American Journal of Molecular Biology, 2025, 15(2), 185-211
https://www.scirp.org/journal/ajmb
ISSN Online: 2161-6663
ISSN Print: 2161-6620

Role of Estrogen Receptor Beta in Monitoring Hormone-Responsive Breast Cancer Cells by Weighing Gene Fusion Process

Noel Dougba Dago[1,2*], Olefongo Dagnogo[3,4], Martial Didier Saraka Yao[1,2], Jordan Aka[1], Dramane Dagnogo[1,5], Hakim Koudouss Souhaleo Cissé[1], Sonia Bénédicte N'Guessan Koffi[1,5], Rebbecca Eliane Ago Eboulé[1], Joseph Allico Djaman[3,4], Nafan Diarrassouba[1,2], Giovanni Malerba[5]

[1]Department of Genetic and Biochemistry, Biological Sciences Research Unit, Peleforo GON COULIBALY University (UPGC), Korhogo, Côte d'Ivoire
[2]African Center for Shea Research and Application (CRAK), Korhogo, Côte d'Ivoire
[3]Biosciences Training and Research Unit (UFR), Biology and Health Laboratory, Felix HOUPHOUET-BOIGNY University, Abidjan, Côte d'Ivoire
[4]Pasteur Institute of Côte d'Ivoire, Department of Clinical and Fundamental Biochemistry, Abidjan, Côte d'Ivoire
[5]Department of Neurosciences, Biomedical and Movement Sciences, Section of Biology and Genetics, University of Verona, Verona, Italy
Email: *dgnoel7@gmail.com

How to cite this paper: Dago, N.D., Dagnogo, O., Saraka Yao, M.D., Aka, J., Dagnogo, D., Souhalio Cissé, H.K., N'Guessan Koffi, S.B., Ago Eboulé, R.E., Djaman, J.A., Diarrassouba, N. and Malerba, G. (2025) Role of Estrogen Receptor Beta in Monitoring Hormone-Responsive Breast Cancer Cells by Weighing Gene Fusion Process. *American Journal of Molecular Biology*, **15**, 185-211.
https://doi.org/10.4236/ajmb.2025.152014

Received: February 8, 2025
Accepted: March 9, 2025
Published: April 15, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

Abstract

One of the most common cancers is Breast Cancer (BC), representing a worldwide public health concern. We previously showed estrogen receptor β (ERβ) oncosuppressor activity as significantly affected estrogen-induced early transcription and mRNA splicing in hormone-responsive MCF-7 human BC cell models. Since gene fusion phenomena are recurrent in cancer cells, we performed a transcriptomic analysis (RNAseq) aiming to assess gene fusion events in MCF-7 human BC cell models that expressed estrogen nuclear receptors α/β under early estradiol (E2) stimulation. The genomic reads sequences were aligned on GRCh38 human genome by using RNA STAR, while the star-fusion was used to detect gene fusion events. Results showed a non-significant variability regarding gene fusion events happening between estradiol-stimulated MCF-7E expressing ERβ (ERβ+/ERα–) and non-stimulated MCF-7noE repressing ERβ expression (ERβ–/ERα+), MCF-7 human BC cell models ($p > 0.05$). Commonly detected gene fusion events between these two (2) BC cell models result in biomarkers of several cancers and as well BC, and are characterized by intra-chromosomal interactions. Findings revealed five (5) gene fusion events specific to MCF-7noE BC cell models in which ERβ gene expression is repressed (ERβ–/ERα+), and recognized as breast cancer

biomarkers. Interestingly, results exhibited ERβ expression as inhibiting gene fusion phenomena specific to BC (BC biomarkers) in replying to estradiol (E2) stimulus in monitoring early hormone-responsive MCF-7 human BC cell models (ERβ+/ERα–; MCF-7E). Overall, even if early estrogen hormone stimulation by inducing nuclear ERβ has non-significant impact on gene fusion variability between MCF-7noE (ERβ–/ERα+) and MCF-7E (ERβ+/ERα–) BC cell line models by contrast to the alternative splicing event, our study highlighted onco-suppressor activity of ERβ in hormone-responsive BC cell line model by potentially silencing several gene fusion expression recognized as BC biomarkers. However, further investigation is necessary to comprehend how ERβ monitors gene fusion in BC.

Keywords

Hormone-Responsive MCF7 BC, Gene Fusion, Estradiol (E2), Estrogen Nuclear Receptor α and β (ERα and ERβ)

1. Introduction

Gene fusions resulting from intra and/or inter chromosomal rearrangements have been well recognized as a main class of genomic aberrations that are key drivers of oncogenesis [1]. The advent of next generation sequencing technologies has boosted the interest in exploring the role of fusion genes in the development and progression of solid tumors. Of note, gene fusion plays an important role in the initial steps of tumor development [2]. In addition, gene fusions, for the most part, are usually disease specific as opposed to several cancer-associated mutations that claim variable and as well heterogeneous nature of occurrence across different tumor types. Breast Cancer (BC) is known to be a heterogeneous disease with a variety of morphological features and clinical manifestations due to genetic, epigenetic and transcriptomic alterations, *i.e.*, gene fusions [3] [4]. This phenotypic diversity rigorously affects the diagnosis and prognosis of BC. The main difficulties in resolving these issues include the complexities of determining specific markers and the lack of a complete understanding of the cellular hierarchy of the mammary epithelium [5]. Numerous reports have demonstrated that the metastatic status, histological grade, tumour stage, size and receptor expression are the main critical determinants of BC treatment [6] [7]. Immunohistochemically, three broad types of breast tumours have been classified by the status of therapeutically significant components, the Estrogen Receptor ER, the Progesterone-Receptor (PR) and the HER2 [8]. Indeed, Danai *et al.* [9] explored the presence of gene fusion and defined their association with breast cancer subtypes, clinical-pathologic characteristics and copy number aberrations, assigning HER2+ samples as exhibiting significantly more fusions than triple negative and luminal BC subtypes suggesting involvement of estrogen receptor in monitoring and characterizing BC subtypes. In breast cancer, most of the detected gene fusions seem to be passenger

events while the presence of recurrent and driver fusions is still under study. Considering the key role of estrogen receptors alpha (α) and (β) (ERα/β) in controlling molecular pathways, *i.e.*, alternative splicing [7] [10] in the breast cancer development process and because of the gene fusions recurrence in that cancer typology, herein, we assessed gene fusions dynamism in MCF-7 human breast cancer models monitored by estrogen receptors (ERα/β) expression and/or repression. In addition, ERβ is an activator of wild-type *P53*-dependent transcription and is thought to interact with *P53*. The upregulation of ERβ or activation with ERβ agonists results in increased nuclear *P53* expression, showing tumor suppressor activity evidence of that estrogen receptor [11] [12]. Of note, the synergistic effect of ERβ and *P53* inactivation functions is an important aspect of the occurrence and development of breast cancer [13]. Moreover, ERβ's anti-proliferative and pro-apoptosis effects in breast cancer cells involve the interaction of *P53* and ERα. ERβ reduces ERα–*P53* binding by interacting with *P53*, resulting in antagonization of ERα–*P53*-mediated transcriptional regulation. Interestingly, the DNA binding domain and ligand-binding domain of the ERβ protein are 96 and 60% homologous with those of ERα, indicating that they may have similar but not identical functions [14]. ERβ is abundant in the majority of normal breast epithelial cells and claims to be present in 20% - 30% of breast cancers [15]. Considering as a whole, numerous molecular functions mediated by ERβ have been exposed. Lattrich *et al.* [16] suggested Erβ be a protective factor in suppressing uncontrolled breast cell proliferation, mediating concentration-dependent and cell line specific effects on cell growth and gene expression, confirming that ERβ can exert its antitumor effect via gene transcription and miRNA. Due to the recent implication and evidence of gene fusions in characterizing BC, continued efforts are needed to understand the nature and function of estrogen receptors in general (ERβ/ERα) and ERβ in particular in the clinical and molecular diagnosis process of breast cancer. For this purpose, we embarked herein in understanding ERβ role in monitoring hormone responsive MCF-7 human BC by analyzing gene fusions expression using transcriptomic data, *i.e.*, RNAseq.

2. Material and Methods

2.1. Sequences Genomic Read

Healthy breast tissues (HBC) and human hormone-responsive MCF-7 breast cancer (BC) model transcriptomic sequences have been used for this study. Human hormone-responsive MCF-7 BC models were maintained hormone free for 5 days, and successively analyzed for estrogen signaling pathway [7]. Analyzed genomic read sequence from healthy breast tissues (HBC) and human MCF-7 BC cell model retrieved from the National Center for Biotechnology Information (NCBI) nucleotide databases were associated with PRJNA531824 and PRJNA271362 bio-projects [7] [17]. From these bio-projects, we used three biological replicates of healthy mammary cell lines (SRR8874287, SRR8874290, and SRR8874293) and three biological replicates for both BC cell models under early estradiol (E2) stimulation

(MCF-7E) (SRR1737052, SRR1737053, and SRR1737054) and BC cell models not stimulated (MCF-7noE) (SRR1737055, SRR1737056, and SRR1737057).

2.2. Genomic Reads Sequence Quality Control and Alignment Process

We checked for genomic reads sequence quality control by running FastQC (Version 0.11.9). Then, we used the Trimmomatic (Version 0.38.0) with the purpose of trimming reads with low quality score (<20). We also used the options "-LEADING 3" and "-TRAILING 3" to remove reads with poor quality and "-MINLEN 36" to remove short reads (length is less than 36 bases). Then, Cut-adapt has been used to cut Illumina adapters.

2.3. Read Sequences Alignment Process Genome Assembling and Gene Fusion Detection

We used human reference genome GRCh38 gencode version 37 and the reference annotation file (gtf) provided by Trinity Cancer Transcriptome Analysis Toolkit (CTAT) to map genomic reads sequence using RNA STAR (Version 2.7.8a). STAR is an ultra-fast universal RNAseq aligner with a large number of options such as junction output files assessing fusion junction point detection compatible with STAR-Fusion for the detection of fusion genes. We therefore configured RNA STAR activating certain options, which are necessary for the optimal detection of fusion genes. The enabled options are --two pass Mode, --chimOutType Junction, --out SAM attributes. The algorithmic parameters were adjusted on the STAR-Fusion platform. Next, STAR-Fusion bioinformatics (Version 1.10.1) was used for the final detection of fusion genes by processing RNA STAR output files, including genomic read sequence files.

2.4. Functional Genomic Analysis

Functional analysis of genes involved in chromosome inter and intra interaction and/or rearrangement provoking gene fusions were carried out on the Gene Cards platform. Of note, Gene Cards is a searchable and integrative human genomics database that provides comprehensive, user-friendly information on all annotated and predicted human genes. The knowledge base automatically integrates source centric data from approximately 150 web sources, including genomic, transcriptomic, proteomic, genetic, clinical, and functional information.

