., SMITS, A. H., KLOET, S. L., WANG, H., SHEPHERD, N. E., STOKES, P. H., BLOBEL, G. A., VERMEULEN, M., GLOVER, D. M., MACKAY, J. P. & LAUE, E. D. 2014. Insight into the architecture of the NuRD complex: structure of the RbAp48-MTA1 subcomplex. *J Biol Chem,* 289**,** 21844-55.

AMBROSIO, S., SACCÀ, C. D. & MAJELLO, B. 2017. Epigenetic regulation of epithelial to mesenchymal transition by the Lysine-specific demethylase LSD1/KDM1A. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms,* 1860**,** 905-910.

ANDRÉS, M. E., BURGER, C., PERAL-RUBIO, M. J., BATTAGLIOLI, E., ANDERSON, M. E., GRIMES, J., DALLMAN, J., BALLAS, N. & MANDEL, G. 1999. CoREST: A functional corepressor required for regulation of neural-specific gene expression. *Proceedings of the National Academy of Sciences,* 96**,** 9873-9878.

ARENTS, G. & MOUDRIANAKIS, E. N. 1993. Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. *Proceedings of the National Academy of Sciences,* 90**,** 10489-10493.

ARNOLD, S. J., HOFMANN, U. K., BIKOFF, E. K. & ROBERTSON, E. J. 2008. Pivotal roles for eomesodermin during axis formation,epithelium-to-mesenchyme transition and endoderm specification in the mouse. *Development,* 135**,** 501-511.

BALLAS, N., GRUNSEICH, C., LU, D. D., SPEH, J. C. & MANDEL, G. 2005. REST and Its Corepressors Mediate Plasticity of Neuronal Gene Chromatin throughout Neurogenesis. *Cell,* 121**,** 645-657.

BANCK, M. S., LI, S., NISHIO, H., WANG, C., BEUTLER, A. S. & WALSH, M. J. 2009. The ZNF217 oncogene is a candidate organizer of repressive histone modifiers. *Epigenetics,* 4**,** 100-106.

BANNISTER, A. J. & KOUZARIDES, T. 2005. Reversing histone methylation. *Nature,* 436**,** 1103-6.

BANNISTER, A. J. & KOUZARIDES, T. 2011. Regulation of chromatin by histone modifications. *Cell Res,* 21**,** 381-95.

BANNISTER, A. J., SCHNEIDER, R. & KOUZARIDES, T. 2002. Histone Methylation: Dynamic or Static? *Cell,* 109**,** 801-806.

BANNISTER, A. J., ZEGERMAN, P., PARTRIDGE, J. F., MISKA, E. A., THOMAS, J. O., ALLSHIRE, R. C. & KOUZARIDES, T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature,* 410**,** 120-124.

BANTSCHEFF, M., HOPF, C., SAVITSKI, M. M., DITTMANN, A., GRANDI, P., MICHON, A.-M., SCHLEGL, J., ABRAHAM, Y., BECHER, I., BERGAMINI, G., BOESCHE, M., DELLING, M., DÜMPELFELD, B., EBERHARD, D., HUTHMACHER, C., MATHIESON, T., POECKEL, D., READER, V., STRUNK, K., SWEETMAN, G., KRUSE, U., NEUBAUER, G., RAMSDEN, N. G. & DREWES, G. 2011. Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nature Biotechnology,* 29**,** 255-265.

BARDOT, E. S. & HADJANTONAKIS, A. K. 2020. Mouse gastrulation: Coordination of tissue patterning, specification and diversification of cell fate. *Mech Dev,* 163**,** 103617.

BARNES, C. 2018. *Investigating the role of Lysine Specific Demethylase 1 in early embryonic stem cell differentiation.* PhD, University of Leicester.

BARNES, C. E., ENGLISH, D. M., BRODERICK, M., COLLINS, M. O. & COWLEY, S. M. 2022. Proximity-dependent biotin identification (BioID) reveals a dynamic LSD1–CoREST interactome during embryonic stem cell differentiation. *Molecular Omics,* 18**,** 31-44.

BARON, R. & VELLORE, N. A. 2012. LSD1/CoREST is an allosteric nanoscale clamp regulated by H3-histone-tail molecular recognition. *Proc Natl Acad Sci U S A,* 109**,** 12509-14.

BARRIOS, A. P., GOMEZ, A. V., SAEZ, J. E., CIOSSANI, G., TOFFOLO, E., BATTAGLIOLI, E., MATTEVI, A. & ANDRES, M. E. 2014. Differential properties of transcriptional complexes formed by the CoREST family. *Mol Cell Biol,* 34**,** 2760-70.

BARROW, J. R., HOWELL, W. D., RULE, M., HAYASHI, S., THOMAS, K. R., CAPECCHI, M. R. & MCMAHON, A. P. 2007. Wnt3 signaling in the epiblast is required for proper orientation of the anteroposterior axis. *Developmental Biology,* 312**,** 312-320.

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

BECCARI, L., GIRGIN, M., TURNER, D. A., BAILLIE-JOHNSON, P., COSSY, A.-C., BECCARI, L., MORIS, N., LUTOLF, M., DUBOULE, D. & MARTINEZ ARIAS, A. 2018a. Generating Gastruloids from Mouse Embryonic Stem Cells. *Protocol Exchange*.

BECCARI, L., MORIS, N., GIRGIN, M., TURNER, D. A., BAILLIE-JOHNSON, P., COSSY, A. C., LUTOLF, M. P., DUBOULE, D. & ARIAS, A. M. 2018b. Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature,* 562**,** 272-276.

BERNSTEIN, B. E., KAMAL, M., LINDBLAD-TOH, K., BEKIRANOV, S., BAILEY, D. K., HUEBERT, D. J., MCMAHON, S., KARLSSON, E. K., KULBOKAS, E. J., 3RD, GINGERAS, T. R., SCHREIBER, S. L. & LANDER, E. S. 2005. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell,* 120**,** 169-81.

BERNSTEIN, B. E., MIKKELSEN, T. S., XIE, X., KAMAL, M., HUEBERT, D. J., CUFF, J., FRY, B., MEISSNER, A., WERNIG, M., PLATH, K., JAENISCH, R., WAGSCHAL, A., FEIL, R., SCHREIBER, S. L. & LANDER, E. S. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell,* 125**,** 315-26.

BLACK, J. C., VAN RECHEM, C. & WHETSTINE, J. R. 2012. Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol Cell,* 48**,** 491-507.

BLINKA, S., REIMER, M. H., JR., PULAKANTI, K. & RAO, S. 2016. Super-Enhancers at the Nanog Locus Differentially Regulate Neighboring Pluripotency-Associated Genes. *Cell Rep,* 17**,** 19-28.

BLUM, M., GAUNT, S. J., CHO, K. W. Y., STEINBEISSER, H., BLUMBERG, B., BITTNER, D. & DE ROBERTIS, E. M. 1992. Gastrulation in the mouse: The role of the homeobox gene goosecoid. *Cell,* 69**,** 1097-1106.

BOGGS, B. A., CHEUNG, P., HEARD, E., SPECTOR, D. L., CHINAULT, A. C. & ALLIS, C. D. 2002. Differentially methylated forms of histone H3 show unique association patterns with inactive human X chromosomes. *Nat Genet,* 30**,** 73-6.

BOSTICK, M., KIM, J. K., ESTÈVE, P. O., CLARK, A., PRADHAN, S. & JACOBSEN, S. E. 2007. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science,* 317**,** 1760-4.

BØTTGER, P., TRACZ, Z., HEUCK, A., NISSEN, P., ROMERO-RAMOS, M. & LYKKE-HARTMANN, K. 2011. Distribution of Na/K-ATPase alpha 3 isoform, a sodium-potassium P-type pump associated with rapid-onset of dystonia parkinsonism (RDP) in the adult mouse brain. *Journal of Comparative Neurology,* 519**,** 376-404.

BOURILLOT, P.-Y. & SAVATIER, P. 2010. Krüppel-like transcription factors and control of pluripotency. *BMC Biology,* 8**,** 125.

BOYER, L. A., LEE, T. I., COLE, M. F., JOHNSTONE, S. E., LEVINE, S. S., ZUCKER, J. P., GUENTHER, M. G., KUMAR, R. M., MURRAY, H. L., JENNER, R. G., GIFFORD, D. K., MELTON, D. A., JAENISCH, R. & YOUNG, R. A. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell,* 122**,** 947-56.

BOYER, L. A., PLATH, K., ZEITLINGER, J., BRAMBRINK, T., MEDEIROS, L. A., LEE, T. I., LEVINE, S. S., WERNIG, M., TAJONAR, A., RAY, M. K., BELL, G. W., OTTE, A. P., VIDAL, M., GIFFORD, D. K., YOUNG, R. A. & JAENISCH, R. 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature,* 441**,** 349-53.

BRACKERTZ, M., GONG, Z., LEERS, J. & RENKAWITZ, R. 2006. p66α and p66β of the Mi-2/NuRD complex mediate MBD2 and histone interaction. *Nucleic Acids Research,* 34**,** 397-406.

BRADLEY, A. & ROBERTSON, E. 1986. Chapter 25 Embryo-Derived Stem Cells: A Tool for Elucidating the Developmental Genetics of the Mouse. *In:* MOSCONA, A. A. & MONROY, A. (eds.) *Current Topics in Developmental Biology.* Academic Press.

