Identification of death-associated protein-like 1 (DAPL1) as a novel prognostic biomarker of breast cancer

CURRENT STATUS: UNDER REVIEW

Haichao Zhang
Tianjin Fourth Central Hospital
Email: zhaic2020@163.com
ORCID: https://orcid.org/0000-0003-0154-2947

Xin Qu
Tianjin Fourth Central Hospital

Lu Han
Tianjin Fourth Central Hospital

DOI:
10.21203/rs.2.24569/v1

SUBJECT AREAS
Pathology  Cancer Biology

KEYWORDS
DAPL1, breast cancer, prognosis, DNA repair, Ubiquitin
Abstract
Background: It is meaningful to identify the potential clinical prognosis-associated oncogenes for cases with breast cancer, considering the complicated pathogenesis of breast cancer.

Methods: We first utilized the bioinformatics approach to investigate the role of the DAPL1 (death-associated protein-like 1) in breast cancer, based on the available datasets of TCGA and GEO.

Results: DAPL1 is lowly expressed in breast cancer tissues compared with the normal tissues. For the breast cancer cases of the TCGA-BRCA cohort, we observed a correlation between lowly expressed DAPL1 gene and poor clinical prognosis of overall survival (P =0.0028). Based on the survival data of GEO, the low DAPL1 expression was associated with a poor prognosis of distant metastasis free survival (P =0.0023), and relapse free survival (P =0.0065). DAPL1 expression was linked to the mutation status or copy number variation of several genes, such as MAP3K1, NUP98, and CCDC59. The infiltration level of immune cells (e.g., M1 macrophage, Follicular B helper T cells, etc.) may be involved in the etiology of breast cancer. Based on the DAPL1-correlated genes, GSEA, GO, and KEGG analysis data indicated the association between DAPL1 expression and a series of biological issues, such as DNA packaging complex, DNA repair complex, nucleotide excision repair, ubiquitin-like protein binding, and ubiquitin proteasome pathway. We also identified several DAPL1-associated phosphorylation kinases, such as MAPK, PRKACA, and GSK3B.

Conclusions: DAPL1 gene is first identified as a prognosis biomarker of breast cancer, and the underlying molecular mechanism involves protein phosphorylation, immune cell infiltration, and DNA repair or protein ubiquitin-associated cellular pathways.

Background
Breast cancer is the most frequently diagnosed heterogeneous tumor in women [1]. There are various clinical types of breast invasive carcinoma, according to the different histopathological characteristics [1]. Considering the complicated pathogenesis of breast cancer, it is meaningful to explore the molecular mechanisms of breast cancer-related oncogenes or epigenetic biomarkers. In this study, we focused on the potential mechanism of the death-associated protein-like 1 (DAPL1) gene in the prognosis or pathogenesis of breast cancer.
The death-associated protein-like 1 (DAPL1) protein, encoded by DAPL1 gene, was reported to be associated with the pathogenesis of age-related macular degeneration (AMD) [2] and the regulation of cell proliferation in the retinal pigment epithelium (RPE) [3, 4]. A whole genome expression profiling data suggested that DAPL1 gene may be linked to the chewing-tobacco-associated oral cancer [5]. However, there are still no reports investigating the potential effect of the DAPL1 gene in the prognosis of human breast cancer diseases. As a public funded project, TCGA (The Cancer Genome Atlas) archives the multiple-genomics data from more than thirteen types of cancer, including expression level, genetic alteration, genome methylation, and clinical information, etc. [6, 7]. In addition, GEO (Gene Expression Omnibus) molecular data sets also provide a large number of clinical cancer-related gene expression data [8, 9]. These available datasets help to identify the prognosis-associated oncogenes for breast cancer cases [10].

In the present study, we first identified that the DAPL1 is lowly expressed in breast cancer, mainly based on the data of TCGA and GEO. Also, we explored the possible molecular mechanisms of DAPL1 in breast carcinogenesis from the perspectives of gene expression, genetic alteration, DNA methylation, immune cell infiltration, and enrichment analysis of DAPL1-correlated genes.

