



University of Southern Denmark

## A Sensitive and Transparent Method for Tumor-Informed Detection of Circulating Tumor DNA in Ovarian Cancer Using Whole-Genome Sequencing

Thusgaard, Christine Fribert; Sadegh, Sepideh; Jochumsen, Kirsten Marie; Kruse, Torben Arvid; Thomassen, Mads

*Published in:*  
International Journal of Molecular Sciences

*DOI:*  
10.3390/ijms252413349

*Publication date:*  
2024

*Document version:*  
Final published version

*Document license:*  
CC BY

*Citation for pulished version (APA):*  
Thusgaard, C. F., Sadegh, S., Jochumsen, K. M., Kruse, T. A., & Thomassen, M. (2024). A Sensitive and Transparent Method for Tumor-Informed Detection of Circulating Tumor DNA in Ovarian Cancer Using Whole-Genome Sequencing. *International Journal of Molecular Sciences*, 25(24), Article 13349.  
<https://doi.org/10.3390/ijms252413349>

Go to publication entry in University of Southern Denmark's Research Portal

### Terms of use

This work is brought to you by the University of Southern Denmark.  
Unless otherwise specified it has been shared according to the terms for self-archiving.  
If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.  
Please direct all enquiries to [puresupport@bib.sdu.dk](mailto:puresupport@bib.sdu.dk)



Article

# A Sensitive and Transparent Method for Tumor-Informed Detection of Circulating Tumor DNA in Ovarian Cancer Using Whole-Genome Sequencing

Christine Fribert Thusgaard<sup>1,2,\*</sup> , Sepideh Sadegh<sup>3,4</sup> , Kirsten Marie Jochumsen<sup>1,2</sup> , Torben Arvid Kruse<sup>3,4</sup> and Mads Thomassen<sup>3,4</sup>

<sup>1</sup> Department of Gynecology and Obstetrics, Odense University Hospital, 5000 Odense, Denmark; kirsten.jochumsen@rsyd.dk

<sup>2</sup> Research Unit of Gynecology and Obstetrics, Department of Clinical Research, University of Southern Denmark, Odense University Hospital, 5000 Odense, Denmark

<sup>3</sup> Department of Clinical Genetics, Odense University Hospital, 5000 Odense, Denmark; sepideh.sadegh@rsyd.dk (S.S.); torben.kruse@rsyd.dk (T.A.K.); mads.thomassen@rsyd.dk (M.T.)

<sup>4</sup> Clinical Genome Center, Department of Clinical Research, University of Southern Denmark, 5000 Odense, Denmark

\* Correspondence: christine.fribert.thusgaard@rsyd.dk

**Abstract:** Circulating tumor DNA (ctDNA) is a biomarker that could potentially improve the survival rate of ovarian cancer (OC), e.g., by monitoring treatment response and early relapse detection. However, an optimal method for ctDNA analysis in OC remains to be established. We developed a method for tumor-informed single-nucleotide variant detection of ctDNA in OC using whole-genome sequencing. Tumor and plasma samples obtained at the time of diagnosis from 10 patients with OC were included. The tested method involved applying basic filters with different cut-offs of read depth, allelic depth, and variant allele frequency of tumor and normal DNA. In addition, we applied a new filtering approach using plasma samples from the other included OC patients (the plasma pool) for specific removal of artefacts. The basic filters with varying cut-offs showed minor improvement in signal-to-noise ratio (S2N). However, the addition of the plasma pool filter resulted in a considerable ctDNA signal improvement, indicated by both S2N and z-score. This study demonstrates a promising method for ctDNA detection in OC patients using a tumor-informed approach for whole-genome sequencing. Despite the limited number of patients involved, the results suggest a significant potential of the method for ctDNA signal detection in patients with OC.

**Keywords:** circulating tumor DNA; epithelial ovarian carcinoma; carcinoma; whole-genome sequencing; liquid biopsy; single-nucleotide variant; single-nucleotide variant detection



**Citation:** Thusgaard, C.F.; Sadegh, S.; Jochumsen, K.M.; Kruse, T.A.; Thomassen, M. A Sensitive and Transparent Method for Tumor-Informed Detection of Circulating Tumor DNA in Ovarian Cancer Using Whole-Genome Sequencing. *Int. J. Mol. Sci.* **2024**, *25*, 13349. <https://doi.org/10.3390/ijms252413349>

Academic Editor: Giovanni Tossetta

Received: 14 October 2024

Revised: 9 December 2024

Accepted: 10 December 2024

Published: 12 December 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Circulating tumor DNA (ctDNA) has been identified as a promising biomarker in certain types of cancer, including ovarian cancer (OC). CtDNA is the fraction of cell-free DNA (cfDNA) in plasma released from a malignant tumor into the circulation [1,2]. A low fraction of ctDNA in cfDNA is a limitation in ctDNA detection. Previously, the consensus has been to increase the sequencing depth of a limited number of targets to improve the detection of ctDNA. However, ctDNA is not always detected by ultra-deep targeted sequencing, especially in cases with a low tumor burden [3]. Recent studies have shown that the breadth of sequencing complements the sequencing depth, and the widest sequencing breadth is achieved through whole-genome sequencing (WGS) [3–7]. By using WGS to detect single-nucleotide variants (SNVs) in ctDNA, a study detected ctDNA with tumor fractions as low as  $10^{-5}$ , compared to  $10^{-3}$  with the use of ultra-deep targeted sequencing [3]. WGS has also been used to detect ctDNA in patients with OC, but the number of studies is limited, and the methods vary among the studies [8–11]. Furthermore,

the use of WGS for ctDNA detection focusing on SNVs in OC has not yet been reported. The limited tools available for ctDNA detection through WGS are often complex and challenging to use.

OC accounted for 3.4% of cancer cases and 4.7% of cancer-related deaths in women in 2020 [12]. High-grade serous carcinoma of the ovaries, fallopian tubes, and peritoneum is the most lethal gynecological cancer. It is the most common type of OC [13,14] and, because of the aggressive nature of the disease, around two-thirds of the patients are diagnosed in advanced stages (FIGO stage III–IV) [14]. The treatment of ovarian cancer is a combination of surgery and chemotherapy, as well as targeted treatment with PARP inhibitors or bevacizumab for specific groups of patients, e.g., patients with somatic and germline *BRCA* mutations or homologous recombination deficiency [14,15]. Despite complete resection of advanced disease and treatment with chemotherapy, the majority of these patients will relapse within a few years [14], and the five-year survival rate is limited to 30–40% [16]. The current biomarker for epithelial OC, CA-125, lacks specificity and sensitivity [14], and the low survival rate indicates that a more specific and sensitive biomarker for OC could potentially increase the survival rate of the disease.

