Prediction of blood-based biomarkers and subsequent design of bisulfite PCR-LDR-qPCR assay for breast cancer detection

Manny D. Bacolod¹*, Jianmin Huang¹, Sarah F. Giardina¹, Philip B. Feinberg¹, Aashiq H. Mirza¹, Alexander Swistel², Steven A. Soper³ and Francis Barany¹*

Abstract

Background: Interrogation of site-specific CpG methylation in circulating tumor DNAs (ctDNAs) has been employed in a number of studies for early detection of breast cancer (BrCa). In many of these studies, the markers were identified based on known biology of BrCa progression, and interrogated using methyl-specific PCR (MSP), a technique involving bisulfite conversion, PCR, and qPCR.

Methods: In this report, we are demonstrating the development of a novel assay (Multiplex Bisulfite PCR-LDR-qPCR) which can potentially offer improvements to MSP, by integrating additional steps such as ligase detection reaction (LDR), methylated CpG target enrichment, carryover protection (use of uracil DNA glycosylase), and minimization of primer-dimer formation (use of ribose primers and RNaseH2). The assay is designed to for breast cancer-specific CpG markers identified through integrated analyses of publicly available genome-wide methylation datasets for 31 types of primary tumors (including BrCa), as well as matching normal tissues, and peripheral blood.

Results: Our results indicate that the PCR-LDR-qPCR assay is capable of detecting ~ 30 methylated copies of each of 3 BrCa-specific CpG markers, when mixed with excess amount unmethylated CpG markers (~ 3000 copies each), which is a reasonable approximation of BrCa ctDNA overwhelmed with peripheral blood cell-free DNA (cfDNA) when isolated from patient plasma. The bioinformatically-identified CpG markers are located in promoter regions of NR5A2 and PRKCB, and a non-coding region of chromosome 1 (upstream of EFNA3). Additional bioinformatic analyses would reveal that these methylation markers are independent of patient race and age, and positively associated with signaling pathways associated with BrCa progression (such as those related to retinoid nuclear receptor, PTEN, p53, pRB, and p27).

Conclusion: This report demonstrates the potential utilization of bisulfite PCR-LDR-qPCR assay, along with bioinformatically-driven biomarker discovery, in blood-based BrCa detection.

Keywords: Breast cancer, Methylation, Early detection, Ligase detection reaction, Biomarker

Background

In 2019, the projected number of new cases of and deaths due to breast cancer (BrCa) in the United States, are 271,270, and 42,260 respectively [1]. Worldwide, the corresponding numbers (2018 estimate) are 2,088,849, and 626,679 respectively [2]. It is the second-leading cause of cancer death in women, one in 8 of whom will acquire the disease in her lifetime. Although genetic predisposition (i.e. BRCA1/2 mutations) is an important contributing factor (5–10%) [3, 4], most BrCa cases are those without clear genetic link (it may still be due to unknown genetic risk, thus considered familial). While Stage I cases have close to 100% 5-year survival rate, those diagnosed at Stage IV have a 5-year relative survival rate of only 22%, accounting for 6–10% of new BrCa cases and 20–30% all of recurrent disease [5]. The early detection of BrCa saves lives and reduces the morbidity associated with the aggressive treatments required.
for treating late-stage cancers. Nevertheless, the primary diagnostic screening method, mammography, has high rates of false positive and false negatives, can result in over-diagnosis, uses harmful radiation, and is an uncomfortable process for patients [6, 7]. This necessitates the pursuit of molecular markers more indicative of a tumor's biological characteristics translatable to a reliable, non-invasive diagnostic assay. Over the years, there have been numerous reports indicating that either blood serum, plasma, or whole blood can harbor molecular biomarkers indicative of a progressing BrCa [3, 8, 9].

These markers include: secreted proteins (e.g. CA15–3, trefoil factors 1, 2, and 3), auto-antibodies (e.g. antibodies against human endogenous retrovirus-K(HML-2) and heterogeneous nuclear ribonucleoprotein F), lipids (e.g. C16:1, C18:3, C18:2), and microRNAs (e.g. miR-21, miR-221, miR-145). In addition to the blood-based markers mentioned above, there is growing field exploring the use of DNA fragments released by cancer cells (referred to as circulating tumor DNAs or ctDNAs) into the patient’s bloodstream as an indicator of cancer [10, 11]. Previous studies proved that genomic signatures (e.g. mutation, copy number variation, CpG methylation) found in cancer tissues are largely concordant with those identified in ctDNAs [12–18]. Already marketed are early cancer diagnostic tests based on interrogating site-specific CpG hypermethylation in ctDNAs isolated from patient plasma. These include: a) Epi proColon, ColoVantage, Realtime mS9, all of which detect methylation in the SEPT9 gene for colon cancer detection [19]; b) Epi proLung which detects methylation of SHOX2 for lung cancer detection [20], and c) Colvera, which detects methylation at BCAT1 and IKZF1 for colon cancer recurrence [21].

There are important considerations in the development of methylation-based early detection assays for BrCa (or any other cancer type). Although the levels of plasma-derived cell free DNA (cfDNA) in serum from cancer patients are indeed abnormally high in early- to late-stage cancers [22–24], only a small percentage are ctDNAs (most cfDNAs are hematological in origin). Another important concern is the selection of appropriate markers. At the very least, the selected CpG sites should be highly methylated in breast primary tumors (PTs) and practically unmethylated in peripheral blood. However, for a marker to be highly specific to BrCa PTs, it needs to have very low levels of methylation in normal breast tissues, and many other tumor types. In this report, we demonstrate a new and more sensitive assay for methylated CpG detection (incorporating various steps including ligase detection reaction), and a comprehensive approach to biomarker discovery using integrated public genomic datasets.

Methods

Public genomic datasets

Analyzed for this study are various publicly available genomic datasets (Additional file 1: Supplement 1) such as those released by the TCGA project (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga) [25] and those deposited in the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/). The primarily Illumina 450 K methylation array-generated TCGA datasets were previously compiled (and processed) in the UCSC Cancer Genomics website (https://genome-cancer.ucsc.edu/) [26, 27]. The TCGA cohorts included in our analyses are: breast invasive carcinoma [BRCA], adrenocortical carcinoma [ACC], bladder urothelial carcinoma [BLCA], cervical squamous cell carcinoma and endocervical adenocarcinoma [CESC], cholangiocarcinoma [CHOL], colon adenocarcinoma [COAD], lymphoid neoplasm diffuse large b-cell lymphoma [DLBC], esophageal carcinoma [ESCA], glioblastoma multiforme [GBM], head and neck squamous cell carcinoma [HNSC], kidney chromophobe carcinoma [KICH], kidney renal clear cell carcinoma [KIRC], kidney renal papillary cell carcinoma [KIRP], brain lower grade glioma [LGG], liver hepatocellular carcinoma [LIHC], lung adenocarcinoma [LUAD], lung squamous cell carcinoma [LUSC], mesothelioma [MESO], pancreatic adenocarcinoma [PAAD], pheochromocytoma and paranganglioma [PCPG], prostate adenocarcinoma [PRAD], rectum adenocarcinoma [READ], sarcoma [SARC], skin cutaneous melanoma [SKCM], stomach adenocarcinoma [STAD], testicular germ cell tumors [TGCT], thymoma [THYM], thyroid carcinoma [THCA], uterine corpus endometrial carcinoma [UCEC], uterine carcinosarcoma [UCS], and uveal melanoma [UVM]. Also crucial to our biomarker identification is the integration of various GEO datasets such as: GSE65820 (ovarian cancer PTs and matching normals) [28], GSE46306 (normal tissues of the cervix) [29], GSE99553 (gastric mucosa), GSE74104 (testis) [30], GSE77871 (adrenal tissues), GSE51954 (dermis and epidermis) [31], GSE64509 (various brain tissues) [32], GSE42861 (peripheral blood) [33], and GSE99553 (gastric mucosa), GSE74104 (testis) [30], GSE77871 (adrenal tissues), GSE51954 (dermis and epidermis) [31], GSE64509 (various brain tissues) [32], GSE42861 (peripheral blood) [33], and GSE99553 (gastric mucosa) [28], GSE46306 (normal tissues of the cervix) [29], GSE99553 (gastric mucosa), GSE74104 (testis) [30], GSE77871 (adrenal tissues), GSE51954 (dermis and epidermis) [31], GSE64509 (various brain tissues) [32], GSE42861 (peripheral blood) [33], and GSE99553 (gastric mucosa) [28].