Methods

Expression analysis

We first obtained the basic information of the DAPL1 gene through Genecards (https://www.genecards.org/) and analyzed the by the Ensembl (http://asia.ensembl.org/index.html). Then, we analyzed the expression profile of the human DAPL1 gene in the breast cancer tissues and corresponding control tissues in the TCGA-BRCA (Breast invasive carcinoma) cohort by the GEPIA 2 [http://gepia2.cancer-pku.cn/#analysis] [11]. Besides, we investigated the expression difference of the DAPL1 gene between normal tissue and breast invasive lobular carcinoma tissues in the dataset of Radvanyi Breast [12], through Oncomine databases [https://www.oncomine.org/resource/login.html]. In addition, we utilized the LinkedOmics tool [13] to analyze the association between DAPL1 expression and a series of clinical factors, including the intrinsic subtype, pathological stage, ER (estrogen receptor) status, HER2 (human epidermal growth
factor receptor-2) status, PR (progesterone receptor) status, race, ethnicity, and radiation therapy [13]. Immunohistochemistry analysis data on the expression of DAPL1 in normal tissue and breast duct or lobular carcinoma tissue were obtained from the online HPA (Human Protein Atlas) database (https://www.proteinatlas.org/pathology).

**Survival curve analysis**

Based on the data of the TCGA-BRCA cohort, we performed the OS (overall survival) analysis by the UALCAN web-portal (http://ualcan.path.uab.edu/) [14]. The factors of cancer type, race, menopause status were included as well for the survival analyses. We utilized the Kaplan-Meier plotter (http://kmplot.com/analysis/) to perform the OS, DMFS (distant metastasis free survival), RFS (relapse free survival), PPS (post progression survival) by the expression level of the DAPL1 gene in a total of 1809 breast cancer cases within the Affymetrix HGU133A and HGU133+2 microarrays of GEO [15]. The auto-select best cutoff was set. The clinical factors, including ER status, PR status, HER2 status, TP53 (Tumor Protein P53) status, grade, intrinsic subtype, lymph node status, and pietenpol subtype, were also considered in the subgroup analysis.

**DNA methylation analysis**

We analyzed the DNA methylation status of DAPL1 in the breast cancer cases of the TCGA-BRCA cohort through the MEXPRESS [16, 17]. We also utilized the UALCAN web-portal (http://ualcan.path.uab.edu/) [14] to analyze the association between the gene expression and promoter methylation of DAPL1 between the normal and primary breast cancer tissues. The factors of pathological stage and intrinsic subtype were considered. In addition, the waterfall plot of the methylation level of DAPL1 gene was obtained through the MethSurv database (https://biit.cs.ut.ee/methsurv/) [18].

**Genetic alteration analysis**

Based on the cBioPortal database (https://www.cbioportal.org/) [19, 20], we analyzed the genetic alteration features of DAPL1 in the TCGA-BRCA cohort, respectively. The overall survival and disease/progression-free survival analyses were conducted by the genetic alteration of DAPL1 in the total cancer cases of the TCGA project. In addition, we performed the Wilcoxon test or Spearman test
to analyze the correlation between the \textit{DAPL1} expression level and the mutation or copy number variation (CNV) status of genes for the breast cancer cases of TCGA-BRCA, through the LinkedOmics (https://www.biostars.org/p/287820/) [18].

\textbf{Immune cell infiltration analysis}

We utilized the GEPIA 2 approach to conduct the pair-wise gene correlation analysis between \textit{DAPL1} expression and the signatures of the following immune cells: Monocytes, M1/2 Macrophage, Tumor-associated macrophages (TAMs), Natural killer cell (NK cell), Dendritic cell, B cell, Follicular B helper T cells (Tfh), naive T cell, effector T cell, effector memory T cell, central memory T cell, resident memory T cell, exhausted T cell, resting Treg T cell, effector Treg T-cell.

\textbf{\textit{DAPL1}-correlated gene enrichment analysis}

We performed the cluster analysis of the \textit{DAPL1}-correlated significant genes, through LinkedOmics. The heat map and association plot targeting the \textit{DAPL1} positively or negatively correlated significant genes were provided, respectively. We also analyzed the expression correlation between \textit{DAPL1} and the selected genes, through the cBioPortal database. In addition the data of GO (Gene ontology) enrichment analysis, KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis and GSEA (Gene Set Enrichment Analysis) profiles were provided, respectively.

\textbf{\textit{DAPL1}-correlated kinase analysis}

We first utilized the PhosphoSitePlus (https://www.phosphosite.org/homeAction.action) to predict the potential phosphorylation sites of DAPL1. Then, the features of these phosphorylation sites (e.g. maximum kinase specificity, hydrophobicity, or P-site similarity score, etc.) was analyzed by the PhosphoNET tool (http://www.phosphonetwork.ca/). The potential kinases were measured by the KinasePhos tool [21]. In addition, we also performed an enrichment analysis of kinase pathway, based on the \textit{DAPL1}-correlated genes.

