



OPEN ACCESS

EDITED BY

Mounir Tilaoui,
Waterford Institute of Technology, Ireland

REVIEWED BY

Raffaella Ergasti,
Agostino Gemelli University Polyclinic
(IRCCS), Italy
Abdelhedi Rania,
Faculty of Science of Sfax, Tunisia

*CORRESPONDENCE

Shan Kang

✉ kangshan@hebmu.edu.cn

RECEIVED 09 July 2024

ACCEPTED 13 December 2024

PUBLISHED 07 January 2025

CITATION

Du N-Y, Li Y, Zheng H, Liu Y-K, Liu L-S,
Xie J and Kang S (2025) Incorporating
genotype information in a precise
prediction model for platinum sensitivity in
epithelial ovarian cancer.
Front. Oncol. 14:1461772.
doi: 10.3389/fonc.2024.1461772

COPYRIGHT

© 2025 Du, Li, Zheng, Liu, Liu, Xie and Kang.
This is an open-access article distributed under
the terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Incorporating genotype information in a precise prediction model for platinum sensitivity in epithelial ovarian cancer

Nai-Yi Du¹, Yan Li², Hui Zheng³, Ya-Kun Liu¹, Lu-Sha Liu¹,
Jianbang Xie³ and Shan Kang^{1*}

¹Department of Gynecology, the Fourth Hospital of Hebei Medical University, Shijiazhuang, Hebei, China, ²Department of Molecular Biology, the Fourth Hospital of Hebei Medical University, Shijiazhuang, Hebei, China, ³Department of Translation Medicine, Shijiazhuang Ninghong Biotechnology Co., Ltd., Shijiazhuang, Hebei, China

Objective: Develop a predicting model that can help stratify patients with epithelial ovarian cancer (EOC) before platinum-based chemotherapy.

Methods: 148 patients with pathologically confirmed EOC and with a minimum 5-year follow-up were retrospectively enrolled. Patients were classified into platinum-sensitive and platinum-resistant groups according to treatment responses. The correlation between clinical factors and drug sensitivity was evaluated using statistical tests. Approximately 700,000 single-nucleotide polymorphism (SNP) sites were assessed for association with drug sensitivity via the Genome-wide Association Study (GWAS). LASSO regression and manual selection were employed to reduce the number of variables. A predicting model based on optimized variables was constructed. The predictive ability of the model was assessed using the Kaplan-Meier curve.

Results: No statistically significant association was found between clinical factors and drug sensitivity. Sixteen SNPs were preserved after the optimization. A predicting model for drug sensitivity was constructed based on those sixteen SNPs. Coefficients of the synergistic effect for each SNP were determined, and an algorithm of the Drug Sensitivity Index (DSI) was built. The DSI score can successfully distinguish the drug-sensitive or drug-resistant patients with sensitivity, specificity, positive predictive value, and accuracy of 94.7%, 83.3%, 90.8%, and 90.5%, respectively. In both the training set and validating samples, the Kaplan-Meier curve showed that the median PFS and mean OS were significantly differentiated between the predicted sensitive and resistant patients (p -value<0.001).

Conclusions: A mathematical model incorporating genotype information could help predict the drug sensitivity of platinum-based chemotherapy before the treatment in EOC patients. A personal chemotherapy could be achieved based on the model.

KEYWORDS

epithelial ovarian cancer, drug sensitivity, predicting model, single-nucleotide polymorphism, clinical factors

1 Introduction

Ovarian cancer (OC) is the leading cause of death from gynecologic cancers in the world, and ninety percent of OCs are epithelial ovarian cancer (EOC) (1). Due to the lack of early symptoms and insufficient sensitive screening methods, most EOC patients are diagnosed at clinically advanced stages. Following the primary cytoreductive surgery, platinum-based chemotherapy is the first-line medication for advanced-stage EOC patients (2, 3). Although the majority of patients respond well to the treatment, approximately 15% of patients will show drug resistance at the initial stage of chemotherapy, and an additional 30% will relapse within six months after the treatment finished (4). Currently, almost all patients will receive platinum-based chemotherapy for the front line of treatment. Unfortunately, those platinum-resistant patients are deprived of their opportunity for early optimal intervention, such as non-platinum-based chemotherapy, targeted therapies, or immunotherapy. Therefore, patient stratification prior to the chemotherapy has a significant clinical impact.

Many clinical and genetic factors were reported to be influential on the sensitivity of EOC patients to platinum-based drugs. A study conducted by Winarno GNA et al. found that the mean age of the platinum-sensitive group was significantly lower than that of the platinum-resistant group (5). In addition, they observed that the stage of EOC may also affect the response to the chemotherapy. Ikeda et al. found that the residual tumor was an independent factor for the response to platinum (6). Regarding genetic factors, single nucleotide polymorphisms (SNPs) may be the main genetic factors responsible for the variations of drug response among individuals. Many researchers, including our group, have identified specific SNPs associated with drug sensitivity in EOC (7–12). Furthermore, other studies suggested that the serum level of lactate dehydrogenase and ratios of platelet/lymphocyte and neutrophil/lymphocyte can be used to predict the drug sensitivity of EOC patients (5, 6). Many models utilizing CT images alone, or using CT images with genetic factors, or combining CT images, clinical factors, and SNPs were also established to predict the sensitivity to platinum-based chemotherapy (13–15).

However, no consensus has been achieved on the best approach for predicting drug sensitivity. More importantly, a commonly

encountered clinical problem was that two similar patients received identical chemotherapies but responded significantly differently. These divergent responses were most likely to be determined by genetics. In most previous reports, only one or a few SNPs were studied in one specific clinical setting. Limited SNPs usually were insufficient for an accurate prediction of the drug response, which is typically a multi-gene process (7–12, 15, 16).

In this study, we systematically assessed the effects of a large amount of SNPs and several clinical factors regarding the drug sensitivity of EOC patients. Based on the results, we built a mathematical model that could predict the patient's response to platinum-based chemotherapy. The algorithm utilizes the genotyping information from 16 SNPs and can help gynecological oncologists stratify patients before the treatment.

2 Materials and methods

2.1 Study subjects

In this retrospective study, 148 EOC patients were recruited from the Fourth Hospital of Hebei Medical University between February 2008 and August 2016. The inclusion criteria for cases were as follows: i) Pathologically confirmed primary EOC at any age. ii) Clinically classified at stage III or IV according to the International Federation of Gynecology and Obstetrics (FIGO) criteria. No restriction on the histological subtypes of tumors. iii) After the standard primary cytoreductive surgery, the first-line treatment was the combination of paclitaxel and platinum; and the patient received the treatment for six to eight cycles. iv) The minimum follow-up period was five years or until the patient was deceased. The exclusion criteria included: i) Diagnosed as other types of cancers. ii) With a history of chemotherapy, radiotherapy, or immunotherapy before the surgery. iii) The patient received a non-standard primary cytoreductive surgery or a non-standard intravenous chemotherapy. The study was approved by the Ethics Committee of the Fourth Hospital of Hebei Medical University (2023KS119), and consent was obtained from all patients enrolled.

