

“To Tea or Not to Tea”

Reanalysis and reinterpretation of NCBI GEO publicly available dataset **GSE124161** by primary author Sara Ud-Din, et al. short title “*Zonal and Direct Topical EGCG for Skin Scarring.*”

By:

*Brandon Thong
and Dawn Layman*

2. Table of Contents:

3. Abstract:	3
4. Statement of Research Question:	3
5. Hypothesis:	3
6. Background:	3
7. Motivation for Problem Addressed:	4
8. Results:	4
Week 1 Zonal Application: EGCG Treatment versus Placebo Control:	4
Figure 1: Volcano Plot of Week 1 EGCG Treatment versus Placebo.	5
Figure 2: Hierarchical Clustering and Silhouette Plot for Selection of “k” Groups.	5
Figure 3a-h: GO Term and KEGG Pathway Enrichment Analysis Week 1 EGCG Treatment versus Placebo.	7
Interaction: Week 1 vs Week 2: Exploration of EGCG Treatment Effects Over Time:	8
Figure 4: Volcano Plot of EGCG Treatment Across Week 1 to Week 2.	8
Figure 5: Hierarchical Clustering and Silhouette Plot for Selection of “k” Groups.	9
Figure 6a-h: GO Term and KEGG Pathway Enrichment Analysis EGCG Treatment Across Week 1 to Week 2.	10
Interaction5: “Global Differential Response” Week 1 vs Week 2:	10
Figure 7: Volcano Plot “Global Differential Response” Week 1 vs Week 2.	11
Figure 8: Hierarchical Clustering and Silhouette Plot for Selection of “k” Groups.	11
Figure 9a-h: GO Term and KEGG Pathway Enrichment Analysis “Global Differential Response” Week 1 vs Week 2.	12
9. Interpretation of Results:	12
10. Discussion/Conclusion:	14
11. Limitations and Alternative Interpretations:	15
12. Future Work:	15
Supplemental Figure S1.....	17
Design Matrix and Contrasts for LIMMA and edgeR Differential Expression Analysis.....	17
Supplemental Figure S1(continued).....	18
Design Matrix and Contrasts for LIMMA and edgeR Differential Expression Analysis.....	18
Supplemental Table S2.....	19
FDR Values of Comparisons/Contrasts Created in LIMMA.....	19
References:	20

3. Abstract:

Epigallocatechin-3-gallate, or EGCG, is a tea-derived compound attributed to having anti-inflammatory, anti-angiogenic, and antioxidant properties that promote wound healing. In this study, we investigated the presumed effects of EGCG on improving scar healing measured by differential gene expression for both Zonal and Direct treatment. This was accomplished by filtering and preprocessing raw reads and performing a limma analysis workflow to obtain differentially expressed genes (DEGs) for several interactions: EGCG vs Control, EGCG effects between Week 1 vs Week 2, and Global Differential Response for Week 1 vs Week 2. For each interaction, biologically significant genes were selected from the DEGs with cutoffs of an FDR=0.2 and Log2FoldChange=1.5, clustered, and were further analyzed after generating heatmaps, GO Term, and KEGG pathway enrichment graphs. In our analysis, we discovered that although results were not significant for the direct treatment weeks, zonal analysis still provided significant gene results. We found that Zonal EGCG application downregulated immune and inflammation responses, and upregulated cellular remodeling, cycle, and apoptosis processes amongst others that were also linked to improved healing outcomes. In conclusion, our study highlights the beneficial effects on scarring outcomes when using zonal EGCG treatment early in the healing process, which may lead to future advancements and insights regarding wound healing and scar treatments.

4. Statement of Research Question:

Does Zonal Priming or Direct Treatment with EGCG improve outcomes for scar-free or minimized scar healing, and can such improvements be captured and identified via differential gene expression analysis?

5. Hypothesis:

Epigallocatechin-3-gallate, EGCG, has been attributed to having anti-inflammatory, anti-angiogenic, and antioxidant properties that promote wound healing. This project investigates the presumed effect of “Direct Application” of topical EGCG compared with “Zonal Priming” on improved outcome of scar healing, as measured by differential gene expression.

6. Background:

Plant-based foods are rich in polyphenols, which are plant derived chemicals that have been linked to many health benefits.¹ Epigallocatechin-3-gallate, abbreviated “EGCG”, is the most abundant polyphenol found in green tea,² and has been linked to anti-inflammatory,^{3,4} antioxidant, and anti-angiogenic activities,⁵ in addition to promoting wound healing and repair processes.⁶

When wound healing processes fail and become disorganized, incomplete healing can occur, leading to difficult to heal wounds and abnormal cutaneous wound repair after injury, which can lead to scarring.^{6,7} Scarring is said to occur on a spectrum,⁸ with injury type, severity, location, genetics, age, and a multitude of other risk factors all contributing to the magnitude of scar formation and presentation.⁹

Multiple studies have linked EGCG as a useful bioactive in preventing tissue destruction, promoting wound healing, and supporting repair processes.⁶ Additionally, the effects of early intervention and immediate treatment in wound care have been implemented to prevent poor scarring outcomes.⁹

7. Motivation for Problem Addressed:

The investigation of the effects of topical EGCG treatment with insights to Zonal Priming (early) versus Direct (late) application on outcomes of scar healing and the underlying mechanisms of healing and repair, as measured by differential gene expression, can provide valuable insights and advancements regarding wound healing and the development of successful scar treatments.

8. Results:

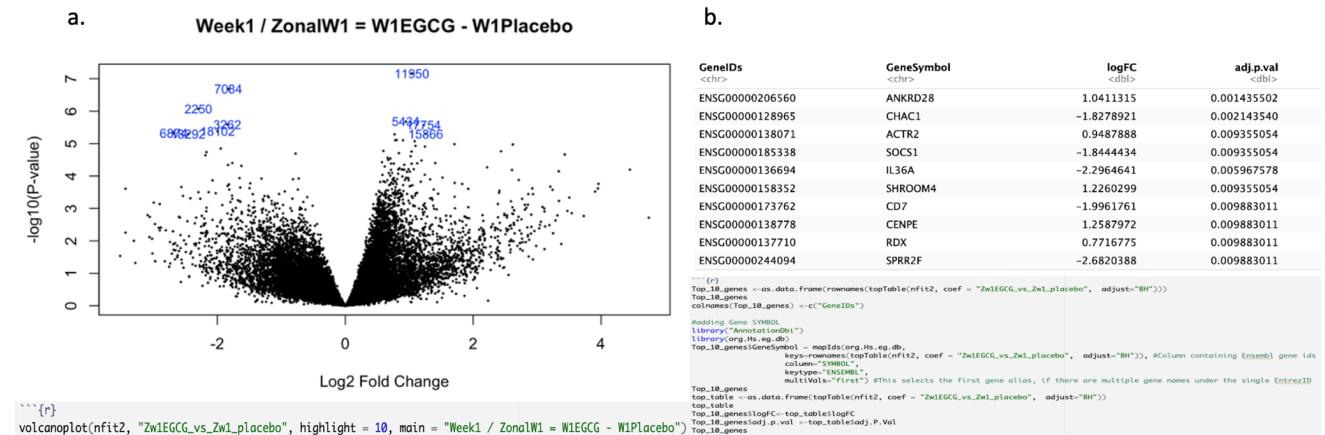
Please refer to Supplemental Figure S1 for limma design matrix and contrasts created to these analyses.

Week 1 Zonal Application: EGCG Treatment versus Placebo Control:

Week 1 was the only week where EGCG treatment proved to be significantly different from the placebo control in all 8 weeks compared (see Supplemental Table S2).

