cancers

MDPI

Article

Benchmarking of Approaches for Gene Copy-Number Variation Analysis and Its Utility for Genetic Aberration Detection in High-Grade Serous Ovarian Carcinomas

Pavel Alekseevich Grebnev [1], Ivan Olegovich Meshkov [1], Pavel Viktorovich Ershov [1], Antonida Viktorovna Makhotenko [1], Valentina Bogdanovna Azarian [1], Marina Vyacheslavovna Erokhina [1], Anastasiya Aleksandrovna Galeta [1], Aleksandr Vladimirovich Zakubanskiy [1], Olga Sergeevna Shingalieva [1], Anna Vasilevna Tregubova [2], Aleksandra Vyacheslavovna Asaturova [2], Vladimir Sergeevich Yudin [1], Sergey Mihaylovich Yudin [1], Valentin Vladimirovich Makarov [1], Anton Arturovich Keskinov [1], Anna Sergeevna Makarova [1], Ekaterina Andreevna Snigir [1,*] and Veronika Igorevna Skvortsova [3]

[1] Federal State Budgetary Institution "Centre for Strategic Planning and Management of Biomedical Health Risks" of the Federal Medical Biological Agency (Centre for Strategic Planning of FMBA of Russia), Bld. 1, Pogodinskaya Street, 10, 119121 Moscow, Russia; pgrebnev@cspfmba.ru (P.A.G.); imeshkov@cspfmba.ru (I.O.M.); pershov@cspfmba.ru (P.V.E.); amahotenko@cspfmba.ru (A.V.M.); vazaryan@cspfmba.ru (V.B.A.); merokhina@cspfmba.ru (M.V.E.); agaleta@cspfmba.ru (A.A.G.); azakubanskiy@cspfmba.ru (A.V.Z.); oshingalieva@cspfmba.ru (O.S.S.); vyudin@cspfmba.ru (V.S.Y.); yudin@cspfmba.ru (S.M.Y.); makarov@cspfmba.ru (V.V.M.); amakarova@cspfmba.ru (A.S.M.)

[2] Federal State Budgetary Institution "National Medical Research Center for Obstetrics, Gynecology and Perinatology Named after Academician V.I. Kulakov", Ministry of Healthcare of The Russian Federation, Oparina Street, Bld. 4, 117997 Moscow, Russia; annyupitrue@mail.ru (A.V.T.); velikova@yandex.ru (A.V.A.)

[3] Federal Medical Biological Agency (FMBA of Russia), Volokolamskoye Shosse, 30, 123182 Moscow, Russia; skvortsova@cspfmba.ru

* Correspondence: esnigir@cspfmba.ru

Simple Summary: Our research integrates two key areas of focus. We first focus on a comparative biostatistical analysis of the agreement between three laboratory methods for genotyping gene copy number variations, and then we explore new characteristics of the distribution of this type of genetic aberration among cancer-associated genes. The results could be used to help molecular oncologists pick the smallest set of methods that will allow for accurate CNV genotyping in personalized cancer prognosis and drug-response prediction. We see the following strengths of our research: (1) We employ various methods for the detection of CNVs, such as CoreExome microarrays, nCounter v2 Cancer CN Assay panel, and digital droplet PCR; (2) This study sets out an interpretable comprehensive biostatistical analysis of the agreement using the PABAK score and Passing–Bablok regression; (3) This study presents the detection of CNVs in 12 paired tumor ovarian cancer/normal samples with clinical data; and (4) The research has translational potential.

Abstract: Objective: The goal of this study was to compare the results of CNV detection by three different methods using 13 paired carcinoma samples, as well as to perform a statistical analysis of the agreement. **Methods**: CNV was studied using NanoString nCounter v2 Cancer CN Assay (Nanostring), Illumina Infinium CoreExome microarrays (CoreExome microarrays) and digital droplet PCR (ddPCR). **Results**: There was a good level of agreement (PABAK score > 0.6) between the CoreExome microarrays and the ddPCR results for finding CNVs. There was a moderate level of agreement (PABAK values ≈ 0.3–0.6) between the NanoString Assay results and microarrays or ddPCR. For 83 out of 87 target genes studied (95%), the agreement between the CoreExome microarrays and NanoString nCounter was characterized by PABAK values < 0.75, except for MAGI3, PDGFRA, NKX2-1 and KDR genes (>0.75). The MET, HMGA2, KDR, C8orf4, PAX9, CDK6, and CCND2 genes had the highest agreement among all three approaches. **Conclusions**: Therefore, to get a better idea of how to genotype an unknown CNV spectrum in tumor or normal tissue samples that are very different molecularly, it makes sense to use at least two CNV detection methods. One of them, like ddPCR, should be able to quantitatively confirm the results of the other.

check for updates

Citation: Grebnev, P.A.; Meshkov, I.O.; Ershov, P.V.; Makhotenko, A.V.; Azarian, V.B.; Erokhina, M.V.; Galeta, A.A.; Zakubanskiy, A.V.; Shingalieva, O.S.; Tregubova, A.V.; et al. Benchmarking of Approaches for Gene Copy-Number Variation Analysis and Its Utility for Genetic Aberration Detection in High-Grade Serous Ovarian Carcinomas. *Cancers* **2024**, *16*, 3252. https://doi.org/10.3390/cancers16193252

Academic Editor: Edward J. Pavlik

Received: 4 September 2024
Revised: 20 September 2024
Accepted: 21 September 2024
Published: 24 September 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/).

Keywords: CoreExome DNA microarray; NanoString CNV panel; digital droplet PCR (ddPCR); benchmarking; gene copy number variation; ovarian cancer; PABAK; high-grade serous carcinoma; HGSC

1. Introduction

There are a group of genetic changes called gene copy number variations (CNVs) that show up at different rates in different parts of the genome of tumor cells [1,2]. Some estimates say that these changes, which mostly involve amplifications, insertions, deletions, and translocations, can rearrange DNA regions that are anywhere from 1 to 5 MB in size [3,4]. CNVs are elements of genomic instability and contribute to the survival of individual clones of neoplastic cells during tumor progression through disruption of transcriptional programs [5,6]. It is thought that CNVs found in tumors are more stable than expression biomarkers [7,8]. This means that studying the signatures of CNVs in different types of malignant neoplasms is a promising area for finding biomarkers that can tell us how a disease will progress. For instance, in a retrospective study [9], the genomic alterations correlated with the amount of circulating free DNA (cfDNA) and the frequency of occurrence of somatic allelic variants as well as the mutation tumor burden. The FDA approved the FoundationOne CDx Targeted Panel Clinical Assay (Foundation Medicine, Cambridge, MA, USA) for clinical use to look for problems in cfDNA taken from blood plasma and to find therapeutic CNV markers of how well the targeted drug therapy is working [10]. The Cancer Genome Atlas (TCGA) project looked at data from 33 different types of cancer and found typical tumor CNV signatures that were present in 97% of samples [2]. These signatures play a role in the development and growth of malignant neoplasms as well as chemotherapy resistance. This underlines the relevance of CNV detection to identify predictive or prognostic clinically relevant biomarkers. Another promising area of application of CNV data is the defining of molecular subtypes of solid tumors, which has particularly been reflected in several recent works [11–13]. However, it must be admitted that the results of similar studies are still far from being implemented in clinical practice.

Whole exome sequencing (WES) and whole genome sequencing (WGS) are NGS-based methods used to find gene number changes and larger chromosomal problems. In recent years, these kinds of analyses have been used a lot in preclinical studies to find CNV biomarkers linked to the development of many diseases [14–16]. Technological challenges for CNV detection in short-read sequencing and probe methods include extreme GC content and pseudogene homology, which make it difficult to unambiguously determine read alignment. The use of NGS panels is extremely convenient but requires high coverage in each position, special expensive analytical equipment, and adapted bioinformatics software for detecting CNVs in NGS 'big data'. In the second case, there are a lot of different algorithms [17,18], which make it harder to use NGS analysis and determine what CNV data mean in everyday clinical practice.

There are several modern methods for CNV analysis, both for detecting aberrations in a single gene of interest and in a panel of genes. These methods can be conditionally divided into two types: a complete study of the genome or a significant part of it, and methods for the detection of clinically relevant CNVs in a number of targeted genome regions. High-throughput methods such as whole genome sequencing [19], microarrays [20], and comparative genomic hybridization (CGH) [21] allow for the detection of CNVs in tens to thousands of genes. Certain techniques, like NanoString CNV panels [22,23], PCR analysis in its different forms, and fluorescence in-situ hybridization (FISH) [24], can be used to specifically find CNVs.

FISH technology is a traditional method for detecting gene deletions and amplifications, which is laborious and time-consuming but still remains the gold standard for the analysis of clinically significant CNVs of gross chromosomal segments in many molecular

genetic laboratories [25]. MLPA, in turn, is the gold standard for studying CNVs at gene size resolution [26].

The technology of microarrays is primarily designed to determine a large number of single nucleotide polymorphisms (SNPs). Illumina, for example, makes microarrays with different densities. The Infinium QC Array-24 has about 15,000 SNPs, while the Infinium Omni5-4 Kit has about 4.2 million SNPs [27]. In our study, CoreExome microarrays, containing about 550,000 SNPs and determining about 20 SNPs per one gene, were used [28]. However, several SNPs within a genome fragment determine its number of copies, potentially limiting the accuracy of such a determination. Thus, microarray technology makes it possible to detect CNVs within the exome with the expected moderate accuracy. Microarray technology is still the most common basic way to genotype a wide range of CNVs. However, not all platforms are equally good at finding different types of CNVs that are of interest [15]. Because of this, it is important to check the results of DNA microarray or NGS analysis using a different, highly repeatable method, like ddPCR [29]. This will lower the amount of false positive data when finding CNVs.

The NanoString nCounter v2 Cancer CN Assay (NanoString CNV) is a specially designed method for CNV detection in 87 genes associated with carcinogenesis [30]. It can be used to analyze a specific set of genes with high precision and, in particular, to validate CNV data obtained with hybridization technologies and microarrays [31].

Digital Droplet PCR (ddPCR) provides an accurate and absolute quantification of molecular targets with high sensitivity and reproducibility. The analytical sensitivity and specificity of the FISH and ddPCR are over 95% and 98%, respectively [32]. Recent studies have demonstrated the utility of different ddPCR versions for CNV detection [24,33].

The application of the above-mentioned methods has certain limitations. The results of high-throughput sequencing by synthesis, such as whole genome sequencing (WGS) and whole exome sequencing (WES), depend on the efficiency of the amplification of various genome regions and the length of reads, which limits the accuracy of regions with high sequence repeatability. The density distribution of polymorphisms in genome regions constrains the use of microarrays for SNP detection in CNV studies. Sequencing and analysis of microarrays make it possible to study a significant part of the genome, but they are highly dependent on bioinformatic methods that are used for CNV calling [34]. Techniques such as NanoString CNV assay and PCR analysis are expected to give higher accuracy of CNV detection in the regions studied, however their results depend on internal controls selected as regions with “normal” copy numbers [29]. In addition, these methods allow for the study of degraded DNA from FFPE blocks [30].

The methodical differences of the technologies on which CNV detection methods are based, along with the different degree of consistency between the results [35–38], actualize benchmarking studies on CNV detection, enabling this type of analysis to be conducted relatively quickly and with high accuracy.

According to the Global Cancer Observatory portal [39], ovarian cancer is the primary cause of death among gynecological cancers, with a five-year survival rate of 40–45% in the early stages (stages I–II) and 5–20% in the late stages (stages III–IV), respectively. High-grade ovarian serous carcinoma (HGSC) is the most aggressive subtype of ovarian cancer and is often diagnosed at advanced stages [40–42].

A bioinformatic analysis of TCGA datasets showed that ovarian cancer samples had the highest rates of CNV burden. There was also a lot of variation in the number of somatic copy number alterations (SCNA events) that happened in each sample [3]. Several studies [43,44] have demonstrated the identification of tumor-associated CNV biomarkers with prognostic or predictive value for ovarian cancer. Thus, the detection of CNVs in ovarian cancer samples is a variant of the analysis of a valuable way to discover specific genetic aberrations with potential clinical significance.

In this study, we want to compare how well three different methods find CNVs in paired samples of conditionally normal and ovarian cancer tumor tissues. We will also use statistics to see how consistent the results are. To detect CNVs, Infinium CoreExome

microarrays (Illumina, Inc., San Diego, CA USA) and an nCounter v2 Cancer CN Assay panel (NanoString Technologies Inc., Seattle, WA, USA) were used. Digital droplet PCR QX200 (Bio-Rad Laboratories Inc., Hercules, CA, USA) was applied to verify the results obtained by these two methods.

2. Materials and Methods

2.1. Clinical Samples and Their Characterization

Thirteen patients were recruited to participate in a scientific study in the Federal State Budgetary Institution "National Medical Research Center for Obstetrics, Gynecology and Perinatology named after academician V.I. Kulakov" the Ministry of Healthcare of the Russian Federation. Of the 13 patients, 12 had a histologically confirmed diagnosis of high-grade serous carcinoma, and one had endometrioid carcinoma. For all the patients, sets of tumor samples and corresponding normal ovarian tissue were obtained. The samples were cryopreserved and then used for DNA extraction and molecular genetic studies. The characterization of clinical cases by disease stage, TNM classification, and immunohistochemical (IHC) results and parameters for assessing IHC markers are presented in Tables A1 and A2 (Appendix A), respectively.