2.5. Statistical Analysis

We verified MCF-7 breast cancer (BC) cell models and healthy breast tissue/cell (HBC) genomic reads sequences normality by applying Shapiro normality test. Then, we checked for variance difference between HBC, MCF-7E and MCF-7noE samples by performing an ANOVA test as suggested in Hsu (1996) and Stevens (2013) [18] [19]. ANOVA and post-hoc test were visualized on the same plot using R ggstatsplot library and by running the following R script:

```
Library (ggstatsplot)
ggbetweenstats (
data= date,
x= Breast_Cancer_cell,
y= Genomic_Read_Number,
type= “parametric”, #ANOVA or Welch ANOVA
plot.type = “box”,
pairwise.display = “significant”,
centrality.plotting = FALSE,
bf.message = FALSE
)
```

Where “*dat*” refers to matrix data that including genomic read sequences for each healthy and MCF-7 human breast cancer cell samples. In this script before applying ANOVA parametric test, conditions of its application have been verified. Genomic and quantitative data have been transformed and normalized, since normalization procedure in genomic [20] and as well in bio-statistical analysis [21] to prevent data variability bias in making decisions in accepting and/or rejecting the null hypothesis. Gene expression differential analysis was used in this study based on the work of Dago *et al.* (2017) [20]. In addition, the statistical test is considered significant at $p \leq 0.05$.

3. Results

3.1. Statistical Analysis and Assessment of Retrieved Genomic Reads Sequences Quality Control from Healthy Breast Tissue/Cells (HBC) and MCF-7 Hormone Responsive Breast Cancer (BC) Cell Models

We check for read quality for each healthy breast tissue and/or cells (HBC) as well as estradiol (E2) stimulated (MCF-7E) and no-stimulated (MCF-7noE) MCF-7 human BC cell models. Analysis revealed high read quality for all processed MCF-7 BC cells and as well for healthy breast tissues and/or cell samples ($Q > 30$). Shapiro normality test revealed normal distribution of raw genomic read sequences from both MCF-7E and MCF-7noE breast cancer models ($p > 0.05$) by contrast to healthy breast tissue ($p \leq 0.05$) (**Table 1** and **Figure 1**). Descriptive statistical analysis showed a high coefficient of variability (CV) for healthy breast cells (HBC) with regard to MCF-7E and MCF-7noE breast cancer (BC) cell models that exhibited weak CV (**Table 1**). Descriptive statistics displayed comparable coefficient of variability between MCF-7E and MCF-7noE breast cancer cell models (**Table 1**). This result is supported by no significant variance difference between MCF-7E and MCF-7noE ($p = 0.89$) MCF-7 human BC cell models as opposite to HBC samples ($p = 0.04$) (**Figure 1A** and **Figure 1B**). MCF-7E breast cancer cell model displayed the highest genomic read sequence number, while healthy breast tissue and/or cell samples (HBC) exhibited the smallest genomic read number (**Table 1** and **Table S1**). As expected and based on previous results, ANOVA text highlighted signifi-

cant variance difference between MCF7 breast cancer cell models and healthy breast tissues (HBC) ($p \ll 0.05$) (**Figure 1B**). In the same tendency, the Fisher test, by comparing MCF-7E, MCF-7noE and healthy breast cells (HBC), exhibited a strong significant variance difference between MCF-7 BC cell models in general and healthy breast tissues and/or cells (HBC) ($p \ll 0.05$) (**Figure 1B**).

Table 1. Healthy breast tissue/cell (HBC), MCF-7E and MCF-7noE MCF-7 hormone responsive human BC cell models genomic read sequences normality and descriptive statistical analysis.

	Healthy Breast Tissue (HBC)	MCF-7 BC cell models stimulated by estradiol (MCF-7E)	MCF-7 BC cell models non-stimulated by estradiol (MCF-7noE)
Genomic Read Sequences Average (Mean)	39035744	49481221	49059367
Genomic Read Sequences Median	44384131	49446600	48997313
Standard Deviation (SD)	8715685	359982.5	371246.7
Minimum	26744336	49058666	48638898
Maximum	45978764	49938398	49541889
CV (%)	22.33%	0.73%	0.76%
Shapiro Coefficient	0.69	0.85	0.84
p	0.005	0.16	0.13

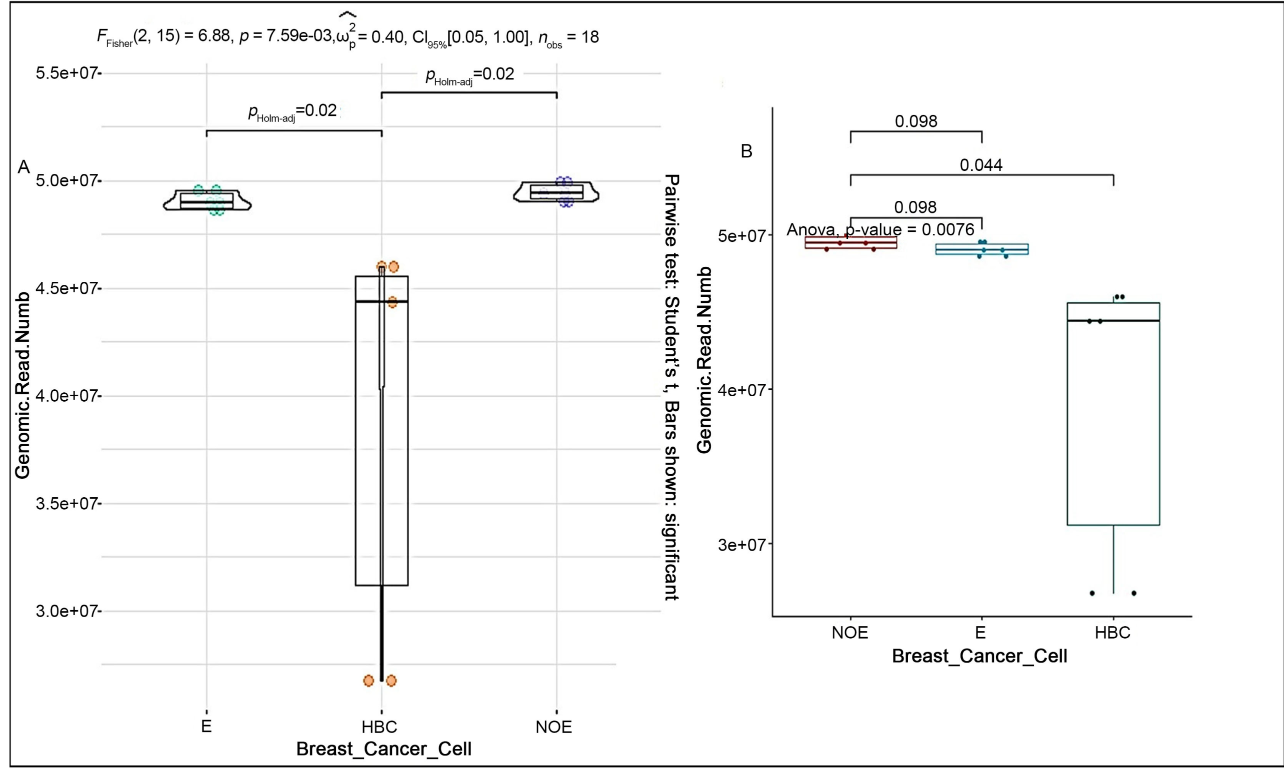

Figure 1. Visualization of post-hoc tests (A) and ANOVA (B) in assessing raw genomic read sequences distribution in healthy breast tissues and/or cells (HBC), estradiol stimulated MCF-7 hormone responsive BC cell models (MCF-7E) and non-estradiol stimulated MCF-7 hormone responsive BC cell models (MCF-7noE). E, NOE and HBC acronyms respectively refer to MCF-7E, MCF-7noE MCF-7 human hormone responsive BC cell models and healthy breast tissues and/or cells.

3.2. Assessment of ESR1 (Estrogen Receptor α ERα) and ESR2 (Estrogen Receptor β ERβ) Gene Expression Profile in MCF-7 Hormone Responsive BC Cell Models

Herein, we checked for gene expression level of ESR1 estrogen receptor α (ERα) and ESR2 estrogen receptor β (ERβ) in MCF-7 hormone responsive breast cancer cell models, stimulated (MCF-7E) and/or no-stimulated (MCF-7noE) by estradiol. Gene expression of ESR1 and ESR2 have been measured in fragment per million kilo-bases, FPKM normalized gene expression measure (Figure 2). MCF-7E and MCF-7noE showed different profiles of expression of ERβ+/ERα– and ERβ–/ERα+. The analysis demonstrated a high-level expression of the ESR2 (ERβ) gene in MCF-7E, in contrast to the expression level of ESR1 (ERα) (Figure 2). The opposite is observed in MCF-7noE, where the expression level of ESR2 (ERβ) decreased, whereas the expression level of ESR1 (ERα) increased (Figure 2). Therefore, MCF-7E and MCF-7noE human hormone responsive BC cell models respectively display ERβ+/ERα– and ERβ–/ERα+ estrogen receptors genes expression profile (Figure 2). However, Chi-square test relatively supported ERβ expression and as well ERα repression in MCF-7E human hormone BC cell models and also suggested a relatively significant repression of ERβ gene expression in MCF-7noE hormone responsive human BC cell models as opposite to ERα expression (calculated Chi-square coefficient = 0.33 < Chi-square coefficient on read Chi table at significant threshold 0.05; degree of freedom 1 = 3.84). Differential expression of gene analysis showed a significant negative Fold Change (FC) value (FC = −0.17, p = 0.04) by measuring ERS1 (ERα) gene expression level in MCF-7E vs. MCF7-noE contrast. The same analysis exhibited a relatively significant positive FC value (FC = 0.22, p = 0.37) in quantifying ERS2 (ERβ) gene expression in MCF-7E vs. MCF7-noE hormone responsive BC cell models contrast. Overall, early estradiol stimulation in hormone responsive MCF-7 human BC models relatively enhanced ERβ gene expression by repressing ERα expression (ERβ+/ERα–).

3.3. Descriptive and Analytical Statistical Analysis of Healthy Breast Tissue (HBC), ERβ+/ERα– MCF-7E and ERβ–/ERα+ MCF-7noE MCF-7 Human Breast Cancer Cell Models Genomic Read Sequences Aligned on Hg38 Human Genome

Analysis revealed adaptor sequences linked to MCF-7 breast cancer (BC) cells genomic read sequences as opposite to healthy breast cells. After adaptor trimming, 70.01% - 91.22% of ERβ+/ERα– MCF-7E and ERβ–/ERα+ MCF-7noE reads were eligible for the subjacent sequences alignment analysis (Table S2). A total of 232.3, 233.7 and 265.2 million read for HBC, ERβ–/ERα+ MCF7noE and ERβ+/ERα– MCF7E respectively have been aligned on the reference genome. Indeed, more than 98% of these genomic read sequences for each sample were correctly aligned on the Hg38 human genome (Table 2). Shapiro test showed normal distribution of the reads HBC (p = 0.17), ERβ+/ERα– MCF-7E (p = 0.75) and ERβ–/ERα+ MCF-7noE (p = 0.36) human breast cancer cell models aligned on the Hg38 human genome (p > 0.05). Although the variability coefficient displayed a relatively high value for

HBC aligned reads (CV = 22.22%) in contrast to those of both ERβ+/ERα– MCF-7E (0.61%) and ERβ–/ERα+ MCF-7noE human MCF-7 BC models samples (CV = 55%), Fisher test suggested no significant variance difference between these samples (p = 0.5) considering genomic reads sequences aligned on Hg38 human genome (**Figure 3**). Considering as a whole, reads resulting correctly aligned to the Hg38 human genome, healthy breast tissue/cell (HBC), ERβ+/ERα– MCF-7E and ERβ–/ERα+ MCF-7noE samples exhibited the right statistical aptitude in terms of data normality for subjacent genomic comparative analysis.

