Original Article
Running Title: An Investigation of *MAGE* Family Genes
Received: September 18, 2024; Accepted: January 8, 2025

An Investigation on the Genetic and Epigenetic Changes of *MAGE* Family Genes in Breast Cancer Metastasis and Chemoresistance

Adam Hermawan♦, PhD, I Made Bayu Kresna Yoga, BSc, Irmasari Irmasari, BSc

Laboratory of Macromolecular Engineering, Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Universitas Gadjah Mada Sekip Utara II, Yogyakarta, Indonesia

♦Corresponding Author
Adam Hermawan, PhD
Laboratory of Macromolecular Engineering,
Department of Pharmaceutical Chemistry,
Faculty of Pharmacy, Universitas Gadjah Mada Sekip Utara II,
Yogyakarta, Indonesia
Tel/Fax: +62-274-543120
Email: adam_apt@ugm.ac.id

Abstract

Background: The investigation of genetic modifications and epigenetic controls within the melanoma antigen gene family (*MAGE*) gene family, the cancer-testis antigen, in breast cancer still remains elusive. The present study aimed to detect genetic and epigenetic alterations in the *MAGE* family genes among breast cancer patients, specifically those with metastatic breast cancer and experienced resistance to chemotherapy.

Method: In this bioinformatics study, *MAGE* family genes were retrieved from HUGO genes database and further analyzed for protein-protein interaction using STRING version 12.0, gene ontology using DAVID v2024q2, genetic alterations using cBioportal, ROC plot using ROC plotter, prognostic value using Kaplan-Meier Plotter, DNA methylation using MethSurv, and the correlation between immune cell infiltration using TIMER 2.0. P <0.05 was considered statistically significant.

Results: Patients with metastatic breast cancer have experienced genetic abnormalities in four specific genes of *MAGEF1*, *NSMCE3*, *MAGEL1*, and *NDN*. The relapse-free survival indicated that *NSMCE3* and *NDN* have an unfavorable prognosis, while *MAGEF1* and *MAGEL1* have a favorable prognosis in breast cancer patients. A moderate association between the mRNA levels of *MAGEF1* and *MAGEL1* and the efficacy of chemotherapy was observed. The DNA methylation analysis revealed two significant CpG sites within the *MAGEL1* gene in which become a poor prognosis of patients with breast cancer.

Conclusion: This work has the potential to pave the way for the creation of immunotherapy and improved treatment strategy for those struggling with metastatic breast cancer and chemoresistance. Further research is guaranteed to authenticate the outcomes of these bioinformatics discoveries.

Keywords: *Melanoma antigen*, Breast neoplasms, Immunotherapy, Drug resistance, Bioinformatics

Introduction

Breast cancer is a highly prevalent kind of cancer among women in both developed and developing countries worldwide. Breast cancer is characterized by the abnormal growth of cells in the breast tissue, particularly in the ducts (the tubes that transport milk to the nipple) and lobules (the glands that create milk).[1] The uncontrolled growth of breast cancer is a result of alterations including genetic and epigenetic alterations.[2] Alterations in the genetic and epigenetic profiles of regulatory genes associated with breast cancer can lead to the proliferation and division of cells without proper regulation.[3] These alterations can lead to the proliferation and advancement of breast cancer, which can then metastasize to other bodily tissues if not appropriately managed.[4] The comprehension of genetic and epigenetic alterations in certain genes motivates researchers to create more focused and efficient therapies for the management of breast cancer.[4]

Melanoma antigen gene family (*MAGE*) families are classified as cancer-testis genes, typically expressed in the testis, yet their expression has been associated with multiple human cancers.[3] The proteins comprising MAGE are encoded by the *MAGE* family genes.[5] As a member of cancer-specific antigens, MAGE have the ability to activate an intense immune reaction.[6] Major histocompatibility complex (MHC) class I and Class II-positive tumors produce cancer testis antigen class MAGE proteins, stimulating immunological responses from T and/or B cells in certain cancer patients.[7] MAGE proteins belong to a specific group of antigens called cancer/testis (CT) antigens.[8] These antigens are shown on the surface of cells by MHC class I molecules, which are located in a region of the immune system that is protected from immunological responses.[9] The MAGE protein family comprises many members. MAGEs can be categorized into two types: type I MAGEs (MAGE-A, -B, and -C subfamilies), which are expressed in the testis and other reproductive tissues, and type II MAGEs (MAGE-D, -E, -F, -L and Necdin), which have a wide expression in other tissues.[10]

Previous studies showed that *MAGE* family genes regulated cancer progression, either as tumor suppressor genes or oncogenes. The expression of *MAGEA12* and *MAGEA3* is significantly elevated in breast cancer cells.[11] Overexpression of *MAGEA12* induces malignancy in cancer cells.[12] The genes *MAGEA5, MAGEA8, MAGEB4*, and *MAGEB6* function as tumor suppressors, while *MAGEB18* and *MAGED4* operate as oncogenes in breast cancer.[13] A prognostic value analysis of *MAGE* family gene members in breast cancer patients was previously performed; the results indicated that *MAGEC3* was an unfavorable prognostic factor, while *MAGEE2, MAGEH1*, and *MAGEL2* were favorable prognostic factors in breast cancer patients.[13] Nevertheless, the investigation of genetic modifications and epigenetic controls within the *MAGE* gene family in breast cancer patients remains unexplored.

The present study aimed to identify genetic and epigenetic changes in the *MAGE* family genes in breast cancer patients, namely those with metastatic breast cancer and chemoresistance, using a bioinformatics method. We proposed that genetic and epigenetic alterations in the *MAGE* family genes correlate with metastatic breast cancer

and chemoresistance; however, the specific relevant genes require further investigation. This study has the potential to result in the development of immunotherapy and improved therapeutic approaches for individuals struggling with metastatic breast cancer. In conclusion, the results of this study have the capacity to enhance patient outcomes and overall quality of life.

Material and Methods

Data mining

The bioinformatics research commenced by extracting data from the publicly available HUGO genes database, which is the authoritative source for authorized human gene nomenclature.[14] HUGO is the primary source for standardized human gene symbols and names. The database was accessed using the URL https://www.genenames.org/ using the keyword "*MAGE*". The generated results are subsequently exported in the form of an Excel file. As we retrieved from HUGO genes, we obtained 40 gene symbol and their synonym (Supplementary Table 1: available at: http://ugm.id/SFMAGE).

Protein-protein interaction (PPI)

PPI study was conducted using the STRING-Db software version 12.0 to identify and understand the interactions between proteins. STRING-Db is a database that combines data from various sources, including experimental data, genomic context, and scientific literature, to provide comprehensive and accurate information. The analysis was performed via the website https://string-db.org/. In summary, a total of 40 genes belonging to the *MAGE* family were inputted into the STRING database using specific criteria, which included *Homo sapiens* organisms and a median confidence level of 0.04, as previously described.[15] Furthermore, the default settings of the database were configured for the analysis.

Gene ontology (GO)

The functional enrichment of the MAGE family of genes was assessed through an analysis conducted on the DAVID Bioinformatics Resources database, a valuable resource for functional annotation and enrichment analysis of gene lists, accessible at https://davidbioinformatics.nih.gov. Briefly, the *MAGE* family genes are commonly designated with the official gene symbol of the *Homo sapiens* species. GO, a functional annotation tool, was then applied to the data in order to extract information regarding molecular functions, cellular components, and biological processes. Furthermore, the default settings of the database were configured for those analyses. Fisher's Exact test is used to assess gene enrichment analysis. *P*-values less than 0.05 are deemed significant, as previously described.[15,16]

Genetic alterations

The cBioportal database was used for analyzing genetic alterations and obtaining information on signaling pathways due to its comprehensive data coverage, interactive visualization tools, and powerful analysis capabilities. The analysis was performed by accessing the database using the website http://www.cbioportal.org/. To summarize, the user inputs the gene symbol as a query in the cbioportal database by choosing the Breast cancer research. The study with the greatest number of genomic alterations from the Cancer Types Summary was chosen for additional research, specifically OncoPrint, Mutual Exclusivity, copy number alteration, and pathways, as previously described.[17] Furthermore, the default settings of the database were configured for those analyses. Mutual exclusivity analysis was performed using Fisher's exact test. *P*-values less than 0.05 are deemed significant.

Prognostic value

The predictive significance of *MAGE* family member genes was assessed using Kaplan-Meier (KM) plotter with breast

cancer patient data obtained from https://kmplot.com. The KM plotter, with its large dataset, offers sufficient statistical power to confidently evaluate the prognostic value of *MAGE* genes. In summary, gene symbols were entered into the KM plotter database of breast cancer, and various parameters were chosen, such as relapse-free survival (RFS) and overall survival (OS). Furthermore, the default settings of the database were configured to restrict analysis to subtypes and selected cohorts. A significance level of $P < 0.05$ was chosen, as previously explained.[18]

Receiver operating characteristic (ROC) plot

The relationship between gene expression and chemotherapy sensitivity in breast cancer patients was examined using a ROC plotter (http://www.rocplot.org). In summary, the gene symbols were submitted to the ROC plotter, and certain criteria were chosen, including ROC plotter for breast cancer, RFS at five years, OS, and patients who underwent different forms of treatment, JetSet only, and no outliers. The two cohorts are compared using the Mann–Whitney test or the ROC test. The threshold for statistical significance was established at a *P*-value of less than 0.05, as previously described.[19]

DNA methylation analysis

We used the MethSurv tool (https://biit.cs.ut.ee/methsurv/) to examine the expression and prognostic trends associated with the methylation of specific CpG sites within the *MAGE* family genes. Survival analysis of CpG sites located within or in close proximity to a target gene is facilitated by the MethSurv tool. In summary, the gene symbols were submitted to the MethSurv, and certain criteria were chosen, including cancer study of breast invasive carcinoma (BRCA) TCGA March 2017, and showing the heatmap. The DNA methylation values were denoted as beta values between 0 and 1, which were computed using the formula M/(M + U + 100), where M and U denote the intensity values of methylated and unmethylated DNA, respectively, in accordance with the methodology described in an earlier investigation.[30] The prognostic significance of a particular CpG site was analyzed using KM plotter. A significance level of $P < 0.05$ was chosen, as previously described.[20]

Analysis of the correlation between immune cell infiltration and the MAGE family genes

The TIMER 2.0 database (http://timer.cistrome.org) was used to calculate the correlation between immune cell infiltration and the *MAGE* family genes. The database incorporates modules designed to explore cancer-related associations within the cohort of The Cancer Genome Atlas (TCGA). In summary, the gene symbols were submitted to the TIMER 2.0 database, and certain criteria were chosen, including BRCA study, immune association, and immune infiltrates such as B cell, T cell CD8+, T cell CD4+, macrophage, neutrophil, and dendritic cells. Spearman's correlation coefficients were employed to perform correlation analyses; the intensity of the correlation was classified into the following categories: very weak (0.00–0.19), weak (0.20–0.39), moderate (0.40–0.69), strong (0.70–0.89), and very strong (0.90–1.0). A positive value signified a direct association, while a negative score indicated an inverse correlation. The threshold for statistical significance was $P < 0.05$.[19]

Results

PPI network

The results of a network analysis examining PPI among 40 *MAGE* gene family members were as follows: 40 nodes, sixty-four edges, an average node degree of 3.46, and an average local clustering coefficient of 0.596 (Figure 1). Furthermore, the PPI enrichment value generated by this investigation was less than 1.0 × 10-16.

