

Original Article

The ovarian carcinoma risk with the polymorphisms of CYP1B1 come from the positive selection

Liyang Zhang, Liyuan Feng, Meng Lou, Xihan Deng, Chuazhong Liu, Li Li

Department of Gynecologic Oncology, The Affiliated Tumor Hospital of Guangxi Medical University, Nanning, China

Received November 22, 2020; Accepted March 18, 2021; Epub May 15, 2021; Published May 30, 2021

Abstract: Ovarian carcinoma is one of the major causes of gynecological cancer. This study aimed to evaluate the association of CYP1 family polymorphism with the risk of ovarian carcinoma and chemotherapy resistance. Positive selection was detected among human CYP1A1, CYP1A2, and CYP1B1, and other species. Several positive sites were detected by site models and branch-site models. Meta-analysis was conducted for the sites rs1056836 (MAF 0.39) and rs1056827 (MAF 0.36) of CYP1B1 to clarify the association between gene polymorphisms and ovarian carcinoma risk. Subgroup analysis showed the association of rs1056836 polymorphism with ovarian cancer risk among Caucasians and Asians, while all the six genetic models showed no association among African-Americans. All the six genetic models showed no association of rs1056827 polymorphism with ovarian cancer risk. The polymorphisms of rs1056836 associated with ovarian cancer risk were detected in chemotherapy-sensitive and drug-resistant ovarian cancer patients. DNA was extracted from 62 chemotherapy resistance Ovarian carcinoma tissue samples and 137 chemotherapy-sensitive ovarian carcinoma tissue samples as controls. Gene polymorphisms were genotyped using the Sequenom MassARRAY SNP approach. There was no significant association between the CYP1B1 rs1056836 polymorphism and chemotherapy resistance of ovarian cancer in all genetic models. The results suggest that rs1056836 polymorphism of gene CYP1B1 under obvious selection pressure had a significantly increased risk for ovarian carcinoma. However, it had no significant correlation with chemotherapy resistance of ovarian cancer.

Keywords: Ovarian carcinoma, CYP1B1, positive selection

Introduction

Ovarian carcinoma is one of the major causes of gynecological cancer and subsequently a reason for mortality among women. Each year, more than 220,000 new cases are diagnosed globally, and it is projected that 14,000 ovarian cancer patients will die annually in the United State of America (U.S.A) [1, 2]. At present, maximum cytoreductive surgery and chemotherapy based on platinum combined with paclitaxel are the standard treatment strategies for patients with epithelial ovarian cancer. However, due to the high degree of heterogeneity of ovarian cancer, most patients undergoing standard treatment relapse within 2-3 years and develop multidrug resistance to chemotherapy drugs. The mechanism of ovarian cancer occurrence and drug resistance still remains unclear. Recently, several studies have revealed that the development of cancer and drug resistance may be an evolutionary pro-

cess [3, 4]. With the development of cancer evolution theory, some scholars have recently used molecular evolution theory to find a new and effective treatment of ovarian cancer.

Studies have shown that the occurrence of ovarian cancer is closely related to the expression of estradiol [5]. Further, the CYP1 family can convert estradiol into a more carcinogenic 4-hydroxyestradiol through the aromatic hydrocarbon pathway [6]. CYP1 family of enzymes is one of the important members of the CYP superfamily. CYP1 family includes three proteins namely; CYP1A1, CYP1A2, and CYP1B1. The family plays a vital role in the bioactivation of the pro-carcinogenic compounds to carcinogenic derivatives. Further, it is also important for endogenous hormone regulation, pharmacokinetics, and the response of drugs that are regulated by AhR [6]. Similarly, studies have linked it to the pathogenesis of various cancers in the reproductive organs such as the endome-

Ovarian cancer risk with CYP1B1 come from the positive selection

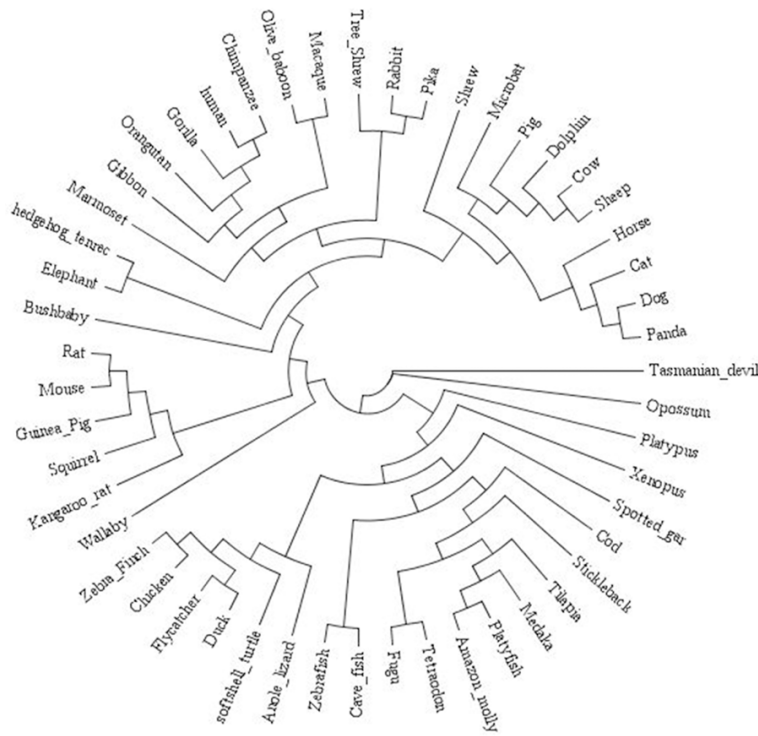


Figure 1. Phylogenetic tree of CYP1A1 gene.

trium, breast, and ovary [7-9]. CYP450 metabolism of foreign chemicals results in either successful detoxification or generation of toxic metabolites that contribute to increased risk of cancers and/or other toxic effects. However, the polymorphism in the CYP1 family gene results in different effects on the conversion efficiency and products of estradiol, resulting in differences in the occurrence of ovarian cancer and its resistance to drugs. This study conducted a selection pressure analysis on the CYP1 family from the perspective of biological evolution and screened out meaningful sites, thereby providing some theoretical basis for the occurrence and drug resistance of ovarian cancer.

Materials and methods

Sequence data collection

DNA sequences and CYP1A1, CYP1A2, and CYP1B1 protein were retrieved from Ensembl (<http://asia.ensembl.org/index.html>) and National Center for Biotechnology Information (NCBI). Subsequently, the verification of retrieved data file was carried out via EST-Blast in NCBI (<http://blast.ncbi.nlm.nih.gov/Blast>).

cgi). Coding sequences utilized in the study included 50 specials of CYP1A1, 51 specials of CYP1A2, and 46 specials of CYP1B1 as indicated in [Supplementary Tables 1, 2, 3](#).

Evolutionary analysis

The sequence alignment of the coding sequence for the protein of the CYP1 family was performed via the MUSCLE website (<https://www.ebi.ac.uk/Tools/msa/muscle/>), followed by translating the coding sequence for codon alignments. Similarly, homologs alignment was performed for each specific exon. The alignment output was examined and manually modified for fitting the software requirements. The Maximum Likelihood (ML) technique was used for the construction of a phylogenetic tree of translat-

able sequences of CYP1 family protein (**Figure 1**; 1000 bootstrap replicates) [10, 11].

Recognition of positive selection

For the quantification of natural selection's accumulative effect on the molecular evolution of the CYP1 family, we applied non-synonymous per synonymous nucleotide substitutions ($\omega = dN/dS$). To analyze the selective pressures, we employed the ML method in the CODEML program of PAML software package 4.4 [12, 13]. In CODEML, the site-models were initially utilized to allow heterogeneous ω across the sites, however, the ω ratios were constant among branches. Furthermore, the explanation was sorted out for the diverse structural and functional constraints experienced through specific site domains [14]. Further, the branch-site models were employed to identify positive selection sites among various species lineages.

The crowd validation

The positive selection sites were changed to DNA sites from searching the SNP sites of the CYP1 family in the SNP database (<https://www>).

Ovarian cancer risk with CYP1B1 come from the positive selection

ncbi.nlm.nih.gov/snp/). The sites with MAF > 0.05 in 1000 genomes website (<http://phase3browser.1000genomes.org/index.html>) were then selected. Only rs1056836 and rs1056827 of CYP1B1 matched the condition.

Meta-analysis

Investigation of eligible studies: A comprehensive systematic literature search was conducted in Cochrane Library, PubMed, Web of Science, EMBASE, Chinese Biomedical Literature Database (SinoMed), Chinese National Knowledge Infrastructure (CNKI), and Wan-Fang database for randomized controlled trials about the rs1056836 and rs1056827 of CYP1B1. Statistical analyses were conducted by using Stata software (version 14.0) and RevMan software (version 5.3).

Inclusion criteria: For study selection, the following inclusion criteria were utilized: (i) case-control study; (ii) The data of genotype frequencies were accessible for both cases as well as controls (iii) published up to 1st October 2020 (iv) full-text articles; (v) literature published in both English and Chinese languages. The authors checked the literature via the exact test.

Data extraction: For the data extraction, the information collected from papers included the name of the authors, publication year, number of total cases as well as control, allele frequencies of the CYP1B1 rs1056827, and rs1056836 polymorphisms, country, and ethnicity. Two authors executed the search independently. In this study, disagreement regarding the eligibility criteria of any paper was resolved via the evaluation by a third reviewer and discussion till reaching a consensus.

Statistical analysis: Odds ratios (ORs) with 95% confidence intervals (CIs) were used for evaluating the strength and size of the association between CYP1B1 polymorphisms and the risk of ovarian carcinoma. Alternatively, OR were obtained from the data of genotype frequencies, and were transformed logarithmically for obtaining normality. OR was measured for homozygous carriers versus 'wild type' and heterozygous carriers versus 'wild type'. Furthermore, OR were measured for the recessive model, i.e., homozygous carriers versus 'wild type' and heterozygous carriers, and the

dominant model, i.e., heterozygous and homozygous carriers versus 'wild type'. Heterogeneity existing between the studies was evaluated via a chi-square-based Q-test and through estimation of I^2 , accordingly. Heterogeneity was considered in the case of P -value < 0.05, and when the DerSimonian Laird random-effects model was fitted. Otherwise, the standard Mantel-Haenszel fixed-effects model was adjusted.

The publication bias was determined via Egger's linear regression method. For the visual inspection of the funnel plot (plots of effect estimate against sample size), asymmetry was carried out. For publication bias, P -value < 0.05 was regarded as statistically significant. Statistical analyses for the data were conducted via statistical software, the stata 17.0. The P -values were for a two-sided test, the P -value < 0.05 was regarded as statistically significant.

Genotyping assay

Patient enrollment and sample collection and processing: In this study, 199 tissue samples of patients with stage I-IV ovarian serous adenocarcinoma were selected from the tissue sample bank of the gynecology and Oncology Department of Affiliated Cancer Hospital of Guangxi Medical University from January 2011 till July 2017. The Ethics Committee of The Affiliated Tumor Hospital of Guangxi Medical University approved the study. All patients received an explanation about the aims of the study and provided signed informed consent. All patients had undergone cytoreductive surgeries and the diagnosis of ovarian serous adenocarcinoma and had been confirmed by two pathologists. Patients were administered platinum-paclitaxel chemotherapy for no less than 4 cycles after surgery. All specimens were divided into two groups: sensitive group (137 cases) and resistant group (62 cases), including 50 cases from fresh frozen tissue, and 149 cases from paraffin tissue. The classification of drug sensitivity and resistance is based on NCCN guidelines: for patients receiving the first treatment, recurrence within half a year after 4-6 months of regular chemotherapy is considered as drug resistance, recurrence after more than half a year is considered as sensitive. In addition, if the disease continues to progress during chemotherapy, it is called

Ovarian cancer risk with CYP1B1 come from the positive selection

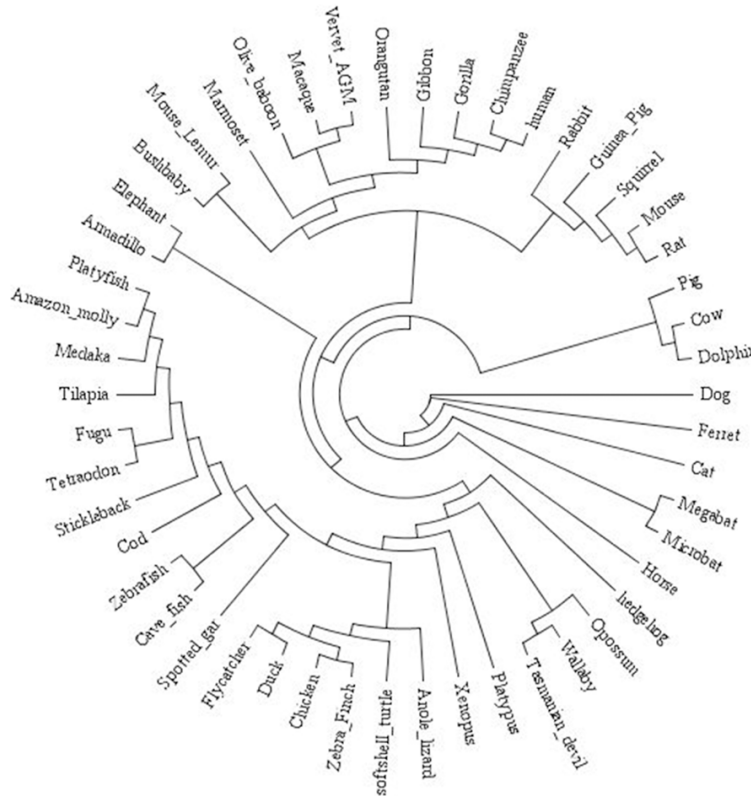


Figure 2. Phylogenetic tree of CYP1A2 gene.

