Signature Genes in Macrodactyly through Transcriptome Network Analysis Reveal their Association of Lipid Metabolism

Jeong-Woo Choi¹, Hyun-Joo Lee², Ji Won Oh¹

¹Department of Anatomy, School of Medicine, Kyungpook National University
²Department of Orthopaedic Surgery, School of Medicine, Kyungpook National University

(Received 14 July 2017, revised 14 September 2017, accepted 19 September 2017)

Abstract : Macrodactyly is one of the most difficult hand anomalies to treat not only surgically but medically as well. Little is known about the molecular pathways and lipid metabolism of this disease. To elucidate the potential mechanism of macrodactyly progress, we used the bioinformatical analysis including quantile normalization, principal component analysis, heatmap and volcano plot. For the functional bioinformatical study, lipid, lipoprotein and phospholipid metabolism of Kyoto Encyclopedia of Genes and Genomes, Wiki Pathways, and Reactome Pathway were utilized to compare the differentially expressed genes in macrodactyly with control group. We found up-regulation of CDK6 and E2F1, which are associated with the mitotic cell cycle of cancer cells. PIK3CG, associated with cancer and lipid metabolism, was also enriched in macrodactyly. In down-regulated genes, PTEN was highlighted in lipid metabolism, phosphatidylinositol signaling system and insulin signaling. ABCD3, related in peroxisomal import of fatty acids, was also down-regulated. In this study, we predicted the pathogenic candidate genes as well as the potential molecular pathways related to macrodactyly by identifying the signature genes. Signature genes through systems bioinformatical analysis can be utilized to catch the insight of the molecular pathogenesis of macrodactyly.

Keywords : Macrodactyly, Gigantism, Computational biology, Lipid metabolism

Introduction

In hand surgery, macrodactyly is one of the most difficult medical conditions [1,2]. Enlargement of all tissues including nerve in digits is the pathoanatomic feature different from other tumorous conditions. Surgery has been widely used and considered as the most reliable treatment modality [3]. However, the etiology of this congenital anomaly has not been established. Many authors suggested the hypertrophied nerve as a causal factor or target for treatment [1,4-7] (Fig. 1).

Macrodactyly is considered as a “rare disease” by the Office of Rare Diseases (ORD) of the National Institutes of Health. Macrodactyly has been known to have an incidence rate of 0.9% [8]. Hardwick et al. described a male predominance (1 : 0.66) in the upper limb in their study [8]. Barsky also reported male predominance (32 : 25) in macrodactyly almost 50 years ago [7]. However, these
studies have a limitation of small number of cases due to the rarity of this disease.

Rios et al. reported somatic gain-of-function mutations in the Phosphatidylinositol 3-kinase (PIK3CA) of the affected nerve in patients with macrodactyly as a causal factor of macrodactyly [9]. Lau et al. described pleiotrophin (PTN), a developmental cytokine, was significantly overexpressed across all their macrodactyly samples [10]. However, previous research has provided the limited pathological insights to describe the molecular mechanism in a better way. In addition, we do not have cellular and genetically modified animal models of macrodactyly, thus making molecular macrodactyly studies difficult.

With advancement in high throughput data techniques including microarray, next generation sequencing (NGS) [11], and proteomics, many researchers in the field of molecular biology have tried to use the interactive network analysis of specific diseases [12]. Recently, our group also found the novel mechanistic insight using meta-analysis of the high throughput data [13,14]. Utilizing the bioinformatical network analysis of differentially expressed genes (DEGs), researcher can find the distinct molecular network connectivities, induced by the unique microenvironments of the disease and physiologic conditions [13-16]. The network analysis of transcription changes of macrodactyly compared with normal fat tissue is helpful to determine the potential pathogenic mechanism of the disease.

In this study, we have identified the potential signature genes of the fat tissue of macrodactyly. At the same time, we analyzed the comprehensive transcriptome meta-network analysis of the macrodactyly using the publicly available data. We predicted several pathogenic candidate genes as well as potential molecular pathways related to this disease.