2.2. DNA Isolation

DNA for analysis was isolated from samples of fresh frozen tissues using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). DNA quantification was performed using the Promega QuantiFluor ONE dsDNA kit (Promega Corporation, Madison, WI, USA) with a Quantus fluorimeter (Promega Corporation, Madison, WI, USA). DNA purity was measured using NanoDrop 8000 (Thermo Fisher Scientific Inc., Waltham, MA, USA). For all the isolated DNA samples, the ratio of absorbance indices 260/280 was 1.8–2.0, the ratio 260/230 was 2.0–2.2.

2.3. CoreExome Microarray-Based Genotyping

Microarray-based genotyping was performed using the Infinium CoreExome-24 v1.3 BeadChip kit (Illumina, Inc., San Diego, CA, USA). The samples were processed using the Freedom EVO automated workstation (Tecan Group Ltd., Männedorf, Switzerland). Microarrays were scanned on the Illumina iScan System (Illumina, Inc., San Diego, CA, USA) with the AutoLoader 2.x autoloading module.

After scanning using the gtc2vcf plugin for bcftools V1.20 [45], vcf files were obtained from IDAT files. CNV was determined using the cnv plugin for bcftools [34]. The results were processed using R packages [46]. Data on the presence of CNV were also obtained using the cnvPartition CNV Analysis Plugin V3.2.0 and CNV Region Report Plug-in for Illumina GenomeStudio 2.0 in accordance with the manufacturer's instructions.

2.4. Estimation of the Gene Copy Number Using a Panel nCounter v2 Cancer CN Assay (NanoString CNV)

The nCounter v2 Cancer CN Assay Kit for the nCounter FLEX Analysis System (NanoString Technologies Inc., Seattle, WA, USA) was used to estimate the number of gene copies. DNA was fragmented with the AluI enzyme and hybridized with the probes according to the manufacturer's protocol MAN-10093-01 (NanoString Technologies Inc., Seattle, WA, USA). A total of 300 ng of DNA was used for hybridization; the incubation time was 18 h at 65 °C. The samples were then placed in the nCounter Prep Station for post-hybridization processing and immobilization on the cartridge.

The cartridge with the applied samples was transferred to the nCounter Digital Analyzer for scanning. A Prep Station and Digital Analyzer were used according to the user guide MAN-C0035 nCounter Analysis System MAX/FLEX. The nCounter data were processed using the nSolver 4.0 Analysis Software (NanoString Technologies Inc., Seattle, WA, USA) in accordance with the user manual MAN-C0019-08. Using the nSolver software, the signals from the samples were normalized, and data were obtained on the number of

gene copies in the studied samples relative to the standard sample from the nCounter v2 Cancer CN kit.

2.5. PCR Analysis

The reference sequences of 15 genes of interest and 2 reference genes were found in the NCBI USA GenBank database [47] (https://www.ncbi.nlm.nih.gov/genbank/ (accessed on 30 August 2024)). The ALB and RPP30 genes, which are commonly used for the quantification of human genomic DNA [48], were selected as references. The primers and probes were designed according to the basic requirements for primers and TaqMan probes [49].

The selected primer and probe sequences were then checked in silico for specificity of annealing using the NCBI Primer BLAST online service (https://www.ncbi.nlm.nih.gov/tools/primer-blast/ (accessed on 30 August 2024)) [50]. The primer and probe sequences are shown in Table A3 (Appendix B). To prepare the TaqMan probes, FAM (fluorescein) or R6G (rhodamine 6G) was incorporated as a fluorophore at the 5′ end of a specific primer, and BHQ1 (Black Hole Quencher 1) was used as a fluorescence quencher at the 3′ end.

The melting and annealing temperatures of the primers and TaqMan probes were calculated using the online tool OligoAnalyzer IDT [51] (https://www.idtdna.com/pages/tools/oligoanalyzer (accessed on 30 August 2024)). This tool was also used to estimate the probability of the formation of secondary structures such as hairpins, homo- and heterodimers.

In-vitro functioning of the selected primer pairs and optimal amplification conditions were checked by symmetric gradient PCR on a CFX96 Touch real-time platform (Bio-Rad Laboratories Inc., Hercules, CA, USA). All the primer pairs were validated using melting curves. The size of the target fragments obtained by PCR analysis and the presence/absence of non-specific products were evaluated using the Agilent 4200 TapeStation system (Agilent Technologies, Inc., Santa Clara, CA, USA).

Digital droplet PCR was performed using the QX200 droplet digital PCR system (Bio-Rad Laboratories, Inc., USA). The ddPCR reaction mixture with probes consisted of 10 µL of 2x ddPCR Supermix for Probes (Bio-Rad Laboratories Inc., Hercules, CA, USA), 10 ng of tested DNA, 900 nM primers, and 250 nM hydrolysis probes for the target and reference genes in a final volume of 20 µL. The ddPCR reaction mixture with EvaGreen consisted of 10 µL of 2× ddPCR EvaGreen Supermix (Bio-Rad Laboratories Inc., Hercules, CA, USA), 10 ng of tested DNA, and 100 nM primers for target and reference genes in a final volume of 20 µL. The entire reaction mixture was loaded into the disposable plastic cartridge (Bio-Rad Laboratories, Inc., USA) along with 70 µL of droplet generation oil (Bio-Rad Laboratories Inc., Hercules, CA, USA) and placed into the droplet generator (Bio-Rad Laboratories Inc., Hercules, CA, USA). After processing, the droplets generated from each sample were transferred to a 96-well PCR plate. After PCR, the plate was loaded into the droplet reader (Bio-Rad Laboratories Inc., Hercules, CA, USA); the acquired data were analyzed using QuantaSoft software V1.0 (Bio-Rad Laboratories, Inc., USA). PCR amplification was carried out on a C1000 Touch thermal cycler (Bio-Rad Laboratories Inc., Hercules, CA USA) using the following conditions described in Tables A4 and A5 (Appendix C). After the end of PCR, the plate was loaded into a droplet reader, and the data obtained were analyzed using the QuantaSoft software.

2.6. Biostatistics Data Analysis

According to the data yielded by the microarray platform, the variable of interest is divided into three categories: deletion, normal copy number state, and amplification, irrespective of the computational protocol (CNVpartition or bcftools). On the same note, NanoString CNV technology or PCR-based methods yield quantitative gene copy number values.

The conversion rule of copy number values from quantitative to integer for the NanoString CNV and PCR-analysis results was defined in keeping with the need to assess the consistency between quantitative and integer variables. It should be noted that often it is difficult to obtain the integer values of the gene copy number in the cancer tissue. This

is caused by intratumor heterogeneity, which results in the existence of several cellular subgroups per tissue sample with differing copy number changes in each one [52].

Since the results were compared with a method that produces only integer values (microarrays), the ranges of this translation were determined for this task in accordance with the developer's recommendations. Data Analysis Guidelines for Copy Number Variation (CNV) recommends considering all values in the range ± 0.4 as the integer, while other values that exceed such a range should be studied on a separate basis or excluded from the analysis [53].

Since it was necessary to compare the results in pairs and each individual value of the copy number played a smaller role, the following rules of conversion were defined: copy number values between 1.5 and 2.5, inclusively, were considered as 2 (conditionally normal gene copy number), and values lower than 1.5 and greater than 2.5 were accepted as the deletion and amplification, respectively. These rules were used to convert the values obtained both by the NanoString CNV method and by PCR-based techniques.

To assess the reliability of consistency between integer or even qualitative variables obtained by two diagnostic methods, we calculated the prevalence-adjusted and bias-adjusted Kappa coefficient (PABAK) [54,55] by the epiR library written for the R programming language (epiR version 2.0.50) [56]. The PABAK value close to zero indicates a chance of consistency, while the PABAK value close to -1 means that nearly all the results of the first diagnostic method are in direct contradiction with the other one and vice versa. The PABAK value close to 1 indicates perfect consistency.

If the results of two diagnostic methods were quantitative variables, then Passing–Bablok regression was used to analyze the consistency. The reliability of consistency between quantitative variables had been assessed by; regression using the mcr library written for the R programming language [57]. This method is one of the most robust linear regression algorithms [58] and can be described as the equation:

$$y_i = k * x_i + b$$

where x_i and y_i are the gene copy number values for i-th observation that were obtained by the first and second method, respectively. The slope coefficient k corresponds to a proportional bias and the intercept b corresponds to a constant bias. The presence of value one inside the 95% confidence interval for k means no proportional difference between two methods and the presence of value zero inside the 95% confidence interval for b means no constant difference between two methods.

The discrepancy between values obtained by both methods could be unacceptably large even without significant proportional and constant biases [59]. It should be noted that consistency and interchangeability between two methods is influenced by the scatter of points around the line of complete consistency, i.e., $y = x$. Therefore, in addition to the Passing–Bablok regression, the following measure was calculated:

$$d_i = \frac{|y_i - x_i|}{\max(y_i,\ x_i)}$$

where x_i and y_i are the gene copy number values that were obtained by both methods.

To evaluate the statistical significance of the difference between the copy number profiles of normal and tumor tissue, we used the Wilcoxon test for the quantitatively paired samples. The same pair of samples had been studied using two or more methods of copy number analysis, so for such a pair, the respective number of statistical testing procedures should be performed. The harmonic mean of the resulting set of p-values was calculated. To counteract the multiple testing problem, all the obtained harmonic means of p-values were adjusted according to the Holm method.

The study utilized R programming language (version 4.2.1) and RStudio IDE (version 2022.02.3.492).

3. Results

3.1. Comparison of Results of Gene Copy-Number Assessment

In the first stage of the study, the gene copy number was determined using the NanoString CNV panel and microarrays. The CNVs on microarrays were determined by the same coordinates, which are targeted by labels in the NanoString CNV panel. Figure A2 (Appendix G) distinguishes at least three groups of genes: MAGI3, PDGFRA, NKX2-1, MET, and KDR are genes with positive consistency (PABAK values closer to one); CRKL, BBC3, BCL2L1, AR, and CDK4 are genes with medium consistency (PABAK values close to zero); and ERBB2, BRCA1, MAPK7, FGFR1, and ITGB4 are genes with negative consistency (PABAK values less than 0).

After analyzing the consistency of the CNV detection results using NanoString CNV panels and microarrays in the second stage of the study, we selected gene panel no. 1 (Table A6 (Appendix D)) for the ddPCR verification of the identified CNVs. The specifics of the primer design and ddPCR settings are described in the Materials and Methods section, as well as in Appendix C (Tables A4 and A5). The research plan is presented in Table A6 (Appendix D), and the results of the experiments are shown in Figure 1 (see also Figure A2 (Appendix G)).

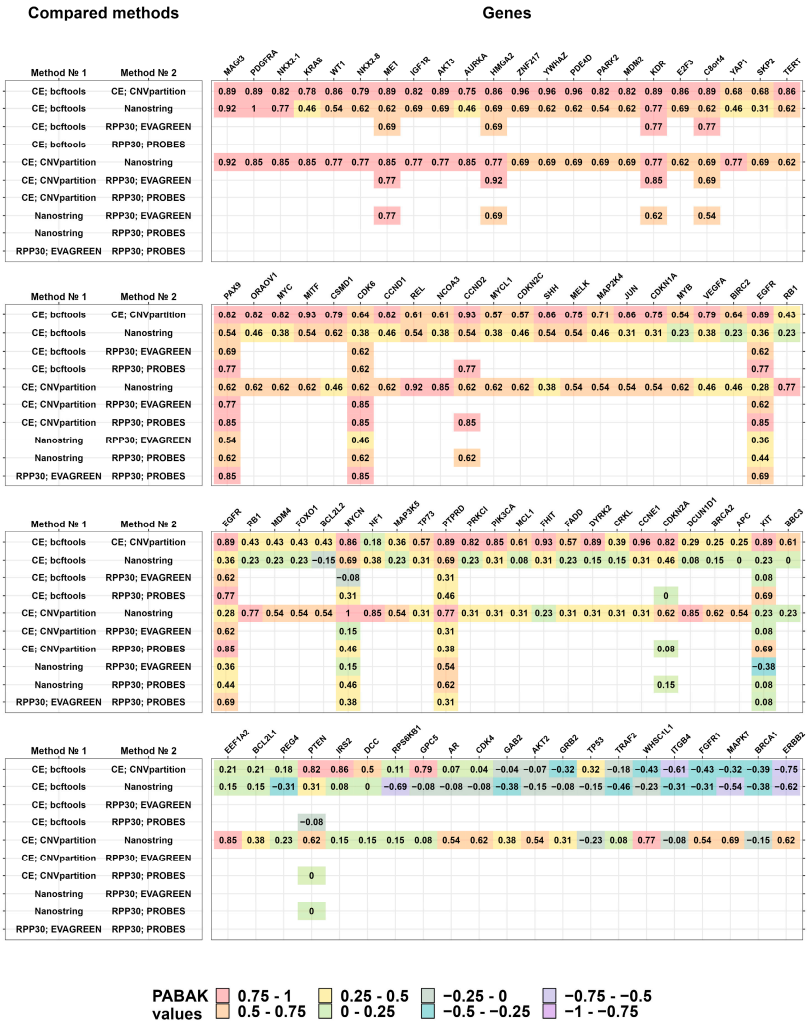


Figure 1. The consistency between the methods of copy number identification measured by prevalence-adjusted and bias-adjusted Kappa coefficients (PABAK). The three wide panels are the tables whose rows correspond to the compared methods and whose columns correspond to the studied genes. The sort order is decreasing according to the median PABAK value per gene, from highest to lowest.

