# INVESTIGATING THE MOLECULAR MECHANISMS OF ESSENTIAL OILS ON HUMAN HEALTH

by

Aneeta Uppal

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Bioinformatics and Computational Biology

Charlotte

2020

Approved by:

_____
Dr. Cory Brouwer

_____
Dr. Jeremy Jay

_____
Dr. Wei Sha

_____
Dr. Weijun Luo

_____
Dr. Matthew Parrow

# ABSTRACT

ANEETA UPPAL. Investigating the molecular mechanisms of essential oils on human health. (Under the direction of DR. CORY BROUWER)

The essential oil (EO) industry continues to grow as consumers search for more alternative and complementary therapies. When possible, EO users often turn to EOs for basic medical ailments instead of traditional medications/pharmaceuticals. With the continually high growth of EO consumers, the scientific research to support their many medical claims is inadequate. Due to the large gap in EO research, consumers do not have enough rigorous scientific sources to make informed decisions in regards to EOs. There is a crucial need for more EO related research. My dissertation work will provide a new resources and information for users to educate themselves on EOs from a scientifically driven stand point. It will also provide new data and insights on the application and molecular mechanisms of *Boswellia carterii* (frankincense) EO for targeting inflammation.

# DEDICATION

I would like to dedicate this dissertation to more than one. Each and everyone of you played a large role in how I got myself here.

To God, my parents (Kashmir and Charanjit Uppal), my brothers (Pardeep and Gurveer Uppal), my dogs (Sophee and Neo), and closest friends, I never would have been able to complete this project without the constant encouragement and never-ending support from all of you.

The teachers that always encouraged me from high school and beyond within the field of science: Siobhan Julian, Matthew Brincka, Dawn Carter, Andre Hudson, Robert Osgood and Ravinder Kaur.

To all my kids (and some in memory) at Levine's children hospital that I have had the blessing to teach for the past 4 years, you have provided me with the courage and perseverance to never give up by displaying it through yourselves.

# ACKNOWLEDGEMENTS

I would like to thank my PI, Cory Brouwer and my committee members: Jeremy Jay, Wei Sha, Weijun Luo, and Matthew Parrow for providing the support, encouragement and insights to complete this dissertation research. I would like to extend a big thank you to all of my colleagues within the boinformatics department for providing me the support and resources needed to complete this work: Robert Reid, Steven Blanchard, Kevin Lambirth, Colby Ford, Elizabeth Cooper and Aaron Trautman. Additional staff and faculty at UNC Charlotte: Ian Marriott, Lisa Russell-Pinson, Elizabeth Malone and Shastra Solomon.

Additional appreciation to Alon Mantel, Tergus Pharmaceuticals and Jean-Philippe Therrien for providing me the resources and mentoring to complete the wet-lab research portion of this project. Richard Carlson at Young Living Essential Oils for providing me with the EOs and other tools used through out this project. Jorge Signes, Oksana Gordon and Heather Holycross at DHMRI. Essential oil University database and Life-science publishing for the permission to use their data in the findings of my work.

Finally, thank you to my PhD program advisor, Cynthia Gibas and the GAANN fellowship, the bioinformatics department at UNC Charlotte and the UNC Charlotte Graduate School for providing me with my stipends and tuition funding.

TABLE OF CONTENTS

LIST OF TABLES xii

LIST OF FIGURES xv

LIST OF ABBREVIATIONS xix

CHAPTER 1: INTRODUCTION 1

1.1. History of essential oils 1

1.2. EOs in the industry today 2

1.3. Chemistry and production of EOs 2

1.4. Testing of EOs 5

1.5. Marketing and FDA regulations 6

1.6. Applications of EOs 7

1.7. The debate of the efficacy of EOs in regards to their health claims 8

1.8. Risks associated with EOs 9

1.9. Current research 11

1.9.1. Clinical studies 12

1.9.2. Placebo controlled trials 12

1.9.3. Animal studies 13

1.9.4. Gene expression studies 13

1.9.5. Cell line studies 14

1.10. Challenges and limitations of current research studies 14

1.10.1. Variation of chemical profiles 14

1.10.2. Lack of oil-based controls or well-established standards 15

1.10.3. Blends versus single oils 15

1.10.4. Limitations of cell line studies 16

1.10.5. EOs versus plant extracts 16

1.11.Using bioinformatics based approaches to further investigate EOs 17

1.12.Main aims of this dissertation 18

CHAPTER 2: THE DESIGN AND IMPLEMENTATION OF THE EO KNOWLEDGEBASE (EOKB) AND ITS UTILIZATION FOR ANALYZING EOS 19

2.1. Introduction 19

2.1.1. Deriving a KB driven hypothesis 20

2.2. Methods 21

2.2.1. Data collection 22

2.2.2. Development and design of the EOKB 27

2.2.3. ETL mapping and programming 32

2.2.4. Structure query language (SQL) joins and analysis 32

2.3. Results 34

2.3.1. EOKB query 34

2.3.2. Stringdb results 37

2.4. Discussion 40

2.4.1. Challenges and limitations 41

CHAPTER 3: STUDYING THE EFFECTS OF EOs ON LPS-STIMULATED EX-VIVO HUMAN SKIN EXPLANTS 42

3.1. Introduction 42

3.1.1. Inflammation 42

3.1.2. EOs of interest for testing purposes 44

3.2. Methods 45

3.2.1. EO acquisition 45

3.2.2. Human skin preparation 46

3.2.3. Topical and systemic applications 47

3.2.4. Controls 48

3.2.5. LPS-stimulation 48

3.2.6. RNA isolation for RT-qPCR 49

3.2.7. RNA isolation for RNA-sequencing 49

3.2.8. cDNA library construction and RNA-sequencing 49

3.2.9. RNA-sequencing 50

3.3. Results 52

3.3.1. GC/MS results of EOs and FAME profile of FCO 52

3.3.2. Statistical analysis of RT-qPCR data 57

3.3.3. DESeq2 results 66

3.4. Discussion 67

3.4.1. Connecting results back to EOKB derived hypotheses 68

3.4.2. Limitations 69

3.4.3. Future studies 70

CHAPTER 4: PREDICTING THE MOLECULAR MECHANISM OF FRANKINCENSE EO 72

4.1. Introduction 72

4.1.1. Pathway analysis 73

4.1.2. Frankincense gene expression data in comparison with NSAIDs and steroids 74

4.1.3. Predicting the mechanism of frankincense EO 74

4.2. Methods 74

4.2.1. Ingenuity pathways analysis 74

4.2.2. Generally applicable gene set enrichment (GAGE) with Pathview 75

4.2.3. Gene set enrichment analysis (GSEA) 75

4.2.4. EO chemical analysis 75

4.2.5. Alternative splice (AS) site detection 76

4.2.6. Weighted gene co-expression network analysis (WGCNA) 76

4.2.7. Analysis of anti-inflammatory pharmaceuticals in comparison to frankincense EO 77

4.2.8. Analysis of anti-inflammatory plant extracts in comparison to frankincense EO 78

4.3. Results 78

4.3.1. Ingenuity pathway analysis 78

4.3.2. Analysis of frankincense systemic results 79

4.3.3. Analysis of frankincense topical results 91

4.3.4. Analysis of FCO topical results 94

4.3.5. Analysis of the untreated results 97

4.3.6. WGCNA 100

4.3.7. Comparison of frankincense to traditional NSAIDs 103

4.3.8. Comparison of frankincense to five non-hematopoietic treated with a methylprednisolone 105

4.3.9. Comparison to frankincense with colchicine 107

4.3.10. Principal component analysis (PCA) and hierarchical clustering of frankincense EO species using chemical profiles 107

4.3.11. Pathway analysis results in comparison to previously generated RT-qPCR results 115

4.3.12. Pathway analysis results for frankincense systemic in comparison to EOKB Predictions 115

4.4. Discussion 117

4.4.1. Pathway analysis 117

4.4.2. PCA and hierarchical clustering 118

4.4.3. Comparisons with pharmaceuticals and OTCs 119

4.4.4. Limitations 119

4.4.5. Conclusion 120

CHAPTER 5: CONTINUATION OF FRANKINCENSE EO ON LPS-INDUCED *EX-VIVO* SKIN EXPLANTS 121

5.1. Introduction 121

5.2. Methods 123

5.2.1. Human skin acquisition and testing 123

5.2.2. EO acquisition 123

5.2.3. Experimental design 124

5.2.4. Topical and systemic applications 124

5.2.5. RNA-isolation and RT-qPCR 125

5.2.6. Statistical analysis 125

5.2.7. Results 125

5.3. Discussion 131

CHAPTER 6: CONCLUSION 133

6.1. Improvements 133

6.2. Future studies 133

6.2.1. Enhancers/vehicles to aid in dermal absorption 133

6.2.2. Synergistic effects of EOs 134

6.2.3. How EOs can effect the microbiome 134

6.2.4. Accessibility of the EOKB 135

6.2.5. Further use of the EOKB 135

6.3. Applications of EO research in current health affairs 136

6.4. Significance in the real world 138

6.5. Social Media Reach 138

6.6. Timeline and reflection 140

REFERENCES 142

APPENDIX A: GITHUB 159

# LIST OF TABLES

TABLE 1.1: Cosmetics and personal care products lead the list of the most common substances implicated in pediatric exposures[1] 10

TABLE 1.2: Pain medications lead the list of the most common substances implicated in adult poison exposures (20 years or older) (NCPC, 2018)[1] 10

TABLE 2.1: Kegg pathway Stringdb results for frankincense 38

TABLE 2.2: Reactome pathway Stringdb results for frankincense 38

TABLE 2.3: Kegg pathway Stringdb results for myrrh 38

TABLE 2.4: Reactome pathway Stringdb results for myrrh 39

TABLE 2.5: Kegg pathway Stringdb results for blue spruce 39

TABLE 2.6: Reactome pathway Stringdb results for blue spruce 39

TABLE 3.1: LPS stimulation study design per donor 48

TABLE 3.2: Samples used for RNA-sequencing from D1 50

TABLE 3.3: GC/MS results of frankincense EO 53

TABLE 3.4: Physical test results of frankincense EO 54

TABLE 3.5: GC/MS results of myrrh EO 54

TABLE 3.6: Physical test results of myrrh EO 54

TABLE 3.7: GC/MS results of blue spruce EO 55

TABLE 3.8: Physical test results of blue spruce EO 55

TABLE 3.9: Fatty acid composition of FCO 56

TABLE 3.10: Composition of FCO by fat type 56

TABLE 3.11: RT-qPCR measurements 57

TABLE 3.12: Results from ANOVA using log transformed fold change data. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001 58

TABLE 3.13: Dunnett's method for D1 (COX2 and IL1$\beta$ based on log transformed fold change data). Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001 59

TABLE 3.14: Dunnett's method for D1 (CXCL10 and IL8 based on log transformed fold change data). Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001 60

TABLE 3.15: Dunnett's Method for D2 (COX2 and IL1$\beta$ based on log transformed fold change data). Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001 60

TABLE 3.16: Dunnett's Method for D2 (CXCL10 and IL8 based on log transformed fold change data). Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001 61

TABLE 4.1: Number of DEGs per sample using DESeq2. Cut off set at p-adjusted value <0.05 and Log2FC -1.5/+1.5 79

TABLE 4.2: Top Canonical Pathways of frankincense systemic 79

TABLE 4.3: Diseases and disorders of frankincense systemic 80

TABLE 4.4: Key genes associated with inflammatory-mediated response by frankincense systemic 80

TABLE 4.5: Subset of predicted activation states for the listed diseases above based on p-values, number of gene associations and z-scores. 82

TABLE 4.6: Subset of upstream analysis regulators and their predicted activation states based on literature findings, pvalues, gene associations and z-scores. 83

TABLE 4.7: Pathview summary results for frankincense systemic 85

TABLE 4.8: Top canonical pathways of FCO topical 94

TABLE 4.9: Diseases and disorders of FCO topical 94

TABLE 4.10: Molecular and cellular functions of FCO topical 95

TABLE 4.11: Top canonical pathways of the untreated applications 97

TABLE 4.12: Diseases and disorders of untreated application 97

TABLE 4.13: Key genes with Log2FC from inflammatory response figure produced by untreated 98

TABLE 4.14: Kegg pathway Stringdb results for frankincense EO (systemic) based on the DEGs calculated from the RNA-seq data 116

TABLE 4.15: Reactome pathway Stringdb results for frankincense EO (systemic) based on the DEGs calculated from the RNA-seq data 116

TABLE 5.1: GC/MS results of the new frankincense EO 124

TABLE 5.2: LPS and EOs continuation study design per donor 3 124

TABLE 5.3: Results from ANOVA (based on log transformed fold change data). Markers where a significant change was observed according to their F-values were indicated by: * <0.05, ** <0.01, *** <0.001 125

TABLE 5.4: Dunnett's method for D3 (COX2 and IL1$\beta$ based on log transformed fold change data). Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001 126

TABLE 5.5: Dunnett's Method for Donor 3 (CXCL10 and IL8 based on log transformed fold change data). Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001 126

TABLE 6.1: Contains running totals of social media page statistics to date (11-11-20) 140

TABLE 6.2: Demographics and age range of social media followers 140

TABLE 6.3: Timeline of dissertation research 141

# LIST OF FIGURES

FIGURE 1.1: Steam distillation process of producing EOs[2] 4

FIGURE 1.2: Number of EO-based scientific publications increases over the last 20 years 11

FIGURE 2.1: Multi-query used for Gene_Chemical_Association table and other related EO tables illustrated within the ER diagram 24

FIGURE 2.2: Entity-relationship diagram for EOKB part 1 28

FIGURE 2.3: Entity-relationship diagram for EOKB part 2 29

FIGURE 2.4: Legend for Crow's foot notation 31

FIGURE 2.5: Utilizing the EOKB to connect chemical-gene associations with gene and pathway information 33

FIGURE 2.6: Utilizing the EOKB to connect chemical-gene associations with other EOs that share similar chemical components with frankincense EO 34

FIGURE 2.7: EOKB driven hypothesis for frankincense 35

FIGURE 2.8: EOKB driven hypothesis for myrrh 36

FIGURE 2.9: EOKB driven hypothesis for blue spruce 37

FIGURE 3.1: Pictured on the left: Franz cell diagram from PermeGear, Ltd. Pictured on the Right: Franz cell used in the lab with EO mixed into media for systemic experimental application. 47

FIGURE 3.2: RNA-sequencing pipeline 52

FIGURE 3.3: RT-qPCR results from COX2 and CXCL10 (D2). Each graph displays the Log2FC values for each application according to the labeled biomarker on the top left corner of each plot. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * $<0.05$, ** $<0.01$, *** $<0.001$ 62

FIGURE 3.4: RT-qPCR results from IL1$\beta$ and IL8 (D1). Each graph displays the Log2FC values for each application according to the labeled biomarker on the top left corner of each plot. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001 63

FIGURE 3.5: RT-qPCR results from COX2 and CXCL10 (D1). Each graph displays the Log2FC values for each application according to the labeled biomarker on the top left corner of each plot. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001 64

FIGURE 3.6: RT-qPCR results from IL1$\beta$ and IL8 (D2). Each graph displays the Log2FC values for each application according to the labeled biomarker on the top left corner of each plot. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001 65

FIGURE 3.7: MA plot results of DEGs from DESeq2. The x-axis displays the log2 mean expression of the data while the y-axis displays the Log2FC. DEGs were filtered before plotting for a Log2FC value of -1.5/+1.5 with a p-adjusted value <0.05. DEGs up regulated displayed in red and DEGs down regulated displayed in blue. Genes that were not statistically significant displayed in grey. 67

FIGURE 4.1: DEGs from frankincense systemic associated with inflammatory-mediated response 84

FIGURE 4.2: Partial image generated by Kegg and Pathview for the frankincense systemic application on the cytokine-cytokine receptor interaction pathway(hsa04060) Log2FC scale displayed in bottom left area of the figure 86

FIGURE 4.3: The graph on the left displays the distribution of the enrichment score (ES) of the LPS controls versus the frankincense systemic. The figure on the right shows the correlation between the ranked gene sets based on the sample conditions. 88

FIGURE 4.4: Heat map of top 50 features for each condition showing the correlation between the ranked genes and the treatment type, using two different gene signature databases (MSigDB/DSigDB). Gene expression is represented by a range of colors (red, pink, light blue, dark blue) shows the range of expression values (high, moderate, low, lowest) 90

FIGURE 4.5: On the left: Heat map of top 50 features for each condition showing the correlation between the ranked genes and the treatment type between LPS controls and frankincense topical (MSigDB: C7). Gene expression is represented by a range of colors (red, pink, light blue, dark blue) shows the range of expression values (high, moderate, low, lowest) On the right: ES scores based on the number of gene lists, and correlation profile of the gene lists in comparison with treatment applications 93

FIGURE 4.6: On the left: Heat map of top 50 features for each condition showing the correlation between the ranked genes and the treatment type between LPS controls and FCO topical (MSigDB: H). Gene expression is represented by a range of colors (red, pink, light blue, dark blue) shows the range of expression values (high, moderate, low, lowest) On the right: ES scores based on the number of gene lists, and correlation profile of the gene lists in comparison with treatment applications 96

FIGURE 4.7: On the left: Heat map of top 50 features for each condition showing the correlation between the ranked genes and the treatment type between LPS controls and untreated (MSigDB: H). Gene expression is represented by a range of colors (red, pink, light blue, dark blue) shows the range of expression values (high, moderate, low, lowest) On the right: ES scores based on the number of gene lists, and correlation profile of the gene lists in comparison with treatment applications 99

FIGURE 4.8: Sample clustering by euclidean distances with WGCNA 100

FIGURE 4.9: The dendrogram displays the clustering of genes with dissimilarity based on topological overlap, together with assigned module colors 101

FIGURE 4.10: Graph displays all genes in common with frankincense systemic from the NSAID treatments with a Log2FC $\leq$ -1.5 104

FIGURE 4.11: This graph displays expression data for genes that had a Log2FC $\leq$ -1.5 between the 5 non-hematopoietic cell lines 106

FIGURE 4.12: This graph displays a PCA plot by looking at the weight of the metabolites of the varying kinds of frankincense EO 109

FIGURE 4.13: This graph displays a PCA plot by adding in the variables (chemical components) 110

FIGURE 4.14: This graph displays a PCA plot of different EOs within the Sapindales order, which contained two families Burseraceae and Rutaceae. 112

FIGURE 4.15: This graph displays two hierarchical clustering graphs. Both are displaying the same data presented with different labels. 114

FIGURE 5.1: RT-qPCR results from COX2 and CXCL10 (D3). Each graph displays the Log2FC values for each application according to the labeled biomarker on the top left corner of each plot. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * $<0.05$, ** $<0.01$, *** $<0.001$ 128

FIGURE 5.2: RT-qPCR results from IL1$\beta$ and IL8 (D3). Each graph displays the Log2FC values for each application according to the labeled biomarker on the top left corner of each plot. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * $<0.05$, ** $<0.01$, *** $<0.001$ 130

FIGURE 6.1: Genes in common between Coronavirus strains with frankincense EO, directly involved in cytokine-cytokine receptor pathway (hsa:04060) In the venn-diagram SARS-CoV-1(Cov1), SARS-CoV-2(cov2), and MERS-CoV(Mers) 137

FIGURE 6.2: Oilyscientist Instagram page providing EO education from scientifically driven resources and peer-reviewed studies 139

# LIST OF ABBREVIATIONS

ANOVA Analysis of variance

Blue spruce *Picea pungens*

D1 Donor 1

D2 Donor 2

D3 Donor 3

DB Database

DE Differentially expressed

DEGs Differentially expressed genes

EOKB Essential Oil Knowledgebase

EOs Essential oils

EOUdb Essential Oil University Database

ER Entity relationship

ES Enrichment score

ETL Extract, transform and load

FA Fatty acids

FCO Fractionated coconut oil

FDA Food and Drug Administration

FDR False discovery rate

Frankincense *Boswellia carterii*

GC Gas chromatography

GEO NCBI Gene Expression Omnibus

GSEA Gene Set Enrichment Analysis

IPA Ingenuity Pathway Analysis

KB Knowledgebase

Log2FC Log2 ratio fold change

LPS Lipopolysaccharide

MS Mass spectrometry

MSigDB Molecular Signature Database

Myrrh *Commiphora myrrha*

NCBI National Centre for Biotechnology Information

NLP Natural language processing

NSAIDs Non-steroidal anti-inflammatory drugs

OTC Over-the-counter

p-adj P-adjusted Value

p-val P-value

PCA Principal component Analysis

PK Primary key

RDMS Relational database management system

RNA-seq RNA-sequencing

RT-qPCR Quantitative reverse transcription polymerase chain reaction

WGCNA Weighted Gene Co-expression Network Analysis

# CHAPTER 1: INTRODUCTION

## 1.1 History of essential oils

The use of essential oils (EOs) and other plant extracts can be traced back to thousands of years ago.[3] An EO is defined as a concentrated liquid of a complex mixture of volatile compounds that are extracted from different parts of a plant.[4] EOs and plant extracts are mentioned in ancient texts such as the bible and Egyptian hieroglyphics.[4] Mummification processes included the use of aromatic plant extracts during the embalming of wealthy people and were poured all over the body.[5] They were also used through out the middle ages and the renaissance, most commonly in religious practices.[6] The first researchers to study plant compounds originated from France in the 1800's.[7] These chemists played important key roles in the isolation and identification of different plant-based molecules that began to unravel our understanding of plant chemistry (M. J. Dumas, M. Berthelot, O. Wallach, Hesse, Gildemeister, Betram, Walbaum and Wienhaus).[7] As time went on, more researchers began to identify different plant molecules and improve the methods and our understandings of chemical extractions, isolation and identification of organic molecules. In 1937, Gattefosse, a french chemist first coined the term "aromatherapy," and its use for medicinal applications.[4] However, many of these medical claims were based on anecdotal evidence and not clinical research studies. During this time the use of EOs and their aromatic compounds were increasing in the perfume industry and for cosmetic applications.[7] As the use of EOs began to increase, as well as their demand, it opened the doors for a market of adulterated and synthetic EOs in the early to mid 1900s.[7] EOs containing contaminants, synthetics and diluted with fatty acids began to rise, and is still an issue today in the EO industry.[7]

## 1.2 EOs in the industry today

The industry for EOs continues to grow at a fast rate. Market researchers have predicted that the EO market will continue to increase with a growth rate of 9% from 2016-2024.[8] In 2015, the market size for EOs was estimated to be valued at about USD 3.36 billion.[8] Today, there are hundreds of varying EOs available for purchase online and in retail businesses with many industry companies now producing them. Industries for food/beverage, spas, cleaning and home products are utilizing EOs and their chemical compounds within their products. Many companies are now introducing EO infused products to the marketplace. Diffusers and EOs are appearing all over stores and online markets worldwide. The Food and Drug Administration (FDA) does not regulate the EO industry.[9] This leaves companies with the freedom to create their own standards on how their EOs are produced and labeled. This includes designing their own protocols for cultivating the crops, distilling the oils, bottling and testing before the product enters the marketplace. Since EOs fall under the category of "aromatherapy" they cannot be used to treat or prevent any illness or disease, and cannot be advertised as such.[9] Many of these oils are being used for basic medical ailments and as complementary therapies. The lack of standards and regulations for EOs imposes risks for consumers when the quality and testing of EOs can vary between different brands. The variance in the chemical profiles of EOs can be caused by the geographical location of the cultivation, species/variety of the plant, time of harvest and distillation practices. It is possible these varying properties could pose a threat to consumers if they are misused or present misleading labels.

## 1.3 Chemistry and production of EOs

An EO is comprised of a complex mixture of volatile organic compounds that have been produced from different parts of the plant.[10] The current industry standards of essential oils are set by the International Standards Organization (ISO) and

the French National Organization for standardization and its International Organization for Standardization member body (AFNOR).[7][3] They state EOs should be produced by some type of water/steam distillation method, or mechanical process (also known as dry distillation or cold-pressed methods) that do not cause significant changes in the plant composition.[3][7] Although these standards are recognized in the EO industry as global standards, no one enforces them across the industry. The process of EO production utilizes three main methods: 1) Hydro or steam distillation method (most common), 2) cold-pressed extraction, and 3) dry-distillation (rarely used).[10] Solvent extractions and supercritical carbon dioxide extractions are widely used for producing oils, but it is argued that these methods of production produce EO absolutes, which are not technically considered EOs, as termed by ISO standards.[7] EOs produced in the methods mentioned above should not contain fatty acids, explained into detail in section 1.10.2.

The traditional method of producing EOs is hydrodistillation, and is still employed for many EOs. Steam distillation has become the general standard method for commercially produced EOs.[11] Hydrodistillation immerses the entire plant into water and brings it to boiling temperatures. Next the condenser collects the vapors of the steam and oil molecules. The oil and water can now be easily separated.[12] Instead of submerging the plant matter into boiling water, steam distillation injects steam into the chamber containing plant material. There, the oil vapors are released within the steam. The steam and oil vapors move into the condenser where the oil and aqueous layer can be separated.[2] This process of steam distillation is outlined in figure 1.1 below.

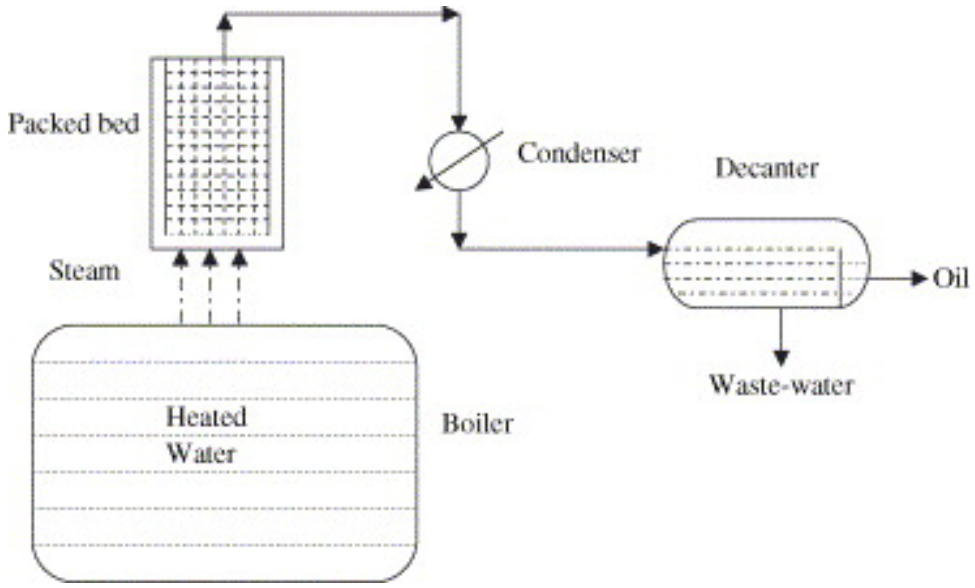


Figure 1.1: Steam distillation process of producing EOs[2]

Certain EOs require cold-pressed methods to release the oils from the rinds of the fruit.[3] Using immense pressure, the rinds of fruits are pressed and the oils are released. The oil is then filtered and bottled.[3][7] EOs commonly produced in this manner include citrus oils such as lemon and orange. As EO production has grown in industry, new methods of plant extraction have been introduced. Mentioned previously as the third method. These methods include: supercritical fluid extraction, subcritical fluid extraction, and solvent free microwave extraction method.[11] Employing different methods of EO production generates different yields and varying chemical profiles. Technically, the oil distiller has the freedom to choose the appropriate method for the production of their EOs.

EOs are comprised of many different chemical compounds. They contain mixtures of terpenic and sesquiterpenic hydrocarbons, alcohols, aldehydes, ketones, alkanes, esters, carboxylic acids, phenols, ethers, and more.[7] Essential Oil University Database (EOUdb) is a publicly available database that has recorded over 4,221 compounds in 4,152 different EOs.[13] The underlying molecular mechanisms inside the body remain unknown for the majority of these chemicals due to the shortage of general EO

research.

EO composition can vary based on the cultivar of the plant in which they are produced from. Two different plants can phenotypically look similar, reside within the same family, yet vary greatly in their chemical profiles.[3] For example, Rosemary EO a common herb can have many different cultivar/chemotypes including: alpha-pinene, 1,8-cineole, camphor, verbenone, camphene, and myrcene.[14][10] These varying chemotypes can cause large variations in the applications of the oil. Not only can EOs vary in chemical profiles based on cultivars, but they can also differ greatly depending on: geographic location, climate (stress/drought), cultivation methods (fertilization, pesticide use, irrigation methods), time and conditions of cultivation/harvesting, primary processing of the harvested plant, methods of distillation and testing, as well as bottling, packaging and storing.[3][10][7]

## 1.4 Testing of EOs

Each company will differ in the production and testing standards of their EOs. Physical characteristics of oils can be determined by measuring the refractive index, specific gravity, melting point, flash point, and viscosity of the EO.[15] EO molecules are small in size and can be hard to identify.[3] Analytical chemical testing methods can be implemented to determine the chemical composition of EOs. Some of the more popular used tests include: Gas chromatography (GC), mass spectrometry (MS), nuclear magnetic resonance (NMR), high performance liquid chromatography (HPLC), HPLC combined with MS/NMR methods, inductively coupled plasma mass spectrometry (ICP-MS), GC-fourier transform infrared spectroscopy (GC-FTIR) and GC-isotope-ratio mass spectrometry (GC-IRMS) .[12][3][10][7] These testing procedures can tell oil makers the chemical composition of their oil and detect the presence of contaminants within the product. Furthermore, analytical chemists are the only trained scientists to understand the results from these testing methods. Each individual oil is said to have a specified range of relative area percentage for each chemical

component that is true to that native species of the plant.[12] There are also concentration standards/limits for known toxic substances set by the FDA and the World Health Organization (WHO), that are allowed in products labeled for consumption.[7] The average consumer will not have the means to test their own samples that they have purchased. Due to the limitations on chemical composition testing paired with the lack of regulations and industry standards, customers are forced to place their trust in the company they choose to buy EOs from

## 1.5 Marketing and FDA regulations

In marketing aspects, companies cannot suggest the use of EOs to diagnose, treat, cure or prevent any illnesses or disease.[9] Products that make such claims have to be approved by the FDA and undergo clinical trials. Most companies will not market their EOs for such uses due to liability. However, there are thousands of websites, blogs, phone applications, and reference guides that contain information on how to use specific EOs to treat, cure and prevent specific diseases/illnesses and related symptoms. Most reference guides have hundreds of user suggestions for EOs relieving symptoms as minor as headaches and rashes to treating symptoms and diseases as severe as cancer, Alzheimer's disease, hepatitis, and many more. These suggestions are not always given by accredited researchers and medical professionals, but there is no way to effectively prevent and regulate the resources being circulated to the public by those not trained in medicine or scientific research. There are a number of EO books written by authors with no scientific or medical background suggesting how users should utilize EOs in human health and disease applications without citing reputable scientific data to support their claims. In a world where many people already have preexisting medical conditions and take medications, this has the potential to be harmful.

## 1.6 Applications of EOs

Most EO users are utilizing EOs in 3 main areas of application: Topical, aromatic, and ingestion. Topical applications of EOs has users apply the EO diluted with a fatty-acid carrier oil to areas of attention/need. For example: those experiencing sore joints due to inflammation may apply oils topically to their knee area in a 1:1 dilution with a fatty acid carrier oil. Some users apply oils neat, but many advise against this as it can cause contact dermatitis and irritation due to the properties of EO constituents.[3] Others will apply oils topically simply for the aromatic properties they provide, similar to perfumes or body care products. Many are utilized in skin issues such as acne/scarring or blemishes.

Aromatic applications involve the inhalation of EOs. Some users may apply the oil in the palm of their hands and inhale it, or inhale the oil from the bottle. Others will diffuse the oil. Although there are different styles of diffusers, the most commonly used diffusers are ultrasonic water-based diffusers. The user adds EOs directly into the water chamber of the diffuser. The ultrasonic vibrations from the diffuser cause the EO molecules to disperse with the water vapors into the environment.[16][17] There are many different types of diffusers sold by industry companies. The most common uses for the diffusion of EOs are for pleasant aromas, relieving negative emotions or feelings, as sleep aids, air purification and for relieving symptoms associated with congestion.

The ingestion of EOs is an area of dispute between many different researchers, medical professionals, and aromatherapists. Many advise against the ingestion of EOs, however there are many users that do ingest EOs. The FDA has listed a number of EOs and natural extractives as substances "generally recognized as safe" (GRAS) in their Code of Federal Regulations.[18] Currently, it is at the discretion of the user on whether they choose to ingest EOs. Not enough clinical research has been conducted in the area of EO ingestion to know specifics about how EOs are digested

and metabolized. There are no current studies in humans studying the effects of EO consumption at different dosages, the rate of excretion/absorption of the EOs, or the mechanisms of how these oils are fully metabolized within the body at the molecular level. As the field of pharmacogenomics grows, there is a deeper need to understand how genetic variations in individuals can play a key role in understanding how these molecules are metabolized and processed within the body. Research continues to grow in areas of the microbiome. With many EOs containing antimicrobial properties, it is also unknown whether the ingestion of EOs can cause alterations of the gut microbiota.[19][20] Some studies are just beginning to investigate the role of herbal medicine on the gut microbiome, but there are very few, if any studying essential oils and their effects on the microbiome.[21]

### 1.7 The debate of the efficacy of EOs in regards to their health claims

Are EO claims true? Can these oils heal various health conditions and diseases? It is very possible they are doing something from the thousands of personal claims and anecdotal evidence consumers are making regarding their own personal user experiences. Many claim they have experienced little to no side effects when using EOs over standard medications. Many researchers believe these claims raise concerns of EOs causing a placebo effect. However, many chemicals extracted from plants have already been re-purposed into products, pharmaceutical applications and many over-the-counter (OTC) products that are widely used today. An example of this would be methyl salicylate, an ester found in wintergreen plants and a major component of its EO. Methyl salicylate is commonly used in topical applications for pain relief associated with acute tissue injuries, sciatica and rheumatic conditions.[22] It is also utilized in some countries as a local analgesic for human and veterinary medicine.[23] One could also argue, consumers would not continue to purchase and use a product if they did not receive favorable results from it. With over 4,000 recorded chemicals in EOs, the list for their possible individual uses has yet to be determined.[13] In the

last decade, there has been a large growth in consumer demands for all natural alternatives in place of traditional pharmaceuticals or synthetic products. Just because an alternative method or complementary therapy is natural, does not mean it is still safe. This has many scientists and researchers currently trying to bridge the gap for the applications of EOs backed up by substantial scientific research.

## 1.8 Risks associated with EOs

Information regarding the effects of EOs in humans at the metabolic level remains widely unknown. The risks associated with the use of EOs seems somewhat unknown. According to the US National Capital Poison Center (NCPC), there is 1 poison exposure roughly every 15 seconds in the USA. As of 2018, the most common substances that are the cause of poison exposures in children are cosmetic and personal care products followed by cleaning substances, analgesics, and foreign bodies/miscellaneous items. Table 1.1 lists pediatric poison exposures according to substance for children less than 6 years old. This age group makes up the majority of poison exposures reported to the NCPC. EOs fall under the dietary supplements/herbals/homeopathic category and make up on a small fraction of the total exposures recorded in children. Most exposures in children are unintentional. Pain medications are reported as the single most frequent cause of pediatric fatalities.[1] Supplements/herbals/homeopathic substances are not on the list for the most frequent cause of poisoning fatalities reported.

The most common substances for poison exposure in adults are recorded in table 1.2. Analgesics and controlled substances are highest on the list. It is important to note most exposures in adults are often intentional. Dietary supplements, herbals, homeopathic or vitamins do not make the list for adult exposures. In 2017, 84% of exposures to the NCPC are typically non-toxic or minimally toxic. According to these statistics, one could argue the risks for EOs and other herbal remedies are quite low, and most likely minimally toxic or non-toxic. This is not meant to take away from

Table 1.1: Cosmetics and personal care products lead the list of the most common substances implicated in pediatric exposures[1]

| **Substance** | **No.** | **%** |
|---|---|---|
| Cosmetics/personal care products | 117,298 | 12.1 |
| Cleaning substances | 103,387 | 10.7 |
| Analgesics | 87,526 | 9.0 |
| Foreign bodies/toys/misc. | 66,519 | 6.9 |
| Topical preparations | 45,397 | 4.7 |
| Antihistamines | 44,734 | 4.6 |
| Vitamins | 41,581 | 4.3 |
| Dietary supplements/herbals/homeopathic | 39,984 | 4.1 |
| Pesticides | 35,015 | 3.6 |
| Gastrointestinal preparations | 25,293 | 2.6 |

previously emphasized statements for the need to further understand these oils. But rather question whether they could be safer products for accidental exposure than traditional standard household items such as cosmetics and cleaners. EOs may also be less prone to substance abuse as seen with alcohol and prescription drugs which make up the majority of the reports. It is important that we learn how EOs could play a role in drug metabolism and interactions.

Table 1.2: Pain medications lead the list of the most common substances implicated in adult poison exposures (20 years or older) (NCPC, 2018)[1]

| **Substance** | **No.** | **%** |
|---|---|---|
| Analgesics | 128,419 | 10.9 |
| Sedatives/hypnotics/antipsychotics | 109,739 | 9.3 |
| Antidepressants | 86,018 | 7.3 |
| Cardiovascular drugs | 76,647 | 6.5 |
| Cleaning products | 64,340 | 5.5 |
| Alcohols | 56,099 | 4.8 |
| Anticonvulsants | 49,730 | 4.2 |
| Stimulants and street drugs | 41,940 | 3.6 |
| Pesticides | 40,760 | 3.5 |
| Antihistamines | 39,921 | 3.4 |

## 1.9 Current research

Upon searching the National Centre for Biotechnology Information (NCBI) there are over 20,000 published articles in regards to EOs, volatile oils or aromatherapy studies. Of these articles, around 1,500 are recorded as a review article. Research in this area continues to grow every month.[24] This number gives the impression that there is an ample amount of scientific literature available on EOs. However, it is actually somewhat lacking when comparing it to the number of published studies around different diseases and other topics in science. When searching for studies based on a disease such as rheumatoid arthritis for example, there are over 133,000 publications. Upon searching for a specific substance, for example, steroids, returns results of over 200,000 publications. When taking into account the high number of different EOs available with their variations in chemical constituents, the amount of scientifically driven data is inadequate. With the growing number of users utilizing EOs, the crucial need for the research increases. Figure 1.2 shows the increasing number of scientific publications involving EOs. NCBI PubMed reported that 261 EO-related articles were published in 2000. Currently, 2019 has the highest recorded number of EO related articles of 1840.

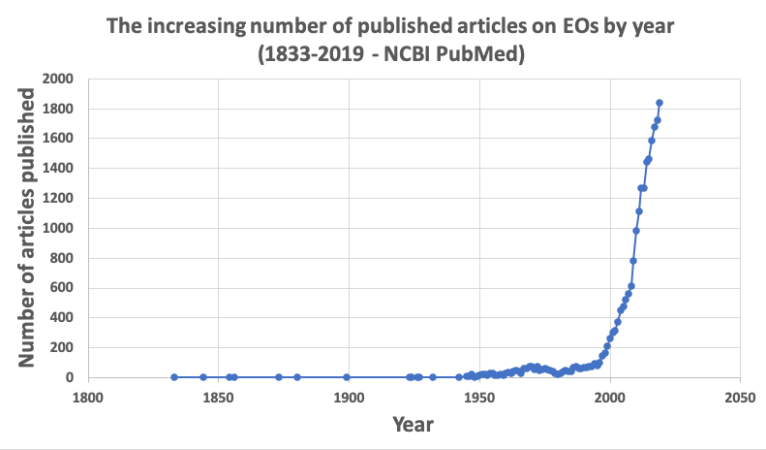


Figure 1.2: Number of EO-based scientific publications increases over the last 20 years

The majority of these scientific studies are done on some of the most common plants and EOs, such as lavender (*lavandula angustifolia*), basil (*Ocimum basilicum*), peppermint (*Mentha piperita*), etc. Search results for plants and oils such as cypress (*Callitris intratopica*), palo santo (*Bursera graveolens*) and elemi (*Canarium luzonicum*), returns many fewer clinical studies investigating their therapeutic properties. Research studies can vary in complexity from basic assays on the biological and chemical properties of oils to cancer studies in mice and even a small subset of clinical studies in humans.

### 1.9.1 Clinical studies

There are some clinical studies on EOs that are focused more towards the aromatherapy application and some based on topical applications. Worldwide there is roughly 248 EO clinical trial studies listed in the clinicaltrials.gov database.[25] Of those total trials, around 81 of them have or are being conducted in the United States. Unfortunately, the majority of these studies do not post the completed results of the study, even though they have been marked as completed. Many are undergoing or still recruiting. NCBI PubMed returns approximately 693 results related to clinical human studies conducted on EOs. Many of those articles conclude the possible use, potential and encouragement of further investigation for EOs as therapeutic treatments.

### 1.9.2 Placebo controlled trials

Most clinical studies for EOs administer a placebo to a number of participants. Placebo controlled trials are important, but finding a realistic placebo for EOs or EO based products is extremely difficult. Taking into consideration the strong aromatic properties of EOs, it is difficult to find a placebo of similar nature. If a study uses coconut oil as a placebo versus the EO treatment, some variation could be seen in the results due to the scent associated with the EO product versus the coconut oil. A

patient receiving a treatment that has different odors associated with it, may think they are receiving the testing treatment versus the placebo. This can be especially challenging in placebo controlled aromatherapy trials. Finding a suitable placebo can be difficult.

### 1.9.3 Animal studies

EO related studies carried out on animal models (excluding humans) account for about 25% ( 5,600) of the results indexed in PubMed. The majority of EO studies have used *ex-vivo/in-vitro* methods for testing. Animal models are mainly used to study the therapeutic/supplemental use of EOs. There is an extensive number of studies analyzing their efficacy as insecticides and anti-microbial agents. The more commonly used animal models are mice *(Mus musculus)* and rats *(Rattus norvegicus)*.[26] Studies on livestock is limited to several hundred articles, but continues to slowly increase.[27] EO research in relation to dog *(Canis familiaris)* and cat *(Felis catus)* models are sparse, but gradually increasing as researchers investigate the possible supplementation of oils for digestive system support, the inhibition of bacterial growth, and lipid oxidation.[28][29] Other studies have analyzed the insecticidal properties and cytotoxic effects of EOs, specifically on mosquito (culicidae) populations.[30] A literature review concluded over 216 different EOs were deemed effective for their larvicidal capabilities.[31]

### 1.9.4 Gene expression studies

As the field of bioinformatics continues to grow there is a continued shift in experiments that are generating genomics related data. Gene expression measurement studies generally use methods of microarrays, sequencing, quantitative reverse transcription PCR (RT-qPCR), northern blots or other similar methods. The RNA is isolated from treatment models and usually requires a step that provides some type of target probe hybridization for microarray studies.[32] RNA-sequencing (RNA-seq)

has been increasing in its use due to its high sensitivity, and varying capabilities to not only look at gene expression but also single nucleotide polymorphisms (SNPs), insertions, deletions and alternative splice sites.[33] Currently, there is a low number of gene expression analysis studies on EOs involving either microarrays or RNA-seq methods. Searching NCBI's Gene Expression Omnibus (GEO) returned only a total of 19 data sets for EOs, with only 5 of them from humans.[34] One of the more recently published data sets is the first gene expression study that looked at the effects of sandalwood EO on skin tissue at multiple time points (9 hours and 24).[35] Studies testing multiple time points are more costly but provide insight on the efficacy of the treatment over time.

