Association of CYP19A1 Polymorphism with Genetic Susceptibility to Lung Cancer in Chinese population:a case-control study

Anqi Li
Xi'an Jiaotong University Medical College First Affiliated Hospital
Meng Li
Xi'an Jiaotong University Medical College First Affiliated Hospital
Ruiqing He
Xi'an Jiaotong University Medical College First Affiliated Hospital
Wenhui Dang
Xi'an Jiaotong University Medical College First Affiliated Hospital
Yang Li
Xi'an Jiaotong University Medical College First Affiliated Hospital
Tianbo Jin
Xi'an Northwest University
Ning Zhang
Xi'an Jiaotong University Medical College First Affiliated Hospital
Shanshan Zhang
Xi'an Northwest University
Mingwei Chen (chenmw36@163.com)
Xi'an Jiaotong University Medical College First Affiliated Hospital https://orcid.org/0000-0002-0418-1510

Research article

Keywords: CYP19A1 polymorphisms; lung cancer; gender; age; pathologic types

Posted Date: November 25th, 2019

DOI: https://doi.org/10.21203/rs.2.17673/v1

Abstract

Background Lung cancer is a kind of cancer with high morbidity and mortality related to genetic factors. Many studies have shown that CYP19A1 gene polymorphism is associated with a variety of cancers, but there are few studies on lung cancer at present. The aim of the study was to explore the correlation between CYP19A1 polymorphisms and lung cancer risk in Chinese population.Methods We enrolled 510 lung cancer patients as the case group and 504 healthy people as the control group. Five single nucleotide polymorphisms determined in CYP19A1 gene were genotyped by MassARRAY, and correlation analysis was performed by Chi square test and logistic regression model.Results The genotypes of rs4646 (OR=0.77, p=0.010), rs6493487 (OR=0.76, p=0.006) and rs17601876 (OR=0.69, p=1.15E-04) in CYP19A1 gene were linked to decreasing the risk of lung cancer, while the rs1062033 (OR=1.49, p=0.029) was linked to increasing lung cancer risk. In gender-stratified analysis, female patients with the GG genotype of the rs6493487 (OR=0.31, p=0.037) and male patients with the rs17601876 (OR=0.52, p=0.012) had lower lung cancer risk. In age-stratified analysis, for patients ≥58 years, decreased lung cancer risk was correlated with the genotypes of rs4646 (OR=0.66, p=0.021), rs6493487 (OR=0.65, p=0.021) and rs17601876 (OR=0.39, p=0.001), and increased risk was associated with the GG genotype of rs1062033 (OR=2.09, p=0.003). In pathologic type-stratified analysis, the AA genotype of rs17601876 (OR=0.41, p=0.048) was associated with decreased risk of small cell lung cancer, and the AA genotype of rs4646 (OR=0.41, p=0.027), the GA genotype of rs6493487 (OR=0.65, p=0.024) and the AA genotype of rs17601876 (OR=0.34, p=0.005) were linked to decreased risk of squamous cell carcinoma.Conclusion CYP19A1 polymorphisms are associated with lung cancer risk, especially in elderly patients and patients with pathologic types of small cell lung cancer and squamous cell carcinoma.

Background

The GLOBOCAN 2018 estimates that there will be 18.1 million new cancer cases and 9.6 million cancer deaths worldwide in 2018, of the cancer cases, Lung cancer (LC) is the most commonly diagnosed cancer (11.6% of the total cases) and it is the leading cause of cancer death (18.4% of the total cancer deaths) [1]. LC results from the interaction of environmental exposure and genetic factors, and most procarcinogens can become carcinogens when they are metabolized in the body.

Cytochromes P450 (CYPs) are proteins of the superfamily of monooxygenases involved in the metabolism of endogenous and exogenous substances. CYP enzymes can covalently bind nucleic acids and proteins to cause genetic mutations, or mediate some signal transduction pathways to induce tumorigenesis and development [2-4]. The *CYP19A1* gene is located on the chromosome 15 at 15q21.2 and it mainly encodes CYPs aromatase, which involves converting testosterone to estradiol and androstenedione to estrone respectively [5, 6]. Most studies of the *CYP19A1* gene are about hormone-related cancers, such as the breast cancer, prostate cancer, and endometrial cancer, and some research of them found that they are directly related to endogenous and exogenous steroid hormones that affect cell proliferation [7-10]. These results suggest that *CYP19A1* gene may be associated with the development of tumors. Besides, there have been studies showing a relation between *CYP19A1* gene and LC. Researchers found that aberrant activation of alternative *CYP19* promoters may lead to upregulation of local aromatase expression in some cases of non-small cell lung cancer (NSCLC) [11]. Immunohistochemical staining tests showed that aromatase was positive in LC specimens and it was mainly distributed in epithelial cells and infiltrating macrophages, suggesting that estrogen release may occur locally in tumor microenvironment [12]. Ikeda K et al. discovered that the rs3764221 on *CYP19A1* gene contributes to the development of multi-centric adenocarcinomas in the peripheral lung by causing higher levels of *CYP19A1* expression [13]. Therefore, we assume that *CYP19A1* polymorphisms might be associated with LC.

So far, the existing studies on *CYP19A1* gene and LC mainly focus on NSCLC. However, the present study aimed to reveal the association between *CYP19A1* gene and all pathologic types of LC by analyzing five single nucleotide polymorphisms (SNPs) in *CYP19A1* gene so as to provide a direction for further study of LC.

