Evaluation of Low Coverage Whole Genome Sequencing as a New Method for Detecting Malignant Ovarian Mass

Chen Ming  
Sun Yat-sen University First Affiliated Hospital

Zhong Pengqiang  
Sun Yat-sen University First Affiliated Hospital

Hong Mengzhi  
Sun Yat-sen University First Affiliated Hospital

Tan Jinfeng  
Sun Yat-sen University First Affiliated Hospital

Yu Xuegao  
Sun Yat-sen University First Affiliated Hospital

Huang Hao  
Sun Yat-sen University First Affiliated Hospital

Ouyang Juan  
Sun Yat-sen University First Affiliated Hospital

Lin Xiaoping  
Sun Yat-sen University Cancer Center

Peisong Chen (chps@mail3.sysu.edu.cn)  
Sun Yat-sen University First Affiliated Hospital  https://orcid.org/0000-0003-0640-3663

Research Article

Keywords: Ovarian cancers, whole genome sequencing, ROMA, RM, ROC

DOI: https://doi.org/10.21203/rs.3.rs-535366/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

To evaluate whether low coverage whole genome sequencing is suitable for the detection of malignant pelvic mass and compare its diagnostic value with traditional tumor markers. We enrolled 63 patients with a pelvic mass suspicious for ovarian malignancy. Each patient underwent low coverage whole genome sequencing (LCWGS) and traditional tumor markers test. The pelvic masses were finally confirmed via pathological examination. The copy number variants (CNVs) of whole genome were detected and the Stouffers Z-scores for each CNV was extracted. The risk of malignancy (RM) of each suspicious sample was calculated based on the CNV counts and Z-scores, which was subsequently compared with ovarian cancer markers CA125 and HE4, and the risk of ovarian malignancy algorithm (ROMA). Receiver Operating Characteristic Curve (ROC) were used to access the diagnostic value of variables. As confirmed by pathological diagnosis, 44 (70%) patients with malignancy and 19 patients with benign mass were identified. Our results showed that CA125 and HE4, the CNV, the mean of Z-scores (Zmean), the max of Z-scores (Zmax), the RM and the ROMA were significantly different between patients with malignant and benign masses. The area under curve (AUC) of CA125, HE4, CNV, Zmax, and Zmean was 0.775, 0.866, 0.786, 0.685 and 0.725 respectively. ROMA and RM showed similar AUC (0.876 and 0.837), but differed in sensitivity and specificity. After all, we develop a LCWGS based method for the identification of pelvic mass of suspicious ovarian cancer. LCWGS shows accurate result and could be complementary with the existing diagnostic methods.

Introduction

According to the latest 2018 global cancer data report, the incidence of ovarian tumors in female reproductive system accounted for 3.4% of all female tumors in China, and the number of women who died of malignant ovarian tumors accounted for 4.4% of all female patients who died of tumors[1]. Ovarian cancer has become the second highest incidence and mortality of female reproductive system tumor following cervical cancer[1, 2]. Because of the small size of the ovary and its position in the pelvic cavity, ovarian tumor itself lacks typical symptoms in early stage[3]. Patients often find that they have ovarian tumor after the pelvic cavity has a huge mass or bleeding in the vagina [4, 5]. At this time, the tumor has developed to the late stage and most of them spread to other pelvic organs, and has missed the best time for treatment[6]. Therefore, the early detection of ovarian tumors is critical for clinical management and prognosis of patients. Multiple efforts have been made to evaluate traditional markers including serum concentration of CA125 and HE4 in the screening of ovarian cancers[7]. However, these markers did not meet the standards required to advocate population-based screening regarding with the diagnostic sensitivity and or specificity[8, 9]. In order to improve the accuracy of diagnosis for ovarian cancer, additional cancer-specific diagnostic methods may be required.

