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).

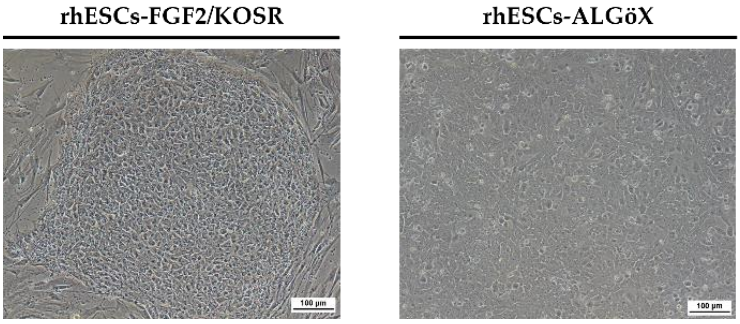


Figure 4.1 Phase contrast images of rhESCs cultured in FGF2/KOSR and ALGöX. Scale bars: 100 µm.

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.

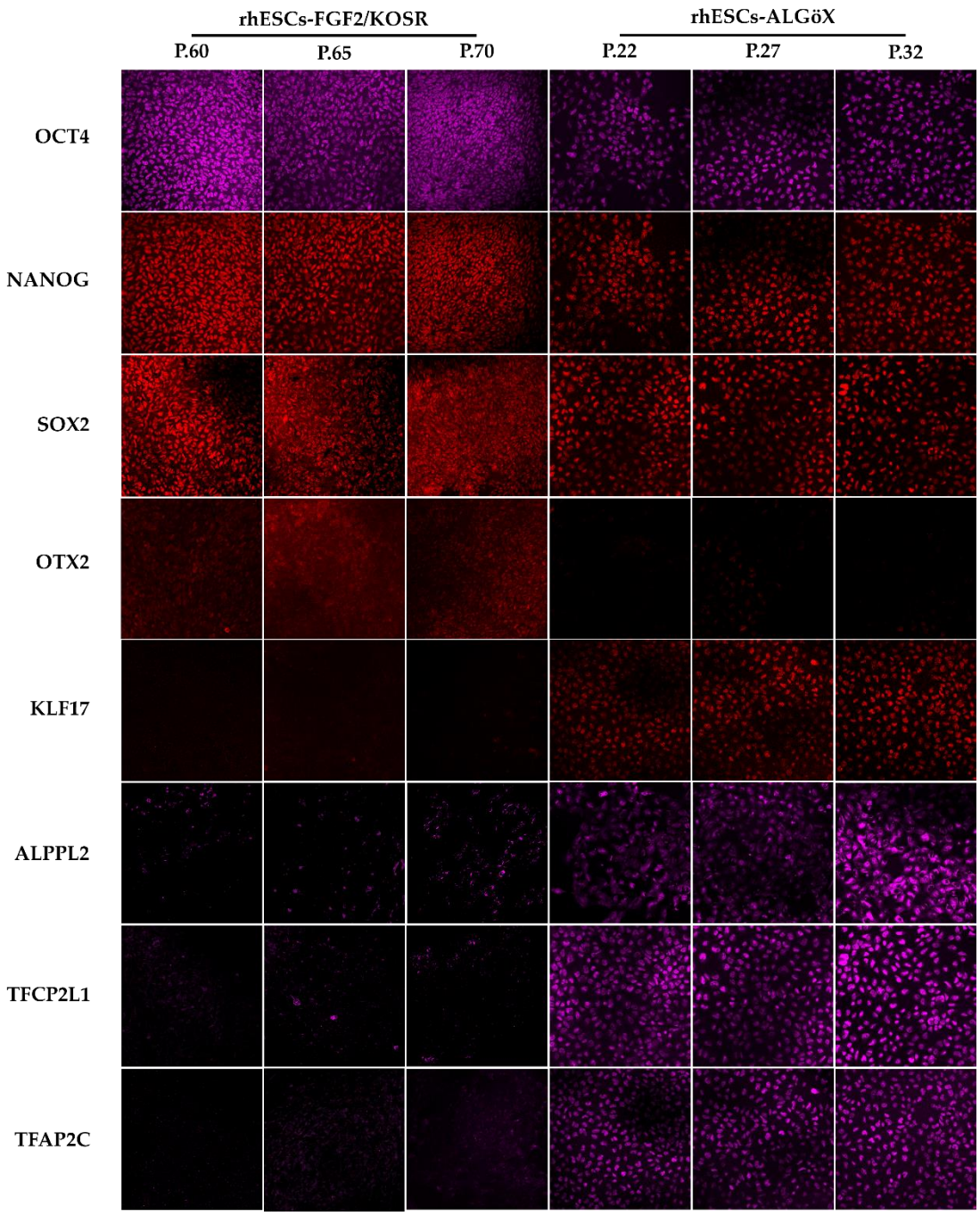


Figure 4.2 Immunolabelling of rhESCs cultured in FGF2/KOSR and ALGöX. Core pluripotency markers: OCT4, NANOG and SOX2. State–specific markers: OTX2 (primed) and KLF17, ALPPL2, TFCP2L1, and TFAP2C (naïve-like).

