Int J Hum Genet, 18(2): 35-42 (2018)
DOI: 10.31901/24566330.2018/18.2.685

Abnormal DNA Methylations Associated with Allergic Asthma Children

Qing-Feng Bai[1], Jin-Xing Feng[2] and Zhao-Yang Yin[1]

[1]Department of Pediatrics, Shangluo Central Hospital, Shangluo, Shaanxi, China
[2]Neonatal Department, Shenzhen Children's Hospital, Shenzhen, Guangdong, China

KEYWORDS Allergic Asthma. DNA Methylation. Enrichment Analysis. Gene Expression Omnibus Database. Hub Gene

ABSTRACT The purpose of this paper was to investigate abnormal DNA methylations in the progression of children with allergic asthma, and to reveal molecular mechanism underlying this disease. To achieve this goal, firstly, DNA methylation data for allergic asthma children were prepared from Gene Expression Omnibus (GEO) database. Secondly, differentially methylated genes were identified between allergic asthma and healthy subjects utilizing *t*-test. Subsequently, Gene Ontology (GO) and pathway enrichment analyses were conducted for differentially methylated genes using the Database for Annotation, Visualization and Integrated Discovery (DAVID) tool. Ultimately, the protein-protein interaction network (PPIN) for differentially methylated genes of allergic asthma children was constructed, following by hub genes identification. As a result, total 3 hub differentially hyper-methylated genes and 7 hub differentially hypo-methylated genes were obtained for allergic asthma. The findings might provide potential targets for allergic asthma diagnosis and treatment, and shed new lights on revealing molecular mechanism underlying this disease.

INTRODUCTION

Asthma is a common chronic inflammatory disorder of the lungs characterized by variable and recurring symptoms, reversible airflow obstruction and bronchospasm (Mintegi et al. 2018). It is well established that a combination of environmental and genetic factors is considered to be responsible for asthma, despite the exact explanations for rapidly changes in prevalence are still unknown (Torgerson et al. 2011). Furthermore, the epigenome of asthma could be altered by specific environmental exposures which increase its risk and severity, and the changes of epigenomic induced by environments lead to persistent changes in gene expression of asthma patients (Yang et al. 2017). Meanwhile, asthma exacerbations are a leading cause of hospitalisation for children in developed countries (Bizzintino et al. 2011), since children are especially vulnerable to airborne pollution because of their narrower airways and longer time of their exposure to air pollutants leading to their more air per pound of body weight than adults (Gasana et al. 2012). For instance, *in utero* exposure to tobacco smoke was related to childhood asthma, and this exposure would alter the expression of gene via DNA methylation (Ferrante et al. 2014).

DNA methylation is a result of the covalent addition of a methyl group at the 52 position of the pyrimidine ring of cytosines within context of CpG dinucleotides (Jimenez-Useche et al. 2014), and has a variety of important functions, including control of gene expression, genomic imprinting, cellular differentiation, and X-chromosome inactivation (Jeltsch and Jurkowska 2014). What's more, DNA methylation can be stable and easily determined qualitatively or quantitatively. Hence it has been taken as the most potential diagnostic marker for early cancer detection (Zhao et al. 2014), related copy number variations (Jiang et al. 2010), SNP/mutation (Guo et al. 2014), and gene/microRNA expression (Zhu and Yao 2009). In the past decades, a few number of researchers paid attention to the abnormal DNA methylations in progression of asthma. Stefanowicz et al. (2012) highlighted the importance of understanding DNA methylation in the epithelium when studying the epithelial contribution to asthma (Stefanowicz

Address for correspondence:
Zhao-Yang Yin
Department of Pediatrics, Shangluo Central Hospital, No.148 on Beixin Street, Shangzhou District, Shangluo,
726000, Shaanxi Province, China
Phone: 86-0914-2313200
Fax: 86-0914-2313200
E-mail: zhaoyangyin2017@163.com

et al. 2012). Most important, Yang et al (2015) demonstrated that allergic asthma was associated with DNA methylation marks in specific genes, which indicated epigenetic changes might be critical for establishing the immune related to children with asthma (Yang et al. 2015). Although great efforts have been made, to the best of knowledge, the molecular mechanisms underlying asthma are still not clear.