BRONS, I. G., SMITHERS, L. E., TROTTER, M. W., RUGG-GUNN, P., SUN, B., CHUVA DE SOUSA LOPES, S. M., HOWLETT, S. K., CLARKSON, A., AHRLUND-RICHTER, L., PEDERSEN, R. A. & VALLIER, L. 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature,* 448**,** 191-5.

BUECKER, C., SRINIVASAN, R., WU, Z., CALO, E., ACAMPORA, D., FAIAL, T., SIMEONE, A., TAN, M., SWIGUT, T. & WYSOCKA, J. 2014. Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell,* 14**,** 838-53.

BURGESS, R., CSERJESI, P., LIGON, K. L. & OLSON, E. N. 1995. Paraxis: A Basic Helix-Loop-Helix Protein Expressed in Paraxial Mesoderm and Developing Somites. *Developmental Biology,* 168**,** 296-306.

BURTSCHER, I. & LICKERT, H. 2009. Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development,* 136**,** 1029-38.

CANDIA, A. F., HU, J., CROSBY, J., LALLEY, P. A., NODEN, D., NADEAU, J. H. & WRIGHT, C. V. 1992. Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially expressed during early mesodermal patterning in mouse embryos. *Development,* 116**,** 1123-1136.

CANO, A., PÉREZ-MORENO, M. A., RODRIGO, I., LOCASCIO, A., BLANCO, M. J., DEL BARRIO, M. G., PORTILLO, F. & NIETO, M. A. 2000. The transcription factor Snail controls epithelial–mesenchymal transitions by repressing E-cadherin expression. *Nature Cell Biology,* 2**,** 76-83.

CAO, J. & YAN, Q. 2012. Histone Ubiquitination and Deubiquitination in Transcription, DNA Damage Response, and Cancer. *Frontiers in Oncology,* 2.

CARONE, D. M. & LAWRENCE, J. B. 2013. Heterochromatin instability in cancer: From the Barr body to satellites and the nuclear periphery. *Seminars in Cancer Biology,* 23**,** 99-108.

CHAL, J. & POURQUIE, O. 2017. Making muscle: skeletal myogenesis in vivo and in vitro. *Development,* 144**,** 2104-2122.

CHAMBERS, I., COLBY, D., ROBERTSON, M., NICHOLS, J., LEE, S., TWEEDIE, S. & SMITH, A. 2003. Functional Expression Cloning of Nanog, a Pluripotency Sustaining Factor in Embryonic Stem Cells. *Cell,* 113**,** 643-655.

CHAMBERS, I., SILVA, J., COLBY, D., NICHOLS, J., NIJMEIJER, B., ROBERTSON, M., VRANA, J., JONES, K., GROTEWOLD, L. & SMITH, A. 2007. Nanog safeguards pluripotency and mediates germline development. *Nature,* 450**,** 1230-1234.

CHAN, SUNNY S.-K., SHI, X., TOYAMA, A., ARPKE, ROBERT W., DANDAPAT, A., IACOVINO, M., KANG, J., LE, G., HAGEN, HANNAH R., GARRY, DANIEL J. & KYBA, M. 2013. Mesp1 Patterns Mesoderm into Cardiac, Hematopoietic, or Skeletal Myogenic Progenitors in a Context-Dependent Manner. *Cell Stem Cell,* 12**,** 587-601.

CHANDRASEKHARAN, M. B., HUANG, F. & SUN, Z.-W. 2010. Histone H2B ubiquitination and beyond. *Epigenetics,* 5**,** 460-468.

CHEN, Y., KIM, J., ZHANG, R., YANG, X., ZHANG, Y., FANG, J., CHEN, Z., TENG, L., CHEN, X., GE, H., ATADJA, P., LI, E., CHEN, T. & QI, W. 2016. Histone Demethylase LSD1 Promotes Adipocyte Differentiation through Repressing Wnt Signaling. *Cell Chem Biol,* 23**,** 1228-1240.

CHEN, Y., YANG, Y., WANG, F., WAN, K., YAMANE, K., ZHANG, Y. & LEI, M. 2006. Crystal structure of human histone lysine-specific demethylase 1 (LSD1). *Proc Natl Acad Sci U S A,* 103**,** 13956-61.

CHOI, H. J., PARK, J. H., PARK, M., WON, H. Y., JOO, H. S., LEE, C. H., LEE, J. Y. & KONG, G. 2015. UTX inhibits EMT-induced breast CSC properties by epigenetic repression of EMT genes in cooperation with LSD1 and HDAC1. *EMBO Rep,* 16**,** 1288-98.

CHOI, J., JANG, H., KIM, H., KIM, S.-T., CHO, E.-J. & YOUN, H.-D. 2010. Histone demethylase LSD1 is required to induce skeletal muscle differentiation by regulating myogenic factors. *Biochemical and Biophysical Research Communications,* 401**,** 327-332.

CHOWDHURY, A. H., RAMROOP, J. R., UPADHYAY, G., SENGUPTA, A., ANDRZEJCZYK, A. & SALEQUE, S. 2013. Differential transcriptional regulation of meis1 by Gfi1b and its co-factors LSD1 and CoREST. *PLoS One,* 8**,** e53666.

CIRUNA, B. & ROSSANT, J. 2001. FGF Signaling Regulates Mesoderm Cell Fate Specification and Morphogenetic Movement at the Primitive Streak. *Developmental Cell,* 1**,** 37-49.

CLAPIER, C. R., IWASA, J., CAIRNS, B. R. & PETERSON, C. L. 2017. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nature Reviews Molecular Cell Biology,* 18**,** 407-422.

CONTRERAS-LEAL, E., HERNÁNDEZ-OLIVERAS, A., FLORES-PEREDO, L., ZARAIN-HERZBERG, Á. & SANTIAGO-GARCÍA, J. 2016. Histone deacetylase inhibitors promote the expression of ATP2A3 gene in breast cancer cell lines. *Molecular Carcinogenesis,* 55**,** 1477-1485.

COSTELLO, I., PIMEISL, I.-M., DRÄGER, S., BIKOFF, E. K., ROBERTSON, E. J. & ARNOLD, S. J. 2011. The T-box transcription factor Eomesodermin acts upstream of Mesp1 to specify cardiac mesoderm during mouse gastrulation. *Nature Cell Biology,* 13**,** 1084-1091.

COWLEY, S. M., IRITANI, B. M., MENDRYSA, S. M., XU, T., CHENG, P. F., YADA, J., LIGGITT, H. D. & EISENMAN, R. N. 2005. The mSin3A chromatin-modifying complex is essential for embryogenesis and T-cell development. *Mol Cell Biol,* 25**,** 6990-7004.

CREYGHTON, M. P., CHENG, A. W., WELSTEAD, G. G., KOOISTRA, T., CAREY, B. W., STEINE, E. J., HANNA, J., LODATO, M. A., FRAMPTON, G. M., SHARP, P. A., BOYER, L. A., YOUNG, R. A. & JAENISCH, R. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A,* 107**,** 21931-6.

DAVID, R., BRENNER, C., STIEBER, J., SCHWARZ, F., BRUNNER, S., VOLLMER, M., MENTELE, E., MÜLLER-HÖCKER, J., KITAJIMA, S., LICKERT, H., RUPP, R. & FRANZ, W. M. 2008. MesP1 drives vertebrate cardiovascular differentiation through Dkk-1-mediated blockade of Wnt-signalling. *Nature Cell Biology,* 10**,** 338-345.

DAVIDSON, K. C., MASON, E. A. & PERA, M. F. 2015. The pluripotent state in mouse and human. *Development,* 142**,** 3090-3099.

DAVIES, O. R., LIN, C. Y., RADZISHEUSKAYA, A., ZHOU, X., TAUBE, J., BLIN, G., WATERHOUSE, A., SMITH, A. J. & LOWELL, S. 2013. Tcf15 primes pluripotent cells for differentiation. *Cell Rep,* 3**,** 472-84.

DELCUVE, G. P., KHAN, D. H. & DAVIE, J. R. 2012. Roles of histone deacetylases in epigenetic regulation: emerging paradigms from studies with inhibitors. *Clinical Epigenetics,* 4**,** 5.

DESCHAMPS, J. & VAN NES, J. 2005. Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development,* 132**,** 2931-2942.

DHALL, A., WELLER, C. E., CHU, A., SHELTON, P. M. M. & CHATTERJEE, C. 2017. Chemically Sumoylated Histone H4 Stimulates Intranucleosomal Demethylation by the LSD1-CoREST Complex. *ACS Chem Biol,* 12**,** 2275-2280.

DI GIOVANNANTONIO, L. G., ACAMPORA, D., OMODEI, D., NIGRO, V., BARBA, P., BARBIERI, E., CHAMBERS, I. & SIMEONE, A. 2021. Direct repression of Nanog and Oct4 by OTX2 modulates the contribution of epiblast-derived cells to germline and somatic lineage. *Development,* 148.

DILLON, N. & FESTENSTEIN, R. 2002. Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. *Trends Genet,* 18**,** 252-8.

DILLON, S. C., ZHANG, X., TRIEVEL, R. C. & CHENG, X. 2005. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biology,* 6**,** 227.

DING, J., ZHANG, Z. M., XIA, Y., LIAO, G. Q., PAN, Y., LIU, S., ZHANG, Y. & YAN, Z. S. 2013. LSD1-mediated epigenetic modification contributes to proliferation and metastasis of colon cancer. *Br J Cancer,* 109**,** 994-1003.

DING, S., WU, X., LI, G., HAN, M., ZHUANG, Y. & XU, T. 2005. Efficient Transposition of the piggyBac (PB) Transposon in Mammalian Cells and Mice. *Cell,* 122**,** 473-483.