Subsequent to the PPI network analysis, an investigation into the GO of 40 nucleotide *MAGE* families commenced. The greater the number of protein interactions, the greater the likelihood that these proteins will influence one another and participate in MAGE-gamily proteins-regulated breast cancer mechanisms.

GO

The purpose of GO analysis is to collect information from the 40 *MAGE* family genes analyzed in the form of cellular components, molecular functions, and biological processes. The acquisition of seven biological process data, one cellular component data, and three molecular function data was accomplished through GO analysis using the DAVID bioinformatics database. *MAGE* family genes were enriched for biological process of negative regulation of transcription from RNA polymerase II promoter; cellular protein catabolic process, positive regulation of ubiquitin-protein transferase activity; cellular component of nucleus; and molecular function of histone deacetylase binding, protein binding, and ubiquitin protein ligase binding cellular component of nucleus (Table 1).

Genetic alterations analysis

The purpose of gene alteration analysis using the cBioPortal for Cancer Genomics database is to investigate the mechanism pathways of genetic changes in 40 *MAGE* family genes that are associated with breast cancer. One study was chosen from the 21 breast cancer studies included in the cBioPortal for Cancer Genomics database: The Metastatic Breast Cancer Project Provisional December 2021, consisting 379 samples from 301 patients with metastatic breast cancer. This selection was based on the study's inclusion of a 40% incidence of genetic alterations in the *MAGE* family genes among the population under investigation (Figure 2A). Using Oncoprints, one can examine the percentage profile of gene alterations that transpire in query genes. The analysis identified the following percentages of genetic alterations that transpired in each query gene: 25% in *MAGEF1*, 18% in *NSMCE3*, 24% in *MAGEL2*, and 23% in *NDN* (Figure 2B). The wingless-related integration site signaling pathway, which is associated with cellular proliferation, is generated through pathway enrichment analysis pertaining to genetic alterations (Figure 2C). The analysis of copie number alterations for all genes examined yielded insignificant results, including diploid, deletion, gain, and amplification (Figure 2D). Additional mutual exclusivity analysis showed six gene pair with co-occurrence of genetic alterations, including *MAGEL2-NDN*, *MAGEF1-NDN*, *MAGEF1-MAGEL2*, *NSMCE3-NDN*, *NSMCE3-MAGEL2*, and *MAGEF1-NSMCE3* (Table 2).

Prognostic value

Using the parameters RFS and OS, the prognostic value of mRNA expressions of *MAGEF1*, *NSMCE3*, *MAGEL1*, and *NDN* was determined. *NSMCE3* and *MAGEL1* are synonymous with *NDNL2* and *MAGEB1*, respectively. No significant results were observed between the mRNA levels of *MAGEF1*, *NSMCE3*, *MAGEL1*, and *NDN* and OS in patients with breast cancer (Figure 3A). Additionally, female breast cancer patients with decreased mRNA expression of *NSMCE3* ($P = 0.013$) and *NDN* ($P = 0.001$) had a significantly better RFS; however, patients with breast cancer who had low mRNA levels of *MAGEF1* ($P = 0.016$) and *MAGEL1* ($P = 0.013$) had a substantially poorer RFS than the other groups, according to RFS analysis (Figure 3B).

ROC plot

Analyses were conducted on the correlation between gene expression levels of *MAGEF1*, *NSMCE3*, *MAGEL1*, and *NDN* and chemotherapy response as measured by RFS and pathological complete response (PCR) using transcriptome data from breast

cancer patients. There was a significant moderate correlation observed between the expression levels of *MAGEF1* (AUC=0.543, *P* = 1.3e-02) and *MAGEL1* (AUC=0.559, *P* = 5.3e-05) and AUC values of PCR (Figure 4A). The expression levels of *MAGEL1* demonstrated significant prognostic capability when the PCR parameter was used; their respective AUC and *P*-values were 0.559 and 1.3e-02, respectively (Figure 4B).

DNA methylation analysis

The heatmap that was produced to represent the DNA methylation patterns and evaluate the predictive value of aggregating the expression levels of *MAGEF1*, *NSMCE3*, *MAGEL1*, and *NDN* in patients with breast cancer (Figure 5A, Supplementary Figure 2-4: available at http://ugm.id/SFMAGE). The results of the DNA methylation expression level analysis indicated that certain CpG sites displayed notably high levels of methylation, which had a significant prognostic significance in relation to breast cancer ($P < 0.05$, as determined by the likelihood ratio test). The significantly identified CpG sites in *MAGEL1* or *MAGEB1* comprised the following cg02714462 ($P = 0.015$) and cg10941721($P = 0.01$) (Figure 5A-5B, Table 3).

Correlation with infiltration of immune cells

To examine the relationship between target gene expression and the degree of immune cell infiltration in breast cancer patients, TIMER 2.0 was implemented. The purity of breast cancer was correlated positively with the mRNA levels of *MAGEF1* (r = 0.213, *P* = 1.17e-11) and negatively correlated with the levels of *NDN* (r = -0.476, *P* = 2.14e-57) (Table 4, Supplementary Figure 5: available at http://ugm.id/SFMAGE). CD8+ cells exhibited a positive correlation with *NDN2* (r = 0.311, *P* = 1.08e-23). CD4+ exhibited positive correlation with *NSMCE3* (r = 0.233, *P* = 9.04e-14). A positive correlation was observed between macrophage and *NDN2* (r = 0.377, *P* = 6.01e-35). *MAGEL1* exhibited positive correlations with neutrophil (r = 0.214, *P* = 8.92e-12), and *NDN2* were positively correlated with dendritic cells (r = 0.209, *P* = 2.58e-11).

Discussion

The results of this study show that, out of the 40 members of the *MAGE* family genes, there are 4 genes that have experienced genetic change, specifically *MAGEF1*, *NSMCE3*, *MAGEL1*, and *NDN*. Furthermore, this study show that mutual exclusivity revealed six pairs of genes that exhibited co-occurrence of genetic changes, namely involving genes *MAGEL2-NDN*, *MAGEF1-NDN*, *MAGEF1-MAGEL2*, *NSMCE3-NDN*, *NSMCE3-MAGEL2*, and *MAGEF1-NSMCE3*. This finding highlights the significant contribution of the four genes of *MAGE* family in the development of metastatic breast cancer cells. Prognostic value analysis demonstrated that based on RFS, *NSMCE3* (*P* = 0.013) and *NDN* (*P* = 0.001) are bad prognosis, while *MAGEF1* (*P* = 0.016) and *MAGEL1* (*P* = 0.013) are good prognosis in patients with breast cancer.

MAGEF1, a gene on chromosome 3, encodes melanoma-associated antigen F1 (MAGE-F1).[3,21]This protein is expressed across a spectrum of tumor tissues, including breast, cervical, ovarian, and melanoma cancers.[32] In addition to its involvement in tumorigenesis, MAGE-F1 plays a role in early neurogenesis, alongside MAGE-D1, D2, D3, E1, E2, G1, H1, and Necdin.[21] In lung cancer, MAGE-F1 interacts with the NSE1 E3 ubiquitin ligase to form a complex, leading to the degradation of MMS19, and this interaction inhibits the cytosolic Fe-S assembly pathway, reducing DNA repair capacity.[22]

Melanoma-associated antigen B1 (MAGE-B1) encoded by *MAGEB1* or

MAGEL1 is expressed primarily in specific tissues such as placenta, embryo, testis, and cancer cells.[23] MAGE-B1 plays a crucial role in the development, progression, and prognosis of lung cancer.[24] Similarly, MAGE-B2 has been implicated in laryngeal cancer progression, showing positive correlations lymphatic metastasis.[24] In an in vivo study, *MAGEB2* knockdown inhibited tumor development and lung metastasis.[24]

NSMCE3, also known as *MAGEG1*, is a gene that produces a component of the SMC5/6 complex, and is essential for repairing DNA damage.[25] *NSMCE3* also plays a role in regulating the cell cycle, promoting neural differentiation, and facilitating apoptosis.[23,26] Necdin interacts with MAGE-G1 and E2F1 during neuronal development both in vitro and in vivo, specifically in cells suffering death.[27] Accordingly, necdin and MAGE-G1 act as inherent proteins that inhibit apoptosis, specifically protecting neuronal precursors and postmitotic neurons from E2F1-induced apoptosis.[28,29]

NDN, a gene that encodes NDN (Necdin) is situated on chromosome 15q11-q13.[30] Necdin interacts with p53 to weaken the p53 signaling pathway in human cells in response to genotoxic stress.[31] Necdin is a suppressor of metastatic breast cancer that regulate the transcription of c-Myc.[32] Yang H et al. conducted a study showing that the reintroduction of *NDN* caused a downregulation of Bcl-2 levels and initiated apoptosis, resulting in a significant suppression of ovarian cancer cell growth in both in vitro and xenografts.[33] Necdin forms a complex with MAGE-G1 by binding to the transactivation domain of E2F1, and thus generate a complex that inhibits the transcriptional activity of E2F1.[28]

A previous study showed that *NSMCE3* missense mutations cause a new autosomal recessive syndrome that damages chromosomes that affects the stability of the SMC5/6 and HR complexes, leading to impaired function of T and B cell lymphocytes, and is known as lung disease, immunodeficiency and chromosome damage syndrome.[34] Currently, there is a lack of research on the impact of mutations in *MAGEF1*, *MAGEL2*, and *NDN* genes and their association with breast cancer, and therefore, deeper investigation is required to understand this relationship.

A previous study showed that MAGE-G1 encoded by *NSMCE3* inhibits apoptosis in sarcoma, colon cancer, breast cancer, and melanoma cells by enhancing the transcription of E2F1 a protein that increased proliferation in various kinds of cancer.[35] Another study demonstrated that the downregulation of *NDN* gene expression is dependent upon the activity of STAT3 in human melanoma, prostate, and breast cancer cell lines, which has been shown to contribute to the pathogenesis of cancer.[36] These findings support our results, that shows *NSMCE3* and *NDN* is a bad prognosis marker in breast cancer patient according to RFS. In head and neck squamous cell carcinomas, high expression of *MAGEF1* is associated with a significantly poorer clinical prognosis.[22] Patients with breast cancer which have high expression of *MAGEF1* showed a significant better OS than those in the lower levels of *MAGEF1*, especially in basal-like, luminal A, and luminal B, but not in HER2+ subtype.[13] A separate study indicated that breast cancer patients with elevated levels of *MAGEB1* mRNA exhibited superior survival rates in comparison with the contrasting group.[13] Taken together, this study offers significant insights; however, additional research is required to confirm these findings, specifically to investigate the potential involvement of *NSMCE3*, *NDN*, *MAGEF1* and *MAGEL1* in breast cancer progression and chemotherapy response.

The ROC analysis, using PCR and RFS parameters, revealed a statistically

significant and moderate association between the mRNA levels of *MAGEF1* and *MAGEL1* with the sensitivity of chemotherapy. A previous study showed that Necdin expression was diminished in melanoma and drug-resistant ovarian cancer cell lines.[36] Thus, *MAGEF1* and *MAGEL1* have the potential to serve as biomarkers for evaluating the sensitivity of metastatic breast cancer cells to chemotherapy.