“uncontrolled” and is considered resistant. The fresh frozen tissue was obtained from the tumor lesion during the operation, which was immediately placed into the cryopreservation tube and stored in the liquid nitrogen tank. The sensitive group was the ovarian epithelial carcinoma tissue sensitive to platinum, and the drug-resistant group was the ovarian epithelial carcinoma tissue resistant to platinum. The preserved FFPE tumor blocks were taken from the Affiliated Tumor Hospital of Guangxi Medical University and kept at $\sim 25^{\circ}\text{C}$ till the processing. Nuclear DNA extraction was performed from FFPE using the GeneJET FFPE DNA Purification Kit (Thermo Scientific, K0881). After the DNA extraction, quantification of DNA in each sample was carried out using nanodrop. The DNA concentration in all samples was in the concentration range from 50 to 500 $\text{ng}/\mu\text{l}$, which fulfilled the need for Sequenom MassARRAY SNP. Extracted DNA was kept at -80°C till the genotyping was carried out. The identification of the genotype was carried out with Sequenom MassARRAY SNP. Fisher’s exact test and Chi-square test were carried out

for the comparison of the variations between the two groups.

Genotyping: The identification for the genotype of CYP1B1 gene rs1056836 was carried out with Sequenom MassARRAY SNP. The allelic frequency for the desired gene was calculated according to the genotypes. For the PCR setup, 2.0 μl of DNA extract (50 to 100 $\text{ng}/\mu\text{l}$) was performed in a 25 μl reaction mixture containing a standard PCR buffer, Taq DNA polymerase (1.0 U), dNTPs (200 μM), MgCl_2 (1.5 mM), and primers (0.4 μM). The PCR setup for the denaturation of DNA included 1 and 35 cycles at 94°C for 5 minutes and 30 s, accordingly. For annealing of primer, the program was as 58°C for 30 seconds and extension at 72°C for 1 minute and 10 minutes, respectively. The primary amplification

primer sequence was ACGTTGGATGTTGTCAACCAGTGGTCTGTG. The secondary amplification primer sequence was ACGTTGGATGGCCATCCTTGTCCTCAAGAATC. The single-base extension primer sequence was GGGTTAGCCACTTCA.

Statistical methods: Fisher’s exact test and Chi-square test were employed to evaluate the differences between each group regarding allelic frequencies and their genotype. Unconditional logistic regression was employed for comparing the odds ratio (OR) and P values to identify the correlation between gene polymorphism and risk of three kinds of tumors. Statistical analysis was carried out via SPSS version 18.0 (USA), and the $P < 0.05$ was regarded as statistically significant.

Results

Positive selection

A phylogenetic tree based on the sequence of amino acids among several Metazoa species was generated for establishing the association of their CYP1 family enzymes (see **Figures 1-3**).

Ovarian cancer risk with CYP1B1 come from the positive selection

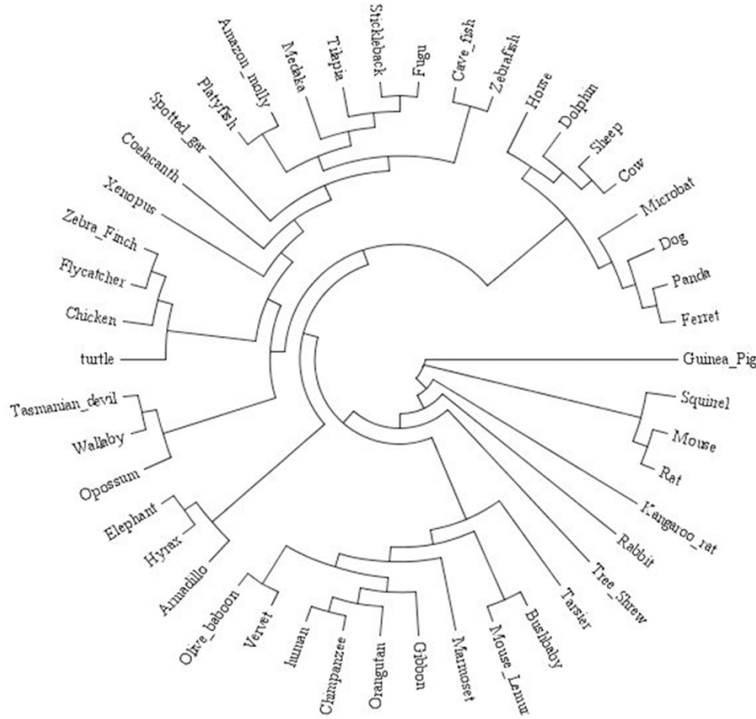


Figure 3. Phylogenetic tree of CYP1B1 gene.

Initially, in the site model, results obtained from the LRT test statistic ($2\Delta\ln L$) of M7-M8 comparisons were 8000.49129 (CYP1A1), 7133.20827 (CYP1A2), and 6608.43877 (CYP1B1). Then the BEB approach was implemented and 31, 33, and 11 sites were detected within the three genes under the positive selection with BPP values > 0.95 (**Table 1**). Secondly, the one-ratio model revealed that the whole CYP1 family genes had gone through the purifying selection. The free-ratio model was analyzed and co-related with the one-ratio model, the obtained results revealed that every branch had its independent u value ($P < 0.01$, see **Table 2**). The branch-site models were used to evaluate positively selected sites in the ancestral lineages. In consequence, 1, 2, and 2 positive sites were detected in CYP1A1, CYP1A2, and CYP1B1, respectively, when branch-site models were used (**Table 3**).

The SNP connection and the positive selection sites

All the positive selection sites were investigated in the SNP database (<https://www.ncbi.nlm.nih.gov/projects/SNP/>) and found 57 SNPs corresponding to them (**Table 4**). The

MAF of rs1056836 and rs-1056827 of CYP1B1 is 0.39 and 0.36, other SNPs are < 0.01 . So, we chose these two SNPs for Meta-analyses to investigate whether they are related to the risk of ovarian cancer.

Meta analysis result

Characteristics of this polymorphism study: In total, 1,655 articles were first retrieved from the database searching until 1st October 2020. Then, only 31 articles were obtained and selected after the duplicated publications were weeded out. Among them, 8 articles were abstracts, review articles, editorials, or meta-analyses, and 16 articles were related to other polymorphisms and diseases or had incomplete genotype data.

Hence, we further checked and excluded unwanted articles from the study following the criteria set for exclusion. Lastly, only 7 articles fulfilled the requirements of our study and were included in the meta-analysis [15-21]. Included seven studies associated with CYP1B1 rs1056836 polymorphism and the other three studies associated with CYP1B1 rs1056827 polymorphism [17, 18, 21]. The related parameters and the baseline feature of the included studies are listed in **Table 5**.

Meta-Analysis of CYP1B1 rs1056836 and rs1056827 polymorphism with a risk of ovarian carcinoma: Meta-analysis was conducted for a total of 7 studies containing 1285 patients samples and 2660 controls to examine the relations between rs1056836 and ovarian cancer risk (shown in **Table 6**). Three studies involving 852 patients' samples and 1202 controls investigated the relations between rs1056827 and the ovarian cancer risk (shown in **Table 7**). In general, collectively our data showed that both rs1056836 and rs1056827 polymorphism in genes have no link with a higher risk of ovarian carcinoma in the six genetic models (**Table 8**).

Ovarian cancer risk with CYP1B1 come from the positive selection

Table 1. Site model tests of Metazoa on CYP1 genes subset

| Gene | Model | lnL | Compared models | np | df | LRT (2ΔlnL) | p value | ω | Positive selected sites |
|--------|-------|--------------|-----------------|-----|----|-------------|---------|--------|--|
| CYP1A1 | M3 | -27301.23484 | | 103 | | | | 0.1696 | |
| | M0 | -28218.98883 | M3-M0 | 99 | 4 | 1835.50798 | 0 | 0.1501 | none |
| | M2a | -27796.90516 | | 102 | | | | 0.2771 | |
| | M1a | -27796.90516 | M2a-M1a | 100 | 2 | 0 | 1 | 0.2771 | none |
| | M8 | -31268.52748 | | 102 | | | | 0.3537 | 6S, 32Q, 35K, 39N, 147I, 172S, 180G, 181P, 217L, 221N, 228V, 246P, 247S, 249N, 250A, 256E, 259Y, 260S, 262M, 271K, 339V, 356R, 363S, 426E, 435P, 442V |
| | M7 | -27268.28184 | M8-M7 | 100 | 2 | 8000.49129 | 0 | 0.1808 | 465W, 483P, 509Q, 510L, 512S |
| CYP1A2 | M3 | -26990.93851 | | 105 | 4 | 1619.60288 | 0 | 0.1522 | |
| | M0 | -27800.73995 | M3-M0 | 101 | | | | 0.1323 | none |
| | M2a | -27550.1175 | | 104 | 2 | 0 | 1 | 0.2474 | |
| | M1a | -27550.1175 | M2a-M1a | 102 | | | | 0.2474 | none |
| | M8 | -30536.17696 | | 104 | 2 | 7133.20827 | 0 | 0.2278 | 5Q, 6S, 19A, 34R, 37K, 41S, 67R, 71R, 149I, 174S, 185H, 190N, 212S, 224H, 229T, 248P, 249A, 252R, 255A, 258Q, 261L, 262W, 267T, 273Q, 300N, 395T, 425S, 434T, 435A, 443P, 444L, 484P |
| | M7 | -26969.57282 | M8-M7 | 102 | | | | 0.1587 | 488K |
| CYP1B1 | M3 | -22636.80457 | | 95 | 4 | 1302.626 | 0 | 0.1256 | none |
| | M0 | -23288.11757 | M3-M0 | 91 | | | | 0.1117 | |
| | M2a | -23104.02312 | | 94 | 2 | 0 | 1 | 0.1799 | none |
| | M1a | -23104.02312 | M2a-M1a | 92 | | | | 0.1799 | |
| | M8 | -25929.03736 | | 94 | 2 | 6608.43877 | 0 | 0.3523 | 136H, 140H, 179A, 194R, 284R, 317L, 367D, 432L, 435P, 502K |
| | M7 | -22624.81797 | M8-M7 | 92 | | | | 0.1296 | 517K |

Subgroup analysis: The findings of the subgroup analysis by ethnicity in African-American, Caucasians, and Asians for all six genetic models are shown in **Figures 4, 5**. The Heterozygous model (GC versus GG: OR = 1.29, 95% CI: 1.03-1.61, P = 0.03), Homozygote model (CC versus GG: OR = 1.30, 95% CI: 1.02-1.65, P = 0.03), and Dominant model (CC+GC versus GG: OR = 1.47, 95% CI: 1.05-1.60, P = 0.016) established the link of rs1056836 polymorphism with ovarian cancer risk in Caucasians ethnicity. The Allele model (G versus C: OR = 2.71, 95% CI: 1.39-5.31), Homozygote model (CC versus GG: OR = 0.22, 95% CI: 0.06-0.77), and Recessive (GG versus GC+CC: OR = 0.22, 95% CI: 0.09-0.59), recognized the co-relations of rs1056836 polymorphism with ovarian cancer risk among Asians. However, all the six genetic models showed no association among African-American. All the six genetic models showed no relation of rs1056827 polymorphism with ovarian cancer risk among African-Americans, Caucasians, and Asians.

To evaluate the publication bias of for included studies, Egger's test was performed. The results for all models (allent model: P = 0.245; dominant model: P = 0.138; recessive model: P = 0.274; additive model: P = 0.224; Homozygous model: P = 0.200; Heterozygous model: P = 0.427) revealed no evidence of publication bias in the meta-analysis of rs1056836. However, two models of The Egger's test results pinpointed publication bias of rs1056827 (allent model: P = 0.718; dominant model: P = 0.151; recessive model: P = 0.277; additive model: P = 0.027; Homozygous model: P = 0.418; Heterozygous model: P = 0.049) in the meta-analysis.

Individual polymorphisms co-relation with drug-resistant of ovarian cancer

To find a link of polymorphism with drug-Resistant in ovarian cancer, the genotype of CYP1B1 (rs1056836) was analyzed via Sequenom MassARRAY SNP technology. We evaluated

Ovarian cancer risk with CYP1B1 come from the positive selection

Table 2. Branch model tests in a subset of Metazoa on CYP1 genes

| Gene | Model | lnL | Compared models | np | df | LRT (2ΔlnL) | p value | ω |
|--------|--------------|--------------|-----------------|-----|----------|-------------|----------------------------|----------------------------|
| CYP1A1 | Fr | -28037.12928 | M0/Fr | 195 | 96 | 363.71911 | 0.00E+00 | |
| | MO | -28218.98883 | | 99 | | | | |
| | Ta | -28215.07649 | M0/Ta | 100 | 1 | 73824688 | 0.005154 | ω0 = 0.1495, ωa = 3.3534 |
| | Tb | -28218.23963 | M0/Tb | 100 | 1 | 1.50E+00 | 0.220919 | ω0 = 0.1494, ωa = 0.4635 |
| | Tc | -28214.53803 | M0/Tc | 100 | 1 | 8.901592 | 0.00285 | ω0 = 0.1488, ωa = 999.0000 |
| | Td | -28218.92125 | M0/Td | 100 | 1 | 1.35E-01 | 0.713147 | ω0 = 0.1498, ωa = 0.2003 |
| | Te | -28218.98455 | M0/Te | 100 | 1 | 0.008556 | 0.926302 | ω0 = 0.1500, ωa = 0.1588 |
| | Tf | -28217.76654 | M0/Tf | 100 | 1 | 2.44E+00 | 0.117932 | ω0 = 0.1499, ωa = 999.0000 |
| CYP1A2 | Fr | -27583.02458 | M0/Fr | 199 | 98 | 435.430748 | 0.00E+00 | |
| | MO | -27800.73995 | | 101 | | | | |
| | Ta | -27799.46573 | M0/Ta | 102 | 1 | 2.54844 | 0.110404 | ω0 = 0.1330, ωa = 0.0473 |
| | Tb | -27798.74531 | M0/Tb | 102 | 1 | 3.99E+00 | 0.045791 | ω0 = 0.1316, ωa = 999.0000 |
| | Tc | -27792.14895 | M0/Tc | 102 | 1 | 17.182014 | 3.40E-05 | ω0 = 0.1300, ωa = 5.2984 |
| | Td | -27798.61324 | M0/Td | 102 | 1 | 4.25E+00 | 0.0391 | ω0 = 0.1316, ωa = 999.0000 |
| | Te | -27800.27495 | M0/Te | 102 | 1 | 0.930009 | 0.334861 | ω0 = 0.1322, ωa = 0.5119 |
| | Tf | -27798.71841 | M0/Tf | 102 | 1 | 4.04E+00 | 0.044353 | ω0 = 0.1321, ωa = 1.0464 |
| CYP1B1 | Tg | -27797.54368 | M0/Tg | 102 | 1 | 6.39E+00 | 0.011461 | ω0 = 0.1317, ωa = 999.0000 |
| | Fr | -23145.36942 | M0/Fr | 179 | 88 | 285.496287 | 0.00E+00 | |
| | MO | -23288.11757 | | 91 | | | | |
| | Ta | -23289.23363 | M0/Ta | 92 | 1 | 2.232128 | 0.135168 | ω0 = 0.1114, ωa = 6.4575 |
| | Tb | -23285.75435 | M0/Tb | 92 | 1 | 4.73E+00 | 0.029703 | ω0 = 0.1110, ωa = 0.9007 |
| | Tc | -23287.41432 | M0/Tc | 92 | 1 | 1.406486 | 0.235641 | ω0 = 0.1115, ωa = 999.0000 |
| | Td | -23282.36142 | M0/Td | 92 | 1 | 1.15E+01 | 6.92E-04 | ω0 = 0.1106, ωa = 1.2861 |
| | Te | -23284.89967 | M0/Te | 92 | 1 | 6.435791 | 0.011185 | ω0 = 0.1114, ωa = 999.0000 |
| | Tf | -23287.15559 | M0/Tf | 92 | 1 | 1.92E+00 | 0.165423 | ω0 = 0.1116, ωa = 999.0000 |
| | Tg | -23288.43365 | M1/Tg | 92 | 1 | 6.32E-01 | 0.426562 | ω0 = 0.1116, ωa = 2.6526 |
| | Th | -23288.04399 | M2/Th | 92 | 1 | 1.47E-01 | 0.701278 | ω0 = 0.1119, ωa = 0.0772 |
| Ti | -23283.84198 | M3/Ti | 92 | 1 | 8.55E+00 | 0.003453 | ω0 = 0.1103, ωa = 999.0000 | |
| Tj | -23288.10274 | M4/Tj | 92 | 1 | 2.97E-02 | 0.863269 | ω0 = 0.1117, ωa = 0.0981 | |
| Tk | -23289.04418 | M5/Tk | 92 | 1 | 1.85E+00 | 0.173409 | ω0 = 0.1116, ωa = 1.9973 | |

