INTRODUCTION

Breast cancer is by far the most common cancer among women of both developed and developing countries, accounting for 22.9% of all female cancers ***(Ferlay J et al., 2012).*** It is also the leading cause of cancer death in females accounting for 13.7% of their cancer-related mortality. In Egypt, breast cancer is estimated to be the most common cancer among females accounting for 37.7% of their total cancer cases ***(The National Cancer Registry Program of Egypt (NCRPE), Reports and Statistics, 2012).*** It is also the leading cause of cancer-related mortality accounting for 29.1% of their cancer related deaths. The incidence to mortality ratio is poor (1.9:1) ***(Zeeneldinet al., 2013)***.

Earlier detection and treatment are thought to improve survival, yet breast cancer can be "an unpredictable disease" because even very small lesions at the limit of detection by mammography, magnetic resonance imaging, or palpation can progress to metastatic disease ***(Misek and kim, 2011).***

A large number of molecules have been investigated as potential prognostic and predictive factors of breast cancer. Well established prognostic factors in breast cancer include ki-67, estrogen receptor, progesterone receptor and HER-2.Other investigational prognostic factors include, mitosin, apoptosis related proteins, cell cycle molecules, plasminogen activators and inhibitors and angiogenesis related proteins ***(Cecchini et al., 2012).***

More and more, the discovery of relevant biomarkers is aided by in silico techniques based on applying data mining and computational chemistry on large molecular databases. However, there is an even larger source of valuable information available that can potentially be utilized through database searching ***(Trugenberger et al., 2013)*.**

Significant progress has been made in informatics to improve the capture and analysis of scientific data. In a complex disease, the expression of many genes can be significantly altered, leading to the appearance of a differentially expressed "disease module". Some of these genes directly correspond to the disease phenotype, (i.e. "driver" genes), while others represent closely-related first-degree neighbors in gene interaction space. The remaining genes consist of further removed "passenger" genes, which are often not directly related to the original cause of the disease. For prognostic and diagnostic purposes, it is crucial to be able to separate the group of "driver" genes and their first-degree neighbors ***(Hahn et al., 2012)*.**

The mammary gland is a dynamic organ that undergoes continuous cycles of proliferation, differentiation, and apoptosis between puberty, pregnancy lactation and menopause. A clear understanding of mammary stem/progenitor regulation and the process by which these cells become fully differentiated has significant implications in the field of breast cancer ***(Jiang et al., 2010)*.**

The initiation of DNA replication in eukaryotic cells is a carefully regulated process requiring the orchestrated assembly of many proteins at origin sites, including the origin recognition complex and minichromosome maintenance (*MCM*) complex. The *MCM* complex consists of six subunits, *MCM2* through *MCM7* (*MCM*2-7), which form a hexamer. Minichromosome Maintenance Complex component 5 (*MCM5*) is structurally very similar to the cyclin dependent kinase 46(CDC46) protein from S. cerevisiae, a protein involved in the initiation of DNA replication. The encoded protein is a member of the *MCM* family of chromatin-binding proteins and can interact with at least two other members of this family. The encoded protein is upregulated in the transition from the G0 to G1/S phase of the cell cycle and may actively participate in cell cycle regulation ***(Cvetic and Walter et al., 2006)*.**

Recent studies suggest that the Mitogenic effects of the up-regulation of minichromosome maintenance proteins play a vital role in development of cancer ***(Guida et al., 2005)*.**

The transcription factor Forkhead box M1 (*FOXM1*) is a member of the winged helix family of transcription factors is a key regulator of cell proliferation. The transcriptional activity of *FOXM1* is dependent on phosphorylation by cell cycle regulated kinases. Target genes include key components of the cell cycle, required for transition to S-phase and successful entry and completion of mitosis. *FOXM1* Plays also a role in DNA breaks repair participating in the DNA damage

checkpoint response. Aberrant expression of *FOXM1* leads to increased tumor growth, invasion, and metastasis. *FOXM1* is a critical regulator of mammary progenitor cells and luminal cell fate through direct transcriptional repression suggesting that *FOXM1*expression showed a stepwise increase with breast tumor grade ***(Park et al., 2011)***.