The DNA methylation analysis revealed significant CpG sites in *MAGEL1* also known as *MAGEB1*, specifically cg02714462 and cg10941721. This indicates that methylation on the *MAGEL1* gene could potentially impact the metastasis of breast cancer cells to other parts of the body and their sensitivity to chemotherapy. A prior study has identified a significant presence of DNA methylation, specifically cg02714462 in *MAGEB1*, in individuals diagnosed with acute myeloid leukemia and diffuse large B cell lymphoma.[37] Other studies also showed that *MAGEB2* gene plays a role in the development of laryngeal carcinoma by being influenced by methylation.[38,39] Accordingly, the role of methylation of *MAGEL1* in metastatic breast cancer and chemoresistance is interesting topic for further investigations.

Analysis of the infiltration of immune cells revealed that the purity of breast cancer had a slight positive connection with *MAGEF1*, but a moderate negative correlation with *NDN*. The *NDN* expression showed a mild positive correlation with CD8+ cells, while *NSMCE3* exhibited a weak positive correlation with CD4+ cells. There was a weak positive correlation between macrophages and *NDN2* levels, a weak positive correlation between neutrophils and *MAGEL1* levels, and a positive correlation between *NDN2* and dendritic cells. A previous study showed that *MAGEB2* demonstrated a strong correlation with immune cell infiltration, particularly CD8+ T cells, in lung cancer and influenced the secretion of chemokines/cytokines and immunogenicity in lung cancer cells.[24] Given the immunogenicity of type I MAGE antigens, ongoing cancer vaccination studies are exploring their potential as targets for immunotherapy.[23] MAGE proteins, like MAGE-B1 and MAGE-B2, hold promise as cancer-testis antigens and targets for immunotherapy due to their distinctive expression patterns in tumors and significant immunogenicity.[23] Taken together, the effect of *MAGEF1*, *NSMCE3*, *MAGEL1*, and *NDN* on immune cell infiltration in patients with metastatic breast cancer and patients experiencing chemoresistance needs to be further clarified.

Although this study provides valuable insights, it has limitations because it relies solely on computational methods. Computational methods require careful consideration to avoid drawing inaccurate conclusions. In addition, the study findings may be affected by patient heterogeneity and the resistance associated with interpreting bioinformatics data into clinical applications. To overcome these limitations, it is necessary to consider patient heterogeneity and conduct further validation studies, which will strengthen the overall analysis and improve the clinical application of the findings of this study.

The results of this study demonstrate that *MAGEF1* and *MAGEL1* serve as good prognostic indicators for breast cancer patients and predict chemotherapy effectiveness. *NSMCE3* and *NDN* are associated with worse prognoses in breast cancer patients. In addition, *MAGEL1* methylation is a negative predictive indicator in breast cancer patients. Future studies should focus on the findings of this study in relation to clinically relevant applications. This involves investigating the potential of *MAGE* family genes as prognostic biomarkers for patient risk stratification and

chemotherapy treatment response prediction. Additionally, exploring these genes as therapeutic targets is essential, including the development of MAGE-specific immunotherapy and the targeting of MAGE-related signaling pathways. Furthermore, clinical trials are necessary to confirm these findings in larger cohorts and to investigate their clinical implications for enhancing the efficacy of breast cancer therapy.

Conclusion

In this study, we found genetic and epigenetic alterations in the *MAGE* family genes among breast cancer patients, specifically those with metastatic breast cancer and experienced resistance to chemotherapy, by employing a bioinformatics approach. We have discovered four genes that have undergone genetic alterations, namely *MAGEF1, NSMCE3, MAGEL1,* and *NDN*. According to RFS, *NSMCE3* and *NDN* have a negative prognosis, while *MAGEF1* and *MAGEL1* have a positive prognosis in breast cancer patients. The ROC analysis, using PCR and RFS data, showed a statistically significant and moderate correlation between the mRNA levels of *MAGEF1* and *MAGEL1* and the effectiveness of chemotherapy. The investigation of DNA methylation identified two important CpG sites, namely cg02714462 and cg10941721, in the gene *MAGEL1*. This study has the potential to lead to the development of immunotherapy and enhanced therapy strategies for those grappling with metastatic breast cancer and chemoresistance. Further studies are needed to validate the results of these bioinformatics findings, including exploration of the potential of *MAGE* family genes as prognostic biomarkers and therapeutic targets for breast cancer treatment. Also, clinical trials are needed for larger cohorts.

Supplementary Files

Supplementary Figure 1. Oncoprint analysis of MAGE family genes, as analyzed by cBioportal from a study entitled The Metastatic Breast Cancer 2021.
Supplementary Figure 2. Expression levels of MAGEF1 DNA methylation in breast cancer cells, as determined by the MethSurv database, depicted as a heatmap.
Supplementary Figure 3. Expression levels of NSMCE3 DNA methylation in breast cancer cells, as determined by the MethSurv database, depicted as a heatmap.
Supplementary Figure 4. Expression levels of NDLN2 DNA methylation in breast cancer cells, as determined by the MethSurv database, depicted as a heatmap.
Supplementary Figure 5. The correlation between the levels of immune cell infiltration and the expression of MAGE family genes, as determined by TIMER 2.0.
Supplementary Files are available at

http://ugm.id/SFMAGE

Acknowledgements

The authors thank Mrs. Ririn Widarti and Mrs. Dian Anita for their administrative assistance, and Badan Penerbit dan Publikasi Universitas Gadjah Mada for their assistance in writing.

Authors' Contributions

A.H: Study design, data gathering, drafting and critical reviewing of the manuscript; I.M.B.K.Y, and I.I : Data gathering, drafting. All authors have read and approved the final manuscript and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding

None declared.

Conflict of Interest

None declared.

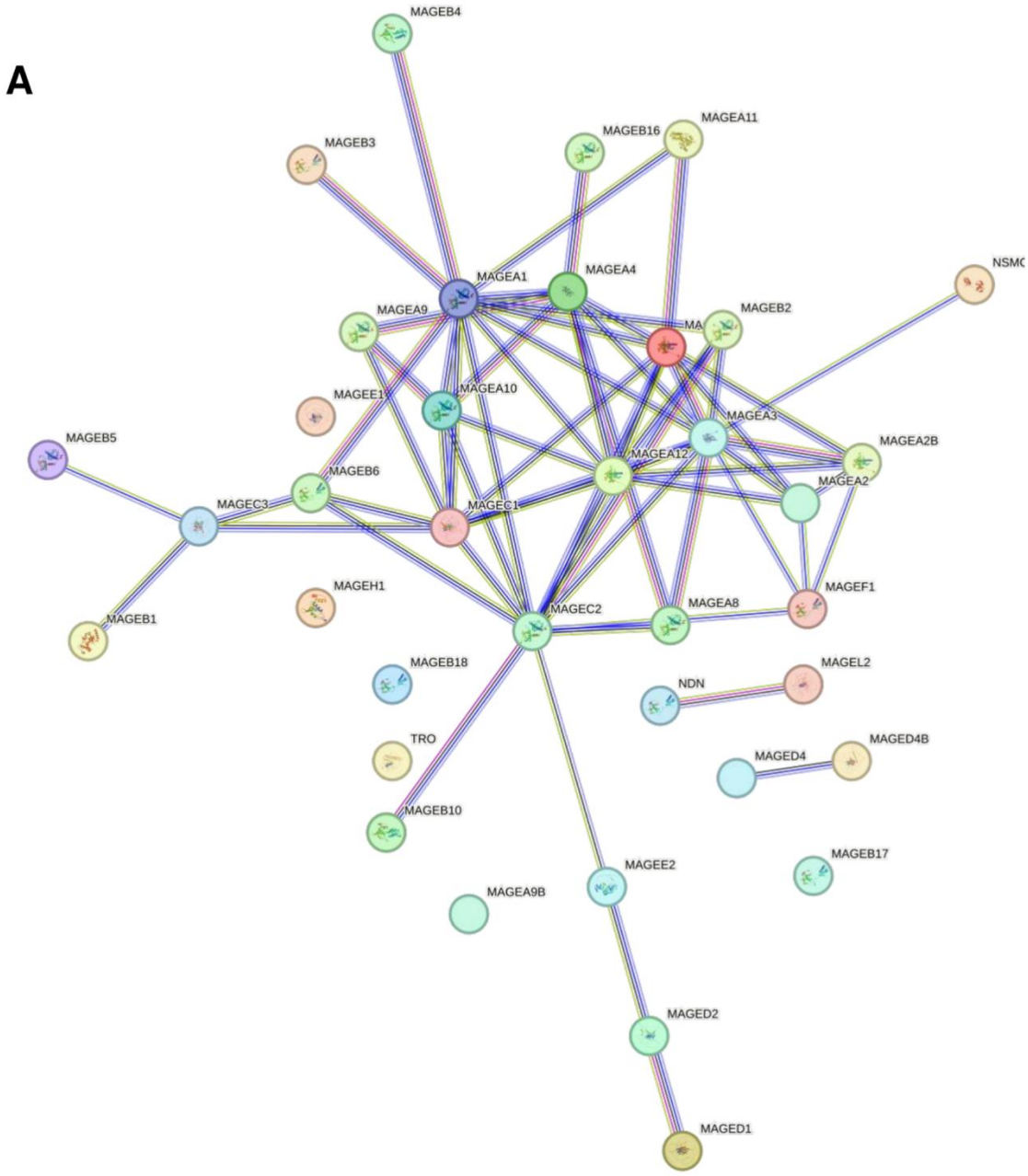
References

1. Feng Y, Spezia M, Huang S, Yuan C, Zeng Z, Zhang L, et al. Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes Dis.* 2018;5(2):77-106. doi: 10.1016/j.gendis.2018.05.001. PMID: 30258937; PMCID: PMC6147049.
2. Coyle KM, Boudreau JE, Marcato P. Genetic mutations and epigenetic modifications: Driving cancer and informing precision medicine. *Biomed Res Int.* 2017;2017:9620870. doi: 10.1155/2017/9620870. Epub 2017 Jun 8. PMID: 28685150; PMCID: PMC5480027.
3. Li S, Shi X, Li J, Zhou X. Pathogenicity of the MAGE family. *Oncol Lett.* 2021;22(6):844. doi: 10.3892/ol.2021.13105. PMID: 34733362; PMCID: PMC8561213.
4. Thakur C, Qiu Y, Fu Y, Bi Z, Zhang W, Ji H, et al. Epigenetics and environment in breast cancer: New paradigms for anti-cancer therapies. *Front Oncol.* 2022;12:971288. doi: 10.3389/fonc.2022.971288. PMID: 36185256; PMCID: PMC9520778.
5. Schooten E, Di Maggio A, van Bergen En Henegouwen PMP, Kijanka MM. MAGE-A antigens as targets for cancer immunotherapy. *Cancer Treat Rev.* 2018;67:54-62. doi: 10.1016/j.ctrv.2018.04.009. PMID: 29763778.
6. Nin DS, Deng LW. Biology of cancer-testis antigens and their therapeutic implications in cancer. *Cells.* 2023;12(6):926. doi: 10.3390/cells12060926. PMID: 36980267; PMCID: PMC10047177.
7. Sohani M, Rastgar A, Kheyrandish S. Harnessing the power of MAGE proteins in cancer immunotherapy for multiple myeloma. I*ran J Blood Cancer.* 2024;16(1):1-20.
8. Zajac P, Schultz-Thater E, Tornillo L, Sadowski C, Trella E, Mengus C, et al. MAGE-A antigens and cancer immunotherapy. *Front Med (Lausanne).* 2017;4:18. doi: 10.3389/fmed.2017.00018. PMID: 28337438; PMCID: PMC5340762.
9. Shim K, Jo H, Jeoung D. Cancer/testis antigens as targets for rna-based anticancer therapy. *Int J Mol Sci.* 2023;24(19):14679. doi: 10.3390/ijms241914679. PMID: 37834126; PMCID: PMC10572814.
10. Abera B, Dinka H. MAGE genes encoding for embryonic development in cattle is mainly regulated by zinc finger transcription factor family and slightly by CpG Islands. *BMC Genom Data.* 2022;23(1):19. doi: 10.1186/s12863-022-01034-0. PMID: 35303799; PMCID: PMC8932067.
11. Oh C, Kim HR, Oh S, Ko JY, Kim Y, Kang K, et al. Epigenetic upregulation of MAGE-A isoforms promotes breast cancer cell aggressiveness. *Cancers (Basel).* 2021;13(13):3176. doi: 10.3390/cancers13133176. PMID: 34202157; PMCID: PMC8268034.
12. Yanagi T, Nagai K, Shimizu H, Matsuzawa SI. Melanoma antigen A12 regulates cell cycle via tumor suppressor p21 expression. *Oncotarget.* 2017;8(40):68448-59. doi: 10.18632/oncotarget.19497. PMID: 28978129; PMCID: PMC5620269.
13. Jia B, Zhao X, Wang Y, Wang J, Wang Y, Yang Y. Prognostic roles of MAGE family members in breast cancer based on KM-Plotter Data. *Oncol Lett.* 2019;18(4):3501-16. doi: 10.3892/ol.2019.10722. PMID: 31516568; PMCID: PMC6733005.