Methods

Study Population

510 LC patients and 504 healthy people were recruited for the case-control comparative studies. The patients all came from the First Affiliated Hospital of Xi'an Jiaotong University and had been diagnosed and histopathologically confirmed to have primary LC, and clinically staged according to the latest edition of the TNM Staging for LC adopted by the International Union Against Cancer (UICC). For the cases recruited, there were no limitations in age, gender, pathologic types and clinical stages of LC, and the patients had no history of cancer, received no radiotherapy and chemotherapy. The control subjects were recruited from the healthy people who received annual health checkup in Medical Examination Center of the First Affiliated Hospital of Xi'an Jiaotong University and were confirmed to have no any chronic or serious endocrine or metabolic diseases.

Genotyping

Genomic Deoxyribonucleic acid (DNA) was extracted from whole blood by using Whole Blood Genome DNA Purification Kit (Xi'an GOLDMAG Biological Company). DNA concentration and purity were determined by using the Nanodrop Lite Ultraviolet Spectrophotometer (Thermo Technology Company). Primers for amplification process and single base extension reactions were designed with Agena MassARRAY Assay Design 3.0 software according to the sequence of the forward strand from the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/). The primer information of five SNPs is shown in Supplementary Table 1. Five SNPs were genotyped on the MassARRAY iPLEX (Agena Bioscience, San Diego, CA, USA) platform by using Matrix-assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF), and the results were output by Agena Bioscience TYPER version 4.0 software. The repeated control samples were set in every genotyping plate and the coincidence rate was >99%. All SNPs were genotyped and the typing rate was> 99.4%.

Statistical Analysis

The Hardy–Weinberg Equilibrium (HWE) analysis of the control group was performed by using the Fisher exact test (p□0.05). Chi square (χ 2) test was used to evaluate the correlation between every SNP and the risk of LC. The odds ratio (OR) and 95% confidence intervals (CI) for each genotype were calculated by logistic regression analyses in the Plink software (http://www.cog-genomics.org/plink2/).

Results

Comparison of baseline characteristics for lung cancer (LC) case and control subjects

General characteristics of the case and control group are listed in the Table 1. In the study, the mean age of case group was 58.0±10.55 years and that of the control group was 57.27±10.85 years. There was no statistical difference between them (p=0.227). For the case group, 75.3% were males and 24.7% were females, and for the control group 75.6% were males and 24.4% were females. There was no statistical difference between the two groups (p=0.911). In the case group, pathologic types were mainly small cell lung cancer (SCLC), squamous cell carcinoma (SQC) and adenocarcinoma (ADC), accounting for 95.1% of the total cases. In addition, the number of patients with lymph node metastasis was 14.4% more than that no lymph node metastasis, and 48.6% of case group had advanced stage (stage III and stage IV).

List of research SNPs, positions, and genotyping data

Table 2 shows the research SNPs of *CYP19A1* and their information. From the table, the research SNPs all fulfil HWE (p>0.05), in addition the table shows the genotype call rates that ranged from 99.4 to 99.9%.

Association of the minimum allele frequencies of CYP19A1 research SNPs with lung cancer (LC) risk

We analyzed five SNPs of *CYP19A1* in the study. Table 3 illustrates the frequency distribution for the minimum alleles in the cases and controls. The minimum allele frequencies of rs4646, rs6493487 and rs17601876 in case group are lower than those in the control group. The allele A of rs4646 (OR=0.77, p=0.010), allele G of rs6493487 (OR=0.76, p=0.006) and allele A of rs17601876 (OR=0.69, p=1.15E-04) are associated with lower LC risk.

Association of CYP19A1 research SNPs Genotype with lung cancer (LC) risk

Table 4 displays the genotypes of rs4646 (p=0.035), rs6493487 (p=0.023), rs1062033 (p=0.009) and rs17601876 (p=0.001) are significantly different between the case and the control. In the table, the AA of rs4646 (OR=0.58, p=0.026), the rs6493487 (GG, OR=0.56, p=0.023; GA, OR=0.77, p=0.047) and the rs17601876(AA, OR=0.44, p=4.75E-04; AG, OR=0.71, p=0.009) of all have shown to link to decreasing the risk of LC, while the GG of rs1062033 (OR=1.49, p=0.029) shows to increase the LC risk, and the rs3751599 shows no association with the risk of LC. There are still same after adjusted for age and sex.

Association of CYP19A1 research SNPs with lung cancer (LC) risk in genetic model analysis

Table 5 summarizes the correlation between research SNPs and lung cancer in different genetic models including dominant, recessive, and additive genetic model. Same as genotype model, rs17601876 was associated with the reduction of LC risk in all three models. However, rs4646 and rs6493487 were only associated with lower LC risk in dominant and additive model, and rs1062033 was only related with higher LC risk in recessive model.

Association of CYP19A1 research SNPs with lung cancer (LC) risk in different gender

Table 6 demonstrates only rs17601876 is related to LC risk in male (AA, OR=0.52, p=0.012; AG, OR=0.74, p=0.047). Moreover, the GG of rs6493487 (OR=0.31, p= 0.037) and the AA of rs17601876 (OR=0.24, p=0.009) are associated with lowering the risk of LC in females.