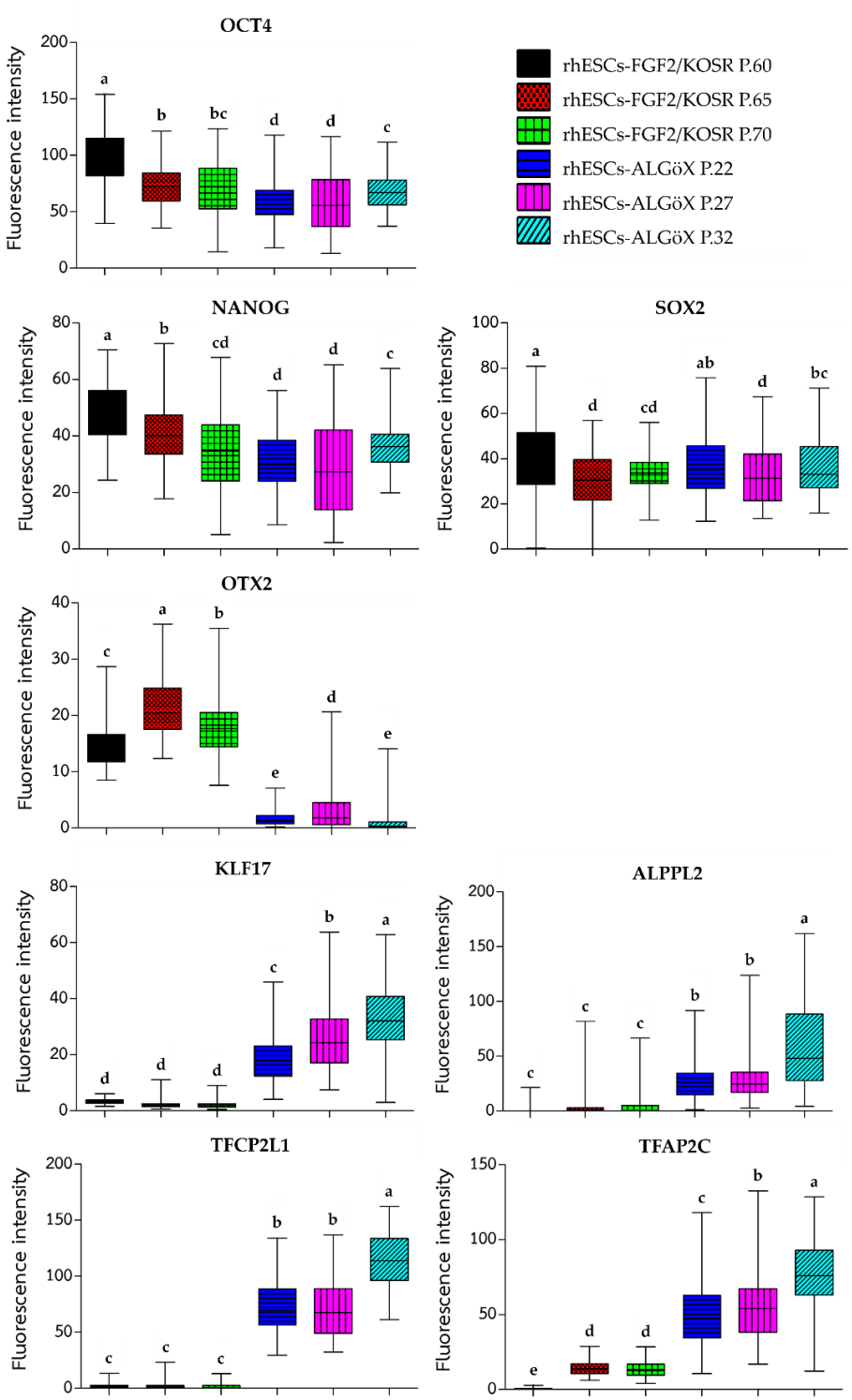


Figure 4.3 Quantification of fluorescence intensity in rhESCs cultured in FGF2/KOSR or ALGöX following immunolabelling. Different letters (a, b, c, d, e) above the bars indicate a statistically significant difference ($p<0.05$) between the groups.

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.

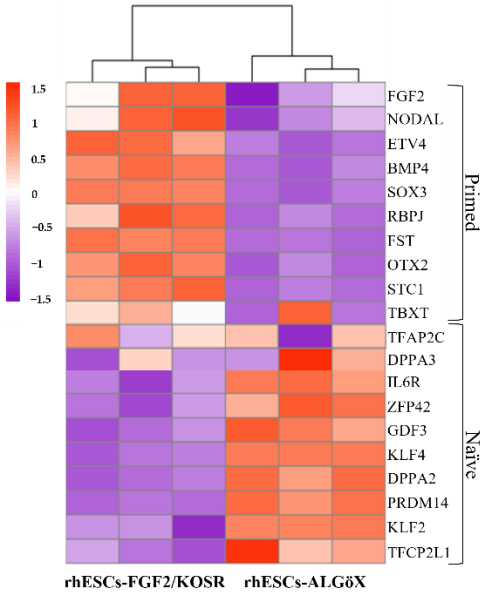


Figure 4.4 Heatmap of gene expression profiles in rhESCs cultured in FGF2/KOSR or ALGöX, based on RNA sequencing data. Genes associated with primed and naïve pluripotency states are shown.

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.

