# 博 士 学 位 論 文

卵巣高異型度漿液性癌では腫瘍内不均一性と
相同組換え修復異常は予後および分子サブタイプ
に関連し治療経過により変化する

近 畿 大 学 大 学 院
医 学 研 究 科 医 学 系 専 攻

髙 矢 寿 光

Doctoral Dissertation

# Intratumor heterogeneity and homologous recombination deficiency of high-grade serous ovarian cancer are associated with prognosis and molecular subtype and change in treatment course

November 2019

Department of Obstetrics and Gynecology, Major in Medical Sciences

Kindai University Graduate School of Medical Sciences

Hisamitsu Takaya

# 同　意　書

2019 年 11 月 11 日

近畿大学大学院
　医学研究科長　殿

| | | |
|---|---|---|
| 共著者 ＿＿＿＿＿ 印 | | 共著者 ＿＿＿＿＿ 印 |
| 共著者 ＿＿＿＿＿ 印 | | 共著者 ＿＿＿＿＿ 印 |
| 共著者 ＿＿＿＿＿ 印 | | 共著者 ＿＿＿＿＿ 印 |
| 共著者 ＿＿＿＿＿ 印 | | 共著者 ＿＿＿＿＿ 印 |

論文題目

Intratumor heterogeneity and homologous recombination deficiency of high-grade serous ovarian cancer are associated with prognosis and molecular subtype and change in treatment course

　下記の学位論文提出者が、標記論文を貴学医学博士の学位論文（主論文）

として使用することに同意いたします。

　また、標記論文を再び学位論文として使用しないことを誓約いたします。

記

1．学位論文提出者氏名　髙木 寿元

2．専攻分野　　医学系　女性機能病態‐周産期医学

# Intratumor heterogeneity and homologous recombination deficiency of high-grade serous ovarian cancer are associated with prognosis and molecular subtype and change in treatment course

Hisamitsu Takaya [a], Hidekatsu Nakai [a], Kazuko Sakai [b], Kazuto Nishio [b], Kosuke Murakami [a], Masaki Mandai [c], Noriomi Matsumura [a]

[a] Department of Obstetrics and Gynecology, Kindai University Faculty of Medicine
[b] Department of Genome Biology, Kindai University Faculty of Medicine
[c] Department of Gynecology and Obstetrics, Kyoto University Graduateschool of Medicine

Abstract

Objective:

High-grade serous ovarian cancers (HGSOC) are genomically characterized by homologous recombination deficiency (HRD) and *TP53* mutations, which lead to intratumor heterogeneity (ITH). This study aimed to reveal the relationship between HRD, ITH and prognosis and analyze their changes during treatment.

Methods:

We obtained 573 SNP array and gene expression array data from The Cancer Genome Atlas. SNP array data were processed to calculate the Clonality Index (CI) and loss of heterozygosity (LOH) scores. Gene expression array data were used for classifying molecular subtypes. Additionally, we obtained 33 samples from 20 HGSOC patients, including 4 samples from interval debulking surgery (IDS) and 9 samples from recurrent surgery.

Results:

We divided HGSOC samples into 2 groups. The high CI group showed a high recurrent risk, and the high LOH group showed a statistically good prognosis. Combining the two factors, the high LOH/low CI group showed a statistically good prognosis. In terms of molecular subtypes, the mesenchymal subtype, which had a poor prognosis, showed a high CI with statisitically significant difference and the immunoreactive subtype, which had a good prognosis, showed a tendency to have a high LOH score. Throughout treatment, the CI decreased to one at the IDS (n=4) and then increased at recurrence (n=3). LOH scores greatly decreased in two cases at the IDS.

Conclusions:

ITH and HRD were associated with prognosis in HGSOC. ITH decreased after neoadjuvant chemotherapy, suggesting that the chemo-resistant cancer clone remains after chemotherapy.

Keywords : Ovarian cancer; Clonality; Loss of heterozygosity; Chemo-resistance

## 1. Introduction

Ovarian cancer has a poor prognosis among gynecologic malignancies [1]. High-grade serous ovarian cancer (HGSOC), which is the most frequent epithelial ovarian cancer, is commonly diagnosed as an advanced cancer [2]. HGSOC is mainly treated by chemotherapy; however, most HGSOCs recur as a chemo-resistant tumor even though the first chemotherapy is usually effective and results in lethal treatment.

The Cancer Genome Atlas (TCGA) has revealed some characteristics of the cancer genome of HGSOC. One of the characteristics is homologous recombination repair deficiency (HRD) [3]. Almost half of HGSOCs have HRD due to *BRCA1/2* mutation or other homologous recombination repair pathway-related gene alterations [4]. HRD is associated with sensitivity to platinum agents [5] and poly-(ADP ribose) polymerase (PARP) inhibitors [6-8], and quantifying the genomic loss of heterozygosity (LOH) provides the possibility of measuring HRD as a biomarker [9]. In the ARIEL2 study, the LOH score was used as a treatment biomarker for the PARP inhibitor, and it was shown that the PARP inhibitor was effective in cases with a high LOH score [10].

The *TP53* mutation can be detected at a high frequency in HGSOC [3]. *TP53* mutations give rise to chromosomal instability due to disruption of the control of the cell cycle and apoptosis [11]. Chromosomal instability causes subclonal evolution, which shows different genomic characteristics in tumorigenesis and tumor progression and results in high intratumor heterogeneity [12, 13], which is associated with resistance to treatment and poor prognosis [14-16]. We have already reported that a single nucleotide polymorphism (SNP) array with formalin-fixed paraffin-embedded specimens can be used to analyze intratumor heterogeneity [17]. With this method, we can evaluate intratumor heterogeneity from various clinical specimens and apply the findings to clinical features, such as prediction of prognosis.

Although each of these main characteristics of HGSOC was investigated in several studies, no reports have examined the association between HRD and intratumor heterogeneity and how these factors changed before and after chemotherapy. In the current study, we first investigated the association between prognosis and HRD or intratumor heterogeneity with the use of TCGA data, and second, we analyzed the changes in these factors throughout treatment of HGSOC with samples from our institution.

## 2. Materials and methods

### 2.1. Data sources

We obtained TCGA CEL-formatted SNP array data from Affymetrix Genome-Wide Human SNP Array 6.0 via the Genomic Data Commons (GDC) Data Portal (https://portal.gdc.cancer.gov) and extracted 573 HGSOC cases with matched tumor-normal data. We obtained TCGA CEL-formatted gene expression array data from the GeneChip HT Human Genome U133A data via the GDC Data Portal and extracted cases having the same samples as the SNP array data. Additionally, we obtained TCGA XML-formatted clinical data for each case via the GDC Data Portal.

## ２．２．Clonality Index（CI）estimation

We previously reported a method to estimate the clonal composition using the OncoScan FFPE Assay Kit [17]. The logR ratio (LRR) and B-allele frequency (BAF) can be calculated with the copy number of tumor cells, proportion of aberrant cells (%AC), and copy number of minor alleles (NOMA), theoretically. When we plot the LRR and logarithm of BAF in 2-dimensional space, if the %AC is the same and NOMA equals zero, the plot makes a straight line. Therefore, we can calculate the %AC from the observed LRR and BAF if the NOMA equals zero in a segment, and we can estimate the number of clones of the tumor as the CI by analyzing the distribution of a set of the %AC.