AIM OF THE WORK

To evaluate the expression of both *MCM5*, *FOXM* genes in relation to clinico-pathological factors of breast cancer and to explore their synergistic expression.

Chapter 1:
MINICHROMOSOME MAINTENANCE 5

Minichromosome maintenance (*MCM*) proteins play vital role in DNA replication, they are related to cell proliferation, and serve as useful markers for cancer screening, surveillance, and prognosis. They are encoded by genes which are parts of the *MCM* genes from *MCM* 2-7 (figure 1) ***(Maiorano et al., 2006)*.**

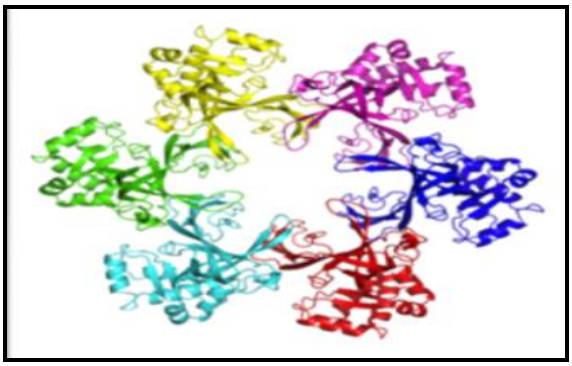

Figure (1): Minichromosome maintenance (*MCM*) proteins hexemer structure ***(Froelich et al., 2014)*.**

Other names of *MCM5* gene is CDC46, MGC5315 and of P1-CDC46. *MCM* proteins were first identified in yeast when mutations in their genes were defective for minichromosome maintenance ***(Maiorano et al., 2006)*.** In eukaryotic cells; six related *MCM* proteins (*MCM2*-7) form a ring-shaped heterohexamer, the *MCM2*-7 complex. Hexameric *MCM* rings act as the replicative DNA helicase encircling the leading strand DNA template at the replication fork (figure 2) ***(Barry et al., 2007)***.

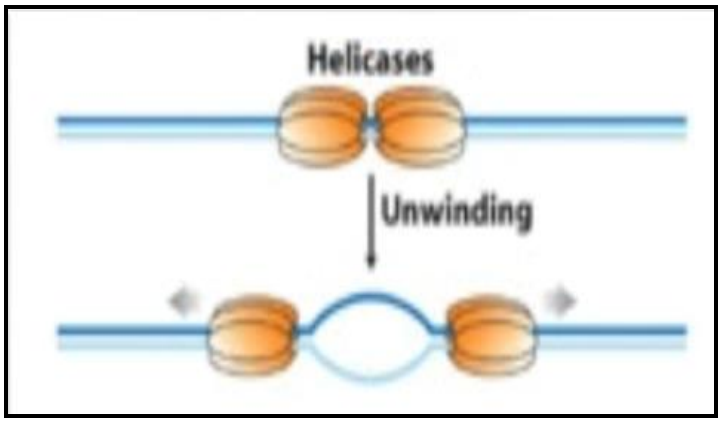

Figure (2): Helicase activity of *MCM* family ***(Barry et al., 2007)***.

MCM5, *MCM*3 and *MCM2* are subunits of the complex that negatively regulate the active *MCM* core subunits (*MCM*4, *MCM*6 and *MCM*7). Actually it is thought that *MCM2* and *MCM5* form a gate in the *MCM* toroid. When the conformation is in a closed status, the dimer *MCM2-MCM5* binds ATP; on the other hand, when the gate is open, the active site of the dimer is empty since no nucleotide is bound, and therefore no helicase activity is observed (figure 3). Further studies suggest that *MCM2*/5 dimer are capable of regulating the helicase activity of the *MCM* complex and/or is essential for the initial loading of the complex onto the origins of replication ***(Bochman et al., 2008)*.**