Association of CYP19A1 research SNPs with lung cancer (LC) risk in different age

Table 7 shows all research SNPs have no significant association with LC in <58 years. But in ≥58, the AC of rs4646 (OR=0.66, p=0.021), the GA of rs6493487 (OR=0.65, p=0.021) and rs17601876 (AA, OR=0.39, p=0.001; AG, OR=0.57, p=0.002) are shown to decrease the risk of LC, while the GG of rs1062033 (OR=2.09, p= 0.003) increases the risk of LC.

Association of CYP19A1 research SNPs with lung cancer (LC) risk in different pathologic type

Table 8 indicates the AA of rs17601876 (OR=0.41, p=0.048) in SCLC, the rs4646 (AA, OR=0.41, p=0.027), rs6493487 (GA, OR=0.65, p=0.024), and rs17601876 (AA, OR=0.34, p=0.005) in SQC are shown to lower the risk of LC. Interestingly, research SNPs have no relationship with LC risk in ADC.

Discussion

The *CYP19A1* gene encodes aromatase, which is involved in the conversion of androstenedione and testosterone to estrone and estradiol respectively as a rate-limiting enzyme [14, 15]. The aromatase is expressed both in gonad and extragonadal tissues including lung, brain, and liver. The activity of aromatase in LC tissues is higher than that in normal lung tissues [16]. The expression of aromatase in NSCLC is associated with estrogen production [17]. *CYP19A1* polymorphisms locally raises the level of estrogen in peripheral lung tissue [18]. Estrogen directly causes cell proliferation and DNA damage of lung tissue [19], and regulate the expression of growth factors such as Vascular Endothelial Growth Factor (VEGF), which promotes microangiogenesis of LC [20], and leads to the beginning and development of LC.

In the present study, we found that the genotypes of rs4646, rs6493487, rs17601876 were linked to lowering the LC risk, while the rs1062033 may increase the LC risk. In the study by Olivo-Marston SE et al. the rs4646 was found to be associated with lowering the levels of serum estrogen among LC patients [16].

Therefore, we conclude that the four SNPs in *CYP19A1* gene have an impact on the risk of LC by affecting local estrogen levels in LC tissues. In addition, it has been shown by Kohno M et al. that high aromatase expression was associated with poor prognosis for both recurrence-free survival and overall survival in lung adenocarcinomas [21]. Hence, we evaluated the correlation between *CYP19A1* gene expression and prognosis in LC tissues through TCGA database (shown in Fig 1). We found that LC with low expression of *CYP19A1* had a higher survival rate than those with high expression of *CYP19A1* (http://kmplot.com/). These results suggest that *CYP19A1* polymorphisms are associated with LC.

In male, the rs17601876 was associated with decreasing the risk of LC, and the other SNPs were not associated with LC. Estrogen receptor is also expressed in male non-reproductive system and regulated by estrogen. Verma MK et al. found that co-expression of estrogen receptor (ER) β and aromatase in male can promote the development of LC, suggesting that rs17601876 may be associated with the risk of LC in male [22, 23]. Overall, the rs6493487 was associated with lowering the risk of LC, but in gender-stratified analysis, the GG of rs6493487 was only correlated with female. Yang SY et al. found that TTTA repeat polymorphism in intron region of *CYP19A1* gene was associated with L858R mutation which is one of epidermal growth factor receptor (EGFR) mutations in female never-smokers [24], and rs6493487 was located in the intron variant of *CYP19A1*, so we assumed that rs6493487 may be related to EGFR mutations and may become a target for future targeted therapy of female LC.

In ≥58 years, the genotypes of rs4646, rs6493487, rs17601876 were associated with decreasing the LC risk, and rs1062033 was linked increasing the risk and all these SNPs had no association with the risk of patients <58 years. The estrogen of postmenopausal women and men is mainly synthesized by aromatase in non-gonadal tissues (e.g. lung) [18], and it may explain the difference in two age groups in the association between *CYP19A1* gene polymorphism and LC risk. Scholars believe that the use of exogenous estrogen in perimenopausal and menopausal women can increase the risk of time-dependent LC, and anti-estrogen therapy can reduce the incidence of secondary lung cancer in breast cancer patients >50 years old [25]. Additionally, studies have shown that elderly women with NSCLC have a longer life span than men and young women, which can be partly explained by lower estradiol (E2) levels[26]. The above studies suggest an important role of estrogen in LC of aged patients, and provide a suitable population for the study of *CYP19A1* gene in LC.

The major pathologic types of LC include SQC, ADC and SCLC. At present, a majority of studies believe that estrogen plays an important role in the beginning and development of NSCLC, especially in lung adenocarcinoma. Estrogen promotes the growth of lung adenocarcinoma cells expressing ERβ receptor, and antagonizing estrogen does the opposite [27]. However, the present study found that there was no correlation between *CYP19A1* gene and lung adenocarcinoma. Interestingly, we found that the genotypes of rs4646, rs6493487, rs17601876 were associated with lowering the risk of SCLC and SQC. Therefore, we believe that there may be other mechanisms between *CYP19A1* gene and LC, and it needs further study.

The present study demonstrated the correlation between *CYP19A1* polymorphisms and LC risk in different genders, age groups and pathologic types, but did not analyze the estrogen level. In the future, we will analyze the estrogen level in different groups, exclude the influence of gender, age and other factors on estrogen level, further clarifying the correlation between *CYP19A1* gene expression, estrogen level and the occurrence of LC.

Conclusions

CYP19A1 gene, encoding the aromatase, is associated with the estrogen level in LC tissues. The genotypes of rs4646, rs6493487 and rs17601876 in *CYP19A1* gene are associated with lowering the risk of LC in the elderly, SCLC and SQC, while rs1062033 has a correlation with increasing the LC risk in elderly patients. It may provide a direction for future research of *CYP19A1* gene used for risk prediction and treatment of LC.

