Article

Investigation of the Role of PUFA Metabolism in Breast Cancer Using a Rank-Based Random Forest Algorithm

Mariia V. Guryleva [1], Dmitry D. Penzar [1,2], Dmitry V. Chistyakov [3,*], Andrey A. Mironov [1,4], Alexander V. Favorov [2,5] and Marina G. Sergeeva [3]

[1] Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, 119234 Moscow, Russia
[2] Vavilov Institute of General Genetics, Russian Academy of Sciences, 119991 Moscow, Russia
[3] Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, 119992 Moscow, Russia;
[4] Kharkevich Institute of Information Transmission Problems, Russian Academy of Sciences, 127051 Moscow, Russia
[5] School of Medicine, Johns Hopkins University, Baltimore, MD 21218, USA
*   Correspondence: chistyakov@belozersky.msu.ru; Tel.: +7-495-939-4332

Simple Summary: Polyunsaturated fatty acids (PUFAs) and their derivatives, oxylipins, are a constant focus of cancer research due to the relationship between cancer and processes of energy metabolism and inflammation, where a PUFA system is an active player. Only recently have methods been developed that allow for studying such complex systems. Using the Rank-based Random Forest (RF) model, we show that PUFA metabolism genes are critical for the pathogenesis of breast cancer (BC); BC subtypes differ in PUFA metabolism gene expression. The enrichment of BC subtypes with various genes associated with oxylipin signaling pathways indicates a different contribution of these compounds to the biology of subtypes.

Abstract: Polyunsaturated fatty acid (PUFA) metabolism is currently a focus in cancer research due to PUFAs functioning as structural components of the membrane matrix, as fuel sources for energy production, and as sources of secondary messengers, so called oxylipins, important players of inflammatory processes. Although breast cancer (BC) is the leading cause of cancer death among women worldwide, no systematic study of PUFA metabolism as a system of interrelated processes in this disease has been carried out. Here, we implemented a Boruta-based feature selection algorithm to determine the list of most important PUFA metabolism genes altered in breast cancer tissues compared with in normal tissues. A rank-based Random Forest (RF) model was built on the selected gene list (33 genes) and applied to predict the cancer phenotype to ascertain the PUFA genes involved in cancerogenesis. It showed high-performance of dichotomic classification (balanced accuracy of 0.94, ROC AUC 0.99) We also retrieved a list of the important PUFA genes (46 genes) that differed between molecular subtypes at the level of breast cancer molecular subtypes. The balanced accuracy of the classification model built on the specified genes was 0.82, while the ROC AUC for the sensitivity analysis was 0.85. Specific patterns of PUFA metabolic changes were obtained for each molecular subtype of breast cancer. These results show evidence that (1) PUFA metabolism genes are critical for the pathogenesis of breast cancer; (2) BC subtypes differ in PUFA metabolism genes expression; and (3) the lists of genes selected in the models are enriched with genes involved in the metabolism of signaling lipids.

Keywords: breast cancer; machine learning; PUFAs; transcriptomics; random forest

1. Introduction

Breast cancer (BC) is the leading cause of cancer death among women worldwide [1]. BC is a heterogeneous disease; such a feature determines the risk of disease progression and its resistance to therapy [2,3]. There are five molecular subtypes of BC: luminal A,

luminal B, HER2-enriched, basal-like, and normal-like [4]. Depending on the subtype, various molecular mechanisms of the pathogenesis of neoplasia are realized [3], while the patterns of metabolic phenotypes of molecular subtypes remain insufficiently studied [5].

The understanding of the relationship between changes in metabolism and cancer development has changed the focus of research several times over past decades [6]. Initially, studies of the mechanism of the "Warburg effect", i.e., a metabolic switch from oxidative to glycolytic metabolism, attracted the greatest interest and, then, its relationship with the metabolism of nucleotides, lipids, and proteins [7]. Currently, metabolic rewiring has been recognized as an important feature to the progression of cancer in light of producing signal molecules [8].

One of the leading roles in the synthesis of signal lipid mediators is ascribed to the polyunsaturated fatty acids (PUFAs). Polyunsaturated fatty acids have more than one double bond in the carbon skeleton and represent a part of fatty acid (FA) metabolism. These acids are divided according to the double-bond position, the main important ones of which are the so-called Omega-3 (e.g., DHA and EPA) and Omega-6 (e.g., AA). Besides being structural components of the membranes and fuel sources for energy production, PUFAs also have a signaling function themselves or via their oxidative derivatives [9]. Both PUFAs and their oxidized derivatives, oxylipins, modulate the intrinsic cell programs and are utilized for communication with neighbor cells [10,11]. The important role of PUFAs and their corresponding oxylipins is an involvement in regulation of inflammatory processes. Omega-3 PUFAs (DHA and EPA) are attributed mainly to anti-inflammatory effects, while Omega-6 PUFAs, such as AA, are thought to be a part of proinflammatory pathways [12]. Oxylipins could be derived from Omega-3 as well as Omega-6 PUFAs [13,14]. Along with their precursors, oxylipins are responsible for inflammation and its subsequent resolution [15–17].

Inflammation is designated as a characteristic among the hallmarks of cancer [6,9,18,19], and unresolved, chronic inflammation, characterized by abnormal oxylipins synthesis, becomes fertile soil for malignant transformation and tumor immune evasion, including colorectal [12], gastrointestinal [20], colon [21,22], breast [23,24], pancreatic [25], prostate, and lung [26] cancers, and in melanoma [27]. Although for decades, it was known that the action of oxylipins is complex and is the result of PUFA metabolism through various enzymatic pathways [13,14,28], only recently did the development of omics technologies and algorithms for analyzing the data open up new opportunities in studying the role of PUFAs and their metabolites in the development of various diseases. Although there is evidence that individual oxylipins or the expression of genes responsible for their metabolism may be characteristics of different subtypes of BC [29,30], no systematic study of PUFA metabolism as a system of interrelated processes in this disease has been carried out. Further research concerning the role of dysregulated PUFA metabolism in the pathobiology of cancer holds great promise in uncovering novel metabolic and signaling nodes for targeted therapies.

Transcriptome analysis is one of the productive ways to study metabolic pathways alteration. It was shown that cancer metabolic reprogramming is regulated on the transcription level [31–33]. The heterogeneity of breast cancer and the variability of metabolic processes accompanying this heterogeneity requires a big amount of data to study. The data have been accumulated with the rise in next-generation sequencing (NGS) techniques and microarrays [34,35]. The joint analysis of such datasets via machine learning (ML) approaches could better resolve PUFAs' roles in breast cancer development.

ML approaches, particularly, the Random Forest (RF) [36] algorithm, have already been successfully applied to expression data analysis for various cancer types [37,38]. Nonetheless, the use of the entire data set is restricted by the differences in the corresponding technologies. Nonparametric methods independent of monotonic normalization can be used to overcome these limits [39,40]. Combining the nonparametric techniques, particularly, the ranking of the expression profile, with Random Forest simplifies the RF application on heterogeneous datasets. RF approaches are effectively implemented to

distinguish heterogeneous groups, while, as far as we know, they have not been run for PUFA pathway analyses.

In this study, we used an approach based on the combination of nonparametric method and RF model. Note that the Boruta feature selection [41,42] and Sequential Feature Selector [43,44] methods have been widely used and shown to be beneficial. The use of the Boruta feature selection algorithm [45] made it possible to identify the genes responsible for PUFA metabolism alterations in BC. The prediction ability of the classifiers was tested on independent datasets. The study has expanded our knowledge about the role of PUFA metabolic pathways in breast cancer pathogenesis and allowed us to identify their specific patterns in different molecular subtypes of breast cancer.

2. Materials and Methods

2.1. Data Source

Transcriptome profiles for samples of breast cancer and normal adjacent tissues were used to train and validate the Random Forest model. For dichotomic classification of breast cancer and normal tissue samples, four datasets from the database Gene Expression Omnibus (GEO) were obtained (GSE65216, GSE29044, GSE1078, and GSE62944) (GEO; http://www.ncbi.nlm.nih.gov/geo/ (accessed on 18 February 2022)) (Supplementary Table S2). Dataset GSE62944 presents TCGA data from the TCGA (The Cancer Genome Atlas) database (Broad GDAC, https://gdac.broadinstitute.org/ (accessed on 18 February 2022)). Three datasets were used (GSE65216, GSE29044, and GSE1078) for the training set and included 221 tumor and 185 normal samples. TCGA data were used for the validation set and consisted of 1082 tumor and 113 normal samples. In order to distinguish among the subtypes of breast cancer by the differences in PUFA metabolism gene expression between five datasets (GSE81538, GSE25066, GSE31448, GSE96058, and GSE21653) with molecular subtypes, annotations were extracted from the GEO database (GEO; http://www.ncbi.nlm.nih.gov/geo/ (accessed on 18 February 2022)) (Supplementary Table S3).