**Materials and Methods**

**Surgical approach of the patient**

This study was approved by the Institutional Review Board of the Kyungpook National University Hospital (2016-01-016). Informed consents were acquired from the patient’s parent to use the photograph of the surgical procedures. The surgical technique was previously described [3]. Briefly, under general anesthesia, a lateral midline incision was made in a zigzag pattern on the convex side (Fig. 1A). The incision also included an adequate amount of nail that needed to be removed. To decrease the width of the finger, debulking was performed. The fat tissue of the convex and, volar sides and the tip of the finger were resected. The skin was removed in a zigzag pattern (Fig. 1A). The nail was also removed along the incision on its convex side and tip. After the digital artery and hypertrophied nerve were identified, the nerve was removed (Fig. 1B). When the enlargement involved the palm, the fat tissue and skin of the palm were disposed.
the finger, a combination of several bone procedures was used.

**Expression microarray data from Gene expression omnibus (GEO)**

In this study, we only analyzed the microarray expression data from publicly accessible database of Gene Expression Omnibus (GEO). We utilized the microarray gene sets of GSE35820 (macrodactyly), GSE15773 (fat group 1), GSE13506 (fat group 2) and GSE14905 (skin group). In macrodactyly expression data (GSE35820), total RNA of 4 pediatric patients, including a 15-month-old boy and an 8-year-old girl who underwent elective surgical debulking at Massachusetts General Hospital or at Children’s Hospital Boston between June 2009 and April 2011, were extracted [10]. In the fat group (GSE15773 [17] and GSE13506 [18]), adipose tissue samples were taken from subcutaneous fat of lower abdominal wall. Whole adipose tissue was dissected to remove fibrotic tissue and obvious vasculature, and frozen in 0.5 cm³ sections in liquid nitrogen. The data of fat group 1 were mainly obtained from University of Massachusetts Memorial Medical Center, a tertiary care academic hospital in a metropolitan setting in central New England [17], while fat group 2 data were obtained from multiple hospitals including Helsinki and Turku University Hospitals and Utrecht Academic University Hospital [18]. In case of skin sample (GSE14905), full thickness skin tissues were obtained from ILSBio and Rockefeller University [19].

**Statistical analysis and visualization**

We employed and modified the methods previously reported by our group [13-15]. First, we sorted the expression data in excel sheet. Then, we calculated the average of expression value in each gene. We considered a specific gene as a DEG if fold change is larger than 2 folds compared with control group. DEG was considered as a statistically significant gene when it has a calculated p-value under 0.05. We utilized R 3.4 program for quantile normalization and Principal Component Analysis (PCA). We also constructed heatmap and volcano plot for the visualization of DEGs. R Studio was used to visualize PCA, heatmap, and volcano plot. The design of PCA and heatmap referenced the previously reported information [14]. Adobe Illustrator CS6 was used to draw a summary figure.

**Pathway analysis with gene-set enrichment analysis**

To draw the bubble chart, we used a Gene Ontology (GO) biological process. Associated genes in lipid, lipoprotein and phospholipid metabolism were utilized to compare DEGs with Gene Set Enrichment Analysis (GSEA) information. Kyoto Encyclopedia of Genes and Genomes (KEGG) 2016, Wiki Pathways 2016, Reactome Pathway and Protein-Protein Interaction (PPI) hub proteins were used for pathway analysis. Network analyses of two DEGs, up-regulated and down-regulated genes, were performed independently. Figure 4 was displayed by Network2Canvas tools [20].

**Results**

**Analysis of differentially expressed genes in macrodactyly compared to control fat tissue**

We used GEO database to identify signature genes in macrodactyly. GSE35820 was used for macrodactyly information, while GSE15773, GSE13506, and GSE14905 were used for fat group 1, 2 and control skin group, respectively. First, we used PCA to confirm the global differences in each group. PCA clearly distinguished control fat groups, skin, and macrodactyly group (Fig. 2A). Next, we investigated DEGs of macrodactyly compared to fat group 1 and fat group 2. Total 781 genes were significantly up-regulated in macrodactyly and 1,094 genes were significantly down-regulated in the macrodactyly group compared to fat group 1. When we compared to fat group 2, total 662 genes were significantly increased and 966 genes were significantly decreased (Fig. 2B). Next, we investigated overlap DEGs of macrodactyly compared to fat group 1 and fat group 2. Venn diagram displayed the overlap up-regulated 456 genes and down-regulated 748 genes (Fig. 2C). We defined these gene groups as the signature up-regulated and down-regulated genes.

