Expression profile of stem cell pathway genes in patients with advanced breast cancer after neoadjuvant therapy

S Dewi1,2, R Karsono2,3, S Prama4 and S I Wanandi2*

1Biomedical Sciences Doctoral Program, Faculty of Medicine, Universitas Indonesia, Jakarta, 10430, Indonesia
2Department of Biochemistry and Molecular Biology, Faculty of Medicine, Universitas Indonesia, Jakarta, 10430, Indonesia
3Department of Surgical Oncology, Dharmais Cancer Hospital, Jakarta, 11420, Indonesia
4Institute of Statistics, Jakarta, Indonesia
5E-mail: septelia.inawati@ui.ac.id

Abstract. Presence of breast cancer stem cells (BCSCs) is one of reasons for therapy resistance and metastatic development. Currently, neoadjuvant therapy is widely applied for local advanced breast cancer prior to tumor resection. This study aimed to analyze existence of BCSCs after neoadjuvant therapy and determine their correlation with clinical characteristics and patient survival. Cancer tissues were collected from 46 patients with stage IIIb or IV breast cancer before and after chemotherapy or hormone neoadjuvant therapy. The expression profiles of stem cell pathway genes were then investigated using next generation sequencing (Miseq, Illumina) and TruSeq Targeted RNA Expression stem cell panel kit (Illumina) comprising 100 genes. Altered stem cell gene expression was analyzed using paired t-test and correlated with clinical characteristics. Twelve stem cell pathway genes were significantly altered after neoadjuvant therapy, comprising six upregulated genes (ALDH1A1, ALDH2, CCND2, CXCL12, FZD7, and IGF1) and six downregulated genes (CCNE1, CDC42, CTNNB1, HDAC2, PSEN1, and PSENEN). We observed a significant increase in ALDH1A1 in the 3-year non-survival group and a poor survival rate of patients in high CDC42 and HDAC2 expression group. It can be concluded that overexpression of ALDH1A1, CDC42, and HDAC2 after neoadjuvant chemotherapy is correlated with poor prognosis in advanced breast cancer.

1. Introduction

Breast cancer treatment has developed rapidly in recent years, and there are many agents to target breast cancer cells. Nevertheless, around 40% of breast cancer survivors are still at risk for relapse, with 60% to 70% of these recurrences being distant metastases [1]. Statistical analyses suggest that approximately 30% to 50% of patients with early stage breast cancer are prone to progression to metastasis despite being administered treatments such as chemotherapy and/or adjuvant therapies [2]. One reason for these recurrences is believed to be the presence of a small population of stem-like cells, known as cancer stem cells (CSCs).

In recent years, many studies have revealed that tumors, including breast cancer, are composed of heterogeneous cell populations with different biological characteristics [3,4]. Tumorigenesis is initiated by CSCs, a small subpopulation of cancer cells that comprise around 1% to 5% of all tumor cells. CSCs have the capability for self-renewal and differentiation, as well as a high capacity for tumor generation in vivo [5]. They can divide asymetrically, producing one stem cell (self-renewal) and one progenitor cell that is able to generate heterogeneous lineages of cancer cells that comprise tumors.
CSCs were initially identified in human breast cancer in 2003 by Al-Hajj et al. They identified a cellular population characterized by cell-surface markers, such as CD44+/CD24−/low and ESA* (epithelial surface antigen) [3]. Breast cancer cells that are enriched with the CD44+/CD24−/low phenotype can initiate tumor formation when injected into NOD/SCID mice [6]. Furthermore, ALDH was identified as a breast CSC marker. Breast cancer cells with positive ALDH activity were able to generate tumors in NOD/SCID mice with phenotypic characteristics similar to the parental tumor. Breast cancer cells with a CD44+/CD24−/ALDH+ phenotype were shown to be more tumorigenic than CD44+/CD24− or ALDH− cells, as it was possible to generate tumors in mice using only 20 cells [7].

