

RESEARCH ARTICLE

Novel estrogen-responsive genes (ERGs) for the evaluation of estrogenic activity

Kentaro Nishi, Wenqiang Fu, Ryoiti Kiyama *

Department of Life Science, Faculty of Life Science, Kyushu Sangyo University Matsukadai, Higashi-ku, Fukuoka, Japan

* kiyama.r@ip.kyusan-u.ac.jp

Abstract

Estrogen action is mediated by various genes, including estrogen-responsive genes (ERGs). ERGs have been used as reporter-genes and markers for gene expression. Gene expression profiling using a set of ERGs has been used to examine statistically reliable transcriptomic assays such as DNA microarray assays and RNA sequencing (RNA-seq). However, the quality of ERGs has not been extensively examined. Here, we obtained a set of 300 ERGs that were newly identified by six sets of RNA-seq data from estrogen-treated and control human breast cancer MCF-7 cells. The ERGs exhibited statistical stability, which was based on the coefficient of variation (CV) analysis, correlation analysis, and examination of the functional association with estrogen action using database searches. A set of the top 30 genes based on CV ranking were further evaluated quantitatively by RT-PCR and qualitatively by a functional analysis using the GO and KEGG databases and by a mechanistic analysis to classify ER α / β -dependent or ER-independent types of transcriptional regulation. The 30 ERGs were characterized according to (1) the enzymes, such as metabolic enzymes, proteases, and protein kinases, (2) the genes with specific cell functions, such as cell-signaling mediators, tumor-suppressors, and the roles in breast cancer, (3) the association with transcriptional regulation, and (4) estrogen-responsiveness. Therefore, the ERGs identified here represent various cell functions and cell signaling pathways, including estrogen signaling, and thus, may be useful to evaluate estrogenic activity.

 OPEN ACCESS

Citation: Nishi K, Fu W, Kiyama R (2022) Novel estrogen-responsive genes (ERGs) for the evaluation of estrogenic activity. PLoS ONE 17(8): e0273164. <https://doi.org/10.1371/journal.pone.0273164>

Editor: Alessandro Weisz, Universita degli Studi di Salerno, ITALY

Received: June 10, 2022

Accepted: August 3, 2022

Published: August 17, 2022

Copyright: © 2022 Nishi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The RNA-seq data are available in the Gene Expression Omnibus database (www.ncbi.nlm.nih.gov/geo/) with Accession No. GSE205784.

Funding: This study was supported partly by a grant from Kyushu Sangyo University for promoting basic technologies, and a Grant-in-Aid for Basic Areas (Kakenhi 20H04343) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Competing interests: The authors have declared that there are no competing interests.

Introduction

Estrogenic chemicals are found abundantly in natural and industrial materials, and they exhibit a variety of cellular and physiological activities [1]. Along with the increasing interest in the estrogenic activity of these chemicals and their beneficial applications, there is growing need to develop new technologies for their detection and characterization. Thus, a variety of assays have been developed, such as ligand-binding assays, reporter-gene assays, yeast two-hybrid assays, transcription assays, protein assays, cell assays, and animal tests, which are based on the mechanisms of estrogen action at the levels of molecules, cells, tissues, and the whole body [2]. Among the assays, gene expression profiling, such as the DNA microarray assay, is a technology used to detect the alterations of gene expression by monitoring the

amount of mRNA or proteins of the estrogen-responsive genes (ERGs), which can detect the estrogenic activity of chemicals and mixtures of chemicals [3–5].

Searching for useful genes was challenging when the technological development was in its infancy. Recently, the development of DNA microarray technology for comprehensive searches of genes that were newly identified by the human genome project accelerated the search for ERGs, where tens to sometimes over a thousand ERGs were identified from more than ten thousand human genes and expressed sequence tags [6–11]. The identified ERGs have been examined for use in various fields such as food, clinical, pharmacological, and environmental applications [1]. For example, better understanding of the gene-regulation networks based on the ERGs is critical for the development of therapeutics for breast cancer [5]. Likewise, gene expression profiling based on ERGs has been used for environmental studies [12].

While the number of ERGs used for gene-expression profiling is important to reliably evaluate the estrogenic activity, the differences in the quality of ERGs due to various factors, such as their amounts in cells, the degrees of their responses, and their different cell functions, can cause unpredictable errors in the assay results. Other factors also need to be considered such as the cost, laboriousness, and the requirements of devices and equipment. The proportion of genes quantitatively showing differences in expression under different stimulations is roughly a few to a few tens of % of all human genes [13]. When considering statistical stability, the number of the genes that can be used to predict specific results, like clinical outcomes, should be no more than 1,000 and preferably less than 100 [13]. For example, MammaPrint, a US Food and Drug Administration-cleared molecular diagnostic test used for predicting the risk of breast cancer recurrence, contained 70 genes [14]. On the other hand, gene expression profiling has been used to predict estrogen activity, where a number of chemicals were examined by DNA microarray-based assays (reviewed by Kiyama & Zhu, 2014 [12]). A total of 120 ERGs, consisting of six functional groups of genes, were used to evaluate estrogenic activity, with comparable reliability to other estrogen assays [15]. However, methods for reliably identifying ERGs with statistical stability within satisfactory limits have not been reported, even though more information about genes and their functions has been obtained. Here, we determined the ERGs that have statistical stability based on RNA sequencing (RNA-seq) analysis and examined their usefulness by focusing on their reliability and applicability to predict the estrogen action of chemicals.

The recent progress in RNA-seq has enabled its wider use in applications such as food quality control, environmental materials, and medicine [16–20]. For estrogen actions, RNA-seq data obtained from estrogen-stimulated MCF-7 cells were compared with reverse transcription (RT)-PCR and DNA microarray data, suggesting mutually consistent or respectively complementary gene expression profiles among them [21]. Furthermore, functional annotations of RNA-seq data with databases, such as the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases, have been used to identify targets of antitumor agents against breast cancer such as diallyl trisulfide (a garlic metabolite) [22], 6-thioguanine [23], shikonin [24], tamoxifen [25], and chickpea isoflavones [26]. Based on RNA-seq analysis, novel mechanisms of cancer progression and metastasis involving non-coding RNAs [27, 28], enhancer RNAs [29], and circular RNAs [30, 31] were demonstrated. Functional regulatory networks examined by comparing chromatin immunoprecipitation-sequencing (ChIP-seq) data and time-course RNA-seq data revealed that transcription factor MYC plays an important role in connecting the networks between promoters and enhancers [32].

Here, we obtained a set of novel ERGs by RNA-seq that can be used for gene expression profiling of potential estrogenic chemicals and demonstrated their usefulness by examining their statistical reliability and their functional relevance.

Materials and methods

Materials

Human breast cancer MCF-7 cells were obtained from the Japanese Collection of Research Bioresources Cell Bank. Antibodies used for Western blotting were those against total Erk1/2 (T-Erk; #9102, Cell Signaling Technology, Ipswich, MA, USA), phospho-Erk1/2 (P-Erk; #9101), total Akt (T-Akt; #4691), and phospho-Akt (P-Akt; #4060).

Sulforhodamine B (SRB) assay

MCF-7 cells were cultured in a phenol red-free RPMI 1640 medium (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 10% fetal bovine serum (FBS) (Gibco, Thermo Fisher Scientific) and maintained at 37°C with 5% CO₂ in an incubator (Thermo Fisher Scientific). The sulforhodamine B (SRB) assay was performed as described by Dong et al. [33, 34]. Before stimulation of MCF-7 cells with chemicals, the cells were cultured at a density of 1.5×10⁴ cells per well in RPMI 1640 containing 10% (v/v) dextran-coated charcoal-treated FBS (DCC-FBS) (Gibco, Thermo Fisher Scientific) for 3 days in 24-well plates. After treatment with 10 nM 17β-estradiol (E₂; Sigma-Aldrich, St. Louis, MO, USA), 1 μM ICI 182,780 (ICI; Sigma-Aldrich), or 0.1% dimethyl sulfoxide (vehicle) (DMSO; FUJIFILM Wako Pure Chemical, Osaka, Japan) for 3 more days, MCF-7 cells were fixed with trichloroacetic acid (Sigma-Aldrich) at 4°C for 30 min and washed with ultrapure water. After washing, samples were stained with acetic acid containing 0.4% SRB (Sigma-Aldrich) at room temperature for 20 min and then washed with acetic acid. The bound protein was dissolved with 10 mM unbuffered Tris-base (pH = 10.5) at room temperature for 10 min and transferred into 96-well plates to measure the absorbance at 490 nm using the Modular-Designed Multimode Reader SH-9000 (Corona Electric, Ibaraki, Japan). Three independent assays were performed for each chemical, and the data were analyzed by *t*-test.

Western blotting

Western blotting was performed as described by Dong et al. [33, 34]. MCF-7 cells at a density of 1.0×10⁵ cell per well were cultured for 2 days in DCC-FBS in 6-well plates and then one more day in serum-free medium. The cells were pretreated with ICI for 1 h and then treated with 10 nM E₂ or vehicle (0.1% v/v DMSO). The total protein was extracted from the cells and examined by SDS-PAGE using e-PAGEL (ATTO, Tokyo, Japan) and electro-transferred onto nitrocellulose membranes (Millipore, Billerica, MA, USA) using a semi-dry transfer cell (Bio-Rad Laboratories, Benicia, CA, USA). The membranes were blocked with EzBlock BSA (ATTO) and incubated with the antibodies against signal proteins (T-Erk, P-Erk, T-Akt or P-Akt) overnight at 4°C after appropriate dilution (1:500 or 1:1000). The antibody-antigen complexes were detected with a horseradish peroxidase-coupled goat antibody against rabbit IgG (Cell Signaling Technology) after dilution (1:1000) and visualized using WSE-6100 Lumi-Graph I (ATTO).

RNA-seq analysis

MCF-7 cells at a density of 1.0×10⁶ cells per well were cultured in DCC-FBS for 3 days at 37°C in 5% CO₂. The cells were treated with 10 nM E₂, 1 μM ICI, or 0.1% DMSO (vehicle) for 2 more days. Total RNA was extracted using a RNeasy Mini-kit (QIAGEN, Venlo, the Netherlands). The DNA libraries for sequencing were constructed using the Truseq Stranded mRNA Library Prep (Illumina, San Diego, CA) and MGIEasy Universal Library Conversion Kit (MGI Tech, Shenzhen, PRC). Then, the libraries were subjected to sequencing using the

DNBSEQ-G400 (MGISEQ-2000RS, MGI Tech) instrument, and the obtained reads were aligned using splitBarcode (<https://github.com/MGI-tech-bioinformatics/splitBarcode>). The read data were downsized to 20 M read pairs using seqkit v0.13.0 [35]. After the adapter sequences were cleaned with cutadapt ver.2.10 [36], the read data were mapped with HISAT v2.2.1 [37]. The value of FPKMs (reads per kilobase of transcript per million reads mapped) for each gene was estimated with Cuffdiff v2.2.1 [38], and the gene expression was analyzed as described previously [15]. After RNA-seq was repeated 6 times for E₂, the standard deviation (SD) and the average (Av) for the 203 genes, which includes ERGs and control genes [15], were determined and used for calculating the log₂ of FPKMs. These genes were evaluated based on the absolute value of SD/Av and used in the correlation analysis with the correlation coefficient (*R* value) between one of the assays (E₂-1) and each of the other five assays (E₂-2~E₂-6).

Real-time RT-PCR

The FPKM values obtained with RNA-seq were log₂-transformed and a total of 300 genes were selected based on the log₂ value (gene expression ratio). Among the 300 genes, 30 genes were selected by ranking the SD/Av values and further analyzed by *real-time* RT-PCR [8]. *Real-time* RT-PCR was conducted with the CFX Connect *real-time* PCR detection system using the iTaq Universal SYBR Green One-Step Kit (Bio-Rad). The conditions for *real-time* RT-PCR were as follows: reverse transcription for cDNA at 42°C for 10 min and denaturation at 95°C for 1 min, followed by 46 cycles of denaturation at 94°C for 10 sec, annealing at 57°C for 30 sec, and extension at 72°C for 20 sec. After PCR, a melting curve was constructed by increasing the temperature from 65 to 95°C to confirm the validity of the reaction. *Real-time* RT-PCR was repeated 3 times for each gene and the Av and SD were calculated using CFX Maestro Software (Bio-Rad). The nucleotide sequences of PCR primers (S1 Table) were obtained from the references for *EGR3* [8], *LOXL2* [39], and *SYNE1* [40], or otherwise by web-based calculation.

Data analysis for gene expression

The FPKM values obtained by RNA-seq were log₂-transformed and *p*-values were calculated. The genes showing statistically significant (*p* < 0.05) up-regulation or down-regulation, or those without significant changes, were visualized by Volcano plotting [26], where log₂-transformed expressional changes (*X*-axis) and -log₁₀-transformed *p*-values or SD/Av values (*Y*-axis) are indicated. The pathway analysis was performed with WebGestalt software (WEB-based Gene Set Analysis Toolkit) (www.webgestalt.org/; [41]) using the gene sets showing statistical significance (*p* < 0.05) from the GO (www.geneontology.org/) and KEGG (www.genome.jp/kegg/) databases. False discovery rates (FDRs: *Q*-values) were estimated using the Benjamini-Hochberg method, and the top ten categories of FDR < 0.05 were visualized.