2.2 Surveillance and categorization of patients

The standard post-chemotherapeutical surveillances included serial physical examinations, the cancer antigen 125 (CA-125) blood test, and the computed tomography (CT) scanning as clinically indicated. According to the National Comprehensive Cancer Network (NCCN) guidelines, the recurrent disease was identified by comprehensive evaluation of supportive evidence such as clinical symptoms (i.e., pelvic pain and weight loss), blood biomarkers (i.e., elevated CA-125 levels), and imaging examinations (17). Patients who did not experience the relapse for not less than six months from the end of the last round of platinum-based therapy were defined as drug-sensitive, while patients who exhibited disease progression during chemotherapy or who experienced relapse within six months from the end of primary chemotherapy were considered to be drug-resistant (18). Progression-free survival (PFS) was defined as the time from the date of surgery to the first recurrence or the last follow-up. Overall survival (OS) was defined as the time from the diagnosis to the all-cause death.

2.3 DNA extraction and genotyping

The whole-blood sample (5ml) from each subject was collected and was used for the genomic DNA extraction as previously described by Miller et. al (19). DNA samples were quantified via spectrometry and were stored at -20°C in TE buffer. Genotypes of $\sim 700,000$ single nucleotide polymorphism (SNP) sites were determined by the Asian Screening Array system (Illumina, San Diego, CA, USA) following the manufacturer's instruction.

2.4 Construction of predictive model

2.4.1 SNP filtrations

All variants with a frequency of less than 0.05 or over 0.95 in the East Asian population were excluded, referring to the GnomAD v2.1.1 database (<https://gnomad.broadinstitute.org/>, last accessed on June. 26, 2024). The Genome-wide Association Study (GWAS) was performed to establish the association between the variant and the sensitivity to the platinum-based regimens. A cutoff value of $p < 0.05$ was used to further downsize the number of initial inputting SNPs.

2.4.2 Model developing strategy

The goal of our study is to build a mathematical model that can predict the sensitivity of epithelial ovarian cancer patients to platinum-based chemotherapeutic regimens. Although we believe that the sensitivity of EOC patients to the drugs is mainly determined by genetics, we cannot exclude the influence of non-genetic factors (age, stage, histology, and residual tumor). Firstly, we evaluated the correlation between the non-genetic factors and drug sensitivity using statistical tests. If any of these factors displayed a significant correlation with drug sensitivity, we would integrate that factor into our model. We also examined the genetic variants. Due to

the significant number of genetic variants (genotypes), we first used the Least Absolute Shrinkage and Selection Operator (LASSO) regression to reduce the total number of potential variants, and we manually checked the biological functions of the gene loci containing those possible variants. Only variants that had meaningful functions to platinum or paclitaxel were preserved. Then, another round of linear regression was applied to establish the model. Of all 148 EOC patients, 104 data sets were randomly chosen for training purposes, and the remaining 44 sets were for validation.

2.4.3 Initial assignment of SNPs

The genotype of each SNP site was transformed into a numerical value as its initial assignment. The human referential genome (version 37, GRCh37) was used as the standard and was referred to as the 'wild type'. The most common alternative sequence from the dbSNP database was referred to as the 'mutant'. A comparison was made between the patient's genotype and the referential sequences. If the sequences of the two alleles were both the same as the wild type, the initial assignment for that SNP would be '0'. If at least one of the sequences of two alleles was the same as the mutant, the initial assignment for that SNP would be '1'. If at least one allelic sequence differed from either the wild type or the mutant, a value of '0' would also be assigned to that site. The initial assigned value for SNP is an integer alternated as either 0 or 1.

2.4.4 Model development

The predictive model was constructed based on the LASSO regression algorithm implemented in the R (version 3.5.1) package 'glmnet'. The optimal regularization parameters were determined using a 10-fold cross-validation technique. Bootstrap analysis was utilized to sample the dataset with replacement 500 times. Subsequently, a model was constructed for each bootstrap cohort. The final model incorporated only markers that were observed in over 50% of all bootstraps.

2.5 Statistical analysis

Statistical analysis was performed with SPSS v24.0 for Windows (Chicago, IL, USA). The association between the patient's characteristics and the drug responses was assessed by the Mann-Whitney U test for continuous variables (i.e. age) and the Fisher's Exact test for categorical variables (i.e. Clinical histology). Survival curves were plotted by Kaplan-Meier analysis with the log-rank test. $P < 0.05$ was considered a significant difference.

3 Results

3.1 Characteristics of patients

As shown in Table 1, out of the 148 patients, 94 (63.5%) exhibited drug-sensitive, and 54 (36.5%) displayed drug-resistant. The median of PFS for drug-sensitive patients was 40 months, and the corresponding time for the drug-resistant patients was only

TABLE 1 The clinical characteristics of 148 EOC patients.

	Patients (n/%)	Median	Range
Age (years)			
<50	52 (35.1%)	45y	(34-49)y
≥50	96 (64.9%)	58y	(50-73)y
Clinical histology			
High grade serous	92 (62.2%)		
Low grade serous	18 (12.1%)		
Endometrioid	38 (25.7%)		
Stage			
III	136 (91.9%)		
IV	12 (8.1%)		
Residual tumor			
≤1cm	98 (66.2%)		
0cm	36 (24.3%)		
0-1cm	62 (41.9%)		
>1cm	50 (33.8%)		
Drug reaction			
Sensitive	94 (63.5%)		
Resistance	54 (36.5%)		
PFS of patients (months)			
PFS of sensitive		40m	(11-60)m
PFS of resistance		7m	(2-13)m
OS of patients (months)			
OS of sensitive		60m	(17-60)m
OS of resistance		22m	(5-53)m

seven months. The median of OS for drug-sensitive patients was 60 months, which was also significantly longer than that of the drug-resistant patients (22 months). In terms of histology, 92 (62.2%) out of the 148 patients were diagnosed with high-grade serous ovarian cancer, 18(12.1%) with low-grade serous ovarian cancer, and 38 (25.7%) with endometrioid ovarian cancer. There were 136 cases (91.9%) in FIGO stage III and 12 cases (8.1%) in FIGO stage IV. A total of 98 patients (66.2%) underwent satisfactory cytoreductive surgery, with 36 (24.3%) and 62 (41.9%) patients having residual tumor sizes of 0cm and 0-1cm, respectively. The remaining 50 patients (33.8%) underwent unsatisfactory cytoreductive surgery (residual tumor>1cm).

3.2 Influence of clinical factors on the drug sensitivity

As shown in Table 2, no clinical factor was significantly associated with drug sensitivity among all tested non-genetic features such as age, clinical histology, stage, and residual tumor. Notably, in our dataset,

11 pairs of patients shared the same clinical characteristics including surgical methods, age, histological type, stage, chemotherapeutical regimen, and residual tumor status yet substantially differed in the drug sensitivity (Supplementary Table S2).