Neoadjuvant (preoperative) therapy of breast cancer is an established safe and effective therapeutic approach for larger primary and locally advanced breast cancer. Neoadjuvant therapy (NAT) offers advantages such as reducing the tumor size and making the patients candidates for surgical resection or breast-conserving surgery rather than mastectomy [8]. The present study investigate the existence of breast cancer cells by analyzing the expression profiles of stemness genes in patients with advanced breast cancer after NAT.

2. Materials and Methods

2.1. Patient population and characteristics

This study was approved by the Health Research Ethics Committee, Faculty of Medicine Universitas Indonesia-Cipto Mangunkusumo Hospital. The study was conducted in the Department of Biochemistry and Molecular Biology at the Universitas Indonesia and the Department of Oncology Surgery at the National Dharmais Cancer Hospital. A total of 46 patients with stage IIIB or VI breast cancer were recruited and treated with NAT for 6 months before mammosurgery. NAT comprised hormone therapy (tamoxifen or aromatase inhibitor) or chemotherapy (5-fluorouracil, anthracyclines, cyclophosphamides; FAC) based on estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) status. Breast cancer specimens were collected at the National Cancer Hospital Dharmais and were obtained from tumor biopsy (pretherapy) and during mammosurgery (posttherapy). We conducted a cohort study to monitor the survival time of patients after NAT.

2.2. Isolation of total RNA

Total RNA was isolated using Tripure isolation reagent (Roche). RNA quantity was measured at an absorbance of 260 nm and verified using a Qubit RNA HS assay kit (Invitrogen), and the purity of RNA was measured using the ratio of the absorbance at 260/280 nm. RNA integrity was verified by observation of the 28S/18S rRNA band on an agarose gel (1%) electrophoresis.

2.3. Preparation of targeted RNA sequencing

2.3.1. cDNA synthesis

Isolated RNA was used to synthesize cDNA using ProtoScript II Reverse Transcriptase (NEB) and random primers. The total RNA input was 50 ng for intact RNA and 200 ng for degraded RNA.

2.4. DNA library construction using total RNA

TruSeq Targeted RNA Expression stem cell panel kits (Illumina) were used to prepare a targeted library. An oligonucleotide pool containing upstream and downstream oligonucleotides specific to targeted regions of interest to cDNA samples was hybridized and bound to paramagnetic streptavidin beads. The hybridized cDNA was washed to remove unbound oligonucleotides prior to the extension–ligation step. DNA polymerase extended from the upstream oligonucleotide through the targeted region, followed by ligation to the 5′-end of the downstream oligonucleotide using a DNA ligase. The resulting products containing the targeted regions of interest were amplified using primers to add index sequences for sample multiplexing and common adapters required for cluster generation. PCR clean up used AMPure XP beads to purify the PCR products and remove other reaction components.

2.5. Quantitation of the library

The library was quantitated using a Library Quantification Kit (Kapa Biosystem) using quantitative real-time PCR (qRT-PCR). The library was diluted according to the Illumina protocol in hybridization buffer for sequencing.
2.6. **Sequence libraries**
TruSeq Targeted RNA Expression libraries were sequenced using an MiSeq sequencing system (Illumina).

2.7. **Data analysis**
The MiSeq Reporter (Illumina Supported Software) was used to process the base cells generated by the MiSeq sequencing system. Data were reported as number of reads for each sample and gene.

3. **Results**

3.1. **Patient characteristics**
Clinical data are shown in Table 1. A total of 46 patients were included in the study. The median age of the patients was 48 (range: 22–70) years. The predominant histological subtype was invasive ductal carcinoma ($n = 26$), and most patients had low-grade ($n = 29$) and high-grade ($n = 17$) tumors. ER positivity was observed in 67% ($n = 31$) and PR positivity in 72% ($n = 33$) of the tumors. HER2 was overexpressed in 30% ($n = 14$) of the cases. Less than half (48%) of patients received hormone therapy, almost the same as the chemotherapy group (52%).