The methylation data for BrCa cell lines were extracted from the GEO datasets GSE57342 [35], GSE68379 [36], GSE78875 [37], and GSE94943.

Bioinformatic and statistical analyses

Programs and other tools
All statistical analyses (comparative statistics, normalization, correlation and regression analyses, multivariate analyses, hierarchical clustering) were performed using JMP Pro 11/ IMP Genomics software (SAS, Cary, NC), and Gene-E (Broad
Institute, Cambridge, MA). Genomic sequence extraction and alignment were performed through the UCSC Genome Browser (https://genome.ucsc.edu/) [38]. The OligoAnalyzer Tool from IDT (https://www.idtdna.com/pages/tools/oligoanalyzer) aided in our primer designs.

**Prediction of blood-based, breast cancer-specific methylation markers** The primary task was to identify CpG markers whose high degree of methylation in blood cfDNA may be indicative of primary BrCa, but (as much as possible) not of normal breast cells, peripheral blood, and many other cancer types. There are multiple approaches which can be employed to accomplish this task. Our approach started with simplifying each subgroup (e.g., PTs, solid normals) of a given cohort (e.g. TCGA-BRCA, TCGA-COADREAD), into arbitrarily defined metrics. In this case, the CpG marker $P$ for a given cohort $C$, and cohort subset $S$ was associated with specific statistical value $V$, such as $\%UM$, $\%IM$, $\%LM$, $\%HM$, $\%UM + \%IM$, and $\%LM + \%HM$, wherein $UM, IM, LM, HM$ respectively refer to UnMethylated ($\beta_p \leq 0.15$), $\beta = \text{fraction of methylation at marker } P$, Indeterminately Methylated ($0.15 < \beta_p \leq 0.3$), Lowly Methylated ($0.3 < \beta_p \leq 0.6$), and Highly Methylated ($\beta_p > 0.6$) at the marker $P$. The candidate markers ($P's$) were dynamically isolated by identifying CpG sites which satisfied multiple criteria in the general form: $V_{P(C, S)} > n; \{0 \leq n \leq 100\}$.

**Assessment of methylation markers’ relationship with various clinico-pathological data** The accompanying clinico-pathological (such as pathological stage, PAM50 molecular subtype, age, race, and ethnicity) were used to assess their influence in the CpG markers’ $\beta$ values.

**Assessment of methylation markers’ roles in epigenetic regulation and other functionalities** The genome-wide transcriptional data (RNASeq-generated) for the BRCA cohort was also integrated into the analyses. First, the combined transcriptional and methylation data enabled us to predict if the CpG sites can potentially influence the transcription of their respective genes. Second, Gene Set Enrichment Analysis [39] was employed to predict the genes or pathways (http://software.broadinstitute.org/gsea/msigdb/gene sets.jsp) that tend to be enriched (or upregulated) when a particular CpG site is highly methylated (such as some of the CpG sites which will eventually be verified experimentally).

**Cell lines and genomic DNAs**

The BrCa cell lines SKBr3, MDA-MB-134VI, and MCF7, which would serve as sources of cancer genomic DNAs (gDNAs), were grown according to culture conditions recommended by ATCC (https://www.atcc.org/). At 80–90% confluence, the cells were washed with Phosphate Buffered Saline ($\times 3$), and collected by centrifugation (500 x g). gDNAs were isolated using the DNeasy Blood & Tissue Kit (Qiagen; Valencia, CA). gDNA ($>50$ kb size) isolated from blood (buffy coat) of healthy individuals was purchased from Roche (Indianapolis, IN) (also referred to as “Roche DNA”), Quant-iT Picogreen Assay (Life Technologies/Thermo Fisher; Waltham, MA) was used to determine gDNA concentration. The isolated gDNAs were then fragmented (50 bp to 1 kb size) using an ultra-sonicator from Covaris (Woburn, Massachusetts). The fragmentation size was assessed using the Agilent Bioanalyzer System.

**Enrichment of methylated genomic DNA**
The gDNA fragments containing CpG methylated fragments were enriched using the EpiMark® Methylated DNA Enrichment Kit (New England BioLabs, Ipswich, MA). This approach uses selective binding of double-stranded methyl-CpG DNA to the methyl-CpG binding domain of human MBD2 protein fused to the Fc tail of human IgG1. The fused IgG1 (MBD2-Fc) antibody is coupled to paramagnetic hydrophilic protein A magnetic beads. The enrichment procedure was carried out according to the manufacturer’s instructions.

**Bisulfite conversion of digested genomic DNA**

Bisulfite conversion of cytosine bases was accomplished using the EZ DNA Methylation-Lightning kit from Zymo Research Corporation (Irvine, CA). In brief, 130 μl of Lightning Conversion Reagent was added to 20 μl of previously enriched gDNA fragments. Subsequent protocol steps (according to the manufacturer’s instructions) led to elution of bisulfite converted DNA fragments in 10 μl of elution buffer.

**PCR-LDR-qPCR**
The assay we developed for detection of plasma-based BrCa methylation markers is divided into several steps described in following subsections. All primers (Additional file 1: Supplement 2) were purchased from Integrated DNA Technologies Inc. (Coralville, IA).

**Linear amplification**

In a 25 μl of reaction volume, the linear amplification step was carried out by mixing: 5.0 μl of corresponding bisulfite converted DNA template (out of 50 μl of eluted DNA after bisulfite conversion), 5 μl of 5x GoTaq Flexi buffer (no Magnesium) (Promega, Madison, WI), 2.5 μl of 25 mM MgCl₂ (Promega, Madison, WI), 0.5 μl of 10 mM dNTPs (dATP, dCTP, dGTP and dTTP) (Promega, Madison, WI), 2.5 μl of the reverse primer (or primers in case of multiplex reaction) (1 μM), 0.625 μl of 20 mM/μl RNAseH2 (diluted in RNAseH2 dilution buffer from IDT) (IDT), and 0.55 μl of KlenTaq polymerase (DNA
Polymerase Technology, St. Louis, MO) mixed with Platinum Taq Antibody (Invitrogen/Thermo Fisher, Waltham, MA). The reactions were run in a ProFlex PCR system thermocycler (Applied Biosystems/Thermo Fisher, Waltham, MA) with the following program: 2 min at 94 °C, 40 cycles of (20 s at 94 °C, 40 s at 60 °C, and 30 s at 72 °C), and a final hold at 4 °C. After the reaction, Platinum Taq antibodies were added in the reaction mixture to inhibit the KlenTaq DNA polymerase. The KlenTaq/Platinum Taq Antibody mixture was prepared by adding 0.02 μl of KlenTaq polymerase at 50 μl to 0.2 μl of Platinum Taq Antibody at 5 μl/μl.

**PCR**

For the PCR reaction, 10 μl of linear amplification product (previous step) was mixed with 2 μl of 5X GoTaq Flexi buffer without Magnesium, 1 μl of 25 mM MgCl₂, 0.4 μl of dNTPs (10 mM each of dATP, dCTP, dGTP and dUTP), 2 μl of 0.5 μM forward primer (or primers in case of multiplex reaction), 0.4 μl of Antarctic Thermolabile UDG (1 U/μl) (New England Biolabs, Ipswich, MA), 0.25 μl of 20 μM/μl RNAseH2, 0.44 μl of KlenTaq polymerase mixed with Platinum Taq Antibody (Invitrogen/Thermo Fisher, Waltham, MA). The KlenTaq/Platinum Taq Antibody mixture was prepared by adding 0.02 μl of 50 U/μl KlenTaq polymerase to 0.2 μl of 5 U/μl Platinum Taq Antibody. The 20 μl-volume reactions were run in a ProFlex PCR system thermocycler, using the following program: 10 min at 37 °C, 40 cycles of (20 s at 94 °C, 40 s at 60 °C and 30 s at 72 °C), 10 min at 99.5 °C, and a final hold at 4 °C.

