약학석사학위논문

Development of Gene Ontology Enrichment Analysis and its Application to Transcriptomic Profiling Analysis on Leptin-treated Human Epidermal Keratinocytes

유전자 온톨로지 기반 전사체 분석법 및 각질형성세포에 대한 렙틴의 전사체 프로파일 연구

2019 년 2 월

서울대학교 약학대학원
약학과 천연물과학전공
이 문 영

ABSTRACT

Development of Gene Ontology Enrichment Analysis and its Application to Transcriptomic Profiling Analysis on Leptin-treated Human Epidermal Keratinocytes

Hypothesis-independent knowledge mining is important in the analysis of omics data as bio-big data. Functional module-based analyses have been extensively used to extract biological insights from bio-big data. Functional modules on whole genome transcripts can be constructed with gene annotation systems in Gene Ontology (GO) terms, biological pathway maps and diverse bioinformatics gene sets. This study was aimed to develop a functional module analysis tool used for the analysis of the microarray or RNA-seq data. The functional annotation of whole genome transcripts was developed based on the gene annotation systems of both Gene Ontology (GO) Consortium and Kyoto Encyclopedia of Gene and Genome (KEGG). Statistical calculations of the GO-based or KEGG-based gene enrichment analysis was performed with *chi*-squared test or Fisher's exact test with a C-based software. The software developed in this study is divided into two parts, 'MYDEGextractor' and 'MYOnto', which are designed to run in succession. MYDEGextractor includes a function that yields DEGs, which is the core gene of transcript analysis, and MYOnto analyzes the functional tendency of DEGs through gene ontology annotations. Compared with currently available gene ontology enrichment analysis softwares, the MYOnto software has the following advantages: (a) it analyzes through multiple functional annotation frameworks simultaneously

and (b) it performs the integrated process from DEG selection to ontology analysis, which can be applied to automated bio-big data analysis.

The functional module analysis with the MYOnto software was performed with the genome scale transcriptional data on the leptin-treated human epidermal keratinocytes. From the oligonucleotide-based microarray data on the leptin-treated keratinocytes, 151 leptin-induced upregulated DEGs and 53 downregulated DEGs were extracted. The gene ontology enrichment analysis with the leptin-treated DEGs showed that inflammation-associated genes were significantly enriched in the leptin treated human keratinocytes. The upregulation of the inflammation-associated genes such as CXCL8 and MMP1 was inhibited by a STAT3 inhibitor, suggesting that leptin induced inflammation is regulated by a STAT3 signaling pathway in human epidermal keratinocytes.

keywords : Functional Genomics, Gene Ontology Enrichment Analysis, Bio-big data, Keratinocytes, Leptin

Student number **:** 2017-23557

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Part I.

Development of Computational Methodology on Gene Ontology Enrichment Analysis

Ⅰ. Introduction

Hypothesis-independent biological insight mining from the structural features of omics data itself is the major issue of functional genomics (1). Genome-wide association studies (GWAS) became common at biological research as the cost of genomic methodologies has decreased (2). For example, oligonucleotide microarray analysis was used to elucidate the molecular mechanism of different skin irritating compounds, showing that genes at chromosomal locus 1q21, which is closely related to cornified envelope, display specific expression pattern to each compounds (3). In addition, bio-big data accumulated by researchers provide an opportunity to discover previously unknown biological meanings even with limited expenses. A new biological insight can be derived by creatively meta-analyzing the various microarray and RNA-seq results provided by bio-big data archives such as National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) (4). For example, identification of genes and proteins commonly involved in epidermal hyperproliferation and skin barrier damage has been performed using databases relating to skin diseases such as atopic dermatitis, psoriasis, allergy, and acne (5).

In order to derive biological insights from whole genome-scale expression data, it is important to define gene ontology (GO) accurately in terms of functional aspects and interaction networks. It is necessary to construct integrated bioinformation by biological pathways for the physiological objects of interest to perform meta-analysis of omics data. Integrated bioinformation for each biological pathway is mainly used to extract data from the open omics database quickly and is essential for

the purpose-oriented extraction of information on biological phenomena. An example of a systematic integrated bioinformatics library is the Kyoto Encyclopedia of Genes and Genomes (KEGG). KEGG has established a system that can be linked to NCBI and EBI within its pathway library, which includes metabolism, genetic information processing, environmental information processing, cellular processes, structures, organismal systems, and human diseases to help understand the molecular interaction networks that are the basis of living organisms (6). In this study, a computational tool for biological insight mining to perform ontology-based bio-big data analysis using integrated bioinformation libraries were developed.

Ⅱ. Materials and Methods

1. Preparation of whole genome-scale expression data

Microarray data were used as samples to evaluate the performance of the software. The test data GSE36700 comparing rheumatoid arthritis to other diseases such as osteoarthritis and systemic lupus erythematosus are downloaded from NCBI GEO (www.ncbi.nlm.nih.gov/geo). Expression data are extracted from raw data image file (.CEL) using Expression Console Software, Version 4.0 (Affymetrix, Santa Clara, CA, USA) based on MAS 5.0 algorithm.

2. Automated selection of differentially expressed genes

Six methods which have been commonly used in combination to select differentially expressed genes (DEGs) were adopted: 1) to exclude genes of fold change (ratio of expression data) below specific value, 2) to exclude genes whose expression value of both control and experimental group are below median, 3) to calculate *p*-value by Student's t test of expression values between control and experimental group and select genes of enough statistical significance, 4) to choose genes with P (present) and M (marginal) call, rather than A (absent) call, 5) to select specific number of genes from top/bottom by fold change, and 6) to select specific percentage from tom/bottom by fold change. In addition, deleting overlap of genes,

corresponding to same gene symbol but different probes in microarray data, is necessary during the process of identifying DEGs. Those processes are made to be performed automatically using C language and Visual Studio®, Version 2010 (Microsoft Corp., Redmond, WA, USA).

3. Gene function annotation frameworks

There are numerous reports to study biological function of genes. However, reflecting all those reports to statistical model is extremely difficult and inefficient. To grasp functional tendency of specific group of genes, each gene was needed to be annotated for classified biological function. In commercial microarray platforms, manufacturers often provide gene ontology (GO) information, yet researchers need to have a system that can utilize GO information on their own in RNA-seq data. A new "functional module" by grouping genes that are selective and correlated with specific biological phenomena was aimed to be provided. Therefore, know-how to organize functional annotation framework provided by Gene Ontology Consortium (http://www.geneontology.org) and Kyoto Encyclopedia of Genes and Genomes (https://www.genome.jp/kegg) was constructed. For example, it is needed to include processes which remove overlap of GO term, considering that some genes are annotated up to dozens of same GO Biological Process term. In this regard, processes which automatically take only one GO term to each genes were constructed. Likewise, the methods to reflect gene function information to transcriptomics data analysis were found and progressed continuously.

4. Statistical modeling for GO enrichment analysis

Statistical modeling adopted in this study evaluates GO information annotated to DEGs with that to the whole genome, determining significance of the degree that each GO is annotated to relatively a number of genes, i.e., "enriched". First, extracted the number of DEGs were annotated to specific GO, not annotated to specific GO, and the same numbers in whole genome. Those would be the elements of 2 × 2 contingency table for further statistical analysis (Table 1). Two statistical methods were adopted: Fisher's exact test and Chi-squared test, both of which calculate *p*-value with null hypothesis that GO annotation is randomly distributed to all genes. Two method yield different values in spite of same tendency, mainly due to the gap between discretized and continuous function. Fisher's exact test is mainly for GO terms containing small (below 5) number of DEGs, whereas Chi-squared can be used to those of large number of DEGs (Fig. 1). Through the *p*-value, it was possible to discriminate whether specific functional unit is enriched in up- or down-regulated DEGs.

Table 1 Concept of contingency table for Gene Ontology enrichment analysis.

In the analysis on DEGs in one transcriptomics data set, contingency table has to be made for each Gene Ontology term. The value *a* and *b* are the number of genes annotated with specific GO term which are included and not included in DEGs, respectively. The value *c* and *d* are the number of genes not annotated with specific GO term which are included and not included in DEGs, respectively. R_1 and R_2 are the number of DEGs, i.e. $a + c$, and the number of genes except DEGs, i.e. $b + d$, respectively. C_1 and C_2 are the number of genes annotated with specific GO, i.e. $a + b$, and the number of genes not annotated, i.e. $c + d$, respectively. N is the number of whole genome (at target platform).

2 × 2 contingency table	# of genes annotated with specific GO	# of genes non-annotated with specific GO	Total
# of DEGs	*a*	*c*	R_1
# of non-DEGs	*b*	*d*	R_2
Total	C_1	C_2	N

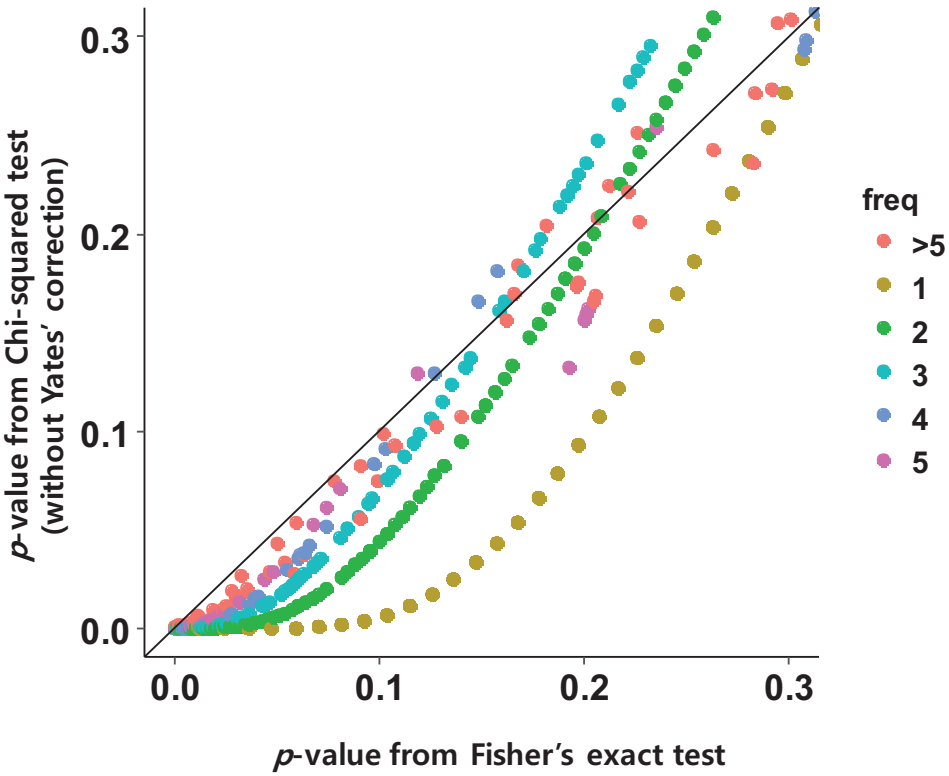


Figure 1. Difference between *p*-values from Fisher's exact test and Chi-squared test.

Colored dots indicate *p*-values calculated by two methods changing the number of DEGs, specific GO-annotated genes, and specific GO-annotated DEGs. x axis is *p*-value calculated from Fisher's exact test, and y axis is that from Chi-squared test. The color of the dots shows the number of GO-annotated DEGs of each dot. Skewed line represents the equation 'y = x', which means two *p*-values are the same if a dot is on this line. Two *p*-values get closer as the number of GO-annotated DEGs get bigger.

5. Automation of statistics on whole annotated GO terms

Applying statistical models based on Fisher's exact test and chi-squared test to all GO terms annotated to DEGs is time-consuming and simple repetitive task. There are thousands of annotated GO terms to count the number of related genes and calculate *p*-value. In order to automate this work, a software were made to search all GO annotation and calculate *p*-value using C language and Visual Studio®, Version 2010 (Microsoft Corp.).

Ⅲ. Results

1. Major features of software developed in this study

An integrated software package was developed to derive biological insights from transcriptomics data (Fig. 2). The package is composed of two parts: MYDEGextractor for DEG selection and MYOnto for gene ontology enrichment analysis. Two programs can be run sequentially, as well as MYDEGextractor can be used alone to get DEG list, and MYOnto can be used to perform GO enrichment analysis from existing DEG list.

The software performs statistical analysis using gene function information from two frameworks: Gene Ontology Consortium, and Kyoto Encyclopedia of Genes and Genomes (KEGG). Both frameworks provide biological pathway, function, and phenomena related to genes, but there are far differences between each other. In order to test the software and identify the differences between two gene information frameworks, a microarray data set from NCBI GEO were downloaded and performed analysis through the software (Table 2). GSE36700, dataset used for the test analysis, was constructed to identify distinct gene expression profiles about rheumatoid arthritis, systemic lupus erythematosus, and osteoarthritis (7). DEGs upregulated in rheumatoid arthritis samples comparing to osteoarthritis ones were determined and analyzed. As a result of analysis by both frameworks, rheumatoid arthritis obviously appeared to activate genes related to immune response. However, detailed Gene Ontology term showed differences in two analysis results. Referring to the information of Gene Ontology Consortium, adaptive immune system was the

most activated in rheumatoid arthritis, while pathways related to specific cytokines as well as immune-related disorders are listed in KEGG pathway analysis. In addition, immunoglobulin elements such as IGLV1-44, IGHM were appeared as major upregulated DEGs in the analysis based on Gene Ontology Consortium, but KEGG pathway analysis indicates other genes such as IL21R, CD3D, and HLA-DOB.

The know-how through the use of the initial version of the software analyzing multiple transcriptomics data was established, which enabled deriving biological insights with strong statistical basis through the use of more sophisticated analytical techniques. These sophisticated analysis techniques include not only software advancement, but also the development of analytical methods. Each microarray and RNA-seq data can be characterized according to the expression value profile, and appropriate analytical methods need to be introduced according to these characteristics. This data interpretation strategy is chosen according to the researcher's experience in the direction that can give the best result. In many cases, it is necessary to analyze the same data several times and check the result and optimize it. However, this type of data analysis can take a long time depending on the nature of the data and the purpose of analysis, and it is also difficult for beginners who do not have bioinformatics background to learn. In addition, due to differences in minor habits that may often be ignored in the data analysis process, or mistakes in the analysis process, researchers may produce slightly different results even with the same data. Therefore, standardizing these processes and constructing an integrated analysis environment for rapid information processing were focused on.

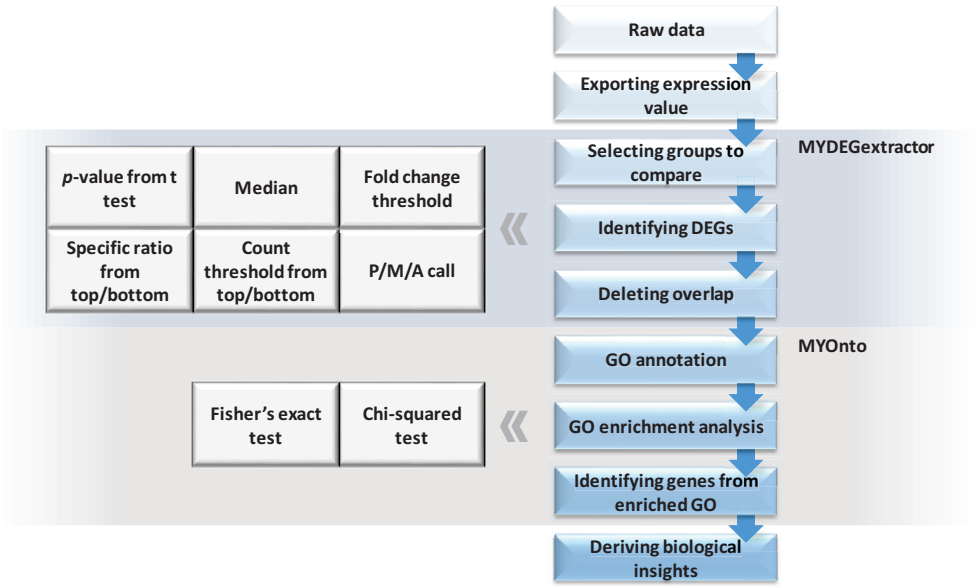


Figure 2. Skim of transcriptomics data analysis and coverage of the software MYDEGextractor and MYOnto.

Expression values have to be prepared from microarray or RNA-seq raw data. After selecting groups to compare, DEGs can be identified through several selection methods. MYDEGextractor was made for processing by the next step, deleting overlaps of same genes. DEG list prepared from MYDEGextractor is needed for operation of the other software. MYOnto reads DEG list file and extract information about Gene Ontology information. Then it performs GO enrichment analysis for all GO terms annotated in DEGs. As MYOnto yields enriched GO list and related gene list, biological insights can be gotten from the analysis.

Table 2 Gene Ontology enrichment analysis on same data based on two different gene function information frameworks.

406 upregulated genes were identified from GSE36700 comparing the gene expression pattern of osteoarthritis and rheumatoid arthritis patients. GO enrichment analysis were performed based on GO Consortium and KEGG information through MYOnto. Top 10 Gene Ontology terms are listed in order of *p*-value by Fisher's exact test. *p*-values were calculated to 10th decimal place, and listed as '<E-10' if below 10^{-10}.

