International Journal of Pediatric Otorhinolaryngology 99 (2017) 30–35

Contents lists available at ScienceDirect

International Journal of Pediatric Otorhinolaryngology

journal homepage: http://www.ijporlonline.com/

Uncovering the pathogenesis of microtia using bioinformatics approach

Liu Lei [a, *], Liu Zhenzhong [a], Lin Lin [b], Pan Bo [b]

[a] Department of Burns and Plastic Surgery, Second Hospital of Shandong University, Jinan, Shandong Province, People's Republic of China
[b] Department of Auricular Reconstruction, Plastic Surgery Hospital, Peking Union Medical College and Chinese Academy of Medical Science, Beijing, People's Republic of China

ARTICLE INFO

Article history:
Received 21 January 2017
Received in revised form
7 May 2017
Accepted 20 May 2017
Available online 22 May 2017

Keywords:
Microtia
Bioinformatics
Protein-protein interaction
Gene ontology (GO)
Kyoto encyclopedia of genes and genomes

ABSTRACT

Objective: Bioinformatics is widely used in the field of cancer research, but in the research of pathogenesis of congenital malformations the situation is different. The aim of this study was to explore the underlying mechanism using bioinformatics approach.
Methods: The data were available from Mouse Genome Informatics and Pubmed. Protein–protein interaction (PPI) network of pathogenic genes was conducted using STRING. Gene ontology and pathway enrichment analyses were also performed to pathogenic genes.
Results: Total 63 genes were identified as pathogenic genes in the study. The PPI networks for pathogenic genes were constructed, which contained 62 nodes and 228 edges with *PAX6*, *FGFR1* and *CTNNB1* as the hub genes. All the genes were linked to 921 pathways in biological processes, 31 pathways in cell component, 41 pathways in molecular function, and 76 pathways in the KEGG. These genes were discovered significantly enriched in embryonic organ development, ear morphogenesis, ear development, and regulation of RNA synthesis and processing.
Conclusions: bioinformatics methods were utilized to analysis pathogenic genes involved in microtia development, including pathogenic genes identifying, PPI network construction and functional analysis. And we also predicted that several potential mechanisms might contribute to occurrence of microtia by disturbing GO terms and pathways. This approach could be useful for the study of the etiology and pathogenesis of microtia.

Microtia(OMIM 600674, OMIM 251800) is a developmental malformation of the external ear, characterized by a small, abnormal shaped auricle. Until now the development of the auricle and the pathogenesis of microtia are still not clear. Both single-gene defects and chromosomal aberrations have been reported in microtia and microtia-associated syndromes [1]. Understanding the genetics of microtia will provide tools for appropriate genetic counseling and molecular genetic testing that can be helpful to confirm an exact diagnosis.

In cancer research area, with the advance of bio-technology in the past decades, an abundance of gene-expression data have been collected and submitted to several databases to date. These data have been widely utilized in bioinformatics analyses and explorations for the molecular mechanism of carcinogenesis, showing many details of the molecular mechanism of carcinogenesis [2,3]. However bioinformatics analyses and explorations have not been widely used in researches of pathogenesis of microtia.

Thus, in the present study, we searched a database of mouse mutants, using the terms 'microtia' and 'anotia', for gene names associated with these terms; we also searched PubMed for genes associated with either of these two terms. We then used the expanded list of gene names to examine a database of known and predicted protein-protein interactions, drawing out other genes which were known experimentally (or by expression, or other lines of evidence) to possibly 'interact' with the primary query genes. Bioinformatics approaches, including functional enrichment analysis and protein–protein interaction (PPI) network construction were utilized in the research. The results of our study might provide new clues for etiology study of microtia and uncover the potential mechanism of the disease.