#### 1.9.5 Cell line studies

After reviewing experimental studies on cell line assays, the majority of studies are carried out on 2D cell lines. Very few studies tested EOs on 3D cell/tissue models. The majority of 2D cell line studies explore the potential therapeutic applications of EOs. Some of these studies used EOs to observe the possible cytotoxic effects of the oils on carcinoma cell lines.[36][37] The majority of EO based cell culture assays test oils on different cancerous/diseased cell lines. Unfortunately, there are not many studies testing the general effects of EOs on healthy cell lines.[38][39][40]

### 1.10 Challenges and limitations of current research studies

#### 1.10.1 Variation of chemical profiles

EOs can vary greatly in their chemical compositions. Depending on the plant, cultivar, geographical location and methods of cultivation/distillation. For these reasons it is important that every EO study provides supplemental information on the chemical composition of the EO that was used. Some studies will list only the main constituents without the full chemical composition. The main components of most EOs include their aromatic compounds. Aromatic compounds may not necessarily be the driving

factor for the favorable effects observed from the EO. Trace amounts of bio-active molecules within an EO, could cause significant differences in the treatment groups. It is crucial that the entirety of all the molecules that were detected/characterized in an EO are listed in the publications (main article or supplementary). This will aid researchers in pinpointing certain effects to specific chemical molecules.

### 1.10.2 Lack of oil-based controls or well-established standards

As EO studies continue to develop, standards and controls need to be established. This is difficult, as it is hard to determine ideal positive and negative controls. Many studies are beginning to use a fatty acid based oils as controls due to their hydrophobic nature. EOs do not contain fatty acids (FA). FA molecules are lost during the distillation process due to their relatively large molecular size.[41] The FA oils serve as a control by providing another oil-based substance. Using FA oils is helpful in any experiments involving cell/tissue culture assays that involve media. If too much EO is added to the media, it can create a film along the surface of the solution, reducing the amount of oxygen available to the cells. Some EOs can even form small precipitates when added to solutions.[42] It is unclear whether FA oils would be the "best ideal" experimental control, considering they differ in composition to EOs. Coconut oil, grapeseed oil, olive oil, vegetable oil and almond oil are all commonly used to dilute EOs for topical applications.

### 1.10.3 Blends versus single oils

Among the studies published, many investigated the effects of a specific blend of EOs. Although, this information may be useful, it makes it even more difficult to pinpoint which oils and compounds were causing the results concluded by the study. Researchers may use more than one oil in hopes of receiving better results by using a blend. This does not aid researchers in understanding the chemical properties of each individual oil. Adding multiple oils only increases the variability within the

study. Researchers should focus on single oils until each one is well investigated and understood. This continues to cause issues in gene expression studies. When a gene is effected by a blends of oils rather than single oils, it becomes difficult to predict which oil or molecules are causing the change.

### 1.10.4 Limitations of cell line studies

Many experiments utilize cell lines in their study designs. These studies play a vital role, serving as a baseline for the preliminary research of EOs. The results from cell line studies are not translational to human applications. When EOs are added to cell culture media, it creates a heavy systemic effect on the cells. EOs are volatile extracts, and generally hydrophobic in nature.[3]Their hydrophobic properties and varying viscosities can limit how well an oil can disperse evenly through the media. With so many people already using EOs in many areas of application, it is necessary for researchers to move EO studies to more clinically applied and translational models.

### 1.10.5 EOs versus plant extracts

EOs vary from plant extracts. Many studies have been conducted on the extracts of the *Boswellia sp.* (frankincense) plant. When an EO is steam or hydro distilled, large molecules of the plant matter are lost due to their heavy molecular weight.[43] Therefore, frankincense EO has variations in its chemical profile compared to frankincense extracts. Boswellic acids from frankincense extracts have been investigated in previous studies. These acids are said to have bio-active properties that can aid in cancer treatment and inflammation.[44][45] These acids are large and are not typically or easily captured in commercial standard distillation methods. There are a limited number of articles where boswellic acids were captured in the EO product.[46][47] If the acids are present in the EO, they are found in very trace amounts, as compared to the extracts and require specific distillation methods that are not set standards in the EO industry.[46][47] The findings from frankincense extracts and boswellic acid

studies are not directly translational for the findings of traditional frankincense EO.

## 1.11 Using bioinformatics based approaches to further investigate EOs

Using a bioinformatics skill set, our goal is to continue expanding our knowledge on the molecular mechanisms of EOs. Using computational based approaches such as text-mining, we built an EO knowledgebase (EOKB) that integrates data surrounding EOs from various data sources. The number of scientific resources available for EOs is low compared to the number of people that are actively using oils. Some of these resources make suggestions on the uses of EOs with no clinical data or scientific evidence. The purpose of the EOKB is to integrate different platforms of EO information from credible sources to better educate the consumer from a scientific perspective. It can also be utilized by researchers and medical professionals to learn more about the general research surrounding EOs. The EOKB can be used to find more information on the chemicals of EOs and their effects based on scientific research published in NCBI PubMed.[24] We hypothesize that we can utilize the EOKB to make a knowledgebase (KB) driven hypothesis, that we can then take to the lab to verify its predictive measures based off of the resources available in the EOKB. We hope to learn more about the molecular mechanisms of EOs and bring more scientifically driven data into the EO world, so users can make more scientifically sound decisions when using and educating others on EOs. Using the EOKB and information derived from different sources, we have chosen three main EOs to study. Some studies have investigated the basic anti-inflammatory potential of frankincense EO. There are fewer studies on myrrh (*Commiphora myrrha*) EO, and even less on the EO of blue spruce(*Picea pungens*). Although frankincense EO is one of the mostly widely known and used EO due to its ancient history, there is only a low number of clinical studies conducted on it. There is a large need to increase gene expression studies on EOs. Despite the high costs, RNA-seq will provide some insights on the capabilities of frankincense EO at the transcriptomic level.

## 1.12 Main aims of this dissertation

This dissertation has been broken down into four main chapters of research. Chapter 2 introduces the structure, design and implementation of the EOKB. Chapter 3 focuses on the testing of EOs on lipopolysaccharide (LPS)-induced *ex-vivo* human skin explants. Three different EOs were applied to human skin explants before inducing inflammation. The RNA was extracted and used for RT-qPCR to measure the effects of the EOs on pro-inflammatory markers. Later, the RNA from 30 samples were sequenced using Illumina HiSeq2500. Chapter 4 applies computational methods to compare the gene expression networks and associated pathways targeted by the frankincense EO treatment. The frankincense expression data will be used to compare its capabilities to other plant extracts, steroids and non-steroidal anti-inflammatory drugs (NSAIDs) investigated in previously published studies for inflammation. Chapter 5 is a continuation on the experimental studies described in chapter 3, with adjustments to the overall study design, with the purpose of improved targeted topical applications on the skin biopsy models.

# CHAPTER 2: THE DESIGN AND IMPLEMENTATION OF THE EO KNOWLEDGEBASE (EOKB) AND ITS UTILIZATION FOR ANALYZING EOS

## 2.1 Introduction

The use of text-mining and KB development have grown in the biological sciences due to their ability to integrate data from various sources. A KB differs from a traditional database (DB) in several ways. A DB is defined as a repository used to store, organize, and easily retrieve core data that belongs to a particular core system. DBs focus on quality assurance and control rather than data interpretation.[48] For example: A DB may contain a genome of an organism, and its sequences, but no further interpretation of the data. Online there are many interpretations of what a KB is comprised of. For biological purposes, in the NIH's strategic plan for data science, a KB is described as a type of DB that links growing bodies of information and require a significantly larger amount of human curation than standard DBs.[49] A KB contains a multitude of different data types. It allows information from different areas to be easily retrievable in one space for the user. As opposed to a traditional DB described above, a KB would include an organism's relative gene expression data, pathway networks, splicing variants, and text-mined data, in addition to its genome and sequences. It would also provide the means to create links and connections between data sources. A KB is often designed from scientific literature, but is curated for human use and interpretation.[49] This has become a fundamental method to collectively store various types of related scientific data in order to find new trends or patterns in the data. The application of a KB can lead to new discoveries or hypothesis generation for scientists. Researchers have created KB specific databases for different areas of research. Some of the most popular ones include: PharmGKB,

UniprotKB, and the Reactome Pathway KB.[50][51][52]

A free and publicly available, centralized KB for EOs does not currently exist. There is a need to provide a more in-depth KB for EOs. The only two scientific EO-based DBs that exist are the Essential Oil University Database (EOUdb) and the ESSential OIL DataBase (EssOilDB).[13][53] These two DBs contain only EOs by their varying names, their chemical composition and what study it was published in. Both are limited with the amount of information they contain pertaining to EOs. In 2018, a new database was published for EO related plants and research known as AromaDB.[54] However, all reported links to the database are unreachable. Three years ago, we began creating the essential oil knowledgebase (EOKB). It serves as an accessible and integrated KB source for researchers to learn and derive more information in regards to EOs. The EOKB can be utilized by anyone with the desire to increase their knowledge about EOs from a scientifically driven stand-point. One of its main benefits is its use to derive KB-driven hypotheses. It can also be applied to find validating/supporting evidence for the possible uses of different EOs with the ability to provide scientifically-driven data. This can provide new insights for unknown oils based on their chemical profiles. Information from numerous data sources have been integrated to develop new hypotheses for the EOs used in this dissertation. The data in the EOKB was updated in August 2020. This included text-mined data from NCBI PubMed, different EOs and their chemical profiles from EOUdb, gene and pathway information, proposed medical applications of EOs, and any available microarray/RNA-seq data of available single EO studies.[24][13][55][56][10][34] The EOKB served as a central location to store and organize all of the data generated and collected through out this dissertation.

### 2.1.1 Deriving a KB driven hypothesis

Using the EOKB we were able to predict whether the oils we tested in the lab would have an effect on inflammation based on the information already available on

each EO and its chemical compounds. Many EOs share similar chemical components. Existing data on an EO can be leveraged to make a prediction on the effects of an unknown EO. We used the EOKB in order to derive information already available on Frankincense EO, Myrrh EO and Blue Spruce EO. Tables 3.3-3.8 in Chapter 3, section 3.3, contain the GC/MS results of all three EOs used in the experimental assay in Chapter 3. Using the EOKB, the chemical compounds present in these oils can be queried to see if any previously published studies have examined these compounds. For example, one of the main chemical components of frankincense EO is alpha-pinene. Based on existing data, one study found alpha-pinene in Roman Chamomile increased gene expression of COX6C after inhalation of the oil.[57] Another study examined mice with cerulein-induced acute pancreatitis that were pretreated with alpha-pinene before being induced with cerulein, displayed decreased expression of pancreatic tumor necrosis factor (TNF), interleukin-1-beta (IL1$\beta$) and interleukin 6 (IL6).[58] All of these study findings could be translational for the application of frankincense EO, due to its high content of alpha-pinene. This information can aid in the future hypotheses for EOs that may share similar mechanisms of action based on their chemical constituents.

## 2.2 Methods

Data for the EOKB was collected from publicly available databases containing any and all relevant scientific information in regards to EOs. An exception to this was the proposed medical applications of EOs. These were provided from a reference book, and consent from the publishing company was given for the information to be used in the EOKB.[10]

### 2.2.1 Data collection

#### 2.2.1.1 Linguamatics for text-mining

Linguamatics, i2E is a text-mining software that provided the tools to text-mine data from published research using natural language processing (NLP).[59] The software takes libraries of information provided to it and indexes it into the software. The user has the freedom to build very complex and nested search queries using built in ontologies to extract information out of the source that was provided to it. Linguamatics was leveraged to text-mine through abstracts indexed on NCBI PubMed for EO based literature for the EOKB. When designing queries, it seemed best to implement each item as its own query. As there are multiple ways to describe EOs, the EO query was designed to look for: essential oil(s), volatile oil(s), aromatherapy, ethereal oil(s), and aethrerolea. The chemical query was designed to find any chemicals related to EOs. The chemical query was given the extracted list of over 4000 EO based chemicals from the EOUdb.[13]. This allows for this query to find any articles related to these chemicals that may have been conducted, even if they did not mention an EO specifically. A gene text-mining query was built onto a previously existing multi-query package designed by Linguamatics within their example query package. Their original query in figure 2.1 was from the subtraction function underneath the gene join, and everything underneath it. This query was previously optimized by Linguamatics to aid in the extraction of Entrez human genes from the material given to it. Their query had many built in layers to reduce the amount of noise while looking for genes. Within this built-in query they nested other queries to reduce errors. As seen in the figure 2.1, these queries searched and filtered for genes that may have contained a Greek letter in the gene symbol, genes that had multiple synonyms, and genes that were listed with other genes. Even with all of these layers/filters for noise reduction, queries were further altered and leveraged into the multi-query searches. Specific words were negated from literature that often got

picked up by the software as a gene in order to increase accuracy of the queries. These queries were all combined in numerous combinations to search the text for different purposes.

The EO query, gene query, and chemical query were combined in a Gene_Chem_ Association multi-query to search the text for all gene-chemical associations found in EO-related studies. The multi-query was given key words to look for, that may help to indicate the type of relationship present between the genes and the chemicals. This included variations of gene expression related key words: Induce, increase/decrease gene expression, regulate, reduce, express, increase/decrease activity, etc. Experimental keywords and their variations it included were: RNA-seq, PCR, microarrays, etc. With the chemical list, key words to aid in extrapolating relationships, and built in queries created by Linguamatics, the Gene_Chemical_ Association table in table 1, figure 2.2, in the Entity-relationship diagram explained in section 2.2.2 was populated. Linguamatics was also used to build the Plant_Chemicals table and the plant_species tables in the ER-diagram (figure 2.2) All of the resulting data sets extrapolated from Linguamatics were built into PostgreSQL, a relational database management system (RDMS).[60]

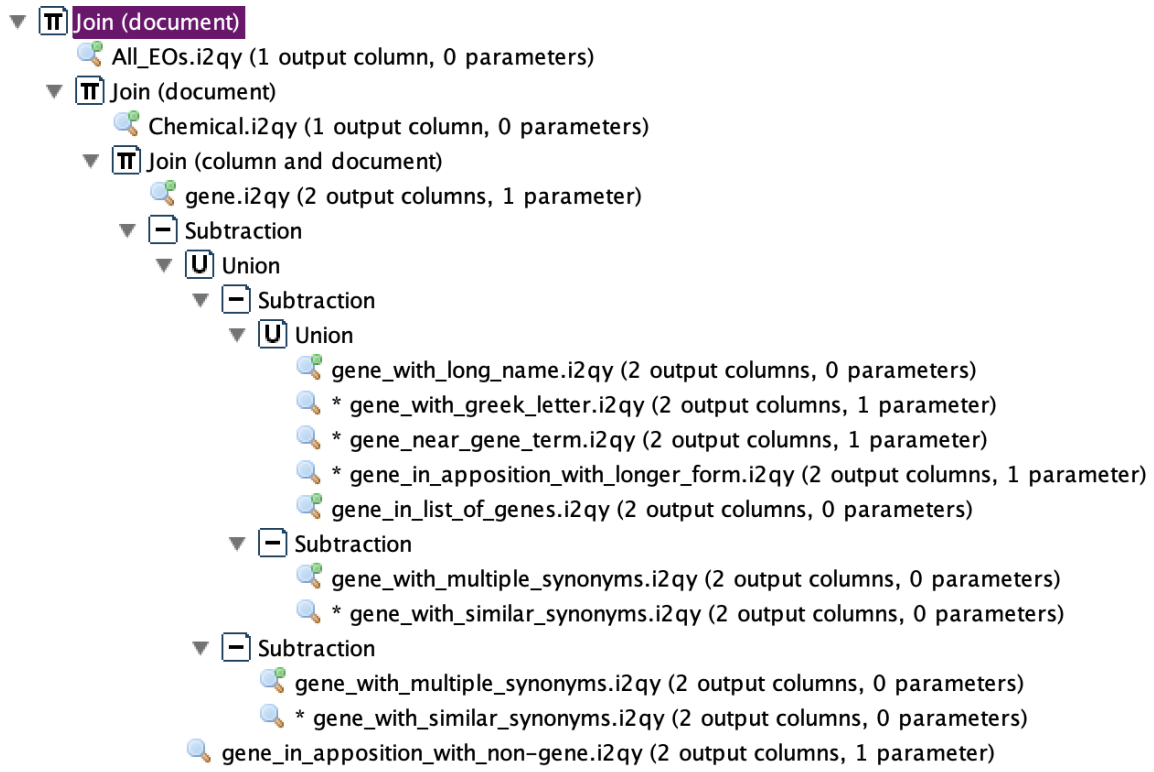


Figure 2.1: Multi-query used for Gene_Chemical_Association table and other related EO tables illustrated within the ER diagram

#### 2.2.1.2 Comparative Toxicogenomics Database

The Comparative Toxicogenomics Database (CTD) is a publicly available database that manually curates its own gene-chemical associations related to human health applications, designed by NC State University.[61] CTD manually curates their own gene-chemical associations from published literature based on their designed hierarchical interaction-type vocabulary list that aims to pull out key words in relation to gene and chemicals.[62] The purpose of CTD is to study the affects of environmental exposures on human health.[62] As these results can differ from what Linguamatics was able to curate, the associations generated from CTD were added to the EOKB to provide a wider variety of possible associations. The same list of EO compounds from the EOUdb was used as the input chemical list for generating the chemical-gene associations from the CTD.[13]

#### 2.2.1.3 Microarray and RNA-sequencing data

Any published microarray or gene expression data set available in NCBI GEO, was added to the EOKB if the study was completed on a single oil and demonstrated on human samples. Any studies using a EO blend was excluded from the EOKB due to the variability within blends. The Log2 fold changes(Log2FC) of the different studies were calculated and added to the EOKB using GEO2R.[34] Access to available gene expression data, allows for the user to connect gene associations from one EO experiment to other gene-chemical associations from other studies to form better predictive methods on the possible mechanisms of certain EOs based on its chemical components and associated data. Three gene expression data sets included in the EOKB thus far were Sandalwood EO treated on 3D skin explants, rosemary supercritical extract tested on colon cancer cells, and frankincense/sandalwood treated on bladder cancer cells.[35][63][64] To minimize complexity of the ER diagram only one general table is displayed for gene expression data sets. These gene expression studies have varying testing conditions in place for their studies. Some are carried out on cancerous cell lines with no controls of using healthy cells. When utilizing them for analysis, it is the users responsibility to keep the experimental conditions in mind.

#### 2.2.1.4 Gene and pathway information

Gene information was downloaded from NCBI gene.[56] The entire gene set of recorded genes in humans on NCBI was downloaded and parsed into the gene-info table 7, figure 2.3. This table contains gene IDs, gene symbols, aliases, description, chromosome number/position and species information, and more. This table can be joined with the Chem_Gene_Association table to find more information on specific genes that may be of interest related to a specific chemical. A three way table can be created connected genes of interest to gene information and pathway information.

#### 2.2.1.5 Pathway information

Pathway information was extracted from David.[65][66] David is a bioinformatics tool that takes an input gene list and has the option to connect it to GO terms, associated pathways from Kegg, functional annotations, and OMIM diseases.[67] [68] The Pathway_information table, table 8, figure 2.3 was curated by David using the Entrez humans gene ID list from NCBI. David then curated all of the pathway information related to the genes given in the query. The results were downloaded and parsed. The results were added to the pathway-information table. This table includes gene ID, gene name, organism, associated pathway and any possible related OMIM diseases.

#### 2.2.1.6 Proposed medical applications

Most EO users rely on published reference materials for medical applications. One of the most popular reference materials is Life Science Publishing's Essential Oil Desk Reference.[10] Proposed medical applications from this book were extracted and added to the EOKB to utilize these applications in conjunction with the available published research for EOs. These applications were added to their own table 10 in the EOKB containing the oil name, the plant family, genus and species, proposed medical properties and their proposed medical uses, seen in figure 2.2.

#### 2.2.1.7 Essential Oil University Database

The Essential Oil University Database (EOUdb) was designed by Dr. Robert Pappas and contains 4,221 chemical compounds recorded in 4,152 different EOs.[13] This database lists the full chemical compositions of different recorded oils and where they were published. With permission, we extracted the chemical compounds and their respective oils from this database using web scraping techniques in python. This data was loaded into the EOKB. The list of chemical compounds recorded in EO studies were utilized to leverage more Linguamatics queries to find all possible EO based

compounds mentioned in any Medline literature. This allows the user to look for a specific compound, see what different oils this compound is found in, and then look back through the Linguamatics curated data to see the published studies where this chemical has been investigated.

#### 2.2.1.8 DrugBank

DrugBank is a database containing detailed worldwide information on drug data and drug targeting. The company and developers for this database are based in Canada. This database provides information on only a handful of EOs. It also has a few entries of EO constituents that are found in OTC medications as well as approved pharmaceuticals.[22] DrugBank has a freely available XML dump of its database contents that can be re-purposed for research applications. The full list of drugs and their respective gene targets were parsed from the XML and added into the EOKB to make comparisons of predicted gene targets for EO chemical compounds.

### 2.2.2 Development and design of the EOKB

#### 2.2.2.1 Entity-relationship schema

Shown below in figures 2.2 and 2.3 is the entity-relationship (ER) diagram of the EOKB in Crow's foot notation generated by Lucid Chart.[69] The ER diagram was split in half into two figures to fit on this document. The rectangles indicate the different tables (entities) contained within the database, and the rows in each table indicate the columns of that specific table known as the table's attributes. The columns on the right of each table indicates the data type and constraints for each specific attribute. For simplicity, the schema only displays one entity of microarray data, and one entity of the RNA-seq data generated from the experiments in chapters 3 and 4. Tables within the ER-diagram have been numbered for simplicity when discussing relevant information specific to each entity.

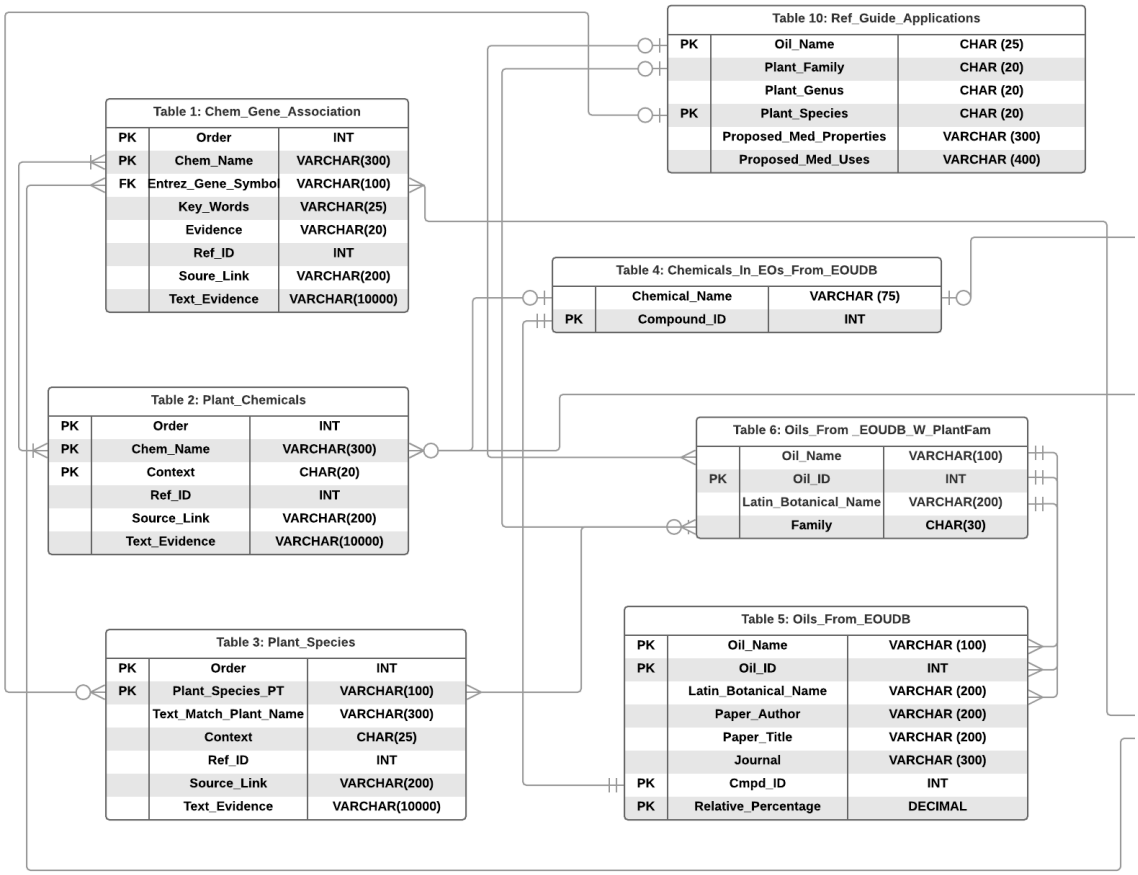


Figure 2.2: Entity-relationship diagram for EOKB part 1

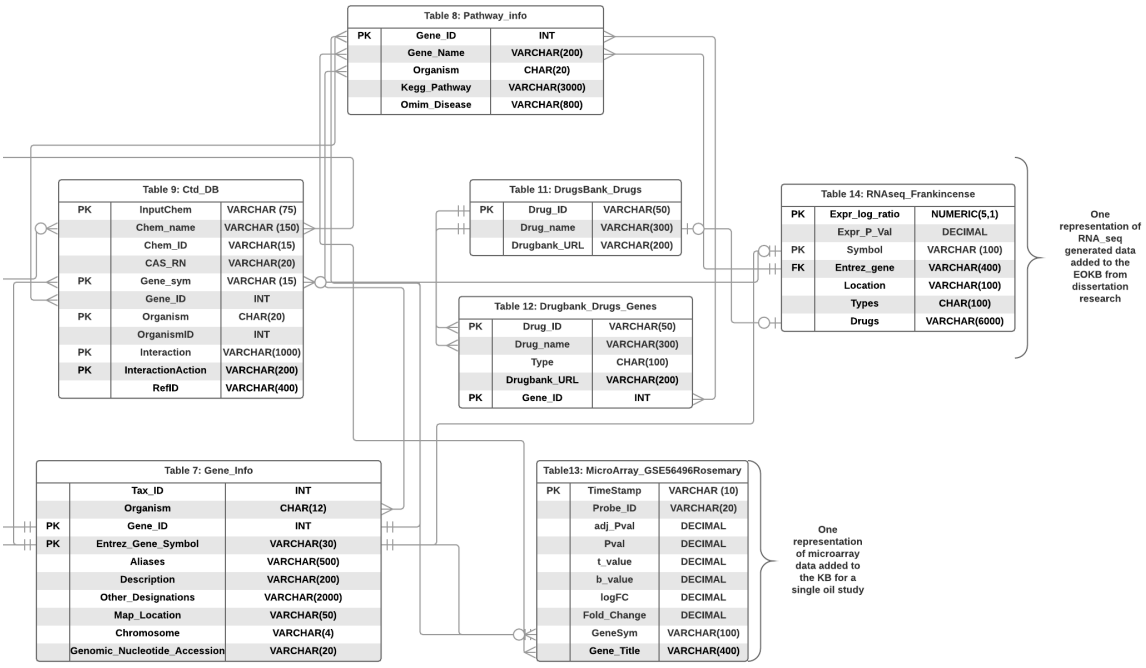


Figure 2.3: Entity-relationship diagram for EOKB part 2

On the left column of each entity displays the appropriate keys for the table. Keys are used in relational-database management systems (RDMS) in order to identify a specific row of data in the table. Tables can have multiple keys depending on what attributes they contain. A primary key (PK) is a type of candidate key (CK) that defines itself as the main key for the table. A PK must uniquely identify each record in its corresponding table and cannot contain null value. A candidate key (CK) is the minimum number of combined fields that can uniquely identify each record within the table, But each key is an independent standalone primary key (PK). A CK is a type of PK when combined with another or multiple PKs. Choosing only 1 PK for each entity was nearly impossible. Due to the complexity of the data, many records were not unique within each column alone. For example, table 1 in figure 2.2 displays the Chem_Gene_Association table. This table was created from the information curated by text-mining. Within this table, chemical name, gene symbol, reference ID, are all attributes of the entity. However, none of these attributes are unique when

they stand alone. The text-mining query will find the same chemical compounds and gene symbols within different studies. However, when you combine the gene name, chemical name and reference ID records, all three of them combined can create a primary key known as a composite key (CompK). A CompK is a key that is created based on the minimum number of fields that can be combined in order to uniquely identify each record in the table.[48] These attributes combined create a key that can uniquely identify each record in the table, but cannot be their own key independently. Due to the repetitive nature of the biological data in the EOKB, many tables used CompK (multiple combined attributes) as their PK. When a primary key of one table is listed in another table but not unique to its according entity, it becomes a foreign key (FK). An example of this would be in figure 2.3 table 7: Gene_Info, its PK is gene symbol. As this attribute is unique to each record in this entity. But when connecting table 7 to table 1, entrez gene symbol within table 1 now becomes a FK to table 7, as it is not uniquely identified in each record in table 1. Keys are important in RDMS as they exhibit the integrity within the design of the database. This often gets difficult in biologically driven databases, such as this one. It is not ideal for an entity to need multiple combined attributes to create PKs/CompKs, however, in the text-mining results it is necessary as the data can become very repetitive, even within each article it pulled the information from.

Relationships describe the associations between the different entities using the various arrows seen in the ER diagram. These arrows, known as Crow's foot notation, describe the type of association or connectivity between two entities. Figure 2.4 explains the different type of relationships that are seen on the ER diagram.

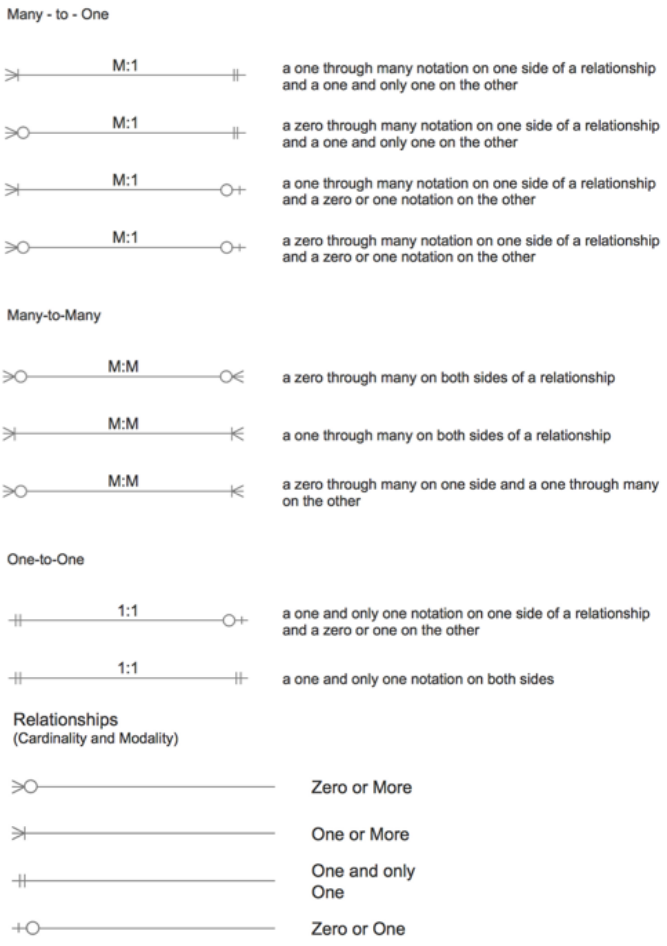


Figure 2.4: Legend for Crow's foot notation

In the ER diagram one can see numerous "M:M" relationships, due to the entities sharing many similar attributes such as chemicals and genes. In figure 2.3 table 7: Gene_Info we can see a M:M relationship with table 8: pathway info. As attributes of the Gene_Info table may have multiple associated relationships to the pathway table because many genes are involved in numerous metabolic pathway and not just 1. For example: Gene A from Gene_Info table 7, can connect to many different metabolic pathways in the Kegg pathway table. Gene B from the Gene_Info table 7, can also connect to the same and different metabolic pathways in table 8, including the same ones as Gene A connected to. In some instances there may be a 0 or 1:1 relationship. Meaning one attribute may connect to another tables attributes 1 time,

or 0 times.

### 2.2.3 ETL mapping and programming

Each set of data had to be parsed and organized according to its purpose and use for querying in the EOKB. Programming in both Python and bash were utilized for the extraction, transformation, and loading (ETL) of the large quantities of data that were collected and added into the EOKB. The data was loaded accordingly into a RDMS in PostgreSQL according to the database design shown in figures 2.2 and 2.3.

### 2.2.4 Structure query language (SQL) joins and analysis

Using SQL coding, table joins were formed to create linkages between genes, their relative information and associated pathways. Instead of the user having to manually go between databases to find more supporting evidence for a gene and its related EO research, the information has already been integrated and is available within the EOKB. Figure 2.5 demonstrates how the EOKB can be useful in finding more supporting evidence.

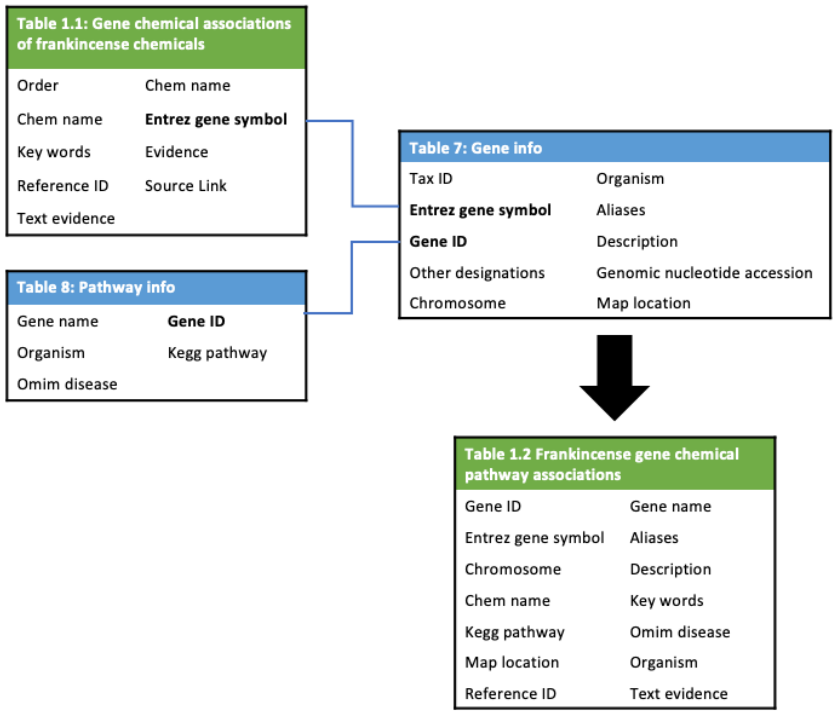


Figure 2.5: Utilizing the EOKB to connect chemical-gene associations with gene and pathway information

In regards to figure 2.5, table 1.1 contains the gene-chemical associations related to the compounds found in frankincense. Each statement in the body of the table describes a column of data found in that table. The columns used for the table joins are bold. Table 1.1, table 7 and table 8 can all be joined to form a new table 1.2 that now contains combined information from the 3 tables specific to frankincense. Table 1.2 now contains the gene-chemical associations of frankincense and additional information on each gene (chromosome, location, name, aliases) joined from table 7. Table 1.2 also contains pathway and OMIM disease associations relative to those genes from table 8. Shown in figure 2.4, table 1.1 can also be used to find out what other oils may share similar chemical compounds with Frankincense EO. Tables 4 and 5 in figure 2.4 can be used in conjunction with table 1.1 to produce table 1.3. Tables 4 and 5 can be joined on compound IDs and join with table 1.1 on chemical names. This join produces table 1.3 which now contains all other oils from previously

published studies that share common chemicals with frankincense along with their specific composition percentages.

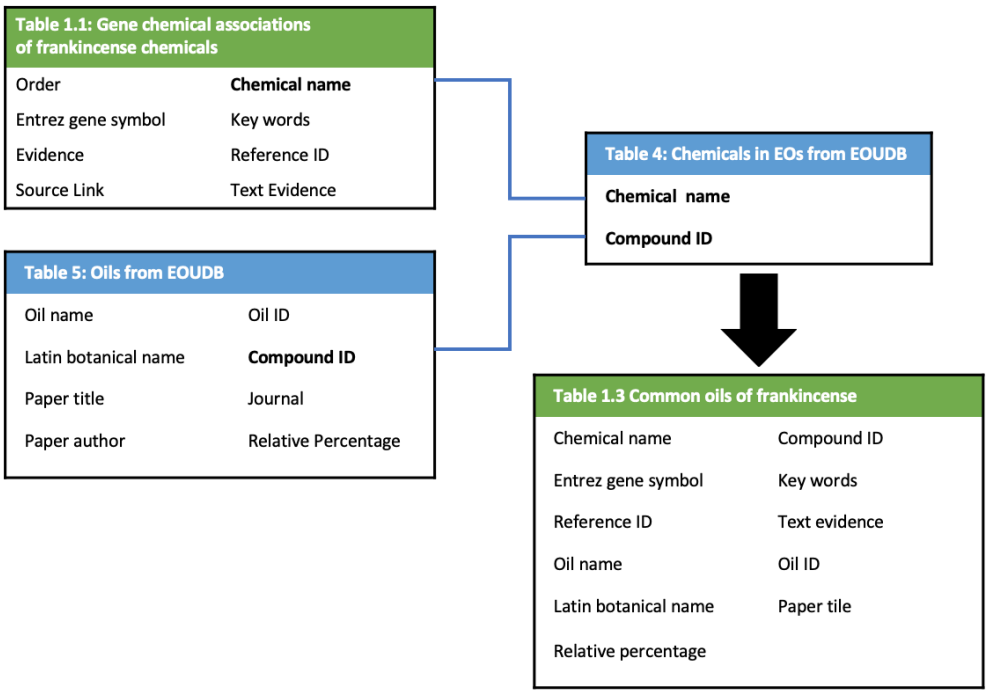


Figure 2.6: Utilizing the EOKB to connect chemical-gene associations with other EOs that share similar chemical components with frankincense EO

Using the EOKB we can learn and discover new trends and patterns in existing data and apply it to oils we know very little about. This could provide more supporting evidence for their uses or provide insight on areas that still need to be researched and explored.

## 2.3 Results

### 2.3.1 EOKB query

Data already available on existing EO compounds can be leveraged to predict the possible effects of the oils used in this project. Figures 2.7, 2.8 and 2.9 display only a subset of chemicals, genes and pathways that the EOKB had previously existing gene-chemical data for. Please refer to the supplementary file for the full query results

for each of these EOs based on their chemical profile.

Each oil is displaying only a subset of their chemical composition with their respective relative area percentages in the blue rectangles. The grey squares are a subset of associated genes for that chemical based on the EOKB findings. To the right of the figures are the listed proposed medical applications of each oil.[10] Below the proposed medical applications displays partial results of the associated pathways.

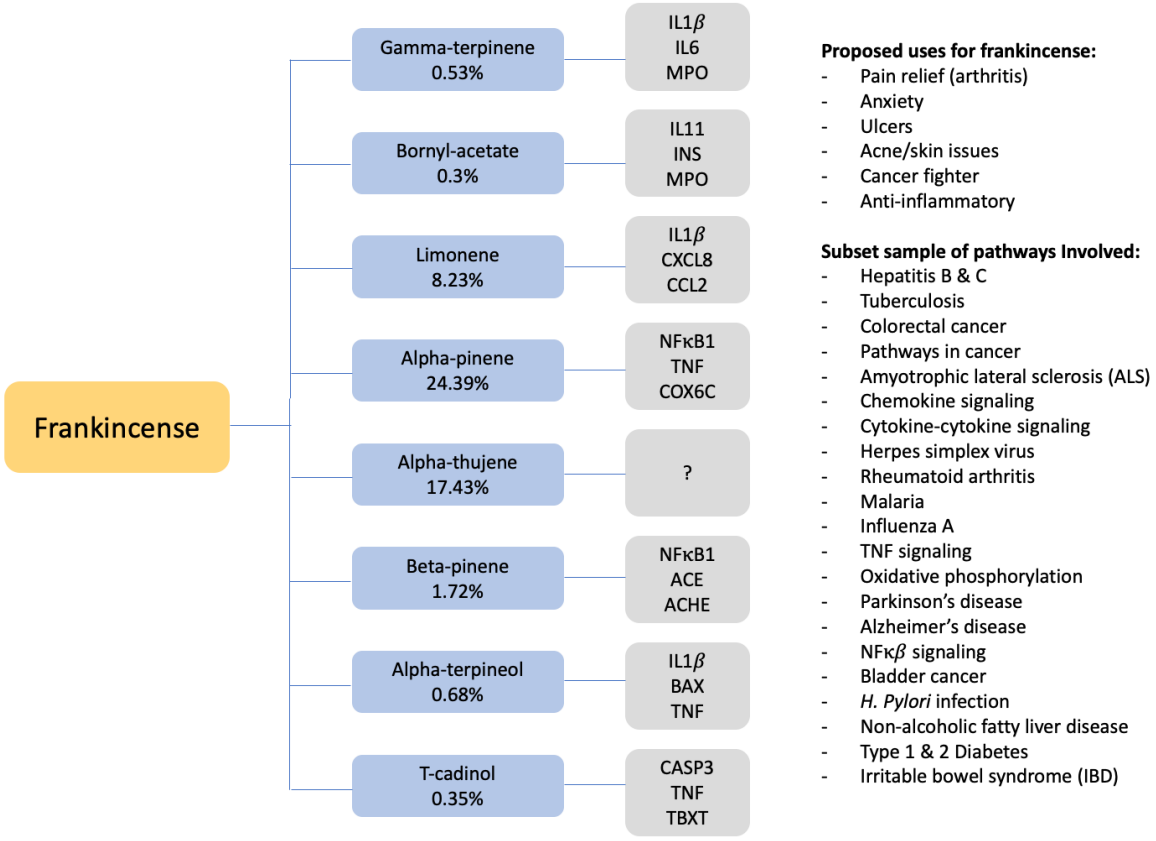


Figure 2.7: EOKB driven hypothesis for frankincense

Frankincense displayed associations with many inflammatory-mediated cytokines. Many of these genes play key roles in inflammatory-mediated pathways.