In this study, we established a sensitive and transparent method for tumor-informed SNV detection of ctDNA in the blood of OC patients using WGS, which to the best of our knowledge has not been done before. A tumor-informed method, using a combination of various allelic and read depth filters with different cut-offs, an additional panel of normals (leucocyte DNA), and a new filtering approach based on forced calling of tumor variants in a pool of plasma samples, was assessed to improve ctDNA signal detection in plasma. Subsequently, the utility of this method for ctDNA signal detection was tested at a reduced sequencing depth in plasma.

## 2. Results

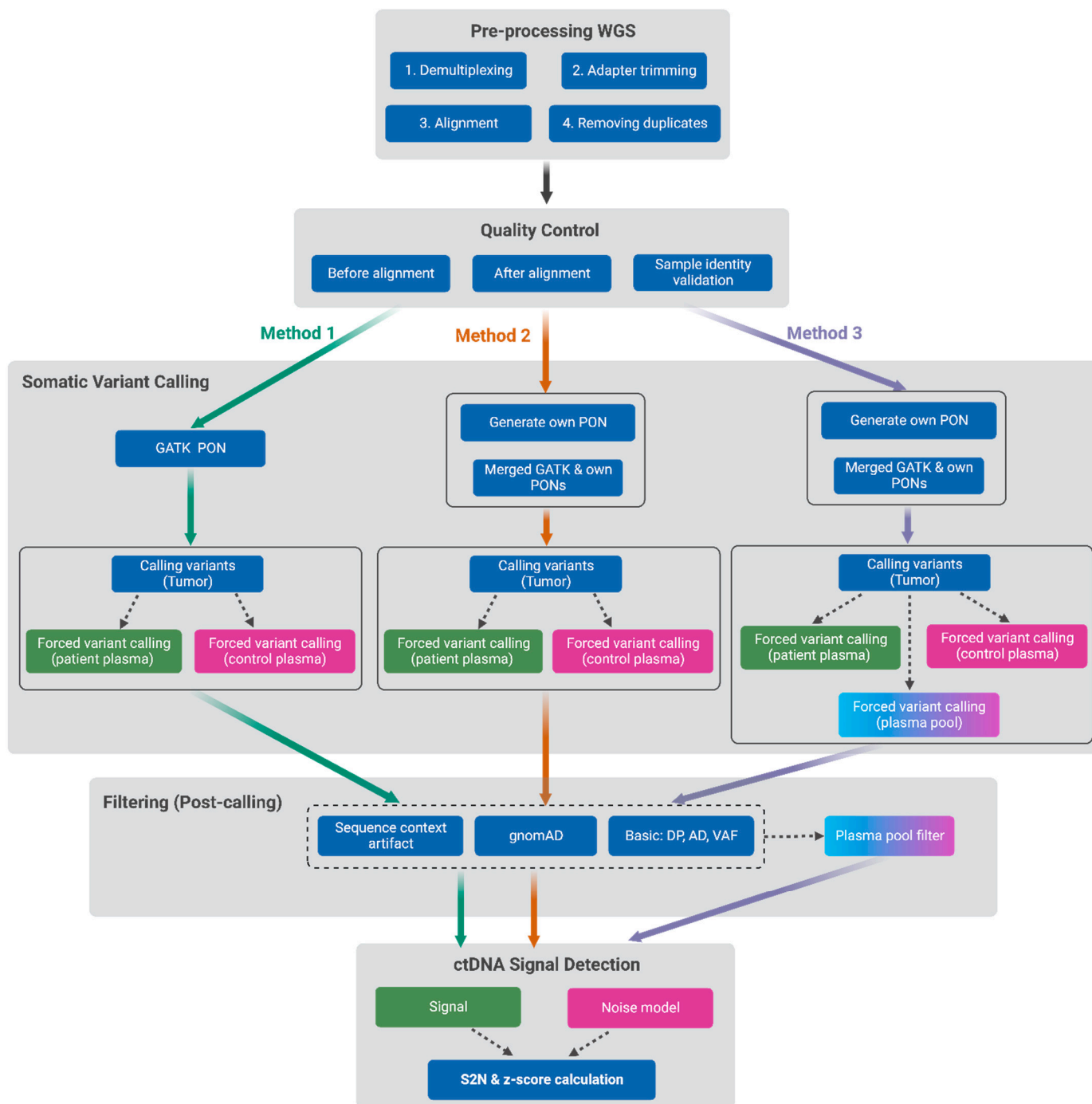
An overview of the WGS data analysis pipeline and the three methods tested in this study is illustrated in Figure 1. A detailed description is included in the Materials and Methods (Section 4).

### 2.1. Clinical Data

Samples from 10 patients with high-grade serous OC were included in this study. Because of the short half-life of ctDNA (less than two hours [17]), blood samples were collected before treatment was initiated. After this, all patients received primary surgery and adjuvant chemotherapy. The tumor samples were collected after surgical removal. Due to contamination of one of the tumor samples, one patient was excluded. Plasma samples from nine patients with benign ovarian tumors were used as control samples. Table 1 summarizes the clinical characteristics of the included patients and controls.

### 2.2. Basic Noise Filtering in Tumor-Informed Plasma Analysis

To establish the best filters for identification of mutations in plasma, signal-to-noise ratios (S2N) were computed. The S2N is the average variant allele frequency (VAF) or alternate allelic depth (AD) in the plasma for mutations identified in the tumor, divided by the average VAF (or AD) of the same mutations in the control plasma. The number of variants detected in the tumor and the S2N for all of the applied filter thresholds are presented in Table S1. Increasing the filter threshold for the read depth (DP) of normal DNA and the DP, AD, and VAF of tumor DNA showed a limited improvement in VAF and AD signal-to-noise ratios (Figures 2 and S1 and Table S1). For the analyses presented in the following paragraphs, we applied the most optimal setting of filter thresholds (Figure 2, Filter E).



**Figure 1.** Overview of the Whole-Genome Sequencing (WGS) data analysis pipeline for ctDNA detection. Method 1: 1000\_g panel of normals (PON); Method 2: merged PON; Method 3: merged PON and plasma pool filter. Read Depth (DP), Allelic Depth (AD), variant allele frequency (VAF).

