Correlation and Meta-Analysis of HER2 in Each Stage of Breast Cancer

K N Ramanto¹, David Agustriawan†, A A Parikesit¹, Rizky Nurdiansyah¹ and Muhammad Z A Nasution¹

¹Department of Bioinformatics, School of Life Sciences, Indonesia International Institute for Life Sciences, Jl. Pulomas Barat Kav.88 Jakarta 13210 Indonesia.

*Corresponding Author: david.agustriawan@i3l.ac.id

Abstract. Human Epidermal growth Receptor 2 (HER2) plays an essential role in the pathogenesis of breast cancer. Previous studies have shown that HER2 overexpression is associated with a higher histological tumor grade. Furthermore, the overexpression of HER2 is associated with miRNAs expression. However, the interaction of HER2 and miRNAs in each stage of breast cancer remain unclear. The objective of this study is to identify the relationship between HER2 and miRNAs expression in each stage of breast cancer. Thirteen miRNAs were selected based on literature evidence that showed their interaction with the overexpression of HER2. The correlation and meta-analysis were done by using the R programming language. The result of this study showed hsa-mir-10b has the highest correlation value in stage two and three breast cancer. On the other hand, hsa-let-7f-2 has the highest correlation value in stage one breast cancer. The result of meta-analysis also showed a significant level of p-value (0.007). These results indicate the downregulation miRNAs only occur in a specific stage of breast cancer.

1. Introduction

Human epidermal growth factor receptor 2 (HER2) or known as ErbB2 is a member of type I transmembrane growth factor receptor family protein that function to activate intracellular signaling pathways in response to extracellular signals, for example Memo-Rhoa-mDia1 signaling pathway that control microtubules, the actin network, and adhesion site formation in migrating cell [1,2]. The HER2 protein product is having tyrosine kinase activity that similar to Epidermal growth receptor (EGFR). Besides, dimerization of the receptor in the autophosphorylation of tyrosine within the cytoplasmic domain of the receptor initiates various signaling pathways leading to cell proliferation and tumorigenesis [3].

HER2 plays an essential role in the pathogenesis of several cancers, including breast cancer. In 1987, Dr. Slamon discovered a genetic link between HER2 and breast cancer for the first time. His research showed that HER2 protein is present at high level in about 15-30% breast cancers [4]. Furthermore, breast cancer cells can have up to 25-50 copies of HER2 gene and up to 40-100 folds increase in HER2 protein expression resulting in up to 2 million receptors expressed at the tumor cell surface [1,2]. Breast cancer with the overexpression of HER2 protein is called HER2-positive (HER2+) in the pathology report.
Microarray-based expression profiling of miRNAs that associated with HER2 in breast cancer has been done among various studies. Mattie et al. identified 43 miRNAs that are significantly lower in HER2+ from 20 stages I-II breast cancer biopsies including 13 cases of HER2 overexpression. Among these miRNAs inversely correlated with HER2 overexpression, let-7f, let-7g, miR-107, miR-10b, miR-126, miR-154 and miR-195 [5]. A similar study conducted by Lowery et al. has identified a signature of five miRNAs (miR-520d, miR-181c, miR-302c, miR-376b, and miR-30ep) [6]. Among them, decreased expression of miR-181c and the increased expression of miR-520d and miR-376b are associated with HER2+ breast cancer.

Another method to analyze the association between miRNAs and HER2 gene is correlation analysis that usually done by using The Cancer Genome Atlas (TCGA) dataset. TCGA is a publicly funded project that aims to catalog and discover significant cancer-causing genome alterations in large cohorts of 30 human tumors through large-scale genome sequencing and integrated multi-dimensional analysis [7]. In 2014, Bailey et al. conducted the correlation analysis with TCGA dataset to find the miRNAs that clinically relevant to the HER2 gene in breast cancer. The result showed miR-let-7c, miR-99a, and miR-125b are correlated with HER2 [8]. Although these studies have identified miRNAs that associated with the overexpression of HER2, the association between HER2 and miRNAs in each stage of breast cancer is remains unclear.