DOETSCHMAN, T. C., EISTETTER, H., KATZ, M., SCHMIDT, W. & KEMLER, R. 1985. The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morphol,* 87**,** 27-45.

DOU, Y., MILNE, T. A., RUTHENBURG, A. J., LEE, S., LEE, J. W., VERDINE, G. L., ALLIS, C. D. & ROEDER, R. G. 2006. Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol,* 13**,** 713-9.

DUBRULLE, J., MCGREW, M. J. & POURQUIÉ, O. 2001. FGF Signaling Controls Somite Boundary Position and Regulates Segmentation Clock Control of Spatiotemporal Hox Gene Activation. *Cell,* 106**,** 219-232.

DYER, N. P., SHAHREZAEI, V. & HEBENSTREIT, D. 2019. LiBiNorm: an htseq-count analogue with improved normalisation of Smart-seq2 data and library preparation diagnostics. *PeerJ,* 7**,** e6222.

EICKBUSH, T. H. & MOUDRIANAKIS, E. N. 1978. The compaction of DNA helices into either continuous supercoils or folded-fiber rods and toroids. *Cell,* 13**,** 295-306.

EVANS, M. J. & KAUFMAN, M. H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature,* 292**,** 154-6.

EWELS, P., MAGNUSSON, M., LUNDIN, S. & KÄLLER, M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics,* 32**,** 3047-8.

FALCO, G., LEE, S.-L., STANGHELLINI, I., BASSEY, U. C., HAMATANI, T. & KO, M. S. H. 2007. Zscan4: A novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells. *Developmental Biology,* 307**,** 539-550.

FANG, R., BARBERA, A. J., XU, Y., RUTENBERG, M., LEONOR, T., BI, Q., LAN, F., MEI, P., YUAN, G. C., LIAN, C., PENG, J., CHENG, D., SUI, G., KAISER, U. B., SHI, Y. & SHI, Y. G. 2010. Human LSD2/KDM1b/AOF1 regulates gene transcription by modulating intragenic H3K4me2 methylation. *Mol Cell,* 39**,** 222-33.

FENG, Q., WANG, H., NG, H. H., ERDJUMENT-BROMAGE, H., TEMPST, P., STRUHL, K. & ZHANG, Y. 2002. Methylation of H3-Lysine 79 Is Mediated by a New Family of HMTases without a SET Domain. *Current Biology,* 12**,** 1052-1058.

FERRARI-AMOROTTI, G., FRAGLIASSO, V., ESTEKI, R., PRUDENTE, Z., SOLIERA, A. R., CATTELANI, S., MANZOTTI, G., GRISENDI, G., DOMINICI, M., PIERACCIOLI, M., RASCHELLÀ, G., CHIODONI, C., COLOMBO, M. P. & CALABRETTA, B. 2013. Inhibiting Interactions of Lysine Demethylase LSD1 with Snail/Slug Blocks Cancer Cell Invasion. *Cancer Research,* 73**,** 235-245.

FESTUCCIA, N., OSORNO, R., HALBRITTER, F., KARWACKI-NEISIUS, V., NAVARRO, P., COLBY, D., WONG, F., YATES, A., TOMLINSON, S. R. & CHAMBERS, I. 2012. Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell Stem Cell,* 11**,** 477-90.

FIRULLI, A. B. 2003. A HANDful of questions: the molecular biology of the heart and neural crest derivatives (HAND)-subclass of basic helix–loop–helix transcription factors. *Gene,* 312**,** 27-40.

FISCHLE, W., TSENG, B. S., DORMANN, H. L., UEBERHEIDE, B. M., GARCIA, B. A., SHABANOWITZ, J., HUNT, D. F., FUNABIKI, H. & ALLIS, C. D. 2005. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature,* 438**,** 1116-22.

FORNERIS, F., BINDA, C., ADAMO, A., BATTAGLIOLI, E. & MATTEVI, A. 2007. Structural basis of LSD1-CoREST selectivity in histone H3 recognition. *J Biol Chem,* 282**,** 20070-4.

FORNERIS, F., BINDA, C., VANONI, M. A., BATTAGLIOLI, E. & MATTEVI, A. 2005a. Human histone demethylase LSD1 reads the histone code. *J Biol Chem,* 280**,** 41360-5.

FORNERIS, F., BINDA, C., VANONI, M. A., MATTEVI, A. & BATTAGLIOLI, E. 2005b. Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process. *FEBS Lett,* 579**,** 2203-7.

FOSTER, C. T., DOVEY, O. M., LEZINA, L., LUO, J. L., GANT, T. W., BARLEV, N., BRADLEY, A. & COWLEY, S. M. 2010. Lysine-specific demethylase 1 regulates the embryonic transcriptome and CoREST stability. *Mol Cell Biol,* 30**,** 4851-63.

FRANCI, G., CIOTTA, A. & ALTUCCI, L. 2014. The Jumonji family: past, present and future of histone demethylases in cancer. *Biomolecular Concepts,* 5**,** 209-224.

GE, W., LIU, Y., CHEN, T., ZHANG, X., LV, L., JIN, C., JIANG, Y., SHI, L. & ZHOU, Y. 2014. The epigenetic promotion of osteogenic differentiation of human adipose-derived stem cells by the genetic and chemical blockade of histone demethylase LSD1. *Biomaterials,* 35**,** 6015-25.

GIANNOTTA, M., TRANI, M. & DEJANA, E. 2013. VE-Cadherin and Endothelial Adherens Junctions: Active Guardians of Vascular Integrity. *Developmental Cell,* 26**,** 441-454.

GOULDING, M., LUMSDEN, A. & PAQUETTE, A. J. 1994. Regulation of Pax-3 expression in the dermomyotome and its role in muscle development. *Development,* 120**,** 957-971.

GRANT, P. A. & BERGER, S. L. 1999. Histone acetyltransferase complexes. *Seminars in Cell & Developmental Biology,* 10**,** 169-177.

GREENBERG, M. V. C. & BOURC'HIS, D. 2019. The diverse roles of DNA methylation in mammalian development and disease. *Nature Reviews Molecular Cell Biology,* 20**,** 590-607.

GREENFELD, H., LIN, J. & MULLINS, M. C. 2021. The BMP signaling gradient is interpreted through concentration thresholds in dorsal–ventral axial patterning. *PLOS Biology,* 19**,** e3001059.

GROTEWOLD, L., PLUM, M., DILDROP, R., PETERS, T. & RÜTHER, U. 2001. Bambi is coexpressed with Bmp-4 during mouse embryogenesis. *Mech Dev,* 100**,** 327-30.

GRZENDA, A., LOMBERK, G., ZHANG, J. S. & URRUTIA, R. 2009. Sin3: master scaffold and transcriptional corepressor. *Biochim Biophys Acta,* 1789**,** 443-50.

GUO, X., WANG, L., LI, J., DING, Z., XIAO, J., YIN, X., HE, S., SHI, P., DONG, L., LI, G., TIAN, C., WANG, J., CONG, Y. & XU, Y. 2015. Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature,* 517**,** 640-644.

HAGEY, D. W., KLUM, S., KURTSDOTTER, I., ZAOUTER, C., TOPCIC, D., ANDERSSON, O., BERGSLAND, M. & MUHR, J. 2018. SOX2 regulates common and specific stem cell features in the CNS and endoderm derived organs. *PLoS Genet,* 14**,** e1007224.

HAKIMI, M. A., BOCHAR, D. A., CHENOWETH, J., LANE, W. S., MANDEL, G. & SHIEKHATTAR, R. 2002. A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proc Natl Acad Sci U S A,* 99**,** 7420-5.

HASSANI, S.-N., MORADI, S., TALEAHMAD, S., BRAUN, T. & BAHARVAND, H. 2019. Transition of inner cell mass to embryonic stem cells: mechanisms, facts, and hypotheses. *Cellular and Molecular Life Sciences,* 76**,** 873-892.

HAYAKAWA, T. & NAKAYAMA, J. 2011. Physiological roles of class I HDAC complex and histone demethylase. *J Biomed Biotechnol,* 2011**,** 129383.

HAYAMI, S., KELLY, J. D., CHO, H.-S., YOSHIMATSU, M., UNOKI, M., TSUNODA, T., FIELD, H. I., NEAL, D. E., YAMAUE, H., PONDER, B. A. J., NAKAMURA, Y. & HAMAMOTO, R. 2011. Overexpression of LSD1 contributes to human carcinogenesis through chromatin regulation in various cancers. *International Journal of Cancer,* 128**,** 574-586.

HAYASHI, K., DE SOUSA LOPES, S. M. C., TANG, F., LAO, K. & SURANI, M. A. 2008. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell,* 3**,** 391-401.

HAYASHI, K., OHTA, H., KURIMOTO, K., ARAMAKI, S. & SAITOU, M. 2011. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell,* 146**,** 519-32.

HEINTZMAN, N. D., STUART, R. K., HON, G., FU, Y., CHING, C. W., HAWKINS, R. D., BARRERA, L. O., VAN CALCAR, S., QU, C., CHING, K. A., WANG, W., WENG, Z., GREEN, R. D., CRAWFORD, G. E. & REN, B. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet,* 39**,** 311-8.

HINO, S., KOHROGI, K. & NAKAO, M. 2016. Histone demethylase LSD1 controls the phenotypic plasticity of cancer cells. *Cancer Sci,* 107**,** 1187-92.

HIVERT, B., LIU, Z., CHUANG, C.-Y., DOHERTY, P. & SUNDARESAN, V. 2002. Robo1 and Robo2 Are Homophilic Binding Molecules That Promote Axonal Growth. *Molecular and Cellular Neuroscience,* 21**,** 534-545.

