CHARACTERIZATION OF RHESUS MONKEY EMBRYONIC STEM CELLS IN PRIMED AND NAÏVE-LIKE STATES OF PLURIPOTENCY USING FOURIER TRANSFORM INFRARED (FTIR) MICROSPECTROSCOPY

JITTANUN SRISUTUSH

A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Biotechnology
Suranaree University of Technology
Academic Year 2025

การตรวจสอบคุณสมบัติของเซลล์ต้นกำเนิดตัวอ่อนลิงวอกระยะไพรม์ และระยะคล้ายนาอีฟ ของสภาวะพลูริโพเทนซี โดยใช้ Fourier Transform Infrared (FTIR) Microspectroscopy

นางสาวจิตตนันท์ ศรีสุทัศน์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชาเทคโนโลยีชีวภาพ
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2568

CHARACTERIZATION OF RHESUS MONKEY EMBRYONIC STEM CELLS IN PRIMED AND NAÏVE-LIKE STATES OF PLURIPOTENCY USING FOURIER TRANSFORM INFRARED (FTIR) MICROSPECTROSCOPY

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

Thesis Examining Committee

.....

(Asst. Prof. Dr. Ruttachuk Rungsiwiwut)
Chairperson

.....

(Prof. Dr. Rangsun Parnpai)
Member (Thesis Advisor)

.....

(Dr. Irene Aksoy)
Member (Thesis Co-Advisor)

Kanjana Thumanu
.....

(Dr. Kanjana Thumanu)
Member (Thesis Co-Advisor)

.....

(Prof. Dr. Pierre Savatier)
Member

.....

(Assoc. Prof. Dr. Yupaporn Ruksakulpiwat)
Acting Vice Rector for Academic Affairs
and Quality Assurance

.....

(Prof. Dr. Neung Teaumroong)
Dean of Institute of Agricultural
Technology

จิตตนันท์ ศรีสุทัศน์: การตรวจสอบคุณสมบัติของเซลล์ต้นกำเนิดตัวอ่อนลิงวอกระยะไพรม์และระยะคล้ายนาอีฟ ของสภาวะพลูริโพเทนซี โดยใช้ Fourier Transform Infrared (FTIR) Microspectroscopy (CHARACTERIZATION OF RHESUS MONKEY EMBRYONIC STEM CELLS IN PRIMED AND NAÏVE-LIKE STATES OF PLURIPOTENCY USING FOURIER TRANSFORM INFRARED (FTIR) MICROSPECTROSCOPY) อาจารย์ที่ปรึกษา: ศาสตราจารย์ ดร.รังสรรค์ พาลพ่าย, 54 หน้า

คำสำคัญ: ฟูเรียร์ทรานส์ฟอร์มอินฟราเรดไมโครสเปกโตรสโคปี/ลิงวอก/เซลล์ต้นกำเนิดตัวอ่อน/ระยะไพรม์/ระยะคล้ายนาอีฟ

งานวิจัยนี้ได้ใช้ Fourier transform infrared (FTIR) microspectroscopy เพื่อศึกษาลักษณะของเซลล์ต้นกำเนิดตัวอ่อนลิงวอก (rhESCs) ซึ่งอยู่ในระยะไพรม์ และระยะคล้ายนาอีฟ ของสภาวะพลูริโพเทนซี เซลล์สองสายพันธุ์มีที่มาจากเซลล์ต้นกำเนิดเดียวกัน (rhESCs-FGF2/KOSR) และนำมาเพาะเลี้ยงในสภาวะที่แตกต่างกันเพื่อให้ได้สภาวะคล้ายนาอีฟ คือ rhESCs-ALGöX Immunocytochemistry และ RNA sequencing จะถูกใช้เปรียบเทียบเพื่อยืนยันผลที่ได้จาก FTIR สเปกตรัม ด้วยเครื่องหมายจำเพาะระยะนาอีฟ (KLF17, ALPPL2, TFCP2L1 และ TFAP2C) และเครื่องหมายจำเพาะระยะไพรม์ (OTX2) ยืนยันสภาวะของเซลล์แต่ละระยะได้อย่างชัดเจน ขณะที่ข้อมูลจากการถอดรหัสพันธุกรรมยังแสดงให้เห็นถึงความแตกต่างของการแสดงออกของยีนที่ชัดเจนเช่นกัน rhESCs-FGF2/KOSR แสดงการแสดงออกของยีนในระยะไพรม์สูงขึ้น ได้แก่ *NODAL*, *OTX2*, *ETV4*, *BMP4*, *FST* และ *SOX3* ในทางตรงกันข้าม rhESCs-ALGöX แสดงการแสดงออกของยีนที่เกี่ยวข้องกับระยะนาอีฟสูงขึ้น ได้แก่ *KLF2*, *DPPA2*, *DPPA3*, *ZFP42*, *PRDM14* และ *TFCP2L1* การวิเคราะห์ด้วย FTIR เผยให้เห็นความแตกต่างทางสเปกตรัมอย่างมีนัยสำคัญระหว่างเซลล์ระยะไพรม์และระยะคล้ายนาอีฟ เซลล์ระยะคล้ายนาอีฟแสดงค่าการดูดกลืนแสงสูงกว่าในบริเวณที่เกี่ยวข้องกับโปรตีน (แถบเอไมด์หนึ่ง และสอง ที่ 1654 และ 1546 เซ็นติเมตร$^{-1}$) และกรดนิวคลีอิก (1240 และ 1080 เซนติเมตร$^{-1}$) ในขณะที่เซลล์ระยะไพรม์ มีค่าการดูดกลืนแสงสูงกว่าในแถบที่เกี่ยวข้องกับไขมัน (สเปกตรัมการยืดหดของพันธะ C–H ที่ 2921-2852 เซ็นติเมตร$^{-1}$ และ การยืดหดของพันธะลิปิดเอสเทอร์ C=O ที่ 1741 เซ็นติเมตร$^{-1}$) การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis; PCA) ของ FTIR สเปกตรัม ของเซลล์ต้นกำเนิดตัวอ่อนลิงวอก ที่ได้จากการเพาะเลี้ยงในสองสภาวะที่แตกต่างกัน แสดงให้เห็นว่าสเปกตรัมของ rhESCs-FGF2/KOSR และ rhESCs-ALGöX สามารถแบ่งแยกออกจากกันด้วยองค์ประกอบหลักที่ 1 โดยสามารถอธิบายความแปรปรวนทั้งหมดในชุดข้อมูลได้ถึงร้อยละ 64 การวิเคราะห์จำแนกความแตกต่างแบบกำลังสองน้อยที่สุดบางส่วน (Partial Least Squares–Discriminant Analysis; PLS–DA) สามารถจำแนกสภาวะของเซลล์ทั้งสองได้อย่างมีประสิทธิภาพด้วยความไว 100% และความจำเพาะ 100% ผลการวิจัยนี้แสดงให้เห็นว่า FTIR microspectroscopy สามารถจำแนกลักษณะทางชีวเคมีเฉพาะของสภาวะพลูริโพเทนต์ ในเซลล์ต้น

กำเนิดตัวอ่อนลิงวอก ซึ่งรวดเร็ว และไม่ต้องใช้สารบ่งชี้ในการตรวจสอบเอกลักษณ์และคุณภาพของเซลล์ต้นกำเนิด

สาขาวิชาเทคโนโลยีชีวภาพ
ปีการศึกษา 2568

ลายมือชื่อนักศึกษา .....
ลายมือชื่ออาจารย์ที่ปรึกษา .....
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม .....
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม Kanjana Thumanu

JITTANUN SRISUTUSH: CHARACTERIZATION OF RHESUS MONKEY EMBRYONIC STEM CELLS IN PRIMED AND NAÏVE-LIKE STATES OF PLURIPOTENCY USING FOURIER TRANSFORM INFRARED (FTIR) MICROSPECTROSCOPY. THESIS ADVISOR: PROF. RANGSUN PARNPAI, PH.D., 54 PP

Keyword: Fourier Transform Infrared/Rhesus macaque/Embryonic Stem Cells/primed states/naïve-like cell states

This research employed Fourier transform infrared (FTIR) microspectroscopy to characterize rhESCs in primed and naïve-like states of pluripotency. Two cell lines, derived from a common parental line (rhESC-FGF2/KOSR), were cultured under distinct conditions to establish naïve-like states: rhESC-ALGöX. Immunocytochemistry and RNA sequencing were used in parallel to validate FTIR spectra findings. Naïve markers (KLF17, ALPPL2, TFCP2L1 and TFAP2C) and primed markers (OTX2) confirmed the respective cell states, while transcriptomic profiling further demonstrated clear gene expression differences. rhESCs-FGF2/KOSR showed elevated expression of primed state genes including *NODAL*, *OTX2*, *ETV4*, *BMP4*, *FST*, and *SOX3*, whereas rhESCs-ALGöX exhibited high expression of naïve–associated genes including *KLF2*, *DPPA2*, *DPPA3*, *ZFP42*, *PRDM14*, and *TFCP2L1*. FTIR analysis revealed significant spectral differences between primed and naïve-like cells. Naïve-like cells showed stronger absorbance in regions associated with protein associated bands (Amide I and II bands at 1654 and 1546 cm^{-1}) and nucleic acids (1240 and 1080 cm^{-1}), while primed cells exhibited stronger absorbance in regions in lipid–associated bands (C–H stretching at 2921-2852 cm^{-1} and lipid ester C=O stretching at 1741 cm^{-1}). Principal Component Analysis (PCA) of FTIR spectra of rhESCs cultured under distinct conditions revealed that the spectra of rhESC-FGF2/KOSR and rhESC-ALGöX could discriminated in scores plots along PC1, which can be explained by the 64% of the total variance in the dataset. Partial Least Squares–Discriminant Analysis (PLS–DA) effectively classified the two cell states with 100% sensitivity and 100% specificity. These findings demonstrate that FTIR microspectroscopy can reliably discriminate pluripotent state specific biochemical

features in rhESCs, providing a rapid and label-free approach for monitoring stem cell identity and quality.

School of Biotechnology
Academic Year 2025

Student's Signature .....
Advisor's Signature .....
Co-Advisor's Signature .....
Co-Advisor's Signature Kanjana Thumanu

ACKNOWLEDGEMENTS

This work was supported by One Research One Graduate (OROG) fellowship, SUT (53/2565) Suranaree University of Technology. Apart from my efforts, the complement of this dissertation would not have been possible without the encouragement, kind support and guidelines of many others.

First and foremost, I wish to express my deepest gratitude and my sincerest thank to my supervisor Prof. Dr. Rangsun Parnpai for his guidance and encouragement. Thank you very much for providing the scholarship, financial supports and laboratory facilities. I also wish to express my deep sense of gratitude to Dr. Irene Aksoy and Dr. Kanjana Thumanu, my co–advisors, for their advice, help and technical support throughout my thesis work. My appreciation is also to Prof. Dr. Pierre Savatier for his advice and impressive experience. I would like to thank Cloé Rognard and Anaïs Amzal, my co–workers, for their help, technical support immunocytochemistry analysis and RNA sequencing.

I gratefully acknowledge Dr. Supatcharee Cael and Dr. Kanokwan Kamkajon for their assistance and guidance in FTIR microspectroscopy analysis. Special thanks also go to Miss. Chunmanus Uthaisar for her valuable support in set up FTIR microspectroscopy machine for analysis. I would like to thank my junior, Ms. Siriphon Suwanna, for her willing help in immunocytochemistry analysis. My thanks and appreciations also go to the members of Embryo Technology and Stem cell Research Center.

A million thanks to my mother, Mrs. Puan Singwankham for warm wishes and financial support.

Finally, I am forever indebted to my beloved families for their understanding, endless love and encouragement which were most required throughout my studies.

Jittanun Srisutush

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF ABBREVIATIONS

PSCs	=	Pluripotent stem cells
ESCs	=	Embryonic stem cells
EpiSCs	=	Epiblast stem cells
iPSCs	=	Induced pluripotent stem cells
NHPs	=	non–human primates
ICM	=	Inner cell mass
IR	=	Infrared
FTIR	=	Fourier transform infrared
mESCs	=	Mouse embryonic stem cells
hESCs	=	Human embryonic stem cells
rhESCs	=	Rhesus macaque embryonic stem cells
IVF	=	*in vitro* fertilization
ICSI	=	Intracytoplasmic sperm injection
HIV	=	Human immunodeficiency virus
AIDS	=	acquired immunodeficiency syndrome
SIV	=	Simian immunodeficiency virus
ICC	=	Immunocytochemistry
RNA–seq	=	RNA sequencing
NGS	=	Next generation sequencing
cDNA	=	Complementary DNA
mRNA	=	messenger RNA
rRNA	=	ribosomal RNA
qPCR	=	Quantitative PCR
RT–PCR	=	Reverse transcription polymerase chain reaction
SNVs	=	Single nucleotide variants
OPD	=	Optical path difference
ATR	=	Attenuated total reflection
HCA	=	Hierarchical cluster analysis
DTGS	=	Deuterated triglycine sulfate
MCT	=	Mercuric cadmium telluride
FPA	=	Focal plane array
PCA	=	Principal component analysis

LIST OF ABBREVIATIONS (Continued)

EMSC	=	Extended multiplicative signal correction
LDA	=	Linear discriminant analysis
PLS–DA	=	Partial least squares–discriminant analysis
ANN	=	Artificial neural networks
MEFs	=	Mouse embryonic fibroblasts
CM	=	Conditioned medium
FGF2	=	Fibroblast growth factor 2
KOSR	=	Knock out serum replacement
P	=	Passage
BaF2	=	Barium fluoride
ml	=	Milliliter
mM	=	Millimolar
mm	=	Millimeter
ng	=	Nanogram
rpm	=	Revolutions per minute
µl	=	Microliter
µg/ml	=	Microgram per milliliter
µm	=	Micrometer

CHAPTER I
INTRODUCTION

1.1 Introduction

Pluripotent stem cells (PSCs), derived from the epiblast of mammalian blastocysts, have the dual capacity for unlimited self–renewal and differentiation into derivatives of all three germ layers (Evans and Kaufman, 1981). In rodents, PSCs can exist in two distinct states: the naïve state, represented by embryonic stem cells (ESCs) derived from inner cell mass (ICM) of the blastocysts, and the primed state, represented by epiblast stem cells (EpiSCs) derived from post implantation epiblast tissue (Nichols and Smith, 2009). These states are also recognized in human and non–human primates (NHPs), including rhesus macaques (*Macaca mulatta*), where PSCs can stabilize in either a naïve and primed configuration depending on culture conditions and signaling inputs (Anwised et al., 2023). Naïve and primed PSCs differ markedly in their transcriptional, epigenetic, and metabolic profiles, which in turn have a profound impact on their functional properties (Weinberger et al., 2016). Notably, only naïve PSCs in the naïve state exhibit the capacity to integrate into host preimplantation embryos and contribute to development, as demonstrated in rodents (Nichols and Smith, 2009), rabbits (Pham et al., 2025), and cynomolgus monkeys (Li et al., 2023). These functional differences are underpinned by distinct regulatory networks and chromatin landscapes: naïve PSCs display global DNA hypomethylation, two active X chromosomes in female cells, and reduced levels of H3K27me3 at developmental gene loci, whereas primed PSCs show higher DNA methylation, one inactive X chromosome in female cells, and accumulation of repressive histone marks (Weinberger et al., 2016). Identifying the pluripotency state of a given PSC line traditionally requires multimodal analyses, including bulk RNA sequencing (Wang et al., 2009), immunostaining (Cregger et al., 2006), state–specific markers (e.g., KLF4, TFCP2L1 for naïve; OTX2 for primed), quantitative PCR (qPCR) (Higuchi et al., 1993; Bustin, 2000), reporter–based functional assays, and epigenomic profiling. These approaches, while informative, are often labor intensive, sample preparation, time consuming procedures, a large number of samples needed, not readily applicable for routine quality control, and in addition, each sample preparation method can available for only one marker at a time. Although qPCR is an accurate method to verify such specific markers, a large number of cells are required for such testing. Moreover, the presence of transcripts

does not guarantee expression at the protein level (Chonanant et al., 2011). In this context, Fourier–transform infrared (FTIR) spectroscopy offers an attractive, non–invasive alternative for cell characterization (Cao et al., 2013b).

FTIR microspectroscopy analyzes vibrational energy absorption in molecular bonds, providing a composite biochemical fingerprint of cells. The technique enables rapid, can measure both transmission and reflection (Thumanu et al., 2011) and label–free quantification of major cellular components—including proteins, lipids, carbohydrates, and nucleic acids—at single cell resolution. Importantly, changes in cell state are accompanied by alterations in molecular composition and structure, which can be detected in specific IR absorption bands. For example, shifts in the amide I and II regions (1700–1500 cm^{-1}) reflect differences in protein secondary structure and abundance; CH_2 and CH_3 stretching modes (3000–2800 cm^{-1}) are indicative of membrane lipid composition and fluidity; and bands in the 1200–800 cm^{-1} region reflect changes in nucleic acid conformation and RNA/DNA content (Ami et al., 2008; Bassan et al., 2009; 1Thumanu et al., 2011; Cao et al., 2013b). Nucleic acid absorption region (1050-850 cm^{-1}) pointed towards high levels of messenger RNA (mRNA) translation and production of specific proteins indicating that the cells harbor a new phenotype. Furthermore, the DNA/RNA hybrid bands at 954 cm^{-1} and 899 cm^{-1} were also observed, suggesting transcriptional switch that started with the differentiation of the cells (Ami et al., 2008). Principal component analysis (PCA) and subsequent linear discriminant analysis (LDA) were employed to effectively segregate stem cell spectra into distinct clusters, facilitating the identification of the most noteworthy spectral changes.

