



Integrative co-expression network analysis of mRNA and miRNA profile identifies miRNA 664 as an inhibitor of key genes in Triple negative breast cancer

Divya Dhananjay Singh¹, Aditya Bardia²

¹.M.D., Master's Student, MMSCI Program, Harvard Medical School

².M.D., M.P.H., Mass General Cancer Center, Massachusetts General Hospital

***Corresponding Author:**

Divya Dhananjay Singh

¹.M.D., Master's Student in MMSCI Program at Harvard Medical School

Type of Publication: Original Research Paper

Conflicts of Interest: Nil

Abstract:

Triple-negative breast cancer (TNBC) is a specific subtype of breast carcinomas that do not express the oestrogen receptor (ER), progesterone receptor (PR), or human epidermal growth factor receptor 2 (HER-2), has clinical features that include high invasiveness, high metastatic potential, proneness to relapse, and poor prognosis. The lack of expression of molecular targets, lack of susceptibility to endocrine therapy or targeted treatment modalities make standardized TNBC treatment regimens lacking with respect to effectiveness. The disease afflicts younger women in larger numbers compared to other breast cancer subtypes and the cancers are more aggressive causing earlier morbidity and decrease survival. Therefore, the development of new TNBC treatment strategies has become an urgent clinical need.

Result:

Co-expression based network screening of messenger RNA (mRNA) profile data of Triple negative breast cancer identifies modules of genes consistently co-expressed. Among the different types of modules, the genes of three module TNBC6, TTNBC2, BCNT16 is enriched with gene ontology of immune response, adrenal lymphoid cells, Natural Killer (NK) cells and shows association with survival of TNBC patients. Also, these genes have a high binding score for the micro RNA (miRNA) regulators such as has-miR-664b-3p, has-miR-186-5p, has-miR-miR-548t-3p. These miRNAs are harbingers of disease aggressiveness in Triple negative breast cancer.

Conclusion: The study reveals the concept of regulation of immune response genes by miRNA, in particular, has-miR-664b-3p which is also involved in better survival of TNBC patients. In vivo validation of the concept obtained might either discover a novel gene target involved in immune response of TNBC aiding improved survival of patients and can help in identifying specific cell models and subtypes which are key for preclinical studies, developing new targeted agents and can inform therapy selection ushering in the often cited dream of personalized treatment regimens in TNBC with improved efficacy and treatment response.

Keywords: Triple Negative Breast Cancer, Chemotherapy, messenger RNA, micro RNA, novel targets.

Introduction

Breast cancer is the most common malignancy in women. Female breast cancer has now surpassed lung cancer as the leading cause of global cancer incidence in 2020, with an estimated 2.3 million new

cases, representing 11.7% of global cancer cases. It is the fifth leading cause of cancer mortality worldwide, with 685,000 deaths. Among women, breast cancer accounts for 1 in 4 cancer cases and for 1 in 6 cancer

deaths, ranking first for incidence in the vast majority of countries (159 of 185 countries) and for mortality in 110 countries. [1] Breast cancer is a heterogeneous group of diseases, each with characteristic aetiologies and specifically designed treatments. Expression of hormone receptors, oestrogen receptor (ER) and progesterone receptor (PR), or human epidermal growth factor receptor 2 (HER2) indicates responsiveness to therapies targeted at these proteins. However, for the approximately 20% of breast cancer patients with tumours negative for such markers, termed triple-negative breast cancer (TNBC), there is presently a lack of effective targeted treatment therapy options [2]. Furthermore, patients with TNBC are presented with worse overall prognoses, necessitating an improved understanding of this disease [3]. Gene expression profiling analysis often classifies TNBC as a subtype of basal-like breast cancer (BLBC). Approximately 56% of TNBC and BLBC gene expression profiles overlap. The overlap ratio can be as high as 60–90% between TNBC and BLBC, compared to only 11.5% between non-TNBC and BLBC [4,5]. Epidemiological data show that TNBC mostly occurs in premenopausal young women under 40 years old, accounting for approximately 15–20% of all breast cancer patients [6]. Compared with other subtypes of breast cancer, the survival time of TNBC patients is shorter, and the mortality rate is 40% within the first 5 years after diagnosis [7]. TNBC is highly invasive, and approximately 46% of TNBC patients will have distant metastasis. The median survival time after metastasis is only 13.3 months, and the recurrence rate after surgery is as high as 25% [8]. The metastasis often involves the brain and visceral organs. Distant metastasis mostly occurs in the 3rd year after diagnosis [8]. The average time to relapse in non-TNBC patients is 35–67 months, while that in TNBC patients is only 19–40 months. The mortality rate of TNBC patients within 3 months after recurrence is as high as 75% [9,10].