Results

Gene expression profiling based on RNA-seq

Gene expression profiling is a method to evaluate the estrogenic activity of chemicals, and a variety of assays have been developed based on transcriptomic and proteomic technologies [1]. Here, we screened ERGs for evaluating estrogenic activity by transcriptomic assays such as RNA-seq and RT-PCR. To obtain a set of ERGs, we adopted a two-step process involving a correlation analysis and a coefficient of variation (CV) analysis (Fig 1A). MCF-7 cells were used as a cell system for the evaluation, which was performed first by two assays, a cell proliferation assay (SRB assay; Fig 1B) and protein assay (Western blotting to detect the

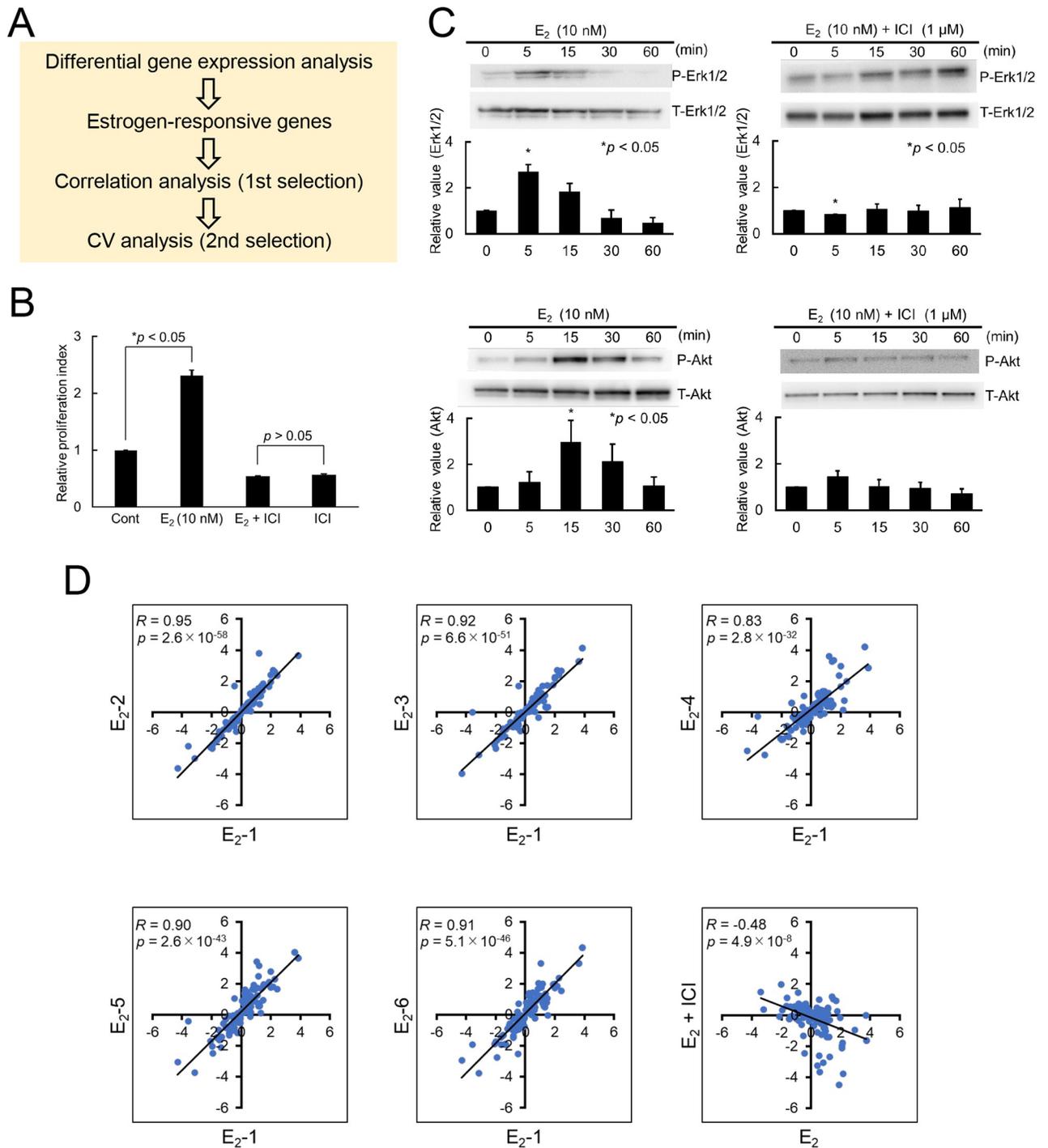


Fig 1. Evaluation of estrogenic activity by gene expression profiling. (A) A strategy adopted here. (B) Cell proliferation assay. MCF-7 cells treated with E₂ (10 nM), E₂ (10 nM) + ICI (1 μM), or vehicle (0.1% DMSO) were subjected to cell proliferation assays with SRB. The relative proliferation index for MCF-7 cells treated with E₂ (10 nM), E₂ (10 nM) + ICI (1 μM), or ICI (1 μM) alone are shown: **p* < 0.05, vs. control (0.1% DMSO). (C) Western blotting. MCF-7 cells were treated with E₂ (10 nM), E₂ (10 nM) + ICI (1 μM), or vehicle (0.1% DMSO) for indicated times and cell extracts were subjected to Western blotting for the evaluation of phosphorylated proteins (P-Erk1/2 and P-Akt) (upper images) and total proteins (T-Erk1/2 and T-Akt) (lower images). (D) Correlation analysis of RNA-seq data. The gene expression profiles for MCF-7 cells treated with E₂ obtained by RNA-seq were compared using a set of the 120 ERGs, and the results are visualized in scatter-plot graphs. The vertical and horizontal axes indicate log₂-values of FPKMs. *R*- and *p*-values were calculated for each graph based on linear regression between two profiles.

<https://doi.org/10.1371/journal.pone.0273164.g001>

phosphorylation of two marker proteins, Erk1/2 and Akt; Fig 1C), which have frequently been used to evaluate estrogenic activity [42–44]. The cells showed the response to estrogen (17 β -estradiol; E₂), and the response was inhibited by treatment with an ER antagonist ICI in both assays in the same way as previously reported [15], suggesting that the cells responded to E₂ through the ER. Then, we performed the RNA-seq analysis to evaluate the cell response at the transcription level (Fig 1D). A total of six sets of gene expression profiles for MCF-7 cells treated with E₂ were obtained by RNA-seq, where a set of the 120 ERGs, which were previously used in DNA microarray assays [15], were used to compare the profiles (first five panels in Fig 1D). The six combinations of profiles showed *R*-values of 0.83 to 0.95, indicating strong and statistically significant correlations among the profiles (see the data for all combinations in S2 Table). In contrast, there was no correlation between the profiles for the E₂ treatment and E₂+ICI treatment (the last panel in Fig 1D), suggesting the response to be mediated by the ER. Therefore, the ERGs that we obtained using gene expression profiling methods and by RNA-seq techniques can be used as markers for the evaluation of estrogenic chemicals.

Evaluation of RNA-seq-based gene expression profiling

Next, we examined the quality of the obtained ERGs by CV analysis (Fig 2). Among the six sets of ERGs obtained by RNA-seq (Fig 1D), the sets of 203 ERGs that were used to evaluate estrogenic activity by DNA microarray assays [15] were again examined by CV analysis (indicated by the bar graph in Fig 2), where the 203 genes included expression standards and 174 original ERGs, and the sets of 150, 120, 90, 60, or 30 genes were selected from the top of the list of genes ranked by CV values (mean of SD/Av). The top-ranked ERGs exhibited low CV values,

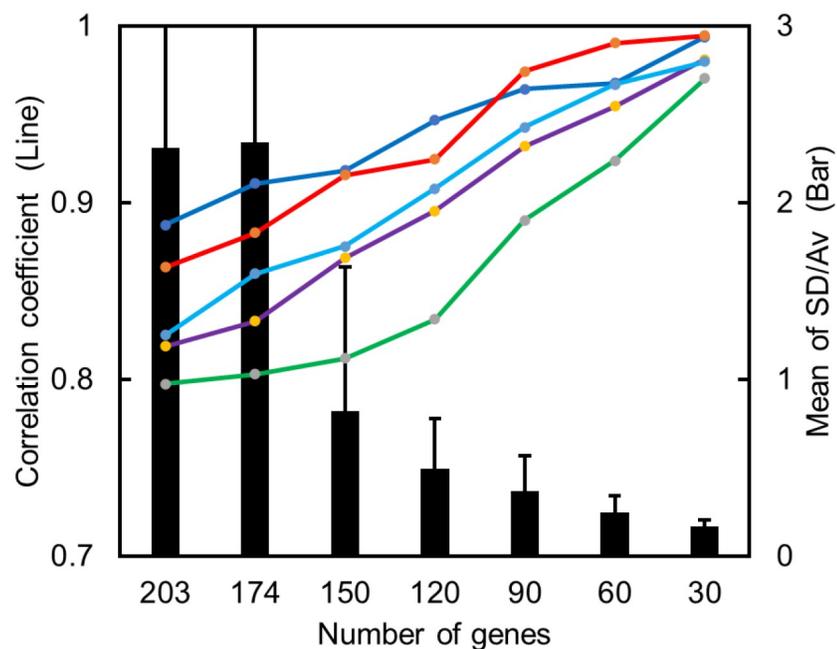


Fig 2. Statistical evaluation of ERGs. The RNA-seq for E₂ was repeated 6 times and gene expression profiles were obtained (E₂-1 to E₂-6). The correlation coefficients (*R*-values) between one profile (E₂-1) and the others (E₂-2 to E₂-6), giving five sets of data, were calculated for respective numbers of gene sets. The *R*-values of the total 203 genes, or the indicated numbers (174, 150, 120, 90, 60, or 30) of genes taken from the top of the list in the order of the genes ranked by the values of SD/Av indicating the statistical stability, are shown for the five sets of data in a line graph. The mean of SD/Av for respective numbers of genes are shown in a bar graph; SD/Av for 203 genes is 2.31 ± 6.24 and that for 174 genes is 2.34 ± 6.6.

<https://doi.org/10.1371/journal.pone.0273164.g002>

0.17 +/- 0.04 for the top 30 ERGs, indicating that gene expression profiling by RNA-seq can provide mean CV values that are lower than those by DNA microarray assays (mean CV value of 0.22; [15]). Furthermore, the correlation coefficients (*R*-values) for the top 30 ERGs in the five data sets obtained by RNA-seq were 0.97 to 0.99, which are also more reliable than those with nine data sets obtained by DNA microarray assays (0.93 to 0.97; [15]), confirming that the statistical stability of data is greater in RNA-seq than in DNA microarray assays.

Selection of new ERGs for gene expression profiling

We selected new ERGs for gene expression profiling based on the approach shown in the previous section. A total of 300 ERGs were selected from the ranks created by CV analysis, from all human genes available for RNA-seq (a total of approx. 26,000 genes), which consist of 150 up-regulated and 150 down-regulated genes (see S3 Table for the complete list). Among the 300 ERGs, the top 30 ERGs (Table 1) were examined by *real-time* RT-PCR (Fig 3A), and the data were compared with those obtained by RNA-seq (Fig 3B). Both datasets were identical, except for minor differences due to slight variation in the locations of the transcripts that were detected by these assays. Notably, the statistical significance levels are lower for some genes, probably due to the lack of expression levels in the statistical consideration.

Functional analysis of ERGs

The newly identified ERGs were examined by signaling pathway analysis using the GO and KEGG databases (Fig 4). The characteristics of ERGs were analyzed first by the statistical evaluation of the relationship of the significance (*p*-values; Fig 4A) or stability (CV values; Fig 4B) of expressional alterations with the degrees of the alterations (\log_2 -transformed fold changes) and then by the functional searches in the GO (Fig 4C and 4D) and KEGG (Fig 4E and 4F) databases. We confirmed that 30 ERGs belong to the groups of genes showing statistically significant alterations of their expression due to the treatment with estrogen (red or green circles in Fig 4A and 4B). From the database searches, significant levels of association were found between the genes of significant alterations of their expression and specific cell functions and

Table 1. List of 30 ERGs.

Rank	Gene symbol	CV	Rank	Gene symbol	CV
1	LINC02593	0.058	16	DOK7	0.125
2	PDLIM3	0.083	17	CYP1A1	0.125
3	RAP1GAP	0.090	18	LOXL2	0.126
4	ACOX2	0.091	19	RAPGEFL1	0.126
5	B4GALT1	0.094	20	PKIB	0.126
6	TMPRSS3	0.096	21	CCDC68	0.126
7	MATN2	0.098	22	GATA4	0.128
8	SUSD3	0.099	23	SPOCD1	0.129
9	FDFT1	0.099	24	CRISP3	0.134
10	ZNF521	0.099	25	IGSF1	0.134
11	FRY	0.104	26	RAB26	0.136
12	BARX2	0.107	27	EGR3	0.139
13	IL20	0.108	28	SRGAP3	0.139
14	CTSD	0.118	29	SYNE1	0.143
15	CSTA	0.124	30	INSYNI	0.145

The top 30 ERGs among the 300 genes stably responding to E₂ are listed.