Objective

In the present study, the aim of the research is to extract differentially methylated genes between allergic asthma patients and healthy subjects dependent on the DNA methylation data. Subsequently, a series of functional enrichment analyses were conducted on these differentially methylated genes. Ultimately, a protein-protein interaction network (PPIN) was constructed and then hub genes were identified. These results might reveal the potential roles of epigenetic mechanisms in children with allergic asthma.

METHODOLOGY

DNA Methylation Data Preparation and Preprocessing

In the present study, DNA methylation data with accessing number GSE40576 (Yang et al. 2015) for childhood with allergic asthma were collected from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) at the National Center for Biotechnology Information (NCBI). Herein, the GEO database is regarded as an international public repository for high-throughput microarray and next-generation sequence functional genomic data sets (Barrett et al. 2011). GSE40576 was comprised of 194 children including 97 allergic asthma cases and 97 healthy controls, and was deposited on the Illumina Human Methylation450 BeadChip (Illumina, California, USA). There were 65,535 CpG sits in the raw DNA methylation data.

Prior to the subsequent analyses, standard preprocessing methods, quality control and normalization, for the raw 65,535 CpG sits were conducted. First of all, probes that satisfied with any one of four conditions were removed: the distance from CpG to SNP < 2; minimum allelic frequency < 0.05; probes on X and Y chromosomes; and cross-hybridising probes. Afterwards, the filtrated methylation data were analyzed and normalized using the beta-mixture quantile normalization method implemented in Lumi package (http://bioconductor.org/packages/release/bioc/html/lumi.html) (Du et al. 2010). The beta-mixture quantile algorithm improves the robustness of the normalization procedure and reduces the technical variation and bias (Teschendorff et al. 2013). In consequence, total 55,102 CpG sits were reserved in the DNA methylation data for in-depth exploitation.

Differentially Methylated Genes Screen

During this step, differentially methylated CpG sites between allergic asthma and healthy controls were identified by *t*-test. Before it, a methylation beta value (percent methylation changes) was computed for each CpG site in asthma group and healthy group, respectively. To the best our knowledge, methylation at individual CpG site is described as a methylation beta value, which is usually used as a quantitative measure of methylation for each CpG site with range from 0 to 1 (Wu et al. 2018). A beta value = 0 referred that the CpG was no methylation, while 1 meant that the CpG was completely methylated. Of note, the beta value for a CpG in different samples across allergic asthma patients and healthy controls were different. Hence, the mean beta value was defined as the value for this CpG in a specific group, and its absolute difference between two groups was also counted, naming score for the CpG. Ultimately, CpG sites which met to the thresholds $P < 0.05$ and Score > 0.05 were considered to be differentially methylated. Besides, genes covered using differentially methylated CpG sites were denoted as differentially methylated genes for allergic asthma cases.

Gene Ontology (GO) Enrichment Analysis

As mentioned above, to further investigate significant biological functions for differentially methylated genes of allergic asthma, GO functional enrichment analysis were conducted utilizing the Database for Annotation, Visualization and Integrated Discovery (DAVID, https://david.ncifcrf.gov/). Specifically, GO analysis has been widely utilized as functional enrichment researches for large-scale genes (Ashburner et al. 2000), and produces a controlled, dynamic vocabulary even as knowledge that effects of gene and protein in cells are changing and accu-

mulating (Ashburner et al. 2000). Besides, the DAVID offers functional annotations for researchers to understand biological meaning behind large list of genes (Huang da et al. 2009). The expression analysis systematic explored (EASE) test implemented in DAVID was used to detect the significant categories, since the EASE indicates biological processes and molecular functions unique to each category (Wang and Simon 2011). Then, P values were corrected using false discovery rate (FDR) using Benjamini and Hochberg method (Benjamini et al. 2001). In addition, the threshold of minimum number of genes (count) for the corresponding term > 10 was regarded as statistically significant for a category. Functional GO terms with $P < 0.001$ and gene count > 10 were regarded as statistically significant for allergic asthma compared with healthy subjects.

