INTRODUCTION

Colorectal cancer (CRC) is a cancer that occurs in the colon or rectum and can occur in any part of the colon, with the rectum and sigmoid colon being the most common. According to the latest global cancer statistics, the incidence (19.5 people per 100,000 people) and mortality (9.0 people per 100,000 people) of CRC rank third and second, respectively. In China, the incidence of CRC increased rapidly, and...
the incidence and mortality of CRC rank second and fifth, respectively. The occurrence of CRC is a multi-factor and multi-step process, which is the result of the interaction between genetic factors and environmental factors. Alcohol is a recognized risk factor for CRC. Drinking alcohol increases the risk of CRC. In terms of mechanism, acetaldehyde is a metabolite of ethanol. Acetaldehyde can cause DNA damage, hinder DNA synthesis and repair, disrupt DNA methylation, and induce inflammation and oxidative stress, leading to lipid peroxidation and further DNA damage. Acetaldehyde can bind glutathione to inhibit the antioxidant defense system and induce tumor formation.

In terms of genetic factors, mismatch repair (MMR) genes, and APC, MUTYH, POLE and POLD1 genes were identified as high-penetrant genes. Results of a meta-analysis of genome-wide association studies (GWASs) in CRC showed that L1TD1, EFCAB2, PPP1R21, SLC02A1, HLA-G, NOTCH4, DENND5B, GNAS, ALDH7A1, PRICKLE1, KL5, WWOX, and GLP2R genes were significantly associated with risk of CRC. In addition, at least 50 low-risk sites for CRC were identified in additional genome-wide association study. A GWAS study suggested that acetaldehyde dehydrogenase 2 (ALDH2) gene is associated with the susceptibility of upper gastrointestinal cancer. ALDH2 belongs to the aldehyde dehydrogenase family and is responsible for metabolizing acetaldehyde into acetic acid. ALDH2 activity level is affected by ALDH2 gene mutations. In the most common ALDH2 SNP rs671 (G>A), glutamate (Glu) becomes Lysine (Lys) at position 504 in the amino acid sequence of the protein encoded by ALDH2, which ALDH2 activity will be greatly affected. The wild-type ALDH2 rs671 G/G has normal catalytic activity, while the enzyme activity of ALDH2 G/A is only 13%-14% of that of the wild-type. ALDH2 A/A basically has no enzyme activity, but the frequency of ALDH2 A/A in Asian people is up to 40%.

Studies have showed that patients with ALDH2 rs671 polymorphism has an increased risk of some cancers, such as esophageal cancer, gastric cancer and colorectal cancer. Studies in different populations have shown inconsistent results. It is of great significance to explore the characteristics and genetic differences between cancer patients and cancer-free people. The Hakka is a Han ethnic group with a unique genetic background formed by the Hakka ancestors from the Han nationality in central China, who migrated southward for many times and fused with the ancient Yue residents in Guangdong, Fujian and Jiangxi. Meizhou is a city located in the northeast of Guangdong Province, is overwhelmingly populated by Hakka people. In this study, the relationship between ALDH2 rs671 and CRC was analyzed in Hakka Chinese population.

2 | MATERIALS AND METHODS

2.1 | Study cohort

This study cohort consisted of 178 CRC patients and 261 controls from Meizhou People’s Hospital, China, between January 2016 and December 2020. The inclusion criteria of patients were: (1) patients diagnosed with histologically confirmed colorectal cancer; (2) patients without other tumors and serious infectious diseases; (3) patients without missing information; (4) ≥18 years old. The inclusion criteria of controls were: (1) from the physical examination Center of Meizhou People’s Hospital and did not develop CRC until 2020; (2) had no history of other tumors of the digestive system or end-stage kidney diseases; (3) ≥18 years old. The subjects included in this study are all Hakka people. This retrospective case-control study was approved by the Human Ethics Committees of Meizhou People’s Hospital. The flow chart of the present study is shown in Figure 1.

2.2 | ALDH2 genotyping

Two milliliters of venous blood from each subject was stored into tube containing EDTA, genomic DNA was extracted from whole blood using a QIAamp DNA Blood Mini Kit (Qiagen GmbH). DNA concentration was measured using a Nanodrop 2000 Spectrophotometer (ThermoFisher Scientific). ALDH2 genotyping was performed by polymerase chain reaction (PCR)-gene chip method (BaiO Technology Co, Ltd.).

2.3 | Covariates

The gender and age of the subjects were collected from medical records. The data of smoking and drinking history were collected through medical records. Smokers include current smokers and ever smokers. Drinkers include current drinkers and ever drinkers. In addition, blood cell parameters of study patients at the time of diagnosis of CRC were collected through medical record data. Blood cell parameters of the controls were collected at the time of the first hospital examination. The inflammation index were calculated according to the following formula: neutrophil-to-lymphocyte ratio (NLR) = neutrophil/lymphocyte, platelet-to-lymphocyte ratio (PLR) = platelet/lymphocyte, lymphocyte-to-monocyte ratio (LMR) = lymphocyte/monocyte.

2.4 | Statistical analysis

Data analysis was performed using SPSS 21.0 (IBM Inc.). Student’s t test or the Mann–Whitney U test was used for continuous data analysis. Genotype composition ratios and allele frequencies between groups were analyzed by the χ² test.