<https://doi.org/10.1371/journal.pone.0273164.t001>

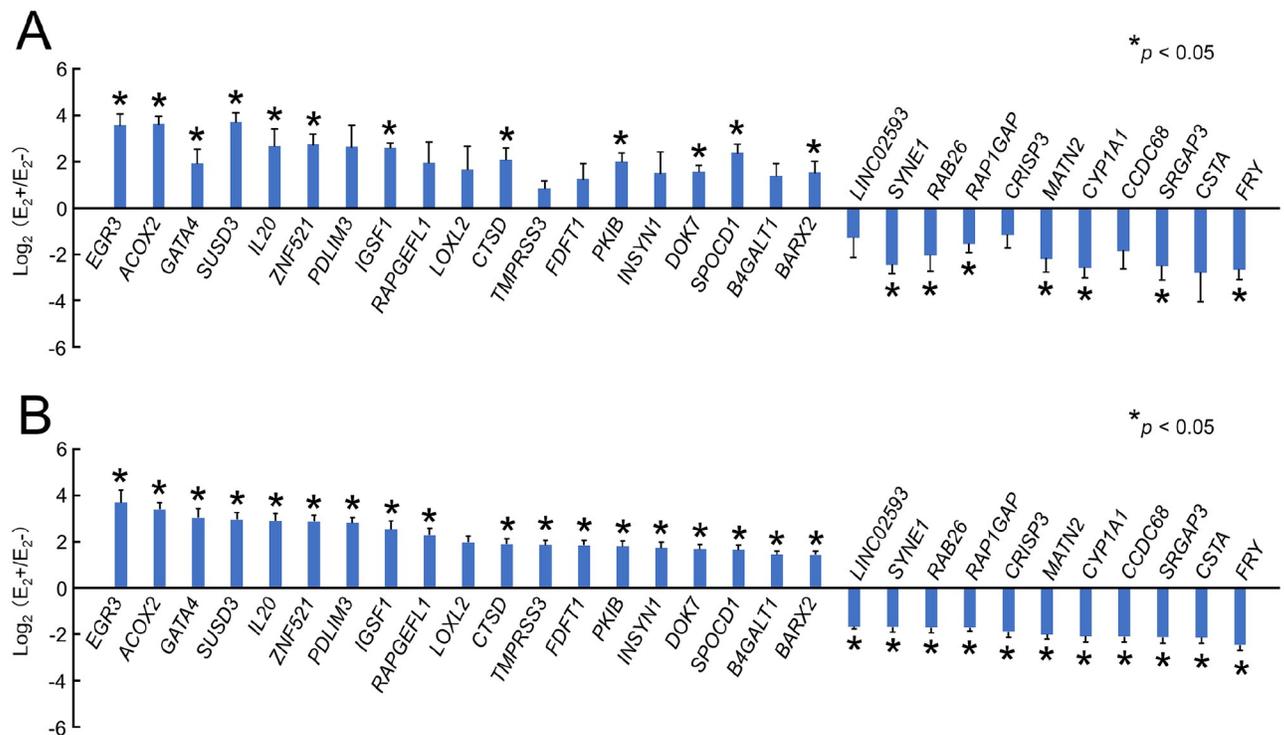


Fig 3. Transcriptional analysis of ERGs. (A) *Real-time* RT-PCR analysis for 30 ERGs. The 30 ERGs (listed in Table 1) that were selected by RNA-seq analysis in Fig 2 were analyzed again by *real-time* RT-PCR. The expression level of each gene was normalized by that of β -actin and indicated in a bar graph, where the bars show the mean \pm SD ($n = 3$) of \log_2 -transformed ratios of E_2+ data and E_2- data (E_2+/E_2-). (B) RNA-seq analysis for 30 ERGs. The values of FPKMs ($n = 6$) obtained by RNA-seq were \log_2 -transformed and indicated in a bar graph, where the bars show the mean \pm SD ($n = 6$) of the ratios, as shown in panel A. The top 30 genes selected according to SD/Av and expression ratios have stable gene expression. * $p < 0.05$: between E_2 and the control (0.1% DMSO).

<https://doi.org/10.1371/journal.pone.0273164.g003>

pathways such as the cellular macromolecule catabolic process (a 262-gene group showing up-regulation), protein localization to organelles (a 220-gene group showing up-regulation), mRNA metabolic process (a 201-gene group showing up-regulation), negative regulation of gene expression (a 290-gene group showing down-regulation), negative regulation of biosynthetic process (a 283-gene group showing down-regulation), negative regulation of cellular biosynthetic process (a 281-gene group showing down-regulation), and negative regulation of macromolecule biosynthetic process (a 276-gene group showing down-regulation) in the GO database (Fig 4C and 4D; S4 Table), and metabolic pathways (a 279-gene group showing up-regulation), protein processing in endoplasmic reticulum (a 58-gene group showing up-regulation), cell cycle (a 55-gene group showing up-regulation), pathways in cancer (a 90-gene group showing down-regulation), human papillomavirus infection (a 66-gene group showing down-regulation), endocytosis (a 46-gene group showing down-regulation), and breast cancer (a 32-gene group showing down-regulation) in the KEGG database (Fig 4E and 4F; S4 Table). Furthermore, significant numbers of the 30 and 300 ERGs that were characterized in Fig 2 were included in the groups of genes found by the database searches (S4 Table), suggesting the association of these ERGs with specific cell functions.

ER subtype-specific interaction of ERGs

Since ERGs are likely controlled by ERs through the interaction at their promoter and enhancer regions, we examined the ChIP-seq data that are available in the Gene Expression

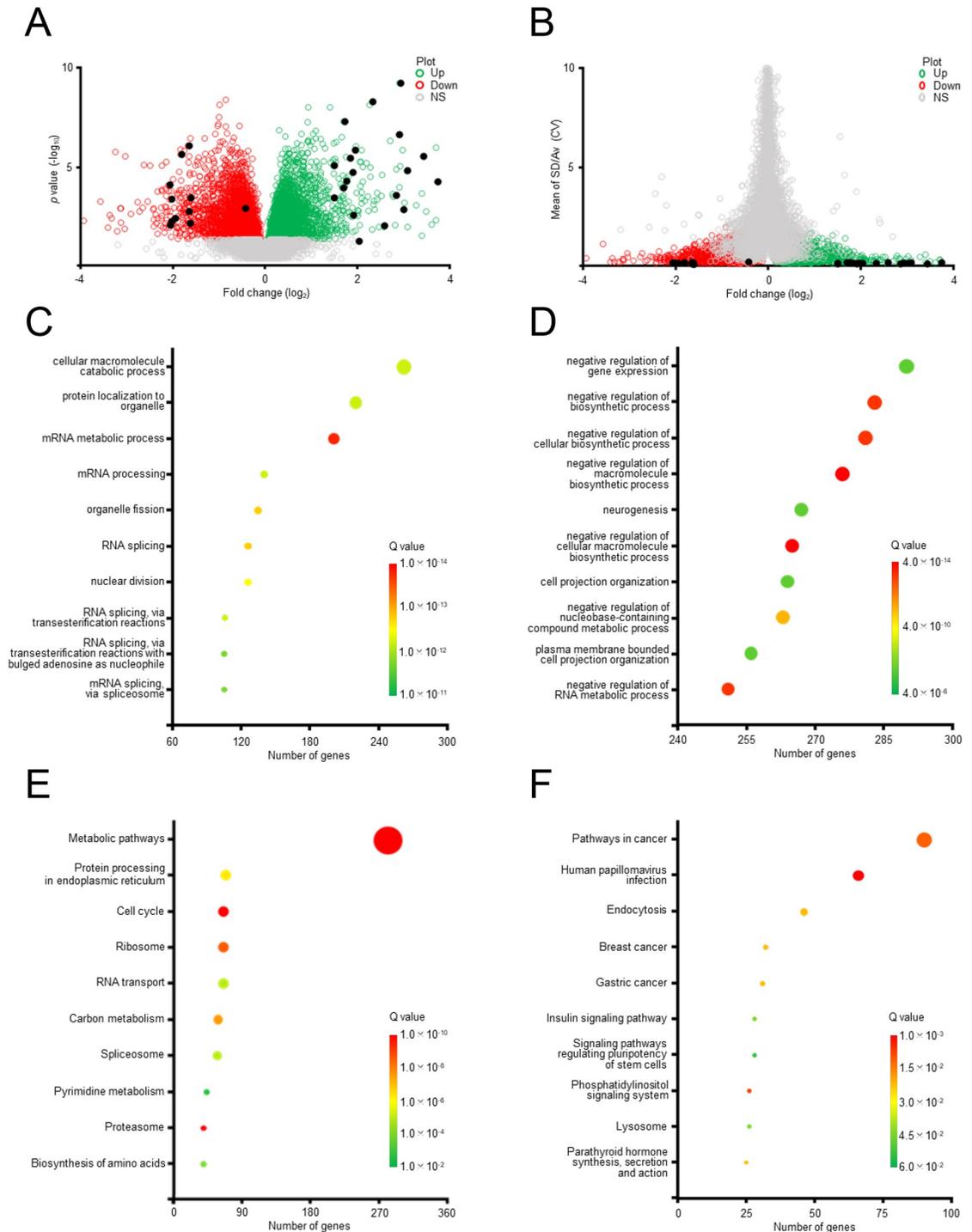


Fig 4. Signaling pathways associated with ERGs. (A) Volcano plots for ERGs in MCF-7 cells. The genes showing significant ($p < 0.05$) up-regulation (red circles) or down-regulation (green circles) are indicated, while those that are not significant are shown by gray circles. The novel 30 ERGs are shown by black dots. The graph shows \log_2 -transformed fold-changes (X-axis) and $-\log_{10}$ of p -values (Y-axis). (B) Volcano plots for the 30 genes that show stable expression in response to E_2 . The graph shows \log_2 -transformed fold-changes (X-axis) and mean of SD/Av values (Y-axis). (C, D) GO analysis for the biological processes (BPs). The top 10 significantly affected ($Q < 0.05$) BPs for the ERGs that exhibit up-regulation (panel C) or down-regulation (panel D) are shown. The number of genes in each group is visualized by the size of the circles, and the Q-values of the group by color. (E, F) KEGG analysis for metabolic pathways. The top 10 significantly affected ($Q < 0.05$) metabolic pathways for the ERGs that exhibit up-regulation (panel C) or down-regulation (panel D) are shown.

<https://doi.org/10.1371/journal.pone.0273164.g004>

Table 2. ER subtype-specific interaction of ERGs.

Gene No.	Gene symbol	ER α ^a		ER α binding ^c	ER β ^a	ER β binding ^c	Receptor subtype
		E ₂ - Q-value ^b (n = 3)	E ₂ + Q-value ^b (n = 3)				
1	<i>EGR3</i>	0	279	+	1.1	+	ER α / β
2	<i>ACOX2</i>	0	0	-	1.4	+	ER β
3	<i>GATA4</i>	133	2133	+	1.2	+	ER α / β
4	<i>SUSD3</i>	449	3026	+	0.7	-	ER α
5	<i>IL20</i>	1307	3091	+	1.1	+	ER α / β
6	<i>ZNF521</i>	0	0	-	0.9	-	-
7	<i>PDLIM3</i>	0	72	+	1.1	+	ER α / β
8	<i>IGSF1</i>	0	0	-	1.0	-	-
9	<i>RAPGEFL1</i>	764	4529	+	0.0	-	ER α
10	<i>LOXL2</i>	0	0	-	8.2	+	ER β
11	<i>CTSD</i>	234	1683	+	4.4	+	ER α / β
12	<i>TMPRSS3</i>	263	2761	+	3.2	+	ER α / β
13	<i>FDFT1</i>	1121	3753	+	1.2	+	ER α / β
14	<i>PKIB</i>	304	1887	+	0.8	-	ER α
15	<i>INSYNI</i>	200	3575	+	1.0	-	ER α
16	<i>DOK7</i>	676	2677	+	32.7	+	ER α / β
17	<i>SPOCD1</i>	0	31	+	1.0	-	ER α
18	<i>B4GALT1</i>	145	1105	+	18.9	+	ER α / β
19	<i>BARX2</i>	89	465	+	1.0	-	ER α
21	<i>SYNE1</i>	0	395	+	1.0	-	ER α
22	<i>RAB26</i>	118	991	+	4.8	+	ER α / β
23	<i>RAP1GAP</i>	45	133	+	2.1	+	ER α / β
24	<i>CRISP3</i>	0	0	-	0.8	-	-
25	<i>MATN2</i>	0	200	+	2.5	+	ER α / β
26	<i>CYP11A1</i>	1203	2489	+	12.3	+	ER α / β
27	<i>CCDC68</i>	0	1490	+	1.2	+	ER α / β
28	<i>SRGAP3</i>	125	1290	+	0.9	-	ER α
29	<i>CSTA</i>	124	840	+	1.0	-	ER α
30	<i>FRY</i>	0	117	+	0.7	-	ER α

Gene No. is based on the order of the data from *real-time* RT-PCR (Fig 3). Gene No. 20 (*LINC02593*) is not listed here because it is a long non-coding RNA and inappropriate for the analysis.

^a The original data analyzed in this table are available in the NCBI's Gene Expression Omnibus with accession numbers GSE117569 (for ER α ; [135]) and GSE149979 (for ER β ; [151]).

^b The Q-value was determined using the ChIP-Atlas database (<https://chip-atlas.org/>; [152]), where each gene was analyzed for the presence of peaks within \pm 10 kb from the transcription start site (TSS).

^c ER α binding is positive (+) when the Q-value of E₂+ is bigger than that of E₂-.

^d The integrative genomics viewer (IGV) was used to visualize the results of ChIP-seq, where each gene was analyzed for the presence of peaks within \pm 10 kb from the TSS. Then, the S/N ratio was calculated, where the average of E₂+/ER β ChIP-seq values, rep 1 and 2, was used as S, and the E₂+/input value was used as N.

^e ER β binding is positive (+) when the S/N ratio is bigger than 1.0.