| Characteristic                      | Number of patients (%) |
|-------------------------------------|------------------------|
| Total number of patients            | 46                     |
| Age, years                          | 48 (22–70)             |
| Histologic findings                 |                        |
| Invasive ductal                     | 26                     |
| Lobular                             | 4                      |
| No special type                     | 15                     |
| Unknown                             | 1                      |
| Tumor grading                       |                        |
| Low-grade                           | 29                     |
| High-grade                          | 17                     |
| Estrogen receptor                   |                        |
| Positive                            | 31                     |
| Negative                            | 15                     |
| Progesterone receptor               |                        |
| Positive                            | 33                     |
| Negative                            | 13                     |
| HER2 status                         |                        |
| Positive                            | 14                     |
| Negative                            | 32                     |
| Therapy                             |                        |
| Hormone therapy                     | 22                     |
| Chemotherapy                        | 24                     |
| Survival                            |                        |
| Alive                               | 21                     |
| Dead                                | 25                     |

3.2. **Expression profiles of stem cell genes after NAT**
We analyzed the expression profiles of stem cell genes by targeted RNA sequencing using a stem cell panel kit (Illumina). The panel contained 100 genes related to stem cell regulation. The expression patterns of 46 paired samples were analyzed before and after NAT using paired $t$-test. Twelve genes were significantly differentially expressed after NAT, comprising six downregulated and six upregulated genes (Table 2).
Table 2. Upregulated and downregulated gene expression in the stem cell panel

| Downregulated | Gene  | Function                        | Fold change average (log₂) | p value |
|---------------|-------|---------------------------------|-----------------------------|---------|
| 1             | CCNE1 | Cell cycle regulation           | −2.0680                     | 0.009   |
| 2             | CDC42 | MAPK signaling pathway          | −0.3705                     | 0.009   |
| 3             | CTNNB1| Wnt signaling pathway           | −0.2037                     | 0.027   |
| 4             | HDAC2 | Cell cycle regulation           | −0.5871                     | 0.001   |
| 5             | PSEN1 | Notch signaling pathway         | −0.3906                     | 0.023   |
| 6             | PSENEN| Notch signaling pathway         | −0.3297                     | 0.016   |

Upregulated

1. ALDH1A1 | Proliferation | 1.1540 | 0.049 |
2. ALDH2  | Proliferation | 1.1296 | 0.042 |
3. CCND2  | Cell cycle regulation | 0.4123 | 0.041 |
4. CXCL12 | PI3K–Akt signaling pathway | 0.7032 | 0.042 |
5. FZD7   | Wnt signaling pathway | 1.0051 | 0.030 |
6. IGF1   | MAPK signaling pathway | 1.6974 | 0.007 |

The fold change represents altered gene expression after NAT expressed as log₂. The downregulated genes included CDC42, CTNNB1, CCNE1, HDAC2, PSEN1, and PSENEN with fold changes of −0.3705, −0.2037, −2.0680, −0.5871, −0.3906, and −0.3297, respectively. Among these downregulated genes, CCNE1 was decreased the most. The upregulated genes included ALDH1A1, ALDH2, CCND2, CXCL12, FZD7, and IGF1, with fold changes of 1.1540, 1.1296, 0.4123, 0.7032, 1.0051, and 1.6974, respectively. Among these upregulated genes, IGF1 was increased the most.

Figure 1 shows a heat map illustrating the changes in expression of the significant genes from each of the 46 samples. Unsupervised hierarchical clustering was used to order all samples according to their similarities in gene expression.

Figure 1. Expression of significant genes (log₂) in the stem cell panel after neoadjuvant therapy. Gene expression from 46 samples was evaluated using the unsupervised hierarchical clustering method and the results were visualized using R statistic software. Green represents downregulation of gene expression and red represent upregulation of gene expression.
3.3. Significant gene expression based on therapy, survival status, histopathologic grade, and molecular intrinsic markers

This study also analyzed the expression of significant genes based on therapy, chemotherapy, and hormone therapy. Figure 2a shows that the pattern of significant gene expression was similar between the chemotherapy and hormone therapy groups, demonstrating that gene regulation in the stem cell pathway did not differ between both therapies.