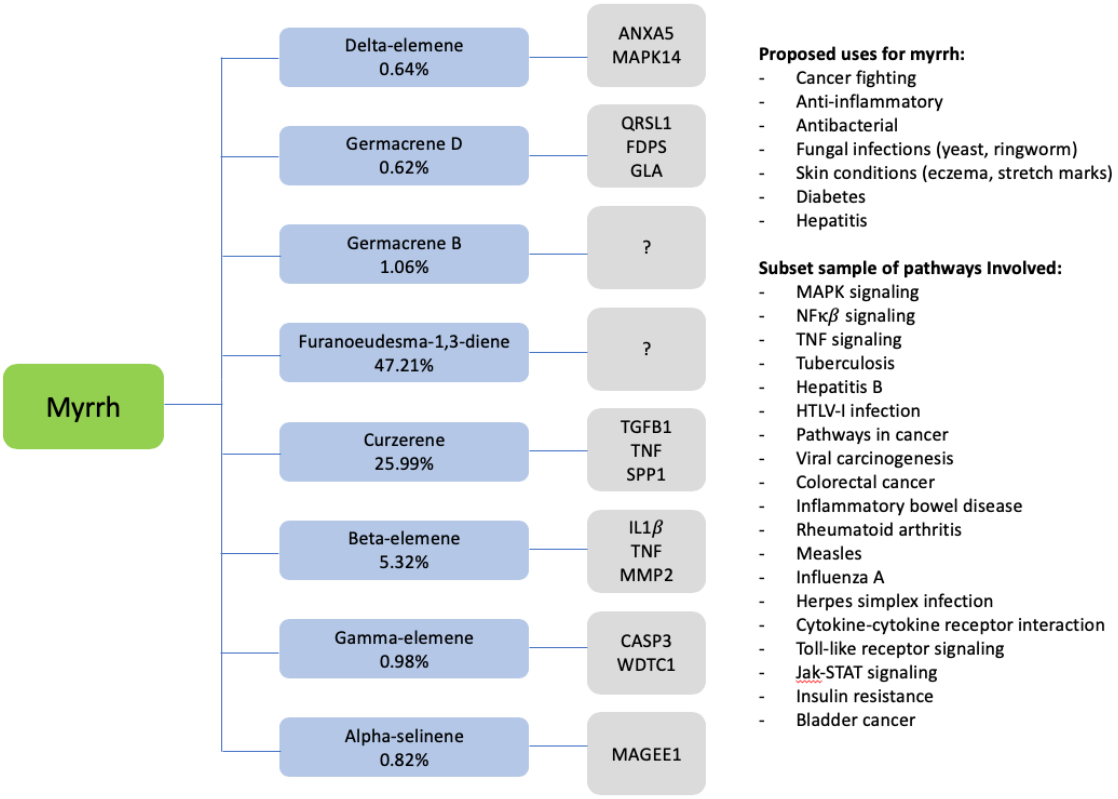


Figure 2.8: EOKB driven hypothesis for myrrh

The EOKB did not contain results for all of Myrrh's chemical components. As stated previously, the chemical profile of myrrh is quite unique. Germacrene B and Furanoedudema-1,3-diene resulted in no gene associations from the text-mined literature. This is not surprising considering the low number of studies published on myrhh. Myrrh did contain chemical components associated with some inflammatory-mediated genes such as IL1$\beta$, TNF, and MMP2. Myrrh has an interesting chemical profile that should be pursued further in EO research studies.

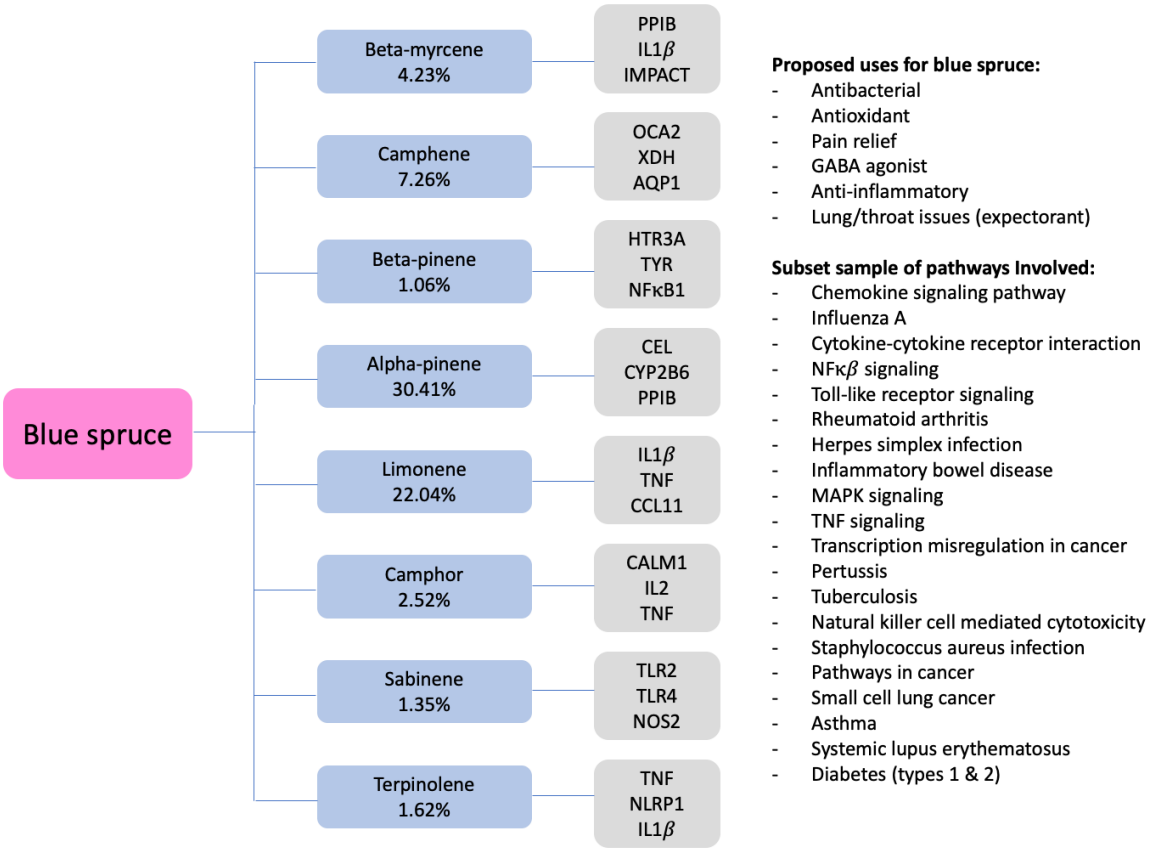


Figure 2.9: EOKB driven hypothesis for blue spruce

Due to the overlap of similar chemical compounds in blue spruce and frankincense, the two queries resulted in very similar gene sets and associated pathways. It will be interesting to observe whether the two will have the same effect on inflammation when they are further studied in Chapter 3.

### 2.3.2 Stringdb results

Stringdb is a publicly available database to study protein-protein interactions based on indirect and direct relationships.[70] Stringdb connects associated pathways based on the gene set given to the program. It counts the number of genes associated with pathways in Kegg and Reactome.[52][55] These pathways are based on the percentage of genes present in related networks from the input data set.[70] The genes generated from the EOKB query were input into Stringdb. The results were used to

visualize the gene interaction networks from the data generated from the EOKB. The results display tables of pathway information, the number of genes from the gene set in association with each relative pathway, and their respective false discovery rates (FDR).

Table 2.1: Kegg pathway Stringdb results for frankincense

| **KEGG pathways** | | | |
|---|---|---|---|
| **Pathway** | **Description** | **Gene count** | **FDR** |
| hsa05161 | Hepatitis B | 35/142 | 1.65E-22 |
| hsa05200 | Pathways in cancer | 56/515 | 3.65E-21 |
| hsa05167 | Kaposi's sarcoma-associated herpes virus infection | 36/183 | 8.83E-21 |
| hsa05152 | Tuberculosis | 35/172 | 1.01E-20 |
| hsa04933 | AGE-RAGE signaling pathway diabetic complications | 28/98 | 6.74E-20 |

Table 2.2: Reactome pathway Stringdb results for frankincense

| **Reactome pathways** | | | |
|---|---|---|---|
| **Pathway** | **Description** | **Gene count** | **FDR** |
| hsa-168256 | Immune system | 106/1925 | 2.04E-18 |
| hsa-1280215 | Cytokine signaling in immune system | 60/654 | 2.04E-18 |
| hsa-6785807 | IL-4 and IL-13 signaling | 28/106 | 2.70E-18 |
| hsa-449147 | Signaling by interleukins | 49/439 | 2.70E-18 |
| hsa-162582 | Signal transduction | 116/2605 | 4.76E-14 |

Table 2.3: Kegg pathway Stringdb results for myrrh

| **KEGG pathways** | | | |
|---|---|---|---|
| **Pathway** | **Description** | **Gene count** | **FDR** |
| hsa05206 | microRNAs in cancer | 17/149 | 5.21E-17 |
| hsa05205 | Proteoglycans in cancer | 17/195 | 1.74E-15 |
| hsa05200 | Pathways in cancer | 23/515 | 1.74E-15 |
| hsa04933 | AGE-RAGE signaling pathway diabetic complications | 13/98 | 4.88E-14 |
| hsa05161 | Hepatitis B | 14/142 | 1.26E-13 |

Table 2.4: Reactome pathway Stringdb results for myrrh

| **Reactome pathways** | | | |
|---|---|---|---|
| **Pathway** | **Description** | **Gene count** | **FDR** |
| hsa-6785807 | IL-4 and IL-13 signaling | 14/106 | 5.36E-14 |
| hsa-449147 | Signaling by interleukins | 17/439 | 1.86E-09 |
| hsa-1280215 | Cytokine signaling in immune system | 19/654 | 7.32E-09 |
| hsa-162582 | Signal transduction | 34/2605 | 6.14E-08 |
| hsa-168256 | Immune system | 29/1925 | 6.84E-08 |

Table 2.5: Kegg pathway Stringdb results for blue spruce

| **KEGG pathways** | | | |
|---|---|---|---|
| **Pathway** | **Description** | **Gene count** | **FDR** |
| hsa05161 | Hepatitis B | 29/142 | 8.03E-19 |
| hsa05152 | Tuberculosis | 29/172 | 4.18E-17 |
| hsa05200 | Pathways in cancer | 45/515 | 7.29E-17 |
| hsa05167 | Kaposi's sarcoma-associated herpes virus infection | 29/183 | 9.37E-17 |
| hsa05210 | Colorectal cancer | 22/85 | 1.51E-16 |

Table 2.6: Reactome pathway Stringdb results for blue spruce

| **Reactome pathways** | | | |
|---|---|---|---|
| **Pathway** | **Description** | **Gene count** | **FDR** |
| hsa-1280215 | Cytokine signaling in immune system | 53/654 | 1.21E-17 |
| hsa-449147 | Signaling by interleukins | 43/439 | 5.93E-17 |
| hsa-6785807 | IL-4 and IL-13 signaling | 24/106 | 4.36E-16 |
| hsa-168256 | Immune system | 82/1925 | 7.75E-13 |
| hsa-8953897 | Cellular responses to external stimuli | 37/459 | 3.18E-12 |

## 2.4 Discussion

Many chemical compounds studied in previous literature found gene associations related inflammatory-mediated pathways. Some of these genes include: Tumor necrosis factor (TNF), toll-like receptor 4 (TLR4), interleukins (IL-1$\beta$, IL-6, IL-11), nuclear factor kappa B1 (NF$\kappa$B1), matrix metallopeptidases (MMP9 and MMP2) and others. (Please see supplementary file for a full list of results from each EOKB query.) In total, the EOKB query found 396 genes associated with the chemical constituents found in frankincense EO, 82 genes associated with the chemicals found in myrrh EO, and 325 genes associated with blue spruce EO. It is interesting to see the number of associated pathways based on the gene associations predicted from the literature. Cytokine-cytokine signaling, chemokine signaling, rheumatoid arthritis, TNF signaling and NF$\kappa$B signaling are all pathways that mediate inflammation.

All 3 EOs displayed associations with inflammatory-based pathways/networks, including cancer pathways. Since frankincense and blue spruce share many chemical compounds, their related pathway associations results are quite similar. Myrrh EO also contained many associated pathways involved with interleukin signaling, and more interestingly many cancer pathway associations. One of the top proposed uses of Myrrh EO is its application for fighting cancers.

This may help support proposed uses for frankincense, blue spruce and myrrh for inflammatory applications. It can also be utilized to discover new purposes for frankincense based on its similarities to other EO studies. Inflammation can effect different genes depending on where it is located. Knowing which genes and pathways are being targeting can provide more specific and targeted applications for EOs rather than broad uses. Using the information already available in this manner can save time and money for researchers. Based on the figures from the query results and Stringdb, these oils could support uses for inflammation and provide enough preliminary data to test more specifically for this application.

### 2.4.1 Challenges and limitations

Due to the large size of the multi-query, the software was often overloaded with too many results. Queries had to be carried out in batches with for example, 500 EO related chemical compounds at each submission. To speed up queries in the future, a general chemical list could be designed and maintained for future EO-related queries. Furthermore, the ever changing research and published literature requires indexes to be updated constantly and queries needing to be rerun or adjusted to have the most recent data available in the EOKB.

With all the specific details and parameters within the text-mining queries, there was still some noise that was hard to overcome. This will always be a challenge due to the complexity of the English language and scientific literature. Some scientific acronyms can be easily be picked up as a gene name due to the structure of the acronym. Due to this, each query result within the table in the EOKB includes the Ref_ID, which is the PubMed ID in order to trace the query back to the article to verify whether the relationship extracted by the query is accurate.

# CHAPTER 3: STUDYING THE EFFECTS OF EOs ON LPS-STIMULATED EX-VIVO HUMAN SKIN EXPLANTS

## 3.1 Introduction

### 3.1.1 Inflammation

One of the more commonly used application of EOs is for inflammation. EOs are utilized by users for both acute inflammation and chronic inflammation relief. Acute inflammation is commonly caused by tissue damage. This type of tissue damage usually arises after trauma, infections, stress, exposure to certain chemical compounds or environmental pollutants.[71] Acute inflammation typically arises quickly but tends to dissipate after a few days. Chronic inflammation occurs when the inflammation is prolonged for periods of months to years.[72] Chronic inflammation can lead to an array of chronic diseases. Some of these include: Alzheimer's disease, arthritis, cancer, Chrohn's disease, multiple sclerosis and Parkinson's disease.[73] Chronic diseases affect more than 133 million Americans.[74] Many often end up with multiple conditions. Due to this high number, the majority of health care costs are due to chronic illnesses.[74] The most common treatments for some of these chronic diseases include steroids, Nonsteroidal anti-inflammatory drugs (NSAIDs), or monoclonal antibody drugs.[75] Unfortunately, many of these drugs can cause complicated side effects for the patient. Many suffering are beginning to look at alternative forms of treatment for pain and symptom management such as life style changes, nutrition, acupuncture, herbal supplements and EOs. Some of the recommended oils for inflammation according to numerous references included: ginger, lavender, thyme, frankincense, patchouli, fennel, peppermint and among many many more.[10][76][77][78][79][80] A

review study published in *Oxidative Medicine and Cellular Longevity*, reviewed over 20 different studies that explored the use of EOs for inflammation.[81] It reviews key EOs that have been explored for inflammation as well as different methods for the testing and application of EOs for inflammation. Some EOs seem to have the potential to block mitogen-activated protein kinase (MAPK) pathways as well as the potential to inhibit NF-KB pathways by reducing oxidative stress.[81]

To investigate the effects of various EOs on inflammation, we utilized a franz cell diffusion system, developed by PermeGear. This system allows for pre-clinical testing for transdermal drug administration.[82] A systemic inflammatory response was induced within these culture models through the supplementation of lipopolysaccharide(LPS) within the medium. LPS is an endotoxin derived from the cell wall of gram negative bacteria and stimulates an innate immune system response.[83] Toll like receptors (TLRs) present on macrophages and dendritic cells can detect different pathogen associated molecular patterns (PAMPs), including the presence of LPS.[84] TLRs will activate the transcription factor nuclear factor-kappaB (NF$\kappa$B).[85] This activation of the NF$\kappa$B signaling pathway regulates the body's immune responses to infections by inducing the expression of various cytokines and chemokines.[86] Incorrect activation of NF$\kappa$B has been linked to the cause of autoimmune diseases, chronic inflammation and cancer.[84] The addition of LPS will induce the release of pro-inflammatory cytokines (IL6, IL4, IL1$\beta$, IL1$\alpha$, IL15, IL17A, IL8 and COX2), pro-inflammatory chemokines (CCL2, CCL3, CCL5, CXCL1, CXCL2 and CXCL10) and matrix metalloproteinases (MMPs).[87] [88][89] Most of these cytokines/chemokines are produced by neutrophils, T cells, dendritic cells, macrophages and natural killer cells.[90][84] LPS is a common method for introducing inflammation in animal, cell and tissue models for testing therapeutics.

### 3.1.2 EOs of interest for testing purposes

Upon conducting a heavy literature review, several EOs were chosen as test subjects for this experiment through the utilization of the EOKB in chapter 2. Frankincense EO was chosen due to its long held high value through out ancient history. Some studies have been done on frankincense oils and frankincense extracts that have suggested the anti-inflammatory effects of frankincense. However, many of these studies investigate the therapeutic use of boswellic acids (BAs). BAs are a chemical compound contained within the resin of the frankincense plant. Due to their large molecular size, they are lost during traditional/standard steam and hydro distillation methods for commercial EO production. BA can be selected for in trace amounts in frankincense oil under very specific distillation methods and identified using HPLC.[91][47] Many studies have claimed that boswellic acids posses powerful therapeutic properties for inflammation and cancer.[92] After the review of these articles, frankincense EO became the main oil of interest for this study to test whether the same potential exists within the EO as the frankincense extract. Myrrh became a secondary oil of interest due to its lack of supportive research but its ample support of human testimonies and recommendations for its use for inflammation. One study investigated the use of myrrh combined with constituents of frankincense for their combined possible anti-inflammatory mechanisms.[93] Myrrh has a chemical profile that is somewhat unique of its own in comparison to other EOs. Due to its minimal scientific research, but similar composition of frankincense EO, blue spruce EO was the third oil chosen for this study. The chemical components they have in common are alpha-pinene, beta-pinene, limonene, camphene, sabinene, myrcene, bornyl acetate and terpinolene. The chemical composition and acquisition of the oils used in this study are shown in section 3.3.

The oils were tested separately for their topical and systemic efficacy. The topical application is more translation to current human use applications. The involves

adding the EO directly onto the skin to target an area of interest and rubbing the oil in. Some users do this by diluting the oil with a fatty-acid based oil before application or by applying the oil neat. Systemic application required adding the EO into the media that resides underneath the skin tissue. This method is not a used or recommended application for EOs*. It was implemented to observe how much the results could differ with two very different styles of application. Fractionated coconut oil (FCO) served as an oil based control for this experiment. FCO is one of the most commonly used carrier oils among EO users due to its stable shelf life, availability and low cost.[94] EOs do not disperse evenly within medium solutions due to their hydrophobic nature. Therefore, we added an oil based control to verify that the results were not dependent on using any oil-based substance. The oils were added on or below the skin tissue *prior* to inducing the LPS response. This may seem counter-intuitive as users of EOs typically apply oils after experiencing some type of inflammation and not as often in preventative measures. The presence of LPS in the skin models will begin to reduce viability within the skin tissue. The viability of skin biopsy samples is not long. This is one main caveats of working with biopsy samples. For this reason, the EOs needed to be added prior to the LPS and incubated for 24 hours for our first set of assays. The tissue was harvested after the experimental trial and the RNA was extracted to measure the effects of the oils on 4 different pro-inflammatory cytokines (CXCL10, COX2, IL1$\beta$ and IL8) with RT-qPCR. Later, 30/100 samples were sent for RNA-seq (the frankincense samples and controls from donor 1 (D1) ).

## 3.2 Methods

### 3.2.1 EO acquisition

Three different EOs used in this study were obtained from Young Living®Essential Oils, an EO company based in Utah. *Boswellia carterii*(Frankincense EO) and *Commiphora myrrha*(Myrrh EO), were both produced from their respective trees culti-

vated in Somalia. *Picea pungens* (Blue Spruce EO) was distilled from blue spruce trees in Idaho, United States. The chemical compositions of these oils were determined by gas chromatography/mass spectometry (GC/MS) and are listed accordingly in section 3.3. The optical rotation, specific gravity and refractive index of each oil was determined as well.

The fatty acid composition of the fractionated *Cocos nucifera* (Coconut) oil (FCO) was determined in collaboration with Creative Proteomics (Shirley, NY, USA) using fatty acid methyl esters (FAME) profiling. The fatty acids were identified via direct methylation in the presence of known amounts of added tridecanoin and methyl tricosonate. The fatty acid methyl esters were prepared using boron trichloride in methanol and heating the methylation tubes in a heating block at 95°C for 30 minutes. The fatty acid methyl esters were then analyzed by a Agilent 7890B gas-liquid chromatograph with a 60-m DB-23 capillary column. Standard mixtures for both qualitative and quantitative analysis along with the known fatty acid components for retention time verification were used. Internal standards were obtained from NuChek Prep (Elysian, MN, USA). Quality control standards were purchased from Millipore-Sigma Canada (Oakville, ON, Canada). Recovery was determined by comparison of peak area between internal standards (90-110%). Results of the analysis are in section 3.3.

### 3.2.2 Human skin preparation

All skin tissue acquisition and testing was completed in collaboration with Tergus Pharma (Durham, NC, USA). The human skin tissue used in these experiments were ethically acquired from healthy patients with signed consent, undergoing abdominoplasty surgery. Once the skin was collected, subcutaneous fat was removed from the skin. It was then transferred to a container containing phosphate buffered saline (PBS) and supplemented with 1% antibiotic/antimycotic. The 12mm punch biopsies were excised using a punch biopsy tool. The skin was then spread on a re-

ceptor chamber, covering the 7mm hole with the skin punch (see Franz cell set up on the left side, below in figure 3.1). The donor chamber was then placed on top and clamped. Using a pasteur pipette to dispense into the sampling port to avoid bubbles, the receptor chamber was filled with 2mL of cornification medium. Franz cells were placed in an incubator at 37°C overnight to allow cells recover/equalize before applying experimental applications.

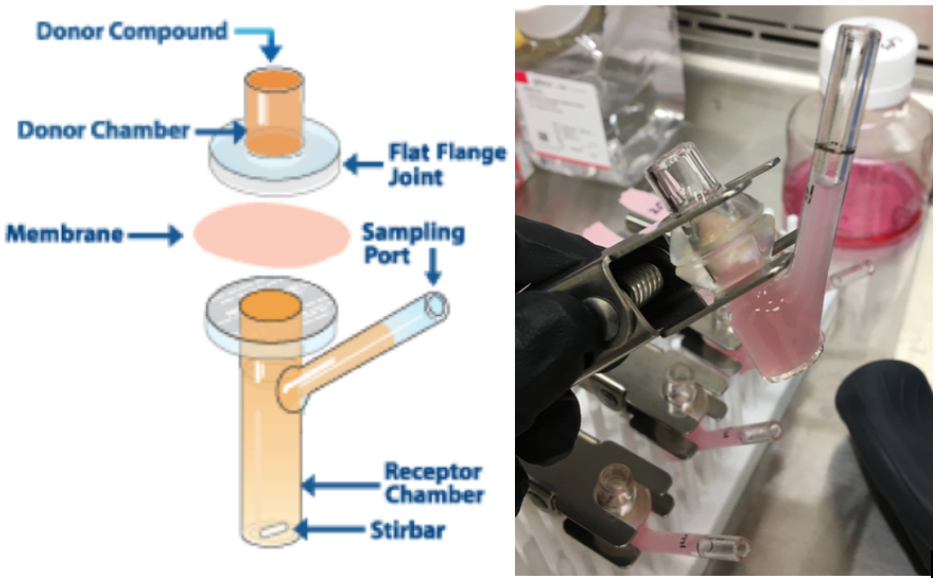


Figure 3.1: Pictured on the left: Franz cell diagram from PermeGear, Ltd. Pictured on the Right: Franz cell used in the lab with EO mixed into media for systemic experimental application.

### 3.2.3 Topical and systemic applications

After the initial 24 hour incubation period, we treated the skin tissue with topical and systemic application of the EOs. **Topical application** To test this topical application, A total of 2uL of EO was added directly onto the top of the skin of each experimental sample while they were still in the Franz cell chamber. Medium was replenished with standard cornification medium. The samples were placed back in the incubator for another 24 hours at 37°C.

**Systemic application** To test the systemic application of the EOs, 2uL of EO was added to 2mL of cornification medium and sonicated to ensure equal dispersion of

the EO molecules. The current medium was removed from the receptor chamber and replenished with the EO supplemented medium. The samples were placed back in the incubator for another 24 hours at 37°C. Displayed above on the right side of figure 3.1 is the SA of Myrrh. The oil can be seen dispersed within the medium that was added underneath the skin within the receptor chamber.

### 3.2.4 Controls

For each donor, each sample had 5 biological replicates. 5 samples served as negative controls (untreated by EOs or LPS). Another 5 replicates served as a positive control (supplemented with LPS and no EOs). FCO served as an oil-based control and was applied to the tissue systemically and topically. During the topical application and systemic application, the untreated controls and LPS-controls were only replenished with new cornification medium. Table 3.1 displays an outline of the number of biological replicates per human donor. There were two donors total (D1 and D2) to yield a total of 100 biological samples.

Table 3.1: LPS stimulation study design per donor

| **Sample numbers** | **Experimental application** | **Total volume of oil used** |
|---|---|---|
| 1-5 | Untreated | N/A |
| 6-10 | LPS | N/A |
| 11-15 | FCO (systemic) + LPS | 10uL |
| 16-20 | Blue spruce (systemic) + LPS | 10uL |
| 21-25 | Myrrh (systemic) + LPS | 10uL |
| 26-30 | Frankincense (systemic) + LPS | 10uL |
| 31-35 | FCO (topical) + LPS | 10uL |
| 36-40 | Blue spruce (topical) + LPS | 10uL |
| 41-45 | Myrrh (topical) + LPS | 10uL |
| 46-50 | Frankincense (topical) + LPS | 10uL |

### 3.2.5 LPS-stimulation

Once the incubation period was completed, the medium was removed from all the Franz cells and replaced with medium supplemented with the addition of LPS

(1ug/mL). The only samples not supplemented with LPS and replenished with cornification medium were the untreated samples. Samples were then incubated an additional 24 hours at 37°C. On day 3, the medium and skin tissue were harvested. The medium was collected and stored at -20°C for each sample. The tissue for each sample was cut into fourths and placed in RNAlater, stored at 4°C overnight and then moved to -20°C.

### 3.2.6 RNA isolation for RT-qPCR

Total RNA isolation was carried out using Precellys tissue homogenizer and RNeasy plus kit (Qiagen). The total RNA quantity and quality were determined using a Thermoscientific nanodrop. Equal amounts of total RNA was subjected to qPCR analysis using QuantStudio™12K Flex Real-Time PCR System. PCR reactions were carried out using Taqman primers designed to detect biomarkers: CXCL10, IL-8, IL-1$\beta$ and COX-2. According to manufacturer's protocol.

### 3.2.7 RNA isolation for RNA-sequencing

The RNA from 30 tissue samples displayed in the table 3.2 from D1, were extracted using the RNeasy Fibrous Tissue Mini Kit from Qiagen, per manufacturers protocol with minor adjustments. For all of the samples, the Proteinase K treatment after homogenization was increased to 60 minutes. Samples were eluted in water and the quality of the RNA preparation was assessed by spectrophotometry (A260/A280; A260/A230). RNA integrity was measured by calculation of the RNA integrity number (RIN) using an Agilent Bioanalyzer.

### 3.2.8 cDNA library construction and RNA-sequencing

Thirty sequencing libraries were created using the TruSeq RNA Library Preparation kit from Illumina following the manufacturers protocol. After the library preparation, sample library size was validated using the Agilent Bioanalyzer DNA 1000 kit and samples were quantitated using qPCR. Libraries were combined in equimolar propor-

Table 3.2: Samples used for RNA-sequencing from D1

| **Samples** | **Experimental application** |
|---|---|
| 1-5 | Untreated |
| 6-10 | LPS |
| 11-15 | FCO (systemic) + LPS |
| 26-30 | Frankincense (systemic) + LPS |
| 31-35 | FCO (topical) + LPS |
| 46-50 | Frankincense (topical) + LPS |

tions and loaded onto a flow cell lane for clustering. The flow cell was a 75 basepair (bp) paired end run and was performed on the Illumina HiSeq2500 instrument. RNA isolation, cDNA library construction and sequencing steps were performed in the Genomics laboratory of the David H. Murdock Research Institute, Kannapolis NC.

### 3.2.9 RNA-sequencing

The 4 main aspects to assemble RNA-sequencing (RNA-seq) data includes: quality control/adapter trimming, sequence alignment, transcript assembly and calculating the differentially expressed genes (DEGs). Figure 3.2 shows the overall pipeline for the RNA-seq assembly. Trimgalore! is a wrapper tool to automate CutAdapt and FastQC for quality control and filtering of RNA-sequencing reads. CutAdapt will trim adapter ends of paired reads based on the phred quality score that is specified by the user.[95][96][97] Figure 3.2 shows the parameters used for each tool used within the pipeline. FastQC will run the output files for quality checks. Next the reads were processed through HiSAT2.[98] HiSAT2 is an alignment program developed by John Hopkins University.[99] It is a fast and sensitive alignment program that utilizes a graph FM index, designed and implemented by the software developers. This is based of off the Burrows-Wheeler Transform (BWT) algorithm for graphs. HiSAT2 utilizes a new index scheme that uses local indexes and several alignment strategies to produce fast and accurate alignments of RNA/DNA sequencing reads. Their indexing method for this is referred to as "hierarchical Graph FM Index." The human index files used as

references for alignment were built using NCBI's RefSeq human reference genome and annotation file (GRCh38).[100] No major issues were seen with any of the samples, all produced alignment rates of over 96%. The files were sorted using Samtools.[101] Next the reads were counted in features using HTseq-count. This program utilizes a python library to develop scripts for processing the RNA-seq data. HTseq took the input SAM files and pre-processed it for differential expression analysis. It uses the same gene annotation file and counts how many reads map to each gene.[102] The count data generated by HTseq was analyzed using DESeq2 in R in order to identify the DEGs among the samples.[103][104][105] In a brief overview, the count data is fitted into a generalized linear model (GLM), that follows a negative binomial distribution.[104] The GLM fitted model compares treatments and control groups and returns coefficients that convey overall expression strength of the genes using the Log2FC.[104] P-values are determined with Wald's test and corrected with the Benjamini-Hochberg method.

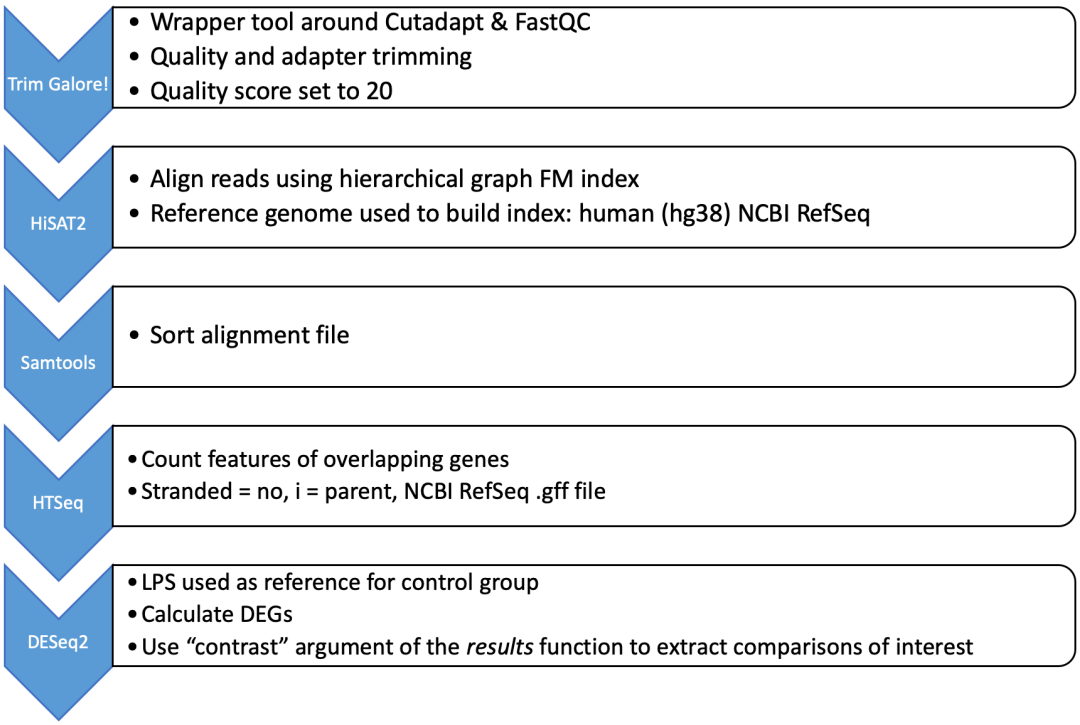


Figure 3.2: RNA-sequencing pipeline

## 3.3 Results

### 3.3.1 GC/MS results of EOs and FAME profile of FCO

Tables 3.3-3.8 below list the GC/MS composition results of each of the EOs used in this experimental assay. Below each GC/MS table, is an additional table for each EO that lists the tested physical properties.

Tables 3.9 and 3.10 display the results from the fatty acid composition of the FCO used in the experiment.

Table 3.3: GC/MS results of frankincense EO

| **Component** | **Relative area percent** | **Component** | **Relative area percent** |
|---|---|---|---|
| Alpha-pinene | 24.39% | Alpha-humulene | 0.53% |
| Alpha-thujene | 17.43% | Gamma-terpinene | 0.53% |
| Limonene | 8.23% | Germacrene D | 0.53% |
| Para-cymene | 5.70% | Trans-pinocarveol | 0.53% |
| Sabinene | 5.50% | Gamma-muurolene | 0.48% |
| Myrcene | 3.96% | Isopropyl-methyl benzene (isomer) | 0.44% |
| Terpinen-4-ol | 3.00% | T-Cadinol | 0.35% |
| Trans-beta-caryophyllene | 2.72% | Incensol acetate | 0.34% |
| Beta-pinene | 1.72% | Bornyl acetate | 0.30% |
| Alpha-phellandrene | 1.71% | Beta-bourbonene | 0.29% |
| Delta-3-carene | 1.44% | Beta-selinene | 0.29% |
| Camphene | 1.15% | Alpha-selinene | 0.21% |
| Isopropyl-methyl benzene | 0.96% | Allo-aromadendrene | 0.20% |
| Eucalyptol + beta-phellandrene | 0.92% | Alpha-muurolene | 0.19% |
| Gamma-cadinene | 0.78% | Alpha-campholene aldehyde | 0.16% |
| Alpha-copaene | 0.74% | Alpha-cubebene | 0.16% |
| Alpha-terpineol | 0.68% | Terpinolene | 0.16% |
| Cis-verbenol | 0.68% | Carveol | 0.14% |
| Caryophyllene oxide | 0.64% | Alpha-gurjunene | 0.13% |
| Beta-elemene | 0.61% | Calamenene | 0.10% |
| Delta-cadinene | 0.60% | Incensol | 0.09% |
| Thujone | 0.60% | Alpha-terpenyl acetate | 0.08% |
| **Total** | | **90.4%** | |

Table 3.4: Physical test results of frankincense EO

| | |
|---|---|
| Optical Rotation | -7.5° |
| Specific Gravity | 0.865 |
| Refractive Index | 1.470 |

Table 3.5: GC/MS results of myrrh EO

| **Component** | **Relative area percent** |
|---|---|
| Furanoeudesma-1,3-diene | 47.21% |
| Curzerene | 25.99% |
| Lindestrene | 13.40% |
| Beta-elemene | 5.3% |
| 2-Methoxy Furanogermacrene | 3.24% |
| Germacrene B | 1.06% |
| Gamma-elemene | 0.98% |
| Alpha-selinene | 0.82% |
| Delta-elemene | 0.64% |
| Germacrene D | 0.62% |
| **Total** | **99.26%** |

Table 3.6: Physical test results of myrrh EO

| | |
|---|---|
| Optical Rotation | +74.3° |
| Specific Gravity | 1.016 |
| Refractive Index | 1.527 |

Table 3.7: GC/MS results of blue spruce EO

| **Component** | **Relative area percent** |
|---|---|
| Alpha-pinene | 30.41% |
| Limonene | 22.04% |
| Beta-pinene | 10.14% |
| Camphene | 7.26% |
| Bornyl Acetate | 6.60% |
| delta-3-carene | 6.34% |
| Myrcene | 4.23% |
| Camphor | 2.52% |
| Beta-phellandrene | 1.80% |
| Terpinolene | 1.62% |
| Sabinene | 1.35% |
| Tricyclene | 0.87% |
| Exomethyl Camphenilol | 0.86% |
| Borneol | 0.82% |
| Cembrene | 0.80% |
| Cembrenol | 0.44% |
| Para-cymene | 0.38% |
| Terpinen-4-ol | 0.41% |
| Gamma-terpinene | 0.33% |
| Citronellol | 0.17% |
| Delta-cadinene | 0.17% |
| **Total** | **99.56%** |

Table 3.8: Physical test results of blue spruce EO

| | |
|---|---|
| Optical Rotation | -28.3° |
| Specific Gravity | 0.870 |
| Refractive Index | 1.473 |

Table 3.9: Fatty acid composition of FCO

| **Units** | **% By weight** | **Units** | **% By weight** |
|---|---|---|---|
| C6:0-Methyl Caproate | 0.49 | C18:3N3-Methyl Lionleate | 0.01 |
| C8:0-Methyl Caprylate | 6.93 | C18:4N3-Methyl Stearidonate | 0.08 |
| C10:0-Methyl Caprate | 8.08 | C20:0-Methyl Arachidate | 0.11 |
| C12:0-Methyl Laurate | 37.31 | C20:1-Methyl Eicosenoate | 0.23 |
| C14:0-Methyl Myristate | 15.47 | C20:2N6-Methyl 11-14 Eicosadienoate | 0.01 |
| C14:1-Methyl Myristoleate | 0.93 | C20:3N6-Methyl Homogamma Linolenate | 0.04 |
| C15:0-Methyl Pentadecanoate | 0.35 | C20:4N6-Methyl Arachidonate | 0.08 |
| C16:0-Methyl Palmitate | 10.41 | C20:3N3-Dihomo-gamma-linoenate | 0.06 |
| C18:0-Methyl Stearate | 3.18 | C22:0-Methyl Behenate | 0.05 |
| C18:1-Methyl Oleate | 12.96 | C22:1-Methyl Erucate | 0.80 |
| C18:2N6-Methyl Linoleate | 2.33 | C22:2N6-Methyl Docosadienoate | 0.01 |
| C22:5N6-Methyl Docosapentaenoate | 0.05 | | |
| **Other** | **0.00** | **Total** | **100.00** |

Table 3.10: Composition of FCO by fat type

| **Type** | **% By weight** |
|---|---|
| Saturated | 82.40 |
| Monounsaturated | 14.92 |
| Polyunsaturated | 2.69 |
| **Total** | **100.00** |

### 3.3.2 Statistical analysis of RT-qPCR data

For the LPS-stimulated ex-vivo assay, there was a total of 50 biological samples per donor. Table 3.11 highlights the number of measurements for the assay. There were two plates total per donor with 4 different experimental biomarkers measured for each biological sample (CXCL10, IL-8, IL-1$\beta$ and COX-2). Housekeeping gene (GADPH) was used as a control, on each of the two RT-qPCR plates, yielding measurements for two more markers. Each biological sample had 2 technical replicates plated. Each donor output a total of 600 measurements. Biological samples per application(5) x each measured marker: CXCL10, IL8, IL1$\beta$, COX2, GADPH, GADPH(6) x technical replicates(2) = 60.

Table 3.11: RT-qPCR measurements

| **Samples** | **Experimental application** | **Total measurements per 5 samples** |
|---|---|---|
| 1-5 | Untreated | 60 |
| 6-10 | LPS | 60 |
| 11-15 | FCO systemic + LPS | 60 |
| 16-20 | Blue spruce systemic + LPS | 60 |
| 21-25 | Myrrh systemic + LPS | 60 |
| 26-30 | Frankincense systemic + LPS | 60 |
| 31-35 | FCO topical + LPS | 60 |
| 36-40 | Blue spruce topical + LPS | 60 |
| 41-45 | Myrrh topical + LPS | 60 |
| 46-50 | Frankincense topical + LPS | 60 |

The technical replicates were averaged and then normalized according to the GADPH control. The data from each donor was analyzed separately. All of the data was log transformed and analyzed using R and the multcomp R package.[103][106] The data was subset for analysis by each donor, per genetic marker. It was checked for any outliers using studentized residual test. When using standardized residuals to calculate outliers, the regression model can get pulled towards the outlier, and thus outliers

may not be as easily detected using standardized residuals. To correct for this issue, studentized residuals detects outliers in an alternative form. The observations are deleted one at a time. During each deletion, the model is refitted on the remaining ($n$-1) observations. The model can then be compared to the observed response values based on their fitted values with the *ith* observation deleted. This will produce deleted residuals, which are standardized and become studentized residuals.[107] If the studentized residuals had a value greater than -/+ 3, they were removed from the data set. No outliers were detected in the data set. Next an analysis of the variance (ANOVA) test was applied to each data set. The ANOVA demonstrated a variance in the means between the experimental groups. These results are shown below in table 3.12.

Table 3.12: Results from ANOVA using log transformed fold change data. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001

| **Marker** | **Donor** | **F-value** | **Pr(>F)** |
|---|---|---|---|
| COX2 | D1 | 2.493 | 0.0289* |
| COX2 | D2 | 4.078 | 1.73E-03** |
| IL1$\beta$ | D1 | 28.4 | 2.46E-13*** |
| IL1$\beta$ | D2 | 11.85 | 4.43E-08*** |
| CXCL10 | D1 | 3.26 | 6.77E-03** |
| CXCL10 | D2 | 3.885 | 2.28E-03** |
| IL8 | D1 | 3.919 | 2.27E-03** |
| IL8 | D2 | 5.292 | 2.01E-04*** |

Dunnett's method was utilized to detect if there was a significant difference in gene expression for each treatment group compared against the LPS-control. Dunnett's method is a multiple comparison test that increases the power of the test by mini-

mizing the number of tests between experimental groups, reducing the false-discovery rates. Tables 3.13- 3.16 highlight the statistical findings for each application, based on donor and marker from the statistical analysis.

