Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN
"MOLECULAR MEDICINE AND PHARMACOLOGY"

CICLO XXX

COORDINATORE Prof. Di Virgilio Francesco

GATA3 IS AN ADJUNCT PROGNOSTIC FACTOR IN BREAST CANCER PATIENTS,

ESPECIALLY WITH LESS AGGRESSIVE DISEASE

Settore Scientifico Disciplinare MED/08-BIO/16

Dottorando
Dott.ssa Kuhn Elisabetta

Tutore
Prof.ssa Secchiero Paola

Co-Tutore
Prof.ssa Querzoli Patrizia

Anni 2014/2017

TABLE OF CONTENTS

INTRODUCTION

Breast cancer epidemiology

Invasive breast cancer (BC) is by far the most frequently diagnosed cancer in women worldwide, and the principal cause of cancer-related death among women (1, 2). Age-standardized incidence rates are 2-fold greater in countries with very high-levels of development when compared with low-level developed countries. On the other hand, the mortality rate is lower in developed countries. Interestingly, whereas incidence has been progressively increasing in most countries of the world, it has peaked and declined over the past decade in a number of highly developed countries. In Italy, there are more than 50.000 reported new cases of BCs yearly and BC accounts for almost 13.000 cancer deaths each year, equal to 31% of all cancer in women and 17% of all cancer deaths in women, respectively (Figure 1) (2).

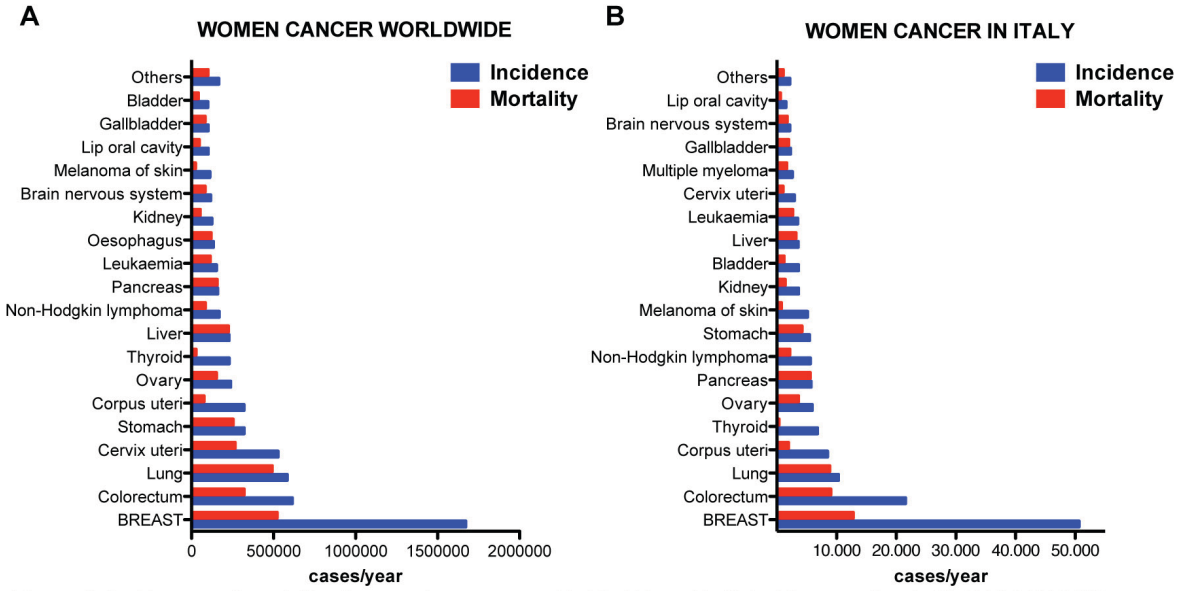

Figure 1. Incidence and mortality of cancer in women worldwide (A) and in Italy (B), according to GLOBOCAN 2012.