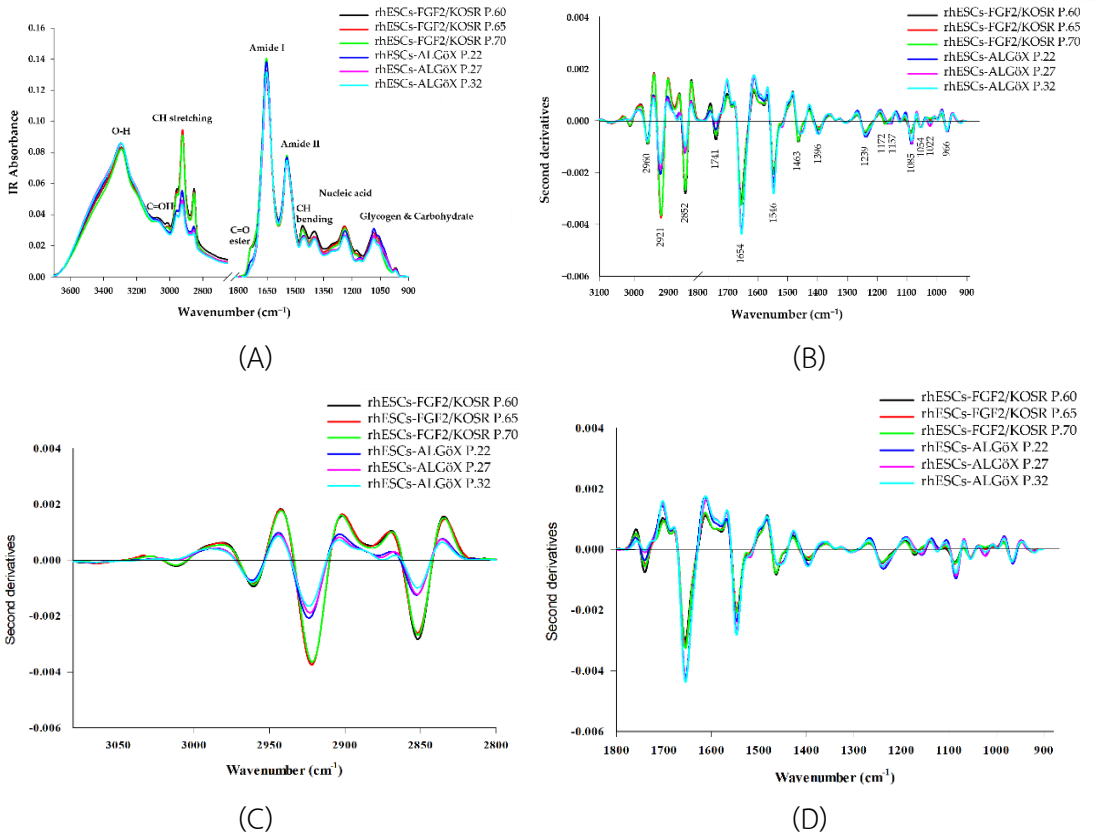


Figure 4.5 FPA–FTIR microspectroscopy of rhESCs cultured in FGF2/KOSR or ALGöX. (A) Smoothed (13–point) and normalized absorbance spectra (4000–800 cm^{-1}). (B) Second–derivative spectra normalized by extended multiplicative signal correction (EMSC). (C) for 300–2800 cm^{-1} regions. (D) for 1800–800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