In recent years, the rapid development in the field of next generation sequencing (NGS) and its application in low coverage whole genome sequencing (LCWGS) makes the detection of tumor-specific copy number alterations (CNA) in cell-free DNA feasible[10, 11]. Evidence has showed that tumor-derived chromosome abnormalities would be detectable in the plasma of patients prior to surgery[10, 12].
Previous studies have reported that occult pelvic cancers can be detected by LCWGS testing but it might cause false positive results [13]. However, the diagnostic accuracy of LCWGS platform and analytic pipeline for ovarian cancer remains unknown. The aim of this study is to investigate whether a clinical LCWGS platform could detect ovarian cancers in patients with pelvic masses based on the abnormal plasma DNA copy number variants (CNVs), and to compare the diagnostic accuracy with traditional screening markers including CA125 and HE4, and the score of risk of ovarian malignancy algorithm (ROMA) [14].

Methods

1. Subjects and samples

Sixty-three patients with a pelvic mass suspicious for ovarian malignancy, who were referred to the gynecology department of the First Affiliated Hospital of Sun Yat-sen university from January 2018 to July 2019 were recruited in this study. In addition, a cohort of 39 healthy female individuals were also recruited. Blood samples were collected using EDTA anticoagulated tube and sent for laboratory within 2 hours. The study approval was obtained from the ethical committee of the First Affiliated Hospital of Sun Yat-sen university (S/55904). All participants submitted their written informed consents.

2. Sample processing and LCWGS

The blood samples were firstly centrifuged at 1600 g for ten minutes at 4°C, and then the supernatant was centrifuged at 16000 g again for ten minutes at 4°C. The plasma was stored − 80°C until analysis. The isolation, purification, library construction and sequencing of cell free DNA from the blood were performed by using a Fetal Aneuploidies Trisomy Detection Kit (Daan Gene Corp, China) on Ion Proton next-generation sequencer (Life Technologies) which was certified by the China Food and Drug Administration. All procedures were performed according to the manufacture’s protocol.

3. Bio-informatics analysis

Raw sequencing reads were mapped to the human reference genome Hg19 using BWA (v0.7.1). Duplicate and low-quality reads were removed by Picard Tools (v1.11) and Samtools (v0.1.18) respectively. TorrentSuit software (v3.6) and a NIPT-plus plugin (provided by the Daan Gene Corp) was used to calculate the Stouffers Z-scores for whole chromosomes and CNV ≥ 5.0MB. |Z-scores|≥3 were marked as high risk. Both CNV counts and |Z-scores| (≥3) were extracted from each sample for further analysis.

4. Analysis of malignant risk

For further analysis of the risk of malignancy, data from 39 healthy females was used to form a baseline. Firstly, we calculated the mean of CNV counts and |Z-scores| (≥3), then the risk of malignancy(RM) of each suspicious sample was calculated as (CNV counts\_suspicious - CNV counts\_mean of healthy) X (|Z-scores|\_suspicious - |Z-scores|\_mean of healthy).

5. Tumor marker detection and ROMA scores
HE4 and CA125 were tested in stored plasma using the ARCHITECT HE4 and CA125 assays (Abbott Diagnostics, Abbott Park, IL, USA) according to the manufacturer's instructions.

6. Pathology diagnosis of pelvic mass

All diagnoses of patients were confirmed via pathological examination by pathologists who were blind to the results of clinical laboratory testing. Tumor staging was performed according to the International Federation of Gynecology and Obstetrics (FIGO) criteria (2010).

7. Statistical analysis

Statistical analysis was carried out by an online statistics tool (http://dxonline.deepwise.com/) and R software (Version 4.0.1) with pROC and Rattle package (5–7). Receiver operating characteristics (ROC) curve was used to evaluate the diagnostic value. A two-tailed P value of less than 0.05 was considered statistically significant.

Results

1. Clinical and pathology data of subjects

This study included 63 patients with a pelvic mass suspicious of ovarian malignancy, who were finally identified as 34 (54%) high grade malignancy, 10 (16%) low grade malignancy and 19 (30%) benign mass by pathological diagnosis. The median age of premenopausal patients were 35 years (range, 16–53 years), and the median age of postmenopausal patients were 62 years (range, 46–83 years). The median age of patients with malignancies was 51 year (range: 21–70) and that of benign diseases was 30 year (range: 18–52). There was a significant difference in age distribution between these 2 groups of patients (P < 0.01). The FIGO stage of ovarian cancers patients included 13 (30%) I stage, 6 (14%) II stage, 18 (41%) III stage and 7 (16%) IV stage. The clinical and pathological data of subjects were listed in Table 1.