We created an algorithm for calculating the CI under the environment of Python 3.6 and R version 3.5.0 as described above. The segmentation data of the LRR and BAF were obtained from CEL-formatted SNP array data using the R package "rawcopy" [18]. We excluded segmentation data whose base pair length was equal to or less than the threshold (default settings = 1000 base pairs). We calculated the weighted Euclidean distance between the coordinates of the observed LRR/observed BAF and the theoretical LRR/theoretical BAF in each %AC, and then the %AC at which the distance was minimized and less than 0.1 was determined as the %AC for that segment. If the number of output %AC was one or less, the CI was not unable to be estimated. Because the distribution of a set of %AC in the samples was unimodal or multimodal, we could estimate the most appropriate shape of the distribution that minimized the Bayesian information criterion using the R package "Mclust", and the number of clusters of the distribution was taken as the CI. If a cluster included less than 1% of the total %AC, that cluster was excluded.

## ２．３．LOH score calculation

CEL-formatted SNP array files were processed by Affymetrix Power Tools and PennCNV software [19] to calculate the LRR and BAF data in each probe, and then segmented copy number data were analyzed using the R package "ASCAT" [20]. The LOH score was calculated as the proportion of the sum of the LOH region whose length was equal to or greater than 15 M base pairs to the length of all chromosomes, as previously reported [10, 21]. We excluded the LOH region, which was more than 90% of the short arm, long arm, or total chromosome, in each chromosome.

## ２．４．Analysis of gene expression array data

CEL-formatted gene expression array data were normalized by the Robust Multiarray Average (RMA) method using the R package "aroma.light". Single sample gene set enrichment analysis (ssGSEA) was performed to classify HGSOC samples into four gene expression subtypes as previously defined [22]. Differential gene expression analysis was performed by Welch's t-test, and p-values were corrected by the Benjamini-Hochberg method. Pathway analysis was performed by Ingenuity Pathway Analysis (IPA; QIAGEN).

## 2．5．Clinical data extraction

XML-formatted clinical data files were read using Python, and the clinical information was extracted, including age at diagnosis, the International Federation of Gynecology and Obstetrics (FIGO) stage, treatment of neoadjuvant chemotherapy (NAC), site of analyzed tumor, residual tumor size, primary outcome, survival data, and vital status.

## 2．6．Analysis of our HGSOC samples

Thirty-three tumor samples from 20 HGSOC cases who underwent primary surgery between 1994 and 2012 in Kindai University Hospital, including 20 samples from primary surgery (14 samples from ovary, 4 samples from peritoneal metastasis, and 2 samples from omentum), 4 samples from interval debulking surgery (IDS) after NAC (3 samples from ovary and 1 sample from omentum), and 9 samples from recurrent surgery (7 samples from peritoneal metastasis and 1 sample from lymph node and brain metastasis), were retrospectively analyzed. The pathological diagnosis of HGSOC was diagnosed at the Central Pathological Department in Kindai University Hospital using either hematoxylin and eosin (HE) -stained slides or additional immunohistochemistry analysis, the results of which were reviewed by at least one gynecologic oncologist. The formalin-fixed, paraffin-embedded (FFPE) tumor specimens stained with HE were reviewed to confirm the presence of more than 50% of viable tumor cells by at least one author specializing in gynecologic pathology and oncology after the confirmation of the Central Pathological Department, and the tumor regions were removed manually. Genomic DNA was extracted from FFPE specimens using an AllPrep DNA/RNA FFPE Kit (QIAGEN). The quality of DNA was analyzed using a NanoDrop 2000/2000c Spectrophotometer (ThermoFisher Scientific), and the quantity of DNA was measured using a Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher Scientific).

DNA extracted from FFPE specimens was analyzed with an OncoScan FFPE Assay Kit (ThermoFisher Scientific). CEL-formatted files were output with Affymetrix GeneChip Command Console software version 4.0 and converted to OSCHP files with OncoScan Console software 1.3. The number of clones was estimated with the OncoClone Composition program [17] using OSCHP files. OSCHP files including the LRR and BAF data in each probe were analyzed by ASCAT to generate copy number segmented data, and the LOH score was calculated. Samples with aberrant cell fractions less than 30% were excluded as previously reported [23]. GISTIC 2.0 [24] was used to identify copy number aberrations in tumors.

## 2．7．Statistical Analysis

The correlation analysis between the CI and the LOH score was performed with Spearman correlation coefficient. The comparison between molecular subtypes was performed with the Kruskal-Wallis test. The CI was dichotomized based on a median split as performed in previous studies, where authors performed a median split to dichotomize samples into two subgroups [25-27] to illustrate the association between the CI

and survival time. The LOH score was also dichotomized based on the Gaussian mixture model. The survival curves were described with Kaplan-Meier estimation, and survival analysis was performed with the log-rank test. To consider confounding factors with respect to the survival analysis, covariates such as age at diagnosis (as a continuous variable), FIGO stage (I/II vs III/IV), and residual tumor size (no macroscopic disease vs 1-10 mm vs > 10 mm) were applied to the Cox proportional hazards model along with the CI and the LOH score to examine the association with the CI or the LOH score on survival time after adjusting these covariates. All statistical tests were performed with R, and a p-value of < 0.05 was considered statistically significant.

## 3. Results

CI and LOH scores could be estimated in 536 and 542 samples out of 573 samples, respectively. Both CI and LOH scores could be calculated in 502 samples, so 502 samples were subjected to analysis. The background data for patients in each sample set are shown in Supplementary Table 1. The averages and standard deviations of the CI and LOH scores were $3.53 \pm 1.95$ and $17.90 \pm 9.41$, respectively (Figure 1A, 1B). The correlation coefficient between the CI and LOH scores was -0.176, which indicated a slight negative correlation (Figure 1C).

Supplementary Table 1.　Background of TCGA patients

|  | All cases (N=573) | Analyzed (N=502) |
|---|---|---|
| Age, median (range) | 59 (26-89) | 59 (30-89) |
| FIGO stage (n, %) | | |
| I | 17 (3.0%) | 14 (2.8%) |
| II | 29 (5.1%) | 29 (5.8%) |
| III | 436 (76.1%) | 386 (76.9%) |
| IV | 86 (15.0%) | 73 (14.5%) |
| N/A | 5 (0.8%) | 0 (0%) |
| Neoadjuvant chemotherapy (n, %) | | |
| No | 571 (99.6%) | 502 (100%) |
| Yes | 1 (0.2%) | 0 (0%) |
| N/A | 1 (0.2%) | 0 (0%) |
| Residual tumor size (n, %) | | |
| No macroscopic disease | 118 (20.6%) | 110 (21.9%) |
| 1-10mm | 250 (43.6%) | 219 (43.6%) |
| 11-20mm | 34 (5.9%) | 30 (6.0%) |
| >20mm | 104 (18.2%) | 92 (18.3%) |
| N/A | 67 (11.7%) | 51 (10.2%) |
| Site of analyzed tumor tissue (n, %) | | |
| Ovary | 567 (99.0%) | 497 (99.0%) |
| Omentum | 4 (0.7%) | 3 (0.6%) |
| Peritoneum | 2 (0.3%) | 2 (0.4%) |

