参赛队员姓名：刘欣然 杨灏芳

中学：中国人民大学附属中学

省份：北京

国家/地区：中国

指导教师姓名：李昱华

论文题目：Analysis of Differentially Expressed Genes and Identification of Pathways Related to Acne Vulgaris Based on Bioinformatics Methods

本参赛团队声明所提交的论文是在指导老师指导下进行的研究工作和取得的研究成果。尽本团队所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果。若有不实之处，本人愿意承担一切相关责任。

参赛队员：刘欣然 杨灏芳　　指导老师：李昱华

2020 年　　**9** 月　　**13** 日

Analysis of Differentially Expressed Genes and Identification of Pathways Related to Acne Vulgaris Based on Bioinformatics Methods

Abstract

Acne vulgaris is one of the most common skin diseases, while the pathogenesis is yet to be fully unveiled. In order to identify the biological pathways or processes that contribute to the formation of Acne vulgaris, we process KEGG and GO enrichment analysis on R platform. Two high throughput RNA sequence datasets, namely GSE124389 and GSE115099, are chosen to be our original datasets. We identified up 41 up-regulated and 1067 down-regulated differentially expressed genes (DEGs) in GSE124389, and 230 up-regulated and 355 down-regulated DEGs in GSE115099, with the threshold of p (p. adjust) less than 0.05 and with the log Fold Change larger than 1 or less than 1. Our results emphasize the crucial role of immunological pathways in the formation of acne vulgaris, and further strengthen the relationship between immune responses and acne formation. FcεRI signaling pathway, neutrophil activation, and NOD-like receptor signaling pathway were identified as significant pathways. Further analysis can be conducted using larger sample bases.

Keywords:

Acne vulgaris, Enrichment analysis, Pathway identification

Table of Contents

Introduction

Acne vulgaris is one of the most interracial and common skin disorders, causing significant psychological and physical morbidity across the world, and the prescribed drugs to cure acne have cost over 4 billion dollars worldwide. In the US alone, between 40 and 50 million individuals are bothered by acne. Besides bringing huge fiscal outlays, the wide existence of acne also leads to other socio-economic problems like unemployment due to the severe symptoms and mental illness (Bataille, 2002). Contrary to people's common sense, the development of acne may not have a great connection with cosmetics, prescribed drugs, and occupation, etc. (Goulden, 1995); rather, it is contributed by other direct endogenous and exogenous elements. Exogenous factors include follicular plugging resulting from the retention of desquamated keratinocytes within the pilosebaceous unit as well as stimulus caused by bacteria (such as Propionibacterium). Endogenous factors include the effect of hormones from estrogen and androgens, and hereditary factors (Davis, 2010). Due to the urgent need for unveiling the causes of the disease, genetic influence has been spotted, whereas the explicit influence of genetic factors has not been thoroughly studied yet.

Li *et al*., 2019 and Chen *et al*., 2019 both used datasets GSE 6475 and GSE 53795(microarray genomic data series), obtained from Gene Expression Omnibus (NCBI-GEO), to conduct bioinformatic analysis on acne-related genes and pathways (Wheeler *et al*., 2001). Their results highlight the effect of proinflammatory cytokines as well as their genes, including IL1β, CXCL1, CXCL2, CXCR4, CXCL8, important for immune response and regulation. Other Differentially Expressed Genes (DEGs) such as FPR2 and CCR1 code for peptides needed in the inflammatory receptor signaling pathway. In addition to genes and cytokines, significant pathways are also noted in their studies. Among them, chemokine signaling pathway, FC gamma R-mediated phagocytosis, and NF-κB pathway are mentioned most frequently. The analysis was thorough, further strengthening the link between the immune system and acne. However, no analysis was done using high throughput data, and the field of acne is sophisticated and not fully understood. The relationship between acne vulgaris and gene pathways needs to be further tested and explained.

Materials and Methods

RNA-seq Data

NCBI-GEO[1] is a public functional genomics data repository; on the website, a GSE accession is a series of GSM samples publicized by researchers worldwide. Two series used in this research were obtained by searching "acne vulgaris homo sapiens" and setting a limit of using high throughput sequencing data. GSE115099(He, 2019) was conducted using platform GPL20795 HiSeq X Ten on 4 acne patients with NCSTN mutation as well as 4 healthy controls. Skin biopsies were made among lesions of patients and skin tissues from

[1] https://www.ncbi.nlm.nih.gov/geo/

healthy people. GSE124389(Jiang, 2018), also a high throughput sequencing, used the platform GPL20301 Illumina HiSeq 4000 and followed up the treatment process of acne flare-up patients. 11 acne patients and 3 healthy patients are divided into 4 groups, as such: flare-up patients, improved patients, no obvious change patients, and healthy control. To be noticed, the gene records of GSE 124389 were not the original expression matrix. The contributor filtered the significant genes in different comparison groups: every patient before and after treatment in group 1, 2, and 3, separately; group 1, 2, and 3 before and after treatment; group 1, 2, and 3 vs healthy control group 4. The treatment was isotretinoin or minocycline and the samples were the peripheral blood of patients instead of skin tissue.