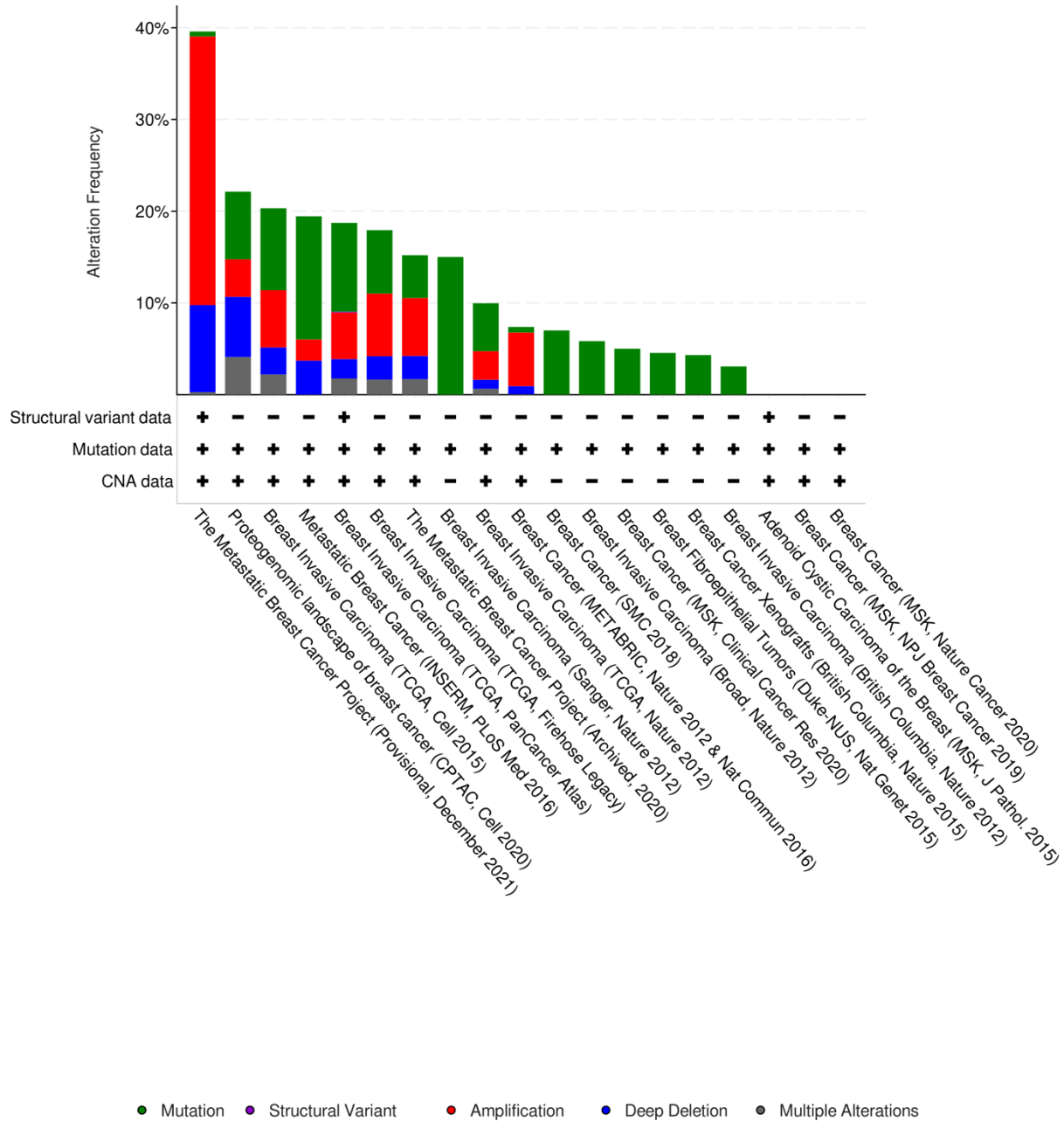
14. Bruford EA, Braschi B, Denny P, Jones TEM, Seal RL, Tweedie S. Guidelines for human gene nomenclature. *Nat Genet.* 2020;52(8):754-58. doi: 10.1038/s41588-020-0669-3. PMID: 32747822; PMCID: PMC7494048.
15. Ali Z, Hermawan A. Bioinformatics analysis uncovers the importance of RTK-RAS-PI3K/Akt regulation by borneol in overcoming breast cancer resistance to tamoxifen. *Indonesian J Pharm*. 2022;33(1):135-46. doi:10.22146/ijp.2346.
16. Hermawan A, Putri H. Bioinformatics analysis of TIMP1, HK2 and IGFBP7 as potential biomarkers and therapeutic targets of paclitaxel resistance in breast cancer. *Middle East J Cancer*. 2021;12(2):198-207. doi:10.30476/mejc.2020.83217.1147.
17. Hermawan A, Putri H. Use of integrative bioinformatics to identify targets of sinensetin and its mechanisms to overcome colorectal cancer resistance. *J Appl Pharm Sci*. 2021;11(1):111-20. doi:10.7324/JAPS.2021.110113.
18. Hermawan A, Satria D, Hasibuan PAZ, Huda F, Tafrihan AS, Fatimah N, et al. Identification of potential target genes of cardiac glycosides from Vernonia amygdalina Delile in HER2+ breast cancer cells. Article. *S Afr J Bot.* 2024;164:401-18. doi:10.1016/j.sajb.2023.12.002.
19. Hermawan A, Putri H. Characterizing excision repair cross-complementing family genes as drug resistance biomarkers in breast cancer. *Beni Suef Univ J Basic Appl Sci.* 2023;12(1):79. doi: 10.1186/s43088-023-00415-3.
20. Hermawan A, Putri H, Fatimah N, Prasetio HH. Transcriptomics analysis reveals distinct mechanism of breast cancer stem cells regulation in mammospheres from MCF-7 and T47D cells. *Heliyon.* 2024;10(2):e24356. doi: 10.1016/j.heliyon.2024.e24356. PMID: 38304813; PMCID: PMC10831612.
21. Zhou Y, Lih TM, Pan J, Höti N, Dong M, Cao L, et al. Proteomic signatures of 16 major types of human cancer reveal universal and cancer-type-specific proteins for the identification of potential therapeutic targets. *J Hematol Oncol.* 2020;13(1):170. doi: 10.1186/s13045-020-01013-x. PMID: 33287876; PMCID: PMC7720039.
22. Weon JL, Yang SW, Potts PR. Cytosolic iron-sulfur assembly is evolutionarily tuned by a cancer-amplified ubiquitin ligase. *Mol Cell.* 2018;69(1):113-25.e6. doi: 10.1016/j.molcel.2017.11.010. PMID: 29225034.
23. Kozakova L, Vondrova L, Stejskal K, Charalabous P, Kolesar P, Lehmann AR, et al. The melanoma-associated antigen 1 (MAGEA1) protein stimulates the E3 ubiquitin-ligase activity of TRIM31 within a TRIM31-MAGEA1-NSE4 complex. *Cell Cycle.* 2015;14(6):920-30. doi: 10.1080/15384101.2014.1000112. PMID: 25590999; PMCID: PMC4614679.
24. Cui J, Chen Y, Ou Y, Liu G, Wen Q, Zhu W, et al. Cancer germline antigen gene MAGEB2 promotes cell invasion and correlates with immune microenvironment and immunotherapeutic efficiency in laryngeal cancer. *Clin Immunol.* 2022;240:109045. doi: 10.1016/j.clim.2022.109045. PMID: 35618211.
25. Tacer KF, Potts PR. Cellular and disease functions of the Prader-Willi Syndrome gene MAGEL2. *Biochem J.* 2017;474(13):2177-90. doi: 10.1042/BCJ20160616. PMID: 28626083; PMCID: PMC5594744.
26. Florke Gee RR, Chen H, Lee AK, Daly CA, Wilander BA, Fon Tacer K, et al. Emerging roles of the MAGE protein family in stress response pathways. *J Biol Chem.* 2020;295(47):16121-55. doi: 10.1074/jbc.REV120.008029. PMID: 32921631; PMCID: PMC7681028.