\textbf{Result}

\textbf{Expression analysis data}

The human \textit{DAPL1} gene is located on chromosome 2q24.1 (Fig. 1). Fig. S1 shows the evolutionary tree of the \textit{DAPL1} gene in different species. Our study aims at exploring the functional links of the
human **DAPL1** gene with breast cancer from different perspectives. We first performed gene expression analysis. As shown in Fig. 1B, based on the dataset of the TCGA-BRCA cohort, we observed a lower expression level of **DAPL1** in the tissues of breast invasive carcinoma \((P<0.01)\), compared with the normal tissues. A similar result was detected in the **DAPL1** expression analysis of Radvanyi breast dataset \((P<0.01)\). The immunohistochemical result further showed that the expression level of DAPL1 protein was higher in the normal tissues than that in the tissues of breast duct carcinoma and lobular carcinoma \((P<0.01)\). Besides, the **DAPL1** gene showed the different expression levels among different intrinsic subtypes and pathologic stages \((P<0.01)\). Nevertheless, there is no significant statistical correlation between **DAPL1** expression and some other factors of ER status, HER2 status, PR status, race, ethnicity, and radiation therapy \((P<0.01)\). Therefore, the above suggested the potential role of the **DAPL1** gene in the etiology of breast cancer.

**Survival curve analysis data**

Subsequently, we performed a correlation analysis between **DAPL1** gene expression and the clinical overall survival outcome of breast cancer patients in TCGA-BRCA cohort. As shown in Fig. 2A, we observed a correlation between high expression of the **DAPL1** gene and better clinical prognosis of overall survival \((P=0.0028)\). When adjusted by the factors of cancer type \((P=0.0043)\) or menopause status \((P=0.0014)\), positive statistical results were detected as well \((P=0.0014)\). Meanwhile, based on the survival data in the GEO database, high **DAPL1** expression was associated with a better prognosis of distant metastasis free survival \((DMFS, P=0.0023)\), relapse free survival \((RFS, P=0.0065)\), but a worse prognosis of post progression survival \((PPS, P =0.023)\). In addition, we further conducted a series of subgroup analyses by the different clinical factors and observed distinct experimental results. For instance, the lowly expressed **DAPL1** is associated with the worse prognosis of RFS \((P=0.04, HR=0.72)\) and DMFS \((P=0.011, HR=0.48)\), but a better prognosis of OS \((P= 0.0042, HR=2.84)\) and PPS \((P= 0.0016, HR=2.87)\) for the breast cancer patients with pathological grade 3. These results provide evidence regarding the association between the **DAPL1** expression and clinical outcomes of breast cancer.

**DNA methylation analysis data**
We attempted to exploit the potential molecular mechanism from the point of *DAPL1* DNA methylation based on the methylation data of TCGA-BRCA. Fig. 3A presents the waterfall plot of *DAPL1* DNA methylation levels for each breast cancer case. We did not observe the statistically significant difference of the promoter methylation levels of *DAPL1* between the normal and tumor tissues, or among the different pathological analyses and tumor types (Fig. 3B, all *P* > 0.05). We also observed a similar negative result using the MEXPRESS method (Fig. S2). Thus, DNA methylation is not involved in the molecular mechanism of *DAPL1* in the pathogenesis of breast cancer.

**Genetic alteration analysis data**

Based on the TCGA-BRCA data, we explore the possible molecular mechanisms from the perspective of genetic alteration. As shown in Fig. 4A, we only observed a genetic alteration frequency of ~ 0.5% for the breast ductal carcinoma cases of TCGA. We also did not observe a correlation between *DAPL1* genetic alteration and the prognosis of overall survival (Fig. 4B, *P* = 0.992), disease/pression free survival (Fig. 4C, *P* = 0.317) for the total cancer patients. We then analyzed the correlation between *DAPL1* gene expression and different mutation status of specific genes. As shown in Fig. 4D, compared with the wildtype status, the *DAPL1* gene was lowly expressed in the mutated genes of *MAP3K1*, *NUP98*, *CEP152M*, and *ZNF300* (*P* < 0.001), whereas it was highly expressed in the mutated *KIF2IB* gene (*P* = 0.003). Furthermore, we observed that the expression of *DAPL1* gene was negatively correlated with the copy number variation (CNV) of the *CCDC59* gene (Fig. 4E, *R* = -0.18, *P* < 0.0001).

**Immune cell infiltration analysis data**

Also, we aimed to investigate whether the *DAPL1* gene is involved in the etiology of breast cancer through immune cell infiltration. As shown in Table 2, the expression of *DAPL1* gene in breast cancer was positively correlated with the infiltration level of immune cells: M1 macrophage, NK cell, B cell, Tfh, naive T cell, effector T cell, effector memory T cell, central memory T cell, resident memory T cell (all *R* > 0, *P* < 0.05), but negatively correlated with the infiltration level of resting or effect Treg T cell (all *R* < 0, *P* < 0.05).

