Investigation of the gene co-expression network and hub genes associated with acute mountain sickness

Yue Chang¹,²†, Jiange He³†, Jiqiang Tang⁴, Kai Chen¹,², Zhenguo Wang¹, Qun Xia³* and Hai Li²,⁵*

Abstract

Background: Acute mountain sickness has become a heavily researched topic in recent years. However, the genetic mechanism and effects have not been elucidated. Our goal is to construct a gene co-expression network to identify the key modules and hub genes associated with high altitude hypoxia.

Results: The GSE46480 dataset of rapidly transported healthy adults with acute mountain sickness was selected and analyzed by weighted gene co-expression network analysis (WGCNA) to construct a co-expression network. The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of the data set were carried out using Database for Annotation Visualization and Integrated Discovery (DAVID), and the hub genes were selected. We found that the turquoise module was most significantly correlated with acute mountain sickness. The functional enrichment analysis showed that the turquoise module was related to the apoptotic process, protein transport, and translation processes. The metabolic pathway analysis identified hsa03010:ribosome and hsa04144: endocytosis as the most important pathways in the turquoise module. Ten top 10 hub genes (MRPL3, PSMC6, AIMP1, HAT1, DPY30, ATP5L, COX7B, UQCRB, DPM1, and COMMD6) for acute mountain sickness were identified.

Conclusion: One module and 10 hub genes were identified, which were related to acute mountain sickness. The reference provided by this module may help to elucidate the mechanism of acute mountain sickness. In addition, the hub genes may be used in the future as a biomarker and therapeutic target for accurate diagnosis and treatment.

Keywords: High altitude, Hypoxia, Gene, Acute mountain sickness, Co-expression network, Pathway

Background

The plateau environment is a special ecological environment with low oxygen and low pressure [1]. Approximately, 40 million people venture into high-altitude regions for leisure or work each year [2]. Even with adequate acclimatization, the low oxygen concentration is an inevitable response that multiple organs of the body are in a state of hypoxia at a new altitude in the first 1–2 days [3]. High altitude illness encompasses a multiple spectrum of hypoxemia from acute mountain sickness to high altitude pulmonary edema (HAPE) or high altitude brain edema (HACE) [4–9]. Acute mountain sickness is characterized by a decrease in oxygen saturation, dizziness, lightheadedness, fatigue, sleep disturbance, gastrointestinal symptoms [4]. Interestingly, people who have lived on the plateau for a long time can adapt to such a low-oxygen environment. These people are thought to be more adapted to such an oxygen-deficient environment after years of genetic selection [10, 11].
adaptation e has been speculated to be related to multiple genes.

In late years, gene expression profile, widely applied in the study of pathogenic genes of a diversity of diseases, is a new popular and useful instrument [12, 13]. Nevertheless, there is still a lack of in-depth analysis of gene expression data related to adaptation to high altitude hypoxia environment, which limits the exploration of key genes for human adaptions to the plateau environment. Weighted gene co-expression network analysis (WGCNA) is an increasingly popular method, which can quickly extract gene co-expression modules related to sample features from complex data for subsequent analysis [14]. Simply put, it clusters genes with expression correlation into a module by calculating the expression correlation between genes. Then the correlation between the selected module and the characteristics of the sample (including clinical features, disease progress, therapeutic effect, etc.) was analyzed. It can be seen that WGCNA has built a bridge between sample characteristics and changes in gene expression.

In this written report, eight co-expression modules are constructed by WGCNA, and the interaction between these modules are analyzed in microarray data of acute mountain sickness. Most importantly, 10 hub genes (MRPL3, PSMC6, AIMP1, HAT1, DPY30, ATP5L, COX7B, UQCRB, DPM1, and COMMD6) may be related to different phenotypes of acute mountain sickness, which may become therapeutic gene targets for the study of altitude hypoxia adaptation in the future.

Results

Microarray data acquisition and gene expression analysis

The keyword "plateau hypoxia or acute mountain sickness" was used to explore GEO datasets in NCBI. The microarray dataset GSE46480 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=gse46480) was chosen for this study as it was most relevant to acute mountain sickness.