Intertumoral heterogeneity within TNBC has been revealed by studies by Lehmann *et al.* [11,12,13], which show that intrinsic molecular subtyping classifies TNBCs into between four and six subtypes variously labelled as basal-like 1 (BL1), basal-like 2

(BL2), mesenchymal (M), mesenchymal stem-like (MSL), immunomodulatory (IM), and luminal androgen receptor (LAR). Evidence has revealed that an abundance of either infiltrating lymphocytes or tumour-associated stromal cells within the sample was the primary determinant specifying the IM or MSL subtype, respectively, resulting in a consensus of four intrinsically-defined TNBC subtypes (BL1, BL2, M and LAR) [12]. Indicating the significant distinctions within TNBC, segregation into these categories yields distinctions in progression with BL1 patients showing significantly greater rates of pathological complete response (pCR) and BL2 patients showing significantly higher rates of distant relapse [12]. Further analysis of the molecular basis for these differences will help to uncover actionable targets to improve outcome.

Also, due to its special molecular phenotype, TNBC is not sensitive to endocrine therapy or molecular targeted therapy. Therefore, chemotherapy is the main systemic treatment, but the efficacy of conventional postoperative adjuvant chemoradiotherapy is poor. The residual metastatic lesions eventually will lead to tumoral recurrence [14]. Bevacizumab has been used in combination with chemotherapeutic drugs to treat TNBC in some countries, but the survival time of patients did not increase significantly [15]. Therefore, it is urgent to develop new treatment regimens and targets.

microRNAs (miRNAs), single-stranded RNA molecules capable of suppressing target gene expression by binding to the 3'UTRs of complementary mRNAs. They have emerged as key regulators of cell phenotype and as a potential therapeutic modality in breast cancer [16,17]. These non-coding RNAs that reduce the abundance and translational efficiency of mRNAs and play a major role in regulatory networks, influencing diverse biological processes through effects of individual miRNAs on translation of multiple mRNAs. Breast cancer imposes significant disruptions to the expression of many miRNAs and dozens of specific regulatory links between microRNAs and tumour suppressing or oncogenic mRNAs have been identified [17,18].

In order to explore these molecular determinants separating TNBC subtypes, we conducted an independent analysis of breast cancer datasets with the aim of characterizing microRNAs and messenger RNAs that significantly contribute to differences in gene expression between TNBC subtypes. The current study was started with aim to identify gene-gene associations conserved in wide range of triple negative breast cancer samples. The identification of module genes that are co-expressed and conserved in large group of samples might improve the robustness of the gene signature obtained as outcome. Stratification of TNBC into subclasses using new markers will identify new screening methods, prognostic factors, methodologies and perhaps targets for personalized therapies. Several recent studies have correlated miRNA expression with outcomes in TNBC using microarray or other high throughput technologies. mRNA expression profiles that subclassify TNBCs have also been reported in association with investigations of outcome, new molecular pathways and possible chemotherapy alternatives [19,20]. We have carried out an independent publicly available dataset analysis of miRNA and cancer-focused mRNA expression in normal and triple negative tumour tissues.

Materials & Methods

The mRNA profile data of TNBC were collected from the expression data base GEO (<https://www.ncbi.nlm.nih.gov/geo/>) and TCGA (<https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>). The probe ID is matched with the corresponding gene names from its respective platform file.

Construction of co-expression

Co-expression network of GSE76275, GSE103091, and TCGA was done by applying Weighted Gene Co-expression Network Analysis (WGCNA) package.

The script followed was as mentioned in WGCNA package. However, the outline of the script is

provided in supplementary methods. The overlapping genes across modules of independent profile data obtained from the network were screened by applying the “IF(ISERROR)” formula in Excel.

Survival Plot:

The survival plots were plotted using the MedCalc tool. The survival plot of the consistent module genes were plotted with the gene expression data GSE103091.