**Enrichment analysis data**

Based on the data of TCGA-BRCA cohort, we utilized the LinkedOmics approach to screen out a group
of DAPL1 expression-correlated positively genes (e.g., MRRP2, KRT5, and LGALS7, etc.) and negatively related genes (e.g., TRPS1, ST3GAL1, ARFGEF1, TPD52, etc.) in Fig. 5A (heat map) and Fig. 5B (association plot). Our correlation analyses showed a strong correlation between the DAPL1 and the selected TRPS1, MRAP2, KRT5 and LGALS7 genes in the TCGA-BRCA cohort (Fig. 5C). The GO analysis data of Fig. 6 identified the DAPL1-associated cellular component (e.g. DNA packaging complex, methyltransferase complex, etc.), molecular function (e.g. ubiquitin-like protein binding, ribonucleoprotein complex binding, etc), and biological process (e.g. ribonucleoprotein complex localization, translational elongation, etc.). KEGG pathway analysis data further showed a series of pathways, such as ubiquitin mediated proteolysis, RNA transports, cell adhesion molecules (CAMs), chemical carcinogenesis, and so on (Fig. 7A). In addition, GSEA data also suggested the correlation between DAPL1 expression and the biological issues of DNA repair complex, nucleotide excision repair, translational elongation, ubiquitin-like protein binding, ubiquitin proteasome pathway, and proteasome (Fig. 7B).

**DAPL1-correlated kinase analysis**

After the prediction of potential phosphorylation sites and kinases, five sites of human DAPL1 protein [NP_001017920.2], namely S10, T35, T77, T86, and Y99, were identified (Fig. 8A). And the CDC2, MAPK were predicted as the potential catalytic kinases of DPAL1 protein at S10 site, while MAPK and PKA were for the T86 site (Fig. 8A). In addition, we performed an enrichment analysis of kinase pathway, and also identified several DAPL1-correlated kinases, such as PRKACA, PRKCA, GSK3B, ROCK, and CHEK1 (Fig. 8B).

**Discussion**

In this study, we focused on the correlation between the DAPL1 gene and breast cancer for the first time. We observed the low expression of DAPL1 in invasive breast cancer tissue, from the levels of both gene and protein, suggesting that DAPL1 may be a kind of tumor suppressor gene. It was found that DAPL1 gene was downregulated in an in vitro cultured ovarian granulosa cell of pig [22], but upregulated in the endometrial stromal and epithelial cells under the treatment of lipopolysaccharide [23]. DAPL1 gene was found to be upregulated in nine hepatoblastomas patients [24]. Based on the
expression dataset of TCGA-LIHC (Liver hepatocellular carcinoma) cohort, we also observed an enhanced DAPL1 gene expression in hepatocellular carcinoma tissues, compared to normal tissues (data not shown). Therefore, the low expression characteristics of the DAPL1 gene in breast cancer may be tumor-specific.

Next, we explored the correlation between DAPL1 expression status and clinical prognosis of breast cancer cases, based on the survival datasets of TCGA and GEO databases, respectively. As stated above, the low expression of DAPL1 is correlated with the poor OS prognosis of the breast cancer patients in the TCGA-BRCA cohort. Besides, the poor prognoses of distance metastasis free survival and relapse free survival, but a good post progression survival prognosis, were observed in the breast cancer cases of GEO with the lowly expressed DAPL1 gene. Considering the impact of clinical information on survival analysis, we performed a series of subgroup analyses, and observed the distinct association conclusions. For the OS prognosis, we observed that low expression of DAPL1 was correlated with the poor OS prognosis of breast cancer patients with pathological grade 1, but a good prognosis of breast cancer patients with pathological grade 3. We also noticed that the sample size in grade 1 group was relatively limited, containing only 26 patients. More breast cancer cases and more detailed clinical information are needed to further investigate the complexity of the correlation between DAPL1 expression and clinical prognosis.

Considering the association between DNA methylation and breast cancerogenesis [25-27], we want to study whether DNA hypermethylation of DAPL1 is associated with the low expression level of DAPL1 in breast cancer tissues. Our finding showed no obvious methylation modification of DAPL1 and the lack of the correlation between promoter methylation status and gene expression level. In addition, for the breast cancer cases of the TCGA-BRCA cohort, DAPL1 did not exhibit a high-frequency of genetic alteration status. Due to the low mutation rate, we performed the survival analysis of TCGA patients with overall tumors and observed the negative conclusion. These suggested that DNA methylation and genetic alteration are not involved in the role of DAPL1 in breast cancer pathogenesis. Even though this, for the first time, we found that the expression of DAPL1 gene is related to the mutation status of several genes (e.g. MAP3K1, NUP98, and KIF21B, etc), hinting that the genetic alteration of
these genes may be a potential upstream regulatory mechanism for DAPL1 expression in breast cancer.