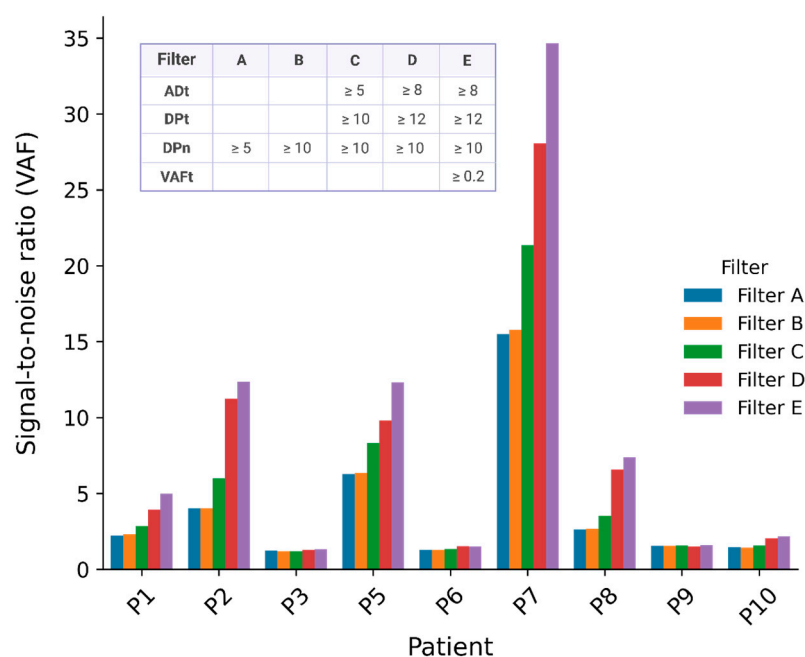
**Table 1.** Clinical characteristics of patients and controls.

Patient Data	Number of Patients (9)	Number of Controls (9)
	Pathology	
HGSC	9	
Serous adenoma		4
Mucinous adenoma		1
Fibroma		1
Other		3

Table 1. Cont.

Patient Data	Number of Patients (9)	Number of Controls (9)
	Age at diagnosis (years)	
<65	6	6
≥65	3	3
	FIGO stage	
I	2	NA
II	1	NA
III	3	NA
IV	3	NA

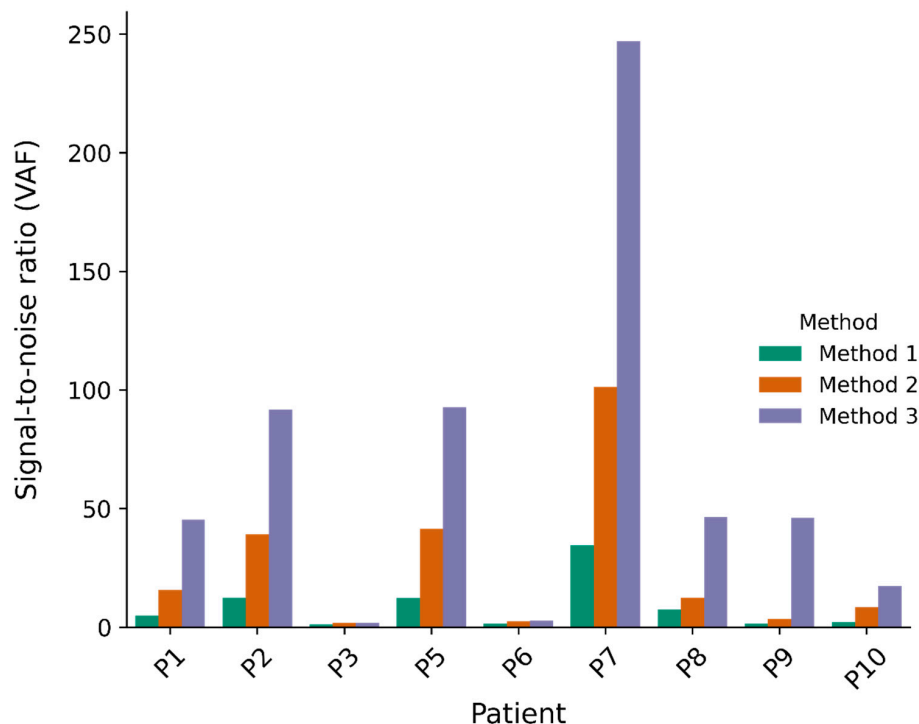
High-grade serous carcinoma (HGSC); not applicable (NA); FIGO stage I: tumor confined to ovaries; FIGO stage II: tumor involves one or both ovaries with pelvic extension (below the pelvic brim) or primary peritoneal cancer; FIGO stage III: tumor involves one or both ovaries with cytologically or histologically confirmed spread to the peritoneum outside the pelvis and/or metastasis to the retroperitoneal lymph nodes; FIGO stage IV: distant metastasis excluding peritoneal metastasis [14].



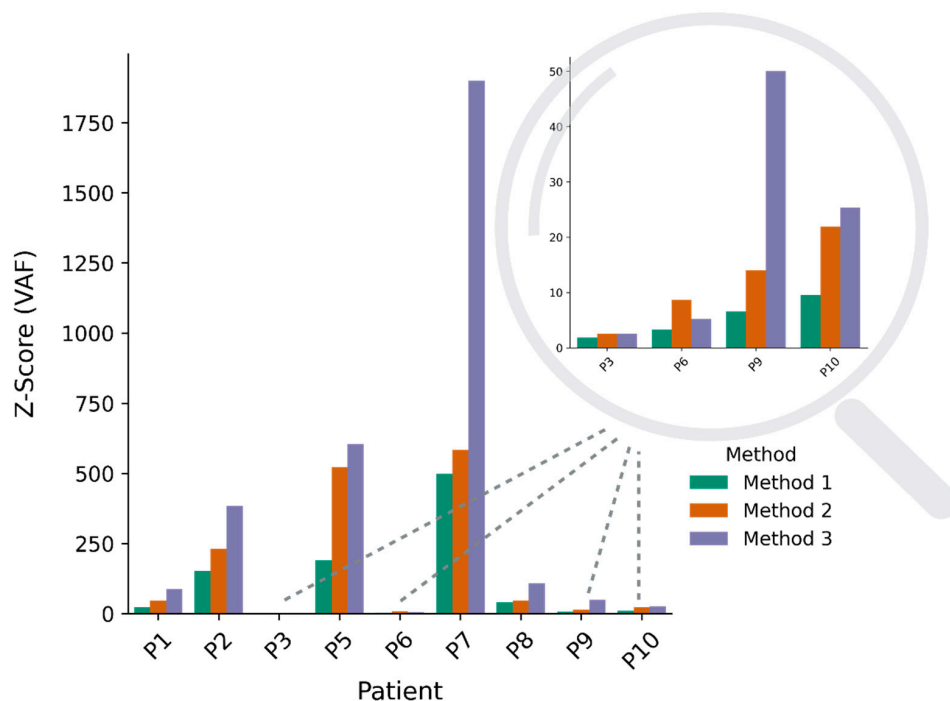
**Figure 2.** Overview of VAF signal-to-noise ratios after applying different filters: The stringency of the filter parameters increases progressively from Filter A to E as the filter threshold is altered for AD and DP of tumor and normal DNA, as well as for VAF of tumor DNA; 1000\_g PON (Method 1) and ADn = 0 are applied to all. ADt (allelic depth in tumor), DPt (read depth in tumor), VAFt (variant allele frequency in tumor), ADn (allelic depth in normal), DPn (read depth in normal).