Volcano plot of the normalized gene data was compiled (Figure 1a), and the top 10 most differentially expressed genes (DEG) were selected for significance by FDR (Figure 1b). The top 10 DEG's identified as affected by EGCG treatment are involved in immunity; inflammation; cellular structure, shape, and motility; cell division; and oxidation. IL36A, a gene involved in inflammation, was similarly identified in the paper as being significantly down-regulated.⁸ The code utilized to generate the following plots and tables are included below the respective images.

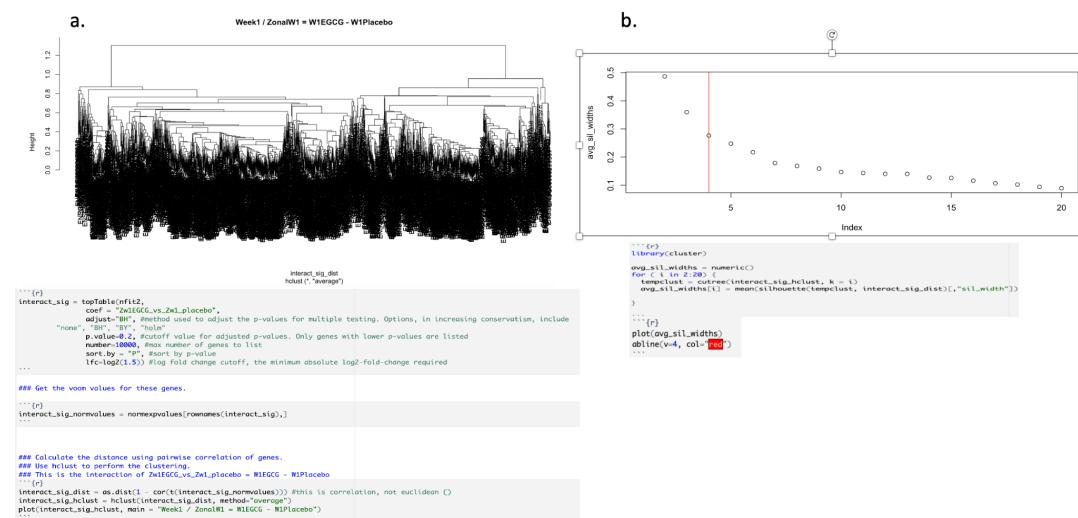
Figure 1: Volcano Plot of Week 1 EGCG Treatment versus Placebo.



Further exploration of Week 1 EGCG treatment for biological relevance utilized hierarchical clustering based on distance as correlation, with group selection determined via Silhouette Plot (Figure 2a+b), based on cutoffs for FDR=0.2 and L2FC =1.5.

Cutoffs were determined based on capturing enough genes for analyzing the interaction of the effect of treatment of EGCG over time from week 1 to week 2, and also to measure the global effect of the difference in differential gene expression (DGE) from week 1 and week 2. Therefore, the use of a silhouette plot for cluster analysis was necessary.

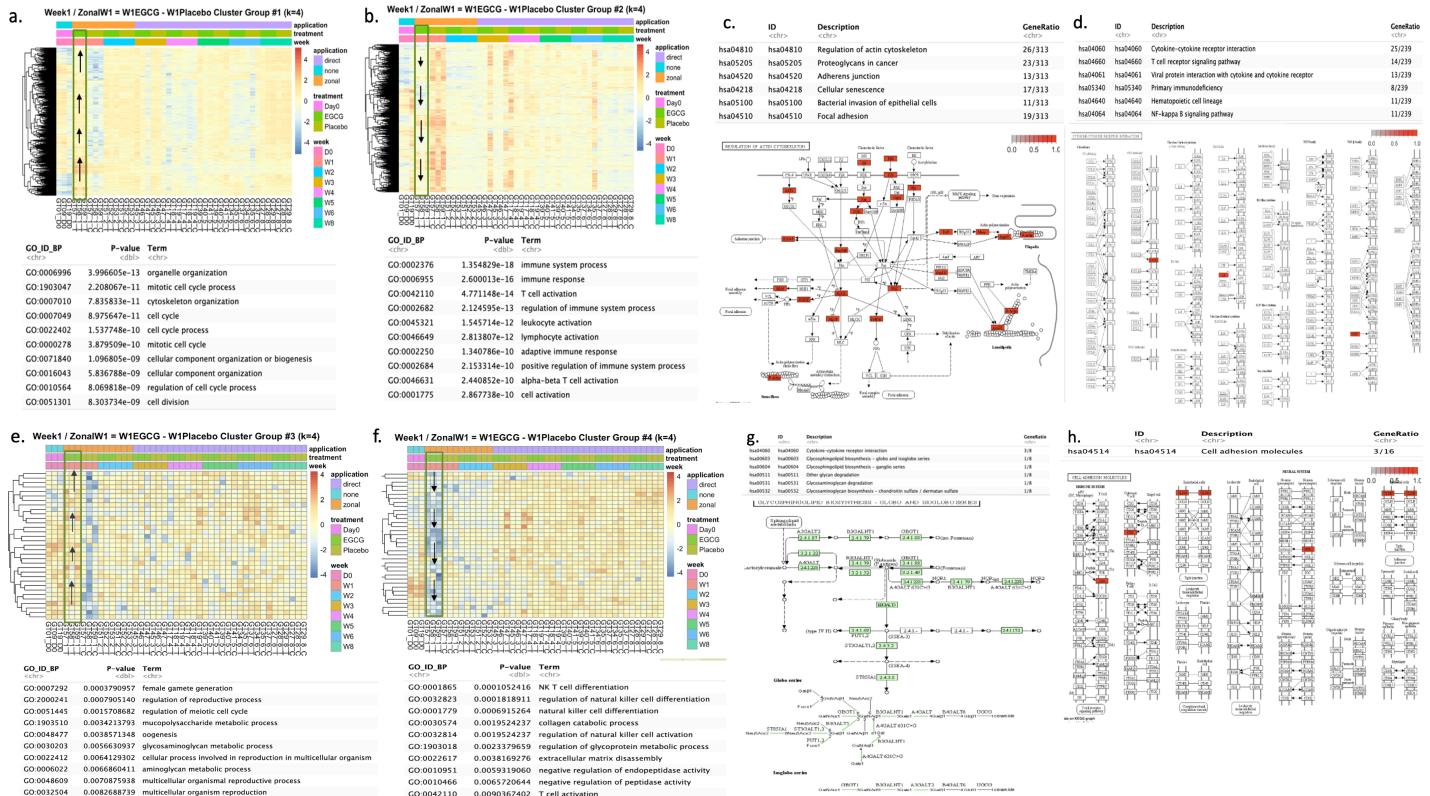
Figure 2: Hierarchical Clustering and Silhouette Plot for Selection of “k” Groups.



GO term enrichment and KEGG pathway analysis (Figure 3a-h) identified upregulated biological processes involved in cell cycle and cellular organization and certain metabolic processes (a,e), and identified a downregulation in biological processes involving the immune system, enzymatic activity, and ECM remodeling (b,f). KEGG Pathways identified with the highest statistical significance and most gene membership involved the “regulation of the actin cytoskeleton”(c), followed by “cytokine-cytokine receptor interaction”(d). A representative example of the critical code utilized to generate the following heatmaps, GO-term enrichment, and KEGG pathways, are included below the respective images.

This Space Left Intentionally Blank for Formatting

Figure 3a-h: GO Term and KEGG Pathway Enrichment Analysis Week 1 EGCG Treatment versus Placebo.



Create the groups. Notice the result is actually a vector of number and the gene names are the labels.

```
```{r}
interact_sig_hclust_4 = cutree(interact_sig_hclust, k=4)

To get the gene names that are in the different groups, use the which command to find out which genes are in the different groups, but then use the names function to get the actual names.