Screening of miRNA and target gene :

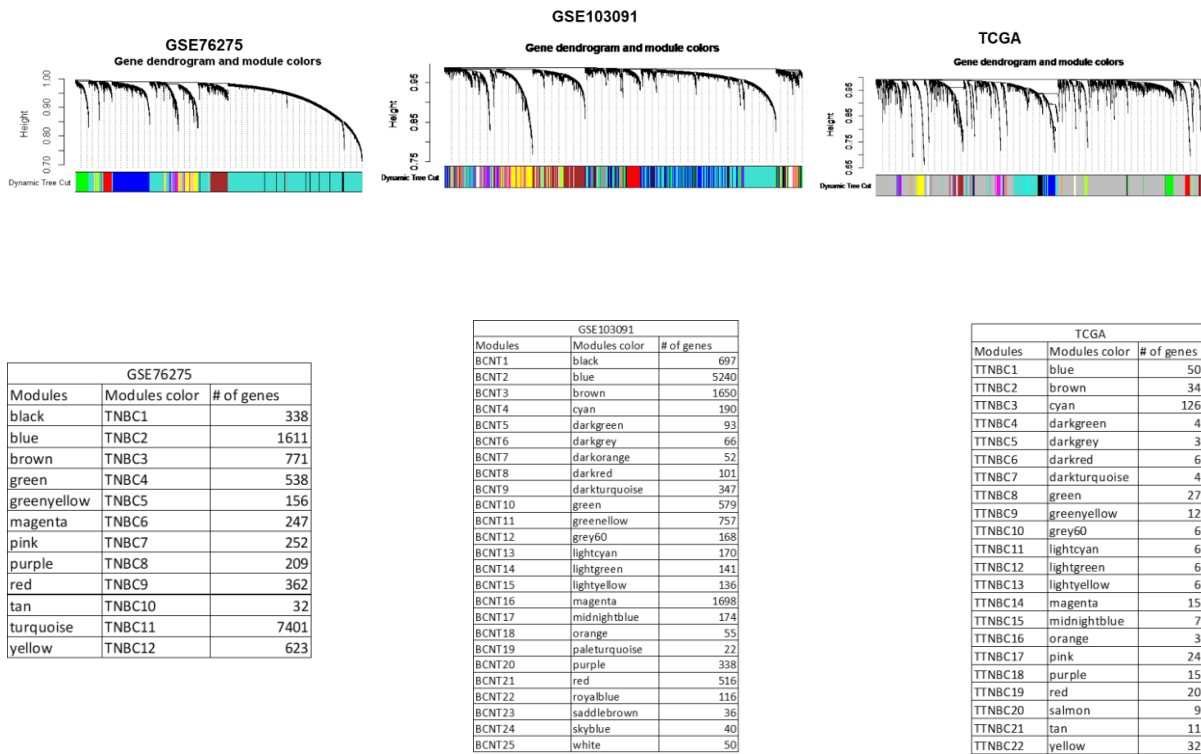
miRNA regulators of the module genes were identified by investigating the genes in the analysis tool miRDB. The module genes were screened during target mining to know the miRNAs of the gene targets.

Results

Co-expressed gene network conserved across multiple profile data of Triple negative breast cancer.

