

RESEARCH ARTICLE

Open Access



# Characteristics of transcriptome and chromatin accessibility in the peripheral blood after acute hypoxia exposure

Kuo Zeng<sup>1†</sup>, Pei-ru Yuan<sup>2†</sup>, Jin-feng Xuan<sup>3,4,5†</sup>, Lai-xi Zhao<sup>3,4,5</sup>, Xiao-na Li<sup>3,4,5\*</sup>, Jun Yao<sup>3,4,5\*</sup>  and Dong Zhao<sup>1\*</sup>

## Abstract

**Background** Human responses and acclimation to the environmental stresses of high altitude and low oxygen are multifaceted and regulated by multiple genes. However, the mechanism of how the body adjusts in a low-oxygen environment is not yet clear.

**Results** Hence, we performed RNA sequencing (RNA-seq) and ATAC sequencing (ATAC-seq) to observe the changes of transcriptome and chromatin accessibility in the peripheral blood of eight individuals at 1 h post adaptation in a simulated plateau environment with 3500 m and 4500 m altitude, respectively. Differential expression analysis and the Boruta algorithm identified differentially expressed genes (DEGs) and differentially accessible regions (DARs) associated with hypoxia adaptation. Specifically, RNA-seq identified 93 and 7 DEGs after 1 h post adaptation with 3500 m altitude and 45 and 8 DEGs after 1 h adaptation with 4500 m. Additionally, ATAC-seq screened 12 and 4 DARs in 3500 m altitude adaption and 15 and 5 DARs in 4500 m altitude adaption. Moreover, the combined analysis of RNA-seq and ATAC-seq revealed that 10 hub genes were independently identified from the protein–protein interaction (PPI) network for each altitude. Gene enrichment analysis displayed that most hub genes were related with hypoxia pathways.

**Conclusions** Our results can provide the reference for the early response of the organism to hypoxic adaptation.

**Keywords** Hypoxia, Transcriptome, Chromatin accessibility, High altitude

<sup>†</sup>Kuo Zeng, Pei-ru Yuan and Jin-feng Xuan are the co-first authors.

Handling editor: Vitor Sousa.

\*Correspondence:

Xiao-na Li

xnli@cmu.edu.cn

Jun Yao

yaojun198717@163.com

Dong Zhao

zhaod@cupl.edu.cn

<sup>1</sup> Key Laboratory of Evidence Science (China University of Political Science and Law), Ministry of Education, Beijing, China

<sup>2</sup> College of Law, Hunan University of Technology, Zhuzhou, China

<sup>3</sup> School of Forensic Medicine, China Medical University, Shenyang, China

<sup>4</sup> Key Laboratory of Forensic Bio-Evidence Sciences, Shenyang, Liaoning Province, China

<sup>5</sup> China Medical University Center of Forensic Investigation, Shenyang, China

## Background

There are a range of exposure factors at high altitude, such as low pressure, low oxygen, cold, and solar radiation, with the oxygen content of the air only about 60% of that at sea level at an altitude of 4000 m above sea level [1]. Among these factors, hypoxic environmental exposures cause high altitude specific diseases in humans [2]. High-altitude hypoxic acclimatization refers to the fact that the low-oxygen environment at high altitude can adversely affect all human systems, especially the respiratory system, and that along with prolonged stays at high altitude, the body will develop a series of adaptations to hypoxia [3]. Alterations in the respiratory system include changes in respiratory control, ventilation, gas exchange, lung mechanics and kinetics, and pulmonary vascular



physiology. However, travelers from lower altitudes to higher altitudes are still subject to hypoxic symptoms and plateau-related pulmonary hypertension, pulmonary edema, and other highland diseases [4]. Therefore, elucidating the biological processes by which the body adapts to hypoxia is of great scientific value.

At the molecular level, human responses and adaptations to the environmental stresses of high altitude and low oxygen are multifaceted and will be regulated by multiple genes in the nuclear genome. As transcription factors, hypoxia-inducible factor (HIF) and nuclear factor erythroid 2-related factor 2 (NRF2) drive genes mediating cellular responses to hypoxia and oxidative stress, respectively [5]. HIF consists of two subunits,  $\alpha$  and  $\beta$ , of which  $\alpha$  subunit has a very short half-life under normoxia, and only under hypoxia will  $\alpha$  dimerize with the building block  $\beta$  subunit to form HIF, which combines with the hypoxia-responsive progenitor on hypoxic response genes and mediates the transcriptional expression of hypoxic response genes to produce adaptive physiological changes [6]. Mitochondria are the major reactive oxygen species (ROS) generators in hypoxic cells. ROS inactivate succinate dehydrogenase, leading to succinate accumulation. Succinate activates molecules to increase the body's resistance to hypoxia [7]. Additionally, Pham et al. found that acute high-altitude exposure might sensitize the Toll-like receptor 4 (TLR4) signaling pathway to subsequent inflammatory stimuli [8]. Collectively, the response of the organism's systems to hypoxia is complex and multifaceted.

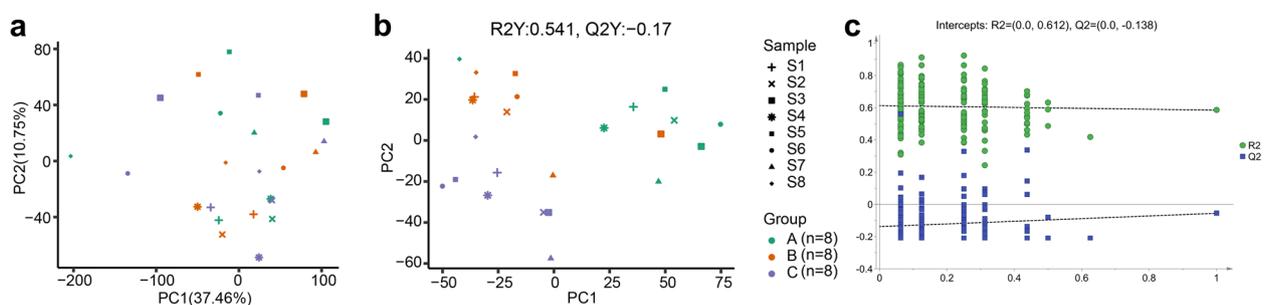
Hence, in order to elucidate the response in the human body after acute hypoxia exposure, we performed RNA sequencing (RNA-seq) and ATAC sequencing (ATAC-seq) to observe the changes of transcriptome and chromatin accessibility in the peripheral blood of individuals exposed to hypoxia. We hypothesized that the differential

expressed genes were associated with hypoxic acclimatization and might be targets for the treatment of diseases associated with hypoxia-exposure.