Identification of Different Expressed Genes (DEGs)

We used Excel 2016 and R (version 4.0.2) in data processing, statistical analysis, and computing. Firstly, we dropped the low-expressed genes. Every gene with the FPKM of 80% samples not less than 1 is counted to have an adequate gene expression level (Benjamini, 2020). Fold change, known as the ratio of the average expression levels between acne patients and healthy control people, indicates the difference of expression. In our case, genes with the log2 fold change greater than 1 are considered as up-regulated, while genes whose log2 fold changes are less than -1 are considered as down-regulated. One-tailed T-test was made in GSE115099 to single out the significantly different expressed genes. A false discovery rate (FDR) control developed by Benjamini and Yekutieli was applied to the series of T-test, to avoid high false positive rate (Benjamini & Hochberg, 1995).

GSE124389 was processed differently than 115099. We chose 5 of all the group comparisons to find the up-regulated genes: group 1 after treatment (acne flaring-up) vs group 1 before treatment, group 2 before treatment vs group 2 after treatment (condition improved), group 1 vs group 4, group 2 vs group 4 and group 3 vs group 4. The contributor applied F-statistics to the series as a statistical test. All significant genes in these comparisons are considered the significant DEGs. Since the size of genes in every comparison is small, FDR was not applied in GSE124389. In order to visualize the DEGs we filtered, TBtools developed by Chen *et al. and* ClusterProfiler by Yu *et al.* are used (2020 and 2012).

KEGG and GO pathway enrichment

Kyoto Encyclopedia of Genes and Genomes (KEGG)[2] is a frequently-used genome sequencing database for accessing molecular-level information (Kanehisa & Goto, updated 2020). The Gene Ontology knowledge base (GO)[3] classifies genes in three categories: Biological Process (BP), Molecular Function (MF), and Cellular Component (CC)(Ashburner *et al*., updated 2020). DEGs from the two databases mentioned are put into statistical tools for pathway enrichment analysis, to identify the significant enriched biological process and

[2] https://www.kegg.jp/

[3] http://geneontology.org/

pathways related to the acne. R package clusterProfiler is an effective analysis tool for GO and KEGG database statistical analysis (Yu *et al*., 2012). In this study, we applied a cutoff of 0.01 for p-value and 0.05 for FDR on GO, whereas both p-value and FDR are set as 0.05 on KEGG.

Results

1. *Summary of Two Datasets: GSE124389 and GSE115099*

Five categories of samples are involved in the experiments and encompassed in GSE124389: control group (healthy people), patients before treatments, patients that show no explicit improvement after treatment, patients that have gotten a worse acne condition and patients with improvements on their symptoms. With comparisons between each of the experimental group and the control group as well as comparisons within each group, gene expression levels from each sample have been recorded and preliminary screening has been processed with a significance level of 0.05. In GSE115099, samples of four patients and four healthy control subjects are gathered to elucidate the relationship between NCSTN mutations (in a family with acne vulgaris) and familial pathogenesis by investigating DEGs. The distribution of the statistical significance and log2 fold change are shown in Fig 1. Because of the screening from the contributor, there is a gap on the plot of GSE 124389.

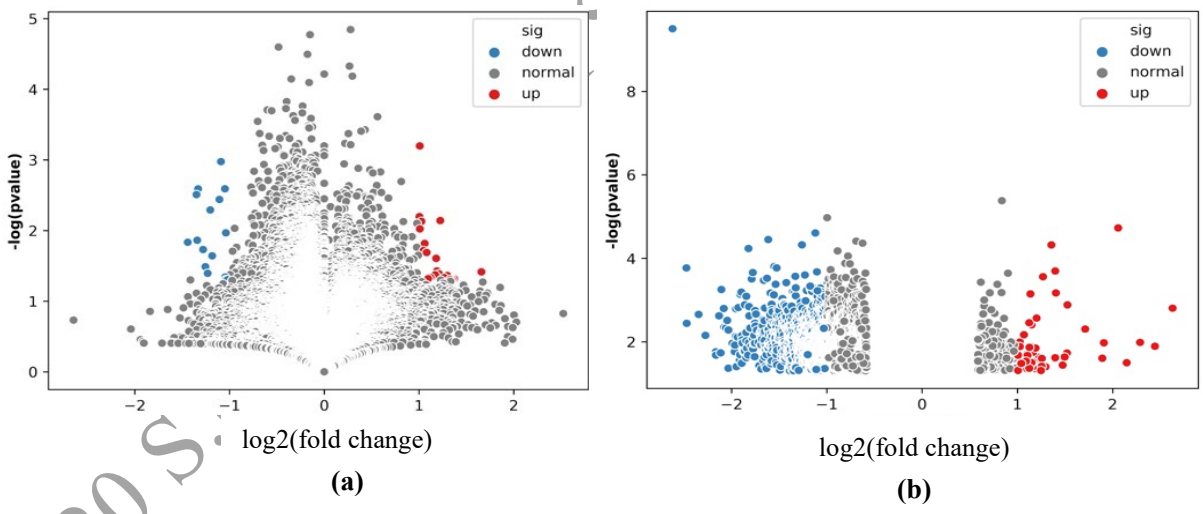


Figure 1 Distribution of the expression level difference and statistical test results of original datasets. (a) GSE115099. (b) GSE124389