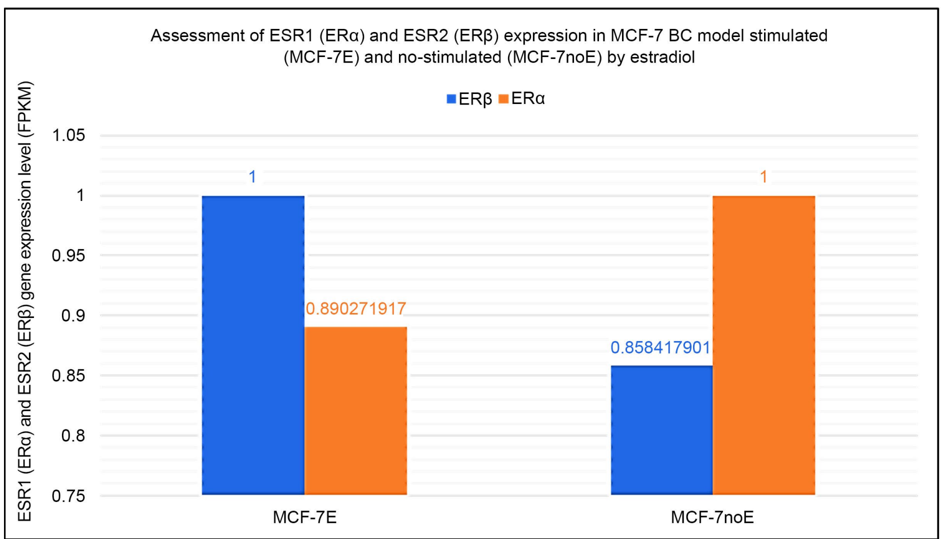

Figure 2. ESR1 or estrogen receptor α (ERα) and ESR2 or estrogen receptor β (ERβ) gene expression profiles in MCF-7 hormone responsive human BC cell models stimulated (MCF-7E) and/or no stimulated (MCF-7noE) by estradiol. For cross comparative analysis between MCF-7E and MCF-7noE hormone responsive BC cell models, we practiced the maximum normalization method [21] on normalized FPKM gene expression values.

Table 2. Descriptive statistic of healthy breast tissues/cells (HBC), MCF-7E (ERβ+/ERα–) and MCF-7noE (ERβ–/ERα–) MCF-7 breast cancer (BC) cell models genomic read sequences aligned on the Hg38 human genome.

Breast healthy tissue (HBC) and MCF-7 breast cancer cell models	Read sequences correctly aligned on Hg38 human genome (Million)	% pair of mapped genomic read sequences	% error in genomic read sequences alignment	Total number of genomic read sequences (Million)
HBC1	91.5	99.50%	0.70%	92.0
HBC2	53.2	99.40%	1.00%	53.5
HBC3	87.6	98.70%	0.30%	88.8
Total HBC	**232.3**			
MCF-7E1	88.3	98.90%	1.40%	89.3
MCF-7E2	89.1	98.90%	1.40%	90.1
MCF-7E3	87.8	98.90%	1.40%	88.7
Total MCF-7E	**265.2**			
MCF-7noE1	78.1	99.10%	3.10%	78.8
MCF-7noE2	78.3	99.20%	3.00%	79.0
MCF-7noE3	77.3	99.10%	3.10%	78.0
Total MCF-7noE	**233.7**			

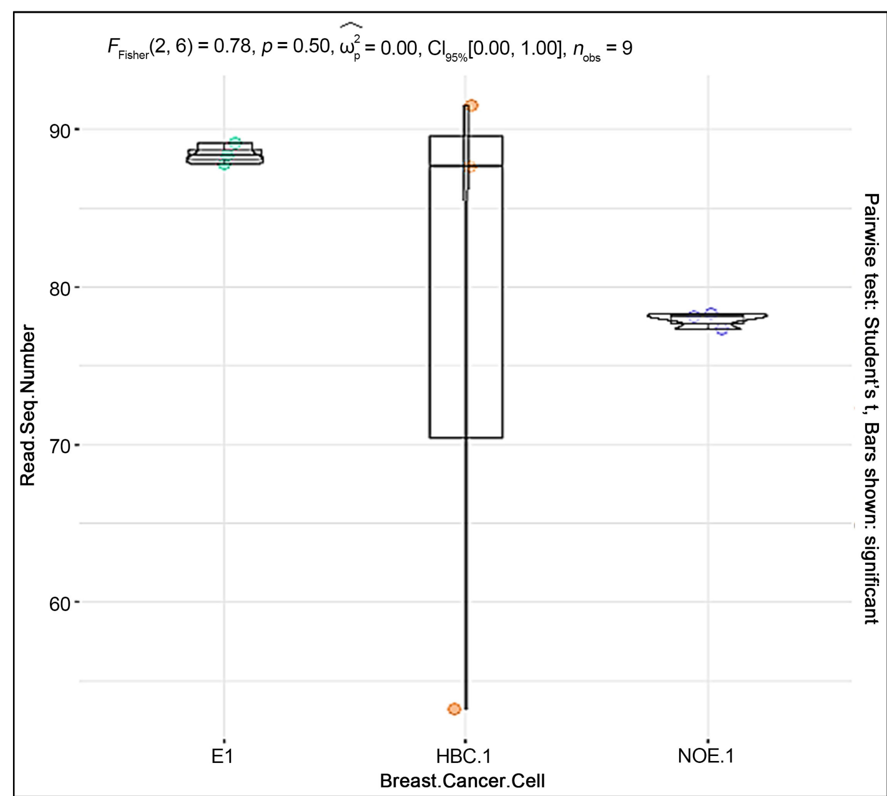

Figure 3. Healthy breast tissues/cells (HBC), MCF-7E expressing ER*β* (ER*β*+/ER*α*–) and MCF-7noE repressing ER*β* expression (ER*β*–/ER*α*+) MCF-7 breast cancer cell models box-plot multi-comparative Fisher test evaluating aligned genomic read sequences variability. E1, NOE.1 and HBC.1 acronyms refer respectively to ER*β*+/ER*α*– MCF-7E, ER*β*–/ER*α*+ MCF-7noE MCF-7 human BC cell models and healthy breast tissues/cells samples.

3.4. Evaluation of the Relationship between Healthy Breast Tissues/Cells (HBC), MCF-7E (ER*β*+/ER*α*–) and MCF-7noE (ER*β*–/ER*α*+) BC Cell Models Samples Biological Replicates

We previously showed weak variability regarding MCF-7 breast cancer (BC) cell samples as opposed to healthy breast tissues/cells (HBC). Indeed, we measured the Euclidian distance between biological replicates of ER*β*+/ER*α*– MCF-7E, ER*β*–/ER*α*+ MCF-7noE human BC cell models and healthy breast tissues/cells (HBC) samples. Results highlighted weak Euclidian distance between MCF-7 breast cancer (BC) cell models biological replicates as opposite to those of healthy breast tissues/cells (HBC) samples (Figure 4). Therefore, Euclidian distance analysis suggested excellent hierarchical clustering between MCF-7E (ER*β*+/ER*α*–) and MCF-7noE (ER*β*–/ER*α*+) human BC cell models biological replicates by contrast to biological replicates of healthy breast tissue/cells (HBC). Of note, Euclidean distance analysis suggested relatively high variability between healthy breast tissues (HBC) biological replicates in comparison to those of MCF-7E and MCF-7noE human BC cell models (Figure 4). Principal component analysis (PCA) showed high variability between HBC samples biological replicates and those of both MCF-7E (ER*β*+/ER*α*–) and MCF-7noE (ER*β*–/ER*α*+) human BC model

samples by its first component explaining 97% of data variability (Figure S1). However, even if healthy breast tissues/cells (HBC) biological replicates exhibited a high Euclidian distance difference between them compared to MCF-7E and MCF-7noE human BC cell model samples, principal component analysis (PCA) by its first component suggested a good clustering ratio regarding HBC biological replicates (Figure S1). Taking together, clustering analysis by Euclidian distance as well as by PCA revealed weak intra-data variability regarding biological replicates of the healthy breast tissues/cells (HBC), MCF7-E (ERβ+/ERα–) and MCF-7noE (ERβ–/ERα+) MCF-7 hormone responsive human BC cell model samples.

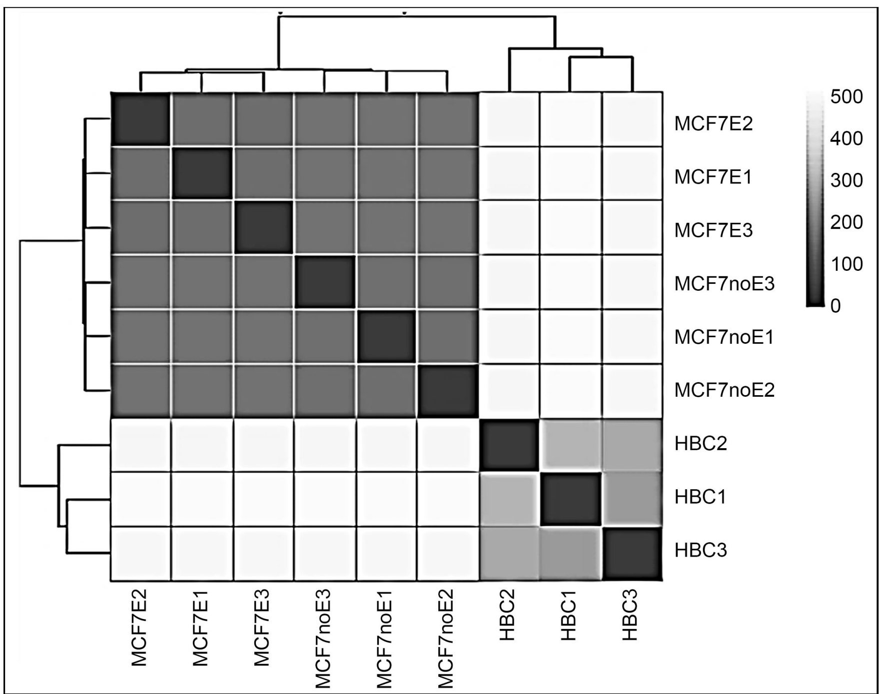

Figure 4. Euclidean distance for hierarchical clustering of healthy breast tissues/cells (HBC), ERβ+/ERα– MCF-7E (MCF-7 BC cell models stimulated by estradiol E2 and expressing ERβ gene) and ERβ–/ERα+ MCF-7noE (MCF-7 human BC cell models no-stimulated by estradiol E2 and repressing ERβ gene expression) MCF-7 human BC cell models biological replicate samples.