The first study that aimed to characterizing specific ESCs derived cells (Ami et al., 2008). The authors used the FTIR method to monitor the spontaneous differentiation of mouse embryonic stem cells (mESCs). Heraud et al. (2010) utilized focal plane array (FPA) FTIR microspectroscopy to examine human embryonic stem cells (hESCs). Their findings suggest that FTIR spectroscopy possesses the capability to characterize the macromolecular composition of cells and effectively discriminate hESCs that differentiate into either ectoderm, endoderm, or mesoderm. This discriminatory potential is further supported by multivariate data analysis techniques such as PCA, partial least squares discriminant analysis (PLS–DA), and artificial neural networks (ANN). Previous bio–spectroscopic studies have successfully applied FTIR to distinguish undifferentiated from differentiated cells, to assess stem cell lineage commitment, and to monitor reprogramming efficiency (Cao et al., 2013a). However, no study has explored its application for distinguishing between naïve and primed states of pluripotency, particularly in non–human primates. Establishing such a methodology

could provide a powerful tool for monitoring cell identity and quality in basic and translational stem cell research. There are 2 types of infrared radiation detectors; 1) mercuric cadmium telluride (MCT) detector measures 1 sample/ 1 spectrum and 2) FPA detectors measure maximum 4,096 spectrums in one time; FPA detectors can create a distribution diagram faster than MCT detectors (Wold et al., 1987). A technique known as FPA infrared imaging with multichannel detectors has recently become available for analyzing cells and tissues with allowing to measure all data points from each detector element simultaneously.

In this study, we applied FPA–FTIR microspectroscopy to rhesus macaque embryonic stem cells (rhESCs) cultured under two distinct conditions: the well–established FGF2/KOSR medium, which supports the primed state, and a novel medium we developed, ALGöX (Pham et al., 2025), designed to support a naïve-like pluripotent state. ALGöX consists of fibroblast conditioned N2B27 medium supplemented with activin A, leukemia inhibitory factor (LIF), PKC inhibitor Gö6983, and tankyrase inhibitor XAV939. This combination aims to inhibit MEK/ERK and WNT signaling while maintaining key self–renewal cues, based on pathways shown to stabilize naïve pluripotency in rodent and human systems. We show that rhESCs cultured in ALGöX exhibit distinct spectral features compared to those in FGF2/KOSR.

1.2 Research objectives

1) To characterize and compare between passage cells in primed and naïve-like states rhESCs by standard methods.
2) To characterize and compare between passage cells in primed and naïve-like states rhESCs by FPA–FTIR microspectroscopy.
3) To correlation analysis between standard characterization methods and FPA–FTIR microspectroscopy.
4) To comparison chemical change from each passage on the biological replication by FPA–FTIR microspectroscopy.

1.3 Research hypotheses

1) rhESCs had different chemical compositions those could characterize and identify primed and naïve-like states.
2) FTIR microspectroscopy could distinguished primed and naïve-like states rhESCs.

1.4 Scope of research

1) rhESCs-FGF2/KOSR cells (primed state) and rhESCs-ALGöX cells (naïve-like state) used in this study were obtained from INSERM, Stem Cell and Brain Research Institute, Lyon, France. Both types of cells were characterized by immunocytochemistry (ICC), RNA sequencing (RNA–seq) and FPA–FTIR microspectroscopy techniques.
2) FTIR spectra of rhESCs in primed and naïve-like states of pluripotency were analyzed used FPA–FTIR microspectroscopy.

CHAPTER II
LITERATURE REVIEW

2.1 Pluripotent stem cells (PSCs)

2.1.1 A defining characteristic of pluripotency

PSCs are a class of stem cells defined by their extraordinary ability to self-renew indefinitely in an undifferentiated state and to differentiate into all three embryonic germ layers: the ectoderm, mesoderm, and endoderm (Thomson et al., 1998a; Yamanaka and Thomson, 2010). This key characteristic of pluripotency makes them invaluable for regenerative medicine, drug discovery, and developmental biology research. The concept of pluripotency is distinct from totipotency, which describes a cell's ability to form all cell types of a complete organism, including extra-embryonic tissues like the placenta (Thomson et al., 1995). PSCs are the progenitors of all cell types within the body proper, but they cannot give rise to extra-embryonic tissues on their own.

PSCs can broadly be classified into two main types based on their origin: ESCs and induced pluripotent stem cells (iPSCs). While ESCs are derived directly from the inner cell mass of a blastocyst-stage embryo, iPSCs are generated artificially by reprogramming somatic cells back to a pluripotent state (Takahashi et al., 2007). Both types of cells share a common set of pluripotency markers, including the transcription factors *OCT4*, *SOX2*, and *NANOG*, which form a core regulatory network that maintains the undifferentiated state (Jaenisch and Young, 2008).

2.1.2 Embryonic stem cells (ESCs)

ESCs are arguably the most well characterized and foundational type of PSCs. They are isolated from the inner cell mass of a pre implantation embryo, typically at the blastocyst stage (Thomson et al., 1998b). The first successful isolation of mESCs was reported in 1981, and this was followed by the landmark isolation of hESCs by Thomson et al. (1998). This discovery unlocked unprecedented opportunities for studying human development and disease.

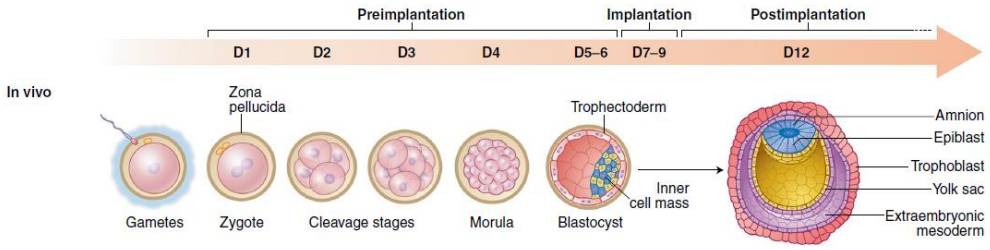

Figure 2.1 Early human development *in vivo*. (Mina et al., 2021).

A key advantage of ESCs is their robust and stable pluripotent state. They can be maintained in culture for long periods without losing their differentiation potential or acquiring significant genetic abnormalities, as long as appropriate culture conditions are maintained (Thomson et al., 1998b). The pluripotency of ESCs is functionally confirmed by their ability to differentiate into all three germ layers *in vitro* and, in the case of *in vivo* assays, by their ability to form teratomas benign tumors containing differentiated cells from all three germ layers when injected into immunocompromised animals (Adewumi et al., 2007).

Despite their scientific promise, the use of ESCs is associated with significant ethical considerations due to their embryonic origin (Thomson et al., 1998b). Additionally, their allogeneic nature means that derivatives of ESCs would subject to immune rejection when transplanted into a patient, necessitating immunosuppressive therapy or the development of immunologically matched cell banks (Depciuch et al., 2018). These challenges have driven the search for alternative sources of PSCs, leading to the discovery of iPSCs.

2.1.3 The relationship between PSCs and ESCs

The term "pluripotent stem cell" is a broader classification that encompasses ESCs. ESCs were the first type of PSCs to isolated and characterized, and they serve as the benchmark against which all other PSCs, particularly iPSCs, are measured (Jaenisch and Young, 2008). While iPSCs have largely overcome the ethical and immunological barriers associated with ESCs, a significant body of research is still dedicated to comparing the two cell types to ensure that iPSCs are functionally equivalent to ESCs (Szymanski et al., 2018). For example, studies often compare the gene expression profiles, epigenetic states, and differentiation efficiencies of iPSCs and ESCs to confirm that the reprogramming process has fully restored pluripotency (Reis et al., 2020).

mESCs were isolated first in 1981 (Evans and Kaufman 1981; Martin 1981). They were obtained by immunosurgically digesting the outer layer (trophectoderm) of mouse blastocysts to isolate the ICM that will adhere to the culture dish and

expand. The ICM is comprised of epiblast cells that give rise to the entire adult mouse. The mESCs harbor a normal karyotype even after long term culture. When initially isolated, mESCs exhibited enhanced growth and showed a higher propensity for differentiation when cultivated on mitotically inactivated mouse embryonic fibroblast (MEFs) feeder cells (Martin and Evans, 1974). The term "feeder" originally was adopted to define a cell providing critical nutrients or trophic factors support for the maintenance of an undifferentiated stem cell state. Furthermore, feeder cells may remove deleterious components from the media or may involve in cell–contact mediated mechanisms to prevent differentiation of the ESCs. The precise composition of the molecules secreted by feeder cells in any ESCs culture is currently unknown. mESCs, when injected into host blastocysts, produce viable chimaeras, by contributing to all fetal tissue lineages including the germline (Bradley et al., 1984).

ESCs from non–human primates (nhpESCs) and human (hESCs) were first derived in the mid–late–1990s by Thomson and colleagues (Thomson et al., 1995; Thomson et al., 1998c).

The hESCs have been shown to harbor a normal karyotype even after long-term culture, to express high levels of telomerase activity, and to express specific cell surface markers including SSEA3, SSEA4, TRA–1–60 and TRA–1–81. Even after extended periods of culture, these cells still maintain their developmental potential in derivatives of all three embryonic germ layers, including 1) endoderm; gut epithelium 2) mesoderm; cartilage, bone, smooth muscle and striated muscle 3) ectoderm; neural epithelium, embryonic ganglia and stratified squamous epithelium. In addition, they can give rise to derivatives of the trophoblast. These cell lines hold great promises in human developmental biology, drug discovery, and regenerative medicine.

The nhpESCs lines from rhesus monkey (*Macaca mulatta*, rhESCs) were first isolated in 1995 (Thomson et al., 1995). These cells, harbor a normal karyotype, and expresses the same cell surface markers as hESCs. The rhESCs remain undifferentiated when grown on MEFs feeder layers but differentiate or die in the absence of fibroblasts. Similarly to mESCs, they require different culture conditions for their maintenance in the pluripotent state. The rhESCs cells allowed to differentiate *in vitro* secrete bioactive chorionic gonadotropin into the medium, express chorionic gonadotropin alpha– and beta–subunit mRNAs, and express alpha fetoprotein mRNA, indicating trophoblast and endoderm differentiation. When injected into severe combined

immunodeficient mice, rhESCs formed teratomas composed of derivatives of all three embryonic germ layers.

The nhpESCs lines from other species, including the common marmoset (*Callithrix jacchus*, Thomson et al., 1996), and the cynomolgus monkey ESCs (cyESCs, Suemori et al., 2001), were also derived from blastocyst stage embryos obtained by *in vitro* fertilization (IVF) and intracytoplasmic sperm injection (ICSI).

In summary, ESCs are a specific and highly important type of PSCs, while the term PSCs describes any cell with the defining characteristics of pluripotency, regardless of its origin. The foundational work with ESCs has been instrumental in the entire field of stem cell biology, and their continued study, alongside iPSCs, is crucial for advancing our understanding of pluripotency and for realizing the full potential of regenerative medicine.

2.2 Characteristics of primed, naïve-like, and naïve pluripotent stem cell states

PSCs are defined by their ability to self-renew and differentiate into all three embryonic germ layers (Thomson et al., 1998a). However, this pluripotency is not a single, static state. Instead, it exists along a continuum, with different states representing varying levels of developmental maturity and functional properties (Nichols and Smith, 2009). The two most well-defined states in this spectrum are naïve and primed pluripotency. More recently, a transitional or intermediate state, often referred to as naïve-like, has also been characterized, primarily in human stem cells. Understanding these distinct states is crucial for their reliable use in research and therapeutic applications.

Table 2.1 The first ESC lines derived from nonhuman primates (NHPs) (Anwised et al., 2023)

Species	Name of cell lines	Culture medium	Cells differentiation analysis	Reference
Marmoset (*C. jacchus*)	Cj11.2, Cj25.1, Cj28, Cj33, Cj35, Cj36, Cj39, Cj62	DMEM culture medium/ feeder–MEF	embryoid bodies	Thomson et al., 1996
Rhesus monkey (*M. mulatta*)	R278, R366, R367, R394, R420, R456, R460	DMEM medium/ feeder–MEF	embryoid bodies teratomas	Thomson et al., 1995
Cynomolgus monkey (*M. fascicularis*)	CMK5, CMK6, CMK7, CMK9	DMEM medium and Ham's nutrient mixture F–12/ feeder–MEF	embryoid bodies teratomas	Suemori et al., 2001

2.2.1 Naïve pluripotency

Naïve pluripotency represents the ground state of pluripotency and is considered the developmental equivalent of cells in the pre implantation ICM of a blastocyst (Boroviak and Nichols, 2017). Naïve PSCs exhibit several defining characteristics:

Morphology: They typically form tightly packed, dome–shaped colonies and are cultured on an inactivated feeder layer or in specific media containing inhibitors of the MAPK/ERK and GSK3 pathways (2i medium) (Nichols and Smith, 2009).

Transcriptional profile: Naïve PSCs express a core set of pluripotency markers, including *OCT4*, *SOX2*, and *NANOG*, at high levels. Critically, they also express unique markers such as *KLF4* and *TFCP2L1*, which are downregulated in primed cells (Marks et al., 2012). The global gene expression profile is characterized by a low level of spontaneous differentiation, and a global hypomethylation of DNA (Theunissen and Jaenisch, 2014).

Functional characteristics: Naïve PSCs possess a more potent differentiation capacity and can contribute to both the embryo proper and the germline when injected into a host blastocyst (Thomson et al., 1995). They also exhibit X–chromosome reactivation in female cells, which is a hallmark of the early embryonic state (Nichols and Smith, 2009).

2.2.2 Primed pluripotency

Primed pluripotency represents a later developmental stage, analogous to the epiblast of a post implantation embryo (Boroviak and Nichols, 2017). hESCs were originally isolated and propagated in the primed state (Thomson et al., 1998b). The key characteristics of primed PSCs are:

Morphology: Primed PSCs grow as flat, compact colonies with more defined cell–cell junctions. They are typically cultured on a feeder layer or an extracellular matrix like Matrigel (Boroviak and Nichols, 2017).

Transcriptional profile: They also express core pluripotency factors like *OCT4* and *NANOG* but lack expression of naïve–specific markers. A key distinguishing feature is the high expression of lineage–priming genes, indicating a predisposition towards differentiation. The overall epigenetic landscape is characterized by global hypermethylation of DNA compared to the naïve state (Theunissen and Jaenisch, 2014).

Functional characteristics: Primed PSCs have a more restricted differentiation potential and cannot contribute efficiently to chimeras when injected into blastocysts (Guo et al., 2017). In female primed cells, one of the two X–chromosomes is inactivated, mirroring the *in vivo* post implantation epiblast state (Nichols and Smith, 2009).

2.2.3 Naïve-like pluripotent stem cells

The concept of a naïve-like state has emerged primarily in the context of human stem cell research, where researchers have successfully reprogrammed or converted human primed ESCs to a naïve state using specific culture conditions (Theunissen et al., 2014). These "naïve-like" cells often share many of the morphological, transcriptional, and functional characteristics of true naïve cells (i.e., mouse ESCs), but their full equivalence to the pre implantation human ICM remains an active area of investigation (Boroviak and Nichols, 2017).

The significance of the naïve-like state is immense, as it offers a more robust and efficient platform for genetic manipulation and disease modeling (Baker et al., 2014). For example, it has been shown that converting primed hESCs to the naïve-like state increases their efficiency for genome editing via CRISPR/Cas9 (Theunissen and Jaenisch, 2014). However, the long–term stability and true biological fidelity of these converted cells are still under scrutiny.

The understanding that pluripotency is not a single state but rather a dynamic spectrum has revolutionized stem cell research. The distinct characteristics of naïve, primed, and naïve-like PSCs provide a crucial framework for guiding research and therapeutic development. While primed cells are valuable for their ability to model post implantation development, naïve cells offer a more flexible and potent tool for fundamental research and advanced genetic engineering. The ability to switch between these states provides an unprecedented level of control over stem cell behavior, paving the way for more effective and targeted applications in regenerative medicine.

Differences between naïve and primed pluripotency have been extensively studied, including culture conditions, transcriptomic, epigenetic and metabolic profiles as outlined in the recent review by Weinberger et al. (2016) (Table 2.2). The naïve state represents the cellular state of the pre–implantation mouse blastocyst epiblast, while the primed state is representative of the post implantation epiblast cells. These two cell types exhibit clearly distinct developmental potential, as evidenced by the fact that naïve cells are able to contribute to blastocyst chimeras, while primed cells cannot. The epigenetic, transcriptomic, metabolic and proteomic differences should lead to a better understanding of the fundamental properties of these two states of pluripotency (Takahashi et al., 2018). Nowadays, many techniques allow deep characterization of pluripotent stem cell states such as immunocytochemistry, RNA–seq, qPCR as well as FTIR microspectroscopy.