Subjects were chosen from healthy volunteers participating in the United States Antarctic Program, and were transported from sea-level to an altitude of 3200 m in less than 4 h. A total of 98 subjects (65 males and 33 females, aged from 26 to 50) with acute mountain sickness were identified. All subjects were of high-altitude non-adapted ancestry. Blood samples were taken 3 days after arrival to altitude. Peripheral blood mononuclear cell gene expressions were analyzed from this gene data set, and the sequencing platform was GPL570 ([HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0Array).

A total of 21,656 gene expression values were obtained from the original data. The top 4121 genes with the highest average expression value were selected for clustering analysis by WGCNA algorithm and iCluster software package (Fig. 1). When the threshold of the clustering height was determined to be 35, there were 5 outlier samples (GSM1131085, GSM1131000, GSM1131066, GSM1130997, and GSM1130998) in all samples. For the reliability of the module, these outlier samples were eliminated. As shown in Fig. 1, the remaining samples analyzed were divided into two groups (36 samples in Cluster 1 and 53 samples in Cluster 2) by Hierarchical cluster clustering. In addition, four samples from the outermost branch (GSM1130986, GSM1130996, GSM1130969, and GSM1130982) were very nearly to two groups also included in this analysis. Finally, only the remaining 89 samples were further analyzed.

Construction of plateau hypoxia gene co-expression module

The WGCNA algorithm was used to construct the co-expression module from the expression values of 4121 genes. First, the power value was screened out that when the value was 7, the scale independence was 0.85 (Fig. 2a). At the same time, it also had a higher average connectivity value (Fig. 2b). Then, we used 4121 genes with a high apparent value to construct a co-expression module (Fig. 2c). Different modules were sorted from high to low according to the number, and distinguished by different colors. As a result, a total of 8 co-expression modules were constructed. Nevertheless, 55 Genes (1.33%) that cannot be included in any module were placed in the gray module and removed in subsequent analysis. At the same time, the number of genes in each module was analyzed. The average number of genes in the 8 modules was 508 and the median was 248. There were 1901 genes in module 1 (turquoise), 854 genes in module 2 (blue), 315 genes in module 3 (brown), 309 genes in module 4 (green), 187 genes in module 5 (red), 176 genes in module 6 (black), 164 genes in module 7 (pink), and 160 genes in module 8 (magenta).

Analysis of interaction between co-expression modules

A network heat map was drawn to show the relationship between 8 modules (Fig. 3). The results showed that each module and the relative of gene expression in each module were all independent. Then, the eigengenes of each module feature were identified. According to the correlation of eigengenes, clustering was carried out to explore the co-expression similarity of all modules (Fig. 4a and b). The result showed that the eight modules were mainly divided into two clusters. The heat map of the module relationship was drawn showed the same results as above (Fig. 4c).

Functional enrichment analyzed by GO and KEGG

The functional enrichment analyzed by GO and KEGG was carried out on the 8 modules. There were significant differences in the results of functional enrichment
analysis among different modules. Moreover, the top 5 enriched GO terms were presented in Table 1. The genes in module 1 (turquoise) were mainly enriched in the process related to apoptotic process and protein transport, and the key terms were GO:0006915 (apoptotic process), GO:0015031 (protein transport), and GO:0006412 (translation). Most of the genes in module 2 (blue) were enriched in apoptosis and protein translation, and the most enriched terms were GO:0006915 (apoptotic process) and GO:0006412 (translation). Most of the genes in module 3 (brown) were enriched in the process related to innate immune response and cell adhesion, and the key terms were GO:0045087 (innate immune response), GO:0098609 (cell-cell adhesion) and GO:0006886 (intracellular protein transport). The genes in module 4 (green) were mainly enriched in signal transduction and inflammatory response, and the most abundant GO terms were GO:0007165 (signal transduction) and GO:0006954 (inflammatory response). The genes in module 6 (black) were mostly enriched in the process of mRNA splicing and intracellular signal transduction, and the most important GO terms were GO:0000398 (mRNA splicing, via spliceosome) and GO:0035556 (intracellular signal transduction). The genes in module 7 (pink) were mainly enriched in GO:0006915 (apoptotic process) and GO:0043066 (negative regulation of apoptotic process). The genes in module 8 (magenta) were mainly enriched in GO:0051607 (defense response to virus), GO:0060337 (type I interferon signaling pathway), and GO:0045087 (innate immune response). Besides, there was no obvious enrichment of terms in module 5 (red).