<https://doi.org/10.1371/journal.pone.0273164.t002>

Omnibus database to determine whether the regions containing these ERGs were identified as ER subtype-specific binding sites (Table 2). The results indicate that 29 ERGs (except *LINC02593*) were classified into four groups; ER α -specific (*SUSD3*, *RAPGEFL1*, *PKIB*, *INSYNI*, *SPOCD1*, *BARX2*, *SYNE1*, *SRGAP3*, *CSTA*, and *FRY*), ER β -specific (*ACOX2* and *LOXL2*), ER α / β -specific (*EGR3*, *GATA4*, *IL20*, *PDLIM3*, *CTSD*, *TMPRSS3*, *FDFT1*, *DOK7*,

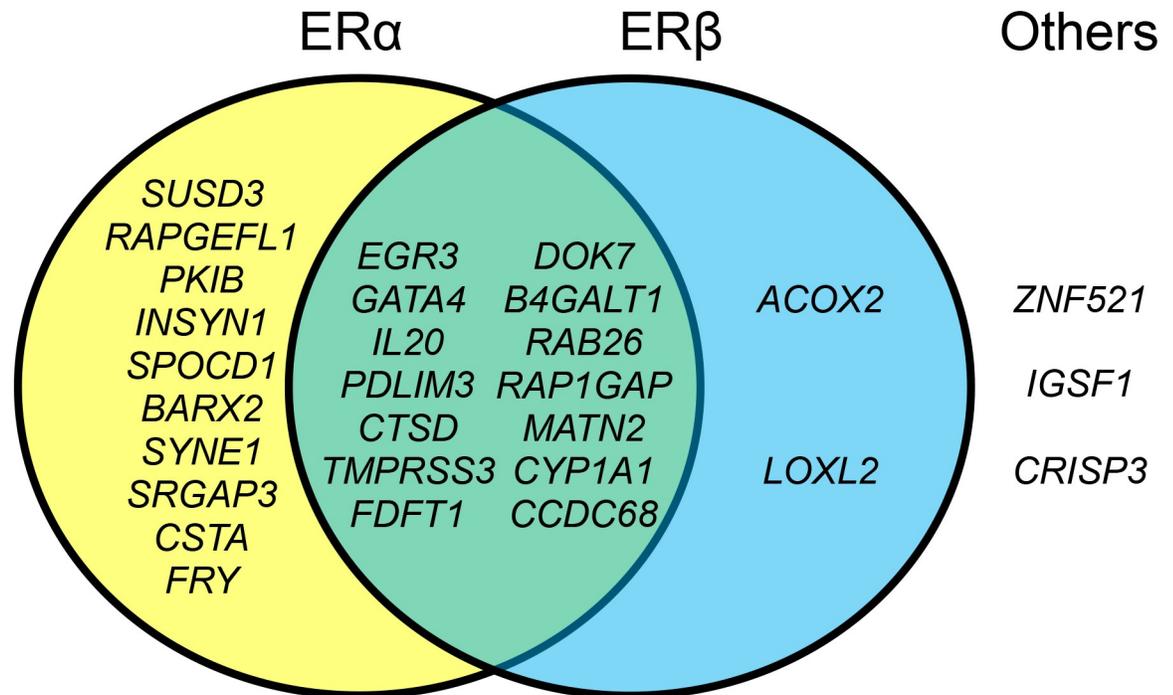


Fig 5. ER subtype-specific interaction of ERGs.

<https://doi.org/10.1371/journal.pone.0273164.g005>

B4GALT1, *RAB26*, *RAP1GAP*, *MATN2*, *CYP1A1*, and *CCDC68*), and other types (*ZNF521*, *IGSF1*, and *CRISP3*) (Fig 5).

Discussion

ERGs for the evaluation of estrogenic activity

Estrogen is a hormone that regulates the female reproductive system and the physiological, neurological, and behavioral activities of females and, to some extent, males. Estrogen action is regulated and mediated by various genes and their protein products, including ERGs. Thus, ERGs can represent various estrogenic cell functions, and therefore, have been used as markers for estrogenic activity (reviewed in Kiyama & Zhu, 2014 [4]). Previously, specific ERGs were used as markers for gene expression or used as reporter-genes, which were later replaced by a set of ERGs in more statistically reliable transcriptomic assays such as DNA microarray assays and RNA-seq. A small number of ERGs, tens to hundreds, were enough for statistical stability, and including the genes without any response would increase the noise level. However, the quality of ERGs has not been extensively explored. Here, we found a set of 300 ERGs newly identified by six sets of RNA-seq data obtained from E_2 -treated and control MCF-7 cells (Figs 1 and 2). A set of the 30 genes with best statistical stability were further evaluated quantitatively by RT-PCR (Fig 3), and then qualitatively by a functional analysis using the GO and KEGG databases (Fig 4) and by a mechanistic analysis using ChIP-seq data (Fig 5). Since the 30 ERGs identified here exhibited alterations in gene expression at significant levels, they are associated with various cell functions and cell signaling pathways. How the genes are functionally associated with estrogen action and how they respond to estrogen are the questions that need to be examined further.

Functional significance of ERGs

The 30 ERGs analyzed here can be classified into several functional categories. Firstly, the genes that encode enzymes, including enzyme activators and inhibitors, such as metabolic enzymes, proteases, and protein kinases. The genes encoding metabolic enzymes are: *ACOX2* encoding branched-chain acyl-CoA oxidase, a peroxisomal enzyme involved in the metabolism of branched-chain fatty acids and bile acid intermediates [45]; *B4GALT1* encoding β -1,4-galactosyltransferase 1, which transfers galactose in a β -1,4-linkage to *N*-acetylglucosamine, glucose, and xylose to form a lactosamine or lactose [46]; *CYP1A1* encoding a cytochrome P450 family 1A enzyme with monooxygenase or oxidoreductase activity [47, 48]; *FDFT1* encoding farnesyl-diphosphate farnesyltransferase 1, or squalene synthase, which catalyzes the conversion of *trans*-farnesyl-diphosphate to squalene, the first specific step in cholesterol biosynthesis [49]; *LOXL2* encoding a lysyl oxidase or protein lysin 6-oxidase, which catalyzes cross-linking collagen and elastin in the extracellular matrix [50–52]. The genes associated with proteolytic activity are as follows: *CRISP3* encoding a protein with homology to plant defense proteins with lytic activity against infectious pathogens [53]; *CSTA* encoding cystatin A, a cysteine proteinase inhibitor that belongs to a large family of the cystatin superfamily [54]; *CTSD* encoding cathepsin D, a lysosomal aspartic protease with many additional functions such as those in cell proliferation, invasion, metastasis, and angiogenesis of cancers [55]; *TMPRSS3* encoding a Type II transmembrane serine protease, a membrane-bound proteolytic enzyme that is associated with biological processes such as poor prognosis in patients with breast cancer [56, 57]. The genes associated with protein kinase activity and related cellular signaling are: *PKIB* encoding a protein acting as a competitive inhibitor of the cAMP-dependent protein kinase, which plays important roles in cell signaling in diseases such as the down-regulation of Akt signaling in irritable bowel syndrome [58], preeclampsia [59], and cancers [60–62]; *PKIB* encoding a competitive inhibitor of cAMP-dependent protein kinase involved in breast cancer development by enhancing ER α action [63]; *RAB26* encoding a RAB-family small GTPase that regulates intercellular vesicle trafficking, including exocytosis, endocytosis, and recycling [64]; *RAP1GAP* encoding a GTPase-activating protein (GAP) that promotes the hydrolysis of GTP in RAP1-GTP and inactivates RAP1, which plays important roles in the regulation of cell adhesion and migration and in the progression and metastasis of several types of cancer [65]; *RAPGEFL1* encoding a protein with predicted guanyl-nucleotide exchange factor (GEF) activity, which promotes RAP1 to load GTP and acquire the active GTP-bound status [66]; *RAPGEFL1*, a gene found in several types of cancer, acting as an estrogen-regulated gene in breast cancer cells [67]; *SRGAP3* encoding a protein highly expressed in the brain, with structural characteristics of BAR/Rac1 GAP/SH3 domains and considered to be a tumor suppressor in breast cancer [68].

Secondly, the genes that are related to specific cell functions and cancer are included in the 30 ERGs. *CRISP3* protein has been known for its androgen-responsive gene expression [69], which can help to avoid apoptosis induced by estrogen deficiency in epithelial cells in the salivary gland [70]. Cathepsins are lysosomal cysteine proteases, and their aberrant expression is related to the malignancy of tumors, and thus, the inhibition of cathepsins by *CSTA* expression has been considered to be a target for cancer treatment [71]. Down-regulation of *CSTA* expression after the treatment of E₂ in ER α -positive breast cancer cells were reported [71, 72]. *CTSD* is an ERG and has been used as a marker of estrogen response [12]. The inhibition of *CTSD* activity is one of the targets for cancer treatment by controlling the expression of *CSTA* [71]. *CTSD* is also considered to be a target in antibody-based therapy for the treatment of triple-negative breast cancer due to its correlation with poor prognosis of breast cancer and providing a tumor-specific extracellular target [73, 74]. *DOK7* is

associated with congenital myasthenic syndromes, a group of neuromuscular disorders caused by impaired neuromuscular transmission, and encodes an adaptor protein which mediates tyrosine kinase signaling to form the neuromuscular synapse [75, 76]. *DOK7* is also associated with malignancy of several types of cancer, including breast cancer, where *DOK7* can inhibit cell proliferation, migration, and invasion of breast cancer by mediating signals through the PI3K/PTEN/AKT pathway, and thus, may act as a tumor suppressor gene [77]. *IGSF1* encodes a transmembrane immunoglobulin superfamily glycoprotein highly expressed in the hypothalamus and in pituitary cells, and has roles in thyroid hormone biosynthesis [78], follicle-stimulating hormone biosynthesis [79], and growth hormone secretion [80]. Although its cellular function has not been clarified, *IGSF1* mutations affecting the symptoms and disease pathologies revealed that extracellular ligands are crucial for its function [81]. *IL20* encodes a proinflammatory cytokine involved in the pathogenesis of inflammatory diseases, such as psoriasis, rheumatoid arthritis, and atherosclerosis, by inducing factors such as TNF- α , IL-1 β , MMP-1/13, and MCP-1, and plays important roles in regulating angiogenesis, osteoblastogenesis, and osteoclastogenesis [82, 83]. *IL20* was found to be an ERG [84], and its expression at the mRNA and protein levels in breast cancer cells was up-regulated by estrogen treatment [85]. *INSYNI* encodes an inhibitory postsynaptic density protein involved in postsynaptic inhibition by decreasing inhibitory currents and increasing excitability in the hippocampus [86]. *MATN2* encodes matrilin-2, a member of the matrilin subfamily of extracellular matrix proteins containing von Willebrand factor A and epidermal growth factor-like domains [87]. *MATN2* is localized downstream of several growth factor receptors, such as IGF-1R, GPER, and TGF β 1, and involved in tumor progression by stimulating the growth, chemotaxis, and migration of cancer cells [88, 89]. *PDLIM3* encodes a protein containing a PDZ and LIM domain, which are used to bind to the actin cytoskeleton, and localizing at the Z-line of mature muscle fibers, and thus, the gene is likely involved in the development and maintenance of muscle cells [90, 91]. *SUSD3* encodes a cell surface protein with the sushi domain, a structure with a characteristic sandwich arrangement of β -strands, and has been known as an ERG [92]. The possibility of *SUSD3* as a prognostic and diagnostic marker of breast cancer was investigated in several studies, including a DNA microarray analysis [93], a pathological study [94] and an epidemiological study [95].

Additionally, several genes are considered to be tumor suppressor genes. *CCDC68* is a putative tumor suppressor gene in several types of cancer [96], partly through regulating the cell cycle by controlling the degradation of CDK4 [97]. *FRY* encodes an evolutionary conserved protein with diverse functions, such as those controlling cell growth and morphogenesis through FRY-NDR kinase signaling, in invertebrates and vertebrates [98]. In mammals, FRY suppresses nuclear localization of YAP, a transcriptional coactivator with a role in promoting cell proliferation, and thus, may act as a tumor suppressor [99]. *FRY* is required for mammary gland development and may have a role in the suppression of breast cancer cell growth and proliferation [100].

Moreover, the genes associated with breast cancer are included in the 30 ERGs. The expression of *B4GALT1* has been known to quickly respond to estrogen treatment [6, 101], and the ER α -dependent activation of cell proliferation through membrane B4GALT1 implicated its crucial role in breast cancer development [102]. *CYP1A1* is involved in estradiol metabolism by the hydroxylation of estradiol to 2-hydroxyestradiol, which may contribute to the anti-estrogenic effects of aryl hydrocarbon receptor ligands such as TCDD [103]. Due to the connection of cholesterol metabolism to cancer, *FDFT1* has been considered as a cancer prognostic marker and a target for anticancer therapy [104]. The involvement of *FDFT1* in breast cancer progression through interaction with PGRMC1, a key protein in lipid metabolism, is considered to be a potential basis for breast cancer treatment [105]. *LOXL2* stimulates the

proliferation of breast cancer cells and contributes to their oncogenic transformation and pathogenesis, suggesting that the development of LOXL2 inhibitors is an effective therapeutic strategy [106]. *SYNE1* encodes nesprin-1, a member of the spectrin family of structural proteins that links the plasma membrane to the actin cytoskeleton [107]. Since *SYNE1* is located 19 kb downstream of *ESR1* (the ER α gene), its association with estrogen action in ovarian cancer [108], endometrial cancer [109], and endometriosis [110] has been reported. *TMPRSS3* was identified as an ERG by DNA microarray analysis, and its transcription level was increased after treatment with E₂ in MCF-7 cells with the help of cohesin, a protein complex that may connect distant transcriptional regulatory elements with gene promoters [111]. Furthermore, among 150 breast cancer risk regions identified by genome-wide association studies [112], several ERGs identified in this study (*BARX2*, *CBX6*, *KCNN4*, *MYEOV*, *PDZK1* in 300 ERGs) were included as the targets for cancer drivers, transcription factors, and signaling proteins. We also found that some of the ERGs identified here are included in the genes listed in cancer genomics databases such as The Cancer Genome Atlas (TCGA; portal.gdc.cancer.gov) and International Cancer Genome Consortium (ICGC; dcc.icgc.org) (12 genes found in TCGA and 29 genes in ICGC; see S5 Table).