It is found that malignant tumor is a complex disease caused by genetic factors and environmental factors. The aim of our study was to determine the association between ALDH2 rs671 and CRC risk after adjusting for other major influencing factors, such as sex, age, smoking history, alcohol consumption, and some blood cell parameters. The optimal cut-off value of continuous data (such as blood cell parameters) for CRC risk prediction was determined by receiver operating characteristic (ROC) curve analysis. All subjects
were grouped according to each parameter greater than, less than or equal to the cut-off value, respectively. Age was divided into old group (>65 years old) and non-old group (≤65 years old). Categorical variables are grouped according to different types, respectively. Logistic regression analysis was applied to assess the interactions between ALDH2 rs671 polymorphism and these covariates in risk assessment of CRC.

3 | RESULTS

3.1 | ALDH2 genotype of the subjects

The study analyzed ALDH2 gene mutations and clinical data from 439 participants, including 178 CRC patients and 261 controls. The distribution of ALDH2 genotypes in CRC patients \( (\chi^2 = 0.802, p = 0.370) \) and controls \( (\chi^2 = 0.267, p = 0.605) \) was consistent with Hardy–Weinberg equilibrium, respectively. The proportion of the ALDH2 rs671 G/G, G/A, and A/A genotype in patients was 48.3%, 44.4%, and 7.3%, respectively. And the proportion of the ALDH2 rs671 G/G, G/A, and A/A genotype in controls was 62.1%, 34.1%, and 3.8%, respectively. The frequencies of G and A alleles were 70.5% and 29.5% in CRC patients, 79.1% and 20.9% in controls, respectively. There was statistical difference in genotype distribution \( (p = 0.011) \) and allele distribution \( (p = 0.004) \) between patients and controls (Table 1).

3.2 | Comparison of characteristics of CRC patients and controls CRC patients grouped by ALDH2 variation

The average age of CRC patients and controls was 68.55±12.98 and 68.99±12.40 years, respectively. The patients had higher proportion of smoking history (27.0% vs. 18.0%, \( p = 0.033 \)), and higher proportion of history of alcohol drinking (14.0% vs. 7.3%, \( p = 0.024 \)), higher level of NLR (7.31±4.15 vs. 5.16±4.15, \( p = 0.009 \)), platelet count (254.67±110.55 vs. 220.05±80.36×10^9/L, \( p < 0.001 \)), and PLR (218.23±163.22 vs. 156.58±100.24, \( p < 0.001 \)) than controls. The patients had lower level of lymphocyte count \( (1.49±1.01×10^9/L, p = 0.001) \), LMR \( (3.15±2.34 vs. 3.95±3.64, p = 0.006) \), and mean hemoglobin concentration \( (323.14±18.92 vs. 330.30±14.47 g/L, p < 0.001) \) than controls. There were no statistically significant differences in the neutrophil count, monocyte
count, platelet distribution width, red cell count, and red cell distribution width (Table 2).

### 3.3 Comparison of characteristics CRC patients grouped by ALDH2 variants

Among CRC patients with ALDH2 rs671 G/G, G/A, and A/A genotype, the proportion of patients with ever or current history of alcohol drinking was 22.1%, 6.3% and 7.7%, respectively, with statistically significant differences (p = 0.011). There was statistically significant differences of neutrophil count (6.51 ± 3.74, 7.57 ± 4.85, and 10.37 ± 7.67 × 10⁹/L, respectively) (p = 0.016) and NLR (5.63 ± 4.65, 8.00 ± 12.72, and 14.21 ± 17.74, respectively) (p = 0.015) levels in CRC patients with G/G, G/A, and A/A genotype. No statistically significant differences were observed in the percentage of history of alcohol drinking and history of smoking, and the levels of blood cell parameters between patients with G and A allele, respectively (Table 3).

### 3.4 Impact of ALDH2 rs671 polymorphism on CRC risk

The optimal cut-off value of blood cell parameters for CRC risk prediction was determined by receiver operating characteristic (ROC) curve analysis. When CRC was taken as the endpoint, the critical value of lymphocyte count was 1.450 (×10⁹/L) (sensitivity 56.7%, specificity 80.5%), the PLR cutoff value was 3.535 (sensitivity 71.3%, specificity 58.2%), the NLR cutoff value was 8.741 (sensitivity 73.2%, specificity 58.2%), and the mean hemoglobin concentration cutoff value was 32.25 (g/L) (sensitivity 74.3%, specificity 74.3%).