HODAWADEKAR, S. C. & MARMORSTEIN, R. 2007. Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design. *Oncogene,* 26**,** 5528-40.

HOFMANN, M., SCHUSTER-GOSSLER, K., WATABE-RUDOLPH, M., AULEHLA, A., HERRMANN, B. G. & GOSSLER, A. 2004. WNT signaling, in synergy with T/TBX6, controls Notch signaling by regulating Dll1 expression in the presomitic mesoderm of mouse embryos. *Genes & Development,* 18**,** 2712-2717.

HOLLIER, B. G., TINNIRELLO, A. A., WERDEN, S. J., EVANS, K. W., TAUBE, J. H., SARKAR, T. R., SPHYRIS, N., SHARIATI, M., KUMAR, S. V., BATTULA, V. L., HERSCHKOWITZ, J. I., GUERRA, R., CHANG, J. T., MIURA, N., ROSEN, J. M. & MANI, S. A. 2013. FOXC2 Expression Links Epithelial–Mesenchymal Transition and Stem Cell Properties in Breast Cancer. *Cancer Research,* 73**,** 1981-1992.

HONG, L., SCHROTH, G. P., MATTHEWS, H. R., YAU, P. & BRADBURY, E. M. 1993. Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 “tail” to DNA. *Journal of Biological Chemistry,* 268**,** 305-314.

HOSHINO, H., SHIOI, G. & AIZAWA, S. 2015. AVE protein expression and visceral endoderm cell behavior during anterior–posterior axis formation in mouse embryos: Asymmetry in OTX2 and DKK1 expression. *Developmental Biology,* 402**,** 175-191.

HU, E., CHEN, Z., FREDRICKSON, T., ZHU, Y., KIRKPATRICK, R., ZHANG, G. F., JOHANSON, K., SUNG, C. M., LIU, R. & WINKLER, J. 2000. Cloning and characterization of a novel human class I histone deacetylase that functions as a transcription repressor. *J Biol Chem,* 275**,** 15254-64.

HU, X., LI, X., VALVERDE, K., FU, X., NOGUCHI, C., QIU, Y. & HUANG, S. 2009. LSD1-mediated epigenetic modification is required for TAL1 function and hematopoiesis. *Proc Natl Acad Sci U S A,* 106**,** 10141-6.

HUANG, H., SABARI, B. R., GARCIA, B. A., ALLIS, C. D. & ZHAO, Y. 2014. SnapShot: histone modifications. *Cell,* 159**,** 458-458 e1.

HUANG, J., SENGUPTA, R., ESPEJO, A. B., LEE, M. G., DORSEY, J. A., RICHTER, M., OPRAVIL, S., SHIEKHATTAR, R., BEDFORD, M. T., JENUWEIN, T. & BERGER, S. L. 2007. p53 is regulated by the lysine demethylase LSD1. *Nature,* 449**,** 105-8.

HYUN, K., JEON, J., PARK, K. & KIM, J. 2017. Writing, erasing and reading histone lysine methylations. *Exp Mol Med,* 49**,** e324.

ITOH, T., FAIRALL, L., MUSKETT, F. W., MILANO, C. P., WATSON, P. J., ARNAUDO, N., SALEH, A., MILLARD, C. J., EL-MEZGUELDI, M., MARTINO, F. & SCHWABE, J. W. 2015. Structural and functional characterization of a cell cycle associated HDAC1/2 complex reveals the structural basis for complex assembly and nucleosome targeting. *Nucleic Acids Res,* 43**,** 2033-44.

IVANOVA, N., DOBRIN, R., LU, R., KOTENKO, I., LEVORSE, J., DECOSTE, C., SCHAFER, X., LUN, Y. & LEMISCHKA, I. R. 2006. Dissecting self-renewal in stem cells with RNA interference. *Nature,* 442**,** 533-538.

JACOBSON, R. H., LADURNER, A. G., KING, D. S. & TJIAN, R. 2000. Structure and Function of a Human TAFII250 Double Bromodomain Module. *Science,* 288**,** 1422-1425.

KALLUNKI, P., EDELMAN, G. M. & JONES, F. S. 1997. Tissue-specific Expression of the L1 Cell Adhesion Molecule Is Modulated by the Neural Restrictive Silencer Element. *Journal of Cell Biology,* 138**,** 1343-1354.

KALLURI, R. & WEINBERG, R. A. 2009. The basics of epithelial-mesenchymal transition. *The Journal of Clinical Investigation,* 119**,** 1420-1428.

KANAI-AZUMA, M., KANAI, Y., GAD, J. M., TAJIMA, Y., TAYA, C., KUROHMARU, M., SANAI, Y., YONEKAWA, H., YAZAKI, K., TAM, P. P. L. & HAYASHI, Y. 2002. Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development,* 129**,** 2367-2379.

KARYTINOS, A., FORNERIS, F., PROFUMO, A., CIOSSANI, G., BATTAGLIOLI, E., BINDA, C. & MATTEVI, A. 2009. A novel mammalian flavin-dependent histone demethylase. *J Biol Chem,* 284**,** 17775-82.

KEATING, S. & EL-OSTA, A. 2013. Transcriptional regulation by the Set7 lysine methyltransferase. *Epigenetics,* 8**,** 361-372.

KELLY, R. D. & COWLEY, S. M. 2013. The physiological roles of histone deacetylase (HDAC) 1 and 2: complex co-stars with multiple leading parts. *Biochem Soc Trans,* 41**,** 741-9.

KERENYI, M. A., SHAO, Z., HSU, Y.-J., GUO, G., LUC, S., O'BRIEN, K., FUJIWARA, Y., PENG, C., NGUYEN, M. & ORKIN, S. H. 2013. Histone demethylase Lsd1 represses hematopoietic stem and progenitor cell signatures during blood cell maturation. *eLife,* 2**,** e00633.

KIM, D., PAGGI, J. M., PARK, C., BENNETT, C. & SALZBERG, S. L. 2019a. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol,* 37**,** 907-915.

KIM, H.-S., NEUGEBAUER, J., MCKNITE, A., TILAK, A. & CHRISTIAN, J. L. 2019b. BMP7 functions predominantly as a heterodimer with BMP2 or BMP4 during mammalian embryogenesis. *eLife,* 8**,** e48872.

KIM, S. A., ZHU, J., YENNAWAR, N., EEK, P. & TAN, S. 2020. Crystal Structure of the LSD1/CoREST Histone Demethylase Bound to Its Nucleosome Substrate. *Mol Cell,* 78**,** 903-914.e4.

KLINOWSKA, T. C. M., IRELAND, G. W. & KIMBER, S. J. 1994. A new in vitro model of murine mesoderm migration: the role of fibronectin and laminin. *Differentiation,* 57**,** 7-19.

KLOSE, R. J., YAMANE, K., BAE, Y., ZHANG, D., ERDJUMENT-BROMAGE, H., TEMPST, P., WONG, J. & ZHANG, Y. 2006. The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature,* 442**,** 312-6.

KORNBERG, R. D. 1997. Chromatin structure- a repeating unit of histones and DNA *Science,* 184**,** 868-871.

KOUZARIDES, T. 2007. Chromatin modifications and their function. *Cell,* 128**,** 693-705.

KUME, T., JIANG, H., TOPCZEWSKA, J. M. & HOGAN, B. L. 2001. The murine winged helix transcription factors, Foxc1 and Foxc2, are both required for cardiovascular development and somitogenesis. *Genes Dev,* 15**,** 2470-82.

KUNATH, T., SABA-EL-LEIL, M. K., ALMOUSAILLEAKH, M., WRAY, J., MELOCHE, S. & SMITH, A. 2007. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development,* 134**,** 2895-902.

LAHERTY, C. D., YANG, W.-M., SUN, J.-M., DAVIE, J. R., SETO, E. & EISENMAN, R. N. 1997. Histone Deacetylases Associated with the mSin3 Corepressor Mediate Mad Transcriptional Repression. *Cell,* 89**,** 349-356.

LAMBROT, R., LAFLEUR, C. & KIMMINS, S. 2015. The histone demethylase KDM1A is essential for the maintenance and differentiation of spermatogonial stem cells and progenitors. *FASEB J,* 29**,** 4402-16.

LAN, F., COLLINS, R. E., DE CEGLI, R., ALPATOV, R., HORTON, J. R., SHI, X., GOZANI, O., CHENG, X. & SHI, Y. 2007. Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature,* 448**,** 718-22.

LAWSON, K. A., DUNN, N. R., ROELEN, B. A. J., ZEINSTRA, L. M., DAVIS, A. M., WRIGHT, C. V. E., KORVING, J. P. W. F. M. & HOGAN, B. L. M. 1999. Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes & Development,* 13**,** 424-436.

LE GUEZENNEC, X., VERMEULEN, M. & STUNNENBERG, H. G. 2006. Molecular characterization of Sin3 PAH-domain interactor specificity and identification of PAH partners. *Nucleic Acids Research,* 34**,** 3929-3937.

LEE, M. G., WYNDER, C., BOCHAR, D. A., HAKIMI, M. A., COOCH, N. & SHIEKHATTAR, R. 2006. Functional interplay between histone demethylase and deacetylase enzymes. *Mol Cell Biol,* 26**,** 6395-402.

LEE, M. G., WYNDER, C., COOCH, N. & SHIEKHATTAR, R. 2005. An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature,* 437**,** 432-5.

