Heliyon 9 (2023) e20226

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

# The role of *EPAS1* polymorphisms on COPD susceptibility in southern Chinese

Yunchao Wang [a,1], Ao Lin [b,1], Ruiqi He [a,1], Cuiyi Chen [c], Xiaobin Zeng [b], Yujie Pan [b], Chun Mao [b], Chenli Xie [d], Dongsheng Huang [e], Yibin Deng [f], Xuhui Zhang [g], Jiachun Lu [b], Xinhua Wang [a,*]

[a] *Institute of Basic Medicine, Institute of Public Health, Gansu University of Chinese Medicine, Lanzhou, Gansu Province, 730030, China*
[b] *The State Key Lab of Respiratory Disease, The First Affiliated Hospital, Institute of Public Health, Guangzhou Medical University, Guangzhou, Guangdong Province, 510000, China*
[c] *Department of Respiratory Medicine, SSL Central Hospital of Dongguan City, Dongguan, Guangdong Province, 523000, China*
[d] *Department of Respiratory and Critical Care Medicine, Dongguan Binwan Central Hospital, Dongguan, Guangdong Province, 523000, China*
[e] *Department of Respiratory and Critical Care Medicine, Shenzhen Longhua District Central Hospital, Shenzhen, Guangdong Province, 515100, China*
[f] *Centre for Medical Laboratory Science, The Affiliated Hospital of Youjiang Medical University for Nationalities, Baise, Guangxi Province, 533000, China*
[g] *Department of Respiratory Medicine, Affiliated Hospital of Gansu University of Chinese Medicine, Lanzhou, Gansu Province, 730000, China*

ARTICLE INFO

*Keywords:*
COPD
*EPAS1*
Polymorphism
Susceptibility
Absolute risk

ABSTRACT

*Objective:* COPD is the most common chronic respiratory disease with complex environmental and genetic etiologies. It was reported that *EPAS1* might participate in the occurrence and development of respiratory diseases. However, the association between *EPAS1* and COPD was unclear.
*Methods:* First, a case-control study enrolling 1130 COPD patients and 1115 healthy controls in Guangzhou was conducted to clarify the association between *EPAS1* polymorphisms and COPD susceptibility. Secondly, a prevalence study recruited 882 participants in Gansu to verify the effect of positive polymorphisms on lung function. Finally, the 10-year absolute risk considering environmental factors and genetic variations was calculated by the method of Gail and Bruzzi.
*Results: EPAS1* rs13419896 AA genotype reduced COPD risk in southern Chinese (AA vs. GG: adjusted OR = 0.689, 95% CI = 0.498–0.955; AA vs. GG/GA: adjusted OR = 0.701, 95% CI = 0.511–0.962). Further, the rs13419896 A allele was significantly associated with higher pre-FEV1/pre-FVC in both the Guangzhou and Gansu populations ($P < 0.05$). Smoking status, coal as fuels, education level, and rs13419896 G > A were finally retained to develop a relative risk model for males. Smoking status, biomass as fuels, and rs13419896 G > A were retained in the female model. The population-attributable risk of the male or female model was 0.457 (0.283–0.632) and 0.421 (0.227–0.616), respectively.
*Conclusions:* This study first revealed that *EPAS1* rs13419896 G > A decreased COPD susceptibility and could be a genetic marker to predict the 10-year absolute risk for COPD.

* Corresponding author. Institute of Basic Medicine, Institute of public health, Gansu University of Chinese Medicine, No. 35 of Dingxi East Road, Lanzhou, Gansu Province, 730030, China.
*E-mail address:* wxh@gszy.edu.cn (X. Wang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2023.e20226
Received 20 October 2022; Received in revised form 10 August 2023; Accepted 14 September 2023
Available online 23 September 2023

## 1. Introduction

Chronic obstructive pulmonary disease (COPD), characterized by persistent respiratory symptoms and irreversible airflow restriction, is the most common chronic respiratory disease [1]. The prevalence of COPD is 10.3% in the world, which raises to 13.7% among Chinese older than 40 [2,3]. It is difficult to diagnose early COPD because of its hidden development. With COPD progress, moderate or severe airflow restriction will seriously affect the quality of life and even cause death [4]. Recently, some studies showed that COPD and lung cancer shared similar etiologies, suggesting that COPD might be a driver of lung cancer [5]. As one of the primary outcomes of COPD, lung cancer has the highest incidence and mortality in Chinese cancers [6]. Therefore, clarifying the etiologies of COPD is of great significance for jointly preventing and treating COPD and lung cancer, which will dramatically reduce the Chinese disease burden.

As we all know, COPD is regulated by genetic and environmental factors, and smoking is the most important environmental risk factor [1,3]. Notedly, COPD only occurs in 10%–20% of long-term smokers and even in some non-smokers, indicating an indispensable role of genetic factors in the occurrence and development of COPD [7]. Single nucleotide polymorphism (SNP), the third-generation genetic marker, is a powerful tool for detecting the genetic etiology of diseases [8]. Although more and more COPD-related genetic loci have been discovered by genome-wide association studies (GWAS) and candidate gene studies, the genetic etiology of COPD remains largely unknown [9].

Endothelial PAS domain protein 1 (*EPAS1*), also known as hypoxia-induced factor-2α (*HIF-2α*), codes a vital subunit of hypoxia-induced factor 2 (HIF-2) and expresses in lung and other organs involved in oxygen metabolism [10,11]. When hypoxia happens, accumulated HIF-2α persistently stimulates the HIF pathway to promote the transcription of downstream genes that involve erythropoiesis, iron homeostasis, metabolism, inflammation, vascularization, and tumorigenesis [12,13]. Some researchers have reported that *EPAS1* is abnormally expressed in COPD and may regulate several COPD-related genes, raising a hypothesis that *EPAS1* would take part in the occurrence and development of COPD [14,15]. However, those hypotheses need to be further verified.