Enrichment analysis based on Gene Ontology Consortium

GO code	up DEGs 406	total 25415	gene ontology biological process	p-value (Fisher's exact test, 2-sided)	gene symbol
GO:0006955	50	339	immune response	<E-10	IGLV1-44(114.85), IGLC1(35.03), CXCL13(30.
GO:0002250	29	157	adaptive immune response	<E-10	IGHM(39.37), CD79A(23.51), TNFRSF17(17.59
GO:0050776	28	148	regulation of immune response	<E-10	IGLV1-44(114.85), IGLC1(35.03), IGKC(24.12)
GO:0050853	18	48	B cell receptor signaling pathway	<E-10	IGHM(39.37), IGLC1(35.03), IGHD(29.45), IGI
GO:0050852	23	127	T cell receptor signaling pathway	<E-10	TRBC1(12.93), CD3D(6.86), LCK(6.52), TRAT1
GO:0007166	22	223	cell surface receptor signaling pathway	<E-10	CXCL13(30.42), CXCL10(11.11), CD27(9.25), C
GO:0070098	13	67	chemokine-mediated signaling pathway	<E-10	CXCL13(30.42), CXCL9(21.65), CXCL10(11.11
GO:0006954	27	369	inflammatory response	1E-10	CXCL13(30.42), CXCL9(21.65), CXCL10(11.11
GO:0006935	15	121	chemotaxis	1E-09	CXCL9(21.65), CXCL10(11.11), CCL5(6.29), CX
GO:0045087	26	398	innate immune response	1.6E-09	IGHM(39.37), IGLC1(35.03), IGHD(29.45), IGI

Enrichment analysis based on KEGG

GO code	up DEGs 406	total 25415	KEGG pathway	p-value (Fisher's exact test, 2-sided)	gene symbol
hsa04659	20	116	Th17 cell differentiation	<E-10	IL21R(6.91), CD3D(6.86), HLA-DOB(6.77), LC
hsa04658	19	105	Th1 and Th2 cell differentiation	<E-10	CD3D(6.86), HLA-DOB(6.77), LCK(6.52), IL2F
hsa04650	19	140	Natural killer cell mediated cytotoxicity	<E-10	LCK(6.52), ITGAL(5.52), GZMB(4.57), ZAP70
hsa05340	12	42	Primary immunodeficiency	<E-10	CD79A(23.51), CD8A(8.11), CD3D(6.86), LCK(
hsa04064	16	102	NF-kappa B signaling pathway	<E-10	LTB(4.30), CSNK2A1(3.89), ZAP70(3.24), CCL
hsa04062	21	205	Chemokine signaling pathway	<E-10	CXCL13(30.42), CXCL9(21.65), CXCL10(11.11
hsa04640	15	101	Hematopoietic cell lineage	1E-10	CD8A(8.11), HLA-DOB(6.77), CD38(5.83), IL7
hsa04672	11	48	Intestinal immune network for IgA productio	2E-10	TNFRSF17(17.59), HLA-DOB(6.77), HLA-DRB
hsa04514	16	149	Cell adhesion molecules (CAMs)	2.4E-09	HLA-DOB(6.77), ITGAL(5.52), SDC1(4.79), HI
hsa05162	15	144	Measles	1.14E-08	CD3D(6.86), IL2RB(5.18), IL2RG(4.86), SH2D1

2. Procedure for operating MYDEGextractor and MYOnto

MYDEGextractor is a program that extracts DEGs from transcriptomics data of microarrays or RNA-seq. Depending on the intention of the researcher, the DEGs selection method and range can be adjusted, and then MYOnto program can be run on its own.

MYDEGextractor demands 3 files: raw data, control, and experimental files (Fig. 3). Each file contains information about sample name, subgroup of the sample, and expression value of all genes. Control and experimental files indicate groups to compare. Control group can be normal cells, sample from healthy volunteers, unaffected tissue, or non-lesion. Experimental group can be specific compound-treated cells, sample from patient, affected tissue, or lesion from same person as control sample. MYDEGextractor read those file and analyze change of expression pattern ultimately yielding DEGs.

The most important characteristic to be considered when selecting DEGs from transcriptomics data is the level how the gene expression pattern of the control and experimental groups is different. As the data of the two groups shows a remarkable difference, the condition of selecting DEGs is complicated. By contrast, if there is only a little difference, a wide range of genes should be classified as DEGs. In order to understand the characteristics of a specific transcript database, it is necessary to find the condition for obtaining the optimal result from GO enrichment analysis by changing condition and obtaining DEG candidates several times. MYDEGextractor has several modes for obtaining DEGs, and it includes functions to quickly select DEG candidates and to confirm or re-select DEGs.

Once the DEG selection has been completed using the MYDEGextractor, the selected DEGs can be found in the up-DEGs and down-DEGs files in .csv format. Each DEG information file contains the conditions applied for DEG selection in order, so researchers can refer to it when wanting to analyze other data under the same conditions or change the conditions to analyze the same data.

After MYDEGextractor is completed, if MYOnto program exists, it can be executed in conjunction with the program. MYOnto performs enrichment analysis by referring to the gene annotations according to the Gene Ontology Consortium and the KEGG pathways for up-DEGs and down-DEGs, respectively, and generates four files. In each result file of MYOnto, the ontology code, the number of DEGs included, and the number of genes containing the corresponding ontology, the *p*-value according to the Fisher's exact test and the chi-squared test, the corresponding gene symbol and fold change, and the functions implied by the ontology are listed.

Researchers may also want to perform GO enrichment analysis on already determined DEGs. In this case, MYOnto can be run alone if the researchers prepare the input data manually. MYOnto reads up-DEGs and down-DEGs file, and ontoinfo file which consists of information about Gene Ontology annotation of all genes in the target microarray or RNA-seq platform. Up- and down-DEGs files should match the standard form of MYDEGextractor result files (Fig. 4). It also needs 3 data files; chisq_cum_define, term_GO, and term_KEGG files, which are included in the software package. Once all files needed are prepared, researchers can get the GO enrichment analysis result files by just single operation.

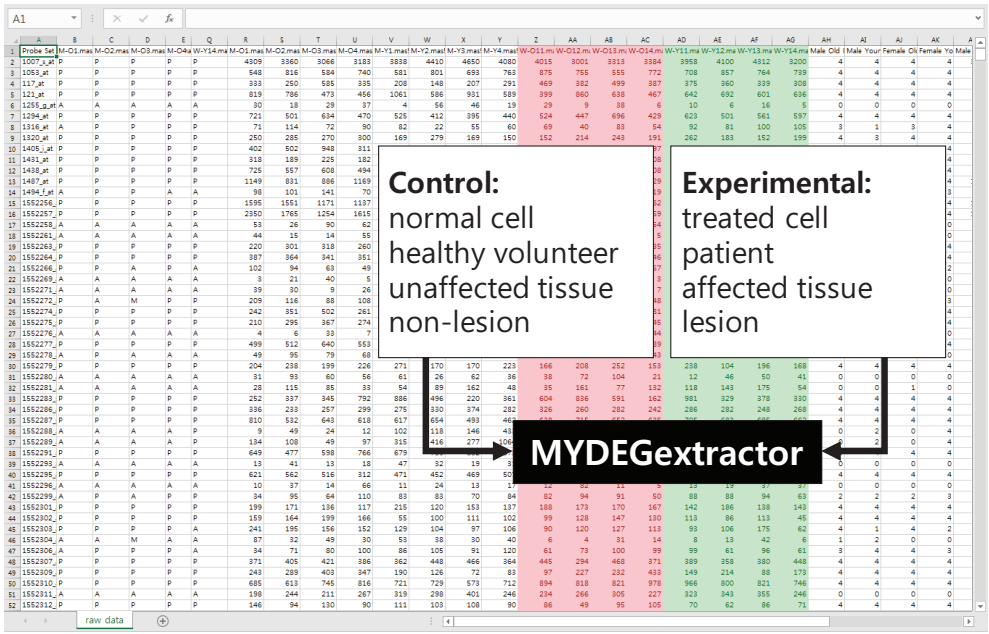


Figure 3. Input data for MYDEGextractor.

Raw data file consists of whole genome-scale expression data of all groups. Among them, control and experimental group should be determined through control and experimental files. MYDEGextractor recognize those files to compare each other and identify differentially expressed genes (DEGs).

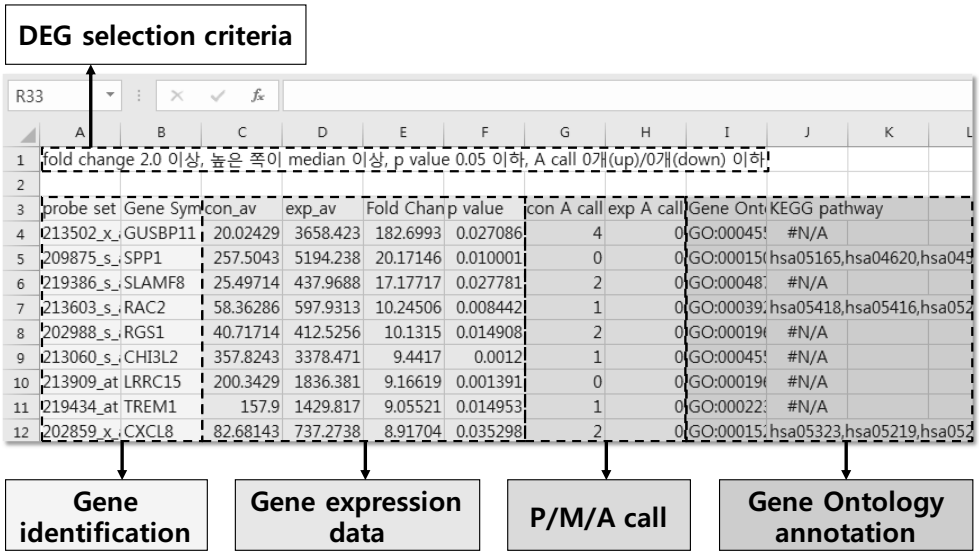


Figure 4. Standard format of up- and down-DEGs list file.

The DEG list can be obtained from the transcriptomics data via MYDEGextractor, which is also the input data of the MYOnto. It contains information about DEG selection criteria, gene identification code and symbol, expression data, present (P) or absent (A) calls, and annotated Gene Ontology.

3. Data interpretation

MYDEGextractor provides DEG list files which can be also read by MYOnto (Fig. 4). In the DEG list files, DEG selection criteria are noted first. They function as the records of the software operation during DEG selection; whenever a researcher adds a criterion on selecting DEGs, MYDEGextractor mentions what process was performed in order of the performance. Gene information, which consists of probe set ID (microarray) or gene ID (RNA-seq), gene symbol, average expression value of control and experimental groups, fold change, *p* value by Student's t test between expression values of two groups, P/M/A calls, and Gene Ontology annotated by Gene Ontology Consortium and KEGG, is listed from line 4. The genes are listed in order (up-DEGs) or reverse order (down-DEGs) of fold change so that researchers can identify the most dramatically changed ones. If different probes in microarray platform indicating same gene were commonly classified as DEGs, repeated ones would be removed except one with fold change far from 1. Researchers should be aware that genes which have many probes in microarray platform can be listed as higher-ranked DEGs than they really are.

The products of MYOnto are also CSV list files, which include list of enriched GO terms (Fig. 5). Gene number information on the top of the list indicates the number of DEGs and whole genome in microarray or RNA-seq platform, which are commonly needed to fill 2 × 2 contingency tables (Table 1) on all GO terms. Then enriched GO terms are listed in order of *p*-value indicating significance of enrichment calculated by Fisher's exact test. The list provides information about GO codes, the numbers of specific GO-annotated genes in DEG group and whole

genome, *p*-values calculated by Fisher's exact test and chi-squared test, gene symbols and corresponding fold changes of the genes which are specific GO-annotated as well as classified as DEGs, Gene Ontology terms, and the class of GO terms if they are based on Gene Ontology Consortium. The *p*-values are provided to the 10^{th} decimal place, and displayed as 0 if they are less than 10^{-10}. Gene symbols are displayed to 100^{th} significant GO terms for fast operation of the software. There are 3 classes of GO terms in Gene Ontology Consortium: Biological Process (BP), Cellular Component (CC), and Molecular Function (MF). If a researcher wants to grasp enriched biological processes only, GO classes can be filtered to show only GO BP terms.

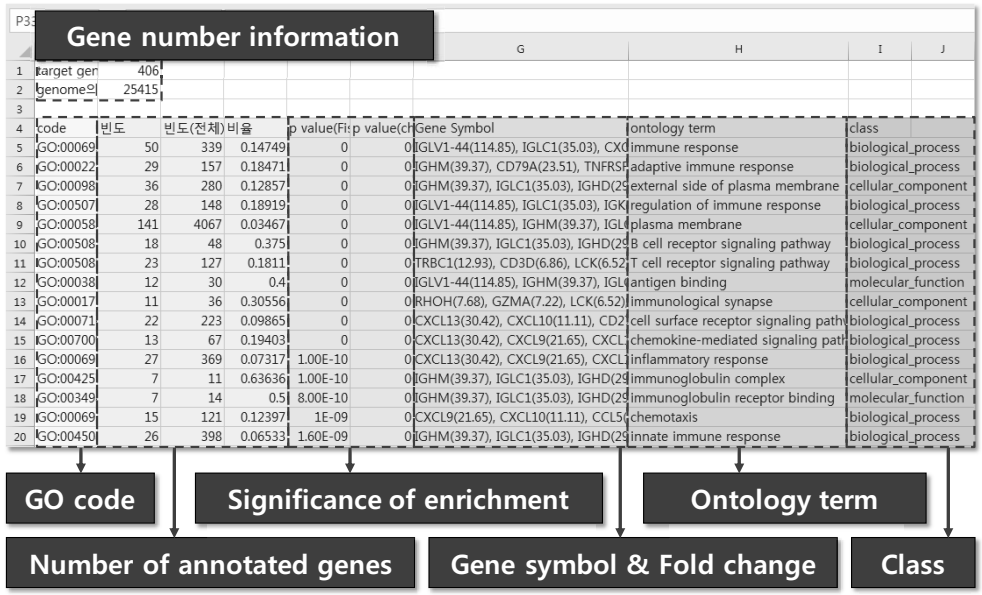


Figure 5. Format of result files yielded from MYOnto.

MYOnto yields four result files based on Gene Ontology Consortium and KEGG, and on up- and down-regulated DEGs. This figure is an example of Gene Ontology Consortium and upregulated GO term list file. It includes gene number information, Gene Ontology information, significance level, and gene information related to enriched GO terms. The other 3 result files also have same formats.

4. Updating functional gene information

Knowledge about biological function is continuously being accumulated. GO Consortium and KEGG websites have updated their functional gene information annotation for many years. In order to reflect the update, new information about functional genomics should always be monitored

The information which is needed to be updated is divided into two kinds. First, GO annotation has to be included in raw data as the latest one. Gene Ontology Consorium provides annotation file of several species including human, and KEGG provides lists of genes included in KEGG pathway maps. Before executing MYDEGextractor, Gene Ontology information should be checked for update from those websites.

Another one to be updated is the relationship between GO codes and names. Gene Ontology information exists in the raw data file as seven (GO Consortium) or five (KEGG) digit GO codes. When MYOnto determines enriched GO terms, GO codes are sufficient to calculate *p*-values based on statistical models. Since pathway name of each GO term is needed as final interpretation of the data, MYOnto matches each GO code and name referring to a correspondence table. The table has saved as data file 'term_GO' and 'term_KEGG'. Each data file can be revised through text editor or spreadsheet programs, and should be saved as text-separated .txt file and then changed the filename extension to .dat format. Original source of the correspondence table can also be extracted from GO Consortium and KEGG websites. Since the update of the correspondence of pathway names and codes is far rarer than that of annotation, update check may be performed every few months or annually.

Ⅳ. Discussion

An integrated package has been aimed at to help discover biological insights from disordered transcriptomics data. Algorithm that explores common functions from specific gene groups and expresses them in order of significant level through statistical modeling has also been developed in bio-big data analysis platforms such as Broad Institute (8, 9). However, statistical analysis from whole genome-scale expression data to select the core gene to be applied and to link it with the statistical analysis program is the first attempt. Currently, the programs are continuously evolving to add features that are needed during transcriptomics data analysis and to develop stable and fast data analysis through optimization processes.

Gene Ontology terms indicate either broad (high level) or narrow (low level) biological processes (10). Specificity problem of Gene Ontology terms is a major barrier of interpreting biological meanings from DEGs. A representative platform of human genome microarray, Affymetrix Human 133 2.0 has a total of 23260 genes except redundancy to which GO biological process were defined. For example, 2485 genes are defined as GO:0006355, and 2330 genes are defined as GO:0006351 (Table 2). Since the public bioinformation analysis system such as DAVID utilizes the incompletely constructed GO BP, it cannot obtain the results for the meaningful functional unit when applied to the experimental system actually studied. Gene Ontology Enrichment Analysis results from the PANTHER Classification System (http://www.pantherdb.org/pathway/) tend to be abstract. As shown below, the top 10 GO BP terms are not specific enough to actually determine the function of an

unknown gene (Table 3). Therefore, statistical methodology need to be applied to extract significantly common feature of DEGs without bias. Nonetheless, extensive care was required treating whole genome-scale expression data using computational method. For example, among the factors to be considered from the GO DB, it is necessary to build a process to remove redundant GO Term considering that several genes are defined redundantly from one to many Gene Ontology Biological Process Term. For example, AGPAT1 gene is defined as 15 GO BP Term. AGPAT2, AGPAT3, and AGPAT4 genes, which are highly similar to AGPAT1, are also defined as multiple GO BP Terms, but GO BP Terms of similar genes are not completely identical. As of December 2018, the GO Consortium DB (http://geneontology.org/) provides GO information on 19675 genes coding for proteins. A process for eliminating redundant GO Term for GO enrichment analysis has been developed.

Gene Ontology annotation is rapidly being updated in line with the progress of knowledge on an ever-increasing individual gene (10). For example, the information on human genes published on the Gene Ontology Consortium (http://www.geneontology.org/) has been updated more than 10 times since 2014 as of the last update on September 26, 2017. Therefore, in analyzing transcript data, it is necessary to always check the gene database information and analyze it with the latest information. If necessary, re-analyze the existing data through a new database to Gene Ontology annotation reflected. In order to reflect the new gene information revealed through recent research, the annotations of Gene Ontology were researched whenever it is updated and have the latest database in the form of genome data analysis.

Since the initial version of this program focuses on calculating the *p*-value, there was lack of function to remove redundant genes in the entire gene group. In Affymetrix GeneChip 133 2.0 microarray platform, annotation information of the same GO term appears in duplicate in one probe set. In this case, there is a problem in calculating the GO term significantly enriched in DEG. This kind problem is occurred mainly if the same biological phenomenon is reported in multiple documents. If the data file provided by the GO Consortium DB is used as it is without removing duplication of the GO BP term for this gene, an error of low *p*-value occurs. Therefore, it was necessary to develop an efficient process that can quickly remove duplication of GO BP term for individual genes from the original data provided by GO Consortium DB. A system that can automatically construct the GO BP term for the entire gene in the genome without duplicate errors or deletion errors every time the database is updated was developed.