27. Yoshikawa K. Necdin: A purposive integrator of molecular interaction networks for mammalian neuron vitality. *Genes Cells*. 2021;26(9):641-83. doi: 10.1111/gtc.12884. PMID: 34338396; PMCID: PMC9290590.
28. Kuwako K, Taniura H, Yoshikawa K. Necdin-related MAGE proteins differentially interact with the E2F1 transcription factor and the p75 neurotrophin receptor. *J Biol Chem*. 2004;279(3):1703-12. doi:10.1074/jbc.M308454200.
29. Li J, Bi W, Lu F, Pan B, Xiong M, Nasifu L, et al. Prognostic role of E2F1 gene expression in human cancer: a meta-analysis. *BMC Cancer*. 2023;23(1):509. doi: 10.1186/s12885-023-10865-8. PMID: 37277745; PMCID: PMC10243032.
30. Napolitano L, Barone B, Morra S, Celentano G, La Rocca R, Capece M, et al. Hypogonadism in patients with Prader Willi syndrome: A narrative review. *Int J Mol Sci.* 2021;22(4):1993. doi: 10.3390/ijms22041993. PMID: 33671467; PMCID: PMC7922674.
31. Lafontaine J, Tchakarska G, Rodier F, Mes-Masson AM. Necdin modulates proliferative cell survival of human cells in response to radiation-induced genotoxic stress. *BMC Cancer*. 2012;12:234. doi: 10.1186/1471-2407-12-234. PMID: 22691188; PMCID: PMC3495902.
32. Lee M, Beggs SM, Gildea D, Bupp S, Lichtenberg J, Trivedi NS, et al. Necdin is a breast cancer metastasis suppressor that regulates the transcription of c-Myc. *Oncotarget.* 2015;6(31):31557-68. doi: 10.18632/oncotarget.5230. PMID: 26384308; PMCID: PMC4741624.
33. Yang H, Das P, Yu Y, Mao W, Wang Y, Baggerly K, et al. NDN is an imprinted tumor suppressor gene that is downregulated in ovarian cancers through genetic and epigenetic mechanisms. *Oncotarget.* 2016;7(3):3018-32. doi: 10.18632/oncotarget.6576. PMID: 26689988; PMCID: PMC4823087.
34. van der Crabben SN, Hennus MP, McGregor GA, Ritter DI, Nagamani SC, Wells OS, et al. Destabilized SMC5/6 complex leads to chromosome breakage syndrome with severe lung disease. *J Clin Invest.* 2016;126(8):2881-92. doi: 10.1172/JCI82890. PMID: 27427983; PMCID: PMC4966312.
35. Peche LY, Ladelfa MF, Toledo MF, Mano M, Laiseca JE, Schneider C, et al. Human MageB2 protein expression enhances E2F transcriptional activity, cell proliferation, and resistance to ribotoxic stress. *J Biol Chem.* 2015;290(49):29652-62. doi: 10.1074/jbc.M115.671982. PMID: 26468294; PMCID: PMC4705963.
36. Haviland R, Eschrich S, Bloom G, Ma Y, Minton S, Jove R, et al. Necdin, a negative growth regulator, is a novel STAT3 target gene down-regulated in human cancer. *PLoS One.* 2011;6(10):e24923. doi: 10.1371/journal.pone.0024923. PMID: 22046235; PMCID: PMC3203112.
37. Dufva O, Pölönen P, Brück O, Keränen MAI, Klievink J, Mehtonen J, et al. Immunogenomic landscape of hematological malignancies. *Cancer Cell.* 2020;38(3):380-399.e13. doi: 10.1016/j.ccell.2020.06.002. Erratum in: *Cancer Cell.* 2020;38(3):424-8. doi: 10.1016/j.ccell.2020.08.019. PMID: 32649887.
38. Cui J, Wang L, Zhong W, Chen Z, Chen J, Yang H, et al. Identification and validation of methylation-driven genes prognostic signature for recurrence of laryngeal squamous cell carcinoma by integrated bioinformatics analysis. *Cancer Cell Int.* 2020;20:472. doi: 10.1186/s12935-020-01567-3. PMID: 33005105; PMCID: PMC7526132.
39. Cui J, Wang L, Zhong W, Chen Z, Chen J, Yang H, et al. Development and validation of epigenetic signature predict survival for patients with laryngeal squamous cell carcinoma. *DNA Cell Biol.*

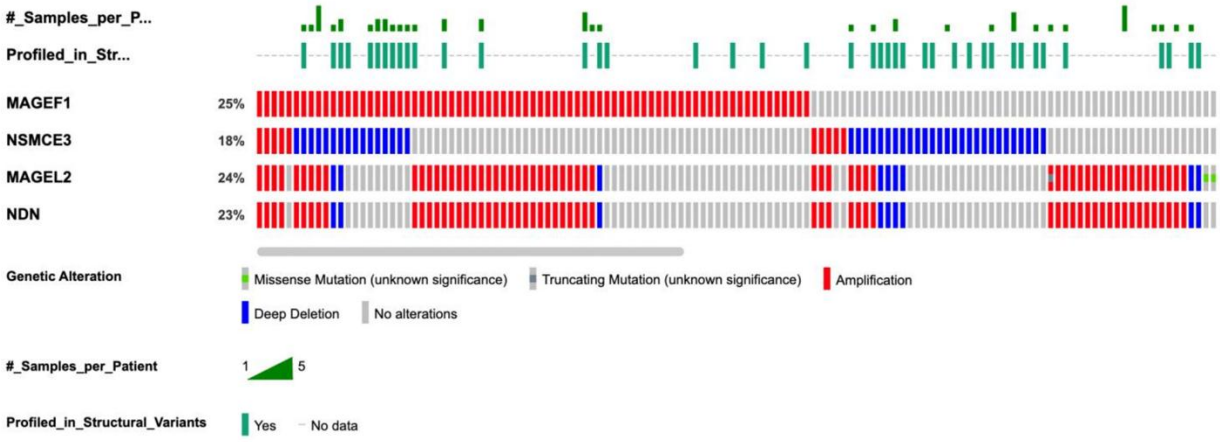
2021;40(2):247-64. doi: 10.1089/dna.2020.5789. PMID: 33481663.


A
MAGEB4
MAGEB3
MAGEB16
MAGEA11
NSMC
MAGEA1
MAGEA4
MAGEA9
MAGEB2
MAGEE1
MAGEA10
MAGEA3
MAGEB5
MAGEA2B
MAGEB6
MAGEA12
MAGEA2
MAGEC3
MAGEC1
MAGEH1
MAGEF1
MAGEB1
MAGEC2
MAGEA8
MAGEB18
MAGEL2
NDN
TRO
MAGED4B
MAGED4
MAGEB10
MAGEB17
MAGEE2
MAGEA9B
MAGED2
MAGED1

B


40%
30%
20%
10%
Alteration Frequency
Structural variant data
Mutation data
CNA data
The Metastatic Breast Cancer Project (Provisional, December 2021)
Proteogenomic landscape of breast cancer (CPTAC, Cell 2020)
Breast Invasive Carcinoma (TCGA, Cell 2015)
Metastatic Breast Cancer (INSERM, PLoS Med 2016)
Breast Invasive Carcinoma (TCGA, PanCancer Atlas)
Breast Invasive Carcinoma (TCGA, Firehose Legacy)
The Metastatic Breast Cancer Project (Archived, 2020)
Breast Invasive Carcinoma (Sanger, Nature 2012)
Breast Invasive Carcinoma (TCGA, Nature 2012)
Breast Cancer (METABRIC, Nature 2012 & Nat Commun 2016)
Breast Cancer (SMC 2018)
Breast Invasive Carcinoma (MSK, Clinical Cancer Res 2020)
Breast Fibroepithelial Tumors (Broad, Nature 2012)
Breast Cancer Xenografts (British Columbia, Nature 2015)
Breast Invasive Carcinoma (Duke-NUS, Nat Genet 2015)
Adenoid Cystic Carcinoma of the Breast (MSK, J Pathol. 2015)
Breast Cancer (MSK, NPJ Breast Cancer 2019)
Breast Cancer (MSK, Nature Cancer 2020)
Mutation
Structural Variant
Amplification
Deep Deletion
Multiple Alterations

C


#_Samples_per_P...
Profiled_in_Str...
MAGEF1
25%
NSMCE3
18%
MAGEL2
24%
NDN
23%
Genetic Alteration
Missense Mutation (unknown significance)
Truncating Mutation (unknown significance)
Amplification
Deep Deletion
No alterations
#_Samples_per_Patient
1
5
Profiled_in_Structural_Variants
Yes
No data