2.2. Random Forest Model

A Random Forest [36] predictor was built using the random Forest [46] R package. Briefly, Random Forest is an assembly of decision trees. It represents the union of two methods: bagging and random subspace method (RSM) [47]. Three main steps can be highlighted when building a Random Forest classifier:

1. From input data N × M (where N is the number of samples and M is the number of used features), k subsets are randomly selected with a return;
2. A decision tree is built for each subset;
3. The final decision is made on the majority vote for classification tasks or by averaging in the regression tasks.

Each decision tree includes a number of comparisons of the feature values and threshold, which is set during the model training. This fact limits the usage of random forest to gene expression data. In order to overcome this limitation, we ranked genes within a sample in both the training and test sets.

After performing a ranking procedure, genes from the PUFA list presented in both the training and test sets (for dichotomic and multi-class classification separately) were selected. The Boruta algorithm (Boruta R package) was implemented on the training sets with the extracted PUFA genes (185 genes for tumor vs. normal samples predictor; 155 genes for molecular subtypes predictor) to shrink the number of studied genes. Final classifiers were built on the genes highlighted by Boruta as important ones (33 for tumor vs. normal samples comparison; 46 genes for molecular subtypes comparison) with the number of trees being 450 (Supplementary Figure S2). The code for the described analysis can be found at the GitHub repository: https://github.com/gurylevamv/PUFA_rRF (accessed on 21 September 2022).

2.3. Boruta Feature Selection Algorithm

The feature selection algorithm Boruta was used to computationally identify the genes for which expression is important to distinguish between biological conditions. The main idea of this algorithm is to compare the features' importance with the randomized version of themselves. The randomized features are referred to as shadows. Technically, a shadow feature is obtained from the initial one through value shuffling in a dataset copy. Two datasets, the initial features and the one with shadow features, are then merged. The Random Forest classifier is built on the merged dataset, and the importance of all features are calculated by the classifier. If the importance of an initial feature is greater than the maximum of shadow feature importance, then it receives 1 score. These operations are repeated a pre-given number (say N) of times. As a result, we obtain the sums of the scores for each feature after N trials. In a null model, these trials are binomially distributed. If the score of a feature is greater than the 99.5% quantile of the distribution, the feature is accepted as the important one. We used the Boruta algorithm as R package Boruta (accessed on 18 February 2022) [45]. The Boruta algorithm was applied to the PUFA gene list both for dichotomic (tumor vs. normal controls) and multi-class classifications. For greater confidence in the selected features, Boruta's algorithm with the default cycles' number parameter N (250) was run 100 times. The genes selected in 90% or more of the runs were finally considered as important genes. In the dichotomic sample comparison, 33 PUFA genes were selected by Boruta as important for classification. In the molecular subtype classification, 46 genes were selected as important.

2.4. Sequential Feature Selector for Minimal Gene Set Selection

The Sequential Feature Selector (SFS) [48] was implemented to reach the minimal set of genes that shows the highest classification performance. This algorithm consequently selects a feature that will maximize the quality criterion function from the space of all features. Additionally, we used a floating extension of the SFS method (SFFS) that allows us to remove features if this step will make the prediction better. SFFS was used from the mlxtend Python package (accessed on 18 February 2022) [49].

2.5. SHAP Values to Identify the Most Important PUFA Genes

SHAP refers to Shapley additive explanations, which is an approach that allows us to reach an explanation for the machine learning models output. It calculates the importance for each feature in each single sample. By applying SHAP values to the built predictors, we enhance their transparency, and moreover, in multi-class classification, it allows us to reach the importance of the feature in the separation of each class. SHAP calculations were performed via the SHAP package in Python (https://shap.readthedocs.io/ (accessed on 18 February 2022)).

2.6. Enrichment Analysis

A GO functional annotation (Biological process, Molecular function) [50], and a KEGG [51] and Wiki pathway [52] enrichment analysis for the important PUFA genes revealed was performed via the Enrichr tool wrapped in python in the GSEApy package (https://gseapy.readthedocs.io/en/latest/introduction.html#gseapy-enrichr-module (accessed on 18 February 2022)). Background gene sets were set as the lists of PUFA genes presented in both the training and test cohorts separately for the classification of tumor and normal samples and for the molecular subtypes. Terms with adjusted *p*-value < 0.001 were considered statistically significant.

2.7. Differential Expression Analysis

The expression levels of the genes that were selected as important features for dichotomic classification were compared between the tumor and normal sample groups with a two-sided Mann–Whitney test, followed by the Benjamini–Hochberg procedure for multiple comparisons. The false discovery rate (FDR) was set as 0.05. To find *PUFA* genes

that were differentially expressed across molecular subtypes, the first one-way ANOVA test was applied to the expressions of the genes that were previously selected as important for the molecular subtype classification. Genes with significantly (adjusted *p*-value < 0.05) different expression means were further compared between subtypes with a one-sided Mann–Whitney test, followed by the Benjamini–Hochberg (FDR = 0.05). Tests were performed with the SciPy library (https://scipy.org/ (accessed on 18 February 2022)) and the statsmodel module in Python (https://www.statsmodels.org/stable/index.html (accessed on 18 February 2022)).

3. Results

3.1. Validation of Machine Learning Nonparametric Approach

This study is based on the Random Forest (RF) machine learning approach. The RF model consists of a number of decision trees with thresholds learned from the training set. A direct combination of different datasets does not work in this framework due to the differences in platforms that do not allow us to learn the common thresholds even for similar biological conditions. This limitation was overcome by the combination of the nonparametric method with RF algorithm (see the Section 2.2). This approach allowed us to use merged datasets to learn and to test the RF model. The model was validated on a dichotomic classification of head and neck cancer and normal tissues based on the 1000 most variable genes (Supplementary Figure S1). Quality metrics (Balanced accuracy 0.99, ROC-AUC 0.99, PR-AUC 0.96) showed high performance and biological relevance of the most important features selected by Boruta [45] from the ranked expression levels of the 1000 genes.

3.2. Rank Model to Identify Most Important PUFA Genes for Breast Cancer vs. Normal Tissues Classification

To assess the role of PUFA metabolism in the pathogenesis of BC, we compiled a list of 202 genes based on known data [13,14,53–55] that was previously described (Genes List in Supplementary Materials) [56]. We performed a systematic search for transcriptomes from open databases using previously developed tool ARGEOS [57], and we selected datasets GSE65216, GSE29044, GSE10780 (*n* = 231 tumor samples, *n* = 185 normal samples), and GSE62944 (*n* = 1082 tumor samples, *n* = 113 normal samples), with the latter representing TCGA data from the TCGA (The Cancer Genome Atlas) database. The datasets included samples of both breast cancer and normal adjacent tissues (Supplementary Table S2). Next, the initial list of PUFA genes selected overlapped with the genes presented in the datasets; 185 genes presented in all datasets were chosen for further analysis (Genes list, Supplementary). The chosen datasets were divided into two groups: training sets (GSE65216, GSE29044, and GSE10780) and test set (GSE62944). The workflow for further studying the differences in PUFA regulation between normal and breast cancer samples is presented in Figure 1.

We used our pipeline based on the Boruta feature selection method. We reran Boruta several times, and on each run, the method selected the genes with ranked expression levels that are reliably more important for the classification than their shuffled ranks; see the Section 2.3. for details. From the 185 PUFA genes, 33 genes (Supplementary Genes list) were chosen as important in the training set (Figure 1, left flowchart).