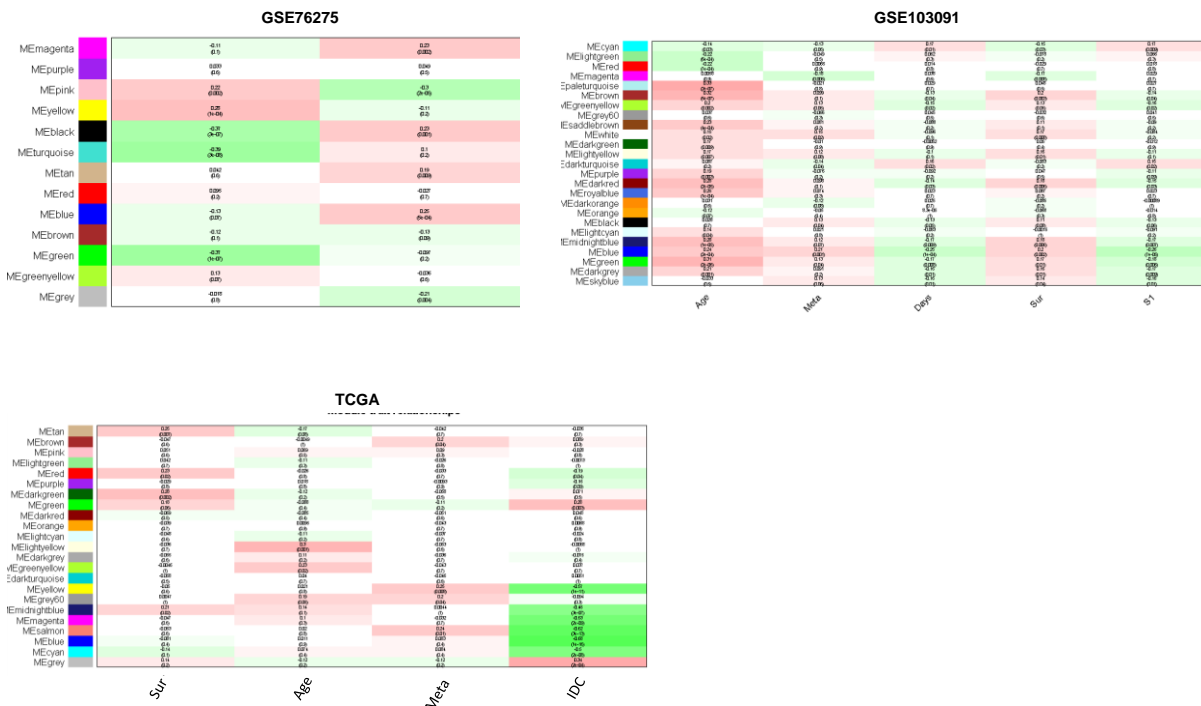
The current study was started with the aim to identify gene-gene associations conserved in wide range of triple negative breast cancer samples. Identification of module genes that are co-expressed and conserved in large group of samples might improve the robustness of the gene signature's obtained as an outcome. The mRNA profile data of triple negative breast cancer from studies such as GSE76275, GSE103091, and TCGA were used for investigation in the study. Gene co-expression network of mRNA profiles GSE76275, GSE103091, and TCGA were constructed by applying weighted gene co-expression network analysis (WGCNA) in R-platform. After deletion of the gene duplication and control probes of the data about 13432 genes from microarray platform were selected for investigation. The genes from each profile were clustered into 12 modules in GSE76275 and named as TNBC1 to TNBC12, followed by 25 modules in GSE103091 named as BCNT1 to BCNT25, followed by 22 modules in TCGA data and named as TTNBC1 to TTNBC22 (Figure1).

Figure1



Understanding the clinical and biological significance of the modules.

To understand the association of the co-expressed modules with the clinical conditions of the samples, module trait correlation was calculated. The module genes expression highly correlated with the clinical trait data were plotted as heat map plot (Fig 2A to C).



The modules BCNT1, BCNT2, BCNT3, BCNT8, BCNT15, BCNT17, BCNT25, show positive correlation with poor survival of patients, on the other side BCNT16 profile shows negative

correlation with poor survival of patients (Fig 2B). Similarly, the modules TTNBC2, TTNBC4, TTNBC8, TTNBC15, TTNBC19 were negatively associated with patients' survival, however TTNBC3, TTNBC6 are negatively correlated with patients' survival (Fig 2c). Modules TNBC1, TNBC2, TNBC6, TNBC10 are positively correlated with state of differentiation - Poorly differentiated samples (Fig 2A).

Screening of the modules of each profile individually obtain significant modules associated with clinical nature of samples. However, to increase the

robustness and accuracy of the findings, the module genes were screened for its consistent co-expression across three independent profile data. Hence the overlapping co-expressed genes of modules across all three independent profile data were sorted. The number of co-expressed genes conserved is from 36 to 222. The conserved co-expressed genes are clustered and named as BC1 to BC7. Among these cluster of modules, the BC2 cluster (consisting of TNBC6, TTNBC2, BCNT16) modules shows conserved co-expressed genes (114) (supplementary Table1)

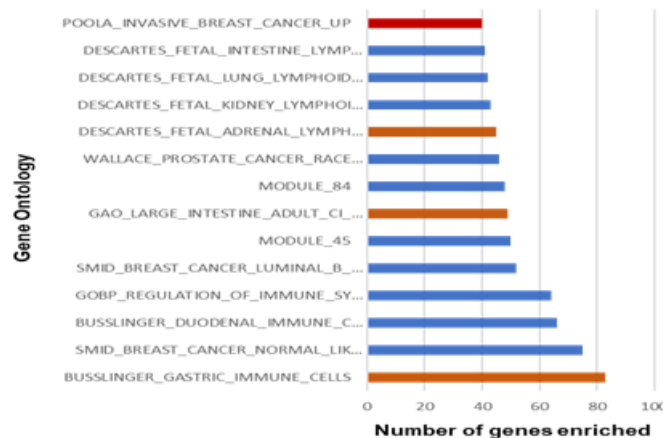
Conserved modules

TNBC2 BCNT21	222
TNBC2 BCNT1	124
TNBC6 TTNBC2 BCNT16	114
TTNBC2 BCNT16	74
TNBC12 BCNT9	70
TNBC3 TTNBC3	58
TNBC6 TTNBC11 BCNT16	56
TNBC5 BCNT3	53
TNBC7 BCNT15	39
TTNBC3, BCNT16	37
TNBC3 Her1	36

They also show consistent negative correlation with poor survival of patients hence these module genes were focused for further screening.