3.3 Association of genetic variants with the drug sensitivity

As aforementioned we could not find a clinical factor that affected the drug sensitivity significantly in our sample set. Consequently, we employed a whole-genome SNP array with approximately 700,000 loci to test the association of those SNPs with drug sensitivity. After the GWAS analysis, a total of 446 SNP loci were identified that exhibited potential associations (*p*-value <0.05) with the drug sensitivity. However, 446 SNPs cannot directly be used to predict drug sensitivity, and we constructed a mathematical model to downsize the number of SNPs for prediction.

3.4 Construction of the model(drug sensitivity index)

Subsequently, we randomly divided 148 cases into the training set (104 cases, 70%) and the validation set (44 cases, 30%). The LASSO regression was applied to the 104 training datasets using 446 SNP loci. Briefly, each SNP site was initially assigned a value depending on its genotype (see method part). Next, the generalized linear model was applied using the R package. The nature of LASSO regression will automatically exclude non-significant variables. Then based on the remaining variables, we manually selected biologically meaningful SNP sites for the refinement. After the optimization, 16 effective SNP loci were selected for the predicting model. Among the 16 sites, seven had

TABLE 2 Statistical comparison of the clinical parameters between sensitive and resistant patients.

	Patients with sensitive (n/%)	Patients with resistance (n/%)	P
Age (range/median)	(34-73) y/55y	(45-73) y/53y	0.8418
Clinical histology			0.2599
High grade serous	54 (58.7%)	38 (41.3%)	
Low grade serous	12 (66.7%)	6 (33.3%)	
Endometrioid	28 (73.7%)	10 (26.3%)	
Stage			0.2116
III	84 (61.8%)	52 (38.2%)	
IV	10 (83.3%)	2 (16.7%)	
Residual tumor			0.2075
≤1cm	66 (67.3%)	32 (32.7%)	
>1cm	28 (56.0%)	22 (44.0%)	

a negative coefficient, and nine were positive, with values varying from about -0.29 to about 0.27 (Figure 1). The exact locations of these 16 sites and their precise coefficients are shown in Supplementary Table S1. By combining the assigned value and coefficient values of each site, a Drug Sensitivity Index (DSI) was defined as: $DSI = 0.50194 + (A_i \times C_i)$. Where: A_i = assigned value of the site i , which equals “0” or “1”, C_i = coefficient value of the site i , the detailed value for each site see Supplementary Table S1.

$$DSI = 0.50194 + (-0.29405 \times DS1value) + (-0.24243 \times DS2value) + (-0.17686 \times DS3value) + \dots + (0.27524 \times DS16value)$$

The threshold of DSI was set to 0.5. If $DSI \geq 0.5$ the patient would be predicted to be positive (drug-sensitive), and if $DSI < 0.5$ the patient would be predicted to be negative (drug-resistant). We checked the training set by this threshold setting. There were 65 true drug-sensitive patients with five misclassified and 39 true drug-resistant patients with six misclassified (Figure 2A). The overall performance of this model on the 104 sets of training data was as follows: sensitivity, 92.3%; specificity, 84.6%; positive predictive value, 90.9%; and accuracy 89.4%.

3.5 Validation of the model

Based on this model, the DSI values were calculated for the 44 validating samples, and the results are illustrated in Figure 2B. 29 true drug-sensitive patients were all correctly classified and three of fifteen true drug-resistant patients were misclassified as drug-sensitive. The sensitivity, specificity, positive predictive value, and accuracy of the model for validating samples were 100%, 80.0%, 90.6%, and 93.2%, respectively. If the two sets of data were combined, the overall sensitivity, specificity, positive predictive value, and accuracy of the model were 94.7%, 83.3%, 90.8%, and 90.5%, respectively.

3.6 Evaluate the predictive ability of the model

As shown in Figure 3, in both the training set and validating set, the Kaplan-Meier curve showed that the median PFS (Figures 3A, C) and mean OS (Figures 3B, D) were significantly longer for the predicted drug-sensitive patients compared to the predicted drug-resistant patients (PFS and OS in training group: 37m vs. 7m, 50.7m vs. 25.8m; PFS and OS in validation group: 31m vs. 8m, 54.0m vs. 26.0m). The log-rank test showed statistically significant differences in PFS and OS between these two predicted groups in both training and validating samples (all four $P < 0.001$).

4 Discussion

In the present study, we investigated the clinical and genetic factors associated with platinum-based drug sensitivity in EOC patients. We established a regularized linear regression model incorporating 16 SNPs for predicting the drug susceptibility of EOC patients. The overall sensitivity, specificity, positive predictive value, and accuracy of the model were 94.7%, 83.3%, 90.8%, and 90.5%, respectively. The Kaplan-Meier curves showed the PFS and OS were significantly different in the predicted sensitive and resistant groups, indicating that the model has high predictive power for the EOC patients' drug susceptibility.

Although Winarno G.N.A et al.'s study reported that the age and the FIGO stage were significant differences between the platinum-sensitive and platinum-resistant groups in EOC patients ($P < 0.05$) (5), our study showed that there was no significant difference between the two groups regarding the age and staging (Table 2). It was consistent with the found of Asami Ikeda et al. that age was not a factor affecting the response to platinum-based drugs (6). We believe the divergence between our study and Winarno G.N.A, et al.'s work was caused by the FIGO staging difference of included patients. In their work, the studied patients were FIGO

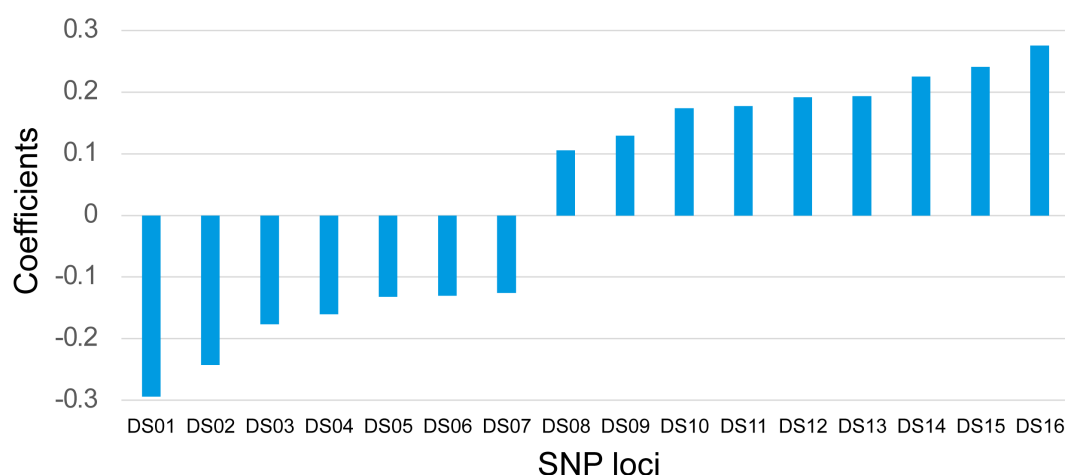


FIGURE 1

Coefficients of 16 SNP Loci. The coefficients of 16 SNP loci were shown for their synergistic powers on the prediction of the platinum-based drug sensitivity. X-axis: IDs of the SNP loci. Y-axis: Correlation coefficients.