### 2.3. Advanced Filtering with Different Approaches and ctDNA Signal Detection

For the first part of this study, a panel of normals (PON) (1000\_g PON) provided by GATK was used (Method 1). To test whether an additional PON would reduce the level of noise, we introduced a self-generated PON based on normal DNA samples from the patients included in this study. Using a merged PON (Method 2) increased the VAF S2N for all patients (Figure 3). The improvement ranged from a factor of 1.4 to 3.9 (Table S2). Similar results were seen for AD S2N (Figure S2). We also calculated the z-score, a good statistical measure used to evaluate the strength of the signal relative to the noise. This is calculated based on the mean and standard deviation of the noise model (control plasmas). Further details can be found in Section 4.6. The z-scores for both VAF and AD increased after applying the merged PON (Figure 4 and Table S2).



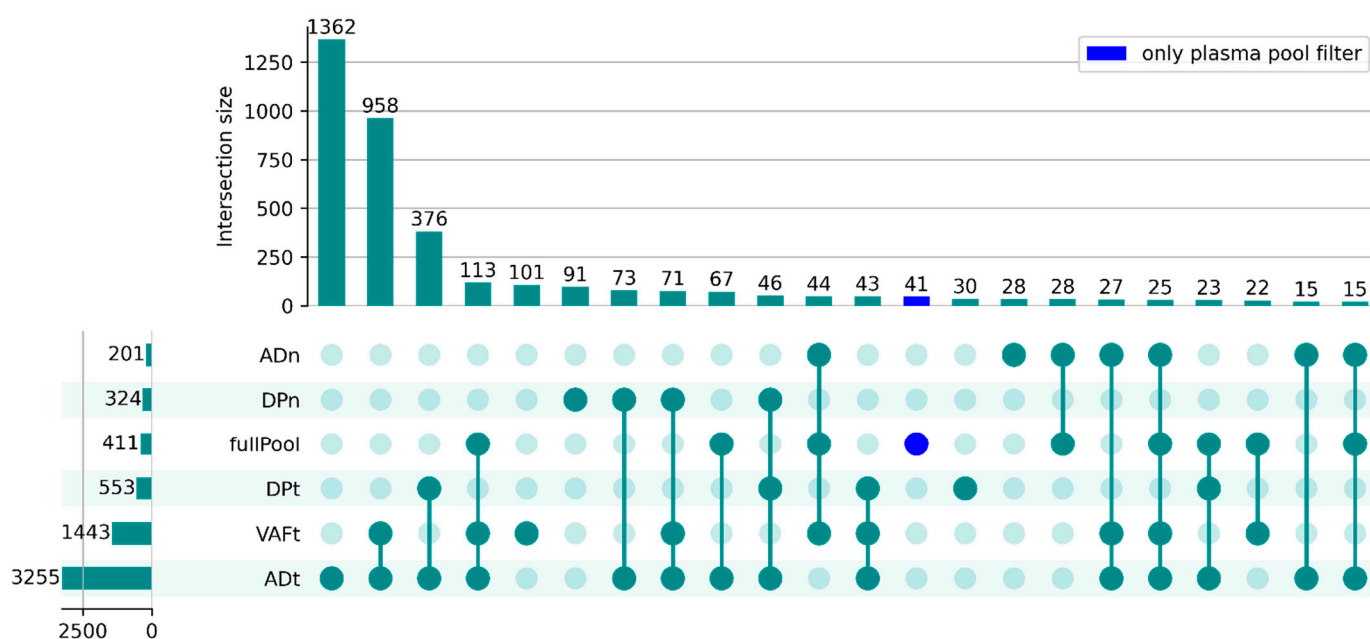
**Figure 3.** Overview of VAF signal-to-noise ratios when different methods for filtering noise are applied. Applied filtering methods: Method 1: 1000\_g PON; Method 2: merged PON; Method 3: merged PON and plasma pool filter. Filtering thresholds  $ADt \geq 8$ ,  $DPt \geq 12$ ,  $Adn = 0$ ,  $DPn \geq 10$ , and  $VAFt \geq 0.2$  were used in all three methods.



**Figure 4.** Overview of z-scores computed based on VAF when different methods for filtering noise are applied. The magnifying glass enlarges the z-score (VAF) in the four patients (P3, P6, P9, and P10) with low z-score results compared to the other five patients. Applied filtering parameters: Method 1: 1000\_g PON; Method 2: merged PON; Method 3: merged PON and plasma pool filter. Filtering thresholds  $ADt \geq 8$ ,  $DPt \geq 12$ ,  $Adn = 0$ ,  $DPn \geq 10$ , and  $VAFt \geq 0.2$  were used in all three methods.

When inspecting the data from Methods 1 and 2 in detail, we expected VAF values close to zero in the control plasmas. However, in a relatively low number of positions we observed outliers noticeably above zero, indicative of general sequencing errors. To remove these errors and thereby decrease the level of noise even further, we introduced a plasma pool filter that removed variants if they were seen with  $AD > 1$  in at least one plasma sample from the plasma pool of the other nine OC patients (Method 3). This further improved the S2N and z-score for both VAF (Figures 3 and 4) and AD (Figures S2 and S3) in seven patients (Table S2). The binary z-scores improved in four patients (Figure S4). Interestingly, the addition of the plasma pool filter increased the S2N and z-scores to the same level when only applying the 1000\_g PON (Table S3).