These genes were further used to learn a rank Random Forest dichotomic model. The model's quality was evaluated on the test samples (Figure 1, right flowchart). The results are shown in Figure 2 and Supplementary Table S4.
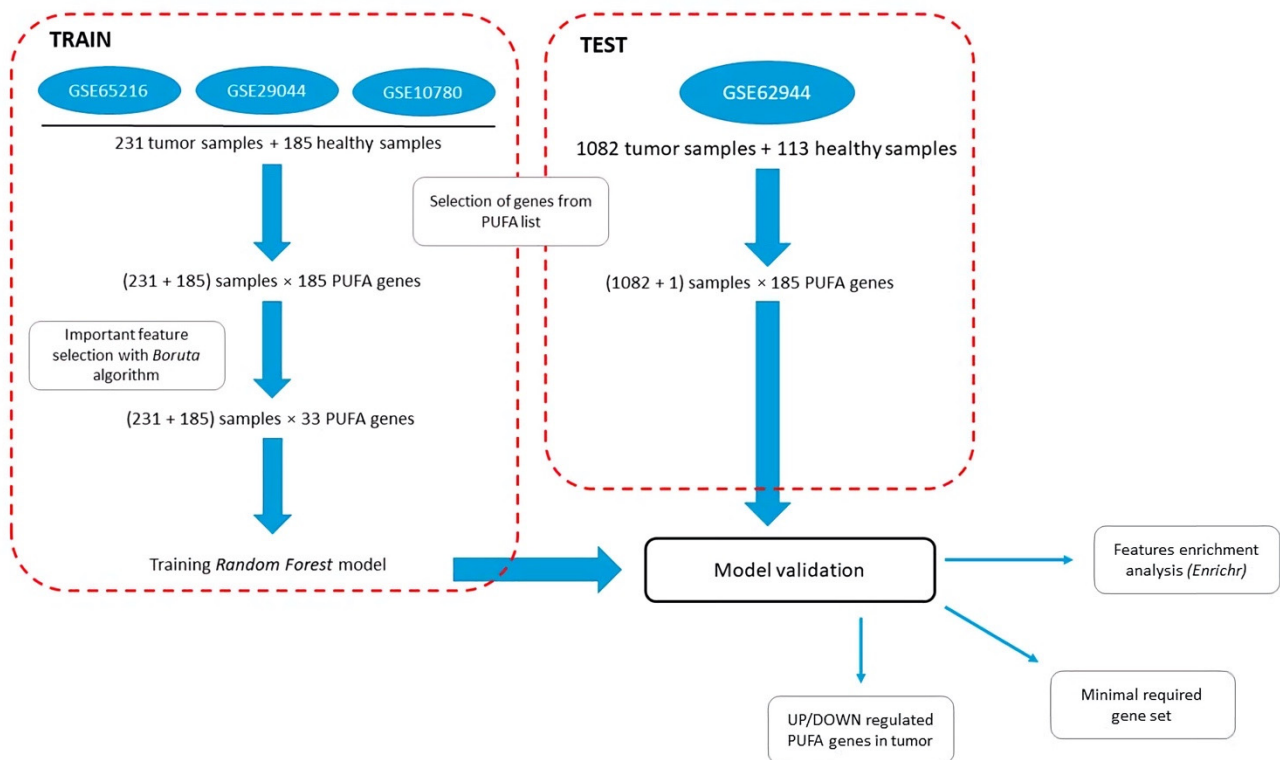
Figure 1. Workflow for studying differences in PUFA regulation between normal and breast cancer samples. Starting from the processed expression matrixes of the corresponding datasets taken from the GEO database, the training and test sets are constructed. To avoid the effect by platform differences, gene expression data within each sample are then converted to the ranks. PUFA genes are selected for the subsequent analysis (See details in the Section 2).
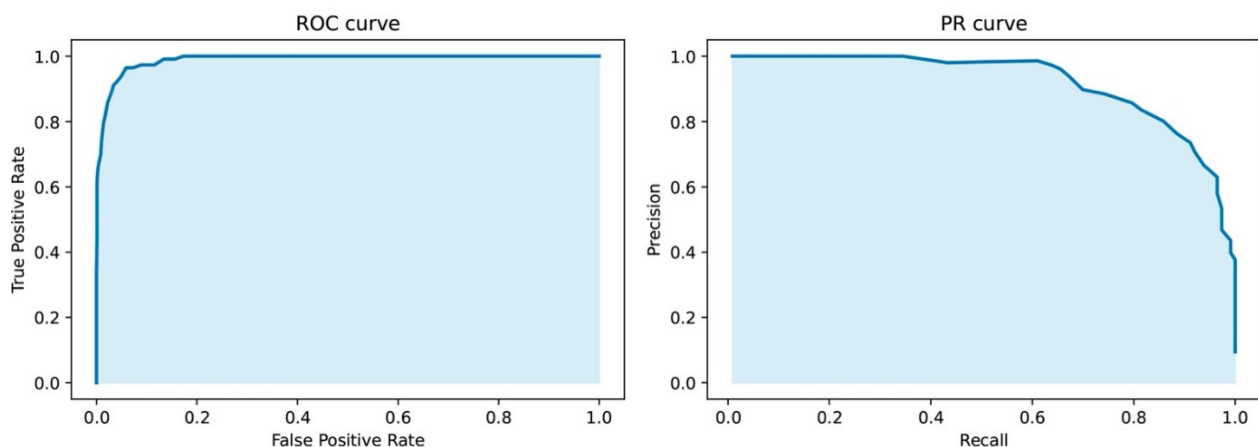

Figure 2. ROC and PR metrics for dichotomic classification of breast cancer samples and normal ones.

The resulting classifier based on 33 *PUFA* genes effectively separates diseased and normal samples (Supplementary Table S4). This indicates that, indeed, the expression profiles of *PUFA* metabolism genes differ between normal and tumor tissues.

Of the 33 selected genes, 6 genes were significantly (p-value < 0.05) upregulated (see Section 2.7) in the breast cancer samples and 24 genes were downregulated in comparison with normal tissues (Supplementary Table S5). To characterize these genes, an analysis of the GO functional and biological pathways, as well as the KEGG and WikiPathways

pathways were performed using the Enrichr method (see the Section 2.6). KEGG enrichment indicated that the linoleic acid metabolic pathway was upregulated in breast cancer, while in normal adjacent tissues, arachidonic acid metabolic processes were the most enriched KEGG pathways (Supplementary Figures S3 and S4). Moreover, eicosanoid metabolism via the cyclooxygenase pathway was found to be downregulated in tumors compared with normal samples according to WikiPathways (Supplementary Figure S3).

At the next stage, we used the Sequential Feature Selector (SFS) method (see the Section 2.4) to identify the minimum set of genes that demonstrates the best quality of the tumor vs. normal tissue separation. The SFS algorithm has determined that the rank RF classifier based on a list of seven genes (ADIPOR1, HADH, ACOT7, PTGER4, PLA2G15, PLA2G1B, and CYP46A1) has the highest predictive efficiency according to ROC-AUC score (ROC-AUC 0.99, ci-bound 0.002) (Figure 3A). The expression of these genes in breast cancer and normal adjacent tissues is shown in Figure 3B.
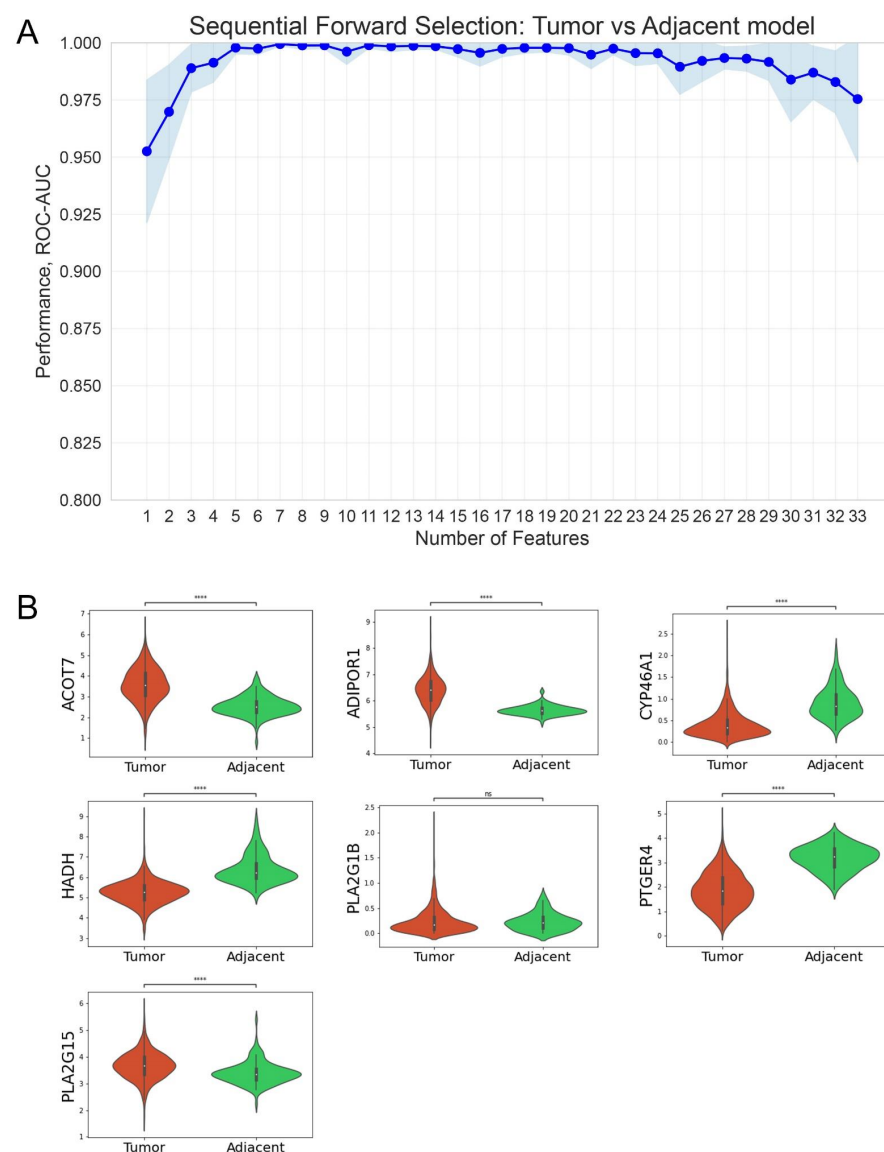