Since the characteristic of breast cancer is different in each stage, then there is a possibility where the downregulation or upregulation of miRNAs only happens in a specific stage. According to the American Joint Committee on Cancer (AJCC) staging guidelines, breast cancer is divided into four stages. Stage I describes that the tumor measures up to 2 centimeters (cm); Stage II describes the tumor is larger than 2 cm but no larger than 5 cm and has not spread to the axillary lymph nodes; Stage II describe the tumor is larger than 5 cm and has spread to 1 to 3 axillary lymph nodes; Stage IV describes that breast cancer that has spread beyond the breast and nearby lymph nodes [9]. Whether there are differences in the regulation of miRNAs in each stage or not need to be further explored.

This study is designed to identify the relationship between HER2 and 13 selected miRNAs in each stage of breast cancer. These 13 miRNAs were selected from different journals article that used microarray-based expression profiling as their method (table 1). The correlation analysis was done by using TCGA dataset and validated by using meta-analysis, miRTARBase, and RNAhybrid.

2. Methods
All datasets are obtained from The Cancer Genome Atlas (https://portal.gdc.cancer.gov/) dataset and downloaded with R studio. R packages that used for the study are: (i) TCGA2STAT is used to download all the datasets [12]; (ii) dplyr is used to manipulate the dataset [13]; and (iii) metacor is used to perform a meta-analysis [15].

2.1. miRNAs Selection
Journals Article were collected from google scholar. The searching was conducted from 9th October 2018 until 21st October 2018 with "miRNAs and HER2 in breast cancer." as the keywords. Each of these journals was implemented microarray-based profiling analysis to determine the association between miRNAs and HER2 gene. The most downregulated miRNAs from each journal were selected for this study. The Details of selected miRNAs shown in table 1.

2.2. Data Preparation
First of all, RNA and miRNA expression datasets were downloaded using TCGA2STAT package [12]. Each of the data objects will contain three different datasets, which are: (i) dat consists of only miRNA or RNA expression dataset; (ii) clinical consist of single clinical dataset and; (iii) merged.dat is the combination of expression and clinical dataset stored in one data object, but normal samples are not available in this dataset [12].
Next, HER2 and 13 miRNAs were extracted from merged.dat by using the dplyr package [13]. All of the datasets were grouped based on the pathological stage from stage I to stage IV. The stage that categorized as X and NA are excluded in this study. Stage X indicates that cancer is detectable, but diagnostician cannot determine the pathological stage of cancer. In the other hand, NA means that the primary tumor cannot be found [14]. Since this study only focuses on analyzing the miRNAs and HER2 in each stage of breast cancer, then the subtype of each stage (a, b, and, c) will be ignored in this study. At the end of data preparation, each stage of RNA and miRNA datasets were matched and merged based on the sample barcode.

Table 1. Selected miRNAs.

| Study            | MiRNAs              | Reference |
|------------------|---------------------|-----------|
| Mattie et al. (2006) | hsa-let-7f-1      | [5]       |
|                  | hsa-let-7f-2       |           |
|                  | hsa-let-7g         |           |
|                  | hsa-mir-107        |           |
|                  | hsa-mir-10b        |           |
|                  | hsa-mir-126        |           |
|                  | hsa-mir-154        |           |
|                  | hsa-mir-195        |           |
| Lowery et al. (2009) | hsa-mir-520d      | [6]       |
|                  | hsa-mir-181c       |           |
|                  | hsa-mir-30e        |           |
| Wu et al. (2012)   | hsa-mir-200a       | [10]      |
| Chen et al. (2009)  | hsa-mir-559        | [11]      |

2.3. Correlation Analysis
Spearman correlation analysis was done in each stage for HER2 and 13 selected miRNAs. After that, each correlation datasets were transposed to tidy up the outlook of the datasets. The sample size and column names were also inputted in correlation dataset. At the end of this step, each stage of breast cancer will have correlation dataset.

2.4. Meta-Analysis
The correlation dataset from stage I to stage III were combined and edited in Microsoft Excel. The meta-analysis was performed by using a combined dataset and metacor package [15]. Since each stage of cancer cannot be treated in the same way, then the random effect model was used to perform the meta-analysis. However, the correlation dataset of stage IV breast cancer was not used because it contains bias correlation values due to limited sample available in stage IV.