Logistic regression analysis indicated that there was significantly high risk of CRC in the presence of history of smoking (Yes vs. No) (p = 0.026), and history of alcohol drinking (Yes vs. No) (p = 0.023). However, history of smoking and alcohol consumption were not risk factors for colorectal cancer after adjustment for gender (Male vs. Female), age (>65 vs. ≤65 years old), blood cell parameters, and ALDH2 genotype. Logistic regression analysis showed that there was significantly high risk of CRC in the presence of high.
### Table 3: Clinical characteristics of cases stratified by ALDH2 variants

| Clinical characteristics | G/G (n = 86) | G/A (n = 79) | A/A (n = 13) | p Values | G allele (G/G + G/A) (n = 165) | A allele (G/A + A/A) (n = 92) | p Values |
|--------------------------|-------------|-------------|-------------|---------|-------------------------------|-------------------------------|---------|
| Age (years)              |             |             |             |         |                               |                               |         |
| <60, n (%)               | 20 (23.3)   | 16 (20.3)   | 0 (0.0)     | 0.207   | 36 (21.8)                     | 16 (17.4)                     | 0.570   |
| 60–75, n (%)             | 44 (51.2)   | 35 (44.3)   | 8 (61.5)    |         | 79 (47.9)                     | 43 (46.7)                     |         |
| >75, n (%)               | 22 (25.6)   | 28 (35.4)   | 5 (38.5)    |         | 50 (30.3)                     | 33 (35.9)                     |         |
| Gender                   |             |             |             |         |                               |                               |         |
| Male, n (%)              | 65 (75.6)   | 63 (79.7)   | 6 (46.2)    | 0.043   | 128 (77.6)                    | 69 (75.0)                     | 0.647   |
| Female, n (%)            | 21 (24.4)   | 16 (20.3)   | 7 (53.8)    |         | 37 (22.4)                     | 23 (25.0)                     |         |
| History of smoking       |             |             |             |         |                               |                               |         |
| Never                    | 59 (68.6)   | 60 (75.9)   | 11 (84.6)   | 0.392   | 119 (72.1)                    | 71 (77.2)                     | 0.459   |
| Ever or Current          | 27 (31.4)   | 19 (24.1)   | 2 (15.4)    |         | 46 (27.9)                     | 21 (22.8)                     |         |
| History of alcohol drinking |         |             |             |         |                               |                               |         |
| Never                    | 67 (77.9)   | 74 (93.7)   | 12 (92.3)   | 0.011   | 141 (85.5)                    | 86 (93.5)                     | 0.068   |
| Ever or Current          | 19 (22.1)   | 5 (6.3)     | 1 (7.7)     |         | 24 (14.5)                     | 6 (6.5)                       |         |
| Neutrophil count, ×10⁹/L | 6.51 ± 3.74 | 7.57 ± 4.85 | 10.37 ± 7.67 | 0.016   | 7.02 ± 4.33                   | 7.97 ± 5.37                   | 0.148   |
| Monocyte count, ×10⁹/L   | 0.57 ± 0.28 | 0.60 ± 0.31 | 0.55 ± 0.38 | 0.776   | 0.59 ± 0.29                   | 0.59 ± 0.32                   | 0.836   |
| Lymphocyte count, ×10⁹/L | 1.54 ± 0.95 | 1.46 ± 0.65 | 1.28 ± 0.58 | 0.494   | 1.50 ± 0.82                   | 1.43 ± 0.64                   | 0.471   |
| Neutrophil to lymphocyte ratio (NLR) | 5.63 ± 4.65 | 8.00 ± 12.72 | 14.21 ± 17.74 | 0.015   | 6.76 ± 9.47                   | 8.88 ± 13.60                  | 0.188   |
| Lymphocyte to monocyte ratio (LMR) | 3.04 ± 1.74 | 3.24 ± 2.86 | 3.36 ± 2.46 | 0.815   | 3.14 ± 2.34                   | 3.26 ± 2.80                   | 0.710   |
| Platelet count, ×10⁹/L   | 256.15 ± 114.09 | 254.68 ± 111.03 | 244.77 ± 88.45 | 0.943   | 255.45 ± 112.29               | 253.28 ± 107.75               | 0.881   |
| Platelet to lymphocyte ratio (PLR) | 210.54 ± 152.55 | 221.81 ± 176.97 | 247.40 ± 151.97 | 0.727   | 215.94 ± 164.28               | 225.43 ± 173.12               | 0.664   |
| Platelet distribution width | 11.74 ± 3.32 | 11.71 ± 3.44 | 11.98 ± 2.39 | 0.962   | 11.73 ± 3.37                  | 11.75 ± 3.30                  | 0.963   |
| Red cell count, ×10¹²/L  | 4.38 ± 0.81  | 4.40 ± 0.89  | 4.10 ± 1.03  | 0.502   | 4.39 ± 0.85                   | 4.35 ± 0.91                   | 0.782   |
| Red cell distribution width | 45.84 ± 7.13 | 46.24 ± 6.62 | 45.92 ± 6.82 | 0.930   | 46.03 ± 6.87                  | 46.20 ± 6.61                  | 0.851   |
| Mean hemoglobin concentration, g/L | 324.66 ± 18.55 | 321.66 ± 19.63 | 322.08 ± 17.48 | 0.585   | 323.22 ± 19.08                | 321.72 ± 19.25                | 0.546   |

Note: *p* < 0.05 was considered statistically significant.
platelet count (>283.5 vs. ≤283.5, ×10⁹/L) (adjusted OR 2.063, 95% CI 1.214–3.507, \( p = 0.007 \)), and low mean hemoglobin concentration (≤322.5 vs. >322.5, g/L) (adjusted OR 1.600, 95% CI 1.030–2.484, \( p = 0.036 \)).