##```

```

```
```{r}
interact_sig_hclust_g1_normexpvalues[names(which(interact_sig_hclust_4==1))],]
interact_sig_hclust_g2_normexpvalues[names(which(interact_sig_hclust_4==2))],]
interact_sig_hclust_g3_normexpvalues[names(which(interact_sig_hclust_4==3))],]
interact_sig_hclust_g4_normexpvalues[names(which(interact_sig_hclust_4==4))],]
```

Create heatmap of each cluster group

Cluster#1

Use "phewmap" to draw cluster. "annot_col" defines how to create the legend. "scale" allows us to see the pattern for each gene. To make it easier to compare the different groups, I asked the columns not to be clustered "cluster_cols = F", and to not show the gene names "show_rownames = F".

```
```{r}
library(phewmap)

annotation <- as.data.frame(cbind(pheno_df$week, pheno_df$treatments, pheno_df$application))
colnames(annotation) <- c("Week", "Treatment", "Application")
rownames(annotation) <- pheno_df$count.colnames

phewmap(interact_sig_hclust_g1_annotation_col = annotation, scale="row", cluster_cols = F, show_rownames = F, main = "Week1 / ZonalW1 = W1EGCG - W1Placebo Cluster Group #1 (k=4)")
```

```
Perform Go-Term Enrichment analysis
(r)
Load the proper packages
library(GOstats)
library(org.Hs.eg.db)
library(Category)
library(org.Hs.eg.db)
See how to perform Part 1
Create Systematic Name

Converting the Incomplete to Express was achieved with this code:
https://www.biostars.org/p/1386/
#(r)
library("Annotation")
Load the proper packages
library(GOstats)
library(org.Hs.eg.db)
library(Category)
library(org.Hs.eg.db)
See how to perform Part 1

Load the proper packages
library(GOstats)
library(org.Hs.eg.db)
library(Category)
library(org.Hs.eg.db)
See how to perform Part 1

dFfreqgenes_names_difftrentrez<- medley$dfFreqgenes[,c("EntrezID", "Category", "Treatment", "Week", "Application")]

dFfreqgenes_names_difftrentrez[["EntrezID"]]<- as.character(dFfreqgenes_names_difftrentrez[["EntrezID"]])
dFfreqgenes_names_difftrentrez[["Treatment"]]<- as.factor(dFfreqgenes_names_difftrentrez[["Treatment"]])
dFfreqgenes_names_difftrentrez[["Week"]]<- as.factor(dFfreqgenes_names_difftrentrez[["Week"]])
dFfreqgenes_names_difftrentrez[["Application"]]<- as.factor(dFfreqgenes_names_difftrentrez[["Application"]])

dFfreqgenes_names_difftrentrez[["EntrezID"]]<- as.character(dFfreqgenes_names_difftrentrez[["EntrezID"]])
dFfreqgenes_names_difftrentrez[["Treatment"]]<- as.factor(dFfreqgenes_names_difftrentrez[["Treatment"]])
dFfreqgenes_names_difftrentrez[["Week"]]<- as.factor(dFfreqgenes_names_difftrentrez[["Week"]])
dFfreqgenes_names_difftrentrez[["Application"]]<- as.factor(dFfreqgenes_names_difftrentrez[["Application"]])

dFfreqgenes_names_difftrentrez[["EntrezID"]]<- as.character(dFfreqgenes_names_difftrentrez[["EntrezID"]])
dFfreqgenes_names_difftrentrez[["Treatment"]]<- as.factor(dFfreqgenes_names_difftrentrez[["Treatment"]])
dFfreqgenes_names_difftrentrez[["Week"]]<- as.factor(dFfreqgenes_names_difftrentrez[["Week"]])
dFfreqgenes_names_difftrentrez[["Application"]]<- as.factor(dFfreqgenes_names_difftrentrez[["Application"]])
```

```
KEGG ENRICHMENT Part1
#(r)
Install libraries needed for KEGG Enrichment Analysis
library(clusterProfiler)
library(pathview)
library(ggplot2)
library(ggpubr)

##Now perform KEGG ENRICHMENT

keggEnrich <- enrichKEGG(
 diffExpGenes_names_difftrentrez,
 organism = "hsa",
 keyType = "kegg",
 pvalueCutoff = 0.05, #adjust this if you are not seeing any results
 pAdjustMethod = "BH",
 ...)

##```
Show results from enrichKEGG
keggEnrich
##```
##Generate a graph for the two KEGG results
##Edit the pathway id to that which is appropriate based on the ID column from the enrichKEGG output
#These will generate images that will be saved to the working directory or the downloads folder
##Repeat for however many results you get from keggEnrich