The signaling pathways analyzed by KEGG were summarized in Table 2. The result showed that a total of 1901 genes were recognized as the genes with high absolute correlations in module 1 and mainly enriched in hsa03010: Ribosome and hsa04144: Endocytosis. Module 2 was mainly enriched in the pathways hsa05016:

![Sample clustering to detect outliers](image-url)
Huntington’s disease and hsa03010: Ribosome. Module 3 was mainly enriched in metabolic pathways hsa04145: Phagosome, hsa05168: Herpes simplex infection, and hsa04144: Endocytosis. Module 4 was mainly enriched in hsa04380: Osteoclast differentiation. Module 6 was mainly enriched in hsa04010: MAPK signaling pathway and hsa03040: Spliceosome. Module 7 was enriched in pathways hsa05200: Pathways in cancer. Module 8 was enriched in hsa05162: Measles and hsa05164: Influenza A. There was no significantly enriched term in module 5.

Identification of hub genes
We visualized the turquoise module as networks in Cytoscape and screened out the top 10 genes ranked by Degree method. Figure 5 shows the top 10 hub genes in the turquoise module. They were MRPL3, PSMC6, AIMP1, HAT1, DPY30, ATP5L, COX7B, UQCRB, DPM1, and COMMD6 in the turquoise module.

Discussion
In this study, we found 8 gene modules related to acute mountain sickness by WGCNA method. Several core pathways and 10 hub genes were found by GO and KEGG. To the best of our knowledge, this is the first study focus on the hub genes related to acute mountain sickness. Our results may provide new help for the prevention and treatment of acute mountain sickness in future research.

Generally, the physical symptoms such as fatigue or headache may develop ascending rapidly to altitudes > 2500 m [15, 16]. Moreover, the physical symptoms and impaired alveolar oxygen diffusion tend to be more severe over 4500 m [17]. A real-world study from Harrison MF et al. reported that subjects with acute mountain sickness had lower oxygen saturation [9]. Indeed, people exposed to high altitudes may have acute hypoxia. Interestingly, not all new entrants to the high altitude develop AMS, which has an incidence of 30 to 69% based on the
altitude [18–20]. Additionally, it deemed to individual and gene differences are genetically related [21].

Instead of focusing only on differentially expressed genes, WGCNA can recognize the set of genes of interest using the information of specific or all genes and makes a significant association analysis with phenotypes [22]. By using 98 samples, 8 co-expression modules with relative independence were constructed. The interaction between modules is co-expressed by the thermographic reaction. Module 1 (turquoise) is considered to be the core module, and we further analyzed it.

Functional enrichment by GO suggested that GO: 0006915 (apoptotic process), GO: 0015031 (protein transport) and GO:0006412 (translation) in module 1 (turquoise) were valuable data and important components in cellular metabolism in high altitude circumstances. Noteworthy, mitochondrial dysfunction and endothelial barrier dysfunction play a key role in hypoxic diseases, especially cute mountain sickness, HAPE and other plateau hypoxia-related diseases. By differential proteomic and immunoassay techniques, 30% drops of glycolytic enzymes and of muscular creatine kinase abundance were found in subjects up to 8848 m. The impaired tricarboxylic acid (TCA) cycle and mitochondrial dysfunction were also found [23]. By reviewing the literature, we found that the apoptotic process [24], protein transport and translation are impaired in hypoxic conditions [25, 26]. Recently, Tsai SH et al. found that
hypoxia leads to hypoxia-inducible factor-1α (HIF-1α) expression and endothelial barrier dysfunction in healthy adult volunteers ascend to 3100 m height [27]. Yang YD et al. suggested that cell apoptosis is related to mitochondria-associated membranes in hypoxia-induced endothelial injury [28]. This evidence verifies the clinical value of functional enrichment by GO in Module 1.