* Corresponding author.
E-mail address: liujilei_666@hotmail.com (L. Lei).

http://dx.doi.org/10.1016/j.ijporl.2017.05.009
0165-5876/

1. Methods

1.1. Data collecting

Two methods were used to collect gene data. First a search was conducted using Mouse Genome Informatics (MGI, version 6.07, http://www.informatics.jax.org/), which is an international database resource for the laboratory mouse, providing integrated genetic, genomic, and biological data to facilitate the study of human health and disease. “microtia” and “anotia” were chosen as the key words to conduct search. Second a systematic literature search was conducted using the Pubmed, among articles published up until December 2016. We first chose “microtia” and “anotia” as the key words to conduct search. Then we chose “gene name” and “microtia” as the key words to conduct search, so as to obtain the further references related to gene and disease.

1.2. Protein–protein interaction network constructing

Functional links between proteins are usually inferred from genomic associations between the genes that encode them: groups of genes that are required for the same function tend to be show similar species coverage, are often located in close proximity on the genome. The PPIs were selected using STRING (http://string-db.org/) [4], which was updated continuously. In STRING, each protein-protein interaction is annotated with one or more 'scores'. Importantly, these scores do not indicate the strength or the specificity of the interaction. Instead, they are indicators of confidence, i.e. how likely STRING judges an interaction to be true, given the available evidence. In the present study, the PPIs with medium Confidence Scores more than 0.4 were selected to maximize the possibility of mining data. In addition, the hub nodes with the most PPIs and highest confidence score were selected, which represent a small proportion of nodes with maximal information exchange with others and are expected to play an important role in biology as well.

1.3. Functional analysis

Functional analysis was also performed using STRING, which included Gene Ontology (GO) [5] and Kyoto encyclopedia of genes and genomes (KEGG) pathway [6] analyses. Biological process, which refers to a biological objective to which the gene or gene product contributes, molecular function, as the biochemical activity (including specific binding to ligands or structures) of a gene product, and cellular component, which refers to the place in the cell where a gene product is active are the three categories of GO. Additionally, KEGG pathway database consists of graphical diagrams of biochemical pathways, including most of the known metabolic pathways and some of the known regulatory pathways. In the present study, these function analyses were performed to all genes, and p values less than 0.05 were set as thresholds.

2. Results

2.1. Identification of pathogenic genes of microtia

Total 62 pathogenic genes involved in occurrence of microtia were selected, including 36 genes from MGI and 27 genes from articles searched in Pubmed (Table 1).

2.2. PPI networks

The PPI networks for pathogenic genes were constructed, which contained 62 nodes and 228 edges with *PAX6*, *FGFR1* and *CTNNB1* as the hub genes (Fig. 1). Network nodes represent proteins and edges represent protein-protein associations. The average node degree is

Table 1
Selected pathogenic genes.

Gene symbol	Reference
BCL2, *BMP5*, *BMPR1A*, *CHRD*, *CTGF*, *DSCAM*, *EH*, *FGF10*, *FGF8*, *FGFR1*, *FLG*, *FOXP3* *FREM2*, *GBX2*, *HIC1*, *HMGA2*, *HOXA1*, *HOXA2*, *IDUA*, *ITPR3*, *LMNA*, *MALF*, *MMP14*, *MP* *MYO6*, *PAX6*, *PAX8*, *PRKDC*, *PRKRA*, *PRRX1* *PRRX2*, *PTPN11*, *RGSC1854*, *RGSC252*, *SBSE*, *SFN*, *TBX1*, *TYRP1*, *WNT5A*, *ORC1*, *FGF3*, *HSPA9*	MGI Direct Data
APC	[5]
CTNNB1	[6]
DLX5	[7]
EGFR	[8]
EDN1	[9]
EDNRA	[10]
EYA1	[11]
FOXI3	[12]
GSC	[13]
MSX1	[14]
HFM	[15]
HMX1	[16]
IRF6	[17]
MAP2K1	[18]
MAPK1	[18]
MAPK3	[18]
MESP1	[19]
PBX1	[20]
SIX1	[21]
SIX4	[21]
FOXG1	[22]
TCOF1	[23]
TFAP2A	[24]
TRP53	[25]
BAPX1	[26]
SALL1	[27]
CHD7	[28]