In addition to redundancy problem, GO Enrichment Analysis should be improved to reflect qualitative evaluation: in other words, among several GO BP terms that are annotated in one gene, weight scores should be applied to a specific GO BP that has been experimentally verified, and a relatively low weight should be assigned to bioinformatively predicted GO BP. Functionally more important is the GO BP term of the gene whose functional background information is supported by well-documented and diverse experimental literature data. There is a need to establish a method for assigning weights to certain individual genes based on duplicate documents. A solution to these issues by updating the software and bioinformatics library is being found.

Another area to be improved is the establishment of a comprehensive gene function information system. Since no Gene Ontology annotation framework perfectly reflects knowledge elucidated about human genes, a system through which researchers can refer to alternative two GO frameworks was developed. However, increasing problems which functional GO frameworks basically involve to give solution through software upgrade have been found. As a representative example, the cholesterol biosynthetic process (GO:0006695) has to be included in the cholesterol metabolic process (GO:0008203) in semantics, and the Gene Ontology Consortium also discloses it. However, cholesterol biosynthetic processes have been annotated in many genes, including Acetyl-CoA Acetyltransferase 2 (ACAT2), but the cholesterol metabolic process has not been commented on. That is, the function genome information database provided by the Gene Ontology Consortium is not systematically layered. Consequently, own functional annotation model for application to GO enrichment analysis is being tried to construct. As part of this attempt, in order to automatically analyze the texts written in natural language after downloading abstracts of many documents from the life science and medical literature search engine PubMed (https://www.ncbi.nlm.nih.gov/pubmed/), GDMiner, a software developed by Electronics and Telecommunications Research Institute, was tried to be utilized. GDMiner extracts the meanings of each sentence from the papers searched by a specific keyword, and then graphs the relation between the papers. Ultimately, text mining and artificial intelligence should be applied to compensate the defects inevitably produced by manual annotation systems.

Table 3 Top 25 Gene Ontology Biological Process terms in Affymetrix Human 133 2.0 microarray platform.

GO BP terms which indicate abstract and general biological functions tend to be annotated to a number of genes. For example, "DNA-dependent regulation of transcription (GO:0006355)" is annotated to 2,485 genes, which is top ranked of all GO BP terms.

Human GeneChip 133 2.0 (Total 54613 probe sets, 23260 GO annoated genes, 9410 No GO annotation)		
Gene Ontology ID #	# of GO annotated genes	Ontology items
GO:0006355	2485	regulation of transcription, DNA-dependent // inferred from electronic annotation
GO:0006351	2330	transcription, DNA-dependent // inferred from electronic annotation
GO:0007165	1862	signal transduction // traceable author statement
GO:0006810	1766	transport // non-traceable author statement
GO:0044281	1477	small molecule metabolic process // traceable author statement
GO:0007275	1078	multicellular organismal development // inferred from mutant phenotype
GO:0006915	897	apoptotic process // inferred from electronic annotation
GO:0045944	848	positive regulation of transcription from RNA polymerase II promoter // inferred from mutant phenotype
GO:0008152	785	metabolic process // inferred from electronic annotation
GO:0016310	779	phosphorylation // inferred from electronic annotation
GO:0055085	723	transmembrane transport // inferred from electronic annotation
GO:0010467	720	gene expression // inferred from mutant phenotype
GO:0055114	709	oxidation-reduction process // inferred from electronic annotation
GO:0045087	703	innate immune response // inferred from mutant phenotype
GO:0030154	690	cell differentiation // inferred from electronic annotation
GO:0015031	680	protein transport // inferred from electronic annotation
GO:0006468	660	protein phosphorylation // inferred from electronic annotation
GO:0006811	645	ion transport // inferred from electronic annotation
GO:0007155	642	cell adhesion // inferred from electronic annotation
GO:0007186	641	G-protein coupled receptor signaling pathway // inferred from electronic annotation
GO:0045893	638	positive regulation of transcription, DNA-dependent
GO:0007049	628	cell cycle // inferred from electronic annotation
GO:0043066	621	negative regulation of apoptotic process // not recorded
GO:0000122	620	negative regulation of transcription from RNA polymerase II promoter // inferred from mutant phenotype
GO:0006508	609	proteolysis // inferred from electronic annotation
GO:0044267	551	cellular protein metabolic process // inferred from electronic annotation
GO:0045892	519	negative regulation of transcription, DNA-dependent // not recorded
GO:0006629	517	lipid metabolic process // inferred from electronic annotation

Part II.

Elucidation of the Role of Leptin on Human Epidermal Keratinocytes

Ⅰ. Introduction

Leptin is a hormone which is known to control appetites and energy homeostasis. It is mainly produced from adipose tissue, and should bind to the leptin receptors in the hypothalamus to activate its function. Leptin receptors are classified as IL-6 receptor family which have no kinase activities. In a cascade initiated by leptin receptors, major signal activation is induced by phosphorylation of tyrosine residues in leptin receptors mediated by JAK2, and STAT3 gene is known to be an important mediator (11). Phosphorylation of leptin receptors is related to activation of PI3K and ERK1/2, and further related to STAT3 or other adaptor molecules which can cause numerous changes according to circumstances (12). Leptin can also affect to various peripheral tissues including bone, immune cells, and wound in an autocrine or paracrine manner (13-16).

Leptin secreted from non-adipose tissue, such as hair follicle papilla cells, may regulate biological function in skin tissue (17). Cultured human fibroblasts also produce leptin and leptin receptor which are regulated by insulin (18). In psoriasis patients, the levels of leptin and leptin receptors are reported to be closely related to the severity of the disease (19). Pro-inflammatory mediators such as IL-6 and TNFα may induce leptin and leptin receptor expression in the severe psoriasis patients (19, 20). In addition, leptin promoted mitosis in epithelial keratinocytes, which indicates the role of leptin in wound healing (21). Meanwhile, obese individuals generally show a degeneration of re-epithelization (22). This is very interesting because leptin in obese individuals tends to be increased (23-25). The leptin resistance of obese

people, especially in the epithelium, cannot be clearly explained since much of the mechanism of leptin to epithelial keratinocytes is not understood.

In order to elucidate the molecular mechanism of leptin to keratinocytes, microarray analysis was adopted and genome-wide change of leptin treatment was observed. As a result of Gene Ontology enrichment analysis, leptin showed upregulation of pro-inflammatory response and downregulation of some metabolic processes. Several genes related to wound healing were identified and confirmed by experimental methods.

Ⅱ. Materials and Methods

1. Primary keratinocyte culture and total RNA isolation

NHKs from neonatal foreskin were obtained from Lonza (Basel, Switzerland) and cultured according to the manufacturer's instructions. Briefly, NHKs were maintained in keratinocyte basal medium (KBM) with a KGM2 complete growth factor cocktail consisting of insulin, human epidermal growth factor, bovine pituitary extract, hydrocortisone, epinephrine, transferrin, and gentamicin/amphotericin B (Lonza). Three independent NHK batches prepared from different donors were purchased, and all experiments were performed in triplicate or quadruplicate. Frozen NHKs were thawed and passaged at 80–90% confluence for cell expansion. At the second passage, leptin (Sigma, St Louis, MO, USA) was used to treat 90% confluent NHKs in all experiments, including the oligonucleotide microarray and validation studies. Total RNA was harvested using Trizol reagent (Invitrogen, Carlsbad, CA, USA). RNA was further purified using an RNeasy Mini Kit (Qiagen Inc., Hilden, Germany). The concentration of total RNA in each sample was measured using a NanoDrop ND-1000 spectrometer (NanoDrop Technologies Inc., Montchanin, DE, USA). RNA integrity was determined with a BioAnalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). For the cell growth kinetic study, the number of cells was counted using a Coulter counter (Beckman Coulter Life Sciences, Indianapolis, IN, USA).

2. Microarray analysis and DEG selection

For the microarray analysis, total RNA was harvested from two independent NHK cultures. Microarray experiments were performed by a local authorized Affymetrix service provider (DNA Link, Inc., Seoul, South Korea). Affymetrix GeneChip Human Genome U133 Plus 2.0 oligonucleotide arrays were used for the genome-scale transcriptional analysis (Affymetrix, Santa Clara, CA, USA). The procedure for performing the Affymetrix GeneChip assay on biotin-labeled RNA by in vitro transcription, hybridization, array washing, and scanning was performed according to the manufacturer's protocol. The hybridized probe array was stained and washed using a Fluidics Station 450 (Affymetrix), and the stained GeneChip probe array was scanned with a GeneChip Scanner 3000+7G (Affymetrix). The signal intensity for each probe set was calculated using Expressin Console Software (Affymetrix) based on the MAS 5.0 algorithm. Differentially expressed genes (DEGs) were selected through MYDEGextractor as follows: (i) selection of "present" Affymetrix probe sets with hybridization partners in control or leptin-treated samples; (ii) selection of Affymetrix probe sets indicating a signal sample/control ratios greater than 2 as "upregulated" and less than 2 as "downregulated"; and (iii) selection of genes targeted by Affymetrix probe sets that resulted in significant *p*-values (threshold, 0.1) by Wilcoxon rank test when compared with genes in vehicle-treated samples.

3. Gene Ontology enrichment analysis

A GO enrichment analysis of leptin-induced DEGs was performed by comparing the frequency of GO biological process (BP) terms assigned to each gene in a group of DEGs to that in the entire set of genes in the Affymetrix Human 133 2.0 GeneChip array. The GO annotation files were downloaded from the Gene Ontology Consortium webpage (http://www.geneontology.org). In the analysis, the August 2016 version of the GO BP terms and MYOnto version 1.02 were used for the GO enrichment analysis. When redundant probe sets in the Affymetrix Human 133 2.0 GeneChip were considered probes recognizing single gene units in the GO enrichment analysis, 23,624 was considered the total number of genes analyzed. A 2 × 2 contingency matrix was constructed to determine the frequency of a specific GO BP term in a group of DEGs compared with that in the full set of 23,624 genes. The 2 × 2 contingency matrix data were analyzed by Fisher's exact test (frequency < 5) or chi-squared test (frequency > 5) to calculate the level of significance.

4. Quantitative real-time reverse transcription polymerase chain reaction

Results of the microarray analysis of DEG expression was validated by Q-RT-PCR. The TaqMan probes (Applied Biosystems, Foster City, CA, USA) used in the Q-RT-PCR analysis targeted matrix metalloproteinase 1 (MMP-1, Hs00899658_m1), fibronectin (FN, Hs01549937_m1), CXCL8/IL-8 (Hs00174103_m1), S100A7 (Hs01923188_u1), MUC1 (Hs00159357_m1), C1S (Hs00156159_m1), CENPA (Hs00156455_m1), CENPM (Hs00894703_g1), CDK1 (Hs00938777_m1),

bleomycin hydrolase (BLMH, Hs00166071_m1), ACER1 (Hs00370322_m1), and HSD11B1 (Hs01547870_m1) genes. To normalize the mRNA expression levels, the relative expression of glyceraldehyde-3-phosphate dehydrogenase (GAPDH, 4333764F) was determined. Quantification of relative expression levels was performed using equations from a mathematical model developed by Pfaffl (26).

5. Protein quantification

Protein expression of the selected genes was confirmed by western blot and ELISA using three independent NHK batches. NHKs (third passages, 2×10^5 cells/well) were seeded into 6-well plates and cultured to 90% confluence. NHKs in 6-well culture plates were treated with leptin, and then lysed with cell lysis buffer (RIPA buffer) containing protease inhibitors (Sigma-Aldrich). The lysate was then subjected to centrifugation at 15,000 $\times$ *g* for 10 min, and the supernatant was used for the protein expression analysis. Proteins (40 μg/well) were fractionated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose membranes. The membranes were blocked with 5% skimmed milk in TBST (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 0.15 % Tween-20) for 1 hour at room temperature and were subsequently probed overnight at 4 °C with anti-phosphoSTAT3 antibody (#9145, Cell Signaling Technology Inc., CST), anti-STAT3 antibody (#9132, CST), anti-phospho AKT antibody (CST), anti-ATK antibody (CST), anti-phospho ERK antibody (#4370, CST), anti-ERK antibody (CST), anti-S100A7 antibody (Abcam, Cambridge, MA, USA), anti-fibronectin

(#SC-9068, Santa Cruz Biotechnology), anti-p53 antibody (#2527, CST), anti-survivin (#2808, CST), anti-Aurora A kinase (#14475, CST), anti- centromere protein A (CENPA) antibody (#2186, CST), anti-CENPM antibody (Abcam), or anti-□-actin (#A5441, Sigma-Aldrich). The blots were washed thrice with TBST and incubated with horseradish peroxidase-conjugated goat anti-rabbit IgG (Bio-Rad, Richmond, CA, USA) at room temperature for 1 hour. Detection was performed using the ECL system (Invitrogen). The concentrations of MMP1 and CXCL8/IL-8 in cell culture supernatants were measured using enzyme immunoassay kits (R&D Systems, Minneapolis, MN, USA) per the manufacturer's instructions.

6. Statistical analysis

Statistical analyses were conducted using SPSS for Windows (SPSS Science). The data are expressed as means ± standard deviations (SD). A Student's t-test was used for comparisons with controls, and one-way analysis of variance (ANOVA) followed by Bonferroni's post-test was used for multiple comparisons. P-values < 0.05 indicate statistical significance.

Ⅲ. Results

1. Microarray expression data and DEG identification on leptin-treated NHKs

First the concentration of leptin for NHKs treatment was determined, which activates intracellular STAT2 and ERK signaling but no AKT phosphorylation, as 10 ng/mL. The microarray data are accessible at National Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO) as raw data by dataset ID GSE90979. 151 upregulated and 53 downregulated DEGs were identified as a result of comparing expression data between leptin-treated and control NHKs (Table 4 and 5).

Table 4 Upregulated 151 genes and their fold change by treatment of leptin.

Probe Set ID (Gene Symbol)	Mean fold change *
213693_s_at (MUC1)	5.29
205916_at (S100A7)	5.14
208747_s_at (C1S)	5.06
204475_at (MMP1)	4.85
242587_at (SLC9A9)	4.83
211478_s_at (DPP4)	4.79
244297_at (ANKRD18A)	4.74
202859_x_at (CXCL8)	4.66
222929_at (FAM160B2)	4.65
217044_s_at (PLEKHG3)	4.58
1552794_a_at (ZNF547)	4.46
212268_at (SERPINB1)	4.35
1552761_at (SLC16A11)	4.25
224920_x_at (MYADM)	4.05
223739_at (PADI1)	3.89
227173_s_at (BACH2)	3.84
205513_at (TCN1)	3.8
201860_s_at (PLAT)	3.73
223204_at (FAM198B)	3.62
205863_at (S100A12)	3.61
205146_x_at (APBA3)	3.6
221328_at (CLDN17)	3.52
237622_at (ACO1)	3.49
211719_x_at (FN1)	3.35
229195_at (MESP1)	3.31
203750_s_at (RARA)	3.31
229718_at (N4BP2L1)	3.31
207592_s_at (HCN2)	3.27
211253_x_at (PYY)	3.26
228876_at (BAIAP2L2)	3.18
1555722_at (SCAMPER)	3.11
211924_s_at (PLAUR)	3.09
204470_at (CXCL1)	3.08
214826_at (PDE12)	3.05
211844_s_at (NRP2)	3.04
201645_at (TNC)	3.02
222075_s_at (OAZ3)	3.02
236499_at (C1orf86)	3
1563646_a_at (TMEM67)	2.99
210827_s_at (ELF3)	2.93

Probe Set ID (Gene Symbol)	Mean fold change *
238505_at (ADPRH)	2.93
1569003_at (VMP1)	2.92
220236_at (PDPR)	2.92
209573_s_at (LDLRAD4)	2.85
207499_x_at (UNC45A)	2.85
1570327_at (C20orf62)	2.84
231046_at (LINC01023)	2.81
211493_x_at (DTNA)	2.78
236088_at (NTNG1)	2.72
237463_at (ZFPM1)	2.7
1553997_a_at (ASPHD1)	2.66
218550_s_at (LRRC20)	2.65
1555529_at (RNH1)	2.64
220664_at (SPRR2C)	2.64
1553454_at (RPTN)	2.61
220351_at (ACKR4)	2.58
1559585_at (DDX60L)	2.58
204437_s_at (FOLR1)	2.57
204803_s_at (RRAD)	2.55
1557123_a_at (CHADL)	2.51
230643_at (WNT9A)	2.5
1568615_a_at (SRD5A3-AS1)	2.48
204464_s_at (EDNRA)	2.46
242579_at (BMPR1B)	2.45
1556505_at (LINC00605)	2.44
209802_at (PHLDA2)	2.42
242809_at (IL1RL1)	2.42
232306_at (CDH26)	2.41
211916_s_at (MYO1A)	2.41
242041_at (CSPP1)	2.4
239139_at (CPNE9)	2.4
205973_at (FEZ1)	2.39
238579_at (C9orf85)	2.39
203003_at (MEF2D)	2.39
239740_at (ETV6)	2.38
1569464_at (PPFIBP1)	2.37
206191_at (ENTPD3)	2.37
236262_at (MMRN2)	2.37
1556202_at (SRGAP2)	2.35
231393_x_at (ZBTB43)	2.34

Table 4 (continued)

Probe Set ID (Gene Symbol)	Mean fold change *
1554690_a_at (TACC1)	2.33
222784_at (SMOC1)	2.32
228871_at (ALG14)	2.31
201650_at (KRT19)	2.31
219926_at (POPDC3)	2.3
222886_at (NSUN3)	2.3
218824_at (PNMAL1)	2.3
234303_s_at (GPR85)	2.29
227178_at (CELF2)	2.28
238551_at (FUT11)	2.28
211261_at (NUP214)	2.28
202435_s_at (CYP1B1)	2.27
1554860_at (PTPN7)	2.27
241709_s_at (DOCK1)	2.27
204570_at (COX7A1)	2.26
221466_at (P2RY4)	2.25
230617_at (OVOL1-AS1)	2.24
226071_at (ADAMTSL4)	2.22
201081_s_at (PIP4K2B)	2.21
242903_at (IFNGR1)	2.21
223604_at (GARNL3)	2.2
229450_at (IFIT3)	2.19
209133_s_at (COMMD4)	2.19
227919_at (UCA1)	2.19
242083_at (ZNF81)	2.18
224917_at (MIR21)	2.18
231541_s_at (SPAG5-AS1)	2.18
214814_at (YTHDC1)	2.17
231095_at (LOC101928045)	2.17
213081_at (ZBTB22)	2.17
1557197_a_at (LGALS3)	2.16
229693_at (TMEM220)	2.16
232110_at (GALNT5)	2.16
214521_at (HES2)	2.16
239203_at (LSMEM1)	2.15
242625_at (RSAD2)	2.15
212190_at (SERPINE2)	2.14
242442_x_at (TRMT10A)	2.14
202237_at (NNMT)	2.14
202328_s_at (LOC101930075)	2.14

Probe Set ID (Gene Symbol)	Mean fold change *
244353_s_at (SLC2A12)	2.14
210650_s_at (PCLO)	2.13
216540_at (YME1L1)	2.12
235173_at (MBNL1-AS1)	2.12
1557116_at (APOL6)	2.11
236954_at (BOLL)	2.11
1552953_a_at (NEUROD2)	2.11
232913_at (TMED8)	2.1
204257_at (FADS3)	2.1
1560070_at (LOC101928191)	2.09
203757_s_at (CEACAM6)	2.09
202790_at (CLDN7)	2.08
232427_at (ZNF224)	2.08
219113_x_at (HSD17B14)	2.08
1552375_at (ZNF333)	2.08
230409_at (MAGI3)	2.08
236840_at (C12orf56)	2.07
228353_x_at (UBASH3B)	2.06
205039_s_at (IKZF1)	2.05
220336_s_at (GP6)	2.05
231291_at (GIPR)	2.05
238689_at (GPR110)	2.05
210511_s_at (INHBA)	2.04
205676_at (CYP27B1)	2.03
206521_s_at (GTF2A1)	2.02
211964_at (COL4A2)	2.02
1553166_at (CDH24)	2.01
210834_s_at (PTGER3)	2.01
221541_at (CRISPLD2)	2.01
234351_x_at (TRPS1)	2.01
201739_at (SGK1)	2

Table 5 Downregulated 53 genes and their fold change by treatment of leptin.