3.2. CNV Detection by Digital Droplet PCR

The next step was to assess the copy number of the target genes relative to the reference genes RPP30 and ALB using Bio-Rad QX200 digital droplet PCR. The ddPCR results were used to evaluate the performance of the primer sets. To accomplish this, the gene copy number values in normal tissue were evaluated. The median value among patients falling within the range of 1.5 to 2.5 was taken as the normal value of the copy number (Figure 2). We detected gene copy number variations using four different versions of PCR analysis: two reference genes and two different PCR protocols. If the median copy number values of a single gene in normal tissue samples were less than 1.5 and more than 2.5, then such a group of observations was considered abnormal and excluded from further consideration (Figure 2). Additionally, we decided not to consider the results of the gene copy number assessment, which used albumin (ALB) as a reference gene, because a significant portion of the gene copy number values in normal tissue samples fall outside the range of 1.5 to 2.5.

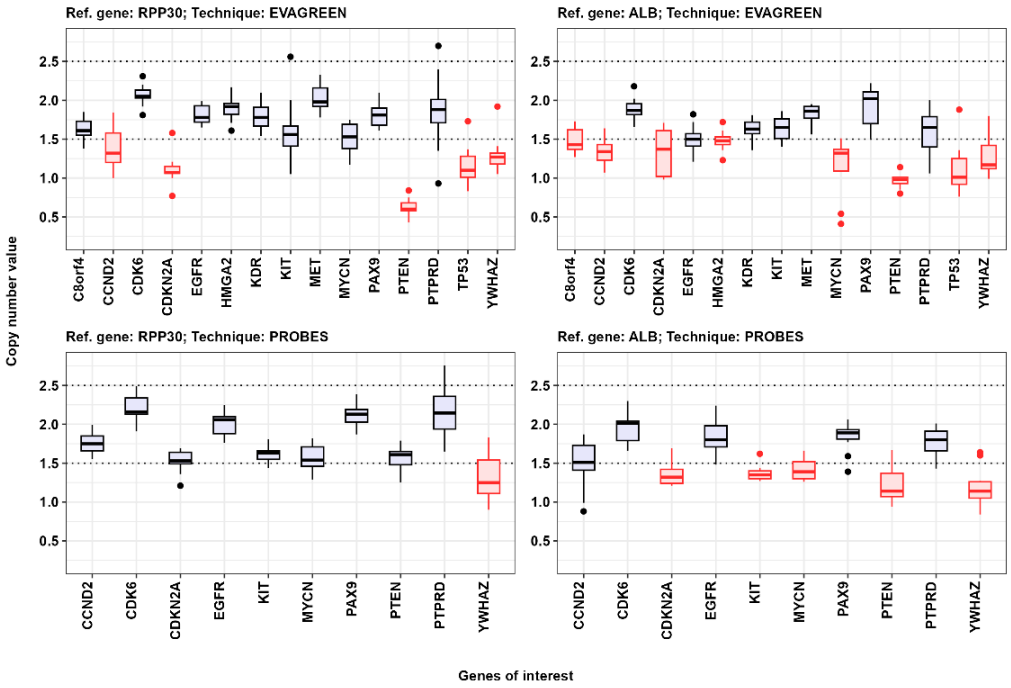


Figure 2. Preliminary descriptive statistics of PCR-analysis results for normal tissue. Copy number values are presented for the genes of interest. The red color depicts groups of observations that were excluded from further consideration.

3.3. The Consistency between Three Methods of CNV Detection

Since the results of the microarray copy number estimation are in an integer format, we began by comparing the NanoString CNV panel and PCR results, which enabled us to quantify the copy number of the genes. Table 1 presents the consistency analysis of CNV detection using the NanoString CNV panel and ddPCR analysis. Only the CCND2 gene showed a statistically significant change in copy number variation. The NanoString CNV panel and PCR analysis detected an increase in the copy number values of this gene in tumor tissues compared to normal tissues. Despite the small sample size, when comparing the results of PCR analysis (tumor/normal), statistically significant differences in the gene copy number were obtained; when comparing the results of the NanoString CNV panel (tumor/normal), statistical differences were noted at the trend level of statistical significance ($0.05 < p < 0.1$). As for the other genes in Table 1, there were no statistically significant changes in the gene copy number variations.

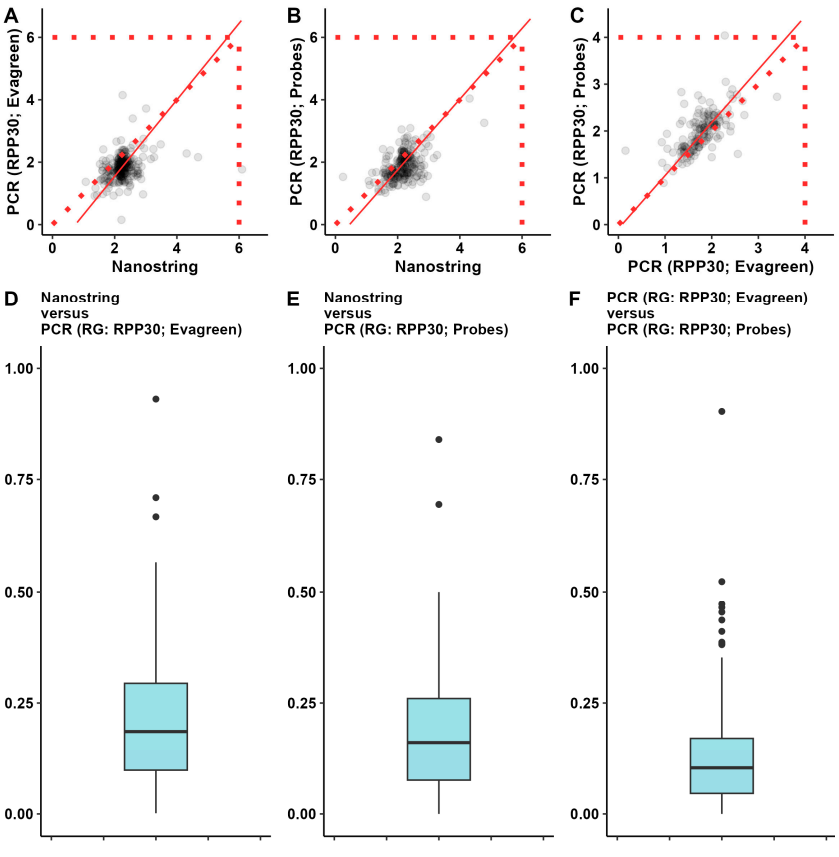
Figure 3 presents three pairwise comparisons of the CNV detection results. The dashed sloping line in each of the graphs indicates the ideal situation in which the application of

both methods on the same set of samples leads to identical results. The solid sloping line represents the resulting regression model. In the table below, there are data indicating the absence (x) or the presence (o) of a statistically significant proportional (k) and systematic errors (b). The results of calculating the magnitude of the relative discrepancy d_i are displayed. We determined the value of d_i for all samples considered in the study, allowing us to judge the consistency of the two methods based on the characteristics of the formed sample. The closer the median, first, and third quartiles of the sample d to zero, the better the consistency of the two methods.

Table 1. Data on copy number variations in tumor and normal tissues *.

Genes	*p*-Value PCR EvaGreen	*p*-Value PCR Probes	*p*-Value NanoString CNV	*p*-Value	*p*-Value (Corrected)
C8orf4	0.972	No data	0.807	0.882	1
CCND2	No data	0.003	0.092	0.006	0.042
CDK6	0.497	0.216	1	0.393	1
HMGA2	0.505	No data	0.168	0.252	1
KDR	0.685	No data	0.363	0.475	1
MET	0.893	No data	0.421	0.573	1
PAX9	0.685	0.542	0.168	0.324	1

* the full description of the statistics is shown in Table A7 (Appendix E).



Method 1	Method 2	Proportional difference (k) with 95% CI	Constant difference (b) with 95% CI
Nanostring	RPP30; EVAGREEN	1.23 [0.9 - 1.55]	**−0.9 [−1.64 - −0.17]**
Nanostring	RPP30; PROBES	1.13 [0.92 - 1.35]	**−0.51 [−0.99 - −0.04]**
RPP30; EVAGREEN	RPP30; PROBES	1.13 [0.97 - 1.3]	−0.09 [−0.42 - 0.25]

Figure 3. The consistency between the methods of copy number identification measured by Passing–Bablock regression and by the magnitude of the relative discrepancy d_i. Three methods were considered:

the NanoString CNV panel, the PCR-based technique with TaqMan probes, and the PCR-based technique with EvaGreen dye. RPP30 has been used for both PCR analyses as a reference gene. Panels (**A**–**C**) are the scatterplots depicting the results of the Passing–Bablock regression. The panels (**D**–**F**) are the boxplots representing the distribution of relative discrepancy d_i. The lower the d_i is to zero, the better the consistency of the two methods. The coefficients of the Passing–Bablock regression models are presented in the bottom table, along with their respective 95% confidence intervals. The shaded cells represent the observed constant differences between the compared results of the two methods. The abbreviation RG means «reference gene».

As you can see in Figure 3, the two types of PCR analysis—the one using the EvaGreen intercalating dye and the one using TaqMan probes (Probes)—are the most consistent and have the smallest amount of difference between them. The NanoString CNV method is characterized by low consistency with methods for detecting copy number variations using PCR analysis. A pairwise comparison of the NanoString CNV panel with the PCR analysis reveals a systematic error. The gene copy numbers obtained by the PCR analysis were generally lower than those obtained by the NanoString CNV panel.

For further pairwise comparisons between three lab approaches and their variations, quantitative values were converted to categorical values according to the scheme described in the Materials and Methods section. A graphical panoramic view of consistency between the results of the CNV detection by three different methods is shown in Figure 1, and more detailed data are presented in Figure A2 (Appendix G). As shown in Figure 1, the results obtained by the two versions of CoreExome microarrays are the most consistent with the bcftools and CNVpartition calling protocols, respectively. In addition, it is noticeable that the results of the PCR analysis are in good consistency with the results of the CoreExome microarrays using the CNVpartition calling protocol. On the contrary, the results obtained by the NanoString CNV panel are satisfactory and consistent with those obtained by all other methods. For pairwise comparisons, the MET, HMGA2, KDR, PAX9, CDK6, and CCND2 genes had the best consistency values. This was true no matter what method was used to find CNVs.

3.4. Results of CNV Detection in Ovarian Cancer Samples Based on ddPCR Data

It follows from Figure 4 that the total number of detected deletions is ≈1.6 (25/16) times higher than the number of amplifications of nine target genes in 12 HGSC samples. The deletions were mostly found in the CDKN2A (cyclin-dependent kinase inhibitor 2A) and PTEN (phosphatase and tensin homolog) genes, with 7 and 8 events per 12 samples, respectively. Copy number amplifications were found mainly in CCND2 (cyclin D2) and PTPRD (protein tyrosine phosphatase receptor type D) genes—3 and 4 events per 12 samples for each gene, respectively (Figure 4).

In tumor samples, the CNVs of other genes were irregularly distributed. For the TCGA serous cystadenocarcinoma ovarian dataset (n=617), cBioPortal data show that 5.9% of CDKN2A genes and 5.7% of PTEN genes have copy number variations. Given that the mutant TP53 gene (which codes for the tumor suppressor p53) is found in about 87% of all ovarian cancer samples, it is likely that the CDKN2A and PTEN genes are also changed. For instance, we found that deep deletions of the CDKN2A gene correlate with TP53 mutational signatures and worse survival in breast cancer subjects [60]. Thus, we attempted to search for the potential associations between the most frequent CNVs of CDKN2A, PTEN, CCND2, and PTPRD genes with clinical data (Tables A1 and A2, Appendix A). For this, a relatively 'homogeneous' subset of nine HGSC samples (SN30, SN32, SN36, SN38, SN44, SN46, SN50, SN54, SN56), having a similar tumor stage, TNM classification (T3N0(1)M0(1)), as well as Ki-67 and p53 IHC-staining patterns (Tables A1 and A2, Appendix A), were analyzed. Notably, deletions of the CDKN2A and PTEN genes were observed in 3 patients (SN38, SN50, and SN54) with late stage cancer (IVb) and metastasis (T3NxM1), which may warrant further investigation. Apparently, the small sample size did not allow us to find more associations between the number of detected CNVs of target genes and other clinical data; for example, the number of CNVs in stage IIIc ovarian cancer cases ranged

from 0 to 4. This was similar to the number of CNVs seen in samples from stages Ia and IIa (Figure 4), which were 2 and 3. Further, there was no correlation (Spearman's rank correlation coefficient = 0.16, p-value > 0.1) between the total number of CNVs of all nine genes under consideration (Figure 4) and the ages of 12 subjects with HGSC.



Figure 4. Changes in copy number values of genes of pane1 according to the PCR analysis. We used a PCR-based technique with TaqMan probes and RPP30 as a reference gene. The horizontal axis corresponds to the patient identifier, and the vertical axis to the genes of interest.

4. Discussion

In general, works on the benchmarking of methods are interesting since it is known that, depending on the source of the data acquisition, their consistency can vary significantly. This was revealed by a pairwise comparison of information on gene copy loss extracted from the Hapmap Project and 1000 Genome Project repositories (less than 30% overlap) [61]. In our study, the NanoString CNV panel was used to genotype CNVs in tumor samples obtained from patients with high-grade serous carcinoma. This commercial product has recently appeared on the biotechnological market and is suitable for targeted multiplex analysis of gene abnormalities associated with oncogenic molecular pathways, so the researchers' interest is directed at clarifying the limits of its application. Recently, it was shown that comparing the results of the NanoString CNV panel (20-gene signature for copy number value detection) with the best ways to measure gene expression, like IHC analysis and in-situ hybridization (FISH), showed a good level of agreement (κ-value = 0.35) in the HER2 copy number variation from biopsy gastric cancer tissues (FFPE blocks) with the results of IHC/FISH, which had sensitivity and specificity values of 66.7% and 85.2%, respectively [62]. Data on benchmarking the results of the study using the NanoString CNV panel, DNA-microarray technologies, or ddPCR to study the CNV profile of HGSC ovarian cancer is not yet available, highlighting the novelty of the obtained results.