UpSet plots [18] were generated to provide an overview of the number of variants removed by each and more than one of the different filters. The plot for patient 1 is illustrated in Figure 5 (the plots for the rest of the patients are in Figures S5–S12). In patient 1, the plasma pool filter removed 411 variants, most of which (370) were also removed by other filters such as AD and VAF in the tumor. However, by filtering out the few variants (41) that were only removed by the plasma pool filter, the S2N and z-score increased considerably. This pattern was also seen in other patients, and the improvement is noticeable from Method 2 to Method 3 in Figure 3, except for patients 3 and 6 (FIGO stage I), who showed no clear improvement.

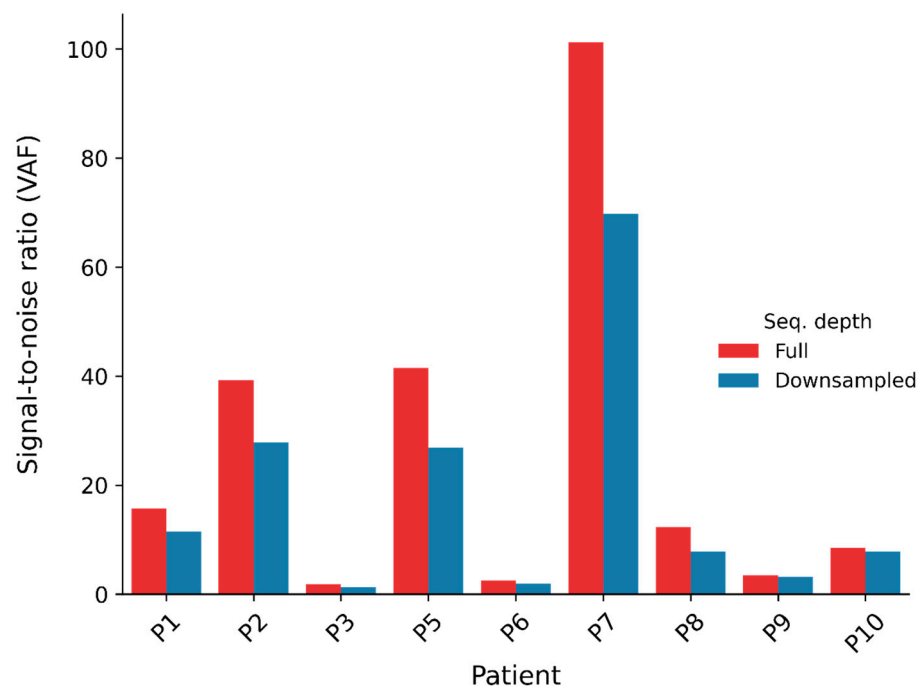


**Figure 5.** Overview of the number of variants removed by filters and the overlap between them for patient 1: UpSet plot shows the number of variants filtered by each filter, as well as the intersection between different filters. Intersection combinations with fewer than 15 variants are not shown. Applied filter thresholds:  $ADt \geq 8$ ,  $DPt \geq 12$ ,  $ADn = 0$ ,  $DPn \geq 10$ ,  $VAFt \geq 0.2$ , and fullPool (plasma pool). The merged PON was applied.

#### 2.4. Downsampling of Plasma

To test whether the ctDNA signal is detectable in plasma sequenced with a lower depth, we performed the SNV calling analysis when the plasma was downsampled to 40 percent. This was equivalent to a sequencing depth of around  $20\times$ . This reduced the VAF S2N and z-scores in all patients (Figure 6 and Table S2). However, these values were still high, indicating identifiable ctDNA signals at a lower sequencing depth in the seven non-stage-I patients (Table S3).





**Figure 6.** Overview of VAF signal-to-noise ratios for sequenced data at full and reduced depth (plasma downsampled to approximately 20 $\times$ ). The merged PON (Method 2) was applied. The plasma pool filter was not used.

### 3. Discussion

This method study demonstrates a tumor-informed approach for SNV detection in ctDNA from the blood of OC patients using WGS. With WGS, we were able to use the combined pattern of all identified SNVs in the plasma to detect ctDNA signal. By focusing on how to mitigate noise through different strategies, we found that the plasma pool filter clearly improved ctDNA signal detection. The ctDNA signal was still detectable even when the sequencing depth of plasma was reduced to around 20 $\times$ .

No somatic variant caller for tumor DNA is perfect. Some variants are missed, and false positive positions (noise) are detected as true variants. We need a method that is useful even when the ctDNA content is low. It is therefore especially important to reduce noise in order to distinguish true variants from noise. The first step of this study was to evaluate the use of basic noise filtering and its ability to reduce noise in plasma. By increasing the filter thresholds in normal and tumor DNA, the number of identified variants dropped dramatically, while only a slight improvement was seen in S2N. The levels were therefore set with the balance between reducing noise and removal of the fewest true tumor variants in mind. These thresholds are not necessarily fixed and can depend on cancer type, tumor content, and the applied sequencing depth.

By applying the merged PON (Method 2), we observed a clear improvement. This indicates the effectiveness of applying a self-generated PON with normal samples that have been sequenced similarly to the tumor and plasma samples to reduce the level of noise. Our self-generated PON included ten normal DNA samples and was able to make a significant improvement. Increasing the number of these samples might improve noise filtering even further.