2.3 Rhesus macaques (*Macaca mulatta*)

Rhesus macaques are a critical model in biomedical research, and their embryonic stem cells serve as a vital tool for understanding human pluripotency and for developing cell–based therapies. The close genetic and physiological similarities

between rhesus macaques and humans make them an ideal translational model for studying human diseases that cannot accurately replicated in rodent models, such as HIV (Human immunodeficiency virus) , Alzheimer's, and Parkinson's disease. The establishment and characterization of rhESCs from this species have paved the way for the derivation of hESCs and have provided a platform for preclinical studies in regenerative medicine.

Table 2.2 Characteristic features of naïve and primed pluripotent stem cell states (Kumari, 2016)

Property	Naïve state	Primed state
1. Colony morphology	Compact dome–shaped	Flattened
2. Expressed genes	High expression of *OCT4, NANOG, SOX2, DPPA3, TFCP2L1, ZFP42, KLF2, KLF4, KLF5, ESRRB, FGF4, CDH1*	*OCT4, SOX2, OTX2, DNMT3B, FGF5, POU3F1, MEIS1, SOX11, GDF3*
3. Growth factor dependence	LIF	Activin, FGF2
4. Single–cell mortality	Low	High
5. XCI status in female cells	XaXa	XaXi
6. H3K27Me3 over developmental regulators	Low	High

2.3.1 Rhesus macaques (*Macaca mulatta*) as a biomedical model

The rhesus macaque (*Macaca mulatta*) is one of the most widely used non–human primates in biomedical research due to its close phylogenetic relationship to humans (Thomson et al., 1995). With a genome that is approximately 93% homologous to the human genome, the rhesus macaque serves as a highly relevant translational model for studying human physiology and disease progression (Szymanski et al., 2018). The species' use has been instrumental in numerous medical breakthroughs, including the development of the polio vaccine and the discovery of the Rh factor in blood typing (Reis et al., 2020).

The value of the rhesus macaque model extends across a wide range of research areas, including infectious diseases, such as HIV/AIDS (acquired immunodeficiency syndrome), where studies in rhesus macaques infected with simian immunodeficiency virus (SIV) have provided invaluable insights into viral pathogenesis

and have guided the development of antiretroviral therapies and vaccines (Wong et al., 2021). The longevity and natural aging processes of rhesus macaques also make them an excellent model for studying age–related diseases like Alzheimer's, as they naturally accumulate the misfolded amyloid–beta proteins observed in human patients (Baker et al., 2014).

2.3.2 Rhesus macaque embryonic stem cells (rhESCs)

rhESCs are pluripotent stem cells isolated from the ICM of rhesus monkey blastocysts. Their derivation by Thomson et al. 1995) was a landmark achievement that preceded the isolation of hESCs in 1998 (Thomson et al., 1995; Thomson et al., 1998a). The establishment of rhESCs provided the foundational knowledge and culture protocols that were essential for the subsequent work with human cells.

rhESCs share key characteristics with their human counterparts, including the ability for self–renewal and pluripotency (Thomson and Marshall, 1998a). They can proliferate indefinitely in an undifferentiated state *in vitro* and maintain the potential to differentiate into cells of all three embryonic germ layers: ectoderm, mesoderm, and endoderm (Thomson and Marshall, 1998a). This pluripotency is confirmed by the expression of a common set of pluripotency markers, such as the transcription factors *OCT4*, *SOX2*, and *NANOG*, and cell surface markers like SSEA–4 and TRA–1–60 (Adewumi et al., 2007).

The similarities between rhESCs and hESCs make the rhesus macaque a crucial model for preclinical studies of cell–based therapies. It allows researchers to evaluate the safety, efficacy, and potential for immune rejection of stem cell derivatives in a physiological context that closely mimics the human body (Miller et al., 2014). This is particularly important for therapies aimed at treating neurodegenerative diseases, as rhesus macaques naturally develop many of the same neurological conditions as humans (Chan and Kazarian, 2013).

2.3.3 Significance and future directions

The study of rhesus macaques and their embryonic stem cells continue to be of paramount importance in advancing our understanding of human biology and disease. The use of this model has not only facilitated the development of stem cell research but has also provided a bridge for translating laboratory discoveries into clinical applications. The ongoing research into rhESCs, including their genetic stability, differentiation potential, and immune compatibility, is essential for ensuring the safety and success of future regenerative medicine therapies.

2.4 Characterization of rhESCs by standard methods

rhESCs are pluripotent stem cells derived from the inner cell mass of blastocyst stage rhesus monkey embryos (Thomson et al., 1995). As a non–human primate model, the rhesus macaque is phylogenetically close to humans, and its stem cells share many biological similarities with hESCs. This makes rhESCs an invaluable model for studying fundamental pluripotency mechanisms, and for preclinical safety and efficacy studies of cell–based therapies before their application in humans (Thomson et al., 1995; Thomson and Marshall, 1998a). To ensure the proper identity, quality, and functionality of these cell lines, a rigorous and standardized set of characterization methods is essential.

The routine characterization of rhESCs involves a combination of morphological, molecular, and functional assays to confirm their key properties: pluripotency, self–renewal, and genetic stability.

2.4.1 Morphological and self–renewal assessment

The most basic step in rhESC characterization is the assessment of their morphology and ability to self–renew in an undifferentiated state (Thomson et al., 1995). Undifferentiated rhESCs typically grow as compact, three–dimensional colonies with sharp borders. The cells within these colonies are small, have a high nuclear–to–cytoplasmic ratio, and possess large, prominent nucleoli (Figure 4.1). This morphology is a key indicator of their undifferentiated state (Suemori et al., 2006). The ability of the cells to maintain this morphology and proliferate for an extended period in culture, without spontaneous differentiation, is evidence of their self–renewal capacity (Thomson and Marshall, 1998a).

2.4.2 Pluripotency marker expression

Pluripotency is the defining characteristic of ESCs, and it is confirmed by the expression of specific molecular markers. These markers can be detected at both the protein and gene levels.

Protein expression: Key pluripotency associated proteins include OCT4, SOX2, and NANOG (Thomson et al., 1995; Thomson and Marshall, 1998c). These transcription factors form a core regulatory network that maintains the pluripotent state. Their presence is commonly verified using immunocytochemistry or flow cytometry (Thomson et al., 1995). Other surface markers, such as SSEA–4 and TRA–1–60, which are found on the cell surface of pluripotent primate cells, are also routinely used to confirm pluripotency (Thomson et al., 1995; Adewumi et al., 2007).

Gene expression: The expression of pluripotency related genes, including *POU5F1* (encoding *OCT4*), *SOX2*, and *NANOG*, is typically measured using reverse

transcription polymerase chain reaction (RT–PCR) or quantitative real time PCR (qRT–PCR). The high expression levels of these genes in undifferentiated rhESCs, followed by their downregulation upon differentiation, is a standard metric for characterization (Suemori et al., 2006).

2.4.3 Pluripotency functional assays

Molecular marker expression is necessary but not sufficient to prove pluripotency. Functional assays are required to demonstrate the cells' ability to differentiate into all three germ layers: ectoderm, mesoderm, and endoderm.

In vitro differentiation: This involves inducing the cells to differentiate into derivatives of all three germ layers in a controlled cell culture environment (Thomson and Marshall, 1998). The formation of different cell types, such as neurons (ectoderm), cardiomyocytes (mesoderm), and gut–like cells (endoderm), can confirmed by analyzing the expression of specific lineage markers using immunocytochemistry or RT–PCR (Thomson et al., 1995).

In vivo teratoma formation: The gold standard for confirming pluripotency is the teratoma formation assay (Adewumi et al., 2007). This involves injecting undifferentiated rhESCs into an immunocompromised mouse. The formation of a benign tumor called a teratoma, which contains differentiated cell types from all three embryonic germ layers, is definitive proof of pluripotency. Histological analysis of the resulting tumor is used to identify the presence of these diverse cell types (Thomson et al., 1995; Thomson and Marshall, 1998).

2.4.4 Genetic stability assessment

Maintaining a stable and normal karyotype is critical for the safe and reliable use of rhESCs (Thomson et al., 1995). Prolonged culture can lead to the acquisition of chromosomal abnormalities, which can compromise the cells' properties or lead to oncogenic potential (Adewumi et al., 2007). Karyotyping is performed to visualize the chromosomes and confirm that the cells have a normal diploid chromosomal complement (2n=42 for rhesus macaque) (Thomson and Marshall, 1998c). This analysis is an essential part of quality control for rhESC lines.

The comprehensive characterization of rhESCs using a combination of morphological, molecular, and functional assays is a critical step in their establishment and application. The standard methods, including the assessment of morphology, expression of pluripotency markers (e.g., OCT4, SSEA–4), functional pluripotency assays (*in vitro* differentiation and teratoma formation), and genetic stability analysis, provide a robust framework for validating the identity and quality of these valuable cell lines.

These methods ensure that rhESCs serve as a reliable and reproducible model for biomedical research and as a crucial stepping stone for human clinical trials.

2.5 Immunocytochemistry: Principles and limitations

2.5.1 Principles of immunocytochemistry

Immunocytochemistry (ICC) is a widely used laboratory technique that harnesses the specific binding of antibodies to visualize the location and distribution of a target protein or antigen within a cell (Coons, 1956). The fundamental principle of ICC is the highly specific interaction between an antibody and its corresponding antigen. The term "immunocytochemistry" is generally used for the analysis of isolated cells or cell cultures, while a similar technique applied to tissue sections is called immunohistochemistry (Haines et al., 2007).

The process of ICC involves several key steps. First, the cells are fixed to a solid support (e.g., a glass slide) to preserve their morphological integrity and immobilize the proteins (Buchwalow et al., 2011). Next, the cell membrane may permeabilized to allow antibodies to access intracellular antigens. The cells are then incubated with a primary antibody that is raised against the target antigen of interest. This antibody binds directly to the protein being studied (Miller et al., 2014).

To visualize this primary antibody antigen complex, a detection system is employed. There are two main approaches:

2.5.1.1 Direct ICC

A primary antibody is directly conjugated to a reporter molecule, such as a fluorophore (e.g., FITC, Cy3) or an enzyme (e.g., horseradish peroxidase, HRP) (Depciuch et al., 2018). While this method is simple and fast, it often suffers from lower signal amplification and is less sensitive than indirect methods.

2.5.1.2 Indirect ICC

This more common and sensitive method uses an unconjugated primary antibody and a secondary antibody that is conjugated to a reporter molecule. The secondary antibody is raised against the species of the primary antibody (e.g., a goat anti rabbit secondary antibody for a primary antibody raised in a rabbit). Multiple secondary antibodies can bind to a single primary antibody, leading to a significant amplification of the signal (Buchwalow et al., 2011). This method is highly sensitive and allows for the use of a wide range of commercially available labeled secondary antibodies.

Following the labeling steps, the cells are typically counterstained to visualize cellular structures, such as the nucleus, and then imaged using a microscope (Ramos-Vara and Miller, 2014).

2.5.2 Limitations of immunocytochemistry

Despite being a powerful and widely used technique, ICC has several inherent limitations that must considered for accurate and reliable results.

2.5.2.1 Antibody specificity and cross reactivity

The reliability of ICC is critically dependent on the specificity of the antibodies used. An antibody might bind to a protein other than the intended target (off target binding), leading to false positive results or a misinterpretation of the localization (Haines et al., 2007). Cross reactivity with related proteins is a common issue, particularly with polyclonal antibodies. Therefore, it is essential to validate antibodies rigorously using controls, such as Western blotting, to confirm that they recognize the correct protein (Buchwalow et al., 2011).

2.5.2.2 Fixation and permeabilization artifacts

The fixation and permeabilization steps, while necessary to preserve cellular morphology and allow antibody access, can introduce artifacts. Fixatives like paraformaldehyde can cross link proteins, potentially masking the epitope (the specific site on the antigen that the antibody recognizes), thereby leading to a weak or absent signal (Ramos-Vara and Miller, 2014). Similarly, detergents used for permeabilization can alter the native structure of the cell, potentially affect protein localization or cause them to washed away (Coons, 1956). The choice of fixative and permeabilization agent must carefully optimized for each specific antigen to minimize these issues (Baker et al., 2014).

2.5.2.3 Quantification challenges

While ICC provides excellent qualitative information on protein localization, its use for precise quantitative analysis is challenging. The intensity of the fluorescent or chromogenic signal is influenced by numerous factors, including antibody concentration, incubation times, and the efficiency of the reporter molecule (Coons, 1956). Direct comparison of signal intensity between different experiments is often unreliable. Therefore, ICC is generally not suitable for determining the absolute quantity of a protein but rather for assessing its relative abundance or cellular distribution (Miller et al., 2014; Depciuch et al., 2018). More sophisticated techniques, such as flow cytometry or western blotting, are often required for accurate quantification.

2.6 RNA sequencing: Principles and limitations

2.6.1 Principles of RNA sequencing

RNA sequencing (RNA–seq) is a high throughput technology that uses next generation sequencing (NGS) to analyze the expression, structure, and function of an organism's transcriptome (Wang et al., 2009). The fundamental principle of RNA–seq is to convert an RNA population, which is highly unstable, into a more stable complementary DNA (cDNA) library. This library is then sequenced to generate millions of short reads that represent the original RNA molecules (Stark et al., 2019). The digital nature of this data provides a more precise and quantitative measure of gene expression compared to traditional methods like microarrays.

The process of RNA–seq generally follows a standardized workflow:

RNA isolation: Total RNA, including mRNA, ribosomal RNA (rRNA), and non–coding RNA, is extracted from the sample. The quality and integrity of the RNA are critical for accurate results (Reis et al., 2020).

Library preparation: This is a crucial step that prepares the RNA for sequencing. It typically involves mRNA enrichment to remove the highly abundant rRNA, which would otherwise dominate the sequencing data. The enriched mRNA is then fragmented, and a reverse transcriptase enzyme is used to synthesize a stable cDNA copy (Mortazavi et al., 2008). Sequencing adapters are then ligated to the ends of the cDNA fragments.

Sequencing: The prepared cDNA library is sequenced using a high throughput platform, such as an Illumina sequencer (Miller et al., 2014). This generates millions of short sequencings reads from the ends of the cDNA fragments.

Data analysis: The short reads are computationally mapped to a reference genome or transcriptome. The number of reads that map to a specific gene or transcript is then counted to quantify its expression level (Wang et al., 2009). This provides a comprehensive gene expression profile of the sample, which can used to identify differentially expressed genes, novel transcripts, gene fusions, and single nucleotide variants (SNVs).

2.6.2 Limitations of RNA sequencing

Despite its widespread use and powerful capabilities, RNA–seq has several limitations that must addressed for reliable data interpretation.

2.6.2.1. Cost and computational demands

RNA–seq can expensive, especially for projects with a large number of samples or a requirement for deep sequencing to detect low abundance transcripts (Krimm and Bandekar, 1986). In addition, the vast amount of data generated by a single

sequencing run requires significant computational resources for storage, processing, and analysis. This necessitates access to high performance computing clusters and specialized bioinformatics expertise, which can a barrier for some research groups (Szymanski et al., 2018).

2.6.2.2 Challenges in quantification and normalization

Quantifying gene expression from RNA–seq data is not straightforward. The number of reads obtained for a gene is influenced not only by its true expression level but also by factors such as gene length and sequencing depth (Wong et al., 2021). To enable accurate comparisons between samples, the data must normalize to account for these technical variations (Depciuch et al., 2018). The choice of normalization method (e.g., TPM, FPKM, or DESeq2) can significantly impact the final results and conclusions (Baker et al., 2014).

2.6.2.3 Sample and technical variability

The quality of the starting RNA material is critical. Degraded RNA can lead to biases in the data, particularly towards the 3' end of transcripts (Miller et al., 2014). Additionally, technical variations introduced during library preparation, such as biases in reverse transcription or PCR amplification, can affect the final gene expression counts. While these biases can partially mitigate through rigorous quality control and careful experimental design, they remain a source of variability (Stark et al., 2019).

2.6.2.4 Difficulty in detecting low abundance transcripts

Sequencing depth directly impacts the ability to detect and quantify transcripts. For genes expressed at very low levels, a shallow sequencing depth may not generate enough reads to accurately represent their expression, leading to false negatives (Depciuch et al., 2018). Conversely, very high sequencing depth can cost prohibitive. Optimizing the balance between sequencing depth and cost is a common challenge in experimental design.

2.7 Characterization of ESCs by FTIR microspectroscopy

2.7.1 Principles, mechanisms and limitations of FTIR microspectroscopy

FTIR microspectroscopy is a powerful analytical technique that combines the principles of FTIR spectroscopy with optical microscopy, allowing for chemical analysis with high spatial resolution (Baker et al., 2014). The fundamental principle relies on the interaction between infrared radiation and matter. Specifically, molecules possess unique vibrational and rotational modes, such as stretching (changes in bond length) and bending (changes in bond angle) (Stuart, 2004). When the frequency of the incident IR radiation matches the natural vibrational frequency of a specific chemical

bond, the bond absorbs the energy, transitioning to a higher vibrational state. This absorption is unique to each type of bond and functional group, creating a distinctive molecular "fingerprint" that can used to identify and characterize the sample's chemical composition (Krimm and Bandekar, 1986).