In the study using CoreExome microarrays, the same raw data were processed using bcftools and CNVpartition tools. The median PABAK value was 0.75 (Figure 5), which gives us an idea of what the theoretical limit of consistency might be for these comparisons. This is because we should obtain more consistent data when we compare different ways of processing the same raw data than when we compare different biological methods.

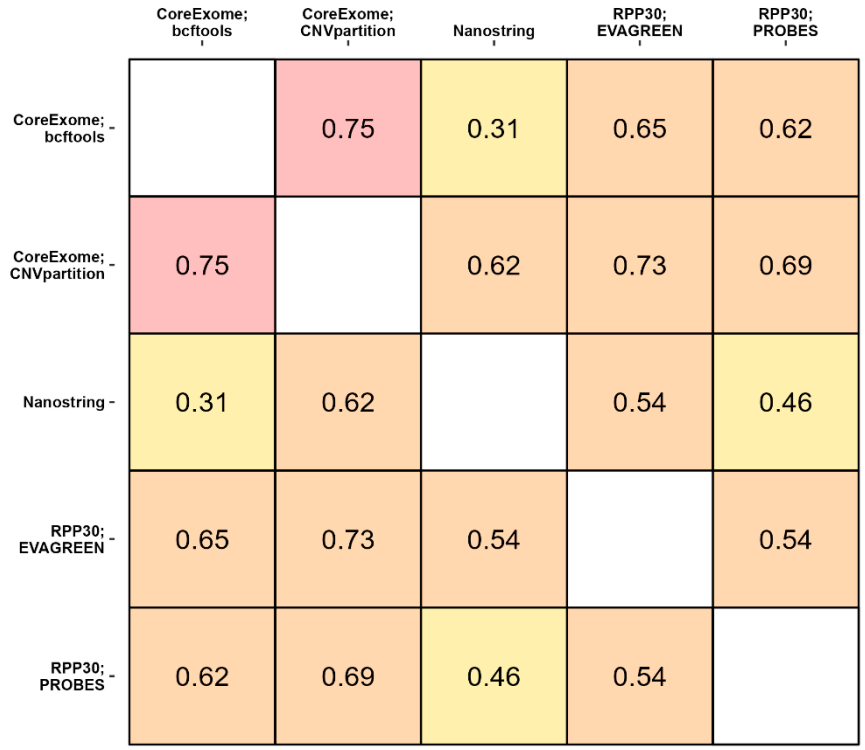


Figure 5. Pairwise consistency between the methods of copy number identification. The prevalence-adjusted and bias-adjusted Kappa coefficient (PABAK) has been used as a measure of consistency.

The results of the copy number assessment using CNVpartition show better consistency with each of the other methods than the values obtained using bcftools. The gene copy number detection using the NanoString CNV panel shows good consistency rates, except for the comparison with bcftools and ddPCR analysis using probes. The results most comparable to all other methods were shown by the CoreExome-CNV partition and ddPCR with intercalating dye. Most comparisons fall within the PABAK value range of 0.61–0.80, which is considered a substantial agreement. In our opinion, values below 0.61 indicate an insufficient level of result consistency. Ideally, the consistency level should exceed PABAK 0.8. In our opinion, this result is extremely difficult to achieve using off-the-shelf panels, but this level of consistency appears realistic when developing customized labels for targeted approaches such as ddPCR.

Copy number assessment can be inconsistent for a number of reasons, such as the genetic structure of the area being studied, the design of the ddPCR probe, and the way the reaction is set up. Another factor we take into account is the small number of comparisons, i.e., ten genes. The consistency scores between samples analyzed with the NanoString CNV panel and various CoreExome microarray processing tools show a significant difference: the median value of PABAK differs by two-fold (0.31) when compared with bcftools and by 0.62 when compared with the CNV partition tool. Relatively low consistency between the two ddPCR methods can be expected due to the small number of compared genes, and as a result, each individual low consistency value contributes more to the median value. It should be noted that the study was carried out on real tumor and normal tissue samples, meaning that there was no known unambiguous copy number value for each gene, in contrast to studies on well-studied cell lines.

Researchers can use the results of gene copy number variations in ovarian cancer tissue samples not only to characterize the genetic profile of tumors, but also to determine the potential prognostic significance of the identified CNVs. In particular, previous studies found that risk assessment based on CNV detection was more valid in predicting overall survival in ovarian serous adenocarcinoma. Thus, the ten-year CNV risk score had an AUC value of 0.747 [43].

Optimization of the gene panel for performing CNV genotyping and data interpretation is critical due to the high initial heterogeneity of tumors from different subjects in the cohort. For instance, a study [63] selected a panel of 34 genes based on tumor heterogeneity to catalog CNVs in a cohort of 96 patients with HGSC. The nCounter v2 Cancer CN Assay panel consists of 87 target genes and is intended for CNV genotyping of tumors with various profiles of aberrations in the PIK3CA, AKT, PTEN, BRCA, ERBB2, and MYC genes. The NanoString CNV panel can be used to test people with different types of cancer because it covers 55% of the genes where mutations are most common, according to data from the Cancer Gene Census [64].

The frequencies of occurrence (>2%) of amplifications and deletions of the genes of the NanoString CNV panel were observed in a cohort of patients with ovary serous carcinoma (568 samples) [65], where 23 and 4 genes, respectively, accounted for 31% of all genes in the panel (27 of 87 genes). Thus, it can be concluded that this NanoString CNV panel is suitable for detecting CNVs in ovarian cancer samples. Our study's statistical analysis revealed that seven target genes (MET, HMGA2, KDR, C8orf4 (TCIM), PAX9, CDK6, CCND2) (a gene panel No. 2) exhibited the highest consistency between the results of CNV detection by different methods (DNA microarray, NanoString CNV, and ddPCR analysis). The biological features of these genes as well as their potential prognostic value in ovarian cancer are discussed in Appendix F, which includes a text subsection, Table A8, Figure A1, and the following literature references [43,65–87].

Our benchmarking study focused on assessing amplifications and deletions as two major categories of CNVs. At the same time, the repertoire of CNVs is represented by a wider spectrum of subcategories due to the high inter- and/or intra-patient tumor heterogeneity and complex nature of CNV combinations [2]. In the current study, we did not differentiate such subcategories by the experimental and bioinformatics methods used, but some of them are mentioned below. The prevalence of CNVs as indicators of genome instability was investigated by Marco and colleagues [88], who identified 201 chromosomal regions and 3300 genes subjected to copy number (CN) alterations in patients with serous ovarian cancer. Out of 576 genes, CN gains comprised 530 low-level and 46 high-level CN gains, while CN loss affected 2724 genes (2454 heterozygous and 270 homozygous deletions, respectively), indicating a 5-fold higher frequency of CN losses compared to CN gains. Cheng and colleagues examined the distribution of homozygous deletions in different cancer types [89]. It was shown for ovarian cancer that deletions ranged from 0 to 6 per case and had a median fragment length of 207 kb per case. Deletion peaks occurred primarily in the tumor suppressor gene CDKN2A, followed by RB1 (RB transcriptional corepressor 1) and then the PTEN gene. This fits with the fact that about 70% of the CDKN2A and PTEN genes found in HGSC samples were deleted, which is higher than the rate of deletion seen in other genes in our study.

In type II high-grade serous ovarian cancer, there is genetic instability and a lot of DNA CN gains or losses. Almost all tumors also have TP53 mutations [90]. Typically, loss of heterozygosity (LoH) of the wild-type allele of tumor suppressor genes (TSGs) can also be caused by CN losses. On average, across all cancers, 16% of genes underwent LoH, with a median of about 30% in ovarian cancer. Rates of LoH were no lower for cell-essential genes, which also included some TSGs, relative to the rest of the genome; tumors harbored an average of 189 essential genes with LoH [91]. Given that the frequency of LoH in tumors is one order of magnitude higher than that of homozygous deletions [88], we can use a specific pattern of LoH to predict disease prognosis. For instance, two relatively small regions of chr 19 (8.0–8.8 and 51.5–53.0 Mbp regions, which include 19 and 37 genes, respectively) exhibited the most significant differences in LoH. Functional loss of the TP53 gene due to mutations in chr 17 found in 32% patients and LoH in chr 19 were associated with serous ovarian cancer recurrence [92]. Also, it is known that turning off both alleles (biallelic inactivation) of TSGs makes ovarian cancer grow and spread a lot [93,94]. Ryland and colleagues [95] suggested several models of non-random accumulation of LoH in ovarian cancer genomes. Among these models are: the classic

two-hit hypothesis, which involves high-frequency biallelic genetic inactivation of TSG; the epigenetic two-hit hypothesis, which involves high-frequency biallelic inactivation through methylation and LoH; the multiple alternate gene biallelic inactivation, which involves low-frequency gene disruption; the hap-lo-insufficiency model, which involves single copy gene disruption; and the modified two-hit hypothesis, which involves a reduction to homozygosity of predisposition alleles. These mean that there are more than one, possibly overlapping, ways in which different types of CN changes can lead to the development of tumorigenic traits. In addition, it is worth noting that the novel version of inter-laboratory classification concordance using the American College of Medical Genetics and Genomics (ACMG) and Clinical Genome Resources (ClinGen) CNV scoring metrics system was adapted for potential clinical use, with an overall complete concordance of up to 85% across laboratories [96].

5. Conclusions

The analysis of statistical consistency between the results of three methods of CNV genotyping allows us to make the following conclusions: (a) the use of one method cannot be considered sufficient to unambiguously detect CNVs; (b) data obtained by panoramic detection of CNVs using microarrays or NanoString CNV panels (or other methods of broad genome analysis for the presence of structural alterations) should be verified with a point quantitative analysis of selected CNVs, for example, with digital droplet PCR; and (c) most pairwise comparisons gave a PABAK value of >0.6, which is considered acceptable [55] and indicates good consistency between the results of the methods used; however, for a more detailed study of the CNVs of a single gene, it is reasonable to rely on consensus results obtained by at least two methods.

Author Contributions: Conceptualization, E.A.S. and A.S.M.; software, I.O.M. and P.A.G.; investigation, M.V.E., A.A.G., A.V.Z., O.S.S., A.V.T. and A.V.A.; resources, V.S.Y., V.V.M., S.M.Y., A.A.K. and V.I.S.; writing—original draft preparation, P.V.E., I.O.M. and P.A.G.; writing—review and editing, A.V.M. and V.B.A.; visualization, I.O.M.; supervision, E.A.S. and A.S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This study received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Ethics Committee of "National Medical Research Center for Obstetrics, Gynecology and Perinatology Named After Academician V.I. Kulakov" (protocol code № 4, 29.04.2021).

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

Data Availability Statement: The NanoString CNV panel raw data results are available under NCBI GEO accession GSE244329. Our code is available at Figshare.com depository: (https://figshare.com/s/b4c44d9f9e3ce7e2a74d (accessed on 20 September 2024)).

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Clinical characteristics of patients with ovarian cancer included in the study.

Patient ID	Age	Stage	TNM Classification			Histological Subtype	Immunohistochemical Markers			
			T	N	M		Ki-67	p53	PAX8	WT1
SN26	45	IIc	2c	0	0	endometrioid carcinoma	high	mut type (+++)	0	2
SN28	58	Ia	1a	0	0	HGSC	high	mut type (−)	1	3

Table A1. *Cont.*

Patient ID	Age	Stage	TNM Classification			Histological Subtype	Immunohistochemical Markers			
			T	N	M		Ki-67	p53	PAX8	WT1
SN42	47	IIIc	3c	x	0	HGSC	high	mut type (+++)	1	3
SN56	43	IIIc	3c	x	0	HGSC	high	mut type (+++)	3	3
SN30	54	IIIc	3c	1	0	HGSC	high	mut type (+++)	1	2
SN32	60	IIIc	3c	0	0	HGSC	high	mut type (+++)	3	0
SN36	60	IIIa	3a	1	0	HGSC	high	mut type (+++)	1	1
SN38	49	IVb	3c	x	1	HGSC	high	mut type (+++)	1	1
SN44	41	IIIc	3c	x	0	HGSC	high	mut type (+++)	1	2
SN46	38	IIIc	3c	1	0	HGSC	high	mut type (−)	3	1
SN50	44	IVb	3c	x	1	HGSC	high	mut type (+++)	2	3
SN52	62	IIa	2a	0	0	HGSC	high	mut type (−)	1	1
SN54	66	IIIc	3c	x	1	HGSC	high	mut type (+++)	1	0

Abbreviations: HGSC—high-grade serous carcinoma); TNM classification of tumors (Tumor–Node–Metastasis); Ki-67—a marker of proliferation). Ki-67 (Ki-67 expression makes it possible to detect tumor cells in the active phase of the cell cycle); p53 marker: strong and negative immunohistochemical staining, mut (+++) and mut (−), respectively; PAX8—paired-box 8 protein (transcription factor); WT1—Wilms tumor protein (transcription factor).

Table A2. The description of immunohistochemical staining gradations used in the histological analysis of tumor samples performed at FSBI «National medical research center for obstetrics, gynecology and perinatology named after academician V.I. Kulakov» Ministry of Healthcare of the Russian Federation.