Pathway Enrichment Analysis

The Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was selected to identify significant gene sets enriched by differentially methylated genes using the online tool DAVID. Here, KEGG is considered as a knowledge database for systematic analysis of gene functions, which links genomic information with higher order functional information (Kanehisa and Goto 2000). KEGG pathways with P value <0.01 were chosen to be significant between children with allergic asthma and healthy subjects based on the EASE test applied in the DAVID. These P values were also adjusted by FDR using Benjamini and Hochberg method (Benjamini et al. 2001). On the basis of the cutoff $P < 0.01$ and count > 10, significant pathways for allergic asthma were determined.

PPIN Construction and Topological Centrality Analysis

After identifying significant gene sets enriched by large scale of differentially methylated genes, the researchers focused on investigating the interactions between any two of them. Thus, a PPIN was constructed based on the Search Tool for the Retrieval of Interacting Genes/proteins (STRING, http://string.embl.de) database. In particular, the STRING database offers an important assessment and integration of PPIs, containing indirect and direct associations (Szklarczyk et al. 2015). Subsequently, to assess the significance of an individual gene in the PPIN, topological centrality analysis was carried out utilizing the degree index. Degree quantifies the local topology of each node in a network through summarizing the number of its adjacent nodes (Haythornthwaite 1996). The genes at the top of degree distribution (e" 99% quantile) in the statistically significant perturbed networks were regarded as hub genes. Furthermore, the molecular complexes for hub differentially methylated genes was extracted from the PPIN by the molecular complex detection (MCODE) algorithm, a theoretically cluster algorithm, which chooses densely connected regions in large complex networks (Bader and Hogue 2003). Finally, the cluster of PPIN was visualized using Cytoscape (http://www. cytoscape. org/). Here, Cytoscape, a free software package, can be used for modeling, and analyzing and visualizing the integration of bimolecular interaction networks with high-throughput expression data (Smoot et al. 2011).

RESULTS

Differentially Methylated Genes

In this paper, there were 65,535 CpG sits in GSE40576 at the time of download. After performing quality control and normalization, a total of 55,102 CpG sites were left in the methylation data of 194 samples for subsequent analyses. Afterwards, a score was computed for each CpG site dependent on the beta value (percentage methylation value), and a *t*-test was used to determine the CpG sites that differentially methylated across allergic asthma cases and healthy controls. A volcano plot clarifying the distribution of the 55,102 analyzed methylated CpG sites was drew for allergic asthma as shown in Figure 1. When setting the criteria as $P < 0.05$ and Score > 0.05, 8,117 CpG sites (representing 3,608 genes) were differentially methylated for allergic asthma patients compared with healthy controls, including 1,513 hyper-methylated CpG sites (covering 658 genes) and 6,604 hypo-methylated CpG sites (referring 2950 genes). What's more, to improve the statistical power of differentially methylated CpG sites and delete the amount of non-variable CpG sites, two conditions were used to refine them, one was that the beta values for the CpG sites in specific sample of any group must ranged from 0.2 to 0.8; and the other condition required that the score for a CpG site should be higher than 0.2. As a result, 1,747 differentially methylated CpG sites (covering 1,204 genes) were de-

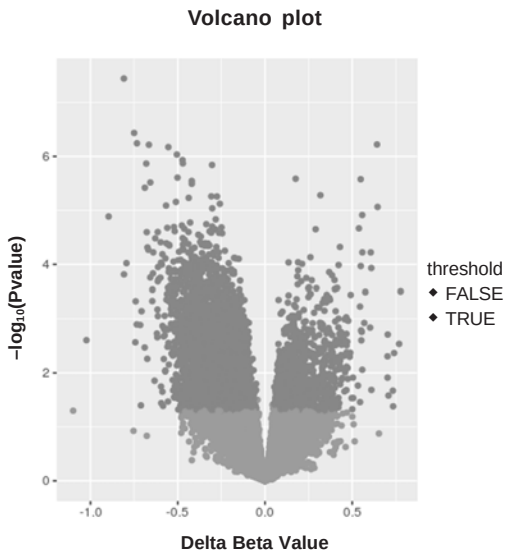

Fig. 1. Volcano plot exhibiting methylation data of children with allergic asthma and healthy subjects. X axis stood for the mean methylation differences between allergic asthma and healthy controls. Y axis represented the log transformed P values
Source: Author

tected. Particular, 489 CpG sites (containing 384 genes) were hyper-methylated, while 1,258 CpG sites (covering 820 genes) were hypo-methylated. Hereinafter, the genes covered by differentially methylated CpG sites were defined as differentially methylated genes. Hence, the researchers obtained total 1,204 differentially methylated genes (384 hyper-methylated and 820 hypo-methylated) between children with and without allergic asthma for further study.