The mechanism of generating the spectrum involves a Michelson interferometer. An incoming broadband IR beam is split into two paths by a beamsplitter, one path directed to a fixed mirror and the other to a movable mirror (Griffiths, 1983). The two beams are recombined, creating an interference pattern called an interferogram, which is a signal of the total IR intensity as a function of the optical path difference (OPD) between the two beams. The raw interferogram is not directly interpretable; therefore, a mathematical operation known as the Fourier transform is applied to convert the time domain interferogram into a frequency domain spectrum (absorbance versus wavenumber) (Stuart, 2004). The integration of this system with a microscope allows for the IR beam to focused onto a small area of the sample, enabling the collection of spectral data point by point, thereby generating a chemical map or image that links chemical information to spatial location (Wong et al., 2021).

2.7.2 Limitations of FTIR microspectroscopy

Despite its significant advantages, FTIR microspectroscopy has several inherent limitations that must carefully considered during experimental design and data interpretation.

2.7.2.1 Spatial resolution

One of the primary limitations is the diffraction limit of light, which restricts the achievable spatial resolution. For mid IR radiation, which is typically used, the wavelength is in the range of 2.5 to 25 μm, meaning the theoretical spatial resolution is limited to several micrometers (Reis et al., 2020). This can a major constraint when analyzing subcellular components or nanoparticles, which are often smaller than the diffraction limit. While techniques like attenuated total reflection (ATR) can improve resolution by using a crystal with a high refractive index, they require direct contact with the sample and are limited by the physical size of the ATR crystal (Chan and Kazarian, 2013).

2.7.2.2 Sample preparation and handling

Proper sample preparation is crucial but can challenging, particularly for biological samples. Samples must very thin (<10 μm) and uniform to avoid strong IR absorption, which can saturate the signal and obscure important spectral features (Miller et al., 2014). The presence of water is a particularly significant issue, as it has a strong and broad IR absorption band that can mask the signals from other key

biomolecules, such as proteins and lipids (Baker et al., 2014). Consequently, biological samples often need to dried, which can potentially alter their native biochemical and structural state.

2.7.2.3. Data processing and analysis

FTIR microspectroscopy generates large, complex datasets that necessitate sophisticated data processing and analysis. The raw spectra often contain noise, baseline artifacts, and overlapping peaks, requiring extensive pre-processing steps, including baseline correction and spectral normalization (Depciuch et al., 2018). Furthermore, the interpretation of the spectral data often requires the use of advanced chemometric methods such as PCA, hierarchical cluster analysis (HCA), and partial least squares (PLS) to extract meaningful information and identify subtle differences between samples (Depciuch et al., 2018). These complex computational demands require specialized software and expertise, which can a barrier for some researchers.

2.8 Diagnostic applications and performance of FTIR microspectroscopy

2.8.1 Diagnostic applications

FTIR microspectroscopy has emerged as a powerful label free diagnostic tool with diverse applications in biotechnology and medicine. Its ability to provide a "molecular fingerprint" of a sample by identifying the presence and relative abundance of biomolecules (e.g., proteins, lipids, carbohydrates, and nucleic acids) makes it suitable for various diagnostic purposes (Baker et al., 2014; Miller et al., 2014).

2.8.1.1 Cancer diagnosis and histopathology

One of the most significant applications is in the diagnosis of cancer. FTIR microspectroscopy can differentiate between healthy and cancerous tissues based on subtle biochemical changes that occur during carcinogenesis, often before morphological changes are evident (Depciuch et al., 2018). For instance, studies have shown that cancerous cells exhibit altered lipid and protein content, changes in nucleic acid conformation, and increased glycolysis, all of which are detectable through their unique IR spectral signatures (Lassiter et al., 2019; Reis et al., 2020). This technology is being explored for the rapid analysis of tissue biopsies, offering an alternative or complement to traditional histopathology, which relies on subjective visual interpretation by pathologists. It has been successfully applied to the diagnosis of various cancers, including breast (Lassiter et al., 2019), colon (Szymanski et al., 2018), and brain tumors (Miller et al., 2014).

2.8.1.2 Microbiology and infectious disease

FTIR microspectroscopy is also a valuable tool in microbiology for the rapid identification and classification of microorganisms, including bacteria, fungi, and viruses. Each microbial species possesses a unique macromolecular composition that yields a distinct IR spectrum (Mendelsohn and Moore, 2012). This allows for the differentiation of closely related species and strains, which is crucial for epidemiological studies and the rapid diagnosis of infectious diseases (Wong et al., 2021). The technique is particularly useful for identifying antibiotic resistant strains, offering a quick method to guide clinical treatment (Depciuch et al., 2018).

2.8.1.3 Stem cell and regenerative medicine

In the field of regenerative medicine, FTIR microspectroscopy provides a non–destructive method to monitor the differentiation and maturation of stem cells. The biochemical changes that accompany stem cell differentiation—such as changes in lipid profiles and protein secondary structures—can tracked in real time. This provides insight into the differentiation pathways and quality control for cell–based therapies (Kazarian and Chan, 2017).

2.8.2 Performance

The performance of FTIR microspectroscopy in a diagnostic context is defined by its sensitivity, specificity, and speed.

2.8.2.1. Sensitivity and specificity

The sensitivity and specificity of FTIR microspectroscopy in distinguishing between different sample types, such as healthy and diseased tissues, are generally very high, often exceeding 90% in well controlled studies (Baker et al., 2014). The high sensitivity is attributed to the fact that the technique measures the collective signal from thousands of molecules within a single measurement, making it highly sensitive to even subtle biochemical changes (Reis et al., 2020). The specificity is derived from the unique spectral fingerprint of different biomolecules and the use of sophisticated chemometric algorithms, such as PCA and LDA, to classify spectra accurately (Depciuch et al., 2018).

2.8.2.2 Speed and throughput

The speed of FTIR microspectroscopy is a significant performance advantage, especially when using modern instrumentation. Unlike many traditional diagnostic methods that require extensive staining and processing time, FTIR provides results in minutes. Advances in technology, such as the use of FPA detectors, have dramatically increased data acquisition speed by allowing for the simultaneous collection of spectra from a large area, making it possible to image entire tissue

sections in a single pass (Miller et al., 2014). This high throughput makes FTIR microspectroscopy a promising tool for high volume clinical and research applications.

2.9 Detectors in fourier transform infrared (FTIR) microspectroscopy

2.9.1 Principles and types of IR detectors

The performance of FTIR microspectroscopy is critically dependent on the type and sensitivity of the detector used to measure the IR radiation transmitted through or reflected from a sample (Griffiths, 1983). IR detectors operate by converting the incident IR radiation into an electrical signal. They are broadly classified into two main categories: thermal detectors and photon detectors (Stuart, 2004).

2.9.1.1 Thermal detectors

These detectors, such as the deuterated triglycine sulfate (DTGS) detector, measure the temperature change caused by the absorption of IR radiation. They are relatively inexpensive and do not require cryogenic cooling, which makes them easy to use (Depciuch et al., 2018). However, their response time is slower and their sensitivity is generally lower than that of photon detectors. For this reason, DTGS detectors are often used in routine or low cost FTIR systems where high speed and sensitivity are not the primary requirements (Miller et al., 2014).

2.9.1.2 Photon detectors

These detectors, such as the MCT detector, measure the change in the electrical properties (e.g., resistance) of a semiconductor material when it absorbs IR photons. MCT detectors are highly sensitive and have a very fast response time, making them ideal for high–speed data acquisition and for analyzing samples with low IR signal intensity (Baker et al., 2014). To achieve maximum performance, MCT detectors typically require cryogenic cooling with liquid nitrogen to minimize thermal noise (Wong et al., 2021). The high sensitivity of MCT detectors is particularly valuable for FTIR microspectroscopy, where the amount of IR radiation reaching the detector from a small sample area is often limited (Reis et al., 2020).

2.9.2 Focal plane array (FPA) detectors

A significant advancement in FTIR microspectroscopy has been the development of FPA detectors. Unlike single element detectors that acquire spectral data sequentially, FPA detectors consist of a two–dimensional array of individual detector elements (pixels) (Miller et al., 2014). This allows for the simultaneous acquisition of an entire IR image, where each pixel in the array collects a full IR spectrum from a corresponding point on the sample (Chan and Kazarian, 2013). This parallel data acquisition dramatically reduces the time required to generate a chemical

map or image of a sample, making it possible to analyze large areas of tissue or cell cultures in minutes, rather than hours (Kazarian and Chan, 2017). The high throughput and high sensitivity of FPA based systems have revolutionized the field, enabling new applications in diagnostic histopathology and high content screening (Lassiter et al., 2019). The use of FPA detectors allows for the rapid visualization of biochemical heterogeneities, which is crucial for distinguishing between different tissue types or identifying disease states.

2.9.3 Performance and limitations

The choice of detector has a direct impact on the performance of the FTIR microspectroscopy system. The signal to noise ratio (SNR), which is a critical measure of spectral quality, is significantly higher for MCT and FPA detectors compared to DTGS detectors, enabling the detection of subtle spectral features (Depciuch et al., 2018). While FPA detectors offer unparalleled speed and spatial information, they can more expensive and require specialized software for data processing due to the large file sizes generated (Wong et al., 2021). The cryogenic cooling required for MCT and FPA detectors also adds complexity and cost to the instrumentation. Despite these limitations, the superior sensitivity and speed of these detectors have made them the standard for advanced research and clinical applications of FTIR microspectroscopy (Reis et al., 2020).

2.10 Multivariate data analysis of FTIR microspectroscopy

2.10.1. The challenge of complex FTIR microspectroscopy data

FTIR microspectroscopy generates large and complex datasets, especially when using modern FPA detectors (Baker et al., 2014). A single chemical map can contain thousands of individual spectra, with each spectrum comprising hundreds or even thousands of data points (wavenumbers). This results in a highly dimensional dataset, where the spectral variations are often subtle and overlapping, making them difficult to interpret by visual inspection alone (Miller et al., 2014). To extract meaningful information from these complex datasets and identify subtle biochemical differences, multivariate data analysis (chemometrics) is essential (Reis et al., 2020). These statistical techniques are designed to handle and simplify multi–dimensional data, revealing underlying patterns and relationships.

2.10.2 Common multivariate data analysis techniques

Several chemometric techniques are routinely applied to FTIR microspectroscopy data. These can broadly be classified into unsupervised and supervised methods.

2.10.2.1. Unsupervised methods: exploring data without prior knowledge

Unsupervised methods are used to explore the data and identify natural groupings or patterns without any prior knowledge or labels.

Principal component analysis (PCA): PCA is the most widely used unsupervised method for analyzing FTIR data (Depciuch et al., 2018). It works by reducing the dimensionality of the dataset by identifying the principal components (PCs), which are new variables that capture the maximum variance in the original data. The first few PCs often represent the most significant sources of variation in the sample, such as differences between healthy and diseased tissues or different cell types (Wong et al., 2021). The results are often visualized in a scores plot to show how samples are clustered and loadings plot to identify which wavenumbers (and thus which chemical bonds) are responsible for the observed clustering (Kazarian and Chan, 2017).

Hierarchical cluster analysis (HCA): HCA is a clustering method that groups similar spectra together based on a measure of spectral distance (Miller et al., 2014). The output is a dendrogram that visually represents the relationships between spectra, with similar spectra located close to each other. HCA is often used in combination with PCA to provide an objective way of classifying different regions within a sample, such as identifying a tumor margin in a tissue section (Chan and Kazarian, 2013).

2.10.2.2. Supervised methods: classification and prediction

Supervised methods are used when there is prior knowledge about the samples (e.g., control vs. treatment, healthy vs. diseased). These methods use a training set of labeled data to build a model that can predict the class of a new, unknown sample.

Linear discriminant analysis (LDA): LDA is a classification technique often used after PCA to enhance the separation between predefined groups (Szymanski et al., 2018). It finds a linear combination of features that best separates two or more classes of objects. When combined with PCA (known as PCA–LDA), it becomes a powerful tool for building robust diagnostic models for applications such as cancer detection (Lassiter et al., 2019).

Partial least squares (PLS): PLS is a regression technique used to model the relationship between the spectral data (X–variables) and an external variable of interest (Y–variable), such as a concentration or a disease score (Depciuch et al., 2018). This method is particularly useful for quantitative analysis, for example, predicting the concentration of a specific metabolite or the degree of differentiation in a cell culture.

2.10.3 Performance and interpretation

The successful application of multivariate data analysis is crucial for the diagnostic and quantitative performance of FTIR microspectroscopy. These methods enable the differentiation of samples with high sensitivity and specificity, often exceeding 90% (Baker et al., 2014). However, proper application requires careful data pre–processing (e.g., baseline correction, normalization) to remove artifacts and a deep understanding of the biological and chemical context to correctly interpret the statistical output (Reis et al., 2020). The results from these analyses can reveal the subtle biochemical changes that are the hallmarks of disease or cellular processes, making chemometrics an indispensable component of modern FTIR microspectroscopy.

CHAPTER III
RESEARCH METHODOLOGY

3.1 Chemicals

All chemicals will purchase from the Sigma Aldrich Corporation (St. Louis, MO, USA). The cell culture media were purchased from Gibco (Paisley, United Kingdom). Plastic cell culture devices were obtained from SPL Life Sciences (Gyeonggi-do, Republic of Korea), unless otherwise indicated.

3.2 Preparation of feeder cells

Mouse embryonic fibroblasts (MEFs) were isolated from 13.5-day-old OF1 mouse embryos (Charles River, Lyon, France) following the protocol previously described (Afanassieff et al., 2014). MEFs were cultured in MEF medium consisting of Dulbecco's modified eagle medium (DMEM, Gibco, Paisley, UK) supplemented with 10% fetal bovine serum (FBS, Gibco), 1% Non-essential amino acids (NEAA, Gibco), 1% penicillin-streptomycin-glutamine (PSG, Gibco). Cells were incubated at 37 °C in a humidified atmosphere of 5% CO_2 in air. Cells were passaged with trypsin-EDTA (Gibco) and frozen at passage 2 (P2) in medium containing 10% dimethyl sulfoxide (DMSO, Sigma-Aldrich, St. Louis, MO, USA). Frozen-thawed MEFs were cultured to 90% confluence and mitotically inactivated by incubation in fresh MEF medium supplemented with 5 µg/mL mitomycin C (Sigma-Aldrich) at 37 °C for 3 h. Cells were then washed five times with Ca^{2+}/Mg^{2+} free PBS (Gibco), trypsinzed, incubated for 5 min at 37 °C, centrifuged at 400 × g for 5 min, and resuspended in MEF medium. Mitomycin-treated MEFs (2.5×10^5 cells) were plated in 35 mm culture dishes before seeding rhESCs.

3.3 Preparation of conditioned medium (CM)

MEFs (4×10^6) were plated in 100 mm dishes (SPL Life Sciences, Gyeonggi, Republic of Korea) coated with 0.1% gelatin. One day after plating, MEF medium was replaced with 25 ml N2B27 basal medium containing 48.7% DMEM/F12 (Gibco), 48.7% Neurobasal medium (Gibco), 1% B27 supplement (Gibco), 0.5% N2 supplement (N2, homemade), 0.02% β-mercaptoethanol (Gibco), 1% PSG, and 20 ng/ml FGF-basic human (bFGF, Sigma-Aldrich). After 24 h, the N2B27 conditioned medium (N2B27-CM) was collected and replaced with an equal volume of fresh complete N2B27 medium.

Conditioned medium was collected daily for three consecutive days, pooled, filtered, and stored at –20°C until use.

Table 3.1 Primary antibodies used for immunocytochemistry.

Marker types	Primary antibodies	Origin	Dilution
Core pluripotency markers	*OCT4* (Santa Cruz Biotechnologies sc–9081)	Rabbit	1:100
	NANOG (R&D system AF1997)	Goat	1:100
	SOX2 (R&D system AF2018)	Goat	1:100
Primed pluripotency marker	*OTX2* (R&D system AF1979)	Goat	1:200
Naïve pluripotency markers	*TFAP2C* (R&D system AF5059)	Goat	1:100
	KLF17 (Atlas Antibodies HPA024629)	Rabbit	1:100
	ALPPL2 (abcam ab96947)	Rabbit	1:100
	TFCP2L1 (abcam ab123354)	Rabbit	1:100

Table 3.2 Secondary antibodies used for immunocytochemistry.

Secondary antibodies	Origin	Dilution
Alexa Fluor 647 (Invitrogen A21447) (Far red)	Donkey anti–Goat	1:500
Alexa Fluor 555 (Invitrogen A21432) (Red)	Donkey anti–Goat	1:500
Alexa Fluor 647 (Invitrogen A31573) (Far red)	Donkey anti–Rabbit	1:500
Alexa Fluor 555 (Invitrogen A31572) (Red)	Donkey anti–Rabbit	1:500

3.4 Culture and expansion of rhESCs-FGF2/KOSR

rhESCs-FGF2/KOSR cells were cultured on feeders in 35 mm dishes with 2 ml ESC medium containing 80% KO–DMEM (Gibco), 20% knockout serum replacement (KOSR, Gibco), 1% NEAA, 1% GlutaMAX (Gibco), 0.1mM β–mercaptoethanol (Sigma–Aldrich), and 5 μg/ml bFGF (Sigma–Aldrich). Frozen–thawed rhESCs were cultured in medium supplemented with 10 μM ROCK inhibitor (Y–27632) (TOCRIS Bioscience, Bristol, UK) for the first 24 h. Cultures were maintained at 37°C, 5% CO_2 and 5% O_2, with daily medium changes. Cells were passaged every 3–4 days by mechanical dissociation.