The infiltration of immune cells is also reported to be associated with the occurrence of breast cancer [28–30]. Our preliminary analysis showed a positive correlation of the DAPL1 gene expression with the infiltration level of M1 macrophage, NK cell, B cell, Tfh, naive T cell, effector T cell, effector memory T cell, central memory T cell, resident memory T, but a negative correlation with resting or effect Treg T cell. Although we observe a statistically significant correlation, the relatively small R-value means a weak correlation. The factor of immune cell infiltration may only be involved, but not the most critical issue affecting the onset of breast cancer.

In the present study, we screened out a number of DAPL1-correlated genes in breast cancer tissues, and then performed a series of enrichment analyses. Our data indicated that the protein ubiquitination and DNA repair may be the essential mechanisms of DAPL1 gene in the pathogenesis of breast cancer, providing an analytical perspective for further investigation. Additionally, we tried to conduct an in-depth analysis from the perspective of protein phosphorylation. Unfortunately, we did not obtain the data of DAPL1 protein expression and phosphorylation modification from the Clinical Proteomic Tumor Analysis Consortium (CPTAC) dataset. Our predictive analysis result identified several potential phosphorylation sites and catalytic kinases of DAPL1 protein. Our Kinase enrichment analysis further showed several DAPL1-correlated kinases. It is possible that the kinases of CDC2, MAPK, and PKA mediate the phosphorylation level of DAPL1 protein, and the DAPL1 gene influences the expression level of other kinases (e.g., PRKACA, PRKCA, GSK3B, etc.), in the occurrence and development of breast invasive carcinoma. These deserve more in-depth molecular research.

Conclusion
In summary, based on the breast cancer cases within TCGA and GEO datasets, we first identified the lowly expressed oncogenic gene DAPL1 as a prognostic gene for breast cancer. Our finding also suggested the involvement of protein phosphorylation, immune cell infiltration, and DNA repair or protein ubiquitin-associated cellular pathways in the role of DAPL1 in the pathogenesis of breast invasive carcinoma.
Additional Files

**Fig. S1.** Data of DAPL1 gene tree.

**Fig. S2.** The DNA methylation status of DAPL1 through MEXPRESS. Based on the data of the TCGA-BRCA, we analyzed the DNA methylation status of DAPL1 in breast cancer cases through the MEXPRESS tool.

**Table S1.** Correlation of DAPL1 expression and DMFS survival of breast cancer cases in GEO (Kaplan-Meier plotter).

**Table S2.** Correlation of DAPL1 expression and OS survival of breast cancer cases in GEO (Kaplan-Meier plotter).

**Table S3.** Correlation of DAPL1 expression and PPS survival of breast cancer cases in GEO (Kaplan-Meier plotter).

**Abbreviations**

DAPL1, death-associated protein-like 1; AMD, age-related macular degeneration; RPE, retinal pigment epithelium; TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus; BRCA, Breast invasive carcinoma; ER, estrogen receptor; HER2, human epidermal growth factor receptor-2; PR, progesterone receptor; HPA, Human Protein Atlas; OS, overall survival; DMFS, distant metastasis free survival; RFS, relapse free survival; PPS, post progression survival; TP53, Tumor Protein P53; CNV, copy number variation; TAMs, Tumor-associated macrophages; NK cell, Natural killer cell; Tfh, Follicular B helper T cells; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSEA, Gene Set Enrichment Analysis; CAMs, cell adhesion molecules; LIHC, Liver hepatocellular carcinoma; CPTAC, Clinical Proteomic Tumor Analysis Consortium.

**Declarations**

**Acknowledgements**

Not applicable.

**Author’s contributions**

Haichao Zhang conceived and designed the study and drafted this manuscript. Haichao Zhang and Xin Qu performed the gene expression, survival curve, and immune cell infiltration analyses. Haichao Zhang and Lu Han performed DNA methylation, genetic alteration, phosphorylation, and enrichment
analyses. All authors have approved the final version of this manuscript.

**Funding**

Not applicable.

**Availability of data and materials**

The data of this manuscript can be downloaded from The Cancer Genome Atlas database (https://portal.gdc.cancer.gov/).

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**References**

1. Mazo C, Kearns C, Mooney C, Gallagher WM: Clinical Decision Support Systems in Breast Cancer: A Systematic Review. Cancers (Basel). 2020; 12(2).