Abbreviations

LC: Lung cancer

CYPs: Cytochromes P450

NSCLC: non-small cell lung cancer

SNPs: single nucleotide polymorphisms

UICC: International Union Against Cancer

DNA: Deoxyribonucleic acid

MALDI-TOF: Matrix-assisted Laser Desorption Ionization-Time of Flight

HWE: Hardy–Weinberg Equilibrium

χ 2: Chi square

OR: odds ratio

CI: confidence intervals

SCLC: small cell lung cancer

SQC: squamous cell carcinoma

ADC: adenocarcinoma

VEGF: Vascular Endothelial Growth Factor

ER: estrogen receptor

EGFR: epidermal growth factor receptor

E2: estradiol

LNM: lymph node metastasis

MA: minimum allele

MAF: minimum alleles frequency

PCRP: polymerase chain reaction primer

UEP_SEQ: unit evolutionary period_ sequences

HR: Hazard Ratio

Declarations

Ethics approval and consent to participate

The present study was approved by the Ethics Committee of the First Affiliated Hospital of Xi'an Jiaotong University and informed consent was obtained from each participant after a full explanation of the study.

Consent for publication

Not applicable

Availability of data and materials

The datasets generated and/or analyzed over the course of the study are

not publicly available but are available from the corresponding author on

reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was funded by the First Affiliated Hospital, Xi'an Jiao tong University. There is no role for the First Affiliated Hospital, Xi'an Jiao tong University in the design of the study; collection, analysis, and interpretation of data; and in writing the manuscript.

Authors' contributions

Anqi Li, Mingwei Chen and Tianbo Jin designed the method study and supervised the study. Anqi Li, Yang Li and Ning Zhang participated in data collection, data analysis and drafted the manuscript. Anqi Li, Meng Li, Ruiqing He, Wenhui Dang and Shanshan Zhang helped with the interpretation, and description of the results. All authors read and approved the final manuscript.

Acknowledgements

The authors thank the Ethics Committee of the First Affiliated Hospital of Xi'an Jiao tong University for approving this study and making the clinical data and samples available for the study.

Authors' information

[1]The Department of Respiratory and Critical Care Medicine, the First Affiliated Hospital, Xi'an Jiaotong University, Xi'an, Shaanxi 710061, China.

[2]Shaanxi Provincial Research Center for the Project of Prevention and Treatment of Respiratory Diseases, Xi'an, Shaanxi 710021, China.