2. *The Filtration of Differentially Expressed Genes (DEGs)*

Five groups of samples are selected from GSE124389, including two pairing sets, and three unpairing sets. This can also be interpreted as two pair sets of internal relativity of patients before treatments and internal relationship relativity of patients with no improvements, and three comparison groups between experimental and control groups (mentioned above). As for

GSE115099, all the valid data were taken into account. Fig 2 shows the expression profile of the two datasets. Since p-value has been provided in the original dataset of GSE124389, the T-test was only applied in GSE115099, acquired adjusted p-value. With the threshold of p (p. adjust) <0.05 and the absolute value of log2 Fold Change larger than 1, we identified 41 up-regulated and 1067 down-regulated DEGs in GSE124389, and 230 up-regulated and 355 down-regulated DEGs in GSE115099 (table 1 and 2). Among all the labelled genes, P13, presenting as an extremely significant protein coding gene, has been demonstrated to have direct connection with keratinization and innate immune system (GeneCards , https://www.genecards.org/.). TINF2 is responsible for encoding telosomes, protecting telomeres, which can be correlated with the development of acne as telomere has great connection with cell senescence (performing inner relationship with acne) (Ribero, 2017) and its relevant disease dyskeratosis congenita (DKC) is also an influential factor of the formation of cutaneous disease (Cole, 1930).

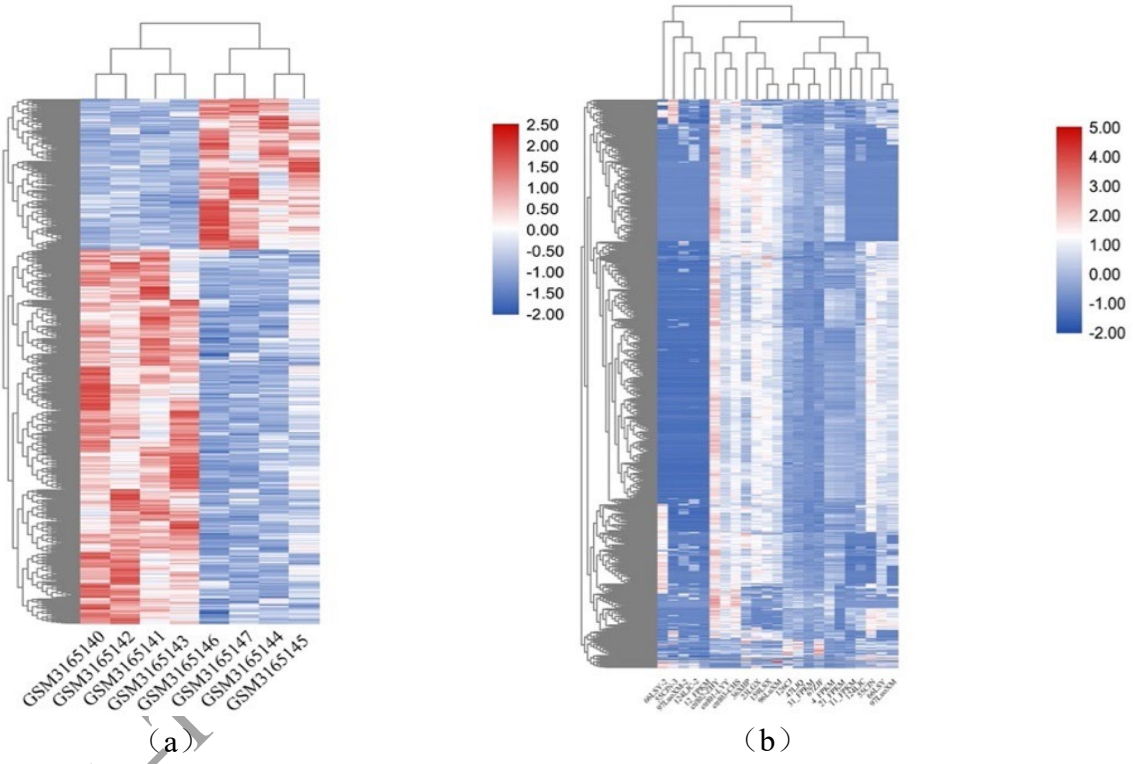
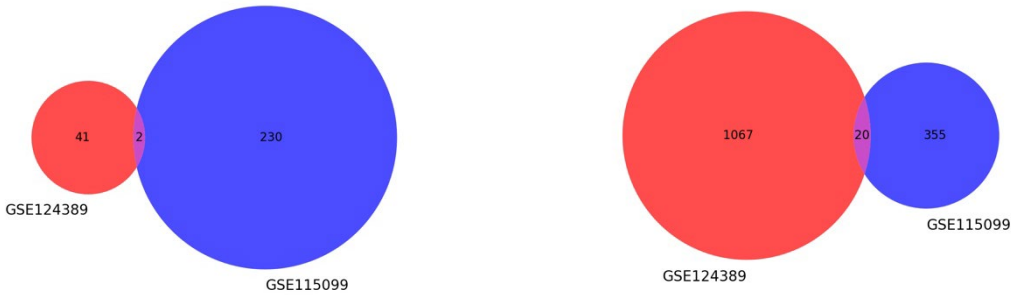