Gene ontology screening of genes of BC2 module cluster identifies the enrichment of immune response genes among the conserved co-expressed genes.

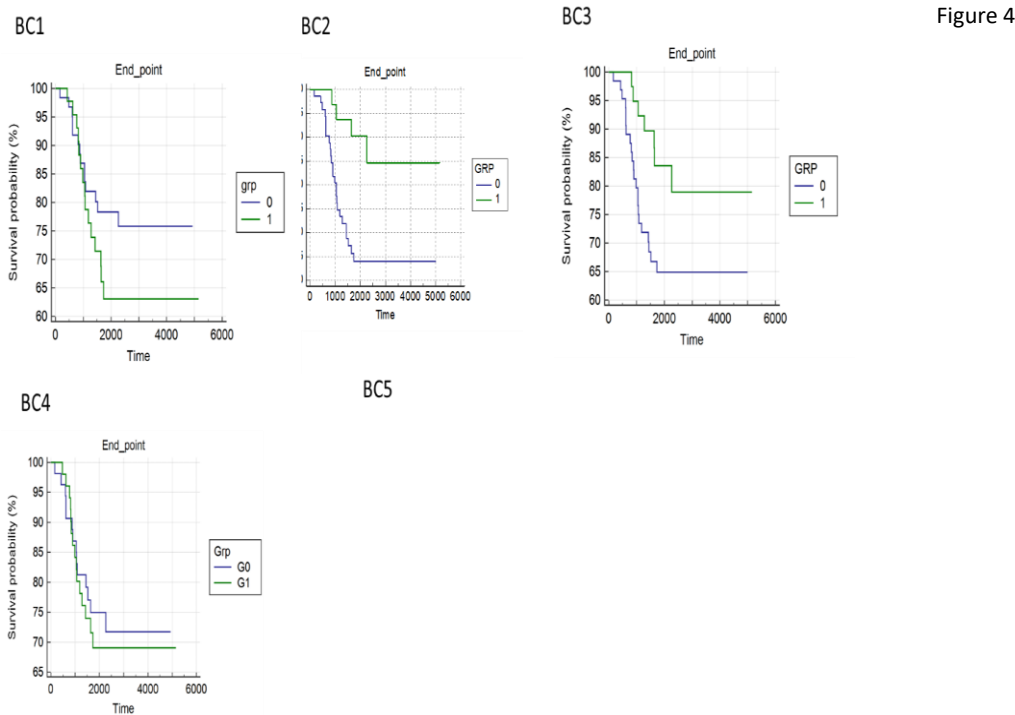
Gene ontology enrichment



Kaplan Meier survival analysis of conserved genes:

Though the focus of the study is fixed on BC2 cluster, the association of conserved genes with survival of patients was screened by applying Kaplan Meier analysis method. The Kaplan Meier analysis of all clusters

BC1 to BC7 was performed to identify the association between the modules and survival of patients. Survival data of GSE103091 was used in Kaplan Meier analysis. Higher expression of BC1 (p value 0.7), BC4, BC5 cluster module genes cause poor survival. However higher expression of BC2 (0.01), BC3 (0.05) genes leads to better survival of the patients. The level of significance is good for BC2 and BC3.