Table 3.13: Dunnett's method for D1 (COX2 and IL1$\beta$ based on log transformed fold change data). Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001

| **Application** | **t-value** | **Pr(>\|t\|)** | **t-value** | **Pr(>\|t\|)** |
|---|---|---|---|---|
| Marker | COX2 | COX2 | IL1$\beta$ | IL1$\beta$ |
| Myrrh systemic - LPS | -1.066 | 0.8462 | -8.469 | <0.001*** |
| Blue spruce systemic - LPS | 0.450 | 0.9988 | -0.794 | 0.961 |
| FCO systemic - LPS | -0.112 | 1 | 0.535 | 0.996 |
| Frankincense systemic - LPS | -3.210 | 0.0179* | -8.837 | <0.001*** |
| Myrrh topical - LPS | 0.222 | 1 | -2.092 | 0.216 |
| Blue spruce topical - LPS | -0.277 | 1 | 1.037 | 0.862 |
| FCO topical - LPS | -0.108 | 1 | -0.038 | 1 |
| Frankincense topical - LPS | -1.073 | 0.8425 | -1.31 | 0.686 |

Table 3.14: Dunnett's method for D1 (CXCL10 and IL8 based on log transformed fold change data). Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001

| **Application** | **t-value** | **Pr(>\|t\|)** | **t-value** | **Pr(>\|t\|)** |
|---|---|---|---|---|
| Marker | CXCL10 | CXCL10 | IL8 | IL8 |
| Myrrh systemic - LPS | -2.99 | 0.0308* | -0.446 | 0.9986 |
| Blue spruce systemic - LPS | -2.981 | 0.0314* | 0.003 | 1 |
| FCO systemic - LPS | 0.04 | 1 | 0.456 | 0.9983 |
| Frankincense systemic - LPS | -2.385 | 0.1214 | -2.998 | 0.0298* |
| Myrrh topical - LPS | -2.279 | 0.1507 | -1.102 | 0.8044 |
| Blue spruce topical - LPS | -1.137 | 0.8035 | 1.855 | 0.3109 |
| FCO topical - LPS | -1.596 | 0.4842 | 0.511 | 0.9964 |
| Frankincense topical - LPS | -3.408 | 0.0107* | -0.143 | 1 |

Table 3.15: Dunnett's Method for D2 (COX2 and IL1$\beta$ based on log transformed fold change data). Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001

| **Application** | **t-value** | **Pr(>\|t\|)** | **t-value** | **Pr(>\|t\|)** |
|---|---|---|---|---|
| Marker | COX2 | COX2 | IL1$\beta$ | IL1$\beta$ |
| Myrrh systemic - LPS | -1.314 | 0.6586 | -4.902 | 1.64E-04*** |
| Blue spruce systemic - LPS | -0.537 | 0.9951 | -0.703 | 0.979929 |
| FCO systemic - LPS | 0.911 | 0.9088 | 0.354 | 0.999794 |
| Frankincense systemic - LPS | -3.392 | 0.0111* | -5.833 | <1E-04*** |
| Myrrh topical - LPS | -2.221 | 0.1611 | 0.703 | 0.979984 |
| Blue spruce topical - LPS | 0.279 | 1 | 0.485 | 0.998079 |
| FCO topical - LPS | -0.223 | 1 | 0.026 | 1 |
| Frankincense topical - LPS | -1.569 | 0.4798 | -0.806 | 0.957727 |

Table 3.16: Dunnett's Method for D2 (CXCL10 and IL8 based on log transformed fold change data). Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001

| **Application** | **t-value** | **Pr(>\|t\|)** | **t-value** | **Pr(>\|t\|)** |
|---|---|---|---|---|
| Marker | CXCL10 | CXCL10 | IL8 | IL8 |
| Myrrh systemic - LPS | -1.133 | 0.809 | -1.764 | 0.377 |
| Blue spruce systemic - LPS | -0.78 | 0.965 | -0.111 | 1 |
| FCO systemic - LPS | 1.723 | 0.406 | -0.476 | 0.998 |
| Frankincense systemic - LPS | -1.648 | 0.453 | -4.422 | <0.001*** |
| Myrrh topical - LPS | 2.155 | 0.194 | 1.038 | 0.862 |
| Blue spruce topical - LPS | 1.881 | 0.316 | -0.344 | 1 |
| FCO topical - LPS | 0.205 | 1 | 0.745 | 0.972 |
| Frankincense topical - LPS | -0.57 | 0.995 | -0.396 | 1 |

Figures 3.3-3.6 display results for each donor according to each individual pro-inflammatory marker that was measured shown in the graph titles. Results display the log2FC values from the RT-qPCR results. The statistically significant applications are emphasized with asterisks on the axis labels described in the figure descriptions. The frankincense systemic application displayed the most significant overall impact on both donors compared to the treatments. The myrrh systemic reduced 3/8 measured markers. Myrrh and frankincense systemic both reduced the expression in IL1$\beta$, most significantly. Blue spruce and myrrh systemic reduced the expression in CXCL10 in D1. Interestingly, frankincense topical only effected one marker in D1 (CXCL10).

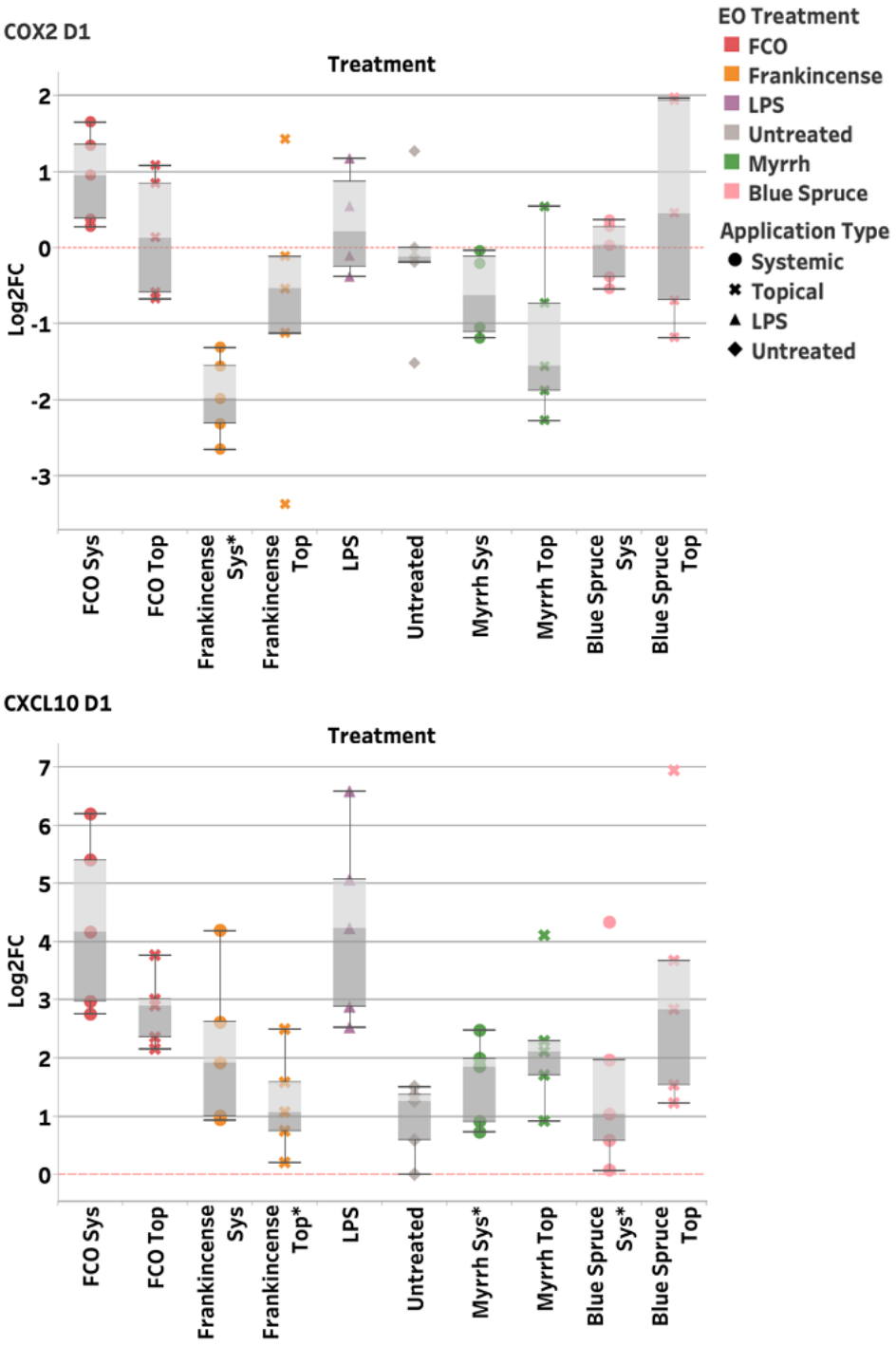


Figure 3.3: RT-qPCR results from COX2 and CXCL10 (D2). Each graph displays the Log2FC values for each application according to the labeled biomarker on the top left corner of each plot. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * $<0.05$, ** $<0.01$, *** $<0.001$

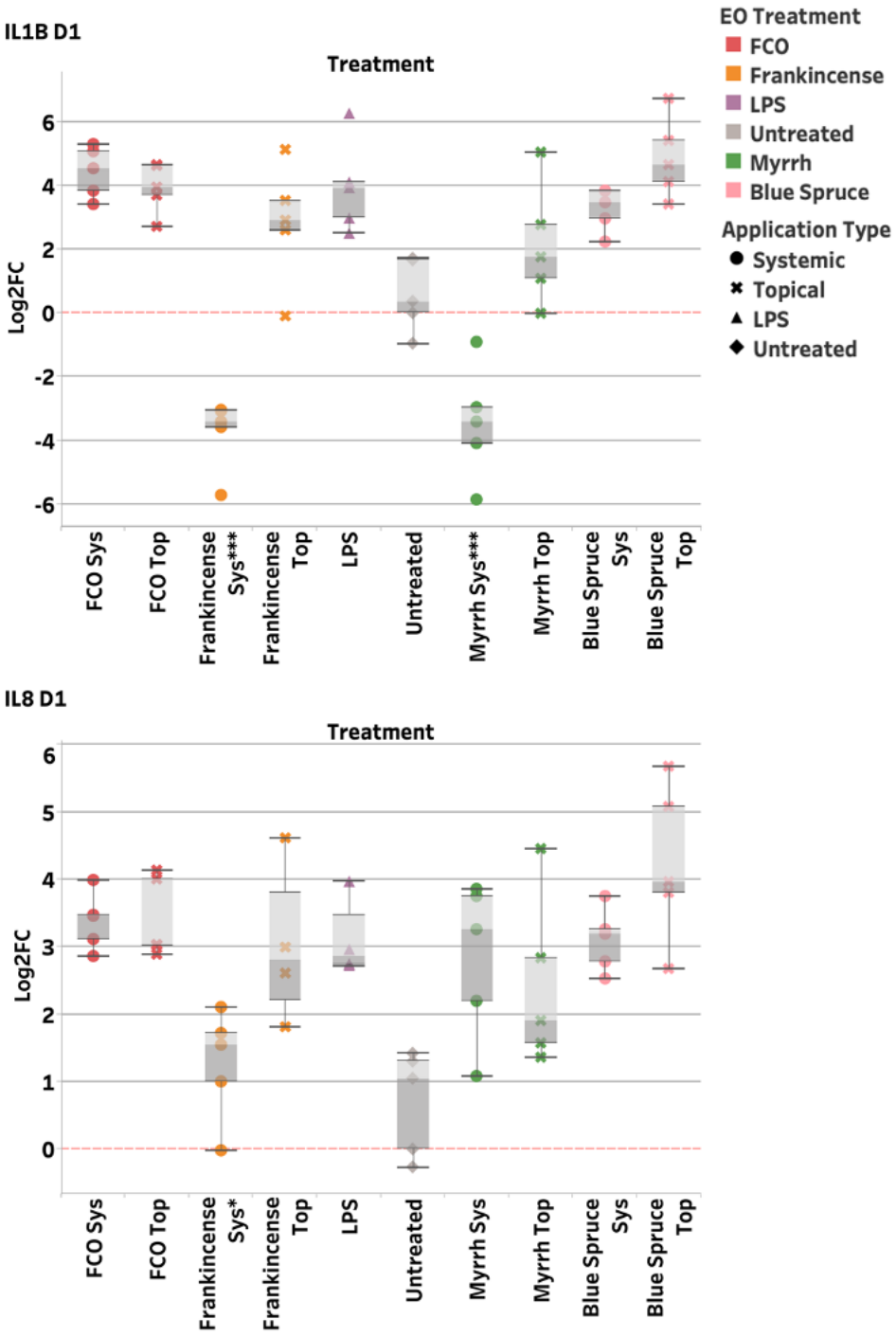


Figure 3.4: RT-qPCR results from IL1$\beta$ and IL8 (D1). Each graph displays the Log2FC values for each application according to the labeled biomarker on the top left corner of each plot. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * $<0.05$, ** $<0.01$, *** $<0.001$

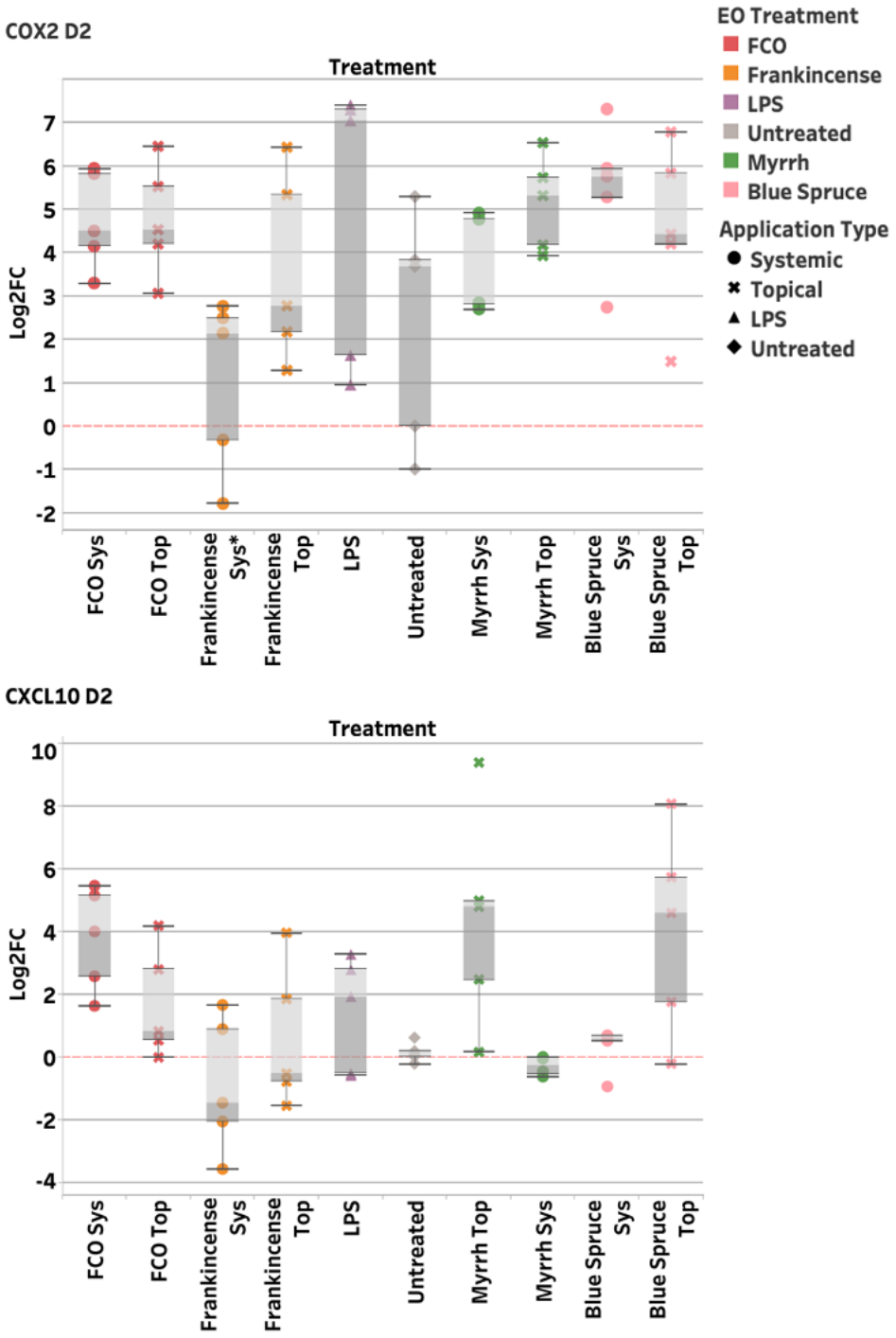


Figure 3.5: RT-qPCR results from COX2 and CXCL10 (D1). Each graph displays the Log2FC values for each application according to the labeled biomarker on the top left corner of each plot. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * $<0.05$, ** $<0.01$, *** $<0.001$

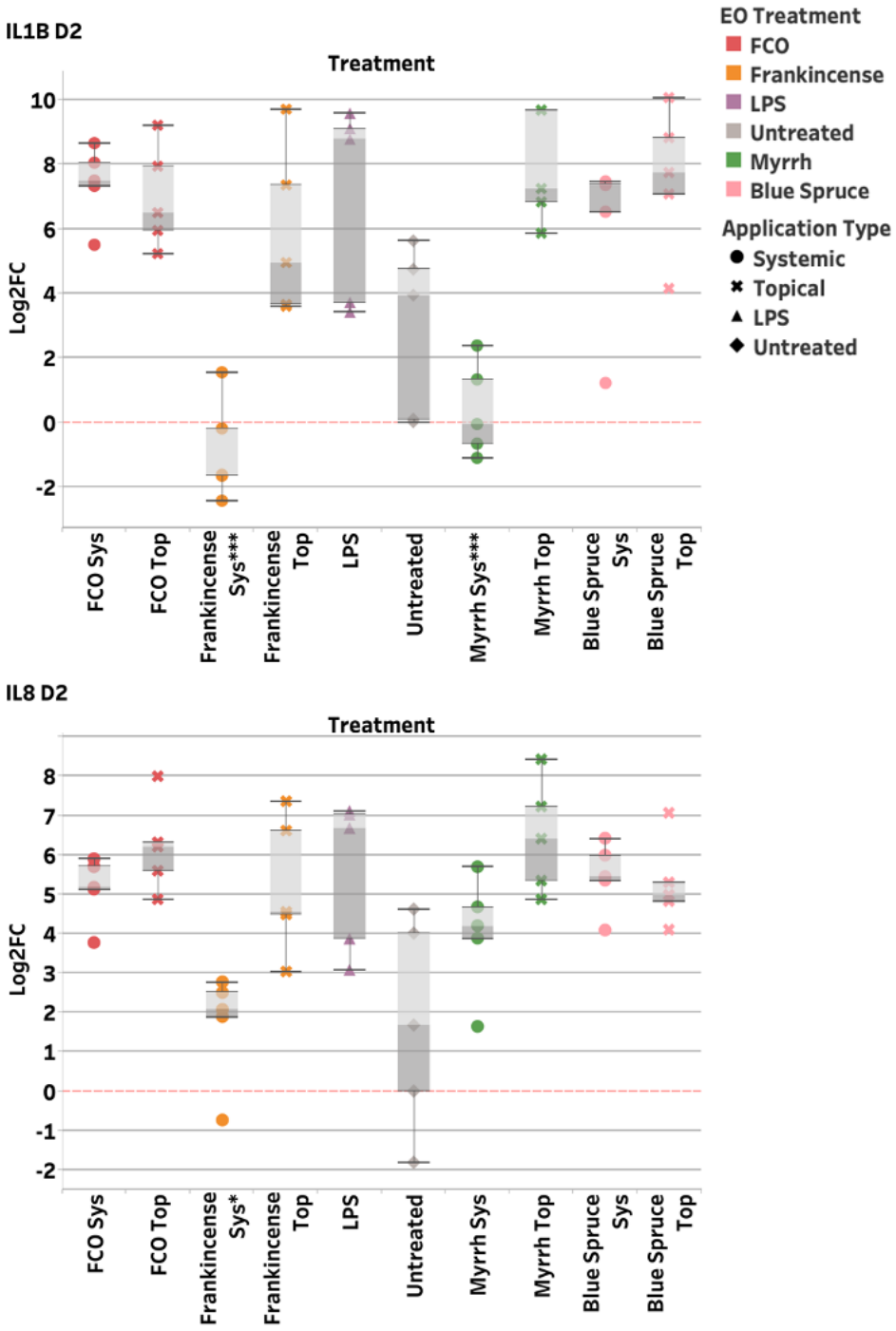


Figure 3.6: RT-qPCR results from IL1$\beta$ and IL8 (D2). Each graph displays the Log2FC values for each application according to the labeled biomarker on the top left corner of each plot. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * $<0.05$, ** $<0.01$, *** $<0.001$

### 3.3.3 DESeq2 results

After the RNA-seq of the subset of samples from D1, referenced in table 3.2, the counted genes from HTSeq were analyzed in DESeq2 to detect the number of DEGs from the experiment. The standard workflow protocol for moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 was followed to calculate the number of DEGs.[104] LPS samples were used as reference values to measure the EO treatments against the positive LPS controls. Shown in figure 3.7 are the summarized results in a mean-average (MA) plot using DESeq2 data with ggpubr R package.[108] Log2FC parameters were set at -1.5/+1.5 with a p-adjusted (p-adj) value of $<0.05$. The frankincense systemic application displayed the highest number of DEGs. The majority of DEGs in frankincense systemic were down regulated, and the majority in the frankincense topical were upregulated. The resulting number of DEGs for frankincense topical was low. The FCO systemic is not displayed in a graph as it did not produce any DEGs.

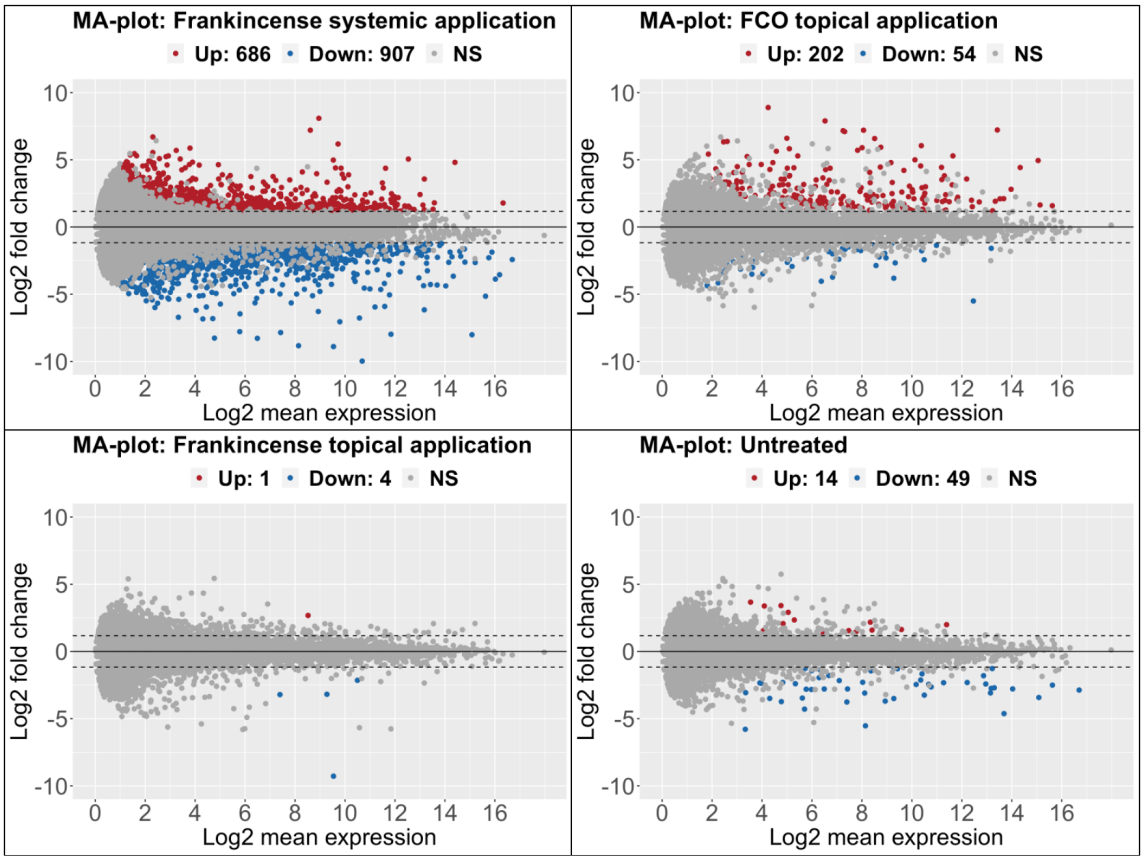


Figure 3.7: MA plot results of DEGs from DESeq2. The x-axis displays the log2 mean expression of the data while the y-axis displays the Log2FC. DEGs were filtered before plotting for a Log2FC value of -1.5/+1.5 with a p-adjusted value $<0.05$. DEGs up regulated displayed in red and DEGs down regulated displayed in blue. Genes that were not statistically significant displayed in grey.

## 3.4 Discussion

The RT-qPCR data displayed varying levels of significant results for the EO applications tested in this experiment. The frankincense systemic application had the most significant impact on both donors compared to other treatments. It reduced expression in IL1$\beta$ in both donors, COX2(D2), and IL8 in (D1). The myrrh systemic application performed well, significantly reducing three out of the eight measured markers. IL1$\beta$ in both donors was reduced, as well as COX2 and CXCL10(D1). The blue spruce systemic achieved only one significant result in CXCL10(D1).

The frankincense topical was the only topical application that had a significant

effect on CXCL10(D1). The plausible reasoning for this could be the inability for the oils to deeply penetrate the skin models. The systemic applications most likely produced higher changes in gene expression due to the specific targeting of the systemic treatments.

After the data was analyzed in DESeq2, the frankincense systemic resulted in the highest number of DEGs. Due to the hydrophobic nature of FCO, we can predict that it may have not dispersed well enough within the media to produce any change for the systemic applications. The EOs seemed to disperse quite evenly through out the media in the systemic applications. This was demonstrated in the franz cell (figure 3.1) where the presence of myrrh gave the media an evenly dispersed cloudy appearance. Overall, these results provide supporting evidence for EOs to display possible anti-inflammatory properties. Although further studies need to be done to verify this, the results of the systemic applications could provide therapeutic alternatives in the treatment of inflammation.

### 3.4.1 Connecting results back to EOKB derived hypotheses

The EOKB predictions discussed in chapter 2 were revisited after seeing the effects the oils had on the given pro-inflammatory markers. For blue spruce EO, previously published studies have noted decreased gene expression in COX2, IL1$\beta$, and IL8(CXCL8) in the presence of limonene.[109][110][111][112] Other studies observed decreased expression levels in IL1$\beta$ in association with plant compounds: gamma-terpinene, p-cymene and terpinolene.[113][114][115] Blue spruce however, only displayed decreased expression levels in CXCL10 (D1). This finding is interesting considering CXCL10 was not in the previously predicted gene set for blue spruce. Considering almost a quarter of blue spruce EO is comprised of limonene, one would have expected it would have confirmed previous findings of limonene showing decreased levels of pro-inflammatory cytokines. There is also evidence of alpha-pinene decreasing expression in pro-inflammatory markers such as TNF, COX2 and NF$\kappa$B.[116][117]

It seems logical, one would expect to see more of an anti-inflammatory response on the skin biopsy samples than what was observed by blue spruce. It is possible that the best application for blue spruce EO may not be for inflammation, which further proves and supports why this area of research need to continue to grow. Even though blue spruce contains chemical compounds that seemed to have anti-inflammatory effects, it is possible that when combined, they may not be producing the same results. Previous studies on chemical compounds related to Myrrh EO, reported compounds such as beta-elemene decreased expression in IL1$\beta$, COX2 and TNF.[118] [119][120] Not many studies have been conducted on curzerene, but one noted its effects on decreasing inflammation in rats experiencing renal failure.[121] Myrrh EO did significantly decrease inflammation in its systemic applications on CXCL10 (D1) and IL1$\beta$ in both donors. This could further support the possibility of beta-elemene providing anti-inflammatory effects within the skin tissue. The EOKB derived hypothesis for frankincense will be further discussed into detail with the RNA-seq results in chapter 4.

### 3.4.2 Limitations

#### 3.4.2.1 Minimal number of skin donors

If 1-2 more donors were added to this experiment, the samples could have been pooled together and analyzed collectively. Since there was only two donors due to cost constraints, this was not possible. There is variation that can occur between different skin donors due to age, ethnicity, gender, etc. In general, the LPS control for COX2 in D1 did not show that high levels of increased expression. Even though the frankincense systemic treatment was still deemed as significant, it raises questions whether this person could have been taking an non-steroidal anti-inflammatory drug (NSAID) which plays a role in the inhibition of COX2. This could explain why the control marker did not display as high levels of inflammation compared to other markers. Variation of skin samples will always be a limiting factor in these assays and

the only way to increase power of these tests is to add more samples (skin donors).

#### 3.4.2.2 Are the results translational?

It appears the topical applications for these EOs did not achieve optimal targeting/penetration. The reason why the systemic applications displayed decreased levels of pro-inflammatory cytokines, could be due to the fact they were specifically targeted to the exact area of inflammation. Many issues around developing effective therapeutics depends on the drugs mechanism for targeting and delivery. Unfortunately, the systemic applications are not directly translational since humans are not injecting EOs underneath the skin. This raises the question of whether EOs are reaching optimal penetration and targeting from traditional topical applications.

#### 3.4.2.3 Dosage and application

Only 2uL of oil was added directly onto the skin for each sample. Human applications typically involve adding 1mL of oil in a topical application. There is the possibility that the dosage of the topical application needs to be increased. When adding oils onto the skin topically, the user typically rubs the oil into the skin. This is almost a second form of application. The stimulation of rubbing the oil into the skin may aid in deeper penetration.

### 3.4.3 Future studies

Future studies should investigate whether applying the oils with stimulation can achieve better penetration into the skin. The RNA-seq results discussed in chapter 3, show that the FCO did penetrate the skin in the topical applications due to the number of DEGs. This raises a new hypothesis that FCO may be able to enhance absorption and penetration of the EOs. Most aromatherapists suggest to use a type of carrier oil for topical application due to possible skin sensitivities and irritation that can arise from applying oils neat. Many users still choose to apply oils neat. It is possible that the carrier oil is also doing an additional important task of not just

relieving irritation and sensitivity but enhancing penetration and absorption. Our next assay will test whether FCO can enhance our topical applications discussed into detail in Chapter 5. A better study design would have been to add samples that were treated with EOs but were not stimulated with LPS. This would allow us to observe the effects of the EO alone on untreated skin. However, adding this factor would have increased our sample size of 50 to 70 samples, per donor. This increases costs and also makes it difficult to source a patient that would provide enough skin for the high number of testing applications.

# CHAPTER 4: PREDICTING THE MOLECULAR MECHANISM OF FRANKINCENSE EO

## 4.1 Introduction

Each EO has its own array of diverse applications. Some of these applications are based on scientific literature and other applications are based on Chinese medicine, historical uses, or anecdotal evidence.[4] With over 4,000 recorded and identified chemicals from EOs and an estimated 500,000 recorded land plant species, the uses for plants and their chemical compounds could be endless.[122] Scientists estimate that more species of plants exist that are still undiscovered.[122] This means there could be many more EOs that could exist for production that have not been discovered or investigated yet. The costs of testing all of these EOs clinically and individually would take an extremely long time and will require large quantities of funding and resources. Leveraging bioinformatics based skills and methods could aid us in learning more about these EOs without having to spend the time and money on studying each of them individually. Using all of the data we have currently available and what we have generated, we hypothesize that we can predict the molecular mechanisms of frankincense to help verify existing claims as well as generate new possible medical-based applications and hypotheses.

The DEGs generated in chapter 3 were first analyzed using Qiagen's Ingenuity pathway analysis (IPA) program to discover what genes/pathways are related to the application of frankincense EO on our inflammatory-induced 3D human skin explants. This provided a data set of genes and pathways that the frankincense EO may be targeting to possibly ameliorate inflammation. There was no further analysis in this chapter for FCO systemic. It appears the FCO systemic application did not mix

well with the media, and did not result in any DEGs. This is most likely due to the fact the FA molecules of FCO are highly hydrophobic in nature and presumably did not disperse within the media and floated back up towards the sampling port (Franz cell set up: figure 3.1 for reference). The FCO topical application did result in a subset of DEGs. FCO is recommended for its use in the dilution of EOs for sensitive skin. For example, one may dilute an EO by adding 1 drop of an EO in a bottle, adding 5 drops of FCO, mixing the solution and then applying 1 drop at the diluted concentration onto the skin. FCO is commonly used for dilution purposes when applying EOs to dogs and children. There are mixed statements on whether a FA oil can aid in transdermal absorption of an EO. Some studies have shown olive oil to have the greatest enhancement for skin penetration and absorption and could be explored for future studies to improve the penetration of EOs.[123]

The chemical profile of frankincense listed in table 3.3 presented the chemical composition of the frankincense EO used in this study. Chemical profiles for frankincense can vary greatly depending on the species of *Boswellia* the EO is produced from as well as the geographical location of cultivation. This makes it uncertain whether all frankincense EOs will ever behave the same, or similarly. This issue also impacts experimental reproducibility in EO studies of all types. A principal component analysis (PCA) was conducted compare the possible variances between the different frankincense oils available online and their chemical profiles to see if any trends or patterns could be spotted between the different species based on their chemical compositions.

### 4.1.1 Pathway analysis

Using the DEG data sets generated by RNA-seq, an in-depth pathway analysis can be conducted on our samples. StringDB and the EOKB was utilized in chapter 2 in order to construct a knowledgebase driven hypothesis on frankincense oil. The results from the pathway analysis can now be used to see how the EOKB hypothesis did with real curated gene expression data. Many genes and pathways overlapped

within previously published studies and our findings in frankincense systemic. These results will be further discussed in this chapter.

#### 4.1.2 Frankincense gene expression data in comparison with NSAIDs and steroids

For mild cases of inflammation, OTC NSAIDs are typically used for pain relief. For more severe cases of inflammation, typically seen in autoimmune disorders, steroids are utilized to suppress the systemic inflammatory response of the immune system. To investigate how frankincense would compare to these drugs, several publicly available gene expression studies were used to compare the genes and pathways associated with NSAIDs and steroids and compare it back to the pathway analysis previously conducted with the frankincense RNA-seq data.

#### 4.1.3 Predicting the mechanism of frankincense EO

We hypothesize that with this data, we can combine different computational methods to integrate and combine data sources in order to verify that frankincense EO holds potential as a therapeutic in the ameliorating inflammation by pinpointing specific mechanisms and pathways involved.

### 4.2 Methods

#### 4.2.1 Ingenuity pathways analysis

The DEG data obtained from DESeq2 was further analyzed using IPA. Each sample set was uploaded and analyzed individually. The LPS samples were used as the reference for calculating the DEGs, and were therefore not analyzed in IPA. When the data was uploaded into IPA, it ran the analysis for each data set using the log2 fold change (Log2FC) values with their respective adjusted p-values of less than 0.05. Results are summarized in section 4.3.

### 4.2.2 Generally applicable gene set enrichment (GAGE) with Pathview

Pathway analysis images were generated by Pathview based on the DEGs calculated from the frankincense samples.[124] The Log2FC data calculated for the samples by DESeq2 were uploaded to Pathview.[125] Pathway selection was set to auto. The bounds for the Log2FC were set at -10/+10. This parameter displayed the DEGs optimally. Some genes had Log2FC lower than -10, but increasing the scale to -20/+20 moved too many genes into the gray scale (mid) bounds. This reduced the overall impact of the treatment on the lower DEGs, that still hold significance.

### 4.2.3 Gene set enrichment analysis (GSEA)

Performing a gene-set enrichment analysis (GSEA) on the RNA-seq data in order to compare the results of related gene sets to those found with Pathview. GSEA, developed by the Broad Institute, takes the normalized count data calculated from DESeq2 to determine which genes show statistically significant, concordant differences between different biological states.[126][127]. The data was uploaded to GSEA and compared against the molecular signatures database (MSigDB) C7 immunologic gene sets. The program output different gene set associations and heat maps based on the different testing groups used in the comparison.

### 4.2.4 EO chemical analysis

Principal component analysis (PCA) was used in R in order to visualize metabolite data of different EOs. Publicly available GC/MS data was utilized from the EOUdb to see if trends could be detected in the data between frankincense EOs of different species.[13] PCA is used to visualize multi-dimensional data. It is a dimension reduction technique that takes correlated variables, transforms them and projects the now de-correlated data into a new coordinate system. Hierarchical clustering will also be employed in R to group different EOs based on the nearest neighbor chain algorithm. This algorithm utilizes Ward's distance method. This method merges clusters based

on the similarity of distances of different clusters based on their chemical profiles using their calculated Euclidean distances.

### 4.2.5 Alternative splice (AS) site detection

One of the many benefits of using RNA-seq, allows for the user to search for AS sites within the data. One gene can produce multiple transcript variations that occur naturally. It is important to investigate whether any of these transcript variants arose from the treatment. The frankincense systemic data contained a little over 200 transcript variants or isoforms that were differentially expressed (DE). From these transcripts, a small subset of genes involved in inflammatory/immunological related pathways were chosen to investigate. Transcripts were visualized using the Integrative Genome Browser (IGB).[128] In IGB all 5 of the BAM alignment files from the frankincense systemic application were uploaded to IGB. Overall coverage and alignment of the reads for each sample was very good. Using IGB we can visualize whether these alternative transcripts are actually within the reads across our replicate samples. Upon further investigation, most of these transcripts were variants that are naturally occurring variations. No transcripts identified as something novel or unique. Some transcripts did however, display some intron retention.

### 4.2.6 Weighted gene co-expression network analysis (WGCNA)

Using the normalized count data generated from HTseq from the RNA-seq analysis, a weighted gene co-expression network analysis (WGCNA) can be performed on the data. This method uses pairwise correlations between the genes to study the relationships between co-expression modules. It then compares the modules to the different EO applications.[129] WGCNA was applied to the data using the R package "WGCNA."[130] One of the added benefits of using WGCNA allows the user to check for correlations in the data between the samples and their trait data. For this analysis, since the samples came from the same human, there was not much trait

data to be added outside of the testing conditions. The software can still be used to identify and cluster samples based on associated genes to find trends across different testing conditions.

### 4.2.7 Analysis of anti-inflammatory pharmaceuticals in comparison to frankincense EO

The ideal data set to compare the anti-inflammatory effects of frankincense would be to use another RNA-seq set that tests an NSAID on skin tissue. Ideally, this would be a topical ointment such as diclofenac gel. Upon searching for this data, the perfect ideal data set is not available. The best approach was to analyze and compare several gene expression data sets that involved the use of different NSAIDs/steroids and use them all as a comparison against the frankincense. One study measured gene expression changes over several time points on human aortic smooth muscle cells treated with 2 different NSAIDs.[131] The main issue with finding related gene expression studies suitable for comparable analysis was the fact most studies test NSAIDs and other related drugs on cancerous cell lines or diseased tissue. This is not an ideal model for comparisons considering healthy skin tissue was used in the previous experiments that we conducted with the EOs. There is a newly available RNA-seq study that stated Glucocorticoids have both direct and indirect effects on gene expression when they tested it on five non-hematopoietic (endothelial cells, fibroblasts, myoblasts, osteoblasts, and preadipocytes).[132] This makes for an excellent comparison tool due to the cells being obtained from a separate cohort of 4 unrelated healthy human donors.[132] Even though steroids tend to have a more direct immunosuppressant effect, it can still be used in comparison with Frankincense EO.

#### 4.2.8 Analysis of anti-inflammatory plant extracts in comparison to frankincense EO

Publicly available data is limited for the testing of NSAIDs on healthy human tissues. However, we can still utilize some data sets of plant extracts also studied for their possible anti-inflammatory effects. One study investigated colchicine, an alkaloid used to treat inflammation. They conducted a microarray of colchicine effects on human umbilical vein endothelial cell lines.[133] There were data sets available that studied plant extracts on different cell lines, however most were carcinoma related cell lines.

## 4.3 Results

### 4.3.1 Ingenuity pathway analysis

The results for the pathway analysis are summarized according to each sample in the sections below. The frankincense systemic had the highest number of DEGs. A cut off value was set for the Log2FC of -1.5/+1.5, for all of the samples. The frankincense topical had a very low number of resulting DEGs after adjusting the results to the p-adj values of the data set. The FCO systemic application did not produce any DEGs, and therefore does not have any summarized results below. The supplementary file contains the PDFs of the full results from IPA for all of the samples within the GitHub repository. The number of DEGs detected by IPA and DESeq2 differ slightly due to duplicate genes based on transcript variants or present isoforms. When multiple transcript IDs of the same gene variant were uploaded, IPA chose the highest valued transcript based on Log2FC/p-adj values and included it in the analysis.

The results are divided by subsection, for each EO application. IPA generated the top canonical pathways respective to each EO application. IPA bases the results on the number of gene associations and their corresponding p-adj values. The canonical

Table 4.1: Number of DEGs per sample using DESeq2. Cut off set at p-adjusted value <0.05 and Log2FC -1.5/+1.5

| **Application** | **Total DEGs** | **Up** | **Down** |
|---|---|---|---|
| Frankincense systemic | 1223 | 500 | 723 |
| Frankincense topical | 3 | 0 | 3 |
| FCO systemic | 0 | 0 | 0 |
| FCO topical | 219 | 177 | 42 |
| Untreated | 48 | 10 | 38 |

pathways and diseases/disorders results do not infer the overall effects of the treatment on the pathways.

### 4.3.2 Analysis of frankincense systemic results

The pathways listed for frankincense systemic in table 4.2 are involved in mainly immune-related responses. This is due to the high number of chemokines, cytokines and interleukins that decreased in expression from the frankincense systemic treatment.

Table 4.2: Top Canonical Pathways of frankincense systemic

| **Name** | **Overlap** |
|---|---|
| Granulocyte adhesion and diapedesis | 18.9% 34/180 |
| Hepatic fibrosis/hepatic stellate cell activation | 17.7% 33/186 |
| Agranulocyte adhesion and diapedesis | 16.6% 32/193 |
| Atherosclerosis signaling | 19.7% 25/127 |
| Hepatic fibrosis signaling pathway | 11.7% 43/368 |

IPA generates the top associated diseases and disorders for the data set based on the highest number of gene associations. Within each one of the 5 categories listed in table 4.3, IPA lists subcategories for each specific disease and the number of associated genes from the data set.

These results align with the high number of inflammatory-mediated genes reported from the input data set. Table 4.4 highlights key genes DE in the frankincense systemic that play key roles in mediating an inflammatory response. These cytokines

Table 4.3: Diseases and disorders of frankincense systemic

| Name | Number of molecules |
|---|---|
| Cancer | 1139 |
| Organismal injury and abnormalities | 1156 |
| Inflammatory response | 396 |
| Gastrointestinal disease | 1022 |
| Cardiovascular disease | 310 |

displayed varying levels of decreased Log2FC values.

Table 4.4: Key genes associated with inflammatory-mediated response by frankincense systemic

| Gene symbol | Log2FC | Gene symbol | Log2FC |
|---|---|---|---|
| CCL2 | -3.330 | CCL13 | -2.491 |
| CCL8 | -3.322 | CCL19 | -2.801 |
| CCL7 | -8.282 | CCL22 | -5.645 |
| IL6 | -4.108 | IL15 | -4.812 |
| IL10 | -5.105 | IL1$\alpha$ | -3.432 |
| IL2RA | -5.993 | IL11 | -3.823 |
| CXCL2 | -3.256 | CXCL6 | -9.978 |
| CXCL3 | -4.367 | CXCL8/IL8 | -4.634 |
| CXCL5 | -8.015 | CXCL10 | -4.984 |
| IL1R1 | -2.566 | CXCL12 | -2.742 |
| MMP9 | -5.908 | MMP7 | -2.034 |

Cytokines are soluble extracellular proteins or glycoproteins that regulate and mobilize cells that engage in adaptive and innate inflammatory host defenses, cell death/growth, differentiation and processes that restore homeostasis.[84] Interleukins (IL) and chemokines (CXC, CC) all belong to the cytokine family. Chemokines are a type of cytokine that stimulate the migration and activation of mainly phagocytic cells and lymphocytes.[84] They play a fundamental role in inflammatory responses and the pathogenesis of autoimmune inflammatory diseases.[134] The CXC subfamily activate neutrophils and members of the CC subfamily attract and activate monocytes and lymphocytes.[135] Transforming growth factor beta 1 (TGF-$\beta$1) did exhibit

a slight decrease (-1.295). TGF$\beta$ is known to suppress cytokine signaling by inhibiting the activity of Th1 and macrophages.[136] Pro-inflammatory cytokines: IL8(CXCL8), CXCL10, IL6, CCL2, and IL1$\alpha$ were decreased from the LPS controls. IL1$\beta$ was significantly lowered in the RT-qPCR assays but missed the cut off value for the p-adj parameters set for the RNA-seq data. Secretion of IL1$\beta$ activates TLR2 and TLR4.[137] TLR3 and TLR4 presented Log2FC values of -2.323 and -3.568 respectively. Of these cytokines, studies investigating the functioning of anti-inflammatory cytokines have found their role to vary based on the biological condition they reside in.[138] IL10, an anti-inflammatory cytokine can produce a pro-inflammatory functions when residing in an inflammatory environment or condition, where endotoxins, such as LPS, are present.[138][139][140] Further analysis of a subset of targeted genes: IL1$\alpha$, IL1R1, IL6, CXCL8/IL8, IL15, CXCL10, MMP9, IL2RA, CCL7, revealed many current and investigational drugs that are being or have been tested to target these genes. These drugs range in use from treatments for autoimmune diseases, dermatological diseases, chronic inflammation and tumor suppression.[22] Due to the high number of cytokines that displayed a decrease in expression, IPA predicted decreased activation states for the diseases and functions listed in table 4.5 below.