FIGO; International Federation of Gynecology and Obstetrics, N/A; not available
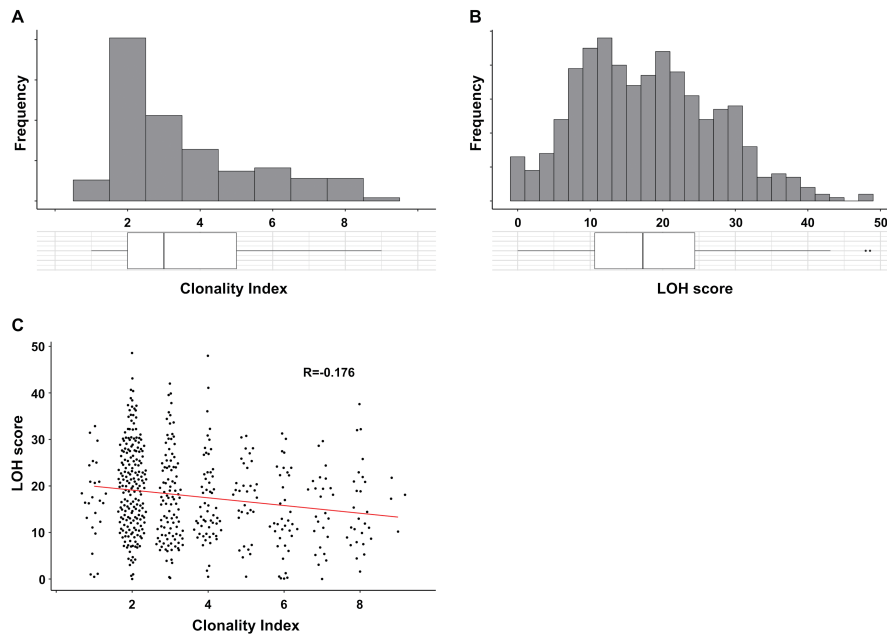
Figure 1. Histogram and boxplot of Clonality Index (CI) (A) and LOH scores (B) and pair plot of CI and LOH scores (C). The mean CI was 3.53, and the median was 3. The mean LOH score was 17.9, and the median was 17.26. Correlation between CI and LOH score. The Spearman correlation coefficient for the CI and LOH scores was -0.176.

We divided samples into 2 groups based on a median split: a high CI group, which had scores of three or higher, and a low CI group, which had scores lower than 3. In the survival analysis, the median progression-free survival (PFS) of the high and low CI groups was 16.6 months vs 20.6 months (p-value = 0.0097), respectively, with a significant difference observed. However, the median overall survival (OS) in the high and low CI groups was 44.5 months and 48.2 months (p-value = 0.419), respectively (Figure 2A, 2B).

We set a cut-off value of 16 for the LOH score based on the Gaussian mixture model and divided the samples into 2 groups: the high LOH group, which had scores of 16 or more, and the low LOH group, which had scores lower than 16. In the survival analysis, the median PFS and OS in the high LOH and low LOH groups was 19.3 months vs 16.3 months (p-value = 0.0056) and 50.0 months vs 38.6 months (p-value < 0.001), respectively, with significant differences observed (Figure 2C, 2D).

In the Cox proportional hazards model adjusting for age, FIGO stage, and residual tumor size, high CI was independently associated with a shorter PFS (adjusting hazard ratio (HR) = 1.473, 95% CI: = 1.140-1.904, p-value = 0.003, Supplementary Table 2) and high LOH score was independently associated with a longer PFS and OS (adjusting HR = 0.698 and 0.644, 95% CI = 0.538-0.905 and 0.501-0.828, p-value = 0.007 and < 0.001, respectively, Supplementary Table 3). When both the CI and the LOH score were simultaneously included in the Cox model with the covariates, the CI remained significantly associated with a shorter PFS (adjusting HR = 1.392, 95% CI = 1.072-1.808, p-value = 0.0132), and the LOH score remained indicative of a better prognosis (OS: adjusting HR = 0.648, 95% CI = 0.502-0.834, p-value < 0.001, PFS: adjusting HR = 0.747, 95% CI = 0.574-0.973, p-value = 0.031).

We stratified samples into 4 groups based on the combination of the CI and LOH scores: a high LOH/high CI group, a high LOH/low CI group, a low LOH/high CI group, and a low LOH/group LOH group. We subsequently performed a survival analysis. Comparing the high LOH/low CI group with the high LOH/high CI, low LOH/high CI, and low LOH/low CI groups, the median PFS was 22.6 months vs 16.8, 16.6, and 15.0 months (p-value < 0.001), and the median OS was 55.1 months vs 49.5, 38.6, and 38.6 months (p-value < 0.001), respectively. There were statistically significant differences (Figure 2E, 2F).

Supplementary Table 2. Cox proportional hazards model for overall survival and progression free survival in TCGA dataset with the Clonality Index adjusted for covariates

| Variable | Overall survival | | | Progression free survival | | |
|---|---|---|---|---|---|---|
| | HR | 95% CI | p-value | HR | 95% CI | p-value |
| CI (high vs low) | 1.095 | 0.853 - 1.406 | 0.475 | 1.473 | 1.140 - 1.904 | 0.003 |
| Age at diagnosis | 1.026 | 1.014 - 1.037 | < 0.001 | 1.007 | 0.995 - 1.018 | 0.244 |
| FIGO stage (I/II vs III/IV) | 2.323 | 0.736 - 7.337 | 0.151 | 7.196 | 1.773 - 29.205 | 0.006 |
| Residual disease (no macroscopic disease vs 1-10 mm vs >10 mm) | 1.420 | 1.184 - 1.703 | < 0.001 | 1.308 | 1.096 - 1.561 | 0.003 |

CI; Clonality Index, FIGO; International Federation of Gynecology and Obstetrics, HR; hazard ratio, 95% CI; 95% confidence interval

Supplementary Table 3. Cox proportional hazards model for overall survival and progression free survival in TCGA data set with the LOH score adjusted for covariates

| Variable | Overall survival | | | Progression free survival | | |
|---|---|---|---|---|---|---|
| | HR | 95% CI | p-value | HR | 95% CI | p-value |
| LOH score (high vs low) | 0.644 | 0.501 - 0.828 | < 0.001 | 0.698 | 0.538 - 0.905 | 0.007 |
| Age at diagnosis | 1.021 | 1.010 - 1.033 | < 0.001 | 1.005 | 0.993 - 1.017 | 0.466 |
| FIGO stage (I/II vs III/IV) | 2.033 | 0.643 - 6.428 | 0.227 | 6.588 | 1.623 - 26.765 | 0.008 |
| Residual disease (no macroscopic disease vs 1-10 mm vs >10 mm) | 1.435 | 1.197 - 1.721 | < 0.001 | 1.273 | 1.066 - 1.520 | 0.008 |