Thirdly, several ERGs are associated with transcriptional regulation. *BARX2* was identified as a homeodomain transcription factor that binds to the transcriptional regulatory regions in the genes encoding cell adhesion molecules (CAMs) [113] and may have important roles in skeletal muscle development, such as digit chondrogenesis, by controlling the expression of CAMs [114]. The coordination of *BARX2* with ER was suggested to regulate the growth, survival, and invasion of breast cancer cells [115]. *EGR3* encodes a transcription factor belonging to a family (the early growth response family) of proteins characterized by conserved zinc finger motifs, which are associated with various roles such as the development of muscle spindles [116], the suppression of excessive immune response [117], and the progression of breast cancer [118]. *EGR3*, known as an ERG [119], has been used for evaluation of estrogenic activity of the constituents in foods and medicines [120]. *GATA4* encodes a transcription factor that binds to the GATA motif as a *cis*-element and transactivates the genes for various cell functions such as the development of heart, liver, lung, and urogenital cells [121, 122]. Due to its suppressive effects against ovarian and colorectal cancer cell proliferation, *GATA4* is considered to act as a potential tumor suppressor gene [123, 124]. *GATA4* is also associated with bone differentiation through the signaling pathways such as the ER α and TGF β pathways [125, 126]. *SPOCD1* encodes a TFIIS-family transcription factor and has been reported to promote the progression and metastasis of several types of cancer, including bladder and ovarian cancers, and osteosarcoma and glioblastoma, by activating signaling pathways such as the EGFR and VEGFR pathways [127–130]. *ZNF521* encodes a zinc-finger transcription factor that has roles in various biological processes involving stem cells, such as erythroid/myeloid/adipocyte/osteoblast differentiation, neural development, and bone formation, and is also implicated in diseases such as pancreatic, hepatic, gastric, bladder, breast, and ovarian cancers [131]. The role of *ZNF521* in ER-positive breast cancer was indicated to be the expressional regulation of *ZNF423*, a homolog of *ZNF521*, which forms a heterodimer with *ZNF521*, through the transcriptional regulation involving ERGs [132].

Fourthly, the ERGs that are explicitly known for estrogen responsiveness include *B4GALT1*, *CSTA*, *CTSD*, *CYP1A1*, *EGR3*, *IL20*, *PKIB*, *RAPGEFL1*, *SUSD3*, and *TMPRSS3* (for *CTSD* and *EGR3*, see Kiyama et al., 2014 [12]). Additionally, the genes known for the association with estrogen signaling and function (some of which are discussed above) are: *BARX2* (breast cancer), *B4GALT1* (breast cancer), *CTSD* (triple-negative breast cancer), *CYP1A1* (estradiol metabolism), *DOK7* (breast cancer), *EGR3* (breast cancer), *FRY* (breast cancer), *GATA4* (bone differentiation), *IL20* (breast cancer and epigenetic regulation), *LOXL2* (breast

cancer), *MATN2* (GPER signaling), *PKIB* (breast cancer), *RAB26* (breast cancer), *RAP1GAP* (GPER signaling), *SUSD3* (breast cancer), *SYNE1* (breast cancer), *TMPRSS3* (breast cancer), and *ZNF521* (breast cancer).

Functions of ERGs through ER-binding

A genome-wide analysis of ERs and RNA polymerase II binding sites by ChIP-seq analysis has revealed a new paradigm of estrogen-mediated gene expression. For example, ChIP-seq analysis revealed the master transcriptional regulator, which may be the best cancer therapeutic target [133]. The mechanisms identified by ChIP-seq include the transcriptions of the genes with E₂-regulated promoters and enhancers, miRNAs, enhancer RNAs, long-ncRNAs, and anti-sense RNAs [134]. The cistrome, or the genome-wide binding landscape, identified by ChIP-seq using ER α phosphorylated at serine 118 revealed the involvement of this site in a genome-wide transcriptional regulation, such as that by the transcription factor GRHL2, in estrogen action [135]. Such cistromic studies can be applied to develop breast cancer treatment [136].

About one third of the ERGs contained functional ER α -binding sites near their transcription start sites [137]. ChIP-seq analysis along with functional annotations of the genes in the vicinity of ER α -binding sites using databases, such as the GO, KEGG, and GeneGo databases, revealed that the genes with specific functions, such as biological regulation, cellular processes, and developmental processes, are the targets of ER α [138]. Furthermore, specific regulatory networks consisting of transcription factors, modulators, and targets, such as GRHL2, RUNX1, SRC1/SRC2/SRC3, and ER β , have been identified as potential ER signaling regulators [135, 139–141]. In our study, the ER-subtype specificity among 29 ERGs was 10 ER α -specific, 2 ER β -specific, and 14 ER α / β -specific genes, along with 3 genes classified as other types (Table 2). Thus, more than 80% of the identified ERGs are ER α -dependent. Interestingly, the ER α non-genomic network is more reliable than the genomic network [141], and thus, may be predominant in MCF-7 cells.

Three ERGs identified here (*ZNF521*, *IGSF1*, and *CRISP3*) were found to be ER α / β -independent, suggesting that they are involved in other signaling pathways. The expression of *ZNF521* is known to potentially be controlled by *ZNF423*, an ER α -dependent ERG [132], and thus, its regulation may be indirect, as observed here. On the other hand, the expressions of *IGSF1* and *CRISP3* are known to be androgen receptor (AR)-dependent [142] and characteristic in prostate cancer cells [143]. The expression and transcriptional regulation of the AR in breast cancer cells, including MCF-7 cells, has been known [144, 145], and the crosstalk between ER α and the AR in breast cancer progression [146, 147] is a likely mechanism of transcriptional regulation of these two genes.

Based on ChIP-seq data, a total of 80,000 potential ER α -binding sites were classified into four groups based on the variations of the presence or absence of both estrogen-responsive elements and the treatment with E₂ [148], which indicates that some ER α -binding sites are not E₂-dependent. Evidently, ER α is present prior to ligand activation in the majority of the binding sites [148]. An independent study also revealed the presence of estrogen-independent ER activity such as enhancer activity and the associated activity of histone modifications and enhancer RNA transcription [149]. Additionally, there are differences in ChIP-seq data for ER α -binding sites detected after the treatment with different ligands, E₂, tamoxifen, and ICI 182,780, suggesting the mechanistic differences of ligand actions [150]. Thus, the ER α / β -independent ERGs may represent the types of transcriptional regulation with more complex mechanisms of estrogen action.

Conclusions

In this paper, we identified the ERGs that show statistical stability based on the CV values. We evaluated their usefulness as markers for estrogenic activity by examining the stability of the data when compared with correlation analysis, RT-PCR, and functional association with estrogen action through database searches. The ERGs identified here have been known for (1) the enzymes, such as metabolic enzymes, proteases, and protein kinases, (2) the genes with specific cell functions, such as cell-signaling mediators, tumor-suppressors, and roles in breast cancer, (3) the association with transcriptional regulation, and (4) estrogen-responsiveness. The top 30 ERGs were further classified as ER α -binding, ER β -binding, ER α/β -binding, or ER-independent types based on ChIP-seq data. Therefore, the ERGs identified here represent various cell functions and cell signaling pathways, including estrogen signaling, and thus, may be useful for evaluations of estrogenic activity.

Supporting information

S1 Table. List of the primers for *real-time* RT-PCR analysis for the top 30 ERGs.

(XLSX)

S2 Table. Correlation analysis of RNA-seq datasets.

(DOCX)

S3 Table. List of 300 ERGs stably responding to E₂ and their expression data. A total of 300 ERGs containing 150 up- and 150 down-regulated genes among ERGs stably responding to E₂.

(XLSX)

S4 Table. Functional analysis of ERGs. (A) up-regulated genes in GO analysis; (B) down-regulated genes in GO analysis; (C) up-regulated genes in KEGG analysis; (D) down-regulated genes in KEGG analysis.

(XLSX)

S5 Table. List of 30 or 300 ERGs listed in the TGCA and ICGC databases.

(DOCX)

S1 Raw images.

(PDF)

Author Contributions

Conceptualization: Ryoiti Kiyama.

Data curation: Kentaro Nishi, Wenqiang Fu, Ryoiti Kiyama.

Formal analysis: Kentaro Nishi, Wenqiang Fu, Ryoiti Kiyama.

Funding acquisition: Ryoiti Kiyama.

Investigation: Kentaro Nishi, Wenqiang Fu, Ryoiti Kiyama.

Methodology: Kentaro Nishi, Wenqiang Fu, Ryoiti Kiyama.

Project administration: Ryoiti Kiyama.

Resources: Kentaro Nishi, Ryoiti Kiyama.

Software: Kentaro Nishi, Ryoiti Kiyama.

Supervision: Ryoiti Kiyama.

Validation: Kentaro Nishi, Wenqiang Fu, Ryoiti Kiyama.

Visualization: Kentaro Nishi, Ryoiti Kiyama.

Writing – original draft: Kentaro Nishi, Ryoiti Kiyama.

Writing – review & editing: Ryoiti Kiyama.

References

1. Kiyama R, Wada-Kiyama Y. Estrogenic endocrine disruptors: Molecular mechanisms of action. *Environ Int.* 2015; 83:11–40. <https://doi.org/10.1016/j.envint.2015.05.012> PMID: 26073844
2. Kiyama R. Nutritional implications of ginger: chemistry, biological activities and signaling pathways. *J Nutr Biochem.* 2020; 86:108486. <https://doi.org/10.1016/j.jnutbio.2020.108486> PMID: 32827666
3. Tanji M, Kiyama R. Expression profiling of estrogen responsive genes using genomic and proteomic techniques for the evaluation of endocrine disruptors. *Curr Pharmacogenomics.* 2004; 2(3):255–266. <https://doi.org/10.2174/1570160043377529>
4. Kiyama R, Zhu Y. DNA microarray-based gene expression profiling of estrogenic chemicals. *Cell Mol Life Sci.* 2014; 71(11):2065–82. <https://doi.org/10.1007/s00018-013-1544-5> PMID: 24399289
5. Ikeda K, Horie-Inoue K, Inoue S. Identification of estrogen-responsive genes based on the DNA binding properties of estrogen receptors using high-throughput sequencing technology. *Acta Pharmacol Sin.* 2015; 36(1):24–31. <https://doi.org/10.1038/aps.2014.123> PMID: 25500870
6. Inoue A, Yoshida N, Omoto Y, Oguchi S, Yamori T, Kiyama R, et al. Development of cDNA microarray for expression profiling of estrogen-responsive genes. *J Mol Endocrinol.* 2002; 29(2):175–92. <https://doi.org/10.1677/jme.0.0290175> PMID: 12370120
7. Lobenhofer EK, Bennett L, Cable PL, Li L, Bushel PR, Afshari CA. Regulation of DNA replication fork genes by 17 β -estradiol. *Mol Endocrinol.* 2002; 16(6):1215–29. <https://doi.org/10.1210/mend.16.6.0858> PMID: 12040010
8. Terasaka S, Aita Y, Inoue A, Hayashi S, Nishigaki M, Aoyagi K, et al. Using a customized DNA microarray for expression profiling of the estrogen-responsive genes to evaluate estrogen activity among natural estrogens and industrial chemicals. *Environ Health Perspect.* 2004; 112(7):773–81. <https://doi.org/10.1289/ehp.6753> PMID: 15159206
9. Cicatiello L, Scafoglio C, Altucci L, Cancemi M, Natoli G, Facchiano A, et al. A genomic view of estrogen actions in human breast cancer cells by expression profiling of the hormone-responsive transcriptome. *J Mol Endocrinol.* 2004; 32(3):719–75. <https://doi.org/10.1677/jme.0.0320719> PMID: 15171711
10. Lin CY, Ström A, Vega VB, Kong SL, Yeo AL, Thomsen JS, et al. Discovery of estrogen receptor α target genes and response elements in breast tumor cells. *Genome Biol.* 2004; 5(9):R66. <https://doi.org/10.1186/gb-2004-5-9-r66> PMID: 15345050
11. Fujimoto N, Igarashi K, Kanno J, Inoue T. Identification of estrogen-responsive genes in the GH3 cell line by cDNA microarray analysis. *J Steroid Biochem Mol Biol.* 2004; 91(3):121–9. <https://doi.org/10.1016/j.jsbmb.2004.02.006> PMID: 15276619
12. Kiyama R, Zhu Y, Kawaguchi K, Iitake N, Wada-Kiyama Y, Dong S. Estrogen-responsive genes for environmental studies. *Environ Technol Innov.* 2014; 1–2:16–28. <https://doi.org/10.1016/j.eti.2014.09.001>
13. Inoue A, Tanji M, Kiyama R. Focused microarray analysis: characterization of phenomes by gene expression profiling. *Curr Pharmacogenomics* 2006; 4:245–260.
14. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature.* 2002; 415(6871):530–6. <https://doi.org/10.1038/415530a> PMID: 11823860
15. Terasaka S, Inoue A, Tanji M, Kiyama R. Expression profiling of estrogen-responsive genes in breast cancer cells treated with alkylphenols, chlorinated phenols, parabens, or bis- and benzoylphenols for evaluation of estrogenic activity. *Toxicol Lett.* 2006; 163(2):130–41. <https://doi.org/10.1016/j.toxlet.2005.10.005> PMID: 16280211
16. Marguerat S, Bähler J. RNA-seq: from technology to biology. *Cell Mol Life Sci.* 2010; 67(4):569–79. <https://doi.org/10.1007/s00018-009-0180-6> PMID: 19859660
17. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for RNA-seq data analysis. *Genome Biol.* 2016; 17:13. <https://doi.org/10.1186/s13059-016-0881-8> PMID: 26813401
18. Xu J, Gong B, Wu L, Thakkar S, Hong H, Tong W. Comprehensive Assessments of RNA-seq by the SEQC Consortium: FDA-Led Efforts Advance Precision Medicine. *Pharmaceutics.* 2016; 8(1):8. <https://doi.org/10.3390/pharmaceutics8010008> PMID: 26999190

