T.C.
DOKUZ EYLUL UNIVERSITY
IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

METABOLOMICS ANALYSIS OF HUMAN INDUCED PLURIPOTENT STEM CELL DERIVED LACRIMAL GLANDS

VEDAT SARI

MOLECULAR BIOLOGY AND GENETICS

MASTER OF SCIENCES THESIS

İZMİR-2022

Thesis ID: DEU.IBG.MSc.2020850022

T.C.
DOKUZ EYLUL UNIVERSITY
IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

METABOLOMICS ANALYSIS OF HUMAN INDUCED PLURIPOTENT STEM CELL DERIVED LACRIMAL GLANDS

MOLECULAR BIOLOGY AND GENETICS

MASTER OF SCIENCES THESIS

VEDAT SARI

Supervisor: Assoc. Prof. Dr. Sinan Güven
Co-Supervisor: Assoc. Prof. Dr. Duygu Sağ Wingender

This research was supported by TÜBİTAK grand number 121C313

THESIS ID: DEU.IBG.MSc. 2020850022

Dokuz Eylül Üniversitesi İzmir Uluslararası Biyotıp ve Genom Enstitüsü Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı, Moleküler Biyoloji ve Genetik Yüksek Lisans programı öğrencisi Vedat Sarı **'METABOLOMICS ANALYSIS OF HUMAN INDUCED PLURIPOTENT STEM CELL DERIVED LACRIMAL GLANDS'** konulu Yüksek Lisans tezini 12.12.2022 tarihinde başarılı olarak tamamlamıştır.

Doç. Dr. Sinan Güven

BAŞKAN

ÜYE

Prof. Dr. Emirhan Nemutlu
Hacettepe Üniversitesi

ÜYE

Doç. Dr. Gökhan Karakülah
Dokuz Eylül Üniversitesi

YEDEK ÜYE

Doç. Dr. Zeynep Fırtına Karagonlar
İzmir Ekonomi Üniversitesi

YEDEK ÜYE

Dr. Öğr. Üyesi Yavuz Oktay
Dokuz Eylül Üniversitesi

Dokuz Eylul University Izmir International Biomedicine and Genome Institute Department of Genomics and Molecular Biotechnology, Molecular Biology and Genetics graduate program Master of Science student Vedat Sarı has successfully completed his Master of Science thesis titled **'METABOLOMICS ANALYSIS OF HUMAN INDUCED PLURIPOTENT STEM CELL DERIVED LACRIMAL GLANDS'** on the date of 12.12.2022.

Assoc. Prof. Sinan Güven
CHAIR

MEMBER

Prof. Dr. Emirhan Nemutlu
Hacettepe University

MEMBER

Assoc. Prof. Gökhan Karakülah
Dokuz Eylül University

SUBSTITUTE MEMBER

Assoc. Prof. Zeynep Fırtına Karagonlar
Izmir University of Ecomomics

SUBSTITUTE MEMBER

Asst. Prof. Yavuz Oktay
Dokuz Eylül University

TABLE OF CONTEXT

INDEX OF FIGURES

INDEX OF TABLES

LIST OF ABBREVIATIONS

(v:v)	Volume per Volume
.csv	Comma-Separated Values
µL	Microliter
µmol	Micromole
2.5D	2.5-Dimensional
2D	2-Dimensional
3D	3-Dimensional
ANOVA	Analysis of Variance
AQP5	Aquaporin-5
aSCs	Adult Stem Cells
B/-W	Better/(Worse)
BiGG	Biochemical Genetic and Genomic
BMP7	Bone Morphogenic Protein 7
BPCA	Bayesian Principal Component Analysis
BSA	Bovine Serum Albumin
cAMP	Adenosine 3',5'-Cyclic Monophosphate
cDNA	Complementary DNA
CK19	Cytokeratin 19
CK5	Cytokeratin 5
c-Myc	C-Myelocytomatosis Oncogene Product
CO2	Carbon Dioxide
CSF	Cerebrospinal Fluid
DAPI	4′,6-Diamidino-2-Phenylindole
DED	Dry Eye Disease
DMEM/F12	Dulbecco's Modified Eagle's Medium and Ham's F-12 Nutrient Mixture
DMSO	Dimethyl Sulfoxide
EBAM	Empirical Bayesion Analysis of Microarray
EDM	Eye Field Differentiation Medium
EGF	Epidermal Growth Factor
ESCs	Embryonic Stem Cells
FDR	False Discovery Rate
FGF	Fibroblast Growth Factor
FGF10	Fibroblast Growth Factor 10
Fisher's LSD	Fisher's Least Significant Difference
GC-MS	Gas Chromatography Mass Spectrometry
GSEA	Gene Set Enrichment Analysis
hiPSCs	Human Induced Pluripotent Stem Cells
HMDB	Human Metabolome Data Base

IBG	Izmir Biomedicine and Genome Center
IDs	Identifications
IF	Immunofluorescence
iPSCs	Induced Pluripotent Stem Cells
KEGG	Kyoto Encyclopedia of Genes and Genomes
Klf4	Kruppel-Like Factor 4
KNN	K-Nearest Neighbours
LC-MS	Liquid Chromatography Mass Spectrometry
LCN	Lipocalin
LFN	Lactoferrin
LG	Lacrimal Gland
log	Logarithmic
MECs	Myoepithelial Cells
MetaCyc	MetaCyc Metabolic Pathway Database
METLIN	Metabolite and Chemical Entity Database
Min	Minute
mL	Milliliter
mM	Millimolar
mmol	Millmole
mRNA	Messenger RNA
MS	Mass Spectrometry
MSD	Mass Spectrometer Dependent
MSEA	Metabolite Set Enrichment Analysis
MTG	Monothioglycerol
mz/rt	Mass Charge / Retention Time
MZOCs	Multi Zonal Ocular Cells
N/A	Non-Available
NaCl	Sodium Chloride
NEAA	Non-Essential Amino Acids
NetCDF	Network Common Data Form
NIST	National Institute of Standards and Technology
NMR/MS	Nuclear Magnetic Resonance / Mass Spectrometer
OCT	Optimum Cutting Temperature
Oct4	Octamer-Binding Transcription Factor 4
ODM	Ocular Cell Differentiation Medium
OPLS-DA	Orthogonal Partial Least Squares Differentiation Analysis
ORA	Over Representation Analysis
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline-Tween 20
PCA	Principal Component Analysis
PCs	Principal Components

PFA	Paraformaldehyde
PLS-DA	Partial Least Squares Differentiation Analysis
plsr	Partial Least-Squares Regression
PPCA	Probablistic Principal Component Analysis
PSCs	Pluripotent Stem Cells
QEA	Quantitative Enrichment Analysis
rpm	Revolutions Per Minute
RT	Room Temperature
SAM	Significance Analysis of Microarray
SMA	Smooth Muscle Actin
SNPs	Single Nucleotide Polymorphism
SOM	Self-Organizing Map
Sox2	SRY-Box Transcription Factor 2
SSP	Single Sample Profiling
SVD	Singular Value Decomposition
SVM	Support Vector Machine
Tukey's HSD	Tukey's Honestly Significant Difference
VIP	Variable Importance in Projection
WT-1	Wild-type 1

ACKNOWLEDGEMENTS

Firstly. I would like to thank my thesis supervisor, Assoc. Prof. Sinan Güven, for giving me opportunities to work on stem cells and allowing me to explore –omics studies. He gave me unending advice and encouragement as I pursued my master's degree. Also, I am truly thankful to Prof. Dr. Emirhan Nemutlu and Dr. Tuba Reçber for their guidance and assistance in metabolomics studies. Of course, I am very grateful and I owe a debt of gratitude to my previous professors and mentors especially, Asst. Prof. Ersin Akıncı, Prof. Dr. Nedim Mutlu and Asst. Prof. Münevver Aksoy for not only helping me get where I am now but also for teaching me how to do science and get ready for what lies ahead.

I would like to thank former and current members of Güven Lab for their support. My special thanks to Sude Uyulgan, Resul Özbilgiç and Tuğba Topbaş for their help during my master's studies. In addition, I would like to thank all members of IBG core facilities for providing every opportunity to us.

Ayşenur Öner, I couldn't be more grateful for having you in my life. Thank you so much for your unconditional love and support and for always being there for me. I wouldn't have gotten this far without you. So glad I have you.

I am thankful to all my friends, who have always stood by me, and shared their support, knowledge and guidance with me. Especially, Mehmet Yıldız, Pelin Ünal and Gamze Badakul for making me love science even more and showing me how it should be done. Also, thanks to Ümit Babacan and Oya Temiz for all the good laughs and for sharing enjoyable moments together.

Finally, I want to show my deepest gratitude to my family, Nazan Sarı, Tahsin Sarı and Sedat Sarı, for their unwavering support and encouragement throughout my life. I dedicate this thesis to my dear family, who provided me with the opportunity and support all the time.

METABOLOMICS ANALYSIS OF HUMAN INDUCED PLURIPOTENT STEM CELL DERIVED LACRIMAL GLANDS

Vedat Sarı, Izmir International Biomedicine and Genome Institute, Dokuz Eylül University, 15 Temmuz Sağlık ve Sanat Yerleşkesi, Balçova, 35340, Izmir/Türkiye

ABSTRACT

The lacrimal gland (LG), commonly known as the tear gland, regulates the ocular surface's homeostasis by secreting proteins and electrolytes that form the tear film's aqueous layer. It produces lacrimal fluid, which lubricates and protects the eye as the major exocrine gland of the eye. Loss of function brought on by LG damage has a negative impact on eye health and visual quality. Studies on human LG formation and development are few, despite the fact that a large percentage of the human population suffers from eye diseases caused by functional disorders in the LG, such as dry eye disease.

The LG's structure and morphology at the molecular level are not well known in the literature. In this study, human induced pluripotent stem cells (hiPSCs) were differentiated into cells with characteristics of LG *in vitro* by recapitulating functional and mature LG morphology at different time points. Differentiation was carried out according to the protocol developed by our lab in previous studies. Additionally, LG's reaction was monitored upon forskolin stimulation to understand secretion functionality better.

In this context, different maturation levels and secretion functionality of LG's upon forskolin stimulation were examined by untargeted gas chromatography-mass spectrometry (GC-MS) metabolomics. The results suggest that the metabolome of LG organoids changed significantly at different maturation levels and upon stimulation with different small molecules. In addition, different -omic analyzes can be performed on the produced LG organoids and can be integrated with other -omic analyzes to investigate LG organoids at the multi-omics level.

Keywords: Human induced pluripotent stem cells, stem cell differentiation, lacrimal gland, metabolomics, organoid

İNSAN İNDÜKLENMİŞ PLURİPOTENT KÖK HÜCRE KAYNAKLI LAKRİMAL BEZ ORGANOİDLERİNİN METABOLOMİK ANALİZİ

Vedat Sarı, İzmir Uluslararası Biyotıp ve Genom Enstitüsü, Dokuz Eylül Üniversitesi, 15 Temmuz Sağlık ve Sanat Yerleşkesi, Balçova, 35340, İzmir/Türkiye

ÖZET

Genellikle gözyaşı bezi olarak bilinen lakrimal bezi (LB), gözyaşı filminin sıvı tabakasını oluşturan proteinleri ve elektrolitleri salgılayarak oküler yüzeyin homeostazını düzenler. Gözün ana ekzokrin bezi olarak gözü kayganlaştıran ve koruyan gözyaşı sıvısı üretir. LB hasarının neden olduğu fonksiyon kaybı, göz sağlığı ve görme kalitesi üzerinde olumsuz bir etkiye sahiptir. İnsan nüfusunun büyük bir yüzdesinin, kuru göz hastalığı gibi LB'deki fonksiyonel bozuklukların neden olduğu göz hastalıklarından muzdarip olmasına rağmen, insan LB oluşumu ve gelişimi üzerine çalışmalar sınırlıdır.

LB'nin moleküler düzeyde yapısı ve morfolojisi literatürde iyi bilinmemektedir. Bu çalışmada, insan kaynaklı pluripotent kök hücreler (hiPSC'ler), farklı zaman noktalarında fonksiyonel ve olgun LG morfolojisini mimik etmek amacıyla *in vitro* LB özelliklerine sahip hücrelere farklılaştırılmıştır. Daha önceki çalışmalarda laboratuvarımız tarafından geliştirilen protokole göre farklılaştırma yapılmıştır. Ek olarak, salgılama işlevselliğini daha iyi anlamak için LB'nin tepkisi forskolin uyarımı yapılarak araştırıldı.

Bu bağlamda, LB'ler farklı olgunluk seviyelerinde ve salgılama işlevselliğini araştırmak için forskolin uyarımıyla hedeflenmemiş gaz kromatografi-kütle spektrometresi (GC-MS) metabolomiks analizler ile incelenmiştir. Sonuçlar, LB organoidlerinin metabolomunun, farklı olgunlaşma seviyelerinde ve farklı küçük moleküller ile uyarımı sonucunda önemli ölçüde değiştiğini göstermektedir. Ayrıca, üretilen LB organoidleri üzerinde farklı -omik analizler yapılabilir ve LB organoidlerini multi-omik düzeyde araştırmak için diğer omik analizler ile entegrasyonu yapılabilir.

Anahtar Kelimeler: İnsan uyarılmış pluripotent kök hücreleri, kök hücre farklılaşması, lakrimal bez, metabolomiks, organoid

1. INTRODUCTION AND GOALS

1.1. Statement and Importance of the Problem

The LG is a tubulo-acinar exocrine gland that secretes the aqueous layer of the tear film (Conrady et al., 2016). Through the fluid substances (water, electrolytes, and proteins) it generates, the LG has a significant impact on the condition of the eyes and the quality of vision. A tear film is produced, promoting and maintaining surface metabolism, corneal transparency, and the quality of vision by maintaining the secretion to the outermost layer of the eye, also known as the ocular surface. (Garg and Zhang, 2017; Örge et al., 2014; Zoukhri, 2010). The dynamic structure of the ocular surface and stabilization of the tear film are two of the most crucial components necessary for healthy vision. The visual sensory system may sustain long-term harm from damages that might compromise its framework (Lu et al., 2017). Despite the LG's crucial physiological activities, we know little about its development, and LG disorders, like DED, need effective therapeutics rather than only makeshift care (Celebi et al., 2014; Tai et al., 2002). The tremendous advancement of regenerative medicine has provided us with enormous potential for organ regeneration and transplantation treatment in the restoration of LG function (Yao and Zhang, 2017).

Regenerative medicine, the most current and expanding discipline of medical research, is concerned with the functional restoration of a particular tissue and/or organs in patients suffering from serious injuries or chronic disease states, when the body's natural regenerative responses are insufficient (Mahla, 2016; Mason and Dunnill, 2008). Functional LG is successfully produced by the organ germ method and transplanted to mice (Hirayama et al., 2013), and the regenerative potential of resident progenitor cells is investigated in a rabbit model with LG main excretory duct ligation-induced injury (Lin et al., 2016).

Organoids, which are miniaturized organs have the characteristics such as self-organization, multicellularity, and usefulness. Organoids of particular tissues may be grown using induced pluripotent stem cells (iPSCs), which have the capacity to develop

into cells from all three germ layers (Takahashi and Yamanaka, 2006; Yu et al., 2007; Thomson et al., 2011). Yet iPSCs have some drawbacks to using regenerative medicine such as tumorigenicity, immune rejection, genetic instability, developing and reproducing primary cultures, etc. (Curry et al., 2015). Organoid technologies have revolutionized how human development and diseases are studied and modeled *in vitro* during the last decade, and these technologies have the potential to overcome some of model organisms' limitations. (Kurmann et al., 2015; Huang et al., 2014; Takebe et al., 2013; Taguchi et al., 2014; Kim et al., 2020; Clevers, 2016). Also, human LG organoids are transplanted into the mouse and recapitulated *in vivo* LG (Bannier-Hélaouët et al., 2021).

Comprehensive metabolite profiling, also known as metabolomics, determines the chemical phenotype of human subjects and animal models. As a result, it has a unique potential for identifying biomarkers that predict disease incidence, severity, and progression (Ishikawa et al., 2016; López-López et al., 2018) as well as shedding new light on underlying mechanistic abnormalities (Newgard, 2017). Metabolomics can be defined as a comprehensive, simultaneous, and systematic identification of metabolite levels in cells, tissues, and/or organisms (Lindon et al., 2007; Hayashi et al., 2011). Metabolites are the byproducts of cellular regulation processes, and their concentrations may be thought of as biological systems' ultimate reactions to genetic or environmental changes (Fiehn, 2002). With over 50 years of defined techniques for metabolite studies, GC-MS is the most standardized approach in metabolomics for different compounds such as sugars, amino acids, sterols, hormones, aromatics, etc. (Fiehn, 2016). Developmental stages of human LG are defined (de la Cuadra-Blanco et al., 2003), however metabolome of LG still needs to be elucidated.

1.2. Purpose of the Research

The purpose of this research is to investigate metabolomic profiles of hiPSCs derived LG organoids using GC-MS based analytical tools. Here we investigate the metabolites obtained from cells and culture media of LG organoids at different time points. Alteration of metabolome of forskolin stimulated LG organoids was also demonstrated.

1.3. Hypothesis of the Research

In this work, we hypothesized that the maturation level and functionality of hiPSCs-derived LG organoids can be monitored through metabolomic analysis. We have analyzed LG organoids at different time points of *in vitro* culture and accessed their response upon chemical stimulation using GC-MS-based metabolomics.

2. GENERAL INFORMATION

2.1. Lacrimal Gland

The LG is an anatomically related tubulo-acinar exocrine gland to the mammary and salivary glands. (Makarenkova and Dartt, 2015; Garg and Zhang, 2017) The LG, which is situated on the top of the eye, creates the aqueous layer of the tear film and covers the cornea and conjunctiva that make up the ocular surface (Figure 1). By causing tears produced, LG keeps the ocular surface microenvironment and physiological function in a state of balance (Hirayama et al., 2016). Tears primarily remove waste materials from corneal and conjunctival cells while transferring oxygen and other nutrients to these cells. Therefore, LG is crucial to maintaining good eye health and vision.

By releasing immunoglobulins, the tear film lubricates the ocular surface while maintaining metabolism, creates a smooth layer over the cornea for appropriate light reflection, and protects the eye against external contaminants and any infections (Garg and Zhang, 2017; Holly and Lemp, 1977; Johnson and Murphy, 2004). Three layers compose this film: the mucous layer, which covers the corneal epithelium, the aqueous layer, which accounts for more than 90% of the tear film's volume, and the lipid layer, which is the outermost layer.

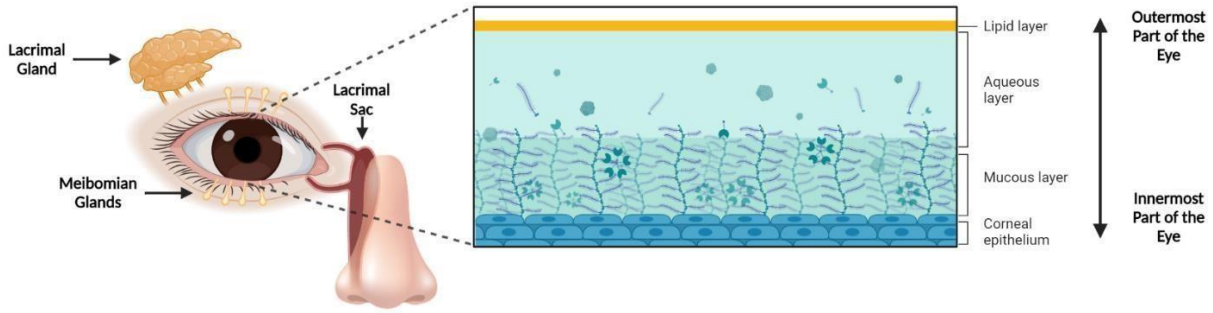


Figure 1. Schematic illustration of ocular glands and structure of tear film.

Goblet cells from the conjunctiva and stratified squamous cells from the cornea generate the mucous layer. This layer connects the ocular surface to the tear film, which has mucins that work to stop bacterial invasion and even out any irregularities in the cornea. The Meibomian glands on the upper and lower eyelids secrete the lipid layer, which maintains the surface's lubricity and stops tears from evaporating (Downie et al., 2021; Makarenkova and Dartt, 2015).

The human LG is found in the upper temporal orbit and mostly secretes water, electrolytes, and a small number of proteins that are essential for eye health (Walcott, 1998; Zoukhri et al., 2008). It is around 20x12x5 mm in length, width, and lobe size.

Acinar, ductal, and myoepithelial cells form the majority of the LG epithelium (Figure 2). Acinar cells, which make up the main secretory mechanism, make up about 80% of the gland's content. The proteins, electrolytes, and water that make up the tear are synthesized and secreted by these cells. The cuboidal duct cells, which make up 10–12% of the LG cell population and provide 30% of the LG secretions, surround the acinar cells (Mircheff, 1989). Unidirectional protein secretion from acinar cells is made possible by the close connections that surround the cells at the lumen, dividing the plasma membrane into apical and basal membranes. The intensely innervated the LG releases the transmitters and peptides near the acinar cell membranes (Paulsen and Berry, 2006). Similar to acinar cells, ductal cells form polarized cells and, like acinar cells, are joined to the apical side via tight junctions. The electrolyte balance is maintained before the tear exits by ductal cells as they transport the tear from the acini to the ocular surface. Although

they have fewer granules than acinar cells, duct cells produce proteins via a similar process.

The acinar and ductal cells are surrounded by myoepithelial cells (MECs), which are in charge of preserving the cellular structure and producing the basal membrane (Downie et al., 2021). Ion channels and transport proteins for electrolyte and water secretion are also present in the basal membranes along with neurotransmitters and neuropeptides that stimulate secretion. On their basal side, MECs carry out several functions, such as contracting the acinar and ductal cells that surround them in order to facilitate the tear secretion process (Hodges and Dartt, 2003). Their main function is to exert pressure in order to facilitate the passage of LG fluid from secretory cells into the duct that has smooth muscle actin (-SMA) on the basal side (Dartt, 2009; Hodges and Dartt, 2003). The delivery granules that convey LG secretory proteins fuse at the apical membrane, which also houses a separate set of ion channels and transport proteins.

The LG comprises fibroblasts, lymphocytes, plasma cells, macrophages, and mast cells that secrete histamine and matrix proteins in addition to the acinar, ductal, and MECs. Thus, the ocular surface's immunological defense can be preserved. According to research, the bulk of the mononuclear cells in the LG are plasma cells, which make up

more than half of the immune system of the ocular surface's immune system and are often immunoglobulin A positive (Wieczorek et al., 1988).

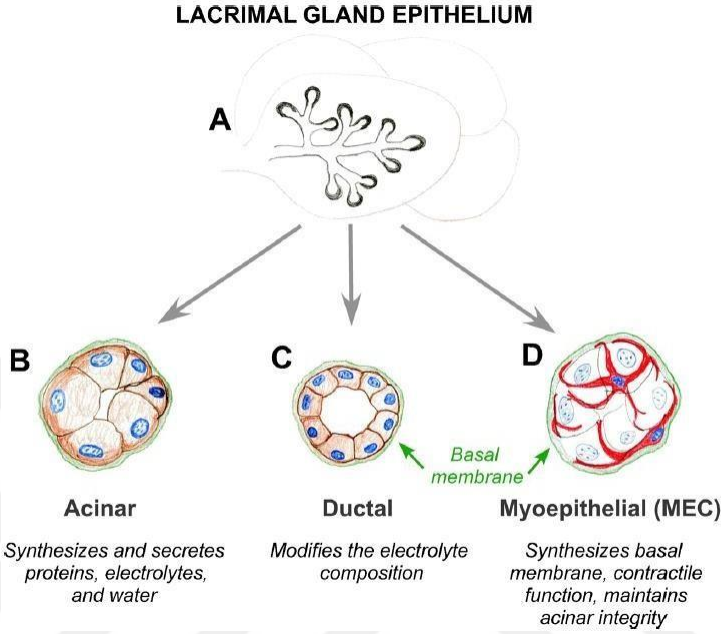


Figure 2. Cell types of lacrimal gland epithelium.

A) The LG epithelium is consisting of, **B)** Acinar cells; **C)** Ductal cells; and **D)** Myoepithelial cells (MECs) (red) (Makarenkova and Dartt, 2015).

2.2. Induced Pluripotent Stem Cells

The essential components of life, stem cells are also known as undifferentiated cells because of their capability for self-renewal and their propensity to develop into any form of the cell (potency) in our body. The developmental potency of stem cells decreases at each stage, and these properties differ across stem cells. The zygote, which has the highest potential for differentiation and totipotency, is the very first cell to develop following fertilization. Cells might continue to differentiate at this time as they work to develop the placenta. The inner cell mass of the blastocyst, which develops via the creation of three germ layers, a structure generated by pluripotent stem cells, contains embryonic stem cells (ESCs) (PSCs). PSCs can develop into all three germ layers, but not into placental cells (Chagastelles and Nardi, 2011; Kolios and Moodley, 2013).

With their ability to differentiate, ESCs have emerged as the cornerstone for research of cell-based therapies for human disorders (Ramirez et al., 2010). However, there are significant limitations to using ESCs for therapeutic purposes, including moral dilemmas surrounding the killing of human embryos and the difficulty in locating HLA-compatible ESCs for use in medical procedures (Bai et al., 2013; Landry and Zucker, 2004).

Although pluripotent cells can only develop spontaneously in blastocysts, it is now feasible to reprogram differentiated cells such that they return to their pluripotent condition. Yamanaka, who won the Nobel Prize, and his colleagues initially described the reprogramming technique in 2006, dedifferentiating mouse fibroblast cells into ESCs. By introducing ectopic expression of the four genes Oct4, Sox2, Klf4, and c-Myc, they were able to induce reprogramming (Takahashi and Yamanaka, 2006). Later, using the same components, they were able to effectively produce hiPSCs from adult human dermal fibroblasts (Takahashi et al., 2007a). In terms of morphology, proliferation and *in vitro* differentiation patterns, gene expression, epigenetic state, telomerase activity, and teratoma development, the created iPSCs are identical to natural pluripotent ECSs (Chin et al., 2010; Doi et al., 2009; Guenther et al., 2010; Mikkelsen et al., 2008; Takahashi et al., 2007b).

There are several ways to convert somatic cells into iPSCs. Retroviruses are frequently utilized to integrate foreign genes into the genome of somatic cells. The nature of the cell changes depending on the conditions once the virus has integrated the cell genome. Adenovirus and Sendai virus, however, are also utilized in some research as non-integrating choices due to their dangers (Ye et al., 2013). Additionally, non-viral infection techniques devoid of genetic integration have been documented. For this, a variety of non-viral techniques have been devised, including transfecting cells with plasmids encoding the factors' cDNAs, repeatedly transfecting cells with non-viral minicircle DNA vectors, or using synthesized mRNA to express the factors (Narsinh et al., 2011; Okita et al., 2008; Warren et al., 2010).

Intense efforts are being undertaken to increase reprogramming's effectiveness and produce iPS cells with as few genomic modifications as feasible. The origin tissue of iPS cells affects reprogramming effectiveness in addition to reprogramming vectors and factors.

2.3. Tissue Engineering and Organoid Models as Therapeutic Approaches

The primary objective of tissue engineering is to develop and produce substitute organs or tissues that may be transplanted under the proper circumstances to hasten the healing and regeneration of the wounded organ or area (Langer and Vacanti, 1993). To further advance the tissue engineering techniques, several investigations have been carried out on animal models *in vivo*, including but not restricted to Drosophila, rat, or zebrafish models, or at the cellular level on cell culture models *in vitro*. However, these investigations may have had limited practical applicability due to the variances across species. *In vitro* models frequently fail to reproduce organ-level functioning, which is necessary for accurately simulating illness. Consequently, more accurate models are needed.

Organoid technologies have been created to meet this demand. Organoids are three-dimensional (3D) structures created from stem cells that mimic the structural and functional complexity of biological organs by self-organizing from cell types that are unique to different organs. As a result, modern organoid technology may "recapitulate in a dish" the development of human organs as well as many human illnesses. Organ-specific adult stem cells, induced pluripotent stem cells, and embryonic stem cells can all be used to create organoids (aSCs). As a result, the creation of organoid models benefits from the seemingly endless capacity for development in the culture of normal stem cells. In order to stimulate the fetal environment and enhance organoid maturation, the differentiation of stem cells can be affected by the addition of biological substances such as exogenous growth factors, ECM substrates, or small compounds (Clevers, 2016). The created models may then be applied to several research projects, including illness modeling, medication testing, and organ replacement (Figure 3).

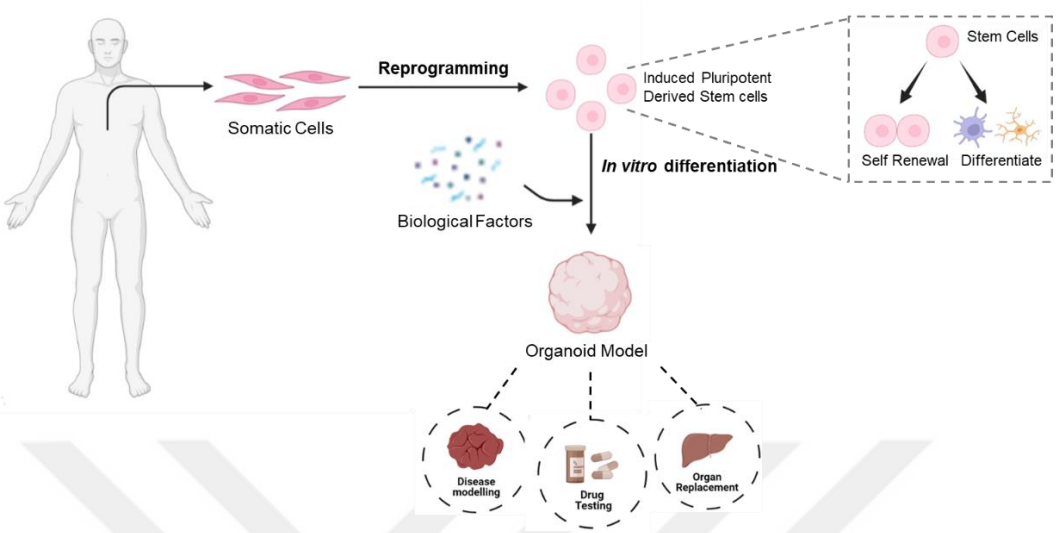


Figure 3. Human-induced pluripotent stem cell-derived organoid models.

2.3.1. Lacrimal Gland Organoid Models

Like the other secretory glands, LG can regenerate itself after being injured. Lifelong regeneration ability is maintained and stem cells are in charge of this (Dietrich and Schrader, 2020; Zoukhri et al., 2008). Stem cells have been used in numerous research in recent years to promote tissue regeneration following damage and to restore the LG's functionality. Regenerative medicine has improved steadily in ophthalmology to treat eye conditions that might impair eyesight.

In order to rejuvenate LG, bioengineered LG germs have recently been created utilizing embryonic stem cells. To recreate LG germs *in vitro*, epithelial and mesenchymal cells from mouse LG progenitors are gathered and put in a collagen gel matrix. With this technique, they were able to successfully mature the LG germs, mimic normal tissue connections between the epithelium and the mesenchyme, and demonstrate tear secretion function following *in vivo* engraftment into an aqueous deficient animal (Hirayama et al., 2013; Nakao et al., 2007).

In a different study that focused on the development of human LG organoids, LG cells were taken from the patients and digested before being cultivated in 3D with the aid of biological agents including epidermal growth factor (EGF) and fibroblast growth factor (FGF). Under certain conditions, they were able to create human microspheres, free-

floating spheres with acinar and ductal cells as well as measurable amounts of tear proteins (Tiwari et al., 2018). This work revealed encouraging results for the potential of LG with a secretory function in 3D culture. Furthermore, earlier research has produced intricate secretory gland architectures enabling a successful 3D restoration of LG using biohybrid materials such as decellularized scaffolds (Lin et al., 2016; Spaniol et al., 2015).

In one of the most recent research, cells were taken from human LG tissues of both healthy and Sjogren's disease patients, and they were then mixed with Matrigel and cultivated in various growth media to promote the creation of organoids. Organoids that had been created had their protein secretion and gene expression analyzed. As a consequence, both the healthy and sick organoids displayed characteristics that were histologically and functionally identical to, respectively, normal and diseased tissues (Jeong et al., 2021). Additionally, a study was recently conducted on the induction of auxiliary molecules to produce 3D LG organoids from various iPSCs. It has been demonstrated that the hiPSCs can develop into LG organoids that resemble native LGs (Hayashi et al., 2022).

The creation of organ-like structures in regenerative medicine and tissue engineering has increased with the introduction of induced pluripotent stem cells. As a result, there is growing interest in creating *in vitro* 3D culture models utilizing iPSCs to get around the drawbacks of existing methods.

2.4. Chemical Stimulation of Secretory Cells

Numerous studies have shown that forskolin has the unique ability to activate adenylyl cyclase directly, quickly, and irreversibly. This increases the level of intracellular cyclic 3′, 5′-adenosine monophosphate (cAMP) in a wide range of mammalian membranes, broken cell preparations, and intact tissues. The unique adenylyl cyclase activation of forskolin, one of the main components, is what underlies its wide range of pharmacological properties. Forskolin is a component of several patented pharmaceutical formulations that are sold over-the-counter to treat a variety of illnesses (Seamon et al., 1981; Fradkin et al., 1982; Alabashi and Melzig, 2010). Forskolin was reported to stimulate

the production of cAMP in rabbit corneal epithelial cells that were cultured, to increase the concentration of cAMP in the aqueous humor by ten times, to cause a sustained decrease in intraocular pressure, and to reduce the production of aqueous humor (Bartels et al., 1987). Additionally, a recent study shows a similar trend. Upon forskolin stimulation, the fluid secretion capacity of LG duct cells increases in rabbits (Katona et al., 2014).

2.5. Metabolomics

The completion of the Human Genome Project and the study of system biology, including genomics, transcriptomics, proteomics, and metabolomics, opened the door for a thorough molecular investigation of disorders (Karczewski and Synder, 2018; Jones, 2001). System biology research may now use omics technology to get relevant findings (Kitano, 2002; Chuang et al., 2010; Nemutlu et al., 2012). Utilizing computational and mathematical modeling of complex biological systems to comprehend whole-system biology is made possible by omics technology (All systems go!, 2006).

In terms of representing phenotypes, metabolomics is regarded as the best omics method (Jewett et al., 2006; Dunn et al., 2011) This is due to the fact that metabolomics analysis may offer information on the level of metabolic activity within a living thing (Rochfort, 2005). For the phenotypic study, comprehensive and quantitative analysis of all metabolites utilizing the metabolomics approach is strongly advised (Roessner and Bowne, 2009). In a nutshell, the goal of metabolomics analysis is to find metabolites that have been significantly changed due to a condition, medication, lifestyle choice, microbiota, etc. Studies on metabolomics are more often utilized to identify biomarkers, diagnose disorders, and assess how therapies are responding (Dunn et al., 2011).

To determine the metabolomic profile of the phenotype, metabolomics investigations analyze a large number of metabolites in great detail at the same time. Changes in metabolite levels can be attributed to genetic or environmental causes, as well as illnesses, changes in the microbiota, medications, toxins, and cellular and systemic changes in lifestyle (Beckonert et al., 2007; Brown et al., 2009; Lewis et al., 2008; Wishart et al., 2009; Kaddurah-Daouk et al., 2008; Nemutlu et al., 2012). The analytical systems

to be employed for metabolomics investigations should be capable of analyzing hundreds of metabolites from complicated biological samples at once and enabling the monitoring of changes in these metabolites (Wishart et al., 2009; Lindon and Nicholson, 2008; Lanza et al., 2010; Dunn, 2008; Nemutlu et al., 2015). The physical variety of these metabolites, including hydrophilic carbohydrates, volatile alcohols and ketones, amino and non-amino organic acids, and hydrophobic lipids, may explain why none of the analytical platforms now in use can properly assess the whole metabolome. Therefore, a combination of several analytical methods is used to thoroughly evaluate the cellular metabolome (Brown et al., 2009; Villas-Boas and Bruheim, 2007; Weckwerth, 2007). However, the majority of the currently known metabolomics research only used one analytical technique (Cansın Zeki et al., 2020).