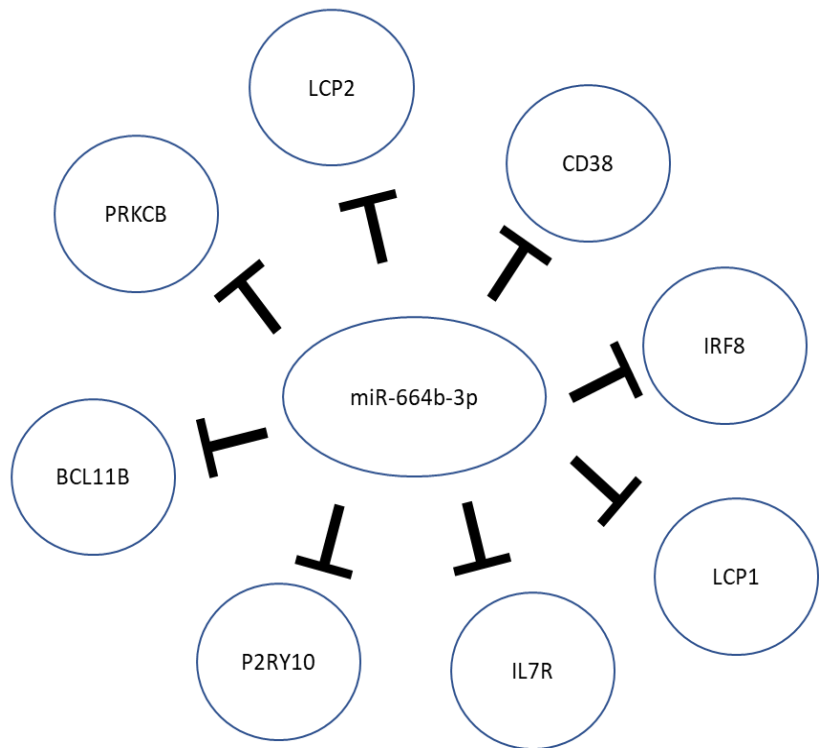
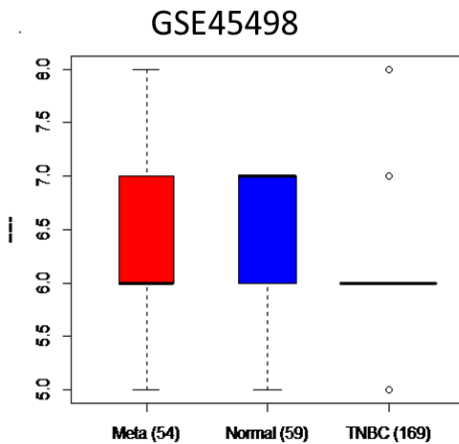


miRNA 664b-3p acting as major regulator of the immune response genes in TNBC.

In order to know the molecular regulatory mechanism of the immune responsive genes in triple negative breast cancer the link between the co-expressed genes and miRNA were screened. The conserved BC2 module genes were screened in miRDB database to know its corresponding miRNAs. Among the 114 conserved co-expressed genes 9 genes were targets of miRNA664b-3p. hsa-miR-186-5p and hsa-miR-548t-3p targets 8 genes. hsa-miR-125a-5p, hsa-miR-125b-5p, hsa-miR-3150b-3p targets 7 genes.

However, hsa-miR-548t-3p has the maximum score 100 and targets 8 genes. On the other hand miRNA664b-3p has maximum score 94.

Target Rank	Target Score	miRNA Name	Gene ID	Gene Sym	
52	94	hsa-miR-664b-3p	64919	BCL11B	
135	89		5579	PRKCB	
181	87		3937	LCP2	
199	86		952	CD38	
200	86		3394	IRF8	
346	80		3936	LCP1	
489	74		3575	IL7R	
541	72		27334	P2RY10	
962	60		3431	SP110	
46	95		hsa-miR-186-5p	3394	IRF8
62	93	10563		CXCL13	
128	90	6402		SELL	
474	75	2124		EVI2B	
538	72	27334		P2RY10	
539	72	1540		CYLD	
560	71	3431		SP110	
810	64	64919		BCL11B	
2	100	hsa-miR-548t-3p		1540	CYLD
180	87			11040	PIM2
247	84		64919	BCL11B	
534	72		3662	IRF4	
600	70		10225	CD96	
601	70		3003	GZMK	
735	66		55423	SIRPG	
888	62		51411	BIN2	



Coincidentally investigation of the level of expression of miRNA 664b-3p in TNBC confirms under-expression of the miRNA 664b-3p in TNBC compared to normal samples.

Discussion

Compared with other breast cancer subtypes, TNBC is a highly invasive breast cancer subtype and has a

high early recurrence rate. Patients usually relapse within 5 years after surgery, with a very poor overall prognosis. Due to negative expression of ER, PR, and HER2, TNBC is insensitive to endocrine treatment and targeted therapies. Only very limited treatment regimens are available for TNBC, with generally poor efficacy. New therapies are urgently needed.