Probe Set ID (Gene Symbol)	Mean fold change *
218741_at (CENPM)	0.07
203213_at (CDK1)	0.12
204962_s_at (CENPA)	0.12
202179_at (BLMH)	0.15
206149_at (CHP2)	0.17
238752_at (GPLD1)	0.17
1561530_at (LOC101927164)	0.19
226438_at (SNTB1)	0.19
213652_at (PCSK5)	0.2
205438_at (PTPN21)	0.2
222848_at (CENPK)	0.21
205404_at (HSD11B1)	0.22
209714_s_at (CDKN3)	0.25
239547_at (HS3ST6)	0.26
206100_at (CPM)	0.26
1552544_at (SERPINA12)	0.28
1553697_at (CCSAP)	0.28
201890_at (RRM2)	0.28
1553929_at (ACER1)	0.29
240420_at (AADACL2)	0.29
230339_at (CCDC138)	0.29
232365_at (SIAH1)	0.29
206177_s_at (ARG1)	0.3
237120_at (KRT77)	0.31
205568_at (AQP9)	0.31
1558822_at (---)	0.31
231930_at (ELMOD1)	0.31
244692_at (CYP4F22)	0.32
207192_at (DNASE1L2)	0.32
1559534_at (LOC100996419)	0.32
206643_at (HAL)	0.33
234039_at (TANC1)	0.33
220414_at (CALML5)	0.34
211603_s_at (ETV4)	0.35
205554_s_at (DNASE1L3)	0.37
201295_s_at (WSB1)	0.38
1564333_a_at (PSAPL1)	0.38
1555497_a_at (CYP4B1)	0.39
220089_at (L2HGDH)	0.41
210505_at (ADH7)	0.41

Probe Set ID (Gene Symbol)	Mean fold change *
242951_at (---)	0.41
220090_at (CRNN)	0.42
243539_at (KIAA1841)	0.42
223700_at (MND1)	0.42
1552532_a_at (ATP6V1C2)	0.43
1558959_at (FAM126B)	0.43
239657_x_at (FOXO6)	0.43
1558719_s_at (RPAIN)	0.44
1554906_a_at (MPHOSPH6)	0.44
239311_at (---)	0.45
230113_at (MBNL3)	0.45
221256_s_at (HDHD3)	0.45
238342_at (---)	0.48

2. GO enrichment analysis of leptin-induced DEGs

Gene Ontology enrichment analysis was performed to grasp the biological meaning of the expressional pattern by treating leptin (Table 6). Among the GO biological process terms based on Gene Ontology Consortium, 'extracellular matrix organization (GO:0030198)' appeared to be the most significantly enriched in upregulated DEGs. The term GO:0030198 is annotated to 332 genes out of total 23,624 genes in Affymetrix Human Microarray 133 2.0 platform, which yields the expectation of 2.12 genes out of 151 upregulated DEGs. Since 9 annotated genes are included in the up-DEG group, the term GO0030198 can be suspected to significantly enriched in upregulated DEGs; the actual *p*-value calculated by Chi-squared test is 0.0008. Using MYOnto software, the list and *p*-values were extracted for all GO biological process terms annotated to at least one DEG. The biological function activated by leptin treatment is mainly about inflammatory response and wound healing process, which corresponds to the formal reports about functional study of leptin (13, 27). In particular, there are studies that elucidated the function of leptin, such as angiogenesis, ECM organization, and response to wound (15, 16).

Table 6 Enriched biological function at leptin-induced genes.

Using MYOnto software, enriched GO biological function terms were extracted from upregulated DEGs based on Gene Ontology Consortium. Here top 10 enriched terms from the list of GO enrichment analysis are shown. All *p*-values are calculated by Chi-squared test. The abbreviations stand for following: NGT for the number of genes annotated to specific GO term in the whole gene set of microarray platform analyzed (total 23,624 genes); NGD for the number of genes annotated to specific GO term in the DEG group (total 151 genes).

GO Term ID	GO biological process terms in the up-regulated DEGs	NGT	NGD	*p*-value	Upregulated DEGs
GO:0030198	extracellular matrix organization	332	9	0.0008	ADAMTSL4, COL4A2, TNC, CRISPLD2, ELF3, FN1, LGALS3, MMP1, SMOC1
GO:0071300	cellular response to retinoic acid	59	4	0.0019	MUC1, RARA, TNC, WNT9A
GO:0034329	cell junction assembly	81	4	0.0060	CDH24, CLDN17, MIR21, VMP1
GO:0007204	positive regulation of cytosolic calcium ion concentration	135	5	0.0073	EDNRA, GIPR, P2RY4, PKD1, PTGER3
GO:0030593	neutrophil chemotaxis	46	3	0.0080	CXCL8, LGALS3, S100A12
GO:0001525	angiogenesis	261	7	0.0091	COL4A2, CXCL8, CYP1B1, FN1, MMRN2, NRP2, S100A7
GO:0050729	positive regulation of inflammatory response	57	3	0.0144	EDNRA, IL1RL1, S100A12
GO:0030334	regulation of cell migration	65	3	0.0204	PHLDA2, SERPINE2, SGK1
GO:0009611	response to wounding	67	3	0.0221	FN1, SERPINE2, TNC
GO:0006805	xenobiotic metabolic process	183	4	0.0243	CYP1B1, CYP27B1, NNMT, S100A12

3. Experimental quantification of major upregulated DEGs

In order to confirm the upregulation of genes related to biological function of inflammation and/or wound healing, quantitative real-time PCR on several major upregulated DEGs was performed (Fig. 6). Six upregulated genes, MMP1, FN1, CXCL8, S100A7, MUC1, and C1S were tested. All those genes showed upregulation at leptin-treated NHKs compared to control cells although MUC1 and C1S did not show enough significance. Protein levels were also examined by experimental methods such as enzyme-linked immunosorbent assay (ELISA) and Western blot (Fig. 7). Increase of MMP1 enzyme and CXCL8 chemokine was validated by concentration-dependent manner through ELISA (Fig. 7A and 7B). Western blot showed S100A7 and FN1 protein were also increased by leptin treatment in NHKs (Fig. 7C)

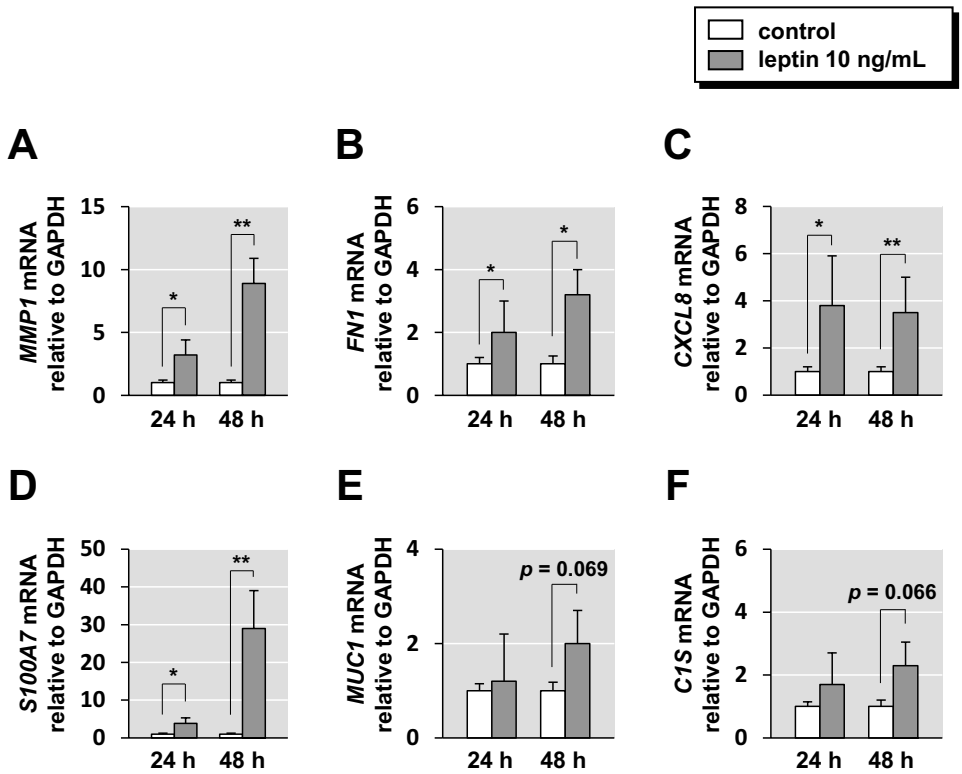


Figure 6. mRNA quantification of major upregulated DEGs.

The upregulation of (A) MMP1, (B) FN1, (C) CXCL8, (D) S100A7, (E) MUC1, and (F) C1S was confirmed through quantitative real-time PCR. Total RNA was extracted from NHKs treated with leptin for 24 and 48 h. Values represent the mean expression level ± SD of the mRNA of various genes relative to that of human GAPDH. Values represent the mean expression level and standard deviation of 3 independent measurements. The graph was reconstituted from the original paper published at the journal *Archives of Dermatological Research* (45). *$p \leq 0.05$ and **$p \leq 0.01$.

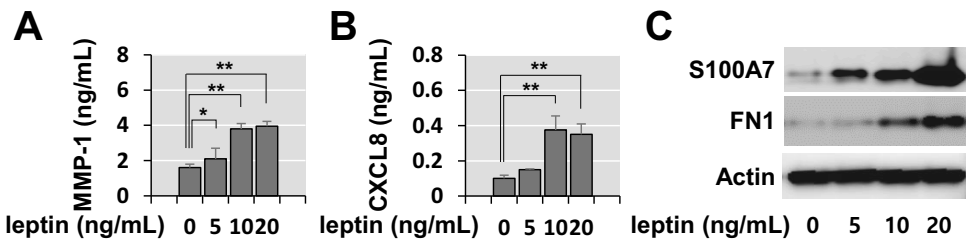


Figure 7. Protein quantification of major upregulated DEGs.

Protein level of MMP1 (A) and CXCL8 (B) was determined by enzyme-linked immunosorbent assay (ELISA). Increase of S100A7 and fibronectin (FN1) protein expression was validated by Western blot (C). The graph was reconstituted from the original paper published at the journal *Archives of Dermatological Research* (45). *$p \leq 0.05$ and **$p \leq 0.01$.

4. Identification of signaling pathways activated by leptin

There are several intracellular signaling pathways reported to mediate the function of leptin (28). Inhibitors of major pathways activated by leptin were used and observed whether each antagonist blocked the pro-inflammatory response by leptin. If JAK2-STAT3 pathway inhibitor AG490 was co-treated with leptin, the upregulation effect of MMP1 and CXCL8 by leptin disappeared (Fig. 8A and 8B). However, ERK pathway inhibitor PD98059 did not show significant effect to those genes in leptin-treated NHKs (Fig. 8C and 8D). This indicated pro-inflammatory effect and wound healing process activated by leptin is mainly mediated by JAK2-STAT3 pathway. Nonetheless, ERK pathway appeared to be also related to induction of several genes such as S100A7 and FN1 in Western blot and its quantification, as well as JAK2-STAT3 pathway did (Fig. 9).

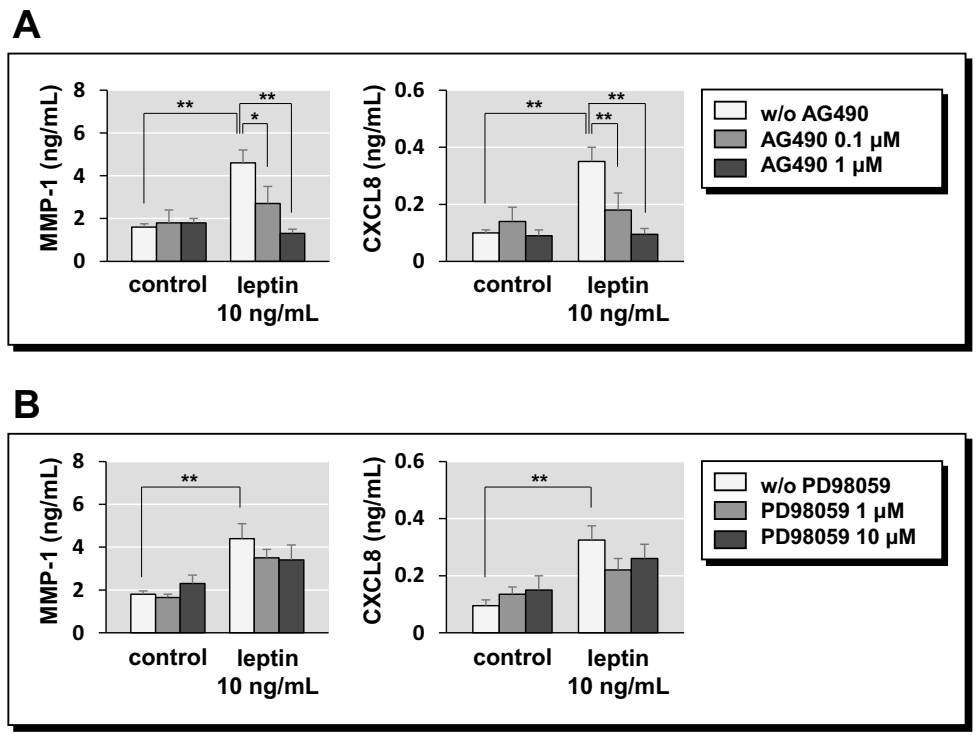


Figure 8. Effects of leptin-related pathway inhibitors to MMP1 and CXCL8.

The mRNA expression of MMP1 and CXCL8 was investigated treating leptin and/or leptin-related pathway inhibitors through q-RT-PCR. AG490 is a well-known JAK2 inhibitor, and PD98059 is a representative ERK pathway inhibitor. Values represent the mean expression level and standard deviation of 3 independent measurements. The graph was reconstituted from the original paper published at the journal *Archives of Dermatological Research* (45). *$p \leq 0.05$ and **$p \leq 0.01$.

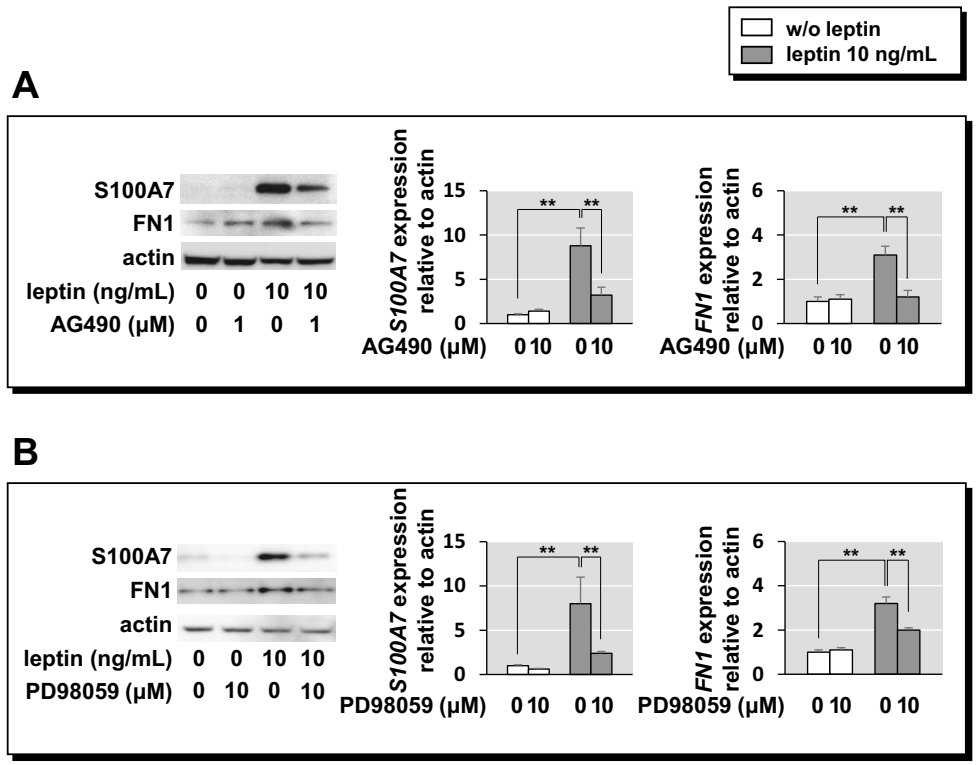


Figure 9. Effects of leptin-related pathway inhibitors to S100A7 and FN1.