Metabolomic research also uses other methods, such as microfluidic platforms. Cells may be grown, worked with, and lysed using extremely effective and regulated microfluidic devices. Though it is regarded as a powerful enabling technology to study the inherent complexity of cellular systems, this approach is used less frequently than other techniques. Precision fluid control, minimal sample consumption, instrument miniaturization, cheap analytical costs, and simple processing of nanoliter quantities are all features of microfluidic systems. Additionally, the possibility of contamination is decreased by the capability to process cells in a closed microfluidic system (Lin and Lin, 2015; Wheeler et al., 2003; Kraly et al., 2009). Microfluidic chips can be utilized to quickly and cheaply carry out the time-consuming, costly, and labor-intensive extraction and derivatization procedures (Jonsson et al., 2004) that are typically employed in the sample preparation stage of metabolomics research. However, some of the microfluidic chip designs are difficult for lab workers to fabricate and their use in metabolomics research is rather challenging. This makes it more difficult to apply microfluidic methods widely in metabolomics research.

2.5.1.1. *Metabolomic Analysis*

Sample preparation, metabolite analysis, and data analysis are the three primary phases of metabolomics analysis (Figure 4). The identification and measurement of

metabolites are part of the metabolomic approach. Targeted analysis and untargeted analysis are the two main metabolomics methods (Fiehn, 2002). In terms of sample collection, preparation procedures, and chromatographic settings, targeted and untargeted metabolomics studies are comparable (stationary and mobile phases). Untargeted research concentrates on repeatability, while targeted studies primarily focus on accuracy (Christians et al., 2011). To better comprehend changes in phenotypic metabolic profiles, both GC-MS and liquid chromatography mass spectrometry (LC-MS) techniques can be used together in metabolomics investigations (Cansın Zeki et al., 2020). Sample size and sample selection are crucial in metabolomics investigations in order to obtain significant results. To decide on the sample size, power analysis must be done at the outset of the investigation. Power analysis is based on changes in the analysis and the prevalence of the disease in the community (Blaise et al., 2016). This procedure improves the study's data dependability and quality. Every metabolomics study requires a certain minimum number of samples, and this number rises as the subject becomes more complicated. 3-6 biological duplicates are employed for cell culture research, 10-20 biological replicates for animal investigations, and 20-30 biological replicates for human studies in order to remove variances resulting from complex system biology during integrated metabolomics analysis (Nyamundanda et al., 2013).

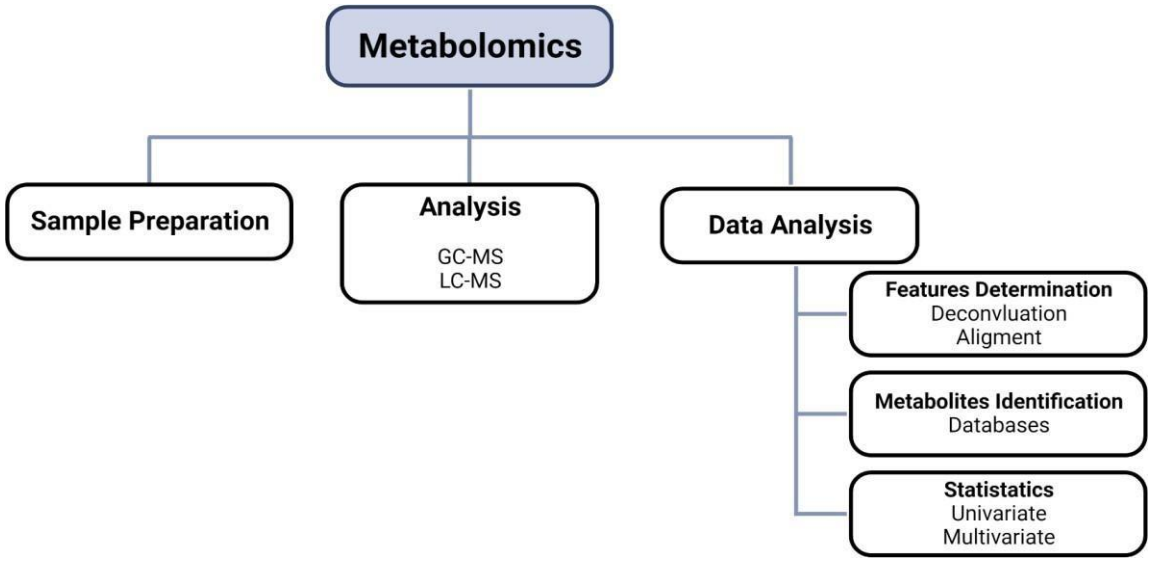


Figure 4. Schematic representation of metabolomics workflow.

Sample preparation, data acquisition and data analysis are the three main steps of metabolomics analyses.

2.5.2. Sample Preparation for Metabolomics

The simple but crucial process of sample preparation is important for the accuracy and quality of the results in a metabolomics study. The kind of metabolite to be examined and the analytical technique affect it (Kim and Verpoorte, 2010). Regardless of the analytical techniques employed, the majority of sample preparation processes are similar.

The main causes of the greatest differences in metabolomics investigations are sample collection and storage. Throughout the course of the research, samples must be collected in the same method, and they must be kept in the same manner before analysis. Because temperature has an impact on the metabolites' stability, sample storage is crucial. While they are only temporarily stable at room temperature (RT), metabolites in plasma and urine samples can stay stable for 30 months at 80 °C and 26 weeks at 25 °C, respectively (Lauridsen et al., 2007; Pinto et al., 2014). To reduce changes in metabolite levels, samples must be kept as cold as possible throughout storage.

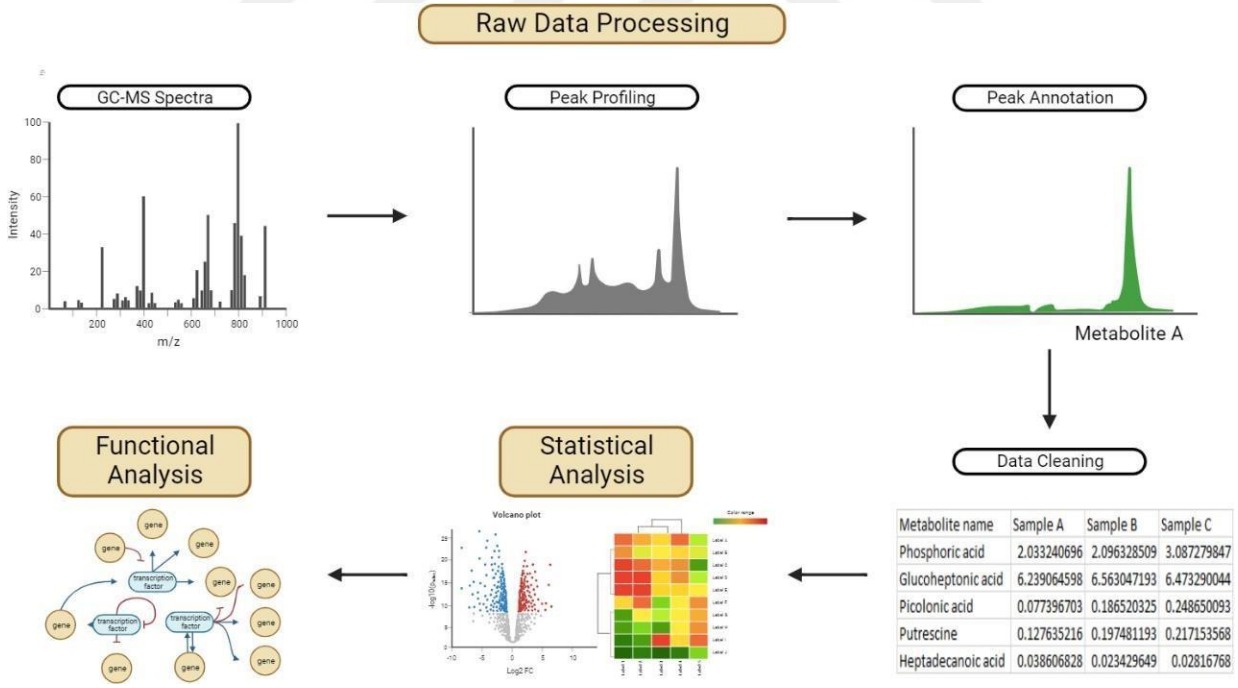
To preserve analyte yield, samples must be carefully collected and then extracted. Additionally, this will improve the efficiency of the instrument and lessen the matrix impact in the experiments. The three steps involved in sample preparation for a metabolomics study are sample collection, metabolite extraction (e.g., protein precipitation, liquid-liquid extraction (Mu et al., 2019)), and extract preparation (e.g., concentration, derivatization, or reconstitution with mobile phase) for the analytical device. Different biological materials (blood, urine, tissue, saliva, cells, plants, etc.) from human and animal origin can be used in metabolomics investigations.

In untargeted metabolomics analysis using GC-MS, the sample preparation phase is divided into two steps: extraction and derivatization. To sustain analyte yield, samples should be adequately collected and extracted, as previously mentioned. Then, unless the metabolites are volatile, all samples, regardless of sample origin, must be derivatized for GC-MS analysis. Derivatization is done to keep molecules from degrading at high

temperatures and to make metabolites with polar functional groups like carboxylic acid, ketone, and alcohol more volatile. Derivatization enhances detector response, sensitivity, thermal stability, and volatility. However, this extra step is one of the main causes of variance in GC-MS analysis and it has an impact on the study's repeatability.

2.5.3. Data Analysis

Applying data processing techniques will assist in comprehending complicated chromatograms. In metabolomics studies, data analysis can be conducted in three steps. The determination of mass features, such as deconvolution and data alignment, comes first. Peak annotation utilizing reference libraries comes next. Univariate and multivariate statistical analysis is used in the third and final step to detect significantly altered metabolites based on phenotype. With the exception of the software and libraries utilized, these processes are identical for GC-MS metabolomics data (Figure 5).



Metabolite name	Sample A	Sample B	Sample C
Phosphoric acid	2.033240696	2.096328509	3.087279847
Glucoheptonic acid	6.239064598	6.563047193	6.473290044
Picolonic acid	0.077396703	0.186520325	0.248650093
Putrescine	0.127635216	0.197481193	0.217153568
Heptadecanoic acid	0.038606828	0.023429649	0.02816768

Figure 5. Workflow of data analysis in metabolomics.

2.5.3.1. *Data Preprocessing and Determination of Mass Features*

The first step to successful data analysis and the production of biologically important results is reliable and efficient data processing. Retention time, m/z values, and density or abundance on each axis are all included in the complicated three-dimensional data format of the raw data files obtained with the GC-MS and LC-MS systems (Mastrangelo et al., 2015). Before using any statistical approaches to analyze the data, this extensive information must be processed. Data processing involves a number of steps, one of which is converting the instrument's raw data into a two-dimensional data matrix (Tian et al., 2017).

2.5.3.2. *Metabolite Annotations*

To characterize and understand the masses acquired in metabolomics research, relevant databases must be used. There are several public databases that may be used to annotate metabolites. Some of the databases that are often utilized are HMDB, METLIN, Fiehn GC-MS Database, Golm Metabolome Database, MetaCyc, and NIST (Vinaixa et al., 2016). In metabolomics research, understanding the retention time and mass of metabolites in untargeted analyses is critical. Compared to LC-MS, the GC-MS approach has far richer libraries and has become the standard for metabolomics investigations because of benefits including excellent chromatographic efficiency and reproducibility. Since hundreds of metabolites must be identified in a single LC-MS chromatogram, identifying metabolites in LC-MS, unlike GC-MS, is the most difficult step in metabolomics studies. The closed formula of each metabolite can be determined with considerable support from high-resolution MS, but the chemical make-up of these metabolites cannot be determined with great accuracy.

2.5.3.3. *Data Normalization*

One sample is saved per row and one variable (bin/peak/metabolite) per column in the data, which is organized as a table. The four categories of normalization techniques used below are listed. Data transformation and scaling are two different ways to make features

more comparable. Row-wise normalization allows general-purpose transformation for variations between samples; sample-specific normalization allows users to manually adjust concentrations based on biological inputs (e.g., volume, mass).

The normalization consists of the following options:

1. Row-wise procedures:
    - Sample-specific normalization (i.e. normalize by dry weight, volume)
    - Normalization by the sum
    - Normalization by the sample median
    - Normalization by a reference sample (probabilistic quotient normalization)
    - Normalization by a pooled or average sample from a particular group
    - Normalization by a reference feature (i.e. creatinine, internal control) Quantile normalization
2. Data transformation:
    - Log transformation (base 10)
    - Square root transformation Cube root transformation
3. Data scaling:
    - Mean-centering (mean-centered only)
    - Auto-scaling (mean-centered and divided by the standard deviation of each variable)
    - Pareto scaling (mean-centered and divided by the square root of the standard deviation of each variable)
    - Range scaling (mean-centered and divided by the value range of each variable)

2.5.3.4. *Statistical Analysis*

Correlation analysis, multivariate analysis, and univariate analysis are used to assess data from metabolomics investigations. Achieving the desired result using univariate analysis may result in time loss and errors, especially when taking into account the number of metabolites and variances from outlier samples in the data. As a result, multivariate analyses rather than univariate analyses are frequently utilized in metabolomics studies. By using a transformation technique called multivariate data

analysis, the size of a data collection with many variables may be decreased while maintaining the data's essential variances. With the least amount of variables possible, these techniques aim to express the supplied data. As a result, the data may be rapidly viewed and assessed with a more comprehensive understanding (Sweetlove and Fernie, 2013). In this context, metabolomic data analysis applies principal component analysis (PCA) and partial least squares-discrimination analysis (PLS-DA). PCA analysis is frequently used to identify metabolites that effectively distinguish across groups. Other multivariate analyses exist as well, including cluster analysis and orthogonal PLS-DA (OPLS-DA). Univariate analysis can be used to locate substantially altered metabolites after multivariate analysis has been used to evaluate the data. Depending on how many samples are in the data set, a parametric or non-parametric test is applied. Additionally, special focus must be given to the sample type to ascertain if they are paired or unpaired samples as in time series analysis (Doerfler et al., 2013).

2.5.3.5. *MSEA Background and Overview*

Quantitative metabolomic data may be used to find physiologically relevant patterns that are markedly enriched using metabolite set enrichment analysis (MSEA). In traditional methods, each metabolite is assessed for its relevance to the study's conditions. The molecules that have reached a particular threshold for significance are then mixed to determine if any significant patterns emerge. MSEA, on the other hand, does not need the preselection of molecules based on some arbitrary cut-off value; instead, it immediately examines whether a group of functionally related metabolites. It may be able to spot minor but persistent changes among a collection of connected chemicals that the traditional methods could miss.

MSEA is essentially a metabolomic adaptation of the well-known gene set enrichment analysis (GSEA) program, complete with a library of metabolite sets and user-friendly online interfaces. GSEA is often employed in the analysis of genomics data and has established itself as a potent substitute for traditional methods (Submarian, 2005; Nam and Kim, 2008).

Data input, data processing, data analysis, and results download are the four phases that make up a metabolite set enrichment analysis. Various analytical techniques are applied based on various input kinds. Users can upload their own self-defined metabolite sets for enrichment analysis in addition to browsing and searching the libraries of metabolite sets. Additionally, users have the option of performing metabolite name mapping between several chemical names, synonyms, and significant database IDs.

3. MATERIALS AND METHODS

3.1. Type of Research

This study is an *in vitro* experimental study.

3.2. Date and Location of Research

The whole research was conducted between 2020 November – 2021 June in Izmir Biomedicine and Genome Center (IBG) and Hacettepe University, Faculty of Pharmacy Department of Analytical Chemistry.

3.3. Universe and Sample of Research

In this thesis scope, primary human samples were not used.

3.4. Materials of Research

Wild-type hiPSCs are reprogrammed in the laboratory of Assoc. Prof. Tamer Önder at Koç University (Istanbul, Turkey) was obtained from Prof. Dr. Esra Erdal at IBG Center (Izmir, Turkey).

3.5. Variables of the Research

Independent variables: Matrigel® concentration, differentiation medium components, duration of differentiation culture, metabolomics sample preparation components, data acquisition protocol

Dependent variables: Heterogeneity of the differentiated cells, ability to be subcultured, cell morphology, metabolite levels

3.6. Data Collection Methods

3.6.1.1. *Human-Induced Pluripotent Stem Cells Culture*

WT1 hiPSCs were grown in complete mTeSR1 media (STEMCELL Technologies, 85851) at 37 °C, 5% CO2 as described by Akbari et al (Akbari et al., 2019). The medium was prepared by combining the mTESR basal medium in a 4:1 (v:v) ratio with its 5X supplement (STEMCELL Technologies, 85852). The media was changed daily to maintain the pluripotent state, and cells were passaged every seven days or lesser, depending on their confluency. Bright-field microscope was used to monitor the quality of hiPSC colonies throughout time. During the culture period, the differentiated colonies were identified by their morphologies, and then they were scraped from the environment using sterile pipette tips. On 6-well sterile culture plates, cells were plated on 1% Matrigel (Corning, 354277) coated wells (Figure 6).

For each well of the six-well plates, 12–13 µL of Matrigel (varies according to the lot) was mixed with 1 mL of serum-free Dulbecco Modified Eagle's Medium and Ham's F-12 Nutrient Mixture (DMEM/F12) (Gibco, Thermofisher Scientific, 31330) and placed on ice. The Matrigel was kept on ice during the coating preparation process since it gels with temperature and thaws at 4 °C. As soon as the 1% Matrigel solution was ready, 1 mL was immediately poured into each well of the 6-well plate. The plate was incubated at 37 °C for 1 hour following the transfer of the mixture to the wells. The plate was used to cultivate cells after an hour of incubation.

The differentiated cells were first scraped out of the culture and then the media was aspirated to obtain passaged hiPSCs. A 1mL DMEM/F12 wash was used to wash the well. The empty well was filled with 1 mL of ReLeSR (STEMCELL Technologies, 05872), which was then incubated at RT for 1 minute. After removing the ReLeSR, the empty well was incubated at 37°C for 4 minutes. In the meanwhile, the pre-matrigel coated well's 1% Matrigel solution was aspirated out and replaced with 1 mL of fresh mTeSR. Cells were taken from the hiPSC-containing well using a 2 mL serological pipette after the incubation time, homogenized, and dissociated by pipetting while being careful not to

dissociate colonies into single cells. Cells were collected using 1-2 mL (depending on the confluency) mTeSR. Replated cells were then added to the Matrigel-coated mTeSR-containing wells in drops (5-10 µL in volume), and the plate was incubated at 37°C with 5% CO2.

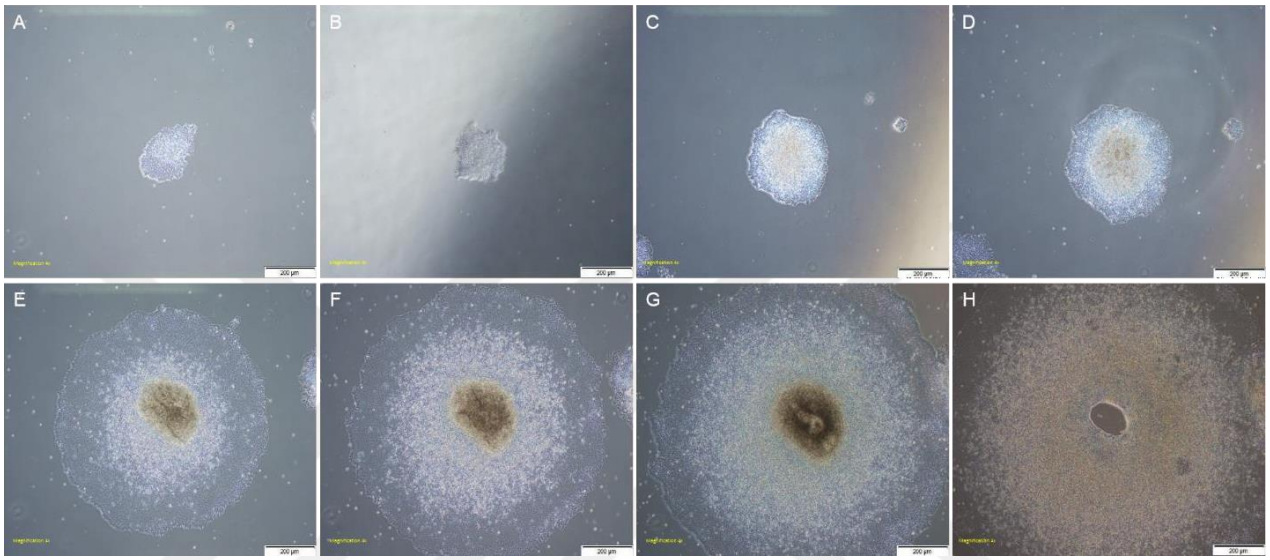


Figure 6. Bright-field images of hiPSCs culture.

A) Bright-field image of hiPSCs on day 1. Scale bar: 200 µm. **B)** Bright-field image of hiPSCs on day 2. Scale bar: 200 µm. **C)** Bright-field image of hiPSCs on day 3. Scale bar: 200 µm. **D)** Bright-field image of hiPSCs on day 4. Scale bar: 200 µm. **E)** Bright-field image of hiPSCs on day 5. Scale bar: 200 µm. **F)** Bright-field image of hiPSCs on day 6. Scale bar: 200 µm. **G)** Bright-field image of hiPSCs on day 7. Scale bar: 200 µm. **H)** Bright-field image of hiPSCs on day 8. Scale bar: 200 µm.

3.6.1.2. *hiPSCs Differentiation into Multi Zonal Ocular Cells (MZOCs)*

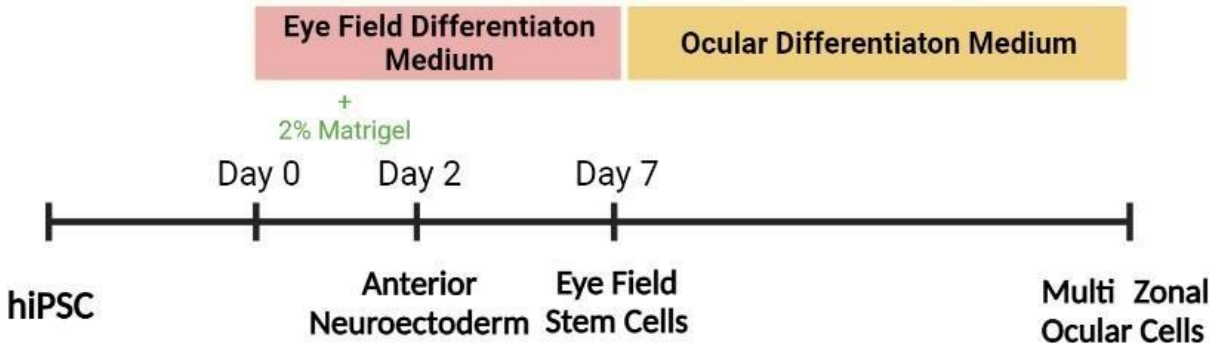


Figure 7. Lacrimal gland differentiation of human induced pluripotent stem cells protocol.

By modifying the Li et al. published protocol, multizonal ocular cells were differentiated from WT1 hiPSCs (Li et al., 2019). hiPSCs were first differentiated into anterior neuroectoderm cells by day 2 and eye field stem cells by day 7 using this protocol. They were subsequently differentiated and matured into multi-zonal ocular cells in about 21 days (Figure 7). A 1:1 (v:v) mixture of DMEM/F-12 and neurobasal medium (Gibco, 21103049) was used as the eye field differentiation medium (EDM), which was additionally supplemented with 2 mM L-glutamine (Gibco, 25030081), 0.1 mM non-essential amino acids (NEAA; Lonza BioWhittaker, 13-114E), 0.1 mM monothioglycerol (MTG; (R&D Systems, AAR009). Ocular differentiation media (ODM) was composed of DMEM/F-12 with 10% knockout serum replacement (Gibco, 10828-028), 2 mM L-Glutamax, 0.1 mM NEAA, and 0.1 mM MTG added as supplements.

Experiment plates were covered with 1% Matrigel before cell seeding, as explained in section 3.6.1.1. hiPSCs that had been in culture for a week were inspected under a microscope, and differentiated cells were marked and taken out of the culture. Ice-cold DMEM/F-12 was used to wash the remaining colonies. The Matrigel is easier to dissolve in cold media. ReLeSR was added, and 1 min was given for it to sit at RT. The empty plate was incubated at 37 °C for 4 minutes after ReLeSR aspiration. A 2 mL serological pipette was used to collect the colonies while being careful not to disturb them. A drop of cells grown in 2% Matrigel containing EDM after being seeded onto wells with pre-matrigel coating. The media was replaced the following day with a new 2% Matrigel mixed with EDM, allowing the 2.5D colonies to develop for 2 days on the 2% Matrigel medium combination. After that, 2% Matrigel was no longer needed, and media was replaced for a further 5 days with fresh EDM. The culture medium was then changed to ODM and renewed daily, except for one weekday when the culture was double-fed (2-3 mL per well of a 6-well culture plate). The culture was expanded as necessary (maximum 45 days in this study).

3.6.2. Cell Characterization

Protein expression levels were used to identify differentiated cells. To detect protein expressions, immunofluorescence staining (IF) was performed. Samples were initially

OCT-embedded to prepare them for IF staining, and then sections were cut using a cryostat.

3.6.2.1. *Immunofluorescent Staining*

After the media was aspirated from the culture plates, the cells that would be stained were washed three times in PBS. Fixation was performed by incubating cells in 4% PFA (Sigma Aldrich) for 20 to 30 minutes at RT. Following the incubation, the PFA solution was aspirated, and three PBS washes were performed on the cells. The fixed sample was placed directly into a freshly made 30% sucrose solution and incubated at 4 °C until the sample sank to the bottom of the tube/well. Alternatively, the fixated sample was incubated with fresh PBS at 4 °C until subsequent operations.

The sample was transferred in Optimal Cutting Temperature (OCT; Sakura; 4583) when it sank, and after embedding, it was instantly frozen by submerging it in liquid nitrogen. Until cryosectioning, samples implanted in OCT were kept at -80 °C. The cryostat was used to cut 7 µm pieces, which were then mounted on polylysine microscope slides (Thermo-scientific). Slides were maintained at -80 °C until staining.

Before the experiment began, all buffers were warmed to RT. Before staining, slides were defrosted at RT for 20 minutes. A Pap pen was used to label and around the samples. The samples were permeabilized by incubating them for 15 mins at RT in a permeabilization buffer containing 0.3% Triton X-100 in PBS (v/v). The samples were washed three times for five minutes with PBS-T containing 0.05% Tween20. The blocking buffer, which contains 1% Bovine Serum Albumin (BSA) in PBS-T (w/v), was applied to the slides after the PBS-T solution had been removed. The samples were treated with a blocking buffer for an hour at RT, followed by three PBS-T wash steps repeated five times. According to the provided table, primary antibodies were diluted with antibody dilution buffer (1% BSA + 0.3% Triton X-100 in PBS) (Table 1). The samples were treated with diluted antibodies overnight at 4 C on the slides. The samples were then rinsed three times with PBS-T for five minutes after the antibody solution had been taken out of the incubation. The antibody dilution buffer is also used to dilute secondary antibodies, as

shown in Table 1. The samples were mixed with diluted secondary antibodies, and the slides were left to incubate at RT without exposure to light for two hours. Slides were washed three times for five minutes with PBS after the incubation. Finally, samples were treated with DAPI diluted to 1 g/mL with PBS and incubated at RT for 10 minutes in the dark. Slides were covered with coverslips after being washed three times with PBS for five minutes. A mounting media was then added. Before being viewed with fluorescence microscopy, slides were kept at 4 °C.

Table 1. Antibody list.

Antibody	Brand and Catalog Number	Host	Dilution
Primary Antibodies			
Anti-Lactoferrin	Abcam, ab109216	Rabbit	1:200
Anti-Lipocalin-2	Abcam, ab41105	Rabbit	1:25
Aquaporin 5	Bioss, bs-1554R	Rabbit	1:200
Cytokeratin 5	Thermo, MA517057	Mouse	1:200
Cytokeratin 19	Cell Signaling, 4558	Mouse	1:200
Secondary Antibodies			
Anti-Mouse	Jackson Immuno 715-585-150	Donkey	1:800
Anti-Rabbit	Jackson Immuno 715-545-152	Donkey	1:800

3.6.3. Forskolin Stimulation

Forskolin stimulation performed by adding 10 µM forskolin (Tocris) into culture media of LG organoids on day 45 for 10 min. After the time is up LG organoids are prepared for the sample preparation step.

3.6.3.1. *Sample Preparation for Metabolomic Analysis*

LG organoids were removed from the incubator at certain intervals (e.g., on days 21, and 45) and instantly rinsed with 3 mL of isotonic sodium chloride (NaCl) solution. The cells were given a 1 mL solution of methanol:water (9:1, v/v) and the petri dish was gently swirled before the cells were submerged in liquid nitrogen to freeze. Using a cell scraper, cells from petri plates maintained in liquid nitrogen were transferred into 2 mL

microcentrifuge tubes. In order to repeat the procedure, 1 mL of a 9:1 methanol:water solution was added to the petri plates in which the cells were scraped. The resultant solution was then transferred to the microcentrifuge tube. After the proteins have precipitated, microcentrifuge tubes are rotated at a speed of 15000 rpm. The supernatants are then transferred to a 2 mL microcentrifuge tube and kept at -80 °C until analysis. Additionally, media samples were taken from LG organoids at certain times (eg, on days 1, 7, 14, 21, 28, 35, 45). Vacuum centrifugation was used to evaporate the collected samples, which were then kept at -80 °C for analysis.

3.6.3.2. *Metabolomic Analysis of Prepared Samples*

Samples kept at -80 °C were thawed at RT and evaporated in a vacuum centrifuge until completely dry (Savant ISS110 SpeedVac Concentrator Thermo Scientific). After being incubated in a 30 °C oven for 90 minutes, dried samples were methoxylated using 20 µL of methoxyamine hydrochloride (20 mg/mL in pyridine). 1% trimethylchlorosilane (Sigma-Aldrich) and 80 µL of N-methyl-N-trimethylsilyl trifluoroacetamide (Sigma-Aldrich) were also added, and samples were then incubated at 37 °C for an additional 30 minutes. Derivatized samples were put into silylated GC-MS vials and analyzed using a DB-5MS stationary phase column (30 m + 10 m DuraGuard 0.25 mm i.d. and 0.25-m film thickness) with a GC-MS system (Shimadzu -QP2010 Ultra). For 5.90 minutes, the solvent delay was set. A gradient of 10 °C/min was used to elevate the oven temperature from 60 °C to 325 °C, which was then maintained for 10 min before cooling down. The temperature of the MSD transfer line was fixed at 290 °C. One mL He/min of flow was maintained via the column. 50–650 Daltons made up the mass range. The analyze lasted 37.5 minutes.

3.6.3.3. *Metabolomic Data Analysis of Samples*

Using the MS-DIAL (ver. 4.0) program, data deconvolution, peak alignment, normalization, and data matrix construction were completed. A commercially available retention index library (Fiehn Retention Index Library) with an identification cut-off score of 70% or above was used for GC-MS metabolite identification. The MS-DIAL data matrix was imported into an Excel work file. The data matrix omitted any metabolite features with

more than 50% of the values missing. The half value of the least concentration in the metabolite group was used to fill in any missing values in the data matrix. The final data matrix was loaded into the multivariate analysis program MetaboAnalyst 5.0 (www.metaboanalyst.ca). Heatmap analysis, PCA, PLS-DA, and enrichment analysis were carried out as part of multivariate analyses. (Figure 8).

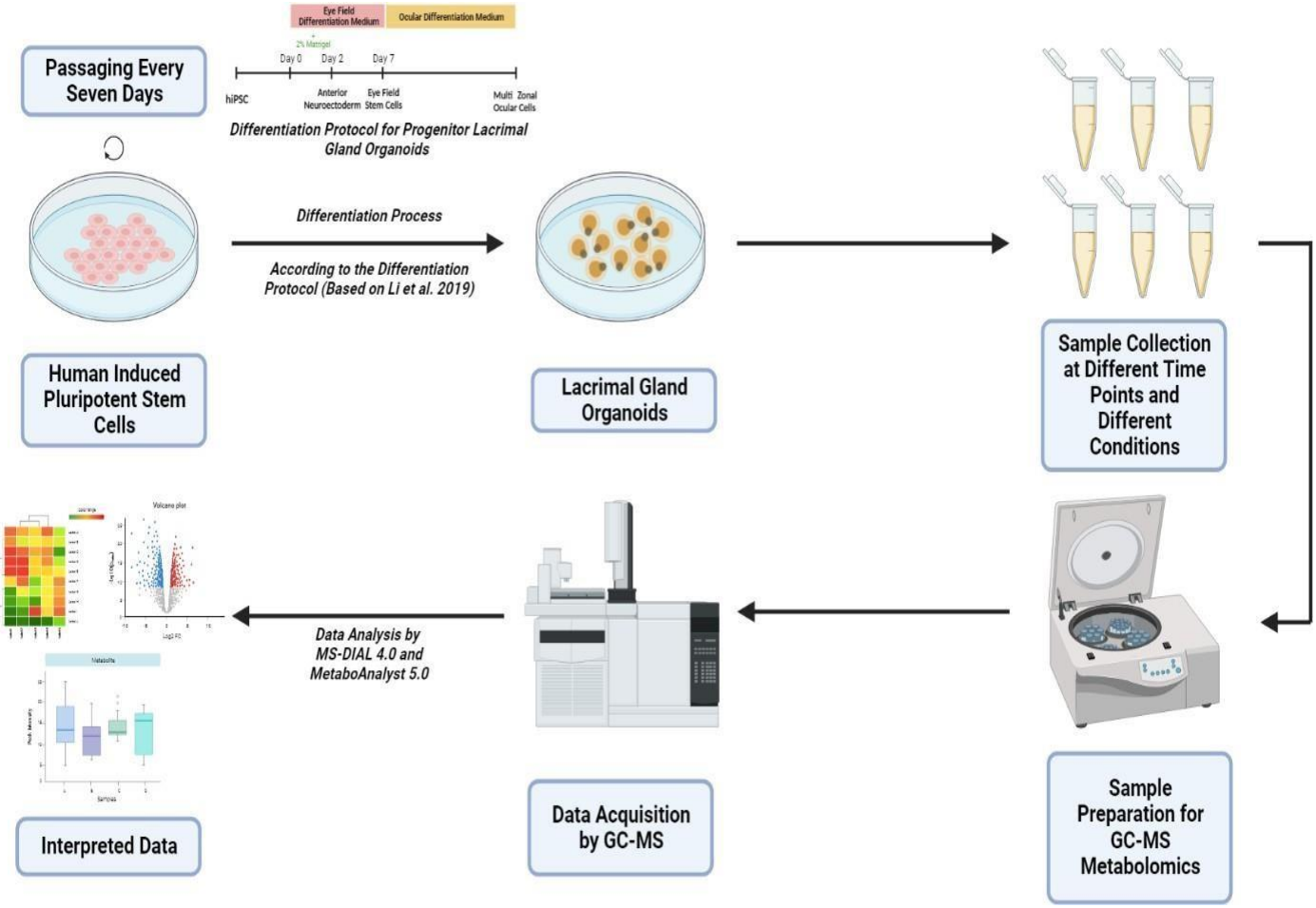


Figure 8. Schematic representation of study design.