variations in the Platinum resistance group and the Platinum sensitive group of CYP1B1 (rs1056836) distribution. The statistical analysis on the CYP1B1 (rs1056836) genotype distribution between the Platinum resistance group and the Platinum sensitive group did not show obvious differences (Table 9).

Discussion

The theory of molecular evolution originated from Darwin's theory of evolution, which refers to the evolution of biological macromolecules in the process of evolution to adapt to environmental changes [22]. The molecular evolutionary analysis employs statistical, mathematical, and computer science methods to study molecular sequence data of DNA and proteins, using

sequence similarity, evolutionary rate, divergence time, and phylogenetic tree to estimate molecular affinity [23]. Besides, the positive selection site analysis of adaptive evolutionary protein sequences can provide information for the identification of important functional sites and amino acid structures. In recent years, comparative genomics and molecular phylogenetic analysis have been developed rapidly and are widely used in the classification of gene families, the incidence and growth of tumors, the evolution of viruses and immune escape, and the study of mechanisms [24-27]. In 1976, Nowell proposed a clonal evolution model, that is, cancer originates from a single mutant normal cell. The gene mutations that occur during the subsequent amplification process gradually enhance the survival advantage of the tumor

Ovarian cancer risk with CYP1B1 come from the positive selection

Table 3. Branch-site model tests in a subset of Metazoa on CYP1 genes

| | | Model | InL | np | df | LRT (2ΔInL) | p value | Positive selected sites |
|--------|--------------|--------------|--------------|-----|----------|-------------|----------|-------------------------|
| CYP1A1 | Ta | Model A | -27796.22707 | 102 | | | | |
| | | Model A null | -27796.41441 | 101 | 1 | 0.374674 | 0.540468 | none |
| | Tc | Model A | -27791.46117 | 102 | | | | 398G |
| | | Model A null | -27794.66746 | 101 | 1 | 6.412593 | 0.011332 | |
| CYP1A2 | Tb | Model A | -27538.26421 | 104 | 1 | 18.321374 | 1.90E-05 | 184G, 418P |
| | | Model A null | -27547.4249 | 103 | | | | |
| | Tc | Model A | -27541.84596 | 104 | 1 | 4.61194 | 0.031751 | none |
| | | Model A null | -27544.15193 | 103 | | | | |
| | Td | Model A | -27548.08974 | 104 | 1 | 4.033103 | 0.044616 | none |
| | | Model A null | -27550.10629 | 103 | | | | |
| | Tf | Model A | -27549.66766 | 104 | 1 | 0.191042 | 0.662051 | none |
| | | Model A null | -27549.76318 | 103 | | | | |
| | Ti | Model A | -27544.84393 | 104 | 1 | 1.999217 | 0.157381 | none |
| | | Model A null | -27545.84354 | 103 | | | | |
| CYP1B1 | Td | Model A | -23095.44335 | 94 | 1 | 2.376804 | 0.12315 | 65A |
| | | Model A null | -23096.63175 | 93 | | | | |
| | Te | Model A | -23102.86315 | 94 | 1 | 0.82245 | 0.364465 | 119A |
| | | Model A null | -23103.27438 | 93 | | | | |
| Ti | Model A | -23100.02028 | 94 | 1 | 1.933311 | 0.164397 | none | |
| | Model A null | -23100.98693 | 93 | | | | | |

Table 4. The SNPs corresponding to the positive selection sites

| gene | positive selection site | SNP | Change of basic group | Change of amino acid | MAF |
|--------|-------------------------|-------------|-----------------------|----------------------|--------|
| CYP1A1 | 6S | --- | --- | --- | --- |
| CYP1A1 | 32Q | --- | --- | --- | --- |
| CYP1A1 | 35K | --- | --- | --- | --- |
| CYP1A1 | 39N | --- | --- | --- | --- |
| CYP1A1 | 116S | --- | --- | --- | --- |
| CYP1A1 | 147I | rs960353942 | As9 | Is9 | < 0.01 |
| CYP1A1 | 172S | --- | --- | --- | --- |
| CYP1A1 | 180G | rs775144476 | Gs7 | Gs7 | < 0.01 |
| CYP1A1 | 181P | rs766480120 | Gs7 | Ps7 | < 0.01 |
| CYP1A1 | 217L | --- | --- | --- | --- |
| CYP1A1 | 221N | --- | --- | --- | --- |
| CYP1A1 | 228V | --- | --- | --- | --- |
| CYP1A1 | 246P | rs760213300 | Cs7 | Ps7 | < 0.01 |
| CYP1A1 | 247S | --- | --- | --- | --- |
| CYP1A1 | 249N | rs771648532 | As7 | Ns7 | < 0.01 |
| CYP1A1 | 249N | rs749731774 | Ts7 | Ns7 | < 0.01 |
| CYP1A1 | 250A | --- | --- | --- | --- |
| CYP1A1 | 256E | rs746238922 | Gs7 | Es7 | < 0.01 |
| CYP1A1 | 259Y | --- | --- | --- | --- |
| CYP1A1 | 260S | --- | --- | --- | --- |
| CYP1A1 | 262M | rs757497550 | As7 | Ms7 | < 0.01 |
| CYP1A1 | 271K | --- | --- | --- | --- |
| CYP1A1 | 339V | --- | --- | --- | --- |

Ovarian cancer risk with CYP1B1 come from the positive selection

| | | | | | |
|--------|------|--------------|--------|-----|--------|
| CYP1A1 | 356R | rs368650547 | Gs3 | Rs3 | < 0.01 |
| CYP1A1 | 356R | rs757872883 | Gs7 | Rs7 | < 0.01 |
| CYP1A1 | 363S | rs759604228 | Cs7 | Ss7 | < 0.01 |
| CYP1A1 | 426E | rs72547510 | -s7 | Es7 | < 0.01 |
| CYP1A1 | 435P | -- | -- | -- | -- |
| CYP1A1 | 442V | -- | -- | -- | -- |
| CYP1A1 | 465W | rs774927292 | (7bp)→ | W7b | < 0.01 |
| CYP1A1 | 483P | rs45500996 | Cs4 | Ps4 | < 0.01 |
| CYP1A1 | 509Q | -- | -- | -- | -- |
| CYP1A1 | 510L | rs578153711 | Gs5 | Ls5 | < 0.01 |
| CYP1A1 | 512S | rs779920114 | Ts7 | Ss7 | < 0.01 |
| CYP1A1 | 398G | -- | -- | -- | -- |
| CYP1A2 | 5Q | -- | -- | -- | -- |
| CYP1A2 | 6S | rs764359866 | Cs7 | Ss7 | < 0.01 |
| CYP1A2 | 19A | rs746300395 | TG74 | AG7 | < 0.01 |
| CYP1A2 | 19A | rs1047643383 | Gs1 | As1 | < 0.01 |
| CYP1A2 | 19A | rs771691950 | Cs7 | As7 | < 0.01 |
| CYP1A2 | 34R | rs201934979 | Cs2 | Rs2 | < 0.01 |
| CYP1A2 | 34R | rs941101723 | Gs9 | Rs9 | < 0.01 |
| CYP1A2 | 37K | rs751602658 | As7 | Ks7 | < 0.01 |
| CYP1A2 | 41S | -- | -- | -- | -- |
| CYP1A2 | 67R | rs760996321 | As7 | Rs7 | < 0.01 |
| CYP1A2 | 71R | rs755565165 | Cs7 | Rs7 | < 0.01 |
| CYP1A2 | 71R | rs779505412 | Gs7 | Rs7 | < 0.01 |
| CYP1A2 | 149I | -- | -- | -- | -- |
| CYP1A2 | 174S | rs1002036121 | Cs1 | Ss1 | < 0.01 |
| CYP1A2 | 185H | rs147248980 | C/G47 | H/G | < 0.01 |
| CYP1A2 | 190N | -- | -- | -- | -- |
| CYP1A2 | 212S | rs758748797 | As7 | Ss7 | < 0.01 |
| CYP1A2 | 224H | rs773105304 | Cs7 | Hs7 | < 0.01 |
| CYP1A2 | 229T | rs755082460 | Cs7 | Ts7 | < 0.01 |
| CYP1A2 | 248P | -- | -- | -- | -- |
| CYP1A2 | 249A | rs575035489 | Cs5 | As5 | < 0.01 |
| CYP1A2 | 252R | rs377683245 | Gs3 | Rs3 | < 0.01 |
| CYP1A2 | 255A | -- | -- | -- | -- |
| CYP1A2 | 258Q | rs775258473 | Cs7 | Qs7 | < 0.01 |
| CYP1A2 | 261L | -- | -- | -- | -- |
| CYP1A2 | 262W | -- | -- | -- | -- |
| CYP1A2 | 267T | rs759584175 | Cs7 | Ts7 | < 0.01 |
| CYP1A2 | 267T | rs756704382 | As7 | Ts7 | < 0.01 |
| CYP1A2 | 273Q | -- | -- | -- | -- |
| CYP1A2 | 300N | -- | -- | -- | -- |
| CYP1A2 | 395T | rs911014711 | As9 | Ts9 | < 0.01 |
| CYP1A2 | 395T | rs149928755 | Cs1 | Ts1 | < 0.01 |
| CYP1A2 | 395T | rs915586110 | Gs9 | Ts9 | < 0.01 |
| CYP1A2 | 425S | rs780344039 | Cs7 | Ss7 | < 0.01 |
| CYP1A2 | 425S | rs969721209 | Ts9 | Ss9 | < 0.01 |
| CYP1A2 | 434T | rs770969636 | As7 | Ts7 | < 0.01 |
| CYP1A2 | 434T | rs200295985 | Cs2 | Ts2 | < 0.01 |

Ovarian cancer risk with CYP1B1 come from the positive selection

| | | | | | |
|--------|------|-------------|-----|-----|--------|
| CYP1A2 | 435A | rs745999166 | Cs7 | As7 | < 0.01 |
| CYP1A2 | 443P | rs889339647 | Cs8 | Ps8 | < 0.01 |
| CYP1A2 | 444L | rs140757511 | Gs1 | Ls1 | < 0.01 |
| CYP1A2 | 484P | rs370476434 | Cs3 | Ps3 | < 0.01 |
| CYP1A2 | 484P | rs148212809 | Gs1 | Ps1 | < 0.01 |
| CYP1A2 | 488K | -- | -- | -- | -- |
| CYP1A2 | 184G | -- | -- | -- | -- |
| CYP1A2 | 418P | -- | -- | -- | -- |
| CYP1B1 | 136H | rs538072907 | Cs5 | HsN | < 0.01 |
| CYP1B1 | 140H | -- | -- | -- | -- |
| CYP1B1 | 179A | rs771076928 | Gs7 | As7 | < 0.01 |
| CYP1B1 | 194R | -- | -- | -- | -- |
| CYP1B1 | 284R | rs368249322 | Cs3 | Rs3 | < 0.01 |
| CYP1B1 | 317L | rs199836011 | Gs1 | Ls1 | < 0.01 |
| CYP1B1 | 367D | rs916479886 | Gs9 | Ds9 | < 0.01 |
| CYP1B1 | 432L | rs1056836 | Cs1 | Ls1 | 0.39 |
| CYP1B1 | 435P | rs113974874 | Cs1 | Ps1 | < 0.01 |
| CYP1B1 | 502K | -- | -- | -- | -- |
| CYP1B1 | 517K | -- | -- | -- | -- |
| CYP1B1 | 65A | rs554295550 | Gs5 | As5 | < 0.01 |
| CYP1B1 | 119A | rs1056827 | Gs1 | As1 | 0.36 |