The association between ALDH2 rs671 genotypes and CRC may be based on three genetic inheritance patterns, including co-dominant mode (G/A vs. G/G, A/A vs. G/G), dominant mode (G/A + A/A vs. G/G), and recessive mode (A/A vs. G/G + G/A). The ALDH2 G/A genotype in co-dominant model (adjusted OR 1.801, 95% CI 1.160–2.794, \( p = 0.009 \)) and A/A genotype in co-dominant model (adjusted OR 2.630, 95% CI 1.041–6.645, \( p = 0.041 \)) were significant risk factors for the presence of CRC. The ALDH2 G/A + A/A genotypes in dominant model (adjusted OR 1.883, 95% CI 1.230–2.881, \( p = 0.004 \)) was risk factor for the presence of CRC (Table 4).

### 4 | DISCUSSION

It is found that malignant tumor is a complex disease caused by genetic factors and environmental factors.²⁴²⁵ CRC is one of the most common malignant tumors of gastrointestinal tract. Direct nodular carcinoma can be effectively diagnosed and evaluated by histopathological and imaging examination.²⁶ However, with the increasing incidence and mortality of CRC, many scholars are devoted to the study of the prevention and risk prediction markers and mechanisms of CRC. With the development of molecular biology, some progress has been made in the study of molecular markers for the risk prediction of CRC, which has important clinical value for the prediction of CRC. China is a multi-ethnic country with a large population, and there are genetic differences among different populations.²⁷²⁸ There are significant regional differences in the incidence of some cancers in our country.²⁹³⁰ Hakka population is one of the Chinese Han populations. Since the Song Dynasty, the Hakka ancestors who migrated southward from the central China began to live in the Hakka area in a relatively stable way, and merged with the local indigenous peoples and became Hakka.²³ Nowadays, most Hakka persons are living in the northeastern part of the Guangdong Province.³¹ Meizhou is a city located in the northeast of Guangdong Province, is overwhelmingly populated by Hakka people. The relationship of ALDH2 rs671 variants and CRC in Hakka population is not well understood. This study provides a retrospective analysis for the results of this relationship among Hakka population.

In this case–control study, the proportion of subjects with drinking and smoking history in the patients was higher than that in the controls. Overall, individuals in the control group had healthier lifestyles than those in patients. Unadjusted logistic regression analysis showed that smoking and alcohol consumption were risk factors for CRC, respectively. But logistic regression analysis after adjustment for covariates showed no significant relationship between smoking, alcohol consumption and CRC risk (Table 4). There are a number of

### TABLE 4 Logistic regression analysis of risk factors associated with colorectal cancer

| Variables | Genotypes | Unadjusted values | Adjusted values |
|-----------|-----------|-------------------|-----------------|
| OR (95% CI) | \( p \) Value | OR (95% CI) | \( p \) Value |
| Age (>65/≤65, years) | 1.161 (0.783–1.722) | 0.457 | 0.915 (0.586–1.428) | 0.696 |
| Gender (Male/Female) | 1.073 (0.692–1.664) | 0.753 | 1.042 (0.619–1.755) | 0.878 |
| History of smoking (Yes/No) | 1.681 (1.064–2.656) | 0.026 | 1.330 (0.729–2.426) | 0.352 |
| History of alcohol drinking (Yes/No) | 2.081 (1.108–3.907) | 0.023 | 1.927 (0.866–4.290) | 0.108 |
| Lymphocyte count (≤322.5/>322.5, g/L) | 1.829 (1.244–2.689) | 0.002 | 1.251 (0.729–2.146) | 0.417 |
| NLR (>8.741/≤8.741) | 1.928 (1.183–3.143) | 0.008 | 1.351 (0.743–2.456) | 0.325 |
| LMR (≤3.535/>3.535) | 1.938 (1.293–2.906) | 0.001 | 1.535 (0.959–2.458) | 0.074 |
| Platelet count (>283.5/≤283.5, ×10⁹/L) | 2.256 (1.462–3.480) | <0.001 | 2.063 (1.214–3.507) | 0.007 |
| PLR (>189.0/≤189.0) | 2.207 (1.470–3.314) | <0.001 | 1.170 (0.650–2.104) | 0.601 |
| Mean hemoglobin concentration (≤322.5/>322.5, g/L) | 1.811 (1.206–2.718) | 0.004 | 1.600 (1.030–2.484) | 0.036 |

Genetic model of ALDH2 gene

| \( p \) Value | Value Adjusted OR (95% CI) |
| G/G | 1.000 (reference) |
| G/A | 1.672 (1.121–2.495) | 0.012 | 1.801 (1.160–2.794) | 0.009 |
| A/A | 2.449 (1.031–5.815) | 0.042 | 2.630 (1.041–6.645) | 0.041 |
| G/G | 1.000 (reference) |
| G/A + A/A | 1.751 (1.190–2.575) | 0.004 | 1.883 (1.230–2.881) | 0.004 |
| G/G + G/A | 1.000 (reference) |
| A/A | 1.978 (0.847–4.615) | 0.115 | 2.003 (0.813–4.931) | 0.131 |

Abbreviations: CI, confidence interval; OR, odds ratio.
studies support that smoking and alcohol consumption are associated with an increased risk of CRC in both men and women, and the risk increases with the duration and amount of smoking and the amount of alcohol consumed. A study on the etiology of CRC in China found that the risks of CRC in China include: smoking, alcohol consumption, obesity, physical inactivity, low vegetable intake, low fruit intake, and high red and processed meat intake. Since our study was based on a retrospective study of hospital patients, we did not collect information on the lifestyle and dietary habits of participants other than smoking and drinking. This is one of the limitations of this study.