3.7. Research Plan

Table 2. Research plan.

Research Plan	September 2020-September 2021	December 2020-February 2022	July 2021-June 2022	July 2022-November 2022
Literature Review	X	X	X	X
Experiment Planning		X	X	
Data Collection and Experimenting		X	X	
Evaluation of Data			X	X
Thesis Writing				X

3.8. Evaluation of Data

MS DIAL 4.0 was used for the annotation of metabolites. Microsoft Excel was used for the data-cleaning process. Metabolomics analyses were carried out by MetaboAnalyst 5.0.

3.9. Limitation of Research

Due to delays occurring during the data acqusition in GC-MS, variations between technical replicates may occur. This effect may cause maybe the reason of numerical fluctuations of identical samples.

3.10. Ethics Approval

This study was approved by the Non-Invasive Research Ethical Committee of IBG with protocol number 2021-022 on 26.07.2021. The thesis approval is presented in section 8.1.

4. RESULTS

4.1. Formation of ocular cells from hiPSCs

According to Section **3.6.1.2**, wild-type hiPSCs were differentiated into ocular cells. Following cell seeding, stimulatory substances were used to begin the differentiation of the hiPSC colonies. Their shifting morphologies and the spread of cells to their surroundings were seen under a microscope. In culture plates, multi-zone development could be visible by day 28. The differentiated cells had quite a lot of heterogeneity, so zones could be seen under a bright-field microscope. Time points on days 1, 7, 14, 21, 28, 35 and 45 (Figure 9).

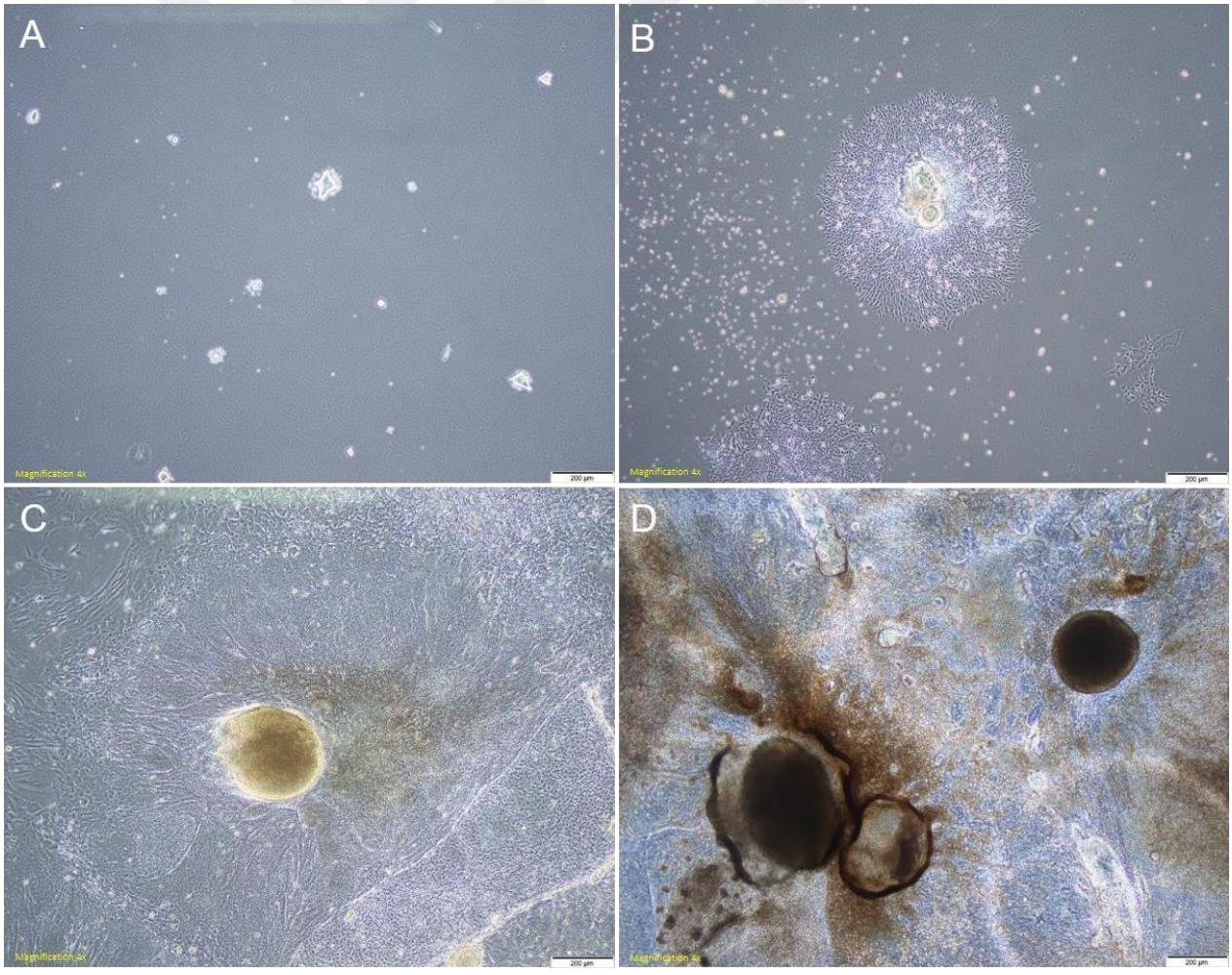


Figure 9. Bright-field images of hiPSCs-derived MZOCs over time.

A) Bright-field image of LG-like structures on day 1. Scale bar: 200 µm. **B)** Bright-field image of LG-like structures on day 7. Scale bar: 200 µm. **C)** Bright-field image of LG-like structures on day 21. Scale bar: 200 µm. **D)** Bright-field image of LG-like structures on day 45. Scale bar: 200 µm.

4.1.2. Immunofluorescent staining of generated LG organoids

Organoids showed AQP5 and lipocalin (LCN) expression, indicating the presence of acinar cells, as well as CK5 and CK19, indicating the presence of myoepithelial cells. Additionally, lactoferrin (LFN) expression was also found in the samples, indicating that they exhibited LG features. The expression of both acinar and myoepithelial cell markers suggests that hiPSCs are differentiated LG organoids (Figure 10).

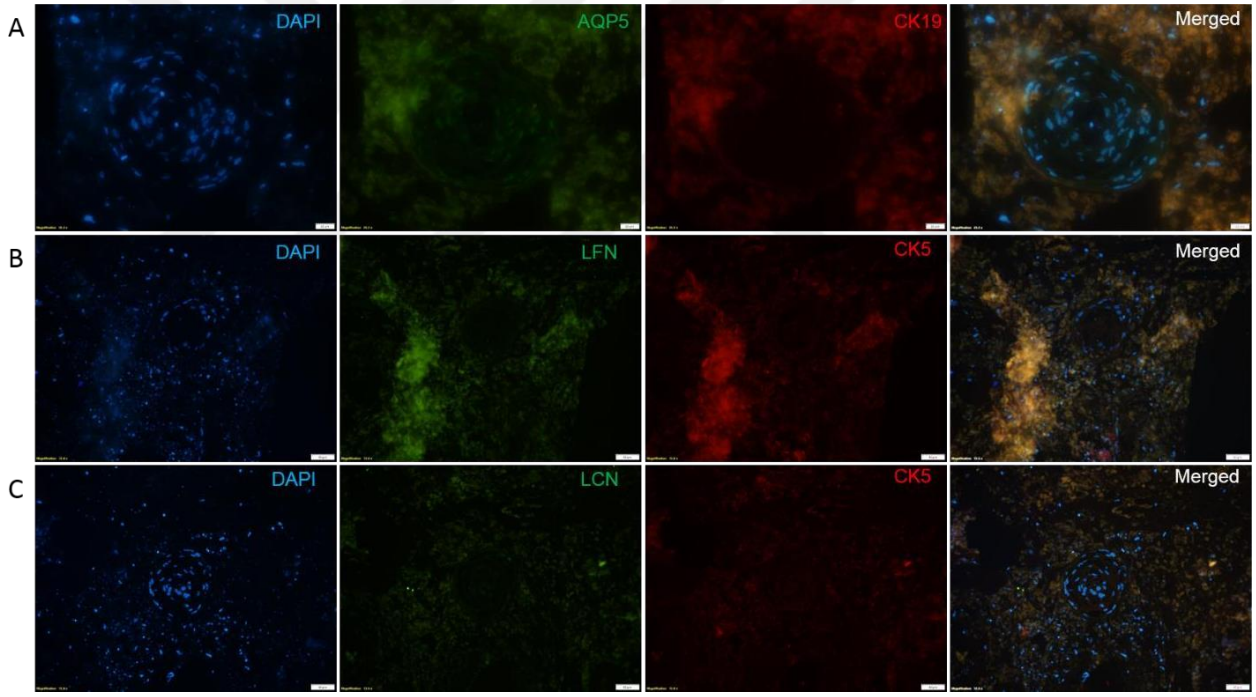


Figure 10. Characterization of LG organoids by immunofluorescent staining.

A) IF images of 45 days old LG-like structure cryosection stained with AQP5 (green) and CK19 (red) antibodies. Nuclei staining was performed via DAPI. Scale bar: 20 µm. **B)** IF images of 45 days old LG-like structure cryosection stained with LFN (green) and CK5 (red) antibodies. Nuclei staining was performed via DAPI. Scale bar: 50 µm. **C)** IF images of 45 days old LG-like structure cryosection stained with LCN (green) and CK5 (red) antibodies. Nuclei staining was performed via DAPI. Scale bar: 50 µm.

4.2. Metabolomics analysis of multi-zonal ocular cells (MZOCs)

4.2.1.1. *Reading and Processing the Raw Data*

Numerous data formats produced by metabolomic research are accepted by MetaboAnalyst, including chemical concentration data, and MS spectra (NetCDF, mzXML, mzDATA). In order for MetaboAnalyst to choose the appropriate algorithm to process the user's data, the data types must be specified when the user uploads their data. The outcomes of the data processing stages are summarized in Table 3 and Table 4.

4.2.1.2. *Reading Peak Intensity Table*

Uploading the peak intensity table in comma separated values (.csv) format is required. Samples could be arranged in columns or rows, with class labels appearing right after the sample identifications (IDs).

Columns include samples, while rows have features. .csv are the format used by the submitted file. The data file that was submitted for the cell sample has a data matrix of 24 samples by 128 metabolites while the data matrix for the media sample is 46 samples by 179 metabolites.

4.2.1.3. *Data Integrity Check*

To ensure that all relevant data has been gathered, a data integrity check is carried out prior to data analysis. Only two classes must be listed in the class labels, which must be present. If samples are paired, the class label for one group must range from -n/2 to -1, and for the other group, it must range from 1 to n/2 (n is the sample number and must be an even number). It is presumed that class labels with the same absolute valuations are pairs. Values for compound concentration or peak intensity must all be positive integers. By default, the least positive value present in the data will be used to replace any missing values, zeros, and negative values.

4.2.1.4. *Missing value imputations*

The downstream analysis may be challenging if there are too many zeros or missing data. For this reason MetaboAnalyst offers a variety of techniques. The default technique assumes that a detection limit is a tiny number, thus it replaces all lacking and zeroes values with small valuations (the other half of the minimum positive values in the original data). This strategy makes the premise that low-abundance metabolites are the primary source of the majority of missing data (i.e.below the detection limit). Additionally, 0 values are also substituted with this small value because they might provide issues for data normalization (i.e., log). To impute the missing values, the user may optionally specify other techniques, such as replacing the missing values with the mean or median or using K-Nearest Neighbours (KNN), Probabilistic PCA (PPCA), Bayesian PCA (BPCA), or Singular Value Decomposition (SVD) (Stacklies et al., 2007). 1/5 of the minimum positive value for each variable was used to replace zero or missing data.

4.2.1.5. *Data Filtering*

Data filtering is used to find and remove factors that are not likely to be useful when modeling the data. Since no phenotypic data were employed in the filtering procedure, the outcome is applicable to any subsequent study. In most cases, this step can enhance the outcomes. When working with datasets that have more than 250 variables, a data filter is highly advised (i.e.chemometrics data). Typically, filtering will enhance your findings (Hackstadt and Hess 2009).

This step will eliminate 5% of the variables from data with fewer than 250 variables and 10% of the variables from data with 250 to 500 variables. Between 500 and 1000 variables, 25% of the variables will be deleted; for data with more than 1000 variables, 40% of the variables will be removed. Only features with fewer than 5000 characteristics can choose none. In addition, the IQR filter will still be used if you select none. Additionally, there can be no more than 10,000 variables. No filtering was implemented.

Table 3. Summary of data processing results of metabolites from cells.

	Features (positive)	Missing/Zero	Features (processed)
iPSC Cell 1	128	0	128
iPSC Cell 2	128	0	128
iPSC Cell 3	128	0	128
iPSC Cell 4	128	0	128
iPSC Cell 5	128	0	128
iPSC Cell 6	128	0	128
MZOC D21 Cell 1	128	0	128
MZOC D21 Cell 2	128	0	128
MZOC D21 Cell 3	128	0	128
MZOC D21 Cell 4	128	0	128
MZOC D21 Cell 5	128	0	128
MZOC D21 Cell 6	128	0	128
MZOC D45 DMSO Cell 1	128	0	128
MZOC D45 DMSO Cell 2	128	0	128
MZOC D45 DMSO Cell 3	128	0	128
MZOC D45 DMSO Cell 4	128	0	128
MZOC D45 DMSO Cell 5	128	0	128
MZOC D45 DMSO Cell 6	128	0	128
MZOC D45 Forskolin Cell 1	128	0	128
MZOC D45 Forskolin Cell 2	127	1	128
MZOC D45 Forskolin Cell 3	128	0	128
MZOC D45 Forskolin Cell 4	128	0	128
MZOC D45 Forskolin Cell 5	128	0	128
MZOC D45 Forskolin Cell 6	128	0	128

Table 4. Summary of data processing results of metabolites from media.

	Features (positive)	Missing/Zero	Features (processed)
MZOC D07 Media 1	179	0	179
MZOC D07 Media 2	179	0	179
MZOC D07 Media 3	179	0	179
MZOC D07 Media 4	179	0	179
MZOC D07 Media 5	179	0	179
MZOC D07 Media 6	179	0	179
MZOC D14 Media 1	179	0	179
MZOC D14 Media 2	179	0	179
MZOC D14 Media 3	179	0	179
MZOC D14 Media 4	179	0	179
MZOC D14 Media 5	179	0	179
MZOC D14 Media 6	179	0	179
MZOC D21 Media 3	179	0	179
MZOC D21 Media 4	178	1	179
MZOC D21 Media 5	179	0	179
MZOC D21 Media 6	179	0	179
MZOC D28 Media 1	178	1	179
MZOC D28 Media 2	179	0	179
MZOC D28 Media 3	179	0	179
MZOC D28 Media 4	179	0	179
MZOC D28 Media 5	178	1	179
MZOC D28 Media 6	179	0	179
MZOC D35 Media 1	179	0	179
MZOC D35 Media 2	179	0	179
MZOC D35 Media 3	179	0	179
MZOC D35 Media 4	179	0	179
MZOC D35 Media 5	178	1	179
MZOC D35 Media 6	179	0	179
MZOC D45 Media 1	179	0	179
MZOC D45 Media 2	179	0	179
MZOC D45 Media 3	178	1	179
MZOC D45 Media 4	179	0	179
MZOC D45 Media 5	178	1	179
MZOC D45 Media 6	179	0	179
MZOC D45 DMSO Media 1	179	0	179
MZOC D45 DMSO Media 2	179	0	179
MZOC D45 DMSO Media 3	179	0	179
MZOC D45 DMSO Media 4	179	0	179
MZOC D45 DMSO Media 5	178	1	179
MZOC D45 DMSO Media 6	179	0	179
MZOC D45 Forskolin Media 1	179	0	179
MZOC D45 Forskolin Media 2	179	0	179
MZOC D45 Forskolin Media 3	178	1	179
MZOC D45 Forskolin Media 4	179	0	179
MZOC D45 Forskolin Media 5	178	1	179
MZOC D45 Forskolin Media 6	179	0	179

4.2.2. Data Scaling

Data scaling (auto-scaling) was performed for both cell and media samples to get better results (Figure 11 and Figure 12).

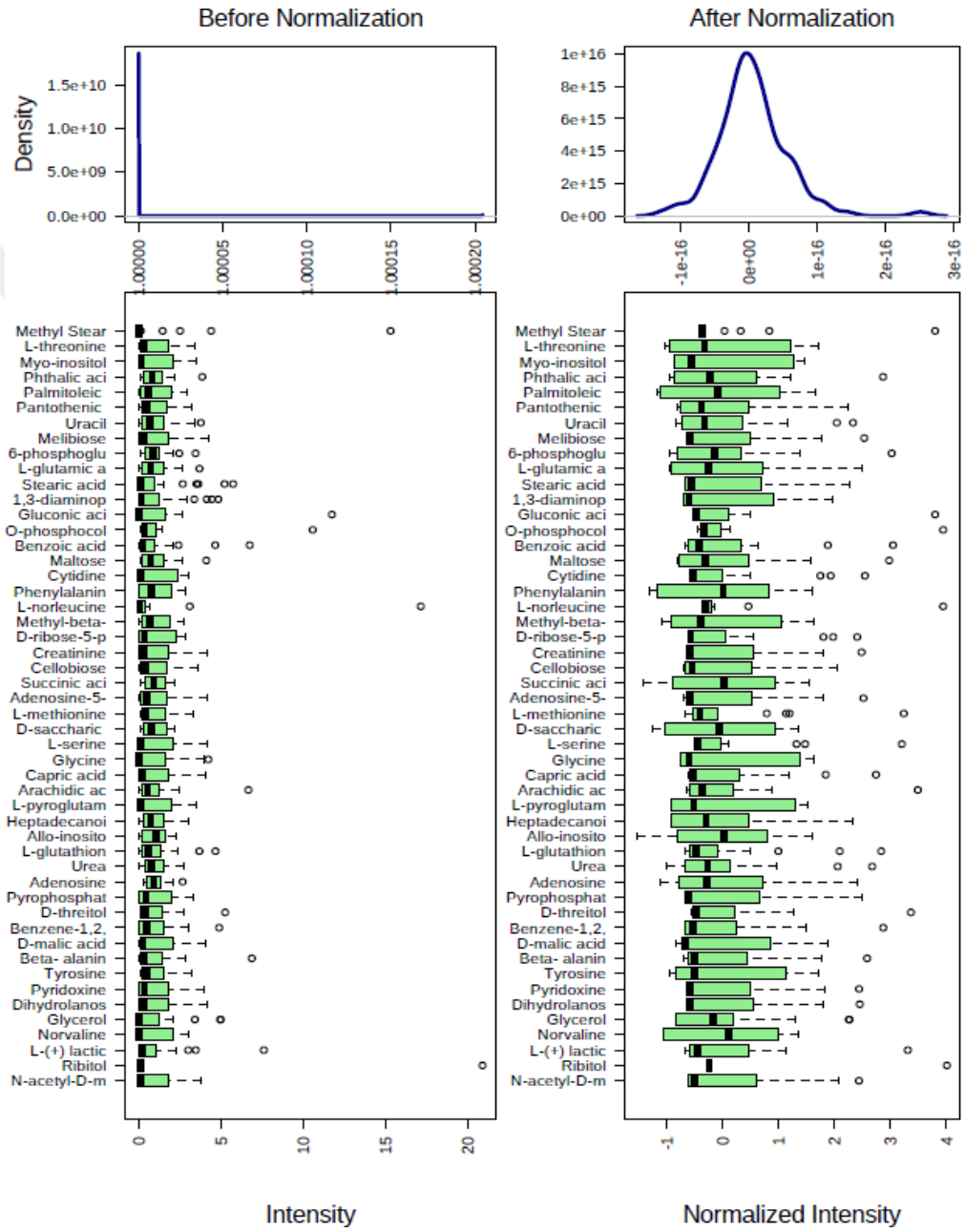


Figure 11. Box plots and kernel density plots before and after data scaling of cell metabolites. The boxplots show at most 50 features due to space limits. The density plots are based on all samples. Selected methods: Row-wise normalization: N/A; Data transformation: N/A; Data scaling: Autoscaling.

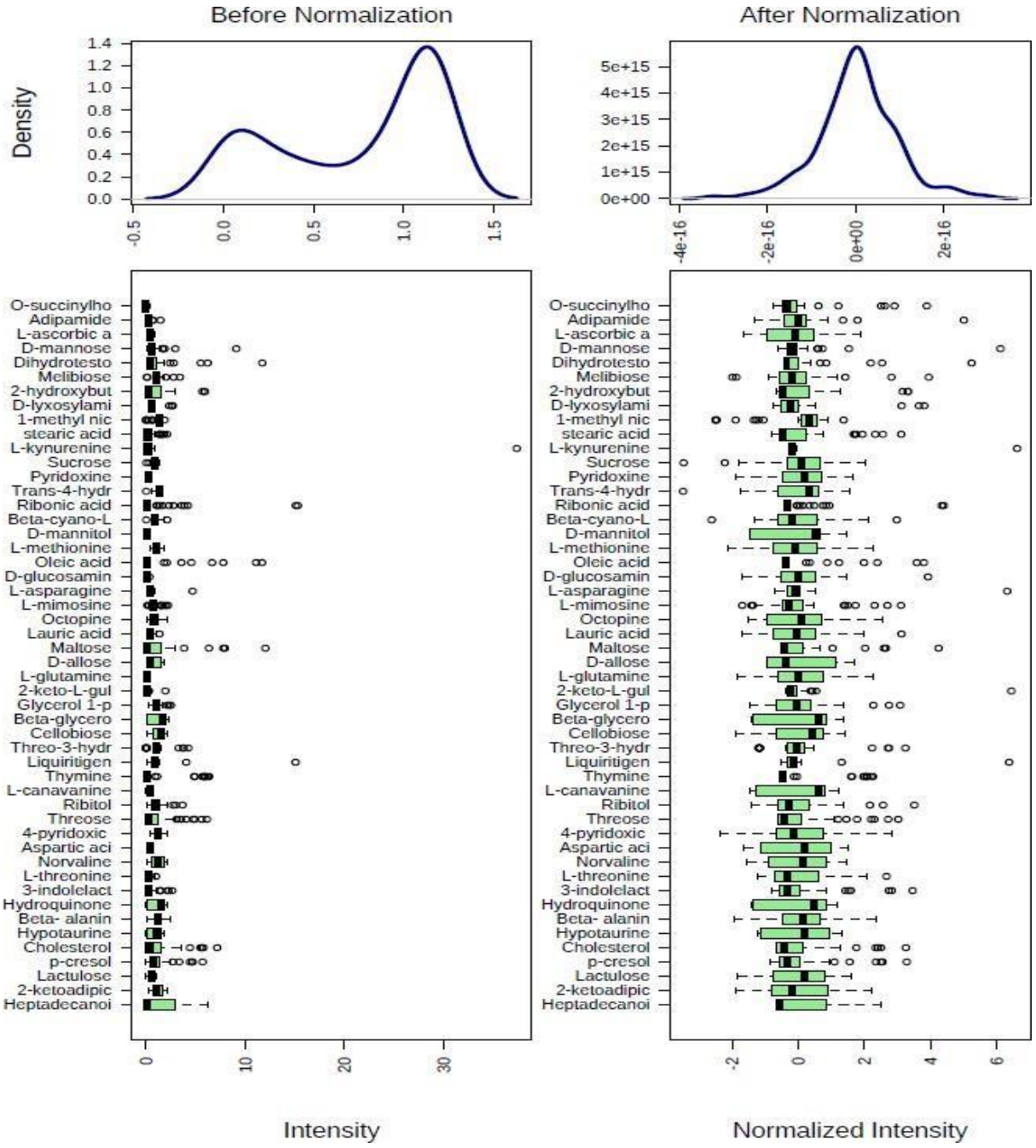


Figure 12. Box plots and kernel density plots before and after data scaling of media metabolites. The boxplots show at most 50 features due to space limits. The density plots are based on all samples. Selected methods: Row-wise normalization: N/A; Data transformation: N/A; Data scaling: Autoscaling.

4.2.3. Statistical Data Analysis

MetaboAnalyst offers a variety of methods commonly used in metabolomic data analyses. They include:

1. Univariate analysis methods:
    - Fold Change Analysis
    - T-tests
    - Volcano Plot
    - One-way ANOVA and post-hoc analysis Correlation analysis
2. Multivariate analysis methods:
    - Principal Component Analysis (PCA)
    - Partial Least Squares - Discriminant Analysis (PLS-DA)
3. Robust Feature Selection Methods in microarray studies
    - Significance Analysis of Microarray (SAM)
    - Empirical Bayesian Analysis of Microarray (EBAM)
4. Clustering Analysis
    - Hierarchical Clustering
    - Dendrogram
    - Heatmap
    - Partitional Clustering
    - K-means Clustering
    - Self-Organizing Map (SOM)
5. Supervised Classification and Feature Selection methods
    - Random Forest
    - Support Vector Machine (SVM)

4.2.3.1. *One-way ANOVA*

The most typical techniques used for exploratory data analysis are univariate methods. MetaboAnalyst offers one-way Analysis of Variance for multigroup analysis (ANOVA). In order to determine whether two levels are different, post-hoc analyses are typically conducted after an ANOVA because it simply indicates if the overall comparison

is significant or not. Fisher's least significant difference technique (Fisher's LSD) and Tukey's Honestly Significant Difference (Tukey's HSD) are the two most popular methods offered by MetaboAnalyst for this purpose. The first review of factors that may be important in differentiating the situations under consideration is provided by the univariate analysis.

Figure 13 and Figure 14 show the important features of both cell and medium samples that were determined using ANOVA analysis. The specifics of these aspects are shown in Table 5 and Table 6. The post-hoc Sig. The comparison column shows the comparisons between different levels that are significant given the p-value threshold.

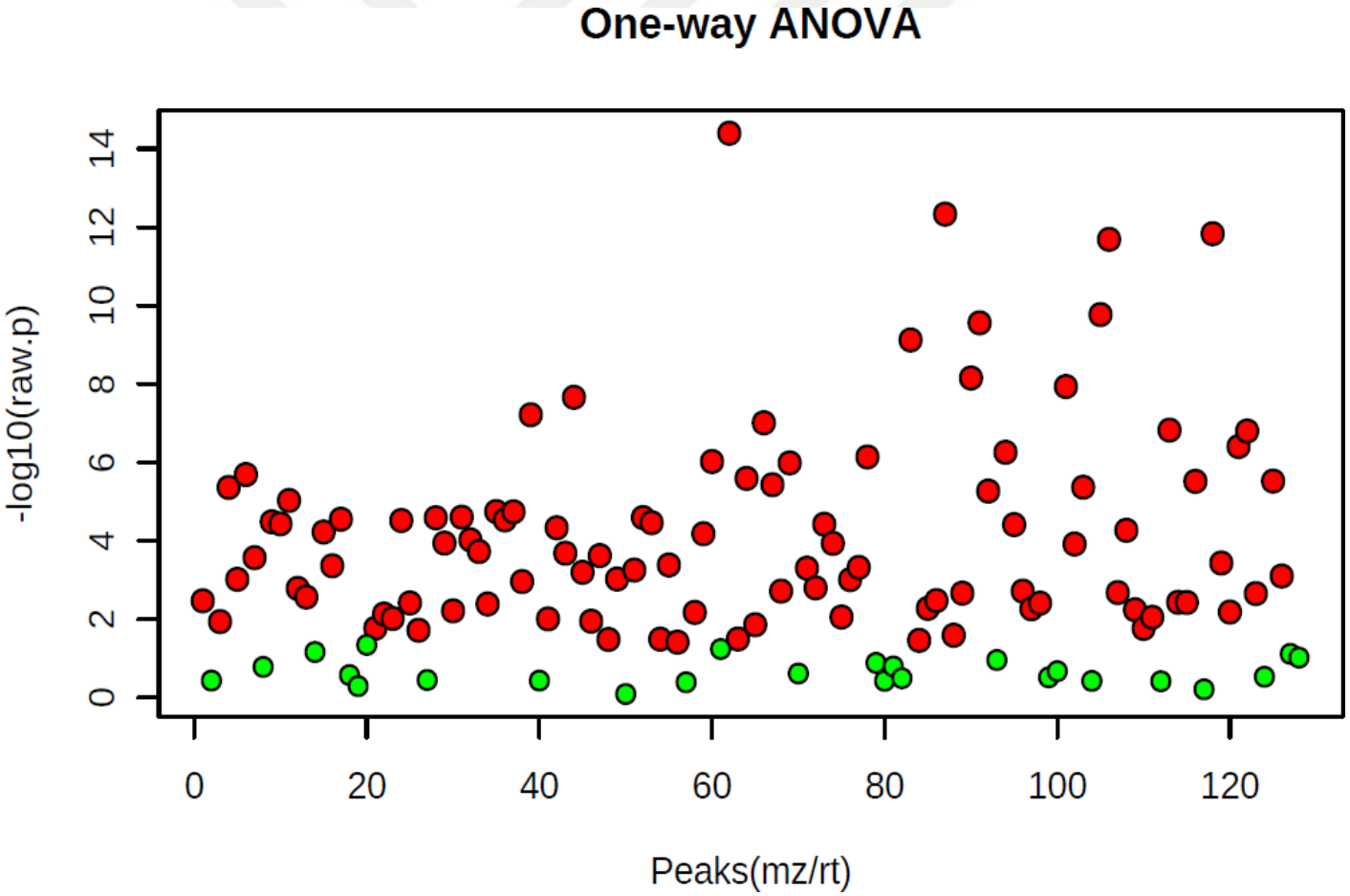


Figure 13. Important features of cell metabolites selected by ANOVA plot with p-value threshold 0.05. Significant metabolites (Green), insignificant metabolites (Red).

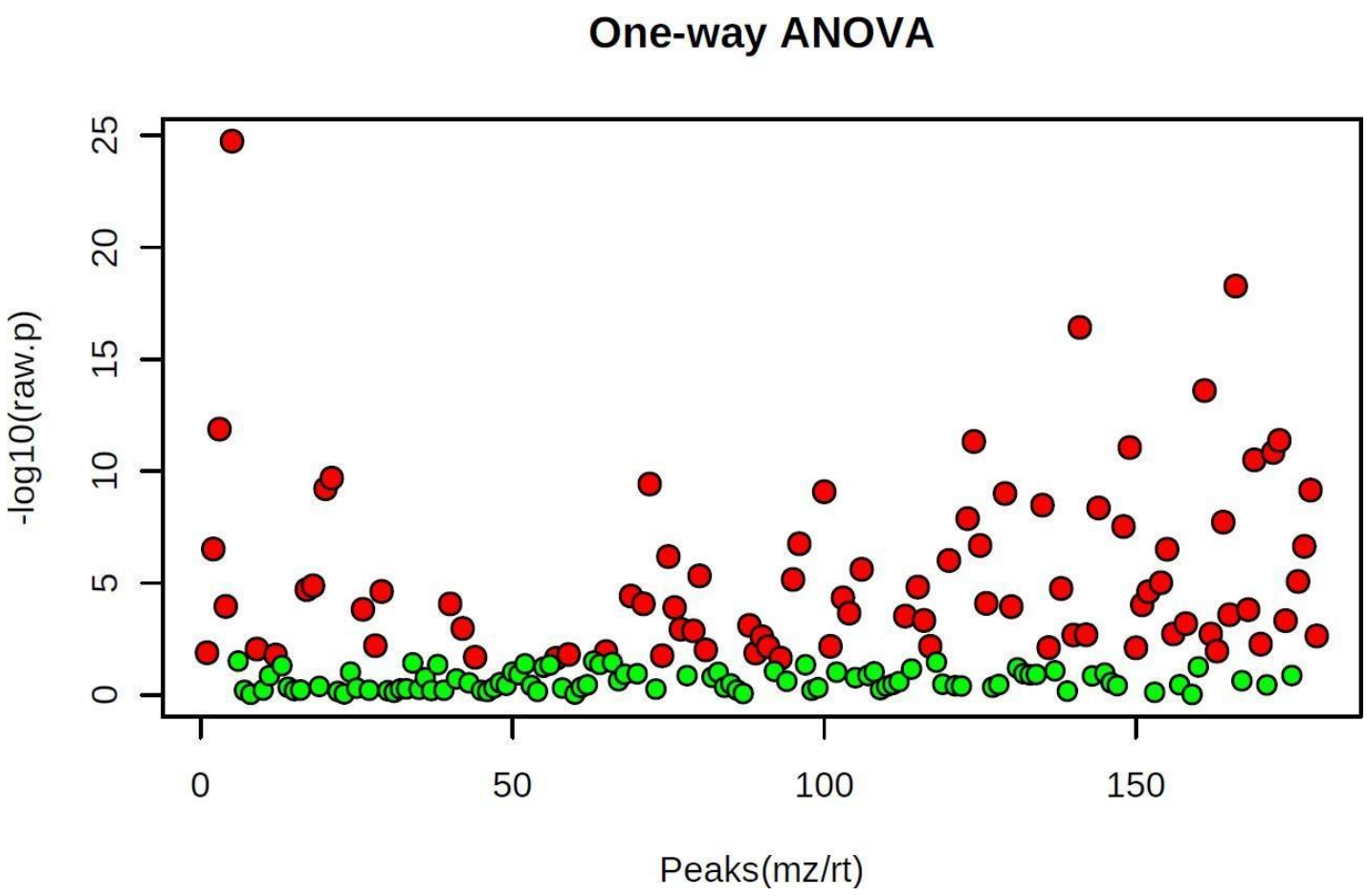


Figure 14. Important features of media metabolites selected by ANOVA plot with p-value threshold 0.05. Significant metabolites (Green), insignificant metabolites (Red).

Table 5. Top 50 features identified by One-way ANOVA and post-hoc analysis for both cell samples.