Western blot was performed to investigate pathways related to leptin-induced increase of protein expression of S100A7 and FN1. A JAK2 inhibitor AG490 (A) and an ERK inhibitor PD98059 (B) are used to co-treatment with leptin. Protein expression levels were quantified by ImageJ software. Values represent the mean expression level and standard deviation of 3 independent measurements. The graph was reconstituted from the original paper published at the journal *Archives of Dermatological Research* (45). *$p \leq 0.05$ and **$p \leq 0.01$.

Ⅳ. Discussion

Leptin resistance is a well-known concept to explain the fact that leptin tends to be higher in people with obesity although it has function of appetite-decreasing (25, 29, 30). In order to grasp the leptin's biochemical effects on skin, whole-genome scale expression profile was investigated. Among the upregulated genes by treating leptin, it was found that pro-inflammatory response and wound healing was mainly activated. Representatively, ECM organization (GO:0030198) and its related genes MMP1 and fibronectin 1 were upregulated by leptin. It corresponds to former studies that investigated the relationship between leptin and wound healing in mice (15, 21) In addition, this can explain people with suspected leptin resistance are susceptible to immunological diseases of skin (22, 31).

Wound healing process of skin includes hemostasis, inflammation, ECM remodeling, resolution of inflammation, and tissue regression (32-35). Multiple reports maintain that peripheral immune response is regulated by leptin in various immune cells (13, 27). MMP1 and CXCL8 are the major DEGs in this study reported to mediate pro-inflammatory response (36-39). It was confirmed that MMP1 and CXCL8 are upregulated by leptin in NHKs, and was elucidated this is mediated mainly by JAK2-STAT3 pathway. ECM organization also plays an important role when physically impaired tissue is remodeled (34, 40). In the microarray data, ECM organization (GO:0030198) was the most significantly enriched GO biological process term. Among 332 genes annotated by Gene Ontology Consortium to be related to ECM organization, 9 genes including MMP1

and CN1 were found as up-regulated genes by treating leptin. It can be inferred that ratio is very significant considering that the number of total DEGs selected is 151 out of 23,624. In addition, it was found that genes related to angiogenesis (GO:0001525) such as COL4A2, CXCL8 are activated by leptin treatment. Since angiogenesis are important in the process of wound healing, this indicates leptin regulates angiogenesis to promote cutaneous remodeling (32). S100A7, one of the major upregulated DEGs in the expression profile, has an anti-microbial activity and is secreted from cutaneous wound (41). It indicates that S100A7 has the protective effects on injured tissue. Research that showed mice lack of leptin (ob/ob) or leptin receptor (db/db) are more susceptible to lipopolysaccharide (LPS) than normal control suggests the protective effect of leptin (42). S100A7 is also known to have function about tight junction, supported by a report that tight junction proteins such as claudins and occludins increased when S100A7 was treated to NHKs (43). GO enrichment analysis listed cell junction assembly (GO:0034329) as one of significantly enriched GO terms in upregulated DEGs. Its related DEGs, such as CDH24 and CLDN17, are well-explained by the upregulation of S100A7. In addition, S100A7 is also known as psoriasin, named after its overexpression in proliferative skin diseases like psoriasis (44). Since the prevalence of psoriasis are significantly higher in obese people than lean control, the expression level of S100A7 and its relationship to obesity should be further researched (22).

This research about the function of leptin in normal human keratinocytes suggests that leptin regulates biological processes related to wound healing, and is related to several immunological skin diseases like psoriasis. This biological insight was

drawn by microarray profile analysis using the software, without depending on pre-established hypothesis. It proved the possibility to derive novel biological knowledge originally by bio-big data.

Ⅴ. Reference

1. Kitsios GD, Zintzaras E (2009) Genome-wide association studies: hypothesis-"free" or "engaged"?. *Transl Res* 154(4):161-164.

2. De R, Bush WS, Moore JH (2014) Bioinformatics challenges in genome-wide association studies (GWAS). *Methods Mol Biol* 1168:63-81.

3. Cheong KA et al. (2014) Retinoic acid and hydroquinone induce inverse expression patterns on cornified envelope-associated proteins: implication in skin irritation. *J Dermatol Sci* 76(2):112-119.

4. Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30(1):207-10.

5. Noh M et al. (2010) MAP17 is associated with the T-helper cell cytokine-induced down-regulation of filaggrin transcription in human keratinocytes. *Exp Dermatol* 19(4):355-362.

6. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*, 28(1):27-30.

7. Nzeusseu TA et al. (2007) Identification of distinct gene expression profiles in the synovium of patients with systemic lupus erythematosus. *Arthritis Rheum* 56(5):1579-1588.

8. Subramanian A et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102(43):15545-15550.

9. Zheng Q, Wang XJ (2008) GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. *Nucleic Acids Res* 36:W358-63.

10. Ashburner M et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25(1):25-29.

11. Poeggeler B et al. (2010) Leptin and the skin: a new frontier. *Exp Dermatol* 19(1):12-18.

12. Peelman F et al. (2006) Techniques: new pharmacological perspectives for the leptin receptor. *Trends Pharmacol Sci* 27(4):218-225.

13. Lord GM et al. (1998) Leptin modulates the T-cell immune response and reverses starvation-induced immunosuppression. *Nature* 394(6696):897-901.

14. Takeda S et al. (2002) Leptin regulates bone formation via the sympathetic nervous system. *Cell* 111(3):305-317.

15. Goren I et al. (2003) Leptin and wound inflammation in diabetic ob/ob mice: differential regulation of neutrophil and macrophage influx and a potential role for the scab as a sink for inflammatory cells and mediators. *Diabetes* 52(11):2821-2832.

16. Murad A et al. (2003) Leptin is an autocrine/paracrine regulator of wound healing. *FASEB J* 17(13):1895-1897.

17. Iguchi M et al. (2001) Human follicular papilla cells carry out nonadipose tissue production of leptin. *J Invest Dermatol* 117(6):1349-1356.

18. Glasow A et al. (2001) Expression of leptin (Ob) and leptin receptor (Ob-R) in human fibroblasts: regulation of leptin secretion by insulin. *J Clin Endocrinol Metab* 86(9):4472-4479.

19. Cerman AA et al. (2008) Serum leptin levels, skin leptin and leptin receptor expression in psoriasis. *Br J Dermatol* 159(4):820-826.

20. Johnston A et al. (2008) Obesity in psoriasis: leptin and resistin as mediators of cutaneous inflammation. *Br J Dermatol* 159(2):342-350.

21. Frank S et al. (2000) Leptin enhances wound re-epithelialization and constitutes a direct function of leptin in skin repair. *J Clin Invest* 106(4):501-509.

22. Yosipovitch G, DeVore A, Dawn A (2007) Obesity and the skin: skin physiology and skin manifestations of obesity. *J Am Acad Dermatol* 56(6):901-916

23. Friedman JM, Halaas JL (1998) Leptin and the regulation of body weight in mammals. *Nature* 395(6704):763-770.

24. Monti V et al. (2006) Adams, Relationship of ghrelin and leptin hormones with body mass index and waist circumference in a random sample of adults. *J Am Diet Assoc* 106(6):822-828

25. Rosenbaum M, Leibel RL (2014) 20 years of leptin: role of leptin in energy homeostasis in humans. *J Endocrinol* 223(1):T83-96.

26. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29(9):e45.

27. Loffreda S et al. (1998) Leptin regulates proinflammatory immune responses. *FASEB J* 12(1):57-65.

28. Otero M et al. (2006) Towards a pro-inflammatory and immunomodulatory emerging role of leptin. *Rheumatology (Oxford)* 45(8):944-950.

29. Crujeiras AB et al. (2015) Leptin resistance in obesity: An epigenetic landscape. *Life Sci* 140:57-63.

30. Sainz N et al. (2015) Leptin resistance and diet-induced obesity: central and peripheral actions of leptin. *Metabolism* 64(1):35-46.

31. Shipman AR, Millington GW, Obesity and the skin. *Br J Dermatol* 165(4):743-750.

32. Tonnesen MG, Feng X, Clark RA (2000), Angiogenesis in wound healing. *J Investig Dermatol Symp Proc* 5(1):40-46.

33. Eming SA, Krieg T, Davidson JM (2007) Inflammation in wound repair: molecular and cellular mechanisms. *J Invest Dermatol* 127(3):514-525.

34. Gurtner GC et al. (2008) Wound repair and regeneration. *Nature* 453(7193):314-321.

35. Guo S, Dipietro LA (2010) Factors affecting wound healing. *J Dent Res* 89(3):219-229.

36. Barker JN et al. (1991) Nickoloff, Keratinocytes as initiators of inflammation. *Lancet* 337(8735):211-214.

37. Biasi D et al. (1998) Neutrophil functions and IL-8 in psoriatic arthritis and in cutaneous psoriasis. *Inflammation* 22(5):533-543.

38. Pasparakis M, Haase I, Nestle FO (2014) Mechanisms regulating skin immunity and inflammation. *Nat Rev Immunol* 14(5):289-301.

39. Bae ON et al. (2015) Keratinocytic vascular endothelial growth factor as a novel biomarker for pathological skin condition. *Biomol Ther* 23(1):12-18.

40. Bonnans C, Chou J, Werb Z (2014) Remodelling the extracellular matrix in development and disease. *Nat Rev Mol Cell Biol* 15(12):786-801.

41. Lee KC, Eckert RL (2007) S100A7 (Psoriasin)--mechanism of antibacterial action in wounds. *J Invest Dermatol* 127(4):945-957.

42. Faggioni R et al. (1999) Leptin deficiency enhances sensitivity to endotoxin-induced lethality. *Am J Physiol* 276(1 Pt 2):R136-142.

43. Hattori F et al. (2014) The antimicrobial protein S100A7/psoriasin enhances the expression of keratinocyte differentiation markers and strengthens the skin's tight junction barrier. *Br J Dermatol* 171(4):742-753.

44. Anderson KS et al. (2009) Detection of psoriasin/S100A7 in the sera of patients with psoriasis. *Br J Dermatol* 160(2):325-332.

45. Lee M et al. (2018) Leptin regulates the pro-inflammatory response in human epidermal keratinocytes. *Arch Dermatol Res* 310(4):351-362.

요약 (국문초록)

유전자 온톨로지 기반 전사체 분석법 및 각질형성세포에 대한 렙틴의 전사체 프로파일 연구

이 문 영

서울대학교 약학대학원

약학과 천연물과학전공

유전체 기술의 진보에 따라 다양한 생리 현상을 규명하기 위한 생물정보학적 방법론이 시도되고 있다. 이러한 방법론에 따라 생성된 오믹스 데이터베이스는 다른 연구자가 이용할 수 있도록 공개되어 바이오빅데이터(Bio-big data)를 구축하고 있으며, 이를 통해 기존에 알려지지 않은 생물학적 지식을 가설에 의존하지 않고 도출할 수 있는 가능성이 주목받고 있다. 그러나 이러한 데이터 자체만으로는 인간에게 유용한 정보를 얻을 수 없으며, 오믹스 데이터의 특성에 맞는 통계적 분석법이 요구된다. 본 연구에서는 전체 유전자 수준에서 시료의 mRNA 발현 양상을 확인할 수 있는 마이크로어레이 또는 RNA-seq 데이터로부터 발현값이 유의하게 변화한 유전자들의 기능적 경향성을

파악할 수 있는 시스템을 구축하고자 하였다. 2 만여 개에 달하는 인간 유전자 각각의 기능 및 연관되어 있는 생체 경로에 대한 정보를 범주화하기 위해 유전자온톨로지컨소시움(Gene Ontology Consortium), 교토유전자및유전체백과사전(KEGG) 등의 유전자 기능 정보 체계를 활용하였다. 특정 유전자 기능정보가 DEGs(Differentially Expressed Genes)에 유의미한 수준으로 공통적으로 포함되어 있는지를 파악하기 위해 카이제곱검정 및 피셔 정확검정을 통한 p 값을 산출하는 방법을 제시한 후, 이를 모든 유전자 기능정보에 적용하기 위하여 데이터 분석 소프트웨어를 개발하였다. 기존에 개발된 유전자 온톨로지 기반 전사체 분석 소프트웨어와 비교할 때, 본 연구에서 개발한 소프트웨어는 다음의 장점을 지닌다. (a) 현재까지의 상용 프로그램은 연구자가 소속되어 있거나 널리 쓰이는 유전자 기능 정보 체계 중 한 가지만을 참조하여 분석하지만, 본 연구에서 개발한 소프트웨어는 복수의 유전자 기능 정보 체계를 통한 분석을 한 번에 진행할 수 있다. (b) 핵심 유전자 선정부터 온톨로지 분석까지의 과정을 한 번에 진행할 수 있어 일관된 분석 결과를 도출하고 분석에 걸리는 시간이 단축되며, 생물정보학 연구 경험이 없는 초심자가 편리하게 분석할 수 있다.

본 연구에서 개발한 소프트웨어는 2 개로 나뉘어 있으며 두 프로그램은 연이어 실행할 수 있도록 설계되어 있다. 각각의 명칭은 ‘MYDEGextractor’와 ‘MYOnto’로, 전자는 전사체 분석의 핵심 유전자인 DEGs 를 구하는 기능을 포함하며 후자는 온톨로지 정보를 통해 DEGs 의 기능적 경향성을 분석한다. 본 연구에서 개발한 소프트웨어를 활용한 생물학적 발견 사례로서 각질세포에서 렙틴이 염증 반응에 미치는

영향을 규명하였다. 렙틴 처리 각질세포에 대한 마이크로어레이 데이터로부터 151 개의 발현 증가 유전자와 53 개의 발현 감소 유전자를 추출하였으며, 온톨로지 기반 기능 정보 분석을 통해 발현이 증가한 유전자 중 염증 개시 조절 유전자가 유의한 수준으로 포함되어 있음을 확인하였다. 이에 해당하는 유전자인 CXCL8, MMP1 등의 발현 증가는 STAT3 길항제에 의하여 억제된다는 것을 확인하여, 각질세포에서 렙틴이 STAT3 대사경로를 통해 상처 치유 과정을 활성화시킨다는 것을 증명하였다.

주요어 : 바이오빅데이터, 기능유전체, 유전자 온톨로지 기반 전사체 분석, 각질형성세포, 렙틴

학번 : 2017-23557

Archives of Dermatological Research
https://doi.org/10.1007/s00403-018-1821-0

ORIGINAL PAPER

Leptin regulates the pro-inflammatory response in human epidermal keratinocytes

Moonyoung Lee[1,2] · Eunyoung Lee[1,2] · Sun Hee Jin[1,2] · Sungjin Ahn[1,2] · Sae On Kim[1,2] · Jungmin Kim[1,2] · Dalwoong Choi[3] · Kyung-Min Lim[4] · Seung-Taek Lee[5] · Minsoo Noh[1,2]

Received: 26 September 2017 / Revised: 22 January 2018 / Accepted: 13 February 2018

Abstract
The role of leptin in cutaneous wound healing process has been suggested in genetically obese mouse studies. However, the molecular and cellular effects of leptin on human epidermal keratinocytes are still unclear. In this study, the whole-genome-scale microarray analysis was performed to elucidate the effect of leptin on epidermal keratinocyte functions. In the leptin-treated normal human keratinocytes (NHKs), we identified the 151 upregulated and 53 downregulated differentially expressed genes (DEGs). The gene ontology (GO) enrichment analysis with the leptin-induced DEGs suggests that leptin regulates NHKs to promote pro-inflammatory responses, extracellular matrix organization, and angiogenesis. Among the DEGs, the protein expression of IL-8, MMP-1, fibronectin, and S100A7, which play roles in which is important in the regulation of cutaneous inflammation, was confirmed in the leptin-treated NHKs. The upregulation of the leptin-induced proteins is mainly regulated by the STAT3 signaling pathway in NHKs. Among the downregulated DEGs, the protein expression of nucleosome assembly-associated centromere protein A (CENPA) and CENPM was confirmed in the leptin-treated NHKs. However, the expression of CENPA and CENPM was not coupled with those of other chromosome passenger complex like Aurora A kinase, INCENP, and survivin. In cell growth kinetics analysis, leptin had no significant effect on the cell growth curves of NHKs in the normal growth factor-enriched condition. Therefore, leptin-dependent downregulation of CENPA and CENPM in NHKs may not be directly associated with mitotic regulation during inflammation.

Keywords Leptin · Epidermal keratinocytes · Pro-inflammatory response · Centromere proteins · Gene ontology enrichment analysis

Introduction

Leptin, a circulating hormone primarily secreted from adipocytes, regulates energy homeostasis by activating leptin receptors in the hypothalamus to suppress appetite. In addition to this central hypothalamic hormonal regulation, leptin has also peripheral functions that affect bone metabolism, T

Moonyoung Lee and Eunyoung Lee have contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00403-018-1821-0) contains supplementary material, which is available to authorized users.

✉ Minsoo Noh
minsoonoh@snu.ac.kr; minsoo@alum.mit.edu

[1] College of Pharmacy, Seoul National University, Seoul, Republic of Korea

[2] Natural Products Research Institute, College of Pharmacy, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea

[3] Department of Public Health Science, Graduate School and College of Public Health Science, Korea University, Seoul 02841, Republic of Korea

[4] College of Pharmacy, Ewha Womans University, Seoul 03760, Republic of Korea

[5] Department of Biochemistry, College of Life Science and Biotechnology, Yonsei University, Seoul 03722, Republic of Korea

Published online: 21 February 2018

cell responses, and wound healing in an autocrine or paracrine manner [13, 22, 26, 37]. Signal transducer and activator of transcription 3 (STAT3) plays a major role in intracellular leptin signaling [29]. Leptin receptor is a member of interleukin-6 (IL-6) receptor family and has no intrinsic kinase activity [28]. Activation of leptin receptors leads to the phosphorylation of receptor-activated Janus kinase 2 (JAK2) and tyrosine residues in the cytoplasmic domain of leptin receptors. This change triggers intracellular responses through multiple signaling pathways. JAK2 activation is coupled with the phosphorylation of STAT3 and other adaptor molecules recruited to the activated leptin receptor complex, which participates in the activation of phosphatidylinositol-3 phosphate kinase (PI3K) and extracellular signal-regulated kinase 1/2 (ERK1/2) [28]. Activation of these intracellular signaling pathways in response to leptin is dependent on various peripheral tissue contexts [28].