3.5. Assessment of Gene Fusion Events in Healthy Breast Cells (HBC), ERβ+/ERα– (MCF-7E) and ERβ–/ERα+ (MCF-7noE) MCF-7 Human BC Cell Models

Genomic analysis revealed several gene fusion events in MCF-7 human BC cell models expressing ERβ (ERβ+/ERα–; MCF-7E) and repressing ERβ expression (ERβ–/ERα+; MCF-7noE) in monitoring estradiol pattern in hormone responsive MCF-7 human BC models by contrast to healthy breast cells (HBC) that reported none gene fusion events (Table 3 and Table S3). Indeed, genomic analysis revealed 30 gene fusion events for MCF-7E expressing ERβ (ERβ+/ERα–) and MCF-7noE BC cell models repressing ERβ expression (ERβ–/ERα+) (Table 3). Of note, 24 genes fusion events out of the above mentioned 30 genes fusion events, resulted to be com-

mon to the estradiol stimulated MCF-7 human BC cell models (ERβ+/ERα–; MCF-7E) and no-estradiol stimulated MCF-7 human BC cell models (ERβ–/ERα+; MCF-7noE) (**Table 3** and **Figure 5**). In addition, the analysis revealed five (5) gene fusion events specific to ERβ–/ERα+ MCF-7noE, while ERβ+/ERα– MCF-7E exhibited only one partial gene fusion event (**Figure 5**, **Table 3** and **Table S3**). Furthermore, the same genomic analysis revealed several partial gene fusion phenomena, *i.e.*, AC099850.1-VMP1, TTC6-AL121790.1 and AL121578.1-SYTL5 commonly expressed in ERβ+/ERα– (MCF-7E) and ERβ–/ERα+ (MCF-7noE) hormone responsive MCF-7 human BC cell models (**Table 3**).

Table 3. Gene fusion events recorded in ERβ+/ERα– MCF-7E and ERβ–/ERα+ MCF-7noE human BC cell models assessing estradiol nuclear receptor beta (ERβ) activity in MCF-7 hormone responsive breast cancer models.

Gene fusion phenomena retrieved in ERβ+/ERα– (MCF-7E) and ERβ–/ERα+ (MCF-7noE) MCF-7 human BC cell models	Gene fusion events recorded exclusively in ERβ–/ERα+(MCF-7noE) MCF-7 human BC cell models	Gene fusion event archived exclusively in ERβ+/ERα– (MCF-7E) MCF-7 human BC cell models
RPS6KB1--VMP1	RPS6KB1--DIAPH3	DENND2C--**AC005914.1***
BCAS4--BCAS3	HIPK1--DENND2C	
ATP1A1--ZFP64	AHCYL1--RAD51C	
ARFGEF2--SULF2	BCAS3--ATXN7	
EVL--EML1	ABCA5--PPP4R1L	
DEPDC1B--ELOVL7		
TANC2--CA4		
ADAMTS19--SLC27A6		
HACD2--MYLK		
TPTE--BAGE2		
ESR1--CCDC170		
TXLNG--SYAP1		
SULF2--PRICKLE2		
MTAP--CDKN2B-AS1		
MYH9--EIF3D		
*AC099850.1--VMP1		
RSBN1--AP4B1-AS1		
POP1--MATN2		
SMARCC1--EMCN		
TTC6--AL121790.1*		
*AL121578.1--SYTL5		
GATAD2B--NUP210L		
EIF3H--RAD21		
PYROXD2--HPSE2		

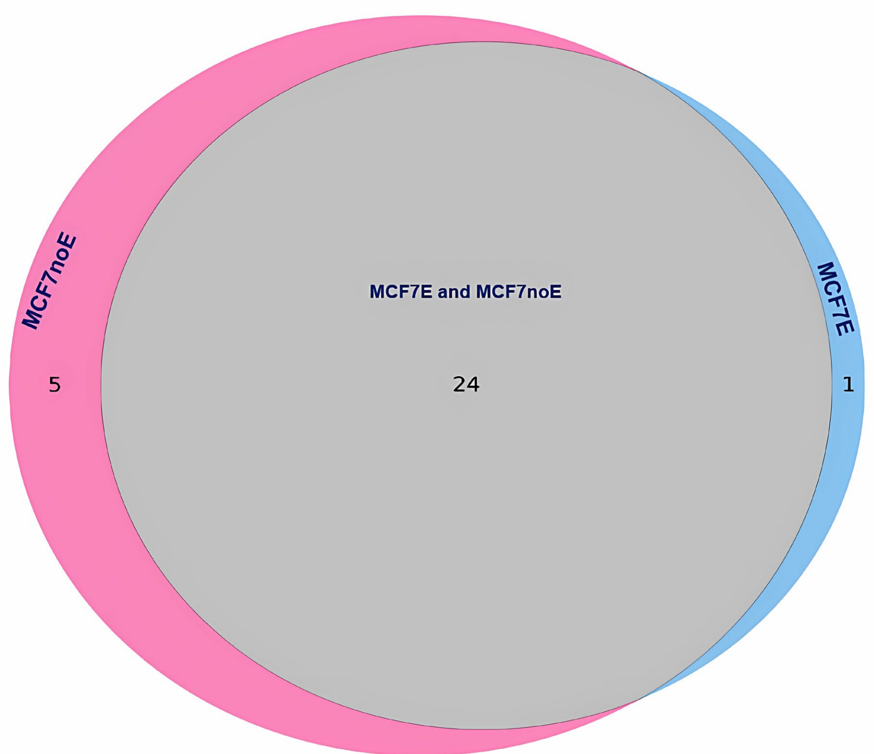

Figure 5. Venn diagram of the 30 gene fusion events retrieved in ERβ+/ERα– MCF-7E and ERβ–/ERα– MCF-7noE human BC cell models in measuring tumor suppressor activity of estrogen receptor beta (ERβ) in hormone responsive MCF-7 BC cell models. MCF-7E and MCF-7noE respectively refer to human MCF-7 BC cell models stimulated by estradiol, expressing ERβ (ERβ+ /ERα+) and to MCF-7 human BC cell models no stimulated by estradiol repressing ERβ gene expression (ERβ–/ERα+).

3.6. Evaluation of Chromosomic Interactions Favoring Gene Fusion Events in MCF-7E and MCF-7noE Hormone Responsive Breast Cancer (BC) Cell Models

Analysis of gene fusion events in MCF-7E expressing ERβ (ERβ+/ERα–) and MCF-7noE repressing ERβ (ERβ–/ERα+) human BC cell models suggested two types of inter-chromosomal interactions (Figure 6A). Indeed, it is noteworthy to underline that among the chromosomal interactions regarding the 24 genes fusion events commune to the MCF-7E (ERβ+/ERα–) and MCF-7noE (ERβ–/ERα+) human BC cell models, 4 claimed to be inter-chromosomal interactions, while the others 20 resulted to be intra-chromosomal interactions (Figure 6A). The inter-chromosomal interactions are established between chromosomes 20 and 1, chromosomes 20 and 3, chromosomes 20 and 17 and between chromosomes 3 and 4 (Figure 6A). Intra-chromosomal interactions involve chromosomes 1, 3, 4, 5, 6, 8, 9, 10, 14, 17, 20, 21, 22 and X (Figure 6A). The same analysis suggested 4 inter-chromosomal interactions regarding the 5 gene fusion biomarkers specific to MCF7noE (Figure 6B). These chromosomal interactions were established between chromosomes 1 and 17, chromosomes 3 and 17, chromosomes 17 and 13 and chromosomes 17 and 20 (Figure 6B). Gene fusion event retrieved in MCF-7E

expressing ERβ is a partial inter-chromosomal fusion and occurs between chromosomes 1 and 20 (**Figure 6B**). Taking together, our results revealed a high level of intra-chromosomal interactions (70%) in comparison to inter-chromosomal interactions (30%) as well as the high involvement of chromosomes 17 and 20 in the intra-chromosomal interactions in evaluating ERβ onco-suppressor activity in monitoring hormone responsive human MCF-7 BC cells.

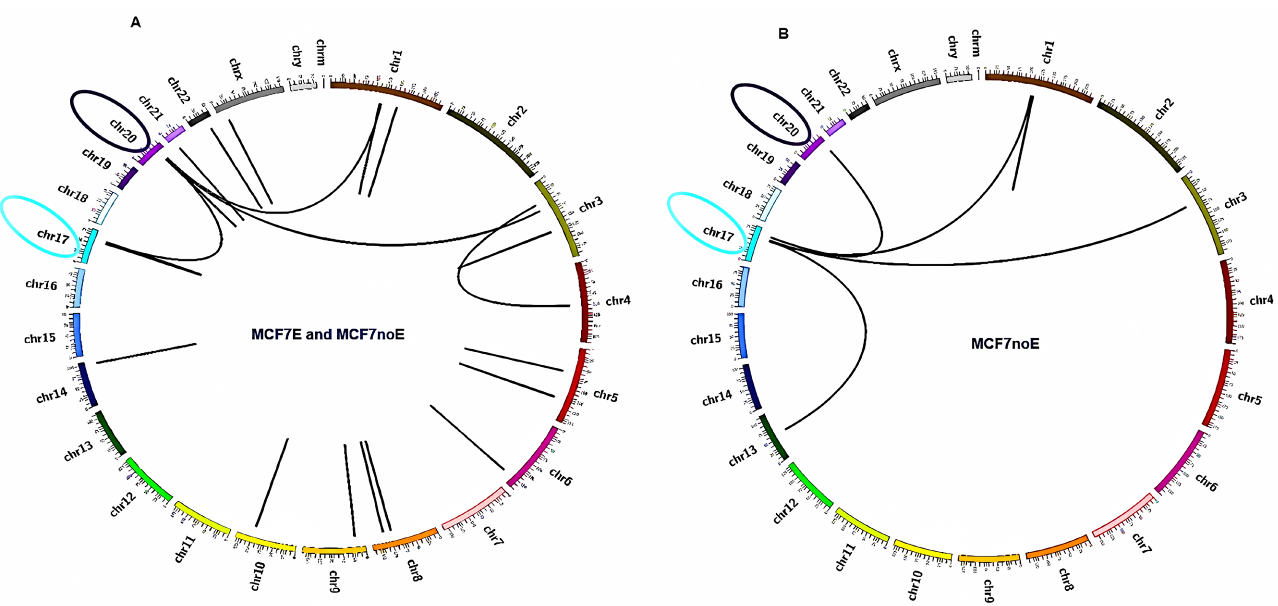

Figure 6. Circlos plot describing intra and inter chromosomal interactions favoring gene fusion phenomena in MCF-7E (ERβ+/ERα–) and MCF7noE (ERβ– ERα+) hormone responsive human BC cell models (A) and in MCF-7noE (ERβ–/ERα+) hormone responsive human BC cell model exclusively (B).

3.7. Gene Fusion Expression Comparative Analysis in MCF-7E (ERβ+/ERα–) and MCF-7noE (ERβ–/ERα+) Breast Cancer (BC) Cell Models Samples

We compared MCF7E (ERβ+) and MCF7noE (ERβ–) breast cancer (BC) cell samples in terms of gene fusion events frequency. Considering the three biological replicates of each MCF7E and MCF7noE BC cell model, the Student test revealed no significant difference between early estradiol (E2) stimulated and no stimulated (MCF7E and MCF7noE) breast cancer cell models ($p = 0.18$). We performed Chi-square test aiming to evaluate the interdependence between early estradiol stimulation of MCF7 breast cancer cells and gene fusion events occurrence. Computed Chi-square test exhibited a calculated Chi coefficient = 0.53. The theoretical Chi coefficient at $\alpha = 0.05$ threshold for 2 degrees of freedom is estimated at 5.99. Because the calculated Chi coefficient is lower than those of theoretical Chi value, we accept the null hypothesis, in other words, there is no significant correlation between estradiol (E2) early stimulation of MCF-7 (ERβ+/ERα–; MCF-7E) breast cancer (BC) cell models and the dynamism of gene fusion process in monitoring breast cancer process. In other words, gene fusions occurring in MCF-7 human

BC cell models do not depend on early estradiol stimulus.