Therefore, we conducted a multicenter case-control study enrolling 1130 COPD patients and 1115 healthy controls in Guangzhou to explore the association between *EPAS1* polymorphisms (rs13419896 G > A, rs59901247 A > C, rs6756667 G > A) and COPD susceptibility in southern Chinese. Furthermore, we performed a prevalence study in Gansu and collected 882 participants with complete records of pulmonary function examination to verify the effect of positive polymorphisms on lung function. Finally, 10-year absolute risks for the southern Chinese with different individual relative risks were calculated and demonstrated in tables.

## 2. Materials and methods

### 2.1. Study population

We totally collected 1130 COPD patients and 1115 healthy controls in the case-control study. They were Southern Chinese who visited the Songshan Lake Central Hospital of Dongguan, the Dongguan Binwan Central Hospital, and the Shenzhen Longhua District Central Hospital from 2015 to 2019 (Table S1).

In 2019, 882 participants from four regions in Gansu were enrolled by multi-stage cluster random sampling to conduct a prevalence study (Chengxian County of Longnan City, Dunhuang City, Zhengning County of Qingyang City, and Zhuoni County of Gannan Tibetan Autonomous Prefecture) (Table S2).

Questionnaires were utilized to collect information about demographic characteristics and environmental exposures. We also asked them to donate 5 ml of peripheral blood for genotyping. The questionnaire details were described in our previous article [16].

The trained staff utilized electronic medical records of the above institutions from 2016 to 2020 to carry out a clinical cohort study and acquire the incidence of COPD. The details were summarized in Table S3.

Our study has obtained participants' informed consent and approbated by the institutional review boards of Guangzhou Medical University and the Gansu University of Chinese Medicine.

### 2.2. Diagnostic criterion for COPD and lung function examination

We diagnosed COPD referring to the Global Initiative for Chronic Obstructive Lung Disease 2019 [17]. Briefly, participants were diagnosed with COPD if they had respiratory symptoms in daily life (coughing, expectoration, dyspnea, wheezing) and the ratio of forced expiratory volume in 1 s ($FEV_1$) to forced vital capacity (FVC) was less than 70% after inhaling 400 μg salbutamol for half an hour. Following the instructions, lung function was examined by the EasyOne Spirometer (NDD Medizintechnik AG, Switzerland).

### 2.3. Polymorphism selection and genotyping

*EPAS1* rs13419896 G > A, rs59901247 A > C, and rs6756667 G > A were screened from the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP) based on the following criteria: SNPs were limited in the region between 2000bp upstream and downstream of *EPAS1*; the minor allele frequencies (MAFs) in the Chinese population were more than 0.05; the linkage disequilibrium of the selected SNPs was low (LD, $R^2 < 0.8$); they were potential functional variations or tag SNPs [18,19]. As reported by previous studies, Rs13419896 G > A and rs6756667 G > A were located in the transcription factor binding regions and might change the expression of *EPAS1*; Rs59901247 A > C was an nsSNP that might cause amino acid substitution [20,21]. LD between three SNPs was exhibited in

Fig. S1.

DNA was extracted from peripheral blood by the DNA Blood Mini Kit (Qiagen, Valencia, CA, USA). ABI7500 System (Applied Biosystems, Foster City, CA, USA) was used to analyze the result of TaqMan real-time polymerase chain reaction (PCR). The information on primers and probes was shown in Table S4. To ensure the reliability and accuracy of genotyping, we not only set up positive and negative controls on each plate but also randomly selected 10% samples for repetition.

### 2.4. Statistical analysis

The $\chi^2$ test was used to compare the differences in the frequency distributions of demographic characteristics between cases and controls. The goodness-of-fit $\chi^2$ test was applied to assess the Hardy-Weinberg equilibrium (HWE) of the selected polymorphisms in controls. To adjust the confounding factors, we adopted a multivariate logistic regression model to evaluate the association between *EPAS1* polymorphisms and COPD susceptibility, as well as the multiplicative interaction of environmental factors and positive polymorphism. The homogeneity of the ORs between different strata was appraised by the Breslow-Day test. The effect of genotypes on lung function was detected by the Kruskal-Wallis test. Furthermore, the false-positive report probabilities (FPRP) of positive findings were calculated by setting the prior probability as 0.1 and the plausible OR as 0.67 [22,23].

The 10-year absolute risk was calculated in the method developed by Gail and Bruzzi (**Appendix**) [24–27]. Specifically, the incidence of COPD and the mortality without COPD were obtained from the clinical cohort study and the 2014–2015 Annual Report on the Cause of Death Surveillance in Guangzhou (Table S3) [28]. Based on the case-control study in Guangzhou, the backward stepwise logistic regression model was used to screen risk factors and estimate relative risks. Adjusted population-attributable risks were produced by IRAP 2.2.0 (Version 2.2.0; National Cancer Institute, Rockville, MD, USA).

Most statistical analyses were performed by SPSS 26.0 (Version 26.0, IBM Corporation, Armonk, NY, USA). All tests were two-sided. Results were considered statistically significant only when *P*-values were less than 0.05.

## 3. Results

### 3.1. Association between EPAS1 polymorphisms and COPD susceptibility

As shown in Table 1, the frequency distribution of genotypes in controls obeyed HWE ($P > 0.05$). *EPAS1* rs13419896 AA genotype was associated with decreased COPD susceptibility (AA vs. GG: adjusted OR = 0.689, 95%CI = 0.498–0.955; and AA vs. GG/GA: adjusted OR = 0.701, 95%CI = 0.511–0.962). Unfortunately, we could not observe that rs59901247 A > C and rs6756667 G > A modified COPD susceptibility.

### 3.2. Stratification analysis

As shown in Table 2, compared with *EPAS1* rs13419896 GG/GA genotypes, the protective effect of the AA genotype was still significant in the strata of age>60 years, female, low education level, non-smokers, and avoiding biomass as fuels ($P < 0.05$).