2. Grassmann F, Friedrich U, Fauser S, Schick T, Milenkovic A, Schulz HL, von Strachwitz CN, Bettecken T, Lichtner P, Meitinger T et al: A Candidate Gene Association Study Identifies DAPL1 as a Female-Specific Susceptibility Locus for Age-Related Macular Degeneration (AMD). Neuromolecular Med. 2015; 17(2):111-120.

3. Ma X, Hua J, Zheng G, Li F, Rao C, Li H, Wang J, Pan L, Hou L: Regulation of cell proliferation in the retinal pigment epithelium: Differential regulation of the death-associated protein like-1 DAPL1 by alternative MITF splice forms. Pigment Cell Melanoma Res. 2018; 31(3):411-422.

4. Ma X, Li H, Wang Y, Wang J, Zheng Q, Hua J, Yang J, Pan L, Lu F, Qu J et al: DAPL1, a susceptibility locus for age-related macular degeneration, acts as a novel suppressor of cell proliferation in the retinal pigment epithelium. Hum Mol Genet. 2017;
5. Chakrabarti S, Multani S, Dabholkar J, Saranath D: Whole genome expression profiling in chewing-tobacco-associated oral cancers: a pilot study. Med Oncol. 2015; 32(3):60.

6. Wang Z, Jensen MA, Zenklusen JC: A Practical Guide to The Cancer Genome Atlas (TCGA). Methods Mol Biol. 2016; 1418:111-141.

7. Tomczak K, Czerwinska P, Wiznerowicz M: The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn). 2015; 19(1a):A68-77.

8. Clough E, Barrett T: The Gene Expression Omnibus Database. Methods Mol Biol. 2016; 1418:93-110.

9. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M et al: NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res. 2013; 41(Database issue):D991-995.

10. Lee E, Moon A: Identification of Biomarkers for Breast Cancer Using Databases. J Cancer Prev. 2016; 21(4):235-242.

11. Tang Z, Kang B, Li C, Chen T, Zhang Z: GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res. 2019; 47(W1):W556-W560.

12. Radvanyi L, Singh-Sandhu D, Gallichan S, Lovitt C, Pedyczak A, Mallo G, Gish K, Kwok K, Hanna W, Zubovits J et al: The gene associated with trichorhinophalangeal syndrome in humans is overexpressed in breast cancer. Proc Natl Acad Sci U S A. 2005; 102(31):11005-11010.

13. Vasaikar SV, Straub P, Wang J, Zhang B: LinkedOmics: analyzing multi-omics data within and across 32 cancer types. Nucleic Acids Res. 2018; 46(D1):D956-d963.

14. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I,
Chakravarthi B, Varambally S: UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. Neoplasia. 2017; 19(8):649-658.

15. Gyorffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, Szallasi Z: An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. Breast Cancer Res Treat. 2010; 123(3):725-731.

16. Koch A, De Meyer T, Jeschke J, Van Criekinge W: MEXPRESS: visualizing expression, DNA methylation and clinical TCGA data. BMC Genomics. 2015; 16:636.

17. Koch A, Jeschke J, Van Criekinge W, van Engeland M, De Meyer T: MEXPRESS update 2019. Nucleic Acids Res. 2019; 47(W1):W561-w565.

18. Modhukur V, Iljasenko T, Metsalu T, Lokk K, Laisk-Podar T, Vilo J: MethSurv: a web tool to perform multivariable survival analysis using DNA methylation data. Epigenomics. 2018; 10(3):277-288.

19. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E et al: Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013; 6(269):pl1.

20. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E et al: The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012; 2(5):401-404.

21. Huang HD, Lee TY, Tzeng SW, Horng JT: KinasePhos: a web tool for identifying protein kinase-specific phosphorylation sites. Nucleic Acids Res. 2005; 33(Web Server issue):W226-229.

22. Kulus M, Sujka-Kordowska P, Konwerska A, Celichowski P, Kranc W, Kulus J, Piotrowska-Kempisty H, Antosik P, Bukowska D, Izycki D et al: New Molecular Markers
Involved in Regulation of Ovarian Granulosa Cell Morphogenesis, Development and Differentiation during Short-Term Primary In Vitro Culture-Transcriptomic and Histochemical Study Based on Ovaries and Individual Separated Follicles. Int J Mol Sci. 2019; 20(16).

23. Salilew-Wondim D, Ibrahim S, Gebremedhn S, Tesfaye D, Heppelmann M, Bollwein H, Pfarrer C, Tholen E, Neuhoff C, Schellander K et al: Clinical and subclinical endometritis induced alterations in bovine endometrial transcriptome and miRNome profile. BMC Genomics. 2016; 17:218.