**LDR**

The LDR step was performed in a 20 μl reaction prepared by combining: 5.82 μl of nuclease-free water (IDT), 2 μl of 10X AK16D ligase reaction buffer 0.5 μl of 40 mM DTT (Sigma-Aldrich, St. Louis, MO), 0.25 μl of 40 mM NAD⁺ (Sigma-Aldrich, St. Louis, MO), 0.5 μl of 20 μM/μl RNAseH2, 0.4 μl of 500 nM LDR upstream probes, 0.4 μl of 500 nM LDR downstream probes, 0.57 μl of purified AK16D ligase (at 0.88 μM), and 4 μl of PCR reaction products from previous step. The AK16D ligase reaction buffer (at 1X) contains the following: 20 mM Tris-HCl at pH 8.5, 5 mM MgCl₂, 50 mM KCl, 10 mM DTT, and 20 μg/ml of BSA (all components purchased from Sigma Aldrich, St. Louis, MO). LDR reactions were run in a ProFlex PCR system thermocycler using the following program: 20 cycles of (10 s at 94 °C, and 4 min at 60 °C) followed by a final hold at 4 °C.

**Taqman real-time qPCR**

The qPCR reaction was performed in a 10 μl of reaction mixture prepared by mixing: 1.5 μl of nuclease-free water (IDT), 5 μl of 2X TaqMan® Fast Universal PCR Master Mix (Fast AmpliTaq, UDG and dUTP)(Applied Biosystems/ThermoFisher; Waltham, MA), 1 μl 2.5 μM forward primer at, 1 μl of 2.5 μM reverse primer, 0.5 μl of 5 μM probe, and 1 μl of LDR reaction products from the previous step. All qPCR reactions were run in a ViiA7 real-time thermo-cycler from Applied Biosystems (Applied Biosystems/ThermoFisher; Waltham, MA), using MicroAmp® Fast-96-Well Reaction 0.1 ml plates sealed with MicroAmp® Optical adhesive film (Applied Biosystems/ThermoFisher; Waltham, MA). The run settings were as follows: fast block, Standard curve as experiment type, ROX as passive reference, TAMRA as reporter, and NFQ-MGB as quencher; program at 2 min at 50 °C, and 40 cycles of (1 s at 95 °C, and 20 s at 60 °C).

**Taqman digital qPCR**

For each digital PCR reaction, a 20 μl mixture was prepared in each of the 96 well digital PCR microplate. The mixture included 2 μl of diluted LDR product (Step 3), 1X Luna Universal ProbeqPCR master mixture, 0.1% tween 20, 0.4 μM RNAseH2, 0.025 U Antarctic Thermolabile UDG, 5 μM each of forward and reverse primers, and Taqman probe. 12 μl of reaction mixture was loaded into the Constellation Digital PCR System (originally FormulaTrix, Bedford, MA; currently Qiagen), and run with the following conditions: 37 °C for 10 min, 95 °C for 20 s, and 45 cycles of 5 s (94 °C), and 20 s (60 °C).

**Results**

**Identification of potential blood-based breast cancer (BrCa) methylation markers**

The BrCa methylation markers were identified (See Fig. 1) by applying the following filters: a) %HM (BrCa, PT) ≥ 30; b) %UM (BrCa, SN) ≥ 40; c) Average (%HM) (C, S) ≥ 40; C = any cancer cohort ≠ BrCa, S = PTs, d) Ave. (%UM) (C, S) ≥ 40; C = any cancer cohort ≠ BrCa; S = Solid normals; and e) % UM (GSE42861,blood) ≥ 98. A total of 229 CpG sites passed these filters. Among these are located in the loci of CPXM1, RASSFIA, and SC3BGA1, which have been reported in the literature to be indicative of BrCa [3]. Selected for further studies are the CpG sites located in the promoter regions of NR5A2 (nuclear receptor subfamily 5 group A member) (referred to as “m_NR5A2) and PRKCB (protein kinase C beta) (m_PRKCB), as well as a CpG site in a non-coding region of chromosome 1 (m_ncr1) (Additional file 1: Supplement 3). The 3 aforementioned CpG sites were selected after further statistical inspections (i.e. the ones most likely to differentiate BrCa with other major cancer types), and assessment of surrounding sequences (i.e. to make sure assay can be designed around the target CpG sites). Of the 3 markers, m_ncr1 appears to be most highly specific to BrCa PTs (Fig. 2, Additional file 1: Supplement 4). The average methylation (β) of m_ncr1, among BRCA PTs is...
0.536 (on the scale 0 to 1, with 1 being 100% methylated). For other major cancer types such as ovarian cancer (OV), endometrial cancer (UCEC), colorectal cancer (COADREAD), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), and pancreatic cancer (PAAD), the \( \beta \) values for PTs (or \( \beta_{PT} \)) range from 0.030 to 0.093. In contrast, the \( \beta \) for normal tissues (or \( \beta_{SN} \)) is 0.206 for normal breast, and ranged from 0.028 to 0.066 in other tissue types. As shown in Fig. 2 and Additional file 1: Supplement 1, m_ncr1 \( \beta_{PT} \) values for Stage I, II, and III are 0.499, 0.534, and 0.590 respectively (Additional file 1: Supplement 5). For m_NR5A2, these respective values are 0.552, 0.610, and 0.662. For the m_PRKCB marker, while Stage III PTs registered the highest \( \beta_{PT} \) (0.611), the \( \beta_{PT} \) value for Stage I (0.562) is slightly higher than that of Stage II (0.548). Nonetheless, the most statistically significant difference, across all markers, is between Stage I and Stage III PTs. Another noteworthy observation is that in all 3 markers, the Basal subtype registered the lowest \( \beta_{PT} \) values. For example, the m_ncr1 \( \beta_{PT} \) values for Basel, Luminal A (LumA), Luminal B (LumB), and HER2 tumors are 0.440, 0.532, 0.628, and 0.630 respectively (Additional file 1: Supplement 6). We did not find a substantial correlation (\( R^2 \) values range from 0.0 to 0.04) between patient age and methylation at each marker (irrespective of whether the tissue sample is a PT or solid normal) (See Additional file 1: Supplement 7). In addition, the methylation at each of the marker is independent of the racial (whether Asian, African American, or White) or ethnic (whether Hispanic or non-Hispanic) classification of the patient (Additional file 1: Supplements 8 and 9).

The predicted biological roles of BrCa-specific methylation markers

The methylation level of either m_NR5A2 or m_PRKCB, apparently does not correlate with its transcript level (the respective \( R \) values are −0.198 and −0.123) (Additional file 1: Supplement 10). Nonetheless, it is still possible to bioinformatically demonstrate the association of these two CpG sites (along with m_ncr1, which is not part of the transcribed gene) with BrCa progression. It was accomplished in several ways. The first approach (See Fig. 3) was to identify the genes whose expression levels are most highly correlated with the methylation of the select markers. Indeed, a number of genes whose transcription positively and negatively correlated with each of the markers are known oncogenes and transcription factors respectively. For example, the expression of the transcription factor FOXA1, is highly (and positively) correlated with the methylation levels at m_PRKCB and m_NR5A2 (Fig. 3a and b). This transcription factor is needed for the transactivation of p27{\textsuperscript{kip1}} [40], which alone or in conjunction with BRCA1 may control gene expression pattern in luminal subtypes of BrCa [41]. Similarly, the expression of SPDEF has a high positive correlation with m_PRKCB and m_ncr1(Fig. 3a and c). Previous studies indicated that SPDEF (also a transcription factor) is upregulated in all al subtypes of BrCa (except basal), and is especially associated with HR