3.5 Culture and expansion of rhESCs-ALGöX

rhESCs-ALGöX cells were cultured in 35 mm dishes pre–coated with 5 µg/ml laminin (LN521, STEMCELL Technologies, Lund, Sweden) for 1 h. Cells were maintained in 2 ml ALGöX medium (Pham et al., 2025). consisting of N2B27–CM supplemented with 10 ng/ml Activin A (Peprotech, Cranbury, NJ, USA), 1,000 U/ml LIF (homemade), 1.25 µM Gö6983 (TOCRIS Bioscience), and 2.5 µM XAV939 (Sigma–Aldrich). Cultures were maintained at 37°C, 5% CO_2 and 5% O_2, with daily medium changes until 80% confluence. Cells were passaged every 3–4 days using 1X TrypLE (Gibco) for single–cell dissociation, followed by addition of 10 µM Y–27632 for the first 24 h to promote cell survival.

3.6 Characterization of rhESC by immunocytochemistry

rhESCs-FGF2/KOSR cells and rhESCs-ALGöX cells were fixed with 4% paraform-aldehyde (PFA, Sigma–Aldrich) in PBS for 20 min at room temperature, washed twice with PBS (5 min each), and permeabilized with PBS containing 0.5% Triton X–100. Non–specific binding sites were blocked in PBS with 10% donkey serum for 1 h at room temperature. Cells were incubated overnight at 4°C with primary antibodies (Table 3.1). After three washes, cells were incubated with secondary antibodies (Table 3.2) for 1 h at room temperature. Nuclei were stained with DAPI (0.5 µg/ml), and cells were mounted with M1289 mounting medium (Sigma–Aldrich). Confocal imaging was performed using a DM 6000 CS SP5 microscope (Leica, Wetzlar, Germany) with a 45×/1.25 oil immersion objective (PL APO HCX, Leica) (Aksoy et al., 2021).

3.7 Characterization of rhESC by RNA sequencing

Total RNA was extracted from 4–5 x 10^6 cells using the RNeasy mini–kit (Qiagen, Hilden, Germany). Libraries were prepared from 200 ng of RNA using the NextFlex Rapid Directional mRNA–Seq kit (Bio–Scientific, Boston, MA, USA), and sequenced on a NextSeq500 platform (Illumina, San Diego, CA, USA) as single–end 75 bp reads. Demultiplexing was performed with bcl2fastq (Illumina), and adapters were trimmed using Cutadapt. Sequencing depth was ~30 million reads per sample. Reads were aligned to the reference genome using HISAT2, and gene counts were generated with HTSeq. RNA–seq data from this study are available under GEO accession number GSE146178 (Amzal et al., 2025).

3.8 Characterization of rhESCs by FPA–FTIR microspectroscopy

Cell pellets were washed three times with 0.9% NaCl, resuspended in 50 µl saline, and 5 µl aliquots were deposited onto IR transparent 2 mm thick barium fluoride (BaF_2) windows. Samples were air–dried and stored in a desiccator until analysis (Figure 3.1). Spectral data was acquired from an FPA–FTIR microscope (Hyperion 3000) with a FPA detector, connected to Tensor 27 FTIR spectrometer (Bruker Optics, Ettlingen, Germany) at the Synchrotron Light Research Institute (Public Organization), Nakhon Ratchasima, Thailand. The acquisition parameters were a 36× objective in transmission mode, 64 scans and all data was measured with 8 × 8 binning at a spectral resolution of 6 cm^{-1} (Heraud et al., 2007). The full area of 64 × 64 pixels equal 4,096 pixels. The field of view area using a 36× objective lens was 70.4 × 70.4 μm^2. The biochemical composition distribution was performed by OPUS 7.5 software (Bruker Optics, Ettlingen, Germany).

3.9 Multivariate data analysis of FTIR spectra

Spectral quality control was performed by visual inspection; spectra with weak absorbance (amide I band maximum absorbance <0.2), acquired from regions of the sample where there were no cells, or spectra with very high absorbance (amide I band maximum absorbance >0.8), acquired from regions where cells may have been clumped or overlaid was reject from the analysis. Spectra from sample groups based on differentiation of cell states and passage cells were analyze using PCA. Preprocessing of the data was conduct by first performing a second derivative by the Savitzky Golay method (13 smoothing points) and then normalized using EMSC using the spectral regions from 3000 to 2800 cm^{-1} and from 1800 to 800 cm^{-1} using The Unscrambler X 10.3 software (CAMO, Oslo, Norway). Score plots was use to visualize any clustering of the data, and the loading plot was used to determine which spectral regions contributed most to the variance in the data set, accounting for any clustering of spectra seen in scores plots. Integrate peak areas, was analyze by OPUS 7.5 software. Spectra were analyze using PLS–DA by The Unscrambler X 10.3 software. PLS–DA employed PCA models derived from calibration sets and was used to test the ability to discriminate different cell states using the independent validation set spectra. Calibration and validation of spectral data was employ using data sets that was randomly select comprising two–thirds and one–third of the spectra, respectively. The dataset can utilize to calculate the percentages of specificity and sensitivity. FTIR microspectroscopy combined with multivariate data analysis, in particular PCA, was apply to explain biochemical changes occurring during cellular differentiation.

3.10 Statistical analysis

All statistical analyses were performed using GraphPad Prism version 5 (GraphPad Software, San Diego, CA, USA). Data are presented as Standard Error of the Mean (SEM). Differences among groups were analyzed using one–way analysis of variance (ANOVA), followed by Tukey–Kramer's Honest Significant Difference (HSD) post hoc test for pairwise comparisons. A p–value of less than 0.05 was considered statistically significant. Different letters (a, b, c, d, e) above the bars indicate a statistically significant difference ($p<0.05$) between the groups. Identical letters indicate no significant difference. Graphs were generated using Sigma Plot version 15 (Grafiti LLC, Palo Alto, CA, USA).

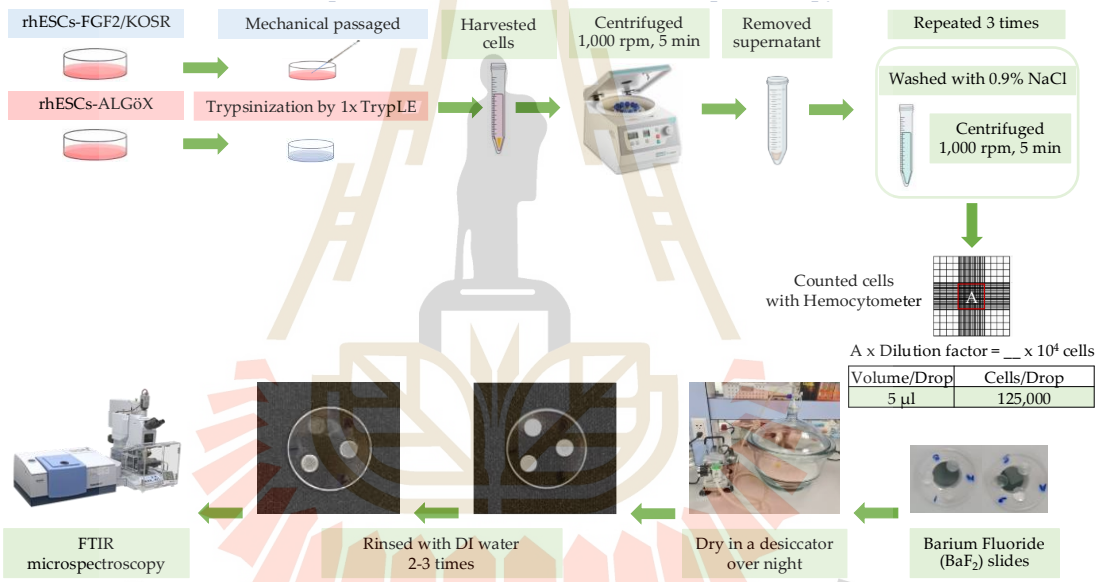

Figure 3.1 Preparation of rhESCs-FGF2/KOSR and rhESCs-ALGöX for characterization by FPA–FTIR Microspectroscopy.

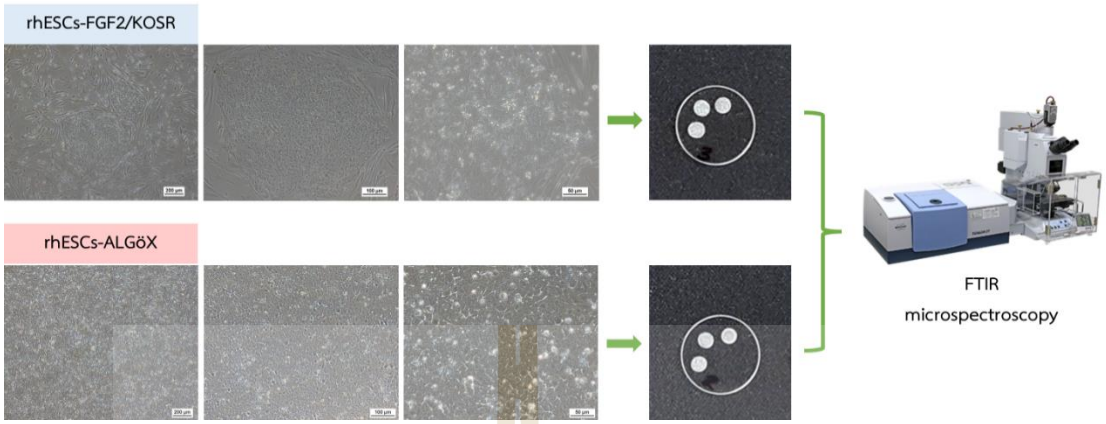

Figure 3.2 Morphology of rhESCs-FGF2/KOSR and rhESCs-ALGöX on culture dishes and morphology of cells on BaF2 slides.

CHAPTER IV
RESULTS AND DISCUSSION

4.1 Evaluation of rhESCs properties by immunohistochemistry

rhESCs were originally derived in FGF2/KOSR (Wianny et al., 2008). Morphological changes were observed when rhESCs-FGF2/KOSR were transferred into ALGöX medium. rhESCs formed more compact colonies (Figure 4.1).

Immunolabelling confirmed that both rhESC-FGF2/KOSR and rhESCs-ALGöX retained expression of core pluripotency markers OCT4, NANOG, and SOX2 (Figure 4.2), indicating stable pluripotent identity across conditions. In rhESC-FGF2/KOSR, the intensity of these markers peaked at passage 60 and declined progressively at later passages (65 and 70). In contrast, rhESC-ALGöX maintained consistent expression levels across passages (Figure 4.3).

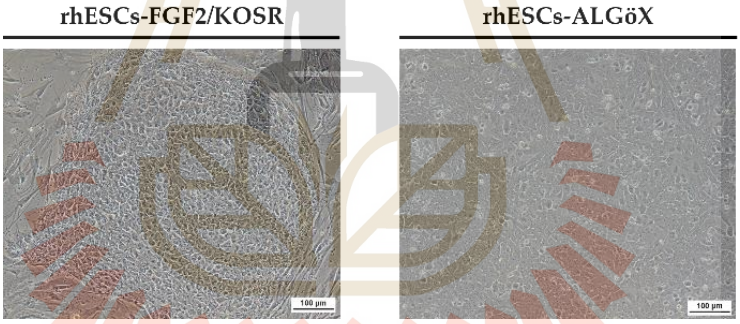

Figure 4.1 Phase contrast images of rhESCs cultured in FGF2/KOSR and ALGöX. Scale bars: 100 µm.

Notably, the primed–state marker OTX2 was detected exclusively in rhESC-FGF2/KOSR, while rhESC-ALGöX cells exhibited robust expression of KLF17, ALPPL2, TFCP2L1, and TFAP2C, consistent with a naïve pluripotency signature (Figure 4.2). Quantitative fluorescence analysis revealed statistically significant differences in marker expression over passages, with the highest levels of OTX2 in primed cells at passage 65 and peak expression of naïve markers in rhESC-ALGöX cells at passage 32 (Figure

4.3). These results indicate that rhESCs cultured in ALGöX medium maintain a naïve-like molecular profile and support stable naïve pluripotency over multiple passages.

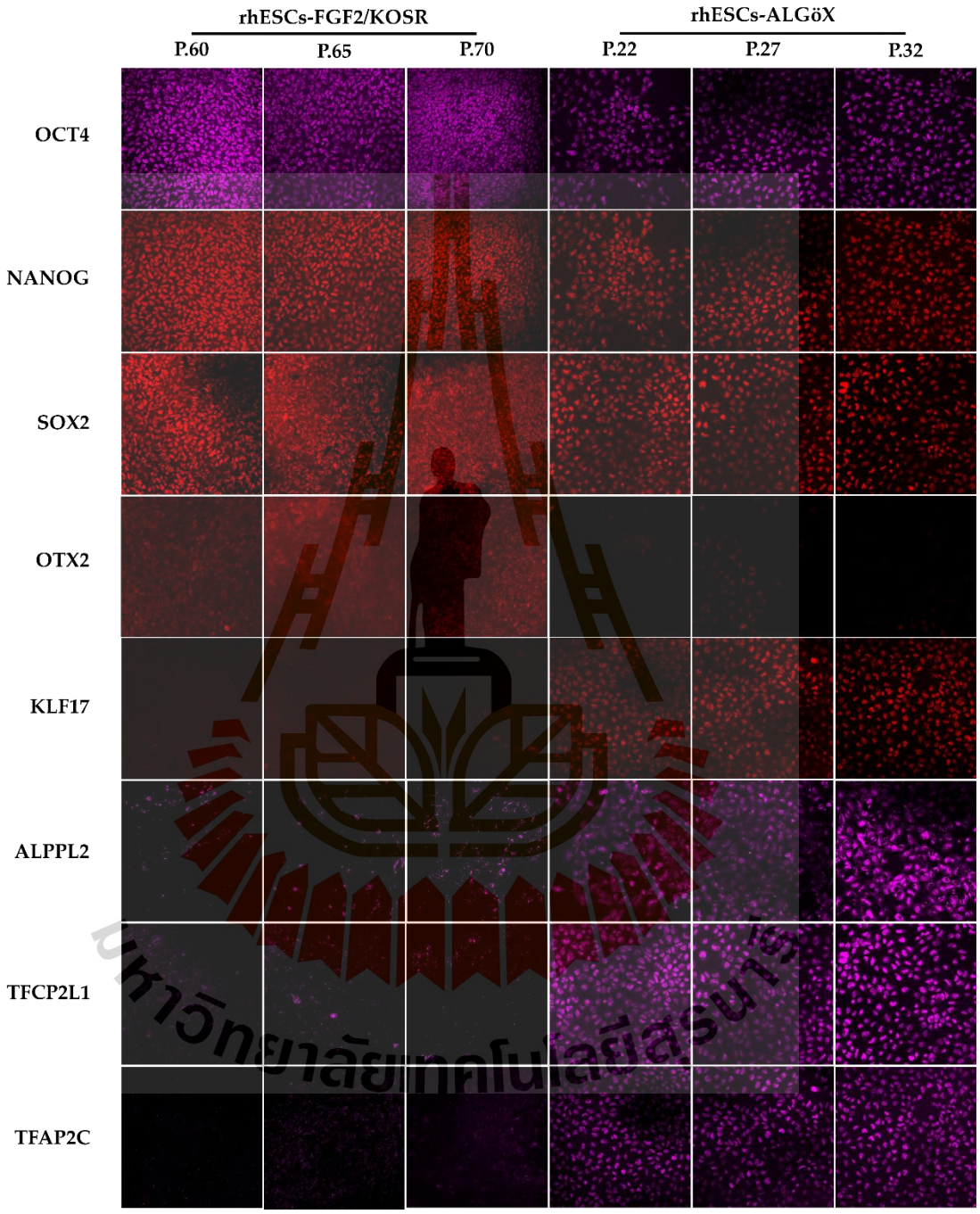

Figure 4.2 Immunolabelling of rhESCs cultured in FGF2/KOSR and ALGöX. Core pluripotency markers: OCT4, NANOG and SOX2. State–specific markers: OTX2 (primed) and KLF17, ALPPL2, TFCP2L1, and TFAP2C (naïve-like).