D

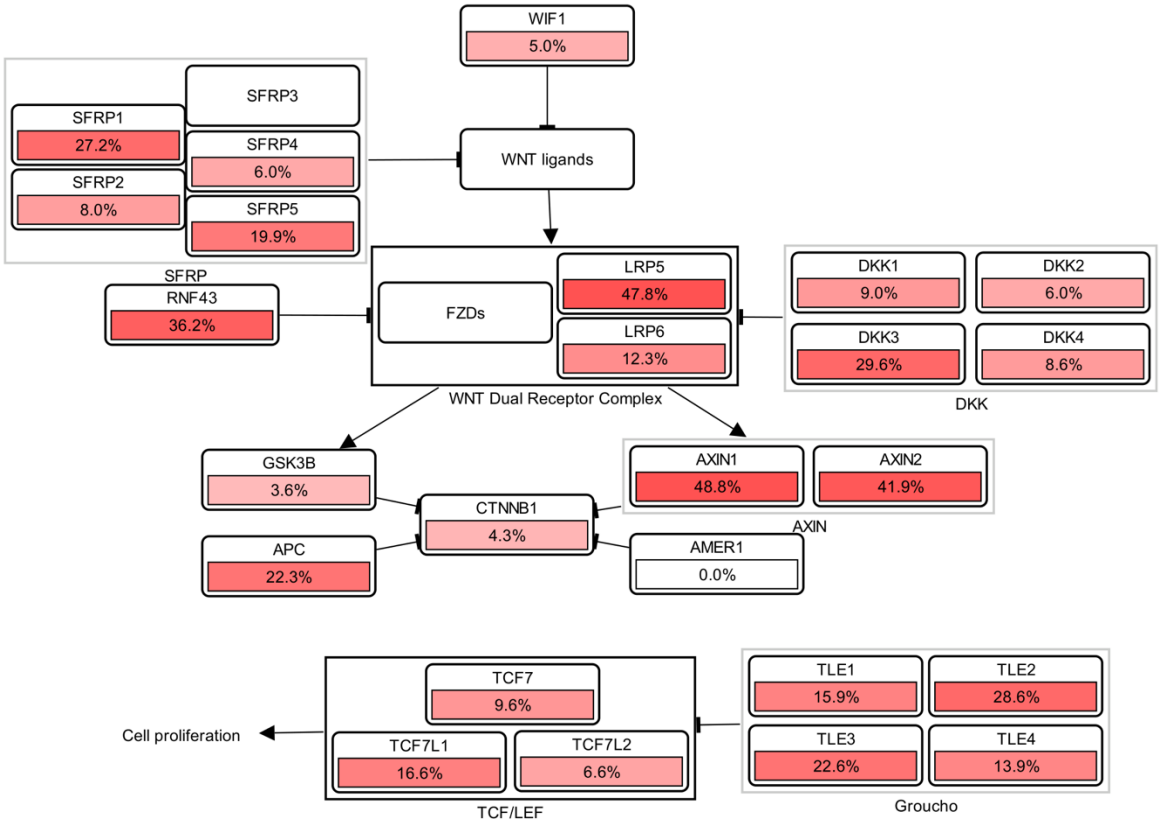
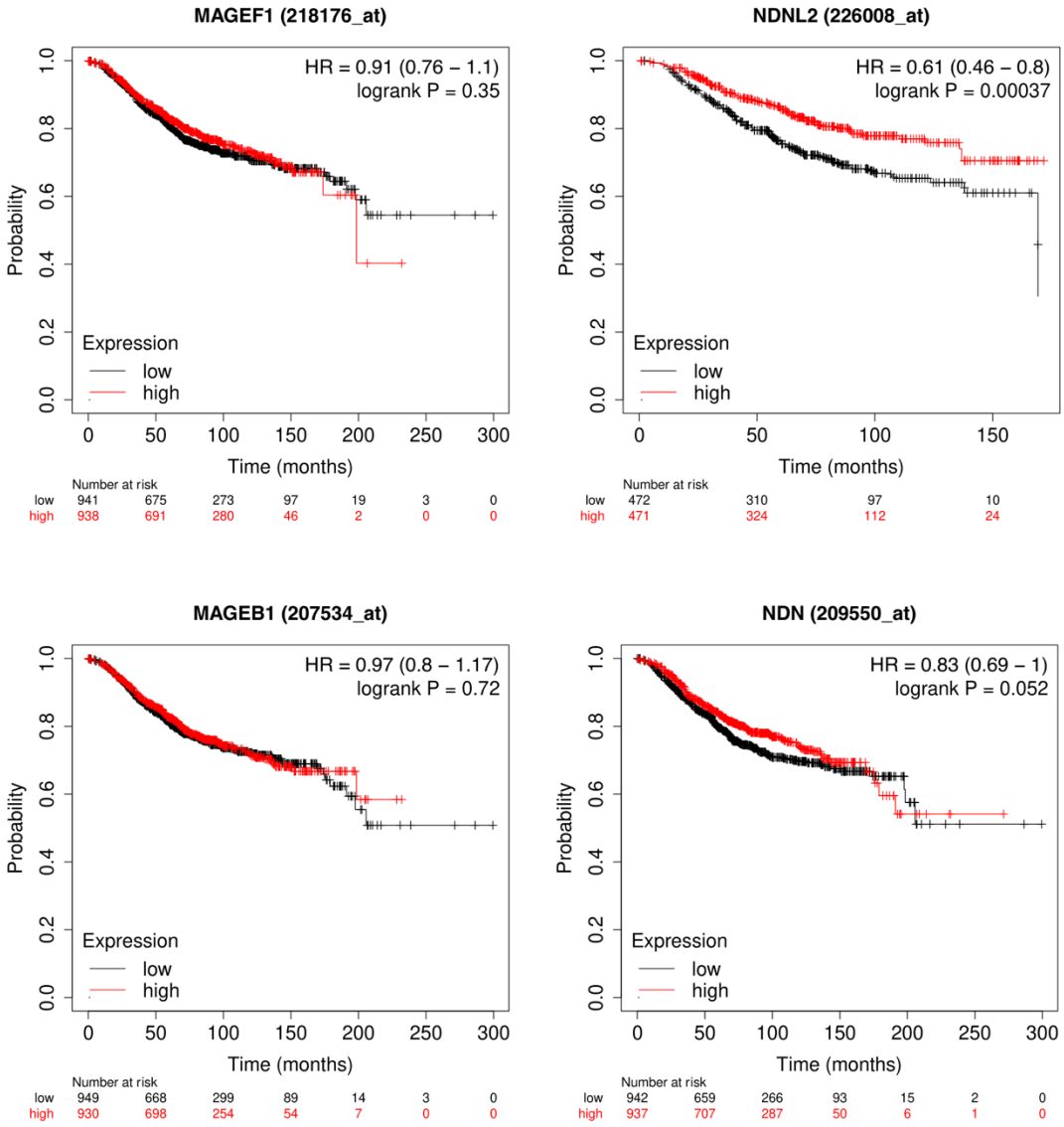
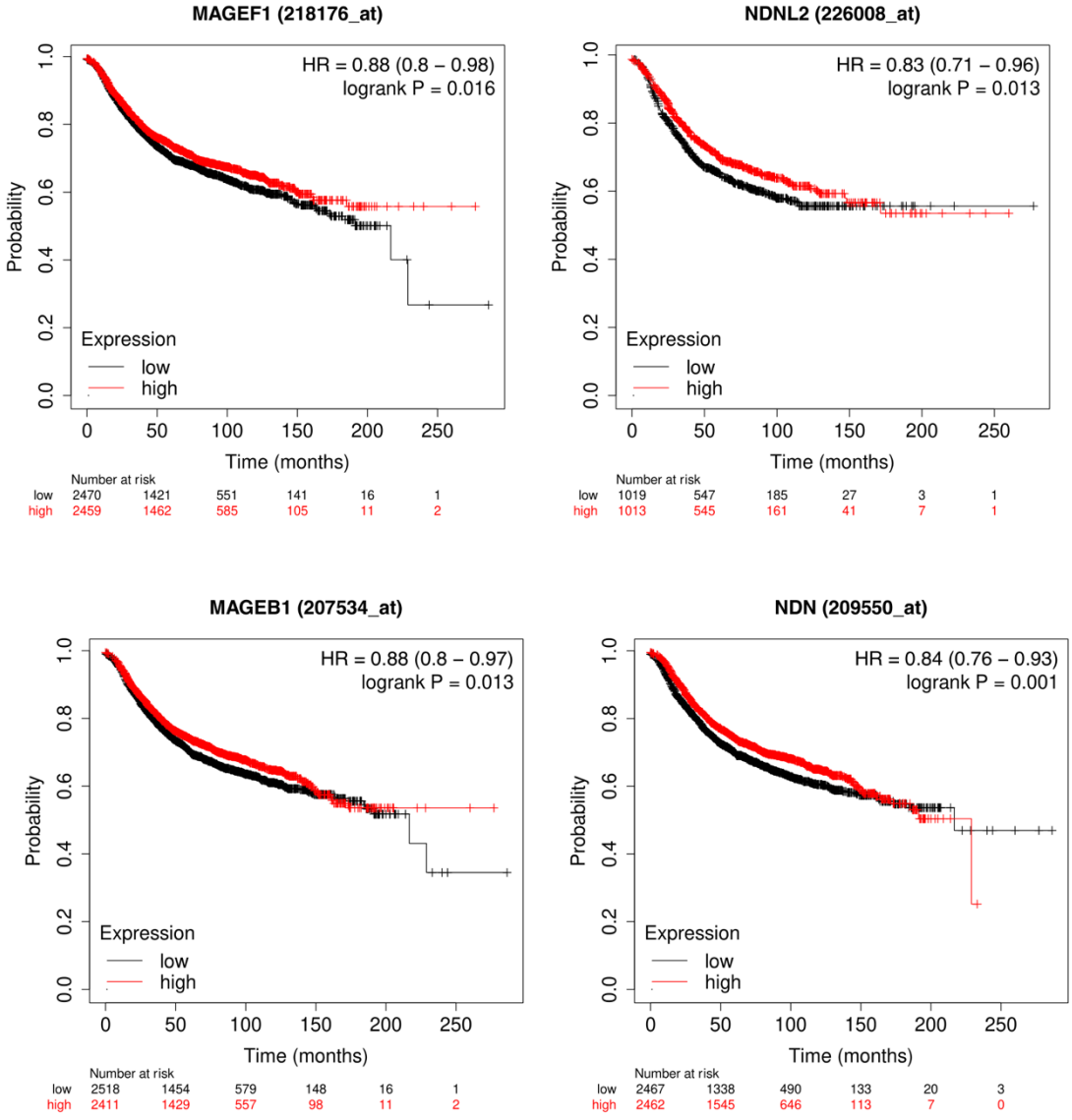

WIF1
5.0%
SFRP3
SFRP1
27.2%
SFRP4
6.0%
SFRP2
8.0%
SFRP5
19.9%
SFRP
WNT ligands
RNF43
36.2%
FZDs
LRP5
47.8%
LRP6
12.3%
WNT Dual Receptor Complex
DKK1
9.0%
DKK2
6.0%
DKK3
29.6%
DKK4
8.6%
DKK
GSK3B
3.6%
AXIN1
48.8%
AXIN2
41.9%
AXIN
CTNNB1
4.3%
APC
22.3%
AMER1
0.0%
TCF7
9.6%
Cell proliferation
TCF7L1
16.6%
TCF7L2
6.6%
TCF/LEF
TLE1
15.9%
TLE2
28.6%
TLE3
22.6%
TLE4
13.9%
Groucho

Figure 1. (A). Protein-protein interaction network of MAGE family genes. (B). Summary alterations of MAGE family genes across breast cancer studies in cBioportal database. (C). Oncoprint analysis of MAGE family genes across The Metastatic Breast Cancer Project Provisional 2021 study in cBioportal showed only four alterations in *MAGEF1*, *NSMCE3*, *MAGEL2* and *NDN*. (D). Pathway related to genetic alterations of *MAGEF1, NSMCE3, MAGEL2* and *NDN* across The Metastatic Breast Cancer Project Provisional 2021 study in cBioportal. (E). Copy number alterations of *MAGEF1, NSMCE3, MAGEL2* and *NDN* across The Metastatic Breast Cancer Project Provisional 2021 study in cBioportal.
MAGE: Melanoma antigen gene

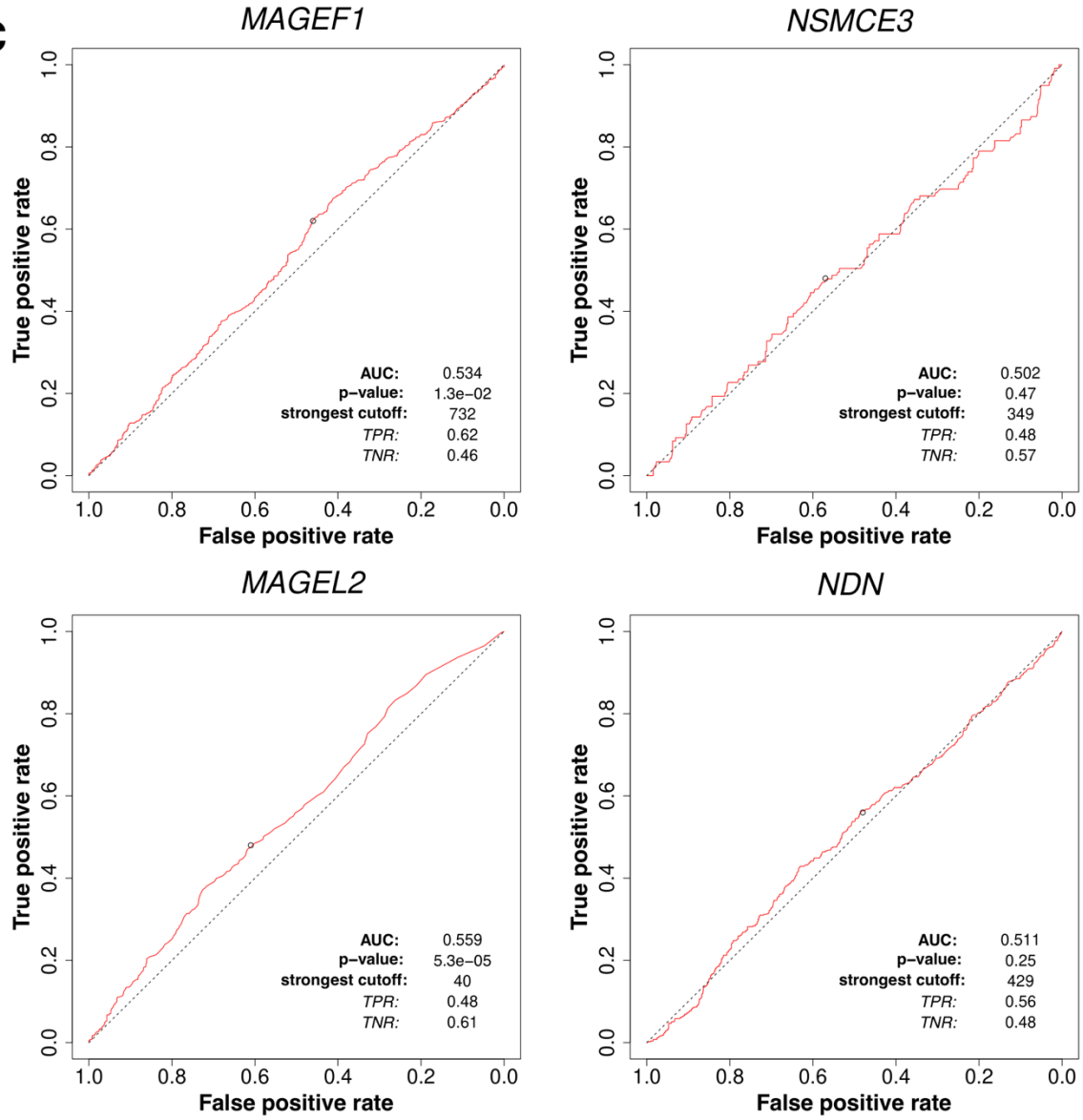
A


MAGEF1 (218176_at)
HR = 0.91 (0.76 – 1.1)
logrank P = 0.35
Probability
Expression
low
high
Time (months)
Number at risk
low 941 675 273 97 19 3 0
high 938 691 280 46 2 0 0
NDNL2 (226008_at)
HR = 0.61 (0.46 – 0.8)
logrank P = 0.00037
Probability
Expression
low
high
Time (months)
Number at risk
low 472 310 97 10
high 471 324 112 24
MAGEB1 (207534_at)
HR = 0.97 (0.8 – 1.17)
logrank P = 0.72
Probability
Expression
low
high
Time (months)
Number at risk
low 949 668 299 89 14 3 0
high 930 698 254 54 7 0 0
NDN (209550_at)
HR = 0.83 (0.69 – 1)
logrank P = 0.052
Probability
Expression
low
high
Time (months)
Number at risk
low 942 659 266 93 15 2 0
high 937 707 287 50 6 1 0