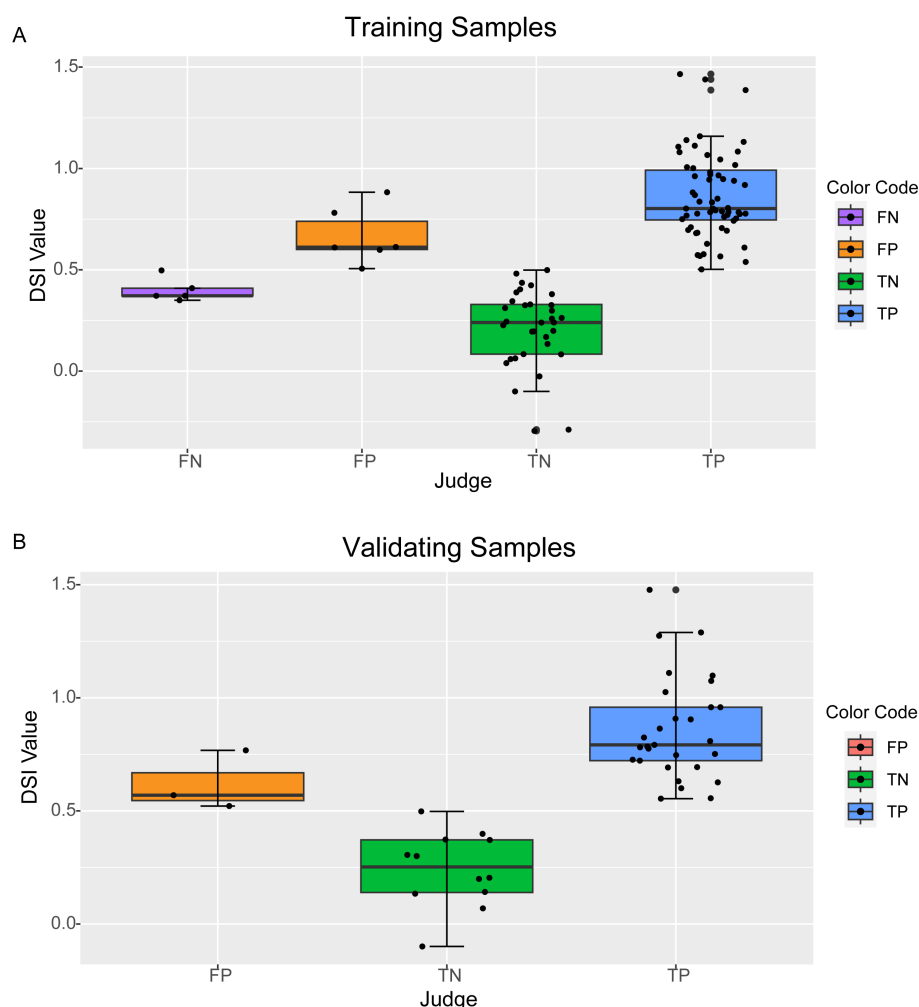


FIGURE 2

Performance of DSI in the training and validating samples. The DSI scores of each sample were calculated according to the LASSO regression model. The categories of prediction results were color indexed (FN= False Negative, purple; FP= False Positive, orange; TN=True Negative, green; TP=True Positive, blue). (A) Training samples. (B) Validating samples. X-axis: Prediction Results. Y-axis: DSI Scores.

stage I to III EOC, and 42.2% of patients presented with stage I in the platinum-sensitive group. But in our study, there were no stage I or II patients. Similarly, Buttitta F. et al. included only stage III and IV patients in their study, and they did not find a significant staging difference between the platinum-sensitive and resistant group either (20), which was in line with our result. In our study, we also found there was no significant difference between the two groups in histology types (Table 2), which was supported by other reports (5, 6, 21). As for the residual tumor, the previous multivariate analysis showed that the residual tumor was an independent predictor of platinum resistance. Compared to patients with the absence of macroscopic residual tumor (R:0cm), those with macroscopic residual tumor (R:0-1cm and >1cm) exhibited an increased risk of developing platinum resistance (HR:2.111, 95% CI: 1.415-3.151, $p<0.001$) (6). Another study reported that compared to patients with residual tumor <2cm, those with residual tumor >2cm had an elevated risk of developing platinum resistance (OR:13.455, 95% CI: 2.624-68.992, $p=0.0018$) (20). However, our study showed that there was no significant

difference between platinum-sensitive and resistant groups in residual tumor factor (Table 2). One possible explanation for this disagreement could be the differences in the criteria employed by different studies to group residual tumors. We adopted a grouped standard that closely aligns with clinical practice, which was whether a satisfactory cytoreductive surgery was conducted ($R \leq 1\text{cm}$ or $R > 1\text{cm}$). This classification criterion was consistent with that utilized by Mendiola M., et al. (22). Another important factor is that although the extent of residual tumor should be documented clearly by the operating gynecological surgeon, it is recognized that there are limitations to the accurate recording of residual tumor, in both the measurement of lesions and the quantification of the tumor residuum (23). The personal subjective factors of an operating gynecological surgeon may affect the accurate recording of residual tumors.

From the above results, we found that the clinical factors (age, stage, residual tumor, and histology types) may not exert a dominant influence on the patient's susceptibility to platinum-based chemotherapy. Furthermore, in our dataset, 11 pairs of