LI, H., HANDSAKER, B., WYSOKER, A., FENNELL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G. & DURBIN, R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* 25**,** 2078-9.

LI, H., ILIN, S., WANG, W., DUNCAN, E. M., WYSOCKA, J., ALLIS, C. D. & PATEL, D. J. 2006. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature,* 442**,** 91-95.

LI, L., LIU, X., HE, L., YANG, J., PEI, F., LI, W., LIU, S., CHEN, Z., XIE, G., XU, B., TING, X., ZHANG, Z., JIN, T., LIU, X., ZHANG, W., YUAN, S., YANG, Z., WU, C., ZHANG, Y., YANG, X., YI, X., LIANG, J., SHANG, Y. & SUN, L. 2017. ZNF516 suppresses EGFR by targeting the CtBP/LSD1/CoREST complex to chromatin. *Nat Commun,* 8**,** 691.

LIN, T., PONN, A., HU, X., LAW, B. K. & LU, J. 2010a. Requirement of the histone demethylase LSD1 in Snai1-mediated transcriptional repression during epithelial-mesenchymal transition. *Oncogene,* 29**,** 4896-904.

LIN, Y., WU, Y., LI, J., DONG, C., YE, X., CHI, Y. I., EVERS, B. M. & ZHOU, B. P. 2010b. The SNAG domain of Snail1 functions as a molecular hook for recruiting lysine-specific demethylase 1. *EMBO J,* 29**,** 1803-16.

LINDSLEY, R. C., GILL, J. G., KYBA, M., MURPHY, T. L. & MURPHY, K. M. 2006. Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. *Development,* 133**,** 3787-96.

LIU, L., SOUTO, J., LIAO, W., JIANG, Y., LI, Y., NISHINAKAMURA, R., HUANG, S., ROSENGART, T., YANG, V. W., SCHUSTER, M., MA, Y. & YANG, J. 2013. Histone lysine-specific demethylase 1 (LSD1) protein is involved in Sal-like protein 4 (SALL4)-mediated transcriptional repression in hematopoietic stem cells. *J Biol Chem,* 288**,** 34719-28.

LOEBEL, D. A., WATSON, C. M., DE YOUNG, R. A. & TAM, P. P. 2003. Lineage choice and differentiation in mouse embryos and embryonic stem cells. *Dev Biol,* 264**,** 1-14.

LOH, C.-Y., CHAI, J. Y., TANG, T. F., WONG, W. F., SETHI, G., SHANMUGAM, M. K., CHONG, P. P. & LOOI, C. Y. 2019. The E-Cadherin and N-Cadherin Switch in Epithelial-to-Mesenchymal Transition: Signaling, Therapeutic Implications, and Challenges. *Cells,* 8**,** 1118.

LOH, K. M. & LIM, B. 2011. A precarious balance: pluripotency factors as lineage specifiers. *Cell Stem Cell,* 8**,** 363-9.

LOH, Y. H., YANG, L., YANG, J. C., LI, H., COLLINS, J. J. & DALEY, G. Q. 2011. Genomic approaches to deconstruct pluripotency. *Annu Rev Genomics Hum Genet,* 12**,** 165-85.

LOPEZ, C. I., SAUD, K. E., AGUILAR, R., BERNDT, F. A., CÁNOVAS, J., MONTECINO, M. & KUKULJAN, M. 2016. The chromatin modifying complex CoREST/LSD1 negatively regulates notch pathway during cerebral cortex development. *Dev Neurobiol,* 76**,** 1360-1373.

LOPEZ-CAMPISTROUS, A., ADEWUYI, E. E., WILLIAMS, D. C. & MCMULLEN, T. P. W. 2021. Gene expression profile of epithelial-mesenchymal transition mediators in papillary thyroid cancer. *Endocrine,* 72**,** 452-461.

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

LUGER, K., MÄDER, A. W., RICHMOND, R. K., SARGENT, D. F. & RICHMOND, T. J. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature,* 389**,** 251-260.

LV, T., YUAN, D., MIAO, X., LV, Y., ZHAN, P., SHEN, X. & SONG, Y. 2012. Over-Expression of LSD1 Promotes Proliferation, Migration and Invasion in Non-Small Cell Lung Cancer. *PLOS ONE,* 7**,** e35065.

LYST, M. J., EKIERT, R., EBERT, D. H., MERUSI, C., NOWAK, J., SELFRIDGE, J., GUY, J., KASTAN, N. R., ROBINSON, N. D., DE LIMA ALVES, F., RAPPSILBER, J., GREENBERG, M. E. & BIRD, A. 2013. Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. *Nat Neurosci,* 16**,** 898-902.

MAIQUES-DIAZ, A., SPENCER, G. J., LYNCH, J. T., CICERI, F., WILLIAMS, E. L., AMARAL, F. M. R., WISEMAN, D. H., HARRIS, W. J., LI, Y., SAHOO, S., HITCHIN, J. R., MOULD, D. P., FAIRWEATHER, E. E., WASZKOWYCZ, B., JORDAN, A. M., SMITH, D. L. & SOMERVAILLE, T. C. P. 2018. Enhancer Activation by Pharmacologic Displacement of LSD1 from GFI1 Induces Differentiation in Acute Myeloid Leukemia. *Cell Rep,* 22**,** 3641-3659.

MARTELLO, G., BERTONE, P. & SMITH, A. 2013. Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. *Embo j,* 32**,** 2561-74.

MARTIN, G. R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences,* 78**,** 7634-7638.

MASUI, S., NAKATAKE, Y., TOYOOKA, Y., SHIMOSATO, D., YAGI, R., TAKAHASHI, K., OKOCHI, H., OKUDA, A., MATOBA, R., SHAROV, A. A., KO, M. S. & NIWA, H. 2007. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol,* 9**,** 625-35.

MCDONALD, O. G., WU, H., TIMP, W., DOI, A. & FEINBERG, A. P. 2011. Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nat Struct Mol Biol,* 18**,** 867-74.

MENDEZ, M. G., KOJIMA, S. & GOLDMAN, R. D. 2010. Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition. *Faseb j,* 24**,** 1838-51.

METZGER, E., WISSMANN, M., YIN, N., MULLER, J. M., SCHNEIDER, R., PETERS, A. H., GUNTHER, T., BUETTNER, R. & SCHULE, R. 2005. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature,* 437**,** 436-9.

MILLARD, C. J., VARMA, N., SALEH, A., MORRIS, K., WATSON, P. J., BOTTRILL, A. R., FAIRALL, L., SMITH, C. J. & SCHWABE, J. W. 2016. The structure of the core NuRD repression complex provides insights into its interaction with chromatin. *Elife,* 5**,** e13941.

MILLARD, C. J., WATSON, P. J., CELARDO, I., GORDIYENKO, Y., COWLEY, S. M., ROBINSON, C. V., FAIRALL, L. & SCHWABE, J. W. 2013. Class I HDACs share a common mechanism of regulation by inositol phosphates. *Mol Cell,* 51**,** 57-67.

MILLARD, C. J., WATSON, P. J., FAIRALL, L. & SCHWABE, J. W. R. 2017. Targeting Class I Histone Deacetylases in a "Complex" Environment. *Trends Pharmacol Sci,* 38**,** 363-377.

MISHINA, Y. 2003. Function of bone morphogenetic protein signaling during mouse development. *Frontiers in Bioscience-Landmark,* 8**,** 855-869.

MITSUI, K., TOKUZAWA, Y., ITOH, H., SEGAWA, K., MURAKAMI, M., TAKAHASHI, K., MARUYAMA, M., MAEDA, M. & YAMANAKA, S. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell,* 113**,** 631-42.

MORARU, M. & SCHALCH, T. 2019. Chromatin fiber structural motifs as regulatory hubs of genome function? *Essays Biochem,* 63**,** 123-132.

MOREY, L., SANTANACH, A. & CROCE, L. D. 2015. Pluripotency and Epigenetic Factors in Mouse Embryonic Stem Cell Fate Regulation. *Molecular and Cellular Biology,* 35**,** 2716-2728.

MUKHOPADHYAY, M., SHTROM, S., RODRIGUEZ-ESTEBAN, C., CHEN, L., TSUKUI, T., GOMER, L., DORWARD, D. W., GLINKA, A., GRINBERG, A., HUANG, S.-P., NIEHRS, C., BELMONTE, J. C. I. & WESTPHAL, H. 2001. Dickkopf1 Is Required for Embryonic Head Induction and Limb Morphogenesis in the Mouse. *Developmental Cell,* 1**,** 423-434.

MURZINA, N. V., PEI, X.-Y., ZHANG, W., SPARKES, M., VICENTE-GARCIA, J., PRATAP, J. V., MCLAUGHLIN, S. H., BEN-SHAHAR, T. R., VERREAULT, A., LUISI, B. F. & LAUE, E. D. 2008. Structural Basis for the Recognition of Histone H4 by the Histone-Chaperone RbAp46. *Structure,* 16**,** 1077-1085.

MUSRI, M. M., CARMONA, M. C., HANZU, F. A., KALIMAN, P., GOMIS, R. & PARRIZAS, M. 2010. Histone demethylase LSD1 regulates adipogenesis. *J Biol Chem,* 285**,** 30034-41.

MUSSELMAN, C. A., KHORASANIZADEH, S. & KUTATELADZE, T. G. 2014. Towards understanding methyllysine readout. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms,* 1839**,** 686-693.

MUSSELMAN, C. A., LALONDE, M. E., CÔTÉ, J. & KUTATELADZE, T. G. 2012. Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol,* 19**,** 1218-27.