Table 5. Characteristics of CYP1B1 rs1056836 and rs1056827 polymorphism studies and the risk for ovarian cancer

| site | Study | year | Country | Ethnicity | Case | Control | P for HWE |
|-----------|------------|------|---------|--------------------------------|------|---------|-----------|
| rs1056836 | Goodman | 2001 | America | Caucasian, Asian, Other | 128 | 144 | Yes |
| | Cecchin | 2004 | Italy | Caucasian | 220 | 280 | yes |
| | Sellers, a | 2005 | America | Caucasian | 454 | 545 | yes |
| | Sellers, b | 2005 | America | African American | 36 | 53 | NO |
| | ZHU ZY | 2006 | China | Asian | 53 | 30 | yes |
| | Holt, a | 2007 | America | Caucasian | 277 | 447 | yes |
| | Holt, b | 2007 | America | African-American | 33 | 127 | yes |
| | Delort | 2008 | France | Caucasian | 51 | 1,000 | yes |
| | Zahid | 2014 | America | African-American and Caucasian | 33 | 34 | yes |
| rs1056827 | Sellers, a | 2005 | America | Caucasian | 453 | 543 | yes |
| | Sellers, b | 2005 | America | African American | 36 | 53 | no |
| | ZHU ZY | 2006 | China | Asian | 53 | 30 | no |
| | Holt, a | 2007 | America | Caucasian | 277 | 450 | yes |
| | Holt, b | 2007 | America | African-American | 33 | 126 | no |

cell population, and eventually lead to the rapid proliferation of aneuploidy, thereby triggering Heterogeneity. This model proposes the mechanism of cancer and heterogeneity from an evolutionary perspective [28]. The poor treatment effect and high recurrence rate of ovarian cancer may be related to the high heterogeneity of cancer cells, which come from an evolutionary perspective. The most commonly

mutated genes in ovarian cancer include p53 and BRCA1/BRCA2. KHAN had conducted an evolutionary analysis of the p53 gene in 26 mammals, showing multiple amino acid positions point under positive selection pressure [29]. LOU used PAML software to analyze the BRCA1 gene of 23 primates. Under the selection pressure analysis, the site model shows that the 10 amino acid sites of the gene are

Ovarian cancer risk with CYP1B1 come from the positive selection

Table 6. Genotype distribution of rs1056836 polymorphism among ovarian cancer cases and controls

| Research | case | | | | Control | | | |
|-----------|------|-----|-----|------|---------|-----|-----|------|
| | GG | GC | CC | HWE | GG | GC | CC | HWE |
| Goodman | 13 | 48 | 67 | 0.62 | 7 | 42 | 95 | 0.71 |
| Cecchin | 35 | 126 | 59 | 0.06 | 57 | 139 | 84 | 1.0 |
| Sellers a | 84 | 230 | 140 | 0.83 | 110 | 269 | 166 | 0.99 |
| Sellers b | 23 | 0 | 13 | 0.00 | 31 | 0 | 22 | 0.00 |
| Zhu ZY | 16 | 25 | 12 | 0.93 | 5 | 8 | 17 | 0.14 |
| Holt, a | 41 | 129 | 107 | 0.98 | 90 | 216 | 141 | 0.91 |
| Holt, b | 18 | 11 | 4 | 0.56 | 80 | 39 | 8 | 0.56 |
| Delort | 6 | 27 | 18 | 0.69 | 203 | 475 | 322 | 0.52 |
| Zahid | 10 | 14 | 9 | 0.69 | 6 | 15 | 12 | 0.94 |

Table 7. Genotype distribution of the rs1056827 polymorphism among ovarian cancer cases and controls

| Research | Case | | | | Control | | | |
|-----------|------|-----|----|-----|---------|-----|-----|-----|
| | GG | GT | TT | HWE | GG | GT | TT | HWE |
| Sellers a | 233 | 178 | 42 | YES | 110 | 269 | 166 | YES |
| Sellers b | 6 | 30 | 0 | NO | 16 | 37 | 0 | NO |
| ZHU ZY | 16 | 25 | 12 | YES | 5 | 8 | 17 | YES |
| Holt, a | 131 | 116 | 30 | YES | 222 | 188 | 40 | YES |
| Holt, b | 12 | 15 | 6 | NO | 49 | 50 | 27 | NO |

under obvious positive selection pressure. The results of the branch model suggest that the positive selection pressure is more obvious on the three branches of human bonobos. The study also performed the same analysis on the BRCCA2 gene, showing that it is also under obvious positive selection pressure [30]. There is no report on the evolutionary analysis of the CYP1 family in metazoans at present. In our research, we revealed 57 positive selection sites in the CYP1 family. Particularly, the MAF of rs1056836 and rs1056827 of CYP1B1 is 0.39 and 0.36. It showed that CYP1B1 was under obvious positive selection pressure. The theory of molecular evolution helps to explain the mechanism of drug resistance of cancer cells in a new direction, to put forward more targeted and effective treatment measures.

Genetic factors significantly contribute to the tendency of ovarian carcinoma development. Though the causal genetic links in a pathological state largely remain unknown, it is assumed that changes in CYP1B1 polymorphic variants in their function to steroid hormones and procarcinogens may enhance the susceptibility to estrogen-dependent carcinomas, for exam-

ple, breast and endometrial cancer [31, 32]. However, the essential role of CYP1B1 polymorphism in ovarian cancer still needs to be explored. On chromosome 2p22-21, the locus of the CYP1B1 gene is present that contains three exons. Human CYP1B1 consists of three exons in size of 371, 1044, and 3707 bp in length, and having two introns in size of 390 and 3032 bp spanning 8.5 kbp of genomic DNA (GenBank accession no. U56438). The coding region starts at the 5'-end of the second exon and ends with the last exon. CYP1B1 is a crucial enzyme, which contributes to the metabolism of exogenous as well as endogenous substrates, many of which include carcinogenic compounds. Gene polymorphisms of CYP1B1 have been characterized in different cancers. The four most frequent polymorphisms of the CYP1B1 occur due to the substitutions in amino acids. The examples include Leu432Val, Asn453Ser, Arg48Gly, and Ala119Ser. P450 1B1 is an efficient catalyst for estrogen hydroxylation. It catalyzes the 4-hydroxylation reaction of 17 β -estradiol (E2) that produces the less active metabolite, 4-hydroxyestradiol.

4-Hydroxyestradiol is believed to cause estrogen-dependent tumors [33]. The CYP1B1 polymorphism is the most relevant one for 4-hydroxyestradiol production which is characterized by a C to G conversion in codon 432, thus leading to the substitution of a Leu with a Val in the protein. The mutated allele is associated with higher efficiency of 4-hydroxyestradiol production. Women carrying the 432GG variant are presumed to be more exposed to the carcinogenic effects of 4-hydroxyestradiol, hence at higher risk of developing estrogen-related cancer [34]. Several epidemiological studies have been carried out to evaluate the correlation between CYP1B1 polymorphisms and the risk of ovarian cancer. Though, the obtained results from these reports were erratic or contradictory [16, 18, 19]. However, the subgroup analysis of this meta-analysis showed the association of rs1056836 polymorphism with ovarian cancer risk among Caucasians and Asians.

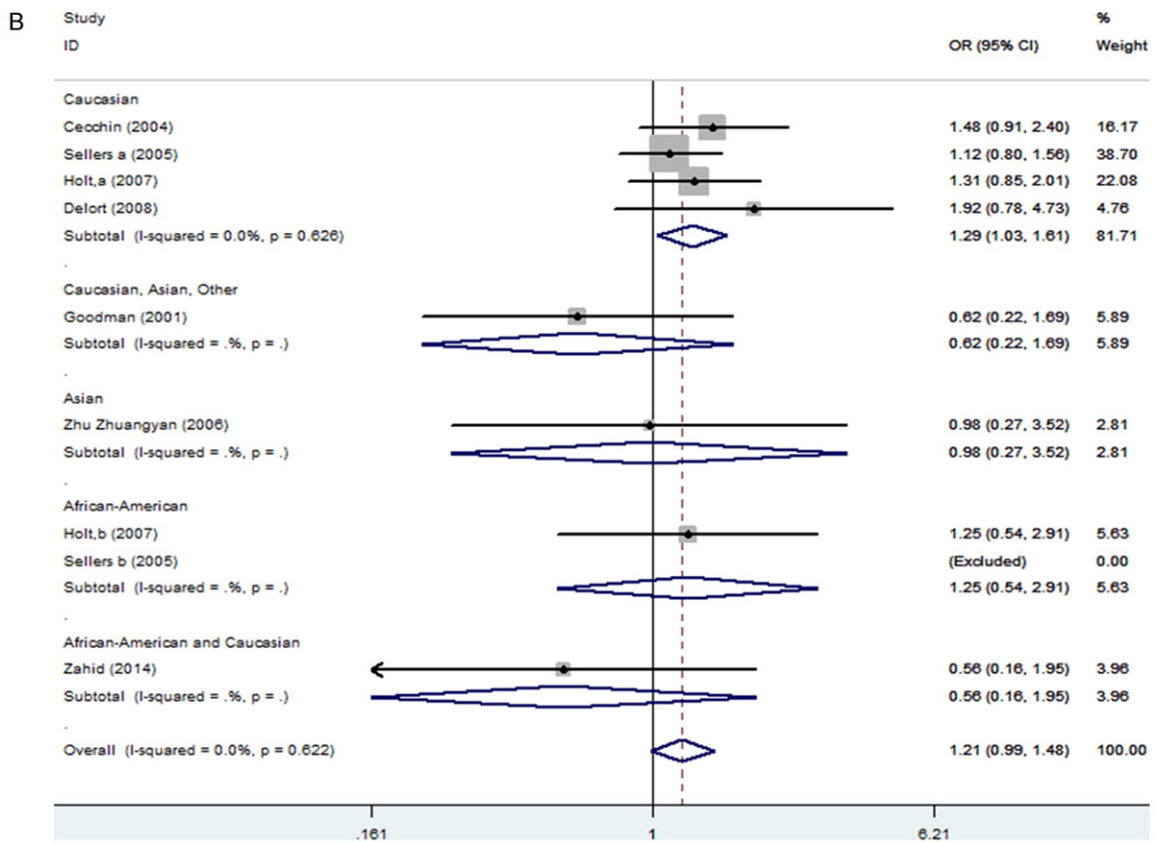
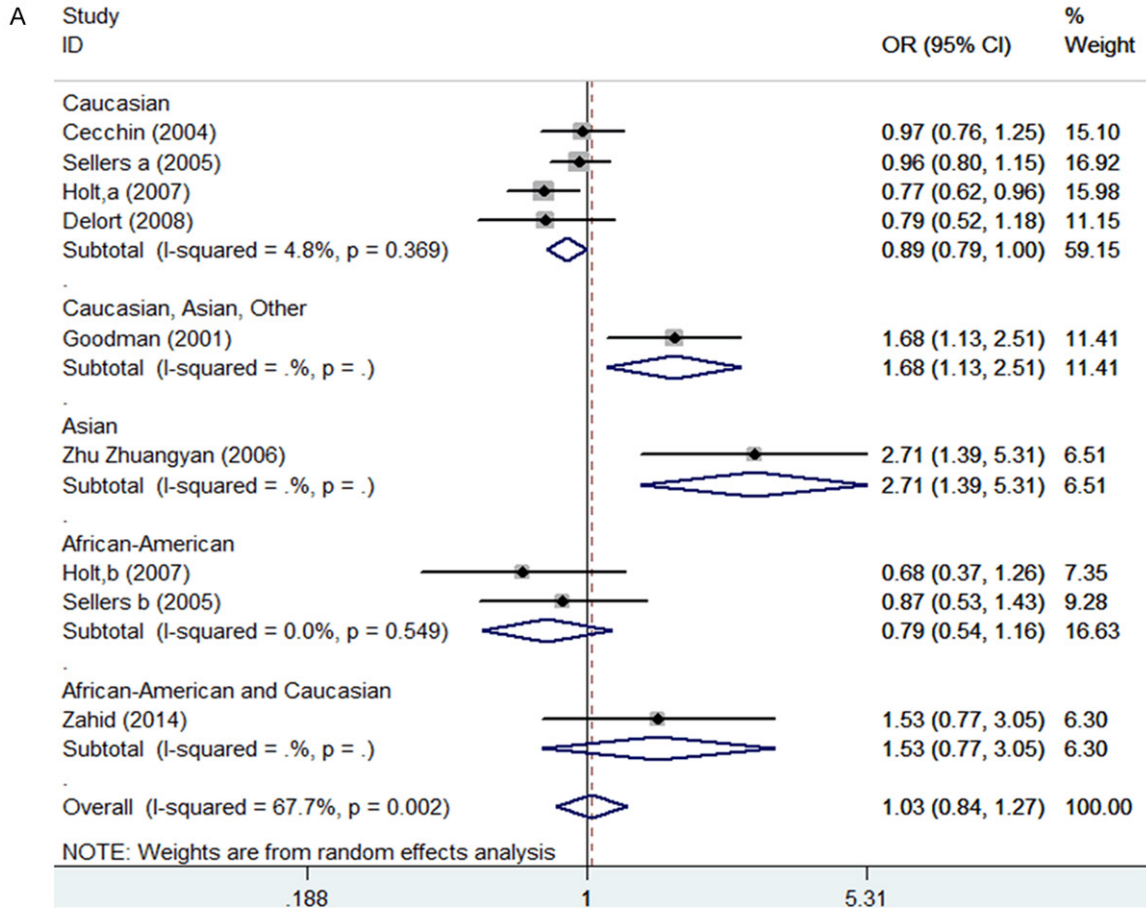
Through database retrieval and meta-analysis, rs1056836 was found strictly associated with the incidence of ovarian cancer. This mechanism may come from molecular evolution. But

