

Integrated bioinformatic approach of Balur therapy for breast cancer subtypes

Sherry Aristyani¹, Saraswati Subagjo¹, Tintrim Rahayu¹, Sutiman B. Sumitro^{1,2*}

¹Lembaga Penelitian Peluruhan Radikal Bebas, Malang, Indonesia

²Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Brawijaya, Malang, Indonesia

Abstract. Balur is a topical complementary alternative medicine that can be applied to breast cancer. Balur eliminates free radicals and heavy metals which it found enhance in cancer patients compared with healthy persons. Balur uses materials that potentially can halt cancer progression and development such as aspirin, coffee, moringa, and tobacco. This study was to describe the complexity of Balur in the biological system of breast cancer subtypes such as luminal A, luminal B, and HER-2, which had been found in high numbers in Indonesia, with an integrative bioinformatic approach. Active compounds found from literature studies were collected and screened by skin permeability variables. For breast cancer subtypes, differential expressed gene (DEG) identification was used to explore the target of Balur compounds. The Cytoscape software was used to construct the network of Balur compounds and target proteins of breast cancer subtypes. Gene Ontology, pathway, docking, hub gene, and survival analysis were applied to investigate the detailed mechanism of Balur to breast cancer in each subtype. The results showed that Balur compounds targeted respectively 68, 86, and 84 genes of luminal A, luminal B, and HER-2. From these target genes, 7 Balur compounds had higher binding energy than chemotherapy compounds when they docked with hub genes from each subtype. In conclusion, this study explained the pharmacological mechanisms of Balur therapy, which has great potential as a complementary therapy for breast cancer subtypes.

1 Introduction

Breast cancer is the most prevalent cancer in the world and it happens in every country. WHO noted that in the end 2020, more than 7.8 million women were diagnosed with breast cancer. In Indonesia, breast cancer is in the first position of cancer cases, with 44% incidence and 15.3% mortality then followed by cervix uteri and lung cancer in 2020. In addition, according to the breast cancer molecular subtypes, luminal and HER-2 were high in Indonesia [1,2]. Luminal subtype is grouped into luminal A which shows high expression of estrogen receptor (ER) and progesterone receptor (PR) and low expression of Ki-67, proliferation protein, and luminal B which is characterized by high expression of ERs and/or PRs, HER2 and Ki-67.

* Corresponding author: sutiman@ub.ac.id

Therefore, luminal B grows faster and more aggressively, with a higher tumour grade than luminal A [3,4]. The HER2 subtype reveals overexpression of HER2 gene and other genes associated with the HER2 pathway but does not present ER and PR hormone receptors. HER2 (ERBB2) is an oncogenic receptor tyrosine kinase that plays a main role in breast development [4,5].

Up to now, chemotherapy has been involved in the primary treatment for breast cancer, but unfortunately, it brings severe side effects and causes the patients to become worse and weak. Chemotherapy targets rapidly dividing cells, but it not only destroys cancer cells but also proliferates healthy cells, including digestive tract cells, reproductive system cells, bone marrow, and hair follicles. Therefore, chemotherapeutic treatments have a serious effect on the body involving mucositis, nausea and vomiting, alopecia, neutropenia, loss of appetite, peripheral neuropathy, thrombocytopenia, and anaemia [6].

In Indonesia, Balur therapy, a topical complementary treatment, is used to neutralize free radicals and reduce heavy metals in the body. The therapy involves the patient lying on a grounded copper table while being treated with a series of applications, including aspirin, moringa leaves, condensed tobacco smoke, and coffee. A previous study demonstrated that Balur therapy can improve the quality of life for patients with degenerative diseases [7]. As described in earlier studies, cancer cells generate excessive free radicals, leading to oxidative stress and cellular redox imbalance. This process increases oxidative DNA damage, a key factor in DNA mutation and cancer development. Additionally, elevated levels of heavy metals in cancer cells promote the formation of free radicals, which damage DNA, lipids, and proteins. Unlike chemotherapy, which can exacerbate free radicals and heavy metal levels, Balur therapy incorporates materials like coffee, moringa leaves, aspirin, and condensed tobacco smoke, which may help inhibit cancer development and progression [8-11]. This study employs an integrated bioinformatics approach to explore the complex relationships between Balur therapy and luminal or HER-2 breast cancer. This method also sheds light on the potential multi-compound and multi-target effects of Balur therapy through large dataset analysis. The workflow of the integrated bioinformatics methods is illustrated in Figure 1.

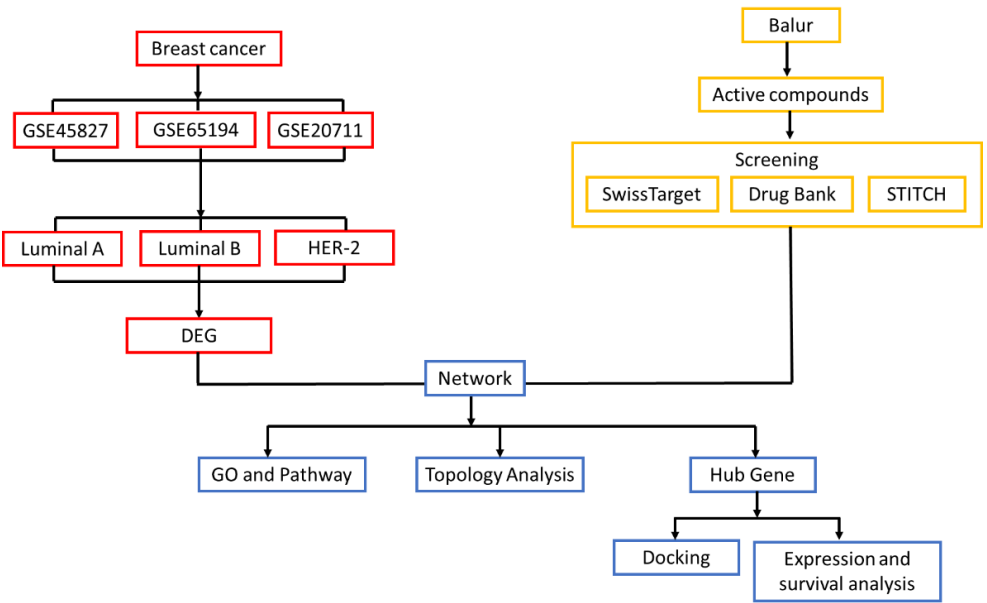


Fig. 1. The framework of Balur target prediction in breast cancer luminal and HER 2 subtypes.

2 Materials and methods

2.1 Microarray dataset and DEG identification

The expression gene data used was obtained from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>). Three datasets were chosen: GSE45827, GSE65194, and GSE20711. The further annotation was shown in Table 1. GEO2R tool (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>) was applied to identify and compare differentially expressed genes (DEG). For determining DEG, $P < 0.05$ and $|\log FC| \geq 2$ were used as the *adjusted P value* and cut off criteria. Venn diagram from FUNRICH software (Version 3.1.3) was applied to find overlapping DEG and overlapping DEG with Balur gene target.

Table 1. GEO datasets information used in this study.

Dataset	Platform	Number of Luminal A samples	Number of Luminal B samples	Number of Her2 samples	Number of control samples
GSE45827	GPL570	29	30	30	11
GSE65194	GPL570	29	30	39	11
GSE20711	GPL570	13	22	26	2

2.2 Balur compound collection, screening and skin permeation analysis

The annotation of Balur compounds was retrieved from several articles. Coffee bean’s active compounds were obtained from previous studies[12-14], moringa leaves active compounds were obtained from some scientific papers [15-19], and tobacco smoke’s active compounds were obtained from Rodgman and Perfetti (2009) [20] To screen the protein targets, web servers were used: Swiss TargetPrediction (<http://www.swisstargetprediction.ch/predict.php>) with probability was higher than 0.60, STITCH (<http://stitch.embl.de/>) with high confidence probability and DrugBank (<https://go.drugbank.com/>). As a topical therapy, SwissADME (<http://www.swissadme.ch/index.php>) was applied to predict the effectivity of active compound absorption through the skin.

2.3 Protein-Protein Interaction (PPI) and network analysis

The network interaction and analysis were constructed with Cytoscape (version 3.8.1) and its plugin features. Search Tool for the Retrieval of Interacting Genes (STRING) plugin was used to set up PPI with 0.7 confidence cut off score. Other plugins, Cytohubba with Maximal Clique Centrality (MCC) technique was selected to predict hub gens and Molecular Complex Detection (MCODE) was applied to identify key nodes and screen significant subnetwork of the PPI with the criteria: degree cutoff=2, node score cutoff=0.2, k-core=2 and maximum depth=100.

2.4 Gene Ontology (GO), KEGG pathway and Hub gene data analysis

For further analysis of gene target function, GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was conducted using WebGestalt (<http://www.webgestalt.org/>). GO term provides information of gene product function, which is classified into including biological process (BP), molecular function (MF), and cellular component (CC) (21). KEGG is a systemic analysis dataset that describes gene functions and

biological pathways [22]. Moreover, to portray gene expression profiling and hub genes’ the clinical data, webserver UALCAN (<http://ualcan.path.uab.edu/>).

2.5 Docking

To predict the accuracy of Balur active compounds on hub gene proteins of breast cancer, the specific docking analysis was led. PyRx software (version 0.9) was selected to assess binding affinity energy and Discovery studio 2016 and PyMOL software were applied to visualize and analyse the docking result. Chemotherapy chemicals such as: gemcitabine (CID: 60750) and doxorubicin (CID: 31703) were selected as controls in docking analysis. Protein prediction structures were taken from RCSB database <https://www.rcsb.org/>, including 1DEH for ADH1B, 3R5N for PPARG, 2HXQ for CHEK1, 2B9R for CCNB1, 3WGW for PCNA and Swiss-Model (<https://swissmodel.expasy.org/>), was used to create protein prediction if the prediction structures were not found in RCSB yet, such as LPL and CDK1.