LOH score; loss of heterozygosity score, FIGO; International Federation of Gynecology and Obstetrics, HR; hazard ratio, 95% CI; 95% confidence interval
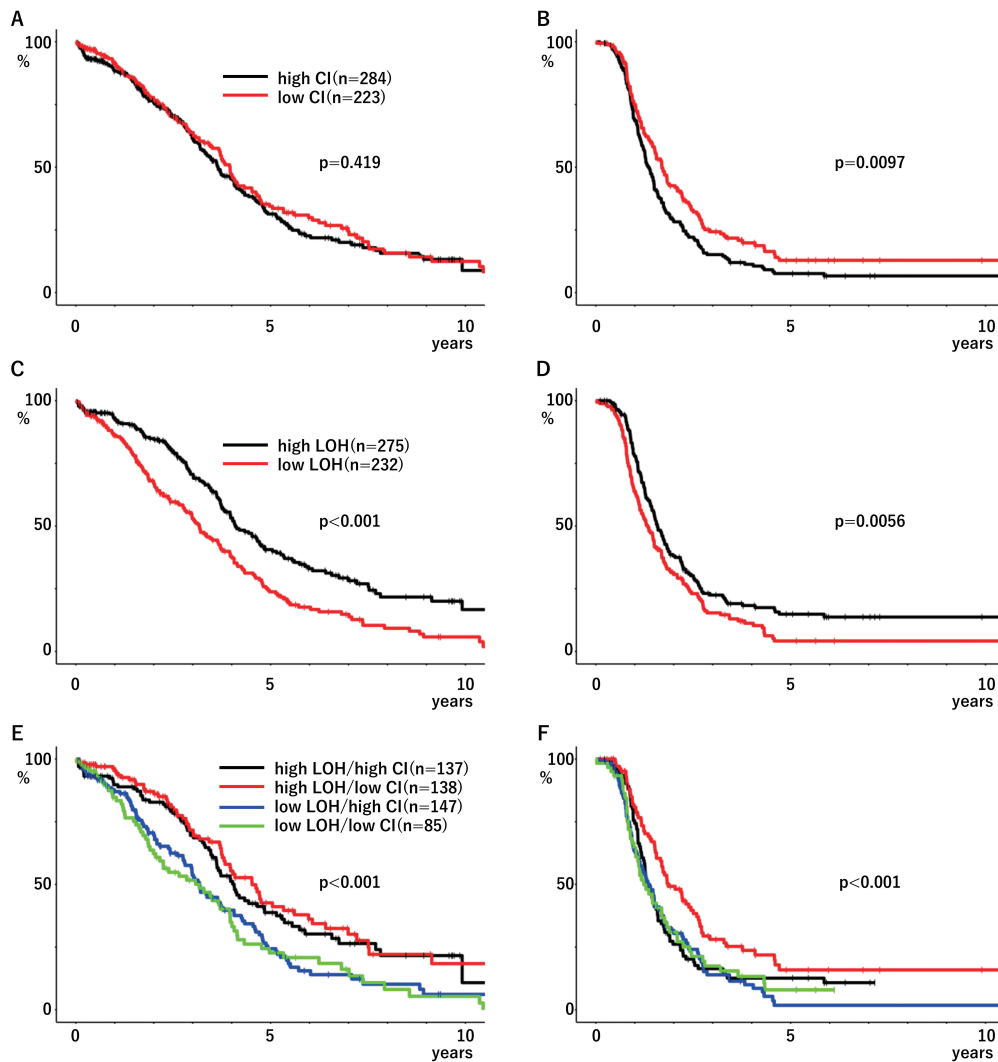
Figure 2. Kaplan-Meier curve of TCGA data stratified with the Clonality Index (CI) value (A, B), LOH score (C, D), and a combination of both (E, F) for overall survival (OS) (A, C, E) and progression-free survival (PFS) (B, D, F). The low CI group showed significantly prolonged PFS but not OS. The high LOH group showed prolonged OS and PFS with significant differences. The combined high LOH and low CI group showed good prognosis compared with the other groups.

Five hundred and twenty-two gene expression array data from the TCGA database were subject to ssGSEA with the CLOVAR gene set to classify 4 molecular subtypes: differentiated, immunoreactive, mesenchymal, and proliferative. Subsequently, 454 samples were analyzed based on the CI, LOH score, and molecular subtype. The average CI values in differentiated, immunoreactive, mesenchymal, and proliferative subtypes were 3.26, 3.69, 3.85, and 3.52, respectively. The mesenchymal subtype, which had a poor prognosis, showed a high CI (Figure 3A, p-value = 0.0403). The average LOH scores in each subtype were 17.5, 19.2, 17.3, and 16.2, respectively. There was no significant difference in the LOH score between molecular subtypes, but the immunoreactive subtype, which had a good prognosis, showed a tendency to have a high LOH score (Figure 3B, p-value = 0.0762).

We performed downstream effect analysis with IPA comparing high CI and low CI, high LOH and low LOH, and high LOH and low CI and other groups (Supplementary Table 5-7). In the high LOH group and high LOH/low CI group, it was predicted that the activation, migration, and adhesion of immune cells were activated.
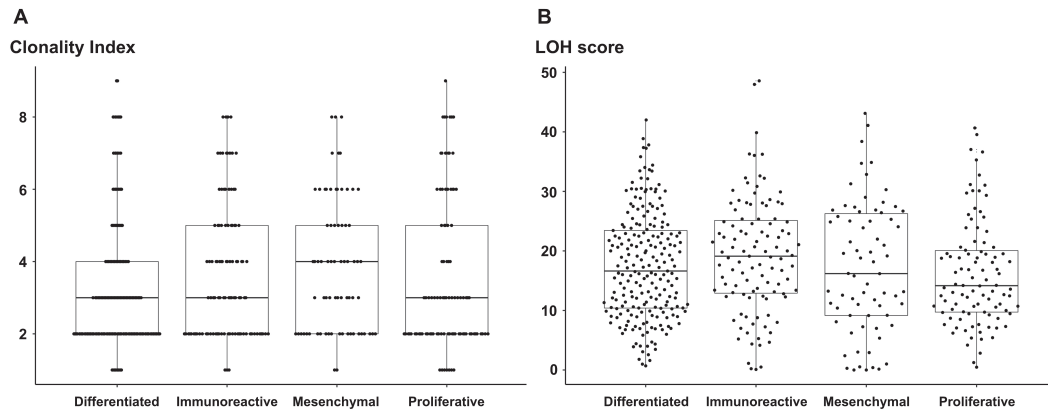


Figure 3. Pair plot and boxplot between each gene expression subtype and CI (A) and LOH scores (B). The mesenchymal subtype had a higher CI than the other subtypes, and the immunoreactive subtype had the highest LOH score.

Next, we analyzed 33 HGSOC samples from 20 cases from Kindai University Hospital. Patient characteristics are shown in Table 1. We divided 20 samples before chemotherapy into a high CI group (CI was 3 or more) and a low CI group (CI less than 3) and performed survival analysis. The low CI group showed a significantly longer PFS and OS (Supplementary Figure 1, p-value = 0.0351, 0.0385, respectively). In the Cox proportional hazards model with adjusting for clinical information, a high CI value was associated with a shorter PFS (adjusting HR = 8.078, 95% CI: = 1.117-58.413, p-value = 0.0385, Supplementary Table 4). We analyzed the change in the CI throughout the time of treatment in 4 cases that could be analyzed for tumor samples at IDS after NAC, including one case that could not be analyzed at recurrence. The CI at primary surgery was 3 in 3 cases and 2 in 1 case, and at IDS, the CI decreased to 1 in all 4 cases, and then the CI increased at the recurrent surgery in all 3 cases (Figure 4A). It was found that the tumor clone consisted of multiple clones before chemotherapy. However, after chemotherapy, the chemo-sensitive clone decreased and appeared to decrease the CI, and at the time of recurrence, the tumor clone increased again. We analyzed changes in the LOH score in the same way, and the LOH score did not change in 2 cases (30.2 → 29.2, 27.3 → 25.4) and greatly decreased in 2 cases from primary surgery to IDS and then increased at recurrent surgery (15.7 → 5.7 → 16.5, 28.6 → 12.0 → 30.0) (Figure 4B). In 6 cases whose samples were analyzed at primary and recurrent surgery, CI did not change in 4 cases and increased and decreased in one case each. The LOH score did not change in 5 cases and decreased in one case (Figure 4C, 4D). With GISTIC analysis, a pattern of the copy number variant compared between primary and recurrent surgery had few differences, but amplification of 8q24 was found at IDS with statistical significance (Figure 5A-C).
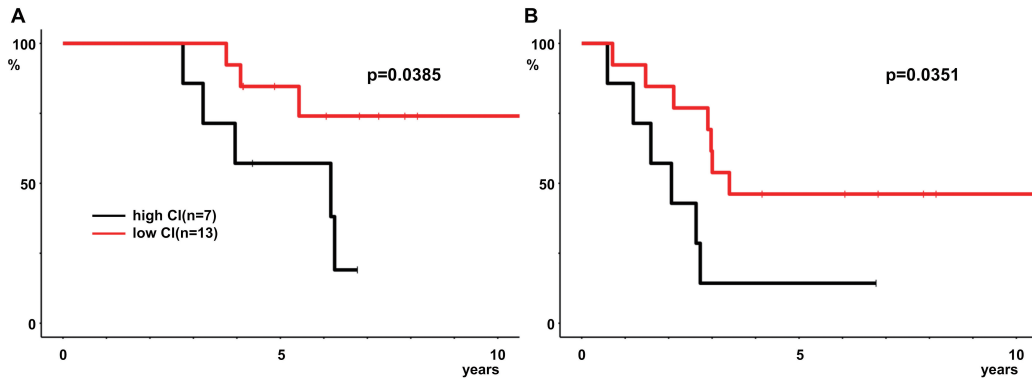
Figure S1.  Kaplan-Meier curve of our institutional data. The high CI group prolonged overall survival（A）and progression-free survival（B）with statistically significant differences.