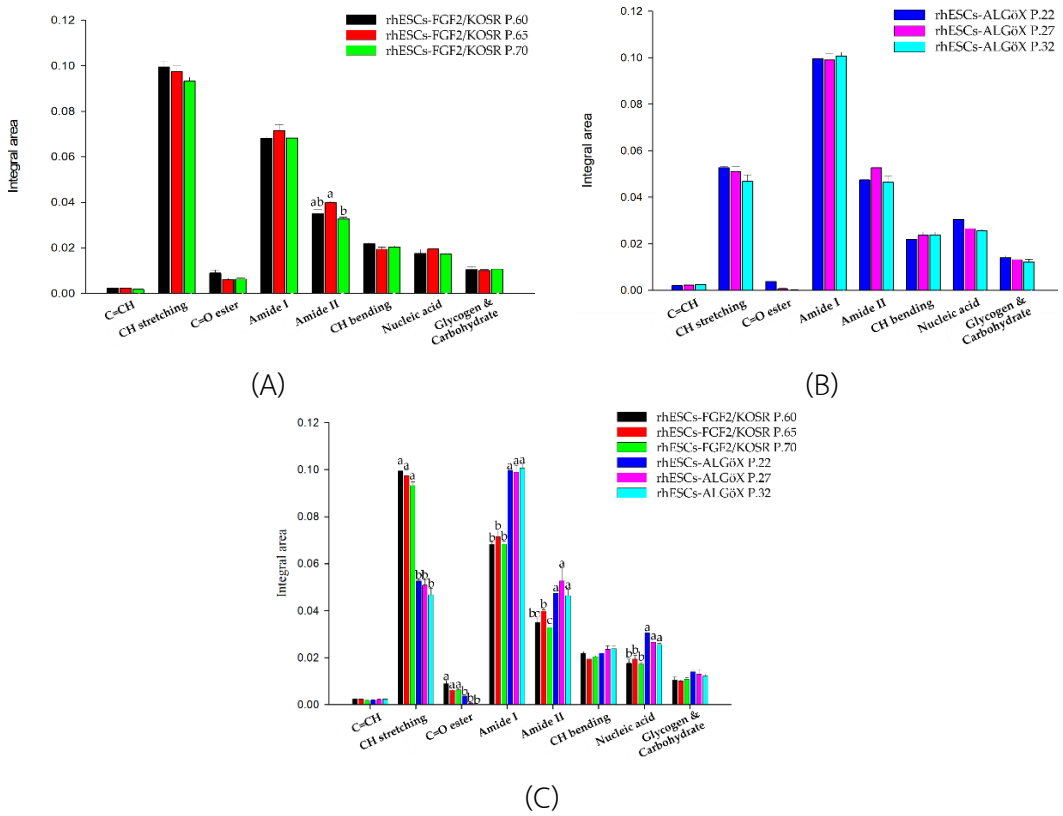


Figure 4.6 Histograms of relative integrated areas of macromolecular components from normalized second-derivative spectra (OPUS 7.5). (A) rhESCs cultured in FGF2/KOSR. (B) rhESCs cultured in ALGöX. (C) rhESCs cultured in FGF2/KOSR or ALGöX. Different letters (a, b, c, d, e) above the bars indicate a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