3 Results and discussions

3.1 Active compound of Balur

We found Balur active compounds consisted of 52 active compounds of tobacco smoke, 18 active compounds of moringa, 12 active compounds of coffee bean and 1 active compound of aspirin. According to the SwissTarget, STITCH and DrugBank database, 83 active compounds targeted 611 proteins, respectively (Table 2). However, after it was matched with DEG breast cancer subtypes, only 198 proteins were targeted by 52 Balur active compounds which estimated skin absorption ability.

Table 2. The number of active compounds and target proteins of Balur.

Balur material	Total prediction		Breast cancer DEG integrated	
	Number of active compounds	Target Protein	Number of active compounds	Target Protein
Tobacco smoke	52	304	29	89
Moringa	18	196	14	74
Coffee bean	12	82	9	30
Aspirin	1	29	1	5
Total	83	611	52	198

Figure 2 showed that most of the compounds had less than 500 g/mol of molecular weight, 100 (Å²) of topological polar surface area (TPSA), 5.00 of lipophilicity and -2.0 of skin permeation (Log Kp). It described that Balur compounds were small molecules that easily pass the skin. Small molecules with less than 500 g/mol and low partition coefficient easily penetrated through epithelial and were commonly used as transdermal drug delivery [23]. TPSA was found as drug prediction pass the dermal barrier and it was also described corelation inversely of drug transport across the brain–blood barrier and intestinal membrane [24]. Furthermore, for skin permeation term demonstrated that the lower of skin permeation coefficient the higher ability molecule can through the skin.

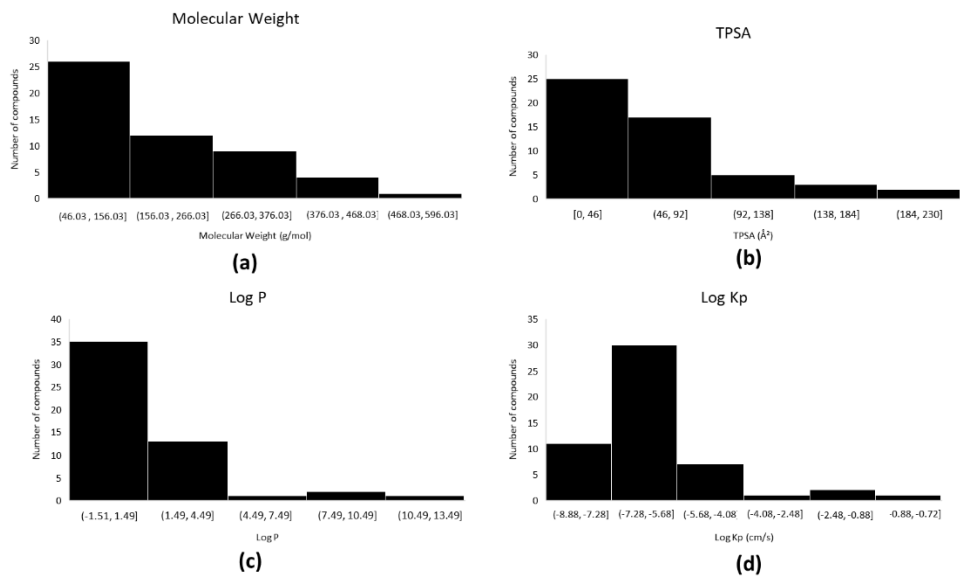


Fig. 2. Skin permeation prediction based on molecular weight (a) Molecular weight (b) topological polar surface area (TPSA) (c) lipophilicity (Log P) and (d) skin permeation (Log Kp).

3.2 Luminal A

In luminal A, there were 145 genes found in three datasets and 1237 genes found in two out of the datasets (Figure 3a). 6 genes were targeted with Balur compounds and overlapped with all datasets. Moreover, 68 genes belonged to two-three out of the dataset genes and Balur compounds target. The most intersection value was tobacco and luminal A containing 43 genes, and was followed by moringa, coffee and aspirin (Figure 3b).

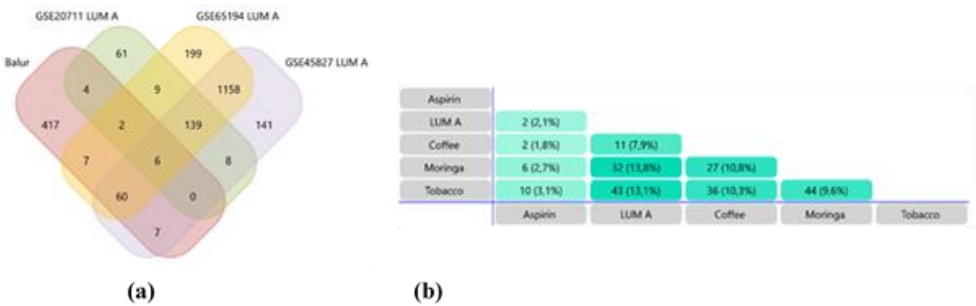


Fig. 3. (a) Luminal A venn diagram of GSE45827, GSE65194, and GSE20711 DEGs (logFC >2) identification and Balur compound targets. (b) Distribution of Balur compounds targets luminal A genes overlapping.

For further analysis, 68 genes which the intersection of DEG from two datasets and Balur, were selected to be constructed protein-protein interaction (PPI) network (Figure 4). It showed that 38 nodes and 39 edges with respectively 19 nodes upregulated and underregulated. According to the MCODE screen (Table 3), four significant cluster modules were found involving 1 cluster downregulated, 2 clusters upregulated and 1 cluster which

consist both of upregulated and downregulated. Top 3 hub genes, ADH1B, LPL, and PPARG, were identified, which all hub genes were downregulated.

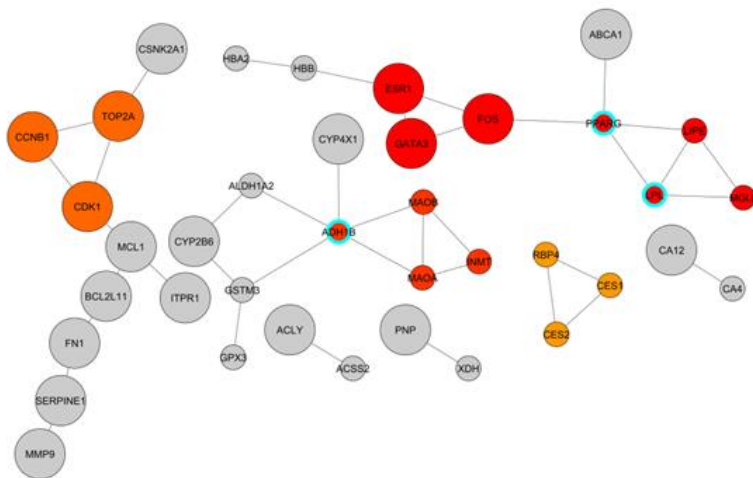


Fig. 4. Protein-protein interaction network of DEG luminal A-Balur. Red-orange nodes present the significant modules. Blue light line circles present hub gene. Big circles present upregulated and small circles present downregulated.

Table 3. MCODE clustering analysis

Cluster	Scores	Edges	Nodes
1	2.25	9	ESR1, GATA3, FOS, PPARG, LIPE, LPL, MGLL
2	2	5	INMT, MAOA, MAOB, ADH1B
3	1.5	3	TOP2A, CDK1, CCNB1
4	1.5	3	CES1, CES2, RBP4

According to GO enrichment analysis (Figure 5), top ten of BP like: developmental process, cell proliferation and growth. metabolic process, response to stimulus and biological regulation were included some cancer characteristics [25] For CC, membrane, endomembrane system, and cytosol were predominate. Protein and ion binding had high number for MF. In addition, nitrogen metabolism, drug metabolism, pathways in cancer, PPAR signaling pathway, tyrosine metabolism, tryptophan metabolism, phenylalanine metabolism, regulation of lipolysis in adipocytes and serotonergic synapse belonged to top ten KEGG pathways, which pathways in cancer was the highest number gene included and nitrogen metabolism was the most significant pathway (Table 4 and Figure 6). A study reported that nitrogen metabolism correlated with cancer development, free amino acids as a nitrogen form is upregulated in extracellular cells, whereas ammonia as nitrogen waste can be assimilated by cancer cells [26].