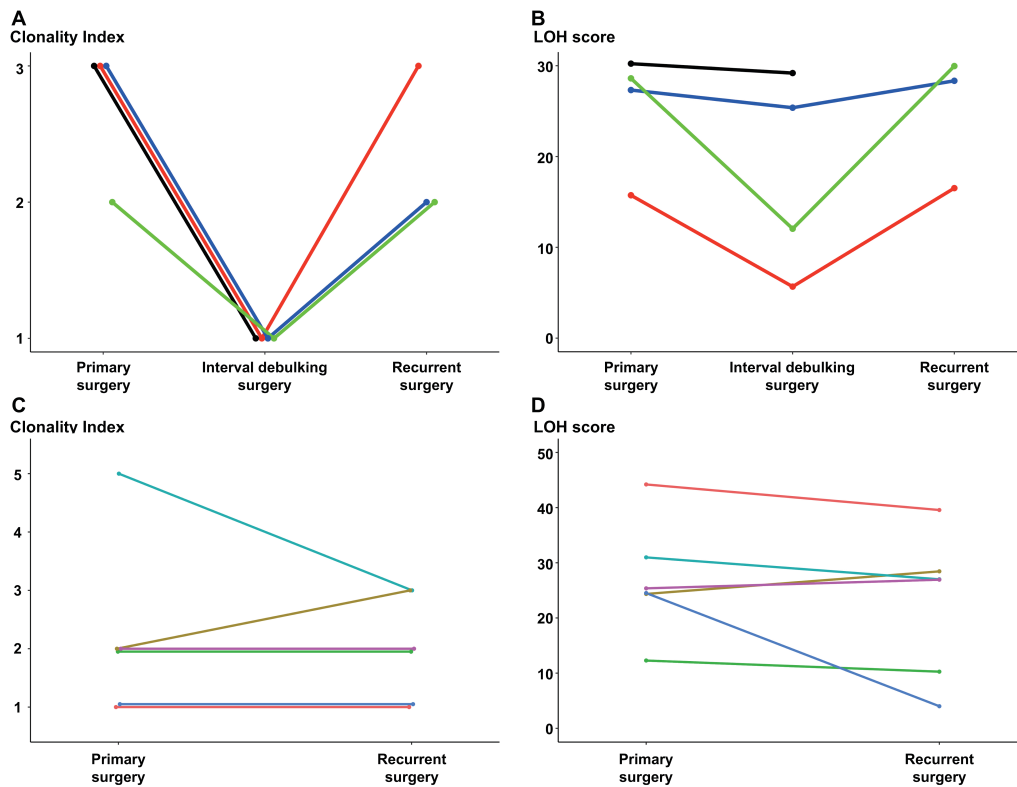Figure 4.  The transition of CI and LOH scores throughout the treatment. The CI was reduced to 1 at interval debulking surgery（IDS）and increased at recurrence in all 4 cases （A）. The LOH score was prominently reduced at IDS and increased at recurrence in 2 cases, but in the other 2 cases, the LOH score was not changed（B）. In 6 cases without IDS data, the CI did not change in 4 cases（C）, and LOH scores were almost the same（D）.
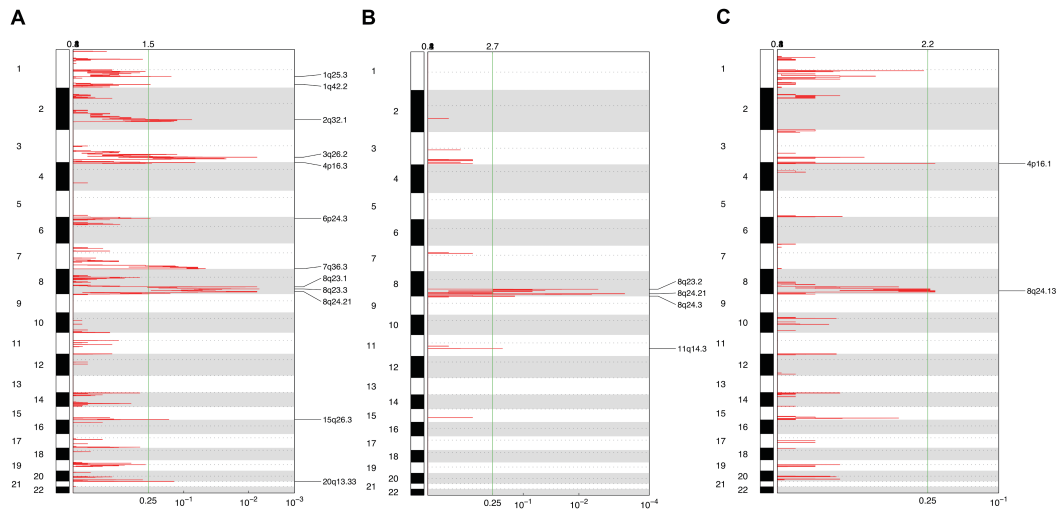
Figure 5. Amplification of HGSOC samples at primary surgery (A), interval debulking surgery (B), and recurrent surgery (C) from GISTIC analysis. Chromosome 8q24 was amplified at all three surgical times.

Table 1. Patient characteristics

|  | (N=20) |
|---|---|
| Age, median (range) | 56 (42-82) |
| FIGO stage | |
| I/II | 2 |
| III/IV | 18 |
| Primary surgery | |
| Maximum cytoreductive | 14 |
| Probe laparotomy/laparoscopy | 6 |
| Size of residual tumors | |
| No macroscopic disease | 7 |
| 1-10mm | 8 |
| >10mm | 5 |
| Sites of recurrence | |
| Peritoneal dissemination | 7 |
| Lymph node | 1 |
| Organ metastasis | 2 |

FIGO; International Federation of Gynecology and Obstetrics

Supplementary Table 4.  Cox proportional hazards model for overall survival and progression free survival in our institutional data with the LOH score adjusted for covariates

| Variable | Overall survival | | | Progression free survival | | |
|---|---|---|---|---|---|---|
| | HR | 95% CI | p-value | HR | 95% CI | p-value |
| CI (high vs low) | 2.862 | 0.517 - 15.831 | 0.228 | 8.078 | 1.117 - 58.413 | 0.0385 |
| Age at diagnosis | 1.019 | 0.934 - 1.112 | 0.672 | 1.068 | 0.968 - 1.178 | 0.189 |
| FIGO stage (I/II vs III/IV) | 0.360 | 0.016 - 7.968 | 0.518 | 0.461 | 0.020 - 10.820 | 0.631 |
| Residual disease (no macroscopic disease vs 1-10 mm vs >10 mm) | 3.336 | 0.861 - 12.921 | 0.081 | 7.483 | 1.234 - 45.363 | 0.0286 |

CI; Clonality Index, FIGO; International Federation of Gynecology and Obstetrics, HR; hazard ratio, 95% CI; 95% confidence interval