p.out.html0 <- pathview(gene.data = diffExpGenes.names_difftrentrez, pathway.id = "hsa04810", species = "hsa")
##Repeat for the second result
p.out.html1 <- pathview(gene.data = diffExpGenes.names_difftrentrez, pathway.id = "hsa05205", species = "hsa")
##```
##Also show the genes involved in the pathway
#These correspond to the elements included in the image of the KEGG pathway generated earlier
p.v.out.html1$plot.data$gene
p.v.out.html1$plot.data$gene
```

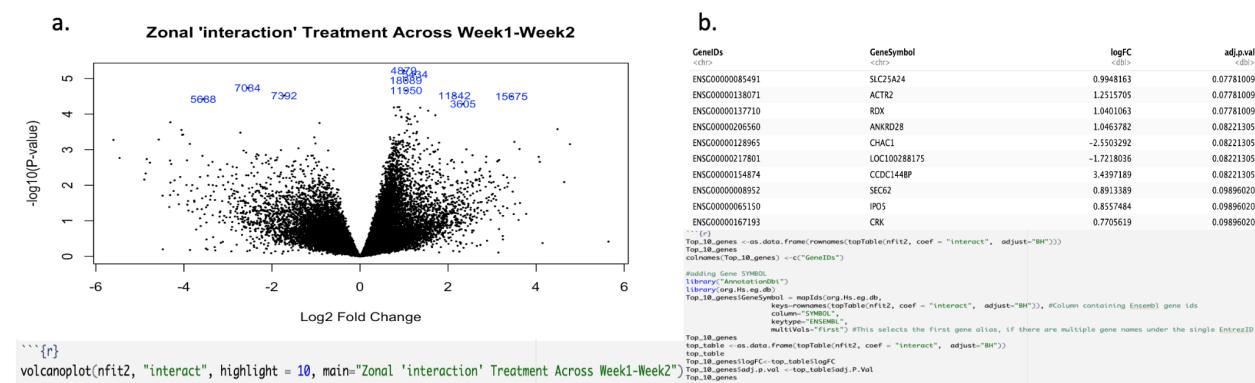
## **Interaction: Week 1 vs Week 2: Exploration of EGCG Treatment Effects Over Time:**

The effect of EGCG treatment from week 1 to week 2 was explored, as treatment with EGCG versus the placebo was effective in week 1, and the interaction of continued treatment into week 2 was assessed.

A volcano plot of the normalized gene data was compiled (Figure 4a), and the top 10 most differentially expressed genes (DEG) were selected for significance by FDR (Figure 4b).

The top 10 DEG's identified as affected by EGCG treatment from week 1 to week 2 are involved in energy, wound healing, ECM remodeling, and protein translocation. Four genes: ACTR2, RDX, ANKRD28, and CHAC1 were common to the week 1 EGCG vs Placebo analysis, and were retained and maintained significance with treatment emerging into week 2. Additionally, the up and down regulation of gene expression for each of the genes were maintained. The code utilized to generate the following plots and tables are included below the respective images.

**Figure 4: Volcano Plot of EGCG Treatment Across Week 1 to Week 2.**

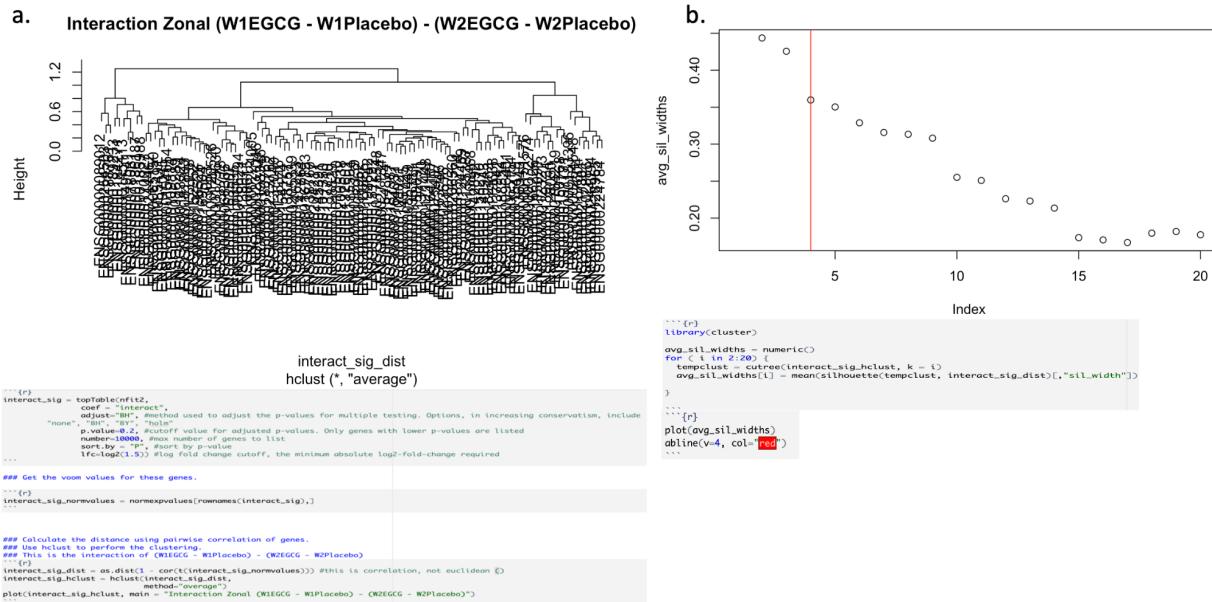


Further exploration of the effect of treatment with EGCG from week 1 to week 2 for biological relevance utilized hierarchical clustering based on distance as correlation, cutoffs utilized for an FDR=0.2 and L2FC =1.5, with group selection determined via Silhouette Plot (Figure 5a+b).

FDR and L2FC cutoffs were determined based on capturing enough genes for analyzing the interaction of the effect of treatment of EGCG over time from week 1 to week 2, as evidenced by the below hierarchical clustering tree. (Figure 5a) The cutoffs needed to be generous, so enough gene membership was captured for analysis of the interaction. We employed the use of the silhouette plot to determine the ideal group selection for clustering, and determined k=4 groups, based on the number of genes

captured. Other interesting clusters appear to be at 9, and 14, but that selection would prove to be difficult, as there was not enough gene membership in each cluster to continue with GO term Enrichment and KEGG Pathway analysis.

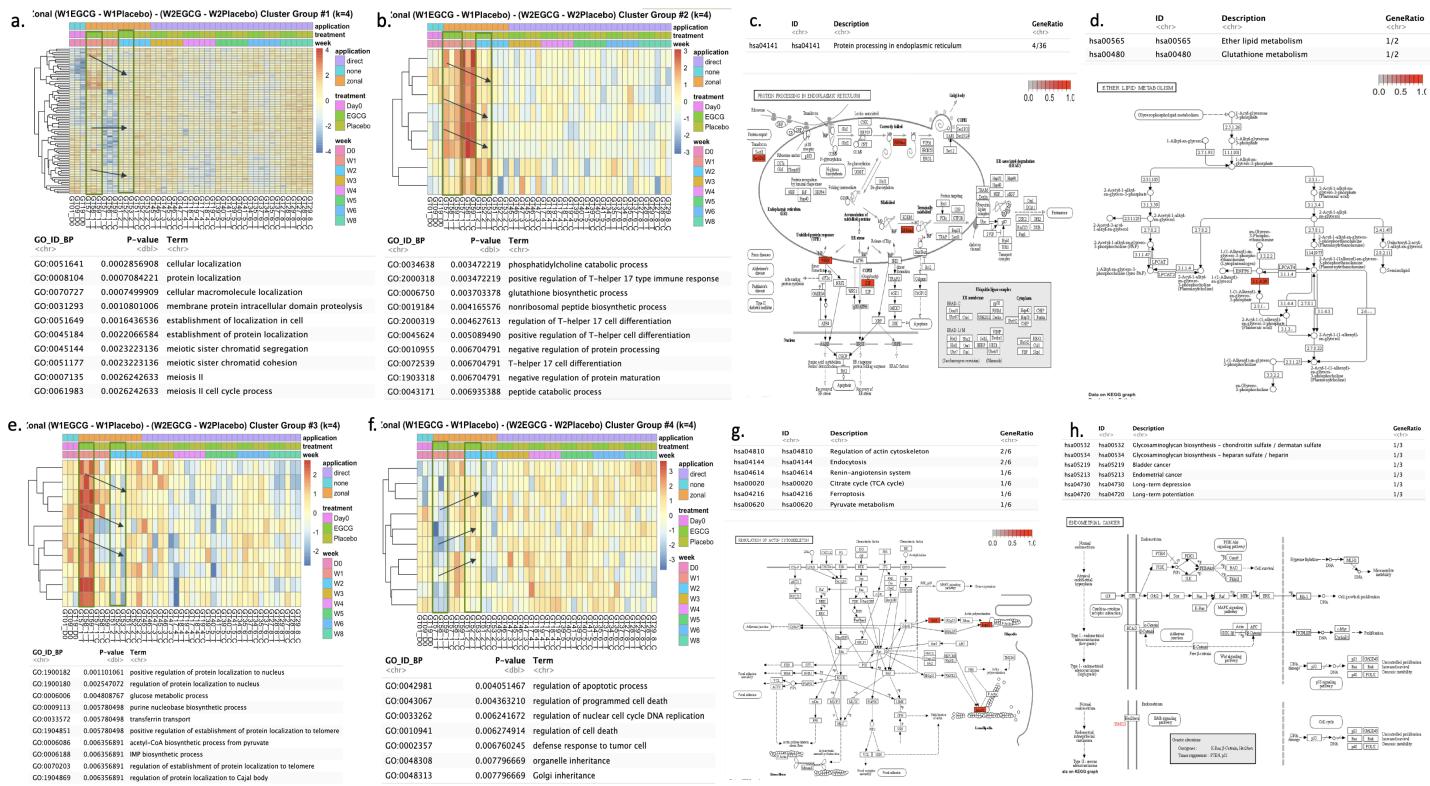
**Figure 5: Hierarchical Clustering and Silhouette Plot for Selection of “k” Groups.**



GO term enrichment and KEGG pathway analysis (Figure 6a-h) identified down-regulated biological processes involved in “cellular and protein localization” and “biosynthesis” (a,e), and identified upregulation in biological processes involving “protein regulation, apoptosis, and cell death” (b,f). KEGG Pathways identified with the highest statistical significance and most gene membership involved the “protein processing in the endoplasmic reticulum”(c). Other significant KEGG pathways identified had low gene membership and will be given less weight to interpretation. (d,g,h)

Critical code regarding the GO term enrichment and KEGG analysis are the same as described previously, so will not be described due to redundancy. The “interact” term (Figure 5a, code), called in the “coeff =” line, calls the proper coefficient for cluster analysis when building the hierarchical tree, which selects the proper contrast for the data and subsequent clustering. Please refer to Supplemental Figure S1 for limma design matrix and contrasts created for these analyses.

**Figure 6a-h: GO Term and KEGG Pathway Enrichment Analysis EGCG Treatment Across Week 1 to Week 2.**



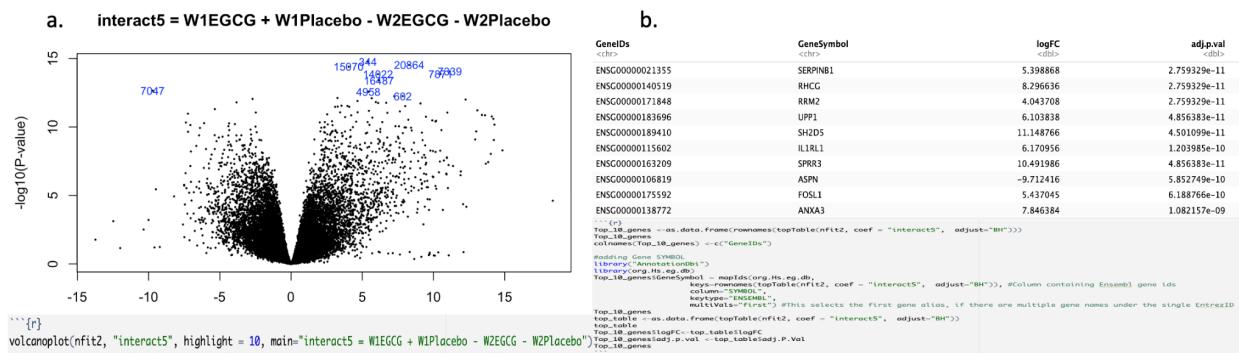
### Interaction5: “Global Differential Response” Week 1 vs Week 2:

The effect of the global differential gene expression trend from week 1 to week 2 was explored. This comparison differs from the prior analysis, as it explores the differential expression trend as a whole, and compares the global trend of gene expression from week 1 to week 2.

A volcano plot of the normalized gene data was compiled (Figure 7a), and the top 10 most differentially expressed genes (DEGs) were selected for significance by FDR (Figure 7b).

The top 10 DEG's identified for this analysis had the largest fold change values of the three comparisons selected for analysis (logFC, Figure 7b). These genes are involved in wound healing; inflammation; ECM remodeling and regeneration; cell proliferation and migration. A cytokine receptor, IL1RL1, was captured as being significantly involved in this comparison, and was similarly identified as a gene of importance in the paper selected for this project.<sup>8</sup> The code utilized to generate the following plots and tables are included below the respective images.

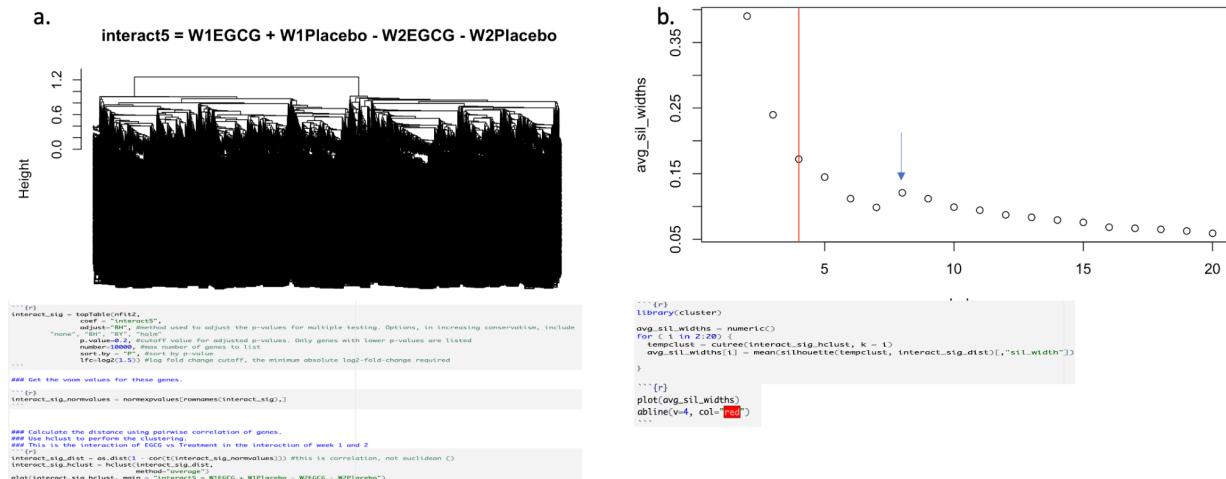
**Figure 7: Volcano Plot “Global Differential Response” Week 1 vs Week 2.**



Further exploration of the effect of “Global Differential Response” from week 1 versus week 2 for biological relevance utilized hierarchical clustering based on distance as correlation, cutoffs utilized for an FDR=0.2 and L2FC =1.5, with group selection determined via Silhouette Plot (Figure 8a+b).