[3]Ministry of Education Key Laboratory of Resource Biology and Biotechnology in Western China, Northwest University, Xi'an, Shaanxi 710069, China.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries**. *CA Cancer J Clin* 2018, **68**(6):394-424.
2. Huang FM, Chen HC, Khan MA, Yang FL, Wan XX, Xu AH, Ou-yang FD, Zhang DZ: **CYP2A6, CYP1A1, and CYP2D6 polymorphisms in lung cancer patients from Central South China**. *Med Oncol* 2013, **30**(2).
3. Hasi S, Yao J, Yu S, Tian Y: **Diversity and distribution of CYP gene family in Bactrian camel**. *Funct Integr Genomics* 2018, **18**(1):23-29.
4. Nebert DW, Dalton TP: **The role of cytochrome P450 enzymes in endogenous signalling pathways and environmental carcinogenesis**. *Nat Rev Cancer* 2006, **6**(12):947-960.
5. Subbaramaiah K, Iyengar NM, Morrow M, Elemento O, Zhou XK, Dannenberg AJ: **Prostaglandin E2down-regulates sirtuin 1 (SIRT1) leading to elevated levels of aromatase, providing insights into the obesity-breast cancer connection**. *J Biol Chem* 2018.
6. Artigalas O, Vanni T, Hutz MH, Ashton-Prolla P, Schwartz IV: **Influence of CYP19A1 polymorphisms on the treatment of breast cancer with aromatase inhibitors: a systematic review and meta-analysis**. *Bmc Med* 2015, **13**.
7. Armamento-Villareal R, Shah VO, Aguirre LE, Meisner ALW, Qualls C, Royce ME: **The rs4646 and rs12592697 Polymorphisms in CYP19A1 Are Associated with Disease Progression among Patients with Breast Cancer from Different Racial/Ethnic Backgrounds**. *Front Genet* 2016, **7**.
8. Magnani L, Frige G, Gadaleta RM, Corleone G, Fabris S, Kempe H, Verschure PJ, Barozzi I, Vircillo V, Hong SP *et al*: **Acquired CYP19A1 amplification is an early specific mechanism of aromatase inhibitor resistance in ER alpha metastatic breast cancer**. *Nat Genet* 2017, **49**(3):444-450.
9. Kanda S, Tsuchiya N, Narita S, Inoue T, Huang MG, Chiba S, Akihama S, Saito M, Numakura K, Tsuruta H *et al*: **Effects of functional genetic polymorphisms in the CYP19A1 gene on prostate cancer risk and survival**. *Int J Cancer* 2015, **136**(1):74-82.
10. Thompson DJ, O'Mara TA, Glubb DM, Painter JN, Cheng T, Folkerd E, Doody D, Dennis J, Webb PM, Gorman M *et al*: **CYP19A1 fine-mapping and Mendelian randomization: estradiol is causal for endometrial cancer**. *Endocr-Relat Cancer* 2016, **23**(2):77-91.
11. Demura M, Demura Y, Ameshima S, Ishizaki T, Sasaki M, Miyamori I, Yamagishi M, Takeda Y, Bulun SE: **Changes in aromatase (CYP19) gene promoter usage in non-small cell lung cancer**. *Lung Cancer* 2011, **73**(3):289-293.
12. Siegfried JM, Stabile LP: **Estrongenic steroid hormones in lung cancer**. *Semin Oncol* 2014, **41**(1):5-16.
13. Ikeda K, Shiraishi K, Eguchi A, Osumi H, Matsuishi K, Matsubara E, Fujino K, Shibata H, Yoshimoto K, Mori T *et al*: **Association of a genetic variant of CYP19A1 with multicentric development of lung adenocarcinomas**. *Ann Surg Oncol* 2014, **21**(3):939-945.
14. Kaewlert W, Sakonsinsiri C, Namwat N, Sawanyawisuth K, Ungarreevittaya P, Khuntikeo N, Armartmuntree N, Thanan R: **The Importance of CYP19A1 in Estrogen Receptor-Positive Cholangiocarcinoma**. *Horm Cancer* 2018, **9**(6):408-419.
15. Siegfried JM: **Smoking out reproductive hormone actions in lung cancer**. *Mol Cancer Res* 2014, **12**(1):24-31.
16. Olivo-Marston SE, Mechanic LE, Mollerup S, Bowman ED, Remaley AT, Forman MR, Skaug V, Zheng YL, Haugen A, Harris CC: **Serum estrogen and tumor-positive estrogen receptor-alpha are strong prognostic classifiers of non-small-cell lung cancer survival in both men and women**. *Carcinogenesis* 2010, **31**(10):1778-1786.
17. Skjefstad K, Grindstad T, Khanehkenari MR, Richardsen E, Donnem T, Kilvaer T, Andersen S, Bremnes RM, Busund LT, Al-Saad S: **Prognostic relevance of estrogen receptor alpha, beta and aromatase expression in non-small cell lung cancer**. *Steroids* 2016, **113**:5-13.
18. Ikeda K, Shiraishi K, Yoshida A, Shinchi Y, Sanada M, Motooka Y, Fujino K, Mori T, Suzuki M: **Synchronous Multiple Lung Adenocarcinomas: Estrogen Concentration in Peripheral Lung**. *PLoS One* 2016, **11**(8):e0160910.
19. Slowikowski BK, Lianeri M, Jagodzinski PP: **Exploring estrogenic activity in lung cancer**. *Mol Biol Rep* 2017, **44**(1):35-50.
20. Applanat MP, Buteau-Lozano H, Herve MA, Corpet A: **Vascular endothelial growth factor is a target gene for estrogen receptor and contributes to breast cancer progression**. *Adv Exp Med Biol* 2008, **617**:437-444.
21. Kohno M, Okamoto T, Suda K, Shimokawa M, Kitahara H, Shimamatsu S, Konishi H, Yoshida T, Takenoyama M, Yano T *et al*: **Prognostic and therapeutic implications of aromatase expression in lung adenocarcinomas with EGFR mutations**. *Clin Cancer Res* 2014, **20**(13):3613-3622.
22. Cooke PS, Nanjappa MK, Ko C, Prins GS, Hess RA: **Estrogens in Male Physiology**. *Physiol Rev* 2017, **97**(3):995-1043.
23. Verma MK, Miki Y, Abe K, Nagasaki S, Niikawa H, Suzuki S, Kondo T, Sasano H: **Co-expression of estrogen receptor beta and aromatase in Japanese lung cancer patients: gender-dependent clinical outcome**. *Life Sci* 2012, **91**(15-16):800-808.
24. Yang SY, Yang TY, Chen KC, Li YJ, Hsu KH, Tsai CR, Chen CY, Hsu CP, Hsia JY, Chuang CY *et al*: **EGFR L858R mutation and polymorphisms of genes related to estrogen biosynthesis and metabolism in never-smoking female lung adenocarcinoma patients**. *Clin Cancer Res* 2011, **17**(8):2149-2158.
25. Chu SC, Hsieh CJ, Wang TF, Hong MK, Chu TY: **Antiestrogen use in breast cancer patients reduces the risk of subsequent lung cancer: A population-based study**. *Cancer Epidemiol* 2017, **48**:22-28.
26. Honma N, Hosoi T, Arai T, Takubo K: **Estrogen and cancers of the colorectum, breast, and lung in postmenopausal women**. *Pathol Int* 2015, **65**(9):451-459.
27. Kohno T, Kakinuma R, Iwasaki M, Yamaji T, Kunitoh H, Suzuki K, Shimada Y, Shiraishi K, Kasuga Y, Hamada GS *et al*: **Association of CYP19A1 polymorphisms with risks for atypical adenomatous hyperplasia and bronchioloalveolar carcinoma in the lungs**. *Carcinogenesis* 2010, **31**(10):1794-1799.