#	Metabolites	f.value	p.value	-log10(p)	FDR
1	4-guanidinobutyric acid	615.51	$7.35x10^{-20}$	19.13	$9.40x10^{-18}$
2	L-proline	332.22	$3.18x10^{-17}$	16.50	$2.04x10^{-15}$
3	Myo-inositol	295.99	$9.85x10^{-17}$	16.01	$4.20x10^{-15}$
4	Glycine	270.16	$2.40x10^{-16}$	15.62	$7.68x10^{-15}$
5	L-pyroglutamic acid	157.08	$4.54x10^{-14}$	13.34	$1.16x10^{-12}$
6	L-threonine	150.17	$6.99x10^{-14}$	13.16	$1.49x10^{-12}$
7	Tyrosine	98.38	$3.81x10^{-12}$	11.42	$6.96x10^{-11}$
8	L-valine	92.80	$6.56x10^{-12}$	11.18	$1.05x10^{-10}$
9	N-acetyl-L-glutamic acid	83.92	$1.67x10^{-11}$	10.78	$2.37x10^{-10}$
10	Cholesterol	81.18	$2.26x10^{-11}$	10.65	$2.89x10^{-10}$
11	1,3-diaminopropane	70.53	$8.19x10^{-11}$	10.09	$9.53x10^{-10}$
12	Norvaline	52.42	$1.20x10^{-9}$	8.93	$1.25x10^{-8}$
13	D-malic acid	49.32	$2.00x10^{-9}$	8.70	$1.97x10^{-8}$
14	Oleic acid	44.82	$4.60x10^{-9}$	8.33	$4.21x10^{-8}$
15	L-lysine	41.71	$8.54x10^{-9}$	8.07	$7.29x10^{-8}$
16	Cholecalciferol	36.23	$2.82x10^{-8}$	7.55	$2.26x10^{-7}$
17	Palmitic acid	34.17	$4.59x10^{-8}$	7.34	$3.45x10^{-7}$
18	Stearic acid	33.00	$6.12x10^{-8}$	7.21	$4.35x10^{-7}$
19	Methyl-beta-D-galactopyranoside	29.90	$1.37x10^{-8}$	6.86	$9.24x10^{-7}$
20	Acetol	27.02	$3.08x10^{-8}$	6.51	$1.97x10^{-6}$
21	Pyrophosphate	26.20	$3.94x10^{-7}$	6.40	$2.40x10^{-6}$
22	Palmitoleic acid	25.05	$5.60x10^{-7}$	6.25	$3.26x10^{-6}$
23	Alpha-D-glucosamine 1-phosphate	24.21	$7.31x10^{-7}$	6.14	$4.07x10^{-6}$
24	Phytanic acid	23.87	$8.14x10^{-7}$	6.09	$4.34x10^{-6}$
25	D-saccharic acid	22.88	$1.13x10^{-6}$	5.95	$5.78x10^{-6}$
26	D-ribose-5-phosphate	22.67	$1.21x10^{-6}$	5.92	$5.96x10^{-6}$
27	Cytidine	20.94	$2.20x10^{-6}$	5.66	$1.04x10^{-5}$
28	Squalene	20.63	$2.46x10^{-6}$	5.61	$1.13x10^{-5}$
29	Myristic acid	20.43	$2.65x10^{-6}$	5.58	$1.17x10^{-5}$
30	Malonic acid	20.26	$2.82x10^{-6}$	5.55	$1.20x10^{-5}$
31	Glucoheptonic acid	20.14	$2.94x10^{-6}$	5.53	$1.21x10^{-5}$
32	L-ascorbic acid	20.05	$3.04x10^{-6}$	5.52	$1.22x10^{-5}$
33	Lauric acid	19.70	$3.46x10^{-6}$	5.46	$1.34x10^{-5}$
34	Talose	18.20	$6.18x10^{-6}$	5.20	$2.33x10^{-5}$
35	Phenylalanine	17.44	$8.38x10^{-6}$	5.08	$3.06x10^{-5}$
36	Methyl palmitoleate	16.99	$1.01x10^{-5}$	5.00	$3.58x10^{-5}$
37	Glycerol 1-phosphate	16.84	$1.07x10^{-5}$	4.97	$3.71x10^{-5}$
38	Adenosine-5-monophosphate	15.61	$1.82x10^{-5}$	4.74	$6.14x10^{-5}$
39	DL-glyceraldehyde 3-phosphate	15.45	$1.95x10^{-5}$	4.71	$6.39x10^{-5}$
40	D-glucose	15.32	$2.07x10^{-5}$	4.68	$6.62x10^{-5}$
41	L-canavanine	14.90	$2.50x10^{-5}$	4.60	$7.80x10^{-5}$
42	Dihydrolanosterol	14.20	$3.45x10^{-5}$	4.46	$1.03x10^{-4}$
43	Lanosterol	14.18	$3.48x10^{-5}$	4.46	$1.03x10^{-4}$
44	2-butyne-1,4-diol	13.85	$4.07x10^{-5}$	4.40	$1.18x10^{-4}$
45	Leucrose	13.73	$4.32x10^{-5}$	4.36	$1.23x10^{-4}$
46	Ribonic acid-gamma-lactone	13.60	$4.59x10^{-5}$	4.34	$1.28x10^{-4}$
47	Tagatose	13.51	$4.79x10^{-5}$	4.32	$1.30x10^{-4}$
48	Aspartic acid	13.18	$5.63x10^{-5}$	4.25	$1.50x10^{-4}$
49	Melibiose	13.10	$5.90x10^{-5}$	4.23	$1.53x10^{-4}$
50	Pyridoxine	12.98	$6.21x10^{-5}$	4.21	$1.59x10^{-4}$

Table 6. Top 50 features identified by One-way ANOVA and post-hoc analysis for both media samples.

#	Metabolites	f.value	p.value	-log10(p)	FDR
1	Heptadecanoic acid	146.24	$1.80x10^{-25}$	24.74	$3.20x10^{-23}$
2	N-acetyl-D-mannosamine	63.37	$5.38x10^{-19}$	18.27	$4.79x10^{-17}$
3	Gluconic acid	49.42	$3.79x10^{-17}$	16.42	$2.25x10^{-15}$
4	Lactose	33.30	$2.52x10^{-14}$	13.59	$1.12x10^{-12}$
5	Picolonic acid	25.85	$1.33x10^{-12}$	11.87	$4.74x10^{-11}$
6	O-succinylhomoserine	23.94	$4.26x10^{-12}$	11.37	$1.21x10^{-10}$
7	Threonine	23.77	$4.75x10^{-12}$	11.32	$1.21x10^{-10}$
8	Glycine	22.81	$8.83x10^{-12}$	11.05	$1.96x10^{-10}$
9	Lactamide	22.04	$1.46x10^{-11}$	10.83	$2.89x10^{-10}$
10	Nicotinamide	20.92	$3.16x10^{-11}$	10.50	$5.62x10^{-10}$
11	4-methyl-5-thiazoleethanol	18.36	$2.06x10^{-10}$	9.69	$3.33x10^{-9}$
12	Galactinol	17.58	$3.79x10^{-10}$	9.42	$5.62x10^{-9}$
13	Thymine	16.98	$6.12x10^{-10}$	9.21	$8.38x10^{-9}$
14	Phenylalanine	16.81	$7.10x10^{-10}$	9.15	$8.99x10^{-9}$
15	Carbamic acid ethyl ester (urethane)	16.60	$8.35x10^{-10}$	9.08	$9.90x10^{-9}$
16	Phenyl-beta-glucopyranoside	16.38	$1.00x10^{-9}$	9.00	$1.16x10^{-8}$
17	Cystine	14.99	$3.32x10^{-9}$	8.48	$3.48x10^{-8}$
18	Gluconic acid lactone	14.67	$4.43x10^{-9}$	8.35	$4.38x10^{-8}$
19	Valine	13.49	$1.33x10^{-8}$	7.88	$1.24x10^{-7}$
20	Allo-inositol	13.11	$1.91x10^{-8}$	7.72	$1.70x10^{-7}$
21	Trehalose	12.65	$3.00x10^{-8}$	7.52	$2.55x10^{-7}$
22	Linoleic acid	10.95	$1.77x10^{-7}$	6.75	$1.43x10^{-6}$
23	Tryptophan	10.79	$2.11x10^{-7}$	6.68	$1.63x10^{-6}$
24	Ribose	10.71	$2.31x10^{-7}$	6.64	$1.72x10^{-6}$
25	Glucoheptonic acid	10.46	$3.04x10^{-7}$	6.52	$2.13x10^{-6}$
26	Isoleucine	10.44	$3.11x10^{-7}$	6.51	$2.13x10^{-6}$
27	6-deoxy-D-glucose	9.78	$6.60x10^{-7}$	6.18	$4.35x10^{-6}$
28	Maltose	9.44	$9.86x10^{-7}$	6.01	$6.27x10^{-6}$
29	Cysteine	8.69	$2.45x10^{-6}$	5.61	$1.50x10^{-5}$
30	Neohesperidin	8.16	$4.80x10^{-6}$	5.32	$2.85x10^{-5}$
31	Sphingosine	7.87	$6.98x10^{-6}$	5.16	$4.01x10^{-5}$
32	Palmitic acid	7.72	$8.54x10^{-6}$	5.07	$4.75x10^{-5}$
33	Hypoxanthine	7.62	$9.74x10^{-6}$	5.01	$5.25x10^{-5}$
34	2-hydroxybutyric acid	7.39	$1.33x10^{-5}$	4.87	$6.97x10^{-5}$
35	Tyrosine	7.29	$1.52x10^{-5}$	4.82	$7.74x10^{-5}$
36	2-ketoisocaproic acid	7.18	$1.77x10^{-5}$	4.75	$8.74x10^{-5}$
37	Ribonic acid-gamma-lactone	7.10	$1.98x10^{-5}$	4.70	$9.52x10^{-5}$
38	Proline	6.96	$2.40x10^{-5}$	4.62	$1.12x10^{-4}$
39	Glycolic acid	6.94	$2.45x10^{-5}$	4.61	$1.12x10^{-4}$
40	Isomaltose	6.62	$3.84x10^{-5}$	4.42	$1.71x10^{-4}$
41	2-ketobutyric acid	6.49	$4.61x10^{-5}$	4.34	$2.00x10^{-4}$
42	2-hydroxycinnamic acid	6.10	$8.15x10^{-5}$	4.10	$3.45x10^{-4}$
43	Raffinose	6.07	$8.53x10^{-5}$	4.07	$3.47x10^{-4}$
44	Acetyl-L-serine	6.07	$8.57x10^{-5}$	4.07	$3.47x10^{-4}$
45	Glycerol 1-phosphate	6.01	$9.41x10^{-5}$	4.03	$3.72x10^{-4}$
46	Putrescine	5.89	$1.11x10^{-4}$	3.95	$4.30x10^{-4}$
47	Norvaline	5.88	$1.14x10^{-4}$	3.94	$4.31x10^{-4}$
48	Citramalic acid	5.81	$1.26x10^{-4}$	3.90	$4.66x10^{-4}$
49	Homoserine	5.69	$1.51x10^{-4}$	3.82	$5.47x10^{-4}$
50	3-indolelactic acid	5.67	$1.56x10^{-4}$	3.80	$5.57x10^{-4}$

4.2.3.2. *Correlation Analysis*

To see the overall correlations between various characteristics, correlation analysis may be utilized. It may be used to determine which features are relevant to an interesting feature. Additionally, correlation analysis may be used to determine if specific traits exhibit particular patterns under various circumstances. The initial step for users is to specify a pattern as a list of hyphenated numbers. In a time series research with four-time points, for instance, a pattern of 1-2-3-4 is used to look for chemicals whose concentration rises over time, whereas a pattern of 3-2-1-3 can be used to look for compounds whose concentration falls initially before rising again. The total correlation heatmaps are displayed in Figure 15 and Figure 16.

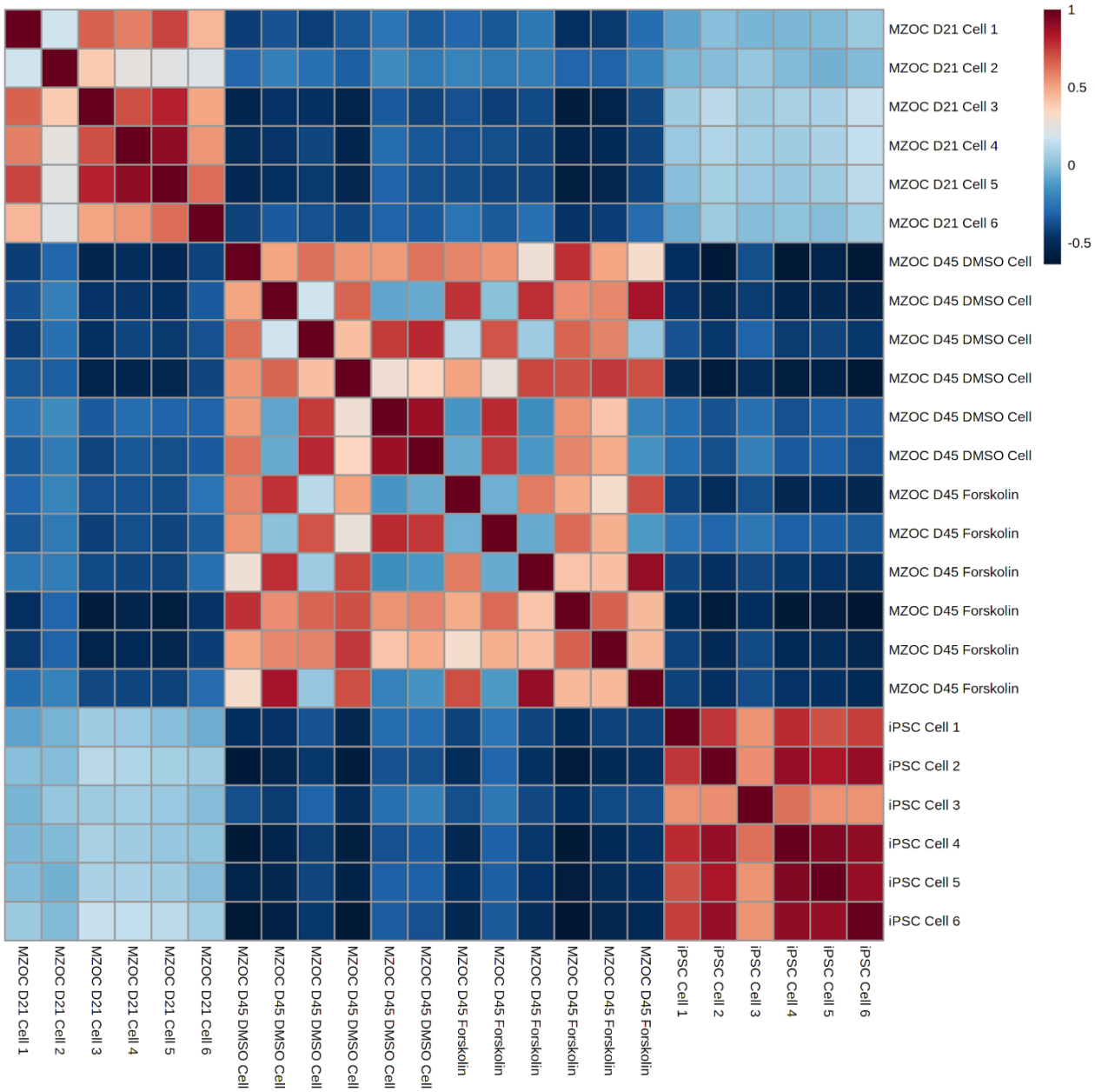


Figure 15. Correlation heatmap of cell samples.

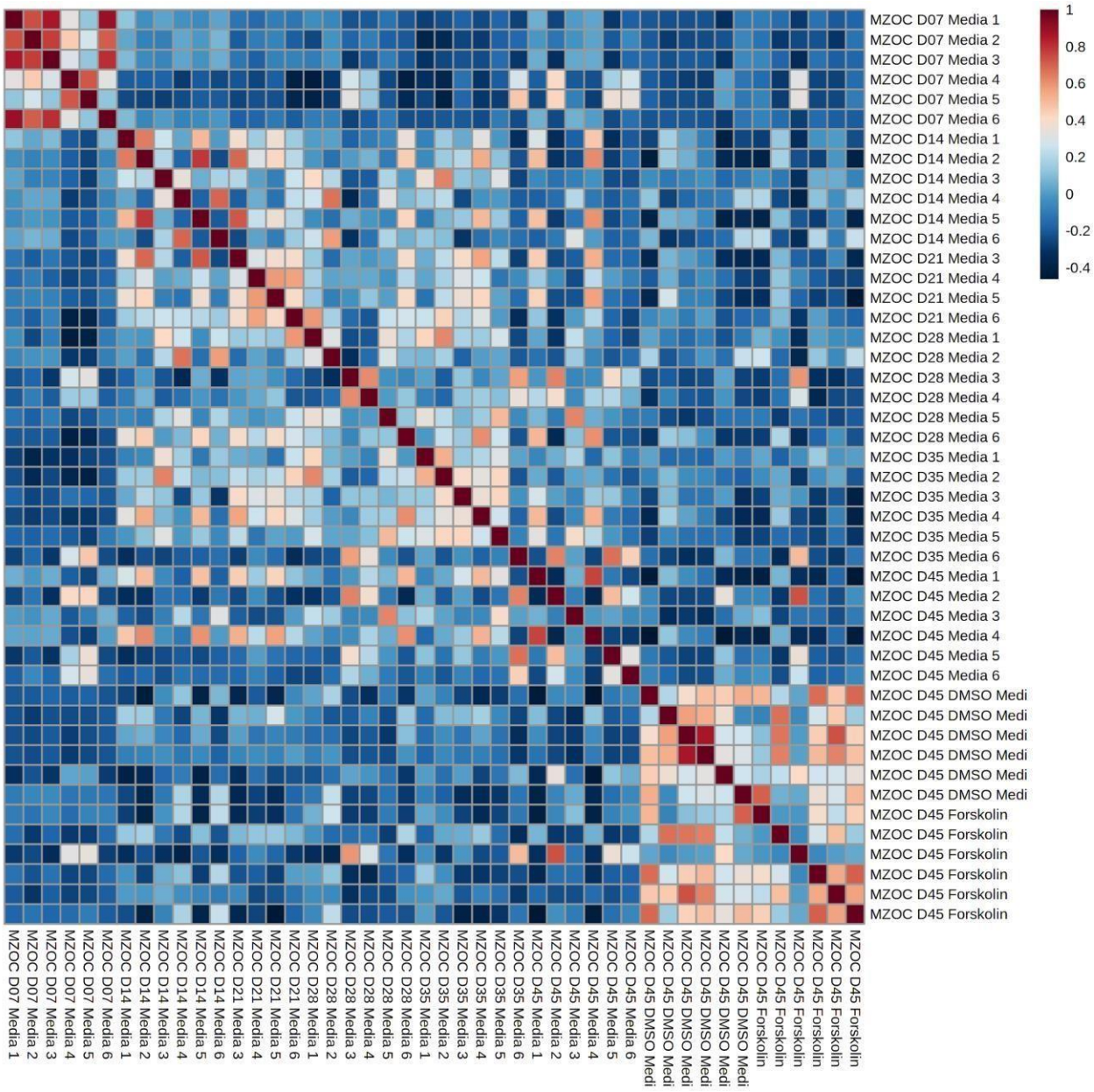


Figure 16. Correlation heatmap of media samples.

4.2.3.3. *Multi Variate Analyses*

4.2.3.3.1. Principal Component Analysis (PCA) Analysis

PCA is an unsupervised technique that considering patterns that best account for divergence in a data set (X) exteriorly using class labels (Y). In scores, which are weighted averages of the original variables, the data are condensed into a considerably smaller number of variables. We refer to the weighting profiles as loadings. Using the prcomp package, the PCA analysis is carried out. On singular value decomposition, the computation is based.

Figure 17 and 18 is pairwise score plots providing an overview of the various separation patterns among the most significant PCs; Figure 19 and 20 shows the 2-D scores plot between selected PCs;

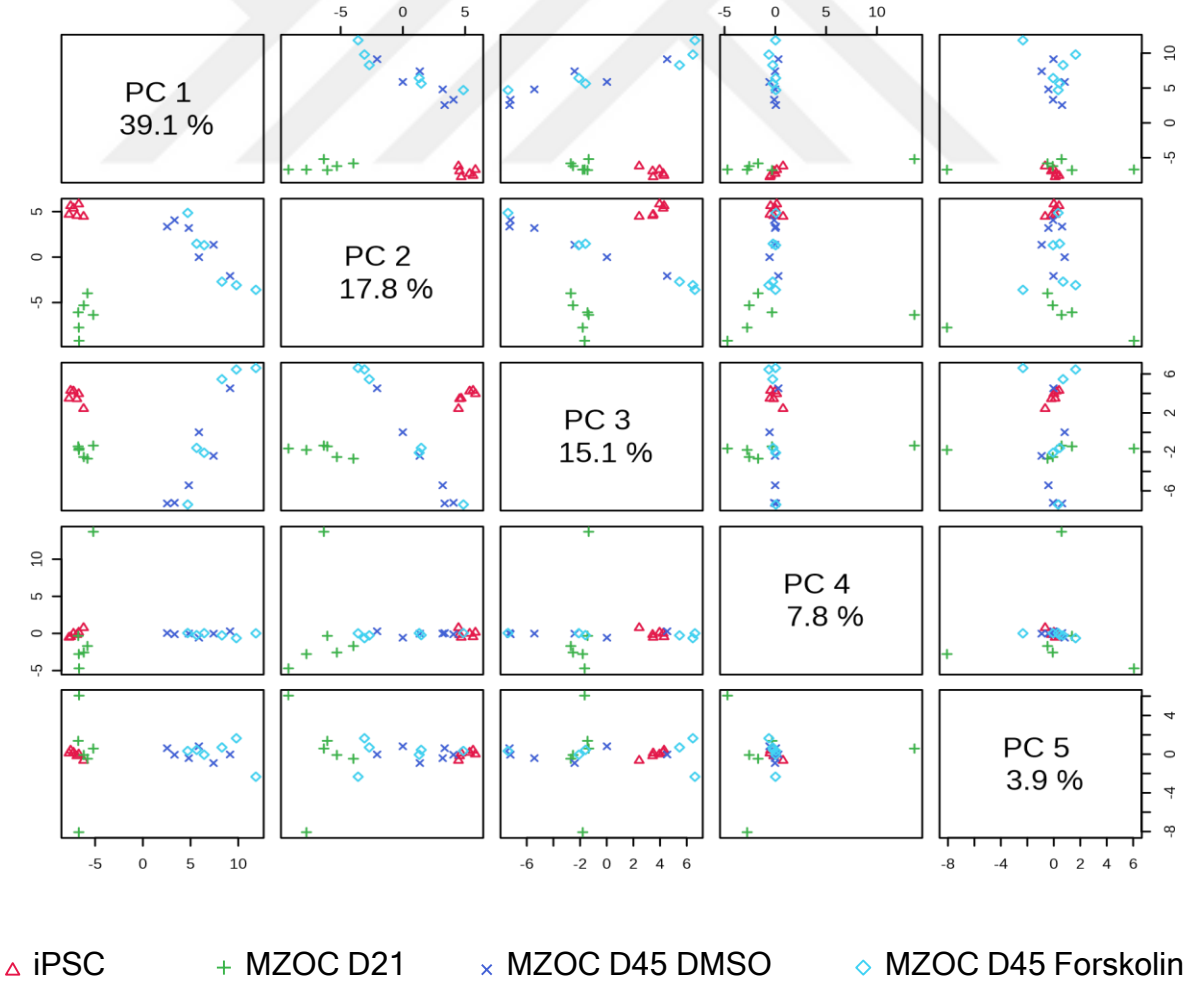


Figure 17. PCA score plots between the selected PCs of cell samples. The explained variance of each PC is shown in the corresponding diagonal cell.

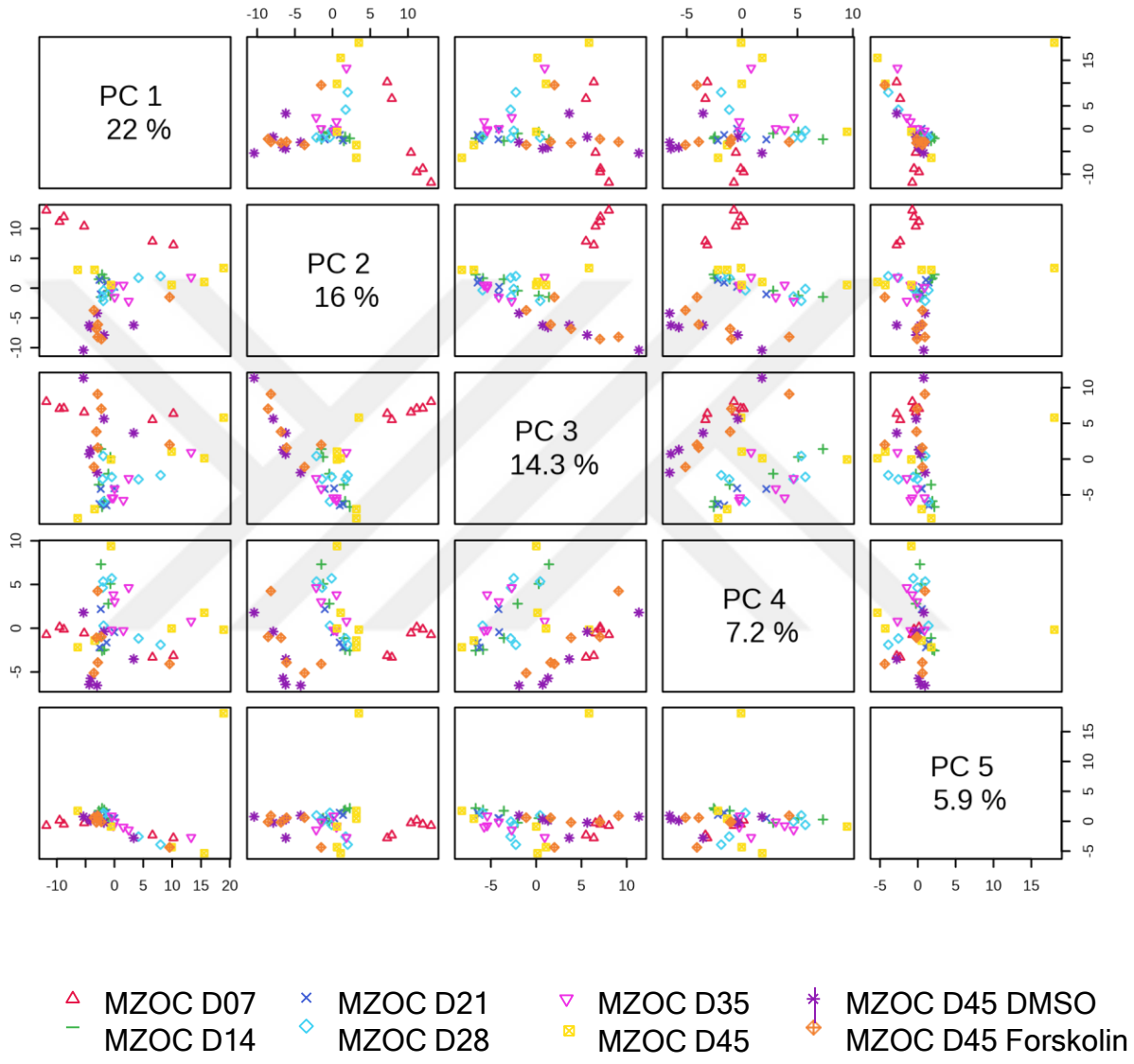


Figure 18. PCA score plots between the selected PCs of media samples. The explained variance of each PC is shown in the corresponding diagonal cell.

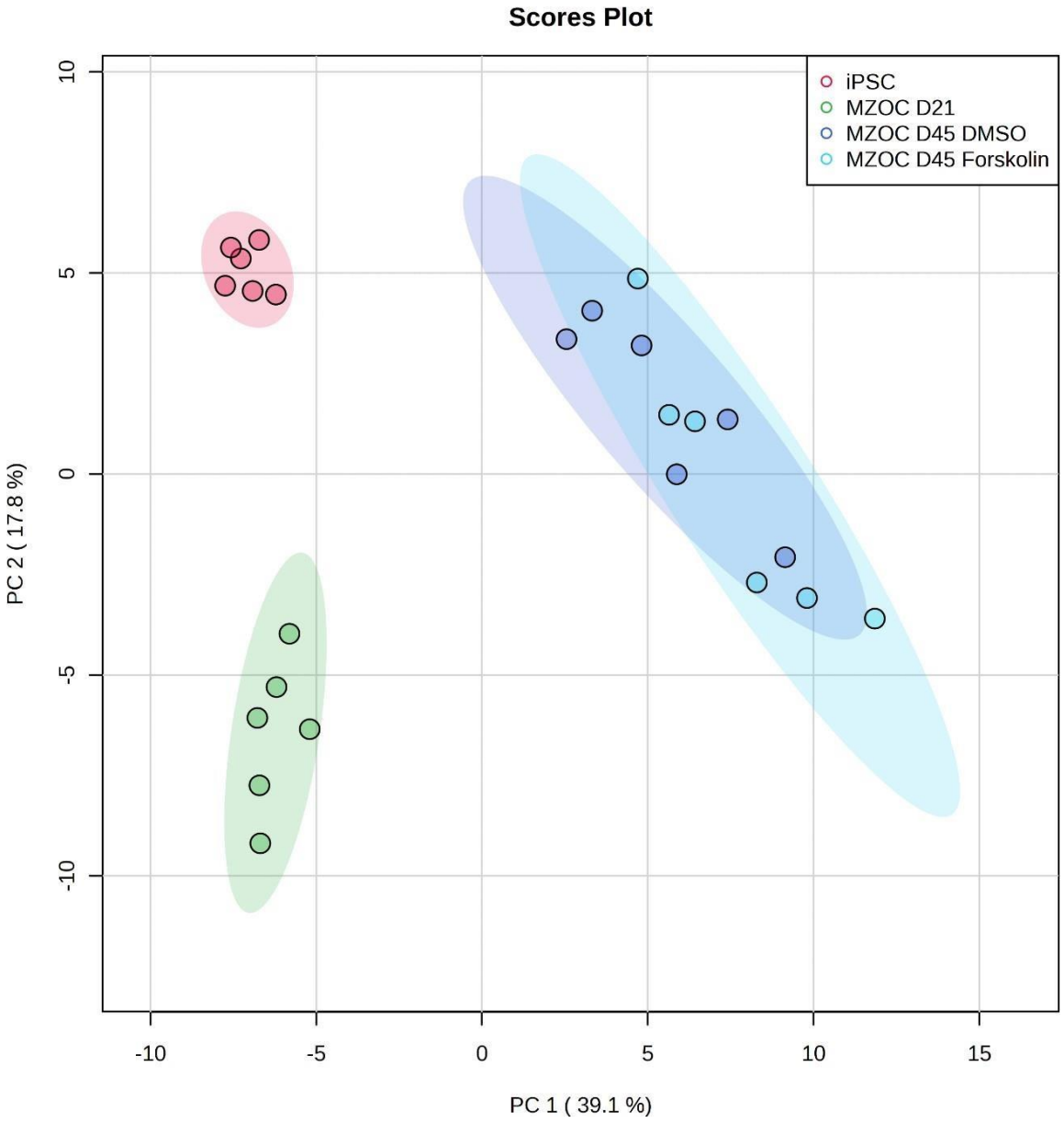


Figure 19. PCA Score plot between the selected PCs of cell samples. The explained variances are shown in brackets.

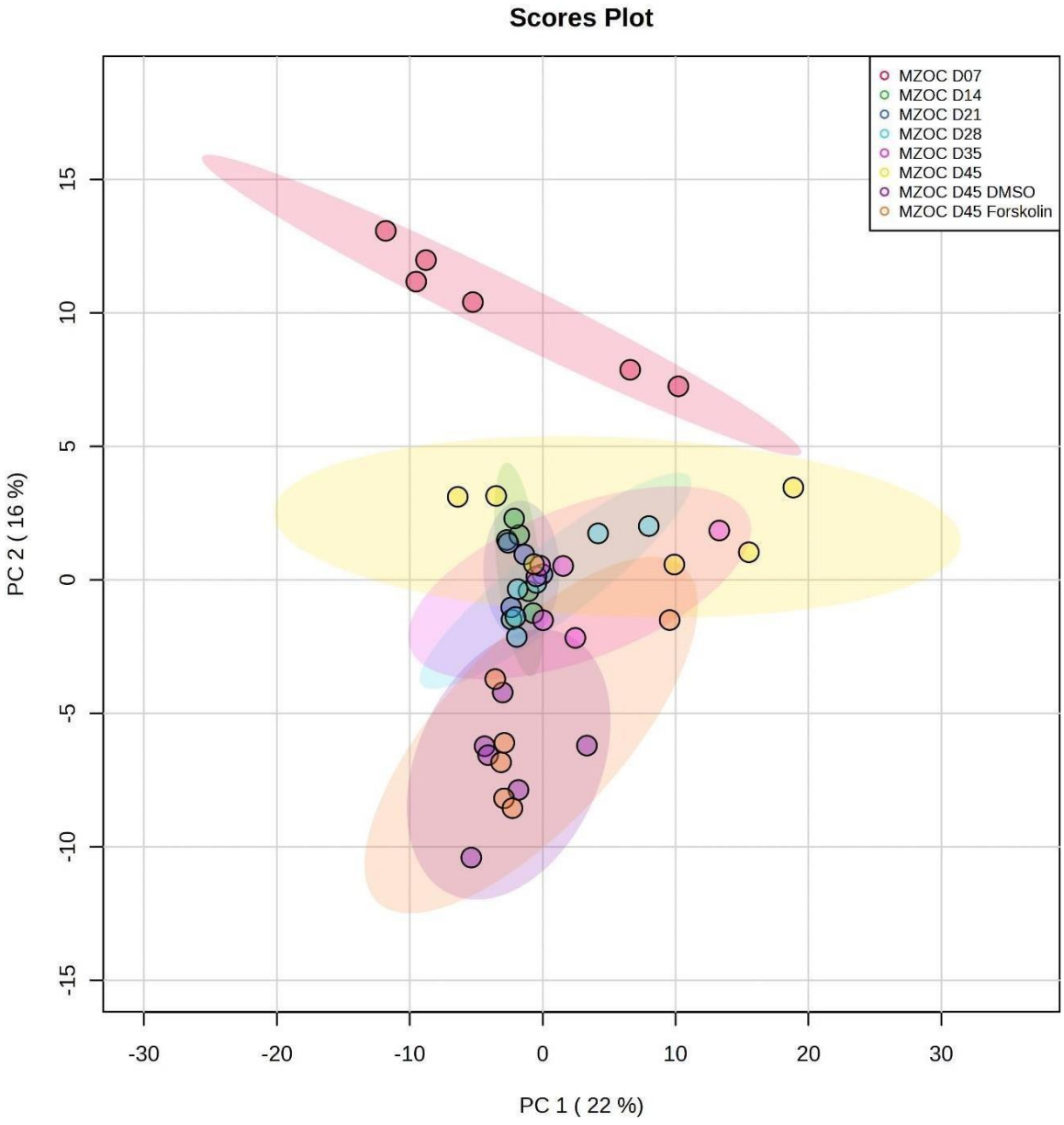


Figure 20. PCA Score plot between the selected PCs of media samples. The explained variances are shown in brackets.

4.2.3.3.2. Partial Least Squares - Discriminant Analysis (PLS-DA)

PLS is a supervised approach that uses multivariate regression techniques to extract data that can predict class membership through a linear combination of the original variables (X) (Y). The plsr function from the R pls package is used to carry out the PLS regression (Wehrens and Mevik, 2007). Using the relevant wrapper function made available by the caret package, the classification and cross-validation are carried out (Kuhn, 2008).

Using a permutation test, the importance of class discrimination was evaluated. The best number of components for the model based on the initial class assignment was found by cross-validation for each permutation, and a PLS-DA model was constructed between the data (X) and the permuted class labels (Y) in each case. Two different test statistics types are supported by MetaboAnalyst to measure class discrimination. The first one is based on prediction accuracy during training. The second one is the separation distance calculated using the B/-W ratio, which compares the total of the squares within and between groups. Class discrimination is not statistically significant if the observed test statistic is a component of the distribution based on permuted class assignments (Bijlsma et al., 2006).

In PLS-DA, there are two measures of varying relevance. The first, called Variable Importance in Projection (VIP), is a weighted sum of squares of the PLS loadings that accounts for how much Y-variation in each dimension has been explained. Please be aware that VIP scores are determined for each component. The average of the VIP scores is applied when determining the feature significance using more than one component. Based on the weighted sum of PLS-regression, the other significance metric is used. The weights are a result of the sums of squares being divided by the number of PLS components being reduced. The same number of predictors will be constructed for each group in a multiple-group (greater than two) analysis, so take note of that. As a result, based on the group you wish to forecast, the coefficient of each attribute will vary. The entire coefficient-based relevance is represented by the feature coefficients' average value.

Figure 21 and Figure 22 depicts an overview of the scores plots, Figure 23 and Figure 24 the 2-D scores plot between chosen components, and Figure 25 and Figure 26 the 3-D scores plot; Figure 27 and Figure 28 display the categorization performance for various component counts; Figure 29 and Figure 30 displays significant features determined using PLS-DA.