Several risk factors have been associated with BC, but the vast majority of women developing BC does not have any identifiable risk factor (3). In particular, hormonal and reproductive factors that prolong the exposure to estrogens, such as early menarche, nulliparity, late age at first childbirth, late menopause, and hormonal intake (either oral contraceptive or menopausal replacement therapy), are the principal risk factors for BC (3). Other established non-modifiable risk factors are previous family history, and breast tissue density. In addition, various modifiable factors contribute to the risk of developing BC. These are mainly lifestyle factors (i.e. alcohol use, high-calorie diets, and physical inactivity) and environmental factors (exposure to ionizing radiation). A minority (5-10%) of the total BCs is due to familiar predisposition, correlated with germline mutation of two high-risk, high penetrance genes: *BRCA1* and *BRCA2* (4, 5). These gene mutations, even if are rare, greatly increase the risk of developing BC. Moreover, several lower-penetrance genes as well as many loci within the genome, for which the specific genes have yet to be

identified, have been correlated with an increased BC risk. Hence, the etiology of BC is multifactorial and implicates in BC development both familiar and personal factors as well as reproductive and lifestyle factors (Figure 2).

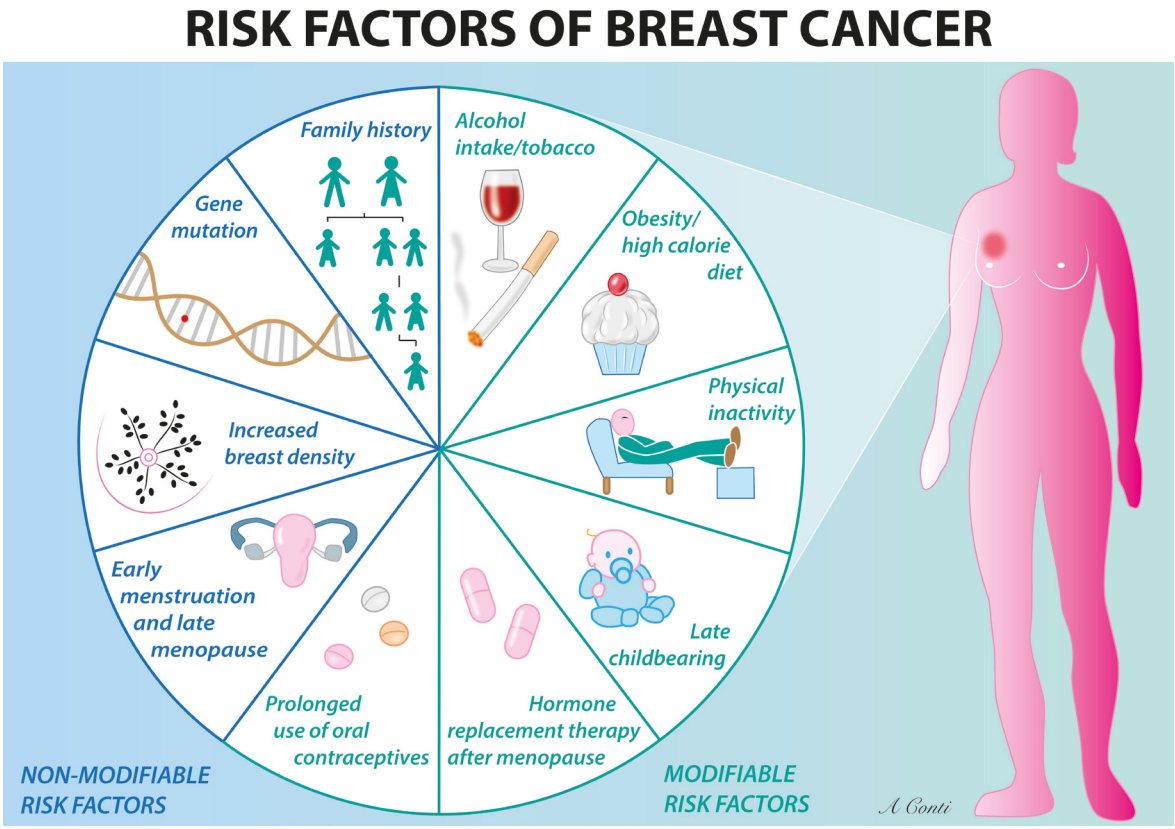

Figure 2. Risk factors that contribute to breast cancer development can be divided into non-modifiable and modifiable.