Tables

Table 1 Comparison of baseline characteristics for lung cancer (LC) case and control subjects

Characteristics	Case(n=510)	Control(n=504)	*p*
Age mean ± SD (years)	58.0±10.55	57.27±10.85	0.227[a]
Sex, n (%)			0.911[b]
Male	384(75.3)	381(75.6)	
Female	126(24.7)	123(24.4)	
Pathologic type, n (%)			
SCLC	97(19.0)		
SQC	169(33.1)		
ADC	161(31.6)		
Others	22(4.3)		
Miss	61(12.0)		
LNM			
Yes	193(37.9)		
No	120(23.5)		
Miss	197(38.6)		
TNM stage			
	62(12.2)		
	67(13.1)		
	90(17.6)		
	158(31.0)		
Miss	133(26.1)		

$p \leq 0.05$ indicates statistical significance

a: p values were calculated from *t* tests

b: p values were calculated by two-side χ 2 test

SCLC: small cell lung cancer

SQC: squamous cell carcinoma

ADC: adenocarcinoma

LNM: lymph node metastasis

Miss indicates data loss

Table 2 List of research SNPs, positions, and genotyping data

SNP	Alleles	Position	Role	p^{HWE}	Call rate (%)
rs4646	C>A^{MA}	51210647	intron variant	1.000	99.4
rs6493487	A>G^{MA}	51221532	intron variant	0.745	99.4
rs1062033	C>G^{MA}	51255741	intron variant	0.084	99.5
rs17601876	G>A^{MA}	51261712	intron variant	0.922	99.6
rs3751599	G>A^{MA}	51281336	intron variant	1.000	99.9

SNP: single nucleotide polymorphism

MA: minimum allele

HWE: Hardy–Weinberg Equilibrium

Table 3 Association of the minimum allele frequencies of *CYP19A1* research SNPs with lung cancer (LC) risk

		MAF			
SNP	MA	Case	Control	p^a	OR (95%CI)
rs4646	A	0.258	0.310	0.010*	0.77□0.64-0.94□
rs6493487	G	0.237	0.292	0.006*	0.76□0.62-0.92□
rs1062033	G	0.475	0.433	0.056	1.19□0.10-1.41□
rs17601876	A	0.269	0.348	1.15E-04*	0.69□0.57-0.83□
rs3751599	A	0.074	0.069	0.721	1.06□0.76-1.49□

* $p \leq 0.05$ indicates statistical significance

a: p values were calculated by two-side χ 2 test

SNP: single nucleotide polymorphism

MA: minimum alleles

MAF: minimum alleles frequency

OR: odds ratio

CI: confidence interval

Table 4 Association of CYP19A1 research SNPs Genotype with lung cancer (LC) risk

SNP	Genotype	Case	Control	P^a	Crude		Adjusted [b]	
					OR (95%CI)	p	OR (95%CI)	p
rs4646	AA	33	48	0.035*	0.58(0.36-0.94)	0.027*	0.58(0.36-0.94)	0.026*
	AC	197	213		0.79(0.61-1.02)	0.069	0.79(0.61-1.02)	0.071
	CC	280	238		1		1	
rs6493487	GG	29	44	0.023*	0.60(0.34-0.92	0.022*	0.56(0.34-0.92	0.023*
	GA	184	203		0.77 0.59-1.00	0.048*	0.77 0.59-0.10	0.047*
	AA	297	252		1		1	
rs1062033	GG	124	84	0.009*	1.50(1.05-2.14)	0.027*	1.49(1.04-2.13)	0.029*
	GC	236	266		0.90(0.68-1.20)	0.467	0.90(0.68-1.20)	0.475
	CC	149	151		1		1	
rs17601876	AA	34	60	0.001*	0.45(0.28-0.71)	0.001*	0.44(0.28-0.70)	4.75 E-04*
	AG	205	230		0.71(0.54-0.92)	0.009*	0.71(0.54-0.92)	0.009*
	GG	269	213		1		1	
rs3751599	AA	2	2	0.893	0.10(0.14-7.12)	0.998	0.96(0.14-6.89)	0.971
	AG	71	66		1.07(0.75-1.54)	0.701	1.08(0.75-1.54)	0.690
	GG	437	436		1		1	

a: p values were calculated by two-side χ 2 test

b: adjusted for age and sex

* $p \leq 0.05$ indicates statistical significance

SNP: single-nucleotide polymorphism

OR: odds ratio

CI: confidence intervals

Table 5 Association of *CYP19A1* research SNPs with lung cancer (LC) risk in genetic model

Genetic model	Genotype	Case	Control	Crude		Adjusted [b]	
				OR (95%CI)	*p*	OR (95%CI)	*p*
Dominant	C/C	280	238	1	0.022*	1	0.022*
	A/A-A/C	230	261	0.75(0.59-0.96)		0.75(0.59-0.96)	
Recessive	A/A	33	48	0.65(0.41-1.03)	0.067	0.65(0.41-1.03)	0.064
	C/C-A/C	477	451	1		1	
Additive	A/A	33	48	0.77(0.64-0.94)	0.010*	0.77(0.64-0.94)	0.010*
	A/C	197	213				
	C/C	280	238				
Dominant	A/A	297	252	1	0.014*	1	0.013*
	G/G-G/A	213	247	0.73(0.57-0.94)		0.73(0.57-0.94)	
Recessive	G/G	29	44	0.62(0.38-1.01)	0.057	0.62(0.38-1.02)	0.058
	G/A-A/A	481	455	1		1	
Additive	G/G	29	44	0.76(0.62-0.92)	0.006*	0.76(0.62-0.92)	0.006*
	G/A	184	203				
	A/A	297	252				
Dominant	C/C	149	151	1	0.763	1	0.762
	G/C-G/G	360	350	1.04(0.80-1.37)		1.04(0.80-1.37)	
Recessive	G/G	124	84	1.60(1.17-2.18)	0.003*	1.59(1.17-2.17)	0.003*
	G/C-C/C	385	417	1		1	
Additive	G/G	124	84	1.19(1.00-1.42)	0.056	1.19(0.99-1.41)	0.060
	G/C	236	266				
	C/C	149	151				
Dominant	G/G	269	213	1	0.001*	1	0.001*
	A/A-A/G	239	290	0.65(0.51-0.84)		0.65(0.51-0.84)	
Recessive	A/A	34	60	0.53(0.34-0.82)	0.005*	0.52(0.34-0.81)	0.004*
	A/G-G/G	474	443	1		1	
Additive	A/A	34	60	0.68(0.56-0.83)	0.000*	0.68(0.56-0.83)	0.000*
	A/G	205	230				
	G/G	269	213				
Dominant	G/G	437	436	1	0.705	1	0.699
	A/A-A/G	73	68	1.07(0.75-1.53)		1.07(0.75-1.53)	
Recessive	A/A	2	2	0.99(0.14-7.04)	0.991	0.96(0.13-6.82)	0.963
	A/G-G/G	508	502	1		1	
Additive	A/A	2	2	1.07(0.76-1.50)	0.719	1.07(0.76-1.50)	0.717
	A/G	71	66				
	G/G	437	436				