Immunohistochemical Markers	Gradations	
PAX8 and WT1	negative	0
	weak	1
	moderate	2
	strong	3
p53	strong	mut (+++)
	negative	mut (−)
	weak	wild type (+)
Ki-67 index	>50%	high
	15–49%	moderate
	<15%	low

Appendix B

Table A3. Primers and probes used for PCR analysis.

Gene	PCR-Product Length, bp	Tm, °C	Primer	Sequence 5′–3′
C8orf4 (TCIM)	191	62.7	Forward	GAAAACTCTTCCGTCCCTGC
		63.7	Reverse	CCTCCTCCCTTCCCCTATCT
		67.9	Probe	[FAM]CACTTTGTCAGGCCCTAGGA CTTAAATCG[BHQ1]

Table A3. *Cont.*

Gene	PCR-Product Length, bp	Tm, °C	Primer	Sequence 5′–3′
CCND2	164	62.7	Forward	CAGCCGAGCAGTGAGAAATC
		63.1	Reverse	CTCCCTTGAACACGCACAAA
		69.5	Probe	[FAM]TGCAATGAAGACCTGG AAATCCCCGA[BHQ1]
CDK6	161	63.3	Forward	GTGGACATGGAACCTGGAGA
		63.1	Reverse	GAAAGACACAACTGGCAGCA
		68.4	Probe	[FAM]TCTCACCTATGACAGGG TTGAGCCA[BHQ1]
CDKN2A	193	63.5	Forward	TGACTCCCTCCCCATTTTCC
		63.3	Reverse	TTTTGGAGAGTCGGACTGCT
		68.9	Probe	[FAM]TTGCCCCAGACAGCCGT TTTACAC[BHQ1]
EGFR	232	63.7	Forward	ATCCTGAGATGGTTTGGGCA
		63.1	Reverse	ATAGAAGTACCCGCCTGCTC
		69.2	Probe	[FAM]CCCTCCAGACCTCTT TCCCCACC[BHQ1]
HMGA2	216	63.2	Forward	GGGGAGAAGGAACACACAGA
		63.5	Reverse	AGTTAGATGCAGCAGTGGGT
		68.9	Probe	[FAM]AGCAGTAACAATGCTCC AAACCACACC[BHQ1]
KDR	188	62	Forward	AGGAAGAACTATCAGTTGACAGAAT
		63.2	Reverse	GGGAGAGAGAACATAAGAGCTACC
		69.4	Probe	[FAM]TCAACCCATGTTTTCCCC TTCTCATAGCAT[BHQ1]
KIT	250	63.5	Forward	AGCCCTACTGCATGTCAAGT
		63.3	Reverse	CCTACTCCAGTGCCCAAGAA
		69	Probe	[FAM]CTCCTGTTACTGTAGCTGG CCTGGG[BHQ1]
MET	154	63.4	Forward	TTGTCTTCCCATCCACCCTC
		63.6	Reverse	TCCCCATTTCTTCCTTCCCC
		68.2	Probe	[FAM]ACTATGAGCTGTGAGAGTCTG GTCATTGAT[BHQ1]
MYCN	161	61.7	Forward	ATGAGTTGTGAAAGTTTTGAGTAGAT
		61.6	Reverse	ACTTTGCATTTACCCAGTTCTATG
		68.1	Probe	[FAM]TGCCTTTTTCCTAGC CTGTTTCTTCCT[BHQ1]
PAX9	165	61.7	Forward	CCAGTCTGGTGAGAAAATAGACT
		62.1	Reverse	ACAGCCACCTAAAAACATTTGATAA
		69.5	Probe	[FAM]AACCATACCATACAGG GACTCTCCTGTCA[BHQ1]
PTEN	202	61.7	Forward	GGTTTAATAGAGGTGAACTGTCTTTC
		62.2	Reverse	AGACTGTACTCTAAAAGCATTTCCT
		67.2	Probe	[FAM]CTTTTCTTTGTTGGTG GCATGAGTCCTAT[BHQ1]
PTPRD	200	61.7	Forward	CCATCAGTGGAAAATTAAGAGCTAC
		61.6	Reverse	TATTAACCATCCAATTATGACAGTGAG
		67	Probe	[FAM]AGATGAAGGGACATAGCATCT GATTATCGT[BHQ1]
YWHAZ	247	63.4	Forward	TCAGCGACAGGTCTCCAAAT
		63.3	Reverse	TCTCCTCCCTTTTGTGGTCC
		68.8	Probe	[FAM]AGGGTCTAAGGAGACCA ATGCCCAG[BHQ1]

Table A3. *Cont.*

Gene	PCR-Product Length, bp	Tm, °C	Primer	Sequence 5′–3′
TP53	192	63.3	Forward	GGACCTCTTAACCTGTGGCT
		63.4	Reverse	AAAGCTGTTCCGTCCCAGTA
		69	Probe	[FAM]CAGAAAGGACAAGGGTGG TTGGGAGTA[BHQ1]
RPP30	62	62.9	Forward	GATTTGGACCTGCGAGCG
		64.2	Reverse	GCGGCTGTCTCCACAAGT
		69.9	Probe	[R6G] TCTGACCTGAAGGCTC TGCGCG[BHQ2]
ALB	94	61.8	Forward	GACTTGCCAAGACATATGAAACC
		61.6	Reverse	TCCAACAATAAACCTACCACTTTG
		69.3	Probe	[R6G] TGCTGTGCCGCT GCAGATCC[BHQ2]

Appendix C

Table A4. Settings for PCR analysis with probes.

Cycling Step	Temperature, °C	Time	Ramp Rate	Number of Cycles
Enzyme activation	95	10 min	2 °C/s	1
Denaturation	94	30 s		40
Annealing/extension	58	1 min		
Enzyme deactivation	98	10 min		1
Hold (optional)	4	∞	1 °C/s	1

Table A5. Settings for PCR analysis with EvaGreen.

Cycling Step	Temperature, °C	Time	Ramp Rate	Number of Cycles
Enzyme activation	95	5 min	2 °C/s	1
Denaturation	95	30 s		40
Annealing/extension	60	1 min		
Signal stabilization	4	5 min		1
	90	5 min		1
Hold (optional)	4	∞	1°C/s	1

Appendix D

Table A6. A research plan for gene panel no. 1 for verification of copy number variations by digital droplet PCR analysis.

Genes	Ref. Gene *RPP30*; EvaGreen *	Ref. Gene *RPP30*; Probes **	Ref. Gene *ALB*; EvaGreen *	Ref. Gene *ALB*; Probes **
CCND2	+	+	+	+
CDK6	+	+	+	+
CDKN2A	+	+	+	+
EGFR	+	+	+	+
KIT	+	+	+	+
MYCN	+	+	+	+
PAX9	+	+	+	+

Table A6. *Cont.*

Genes	Ref. Gene *RPP30*; EvaGreen *	Ref. Gene *RPP30*; Probes **	Ref. Gene *ALB*; EvaGreen *	Ref. Gene *ALB*; Probes **
PTEN	+	+	+	+
PTPRD	+	+	+	+
YWHAZ	+	+	+	+
KDR	+		+	
C8orf4 (TCIM)	+		+	
HMGA2	+		+	
MET	+		+	
TP53	+		+	

* ddPCR analysis in the presence of EvaGreen intercalating dye. ** ddPCR analysis using probes.

Appendix E

Table A7. Data on copy number variations in normal tissue and ovarian cancer tissue */**.

Genes	Pcr_*RPP30* EvaGreen (Normal)	Pcr_*RPP30* EvaGreen (Tumor)	Pcr_*RPP30* Probes (Normal)	Pcr_*RPP30* Probes (Tumor)	NanoString CNV (Normal)	NanoString CNV (Tumor)	*p*-Value Pcr EvaGreen	*p*-Value Pcr Probes	*p*-Value NanoString CNV	*p*-Value ***	*p*-Value ****
C8orf4 (TCIM)	1.61 [1.55–1.73] #	1.59 [1.51–1.81]	No data	No data	2.16 [2.11–2.22]	2.2 [2.18–2.26]	0.972	No data	0.807	0.882	1
CCND2	No data	No data	1.75 [1.66–1.85]	2.05 [1.91–2.57]	2.06 [1.94–2.17]	2.27 [2.15–2.44]	No data	0.003	0.092	0.006	0.042
CDK6	2.05 [2.03–2.13]	2.15 [2.02–2.41]	2.16 [2.13–2.34]	2.29 [2.13–2.48]	2.23 [2.2–2.35]	2.42 [2.21–2.51]	0.497	0.216	1	0.393	1
HMGA2	1.92 [1.82–1.96]	2.03 [1.8–2.1]	No data	No data	2.27 [2.17–2.31]	2.33 [2.28–2.45]	0.505	No data	0.168	0.252	1
KDR	1.78 [1.67–1.91]	1.88 [1.7–1.98]	No data	No data	2.19 [2.14–2.27]	2.3 [2.12–2.37]	0.685	No data	0.363	0.475	1
MET	1.98 [1.92–2.16]	2.06 [1.84–2.21]	No data	No data	2.28 [2.14–2.36]	2.22 [2.2–2.28]	0.893	No data	0.421	0.573	1
PAX9	1.81 [1.68–1.9]	1.84 [1.83–1.92]	2.13 [2.03–2.19]	2.09 [1.93–2.24]	2.23 [2.15–2.36]	2.35 [2.15–2.69]	0.685	0.542	0.168	0.324	1

* The consistency between methods was calculated using modified Kappa statistic (PABAK). To transform the quantitative variables to qualitative, the following rule has been used everywhere: copy number values between 1.5 and 2.5, inclusively, were considered as 2 and values lower than 1.5 and greater than 2.5 were accepted as the deletion and amplification, respectively. The genes presented in the table are characterized by relatively high PABAK values according to the pairwise comparison of several methods of copy number profiling yielding qualitative data. The descriptive statistics of the copy number data are represented as median, first, and third quartile. The difference between the copy number profiles of normal and tumor tissue has been studied using two or more methods of copy number analysis. Two or more methods of copy number analysis have been used for one pair of samples, so for such a pair the respective number of statistical testing procedures should be performed. We used the Wilcoxon test for the quantitative paired samples. The harmonic mean of the resulting set of *p*-values was calculated. To counteract the multiple testing problem, all the obtained harmonic means of *p*-values were adjusted according to the Holm method. ** results obtained by CoreExome microarray platform are qualitative variables and have thus been excluded from the analysis. *** harmonic mean *p*-value **** adjusted harmonic mean *p*-value according to Holm's method # median [first quartile–third quartile].

Appendix F

Biological features and prognostic value of target genes

Table A8 reveals that seven known oncogenes, namely MET, HMGA2, KDR, C8orf4 (TCIM), PAX9, CDK6, and CCND2, are oncogenes. The frequency of gene amplifications in ovarian cancer tissue ranges from 2.1 to 11.6%, while the frequency of deletions does not surpass 1%. The potential clinical significance of these genes can be discussed in terms of studying the corresponding protein products as potential drug targets. The prognostic value of each of the seven genes was determined primarily by the correlation with progression-free survival (PFS) and overall survival (OS). Based on CNV signatures, the survival rate of people with ovarian cancer may not depend on how genes work, giving them an edge over other expression biomarkers. Therefore, we looked into the TCGA Ovarian Serous Cystadenocarcinoma dataset in the Broad GDAC Firehose (n = 617) [67] using the cBioPortal [65,66] resource to see if there were any links between CNVs and how well the cancer would do. The signature of three genes, TCIM, PAX9, and CCND2, with amplifications was found to correlate with a much lower OS in ovarian cancer patients (Figure A1) of 36.89 months (32.98–44.48, 95% confidence interval, CI) compared

to the group without amplifications of these genes, 48.29 months (44.51–54.07, 95% CI) at *p*-value = 0.009 and q-value = 0.02. The other four genes (MET, HMGA2, KDR, and CDK6) that had different copy numbers in ovarian cancer samples were not linked to either PFS or OS.

We extracted the panoramic expression data of proteins encoded by TCIM, PAX9, and CCND2 genes from the Human Proteome Atlas database [68] According to HPA, moderate and low levels of TCIM protein immunoreactivity were observed in almost all the ovarian cancer samples (11 out of 12 HPA samples), while CCND2 protein was detected in only 33% of all the samples (4 out of 12 HPA samples). There are no data on PAX9 protein immunoreactivity. The BioGrid resource [68] reports no direct interactions between these three proteins, but identifies PFDN5 (prefoldin subunit 5 alpha) as a common protein partner for PAX9 and CCND2 proteins. There are no direct interactions between these three proteins; however, one common protein partner, PFDN5 (prefoldin subunit 5 alpha), has been identified for PAX9 and CCND2 proteins. The PFDN5 protein is part of a hexamolecular chaperone complex that stabilizes newly synthesized polypeptides, creating conditions for proper protein folding.

The TCIM protein is overexpressed in solid tumors and coordinates the oncogenic Wnt/β catenin and FGFR2 signaling pathways, which influence the formation of an aggressive tumor phenotype [70–72]. Transcription factor PAX9 belongs to the PAX (the paired-box) family of highly homologous transcription factors overexpressed in ovarian cancer [73]. Some members of the PAX gene family are used as immunohistochemical biomarkers, e.g., PAX8 protein [74]. Wang et al. found that PAX9 interacts and co-functions with the nucleosome remodeling and deacetylase complex NuRD (nucleosome remodeling and deacetylase complex) by suppressing the expression of downstream signaling cascade genes [75]. PAX9 regulates the activity of the Wnt/β-catenin signaling pathway by controlling the expression of the Dkk1/Dkk2 and Wnt9b/Wnt3 components, as evidenced by PAX9 gene knockout experiments in mouse models [76]. Cyclin D2, encoded by the CCND2 gene, is a regulator of cyclin-dependent kinases CDK4 and CDK6, the functional activity of which is critical for the control of the G1/S phase cell cycle [77].