Significant GO Terms for Differentially Methylated Genes

To better understand the biological functions of the differentially methylated genes, differentially hyper-methylated and hypo-methylated were annotated using GO annotation in the DAVID, respectively. Most important, GO terms with count > 10 and P < 0.001 were thought as statistically significant for allergic asthma children. Consequently, 8 and 17 significant GO terms were obtained for differentially hyper-methylated and hypo-methylated genes, respectively, as shown in Table 1. Of which, positive regulation of transcription from RNA polymerase II promoter (P =

Table 1: Significant GO terms with P < 0.01 and count > 10

Type	*Term*	*GO term*	*Count*	*P Value*
hyper-methylated	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	23	5.99E-08
	GO:0006915	apoptotic process	14	4.57E-06
	GO:0006366	transcription from RNA polymerase II promoter	13	6.38E-05
	GO:0045893	positive regulation of transcription, DNA-templated	13	6.61E-05
	GO:0010628	positive regulation of gene expression	18	4.31E-04
	GO:0042110	T cell activation	13	5.09E-04
	GO:0007165	signal transduction	14	5.73E-04
	GO:0006959	humoral immune response	13	6.65E-04
hypo-methylated	GO:0000122	negative regulation of transcription from RNA polymerase II promoter	20	8.33E-09
	GO:0035023	regulation of Rho protein signal transduction	40	1.32E-08
	GO:0006351	transcription, DNA-templated	84	6.39E-08
	GO:0007155	cell adhesion	26	4.74E-07
	GO:0006366	transcription from RNA polymerase II promoter	28	8.03E-07
	GO:0090002	establishment of protein localization to plasma membrane	30	9.45E-07
	GO:0043547	positive regulation of GTPase activity	22	1.01E-05
	GO:0030324	lung development	18	1.59E-05
	GO:0048538	thymus development	14	2.28E-05
	GO:0006325	chromatin organization	16	3.41E-05
	GO:0007275	multicellular organism development	28	4.47E-05
	GO:0035556	intracellular signal transduction	23	7.79E-05
	GO:0001558	regulation of cell growth	18	3.87E-04
	GO:0007155	cell adhesion	26	5.63E-04
	GO:0035556	intracellular signal transduction	23	7.84E-04
	GO:0008285	negative regulation of cell proliferation	22	8.21E-04
	GO:0043065	positive regulation of apoptotic process	16	9.79E-04

5.99E-08, count = 23), apoptotic process (P = 4.57E-06, count = 14) and transcription from RNA polymerase II promoter (P = 6.38E-05, count = 13) were the most three significant gene sets for 384 differentially hyper-methylated genes of allergic asthma children. Meanwhile, negative regulation of transcription from RNA polymerase II promoter (P = 8.33E-09, count = 20) was the most significant terms for 820 differentially hypo-methylated genes, and the next two were regulation of Rho protein signal transduction (P = 1.32E-08, count = 40) and transcription, DNA-templated (P = 6.39E-08, count = 84). Interestingly, several GO terms for two kinds of genes (hyper- and hypo-methylated) were corresponding, such as negative regulation of transcription from RNA polymerase II promoter and positive regulation of transcription from RNA polymerase II promoter.