2. LCWGS on CNVs

LCWGS used a whole genome low coverage strategy to analyze the CNVs. For each sample, more than 5M (5.9 ± 0.68 for all samples) reads was obtained. The coverage of each sample is about 0.35×. A representative LCWGS figure for ovarian cancer and benign disease was shown in Fig. 1. The results from a patient with FIGO Stage III serous cystadenocarcinoma showed multiple regions of CNV (Fig. 1A). And the results from a patient with teratoma showed that no CNV (Fig. 1B). In this study, only 7 patients with malignancy showed trisomy or monosomy as indicated by LCWGS. To further investigate the diagnostic performance of LCWGS, CNV counts, max of Z scores (Zmax) of all CNVs, mean of Z scores (Zmean) and RM was calculated from each sample. Significant difference of LCWGS based index was found between patients with malignant and benign tumors. Patients with malignancy showed higher level in LCWGS based index than patients with benign disease. In addition, these indexes were closely related to different FIGO stage (Figure 2 and Table 2).

3. Traditional tumor markers
The serum concentration of CA125 was $416.457 \pm 747.887$ (Mean ± SD), HE4 was $219.192 \pm 457.614$ and ROMA was $0.534 \pm 0.422$ in all subjects. There were significant differences between the concentration of CA125($560.282 \pm 854.994$ VS $83.387 \pm 112.353$) and HE4($286.382 \pm 534.32$ VS $63.595 \pm 51.849$) in patients with malignant and benign diseases. Besides, we compared the serum concentration of CA125 and HE4 in different FIGO stage, and the results indicated that the level of both markers were correlated with different FIGO stage(Figure 2 and Table 2).

4. Correlation between traditional tumor markers and LCWGS index

Spearman correlation was used to investigate the relationship between tumor markers and LCWGS index. As shown in Fig. 3 and table 3, all indexes were statistically correlated ($P < 0.01$). However, the correlation between traditional tumor markers and LCWGS index was weak ($r$ value range from 0.38 to 0.77).

5. Comparison of the diagnostic value of LCWGS and traditional tumor markers

Firstly, we evaluated the diagnostic value of single index in the research subjects. The AUC of CA125 and HE4 was 0.775 and 0.866 respectively. HE4 showed better diagnostic accuracy than other markers. Then the integrated indexes were evaluated. The AUC of ROMA and RM was 0.876 and 0.837, respectively. Both ROMA and RM showed higher diagnostic accuracy than single index. However, no significant difference was found between ROMA and RM (Delong test: $P = 0.476$), which indicated that ROMA and RM had similar diagnostic value between ovarian cancers and benign diseases. With the cutoff of 0.085, the sensitivity and specificity of ROMA was 0.684 and 0.909 respectively. With the cutoff of 1.25, the sensitivity and specificity of ROMA was 0.895 and 0.773 respectively(Figure 4 and Table 4).

Discussion

As the second highest incidence and mortality of female reproductive system tumor following cervical cancer, ovarian cancer has the early clinical presentation that are difficult to be differentiated from digestive tract diseases, such as bloating or abdominal pain[15, 16]. When ovarian cancer develops and spreads to the abdominal cavity, abdominal mass may appear[17]. Therefore, distinguishing between benign and malignant abdominal masses is very important for the early diagnosis of ovarian cancer.