In patients with localized primary or recurrent cancer, it is likely that the ctDNA level is low; hence, a useful method for ctDNA detection should be sensitive enough to identify these cases. To obtain sensitive ctDNA analysis, it is vital that the applied method is able to remove background noise. To reduce the level of noise even further, we introduced a plasma pool filter consisting of plasma samples from the other OC patients included in the study, and we removed variants seen in at least one of the plasma pool samples (with AD > 1) from the patient plasma and from the control plasmas before further computation



(Method 3). The plasma pool filter focuses on the variants identified in the tumor and assesses the level of noise in each variant position by performing forced variant calling of the tumor variants in all of the plasma pool samples. This allows for more sensitive noise reduction compared to other methods that use a general list of artefactual variants for noise removal, such as the plasma blacklist filter described as part of MRD-EDGE [7]. Even though the used plasma pool consisted of only nine patients, this filter showed the greatest improvement in ctDNA signal in seven out of nine patients, while removing relatively few variants. This was independent of whether the filter was applied before or after the addition of the merged PON (Method 2). When applying the plasma pool filter, it is important that the control plasma samples are treated in an identical way. All plasma samples were sequenced in the same way, and the plasma samples from the OC patients even in the same sequencing run. The plasma samples were therefore comparable, and there was only a slight variation in sequencing depth among them. The average sequencing depth was  $52\times$  among the plasmas from OC patients and  $49\times$  for the control plasmas. Furthermore, the computation and the assessment of variants identified in the patient's own plasma, the plasma pool, and the control plasmas were run in identical manners. There is a slight risk that a true recurrent mutation could be removed by the plasma pool filter, but as recurrent mutations are rare, this would only lead to a minimal reduction in the signal.

In the current study, the z-score was calculated based on VAF, AD, and in a binary version. The binary z-score has been used in some other studies on ctDNA [3,7]. Our results showed that the introduction of plasma pool filtering in Method 3 improved the S2N, VAF, and AD z-scores. In contrast, the binary z-score showed no clear improvement (Table S2). This addition of the plasma pool filter removed noisy outliers; thus, the results with the least amount of noise were extracted after the application of Method 3.

The sensitivity of the method was mainly determined by the average VAF in the noise model (Table S2). After application of the plasma pool filter (Method 3), the level of noise was calculated as average VAF with values 3–4 times  $10^{-4}$  in all patients, except patient 9. This was independent of the set of variants called, giving an indication of the maximum obtainable sensitivity. Seven of the samples showed signals far above the noise level. In the last two samples, the signals were two- and three-fold the level of noise, which was still significant at the 5% level (z-score higher than 1.895, one-sided). For patient 6, it was even significant at the 1% level.

The limitations of this study are the small number of patients included and the fact that all tumor samples had a very favorable tumor cell fraction of at least 75%. The method therefore needs to be validated on a higher number of patients with different tumor cell fractions. Furthermore, the method should be tested for different clinical applications. A tumor-informed approach such as the method applied in this study could potentially be useful for monitoring of treatment response, and in early detection of recurrence.

In conclusion, this method study is the first to demonstrate a successful method for ctDNA detection in OC patients using a tumor-informed approach for WGS focusing on SNVs. Despite the relatively low number of patients involved in this study, the results suggest a significant potential of our method in noise removal and ctDNA signal detection. It is therefore promising in the search for a biomarker in OC that might improve the survival rate of the disease. Furthermore, this method is readily comprehensible and could easily be applied and tested on other cancer types.

## 4. Materials and Methods

### 4.1. Inclusion of Patients

Patients with suspected primary OC were enrolled in the study from January 2016 to May 2017 at the Department of Gynecology and Obstetrics, Odense University Hospital, after giving written informed consent. From each patient, a pretreatment venous blood sample was collected in two Cell-Free DNA BCT tubes (STRECK, 230470, La Vista, NE, USA), and a frozen tissue sample was obtained from the tumor. Among the enrolled

patients, 10 patients with high-grade serous OC and a tissue sample with a tumor cell fraction of at least 75%, as assessed by the pathologist, were selected for further analysis.

#### 4.2. Sample Processing

Plasma and buffy coat (normal peripheral blood cells) were isolated from the blood samples and frozen within three days of collection. CfDNA was extracted from approximately 4 mL of plasma using the QIAamp Circulating Nucleic Acid Kit (QIAGEN, 55114, Venlo, The Netherlands), normal DNA was isolated using the Maxwell RSC Blood DNA (Promega, AS1400, Madison, WI, USA), and tumor DNA was extracted using the Maxwell RSC Tissue DNA Kit (Promega, AS1610, Madison, WI, USA), according to the manufacturers' instructions. The Qubit dsDNA HS assay kit (Thermo Fisher Scientific, Q32854, Waltham, MA, USA) was used to assess the concentration of isolated DNA. To prepare samples for NGS, the ThruPLEX TagSeq HV Kit (Takara Bio, R400743, Kusatsu, Japan) was used on extracted cfDNA, tumor DNA, and normal leucocyte DNA, following the manufacturer's instructions. Unique dual indices (ThruPLEX HV UDI Set A (Takara Bio, 400738, Kusatsu, Japan) were used to enable the removal of hopped reads from the downstream analysis.

#### 4.3. Pre-Processing of WGS Data

Samples were sequenced on an Illumina NovaSeq 6000 with paired-end  $2 \times 150$  base pairs. The yielded average sequencing coverage was  $17\times$  for normal DNA,  $24\times$  for tumor DNA,  $52\times$  for patient plasma, and  $49\times$  for control plasma. An overview of the computational analysis pipeline is illustrated in Figure 1. The raw sequenced data were pre-processed by first demultiplexing using the bcl2fastq tool (version 2.19.0) [19] and adapter trimming with Skewer software (version 0.2.2) [20] before the alignment to the human reference genome (GRCh38) using the BWA-MEM tool (version 0.7.17) [21]. Next, SortSam from Picard (version 4.2.6.1) [22] was used to sort and index the aligned files. This was followed by removing duplicates from the BAM files using MarkDuplicates from the GATK package (version 4.2.6.1) [23]. In addition to performing the quality control of the samples' data with standard tools such as FastQC [24], Qualimap [25], and MultiQC [26], the NGSCheckmate tool (version 1.0.1) [27] was used to validate sample identity. This indicated misidentification or contamination of patient 4's tumor sample. Therefore, this sample was eliminated from the study. However, the plasma from this patient passed the identity check and was included in the plasma pool filter (see Section 4.5). To simulate a reduced sequencing depth (around  $20\times$ ) in plasma, a downsampling tool (DownsampleSam from PICARD (version 3.0.0) [28]) was used on the BAM files to randomly include 40 percent of the reads from the original WGS data ( $\sim 50\times$ ).

#### 4.4. Somatic Variant Calling

Mutect2 from the GATK package (version 4.2.6.1) [29] was used for SNV calling. First, Mutect2 was run on the tumor and matched normal BAM files to call somatic variants while using the PON provided by GATK (1000\_g PON) for pre-calling filtering. Using the FilterMutectCalls method from GATK, we filtered out variants that were due to contamination (using the contamination table generated by GATK's CalculateContamination function) and sequence context artefacts (generated by GATK's LearnReadOrientationModel function). To implement the tumor-informed strategy for variant calling in plasma, Mutect2 was executed utilizing the variants identified in the corresponding tumor as the input set of alleles for force calling. To generate the noise model per patient, the variants detected in the patient's tumor were force-called on every control plasma sample.