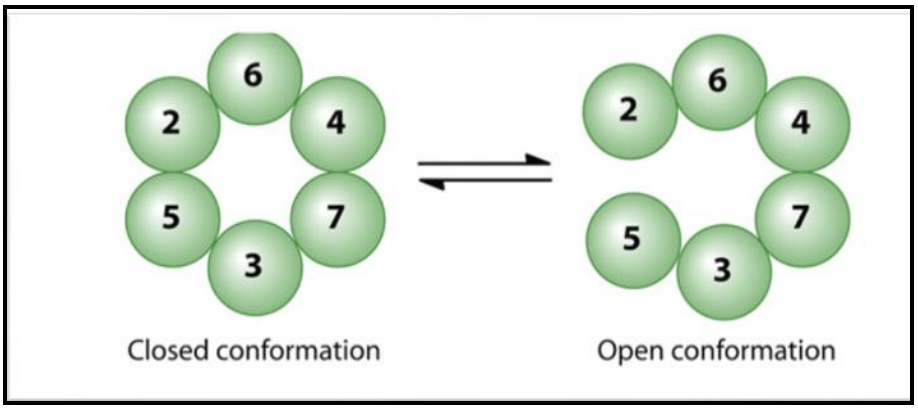

Figure (3): Closed form of *MCM* family which represents active form and open form which represent inactive form of *MCM* ***(Bochman et al., 2008).***

MCM5 Structure:

MCM 5 gene has Spanning 24,4 kb of genomic DNA, it consists of 17 exons and 16 intervening introns. The *MCM5* protein is localized to the nucleus. The unique transcript of the *MCM5* gene is 2546 nt(figure 4)**.**The *MCM5* protein is composed of 734 amino acid residues, with a calculated molecular mass of 82, 3 kDa and a basal isoelectric point of 8, 64 ***(Paul et al., 1996)***.

Figure (4): Schematic representation of the *MCM5* gene. Exons are shown as boxes and introns as connecting lines. The coding sequence is highlighted as red, while 5' and 3' untranslated regions (UTRs) are shown in white. The numbers inside boxes indicate exon lengths and the vertical connecting lines show the intron lengths. The arrows show the position of the start codon (ATG) and stop codon (TGA), and the asterisk shows the position of the polyadenylation signal (AATAAA). Roman numerals

indicate intron phases. The intron phase refers to the location of the intron within the codon; I denote that the intron occurs after the first nucleotide of the codon, II denotes that the intron occurs after the second nucleotide, and 0 means that the intron occurs between distinct codons ***(Paul et al., 1996).***

MCM5 subunit contains three domains. The N-terminal domain possesses an OB (oligonucleotide/oligosaccharide binding)-fold and a zinc-binding motif, this domain mediates the head-to-head interaction of the two hexamers ***(Evrin et al., 2009).*** The second domain contains a conserved ATPase AAA+ fold, which binds and hydrolyzes ATP at subunit interfaces around the hexameric ring and is, required for DNA unwinding ***(Bochman et al., 2008).*** The structural characteristic of *MCM5* is an *MCM* box consisting of approximately 200 amino acids of the catalytic domain. This includes a Walker A motif containing the P-loop (phosphate-binding loop) of the active site and the invariant lysine residue found in all ATP-binding proteins, a hydrophobic Walker B element that is responsible for ATP hydrolysis, and an Arginine finger. The Walker B motif is part of the sequence IDEFDKM, which is conserved in all *MCM* proteins and defines the *MCM* family. A short domain at the C-terminus includes a winged helix fold (figure 5). *MCM* hexamers demonstrate a two-tiered ring architecture in electron microscopy studies with an N-terminal domain tier and an ATPase domain tier ***(Aravind and Koonin, 1997; Bochman et al, 2008; Remus et al., 2009; Costa et al., 2011).***