3.8. Functional Genomic Analysis of Genes Fusion Events Commonly Retrieved in MCF-7E (ERβ+/ERα–) and MCF-7noE (ERβ–/ERα+) Breast Cancer Cell Models

Functional genomic analysis revealed that among gene fusion commonly expressed in both hormone responsive MCF-7E (ERβ+/ERα–) and MCF-7noE (ERβ–/ERα+) BC cell models, six (6) are recognized as breast cancer biomarkers (ARFGEF2--SULF2, BCAS4--BCAS3, DEPDC1B--ELOVL7, ESR1--CCDC170, EIF3H--RAD21, SULF2--PRICKLE2) (Table S3). Three (3) gene fusions out of these six (6) BC biomarkers, used chromosome 20 in their inter-chromosomal interaction. In addition, biological functions associated with the above-mentioned breast cancer biomarkers are as follows: activation of ribosylation factors, vesicular transport, lipid and protein metabolism, necrosis tumor, cell signaling, growth factor co-receptors, nucleic acid regulation, cell migration, translational control, DNA break repair, ectoderm differentiation, apoptosis (Table S3). Among those gene fusions, 1) EIF3H--RAD21 and 2) ESR1--CCDC170 gene interactions were recognized as prostate, ovarian and endometrial cancers biomarkers respectively. In addition, retrieved RPS6KB1--VMP1 gene fusion event in both MCF-7E (ERβ+/ERα–) and MCF-7noE (ERβ–/ERα+) BC cell models claimed to be several cancer types molecular biomarkers and result to be associated to the following biological functions: 1) cell signaling, 2) cell growth, 3) cell proliferation, 4) translational control and 5) cell junction formation (Table S3). The majority of the gene fusion events recorded in MCF-7E (ERβ+/ERα–) and MCF-7noE (ERβ–/ERα+) human BC cell models result in several cancer types biomarkers. Thus, the following gene fusion events AL121578.1--SYTL5, EVL--EML1, MTAP--CDKN2B-AS1, MYH9--EIF3D, SMARCC1--EMCN and TPTE--BAGE2 detected in the present study are respectively associated as biomarkers to bladder, gastrointestinal adenoma, bone, colon, and skin cancers. In addition, genomic analysis revealed 3 gene fusion phenomena, *i.e.*, AC099850.1--VMP1, TTC6--AL121790.1 and AL121578.1--SYTL5 that are not reported in the genomic databases (Table S3).

3.9. Genomic Characterization and Analysis of Gene Fusion Events Specific to MCF-7E (ERβ+/ERα–) and MCF-7noE (ERβ–/ERα+) Breast Cancer Cell Models

Analysis revealed 4 gene fusion events, *i.e.*, ABCA5--PPP4R1L, AHCYL1--RAD51C, BCAS3--ATXN7 and RPS6KB1--DIAPH3 specific to MCF-7noE (ERβ–/ERα+) BC cell models that have been recognized as breast cancer biomarkers (Table 3 and Table S3), while HIPK1--DENND2C gene fusion event specific the same breast cancer cell model (MCF-7noE) claim to be a biomarker of several cancer types (Table S3). AHCYL1--RAD51C gene fusion event recorded in MCF-7noE (ERβ–/ERα+) BC cell model results to be associated with ovarian cancer (Table S3). In addition, biological functions linked to gene involved in the following

chromosomic fusion events: 1) ABCA5--PPP4R1L, 2) HIPK1--DENND2C and 3) RPS6KB1--DIAPH3, specific to MCF-7noE can be classified into two groups as follows: 1) classical functions such as molecular transport, l phosphorylation activity, transcription regulation, catalytic regulation, protein regulation, translational control, actin regulation, cellular regulation and 2) functions involved in tumorigenesis such as apoptosis, angiogenesis, squamous cell tumor formation, cell growth and proliferation (**Table S3**). All gene fusion was specific to MCF-7noE (ERβ–/ERα+) BC cell model resulting in being associated with genetic diseases (*i.e.*, neuropathy, ataxia, etc.) (**Table S4**). However, functional genomic analysis revealed DENND2C--AC005914.1, a partial fusion specific to MCF-7E human BC cell models expressing ERβ (ERβ+/ERα–). The head of this partial chromosomic fusion event, DENND2C is involved in the following biological functions: regulation of catalytic activity, guanine nucleotide exchange factor, vesicle-mediated transporter, and regulation of rab proteins (**Table S3**).

4. Discussion

Herein, in this study, we evaluated estrogen receptor beta (ERβ) involvement in monitoring hormone responsive MCF-7 breast cancer (BC) cell models by weighing the gene fusion process. We previously demonstrated that ERβ expression (ERβ+/ERα–) significantly affects estrogen induced early transcription and mRNA splicing in hormone responsive MCF-7 BC cells, providing novel information on the biological role of ERβ in these tumors [7] [10]. Several lines of evidence suggest that the inhibitory effect of ERβ is due to blockade of cell cycle progression. ERβ knockdown (ERβ–/ERα+) results in significant growth of several breast cancer cells, accompanied by elevated cyclin A2 expression [22] and mitofusin 2 (mfn2) [23], suggesting tumor suppressor activity of this (ERβ) estrogen receptor. Because genomic arrangement analysis by the advancement in next generation sequencing technology has led to improving various genomic and transcriptomic studies [24] [25] and, in particular, highlighting gene fusion events in characterizing and monitoring breast cancer [26], we embarked here in understanding the role of ERβ in monitoring hormone responsive MCF-7 BC cells, by assessing gene fusion pattern. The first studies relating to the process of gene fusion in breast cancer date back to 2011 [26]. For many years, it was assumed that gene fusions were a type of mutation confined largely to leukemias and sarcomas. However, gene fusion is now known to be important in several epithelial cancers and a number have been described in breast cancers [27]. Gene fusion events are preferentially more frequent in aggressive breast cancers such as luminal B, basal-like or endocrine-resistant breast cancer, serving as an additional key driver event [28]. Recurrent fusion genes have been identified in different subtypes of breast cancer. However, further investigations are required to establish their significant role in oncogenesis [27]. Genomic read quality control is an important step for subsequent genomic analysis. Indeed, genomic read sequences of HBC, MCF-7E (ERβ+/ERα–) and MCF-7noE (ERβ–/ERα+) BC cell models used in this study exhibited high

quality score as well as high alignment accuracy guarantying exactness of subjacent biological, molecular and genomic analysis [29]. The RNA-STAR and STAR-Fusion analysis modules produced statistically stable results with high precision, high sensitivity for mapping and detection of fusion genes from RNA-seq data [30] [31]. We performed hierarchical clustering analysis between healthy breast and breast cancer cells biological replicates by performing a computational statistical analysis based on 1) principal component analysis, 2) Euclidean distance of Pearson correlation and 3) ANOVA test. Findings showed variance homogeneity between HBC, MCF-7E and MCF-7noE cells, suggesting no bias in the comparative genomic analysis in assessing gene fusion events in hormone responsive induced MCF-7 BC cell models. Indeed, numerous studies supported the integration between bioinformatics and biostatistics approaches as a right practice in improving complex genomic and transcriptomic analysis [20] [29] [32] [33]. As expected, no gene fusion expression was revealed in healthy breast tissues (HBC) in contrast to MCF-7 hormone-responsive breast cancer cells. In other words, several gene fusion events were recorded in the human MCF-7 stable BC cell models expressing ERβ (ERβ+/ERα–; MCF-7E) as well as not expressing ERβ (ERβ–/ERα+; MCF-7noE). In this study, we associated ERβ expression (ERβ+/ERα–) to MCF-7 BC cell models under early estradiol stimulation referred as MCF-7E, while the repression of that receptor (ERβ–/ERα+) is linked to no estradiol stimulated MCF-7 BC cell models named MCF-7noE. Based on this evidence, findings suggested ERβ expression (ERβ+/ERα–) as unable to monitor differential gene fusion expression in breast cancer [34] [35]. However, several studies have shown the major role of estrogens in the molecular development of breast cancer [36]-[38]. Thus, Dago *et al.* [7] showed for the first time the recurrence of the alternative splicing process in MCF-7 BC cells expressing estrogen receptor beta (ERβ+/ERα–) in response to early stimulation to estradiol (E2) as opposite to wild type MCF-7 cells, which do not express endogenous ERβ (ERβ–/ERα+). By contrast to the alternative splicing process, expression of ERβ in human MCF-7 cells (MCF-7 expressing ERβ+ referred to as MCF-7E) cannot significantly regulate gene fusion differential expression between MCF-7 ERβ+/ERα– and ERβ–/ERα+ BC cells. However, our findings revealed two typologies of chromosomal interactions as characterizing gene fusion patterns in ERβ+/ERα– and ERβ–/ERα+ MCF-7 BC cells and suggested an abundance of intra-chromosomal interactions supporting a loss of inter-chromosomal regulation in that human BC cells. Indeed, the loss of inter-chromosomal regulation in breast cancer has been shown in several studies [39] [40]. In addition, genomic analysis revealed the recurrence of chromosomes 17 and 20 in retrieved chromosomal interactions characterizing ERβ+ and ERβ– MCF-7 BC cells. This result is in agreement with several studies, which have mentioned the involvement of chromosomes 17 and 20 in monitoring breast cancer [41]-[43]. Kytölä *et al.* (2000) [42] study analyzing 15 BC cells including human MCF-7 BC exhibited chromosomes 8, 1, 17, 16 and 20 as significantly involved in the translocation process. However, our findings revealed several gene fusion events, *i.e.*, ARFGEF2--

SULF2, BCAS4--BCAS3, DEPDC1B--ELOVL7, ESR1--CCDC170, EIF3H--RAD21 commonly expressed in ERβ+/ERα– and ERβ–/ERα+ BC cell lines recognized as breast cancer biomarkers [44] [45]. Although SULF2--PRICKLE2 gene fusion event commonly expressed in MCF-7 ERβ+/ERα– and ERβ–/ERα+ BC models are not formally identified as a breast cancer biomarker, certain studies revealed its implication in characterizing the latter [46] [47], suggesting SULF2--PRICKLE2 as novel breast cancer specific gene fusion with great expression potential. In addition, analysis revealed AHCYL1--RAD51C and BCAS3--ATXN7 gene fusion expressions in MCF-7 ERβ–/ERα+ breast cancer model. Indeed, Kim *et al.* (2019) [42] identified AHCYL1--RAD51C as a recurrent gene fusion event in breast cancer while AHCYL1--RAD51C gene fusion has not been clearly recognized as such, even if AHCYL1 and RAD51C genes involved in this fusion process were recognized as breast cancer biomarkers [48] [49]. HIPK1--DENND2C gene fusion specific to MCF-7 ERβ–/ERα+ BC model results to be a biomarker of several tumors [50], while ABCA5--PPP4R1L and RPS6KB1--DIAPH3 gene fusion specifically expressed in the same BC model (MCF-7 ERβ–/ERα+) are specific breast cancer molecular biomarkers [51]. Interestingly, genes contributing to gene fusion processes specific to MCF-7 ERβ–/ERα+ BC model, meanly result involved in biological function favoring breast cancer development [44] [45], as opposite to DENND2C gene involved in the partial gene fusion event characterizing MCF-7 ERβ+/ERα– BC model. Biegin-Rolett *et al.* (2020) [50] study revealed DENND2C gene as a potential tumor suppressor. These results clearly revealed functional differences in terms of gene fusion events between MCF-7 ERβ+/ERα– and MCF-7 ERβ–/ERα+ BC cell models. Overall, these results support the role of nuclear ERβ in monitoring and repressing breast cancer development, under early estradiol stimulation. Interestingly, 75% of gene fusion events discriminated in our study were recognized as breast cancer biomarkers. Genes contributing to those chromosomal fusion events are mainly involved in biological functions, *i.e.*, 1) cell proliferation, 2) cell growth, 3) angiogenesis, apoptosis evasion, cell signaling, nucleic acid regulation, carbohydrate, lipid and protein regulation, which result in being function contributing and monitoring several cancers and as well specifically breast cancer [52] [53].