The expression patterns of significant genes according to the 3-year survival and non-survival group are shown in Figure 2b. When these groups were compared using Student t-test, ALDH1A1 expression was significantly different between the two groups (p = 0.044). ALDH1A1 expression changes were lower in the 3-year survival group (fold change: −0.0899) than in the 3-year non-survival group (fold change: 2.1990).

This study also compared the expression of significant genes based on histological grade (Fig. 2c). ALDH2 expression was significantly different between the low and high groups (t-test, p = 0.002). ALDH2 expression changes were higher in the high group (fold change: 2.355) than in the low group (fold change: −0.5170).
Comparison of the expression of significant genes based on ER, PR, and HER2 status (Fig. 2) demonstrated a significantly different change in *CCND2* expression between the ER-positive and -negative groups (*t*-test, *p* = 0.011), as well as the PR-positive and -negative groups (*t*-test, *p* = 0.031). *CCND2* expression was lower in the ER-negative group (fold change: −0.2877) than the ER-positive group (fold change: 0.7511). Similarly, *CCND2* expression was lower (fold change: −0.2553) in the PR-negative group than in the PR-positive group (fold change: 0.6753). Besides *CCND2*, the expressions of *CDC42*, *FZD7*, and *PSENEN* were significantly different between the PR status groups (*t*-test, *p* = 0.044, 0.012, and 0.025, respectively). Changes in the expressions of *CDC42* and *PSENEN* were higher in the PR-negative group (fold change: 0.0606 and 0.1366) than the PR-positive group (fold change: −0.5404 and −0.5134). However, changes in *FZD7* expression in the PR-negative group were lower (fold change: −0.7641) than the PR-positive group (fold change: 2.8807). Changes in expression based on HER-2 status were not significantly different between the HER2-positive and -negative groups.

### 3.4. Survival Analysis

This study analyzed the survival rate of patients with advanced breast cancer after NAT. The survival rate was compared according to the expression levels of all significant genes, and the expression levels were classified into two groups: high (fold change > 0) and low (fold change < 0). The differences in survival rate was analyzed based on expression levels using log rank test analysis. It is shown that the survival rate of patients based on *ALDH1A1, ALDH2, CCND2, CCNE1, CDC42, CTNNB1, CXCL12, FZD7, HDAC2, IGF1, PSEN1*, and *PSENEN* gene expressions was not significantly different between the high and low expression groups. When the samples were classified according to the type of therapy (Fig. 3), this study found a significant survival rate between the low and high expression groups for *CDC42* and *HDAC2* levels (log rank test, *p* = 0.025 and 0.039, respectively), especially in the chemotherapy group. Classification based on HER2 status (Fig. 3) showed a significant survival rate between the low and high expression groups for *CDC42* and *HDAC2* levels (log rank test, *p* = 0.022 and 0.010, respectively), especially in the HER2-negative group. In the chemotherapy and HER2-negative groups, the survival rate for low *CDC42* and *HDAC2* gene expression was longer than that for high gene expression.
4. Discussion

In the present study, we analyzed the expression profiles of stem cell genes in patients with advanced breast cancer after NAT. NAT was given either as chemotherapy (FAC) or hormone therapy (tamoxifen or aromatase inhibitor). The aim of the study was to identify prognostic factors for chemotherapy or hormone therapy as NAT in advanced breast cancer. Alterations in gene expression in breast cancer induced by NAT were identified by comparing gene expressions before and after treatment.

The expression levels of 12 genes were significantly different after NAT. Their functions included cell cycle regulation, Wnt signaling pathway, Notch signaling pathway, and MAPK signaling pathway, all of which can stimulate increased cell proliferation. Among these significant genes, six were upregulated and six were downregulated. However, the gene functions did not differ between upregulated and downregulated genes.