Blood cell parameters reflect the formation and survival of blood cells, and reflect some metabolic abnormalities of the body, such as oxidative stress, inflammation, malnutrition, dyslipidemia, hypertension changes. Abnormal blood cell parameters are associated with a variety of diseases. Red cell distribution width (RDW) was statistically different in patients with cancer and in those without cancer. The study by Spell et al. found that 84% of patients with right-sided colon cancer had elevated RDW. RDW is also correlated with pathological features of breast cancer patients. Koma et al. found that high RDW levels were associated with lung cancer stage, and increased RDW level was associated with worse prognosis. Among multiple myeloma patients, patients with elevated RDW have a worse prognosis. The relationship of inflammation-related blood cell biomarkers has been reported in many types of cancer, including NLR, PLR, LMR, and so on. But the relationship between these markers and cancer risk has been relatively little studied. Elevated NLR is associated with increased risk of colorectal adenoma. In addition, fibrinogen to pre-albumin ratio is a biomarker for diagnosis of CRC. Systemic inflammatory cell ratios could be useful in early diagnosis of CRC. NLR and PLR are useful markers in diagnostic and early recognition of different stages of CRC. Monocyte-to-lymphocyte ratio (MLR), NLR and PLR are valuable predictive markers of colorectal cancer. Zhu et al. found that elevated platelet count, and platelet distribution width levels might serve as potential biomarkers for the diagnosis of CRC. PLR and mean platelet volume can predict the risk of CRC. In this study, the patients had higher level of NLR, platelet count, and PLR and lower level of LMR than controls. Our study supports that these inflammation-related blood cell biomarkers may be of value in the diagnosis of CRC.

In this study, the proportion of persons with ever or current history of alcohol consumption among persons with ALDH2 rs671 G/G, G/A, and A/A genotype was statistically significant difference. Participants with the ALDH2 rs671 G/G genotype had a higher proportion of alcohol consumption. It may be due to the fact that ALDH2 rs671 G/G genotype encodes highly active ALDH2, which metabolizes acetaldehyde quickly in vivo and is not easy to produce post-drinking reaction. Therefore, individuals carrying this genotype are prone to drinking alcohol. On the contrary, ALDH2 encoded by G/A and A/A genotype have relatively low activity, slow metabolism of acetaldehyde in the body, and prone to blush reaction, resulting in individuals carrying this genotype drinking less or even not drinking, and have certain resistance to drinking behavior. In this retrospective study, information on the frequency and amount of alcohol consumption of subjects were not collected. This is another limitation of this study.

It has been reported that ALDH2 rs671 polymorphism may be related to the susceptibility of some cancers, such as CRC, esophageal cancer, and liver cancer. A meta-analysis showed that the ALDH2 rs671 significantly increases the risk of CRC in East Asians. In individuals with ALDH2 rs671 A allele, the risk of CRC was higher on the hypomethylated diet than on the hypermethylated diet, suggesting that the hypomethylated diet and the ALDH2 rs671 A allele are risk factors for CRC. ALDH2 rs671 polymorphism increase the susceptibility to CRC. In contrast, a Korean study showed that men with the ALDH2 rs671 G/A or A/A genotype had a significantly lower risk of CRC than those with the ALDH2 G/G genotype, but there was no association among women. A study from China and a study in Japanese agree with the study in Korean. However, some studies in Japanese showed no association between ALDH2 rs671 polymorphism and the risk of CRC. A study from China also showed no association between ALDH2 rs671 and CRC. It showed that studies in different populations have shown inconsistent results. It may be important to explore the relationship between ALDH2 rs671 and CRC based on a large population of multicenter studies.

More and more attention has been paid to the biological characteristics of ALDH2. More attention should be paid to the upstream and downstream molecules of ALDH2 and their related pathways as well as their mechanism and significance in the process of tumor genesis and development. Mechanistically, PI3K/AKT/mTOR and MEK/ERK signaling pathways are involved in the regulation of ALDH2, and ALDH2 is also associated with DNA damage repair, autophagy, and immune system dysfunction. Acetaldehyde is a metabolic product of ethanol, it induces DNA damage and genome instability. Accumulation of acetaldehyde due to alcohol consumption or ALDH2 deficiency increases the risks of various types of cancers. Moreover, ALDH2 is linked to autophagy regulation in some diseases, and it was speculated that autophagy is one of the mechanisms that ALDH2 uses to regulate tumor occurrence and development. In addition, ALDH2 indirectly regulates the immune system due to its role in aldehydes metabolism and acetaldehyde adducts, and it may also be one of the mechanisms by which ALDH2 is involved in tumor development.