2.5. Result Validation
The selected miRNAs were validated using miRTARBase (http://mirtarbase.mbc.nctu.edu.tw); it’s an open source website that records all validated mRNA and miRNA interaction [16]. Furthermore, the interaction between HER2 and selected miRNAs were simulated by using RNAhybrid (https://bibiserv2.ccebitec.uni-bielefeld.de/); a website that can predict the minimum free energy (mfe) required for miRNA and mRNA to form [17]. The sequence of HER2 and selected miRNAs are needed to operate RNAhybrid. The sequence of HER2 and miRNAs are obtained from NCBI (https://www.ncbi.nlm.nih.gov).

3. Result and Discussions
In this study, 452 samples used and divided into four stages of breast cancer: 84 samples categorized as stage I; 270 samples classified as stage II; 95 samples organized as stage III and; 3 samples
categorized as stage IV. Unfortunately, the sample size for stage IV is less than 20 and give a biased result in correlation and meta-analysis cannot be performed in stage IV.

3.1. Correlation Analysis
As shown in table 2, the correlation of 13 selected miRNAs was not always negative in all stages. There are only four miRNAs (miR-107, miR-30e, miR-559, and miR-520d) that are negatively correlated with HER2 from stage I to stage I

In stage I, let-7f-2 showed the highest correlation compare to other miRNAs with the rho value of 0.156, but significantly decreased in stage II. let-7 family is known as miRNAs that act as tumor suppressor in breast cancer [18]. However, the expression of the let-7f family is lower in HER2+ breast cancer. The overexpression of HER2 activate the Extracellular Regulated Kinase (ERK) signaling pathway that lead to the downregulation of let-7f expression [19]. The rho value of let-7f indicates that the activation of the ERK signaling pathway and inhibition of let-7f mostly occur in stage II of breast cancer. Shao et al. previously reported the Stage II of breast cancer has the highest rho value with ERK signaling pathway which leads to inhibition of let-7f family [20].

### Table 2. The result of correlation analysis in stage I to stage III in breast cancer.

| miRNAs          | Stage I n = 84 | Stage II n = 270 | Stage III n = 95 |
|-----------------|---------------|-----------------|-----------------|
| hsa-let-7f-1    | 0.1           | -0.1            | 0.0             |
| hsa-let-7f-2    | 0.1           | -0.1            | 0.1             |
| hsa-let-7g      | -0.2          | -0.2            | 0.0             |
| hsa-mir-107     | -0.2          | 0.1             | -0.1            |
| hsa-mir-10b     | 0.0           | 0.1             | 0.1             |
| hsa-mir-126     | -0.1          | 0.0             | -0.0            |
| hsa-mir-154     | 0.1           | 0.1             | -0.0            |
| hsa-mir-181c    | 0.1           | -0.1            | -0.2            |
| hsa-mir-195     | 0.1           | -0.1            | -0.2            |
| hsa-mir-200a    | 0.1           | -0.2            | -0.1            |
| hsa-mir-30c     | -0.1          | -0.2            | -0.2            |
| hsa-mir-559     | -0.1          | -0.1            | -0.2            |
| hsa-mir-520d    | -0.1          | -0.1            | -0.1            |

Meanwhile, mir-10b was showed high correlation in stage II and stage III of breast cancer with rho value of 0.137 and 0.166 respectively. Furthermore, there is no negative correlation found in any stage of breast cancer. These findings were supported by research conducted by Zhang et al., that saw the higher expression of mir-10b in HER2+ sample. The expression of mir-10b is significantly increased in stage II to stage IV of breast cancer [21]. However, our result is different with Mattie et al., and Guo et al. that conducted microarray-based profiling of mir-10b and found a negative correlation between HER2 and mir-10b [5,22]. Thus, more functional and clinical study is needed to explore the role of mir-10b in breast cancer.