a: adjusted for age and sex

* $p \leq 0.05$ indicates statistical significance

SNP: single-nucleotide polymorphism

OR: odds ratio

CI: confidence intervals

Table 6 Association of *CYP19A1* research SNPs with lung cancer (LC) risk in different gender

rs4646			rs6493487			rs1062033			rs17601876			rs3751599		
Genotype	OR (95%CI)	p^a	Genotype	OR (95%CI)	p^a	Genotype	OR (95%CI)	p^a	Genotype	OR (95%CI)	p^a	Genotype	OR (95%CI)	p^a
AA	0.63(0.36-1.08)	0.091	GG	0.67(0.38-1.17)	0.159	GG	1.43(0.95-2.15)	0.085	AA	0.52(0.31-0.86)	0.012*	AA	NA	NA
AC	0.82(0.61-1.11)	0.196	GA	0.76(0.56-1.03)	0.077	GC	0.79(0.57-1.10)	0.168	AG	0.74(0.55-1.00)	0.047*	AG	0.96(0.63-1.46)	0.841
CC	1.00		AA	1.00		CC	1.00		GG	1.00		GG	1.00	
AA	0.46(0.17-1.24)	0.124	GG	0.31(0.11-0.93)	0.037*	GG	1.66(0.77-3.56)	0.194	AA	0.24(0.08-0.70)	0.009*	AA	NA	NA
AC	0.69(0.41-1.17)	0.170	GA	0.78(0.46-1.33)	0.363	GC	1.32(074-2.35)	0.345	AG	0.62(0.37-1.05)	0.076	AG	1.48(0.73-3.02)	0.277
CC	1.00		AA	1.00		CC	1.00		GG	1.00		GG	1.00	

* p ≤ 0.05 indicates statistical significance

a: adjusted for age and sex

OR: odds ratio

CI: confidence intervals

NA indicates data loss

Table 7 Association of *CYP19A1* research SNPs with lung cancer (LC) risk in different age

rs4646			rs6493487			rs1062033			rs17601876			rs3751599		
Genotype	OR(95%CI)	p^a	Genotype	OR(95%CI)	p^a	Genotype	OR(95%CI)	p^a	Genotype	OR(95%CI)	p^a	Genotype	OR(95%CI)	p^a
AA	0.62(0.30-1.30)	0.209	GG	0.52(0.24-1.12)	0.093	GG	0.98(0.57-1.69)	0.937	AA	0.52(0.24-1.13)	0.096	AA	NA	NA
AC	0.96(0.66-1.41)	0.840	GA	0.93(0.63-1.36)	0.697	GC	0.82(0.54-1.26)	0.373	AG	0.90(0.61-1.31)	0.572	AG	1.01(0.60-1.70)	0.982
CC	1.00		AA	1.00		CC	1.00		GG	1.00		GG	1.00	
AA	0.55(0.29-1.03)	0.062	GG	0.59(0.30-1.14)	0.062	GG	2.09(1.29-3.38)	0.003*	AA	0.39(0.22-0.69)	0.001*	AA	0.47(0.04-5.21)	0.535
AC	0.66(0.46-0.94)	0.021*	GA	0.65(0.46-0.94)	0.021*	GC	0.96(0.65-1.42)	0.848	AG	0.57(0.40-0.82)	0.002*	AG	1.12(0.67-1.85)	0.670
CC	1.00		AA	1.00		CC	1.00		GG	1.00		GG	1.00	