The signaling pathways identified as hsa03010:Ribosome and hsa04144:Endocytosis. This conclusion is supported by a recent plasma proteomic study in the plateau of volunteers [29]. After 9 h of hypoxia, proteins related to the TCA cycle and ribosomes were significantly reduced in acute mountain sickness resistant individuals [30]. Furthermore, autophagy increase is related to endocytosis after the intestinal failure in Wistar rats acutely exposed to hypobaric hypoxia in plateau stress. Therefore, it is not queer to conclude that the mTOR signaling pathway participates in hypoxic memory injury in plateau and may be a potentially important therapeutic target for low-pressure and low-oxygen [31].

Moreover, 10 genes were screened out as hub genes. Many of them have been well documented to be associated with hypoxia such as AIMP1 [32], HAT 1 [33], and DPY30 [34] as well. Hub genes are highly connected genes by the WGCNA and the by Cytoscape software considered as functionally significant. So the hub genes can be different from differential genes. These novel hubs gene may use as a biomarker or therapeutic target for accurate diagnosis and treatment in acute mountain sickness in the future.

**Conclusion**

All in all, our fruitful work gives an explanation of the gene co-expression network and hub genes in acute
mountain sickness by functional enrichment and metabolic pathway analysis. The turquoise module was most significantly correlated with acute mountain sickness. And the hsa03010: Ribosome and hsa04144: Endocytosis were screened as the most important pathways while 10 hub genes were identified. The hub genes may be used in the future as a biomarker and therapeutic target for
accurate diagnosis and treatment. Of course, more research with acute mountain sickness is still needed in the future, especially in terms of the function of different genes.

Materials and methods
Gene expression analysis of acute mountain sickness
The microarray data were downloaded from the genome expression summary dataset on the https://www.ncbi.nlm.nih.gov/geo/ website with the keyword “altitude hypoxia or acute mountain sickness”, meanwhile “Homo sapiens”, “Expression profiling by array” were selected. The signal information of the probe provided in the dataset was downloaded. Then, matched the probe with the corresponding gene used the annotation information provided by the record. The number of genes was counted under different gene expression thresholds. WGCNA was used to evaluate the gene expression value and selected the appropriate threshold. In addition, the hclust toolkit in R language was used to cluster the samples under the appropriate threshold [35].

Construction of gene co-expression module at in acute mountain sickness
The weighting value in the construction of the module was screened out by the WGCNA algorithm. Gradient tests were used to analyze the scale independence and average connectivity of the module at different power values (from 1 to 20). The soft-thresholding power was selected when the degree of scale independence was set to 0.85.

Then at the soft-thresholding power value selected above, the co-expression module was built using the WGCNA algorithm. In addition, the corresponding gene information in each module was extracted. In order to obtain more accurate results, the minimum number of genes in the module was set to 40.

Interaction analysis of hypoxia gene co-expression module in acute mountain sickness
The relationship between the co-expression of acute mountain sickness was analyzed by the WGCNA algorithm. WGCNA was carried out by R language (http://R-project.org/), and the intensity of these interactions was reflected by heat maps.
Functional enrichment analysis of genes in hypoxia co-expression module of acute mountain sickness
The co-expression modules were sorted according to the number of genes in them. After that, the functional enrichment analysis of these genes was carried out. The genetic information in the corresponding module was used in Kyoto Encyclopedia of Genes and Genomes (KEGG) and gene ontology (GO) path analysis. The related annotation, visualization, and integration discovery were carried out if the corrected $p < 0.05$. If the number of records exceeds five, the top five records were selected for further analysis.

Identification of hub genes
Hub genes were visualized in networks by Cytoscape (version 3.7.0). Additionally, the top 10 hub genes in networks were identified by the degree and considered as functionally significant.

Abbreviations
DAVID: Database for Annotation Visualization and Integrated Discovery; GO: Gene Ontology; HACE: High altitude brain edema; HAPIE: High altitude pulmonary edema; KEGG: Kyoto Encyclopedia of Genes and Genomes; TCA: Tricarboxylic acid; WGCNA: Weighted gene co-expression network analysis

Acknowledgments
We would like to express our appreciation to the contributors of profiles of GSE46480 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=gse46480) in GEO datasets.