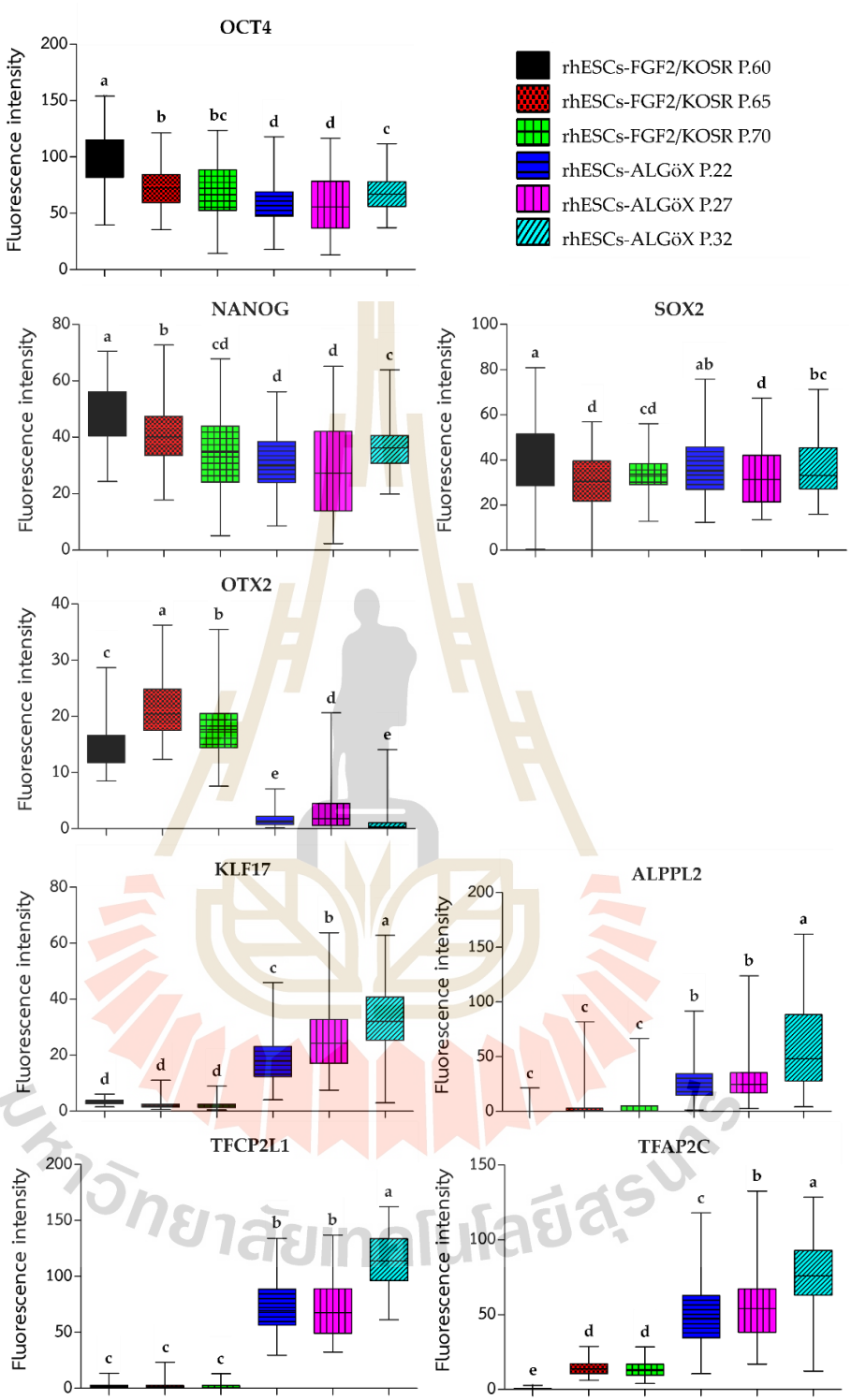

Figure 4.3 Quantification of fluorescence intensity in rhESCs cultured in FGF2/KOSR or ALGöX following immunolabelling. Different letters (a, b, c, d, e) above the bars indicate a statistically significant difference ($p<0.05$) between the groups.

4.2 Evaluation of rhESCs properties by RNA sequencing

RNA–seq confirmed the transcriptional divergence between rhESC-FGF2/KOSR and rhESCs-ALGöX. Hierarchical clustering of Z–score–normalized expression profiles segregated all samples into two distinct clusters corresponding to their respective culture conditions (Figure 4.4). rhESCs-FGF2/KOSR showed elevated expression of primed–state genes including *NODAL*, *OTX2*, *ETV4*, *BMP4*, *FST*, and *SOX3*, whereas rhESC-ALGöX exhibited high expression of naïve–associated genes including *KLF2*, *DPPA2*, *DPPA3*, *ZFP42*, *PRDM14*, and *TFCP2L1*. These transcriptomic data validate the immunolabelling data and strongly suggest that ALGöX conditions reprogram primed rhESCs to a molecularly distinct, naïve-like pluripotent state.

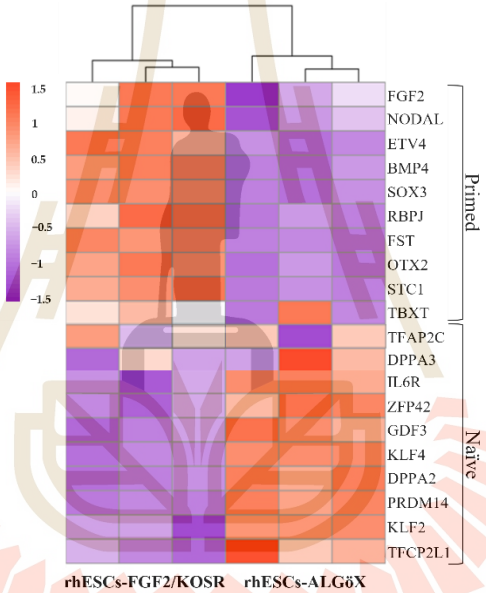

Figure 4.4 Heatmap of gene expression profiles in rhESCs cultured in FGF2/KOSR or ALGöX, based on RNA sequencing data. Genes associated with primed and naïve pluripotency states are shown.

4.3 Evaluation of rhESCs properties by FTIR microspectroscopy

4.3.1 Biochemical differences between cell states

We used a FPA–based FTIR imaging system to analyze rhESC-FGF2/KOSR and rhESC-ALGöX cells at the high pixel resolution (9.6 × 9.6 μm). Three biological replicates per condition were analyzed: passages 60, 65, and 70 for primed rhESC-FGF2/KOSR, and passages 22, 27, and 32 for naïve-like rhESC-ALGöX. The amide I (1700–1600 cm^{-1}) and amide II (1600–1500 cm^{-1}) bands provide insight into protein secondary structures (Figure 4.5A). The amide I band, arising primarily from C=O stretching, resolves into

features attributed to **α**–helices (~1654 cm^{-1}), **β**–sheets (~1635 cm^{-1}), and **β**–turns (~1685 cm^{-1}) (Cao et al., 2013b). The amide II band originates from N–H bending and C–N stretching and typically peaks around 1546 cm^{-1}. To resolve overlapping peaks, we applied second derivative spectroscopy, which enhanced the identification of individual vibrational bands. The second derivative spectra of primed and naïve-like cells are shown in Figure 4.5B.

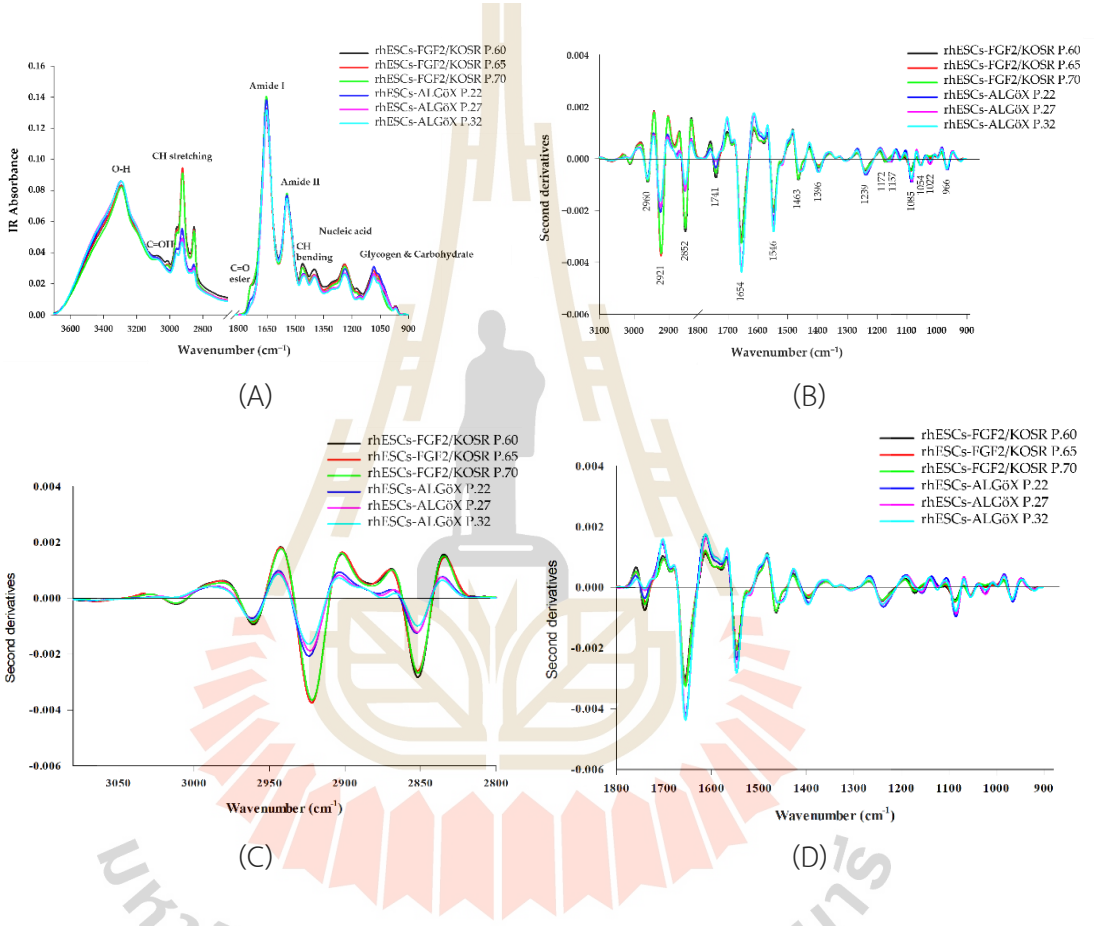

Figure 4.5 FPA–FTIR microspectroscopy of rhESCs cultured in FGF2/KOSR or ALGöX. (A) Smoothed (13–point) and normalized absorbance spectra (4000–800 cm^{-1}). (B) Second–derivative spectra normalized by extended multiplicative signal correction (EMSC). (C) for 300–2800 cm^{-1} regions. (D) for 1800–800 cm^{-1} regions.

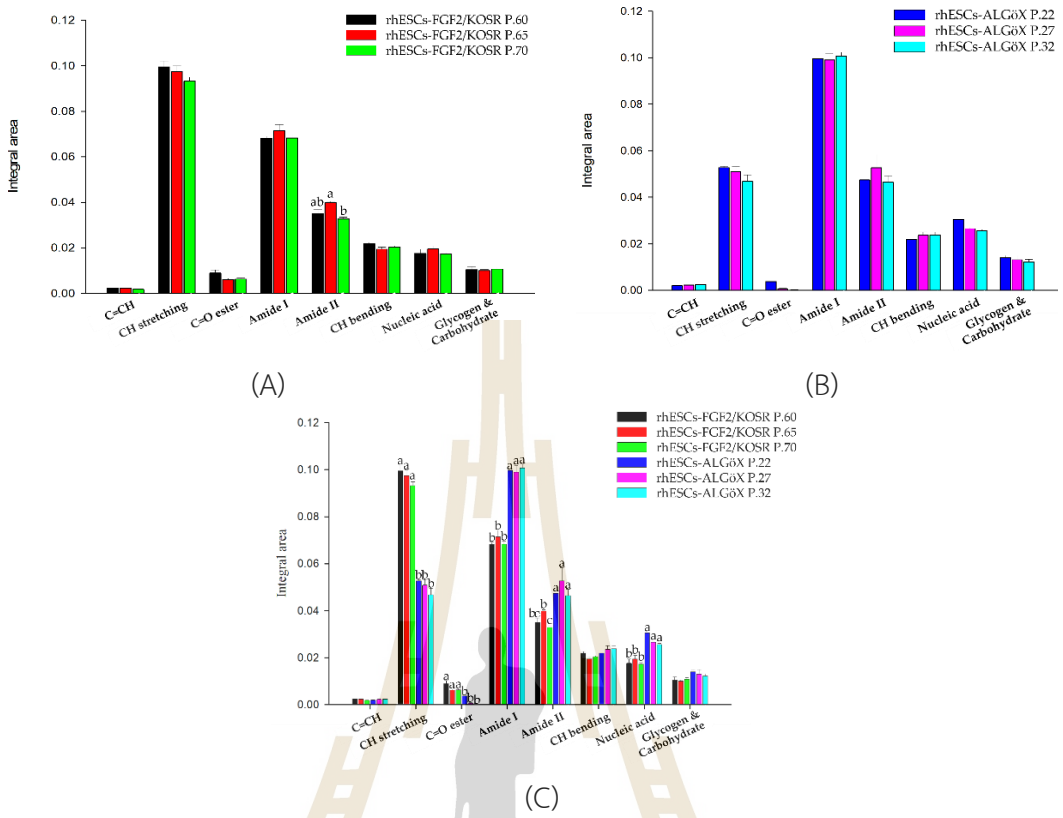

Figure 4.6 Histograms of relative integrated areas of macromolecular components from normalized second-derivative spectra (OPUS 7.5). (A) rhESCs cultured in FGF2/KOSR. (B) rhESCs cultured in ALGöX. (C) rhESCs cultured in FGF2/KOSR or ALGöX. Different letters (a, b, c, d, e) above the bars indicate a statistically significant difference ($p<0.05$) between the groups.

Naïve-like rhESCs-ALGöX displayed significantly higher integrated absorbance in the amide I, amide II, and nucleic acid regions compared to primed rhESC-FGF2/KOSR cells ($p < 0.05$, ANOVA) (Figure 4.6). Notably, naïve-like cells exhibited prominant bands at 1654 and 1546 cm^{-1}, corresponding to amide I and II, and strong nucleic acid peaks at 1239 and 1085 cm^{-1}. In contrast, primed cells showed stronger lipid-associated absorbance, particularly in the CH_2/CH_3 stretching region (2921–2852 cm^{-1}) and the ester carbonyl stretch at 1741 cm^{-1} (Figure 4.6C). These spectral differences suggest higher protein synthesis and nucleic acid content in naïve-like cells, while primed cells are enriched in membrane lipid signatures (Crowe et al., 1996; Dunkhunthod et al., 2017; Cornacchia et al., 2019; Yousefi et al., 2019; Aksoy et al., 2021). No statistically significant spectral variation was observed between passages within either state (Figure 4.6C).

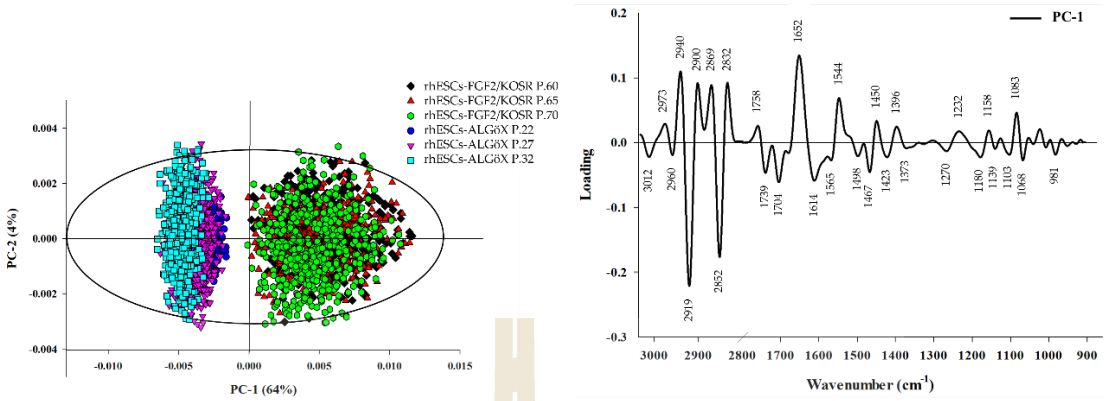

Figure 4.7 Principal component analysis of rhESCs based on FTIR spectra. (A) PCA of the full spectral range (800–4000 cm^{-1}). (B) PC–1 loading plots from independent spectra.

4.3.2 Principal component analysis (PCA) of rhESCs based on FTIR spectra

PCA was performed on the second derivative spectra from rhESC-FGF2/KOSR and rhESC-ALGöX cells in order to visualize of clustering of similar spectra within datasets in scatter plots; and identification of variables (spectral bands representing various molecular groups within the samples) in loading plots. Unsupervised PCA of the full spectral dataset revealed clear segregation between cell states. PC–1 (64% variance) and PC–2 (4% variance) of the total variance effectively separated naïve-like rhESC-ALGöX and primed rhESC-FGF2/KOSR cells (Figure 4.7A). Loadings on PC–1 indicated strong contributions from lipid–associated bands (3000–2800 cm^{-1}), Lipid ester carbonyl (1750–1700 cm^{-1}), protein Amide I and Amide II–related bands (1700–1500 cm^{-1}), Phosphodiester bond from nucleic acid (1240 and 1080 cm^{-1}) (Figure 4.7B). Primed rhESC-FGF2/KOSR cells showed positive PC–1 scores associated with negative loading from CH–stretching (2852 and 2919 cm^{-1}) and ester carbonyls (1739 and 1704 cm^{-1}), while negative score plot from naïve-like rhESC-ALGöX cells were associated with positive loading from strong amide I absorption (~1652 cm^{-1}), amide II absorption (~1544 cm^{-1}) and phosphodiester bond from nucleic acid at 1232 and 1083 cm^{-1} (Table 4.1).