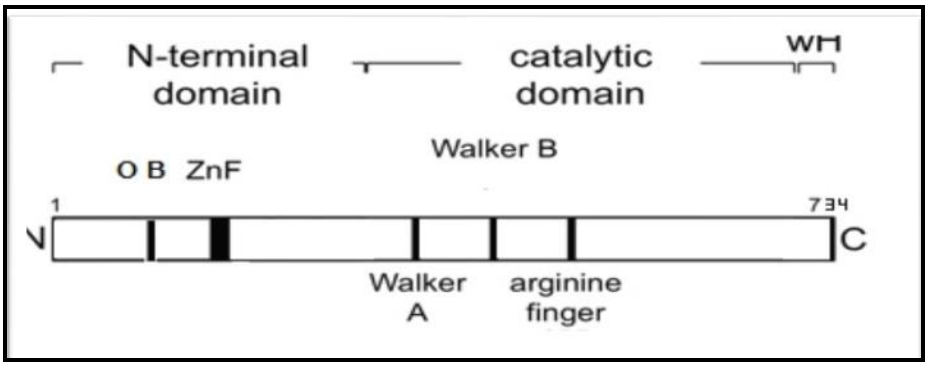

Figure (5): *MCM5* Domains consist of three domains. The N-terminal domain has an OB (oligonucleotide/oligosaccharide binding)-fold and a zinc-binding finger, the catalytic active domain contain Walker A, Walker Band arginine finger and N-terminal domain ***(Remus et al., 2009).***

MCM 5 expression:

In human cells *MCM5* gene was shown to be expressed widely in many normal tissues, but its mRNA levels vary a lot. Expression of all human genes of the *MCM* family is induced by growth stimulation and their mRNA levels peak at G1/S transition. This protein is mainly expressed in bone marrow hematopoietic cells, lymphocytes in tonsil, Liver, and trophoblastic cells in placenta. This DNA replication licensing factor is also expressed in a few other cell types, including colorectal glandular cells, epidermal cells of the skin and bronchus, urothelial cells of the urinary bladder, decidual cells of placenta, and glandular cells of the pre-menopause uterus, though at lower intensity. The growth-regulated expression of *MCM5* is primarily regulated by members of the E2F family through binding to multiple E2F sites of the *MCM5* gene promoter ***(Kearsey et al., 1998).***

MCM 5 and their role in DNA replication:

DNA Replication is the process of producing two identical DNA strand from one original DNA strand. Eukaryotic DNA replication is a conserved mechanism that controls DNA replication only once per cell cycle.Cell cycle contain four phases Gap1 (G1phase), S (synthesis) phase, Gap2 (G2phase) and mitosis (M). The vast majority of DNA synthesis occurs during S phase of the cell cycle ***(Alberts et al., 2002)*.**

To initiate DNA replication, multiple replicative proteins assemble to these replicative origins called Pre-replication complex. This process takes place at the G1 stage of the cell cycle. During G1phase, Pre-replication complex assemble including the origin recognition complex 1-6 (ORC), cell division cycle 6(Cdc6) protein, chromatin licensing and DNA replication factor 1(Cdt1), and minichromosome maintenance proteins (*MCM2*-7) figure (6) ***(Araki, 2011)***.

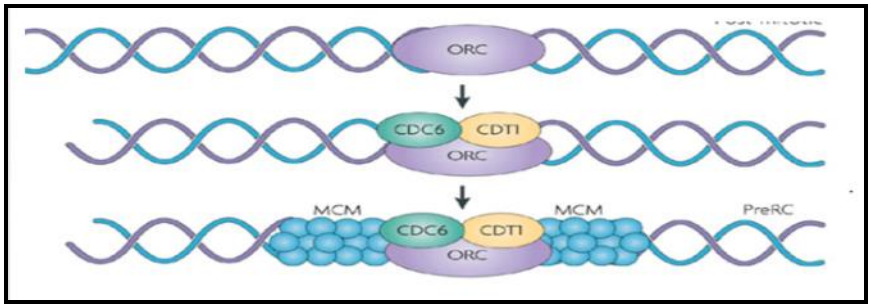

Figure (6): Prereplicative complex including the origin recognition complex (ORC), cell division cycle 6 (Cdc6) protein, chromatin licensing and DNA replication factor 1(Cdt1), and minichromosome maintenance proteins (*MCM2*-7**) *(Aladjem, 2007)*.**

The crucial event during DNA replication initiation is the loading of the *MCM2*-7 hetero hexameric complex that constitutes the core replicative helicase onto ORC at origins of replication. The *MCM2*-7 proteins do not have inherent affinity for DNA and so must be actively brought to origins. This is achieved through the action of the proteins Cdc6 and Cdt1 and is critically dependent on their ability to utilize ATP (figure 7) ***(Randell et al., 2006)*.**