Authors’ contributions
YC, JGH, and Treatment, Tianjin 300162, China. 3Institute of Special War Trauma Technology, Characteristic Medical Center of People’s Armed Police Force, Tianjin 300162, China. 4Department of Orthopaedics, Characteristic Medical Center of People’s Armed Police Force, 220 Chenglin Road, Hedong District, Tianjin 300162, China. 5Department of Hepatopancreatobiliary and Splenic Medicine, Characteristic Medical Center of People’s Armed Police Force, Tianjin 300162, China. 6Department of Hepatopancreatobiliary and Splenic Medicine, Characteristic Medical Center of People’s Armed Police Force, Tianjin 300162, China. 7Division of Gastroenterology and Hepatology, Tianjin Xiqing Hospital, No.403 Xiqing Road, Xiqing District, Tianjin 300380, China.

Received: 26 November 2019 Accepted: 24 March 2020
Published online: 16 April 2020

References
1. Pan S, Zhang T, Rong Z, Hu L, Gu Z, Wu Q, et al. Population transcriptomes reveal synergistic responses of DNA polymorphism and RNA expression to extreme environments on the Qinghai-Tibetan plateau in a predatory bird. Mol Ecol. 2017;26:2993–3010.
2. Murray AJ. Energy metabolism and the high-altitude environment. Exp Physiol. 2016;101:23–7.
3. Bárttsch P, Swenson ER. Acute high-altitude illnesses. N Engl J Med. 2013;369:1666–7.
4. Davranche K, Casini L, Arnal PJ, Rupp T, Perrey S, Verges S. Cognitive functions and cerebral oxygenation changes during acute and prolonged hypoxic exposure. Physiol Behav. 2016;164:189–97.
5. Davis C, Hackett P. Advances in the prevention and treatment of high altitude illness. Emerg Med Clin North Am. 2017;35:241–60.
6. Du H, Zhao J, Su Z, Liu Y, Yang Y. Sequencing the exons of human glucocorticoid receptor ( NR3C1 ) gene in Han Chinese with high-altitude pulmonary edema. J Physiol Anthropol. 2018;37:7.
7. Gong G, Yin L, Yuan L, Sui D, Sun Y, Fu H, et al. Ganglioside GM1 protects against high altitude cerebral edema in rats by suppressing the oxidative stress and inflammatory response via the PI3K/AKT-NF-$\kappa$B pathway. Mol Immunol. 2018;95:91–8.
8. Li X, Jin T, Zhang M, Yang H, Huang X, Zhou X, et al. Genome-wide association study of high-altitude pulmonary edema in a Han Chinese population. Oncotarget. 2017;8:15168–80.
9. Harrison MF, Anderson PJ, Johnson JB, Richert M, Miller AD, Johnson BD. Acute mountain sickness symptom severity at the south pole: the influence of self-selected prophylaxis with acetazolamide. PLoS One. 2016;11: e0148206.
10. Loma G. Moore. Human genetic adaptation to high altitudes: current status and future prospects. Quat Int. 2017;461–13.
11. Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZXP, Pool JE, et al. Sequencing of 50 human exomes reveals adaptation to high altitude. Science. 2010;329:75–8.
12. Lev A, Simon AJ, Amariullo N, Rechavi G, Somech R. Thymic functions and gene expression profile distinct double-negative cells from single positive cells in the autoimmune lymphoproliferative syndrome. Autoimmun Rev. 2012;11:723–30.
13. Jha PK, Sahu A, Prabhabkar A, Tyagi T, Chatterjee T, Arvind P, et al. Genome-wide expression analysis suggests hypoxia-triggered hyper-coagulation leading to venous thrombosis at high altitude. Thromb Haemost. 2018;118:1279–95.
14. Prom-On S, Chanthaphan A, Chan JH, Meechai A. Enhancing biological relevance of a weighted gene co-expression network for functional module identification. J Bioinforma Comput Biol. 2011;9:111–29.
15. Bárttsch P, Swenson ER. Clinical practice: acute high-altitude illnesses. N Engl J Med. 2013;368:2294–302.
16. Boas CJ, Bass M, O’Hara JP, Vincent E, Mollor A, Severy L, et al. The relationship between anxiety and acute mountain sickness. PLoS One. 2018;13:e0197147.
17. Sagawa S, Torii R, Nagaya K, Wada F, Endo Y, Shiraki K. Carotid baroreflex sensitivity and its relationship to arterial stiffness at high altitude. J Hypertension. 2012;30:757–64.
18. Sagawa S, Torii R, Nagaya K, Wada F, Endo Y, Shiraki K. Carotid baroreflex sensitivity and its relationship to arterial stiffness at high altitude. J Hypertension. 2012;30:757–64.
19. Montgomery AB, Mills J, Luce JM. Incidence of acute mountain sickness at intermediate altitude. JAMA. 1989;261:732–4.
20. Onopa J, Haley A, Yewo ME. Survey of acute mountain sickness on Mauna Kea. High Alt Med Biol. 2007;8:200–5.
21. Hennis PJ, O’Doherty AF, Levett DZH, Grocott MPW, Montgomery HM. Genetic factors associated with exercise performance in atmospheric hypoxia. Sports Med. 2015;45:745–61.
22. Subramanian J, Simon R. Gene expression-based prognostic signatures in lung cancer: ready for clinical use? J Natl Cancer Inst Monogr. 2010;102:464–74.