## Results

### Differences in transcriptional profiles of three groups were not obvious

To determine systematic differences in the transcriptome landscape before and after entering into the simulated high-altitude environment, RNA-seq was performed on blood samples derived from 8 volunteers, including after adapting to an altitude equivalent to sea level for 1 h (baseline, group A,  $n=8$ ), after adapting to an altitude 3500 m for 1 h (group B,  $n=8$ ), and after adapting to an altitude 4500 m for 1 h (group C,  $n=8$ ). At first, PCA was carried out based on the gene expression landscape. As shown in the score plot, samples from three groups were indistinguishable, which was further confirmed by PERMANOVA analysis (sample:  $R^2=0.448$ ,  $P=0.009$ ; group:  $R^2=0.057$ ,  $P=0.611$ ) (Fig. 1a). To separate the groups, supervised PLS-DA was carried out by fitting these samples. The PLS-DA model organized the replicates into three distinct clusters corresponding to the A, B, and C groups, with few outlier samples (Fig. 1b). Although the model fitted relatively well to the data set (R<sup>2</sup>Y value 0.541), the predictive performance of the model was poor (Q<sup>2</sup>Y value  $-0.17$ ). Besides, the result of permutation test indicated that the model was overfitting (Fig. 1c). These findings illustrated that the global transcription profiles of the three groups might be not significantly different. Therefore, the difference between baseline (group A) and after entering the simulated high-altitude environment (group B or C) was further explored. Similar to the previous results, although clear between-group differences were observed in the score plots and the models possessed a reasonable fit, the models were overfitting with



**Fig. 1** Plots of the multivariate statistical comparisons among groups based on transcriptional profiles. **a** The score plot of the principal component analysis (PCA). Each point represents an individual sample. **b** The PLS-DA score scatter plot. **c** Permutation test for the PLS-DA model. Abscissa represents the permutation retention of permutation test, ordinate indicates the value of R<sup>2</sup> (green dot) and Q<sup>2</sup> (blue square) permutation test, and the two dotted lines represent the regression lines of R<sup>2</sup> and Q<sup>2</sup> respectively. R<sup>2</sup> indicates the proportion of variance in the data explained by the model, and Q<sup>2</sup> indicates the predictive accuracy of the model. Group A, after adapting to an altitude equivalent to sea level for 1 h. Group B, after adapting to an altitude equivalent to 3500 m for 1 h. Group C, after adapting to an altitude equivalent to 4500 m for 1 h

low predictive performance (Additional file 1: Fig. S1). Additionally, different groups were indistinguishable in the RF models (Fig. 2).

Nevertheless, pairwise differential expression analysis was further performed with the three groups (i.e., comparison between baseline [group A] and after adapting to an altitude equivalent to 3500 m for 1 h [group B]; comparison between baseline [group A] and after adapting to an altitude equivalent to 4500 m for 1 h [group C]) to identify differentially expressed genes (DEGs). A total of 93 differential genes were generated after 1 h post adaptation in a simulated plateau environment with 3500 m altitude, including 67 downregulated and 26 upregulated genes (Fig. 3a). Moreover, a total of 45 differential genes for group C was identified, of which three were expressed upregulated and 42 were expressed downregulated (Fig. 3b). The expression levels of these genes were provided as a heatmap (Additional file 2: Fig. S2a-b). Then, the Boruta algorithm, based on the RF machine learning algorithm, further confirmed seven (comparison between group A and B, Fig. 3c) and eight (comparison between group A and C, Fig. 3d) target genes, respectively, most of which were downregulated genes (Additional file 2: Fig. S2c-d). Moreover, we observed that inter-individual variability of these important genes was low. Taken together, the above results revealed that there were few differences in blood transcriptional profiles before and after entering into the simulated high-altitude environment.

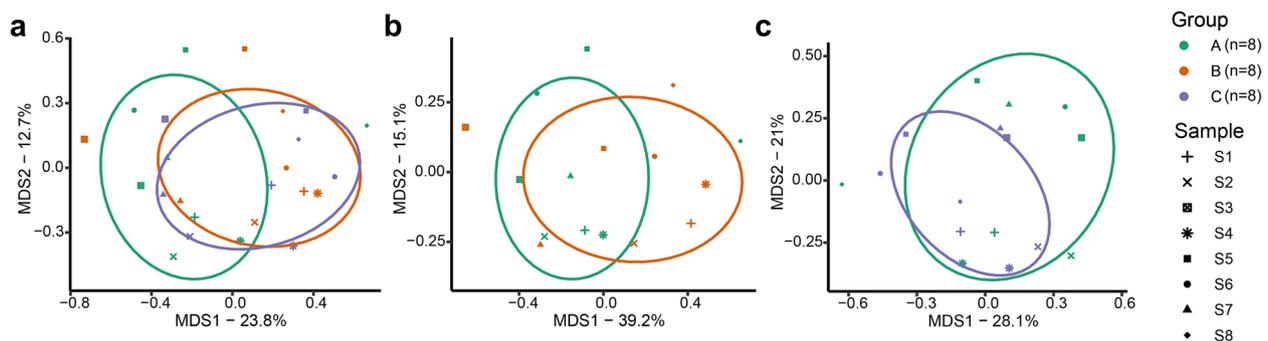
#### ATAC-seq analysis showed less noticeable differences between three groups

An ATAC-seq analysis was performed to explore changes in accessible regions in the genome after adaptation to the simulated high-altitude environment with the three groups. We focused on the peaks located on the promoter region. A PCA plot showed that the three groups could not be distinguished from each other on the basis

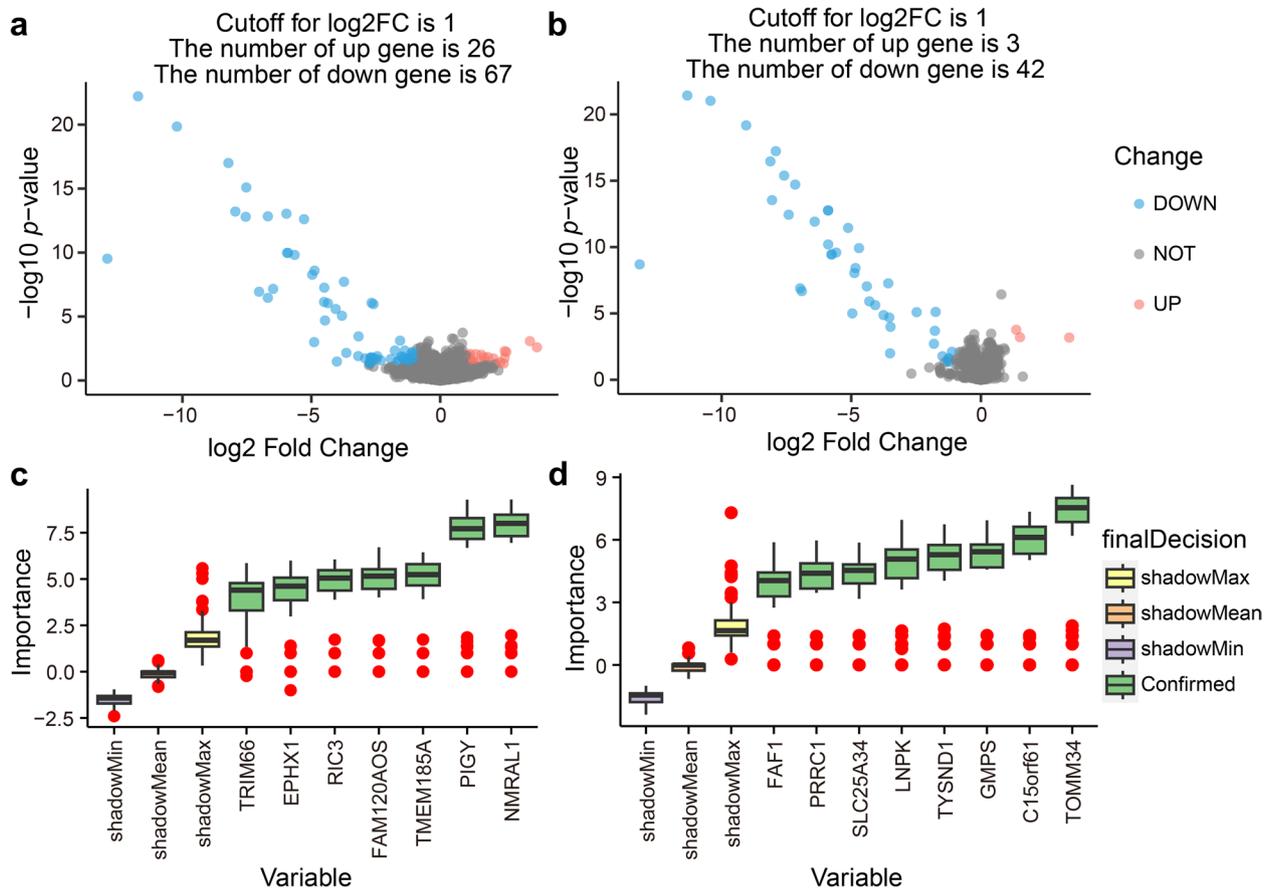
of PC1 (13.53%) and PC2 (6.56%), which was further confirmed by PERMANOVA analysis (sample:  $R^2=0.392$ ,  $P=0.001$ ; group:  $R^2=0.074$ ,  $P=0.482$ ) (Fig. 4a). Different groups were clearly discriminated by the PLS-DA plot; however, these models were overfitting (Fig. 4b, c, Additional file 3: Fig. S3). Similar results were observed in the RF models (Fig. 5). Subsequently, pairwise comparative analysis based on the ATAC-seq data obtained from the three groups were performed to evaluate the differentially accessible regions (DARs), which resulted in 12 (comparison between group A and B, Fig. 6a) and 15 (comparison between group A and C, Fig. 6b) DARs, respectively. We then plotted the expression levels of these DARs as a heatmap (Additional file 4: Fig. S4a-b). Additionally, using the Boruta algorithm, four (comparison between group A and B, Fig. 6c) and nine (comparison between group A and C, Fig. 6d) peaks were identified as important. The chromatin accessibility of the genes related to these important peaks exhibited large heterogeneity between groups (Additional file 4: Fig. S4c-d). These findings illustrated that the ATAC-Seq profiles in blood associated with the promoter region were less impacted by a short period of adaptation to the simulated high-altitude environment.