19. Lamas A, Regal P, Vázquez B, Miranda JM, Franco CM, Cepeda A. Transcriptomics: A powerful tool to evaluate the behavior of foodborne pathogens in the food production chain. *Food Res Int*. 2019; 125:108543. <https://doi.org/10.1016/j.foodres.2019.108543> PMID: 31554082
20. Niazi M. Application of genetics and biotechnology for improving medicinal plants. *Planta*. 2019; 249(4):953–973. <https://doi.org/10.1007/s00425-019-03099-1> PMID: 30715560
21. Wang S, Li X, Zhang W, Gao Y, Zhang K, Hao Q, et al. Genome-wide investigation of genes regulated by ER α in breast cancer cells. *Molecules*. 2018; 23(10):2543. <https://doi.org/10.3390/molecules23102543> PMID: 30301189
22. Hahm ER, Kim SH, Mathan SV, Singh RP, Singh SV. Mechanistic targets of diallyl trisulfide in human breast cancer cells identified by RNA-seq analysis. *J Cancer Prev*. 2021; 26(2):128–136. <https://doi.org/10.15430/JCP.2021.26.2.128> PMID: 34258251
23. Li H, An X, Li Q, Yu H, Li Z. Construction and analysis of competing endogenous RNA network of MCF-7 breast cancer cells based on the inhibitory effect of 6-thioguanine on cell proliferation. *Oncol Lett*. 2021; 21(2):104. <https://doi.org/10.3892/ol.2020.12365> PMID: 33376537
24. Lin KH, Huang MY, Cheng WC, Wang SC, Fang SH, Tu HP, et al. RNA-seq transcriptome analysis of breast cancer cell lines under shikonin treatment. *Sci Rep*. 2018; 8(1):2672. <https://doi.org/10.1038/s41598-018-21065-x> PMID: 29422643
25. Men X, Ma J, Wu T, Pu J, Wen S, Shen J, et al. Transcriptome profiling identified differentially expressed genes and pathways associated with tamoxifen resistance in human breast cancer *Oncotarget*. 2017; 9(3):4074–4089. <https://doi.org/10.18632/oncotarget.23694> PMID: 29423105
26. Wang J, Yu H, Yili A, Gao Y, Hao L, Aisa HA, et al. Identification of hub genes and potential molecular mechanisms of chickpea isoflavones on MCF-7 breast cancer cells by integrated bioinformatics analysis. *Ann Transl Med*. 2020; 8(4):86. <https://doi.org/10.21037/atm.2019.12.141> PMID: 32175379
27. Ferraro L, Ravo M, Nassa G, Tarallo R, De Filippo MR, Giurato G, et al. Effects of oestrogen on micro-RNA expression in hormone-responsive breast cancer cells. *Horm Cancer*. 2012; 3(3):65–78. <https://doi.org/10.1007/s12672-012-0102-1> PMID: 22274890
28. Li MX, Wang HY, Yuan CH, Ma ZL, Jiang B, Li L, et al. Establishment of a macrophage phenotypic switch related prognostic signature in patients with pancreatic cancer. *Front Oncol*. 2021; 11:619517. <https://doi.org/10.3389/fonc.2021.619517> PMID: 33747931
29. Hah N, Murakami S, Nagari A, Danko CG, Kraus WL. Enhancer transcripts mark active estrogen receptor binding sites. *Genome Res*. 2013; 23(8):1210–23. <https://doi.org/10.1101/gr.152306.112> PMID: 23636943
30. Coscujuela Tarrero L, Ferrero G, Miano V, De Intinis C, Ricci L, Arigoni M, et al. Luminal breast cancer-specific circular RNAs uncovered by a novel tool for data analysis. *Oncotarget*. 2018; 9(18):14580–14596. <https://doi.org/10.18632/oncotarget.24522> PMID: 29581865
31. Smid M, Wilting SM, Uhr K, Rodríguez-González FG, de Weerd V, Prager-Van der Smissen WJC, van der Vlugt-Daane M, van Galen A, et al. The circular RNome of primary breast cancer. *Genome Res*. 2019; 29(3):356–366. <https://doi.org/10.1101/gr.238121.118> PMID: 30692147
32. Chen X, Gu J, Wang X, Jung JG, Wang TL, Hilakivi-Clarke L, et al. CRNET: an efficient sampling approach to infer functional regulatory networks by integrating large-scale ChIP-seq and time-course RNA-seq data. *Bioinformatics*. 2018; 34(10):1733–1740. <https://doi.org/10.1093/bioinformatics/btx827> PMID: 29280996
33. Dong S, Terasaka S, Kiyama R. Bisphenol A induces a rapid activation of Erk1/2 through GPR30 in human breast cancer cells. *Environ Pollut*. 2011; 159(1):212–218. <https://doi.org/10.1016/j.envpol.2010.09.004> PMID: 20875696
34. Dong S, Furutani Y, Kimura S, Zhu Y, Kawabata K, Furutani M, et al. Brefeldin A is an estrogenic, Erk1/2-activating component in the extract of *Agaricus blazei* mycelia. *J Agric Food Chem*. 2013; 61(1):128–36. <https://doi.org/10.1021/jf304546a> PMID: 23215459
35. Shen W, Le S, Li Y, Hu F. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One*. 2016; 11(10):e0163962. <https://doi.org/10.1371/journal.pone.0163962> PMID: 27706213
36. Marcel M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 2011; 17(1):10–12. <https://doi.org/10.14806/ej.17.1.200>
37. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015; 12(4):357–60. <https://doi.org/10.1038/nmeth.3317> PMID: 25751142
38. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010; 28(5):511–5. <https://doi.org/10.1038/nbt.1621> PMID: 20436464
39. Weidenfeld K, Schif-Zuck S, Abu-Tayeh H, Kang K, Kessler O, Weissmann M, et al. Dormant tumor cells expressing LOXL2 acquire a stem-like phenotype mediating their transition to proliferative

- growth. *Oncotarget*. 2016; 7(44):71362–71377. <https://doi.org/10.18632/oncotarget.12109> PMID: 27655685
40. Duong NT, Morris GE, Lam le T, Zhang Q, Sewry CA, Shanahan CM, et al. Nesprins: tissue-specific expression of epsilon and other short isoforms. *PLoS One*. 2014; 9(4):e94380. <https://doi.org/10.1371/journal.pone.0094380> PMID: 24718612
 41. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res*. 2019; 47(W1):W199–W205. <https://doi.org/10.1093/nar/gkz401> PMID: 31114916
 42. Zhu Y, Kawaguchi K, Kiyama R. Differential and directional estrogenic signaling pathways induced by enterolignans and their precursors. *PLoS One*. 2017; 12(2):e0171390. <https://doi.org/10.1371/journal.pone.0171390> PMID: 28152041
 43. Zhu Y, Kiyama R. Capsaicinoids are silent estrogens, a class of estrogenic chemicals without cell-proliferation activity. *Environ Technol Innov*. 2017; 7(1–2):182–193. <https://doi.org/10.1016/j.eti.2017.02.006>
 44. Kiyama R, Furutani Y, Kawaguchi K, Nakanishi T. Genome sequence of the cauliflower mushroom *Sparassis crispa* (Hanabiratake) and its association with beneficial usage. *Sci Rep*. 2018; 8(1):16053. <https://doi.org/10.1038/s41598-018-34415-6> PMID: 30375506
 45. Vilarinho S, Sari S, Mazzacuva F, Bilgüvar K, Esendagli-Yilmaz G, Jain D, et al. ACOX2 deficiency: A disorder of bile acid synthesis with transaminase elevation, liver fibrosis, ataxia, and cognitive impairment. *Proc Natl Acad Sci U S A*. 2016; 113(40):11289–11293. <https://doi.org/10.1073/pnas.1613228113> PMID: 27647924
 46. Lee-Sundlov MM, Stowell SR, Hoffmeister KM. Multifaceted role of glycosylation in transfusion medicine, platelets, and red blood cells. *J Thromb Haemost*. 2020; 18(7):1535–1547. <https://doi.org/10.1111/jth.14874> PMID: 32350996
 47. Santes-Palacios R, Ornelas-Ayala D, Cabañas N, Marroquín-Pérez A, Hernández-Magaña A, Del Rosario Olguín-Reyes S, et al. Regulation of human cytochrome P4501A1 (hCYP1A1): A plausible target for chemoprevention? *Biomed Res Int*. 2016; 2016:5341081. <https://doi.org/10.1155/2016/5341081> PMID: 28105425
 48. Sridhar J, Goyal N, Liu J, Foroozesh M. Review of ligand specificity factors for CYP1A subfamily enzymes from molecular modeling studies reported to-date. *Molecules*. 2017; 22(7):1143. <https://doi.org/10.3390/molecules22071143> PMID: 28698457
 49. Shechter I, Conrad DG, Hart I, Berger RC, McKenzie TL, Bleskan J, et al. Localization of the squalene synthase gene (FDFT1) to human chromosome 8p22-p23.1. *Genomics*. 1994; 20(1):116–8. <https://doi.org/10.1006/geno.1994.1135> PMID: 8020937
 50. Moon HJ, Finney J, Xu L, Moore D, Welch DR, Mure M. MCF-7 cells expressing nuclear associated lysyl oxidase-like 2 (LOXL2) exhibit an epithelial-to-mesenchymal transition (EMT) phenotype and are highly invasive *in vitro*. *J Biol Chem*. 2013; 288(42):30000–30008. <https://doi.org/10.1074/jbc.C113.502310> PMID: 24014025
 51. Wu L, Zhu Y. The function and mechanisms of action of LOXL2 in cancer (Review). *Int J Mol Med*. 2015; 36(5):1200–4. <https://doi.org/10.3892/ijmm.2015.2337> PMID: 26329904
 52. Wen B, Xu LY, Li EM. LOXL2 in cancer: regulation, downstream effectors and novel roles. *Biochim Biophys Acta Rev Cancer*. 2020; 1874(2):188435. <https://doi.org/10.1016/j.bbcan.2020.188435> PMID: 32976981
 53. Pfisterer P, König H, Hess J, Lipowsky G, Haendler B, Schleuning WD, et al. CRISP-3, a protein with homology to plant defense proteins, is expressed in mouse B cells under the control of Oct2. *Mol Cell Biol*. 1996; 16(11):6160–8. <https://doi.org/10.1128/MCB.16.11.6160> PMID: 8887646
 54. Keppler D. Towards novel anti-cancer strategies based on cystatin function. *Cancer Lett*. 2006; 235(2):159–76. <https://doi.org/10.1016/j.canlet.2005.04.001> PMID: 15893421
 55. Benes P, Vetvicka V, Fusek M. Cathepsin D—many functions of one aspartic protease. *Crit Rev Oncol Hematol*. 2008; 68(1):12–28. <https://doi.org/10.1016/j.critrevonc.2008.02.008> PMID: 18396408
 56. Rui X, Li Y, Jin F, Li F. Tmprss3 is a novel poor prognostic factor for breast cancer. *Int J Clin Exp Pathol*. 2015; 8(5):5435–42. PMID: 26191247
 57. Luo P, Lu G, Fan LL, Zhong X, Yang H, Xie R, et al. Dysregulation of Tmprss3 and TNFRSF11B correlates with tumorigenesis and poor prognosis in patients with breast cancer. *Oncol Rep*. 2017; 37(4):2057–2062. <https://doi.org/10.3892/or.2017.5449> PMID: 28260080
 58. Fei L, Wang Y. microRNA-495 reduces visceral sensitivity in mice with diarrhea-predominant irritable bowel syndrome through suppression of the PI3K/AKT signaling pathway via PKIB. *IUBMB Life*. 2020; 72(7):1468–1480. <https://doi.org/10.1002/iub.2270> PMID: 32187820