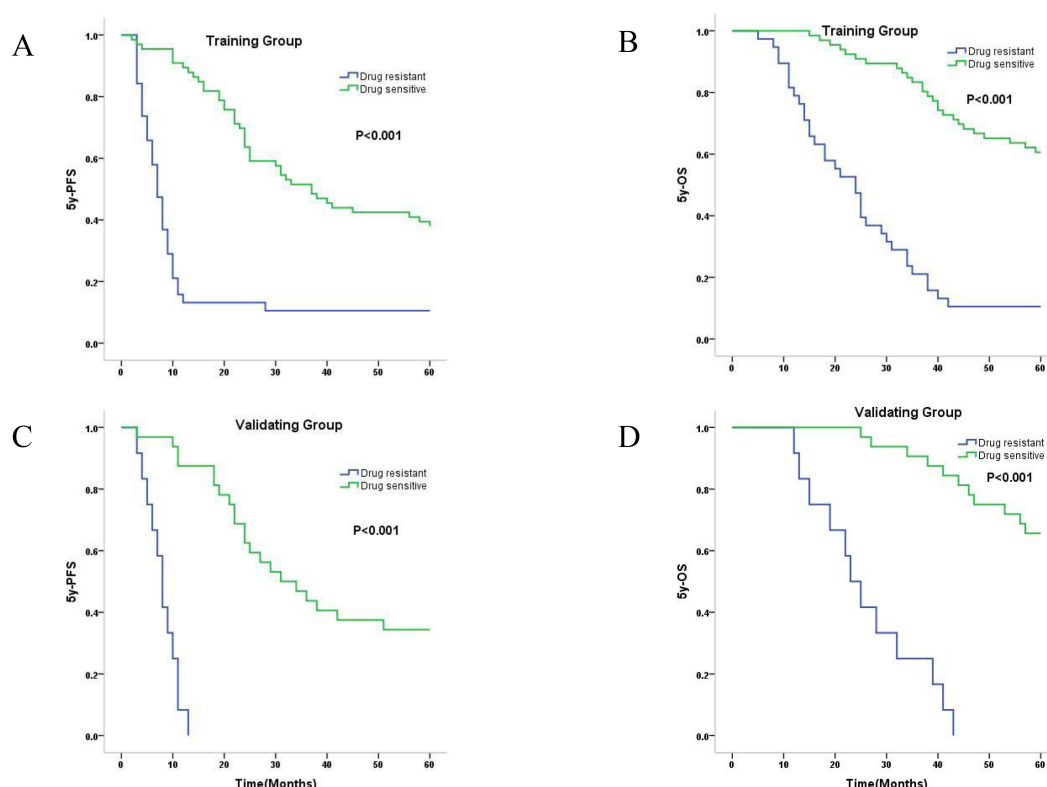


FIGURE 3

Kaplan-Meier survival curves of predicted drug-sensitive and resistant patients. (A) Median 5y-PFS of predicted drug-sensitive and resistant patients in training group were 37.0, and 7.0 months, respectively ($P < 0.001$). (B) Mean 5y-OS of predicted drug-sensitive and resistant patients in training group were 50.7, and 25.8 months, respectively ($P < 0.001$). (C) Median 5y-PFS of predicted drug-sensitive and resistant patients in validating group were 31.0, and 8.0 months, respectively ($P < 0.001$). (D) Mean 5y-OS of predicted drug-sensitive and resistant patients in validating group were 54.0, and 26.0 months, respectively ($P < 0.001$).

patients shared identical surgical methods, age, histological type, stage, chemotherapy regimen, and residual tumor status, yet their drug responses differed significantly, which strongly implied that the inter-individual genetic variants may exert a significant influence on drug sensitivity. The most prevalent genetic variation in the human genome arises from single nucleotide polymorphisms (24–26). Previous studies have identified specific SNPs associated with the response to platinum-based chemotherapy in EOC (7–12). However, these studies focused only on very few variants and lacked a comprehensive assessment of the synergistic effect of multiple variants. The isolated results from these studies could barely be used for clinical guidance while the complexity of the human genome brought many difficulties for solving the problem methodologically. To overcome this drawback, the whole-genome-wide SNP array was employed in this study to comprehensively investigate approximately 700,000 SNP sites within the human genome. The goals of the study were to: 1.) discover as many as possible variants that are relevant to the drug sensitivity; 2.) quantitatively determine the effects of these effective variants; 3.) build an accurate model for the drug sensitivity prediction that can be easily used by the gynecological oncologists in the clinic. Although the epigenetics may also impact the chemotherapy resistance in EOC (27), we did not compare our discoveries with epigenetics. Because in terms of platinum resistance in ovarian cancer, there are two types of

resistance: primary and acquired. The primary resistance means the patient showed resistance at the early stage of the platinum therapy (less than 6 months) and acquired resistance means the resistance shows up after a long period (over 12 months) (28). Generally, the primary resistance is determined by genetic variants and the acquired resistance is affected by alterations in multiple signaling pathways and usually is related to epigenetic modifications. Our study focuses on the primary resistance; therefore, we did not compare our discoveries with epigenetics.

One common problem for model construction is that the observed sample size (n) is much smaller than the number of predictive variables (p), and usually the variable selection (feature selection) is required. In our data, there is a small sample size (148 samples) and a high number of independent variables (about 700,000). It is obvious that $n \ll p$. To solve this problem the LASSO regression was introduced by Robert Tibshirani in the year of 1996 (29). The LASSO regression is widely employed as a regularization method for linear models. As an alternative to the least squares regression, LASSO regression introduces an L1 penalty function as a regularization term, effectively addressing the problem of multicollinearity, promoting variable selection (feature selection), and helping to improve the predictive ability and interpretability of the model (30–33). In GWAS analysis, when multiple SNP loci are used as independent variables, the L1 norm-

based feature selection can not only solve the problem of overfitting but also select important SNP loci (34). Combined with LASSO regression and biological feature selection, we successfully shrank the variables to 16 SNP sites. Then next round of least squares regression was applied. In this new scenario, the sample size (n) greatly exceeds the number of predictive variables (p), and the least squares regression worked well and showed highly accurate predicting capability (Figure 3). In the construction of the model, we also tested the elastic net regression which combined LASSO and Ridge regressions. However, the elastic net regression didn't show any better performance when checked by the ROAUC curve and also resulted in more leftover variants. And we also tried other models including binary logistic regression. However, similar to the elastic net regression, the performance is no better than the LASSO regression.

In the manual selection, 16 SNPs were preserved based on their biological functions (Supplementary Table S1). 12 SNPs were located within the exonic region, while one each was found in the promoter, enhancer, 5' UTR, and upstream regions. Notably, out of the 12 exonic SNPs, ten resulted in missense mutations, one led to a synonymous mutation, and another caused a nonsense mutation. When we searched the function of these 16 genes, three genes, including *LIG1* (DNA ligase 1), *VEGFR2* (one reception for VEGF), and *PPP4R3A* (a highly conserved member of the phosphatase family of serine/threonine phosphatases) have been demonstrated to be associated with the response to platinum and paclitaxel chemotherapy in EOC (35–41). Although the remaining 13 genes have not been conclusively associated with drug response in EOC they all have relevant bio-functions. The *L3MBTL3*, *PLA2G7*, *KIAA1614*, and *C3orf33* genes were implicated in DNA repair, cell proliferation, apoptosis, and the cell cycle, and may be involved in the development of cancers (42–46). *LAP3*, *ADRA1D*, *KIF25*, *GPS2*, and *MCF2L* genes were associated with proliferation, movement, and invasion of tumor cells (47–52). *CLEC7A* and *XIRP2* genes can contribute to cellular motility through their involvement with actin (53, 54). The *CEP128* gene encodes a centrosomal protein that is a fundamental subdistal appendage protein and functions on the mother centriole for the organization of the centriolar microtubules (55, 56). Finally, the *SLC25A39* gene was necessary for mitochondrial glutathione import in mammalian cells (57). Therefore, all those 16 SNPs may play some role in the pharmacology of platinum and paclitaxel combined chemotherapy.