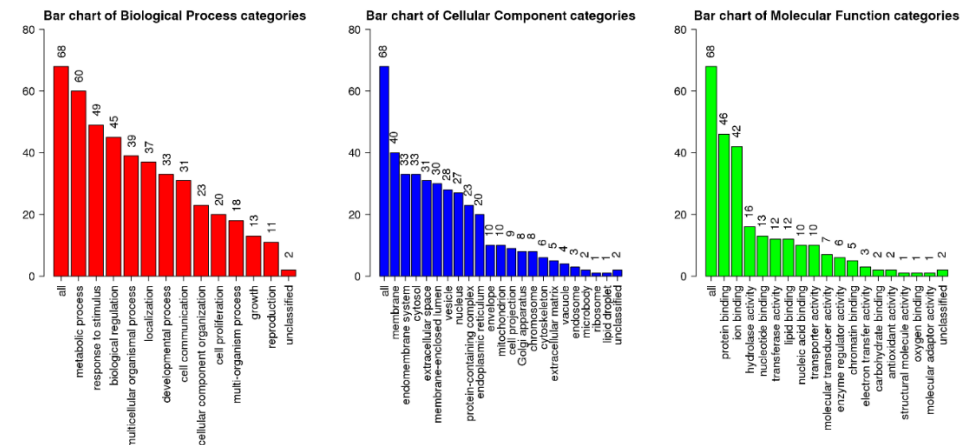


Fig. 5. GO enrichment analysis of 68 DEG luminal A-Balur

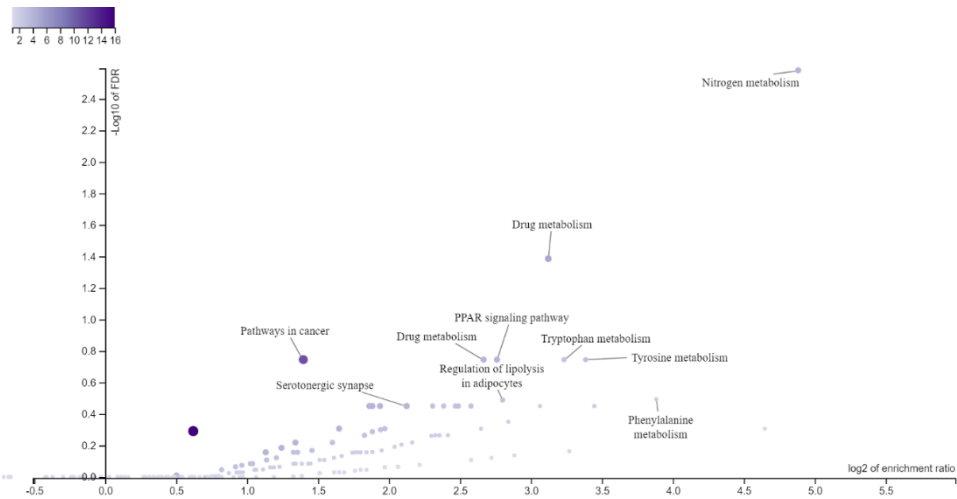


Fig. 6. KEGG enrichment of 68 DEG luminal A-Balur. Each purple point represented a pathway. The darker the color of a point, the more the overlap genes contained in the pathway.

Table 4. Top 10 KEGG pathway.

Gene Set	Description	Number of genes	Ratio	P-value	FDR
hsa00910	Nitrogen metabolism	4	29.497	0.0000080311	0.0026182
hsa00982	Drug metabolism	4	8.7057	0.00025217	0.041104
hsa05200	Pathways in cancer	11	2.6316	0.0024956	0.17979
hsa03320	PPAR signaling pathway	4	6.7763	0.0027696	0.17979
hsa00350	Tyrosine metabolism	3	10.447	0.0028535	0.17979
hsa00983	Drug metabolism	4	6.3474	0.0035115	0.17979
hsa00380	Tryptophan metabolism	3	9.4022	0.0038605	0.17979
hsa00360	Phenylalanine metabolism	2	14.748	0.0078759	0.32094
hsa04923	Regulation of lipolysis in adipocytes	3	6.9646	0.0089585	0.32450
hsa04726	Serotonergic synapse	4	4.3604	0.013077	0.35440

3.3 Luminal B

From three datasets, it was found 251 genes overlapped which 21 genes were targeted by Balur active compounds. In addition, 1548 genes were intersected with two of three datasets which 65 genes were overlapped with Balur genes (Figure 7a). For further analysis, 86 genes consisted of two-three out of three sets DEG and Balur active compounds target were selected. According to Figure 7b, tobacco had the most number overlapping genes with luminal B and continued by moringa, coffee and aspirin. In PPI network construction, 56 nodes and 96 edges were established involved 30 upregulated and 25 downregulated (Figure 8). MCODE screening result showed that 5 clusters were identified, which the most significant subnetwork consisted of three upregulated hub genes, CDK1, PCNA, and CCNB1(Figure 8, Table 5).

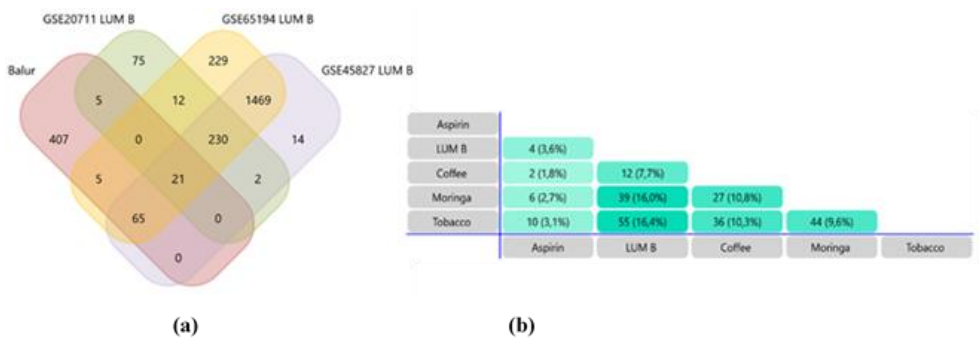


Fig. 7. (a) Luminal B venn diagram of GSE45827, GSE65194, and GSE20711 DEGs (logFC >2) identification and Balur compound targets. (b) Distribution of Balur compounds targets luminal B genes overlapping.

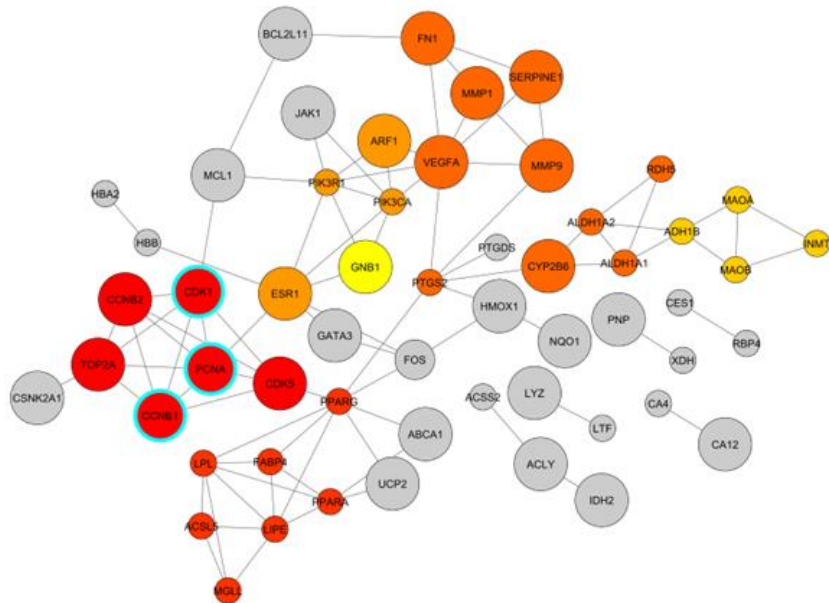


Table 5. MCODE clustering analysis.

Cluster	Scores	Edges	Nodes
1	4	14	TOP2A, CDK1, PCNA, CDK5, CCNB1, CCNB2
2	3.5	14	PPARA, PPARG, FABP4, LIPE, LPL, MGLL, ACSL5
3	2.909	16	VEGFA, MMP9, MMP1, FN1, SERPINE1, PTGS2, CYP2B6, ALDH1A1, RDH5, ALDH1A2
4	2.667	8	ESR1, ARF1, PIK3R1, PIK3CA, GNB1
5	2	5	MAOA, MAOB, ADH1B, INMT

GO analysing described that developmental process, cell proliferation, and growth belonged to top ten BP. Importantly, the hub genes identified such as were pertained to tumour cells proliferation and growth [27,28]. In CC term, membrane and cytosol had the number while protein and ion binding were dominantly in MF (Figure 9). Based on KEGG identification result, top ten pathways were PPAR signaling pathway, lipolysis in adipocytes regulation, pathways in cancer, relaxin signaling pathway, fluid shear stress and atherosclerosis, HIF-1 signaling pathway, retinol metabolism, nitrogen metabolism, tryptophan metabolism, and MicroRNAs in cancer. PPAR signaling pathway became the most significant pathway. PPAR signaling responsible to metabolic cell regulation and cell fate, but dysregulation PPAR signaling leads chronic diseases involving cancer, diabetes, and autoimmune [29]. Moreover, the most plentiful number of genes implicated was pathway in cancer (Table 6, Figure 10).

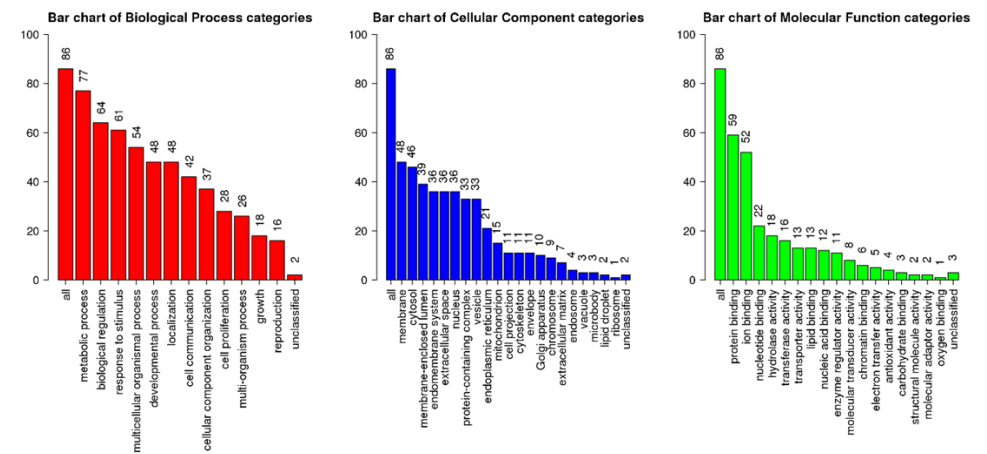


Fig. 9. GO enrichment analysis of 86 DEG luminal B-Balur.