**Bioinformatical analysis of macrodactyly gene set**

Volcano plots were used to confirm the representative DEGs in each group (Fig. 3A and 3B). We have marked
the genes in fat group 1 under the conditions of $\log_{10}(p\text{-value}) > 6$ and fold change (under $-4$ or over 4) and in fat group 2 under the conditions of $\log_{10}(p\text{-value}) > 30$ and fold change (under $-4$ or over 4). Overlap DEGs were used to investigate GO biological process and to visualize bubble chart. The up-regulated genes showed phosphatidylinositol 3-kinase signaling, negative regulation of cell proliferation, type B pancreatic cell development, regulation of transforming growth factor beta receptor, and G1/S transition of mitotic cell cycle. In down-regulated genes, positive regulation of axon extension, cell migration, negative regulation of cell proliferation, apoptotic process, neutrophil degranulation, negative regulation of ERK1 and ERK2 cascade were enriched (Fig. 3C). Then, we used GSEA information to look for the genes that are related to lipid metabolism. We identified total 18 up-regulated genes and 45 down-regulated genes (Fig. 3D).
Comparative pathway enrichment analysis of macrodactyly signature gene sets

Next, we conducted pathway analysis to compare the overlap DEGs (Fig. 4A and 4B). Using KEGG, Wiki and Reactome pathways, we identified the potential molecular mechanism of macrodactyly. Especially, in up-regulated gene sets, polyunsaturated fatty acid biosynthesis as well as Wnt signaling pathway were highlighted. Concurrently, mitotic cell cycle and DNA replication were also enriched, implying that cellular proliferation was activated.

We analyzed the specific genes of macrodactyly group compared with fat group 1 as well as fat group 2. Intriguingly, we found that CDK6 and E2F1 genes, up-regulated in macrodactyly, are associated with mitotic cell cycle in cancer cell (Chronic myeloid leukemia, pancreatic cancer, melanoma, non-small cell lung cancer, small cell lung cancer, bladder cancer, and prostate cancer). In down-regulated genes, CSNK1A1, ABCD3, and MKNK1 were involved in the Hedgehog signaling pathway as well as nuclear receptors in lipid metabolism and toxicity. Based on fat group 2 comparison, in down-regulated group, MAGI1, ABCD3, and MKNK1 were involved in Dentatorubral-palidolouysian atrophy. ABC transporters, P38 MAPK signaling pathway were also highlighted in down-regulated group.
Discussion

Using meta-network analysis followed by volcano plot, bubble chart and GSEA comparison, we summarized the genes explaining the potential pathogenesis of macrodactyly (Fig. 5). The expression patterns of soft tissue of macrodactyly are very prominent in the global scale, and it leads us to identify the signature genes specifically enhanced only in macrodactyly compared with normal subcutaneous fat and skin.

Among them, network analysis using up-regulated signature genes reveals that the mitotic cell cycle, PI3K signaling pathway [9] and growth factor beta receptor transforming regulation are enriched in macrodactyly status. Intriguingly, negative regulation of cellular proliferation is also highlighted. It seems to be paradoxical considering the morphological feature of the disease, enlargement. This might be explained by the negative feedback loop of critical genes involved in the chronic proliferation of macrodactyly. Especially, in the chronic status of liver and colon cancer, there is a high negative feedback turning-on Wnt signaling, leading to cancer development and growth [21].

Macrodactyly is classified into 4 types by Flatt classification [22]. Type I is the most common and representative type of macrodactyly, manifested by gigantism and lipofibromatosis. However, other diseases including neurofibromatosis, hyperostosis, and hemihypertrophy also have roles for generation of macrodactyly, which correspond to type II, II, and IV respectively [22].

In molecular level, EVOVL6 previously known as LCE and FACE, the up-regulated signature gene in macrodactyly...
tyly, is one of the family member of ELOVL fatty acid elongase. This gene is considered as a working member of de novo lipogenesis through the elongation of saturated long-chain [23]. In a recent study, ELOVL6 gene polymorphism is associated with type 2 diabetes mellitus [24]. Meanwhile, animal studies have indicated that overexpression of murine Elovl6 gene can cause insulin resistance [25]. Combining all the recent studies, enlargement of finger in macrodactyly can be associated with insulin signaling mechanism considering the role of ELOVL6 gene in de novo lipogenesis and insulin resistance like pseudo-acromegaly [26]. Interestingly, the fat tissue was not distinguishable from normal fat tissue, except for the increased size of fat lobules in our case (Fig. 1B).