24. Rodrigues TC, Fidalgo F, da Costa CM, Ferreira EN, da Cunha IW, Carraro DM, Krepischi AC, Rosenberg C: Upregulated genes at 2q24 gains as candidate oncogenes in hepatoblastomas. Future Oncol. 2014; 10(15):2449-2457.

25. Kresovich JK, Gann PH, Erdal S, Chen HY, Argos M, Rauscher GH: Candidate gene DNA methylation associations with breast cancer characteristics and tumor progression. Epigenomics. 2018; 10(4):367-378.

26. Pasculli B, Barbano R, Parrella P: Epigenetics of breast cancer: Biology and clinical implication in the era of precision medicine. Seminars in cancer biology. 2018; 51:22-35.

27. Wang S, Huang Y, Mu X, Qi T, Qiao S, Lu Z, Li H: DNA methylation is related to the occurrence of breast cancer and is not affected by culture conditions. Molecular medicine reports. 2018; 17(5):7365-7371.

28. Cimino-Mathews A, Foote JB, Emens LA: Immune targeting in breast cancer. Oncology (Williston Park, NY). 2015; 29(5):375-385.

29. Desmedt C, Salgado R, Fornili M, Pruneri G, Van den Eynden G, Zoppoli G, Rothé F, Buisseret L, Garaud S, Willard-Gallo K et al: Immune Infiltration in Invasive Lobular Breast Cancer. Journal of the National Cancer Institute. 2018; 110(7):768-776.
30. Karn T, Jiang T, Hatzis C, Sänger N, El-Balat A, Rody A, Holtrich U, Becker S, Bianchini G, Pusztai L: Association Between Genomic Metrics and Immune Infiltration in Triple-Negative Breast Cancer. JAMA oncology. 2017; 3(12):1707-1711.

Tables

Table 1. Correlation of DAPL1 expression and RFS survival of breast cancer cases in GEO Kaplan-Meier plotter.

| Factor              | Subgroup               | Sample size (High/Low) | HR (95% CI)   | logRank_P |
|---------------------|------------------------|------------------------|---------------|-----------|
| ER status           | ER positive            | 292/470                | 0.71(0.52-0.97) | 0.032     |
|                     | ER negative            | 183/164                | 0.64(0.46-0.9)  | 0.0093    |
| PR status           | PR positive            | 309/180                | 0.69(0.47-1.01) | 0.056     |
|                     | PR negative            | 177/195                | 0.73(0.51-1.05) | 0.089     |
| HER2 status         | HER2 positive          | 78/72                  | 0.51(0.29-0.89) | 0.016     |
|                     | HER2 negative          | 455/180                | 0.58(0.43-0.79) | 0.00047   |
| TP53 status         | mutated                | 42/90                  | 1.32(0.72-2.43) | 0.37      |
|                     | Wild type              | 41/41                  | 0.44(0.18-1.08) | 0.086     |
|                     | Grade 1                | 50/58                  | 0.29(0.08-1.5)  | 0.046     |
|                     | Grade 2                | 129/98                 | 0.52(0.31-0.86) | 0.01      |
|                     | Grade 3                | 222/222                | 0.72(0.53-0.99) | 0.04      |
| Intrinsic subtype   | Basal                  | 147/213                | 1.24(0.89-1.71) | 0.2       |
|                     | Luminal A              | 632/209                | 0.66(0.5-0.85)  | 0.0015    |
|                     | Luminal B              | 281/126                | 0.8(0.58-1.11)  | 0.18      |
|                     | HER2+                  | 95/61                  | 1.18(0.74-1.88) | 0.49      |
| Lymph node status   | Lymph node positive    | 528/196                | 0.7(0.54-0.91)  | 0.0071    |
|                     | Lymph node negative    | 207/289                | 0.55(0.36-0.84) | 0.0051    |
| Pietenpol subtype   | Basal-like 1           | 30/75                  | 0.6(0.3-1.21)   | 0.15      |
|                     | Basal-like 2           | 18/34                  | 1.58(0.68-3.66) | 0.28      |
|                     | Immunomodulatory       | 82/28                  | 2.25(0.67-7.6)  | 0.18      |
|                     | Mesenchymal stem-like   | 19/14                  | 0.31(0.11-0.88) | 0.021     |
|                     | Mesenchymal            | 26/75                  | 1.41(0.78-2.54) | 0.26      |
|                     | Luminal androgen       | 65/35                  | 0.53(0.3-0.89)  | 0.015     |

HR, hazard ratio; CI, confidence interval; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor-2; TP53, Tumor Protein P53.