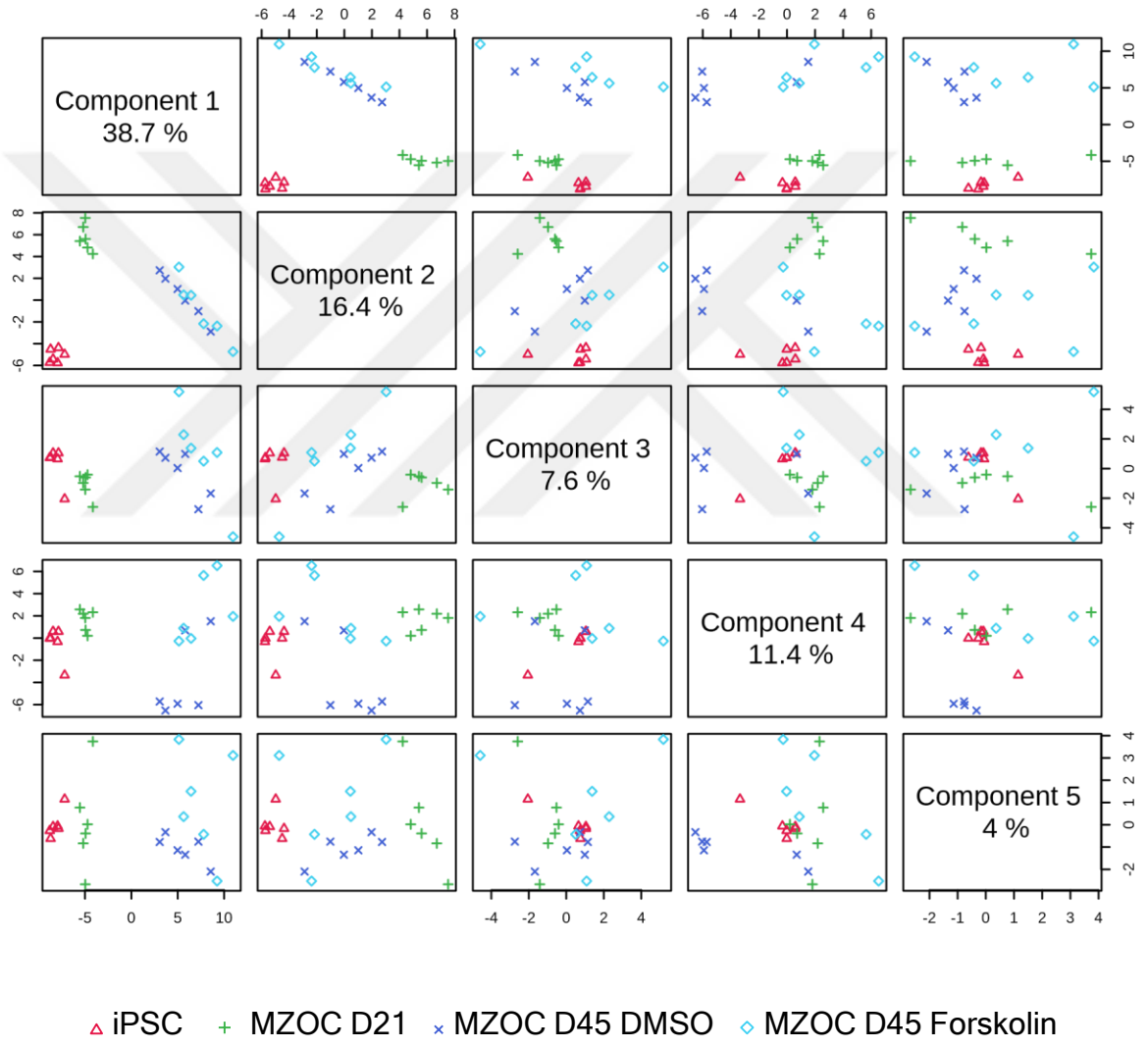


Figure 21. Pairwise PLS-DA score plots between the selected components of cell samples. The explained variance of each component is shown in the corresponding diagonal cell.

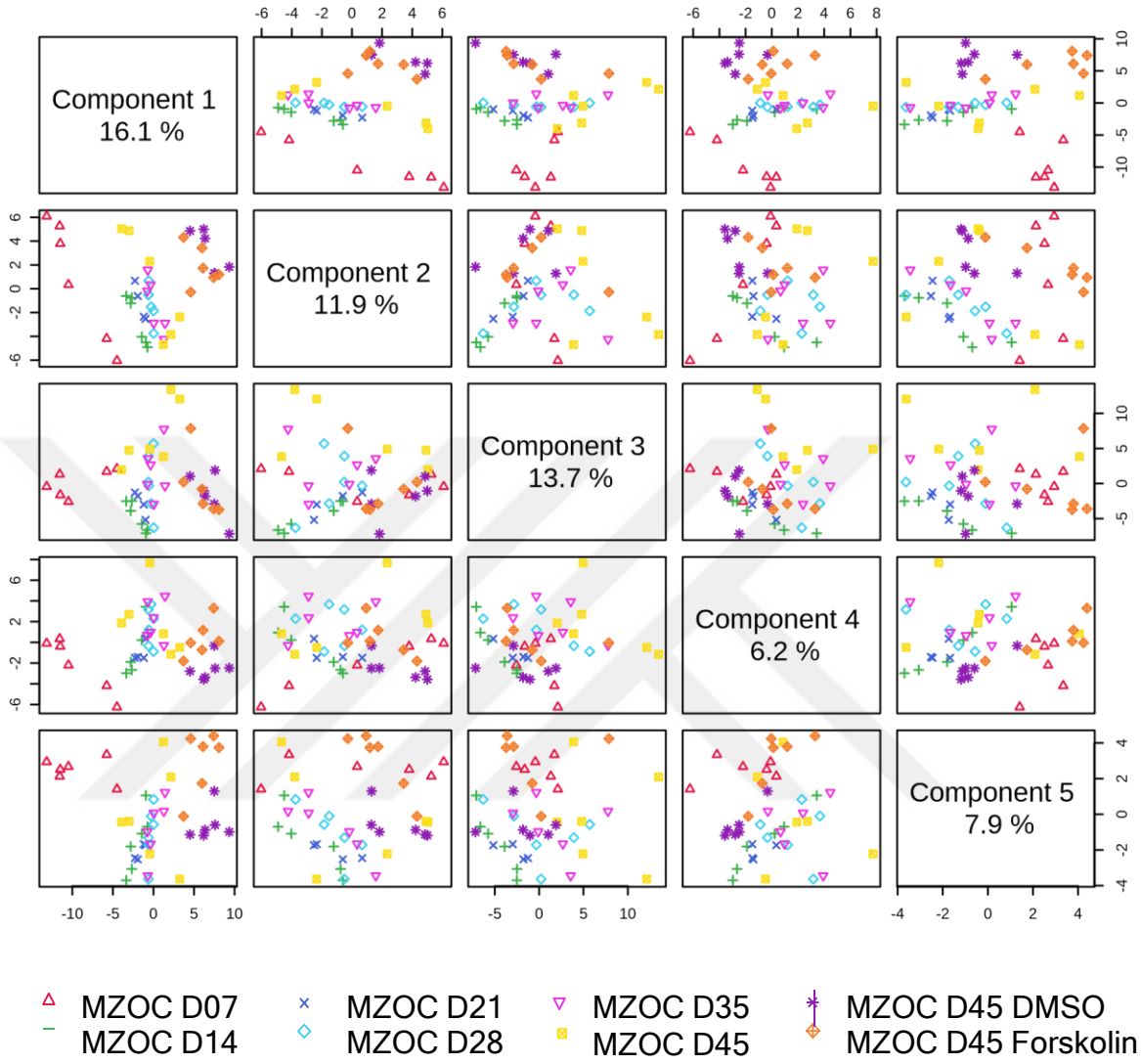


Figure 22. Pairwise PLS-DA score plots between the selected components of media samples. The explained variance of each component is shown in the corresponding diagonal cell.

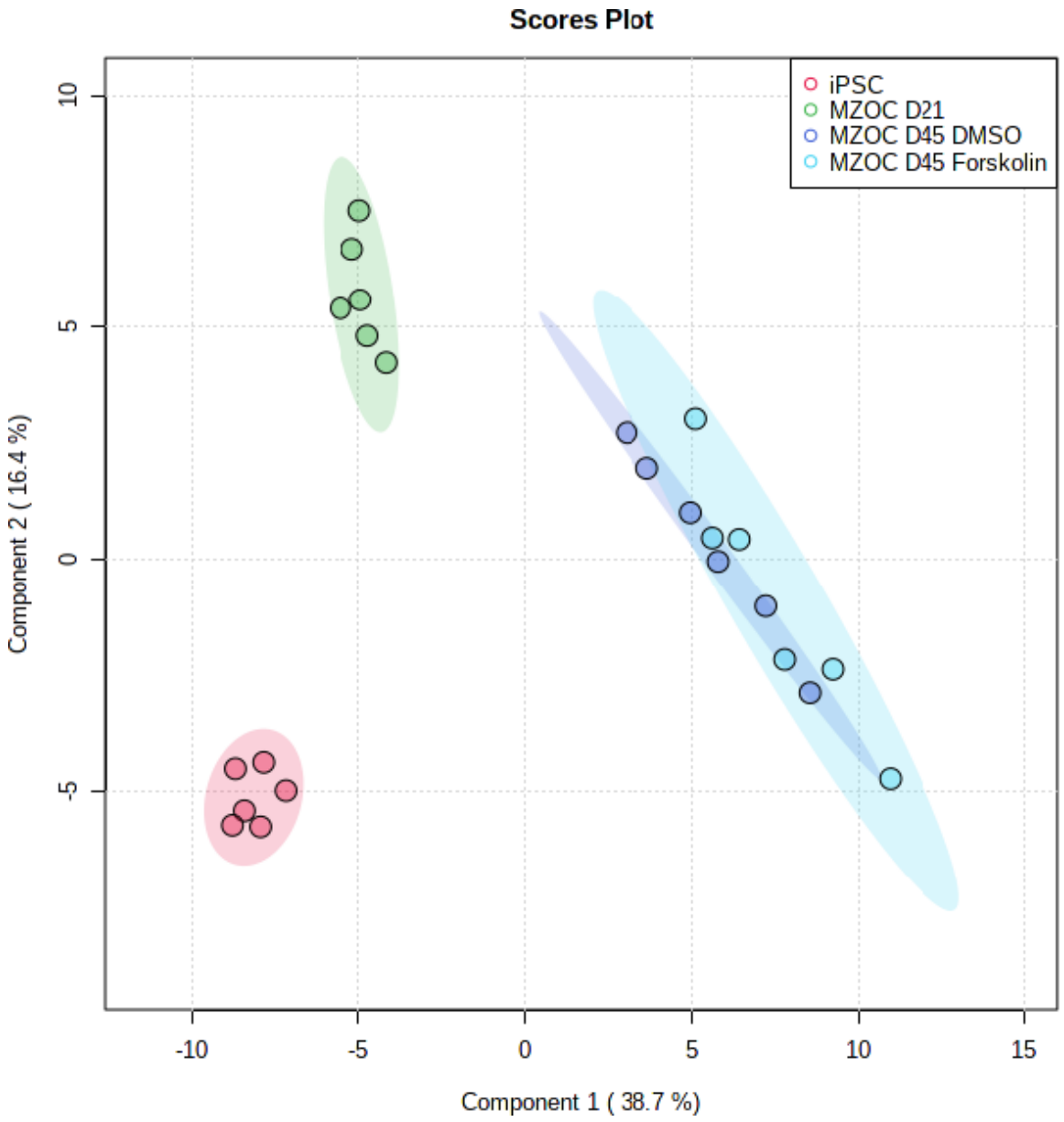


Figure 23. PLS-DA score plot between the selected PCs of cell samples. The explained variances are shown in brackets.

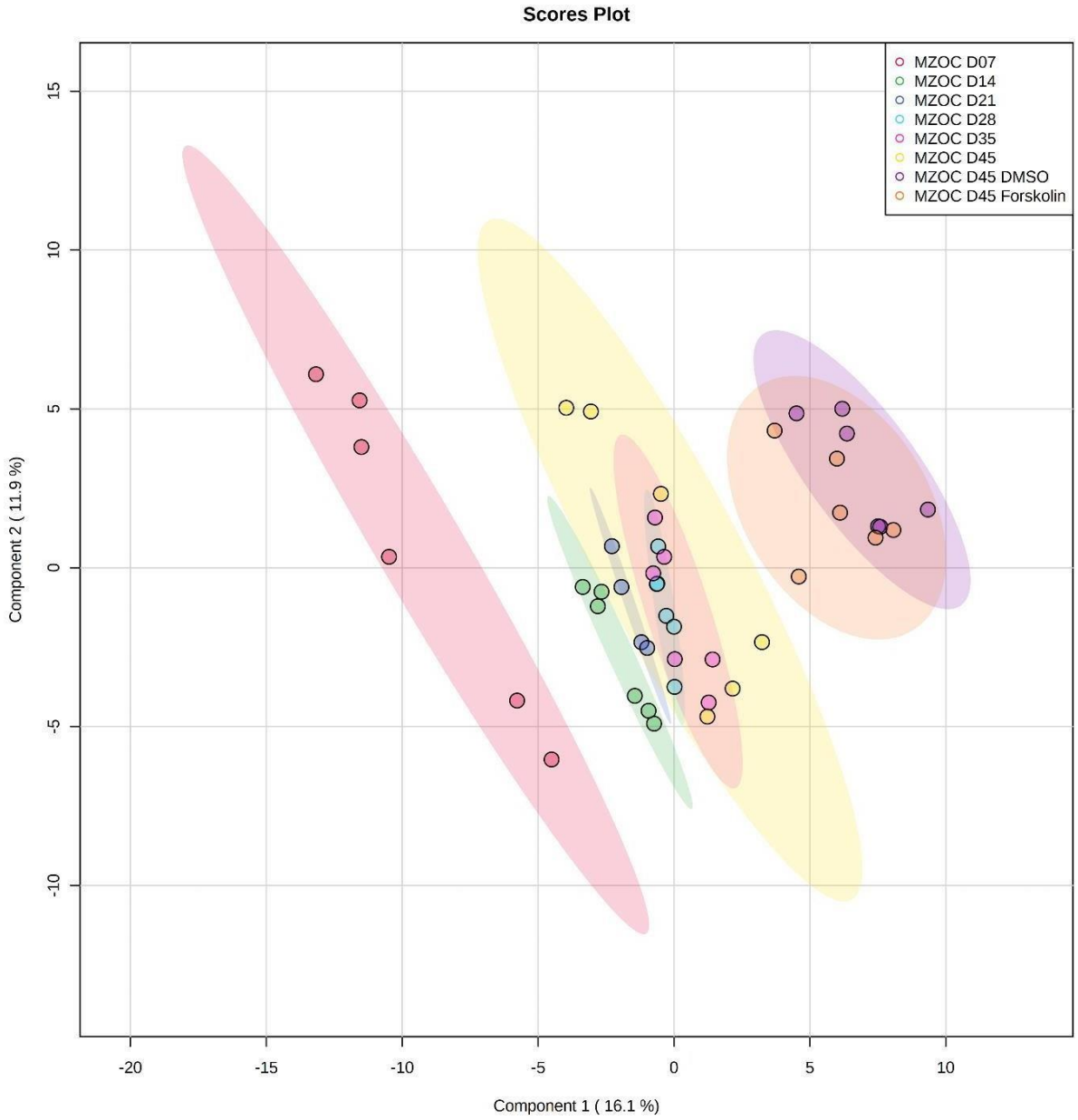


Figure 24. PLS-DA score plot between the selected PCs of media samples. The explained variances are shown in brackets.

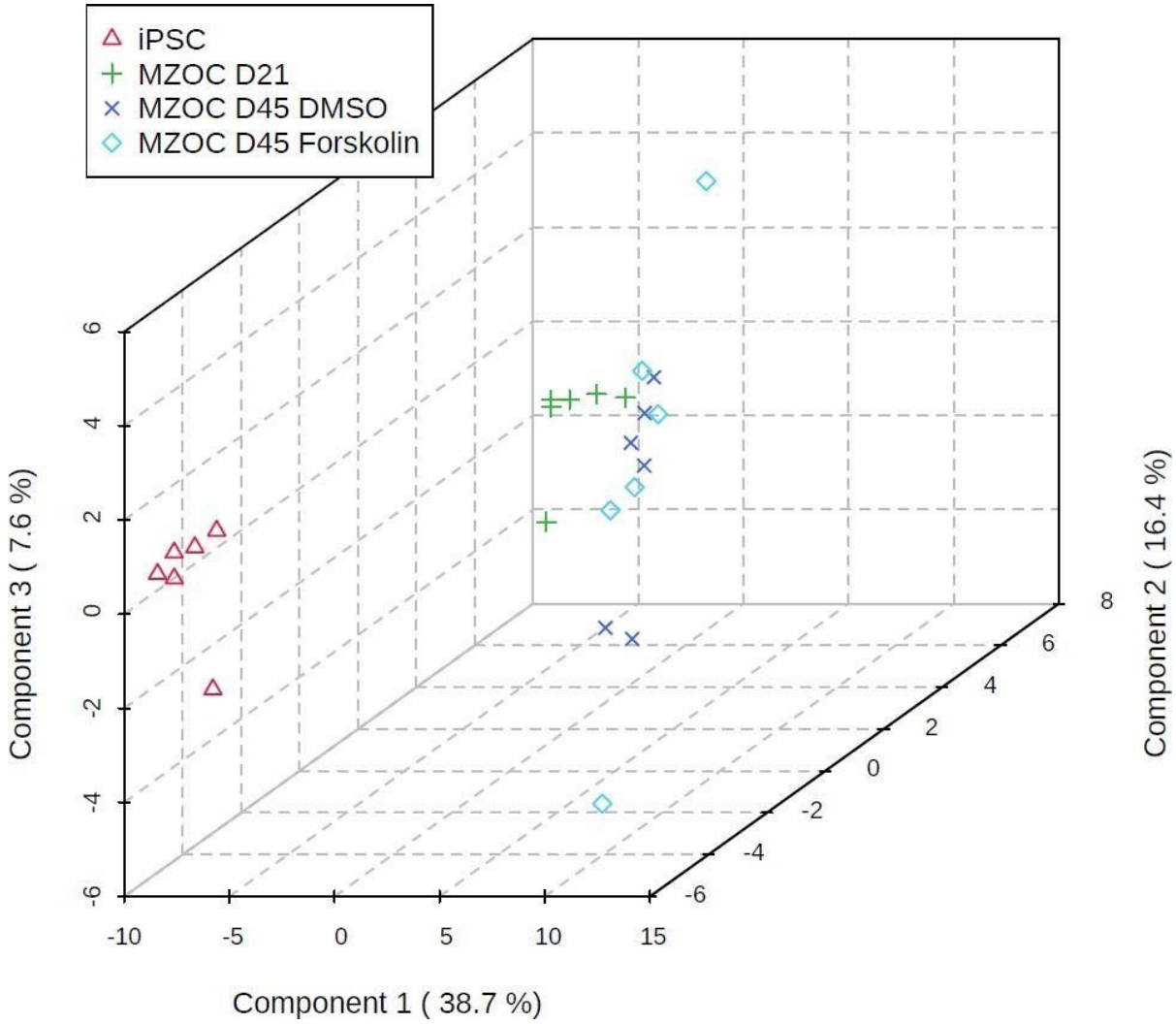


Figure 25. 3D PLS-DA score plot between the selected PCs of cell samples. The explained variances are shown in brackets.

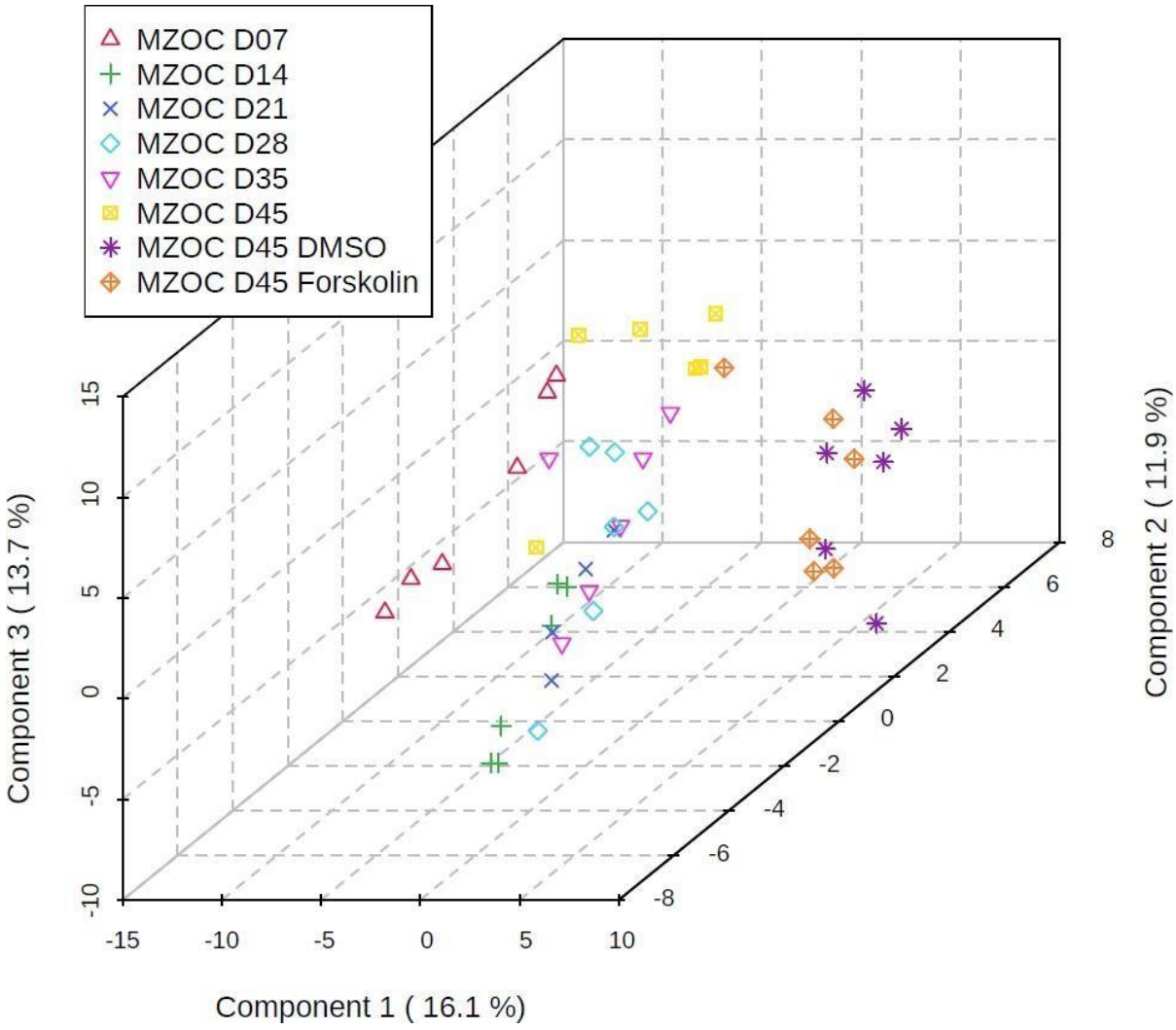


Figure 26. 3D PLS-DA score plot between the selected PCs of media samples. The explained variances are shown in brackets.

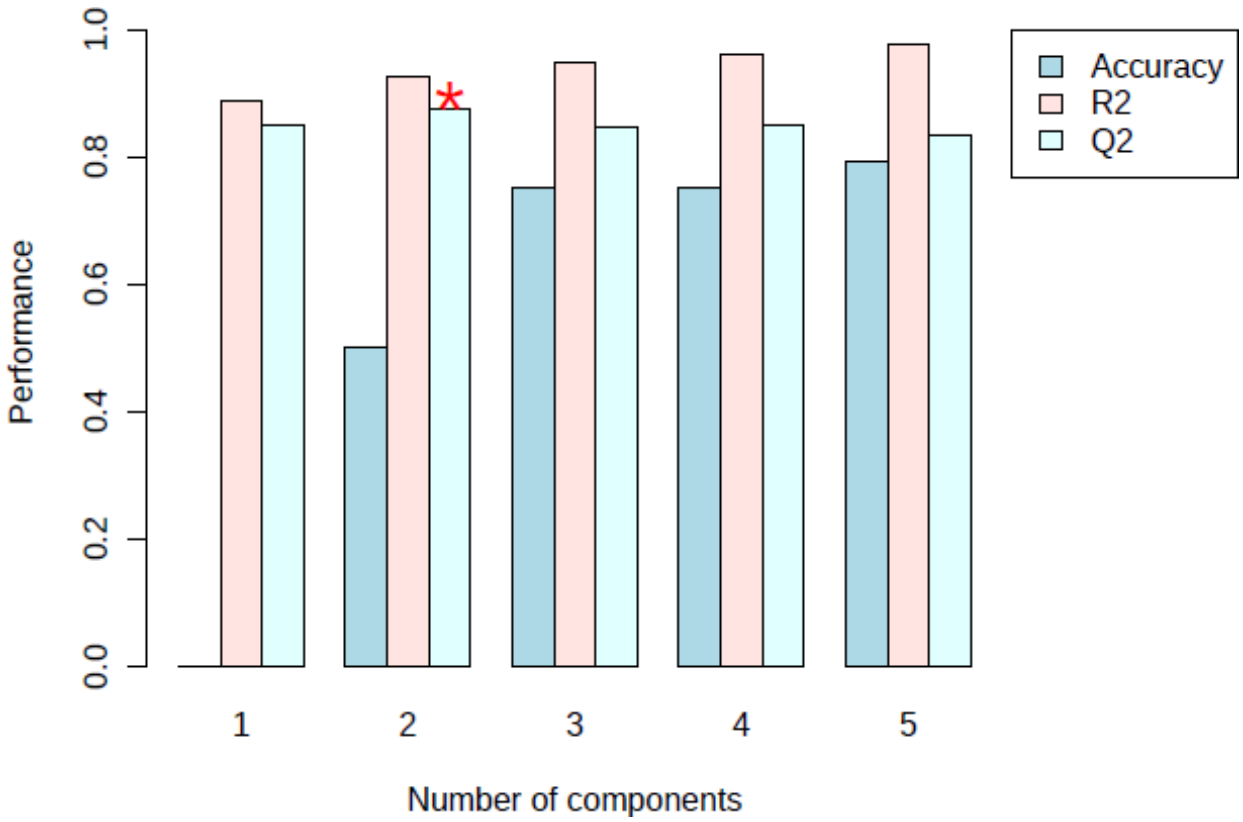


Figure 27. PLS-DA classification using different numbers of components. The red star indicates the best classifier for cell samples. R^2 value: 0.92689 Q^2 value: 0.87153

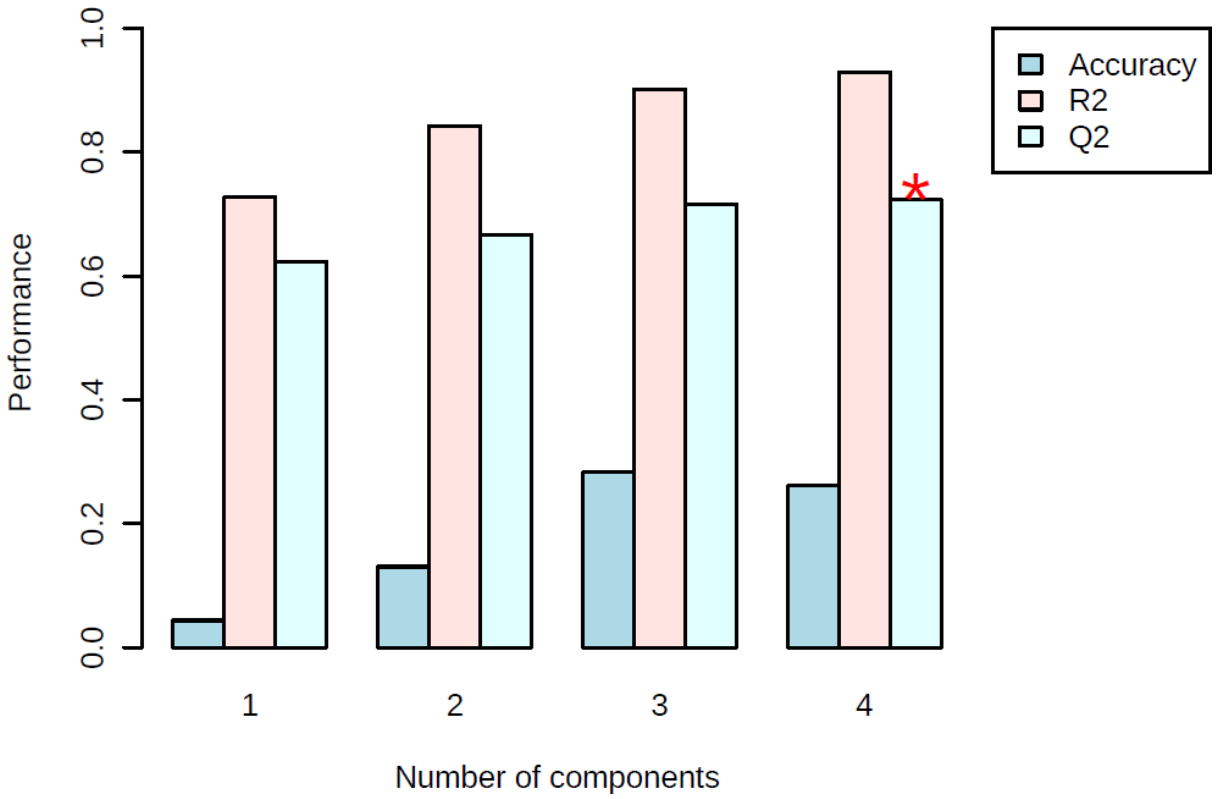


Figure 28. PLS-DA classification using different numbers of components. The red star indicates the best classifier for media samples. R^2 value: 0.87835 Q^2 value: 0.7006

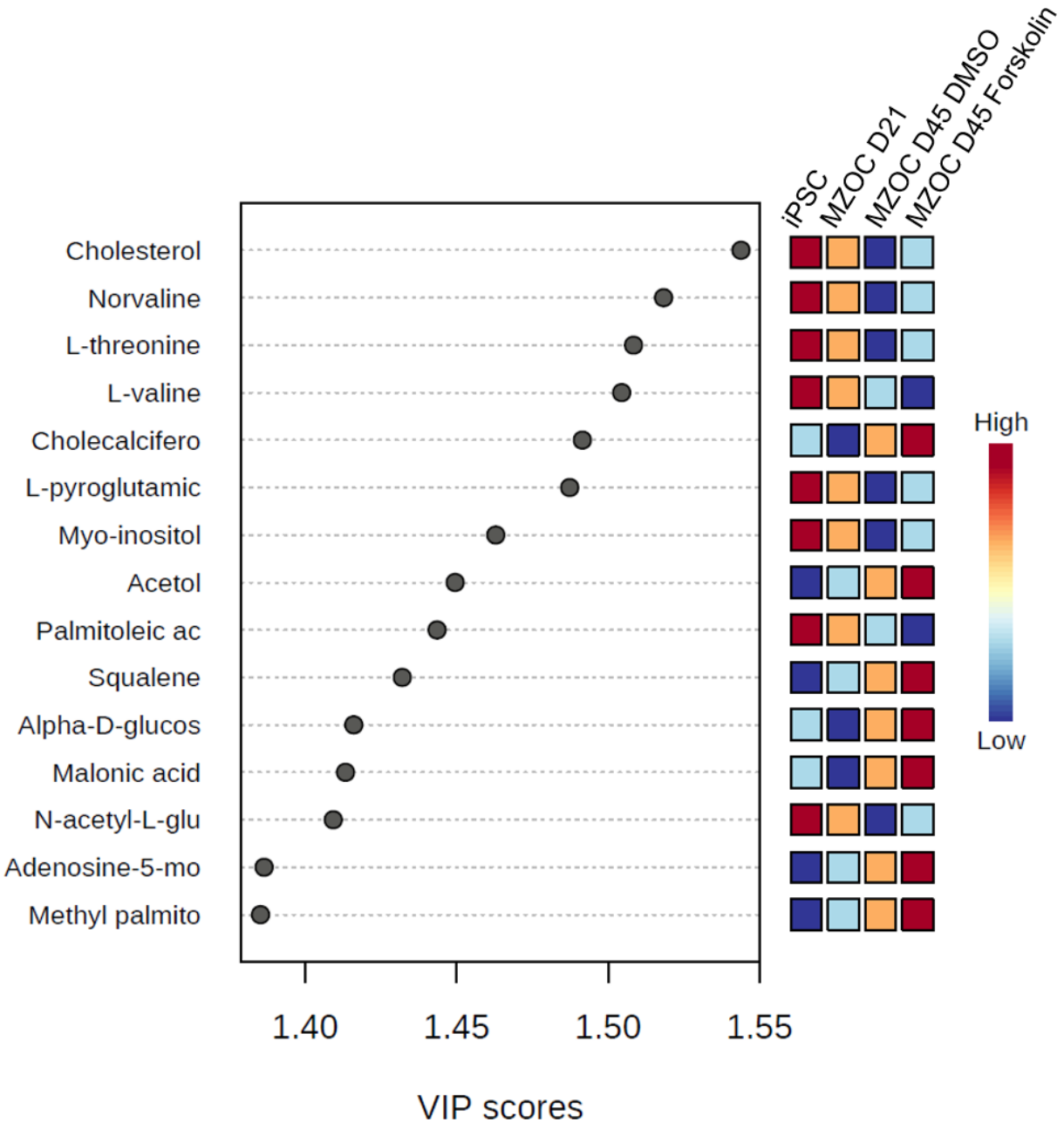


Figure 29. Important features of cell samples identified by PLS-DA. The colored boxes on the right indicate the relative concentrations of the corresponding metabolite in each group under study.

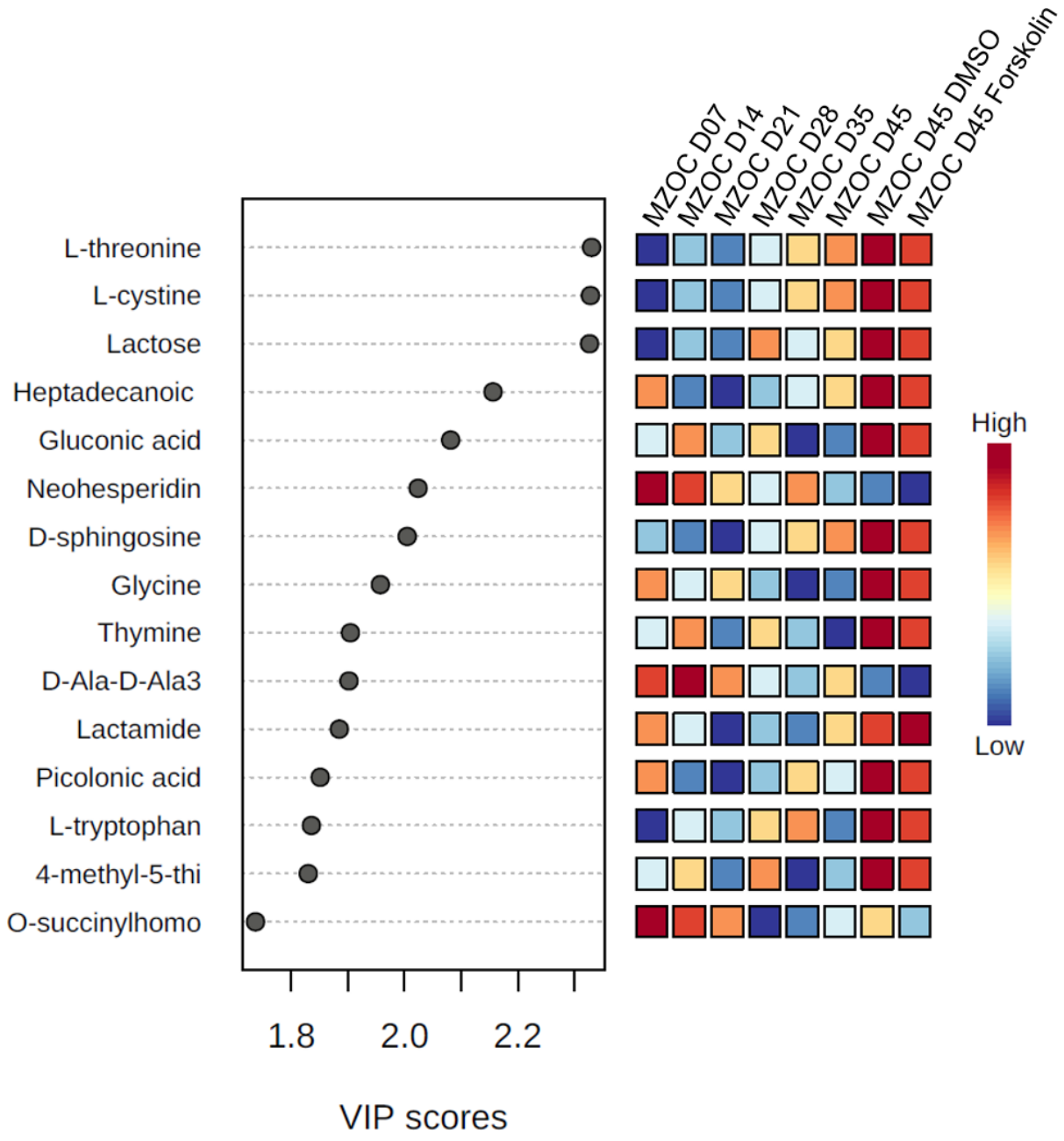


Figure 30. Important features of media samples identified by PLS-DA. The colored boxes on the right indicate the relative concentrations of the corresponding metabolite in each group under study.