Since FDR and L2FC cutoffs were determined based on capturing enough genes for analyzing the previous interaction, this analysis also produced a hierarchical clustering that was extremely congested, therefore the reliance of a silhouette plot for the determination of cluster groups was a necessity. Examining the silhouette plot below (b), it was determined that clutter groups k=4 would be utilized for consistency across analyses, however, k=8 groups would be another ideal cluster to explore in future analysis, as indicated with the blue arrow.

**Figure 8: Hierarchical Clustering and Silhouette Plot for Selection of “k” Groups.**

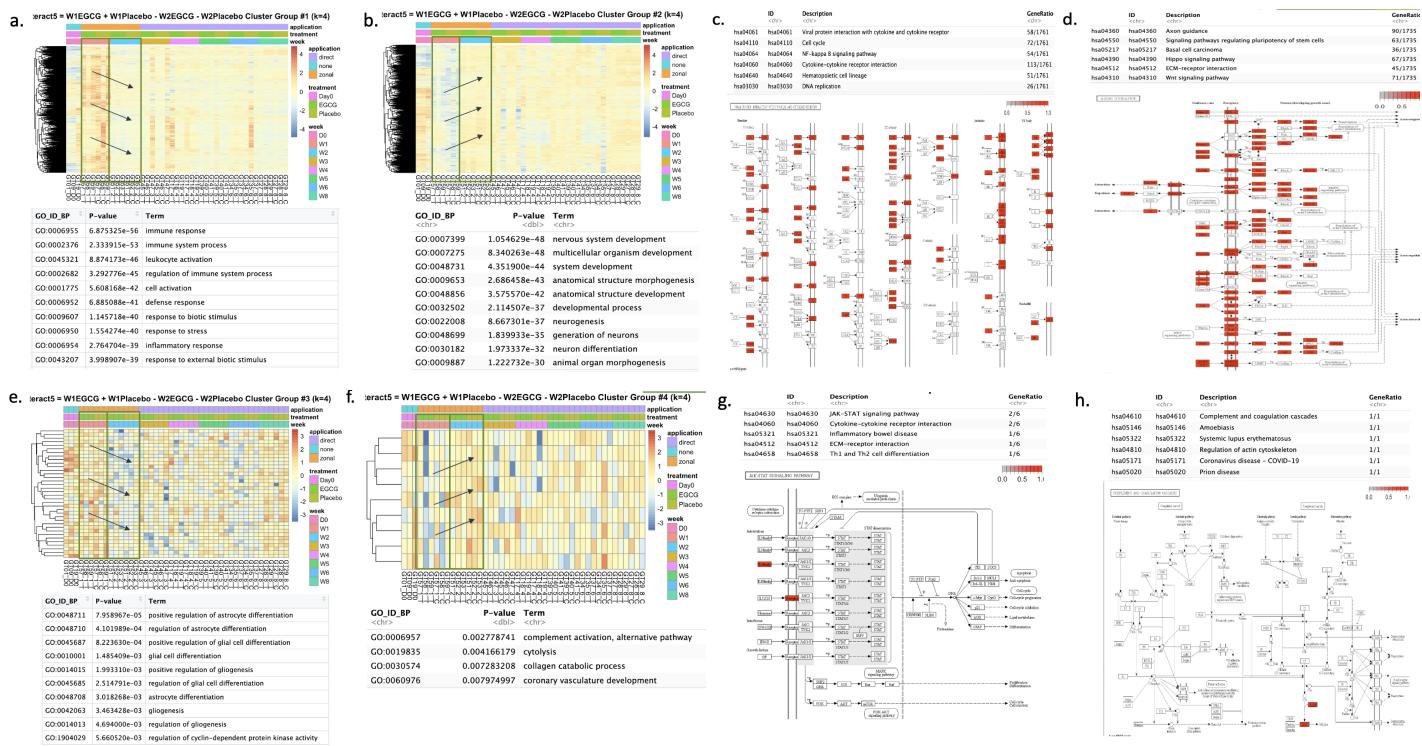


GO term enrichment and KEGG pathway analysis (Figure 9a-h) identified down-regulated biological processes involved in “immune, defense, and stress

responses” and “neural cell involvement” (a,e), and identified upregulation in biological processes involving “neurological components, structural morphogenesis” and “collagen catabolism, vascular involvement” (b,f). KEGG Pathways identified with the highest statistical significance and most gene membership involved “Viral protein interaction with cytokine and cytokine receptor”, and “Axon guidance”.(c,d) Other significant KEGG pathways identified had low gene membership and will be given less weight to interpretation.(g,h)

Critical code regarding the GO term enrichment and KEGG analysis are the same as described previously, so will not be described due to redundancy. The “interact5” term (Figure 8a, code), called in the “coeff =” line, calls the proper coefficient for cluster analysis when building the hierarchical tree, which selects the proper contrast for the data and subsequent clustering. Please refer to Supplemental Figure S1 for limma design matrix and contrasts created for these analyses.

**Figure 9a-h: GO Term and KEGG Pathway Enrichment Analysis “Global Differential Response” Week 1 vs Week 2.**



## 9. Interpretation of Results:

Our analysis of the NCBI GEO publicly available dataset GSE124161, was able to confirm that early treatment with Epigallocatechin-3-gallate, EGCG, significantly impacted genes and biological processes involved in wound healing. Data that was

analyzed beyond the first week of treatment, both Zonal and Direct applications with EGCG, were not found to be significantly different compared to the placebo control (refer to Supplementary Table S2).

In our analysis of the data regarding the effects of wound treatment with Zonal priming for 1 week with EGCG treatment was found to significantly regulate genes (Figure 1a,b) and biological processes (Figure 3a-h) involved in immunity, inflammation, ECM remodeling, and cellular organization, particularly relating to the actin cytoskeleton.

Regarding inflammation, similar to the paper, our analysis also captured down regulated IL36A,<sup>8</sup> a cytokine involved in promoting inflammation and the secretion of profibrotic mediators, which indicate a role in fibrotic disorders.<sup>10</sup> The downregulation and modulation of this cytokine may prove to be beneficial in the treating of fibrosis directly relating to the formation of scar tissue.<sup>10</sup>

Activities of wound repair include extracellular matrix (ECM) alterations, cellular organization, and cytoskeletal remodeling which involve the assembly and contraction of different actin architectures and contractile molecules involved in all stages of wound repair.<sup>11</sup> One of the most prominent pathways identified in KEGG as “hsa04810” which involves the “Regulation of the Actin Cytoskeleton” (Figure 3c), captures the genes related to the hierarchical clustering group 1 that were found to be upregulated with treatment (Figure 3a).

Regarding the comparison of EGCG treatment over time from week 1 to week 2, our week 1 analysis with EGCG treatment also captured 4 genes ACTR2, RDX, ANKRD28 and CHAC1 as also being significantly regulated with EGCG treatment during the first week (Figure 2b), which remained significant with EGCG treatment progression from week 1 into week 2 (Figure 4b). Our analyses for both comparisons detected an upregulation with treatment of ACTR2, RDX, ANKRD28 which activities encompass cytoskeletal remodeling, cell motility and migration;<sup>12,13</sup> cell adhesion;<sup>13</sup> and involvement in wound healing and tissue injury;<sup>14</sup> and a downregulation in CHAC1, which has been described to have a role in glutathione degradation, which leads to high levels of oxidative stress ferroptosis, when activated.<sup>15</sup>

Our analysis also investigated the global differential response patterns in gene expression from week 1 to week 2. The differential expression analysis captured genes that had the highest magnitude of fold change as compared to the other comparisons (logFC, Figure 7b). These genes were found to be involved in wound healing; inflammation; ECM remodeling and regeneration; cell proliferation and migration, following the same theme of wound healing processes as the prior comparisons.

Immunity and immune regulation continued to involve a downward trend (Figure 9a), with KEGG pathway containing the most gene membership involving the theme of immunity and cytokine interactions.(Figure 9c) Interestingly, many neuronal themes regarding cellularity and biological functions were captured in the GO-term enrichment analysis (Figure 9b,d,e), which may be indicative of what we theorize can be attributable to melanocytes shared derivation from neural crest cells, and or the involvement of the skin's cutaneous neuroendocrine system.<sup>16, 19</sup>