Supplementary Table 5.  Downstream effect analysis of high CI vs. low CI

| Disease and Functions Annotation | p-value | Predicted Activation State | Activation z-score |
|---|---|---|---|
| Accumulation of myeloid cells | 0.00119 | Decreased | -2.387 |
| Accumulation of leukocytes | 0.00208 | Decreased | -2.387 |
| Growth of malignant tumor | 0.00316 | Decreased | -2.336 |
| Proliferation of cancer cells | 0.00913 | Decreased | -2.201 |
| Endotoxin shock response | 0.00165 | Decreased | -2 |
| Angiogenesis of tumor | 0.00697 | Decreased | -2 |

Supplementary Table 6. Downstream effect analysis of high LOH vs. low LOH

| Disease and Functions Annotation | p-value | Predicted Activation State | Activation z-score |
|---|---|---|---|
| Cell proliferation of fibroblasts | 0.00066 | Increased | 2.945 |
| Migration of mononuclear leukocytes | 0.00138 | Increased | 2.901 |
| Accumulation of granulocytes | 0.000114 | Increased | 2.595 |
| Accumulation of phagocytes | 0.00105 | Increased | 2.589 |
| Lymphocyte migration | 0.00254 | Increased | 2.583 |
| Cell movement of muscle cells | 0.00114 | Increased | 2.566 |
| Cell movement of smooth muscle cells | 0.0024 | Increased | 2.566 |
| Migration of phagocytes | 0.000299 | Increased | 2.444 |
| Antimicrobial response | 0.00133 | Increased | 2.433 |
| Accumulation of myeloid cells | 0.000147 | Increased | 2.421 |
| Accumulation of neutrophils | 0.00011 | Increased | 2.413 |
| Degeneration of connective tissue | 0.000486 | Increased | 2.382 |
| Cell movement of T lymphocytes | 0.000285 | Increased | 2.375 |
| Activation of neutrophils | 0.000143 | Increased | 2.366 |
| Chemotaxis of granulocytes | 0.000686 | Increased | 2.327 |
| Morbidity or mortality | 0.00192 | Increased | 2.303 |
| Organismal death | 0.00281 | Increased | 2.279 |
| Accumulation of leukocytes | 0.00279 | Increased | 2.242 |
| Inflammatory response | 0.0000403 | Increased | 2.24 |
| Mobilization of leukocytes | 0.0000429 | Increased | 2.219 |
| Cell proliferation of tumor cell lines | 0.00000266 | Increased | 2.211 |
| Mobilization of cells | 0.00011 | Increased | 2.198 |
| Interaction of lymphoma cell lines | 0.00263 | Increased | 2.197 |
| Homing of T lymphocytes | 0.001 | Increased | 2.193 |
| Chemotaxis of T lymphocytes | 0.00232 | Increased | 2.193 |
| Abnormality of cartilage tissue | 0.000163 | Increased | 2.173 |
| Damage of cartilage tissue | 0.00185 | Increased | 2.173 |
| Deterioration of connective tissue | 0.000901 | Increased | 2.157 |
| Chemotaxis of neutrophils | 0.00141 | Increased | 2.107 |
| Inflammation of joint | 0.000553 | Increased | 2.104 |
| Migration of neutrophils | 0.000188 | Increased | 2.01 |
| Interphase of breast cancer cell lines | 0.00323 | Decreased | -2 |
| Growth of bacteria | 0.00108 | Decreased | -2.207 |
| Locomotion | 0.000608 | Decreased | -2.395 |

Supplementary Table 7. Downstream effect analysis of high LOH/low CI vs. others

| Disease and Functions Annotation | p-value | Predicted Activation State | Activation z-score |
|---|---|---|---|
| Inflammatory response | 8.52E-08 | Increased | 3.167 |
| Activation of neutrophils | 1.05E-08 | Increased | 3.046 |
| Activation of granulocytes | 1.22E-08 | Increased | 3.039 |
| Migration of carcinoma cell lines | 0.0000888 | Increased | 2.945 |
| Migration of neutrophils | 0.00000451 | Increased | 2.933 |
| Replication of HIV | 0.0000876 | Increased | 2.775 |
| Activation of myeloid cells | 0.000000351 | Increased | 2.654 |
| Endotoxin shock response | 0.0000273 | Increased | 2.611 |
| Angiogenesis of lesion | 0.000527 | Increased | 2.611 |
| Septic shock | 0.000812 | Increased | 2.611 |
| Activation of phagocytes | 5.51E-08 | Increased | 2.6 |
| Invasion of cells | 0.000295 | Increased | 2.6 |
| Replication of HIV-1 | 0.000286 | Increased | 2.592 |
| Differentiation of bone cells | 0.00138 | Increased | 2.426 |
| Chemotaxis of granulocytes | 0.000197 | Increased | 2.411 |
| Damage of cartilage tissue | 0.000327 | Increased | 2.382 |
| Abnormality of cartilage tissue | 0.00116 | Increased | 2.382 |
| Deterioration of connective tissue | 0.00127 | Increased | 2.377 |
| Uptake of 2-deoxyglucose | 0.000977 | Increased | 2.338 |
| Development of vasculature | 0.00000349 | Increased | 2.27 |
| Migration of granulocytes | 0.000000834 | Increased | 2.259 |
| Synthesis of nitric oxide | 0.000397 | Increased | 2.251 |
| Production of reactive oxygen species | 0.000141 | Increased | 2.25 |
| Invasion of tumor | 0.00001 | Increased | 2.249 |
| Chemotaxis of leukocytes | 0.000022 | Increased | 2.236 |
| Production of hydrogen peroxide | 0.0000837 | Increased | 2.236 |
| Migration of squamous cell carcinoma cell lines | 0.000529 | Increased | 2.236 |
| Angiogenesis | 0.00011 | Increased | 2.227 |
| Biosynthesis of hydrogen peroxide | 0.00000825 | Increased | 2.206 |
| Chemotaxis of neutrophils | 0.000379 | Increased | 2.196 |
| Cell movement of neutrophils | 0.0000928 | Increased | 2.192 |
| Proliferation of connective tissue cells | 0.0000626 | Increased | 2.186 |
| Chemotaxis of myeloid cells | 0.00000685 | Increased | 2.182 |
| Activation of antigen presenting cells | 0.00000348 | Increased | 2.177 |
| Fever | 0.000199 | Increased | 2.155 |
| Synthesis of reactive oxygen species | 0.000101 | Increased | 2.152 |
| Influx of neutrophils | 4.33E-08 | Increased | 2.138 |
| Invasion of malignant tumor | 0.000151 | Increased | 2.138 |
| Accumulation of myeloid cells | 0.00000668 | Increased | 2.109 |
| Advanced malignant tumor | 0.000137 | Increased | 2.101 |
| Synthesis of DNA | 0.000254 | Increased | 2.086 |
| Cell movement of smooth muscle cells | 0.0000235 | Increased | 2.08 |
| Cell movement of granulocytes | 0.0000156 | Increased | 2.07 |
| Cell proliferation of vascular endothelial cells | 0.00056 | Increased | 2.07 |
| Growth of connective tissue | 0.0000263 | Increased | 2.067 |
| Invasion of tumor cells | 0.00000234 | Increased | 2.051 |
| Adhesion of neutrophils | 0.0000502 | Increased | 2.049 |
| Accumulation of cells | 0.000569 | Increased | 2.048 |
| Migration of myeloid cells | 0.000000223 | Increased | 2.027 |
| Adhesion of myeloid cells | 0.0000785 | Increased | 2.019 |
| Quantity of blood cells | 0.000447 | Increased | 2.017 |
| Adhesion of phagocytes | 0.0000356 | Increased | 2.009 |
| Secretion of peptide | 0.000576 | Increased | 2 |
| Blood pressure | 0.00016 | Decreased | -2.154 |
| Failure of kidney | 0.000032 | Decreased | -2.159 |
| Contraction of muscle cells | 0.000127 | Decreased | -2.196 |
| Contraction of cells | 0.000275 | Decreased | -2.196 |

4．Discussion

We analyzed intratumor heterogeneity and HRD in HGSOC with SNP array data in the current study. Some methods to analyze the intratumor heterogeneity of malignant tumors, including ovarian cancer, have previously been reported, such as multi-region sequencing [28], deep sequencing [29], and single-cell sequencing [30]. We could analyze intratumor heterogeneity from a single sampling of a tumor by applying a previous method [17], which we reported in the current study. The merit of this method is as follows: 1) public data such as TCGA data can be applied, 2) residual tumors after chemotherapy can be analyzed as FFPE tumor specimens, and 3) easy application in clinical settings is possible because the cost of analysis is relatively low compared with that of sequencing methods.