#### Combined analysis of RNA-seq and ATAC-seq screened gene interaction

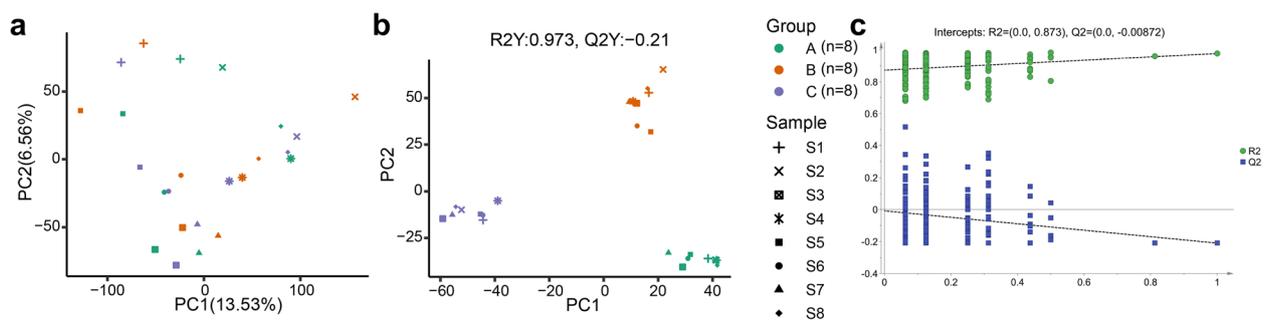
To establish the relationship between changes in gene expression and chromatin accessibility, we combined the results of RNA-seq with the results of ATAC-seq. A PPI analysis was performed on the 100 DEGs and 16 DARs identified in 3500 m altitude adaptation as well as 53 DEGs and 20 DARs identified in 4500 m altitude adaptation (Fig. 7a, Fig. 8a). The PPI results indicate interactions between DEGs identified by RNA-seq and DARs identified by ATAC-seq. Additionally, interactions between DEGs and DARs identified through



**Fig. 2** Random forest models for the exploration of group differences based on transcriptional profiles. **a** Multi-dimensional scaling (MDS) plot shows no differentiation among the three groups by random forest model. **b** MDS plot of random forest model to differentiate groups A and B. **c** MDS plot of random forest model to differentiate groups A and C



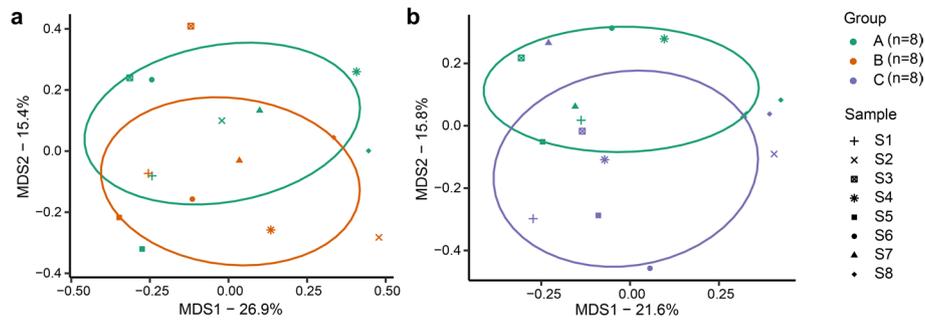
**Fig. 3** Differentially expressed genes were identified by differential expression analysis and the Boruta algorithm. **a** The volcano plot indicates the results of the pairwise comparisons of genes in the blood samples from groups A and B. **b** Volcano plot presents the differences between groups A and C. **c** Relative importance of important genes between groups A and B computed using the Boruta algorithm. The horizontal axis is the names of genes, and the vertical axis is the Z-value of each gene. The box plot shows the Z-value of each gene during model calculation. **d** Relative importance of important genes between groups A and C