Histological classification

It is now emphasized that BC is not a single disease, but a heterogeneous group of diseases, both morphologically and clinically. The last WHO Classification of Tumours of the Breast reports more than 20 different subtypes of BC. The majority of BCs derives from epithelial cells and are roughly subdivided into in situ and invasive carcinomas. Carcinoma in situ (CIS) is the preinvasive neoplasia and is composed of malignant epithelial cells that are confined to the ductal-lobular structures of the breast, without penetration of the basement membrane, that appears integral. CISs are further subclassified in ductal (DCIS) and lobular (LCIS) histotypes. DCIS and LCIS differ in cytological features and architectural growth of the cells, but also in clinical characteristics, such as anatomical distribution, bilaterality, and clinical outcome. Specifically, LCIS frequently is multifocal and multicentric and involves bilaterally the breast, differently from DCIS. Moreover, LCIS shows a very low tendency to transformation and invasion. However, a more aggressive variant has been described, pleomorphic LCIS, that seems to carry a greater tendency to transformation. DCIS is further graded according to a three-tiered system mainly based on cytological features (6). Low-grade and high-grade DCISs are distinct disorders, of which high-grade DCIS progress more rapidly and frequently to invasive BC. Both DCIS and LCIS origin from the terminal duct lobular unit of the breast, that is the microanatomical and functional unit of the mammary gland; therefore, the terminology ductal and lobular does not refer to distinct sites of origin (duct vs lobule), but rather to intrinsic differences between these two neoplasias that manifest morphologically and dictate their contradistinctive biological behaviors (7).

Invasive carcinoma named “no special type” (previously known as invasive ductal carcinoma) is the most common BC, accounting for more than 70% of BC cases (Figure 3). This histotype is characterized by lack of specific morphological features as opposed to special type BCs. The most common special types of BC include lobular carcinoma, metaplastic, mucinous, and micropapillary carcinoma, and carcinoma with apocrine differentiation (Table 1). In addition to histotype, invasive BCs are routinely graded according to a semi-quantitative method based on three morphological features: 1) tubule/gland formation; 2) nuclear pleomorphism; and 3) mitotic count (Table 2) (8).

Importantly, histological grade is a powerful prognostic factor, that significantly correlates with BC patient survival.