Oncogenesis involves many types of genomic variation, such as point mutation, copy number variation and gene fusion [18]. Tumors are different from genetic diseases, and their genomic variation is frequently acquired[19]. The development of ovarian cancer is a complex process involving the changes of DNA, RNA, and proteins[20, 21]. The abnormal DNA of cancers could release from cancer tissues and be detected in blood samples in the form of cell free DNA[22]. Therefore, the detection of CNVs would be a promising method for the identification of malignant abdominal masses.
In this study, we evaluated whether CNVs detected by LCWGS platform could accurately predict the existence of malignancy. In our study cohort, the number of patients with malignant (43 cases) was higher than the patients with benign disease (19 cases). Our results showed that, chromosome variation could be detected in cell free DNA in patients with malignancy. However, only a few cases with malignant mass showed trisomy or monosomy. Despite that chromosome instability was common in tumor cells, owing to the low concentration of tumor derived cell free DNA, detection of trisomy or monosomy might lack sensitivity for clinical diagnosis[23]. We set our detection target to CNVs at the resolution of 5MB. With this strategy, more chromosome instabilities could found in the subjects, however, the specificity might reduce. To solve this problem, we extracted more indexes from the LCWGS results and a healthy cohort was used to calibrate our results. Our results indicate that LCWGS based indexes were significantly different between patients with malignant and benign diseases and closely related to FIGO Stage, which would be valuable in the diagnosis of malignant mass. The diagnostic value of LCWGS based indexes were evaluated by ROC curve. Despite that CNV, Zmax and Zmean were useful for the diagnosis of malignant mass, however, the AUCs were less than 0.80. An integrated RM index which is calculated by CNV and Zmean and calibrated by a healthy cohort, showed better diagnostic performance with a AUC of 0.837. With the cut-off value of 1.25, RM is highly sensitive in the detection of malignant mass with all stage.

Both CA125 and HE4 were the most widely used markers in ovarian cancer diagnosis [24]. In our study, CA125 and HE4 showed significant difference between the malignant mass and benign disease, which is consistent with previous reports. In 2009, Moore proposed ROMA as a new algorithm. He correlated HE4 and CA125 levels with menopausal status, which was defined as 6 months of menopause without menstruation or clinical symptoms. The ROMA corresponds to the predicted probability [PP], expressed as a percentage[14]. The sensitivity of ROMA for ovarian cancer diagnosis varies from 75–97%, however, the detection of early stage malignancy was still a problem [25–27]. We compared the diagnostic value between RM and ROMA, despite that ROMA showed higher AUC than RM, however, the difference was not statistically significant. The sensitivity of RM (0.895) is superior to that of ROMA (0.684), while the specificity of RM (0.773) is inferior to that of ROMA (0.909). The CA125 and HE4 were correlated with LCWGS based index. However, the correlation was weak. Therefore, RM and ROMA could only be used as complementary in the diagnosis of pelvic malignant mass.

Low specificity of RM may originate from the bio-informatics pipeline in LCWGS. All CNVs in whole genome were used for further analysis. Ovarian cancers showed specific gain or loss of chromosomes in tissues as demonstrated by other studies, however,there was no widely accepted specific CNVs in cell free DNAs [28]. Further studies should be developed and focus on ovarian cancer specific CNVs to improve the diagnostic specificity. In addition, the increase of sequencing depth would be helpful in increasing the diagnostic value. Further studies could try to ascertain the sequencing depth regarding with the cost and effect.

A limitation of this study was that the number of patients was small. A larger sample size is needed to validate our findings, and to conduct further studies on different FIGO stages of ovarian cancer or in
patients with pre- and post-menopause.

In conclusion, our study provided a new methodology with high accuracy for the diagnosis of ovarian cancers, which could be a supplement to the existing diagnostic methods.

**Declarations**

**Ethics approval and consent to participate**

The Research Ethics Committee of the Sun Yat-sen University approved the study.

**Consent for publication**

All of the authors agreed for publication.

** Availability of data and material**

The data and material in our studies were availability.

**Authors’ contributions**

Chen Peisong and Chen Ming contributed to the study design and bioinformatics analysis. Zhong Pengqiang and Hong Mengzhi drafted the manuscript. Tan Jinfeng and Ouyang Juan for the discussion. Yu Xuegao and Huang Hao performed laboratory test. Lin Xiaoping contributed to the statistic analysis. All authors read and approved the final manuscript.

**Competing interests**

All authors declare no conflict of interest.

**Acknowledgments**

We would like to thank Dr Liang from Daan Biotechnology for his assistance in the NGS data analysis.

**Funding**

This study was funded by the National Natural Science Foundation of China 81602261, CSCO Cancer Research Foundation Y-sy2018-120, and Guangdong Basic and Applied Basic Research Foundation 2019A1515011663.