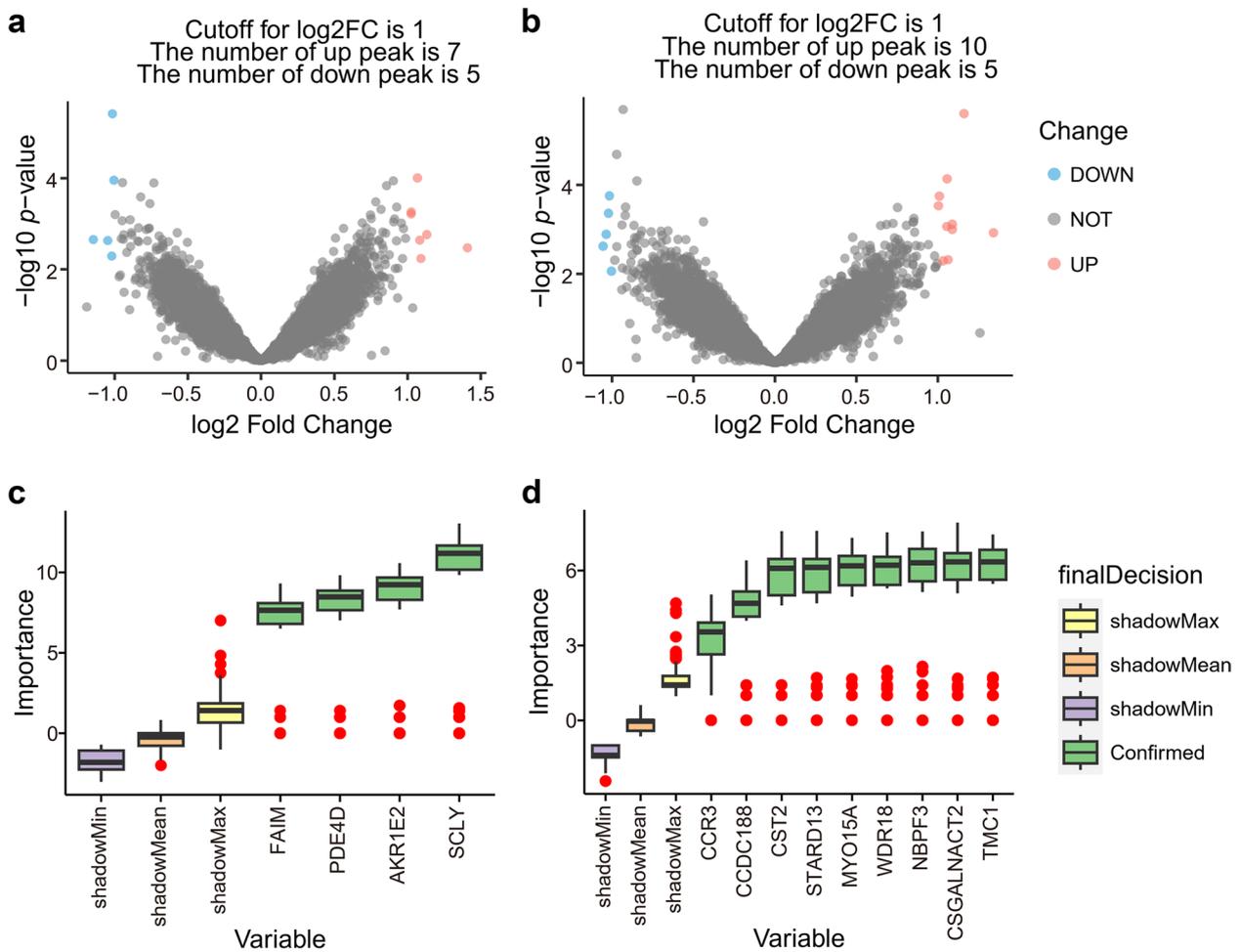
**Fig. 4** Plots of the multivariate statistical comparisons among groups based on the ATAC-seq data. **a** The score plot of PCA. **b** The PLS-DA score scatter plot. **c** Cross-validation plot with a permutation test repeated 200 times

differential expression analysis and the Boruta algorithm were observed. Then, CytoHubba was used to screen top 10 hub genes for each of the 3500 m and 4500 m altitude adaptations (Fig. 7b, Fig. 8b). Among

these, CREB-binding protein (CREBBP), tumor necrosis factor receptor-associated protein (TRAP1), tubulin (TUB), and DnaJ (Hsp40) homolog subfamily A member 3 (DNAJA3) were the shared hub genes in adaptation to altitudes in 3500 m and 4500 m.



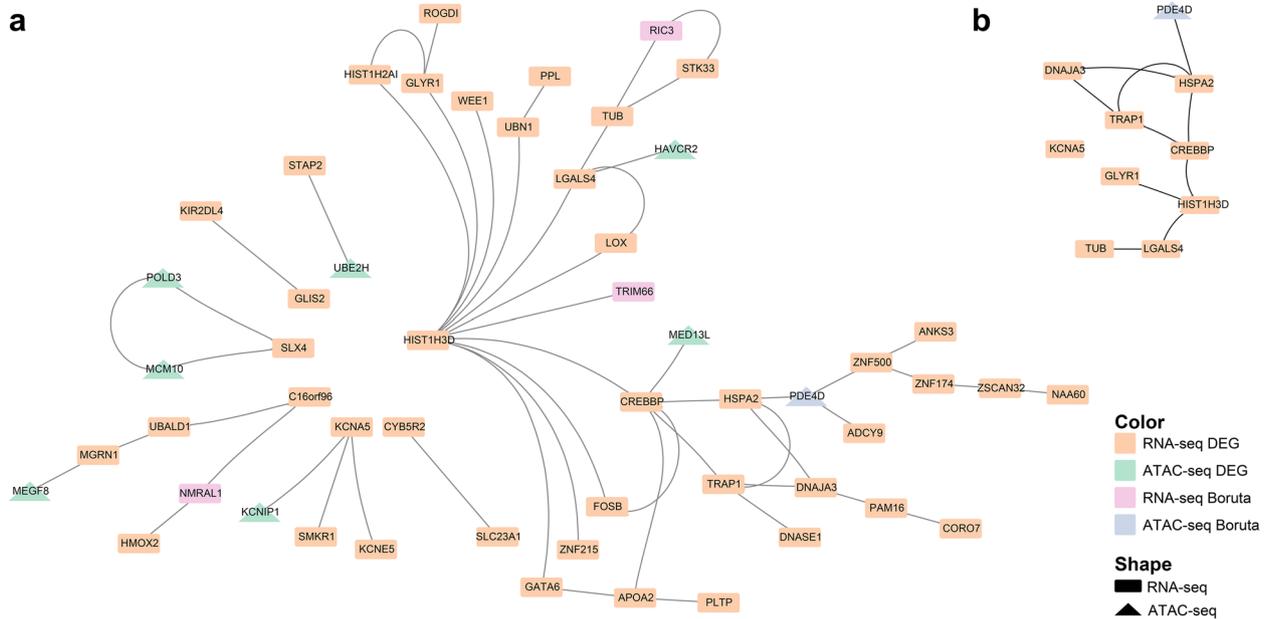
**Fig. 5** Random forest models based on the ATAC-seq data. **a** MDS plot of random forest model to differentiate groups A and B. **b** MDS plot of random forest model to differentiate groups A and C



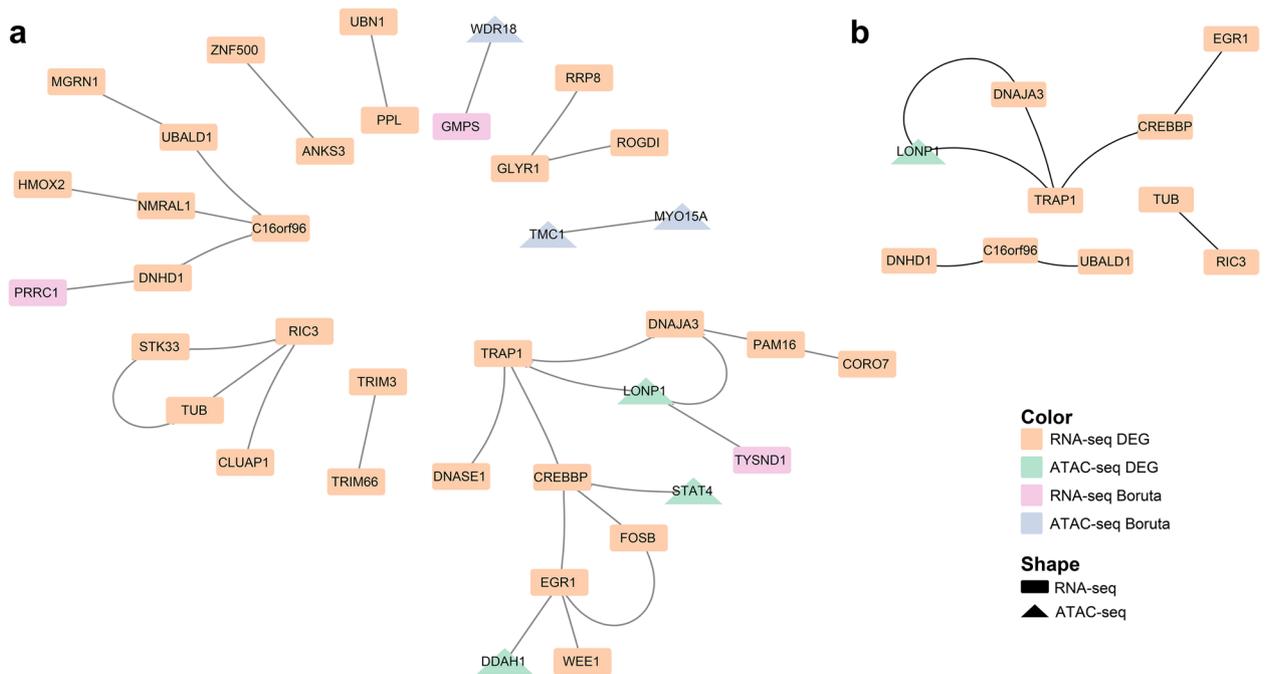
**Fig. 6** Differentially accessible regions were identified by differential expression analysis and the Boruta algorithm. **a** Volcano plot showing differential accessible regions between groups A and B. **b** Volcano plot showing differential accessible regions between groups A and C. **c** Relative importance of important peaks between groups A and B. **d** Relative importance of important peaks between groups A and C