#### 4.5. Different Methods for Filtering Variants

The vcf files from Mutect2 were imported to VarSeq (version 2.4.0) (Golden Helix Inc., Bozeman, MT, USA) [30]. After applying a gnomAD filter (gnomAD Genomes and Exomes variant frequencies 2.0.1) with  $AD \leq 0.00001$  in VarSeq, a concatenated table of all vcf

files, including the information on the read depth (DP), AD, and VAF for their own tumor, normal, plasma, control plasmas, and the plasma pool (plasma from other OC-patients), was exported per patient. To assess the effects of basic filters, varying filter thresholds on DP, AD, and VAF for the tumor and normal DNA were applied to the Mutect2 results using Python scripts developed in-house. For this, the 1000\_g PON was used in the variant calling (Method 1).

For more effective capturing and removal of technical artefacts, an additional PON from the normal DNA of the 10 OC patients, prepared by the same library kit and sequenced by the same technology as used for tumors and plasma, was generated. This self-generated PON was merged with the PON provided by Mutect2 and used in the first step of the variant calling pipeline (Method 2). To reduce the noise level further, we introduced the plasma pool filter that removed variants detected in at least one plasma sample from the nine other OC patients with AD > 1 (Method 3) (Figure 1).

#### 4.6. ctDNA Signal Detection

To assess the ctDNA signal in plasma, we calculated the S2N and z-scores. S2N is calculated as the average of VAF (or AD) in the plasma for mutations identified in the tumor (after applying filters), divided by the average VAF (or AD) of the same mutations in the control plasmas (constituting our noise model):

$$S2N = \frac{Ave_{VarDet}}{\mu_{noise}} \quad (1)$$

where  $Ave_{VarDet}$  is the average VAF (or AD) of detected variants in the patient's plasma, and  $\mu_{noise}$  is the mean VAF (or AD) in controls. Similarly, z-scores were calculated based on the mean and standard deviation of the noise model as follows:

$$z - score = \frac{Ave_{VarDet} - \mu_{noise}}{\sigma_{noise}} \quad (2)$$

where  $\sigma_{noise}$  is the standard deviation of VAF (or AD) in controls. We also computed a binary z-score, i.e., the values of VAF and AD of the variants were not included in the calculation, only the number of detected variants.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms252413349/s1>.

**Author Contributions:** Conceptualization, C.F.T., S.S., K.M.J., T.A.K. and M.T.; data curation, C.F.T. and S.S.; formal analysis, C.F.T. and S.S.; funding acquisition, C.F.T., K.M.J., T.A.K. and M.T.; investigation, C.F.T., S.S., K.M.J., T.A.K. and M.T.; methodology, C.F.T., S.S., T.A.K. and M.T.; project administration, C.F.T.; software, S.S.; supervision, K.M.J., T.A.K. and M.T.; visualization, C.F.T. and S.S.; writing—original draft, C.F.T. and S.S.; writing—review and editing, C.F.T., S.S., K.M.J., T.A.K. and M.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Region of Southern Denmark (18/17528 and 20/14090), Odense University Hospital (R66-A3026 and R110-A4580), Holms Mindelegat (20035), KV Fonden (2020-2665), Fonden til Lægevidenskabens Fremme (part of The AP Moller Foundation, 20-L-0198), Aase & Ejnar Danielsens Fond (21-10-0058), Tornøes & Høyrups Legat (32120), and Dagmar Marshalls Fond.

**Institutional Review Board Statement:** This study was conducted in accordance with the Declaration of Helsinki and approved by the Regional Committee on Health Research Ethics in the Region of Southern Denmark (Project-ID: S-20150146; approved 20 November 2015).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in this study.

**Data Availability Statement:** Our ethical approval does not allow for the sharing of supporting sequencing data. The scripts used for bioinformatics analysis are available upon reasonable request.