Figure 2 Expression profile of the DEGs between acne patients and control groups. (a) GSE115099. (b) GSE124389)

Table 1 Partial differentially expressed genes in paired and unpaired groups of GSE124389

DEGs	Paired	Unpaired
GSE124389 up-regulated genes	*LY96, CD52, S100A12, TICAM2, AC073111.5, MRFAP1, KLRK1*	*PI3, SNU13, NDST1, HLA-G, CASP10, HES4, ARHGAP30, AL662899.3, NDST1, STK4, CHKB-CPT1B, TBKBP1, DCTN5, IL18RAP, TNFSF12-TNFSF13, AC120057.3, HBG1, HBG2, AC099811.2, CD177, PGLYRP1, HBB, LGALS9B, VMO1, NFIX, JSRP1, MMP9, TBKBP1, CHKB-CPT1B, AC018755.2,*

		ADGRG3, COL18A1, KLC3, ARHGAP30, CDC42EP2, NDST1, ALPL, TMEM265, AGER, LRWD1, RCN3, PHC2, CYP27A1, SMIM3
GSE124389 down-regulated genes	*AC010441.1, TGIF2, BIVM-ERCC5, NKTR, ZBTB37, AGO1, SPATA13, KMT2A, CDKN2AIP*	*UHMK1, B2M, HIST1H4C, ARL8B, HIST1H4C, SKIL, SPATA13, NR2C2, ZDHHC20, CCR5, ATM, GATA2, UHMK1, AC138811.2, IL7R, SPOPL, TNPO1, PRKACB, FCER1A, ZKSCAN8, BCL11B, MBNL1, P2RY10, BMI1, OSBPL8, BTAF1, ZBTB1, TCF7, CMPK1, MYADM, UFM1, TRAT1, TSPAN18, MIER1, GPR183, ITGA4, SLC7A6, PTAR1, ZNF24, RASGRP1, RCAN3, BIRC3, etc.* (1059 in total)



(a) (b)

Figure 3 Overlapping of DEGs from GSE124389 and GSE115099. (a) up-regulated genes. (b) down-regulated genes.

Comparing the DEGs of two datasets, two common DEGs are found in upregulated genes of GSE115099 and GSE124389, and twenty common DEGs in downregulated groups. The groups of up-regulated genes have overlap on genes PI3 and COL18A1, while down-regulated genes have intersection on CLIP4, TOB1, PGRMC1, RNF141, IPMK, FCER1A, OGFRL1, CHURC1-FNTB, MPP7, PEBP, CRY2, POLR2M, PIK3IP1, TC2N, JADE1, TBC1D4, RCAN3, EIF1AX, DENND4C, and KDM6A.

Table 2 partial differentially expressed genes of GSE115099

DEGs	Gene Symbol
GSE115099 up-regulated genes	*TINF2, CFI, CDH5, LINC01503, TSHZ3, JAML, ADA, CCDC117, OAS2, GYPC, ST6GALNAC4, ZNF672, HOXD4, MS4A6A, SLC35C1, GMFG, CFB, GRK5, NEU3, P3H1, DRAM1, ARID3A, TOR4A, DNM3OS, NPIPP1, BCL6B, EPSTI1, CST7, PLEKHO1, PLPP5, CD163L1, KCNJ8, CCL5, ISG15, NRP2, APBB1IP, DENND5B, FMNL1, RN7SL116P, SPRR2F, WDFY4, BTK, C1QTNF5, EVI2B, C16orf54, SEMA6B, NCF1C, IGHV3-30, IGHV3-53, IGHV3-23, IGHV3-21, IGLV1-51, IGHV3-74, etc.* (230 in total)
GSE115099 down-regulated	*FZD8, CCND1, MYH14, MACROD1, ENO1P4, CNTNAP3P2, NFIB, ECHDC3, RTN4, EFNA1, SYNGR1, RNF180, CAT, SPRY2, EFNB2,*

genes	*TBC1D4, NFIA, TGFBR3, SLC9A7, ST6GALNAC2, ABHD6, EFNA5, TMEM237, ZNF658, CEBPA, BNIPL, EMP2, CSPG4, RNF141, NEBLl, USP54, PGRMC2, RTN3, HSF2, IL20RA, COL21A1, HNRNPA1P33, OCLN, L3MBTL4, FRS2, ELL3, ABTB2, KANTR, TFDP2, HS3ST6, LINC01290, DEGS1,etc.* (355 in total)

3. KEGG Enrichment Analysis

KEGG enrichment analysis was run on both datasets. 29 pathways (all down-regulated) and 14 pathways (13 up-regulated and 1 down-regulated) were acquired by analyzing DEGs of GSE124389 and GSE115099, respectively (Table 3. Complete table in supplementary Table 2). Comparing to the findings of Chen *et al.* 2019 and Li *et al.* 2019 who used microarray data GSE6475 and GSE53795 on acne analysis, there are 8 overlapping KEGG pathways: hsa04664: Fc epsilon RI signaling pathway, hsa04666: Fc gamma R-mediated phagocytosis, hsa04621: NOD-like receptor signaling pathway, hsa04670: Leukocyte transendothelial migration, hsa04380: Osteoclast differentiation, hsa04062: Chemokine signaling pathway, hsa05202: Transcriptional misregulation in cancer, and hsa04064: NF-kappa B signaling pathway. The overlapping pathways are similar and can be classified as immunological response and signaling pathways, consistent with current knowledge of acne formation. For example, one of the examples, FcεRI signaling pathway, has already been proven to be active in recruiting and activating cytokines and chemokines in inflammatory cells (Shin, 2015), which may lead to further aggravation of acne vulgaris.