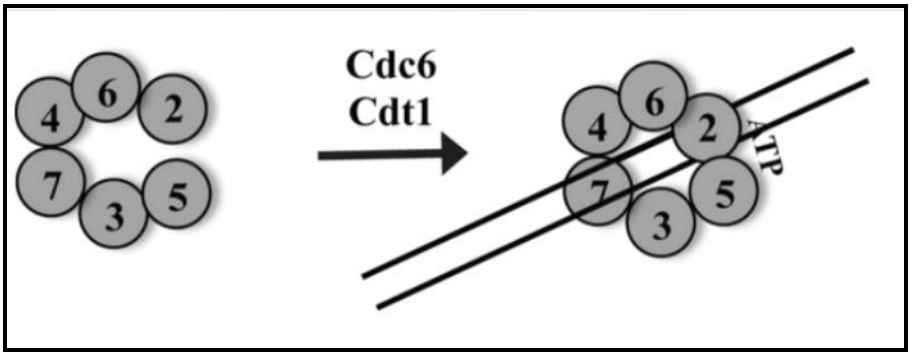

Figure (7): *MCM2*-7 hexamer assembles as an open complex that can be closed through ATP binding. The hexamer is loaded on to Orc1-6 bound ORI duplex DNA. This loading is dependent upon the loading factors Cdc6 and Cdt1. The hexamer can be closed loosely around the duplex through binding to ATP ***(Stuart, 2013)*.**

MCM2-7 are initially only loosely associated with DNA at the origin but ATP hydrolysis results in the complex being tightly bound, or "loaded". Once the *MCM2*-7 complex is loaded onto DNA the origin is "licensed" for replication ***(Sclafani and Holzen, 2007)*.**

The loaded complex is then activated during S phase by recruitment of the Cdc45 protein and the GINS complex to

form the active Cdc45–*MCM*–GINS (CMG) helicase at DNA replication forks. The full CMG complex is required for DNA unwinding, and the complex of CDC45-*MCM*-GINS is the functional DNA helicase in eukaryotic cells ***(Moyer et al., 2006)*.**

Licensed origins become active replication forks during origin firing. This depends on cell cycle regulated kinases that makes phosphorylation. *MCM* 2-7 complex is one of the targets of phosphorylation activity during origin firing. The most important of these kinases regulating the initiation of replication are cyclin A-dependent kinases (CDKs), cyclin E, the Dbf4- and Drf1-dependent kinase Cdc7 (also called DDK), all of which are essential for proper duplication of the genome ***(Labib, 2010)*.**

During the G1 phase of the cell cycle, cyclin-dependent kinase (CDK) levels are low. This condition allows formation of new pre-replication complexes (pre-RCs) but is insufficient to activate initiation of replication from these sites. During the S phase of the cell cycle, CDK levels increase dramatically, causing pre-RCs to direct the initiation of DNA replication. At the same time, the increased levels of CDK activity prevent new pre-RC formation. High levels of CDK activity are maintained until cell division, preventing new pre-RC formation until after cell division is complete (figure 8). Thus, the sequential formation and activation of pre-RCs can only occur in the context of a complete cell cycle ***(Araki et al., 2010)*.**