In the biological process category, GO enrichment results revealed that hub genes primarily participated in response to hypoxia, response to decreased oxygen

levels, response to oxygen levels, response to oxidative stress, and regulation of cellular response to oxidative stress. In the cellular component category, hub genes



**Fig. 7** PPI network construction and hub gene identification from groups A and B. **a** PPI network of all DEGs and DARs. **b** PPI network of the ten hub genes. Different colors represent different sequencing using different screening methods. Different shapes represent different sequencing techniques



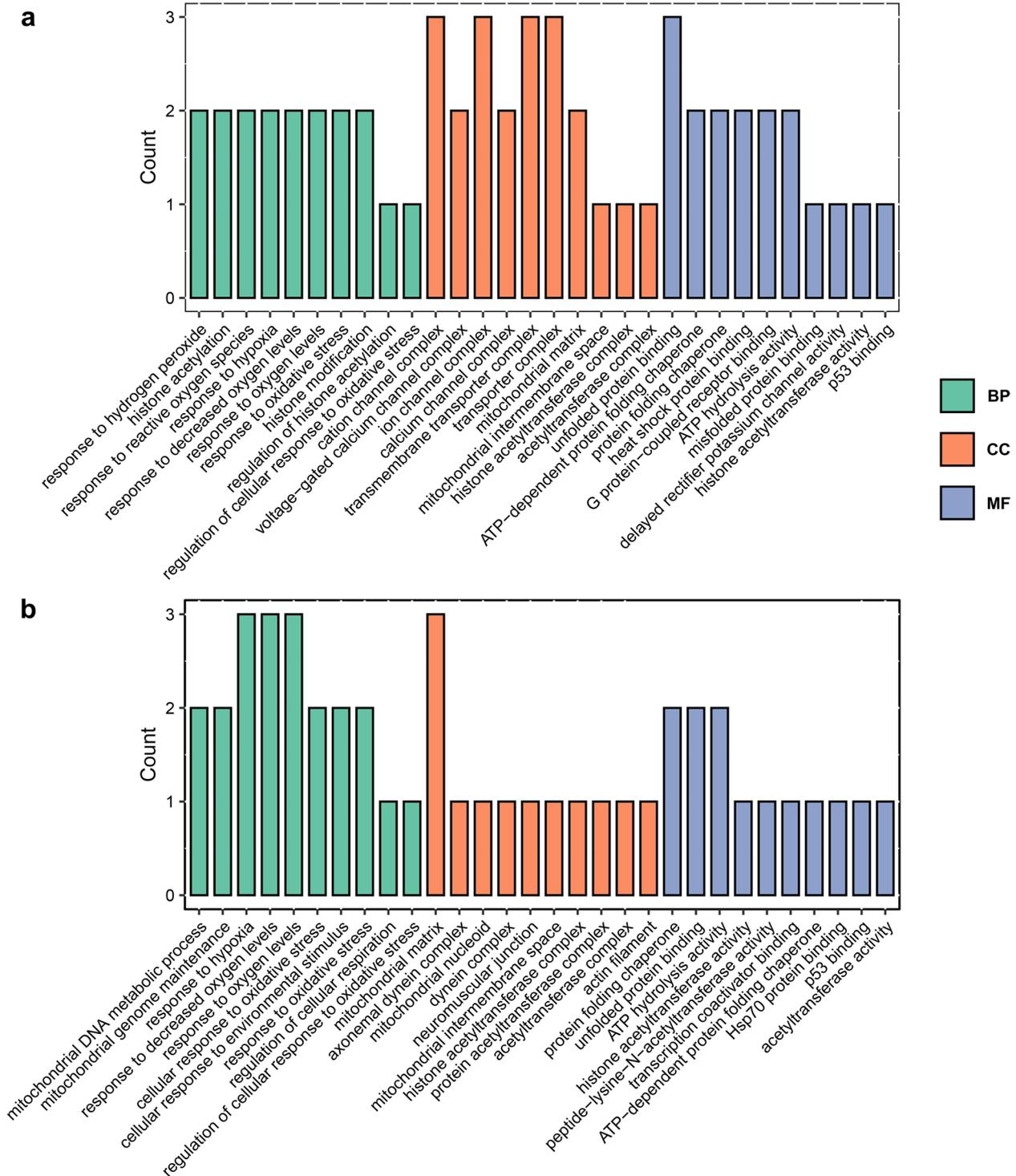
**Fig. 8** PPI network construction and hub gene selection from groups A and C. **a** PPI network of all DEGs and DARs. **b** PPI network of the ten hub genes

were primarily associated with the mitochondrial matrix, mitochondrial intermembrane space, histone acetyltransferase complex, and acetyltransferase complex. In terms

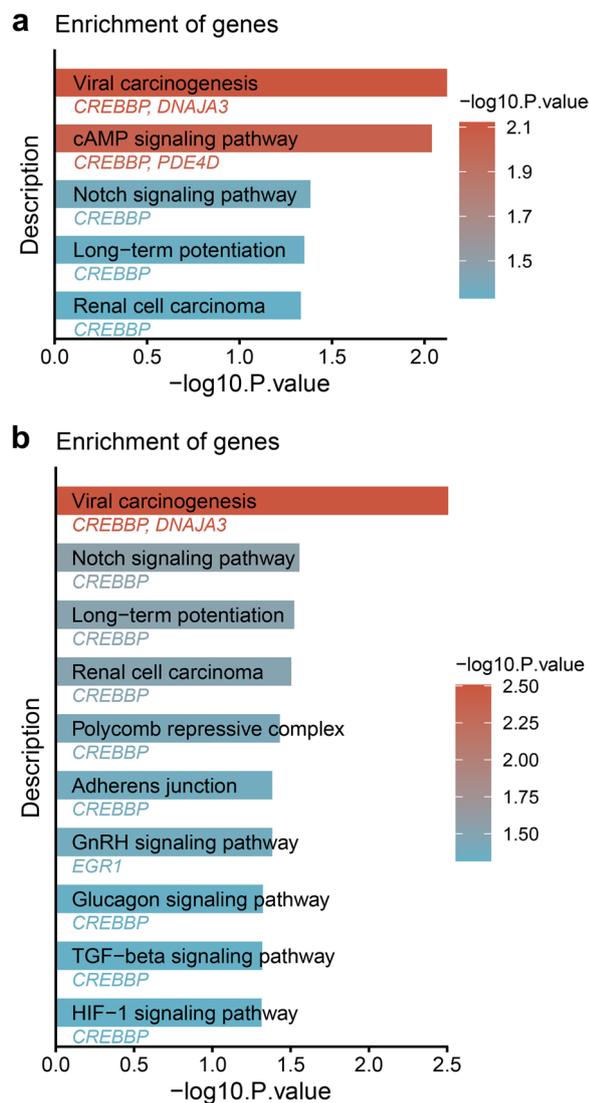
of molecular function, hub genes were primarily involved in protein folding chaperone, unfolded protein binding, adenosine triphosphate (ATP) hydrolysis activity, histone

