Current methods of breast cancer diagnosis
Métodos actuales en el diagnóstico del cáncer de mama

Keila Gómez, Juan Carlos Laglaguano, Raúl Rodríguez

ABSTRACT: Breast cancer is the most frequent cancer among women, and represents the second highest number of deaths around the world. It’s difficult to diagnose because there are different types of breast cancer. The new methods for the detection of breast cancer are very promising because they guarantee that the patients received the specific kind of treatment. At present, there are breast cancer detection techniques which are very expensive, traditional and complicated, which leads us to the need to develop easy-to-use and more sensitive devices. In this review, we will analyze more recent methods of breast cancer diagnosis such as molecular signatures, micro-RNA, and bioinformatic methods.

Keywords: Breast cancer, diagnostic methods

RESUMEN: El cáncer de mama es el más frecuente entre las mujeres y representa el segundo mayor número de muertes en todo el mundo, es difícil de diagnosticar porque hay diferentes tipos de cáncer de mama. Los nuevos métodos para la detección de cáncer de mama son muy prometedores porque garantizan que los pacientes recibieron el tipo adecuado de tratamiento. En la actualidad, existen técnicas de detección de cáncer de mama que son muy costosas, tradicionales y complicadas, lo que nos lleva a la necesidad de desarrollar dispositivos más fáciles de usar y más sensibles. En esta revisión, analizaremos los métodos de diagnóstico de cáncer de mama más recientemente, como las firmas moleculares, el microARN y los métodos bioinformáticos.

Palabras claves: Cáncer de mama, métodos diagnósticos.

Introduction

Breast cancer is the most common cancer diagnosed in women around the world; Ecuador is not an exception (Tab. 1). Cancer starts when breast cells begin to grow in an uncontrolled manner way. To understand breast cancer, we must know the basic anatomical structure of the breast. Each breast is composed of lobules, lobes, and bulbs that are connected by ducts. Lobes are divided into smaller structures called lobules that end in tiny bulbs that can produce milk. Breast cancer is classified based on the anatomical structure and according to this classification, the most common types of breast cancer are the ductal carcinoma and invasive lobular carcinoma. Differentiating all breast tumor types is relevant in the clinical treatment case to guarantee that patients received the most appropriate kind of therapy. Currently, there are breast cancer detection techniques, which are very expensive, traditional, and complicated, which leads us to the need to develop innovative, determined, and ultrasensitive devices. In this review, we will conduct a study on the new methods of breast cancer diagnosis currently in use such as the use of molecular signatures, biosensors, micro RNA, and bioinformatics methods.

Molecular signatures

The acceptance of molecular profiles of DNA, RNA, and proteins reveals significant differences in fundamental biology, and this begins to have an impact on patients who are handling this technique. Diverse bioinformatics devices have been refined using DNA or RNA-based signatures to classify the illness into biologically and clinically essential subgroups. A characterization of cancer will obtain somatic mutations in the tumor genome, which include base substitutions, insertions, and deletions, changes in the number of copies and structural rearrangements. Nowadays, some mutations accumulate in the cell lineage, and individual or collective work gives a significant selective advantage to the tumor cell. Its form of distribution serves as a trace to discover underlying mutational processes that aid the development of the tumor. A characterization of cancer is that it obtains somatic mutations in the tumor genome, which includes substitutions of bases, insertions, and deletions, changes in the number of copies and structural rearrangements. A large number of classifiers are found which based on the number of copies for breast cancer. These classifiers can be the use of random forest variants, logistic regression and group loop, fused support vector machine, and the union of functions that have supervision or not which use silhouette techniques to estimate clusters. In the supervised as in the unsupervised FC, excellent performance found when ordering the CGH data of the tumor samples, due to the ability to exclude the unwanted correlation bias.