5. Conclusion

Gene's fusions expressed in cancer result in the right feature in characterizing and as well as monitoring breast cancer. Our study highlighted nuclear ERβ involvement as a tumor suppressor in monitoring MCF-7 breast cancer progression by silencing some gene fusion expressions recognized as breast cancer biomarkers. In other words, although the analysis was focused on the MCF-7 BC models and that could constitute a limitation to the study, nuclear ERβ expression in hormone responsive breast cancer cells, in response to early estradiol stimulation by monitoring gene fusion processes exhibited an antagonist activity against breast cancer progression.

Author's Contribution

DDN proposed the topic and set up the experiment as well as the bioinformatics and biostatistics workflow. DDN performed and supervised bioinformatics and biostatistics analysis and wrote the paper. AJ and DO made a considerable contribution to the manuscript elaboration. DD, AJ, CSKH and DO participated in the bioinformatics and biostatistics analysis. AJ, DD, and DO contributed to preparing tables and figures and organizing the reference chapter. All authors participated in the conception and design of the study, as well as manuscript drafting and revision. All authors read and approved the final version of the paper.

Acknowledgements

Thank you to Professor Alessandro Weisz's laboratory and research team (Genomic and Molecular Medicine Laboratory) at the medicine and chirurgic department of the University of Salerno, Italy, for providing the transcriptomic sequences for the present study.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Mitelman, F., Johansson, B. and Mertens, F. (2007) The Impact of Translocations and Gene Fusions on Cancer Causation. *Nature Reviews Cancer*, **7**, 233-245. https://doi.org/10.1038/nrc2091

[2] Futreal, P.A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., *et al.* (2004) A Census of Human Cancer Genes. *Nature Reviews Cancer*, **4**, 177-183. https://doi.org/10.1038/nrc1299

[3] Geyer, F.C., Marchiò, C. and Reis-Filho, J.S. (2009) The Role of Molecular Analysis in Breast Cancer. *Pathology*, **41**, 77-88. https://doi.org/10.1080/00313020802563536

[4] Weigelt, B. and Reis-Filho, J.S. (2009) Histological and Molecular Types of Breast Cancer: Is There a Unifying Taxonomy? *Nature Reviews Clinical Oncology*, **6**, 718-730. https://doi.org/10.1038/nrclinonc.2009.166

[5] Stingl, J. and Caldas, C. (2007) Molecular Heterogeneity of Breast Carcinomas and the Cancer Stem Cell Hypothesis. *Nature Reviews Cancer*, **7**, 791-799. https://doi.org/10.1038/nrc2212

[6] Weigelt, B., Geyer, F.C. and Reis-Filho, J.S. (2010) Histological Types of Breast Cancer: How Special Are They? *Molecular Oncology*, **4**, 192-208. https://doi.org/10.1016/j.molonc.2010.04.004

[7] Dago, D.N., Scafoglio, C., Rinaldi, A., Memoli, D., Giurato, G., Nassa, G., *et al.* (2015) Estrogen Receptor Beta Impacts Hormone-Induced Alternative mRNA Splicing in Breast Cancer Cells. *BMC Genomics*, **16**, Article No. 367. https://doi.org/10.1186/s12864-015-1541-1

[8] Bauer, K.R., Brown, M., Cress, R.D., Parise, C.A. and Caggiano, V. (2007) Descriptive Analysis of Estrogen Receptor (ER)-Negative, Progesterone Receptor (PR)-Negative, and HER2-Negative Invasive Breast Cancer, the So-Called Triple-Negative Phenotype: A Population-Based Study from the California Cancer Registry. *Cancer*, **109**,

1721-1728. https://doi.org/10.1002/cncr.22618

[9] Fimereli, D., Fumagalli, D., Brown, D., Gacquer, D., Rothé, F., Salgado, R., *et al.* (2018) Genomic Hotspots but Few Recurrent Fusion Genes in Breast Cancer. *Genes, Chromosomes and Cancer*, **57**, 331-338. https://doi.org/10.1002/gcc.22533

[10] Noel, D.D., Nafan, D., Inza, J.F., Jean-Luc, A.M., Hermann-Desire, L., Didier, M.Y.S., *et al.* (2017) DEXseq and Cuffdiff Approaches Weighing Differential Spliced Genes Exons Modulation in Estrogen Receptor (ErÎ²) Breast Cancer Cells. *African Journal of Biotechnology*, **16**, 1404-1427. https://doi.org/10.5897/ajb2016.15860

[11] Cotrim, C.Z., Fabris, V., Doria, M.L., Lindberg, K., Gustafsson, J., Amado, F., *et al.* (2012) Estrogen Receptor β Growth-Inhibitory Effects Are Repressed through Activation of MAPK and PI3K Signalling in Mammary Epithelial and Breast Cancer Cells. *Oncogene*, **32**, 2390-2402. https://doi.org/10.1038/onc.2012.261

[12] Bado, I., Pham, E., Soibam, B., Nikolos, F., Gustafsson, J. and Thomas, C. (2018) Erβ Alters the Chemosensitivity of Luminal Breast Cancer Cells by Regulating P53 Function. *Oncotarget*, **9**, 22509-22522. https://doi.org/10.18632/oncotarget.25147

[13] Bado, I., Nikolos, F., Rajapaksa, G., Wu, W., Castaneda, J., Krishnamurthy, S., *et al.* (2017) Somatic Loss of Estrogen Receptor β and p53 Synergize to Induce Breast Tumorigenesis. *Breast Cancer Research*, **19**, Article No. 79. https://doi.org/10.1186/s13058-017-0872-z

[14] Hall, J.M., McDonnell, D.P. and Korach, K.S. (2002) Allosteric Regulation of Estrogen Receptor Structure, Function, and Coactivator Recruitment by Different Estrogen Response Elements. *Molecular Endocrinology*, **16**, 469-486. https://doi.org/10.1210/mend.16.3.0814

[15] Hawse, J.R., Carter, J.M., Aspros, K.G.M., Bruinsma, E.S., Koepplin, J.W., Negron, V., *et al.* (2019) Optimized Immunohistochemical Detection of Estrogen Receptor Beta Using Two Validated Monoclonal Antibodies Confirms Its Expression in Normal and Malignant Breast Tissues. *Breast Cancer Research and Treatment*, **179**, 241-249. https://doi.org/10.1007/s10549-019-05441-3

[16] Lattrich, C., Stegerer, A., Häring, J., Schüler, S., Ortmann, O. and Treeck, O. (2013) Estrogen Receptor β Agonists Affect Growth and Gene Expression of Human Breast Cancer Cell Lines. *Steroids*, **78**, 195-202. https://doi.org/10.1016/j.steroids.2012.10.014

[17] Punzi, S., Meliksetian, M., Riva, L., Marocchi, F., Pruneri, G., Criscitiello, C., *et al.* (2019) Development of Personalized Therapeutic Strategies by Targeting Actionable Vulnerabilities in Metastatic and Chemotherapy-Resistant Breast Cancer PDXs. *Cells*, **8**, Article 605. https://doi.org/10.3390/cells8060605

[18] Hsu, J. (1996) Multiple Comparisons: Theory and Methods. Chapman and Hall/CRC. https://doi.org/10.1201/b15074

[19] Stevens, J.P. (2013) Intermediate Statistics: A Modern Approach. Routledge.

[20] Dago, N., Saraka, M., Diarrassouba, N., Mori, A., Lallié, H., N'Goran, E., *et al.* (2017) RNA-Seq Evaluating Several Custom Microarrays Background Correction and Gene Expression Data Normalization Systems. *Biotechnology Journal International*, **19**, 1-14. https://doi.org/10.9734/bji/2017/36345

[21] Noel, D.D. (2021) Normality Assessment of Several Quantitative Data Transformation Procedures. *Biostatistics and Biometrics Open Access Journal*, **10**, 53-67. https://doi.org/10.19080/bboaj.2021.10.555786

[22] Treeck, O., Lattrich, C., Springwald, A. and Ortmann, O. (2009) Estrogen Receptor Beta Exerts Growth-Inhibitory Effects on Human Mammary Epithelial Cells. *Breast*

Cancer Research and Treatment, **120**, 557-565.
https://doi.org/10.1007/s10549-009-0413-2

[23] Ma, L., Liu, Y., Geng, C., Qi, X. and Jiang, J. (2013) Estrogen Receptor Β Inhibits Estradiol-Induced Proliferation and Migration of MCF-7 Cells through Regulation of Mitofusin 2. *International Journal of Oncology*, **42**, 1993-2000.
https://doi.org/10.3892/ijo.2013.1903

[24] Corchete, L.A., Rojas, E.A., Alonso-López, D., De Las Rivas, J., Gutiérrez, N.C. and Burguillo, F.J. (2020) Systematic Comparison and Assessment of RNA-Seq Procedures for Gene Expression Quantitative Analysis. *Scientific Reports*, **10**, Article No. 19737. https://doi.org/10.1038/s41598-020-76881-x

[25] Noel, D.D., Marinella, P., Mauro, G., Tripodi, S.I., Pin, A., Serena, A., *et al.* (2021) Genetic Variants Assessing Crohn's Disease Pattern in Pediatric Inflammatory Bowel Disease Patients by a Clinical Exome Survey. *Bioinformatics and Biology Insights*, **15**.
https://doi.org/10.1177/11779322211055285

[26] Edwards, P.A. and Howarth, K.D. (2012) Are Breast Cancers Driven by Fusion Genes? *Breast Cancer Research*, **14**, Article No. 303. https://doi.org/10.1186/bcr3122

[27] Fimereli, D., Fumagalli, D., Brown, D., Gacquer, D., Rothé, F., Salgado, R., *et al.* (2018) Genomic Hotspots but Few Recurrent Fusion Genes in Breast Cancer. *Genes, Chromosomes and Cancer*, **57**, 331-338. https://doi.org/10.1002/gcc.22533