**Table 1**
Frequency distributions of *EPAS1* polymorphisms and their associations with COPD susceptibility.

| Models | Genotypes | Case (n = 1130) n (%) | Control (n = 1115) n (%) | Adjusted OR (95% CI)[a] | *P*[a] | AIC |
|---|---|---|---|---|---|---|
| rs13419896 G > A (HWE = 0.603) | | | | | | |
| | GG | 587 (51.9) | 550 (49.3) | 1.000 (ref.) | | |
| | GA | 469 (41.5) | 461 (41.3) | 0.963 (0.808–1.149) | 0.679 | |
| Codominant | AA | 74 (6.5) | 104 (9.3) | **0.689(0.498–0.955)** | **0.025** | 3030.394 |
| Additive | GG vs. GA vs. AA | | | 0.887 (0.777–1.013) | 0.077 | 3032.349 |
| Dominant | GA + AA vs. GG | 543 (48.1) | 565 (50.7) | 0.917 (0.772–1.081) | 0.294 | 3034.373 |
| Recessive | AA vs. GG + GA | 74 (6.5) | 104 (9.3) | **0.701(0.511–0.962)** | **0.028** | 3030.565 |
| rs59901247 A > C (HWE = 0.760) | | | | | | |
| | AA | 895 (79.2) | 878 (78.7) | 1.000 (ref.) | | |
| | AC | 217 (19.2) | 224 (20.1) | 0.986 (0.797–1.219) | 0.895 | |
| Codominant | CC | 18 (1.6) | 13 (1.2) | 1.403 (0.674–2.920) | 0.365 | 3034.609 |
| Additive | AA vs. AC vs. CC | | | 1.031 (0.854–1.244) | 0.754 | 3035.376 |
| Dominant | AC + CC vs. AA | 235 (20.8) | 237 (21.3) | 1.009 (0.820–1.241) | 0.935 | 3035.467 |
| Recessive | CC vs. AA + AC | 18 (1.6) | 13 (1.2) | 1.407 (0.677–2.925) | 0.360 | 3034.626 |
| rs6756667 G > A (HWE = 0.552) | | | | | | |
| | GG | 878 (77.7) | 846 (75.9) | 1.000 (ref.) | | |
| | GA | 236 (20.9) | 253 (22.7) | 0.901 (0.734–1.105) | 0.318 | |
| Codominant | AA | 16 (1.4) | 16 (1.4) | 1.029 (0.505–2.099) | 0.936 | 3034.455 |
| Additive | GG vs. GA vs. AA | | | 0.926 (0.772–1.111) | 0.410 | 3034.794 |
| Dominant | GA + AA vs. GG | 252 (22.3) | 269 (24.1) | 0.908 (0.744–1.109) | 0.345 | 3034.583 |
| Recessive | AA vs. GG + GA | 16 (1.4) | 16 (1.4) | 1.053 (0.517–2.144) | 0.886 | 3035.454 |

OR, odds ratio; CI, confidence interval; HWE, Hardy-Weinberg equilibrium; AIC, Akaike Information Criterion.
[a] ORs were adjusted for gender, age, smoking status, education level, coal or biomass as fuels by the logistic regression model.

Furthermore, the homogeneity test did not discover any heterogeneity between the ORs from different strata ($P > 0.05$). Also, no remarkable interaction between environmental factors and *EPAS1* rs13419896 G > A was identified ($P > 0.05$).

### 3.3. Effect of EPAS1 rs13419896 G > A on lung function in the Guangzhou and Gansu population

We further explored the effect of *EPAS1* rs13419896 G > A on lung function in the Guangzhou population. A total of 1656 participants with eligible lung function estimation were successfully genotyped. The number of participants with GG, GA, and AA genotypes was 826, 700, and 130, respectively. Although there were no statistical differences in pre-$FEV_1$ and pre-FVC among different genotypes ($P > 0.05$), the A allele was significantly associated with higher pre-$FEV_1$/pre-FVC (Fig. 1C). This result was verified in the Gansu population (Fig. 2C).

### 3.4. Relative risks and population-attributable risks

The backward stepwise logistic regression model retained smoking status, biomass as fuels, education level, and rs13419896 G > A to construct the male model; smoking status, biomass as fuels, and rs13419896 G > A were retained for the female model (Table 3). The population-attributable risks of the male and female models were 0.457 (0.283–0.632) and 0.421 (0.227–0.616).

### 3.5. The 10-year absolute risk for COPD among southern Chinese

Combining rs13419896 G > A with the environmental factors, we calculated the 10-year absolute risk for the southern people with different individual relative risks (**Appendix**) [24–27]. For example, a 40 years old woman who smokes, uses biomass as fuels, and carries rs13419896 AA genotype has an individual relative risk of 4.527 ($r = 1.800 \times 2.515 \times 1$). Her probability of developing COPD is about 0.149 at 50 years old (Table 4).

## 4. Discussion

This study revealed that compared with *EPAS1* rs13419896 GG/GA genotype, the AA genotype significantly reduced COPD risk in southern Chinese. The protective effect of the AA genotype remained remarkable in the strata of age>60 years, females, low education level, non-smokers, and avoiding biomass as fuels. Moreover, the A allele was associated with a higher pre-$FEV_1$/pre-FVC in the Guangzhou and Gansu population. Therefore, we calculated the 10-year absolute risks for COPD by combining the rs13419896 G > A with environmental factors. For the first time, this study investigated the association between *EPAS1* polymorphisms and COPD susceptibility and estimated the absolute 10-year risk for COPD.

*EPAS1*, located at 2p16-21, encodes a protein named HIF-2α that dimerizes with HIF-1β to form HIF-2 [10]. The degradation of HIF-2α depends on the ubiquitination mediated by the von Hippel-Lindau protein (pVHL). When hypoxia happens, the hydroxylation of HIF-2α is attenuated, thus hindering the binding of HIF-2α to pVHL. Accumulated HIF-2α continually activates the HIF pathway, which may affect the expression of downstream genes involving various physiological and biochemical processes [29]. Using the Lung