Table 3: KEGG pathway and GO term enrichment analyses of DEGs in acne

Category	Term	Count	P Value	FDR
Gene Ontology	GO:0006909-phagocytosis	34	1.23E-21	3.24E-18
	GO:0038094-Fc-gamma receptor signaling pathway	23	1.84E-20	1.99E-17
	GO:0050900-leukocyte migration	37	2.98E-20	1.99E-17
	GO:0006959-humoral immune response	31	4.60E-19	1.34E-16
	GO:0006958-complement activation, classical pathway	20	5.20E-17	9.77E-15
	GO:0002697-regulation of immune effector process	29	3.93E-14	5.17E-12

	GO:0001533-cornified envelope	11	1.13E-10	1.47E-08
	GO:0042119-neutrophil activation	25	3.11E-10	2.82E-08
	GO:0018149-peptide cross-linking	10	1.05E-09	8.37E-08
	GO:0043312-neutrophil degranulation	23	4.95E-09	3.52E-07
	GO:0002283-neutrophil activation involved in immune response	23	5.56E-09	3.85E-07
	GO:0002446-neutrophil mediated immunity	23	8.44E-09	5.69E-07
	GO:0001906-cell killing	6	7.56E-07	5.56E-03
	GO:0070268-cornification	9	4.36E-06	2.25E-04
	GO:0002703-regulation of leukocyte mediated immunity	10	1.00E-04	3.14E-03
KEGG	hsa05340: Primary immunodeficiency	5	1.06E-04	2.73E-03
	hsa04662: B cell receptor signaling pathway	7	7.37E-05	2.28E-03
	hsa04670: Leukocyte transendothelial migration	7	5.74E-04	9.88E-03
	hsa04380: Osteoclast differentiation	7	1.14E-03	1.77E-02
	hsa04664: Fc epsilon RI signaling pathway	8	1.96E-06	3.05E-04
	hsa04666: Fc gamma R-mediated phagocytosis	8	2.09E-05	9.59E-04
	hsa04062: Chemokine signaling pathway	8	2.90E-03	4.09E-02

hsa04611: Platelet activation	9	2.47E-05	9.59E-04
hsa04621: NOD-like receptor signaling pathway	9	4.52E-04	8.75E-03
hsa04218: Cellular senescence	24	1.78E-06	4.99E-04

In addition to the overlapping pathway, some additional findings have been recognized in our study as well. Activation of platelets in P.acne patients is found by Sidra *et al*. (Younis *et al*., 2017), which could explain the outcome of up regulation of hsa04611: platelet activation. hsa05340 is the primary immunodeficiency pathway; the up regulation of this disorder will affect the host defense system, and increase the chance of infection of microbes, leading to acne formation. One of the phenomena of immunodeficiency is affected B-cell differentiation (GenomeNet[4]). Decreased ability of B-cell differentiation in acne vulgaris patients offers new insights to the pathways also correlate to existing study and potential research areas. Hsa04218: cellular senescence is a down regulated pathway obtained from GSE124389. The downregulation corresponds to the experiment by Ribero *et al.* in 2016, whose results presented reduced skin aging and longer telomere length in acne patients after conducting a genomic experiment on twins with control and acne-patient group. Figure 4 shows that the pathways of up regulated genes in GSE 115099 are clustered mainly in B-cell receptor signaling and protein processing, which are believed to be related to immune response and the secretion of inflammatory factors, both of which play an considerable role in the flaring-up of acne.

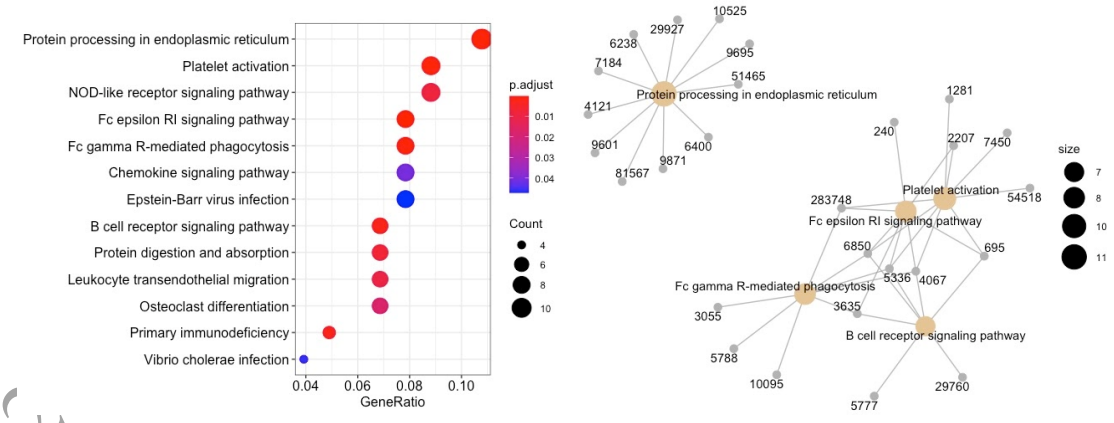


Figure 4 Significantly enriched KEGG pathways and their relationship of up-regulated DEGs from GSE115099