Ovarian cancer risk with CYP1B1 come from the positive selection

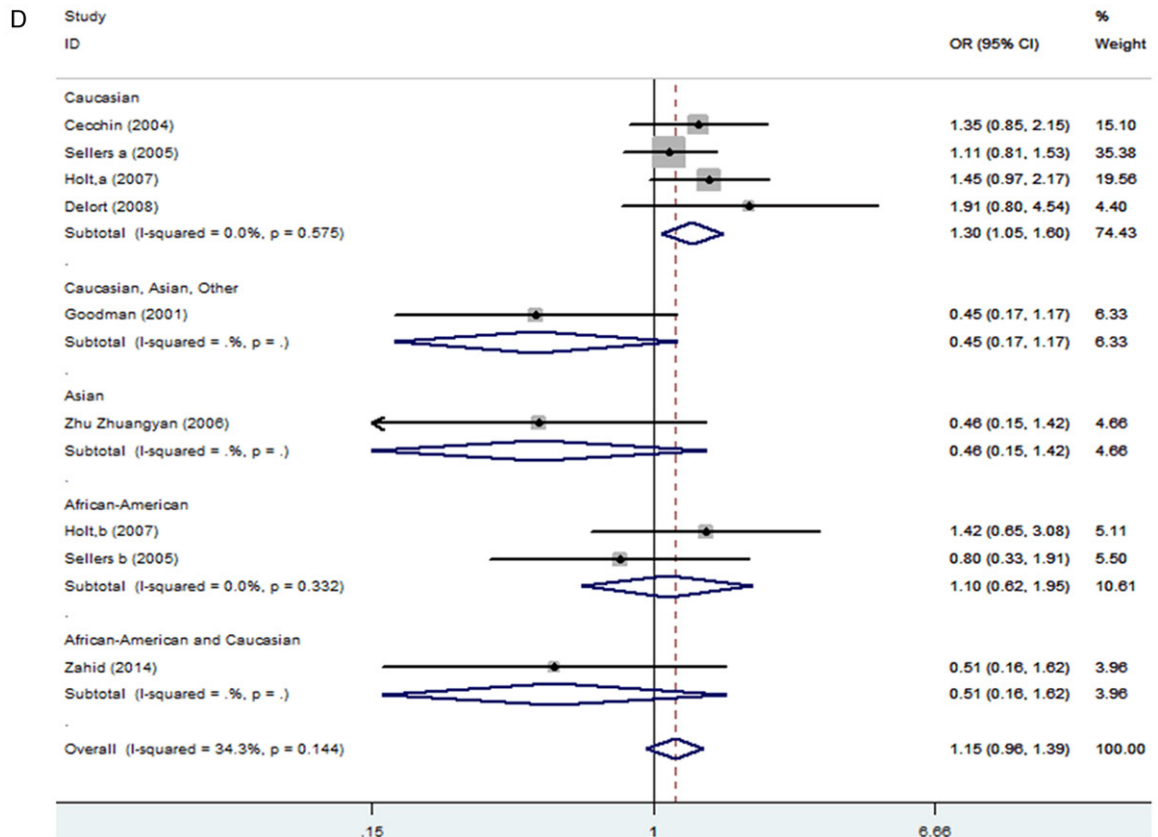
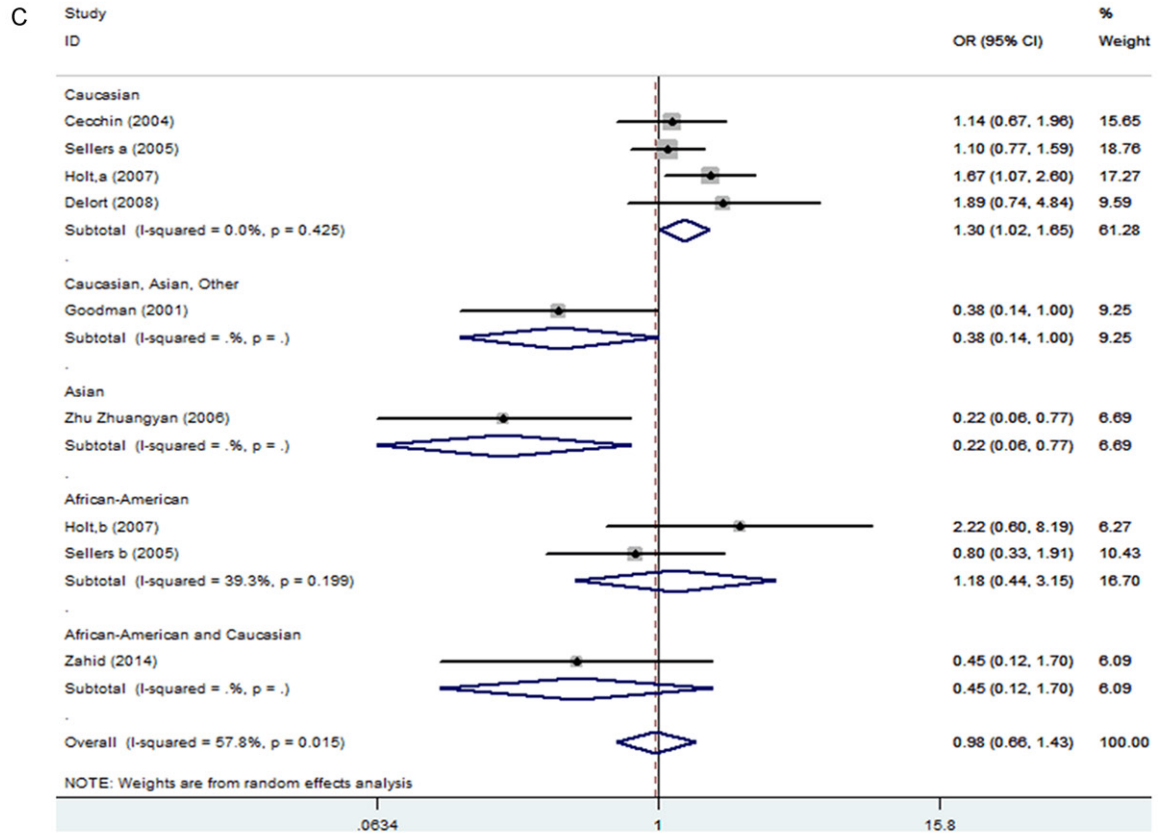
Table 8. Pooled odds ratios for heterozygous, homozygous carriers, dominant and recessive model for the polymorphisms CYP1B1 rs1056836 and rs1056827

| Polymorphisms | Allele model | | Heterozygous | | Homozygous | | Dominant model | | Recessive model | | F. additive model | |
|---------------|------------------|------------------------|------------------|------------------------|------------------|------------------------|------------------|------------------------|------------------|------------------------|-------------------|------------------------|
| | OR (95% CI) | Test for heterogeneity | OR (95% CI) | Test for heterogeneity | OR (95% CI) | Test for heterogeneity | OR (95% CI) | Test for heterogeneity | OR (95% CI) | Test for heterogeneity | OR (95% CI) | Test for heterogeneity |
| rs1056836 | 1.03 (0.84-1.27) | 0.002 | 1.21 (0.99-1.48) | 0.622 | 0.98 (0.66-1.43) | 0.015 | 1.15 (0.96-1.39) | 0.144 | 0.87 (0.65-1.17) | 0.007 | 0.88 (0.76-1.02) | 0.462 |
| rs1056827 | 0.94 (0.82-1.08) | 0.935 | 1.09 (0.90-1.32) | 0.707 | 1.06 (0.79-1.45) | 0.570 | 1.09 (0.91-1.31) | 0.776 | 1.04 (0.77-1.40) | 0.762 | 0.93 (0.77-1.11) | 0.624 |

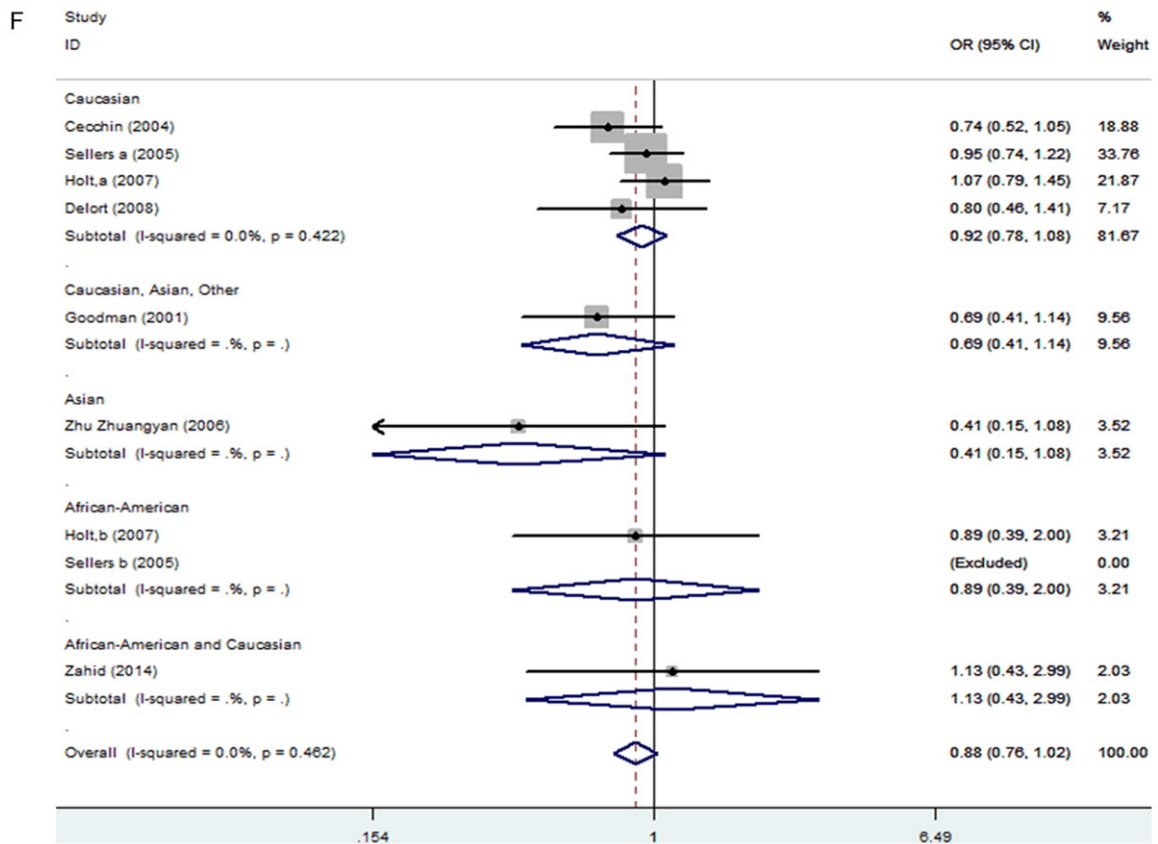
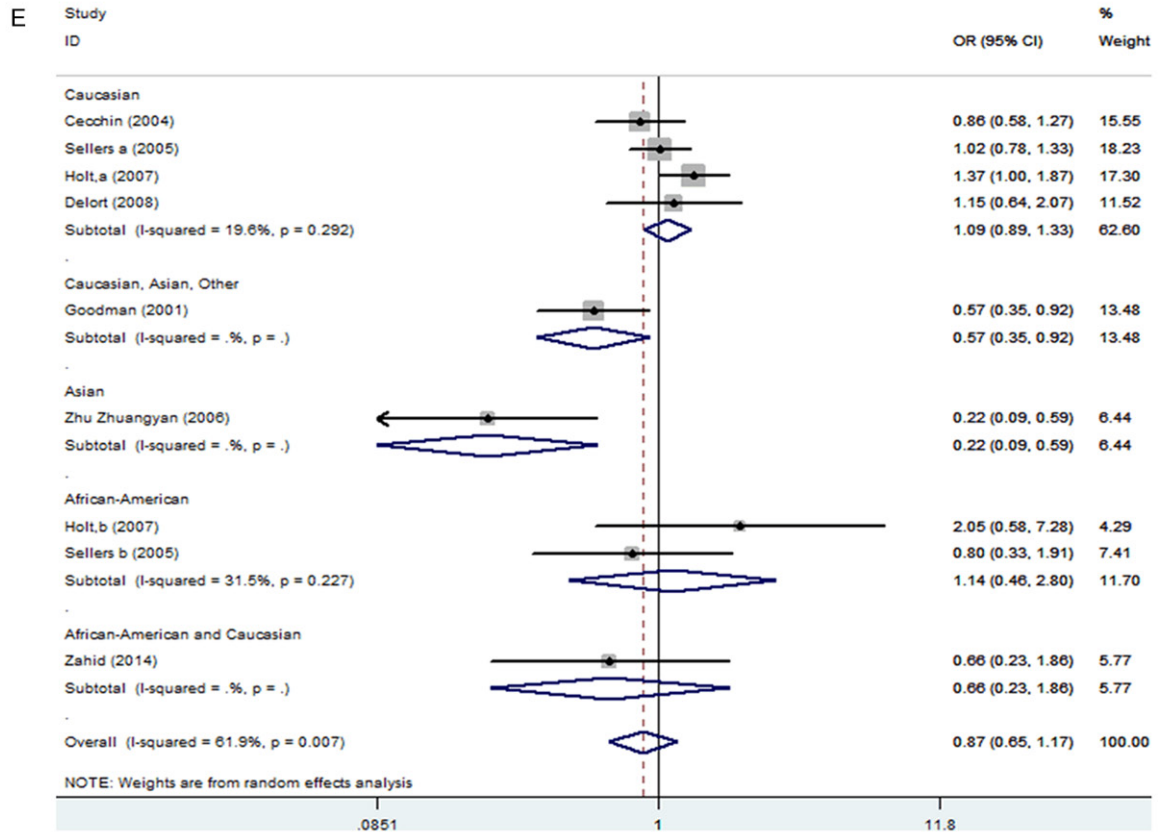
Ovarian cancer risk with CYP1B1 come from the positive selection