5 | LIMITATIONS

This research has some limitations. First, the association between this polymorphism and the clinicopathologic features of CRC patients (tumor site, tumor maximum diameter, TNM stage, total clinical stage, et al) was not investigated in this study because some medical records of some CRC patients were incomplete. Second, the association between the most common polymorphism of ALDH2 gene (ALDH2 rs671) and CRC was analyzed, but this study did not investigate the relationship between the full-length variation of
ALDH2 gene, gene expression level and the risk of CRC. Third, this study lacks external validation by external data sources from other places as validation datasets. External validation is needed to better assess the predictive power of the ALDH2 rs671 A allele for CRC risk. Hence, future studies with larger sample sizes, inclusion of more polymorphisms, and comprehensive analyses with external data are needed to investigate this relationship.

6 | CONCLUSION

Individuals carrying ALDH2 rs671 A allele may be at increased risk of colorectal cancer among Hakka population. Our study is the first report of this population and is a valuable addition to data on the role of ALDH2 polymorphism in diseases.

AUTHOR CONTRIBUTIONS

Yijin Chen designed the study. Zhuoxin Zhang, Yijin Chen, Qingqing Zhuo, Changqing Deng, Yang Yang, Wen Luo, and Shixun Lai collected clinical data. Zhuoxin Zhang, Yijin Chen, and Hui Rao analyzed the data. Yijin Chen prepared the manuscript. All authors were responsible for critical revisions, and all authors read and approved the final version of this work.

ACKNOWLEDGEMENTS

This work was supported by the Guangdong Provincial Key Laboratory of Precision Medicine and Clinical Translation Research of Hakka Population (Grant No.: 2018B030322003), Science and Technology Program of Meizhou (Grant No.: 2019B0202001).

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

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

ORCID

Yijin Chen https://orcid.org/0000-0003-0218-7115

REFERENCES

1. Birgisson H, Olausdottir EJ, Sverrisdottir A, Einarsson S, Smaradottir A, Tryggvadottir L. Screening for cancer of the colon and rectum a review on incidence, mortality, cost and benefit. Laeknabladid. 2021;107(9):398-405.
2. Sung H, Ferlay J, Siegel RL. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209-249.
3. Cao W, Chen HD, Yu YW, Li N, Chen WQ. Changing profiles of cancer burden worldwide and in China: a secondary analysis of the global cancer statistics 2020. Chin Med J (Engl). 2021;134(7):783-791.
4. Keum N, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. Nat Rev Gastroenterol Hepatol. 2019;16(12):713-732.
5. Vernia F, Longo S. Dietary factors modulating colorectal carcinogenesis. Nutrients. 2021;13(1):143.
6. McNabb S, Harrison TA, Albanes D, et al. Meta-analysis of 16 studies of the association of alcohol with colorectal cancer. Int J Cancer. 2020;146(3):861-873.
7. Rumgay H, Murphy N, Ferrari P, Soerjomataram I. Alcohol and cancer: epidemiology and biological mechanisms. Nutrients. 2021;13(9):3173.
8. Waris S, Patel A, Ali A, Mahmood R. Acetaldehyde-induced oxidative modifications and morphological changes in isolated human erythrocytes: an in vitro study. Environ Sci Pollut Res Int. 2020;27(14):16268-16281.
9. Chubb D, Broderick P, Frampton M, et al. Genetic diagnosis of high-penetrance susceptibility for colorectal cancer (CRC) is achievable for a high proportion of familial CRC by exome sequencing. J Clin Oncol. 2015;33(5):426-432.
10. Lu Y, Kweon SS, Tanikawa C, et al. Large-scale genome-wide association study of east Asians identifies loci associated with risk for colorectal cancer. Gastroenterology. 2019;156(5):1455-1466.
11. Schumacher FR, Schmit SL, Jiao S, et al. Genome-wide association study of colorectal cancer identifies six new susceptibility loci. Nat Commun. 2015;6:7138.
12. McKay JD, Truong T, Gaborieau V, et al. A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE consortium. PLoS Genet. 2011;7(3):e1001333.
13. Omran Z. Development of new disulfiram analogues as ALDH1a1-selective inhibitors. Biomed Med Chem Lett. 2021;40:127958.
14. Mizoi Y, Yamamoto K, Ueno Y, Fukunaga T, Harada S. Involvement of genetic polymorphism of alcohol and aldehyde dehydrogenases in individual variation of alcohol metabolism. Alcohol Alcohol. 1994;29(6):707-710.
15. Matsumoto A. The bidirectional effect of defective ALDH2 polymorphism and disease prevention. Adv Exp Med Biol. 2019;1193:69-87.
16. Seo W, Gao Y, He Y, et al. ALDH2 deficiency promotes alcohol-associated liver cancer by activating oncogenic pathways via oxidized DNA-enriched extracellular vesicles. J Hepatol. 2019;71(5):1000-1011.
17. Koyanagi YN, Suzuki E, Imoto I, et al. Across-site differences in the mechanism of alcohol-induced digestive tract carcinogenesis: An evaluation by mediation analysis. Cancer Res. 2020;80(7):1601-1610.
18. Yu C, Guo Y, Bian Z, et al. Association of low-activity ALDH2 and alcohol consumption with risk of esophageal cancer in Chinese adults: a population-based cohort study. Int J Cancer. 2018;143(7):1652-1661.
19. Yang H, Zhou Y, Zhou Z, et al. A novel polymorphism rs1329149 of CYP2E1 and a known polymorphism rs671 of ALDH2 of alcohol metabolizing enzymes are associated with colorectal cancer in a southwestern Chinese population. Cancer Epidemiol Biomarkers Prev. 2009;18(9):2522-2527.
20. Guo XF, Wang J, Yu SJ, et al. Meta-analysis of the ADH1B and ALDH2 polymorphisms and the risk of colorectal cancer in east Asians. Intern Med. 2013;52(24):2693-2699.
21. Yin G, Kono S, Toyomura K, et al. Alcohol dehydrogenase and aldehyde dehydrogenase polymorphisms and colorectal cancer: the Fukuoka colorectal cancer study. Cancer Sci. 2007;98(8):1248-1253.
22. Choi CK, Shin MH, Cho SH, et al. Association between ALDH2 polymorphisms and the risk of colorectal cancer in Koreans. Cancer Res Treat. 2021;53(3):754-762.
23. Wang WZ, Wang CY, Cheng YT, et al. Tracing the origins of Hakka and Chaoshanese by mitochondrial DNA analysis. Am J Phys Anthropol. 2010;141(1):124-130.
24. Shah D, Bentrem D. Environmental and genetic risk factors for gastric cancer. J Surg Oncol. 2022;125(7):1096-1103.
25. Gomez-Quiroz LE, Roman S. Influence of genetic and environmental risk factors in the development of hepatocellular carcinoma in Mexico. Ann Hepatol. 2022;27(Suppl 1):100649.
31. Han X, Shen A, Yao T, et al. Genetic diversity of 17 autosomal STR loci.