Figure 3. **** $p < 0.0001$. Identification of the minimal PUFA gene set required for efficient classification of breast cancer and normal samples. (**A**) Light blue zones refer to 95% confidence intervals. (**B**) Expression of *ADIPOR1*, *HADH*, *ACOT7*, *PTGER4*, *PLA2G15*, *PLA2G1B*, and *CYP46A1* genes in breast cancer and normal adjacent tissues.

3.3. Rank Model to Identify Most Important PUFA Genes for Breast Cancer Classification

Breast cancer is a heterogeneous oncological disorder [58]. Since the emergence of high-throughput sequencing intrinsic molecular subtypes of breast cancer became widely used. Sørlie et al. distinguished five molecular subtypes: luminal A, luminal B, normal-like, HER2-enriched, and basal-like tumors [59]. These subgroups differ in prognosis and therapeutic strategies [60,61]. Thus, it is worth investigating the differences in PUFA metabolism not only between normal and cancer tissues but also between molecular subtypes.

Aiming to address this question, we used five datasets from the GEO database: training sets (GSE81538, GSE25066, and GSE31448) and test sets (GSE96058 and GSE21653). The workflow for further studying the differences in PUFA regulation between tumor subtypes is presented in Figure 4. Due to the platform differences, only 155 genes from the full PUFA list (202 genes) were present in both sets and selected for further study (Supplementary Genes list). No normal-like subtype was considered due to the small number of samples presented in datasets. Our feature selection Boruta-based pipeline (see the Section 2.3.) marked 46 genes as important for separation of four molecular subtypes of breast cancer (Supplementary Genes list). Genes highlighted as important were further used for building the rank Random Forest classifier. The multi-class model had a balanced accuracy of 0.82 and an ROC-AUC of 0.85. As the test was not balanced between classes, it was worth looking at the quality metric for multi-class prediction F1-score, which was 0.75.
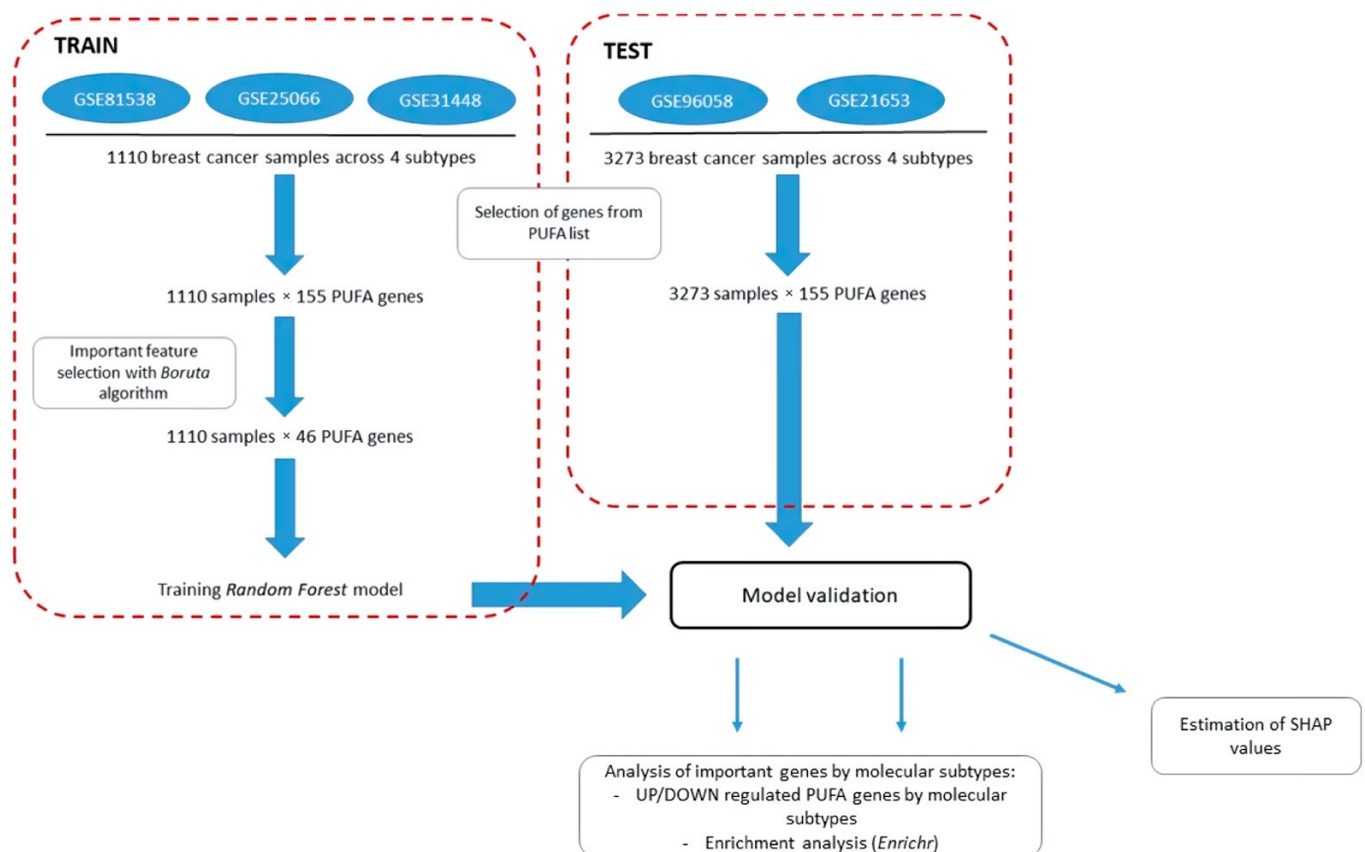

Figure 4. Workflow for studying differences in PUFA metabolism across four molecular subtypes of breast cancer. Processed expression matrixes for GSE81538, GSE25066, GSE31448, GSE21653, and GSE96058 datasets are collected from GEO database. Gene expression levels are ranked, and *PUFA* genes (155 genes) selected in training and test sets. Final model is built on the list of 46 important *PUFA* genes extracted by Boruta selection algorithm. (See details in the Section 2).

The quality descriptors (balanced accuracy, ROC-AUC, and F1 score) of the constructed model show that the expression profiles of PUFA genes differ between the molecular subtypes. The largest number of misclassifications (Figure 5) falls on the luminal subtype (luminal A and luminal B) separation.
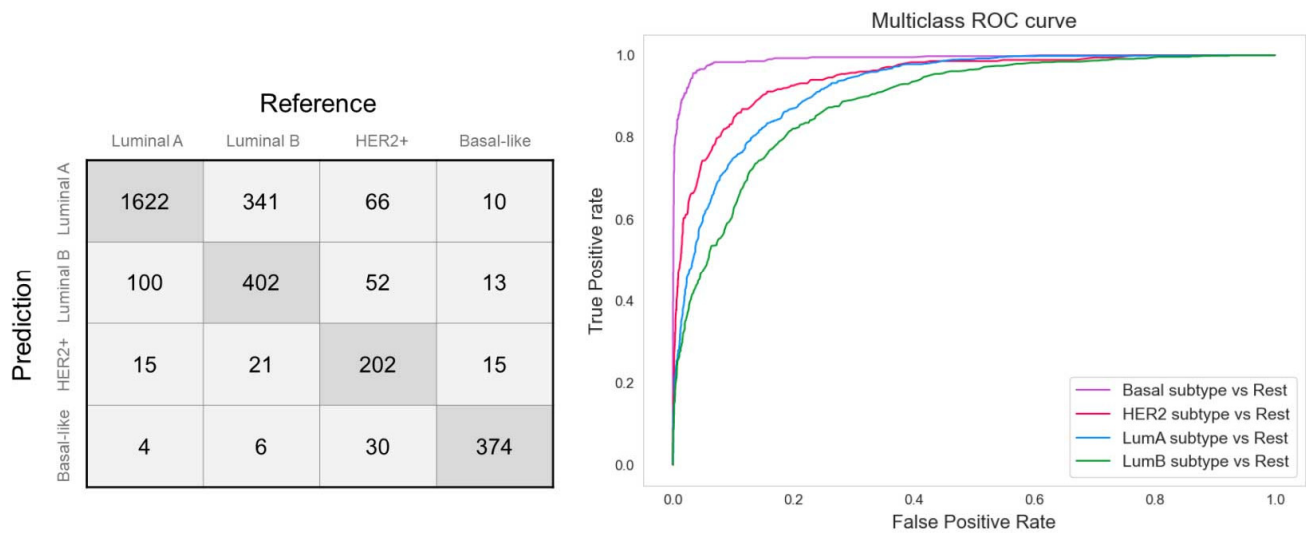

Figure 5. Confusion matrix for breast cancer molecular subtype predictor based on expression of PUFA genes.