**Table 2**
Stratification analysis of the association between *EPAS1* rs13419896 G > A and COPD susceptibility.

| Variables | Case (n = 1130) | | | Control (n = 1115) | | | AA vs. GG + GA | $P_{homo}$[b] | $P_{inter}$[c] |
|---|---|---|---|---|---|---|---|---|---|
| | GG n (%) | GA n (%) | AA n (%) | GG n (%) | GA n (%) | AA n (%) | OR (95% CI)[a] | | |
| Age group (years) | | | | | | | | | |
| ≤60 | 223 (47.8) | 208 (44.5) | 36 (7.7) | 246 (49.0) | 208 (41.4) | 48 (9.6) | 0.794 (0.494–1.278) | 0.399 | 0.232 |
| >60 | 364 (54.9) | 261 (39.4) | 38 (5.7) | 304 (49.6) | 253 (41.3) | 56 (9.1) | **0.556(0.356–0.869)** | | |
| Gender | | | | | | | | | |
| Male | 352 (50.0) | 300 (42.6) | 52 (7.4) | 360 (50.2) | 294 (41.0) | 63 (8.8) | 0.853 (0.572–1.271) | 0.097 | 0.121 |
| Female | 235 (55.2) | 169 (39.7) | 22 (5.2) | 190 (47.7) | 167 (42.0) | 41 (10.3) | **0.425(0.244–0.740)** | | |
| Education level | | | | | | | | | |
| Middle school or below | 458 (53.8) | 344 (40.4) | 49 (5.8) | 368 (48.8) | 313 (41.5) | 73 (9.7) | **0.561(0.383–0.823)** | 0.073 | 0.067 |
| High school or above | 129 (46.2) | 125 (44.8) | 25 (9.0) | 182 (50.4) | 148 (41.0) | 31 (8.6) | 1.063 (0.603–1.876) | | |
| Smoking status | | | | | | | | | |
| No | 302 (53.6) | 230 (40.9) | 31 (5.5) | 321 (49.3) | 266 (40.9) | 64 (9.8) | **0.543(0.346–0.852)** | 0.130 | 0.208 |
| Yes | 285 (50.3) | 239 (42.2) | 43 (7.6) | 229 (49.4) | 195 (42.0) | 40 (8.6) | 0.873 (0.549–1.388) | | |
| Coal as fuels | | | | | | | | | |
| No | 504 (51.4) | 407 (41.5) | 70 (7.1) | 508 (49.5) | 422 (41.1) | 97 (9.4) | 0.738 (0.534–1.022) | 0.197 | 0.260 |
| Yes | 83 (55.7) | 62 (41.6) | 4 (2.7) | 42 (47.7) | 39 (44.3) | 7 (8.0) | 0.291 (0.074–1.143) | | |
| Biomass as fuels | | | | | | | | | |
| No | 487 (51.8) | 393 (41.8) | 61 (6.5) | 489 (48.6) | 421 (41.8) | 96 (9.5) | **0.673(0.480–0.944)** | 0.480 | 0.511 |
| Yes | 100 (52.9) | 76 (40.2) | 13 (6.9) | 61 (56.0) | 40 (36.7) | 8 (7.3) | 1.143 (0.434–3.012) | | |

OR, odds ratio; CI, confidence interval.
[a] ORs were adjusted for age, gender, education level, smoking status, coal or biomass as fuels by the logistic regression model.
[b] *P*-value of the Breslow-Day homogeneity test for the ORs between strata.
[c] *P*-value of test for the multiplicative interaction between *EPAS1* rs13419896 genotypes and the factors.

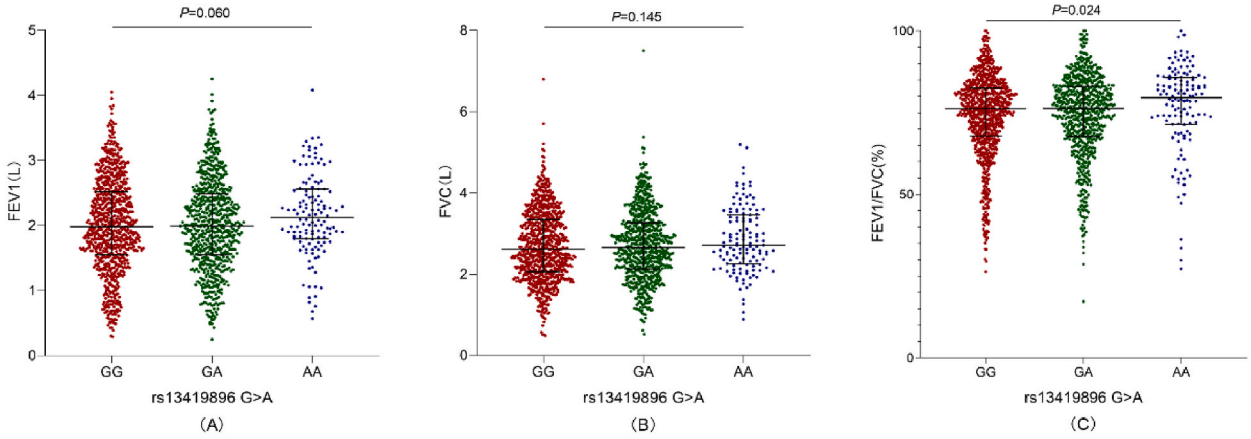


**Fig. 1.** The effect of the *EPAS1* rs13419896 G > A on pulmonary function in the Guangzhou population was estimated by the Kruskal-Wallis test. **(A)** Pre-forced expiratory volume in 1s (pre-$FEV_1$) [GG, GA, AA: 1.98 L (1.55, 2.52), 1.99 L (1.55, 2.488), 2.12 L (1.79, 2.56); $P = 0.06$]**; (B)** Pre-forced vital capacity ratio (pre-FVC) [GG, GA, AA: 2.61 L (2.06, 3.34), 2.66 L (2.13, 3.27), 2.72 L (2.26, 3.46); $P = 0.145$]; **(C)** Pre-$FEV_1$/FVC [GG, GA, AA: 76.23% (67.86, 82.45), 76.27% (67.73, 83.00), 79.56% (71.39, 85.72); $P = 0.024$].