[28] Veeraraghavan, J., Ma, J., Hu, Y. and Wang, X. (2016) Recurrent and Pathological Gene Fusions in Breast Cancer: Current Advances in Genomic Discovery and Clinical Implications. *Breast Cancer Research and Treatment*, **158**, 219-232.
https://doi.org/10.1007/s10549-016-3876-y

[29] Loman, N.J., Misra, R.V., Dallman, T.J., Constantinidou, C., Gharbia, S.E., Wain, J., *et al.* (2012) Performance Comparison of Benchtop High-Throughput Sequencing Platforms. *Nature Biotechnology*, **30**, 434-439. https://doi.org/10.1038/nbt.2198

[30] Dobin, A. and Gingeras, T.R. (2015) Mapping RNA-Seq Reads with Star. *Current Protocols in Bioinformatics*, **51**, 11.14.1-11.14.19.
https://doi.org/10.1002/0471250953.bi1114s51

[31] Haas, B. J., Dobin, A., Stransky, N., Li, B., Yang, X., Tickle, T., *et al.* (2017). STAR-Fusion: Fast and Accurate Fusion Transcript Detection from RNA-Seq. BioRxiv.
https://doi.org/10.1101/120295

[32] Eswaran, J., Cyanam, D., Mudvari, P., Reddy, S.D.N., Pakala, S.B., Nair, S.S., *et al.* (2012) Transcriptomic Landscape of Breast Cancers through mRNA Sequencing. *Scientific Reports*, **2**, Article No. 264. https://doi.org/10.1038/srep00264

[33] Noel, D.D., Alberto, F., Luciano, X., Antonio, M., Massimo, D. and Giovanni, M. (2016) Heterogeneity of Global Gene Expression Microarray Designs in Detecting Differentially Expressed Genes. *International Journal of Bioinformatics Research*, **7**, 349-357.

[34] Remoué, A., Conan-Charlet, V., Le Flahec, G., Lambros, L., Marcorelles, P. and Uguen, A. (2020) Un cancer mammaire présentant un réarrangement de NTRK1. *Annales de Pathologie*, **40**, 42-45. https://doi.org/10.1016/j.annpat.2019.05.009

[35] Rouault, A. (2013) Etude génomique des cancers du sein familiaux liés à une mutation constitutionnelle du gène BRCA2. Doctorat Thesis, Bordeaux 2 University.

[36] Saha Roy, S. and Vadlamudi, R.K. (2012) Role of Estrogen Receptor Signaling in Breast Cancer Metastasis. *International Journal of Breast Cancer*, **2012**, Article ID: 654698. https://doi.org/10.1155/2012/654698

[37] Yue, W., Wang, J., Li, Y., Fan, P., Liu, G., Zhang, N., *et al.* (2010) Effects of Estrogen

on Breast Cancer Development: Role of Estrogen Receptor Independent Mechanisms. *International Journal of Cancer*, **127**, 1748-1757. https://doi.org/10.1002/ijc.25207

[38] McDonnell, D.P. and Norris, J.D. (2002) Connections and Regulation of the Human Estrogen Receptor. *Science*, **296**, 1642-1644. https://doi.org/10.1126/science.1071884

[39] Anda-Jáuregui, G.D., Fresno, C., García-Cortés, D., Enríquez, J.E. and Hernández-Lemus, E. (2019) Intrachromosomal Regulation Decay in Breast Cancer. *Applied Mathematics and Nonlinear Sciences*, **4**, 223-230. https://doi.org/10.2478/amns.2019.1.00020

[40] Espinal-Enríquez, J., Fresno, C., Anda-Jáuregui, G. and Hernández-Lemus, E. (2017) RNA-Seq Based Genome-Wide Analysis Reveals Loss of Inter-Chromosomal Regulation in Breast Cancer. *Scientific Reports*, **7**, Article No. 1760. https://doi.org/10.1038/s41598-017-01314-1

[41] Gunn, S., Yeh, I., Lytvak, I., Tirtorahardjo, B., Dzidic, N., Zadeh, S., *et al.* (2010) Clinical Array-Based Karyotyping of Breast Cancer with Equivocal HER2 Status Resolves Gene Copy Number and Reveals Chromosome 17 Complexity. *BMC Cancer*, **10**, Article No. 396. https://doi.org/10.1186/1471-2407-10-396

[42] Kytölä, S., Rummukainen, J., Nordgren, A., Karhu, R., Farnebo, F., Isola, J., *et al.* (2000) Chromosomal Alterations in 15 Breast Cancer Cell Lines by Comparative Genomic Hybridization and Spectral Karyotyping. *Genes, Chromosomes and Cancer*, **28**, 308-317. https://doi.org/10.1002/1098-2264(200007)28:3<308::aid-gcc9>3.0.co;2-b

[43] Marchiò, C., Lambros, M.B., Gugliotta, P., Di Cantogno, L.V., Botta, C., Pasini, B., *et al.* (2009) Does Chromosome 17 Centromere Copy Number Predict Polysomy in Breast Cancer? A Fluorescence *in Situ* Hybridization and Microarray-Based CGH Analysis. *The Journal of Pathology*, **219**, 16-24. https://doi.org/10.1002/path.2574

[44] Kim, P., Jang, Y.E. and Lee, S. (2019) FusionScan: Accurate Prediction of Fusion Genes from RNA-Seq Data. *Genomics & Informatics*, **17**, e26. https://doi.org/10.5808/gi.2019.17.3.e26

[45] Fotakis, G., Rieder, D., Haider, M., Trajanoski, Z. and Finotello, F. (2019) NeoFuse: Predicting Fusion Neoantigens from RNA Sequencing Data. *Bioinformatics*, **36**, 2260-2261. https://doi.org/10.1093/bioinformatics/btz879

[46] Zhu, C., He, L., Zhou, X., Nie, X. and Gu, Y. (2015) Sulfatase 2 Promotes Breast Cancer Progression through Regulating Some Tumor-Related Factors. *Oncology Reports*, **35**, 1318-1328. https://doi.org/10.3892/or.2015.4525

[47] Chen, Z., Huang, J., Feng, Y., Li, Z. and Jiang, Y. (2021) Profiling of Specific Long Non-Coding RNA Signatures Identifies ST8SIA6-AS1 as a Novel Target for Breast Cancer. *The Journal of Gene Medicine*, **23**, e3286. https://doi.org/10.1002/jgm.3286

[48] Milne, R.L., Burwinkel, B., Michailidou, K., Arias-Perez, J.I., Zamora, M.P., Menéndez-Rodríguez, P., *et al.* (2014) Common Non-Synonymous SNPs Associated with Breast Cancer Susceptibility: Findings from the Breast Cancer Association Consortium. *Human Molecular Genetics*, **23**, 6096-6111.

[49] Zhou, Z., Qiu, R., Liu, W., Yang, T., Li, G., Huang, W., *et al.* (2021) BCAS3 Exhibits Oncogenic Properties by Promoting CRL4A-Mediated Ubiquitination of P53 in Breast Cancer. *Cell Proliferation*, **54**, e13088. https://doi.org/10.1111/cpr.13088

[50] Bieging-Rolett, K.T., Kaiser, A.M., Morgens, D.W., Boutelle, A.M., Seoane, J.A., Van Nostrand, E.L., *et al.* (2020) Zmat3 Is a Key Splicing Regulator in the p53 Tumor Suppression Program. *Molecular Cell*, **80**, 452-469.e9. https://doi.org/10.1016/j.molcel.2020.10.022

[51] Asmann, Y.W., Hossain, A., Necela, B.M., Middha, S., Kalari, K.R., Sun, Z., *et al.* (2011) A Novel Bioinformatics Pipeline for Identification and Characterization of Fusion Transcripts in Breast Cancer and Normal Cell Lines. *Nucleic Acids Research*, **39**, e100-e100. https://doi.org/10.1093/nar/gkr362

[52] Dupont, V. (2011) Signalisation par les récepteurs des œstrogènes: Mécanismes de reconnaissance de l'ADN et nouvelles approches pharmacologiques d'inhibition. Master's Thesis, Université de Montréal.

[53] Shi, X., Singh, S., Lin, E. and Li, H. (2021) Chimeric RNAs in Cancer. *Advances in Clinical Chemistry*, **100**, 1-35. https://doi.org/10.1016/bs.acc.2020.04.001

Supplement

Table S1. Raw genomic reads sequences statistics retrieved from sequencing processes.

Sample	Genomic reads sequences size (pb)	Total genomic reads sequences
HBC1_R1	51	45978764
HBC1_R2	51	45978764
HBC2_R1	51	26744336
HBC2_R2	51	26744336
HBC3_R1	51	44384131
HBC3_R2	51	44384131
MCF7noE1_R1	101	49446600
MCF7noE1_R2	101	49446600
MCF7noE2_R1	101	49938398
MCF7noE2_R2	101	49938398
MCF7noE3_R1	101	49058666
MCF7noE3_R2	101	49058666
MCF7E1_R1	101	48997313
MCF7E1_R2	101	48997313
MCF7E2_R1	101	49541889
MCF7E2_R2	101	49541889
MCF7E3_R1	101	48638898
MCF7E3_R2	101	48638898

Table S2. Genomic read sequences frequency after adaptor sequences trimming in MCF7 breast cancer cells samples for subjacent read alignment survey. R1 and R2 refer to pair end sequences from sequencing machine for each sample.

Breast Cancer Cell Line Models	Genomic read sequence with adaptor sequences	Genomic read sequences after adaptor sequences trimming	% of genomic read sequences selected for alignment
MCF7noE1 (R1, R2)	49446600.0	39387090.0	79.66
MCF7noE2 (R1, R2)	49938398.0	39486747.0	79.01
MCF7noE3 (R1, R2)	49058666.0	38995993.0	79.49
MCF7E1 (R1, R2)	48997313.0	44668816.0	91.17
MCF7E2 (R1, R2)	49541889.0	45062691.0	90.96
MCF7E3 (R1, R2)	48638898.0	44370387.0	91.22

Table S3. Functional analysis of genes involved in inter and/or intra-chromosomic fusion phenomena retrieved in ERβ+/ERα– (MCF7E) and ERβ–/ERα+ (MCF7noE) MCF-7 human BC cell models.

Head Gene	Related Functions	Associated Diseases	Tail Gene	Related Functions	Associated Diseases
AC099850.1	Unknown	No	VMP1	Regulation of the autophagy and zymophagy process; Lipid homeostasis; Membrane dynamics process; Formation of cell junctions.	Pancreatic.