Leptin is mainly produced from intradermal and subcutaneous adipocytes in human skin. Interestingly, it has been suggested that non-adipose leptin regulates cutaneous biological functions. For example, hair follicle papilla cells produce leptin and functional leptin receptor [17], and insulin stimulates the expression of leptin and leptin receptor in cultured human fibroblasts [12]. In addition, the levels of leptin and leptin receptor expression in psoriatic human skin correlates with disease severity [7]. This correlation between leptin levels and disease severity may be related to leptin-induced secretion of pro-inflammatory mediators such as IL-6 and tumor necrosis factor α, which are produced by immune cells [7, 18]. One important peripheral function of leptin is participation in wound healing. The importance of leptin is apparent in the impaired wound healing of leptin-deficient *ob*/*ob* mice [27]. The effect of leptin on re-epithelialization has been explained mechanistically by an examination of the mitotic activity of leptin in a bromodeoxyuridine (BrDU) incorporation assay of epidermal keratinocytes [10]. However, obese humans have impaired skin wound healing [41], and leptin levels in obese humans generally correlate with an increase in body mass index (BMI) [11, 23, 24, 32]. Leptin resistance in obese individuals may lead to discrepant data in human epidemiological and mouse experimental studies and further studies are needed to clarify this difference. Although the significance of leptin in epidermal regulation and pathogenesis has been shown, the molecular mechanism underlying the regulation human keratinocyte functions by leptin is not fully understood.

In this study, to elucidate the molecular mechanisms underlying the effect of leptin on epidermal keratinocytes, we analyzed genome-wide transcriptional responses of normal human keratinocytes (NHKs). We found that leptin upregulated pro-inflammatory responses in NHKs, as it did in immune cells. Leptin also stimulated NHKs to produce many essential proteins in the extracellular matrix (ECM) remodeling process. Notably, leptin suppressed the expression of centromere proteins and proteins associated with metabolic processes.

Materials and methods

Primary keratinocyte culture and total RNA isolation

NHKs from neonatal foreskin were obtained from Lonza (Basel, Switzerland) and cultured as previously described [19]. Briefly, NHKs were maintained in keratinocyte basal medium (KBM) with a KGM2 complete growth factor cocktail consisting of insulin, human epidermal growth factor, bovine pituitary extract, hydrocortisone, epinephrine, transferrin, and gentamicin/amphotericin B (Lonza). Three independent NHK batches prepared from different donors were purchased, and all experiments were performed in triplicate or quadruplicate. Frozen NHKs were thawed and passaged at 80–90% confluence for cell expansion. For the analysis of cell growth kinetics, the second-passage NHKs were plated on 6-well culture plates with or without the treatment of leptin (Sigma, St Louis, MO, USA). Cells were counted at 24, 48, and 96 h after leptin treatment using a Z-Series Coulter Counter (Indianapolis, IN, USA). For the oligonucleotide microarray and validation studies, leptin was treated in 90% confluent NHKs at the second passage. Total RNA was harvested using Trizol reagent (Invitrogen, Carlsbad, CA, USA). RNA integrity was determined with a BioAnalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). For the cell growth kinetic study, the number of cells was counted using a Coulter counter (Beckman Coulter Life Sciences, Indianapolis, IN, USA).

Microarray analysis and gene ontology (GO) enrichment analysis

Affymetrix GeneChip Human Genome U133 Plus 2.0 oligonucleotide arrays were used for the genome-scale transcriptional analysis (Affymetrix, Santa Clara, CA, USA). Microarray experiments were performed and the differentially expressed genes (DEGs) were selected as described [19]. A GO enrichment analysis of leptin-induced DEGs was performed by comparing the frequency of GO biological process (BP) terms assigned to each gene in a group of DEGs to that in the entire set of genes in the Affymetrix Human 133 2.0 GeneChip array. The GO annotation files were downloaded from the Gene Ontology Consortium webpage (http://www.geneontology.org). In the analysis, the August 2016 version of the GO BP terms was used for the GO enrichment analysis. When redundant probe sets in the Affymetrix Human 133 2.0 GeneChip were considered

probes recognizing single gene units in the GO enrichment analysis, 23,624 was considered the total number of genes analyzed. A 2×2 contingency matrix was constructed to determine the frequency of a specific GO BP term in a group of DEGs compared with that in the full set of 23,624 genes. The 2×2 contingency matrix was analyzed by Fisher's exact test (frequency < 5) or χ^2 test (frequency ≥ 5) using SPSS for Windows (SPSS Science, Chicago, IL, USA) to calculate the level of significance.

Quantitative real-time reverse transcription polymerase chain reaction (Q-RT-PCR)

The TaqMan probes (Applied Biosystems, Foster City, CA, USA) used in the Q-RT-PCR analysis targeted matrix metalloproteinase 1 (MMP-1, Hs00899658_m1), fibronectin (FN, Hs01549937_m1), CXCL8/IL-8 (Hs00174103_m1), S100A7 (Hs01923188_u1), MUC1 (Hs00159357_m1), C1S (Hs00156159_m1), CENPA (Hs00156455_m1), CENPM (Hs00894703_g1), CDK1 (Hs00938777_m1), bleomycin hydrolase (BLMH, Hs00166071_m1), ACER1 (Hs00370322_m1), and HSD11B1 (Hs01547870_m1) genes. To normalize the mRNA expression levels, the relative expression of glyceraldehyde-3-phosphate dehydrogenase (GAPDH, 4333764F) was determined.

Western blot analysis and enzyme-linked immunosorbent assay (ELISA)

Protein levels (third passages, 2×10^5 cells/well) in 6-well culture plates were determined as previously described [16] with anti-phosphoSTAT3 antibody (#9145, Cell Signaling Technology Inc., CST), anti-STAT3 antibody (#9132, CST), anti-phospho-AKT antibody (CST), anti-ATK antibody (CST), anti-phospho-ERK antibody (#4370, CST), anti-ERK antibody (CST), anti-S100A7 antibody (Abcam, Cambridge, MA, USA), anti-fibronectin (#SC-9068, Santa Cruz Biotechnology), anti-p53 antibody (#2527, CST), anti-survivin (#2808, CST), anti-Aurora A kinase (#14475, CST), anti- centromere protein A (CENPA) antibody (#2186, CST), anti-CENPM antibody (Abcam), or anti-β-actin (#A5441, Sigma-Aldrich). The concentrations of MMP-1 and CXCL8/IL-8 in cell culture supernatants were measured using enzyme immunoassay kits (R&D Systems, Minneapolis, MN, USA) as per the manufacturer's instructions.

Statistical analysis

Statistical analyses were conducted using SPSS for Windows (SPSS Science). A Student's *t* test was used for comparisons with controls, and one-way analysis of variance (ANOVA) followed by Bonferroni's post-test was used for multiple comparisons. p values < 0.05 indicate statistical significance.

Results

Genome-scale transcriptional responses of NHKs to leptin

For the genome-scale microarray analysis of leptin-induced biological responses in NHKs, we first confirmed leptin-induced biological responses at a concentration that activated intracellular STAT3 and ERK signaling in parallel NHK cultures. Phosphorylation of STAT3 and ERK was significantly increased in NHKs at 30 min and 24 h after treatment with 10 ng/ml leptin, whereas AKT was not phosphorylated (Fig. 1A). The raw microarray data in the form of Affymetrix CEL files can be accessed in the National Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO) database using the accession number GSE90979. We identified 151 upregulated and 53 downregulated DEGs in leptin-treated NHKs (Fig. 1b and Supplementary Table 1), and the top 20 DEGs are listed in the Table 1.

GO BP enrichment analysis of leptin-induced DEGs

To determine functional phenotypes that were significantly overrepresented in leptin-treated NHKs, the frequency of each GO BP term in the DEGs was compared with that in the full set of genes represented in the Affymetrix human 133 2.0 oligonucleotide array (Table 2). Specifically, 2×2 contingency tables for all GO BP terms annotated in the leptin-induced DEGs were produced for comparison with those in a total of 23,624 genes. The p values were calculated from each contingency table using Fisher's exact test- or $\chi\chi^2$ test-based statistics. When the frequencies of all GO BP terms for the 151 upregulated DEGs were compared with those in the total gene set, ECM organization (GO:0030198) was identified as the most significantly overrepresented GO BP in the leptin-treated NHKs (Table 2). The contingency table for the ECM organization GO BP was constructed, and nine genes in the 151 upregulated DEGs were annotated as ECM organization genes, whereas 332 genes in a total of 23,624 genes were associated with this GO BP term. The p value calculated by $\chi\chi^2$ test using the contingency table was 0.0008. This procedure was applied to all other GO BP terms annotated in the 151 DEGs, and the top 10 GO BP terms significantly represented in the leptin-induced DEGs are listed in Table 2. Notably, genes associated with inflammatory responses were the major gene expression signature of

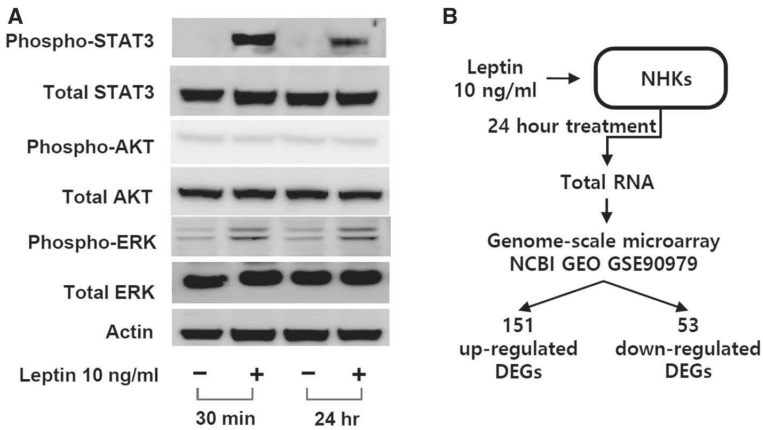


Fig. 1 Phosphorylation profile of STAT3, AKT, and ERK in leptin-treated NHKs and the scheme for whole-genome transcription profiling. NHKs were treated with leptin for 30 min and 24 h, and protein samples were prepared for western blotting of phospho-STAT3, total STAT3, phospho-AKT, total AKT, phospho-ERK, and total ERK (A). The microarray study was performed with leptin-treated NHKs using an Affymetrix Human 133 2.0 GeneChip platform. The raw data from this microarray study can be accessed in the NCBI GEO database under accession number GSE90979 (B)

the leptin-induced upregulated DEGs (Table 2). In addition, other significant inflammation-associated GO BPs in the upregulated DEGs were neutrophil chemotaxis (GO:0030593), angiogenesis (GO:0001525), positive regulation of inflammatory response (GO:0050729), regulation of cell migration (GO:0030334), and response to wounding (GO:0009611). This gene expression signature in leptin-treated NHKs is consistent with the reported functional role of leptin in various BPs such as inflammation and wound healing [21, 22]. The leptin-enriched GO BP terms identified, namely ECM organization (GO:0030198), angiogenesis, and response to wounding (GO:0009611), support the functional role of leptin in wound healing that has been reported in the literature [13, 26].

The GO BP enrichment analysis of 53 downregulated DEGs showed that nucleosome assembly processes were significantly downregulated in leptin-treated NHKs (Table 2). The downregulated DEGs *CENPA, CENPK*, and *CENPM* are associated with nucleosome assembly and mitotic cell cycle (GO:0034080, GO:0006334, and GO:0000278). In addition, leptin also downregulated DEGs annotated as genes associated with metabolic processes such as insulin receptor signaling (GO:0008286) and other metabolic processes (GO:0008152, GO:0044281).

Validation of the leptin-induced upregulated DEGs

Expression levels in microarray data are calculated based on several statistical assumptions using analytical software algorithm. Therefore, the expression of DEGs should be validated by independent measurements such as Q-RT-PCR or western blot. We performed validation experiments for the DEGs upregulated in leptin-treated NHKs. We validated the upregulation of six genes by Q-RT-PCR, namely *MMP-1*, fibronectin (*FN1*), *CXCL8/IL-8, S100A7, MUC1*, and *C1S*, because they are associated with the most highly representing GO BP enrichment processes such as inflammation, ECM organization, and response to wounding (Fig. 2). The gene transcription of *MMP-1, FN1, CXCL8/IL-8*, and *S100A7* was significantly upregulated in NHKs (Fig. 2a–d). However, the mRNA expression levels of *MUC1* and *C1S* increased to some extent with no significant difference compared to the control (Fig. 2E, F). We next confirmed leptin-induced protein expression by ELISA or Western blot analysis (Fig. 3). The protein levels of MMP-1 and CXCL8/IL-8 were significantly higher than the control, correlated to their mRNA levels (Fig. 3a, b). In the Western blot analysis, leptin significantly increased, in a concentration-dependent manner, the protein expression levels of both S100A7 and fibronectin (Fig. 3c).

Next, to identify the intracellular signaling pathways associated with changes in these validated DEGs, we evaluated the effects of pharmacological inhibitors against components of the STAT3 and ERK signaling pathways (Fig. 4). AG490, a JAK2 inhibitor that attenuates activation of the STAT3 pathway inhibited leptin-induced upregulation of MMP-1 and CXCL8/IL-8 in NHK culture supernatants (Fig. 4a, b). In contrast, there was no significant effect of PD98059, an ERK pathway inhibitor, on the leptin-dependent upregulation of MMP-1 and CXCL8/IL-8 (Fig. 4c, d). We also investigated the effects of AG490 and

Table 1 Top 20 upregulated and downregulated DEGs in the leptin-treated NHKs

Probe Set ID	Gene Title	Gene symbol	Mean fold change[a]
Upregulated genes			
213693_s_at	Mucin 1, cell surface associated	MUC1	5.29
205916_at	S100 calcium-binding protein A7	S100A7	5.14
208747_s_at	Complement component 1, s subcomponent	C1S	5.06
204475_at	Matrix metallopeptidase 1 (interstitial collagenase)	MMP1	4.85
242587_at	Solute carrier family 9, subfamily A, member 9	SLC9A9	4.83
211478_s_at	Dipeptidyl-peptidase 4	DPP4	4.79
244297_at	Ankyrin repeat domain 18A	ANKRD18A	4.74
202859_x_at	Chemokine (C-X-C motif) ligand 8	CXCL8	4.66
222929_at	Family with sequence similarity 160, member B2	FAM160B2	4.65
217044_s_at	Pleckstrin homology domain containing, family G member 3	PLEKHG3	4.58
1552794_a_at	Zinc finger protein 547	ZNF547	4.46
212268_at	Serpin peptidase inhibitor, clade B (ovalbumin), member 1	SERPINB1	4.35
1552761_at	solute carrier family 16, member 11	SLC16A11	4.25
224920_x_at	Myeloid-associated differentiation marker	MYADM	4.05
223739_at	Peptidyl arginine deiminase, type I	PADI1	3.89
227173_s_at	BTB and CNC homology 1, basic leucine zipper transcription factor 2	BACH2	3.84
205513_at	Transcobalamin I (vitamin B12 binding protein, R binder family)	TCN1	3.80
201860_s_at	Plasminogen activator, tissue	PLAT	3.73
223204_at	Family with sequence similarity 198, member B	FAM198B	3.62
205863_at	S100 calcium-binding protein A12	S100A12	3.61
Downregulated genes			
218741_at	Centromere protein M	CENPM	0.07
203213_at	Cyclin-dependent kinase 1	CDK1	0.12
204962_s_at	Centromere protein A	CENPA	0.12
202179_at	Bleomycin hydrolase	BLMH	0.15
206149_at	Calcineurin-like EF-hand protein 2	CHP2	0.17
238752_at	Glycosylphosphatidylinositol specific phospholipase D1	GPLD1	0.17
1561530_at	Uncharacterized LOC101927164	LOC101927164	0.19
226438_at	Syntrophin, beta 1	SNTB1	0.19
213652_at	proprotein convertase subtilisin/kexin type 5	PCSK5	0.20
205438_at	Protein tyrosine phosphatase, non-receptor type 21	PTPN21	0.20
222848_at	Centromere protein K	CENPK	0.21
205404_at	Hydroxysteroid (11-beta) dehydrogenase 1	HSD11B1	0.22
209714_s_at	Cyclin-dependent kinase inhibitor 3	CDKN3	0.25
239547_at	Heparan sulfate (glucosamine) 3-O-sulfotransferase 6	HS3ST6	0.26
206100_at	Carboxypeptidase M	CPM	0.26
1552544_at	Serpin peptidase inhibitor, clade A (alpha-1 antitrypsin), member 12	SERPINA12	0.28
1553697_at	Centriole, cilia, and spindle-associated protein	CCSAP	0.28
201890_at	Ribonucleotide reductase M2	RRM2	0.28
1553929_at	Alkaline ceramidase 1	ACER1	0.29
240420_at	Arylacetamide deacetylase-like 2	AADACL2	0.29

[a]Values represent the ratio of leptin-treated-to-vehicle-treated control expression values

Table 2 Top ten enriched gene ontology biological processes in the DEGs of leptin-treated NHKs