Relationship of select CpG markers with other clinicopathological profiles

It is also possible to examine the various clinicopathological information which may be associated with the methylation markers’ \( \beta \) values. As expected, more progressive tumors tend to have higher \( \beta_{PT} \) values. For example, the m_ncr1 \( \beta_{PT} \) values for Stage I, II, and III are 0.499, 0.534, and 0.590 respectively (Additional file 1: Supplement 5). For m_NR5A2, these respective values are 0.552, 0.610, and 0.662. For the m_PRKCB marker, while Stage III PTs registered the highest \( \beta_{PT} \) (0.611), the \( \beta_{PT} \) value for Stage I (0.562) is slightly higher than that of Stage II (0.548). Nonetheless, the most statistically significant difference, across all markers, is between Stage I and Stage III PTs. Another noteworthy observation is that in all 3 markers, the Basal subtype registered the lowest \( \beta_{PT} \) values. For example, the m_ncr1 \( \beta_{PT} \) values for Basel, Luminal A (LumA), Luminal B (LumB), and HER2 tumors are 0.440, 0.532, 0.628, and 0.630 respectively (Additional file 1: Supplement 6). We did not find a substantial correlation (\( R^2 \) values range from 0.0 to 0.04) between patient age and methylation at each marker (irrespective of whether the tissue sample is a PT or solid normal) (See Additional file 1: Supplement 7). In addition, the methylation at each of the marker is independent of the racial (whether Asian, African American, or White) or ethnic (whether Hispanic or non-Hispanic) classification of the patient (Additional file 1: Supplements 8 and 9).
phenotype and nodal metastasis [42]. ERBB3 expression, which is positively correlated with methylation at m_PRKCB and m_NR5A2 (Fig. 3a and b), can form a heterodimer with ERBB2, and subsequently activates downstream signaling leading to cell cycle progression [43]. Another transcription factor, GATA3 is positively correlated with both m_PRKCB and m_NR5A2 (Fig. 3a and b). Upregulated in all BrCA subtypes (except basal), GATA3 is associated with ER positive BrCa (shown to regulate the expression of ER gene itself) [44]. In contrast, many of the genes whose transcript levels are negatively correlated with the methylation at the 3 CpG sites are known tumor suppressor genes. These include the genes CDC14B and SPRY2. CDC14B is a protein tyrosine phosphatase which can regulate RNA Pol II and repress cell cycle transcription [45]. SPRY2 acts as tumor suppressor in BrCA by inhibiting the Ras/Mitogen-Activated Protein Kinase Pathway [46], and in ovarian cancer, through inhibition of Amphiregulin (AREG)-induced cell invasion [47].

Another approach to bioinformatically justify the potential of methylation markers in cancer detection is through genome-wide transcriptional comparison between two BrCa PT subsets: those highly methylated (H), and those lowly methylated (L) at a particular CpG site p. The two subsets are defined as the follows: a) H subset includes PTs with $\beta_p$ values higher by at least 1 $\sigma$ from the mean ($\beta_{mean}$), and b) L subset includes BrCa PTs with $\beta_p$ values lower by at least 1 $\sigma$ from the mean ($\beta_{mean}$). Gene Set Enrichment Analysis (GSEA) [39] was then employed to identify the molecular pathways, as defined in Biocarta (https://www.gsea-msigdb.org/gsea/msigdb/gene sets.jsp?collection=CP:BIOCARTA) or Reactome (https://reactome.org/) [48] databases, most likely enriched in H over L subsets. Results from this analysis would indicate that gene sets associated with retinoid nuclear receptor signaling, are highly enriched in H over L subsets. These include: RARRXR (Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription in carcinoma cells), VDR (Control of Gene Expression by Vitamin D Receptor), and CARM1 (Transcription Regulation by Methyltransferase of CARM1). In the RARRXR pathway, the binding of retinoids to RARA/RXRA nuclear receptor heterodimer initiates the transcriptional activation of cell proliferation-associated phosphatidylinositol 3-kinase (PI3K) [49]. Other studies point to RARRXR pathway’s association with estrogen signaling [50, 51]. RARA/RXRA nuclear receptor plays a role in epigenetic regulation of Vitamin D Receptor (VDR). Previous studies have demonstrated VDR pathway’s association with BrCa proliferation [52].
positive association between the methylation markers and CARM1 pathway is consistent with previous studies indicating that CARM1 (which codes for arginine methyltransferase) is involved in epigenetic transactivation of many nuclear receptors (NRs) including ERα [53]. CARM1 is also associated with poor prognosis in BrCa [54].

Also positively associated with at least one of the 3 methylation markers is the upregulation of cancer proliferation signaling pathways including: PTEN (PTEN-dependent cell cycle arrest and apoptosis), P53 (p53 Signaling Pathway), P27 (Regulation of p27 Phosphorylation during Cell Cycle Progression), RB (RB Tumor Suppressor/Checkpoint Signaling in response to DNA damage), ATRBRCA (Role of BRCA1, BRCA2 and ATR in Cancer Susceptibility), and mTOR (mTOR Signaling Pathway). The enrichment of the aforementioned cancer pathways (in H over L PT subsets) may be explained by the upregulated expression of genes comprising these gene sets. These genes include: a) MAPK3, PDK2, and PTEN, of the PTEN gene set; b) BCL2, CCND1, RB1, TIMP3, and MDM2 of the P53 gene set; c) E2F1 and CKS1B, of the P27 gene set; d) CDC25C, MYT1, CDK1, and CD25B of the RB gene set; e) RAD51, BRCA1, FANCd2, and BRCA2, of the ATRBRCA gene set; and f) EIF4EBP1 of the mTOR gene set.

**Multiplex bisulfite PCR-LDR-qPCR assay to validate the 3 breast cancer methylation markers**

In essence, the analyses described above rationalized that the methylation level at 3 CpG sites (m_ncR1, m_NR5A2, and m_PRKCB) can serve as serum markers for progressing BrCa. With the aim of translating these findings to clinical use, we tested a homegrown procedure (bisulfite PCR-LDR-qPCR assay) (Fig. 4) designed to interrogate these methylated markers present in minuscule amount (i.e. down to less than 50 copies), as it is expected in cfDNA isolated from ~10-ml patient blood. The assay was tested on simulated cfDNA sample consisting of a mixture of gDNA fragments from BrCa cell line and peripheral blood from healthy patients, with the fragments from the latter present at ~100-fold relative to the former. It is a fair assumption that many of the BrCa cell
lines will be highly methylated at the 3 CpG sites. Nonetheless, given the availability of genome-wide methylation datasets for many of the commonly used cancer cell lines, this assumption can easily be verified. According to the datasets GSE57342, GSE68379, GSE78875, and GSE94943, the average $\beta$ values for the CpG markers m_NR5A2, m_PRKCB, and m_ncr1 in the BrCa cell line MCF7 are 0.96, 0.97, and 0.98 respectively (Fig. 5). For the cell line MDA-MB-134-VI, the respective numbers are 0.98, 0.74, and 0.99. With this information, the gDNAs from these cell lines were isolated, fragmented, and enriched for methylation (through the use of fusion protein with the methyl-CpG binding domain, which can selectively bind DNA fragments containing
methylated CpG sites). As shown in the scheme (Fig. 4), the simulated cfDNA (~30 and 3000 copies of cell line and peripheral DNA fragments respectively), was bisulfite-converted, then subjected to two-step PCR. The resulting amplicon would then serve as template for ligation detection reaction (LDR). The final step to detect the methylation signal was Taqman-based real time PCR, using the LDR products (from the previous step) as templates. The Ct values for m_NR5A2, m_PRKCB, and m_ncr1 were 8.96, 11.38, and 14.75 respectively for MCF7, and 11.38, 11.73, and 12.70 respectively for MDA-MB-134-VI. No Ct values registered for the unmixed peripheral blood for each of the marker. Indeed, the designed assays were capable of detecting minute amount of the BrCa methylation markers. These assays were designed to reduce false positive signals by minimizing non-specific amplifications. Among the important features are the 8–11 base tails in the reverse PCR primers, meant to reduce the possibility of primer dimer formation. Also present were the ribose bases at the 3’ end of the PCR and LDR primers. These sequences were removed by RNaseH2 when the primers were bound to their targets, ensuring highly specific extension (for PCR) and ligation (for LDR). Carryover prevention was the objective of the timely addition of UDG, dUTP, or dTTP in the dNTP mixture [55]. Also, the tags in LDR primers (Ai, Ci) allowed for uniformity (in terms of Tm) in qPCR reactions.