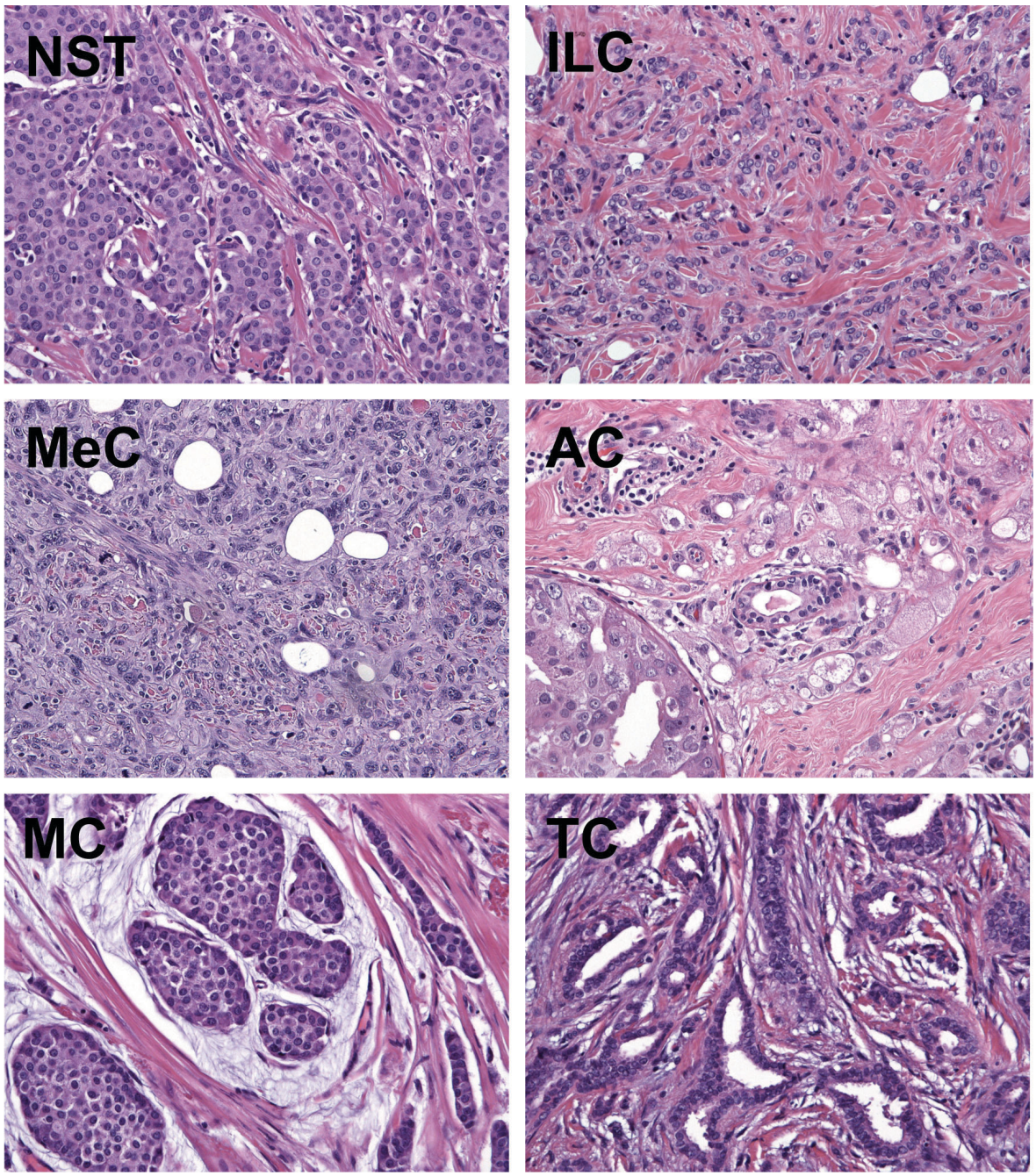

Figure 3. Representative images of the most common BC histotypes: invasive carcinoma of no special type (NST), invasive lobular carcinoma (ILC), metaplastic carcinoma (MeC), apocrine carcinoma (AC), mucinous carcinoma (MC) and tubular carcinoma (TC).

Table 1. Most common histotypes of Invasive Breast Carcinomas according to 2012 WHO classification and frequency (1)

Histological type	Frequency
Invasive carcinoma of no special type (NST)	50-60 %
Invasive lobular carcinoma	5–15 %
Metaplastic carcinoma	<5 %
Carcinoma with apocrine differentiation	4 %
Mucinous carcinoma	2 %
Tubular carcinoma	2 %
Invasive micropapillary carcinoma	1–2 %
Carcinoma with medullary features	<1%
Carcinoma with neuroendocrine features	<1%
Cribriform carcinoma	0.3-0.8 %

Table 2. Nottingham histologic score system

Feature	Score
Tubule and gland formation	
>75%	1
10-75%	2
<10%	3
Nuclear pleomorphism	
Small regular uniform cells	1
Moderate increased size and variability	2
Marked variation	3
Mitotic count	
Low ≤4	1
Moderate 5-9	2
High ≥10	3
Final grading	
Grade 1	3-5
Grade 2	6-7
Grade 3	8-9

Prognostic biological markers

Currently, the established biological biomarkers predicting BC prognosis and response to therapy are estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2), while Ki-67 remains yet controversial. These markers are routinely tested in all primary BCs by immunohistochemistry (IHC) and scored according to the American Society of Clinical Oncology and the College of American Pathologists (ASCO/CAP) recommendations (9, 10). These guidelines have fostered standardization and reproducibility of IHC scoring. Based on ASCO/CAP guidelines, hormonal receptor IHC must be reported as the percentage of immunoreactive invasive BC cells and BC is regarded as positive when shows positive cells ≥1%. Notably, normal breast cells should be used as an internal positive control. Approximately 80% of BCs are ER-positive and ER IHC positivity correlates with histologic grade 1-2 and lobular, mucinous, and tubular histotypes. On the contrary, ER negativity associates with grade 3 and metaplastic, medullary, and apocrine carcinomas (1). Molecularly, ER-positive BCs lack both *HER2* amplification and *TP53* mutation (11-13). PR strongly correlates with ER, but demonstrates a less predictive power.