Ovarian cancer risk with CYP1B1 come from the positive selection

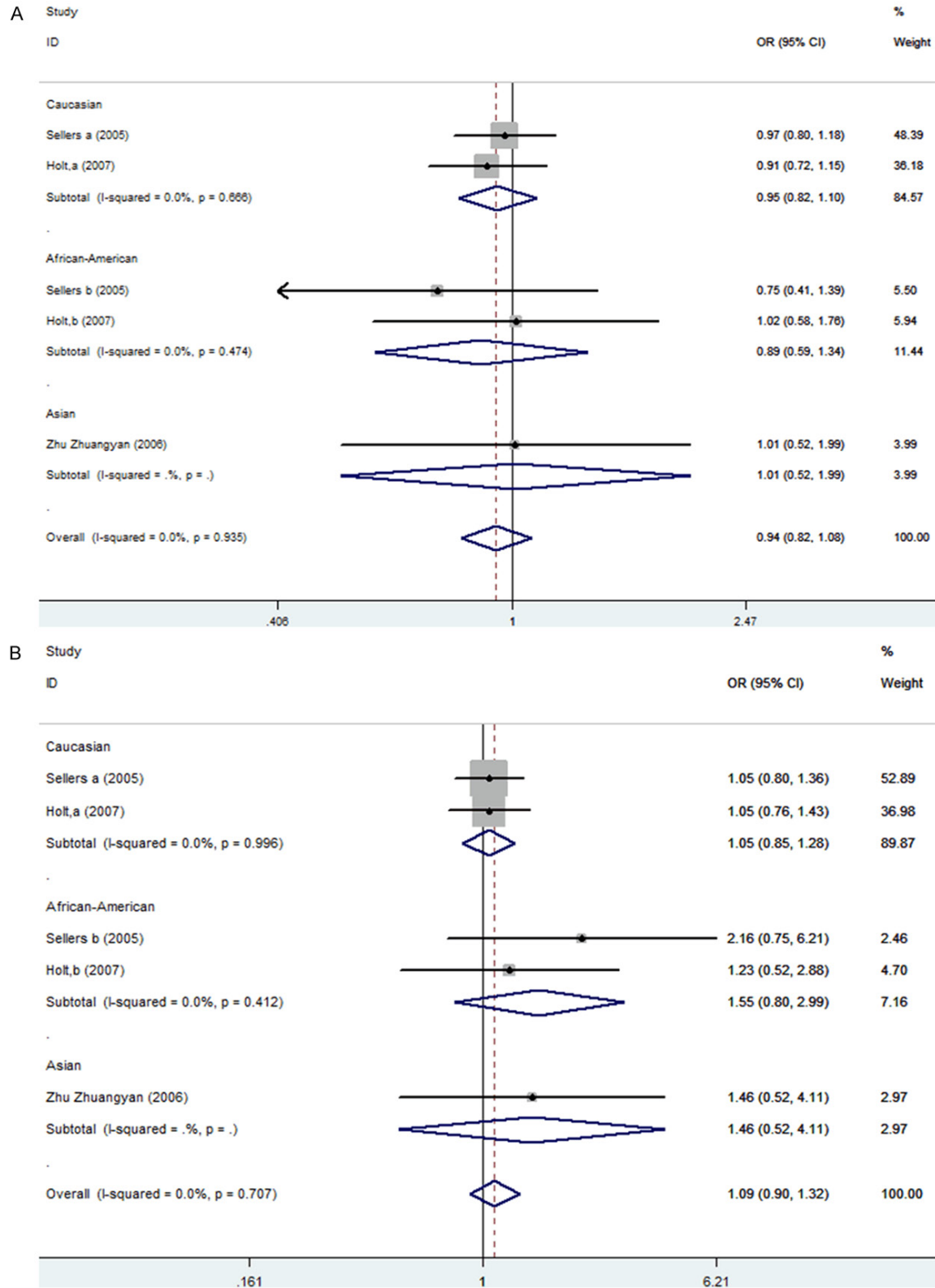


Ovarian cancer risk with CYP1B1 come from the positive selection

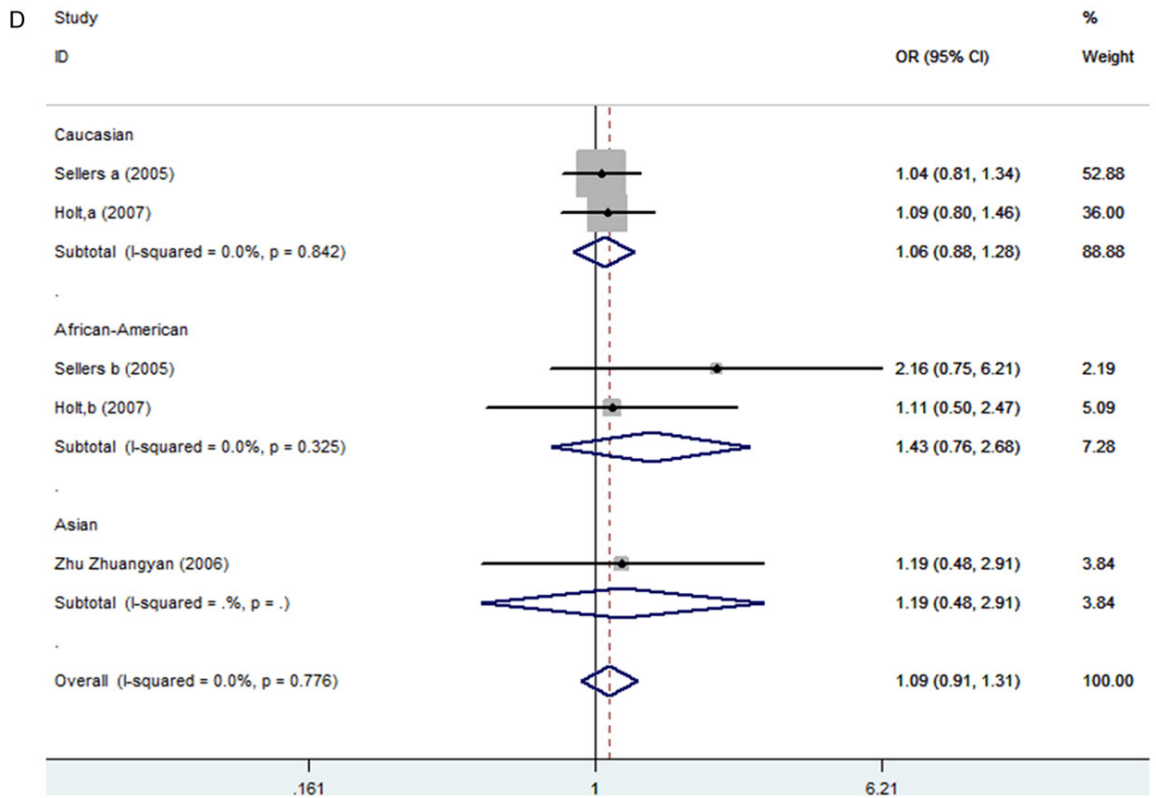
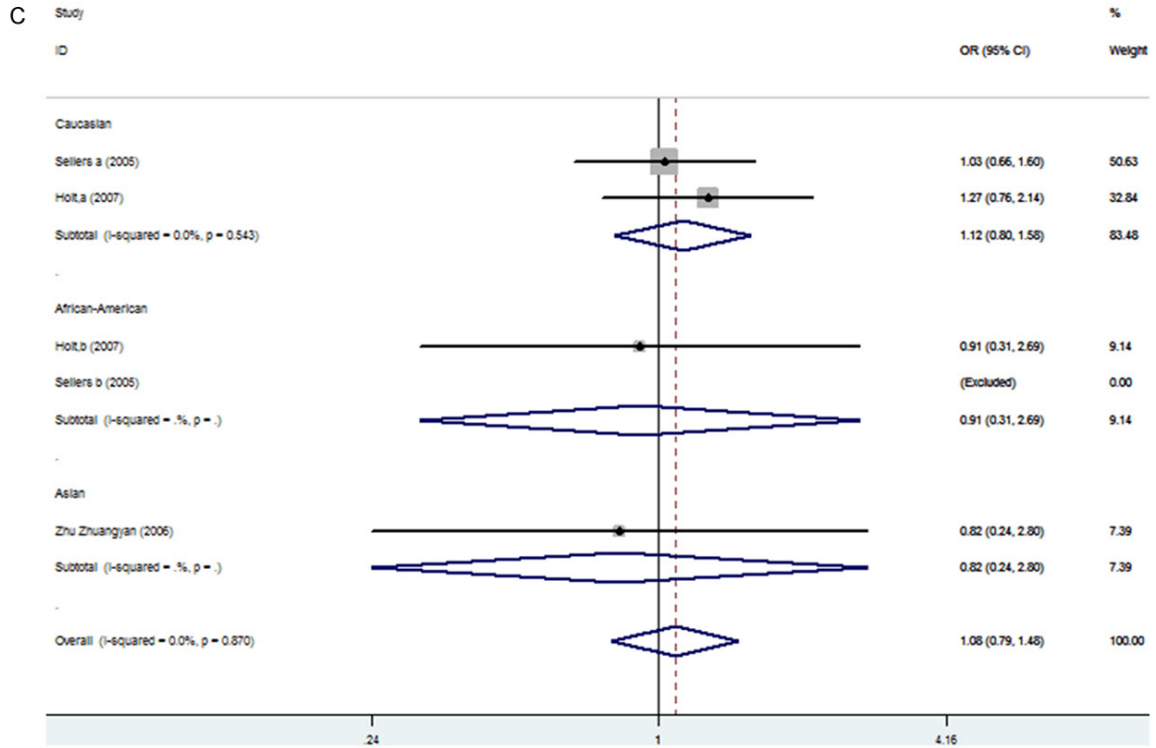


Ovarian cancer risk with CYP1B1 come from the positive selection

Figure 4. Forest plot for the inclusive correlation between V432L polymorphism and the risk for ovarian cancer. Shows the Allele model (A) Heterozygous model (B) Homozygote model (C) Dominant model (D) Recessive model (E) and additive model (F) depicts the recessive model. The fixed-effects model is implemented for a 95% confidence interval and to derive the pooled OR. The size of the square has been proportional to the weight of all studies and each study has been indicated via point estimate of the OR with 95% CI (extended lines).



Ovarian cancer risk with CYP1B1 come from the positive selection



Ovarian cancer risk with CYP1B1 come from the positive selection

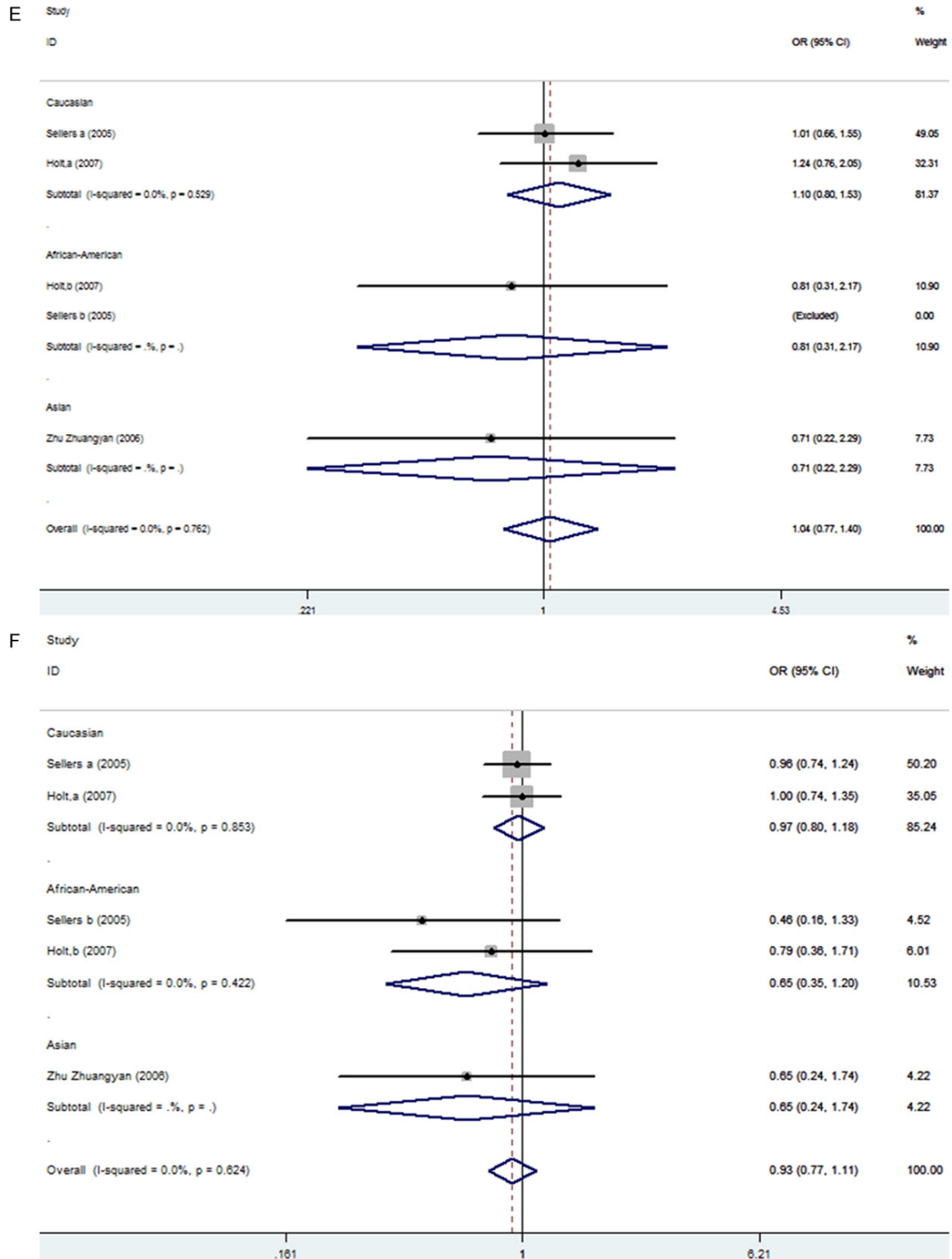


Figure 5. Forest plot for the inclusive relationship of A119S polymorphism and risk for ovarian carcinoma that reveals the Allele model (A) Heterozygous model (B) Homozygote model (C) Dominant model (D) Recessive model (E) and additive model (F) shows the recessive model. The fixed-effects model is implemented for a 95% confidence interval and to derive the pooled OR. The size of the square has been proportional to the weight of all studies and each study has been revealed via point estimate of the OR with 95% CI (extended lines).

Ovarian cancer risk with CYP1B1 come from the positive selection

Table 9. Association link of rs1056836 polymorphisms with drug-Resistant of ovarian cancer

| Genotype/allele CYP1B1 rs1056836 | Sensitive N = 137 | Resistant N = 62 | OR (95% CI) | P-value |
|----------------------------------|-------------------|------------------|---------------------|---------|
| Genotypic frequencies | | | | |
| CC | 3 | 0 | 1 | |
| GC | 16 | 8 | 0.842 (0.693-1.023) | 0.532 |
| GG | 118 | 54 | 0.975 (0.948-1.003) | 0.553 |
| Recessive model | | | | |
| Others | 19 | 8 | 1 | |
| GG | 118 | 54 | 0.920 (0.379-2.233) | 1.000 |
| Dominant model | | | | |
| CC | 3 | 0 | 1 | |
| Others | 134 | 62 | 1.463 (1.330-1.609) | 0.554 |
| Allelic frequencies | | | | |
| C | 22 | 8 | 1 | |
| G | 252 | 116 | 1.266 (0.547-2.928) | 0.684 |

genotyping assay results revealed that there was no considerable correlation between this locus and drug resistance of ovarian cancer. Overall results show that the SNP (rs1056836) of the CYP1B1 gene might be used as an indicator of ovarian cancer. Furthermore, structural and functional studies should be conducted for this SNP.

Acknowledgements

This study was supported by the Subject of Guangxi Scientific Research and Technology Development Program (No. 14124004); Guangxi Zhuang Autonomous Region Health and Family Planning Commission self-funded research project (No. Z2016288).

Disclosure of conflict of interest

None.