Two different CNV detection approaches confirm that ovarian cancer tissue significantly amplifies the CCND2 gene. In this case, presumably, in the presence of co-amplification of the CDK6 gene in ovarian cancer, a synergistic effect of stimulating the proliferation of tumor cells can take place. Myant & Sansom [78] observed an interesting fact: the productive proliferation of colorectal adenoma cells, characterized by impaired transduction of the Wnt signaling pathway, necessitates high levels of cyclin D2. Conversely, the proliferation of normal cells does not rely on cyclin D2, allowing us to concentrate on the pharmacological effect on the CDK4/6 and CCND2 complex. The presence of amplification of the TCIM, PAX9, and CCND2 genes is linked to a lower overall survival rate in people with ovarian cancer (Figure A1). This can change the activity of the Wnt/β-catenin signaling pathway, which is seen in tumors and is linked to a more aggressive course of cancer.

Table A8. Biological features and potential prognostic value of a panel of genes for which CNVs were detected in ovarian cancer samples.

Genes *	Frequency of CNVs **	Oncogene ***/ Chromosomal Localization	Potential Prognostic Value of a Gene	Known Drug Target	References
MET	Amplification, 6.7%	Oncogene/Chr 7th	Prognostic biomarker	Drug target	[79,80]
HMGA2	Amplification, 3.6%	Oncogene/Chr 14th	Prognostic biomarker	Drug target	[81,82]
KDR	Amplification, 2.1%	Oncogene/Chr 12th	Prognostic biomarker	Not found	[83]

Table A8. *Cont.*

Genes *	Frequency of CNVs **	Oncogene ***/ Chromosomal Localization	Potential Prognostic Value of a Gene	Known Drug Target	References
C8orf4 (TCIM)	Amplification, 2.9%	No data/Chr 8th	Prognostic biomarker	Not found	[84]
PAX9	Amplification, 1.4%	No data/Chr 14th	Prognostic biomarker	Not found	[85]
CDK6	Amplification, 4.2%	Oncogene/Chr 21th	Not found	Drug target	[86]
CCND2	Amplification, 11.6%	Oncogene/Chr13th	Not found	Drug target	[87]

* 'MET'—MET proto-oncogene, receptor tyrosine kinase (UniProtKB ID P08581); 'HMGA2'—high mobility group AT-hook 2 (UniProtKB ID P52926); 'KDR'—kinase insert domain receptor (UniProtKB ID P35968); 'C8orf4'—chromosome 8 open reading frame 4 (UniProtKB ID Q9NR00); 'PAX9'—paired-box 9 (UniProtKB ID P55771); 'CDK6'—cyclin-dependent kinase 6 (UniProtKB ID Q00534); 'CCND2'—cyclin D2 (UniProtKB ID P30279). ** according to cBioPortal database (TCGA Ovarian Serous Cystadenocarcinoma, 617 samples) *** according to Cancer Gene Census Database (https://cancer.sanger.ac.uk/census (accessed on 30 August 2024)).

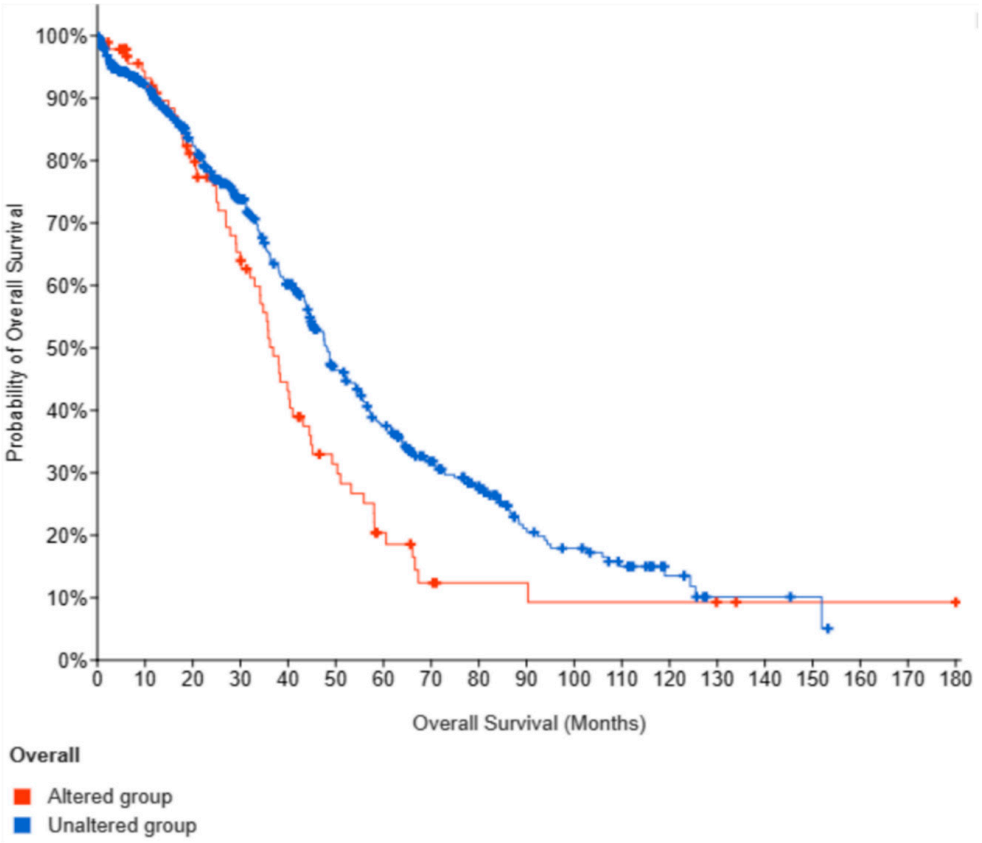


Figure A1. Overall survival rates of patients with ovarian cancer depending on the amplification of any of the three target genes TCIM, PAX9, and CCND2. A survival assessment was conducted using cBioPortal [97] based on public data from ovarian cancer studies (dataset: ovarian serous cystadenocarcinoma, TCGA). The log-rank test is used to test the null hypothesis that there is no difference between the groups in the probability of an event at any time point. Hazard ratios (HR) are derived from the log-rank test. The altered group (red color) represents the samples, n = 93, with at least one amplification in the queried genes. The unaltered group (blue color) includes samples (n = 471) without any alteration in the queried genes.

Appendix G

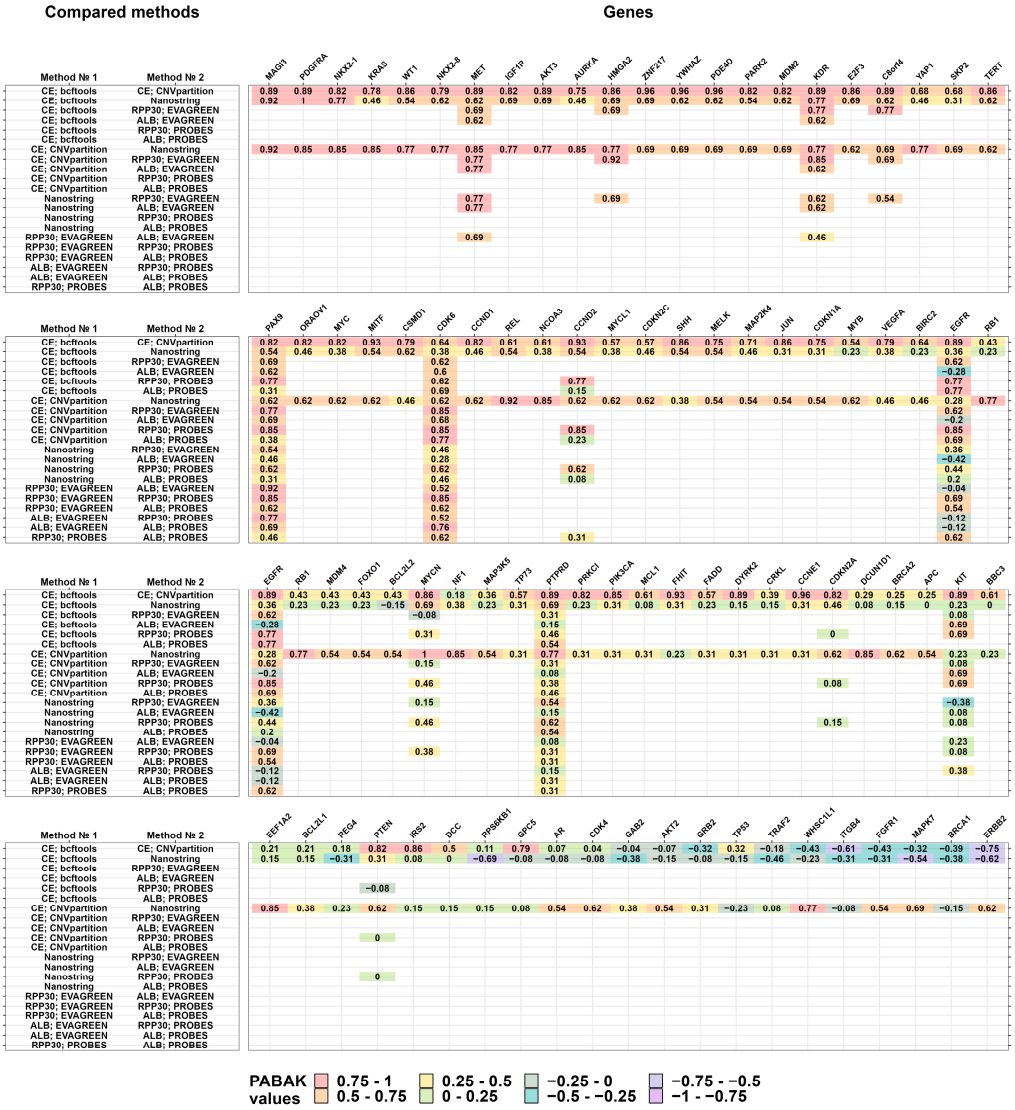


Figure A2. The consistency between the methods of copy number identification measured by prevalence-adjusted and bias-adjusted Kappa coefficients (PABAK). Rows correspond to the compared methods and columns correspond to the studied genes. The sort order is decreasing according to the median PABAK value per gene, from highest to lowest.

References

1. PCAWG Transcriptome Core Group; Calabrese, C.; Davidson, N.R.; Demircioğlu, D.; Fonseca, N.A.; He, Y.; Kahles, A.; Lehmann, K.-V.; Liu, F.; Shiraishi, Y.; et al. Genomic Basis for RNA Alterations in Cancer. *Nature* **2020**, *578*, 129–136. [CrossRef]
2. Steele, C.D.; Abbasi, A.; Islam, S.M.A.; Bowes, A.L.; Khandekar, A.; Haase, K.; Hames-Fathi, S.; Ajayi, D.; Verfaillie, A.; Dhami, P.; et al. Signatures of Copy Number Alterations in Human Cancer. *Nature* **2022**, *606*, 984–991. [CrossRef] [PubMed]
3. Harbers, L.; Agostini, F.; Nicos, M.; Poddighe, D.; Bienko, M.; Crosetto, N. Somatic Copy Number Alterations in Human Cancers: An Analysis of Publicly Available Data From The Cancer Genome Atlas. *Front. Oncol.* **2021**, *11*, 700568. [CrossRef] [PubMed]
4. Huang, T.; Li, J.; Jia, B.; Sang, H. CNV-MEANN: A Neural Network and Mind Evolutionary Algorithm-Based Detection of Copy Number Variations From Next-Generation Sequencing Data. *Front. Genet.* **2021**, *12*, 700874. [CrossRef]
5. Andor, N.; Maley, C.C.; Ji, H.P. Genomic Instability in Cancer: Teetering on the Limit of Tolerance. *Cancer Res.* **2017**, *77*, 2179–2185. [CrossRef]
6. Grist, E.; Friedrich, S.; Brawley, C.; Mendes, L.; Parry, M.; Ali, A.; Haran, A.; Hoyle, A.; Gilson, C.; Lall, S.; et al. Accumulation of Copy Number Alterations and Clinical Progression across Advanced Prostate Cancer. *Genome Med.* **2022**, *14*, 102. [CrossRef]
7. Pan, X.; Hu, X.; Zhang, Y.-H.; Chen, L.; Zhu, L.; Wan, S.; Huang, T.; Cai, Y.-D. Identification of the Copy Number Variant Biomarkers for Breast Cancer Subtypes. *Mol. Genet. Genom.* **2019**, *294*, 95–110. [CrossRef]