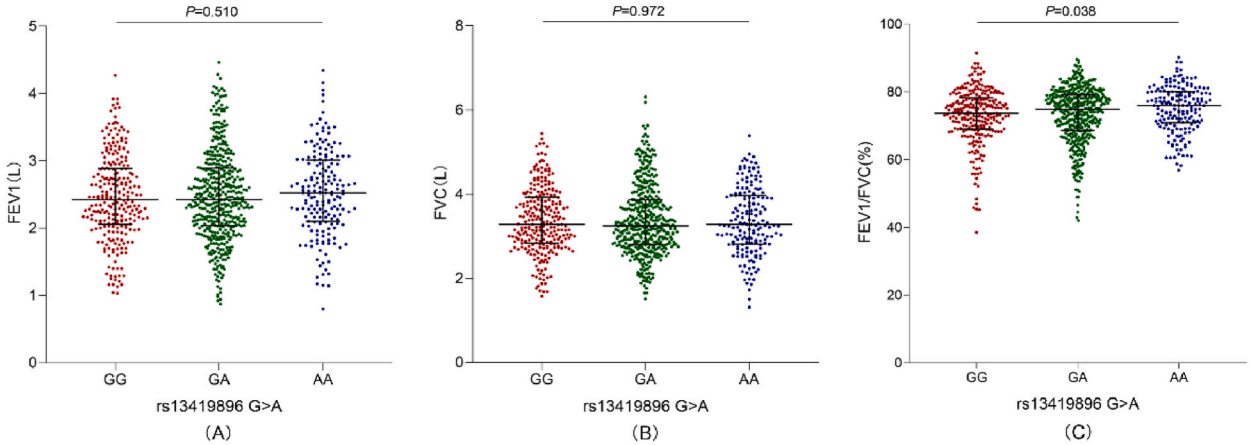


**Fig. 2.** The effect of the *EPAS1* rs13419896 G > A on pulmonary function in the Gansu population was estimated by the Kruskal-Wallis test. **(A)** Pre-forced expiratory volume in 1s (pre-$FEV_1$) [GG, GA, AA: 2.42 L (2.06, 2.89), 2.42 L (2.04, 2.89), 2.52 L (2.10, 3.01); $P = 0.510$]**; (B)** Pre-forced vital capacity ratio (pre-FVC) [GG, GA, AA: 3.28 L (2.84, 3.93), 3.25 L (2.80, 3.86), 3.28 L (2.81, 3.97); $P = 0.972$]; **(C)** Pre-$FEV_1$/FVC [GG, GA, AA: 73.72% (68.98, 78.29), 74.87% (68.57, 79.08), 76.00% (70.87, 80.00); $P = 0.038$].

Genomics Research Consortium (LTRC), Yoo et al. recently observed a decrease in *EPAS1* expression due to elevated methylation in COPD patients. At the same time, a similar effect was repeated in tobacco-treated mice. In addition, the motif enrichment analysis found that some downstream genes of *EPAS1* overlapped with multiple COPD-related genes, such as *VEGF*, *AP-1*, *C/EBPβ*, *XPB-1,* and *Mir-145* [14]. It was reported that these genes play critical roles in lung development, chronic inflammation, fibrosis, airway remodeling, and lung function degenerating [10,30–34]. A study with a small sample size (55 cases and 25 controls) also found increased expression of *EPAS1* in COPD [15]. The previous studies indicated that *EPAS1* might participate in the occurrence and development of COPD. However, the hypothesis has not reached a consistent conclusion.

Therefore, we conducted this study to elucidate the association between *EPAS1* polymorphisms and COPD susceptibility. As reported by previous studies, rs13419896 G > A and rs6756667 G > A were located in the transcription factor binding regions and could change the expression of *EPAS1*; rs59901247 A > C was an nsSNP that could cause amino acid substitution [20,21]. Initially, *EPAS1* attracted widespread attention for its important role in Tibetans' adaptation to hypoxia [35,36]. Subsequently, several studies reported that *EPAS1* polymorphisms modified the susceptibility to multiple acute altitude sicknesses when Han Chinese rapidly entered the plateau [20,37–40]. For example, the rs6756667 A allele reduced the risk of acute high-altitude headache and acute appetite loss [38, 40], while the rs13419896 A allele made Han Chinese susceptible to acute cardiorespiratory fitness impairment [37]. As the research continues to deepen, researchers also observed that the rs13419896 A allele is a risk factor for premature retinopathy, lung cancer, hepatitis B, and liver cirrhosis [21,41–43]. Exiting evidence suggested that *EPAS1* polymorphisms could indicate susceptibility to disorders, but their biological functions were quite heterogeneous in different ethnic and diseases.

**Table 3**
Estimated relative risks based on a case-control study.

| Factors | RR (95% CI)[a] |
|---|---|
| Male | |
| Smoking status | |
| No | 1.000 (ref.) |
| Yes | 1.905 (1.506–2.410) |
| Coal as fuels | |
| No | 1.000 (ref.) |
| Yes | 1.903 (1.330–2.722) |
| Education level | |
| High school or above | 1.000 (ref.) |
| Middle school or below | 1.641 (1.282–2.100) |
| rs13419896 | |
| AA | 1.000 (ref.) |
| GG + GA | 1.172 (0.786–1.747) |
| Population-attributable risk | 0.457 (0.283–0.632) |
| **Female** | |
| Smoking status | |
| No | 1.000 (ref.) |
| Yes | 1.800 (1.018–3.182) |
| Biomass as fuels | |
| No | 1.000 (ref.) |
| Yes | 2.515 (1.624–3.896) |
| rs13419896 | |
| AA | 1.000 (ref.) |
| GG + GA | 2.329 (1.340–4.047) |
| Population-attributable risk | 0.421 (0.227–0.616) |

RR, relative risk.
[a] Backward logistic regression, adjusting for age.