The 46 genes selected were analyzed to identify the subtype in which they are significantly (p-value < 0.05) differentially expressed. The analysis was carried out on the largest dataset (GSE96058) from the test set. Table 1 shows the genes in which the expression was significantly increased in the corresponding molecular subtype of cancer. It can be seen that for each subtype, a characteristic set of genes is revealed, most of them attributed to the group of genes responsible for ensuring the functioning of the signaling oxylipin system. The expression values for individual genes are shown in Supplementary Figure S5.

Table 1. Genes that are upregulated in the respective subtype.

Luminal A	Luminal B	HER2+	Basal-Like
ELOVL5	PTGES3	FASN *	AKR1B1
ACAA1 *	ADIPOR1	FABP6	CYP39A1
PLD2	MBOAT7 *	MGLL	PLD1
ACAD8 *	ACOT8 *	ALOX15B	PLA2G4A
PLCL1	CYP2B6	FADS2	FPR2
HPGDS	FAAH		PLCG2
CYP4F11			CYP7B1
PTGER3			FABP5
CYP4F8			PLA2G7
ELOVL2			CBR1
EPHX2			PLAA
LPCAT3 *			ACOT9 *
LTC4S			HSD17B12
FABP4			CYP39A1
			PLA2G2D
			PLCH1

*—assigned to the group of genes responsible for the energy and structural functions of fatty acids; the rest can be attributed to the genes of the signaling oxylipin system.

We investigated the impact of the 46 genes utilized in the subtype classification using SHAP values (see the Section 2.5). The summary plot in Figure 6 shows the top 20 most influential genes. The color bar represents the features' impact on separating the corresponding class from the others. ELOVL5 was the most important gene for overall classification, particularly, for the basal and luminal A subtypes. Rank FABP7 expression made the biggest impact on luminal B separation, while ELOVL2 expression made the biggest impact on the HER2-enriched subtype (Figures 6 and S6).
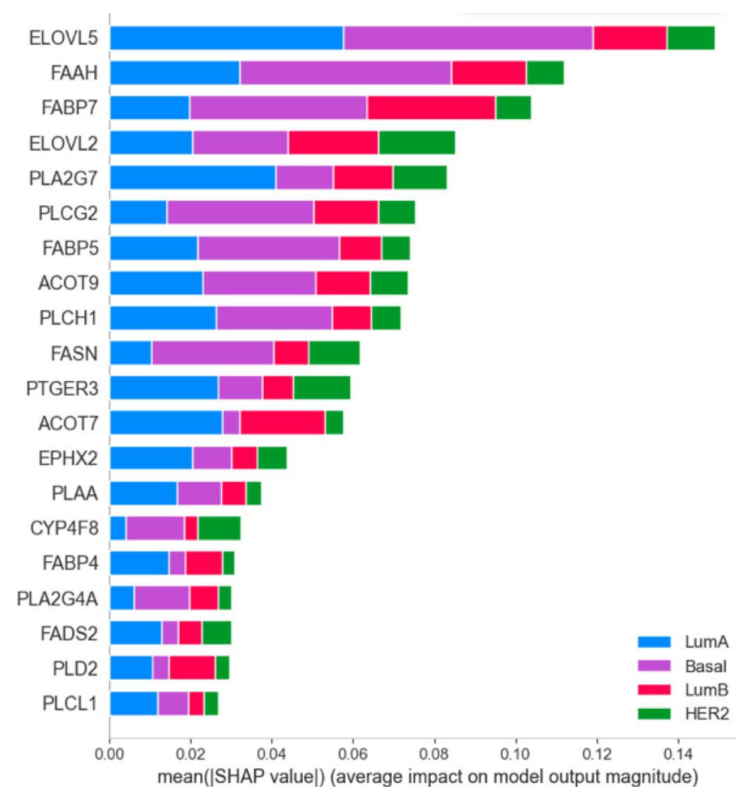

Figure 6. Most important features for classification between molecular subtypes of breast cancer. Importance was evaluated by SHAP values. Each color represents the importance of the separation of the corresponding class from the others.

4. Discussion

Here, we applied a rank Random Forest to the expression data of *PUFA* genes in BRCA to investigate their role in BRCA pathogenesis and in subtype phenotype differences. Our analysis shows that changes in the energy metabolism of PUFA, particularly, in the metabolism of signaling messenger oxylipins are important characteristics that can even be biomarkers for separating patients with BC from healthy people, as well as can determine the nature of molecular subtypes. The use of the feature selection RF-based algorithm Boruta made it possible to identify 33 PUFA metabolism genes that distinguish BC samples from normal tissues and 46 genes that differ between BC subtypes.

It should be noted that the search for biomarkers (signatures) that allows for classifying breast cancer subtypes was carried out earlier (see, for example [62], where using copy number variant data can identify some biomarkers). The focus of our work was to evaluate the role of PUFA metabolic pathways in the biology of subtypes. The impetus for this study was previous work in which we compared the blood profile signatures of oxylipins and PUFAs in 152 healthy volunteers (HC) and 169 patients with various stages of BC [56]. Blood oxylipin signatures reflect the organism's level of response to the disease. We also analyzed the DEGs of ten transcriptome datasets, and 19 genes for oxylipins biosynthesis were among the DEGs [56]. The SNP data for 33 genes related to oxylipin metabolism

analysis reveal that CYP2C19, PTGS2, HPGD, and FAAH were on the list of DEGs in the analysis of transcriptomes and the list of SNPs associated with BC [56]. There is no doubt that PUFA metabolism is involved in BC manifestation, but further research is required to understand the mechanisms of interactions within PUFA metabolic cascades.

The rank RF model built on 33 selected genes showed high performance classifying breast cancer and normal adjacent tissues. The minimal number of genes required for the best performance (ROC-AUC 0.99, ci-bound 0.002; Figure 3) included seven genes (*ADIPOR1*, *HADH*, *ACOT7*, *PTGER4*, *PLA2G15*, *PLA2G1B*, and *CYP46A1*). In this list, *HADH* and *ACOT7* belong to the fatty acid beta-oxidation (FAO) pathway. Previously, for various malignant neoplasms, the so-called "lipolytic phenotype" was shown, in which the FAO pathway was reprogrammed [63,64]. Cancer cells can use changes in FAO metabolism for proliferation, survival, stemming, and metastasis [65]. The other four genes (*ADIPOR1*, *HADH*, *ACOT7*, *PTGER4*, *PLA2G15*, *PLA2G1B*, and *CYP46A1*) can be signed to the PUFA signaling function.

Breast cancer is a highly heterogeneous disease; therefore, it is also essential to understand the diversity in PUFA metabolism across molecular subtypes for further research on the possibility of their use, both as a biomarker and target for therapy. We defined a list of 46 *PUFA* genes differentially across four molecular subtypes. Some of the genes identified in our study in "the distinguishing list" have previously been linked to cancer. It was shown that ELOVL5 (elongase, responsible for elongation of long chain fatty acids) is upregulated in breast cancer (BC) vs. normal adjacent tissue, with the expression correlated with changes in blood lipid species [66]. ELOVL2 expression was associated with malignant phenotypes and suggested as a novel prognostic biomarker in breast cancer [30]. On a cellular model, it was shown that ELOVL2 downregulation is associated with an increased likelihood of metastasis in breast cancer [30]. This is consistent with the data obtained suggesting that the level of ELOVL2 expression has the highest expression in the luminal A subtype, which has the best survival prognosis [4].