#### **10. Discussion/Conclusion:**

In this reanalysis of the NCBI GEO publicly available dataset GSE124161, our analysis was able to confirm that Zonal treatment, which includes applying treatment immediately upon injury with topical application of Epigallocatechin-3-gallate, EGCG, significantly impacted genes and biological processes involved in wound healing.

EGCG is reported in the original study to significantly downregulate mast cell, angiogenic, inflammatory, and antioxidant genes, as well as immune system and cytokine binding processes.<sup>8</sup> With our analysis we were able to confirm this interpretation, with the addition of genes and biological processes involved wound healing.

Both the immune system and inflammation play critical roles in coordinating and leading the tissue healing process. The immune response is activated by tissue injury, and immune responses lead to many downstream mechanistic effects such as the mobilization of cells, extracellular matrix deposition, and the genesis of new factors and structures to promote healing.<sup>17</sup> All of these processes were supported by our analysis, and were impacted with EGCG treatment, with the identification of genes and biological processes involved in ECM remodeling, cell death and apoptosis, cell proliferation, migration, protein and cellular localization, cytokine interaction, inflammation and immunity.

Our analysis of Zonal application with EGCG treatment provided results of upregulated biological processes involving the cytoskeleton (actin); cellular organization and cell cycle processes; cell death and apoptosis pathways. Cell death and apoptosis is a critical mechanism in reducing scar formation as most recruited endothelial cells, macrophages, and myofibroblasts initially recruited for repair must exit the wound or undergo apoptosis to clear the way for small scar formation.<sup>17</sup>

Alternately, Zonal application with EGCG down-regulated biological processes involved immune system responses, stress responses, inflammation, collagen catabolic process

and matrix disassembly. Inflammation, in excess, is a key factor in the dysregulation of normal wound healing which can lead to outcomes on either spectrum in the healing process with either the lack of proper wound closure, or hypertrophic scar formation.<sup>17</sup> Excess of both inflammation and immune system activity can contribute to non-healing wounds and limiting inflammation reduces scarring.<sup>17</sup>

Although the data was not significant for Direct application, Zonal priming with treatment of EGCG supports the hypothesis as evidenced with the results of differential gene expression, and effect on biological pathways via GO term enrichment and KEGG pathway analysis, which supports the expectation of an improved outcome of wound healing and scar formation, verified with the additional *in-vitro* and *in-vivo* experiments performed in the supporting paper.<sup>8</sup>

### **11. Limitations and Alternative Interpretations:**

One limitation/hurdle was statistical significance, perhaps natural wound healing swamped the nuances of treatment, therefore the effects of treatment after the second week was not discernible from the noise or strength of natural wound healing. This created a stopping point in our analysis where we could not continue examining the effect of Direct EGCG treatment versus the placebo for gene expression, or rely on this data for insights into wound healing or scar formation. Therefore, we can make no interpretation or determination, forgoing the analysis and comparison of 6 weeks of data when Direct treatment with EGCG yielded no significant interaction for comparison.

Alternative interpretations of this data could revolve around a deeper dive into the GO term biological processes and KEGG pathway enrichment around the neuronal processes and nerve cell activities identified in the data. Melanocytes originate from the neural crest, and share similar origins as neurons and astrocytes.<sup>16</sup> Melanocytes are best known for their production of melanin, a pigment which gives skin its color, but melanocytes have many other regulatory roles in wound healing through regulation of immune response and inflammation, and secrete many other substances that facilitate communication with the different cell types in the epidermis.<sup>18</sup> In addition melanocytes participate in a signaling cascade through the cutaneous neuroendocrine system in the skin that is a “homologue of the HPA axis (hypothalamo-pituitary-adrenal axis) to regulate local skin responses”.<sup>19</sup>

### **12. Future Work:**

So, in light of this information, experiments investigating the role of melanocytes, and broadening the role of wound healing and scar formation to include the interaction of the skin HPA homologue with connection to the central and peripheral nervous system,

may be helpful in understanding how early zonal treatment with EGCG influences healing and scar formation.

Additionally, future work for this dataset could also include analysis of the “interact5” investigating the “Global Differential Response” Week 1 vs Week 2 regarding the selection of k=8 groups, to further investigate interesting sub-groups within the data.

Finally, both the “Zonal Week 1 EGCG/Placebo” and “Interact 5” contrasts had statistically significant data, where the FDR could have been lowered to 0.05 and below for these contrasts. This would allow for the interrogation of clusters that had similar differential expression patterns facilitating the extraction of meaningful biological trends from the data.

## **Supplemental Figure S1**

### **Design Matrix and Contrasts for LIMMA and edgeR Differential Expression Analysis**