Precision medicine, also known as personalized medicine, is a kind of medical care based on an individual's genetic information and is the main direction of future medicine (58, 59). The fundamental concept for it is that each therapy should be customized based on personal characteristics. Genetic features are the most predominantly used personal characteristics for precision medicine. In clinical practice, both congenital genetic traits (e.g. germline mutations of breast and ovarian cancers) and acquired genetic traits (e.g. somatic mutations of lung cancer, etc.) have been employed in targeted therapy and immunotherapy interventions (60–63). In this study, we aim to develop a precision medicine approach that enables gynecologic oncologists to effectively utilize genetic features. Compared with other methods, our approach is

relatively simplified and more clinically feasible. It just needs the blood sample, which can be easily obtained, and the genotype information, which can also be easily acquired either by fluorescent PCR, gene chip genotyping, or sequencing.

However, there were several limitations in our study. Firstly, the retrospective nature of this study may introduce selection bias. Secondly, the sample size in our study was modest, potentially impacting the stability and reliability of our model. Thirdly, this study is limited by its reliance on data from a single institution for both the training and validation cohorts. External validation using an independent cohort from other institutions will be necessary for the performance of our model. In future research, a prospective multicenter study should be conducted. Fourthly, we did not recruit any patients with other histotypes such as ovarian clear cell ovarian cancer in this study in addition to high-grade serous ovarian cancer, low-grade serous ovarian cancer and endometrioid ovarian cancer, although we did not restrict the histotypes in the inclusion criteria. Therefore, we restrict our conclusion within the histotypes included in our study. In future research, other histotypes may be concluded. Fifthly, the ethnic variances could compromise our conclusions in Western populations.

In summary, clinical factors may not exert a predominant influence on predicting the sensitivity of epithelial ovarian cancer to platinum-based chemotherapy, while a mathematical model that incorporates multiple genetic variations can achieve good outcomes in this context. Future prospective multicenter studies with a larger sample size should be conducted to validate our prediction model and to optimize it for clinical practice.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

The studies involving humans were approved by the Ethics Committee of the Fourth Hospital of Hebei Medical University (2023KS119). The participants provided their written informed consent to participate in this study.

Author contributions

N-YD: Data curation, Formal analysis, Software, Writing – original draft, Writing – review & editing. YL: Investigation, Resources, Writing – review & editing. HZ: Investigation, Supervision, Writing – original draft, Writing – review & editing. Y-KL: Data curation, Resources, Writing – review & editing. L-SL: Data curation, Writing – original draft. JX: Data curation, Writing – review & editing. SK: Funding acquisition, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by Hebei Province Medical Science Research Key Project (No. 20230771) and the Natural Science Foundation of Hebei Province (No. H2020206535).

Acknowledgments

The authors acknowledge the doctors at the Department of Gynaecology, the Fourth Hospital of Hebei Medical University, for their assistance in recruiting study subjects.

Conflict of interest

Authors HZ and JX were employed by the company Shijiazhuang Ninghong Biotechnology Co., Ltd.

References

- Torre LA, Trabert B, DeSantis CE, Miller KD, Samimi G, Runowicz CD, et al. Ovarian cancer statistics, 2018. *CA: Cancer J Clin.* (2018) 68:284–96. doi: 10.3322/caac.21456
- Jayson GC, Kohn EC, Kitchener HC, Ledermann JA. Ovarian cancer. *Lancet (London England).* (2014) 384:1376–88. doi: 10.1016/S0140-6736(13)62146-7
- Querleu D, Planchamp F, Chiva L, Fotopoulou C, Barton D, Cibula D, et al. European society of gynaecological oncology (ESGO) guidelines for ovarian cancer surgery. *Int J gynecological cancer: Off J Int Gynecological Cancer Society.* (2017) 27:1534–42. doi: 10.1097/IGC.0000000000001041
- Christie EL, Bowtell DDL. Acquired chemotherapy resistance in ovarian cancer. *Ann oncology: Off J Eur Soc Med Oncol.* (2017) 28:viii13–viii5. doi: 10.1093/annonc/mdx446
- Winarno GNA, Pasaribu M, Susanto H, Nisa AS, Harsono AB, Yuseran H, et al. The platelet to lymphocyte and neutrophil to lymphocyte ratios in predicting response to platinum-based chemotherapy for epithelial ovarian cancer. *Asian Pacific J Cancer prevention: APJCP.* (2021) 22:1561–6. doi: 10.31557/APJCP.2021.22.5.1561