**Digital PCR as the final detection step**

Instead of Taqman PCR, digital PCR may also be used as detection step. For this experiment, we performed the 2-step PCR and LDR reactions as described above, using the primers for the detection of methylation at the CpG site located in the promoter region of the gene GRK7 (m_GRK7 or cg18768784; Chr3:141516271–141,516,272), although highly methylated in the BrCa cohort, has low BRCA-specificity. Nonetheless, there is a high degree of concordance between detection based in qPCR (Fig. 6a) and detection using the Constellation (Formula- trix) digital PCR System (Fig. 6b). The bisulfite-converted DNA fragments consisting of ~3000 copies of unmethylated m_GRK7 (from peripheral blood) and ~
30 copies of methylated m_GRK7 yielded Ct values of 7.8 and 9.4 for MCF7 and MDA-MB-134VI respectively. The unmixed template (DNA fragments from peripheral blood only) did not register readable Ct. In the digital PCR detection system, the corresponding readings are 8164, 4986, and 805 for the MCF7, MDA-MB-134VI, and control (peripheral blood only) respectively.

**Discussion**

The limitations of mammography are what drives the persistent efforts towards developing non-invasive screening approaches for early BrCa detection. Falling under the term “liquid biopsy”, many of the methods under investigation are technologies which aim to detect blood-based molecular markers originating from BrCa cells. The molecular markers can include cfDNA fragments, exosome-enclosed or naked RNA molecules, secreted proteins and metabolites [8, 9].

Of particular interest in the early-cancer detection field are circulating tumor DNAs (ctDNAs), which apoptotic and necrotic cancer cells release into patient plasma [56]. As expected, ctDNA fragments possess the same molecular signatures (somatic mutations, methylation, copy number variation/aberration, SNPs) present in gDNAs isolated from the tumor tissue samples. Hence, molecular characterization tools normally used to investigate cancer gDNAs (such as exome or genome-wide sequencing, PCR, DNA arrays, methylation arrays) have also been applied in ctDNA analysis [57]. What makes ctDNA analysis especially challenging is the fact that when isolated from patient plasma, ctDNAs are mixed with an overwhelming amount of DNA fragments that are hematopoietic in origin [58–60]. All of the fragments are collectively referred to as cell-free DNAs (or cfDNAs). According to a recent study, the ctDNAs originating from BrCa cells is just a small fraction of total cfDNAs [24]. This is based on the observation that the mutant allele fraction (MAF; from sequencing 58 cancer-related genes) of cfDNAs isolated from BrCa patients is less than 1% [24]. It is imperative that the assay...
employed to analyze cfDNAs is capable of distinguishing between the positive (several copies of cfDNAs) and mostly negative (from non-cancer cfDNAs) signals. This limitation of ctDNA analyses can be circumvented through the identification of more appropriate molecular biomarkers, along with the modification of assay biochemistry towards higher sensitivity and specificity. Although plasma-based ctDNA markers may include markers for mutations, methylation states, or copy number variations (most reports interrogate methylation and mutation markers), methylation markers have several inherent advantages. First, methylation changes are tissue-specific [61], thus as markers, would make them highly capable of distinguishing one cancer type from another. Another advantage of CpG methylation over mutation is that oftentimes the methylation changes adjacent CpG sites in promoter regions, are concordant. Methylation-dependent procedures (such as the use of methyl DNA-binding antibodies) would then be more effective in enriching the fragments containing the highly methylated markers.

Identification of appropriate methylation markers (i.e. particular CpG sites) is very crucial. To pinpoint the specific CpG sites that can easily distinguish BrCa tissues from peripheral blood and other types of cancer, we took advantage of the availability of various genomewide methylation datasets. As previously pointed out, these calculations resulted in identification of 229 potential CpG markers which included CpG sites at the locus of RASSF1A (Ras association domain-containing protein 1), which happens to be the most highly reported blood-based methylation markers for breast cancer [62–69]. Additional statistical inspections and assay design considerations would then point to the selection of the 3 CpG markers we focused on for this manuscript. Two of the CpG sites (m_NR5A2 and m_PRKCB) are located in promoter regions of genes, with reported link to breast cancer. NR5A2 (or LRH1) is a zinc finger transcription factor which can regulate CDKN1A expression in BrCa [70], and has been positively associated with BrCa proliferation [71], drug resistance [72], aggressiveness [73], high grade, and poor outcome [74]. On the other hand, the role of PRKCB in breast cancer progression is still not clearly defined. While there are reports indicating that PRKCB can promote mammary tumorigenesis [75], enhance breast cancer cells growth and cyclin D1 expression [76], and has the potential as therapeutic target [77], there is also a report indicating it may inhibit tumor growth and metastasis [78]. The third CpG site interrogated by our assay (m_nrc1) is located less than 8000 bp upstream of the exon 1 (according to GENCODE v31 annotation) of the protein coding gene EFNA3, a member of the ephrin (EPH) family. Whether this particular CpG site influences the expression of EFNA3 protein, or the hypoxia-related EFNA3 lncRNA [79, 80] is not clear at this point.

Interestingly, the methylation level at m_NR5A2 and m_PRKCB did not correlate with the transcription of the corresponding genes (Additional file 1: Supplement 10). However, it is important to note that CpG methylation (at the promoter region) is not the only factor that influences gene transcription. It is quite possible that histone modification [81], regulatory miRNA or ncRNAs [82], as well as transcription factors can supersede CpG methylation in influencing transcription. The competition between mRNA transcription and mRNA degradation is a dynamic process that can determine the transcript level of a gene at any given time [83]. Regardless of m_NR5A2 and m_PRKCB CpG sites’ influence on their respective transcript levels, their association with BrCa progression is quite clear. This is further demonstrated through comparative genome-wide transcription analyses (which is essentially what GSEA is) of BrCa samples that are highly and lowly methylated at each CpG site. As shown in our analyses, the methylation level at each of the three methylation markers (m_nr51, m_NR5A2, and m_PRKCB), is positively associated with genes, processes, and pathways indicative of BrCa progression. These include processes (and much of the component genes) associated with the retinoid nuclear receptor, PTEN, p53, p27, RB, and MTOR signaling pathways.

A great majority of reports on the interrogation of CpG methylation in cfDNA for BrCa detection employed the methylation-specific PCR (MSP) approach. Aside from RASSFI, other genes whose CpG sites were observed to be hypermethylated in BrCa patient-derived cfDNAs (through MSP approach) are: AKR1B1, ARHGEF7, BRCA1, BRMS1, COL6A2, CST6, CDKN2A, CCND2, DKK3, ESRI, GATA3, GPX7, GSTP1, HOXD13, HIST1H3C, HOX84, ITIH5, KLK10, MSH2, MLH1, NBP1I, P16, PCDHB7, RARB, RBGRF2, SOX17, SLIT2, SFN, SFRP1, SOX17, TM6SF1, TMEFF2, TRIM9, and WNTSA [84] [62, 64–69, 85–92]. The aforementioned CpG markers were selected because the genes have known roles in BrCa progression (primarily as tumor suppressors), or were previously identified from the use of earlier, much less dense version of Illumina methylation array (27 K).

Bisulfite conversion is perhaps the most crucial step in MSP. However, bisulfite conversion can cause the degradation of around 84–96% of the input cfDNA, and is thus a significant contributing factor to MSP’s limitations in liquid biopsy [93]. This is not an issue in analyzing gDNAs extracted from tissues and cell lines, which the MSP assay was originally intended for. In some reports, BrCa patient cfDNAs were analyzed through methylated CpG digestion (e.g. BstUI enzyme), followed by qPCR, with no bisulfite conversion step in the
protocols [94–96]. However, results using this approach are not reliable (higher rates of false positives) if there is incomplete digestion of unmethylated CpG sites.

The assay we are proposing incorporated several features which can collectively improve the MSP approach. These include the following: a) selective enrichment of methylated DNA, through the methylated CpG capture by using methyl-DNA binding protein, b) signal amplification of the targeted CpG site through successive steps of bisulphite PCR and LDR, (c) prevention of non-specific primer extension by incorporating RNaseH2-targeted ratiocination of the targeted CpG site through successive steps by using methyl-DNA binding protein, b) signal amplification of methylated DNA, through the methylated CpG capture tures which can collectively improve the MSP approach. b) signal amplification of methylated DNA, through the methylated CpG capture tures which can collectively improve the MSP approach.