HER2/ERBB2 is an oncogene that encodes for a transmembrane tyrosine kinase, a component of the family of epidermal growth factor receptors (14-16). Globally, 15–20% of invasive BCs are HER2-positive and show HER2 protein expression and gene amplification (10). The introduction of anti-HER2 targeted therapy has revolutionized BC treatment, improving considerably the outcome of BC women (17-21). Routinely, two methods are used to test HER2 status, either immunohistochemistry (IHC) or in situ hybridization (ISH), both assessed following the ASCO/CAP guidelines (Table 3 and Figure 4)(10, 22). A good correlation has been found between these two tests (23). Of note, anti-HER2 therapy provides noticeable benefits only in patients with HER2-positive BC (24). HER2-positivity histologically associates with grade 3 and lack of ER and PR

expression (13).

Table 3. HER2 classification tested by IHC and dual-color fluorescence in situ hybridization according to 2013 ASCO/CAP guidelines

HER2 IHC (membrane staining)			
0	**1+**	**2+**	**3+**
no staining or incomplete, faint/barely perceptible ≤10%	incomplete, faint/barely perceptible >10%	incomplete, weak/moderate > 10% or complete, intense ≤ 10%	complete, intense >10%
***HER2* dual-color FISH**			
Negative	**Equivocal**		**Positive**
HER2/CEP17 <2 & *HER2* GCN <4	*HER2*/CEP17 <2 & *HER2* GCN ≥4 & <6		*HER2*/CEP17 ≥2 or HER2 GCN ≥6

IHC, immunohistochemistry; FISH, fluorescence in situ hybridization.

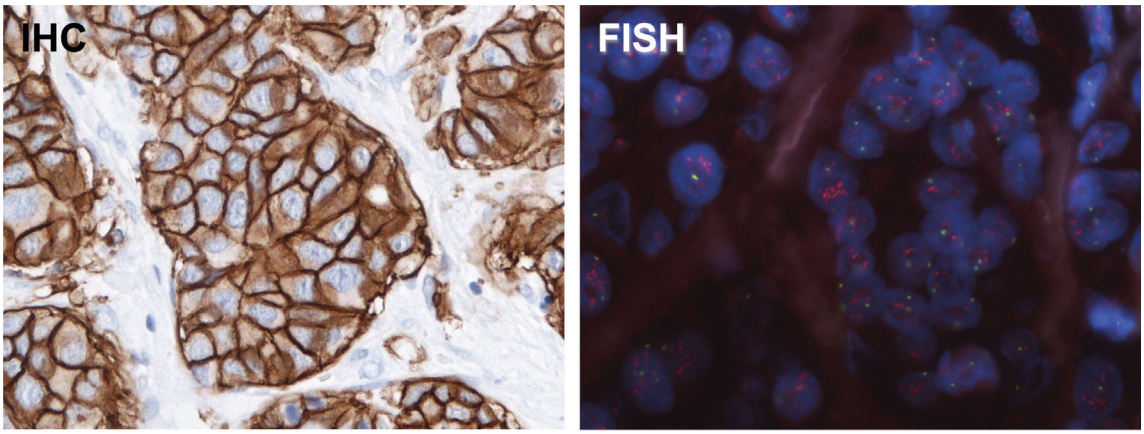

Figure 4. HER2-positive BC as demonstrated by immunohistochemistry (IHC) and dual-color fluorescence in situ hybridization (FISH), where red dots are *HER2* gene copies and green dots the centromeric chromosome 17 probes.

Molecular classification

In 2000 Perou *et al.* proposed in a revolutionary study to subclassify BCs according to the gene expression profiles obtained by microarray platform (25). They initially recognized 5 different molecular subtypes: type A luminal, type B luminal, HER2 enriched, basal-like and normal breast-like. This BC subtyping has progressively evolved into a molecular classification, in which every specific subtype not only differs in gene expression but also in clinical, biologic, histologic features, and treatment response (Table 4) (25-27).