The current study suggested that the rs13419896 A allele was associated with decreased COPD susceptibility and better lung function. Some studies have explored how rs13419896 G > A regulated the expression of *EPAS1* [21,44]. Li et al. observed that the rs13419896 AA genotype carrier had the lowest expression of *EPAS1* than those with the GA or GG genotypes. However, the expression of *EPAS1* changed more dramatically when the former rapidly entered high-altitude areas [44]. Putra et al. reported that the rs13419896 A allele enhanced the expression of *EPAS1* in lung cancer and was associated with a poor prognosis [21]. Interestingly, in disagreement with Yoo et al. Putra et al. suggested that AP-1, C/EBPβ, and C-MYC could regulate the expression of *EPAS1* by binding to the region where the rs13419896 G > A located, raising a possibility that there might be bidirectional regulation between *EPAS1* and COPD-related genes [14,21]. Although the regulatory mechanism is not clear, we reasonably speculated that rs13419896 G > A affected the affability of transcription factors to *EPAS1*, thus changing the expression of *EPAS1*. In this way, rs13419896 G > A might participate in the development of COPD by the HIF signaling pathway [21,43].

The incidence of COPD and the mortality without COPD were different between males and females [28,45]. Thus, the 10-year absolute risk model was conducted by gender. However, coal as fuels and education level, the vital factors of COPD, did not remain in the female models, which might attribute to the lessened sample size after stratification [1,3]. In addition, the absence of coal as fuels might be partly interpreted by the correlation with biomass as fuels [24]. The population-attributable risks were 0.457 (0.283–0.632) for the male model and 0.421 (0.227–0.616) for the female model, indicating that more factors should be considered in models to capture much of COPD risk [46]. We also tried to remove the positive SNP from the male or female models for sensitive analysis. As shown in Fig. S2, the prediction accuracy declined to some degree after removing rs13419896 G > A, indicating an important contribution of rs13419896 G > A to the models.

However, there were some other limitations in this study. 1. As a hospital-based retrospective study, selection and recall biases were inevitable. 2. The sample size was considerably large. We had 0.8 power to detect variations with ORs>1.5. But most polymorphisms were low-risk (ORs<1.5), which needed a larger sample size to improve the detection efficiency and decrease the false-positive report probability, especially in the stratification analysis (Table S5) [47]. 3. We should have performed more experiments to clarify the function of *EPAS1*. 4. Our findings and models needed external validation in the other population.

## 5. Conclusion

*EPAS1* rs13419896 G > A significantly reduced COPD susceptibility in southern Chinese. The A allele was associated with better lung function. Combined with environmental factors, *EPAS1* rs13419896 G > A could be a genetic marker to predict the 10-year absolute risk for COPD.

## Funding

This work was supported by the following grants: National Key Research and Development Project of China [grant number

**Table 4**
The 10-year absolute risk for COPD in southern China.

| Initial Age (years) | r (individual relative risk) | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | male | | | | | | | | female | | | | | | | |
| | 0.0 | 1.0 | 2.0 | 3.0 | 4.0 | 5.0 | 6.0 | 7.0 | 0.0 | 2.0 | 4.0 | 6.0 | 8.0 | 10.0 | 12.0 | 14.0 |
| 40–49 | 0.000 | 0.074 | 0.143 | 0.207 | 0.265 | 0.320 | 0.370 | 0.417 | 0.000 | 0.077 | **0.149** | 0.214 | 0.275 | 0.331 | 0.382 | 0.430 |
| 50–59 | 0.000 | 0.145 | 0.268 | 0.374 | 0.464 | 0.540 | 0.605 | 0.661 | 0.000 | 0.140 | 0.261 | 0.364 | 0.453 | 0.529 | 0.595 | 0.651 |
| 60–69 | 0.000 | 0.255 | 0.442 | 0.581 | 0.682 | 0.758 | 0.813 | 0.855 | 0.000 | 0.224 | 0.397 | 0.531 | 0.634 | 0.714 | 0.776 | 0.823 |

2017YFC0907202]; National Natural Science Foundation of China [grant number 81872694, 82173609, 81460123, 82260889]; Local Innovative and Research Teams Project of Guangdong Pearl River Talents Program [grant number 2017BT01S155]; Guangdong Medical Science and Technology Research Fund Project [grant number A2020324]; Guangxi Natural Foundation [grant number 2018GXNSFAA281187]; the Science and Technology Project of Guangzhou [grant number 202103000073].

## Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Human/Animal ethics approval declaration

This study was performed in accordance with the Declaration of Helsinki and approbated by the institutional review boards of Xi'an Jiaotong University Health Science Center (approval: XJTU 2016-411) and Guangzhou Medical University (approval: GZMC2007-07-0676).

## Author contributions

Jiachun Lu, Xinhua Wang, Yunchao Wang conceived and designed the experiments; Ruiqi He, Yujie Pan, Chun Mao, Yunchao Wang performed the experiments; Ao Lin, Xiaobin Zeng analyzed and interpreted the data; Xinhua Wang, Jiachun Lu, Cuiyi Chen, Chenli Xie, Dongsheng Huang, Yibin Deng, Xuhui Zhang contributed reagents, materials, analysis tools or data; Ao Lin, Yunchao Wang wrote the paper. All authors have read and agreed to the published version of the manuscript.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

Not applicable.

## Appendix

The backward stepwise logistic regression model was used to screen risk factors and estimated the RR of each variate. Individual relative risk (r) was calculated by formula 1. Population-attributable risk (PAR) could be obtained by IRAP 2.2.0. $h_1$ was calculated by formula 2, where IR was the incidence of COPD. Formula 3 was used to predict the 10-year absolute risk (AR), where $h_2$ was the mortality without COPD and a was the initial age.

$$r = \Pi_{i=1}^{n} r_i \tag{1}$$

$$h_1 = IR\ (1 - PAR) \tag{2}$$

$$P(a, r) = \{h_1\ r/(h_1\ r + h_2)[1\text{-}exp\{\text{-}10(h_1\ r + h_2)\}] \tag{3}$$

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e20226.

