INTRODUCTION

Breast cancer (BC) is the most common cancer among women worldwide, with an estimated 2.4 million new cancer cases diagnosed in 2015 ***(Fitzmaurice et al., 2017)***.

In United States, an estimated 252,710 new cases of invasive breast cancer and 63,410 new cases of *in situ* breast carcinoma are diagnosed among women in 2017 (***American Cancer Society, 2017)***.

In Africa, BC is the most commonly diagnosed cancer among women annually, with an estimated 63,100 cancer deaths in 2012. The incidence rates of breast cancer vary considerably between African countries, with a rate of 27 per 100,000 women diagnosed with BC in Middle Africa ***(Kantelhardt et al., 2015)***.

In Egypt, BC accounts for 32 % of total malignancies among females **(*Ibrahim et al., 2014)***. It also occupies the second rank (18.3%) of total cancer cases ***(Fattah et al., 2018)***.

Breast cancer is curable when detected at its early stage. Oncologists encounter an exceptionally challenging task when it comes to taking clinical decisions on breast cancer treatment ***(Hayes et al., 2001)***. Molecular techniques, especially gene

and non-coding RNA (ncRNA) expression profiling, have been used increasingly to improve breast cancer classification and to evaluate patient prognosis and response to therapy ***(Eo et al., 2012)***.

Recent acquisitions on human carcinogenesis suggest that small populations of tumor stem cells can influence and modify neoplastic cell behavior and aggressiveness as well as therapeutic response ***(Geng et al., 2014)***. Stem cells undergo symmetric or asymmetric cell division producing identical stem cell progeny or cells that are more differentiated. Moreover, the cycling of stem cells varies considerably, facilitating so-called kinetic resistance, whereby slow cycling or quiescent stem cells are unaffected by DNA damaging agents or radiation comparing to the more rapidly dividing cells ***(Dalerba et al., 2007)***. Cancer stem cells (CSCs) also show great drug resistance through other mechanisms, including drug effluxing ***(Phi et al., 2018)***.

Resistance to radiation can occur through the repair of damaged DNA, redistribution of cycling cells, re-oxygenation of hypoxic tumor regions and repopulation of the tissue, commonly considered the “four Rs”. The role of ncRNAs in regulating these mechanisms of resistance, in addition to

intrinsic resistance that CSCs may have, is an exciting area of research ***(Valent et al., 2012)***.

Many observations suggest that breast cancer ability to proliferate, progress and spread is based on a limited subpopulation of cells with properties similar to stem cells, defined as breast cancer stem cells (BCSCs) ***(Kagara et al., 2012)***. Several stem cell markers have been described for identification of BCSCs in cancer subtypes, such as CD44, CD24, CD133, EpCAM, CD166, Lgr5, CD47, ALDH1 and ABCG2 ***(Ahmed et al., 2012)***. The mechanisms that regulate CSCs growth, in conjunction with factors that facilitate resistance to radiation and chemotherapy, are of particular clinical importance given the role of CSCs in tumorigenesis and recurrence ***(Chiba et al., 2013)***.

It has become increasingly apparent that non coding RNAs are of crucial importance for human cancer. The functional relevance of ncRNAs is particularly evident for microRNAs (miRNAs) and long non-coding RNAs (lncRNAs). MicroRNAs are endogenously expressed as small RNA sequences that act as post-transcriptional regulators of gene expression ***(Gibb et al., 2011)***.

LncRNAs regulate important cellular signaling pathways in cancer cells at transcriptional, post-transcriptional and epigenetic levels and are often aberrantly expressed in a variety of human cancers ***(Spizzo et al., 2012)***. In addition to effects on tumor cell growth, survival and cell signaling, lncRNA can modulate CSC behavior, including the expression of pluripotency factors ***(Li et al., 2015)***.

The identification of lncRNA that are mechanistically linked to cancer stem cell self-renewal and differentiation or aberrant signaling pathways associated with tumor growth or progression offer new opportunities for diagnostic and therapeutic intervention ***(Li et al., 2015)***.

Among the oncogenic lncRNAs is lncRNA-HOTAIR (HOX transcript antisense RNA), which is located in the mammalian homeobox C (HOXC) locus on chromosome 12q13.13. HOTAIR was highly up-regulated in primary and metastatic breast tumors ***(Cheetham et al., 2013)***. Moreover, *Alves et al., 2013* investigated role of HOTAIR in epithelial-to-mesenchymal transition (EMT) and its role in arising and maintenance of CSCs. They revealed that HOTAIR plays an important role in the process of tumorigenicity by triggering EMT and acquiring stemness ***(Hajjari* & Salavaty, *2015)***.

AIM OF THE WORK

The core of the study is to:

1. Retrieve non coding RNAs that are mechanistically linked to breast cancer stem cell self-renewal and differentiation.
2. Validate the results in a group of breast cancer patients versus control groups to evaluate their usefulness as a potential biomarker in breast cancer diagnosis.