GO Term ID	Enriched GO biological process terms in DEGs	*p* value (*chi* test)	# of GO BP in the DEGs	# of GO BP in the microarray (total 23,624)	DEGs (Fold change, leptin to control)
Upregulated DEGs (Total **151**)					
GO:0030198	extracellular matrix organization	0.0008	9	332	MMP1 (4.85), FN1 (3.35), TNC (3.02), ELF3 (2.93), SMOC1 (2.32), ADAMTSL4 (2.22), LGALS3 (2.16), COL4A2 (2.02), CRISPLD2 (2.01)
GO:0071300	cellular response to retinoic acid	0.0019	4	59	MUC1 (5.29), RARA (3.31), TNC (3.02), WNT9A (2.50)
GO:0034329	cell junction assembly	0.0060	4	81	CLDN17 (3.52), VMP1 (2.92), MIR21 (2.18), CDH24 (2.01)
GO:0007204	positive regulation of cytosolic calcium ion concentration	0.0073	5	135	EDNRA (2.46), P2RY4 (2.25), PKD1 (2.14), GIPR (2.05), PTGER3 (2.01)
GO:0030593	neutrophil chemotaxis	0.0080	3	46	CXCL8 (4.66), S100A12 (3.61), LGALS3 (2.16)
GO:0001525	angiogenesis	0.0091	7	261	S100A7 (5.15), CXCL8 (4.66), FN1 (3.35), NRP2 (3.04), MMRN2 (2.37), CYP1B1 (2.28), COL4A2 (2.02)
GO:0050729	positive regulation of inflammatory response	0.0144	3	57	S100A12 (3.61), EDNRA (2.46), IL1RL1 (2.42)
GO:0030334	regulation of cell migration	0.0204	3	65	PHLDA2 (2.42), SERPINE2 (2.14), SGK1 (2.00)
GO:0009611	response to wounding	0.0221	3	67	FN1 (3.35), TNC (3.02), SERPINE2 (2.14)
GO:0006805	xenobiotic metabolic process	0.0243	4	183	S100A12 (3.61), CYP1B1 (2.28), NNMT (2.14), CYP27B1 (2.03)
Downregulated DEGs (total 53)					
GO:0034080	centromere-specific nucleosome assembly	< 0.0001	3	26	CENPM (0.07), CENPA (0.12), CENPK (0.21)
GO:0071277	cellular response to calcium ion	< 0.0001	3	32	CHP2 (0.17), GPLD1 (0.17), ACER1 (0.29)
GO:0008286	insulin receptor signaling pathway	0.0004	4	155	CDK1 (0.12), GPLD1 (0.17), ATP6V1C2 (0.43), FOXO6 (0.43)
GO:0006334	nucleosome assembly	0.0017	3	102	CENPM (0.07), CENPA (0.12), CENPK (0.21)
GO:0000278	mitotic cell cycle	0.0022	5	397	CENPM (0.07), CENPA (0.12), CDK1 (0.12), CENPK (0.21), RRM2 (0.28)
GO:0055114	oxidation–reduction process	0.0048	6	687	HSD11B1 (0.22), RRM2 (0.28), CYP4F22 (0.32), CYP4B1 (0.39), ADH7 (0.41), L2HGDH (0.41)
GO:0044281	small molecule metabolic process	0.0063	9	1480	HSD11B1 (0.22), HS3ST6 (0.26), RRM2 (0.28), ACER1 (0.29), ARG1 (0.30), HAL (0.33), CYP4B1 (0.39), ADH7 (0.41), L2HGDH (0.41)
GO:0008152	metabolic process	0.0068	6	739	HSD11B1 (0.22), CPM (0.26), HS3ST6 (0.26), AADACL2 (0.29), AQP9 (0.31), HDHD3 (0.45)
GO:0006508	proteolysis	0.0477	4	594	BLMH (0.15), PCSK5 (0.20), CPM (0.26), SIAH1 (0.29)

Table 2 (continued)

GO Term ID	Enriched GO biological process terms in DEGs	*p* value (*chi* test)	# of GO BP in the DEGs	# of GO BP in the microarray (total 23,624)	DEGs (Fold change, leptin to control)
GO:0042493	response to drug	0.0498	3	359	CDK1 (0.12), BLMH (0.15), ARG1 (0.30)

*#Denotes 'number'

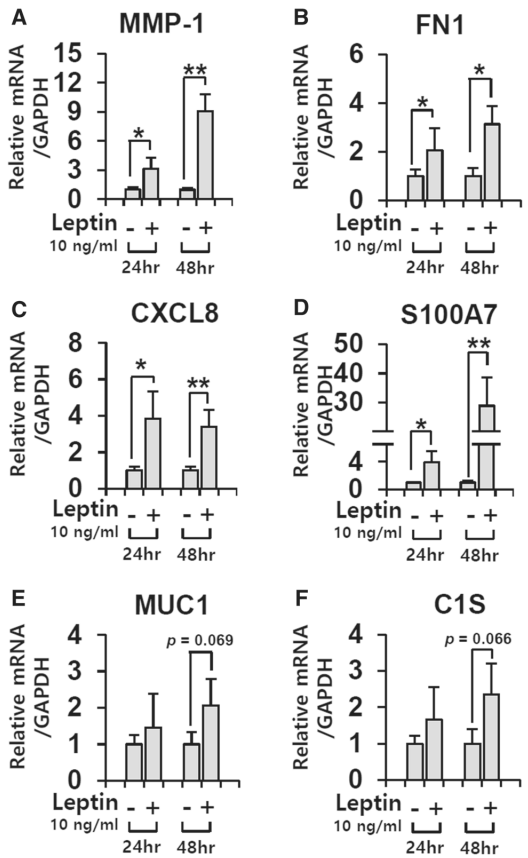


Fig. 2 Validation of upregulated differentially expressed genes by Q-RT-PCR. To determine the mRNA levels of MMP-1 (**a**), FN1 (**b**), CXCL8/IL8 (**c**), S100A7 (**d**), MUC1 (**e**), and C1S (**f**), total RNA was extracted from NHKs treated with leptin for 24 and 48 h. Values represent the mean expression level ± SD of the mRNA of various genes relative to that of human GAPDH. Error bars represent the SD of three independent measurements ($n = 3$). *$p \leq 0.05$ and **$p \leq 0.01$

PD98059 on the intracellular protein levels of S100A7 and fibronectin. Both AG490 and PD98059 significantly inhibited leptin-dependent upregulation of both S100A7 and

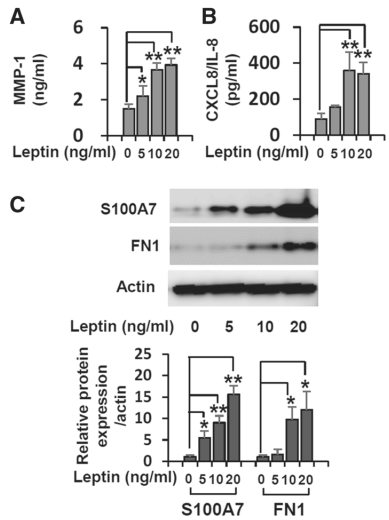


Fig. 3 Confirmation of protein expression levels of the validated upregulated differentially expressed genes by ELISA and western blot. The protein levels of MMP-1 (**a**) and CXCL8/IL8 (**b**) in the culture supernatants of leptin-treated NHKs were measured by ELISA. The levels (**c**) of S100A7 and fibronectin in leptin-treated NHKs were analyzed by western blot. The western blot results were quantified relative to β-actin levels using ImageJ software. Values represent the mean expression level ± SD (n = 3). *$p \leq 0.05$ and ** $p \leq 0.01$

fibronectin (Fig. 4E, F). In the validation experiments, the upregulation of the leptin-induced DEGs MMP-1, CXCL8/IL-8, S100A7, and fibronectin was confirmed at the protein level in leptin-treated NHKs.

Validation of the leptin-induced downregulated DEGs

Based on the GO BP enrichment analysis of downregulated genes, the leptin-treated NHKs showed transcriptional phenotypes that suggested changes in nucleosomal chromatin reassembly, cell cycling, and metabolic processes. We validated the six downregulated DEGs *CENPA, CENPM, CDK1, BLMH, ACER1*, and *HSD11B1* (Fig. 5). At both 24 and 48 h, transcription of *CENPA* and *CENPM* was significantly decreased (Fig. 5a, b). Significant changes in the mRNA

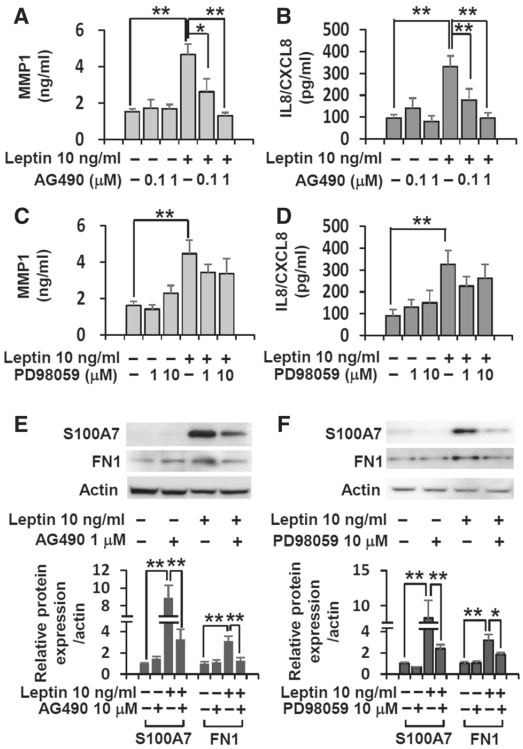


Fig. 4 Effects of AG490 and PD98059 on the inflammatory response in leptin-treated NHKs. The contribution of STAT signaling to leptin-induced upregulation was evaluated by measuring MMP-1 (**a**) and CXCL8/IL8 (**b**) in culture supernatants of NHKs co-treated with JAK2 inhibitor AG490. The contribution of ERK signaling to leptin-dependent upregulation was evaluated by treatment with ERK pathway inhibitor PD98059 (**c, d**). The effect of AG490 (**e**) and PD98059 (**f**) on S100A7 protein expression was determined by western blot analysis. Results were quantified relative to β-actin levels using ImageJ software. Values represent the mean expression level ± SD ($n = 3$). *$p \leq 0.05$ and **$p \leq 0.01$

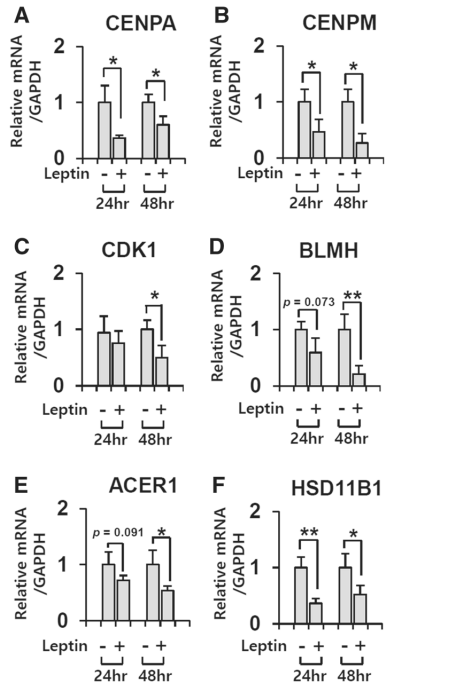


Fig. 5 Validation of the downregulated differentially expressed genes by Q-RT-PCR. To determine the mRNA levels of *CENPA* (**a**), *CENPM* (**b**), *CDK1* (**c**), *BLMH* (**d**), *ACER1* (**e**), and *HSD11B1* (**f**), total RNA samples were extracted from NHKs treated with leptin for 24 h. Values represent the mean expression level ± SD of these genes relative to that of human GAPDH. Error bars represent the SD from three independent measurements ($n = 3$). *$p \leq 0.05$ and **$p \leq 0.01$

levels of *CDK1, BLMH*, and *ACER1* were observed at 48 h after leptin treatment (Fig. 5c–E). *HSD11B1* transcription was also downregulated at both 24 and 48 h (Fig. 5f). To validate the protein expression results, we performed a western blot of the two centromere proteins, because they were the primary contributors to the result of $\chi\chi^2$ testing (Table 2). Leptin significantly decreased the expression of CENPA and CENPM proteins in NHKs (Fig. 6). CENPA and CENPM affect diverse cellular functions such as epigenetic control, chromosome segregation, and cell division cycle [5, 38]. Leptin was reported to increase re-epithelialization in a BrDU incorporation study in human keratinocytes [10, 36]. In this regard, the downregulation of CENPA and CENPM found in this study contrasts with the results in other reports. However, these studies did not directly measure the effect of leptin on cell growth. To resolve this discrepancy, a cell growth curve analysis of the leptin-treated NHKs was performed (Fig. 7). When NHKs were maintained in KGM-2 media with the growth factor cocktail, the major cell growth parameters were not significantly different in NHKs treated with 5 or 10 ng/ml of leptin (Fig. 7A). At 20 ng/ml leptin, the doubling time of NHKs increased slightly compared to the control; however, this difference was not statistically significant. In KBM without major growth factors, 5 ng/ml of leptin tended to promote proliferation due to the decrease in the doubling time compared to that of the control, although this effect was not statistically significant (Fig. 7B). Notably, in NHKs treated with the higher concentration of leptin, 20 ng/ml, the doubling time of cell growth increased. Therefore, leptin is likely to enhance cell proliferation at lower concentrations and to decrease it at higher concentrations.

Along with centromere proteins like CENPA, chromosome passenger proteins such as survivin, Aurora A kinase, and inner centromere protein (INCENP) participate

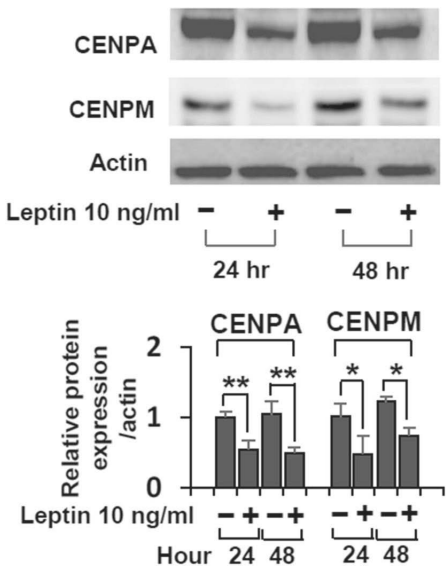
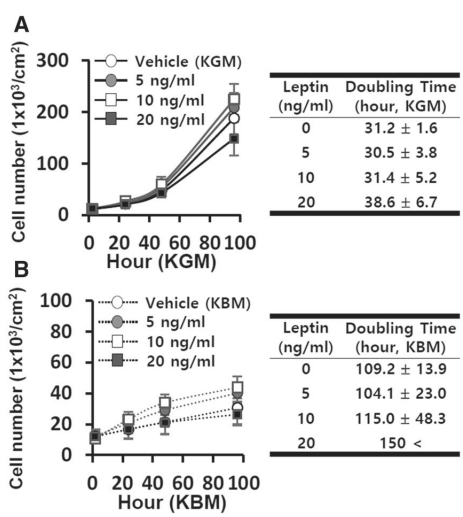


Fig. 6 Confirmation of CENPA and CENPM protein expression in leptin-treated NHKs. NHKs were treated with leptin for 24 and 48 h and then prepared for western blot analysis of CENPA and CENPM. Results were quantified relative to β-actin levels using ImageJ software. Values represent the mean expression level ± SD ($n = 3$). *$p \leq 0.05$ and **$p \leq 0.01$

in chromosome segregation during mitosis [39]. In general, chromosome passenger protein expression is affected by p53 expression. To study the effect of leptin on cell cycle regulation of NHKs, we investigated whether leptin also affected the expression of p53 and other chromosome passenger proteins (Fig. 7c). In contrast with that observed for CENPA and CENPM, leptin did not change the expression levels of p53 or the other chromosome passenger proteins survivin, Aurora A, or INCENP in NHKs. Therefore, leptin-induced downregulation of centromere proteins may not be directly associated with cell cycle regulation and proliferation.

Discussion

Obesity is associated with greater severity of human dermatologic conditions like psoriasis [35, 40]. Serum and tissue leptin levels are generally higher in obese individuals than in lean controls [23, 32]. Although leptin is secreted from adipose tissue and acts on brain to reduce appetite, the resistance to leptin is common in obese individuals, which shows that obesity and leptin are closely related to each other [25]. To understand the association between obesity and various dermatological outcomes like psoriasis, it is essential to elucidate the molecular responses induced by leptin in human epidermal keratinocytes. In this study, we determined 151 upregulated and 53 downregulated DEGs in leptin-treated NHKs in a genome-scale transcriptional analysis. The GO BP enrichment analysis of the 151 upregulated DEGs showed that leptin induces phenotypic changes to promote cutaneous inflammation and wound healing (GO:0009611, GO:0030198, GO:0030334, GO:0030593, and GO:0050729). In the validation study of upregulated DEGs, we confirmed the expression of the genes associated with inflammation and progressive wound healing. Among the upregulated DEGs, MMP-1 is an interstitial enzyme which shows a strong expression during skin inflammation



Leptin (ng/ml)	Doubling Time (hour, KGM)
0	31.2 ± 1.6
5	30.5 ± 3.8
10	31.4 ± 5.2
20	38.6 ± 6.7

Leptin (ng/ml)	Doubling Time (hour, KBM)
0	109.2 ± 13.9
5	104.1 ± 23.0
10	115.0 ± 48.3
20	150 <

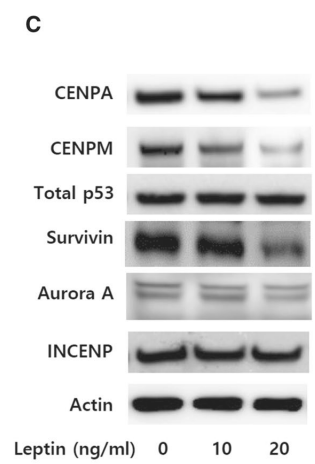


Fig. 7 Effects of leptin on cell growth and the expression of chromosome passenger protein-related proteins in NHKs. To evaluate the effects of leptin on cell growth kinetics, NHKs were cultured in KGM (**a**) and KBM with leptin (**b**). Cells were counted at 24, 48, and 96 h after leptin treatment using a Coulter counter. Values represent the mean expression level ± SD ($n = 4$). **c** To determine the effect of leptin on mitotic cell cycle regulation, protein samples were prepared for western blot of CENPA, CENPM, p53, survivin, Aurora A, and INCENP

resolution procedure [33]. CXCL8/IL has been reported as a powerful attractant of polymorphonuclear leukocytes [30], and S100A7 plays a role of preventing wound from bacterial infection during wound healing [20]. We observed an increase of mRNA expression of the genes related to wound healing, MMP-1, CXCL8/IL, and S100A7, in leptin-induced keratinocytes by Q-RT-PCR.