Biomarkers

In recent years, research in the field of molecular biology has contributed to the improvement in the diagnosis of can-
Tumor markers are biochemical indicators when a tumor is present: this incorporates cell surface antigens, cytoplasmic proteins, enzymes, and hormones. In clinical practice, the term is used to refer to molecules that are available in plasma, body fluids, solid tumors, circulating tumor cells, lymph nodes and bone marrow. The technique of immunohistochemical detection of specific antigens of the cell cycle is used to evaluate the proliferative activity of cells.

For example, Ki67 is located in the cell nucleus where it only binds to the perichromosomal layer in cells that divide and grow actively; this is used as a marker to determine cell proliferation. In consequence, changes in Ki67 expression in growing tumor cells compared to healthy cells can be used as an early predictor for treatment efficiency and can also be used as a prognostic factor for long-term outcomes in cancer patients.

Salivary biomarkers are being used to characterize breast cancer, correlating total salivary sialic acid and breast cancer. The sialic acid found in the final residues of the carbohydrate chains are biologically necessary and essential for the functioning of the glycoconjugates. In patients who have cancer, these markers are altered. The increase of sialylation of the glycoconjugates of the cell surface is localized between the critical molecular changes associated with malignant transformation and cancer progression. Studies have found higher levels of salivary sialic acid in patients who have breast cancer compared to the healthy control group. Finally, the future of tumor biomarkers is promising, because they provide information on the biological behavior of the tumor and this is of great importance because they reduce the mortality of patients who have cancer.

**Micro-RNA**

Micro-RNA are small noncoding RNA molecules around of 18-25 nucleotides in length. They regulate gene expression by repress or promote mRNA degradation. They can control multiple critical pathways that are involved in physiological as well as pathological processes. The downregulation of miRNA genes could result from aberrant hypermethylation, as in the case of miR-9-1 gene in breast cancer, or from histone deacetylation and tri-methylation, as in the case of miR-29 in B-cell lymphomas. miRNA is present in several biological fluids, among them breast milk.

Breast cancer could be classified according to different parameters such as histology, immunopathology, mRNA expression profiling, and miRNA expression signature. miRNA signature can sub-classify breast cancer, and can even determine new subtypes, as recently reported. MiRNA is expressed in breast cancer, as manifest by microarray profiling of tumor and normal breast tissues. Various studies reveal that diverse functions of the dysregulated miRNA in malignant breast transformation whereby they act as oncogenes (oncomiRNA) or as tumor suppressors. In breast cancer, miRNA has been shown to regulate cell cycle progression, apoptosis, angiogenesis, epithelial-mesenchymal transition, tumor microenvironment, invasion, metastasis and drug resistance, as well as the differentiation and self-renewal of breast cancer stem cells. For strength to endocrine treatments, the serum miRNA has the potential to serve as a biomarker of EBC.

**Diagnostic methods of microRNA in breast cancer**

Studies carried out had shown that the miR-21 that is the most researched onco miRNAs in this malignant disease serves as a diagnostic biomarker for EBC. The studies detected miR-21 in breast tissues using experimental technology as microarray/sequencing/RT-qPCR followed by validation in serum and plasma. Other studies were carried out in array panels on plasma samples followed by verification in plasma, and the technique that implemented was RT-qPCR or started with RT-qPCR on tissues then direct serum RT-qPCR. Signally, serum miR-21 exhibited higher sensitivity in EBC diagnosis than other conventional cancer markers, such as CA153 and CEA. In another study with the same method, the miR-155 shown to upregulated in serum and tissues of breast cancer patients. MiR-12b and miR-145 were the first to be downregulated and miR-21 to be upregulated in BC tumors compared to normal breast tissues. miR-125b, miR-145, and miR-21 both in human breast cancers and in breast cancer cell line. Moreover, all Northern blots confirmed results obtained by the microarray analysis. The miRNAs: miR-10b, miR-21 was used in diagnostic and prediction analysis of software microarrays, these algorithms used too for cross-validation. Another method of diagnostic is the Northern blotting analysis that were done in RNA sample (10 mg each) were electrophoresed on 15% acrylamide, 7 mol/L urea criterion precast gels and transferred onto Hybond-N+ membrane. The Hybridization was done at a temperature of 37° C in 7 % SDS/0.2 mol/L Na2PO4 (pH 7.0) for a time of sixteen hours. The membranes were washed at 42°C twice with two standard saline phosphate one mmol/L EDTA, and 0.1% SDS and again twice with 0.5 SSPE/0.1% SDS. The oligonucleotides used as probes are the complementary sequences of the mature miRNA, used to normalize expression levels. In the results the microarray analysis they carried out the study of the Northern blot of the differentially expressed miRNAs. They analyzed the expression of miR-125b, miR-145, and miR-21 both in human breast cancers and in breast cancer cell line. Moreover, all Northern blots confirmed results obtained by the microarray analysis. The miRNAs: miR-10b, miR-225b, miR-145, miR-21, and miR-155 were analyzed using three algorithms, miRanda, TargetScan, and PicTar, that are commonly used to predict human miRNA gene target.