Table 2. Correlation between DAPL1 expression and infiltration level of immune cells.
| Immune cells       | signature genes.                      | R     |
|--------------------|---------------------------------------|-------|
| Monocytes          | CD86/CCR2/CSF1R                       | 0.013 |
| M1 Macrophage      | NOS2/IRF5/PTGS2                       | 0.13  |
| M2 Macrophage      | VSIG4/MRC1/CD163/MSR1                 | -0.039|
| TAMs               | CCL2/CD68/IL10                        | 0.022 |
| NK cell            | FCGR3A/NCAM1/CD94/KIR2DL3/CD161       | 0.071 |
| Dendritic cell     | CD1C/ITGAX/CD83/HLA-DPB1/HLA-DRA     | 0.055 |
| B cell             | CD19/CR2/CD79A                        | 0.13  |
| Tfh                | BCL6/IL21/CXCR5                       | 0.14  |
| Naive T cell       | CCR7/LEF1/TCF7/SELL                   | 0.12  |
| Effector T cell    | CX3CR1/FGFBP2/FCGR3A                  | 0.11  |
| Effector memory T cell | PDCD1/DUSP4/GZMK/GZMA/IFNG          | 0.079 |
| Central memory T cell | CCR7/SELL/IL7R                    | 0.087 |
| Resident memory T cell | CD69/ITGAE/CXCR6/MYADM               | 0.13  |
| Exhausted T cell   | HAVCR2/TIGIT/LAG3/PDCD1/CXCL13/LAYN  | 0.0059|
| Resting Treg T cell | FOXP3/IL2RA                          | -0.084|
| Effector Treg T-cell | FOXP3/CTLA4/CCR8/TNFRSF9          | -0.11 |

TAMs, Tumor-associated macrophages; NK cell, Natural killer cell; Tfh, Follicular B helper T cells.

Figures
Expression analysis of DAPL1 in breast cancer. The genetic location of the DAPL1 gene (A).

We analyzed the expression level of the DAPL1 gene in the breast cancer tissues and corresponding control tissues in the TCGA-BRCA cohort (B) and the dataset of Radvanyi Breast in the Oncomine databases (C). Immunohistochemistry analysis data on the expression of DAPL1 in normal tissue and breast duct or lobular carcinoma tissue were also provided (D). Besides, we analyzed the association between DAPL1 expression and a series of clinical factors (e.g., intrinsic subtype, pathological stage, ER status, etc.) analyzed (E).
Survival curve analysis of DAPL1 for breast cancer cases. We utilized the UALCAN web-portal to perform the overall survival analyses by the factors of cancer type, race, menopause status (A). Based on the data of GEO, we also performed the survival analyses of OS, DMFS, RFS, and PPS, using the Kaplan-Meier plotter (B).
Correlation analysis between DNA methylation status and expression level of DAPL1. Based on the data of the TCGA-BRCA, we provided the waterfall plot of the methylation level of DAPL1 gene (A). We also analyzed the association between the gene expression and promoter methylation of DAPL1 between the normal and primary breast tumor tissues, by the factors of pathological stage and intrinsic subtype (B). TNBC, triple negative breast cancer.
Figure 4

Genetic alteration analysis of DAPL1. We analyzed the genetic alteration features of DAPL1 (A), and perform the OS (B) and disease/progression-free survival (C) analyses, based on the total cancer data of the TCGA project, respectively. We also analyzed the correlation between the DAPL1 expression level and the mutation (D) or CNV (E) status of genes for the breast cancer cases.
DAPL1-correlated gene analysis. We first obtained the DAPL1 positively or negatively correlated significant genes by the LinkedOmics, and provided the data of heat maps with top 10 genes (A) and association plot (B). We also analyzed the expression correlation between DAPL1 and the selected genes, including TRPS1, MRAP2, KRT5, and LGALS7.
Figure 6

Data of GO analysis. Based on the DAPL1-correlated genes, the enriched information of biological process, cellular component, and molecular function in GO analysis was provided.
Figure 7

Data of KEGG and GSEA analysis. Based on the DAPL1-correlated genes, the enriched information of KEGG pathway (A), and the data of GSEA (B) were provided, respectively.
DAPL1-correlated kinase analysis. We first predicted the phosphorylation sites and the potential kinases by means of PhosphoSitePlus, PhosphoNET and KinasePhos tool (A). We then performed the enriched analysis for the kinases, based on the DAPL1-correlated genes (B).

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.
Figure S2.tif
Table S1-S3.docx
Figure S1.tif