## References

[1] J.B. Soriano, P.J. Kendrick, K.R. Paulson, V. Gupta, E.M. Abrams, R.A. Adedoyin, et al., Prevalence and attributable health burden of chronic respiratory diseases, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017, Lancet Respir. Med. 8 (6) (2020) 585–596.
[2] D. Adeloye, P. Song, Y. Zhu, H. Campbell, A. Sheikh, I. Rudan, Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis, Lancet Respir. Med. 10 (5) (2022) 447–458.
[3] C. Wang, J. Xu, L. Yang, Y. Xu, X. Zhang, C. Bai, et al., Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study, Lancet 391 (10131) (2018) 1706–1717.
[4] T. Welte, C. Vogelmeier, A. Papi, COPD: early diagnosis and treatment to slow disease progression, Int. J. Clin. Pract. 69 (3) (2015) 336–349.
[5] A.L. Durham, I.M. Adcock, The relationship between COPD and lung cancer, Lung Cancer 90 (2) (2015) 121–127.
[6] C. Xia, X. Dong, H. Li, M. Cao, D. Sun, S. He, et al., Cancer statistics in China and United States, 2022: profiles, trends, and determinants, Chin. Med. J. (Engl.) 135 (5) (2022) 584–590.

[7] C. Yuan, Lu G. Chang, X. Deng, Genetic polymorphism and chronic obstructive pulmonary disease, Int. J. Chronic Obstr. Pulm. Dis. 12 (2017) 1385–1393.
[8] D.G. Wang, J.B. Fan, C.J. Siao, A. Berno, P. Young, R. Sapolsky, et al., Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome, Science 280 (5366) (1998) 1077–1082.
[9] M.F. Ragland, C.J. Benway, S.M. Lutz, R.P. Bowler, J. Hecker, J.E. Hokanson, et al., Genetic advances in chronic obstructive pulmonary disease. Insights from COPDGene, Am. J. Respir. Crit. Care Med. 200 (6) (2019) 677–690.
[10] M. Ema, S. Taya, N. Yokotani, K. Sogawa, Y. Matsuda, Y. Fujii-Kuriyama, A novel bHLH-PAS factor with close sequence similarity to hypoxia-inducible factor 1alpha regulates the VEGF expression and is potentially involved in lung and vascular development, Proc. Natl. Acad. Sci. U.S.A. 94 (9) (1997) 4273–4278.
[11] S.A. Patel, M.C. Simon, Biology of hypoxia-inducible factor-2alpha in development and disease, Cell Death Differ. 15 (4) (2008) 628–634.
[12] M.C. van Patot, M. Gassmann, Hypoxia: adapting to high altitude by mutating EPAS-1, the gene encoding HIF-2alpha, High Alt. Med. Biol. 12 (2) (2011) 157–167.
[13] L. Davis, M. Recktenwald, E. Hutt, S. Fuller, M. Briggs, A. Goel, et al., Targeting HIF-2alpha in the tumor microenvironment: redefining the role of HIF-2alpha for solid cancer therapy, Cancers 14 (5) (2022) 1259.
[14] S. Yoo, S. Takikawa, P. Geraghty, C. Argmann, J. Campbell, L. Lin, et al., Integrative analysis of DNA methylation and gene expression data identifies EPAS1 as a key regulator of COPD, PLoS Genet. 11 (1) (2015), e1004898.
[15] H. Li, T. Yang, F.Y. Li, Q. Ning, Z.M. Sun, TLR4 overexpression inhibits endothelial PAS domain-containing protein 1 expression in the lower respiratory tract of patients with chronic COPD, Cell. Physiol. Biochem. 39 (2) (2016) 685–692.
[16] L. Yang, X. Lu, J. Deng, Y. Zhou, D. Huang, F. Qiu, et al., Risk factors shared by COPD and lung cancer and mediation effect of COPD: two center case-control studies, Cancer Causes Control 26 (1) (2015) 11–24.
[17] D. Singh, A. Agusti, A. Anzueto, P.J. Barnes, J. Bourbeau, B.R. Celli, et al., Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease: the GOLD science committee report 2019, Eur. Respir. J. 53 (5) (2019).
[18] A. Lin, M. Zhou, R.X. Hua, J. Zhang, H. Zhou, S. Li, et al., METTL3 polymorphisms and Wilms tumor susceptibility in Chinese children: a five-center case-control study, J. Gene Med. 22 (11) (2020), e3255.
[19] J. He, L. Yuan, H. Lin, A. Lin, H. Chen, A. Luo, et al., Genetic variants in m(6)A modification core genes are associated with glioma risk in Chinese children, Mol. Ther. Oncolytics 20 (2021) 199–208.
[20] L.I. Guo, J. Zhang, J. Jin, X. Gao, J. Yu, Q. Geng, et al., Genetic variants of endothelial PAS domain protein 1 are associated with susceptibility to acute mountain sickness in individuals unaccustomed to high altitude: a nested case-control study, Exp. Ther. Med. 10 (3) (2015) 907–914.
[21] A.C. Putra, H. Eguchi, K.L. Lee, Y. Yamane, E. Gustine, T. Isobe, et al., The A allele at rs13419896 of EPAS1 is associated with enhanced expression and poor prognosis for non-small cell lung cancer, PLoS One 10 (8) (2015), e0134496.
[22] J. He, Y. Zou, X. Liu, J. Zhu, J. Zhang, R. Zhang, et al., Association of common genetic variants in pre-microRNAs and neuroblastoma susceptibility: a two-center study in Chinese children, Mol. Ther. Nucleic Acids 11 (2018) 1–8.
[23] S. Wacholder, S. Chanock, M. Garcia-Closas, L. El Ghormli, N. Rothman, Assessing the probability that a positive report is false: an approach for molecular epidemiology studies, J. Natl. Cancer Inst. 96 (6) (2004) 434–442.
[24] T.R. Fears, D.t. Guerry, R.M. Pfeiffer, R.W. Sagebiel, D.E. Elder, A. Halpern, et al., Identifying individuals at high risk of melanoma: a practical predictor of absolute risk, J. Clin. Oncol. 24 (22) (2006) 3590–3596.
[25] J. Xin, M. Du, D. Gu, Y. Ge, S. Li, H. Chu, et al., Combinations of single nucleotide polymorphisms identified in genome-wide association studies determine risk for colorectal cancer, Int. J. Cancer 145 (10) (2019) 2661–2669.