Table 4.1 Band maxima distinguishing primed and naïve-like pluripotent states in rhESCs as identified by FPA–FTIR microspectroscopy.

Band maxima second derivative spectra (cm^{-1})		PC–1 loading (cm^{-1})		Band assignments (Movasaghi et al., 2008; Cao et al., 2013c; Cao et al., 2014; Dunkhunthod et al., 2017; Cornacchia et al., 2019)
rhESCs-FGF2/KOSR P60, P65, P70	rhESCs-ALGöX P22, P27, P32	Negative loading	Positive loading	
2960	2960		2940	CH_3 asymmetric stretch due to methyl terminal of membrane phospholipids: mainly lipids
2921	2921	2919	2900	CH_2 asymmetric stretch of the methylene group of membrane phospholipids: mainly lipids, with some contribution from proteins, carbohydrates, nucleic acids
2852	2852	2852	2869, 2832	CH_2 symmetric stretching: mainly lipids, with some contribution from proteins, carbohydrates, nucleic acids
1741	1741	1739, 1704		C=O stretching vibrations of lipids (triglycerides and cholesterol esters)
1654	1654	1614	1652	Amide I: C=O (80%) and C—N (10%) stretching, N—H (10%) bending vibrations: proteins α–helix
1546	1546		1544	Amide II: N—H (60%) bending and C—N (40%) stretching vibrations: proteins α helix
1463	1463	1467	1450	CH_2 bending vibrations: lipids and proteins, Cholesterol–methyl band
1396	1396		1396	COO^- stretching vibrations of amino acid side chains
1239	1239		1232	PO_2^- asymmetric stretching vibrations: RNA, DNA, and phospholipids
1172	1157, 1022	1180	1158	C–O–C vibrations from glycogen and other carbohydrates
1085	1085		1083	PO_2^- symmetric stretching vibrations: RNA, DNA
1054	1054	1068		C—O vibrations from glycogen and other carbohydrates
966	966	981		C—O deoxyribose, C–C DNA

4.3.3 Partial least squares–discriminant analysis (PLS–DA) of rhESCs based on FTIR spectra

Primed or naïve-like spectra from each group were randomly separated into calibration and validation sets, comprising approximately two–thirds and one–third of spectra, respectively. A total of 1581 spectra from primed cells and 3317 spectra from naïve-like cells were used for the analysis. The calibration data matrix employed for PLS–DA consisted of the spectral dataset (multivariate X) and two Y variables with integer values of 0 or 1 coding for the each of the two modelled spectral classes. Classification of the dataset was then carried out by predicting a Y value for each spectrum in the independent validation using PLS models that had been generated from the calibration sets. Correct classification of each class was arbitrarily assigned to samples with predicted Y > 0.5 for respective spectra. The resulting model demonstrated strong predictive performance, with a correlation coefficient of R = 0.95 (Figure 4.8A, 4.8B). Classification accuracy reached 100% specificity and 100% sensitivity in identifying primed and naïve-like samples, respectively.

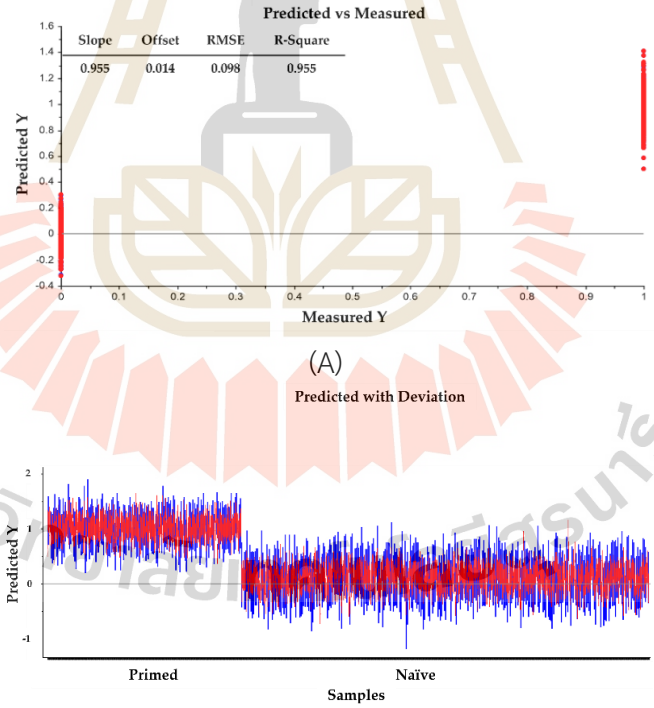

Figure 4.8 PLS–DA modeling of rhESCs based on FTIR spectra. (A) Calibration (training) set showing measured versus predicted Y values, with primed state = +1 and naïve-like state = 0. (B) Validation set predictions using the PLS–DA model.

4.4 Discussion

Our study demonstrates that FTIR microspectroscopy is a robust, label–free, and non–invasive approach for distinguishing primed and naïve-like states in rhesus macaque ESCs. FPA detectors use multiple elements that allows for the simultaneous measurement of all data points from each detector element in the spectral interval recorded, with each detector pixel recording independently. Each detector pixel functions as an aperture and records the entire spectrum. Using FPA detector, we identified reproducible spectral signatures that correspond to characteristic differences in protein, lipid, and nucleic acid composition between the two pluripotent states (Bassan et al., 2010; Whelan et al., 2011).

Naïve-like rhESCs cultured in ALGöX medium exhibited stronger absorbance in the amide I (~1654 cm^{-1}) and amide II (~1546 cm^{-1}) regions. Prominent peaks at 1240 and 1080 cm^{-1}, attributed to phosphodiester groups in nucleic acids, suggest increased transcriptional activity and molecular complexity. These biochemical features align with the molecular hallmarks of naïve pluripotency, including greater developmental plasticity and retention of an earlier embryonic identity (Aksoy et al., 2021).

In contrast, primed rhESCs maintained in FGF2/KOSR medium displayed stronger lipid–associated absorbance, including CH_2/CH_3 stretching vibrations (2921–2852 cm^{-1}) and ester carbonyl peaks (1741 cm^{-1}). Similar lipid enrichment has been reported in primed human ESCs analyzed by FPA–FTIR microspectroscopy, where high CH_2/CH_3 intensity was linked to membrane remodeling and metabolic adaptations (Dunkhunthod et al., 2017). Metabolically, naïve PSCs rely predominantly on oxidative phosphorylation, whereas primed PSCs shift toward glycolysis and increased lipid utilization, which may underlie the lipid signatures observed here (Cornacchia et al., 2019; Yousefi et al., 2019).

Multivariate analyses reinforced these findings: PCA revealed clear separation of primed and naïve-like spectra, while PLS–DA achieved 100% specificity and 100% sensitivity in classifying cell states. These FTIR–based distinctions were fully consistent with immunocytochemistry—showing mutually exclusive expression of OTX2 in primed cells and KLF17, TFCP2L1, ALPPL2, and TFAP2C in naïve-like cells—and with transcriptomic data that confirmed state–specific gene expression patterns (Pham et al., 2025; Aksoy et al., 2021; Amzal et al., 2025).

This work represents, to our knowledge, the first FTIR–based biochemical profiling of primed and naïve-like pluripotent states in a non–human primate model. It establishes FTIR microspectroscopy as a new approach for pluripotent stem cell phenotyping at the single–cell level (Bassan et al., 2010; Heraud et al., 2010). A key

question is whether this approach can reliably distinguish the naïve pluripotent state induced under diverse culture conditions and in multiple species. The discovery of a conserved and distinctive molecular signature would greatly enhance our ability to identify and validate the naïve state across experimental systems. With developing and validating PLS–DA models based on much higher sample numbers this technique might further tested and ultimately applied as a practical tool for optimization for identifying stem cells. In our opinion, this study represents the first steps toward achieving this aim. We employed the QUASAR 1.11.1 software (Toplak et al., 2021) for cell classification. The software was selected for its ability to analyze a large number of spectra specifically, more than 3,000 spectra simultaneously a significant advantage over the Unscrambler X 10.3 software, which has limitations in this regard. The results demonstrated that the QUASAR software achieved 99% specificity and 99% sensitivity (Appendix, Figure A1), shows cross–validated prediction quality of 3 methods: Tree, Random Forest and support vector machine (SVM) Confusion matrix from SVM model, matching the performance of Unscrambler X. Furthermore, the QUASAR software offers platform flexibility and high processing speed. For future work, we plan to compile a comprehensive database of spectra from various types of stem cells and integrate machine learning for cell classification. This approach is intended to substantially reduce the high costs, complex procedures, and extensive time currently associated with traditional analysis method.

CHAPTER V
CONCLUSION

This work represents the first application of FTIR microspectroscopy to characterize and distinguish primed from naïve-like rhESCs. The FTIR–based findings were validated against two established methods immunocytochemistry and RNA sequenching which confirmed that naïve-like cells possess higher protein content, while primed cells are enriched in lipids. The three analytical approaches were complementary: immunocytochemistry localized proteins, RNA sequencing provided gene–level resolution, and FTIR offered a rapid, cost–effective overview of cellular biochemical composition.

Immunolabeling confirmed that rhESCs in both culture conditions expressed core pluripotency markers; OCT4, NANOG, and SOX2. However, the expression of these markers was more stable in rhESC-ALGöX cells across passages, while expression in rhESC-FGF2/KOSR peaked at passage 60 and declined thereafter. Critically, a key distinction was the exclusive detection of the primed–state marker OTX2 in rhESC-FGF2/KOSR cells. Conversely, rhESC-ALGöX cells showed robust expression of naïve markers, including KLF17, ALPPL2, TFCP2L1, and TFAP2C, confirming a naïve pluripotency signature. These findings were further validated by RNA sequencing, which demonstrated a clear transcriptional divergence between the two cell populations. The rhESC-FGF2/KOSR cells exhibited elevated expression of primed–state genes (*NODAL*, *OTX2*, *ETV4*, *BMP4*, *FST*, and *SOX3*). In contrast, rhESC-ALGöX cells showed high expression of naïve–associated genes (*KLF2*, *DPPA2*, *DPPA3*, *ZFP42*, *PRDM14*, and *TFCP2L1*). These results collectively indicate that the ALGöX medium successfully reprogrammed primed rhESCs to a stable, naïve-like pluripotent state.

FTIR microspectroscopy served as a robust, label–free method for distinguishing between the two pluripotent states. Analysis of the second–derivative spectra revealed distinct biochemical profiles. Naïve-like rhESCs exhibited significantly higher integrated absorbance in the amide I and amide II regions, indicating higher protein synthesis. They also showed strong nucleic acid peaks at 1239 and 1085 cm^{-1}, suggesting increased transcriptional activity. In contrast, primed rhESCs showed stronger lipid–associated absorbance, particularly in the CH_2/CH_3 stretching region (2921–2852 cm^{-1}) and the ester carbonyl stretch at 1741 cm^{-1}, consistent with a metabolic shift towards increased lipid utilization.

PCA effectively segregated the primed and naïve-like cells into two distinct clusters based on their spectral signatures, with PC–1 accounting for 64% of the variance. Additionally, a PLS–DA model was developed, which achieved 100% specificity and 100% sensitivity in classifying the two cell states. This demonstrates the high predictive power of FTIR microspectroscopy for identifying rhESCs. These FTIR–based biochemical distinctions were consistent with the results from immunocytochemistry and transcriptomic data, establishing FTIR microspectroscopy as a powerful new approach for monitoring stem cell identity and quality.

REFERENCES

Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P.W., Beighton, G., Bello, P.A., Benvenisty, N., Berry, L.S., Bevan, S., Blum, B., Brooking, J., Chen, K.G., Choo, A.B.H., Churchill, G.A., Corbel, M., Damjanov, I., Draper, J.S., Dvorak, P., Emanuelsson, K., Fleck, R.A., Ford, A., Gertow, K., Gertsenstein, M., Gokhale, P.J., Hamilton, R.S., Hampl, A., Healy, L.E., Hovatta, O., Hyllner, J., Imreh, M.P., Itskovitz-Eldor, J., Jackson, J., Johnson, J.L., Jones, M., Kee, K., King, B.L., Knowles, B.B., Lako, M., Lebrin, F., Mallon, B.S., Manning, D., Mayshar, Y., McKay, R.D.G., Michalska, A.E., Mikkola, M., Mileikovsky, M., Minger, S.L., Moore, H.D., Mummery, C.L., Nagy, A., Nakatsuji, N., O'Brien, C.M., Oh, S.K.W., Olsson, C., Otonkoski, T., Park, K., Passier, R., Patel, H., Patel, M., Pedersen, R., Pera, M.F., Piekarczyk, M.S., Reijo Pera, R.A., Reubinoff, B.E., Robins, A.J., Rossant, J., Rugg-Gunn, P., Schulz, T.C., Semb, H., Sherrer, E.S., Siemen, H., Stacey, G.N., Stojkovic, M., Suemori, H., Szatkiewicz, J., Turetsky, T., Tuuri, T., van den Brink, S., Vintersten, K., Vuoristo, S., Ward, D., Weaver, T.A., Young, L.A., and Zhang, W. (2007). Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. **Nature Biotechnology**, *25*(7), 803–816.

Afanassieff, M., Tapponnier, Y., and Savatier, P. (2014). Generation of Induced Pluripotent Stem Cells in Rabbits. Induced Pluripotent Stem (iPS) Cells**. Methods and Protocols**, *1357*, 149-172.

Aksoy, I., Rognard, C., Moulin, A., Marcy, G., Masfaraud, E., Wianny, F., Cortay, V., Bellemin-Ménard, A., Doerflinger, N., Dirheimer, M., Mayère, C., Bourillot, P.Y., Lynch, C., Raineteau, O., Joly, T., Dehay, C., Serrano, M., Afanassieff, M., and Savatier, P. (2021). Apoptosis, G1 phase stall, and premature differentiation account for low chimeric competence of human and rhesus monkey naïve pluripotent stem cells. **Stem cell reports**, *16*(1), 56-74.

Ami, D., Neri, T., Natalello, A., Mereghetti, P., Doglia, S.M., Zanoni, M., Zuccotti, M., Garagna, S., and Redi, C.A. (2008). Embryonic stem cell differentiation studied by FT-IR spectroscopy. **Biochim Biophys Acta**, *1783*(1), 98-106.

Amzal, A., Pijoff, Y., Rognard, C., Marcy, G., Doerflinger, N., Bréhier, C., Anwised, P., Srisutush, J., Chimngam, M., Chatdarong, K., Cazotte, E., Moyret, C., Kuorome, M., Cortay, V., Dehay, C., Afanassieff, M., Joly, T., Rougeulle, C., Gillet, G., Wolf, E., Parnpai, R., and Savatier, P., and Aksoy, I. (2025). MEK and AKT signaling

Determine the Potential of Pluripotent stem Cells to Colonize Heterologous Preimplantation Embryos. **Nature cell biology**

Anwised, P., Moorawong, R., Samruan, W., Somredngan, S., Srisutush, J., Laowtammathron, C., Aksoy, I., Parnpai, R., and Savatier, P. (2023). An expedition in the jungle of pluripotent stem cells of non-human primates. **Stem Cell Reports**, *18*(11), 2016-2037.

Baker, M.J., Trevisan, J., Bassan, P., Bhargava, R., Butler, H.J., Dorling, K.M., Fielden, P.R., Fogarty, S.W., Fullwood, N.J., Heys, K.A., Hughes, C., Lasch, P., Martin-Hirsch, P.L., Obinaju, B., Sockalingum, G.D., Sulé-Suso, J., Strong, R.J., Walsh, M.J., Wood, B.R., Gardner, P., and Martin, F.L. (2014). Using Fourier transform infrared (FTIR) spectroscopy to analyse biological materials. **Nature Protocols**, *9*(8), 1771–1791.

Bassan, P., Kohler, A., Martens, H., Lee, J., Byrne, H.J., Dumas, P., Gazi, E., Brown, M., Clarke, N., and Gardner, P. (2010) Resonant Mie scattering (RMieS) correction of infrared spectra from highly scattering biological samples. **Analyst**, *135*(2), 268-277.

Bassan, P., Byrne, H.J., Bonnier, F., Lee, J., Dumas, P., and Gardner, P. (2009). Resonant Mie scattering in infrared spectroscopy of biological materials–understanding the ‘dispersion artefact’. **Analyst**, *134*(8), 1586-1593.

Boroviak, T., and Nichols, J. (2017). The pluripotency continuum: Naive, primed and beyond. **Development**, *144*(8), 1361–1371.

Bradley, M., Jarrell, G.A., and Kim, E.H. (1984). On the existence of an optimal capital structure: Theory and evidence. **The journal of Finance**, *39*(3), 857-878.

Buchwalow, I. B., Böcker, W., and Samoilova, V. (2011). Immunohistochemistry: History and principles. **In Immunohistochemistry** (pp. 1-10). Springer.

Bustin S.A. (2000). Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. **J Mol Endocrinol**. *25*(2), 169-93.

Cao, J., Ng, E. S., McNaughton, D., Stanley, E. G., Elefanty, A. G., Tobin, M. J., and Heraud, P. (2014a). Fourier transform infrared microspectroscopy reveals unique phenotypes for human embryonic and induced pluripotent stem cell lines and their progeny. **Journal of Biophotonics**, *7*(10), 767-781.