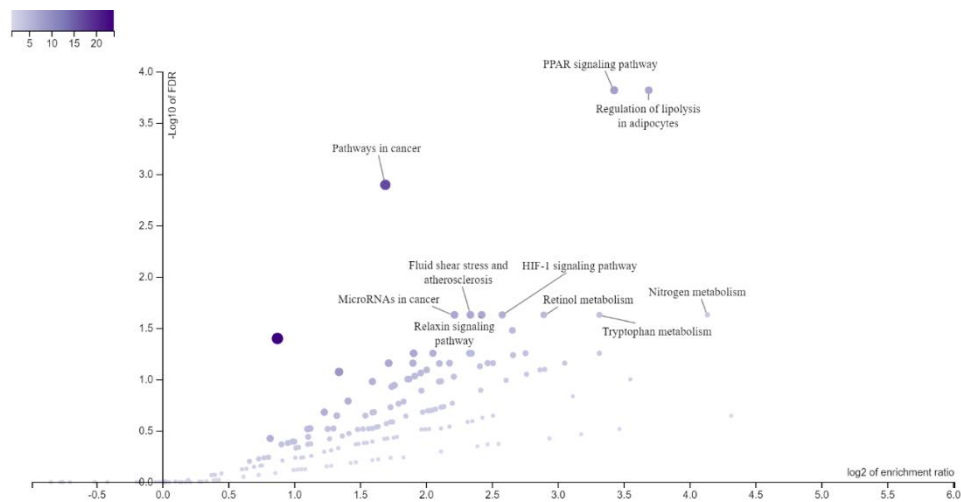


Fig.10. KEGG enrichment 86 luminal B-Balur DEG. Each purple point represented a pathway. The darker the color of a point, the more the number genes contained in the pathway.

Table 6. Top 10 KEGG pathway of 86 DEG luminal B-Balur

Gene Set	Description	Number of genes	Ratio	P Value	FDR
hsa03320	PPAR signaling pathway	8	10.768	6.1942e-7	0.00015207
hsa04923	Regulation of lipolysis in adipocytes	7	12.911	9.3297e-7	0.00015207
hsa05200	Pathways in cancer	17	3.2314	0.000011696	0.0012709
hsa04926	Relaxin signaling pathway	7	5.3632	0.00030364	0.023504
hsa05418	Fluid shear stress and atherosclerosis	7	5.0523	0.00043698	0.023504
hsa04066	HIF-1 signaling pathway	6	5.9762	0.00047082	0.023504
hsa00830	Retinol metabolism	5	7.4330	0.00052880	0.023504
hsa00910	Nitrogen metabolism	3	17.577	0.00059683	0.023504
hsa00380	Tryptophan metabolism	4	9.9603	0.00065031	0.023504
hsa05206	MicroRNAs in cancer	7	4.6481	0.00072098	0.023504

3.4 HER2

In term of HER2, 258 genes were appertained to three datasets which 15 were intersected with Balur active compound targets. From two of three datasets, 1692 were found and 69 genes were involved in Balur active compound targets (Figure 11 a). In addition, Figure 11b elucidated that tobacco, moringa, coffee, and aspirin had the high number of target distribution sequentially. Further analysis, two of three datasets and Balur target overlapping, 84 genes were selected.

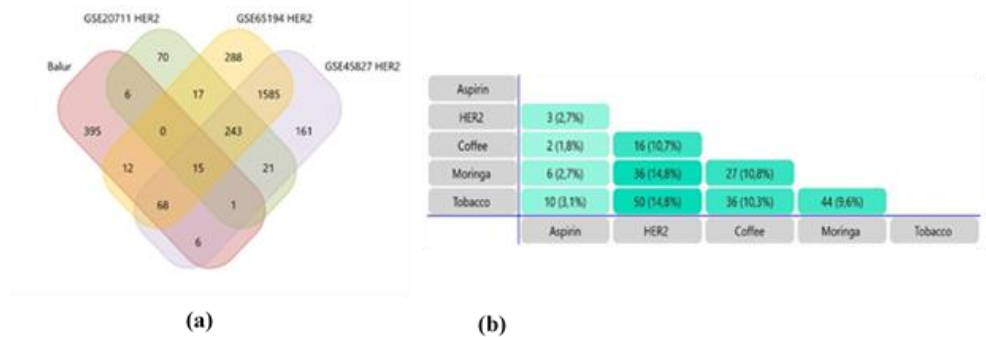


Fig. 11. (a) HER 2 venn diagram of GSE45827, GSE65194, and GSE20711 DEGs (logFC >2) identification and Balur compound targets. (b) Distribution of Balur compounds targets HER 2 genes overlapping.

Construction of 84 DEGS PPI network described which 54 nodes and 106 edges were connected and consisted of 32 expressed genes were upregulated otherwise 21 were downregulated (Figure 12). MCODE screening resulted six clusters involving 6 upregulated nodes and three hub genes, CDK1, PCNA, and CHEK1 as the highest score of significant subnetworks (Table 7).

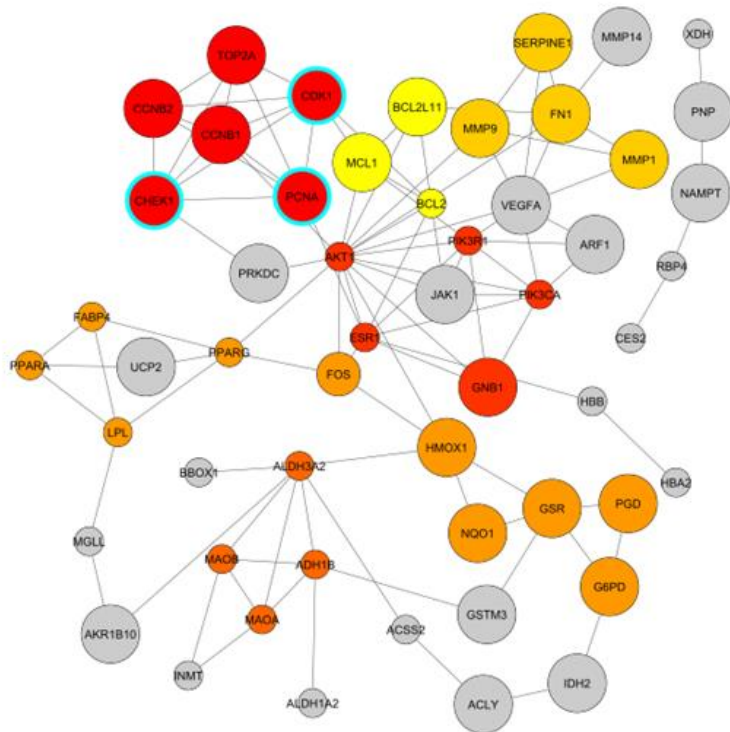


Fig. 12. Protein-protein interaction network of DEG HER 2-Balur. Red-orange nodes present the significant modules. Blue light line circles present hub gene. Big circles present upregulated and small circles present downregulated.

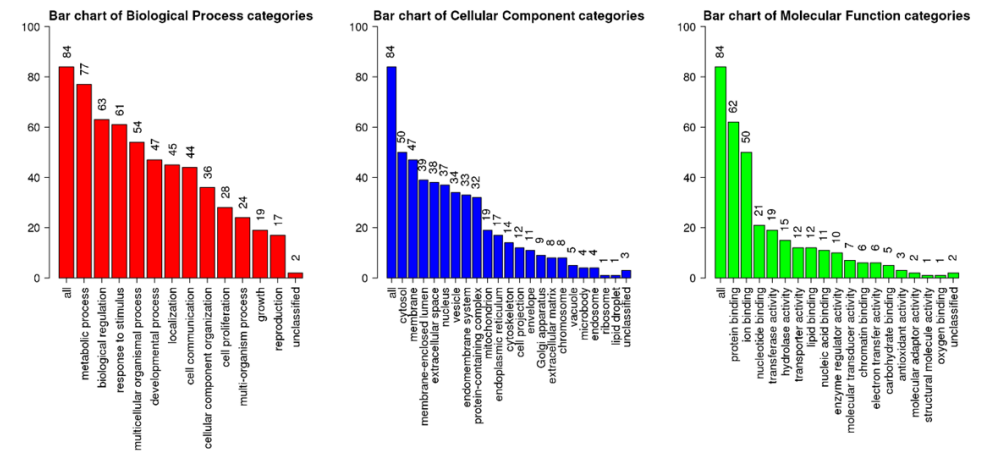


Fig. 13. GO enrichment analysis of 84 DEG HER 2-Balur.

Table 7. MCODE clustering analysis

Cluster	Scores	Edges	Nodes
1	4.286	15	TOP2A, PCNA, CCNB2, CDK1, CHEK1, CCNB1
2	3.333	10	AKT1, ESR1, GNB1, PIK3CA, PIK3R1
3	2.4	6	ADH1B, MAOA, MAOB, ALDH3A2
4	2.364	13	G6PD, PGD, GSR, NQO1, HMOX1, FOS, PPARG, LPL, FABP4, PPARA
5	1.6	4	MMP1, FN1, MMP9, SERPINE1
6	1.6	3	BCL2L11, BCL2, MCL1

Figure 13 informed that same as luminal A and B, in top ten of BP were included development process, cell proliferation, and growth which were linked to role of hub genes identified. As same as the previous results, luminal A and B, cytosol and membrane were primary in MF whereas protein and ion binding were dominantly in CC. Based on KEGG pathway (Table 8, Figure 14), pathway in cancer was the most significant and the highest gene involved from top ten pathway analysis and was followed by HIF-1 signalling pathway and fluid shear stress and atherosclerosis as a second significant pathway and number of genes involved. Both pathways were reported that had correlated with development and metastases of cancer [30, 31].