Significant Pathways for Differentially Methylated Genes

As mentioned above, KEGG pathway enrichment analysis was carried out for differentially hyper-methylated and hypo-methylated genes, respectively. When setting the criteria as P < 0.01 and count > 10, a total of 9 and 15 significant pathways were explored for the two types of differentially methylated genes, respectively. In order to illustrate these pathways more directly and clearly, a graph for the first third of significant pathways dependent on their P values and count values was visualized in Figure 2. Differentially hyper-methylated genes were remarkably enriched in transmembrane receptor protein tyrosine kinase activity (P = 3.53E-04, count = 11), transmembrane receptor protein kinase activity (P = 6.29E-03, count = 14) and growth factor binding (P = 7.99E-03, count = 18). For differentially hypo-methylated genes, 5 most significant pathways were mapped to the figure, of which transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific binding (P = 6.54E-06, count = 58) was the most significant one in allergic asthma children compared with healthy subjects. Particularly, 6 of 8 significant gene sets were correlated to binding, which indicated the importance of binding in the progression of allergic asthma.

Hub Differentially Methylated Genes

For purpose of revealing interactions and connections between any two of differentially methylated genes, hypo-methylated genes and PPINs for differentially hyper-methylated genes were constructed, respectively. For 384 up-methylated genes, 304 nodes involved in 1,587 edges were mapped to the PPIN for hyper-methylated condition. However, 734 of 820 differentially hypo-methylated genes were deposited in the PPIN for them. By conducting the topological degree centrality analysis, total 3 hub hyper-methylated genes (LSM2, CD247 and CPSF1) and 7 hub hypo-methylated genes (RAC1, ADCY3, GNA12, GNAI2, TRIO, EIF4G1 and CDKN1A) were obtained for allergic asthma. Afterwards, clusters or sub-networks for hub methylated genes were captured from their PPINs through the MCODE algorithm. Unfortunately, the 3 hub differentially hyper-methylated genes couldn't form a cluster. At the same time, 5 of the 7 hub differentially hypo-methylated genes constructed a cluster, including RAC1, ADCY3,

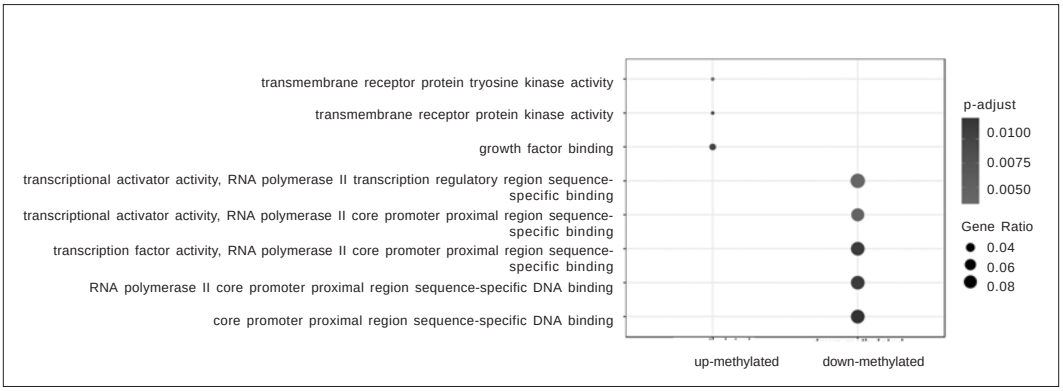

Fig. 2. The first third of significant pathways for differentially hyper (up) - and hypo (down) - methylated genes of allergic asthma children
Source: Author

GNA12, GNAI2 and TRIO, as described in Figure 3. This result might indicate that good correlations presented among the 5 hub differentially hypo-methylated genes.

DISCUSSION

The analysis of DNA methylation data has been widely applied to investigate the abnormally methylated genes associated with allergic asthma children, and enhance the feasibility and confidence of identifying targets for therapeutic strategies. Hence, in this study, the molecular mechanism of allergic asthma was uncovered by bioinformatics methods, containing identification of differentially methylation genes, GO functional and KEGG pathway enrichment analysis of differentially methylated genes, construction of PPIN based on differentially methylated genes and exploration of hub differentially methylated genes. Based on these results, the potential molecular mechanism underlying allergic asthma was inferred.