59. Liu C, Wang H, Yang M, Liang Y, Jiang L, Sun S, et al. Downregulation of cAMP-dependent protein kinase inhibitor- β promotes preeclampsia by decreasing phosphorylated Akt. *Reprod Sci*. 2021; 28(1):178–185. <https://doi.org/10.1007/s43032-020-00258-8> PMID: 32676926
60. Chung S, Furihata M, Tamura K, Uemura M, Daigo Y, Nasu Y, et al. Overexpressing PKIB in prostate cancer promotes its aggressiveness by linking between PKA and Akt pathways. *Oncogene*. 2009; 28(32):2849–59. <https://doi.org/10.1038/onc.2009.144> PMID: 19483721
61. Dou P, Zhang D, Cheng Z, Zhou G, Zhang L. PKIB promotes cell proliferation and the invasion-metastasis cascade through the PI3K/Akt pathway in NSCLC cells. *Exp Biol Med (Maywood)*. 2016; 241(17):1911–1918. <https://doi.org/10.1177/1535370216655908> PMID: 27325557
62. Zhang JB, Song W, Wang YY, Liu MG, Sun MM, Liu H. Study on correlation between PKIB and pAkt expression in breast cancer tissues. *Eur Rev Med Pharmacol Sci*. 2017; 21(6):1264–1269. PMID: 28387904
63. Böckers M, Paul NW, Efferth T. Indeno[1,2,3-*cd*]pyrene and picene mediate actions via estrogen receptor α signaling pathway in *in vitro* cell systems, altering gene expression. *Toxicol Appl Pharmacol*. 2020; 396:114995. <https://doi.org/10.1016/j.taap.2020.114995> PMID: 32251684
64. Seki N, Yoshikawa T, Hattori A, Miyajima N, Muramatsu M, Saito T. cDNA cloning of a human RAB26-related gene encoding a Ras-like GTP-binding protein on chromosome 16p13.3 region. *J Hum Genet*. 2000; 45(5):309–14. <https://doi.org/10.1007/s100380070023> PMID: 11043516
65. Shah S, Brock EJ, Ji K, Mattingly RR. Ras and Rap1: A tale of two GTPases. *Semin Cancer Biol*. 2019; 54:29–39. <https://doi.org/10.1016/j.semcancer.2018.03.005> PMID: 29621614
66. Rebhun JF, Castro AF, Quilliam LA. Identification of guanine nucleotide exchange factors (GEFs) for the Rap1 GTPase. Regulation of MR-GEF by M-Ras-GTP interaction. *J Biol Chem*. 2000; 275(45):34901–8. <https://doi.org/10.1074/jbc.M005327200> PMID: 10934204
67. Wardell SE, Kazmin D, McDonnell DP. Research resource: Transcriptional profiling in a cellular model of breast cancer reveals functional and mechanistic differences between clinically relevant SERM and between SERM/estrogen complexes. *Mol Endocrinol*. 2012; 26(7):1235–48. <https://doi.org/10.1210/me.2012-1031> PMID: 22570330
68. Lahoz A, Hall A. A tumor suppressor role for srGAP3 in mammary epithelial cells. *Oncogene*. 2013; 32(40):4854–60. <https://doi.org/10.1038/onc.2012.489> PMID: 23108406
69. Schwidetzky U, Haendler B, Schleuning WD. Isolation and characterization of the androgen-dependent mouse cysteine-rich secretory protein-3 (CRISP-3) gene. *Biochem J*. 1995; 309(Pt 3):831–6. <https://doi.org/10.1042/bj3090831> PMID: 7639699
70. Trokovic N, Pöllänen R, Porola P, Stegaev V, Hetzel U, Tivesten Å, et al. Exosomal secretion of death bullets: a new way of apoptotic escape? *Am J Physiol Endocrinol Metab*. 2012; 303(8):E1015–24. <https://doi.org/10.1152/ajpendo.00139.2012> PMID: 22912365
71. John Mary DJS, Manjgowda MC, Kumar A, Dutta S, Limaye AM. The role of cystatin A in breast cancer and its functional link with ER α . *J Genet Genomics*. 2017; 44(12):593–597. <https://doi.org/10.1016/j.jgg.2017.10.001> PMID: 29246862
72. Vendrell JA, Magnino F, Danis E, Duchesne MJ, Pinloche S, Pons M, et al. Estrogen regulation in human breast cancer cells of new downstream gene targets involved in estrogen metabolism, cell proliferation and cell transformation. *J Mol Endocrinol*. 2004; 32(2):397–414. <https://doi.org/10.1677/jme.0.0320397> PMID: 15072547
73. Ashraf Y, Mansouri H, Laurent-Matha V, Alcaraz LB, Roger P, Guiu S, et al. Immunotherapy of triple-negative breast cancer with cathepsin D-targeting antibodies. *J Immunother Cancer*. 2019; 7(1):29. <https://doi.org/10.1186/s40425-019-0498-z> PMID: 30717773
74. Alcaraz LB, Mallavialle A, David T, Derocq D, Delolme F, Dieryckx C, et al. A 9-kDa matricellular SPARC fragment released by cathepsin D exhibits pro-tumor activity in the triple-negative breast cancer microenvironment. *Theranostics*. 2021; 11(13):6173–6192. <https://doi.org/10.7150/thno.58254> PMID: 33995652
75. Finsterer J. Congenital myasthenic syndromes. *Orphanet J Rare Dis*. 2019; 14(1):57. <https://doi.org/10.1186/s13023-019-1025-5> PMID: 30808424
76. Herbst R. MuSk function during health and disease. *Neurosci Lett*. 2020; 716:134676. <https://doi.org/10.1016/j.neulet.2019.134676> PMID: 31811897
77. Yue C, Bai Y, Piao Y, Liu H. DOK7 Inhibits Cell Proliferation, Migration, and Invasion of Breast Cancer via the PI3K/PTEN/AKT Pathway. *J Oncol*. 2021; 2021:4035257. <https://doi.org/10.1155/2021/4035257> PMID: 33552156
78. Schoenmakers N, Alatzoglou KS, Chatterjee VK, Dattani MT. Recent advances in central congenital hypothyroidism. *J Endocrinol*. 2015; 227(3):R51–71. <https://doi.org/10.1530/JOE-15-0341> PMID: 26416826

79. Bernard DJ, Smith CL, Br ul e E. A Tale of Two Proteins: Betaglycan, IGSF1, and the continuing search for the inhibin B receptor. *Trends Endocrinol Metab.* 2020; 31(1):37–45. <https://doi.org/10.1016/j.tem.2019.08.014> PMID: 31648935
80. Joustra SD, Roelfsema F, van Trotsenburg ASP, Schneider HJ, Kosilek RP, Kroon HM, et al. IGSF1 deficiency results in human and murine somatotrope neurosecretory hyperfunction. *J Clin Endocrinol Metab.* 2020; 105(3):e70–84. <https://doi.org/10.1210/clinem/dgz093> PMID: 31650157
81. Faucz FR, Horvath AD, Azevedo MF, Levy I, Bak B, Wang Y, et al. Is IGSF1 involved in human pituitary tumor formation? *Endocr Relat Cancer.* 2015; 22(1):47–54. <https://doi.org/10.1530/ERC-14-0465> PMID: 25527509
82. Chen J, Caspi RR, Chong WP. IL-20 receptor cytokines in autoimmune diseases. *J Leukoc Biol.* 2018; 104(5):953–959. <https://doi.org/10.1002/JLB.MR1117-471R> PMID: 30260500
83. Wei H, Li B, Sun A, Guo F. Interleukin-10 family cytokines immunobiology and structure. *Adv Exp Med Biol.* 2019; 1172:79–96. https://doi.org/10.1007/978-981-13-9367-9_4 PMID: 31628652
84. Su CH, Lin IH, Tzeng TY, Hsieh WT, Hsu MT. Regulation of IL-20 Expression by estradiol through KMT2B-mediated epigenetic modification. *PLoS One.* 2016; 11(11):e0166090. <https://doi.org/10.1371/journal.pone.0166090> PMID: 27806114
85. Lee JY, Park YJ, Oh N, Kwack KB, Park KS. A transcriptional complex composed of ER(α), GATA3, FOXA1 and ELL3 regulates IL-20 expression in breast cancer cells. *Oncotarget.* 2017; 8(26):42752–42760. <https://doi.org/10.18632/oncotarget.17459> PMID: 28514748
86. Uezu A, Kanak DJ, Bradshaw TW, Soderblom EJ, Catavero CM, Burette AC, et al. Identification of an elaborate complex mediating postsynaptic inhibition. *Science.* 2016; 353(6304):1123–9. <https://doi.org/10.1126/science.aag0821> PMID: 27609886
87. De ak F, Wagener R, Kiss I, Paulsson M. The matrilins: a novel family of oligomeric extracellular matrix proteins. *Matrix Biol.* 1999; 18(1):55–64. [https://doi.org/10.1016/s0945-053x\(98\)00006-7](https://doi.org/10.1016/s0945-053x(98)00006-7) PMID: 10367731
88. Avino S, De Marco P, Cirillo F, Santolla MF, De Francesco EM, Perri MG, Rigracciolo D, et al. Stimulatory actions of IGF-I are mediated by IGF-IR cross-talk with GPER and DDR1 in mesothelioma and lung cancer cells. *Oncotarget.* 2016; 7(33):52710–52728. <https://doi.org/10.18632/oncotarget.10348> PMID: 27384677
89. Tang H, Leung L, Saturno G, Viros A, Smith D, Di Leva G, et al. Lysyl oxidase drives tumour progression by trapping EGF receptors at the cell surface. *Nat Commun.* 2017; 8:14909. <https://doi.org/10.1038/ncomms14909> PMID: 28416796
90. Xia H, Winokur ST, Kuo WL, Altherr MR, Bredt DS. Actinin-associated LIM protein: identification of a domain interaction between PDZ and spectrin-like repeat motifs. *J Cell Biol.* 1997; 139(2):507–15. <https://doi.org/10.1083/jcb.139.2.507> PMID: 9334352
91. Bouju S, Pi etu G, Le Cunff M, Cros N, Malzac P, Pellissier JF, et al. Exclusion of muscle specific actinin-associated LIM protein (ALP) gene from 4q35 facioscapulohumeral muscular dystrophy (FSHD) candidate genes. *Neuromuscul Disord.* 1999; 9(1):3–10. [https://doi.org/10.1016/s0960-8966\(98\)00087-x](https://doi.org/10.1016/s0960-8966(98)00087-x) PMID: 10063829
92. Yu Z, Jiang E, Wang X, Shi Y, Shangguan AJ, Zhang L, et al. Sushi domain-containing protein 3: A potential target for breast cancer. *Cell Biochem Biophys.* 2015; 72(2):321–4. <https://doi.org/10.1007/s12013-014-0480-9> PMID: 25556073
93. Moy I, Todorovi c V, Dubash AD, Coon JS, Parker JB, Buranapramest M, et al. Estrogen-dependent sushi domain containing 3 regulates cytoskeleton organization and migration in breast cancer cells. *Oncogene.* 2015; 34(3):323–33. <https://doi.org/10.1038/onc.2013.553> PMID: 24413080
94. Zhao S, Chen SS, Gu Y, Jiang EZ, Yu ZH. Expression and clinical significance of sushi domain-containing protein 3 (SUSD3) and insulin-like growth factor-I receptor (IGF-IR) in breast cancer. *Asian Pac J Cancer Prev.* 2015; 16(18):8633–6. <https://doi.org/10.7314/apjcp.2015.16.18.8633> PMID: 26745129
95. Aushev VN, Lee E, Zhu J, Gopalakrishnan K, Li Q, Teitelbaum SL, et al. Novel predictors of breast cancer survival derived from miRNA activity analysis. *Clin Cancer Res.* 2018; 24(3):581–591. <https://doi.org/10.1158/1078-0432.CCR-17-0996> PMID: 29138345
96. Radulovich N, Leung L, Ibrahimov E, Navab R, Sakashita S, Zhu CQ, et al. Coiled-coil domain containing 68 (CCDC68) demonstrates a tumor-suppressive role in pancreatic ductal adenocarcinoma. *Oncogene.* 2015; 34(32):4238–47. <https://doi.org/10.1038/onc.2014.357> PMID: 25381825
97. Wang C, Li H, Wu L, Jiao X, Jin Z, Zhu Y, et al. Coiled-coil domain-containing 68 downregulation promotes colorectal cancer cell growth by inhibiting ITCH-mediated CDK4 degradation. *Front Oncol.* 2021; 11:668743. <https://doi.org/10.3389/fonc.2021.668743> PMID: 33968776
98. Nagai T, Mizuno K. Multifaceted roles of Furry proteins in invertebrates and vertebrates. *J Biochem.* 2014; 155(3):137–46. <https://doi.org/10.1093/jb/mvu001> PMID: 24403109

99. Irie K, Nagai T, Mizuno K. Furry protein suppresses nuclear localization of yes-associated protein (YAP) by activating NDR kinase and binding to YAP. *J Biol Chem.* 2020; 295(10):3017–3028. <https://doi.org/10.1074/jbc.RA119.010783> PMID: 31996378
100. Liu Y, Chen X, Gong Z, Zhang H, Fei F, Tang X, et al. *Fry* is required for mammary gland development during pregnant periods and affects the morphology and growth of breast cancer cells. *Front Oncol.* 2019; 9:1279. <https://doi.org/10.3389/fonc.2019.01279> PMID: 31824855
101. Wang DY, Fulthorpe R, Liss SN, Edwards EA. Identification of estrogen-responsive genes by complementary deoxyribonucleic acid microarray and characterization of a novel early estrogen-induced gene: EEIG1. *Mol Endocrinol.* 2004; 18(2):402–11. <https://doi.org/10.1210/me.2003-0202> PMID: 14605097
102. Choi HJ, Chung TW, Kim CH, Jeong HS, Joo M, Youn B, et al. Estrogen induced beta-1,4-galactosyltransferase 1 expression regulates proliferation of human breast cancer MCF-7 cells. *Biochem Biophys Res Commun.* 2012; 426(4):620–5. <https://doi.org/10.1016/j.bbrc.2012.08.140> PMID: 22982306
103. Mescher M, Haarmann-Stemmann T. Modulation of CYP1A1 metabolism: From adverse health effects to chemoprevention and therapeutic options. *Pharmacol Ther.* 2018; 187:71–87. <https://doi.org/10.1016/j.pharmthera.2018.02.012> PMID: 29458109
104. Ha NT, Lee CH. Roles of farnesyl-diphosphate farnesyltransferase 1 in tumour and tumour microenvironments. *Cells.* 2020; 9(11):2352. <https://doi.org/10.3390/cells9112352> PMID: 33113804
105. Asperger H, Stamm N, Gierke B, Pawlak M, Hofmann U, Zanger UM, et al. Progesterone receptor membrane component 1 regulates lipid homeostasis and drives oncogenic signaling resulting in breast cancer progression. *Breast Cancer Res.* 2020; 22(1):75. <https://doi.org/10.1186/s13058-020-01312-8> PMID: 32660617
106. Ferreira S, Saraiva N, Rijo P, Fernandes AS. LOXL2 inhibitors and breast cancer progression. *Antioxidants (Basel).* 2021; 10(2):312. <https://doi.org/10.3390/antiox10020312> PMID: 33669630
107. Schuurs-Hoeijmakers JH, Vulto-van Silfhout AT, Vissers LE, van de Vondervoort II, van Bon BW, de Ligt J, et al. Identification of pathogenic gene variants in small families with intellectually disabled siblings by exome sequencing. *J Med Genet.* 2013; 50(12):802–11. <https://doi.org/10.1136/jmedgenet-2013-101644> PMID: 24123876
108. Doherty JA, Rossing MA, Cushing-Haugen KL, Chen C, Van Den Berg DJ, Wu AH, et al; Australian Ovarian Cancer Study Management Group; Australian Cancer Study (Ovarian Cancer), Pearce CL; Ovarian Cancer Association Consortium (OCAC). ESR1/SYNE1 polymorphism and invasive epithelial ovarian cancer risk: an Ovarian Cancer Association Consortium study. *Cancer Epidemiol Biomarkers Prev.* 2010; 19(1):245–50. <https://doi.org/10.1158/1055-9965.EPI-09-0729> PMID: 20056644
109. O'Mara TA, Glubb DM, Painter JN, Cheng T, Dennis J; Australian National Endometrial Cancer Study Group (ANECs), et al. Comprehensive genetic assessment of the ESR1 locus identifies a risk region for endometrial cancer. *Endocr Relat Cancer.* 2015; 22(5):851–61. <https://doi.org/10.1530/ERC-15-0319> PMID: 26330482
110. Sapkota Y, Steinhorsdottir V, Morris AP, Fassbender A, Rahmioglu N, De Vivo I, et al. Meta-analysis identifies five novel loci associated with endometriosis highlighting key genes involved in hormone metabolism. *Nat Commun.* 2017; 8:15539. <https://doi.org/10.1038/ncomms15539> PMID: 28537267
111. Antony J, Dasgupta T, Rhodes JM, McEwan MV, Print CG, O'Sullivan JM, et al. Cohesin modulates transcription of estrogen-responsive genes. *Biochim Biophys Acta.* 2015; 1849(3):257–69. <https://doi.org/10.1016/j.bbagr.2014.12.011> PMID: 25542856
112. Fachal L, Aschard H, Beesley J, Barnes DR, Allen J, et al. Fine-mapping of 150 breast cancer risk regions identifies 191 likely target genes. *Nat Genet.* 2020; 52(1):56–73. <https://doi.org/10.1038/s41588-019-0537-1> PMID: 31911677
113. Jones FS, Kioussi C, Copertino DW, Kallunki P, Holst BD, Edelman GM. Barx2, a new homeobox gene of the Bar class, is expressed in neural and craniofacial structures during development. *Proc Natl Acad Sci U S A.* 1997; 94(6):2632–7. <https://doi.org/10.1073/pnas.94.6.2632> PMID: 9122247
114. Montero JA, Hurlé JM. Deconstructing digit chondrogenesis. *Bioessays.* 2007; 29(8):725–37. <https://doi.org/10.1002/bies.20607> PMID: 17621635
115. Stevens TA, Meech R. BARX2 and estrogen receptor- α (ESR1) coordinately regulate the production of alternatively spliced ESR1 isoforms and control breast cancer cell growth and invasion. *Oncogene.* 2006; 25(39):5426–35. <https://doi.org/10.1038/sj.onc.1209529> PMID: 16636675
116. O'Donovan KJ, Tourtellotte WG, Millbrandt J, Baraban JM. The EGR family of transcription-regulatory factors: progress at the interface of molecular and systems neuroscience. *Trends Neurosci.* 1999; 22(4):167–73. [https://doi.org/10.1016/s0166-2236\(98\)01343-5](https://doi.org/10.1016/s0166-2236(98)01343-5) PMID: 10203854
117. Morita K, Okamura T, Sumitomo S, Iwasaki Y, Fujio K, Yamamoto K. Emerging roles of Egr2 and Egr3 in the control of systemic autoimmunity. *Rheumatology (Oxford).* 2016; 55(suppl 2):ii76–ii81. <https://doi.org/10.1093/rheumatology/kew342> PMID: 27856665