The *ACOT* (Acyl-CoA thioesterase) genes were mostly expressed in the basal-like and luminal B molecular subtypes. ACOT enzymes catalyze the hydrolysis of coenzyme A (CoA) esters to free fatty acids and CoA. Further pathways of these fatty acid's metabolism are not completely clear. Additionally, acyl-CoA esters have more functions than simply an energy source, and modulation of their levels via ACOT enzymes activities is important for various pathways of lipid metabolism [67]. It was shown that an increased expression of ACOT1 was correlated with pivotal clinicopathological parameters and poor prognosis in gastric adenocarcinoma [68]. ACOT7 expression increased in lung and breast carcinoma, and low levels of its expression were associated with better survival prognosis [69]. This is also confirmed by the data obtained in the present work. A bar plot with the SHAP values shows the importance of the *ACOT7* gene expression for distinguishing luminal subtypes (Figure 6). Increased expression values of this gene are more likely to indicate the luminal B subtype, which is more aggressive than luminal A (Figure S6).

Fatty acid-binding proteins (FABPs) are involved in binding, storing, and transporting to the appropriate compartments in the cell various fatty acids and other lipophilic ligands such as oxylipins and retinoids. This group of protein is tightly involved in inflammatory processes. Previous studies have revealed that FABP5 [70] and FABP7 [71] might regulate lipid quality and/or quantity to promote aggressiveness such as cell growth, invasiveness, survival, and inflammation in breast cancer cells. FABP7 was suggested as a potential target for the treatment complications of HER2 in breast cancer patients [71]. In our study, we found that a lower expression of this gene is the most important feature for determining luminal B subtype, while its higher expression levels make up the top five important features for basal-like breast cancer (Figure S6). FABP4 was also previously linked to the invasion and migration of colon cancer cells and obesity-associated breast cancer development [72,73]. We showed that the expressions of FABP4 are found in the luminal subtypes of breast cancer; nonetheless, it was not included in the most important hallmark of any subtype (Table 1 and Figure S6).

Twenty of the most important features for classification between molecular subtypes of BC include genes that could be combined into groups of PUFA elongation or desaturation (ELOVL5, ELOVL2, and FADS2); intracellular transport (FABP4, FABP5, and FABP7); release of fatty acids from CoA esters (ACOT7 and ACOT9) and from more complex lipids (phospholipases PLA2G7, PLAA, PLA2G4A, PLCL1, PLCG2, PLCH1, and PLD2); and others, which include six genes attributed to various pathways *FASN* (fatty acid synthase catalyzes elongation of saturated fatty acids), *FAAH* (fatty acid amide hydrolase), *PTGER3* (prostaglandin EP3 receptor), *EPHX2* (soluble epoxide hydrolase), and *CYP4F8* (one of the monooxygenases that is specialized in the metabolism of PUFAs). The list of 185 genes took into account the processes of synthesis and degradation of fatty acids (both saturated and unsaturated), their transformation into oxylipins, and various oxylipin receptors. Interesting to note is that, besides *PTGER3* and *EPHX2*, all other genes from the list in Figure 6 can be attributed to processes that regulate the amounts and species of free fatty acids within cells. It is currently difficult to say why differences between these genes lead to differences in BC subtypes. All of the enzymes corresponding to these genes have been studied in different processes and have not previously been considered as a whole system. It is important that the study indicates the need for such a consideration.

5. Conclusions

Thus, BC subtypes can be discriminated by genes for fatty acid metabolism. A significant part of the genes that differ between subtypes refers specifically to the metabolism of PUFAs and regulatory oxylipins. This supposes that changes in PUFA metabolism are decisive in the manifestation of the subtype phenotypes. The use of rank RF has demonstrated the effectiveness of this approach and has yielded promising results. These results indicate that the genes found for FA metabolites may be potential biomarkers and therapeutic targets for different BC subtypes.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cancers14194663/s1, Figure S1. Validation of the model was provided by the binary classification of head and neck cancer and healthy tissues based on 1000 most variable genes; Table S1. Genes_supplementary.xlsx; Table S2. Datasets selected for binary classification of healthy and tumor breast tissues; Table S3. Datasets selected for multiclass subtype classification of breast cancer; Table S4. Quality of rank RF for binary classification of healthy and tumor breast tissues; Table S5. Genes significantly upregulated in tumor samples (left column) and upregulated in healthy samples (right column); Figure S2. OOB error according to set number of trees (ntree) in the model for binary classification of normal and breast cancer tissues; Figure S3. Enrichment analysis of GO functional and biological pathways, as well as KEGG and WikiPathways pathways by genes important for classification of healthy and tumor breast tissues and upregulated in cancer samples; Figure S4. Enrichment analysis of GO functional and biological pathways, as well as KEGG and WikiPathways pathways by genes important for classification of healthy and tumor breast tissues and upregulated in healthy samples; Figure S5. Expression of FABP6, PLA2G7, ACOT7, FAAH, EPHX2 genes across subtypes; Figure S6. Top-5 most important genes for defining each of four molecular subtypes of breast cancer revealed by SHAP values.

Author Contributions: M.V.G. performed the statistical analysis, the transcriptome analysis, and the ML approaches; writing—original draft preparation, M.V.G., D.V.C. and M.G.S.; writing—review and editing, M.G.S., A.A.M. and A.V.F.; supervision—A.A.M., A.V.F., D.D.P. and M.G.S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are available upon request.