The significance of microRNAs in cancer cell regulation is still a widely unexplored area. The Genomic Data Commons database is a monumental collection of genetic data for cancer research, encompassing The Cancer Genome Atlas (TCGA) and other projects, creating an opportunity for discovering new microRNA-mRNA pairs impacting cell proliferation. Indeed, attempts have been made in the past to build tools that could, automatize the search and were applied to TCGA datasets [22,23]. However, identification of candidate pairs is a challenging task due to the regulatory complexity and inter-dependence of mRNAs and microRNAs. Performing only correlation analysis between differentially expressed mRNAs and microRNAs followed by a network analysis might not be a satisfactory approach. It is worthwhile to note that the mRNA-microRNA pairs of therapeutic interest are not necessarily the most differentially expressed ones or the ones with highest anti-correlation or the ones in the centre of the target network.

In this study, we have combined correlation analysis and target analysis together with survival analysis, thus integrating statistical and biological relevance with practical relevance. This approach allowed us to perform the final selection of candidate pairs based on less stringent thresholds in each factor while still achieving a reasonable count of the candidates, which are additionally interesting from the therapeutic perspective for their possible impact on survival rates. A very recent publication analysing TCGA data [24] also performs survival analysis for selection of candidate mRNA-microRNA pairs although differentially expressed mRNAs were pre-filtered and only around 1% of statistically significant ones were analysed.

Many studies demonstrated that aberrant expression of miRNAs has been reported in many cancer types.

Considering the significant role of miRNA in regulation of cancer hallmarks, researchers are very much interested in targeting miRNA for cancer treatment. Interestingly the first RNAi drug specifically inhibiting hepatic synthesis of transthyretin in patients with hereditary transthyretin amyloidosis, has been successfully completed [25,26]. To these emerging concepts about the role of miRNA 664b-3p in TNBC, the current results add another important fact that a set of genes involved in immune response are targets of miRNA-664b-3p. It would be worthwhile to work on the concept of immune response gene and miRNA-664b-3p in future and identify a novel target for miRNA therapy in TNBC.

Also the result obtained raises the question why and how the target genes of miRNA664b-3p are involved in immune response is supporting better survival. Though the work has identified the link between the miRNA-664b-3p and immune genes in TNBC, the importance and mechanism behind the better survival of the patients requires validation by multiple screening strategies.

Undoubtedly, recently huge strides have been made in understanding TNBC as a disease with intrinsic molecular subtypes and immunological heterogeneity, recognizing the variety of clinical phenotypes. This new scenario demands an urgent comprehensive sub-classification that incorporates immune-molecular signatures for a more targeted and effective treatment regimen. Although, targeted inhibitors and checkpoint inhibitors have recently been incorporated in some settings, cytotoxic chemotherapy remains the mainstay therapy against TNBC, resulting in different outcomes for patients with similar clinicopathologic features.

A more complete accessible panel of immunohistochemical molecular subtypes has improved decision making in the treatment of TNBC. Additionally, in many cases, more precise molecular classification of tumours has been proposed to predict survival and response to chemotherapy, allowing for

personalized approaches, such as the need for dose escalation and incorporation of new antitumor agents into the standard regimen, and for new treatment options, such as CAR-T immune cell therapy, checkpoint inhibitors, and molecular targeted inhibitors.

The study has its limitations based on the fact that Publicly available databases are unwieldy to work with as data is stored in the raw form and is not analysis-ready. This causes the utilization of significant resources and expertise for data cleaning and management. Most of the data sources lacked follow-up and outcome data and the available information was of limited granularity and de-identification precluded the collection of additional variables of interest, increasing the risk of unmeasured confounding.

Formerly considered a disease unapproachable with molecular therapy, TNBC has recently been the center of successful investigations for the incorporation of new targeted therapies due to intrinsic molecular TNBC subtyping and accurate classification and prediction of prognosis improvements. Considering the proposed subtypes and their molecular variations as defined by specific biomarkers and the current chemotherapy, immunotherapy, and targeted inhibitor combination options, great advances have been achieved in TNBC treatment.