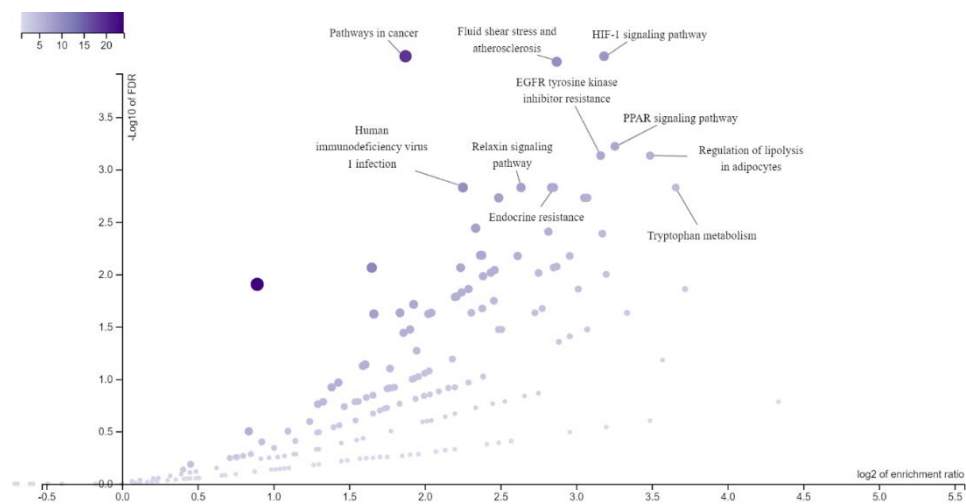


Fig. 14. KEGG enrichment 84 HER2-Balur DEG. Each purple point represented a pathway. The darker the color of a point, the more the number genes contained in the pathway.

Table 8. Top 10 KEGG pathway of 84 DEG HER2-Balur.

Gene Set	Description	Number of genes	Ratio	P Value	FDR
hsa05200	Pathways in cancer	19	3.6617	4.5952e-7	0.000082917
hsa04066	HIF-1 signaling pathway	9	9.0887	5.0870e-7	0.000082917
hsa05418	Fluid shear stress and atherosclerosis	10	7.3178	8.6122e-7	0.000093586
hsa03320	PPAR signaling pathway	7	9.5527	0.0000073666	0.00060038
hsa01521	EGFR tyrosine kinase inhibitor resistance	7	8.9481	0.000011409	0.00073590
hsa04923	Regulation of lipolysis in adipocytes	6	11.221	0.000013544	0.00073590
hsa04926	Relaxin signaling pathway	8	6.2145	0.000038550	0.0014841
hsa05170	Human immunodeficiency virus 1 infection	10	4.7635	0.000039668	0.0014841
hsa00380	Tryptophan metabolism	5	12.623	0.000041578	0.0014841
hsa01522	Endocrine resistance	7	7.2133	0.000046904	0.0014841

3.5 Hub gene analysis

According to UALCAN analysis, ADH1B, LPL, and PPARG were down expressed in various type of breast cancer, otherwise CCNB1, CDK1, PCNA, and CHEK1 were over expressed (Figure 15). All hub genes showed better prognosis average, except in HER2 subtype, CHEK1, CDK1, and CCNB1 had below 50% survival probability when in high express term (Figure 16).

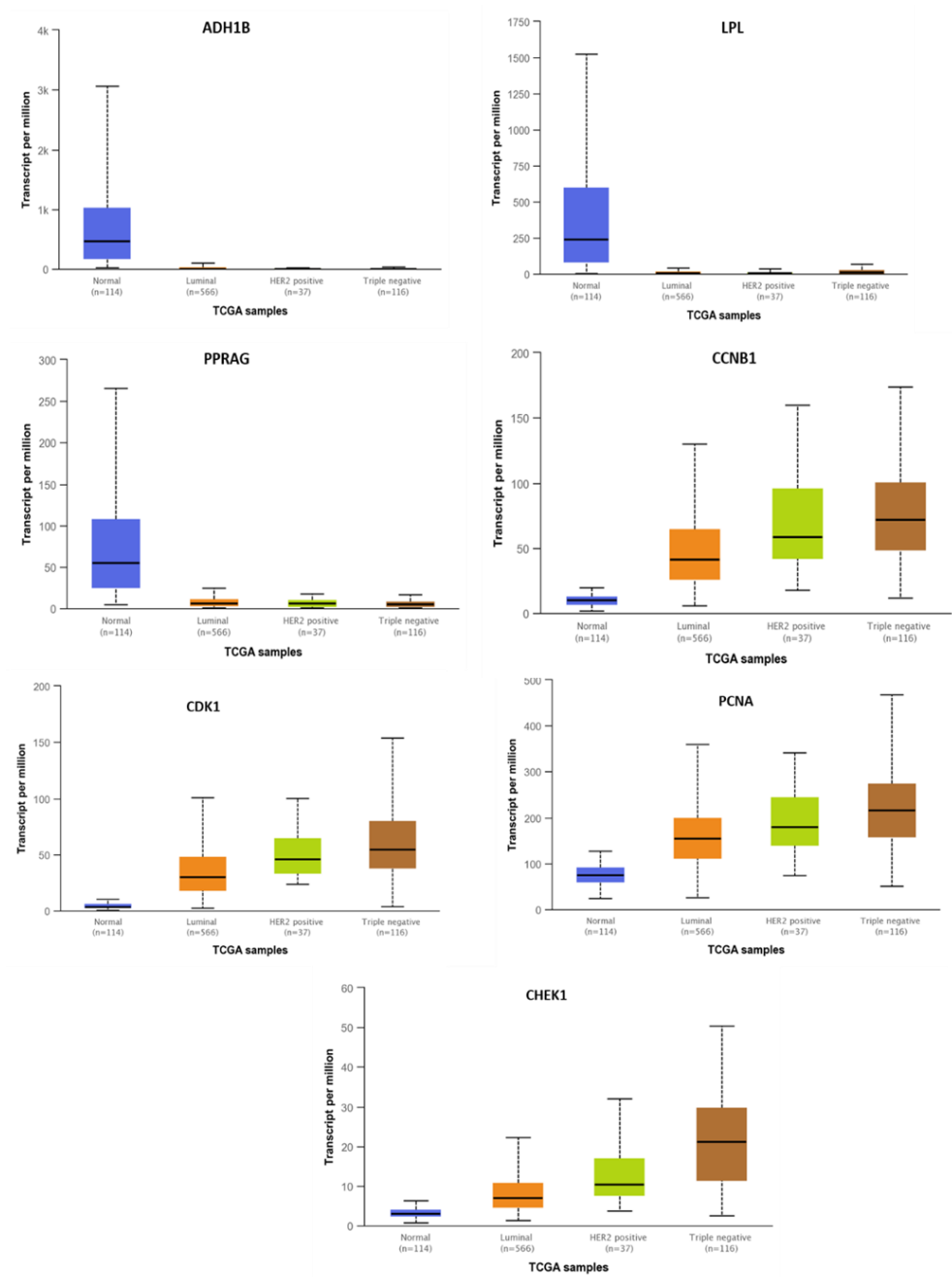


Fig. 15. Boxplot result of gene hubs expression by UALCAN.

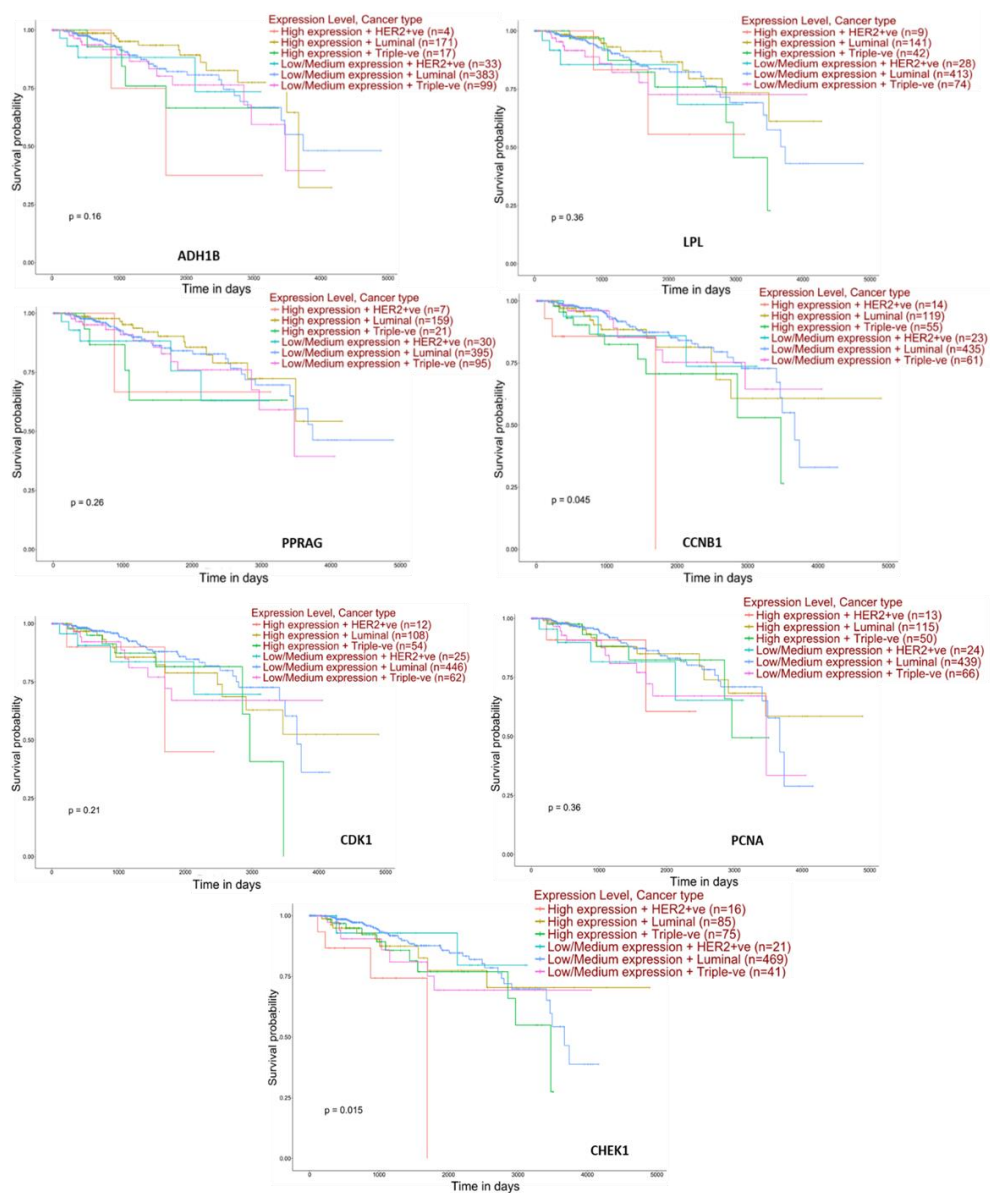


Fig. 16. Survival analysis of hub genes.