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

References

1. Global Burden of Disease Cancer Collaboration. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol.* **2019**, *5*, 1749–1768. [CrossRef] [PubMed]
2. Polyak, K. Heterogeneity in breast cancer. *J. Clin. Investig.* **2011**, *121*, 3786–3788. [CrossRef] [PubMed]
3. Harbeck, N.; Penault-Llorca, F.; Cortes, J.; Gnant, M.; Houssami, N.; Poortmans, P.; Ruddy, K.; Tsang, J.; Cardoso, F. Breast cancer. *Nat. Rev. Dis. Prim.* **2019**, *5*, 66. [CrossRef] [PubMed]
4. Kensler, K.H.; Sankar, V.N.; Wang, J.; Zhang, X.; Rubadue, C.A.; Baker, G.M.; Parker, J.S.; Hoadley, K.A.; Stancu, A.L.; Pyle, M.E.; et al. PAM50 molecular intrinsic subtypes in the nurses' health Study cohorts. *Cancer Epidemiol. Biomarkers Prev.* **2019**, *28*, 798–806. [CrossRef]
5. Yu, T.J.; Ma, D.; Liu, Y.Y.; Xiao, Y.; Gong, Y.; Jiang, Y.Z.; Shao, Z.M.; Hu, X.; Di, G.H. Bulk and single-cell transcriptome profiling reveal the metabolic heterogeneity in human breast cancers. *Mol. Ther.* **2021**, *29*, 2350–2365. [CrossRef]
6. Frezza, C. Metabolism and cancer: The future is now. *Br. J. Cancer* **2020**, *122*, 133–135. [CrossRef]
7. Heiden, M.G.V.; Cantley, L.C.; Thompson, C.B. Understanding the warburg effect: The metabolic requirements of cell proliferation. *Science* **2009**, *324*, 1029–1033. [CrossRef]
8. Wang, Y.P.; Li, J.T.; Qu, J.; Yin, M.; Lei, Q.Y. Metabolite sensing and signaling in cancer. *J. Biol. Chem.* **2020**, *295*, 11938–11946. [CrossRef]
9. Koundouros, N.; Poulogiannis, G. Reprogramming of fatty acid metabolism in cancer. *Br. J. Cancer* **2020**, *122*, 4–22. [CrossRef]
10. Sampath, H.; Ntambi, J.M. Polyunsaturated fatty acid regulation of gene expression. *Nutr. Rev.* **2004**, *62*, 727–739. [CrossRef]
11. Jabbour, H.N.; Sales, K.J. Prostaglandin receptor signalling and function in human endometrial pathology. *Trends Endocrinol. Metab.* **2004**, *15*, 398–404. [CrossRef] [PubMed]
12. Pakiet, A.; Kobiela, J.; Stepnowski, P.; Sledzinski, T.; Mika, A. Changes in lipids composition and metabolism in colorectal cancer: A review. *Lipids Health Dis.* **2019**, *18*, 1–21. [CrossRef] [PubMed]
13. Gabbs, M.; Leng, S.; Devassy, J.G.; Monirujjaman, M.; Aukema, H.M. Advances in Our Understanding of Oxylipins Derived from Dietary PUFAs. *Adv. Nutr.* **2015**, *6*, 513–540. [CrossRef]
14. Buczynski, M.W.; Dumlao, D.S.; Dennis, E.A. An integrated omics analysis of eicosanoid biology. *J. Lipid Res.* **2009**, *50*, 1015–1038. [CrossRef]
15. Schmid, T.; Brüne, B. Prostanoids and Resolution of Inflammation—Beyond the Lipid-Mediator Class Switch. *Front. Immunol.* **2021**, *12*, 714042. [CrossRef] [PubMed]
16. Chistyakov, D.V.; Astakhova, A.A.; Sergeeva, M.G. Resolution of inflammation and mood disorders. *Exp. Mol. Pathol.* **2018**, *105*, 190–201. [CrossRef]
17. Kotas, M.E.; Medzhitov, R. Homeostasis, inflammation, and disease susceptibility. *Cell* **2015**, *160*, 816–827. [CrossRef]
18. Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: The next generation. *Cell* **2011**, *144*, 646–674. [CrossRef]
19. Johnson, A.M.; Kleczko, E.K.; Nemenoff, R.A. Eicosanoids in Cancer: New Roles in Immunoregulation. *Front. Pharmacol.* **2020**, *11*, 595498. [CrossRef]
20. Wang, D.; DuBois, R.N. Role of prostanoids in gastrointestinal cancer. *J. Clin. Investig.* **2018**, *128*, 2732–2742. [CrossRef]
21. Guillem-Llobat, P.; Dovizio, M.; Bruno, A.; Ricciotti, E.; Cufino, V.; Sacco, A.; Grande, R.; Alberti, S.; Arena, V.; Cirillo, M.; et al. Aspirin prevents colorectal cancer metastasis in mice by splitting the crosstalk between platelets and tumor cells. *Oncotarget* **2016**, *7*, 32462–32477. [CrossRef] [PubMed]
22. Patrignani, P.; Sacco, A.; Sostres, C.; Bruno, A.; Dovizio, M.; Piazuelo, E.; Di Francesco, L.; Contursi, A.; Zucchelli, M.; Schiavone, S.; et al. Low-Dose Aspirin Acetylates Cyclooxygenase-1 in Human Colorectal Mucosa: Implications for the Chemoprevention of Colorectal Cancer. *Clin. Pharmacol. Ther.* **2017**, *102*, 52–61. [CrossRef] [PubMed]
23. Kundu, N.; Ma, X.; Kochel, T.; Goloubeva, O.; Staats, P.; Thompson, K.; Martin, S.; Reader, J.; Take, Y.; Collin, P.; et al. Prostaglandin E receptor EP4 is a therapeutic target in breast cancer cells with stem-like properties. *Breast Cancer Res. Treat.* **2014**, *143*, 19–31. [CrossRef] [PubMed]
24. Markosyan, N.; Chen, E.P.; Smyth, E.M. Targeting COX-2 abrogates mammary tumorigenesis: Breaking cancer-associated suppression of immunosurveillance. *Oncoimmunology* **2014**, *3*, e29287. [CrossRef] [PubMed]
25. Markosyan, N.; Li, J.; Sun, Y.H.; Richman, L.P.; Lin, J.H.; Yan, F.; Quinones, L.; Sela, Y.; Yamazoe, T.; Gordon, N.; et al. Tumor cell-intrinsic EPHA2 suppresses anti-tumor immunity by regulating PTGS2 (COX-2). *J. Clin. Investig.* **2019**, *129*, 3594–3609. [CrossRef]
26. Hanaka, H.; Pawelzik, S.C.; Johnsen, J.I.; Rakonjac, M.; Terawaki, K.; Rasmuson, A.; Sveinbjörnsson, B.; Schumacher, M.C.; Hamberg, M.; Samuelsson, B.; et al. Microsomal prostaglandin E synthase 1 determines tumor growth in vivo of prostate and lung cancer cells. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18757–18762. [CrossRef]
27. Zelenay, S.; Van Der Veen, A.G.; Böttcher, J.P.; Snelgrove, K.J.; Rogers, N.; Acton, S.E.; Chakravarty, P.; Girotti, M.R.; Marais, R.; Quezada, S.A.; et al. Cyclooxygenase-Dependent Tumor Growth through Evasion of Immunity. *Cell* **2015**, *162*, 1257–1270. [CrossRef]
28. Chistyakov, D.V.; Grabeklis, S.; Goriainov, S.V.; Chistyakov, V.V.; Sergeeva, M.G.; Reiser, G. Astrocytes synthesize primary and cyclopentenone prostaglandins that are negative regulators of their proliferation. *Biochem. Biophys. Res. Commun.* **2018**, *500*, 204–210. [CrossRef]

29. Wolf, I.; O'Kelly, J.; Rubinek, T.; Tong, M.; Nguyen, A.; Lin, B.T.; Tai, H.H.; Karlan, B.Y.; Koeffler, H.P. 15-hydroxyprostaglandin dehydrogenase is a tumor suppressor of human breast cancer. *Cancer Res.* **2006**, *66*, 7818–7823. [CrossRef]

30. Kang, Y.P.; Yoon, J.H.; Long, N.P.; Koo, G.B.; Noh, H.J.; Oh, S.J.; Lee, S.B.; Kim, H.M.; Hong, J.Y.; Lee, W.J.; et al. Spheroid-induced epithelial-mesenchymal transition provokes global alterations of breast cancer lipidome: A multi-layered omics analysis. *Front. Oncol.* **2019**, *9*, 145. [CrossRef]

31. Desvergne, B.; Michalik, L.; Wahli, W. Transcriptional regulation of metabolism. *Physiol. Rev.* **2006**, *86*, 465–514. [CrossRef] [PubMed]

32. Rodríguez-Enríquez, S.; Marín-Hernández, Á.; Gallardo-Pérez, J.C.; Pacheco-Velázquez, S.C.; Belmont-Díaz, J.A.; Robledo-Cadena, D.X.; Vargas-Navarro, J.L.; de la Peña, N.A.C.; Saavedra, E.; Moreno-Sánchez, R. Transcriptional Regulation of Energy Metabolism in Cancer Cells. *Cells* **2019**, *8*, 1225. [CrossRef] [PubMed]

33. Soga, T. Cancer metabolism: Key players in metabolic reprogramming. *Cancer Sci.* **2013**, *104*, 275–281. [CrossRef] [PubMed]

34. Mardis, E.R. The challenges of big data. *Dis. Model. Mech.* **2016**, *9*, 483–485. [CrossRef]

35. Schmidt, B.; Hildebrandt, A. Next-generation sequencing: Big data meets high performance computing. *Drug Discov. Today* **2017**, *22*, 712–717. [CrossRef]

36. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

37. Hossain, M.A.; Saiful Islam, S.M.; Quinn, J.M.W.; Huq, F.; Moni, M.A. Machine learning and bioinformatics models to identify gene expression patterns of ovarian cancer associated with disease progression and mortality. *J. Biomed. Inform.* **2019**, *100*, 103313. [CrossRef]

38. Malta, T.M.; Sokolov, A.; Gentles, A.J.; Burzykowski, T.; Poisson, L.; Weinstein, J.N.; Kamińska, B.; Huelsken, J.; Omberg, L.; Gevaert, O.; et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* **2018**, *173*, 338–354.e15. [CrossRef]

39. Geman, D.; D'Avignon, C.; Naiman, D.Q.; Winslow, R.L. Classifying gene expression profiles from pairwise mRNA comparisons. *Stat. Appl. Genet. Mol. Biol.* **2004**, *3*, 1–19. [CrossRef]

40. Tan, A.C.; Naiman, D.Q.; Xu, L.; Winslow, R.L.; Geman, D. Simple decision rules for classifying human cancers from gene expression profiles. *Bioinformatics* **2005**, *21*, 3896–3904. [CrossRef]

41. Acharjee, A.; Larkman, J.; Xu, Y.; Cardoso, V.R.; Gkoutos, G.V. A random forest based biomarker discovery and power analysis framework for diagnostics research. *BMC Med. Genomics* **2020**, *13*, 1–14. [CrossRef] [PubMed]

42. Fortino, V.; Kinaret, P.; Fyhrquist, N.; Alenius, H.; Greco, D. A robust and accurate method for feature selection and prioritization from multi-class OMICs data. *PLoS ONE* **2014**, *9*, e107801. [CrossRef] [PubMed]

43. Liu, S.; Xu, C.; Zhang, Y.; Liu, J.; Yu, B.; Liu, X.; Dehmer, M. Feature selection of gene expression data for Cancer classification using double RBF-kernels. *BMC Bioinform.* **2018**, *19*, 1–14. [CrossRef] [PubMed]

44. Källberg, D.; Vidman, L.; Rydén, P. Comparison of Methods for Feature Selection in Clustering of High-Dimensional RNA-Sequencing Data to Identify Cancer Subtypes. *Front. Genet.* **2021**, *12*, 632620. [CrossRef]

45. Kursa, M.B.; Rudnicki, W.R. Feature Selection with the Boruta Package. *J. Stat. Softw.* **2010**, *36*, 1–13. [CrossRef]

46. Liaw, A.; Wiener, M. Classification and Regression by randomForest. *R News* **2002**, *2*, 18–22.

47. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]

48. Raschka, S. MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack. *J. Open Source Softw.* **2018**, *3*, 638. [CrossRef]

49. SequentialFeatureSelector: The Popular Forward and Backward Feature Selection Approaches Incl. Floating Variants—Mlxtend. Available online: http://rasbt.github.io/mlxtend/user_guide/feature_selection/SequentialFeatureSelector/ (accessed on 29 June 2022).