Many different metabolic pathways are linked to cancer due to the numerous types that exist. Most inflammatory markers play roles in cancer-associated pathways, resulting in the high number of associated genes in cancer. Further upstream analysis for the frankincense systemic revealed inhibited activation states for the upstream regulators listed in table 4.6. Of these upstream regulators, it was intriguing to see that LPS was listed as a predicted inhibited chemical drug due to the high number of associated inflammatory genes and over 39,000 literature findings. The presence of LPS increases expression in TNF, IL6, IL1$\beta$, NOS2, PTGS2, NF$\kappa$B (complex), IL10, IL12B, CXCL8(IL8), nitric oxide, TLR4, CCl2, IL12 (complex), CXCL10 and CXCL3. All of which were significantly decreased in the frankincense systemic samples. Activa-

Table 4.5: Subset of predicted activation states for the listed diseases above based on p-values, number of gene associations and z-scores.

| **Categories** | **Diseases/ functions annotation** | **p-value** | **Predict. activation state** | **Z-score** | **# of genes** |
|---|---|---|---|---|---|
| Organismal injury and abnormalities | Cancer | 6.25E-56 | Decreased | -2.52 | 1138 |
| Immune cell trafficking | Leukocyte migration | 1.75E-26 | Decreased | -3.722 | 182 |
| Immune cell trafficking | Cell movement of leukocytes | 6.25E-24 | Decreased | -3.763 | 155 |
| Cell-to-cell signaling and interaction | Activation of cells | 3.12E-23 | Decreased | -3.968 | 176 |
| Inflammatory response | Inflammatory response | 1.20E-21 | Decreased | -4.438 | 150 |

tion of tumor necrosis factor (TNF) and/or nuclear factor-kappa-B (NF$\kappa$B) regulate signaling genes that mediate and signal inflammation in innate and adaptive immune responses. IPA predicted the inhibition of the NF$\kappa$B complex, IL1$\beta$ and LPS based on the upstream analysis. Myrrh and frankincense systemic treatments both had significant effects on reducing IL1$\beta$ signaling in the initial RT-qPCR assay. Further investigation of cardiovascular disease showed the decreased activity of occlusion of blood vessel and artery based on the downstream analysis of the 90 associated genes from the data set. Membrane metallo-endopeptidase (MME) presented a Log2FC value of -3.549. Increased MME expression within the neutrophils is detected in patients with early phase acute myocardial infarction (AMI). These increased expression levels of MME may play a role in the pathophysiology of ischaemia/reperfusion myocardial injuries by increasing degradation in neutrophils.[141]

The IPA software curated the results into different networks based on associations with other related networks and genes. Figure 4.1 illustrates the number of genes associated with an inflammatory response and the cellular locations of each of the

Table 4.6: Subset of upstream analysis regulators and their predicted activation states based on literature findings, pvalues, gene associations and z-scores.

| **Upstream regulator** | **Molecule type** | **Predict. activation state** | **Z-score** | **P-val of overlap** | **Mechanistic network** |
|---|---|---|---|---|---|
| TNF | Cytokine | Inhibited | -6.111 | 2.94E-37 | 431 (11) |
| LPS | Chemical drug | Inhibited | -6.758 | 1.29E-25 | 410 (13) |
| IL1$\beta$ | Cytokine | Inhibited | -6.376 | 7.81E-23 | 370 (12) |
| P38 MAPK | Group | Inhibited | -4.583 | 3.04E-19 | 349 (13) |
| NF$\kappa$B (complex) | Complex | Inhibited | -6.684 | 2.43E-13 | 430 (13) |

targeted genes. Green shapes indicate decrease measurements of gene expression. The red shapes indicated increased measurements of gene expression. Refer to the IPA legend for more detailed descriptions. Numerous cytokines existing within the extracellular space were decreased from the LPS controls. Specific Log2FC values were excluded from the image due to complexity, but can be found in the supplementary files under chapter 4 results for IPA.

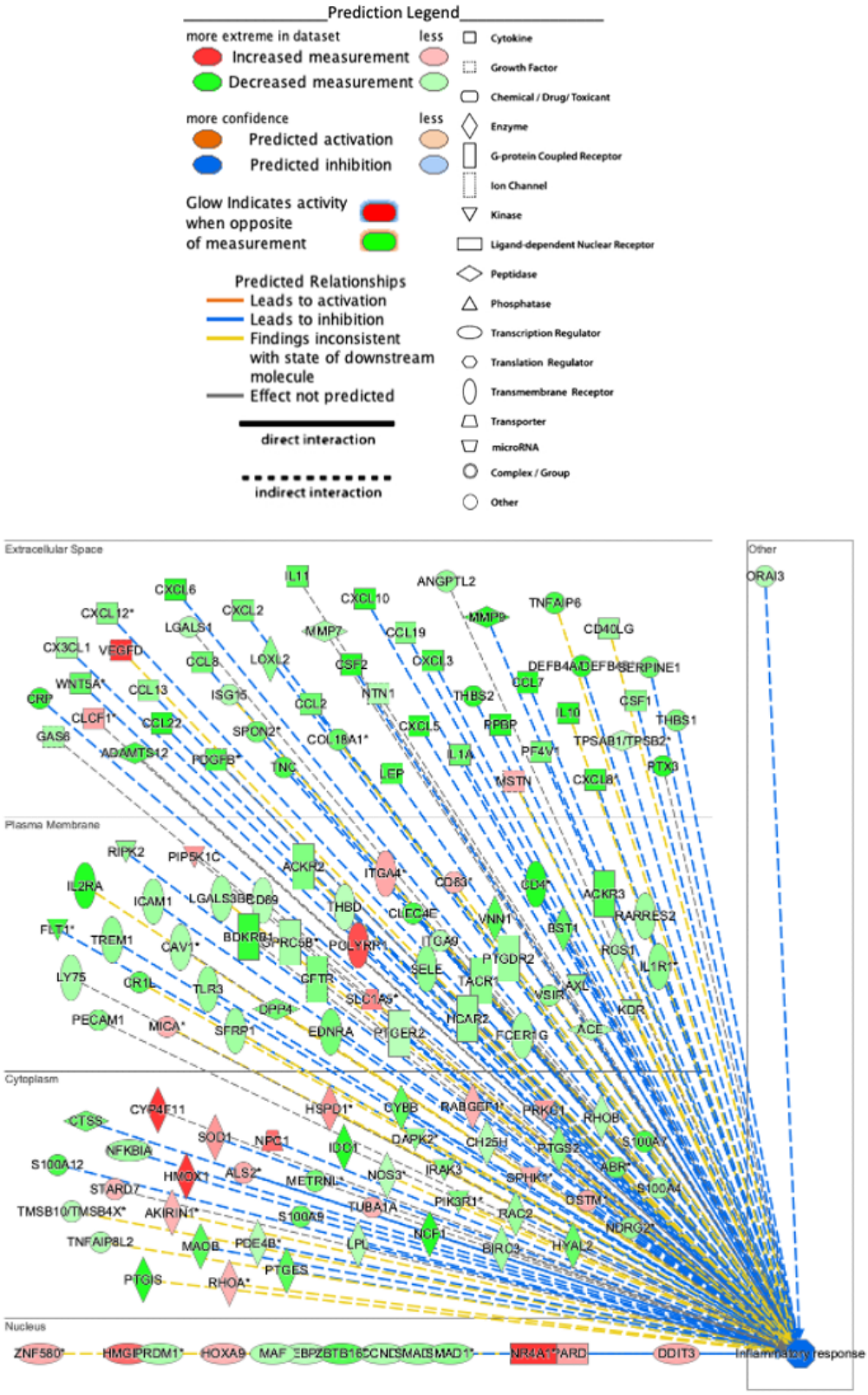


Figure 4.1: DEGs from frankincense systemic associated with inflammatory-mediated response

#### 4.3.2.1 Generally applicable gene-set enrichment of frankincense systemic in pathview

Pathview applied the generally applicable gene-set enrichment (GAGE) method to the frankincense systemic data set. The software auto-detects KEGG pathways from genes that are co-regulated towards a single direction. The canonical pathways generated by IPA allow for changes in both directions of the genes. Pathview generated a total of 6 associated Kegg pathways for the frankincense systemic data listed in the table below.

Table 4.7: Pathview summary results for frankincense systemic

| **KEGG pathways** | |
|---|---|
| **Pathway** | **Description** |
| hsa04060 | Cytokine-cytokine receptor interaction |
| hsa04062 | Chemokine signaling pathway |
| hsa04510 | Focal adhesion |
| hsa04630 | Jak-stat signaling |
| hsa04640 | Hematopoietic cell lineage |
| hsa04668 | TNF-signaling pathway |

The majority of the DEGs from the frankincense systemic were involved with the cytokine-cytokine receptor interaction pathway. Partial results of the cytokine-cytokine pathway are pictured in the figure below. Due to the complexity of the pathway, the image displays only part of the pathway containing the majority of the genes residing in the chemokine and cytokine families. (Full images can be found in GitHub repository) Diseases associated with this pathway include: Type 1 diabetes, Chron's disease, inflammatory bowel disease, ulcerative colitis, juvenile idiopathic arthritis, chronic mucocutaneous candidiasis and many others. There are 414 listed drugs for this pathway. The genes are colored in the figure below based on a log2FC scale of -10/+10.

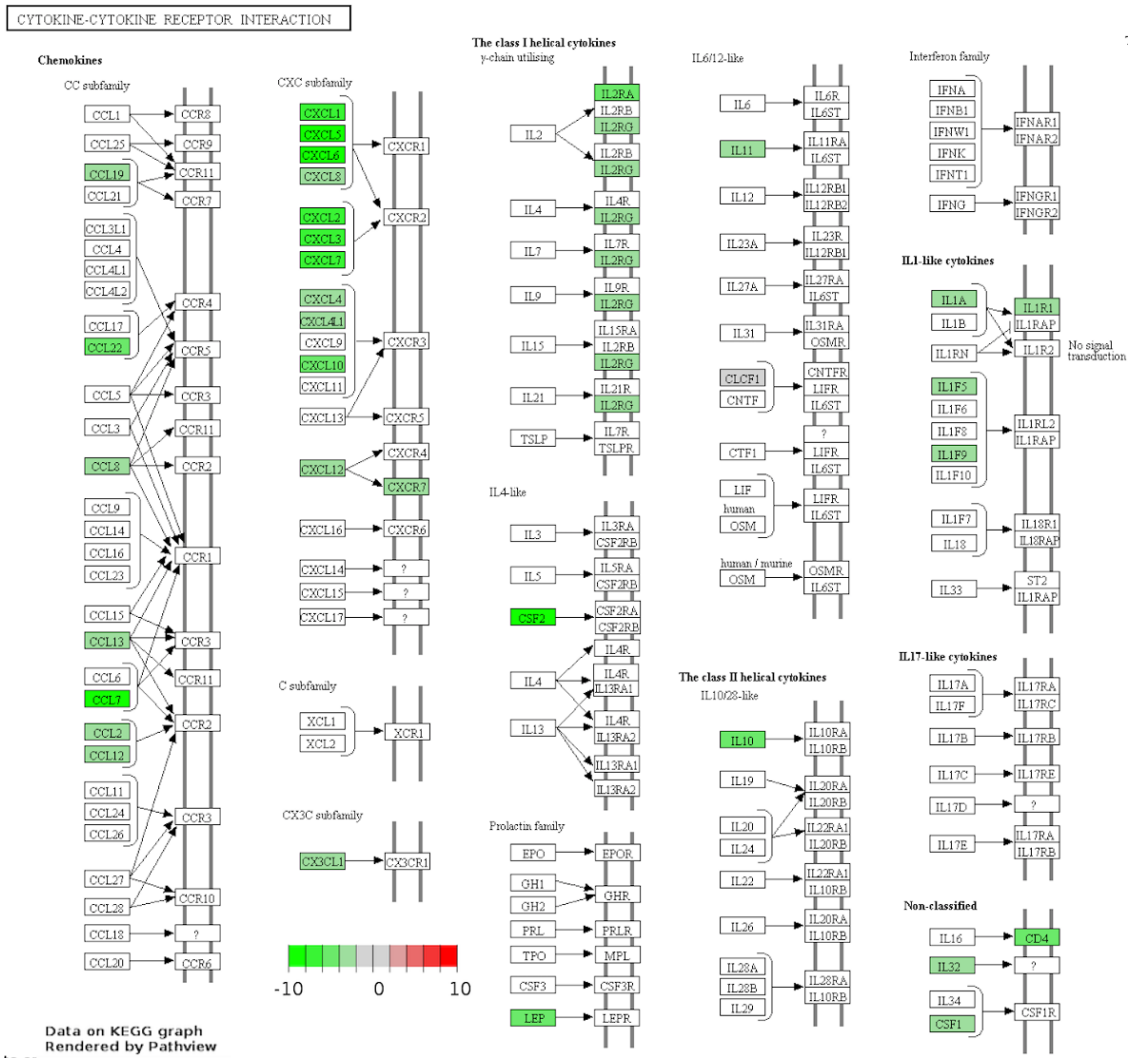


Figure 4.2: Partial image generated by Kegg and Pathview for the frankincense systemic application on the cytokine-cytokine receptor interaction pathway(hsa04060) Log2FC scale displayed in bottom left area of the figure

An intriguing finding was the decreased expression values for IL2 receptors IL2RA and IL2RG under the class I helical cytokines which activates JAK3 and activates tyrosine phosphorylation. JAK3 is associated with a variety of immunodeficient-related diseases. The highest expression change was seen in the CXC subfamily, also seen in table 4.4. CXC chemokines bind to their receptors and can induce a wide array of responses including: allergic responses, inflammation, infections, autoimmune diseases and tumor growth.

#### 4.3.2.2 Gene set enrichment analysis (GSEA) of frankincense systemic

The overarching goal of GSEA is to determine whether the members of a given gene set, tend to occur towards the top or bottom of the list in correlation with the distinctive treatment types.[126] The frankincense systemic was compared against the LPS controls with the chosen MSigDB C7 immunologic gene data sets. This gene set contains a variety of expression data from different cell types and states within the human and mouse immune system. Since it is known that most of the associated pathways from frankincense systemic were from the immune system, and since samples were target with LPS, it was suitable to use this MSigDB gene set. Results displayed that the frankincense systemic displayed mostly negative enrichment scores (ES) in association with the gene sets. Figure 4.3 below displays 2 graphs. The first graph on the left summarizes the results of the enrichment score given by GSEA. ES is scored based on a running sum statistic that depends on the magnitude of correlations between the given genes and the phenotypes of the data. It reflects the degree to which a gene set is over-represented at the extremes of the entire ranked list.[126] Next, its level of significance is calculated by creating a correlation structure based on the data by using permutations of the phenotype labels. Lastly, the score is adjusted using for multiple hypothesis testing using FDR.[126]

Since GSEA ranks based on genes upregulated within the gene sets, the LPS control samples received a higher ES due to the highly expressed immunologic-related genes. Frankincense received a negative correlation score with genes seen as upregulated in LPS controls. This is due to the majority of genes that were down-regulated due to the frankincense EO treatment application. The graph on the left in figure 4.3 displays the distribution of the ES against the number of enriched gene sets. The correlation profile on the right side of the figure, demonstrates a positive correlation of upregulated genes in the positive LPS controls, as well as a high enrichment score. The frankincense systemic displays a low enrichment score and is negatively correlated

with the ranked gene list due to the number of down regulated genes measured in those samples.

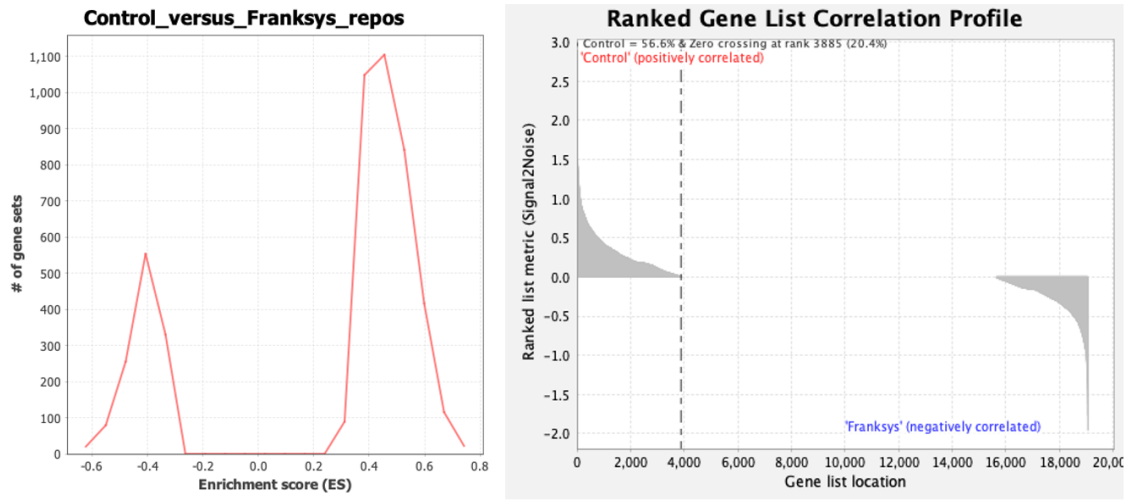


Figure 4.3: The graph on the left displays the distribution of the enrichment score (ES) of the LPS controls versus the frankincense systemic. The figure on the right shows the correlation between the ranked gene sets based on the sample conditions.

Two heat maps are shown in figure 4.4. These heat maps rank the top 50 genes correlated with each condition based on expression values. Dark blue indicates lowest expression values and red indicates highest expression values. The first heat map on the left are the results from running GSEA on the data with the MSigDB C7 gene sets. When GSEA was run a second time, MSigDB was changed out and replaced with a drug signature database gene set (DSigDB).[142] This gene set contains target genes of FDA-approved drugs. GSEA applies scores for each drug based on the correlation of the disease signature genes from the expression changes in the samples. This can be used to compare experimental/investigational drugs to already existing approved drugs.[143] The results from this run, presented extremely similar results even when using two different gene signature data sets. Frankincense systemic presented ES negatively correlated with upregulated genes in the LPS controls using either DSigDB/MSigDB. It is clear here based on the figure, that there is a

high number of completely separate genes correlated with each of the two condition types. Cytokines CXCL6, CCL7, CCL22, CXCL5, CXCL10, CXCL8/IL8, displaying opposite gene expression values versus the LPS controls in both GSEA runs.

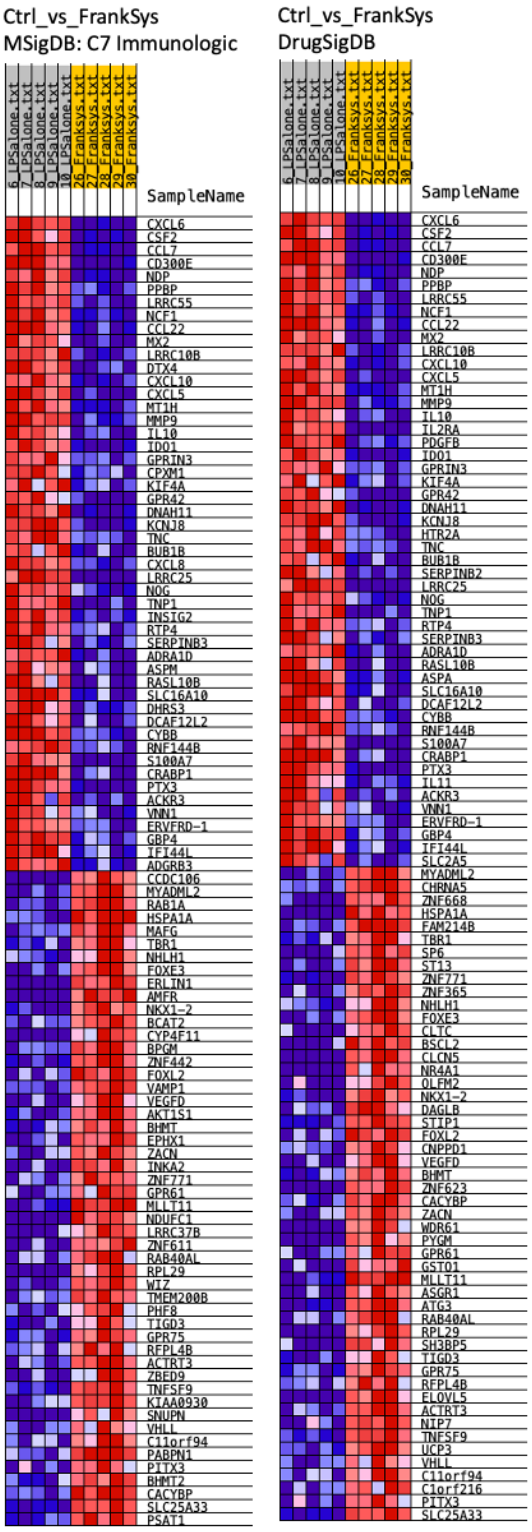


Figure 4.4: Heat map of top 50 features for each condition showing the correlation between the ranked genes and the treatment type, using two different gene signature databases (MSigDB/DSigDB). Gene expression is represented by a range of colors (red, pink, light blue, dark blue) shows the range of expression values (high, moderate, low, lowest)

### 4.3.3 Analysis of frankincense topical results

The frankincense topical application resulted in a very low number of DEGs with the tight cut off parameters set for IPA (p-adj <0.05). These genes and their respective Log2FC values were SPRR2A(-9.282), CXCL10(-3.211), and IFIT2(-2.148). The results from the RNA-seq and the RT-qPCR for CXCL10 from frankincense topical donor 1 were in agreement. CXCL10 binds to CXCR3 in order to regulate immune cell activation, signaling, migration and differentiation.[144] It is highly expressed in a wide variety of human diseases, including cancer and autoimmune diseases. Its increased expression proliferates inflammation and can cause tissue damage.[145] Decreasing the expression of CXCL10 can interfere with its binding to the CXCR3 receptor and could slow the progression of central nervous system (CNS) associated diseases such as epilepsy, multiple sclerosis (MS), and Alzheimer's disease.[146] Eldelumab is a human antibody drug currently in the investigational phase that binds to CXCL10 in order to block calcium flux, cell migration and reduce the production of pro-inflammatory cytokines.[22] It is being studied on multiple inflammatory diseases: rheumatoid arthritis, inflammatory bowel disease (IBD) and MS. It is possible that the topical application of frankincense could provide a similar mechanism of action, however more studies need to be continued for its use in therapeutics. A full pathway analysis could not be conducted on the frankincense topical due to the low quantity of DEGs.

#### 4.3.3.1 GSEA of frankincense topical

Although the resulting DEGs from frankincense topical were low, the normalized counts generated by DESeq2 were still analyzed in GSEA. The results from the GSEA for frankincense topical, mimicked similar results to those seen in the frankincense systemic. The frankincense topical displayed mainly negatively correlated gene lists and enrichment scores compared to the LPS controls seen in figure 4.5. Figure 4.5 also

displays the same heat map explained in section 4.3.2.2. A trend can still be observed in the gene sets expressing opposing correlations between the LPS and frankincense topical. However, the results here are not as distinctly separated as seen previously in the frankincense systemic heat map. As noted earlier, CXCL10 is still heavily down-regulated in comparison to the LPS samples as presented in the heat map. The majority of these genes, did not make the cut off parameters for the analysis that was completed in IPA.

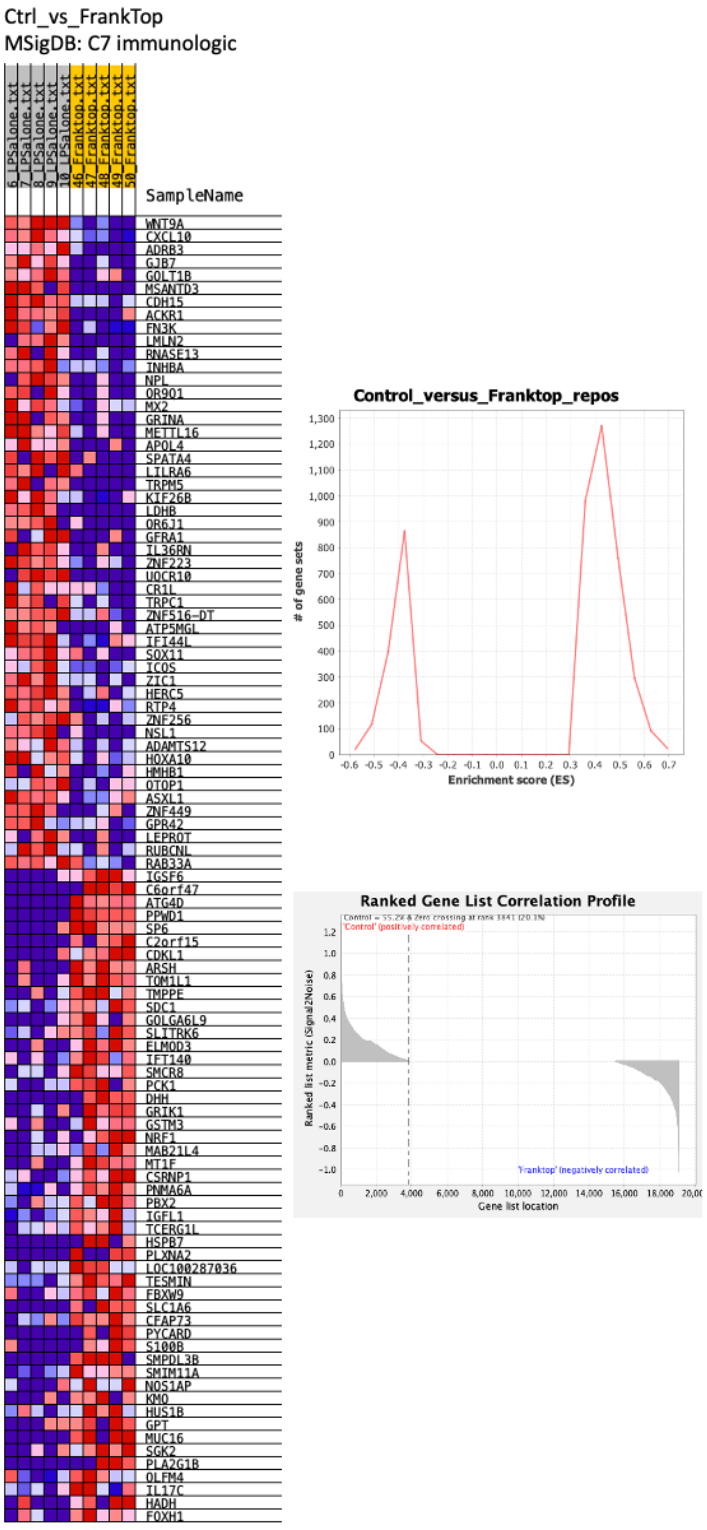


Figure 4.5: On the left: Heat map of top 50 features for each condition showing the correlation between the ranked genes and the treatment type between LPS controls and frankincense topical (MSigDB: C7). Gene expression is represented by a range of colors (red, pink, light blue, dark blue) shows the range of expression values (high, moderate, low, lowest) On the right: ES scores based on the number of gene lists, and correlation profile of the gene lists in comparison with treatment applications

### 4.3.4 Analysis of FCO topical results

The DEGs in the FCO topical samples were not mainly associated with signaling or mediating inflammation. This implies that the LPS was able to successfully induce inflammation within the tissue. Therefore, most inflammatory signaling molecules did not change significantly from the controls. Genes from the FCO topical were involved in the binding, transporting, and metabolizing of fatty acids. These fatty-acid related genes and proteins prompted a significant increase in expression versus the LPS controls, therefore supporting effective absorption of the FCO in the skin biopsy models. Genes from the FCO topical had increased expression levels in the listed diseases in table 4.9, but were not main targets in mediating or signaling inflammation.

Table 4.8: Top canonical pathways of FCO topical

| **Name** | **Overlap** |
|---|---|
| Role in IL-17A in Psoriasis | 30.8% 4/13 |
| Intristic prothrombin activation pathway | 9.8% 4/41 |
| SPINK1 pancreatic cancer pathway | 7.3% 4/55 |
| Fatty acid activation | 15.4% 2/13 |
| Cholesterol biosynthesis I | 15.4% 2/13 |
| SPINK1 pancreatic cancer pathway | 6.7% 4/60 |

Table 4.9: Diseases and disorders of FCO topical

| **Name** | **Number of molecules** |
|---|---|
| Dermatological diseases and conditions | 148 |
| Organismal injury and abnormalities | 208 |
| Cancer | 206 |
| Skeletal and muscular disorders | 40 |
| Immunological disease | 42 |

#### 4.3.4.1 GSEA of FCO topical

A GSEA was also conducted on the FCO topical samples. The MSigDB used for the FCO samples was the Hallmark (H) gene sets. This signature database was chosen

Table 4.10: Molecular and cellular functions of FCO topical

| Name | Number of molecules |
|---|---|
| Cell morphology | 26 |
| Lipid metabolism | 26 |
| Small molecule biochemistry | 30 |
| Cellular development | 16 |
| Post-translational modification | 12 |

to look at a more overall effect of the FCO topical, since many DEGs of the FCO did not play direct roles in immunologic-related pathways. The results from the GSEA produced very different gene sets due to the fact FCO topical was associated with different pathways than the LPS controls. Cytokines CCL21 and IL17A displayed in the heat map in figure 4.6 were observed to have decreased gene expression levels. The fact that many cytokines are not seen on the heat map indicate that they were the same in both the LPS controls and the FCO topical treatments, since they were not DE from one another.

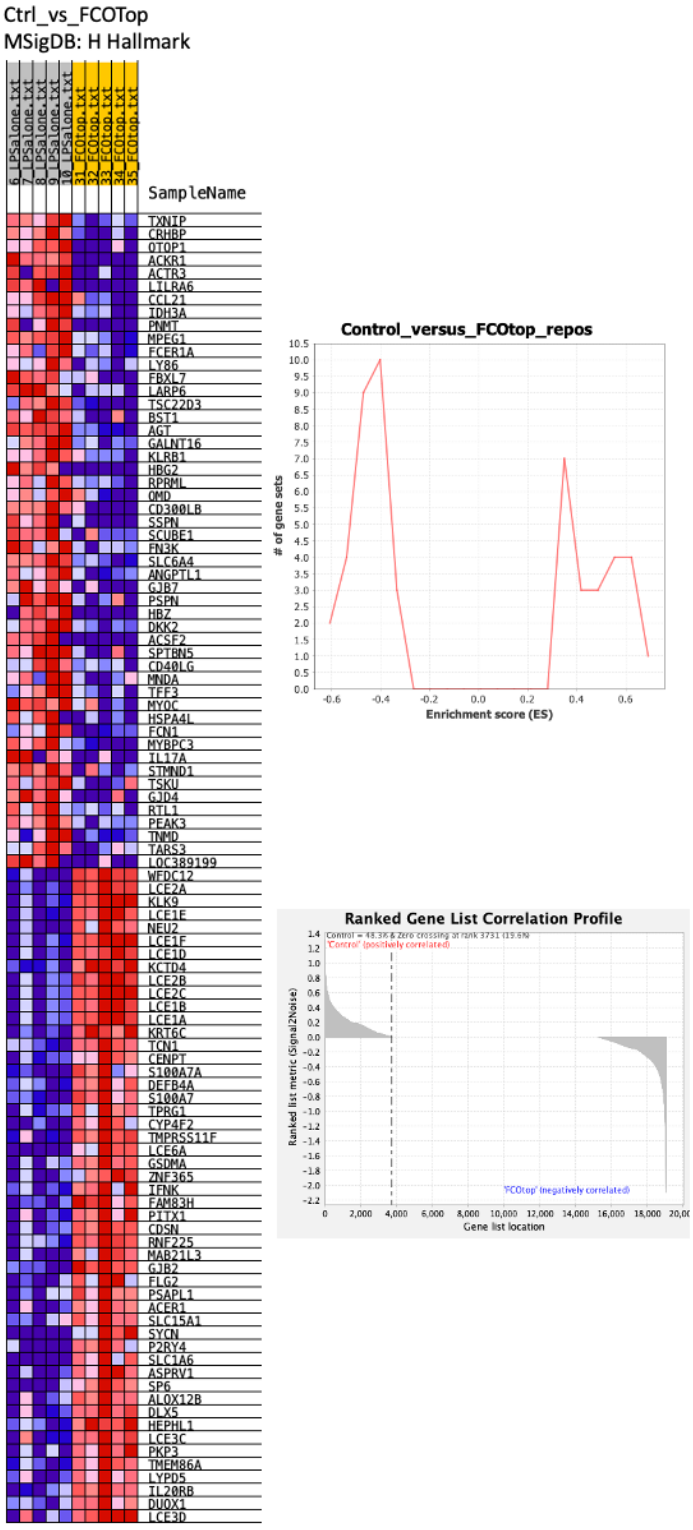


Figure 4.6: On the left: Heat map of top 50 features for each condition showing the correlation between the ranked genes and the treatment type between LPS controls and FCO topical (MSigDB: H). Gene expression is represented by a range of colors (red, pink, light blue, dark blue) shows the range of expression values (high, moderate, low, lowest) On the right: ES scores based on the number of gene lists, and correlation profile of the gene lists in comparison with treatment applications

### 4.3.5 Analysis of the untreated results

The untreated samples were not stimulated with the addition of LPS or EOs. The untreated samples served as a negative control and provided a solid baseline in regards to inflammation. The expression data for the untreated samples were analyzed in IPA. The small subset of resulting genes from the untreated samples were the expected inflammatory-mediated genes that significantly increased within the positive controls after the addition of LPS. Most of these genes are associated with similar inflammatory-related pathways as seen in the frankincense systemic samples.

Table 4.11: Top canonical pathways of the untreated applications

| **Name** | **Overlap** |
|---|---|
| Granulocyte adhesion and diapedesis | 5.6% 10/180 |
| Agranulocyte adhesion and diapedesis | 5.2% 10/193 |
| Role of IL-17F in allergic inflammatory airway diseases | 14.3% 6/42 |
| TREM1 signaling | 8.0% 6/75 |
| Role of IL-17A in arthritis | 9.3% 5/54 |

The number of resulting molecules seen in these diseases are fewer than those DE in the frankincense systemic.

Table 4.12: Diseases and disorders of untreated application

| **Name** | **Number of molecules** |
|---|---|
| Inflammatory response | 33 |
| Connective tissue disorders | 24 |
| Inflammatory disease | 26 |
| Organismal injury and abnormalities | 44 |
| Skeletal and muscular disorders | 24 |

Table 4.13 highlights key genes from the untreated samples that are the cause of the resulting canonical pathways and related diseases seen in the table above. These cytokines in the untreated samples were identified as DEGs from the LPS controls. These are ideal results for the untreated samples, as we can confirm the addition of

LPS in the media did produce a substantial inflammatory response within the skin tissues. It is also important to note that all the skin biopsy samples were treated with 1% antibiotic/antimycotic solution upon arrival before they began experimental testing to reduce risk of bacterial/fungal contamination in the skin culture assay. Since the untreated samples were never treated with anything else but fresh media, this could be why the baseline began at a relatively lower level of inflammation.

Table 4.13: Key genes with Log2FC from inflammatory response figure produced by untreated

| Gene symbol | Log2FC | Gene symbol | Log2FC |
|---|---|---|---|
| CCL2 | -2.331 | CCL7 | -5.533 |
| CXCL2 | -2.703 | CXCL3 | -3.087 |
| CXCL5 | -3.427 | CXCL8(IL8) | -2.876 |
| CXCL10 | -3.762 | CXCL13 | -3.087 |

#### 4.3.5.1 GSEA of the Untreated Samples

The untreated samples were added in GSEA with MSigDB Hallmark (H) gene set. The GSEA with the untreated samples produced results similar to those derived with IPA. IL10, an anti-inflammatory cytokine can be seen with lower expression levels that those in the control. It is interesting to note most of the genes seen as DE in the untreated samples were also DE in the frankincense systemic samples as well as additional genes. The addition of frankincense EO at the systemic level, may have been able to retain these low expression levels as well as induce others, before the addition of LPS.

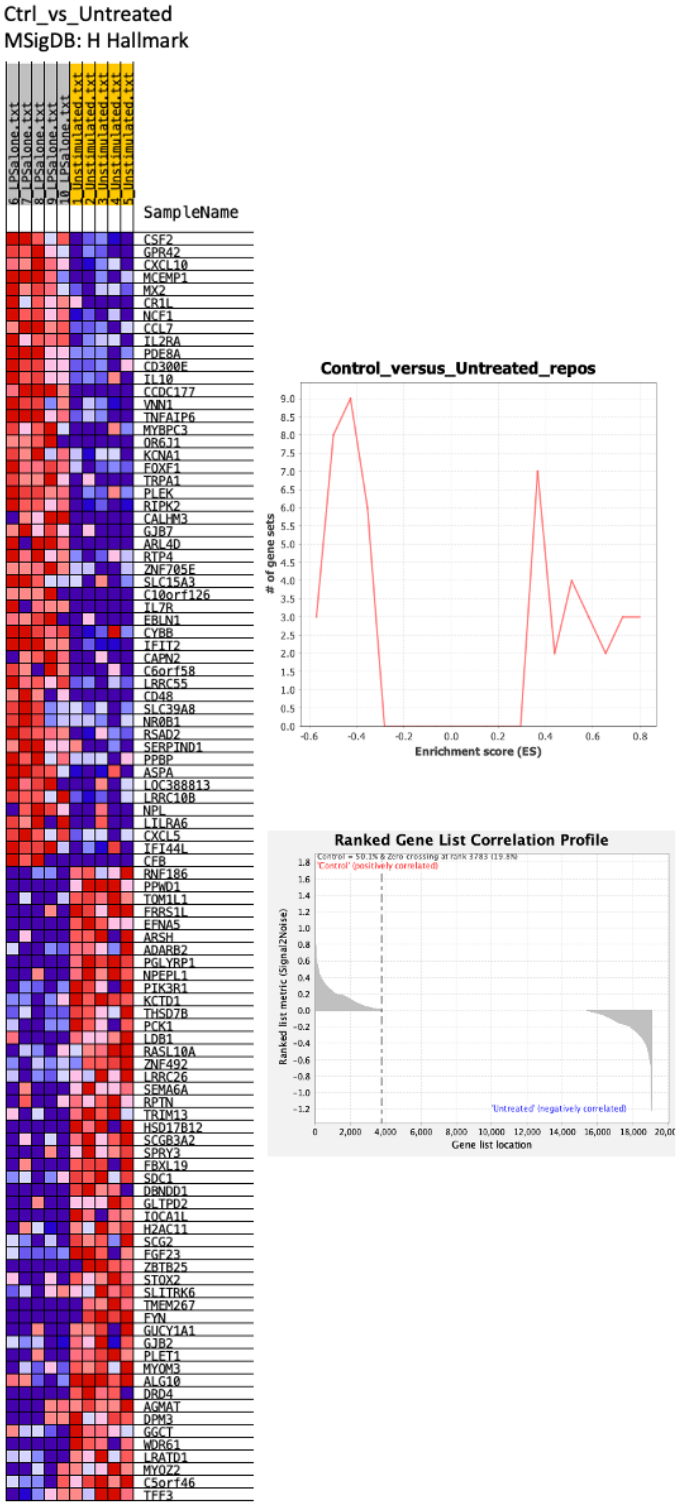


Figure 4.7: On the left: Heat map of top 50 features for each condition showing the correlation between the ranked genes and the treatment type between LPS controls and untreated (MSigDB: H). Gene expression is represented by a range of colors (red, pink, light blue, dark blue) shows the range of expression values (high, moderate, low, lowest) On the right: ES scores based on the number of gene lists, and correlation profile of the gene lists in comparison with treatment applications

### 4.3.6 WGCNA

Using WGCNA, data was first clustered according to euclidean distances within the data set in order to detect any outliers. Next it was linked to the trait data to visualize whether the trait data had any connection with the illustrated dendrogram. The frankincense systemic samples clustered together. The sample names are labeled according to their testing condition. The FCO topical samples also clustered closely together. This would support our findings as the DEGs in the FCO topical applications were much different than the findings in the frankincense systemic.

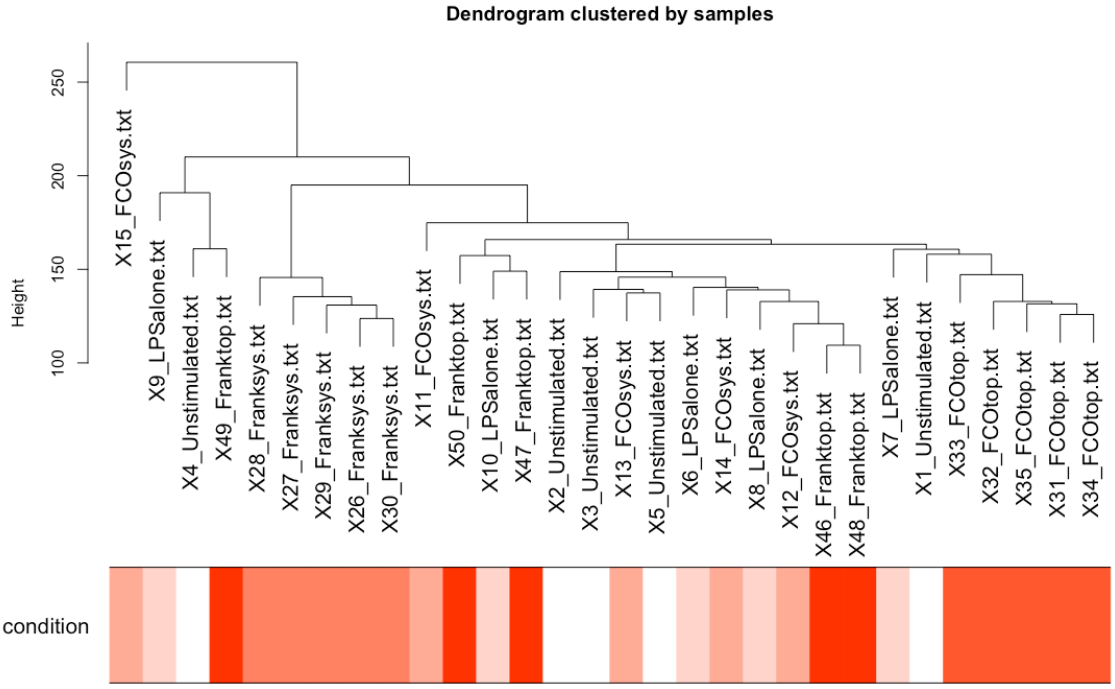


Figure 4.8: Sample clustering by euclidean distances with WGCNA

WGCNA detected a total of 22 different modules. A total of 1504 genes were seen in the module with the highest number of genes, and 39 genes were detected in the module with the lowest number of genes. Next the genes within the data set are clustered according to genes with dissimilarity based on topological overlap, together with the assigned modules. First the dissimilarity between the genes is calculated using the topical overlap measure (TOM). TOM measures the pairwise

similarity measure between the genes. If TOM is high, then genes have similar overlap between neighbor genes. This implies the genes have similar expression patterns across the samples.[130] Although, TOM is measuring similarity between the genes and their networks 1-TOM will produce the dissimilarity seen in the dendrogram. The dendrogram produces hierarchical clustering results of the genes based on module detection. All 8390 genes were split into 2 total gene blocks which produced 2 different cluster dendrograms. Pictured on the top half of figure 4.9.