8. Dharanipragada, P.; Parekh, N. Genome-Wide Characterization of Copy Number Variations in Diffuse Large B-Cell Lymphoma with Implications in Targeted Therapy. *Precis. Clin. Med.* **2019**, *2*, 246–258. [CrossRef]
9. Zhou, C.; Yuan, Z.; Ma, W.; Qi, L.; Mahavongtrakul, A.; Li, Y.; Li, H.; Gong, J.; Fan, R.R.; Li, J.; et al. Clinical Utility of Tumor Genomic Profiling in Patients with High Plasma Circulating Tumor DNA Burden or Metabolically Active Tumors. *J. Hematol. Oncol.* **2018**, *11*, 129. [CrossRef]
10. Premarket Approval (PMA). Available online: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P170019S017 (accessed on 30 August 2024).
11. Lalonde, E.; Ewens, K.; Richards-Yutz, J.; Ebrahimzedeh, J.; Terai, M.; Gonsalves, C.F.; Sato, T.; Shields, C.L.; Ganguly, A. Improved Uveal Melanoma Copy Number Subtypes Including an Ultra–High-Risk Group. *Ophthalmol. Sci.* **2022**, *2*, 100121. [CrossRef]
12. Tavares, I.; Martins, R.; Ribeiro, I.P.; Esteves, L.; Caramelo, F.; Abrantes, A.M.; Neves, R.; Caetano-Oliveira, R.; Botelho, M.F.; Barbosa De Melo, J.; et al. Development of a Genomic Predictive Model for Cholangiocarcinoma Using Copy Number Alteration Data. *J. Clin. Pathol.* **2022**, *75*, 274–278. [CrossRef] [PubMed]
13. Takahashi, N.; Hatakeyama, K.; Nagashima, T.; Ohshima, K.; Urakami, K.; Yamaguchi, K.; Hirashima, Y. Characterization of Rare Histological Subtypes of Ovarian Cancer Based on Molecular Profiling. *Cancer Med.* **2023**, *12*, 387–395. [CrossRef]
14. Zampaglione, E.; Kinde, B.; Place, E.M.; Navarro-Gomez, D.; Maher, M.; Jamshidi, F.; Nassiri, S.; Mazzone, J.A.; Finn, C.; Schlegel, D.; et al. Copy-Number Variation Contributes 9% of Pathogenicity in the Inherited Retinal Degenerations. *Genet. Med.* **2020**, *22*, 1079–1087. [CrossRef] [PubMed]
15. Van Roy, N.; Van Der Linden, M.; Menten, B.; Dheedene, A.; Vandeputte, C.; Van Dorpe, J.; Laureys, G.; Renard, M.; Sante, T.; Lammens, T.; et al. Shallow Whole Genome Sequencing on Circulating Cell-Free DNA Allows Reliable Noninvasive Copy-Number Profiling in Neuroblastoma Patients. *Clin. Cancer Res.* **2017**, *23*, 6305–6314. [CrossRef] [PubMed]
16. Grond-Ginsbach, C.; Erhart, P.; Chen, B.; Kloss, M.; Engelter, S.T.; Cole, J.W. Copy Number Variation and Risk of Stroke. *Stroke* **2018**, *49*, 2549–2554. [CrossRef]
17. Gordeeva, V.; Sharova, E.; Babalyan, K.; Sultanov, R.; Govorun, V.M.; Arapidi, G. Benchmarking Germline CNV Calling Tools from Exome Sequencing Data. *Sci. Rep.* **2021**, *11*, 14416. [CrossRef]
18. Gabrielaite, M.; Torp, M.H.; Rasmussen, M.S.; Andreu-Sánchez, S.; Vieira, F.G.; Pedersen, C.B.; Kinalis, S.; Madsen, M.B.; Kodama, M.; Demircan, G.S.; et al. A Comparison of Tools for Copy-Number Variation Detection in Germline Whole Exome and Whole Genome Sequencing Data. *Cancers* **2021**, *13*, 6283. [CrossRef]
19. Shen, W.; Szankasi, P.; Durtschi, J.; Kelley, T.W.; Xu, X. Genome-Wide Copy Number Variation Detection Using NGS: Data Analysis and Interpretation. In *Tumor Profiling*; Murray, S.S., Ed.; Methods in Molecular Biology; Springer: New York, NY, USA, 2019; Volume 1908, pp. 113–124. ISBN 978-1-4939-9002-3.
20. Tsyganov, M.M.; Ibragimova, M.K.; Garbukov, E.Y.; Bragina, O.D.; Karchevskaya, A.A.; Usynin, E.A.; Litvyakov, N.V. Determination of BRCAness Phenotype in Breast Tumors for the Appointment of Neoadjuvant Chemotherapy Based on Platinum and Taxanes. *Int. J. Mol. Sci.* **2022**, *24*, 207. [CrossRef]
21. Sagath, L.; Lehtokari, V.-L.; Pelin, K.; Kiiski, K. Array Comparative Genomic Hybridisation and Droplet Digital PCR Uncover Recurrent Copy Number Variation of the TTN Segmental Duplication Region. *Genes* **2022**, *13*, 905. [CrossRef]
22. Sasa, K.; Saito, T.; Kurihara, T.; Hasegawa, N.; Sano, K.; Kubota, D.; Akaike, K.; Okubo, T.; Hayashi, T.; Takagi, T.; et al. Establishment of Rapid and Accurate Screening System for Molecular Target Therapy of Osteosarcoma. *Technol. Cancer Res. Treat.* **2022**, *21*, 153303382211382. [CrossRef]
23. Norton, N.; Advani, P.P.; Serie, D.J.; Geiger, X.J.; Necela, B.M.; Axenfeld, B.C.; Kachergus, J.M.; Feathers, R.W.; Carr, J.M.; Crook, J.E.; et al. Assessment of Tumor Heterogeneity, as Evidenced by Gene Expression Profiles, Pathway Activation, and Gene Copy Number, in Patients with Multifocal Invasive Lobular Breast Tumors. *PLoS ONE* **2016**, *11*, e0153411. [CrossRef] [PubMed]
24. O'Hern, K.; Barney, R.; Chambers, M.; Baker, C.; Stevanovic, M.; Tsongalis, G.J.; Hughes, E.; Sriharan, A. A Novel Method to Assess Copy Number Variation in Melanoma: Droplet Digital PCR for Precise Quantitation of the RREB1 Gene in Formalin-fixed, Paraffin-embedded Melanocytic Neoplasms, a PROOF-OF-CONCEPT Study. *J. Cutan. Pathol.* **2023**, *50*, 169–177. [CrossRef] [PubMed]
25. Levsky, J.M.; Singer, R.H. Fluorescence in Situ Hybridization: Past, Present and Future. *J. Cell Sci.* **2003**, *116*, 2833–2838. [CrossRef] [PubMed]
26. Schouten, J.P.; McElgunn, C.J.; Waaijer, R.; Zwijnenburg, D.; Diepvens, F.; Pals, G. Relative Quantification of 40 Nucleic Acid Sequences by Multiplex Ligation-Dependent Probe Amplification. *Nucleic Acids Res.* **2002**, *30*, e57. [CrossRef]
27. Microarrays | Microarray Analysis Techniques and Products. Available online: https://emea.illumina.com/techniques/microarrays.html (accessed on 30 August 2024).
28. Infinium Exome-24 Kit. Available online: https://emea.illumina.com/products/by-type/microarray-kits/infinium-exome.html#tabs-6c809a704c-item-8cfede5c7e-documentation (accessed on 30 August 2024).
29. Rentas, S.; Abou Tayoun, A. Utility of Droplet Digital PCR and NGS-Based CNV Clinical Assays in Hearing Loss Diagnostics: Current Status and Future Prospects. *Expert Rev. Mol. Diagn.* **2021**, *21*, 213–221. [CrossRef]
30. nCounter v2 Cancer CN Assay PRODUCT BULLETIN. Available online: https://nanostring.com/wp-content/uploads/v2-Cancer-CN-Assay.pdf (accessed on 30 August 2024).

31. Kutasovic, J.R.; McCart Reed, A.E.; Males, R.; Sim, S.; Saunus, J.M.; Dalley, A.; McEvoy, C.R.; Dedina, L.; Miller, G.; Peyton, S.; et al. Breast Cancer Metastasis to Gynaecological Organs: A Clinico-pathological and Molecular Profiling Study. *J. Pathol. Clin. Res.* **2019**, *5*, 25–39. [CrossRef]
32. Cusenza, V.Y.; Bisagni, A.; Rinaldini, M.; Cattani, C.; Frazzi, R. Copy Number Variation and Rearrangements Assessment in Cancer: Comparison of Droplet Digital PCR with the Current Approaches. *Int. J. Mol. Sci.* **2021**, *22*, 4732. [CrossRef]
33. Sagath, L.; Lehtokari, V.-L.; Wallgren-Pettersson, C.; Pelin, K.; Kiiski, K. A Custom ddPCR Method for the Detection of Copy Number Variations in the Nebulin Triplicate Region. *PLoS ONE* **2022**, *17*, e0267793. [CrossRef]
34. Danecek, P.; McCarthy, S.A.; HipSci Consortium; Durbin, R. A Method for Checking Genomic Integrity in Cultured Cell Lines from SNP Genotyping Data. *PLoS ONE* **2016**, *11*, e0155014. [CrossRef]
35. Heydt, C.; Becher, A.-K.; Wagener-Ryczek, S.; Ball, M.; Schultheis, A.M.; Schallenberg, S.; Rüsseler, V.; Büttner, R.; Merkelbach-Bruse, S. Comparison of in Situ and Extraction-Based Methods for the Detection of MET Amplifications in Solid Tumors. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 1339–1347. [CrossRef]
36. Peng, H.; Lu, L.; Zhou, Z.; Liu, J.; Zhang, D.; Nan, K.; Zhao, X.; Li, F.; Tian, L.; Dong, H.; et al. CNV Detection from Circulating Tumor DNA in Late Stage Non-Small Cell Lung Cancer Patients. *Genes* **2019**, *10*, 926. [CrossRef] [PubMed]
37. Cristescu, R.; Liu, X.Q.; Arreaza, G.; Chen, C.; Albright, A.; Qiu, P.; Marton, M.J. Concordance between Single-Nucleotide Polymorphism–Based Genomic Instability Assays and a next-Generation Sequencing–Based Homologous Recombination Deficiency Test. *BMC Cancer* **2022**, *22*, 1310. [CrossRef] [PubMed]
38. Rafter, P.; Gormley, I.C.; Parnell, A.C.; Kearney, J.F.; Berry, D.P. Concordance Rate between Copy Number Variants Detected Using Either High- or Medium-Density Single Nucleotide Polymorphism Genotype Panels and the Potential of Imputing Copy Number Variants from Flanking High Density Single Nucleotide Polymorphism Haplotypes in Cattle. *BMC Genom.* **2020**, *21*, 205. [CrossRef]
39. 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]
40. Li, X.; Ding, Y.; Liu, Y.; Yang, M. Differences between Complex Epithelial Neoplasms of the Ovary and High-Grade Serous Ovarian Cancer: A Retrospective Observational Cohort Study. *J. Ovarian Res.* **2022**, *15*, 125. [CrossRef]
41. Khashaba, M.; Fawzy, M.; Abdel-Aziz, A.; Eladawei, G.; Nagib, R. Subtyping of High Grade Serous Ovarian Carcinoma: Histopathological and Immunohistochemical Approach. *J. Egypt. Natl. Cancer Inst.* **2022**, *34*, 6. [CrossRef]
42. Ritch, S.J.; Noman, A.S.M.; Goyeneche, A.A.; Telleria, C.M. The Metastatic Capacity of High-Grade Serous Ovarian Cancer Cells Changes along Disease Progression: Inhibition by Mifepristone. *Cancer Cell Int.* **2022**, *22*, 397. [CrossRef]
43. Graf, R.P.; Eskander, R.; Brueggeman, L.; Stupack, D.G. Association of Copy Number Variation Signature and Survival in Patients With Serous Ovarian Cancer. *JAMA Netw. Open* **2021**, *4*, e2114162. [CrossRef]
44. Engler, D.A.; Gupta, S.; Growdon, W.B.; Drapkin, R.I.; Nitta, M.; Sergent, P.A.; Allred, S.F.; Gross, J.; Deavers, M.T.; Kuo, W.-L.; et al. Genome Wide DNA Copy Number Analysis of Serous Type Ovarian Carcinomas Identifies Genetic Markers Predictive of Clinical Outcome. *PLoS ONE* **2012**, *7*, e30996. [CrossRef]
45. Genovese, G. Freeseek/Gtc2vcf 2024. Available online: https://github.com/freeseek/gtc2vcf (accessed on 30 August 2024).
46. GenomicRanges. Available online: http://bioconductor.org/packages/GenomicRanges/ (accessed on 30 August 2024).
47. Benson, D.A.; Cavanaugh, M.; Clark, K.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Sayers, E.W. GenBank. *Nucleic Acids Res.* **2012**, *41*, D36–D42. [CrossRef]
48. Oscorbin, I.; Kechin, A.; Boyarskikh, U.; Filipenko, M. Multiplex ddPCR Assay for Screening Copy Number Variations in BRCA1 Gene. *Breast Cancer Res. Treat.* **2019**, *178*, 545–555. [CrossRef] [PubMed]
49. Real-Time PCR Application Guide. Available online: https://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_5279.pdf (accessed on 30 August 2024).
50. Ye, J.; Coulouris, G.; Zaretskaya, I.; Cutcutache, I.; Rozen, S.; Madden, T.L. Primer-BLAST: A Tool to Design Target-Specific Primers for Polymerase Chain Reaction. *BMC Bioinform.* **2012**, *13*, 134. [CrossRef] [PubMed]
51. OligoAnalyzer Tool—Primer Analysis and Tm Calculator | IDT. Available online: https://eu.idtdna.com/pages/tools/oligoanalyzer (accessed on 30 August 2024).
52. Bell, A.D.; Usher, C.L.; McCarroll, S.A. Analyzing Copy Number Variation with Droplet Digital PCR. In *Digital PCR*; Karlin-Neumann, G., Bizouarn, F., Eds.; Methods in Molecular Biology; Springer: New York, NY, USA, 2018; Volume 1768, pp. 143–160. ISBN 978-1-4939-7776-5.
53. Data Analysis Guidelines for Copy Number Variation (CNV). Available online: https://nanostring.com/wp-content/uploads/MAN-C0014-02_nCounter_CNV_Data_Analysis_Guidelines.pdf (accessed on 30 August 2024).
54. McHugh, M.L. Interrater Reliability: The Kappa Statistic. *Biochem. Medica* **2012**, *22*, 276–282. [CrossRef]
55. Byrt, T.; Bishop, J.; Carlin, J.B. Bias, Prevalence and Kappa. *J. Clin. Epidemiol.* **1993**, *46*, 423–429. [CrossRef] [PubMed]
56. Stevenson, M.; Sergeant, E.; Heuer, C.; Nunes, T.; Heuer, C.; Marshall, J.; Sanchez, J.; Thornton, R.; Reiczigel, J.; Robison-Cox, J.; et al. *epiR: Tools for the Analysis of Epidemiological Data*; The Comprehensive R Archive Network: New York, NY, USA, 2024.
57. Potapov, S.; Model, F.; Schuetzenmeister, A.; Manuilova, E.; Dufey, F.; Raymaekers, J.; Seshan, V.E. *Roche Mcr: Method Comparison Regression*; The Comprehensive R Archive Network: New York, NY, USA, 2023.
58. Bilic-Zulle, L. Comparison of Methods: Passing and Bablok Regression. *Biochem. Medica* **2011**, *21*, 49–52. [CrossRef]
59. Mayer, B.; Gaus, W.; Braisch, J.U. The Fallacy of the Passing-Bablok-Regression. *Jökull J.* **2016**, *66*, 95–106.