[26] M.H. Gail, L.A. Brinton, D.P. Byar, D.K. Corle, S.B. Green, C. Schairer, et al., Projecting individualized probabilities of developing breast cancer for white females who are being examined annually, J. Natl. Cancer Inst. 81 (24) (1989) 1879–1886.
[27] P. Bruzzi, S.B. Green, D.P. Byar, L.A. Brinton, C. Schairer, Estimating the population attributable risk for multiple risk factors using case-control data, Am. J. Epidemiol. 122 (5) (1985) 904–914.
[28] H.Z. Liu, G.Z. Lin, Annual Report on the Cause of Death Surveillance in Guangzhou, 2014-2015, Yangcheng Evening News Publishers, Guangzhou, China, 2018.
[29] M.E. Pamenter, J.E. Hall, Y. Tanabe, T.S. Simonson, Cross-species insights into genomic adaptations to hypoxia, Front. Genet. 11 (2020) 743.
[30] Y.H. Tu, Y. Guo, S. Ji, J.L. Shen, G.H. Fei, The influenza A virus H3N2 triggers the hypersusceptibility of airway inflammatory response via activating the lncRNA TUG1/miR-145-5p/NF-kappaB pathway in COPD, Front. Pharmacol. 12 (2021), 604590.
[31] W. Gu, Y. Yuan, L. Wang, H. Yang, S. Li, Z. Tang, et al., Long non-coding RNA TUG1 promotes airway remodelling by suppressing the miR-145-5p/DUSP6 axis in cigarette smoke-induced COPD, J. Cell Mol. Med. 23 (11) (2019) 7200–7209.
[32] A. Birnhuber, V. Biasin, D. Schnoegl, L.M. Marsh, G. Kwapiszewska, Transcription factor Fra-2 and its emerging role in matrix deposition, proliferation and inflammation in chronic lung diseases, Cell. Signal. 64 (2019), 109408.
[33] M. Mori, L. Bjermer, J.S. Erjefalt, M.R. Stampfli, A.B. Roos, Small airway epithelial-C/EBPbeta is increased in patients with advanced COPD, Respir. Res. 16 (2015) 133.
[34] E.J. Becker, A. Faiz, M. van den Berge, W. Timens, P.S. Hiemstra, K. Clark, et al., Bronchial gene expression signature associated with rate of subsequent FEV1 decline in individuals with and at risk of COPD, Thorax 77 (1) (2022) 31–39.
[35] C.M. Beall, G.L. Cavalleri, L. Deng, R.C. Elston, Y. Gao, J. Knight, et al., Natural selection on EPAS1 (HIF2alpha) associated with low hemoglobin concentration in Tibetan highlanders, Proc. Natl. Acad. Sci. U.S.A. 107 (25) (2010) 11459–11464.
[36] Z. Basang, B. Wang, L. Li, L. Yang, L. Liu, C. Cui, et al., HIF2A variants were associated with different levels of high-altitude hypoxia among native Tibetans, PLoS One 10 (9) (2015), e0137956.
[37] J. Yang, H. Tan, M. Sun, R. Chen, J. Zhang, C. Liu, et al., Prediction of high-altitude cardiorespiratory fitness impairment using a combination of physiological parameters during exercise at sea level and genetic information in an integrated risk model, Front. Cardiovasc. Med. 8 (2021), 719776.
[38] Y. Shen, J. Zhang, J. Yang, C. Liu, S. Bian, C. Zhang, et al., Association of EPAS1 and PPARA gene polymorphisms with high-altitude headache in Chinese han population, BioMed Res. Int. 2020 (2020), 1593068.
[39] J.H. Zhang, Y. Shen, C. Liu, J. Yang, Y.Q. Yang, C. Zhang, et al., EPAS1 and VEGFA gene variants are related to the symptoms of acute mountain sickness in Chinese Han population: a cross-sectional study, Mil. Med. Res. 7 (1) (2020) 35.
[40] W. Pan, C. Liu, J. Zhang, X. Gao, S. Yu, H. Tan, et al., Association between single nucleotide polymorphisms in PPARA and EPAS1 genes and high-altitude appetite loss in Chinese young men, Front. Physiol. 10 (2019) 59.
[41] S. Ilguy, O. Cilingir, M.D. Bilgec, O. Ozalp, E. Erzurumluoglu Gokalp, S. Arslan, et al., The relationship of retinopathy of prematurity with brain-derivated neurotrophic factor, vascular endotelial growth factor-A, endothelial PAD domain protein 1 and nitric oxide synthase 3 gene polymorphisms, Ophthalmic Genet. 42 (6) (2021) 725–731.
[42] Y. Yamamoto, C. Kiyohara, S. Ogata-Suetsugu, N. Hamada, Y. Nakanishi, Association between genetic polymorphisms involved in the hypoxia-inducible factor pathway and lung cancer risk: a case-control study in Japan, Asia Pac. J. Clin. Oncol. 13 (3) (2017) 234–242.
[43] L. Huang, C. Liu, Y. Deng, Y. Liu, J. Zhao, X. Huang, et al., Association of hypoxia-inducible factor-2 alpha gene polymorphisms with the risk of hepatitis B virus-related liver disease in Guangxi Chinese: a case-control study, PLoS One 11 (7) (2016), e0158241.
[44] C. Li, X. Li, J. Xiao, J. Liu, X. Fan, F. Fan, et al., Genetic changes in the EPAS1 gene between Tibetan and Han ethnic groups and adaptation to the plateau hypoxic environment, PeerJ 7 (2019), e7943.
[45] L.A. Garcia Rodriguez, M.A. Wallander, L.B. Tolosa, S. Johansson, Chronic obstructive pulmonary disease in UK primary care: incidence and risk factors, COPD 6 (5) (2009) 369–379.
[46] P. Burney, J. Patel, C. Minelli, L. Gnatiuc, A.F.S. Amaral, A. Kocabas, et al., Prevalence and population-attributable risk for chronic airflow obstruction in a large multinational study, Am. J. Respir. Crit. Care Med. 203 (11) (2021) 1353–1365.
[47] J. Yang, B. Benyamin, B.P. McEvoy, S. Gordon, A.K. Henders, D.R. Nyholt, et al., Common SNPs explain a large proportion of the heritability for human height, Nat. Genet. 42 (7) (2010) 565–569.