acetyltransferase activity, ATP-dependent protein folding chaperone, and p53 binding (Fig. 9). KEGG pathway enrichment analysis showed that hub genes were mainly

enriched in the viral carcinogenesis, Notch signaling pathway, long-term potentiation, and renal cell carcinoma (Fig. 10). Notably, HIF-1 signaling pathway was



**Fig. 9** GO enrichment analysis of hub genes. **a** GO enrichment analysis between groups A and B. **b** GO enrichment analysis between groups A and C. The hub genes were enriched into three classifications of biological process (BP), molecular function (MF), and cellular components (CC)



**Fig. 10** KEGG enrichment analysis of ten hub genes. **a** KEGG enrichment analysis between groups A and B. **b** KEGG enrichment analysis between groups A and C

enriched in 4500 m altitude adaptations, not in 3500 m altitude adaptations. Accordingly, GO and KEGG enrichment analyses both showed that hub genes were mainly enriched in biological processes and pathways closely related to hypoxia adaptation.

## Discussion

Acute hypoxia typically occurs within a timescale of minutes to hours, during which the body has evolved exquisite mechanisms for adaptation. During hypoxia, transcription is mainly regulated by the HIFs. HIFs activates genes that control cellular oxygen homeostasis, including genes involved in oxygen consumption,

erythrocyte production, angiogenesis, and mitochondrial metabolism [9]. However, recently, it has been demonstrated that epigenetic regulators and chromatin reprogramming are involved in the hypoxic stress response [10]. Thus, the transcriptome and chromatin accessibility changes in the peripheral blood of eight individuals were analyzed 1 h after adapting to simulated altitudes of 3500 m and 4500 m to better characterize the mechanisms of acute hypoxia adaptation. RNA-seq identified 100 DEGs for the 3500 m adaptation and 53 DEGs for the 4500 m adaptation. ATAC-seq revealed 16 DARs at 3500 m and 20 DARs at 4500 m. Additionally, ten hub genes were identified for each altitude, with functional enrichment analysis indicating their involvement in hypoxia-related biological processes.

The lack of significant differences in RNA and ATAC profiles between group A and group B, or group A and group C, can be attributed to several interrelated factors. Firstly, our data demonstrate substantial interindividual variability, as confirmed by PERMANOVA analysis. This high degree of variability can obscure clear clustering patterns in PCA. Secondly, the response to hypoxia involves a complex interplay of molecular mechanisms, including the stabilization and activation of HIFs and other downstream pathways. These intricate processes can vary widely between individuals, contributing to the observed dispersion in the PCA. Thirdly, the 1 h exposure to hypoxia may not be sufficient to elicit strong, consistent effects across all individuals. Short-term hypoxia exposure might result in transient and variable responses, thereby impacting the clustering observed in the PCA analysis. These limitations underscore the need for further studies with extended exposure times and larger sample sizes to better understand individual-specific responses to hypoxia. Future research should also delve into the molecular mechanisms underlying these responses to elucidate the observed variability.

In our study, combined analysis of RNA-seq and ATAC-seq identified four shared hub genes—*CREBBP*, *TRAP1*, *TUB*, and *DNAJA3*—in adaptation to altitudes of 3500 m and 4500 m. *CREBBP*, which belongs to the type 3 family of lysine acetyl transferases (KAT3), induces histone acetylation to relax the chromatin as epigenetic writers [11]. It binds to the transactivation domain of HIF- $\alpha$  to coactivate HIF-mediated transactivation and is responsible for expression of about 30–50% of global HIF-1 downstream target genes [12]. *TRAP1* is a mitochondrial chaperone from the Hsp90 family that regulates mitochondrial respiration by inhibiting the succinate dehydrogenase complex [13]. This inhibition leads to succinate accumulation, which stabilizes HIF1 $\alpha$  and induces a “pseudo-hypoxic” state, promoting a shift from oxidative phosphorylation to glycolysis and maintaining

mitochondrial homeostasis, also providing cytoprotection [14]. Thus, our results coincide with the previous study, further supporting the significance of these genes in high-altitude adaptation. As for the other two genes, TUB and DNAJA3, although there is no clear evidence at present indicating their connection to hypoxia, their roles in sonic hedgehog signaling and mitochondrial translocation, respectively, may provide valuable insights for future research [15, 16].

Our study systematically elucidates the multifaceted biological functions and potential mechanisms of hub genes in hypoxia adaptation through GO and KEGG enrichment analyses. In the category of biological processes, hub genes are primarily involved in response to hypoxia, oxygen levels, and oxidative stress. A decline in oxygen levels can alter gene transcription or result in post-translational modifications of proteins, leading to changes in cellular metabolism [17]. Except hypoxia, oxidative stress also promotes the activity of HIFs [18, 19]. In the category of cellular components, the enriched GO terms mainly focus on mitochondrial components and acetyltransferase complexes. Mitochondria serve as key oxygen sensors and important signaling organelles, signaling the onset of hypoxia by generating ROS through the electron transport chain [20]. Meanwhile, histone acetyltransferases are defined by a catalytic domain that facilitates the transfer of acetyl groups to lysine residues within the N-terminal tails of histones as well as other protein substrates [21]. Histone acetyltransferases directly acetylate HIF- $\alpha$  to modulate HIF transcriptional activity [22]. Molecular function analysis reveals that these hub genes not only play a role in maintaining the correct folding and functional stability of proteins but also provide energy through ATP hydrolysis, supporting various metabolic and stress responses [23–25]. In the KEGG pathway enrichment analysis, the Notch pathway is a key player in the complex interaction of regulatory pathways mediating hypoxic adaptation and is known to mediate numerous processes that are linked to cellular and tissue remodeling upon high-altitude exposure [26]. These findings are consistent with previous research, showing that changes in gene expression under hypoxia are intimately associated with alterations in chromatin structure. This unveils the multi-layered adaptive mechanisms of cells under hypoxic conditions and provides new perspectives for further exploration of the molecular mechanisms underlying related diseases.

There are several limitations in our study. Firstly, we only investigated changes in the transcriptome and chromatin in the peripheral blood of individuals exposed to hypoxic conditions for 1 and 2 h. These changed factors are more likely to act as early factors in response to hypoxic conditions to further activate downstream

pathways to adapt to the hypoxic environment. Longer exposure time was likely to identify the meaningful differences. However, it is unethical to increase the risk of individuals. Changes in mRNA transcriptome levels do not adequately respond to organismal adaptation to hypoxia. As an example, the stability and activity of HIF-1 $\alpha$  subunit are regulated by its post-translational modifications. HIF-1 $\alpha$  is degraded via ubiquitin–proteasome pathway in normoxia and is stable and interacts with coactivators and regulates the expression of target genes in hypoxia. Therefore, proteomics and metabolomics assays are also needed for future studies. Finally, only eight individuals were included in our study. This sample size is insufficient considering the randomization and bias of transcriptome analysis.