In addition, to predict the effectivity of Balur compound, specific docking of hub genes and Balur active compounds which targeting hub genes protein was applied, and the position of control, chemotherapy compounds (gemcitabine and doxorubicin) as the criterion to determine docking site. It was found that baluar active compounds like α -tocopherol, γ -tocopherol and apigenin from moringa had higher binding affinity than doxorubicin but it was not in gemcitabine, otherwise retinol, kaempferol, myricetin, and quercetin from moringa had higher score of binding affinity than gemcitabine but not in doxorubicin. However, apigenin from moringa had higher binding affinity than doxorubicin and gemcitabine when respectively bound with CCNB1 and CDK1 (Table 9).

Table 9. Binding affinity of hub gene proteins, Balur active compounds, and chemotherapy compounds.

Protein	Balur compound	Binding affinity (kcal/mol)	
		Gemcitabine	Doxorubicin
ADHB1	control	-6.7	-9.6
	Glycerol (tobacco)	-3.4	-3.7
	Acetone (tobacco)	-2.9	-2.9
	Retinol (moringa)	-7.2	-8.1
LPL	control	-5.9	-7.4
	Formic acid (tobacco)	-2.2	-2.6
PPARG	control	-6.1	-8.4
	α -tocopherol (moringa)	-5.7	-9.1
	Capric acid / decanoic acid (tobacco)	-4.5	-5.8
	Nonanoic acid (tobacco)	-4.4	-6.0
	γ -tocopherol (moringa)	-5.8	-9.2
PCNA	Control	-7.8	-5.7
	Aspirin	-5.1	-5.5
CCBN1	Control	-7.4	-6.8
	apigenin (moringa)	-3.4	-7.0
CHECK1	Control	-6.9	-9.4
	Caffeine (Coffee)	-5.9	-5.9
CDK1	Control	-6.5	-9.3
	Apigenin (moringa)	-8.7	-8.8
	Kaempferol (moringa)	-8.2	-8.3
	Myricetin (moringa)	-8.5	-8.5
	Quercetin (moringa)	-8.3	-8.3

3.6 Discussion

The complexity of natural medicine for healing diseases is too complicated to be explained, a large amount of active compounds provides massive possibilities happened. Therefore, integrative biological computation method is one of the methods that can be adopted to illustrate possibly perspective the mechanism of natural active compounds to supress the development of disease. In this study we revealed the prediction of the complexity Balur mechanism as breast cancer therapy through bioinformatics approach. We compared and made the network analysis of Balur therapy role in luminal A, luminal B and HER-2 breast cancer subtypes (Figure 17).

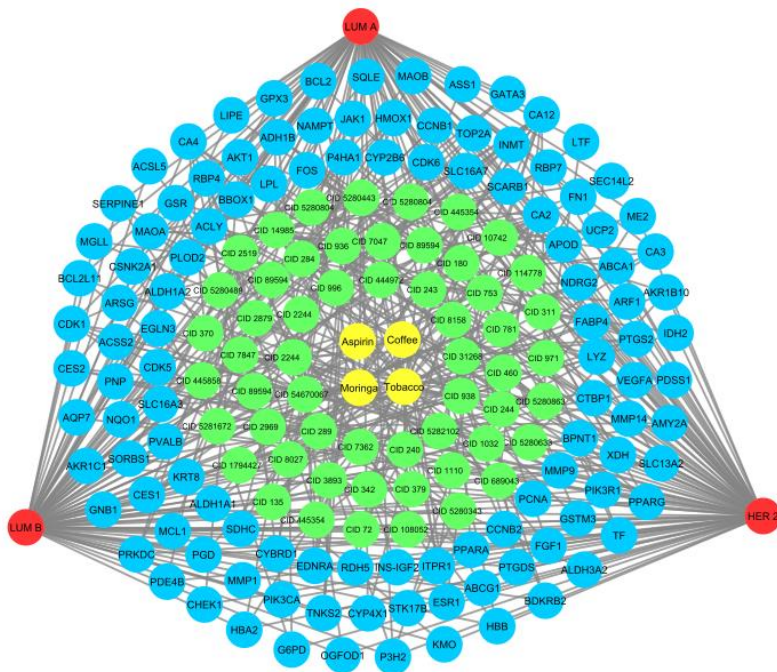


Fig. 17. The network of Balur compounds to luminal and HER-2 breast cancer subtype

Balur consists of a series of medicinal plants topological processes that has been used to treat breast cancer. Moringa (*Moringa olifera*) leaves, coffee (*Coffea arabica* L.) beans, and tobacco (*Nicotiana tabacum*) condensed fume leaves are the part of Balur material. All these materials have been reported that loaded anti-tumour promoting compounds and can be applied for treating breast cancer [32-34]. Additionally, Balur active compounds majority were not difficult to penetrate the skin, thus it might be easy to relieve the disease on direct and proper target. This is opposite with chemotherapy remaining unexpected target which kills all rapid proliferation rate cells without distinguish normal cells or not, thus becomes more severe [35].

According to the DEG result, luminal B had big number of genes targeted with Balur compounds and followed by HER-2 and luminal A. GO and pathways analysis results informed that all genes targeted in three breast cancer subtypes executed pathways in cancer and it was predicted through cancer development mechanism processes such as: development process, proliferation, and growth, which were explained by biological process result. A study described that breast cancer processes are frequently hijack the pathway of mammary remodelling development. The development process signalling pathways involve cellular basic process like proliferation, differentiation, and migration which took part a major role in transformation embryo organism and adult tissues and organs maintenance [36]. Therefore, dysregulation of these mechanisms triggers the emergence of cancer [37]. Rapid proliferation is the main part for cancer development and progression which lead by the unbalancing of expression and activity proteins related cell cycle regulation. Inadequacy of suppressor protein to control cell division brings cancer cells never stop to mitosis and become infinitely. Besides, cancer cells also manipulate telomerase to add telomeric section in the end of DNA during replication thus make cancer cells have longer live than normal cell [38]. On the other side, cancer cells lose to control growth factor expression, producing massive growth factor ligands by themselves and reducing growth factor suppressors and as the consequence it becomes unlimited growth [25].

In hub genes term, three main key genes targeted by Balur active compounds in each subtype were discovered. Luminal A had under expressed hub genes like: *PPARG*, *LPL*, and *ADH1B*. *PPAR γ* and *LPL* are elaborated in *PPAR* signaling pathway which played as lipid processing metabolism. *PPAR γ* as a nuclear receptor induce the transcription of *LPL*, the *LPL* enzyme is responsible to regulate fatty acids supply in various tissues, lipoprotein particles movement and lipid-lipoprotein shifting [39]. *PPAR γ* , a tumor suppressor, and *LPL* were found down regulated in breast cancer [29]. Nevertheless, both proteins are potentially as target for breast cancer therapy, *PPAR γ* leads autophagy by inducing *HIF1 α* and *BNIP3* [40], whereas *LPL* has ability as anti-inflammation by depleting inflammation proteins like *TNF α* and *IFN γ* and trigger apoptosis by inhibiting *COX-2* [41]. Moreover, activation of *ADH1B* increase acetaldehyde levels in blood inducing DNA damage and arising cancer. The polymorphism of *ADH1B* has been showed to be breast cancer risk [42,43].

In term of luminal B and HER-2 hub genes, both of subtypes had two similar genes, *CDK1* and *PCNA*. *CDK1* belongs to serine/threonine protein kinases family which regulating the alteration each phase of cell cycle. The combination of *CDK1* and its inhibitor, *p21*, hinder cell cycle into S phase and arrest G2/M transition. *p21/cyclin/CDK* complex attains to prevent interaction of proliferating cell nuclear antigen (*PCNA*) with DNA polymerase δ to stop DNA replication [44]. Therefore, the dysregulation of *CDK1* and *PCNA* potentially disturb the cell cycle mechanism and lead to unstoppable proliferation as a cancer progression. In line with previous studies, *CDK1* and *PCNA* were found to overexpressed in breast cancer [45]. Further, *PCNA* has been already to be a potential proliferation biomarker for breast cancer and *CDK1* has been realized to be breast cancer therapy [46] Moreover, the survival analysis described that the rise of *CDK1* was associated with short survival opportunity. The high expression of *CDK1* showed worse clinicopathologic features, such as high histologic grade, large lymph node metastases and tumor size [47].

CCNB1 and *CHEK1* respectively were identified as another Balur targeted hub genes of luminal B and HER-2. *CCNB1* (cyclin B1) a member of the cyclin family, took part as an initiator of cell cycle controlling and checkpoint. The complex of *CCNB1* and *CDK1* allows G2 to M phase transition. The dysregulation of cyclin B1 associate with breast cancer and has been utilized as prospective prediction breast cancer prognosis biomarker and survival [45,48]. Checkpoint kinase1 (*Chk1*), the enzyme of *CHEK1*, regulates G2/M switch such as *CDK1* and cyclin B1. *Chk1* is also take part in activation of DNA damage checkpoint proteins resulting cell cycles arrestment, DNA repairing and determining cell death or not. As the cancer cell has drawback DNA repairing mechanism, thus *Chk1* has been utilized as target of breast cancer therapy [49, 50] Besides, in colon cancer case, *Chk1* induces the activation of *CCNB1* and leads proliferation [27]. In addition, based on the survival analysis, high expression of *CCNB1* and *CHECK1* showed low survival opportunity in HER-2 subtypes, which was in line with other previous studies [51,52]. The elevation of *CCNB1* was correlated with poor prognosis of HER-2 subtypes: large tumor size, positive nodal status, and high grade. Therefore, the high *CCNB1* coincide with the aggressive phenotype even in patients who received adjuvant therapy after surgery [53]. In a high expression of *CHECK1* showed a worse prognosis and short survival rather than under expression [52].