Bisulphite sequencing is capable of interrogating more CpG markers compared to site-specific bisulphite conversion assays [97–102]. However, we can only assume that the primary problems in MSP assays (the low abundance of cfDNAs and of target methylated CpG markers) are also encountered in bisulphite sequencing approaches. These factors, along with high cost, limits the recovery of information from bisulphite sequencing of cfDNA fragments [103].

Conclusion
This report demonstrates the steps which can be utilized to improve blood-based early detection assays for BrCa detection, including bioinformatic identification and characterization of the biomarkers and improvements in assay biochemistry. Understandably, the assays were only tested on simulated cfDNA samples. However, in future studies, we will evaluate the non-invasive BrCa diagnostic capability of our multiplex PCR-LDR-qPCR assay through analysis of patient-derived cfDNAs.

Supplementary information
Supplementary information accompanies this paper at https://doi.org/10.1186/s12885-020-6574-4.

Abbreviations
BrCa: Breast cancer; cfDNA: Cell-free DNA; ctDNA: Circulating tumor DNA; gDNA: Genomic DNA; LDR: Ligase detection reaction; PT: Primary tumor

Acknowledgements
The authors thank David Geifland, Bill Efcavitch, Olivier Elemento, Bert Gold, Steve Lipkin, Cristian Ruiz Rueda, Chris Mason, Ken Oftt, Pat Paty, Bernard Peperstraete, Mark Pochapin, and Eugene Spier for helpful discussions.

Authors’ contributions
MDB is the primary writer, conducted data analysis, and contributed to conceptualization and design of the project. JW conducted the experiments and contributed to project design and manuscript writing. SFG contributed to manuscript writing. PBF contributed to experimental design and data analysis. AHM contributed to data analysis. AS contributed to conceptualization and design of the project. SAS contributed to conceptualization and design of the project, as well as manuscript writing. All of the authors have read and approved the manuscript.

Authors’ information
Not applicable.

Funding
This study was supported by the following: (i) Research Grant from Earlier.org, Friends for an Earlier Breast Cancer Test (MDB); (ii) Weill Cornell Medicine funding through distribution of royalties from intellectual property generated by the Barany laboratory; (iii) A Sponsored Research Agreement between AcuamarkDx and Weill Cornell Medicine; (iv) The Biotechnology Resource Center of Biomedical Multi Scale Systems CBM2 for Precision Molecular Diagnostics, NIH grant P41 EB020594; and (v) Research Grant from Laura Candall Brown Foundation and The Foundation for Women’s Cancer (SFG). All of the aforementioned funding bodies were not involved in the design of the study, collection, analysis, and interpretation of data; and writing of the manuscript.

Availability of data and materials
We did not generate new datasets for this article.

Ethics approval and consent to participate
The study did not involve human participation and is not subjected to ethical approval and consent.

Consent for publication
The study did not involve human participation and is not subjected to consent for publication.

Competing interests
MDB, AHM, JH, PBF, SFG and FB are shareholders in AcuamarkDx.

Author details
1Department of Microbiology and Immunology, Weill Cornell Medicine, New York, NY 10065, USA. 2Department of Surgery, Weill Cornell Medicine, New York, NY 10065, USA. 3Department of Mechanical Engineering, The University of Kansas, Lawrence, KS 66047, USA.

Received: 23 September 2019 Accepted: 23 January 2020
Published online: 31 January 2020

References
1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019; 69(1):7–34.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.
3. Tang Q, Cheng J, Cao X, Surowy H, Burewinkel B. Blood-based DNA methylation as biomarker for breast cancer: a systematic review. Clin Epigenetics. 2016;8:115.
4. Antoniou A, Pharoah PD, Narod S, Risch HA, Eystjørd JE, Hopper JL, Loman N, Olsson H, Johannsson O, Borg A, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series.
unselected for family history: a combined analysis of 22 studies. Am J Hum Genet. 2003;73(2):1117–30.

5. O'Shaughnessy J. Extending survival with chemotherapy in metastatic breast cancer. Oncologist. 2005;10(Supp1):2–9.

6. Welch HG, Prikol PC, O'Malley AJ, Kramer BS. Breast-Cancer tumor size, Overdiagnosis, and mammography screening effectiveness. N Engl J Med. 2016;375(11):1438–47.

7. Nelson HD, Pappas M, Cantor A, Griffin J, Daeges M, Humphrey LH. Harms of Breast Cancer Screening. Systematic Review to Update the 2009 U.S. preventive services task force recommendation. Ann Intern Med. 2016;164(2):256–67.

8. Loke SY, Lee ASG. The future of blood-based biomarkers for the early detection of breast cancer. Eur J Cancer. 2018;92:54–68.

9. Alimirzaie S, Bagherzadeh M, Akbari MR. Liquid biopsy in breast cancer: a comprehensive review. Clin Genet. 2019.

10. Rohanizadegan M. Analysis of circulating tumor DNA in breast cancer as a diagnostic and prognostic biomarker. Cancer Genet. 2018;228:29159–68.

11. Buono G, Gerratana L, Bulfoni M, Provinciali N, Basile D, Giuliano M, Corvaja C, Apino G, Del Mastro L, De Pacido S, et al. Circulating tumor DNA analysis in breast cancer: is it ready for prime-time? Cancer Treat Rev. 2019;73:73–83.

12. Kato S, Schwaedler MC, Fanta PT, Okamura R, Leichman L, Lippman SM, Lannan RB, Raymond VM, Talasaz A, Kurockz N. Genomic assessment of blood-derived circulating tumor DNA in patients with colorectal cancers: correlation with tissue sequencing, therapeutic response, and survival. JCO Precis Oncol. 2019;3.

13. Zill OA, Banki L, Firthough SR, Mortimer SA, Voskett JV, Mokhtar H, Gandara DR, Mack PC, Odegaard JI, Nagy RI, et al. The landscape of actionable genomic alterations in cell-free circulating tumor DNA from 21,807 advanced Cancer patients. Clin Cancer Res. 2018;24(15):3528–38.

14. Mansukhani S, Barber LJ, Kleiflogianis D, Moorcraft SY, Davidson M, Woolston A, Proszek PZ, Griffiths B, Fenwick K, Herman B, et al. Ultra-sensitive mutation detection and genome-wide DNA copy number reconstruction by error-corrected circulating tumor DNA sequencing. Clin Chem. 2018;64(11):1626–35.

15. Kato S, Okamura R, Baumgartner JM, Patel H, Leichman L, Kelly K, Siddick JK, Fanta PT, Lippman SM, Kurockz R. Analysis of circulating tumor DNA and clinical correlates in patients with esophageal, Gastroesophageal junction, and gastric adenocarcinoma. Clin Cancer Res. 2018;24(24):6248–56.

16. Wyatt AW, Annala M, Aggarwal R, Beja K, Feng F, Youngren J, Foye A, Lloyd MJ, Luoy X, Raymond VM, Talasaz A, Kurzrock R. Genomic assessment of blood-derived circulating tumor DNA in patients with colorectal cancer. PLoS One. 2015;10(5):e0126417.

17. Danese E, Minicozzi AM, Benati M, Montagnana M, Paviati E, Salvagno GL, Mundt MW, Cole SR, Gopalsamy G, Mangira D, et al. Evaluation of an assay for noninvasive blood-derived circulating tumor DNA in patients with colorectal cancer. PLoS One. 2015;10(5):e0126417.

18. Subramanian A, Tamayo P, Mukherjee S, Sood AK, Ebert BL, Gillette MA, Jeong JH,鳞大hoff MR, K M, K Youngren J, Foye A, Lloyd MJ, Luoy X, Raymond VM, Talasaz A, Kurzrock R. Genomic assessment of blood-derived circulating tumor DNA in patients with colorectal cancer. PLoS One. 2015;10(5):e0126417.