32. Park SY, Wilkens LR, Setiawan VW, Monroe KR, Haiman CA, Le.

34. Lee S, Woo H, Lee J, Oh JH, Kim J, Shin A. Cigarette smoking, alcohol consumption, and risk of colorectal cancer in South Korea: a prospective investigation into cancer and nutrition (EPIC)-Italy cohort. Eur J Gastroenterol Hepatol. 2020;32(4):475-483.

35. Gu MJ, Huang QC, Bao CZ, et al. Attributable causes of colorectal cancer in China. BMC Cancer. 2018;18(1):38.

36. Baicus C, Caraiola S, Rimbas M, Patrascu R, Baicus A. Utility of routine hematological and inflammation parameters for the diagnosis of cancer in involuntary weight loss. Int J Lab Hematol. 2015;37(2):121-126.

37. Spell DW, Jones DV Jr, Harper WF, David BJ. The value of a complete blood count in predicting cancer of the colon. Cancer Detect Prev. 2004;28(1):37-42.

38. Seretis C, Seretis F, Lagoudianakis E, Gemenetzis G, Salemis NS. Is red cell distribution width a novel biomarker of breast cancer activity? Data from a pilot study. J Clin Med Res. 2013;5(2):121-126.

39. Koma Y, Onishi A, Matsuoka H, et al. Increased red blood cell distribution width associates with cancer stage and prognosis in patients with lung cancer. PloS One. 2013;8(11):e80240.

40. Lee H, Kong SY, Sohn JY, Shim H, Youn HS, Lee S, Kim HJ, Eom HS. Elevated red blood cell distribution width as a simple prognostic factor in patients with symptomatic multiple myeloma. Biomed Res Int. 2014;2014:1-8.

41. Yamamoto T, Kawada K. Inflammation-related biomarkers for the prediction of prognosis in colorectal cancer patients. Int J Mol Sci. 2021;22(15):8002.

42. Kim JH, Cho IK, Kim YA, Park SJ. Elevated neutrophil-to-lymphocyte ratio in metabolic syndrome is associated with increased risk of colorectal adenoma. Metab Syndr Relat Disord. 2017;15(8):393-399.

43. Sun F, Tan YA, Gao QF, et al. Circulating fibrinogen to pre-albumin ratio is a promising biomarker for diagnosis of colorectal cancer. J Clin Lab Anal. 2019;33(1):e22635.

44. Hernandez-Ainsa M, Velamazan R, Lanas A, Carrera-Lasfuentes P, Piazzuelo E. Blood-cell-based inflammatory markers as a useful tool for early diagnosis in colorectal cancer. Front Med. 2022;9:843074.

45. Stojkovic Lalosevic M, Pavlovic Markovic A, Stankovic S, et al. Combined diagnostic efficacy of neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), and mean platelet volume (MPV) as biomarkers of systemic inflammation in the diagnosis of colorectal cancer. Dis Markers. 2019;2019:6036979-6036977.

46. Kang Y, Zhu X, Lin Z, et al. Compare the diagnostic and prognostic value of MLR, NLR and PLR in CRC patients. Clin Lab. 2021;67(9). doi:10.7754/Clin.Lab.2021.201130

47. Zhu X, Cao Y, Lu P, et al. Evaluation of platelet indices as diagnostic biomarkers for colorectal cancer. Sci Rep. 2018;8(1):11814.