MYRICK, D. A., CHRISTOPHER, M. A., SCOTT, A. M., SIMON, A. K., DONLIN-ASP, P. G., KELLY, W. G. & KATZ, D. J. 2017. KDM1A/LSD1 regulates the differentiation and maintenance of spermatogonia in mice. *PLOS ONE,* 12**,** e0177473.

NAKATAKE, Y., FUKUI, N., IWAMATSU, Y., MASUI, S., TAKAHASHI, K., YAGI, R., YAGI, K., MIYAZAKI, J.-I., MATOBA, R., KO, M. S. H. & NIWA, H. 2006. Klf4 Cooperates with Oct3/4 and Sox2 To Activate the Lefty1 Core Promoter in Embryonic Stem Cells. *Molecular and Cellular Biology,* 26**,** 7772-7782.

NISHIMURA, R., HATA, K., IKEDA, F., MATSUBARA, T., YAMASHITA, K., ICHIDA, F. & YONEDA, T. 2003. The role of Smads in BMP signaling. *Front Biosci,* 8**,** s275-84.

NIWA, H., BURDON, T., CHAMBERS, I. & SMITH, A. 1998. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev,* 12**,** 2048-60.

NIWA, H., MIYAZAKI, J. & SMITH, A. G. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet,* 24**,** 372-6.

NOWAK, S. J. & CORCES, V. G. 2004. Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet,* 20**,** 214-20.

NOWOTSCHIN, S., FERRER-VAQUER, A., CONCEPCION, D., PAPAIOANNOU, V. E. & HADJANTONAKIS, A.-K. 2012a. Interaction of Wnt3a, Msgn1 and Tbx6 in neural versus paraxial mesoderm lineage commitment and paraxial mesoderm differentiation in the mouse embryo. *Developmental Biology,* 367**,** 1-14.

NOWOTSCHIN, S., FERRER-VAQUER, A., CONCEPCION, D., PAPAIOANNOU, V. E. & HADJANTONAKIS, A. K. 2012b. Interaction of Wnt3a, Msgn1 and Tbx6 in neural versus paraxial mesoderm lineage commitment and paraxial mesoderm differentiation in the mouse embryo. *Dev Biol,* 367**,** 1-14.

OKANO, M., BELL, D. W., HABER, D. A. & LI, E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell,* 99**,** 247-57.

PARANJAPE, S. M., KAMAKAKA, R. T. & KADONAGA, J. T. 1994. Role of chromatin structure in the regulation of transcription by RNA polymerase II. *Annu Rev Biochem,* 63**,** 265-97.

PENA, P. V., DAVRAZOU, F., SHI, X., WALTER, K. L., VERKHUSHA, V. V., GOZANI, O., ZHAO, R. & KUTATELADZE, T. G. 2006. Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature,* 442**,** 100-3.

PEREA-GOMEZ, A., LAWSON, K. A., RHINN, M., ZAKIN, L., BRULET, P., MAZAN, S. & ANG, S. L. 2001. Otx2 is required for visceral endoderm movement and for the restriction of posterior signals in the epiblast of the mouse embryo. *Development,* 128**,** 753-765.

PETERS, A. H. F. M., O'CARROLL, D., SCHERTHAN, H., MECHTLER, K., SAUER, S., SCHÖFER, C., WEIPOLTSHAMMER, K., PAGANI, M., LACHNER, M., KOHLMAIER, A., OPRAVIL, S., DOYLE, M., SIBILIA, M. & JENUWEIN, T. 2001. Loss of the Suv39h Histone Methyltransferases Impairs Mammalian Heterochromatin and Genome Stability. *Cell,* 107**,** 323-337.

PIJUAN-SALA, B., GRIFFITHS, J. A., GUIBENTIF, C., HISCOCK, T. W., JAWAID, W., CALERO-NIETO, F. J., MULAS, C., IBARRA-SORIA, X., TYSER, R. C. V., HO, D. L. L., REIK, W., SRINIVAS, S., SIMONS, B. D., NICHOLS, J., MARIONI, J. C. & GÖTTGENS, B. 2019. A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature,* 566**,** 490-495.

RADICE, G. L., RAYBURN, H., MATSUNAMI, H., KNUDSEN, K. A., TAKEICHI, M. & HYNES, R. O. 1997. Developmental Defects in Mouse Embryos Lacking N-Cadherin. *Developmental Biology,* 181**,** 64-78.

RAMÓN-MAIQUES, S., KUO, A. J., CARNEY, D., MATTHEWS, A. G. W., OETTINGER, M. A., GOZANI, O. & YANG, W. 2007. The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. *Proceedings of the National Academy of Sciences,* 104**,** 18993-18998.

REA, S., EISENHABER, F., O'CARROLL, D., STRAHL, B. D., SUN, Z.-W., SCHMID, M., OPRAVIL, S., MECHTLER, K., PONTING, C. P., ALLIS, C. D. & JENUWEIN, T. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature,* 406**,** 593-599.

ROBERTSON, E., BRADLEY, A., KUEHN, M. & EVANS, M. 1986. Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. *Nature,* 323**,** 445-448.

ROGERS, M. B., HOSLER, B. A. & GUDAS, L. J. 1991. Specific expression of a retinoic acid-regulated, zinc-finger gene, Rex-1, in preimplantation embryos, trophoblast and spermatocytes. *Development,* 113**,** 815-824.

ROJAS, A., SCHACHTERLE, W., XU, S. M., MARTÍN, F. & BLACK, B. L. 2010. Direct transcriptional regulation of Gata4 during early endoderm specification is controlled by FoxA2 binding to an intronic enhancer. *Dev Biol,* 346**,** 346-55.

ROSSETTO, D., AVVAKUMOV, N. & CÔTÉ, J. 2012. Histone phosphorylation. *Epigenetics,* 7**,** 1098-1108.

SALEQUE, S., KIM, J., ROOKE, H. M. & ORKIN, S. H. 2007. Epigenetic regulation of hematopoietic differentiation by Gfi-1 and Gfi-1b is mediated by the cofactors CoREST and LSD1. *Mol Cell,* 27**,** 562-72.

SAMATOV, T. R., WICKLEIN, D. & TONEVITSKY, A. G. 2016. L1CAM: Cell adhesion and more. *Progress in Histochemistry and Cytochemistry,* 51**,** 25-32.

SARKAR, A. & HOCHEDLINGER, K. 2013. The Sox Family of Transcription Factors: Versatile Regulators of Stem and Progenitor Cell Fate. *Cell Stem Cell,* 12**,** 15-30.

SAUNDERS, A., HUANG, X., FIDALGO, M., REIMER, M. H., FAIOLA, F., DING, J., SÁNCHEZ-PRIEGO, C., GUALLAR, D., SÁENZ, C., LI, D. & WANG, J. 2017. The SIN3A/HDAC Corepressor Complex Functionally Cooperates with NANOG to Promote Pluripotency. *Cell Reports,* 18**,** 1713-1726.

SCHEIBNER, K., SCHIRGE, S., BURTSCHER, I., BÜTTNER, M., STERR, M., YANG, D., BÖTTCHER, A., ANSARULLAH, IRMLER, M., BECKERS, J., CERNILOGAR, F. M., SCHOTTA, G., THEIS, F. J. & LICKERT, H. 2021. Epithelial cell plasticity drives endoderm formation during gastrulation. *Nature Cell Biology,* 23**,** 692-703.

SCHUBELER, D., MACALPINE, D. M., SCALZO, D., WIRBELAUER, C., KOOPERBERG, C., VAN LEEUWEN, F., GOTTSCHLING, D. E., O'NEILL, L. P., TURNER, B. M., DELROW, J., BELL, S. P. & GROUDINE, M. 2004. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev,* 18**,** 1263-71.

SEHRAWAT, A., GAO, L., WANG, Y., BANKHEAD, A., 3RD, MCWEENEY, S. K., KING, C. J., SCHWARTZMAN, J., URRUTIA, J., BISSON, W. H., COLEMAN, D. J., JOSHI, S. K., KIM, D. H., SAMPSON, D. A., WEINMANN, S., KALLAKURY, B. V. S., BERRY, D. L., HAQUE, R., VAN DEN EEDEN, S. K., SHARMA, S., BEARSS, J., BEER, T. M., THOMAS, G. V., HEISER, L. M. & ALUMKAL, J. J. 2018. LSD1 activates a lethal prostate cancer gene network independently of its demethylase function. *Proc Natl Acad Sci U S A,* 115**,** E4179-E4188.

SHARMA, R., SHAFER, M. E. R., BAREKE, E., TREMBLAY, M., MAJEWSKI, J. & BOUCHARD, M. 2017. Bmp signaling maintains a mesoderm progenitor cell state in the mouse tailbud. *Development,* 144**,** 2982-2993.

SHEN, H., XU, W., GUO, R., RONG, B., GU, L., WANG, Z., HE, C., ZHENG, L., HU, X., HU, Z., SHAO, Z. M., YANG, P., WU, F., SHI, Y. G., SHI, Y. & LAN, F. 2016. Suppression of Enhancer Overactivation by a RACK7-Histone Demethylase Complex. *Cell,* 165**,** 331-42.

SHI, X., HONG, T., WALTER, K. L., EWALT, M., MICHISHITA, E., HUNG, T., CARNEY, D., PEÑA, P., LAN, F., KAADIGE, M. R., LACOSTE, N., CAYROU, C., DAVRAZOU, F., SAHA, A., CAIRNS, B. R., AYER, D. E., KUTATELADZE, T. G., SHI, Y., CÔTÉ, J., CHUA, K. F. & GOZANI, O. 2006. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature,* 442**,** 96-99.