## Conclusions

In summary, we investigated the transcriptome and chromatin profiles in peripheral blood of individuals exposed to hypoxia for 1 and 2 h. Our results can provide the reference for the early response of the organism to hypoxic adaptation.

## Methods

### Sample collection

In this study, hypobaric chambers were used to reduce pressure from ambient barometric pressure to simulate high-altitude conditions. Eight male volunteers were recruited after signing written informed consent. They were healthy unrelated individuals with normal blood tests and no other hereditary disease. The mean age was  $21 \pm 3$  years, and body mass index was  $23.4 \pm 3.7$  kg/m<sup>2</sup>. Exclusion criteria were as follows: travel > 2500 m within 3 month of the first measurements, a previous history of high-altitude pulmonary or cerebral oedema, smoking and pregnancy according to Path's study [8]. All experiments were approved by the Biological and Medical Ethics Committee of Beijing University of Aeronautics and Astronautics.

Peripheral blood samples from eight male volunteers were obtained at three distinct time points. Initially, blood samples were taken after maintaining an atmospheric pressure of 760 mmHg (equivalent to sea level) for 1 h (baseline, group A). Subsequently, the hypobaric chambers were programmed to reduce the atmospheric pressure from 760 to 493 mmHg (equivalent to an altitude of 3500 m) at a rate of 5 m/s, which was then maintained for 1 h (group B). Finally, the atmospheric pressure was further decreased to 448 mmHg (equivalent to an altitude of 4500 m) at the same rate of 5 m/s and maintained for another hour (group C). Blood sample was snap-frozen in liquid nitrogen and stored at  $-80$  °C until further processing. The study was conducted in a

single-blind manner (with volunteers), and the sample collection, experimental procedures, and analyses were carried out by three independent teams of researchers.

#### RNA isolation, library preparation, and sequencing

The total RNA in blood samples was extracted according to the instruction manual of the TRIzol Reagent (Life technologies, California, USA). RNA concentration, purity, and integrity was assessed according to the published study [27]. Sequencing libraries were generated using 1 µg qualified RNA by a NEBNext Ultra™ RNA Library Prep Kit for Illumina (NEB, USA) following the manufacturer's recommendations, and index codes were added to attribute sequences to each sample. Clustering of the index-coded samples was performed on a cBot Cluster Generation System using a TruSeq PE Cluster Kit v4-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina platform and 150 bp paired-end reads were generated.

#### ATAC library preparation and sequencing

About 50,000 cells were subjected to centrifugation for a duration of 5 min (500 g, 4 °C), followed by the removing of supernatant. The cells were washed with cold PBS once and centrifuged for 5 min (500 g, 4 °C), after which the supernatant was removed. Then, cells were suspended with cold lysis buffer and centrifuged again (10 min, 500 g, 4 °C) to remove the supernatant. The transposing reaction system was configured with the Tn5 Transposase. The cell nuclei were suspended with the transposing reaction system, and the DNA was purified after incubation at 37 °C for 30 min. The PCR reaction system was configured with the purified DNA, and then the PCR amplification reaction was performed. The final DNA libraries were sequenced on Illumina NovaSeq platform after the DNA was purified.

#### RNA-seq data analysis

Raw data (raw reads) were firstly processed through in-house perl scripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter, reads containing ploy-N, and low-quality reads from raw data. All the subsequent analyses were based on clean data. Paired-end reads were aligned to the human reference genome by Hisat2 with default parameters, followed by HTSeq-count to count the reads mapped to the genome [28, 29]. Then, a count matrix was used as the input to identify the differentially expressed genes (DEGs) between different groups by the DESeq2 package [30]. Candidate genes with greater than twofold changes at  $P$  values < 0.05 were considered to be significant DEGs.

#### ATAC-seq data processing

Raw reads were filtered to remove adapters and low-quality reads, and then high-quality clean reads provided in FASTQ format were obtained for subsequent information analysis. The Bowtie2 software was used to compare the high-quality reads obtained from the sequencing of each sample with the reference genome. Properly paired-end reads with high mapping quality score (MAPQ) were retained in analysis with Samtools [31]. Low-quality mapping reads (MAPQ < 10) and duplicated reads were omitted for further analysis. ATAC-Seq peaks were called using MACS2 [32]. Differential accessible regions (DARs) analysis of two conditions/groups was performed using DESeq2 package [30]. Genes with greater than two-fold changes at  $P$  values < 0.05 found by DESeq2 were assigned as DARs.

#### Multivariate statistics and biomarker selection

The processed data were analyzed by R package (ropIs), where it was subjected to multivariate statistical analysis, including principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA). For the PLS-DA models, the R2Y (goodness of fit) and Q2Y (goodness of prediction) values were used to assess the performance of each model. Models were further validated for fit (R2) and predictability (Q2) by a permutation test performed with 200 random permutations. To identify the effect of factors affecting the differences between RNA-seq or ATAC-seq profiles from different individuals, permutational multivariate analysis of variance (PERMANOVA) will be used.

Random forest (RF) ensemble learners provide estimates of variables (genes or peaks) importance [33]. Higher importance represents a more important variable in distinguishing different groups. Here, the Boruta feature selection method was adopted to identify the most important variables [34].

All the differentially expressed genes and features were used to construct the protein-protein interaction (PPI) network by STRING (<https://string-db.org/>). The hub genes in the PPI network were identified by using CytoHubba plugin [35], based on maximal clique centrality (MCC) algorithms. Finally, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed for functional gene annotation of these hub genes.

#### Abbreviations

RNA-seq	RNA sequencing
ATAC-seq	ATAC sequencing
DEGs	Differentially expressed genes
DARs	Differentially accessible regions
PPI	Protein-protein interaction
HIF	Hypoxia-inducible factor

NRF2	Nuclear factor erythroid 2-related factor 2
ROS	Reactive oxygen species
TLR4	Toll-like receptor 4
MAPQ	Mapping quality score
PCA	Principal component analysis
PLS-DA	Partial least squares discriminant analysis
R2Y	Goodness of fit
Q2Y	Goodness of prediction
R2	Fit
Q2	Predictability
RF	Random forest
PERMANOVA	Permutational multivariate analysis of variance
MCC	Maximal clique centrality
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
CREBBP	CREB-binding protein
TRAP1	Tumor necrosis factor receptor-associated protein
TUB	Tubulin
DNAJA3	DnaJ (Hsp40) homolog subfamily A member 3
ATP	Adenosine triphosphate
KAT3	Type 3 family of lysine acetyl transferases

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12915-025-02123-z>.

Additional file 1: Fig. S1. The score scatter plots and permutation tests of PLS-DA models based on transcriptional profiles.

Additional file 2: Fig. S2. The score scatter plots and permutation tests of PLS-DA models based on the ATAC-seq data.

Additional file 3: Fig. S3. Heat map displaying differentially expressed genes identified by differential expression analysis and the Boruta algorithm.

Additional file 4: Fig. S4. Heat map displaying differentially accessible regions identified by differential expression analysis and the Boruta algorithm.

## Acknowledgements

We would like to thank editor and anonymous reviewers for their careful reading of our manuscript as well as for their invaluable constructive comments and suggestions.