**References**

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018, 68(6):394-424.
2. Xu C, Wang Y, Yang H, Hou J, Sun L, Zhang X, Cao X, Hou Y, Wang L, Cai Q et al: Association Between Cancer Incidence and Mortality in Web-Based Data in China: Infodemiology Study. J Med Internet Res 2019, 21(1):e10677.

3. Stewart C, Ralyea C, Lockwood S: Ovarian Cancer: An Integrated Review. Semin Oncol Nurs 2019, 35(2):151-156.

4. Straubhar AM, Wolf JL, Zhou MQC, Iasonos A, Cham S, Wright JD, Long Roche K, Chi DS, Zivanovic O: Advanced ovarian cancer and cytoreductive surgery: Independent validation of a risk-calculator for perioperative adverse events. Gynecol Oncol 2021, 160(2):438-444.

5. Cham S, Chen L, St Clair CM, Hou JY, Tergas AI, Melamed A, Ananth CV, Neugut AI, Hershman DL, Wright JD: Development and validation of a risk-calculator for adverse perioperative outcomes for women with ovarian cancer. American journal of obstetrics and gynecology 2019, 220(6):571.e571-571.e578.

6. Lin JJ, Egorova N, Franco R, Prasad-Hayes M, Bickell NA: Ovarian Cancer Treatment and Survival Trends Among Women Older Than 65 Years of Age in the United States, 1995-2008. Obstetrics and gynecology 2016, 127(1):81-89.

7. Williams RM, Lee C, Galassi TV, Harvey JD, Leicher R, Sirenko M, Dorso MA, Shah J, Olvera N, Dao F et al: Noninvasive ovarian cancer markers detection via an optical nanosensor implant. Science advances 2018, 4(4):eaao1090.

8. Gentry-Maharaj A, Burnell M, Dilley J, Ryan A, Karpinskyj C, Gunu R, Mallett S, Deeks J, Campbell S, Jacobs I et al: Serum HE4 and diagnosis of ovarian cancer in postmenopausal women with adnexal masses. American journal of obstetrics and gynecology 2020, 222(1):56.e51-56.e17.

9. Janas L, Głowacka E, Wilczyński JR, Malinowski A, Nowak M: [Evaluation of applicability of HE4 and ROMA in the preoperative diagnosis of adnexal masses]. Ginekologia polska 2015, 86(3):193-197.

10. Cohen PA, Flowers N, Tong S, Hannan N, Pertile MD, Hui L: Abnormal plasma DNA profiles in early ovarian cancer using a non-invasive prenatal testing platform: implications for cancer screening. BMC medicine 2016, 14(1):126.

11. Kulasingam V, Diamandis EP: Genomic profiling for copy number changes in plasma of ovarian cancer patients - a new era for cancer diagnostics? BMC medicine 2016, 14(1):186.

12. Nakabayashi M, Kawashima A, Yasuhara R, Hayakawa Y, Miyamoto S, lizuka C, Sekizawa A: Massively parallel sequencing of cell-free DNA in plasma for detecting gynaecological tumour-associated copy number alteration. Scientific reports 2018, 8(1):11205.

13. Bianchi DW, Chudova D, Sehnert AJ, Bhatt S, Murray K, Prosen TL, Garber JE, Wilkins-Haug L, Vora NL, Warsof S et al: Noninvasive Prenatal Testing and Incidental Detection of Occult Maternal Malignancies. JAMA 2015, 314(2):162-169.

14. Moore RG, McMeekin DS, Brown AK, DiSilvestro P, Miller MC, Allard WJ, Gajewski W, Kurman R, Bast RC, Jr., Skates SJ: A novel multiple marker bioassay utilizing HE4 and CA125 for the prediction of ovarian cancer in patients with a pelvic mass. Gynecol Oncol 2009, 112(1):40-46.
15. Orr B, Edwards RP: Diagnosis and Treatment of Ovarian Cancer. Hematology/oncology clinics of North America 2018, 32(6):943-964.