```

Load limma and edgeR

```{r}
library(limma)
library(edgeR)
```

Create a design matrix for lm.

First we create the levels that we are potentially interested in. Some of these we may not use, but we have them in case we want to make a different comparison.

```{r}
week_mm <- factor(pheno_df$week, levels = c("D0", "W1", "W2", "W3", "W4", "W5", "W6", "W8"))
treatments_mm <- factor(pheno_df$treatments, levels = c("Day0", "Placebo", "EGCG"))
status_mm <- factor(pheno_df$status, levels = c("uninjured", "injured"))
application_mm <- factor(pheno_df$application, levels = c("none", "zonal", "direct"))
appl_by_wk_mm <- factor(pheno_df$appl_by_wk, levels = c("W0", "Zw1", "Zw2", "Dw1", "Dw2", "Dw3", "Dw4", "Dw6"))
```

This model matrix defines all of the things that you are interested in comparing. It creates a matrix that defines the various experimental categories of the samples in your experiment that you want to compare. We will use the week_mm and treatments_mm as we are interested in comparing treatments across the weeks.

```{r}
weektreat = factor(paste(week_mm,treatments_mm, sep=""))
weektreat
```

```{r}
colnames(design) = levels(weektreat)
design
```

attr(,"contrasts")$weektreat
[1] "contr.treatment"

	D0Day0	W1EGCG	W1Placebo	W2EGCG	W2Placebo	W3EGCG	W3Placebo	W4EGCG	W4Placebo	W5EGCG	W5Placebo	W6EGCG	W6Placebo	W8EGCG	W8Placebo
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0


```{r}
dge = DGEList(counts = GSE124161_readcount)
dim(dge$counts) #before filtering

keep = filterByExpr(dge, design)
dge = dge[,keep,,keep.lib.sizes=FALSE]

dim(dge$counts) #after filtering there are 20,935 genes, this is a lot stricter than the 25,995 genes we kept by filtering using genesum which was arbitrary and selected by judgement.
```

```

## Supplemental Figure S1(continued)

### Design Matrix and Contrasts for limma and edgeR Differential Expression Analysis

```
Voom normalization
```

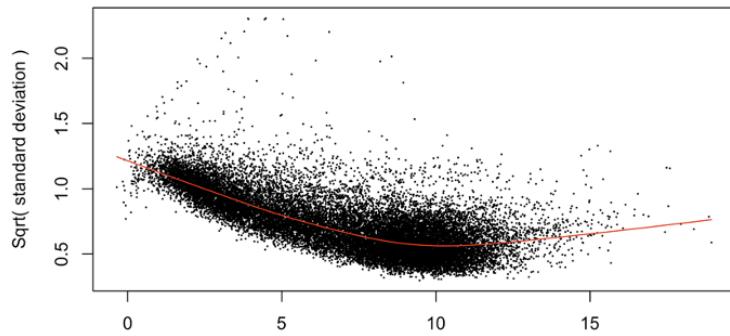
Voom provides data in a format that can be used for standard limma methods. In the Limma manual, another normalization process is called "eset", which is normalization through the AFFY package, however we are normalizing this data with "voom", so "v" is the object we are storing the normalized data in.

Voom is the normalization method that allows us to use the data in downstream analyses. For this expression, voom is acting on the 20,935 genes captured in "dge", using the design outlined

```
```{r}
v = voom(dge, design, plot=TRUE, normalize="quantile")
```

```
...
```

voom: Mean-variance trend



```
## Then create the lmfit ( this calculates the "within" variance). This fits a linear model to the data.
```

```
```{r}
nfit = lmFit(v,design)
```

```

```
## Now specifically compare the different coefficients for the comparison
```

This gives us a lot more control to make the specific comparisons we want. This gives us a lot of control over a complex data set, one with a lot of levels, time-series data.

```
```{r}
```

```
newcontrasts = makeContrasts(Zw1EGCG_vs_Zw1_placebo = W1EGCG - W1Placebo,#these are comparing the Treatment to the control at a single time point
```

```
Zw2EGCG_vs_Zw2_placebo = W2EGCG - W2Placebo,
Dw1_EGCG_vs_Dw1_placebo = W3EGCG - W3Placebo,
Dw2_EGCG_vs_Dw2_placebo = W4EGCG - W4Placebo,
Dw3_EGCG_vs_Dw3_placebo = W5EGCG - W5Placebo,
Dw4_EGCG_vs_Dw4_placebo = W6EGCG - W6Placebo,
Dw6_EGCG_vs_Dw6_placebo = W8EGCG - W8Placebo,
```

```
interact = (W1EGCG - W1Placebo) - (W2EGCG - W2Placebo), #Change in expression levels from Zonal week1 to wk2 differs between the EGCG-treated group and the placebo-treated group. If statistically significant, it would suggest that the change in expression levels over time (from week1 --> week 2) differs between the EGCG-treated group and the placebo-treated group.
```