## Authors' contributions

Z.D. and Y.J. conceived and designed the experiments. Z.K. and Y.P.R. collected the sample. X.J.F. and Z.L.X. performed the experiments. Y.J. and L.X.N. analyzed the data. Z.D. wrote the paper. All authors reviewed and approved the final manuscript.

## Funding

This study was supported by Defense Industrial Technology Development Program (JCKY2024601C018).

## Data availability

All data generated or analyzed during this study are included in this published article, its supplementary information files, and publicly available repositories. The sequencing data (RNA-seq and ATAC-seq) generated and analyzed for the study are available in the SRA database with the accession ID PRJNA1195337 (<https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA1195337>). The main code for the analysis used in this study is publicly accessible on Zenodo (<https://zenodo.org/records/14533626>).

## Declarations

### Ethics approval and consent to participate

The study was approved by the Biological and Medical Ethics Committee of Beijing University of Aeronautics and Astronautics (BM20210149). All participants were included after providing written informed consent.

### Consent for publication

We have obtained consent to publish from the participants to report their data.

### Competing interests

The authors declare that they have no competing interests.

Received: 17 April 2024 Accepted: 8 January 2025

Published online: 21 January 2025

## References

- Furian M, Bitos K, Hartmann SE, Murali L, Lichtblau M, Bader PR, Rawling JM, Ulrich S, Poulin MJ, Bloch KE. Acute high altitude exposure, acclimatization and re-exposure on nocturnal breathing. *Front Physiol.* 2022;13:965021.
- Idrose AM, Juliana N, Azmani S, Yazit NAA, Muslim MSA, Ismail M, Amir SN. Singing improves oxygen saturation in simulated high-altitude environment. *J Voice.* 2022;36(3):316–21.
- Storz JF, Bautista NM. Altitude acclimatization, hemoglobin-oxygen affinity, and circulatory oxygen transport in hypoxia. *Mol Aspects Med.* 2022;84:101052.
- Shen Z, Huang D, Jia N, Zhao S, Pei C, Wang Y, Wu Y, Wang X, Shi S, Wang F, et al. Protective effects of Eleutheroside E against high-altitude pulmonary edema by inhibiting NLRP3 inflammasome-mediated pyroptosis. *Biomed Pharmacother.* 2023;167:115607.
- Mallet RT, Burtscher J, Pialoux V, Pasha Q, Ahmad Y, Millet GP, Burtscher M. Molecular mechanisms of high-altitude acclimatization. *Int J Mol Sci.* 2023;24(2):1698.
- Jaskiewicz M, Moszynska A, Krolczewski J, Cabaj A, Bartoszewska S, Charzynska A, Gebert M, Dabrowski M, Collawn JF, Bartoszewski R. The transition from HIF-1 to HIF-2 during prolonged hypoxia results from reactivation of PHDs and HIF1A mRNA instability. *Cell Mol Biol Lett.* 2022;27(1):109.
- Lukyanova LD, Dudchenko AM, Tsybina TA, Germanova EL, Tkachuk EN, Erenburg IV. Effect of intermittent normobaric hypoxia on kinetic properties of mitochondrial enzymes. *Bull Exp Biol Med.* 2007;144(6):795–801.
- Pham K, Frost S, Parikh K, Puvvula N, Oeung B, Heinrich EC. Inflammatory gene expression during acute high-altitude exposure. *J Physiol.* 2022;600(18):4169–86.
- Choudhry H, Harris AL. Advances in hypoxia-inducible factor biology. *Cell Metab.* 2018;27(2):281–98.
- Chen Y, Liu M, Niu Y, Wang Y. Romance of the three kingdoms in hypoxia: HIFs, epigenetic regulators, and chromatin reprogramming. *Cancer Lett.* 2020;495:211–23.
- Dutto I, Scalera C, Prosperi E. CREBBP and p300 lysine acetyl transferases in the DNA damage response. *Cell Mol Life Sci.* 2018;75(8):1325–38.
- Kasper LH, Boussouar F, Boyd K, Xu W, Biesen M, Rehg J, Baudino TA, Cleveland JL, Brindle PK. Two transactivation mechanisms cooperate for the bulk of HIF-1-responsive gene expression. *EMBO J.* 2005;24(22):3846–58.
- Sciacovelli M, Guzzo G, Morello V, Frezza C, Zheng L, Nannini N, Calabrese F, Laudiero G, Esposito F, Landriscina M, et al. The mitochondrial chaperone TRAP1 promotes neoplastic growth by inhibiting succinate dehydrogenase. *Cell Metab.* 2013;17(6):988–99.
- Joshi A, Ito T, Picard D, Neckers L. The mitochondrial HSP90 paralog TRAP1: structural dynamics, interactome, role in metabolic regulation, and inhibitors. *Biomolecules.* 2022;12(7):880.
- Mukhopadhyay S, Jackson PK. The tubby family proteins. *Genome Biol.* 2011;12(6):225.

16. Banerjee S, Chaturvedi R, Singh A, Kushwaha HR. Putting human Tid-1 in context: an insight into its role in the cell and in different disease states. *Cell Commun Signal*. 2022;20(1):109.
17. Batié M, Fasanya T, Kenneth NS, Rocha S. Oxygen-regulated post-translational modifications as master signalling pathway in cells. *EMBO Rep*. 2023;24(12): e57849.
18. Semenza GL. Hypoxia-inducible factors in physiology and medicine. *Cell*. 2012;148(3):399–408.
19. Snyder CM, Chandel NS. Mitochondrial regulation of cell survival and death during low-oxygen conditions. *Antioxid Redox Signal*. 2009;11(11):2673–83.
20. Waypa GB, Smith KA, Schumacker PT. O<sub>2</sub> sensing, mitochondria and ROS signaling: the fog is lifting. *Mol Aspects Med*. 2016;47–48:76–89.
21. Perez-Perri JJ, Acevedo JM, Wappner P. Epigenetics: new questions on the response to hypoxia. *Int J Mol Sci*. 2011;12(7):4705–21.
22. Luo W, Wang Y. Epigenetic regulators: multifunctional proteins modulating hypoxia-inducible factor- $\alpha$  protein stability and activity. *Cell Mol Life Sci*. 2018;75(6):1043–56.
23. Lee P, Chandel NS, Simon MC. Cellular adaptation to hypoxia through hypoxia inducible factors and beyond. *Nat Rev Mol Cell Biol*. 2020;21(5):268–83.
24. Li C, Hao B, Yang H, Wang K, Fan L, Xiao W. Protein aggregation and biomolecular condensation in hypoxic environments (Review). *Int J Mol Med*. 2024;53(4):33.
25. Wheaton WW, Chandel NS. Hypoxia. 2. Hypoxia regulates cellular metabolism. *Am J Physiol Cell Physiol*. 2011;300(3):C385–393.
26. O'Brien KA, Murray AJ, Simonson TS. Notch signaling and cross-talk in hypoxia: a candidate pathway for high-altitude adaptation. *Life (Basel)*. 2022;12(3):437.
27. Wilfinger WW, Eghbalnia HR, Mackey K, Miller R, Chomczynski P. Whole blood RNA extraction efficiency contributes to variability in RNA sequencing data sets. *PLoS ONE*. 2023;18(11): e0291209.
28. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357–60.
29. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166–9.
30. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
31. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. Genome Project Data Processing S: The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.
32. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137.
33. Liaw A, Wiener M. Classification and regression by randomForest. *R news*. 2002;2(3):18–22.
34. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. #N/A. 2010;36:1–13.
35. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014;8 Suppl 4(Suppl 4):S11.

## Publisher's Note

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