4.2.3.4. *Hierarchical Clustering*

Each sample in (agglomerative) hierarchical cluster analysis starts off as a distinct cluster, and the algorithm then combines the samples until they all belong to the same cluster. When doing hierarchical clustering, two factors must be taken into account. Euclidean distance, Pearson's correlation, and Spearman's rank correlation are the first two resemblance test. The other factor is clustering algorithms, such as average linkage, complete linkage, single linkage, and Ward's linkage. The centroids of the observations are used for average linkage, the farthest pair of observations between the two groups are used for full linkage, and the closest pair of observations are used for single linkage (clustering to minimize the sum of squares of any two clusters).The dendrogram is frequently given alongside a heatmap as visual assistance. With the hclust function in the stat package, hierarchical clustering is accomplished. The heatmap representation of the clustering results is shown in Figure 31, Figure 32, Figure 33 and Figure 34.

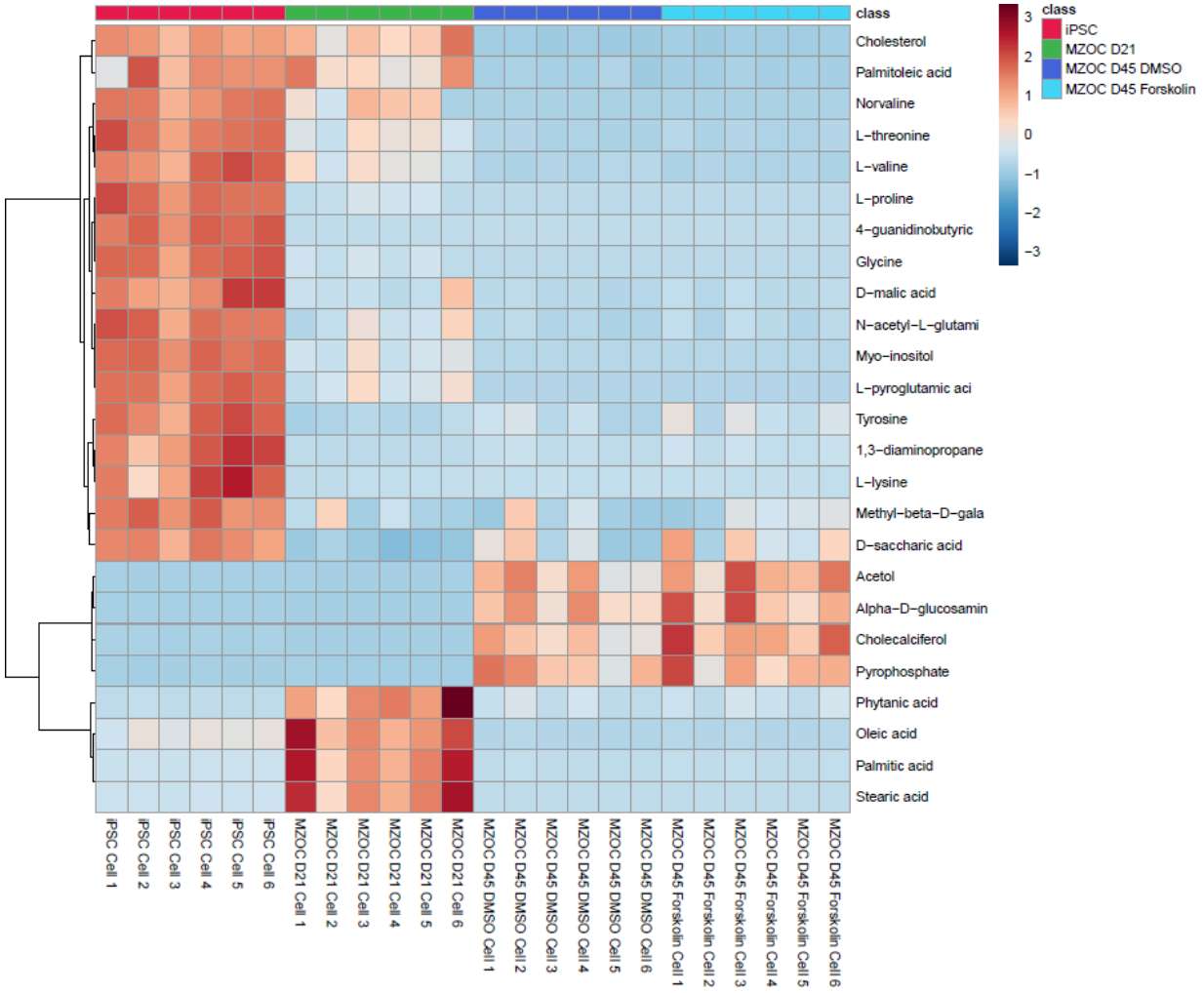


Figure 31. Clustering result of top 25 metabolites from cells shown as heatmap. (distance measure using euclidean, and clustering algorithm using ward.D).

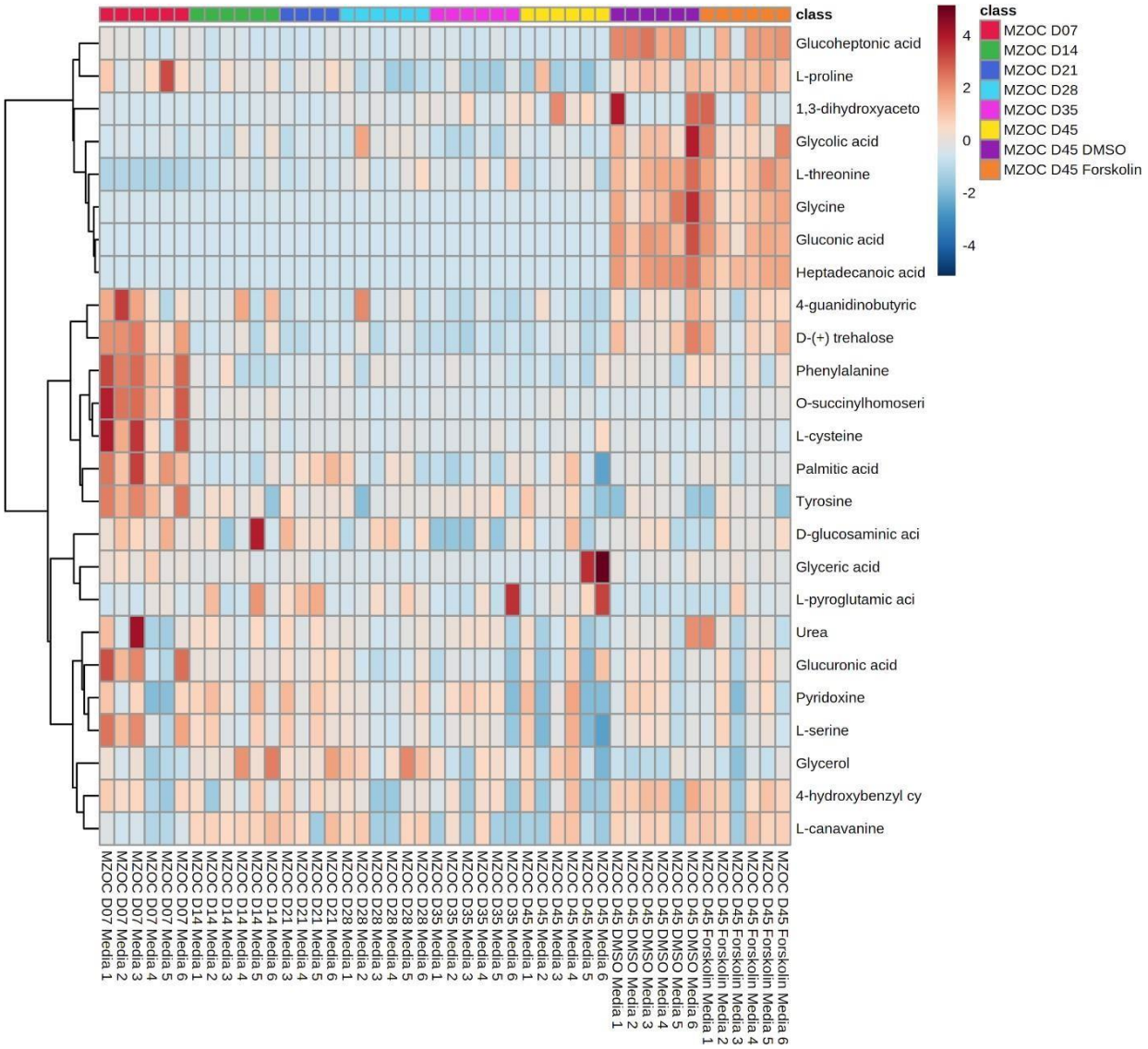
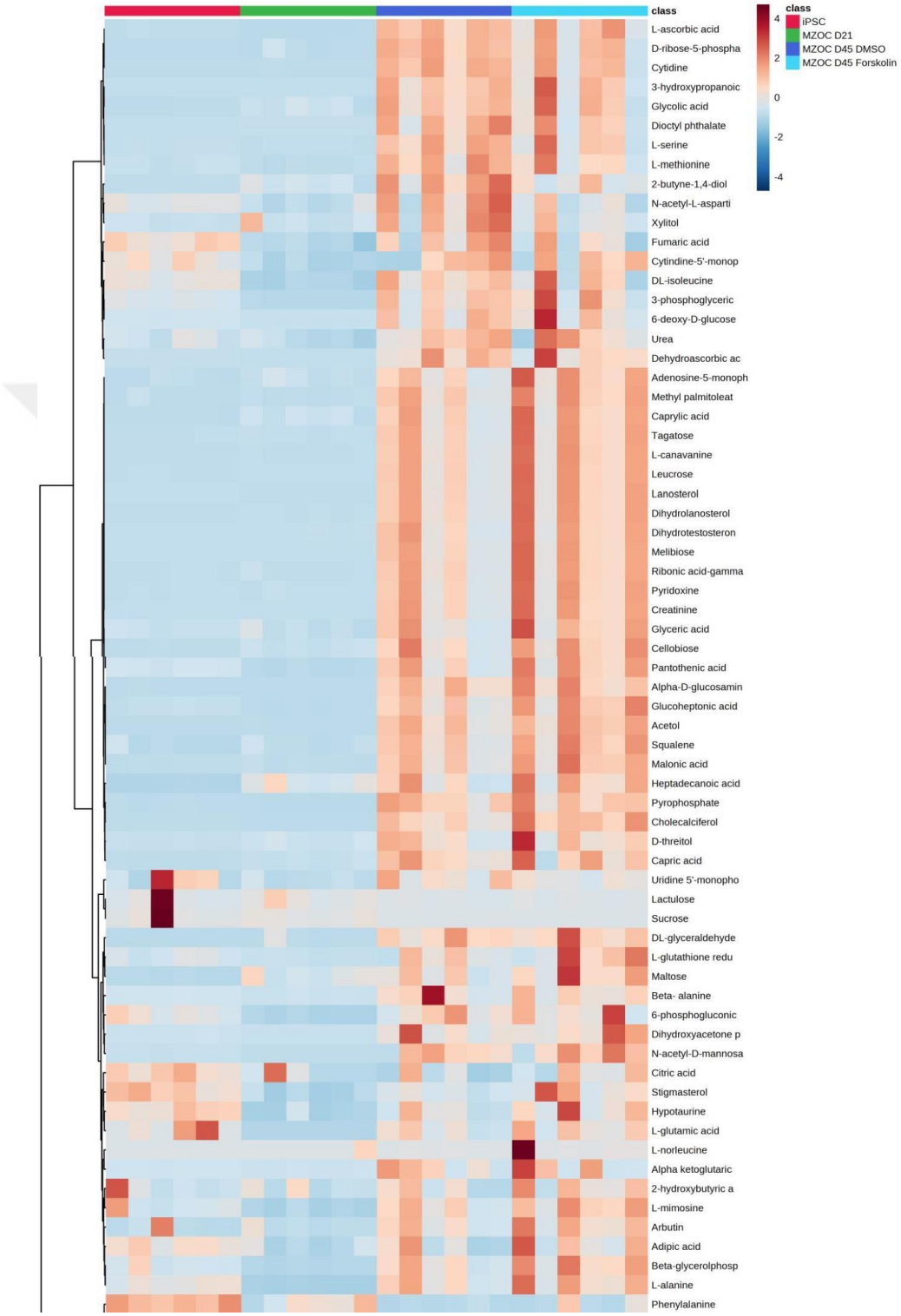


Figure 32. Clustering result of top 25 metabolites from media shown as heatmap. (distance measure using euclidean, and clustering algorithm using ward.D).


class
L-ascorbic acid
D-ribose-5-phospha
Cytidine
3-hydroxypropanoic
Glycolic acid
Dioctyl phthalate
L-serine
L-methionine
2-butyne-1,4-diol
N-acetyl-L-asparti
Xylitol
Fumaric acid
Cytindine-5'-monop
DL-isoleucine
3-phosphoglyceric
6-deoxy-D-glucose
Urea
Dehydroascorbic ac
Adenosine-5-monoph
Methyl palmitoleat
Caprylic acid
Tagatose
L-canavanine
Leucrose
Lanosterol
Dihydrolanosterol
Dihydrotestosteron
Melibiose
Ribonic acid-gamma
Pyridoxine
Creatinine
Glyceric acid
Cellobiose
Pantothenic acid
Alpha-D-glucosamin
Glucoheptonic acid
Acetol
Squalene
Malonic acid
Heptadecanoic acid
Pyrophosphate
Cholecalciferol
D-threitol
Capric acid
Uridine 5'-monopho
Lactulose
Sucrose
DL-glyceraldehyde
L-glutathione redu
Maltose
Beta- alanine
6-phosphogluconic
Dihydroxyacetone p
N-acetyl-D-mannosa
Citric acid
Stigmasterol
Hypotaurine
L-glutamic acid
L-norleucine
Alpha ketoglutaric
2-hydroxybutyric a
L-mimosine
Arbutin
Adipic acid
Beta-glycerolphosp
L-alanine
Phenylalanine
class
iPSC
MZOC D21
MZOC D45 DMSO
MZOC D45 Forskolin
4
2
0
-2
-4

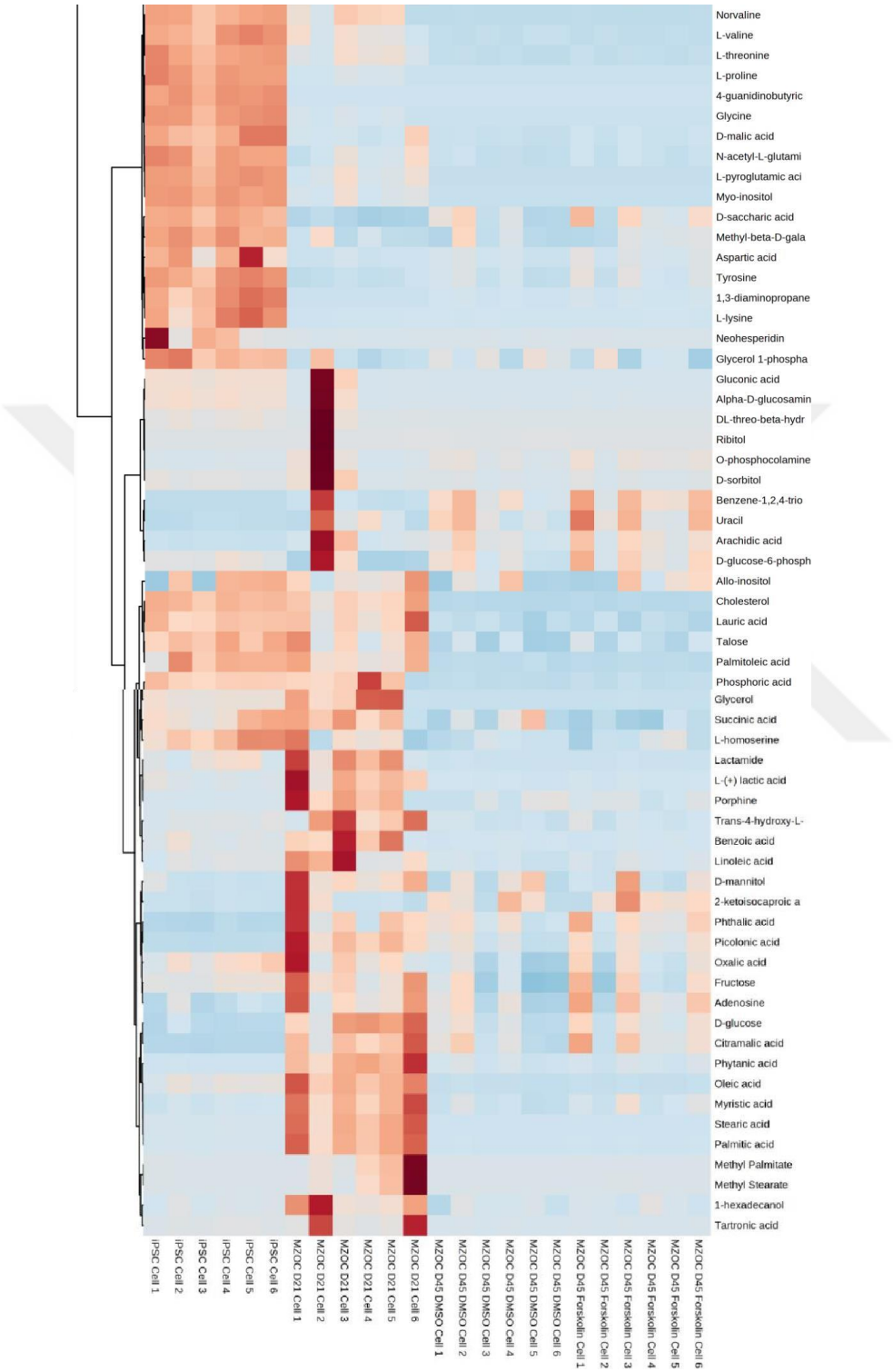
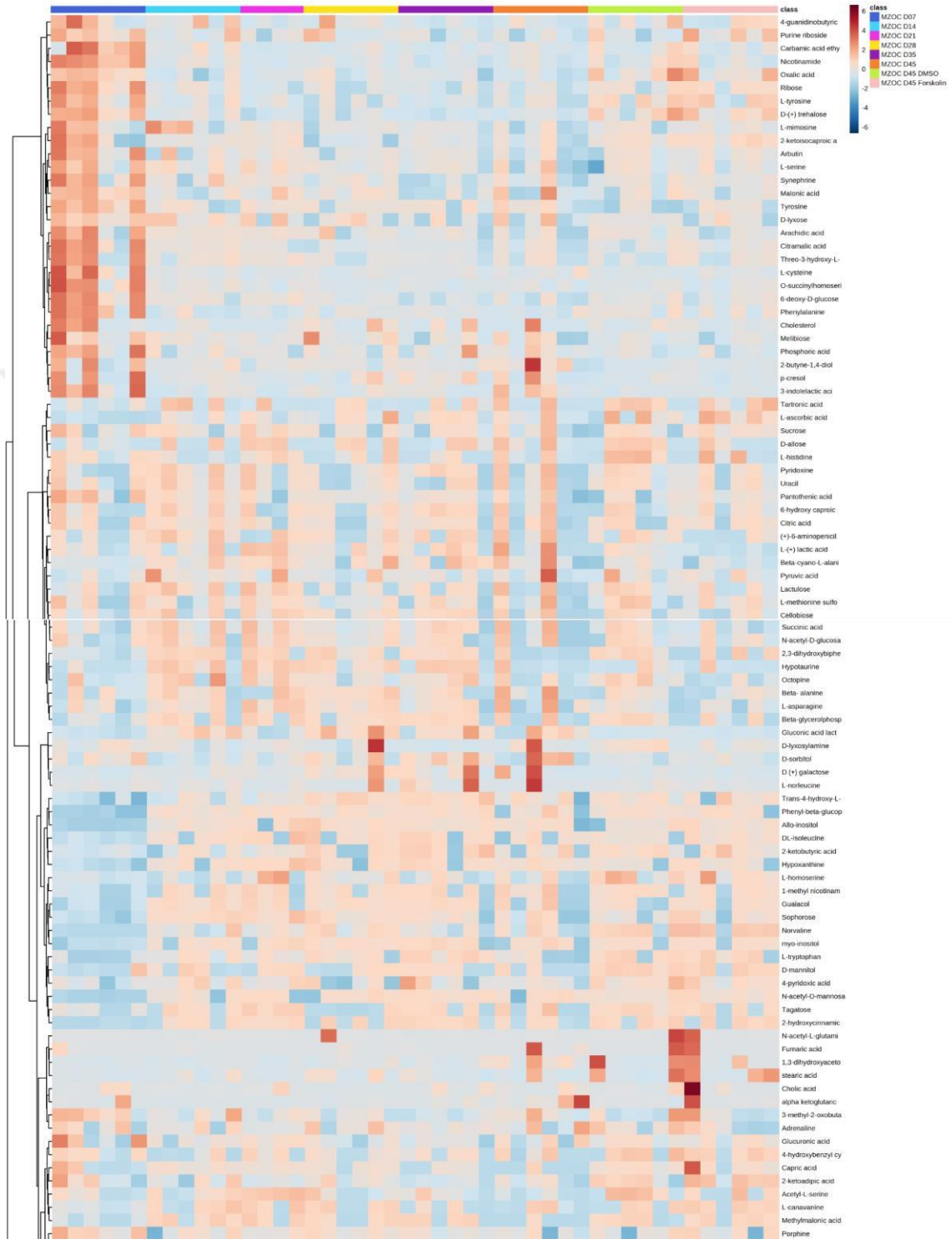


Figure 33. Clustering result of all metabolites from cells shown as heatmap. (distance measure using euclidean, and clustering algorithm using ward.D).


class
4-guanidinobutyric
Purine riboside
Carbamic acid ethy
Nicotinamide
Oxalic acid
Ribose
L-tyrosine
D-(+) trehalose
L-mimosine
2-ketoisocaproic a
Arbutin
L-serine
Synephrine
Malonic acid
Tyrosine
D-lyxose
Arachidic acid
Citramalic acid
Threo-3-hydroxy-L-
L-cysteine
O-succinylhomoseri
6-deoxy-D-glucose
Phenylalanine
Cholesterol
Melibiose
Phosphoric acid
2-butyne-1,4-diol
p-cresol
3-indolelactic aci
Tartronic acid
L-ascorbic acid
Sucrose
D-allose
L-histidine
Pyridoxine
Uracil
Pantothenic acid
6-hydroxy caproic
Citric acid
(+)-6-aminopenicil
L-(+) lactic acid
Beta-cyano-L-alani
Pyruvic acid
Lactulose
L-methionine sulfo
Cellobiose
Succinic acid
N-acetyl-D-glucosa
2,3-dihydroxybiphe
Hypotaurine
Octopine
Beta- alanine
L-asparagine
Beta-glycerolphosp
Gluconic acid lact
D-lyxosylamine
D-sorbitol
D (+) galactose
L-norleucine
Trans-4-hydroxy-L-
Phenyl-beta-glucop
Allo-inositol
DL-isoleucine
2-ketobutyric acid
Hypoxanthine
L-homoserine
1-methyl nicotinam
Gualacol
Sophorose
Norvaline
myo-inositol
L-tryptophan
D-mannitol
4-pyridoxic acid
N-acetyl-D-mannosa
Tagatose
2-hydroxycinnamic
N-acetyl-L-glutami
Fumaric acid
1,3-dihydroxyaceto
stearic acid
Cholic acid
alpha ketoglutaric
3-methyl-2-oxobuta
Adrenaline
Glucuronic acid
4-hydroxybenzyl cy
Capric acid
2-ketoadipic acid
Acetyl-L-serine
L-canavanine
Methylmalonic acid
Porphine
class
MZOC D07
MZOC D14
MZOC D21
MZOC D28
MZOC D35
MZOC D45
MZOC D45 DMSO
MZOC D45 Forskolin
6
4
2
0
-2
-4
-6

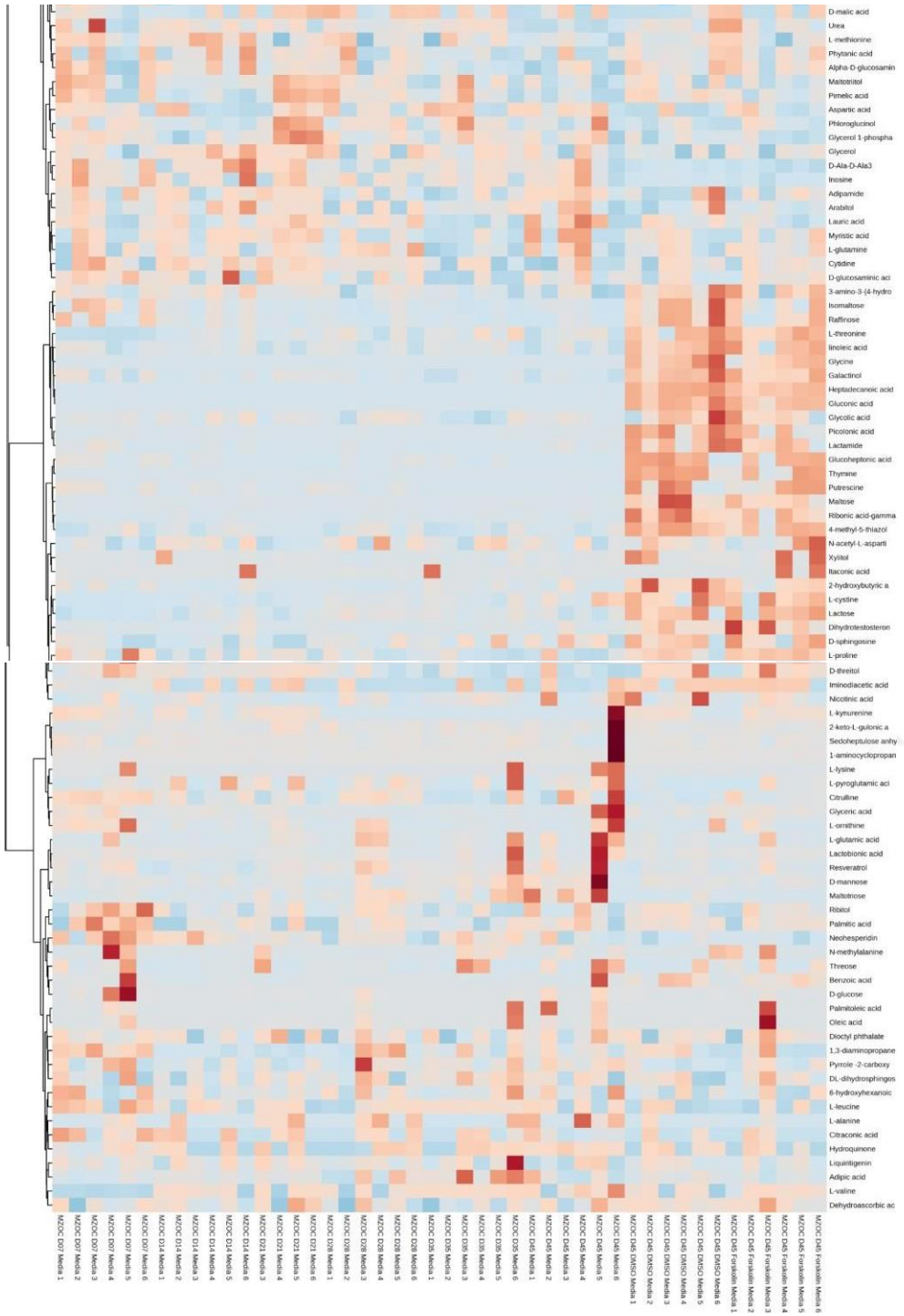


Figure 34. Clustering result of all metabolites from media shown as heatmap. (distance measure using euclidean, and clustering algorithm using ward.D)

4.2.4. Enrichment Analysis

4.2.4.1. *Data Input*

There are three enrichment analysis algorithms offered by MSEA. Accordingly, three different types of data inputs are required by these three approaches:

- A list of important compound names - entered as a one-column data (Over Representation Analysis (ORA));
- A single measured biofluid (urine, blood, CSF) sample- entered as tab-separated two-column data with the first column for compound name, and the second for concentration values (Single Sample Profiling (SSP));
- A compound concentration table - entered as a .csv file with each sample per row and each metabolite concentration per column. The first column is for sample names and the second column is for sample phenotype labels (Quantitative Enrichment Analysis (QEA))

4.2.4.2. *Data Processing*

The compound labels must be standardized as a preliminary step. Since the compound labels will later be compared to compounds in the metabolite set library, it is a crucial step. To transform between compound common names, synonyms, and identifiers used in HMDB ID, PubChem, ChEBI, BiGG, METLIN, KEGG, or Reactome, MSEA includes an integrated tool. The results of the conversion are shown in Table 7-12 for both cell and media metabolites. Note: 0 indicates no match, 1 represents a precise match, 2 suggests an approximate match. You may get a text file with the results by downloading the file with the name map .csv

Checking the concentration readings is the next step. Blood and CSF samples' concentrations must be calculated in umol for SSP analysis. To compare the urine concentrations with published amounts in the literature, the urinary concentrations must first be converted to umol/mmol creatinine. In SSP analysis, no values that are missing or negative are permitted. For QEA analysis, the concentration data is more adaptable.

Users have the option of uploading normalized data or the original concentration data. For QEA, missing or negative values are acceptable (coded as NA).

4.2.4.3. *Selection of Metabolite Set Library*

Before proceeding to enrichment analysis, a metabolite set library has to be chosen. There are seven built-in libraries offered by MSEA:

- Metabolic pathway associated metabolite sets (currently contains 99 entries);
- Disease associated metabolite sets (reported in blood) (currently contains 344 entries);
- Disease associated metabolite sets (reported in urine) (currently contains 384 entries)
- Disease associated metabolite sets (reported in CSF) (currently contains 166 entries)
- Metabolite sets associated with SNPs (currently contains 4598 entries)
- Predicted metabolite sets based on computational enzyme knockout model (currently contains 912 entries)
- Metabolite sets based on locations (currently contains 73 entries)
- Drug pathway associated metabolite sets (currently contains 461 entries)

Additionally, MSEA enables the upload of user-defined metabolite sets to do enrichment analysis on any assortment of groups of compounds that scientists choose to examine. The metabolite set library is only a two-column, .csv text file that lists the names of the metabolite sets in the first column and their compounds in the second column (HMDB compound names must be used). Please be aware that the built-in libraries mostly include research on humans. The functional classification of metabolites might not be accurate. Therefore, users are advised to upload their own self-defined metabolite set libraries for enrichment analysis for data from subjects other than humans.

When a list of compound names is given, Over Representation Analysis (ORA) is conducted. To see if any biologically significant patterns can be found, the list of chemicals

can be compiled using standard feature selection techniques, a clustering tool, or substances with aberrant amounts found in SSP.

To determine if a certain metabolite set is more strongly represented in the supplied compound list than would be predicted by chance, ORA was developed using the hypergeometric test. After accounting for repeated testing, one-tailed p values are given. Figure 35-40 summarizes the outcome for both cell and media metabolites.

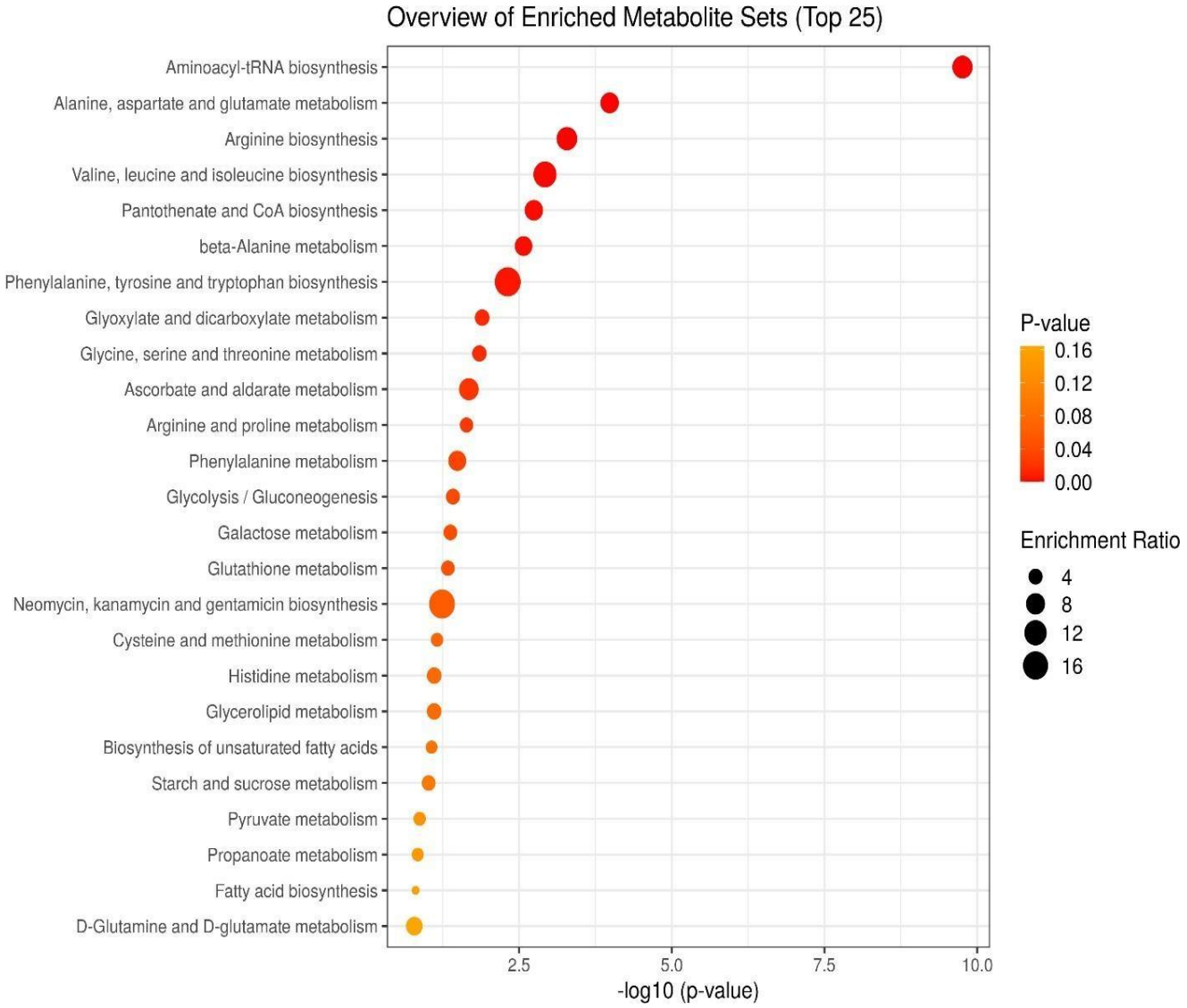


Figure 35. Enrichment analysis results of iPSC vs D21 cell metabolites.