In this study, we report two findings regarding the association between intratumor heterogeneity and the prognosis of HGSOC: PFS in the high intratumor heterogeneity group was shorter than that observed in the low intratumor heterogeneity group as an independent factor, and there was no difference in OS with respect to intratumor heterogeneity. Several studies have reported that a high degree of intratumor heterogeneity-induced resistance to treatment leads to poor prognosis in various types of cancer. Oh et al. reported that tumors with high intratumor heterogeneity of colorectal cancer had a shorter PFS [16]. Andor et al. analyzed the relationship between the number of clones and prognosis using 12 types of cancer with TCGA data and reported that a high clone number was related to poor prognosis across cancer types [31]. Morris et al. analyzed the relationship between intratumor heterogeneity and prognosis using 9 types of cancer with TCGA data and reported high intratumor heterogeneity related to shorter OS in several types of cancer [14]. In the current study, our results on PFS were similar with those of previous reports, although the finding regarding the OS was different from that observed in previous studies. However, pan-cancer studies reported by Andor et al. and Morris et al. did not include ovarian cancer, so it is difficult to directly compare their reports with our study. Moreover, Andor et al. revealed the association between OS and the number of clones across all cancer types but not individual cancer type. Additionally, the association between intratumor heterogeneity and OS was not shown in bladder cancer, lung adenocarcinoma, and lung squamous cell carcinoma in the study from Morris et al.. Thus, from these previous studies, the association between intratumor heterogeneity and OS might be different by cancer type.

HRD related to platinum sensitivity and prognosis, and the proportion of the LOH could be a biomarker of HRD [5, 32, 33]. In the current study, HGSOC cases with high LOH scores were associated with a good prognosis（Figure 2C, D）as an independent factor. Furthermore, molecular subtypes of HGSOC had different prognoses [22]; the CI was high in the mesenchymal subtype, which had a poor prognosis, and LOH tended to be high in the immunoreactive subtype, which had a good prognosis（Figure 3A, 3B）. Additionally, the pathways related to immune function were upregulated in cases with high LOH scores（Supplementary Tables 6, 7）. Consequently, intratumor heterogeneity and HRD status are partially related to differences in prognosis between molecular subtypes. In a previous study, cases in the immunoreactive subtype had a high

frequency of *BRCA1* mutations [34].

This study is the first to analyze the changes in intratumor heterogeneity before and after NAC in HGSOC. We revealed that the CI of the tumor that remained at the IDS decreased; hence, chemo-sensitive clones are thought to disappear and chemo-resistant clones are thought to remain after chemotherapy. Additionally, chromosome 8q24 amplification was detected in the tumors remaining after chemotherapy (Figure 5B). Several tumor-related genes, such as *MYC*, *PVT-1*, *PSCA*, and *TNFRSF11B*, were located on chromosome 8q24, and amplification of this region was related to poor prognosis in gastric cancer, breast cancer, colorectal cancer, and ovarian cancer in a previous study [35–38]. The LOH score prominently decreased after chemotherapy in two cases (Figure 4B). This result indicated that intratumor heterogeneity of HRD status decreased after chemotherapy, that is, non-HRD clones were selected. On the other hand, in another two cases, the LOH score did not decrease after chemotherapy. Although these cases had high LOH scores, which represent HRD, the LOH score did not change; hence, this result showed that the chemosensitivity was not always consistent with HRD status. We analyzed only 4 cases in the current study, so we need to increase the number of cases in future analyses.

Comparing the tumor at primary and recurrent surgery, there were few changes in the CI, LOH score, and patterns of copy number variation. In a previous study, mutation status or copy number variation detected by target sequencing [29] or whole exome sequencing [39] drastically changed between primary and recurrent tumors; however, one study reported that HRD status did not change [39]. The current study indicated that the tumor that remained after chemotherapy was very different from the recurrent tumor. Therefore, we should analyze the tumor that remained after chemotherapy precisely to investigate the mechanism of chemo-resistance.

There are some limitations of this study. First, the method developed to calculate the CI in this study cannot be applied to any cancer, since its use depends on amount of LOH region of the cancer genome. When a cancer has few copy number variations, the CI of the cancer can be underestimated. Second, the number of samples was limited in our own data. Both the total number of cases and NAC samples was inadequate to draw any statistically meaningful conclusions. Thus, further studies with additional samples are needed to confirm the current conclusion. Third, when we compared the CI values and the LOH scores before and after chemotherapy, the analyzed tumor tissue was obtained from different site of lesions in the same case. Because genetic or phenotypic variations between tumors would occur at different sites in the same patients, which is known to be intertumoral heterogeneity [40], comparing the CI values of different tumors before and after chemotherapy in the same patients many not adequately allow the effect of chemotherapy on the tumors to be evaluated.


5．Conclusion

We have shown that both the CI and LOH in HGSOC is related to prognosis using SNP array data from

a single sample. Furthermore, we have shown that the intratumor heterogeneity of chemo-resistant tumors that remained at IDS after NAC decreased compared with that of primary tumors. This study indicated that for analysis of tumors that remain after chemotherapy, investigation of the mechanism underlying the development of chemoresistance is important; thus, further studies are needed.

References

1. Siegel RL, Miller KD, Jemal A（2019）Cancer statistics, 2019. CA Cancer J Clin 69:7–34. doi:10.3322/caac.21551

2. Seidman JD, Horkayne-Szakaly I, Haiba M, et al（2004）The Histologic Type and Stage Distribution of Ovarian Carcinomas of Surface Epithelial Origin. Int J Gynecol Pathol 23:41–44. doi:10.1097/01.pgp.0000101080.35393.16

3. Cancer Genome Atlas Research Network（2011）Integrated genomic analyses of ovarian carcinoma. Nature 474:609–615. doi:10.1038/nature10166

4. Konstantinopoulos PA, Ceccaldi R, Shapiro GI, D???Andrea AD（2015）Homologous recombination deficiency:Exploiting the fundamental vulnerability of ovarian cancer. Cancer Discov 5:1137–1154. doi:10.1158/2159-8290.CD-15-0714

5. Zhao EY, Shen Y, Pleasance E, et al（2017）Homologous recombination deficiency and platinum-based therapy outcomes in advanced breast cancer. Clin Cancer Res 23:7521–7530. doi:10.1158/1078-0432.CCR-17-1941

6. Ledermann J, Harter P, Gourley C, et al（2014）Olaparib maintenance therapy in patients with platinum-sensitive relapsed serous ovarian cancer:a preplanned retrospective analysis of outcomes by BRCA status in a randomised phase 2 trial. Lancet Oncol 15:852–861. doi:10.1016/S1470-2045（14）70228-1