There were no significant differences in gene expression between the chemotherapy and hormone therapy groups. However, when we compared gene expression based on 3-year survival and non-survival, we found that ALDH1A1 expression was significantly decreased in the 3-year survival group compared with the non-survival group. ALDH1A1 is a breast cancer stem cell marker [7]. It is a
detoxifying enzyme responsible for the oxidation of intracellular aldehydes [9] and may play a role in the early differentiation of stem cells via its role in oxidizing retinol to retinoic acid [10]. Some studies have demonstrated that cells with ALDH activity isolated from normal human breast have phenotypic and functional characteristics of mammary stem cells [11,12], and they are associated with poor clinical outcome [7]. We also indicated that ALDH2 expression was significantly higher in high-grade breast cancer than low-grade breast cancer.

Moreover, the present study also analyzed the alteration of gene expressions based on ER, PR, and HER2 status. We demonstrated an increased expression of CCND2 in the ER- and PR-positive groups and increased expression of FZD7 in the PR-positive group. This could be due to hormone stimulation that activates FZD7 involved in the Wnt signaling pathway, leading to cell cycle activation by CCND2. Other genes also affected by PR status included CDC42 and PSENEN. Expressions of CDC42 and PSENEN were decreased in the PR-positive group compared with the PR-negative group. PR-negative breast cancer usually presents as triple negative breast cancer. In this subtype, the MAPK and Notch signaling pathways are usually activated. Furthermore, CDC42 and PSENEN are involved in the MAPK and Notch signaling pathways.

In the survival analysis, there were no significant differences between high and low expression groups of all genes. However, when the chemotherapy and HER2-negative groups were separated, we demonstrated a significant difference in the survival rate between the high and low CDC42 and HDAC2 expression group. Expression of these genes could predict prognosis in the chemotherapy and HER2-negative groups only.

Cell division cycle 42 (CDC42) is a member of the RAS homolog family of small guanine nucleotide triphosphatases. It regulates crucial cellular processes, including cell cycle and cell cytoskeleton organization [13]. Studies have reported that CDC42 is overexpressed and hyperactivated in human breast invasive ductal carcinomas [14,15]. Other studies in other cell and tissue types have demonstrated that CDC42 regulates cell cycle progression, polarity, migration, and differentiation, which are essential for mammary gland development and become disturbed during tumor formation [16,17]. Pichot et al. reported that CDC42 interacting protein 4 plays a role in promoting migration and invasion of breast cancer cell lines [18].

Histone deacetylation is an epigenetic modification and is regulated by a family of histone deacetylases (HDACs). HDACs remove acetyl groups from lysine residues of histones, increasing the ionic interactions between histones and DNA, resulting in a closed chromatin conformation, repression of transcription and silencing of gene expression [19], and cell cycle regulation, differentiation, and tumorigenesis. [20,21] Recent studies have reported that HDACs are expressed in breast cancer [22,23], and overexpression of HDAC2 is correlated with metastasis and poor prognosis [24].

5. Conclusion
In conclusion, the present study demonstrated altered expression of stem cell pathway genes after NAT in patients with advanced breast cancer. Overexpression of ALDH1A1 was significant in the 3-year non-survival group after NAT. Survival rate of patients after neoadjuvant chemotherapy was longer in the CDC42 and HDAC2 low expression group than the high expression group.

References
[1] Smalley M, Piggott L and Clarkson R 2013 Breast cancer stem cells: obstacles to therapy Cancer Lett. 338 57–62.
[2] Gupta P B, Onder T T, Jiang G, et al 2009 Identification of Selective Inhibitors of Cancer Stem Cells by High-Throughput Screening Cell. 138 645–59.
[3] Al-Hajj M, Wicha M S, Benito-Hernandez, Adalberto Morrison S J and Clarke M F 2003 Prospective identification of tumorigenic breast cancer cells PNAS. 100 3983–8.
[4] Harrison H, Farnie G, Howell S J, et al 2010 Regulation of Breast Cancer Stem Cell Activity by Signaling through the Notch4 Receptor Cancer Res. 70 709.
[5] Jordan C T, Guzman M L and Noble M 2006 Cancer Stem Cells N. Engl. J. Med. 355 1253–61.
[6] Ponti D, Costa A, Zaffaroni N, et al 2005 solation and &lt;em&gt;in vitro&lt;/em&gt; Propagation of Tumorigenic Breast Cancer Cells with Stem/Progenitor Cell Properties Cancer Res. 65 5506.