48. Mo CJ, Hu ZJ, Qin SZ, Chen HP, Huang L, Li S. Diagnostic value of platelet-lymphocyte ratio and hemoglobin-platelet ratio in patients with rectal cancer. J Clin Lab Anal. 2020;34(4):e23153.

49. Zeng D, Huang Q, Yu Z, Wu H. Association between aldehyde dehydrogenase 2 gene rs671 G→a polymorphism and alcoholic liver cirrhosis in southern Chinese Hakka population. J Clin Lab Anal. 2021;35(7):e23855.

50. Chen Y, Liu H, Yu Z, et al. ALDH2 polymorphism rs671 *1/*2 genotype is a risk factor for the development of alcoholic liver cirrhosis in Hakka alcoholics. Int J Gen Med. 2022;15:4067-4077.

51. Li R, Zhao Z, Sun M, Luo J, Xiao Y. ALDH2 gene polymorphism in different types of cancers and its clinical significance. Life Sci. 2016;147:59-66.

52. Chang JS, Hsiao JR, Chen CH. ALDH2 polymorphism and alcohol-related cancers in Asians: a public health perspective. J Biomed Sci. 2017;24(1):19.

53. Seol JE, Kim J, Lee BH, et al. Folate, alcohol, ADH1B and ALDH2 and colorectal cancer risk. Public Health Nutr. 2020;6:1-8.

54. Gao CM, Takezaki T, Wu JZ, et al. Polymorphisms of alcohol dehydrogenase 2 and aldehyde dehydrogenase 2 and colorectal cancer risk in Chinese males. World J Gastroenterol. 2008;14(32):5078-5083.

55. Im PK, Yang L. Alcohol metabolism genes and risks of site-specific cancers in Chinese adults: an 11-year prospective study. Int J Cancer. 2022;150(10):1627-1639.

56. Chen B, Hu KW, Zhang JW, Wei ZJ, Meng XL, Xiong MM. A critical analysis of the relationship between aldehyde dehydrogenases-2 Glu487Lys polymorphism and colorectal cancer susceptibility. Pathol Oncol Res. 2015;21(3):727-733.

57. Hirose M, Kono S, Tabata S, et al. Genetic polymorphisms of methyleneetetrahydrofolate reductase and aldehyde dehydrogenase 2, alcohol use and risk of colorectal adenomas: self-defense forces health study. Cancer Sci. 2005;96(8):513-518.

58. Kuriki K, Hamajima N, Chiba H, et al. Relation of the CD36 gene A52C polymorphism to the risk of colorectal cancer among Japanese, with reference to with the aldehyde dehydrogenase 2 gene Glu487Lys polymorphism and drinking habit. Asian Pac J Cancer Prev. 2005;6(1):62-68.

59. Matsuo K, Wakai K, Hirose K, et al. A gene-gene interaction between ALDH2 Glu487Lys and ADH2 His47Arg polymorphisms regarding the risk of colorectal cancer in Japan. Carcinogenesis. 2006;27(5):1018-1023.

60. Matsuo K, Hamajima N, Hirai T, et al. Aldehyde dehydrogenase 2 (ALDH2) genotype affects rectal cancer susceptibility due to alcohol consumption. J Epidemiol. 2002;12(2):70-76.

61. Zhong Q, Wu RR, Zeng ZM. Association of ADH1B Arg47His and ALDH2 Glu487Lys polymorphisms with risk of colorectal cancer and their interaction with environmental factors in a Chinese population. Genet Mol Res. 2016;15(3):1-8.

62. Zhang H, Fu L. The role of ALDH2 in tumorigenesis and tumor progression: targeting ALDH2 as a potential cancer treatment. Acta Pharm Sin B. 2021;11(6):1400-1411.

63. Matsuoka K, Kumatoriya K, Tando M, Kometani T, Shinhara M. Polyphenols from persimmon fruit attenuate acetaldehyde-induced DNA double-strand breaks by scavenging acetaldehyde. Sci Rep. 2022;12(1):10300.

64. Wang S, Wang L, Qin X, et al. ALDH2 contributes to melatonin-induced protection against APP/PS1 mutation-prompted cardiac anomalies through cGAS-STING-TBK1-mediated regulation of miR146a. Signal Transduct Target Ther. 2020;5(1):119.
65. Hu J, Yang L, Peng X, et al. ALDH2 hampers immune escape in liver hepatocellular carcinoma through ROS/Nrf2-mediated autophagy. *Inflammation*. 2022;45:2309-2324.

66. Zhang H, Xia Y, Wang F, et al. Aldehyde dehydrogenase 2 mediates alcohol-induced colorectal cancer immune escape through stabilizing PD-L1 expression. *Adv Sci*. 2021;8(10):2003404.

67. Yao S, Yin X, Chen T, et al. Exploring ALDH2 expression and immune infiltration in HNSC and its correlation of prognosis with gender or alcohol intake. *Sci Rep*. 2022;12(1):2504.

How to cite this article: Zhang Z, Chen Y, Zhuo Q, et al. ALDH2 gene rs671 G > a polymorphism and the risk of colorectal cancer: A hospital-based study. *J Clin Lab Anal*. 2022;36:e24789. doi:10.1002/jcla.24789