23. Levetz DZH, Viganò A, Capitanio D, Vasso M, De Palma S, Moriggi M, et al. Changes in muscle proteomics in the course of the Caudwell research expedition to Mt. Everest Proteomics. 2015;15:160–71.

24. Gyongyosi A, Terraneo L, Bianciardi P, Tosaki A, Lekli I, Samaja M. The impact of moderate chronic hypoxia and hyperoxia on the level of apoptotic and autophagic proteins in myocardial tissue. Oxidative Med Cell Longev. 2018;2018:786742.

25. Chang Y, Zhang W, Chen K, Wang Z, Xia S, Li H. Metabonomics window into plateau hypoxia. J Int Med Res. 2019;47:5441–52.

26. Zhou Y, Huang X, Zhao T, Qiao M, Zhao X, Zhao M, et al. Hypoxia augments LPS-induced inflammation and triggers high altitude cerebral edema in mice. Brain Behav Immun. 2017;64:266–75.

27. Tsai SH, Huang PH, Tsai HY, Hsu YJ, Chen YW, Wang JC, et al. Roles of the hypoximir microRNA-424/322 in acute hypoxia and hypoxia-induced pulmonary vascular leakage. FASEB J. 2019;33(11):12565 R(201900564RR).

28. Yang YD, Li MM, Xu G, Zhang EL, Chen J, Sun B, et al. Targeting mitochondria-associated membranes as a potential therapy against endothelial injury induced by hypoxia. J Cell Biochem. 2019;120:18967–78.

29. Zhang F, Deng Z, Li W, Zheng X, Zhang J, Deng S, et al. Activation of autophagy in rats with plateau stress-induced intestinal failure. Int J Clin Exp Pathol. 2015;8:1816–21.

30. Lu H, Wang R, Lu W, Xie H, Wang C, Hao Y, et al. Plasma proteomic study of acute mountain sickness susceptible and resistant individuals. Sci Rep. 2018;8:1265.

31. Li M, Zhu Y, Li J, Chen L, Tao W, Li X, et al. Effect and mechanism of verbascoside on hypoxic memory injury in plateau. Phytother Res. 2019;33:2692–701.

32. Bronkhorst IH, Jehs TM, Dijkgraaf EM, Luyten GP, Van der Velden PA, Van der Burg SH, et al. Effect of hypoxic stress on migration and characteristics of monocytes in uveal melanoma. JAMA Ophthalmol. 2014;132:614–21.

33. Vrtacnik P, Zupan J, Mlakar V, Kranjc T, Marc J, Kem B, et al. Epigenetic enzymes influenced by oxidative stress and hypoxia mimetic in osteoblasts are differentially expressed in patients with osteoporosis and osteoarthritis. Sci Rep. 2018;8:16215.

34. Liu C, Zhang Y, Hou Y, Shen L, Li Y, Guo W, et al. PAQR3 modulates H3K4 trimethylation by spatial modulation of the regulatory subunits of COMPASS-like complexes in mammalian cells. Biochem J. 2015;467:415–24.

35. Chan BKC. Data analysis using R programming. Adv Exp Med Biol. 2018;1082:47–122.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.