Address correspondence to: Li Li, Department of Gynecologic Oncology, Affiliated Tumor Hospital of Guangxi Medical University, 71 Hedi Road, Nanning 530021, China. E-mail: lili@gxmu.edu.cn

References

- [1] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* 2020; 70: 7-30.
- [2] Cornelison R, Llaneza DC and Landen CN. Emerging therapeutics to overcome chemoresistance in epithelial ovarian cancer: a mini-review. *Int J Mol Sci* 2017; 18: 2171.
- [3] Friedman R. Drug resistance in cancer: molecular evolution and compensatory proliferation. *Oncotarget* 2016; 11: 11746-55.

- [4] Turajlic S and Swanton C. Metastasis as an evolutionary process. *Science* 2016; 352: 169-175.
- [5] Langdon SP, Gourley C, Gabra H and Stanley B. Endocrine therapy in epithelial ovarian cancer. *Expert Rev Anticancer Ther* 2017; 17: 109-117.
- [6] Go RE, Hwang KA and Choi KC. Cytochrome P450 1 family and cancers. *J Steroid Biochem Mol Biol* 2015; 147: 24-30.
- [7] Sergeantanis TN, Economopoulos KP, Choussein S and Vlahos NF. Cytochrome P450 1A1 (CYP1A1) gene polymorphisms and ovarian cancer risk: a meta-analysis. *Mol Biol Rep* 2012; 39: 9921-9930.
- [8] Tian Z, Li YL, Zhao L and Zhang CL. Role of CYP1A2 1F polymorphism in cancer risk: evidence from a meta-analysis of 46 case-control studies. *Gene* 2013; 524: 168-174.
- [9] Liu JY, Yang Y, Liu ZZ, Xie JJ, Du YP and Wang W. Association between the CYP1B1 polymorphisms and risk of cancer: a meta-analysis. *Mol Genet Genomics* 2015; 290: 739-765.
- [10] Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W and Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010; 59: 307-321.
- [11] Guindon S, Delsuc F, Dufayard JF and Gascuel O. Estimating maximum likelihood phylogenies with PhyML. *Methods Mol Biol* 2009; 537: 113-137.
- [12] Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 2007; 24: 1586-1591.
- [13] Bielawski JP. Detecting the signatures of adaptive evolution in protein-coding genes. *Curr Protoc Mol Biol* 2013; Chapter 19: Unit 19.1.
- [14] Stadler T and Yang Z. Dating phylogenies with sequentially sampled tips. *Syst Biol* 2013; 62: 674-688.

Ovarian cancer risk with CYP1B1 come from the positive selection

- [15] Goodman MT, McDuffie K, Kolonel LN, Terada K, Donlon TA, Wilkens LR, Guo C and Le Marchand L. Case-control study of ovarian cancer and polymorphisms in genes involved in catecholesterogen formation and metabolism. *Cancer Epidemiol Biomarkers Prev* 2001; 10: 209-216.
- [16] Cecchin E, Russo A, Campagnutta E, Martella L and Toffoli G. Lack of association of CYP1B1*3 polymorphism and ovarian cancer in a Caucasian population. *Int J Biol Markers* 2004; 19: 160-163.
- [17] Sellers TA, Schildkraut JM, Pankratz VS, Vierkant RA, Fredericksen ZS, Olson JE, Cunningham J, Taylor W, Liebow M, McPherson C, Hartmann LC, Pal T and Adjei AA. Estrogen bioactivation, genetic polymorphisms, and ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 2005; 14: 2536-2543.
- [18] Holt SK, Rossing MA, Malone KE, Schwartz SM, Weiss NS and Chen C. Ovarian cancer risk and polymorphisms involved in estrogen catabolism. *Cancer Epidemiol Biomarkers Prev* 2007; 16: 481-489.
- [19] Delort L, Chalabi N, Satih S, Rabiau N, Kwiatkowski F, Bignon YJ and Bernard-Gallon DJ. Association between genetic polymorphisms and ovarian cancer risk. *Anticancer Res* 2008; 28: 3079-3081.
- [20] Zahid M, Beseler CL, Hall JB, LeVan T, Cavalieri EL and Rogan EG. Unbalanced estrogen metabolism in ovarian cancer. *Int J Cancer* 2014; 134: 2414-2423.
- [21] Zhu ZY, Mi RR and Liu J. Polymorphism of CYP1B1 gene and its susceptibility to ovarian cancer. *Progress in Modern Obstetrics and Gynecology* 2006; 184-187.
- [22] Liu Y. Natural selection and pangensis: the darwinian synthesis of evolution and genetics. *Adv Genet* 2018; 102: 121-142.
- [23] Shakya M, Ahmed SA, Davenport KW, Flynn MC, Lo CC and Chain PSG. Standardized phylogenetic and molecular evolutionary analysis applied to species across the microbial tree of life. *Sci Rep* 2020; 10: 1723.
- [24] Mohapatra DP, Singh SK, Sahoo M, Patole S, Mishra M, Debata NK and Mohapatra H. Retrospective study on clonal relationship of multi-drug-resistant klebsiella spp. indicates closed circulation and initiation of clonal divergence. *J Med Microbiol* 2018; 67: 611-619.
- [25] Benvenuto D, Giovanetti M, Salemi M, Prosperi M, De Flora C, Junior Alcantara LC, Angeletti S and Ciccozzi M. The global spread of 2019-nCoV: a molecular evolutionary analysis. *Pathog Glob Health* 2020; 114: 64-67.
- [26] Yohe LR, Liu L, Davalos LM and Liberles DA. Protocols for the molecular evolutionary analysis of membrane protein gene duplicates. *Methods Mol Biol* 2019; 1851: 49-62.
- [27] De Grazia S, Lanave G, Bonura F, Urone N, Cappa V, Li Muli S, Pepe A, Gellert A, Banyai K, Martella V and Giammanco GM. Molecular evolutionary analysis of type-1 human astroviruses identifies putative sites under selection pressure on the capsid protein. *Infect Genet Evol* 2018; 58: 199-208.
- [28] Martincorena I and Campbell PJ. Somatic mutation in cancer and normal cells. *Science* 2015; 349: 1483-1489.
- [29] Khan MM, Ryden AM, Chowdhury MS, Hasan MA and Kazi JU. Maximum likelihood analysis of mammalian p53 indicates the presence of positively selected sites and higher tumorigenic mutations in purifying sites. *Gene* 2011; 483: 29-35.
- [30] Lou DI, McBee RM, Le UQ, Stone AC, Wilkerson GK, Demogines AM and Sawyer SL. Rapid evolution of BRCA1 and BRCA2 in humans and other primates. *BMC Evol Biol* 2014; 14: 155.
- [31] Abdul Aziz AA, Md Salleh MS, Mohamad I, Krishna Bhavaraju VM, Mazuwin Yahya M, Zakaria AD, Hua Gan S and Ankathil R. Single-nucleotide polymorphisms and mRNA expression of CYP1B1 influence treatment response in triple negative breast cancer patients undergoing chemotherapy. *J Genet* 2018; 97: 1185-1194.
- [32] Alsubait A, Aldossary W, Rashid M, Algamdi A and Alrfaei BM. CYP1B1 gene: implications in glaucoma and cancer. *J Cancer* 2020; 11: 4652-4661.
- [33] Sissung TM, Price DK, Sparreboom A and Figg WD. Pharmacogenetics and regulation of human cytochrome P450 1B1: implications in hormone-mediated tumor metabolism and a novel target for therapeutic intervention. *Mol Cancer Res* 2006; 4: 135-150.
- [34] Shimada T, Watanabe J, Kawajiri K, Sutter TR, Guengerich FP, Gillam EM and Inoue K. Catalytic properties of polymorphic human cytochrome P450 1B1 variants. *Carcinogenesis* 1999; 20: 1607-1613.

Ovarian cancer risk with CYP1B1 come from the positive selection

Supplementary Table 1. CYP1A1 gene accession number from species, which are included in the selection and phylogenetic analyses

| order | Common name | Species name | DNA sequence | protein sequence | Abbreviation |
|-------|--------------------------|----------------------------|-----------------------|----------------------|------------------|
| 1 | human | Homo sapiens | ENST00000379727.7 | ENSP00000369050 | human |
| 2 | Chimpanzee | Pan troglodytes | ENSPTRT00000066960.2 | ENSPTRP00000058539 | Chimpanzee |
| 3 | Gorilla | Gorilla | ENSGGOT00000016584.2 | ENSGGOP00000016126 | Gorilla |
| 4 | Orangutan | Pongo abelii | ENSPPYT00000007843.1 | ENSPPYP00000007536 | Orangutan |
| 5 | Gibbon | Nomascus leucogenys | ENSNLET00000018631.2 | ENSNLEP00000017745 | Gibbon |
| 6 | Macaque | Macaca mulatta | ENSMMUT00000031615.3 | ENSMMUP00000029575 | Macaque |
| 7 | Olive baboon | Papio anubis | ENSPANT00000007574.1 | ENSPANP00000014273 | Olive_baboon |
| 8 | Vervet-AGM | Chlorocebus sabaeus | ENSCSAT00000012341.1 | ENSCSAP00000010381 | Vervet_AGM |
| 9 | Marmoset | Callithrix jacchus | ENSCJAT00000034574.2 | ENSCJAP00000032713 | Marmoset |
| 10 | Bushbaby | Otolemur garnettii | ENSOGAT00000014451.2 | ENSOGAP00000012947 | Bushbaby |
| 11 | Mouse Lemur | Microcebus murinus | ENSMICT00000048712.1 | ENSMICP00000033975 | Mouse_Lemur |
| 12 | Elephant | Loxodonta africana | ENSLAFT00000015554.2 | ENSLAFP00000018088 | Elephant |
| 13 | Microbat | Myotis lucifugus | ENSMLUT00000008383.2 | ENSMLUP00000007650 | Microbat |
| 14 | Armadillo | Dasypus novemcinctus | ENSNDOT00000052025.1 | ENSNDOP00000026692 | Armadillo |
| 15 | Cat | Felis catus | ENSFCAT00000002015.2 | ENSFCAP00000001863 | Cat |
| 16 | Dog | Canis lupus familiaris | ENSCAFT00000028474.3 | ENSCAFP00000026483 | Dog |
| 17 | Dolphin | Tursiops truncatus | ENSTTRT00000007994.1 | ENSTTRP00000007564 | Dolphin |
| 18 | Horse | Equus caballus | ENSECAT00000015233.1 | ENSECAP00000012229 | Horse |
| 19 | Pig | Sus scrofa | ENSSSCT00000002135.2 | ENSSSCP00000002085 | Pig |
| 20 | Cow | Bos taurus | ENSBTAT00000061300.2 | ENSBTAP00000005321.2 | Cow |
| 21 | Mouse | Mus musculus | ENSMUST00000034865.5 | ENSMUSP0000003486225 | Mouse |
| 22 | Squirrel | Ictidomys tridecemlineatus | ENSSTOT00000013941.2 | ENSSTOP00000012492 | Squirrel |
| 23 | Megabat | Pteropus vampyrus | ENSPVAT00000015845.1 | ENSPVAP00000014951 | Megabat |
| 24 | Rat | Rattus norvegicus | ENSRNOT00000026473.4 | ENSRNOP00000026473 | Rat |
| 25 | Guinea Pig | Cavia porcellus | ENSCPOT00000015748.2 | ENSCPOP00000014064 | Guinea_Pig |
| 26 | hedgehog | Erinaceus europaeus | ENSETET00000010192.1 | ENSETEP00000008275 | hedgehog |
| 27 | Rabbit | Oryctolagus cuniculus | ENSOCUT00000017743.3 | ENSOCUP00000015239 | Rabbit |
| 28 | Ferret | Mustela putorius furo | ENSMPUT00000017441.1 | ENSMPUP00000017185 | Ferret |
| 29 | Opossum | Monodelphis domestica | ENSMODT00000012011.3 | ENSMODP00000011788 | Opossum |
| 30 | Wallaby | Notamacropus eugenii | ENSMEUT0000000131.1 | ENSMEUP00000000119 | Wallaby |
| 31 | Tasmanian devil | Sarcophilus harrisii | ENSSHAT00000018234.1 | ENSSHAP00000018085 | Tasmanian_devil |
| 32 | Platypus | Ornithorhynchus anatinus | ENSOANT00000005633.2 | ENSOANP00000005631 | Platypus |
| 33 | Chinese softshell turtle | Pelodiscus sinensis | ENSPSIT00000008339.1 | ENSPSIP00000008296 | softshell_turtle |
| 34 | Flycatcher | Ficedula albicollis | ENSFALT00000004745.1 | ENSFALP00000004721 | Flycatcher |
| 35 | Duck | Anas platyrhynchos | ENSAPLT00000011431.1 | ENSAPLP00000010714 | Duck |
| 36 | Xenopus | Xenopus tropicalis | ENSXETT000000061279.1 | ENSXETP000000063291 | Xenopus |
| 37 | Zebra Finch | Taeniopygia guttata | ENSTGUT00000004161.1 | ENSTGUP00000004116 | Zebra_Finch |
| 38 | Chicken | Gallus gallus | ENSGALT00000002018.5 | ENSGALP00000002016 | Chicken |
| 39 | Anole lizard | Anolis carolinensis | ENSACAT00000014825.3 | ENSACAP00000014530 | Anole_lizard |
| 40 | Spotted gar | Lepisosteus oculatus | ENSLOCT00000018040.1 | ENSLOCP00000018008 | Spotted_gar |
| 41 | Tilapia | Oreochromis niloticus | ENSONIT00000003676.1 | ENSONIP00000003675 | Tilapia |
| 42 | Amazon molly | Poecilia formosa | ENSPFOT00000015376.1 | ENSPFOP00000015354 | Amazon_molly |
| 43 | Cave fish | Astyanax mexicanus | ENSAMXT00000021404.1 | ENSAMXP00000021404 | Cave_fish |
| 44 | Cod | Gadus morhua | ENSGMOT00000000331.1 | ENSGMOP00000000312 | Cod |
| 45 | Fugu | Takifugu rubripes | ENSTRUT00000002337.1 | ENSTRUP00000002327 | Fugu |
| 46 | Stickleback | Gasterosteus aculeatus | ENSGACT00000019429.1 | ENSGACP00000019391 | Stickleback |
| 47 | Tetraodon | Tetraodon nigroviridis | ENSTNIT00000015092.1 | ENSTNIP00000014891 | Tetraodon |
| 48 | Medaka | Oryzias latipes | ENSORLT00000018074.1 | ENSORLP00000018073 | Medaka |
| 49 | Platyfish | Xiphophorus maculatus | ENSXMAT00000016834.1 | ENSXMAP00000016810 | Platyfish |
| 50 | Zebrafish | Danio rerio | ENSDART000000161538.1 | ENSDARP000000139599 | Zebrafish |