Chapter 1

BREAST CANCER

Cancer is a group of related diseases that cause cells in the body to divide enormously and spread into surrounding tissues. Many cancers form solid tumors, which are masses of tissue **(*National Cancer Institute, 2015)***.

Breast cancers can start from different parts of the breast. Most breast cancers begin in the ducts that carry milk to the nipple (ductal cancers), some start in the glands that make breast milk (lobular cancers) and small number of cancers start in other tissues in the breast ***(American Cancer Society, 2017)***.

I-Breast cancer epidemiology:

Breast cancer is the most common cancer in women in both developed and developing countries and the second most common cancer in the world ***(Rivera-Franco & Leon-Rodriguez, 2018)***, with an estimation of 2.4 million incident cases in 2015. It is the cause of death for 533,000 cases making it the leading cause of cancer deaths in women ***(Fitzmaurice et al., 2017)***.

The incidence and mortality rates of BC vary greatly worldwide. High-income countries represent most of the countries with top incidence rates, whereas low- and middle-income countries represent most of those with the highest mortality rates ***(Rivera-Franco & Leon-Rodriguez, 2018)***.

In Egypt, based on the National Population-Based Cancer Registry Program, breast cancer is the most frequent cancer among females accounting for 32 % of their total malignancies. The frequencies of breast cancer among females were 33.8%, 26.8% and 38.7% in lower, middle and upper Egypt, respectively ***(Ibrahim et al., 2014)***. It also occupies the second rank (18.3%)of total cancer cases ***(Fattah et al., 2018)***.

II-Risk factors for breast cancer:

1-Gender and Age:

Many risk factors for developing breast cancer have been established, but increasing age is the single most important risk factor after female gender ***(Balekouzou et al., 2017-b)***. Breast cancer is the most frequently diagnosed cancer among women aged 55–64 years with the median age of death at 68 years ***(Shah & Guraya, 2017)***. While male breast cancer is rare and accounts for around 1% of all diagnosed breast cancer cases ***(Humphries et al., 2017)***.

2-Race and ethnicity:

African American and Hispanic women have a lower incidence of breast cancer compared to white women but they are more likely to be diagnosed with breast cancer at a younger age and they have a higher mortality rate. The median age at diagnosis for African American women is 59 compared to 63 for white women ***(Yedjou et al., 2017)***.

East Asian women have a higher risk of triple-negative and her2/neu enriched tumors. Luminal tumors are more commonly diagnosed among postmenopausal Caucasian women in Western countries, but conversely are more prevalent among premenopausal women in East Asia ***(Li et al., 2017)***.

3-Obesity:

Obesity is associated with both; increased risk of developing postmenopausal breast cancer and with a worse disease outcome for women of all ages ***(James et al., 2015)***. The Million Women Study followed 1.2 million UK women ages 50 to 64 years for a mean of 5.4 years, including 45,037 cases with breast cancer, and identified a nearly 30% higher risk of developing postmenopausal breast cancer with obesity ***(Picon-Ruiz et al., 2017)***.

Overweight women have been found to have higher estrogen levels than normal weight women, providing a possible explanation for positive associations observed between body mass index (BMI) and breast cancer risk in postmenopausal women ***(Guo et al., 2016)***. Obesity has been associated with increased risk of postmenopausal hormone receptor positive breast cancer ***(Goodwin, 2017)***.

4-Genetic risk factors:

Hereditary breast cancer, which is usually caused by a mutation in BRCA1 or BRCA2 genes, is responsible for 5% to 10% of all breast cancer cases. BRCA1 mutations occur more often in certain ethnic groups such as Jewish population ***(Mehrgou & Akouchekian, 2016)***.

Recent estimates suggest that 55 to 65% of BRCA1 mutation carriers and approximately 45% of BRCA2 mutation carriers will develop breast cancer by age 70 ***(Godet & Gilkes, 2017)***.Other inherited conditions associated with smaller breast cancer risk include Li-Fraumeni (*TP53*) and Cowden syndromes (*PTEN*) ***(O'Leary et al., 2017)***.

5-Family and Personal history of breast cancer:

First-degree family history was associated with an increased risk for developing breast cancer among women

(Braithwaite et al., 2018).The risk is further increased if they had at least one first-degree relative with breast cancer before the age of 50 years or two or more relatives with breast cancer, with at least one being a first-degree relative ***(Visscher et al., 2017)***. In addition, women with a personal history of breast cancer have a significantly elevated risk for future cancer recurrence ***(Lehman et al., 2016)***.

6-Hormonal factors:

There is substantial evidence for a role of female hormones in the etiology of breast cancer. Reproductive factors, such as early age at menarche, nulliparity and late age at menopause are all believed to be associated with breast cancer risk through hormonal mechanisms. Also the use of oral contraceptives or combined postmenopausal hormone therapy is associated with some increased risk of breast cancer ***(Ellingjord-Dale et al., 2017)***.