**Bioinformatic methods**

There are many data about breast cancer such as medical records, clinical data or biomedical images are collected and available on databases. Some of the famous databases about breast cancer are the Wisconsin Breast Cancer Database and ArrayExpress. Wisconsin Breast Cancer Database is renowned because it offers information of computed features of the breast cancer cells nuclei. From digitized images, Wisconsin Breast Cancer Database offers an evaluation features such as radius, texture, perimeter, area, smoothness, compactness, concavity, hollow points, symmetry fractal dimension and diagnosis (malignant or benign) from digitized images of a fine-needle aspirate of a breast mass. Another database is Array Express, that is a public repository for microarray data. Their databases not only contain data about cancer but also from functional genomics experiments of thousands of clinical studies, and provides these data for the reuse to the research community.

Despite the existence of available data about breast cancer, there are no suitable methods to analyze the significant
amounts of data. It will be impossible for humans to process vast amounts of information so that computers have been fundamental for the development of science in general. Bioinformatics combines computer sciences, biological sciences, mathematics and engineering to process and interpret data. One of the most used fields of bioinformatics is artificial intelligence. In simple terms, artificial intelligence would be defined as the use of a computer to simulate human intelligence and ability to learn[46]. Data mining and machine learning are two related artificial intelligence technologies; they can be related to each other. Machine Learning refers to the study, design, and development of the algorithms that give computers the capability to learn without being explicitly programmed[47]. Data Mining is the process that starts from apparently unstructured data tries to extract knowledge and interesting unknown patterns. During this process, the use of machine learning techniques is necessary for most of the data mining techniques.

Data mining is the automated or convenient extraction of information patterns from data stored in massive information repositories, or data streams[48]. The primary medical applications of data mining are patient phenotype cohorts[49], pharmacovigilance[50], clinical pathways[51], healthcare process[52], disease progression[53], and deep learning for precision medicine and Human-computer interaction and knowledge discovery in dataset approach in biomedical informatics. Data mining techniques are based on training and testing the system regarding the characteristic or feature that have to be analyzed. The complexity of the data mining process depends on how extensive are the data to be examined and the number of elements to integrate into the mining process.

In the data mining is a multistep process[53]. The first step implies the collection, preprocessing, and normalization of data. Data is collected from heterogeneous sources and converted into homogenous. The second step is training and testing of the algorithms to obtain some meaningful information in an automatized way. The third step analyzes the processed data and represents it in a standardized format. Finally, the upshots of data mining are patient phenotype cohorts[48], analyzing DNA microarrays[55]. In general, data mining and machine learning are two related artificial intelligence technologies that are traumatic for the patients.

**References**

45. ArrayExpress < EMBL-EBI [Internet]. [cited 2017 Nov 27]. Available from: https://www.ebi.ac.uk/arrayexpress/

Received: 18 november 2017
Approved: 20 january 2018