7. Gelmon KA, Tischkowitz M, Mackay H, et al（2011）Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-negative breast cancer:a phase 2, multicentre, open-label, non-randomised study. Lancet Oncol 12:852–861. doi:10.1016/S1470-2045（11）70214-5

8. Mateo J, Carreira S, Sandhu S, et al（2015）DNA-Repair Defects and Olaparib in Metastatic Prostate Cancer. N Engl J Med 373:1697–1708. doi:10.1056/NEJMoa1506859

9. Abkevich V, Timms KM, Hennessy BT, et al（2012）Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. Br J Cancer 107:1776–1782. doi:10.1038/bjc.2012.451

10. Swisher EM, Lin KK, Oza AM, et al（2017）Rucaparib in relapsed, platinum-sensitive high-grade ovarian carcinoma（ARIEL2 Part 1）:an international, multicentre, open-label, phase 2 trial. Lancet Oncol 18:75–87. doi:10.1016/S1470-2045（16）30559-9

11. Hanel W, Moll UM（2012）Links between mutant p53 and genomic instability. J Cell Biochem 113:433–439. doi:10.1002/jcb.23400

12. Bakhoum SF, Cantley LC（2018）The Multifaceted Role of Chromosomal Instability in Cancer and Its Microenvironment. Cell 174:1347–1360. doi:10.1016/j.cell.2018.08.027

13. Sansregret L, Vanhaesebroeck B, Swanton C（2018）Determinants and clinical implications of chromosomal instability in cancer. Nat Rev Clin Oncol 15:139–150. doi:10.1038/nrclinonc.2017.198

14. Morris LGT, Riaz N, Desrichard A, et al（2016）Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. Oncotarget 7:10051–63. doi:10.18632/oncotarget.7067

15. Oltmann J, Heselmeyer-Haddad K, Hernandez LS, et al（2018）Aneuploidy, TP53 mutation, and amplification of MYC correlate with increased intratumor heterogeneity and poor prognosis of breast cancer patients. Genes Chromosomes Cancer 57:165–175. doi:10.1002/gcc.22515

16. Oh BY, Shin H-T, Yun JW, et al（2019）Intratumor heterogeneity inferred from targeted deep sequencing as a prognostic indicator. Sci Rep 9:4542. doi:10.1038/s41598-019-41098-0

17. Sakai K, Ukita M, Schmidt J, et al（2017）Clonal composition of human ovarian cancer based on copy number analysis reveals a reciprocal relation with oncogenic mutation status. Cancer Lett 405:22–28. doi:10.1016/j.canlet.2017.07.013

18. Mayrhofer M, Viklund B, Isaksson A（2016）Rawcopy:Improved copy number analysis with Affymetrix arrays. Sci Rep 6:36158. doi:10.1038/srep36158

19. Wang K, Li M, Hadley D, et al（2007）PennCNV:An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res 17:1665–1674. doi:10.1101/gr.6861907

20. Van Loo P, Nordgard SH, Lingjaerde OC, et al（2010）Allele-specific copy number analysis of tumors. Proc Natl Acad Sci 107:16910–16915. doi:10.1073/pnas.1009843107

21. Coleman RL, Oza AM, Lorusso D, et al（2017）Rucaparib maintenance treatment for recurrent ovarian carcinoma after response to platinum therapy（ARIEL3）:a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet 390:1949–1961. doi:10.1016/S0140-6736（17）32440-6

22. Verhaak RGW, Tamayo P, Yang J-Y, et al（2012）Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. J Clin Invest 123:517–25. doi:10.1172/JCI65833

23. Marquard AM, Eklund AC, Joshi T, et al（2015）Pan-cancer analysis of genomic scar signatures associated with homologous recombination deficiency suggests novel indications for existing cancer drugs. Biomark Res 3:9. doi:10.1186/s40364-015-0033-4

24. Mermel CH, Schumacher SE, Hill B, et al（2011）GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol 12:R41. doi:10.1186/gb-2011-12-4-r41

25. Stickgold R, Hobson JA, Fosse R, Fosse M（2001）Sleep, learning, and dreams:off-line memory reprocessing. Science 294:1052–7. doi:10.1126/science.1063530

26. Rand DG, Greene JD, Nowak MA（2012）Spontaneous giving and calculated greed. Nature 489:427–430. doi:10.1038/nature11467

27. Jamal-Hanjani M, Wilson GA, McGranahan N, et al（2017）Tracking the Evolution of Non–Small-Cell Lung Cancer. N Engl J Med NEJMoa1616288. doi:10.1056/NEJMoa1616288

28. Bashashati A, Ha G, Tone A, et al（2013）Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling. J Pathol 231:21–34. doi:10.1002/path.4230

29. Beltrame L, Di Marino M, Fruscio R, et al（2015）Profiling cancer gene mutations in longitudinal epithelial ovarian cancer biopsies by targeted next-generation sequencing:a retrospective study. Ann Oncol 26:1363–1371. doi:10.1093/annonc/mdv164

30. McPherson A, Roth A, Laks E, et al（2016）Divergent modes of clonal spread and intraperitoneal mixing in high-grade serous ovarian cancer. Nat Genet 48:758–67. doi:10.1038/ng.3573

31. Andor N, Graham TA, Jansen M, et al（2016）Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. Nat Med 22:105–113. doi:10.1038/nm.3984

32. Tumiati M, Hietanen S, Hynninen J, et al（2018）A Functional Homologous Recombination Assay Predicts Primary Chemotherapy Response and Long-Term Survival in Ovarian Cancer Patients. Clin Cancer Res 24:4482–4493. doi:10.1158/1078-0432.CCR-17-3770

33. Norquist BM, Brady MF, Harrell MI, et al（2018）Mutations in homologous recombination genes and outcomes in ovarian carcinoma patients in GOG 218:An NRG oncology/Gynecologic oncology group study. Clin Cancer Res 24:777–783. doi:10.1158/1078-0432.CCR-17-1327

34. George J, Alsop K, Etemadmoghadam D, et al（2013）Nonequivalent Gene Expression and Copy Number Alterations in High-Grade Serous Ovarian Cancers with BRCA1 and BRCA2 Mutations. Clin Cancer Res 19:3474–3484. doi:10.1158/1078-0432.CCR-13-0066

35. Meng F, Liu B, Xie G, et al（2017）Amplification and overexpression of PSCA at 8q24 in invasive micropapillary carcinoma of breast. Breast Cancer Res Treat 166:383–392. doi:10.1007/s10549-017-4407-1

36. Wang X, Liu Y, Shao D, et al（2016）Recurrent amplification of MYC and TNFRSF11B in 8q24 is associated with poor survival in patients with gastric cancer. Gastric Cancer 19:116–127. doi:10.1007/s10120-015-0467-2

37. Takahashi Y, Sawada G, Kurashige J, et al（2014）Amplification of PVT-1 is involved in poor prognosis via apoptosis inhibition in colorectal cancers. Br J Cancer 110:164–171. doi:10.1038/bjc.2013.698

38. Guan Y, Kuo W-L, Stilwell JL, et al（2007）Amplification of PVT1 Contributes to the Pathophysiology of Ovarian and Breast Cancer. Clin Cancer Res 13:5745–5755. doi:10.1158/1078-0432.CCR-06-2882

39. Lambrechts S, Smeets D, Moisse M, et al（2016）Genetic heterogeneity after first-line chemotherapy in high-grade serous ovarian cancer. Eur J Cancer 53:51–64. doi:10.1016/j.ejca.2015.11.001

40. Burrell RA, McGranahan N, Bartek J, Swanton C（2013）The causes and consequences of genetic heterogeneity in cancer evolution. Nature 501:338–345. doi:10.1038/nature12625