Conclusion

The concept of regulation of immune response genes by miRNA helps further classify and define the disease as well as show correlation with aggressiveness of cancer progression which can be identified early in individual patients. Considering the significant role of miRNA in regulation of cancer hallmarks, researchers are very much interested in targeting miRNA for the dual role it plays, that of the Diagnostic Biomarkers as well as Therapeutic Target. Future functional analysis studies with exploratory work done “in silico” will help in confirmation of the same. The discovery of the immune response targets

in TNBC presents an interesting focus for further elucidation.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021 May;71(3):209-249.
2. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. *N Engl J Med.* 2010;363(20):1938–48.
3. Liedtke C, Mazouni C, Hess KR, Andre F, Tordai A, Mejia JA, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol.* 2008;26(8):1275–81.
4. Perou CM, et al. Molecular portraits of human breast tumours. *Nature.* 2000;406(6797):747–52.
5. Prat A, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast.* 2015;24(Suppl 2):S26–35.
6. Morris GJ, et al. Differences in breast carcinoma characteristics in newly diagnosed African-American and Caucasian patients: a single-institution compilation compared with the National Cancer Institute’s Surveillance, Epidemiology, and End Results database. *Cancer.* 2007;110(4):876–84.
7. Dent R, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res.* 2007;13(15 Pt 1):4429–34.
8. Lin NU, et al. Sites of distant recurrence and clinical outcomes in patients with metastatic triple-negative breast cancer: high incidence of central nervous system metastases. *Cancer.* 2008;113(10):2638–45.
9. Zhang L, et al. Androgen receptor, EGFR, and BRCA1 as biomarkers in triple-negative breast cancer: a meta-analysis. *Biomed Res Int.* 2015;2015:357485.
10. Gluz O, et al. Triple-negative breast cancer--current status and future directions. *Ann Oncol.* 2009;20(12):1913–27.
11. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest.* 2011;121(7):2750–67.
12. Lehmann BD, Jovanovic B, Chen X, Estrada MV, Johnson KN, Shyr Y, et al. Refinement of triple-negative breast cancer molecular subtypes: implications for neoadjuvant chemotherapy selection. *PLoS One.* 2016;11(6):e0157368.
13. Jiang YZ, Ma D, Suo C, Shi J, Xue M, Hu X, et al. Genomic and transcriptomic landscape of triple-

- negative breast cancers: subtypes and treatment strategies. *Cancer Cell*. 2019;35(3):428–40.e5.
14. Chaudhary LN, Wilkinson KH, Kong A. Triple-negative breast cancer: who should receive neoadjuvant chemotherapy? *Surg Oncol Clin N Am*. 2018;27(1):141–53.
 15. Collignon J, et al. Triple-negative breast cancer: treatment challenges and solutions. *Breast Cancer (Dove Med Press)*. 2016;8:93–107.
 16. Rupaimoole R, Slack FJ. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nat Rev Drug Discov*. 2017;16(3):203–22.
 17. Kaboli PJ, Rahmat A, Ismail P, Ling KH. MicroRNA-based therapy and breast cancer: a comprehensive review of novel therapeutic strategies from diagnosis to treatment. *Pharmacol Res*. 2015;97:104–21.
 18. Goh JN, Loo SY, Datta A, Siveen KS, Yap WN, Cai W, et al. microRNAs in breast cancer: regulatory roles governing the hallmarks of cancer. *Biol Rev Camb Philos Soc*. 2016;91(2):409–28.
 19. Carey LA, Dees EC, Sawyer L, Gatti L, Moore DT, et al. (2007) The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. *Clin Cancer Res* 13: 2329–2334.
 20. Martin M, Romero A, Cheang MC, López García-Asenjo JA, García-Saenz JA, et al. (2011) Genomic predictors of response to doxorubicin versus docetaxel in primary breast cancer. *Breast Cancer Res Treat* 128: 127–136.
 21. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
 22. Vila-Casadesus M, Gironella M, Lozano JJ. MiRComb: an R package to analyse miRNA-mRNA interactions. Examples across five digestive cancers. *PLoS One*. 2016;11(3):e0151127.
 23. Bai YS, Ding LZ, Baker S, Bai JM, Rath E, Jiang F, et al. Dissecting the biological relationship between TCGA miRNA and mRNA sequencing data using MmiRNA-Viewer. *BMC Bioinformatics*. 2016;17(Suppl 13):336.
 24. Wu P, Xiao YY, Guo T, Wang YT, Liao SS, Chen L, et al. Identifying miRNA-mRNA pairs and novel miRNAs from hepatocellular carcinoma miRNomes and TCGA database. *J Cancer*. 2019;10(11):2552–9.
 25. Cortez MA, Anfossi S, Ramapriyan R, et al. Role of miRNAs in immune responses and immunotherapy in cancer. *Genes Chromosomes Cancer*. 2019;58(4):244–253.
 26. Jin, Qiu, et al. “Mir-664B-3P Inhibited Triple-Negative Breast Cancer Cell Growth via Targeting BRIP1.” 2021.