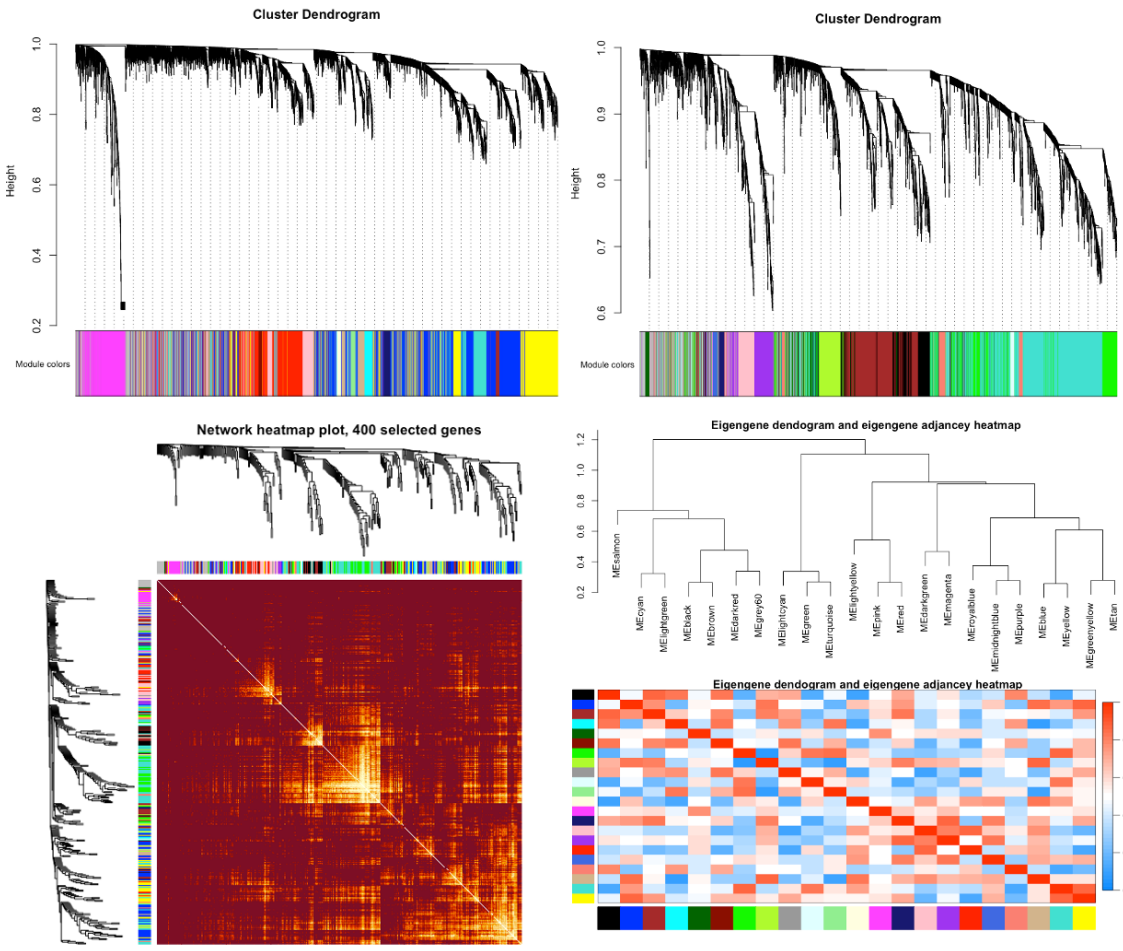


Figure 4.9: The dendrogram displays the clustering of genes with dissimilarity based on topological overlap, together with assigned module colors

The heat map on the bottom left of figure 4.9 displays genes that are highly correlated due to high co-expression interconnections. The more saturated the yellow and

red colors appear, the more the modules correspond to highly interconnected genes. The tips of each of the branches meet the modules where genes display the highest interconneectedness with the rest of the genes in the module.[130] The heat map was limited to the top 400 genes, as suggested by the developers due to high complexity and time of generating a heat map with all of the genes. Unfortunately, no clinical trait data could be used to see if there were any patterns between different traits and the gene expression networks. They are typically utilized in the graph pictured on the bottom right in figure 4.9. The traits will typically be intermixed within the module colors to detect modules with close relatedness to the gene cluster modules. The clustered modules were extracted from WGCNA and added into Stringdb to detect whether there was a general trend based on the genes present in each module. The larger modules did display more connected genes within the network, however, just because they displayed relativity with each other, they did not produce many associated pathways that provided much meaning. This could be due to the fact that the testing applications did vary greatly between samples. The genes produced by the FCO, were quite different than those generated by the frankincense. The genes from each of these modules were further analyzed using Stringdb to detect any trends within the modules. Some of the smaller modules failed to produce results that clearly indicated any clear pathway associations within the gene groups. Some of the largest modules produced many gene connections, but not necessarily coherent networks of genes that play a biological role in many related/clustered pathways. In the bottom image of 4.9, displaying the hierarchical clustering dendrogram of the modules, MEbrown and MEblack share the same branch. Upon further investigation of these modules in Stringdb, the black module played a role in chemokine signaling pathways, TNF signaling pathways, and IL17 signaling pathway and rheumatoid arthritis. The brown module had genes involved in PI3K-AKt signaling and cancer. The resulting brown and black modules are most likely due to genes involved with cancer and

inflammation associated with the frankincense EOs, the untreated samples and the LPS controls.

### 4.3.7 Comparison of frankincense to traditional NSAIDs

A microarray study had previously pre-treated celecoxib and rofecoxib on human aortic smooth muscle cells (3F1243) and exposed them to pro-inflammatory cytokine, IL1$\beta$. (GSE59671) The transcription profiling was compared to the results generated by frankincense systemic. Genes from each treatment were separated by their trend (increased/decreased expression) and compared. Celecoxib originally resulted in a higher number of DEGs than Rofecoxib. Frankincense systemic shared 44 genes that decreased in expression levels with the Celecoxib. Of these related genes, the majority of them were involved in IL-17 signaling, TNF signaling, cytokine-cytokine receptor interactions and chemokine signaling.[147] Both Celecoxib and Rofecoxib are known to target COX2 and are used to treat acute pain and inflammation in various types of arthritis.[22] Celecoxib and frankincense systemic both shared lower expression values for COX2 (-3.82, -2.60) respectively. Celecoxib has several other targets, but none were DE in the data set nor within the frankincense data set. Since Rofecoxib resulted in a low number of DEGs to begin with, those that could be used in comparison with frankincense were low. From the graph in figure 4.10, not many genes from Rofecoxib were DE and in common with frankincense systemic. Rofecoxib was commonly used in cases of arthritis but was voluntarily withdrawn from the market by the manufacturers due to possible increased risk of heart attack or stroke with long term usage of the drug.[22]

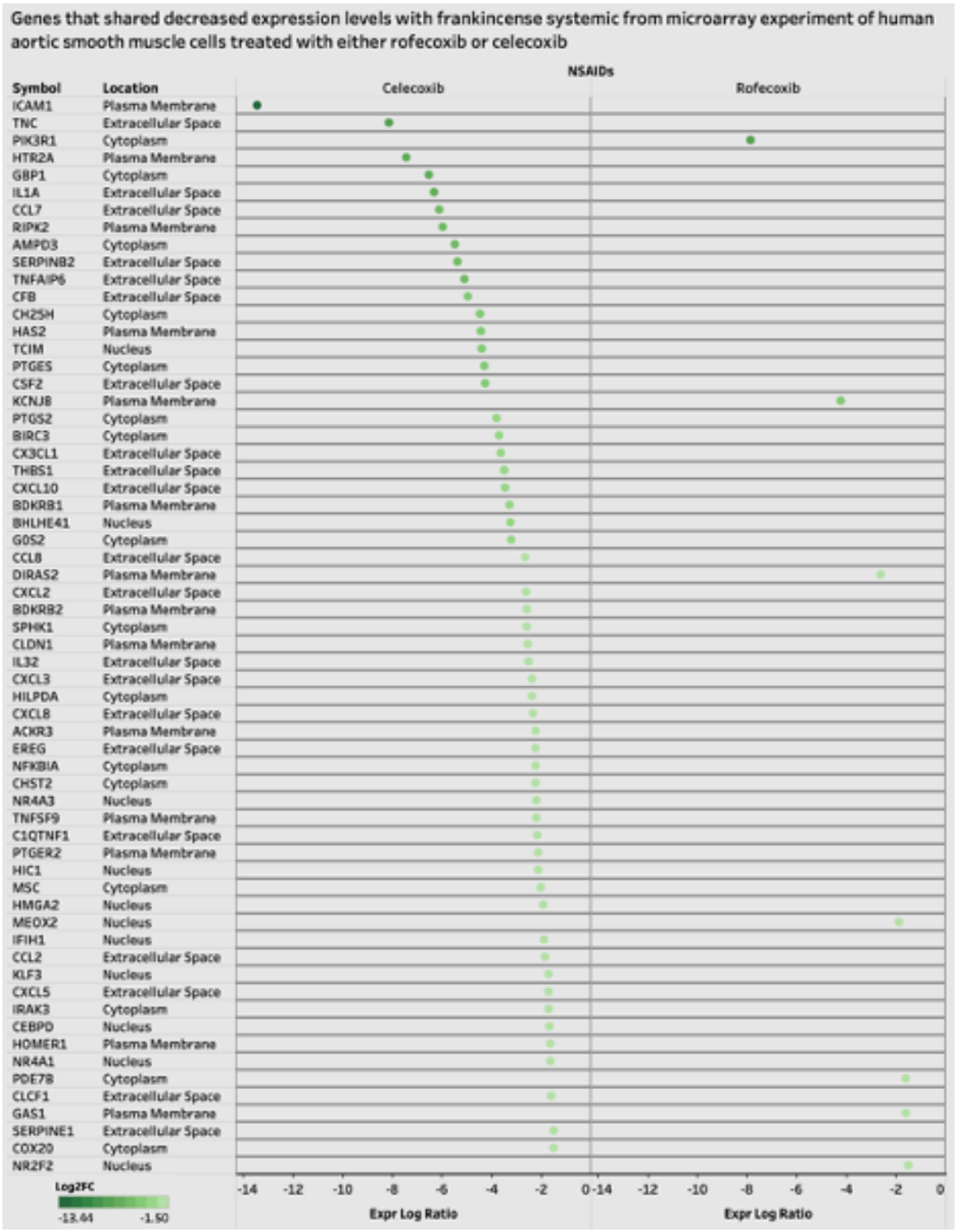


Figure 4.10: Graph displays all genes in common with frankincense systemic from the NSAID treatments with a Log2FC $\leq$ -1.5

### 4.3.8 Comparison of frankincense to five non-hematopoietic treated with a methylprednisolone

A newly available RNA-seq data set that tested methylprednisolone, a glucocorticoid on 9 different cell lines.[132] This data was used as a comparison tool for examining the inflammatory-related gene associations between the glucocoticoid and the frankincense systemic. The five non-hematopoietic (endothelial cells, fibroblasts, myoblasts, osteoblasts, and preadipocytes) were treated with the steroid. Samples were taken at two time points (2h and 6h). The data from the 6h time point was used in this comparison analysis. The graph shows genes in common that decreased in gene expression in both the steroid treatments and the frankincense systemic treatments. Genes that presented a Log2FC value of $\leq$ -1.5 were compared against frankincense systemic genes that also presented a Log2FC value of $\leq$ -1.5. The Log2FC values are colored in varying shades of green based on the Log2FC value of that individual gene in the steroid data set.

The graph in figure 4.11 displays these genes that were in common with those found in frankincense.

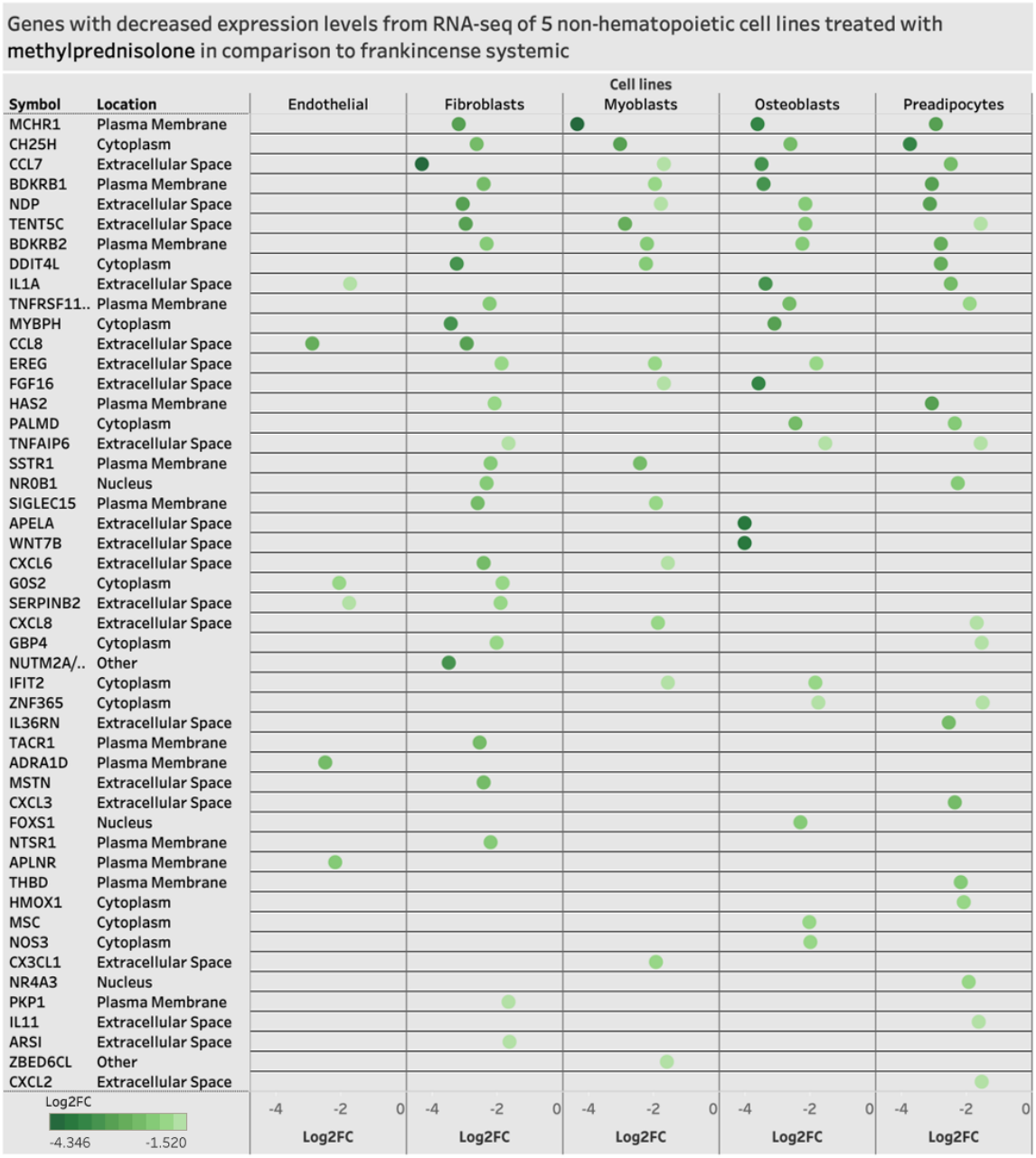


Figure 4.11: This graph displays expression data for genes that had a Log2FC $\leq$ -1.5 between the 5 non-hematopoietic cell lines

Methylprednisolone is known to target the glucocorticoid receptor (NR3C2). NR3C2 is involved in regulating an inflammatory response. NR3C2 was recorded in the data as a DEG in only one of the 5 non-hematopoietic: the myoblast cell line. Annexin A1 (ANXA1) is the second recorded target for methylprednisolone. ANXA1 was not seen

in the 5 non-hematopoietic expression data as a targeted gene. It is clear that both this steroid treatment and frankincense decreased expression in similar gene sets. The highest number of genes in common with frankincense are seen in the fibroblasts cell line. Many cytokines are shared amongst frankincense and these cell lines. CCL7, CCL8, CXCL6, CXCL8(IL8), IL1$\alpha$, are some of the main cytokines showing a decreased expression values after the treatment of methylprednisolone, and are also seen in the frankincense systemic treatments. The large overlap of similar DEGs from the data proposes the question of whether frankincense could be capable and further utilized as a therapeutic with steroid-like suppression capabilities.

#### 4.3.9 Comparison to frankincense with colchicine

The number of DEGs from this data set was quite low in expression level values to be used for comparison against the frankincense. Furthermore, many of the genes described as producing an anti-inflammatory mediated effect by the colchicine were not many in relation to the cytokine family. Some were associated with cell growth and differentiation, but very few were involved in inflammation.

#### 4.3.10 Principal component analysis (PCA) and hierarchical clustering of frankincense EO species using chemical profiles

The data scraped from the EOUdb, was used in R in order to plot a PCA based on the chemical profiles of different EOs, with the permission of EOUdb. The data was manipulated into the correct format to run this analysis and was loaded into R. R packages utilized for this analysis were: Readr, plyr, devtools, dplyr, mclust, and ggbiplot.[148][149][150][151][152][153][154]

At first, all the EOs in the EOUdb were used in the first analysis. This did not work well, considering the high dimensionality of the data. Many chemical components seen in some EOs are not seen in others. Making their scores unhelpful in drawing any conclusions about the data. Next based on reference guides, those suggested for uses

in inflammation were used next. But also proved to have the same issues, chemical profiles were too varied to use that large of a subset of oils. This would have been easier if the data set contained more numerical variable measurements seen across all oils, rather than certain ones. For these data sets all that was available was, chemical name, relative percentage (amount of that chemical present in the oil), oil name and the family the oil belongs to. Typically data sets used in these types of analyses have more phenotype or trait data that can be used in tandem with PCA calculations. First we used PCA to generate a graph based on different frankincense EOs and their chemical profiles based on chemicals and their respective relative percentages, seen in figures 4.12. In the figure the species of frankincense are color coded by their groupings shown in the figure legend. The principal component axes, displays the scores which are a linear combination of all the measurements. PC1 accounted for 35.5% of variance within the data, and PC2 accounted for 12.8%. The bottom image, displaying the projected scores on PC2 and PC3 axis, allow for a better visual for separation/grouping of the frankincense EO by family. Figure 4.13 projected the different chemical component variables in relation to the species of frankincense.

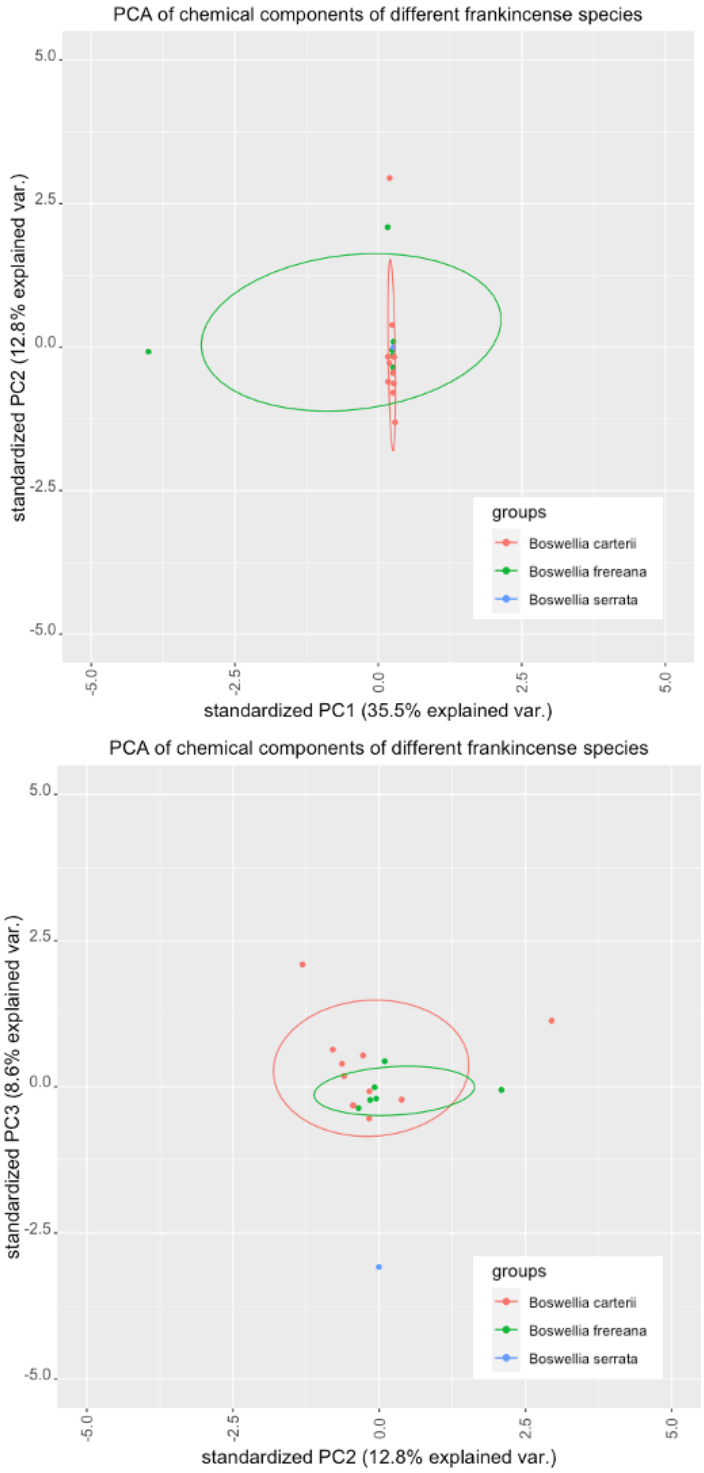


Figure 4.12: This graph displays a PCA plot by looking at the weight of the metabolites of the varying kinds of frankincense EO

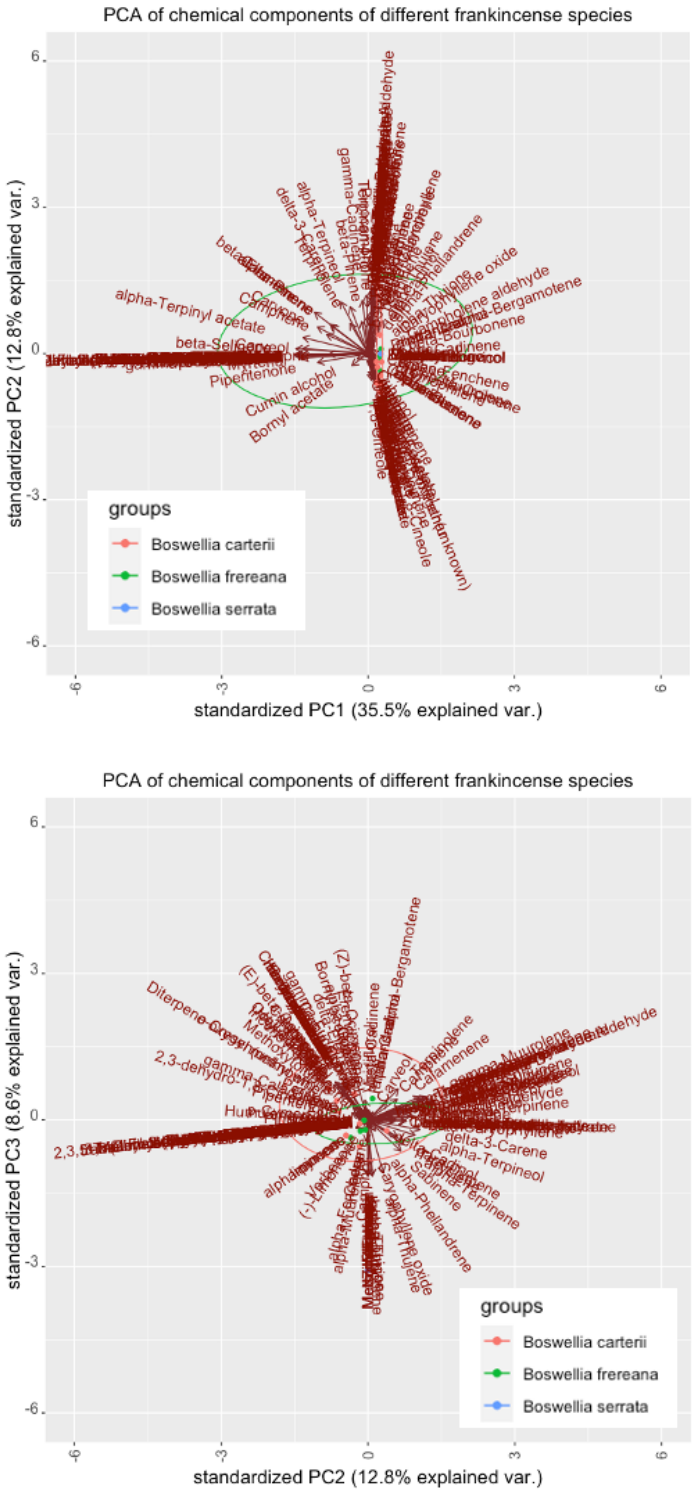


Figure 4.13: This graph displays a PCA plot by adding in the variables (chemical components)

Trying to perform a general PCA on the oils with proposed inflammatory properties (864 EOs) was too difficult to see any distinct variation between the types of EOs or their respective plant families. The generated variances from the PCA were too low. Frankincense belongs to the Burseraceae family, also known as the Torchwood family.[155] This family resides in the order of Sapindales. Sapindales is an order of flowering plants and contains many trees, as well as citrus producing trees.[155] It was intriguing to find out that maybe there was distinct groupings of the chemical components of different families within the Sapindales order. The Sapindales order contains many different families of plants. Only EOs belonging to 2 of families of the Sapindales order them were logged into EOUdb. This included mainly citrus EOs (Rutaceae family), frankincense EOs, and myrrh EOs(Burseraceae). They were used to perform the PCA in figure 4.14. EOs within the Burseraceae family seemed to show a slight distinction with EOs in the Rutaceae family, but there was not enough explained variance to support it.

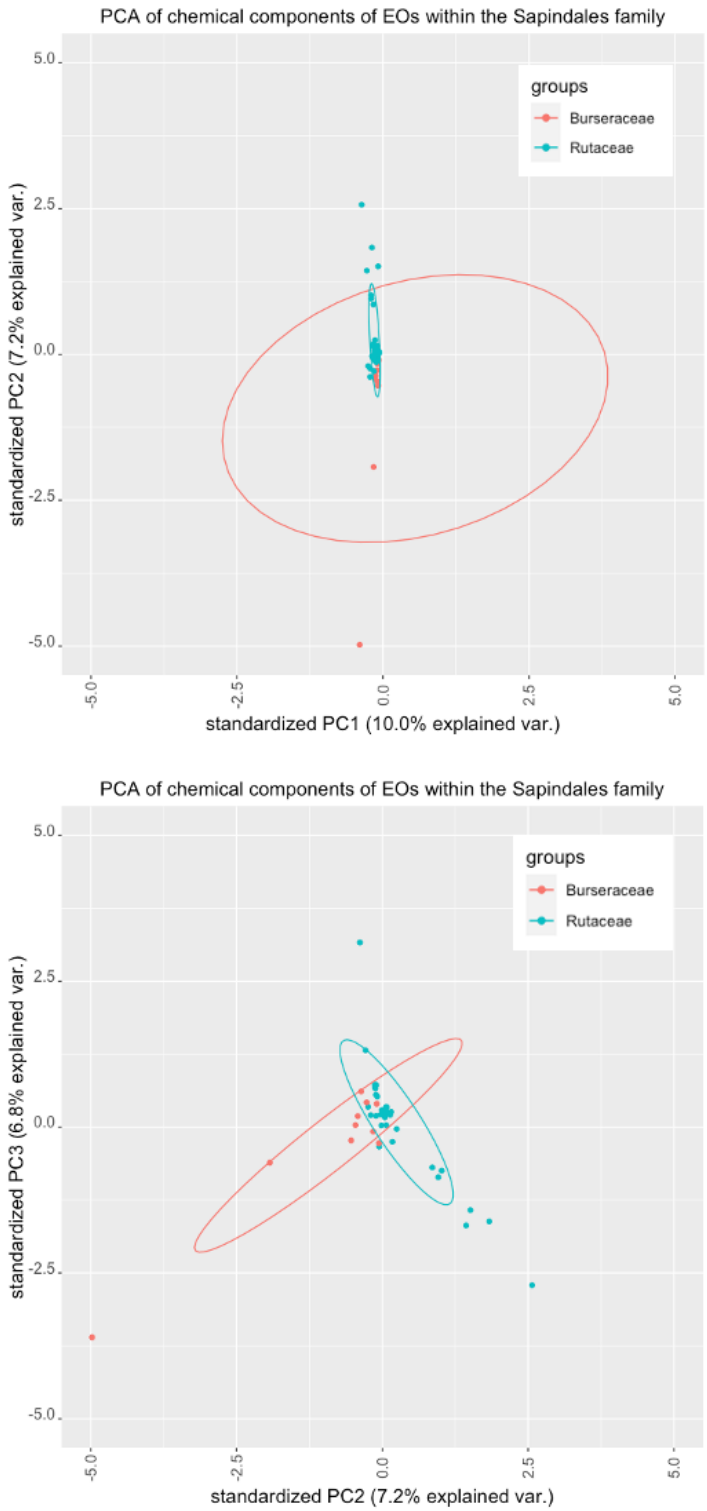


Figure 4.14: This graph displays a PCA plot of different EOs within the Sapindales order, which contained two families Burseraceae and Rutaceae.

Hierarchical clustering was employed in R using the mclust package.[149][152] The groupings displayed in figure 4.15 display some general grouping of *Boswellia carterii*, however other samples of the same species did not group in the same branch. The bottom figure displays the same data and graph, presented with different labels. Labels were replaced with those containing the geographical location for each sample. There are no clear groupings based on species, however, the groupings on the far right side did display a general Africa trend, even though they present different species of frankincense.

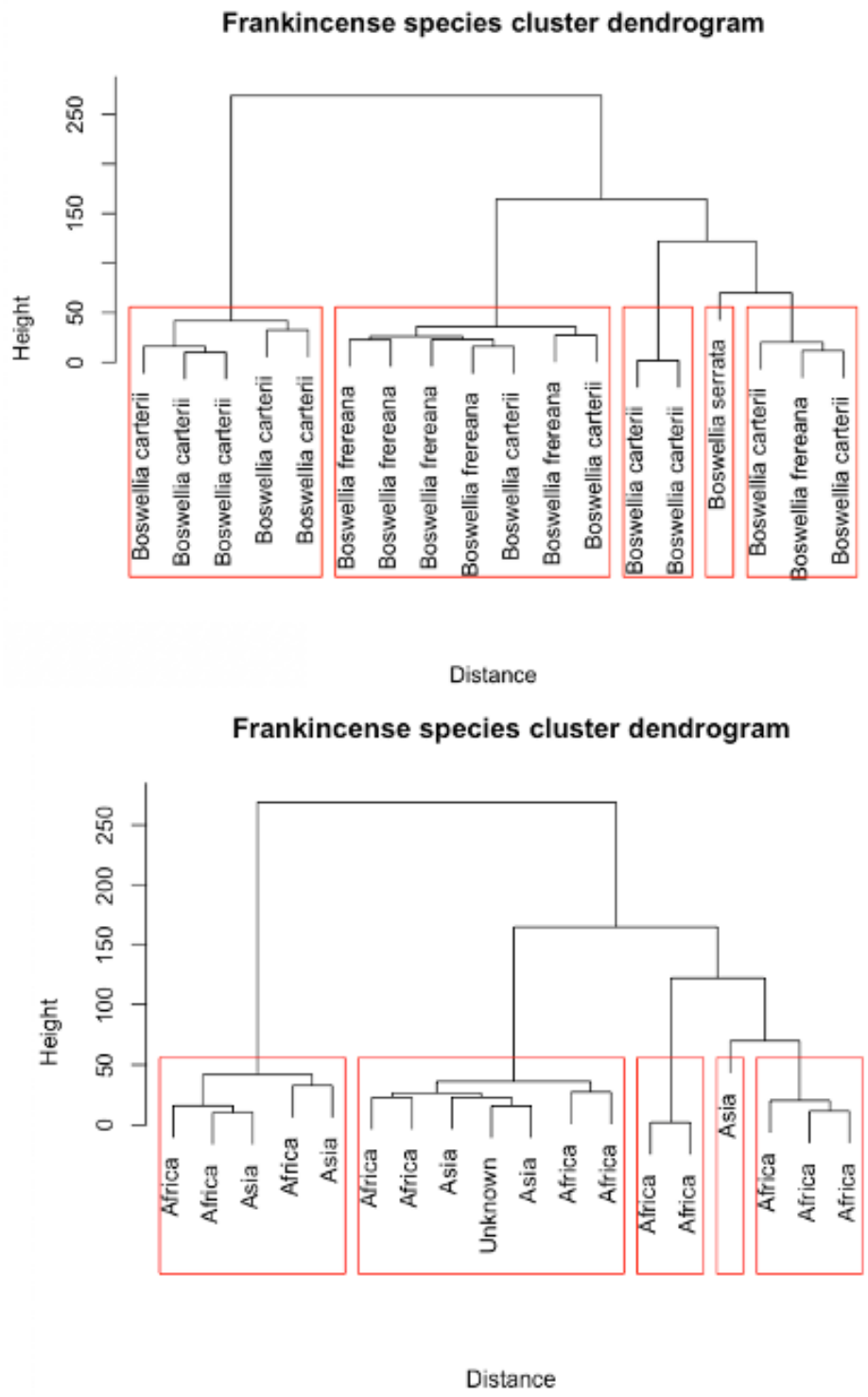


Figure 4.15: This graph displays two hierarchical clustering graphs. Both are displaying the same data presented with different labels.

### 4.3.11 Pathway analysis results in comparison to previously generated RT-qPCR results

#### 4.3.11.1 Frankincense systemic

The RT-qPCR results discussed in chapter 3, was generally in agreement with the frankincense systemic results from the RNA-seq. Pro-inflammatory biomarker IL1$\beta$ was significantly decreased in both donors and had a Log2FC of -4.197. However, IL1$\beta$ did not meet the cut off parameters set for the p-adj values in the RNA-seq results. IL8 (CXCL8) was decreased in both donors and confirmed in the RNA-seq results. CXCL10 expression was reduced in the RT-qPCR results in both donors but not at a statistically significant level. In the RNA-seq data, the Log2FC of CXCL10 was -4.984 with a p-adj of 2.76E-08. COX2 was decreased in both donors, and identified as a DEG in the RNA-seq data. Reported as its alias, PTGS2 with a Log2FC value of -2.64, p-adj value of 3.83E-03.

#### 4.3.11.2 Frankincense topical

In the RT-qPCR results, frankincense topical decreased 1 marker significantly in D1, CXCL10. This was further confirmed in the RNA-seq data as one of the few DEGs from the frankincense topical.

### 4.3.12 Pathway analysis results for frankincense systemic in comparison to EOKB Predictions

The EOKB derived 421 genes from various studies that were associated with the chemical compounds found in frankincense EO. The 1224 DEGs were also added into Stringdb to see the comparison in the associated Kegg and Reactome pathways previously predicted for frankincense EO. Chapter2, Stringdb found the top 5 Kegg pathways to be different from the top 5 predicted associated pathways using the EOKB data (table 2.1 for reference).

However, the top 5 generated Reactome pathways from the DEG data yielded

Table 4.14: Kegg pathway Stringdb results for frankincense EO (systemic) based on the DEGs calculated from the RNA-seq data

| **KEGG pathways** | | | |
|---|---|---|---|
| **Pathway** | **Description** | **Gene count** | **FDR** |
| hsa04060 | Cytokine-cytokine receptor interaction | 49/263 | 7.22E-07 |
| hsa05323 | Rheumatoid arthritis | 22/84 | 8.12E-05 |
| hsa04668 | TNF signaling pathway | 25/108 | 8.12E-05 |
| hsa05144 | Malaria | 14/47 | 0.0014 |
| hsa04657 | IL-17 signaling pathway | 20/92 | 0.0014 |

pathway associations similar for 4/5 predicted by the EOKB data

Table 4.15: Reactome pathway Stringdb results for frankincense EO (systemic) based on the DEGs calculated from the RNA-seq data

| **Reactome pathways** | | | |
|---|---|---|---|
| **Pathway** | **Description** | **Gene count** | **FDR** |
| hsa-6783783 | IL-10 signaling | 20/45 | 1.95E-06 |
| hsa-449147 | Signaling by interleukins | 61/439 | 3.2E-04 |
| hsa-1280215 | Cytokine signaling in immune system | 79/654 | 7.4E-04 |
| hsa-168256 | Immune system | 177/1925 | 0.0073 |
| hsa-6785807 | IL-4 and IL-13 signaling | 21/106 | 0.0111 |

These findings were similar pathways predicted by pathview, based on the expression data.[124] When comparing the number of DEGs detected in the RNA-seq analysis (1224) to those generated by the EOKB (421), 39 of those genes were found in common between the two data sets. The genes in common were: UBB, LEP, CCL19, MFSD11, CRP, GLA, CDSN, SNAI1, IL1$\alpha$, FST, PTGS2, XYLT2, NR4A3, MRC1, HMOX1, FLT1, IL10, SOD1, STIP1, NF$\kappa$BIA, NOD2, TREM1, TLR2, CD55, MAGEE1, PRDX1, MERTK, CYP4F3, CXCL8(IL8), HOXA10, CCL2, IL6, RNY5, CLU, ICAM1, ST13, IL11, F3, and MMP9. These genes are associated with Kegg Pathways: Rheumatoid arthritis, malaria, TNF-signaling, cytokine-cytokine receptor interaction, Leishmaniasis, IL-17 signaling, and NF$\kappa$B signaling. These results reveal the possible validation of the genes generated by the EOKB in conjunction with

those produced by the RNA-seq. This poses more potential for the possible therapeutic applicability for frankincense EO.

## 4.4 Discussion

### 4.4.1 Pathway analysis

The FCO systemic application did not yield any DEGs and it did not show any promising results in the RT-qPCR experiments. Due to its hydrophobic nature, it most likely did not mix well enough during sonification with the medium. Once the fatty-acid molecules settled, they most likely aggregated and floated back into the top of the receptor chamber. The FCO topical application, did penetrate the skin and produced 291 DEGs. Looking at the initial results in table 4.9, it appears the FCO topical application was involved with diseases and disorders similar to the frankincense systemic application. However, upon further investigation of some of the pathways listed (cancer, immunological disease, neurological disease) it appeared that the FCO topical application did not have much of a profound impact on any crucial cancerous markers or inflammatory markers. The frankincense systemic application had two more added tables than the FCO topical application. This was the "top regulator effect networks" and the "top upstream regulators" tables ( 4.5 and 4.6). This is because the FCO did not have strong enough data for the IPA program to predict an inhibition or activation action of any of the listed upstream regulators. Given a set of observed DEGs, the activation state of the upstream regulator is determined as "predicted" or "inhibited" by IPA.[156] This prediction is based on the associated relationship of the regulator to multiple associated genes in the data set using their relative z-scores to determine whether the regulator has significantly more "activated" predictions rather than "inhibited" predictions, or vice versa.[156] Results from IPA showed implications of possible hepatotoxicity of the frankincense systemic, due to the number of associated genes with their relative pathways. Upon further investigation of these targets, IPA displays genes interacting in liver inflammation and

viral hepatitis as the ones displayed in the inflammatory network above. As seen in the figure the majority of these genes are being down regulated. Similar observations can be made in networks associated with cardiotoxicity. This should continue to be investigated in a dosage dependent matter. Another interesting finding, not associated with inflammation was the networks affected that were associated with cancer-related pathways. The number of genes involved in cancer mediated pathways could be a future area of research for targeted cancer applications for frankincense. Results cannot be interpreted in regards to cancer applications due to the fact non-cancerous skin tissue was used in these models.

### 4.4.2 PCA and hierarchical clustering

Using PCA to possibly predict general trends in EO composition based on their uses, did not produce such results. This is most likely due to the high variability in these oils, at the species level, geographical areas of cultivation, and other factors. Variation in results can also be due to the fact some studies post general EO composition profiles and others state and identify each individual compound, even those found in trace amounts (>1%). This is an issue across research studies when publishing compositions of EO profiles. When looking using PCA to visualize the frankincense samples at a species level did explain more of the seen variance. However, more samples would be ideal before any evident trends can be detected. Employing hierarchical clustering for the frankincense data did not produce profound results either. Although, one grouping displayed *Boswellia carterii* species group together, there were still other branches also containing other samples of *Boswellia carterii* as well. When looking at the groupings by continent or country, no general trend could be seen based on the geographical location that were similar with frankincense species. This could be improved as the sample size for GC/MS results for frankincense oils increases. *Boswellia serrata* had a very limited number of samples. If the sample size could increase for this particular species, it may be easier to see more distant groups

from a clustering analysis of this frankincense species of EOs.

### 4.4.3 Comparisons with pharmaceuticals and OTCs

The results from the comparison analysis of gene expression studies utilizing NSAIDs and steroids versus frankincense EO were only used to investigate whether similar mechanisms and pathways may be at play. The conclusions of these data analyses cannot be relied on too heavily considering the models for testing applications were very different between the experiments. Further more, one of the bigger key differences were much shorter incubation times for the drugs tested (0hrs -6hrs) where as the frankincense EO incubated for 24 hours. The celecoxib/rofecoxib was a more comparable data set due to the fact the cells were pre-treated with the NSAIDs, similar to the EO and then exposed to IL1$\beta$.[131] In the study that tested methylprednisolone on non-hematopoietic cells, used cells derived from healthy donors and never exposed the cells to any type of inflammation. Comparisons were not made to the 4 hematopoietic cells also used in this study. This was due to the fact they circulate within the blood and bone marrow and would not be an ideal comparison model against the skin biopsy samples.[157] It was interesting to note how many genes in common that were DE in the frankincense systemic as well as in cell samples treated with either NSAIDs or glucocorticoids.

### 4.4.4 Limitations

For the experiment ran in Chapter 3, the incubation was 24 hours for the EO, and then following the EO treatment, they were stimulated with LPS and incubated an additional 24 hours. Due to the long period of incubation time, it becomes somewhat unclear on whether the results seen from the frankincense EO produced a direct response seen in ALL of these genes. It is very possible that frankincense targeted only a few genes during the initial phase of incubation, which then induced secondary associated responses induced by the chemokines as time increased. Ideally, this ex-

periment should be redone with skin samples that are tested with just EOs and no LPS. This would help determine what direct effects (if any) the frankincense itself may be causing. It would also deepen our understanding on the possible secondary response that may be initiated after triggering the immune system. The protocols used in the experiments in chapters 3 and 5 were based on Tergus pharmaceuticals and their protocols based on their internal studies. After further investigation, they noted similar results are seen in applications where incubation can be completed for only 6 hours after supplementation of LPS at a higher dosage to receive the same inflammatory mediated effects. Continuing to improve these protocols could also yield in more accurate results for future experiments.