B


MAGEF1 (218176_at)
HR = 0.88 (0.8 – 0.98)
logrank P = 0.016
Probability
Expression
low
high
Time (months)
Number at risk
low 2470 1421 551 141 16 1
high 2459 1462 585 105 11 2
NDNL2 (226008_at)
HR = 0.83 (0.71 – 0.96)
logrank P = 0.013
Probability
Expression
low
high
Time (months)
Number at risk
low 1019 547 185 27 3 1
high 1013 545 161 41 7 1
MAGEB1 (207534_at)
HR = 0.88 (0.8 – 0.97)
logrank P = 0.013
Probability
Expression
low
high
Time (months)
Number at risk
low 2518 1454 579 148 16 1
high 2411 1429 557 98 11 2
NDN (209550_at)
HR = 0.84 (0.76 – 0.93)
logrank P = 0.001
Probability
Expression
low
high
Time (months)
Number at risk
low 2467 1338 490 133 20 3
high 2462 1545 646 113 7 0

C


MAGEF1
True positive rate
False positive rate
AUC: 0.534
p−value: 1.3e−02
strongest cutoff: 732
TPR: 0.62
TNR: 0.46
NSMCE3
AUC: 0.502
p−value: 0.47
strongest cutoff: 349
TPR: 0.48
TNR: 0.57
MAGEL2
AUC: 0.559
p−value: 5.3e−05
strongest cutoff: 40
TPR: 0.48
TNR: 0.61
NDN
AUC: 0.511
p−value: 0.25
strongest cutoff: 429
TPR: 0.56
TNR: 0.48

D

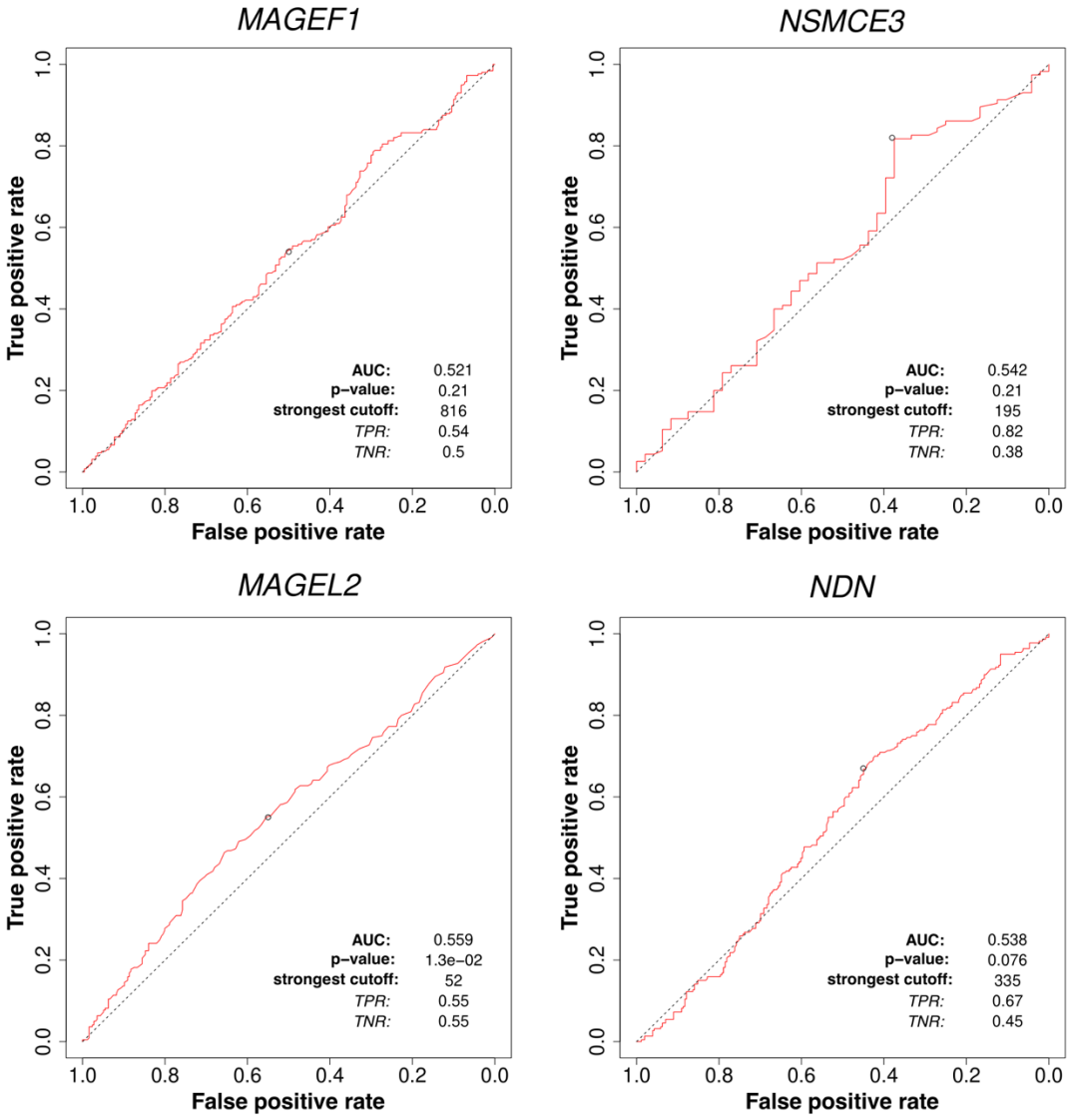
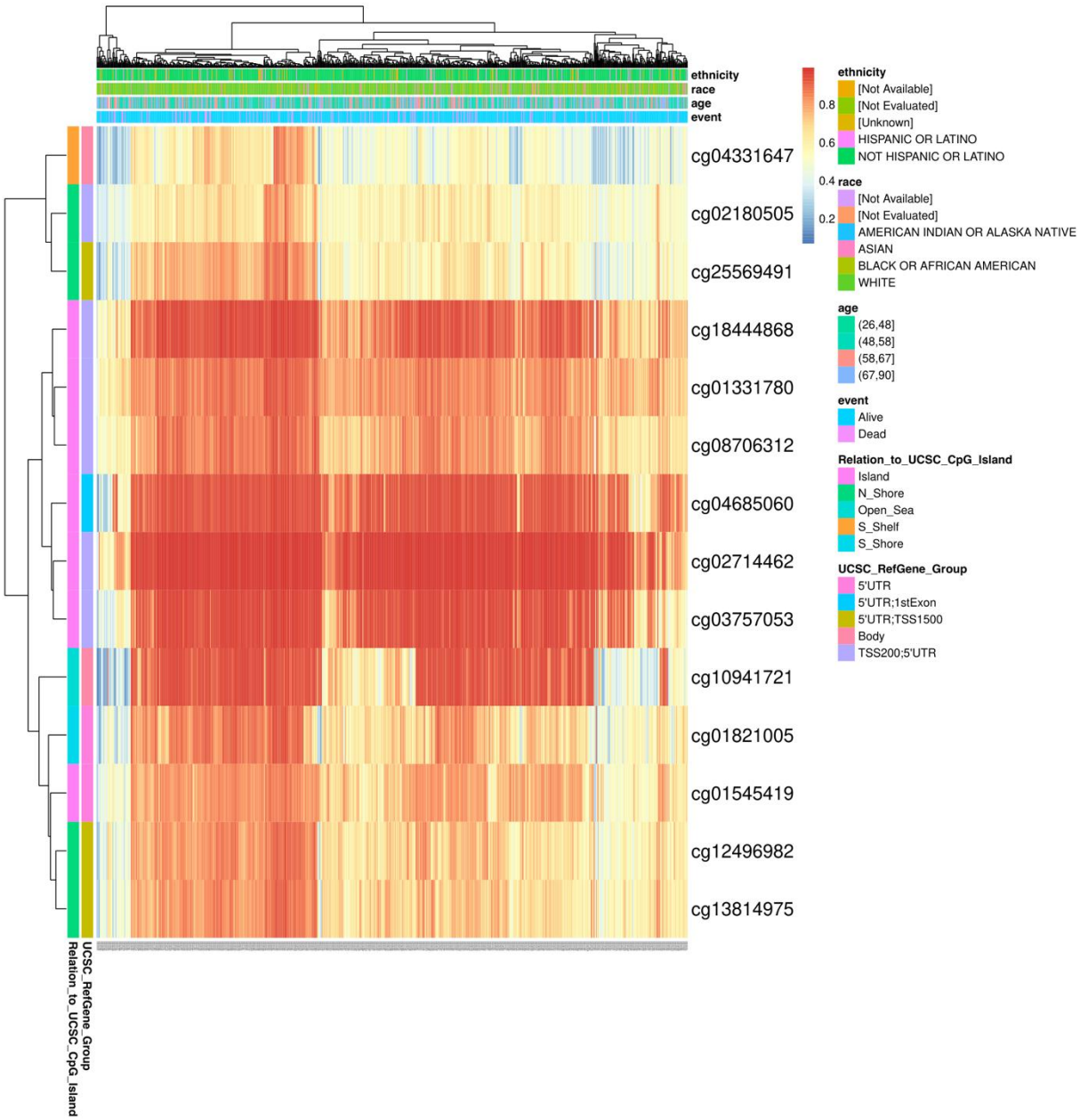


Figure 2. Prognostic value related to the mRNA levels of *MAGEF1*, *NSMCE3*, *MAGEL1*, and *NDN* in patients with breast cancer, as analyzed using KM plotter based on (A). OS, and (B). RFS. ROC plot related to the mRNA levels of *MAGEF1*, *NSMCE3*, *MAGEL1*, *NDN* and chemotherapy efficiency in patients with breast cancer, as analyzed using ROCPlotter based on (C). PCR and (D). 5-years RFS.