Particularly, results of differentially methylation analysis showed that a total of 1,204 differentially methylated genes between allergic asthma and healthy subjects were obtained, of which 384 were hyper-methylated and 820 were hypo-methylated. Furthermore, based on topological centrality analysis for PPINs constructed on hyper- and hypo-methylated genes, 3 and 7 hub differentially hyper- and hypo-methylated genes were extracted, respectively. Taking RAC1 as an example, RAC1 (ras-related C3 botulinum toxin

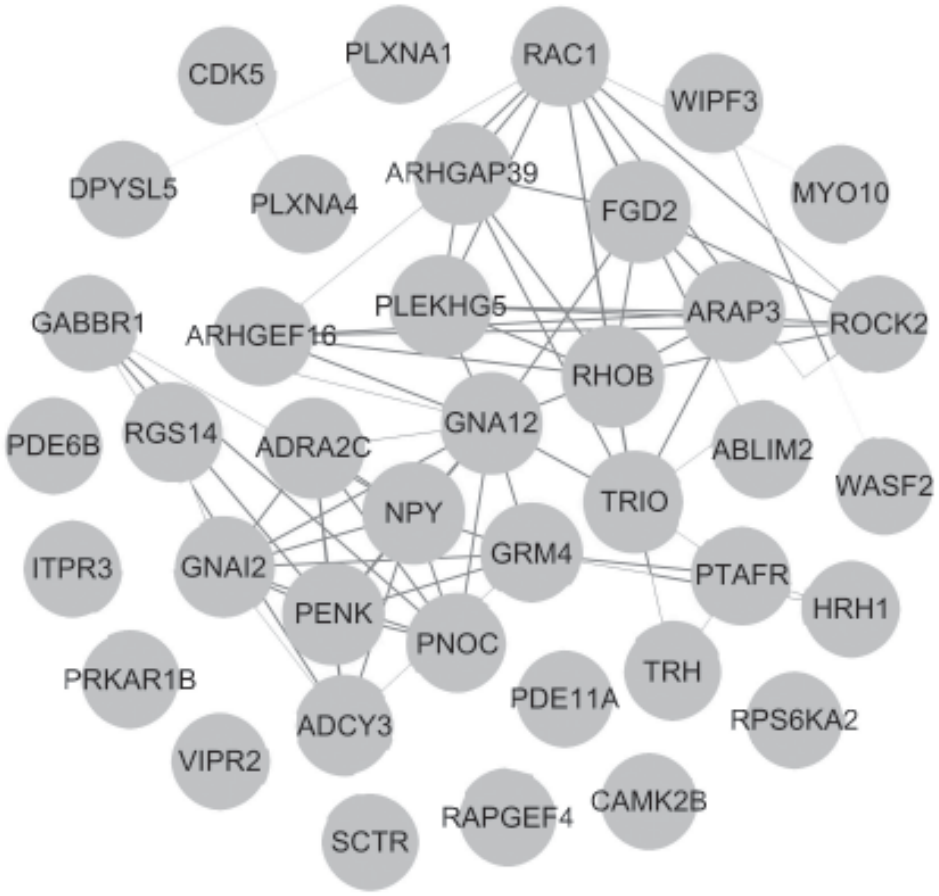

Fig. 3. A cluster extracted for the protein-protein interaction network of differentially hypo-methylated genes. Nodes stood for genes, and the edges represented the interactions between any two genes
Source: Author

substrate 1) is a small signaling G protein that is more specifically a GTPase, as well as a member of the Rac subfamily of the family Rho family of GTPases. Specifically, Rho family GTPases adjusts various cell functions mainly including proliferation, motility, apoptosis, redox signaling, and gene transcription (Riching and Keely 2015). It had been suggested that deletion of RAC1 possessed modest effects on T and B cell function and development, and RAC1 had distinct functions in hematopoietic cells where it was critical for proliferation and entry into cell cycle (Guo et al. 2008; Saci et al. 2011). If RAC1 combined with activators of transcription and signal transducers as functional binding partners, it would play critical roles in airway smooth muscle cells proliferation downstream of growth factor stimulation in asthma (Simeone-Penney et al. 2008). Above all, targeting RAC1 can be a potential therapeutic method for people with severe asthma and airway remodeling.