### 4.4.5 Conclusion

Based on all of the data no solid conclusions can be made about the molecular mechanism of frankincense EO. However, indications from the data imply that frankincense EO has the potential to play a key role in mediating inflammation via initiation of the cytokine-cytokine receptor pathway as well as the TNF/NF$\kappa$B pathway. Because the frankincense was added prior to inducing inflammation, it is possible the frankincense EO is playing more of a protective role in response to the LPS within the skin tissue. It could instead be behaving as an antagonist on TNF$\alpha$, and toll-like-receptors (TLRs) which mediate inflammatory responses induced by LPS. Or it could be playing a role as a precursor to recruiting protective immune cells. Inhibition of TNF$\alpha$ or TLRs will decrease inflammatory signaling, and further prevent NF$\kappa$B signaling.[158] It is important to keep in mind that the RNA-seq data is looking at expression from the mRNA level. This does not necessarily mean that these are the changes that will be seen in the protein level. Further studies could be used to test the effect of frankincense EO on protein levels by running enzyme-linked immunosorbent assays (ELISA). These results reinforce the need to continue exploring this area of EO research in therapeutic applications.

# CHAPTER 5: CONTINUATION OF FRANKINCENSE EO ON LPS-INDUCED *EX-VIVO* SKIN EXPLANTS

## 5.1 Introduction

In order to increase the statistical power of the results from chapter 4, it was ideal to add a third donor to our experimental data set. However, in this continuation of experiments, it was decided to focus on solely frankincense EO LPS-induced *ex-vivo* skin explants. The experimental design differed slightly from our previous experimental design described in chapter 4. The previous studies conducted in chapter 3, tested a frankincense EO that was produced in 2016. EOs will degrade over time due to oxidation and their sensitivities to light and temperature.[7] The rate of oxidation will vary depending on the chemical composition of the EO. Oxidation tends to occur faster in EOs with higher concentrations of monoterpenes and other volatile compounds.[3] This would include EOs rich in alpha/beta-pinene or limonene. As EOs oxidize over time, its been shown in some studies to have an effect on efficacy of the EO.[159] EOs that have anti-fungal/antimicrobial properties can be effected as EOs oxidize over time.[160] EOs containing high concentrations of antioxidants (typically phenolic compounds) are less prone to degradation, such as eugenol, thymol, carvacrol and 1,8-cineole.[161][3] Degraded EOs have a higher probability of inducing contact dermatitis and other skin sensitivities.[162][163][164] The EOs used during these experiments were stored properly, out of direct sunlight and/or warm temperatures. It is not clear whether frankincense EO will lose or vary in efficacy due to these factors and oxidation. Based on the chemical composition of frankincense, it has some potential to degrade, but there are no current existing studies to verify how much or to what extent. For this reason, it was ideal to add a newer frankincense

EO in addition to the old frankincense EO for testing purposes. They are both manufactured by the same company which provides more consistency in the cultivation, harvesting and distillation between the two batches. The constant variability in EO profiles is a downside of testing any natural plant extracts. It makes reproducibility and standardization between experimental assays difficult.

Fatty acid (FA) oils are commonly used in cosmetic and beauty care products as emulsifiers, stabilizers and to enhance absorption of the product.[123] FAs and their ability to penetrate the skin has not been incredibly well documented. More research focuses on FA uptake and lipid metabolism inside the body/cells. This area of research is beginning to pick up traction due to the increased use of FA oils in beauty products. Dupont studied the penetration of fatty acid isomers on skin in 1989 and did not report significant findings for any associations between unsaturated fatty acids or saturated fatty acid differences as enhancers.[165] They also reported these fatty acid oils had minimal skin irritation potential.[165] Most trained aromatherapists recommend always adding a FA carrier oil to EOs before topical applications in order to dilute the EOs to reduce the risk of contact dermatitis.[3] This can include: grapeseed oil, FCO, almond oil, olive oil, jojoba oil, and more. FCO is commonly chosen as a carrier oil of chose due to its stable shelf life. There is some variation on what FCOs are comprised of. There are different methods for producing FCO. Some may be refined, bleached and deodorized in order to elongate their shelf life, which makes it more valuable for use with EOs.[94] The coconut oil industry, will state that FCOs "should" undergo processes to remove long chain fatty acids, leaving mainly medium-chain triglycerides (MCTs) in the final product.[166] However, like all natural products, there are no federally regulated guidelines on the composition and production of FCO and other coconut oil products. Furthermore, composition of FCO can only be determined through FA testing methods. The FCO used in our assays was advertised as containing mainly MCTs, however, composition testing in table 3.9,

revealed that the FCO was not mainly comprised of MCTs and carried a profile very similar to standard coconut oil. It was seen in the previous results that the FCO topical did penetrate the skin successfully. We hypothesized that FCO may be able to enhance penetration and absorption of the topical applications of the frankincense EOs. Of the DEGs expressed by FCO topical, most of them did not play key roles in mediating or ameliorating inflammation, but it may help to improve the absorption of the EO into the skin. Many users utilize FCO to dilute EOs for dermal application purposes. Many roll-on oil bottles typically contain components from FCO such as caprylic-capric triglycerides (CCTs). Previous studies have looked at coconut oil for enhancing skin absorption.[167] The FCO was comprised of mainly saturated triglycerides and a small portion of monounsaturated triglycerides. There are studies investigating CCTs for use in conjunction with nanoparticles and active therapeutic molecules for drug for drug targeting and delivery.[168] It is unclear whether the FCO will deepen the penetration of frankincense EO based on the literature and previous studies.

## 5.2 Methods

### 5.2.1 Human skin acquisition and testing

This skin testing was completed in collaboration with Tergus Pharma. The methods used for acquiring, prepping and testing on the skin were the same as described in chapter 3, section 3.2.

### 5.2.2 EO acquisition

The same *Boswellia carterri* (frankincense) EO and FCO used in chapter 3, section 3.2 were also used in this continuation study. The addition of a new frankincense EO was added to this assay. This new frankincense EO was also cultivated and distilled in Somalia by the same company that produced the previously used frankincense EO. The chemical profile for the new frankincense was determined using GC/MS and is

displayed in table 5.1.

Table 5.1: GC/MS results of the new frankincense EO

| **Component** | **Relative area percent** |
|---|---|
| Alpha-pinene | 47.4% |
| Limonene | 13.0% |
| Myrcene | 5.4% |
| Sabinene | 3.6% |
| Alpha-thujene | 3.3% |
| Octyl acetate | 0.2% |
| **Total** | **76.2%** |

### 5.2.3 Experimental design

This experimental design differs slightly from chapter 3. This experiment produced the same number of samples as the previous donors have (50 total). Instead of testing blue spruce and myrrh EO, we used 2 different frankincense EOs (old and new). The experimental design is displayed below in figure 5.2.

Table 5.2: LPS and EOs continuation study design per donor 3

| **Samples** | **Experimental application** | **Total vol. of oil used** |
|---|---|---|
| 1-5 | Untreated | N/A |
| 6-10 | LPS | N/A |
| 11-15 | FCO (systemic) + LPS | 10uL |
| 16-20 | New frankincense (systemic) + LPS | 10uL |
| 21-25 | Old frankincense (systemic) + LPS | 10uL |
| 26-30 | FCO (topical) + LPS | 10uL |
| 31-35 | New frankincense (topical) + LPS | 10uL |
| 36-40 | Old frankincense (topical) + LPS | 10uL |
| 41-45 | New frankincense blend 1:1 FCO (topical) + LPS | 20uL |
| 46-50 | Old frankincense blend 1:1 FCO (topical) + LPS | 20uL |

### 5.2.4 Topical and systemic applications

Topical and systemic applications were treated the same was as stated in chapter 5, section 3.2. In addition to the topical and systemic applications we added a 1:1 blend,

stated in samples 40-50 (figure 5.2). The frankincense blend application consisted of a 1:1 (Frankincense:FCO) mixture ratio (4uL of total volume, per sample) added topically onto the skin.

### 5.2.5 RNA-isolation and RT-qPCR

The RNA extraction and RT-qPCR were conducted in the same exact manner described in chapter 3, section 3.2. Taqman primers were designed to detect the same biomarkers in chapter 3(CXCL10, IL8, IL-1$\beta$ and COX-2).

### 5.2.6 Statistical analysis

The same statistical analysis was conducted in this section as performed in chapter 3, section 3.2. Studentized residual test was used to check for outliers and 8 were eliminated from the data set based on an r-student values of more that +3/-3.

### 5.2.7 Results

The ANOVA test produced significant results for each inflammatory marker. However, upon further analysis, it was seen some of the FCO applications increased inflammation significantly from the LPS controls.

Table 5.3: Results from ANOVA (based on log transformed fold change data). Markers where a significant change was observed according to their F-values were indicated by: * <0.05, ** <0.01, *** <0.001

| **Marker** | **Donor** | **F-value** | **Pr(>F)** |
|---|---|---|---|
| COX2 | D3 | 50.23 | 2E-16*** |
| IL1$\beta$ | D3 | 61.05 | 2E-16*** |
| CXCL10 | D3 | 58.4 | 2E-16*** |
| IL8 | D3 | 82.11 | 2E-16*** |

Tables 5.4 and 5.5 displayed significant results in D3 based on application. Which will be discussed into further detail with figures 5.1 and 5.2 below.

Table 5.4: Dunnett's method for D3 (COX2 and IL1$\beta$ based on log transformed fold change data). Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001

| **Application** | **t-value** | **Pr(>\|t\|)** | **t-value** | **Pr(>\|t\|)** |
|---|---|---|---|---|
| Marker | COX2 | COX2 | IL1$\beta$ | IL$\beta$ |
| FCO systemic - LPS | 5.143 | <0.001*** | 3.188 | 0.0193* |
| Old frank systemic - LPS | -7.769 | <0.001*** | -11.544 | <0.001*** |
| New frank systemic - LPS | -7.231 | <0.001*** | -10.636 | <0.001*** |
| FCO topical - LPS | 3.061 | 0.02714* | -0.355 | 0.9998 |
| Old frank topical - LPS | -2.533 | 0.0866 | 0.560 | 0.9951 |
| New frank topical - LPS | 3.891 | 0.00320** | 1.774 | 0.3751 |
| Old frank blend topical - LPS | 5.895 | <0.001*** | 2.666 | 0.0672 |
| New frank blend topical - LPS | 3.841 | 0.00359** | 0 | 1 |

Table 5.5: Dunnett's Method for Donor 3 (CXCL10 and IL8 based on log transformed fold change data). Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * <0.05, ** <0.01, *** <0.001

| **Application** | **t-value** | **Pr(>\|t\|)** | **t-value** | **Pr(>\|t\|)** |
|---|---|---|---|---|
| Marker | CXCL10 | CXCL10 | IL8 | IL8 |
| FCO systemic - LPS | 4.010 | 0.0023** | 3.080 | 0.0257* |
| Old frank systemic - LPS | -10.438 | <0.001*** | -12.818 | <0.001*** |
| New frank systemic - LPS | -11.794 | <0.001*** | -10.740 | <0.001*** |
| FCO topical - LPS | 2.642 | 0.0731 | 0.837 | 0.9507 |
| Old frank topical - LPS | -6.550 | <0.001*** | -0.289 | 1 |
| New frank topical - LPS | -1.115 | 0.8251 | 3.052 | 0.0277* |
| Old frank blend topical - LPS | -1.974 | 0.2747 | 5.756 | <0.001*** |
| New frank blend topical - LPS | -0.268 | 1 | 2.601 | 0.0787 |

Displayed in figure 5.1 the FCO produced significant increases in LPS signaling in COX2 and CXCL10. The blends with added FCO also increased inflammation significantly. Both the frankincense systemic applications decreased inflammation significantly. The old frankincense topical neat application, did display a decrease in CXCL10, as seen in D1, both the RT-qPCR results and the RNA-seq results.

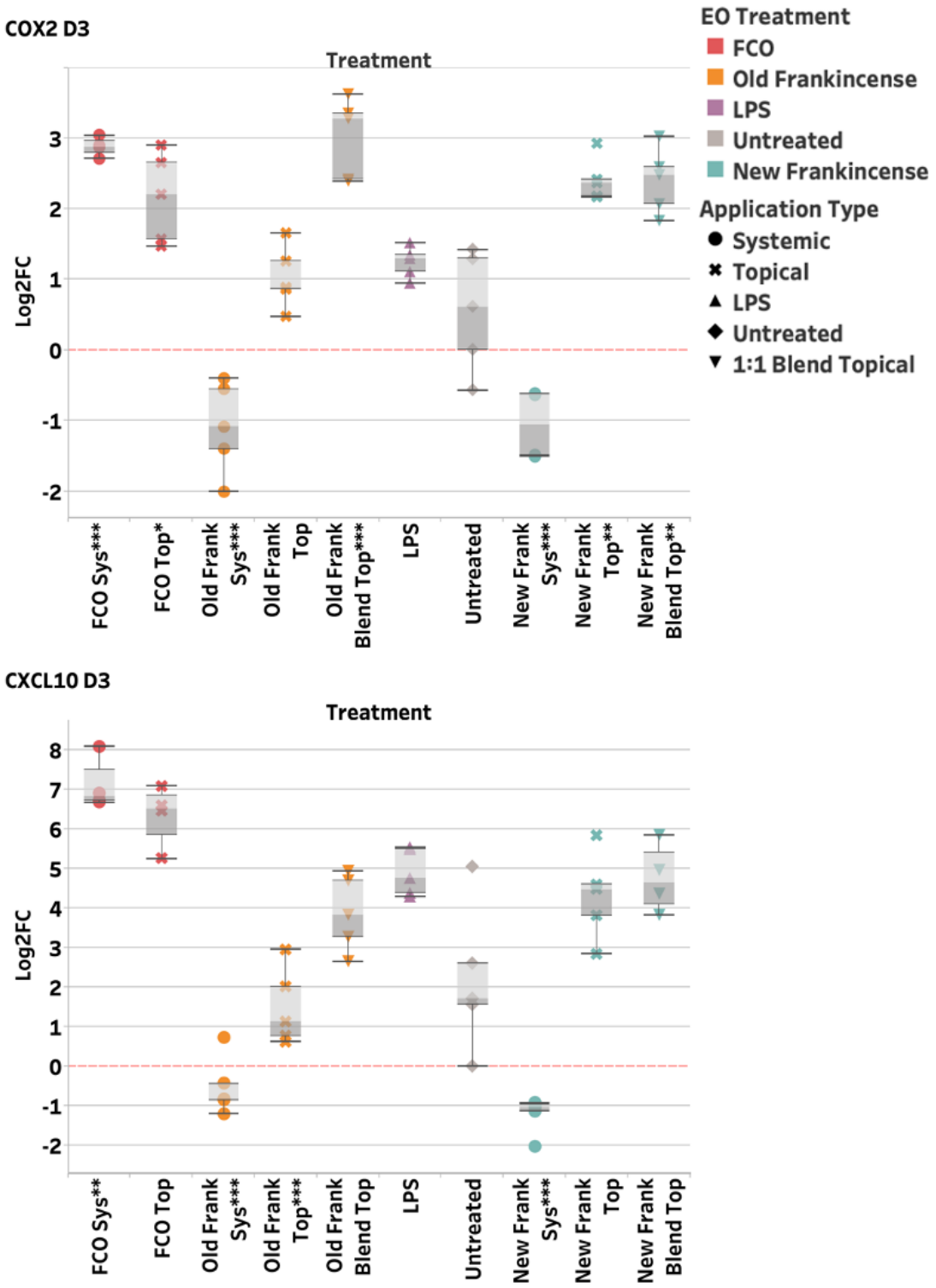


Figure 5.1: RT-qPCR results from COX2 and CXCL10 (D3). Each graph displays the Log2FC values for each application according to the labeled biomarker on the top left corner of each plot. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * $<0.05$, ** $<0.01$, *** $<0.001$

The old and new frankincense systemic applications continued to reduce inflammatory markers IL1$\beta$ and IL8. The new frankincense topical application in IL8, when applied to the skin neatly, seemed to also increase inflammation significantly versus the LPS controls alone.

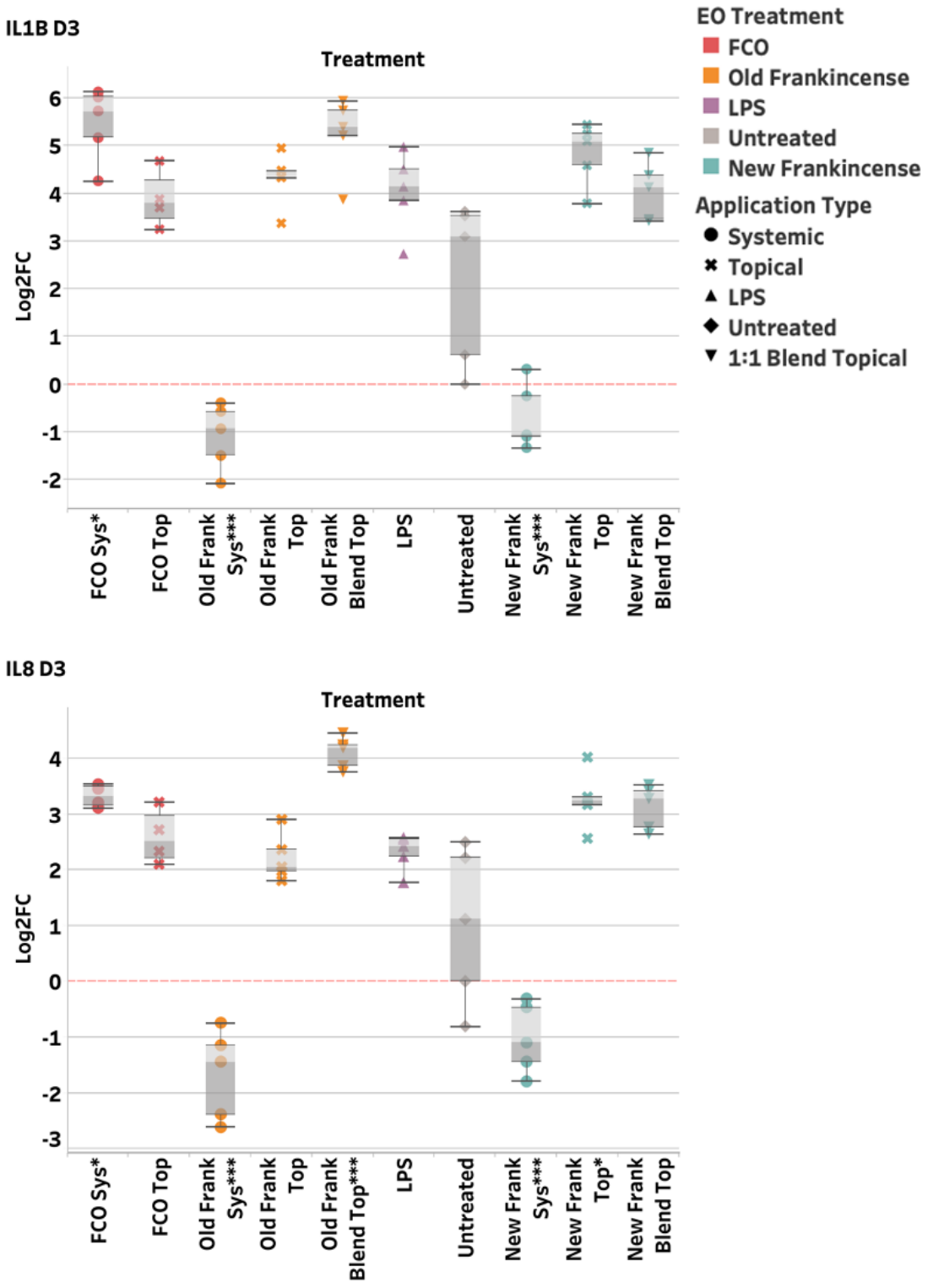


Figure 5.2: RT-qPCR results from IL1$\beta$ and IL8 (D3). Each graph displays the Log2FC values for each application according to the labeled biomarker on the top left corner of each plot. Applications with significant results are displayed on the axis labels using asterisks based on their p-values as follows: * $<0.05$, ** $<0.01$, *** $<0.001$

## 5.3 Discussion

After observing the results from the first assay in our EO studies, we hypothesized that the FCO could help to enhance penetration of the EO into the skin. However, we did not achieve those results in this third assay. In fact, the addition of FCO actually increased inflammation significantly. The FCO applications in chapter 3 seemed to not aid nor enhance the LPS response in the skin tissue systemically or topically. After the RNA-seq analysis, results suggested FCO penetrated the skin and produced DEGs involved in FA transport and metabolism. This sparked the curiosity for FCO to serve as a vehicle for the EO without further increasing the inflammatory signaling induced by the addition of LPS. Studies suggest that the coconut oil aids in alleviating skin inflammation, which has given it its rise in popularity in cosmetic and personal care products.[169][170] Lipids are typically non-immunogenic, and but have the possibility of being haptens.[84] Haptens can bind to proteins within the body and produce mild allergic reactions. Allergies to coconut oil are plausible, but quite rare. The Food Allergy and Anaphylaxis Network did a survey for nut allergies and found only 4 out of 5149 participants reported a coconut allergy.[171] Statistics on coconut oil allergies are not very precise or readily available. Coc n2, a 7S globulin and Coc n4, 11S globulin have been identified as allergens of coconut.[172] It is still possible for coconut oil to induce mild contact dermatitis or anaphylaxis, and symptoms should not be ignored due to low reported numbers of coconut allergies.[172] More doctors are noting the increase in children experiencing coconut allergies.[173] Many blogs have also reported increased anecdotal evidence of sensitivities and irritations to coconut oil when applied on the face.[174][175] It is unclear whether the patients skin samples that were used for this assay experienced any irritation or contact dermatitis from the FCO.

Another possibility could be the quality of the FCO may have degraded over time due to oxidation. Rancid oils have an increased likelihood of inducing irritation or

contact dermatitis.[3] A very recent FA composition test was carried out on the FCO in May 2020 before this FCO was used in these continued assays. No contaminants were found in the oil. The refinement process of coconut oil to produce FCO is supposed to increase its shelf life. The FCO itself appeared clear, colorless and odorless. However, it remains uncertain why in this assay it increased inflammation.

The main reasoning for using two different frankincense EOs was due to the possibility of the older frankincense losing efficacy or increasing inflammation due to possible oxidation and degradation of the EO. It was interesting to see that the two frankincense EOs did not differ immensely between their results. It is also interesting to note that the older frankincense EO topical application, decreased CXCL10 as seen in our previous experiments and RNA-seq analysis. However, the newer frankincense EO did not produce that same effect on CXCL10 in its topical applications. Both frankincenses performed similarly in their systemic applications for the various pro-inflammatory markers. The slight differences seen in their results could be due to the fact they have slightly different chemical profiles, rather than the age of the EO itself. This is another area that should be investigated further to fully understand the difference in effects of EOs that have had a substantial amount of time to oxidize and degrade.

If future experiments are conducted it would be wise to test a pharmaceutical enhancer as a vehicle to guarantee successful penetration of the EO topically for results.[176] A clear limitation of using skin biopsy samples resides in variability within the skin. As the human donor can vary in race/ethnicity and age, how the skin reacts to treatments can vary between some of these factors. Therefore, it would be ideal to continue adding more sample donors of various sex/age and ethnicity to increase statistical power of the results.

# CHAPTER 6: CONCLUSION

## 6.1 Improvements

The experiments carried out in chapters 3 and 5 could have been improved by increasing the number of samples and harvesting the skin of a subset of samples at earlier time points. This would have provided a more in depth look at the initial effects of LPS, rather than possible targets that were secondary effects from the longer incubation period and signaling of cytokines.

Another improvement to the study would have been to add skin samples that were tested with the EO and not stimulated with the LPS. This would have provided insights on the effects of the EO alone without inducing the inflammatory response. Adding these samples would also provide data for more complex statistical analyses.

## 6.2 Future studies

The application of frankincense EO for inflammatory-mediated diseases is a subject worth researching. If the capabilities are there for frankincense EO to hit gene targets associated with major inflammatory-mediated pathways, it should be further investigated on how to mimic these results through topical applications. If the systemic results can be replicated in topical applications, it could be used to provide relief in targeted areas of inflammation. For example: Joint pain from rheumatoid arthritis, skin inflammation inflicted by psoriasis or eczema, or localized areas of tissue damage from trauma.

### 6.2.1 Enhancers/vehicles to aid in dermal absorption

The use of pharmaceutical enhancements to drive penetration deeper should be tested in future studies. Dimethyl sulfoxide (DMSO) could be tested on skin tissue

as a carrier to enhance penetration and absorption of EOs into the skin.[177]

Nanoparticles have grown in popularity for their application in drug delivery. Using nanotechnology, researchers are now able to encapsulate drugs within nanoparticles for more effective targeted delivery of the drug, while providing the drug at lower dosages.[178] This field of nanotechnology in medical applications is now coined as nanomedicine, and can soon be seen on the forefront as personalized medicine continues to develop.[179] A nanoparticle delivery system for these EOs could provide insight on whether these systemic effects can be seen through topical applications.

### 6.2.2 Synergistic effects of EOs

Due to the fact EOs are comprised of many different chemical compounds, it remains unclear whether the effects can be pinpointed to one molecule alone, or rather the effects are due to molecules working together in synergy. Further studies could isolate these compounds and test them individually. The possible synergistic effects of EOs could be utilized in areas of polypharmacology. Collective data on several studies have noted the possible polypharmacological properties of just 1 EO chemical alone.[180] Further investigation of EOs and their chemical compounds separately and cohesively, could aid in the development of multi-targeting drug design.

### 6.2.3 How EOs can effect the microbiome

As more researchers begin to unravel the details of the microbiome, we are learning just how many factors may be influencing the gut flora. Researchers are still trying to understand the relationship between nutrition and different supplements can have on the microbiome. With people ingesting EOs as well as applying them topically, it raises concerns on whether these oils can alter the gut or skin flora. As mentioned earlier, there are many studies investigating the antimicrobial properties of EOs.[181] It currently remains unknown on whether these EOs can have an effect of the microbiome either through ingestion or in areas where applying topically. Studies have

suggested our skin flora can play a vital role in skin disorders and the functioning of immune systems.[182]

### 6.2.4 Accessibility of the EOKB

The EOKB is a fascinating tool to learn more about EOs and their chemical components. It was built in PostgreSQL but is lacking a graphical user interface (GUI). To access the EOKB, has to be done on the back end which requires basic knowledge of coding and databases. A GUI would provide accessibility for anyone without that knowledge or experience to access and use it. The most ideal version of this EOKB would reside in a phone application. Where users can look up basic information and explore current research that is available on EOs. Making the EOKB publicly available and accessible would also require general maintenance in order to repopulate entities and update them with the most recent literature findings. This could be a fairly reasonable internship project for a computer science student or web developer.

Currently, there are several PostgreSQL client cross-platforms that allow a user to securely access their databases and query from the database right through the phone application. However, query designing/writing (although not substantially difficult) is not something the average person will know how to do. PostgreSQL is a widely used database management system, and there are many readily available tools and resources to build an interface for it. PgAdmin, DBbeaver, Navicat and Datagrip are some of the most commonly used GUI tools for PostgreSQL.[183]

### 6.2.5 Further use of the EOKB

Many companies are now recommending personalized supplement plans to clients based on 23&me data. Using 23&me data with permission from the user, allows companies to look at individual single nucleotide polymorphisms (SNPs) and make health recommendations based on those SNPs. Utilizing this data within the EOKB could be useful to make connections between certain genes/SNPs with associated

studies of EOs based on these SNPs. This could be a potential area of further research.

## 6.3 Applications of EO research in current health affairs

The year 2019/2020 will forever be remembered by the year of the global pandemic inflicted by novel Coronavirus, SARS-CoV-2. Many health care systems had to gear up in preparations for the high influx of patients they were going to begin treating with Coronavirus disease of 2019 (COVID-19). FDA-approved drugs were being investigated for re-purposing measures and newly designed drugs were undergoing clinical trials for their uses in COVID-19. As we learned more about patients with COVID-19, it was noted that many patients with severe cases of COVID-19 began experiencing a "cytokine storm." A cytokine storm occurs when the immune system goes into an over-drive state trying to fight the virus, that it ends up inflicting more damage in the organism itself due to the high level of systemic inflammation. It has been theorized that elevated cytokines IL1, IL2, IL6, IFN$\gamma$, and TNF, can induced disseminated intravascular coagulation, multi-organ failure, acute respiratory distress syndrome and the potential of viral sepsis.[184] This caused research to begin investigating whether anti-inflammatory medications could help fight the infection. Other studies argue, immunosuppressive or anti-inflammatory drugs could be more damaging and detrimental to those with weakened immune systems, and may cause them to become unable to fight off the virus altogether.[185] However, it is important to note that the research and theories surrounding COVID-19 is a rapidly evolving topic, and most findings still remain unclear. Two clinical studies are currently being tried to test whether monoclonal antibodies that target IL6 can be beneficial to those with COVID-19, experiencing elevated levels of cytokines. (NCT04306705, NCT04322773)[186][25] It could be interesting to investigate whether frankincense EO could target any of those specific elevated cytokines.

Linguamatics was utilized by fellow PhD student, Aaron Trautman, to text mine associated genes from published studies in relation to corona viruses: SARS-CoV-

1, SARS-CoV-2, and MERS-CoV. Of these genes, those that played a role in the cytokine-cytokine receptor pathway were pulled aside. These genes were compared against DEGs associated in the cytokine receptor pathway from the frankincense systemic EO application discussed in chapters 3 and 5. IL6 was not targeted (directly or indirectly) by frankincense EO in our previous assays. However, when looking at figure 6.1, 5 genes in total were seen in common between all 3 Coronavirus strains and frankincense, CXCL10 being one of those genes, that was also seen in the frankincense topical application. CXCL8/IL8 was seen in both SARS-CoV-1 and SARS-CoV-2.

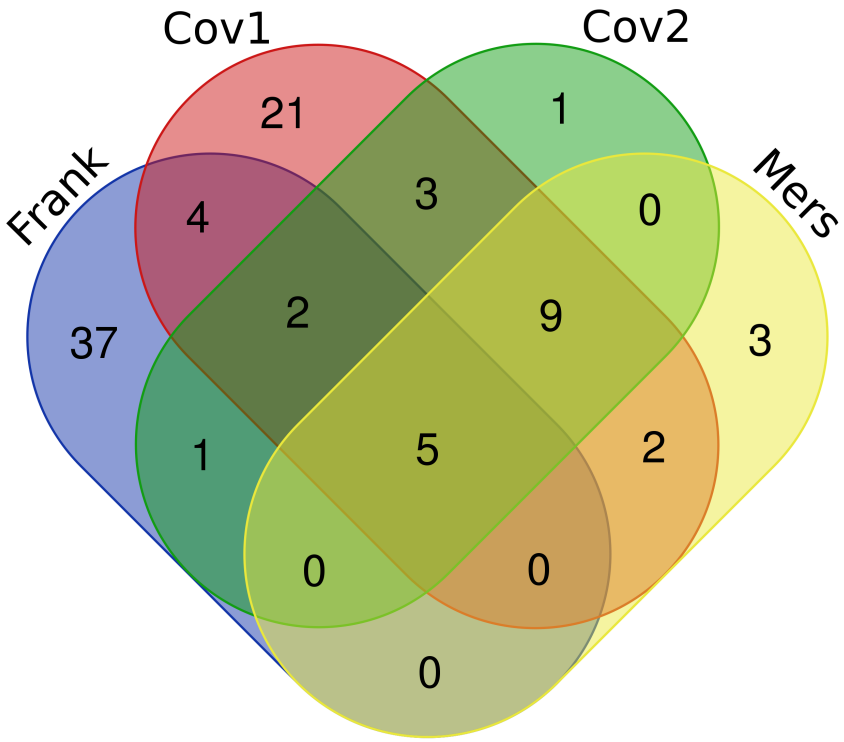


Figure 6.1: Genes in common between Coronavirus strains with frankincense EO, directly involved in cytokine-cytokine receptor pathway (hsa:04060) In the venn-diagram SARS-CoV-1(Cov1), SARS-CoV-2(cov2), and MERS-CoV(Mers)

It is clear there is still much more to learn about Corona viruses and frankincense EO before the two can be investigated together. However, this poses another potential area of application for it, proving that EOs would be worth exploring more.

## 6.4 Significance in the real world

This dissertation research served as a needed source for not only the public but also for researchers. The views/thoughts that were shared by other scientists when they heard this project focused on EOs was doubt, criticism and sometimes comedic jokes of pseudoscience. The perception of studying plant based medicine in the sciences if often not taken seriously and far too often overlooked. The excuse for lack of control and variability in natural substances, is a poor excuse for the scientific world to not pursue research topics in such areas. It should be a duty as researchers to push studies into areas that are unknown. We should have the desire to gain new insights and understanding in areas we know very little about. The fact that frankincense oils/extracts have their first reported uses dating back to BC eras, but to this day in 2020, NCBI only has a total of ≈227 research studies indexed on it, almost seems slightly unbelievable.[24] Noting that ≈39% of these total studies involve some type of cancer or inflammatory application. Currently two substantially large world wide health issues. At the end of the day there are millions of consumers (and growing) in the EO market. There are growing numbers of industry companies, producing and utilizing EOs and their chemical components(synthetic or natural) in body care products, cosmetics, cleaners, fragrances, supplements and foods, without having the extensive amount of research that is needed to be using these products and utilizing them fully.

## 6.5 Social Media Reach

While creating a dissertation that would provide more scientific evidence and education for the use of EOs, I created an Instagram platform in August, 2020 for just that. The Instagram page: oilyscientist shown in figure 6.2, has gained over 4,300 followers and continues to grow daily. Topics of discussion have included: Whether EOs contain vitamins, if EO molecules can accumulate within the body, whether one

can experience an allergic reaction to EOs even though EOs do not contain antigens, using EOs safely, interpreting scientific studies, etc. All discussions and information provided cites peer-reviewed scientific studies. The purpose of this platform is to break down and simplify the scientific concepts used in these studies for a general audience to understand. Followers are taught different aspects of judging the findings of scientific studies based on models used, experimental designs, testing methods, analytical methods and more.

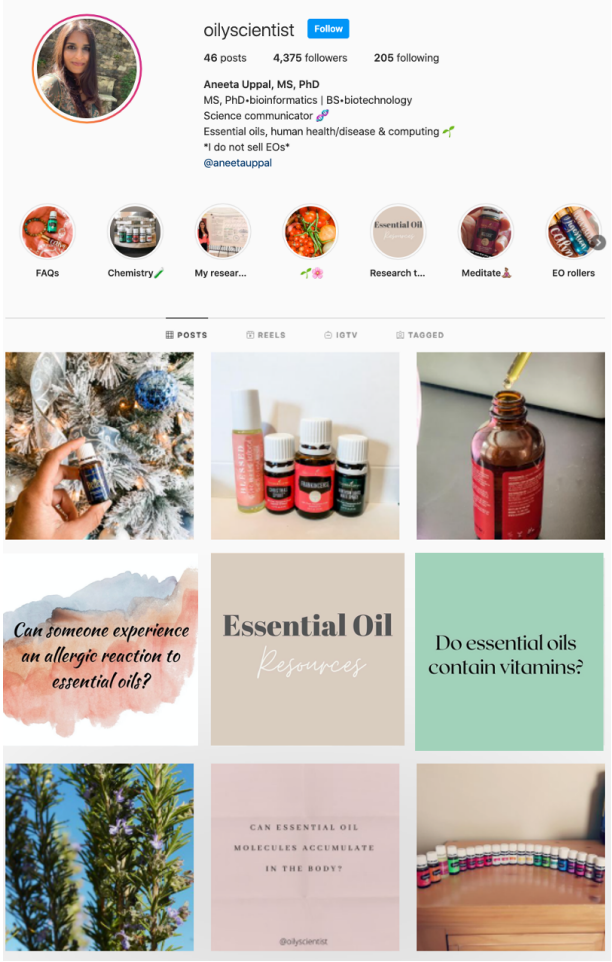


Figure 6.2: Oilyscientist Instagram page providing EO education from scientifically driven resources and peer-reviewed studies

Basic insights/statistics of overall page totals are listed in table 6.1. The demographics of followers in table 6.2 display followers spread across multiple continents

such as South Africa, the USA and Australia. Although the majority of the audience is females between the ages of 25-34, it is interesting to see some followers in the youngest age bracket (18-24) as well as the oldest age bracket of 65+.

Table 6.1: Contains running totals of social media page statistics to date (11-11-20)

| | |
|---|---|
| Posts | 39 |
| Videos | 4 |
| Video views | 15765 |
| Likes | 5721 |
| Comments | 1209 |
| Sharing content in stories/messages | 1643 |
| Saves | 1520 |

Table 6.2: Demographics and age range of social media followers

| **Demographics** | | **Ages** | |
|---|---|---|---|
| **Location** | **% of followers** | **Age Range** | **% of followers** |
| USA | 75% | 18-24 | 8.2% |
| United Kingdom | 4.9% | 25-34 | 51% |
| Canada | 4.0% | 45-54 | 8.2% |
| South Africa | 3.2% | 55-65 | 2.7% |
| Australia | 2.7% | 65+ | 1.2% |

## 6.6 Timeline and reflection

All of the goals were met from the first proposed timeline of the dissertation proposal. An additional donor was added in chapter 5, as requested by the dissertation committee. Additional tasks were completed such as submission of an NIH R21 grant to further the studies seen in chapters 3 and 5. The LPS paper for the results seen in chapters 3 and 4 is currently undergoing edits by collaborators before being submitted for publication, but is expected to be submitted by this fall (2020).

Table 6.3: Timeline of dissertation research

| Aim | Year 1 | Year 2 | Year 3 | Year 4 |
|---|---|---|---|---|
| 1.1 Text mining/data collection | | | | |
| 1.2 EOKB construction | | | | |
| 1.3 EOKB analysis | | | | |
| 1.4 EOKB updates | | | | |
| 2.1 Inflammatory skin assay | | | | |
| 2.2 RNA extraction for RT-qPCR | | | | |
| 2.3 Analysis for RT-qPCR | | | | |
| 2.4 RNA extraction for RNA-seq | | | | |
| 2.5 RNA-seq assembly | | | | |
| 3.1 Gene expression analysis | | | | |
| 3.2 Pathway and GSE analysis | | | | |
| 3.2 Analysis of metabolites of Frankincense | | | | |
| 3.3 Comparison to other NSAIDs | | | | |
| 3.4 Comparison to steroid/plant extract | | | | |
| 3.5 Additional added donor | | | | |
| 3.6 Submission of NIH R21 grant | | | | |
| 3.7 Publication for LPS Study | | | | |

# REFERENCES


[1] U. N. C. P. Center, "Poison Statistics, National Data 2018," tech. rep.

[2] P. Masango, "Cleaner production of essential oils by steam distillation," *Journal of Cleaner Production*, vol. 13, no. 8, pp. 833–839, 2005.

[3] R. Tisserand, R. Young, and E. Williamson, *Essential Oil Safety: A Guide for Health Care Professionals*. Churchill Livingstone, second ed., 2014.

[4] PDQ Integrative, Alternative, and Complementary Therapies Editorial Board, "Aromatherapy With Essential Oils (PDQ®): Health Professional Version," in *PDQ Cancer Information Summaries*, Bethesda (MD): National Cancer Institute (US), 2002.

[5] S. C. Knight, "Egyptian Funerary Practices," p. 5, 2009.

[6] R.-M. Gattefossé, R. Tisserand, and L. Davies, *Gattefossé's aromatherapy*. Saffron Walden: C.W. Daniel, 1993. OCLC: 28181433.

[7] K. Husnu Can Baser and G. Buchbauer, eds., *Handbook of Essential Oils: Science, Technology and Applications*. CRC Press, Taylor & Francis Group, 1 ed., Dec. 2009.

[8] G. V. Research, "Essential Oils Market Size, Share & Trends Analysis Report By Application (Cleaning & Home, Medical, Food & Beverages, Spa & Relaxation), By Product, By Sales Channel, And Segment Forecasts, 2019 - 2025."

[9] U. F. D. Administration, *Aromatherapy*. Aug. 2014.

[10] L. S. Publishing, *Essential Oils Desk Reference 7th Edition*. The Oily Essentials Purely Naturals, 2016. OCLC: 988579756.

[11] Z. A. A. Aziz, A. Ahmad, S. H. M. Setapar, A. Karakucuk, M. M. Azim, D. Lokhat, M. A. Kamal, and G. M. Ashraf, "Essential Oils: Extraction techniques, Pharmaceutical and Therapeutic potential - A review," *Current Drug Metabolism*, July 2018.

[12] C. R. Manion and R. M. Widder, "Essentials of essential oils," *American Journal of Health-System Pharmacy*, vol. 74, pp. e153–e162, May 2017.

[13] R. Pappas, "Essential Oil University Database," 1998.

[14] P. Satyal, T. H. Jones, E. M. Lopez, R. L. McFeeters, N. A. A. Ali, I. Mansi, A. G. Al-kaf, and W. N. Setzer, "Chemotypic Characterization and Biological Activity of Rosmarinus officinalis," *Foods*, vol. 6, Mar. 2017.

[15] Y. Endo, "Analytical Methods to Evaluate the Quality of Edible Fats and Oils: The JOCS Standard Methods for Analysis of Fats, Oils and Related Materials (2013) and Advanced Methods," *Journal of Oleo Science*, vol. 67, pp. 1–10, Jan. 2018.

[16] "Which Essential Oil Diffuser Should I Use? | dōTERRA Essential Oils." Library Catalog: www.doterra.com.

[17] C. Wong and R. Collins, "The Different Types of Essential Oil Diffusers," Dec. 2019. Library Catalog: www.verywellhealth.com Section: Verywell.

[18] "Substances Generally Recognized As Safe," Mar. 1977. Library Catalog: www.accessdata.fda.gov.

[19] M. Nevas, A.-R. Korhonen, M. Lindström, P. Turkki, and H. Korkeala, "Antibacterial efficiency of Finnish spice essential oils against pathogenic and spoilage bacteria," *Journal of Food Protection*, vol. 67, pp. 199–202, Jan. 2004.

[20] A. Man, L. Santacroce, R. Jacob, A. Mare, and L. Man, "Antimicrobial Activity of Six Essential Oils Against a Group of Human Pathogens: A Comparative Study," *Pathogens*, vol. 8, Jan. 2019.

[21] T. A. Thumann, E.-M. Pferschy-Wenzig, C. Moissl-Eichinger, and R. Bauer, "The role of gut microbiota for the activity of medicinal plants traditionally used in the European Union for gastrointestinal disorders," *Journal of Ethnopharmacology*, vol. 245, p. 112153, Dec. 2019.

[22] D. S. Wishart, Y. D. Feunang, A. C. Guo, E. J. Lo, A. Marcu, J. R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, N. Assempour, I. Iynkkaran, Y. Liu, A. Maciejewski, N. Gale, A. Wilson, L. Chin, R. Cummings, D. Le, A. Pon, C. Knox, and M. Wilson, "DrugBank 5.0: a major update to the DrugBank database for 2018," *Nucleic Acids Research*, vol. 46, no. D1, pp. D1074–D1082, 2018.

[23] "Methyl Salicylate." Library Catalog: www.drugs.com.

[24] N. C. for Biotechnology Information, "PubMed," 1996.