- Ikeda A, Yamaguchi K, Yamakage H, Abiko K, Satoh-Asahara N, Takakura K, et al. Serum lactate dehydrogenase is a possible predictor of platinum resistance in ovarian cancer. *Obstetrics gynecology science.* (2020) 63:709–18. doi: 10.5468/ogs.20117
- Smith S, Su D, Rigault de la Longrais IA, Schwartz P, Puopolo M, Rutherford TJ, et al. ERCC1 genotype and phenotype in epithelial ovarian cancer identify patients likely to benefit from paclitaxel treatment in addition to platinum-based therapy. *J Clin oncology: Off J Am Soc Clin Oncol.* (2007) 25:5172–9. doi: 10.1200/JCO.2007.11.8547
- Steffensen KD, Waldström M, Jeppesen U, Brandslund I, Jakobsen A. Prediction of response to chemotherapy by ERCC1 immunohistochemistry and ERCC1 polymorphism in ovarian cancer. *Int J gynecological cancer: Off J Int Gynecological Cancer Society.* (2008) 18:702–10. doi: 10.1136/ijgc-00009577-200807000-00014
- Marmé F, Hielscher T, Hug S, Bondong S, Zeillinger R, Castillo-Tong DC, et al. Fibroblast growth factor receptor 4 gene (FGFR4) 388Arg allele predicts prolonged survival and platinum sensitivity in advanced ovarian cancer. *Int J cancer.* (2012) 131:E586–91. doi: 10.1002/ijc.27329
- Malisic E, Jankovic R, Jakovljevic K, Radulovic S. TP53 gene status and human papilloma virus infection in response to platinum plus taxane-based chemotherapy of epithelial ovarian carcinomas. *J BUON: Off J Balkan Union Oncol.* (2011) 16:701–7.
- Yan L, Shu-Ying Y, Shan K, Yip BH, Rong-Miao Z, Na W, et al. Association between polymorphisms of ERCC1 and survival in epithelial ovarian cancer patients with chemotherapy. *Pharmacogenomics.* (2012) 13:419–27. doi: 10.2217/pgs.11.181
- Zhao J, Wu J, Zuo W, Kang S, Li Y. A functional polymorphism in the poly (ADP-ribose) polymerase-1 gene is associated with platinum-based chemotherapeutic response and prognosis in epithelial ovarian cancer patients. *Eur J obstetrics gynecology Reprod Biol.* (2020) 255:183–9. doi: 10.1016/j.ejogrb.2020.10.021
- Wei W, Rong Y, Liu Z, Zhou B, Tang Z, Wang S, et al. Radiomics: a novel CT-based method of predicting postoperative recurrence in ovarian cancer. *Annu Int Conf IEEE Eng Med Biol Soc.* (2018) 2018:4130–3. doi: 10.1109/EMBC.2018.8513351
- Veeraraghavan H, Vargas HA, Jimenez-Sanchez A, Micco M, Mema E, Lakhman Y, et al. Integrated multi-tumor radio-genomic marker of outcomes in patients with high serous ovarian carcinoma. *Cancers.* (2020) 12:3403. doi: 10.3390/cancers12113403
- Yi X, Liu Y, Zhou B, Xiang W, Deng A, Fu Y, et al. Incorporating SULF1 polymorphisms in a pretreatment CT-based radiomic model for predicting platinum resistance in ovarian cancer treatment. *Biomedicine pharmacotherapy = Biomedecine pharmacotherapie.* (2021) 133:111013. doi: 10.1016/j.biopha.2020.111013
- Bolton KL, Ganda C, Berchuck A, Pharaoh PD, Gayther SA. Role of common genetic variants in ovarian cancer susceptibility and outcome: progress to date from the Ovarian Cancer Association Consortium (OCAC). *J Internal Med.* (2012) 271:366–78. doi: 10.1111/j.1365-2796.2011.02509.x
- Network NCCN. Ovarian cancer 2023 (Version 1.2023). Available online at: https://www.nccn.org/professionals/physician_gls/f_guidelines.asp (Accessed June 1, 2024).
- Gordon AN, Fleagle JT, Guthrie D, Parkin DE, Gore ME, Lacave AJ. Recurrent epithelial ovarian carcinoma: a randomized phase III study of pegylated liposomal doxorubicin versus topotecan. *J Clin oncology: Off J Am Soc Clin Oncol.* (2001) 19:3312–22. doi: 10.1200/JCO.2001.19.14.3312
- Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* (1988) 16:1215. doi: 10.1093/nar/16.3.1215
- Butti F, Pellegrini C, Marchetti A, Gadducci A, Cosio S, Felicioni L, et al. Human telomerase reverse transcriptase mRNA expression assessed by real-time reverse transcription polymerase chain reaction predicts chemosensitivity in patients with ovarian carcinoma. *J Clin oncology: Off J Am Soc Clin Oncol.* (2003) 21:1320–5. doi: 10.1200/JCO.2003.09.065
- Raunkaewmanee S, Tangjitgamol S, Manusirivithaya S, Srijaipracharoen S, Thavaramara T. Platelet to lymphocyte ratio as a prognostic factor for epithelial ovarian cancer. *J gynecologic Oncol.* (2012) 23:265–73. doi: 10.3802/jgo.2012.23.4.265
- Mendiola M, Redondo A, Heredia-Soto V, Herranz J, Berjón A, Hernández A, et al. Predicting response to standard first-line treatment in high-grade serous ovarian carcinoma by angiogenesis-related genes. *Anticancer Res.* (2018) 38:5393–400. doi: 10.21873/anticancer.12869
- Karam A, Ledermann JA, Kim JW, Sehouli J, Lu K, Gourley C, et al. Fifth Ovarian Cancer Consensus Conference of the Gynecologic Cancer InterGroup: first-line interventions. *Ann oncology: Off J Eur Soc Med Oncol.* (2017) 28:711–7. doi: 10.1093/annonc/mdx011
- Gibbs RA, Belmont JW, Hardenbol P, Willis TD, Yu F, Yang H, et al. The international hapMap project. *Nature.* (2003) 426:789–96. doi: 10.1038/nature02168
- Gibbs RA, Belmont JW, Boudreau A, Leal SM, Hardenbol P, Pasternak S, et al. A haplotype map of the human genome. *Nature.* (2005) 437:1299–320. doi: 10.1038/nature04226
- Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature.* (2007) 449:851–61. doi: 10.1038/nature06258