19. Lippman SM, Ray EE, Alningar A, Bagherzadeh M, Akbari MR. Liquid biopsy in breast cancer: a comprehensive review. Clin Genet. 2019.

20. Rohanizadegan M. Analysis of circulating tumor DNA in breast cancer as a diagnostic and prognostic biomarker. Cancer Genet. 2018;228:29159–68.

21. Buono G, Gerratana L, Bulfoni M, Provinciali N, Basile D, Giuliano M, Corvaja C, Apino G, Del Mastro L, De Pacido S, et al. Circulating tumor DNA analysis in breast cancer: is it ready for prime-time? Cancer Treat Rev. 2019;73:73–83.

22. Perlin E, Moquin RB. Serum DNA levels in patients with malignant disease. J Thorac Oncol. 2017;12(1):77–91.

23. Phallan J, Sausies M, Gopalsamy G, Margaria D, et al. Evaluation of an assay for noninvasive blood-derived circulating tumor DNA in patients with colorectal cancer. PLoS One. 2015;10(5):e0126417.

24. Phallan J, Sausies M, Gopalsamy G, Margaria D, et al. Evaluation of an assay for noninvasive blood-derived circulating tumor DNA in patients with colorectal cancer. PLoS One. 2015;10(5):e0126417.

25. Kaiser J. National Institutes of Health. NCI gears up for cancer genome project. Science. 2005;307(5713):1182.
46. Lo TL, Yousif P, Fong CW, Guo K, McCaw BJ, Phillips WA, Yang H, Wong ES, Leong HF, Zeng Q, et al. The ras/mitogen-activated protein kinase pathway inhibitor and likely tumor suppressor proteins, sprouty 1 and sprouty 2 are deregulated in breast cancer. Cancer Res. 2004;64(7):7612–76.

47. Cheng JC, Chang HM, Xiong S, So WK, Leung PC. Sprouty2 inhibits amphiregulin-induced down-regulation of E-cadherin and cell invasion in human ovarian cancer cells. Oncotarget. 2016;7(9):18465–18470.

48. Joshi-Tope G, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, de Bono B, Jassal B, Gopinath GR, Wu GR, Matthews L, et al. Reactome: a knowledgebase of biological pathways. Nucleic Acids Res. 2005;33(Database issue):D428–D433.

49. Rossetti S, Ren M, Visconti N, Corlazzoli F, Corlazzoli H, Koller A, Samet Y, Takahashi H, Kagara N, Tanei T, Naoi Y, Shimoda M, Shimazu K, Salvi S, Gurioli G, De Giorgi U, Conteduca V, Tedaldi G, Calistri D, Casadio V, Longo MC, Berninger MS, Hartley JL. Use of uracil DNA glycosylase to control carry-over contamination in polymerase chain reactions. Genet. 1990;93(1):25–30.

50. Salvi S, Gurioli G, De Giorgi U, Conteduca V, Tedaldi G, Calistri D, Casadio V. OncoTargets Ther. 2016;9:6549–6556.

51. Habashy HO, Rakha EA, Ellis IO, Powe DG. The oestrogen receptor coactivator CARMA1 has an oncogenic effect and is associated with poor prognosis in breast cancer. Breast Cancer Res Treat. 2013;140(2):307–16.

52. Longo MC, Biningger MS, Hartley JL. Use of uracil DNA glycosylase to control carry-over contamination in polymerase chain reactions. Gene. 1990;93(1):25–30.

53. Lui X, Ren J, Luo N, Guo H, Zheng Y, Li J, Tang F, Wen L, Peng J. Comprehensive DNA methylation analysis of tissue of origin of plasma cell-free DNA by methylated CpG tandem amplification and sequencing (MCTA-Seq). Clin Epigenetics. 2019;11(1):93.

54. Kides E, Jeffrey SS. Circulating tumor cells versus tumor-derived cell-free DNA: rivals or partners in cancer care in the era of single-cell analysis? Genome Med. 2013;5(8):70.

55. Issa JP. DNA methylation as a therapeutic target in cancer. Clin Cancer Res. 2007;13(6):1634–7.

56. Visvanathan K, Fackler MS, Zhang Z, Loper-Bujanda ZA, Jeter SC, Sokoll LJ, Garrett-Mayer E, Cope JM, Umbrecht CB, Ehsani DM, et al. Monitoring of serum DNA methylation as an early independent marker of response and survival in metastatic breast cancer. JBCRC 005 prospective biomarker study. J Clin Oncol. 2017;35(7):751–8.

57. Takahashi H, Kagara N, Tanii T, Naoi Y, Shimoda M, Shimomura A, Shimazu K, Kim SJ, Noguchi S, Correlation of methylated circulating tumor DNA with response to Neoadjuvant chemotherapy in breast cancer patients. Clin Breast Cancer. 2017;17(1):61–9 e63.

58. Shon M, Yin H, Li J, Li X, Wang D, Su Y, Niu M, Zhong Z, Wang J, Zhang X, et al. Detection of aberrant methylation of a six-gene panel in serum DNA for diagnosis of breast cancer. Oncotarget. 2016;7(14):18485–18495.

59. Fackler MJ, Lopez-Bujanda Z, Umbrecht C, Teo WW, Cho S, Zhang Z, Visvanathan K, Jeter S, Arangio P, Wang C, et al. Novel methylated biomarkers and a robust assay to detect circulating tumor DNA in metastatic breast cancer. Cancer Res. 2014;74(8):2160–70.

60. Kloten V, Becker B, Winner K, Schrauder MG, Fasching PA, Anzeneder T, Veeck J, Hartmann A, Knuchel R, Dahl E. Promoter hypermethylation of the tumor-suppressor genes ITHS, DKK3, and RASSF1A as novel biomarkers for blood-based breast cancer screening. Breast Cancer Res. 2013;15(1):R4.

61. Agostini M, Enzo MV, Bedin C, Belardinelli V, Goldin E, Del Bianco P, Masciotti E, D'Angelo E, Izi L, Saccani A, et al. Circulating cell-free DNA: a promising marker of regional lymphnode metastasis in breast cancer patients. Cancer Biomark. 2012;11(2–3):89–98.

62. Brooks JD, Carins P, Shore RE, Klein CB, Wingin I, Afanasyeva Y, Zeleniuch-Jacquotte A. DNA methylation in pre-diagnostic serum samples of breast cancer cases: results of a nested case-control study. Cancer Epidemiol. 2010;34(6):717–23.

63. Shukla S, Mirza S, Sharma G, Panah R, Gupta SD, Ralhan R. Detection of RASSF1A and RARbeta hypermethylation in serum DNA from breast cancer patients. Epigenetics. 2006;1(2):88–93.

64. Bianco S, Jangal M, Garneau D, Gevry N. LRH-1 controls proliferation in breast tumor cells by regulating CDK1A gene expression. Oncogene. 2015;34(39):4509–18.

65. Chand AL, Wijayakumara DD, Knower KC, Herridge KA, Howard TL, Lazarus KA, Clyne CD. The orphan nuclear receptor LRH-1 and ERalpha activate GREB1 expression to induce breast cancer cell proliferation. PLoS One. 2012;7(2):e31593.

66. Wang S, Zou Z, Luo X, Mi Y, Wang C, Xing D. LRH1 enhances cell resistance to chemotherapy by transcriptionally activating MDC1 expression and attenuating DNA damage in human breast cancer. Oncogene. 2018;37(24):3243–59.

67. Pang JB, Molan A, Chang A, Knower K, Takano EA, Byrne DJ, Mikeska T, Millar ERA, Lee CS, O'Toole SA, et al. LRH-1 expression patterns in breast cancer tissues are associated with tumour aggressiveness. Oncotarget. 2017;8(48):86256–36.

68. Bianco S, Brunelle M, Jangal M, Magnani L, Gevry N. LRH-1 governs vital transcriptional programs in endocrine-sensitive and -resistant breast cancer cells. Cancer Res. 2014;74(7):7415–25.

69. Wallace JA, Pitarresi JR, Sharma N, Palettas M, Cuitino MC, Sizemore ST, Yu L, Sardella R, Rosol TJ, Mehta KD, et al. Protein kinase C beta in the tumor microenvironment promotes mammary tumorigenesis. Front Oncol. 2014;4:87.