[4] https://www.genome.jp/dbget-bin/www_bget?hsa05340#:~:text=KEGG%20PATHWAY%3A%20hsa05340&text=Primary%20immunodeficiencies%20(PIs)%20are%20a,natural%20killer%20(NK)%20cells

4. GO Enrichment Analysis

108 GO terms (42 up regulated and 66 down regulated) were obtained from GSE124389 and 213 pathways (158 upregulated and 55 down regulated) from GSE115099 (Table 3. Complete table in supplementary Table 2). Divided into three categories of GO, the upregulated pathways of GSE124389 have 29, 4, 9 for BP, MF, and CC respectively, whereas those of GSE115099 have 135, 17 and 6. One of the upregulated pathways, GO:0050900-leukocyte migration, is consistent with the outcome of Chen *et al*. 2019. This pathway refers to the movement of leukocyte between tissues, contributing to the inflammatory process.

In order to show the distribution of pathways in terms of features, we manually classify the upregulated pathways into three mutually exclusive subsets: immunological pathways (23 for GSE124389 and 107 for GSE115099), non-immunological pathways (19 for GSE124389 and 47 for GSE115099) and skin related pathways (0 for GSE124389 and 4 for GSE115099). There are 6 common up-regulated pathways of the two datasets, four of which are related to neutrophil (GO:0042119, GO:0043312, GO:0002283, GO:0002446); the others are GO:0002697-regulation of immune effector process and GO:0002703-regulation of leukocyte mediated immunity. All of them are immunological pathways. The skin related pathways include GO:0070268-cornification and GO:0031424-keratinization. For the significance and correlation between pathways, see Figure 5 (a,b,c,d).