Table 7. Results of over representation analysis of iPSC vs D21 cell metabolites.

Metabolite Set	Total	Hits	Expect	P value	FDR
Aminoacyl-tRNA biosynthesis	48	13	1.41	1.74×10^{-10}	1.46×10^{-8}
Alanine, aspartate and glutamate metabolism	28	6	0.82	1.04×10^{-4}	4.39×10^{-2}
Arginine biosynthesis	14	4	0.41	5.20×10^{-4}	0.01
Valine, leucine and isoleucine biosynthesis	8	3	0.23	1.19×10^{-2}	0.02
Pantothenate and CoA biosynthesis	19	4	0.56	1.81×10^{-2}	0.03
beta-Alanine metabolism	21	4	0.61	2.67×10^{-2}	0.04
Phenylalanine, tyrosine and tryptophan biosynthesis	4	2	0.18	4.85×10^{-2}	0.06
Glyoxylate and dicarboxylate metabolism	32	4	0.94	0.01	0.13
Glycine, serine and threonine metabolism	33	4	0.97	0.01	0.13
Ascorbate and aldarate metabolism	8	2	0.23	0.02	0.17
Arginine and proline metabolism	38	4	1.11	0.02	0.17
Phenylalanine metabolism	10	2	0.29	0.03	0.23
Glycolysis / Gluconeogenesis	26	3	0.76	0.04	0.25
Galactose metabolism	27	3	0.79	0.04	0.25

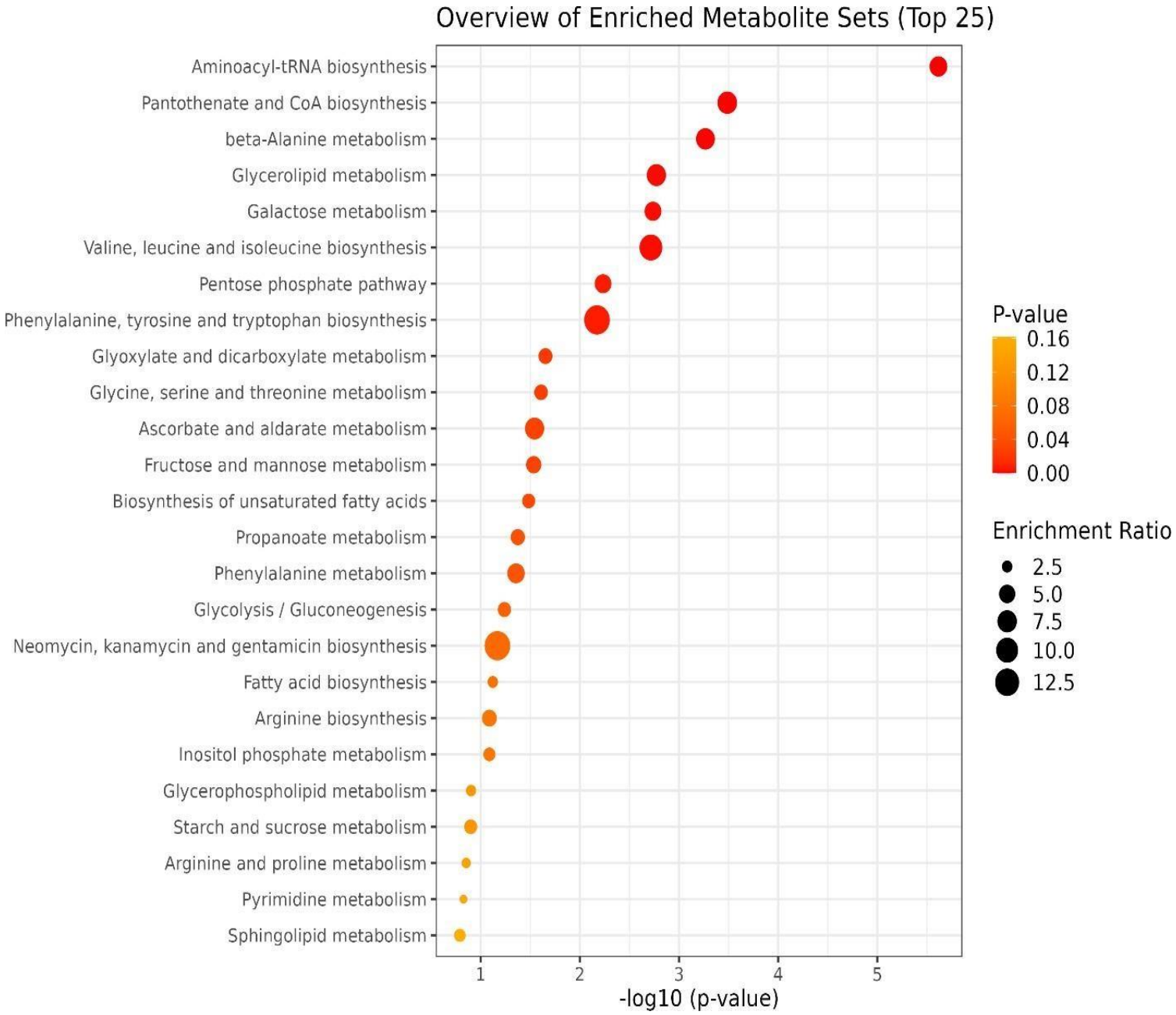


Figure 36. Enrichment analysis results of iPSC vs D45 forskolin cell metabolites.

Table 8. Results of over representation analysis of iPSC vs D45 forskolin cell metabolites.

Metabolite Set	Total	Hits	Expect	P value	FDR
Aminoacyl-tRNA biosynthesis	48	10	1.66	2.42×10^{-6}	2.03×10^{-4}
Pantothenate and CoA biosynthesis	19	5	0.67	3.26×10^{-4}	0.01
beta-Alanine metabolism	21	5	0.72	5.41×10^{-4}	0.01
Glycerolipid metabolism	16	4	0.55	1.69×10^{-2}	0.03
Galactose metabolism	27	5	0.93	1.83×10^{-2}	0.03
Valine, leucine and isoleucine biosynthesis	8	3	0.28	1.92×10^{-2}	0.03
Pentose phosphate pathway	22	4	0.76	5.83×10^{-2}	0.07
Phenylalanine, tyrosine and tryptophan biosynthesis	4	2	0.14	6.71×10^{-2}	0.07
Glyoxylate and dicarboxylate metabolism	32	4	1.10	0.02	0.20
Glycine, serine and threonine metabolism	33	4	1.14	0.02	0.20
Ascorbate and aldarate metabolism	8	2	0.27	0.03	0.20
Fructose and mannose metabolism	20	3	0.69	0.03	0.20
Biosynthesis of unsaturated fatty acids	36	4	1.24	0.03	0.21
Propanoate metabolism	23	3	0.8	0.04	0.25
Phenylalanine metabolism	10	2	0.34	0.04	0.25

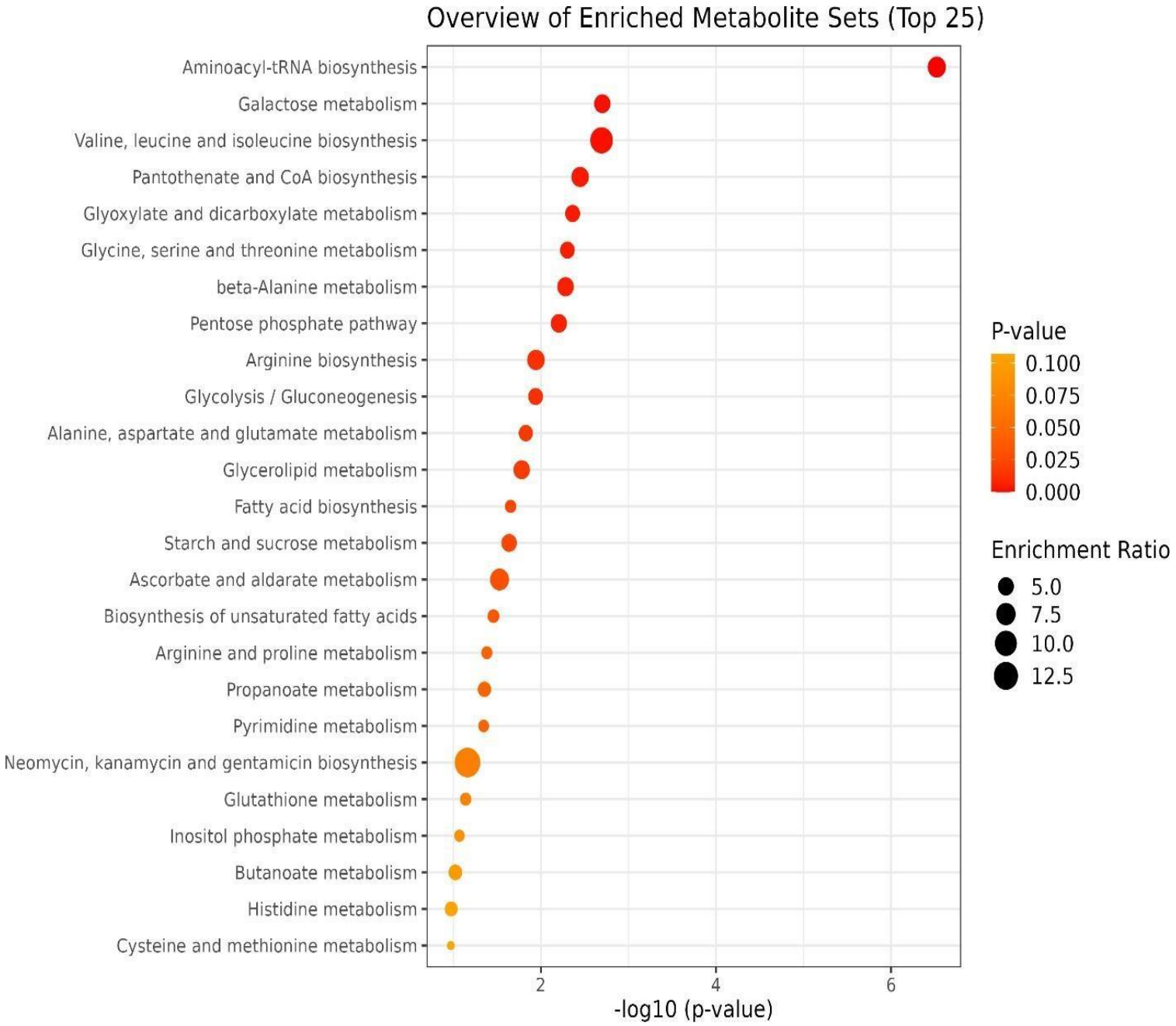


Figure 37. Enrichment analysis results of D21 vs D45 forskolin cell metabolites.

Table 9 Results of over representation analysis of D21 vs D45 forskolin cell metabolites.

Metabolite Set	Total	Hits	Expect	P value	FDR
Aminoacyl-tRNA biosynthesis	48	11	1.69	$3.00x10^{-7}$	$2.52x10^{-5}$
Galactose metabolism	27	5	0.95	$2.00x10^{-2}$	0.05
Valine, leucine and isoleucine biosynthesis	8	3	0.28	$2.00x10^{-2}$	0.05
Pantothenate and CoA biosynthesis	19	4	0.67	$3.00x10^{-2}$	0.06
Glyoxylate and dicarboxylate metabolism	32	5	1.12	$4.00x10^{-2}$	0.06
Glycine, serine and threonine metabolism	33	5	1.16	$5.00x10^{-2}$	0.06
beta-Alanine metabolism	21	4	0.74	$5.00x10^{-2}$	0.06
Pentose phosphate pathway	22	4	0.77	$6.00x10^{-2}$	0.06
Arginine biosynthesis	14	3	0.49	0.01	0.10
Glycolysis / Gluconeogenesis	26	4	0.91	0.01	0.10
Alanine, aspartate and glutamate metabolism	28	4	0.98	0.01	0.11
Glycerolipid metabolism	16	3	0.56	0.02	0.12
Fatty acid biosynthesis	47	5	1.65	0.02	0.14
Starch and sucrose metabolism	18	3	0.63	0.02	0.14
Ascorbate and aldarate metabolism	8	2	0.28	0.03	0.17
Biosynthesis of unsaturated fatty acids	36	4	1.27	0.03	0.18
Arginine and proline metabolism	38	4	1.34	0.04	0.20
Propanoate metabolism	23	3	0.80	0.04	0.20
Pyrimidine metabolism	39	4	1.37	0.04	0.20

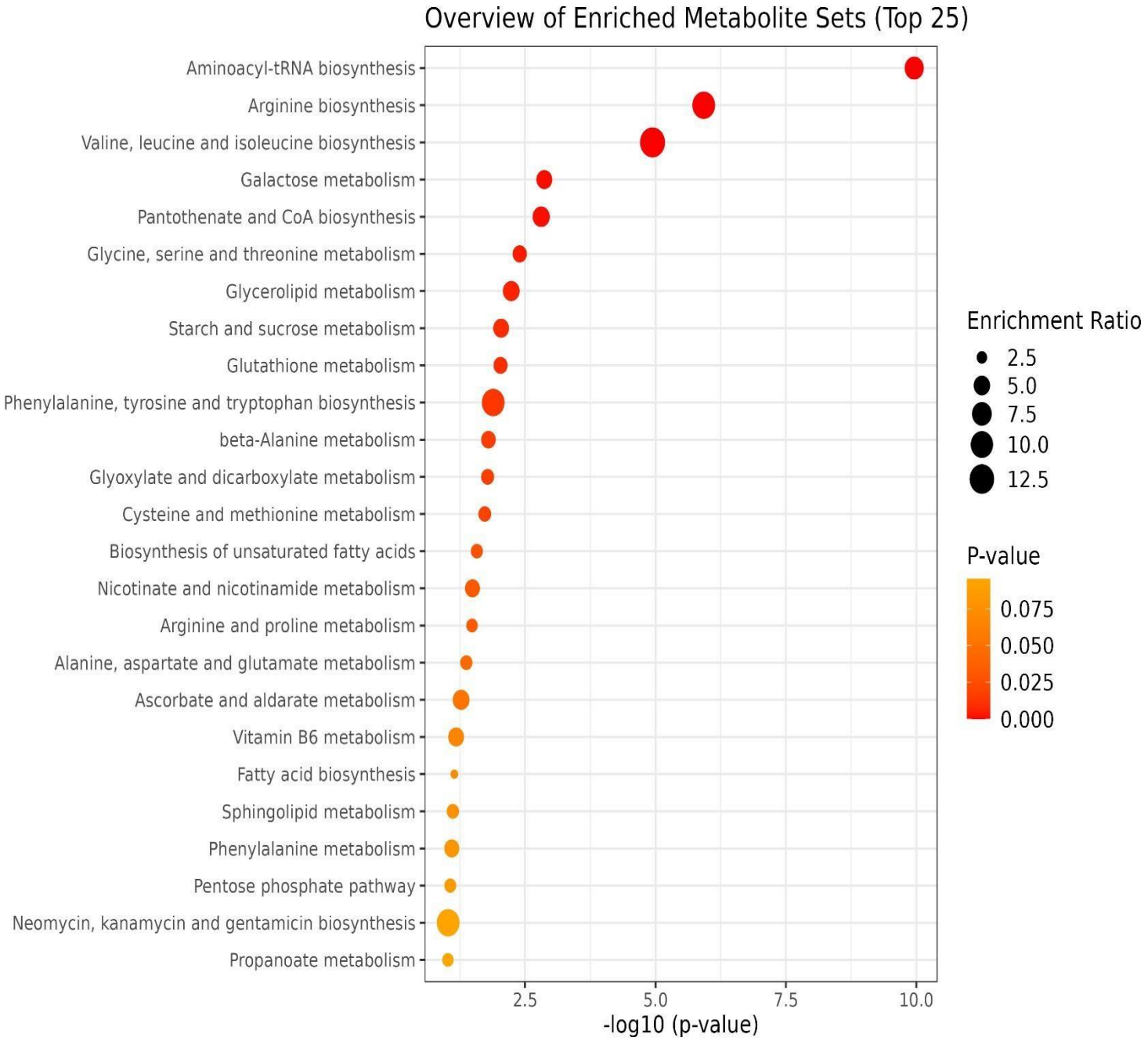


Figure 38. Enrichment analysis results of iPSCs vs D21 media metabolites.

Table 10. Results of over representation analysis of iPSCs vs D21 media metabolites.

Metabolite Set	Total	Hits	Expect	P value	FDR
Aminoacyl-tRNA biosynthesis	48	16	2.31	1.10×10^{-10}	9.23×10^{-9}
Arginine biosynthesis	14	7	0.67	1.19×10^{-6}	5.00×10^{-5}
Valine, leucine and isoleucine biosynthesis	8	5	0.38	1.14×10^{-5}	3.18×10^{-4}
Galactose metabolism	27	6	1.3	1.35×10^{-3}	0.03
Pantothenate and CoA biosynthesis	19	5	0.91	1.55×10^{-3}	0.03
Glycine, serine and threonine metabolism	33	6	1.59	4.00×10^{-2}	0.06
Glycerolipid metabolism	16	4	0.78	5.00×10^{-2}	0.07
Starch and sucrose metabolism	18	4	0.87	9.10×10^{-2}	0.08
Alanine, aspartate and glutamate metabolism	28	5	1.35	9.30×10^{-2}	0.08
Glutathione metabolism	28	5	1.35	9.30×10^{-2}	0.08
Phenylalanine, tyrosine and tryptophan biosynthesis	4	2	0.19	0.01	0.01
beta-Alanine metabolism	21	4	1.01	0.01	0.10
Glyoxylate and dicarboxylate metabolism	32	5	1.54	0.02	0.10
Cysteine and methionine metabolism	33	5	1.59	0.02	0.11
Biosynthesis of unsaturated fatty acids	36	5	1.73	0.02	0.15
Nicotinate and nicotinamide metabolism	15	3	0.72	0.03	0.16
Arginine and proline metabolism	38	5	1.83	0.03	0.16

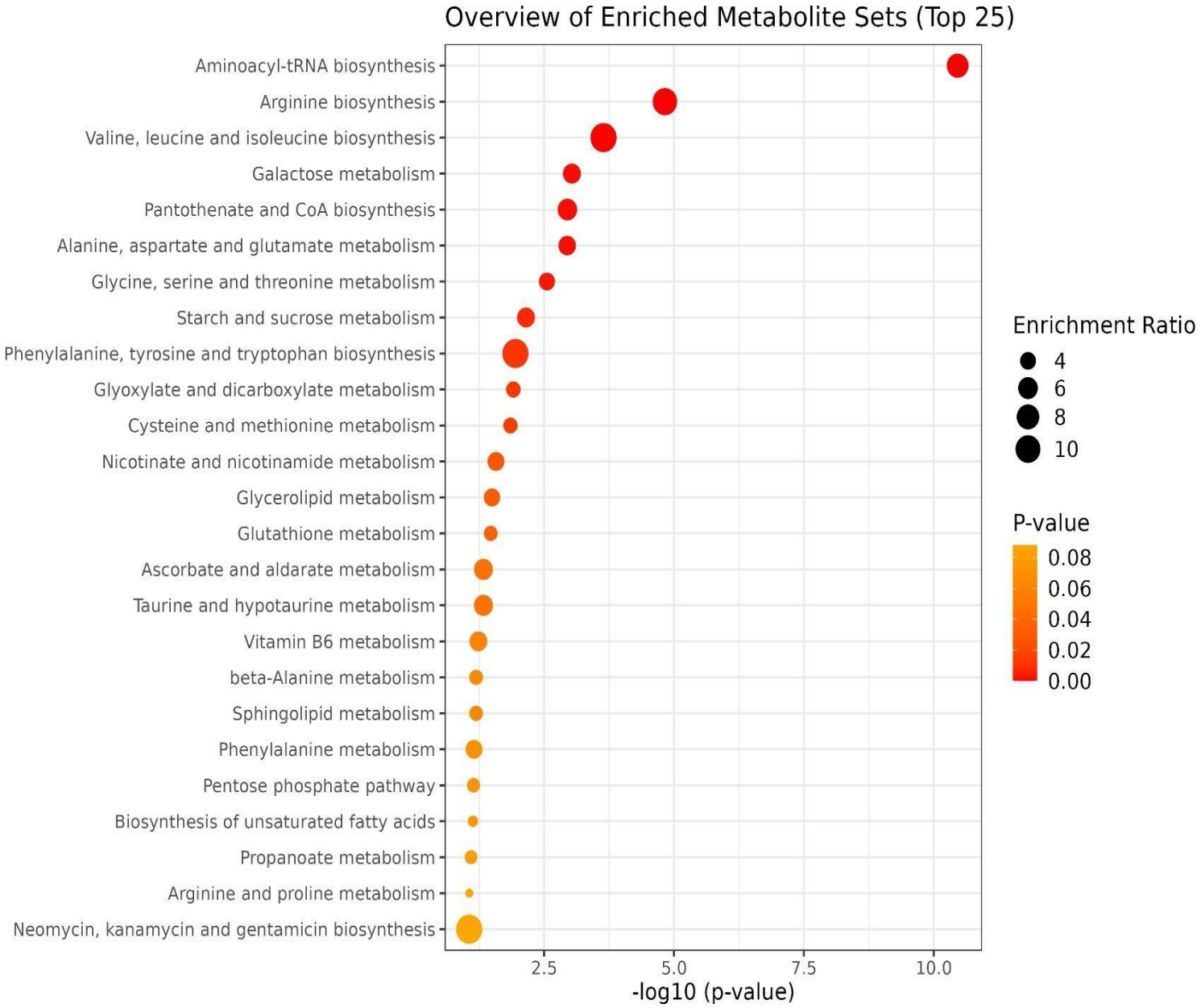


Figure 39. Enrichment analysis results of iPSCs vs D45 forskolin media metabolites.

Table 11. Results of over representation analysis of iPSCs vs D45 forskolin media metabolites.

Metabolite Set	Total	Hits	Expect	P value	FDR
Aminoacyl-tRNA biosynthesis	48	16	2.12	$2.72x10^{-11}$	$2.28x10^{-9}$
Arginine biosynthesis	14	6	0.62	$1.37x10^{-4}$	$5.76x10^{-5}$
Valine, leucine and isoleucine biosynthesis	8	4	0.35	$2.15x10^{-4}$	$6x10^{-2}$
Galactose metabolism	27	6	1.2	$8.55x10^{-4}$	0.01
Alanine, aspartate and glutamate metabolism	28	6	1.24	$1.05x10^{-3}$	0.01
Pantothenate and CoA biosynthesis	19	5	0.84	$1.05x10^{-3}$	0.01
Glycine, serine and threonine metabolism	33	6	1.46	$2.59x10^{-3}$	0.03
Starch and sucrose metabolism	18	4	0.79	$6.72x10^{-3}$	0.07
Phenylalanine, tyrosine and tryptophan biosynthesis	4	2	0.17	$1.09x10^{-2}$	0.10
Glyoxylate and dicarboxylate metabolism	32	5	1.42	$1.16x10^{-2}$	0.10
Cysteine and methionine metabolism	33	5	1.46	0.01	0.10
Nicotinate and nicotinamide metabolism	15	3	0.66	0.02	0.18
Glycerolipid metabolism	16	3	0.70	0.03	0.19
Glutathione metabolism	28	4	1.24	0.03	0.19
Ascorbate and aldarate metabolism	8	2	0.35	0.04	0.23
Taurine and hypotaurine metabolism	8	2	0.35	0.04	0.24

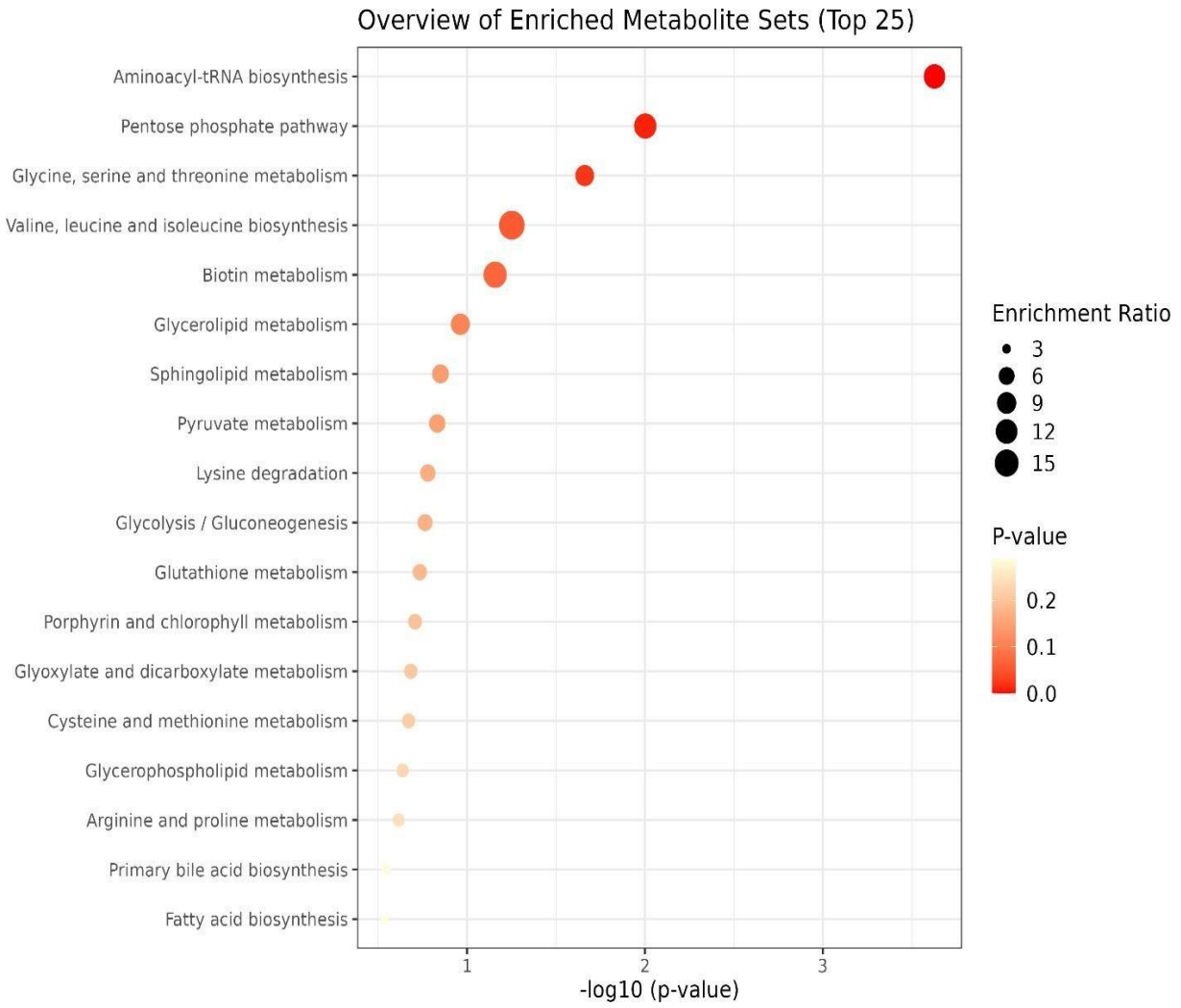


Figure 40. Enrichment analysis results of D21 vs D45 forskolin media metabolites.

Table 12. Results of over representation analysis of D21 vs D45 forskolin media metabolites.

Metabolite Set	Total	Hits	Expect	P value	FDR
Aminoacyl-tRNA biosynthesis	48	4	0.34	2.36×10^{-4}	0.02
Pentose phosphate pathway	22	2	0.15	9.96×10^{-3}	0.41
Glycine, serine and threonine metabolism	33	2	0.23	0.02	0.61

5. DISCUSSION

Today, omics techniques are routinely used as high-throughput methods in the diagnosis and prognosis of diseases, and the study of biological processes in basic research. The newest omics technology that can perform the analysis of all low molecular weight compounds in a sample is metabolomics. The biggest advantage of this method is that it allows the establishment of a closer link between the cell/tissue phenotype under certain environmental conditions at a specified time.

Previously, ocular cells were generated from hiPSCs by our lab. However, the phenotype of the generated LG organoids is not well known. In this thesis, methods have been developed in order to understand LG organoids better. So, we performed metabolomics analyses on the generated LG organoids. Matrigel was used in all of the experiments as a supportive ECM component to help better enhance stem cell development and imitate the natural environment. In metabolomics analyses, the same protocol was used for all samples.

5.1. Differentiation of hiPSCs into Lacrimal Gland Organoids

First, in 2016 Hayashi et al. generated self-formed ectodermal autonomous multi-zone (SEAM) from hiPSCs. They showed Concentric SEAM imitates whole-eye development because cell placement within various zones, including the ocular surface ectoderm, lens, neuro-retina, and retinal pigment epithelium, is suggestive of lineage (Hayashi et al., 2016). Then in 2019 Li et al. developed a different differentiation protocol for hiPSCs into MZOCs. Their protocol showed instead of four zones appearing during SEAM induction, their induction approach generated five zones, comprising the neural retina, RPE, surface ectoderm, neural crest, and lens. Each of the five zones displayed distinct differentiation fates for ocular cells, and all five showed separate distribution and apparent boundaries. Additionally, while all five zones of the MZOCs in this study were defined within four weeks, the four-zone cells of SEAM took two to six weeks to characterize. However, in both studies, they cultured the generated organoids for up to

four weeks. In this study, we go for further time points like 45 days in order to observe the LG organoids at later maturation levels (Figure 9).

5.2. Metabolomic Analysis of Lacrimal Gland Organoids

After differentiation of hiPSCs to LG organoids, generated organoids were taken for sample preparation at certain time points (Days 21, and 45 from cell samples for intracellular metabolites and days 7, 14, 21, 28, 35, and 45 from media samples for extracellular metabolites). Additionally, forskolin stimulation is done on day 45 to investigate the effect on the secretion of the LG organoids. Different stimulants or small molecules such as carbachol, FGF10, BMP7, etc. can alter the metabolome of LG organoids variously. Sample preparation was performed according to GC-MS-based metabolomics sample preparation.

In order to analyze the metabolic phenotypes of samples we performed untargeted metabolomics both on cell and media samples. A total of 571 metabolites were found in cell samples and 1186 metabolites were found in media samples. 128 of them were identified in cell samples and 179 of them were identified in media samples by using databases (Table 3 and 4). Multivariate data analyses were performed on the normalized data. First, whether there is a systematic error or an outlier in the data set was observed with the PCA score graph (Figures 19 and 20). PLS-DA analyses were performed separately for both conditions in order to examine the differentiation between groups in more detail. PLS-DA plots showed that LG organoids had different metabolomic profiles in both conditions, depending on their maturation process. Metabolic profile alteration at different time points occurs due to the development process of organoids. Because, the energy need of organoids increases during the development process and different metabolisms, especially energy metabolisms and metabolites regarding those metabolisms are increasing. (Figures 23 and 24). VIP score graphs were used to find the metabolites causing this differentiation (Figures 29 and 30). Since the PLS-DA method is a forced analysis method, its accuracy needs to be demonstrated. For this purpose, the reliability (R^2) and predictability (Q^2) values of the PLS-DA models are used. The R^2 and Q^2 values obtained for the different models R^2: 0.9268 and Q^2: 0.8715 (cell samples) and

R^2: 0.8783 and Q^2: 0.7006 (media samples) are given in Figures 27 and 28. The fact that the obtained R^2 and Q^2 values are greater than 0.5 indicates the validity of the methods.

Enrichment analyses show us several metabolic pathways underwent significant alteration in both cell and media samples. Such as valine, leucine, and isoleucine (BCAAs) biosynthesis or glycine, serine, and threonine metabolism. The metabolic flux of purine biosynthesis and the Krebs cycle are both positively and negatively influenced by glycine (Cheng et al., 2019). The mammalian target of rapamycin (mTOR) signaling pathway is primarily activated by the BCAAs to carry out a variety of metabolic and signaling tasks. Additionally, they act as substrates for the production of energy or proteins (Nie et al., 2018). As LG organoids maturate over time, their energy metabolisms increased as expected (Figure 35-40).

In the literature, several studies show the substances of the human tear. As their results indicates, healthy human tears consists of amino acids (1-methylhistidine/ 3-methylhistidine, L-(+)-arginine, dimethylarginine, citrullina, creatine, glutamic acid, glutamine, homo-L-arginine, oxyproline, phenylalanine, proline, pyroglutamic acid, serine, aminoethylsulfonic acid, threonine, tryptophan, tyrosine, urocanate, valine), amino alcohols (bepantol), amino ketones (alantan, Kreatinin), aromatic acids (cinnamate, 2-coumarinate / m-hydroxycinnamate / p-coumaric acid), carbohydrates (aceneuramate), nucleotides (adenosine pyrophosphate, 3'-adenylic acid, cytidine monophosphate, cytidine diphosphate choline, guanosine monophosphate, inosine monophosphate, uridine diphosphate, uridine monophosphate, UDP-N-acetyl-D-mannosamine / UDP-N-Acetyl-D-galactosamine), carnitines (branigen, carnitine, hexanoic acid, L-palmitoylcarnitine), cyclic amines (3-pyridinecarboxamide), dicarboxylic acids (lichenate / toxilicate), nucleosides (1-methyladenosine, boniton, posilent, guanosine, inosine, formycinylhomocysteine, s-adenosyl-methionine, uridine, 9-D-ribofuranosylxanthine), peptides (glutathione disulphide), phospholipids (L-palmitoyllysolecithin), purines and derivatives (inosine, 3,7-dimethylxanthine, lithate, xanthine), pyridoxals and derivatives (4-pyridoxinsaeure), quaternary amines (choline acetic acid, gliatilin, phosphocholine) and finally tricarboxylic acids (aciletten). (Khanna et al., 2022; Chen et al., 2011; Nakatsukasa

et al., 2011; Choy et al., 2001; Gogia et al., 1998; Mendelsohn et al., 1998). In our study, we found similar results. 26 metabolites (pyridoxine, hypotaurine, l-homoserine, l-norleucine, adenosine-5-monophosphate, citric acid, l-glutamic acid, fumaric acid, palmitoleic acid, l-ascorbic acid, creatinine, l-serine, cytidine, l-lysine, uridine 5'-monophosphate, tyrosine, adenosine, l-proline, l-valine, l-threonine, n-acetyl-l-aspartic acid, n-acetyl-l-glutamic acid, l-pyroglutamic acid, glycine, dl-isoleucine, phenylalanine) out of 128 cell metabolites shows parallel outcome with the literature (Figure 33).

6. CONCLUSION and FUTURE DIRECTIONS

In conclusion, it has been shown by this research that, hiPSCs can be differentiated into LG organoids *in vitro* conditions. Characteristics of developed organoid samples showed similarity with human LG, in terms of protein expressions, as reported in the literature on human tear samples in terms of metabolite profile and secretion function. Untargeted GC-MS-based metabolomic analyses can be performed on generated LG organoids and culture media at different maturation levels successfully. Additionally, the secretion profile of generated LG organoids was further investigated with forskolin stimulation. Forskolin alters the metabolome of organoids significantly. The results might vary with different components and analysis methods. Besides, although only the WT1-hiPSC cell line was used in this study, other iPSC cell types can also be investigated. Furthermore, we could not perform metabolomics analysis on the native LG tissue to compare our generated LG organoids. Despite the performed analyses, there are still many unknown metabolites waiting to be annotated. More studies could be performed focusing on this issue such as the integration of multi-omics and utilizing different more advanced instrumentation could be used in future studies.