SHI, Y., LAN, F., MATSON, C., MULLIGAN, P., WHETSTINE, J. R., COLE, P. A., CASERO, R. A. & SHI, Y. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell,* 119**,** 941-53.

SHI, Y., SAWADA, J.-I., SUI, G., AFFAR, E. B., WHETSTINE, J. R., LAN, F., OGAWA, H., PO-SHAN LUKE, M., NAKATANI, Y. & SHI, Y. 2003. Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature,* 422**,** 735-738.

SHI, Y. J., MATSON, C., LAN, F., IWASE, S., BABA, T. & SHI, Y. 2005. Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell,* 19**,** 857-64.

SHIIO, Y. & EISENMAN, R. N. 2003. Histone sumoylation is associated with transcriptional repression. *Proc Natl Acad Sci U S A,* 100**,** 13225-30.

SHOGREN-KNAAK, M., ISHII, H., SUN, J. M., PAZIN, M. J., DAVIE, J. R. & PETERSON, C. L. 2006. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science,* 311**,** 844-7.

SMITH, A. G., HEATH, J. K., DONALDSON, D. D., WONG, G. G., MOREAU, J., STAHL, M. & ROGERS, D. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature,* 336**,** 688-690.

SMITH, S. G. & ZHOU, M.-M. 2016. The Bromodomain: A New Target in Emerging Epigenetic Medicine. *ACS Chemical Biology,* 11**,** 598-608.

SONG, Y., DAGIL, L., FAIRALL, L., ROBERTSON, N., WU, M., RAGAN, T. J., SAVVA, C. G., SALEH, A., MORONE, N., KUNZE, M. B. A., JAMIESON, A. G., COLE, P. A., HANSEN, D. F. & SCHWABE, J. W. R. 2020. Mechanism of Crosstalk between the LSD1 Demethylase and HDAC1 Deacetylase in the CoREST Complex. *Cell Rep,* 30**,** 2699-2711 e8.

STAVROPOULOS, P., BLOBEL, G. & HOELZ, A. 2006. Crystal structure and mechanism of human lysine-specific demethylase-1. *Nat Struct Mol Biol,* 13**,** 626-32.

STERNER, D. E. & BERGER, S. L. 2000. Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev,* 64**,** 435-59.

SUMARA, I., QUADRONI, M., FREI, C., OLMA, M. H., SUMARA, G., RICCI, R. & PETER, M. 2007. A Cul3-Based E3 Ligase Removes Aurora B from Mitotic Chromosomes, Regulating Mitotic Progression and Completion of Cytokinesis in Human Cells. *Developmental Cell,* 12**,** 887-900.

SUN, G., ALZAYADY, K., STEWART, R., YE, P., YANG, S., LI, W. & SHI, Y. 2010. Histone demethylase LSD1 regulates neural stem cell proliferation. *Mol Cell Biol,* 30**,** 1997-2005.

SUN, X., MEYERS, E. N., LEWANDOSKI, M. & MARTIN, G. R. 1999. Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. *Genes & Development,* 13**,** 1834-1846.

SUTER, D. M., TIREFORT, D., JULIEN, S. & KRAUSE, K.-H. 2009. A Sox1 to Pax6 Switch Drives Neuroectoderm to Radial Glia Progression During Differentiation of Mouse Embryonic Stem Cells. *Stem Cells,* 27**,** 49-58.

TADDEI, A., GIAMPIETRO, C., CONTI, A., ORSENIGO, F., BREVIARIO, F., PIRAZZOLI, V., POTENTE, M., DALY, C., DIMMELER, S. & DEJANA, E. 2008. Endothelial adherens junctions control tight junctions by VE-cadherin-mediated upregulation of claudin-5. *Nat Cell Biol,* 10**,** 923-34.

TAM, P. P. & BEHRINGER, R. R. 1997. Mouse gastrulation: the formation of a mammalian body plan. *Mech Dev,* 68**,** 3-25.

TAM, P. P. & LOEBEL, D. A. 2007. Gene function in mouse embryogenesis: get set for gastrulation. *Nat Rev Genet,* 8**,** 368-81.

TAMASHIRO, D. A., ALARCON, V. B. & MARIKAWA, Y. 2012. Nkx1-2 is a transcriptional repressor and is essential for the activation of Brachyury in P19 mouse embryonal carcinoma cell. *Differentiation,* 83**,** 282-92.

TANG, F., BARBACIORU, C., WANG, Y., NORDMAN, E., LEE, C., XU, N., WANG, X., BODEAU, J., TUCH, B. B., SIDDIQUI, A., LAO, K. & SURANI, M. A. 2009. mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods,* 6**,** 377-382.

TESAR, P. J., CHENOWETH, J. G., BROOK, F. A., DAVIES, T. J., EVANS, E. P., MACK, D. L., GARDNER, R. L. & MCKAY, R. D. G. 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature,* 448**,** 196-199.

THAMBYRAJAH, R., MAZAN, M., PATEL, R., MOIGNARD, V., STEFANSKA, M., MARINOPOULOU, E., LI, Y., LANCRIN, C., CLAPES, T., MOROY, T., ROBIN, C., MILLER, C., COWLEY, S., GOTTGENS, B., KOUSKOFF, V. & LACAUD, G. 2016. GFI1 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1. *Nat Cell Biol,* 18**,** 21-32.

THOMPSON, S., CLARKE, A. R., POW, A. M., HOOPER, M. L. & MELTON, D. W. 1989. Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. *Cell,* 56**,** 313-321.

TOCHIO, N., UMEHARA, T., KOSHIBA, S., INOUE, M., YABUKI, T., AOKI, M., SEKI, E., WATANABE, S., TOMO, Y., HANADA, M., IKARI, M., SATO, M., TERADA, T., NAGASE, T., OHARA, O., SHIROUZU, M., TANAKA, A., KIGAWA, T. & YOKOYAMA, S. 2006. Solution structure of the SWIRM domain of human histone demethylase LSD1. *Structure,* 14**,** 457-68.

TOSIC, J., KIM, G.-J., PAVLOVIC, M., SCHRÖDER, C. M., MERSIOWSKY, S.-L., BARG, M., HOFHERR, A., PROBST, S., KÖTTGEN, M., HEIN, L. & ARNOLD, S. J. 2019. Eomes and Brachyury control pluripotency exit and germ-layer segregation by changing the chromatin state. *Nature Cell Biology,* 21**,** 1518-1531.

TOSIC, M., ALLEN, A., WILLMANN, D., LEPPER, C., KIM, J., DUTEIL, D. & SCHULE, R. 2018. Lsd1 regulates skeletal muscle regeneration and directs the fate of satellite cells. *Nat Commun,* 9**,** 366.

TOYOOKA, Y., SHIMOSATO, D., MURAKAMI, K., TAKAHASHI, K. & NIWA, H. 2008. Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development,* 135**,** 909-918.

TSUKADA, Y., FANG, J., ERDJUMENT-BROMAGE, H., WARREN, M. E., BORCHERS, C. H., TEMPST, P. & ZHANG, Y. 2006. Histone demethylation by a family of JmjC domain-containing proteins. *Nature,* 439**,** 811-6.

TSUNEOKA, Y. & FUNATO, H. 2020. Modified in situ Hybridization Chain Reaction Using Short Hairpin DNAs. *Frontiers in Molecular Neuroscience,* 13.

UPADHYAY, G., CHOWDHURY, A. H., VAIDYANATHAN, B., KIM, D. & SALEQUE, S. 2014. Antagonistic actions of Rcor proteins regulate LSD1 activity and cellular differentiation. *Proc Natl Acad Sci U S A,* 111**,** 8071-6.

VENKATESH, K., CHIVATAKARN, O., LEE, H., JOSHI, P. S., KANTOR, D. B., NEWMAN, B. A., MAGE, R., RADER, C. & GIGER, R. J. 2005. The Nogo-66 receptor homolog NgR2 is a sialic acid-dependent receptor selective for myelin-associated glycoprotein. *J Neurosci,* 25**,** 808-22.

WANG, H., WANG, L., ERDJUMENT-BROMAGE, H., VIDAL, M., TEMPST, P., JONES, R. S. & ZHANG, Y. 2004. Role of histone H2A ubiquitination in Polycomb silencing. *Nature,* 431**,** 873-8.

WANG, J., HEVI, S., KURASH, J. K., LEI, H., GAY, F., BAJKO, J., SU, H., SUN, W., CHANG, H., XU, G., GAUDET, F., LI, E. & CHEN, T. 2009a. The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nat Genet,* 41**,** 125-9.

WANG, J., SCULLY, K., ZHU, X., CAI, L., ZHANG, J., PREFONTAINE, G. G., KRONES, A., OHGI, K. A., ZHU, P., GARCIA-BASSETS, I., LIU, F., TAYLOR, H., LOZACH, J., JAYES, F. L., KORACH, K. S., GLASS, C. K., FU, X. D. & ROSENFELD, M. G. 2007. Opposing LSD1 complexes function in developmental gene activation and repression programmes. *Nature,* 446**,** 882-7.

WANG, M., LIU, X., GUO, J., WENG, X., JIANG, G., WANG, Z. & HE, L. 2015. Inhibition of LSD1 by Pargyline inhibited process of EMT and delayed progression of prostate cancer in vivo. *Biochemical and Biophysical Research Communications,* 467**,** 310-315.