Ovarian cancer risk with CYP1B1 come from the positive selection

Supplementary Table 2. CYP1A2 gene accession number from species that are included in the selection and phylogenetic analyses

| Mammalian order | Common name | Species name | DNA sequence | protein sequence | Abbreviation |
|-----------------|--------------------------|----------------------------|----------------------|--------------------|------------------|
| 1 | human | Homo sapiens | ENST00000343932.4 | ENSP00000342007 | human |
| 2 | Gorilla | Gorilla gorilla gorilla | ENSGGOT00000010896.2 | ENSGGOP00000010581 | Gorilla |
| 3 | Gibbon | Nomascus leucogenys | ENSNLET00000018654.1 | ENSNLEP00000017767 | Gibbon |
| 4 | Olive baboon | Papio anubis | ENSPANT00000004824.1 | ENSPANP00000014274 | Olive_baboon |
| 5 | Macaque | Macaca mulatta | ENSMMUT00000030720.3 | ENSMMUP00000028750 | Macaque |
| 6 | Orangutan | Pongo abelii | ENSPPYT00000007844.2 | ENSPPYP00000007537 | Orangutan |
| 7 | Marmoset | Callithrix jacchus | ENSCJAT00000034599.1 | ENSCJAP00000032736 | Marmoset |
| 8 | Bushbaby | Otolemur garnettii | ENSOGAT00000014453.2 | ENSOGAP00000017637 | Bushbaby |
| 9 | Chimpanzee | Pan troglodytes | ENSPTRT00000013446.3 | ENSPTRP00000012463 | Chimpanzee |
| 10 | Horse | Equus caballus | ENSECAT00000009040.1 | ENSECAP00000006831 | Horse |
| 11 | Cow | Bos taurus | ENSBTAT00000000094.5 | ENSBTAP00000000094 | Cow |
| 12 | Squirrel | Ictidomys tridecemlineatus | ENSSTOT00000029961.1 | ENSSTOP00000021252 | Squirrel |
| 13 | Microbat | Myotis lucifugus | ENSMLUT00000014050.2 | ENSMLUP00000012780 | Microbat |
| 14 | Dolphin | Tursiops truncatus | ENSTTRT00000007995.1 | ENSTTRP00000007565 | Dolphin |
| 15 | Panda | Ailuropoda melanoleuca | ENSAMET00000009752.1 | ENSAMEP00000009352 | Panda |
| 16 | Dog | Canis lupus familiaris | ENSCAFT00000039456.2 | ENSCAFP00000035314 | Dog |
| 17 | Elephant | Loxodonta africana | ENSLAFT00000035152.1 | ENSLAFP00000027591 | Elephant |
| 18 | Sheep | Ovis aries | ENSOART00000003636.1 | ENSOARP00000003572 | Sheep |
| 19 | Pig | Sus scrofa | ENSSSCT00000002129.2 | ENSSSCP00000002079 | Pig |
| 20 | hedgehog tenrec | Erinaceus europaeus | ENSETET00000002401.1 | ENSETEP00000001954 | hedgehog_tenrec |
| 21 | Cat | Felis catus | ENSFCAT00000000343.2 | ENSFCAP00000000315 | Cat |
| 22 | Tree Shrew | Tupaia belangeri | ENSTBEP00000005505 | ENSTBEP00000005505 | Tree_Shrew |
| 23 | Rabbit | Oryctolagus cuniculus | ENSOCUT00000010056.3 | ENSOCUP00000008665 | Rabbit |
| 24 | Shrew | Sorex araneus | ENSSART00000000304.1 | ENSSARP00000000273 | Shrew |
| 25 | Guinea Pig | Cavia porcellus | ENSCPOT00000001250.2 | ENSCPOP00000017610 | Guinea_Pig |
| 26 | Pika | Ochotona princeps | ENSOPRT00000010487.1 | ENSOPRP00000009583 | Pika |
| 27 | Rat | Rattus 28 norvegicus | ENSRNOT00000021653.7 | ENSRNOP00000021653 | Rat |
| 28 | Mouse | Mus musculus | ENSMUST00000034860.4 | ENSMUSP00000034860 | Mouse |
| 29 | Wallaby | Notamacropus eugenii | ENSMEUT00000015717.1 | ENSMEUP00000014310 | Wallaby |
| 30 | Kangaroo rat | Dipodomys ordii | ENSDORT00000013132.1 | ENSDORP00000012346 | Kangaroo_rat |
| 31 | Opossum | Monodelphis domestica | ENSMODT00000012011.3 | ENSMODP00000011788 | Opossum |
| 32 | Tasmanian devil | Sarcophilus harrisii | ENSSHAT00000018234.1 | ENSSHAP00000018085 | Tasmanian_devil |
| 33 | Chinese softshell turtle | Pelodiscus sinensis | ENSPSIT00000008339.1 | ENSPSIP00000008296 | softshell_turtle |

Ovarian cancer risk with CYP1B1 come from the positive selection

| | | | | | |
|----|--------------|------------------------|------------------------|----------------------|--------------|
| 34 | Chicken | Gallus gallus | ENSGALT00000002018.5 | ENSGALP00000002016 | Chicken |
| 35 | Duck | Anas platyrhynchos | ENSAPLT00000011431.1 | ENSAPLP00000010714 | Duck |
| 36 | Zebra Finch | Taeniopygia guttata | ENSTGUT00000004161.1 | ENSTGUP00000004116 | Zebra_Finch |
| 37 | Anole lizard | Taeniopygia guttata | ENSACAT00000014825.3 | ENSACAP00000014530 | Anole_lizard |
| 38 | Flycatcher | Ficedula albicollis | ENSFALT00000004745.1 | ENSFALP00000004721 | Flycatcher |
| 39 | Xenopus | Xenopus tropicalis | ENSXETT00000061279.1 | ENSXETP00000063291 | Xenopus |
| 40 | Spotted gar | Lepisosteus oculatus | ENSLOCT00000018040.1 | ENSLOCP00000018008 | Spotted_gar |
| 41 | Chicken | Gallus gallus | ENSGALT00000040012.2 | ENSGALP00000039219 | Chicken |
| 42 | Stickleback | Gasterosteus aculeatus | ENSGACT00000019429.1 | ENSGACP00000019391 | Stickleback |
| 43 | Tetraodon | Tetraodon nigroviridis | ENSTNIT00000015092.1 | ENSTNIP00000014891 | Tetraodon |
| 44 | Tilapia | Oreochromis niloticus | ENSONIT00000003676.1 | ENSONIP00000003675 | Tilapia |
| 45 | Zebrafish | Danio rerio | ENSDDART000000161538.1 | ENSDDARP000000139599 | Zebrafish |
| 46 | Cod | Gadus morhua | ENSGMOT00000000331.1 | ENSGMOP00000000312 | Cod |
| 47 | Cave fish | Astyanax mexicanus | ENSAMXT00000021404.1 | ENSAMXP00000021404 | Cave_fish |
| 48 | Medaka | Oryzias latipes | ENSORLT00000018074.1 | ENSORLP00000018073 | Medaka |
| 49 | Fugu | Takifugu rubripes | ENSTRUT00000002337.1 | ENSTRUP00000002327 | Fugu |
| 50 | Amazon molly | Poecilia formosa | ENSPFOT00000015376.1 | ENSPFOP00000015354 | Amazon_molly |
| 51 | Platyfish | Xiphophorus maculatus | ENSXMAT00000016834.1 | ENSXMAP00000016810 | Platyfish |

Supplementary Table 3. CYP1b1 gene accession number from species that are included in the selection and phylogenetic analyses

| Mammalian order | Common name | Species name | DNA sequence | protein sequence | Abbreviation |
|-----------------|--------------|------------------------|----------------------|--------------------|--------------|
| 1 | human | Homo sapiens | ENST00000610745.4 | ENSP00000478561 | human |
| 2 | Chimpanzee | Pan troglodytes | ENSPTRT00000022042.3 | ENSPTRP00000020335 | Chimpanzee |
| 3 | Orangutan | Pongo abelii | ENSPPYT00000014510.2 | ENSPPYP00000013944 | Orangutan |
| 4 | Gibbon | Nomascus leucogenys | ENSNLET00000020780.2 | ENSNLEP00000019784 | Gibbon |
| 5 | Vervet-AGM | Chlorocebus sabaues | ENSCSAT00000010318.1 | ENSCSAP00000008414 | Vervet |
| 6 | Olive baboon | Papio anubis | ENSPAN00000007322.1 | ENSPANP00000011533 | Olive_baboon |
| 7 | Marmoset | Callithrix jacchus | ENSCJAT00000007727.2 | ENSCJAP00000007313 | Marmoset |
| 8 | Tree Shrew | Tupaia belangeri | ENSTBET00000001843.1 | ENSTBEP00000001596 | Tree_Shrew |
| 9 | Dolphin | Tursiops truncatus | ENSTTRT00000007063.1 | ENSTTRP00000006681 | Dolphin |
| 10 | Horse | Equus caballus | ENSECAT00000020731.1 | ENSECAP00000017028 | Horse |
| 11 | Mouse Lemur | Microcebus murinus | ENSMICT00000042808.1 | ENSMICP00000020375 | Mouse_Lemur |
| 12 | Ferret | Mustela putorius furo | ENSMPUT00000010354.1 | ENSMPUP00000010192 | Ferret |
| 13 | Rabbit | Oryctolagus cuniculus | ENSOCUT00000016155.3 | ENSOCUP00000013885 | Rabbit |
| 14 | Panda | Ailuropoda melanoleuca | ENSAMET00000019372.1 | ENSAMEP00000018626 | Panda |

Ovarian cancer risk with CYP1B1 come from the positive selection

| | | | | | |
|----|--------------------------|----------------------------|----------------------|--------------------|-----------------|
| 15 | Dog | Canis lupus familiaris | ENSCAFT00000047464.2 | ENSCAFP00000042136 | Dog |
| 16 | Bushbaby | Otolemur garnettii | ENSOGAT00000014191.2 | ENSOGAP00000012713 | Bushbaby |
| 17 | Squirrel | Ictidomys tridecemlineatus | ENSSTOT00000001645.2 | ENSSTOP00000001475 | Squirrel |
| 18 | Cow | Bos taurus | ENSBTAT00000013922.2 | ENSBTAP00000013922 | Cow |
| 19 | Armadillo | Dasypus novemcinctus | ENSDNOT00000047745.1 | ENSDNOP00000029542 | Armadillo |
| 20 | Mouse | Mus musculus | ENSMUST00000024894.1 | ENSMUSP00000024894 | Mouse |
| 21 | Sheep | Ovis aries | ENSOART00000009894.1 | ENSOARP00000009752 | Sheep |
| 22 | Guinea Pig | Cavia porcellus | ENSCPOT00000014371.2 | ENSCPOP00000012819 | Guinea_Pig |
| 23 | Elephant | Loxodonta africana | ENSLAFT00000031675.1 | ENSLAFP00000028088 | Elephant |
| 24 | Hyrax | Procavia capensis | ENSPCAT00000011200.1 | ENSPCAP00000010459 | Hyrax |
| 25 | Kangaroo rat | Dipodomys ordii | ENSDORT00000007803.1 | ENSDORP00000007316 | Kangaroo_rat |
| 26 | Microbat | Myotis lucifugus | ENSMLUT00000013788.2 | ENSMLUP00000012545 | Microbat |
| 27 | Rat | Rattus norvegicus | ENSRNOT00000082017.1 | ENSRNOP00000071724 | Rat |
| 28 | Opossum | Monodelphis domestica | ENSMODT00000010756.1 | ENSMODP00000010550 | Opossum |
| 29 | Wallaby | Notamacropus eugenii | ENSMEUT00000007415.1 | ENSMEUP00000006754 | Wallaby |
| 30 | Tasmanian devil | Sarcophilus harrisii | ENSSHAT00000020730.1 | ENSSHAP00000020566 | Tasmanian_devil |
| 31 | Tarsier | Carlito syrichta | ENSTSYT00000008735.1 | ENSTSYT00000008021 | Tarsier |
| 32 | Chinese softshell turtle | Pelodiscus sinensis | ENSPSIT00000005914.1 | ENSPSIP00000005879 | turtle |
| 33 | Chicken | Gallus gallus | ENSGALT00000047969.1 | ENSGALP00000048510 | Chicken |
| 34 | Zebra Finch | Taeniopygia guttata | ENSTGUT00000009158.1 | ENSTGUP00000009061 | Zebra_Finch |
| 35 | Flycatcher | Ficedula albicollis | ENSFALT00000012490.1 | ENSFALP00000012440 | Flycatcher |
| 36 | Coelacanth | Latimeria chalumnae | ENSLACT00000018774.2 | ENSLACP00000018641 | Coelacanth |
| 37 | Spotted gar | Lepisosteus oculatus | ENSLOCT00000020495.1 | ENSLOCP00000020460 | Spotted_gar |
| 38 | Xenopus | Xenopus tropicalis | ENSXETT00000054079.2 | ENSXETP00000054079 | Xenopus |
| 39 | Zebrafish | Danio rerio | ENSDART00000131147.2 | ENSDARP00000107132 | Zebrafish |
| 40 | Stickleback | Gasterosteus aculeatus | ENSGACT00000003949.1 | ENSGACP00000003935 | Stickleback |
| 41 | Cave fish | Astyanax mexicanus | ENSAMXT00000021545.1 | ENSAMXP00000021545 | Cave_fish |
| 42 | Amazon molly | Poecilia formosa | ENSPFOT00000010612.1 | ENSPFOP00000010597 | Amazon_molly |
| 43 | Fugu | Takifugu rubripes | ENSTRUT00000017470.1 | ENSTRUP00000017396 | Fugu |
| 44 | Platyfish | Xiphophorus maculatus | ENSXMAT00000007749.1 | ENSXMAP00000007741 | Platyfish |
| 45 | Medaka | Oryzias latipes | ENSORLT00000001045.1 | ENSORLP00000001044 | Medaka |
| 46 | Tilapia | Oreochromis niloticus | ENSONIT00000013064.1 | ENSONIP00000013054 | Tilapia |