60. Han, M.; Yamaguchi, S.; Onishi, M.; Fujii, T.; Hosoya, M.; Wen, X.; Kido, H.; Kato, S. The *MDM2* and *CDKN2A* Copy-Number-Variation Influence the *TP53* -Signature-Score in Wild-Type *TP53* Luminal Type Breast Cancer. *Anticancer Res.* **2022**, *42*, 2277–2288. [CrossRef]
61. Li, S.; Dou, X.; Gao, R.; Ge, X.; Qian, M.; Wan, L. A Remark on Copy Number Variation Detection Methods. *PLoS ONE* **2018**, *13*, e0196226. [CrossRef]
62. Ahn, S.; Hong, M.; Van Vrancken, M.; Lyou, Y.J.; Kim, S.T.; Park, S.H.; Kang, W.K.; Park, Y.S.; Jung, S.-H.; Woo, M.; et al. A nCounter CNV Assay to Detect HER2 Amplification: A Correlation Study with Immunohistochemistry and In Situ Hybridization in Advanced Gastric Cancer. *Mol. Diagn. Ther.* **2016**, *20*, 375–383. [CrossRef]
63. Kutilin, D.S.; Tsandekova, M.R.; Porkhanova, N.V. Features of the Copy Number Variation of Certain Genes in Tumor Cells in Patients with Serous Ovarian Adenocarcinoma. *Bull. Exp. Biol. Med.* **2021**, *170*, 332–339. [CrossRef]
64. Cosmic Cancer Gene Census. Available online: http://cancer.sanger.ac.uk/census (accessed on 30 August 2024).
65. Cerami, E.; Gao, J.; Dogrusoz, U.; Gross, B.E.; Sumer, S.O.; Aksoy, B.A.; Jacobsen, A.; Byrne, C.J.; Heuer, M.L.; Larsson, E.; et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov.* **2012**, *2*, 401–404. [CrossRef] [PubMed]
66. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal.* **2013**, *6*, pl1. [CrossRef] [PubMed]
67. Broad GDAC Firehose. Available online: https://gdac.broadinstitute.org/ (accessed on 30 August 2024).
68. Uhlén, M.; Pontén, F.; Lindskog, C. Charting the Human Proteome: Understanding Disease Using a Tissue-Based Atlas. *Science* **2015**, *347*, 1274. [CrossRef]
69. Oughtred, R.; Rust, J.; Chang, C.; Breitkreutz, B.; Stark, C.; Willems, A.; Boucher, L.; Leung, G.; Kolas, N.; Zhang, F.; et al. The BIOGRID Database: A Comprehensive Biomedical Resource of Curated Protein, Genetic, and Chemical Interactions. *Protein Sci.* **2021**, *30*, 187–200. [CrossRef]
70. Yang, Z.; Moffa, A.B.; Haddad, R.; Streicher, K.L.; Ethier, S.P. Transforming Properties of TC-1 in Human Breast Cancer: Interaction with FGFR2 and B-catenin Signaling Pathways. *Int. J. Cancer* **2007**, *121*, 1265–1273. [CrossRef]
71. Kim, B.; Koo, H.; Yang, S.; Bang, S.; Jung, Y.; Kim, Y.; Kim, J.; Park, J.; Moon, R.T.; Song, K.; et al. TC1(C8orf4) Correlates with Wnt/β-Catenin Target Genes and Aggressive Biological Behavior in Gastric Cancer. *Clin. Cancer Res.* **2006**, *12*, 3541–3548. [CrossRef]
72. De Melo Martins, P.C.; Parise Junior, O.; Pereira Hors, C.; Villela Miguel, R.E.; Da Costa Andrade, V.C.; Garicochea, B. *C8orf4/TC-1(Thyroid Cancer-1)* Gene Expression in Thyroid Cancer and Goiter. *ORL* **2007**, *69*, 127–130. [CrossRef]
73. Muratovska, A.; Zhou, C.; He, S.; Goodyer, P.; Eccles, M.R. Paired-Box Genes Are Frequently Expressed in Cancer and Often Required for Cancer Cell Survival. *Oncogene* **2003**, *22*, 7989–7997. [CrossRef]
74. Tacha, D.; Zhou, D.; Cheng, L. Expression of PAX8 in Normal and Neoplastic Tissues: A Comprehensive Immunohistochemical Study. *Appl. Immunohistochem. Mol. Morphol.* **2011**, *19*, 293–299. [CrossRef]
75. Zhao, Z.; Szczepanski, A.P.; Tsuboyama, N.; Abdala-Valencia, H.; Goo, Y.A.; Singer, B.D.; Bartom, E.T.; Yue, F.; Wang, L. PAX9 Determines Epigenetic State Transition and Cell Fate in Cancer. *Cancer Res.* **2021**, *81*, 4696–4708. [CrossRef]
76. Jia, S.; Zhou, J.; D'Souza, R.N. Pax9's Dual Roles in Modulating Wnt Signaling during Murine Palatogenesis. *Dev. Dyn.* **2020**, *249*, 1274–1284. [CrossRef] [PubMed]
77. Hume, S.; Dianov, G.L.; Ramadan, K. A Unified Model for the G1/S Cell Cycle Transition. *Nucleic Acids Res.* **2020**, *48*, 12483–12501. [CrossRef] [PubMed]
78. Myant, K.; Sansom, O. Efficient Wnt Mediated Intestinal Hyperproliferation Requires the Cyclin D2-CDK4/6 Complex. *Cell Div.* **2011**, *6*, 3. [CrossRef] [PubMed]
79. Bååth, M.; Jönsson, J.-M.; Westbom Fremer, S.; Martín de la Fuente, L.; Tran, L.; Malander, S.; Kannisto, P.; Måsbäck, A.; Honeth, G.; Hedenfalk, I. MET Expression and Cancer Stem Cell Networks Impact Outcome in High-Grade Serous Ovarian Cancer. *Genes* **2021**, *12*, 742. [CrossRef] [PubMed]
80. Huang, L.; Xie, K.; Li, H.; Wang, R.; Xu, X.; Chen, K.; Gu, H.; Fang, J. Suppression of C-Met-Overexpressing Tumors by a Novel c-Met/CD3 Bispecific Antibody. *Drug Des. Dev. Ther.* **2020**, *14*, 3201–3214. [CrossRef]
81. Song, Z.; Liao, C.; Yao, L.; Xu, X.; Shen, X.; Tian, S.; Wang, S.; Xing, F. miR-219-5p Attenuates Cisplatin Resistance of Ovarian Cancer by Inactivating Wnt/β-Catenin Signaling and Autophagy via Targeting HMGA2. *Cancer Gene Ther.* **2022**, *30*, 596–607. [CrossRef]
82. Wei, J.-J. HMGA2: A Biomarker in Gynecologic Neoplasia. *J. Clin. Transl. Pathol.* **2022**, *2*, 3–7. [CrossRef]
83. Si, M.; Zhang, J.; Cao, J.; Xie, Z.; Shu, S.; Zhu, Y.; Lang, J. Integrated Analysis To Identify Molecular Biomarkers Of High-Grade Serous Ovarian Cancer. *OncoTargets Ther.* **2019**, *12*, 10057–10075. [CrossRef]
84. Xu, H.-T.; Liu, Y.; Liu, S.-L.; Miao, Y.; Li, Q.-C.; Wang, E.-H. TC-1 (C8orf4) Expression Is Correlated with Differentiation in Ovarian Carcinomas and Might Distinguish Metastatic Ovarian from Metastatic Colorectal Carcinomas. *Virchows Arch. Int. J. Pathol.* **2013**, *462*, 281–287. [CrossRef]
85. Chen, X.; Li, Y.; Paiboonrungruang, C.; Li, Y.; Peters, H.; Kist, R.; Xiong, Z. PAX9 in Cancer Development. *Int. J. Mol. Sci.* **2022**, *23*, 5589. [CrossRef]

86. Dall'Acqua, A.; Bartoletti, M.; Masoudi-Khoram, N.; Sorio, R.; Puglisi, F.; Belletti, B.; Baldassarre, G. Inhibition of CDK4/6 as Therapeutic Approach for Ovarian Cancer Patients: Current Evidences and Future Perspectives. *Cancers* **2021**, *13*, 3035. [CrossRef] [PubMed]
87. Chen, G.; Yan, Y.; Qiu, X.; Ye, C.; Jiang, X.; Song, S.; Zhang, Y.; Chang, H.; Wang, L.; He, X.; et al. miR-93-5p Suppresses Ovarian Cancer Malignancy and Negatively Regulate CCND2 by Binding to Its 3′UTR Region. *Discov. Oncol.* **2022**, *13*, 15. [CrossRef]
88. De Marco, C.; Zoppoli, P.; Rinaldo, N.; Morganella, S.; Morello, M.; Zuccalà, V.; Carriero, M.V.; Malanga, D.; Chirillo, R.; Bruni, P.; et al. Genome-Wide Analysis of Copy Number Alterations Led to the Characterisation of PDCD10 as Oncogene in Ovarian Cancer. *Transl. Oncol.* **2021**, *14*, 101013. [CrossRef] [PubMed]
89. Cheng, J.; Demeulemeester, J.; Wedge, D.C.; Vollan, H.K.M.; Pitt, J.J.; Russnes, H.G.; Pandey, B.P.; Nilsen, G.; Nord, S.; Bignell, G.R.; et al. Pan-Cancer Analysis of Homozygous Deletions in Primary Tumours Uncovers Rare Tumour Suppressors. *Nat. Commun.* **2017**, *8*, 1221. [CrossRef] [PubMed]
90. Testa, U.; Petrucci, E.; Pasquini, L.; Castelli, G.; Pelosi, E. Ovarian Cancers: Genetic Abnormalities, Tumor Heterogeneity and Progression, Clonal Evolution and Cancer Stem Cells. *Medicines* **2018**, *5*, 16. [CrossRef] [PubMed]
91. Nichols, C.A.; Gibson, W.J.; Brown, M.S.; Kosmicki, J.A.; Busanovich, J.P.; Wei, H.; Urbanski, L.M.; Curimjee, N.; Berger, A.C.; Gao, G.F.; et al. Loss of Heterozygosity of Essential Genes Represents a Widespread Class of Potential Cancer Vulnerabilities. *Nat. Commun.* **2020**, *11*, 2517. [CrossRef]
92. Skirnisdottir, I.; Mayrhofer, M.; Rydåker, M.; Åkerud, H.; Isaksson, A. Loss-of-Heterozygosity on Chromosome 19q in Early-Stage Serous Ovarian Cancer Is Associated with Recurrent Disease. *BMC Cancer* **2012**, *12*, 407. [CrossRef]
93. Zhai, Y.; Wu, R.; Kuick, R.; Sessine, M.S.; Schulman, S.; Green, M.; Fearon, E.R.; Cho, K.R. High-grade Serous Carcinomas Arise in the Mouse Oviduct via Defects Linked to the Human Disease. *J. Pathol.* **2017**, *243*, 16–25. [CrossRef]
94. Szabova, L.; Yin, C.; Bupp, S.; Guerin, T.M.; Schlomer, J.J.; Householder, D.B.; Baran, M.L.; Yi, M.; Song, Y.; Sun, W.; et al. Perturbation of Rb, P53, and Brca1 or Brca2 Cooperate in Inducing Metastatic Serous Epithelial Ovarian Cancer. *Cancer Res.* **2012**, *72*, 4141–4153. [CrossRef]
95. Ryland, G.L.; Doyle, M.A.; Goode, D.; Boyle, S.E.; Choong, D.Y.H.; Rowley, S.M.; Li, J.; Australian Ovarian Cancer Study Group; Bowtell, D.D.; Tothill, R.W.; et al. Loss of Heterozygosity: What Is It Good For? *BMC Med. Genom.* **2015**, *8*, 45. [CrossRef]
96. Zhang, K.; Lin, G.; Han, D.; Han, Y.; Peng, R.; Li, J. Adaptation of ACMG-ClinGen Technical Standards for Copy Number Variant Interpretation Concordance. *Front. Genet.* **2022**, *13*, 829728. [CrossRef] [PubMed]
97. cBioPortal for Cancer Genomics. Available online: https://www.cbioportal.org/ (accessed on 30 August 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.