118. Drabovich AP, Pavlou MP, Schiza C, Diamandis EP. Dynamics of protein expression reveals primary targets and secondary messengers of estrogen receptor alpha signaling in MCF-7 breast cancer cells. *Mol Cell Proteomics*. 2016; 15(6):2093–107. <https://doi.org/10.1074/mcp.M115.057257> PMID: 27067054
119. Inoue A, Omoto Y, Yamaguchi Y, Kiyama R, Hayashi SI. Transcription factor EGR3 is involved in the estrogen-signaling pathway in breast cancer cells. *J Mol Endocrinol*. 2004; 32(3):649–61. <https://doi.org/10.1677/jme.0.0320649> PMID: 15171706
120. Liu M, Fan J, Wang S, Wang Z, Wang C, Zuo Z, et al. Transcriptional profiling of Chinese medicinal formula Si-Wu-Tang on breast cancer cells reveals phytoestrogenic activity. *BMC Complement Altern Med*. 2013; 13:11. <https://doi.org/10.1186/1472-6882-13-11> PMID: 23305139
121. Lentjes MH, Niessen HE, Akiyama Y, de Bruïne AP, Melotte V, van Engeland M. The emerging role of GATA transcription factors in development and disease. *Expert Rev Mol Med*. 2016; 18:e3. <https://doi.org/10.1017/erm.2016.2> PMID: 26953528
122. Välimäki MJ, Ruskoaho HJ. Targeting GATA4 for cardiac repair. *IUBMB Life*. 2020; 72(1):68–79. <https://doi.org/10.1002/iub.2150> PMID: 31419020
123. Wakana K, Akiyama Y, Aso T, Yuasa Y. Involvement of GATA-4/-5 transcription factors in ovarian carcinogenesis. *Cancer Lett*. 2006; 241(2):281–8. <https://doi.org/10.1016/j.canlet.2005.10.039> PMID: 16337738
124. Hellebrekers DM, Lentjes MH, van den Bosch SM, Melotte V, Wouters KA, Daenen KL, et al. GATA4 and GATA5 are potential tumor suppressors and biomarkers in colorectal cancer. *Clin Cancer Res*. 2009; 15(12):3990–7. <https://doi.org/10.1158/1078-0432.CCR-09-0055> PMID: 19509152
125. Miranda-Carboni GA, Guemes M, Bailey S, Anaya E, Corselli M, Peault B, et al. GATA4 regulates estrogen receptor- α -mediated osteoblast transcription. *Mol Endocrinol*. 2011; 25(7):1126–36. <https://doi.org/10.1210/me.2010-0463> PMID: 21566084
126. Güemes M, Garcia AJ, Rigueur D, Runke S, Wang W, Zhao G, et al. GATA4 is essential for bone mineralization via ER α and TGF β /BMP pathways. *J Bone Miner Res*. 2014; 29(12):2676–87. <https://doi.org/10.1002/jbmr.2296> PMID: 24932701
127. van der Heijden AG, Mengual L, Lozano JJ, Ingelmo-Torres M, Ribal MJ, Fernández PL, et al. A five-gene expression signature to predict progression in T1G3 bladder cancer. *Eur J Cancer*. 2016; 64:127–36. <https://doi.org/10.1016/j.ejca.2016.06.003> PMID: 27414486
128. Liang J, Zhao H, Hu J, Liu Y, Li Z. SPOCD1 promotes cell proliferation and inhibits cell apoptosis in human osteosarcoma. *Mol Med Rep*. 2018; 17(2):3218–3225. <https://doi.org/10.3892/mmr.2017.8263> PMID: 29257309
129. Liu D, Yang Y, Yan A, Yang Y. SPOCD1 accelerates ovarian cancer progression and inhibits cell apoptosis via the PI3K/AKT pathway. *Oncotargets Ther*. 2020; 13:351–359. <https://doi.org/10.2147/OTT.S200317> PMID: 32021280
130. Kulasekaran G, Chaineau M, Piscopo VEC, Verginelli F, Fotouhi M, Girard M, et al. An Arf/Rab cascade controls the growth and invasiveness of glioblastoma. *J Cell Biol*. 2021; 220(2):e202004229. <https://doi.org/10.1083/jcb.202004229> PMID: 33443570
131. Chiarella E, Aloisio A, Scicchitano S, Bond HM, Mesuraca M. Regulatory role of microRNAs targeting the transcription co-factor ZNF521 in normal tissues and cancers. *Int J Mol Sci*. 2021; 22(16):8461. <https://doi.org/10.3390/ijms22168461> PMID: 34445164
132. Bond HM, Scicchitano S, Chiarella E, Amodio N, Lucchino V, Aloisio A, et al. ZNF423: a new player in estrogen receptor-positive breast cancer. *Front Endocrinol (Lausanne)*. 2018; 9:255. <https://doi.org/10.3389/fendo.2018.00255> PMID: 29867779
133. Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoutte J, et al. Genome-wide analysis of estrogen receptor binding sites. *Nat Genet*. 2006; 38(11):1289–97. <https://doi.org/10.1038/ng1901> PMID: 17013392
134. Hah N, Kraus WL. Hormone-regulated transcriptomes: lessons learned from estrogen signaling pathways in breast cancer cells. *Mol Cell Endocrinol*. 2014; 382(1):652–664. <https://doi.org/10.1016/j.mce.2013.06.021> PMID: 23810978
135. Helzer KT, Szatkowski Ozers M, Meyer MB, Benkusky NA, Solodin N, Reese RM, et al. The phosphorylated estrogen receptor α (ER) cistrome identifies a subset of active enhancers enriched for direct ER-DNA binding and the transcription factor GRHL2. *Mol Cell Biol*. 2019; 39(3):e00417–18. <https://doi.org/10.1128/MCB.00417-18> PMID: 30455249
136. Flach KD, Zwart W. The first decade of estrogen receptor cistromics in breast cancer. *J Endocrinol*. 2016; 229(2):R43–56. <https://doi.org/10.1530/JOE-16-0003> PMID: 26906743
137. Yamaga R, Ikeda K, Horie-Inoue K, Ouchi Y, Suzuki Y, Inoue S. RNA sequencing of MCF-7 breast cancer cells identifies novel estrogen-responsive genes with functional estrogen receptor-binding sites

- in the vicinity of their transcription start sites. *Horm Cancer*. 2013; 4(4):222–32. <https://doi.org/10.1007/s12672-013-0140-3> PMID: 23526455
138. Ding M, Wang H, Chen J, Shen B, Xu Z. Identification and functional annotation of genome-wide ER-regulated genes in breast cancer based on ChIP-Seq data. *Comput Math Methods Med*. 2012; 2012:568950. <https://doi.org/10.1155/2012/568950> PMID: 23346221
 139. Zwart W, Theodorou V, Kok M, Canisius S, Linn S, Carroll JS. Oestrogen receptor-co-factor-chromatin specificity in the transcriptional regulation of breast cancer. *EMBO J*. 2011; 30(23):4764–76. <https://doi.org/10.1038/emboj.2011.368> PMID: 22002538
 140. Barutcu AR, Hong D, Lajoie BR, McCord RP, van Wijnen AJ, Lian JB, et al. RUNX1 contributes to higher-order chromatin organization and gene regulation in breast cancer cells. *Biochim Biophys Acta*. 2016; 1859(11):1389–1397. <https://doi.org/10.1016/j.bbagr.2016.08.003> PMID: 27514584
 141. Wu HY, Zheng P, Jiang G, Liu Y, Nephew KP, Huang TH, et al. A modulator based regulatory network for ER α signaling pathway. *BMC Genomics*. 2012; 13 (Suppl 6):S6. <https://doi.org/10.1186/1471-2164-13-S6-S6> PMID: 23134758
 142. Pathak BR, Breed AA, Deshmukh P, Mahale SD. Androgen receptor mediated epigenetic regulation of CRISP3 promoter in prostate cancer cells. *J Steroid Biochem Mol Biol*. 2018; 181:20–27. <https://doi.org/10.1016/j.jsbmb.2018.02.012> PMID: 29477539
 143. Díaz de la Guardia-Bolívar E, Barrios-Rodríguez R, Zwir I, Jiménez-Moleón JJ, Del Val C. Identification of novel prostate cancer genes in patients stratified by Gleason classification: Role of antitumoral genes. *Int J Cancer*. 2022; 151(2):255–264. <https://doi.org/10.1002/ijc.33988> PMID: 35234293
 144. Migliaccio A, Di Domenico M, Castoria G, Nanayakkara M, Lombardi M, de Falco A, et al. Steroid receptor regulation of epidermal growth factor signaling through Src in breast and prostate cancer cells: steroid antagonist action. *Cancer Res*. 2005; 65(22):10585–93. <https://doi.org/10.1158/0008-5472.CAN-05-0912> PMID: 16288052
 145. Susa T, Ikaga R, Kajitani T, Iizuka M, Okinaga H, Tamamori-Adachi M, et al. Wild-type and specific mutant androgen receptor mediates transcription via 17 β -estradiol in sex hormone-sensitive cancer cells. *J Cell Physiol*. 2015; 230(7):1594–606. <https://doi.org/10.1002/jcp.24906> PMID: 25536295
 146. Severson TM, Kim Y, Joosten SEP, Schuurman K, van der Groep P, Moelans CB, et al. Characterizing steroid hormone receptor chromatin binding landscapes in male and female breast cancer. *Nat Commun*. 2018; 9(1):482. <https://doi.org/10.1038/s41467-018-02856-2> PMID: 29396493
 147. Giovannelli P, Di Donato M, Galasso G, Di Zazzo E, Bilancio A, Migliaccio A. The androgen receptor in breast cancer. *Front Endocrinol (Lausanne)*. 2018; 9:492. <https://doi.org/10.3389/fendo.2018.00492> PMID: 30210453
 148. Bojcsuk D, Nagy G, Balint BL. Inducible super-enhancers are organized based on canonical signal-specific transcription factor binding elements. *Nucleic Acids Res*. 2017; 45(7):3693–3706. <https://doi.org/10.1093/nar/gkw1283> PMID: 27994038
 149. Ferrero G, Miano V, Beccuti M, Balbo G, De Bortoli M, Cordero F. Dissecting the genomic activity of a transcriptional regulator by the integrative analysis of omics data. *Sci Rep*. 2017; 7(1):8564. <https://doi.org/10.1038/s41598-017-08754-9> PMID: 28819152
 150. Welboren WJ, van Driel MA, Janssen-Megens EM, van Heeringen SJ, Sweep FC, Span PN, et al. ChIP-seq of ER α and RNA polymerase II defines genes differentially responding to ligands. *EMBO J*. 2009; 28(10):1418–28. <https://doi.org/10.1038/emboj.2009.88> PMID: 19339991
 151. Indukuri R, Jafferli MH, Song D, Damdimopoulos A, Hases L, Zhao C, et al. Genome-wide estrogen receptor β chromatin binding in human colon cancer cells reveals its tumor suppressor activity. *Int J Cancer*. 2021; 149(3):692–706. <https://doi.org/10.1002/ijc.33573> PMID: 33754337
 152. Oki S, Ohta T, Shioi G, Hatanaka H, Ogasawara O, Okuda Y, et al. ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. *EMBO Rep*. 2018; 19(12):e46255. <https://doi.org/10.15252/embr.201846255> PMID: 30413482