In terms of soft tissue regeneration, fat is highly linked with the skin appendage, especially the hair follicles. There is a small number of hair follicles with very sparse density in the normal hand and foot, compared with the rest of the body. We recently reported that fat regeneration can be committed through hair follicles and finely regulated by myofibroblasts [13]. Simultaneously, in human, subcutaneous fat regeneration is highly synchronized with the cyclic patterns of hair follicles [27]. Along these lines, we can find Wnt signaling pathway, a critical developmental signaling mechanism of fat and hair, is also highlighted in up-regulated DEGs in macrodactyly. Further study including three dimension tissue scaffold experiment [28] or organotypic skin culture with adipose tissue [29] can elucidate the novel mechanistic correlation between soft tissue enlargement of macrodactyly finger and sparse density of hair follicles at the finger.

In this study, we predicted the pathogenic candidate genes as well as the potential molecular pathways related to macrodactyly by identifying the signature genes of the fat tissue of macrodactyly. Signature genes through systems bioinformatical analysis can be utilized to catch an insight to the molecular pathogenesis of macrodactyly. We believe the future research applying in vitro cultivation of fat from macrodactyly compared with normal pediatric adipose tissues and genetically modified animal models of macrodactyly will shed the light on the mechanistic study of this rare disease. Our study will help to broaden our knowledge with the molecular mechanism of the macrodactyly.

**Conflict of Interest**

The authors declare no potential conflicts of interest.

**Acknowledgments**

Authors thank Buu Le Dao, Dong-Woo Yang, Nanda Maya Mali, Hyeon Sung Lee and Brian Vu for critical reading of the manuscript.

**REFERENCES**

4. Inglis K. Local Gigantism (A Manifestation of Neurofibromatosis): Its Relation to General Gigantism and to Acro-
전사체 메타분석을 통한 큰손가락증과 지질 대사 간의 관련성 연구 및 signature genes 발굴

최정우 1, 이현주 2, 오지원 1

1 경북대학교 의학전문대학원 해부학교실, 2 경북대학교 의학전문대학원 정형외과학교실

간추림 : 큰손가락증(Macroductyly)은 치료하기 어려운 선천성 기형 중의 하나로 그 원인이 정확히 밝혀져 있지 않다. 본 연구에서는 Gene Expression Omnibus (GEO)에 공개된 microarray 정보를 이용하여 큰손가락증에서 비정상적으로 발현되는 유전자를 찾고, 이 유전자들이 지방 대사와 어떻게 연관되어 있는지 조사하였다. 또한, 메타분석 방법을 사용하여 네트워크를 아우르는 종합적인 분석을 진행하였다. GEO 정보를 이용하여 대조군 지방 전사체 정보와 큰손가락증 전사체 정보를 분석하고, R 프로그램을 사용하여 quantile normalization과 principal component analysis를 시행한 후, heatmap과 volcano plot으로 differentially expressed genes (DEGs)를 분석하고 시각화하였다. 큰손가락증 유전자 GO biological process의 정보를 정리하여 bubble chart로 표현하였고, GSEA의 지방 대사 유전자군을 이용하여 큰손가락증 유의 유전자와 비교하였다. 그 결과, 유사분열 세포주기와 혈중글로벌에 작용하는 CDK6, E2F1 유전자와 세포주기와 총량에 영향을 주면서 지방 대사에도 관여하는 PIK3CG 유전자와 유의적으로 큰손가락증에서 높게 발현되는 것을 확인했다. 또한 대조군에 비해 큰손가락증에서 유의적으로 높게 발현되는 유전자 중 PTEN과 PIK3C2A는 지방 대사, phosphatidylinositol signaling system과 insulin signaling에 유의적으로 포함된 것을 확인했다. ABCD3는 peroxisomal import of fatty acids에 작용하는 것을 확인했다. 본 연구에서는 큰손가락증에서 비정상적으로 발현되는 유전자들 중, 이 질병에 관여할 가능성이 높은 유전자(signature genes)를 발굴하였고 이는 희귀질환 중 하나인 큰손가락증의 기전을 이해하는 데 도움이 될 수 있을 것이다.

 찾아보기 낱말 : 큰손가락증, 메타분석, 경로분석

교신저자 : 오지원(경북대학교 의학전문대학원 해부학교실)
이현주(경북대학교 의학전문대학원 정형외과학교실)
전자우편 : ohjiwon@knu.ac.kr, hjlee@knu.ac.kr