WANG, Y., ZHANG, H., CHEN, Y., SUN, Y., YANG, F., YU, W., LIANG, J., SUN, L., YANG, X., SHI, L., LI, R., LI, Y., ZHANG, Y., LI, Q., YI, X. & SHANG, Y. 2009b. LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer. *Cell,* 138**,** 660-72.

WANG, Z., ZANG, C., ROSENFELD, J. A., SCHONES, D. E., BARSKI, A., CUDDAPAH, S., CUI, K., ROH, T. Y., PENG, W., ZHANG, M. Q. & ZHAO, K. 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet,* 40**,** 897-903.

WANG, Z. A., MILLARD, C. J., LIN, C.-L., GURNETT, J. E., WU, M., LEE, K., FAIRALL, L., SCHWABE, J. W. R. & COLE, P. A. 2020. Diverse nucleosome Site-Selectivity among histone deacetylase complexes. *eLife,* 9**,** e57663.

WASSON, J. A., SIMON, A. K., MYRICK, D. A., WOLF, G., DRISCOLL, S., PFAFF, S. L., MACFARLAN, T. S. & KATZ, D. J. 2016. Maternally provided LSD1/KDM1A enables the maternal-to-zygotic transition and prevents defects that manifest postnatally. *Elife,* 5.

WASYLYK, B. & CHAMBON, P. 1979. Transcription by Eukaryotic RNA Polymerases A and B of Chromatin Assembled in vitro. *European Journal of Biochemistry,* 98**,** 317-327.

WEINBERGER, L., AYYASH, M., NOVERSHTERN, N. & HANNA, J. H. 2016. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat Rev Mol Cell Biol,* 17**,** 155-69.

WELCKER, J. E., HERNANDEZ-MIRANDA, L. R., PAUL, F. E., JIA, S., IVANOV, A., SELBACH, M. & BIRCHMEIER, C. 2013. Insm1 controls development of pituitary endocrine cells and requires a SNAG domain for function and for recruitment of histone-modifying factors. *Development,* 140**,** 4947-58.

WHITE, P. H., FARKAS, D. R., MCFADDEN, E. E. & CHAPMAN, D. L. 2003. Defective somite patterning in mouse embryos with reduced levels ofTbx6. *Development,* 130**,** 1681-1690.

WHITMAN, M. 2001. Nodal Signaling in Early Vertebrate Embryos: Themes and Variations. *Developmental Cell,* 1**,** 605-617.

WHYTE, W. A., BILODEAU, S., ORLANDO, D. A., HOKE, H. A., FRAMPTON, G. M., FOSTER, C. T., COWLEY, S. M. & YOUNG, R. A. 2012. Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature,* 482**,** 221-5.

WILLIAMS, R. L., HILTON, D. J., PEASE, S., WILLSON, T. A., STEWART, C. L., GEARING, D. P., WAGNER, E. F., METCALF, D., NICOLA, N. A. & GOUGH, N. M. 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature,* 336**,** 684-687.

WRAY, J., KALKAN, T., GOMEZ-LOPEZ, S., ECKARDT, D., COOK, A., KEMLER, R. & SMITH, A. 2011. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat Cell Biol,* 13**,** 838-45.

WU, H., MATHIOUDAKIS, N., DIAGOURAGA, B., DONG, A., DOMBROVSKI, L., BAUDAT, F., CUSACK, S., DE MASSY, B. & KADLEC, J. 2013. Molecular Basis for the Regulation of the H3K4 Methyltransferase Activity of PRDM9. *Cell Reports,* 5**,** 13-20.

WU, M., HAYWARD, D., KALIN, J. H., SONG, Y., SCHWABE, J. W. R. & COLE, P. A. 2018. Lysine-14 acetylation of histone H3 in chromatin confers resistance to the deacetylase and demethylase activities of an epigenetic silencing complex. *eLife,* 7**,** e37231.

XU, H. & VAN REMMEN, H. 2021. The SarcoEndoplasmic Reticulum Calcium ATPase (SERCA) pump: a potential target for intervention in aging and skeletal muscle pathologies. *Skeletal Muscle,* 11**,** 25.

YAMAGUCHI, T. P., TAKADA, S., YOSHIKAWA, Y., WU, N. & MCMAHON, A. P. 1999. T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. *Genes Dev,* 13**,** 3185-90.

YAMAMOTO, M., BEPPU, H., TAKAOKA, K., MENO, C., LI, E., MIYAZONO, K. & HAMADA, H. 2009. Antagonism between Smad1 and Smad2 signaling determines the site of distal visceral endoderm formation in the mouse embryo. *Journal of Cell Biology,* 184**,** 323-334.

YANG, J.-H., WYLIE-SEARS, J. & BISCHOFF, J. 2008. Opposing actions of Notch1 and VEGF in post-natal cardiac valve endothelial cells. *Biochemical and Biophysical Research Communications,* 374**,** 512-516.

YANG, M., CULHANE, J. C., SZEWCZUK, L. M., GOCKE, C. B., BRAUTIGAM, C. A., TOMCHICK, D. R., MACHIUS, M., COLE, P. A. & YU, H. 2007. Structural basis of histone demethylation by LSD1 revealed by suicide inactivation. *Nat Struct Mol Biol,* 14**,** 535-9.

YANG, M., GOCKE, C. B., LUO, X., BOREK, D., TOMCHICK, D. R., MACHIUS, M., OTWINOWSKI, Z. & YU, H. 2006. Structural basis for CoREST-dependent demethylation of nucleosomes by the human LSD1 histone demethylase. *Mol Cell,* 23**,** 377-87.

YANG, X. J. & SETO, E. 2008. The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. *Nat Rev Mol Cell Biol,* 9**,** 206-18.

YANG, Y., YIN, X., YANG, H. & XU, Y. 2015. Histone Demethylase LSD2 Acts as an E3 Ubiquitin Ligase and Inhibits Cancer Cell Growth through Promoting Proteasomal Degradation of OGT. *Molecular Cell,* 58**,** 47-59.

YAP, K. L. & ZHOU, M. M. 2011. Structure and mechanisms of lysine methylation recognition by the chromodomain in gene transcription. *Biochemistry,* 50**,** 1966-80.

YIN, Y., MORGUNOVA, E., JOLMA, A., KAASINEN, E., SAHU, B., KHUND-SAYEED, S., DAS, P. K., KIVIOJA, T., DAVE, K., ZHONG, F., NITTA, K. R., TAIPALE, M., POPOV, A., GINNO, P. A., DOMCKE, S., YAN, J., SCHÜBELER, D., VINSON, C. & TAIPALE, J. 2017. Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science,* 356**,** eaaj2239.

YING, Q.-L., NICHOLS, J., CHAMBERS, I. & SMITH, A. 2003. BMP Induction of Id Proteins Suppresses Differentiation and Sustains Embryonic Stem Cell Self-Renewal in Collaboration with STAT3. *Cell,* 115**,** 281-292.

YING, Q.-L., WRAY, J., NICHOLS, J., BATLLE-MORERA, L., DOBLE, B., WOODGETT, J., COHEN, P. & SMITH, A. 2008. The ground state of embryonic stem cell self-renewal. *Nature,* 453**,** 519-523.

YOU, A., TONG, J. K., GROZINGER, C. M. & SCHREIBER, S. L. 2001. CoREST is an integral component of the CoREST- human histone deacetylase complex. *Proc Natl Acad Sci U S A,* 98**,** 1454-8.

ZALZMAN, M., FALCO, G., SHAROVA, L. V., NISHIYAMA, A., THOMAS, M., LEE, S.-L., STAGG, C. A., HOANG, H. G., YANG, H.-T., INDIG, F. E., WERSTO, R. P. & KO, M. S. H. 2010. Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature,* 464**,** 858-863.

ZENG, L., ZHANG, Q., LI, S., PLOTNIKOV, A. N., WALSH, M. J. & ZHOU, M.-M. 2010. Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature,* 466**,** 258-262.

ZHANG, Q., QI, S., XU, M., YU, L., TAO, Y., DENG, Z., WU, W., LI, J., CHEN, Z. & WONG, J. 2013a. Structure-function analysis reveals a novel mechanism for regulation of histone demethylase LSD2/AOF1/KDM1b. *Cell Research,* 23**,** 225-241.

ZHANG, W., TYL, M., WARD, R., SOBOTT, F., MAMAN, J., MURTHY, A. S., WATSON, A. A., FEDOROV, O., BOWMAN, A., OWEN-HUGHES, T., EL MKAMI, H., MURZINA, N. V., NORMAN, D. G. & LAUE, E. D. 2013b. Structural plasticity of histones H3–H4 facilitates their allosteric exchange between RbAp48 and ASF1. *Nature Structural & Molecular Biology,* 20**,** 29-35.

ZHANG, Y., NG, H.-H., ERDJUMENT-BROMAGE, H., TEMPST, P., BIRD, A. & REINBERG, D. 1999. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes & Development,* 13**,** 1924-1935.

ZHAO, D., LI, Y., XIONG, X., CHEN, Z. & LI, H. 2017. YEATS Domain—A Histone Acylation Reader in Health and Disease. *Journal of Molecular Biology,* 429**,** 1994-2002.

ZHAO, G. Q. 2003. Consequences of knocking out BMP signaling in the mouse. *Genesis,* 35**,** 43-56.

ZHONG, W., XU, Z., WEN, S., XIE, T., WANG, F., WANG, Q. & CHEN, J. 2019. Long non-coding RNA myocardial infarction associated transcript promotes epithelial-mesenchymal transition and is an independent risk factor for poor prognosis of tongue squamous cell carcinoma. *Journal of Oral Pathology & Medicine,* 48**,** 720-727.