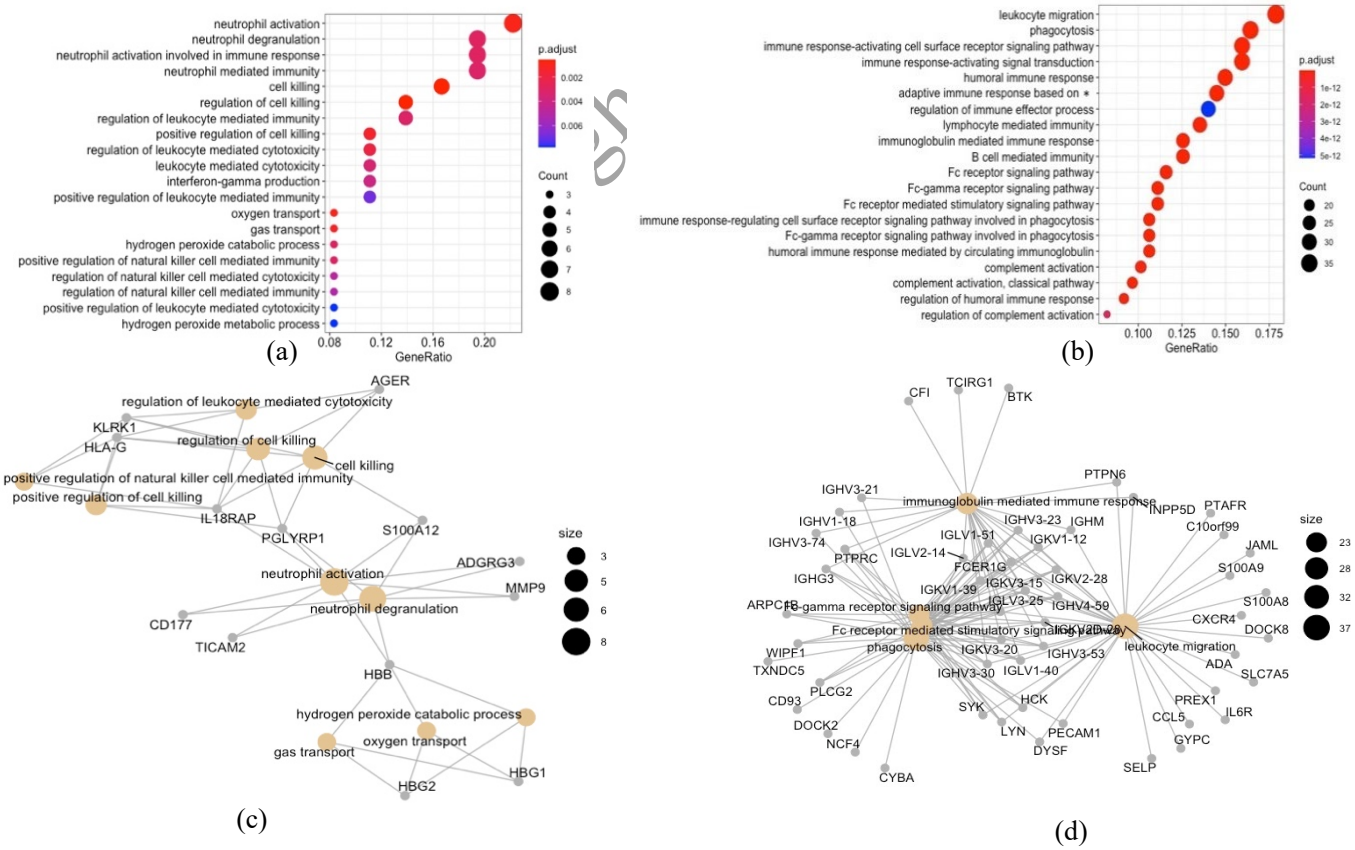


Figure 5 Significantly enriched GO terms and their relationship of up-regulated DEGs. (a.c) GSE124389 (b.d) GSE115099

Discussion

Acne-related pathways and genes were usually tested along with other diseases in experimental trials, and were not well studied independently. With more and more public genomic data uploaded online, we can reveal the underlying reasons for acne formation and aggravation that has affected people's life quality. In our analysis, we have revealed the relationship between acne and some significant DEGs mentioned in Results: P13 has correlation with keratinization and alteration of keratinization may lead to severe acne; TOB1 has the potential to regulate cell growth; MPP7 facilitates epithelial cell polarity and tight junction formation; PEBP1 has been implicated in numerous human cancers, suppressing metastasis. All of the above may have potential correlation with acne from immunological, epidermal and cancerous development (GeneCards, https://www.genecards.org/). However, we do realize that the overlapping of DEGs between the two datasets is small. The small amount of common DEGs might imply the intricacy of the pathogenesis of acne vulgaris. (see figure 3) As it has a great reliance on exogenous factors such as desquamated keratinocytes, the result might fluctuate, hence, more sources of data should be considered in later study. The different categories of the samples from the two datasets may lead to different expression profiles. What's more, the drugs used in the three patient groups in GSE124389 might also have an influence on the gene expression, since the drugs that aim to treat acne are based on the repression of specific pathways, resulting in fluctuation of gene expression. Due to the poor distribution of the intersection of DEGs, gene enrichment analysis is executed respectively. This could be a weakness of our study and could be later improved after more datasets come into public.

We also obtained several important pathways related to acne, including phagocytosis, Fc-gamma receptor signaling pathway, and cellular senescence. As is all known, acne is an inflammatory skin disease. Most of the significant pathways are clustered at the synthesis of inflammatory cytokines and immune reaction, including IL1β, IL6, NF-κB pathway, and CXCL8(Erdei *et al*. 2018). Of the six overlapping GOs obtained from GSE124389 and GSE115099, neutrophil related terms take up four of the position, ranging from neutrophil activation to degranulation. Neutrophil ingests and kills microbes like Cutibacterium acnes, forming pustules and later acnes. Most of the pathways that we obtain confirm the current belief of the correlation between immune system, inflammation and acne. The downregulation of cellular senescence in acne patients provides potential areas of further research in the future. Despite of the chemokine signaling pathway and cytokine-cytokine receptor signaling pathway that Chen has mentioned and defense response and NF-κB pathway that Li has mentioned, we found immunoglobulin mediated immune response and cornified envelope that also play an important role in acne. Additionally, complement activation, classical pathway is significant in the formation and deterioration of acne, too.

Apart from the fact that the relationship between most pathways and acne are understood well, there are still some pathways whose roles remain to be exactly located in acne formation. For instance, a series of respiration related GO terms, including haptoglobin binding and oxygen carrier activity, were tested to be significant in GSE124389. Possible

explanations include the respiratory burst-rapid release of reactive oxygen species-in neutrophilic phagocytosis. The release of reactive oxygen would be useful in anti-acne processes. An example of this would be p53 effector. It is noted by Melnik that release of active oxygen species can induce p53 and regulate expression of negative regulators of p53 like miRNA-125b (2017).

This study inspired more specific treatments in future medical research. It is usually recommended to use topical medicine, such as salicylic acid, for sterilization and anti-inflammatory in most cases and antibiotics, retinoids or isotretinoin taken by mouth for severe acne (Titus & Hodge, 2012). However, we noticed the extreme expression difference between the patients with NCSTN mutation and the patients in GSE124389, which 3 groups of patients in GSE124389 turned out to have contrast responses to the same treatment. Some of them even got worse after treatment, though isotretinoin is almost the last plan. If different type of the expression profile from different patients are further studied, treatments could be more targeted at particular pathways, so that the treatment solution scheme would be more effective.