16. Ebell MH, Culp MB, Radke TJ: A Systematic Review of Symptoms for the Diagnosis of Ovarian Cancer. American journal of preventive medicine 2016, 50(3):384-394.

17. Pradeep S, Kim SW, Wu SY, Nishimura M, Chaluvally-Raghavan P, Miyake T, Pecot CV, Kim SJ, Choi HJ, Bischoff FZ et al: Hematogenous metastasis of ovarian cancer: rethinking mode of spread. Cancer cell 2014, 26(1):77-91.

18. Kar SP, Berchuck A, Gayther SA, Goode EL, Moysich KB, Pearce CL, Ramus SJ, Schildkraut JM, Sellers TA, Pharoah PDP: Common Genetic Variation and Susceptibility to Ovarian Cancer: Current Insights and Future Directions. Cancer epidemiology, markers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 2018, 27(4):395-404.

19. Previs RA, Sood AK, Mills GB, Westin SN: The rise of genomic profiling in ovarian cancer. Expert review of molecular diagnostics 2016, 16(12):1337-1351.

20. Asante DB, Calapre L, Ziman M, Meniawy TM, Gray ES: Liquid biopsy in ovarian cancer using circulating tumor DNA and cells: Ready for prime time? Cancer letters 2020, 468:59-71.

21. Kroeger PT, Jr., Drapkin R: Pathogenesis and heterogeneity of ovarian cancer. Current opinion in obstetrics & gynecology 2017, 29(1):26-34.

22. Kurman RJ, Shih le M: Molecular pathogenesis and extraovarian origin of epithelial ovarian cancer—shifting the paradigm. Human pathology 2011, 42(7):918-931.

23. Biesecker LG, Spinner NB: A genomic view of mosaicism and human disease. Nature reviews Genetics 2013, 14(5):307-320.

24. Henderson JT, Webber EM, Sawaya GF: Screening for Ovarian Cancer: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. Jama 2018, 319(6):595-606.

25. Chen X, Zhou H, Chen R, He J, Wang Y, Huang L, Sun L, Duan C, Luo X, Yan H: Development of a multimarker assay for differential diagnosis of benign and malignant pelvic masses. Clin Chim Acta 2015, 440:57-63.

26. Dochez V, Caillon H, Vaucel E, Dimet J, Winer N, Ducarme G: markers and algorithms for diagnosis of ovarian cancer: CA125, HE4, RMI and ROMA, a review. J Ovarian Res 2019, 12(1):28.

27. Al Musalhi K, Al Kindi M, Al Aisary F, Ramadhan F, Al Rawahi T, Al Hatali K, Mula-Abed WA: Evaluation of HE4, CA-125, Risk of Ovarian Malignancy Algorithm (ROMA) and Risk of Malignancy Index (RMI) in the Preoperative Assessment of Patients with Adnexal Mass. Oman medical journal 2016, 31(5):336-344.

28. Concolino P, Capoluongo E: Detection of BRCA1/2 large genomic rearrangements in breast and ovarian cancer patients: an overview of the current methods. Expert review of molecular diagnostics 2019, 19(9):795-802.

Tables
Due to technical limitations, table 1, 2, 3 and 4 is only available as a download in the Supplemental Files section.

Figures

A

P5 (4,887,963 reads)

B

22 (6,044,700 reads)

Figure 1
a representative figure for ovarian cancer and benign disease. (A) NIPT results from a patient with FIGO Stage III Serous cystadenocarcinoma which showed multiple regions of copy number variants. The blue regions indicate duplications while the red regions indicate deletions. (B) NIPT results from a patient with Teratoma. No copy number variants was found in this patient.

Figure 2

CNV, RM, ZMAX, Zmean, CA125 and HE4 in different FIGO stage.

Figure 3

Spearman correlation analysis of variables.

Figure 4
ROC of ROMA and RM. The ROC included Age, CA125, HE4, CNV, Zmax, Zmean, ROMA and RM. There were no significant differences in AUC of ROMA and RM (Delong test: p = 0.476).

**Supplementary Files**

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

- Table1.pdf
- Table2.pdf
- Table3.pdf
- Table4.pdf