```
interact2 = (W1EGCG - W1Placebo) - (W3EGCG - W3Placebo), #Change in expression levels for Zonal wk 1 to Direct wk 1. If statistically significant, it would suggest that the change in expression levels over time (from week1 --> week 2) differs between the EGCG-treated group and the placebo-treated group.
```

```
interact3 = (W2EGCG - W2Placebo) - (W4EGCG - W4Placebo), #Change in Expression levels for Zonal wk 2 and Direct wk 2
```

```
interact4 = W2EGCG - W1EGCG - W2Placebo + W1Placebo, # EGCG vs Placebo, Significant means that EGCG is having a statistically differential response between the two time points W1-W2
```

```
interact5 = W1EGCG + W1Placebo - W2EGCG - W2Placebo, # Significant means that EGCG is having a statistically differential response between the two time points, relative to the control
```

```
levels = weektreat)
```

```
...
```

**Supplemental Table S2**

**FDR Values of Comparisons/Contrasts Created in limma.**

Group	Contrast terms in limma “new Contrasts”	FDR minimum limit for clustering
Zonal Wk1	<b>Zw1EGCG_vs_Zw1_placebo = W1EGCG - W1Placebo</b>	<b>0.2 (but, can go lower to 0.05)</b>
Zonal Wk2	Zw2EGCG_vs_Zw2_placebo = W2EGCG - W2Placebo	1
Direct Wk1	Dw1_EGCG_vs_Dw1_placebo = W3EGCG – W3Placebo	1
Direct Wk2	Dw2_EGCG_vs_Dw2_placebo = W4EGCG – W4Placebo	1
Direct Wk3	Dw3_EGCG_vs_Dw3_placebo = W5EGCG – W5Placebo	1
Direct Wk4	Dw4_EGCG_vs_Dw4_placebo = W6EGCG – W6Placebo	0.9
Direct Wk6	Dw1_EGCG_vs_Dw1_placebo = W8EGCG – W8Placebo	1
<b>Interaction of EGCG treatment Wk1 vs Wk2</b>	<b>Interact = (W1EGCG - W1Placebo) - (W2EGCG - W2Placebo),</b>	<b>0.2 (cannot go any lower)</b>
<b>Global Expression Wk1 vs Wk2</b>	<b>Interact5 = W1EGCG + W1Placebo - W2EGCG - W2Placebo</b>	<b>0.2 (but, can go lower to 0.05 and below.)</b>

**References:**

1. Tsao R. Chemistry and biochemistry of dietary polyphenols. *Nutrients*. 2010;2(12):1231-1246. doi:10.3390/nu2121231
2. Mokra D, Joskova M, Mokry J. Therapeutic Effects of Green Tea Polyphenol (-)-Epigallocatechin-3-Gallate (EGCG) in Relation to Molecular Pathways Controlling Inflammation, Oxidative Stress, and Apoptosis. *International Journal of Molecular Sciences*. 2023; 24(1):340. <https://doi.org/10.3390/ijms24010340>
3. Zhong Y, Chiou YS, Pan MH, Shahidi F. Anti-inflammatory activity of lipophilic epigallocatechin gallate (EGCG) derivatives in LPS-stimulated murine macrophages. *Food Chem*. 2012;134(2):742-748. doi:10.1016/j.foodchem.2012.02.172
4. Riegsecker S, Wiczynski D, Kaplan MJ, Ahmed S. Potential benefits of green tea polyphenol EGCG in the prevention and treatment of vascular inflammation in rheumatoid arthritis. *Life Sci*. 2013;93(8):307-312
5. Rashidi B, Malekzadeh M, Goodarzi M, Masoudifar A, Mirzaei H. Green tea and its anti-angiogenesis effects. *Biomedicine & Pharmacotherapy*. 2017;89:949-956. doi: <https://doi.org/10.1016/j.biopha.2017.01.161>.
6. Li, M., Xu, J., Shi, T., Yu, H., Bi, J. and Chen, G. (2016), Epigallocatechin-3-gallate augments therapeutic effects of mesenchymal stem cells in skin wound healing. *Clin Exp Pharmacol Physiol*, 43: 1115-1124. <https://doi.org/10.1111/1440-1681.12652>
7. Song, Y, Wang, T, Yang, L, et al. EGCG inhibits hypertrophic scar formation in a rabbit ear model. *J Cosmet Dermatol*. 2023; 22: 1382- 1391. doi:10.1111/jocd.15587
8. Ud-Din S, Foden P, Mazhari M, et al. A Double-Blind, Randomized Trial Shows the Role of Zonal Priming and Direct Topical Application of Epigallocatechin-3-Gallate in the Modulation of Cutaneous Scarring in Human Skin. *J Invest Dermatol*. 2019;139(8):1680-1690.e16. doi:10.1016/j.jid.2019.01.030
9. Frech FS, Hernandez L, Urbonas R, Zaken GA, Dreyfuss I, Nouri K. Hypertrophic Scars and Keloids: Advances in Treatment and Review of Established Therapies. *Am J Clin Dermatol*. 2023;24(2):225-245. doi:10.1007/s40257-022-00744-6
10. Elias M, Zhao S, Le HT, et al. IL-36 in chronic inflammation and fibrosis - bridging the gap?. *J Clin Invest*. 2021;131(2):e144336. doi:10.1172/JCI144336
11. Ajeti V, Tabatabai AP, Fleszar AJ, et al. Wound Healing Coordinates Actin Architectures to Regulate Mechanical Work. *Nat Phys*. 2019;15:696-705. doi:10.1038/s41567-019-0485-9
12. Schlieben P, Meyer A, Weise C, Bondzio A, Gruber AD, Klopffleisch R. Tandem duplication of KIT exon 11 influences the proteome of canine mast cell tumours. *J Comp Pathol*. 2013;148(4):318-322. doi:10.1016/j.jcpa.2012.07.006
13. Chen H, Dai J. miR-409-3p suppresses the proliferation, invasion and migration of tongue squamous cell carcinoma via targeting RDX. *Oncol Lett*. 2018;16(1):543-551. doi:10.3892/ol.2018.8687

14. Samaras SE, Almodóvar-García K, Wu N, Yu F, Davidson JM. Global deletion of Ankrd1 results in a wound-healing phenotype associated with dermal fibroblast dysfunction. *Am J Pathol.* 2015;185(1):96-109.  
doi:10.1016/j.ajpath.2014.09.018
15. Chen MS, Wang SF, Hsu CY, et al. CHAC1 degradation of glutathione enhances cystine-starvation-induced necroptosis and ferroptosis in human triple negative breast cancer cells via the GCN2-eIF2α-ATF4 pathway. *Oncotarget.* 2017;8(70):114588-114602. Published 2017 Dec 9.  
doi:10.18632/oncotarget.23055
16. Sommer L. Generation of melanocytes from neural crest cells. *Pigment Cell Melanoma Res.* 2011;24(3):411-421. doi:10.1111/j.1755-148X.2011.00834.x
17. Larouche J, Sheoran S, Maruyama K, Martino MM. Immune Regulation of Skin Wound Healing: Mechanisms and Novel Therapeutic Targets. *Adv Wound Care (New Rochelle)*. 2018;7(7):209-231. doi:10.1089/wound.2017.0761
18. D'Mello SA, Finlay GJ, Baguley BC, Askarian-Amiri ME. Signaling Pathways in Melanogenesis. *Int J Mol Sci.* 2016;17(7):1144. Published 2016 Jul 15.  
doi:10.3390/ijms17071144
19. Slominski AT, Zmijewski MA, Skobowiat C, Zbytek B, Slominski RM, Steketee JD. Sensing the environment: regulation of local and global homeostasis by the skin's neuroendocrine system. *Adv Anat Embryol Cell Biol.* 2012;212:v-115.  
doi:10.1007/978-3-642-19683-6\_1