By accessing the GO functional enrichment analysis for differentially methylated genes, positive regulation of transcription from RNA polymerase II promoter was the most significant terms for differentially hyper-methylated genes, whereas the negative regulation of transcription from RNA polymerase II promoter was the most significant one for differentially hypo-methylated genes. Interestingly, in this study, the results demonstrated that the two GO terms were corresponding, just like the hyper- and hypo-methylated genes. Meanwhile, RNA polymerase II transcription regulatory region sequence-specific binding was significant pathways for differentially hypo-methylated genes, and the other significant pathways no matter for hyper- or hypomethylated genes were associated with binding. RNA polymerase II, the most studied type of RNA polymerase, catalyzes the transcription of DNA to synthesize precursors of mRNA and most snRNA and microRNA, which requires multiple transcription factors to bind to upstream gene promoters and begin transcription (Sainsbury et al. 2015; Wang et al. 2018).

CONCLUSION

In summary, the researchers have identified hub differentially hyper- and hypo-methylated genes and their enriched significant GO terms and pathways. The findings might provide potential targets for allergic asthma diagnosis and treatment, and shed new lights on revealing molecular mechanism underlying this disease.

RECOMMENDATIONS

Results from the current study will provide the groundwork for the understanding of allergic asthma pathogenesis and provide potential targets for allergic asthma diagnosis and treatment.

REFERENCES

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G 2000. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*, 25: 25-29.

Bader GD, Hogue CW 2003. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*, 4: 2.

Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Muertter RN, Holko M, Ayanbule O, Yefanov A, Soboleva A 2011. NCBI GEO: Archive for functional genomics data sets-10 years on. *Nucleic Acids Res*, 39: D1005-1010.

Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I 2001. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res*, 125: 279-284.

Bizzintino J, Lee WM, Laing IA, Vang F, Pappas T, Zhang G, Martin AC, Khoo SK, Cox DW, Geelhoed GC, McMinn PC, Goldblatt J, Gern JE, Le Souef PN 2011. Association between human rhinovirus C and severity of acute asthma in children. *Eur Respir J*, 37: 1037-1042.

Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, Lin SM 2010. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, 11: 587.

Ferrante G, Antona R, Malizia V, Montalbano L, Corsello G, La Grutta S 2014. Smoke exposure as a risk factor for asthma in childhood: A review of current evidence. *Allergy Asthma Proc*, 35: 454-461.

Gasana J, Dillikar D, Mendy A, Forno E, Ramos Vieira E 2012. Motor vehicle air pollution and asthma in children: A meta-analysis. *Environ Res*, 117: 36-45.

Guo F, Cancelas JA, Hildeman D, Williams DA, Zheng Y 2008. Rac GTPase isoforms Rac1 and Rac2 play a redundant and crucial role in T-cell development. *Blood*, 112: 1767-1775.

Guo S, Wang YL, Li Y, Jin L, Xiong M, Ji QH, Wang J 2014. Significant SNPs have limited prediction ability for thyroid cancer. *Cancer Med*, 3: 731-735.

Haythornthwaite C 1996. Social network analysis: An approach and technique for the study of information exchange &. *Library and Information Science Research*, 18: 323-342.

Huang da W, Sherman BT, Lempicki RA 2009. Systematic and integrative analysis of large gene lists using