3.2. Meta-analysis
Statistical results are determining the significant level of HER2 and 13 selected miRNAs in stage I to stage III of breast cancer (table 3).
Table 3. Statistical results from the meta-analysis.

| Statistical Analysis | Results  |
|----------------------|----------|
| z mean               | -0.046   |
| r mean               | -0.046   |
| z mean se            | 0.019    |
| z mean lower         | -0.009   |
| z mean upper         | -0.084   |
| r mean lower         | -0.009   |
| r mean upper         | -0.084   |
| p                    | 0.007    |

The statistical result can be seen in table 3, where the p-value is less than 0.05 meaning the interaction between HER2 and 13 selected miRNAs in this study is significant. The mean if r value was indicated that overall of miRNAs are downregulated in breast cancer. Standard estimate error was also reporting that the meta-analysis is accurate [23]. These meta-analysis results are showed that there is heterogeneity in this study meaning the various outcomes of correlation analysis is diverse among each stage. The results also indicate that the overall correlation is statistically significant.

3.3. Result Validation
There is a limited number of experiments that validate the interaction between miRNAs and HER2 in miRTARBase. From 13 selected miRNAs, Only the interaction between hsa-mir-559 and HER2 recorded in the miRTARBase [16]. Therefore, further validation is conducted with RNAhybrid.

The validation with RNAhybrid was done to only miRNAs that have a negative correlation at least in one stage, so hsa-mir-10b was excluded from RNAhybrid validation. This was done to confirm the downregulated miRNAs in breast cancer as the result of their interaction with HER2. The lowest prediction of mfe for HER2 is with hsa-mir-181c with the mfe of -104.3 kcal/mol. Although all miRNAs with HER2 mfe are less than zero that indicating the spontaneous reaction, the significance value is all equal to one [17]. Further wet lab study is needed to confirm the association between selected miRNAs and HER2 in each stage.

3.4. Limitations
This study showed the interaction between HER2 and 13 selected miRNAs expression is significant in each stage of breast cancer. However, there are some limitation that needs to consider: (i) the result of correlation analysis in stage IV is bias due to limited sample size; (ii) the sample size in each stage of breast cancer is different from each other; (iii) the dataset only limited to TCGA database and; (iv) there not enough dataset for normal sample, so the author cannot compare the result between a normal sample and tumor sample.

4. Conclusion
The result of correlation analysis found only four miRNAs that have a negative correlation with HER2 in all stages and indicated that the downregulation of miRNAs in breast cancer occurs in a specific stage of cancer. Besides, the result validation with meta-analysis and RNAhybrid also suggest a possible interaction of HER2 and 13 selected miRNAs. Even though this study found a significant result, further research is needed to confirm the interaction of HER2 and miRNAs in each stage.

Acknowledgment
The authors would like to thanks to the Institute for Research and Community Services of Indonesia International Institute for Life Sciences (I3L) for funding this project.

References
[1] Moasser, M. M. (2007). The oncogene HER2: Its signaling and transforming functions and its role in human cancer pathogenesis. *Oncogene*. 26(45) pp.6469-6487.
[2] Mattie MD, Benz CC, Bowers J, et al. (2006) Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies. *Methods of Cancer Diagnosis, Therapy, and Prognosis, 5*(24).

[3] Iqbal, N., & Iqbal, N. (2014). Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications. *Molecular Biology International, 2014*, pp.1-9.

[4] Zaouï, K., Honoré, S., Isnardon, D., Braguer, D., & Badache, A. (2008). Memo–RhoA–mDia1 signaling controls microtubules, the actin network, and adhesion site formation in migrating cells. *The Journal of Cell Biology, 183*(3), 401-408.

[5] Mattie MD, Benz CC, Bowers J, et al. (2006) Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies. *Methods of Cancer Diagnosis, Therapy, and Prognosis, 5*(24).

[6] Lowery, A. J., Miller, N., Devaney, A., Mcneill, R. E., Davoren, P. A., Lemetre, C., Benes, V., Schmidt, S., Blake, J., Ball, G., Kerin, M. J. (2009). MicroRNA signatures predict estrogen receptor, progesterone receptor and HER2/neu receptor status in breast cancer. *Breast Cancer Research, 11*(3).