Cao, J., Ng, E. S., McNaughton, D., Stanley, E. G., Elefanty, A. G., Tobin, M. J., and Heraud, P. (2013b). Fourier transform infrared microspectroscopy reveals that tissue culture conditions affect the macromolecular phenotype of human embryonic stem cells. **Analyst**, *138*(14), 4147-4160.

Cao, J., Ng, E. S., McNaughton, D., Stanley, E. G., Elefanty, A. G., Tobin, M. J., and Heraud, P. (2013c). The characterisation of pluripotent and multipotent stem cells using Fourier transform infrared microspectroscopy. **Int J Mol Sci**, *14*(9), 17453-76.

Chan, K. L. A., and Kazarian, S. G. (2013). Attenuated total reflection (ATR) - FTIR spectroscopy and imaging of biological samples. **Analyst**, *138*(14), 3698-3707.

Chonanant, C., Jearanaikoon, N., Leelayuwat, C., Limpaiboon, T., Tobin, M. J., Jearanaikoon, P., and Heraud, P. (2011). Characterisation of chondrogenic differentiation of human mesenchymal stem cells using synchrotron FTIR microspectroscopy. **Analyst**, *136*(12), 2542-2551.

Coons, A. H. (1956). Histochemistry with labelled antibody. **International Review of Cytology**, *5*, 1–23.

Cornacchia, D., Zhang, C., Zimmer, B., Chung, S.Y., Fan, Y., Soliman, M.A., Tchieu, J., Chambers, S.M., Shah, H., Paull, D., Konrad, C., Vincendeau, M., Noggle, S.A., Manfredi, G., Finley, L.W.S., Cross, J.R., Betel, D., and Studer, L. (2019). Lipid Deprivation Induces a Stable, Naïve-to-Primed Intermediate State of Pluripotency in Human PSCs. **Cell stem cell**, *25*(1), 120-136.

Cregger, M., Berger, A. J., and Rimm, D.L. (2006). Immunohistochemistry and Quantitative Analysis of Protein Expression. **Arch Pathol Lab**, *130*, 1026–1030.

Crowe, J. H., Hoekstra, F. A., Crowe, L. M., Anchordoguy, T. J., and Drobnis, E. (1996). Lipid phase transitions measured in intact cells with Fourier transform infrared spectroscopy. **Cryobiology**, *33*(4), 463-477.

Depciuch, J., Parcheta, R., Depciuch, G., and Kasprzyk, A. (2018). An introduction to FTIR spectroscopy in biomedical applications. **Journal of Applied Spectroscopy**, *85*(2), 169–181.

Dunkhunthod, B.; Thumanu, K., and Eumkeb, G. (2017). Application of FTIR microspectroscopy for monitoring and discrimination of the anti-adipogenesis activity of baicalein in 3T3-L1 adipocytes. **Vibrational Spectroscopy**, *89*, 92-101.

Evans, M. J., and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. **Nature**, *292*(5819), 154-156.

Griffiths, P. R. (1983). Fourier Transform Infrared Spectrometry. **Science**, *222*(4621), 297-302.

Guo, G., von Grotthuss, M., and Yang, B. (2017). The molecular basis of naïve and primed pluripotency. **Current Opinion in Cell Biology**, 45, 63-70.

Haines, D. M., and Woytowych, J. C. (2007). A brief review of immunohistochemistry and the use of immunocytochemistry in the detection of viral antigens in

veterinary medicine. **The Canadian Journal of Veterinary Research**, *71*(1), 1-10.

Heraud, P., Ng, E.S., Caine, S., Yu, Q.C., Hirst, C., Mayberry, R., Bruce, A., Wood, B.R., McNaughton, D., Stanley, E.G., and Elefanty, A.G. (2010). Fourier transform infrared microspectroscopy identifies early lineage commitment in differentiating human embryonic stem cells. **Stem Cell Research**, *4*(2), 140-147.

Heraud, P., Caine, S., Sanson, G., Gleadow, R., Wood, B.R., and McNaughton, D. (2007). Focal plane array infrared imaging: a new way to analyse leaf tissue. **New Phytologist**, *173*(1), 216-225.

Higuchi, R., Fockler, C., Dollinger, G., and Watson, R. (1993). Kinetic PCR analysis: Real-time monitoring of DNA amplification reactions. **Biotechnology (N Y)**, *11*(9), 1026-1030.

Jaenisch, R., and Young, R. (2008). Stem cells, the pluripotent state, and induced pluripotency. **Cell**, *132*(4), 567-582.

Kazarian, S. G., and Chan, K. L. A. (2017). Applications of FTIR microspectroscopy in polymer science. **Applied Spectroscopy Reviews**, *52*(1), 1-32.

Krimm, S., and Bandekar, J. (1986). Vibrational analysis of peptides, polypeptides, and proteins. **Advances in Protein Chemistry**, 38, 181-364.

Kumari, D. (2016). States of pluripotency: Naïve and primed pluripotent stem cells. **Pluripotent Stem Cells-from the bench to the clinic**, 31-45.

Lassiter, R. M., Ma, Z., Nallapeta, S., and Bhargava, R. (2019). Rapid and high-throughput infrared spectroscopic imaging for breast cancer diagnosis. **Journal of Biophotonics**, *12*(6), e201800366.

Li, J., Zhu, Q., Cao, J., Liu, Y., Lu, Y., Sun, Y., Li, Q., Huang, Y., Shang, S., Bian, X., Li, C., Zhang, L., Wang, Y., Nie, Y., Fu, J., Li, W., Mazid, M.A., Jiang, Y., Jia, W., Wang, X., Sun, Y., Esteban, M.A., Sun, Q., Zhou, F., and Liu, Z. (2023). Cynomolgus monkey embryo model captures gastrulation and early pregnancy. **Cell Stem Cell**, *30*(4), 362-377.

Marks, H., Kalkan, T., and Stunnenberg, H. G. (2012). The transcriptional programs of naïve and primed pluripotency. **Nature Reviews Genetics**, *13*(10), 717-729.

Martin, G. R., and Evans, M. J. (1974). The morphology and growth of a pluripotent teratocarcinoma cell line and its derivatives in tissue culture. **Cell**, *2*(3), 163-172.

Mendelsohn, R., and Moore, D. J. (2012). FTIR spectroscopy in microbiology: a review. **Biochemical Spectroscopy**, *1*(1), 1-13.

Miller, L. M., Sowa, M. J., and Riemer, R. (2014). Infrared spectroscopy of brain tumors. **Journal of Biophotonics**, *7*(11-12), 978-984.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. **Nature Methods**, *5*(7), 621–628.

Movasaghi, Z., Rehman, S., and Rehman, D. I. (2008). Fourier transform infrared (FTIR) spectroscopy of biological tissues. **Applied Spectroscopy Reviews**, *43*, 134-179.

Nichols, J., and Smith, A. (2009). Naïve and primed pluripotent states. **Cell stem cell**, *4*, 487-492.

Pham, H.T., Perold, F., Pijoff, Y., Doerflinger, N., Rival-Gervier, S., Givelet, M., Moulin, A., Ressaire, M., Da Silva Fernandes, E., Bidault, V., Jouneau, L., Duranthon, V., Wianny, F., Pain, B., Plotton, I., Joly, T., Afanassieff, M., Savatier, P., and Beaujean, N. (2025). Efficient generation of germline chimeras in a non-rodent species using rabbit induced pluripotent stem cells. **Nature Communications**, *16*(1), 5165.

Ramos-Vara, J. A., and Miller, M. A. (2014). When is a negative immunohistochemical result a true negative? **Veterinary Pathology**, *51*(2), 329-335.

Reis, C., Sousa, R., Reis, R. L., and Gomes, M. E. (2020). FTIR spectroscopy and imaging as a tool to assess the chemical composition of biological tissues. **Current Opinion in Biotechnology**, *63*, 211-220.

Stark, R., Grzelak, M., and Hadfield, L. J. (2019). RNA sequencing: The teenage years. **Nature Reviews Genetics**, *20*(6), 337-352.

Stuart, B. H. (2004). Infrared Spectroscopy: Fundamentals and Applications. John Wiley & Sons.

Suemori, H., Yasuchika, K., and Nakatsuji, N. (2006). Maintenance of pluripotency in primate ES cells. **Cloning and Stem Cells**, *8*(2), 107-112.

Suemori, H., Tada, T., Torii, R., Hosoi, Y., Kobayashi, K., Imahie, H., and Nakatsuji, N. (2001). Establishment of embryonic stem cell lines from cynomolgus monkey blastocysts produced by IVF or ICSI. **Dev Dyn**, *222*(2), 273-279.

Szymanski, T., Piekarski, J., and Depciuch, J. (2018). FTIR spectroscopy for the diagnosis of colon cancer. **Journal of Molecular Structure**, *1157*, 181-187.

Takahashi, S., Kobayashi, S., and Hiratani, I. (2018). Epigenetic differences between naïve and primed pluripotent stem cells. **Cellular and Molecular Life Sciences**, *75*(7), 1191-1203.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. **Cell**, *131*(5), 861-872.

Theunissen, T. W., Powell, B. E., and Jaenisch, R. (2014). Molecular criteria for pluripotent stem cells: a new tool for regenerative medicine. **Stem Cell**, *32*(4), 1017-1026.

Theunissen, T. W., and Jaenisch, R. (2014). Molecular criteria for pluripotent stem cells. **Nature Biotechnology**, *32*(2), 165-174.

Thomson, J. A., and Marshall, J. E. (1998a). Pluripotent stem cell lines from human and non-human primate embryos. **Stem Cells**, *16*(2), 79-87.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., and Jones, J. M. (1998b). Embryonic stem cell lines derived from human blastocysts. **Science**, *282*(5391), 1145-1147.

Thomson, J. A., and Marshall, V. S. (1998c). Primate embryonic stem cells. **Current Topics in Developmental Biology**, *38*, 133-165.

Thomson, J. A., Kalishman, J., Golos, T. G., Durning, M., Harris, C. P., and Hearn, J. P. (1996). Pluripotent cell lines derived from common marmoset (Callithrix jacchus) blastocysts. **Biol Reprod**, *55*(2), 254-259.

Thomson, J. A., Kalishman, J., Golos, T. G., Durning, M., Harris, C. P., Becker, R. A., and Hearn, J. P. (1995). Isolation of a primate embryonic stem cell line. **Proceedings of the National Academy of Sciences**, *92*(17), 7844-7848.

Thumanu, K., Tanthanuch, W., Ye, D., Sangmalee, A., Lorthongpanich, C., Parnpai, R., ... and Heraud, P. (2011). Spectroscopic signature of mouse embryonic stem cell–derived hepatocytes using synchrotron Fourier transform infrared microspectroscopy. **Journal of Biomedical Optics**, *16*(5), 057005-057005.

Toplak, M., Read, S. T., Sandt, C., and Borondics, F. (2021). Quasar: easy machine learning for Biospectroscopy. **Cells**, *10*(9), 2300.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. **Nature Reviews Genetics**, *10*(1), 57–63.

Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J. H. (2016). Dynamic stem cell states: naïve to primed pluripotency in rodents and humans. **Nature reviews Molecular cell biology**, *17*(3), 155-169.

Whelan, D. R., Bambery, K. R., Heraud, P., Tobin, M. J., Diem, M., McNaughton, D., and Wood, B. R. (2011). Monitoring the reversible B to A-like transition of DNA in eukaryotic cells using Fourier transform infrared spectroscopy. **Nucleic acids research**, *39*(13), 5439-5448.

Wianny, F., Bernat, A., Huissoud, C., Marcy, G., Markossian, S., Cortay, V., Giroud, P., Leviel, V., Kennedy, H., Savatier, P., and Dehay, C. (2008). Derivation and cloning of a novel rhesus embryonic stem cell line stably expressing tau-green fluorescent protein. **Stem Cells**, *26*(6), 1444-1453.

Wold, S., Esbensen, K., and Geladi, P. (1987). Principal component analysis. **Chemometrics and Intelligent Laboratory Systems**, *2*(1-3), 37-52.

Wong, K., Lee, M. A., and Chen, J. (2021). Fourier-transform infrared microspectroscopy: A powerful tool for the study of cellular and subcellular components. **Biotechnology Journal**, *16*(5), e2000287.

Yamanaka, S., and Thomson, J. (2010). Pluripotent stem cells: a new tool for regenerative medicine. **Cell Stem Cell**, *6*(6), 555–556.

Yousefi, M., Marashi, S.A., Sharifi, Z.A., and Taleahmad, S. (2019). The metabolic network model of primed/naïve human embryonic stem cells underlines the importance of oxidation-reduction potential and tryptophan metabolism in primed pluripotency. **Cell Biosci**, *9*, 71.

APPENDIX

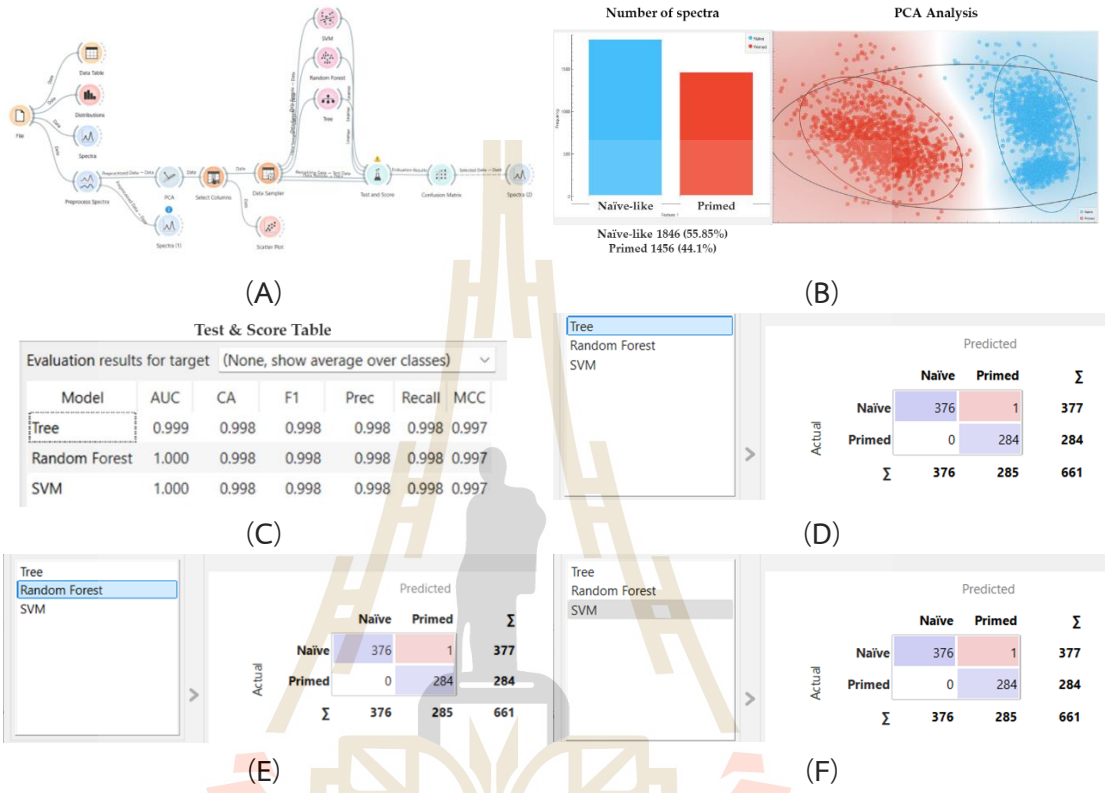

Figure A1. QUASAR: Machine Learning for cell classification. (A) A workflow for supervised analysis for rhESCs-FGF2/KOSR cells and rhESCs-ALGöX cells. (B) Number of spectra from rhESCs-FGF2/KOSR cells and rhESCs-ALGöX cells for PCA analysis. (C) The test and score table, shows cross–validated prediction quality of 3 methods. (D) Confusion matrix from Tree model. (E) Confusion matrix from Random Forest model. (F) Confusion matrix from SVM model.

BIOGAPHY

Jittanun Srisutush was born in Sakon Nakhon, Thailand on July 22th, 1992. She finished her high school from Nawaminthrachinuthit Suankularb Wittayalai Samutprakarn School in Samutprakarn. In 2015, she received her B.Sc. (Biology) from Faculty of Science, Mahasarakham University, Mahasarakham, Thailand. Then in 2020, she received her M.Sc. (Medical Microbiology) from Faculty of Medicine, Khonkaen University, Khonkaen, Thailand. In 2022, commenced Ph.D. course at School of Biotechnology, Institute of Agricultural technology, Suranaree University of Technology under supervision of Prof. Dr. Rangsun Parnpai. During her study, she had received a One Research One Graduate (OROG) scholarship from Suranaree University of Technology, Grant Number: SUT (53/2565) and the National Science, Research and Innovation Fund (NSRF) through the Program Management Unit for Human Resources & Institutional Development, Research and Innovation (PMU-B), Government of Thailand, Grant Number: B16F640104, for 6 months of Internship student at INSERM, Stem Cell and Brain Research Institute, Lyon, France.

The results of this work have been publication on International Journal of Molecular Sciences “Characterization of Rhesus Macaque Embryonic Stem Cells in Primed and Naïve-like cell States of Pluripotency Using Fourier Transform Infrared (FTIR) Microspectroscopy” in 2025.