DAVID bioinformatics resources. *Nat Protoc*, 4: 44-57.
Jeltsch A, Jurkowska RZ 2014. New concepts in DNA methylation. *Trends Biochem Sci*, 39: 310-318.
Jiang F, Todd NW, Li R, Zhang H, Fang H, Stass SA 2010. A panel of sputum-based genomic marker for early detection of lung cancer. *Cancer Prev Res (Phila)*, 3: 1571-1578.
Jimenez-Useche I, Shim D, Yu J, Yuan C 2014. Unmethylated and methylated CpG dinucleotides distinctively regulate the physical properties of DNA. *Biopolymers*, 101: 517-524.
Kanehisa M, Goto S 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*, 28: 27-30.
Mintegi S, Paniagua N, Pijoan JI, Gorostizaga Z, Gonzalez M, Benito J 2018. Dexamethasone for pediatric asthma exacerbations: A quality improvement intervention. *Am J Med Qual*: 1062860618765676.
Riching KM, Keely PJ 2015. Rho family GTPases: Making it to the third dimension. *Int J Biochem Cell Biol*, 59: 111-115.
Saci A, Cantley LC, Carpenter CL 2011. Rac1 regulates the activity of mTORC1 and mTORC2 and controls cellular size. *Mol Cell*, 42: 50-61.
Sainsbury S, Bernecky C, Cramer P 2015. Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol*, 16: 129-143.
Simeone-Penney MC, Severgnini M, Rozo L, Takahashi S, Cochran BH, Simon AR 2008. PDGF-induced human airway smooth muscle cell proliferation requires STAT3 and the small GTPase Rac1. *Am J Physiol Lung Cell Mol Physiol*, 294: L698-704.
Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T 2011. Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics*, 27: 431-432.
Stefanowicz D, Hackett TL, Garmaroudi FS, Gunther OP, Neumann S, Sutanto EN, Ling KM, Kobor MS, Kicic A, Stick SM, Pare PD, Knight DA 2012. DNA methylation profiles of airway epithelial cells and PBMCs from healthy, atopic and asthmatic children. *PLoS One*, 7: e44213.
Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von Mering C 2015. STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*, 43: D447-452.
Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S 2013. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics*, 29: 189-196.
Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, Himes BE, Levin AM, Mathias RA, Hancock DB, Baurley JW, Eng C, Stern DA, Celedon JC, Rafaels N, Capurso D, Conti DV, Roth LA, Soto-Quiros M, Togias A, Li X, Myers RA, Romieu I, Van Den Berg DJ, Hu D, Hansel NN, Hernandez RD, Israel E, Salam MT, Galanter J, Avila PC, Avila L, Rodriquez-Santana JR, Chapela R, Rodriguez-Cintron W, Diette GB, Adkinson NF, Abel RA, Ross KD, Shi M, Faruque MU, Dunston GM, Watson HR, Mantese VJ, Ezurum SC, Liang L, Ruczinski I, Ford JG, Huntsman S, Chung KF, Vora H, Li X, Calhoun WJ, Castro M, Sienra-Monge JJ, del Rio-Navarro B, Deichmann KA, Heinzmann A, Wenzel SE, Busse WW, Gern JE, Lemanske RF, Jr., Beaty TH, Bleecker ER, Raby BA, Meyers DA, London SJ, Gilliland FD, Burchard EG, Martinez FD, Weiss ST, Williams LK, Barnes KC, Ober C, Nicolae DL 2011. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet*, 43: 887-892.
Wang W, Walmacq C, Chong J, Kashlev M, Wang D 2018. Structural basis of transcriptional stalling and bypass of abasic DNA lesion by RNA polymerase II. *Proc Natl Acad Sci U S A*, 115: E2538-e2545.
Wang X, Simon R 2011. Microarray-based cancer prediction using single genes. *BMC Bioinformatics*, 12: 391.
Wu P, Farrell WE, Haworth KE, Emes RD, Kitchen MO, Glossop JR, Hanna FW, Fryer AA 2018. Maternal genome-wide DNA methylation profiling in gestational diabetes shows distinctive disease-associated changes relative to matched healthy pregnancies. *Epigenetics*, 13 (2): 122-128.
Yang IV, Lozupone CA, Schwartz DA 2017. The environment, epigenome, and asthma. *J Allergy Clin Immunol*, 140: 14-23.
Yang IV, Pedersen BS, Liu A, O'Connor GT, Teach SJ, Kattan M, Misiak RT, Gruchalla R, Steinbach SF, Szefler SJ, Gill MA, Calatroni A, David G, Hennessy CE, Davidson EJ, Zhang W, Gergen P, Togias A, Busse WW, Schwartz DA 2015. DNA methylation and childhood asthma in the inner city. *J Allergy Clin Immunol*, 136: 69-80.
Zhao Y, Sun J, Zhang H, Guo S, Gu J, Wang W, Tang N, Zhou X, Yu J 2014. High-frequency aberrantly methylated targets in pancreatic adenocarcinoma identified via global DNA methylation analysis using methylCap-seq. *Clin Epigenetics*, 6: 18.
Zhu J, Yao X 2009. Use of DNA methylation for cancer detection: Promises and challenges. *Int J Biochem Cell Biol*, 41: 147-154.

Paper received for publication on March 2018
Paper accepted for publication on May 2018