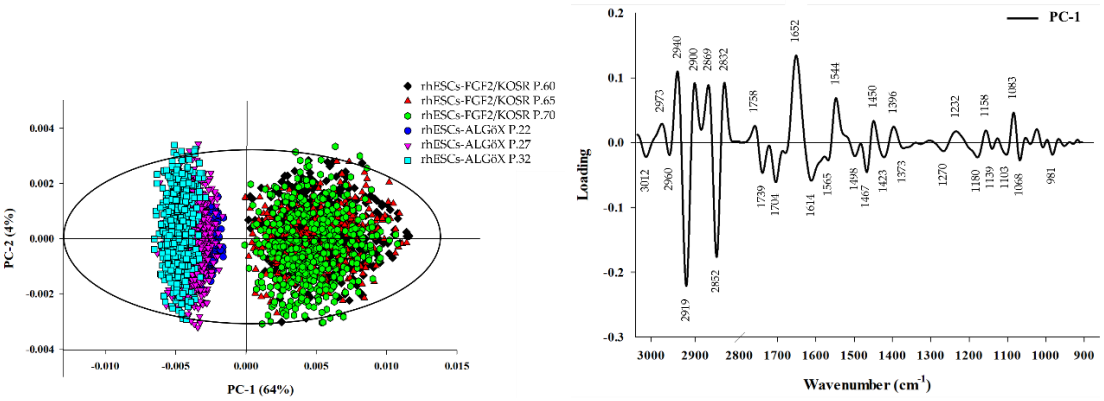


Figure 4.7 Principal component analysis of rhESCs based on FTIR spectra. (A) PCA of the full spectral range (800–4000 cm^{-1}). (B) PC–1 loading plots from independent spectra.

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

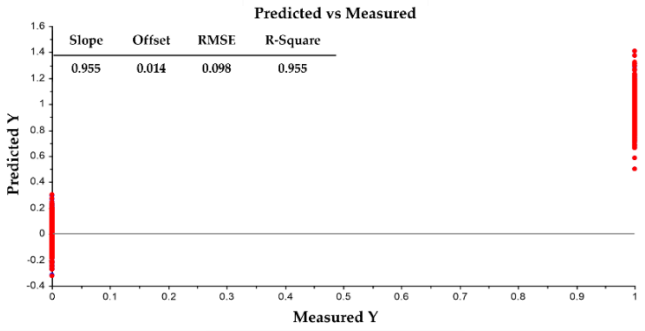
PCA was performed on the second derivative spectra from rhESC-FGF2/KOSR and rhESC-ALGöX cells in order to visualize of clustering of similar spectra within datasets in scatter plots; and identification of variables (spectral bands representing various molecular groups within the samples) in loading plots. 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.

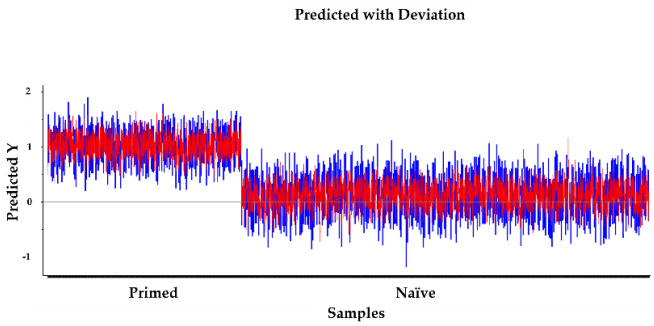
Band maxima second derivative spectra (cm^{-1})		PC–1 loading (cm^{-1})		Band assignments (Movasaghi et al., 2008; Cao et al., 2013c; Cao et al., 2014; Dunkhunthod et al., 2017; Cornacchia et al., 2019)
rhESCs-FGF2/KOSR P60, P65, P70	rhESCs-ALGöX P22, P27, P32	Negative loading	Positive loading	
2960	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.



(A)



(B)

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.