OS: Overall survival; RFS: Relapse-free survival; PCR: Pathological complete response

A


ethnicity
race
age
event
cg04331647
cg02180505
cg25569491
cg18444868
cg01331780
cg08706312
cg04685060
cg02714462
cg03757053
cg10941721
cg01821005
cg01545419
cg12496982
cg13814975
UCSC_RefGene_Group
Relation_to_UCSC_CpG_Island
0.8
0.6
0.4
0.2
ethnicity
[Not Available]
[Not Evaluated]
[Unknown]
HISPANIC OR LATINO
NOT HISPANIC OR LATINO
race
[Not Available]
[Not Evaluated]
AMERICAN INDIAN OR ALASKA NATIVE
ASIAN
BLACK OR AFRICAN AMERICAN
WHITE
age
(26,48]
(48,58]
(58,67]
(67,90]
event
Alive
Dead
Relation_to_UCSC_CpG_Island
Island
N_Shore
Open_Sea
S_Shelf
S_Shore
UCSC_RefGene_Group
5'UTR
5'UTR;1stExon
5'UTR;TSS1500
Body
TSS200;5'UTR

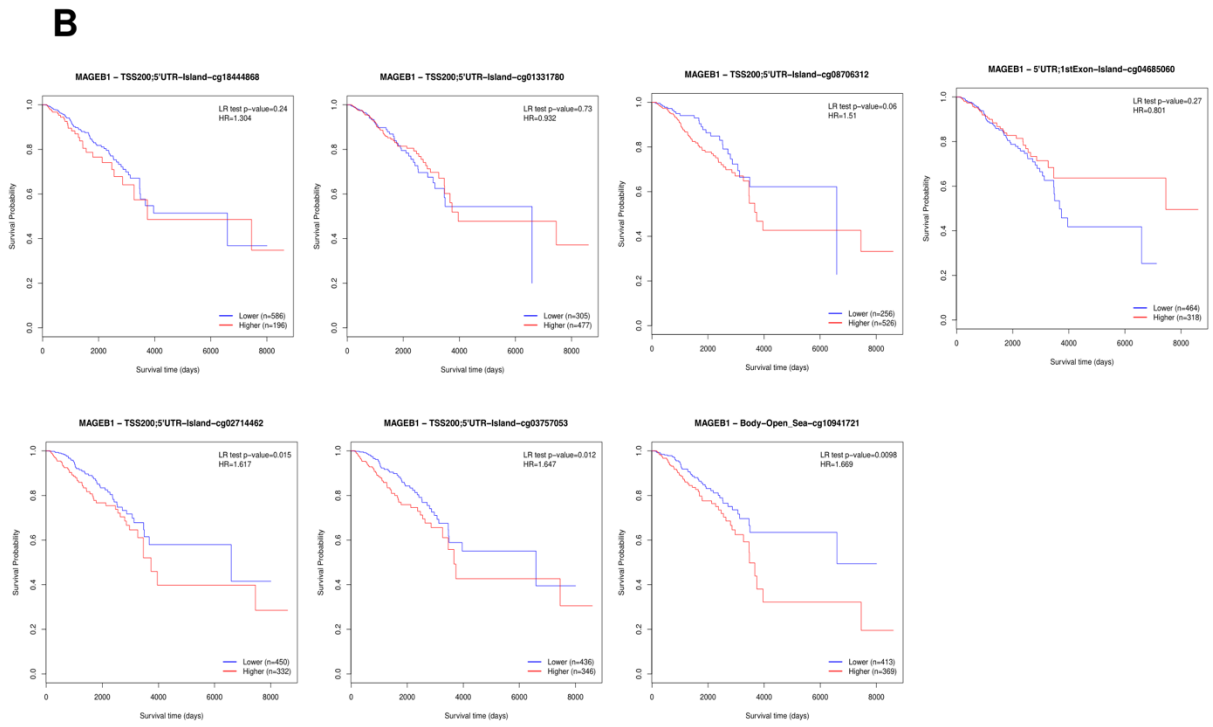


Figure 3. (A). Expression levels of *MAGEL1* or *MAGEB1* DNA methylation in breast cancer patients from TCGA study (2017), as determined by the MethSurv database, depicted as a heatmap. The heatmap was classified based on ethnicity, race, age, event, relation to UCSC CpG island, and UCSC reference gene group. (B). The prognostic significance of a particular CpG site within the *MAGEL1* or *MAGEB1* genes in patients with breast cancer from TCGA study (2017), assessed using the MethSurv platform.

TCGA: The Cancer Genome Atlas; UCSC: University of California, Santa Cruz

Table 1. Gene Ontology enrichment analysis of *MAGE* family genes, as analyzed using DAVID

No.	Biological process	Count	*P*-value	Genes
1	GO:0000122~negative regulation of transcription from RNA polymerase II promoter	39	9.82E-50	*MAGEH1, MAGEL2, MAGEA12, MAGEB18, MAGED1, MAGEB17, MAGEB16, MAGEF1, MAGEA9B, MAGEA9, TRO, MAGEA8, MAGED4B, MAGEA1, MAGEC3, MAGEA2, NDN, MAGEC1, MAGEC2, MAGEA11, MAGEB10, MAGEA6, MAGEA10, MAGEA3, MAGEA4, MAGEE1, MAGEE2, NSMCE3, MAGEA2B, MAGEB2, MAGED4, MAGEA5P, MAGEB3, MAGED2, MAGEB1, MAGEB6, MAGEA13P, MAGEB4, MAGEB5*
2	GO:0044257~cellular protein catabolic process	3	4.96E-04	*MAGEA2, MAGEC2, MAGEA2B*
3	GO:0051443~positive regulation of ubiquitin-protein transferase activity	3	8.37E-04	*MAGEA2, MAGEC2, MAGEA2B*
4	GO:0072331~signal transduction by p53 class mediator	2	0.00779144	*MAGEA2, MAGEA2B*
5	GO:0030163~protein catabolic process	3	0.00942586	*MAGEA2, MAGEC2, MAGEA2B*
6	GO:1901984~negative regulation of protein acetylation	2	0.01166496	*MAGEA2, MAGEA2B*
7	GO:0033234~negative regulation of protein sumoylation	2	0.02128446	*MAGEA2, MAGEA2B*
No.	**Cellular component**	**Count**	***P*-value**	**Genes**
1	GO:0005634~nucleus	39	6.96E-21	*MAGEH1, MAGEL2, MAGEA12, MAGEB18, MAGED1, MAGEB17, MAGEB16, MAGEF1, MAGEA9B, MAGEA9, TRO, MAGEA8, MAGED4B, MAGEA1, MAGEC3, MAGEA2, NDN, MAGEC1, MAGEC2, MAGEA11, MAGEB10, MAGEA6, MAGEA10, MAGEA3, MAGEA4, MAGEE1, MAGEE2, NSMCE3, MAGEA2B, MAGEB2, MAGED4, MAGEA5P, MAGEB3, MAGED2, MAGEB1, MAGEB6, MAGEA13P, MAGEB4, MAGEB5*
No.	**Molecular function**	**Count**	***P*-value**	**Genes**
1	GO:0042826~histone deacetylase binding	13	1.17E-18	*MAGEA12, MAGEA9B, MAGEA2B, MAGEA9, MAGEA8, MAGEA1, MAGEA2, MAGEA5P, MAGEA11,*

				MAGEA6, MAGEA10, MAGEA3, MAGEA4
2	GO:0005515~protein binding	30	4.69E-04	*MAGEH1, MAGEL2, MAGEA12, MAGEB18, MAGED1, MAGEF1, MAGEA9B, MAGEA9, TRO, MAGEA8, MAGED4B, MAGEA1, MAGEA2, NDN, MAGEC1, MAGEC2, MAGEA11, MAGEB10, MAGEA6, MAGEA3, MAGEA4, MAGEE1, NSMCE3, MAGEA2B, MAGEB2, MAGED4, MAGEB3, MAGED2, MAGEB6, MAGEB4*
3	GO:0031625~ubiquitin protein ligase binding	3	0.09334344	*MAGEA2, MAGEC2, MAGEA2B*

No.: Number; GO: Gene ontology, *MAGE*: Melanoma antigen gene; DAVID: The Database for annotation, visualization and integrated discovery

Table 2. Mutual exclusivity analysis of the *MAGE* family genes in breast cancer patients from The Metastatic Breast Cancer 2021 study, analysis was conducted using cBioportal

A	B	*P*-value	Tendency
MAGEL2	*NDN*	<0.001	Co-occurrence
MAGEF1	*NDN*	<0.001	Co-occurrence
MAGEF1	*MAGEL2*	<0.001	Co-occurrence
NSMCE3	*NDN*	<0.001	Co-occurrence
NSMCE3	*MAGEL2*	0.001	Co-occurrence
MAGEF1	*NSMCE3*	0.009	Co-occurrence

MAGE: Melanoma antigen gene

Table 3. The prognostic significance of a particular CpG site within the *MAGEL1* or *MAGEB1* genes in patients with breast cancer from TCGA study (2017), assessed using the MethSurv platform

Name	HR	CI	Wald_*P* value
MAGEB1-TSS200; 5'UTR-island-cg18444868	1.304	(0.845;2.011)	0.23
MAGEB1-TSS200; 5'UTR-island-cg01331780	0.932	(0.628;1.384)	0.73
MAGEB1-TSS200; 5'UTR-island-cg08706312	1.51	(0.971;2.349)	0.067
MAGEB1- 5'UTR;1stExon-island-cg04685060	0.801	(0.537;1.195)	0.28
MAGEB1-TSS200; 5'UTR-island-cg02714462	1.617	(1.097;2.384)	**0.015**
MAGEB1-TSS200; 5'UTR-island-cg03757053	1.647	(1.116;2.43)	**0.012**
MAGEB1-Body-Open_Sea-cg10941721	1.669	(1.129;2.468)	**0.01**

Bold indicates significant value ($P < 0.05$); TCGA: The Cancer Genome Atlas; HR: Hazard ratio; CI: Confidence interval

Table 4. The correlation between immune cell infiltration and the levels of *MAGEF1, NSMCE3, MAGEL1,* and *NDN* expression in breast cancer patients

Gene Name	Parameters	Purity	B cell	CD8+	CD4+	Macrophage	Neutrophil	Dendritic cell
MAGEF1	R	**0.213**	0.03	-0.141	-0.003	-0.021	0	-0.034
	P value	**1.17e-11**	3.37e-01	8.10e-06	9.27e-01	5.12e-01	9.94e-01	2.77e-01
NSMCE3	R	-0.116	-0.065	0.182	**0.233**	-0.04	0.128	0.011
	P value	2.38e-04	4.09e-02	7.74e-09	**9.04e-14**	2.10e-01	5.42e-05	7.36e-01
MAGEL1	R	0.007	-0.17	-0.108	0.142	-0.126	**0.214**	0.128
	P value	8.32e-01	7.31e-08	6.77e-04	7.41e-06	6.48e-05	**8.92e-12**	4.90e-05
NDN	R	**-0.476**	-0.158	**0.311**	-0.036	**0.377**	0.02	**0.209**
	P value	**2.14e-57**	5.15e-07	**1.08e-23**	2.62e-01	**6.01e-35**	5.37e-01	**2.58e-11**

The Spearman's Rank Correlation Coefficient values (R) indicated the following degrees of correlation strength: extremely weak (0.00-0.19), weak (0.20-0.39), moderate (0.40-0.69), strong (0.70-0.89), and very strong (0.90-1.0). A correlation and a significant value ($P < 0.05$) are denoted in bold. CD8: Cluster of differentiation 8; CD4: Cluster of differentiation 4