7. REFERENCES

Akbari, S., Sevinç, G. G., Ersoy, N., Basak, O., Kaplan, K., Sevinç, K., Ozel, E., Sengun, B., Enustun, E., Ozcimen, B., Bagriyanik, A., Arslan, N., Önder, T. T., & Erdal, E. (2019). Robust, Long-Term Culture of Endoderm-Derived Hepatic Organoids for Disease Modeling. *Stem Cell Reports*, *13*(4), 627-641. https://doi.org/10.1016/j.stemcr.2019.08.007

Alasbahi RH, Melzig MF. Plectranthus barbatus: a review of phytochemistry, ethnobotanical uses and pharmacology - part 2. Planta Med. 2010 May;76(8):753-65. doi: 10.1055/s-0029-1240919

All systems go!. *Nat Cell Biol* **8**, 1179 (2006). https://doi.org/10.1038/ncb1106-1179

Bai, Q., Desprat, R., Klein, B., Lemaitre, J.-M., & De Vos, J. (2013). Embryonic stem cells or induced pluripotent stem cells? A DNA integrity perspective. *Current Gene Therapy*, *13*(2), 93-98.

Bannier-Hélaouët, M., Post, Y., Korving, J., Trani Bustos, M., Gehart, H., Begthel, H., Bar-Ephraim, Y. E., van der Vaart, J., Kalmann, R., Imhoff, S. M., & Clevers, H. (2021). Exploring the human lacrimal gland using organoids and single-cell sequencing. *Cell Stem Cell*, *28*(7), 1221-1232.e7. https://doi.org/10.1016/j.stem.2021.02.024

Bartels SP, Lee SR, Neufeld AH. The effects of forskolin on cyclic AMP, intraocular pressure and aqueous humor formation in rabbits. Curr Eye Res. 1987 Feb;6(2):307-20. doi: 10.3109/02713688709025183

Beckonert, O., Keun, H., Ebbels, T. *et al.* Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat Protoc* 2, 2692-2703 (2007). https://doi.org/10.1038/nprot.2007.376

Bijlsma et al.Large-Scale Human Metabolomics Studies: A Strategy for Data (Pre-) Processing and Validation, Anal Chem. 2006, 78 567 – 574

Blaise BJ, Correia G, Tin A, Young JH, Vergnaud AC, Lewis M, Pearce JT, Elliott P, Nicholson JK, Holmes E, Ebbels TM. Power Analysis and Sample Size

Determination in Metabolic Phenotyping. Anal Chem. 2016 May 17;88(10):5179-88. doi: 10.1021/acs.analchem.6b00188

Brown M, Dunn WB, Dobson P, Patel Y, Winder CL, Francis-McIntyre S, Begley P, Carroll K, Broadhurst D, Tseng A, Swainston N, Spasic I, Goodacre R, Kell DB. Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. Analyst. 2009 Jul;134(7):1322-32. doi: 10.1039/b901179j

Celebi, A. R. C., Ulusoy, C., & Mirza, G. E. (2014). The efficacy of autologous serum eye drops for severe dry eye syndrome: A randomized double-blind crossover study. *Graefe's Archive for Clinical and Experimental Ophthalmology = Albrecht Von Graefes Archiv Fur Klinische Und Experimentelle Ophthalmologie*, *252*(4), 619–626. https://doi.org/10.1007/s00417-014-2599-1

Chagastelles, P. C., & Nardi, N. B. (2011). Biology of stem cells: An overview. *Kidney International Supplements*, *1*(3), 63–67. https://doi.org/10.1038/kisup.2011.15

Chen L, Zhou L, Chan EC, Neo J, Beuerman RW. Characterization of the human tear metabolome by LC-MS/MS. J Proteome Res. 2011 Oct 7;10(10):4876-82. doi: 10.1021/pr2004874

Cheng ZX, Guo C, Chen ZG, Yang TC, Zhang JY, Wang J, Zhu JX, Li D, Zhang TT, Li H, Peng B, Peng XX. Glycine, serine and threonine metabolism confounds efficacy of complement-mediated killing. Nat Commun. 2019 Jul 25;10(1):3325. doi: 10.1038/s41467-019-11129-5

Chin, M. H., Pellegrini, M., Plath, K., & Lowry, W. E. (2010). Molecular analyses of human induced pluripotent stem cells and embryonic stem cells. *Cell Stem Cell*, *7*(2), 263–269. https://doi.org/10.1016/j.stem.2010.06.019

Choy CK, Cho P, Chung WY, Benzie IF. Water-soluble antioxidants in human tears: effect of the collection method. Invest Ophthalmol Vis Sci. 2001 Dec;42(13):3130-4

Christians U, Klawitter J, Hornberger A, Klawitter J. How unbiased is non-targeted metabolomics and is targeted pathway screening the solution? Curr Pharm Biotechnol. 2011 Jul;12(7):1053-66. doi: 10.2174/138920111795909078

Chuang HY, Hofree M, Ideker T. A decade of systems biology. Annu Rev Cell Dev Biol. 2010;26:721-44. doi: 10.1146/annurev-cellbio-100109-104122

Clevers H. Modeling Development and Disease with Organoids. Cell. 2016 Jun 16;165(7):1586-1597. doi: 10.1016/j.cell.2016.05.082

Conrady, C. D., Joos, Z. P., & Patel, B. C. K. (2016). Review: The Lacrimal Gland and Its Role in Dry Eye. Journal of Ophthalmology, 2016, 7542929. https://doi.org/10.1155/2016/7542929

Curry EL, Moad M, Robson CN, Heer R. Using induced pluripotent stem cells as a tool for modelling carcinogenesis. World J Stem Cells. 2015 Mar 26;7(2):461-9. doi: 10.4252/wjsc.v7.i2.461

de la Cuadra-Blanco C, Peces-Peña MD, Mérida-Velasco JR. Morphogenesis of the human lacrimal gland. J Anat. 2003 Nov;203(5):531-6. doi: 10.1046/j.1469-7580.2003.00233.x

Dietrich, J., & Schrader, S. (2020). Towards Lacrimal Gland Regeneration: Current Concepts and Experimental Approaches. *Current Eye Research*, *45*(3), 230–240. https://doi.org/10.1080/02713683.2019.1637438

Doi, A., Park, I.-H., Wen, B., Murakami, P., Aryee, M. J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., Miller, J., Schlaeger, T., Daley, G. Q., & Feinberg, A. P. (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nature Genetics*, *41*(12), 1350–1353. https://doi.org/10.1038/ng.471

Downie, L. E., Bandlitz, S., Bergmanson, J. P. G., Craig, J. P., Dutta, D., Maldonado-Codina, C., Ngo, W., Siddireddy, J. S., & Wolffsohn, J. S. (2021). CLEAR - Anatomy and physiology of the anterior eye. *Contact Lens & Anterior Eye: The Journal of the British Contact Lens Association*, *44*(2), 132–156. https://doi.org/10.1016/j.clae.2021.02.009

Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N, Brown M, Knowles JD, Halsall A, Haselden JN, Nicholls AW, Wilson ID, Kell DB, Goodacre R; Human Serum Metabolome (HUSERMET) Consortium. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat Protoc. 2011 Jun 30;6(7):1060-83. doi: 10.1038/nprot.2011.335

Dunn WB. Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes. Phys Biol. 2008 Feb 20;5(1):011001. doi: 10.1088/1478-3975/5/1/011001

Fiehn O. Metabolomics by Gas Chromatography-Mass Spectrometry: Combined Targeted and Untargeted Profiling. Curr Protoc Mol Biol. 2016 Apr 1;114:30.4.1-30.4.32. doi: 10.1002/0471142727.mb3004s114

Fiehn O. Metabolomics--the link between genotypes and phenotypes. Plant Mol Biol. 2002 Jan;48(1-2):155-71

Fradkin JE, Cook GH, Kilhoffer MC, Wolff J. Forskolin stimulation of thyroid adenylate cyclase and cyclic 3',5'-adenosine monophosphate accumulation. Endocrinology. 1982 Sep;111(3):849-56. doi: 10.1210/endo-111-3-849

Garg, A., & Zhang, X. (2017). Lacrimal gland development: From signaling interactions to regenerative medicine. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *246*(12), 970–980. https://doi.org/10.1002/dvdy.24551

Gogia R, Richer SP, Rose RC. Tear fluid content of electrochemically active components including water soluble antioxidants. Curr Eye Res. 1998 Mar;17(3):257-63. doi: 10.1076/ceyr.17.3.257.5213

Guenther, M. G., Frampton, G. M., Soldner, F., Hockemeyer, D., Mitalipova, M., Jaenisch, R., & Young, R. A. (2010). Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. *Cell Stem Cell*, *7*(2), 249–257. https://doi.org/10.1016/j.stem.2010.06.015

H. Doerfler, D. Lyon, T. Nägele, X. Sun, L. Fragner, F. Hadacek, V. Egelhofer, W.Weckwerth, Granger causality in integrated GC–MS and LC–MSmetabolomics data reveals the interface of primary and secondarymetabolism, Metabolomics 9 (3) (2013) 564–574

Hackstadt AJ, Hess AM.Filtering for increased power for microarray data analysis, BMC Bioinformatics. 2009; 10:11

Hayashi R, Ishikawa Y, Sasamoto Y, Katori R, Nomura N, Ichikawa T, Araki S, Soma T, Kawasaki S, Sekiguchi K, Quantock AJ, Tsujikawa M, Nishida K. Co-ordinated ocular development from human iPS cells and recovery of corneal function. Nature. 2016 Mar 17;531(7594):376-80. doi: 10.1038/nature17000

Hayashi S, Yoshida M, Fujiwara T, Maegawa S, Fukusaki E. Single-embryo metabolomics and systematic prediction of developmental stage in zebrafish. Z Naturforsch C J Biosci. 2011 Mar-Apr;66(3-4):191-8. doi: 10.1515/znc-2011-3-414

Hayashi, R., Okubo, T., Kudo, Y., Ishikawa, Y., Imaizumi, T., Suzuki, K., Shibata, S., Katayama, T., Park, S.-J., Young, R. D., Quantock, A. J., & Nishida, K. (2022). Generation of 3D lacrimal gland organoids from human pluripotent stem cells. *Nature*, *605*(7908), 126–131. https://doi.org/10.1038/s41586-022-04613-4

Hirayama, M., Kawakita, T., Tsubota, K., & Shimmura, S. (2016). Challenges and Strategies for Regenerating the Lacrimal Gland. *The Ocular Surface*, *14*(2), 135–143. https://doi.org/10.1016/j.jtos.2015.11.005

Hirayama, M., Ogawa, M., Oshima, M., Sekine, Y., Ishida, K., Yamashita, K., Ikeda, K., Shimmura, S., Kawakita, T., Tsubota, K., & Tsuji, T. (2013). Functional lacrimal gland regeneration by transplantation of a bioengineered organ germ. *Nature Communications*, *4*(1), 2497. https://doi.org/10.1038/ncomms3497

Hodges, R. R., & Dartt, D. A. (2003). Regulatory pathways in lacrimal gland epithelium. *International Review of Cytology*, *231*, 129–196. https://doi.org/10.1016/s0074-7696(03)31004-6

Holly FJ, Lemp MA. Tear physiology and dry eyes. Surv Ophthalmol. 1977 Sep-Oct;22(2):69-87. doi: 10.1016/0039-6257(77)90087-x

Huang SX, Islam MN, O'Neill J, Hu Z, Yang YG, Chen YW, Mumau M, Green MD, Vunjak-Novakovic G, Bhattacharya J, Snoeck HW. Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. Nat Biotechnol. 2014 Jan;32(1):84-91. doi: 10.1038/nbt.2754

Ishikawa S, Sugimoto M, Kitabatake K, Sugano A, Nakamura M, Kaneko M, Ota S, Hiwatari K, Enomoto A, Soga T, Tomita M, Iino M. Identification of salivary metabolomic biomarkers for oral cancer screening. Sci Rep. 2016 Aug 19;6:31520. doi: 10.1038/srep31520

Jeong, S. Y., Choi, W. H., Jeon, S. G., Lee, S., Park, J.-M., Park, M., Lee, H., Lew, H., & Yoo, J. (2021). Establishment of functional epithelial organoids from human lacrimal glands. *Stem Cell Research & Therapy*, *12*(1), 247. https://doi.org/10.1186/s13287-021-02133-y

Jewett MC, Hofmann G, Nielsen J. Fungal metabolite analysis in genomics and phenomics. Curr Opin Biotechnol. 2006 Apr;17(2):191-7. doi: 10.1016/j.copbio.2006.02.001

Johnson ME, Murphy PJ. Changes in the tear film and ocular surface from dry eye syndrome. Prog Retin Eye Res. 2004 Jul;23(4):449-74. doi: 10.1016/j.preteyeres.2004.04.003

Jones, P. (2001). *Bioinformatics in the post-genomic age. World Patent Information, 23(4), 349–354.* doi:10.1016/s0172-2190(01)00043-6

Jonsson P, Gullberg J, Nordström A, Kusano M, Kowalczyk M, Sjöström M, Moritz T. A strategy for identifying differences in large series of metabolomic samples analyzed by GC/MS. Anal Chem. 2004 Mar 15;76(6):1738-45. doi: 10.1021/ac0352427

Kaddurah-Daouk R, Kristal BS, Weinshilboum RM. Metabolomics: a global biochemical approach to drug response and disease. Annu Rev Pharmacol Toxicol. 2008;48:653-83. doi: 10.1146/annurev.pharmtox.48.113006.094715

Karczewski KJ, Snyder MP. Integrative omics for health and disease. Nat Rev Genet. 2018 May;19(5):299-310. doi: 10.1038/nrg.2018.4

Katona M, Vizvári E, Németh L, Facskó A, Venglovecz V, Rakonczay Z Jr, Hegyi P, Tóth-Molnár E. Experimental evidence of fluid secretion of rabbit lacrimal gland duct epithelium. Invest Ophthalmol Vis Sci. 2014 Jun 12;55(7):4360-7. doi: 10.1167/iovs.14-14025

Khanna RK, Catanese S, Emond P, Corcia P, Blasco H, Pisella PJ. Metabolomics and lipidomics approaches in human tears: A systematic review. Surv Ophthalmol. 2022 Jul-Aug;67(4):1229-1243. doi: 10.1016/j.survophthal.2022.01.010

Kim HK, Verpoorte R. Sample preparation for plant metabolomics. Phytochem Anal. 2010 Jan-Feb;21(1):4-13. doi: 10.1002/pca.1188

Kim J, Koo BK, Knoblich JA. Human organoids: model systems for human biology and medicine. Nat Rev Mol Cell Biol. 2020 Oct;21(10):571-584. doi: 10.1038/s41580-020-0259-3

Kitano H. Systems biology: a brief overview. Science. 2002 Mar 1;295(5560):1662-4. doi: 10.1126/science.1069492

Kolios, G., & Moodley, Y. (2013). Introduction to stem cells and regenerative medicine. *Respiration; International Review of Thoracic Diseases*, *85*(1), 3-10. https://doi.org/10.1159/000345615

Kraly JR, Holcomb RE, Guan Q, Henry CS. Review: Microfluidic applications in metabolomics and metabolic profiling. Anal Chim Acta. 2009 Oct 19;653(1):23-35. doi: 10.1016/j.aca.2009.08.037

Kurmann AA, Serra M, Hawkins F, Rankin SA, Mori M, Astapova I, Ullas S, Lin S, Bilodeau M, Rossant J, Jean JC, Ikonomou L, Deterding RR, Shannon JM, Zorn AM, Hollenberg AN, Kotton DN. Regeneration of Thyroid Function by Transplantation of Differentiated Pluripotent Stem Cells. Cell Stem Cell. 2015 Nov 5;17(5):527-42. doi: 10.1016/j.stem.2015.09.004

L.J. Sweetlove, A.R. Fernie, The spatial organization of metabolism within theplant cell, Annu. Rev. Plant Biol. 64 (2013) 723–746

Landry, D. W., & Zucker, H. A. (2004). Embryonic death and the creation of human embryonic stem cells. *The Journal of Clinical Investigation*, *114*(9), 1184–1186. https://doi.org/10.1172/JCI23065

Langer, R., & Vacanti, J. P. (1993). Tissue engineering. *Science (New York, N.Y.)*, *260*(5110), 920–926. https://doi.org/10.1126/science.8493529

Lanza IR, Zhang S, Ward LE, Karakelides H, Raftery D, Nair KS. Quantitative metabolomics by H-NMR and LC-MS/MS confirms altered metabolic pathways in diabetes. PLoS One. 2010 May 10;5(5):e10538. doi: 10.1371/journal.pone.0010538

Lauridsen M, Hansen SH, Jaroszewski JW, Cornett C. Human urine as test material in 1H NMR-based metabonomics: recommendations for sample preparation and storage. Anal Chem. 2007 Feb 1;79(3):1181-6. doi: 10.1021/ac061354x

Lewis GD, Asnani A, Gerszten RE. Application of metabolomics to cardiovascular biomarker and pathway discovery. J Am Coll Cardiol. 2008 Jul 8;52(2):117-23. doi: 10.1016/j.jacc.2008.03.043

Li, Z., Duan, H., Li, W., Hu, X., Jia, Y., Zhao, C., Zhang, S., Zhou, Q., & Shi, W. (2019). Rapid Differentiation of Multi-Zone Ocular Cells from Human Induced Pluripotent Stem Cells and Generation of Corneal Epithelial and Endothelial Cells. *Stem Cells and Development*, *28*(7), 454–463. https://doi.org/10.1089/scd.2018.0176

Lin L, Lin JM. Development of cell metabolite analysis on microfluidic platform. J Pharm Anal. 2015 Dec;5(6):337-347. doi: 10.1016/j.jpha.2015.09.003

Lin, H., Sun, G., He, H., Botsford, B., Li, M., Elisseeff, J. H., & Yiu, S. C. (2016). Three-Dimensional Culture of Functional Adult Rabbit Lacrimal Gland Epithelial Cells on Decellularized Scaffold. *Tissue Engineering. Part A*, *22*(1–2), 65–74. https://doi.org/10.1089/ten.TEA.2015.0286

Lindon JC, Holmes E, Nicholson JK. Metabonomics in pharmaceutical R&D. FEBS J. 2007 Mar;274(5):1140-51. doi: 10.1111/j.1742-4658.2007.05673.x

Lindon JC, Nicholson JK. Spectroscopic and statistical techniques for information recovery in metabonomics and metabolomics. Annu Rev Anal Chem (Palo Alto Calif). 2008;1:45-69. doi: 10.1146/annurev.anchem.1.031207.113026

López-López JA, Page MJ, Lipsey MW, Higgins JPT. Dealing with effect size multiplicity in systematic reviews and meta-analyses. Res Synth Methods. 2018 Jul 3. doi: 10.1002/jrsm.1310

Lu, Q., Yin, H., Grant, M. P., & Elisseeff, J. H. (2017). An In Vitro Model for the Ocular Surface and Tear Film System. Scientific Reports, 7(1), 6163. https://doi.org/10.1038/s41598-017-06369-8

M. Vinaixa, E.L. Schymanski, S. Neumann, M. Navarro, R.M. Salek, O. Yanes,Mass spectral databases for LC/MS-and GC/MS-based metabolomics: state ofthe field and future prospects, Trends Anal. Chem. 78 (2016) 23–35

Mahla RS. Stem Cells Applications in Regenerative Medicine and Disease Therapeutics. Int J Cell Biol. 2016;2016:6940283. doi: 10.1155/2016/6940283

Makarenkova, H. P., & Dartt, D. A. (2015). Myoepithelial Cells: Their Origin and Function in Lacrimal Gland Morphogenesis, Homeostasis, and Repair. *Current Molecular Biology Reports*, *1*(3), 115–123.

Mason C, Dunnill P. A brief definition of regenerative medicine. Regen Med. 2008 Jan;3(1):1-5. doi: 10.2217/17460751.3.1.1

Mastrangelo A, Ferrarini A, Rey-Stolle F, García A, Barbas C. From sample treatment to biomarker discovery: A tutorial for untargeted metabolomics based on GC-(EI)-Q-MS. Anal Chim Acta. 2015 Nov 5;900:21-35. doi: 10.1016/j.aca.2015.10.001

Max Kuhn. Contributions from Jed Wing and Steve Weston and Andre Williams. caret: Classification and Regression Training, 2008, R package version 3.45

Mendelsohn ME, Abramson DH, Senft S, Servodidio CA, Gamache PH. Uric acid in the aqueous humor and tears of retinoblastoma patients. J AAPOS. 1998 Dec;2(6):369-71. doi: 10.1016/s1091-8531(98)90037-4

Mikkelsen, T. S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B. E., Jaenisch, R., Lander, E. S., & Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. *Nature*, *454*(7200), 49–55. https://doi.org/10.1038/nature07056

Mircheff, A. K. (1989). Lacrimal fluid and electrolyte secretion: A review. *Current Eye Research*, *8*(6), 607–617. https://doi.org/10.3109/02713688908995761

Mu L, Niu Z, Blair RH, Yu H, Browne RW, Bonner MR, Fanter T, Deng F, Swanson M. Metabolomics Profiling before, during, and after the Beijing Olympics: A Panel Study of Within-Individual Differences during Periods of High and Low Air Pollution. Environ Health Perspect. 2019 May;127(5):57010. doi: 10.1289/EHP3705

Nakao, K., Morita, R., Saji, Y., Ishida, K., Tomita, Y., Ogawa, M., Saitoh, M., Tomooka, Y., & Tsuji, T. (2007). The development of a bioengineered organ germ method. *Nature Methods*, *4*(3), 227–230. https://doi.org/10.1038/nmeth1012

Nakatsukasa M, Sotozono C, Shimbo K, Ono N, Miyano H, Okano A, Hamuro J, Kinoshita S. Amino Acid profiles in human tear fluids analyzed by high-performance liquid chromatography and electrospray ionization tandem mass spectrometry. Am J Ophthalmol. 2011 May;151(5):799-808.e1. doi: 10.1016/j.ajo.2010.11.003

Nam D, Kim SY. Gene-set approach for expression pattern analysis, Briefings in Bioinformatics. 2008 9(3): 189-197.

Narsinh, K. H., Jia, F., Robbins, R. C., Kay, M. A., Longaker, M. T., & Wu, J. C. (2011). Generation of adult human induced pluripotent stem cells using nonviral minicircle DNA vectors. *Nature Protocols*, *6*(1), 78–88. https://doi.org/10.1038/nprot.2010.173

Nemutlu E, Zhang S, Gupta A, Juranic NO, Macura SI, Terzic A, Jahangir A, Dzeja P. Dynamic phosphometabolomic profiling of human tissues and transgenic models

by 18O-assisted ^{31}P NMR and mass spectrometry. Physiol Genomics. 2012 Apr 2;44(7):386-402. doi: 10.1152/physiolgenomics.00152.2011

Nemutlu E, Zhang S, Juranic NO, Terzic A, Macura S, Dzeja P. 18O-assisted dynamic metabolomics for individualized diagnostics and treatment of human diseases. Croat Med J. 2012 Dec;53(6):529-34. doi: 10.3325/cmj.2012.53.529

Nemutlu E, Zhang S, Xu YZ, Terzic A, Zhong L, Dzeja PD, Cha YM. Cardiac resynchronization therapy induces adaptive metabolic transitions in the metabolomic profile of heart failure. J Card Fail. 2015 Jun;21(6):460-9. doi: 10.1016/j.cardfail.2015.04.005

Newgard CB. Metabolomics and Metabolic Diseases: Where Do We Stand? Cell Metab. 2017 Jan 10;25(1):43-56. doi: 10.1016/j.cmet.2016.09.018

Nie C, He T, Zhang W, Zhang G, Ma X. Branched Chain Amino Acids: Beyond Nutrition Metabolism. Int J Mol Sci. 2018 Mar 23;19(4):954. doi: 10.3390/ijms19040954

Nyamundanda, G., Gormley, I.C., Fan, Y. *et al.* MetSizeR: selecting the optimal sample size for metabolomic studies using an analysis based approach. *BMC Bioinformatics* 14, 338 (2013). https://doi.org/10.1186/1471-2105-14-338

Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T., & Yamanaka, S. (2008). Generation of mouse induced pluripotent stem cells without viral vectors. *Science (New York, N.Y.)*, *322*(5903), 949–953. https://doi.org/10.1126/science.1164270

Örge FH, Boente CS. The lacrimal system. Pediatr Clin North Am. 2014 Jun;61(3):529-39. doi: 10.1016/j.pcl.2014.03.002

Paulsen, F. P., & Berry, M. S. (2006). Mucins and TFF peptides of the tear film and lacrimal apparatus. *Progress in Histochemistry and Cytochemistry*, *41*(1), 1–53. https://doi.org/10.1016/j.proghi.2006.03.001

Pinto J, Domingues MR, Galhano E, Pita C, Almeida Mdo C, Carreira IM, Gil AM. Human plasma stability during handling and storage: impact on NMR metabolomics. Analyst. 2014 Mar 7;139(5):1168-77. doi: 10.1039/c3an02188b

Ramirez, J.-M., Bai, Q., Dijon-Grinand, M., Assou, S., Gerbal-Chaloin, S., Hamamah, S., & Vos, J. D. (2010). Human pluripotent stem cells: From biology to cell therapy. *World Journal of Stem Cells*, *2*(2), 24–33. https://doi.org/10.4252/wjsc.v2.i2.24

Rochfort S. Metabolomics reviewed: a new "omics" platform technology for systems biology and implications for natural products research. J Nat Prod. 2005 Dec;68(12):1813-20. doi: 10.1021/np050255w

Roessner U, Bowne J. What is metabolomics all about? Biotechniques. 2009 Apr;46(5):363-5. doi: 10.2144/000113133.

Ron Wehrens and Bjorn-Helge Mevik.pls: Partial Least Squares Regression (PLSR) and Principal Component Regression (PCR), 2007, R package version 2.1-0

Seamon KB, Padgett W, Daly JW. Forskolin: unique diterpene activator of adenylate cyclase in membranes and in intact cells. Proc Natl Acad Sci U S A. 1981 Jun;78(6):3363-7. doi: 10.1073/pnas.78.6.3363

Spaniol, K., Metzger, M., Roth, M., Greve, B., Mertsch, S., Geerling, G., & Schrader, S. (2015). Engineering of a Secretory Active Three-Dimensional Lacrimal Gland Construct on the Basis of Decellularized Lacrimal Gland Tissue. *Tissue Engineering. Part A*, *21*(19–20), 2605–2617. https://doi.org/10.1089/ten.TEA.2014.0694

Stacklies W, Redestig H, Scholz M, Walther D, Selbig J. pcaMethods: a bioconductor package, providing PCA methods for incomplete data., Bioinformatics 2007 23(9):1164-1167

Subramanian Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles., Proc Natl Acad Sci USA. 2005 102(43): 15545-50

Taguchi A, Kaku Y, Ohmori T, Sharmin S, Ogawa M, Sasaki H, Nishinakamura R. Redefining the in vivo origin of metanephric nephron progenitors enables generation of complex kidney structures from pluripotent stem cells. Cell Stem Cell. 2014 Jan 2;14(1):53-67. doi: 10.1016/j.stem.2013.11.010

Tai MC, Cosar CB, Cohen EJ, Rapuano CJ, Laibson PR. The clinical efficacy of silicone punctal plug therapy. Cornea. 2002 Mar;21(2):135-9. doi: 10.1097/00003226-200203000-00001

Takahashi, K., & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, *126*(4), 663-676. https://doi.org/10.1016/j.cell.2006.07.024

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007a). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, *131*(5), 861-872. https://doi.org/10.1016/j.cell.2007.11.019

Takebe, T., Sekine, K., Enomura, M. *et al.* Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature* 499, 481–484 (2013). https://doi.org/10.1038/nature12271

Thomson M, Liu SJ, Zou LN, Smith Z, Meissner A, Ramanathan S. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. Cell. 2011 Jun 10;145(6):875-89. doi: 10.1016/j.cell.2011.05.017

Tian, H., Li, B. & Shui, G. Untargeted LC-MS Data Preprocessing in Metabolomics. *J. Anal. Test.* **1**, 187–192 (2017). https://doi.org/10.1007/s41664-017-0030-8

Tiwari, S., Nair, R. M., Vamadevan, P., Ali, M. J., Naik, M. N., Honavar, S. G., & Vemuganti, G. K. (2018). Establishing and characterizing lacrispheres from human lacrimal gland for potential clinical application. *Graefe's Archive for Clinical and Experimental Ophthalmology = Albrecht Von Graefes Archiv Fur Klinische Und Experimentelle Ophthalmologie*, *256*(4), 717-727. https://doi.org/10.1007/s00417-018-3926-8

Villas-Bôas SG, Bruheim P. The potential of metabolomics tools in bioremediation studies. OMICS. 2007 Fall;11(3):305-13. doi: 10.1089/omi.2007.0005

W. Weckwerth, Metabolomics: methods and protocols, Humana Pr Inc(2007).

Walcott, B. (1998). The Lacrimal Gland and Its Veil of Tears. *News in Physiological Sciences: An International Journal of Physiology Produced Jointly by the*

International Union of Physiological Sciences and the American Physiological Society, *13*, 97–103. https://doi.org/10.1152/physiologyonline.1998.13.2.97

Warren, L., Manos, P. D., Ahfeldt, T., Loh, Y.-H., Li, H., Lau, F., Ebina, W., Mandal, P. K., Smith, Z. D., Meissner, A., Daley, G. Q., Brack, A. S., Collins, J. J., Cowan, C., Schlaeger, T. M., & Rossi, D. J. (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell*, *7*(5), 618–630. https://doi.org/10.1016/j.stem.2010.08.012

Wheeler AR, Throndset WR, Whelan RJ, Leach AM, Zare RN, Liao YH, Farrell K, Manger ID, Daridon A. Microfluidic device for single-cell analysis. Anal Chem. 2003 Jul 15;75(14):3581-6. doi: 10.1021/ac0340758

Wieczorek, R., Jakobiec, F. A., Sacks, E. H., & Knowles, D. M. (1988). The immunoarchitecture of the normal human lacrimal gland. Relevancy for understanding pathologic conditions. *Ophthalmology*, *95*(1), 100–109. https://doi.org/10.1016/s0161-6420(88)33228-8

Wishart DS, Knox C, Guo AC, Eisner R, Young N, Gautam B, Hau DD, Psychogios N, Dong E, Bouatra S, Mandal R, Sinelnikov I, Xia J, Jia L, Cruz JA, Lim E, Sobsey CA, Shrivastava S, Huang P, Liu P, Fang L, Peng J, Fradette R, Cheng D, Tzur D, Clements M, Lewis A, De Souza A, Zuniga A, Dawe M, Xiong Y, Clive D, Greiner R, Nazyrova A, Shaykhutdinov R, Li L, Vogel HJ, Forsythe I. HMDB: a knowledgebase for the human metabolome. Nucleic Acids Res. 2009 Jan;37(Database issue):D603-10. doi: 10.1093/nar/gkn810

Yao, Y., & Zhang, Y. (2017). The lacrimal gland: Development, wound repair and regeneration. *Biotechnology Letters*, *39*(7), 939–949. https://doi.org/10.1007/s10529-017-2326-1

Ye, L., Swingen, C., & Zhang, J. (2013). Induced Pluripotent Stem Cells and Their Potential for Basic and Clinical Sciences. *Current Cardiology Reviews*, *9*(1), 63–72. https://doi.org/10.2174/157340313805076278

Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II, Thomson JA. Induced pluripotent

stem cell lines derived from human somatic cells. Science. 2007 Dec 21;318(5858):1917-20. doi: 10.1126/science.1151526

Zeki ÖC, Eylem CC, Reçber T, Kır S, Nemutlu E. Integration of GC-MS and LC-MS for untargeted metabolomics profiling. J Pharm Biomed Anal. 2020 Oct 25;190:113509. doi: 10.1016/j.jpba.2020.113509

Zoukhri D. Mechanisms involved in injury and repair of the murine lacrimal gland: role of programmed cell death and mesenchymal stem cells. Ocul Surf. 2010 Apr;8(2):60-9. doi: 10.1016/s1542-0124(12)70070-8. Erratum in: Ocul Surf. 2010 Jul;8(3):134

Zoukhri, D., Fix, A., Alroy, J., & Kublin, C. L. (2008). Mechanisms of Murine Lacrimal Gland Repair after Experimentally Induced Inflammation. *Investigative Ophthalmology & Visual Science*, *49*(10), 4399–4406. https://doi.org/10.1167/iovs.08-1730

8. APPENDIX

8.1. Ethics Committee Approval

İZMİR BİYOTIP VE GENOM MERKEZİ
GİRİŞİMSEL OLMAYAN ARAŞTIRMALAR ETİK KURULU (İBG-GOEK)
KARARI

Toplantı Tarihi	: 26/07/2021	**Toplantı Günü**	: Pazartesi
Toplantı Sayısı	: 8	**Toplantı Saati**	: 10:30

Sayın Doç. Dr. Sinan GÜVEN,

2021-022 Protokol No'lu; sorumlusu olduğunuz "İnsan indüklenmiş pluripotent kök hücre kökenli lakrimal bezlerin metabolomik analizi" başlıklı araştırmanın uygulanmasında etik açıdan sakınca olmadığına oy birliği ile karar verilmiştir.

Bilgilerinizi ve gereğini rica ederiz.

Prof. Dr. H. Alper BAĞRIYANIK **Başkan** (Katılmadı)	**Dr. Serap ERKEK ÖZHAN** **Başkan Yardımcısı**
Prof. Dr. Sedef AKGÜNGÖR **Üye**	**Prof. Dr. İnci ALACACIOĞLU** **Üye**
Prof. Dr. Gülgün OKTAY **Üye** (Katılmadı)	**Prof. Dr. Hilmi ORHAN** **Üye**
Prof. Dr. Şermin GENÇ **Üye** (Katılmadı)	**Doç. Dr. Canan Aslı YILDIRIM** **Üye** (Katılmadı)
Doç. Dr. Can KÜÇÜK **Üye**	**Dr. Öğr. Üyesi Yavuz OKTAY** **Üye**

8.2. Curriculum Vitae

Vedat Sarı

ID Number:	1400
Date of Birth:	
Contact Address:	
Phone :	+90 50
e-mail:	ve

EDUCATIONAL BACKGROUND

Country	University	Faculty/ Institute	Department	Degree	Year of Graduation
Turkey	Akdeniz University	Agriculture	Agricultural Biotechnology	Bachelor's	2020
Turkey	Dokuz Eylul University	Izmir Biomedicine and Genome Institute	Molecular Biology and Genetics	Master of Science	2022

ACADEMIC / PROFESSIONAL EXPERIENCE

Institution	Country	City	Department	Position	Year
Izmir Biomedicine and Genome Center	Turkey	Izmir	Molecular Biology and Genetics	Internship	August 2019

SPECIALIZED FIELDS

Lacrimal Gland, Induced Pluripotent Stem Cells, Metabolomics, Regenerative Medicine

AWARDS

Name of the Reward	Organization	Year
Poster Award, 1st Place	The International Symposium on Health Informatics and Bioinformatics (HIBIT)	2022

ABSTRACTS PUBLISHED IN CONFERENCE PROCEEDINGS

INTERNATIONAL CONFERENCES
Sarı V., Koçak G., Reçber T., Karakülah G., Utine C.A., Nemutlu E., Güven S., 'Metabolomics analysis of iPSCs derived organoids' The International Symposium on Health Informatics and Bioinformatics (HIBIT), Erdemli, Mersin, Türkiye, 20-21 October 2022