References

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). The Gene Ontology Consortium. GO(updated 2020). http://geneontology.org (accessed 2020)

Bataille, V., Snieder, H., Macgregor, A., Sasieni, P., & Spector, T. (2002). The Influence of Genetics and Environmental Factors in the Pathogenesis of Acne: A Twin Study of Acne in Women. *Journal of Investigative Dermatology*, *119*(6), 1317–1322. https://doi.org/10.1046/j.1523-1747.2002.19621.x

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Chen, B., Zheng, Y., & Liang, Y. (2019). Analysis of Potential Genes and Pathways Involved in the Pathogenesis of Acne by Bioinformatics. *BioMed Research International*, *2019*, 1–8. https://doi.org/10.1155/2019/3739086

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., & Xia, R. (2020). TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular Plant*, *13*(8), 1194–1202. https://doi.org/10.1016/j.molp.2020.06.009

Cole, H. N. (1930). Dyskeratosis Congenita with Pigmentation, Dystrophia Unguis And Leukokeratosis Oris. *Archives of Dermatology*, *21*(1), 71. https://doi.org/10.1001/archderm.1930.01440070079008

Database, G. C. H. G. *GeneCards®: The Human Gene Database*. https://www.genecards.org/.

Davis, E. C., & Callender, V. D. (2010). A review of acne in ethnic skin: pathogenesis, clinical manifestations, and management strategies. The Journal of clinical and aesthetic dermatology, 3(4), 24.

Erdei, L., Bolla, B. S., Bozó, R., Tax, G., Urbán, E., Kemény, L., & Szabó, K. (2018). TNIP1 Regulates Cutibacterium Acnes-Induced Innate Immune Functions in Epidermal Keratinocytes. Frontiers in immunology, 9, 2155. https://doi.org/10.3389/fimmu.2018.02155

Goulden, V., Stables, G., & Cunliffe, W. (1999). Prevalence of facial acne in adults. *Journal of the American Academy of Dermatology*, *41*(4), 577–580. https://doi.org/10.1016/s0190-9622(99)80056-5

He Y. GSE115099: Transcriptome sequencing of differentially expressed genes in acne inversa (AI) patients with NCSTN mutation and healthy individuals (updated 2019). Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE115099(accessed 2020)

Jiang Y. GSE124389: Acne flare-up in treatment process: investigating the difference of gene expression in peripheral blood (updated 2018). Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE124389(accessed 2020)

Kanehisa, M., Goto, S. Kyoto Encyoclopedia of Genes and Genomes. KEGG(updated 2020). https://www.kegg.jp (accessed 2020)

Kanehisa, M., Goto, S. GenomeNet(release 2011). https://www.genome.jp (accessed 2020)

Li, X., Jia, Y., Wang, S., Meng, T., & Zhu, M. (2019). Identification of Genes and Pathways Associated with Acne Using Integrated Bioinformatics Methods. Dermatology, 235(6), 445–455. https://doi.org/10.1159/000502203

Melnik B. C. (2017). p53: key conductor of all anti-acne therapies. Journal of translational medicine, 15(1), 195. https://doi.org/10.1186/s12967-017-1297-2

Ribero, S., Sanna, M., Visconti, A., Navarini, A., Aviv, A., Glass, D., … Bataille, V. (2017). Acne and Telomere Length: A New Spectrum between Senescence and Apoptosis Pathways. *Journal of Investigative Dermatology*, *137*(2), 513–515. https://doi.org/10.1016/j.jid.2016.09.014

Shin, J.-S., & Greer, A. M. (2015). The role of FcεRI expressed in dendritic cells and monocytes. *Cellular and Molecular Life Sciences*, *72*(12), 2349–2360. https://doi.org/10.1007/s00018-015-1870-x

Titus S, Hodge J (October 2012). "Diagnosis and treatment of acne". American Family Physician (Review). 86 (8): 734–40. PMID 23062156. Archived from the original on 18 February 2015.

Wheeler, D. L., Church, D. M., Lash, A. E., Leipe, D. D., Madden, T. L., Pontius, J. U., Schuler, G. D., Schriml, L. M., Tatusova, T. A., Wagner, L., & Rapp, B. A. (2001). Database resources of the National Center for Biotechnology Information. Nucleic acids research, 29(1), 11–16. https://doi.org/10.1093/nar/29.1.11

Younis, S., Rana, F., Blumenberg, M., & Javed, Q. (2017). Role of activated platelets in severe acne scarring and adaptive immunity activation. *Clinical Chemistry and Laboratory Medicine (CCLM)*, *55*(7). https://doi.org/10.1515/cclm-2016-0743

Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology*, *16*(5), 284–287. https://doi.org/10.1089/omi.2011.0118

Acknowledgement

Acne is an illness commonly seen in teenagers. As Chinese high school students, we see how much Acne Vulgaris has affected the mindset of teenagers, who not only suffer from pain from the outside but also become less confident on the inside. Studies about acne are frequently related to other diseases and the issue has not been valued. To get a better understanding of the issue, and to try to resolve this question with our efforts, we decided to approach this problem using bioinformatics technology.

Haofang Yang focused on the statistical computing of genomic data and pathway analysis, while Xinran Liu performed genes and pathways recognition and data visualization. Both of the students took part in writing the essay. The teacher Yuhua Li commented on the essay and provided useful opinions to the students without being paid.

We want to thank for Yuhua Li's guidance; without her, we won't be able to finish the paper.

We want to extend our gratitude to The High School Affiliated to the Renmin University of China for its support and understanding, giving us the opportunity to attend this competition.

Thank You
Liu. X & Yang. H