[7] Ginestier C, Hur M H, Charafe-Jauffret E, et al 2007 ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome Cell Stem Cell. 1 555–67.

[8] Thompson A M and Moulder-Thompson S L 2012 Neoadjuvant treatment of breast cancer Ann. Oncol. 23 x231–x6.

[9] Sophos N A, Pappa A, Ziegler T L and Vasiliou V 2001 Aldehyde dehydrogenase gene superfamily: the 2000 update Chem. Biol. Interact. 130–132 323–37.

[10] Chute J P, Muramoto G G, Whitesides J, et al 2006 Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells Proc. Natl. Acad. Sci. 103 11707–12.

[11] Marcato P, Dean C A, Giacomantonio C A and Lee P W K 2011 Aldehyde dehydrogenase: Its role as a cancer stem cell marker comes down to the specific isoform Cell Cycle. 10 1378–84.

[12] Liu S, Cong Y, Wang D, et al 2014 Breast Cancer Stem Cells Transition between Epithelial and Mesenchymal States Reflective of their Normal Counterparts Stem Cell Reports 2 78–91.

[13] Arias-Romero L E and Chernoff J 2013 Targeting Cdc42 in cancer Expert Opin. Ther. Targets. 17 1263–73.

[14] Ellenbroek S I J and Collard J G 2007 Rho GTPases: functions and association with cancer 24 657–72.

[15] Fritz G, Brachetti C, Bahlmann F, Schmidt M and Kaina B 2002 Rho GTPases in human breast tumours: expression and mutation analyses and correlation with clinical parameters Br. J. Cancer. 87 435–44.

[16] Stengel K and Zheng Y 2011 Cdc42 in oncogenic transformation, invasion, and tumorigenesis Cell Signal. 23 1415–23.

[17] Pedersen E and Brakebusch C 2012 Rho GTPase function in development: How in vivo models change our view Exp. Cell. Res. 318 1779–87.

[18] Pichot CS, Arvanitis C, Hartig S M, et al 2010 Cdc42 Interacting Protein 4 promotes breast cancer cell invasion and formation of invadopodia through activation of N-WASp Cancer Res. 70 8347–56.

[19] Khan O and La Thangue N B 2012 HDAC inhibitors in cancer biology: emerging mechanisms and clinical applications Immunol. Cell Biol. 90 85–94.

[20] Thiagalingam Sam, Cheng K H, Lee H J, Mineva N, Thiagalingam A and Ponte JF 2003 Histone Deacetylases: Unique Players in Shaping the Epigenetic Histone Code Ann. N Y Acad. Sci. 983 84–100.

[21] Lagger G, Doetzhofer A and Schuettengruber B, et al 2003 The Tumor Suppressor p53 and Histone Deacetylase 1 Are Antagonistic Regulators of the Cyclin-Dependent Kinase Inhibitor p21/MDM2/CIP1 Gene Mol. Cell Biol. 23 2669–79.

[22] Zhang Z, Yamashita H and Toyama T, et al 2005 Quantitation of HDAC1 mRNA Expression in Invasive Carcinoma of the Breast Breast Cancer Res. Treat. 94 11–6. doi:10.1007/s10549-005-6001-1.

[23] Ververis K and Karagiannis T C 2012 An atlas of histone deacetylase expression in breast cancer: fluorescence methodology for comparative semi-quantitative analysis Am. J. Transl. Res. 4 24–43.

[24] Zhao H, Yu Z and Zhao L, et al 2016 HDAC2 overexpression is a poor prognostic factor of breast cancer patients with increased multidrug resistance-associated protein expression who received anthracyclines therapy Jpn. J. Clin. Oncol. 46 893–902.