**Acknowledgments:** The authors would like to thank the patients who participated in this study. Furthermore, we would like to thank the doctors and nurses at the Department of Gynecology and Obstetrics, Odense University Hospital; the staff at the Department of Pathology, Odense University Hospital; and Kristina Magaard Koldby and the laboratory technicians at the Department of Clinical Genetics, Odense University Hospital. Figures 1, 2 and 4 were generated using Biorender.com.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Crowley, E.; Di Nicolantonio, F.; Loupakis, F.; Bardelli, A. Liquid biopsy: Monitoring cancer-genetics in the blood. *Nat. Rev. Clin. Oncol.* **2013**, *10*, 472–484. [CrossRef] [PubMed]
2. Siravegna, G.; Marsoni, S.; Siena, S.; Bardelli, A. Integrating liquid biopsies into the management of cancer. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 531–548. [CrossRef] [PubMed]
3. Zviran, A.; Schulman, R.C.; Shah, M.; Hill, S.T.K.; Deochand, S.; Khamnei, C.C.; Maloney, D.; Patel, K.; Liao, W.; Widman, A.J.; et al. Genome-wide cell-free DNA mutational integration enables ultra-sensitive cancer monitoring. *Nat. Med.* **2020**, *26*, 1114–1124. [CrossRef] [PubMed]
4. Wan, J.C.M.; Heider, K.; Gale, D.; Murphy, S.; Fisher, E.; Mouliere, F.; Ruiz-Valdepenas, A.; Santonja, A.; Morris, J.; Chandrananda, D.; et al. ctDNA monitoring using patient-specific sequencing and integration of variant reads. *Sci. Transl. Med.* **2020**, *12*, eaaz8084. [CrossRef] [PubMed]
5. Santonja, A.; Cooper, W.N.; Eldridge, M.D.; Edwards, P.A.W.; Morris, J.A.; Edwards, A.R.; Zhao, H.; Heider, K.; Couturier, D.L.; Vijayaraghavan, A.; et al. Comparison of tumor-informed and tumor-naïve sequencing assays for ctDNA detection in breast cancer. *EMBO Mol. Med.* **2023**, *15*, e16505. [CrossRef]
6. Bae, M.; Kim, G.; Lee, T.R.; Ahn, J.M.; Park, H.; Park, S.R.; Song, K.B.; Jun, E.; Oh, D.; Lee, J.W.; et al. Integrative modeling of tumor genomes and epigenomes for enhanced cancer diagnosis by cell-free DNA. *Nat. Commun.* **2023**, *14*, 2017. [CrossRef]
7. Widman, A.J.; Shah, M.; Frydendahl, A.; Halmos, D.; Khamnei, C.C.; Øgaard, N.; Rajagopalan, S.; Arora, A.; Deshpande, A.; Hooper, W.F.; et al. Ultrasensitive plasma-based monitoring of tumor burden using machine-learning-guided signal enrichment. *Nat. Med.* **2024**, *30*, 1655–1666. [CrossRef]
8. Bak, S.E.; Kim, H.; Ho, J.Y.; Cho, E.-H.; Lee, J.; Youn, S.M.; Park, S.-W.; Han, M.-R.; Hur, S.Y.; Lee, S.J.; et al. Shallow Whole-Genome Sequencing of Cell-Free DNA (cfDNA) Detects Epithelial Ovarian Cancer and Predicts Patient Prognosis. *Cancers* **2023**, *15*, 530. [CrossRef]
9. Braicu, E.I.; du Bois, A.; Sehouli, J.; Beck, J.; Prader, S.; Kulbe, H.; Eiben, B.; Harter, P.; Traut, A.; Pietzner, K.; et al. Cell-Free-DNA-Based Copy Number Index Score in Epithelial Ovarian Cancer-Impact for Diagnosis and Treatment Monitoring. *Cancers* **2021**, *14*, 168. [CrossRef]
10. Sabatier, R.; Garnier, S.; Guille, A.; Carbuccia, N.; Pakradouni, J.; Adelaide, J.; Provansal, M.; Cappiello, M.; Rousseau, F.; Chaffanet, M.; et al. Whole-genome/exome analysis of circulating tumor DNA and comparison to tumor genomics from patients with heavily pre-treated ovarian cancer: Subset analysis of the PERMED-01 trial. *Front. Oncol.* **2022**, *12*, 946257. [CrossRef]
11. Zhou, H.; Zhang, X.; Liu, Q.; Yang, J.; Bai, J.; Yin, M.; Cao, D.; Zhang, Q.; Zheng, L. Can circulating cell free DNA be a promising marker in ovarian cancer?—a genome-scale profiling study in a single institution. *J. Ovarian Res.* **2023**, *16*, 11. [CrossRef] [PubMed]
12. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [CrossRef] [PubMed]
13. Momenimovahed, Z.; Tiznobaik, A.; Taheri, S.; Salehiniya, H. Ovarian cancer in the world: Epidemiology and risk factors. *Int. J. Women's Health* **2019**, *11*, 287–299. [CrossRef] [PubMed]
14. Berek, J.S.; Renz, M.; Kehoe, S.; Kumar, L.; Friedlander, M. Cancer of the ovary, fallopian tube, and peritoneum: 2021 update. *Int. J. Gynecol. Obstet.* **2021**, *155*, 61–85. [CrossRef]
15. Ray-Coquard, I.; Leary, A.; Pignata, S.; Cropet, C.; González-Martín, A.; Marth, C.; Nagao, S.; Vergote, I.; Colombo, N.; Mäenpää, J.; et al. Olaparib plus bevacizumab first-line maintenance in ovarian cancer: Final overall survival results from the PAOLA-1/ENGOT-ov25 trial. *ESMO Ann. Oncol.* **2023**, *34*, 681–692. [CrossRef]
16. World Ovarian Cancer Coalition. The World Ovarian Cancer Coalition Atlas. Available online: <https://worldovariancancercoalition.org/wp-content/uploads/2018/10/THE-WORLD-OVARIAN-CANCER-COALITION-ATLAS-2018.pdf> (accessed on 15 June 2023).
17. Diehl, F.; Schmidt, K.; Choti, M.A.; Romans, K.; Goodman, S.; Li, M.; Thornton, K.; Agrawal, N.; Sokoll, L.; Szabo, S.A.; et al. Circulating mutant DNA to assess tumor dynamics. *Nat. Med.* **2008**, *14*, 985–990. [CrossRef]
18. Lex, A.; Gehlenborg, N.; Strobel, H.; Vuillemot, R.; Pfister, H. UpSet: Visualization of Intersecting Sets. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 1983–1992. [CrossRef]
19. Illumina. Available online: [https://support.illumina.com/sequencing/sequencing\\_software/bcl2fastq-conversion-software.html](https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html) (accessed on 9 December 2024).

20. Jiang, H.; Lei, R.; Ding, S.-W.; Zhu, S. Skewer: A fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinform.* **2014**, *15*, 182. [CrossRef]
21. Burrows-Wheeler Aligner. Available online: <https://bio-bwa.sourceforge.net/> (accessed on 9 December 2024).
22. GATK. Available online: <https://gatk.broadinstitute.org/hc/en-us/articles/5358884334747-SortSam-Picard> (accessed on 9 December 2024).
23. GATK. Available online: <https://gatk.broadinstitute.org/hc/en-us/articles/5358880192027-MarkDuplicates-Picard> (accessed on 9 December 2024).
24. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. Available online: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed on 9 December 2024).
25. Okonechnikov, K.; Conesa, A.; García-Alcalde, F. Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **2015**, *32*, 292–294. [CrossRef]
26. Ewels, P.; Magnusson, M.; Lundin, S.; Käller, M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **2016**, *32*, 3047–3048. [CrossRef]
27. Lee, S.; Lee, S.; Ouellette, S.; Park, W.-Y.; Lee, E.A.; Park, P.J. NGSCheckMate: Software for validating sample identity in next-generation sequencing studies within and across data types. *Nucleic Acids Res.* **2017**, *45*, e103. [CrossRef] [PubMed]
28. GATK. Available online: <https://gatk.broadinstitute.org/hc/en-us/articles/5358860680987-DownsampleSam-Picard> (accessed on 9 December 2024).
29. Benjamin, D.; Sato, T.; Cibulskis, K.; Getz, G.; Stewart, C.; Lichtenstein, L. Calling Somatic SNVs and Indels with Mutect2. *bioRxiv* **2019**. [CrossRef]
30. GoldenHelix. Available online: <https://www.goldenhelix.com/products/VarSeq/> (accessed on 9 December 2024).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.