Normal wound healing after tissue injury involves biological responses such as hemostasis, inflammation, ECM remodeling, resolution of inflammation, and tissue regeneration [8, 14, 15, 38]. In response to leptin, NHKs exhibited a GO BP gene expression signature associated with wound healing. This relationship between leptin and cutaneous wound healing was noted in a study of leptin-deficient mice [10, 13]. Similarly, the leptin-induced upregulated DEGs indicated an active role of human keratinocytes in regulating cutaneous wound healing. Especially, the early stage of wound healing is described as an inflammatory phase [14]. A potential role for leptin in the regulation of pro-inflammatory immune responses was previously suggested in a study of murine macrophages [21]. Roles of leptin in regulating inflammation and immune responses in the periphery, especially in T cells, macrophages, and monocytes, have been reported [21, 22]. Among the leptin-induced DEGs, *CXCL8*/*IL8* and *MMP-1* are known as key pro-inflammatory mediators during the early acute phase of cutaneous inflammation [2–4, 27]. Leptin-induced expression of acute phase pro-inflammatory mediators like CXCL8/IL8 and MMP-1 was shown to be regulated by the JAK2-STAT3 signaling pathway in NHKs. In addition, neutrophil chemotaxis, a major component of an acute inflammatory response [40], was identified as a significantly enriched BP (GO:0030593). Therefore, leptin triggers intracellular signaling in NHKs to promote cutaneous inflammation, which is related to the pathophysiological outcomes of dermatological diseases in obese humans.

During wound healing, ECM is dynamically remodeled to promote cell migration and other regenerative processes in injured tissue [6, 15]. The expression of fibronectin and MMP-1, important in ECM remodeling and cell migration, was confirmed in leptin-treated NHKs. In addition to *FN1* and *MMP-1*, seven other DEGs, *ADAMTSL4, COL4A2, CRISPLD2, ELF3, LGALS3*, SMOC1, and *TNC*, were functionally annotated as genes involved in ECM organization, making ECM organization (GO:0030198), the most significant GO BP term in the χ^2 test-based GO BP enrichment analysis for leptin-induced upregulated DEGs. Angiogenesis also plays an important role in wound healing [38]. Seven upregulated DEGs, *COL4A2, CXCL8*/*IL8, CYP1B1, FN1, MMRN2, NRP2*, and *S100A7*, were annotated as genes associated with angiogenesis (GO:0001525), suggesting that leptin regulates NHKs to promote dermal angiogenesis. Therefore, further studies should be aimed to elucidate the molecular mechanism underlying coordination of cutaneous wound healing by leptin-dependent ECM remodeling and angiogenesis in response to NHKs.

Leptin-deficient (*ob*/*ob*) and leptin receptor-deficient (*db*/*db*) mice were found to be more susceptible to lipopolysaccharide (LPS)-induced death than their lean littermates, supporting the idea that leptin-induced inflammation may be protective for an organism [9]. Regarding the protective effect of a leptin-induced inflammatory response, it is noteworthy that leptin significantly upregulated S100A7 in NHKs. S100A7 has anti-microbial activity and is secreted in wound exudate, indicating that S100A may play a protective role in cutaneous wound healing [20]. The leptin-induced increase in S100A7 was attenuated by the suppression of either JAK2-STAT3 or ERK signaling in NHKs. Treatment of S100A7 in NHKs was recently reported to increase the expression of tight junction proteins such as claudins and occludin in NHKs [16]. The GO BP enrichment analysis in this study showed that cell junction assembly (GO:0034329) was identified as a significant BP in leptin-induced DEGs, with expression of *CDH24, CLDN17, MIR21*, and *VMP1* contributing, based on the statistical analysis. The leptin-dependent upregulation of S100A7 protein may be responsible for the increase in transcription of these cell junction assembly-associated genes. S100A7, also known as psoriasin, is significantly overexpressed in hyperproliferative skin diseases such as psoriasis and skin tumors [1]. In patients with severe psoriasis, levels of leptin and leptin receptors are significantly higher in skin and serum compared to these levels in mild psoriasis patients [7]. Currently, it is unclear whether the levels of cutaneous S100A7 expression in obese individuals are different from those in the non-obese. Future study should be examined in relationship between leptin-dependent upregulation of S100A7 and psoriatic pathogenesis in obese people.

Three centromere protein-associated genes, *CENPA, CENPK*, and *CENPM*, were included among the downregulated DEGs in leptin-treated NHKs. Because of the contribution of these three genes in the χ^2-based statistical assessment, the GO BP enrichment analysis of the 53 downregulated DEGs may have been biased toward an association with the mitotic cell cycle. Leptin was reported to be a mitogen in keratinocytes, which was supported by a leptin-induced increase in BrDU incorporation in human keratinocytes [36]. This BrDU incorporation assay was performed in KBM with most essential growth factors for keratinocytes excluded. In the cell growth analysis performed in KGM-2 with complete growth factor cocktail, no significant difference in cell growth was observed in the leptin-treated NHKs and control cells. In the validation study, we confirmed that leptin significantly decreased the expression of CENPA and CENPM proteins. During mitosis, CENPA interacts with chromosome passenger

complex proteins to mediate chromosome segregation [5, 39]. However, the expression levels of chromosome passenger complex proteins, survivin, Aurora A kinase, and INCENP were unchanged in this study. These results suggested that leptin-dependent downregulation of CENPA in NHKs was not coupled with the downregulation of other chromosome passenger complex proteins at the early stage wound healing. Our genome-wide transcriptional analysis was performed at 24 h after leptin treatment, so that the transcriptional phenotype of the leptin-induced DEGs would represent the inflammatory phase of the wound healing process, generally followed by the proliferative phase, although the two phases briefly overlap [14]. As inflammation is resolved in the wound healing process, cell cycle progression and epithelial proliferation may be essential to cutaneous re-epithelialization. However, after injury, DNA replication and cell cycle progression may be temporarily delayed for acute inflammation to protect parenchymal cells from mutations, because cell cycle checkpoints are generally activated in response to external stresses such as those that cause wounds [15, 31]. As inflammation is resolved, injured parenchymal or fibrotic cells proliferate for tissue regeneration [14]. As wound healing advances from the inflammatory to proliferative phase, the cytokines and growth factors at the wound site vary, modulating the cellular response [14]. Therefore, because of the different cytokines present in each phase, it is possible that leptin has different effects in the inflammatory and proliferative phases of wound healing. To explore this possibility, further studies should be directed at evaluating the long-term treatment effects of leptin in a more complex culture system that mimics in vivo cutaneous wound healing conditions.

A potential limitation of this study is that cutaneous would healing is often defective in obese individuals whose serum or tissue leptin levels are generally higher than those in non-obese individuals [11, 24, 32]. However, some obese patients show leptin resistance [34]. Therefore, the possibility that leptin, an important regulator of wound healing, is not fully active in obese individuals should also be examined.

In conclusion, our genome-scale transcriptional study showed that leptin induces a pro-inflammatory response in NHKs. In addition, levels of S100A7, an important psoriatic protein with anti-microbial and immunomodulatory functions, increased in a concentration-dependent manner in NHKs in response to leptin. Leptin also upregulates the level of ECM components, which are important in wound healing, in NHKs. Therefore, we conclude that leptin regulates the pro-inflammatory response of human epidermal keratinocytes.

Acknowledgements This study was supported by a grant from the Korea Healthcare Technology R&D Project, Ministry of Health & Welfare, Republic of Korea (Grant No. HN13C0077), a National Research Foundation of Korea (NRF) grant funded by the Ministry of Science, ICT and Future Planning (Grant No. 2014M3C9A2064603), a NRF Grant (2015R1A2A2A01008408), and a grant by Promising-Pioneering Researcher Program through the Seoul National University (SNU) in 2015.

Compliance with ethical standards

Conflict of interest The authors have no conflicts of interest.

References

1. Anderson KS, Wong J, Polyak K, Aronzon D, Enerback C (2009) Detection of psoriasin/S100A7 in the sera of patients with psoriasis. Br J Dermatol 160:325–332. https://doi.org/10.1111/j.1365-2133.2008.08904.x
2. Bae ON, Noh M, Chun YJ, Jeong TC (2015) Keratinocytic vascular endothelial growth factor as a novel biomarker for pathological skin condition. Biomol Ther 23:12–18. https://doi.org/10.4062/biomolther.2014.102
3. Barker JN, Mitra RS, Griffiths CE, Dixit VM, Nickoloff BJ (1991) Keratinocytes as initiators of inflammation. Lancet 337:211–214
4. Biasi D, Carletto A, Caramaschi P, Bellavite P, Maleknia T, Scambi C, Favalli N, Bambara LM (1998) Neutrophil functions and IL-8 in psoriatic arthritis and in cutaneous psoriasis. Inflammation 22:533–543
5. Black BE, Cleveland DW (2011) Epigenetic centromere propagation and the nature of CENP-a nucleosomes. Cell 144:471–479. https://doi.org/10.1016/j.cell.2011.02.002
6. Bonnans C, Chou J, Werb Z (2014) Remodelling the extracellular matrix in development and disease. Nat Rev Mol Cell Biol 15:786–801. https://doi.org/10.1038/nrm3904
7. Cerman AA, Bozkurt S, Sav A, Tulunay A, Elbasi MO, Ergun T (2008) Serum leptin levels, skin leptin and leptin receptor expression in psoriasis. Br J Dermatol 159:820–826. https://doi.org/10.1111/j.1365-2133.2008.08742.x
8. Eming SA, Krieg T, Davidson JM (2007) Inflammation in wound repair: molecular and cellular mechanisms. J Investig Dermatol 127:514–525. https://doi.org/10.1038/sj.jid.5700701
9. Faggioni R, Fantuzzi G, Gabay C, Moser A, Dinarello CA, Feingold KR, Grunfeld C (1999) Leptin deficiency enhances sensitivity to endotoxin-induced lethality. Am J Physiol 276:R136-R142
10. Frank S, Stallmeyer B, Kampfer H, Kolb N, Pfeilschifter J (2000) Leptin enhances wound re-epithelialization and constitutes a direct function of leptin in skin repair. J Clin Investig 106:501–509. https://doi.org/10.1172/JCI9148
11. Friedman JM, Halaas JL (1998) Leptin and the regulation of body weight in mammals. Nature 395:763–770. https://doi.org/10.1038/27376
12. Glasow A, Kiess W, Anderegg U, Berthold A, Bottner A, Kratzsch J (2001) Expression of leptin (Ob) and leptin receptor (Ob-R) in human fibroblasts: regulation of leptin secretion by insulin. J Clin Endocrinol Metab 86:4472–4479. https://doi.org/10.1210/jcem.86.9.7792
13. Goren I, Kampfer H, Podda M, Pfeilschifter J, Frank S (2003) Leptin and wound inflammation in diabetic ob/ob mice: differential regulation of neutrophil and macrophage influx and a potential role for the scab as a sink for inflammatory cells and mediators. Diabetes 52:2821–2832

14. Guo S, Dipietro LA (2010) Factors affecting wound healing. J Dent Res 89:219–229. https://doi.org/10.1177/0022034509359125
15. Gurtner GC, Werner S, Barrandon Y, Longaker MT (2008) Wound repair and regeneration. Nature 453:314–321. https://doi.org/10.1038/nature07039
16. Hattori F, Kiatsurayanon C, Okumura K, Ogawa H, Ikeda S, Okamoto K, Niyonsaba F (2014) The antimicrobial protein S100A7/psoriasin enhances the expression of keratinocyte differentiation markers and strengthens the skin's tight junction barrier. Br J Dermatol 171:742–753. https://doi.org/10.1111/bjd.13125
17. Iguchi M, Aiba S, Yoshino Y, Tagami H (2001) Human follicular papilla cells carry out nonadipose tissue production of leptin. J Investig Dermatol 117:1349–1356. https://doi.org/10.1046/j.0022-202x.2001.01606.x
18. Johnston A, Arnadottir S, Gudjonsson JE, Aphale A, Sigmarsdottir AA, Gunnarsson SI, Steinsson JT, Elder JT, Valdimarsson H (2008) Obesity in psoriasis: leptin and resistin as mediators of cutaneous inflammation. Br J Dermatol 159:342–350. https://doi.org/10.1111/j.1365-2133.2008.08655.x
19. Lee E, Kim HJ, Lee M, Jin SH, Hong SH, Ahn S, Kim SO, Shin DW, Lee ST, Noh M (2016) Cystathionine metabolic enzymes play a role in the inflammation resolution of human keratinocytes in response to sub-cytotoxic formaldehyde exposure. Toxicol Appl Pharmacol 310:185–194. https://doi.org/10.1016/j.taap.2016.09.017
20. Lee KC, Eckert RL (2007) S100A7 (Psoriasin)-mechanism of antibacterial action in wounds. J Investig Dermatol 127:945–957. https://doi.org/10.1038/sj.jid.5700663
21. Loffreda S, Yang SQ, Lin HZ, Karp CL, Brengman ML, Wang DJ, Klein AS, Bulkley GB, Bao C, Noble PW, Lane MD, Diehl AM (1998) Leptin regulates proinflammatory immune responses. FASEB J 12:57–65
22. Lord GM, Matarese G, Howard JK, Baker RJ, Bloom SR, Lechler RI (1998) Leptin modulates the T-cell immune response and reverses starvation-induced immunosuppression. Nature 394:897–901. https://doi.org/10.1038/29795
23. Mitsuyama S, Abe F, Kimura M, Yoshida M, Higuchi T (2015) Association between leptin gene expression in subcutaneous adipose tissue and circulating leptin levels in obese patients with psoriasis. Arch Dermatol Res 307:539–544. https://doi.org/10.1007/s00403-015-1581-z
24. Monti V, Carlson JJ, Hunt SC, Adams TD (2006) Relationship of ghrelin and leptin hormones with body mass index and waist circumference in a random sample of adults. J Am Diet Assoc 106:822–828. https://doi.org/10.1016/j.jada.2006.03.015; quiz 829–830.
25. Munzberg H, Myers MG Jr (2005) Molecular and anatomical determinants of central leptin resistance. Nat Neurosci 8:566–570. https://doi.org/10.1038/nn1454
26. Murad A, Nath AK, Cha ST, Demir E, Flores-Riveros J, Sierra-Honigmann MR (2003) Leptin is an autocrine/paracrine regulator of wound healing. FASEB J 17:1895–1897. https://doi.org/10.1096/fj.03-0068fje
27. Pasparakis M, Haase I, Nestle FO (2014) Mechanisms regulating skin immunity and inflammation. Nat Rev Immunol 14:289–301. https://doi.org/10.1038/nri3646
28. Peelman F, Couturier C, Dam J, Zabeau L, Tavernier J, Jockers R (2006) Techniques: new pharmacological perspectives for the leptin receptor. Trends Pharmacol Sci 27:218–225. https://doi.org/10.1016/j.tips.2006.02.009
29. Poeggeler B, Schulz C, Pappolla MA, Bodo E, Tiede S, Lehnert H, Paus R (2010) Leptin and the skin: a new frontier. Exp Dermatol 19:12–18. https://doi.org/10.1111/j.1600-0625.2009.00930.x
30. Rennekampff HO, Hansbrough JF, Kiessig V, Dore C, Sticherling M, Schroder JM (2000) Bioactive interleukin-8 is expressed in wounds and enhances wound healing. J Surg Res 93:41–54. https://doi.org/10.1006/jsre.2000.5892
31. Rodier F, Campisi J (2011) Four faces of cellular senescence. J Cell Biol 192:547–556. https://doi.org/10.1083/jcb.201009094
32. Rosenbaum M, Leibel RL (2014) 20 years of leptin: role of leptin in energy homeostasis in humans. J Endocrinol 223:T83-96. https://doi.org/10.1530/JOE-14-0358
33. Saarialho-Kere UK, Chang ES, Welgus HG, Parks WC (1992) Distinct localization of collagenase and tissue inhibitor of metalloproteinases expression in wound healing associated with ulcerative pyogenic granuloma. J Clin Investig 90:1952–1957. https://doi.org/10.1172/JCI116073
34. Sainz N, Barrenetxe J, Moreno-Aliaga MJ, Martinez JA (2015) Leptin resistance and diet-induced obesity: central and peripheral actions of leptin. Metabolism 64:35–46. https://doi.org/10.1016/j.metabol.2014.10.015
35. Shipman AR, Millington GW (2011) Obesity and the skin. Br J Dermatol 165:743–750. https://doi.org/10.1111/j.1365-2133.2011.10393.x
36. Stallmeyer B, Kampfer H, Podda M, Kaufmann R, Pfeilschifter J, Frank S (2001) A novel keratinocyte mitogen: regulation of leptin and its functional receptor in skin repair. J Investig Dermatol 117:98–105. https://doi.org/10.1046/j.0022-202x.2001.01387.x
37. Takeda S, Elefteriou F, Levasseur R, Liu X, Zhao L, Parker KL, Armstrong D, Ducy P, Karsenty G (2002) Leptin regulates bone formation via the sympathetic nervous system. Cell 111:305–317
38. Tonnesen MG, Feng X, Clark RA (2000) Angiogenesis in wound healing. J Investig Dermatol Symp Proc 5:40–46. https://doi.org/10.1046/j.1087-0024.2000.00014.x
39. Verdaasdonk JS, Bloom K (2011) Centromeres: unique chromatin structures that drive chromosome segregation. Nat Rev Mol Cell Biol 12:320–332. https://doi.org/10.1038/nrm3107
40. Wright HL, Moots RJ, Bucknall RC, Edwards SW (2010) Neutrophil function in inflammation and inflammatory diseases. Rheumatology 49:1618–1631. https://doi.org/10.1093/rheumatology/keq045
41. Yosipovitch G, DeVore A, Dawn A (2007) Obesity and the skin: skin physiology and skin manifestations of obesity. J Am Acad Dermatol 56:901–916. https://doi.org/10.1016/j.jaad.2006.12.004