7-Dense breast tissue:

Women with high breast density (a mammographic indicator of the amount of breast tissue relative to fatty tissue in the breast) have a 4 to 6 fold increased risk of breast cancer ***(Baglietto et al., 2013)***.

8-Benign breast diseases and Lobular carcinoma in situ:

Benign breast diseases (BBD) are divided into: non proliferative disease (fibrocystic changes), proliferative disease without atypia (fibroadenoma) and proliferative disease with atypia (atypical ductal and lobular hyperplasia). All the 3 subtypes of BBD increase the risk of breast cancer, but the risk degree is different in each of them with the highest risk being in atypical proliferative lesions ***(Zendehdel et al., 2018)***.

Lobular carcinoma in situ (LCIS) is typically confined to lobules and terminal ducts of the breast and is usually found incidentally in biopsy specimens. Women with LCIS showed a 7 to 10 fold increase in the risk of developing subsequent breast cancers compared with the general population ***(Mao et al., 2017)***.

9-Alcohol consumption:

Excessive alcohol consumption can elevate the level of estrogen-related hormones in the blood and trigger the estrogen receptor pathways increasing the risk of breast cancer. The risk of breast cancer is also elevated in women who both smoke and drink ***(Sun et al., 2017)***.

10-Previous chest radiation

The breast is one of the most sensitive organs to radiation effects; therefore, radiation exposure is highly associated with breast cancer. Exposure of breast tissue to ionizing radiation before 40 years old is associated with a three-fold risk of breast cancer for an exhibition valued at 1 gray ***(Balekouzou et al., 2017-a)***.

III-Histological classification of breast cancer:

Breast cancer can be broadly categorized into: *in situ* and invasive (infiltrating) carcinomas. Carcinoma *in situ* is a growth of low-grade cancerous cells that are completely confined within breast lobules and/or ducts without invasion of the surrounding tissue. In contrast, invasive carcinoma does not confine itself to the initial tissue compartment from which it began ***(Malhotra et al., 2010)***.

Two forms of carcinoma *in situ* in the breast pathology have been identified, ductal carcinoma in situ and lobular carcinoma in situ. Ductal carcinoma in situ (DCIS) is the most frequently diagnosed *in situ* cancer of breast and considered as a true cancer precursor; whereas lobular carcinoma *in situ* (LCIS), which is also known as lobular neoplasia, is primarily

viewed as an indicator of increased breast cancer risk ***(Wang et al., 2017-e)***.

Invasive breast carcinomas include invasive ductal carcinoma (IDC, the most common type representing 70–80% of all cases) and invasive lobular carcinoma (ILC, representing 5–10% of breast carcinomas) ***(Liu, 2014)***.

Rarer histological types of breast cancer that comprise less than 10% of breast tumors include mucinous, tubular, medullary, papillary carcinomas and other epithelial tumors ***(Gudaviciene et al., 2015)***.

IV-The American Joint Committee on Cancer (AJCC) staging:

TNM system:

The TNM staging system is the worldwide basis for breast cancer staging. It is maintained by American Joint Committee on Cancer (AJCC). It is based on the TNM principle: tumor size, lymph node involvement and metastases ***(Giuliano et al., 2017)***.

The TNM staging system classifies cancers based on their T, N and M score:

- The letter T followed by a number from 0 to 4 describes the tumor's size and spread to the skin or to the chest wall under the breast. Higher T numbers mean a larger tumor and/or wider spread to tissues near the breast.
- The letter N followed by a number from 0 to 3 indicates whether the cancer has spread to lymph nodes near the breast and, if so, how many lymph nodes are affected.
- The letter M followed by a 0 or 1 indicates whether the cancer has metastasized to distant organs e.g. lungs or bones or not ***(American Joint Committee on Cancer, 2016)***.

Breast cancer stage grouping:

Once the T, N and M categories have been determined, this information is combined in a process called stage grouping. Cancers with similar stages tend to have a similar outlook and thus are often treated in a similar way. Stage is expressed in Roman numerals from stage I (the least advanced stage) to stage IV (the most advanced stage). Non-invasive cancer is listed as stage 0 as in table 1 ***(American Joint Committee on Cancer, 2016)***.

Table (1): TNM Stage Grouping of Breast Cancer ***(American Joint Committee on Cancer, 2016)***.

Occult Carcinoma	TX	N0	M0
Stage 0	Tis	N0	M0
Stage IA	T1	N0	M0
Stage IB	T0	N1mi	M0
Stage IB	T1	N1mi	M0
Stage IIA	T0	N1	M0
Stage IIA	T1	N1	M0
Stage IIA	T2	N0	M0
Stage IIB	T2	N1	M0
Stage IIB	T3	N0	M0
Stage IIIA	T1	N2	M0
Stage IIIA	T2	N2	M0
Stage IIIA	T3	N1	M0
Stage IIIA	T3	N2	M0
Stage IIIB	T4	N0	M0
Stage IIIB	T4	N1	M0
Stage IIIB	T4	N2	M0
Stage IIIC	Any T	N3	M0
Stage IV	Any T	Any N	M1