[25] N. L. of Medicine and N. I. of Health, "ClinicalTrials.gov."

[26] T. Sakurada, H. Kuwahata, S. Katsuyama, T. Komatsu, L. A. Morrone, M. T. Corasaniti, G. Bagetta, and S. Sakurada, "Intraplantar injection of bergamot essential oil into the mouse hindpaw: effects on capsaicin-induced nociceptive behaviors," *International Review of Neurobiology*, vol. 85, pp. 237–248, 2009.

[27] M. Sabino, V. A. O. Carmelo, G. Mazzoni, K. Cappelli, S. Capomaccio, P. Ajmone-Marsan, A. Verini-Supplizi, M. Trabalza-Marinucci, and H. N. Kadarmideen, "Gene co-expression networks in liver and muscle transcriptome reveal sex-specific gene expression in lambs fed with a mix of essential oils," *BMC Genomics*, vol. 19, p. 236, Apr. 2018.

[28] P. E. Simitzis, "Enrichment of Animal Diets with Essential Oils—A Great Perspective on Improving Animal Performance and Quality Characteristics of the Derived Products," *Medicines*, vol. 4, June 2017.

[29] C. Meason-Smith, C. E. Older, R. Ocana, B. Dominguez, S. D. Lawhon, J. Wu, A. P. Patterson, and A. Rodrigues Hoffmann, "Novel association of Psychrobacter and Pseudomonas with malodour in bloodhound dogs, and the effects of a topical product composed of essential oils and plant-derived essential fatty acids in a randomized, blinded, placebo-controlled study," *Veterinary Dermatology*, vol. 29, pp. 465–e158, Dec. 2018.

[30] A. Ali, M. M. Radwan, A. S. Wanas, and I. A. Khan, "Repellent Activity of Carrot Seed Essential Oil and Its Pure Compound, Carotol, Against Mosquitoes," *Journal of the American Mosquito Control Association*, vol. 34, no. 4, pp. 272–280, 2018.

[31] C. N. Dias and D. F. C. Moraes, "Essential oils and their compounds as Aedes aegypti L. (Diptera: Culicidae) larvicides: review," *Parasitology Research*, vol. 113, pp. 565–592, Feb. 2014.

[32] H. Koltai and C. Weingarten-Baror, "Specificity of DNA microarray hybridization: characterization, effectors and approaches for data correction," *Nucleic Acids Research*, vol. 36, pp. 2395–2405, Apr. 2008.

[33] Z. Wang, M. Gerstein, and M. Snyder, "RNA-Seq: a revolutionary tool for transcriptomics," *Nature reviews. Genetics*, vol. 10, pp. 57–63, Jan. 2009.

[34] T. Barrett, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C. L. Robertson, N. Serova, S. Davis, and A. Soboleva, "NCBI GEO: archive for functional genomics data sets—update," *Nucleic Acids Research*, vol. 41, pp. D991–D995, Jan. 2013. Publisher: Oxford Academic.

[35] P. Moretta, R. Clero, P. Benech, and T. Bader, "Gene Expression Profile Analysis and Identification of the Effects Triggered by Essential Sandalwood Oil on Human Skin Explants," *Clin Dermatol Res J*, vol. 2, no. 2, 2017.

[36] P. Cn, O. Jl, M. Rl, B. Cb, O. El, M. Dm, S. P, and S. Wn, "Antifungal and Cytotoxic Activities of Sixty Commercially-Available Essential Oils," June 2018. ISSN: 1420-3049 Issue: 7 Library Catalog: pubmed.ncbi.nlm.nih.gov Publisher: Molecules Volume: 23.

[37] H. Y, T. M, I. A, F. J, B. S, A. S, M. Ha, and Z. A, "Phytochemical Analysis, Cytotoxic and Antioxidant Activities of Myrtus Communis Essential Oil From Morocco," Jan. 2019. ISSN: 1553-3840 Issue: 3 Library Catalog: pubmed.ncbi.nlm.nih.gov Publisher: J Complement Integr Med Volume: 16.

[38] B. Najar, J. E. Shortrede, L. Pistelli, and J. Buhagiar, "Chemical Composition and in Vitro Cytotoxic Screening of Sixteen Commercial Essential Oils on Five Cancer Cell Lines," *Chemistry & Biodiversity*, Nov. 2019.

[39] D. Thapa, A. J. Richardson, B. Zweifel, R. J. Wallace, and S. W. Gratz, "Genoprotective Effects of Essential Oil Compounds Against Oxidative and Methylated DNA Damage in Human Colon Cancer Cells," *Journal of Food Science*, vol. 84, pp. 1979–1985, July 2019.

[40] P. a. S. R. Santos, G. B. Avanço, S. B. Nerilo, R. I. A. Marcelino, V. Janeiro, M. C. Valadares, and M. Machinski, "Assessment of Cytotoxic Activity of Rosemary (Rosmarinus officinalis L.), Turmeric (Curcuma longa L.), and Ginger (Zingiber officinale R.) Essential Oils in Cervical Cancer Cells (HeLa)," *TheScientificWorldJournal*, vol. 2016, p. 9273078, 2016.

[41] F. a. A. O. of the United Nations and W. H. Organization, *Fat and oils in Human Nutrition: Report of a Joint Expert Consultation*. Oct. 1993.

[42] T. Weicker, ed., *The Merck Report: A practical Journal of Pharmacy, Materia Medica and Chemistry*, vol. 10. Merck & Company, 1901. Google-Books-ID: hzpHAQAAMAAJ.

[43] L. L. Dilworth, C. K. Riley, and D. K. Stennett, "Chapter 5 - Plant Constituents: Carbohydrates, Oils, Resins, Balsams, and Plant Hormones," in *Pharmacognosy* (S. Badal and R. Delgoda, eds.), pp. 61–80, Boston: Academic Press, Jan. 2017.

[44] H. P. T. Ammon, "Boswellic Acids and Their Role in Chronic Inflammatory Diseases," *Advances in Experimental Medicine and Biology*, vol. 928, pp. 291–327, 2016.

[45] N. K. Roy, A. Deka, D. Bordoloi, S. Mishra, A. P. Kumar, G. Sethi, and A. B. Kunnumakkara, "The potential role of boswellic acids in cancer prevention and treatment," *Cancer Letters*, vol. 377, no. 1, pp. 74–86, 2016.

[46] X. Ni, M. M. Suhail, Q. Yang, A. Cao, K.-M. Fung, R. G. Postier, C. Woolley, G. Young, J. Zhang, and H.-K. Lin, "Frankincense essential oil prepared from hydrodistillation of Boswellia sacra gum resins induces human pancreatic cancer cell death in cultures and in a xenograft murine model," *BMC Complementary and Alternative Medicine*, vol. 12, p. 253, Dec. 2012.

[47] M. M. Suhail, W. Wu, A. Cao, F. G. Mondalek, K.-M. Fung, P.-T. Shih, Y.-T. Fang, C. Woolley, G. Young, and H.-K. Lin, "Boswellia sacra essential oil induces tumor cell-specific apoptosis and suppresses tumor aggressiveness in cultured human breast cancer cells," *BMC complementary and alternative medicine*, vol. 11, p. 129, Dec. 2011.

[48] C. Coronel, S. Morris, and P. Rob, *Database Systems: Design, Implementation and Management*. Cengage Learning, Sabatino, Joe, ninth ed., 2011.

[49] H. H. a. S. (HHS), "NIH Strategic Plan for Data Science | Data Science at NIH," Aug. 2019.

[50] C. F. Thorn, T. E. Klein, and R. B. Altman, "PharmGKB: the pharmacogenetics and pharmacogenomics knowledge base," *Methods in Molecular Biology (Clifton, N.J.)*, vol. 311, pp. 179–191, 2005.

[51] "UniProt: a worldwide hub of protein knowledge," *Nucleic Acids Research*, vol. 47, pp. D506–D515, Jan. 2019. Publisher: Oxford Academic.

[52] A. Fabregat, S. Jupe, L. Matthews, K. Sidiropoulos, M. Gillespie, P. Garapati, R. Haw, B. Jassal, F. Korninger, B. May, M. Milacic, C. D. Roca, K. Rothfels, C. Sevilla, V. Shamovsky, S. Shorser, T. Varusai, G. Viteri, J. Weiser, G. Wu, L. Stein, H. Hermjakob, and P. D'Eustachio, "The Reactome Pathway Knowledgebase," *Nucleic Acids Research*, vol. 46, no. D1, pp. D649–D655, 2018.

[53] "Essoil database."

[54] Y. Kumar, O. Prakash, H. Tripathi, S. Tandon, M. M. Gupta, L.-U. Rahman, R. K. Lal, M. Semwal, M. P. Darokar, and F. Khan, "AromaDb: A Database of Medicinal and Aromatic Plant's Aroma Molecules With Phytochemistry and Therapeutic Potentials," *Frontiers in Plant Science*, vol. 9, Aug. 2018.

[55] M. Kanehisa, Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe, "KEGG as a reference resource for gene and protein annotation," *Nucleic Acids Research*, vol. 44, pp. D457–D462, Jan. 2016.

[56] "NCBI Gene," 2004.

[57] Y. Kong, T. Wang, R. Wang, Y. Ma, S. Song, J. Liu, W. Hu, and S. Li, "Inhalation of Roman chamomile essential oil attenuates depressive-like behaviors in Wistar Kyoto rats," *Science China. Life Sciences*, vol. 60, pp. 647–655, June 2017.

[58] G.-S. Bae, K.-C. Park, S. B. Choi, I.-J. Jo, M.-O. Choi, S.-H. Hong, K. Song, H.-J. Song, and S.-J. Park, "Protective effects of alpha-pinene in mice with cerulein-induced acute pancreatitis," *Life Sciences*, vol. 91, pp. 866–871, Oct. 2012.

[59] Linguamatics, "I2E," 2001.

[60] A. Yu and J. Chen, "The Postgres95 User Manual," Sept. 1995.

[61] A. Davis, C. Grondin, R. Johnson, D. Sciaky, R. McMorran, J. Wiegers, T. Wiegers, and C. Mattingly, "The Comparative Toxicogenomics Database: update 2019," *Nucleic Acids Res*, Sept. 2018.

[62] A. P. Davis, C. J. Grondin, R. J. Johnson, D. Sciaky, B. L. King, R. McMorran, J. Wiegers, T. C. Wiegers, and C. J. Mattingly, "The Comparative Toxicogenomics Database: update 2017," *Nucleic Acids Research*, vol. 45, pp. D972–D978, Jan. 2017.

[63] M. González-Vallinas, S. Molina, G. Vicente, V. Zarza, R. Martín-Hernández, M. R. García-Risco, T. Fornari, G. Reglero, and A. Ramírez de Molina, "Expression of microRNA-15b and the glycosyltransferase GCNT3 correlates with antitumor efficacy of Rosemary diterpenes in colon and pancreatic cancer," *PloS One*, vol. 9, no. 6, p. e98556, 2014.

[64] M. G. Dozmorov, Q. Yang, W. Wu, J. Wren, M. M. Suhail, C. L. Woolley, D. G. Young, K.-M. Fung, and H.-K. Lin, "Differential effects of selective frankincense (Ru Xiang) essential oil versus non-selective sandalwood (Tan Xiang) essential oil on cultured bladder cancer cells: a microarray and bioinformatics study," *Chinese Medicine*, vol. 9, p. 18, 2014.

[65] D. W. Huang, B. T. Sherman, Q. Tan, J. R. Collins, W. G. Alvord, J. Roayaei, R. Stephens, M. W. Baseler, H. C. Lane, and R. A. Lempicki, "The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists," *Genome Biology*, vol. 8, no. 9, p. R183, 2007.

[66] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources," *Nature Protocols*, vol. 4, no. 1, pp. 44–57, 2009.

[67] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima, "KEGG: new perspectives on genomes, pathways, diseases and drugs," *Nucleic Acids Research*, vol. 45, pp. D353–D361, Jan. 2017.

[68] M.-N. I. of Genetic Medicine and M. John Hopkins University (Baltimore, "OMIM®."

[69] "Lucid Chart."

[70] D. Szklarczyk, J. H. Morris, H. Cook, M. Kuhn, S. Wyder, M. Simonovic, A. Santos, N. T. Doncheva, A. Roth, P. Bork, L. J. Jensen, and C. von Mering, "The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible," *Nucleic Acids Research*, vol. 45, no. D1, pp. D362–D368, 2017.

[71] L. Chen, H. Deng, H. Cui, J. Fang, Z. Zuo, J. Deng, Y. Li, X. Wang, and L. Zhao, "Inflammatory responses and inflammation-associated diseases in organs," *Oncotarget*, vol. 9, pp. 7204–7218, Dec. 2017.

[72] R. Pahwa and I. Jialal, "Chronic Inflammation," in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2019.

[73] P. Hunter, "The inflammation theory of disease," *EMBO Reports*, vol. 13, pp. 968–970, Nov. 2012.

[74] E. A. McGlynn, S. M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E. A. Kerr, "The Quality of Health Care Delivered to Adults in the United States," *New England Journal of Medicine*, vol. 348, pp. 2635–2645, June 2003.

[75] M. G. Netea, F. Balkwill, M. Chonchol, F. Cominelli, M. Y. Donath, E. J. Giamarellos-Bourboulis, D. Golenbock, M. S. Gresnigt, M. T. Heneka, H. M. Hoffman, R. Hotchkiss, L. A. B. Joosten, D. L. Kastner, M. Korte, E. Latz, P. Libby, T. Mandrup-Poulsen, A. Mantovani, K. H. G. Mills, K. L. Nowak, L. A. O'Neill, P. Pickkers, T. van der Poll, P. M. Ridker, J. Schalkwijk, D. A. Schwartz, B. Siegmund, C. J. Steer, H. Tilg, J. W. M. van der Meer, F. L. van de Veerdonk, and C. A. Dinarello, "A guiding map for inflammation," *Nature Immunology*, vol. 18, pp. 826–831, Aug. 2017.

[76] K. Jeena, V. B. Liju, and R. Kuttan, "Antioxidant, anti-inflammatory and antinociceptive activities of essential oil from ginger," *Indian Journal of Physiology and Pharmacology*, vol. 57, pp. 51–62, Mar. 2013.

[77] G. L. d. Silva, C. Luft, A. Lunardelli, R. H. Amaral, D. A. d. S. Melo, M. V. F. Donadio, F. B. Nunes, M. S. de Azambuja, J. C. Santana, C. M. B. Moraes, R. O. Mello, E. Cassel, M. A. d. A. Pereira, and J. R. de Oliveira, "Antioxidant, analgesic and anti-inflammatory effects of lavender essential oil," *Anais Da Academia Brasileira De Ciencias*, vol. 87, pp. 1397–1408, Aug. 2015.

[78] G. Sandner, M. Heckmann, and J. Weghuber, "Immunomodulatory Activities of Selected Essential Oils," *Biomolecules*, vol. 10, Aug. 2020.

[79] M. Hotta, R. Nakata, M. Katsukawa, K. Hori, S. Takahashi, and H. Inoue, "Carvacrol, a component of thyme oil, activates PPAR alpha and gamma and suppresses COX-2 expression," *Journal of Lipid Research*, vol. 51, pp. 132–139, Jan. 2010.

[80] R. de Cássia da Silveira e Sá, L. N. Andrade, and D. P. de Sousa, "A Review on Anti-Inflammatory Activity of Monoterpenes," *Molecules*, vol. 18, pp. 1227–1254, Jan. 2013.

[81] E. M. de Lavor, A. W. C. Fernandes, R. B. de Andrade Teles, A. E. B. P. Leal, R. G. de Oliveira Junior, M. Gama E Silva, A. P. de Oliveira, J. C. Silva, M. T. de Moura Fontes Araujo, H. D. M. Coutinho, I. R. A. de Menezes, L. Picot, and J. R. G. da Silva Almeida, "Essential Oils and Their Major Compounds in the Treatment of Chronic Inflammation: A Review of Antioxidant Potential in Preclinical Studies and Molecular Mechanisms," *Oxidative Medicine and Cellular Longevity*, vol. 2018, p. 6468593, 2018.

[82] H.-H. Hsu, J.-K. Kracht, L. E. Harder, K. Rudnik, G. Lindner, K. Schimek, U. Marx, and R. Pörtner, "A Method for Determination and Simulation of

Permeability and Diffusion in a 3D Tissue Model in a Membrane Insert System for Multi-well Plates," *Journal of Visualized Experiments : JoVE*, Feb. 2018.

[83] C. Alexander and E. T. Rietschel, "Invited review: Bacterial lipopolysaccharides and innate immunity," *Journal of Endotoxin Research*, vol. 7, pp. 167–202, June 2001.

[84] K. Murphy, *Janeway's Immunobiology*. Garland Science, 8th ed., 2012.

[85] M. Matsuura, "Structural Modifications of Bacterial Lipopolysaccharide that Facilitate Gram-Negative Bacteria Evasion of Host Innate Immunity," *Frontiers in Immunology*, vol. 4, 2013.

[86] T. Liu, L. Zhang, D. Joo, and S.-C. Sun, "NF-KB signaling in inflammation," *Signal Transduction and Targeted Therapy*, vol. 2, pp. 1–9, July 2017. Number: 1 Publisher: Nature Publishing Group.

[87] J.-S. Cho, J.-H. Kang, J.-Y. Um, I.-H. Han, I.-H. Park, and H.-M. Lee, "Lipopolysaccharide induces pro-inflammatory cytokines and MMP production via TLR4 in nasal polyp-derived fibroblast and organ culture," *PloS One*, vol. 9, no. 11, p. e90683, 2014.

[88] X. Liu, S. Yin, Y. Chen, Y. Wu, W. Zheng, H. Dong, Y. Bai, Y. Qin, J. Li, S. Feng, and P. Zhao, "LPS-induced proinflammatory cytokine expression in human airway epithelial cells and macrophages via NF-KB, STAT3 or AP-1 activation," *Molecular Medicine Reports*, vol. 17, pp. 5484–5491, Apr. 2018.

[89] W. Li, S. Yang, S. O. Kim, G. Reid, J. R. G. Challis, and A. D. Bocking, "Lipopolysaccharide-Induced Profiles of Cytokine, Chemokine, and Growth Factors Produced by Human Decidual Cells Are Altered by Lactobacillus rhamnosus GR-1 Supernatant," *Reproductive Sciences*, vol. 21, pp. 939–947, July 2014.

[90] A. Zlotnik and O. Yoshie, "The Chemokine Superfamily Revisited," *Immunity*, vol. 36, pp. 705–716, May 2012.

[91] A. Al-Harrasi and S. Al-Saidi, "Phytochemical Analysis of the Essential Oil from Botanically Certified Oleogum Resin of Boswellia sacra (Omani Luban)," *Molecules*, vol. 13, pp. 2181–2189, Sept. 2008.

[92] M. Schmiech, S. J. Lang, J. Ulrich, K. Werner, L. J. Rashan, T. Syrovets, and T. Simmet, "Comparative Investigation of Frankincense Nutraceuticals: Correlation of Boswellic and Lupeolic Acid Contents with Cytokine Release Inhibition and Toxicity against Triple-Negative Breast Cancer Cells," *Nutrients*, vol. 11, Oct. 2019.

[93] S. Su, Y. Hua, Y. Wang, W. Gu, W. Zhou, J.-a. Duan, H. Jiang, T. Chen, and Y. Tang, "Evaluation of the anti-inflammatory and analgesic properties

of individual and combined extracts from Commiphora myrrha, and Boswellia carterii," *Journal of Ethnopharmacology*, vol. 139, pp. 649–656, Jan. 2012.

[94] N. Parris, R. F. Joubran, and D. P. Lu, "Triglyceride Microemulsions: Effect of Nonionic Surfactants and the Nature of the Oil," *Journal of Agricultural and Food Chemistry*, vol. 42, pp. 1295–1299, June 1994.

[95] F. Krueger, "Trimgalore."

[96] S. Andrews, "FastQC."

[97] M. Martin, "Cutadapt removes adapter sequences from high-throughput sequencing reads," *EMBnet.journal*, vol. 17, pp. 10–12, May 2011.

[98] M. Pertea, D. Kim, G. Pertea, J. T. Leek, and S. L. Salzberg, "Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie, and Ballgown," *Nature protocols*, vol. 11, pp. 1650–1667, Sept. 2016.

[99] D. Kim, B. Langmead, and S. L. Salzberg, "HISAT: a fast spliced aligner with low memory requirements," *Nature Methods*, vol. 12, pp. 357–360, Apr. 2015. Number: 4 Publisher: Nature Publishing Group.

[100] N. A. O'Leary, M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B. Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y. Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M. Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K. Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O'Neill, S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz, H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully, A. R. Vatsan, C. Wallin, D. Webb, W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T. D. Murphy, and K. D. Pruitt, "Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation," *Nucleic Acids Research*, vol. 44, pp. D733–745, Jan. 2016.

[101] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup, "The Sequence Alignment/Map format and SAMtools," *Bioinformatics (Oxford, England)*, vol. 25, pp. 2078–2079, Aug. 2009.

[102] S. Anders, P. T. Pyl, and W. Huber, "HTSeq—a Python framework to work with high-throughput sequencing data," *Bioinformatics*, vol. 31, pp. 166–169, Jan. 2015.

[103] R. C. Team, "R: A Language and Environment for Statistical Computing," 2013.

[104] M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2," *Genome Biology*, vol. 15, p. 550, Dec. 2014.

[105] S. Anders and W. Huber, "Differential expression analysis for sequence count data," *Genome Biology*, vol. 11, p. R106, Oct. 2010.

[106] T. Hothorn, F. Bretz, and P. Westfall, "Simultaneous Inference in General Parametric Models," *Biometrical Journal*, vol. 50, no. 3, pp. 346–363, 2008.

[107] I. Pardoe, L. Simon, and D. Young, "9.4 - Studentized Residuals | STAT 462," 2018.

[108] A. Kassambara, "ggpubr: 'ggplot2' Based Publication Ready Plots," 2020.

[109] M. U. Rehman, M. Tahir, A. Q. Khan, R. Khan, n. Oday-O-Hamiza, A. Lateef, S. K. Hassan, S. Rashid, N. Ali, M. Zeeshan, and S. Sultana, "D-limonene suppresses doxorubicin-induced oxidative stress and inflammation via repression of COX-2, iNOS, and NFKB in kidneys of Wistar rats," *Experimental Biology and Medicine (Maywood, N.J.)*, vol. 239, pp. 465–476, Apr. 2014.

[110] X. Wang, G. Li, and W. Shen, "Protective effects of D-Limonene against transient cerebral ischemia in stroke-prone spontaneously hypertensive rats," *Experimental and Therapeutic Medicine*, vol. 15, pp. 699–706, Jan. 2018.

[111] Y. Qian, S. Yu, M.-Q. Shan, W.-F. Yao, and L. Zhang, "[Comparison between traditional processing and integration processing for Schizonepetae Herba based on chemical constituents and pharmacological effect]," *Zhongguo Zhong Yao Za Zhi = Zhongguo Zhongyao Zazhi = China Journal of Chinese Materia Medica*, vol. 41, pp. 2081–2086, June 2016.

[112] M. C. de Souza, A. J. Vieira, F. P. Beserra, C. H. Pellizzon, R. H. Nóbrega, and A. L. Rozza, "Gastroprotective effect of limonene in rats: Influence on oxidative stress, inflammation and gene expression," *Phytomedicine: International Journal of Phytotherapy and Phytopharmacology*, vol. 53, pp. 37–42, Feb. 2019.

[113] T. R. Ramalho, L. R. Filgueiras, M. T. Pacheco de Oliveira, A. L. d. A. Lima, C. R. Bezerra-Santos, S. Jancar, and M. R. Piuvezam, "Gamma-Terpinene Modulation of LPS-Stimulated Macrophages is Dependent on the PGE2/IL-10 Axis," *Planta Medica*, vol. 82, pp. 1341–1345, Oct. 2016.

[114] E. Games, M. Guerreiro, F. R. Santana, N. M. Pinheiro, E. A. de Oliveira, F. D. T. Q. S. Lopes, C. R. Olivo, I. F. L. C. Tibério, M. A. Martins, J. H. G. Lago, and C. M. Prado, "Structurally Related Monoterpenes p-Cymene, Carvacrol and Thymol Isolated from Essential Oil from Leaves of Lippia sidoides Cham. (Verbenaceae) Protect Mice against Elastase-Induced Emphysema," *Molecules (Basel, Switzerland)*, vol. 21, Oct. 2016.

[115] A. Kumar, K. Agarwal, M. Singh, A. Saxena, P. Yadav, A. K. Maurya, A. Yadav, S. Tandon, D. Chanda, and D. U. Bawankule, "Essential oil from waste leaves of Curcuma longa L. alleviates skin inflammation," *Inflammopharmacology*, vol. 26, pp. 1245–1255, Oct. 2018.

[116] X.-J. Li, Y.-J. Yang, Y.-S. Li, W. K. Zhang, and H.-B. Tang, "Alpha-Pinene, linalool, and 1-octanol contribute to the topical anti-inflammatory and analgesic activities of frankincense by inhibiting COX-2," *Journal of Ethnopharmacology*, vol. 179, pp. 22–26, Feb. 2016.

[117] R. Karthikeyan, G. Kanimozhi, N. R. Madahavan, B. Agilan, M. Ganesan, N. R. Prasad, and P. Rathinaraj, "Alpha-pinene attenuates UVA-induced photoaging through inhibition of matrix metalloproteinases expression in mouse skin," *Life Sciences*, vol. 217, pp. 110–118, Jan. 2019.

[118] C. Pan, Y. Si, Q. Meng, L. Jing, L. Chen, Y. Zhang, and H. Bao, "Suppression of the RAC1/MLK3/p38 Signaling Pathway by $\beta$-Elemene Alleviates Sepsis-Associated Encephalopathy in Mice," *Frontiers in Neuroscience*, vol. 13, p. 358, 2019.

[119] C.-P. Zheng, X.-M. Tong, H.-P. Yao, J. Yang, J. Xu, X.-P. Cai, and Z. Liu, "[beta-elemene enhances aclarubicin-induced apoptotic effect in HL-60 cells and its mechanism.]," *Zhonghua Xue Ye Xue Za Zhi = Zhonghua Xueyexue Zazhi*, vol. 30, pp. 821–824, Dec. 2009.

[120] J. Liu, Z. Zhang, J. Gao, J. Xie, L. Yang, and S. Hu, "Downregulation effects of beta-elemene on the levels of plasma endotoxin, serum TNF-alpha, and hepatic CD14 expression in rats with liver fibrosis," *Frontiers of Medicine*, vol. 5, pp. 101–105, Mar. 2011.

[121] C.-Y. Zhang, J.-Y. Zhu, Y. Ye, M. Zhang, L.-J. Zhang, S.-J. Wang, Y.-N. Song, and H. Zhang, "Erhuang Formula ameliorates renal damage in adenine-induced chronic renal failure rats via inhibiting inflammatory and fibrotic responses," *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie*, vol. 95, pp. 520–528, Nov. 2017.

[122] R. T. Corlett, "Plant diversity in a changing world: Status, trends, and conservation needs," *Plant Diversity*, vol. 38, pp. 10–16, May 2016.

[123] V. Čižinauskas, N. Elie, A. Brunelle, and V. Briedis, "Fatty acids penetration into human skin ex vivo: A TOF-SIMS analysis approach," *Biointerphases*, vol. 12, no. 1, p. 011003, 2017.

[124] W. Luo, M. S. Friedman, K. Shedden, K. D. Hankenson, and P. J. Woolf, "GAGE: generally applicable gene set enrichment for pathway analysis," *BMC Bioinformatics*, vol. 10, p. 161, May 2009.

[125] W. Luo, G. Pant, Y. K. Bhavnasi, S. G. Blanchard, and C. Brouwer, "Pathview Web: user friendly pathway visualization and data integration," *Nucleic Acids Research*, vol. 45, pp. W501–W508, July 2017. Publisher: Oxford Academic.

[126] A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, and J. P.

Mesirov, "Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles," *Proceedings of the National Academy of Sciences*, vol. 102, pp. 15545–15550, Oct. 2005.

[127] S.-Y. Kim and D. J. Volsky, "PAGE: Parametric Analysis of Gene Set Enrichment," *BMC Bioinformatics*, vol. 6, p. 144, June 2005.

[128] N. H. Freese, D. C. Norris, and A. E. Loraine, "Integrated genome browser: visual analytics platform for genomics," *Bioinformatics (Oxford, England)*, vol. 32, no. 14, pp. 2089–2095, 2016.

[129] B. Zhang and S. Horvath, "A general framework for weighted gene co-expression network analysis," *Statistical Applications in Genetics and Molecular Biology*, vol. 4, p. Article17, 2005.

[130] P. Langfelder and S. Horvath, "WGCNA: an R package for weighted correlation network analysis," *BMC Bioinformatics*, vol. 9, p. 559, Dec. 2008.

[131] C. S. Greene, A. Krishnan, A. K. Wong, E. Ricciotti, R. A. Zelaya, D. S. Himmelstein, R. Zhang, B. M. Hartmann, E. Zaslavsky, S. C. Sealfon, D. I. Chasman, G. A. FitzGerald, K. Dolinski, T. Grosser, and O. G. Troyanskaya, "Understanding multicellular function and disease with human tissue-specific networks," *Nature Genetics*, vol. 47, pp. 569–576, June 2015.

[132] L. M. Franco, M. Gadkari, K. N. Howe, J. Sun, L. Kardava, P. Kumar, S. Kumari, Z. Hu, I. D. C. Fraser, S. Moir, J. S. Tsang, and R. N. Germain, "Immune regulation by glucocorticoids can be linked to cell type-dependent transcriptional responses," *The Journal of Experimental Medicine*, vol. 216, pp. 384–406, Feb. 2019.

[133] E. Ben-Chetrit, S. Bergmann, and R. Sood, "Mechanism of the anti-inflammatory effect of colchicine in rheumatic diseases: a possible new outlook through microarray analysis," *Rheumatology (Oxford, England)*, vol. 45, pp. 274–282, Mar. 2006.

[134] A. Matejuk, A. A. Vandenbark, G. G. Burrows, B. F. Bebo, and H. Offner, "Reduced Chemokine and Chemokine Receptor Expression in Spinal Cords of TCR BV8S2 Transgenic Mice Protected Against Experimental Autoimmune Encephalomyelitis with BV8S2 Protein," *The Journal of Immunology*, vol. 164, pp. 3924–3931, Apr. 2000. Publisher: American Association of Immunologists Section: Clinical Immunology.

[135] M. Baggiolini, B. Moser, and I. Clark-Lewis, "Interleukin-8 and Related Chemotactic Cytokines: The Giles Filley Lecture," *CHEST*, vol. 105, pp. 95S–98S, Mar. 1994. Publisher: Elsevier.

[136] J.-M. Zhang and J. An, "Cytokines, Inflammation and Pain," *International anesthesiology clinics*, vol. 45, no. 2, pp. 27–37, 2007.

[137] C. Garlanda and S. Jaillon, "The Interleukin-1 Family," in *Encyclopedia of Immunobiology* (M. J. H. Ratcliffe, ed.), pp. 438–446, Oxford: Academic Press, Jan. 2016.

[138] R. Kowsar, B. Keshtegar, and A. Miyamoto, "Understanding the hidden relations between pro- and anti-inflammatory cytokine genes in bovine oviduct epithelium using a multilayer response surface method," *Scientific Reports*, vol. 9, p. 3189, Feb. 2019. Number: 1 Publisher: Nature Publishing Group.

[139] H. Mühl, "Pro-Inflammatory Signaling by IL-10 and IL-22: Bad Habit Stirred Up by Interferons?," *Frontiers in Immunology*, vol. 4, 2013. Publisher: Frontiers.

[140] F. N. Lauw, D. Pajkrt, C. E. Hack, M. Kurimoto, S. J. H. v. Deventer, and T. v. d. Poll, "Proinflammatory Effects of IL-10 During Human Endotoxemia," *The Journal of Immunology*, vol. 165, pp. 2783–2789, Sept. 2000. Publisher: American Association of Immunologists Section: Inflammation.

[141] T. Matsumura, K. Kugiyama, S. Sugiyama, M. Ohgushi, K. Amanaka, M. Suzuki, and H. Yasue, "Neutral endopeptidase 24.11 in neutrophils modulates protective effects of natriuretic peptides against neutrophils-induced endothelial cytotoxity.," *Journal of Clinical Investigation*, vol. 97, pp. 2192–2203, May 1996.

[142] M. Yoo, J. Shin, J. Kim, K. A. Ryall, K. Lee, S. Lee, M. Jeon, J. Kang, and A. C. Tan, "DSigDB: drug signatures database for gene set analysis," *Bioinformatics*, vol. 31, pp. 3069–3071, Sept. 2015.

[143] X. Han, D. Rodriguez, and T. L. Parker, "Biological activities of frankincense essential oil in human dermal fibroblasts," *Biochimie Open*, vol. 4, pp. 31–35, Feb. 2017.

[144] R. Tokunaga, W. Zhang, M. Naseem, A. Puccini, M. D. Berger, S. Soni, M. McSkane, H. Baba, and H.-J. Lenz, "CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation - A target for novel cancer therapy," *Cancer Treatment Reviews*, vol. 63, pp. 40–47, Feb. 2018.

[145] E. Y. Lee, Z.-H. Lee, and Y. W. Song, "CXCL10 and autoimmune diseases," *Autoimmunity Reviews*, vol. 8, pp. 379–383, Mar. 2009.

[146] J. Vinet, E. K. de Jong, H. W. G. M. Boddeke, V. Stanulovic, N. Brouwer, I. Granic, U. L. M. Eisel, R. S. B. Liem, and K. Biber, "Expression of CXCL10 in cultured cortical neurons," *Journal of Neurochemistry*, vol. 112, pp. 703–714, Feb. 2010.

[147] M. Kanehisa and S. Goto, "KEGG: kyoto encyclopedia of genes and genomes," *Nucleic Acids Research*, vol. 28, pp. 27–30, Jan. 2000.

[148] H. Wickham, R. François, L. Henry, and K. Muller, "dplyr: A Grammar of Data Manipulation," 2020.

[149] C. Fraley, A. E. Raftery, B. T. Murphy, and L. Scrucca, "mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation," p. 57.

[150] H. Wickham, J. Hester, and R. Francois, "readr: Read Rectangular Text Data," 2018.

[151] H. Wickham, "The Split-Apply_combine Strategy for Data Analysis," *The Journal of Statistical Software*, vol. 40, no. 1, pp. 1–29, 2011.

[152] L. Scrucca, M. Fop, T. Brendan Murphy, and A. Raftery, E, "mclust 5: clustering, classification and density estimation using Gausian finite mixture models," *The R Journal*, vol. 8, no. 1, pp. 289–317, 2016.

[153] V. Vu, "ggbiplot: A ggplot2 based biplot," 2011.

[154] H. Wickham, J. Hester, and W. Chang, "devtools: Tools to Make Developing R Packages Easier," 2020.

[155] The Angiosperm Phylogeny Group, "An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV," *Botanical Journal of the Linnean Society*, vol. 181, pp. 1–20, May 2016.

[156] Qiagen, "Ingenuity Pathway Analysis."

[157] I. G. Winkler, V. Barbier, R. Wadley, A. C. W. Zannettino, S. Williams, and J.-P. Lévesque, "Positioning of bone marrow hematopoietic and stromal cells relative to blood flow in vivo: serially reconstituting hematopoietic stem cells reside in distinct nonperfused niches," *Blood*, vol. 116, pp. 375–385, July 2010. Publisher: American Society of Hematology.

[158] K. V. Korneev, K.-S. N. Atretkhany, M. S. Drutskaya, S. I. Grivennikov, D. V. Kuprash, and S. A. Nedospasov, "TLR-signaling and proinflammatory cytokines as drivers of tumorigenesis," *Cytokine*, vol. 89, pp. 127–135, Jan. 2017.

[159] L. O. Orafidiya, "The effect of autoxidation of lemon-grass oil on its antibacterial activity," *Phytotherapy Research*, vol. 7, no. 3, pp. 269–271, 1993. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ptr.2650070315.

[160] N. Kishore, J. P. N. Chansouria, and N. K. Dubey, "Antidermatophytic Action of the Essential Oil of Chenopodium ambrosioides and an Ointment Prepared from it," *Phytotherapy Research*, vol. 10, no. 5, pp. 453–455, 1996. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291099-1573%28199608%2910%3A5%3C453%3A%3AAID-PTR874%3E3.0.CO%3B2-A.

[161] P. L. Teissedre and A. L. Waterhouse, "Inhibition of Oxidation of Human Low-Density Lipoproteins by Phenolic Substances in Different Essential Oils Varieties," *Journal of Agricultural and Food Chemistry*, vol. 48, pp. 3801–3805, Sept. 2000. Publisher: American Chemical Society.

[162] G. Deza, B. García-Bravo, J. F. Silvestre, M. A. Pastor-Nieto, R. González-Pérez, F. Heras-Mendaza, P. Mercader, V. Fernández-Redondo, B. Niklasson, A. M. Giménez-Arnau, and GEIDAC, "Contact sensitization to limonene and linalool hydroperoxides in Spain: a GEIDAC* prospective study," *Contact Dermatitis*, vol. 76, pp. 74–80, Feb. 2017.

[163] A. Swerdlin, D. Rainey, and F. J. Storrs, "Fragrance mix reactions and lime allergic contact dermatitis," *Dermatitis: Contact, Atopic, Occupational, Drug*, vol. 21, pp. 214–216, Aug. 2010.

[164] A. Sarkic and I. Stappen, "Essential Oils and Their Single Compounds in Cosmetics—A Critical Review," *Cosmetics*, vol. 5, p. 11, Mar. 2018. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

[165] B. J. Aungst, "Structure/effect studies of fatty acid isomers as skin penetration enhancers and skin irritants," *Pharmaceutical Research*, vol. 6, pp. 244–247, Mar. 1989.

[166] "What Is Fractionated Coconut Oil?," May 2018.

[167] J. M. Viljoen, A. Cowley, J. du Preez, M. Gerber, and J. du Plessis, "Penetration enhancing effects of selected natural oils utilized in topical dosage forms," *Drug Development and Industrial Pharmacy*, vol. 41, no. 12, pp. 2045–2054, 2015.

[168] E. Esposito, M. Sguizzato, M. Drechsler, P. Mariani, F. Carducci, C. Nastruzzi, and R. Cortesi, "Progesterone lipid nanoparticles: Scaling up and in vivo human study," *European Journal of Pharmaceutics and Biopharmaceutics: Official Journal of Arbeitsgemeinschaft Fur Pharmazeutische Verfahrenstechnik e.V*, vol. 119, pp. 437–446, Oct. 2017.

[169] P. Tiwari, T. Nagatake, S.-i. Hirata, K. Sawane, A. Saika, Y. Shibata, S. Morimoto, T. Honda, J. Adachi, Y. Abe, J. Isoyama, T. Tomonaga, H. Kiyono, K. Kabashima, and J. Kunisawa, "Dietary coconut oil ameliorates skin contact hypersensitivity through mead acid production in mice," *Allergy*, vol. 74, no. 8, pp. 1522–1532, 2019. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/all.13762.

[170] S. R. Varma, T. O. Sivaprakasam, I. Arumugam, N. Dilip, M. Raghuraman, K. B. Pavan, M. Rafiq, and R. Paramesh, "In vitro anti-inflammatory and skin protective properties of Virgin coconut oil," *Journal of Traditional and Complementary Medicine*, vol. 9, pp. 5–14, Jan. 2019.

[171] S. H. Sicherer, T. J. Furlong, A. Muñoz-Furlong, A. W. Burks, and H. A. Sampson, "A voluntary registry for peanut and tree nut allergy: characteristics of the first 5149 registrants," *The Journal of Allergy and Clinical Immunology*, vol. 108, pp. 128–132, July 2001.

[172] K. Anagnostou, "Coconut Allergy Revisited," *Children*, vol. 4, Sept. 2017.

[173] A. Lyttle, "A Caution Against Coconut Oil," Apr. 2017.

[174] E. T. Grant, "7 Little-Known Risks Of Using Coconut Oil."

[175] K. Martin, "No, You Should Not Use Coconut Oil On Your Face."

[176] R. Notman, W. K. den Otter, M. G. Noro, W. J. Briels, and J. Anwar, "The Permeability Enhancing Mechanism of DMSO in Ceramide Bilayers Simulated by Molecular Dynamics," *Biophysical Journal*, vol. 93, pp. 2056–2068, Sept. 2007.

[177] C. A. Kirker-Head and H. Feldmann, "23 - Pharmacotherapy of joint and tendon disease," in *Equine Sports Medicine and Surgery (Second Edition)* (K. W. Hinchcliff, A. J. Kaneps, and R. J. Geor, eds.), pp. 473–502, W.B. Saunders, Jan. 2014.

[178] J. H. Thrall, "Nanotechnology and Medicine," *Radiology*, vol. 230, pp. 315–318, Feb. 2004. Publisher: Radiological Society of North America.

[179] P. Nasimi and M. Haidari, "Medical Use of Nanoparticles: Drug Delivery and Diagnosis Diseases," *International Journal of Green Nanotechnology*, vol. 1, Jan. 2013.

[180] C. Sharma, J. M. Al Kaabi, S. M. Nurulain, S. N. Goyal, M. A. Kamal, and S. Ojha, "Polypharmacological Properties and Therapeutic Potential of Beta-Caryophyllene: A Dietary Phytocannabinoid of Pharmaceutical Promise," *Current Pharmaceutical Design*, vol. 22, no. 21, pp. 3237–3264, 2016.

[181] M. K. Swamy, M. S. Akhtar, and U. R. Sinniah, "Antimicrobial Properties of Plant Essential Oils against Human Pathogens and Their Mode of Action: An Updated Review," *Evidence-Based Complementary and Alternative Medicine: eCAM*, vol. 2016, p. 3012462, 2016.

[182] E. A. Grice and J. A. Segre, "The skin microbiome," *Nature reviews. Microbiology*, vol. 9, pp. 244–253, Apr. 2011.

[183] "Which Is the Best PostgreSQL GUI? 2019 Comparison," Sept. 2019.

[184] N. Mangalmurti and C. A. Hunter, "Cytokine Storms: Understanding COVID-19," *Immunity*, vol. 53, pp. 19–25, July 2020.

[185] K. E. Remy, M. Mazer, D. A. Striker, A. H. Ellebedy, A. H. Walton, J. Unsinger, T. M. Blood, P. A. Mudd, D. J. Yi, D. A. Mannion, D. F. Osborne, R. S. Martin, N. J. Anand, J. P. Bosanquet, J. Blood, A. M. Drewry, C. C. Caldwell, I. R. Turnbull, S. C. Brakenridge, L. L. Moldwawer, and R. S. Hotchkiss, "Severe immunosuppression and not a cytokine storm characterizes COVID-19 infections," *JCI Insight*, vol. 5, Sept. 2020. Publisher: American Society for Clinical Investigation.

[186] P. Sinha, M. A. Matthay, and C. S. Calfee, "Is a "Cytokine Storm" Relevant to COVID-19?," *JAMA Internal Medicine*, June 2020.

# APPENDIX A: GITHUB

The first repository (dissertation research) contains sub directories broken up by chapter. Each sub-directory contains the codes used for different aspects of the project. https://github.com/aneetauppal/DissertationResearch

The LPS study repository contains the codes used in chapters 3 and 5 for RT-qPCR analysis and the RNA-sequencing assembly. https://github.com/aneetauppal/LPSstudy