[7] Grossman, Robert L., Heath, Allison P., Ferretti, Vincent, Varmus, Harold E., Lowy, Douglas R., Kibbe, Warren A., Staudt, Louis M. (2016) Toward a Shared Vision for Cancer Genomic Data. *Methods of Cancer Diagnosis, Therapy, and Prognosis, 5*(24).

[8] Bailey, S. T., Westerling, T., & Brown, M. (2014). Loss of Estrogen-Regulated microRNA Expression Increases HER2 Signaling and Is Prognostic of Poor Outcome in Luminal Breast Cancer. *Cancer Research, 75*(2), pp.436-445.

[9] Breast. (2010). *AJCC Cancer Staging Handbook*,419-460.

[10] Wu, X., Somlo, G., Palomares, M., Yen, Y., Rossi, J., Gao, H., & Wang, S. E. (2012). Abstract A9: De novo sequencing of circulating microRNAs identifies novel markers predicting the clinical outcome of locally advanced breast cancer. *Cancer Research, 72*(2 Supplement).

[11] Chen, H., Sun, J. G., Cao, X., Ma, X., Xu, J., Luo, F., & Chen, Z. (2009). Preliminary validation of ERBB2 expression regulated by miR-548d-3p and miR-559. *Biochemical and Biophysical Research Communications, 385*(4), pp.596-600.

[12] Ying-Wooi Wan, Genevera I. Allen, Matthew L. Anderson and Zhandong Liu (2015). TCGA2STAT: Simple TCGA Data Access for Integrated Statistical Analysis in R. R package version 1.2. https://CRAN.R-project.org/package=TCGA2STAT

[13] Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2018). dplyr: A Grammar of Data Manipulation. R package version 0.7.8. https://CRAN.R-project.org/package=dplyr

[14] National Cancer Institute. (2015, June 20). Staging. Retrieved February 13, 2019, from https://www.cancer.gov/about-cancer/diagnosis-staging/staging

[15] Etienne Laliberté (2011). metacor: Meta-analysis of correlation coefficients. R package version 1.0-2. https://CRAN.R-project.org/package=metacor

[16] Chou, C., Chang, N., Shrestha, S., Hsu, S., Lin, Y., Lee, W., , Huang, H. (2015). MiRTarBase 2016: Updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Research, 44*(D1).

[17] Kruger, J., & Rehmsmeier, M. (2006). RNAhybrid: MicroRNA target prediction easy, fast and flexible. *Nucleic Acids Research, 34* pp.W451-454.

[18] Thammaiah, C. K., & Jayaram, S. (2016). Role of let-7 family microRNA in breast cancer. *Non-coding RNA Research,1*(1), 77-82.

[19] Liu, D., Deng, Q., Sun, L., Wang, T., Yang, Z., Chen, H., Guo, L., Liu, Y., Ma, Y., Guo, N., Shi, M. (2015). A Her2let-7-β2-AR circuit affects prognosis in patients with Her2-positive breast cancer. *BMC Cancer, 15*(1).

[20] Shao, G., Wang, M., Fan, X., Zhong, L., Ji, S., Sang, G., & Wang, S. (2018). Correlation Between Raf/MEK/ERK Signaling Pathway and Clinicopathological Features and Prognosis for Patients With Breast Cancer Having Axillary Lymph Node Metastasis. *Technology in
[21] Zhang, J., Yang, J., Zhang, X., Xu, J., Sun, Y., & Zhang, P. (2018). MicroRNA-10b expression in breast cancer and its clinical association. *Plos One, 13*(2) pp.1-11.

[22] Guo, C., Fu, M., Dilimina, Y., Liu, S., & Guo, L. (2018). MicroRNA-10b expression and its correlation with molecular subtypes of early invasive ductal carcinoma. *Experimental and Therapeutic Medicine. 15*(3), pp.2851-2859.

[23] Cleophas, T. J., & Zwinderman, A. H. (2017). Ensembled Correlation Coefficients. *Modern Meta-Analysis*, pp.195-204.