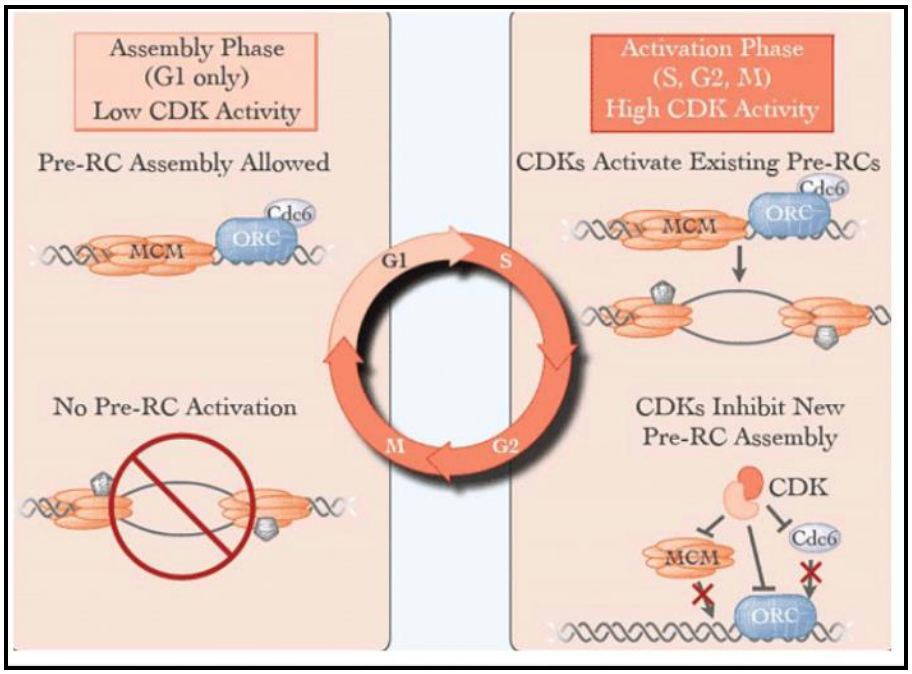

Figure (8): *MCM* is a target of CDKs phosphorylation which activate Existing Prereplication complex during S phase and inhibit new Prereplication complex Assembly during G_2 phase ***(Araki et al., 2010)*.**

Another control mechanisms are applied to ensure that origins are licensed once and only once during each passage through the cell cycle. After the start of S phase, Cdt1 is targeted for proteolytic degradation by phosphorylation and ubiquitination. And control mechanism involves the action of the novel regulator geminin. Geminin tightly binds Cdt1and through steric hindrance is able to stop binding of Cdt1 to *MCM*, thereby sequestering it away from the assembling pre-.RC (figure 9) ***(Lee et al., 2004)*.**

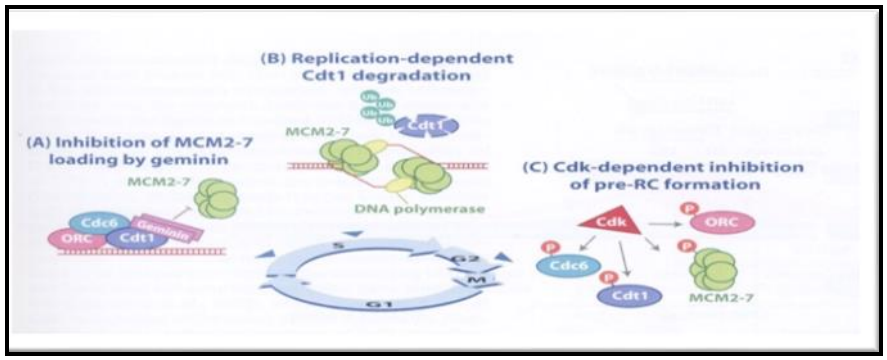

Figure (9): Control mechanisms of Eukaryotic DNA Replication including cyclin A-dependent kinases (CDKs) of prereplication complex ,Cdt1 is targeted for proteolytic degradation and Geminin that tightly binds Cdt1 and stop binding of Cdt1 to *MCM*, thereby inhibition of prerelication complex ***(Bell, 2009)*.**

MCM 5 and their role in Transcription:

*MCM*3-*MCM5* interacts with the transcription factor STAT1a (STAT1 alpha isoform), therefore *MCM5* has vital role also in transcription regulation. In fact, the integrity of the *MCM* hetero hexameric complex and the DNA helicase domain of *MCM5* are essential for the process of transcription. Increased levels of *MCM5* are associated with activation of transcription ***(Da Fonseca et al., 2001)*.**

Another study showed that the *MCM* complex is co-localized with RNA polymerase II (RNA Pol II) on chromatin of genes being constitutively transcribed, and that *MCM5* is required for transcription elongation of RNA Pol II. In fact, the integrity of the *MCM* hetero hexameric complex and the DNA helicase domain of *MCM5* are essential for the process of transcription. Additionally, human minichromosome maintenance proteins