* p ≤ 0.05 indicates statistical significance

a: adjusted for age and sex

OR: odds ratio

CI: confidence interval

NA indicates data loss

Table 8 Association of *CYP19A1* research SNPs with lung cancer (LC) risk in different pathologic type

rs4646			rs6493487			rs1062033			rs17601876			rs3751599		
Genotype	OR(95%CI)	p^a	Genotype	OR(95%CI)	p^a	Genotype	OR(95%CI)	p^a	Genotype	OR(95%CI)	p^a	Genotype	OR(95%CI)	p^a
AA	0.37(0.13-1.06)	0.065	GG	0.41(0.14-1.18)	0.098	GG	1.41(0.76-2.60)	0.272	AA	0.41(0.17-0.99)	0.048*	AA	3.27(0.29-37.27)	0.339
AC	0.78(0.50-1.23)	0.284	GA	0.79(0.50-1.26)	0.324	GC	0.84(0.50-1.40)	0.501	AG	0.64(0.40-1.01)	0.055	AG	1.12(0.60-2.09)	0.721
CC	1.00		AA	1.00		CC	1.00		GG	1.00		GG	1.00	
AA	0.41(0.18-0.90)	0.027*	GG	0.46(0.21-1.02)	0.057	GG	1.53(0.93-2.50)	0.092	AA	0.34(0.16-0.73)	0.005*	AA	2.15(0.19-24.41)	0.538
AC	0.76(0.52-1.10)	0.140	GA	0.65(0.44-0.94)	0.024*	GC	0.86(0.57-1.30)	0.467	AG	0.75(0.52-1.08)	0.121	AG	1.08(0.64-1.81)	0.774
CC	1.00		AA	1.00		CC	1.00		GG	1.00		GG	1.00	
AA	0.83(0.43-1.59)	0.574	GG	0.71(0.35-1.45)	0.349	GG	1.30(0.77-2.19)	0.318	AA	0.52(0.27-1.03)	0.060	AA	NA	NA
AC	0.83(0.57-1.21)	0.329	GA	0.90(0.62-1.31)	0.585	GC	0.94(0.62-1.43)	0.789	AG	0.73(0.50-1.07)	0.106	AG	1.35(0.83-2.20)	0.228
CC	1.00		AA	1.00		CC	1.00		GG	1.00		GG	1.00	

* $p \leq 0.05$ indicates statistical significance

a: adjusted for age and sex

OR: odds ratio

CI: confidence interval

NA indicates data loss

SCLC: small cell lung cancer

SQC: squamous cell carcinoma

ADC: Adenocarcinoma

Figures

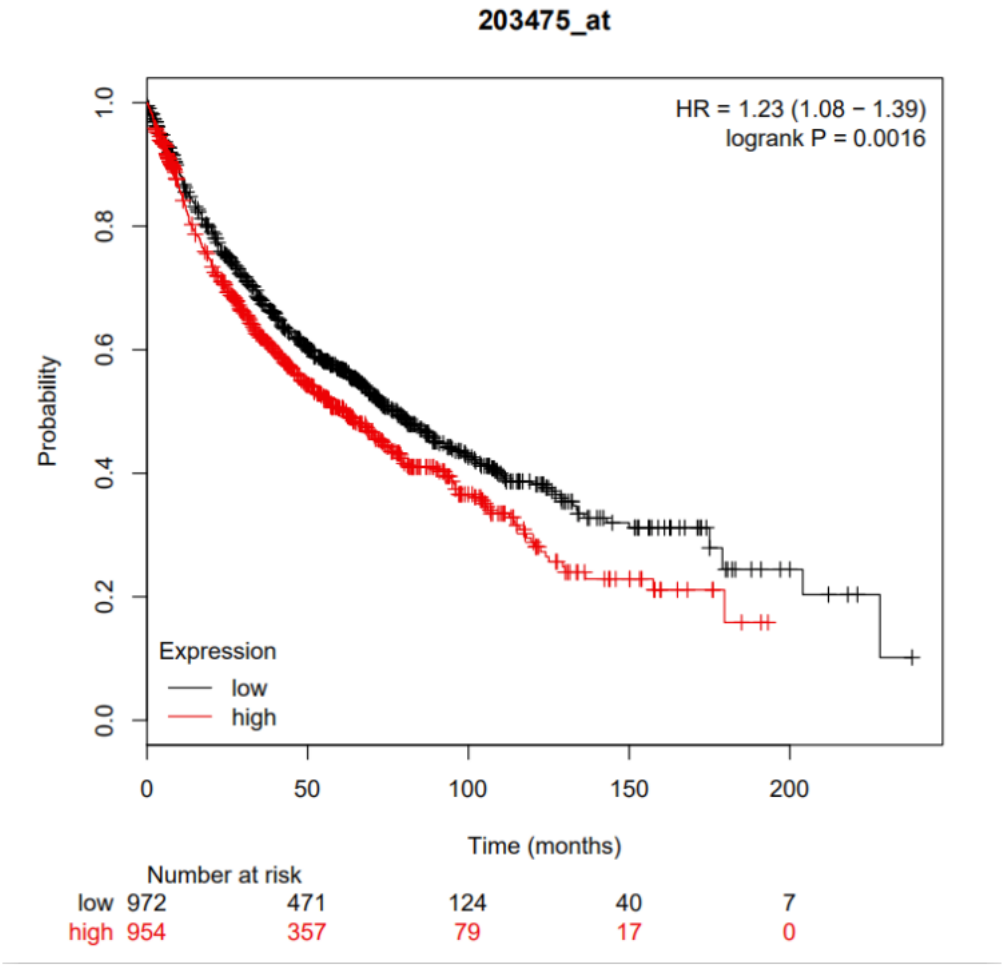

Figure 1

Survival Percentage of CYP19A1 Gene at Different Expression Levels: Black line represents survival curve of LC patients with low expression of CYP19A1; red line represents survival curve of LC patients with high expression of CYP19A1. Log-rank test shows that $p<0.05$, suggesting that two survival curves have statistical significance, and high expression of CYP19A1 group has higher death risk than low expression of CYP19A1 group (HR=1.23).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable1.docx