Continued

ADAMTS19	Protein metabolism.	Discrete sub-aortic stenosis and subvalvular aortic stenosis.	SLC27A6	Translocation of long-chain fatty acids across the plasma membrane.	Lethal restrictive dermopathy; IV platelet glycoprotein deficiency.
AL121578.1	Unknown	No	SYTL5	Liaison site for Rab27A; Protein transport; Bladder traffic. Phospholipid binding.	Bladder cancer
ARFGEF2	Activation of ribosylating factors; Intracellular vesicular trafficking Golgi Transport Regulations Guanine-nucleotide exchange activity and brefeldin A inhibition; Chaperonin-mediated protein folding; Protein metabolism; Transferrin receptor recycling from endosome recycling to plasma membrane; Tumor necrosis activity	Periventricular heterotopia with microcephaly; Autosomal recessive and periventricular nodular heterotopia.	SULF2	Growth factor coreceptors and cytokines; Cell signaling;	Adenocarcinoma of the breast.
ATP1A1	Establishment and maintenance of electrochemical gradients of Na and K ions; Provides energy for active transportation.	Charcot-Marie-Tooth disease; Axonal; T type 2Dd; Hypomagnesemia; Seizures and Mental Retardation 2.	ZFP64	DNA and metal ion binding activity; Upregulation of cytokine production and transcription by RNA polymerase; Upregulation of mRNA splicing.	Autosomal recessive congenital cerebellar ataxia.
BCAS4	Amplified Sequence of Breast Carcinoma 4; Biomarker of breast cancer.	Breast cancer	BCAS3	Amplified Sequence of Breast Carcinoma 3; Cellular response to estrogen stimulus; Upregulation of catalytic activity and transcription by RNA polymerase II; Differentiation of ectoderms.	Hengel-Maroofian-Schols syndrome; Non-specific syndromic intellectual disability; Breast cancer.
DEPDC1B	Activation of GTPase activity; Cell migration; Upregulation of the Wnt signaling pathway.	Breast cancer	ELOVL7	Biosynthesis and metabolism of fatty acyl-CoA; Involvement in multiple biological processes as precursors of membrane lipids and lipid mediators.	No
EIF3H	Deubiquitinase activity; Involved in the downregulation of ubiquitin-dependent proteasomal catabolic process and translational initiation; Biomarker of prostate cancer. Translational Control and Peptide Chain Elongation. Cell proliferation; Involved in cell cycle, differentiation, and apoptosis.	Prostate cancer; Breast cancer; Prostate carcinoma; Trichorhinophalangeal syndrome, type II.	RAD21	Repair of DNA double-strand breaks; Chromatid cohesion during mitosis;	Cornelia De Lange syndrome 4 with or without median brain abnormalities; Mungan syndrome.

Continued

ESR1	Estrogen receptor alpha; Involve in breast, ovarian, endometrial and other cancers; Homodimerization and heterodimerization activity; Independent activation of the ESR1 and ESR2 development ligand; Regulation of the cell cycle of the G1/S transition.	Breast cancer; Ovarian cancer; Endometrial cancer; Cancers.	CCDC170	Organization and stabilization of Golgi-associated microtubules; Probably involved in the breast cancer process	Breast cancer, Diseases associated with estrogen resistance; Van Maldergem syndrome 2.
EVL	Involved in the downregulation of epithelial cell migration; Regulation of SLIT and ROBO expression; Cytoskeletal signaling.	gastrointestinal adenoma; Focal segmental glomerulosclerosis 3.	EML1	Assembly and organization of the microtubule cytoskeleton; Regulation of mitotic spindle orientation and cell division plane orientation; Normal proliferation of neuronal progenitor cells.	Band heterotopia; Usher syndrome.
GATAD2B	Transcriptional repressor of zinc finger protein; Regulation of tp53 activity by acetylation; Opening of the RNA polymerase I promoter.	Ghent syndrome; Neurodevelopmental disorder associated with Gatad2b.	NUP210L	Development of Sertoli cells; Spermatid development.	Severe congenital neutropenia 5.
HACD2	Biosynthesis and metabolism of fatty acyl-CoA; Precursors of membrane lipids and lipid mediators.	Spinocerebellar ataxia 34; Sarcocystosis.	MYLK	Nucleation of actin by the ARP-WASP complex; ERK Signaling.	Megacystis-microcolon-intestinal hypoperistalsis syndrome 1; Aortic aneurysm; Familial thoracic 7.
MTAP	Metabolism of sulphur-containing amino acids; Cytokine signaling in the immune system. Deficient in many Cancers.	Diaphyseal bone marrow stenosis with malignant fibrous histiocytoma; Bone sarcoma.	CDKN2B-AS1	Epigenetic silence of other genes in the same group;	Cardiovascular disease, and also; Several cancers, Intracranial aneurysm Type 2 diabetes; Alzheimer's disease; Glaucoma. Periodontitis; Endometriosis.
MYH9	Cytokinesis; Cellular motility; Maintenance of cell shape; Regulation of actin dynamics for the formation of phagocytic sections; Nucleation of actin by the ARP-WASP complex.	Macrothrombocytopenia; Granulocyte inclusions with or without nephritis or sensorineural hearing loss; Deafness, autosomal dominant 17	EIF3D	Role in the formation of the 40S initiation complex; Peptide chain elongation and Apoptotic pathways in synovial fibroblasts.	Squamous cell carcinoma of the colon.
POP1	TRNA processing; Treatment of pre-mRNA containing capped intron.	Anauxetic dysplasia 1 and 2.	MATN2	Formation of filamentous networks in the extracellular matrices of various tissues.	Pilocytic astrocytoma
PYROXD2	Oxidoreductase activity; Organization of mitochondria.	No	HPSE2	Remodeling of the extracellular matrix; Angiogenesis; Tumor progression; Metabolism of glycosaminoglycans.	Urofacial syndrome 1; Cakut.

Continued

RPS6KB1	Response to mTOR (mammalian target of rapamycin) signaling to promote protein synthesis; Cell growth and cell proliferation; Translational Control	Biomarker of cancers; Tuberous sclerosis; Lymphangioleiomyomatosis	VMP1	Regulation of the autophagy and zymophagy process; Lipid homeostasis; Membrane dynamics process; Formation of cell junctions.	Pancreatic.
RSBN1	Dioxygenase activity; Metal ion binding activity; Chromatin organization.	No	AP4B1-AS1	Antisense RNA 1	Spastic paraplegia 47; Autosomal recessive and congenital abnormality of the nervous system.
SMARCC1	Transcriptional activation; Repression of certain genes by chromatin remodeling; Alteration of DNA-nucleosome topology; Glucocorticoid regulation mediated by the MIC and BRCA1 pathway.	Congenital hydrocephalus; Rhabdoid cancer.	EMCN	Inhibition of cell-extracellular matrix interaction; Biomarker of gastric cancer.	Gastric cancer; Nephrotic syndrome; Type 10; Hypotrichosis-lymphedema-telangiectasia-kidney malformation syndrome.
SULF2	Growth factor coreceptors and cytokines; Cell signaling;	Adenocarcinoma of the breast.	PRICKLE2	Wnt signaling in hepatocellular carcinoma.	Progressive myoclonic epilepsy type 5; Ataxic sensory neuropathy; Dysarthria; Ophthalmoparesis; Benign epilepsy with centrotemporal peaks.
TANC2	Cytoskeletal transport of dense-core granules; Regulation of dendritic spine development; Regulation of dendritic morphogenesis of the spine; Embryonic development in utero.	Intellectual development disorder with autistic features and language delay, with or without seizures and non-specific syndromic intellectual disability.	CA4	Breathing; Calcification; Acid-base balance; Bone resorption; Formation of aqueous humor of cerebrospinal fluid, saliva, and stomach acid; Role in inherited renal abnormalities of bicarbonate transport.	Retinitis pigmentosa 17 and retinitis pigmentosa.
TPTE	Signaling pathways of endocrine or spermatogenic function of the testis; Metabolism pi; Glycerophospholipid biosynthesis.	No.	BAGE2	Cancer antigen; Code for tumor antigens.	Melanoma.
TTC6	Tetratricopeptide repeat protein 6.	No.	AL121790.1	Not recognized.	No.
TXLNG	Intracellular vesicle trafficking; ATF4-mediated transcription inhibition; Regulation of bone mass density; Cell Cycle Progression	Isolated familial hypoparathyroidism.	SYAP1	PI3K/Akt signaling; Cellular response to growth factor stimulus; Cellular response to peptide hormone stimulus. Adipocyte differentiation.	No.

Table S4. Functional analysis of genes involved in inter and/or intra-chromosomic interaction in ERβ–/ERα+ MCF-7 BC cell models (MCF7noE).

Head Gene	Related Functions	Associated Diseases	Tail Gene	Related Functions	Associated Diseases
ABCA5	ATP Link Cassette Transport Cholesterol Transport Cation/anion transport Membrane transport	Hypertrichosis; Lysosomal disease.	PPP4R1L	Protein de-phosphorylation; Regulation of serine/threonine-protein phosphatase 4.	No.
AHCYL1	Cellular regulation; CLEC7A (Dectin-1) signaling; Ca^{2+} mobilization mediated by FCERI.	Intestinal impaction Spinocerebellar ataxia 29.	RAD51C	Response to elevated platelet cytosolic Ca^{2+}; Homology-directed repair.	Breast and Ovarian Cancer Familial Anemia 3 and Fanconi Anemia Complementation Groups O
BCAS3	Amplified Sequence of Breast Carcinoma 3; Cellular response to estrogen stimulus; Upregulation of catalytic activity and transcription by RNA polymerase II; Differentiation of ectoderms.	Hengel-Maroofian-Schols syndrome; Non-specific syndromic intellectual disability.	ATXN7	Protein metabolism and chromatin organization.	Spinocerebellar Ataxia 7 and Retinal Degeneration
HIPK1	Regulation of TNF-mediated transcription and cellular apoptosis; Role as a corepressor for homeodomain transcription factors; Promotes angiogenesis and involved in erythroid differentiation; Implicated in the formation of squamous cell malignant tumors.	Malignant squamous cell tumors.	DENND2C	Regulation of catalytic activity; Guanine nucleotide exchange factor; Vesicle-mediated transport and traffic regulations.	Neuropathy; Hereditary sensory and autonomic diseases; Type V; Noonan syndrome 6.
RPS6KB1	Response to mTOR (mammalian target of rapamycin) signaling to promote protein synthesis; Cell growth and cell proliferation; Translational Control	Biomarker of cancers; Tuberous sclerosis; Lymphangioleiomyo matosis	DIAPH3	Actin remodeling; Regulation of cell movement and adhesion; Signaling of the G RAC1 protein in the cellular process and the RAC2 GTPase cycle.	Auditory neuropathy; Autosomal dominant hearing loss 1 and autosomal dominant nonsyndromic sensorineural hearing loss of the DFNA type.

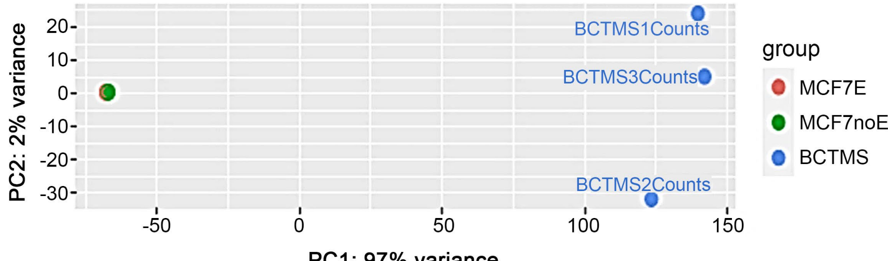

Figure S1. Principal Component Analysis (PCA) evaluating healthy breast cells, ERβ+ (MCF7E), ERβ– (MCF7noE) BC cell models biological replicate samples. BCTM acronym is for healthy breast cells (HBC).