50. The Gene Ontology Consortium. The Gene Ontology resource: Enriching a GOld mine. *Nucleic Acids Res.* **2021**, *49*, D325–D334. [CrossRef]

51. Kanehisa, M.; Furumichi, M.; Sato, Y.; Ishiguro-Watanabe, M.; Tanabe, M. KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* **2021**, *49*, D545–D551. [CrossRef]

52. Martens, M.; Ammar, A.; Riutta, A.; Waagmeester, A.; Slenter, D.N.; Hanspers, K.; Miller, R.A.; Digles, D.; Lopes, E.N.; Ehrhart, F.; et al. WikiPathways: Connecting communities. *Nucleic Acids Res.* **2021**, *49*, D613–D621. [CrossRef] [PubMed]

53. Tejera, N.; Boeglin, W.E.; Suzuki, T.; Schneider, C. COX-2-dependent and -independent biosynthesis of dihydroxy-arachidonic acids in activated human leukocytes. *J. Lipid Res.* **2012**, *53*, 87–94. [CrossRef] [PubMed]

54. Hajeyah, A.A.; Griffiths, W.J.; Wang, Y.; Finch, A.J.; O'Donnell, V.B. The Biosynthesis of Enzymatically Oxidized Lipids. *Front. Endocrinol.* **2020**, *11*, 591819. [CrossRef] [PubMed]

55. Bryk, M.; Chwastek, J.; Kostrzewa, M.; Mlost, J.; Pędracka, A.; Starowicz, K. Alterations in anandamide synthesis and degradation during osteoarthritis progression in an animal model. *Int. J. Mol. Sci.* **2020**, *21*, 7381. [CrossRef] [PubMed]

56. Chistyakov, D.V.; Guryleva, M.V.; Stepanova, E.S.; Makarenkova, L.M.; Ptitsyna, E.V.; Goriainov, S.V.; Nikolskaya, A.I.; Astakhova, A.A.; Klimenko, A.S.; Bezborodova, O.A.; et al. Multi-Omics Approach Points to the Importance of Oxylipins Metabolism in Early-Stage Breast Cancer. *Cancers* **2022**, *14*, 2041. [CrossRef] [PubMed]

57. Gavrish, G.E.; Chistyakov, D.V.; Sergeeva, M.G. ARGEOS: A new bioinformatic tool for detailed systematics search in GEO and arrayexpress. *Biology* **2021**, *10*, 1026. [CrossRef]

58. The Cancer Genome Atlas (TCGA) Research Network. Comprehensive molecular portraits of human breast tumours. *Nature* **2012**, *490*, 61–70. [CrossRef]

59. Sørlie, T.; Perou, C.M.; Tibshirani, R.; Aas, T.; Geisler, S.; Johnsen, H.; Hastie, T.; Eisen, M.B.; Van De Rijn, M.; Jeffrey, S.S.; et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10869–10874. [CrossRef]

60. Perou, C.M.; Sørlie, T.; Eisen, M.B.; Van De Rijn, M.; Jeffrey, S.S.; Ress, C.A.; Pollack, J.R.; Ross, D.T.; Johnsen, H.; Akslen, L.A.; et al. Molecular portraits of human breast tumours. *Nature* **2000**, *406*, 747–752. [CrossRef]

61. Weigelt, B.; Baehner, F.L.; Reis-Filho, J.S. The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: A retrospective of the last decade. *J. Pathol.* **2010**, *220*, 263–280. [CrossRef]

62. Pan, X.; Hu, X.H.; Zhang, Y.H.; Chen, L.; Zhu, L.C.; Wan, S.B.; Huang, T.; Cai, Y.D. Identification of the copy number variant biomarkers for breast cancer subtypes. *Mol. Genet. Genomics* **2019**, *294*, 95–110. [CrossRef] [PubMed]

63. Shao, H.; Mohamed, E.M.; Xu, G.G.; Waters, M.; Jing, K.; Ma, Y.; Zhang, Y.; Spiegel, S.; Idowu, M.O.; Fang, X. Carnitine palmitoyltransferase 1A functions to repress FoxO transcription factors to allow cell cycle progression in ovarian cancer. *Oncotarget* **2016**, *7*, 3832–3846. [CrossRef] [PubMed]

64. Liu, P.P.; Liu, J.; Jiang, W.Q.; Carew, J.S.; Ogasawara, M.A.; Pelicano, H.; Croce, C.M.; Estrov, Z.; Xu, R.H.; Keating, M.J.; et al. Elimination of chronic lymphocytic leukemia cells in stromal microenvironment by targeting CPT with an antiangina drug perhexiline. *Oncogene* **2016**, *35*, 5663–5673. [CrossRef] [PubMed]

65. Ma, Y.; Temkin, S.M.; Hawkridge, A.M.; Guo, C.; Wang, W.; Wang, X.Y.; Fang, X. Fatty acid oxidation: An emerging facet of metabolic transformation in cancer. *Cancer Lett.* **2018**, *435*, 92–100. [CrossRef]

66. Tomida, S.; Goodenowe, D.B.; Koyama, T.; Ozaki, E.; Kuriyama, N.; Morita, M.; Yamazaki, Y.; Sakaguchi, K.; Uehara, R.; Taguchi, T. Plasmalogen deficiency and overactive fatty acid elongation biomarkers in serum of breast cancer patients pre-and post-surgery—new insights on diagnosis, risk assessment, and disease mechanisms. *Cancers* **2021**, *13*, 4170. [CrossRef]

67. Hunt, M.C.; Alexson, S.E.H. The role Acyl-CoA thioesterases play in mediating intracellular lipid metabolism. *Prog. Lipid Res.* **2002**, *41*, 99–130. [CrossRef]

68. Wang, F.; Wu, J.; Qiu, Z.; Ge, X.; Liu, X.; Zhang, C.; Xu, W.; Wang, F.; Hua, D.; Qi, X.; et al. ACOT1 expression is associated with poor prognosis in gastric adenocarcinoma. *Hum. Pathol.* **2018**, *77*, 35–44. [CrossRef]

69. Jung, S.H.; Lee, H.C.; Hwang, H.J.; Park, H.A.; Moon, Y.A.; Kim, B.C.; Lee, H.M.; Kim, K.P.; Kim, Y.N.; Lee, B.L.; et al. Acyl-CoA thioesterase 7 is involved in cell cycle progression via regulation of PKCζ-p53-p21 signaling pathway. *Cell Death Dis.* **2017**, *8*, e2793. [CrossRef]

70. Senga, S.; Kobayashi, N.; Kawaguchi, K.; Ando, A.; Fujii, H. Fatty acid-binding protein 5 (FABP5) promotes lipolysis of lipid droplets, de novo fatty acid (FA) synthesis and activation of nuclear factor-kappa B (NF-κB) signaling in cancer cells. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **2018**, *1863*, 1057–1067. [CrossRef]

71. Cordero, A.; Kanojia, D.; Miska, J.; Panek, W.K.; Xiao, A.; Han, Y.; Bonamici, N.; Zhou, W.; Xiao, T.; Wu, M.; et al. FABP7 is a key metabolic regulator in HER2+ breast cancer brain metastasis. *Oncogene* **2019**, *38*, 6445–6460. [CrossRef]

72. Tian, W.; Zhang, W.; Zhang, Y.; Zhu, T.; Hua, Y.; Li, H.; Zhang, Q.; Xia, M. FABP4 promotes invasion and metastasis of colon cancer by regulating fatty acid transport. *Cancer Cell Int.* **2020**, *20*, 512. [CrossRef] [PubMed]

73. Zeng, J.; Sauter, E.R.; Li, B. FABP4: A New Player in Obesity-Associated Breast Cancer. *Trends Mol. Med.* **2020**, *26*, 437–440. [CrossRef] [PubMed]