Hereinafter, the docking technique was applied to explore Balur active compounds to the hub genes protein, and it was compared with chemotherapy compounds like gemcitabine and doxorubicin. Gemcitabine is a pyrimidine nucleoside antimetabolite compound that can inhibit DNA synthesis and cell proliferation. It has been approved to treat metastatic and advance breast cancer. It applies in second line therapy to overcome the first line drug resistance [54,55]. Nevertheless, the toxicity of gemcitabine lead to severe dyspnea, neutropenia (grade 3) and vomiting [54]. Balur compounds such as quercetin, apigenin, kaempferol, myricetin, and retinol targeting *CDK1* and *ADH1B* shown that had higher binding affinity score than gemcitabine in its binding position. In term of *CDK1*, flavonoid

compounds involving apigenin, kaempferol, quercetin and myricetin have been reported that trigger downregulation of CDK1 for arresting G2/M phase in cancer cells [56-58] Indeed, kaempferol and apigenin reduce kinase activity of CDK1 by 50-70% [56]. Besides, ADH1B acts in retinol metabolism. It has been found that ADHB1 is downregulated in cancer cells. Still, adding retinol could abolish cancer development by restraining glycolysis capacity, activating the AMPK pathway, and inhibiting the expression of ARG1 on these cells to improve the antitumor T-cell response [59]. However, the docking result also showed that apigenin targeting cyclin B1 and α -tocopherol had higher binding affinity than doxorubicin in its binding position. Doxorubicin, an anthracycline antibiotic, is widely used and claimed as one of the most effective compounds for treating early and advance breast cancer [60,61], but the contrary, it still causes suffered hematological abnormalities and acute renal injury [62,63]. Apigenin found had higher binding affinity than chemotherapy agent not only gemcitabine but also doxorubicin, which acted G2/M phase by downregulating cyclin B [64]. Moreover, previous studies explained that tocopherols suppress breast cancer development by increasing the PPAR γ expression and activating PPAR γ indirectly [65].

This study reveals the specific probabilities target compounds and mechanisms of Balur to breast cancer, yet it is still having vast limitation to understand because of the nature complexity of herbal compounds itself which is too complicated to be analysed and explained. Indeed, the complex mixture of herbal medicine have considerable effect rather than single compounds and lead to have multiple targets and activities. Therefore, analysing the single isolated compounds from herbal medicine emerges possible unknown or unstable active constituents.

4 Conclusions

By combining DEG analysis, pharmacological network, and docking, this study demonstrated the pharmacological mechanism of Balur therapy to breast cancer subtypes. This study described those multi-compounds of Balur might target 7 hub genes from luminal A, luminal B, and HER-2 with high binding energy. Although this study verified and predicted the molecular mechanism of Balur therapy, but it still needs further integrated investigation to validate the complexity of Balur.

References

1. Ng CH, Pathy NB, Taib NA, Teh YC, Mun KS, Amiruddin A, Evlina S, Rhodes A, Yip CH. Comparison of breast cancer in Indonesia and Malaysia--a clinico-pathological study between Dharmais Cancer Centre Jakarta and University Malaya Medical Centre, Kuala Lumpur. *Asian Pac J Cancer Prev.* **12**(11), 2943-6 (2011).
2. Paramita S, Raharjo EN, Niasari M, Azizah F, Hanifah NA. Luminal B is the Most common intrinsic molecular subtypes of invasive ductal breast carcinoma patients in East Kalimantan, Indonesia. *Asian Pac J Cancer Prev.* **20**(8), 2247-2252 (2019).
3. Le-Petross, H., J. Stafford, I. Bedrosian, P.B. Garvey, W.A. Woodward, and S.L. Moulder-Thompson: Chapter 27 - Breast Cancer; In: *Oncologic Imaging: A Multidisciplinary Approach*; edited by M.P. Silverman, USA: Elsevier. 497-507 (2012).
4. Moyal L. Navadyan, and B. Bonavida, Chapter 18 - New therapeutic modalities in breast cancer by targeting NK cell inhibitory and activating receptors; In: *Successes and Challenges of NK Immunotherapy*, edited by B. Bonavida and A. Jewett, USA: Elsevier. 387-402 (2021).

5. Yersal O, Barutca S., Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World J Clin Oncol.* **5**(3), 412-24 (2014).
6. Smoot B, Wampler M. Topp K.S. Breast Cancer Treatments and Complications: Implications for Rehabilitation. *Rehabil Oncol.* **27**, 16-26 (2019).
7. Jayanti GE, Subagjo S., Balur and improving quality of life. *The journal of Tropical Life Science.***6**, 35-40 (2016).
8. Australian Associated Press, Tobacco plant may light the way to beating cancer, Australian scientists say. <https://www.theguardian.com/science/2014/apr/02/tobacco-plant-may-light-the-way-to-beating-cancer-australian-scientists-say>. Accessed June 13, 2022
9. Bashir AIJ, Kankipati CS, Jones S, Newman RM, Safrany ST, Perry CJ, Nicholl ID, A novel mechanism for the anticancer activity of aspirin and salicylates. *Int J Oncol.* **54**, 1256-1270 (2019).
10. Nigra AD, de Almeida Bauer Guimarães D, Pucca CG, Freitas-Silva O, Teodoro AJ, Gil GA. Antitumor effects of freeze-dried robusta coffee (*Coffea canephora*) extracts on breast cancer cell lines. *Oxid Med Cell Longev.* **2021**, 5572630 (2021).
11. Kitzberger CSG, Scholz MBS, Benassi MDT, Bioactive compounds content in roasted coffee from traditional and modern *Coffea arabica* cultivars grown under the same edapho-climatic conditions. *Food Res Int.* **2024**, 61-64 (2024).
12. Affonso RC, Voytena AP, Fanan S, Pitz H, Coelho DS, Horstmann AL, Pereira A, Uarrota VG, Hillmann MC, Varela LA, Ribeiro-do-Valle RM, Maraschin M., Phytochemical composition, antioxidant activity, and the effect of the aqueous extract of coffee (*Coffea arabica* L.) bean residual press cake on the skin wound healing. *Oxid Med Cell Longev.* **2016**, 1923754 (2016).
13. Król K, Gantner M, Tatarak A, Hallmann E., The content of polyphenols in coffee beans as roasting, origin and storage effect. *Eur Food Res Technol.* **246**, 33–39 (2020).
14. Ferreira PMP, Farias DF, Olivera JTDA, Carvalho ADFU, *Moringa oleifera*: bioactive compounds and nutritional potential. *Rev de Nutr.* **21**, 431-437 (2018).
15. Muhammad AA, Pauzi NA, Arulselvan P, Abas F, Fakurazi S., In vitro wound healing potential and identification of bioactive compounds from *Moringa oleifera* Lam. *Biomed Res Int.* **2013**, 974580 (2013).
16. Vongsak B, Sithisarn P, Gritsanapan W., Simultaneous HPLC quantitative analysis of active compounds in leaves of *Moringa oleifera* Lam. *J Chromatogr Sci.* **52**, 641-5 (2014).
17. Valdez-Solana MA, Mejía-García VY, Téllez-Valencia A, García-Arenas G, Salas-Pacheco J, Alba-Romero JJ, Sierra-Campos E., Nutritional content and elemental and phytochemical analyses of *Moringa oleifera* grown in Mexico. *J Chem.* 860381 (2015).
18. Vergara-Jimenez M, Almatrafi MM, Fernandez ML., Bioactive components in *Moringa oleifera* leaves protect against chronic disease. *Antioxidants (Basel).* **6**, 91 (2017).
19. Rodgman A, Perfetti TA., The chemical components identified in tobacco and tobacco smoke prior to 1954: a chronology of classical chemistry. *Beitr. Tab. Int./ Contrib. Tob. Res.* **23**, 277-333 (2014).
20. Milano M., Gene prioritization tools in Shoba Ranganathan; In: *Encyclopedia of Bioinformatics and Computational Biology*, edited by Gribskov M., Nakai K., and Schönbach C., USA: Academic Press. 907-914 (2019).
21. Kanehisa M, Goto S, KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27-30 (2000).