70. Li H, Weinstein IB. Protein kinase C beta enhances growth and expression of cyclin D1 in human breast cancer cells. Cancer Res. 2006;66(23):1399–408.

71. Grosson VC, Todaro LB, Kazanietz MG, de Kier Joffe ED, Ungerer AJ. Opposite effects of protein kinase C beta1 (PKCbeta1) and PKCepsilon in the metastatic potential of a breast cancer murine model. Breast Cancer Res Treat. 2009;118(3):469–80.

72. Sledge GW Jr, Okmen-Polar Y. Protein kinase C-beta as a therapeutic target in breast cancer. Semin Oncol. 2006;33(3 Suppl 9):515–8.

73. Grossoni VC, Todaro LB, Kazanietz MG, de Kier Joffe ED, Ungerer AJ. Opposite effects of protein kinase C beta1 (PKCbeta1) and PKCepsilon in the metastatic potential of a breast cancer murine model. Breast Cancer Res Treat. 2009;118(3):469–80.

74. Gomez-Maldonado L, Tiana M, Roche O, Prado-Cabrero A, Jensen L, Fernandez-Barral A, Gijurko-Munoz I, Favaro E, Moreno-Bueno G, Sanz L, et al. ENFAS long noncoding RNAs induced by hypoxia promote metastatic dissemination. Oncogene. 2015;34(20):2609–20.

75. Gong W, Qie S, Huang P, Xu J, Deletion of long noncoding RNA ENFAS aggravates hypoxia-induced injury in PC-12 cells by upregulation of miR-101A. J Cell Biochem. 2019;120(1):836–47.

76. Swierczynski S, Klieser E, Illig R, Alinger-Scharinger B, Kieslich T, Neureiter D. Histone deacetylation meets miRNA: epigenetics and post-transcriptional regulation in cancer and chronic diseases. Expert Opin Biol Ther. 2015;15(5):651–64.

77. Wang S, Wu W, Claret FX, Mutual regulation of micro-RNAs and DNA methylation in human cancers. Epigenetics. 2017;12(3):187–97.

78. Haemovich MG, Medina DA, Causse SZ, Garber M, Millan-Zambrano G, Barkai O, Chavez S, Perez-Ortin JE, Darzak X, Choder M. Gene expression is circular factors for mRNA degradation also foster mRNA synthesis. Cell. 2013;153(5):1000–11.

79. Panagopoulou M, Karaglini M, Bouloukaki I, Biziota E, Koukati K, Karamitrousis E, Nena E, Tsimardinos I, Kollios G, Lianidou E, et al. Circulating cell-free DNA in breast cancer: size profiling, levels, and methylation patterns lead to prognostic and predictive classifiers. Oncogene. 2019;38(18):3387–401.

80. Mishima C, Kagara N, Matsui S, Tanii T, Naoi Y, Shimoda M, Shimomura A, Shimazu K, Kim SJ, Noguchi S. Promoter methylation of TRIM9 as a marker for detection of circulating tumor DNA in breast cancer patients. Springerplus. 2015;4:635.

81. Parisi C, Mastoraki S, Markou A, Strati A, Chimonidou M, Georgoulia V, Lianidou ES. Development and validation of a multiplex methylation specific PCR-coupled liquid bead array for liquid biopsy analysis. Clin Chim Acta. 2016;461:156–64.
87. Fu D, Ren C, Tan H, Wei J, Zhu Y, He C, Shao W, Zhang J. Sox17 promoter methylation in plasma DNA is associated with poor survival and can be used as a prognostic factor in breast cancer. Medicine. 2015;94(11):e637.

88. Mastoraki S, Strati A, Tzanikou E, Chimonidou M, Politaki E, Voutsina A, Psyrri A, Georgoulas V, Lianidou E. ESR1 methylation: a liquid biopsy-based epigenetic assay for the follow-up of patients with metastatic breast cancer receiving endocrine treatment. Clin Cancer Res. 2018;24(6):1500–10.

89. Chimonidou M, Strati A, Malamos N, Georgoulas V, Lianidou ES. SOX17 promoter methylation in circulating tumor cells and matched cell-free DNA isolated from plasma of patients with breast cancer. Clin Chem. 2013;59(1):270–9.

90. Bosviel R, Garcia S, Lavediaux G, Michard E, Dravers M, Kwiatkowski F, Bignon VJ, Bernard-Gallon DJ. BRCA1 promoter methylation in peripheral blood DNA was identified in sporadic breast cancer and controls. Cancer Epidemiol. 2012;36(3):e177–82.

91. Sharma G, Mirza S, Prasad CP, Srivastava A, Gupta SD, Ralhan R. Promoter hypermethylation of p16INK4A, p14ARF, CyclinD2 and Sirt2 in serum and tumor DNA from breast cancer patients. Life Sci. 2007;80(20):1873–81.

92. Martinez-Galan I, Torres B, Del Moral R, Munoz-Gamez JA, Martin-oliva D, Villabobos M, Nuñez MI, Luna I, De D, Oliver FJ, Ruiz de Almodovar JM. Quantitative detection of methylated ESR1 and 14-3-3-sigma gene promoters in serum as candidate biomarkers for diagnosis of breast cancer and evaluation of treatment efficacy. Cancer Biol Ther. 2008;7(6):958–65.

93. Grunau C, Clark SJ, Rosenthal A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. Nucleic Acids Res. 2001;29(13):E65.

94. Lee KH, Shin TJ, Kim WH, Lee SY, Cho JY. Methylation of LINE-1 in cell-free DNA serves as a liquid biopsy biomarker for human breast cancers and dog mammary tumors. Sci Rep. 2019;9(1):175.

95. Kristiansen S, Nielsen D, Schettopoulou F, Detection and monitoring of hypermethylated RASSF1A in serum from patients with metastatic breast cancer. Clin Epigenetics. 2016;8:35.

96. Ng EK, Leung CP, Shin YY, Wong CL, Ma ES, Jin HC, Chu KM, Kwong A. Quantitative analysis and diagnostic significance of methylated SLC19A3 DNA in the plasma of breast and gastric cancer patients. PLoS One. 2011;6(7):e22333.

97. Shen SY, Singhania R, Fehringer G, Chakravarthy A, Roehl MHA, Chadwick D, Zuanese PC, Borgia A, Wang TT, Li T, et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. Nature. 2018;563(7732):579–83.

98. Maggi EC, Gravina S, Cheng H, Piperdì B, Yuan Z, Dong X, Libutti SK, Vilj J, Montagna C. Development of a method to implement whole-genome bisulfite sequencing of cfDNA from Cancer patients and a mouse tumor Model. Front Genet. 2018;9:56.

99. Liu J, Young JM, Jassowicz AF, Vijayaraghavan R, Kang H, Zhang R, Kryukyuk KM, Huang HJ, Hinoue T, Shen H, et al. Targeted methylation sequencing of plasma cell-free DNA for cancer detection and classification. Ann Oncol. 2018;29(6):1445–53.

100. Holmila R, Sillesen A, Muller OC, Degli Esposti D, Guillotereau P, McKay J, Sargrajrang S, Srivatanakul P, Hainaut P, Merle P, et al. Targeted deep sequencing of plasma circulating cell-free DNA reveals Vimentin and F比lin 1 as potential epigenetic biomarkers for hepatocellular carcinoma. PLoS One. 2017;12(3):e0174265.

101. Vaca-Paniagua F, Oliver J, Nogueira da Costa a, Merle P, McKay J, Hercek Z, Holmila R. Targeted deep DNA methylation analysis of circulating cell-free DNA in plasma using massively parallel semiconductor sequencing. Epigenomics. 2015;7(3):353–62.

102. Sun K, Jiang P, Chan KC, Wong J, Cheng YK, Liang RH, Chan WK, Ma ES, Chan Si, Cheng SH, et al. Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. Proc Natl Acad Sci U S A. 2015;112(40):E5503–12.

103. Lissa D, Robles AI. Methylation analyses in liquid biopsy. Transl Lung Cancer Res. 2016;5(5):545–504.