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2024.1461772/full#supplementary-material>

27. Feng LY, Yan BL, Huang YZ, Li L. Abnormal methylation characteristics predict chemoresistance and poor prognosis in advanced high-grade serous ovarian cancer. *Clin Epigenet.* (2021) 13:141. doi: 10.1186/s13148-021-01133-2
28. Slaughter K, Holman L, Thomas EL, Gunderson CC, Lauer JK, Ding K, et al. Primary and acquired platinum-resistance among women with high grade serous ovarian cancer. *Gynecol Oncol.* (2016) 142:225–30. doi: 10.1016/j.ygyno.2016.05.020
29. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B (Statistical Methodology).* (1996) 58:267–88. doi: 10.1111/j.2517-6161.1996.tb02080.x
30. Fu W, Knight K. Asymptotics for lasso-type estimators. *Ann Statist.* (2000) 28:1356–78. doi: 10.1214/aos/1015957397
31. Tibshirani R, Saunders M, Rosset S, Zhu J, Knight K. Sparsity and smoothness via the fused LASSO. *J R Stat Soc Ser B (Statistical Methodology).* (2005) 67:91–108. doi: 10.1111/j.1467-9868.2005.00490.x
32. Leng C, Lin Y, Wahba G. A note on the lasso and related procedures in model selection. *Statistica Sinica.* (2004) 16:1273–84. doi: 10.1007/s11135-006-3960-6
33. Meinshausen N, Bühlmann P. High-dimensional graphs and variable selection with the Lasso. *Ann Of Statistics.* (2006) 34:1436–62. doi: 10.1214/009053606000000281
34. Yang S, Wen J, Eckert ST, Wang Y, Liu DJ, Wu R, et al. Prioritizing genetic variants in GWAS with lasso using permutation-assisted tuning. *Bioinf (Oxford England).* (2020) 36:3811–7. doi: 10.1093/bioinformatics/btaa229
35. Ali R, Alabdullah M, Algethami M, Alblihy A, Miligy I, Shoaqfi A, et al. Ligase 1 is a predictor of platinum resistance and its blockade is synthetically lethal in XRC1 deficient epithelial ovarian cancers. *Theranostics.* (2021) 11:8350–61. doi: 10.7150/thno.51456
36. Nusrat O, Belotte J, Fletcher NM, Memaj I, Saed MG, Diamond MP, et al. The role of angiogenesis in the persistence of chemoresistance in epithelial ovarian cancer. *Reprod Sci (Thousand Oaks Calif).* (2016) 23:1484–92. doi: 10.1177/1933719116645191
37. Hastie CJ, Vázquez-Martín C, Philp A, Stark MJ, Cohen PT. The *Saccharomyces cerevisiae* orthologue of the human protein phosphatase 4 core regulatory subunit R2 confers resistance to the anticancer drug cisplatin. *FEBS J.* (2006) 273:3322–34. doi: 10.1111/j.1742-4658.2006.05336.x
38. Chen GI, Tisayakorn S, Jorgensen C, D'Ambrosio LM, Goudreau M, Gingras AC. PP4R4/KIAA1622 forms a novel stable cytosolic complex with phosphoprotein phosphatase 4. *J Biol Chem.* (2008) 283:29273–84. doi: 10.1074/jbc.M803443200
39. Chowdhury D, Xu X, Zhong X, Ahmed F, Zhong J, Liao J, et al. A PP4-phosphatase complex dephosphorylates gamma-H2AX generated during DNA replication. *Mol Cell.* (2008) 31:33–46. doi: 10.1016/j.molcel.2008.05.016
40. Nakada S, Chen GI, Gingras AC, Durocher D. PP4 is a gamma H2AX phosphatase required for recovery from the DNA damage checkpoint. *EMBO Rep.* (2008) 9:1019–26. doi: 10.1038/embor.2008.162
41. Byun HJ, Kim BR, Yoo R, Park SY, Rho SB. sMEK1 enhances gemcitabine anticancer activity through inhibition of phosphorylation of Akt/mTOR. *Apoptosis: an Int J programmed Cell death.* (2012) 17:1095–103. doi: 10.1007/s10495-012-0751-0
42. Kang K, Bae JH, Han K, Kim ES, Kim TO, Yi JM. A genome-wide methylation approach identifies a new hypermethylated gene panel in ulcerative colitis. *Int J Mol Sci.* (2016) 17:1291. doi: 10.3390/ijms17081291
43. Stafforini DM. Diverse functions of plasma PAF-AH in tumorigenesis. *Enzymes.* (2015) 38:157–79. doi: 10.1016/bs.enz.2015.09.005
44. Hall D, Giaimo BD, Park SS, Hemmer W, Friedrich T, Ferrante F, et al. The structure, binding and function of a Notch transcription complex involving RBPJ and the epigenetic reader protein L3MBTL3. *Nucleic Acids Res.* (2022) 50:13083–99. doi: 10.1093/nar/gkac1137
45. Hao D, Gao P, Liu P, Zhao J, Wang Y, Yang W, et al. AC3-33, a novel secretory protein, inhibits Elk1 transcriptional activity via ERK pathway. *Mol Biol Rep.* (2011) 38:1375–82. doi: 10.1007/s11033-010-0240-x
46. Eferl R, Wagner EF. AP-1: a double-edged sword in tumorigenesis. *Nat Rev Cancer.* (2003) 3:859–68. doi: 10.1038/nrc1209
47. Colciago A, Mornati O, Ferri N, Castelnovo LF, Fumagalli L, Bolchi C, et al. A selective alpha1D-adrenoreceptor antagonist inhibits human prostate cancer cell proliferation and motility “in vitro”. *Pharmacol Res.* (2016) 103:215–26. doi: 10.1016/j.phrs.2015.11.017
48. Wang X, Shi L, Deng Y, Qu M, Mao S, Xu L, et al. Inhibition of leucine aminopeptidase 3 suppresses invasion of ovarian cancer cells through down-regulation of fascin and MMP-2/9. *Eur J Pharmacol.* (2015) 768:116–22. doi: 10.1016/j.ejphar.2015.10.039
49. Mosley JD, Van Driest SL, Larkin EK, Weeke PE, Witte JS, Wells QS, et al. Mechanistic phenotypes: an aggregative phenotyping strategy to identify disease mechanisms using GWAS data. *PloS One.* (2013) 8:e81503. doi: 10.1371/journal.pone.0081503
50. Okamoto S, Matsushima M, Nakamura Y. Identification, genomic organization, and alternative splicing of KNSL3, a novel human gene encoding a kinesin-like protein. *Cytogenetics Cell Genet.* (1998) 83:25–9. doi: 10.1159/000015159
51. Si Y, Zhang H, Peng P, Zhu C, Shen J, Xiong Y, et al. G protein pathway suppressor 2 suppresses gastric cancer by destabilizing epidermal growth factor receptor. *Cancer science.* (2021) 112:4867–82. doi: 10.1111/cas.v112.12
52. Liu Z, Adams HC 3rd, Whitehead IP. The rho-specific guanine nucleotide exchange factor Dbs regulates breast cancer cell migration. *J Biol Chem.* (2009) 284:15771–80. doi: 10.1074/jbc.M901853200
53. Choraghe RP, Kołodziej T, Buser A, Rajfur Z, Neumann AK. RHOA-mediated mechanical force generation through Dectin-1. *J Cell Sci.* (2020) 133:jcs236166. doi: 10.1242/jcs.236166
54. Huang L, Wu KH, Zhang L, Wang Q, Tang S, Wu Q, et al. Critical roles of xirp proteins in cardiac conduction and their rare variants identified in sudden unexplained nocturnal death syndrome and brugada syndrome in chinese han population. *J Am Heart Assoc.* (2018) 7. doi: 10.1161/JAHA.117.006320
55. Mazo G, Soplop N, Wang WJ, Uryu K, Tsou MF. Spatial control of primary ciliogenesis by subdistal appendages alters sensation-associated properties of cilia. *Dev Cell.* (2016) 39:424–37. doi: 10.1016/j.devcel.2016.10.006
56. Kashiwara H, Chiba S, Kanno SI, Suzuki K, Yano T, Tsukita S. Cep128 associates with Odf2 to form the subdistal appendage of the centriole. *Genes to cells: devoted to Mol Cell mechanisms.* (2019) 24:231–43. doi: 10.1111/gtc.2019.24.issue-3
57. Wang Y, Yen FS, Zhu XG, Timson RC, Weber R, Xing C, et al. SLC25A39 is necessary for mitochondrial glutathione import in mammalian cells. *Nature.* (2021) 599:136–40. doi: 10.1038/s41586-021-04025-w
58. Ashley EA. The precision medicine initiative: a new national effort. *Jama.* (2015) 313:2119–20. doi: 10.1001/jama.2015.3595
59. Collins FS, Varmus H. A new initiative on precision medicine. *New Engl J Med.* (2015) 372:793–5. doi: 10.1056/NEJMp1500523
60. Zaman A, Bivona TG. Emerging application of genomics-guided therapeutics in personalized lung cancer treatment. *Ann Trans Med.* (2018) 6:160. doi: 10.21037/atm.2018.05.02
61. Raez LE, Cardona AF, Santos ES, Catoe H, Rolfo C, Lopes G, et al. The burden of lung cancer in Latin-America and challenges in the access to genomic profiling, immunotherapy and targeted treatments. *Lung Cancer (Amsterdam Netherlands).* (2018) 119:7–13. doi: 10.1016/j.lungcan.2018.02.014
62. Cortesi L, Rugo HS, Jackisch C. An overview of PARP inhibitors for the treatment of breast cancer. *Targeted Oncol.* (2021) 16:255–82. doi: 10.1007/s11523-021-00796-4
63. Walsh CS. Two decades beyond BRCA1/2: Homologous recombination, hereditary cancer risk and a target for ovarian cancer therapy. *Gynecologic Oncol.* (2015) 137:343–50. doi: 10.1016/j.ygyno.2015.02.017