22. Supe S, Takudage P., Methods for evaluating penetration of drug into the skin: A review. *Skin Res Technol.* **27**, 299-308 (2021).
23. Maciel Tabosa MA, Hoppel M, Bunge AL, Guy RH, Delgado-Charro MB., Predicting topical drug clearance from the skin. *Drug Deliv Transl Res.* **11**, 729-740 (2021).
24. Hanahan D, Weinberg RA, Hallmarks of cancer: the next generation. *Cell.* **144**, 646-74 (2011).
25. Kurmi K, Haigis MC., Nitrogen metabolism in cancer and immunity. *Trends Cell Biol.* **30**, 408-424 (2020).
26. Fang Y, Yu H, Liang X, Xu J, Cai X., Chk1-induced CCNB1 overexpression promotes cell proliferation and tumor growth in human colorectal cancer. *Cancer Biol Ther.* **15**, 1268-79 (2014).
27. Cardano M, Tribioli C, Prosperi E., Targeting proliferating cell nuclear antigen (PCNA) as an effective strategy to inhibit tumor cell proliferation. *Curr Cancer Drug Targets.* **20**, 240-252 (2020).
28. Cheng HS, Yip YS, Lim EKY, Wahli W, Tan NS., PPARs and tumor microenvironment: the emerging roles of the metabolic master regulators in tumor stromal-epithelial crosstalk and carcinogenesis. *Cancers (Basel).* **13**, 2153 (2021).
29. Semenza GL., Defining the role of hypoxia-inducible factor 1 in cancer biology and therapeutics. *Oncogene.* **29**, 625-34 (2010).
30. Huang Q, Hu X, He W, Zhao Y, Hao S, Wu Q, Li S, Zhang S, Shi M., Fluid shear stress and tumor metastasis. *Am J Cancer Res.* **8**, 763-777 (2018).
31. Levitsky DO, Dembitsky VM., Anti-breast cancer agents derived from plants. *Nat Prod Bioprospect.* **5**, 1–16 (2014).
32. Tiloke C, Anand K, Gengan RM, Chuturgoon AA., *Moringa oleifera* and their phytonanoparticles: Potential antiproliferative agents against cancer. *Biomed Pharmacother.* **108**, 457-466 (2018).
33. Saito Y, Takizawa H, Konishi S, Yoshida D, Mizusaki S., Identification of cembratriene-4,6-diol as antitumor-promoting agent from cigarette smoke condensate. *Carcinogenesis.* **6**, 1189-94 (1985).
34. Pearce A, Haas M, Viney R, Pearson SA, Haywood P, Brown C, Ward R., Incidence and severity of self-reported chemotherapy side effects in routine care: A prospective cohort study. *PLoS One.* **12**, e0184360 (2017).
35. Fantozzi A, Christofori G., Mouse models of breast cancer metastasis. *Breast Cancer Res.* **8**, 212 (2006).
36. Kumar V, Vashishta M, Kong L, Wu X, Lu JJ, Guha C, Dwarakanath BS., The role of notch, hedgehog, and Wnt signaling pathways in the resistance of tumors to anticancer therapies. *Front Cell Dev Biol.* **9**, 650772 (2021).
37. Guilloux G, Gibeaux R., Mechanisms of spindle assembly and size control. *Biol Cell.* **112**, 369-382 (2020).
38. Bozina T, Simić I, Lovrić J, Pećin I, Jelaković B, Sertić J, Reiner Z., Effects of lipoprotein lipase and peroxisome proliferator-activated receptor-gamma gene variants on metabolic syndrome traits. *Coll Antropol.* **37**, 801-8 (2013).
39. Zhou J, Zhang W, Liang B, Casimiro MC, Whitaker-Menezes D, Wang M, Lisanti MP, Lanza-Jacoby S, Pestell RG, Wang C., PPARgamma activation induces autophagy in breast cancer cells. *Int J Biochem Cell Biol.* **41**, 2334-42 (2009).
40. Takasu S, Mutoh M, Takahashi M, Nakagama H., Lipoprotein lipase as a candidate target for cancer prevention/therapy. *Biochem Res Int.* **2012**, 398697 (2012).

41. Kawase T, Matsuo K, Hiraki A, Suzuki T, Watanabe M, Iwata H, Tanaka H, Tajima K., Interaction of the effects of alcohol drinking and polymorphisms in alcohol-metabolizing enzymes on the risk of female breast cancer in Japan. *J Epidemiol.* **19**, 244-50 (2009).
42. Park B, Kim JH, Lee ES, Jung SY, Lee SY, Kang HS, Lee EG, Han JH., Role of aldehyde dehydrogenases, alcohol dehydrogenase 1B genotype, alcohol consumption, and their combination in breast cancer in East-Asian women. *Sci Rep.* **10**, 6564 (2020).
43. Karimian A, Ahmadi Y, Yousefi B., Multiple functions of p21 in cell cycle, apoptosis and transcriptional regulation after DNA damage. *DNA Repair (Amst).* **42**, 63-71 (2016).
44. Xing Z, Wang X, Liu J, Zhang M, Feng K, Wang X., Expression and prognostic value of CDK1, CCNA2, and CCNB1 gene clusters in human breast cancer. *J Int Med Res.* **49**, 300060520980647 (2021).
45. Izadi S, Nikkhoo A, Hojjat-Farsangi M, Namdar A, Azizi G, Mohammadi H, Yousefi M, Jadidi-Niaragh F., CDK1 in breast cancer: Implications for theranostic potential. *Anticancer Agents Med Chem.* **20**, 758-767 (2020).
46. Kim SJ, Nakayama S, Miyoshi Y, Taguchi T, Tamaki Y, Matsushima T, Torikoshi Y, Tanaka S, Yoshida T, Ishihara H, Noguchi S., Determination of the specific activity of CDK1 and CDK2 as a novel prognostic indicator for early breast cancer. *Ann Oncol.* **19**, 68-72 (2008).
47. Niméus-Malmström E, Koliadi A, Ahlin C, Holmqvist M, Holmberg L, Amini RM, Jirström K, Wärnberg F, Blomqvist C, Fernö M, Fjällskog ML., Cyclin B1 is a prognostic proliferation marker with a high reproducibility in a population-based lymph node negative breast cancer cohort. *Int J Cancer.* **127**, 961-7 (2010).
48. Patil M, Pabla N, Dong Z., Checkpoint kinase 1 in DNA damage response and cell cycle regulation. *Cell Mol Life Sci.* **70**, 4009-21 (2013).
49. Bryant C, Rawlinson R, Massey AJ., Chk1 inhibition as a novel therapeutic strategy for treating triple-negative breast and ovarian cancers. *BMC Cancer.* **14**, 570 (2014).
50. Aaltonen K, Amini RM, Heikkilä P, Aittomäki K, Tamminen A, Nevanlinna H, Blomqvist C., High cyclin B1 expression is associated with poor survival in breast cancer. *Br J Cancer.* **100**, 1055-60 (2009).
51. Al-Kaabi MM, Alshareeda AT, Jerjees DA, Muftah AA, Green AR, Alsubhi NH, Nolan CC, Chan S, Cornford E, Madhusudan S, Ellis IO, Rakha EA., Checkpoint kinase1 (CHK1) is an important biomarker in breast cancer having a role in chemotherapy response. *Br J Cancer.* **112**, 901-11 (2015).
52. Suzuki T, Urano T, Miki Y, Moriya T, Akahira J, Ishida T, Horie K, Inoue S, Sasano H., Nuclear cyclin B1 in human breast carcinoma as a potent prognostic factor. *Cancer Sci.* **98**, 644-51 (2007).
53. Toschi L, Finocchiaro G, Bartolini S, Gioia V, Cappuzzo F., Role of gemcitabine in cancer therapy. *Future Oncol.* **1**, 7-17 (2005).
54. Xie Z, Zhang Y, Jin C, Fu D., Gemcitabine-based chemotherapy as a viable option for treatment of advanced breast cancer patients: a meta-analysis and literature review. *Oncotarget.* **9**, 7148-7161 (2017).
55. Casagrande F, Darbon JM., Effects of structurally related flavonoids on cell cycle progression of human melanoma cells: regulation of cyclin-dependent kinases CDK2 and CDK1. *Biochem Pharmacol.* **61**, 1205-15 (2001).

56. Iizumi Y, Oishi M, Taniguchi T, Goi W, Sowa Y, Sakai T., The flavonoid apigenin downregulates CDK1 by directly targeting ribosomal protein S9. *PLoS One*. **8**, e73219 (2013).
57. Huang WW, Tsai SC, Peng SF, Lin MW, Chiang JH, Chiu YJ, Fushiya S, Tseng MT, Yang JS., Kaempferol induces autophagy through AMPK and AKT signaling molecules and causes G2/M arrest via downregulation of CDK1/cyclin B in SK-HEP-1 human hepatic cancer cells. *Int J Oncol*. **42**, 2069-77 (2013).
58. Sun HW, Chen J, Wu WC, Yang YY, Xu YT, Yu XJ, Chen HT, Wang Z, Wu XJ, Zheng L., Retinoic acid synthesis deficiency fosters the generation of polymorphonuclear myeloid-derived suppressor cells in colorectal cancer. *Cancer Immunol Res*. **9**, 20-33 (2021).
59. Zheng RR, Hu W, Sui CG, Ma N, Jiang YH., Effects of doxorubicin and gemcitabine on the induction of apoptosis in breast cancer cells. *Oncol Rep*. **32**, 2719-25 (2014).
60. Shafei A, El-Bakly W, Sobhy A, Wagdy O, Reda A, Aboelenin O, Marzouk A, El Habak K, Mostafa R, Ali MA, Ellithy M., A review on the efficacy and toxicity of different doxorubicin nanoparticles for targeted therapy in metastatic breast cancer. *Biomed Pharmacother*. **95**, 1209-1218 (2017).
61. Yemm KE, Alwan LM, Malik AB, Salazar LG., Renal toxicity with liposomal doxorubicin in metastatic breast cancer. *J Oncol Pharm Pract*. **25**, 1738-1742 (2019).
62. Gadisa DA, Assefa M, Wang SH, Yimer G., Toxicity profile of Doxorubicin-Cyclophosphamide and Doxorubicin-Cyclophosphamide followed by Paclitaxel regimen and its associated factors among women with breast cancer in Ethiopia: A prospective cohort study. *J Oncol Pharm Pract*. **26**, 1912-1920 (2020).
63. Choi EJ, Kim GH., Apigenin causes G(2)/M arrest associated with the modulation of p21(Cip1) and Cdc2 and activates p53-dependent apoptosis pathway in human breast cancer SK-BR-3 cells. *J Nutr Biochem*. **20**, 285-90 (2009).
64. Smolarek AK, Suh N., Chemopreventive activity of vitamin E in breast cancer: a focus on γ - and δ -tocopherol. *Nutrients*. **3**, 962-86 (2011).
65. Carmona, F., and A.M.S. Pereira: Herbal medicines: old and new concepts, truths and misunderstandings. *Rev Bras Farmacogn*. **23**, 379-385 (2013).