Table 4. Molecular subtypes based on gene expression

CHARACTERISTICS	LUMINAL A	LUMINAL B	HER2-ENRICHED	BASAL-LIKE*
Gene expression	LMWCKs and high hormone receptors and related genes	LMWCKs and moderate-weak PR and related genes	High HER2 and other genes in amplicon on 17q12 Low ER and related genes	High basal epithelial genes Basal cytokeratins Low ER and related genes Low HER2 related genes
Prevalence	~60%	~10%	~15%	~15%
Biologic features	ER/PR positive HER2 negative Low proliferation rate	ER positive, PR low positive HER2 expression variable Intermediate-high proliferation rate Luminal B tends to be higher histologic grade than luminal A	ER/PR negative HER2 positive High proliferation rate *TP53* mutation common Node positive	Triple negative High proliferation rate *TP53* mutation common *BRCA1* dysfunction African–American women
Histology	Tubular carcinoma Cribriform carcinoma Low-grade NST Classic lobular carcinoma	NST Micropapillary carcinoma	High-grade NST	High-grade NST Metaplastic carcinoma Carcinoma with medullary features
Target therapy	Endocrine therapy	Endocrine therapy Response may not be as good as for luminal A	Trastuzumab	No response to endocrine therapy/trastuzumab
Chemotherapy response	Generally not indicated	Variable (greater than luminal A)	Anthracycline-based	Platinum-based and PARP inhibitors
Prognosis	Good	Not as good as for luminal A	Generally poor, but improved with anti-HER2	Generally poor (but not uniformly poor)

LMWCKs, low molecular weight cytokeratins; *ER*, estrogen receptor; *PR*, progesterone receptor; *NST*, no special type, *PARP*, poly-adenosine diphosphate-ribose polymerase.
*Some low-grade BC special types, such as adenoid cystic carcinoma, secretory carcinoma, are basal-like and triple negative, but with low proliferation .

The normal breast-like subtype is currently considered an artifact due to lack of tumor cells in the examined samples.

Much effort has been spent in order to translate molecular classification into the clinic. To this end, clinico-pathological surrogates of molecular subtypes based on IHC have been proposed (Table 5) (26-28). However, still many controversies regarding the IHC cutoffs are yet unresolved and discordant cases, between the molecular subtypes according to gene expression versus IHC, are relatively common.

It is important to underline that some special types of BC, such as acinic, adenoid cystic, apocrine, and secretory carcinomas, despite a triple-negative immunoprofile display an indolent behavior and a favorable outcome. Therefore, molecular subtypes should be always considered in view of the histological features.

Table 5. Surrogate definitions of intrinsic subtypes of breast cancer

SURROGATE DEFINITIONS	LUMINAL A-LIKE	LUMINAL B-LIKE (HER2-)	LUMINAL B-LIKE (HER2+)	HER2-POSITIVE	TRIPLE NEGATIVE
Intrinsic subtype	Luminal A	Luminal B		HER2-enriched	Basal-like
Immunophenotype	HER2 negative ER positive PR positive Ki67 low (<20%)	HER2 negative ER positive *Either one:*-PR low (<20%) -Ki67 high (≥20%)	HER2 positive ER positive Any PR Any Ki67	HER2 positive ER negative PR negative	HER2 negative ER negative PR negative

ER, estrogen receptor; PR, progesterone receptor. ER/PR positive are ≥1%.

Lastly, over last decade many genomic prognostic tests (e.g. Oncotype Dx, MammaPrint, PAM50), that predict the risk of recurrence in luminal-type BCs, and, coherently, indicate the utility of adjuvant chemotherapy, have been commercialized (28-30). These multigene assays intend to identify the group of patients with node-negative, luminal-like BC that would not benefit from the chemotherapy based on the biology of their cancer.

Molecular genetics

The rapid widespread use of next-generation sequencing techniques has allowed for the genomic characterization of many thousands of BCs in record time (31, 32). The emerged genomic landscape is highly fragmented with the majority of genes mutated in a small percentage of BC totality. The genes consistently affected in BC are *AKT1, CDH1, CDKN1B, GATA3*, *MAP3K1, MLL3*, *PIK3CA*, *RB1*, *PTEN*, and *TP53*, with only 3 genes (i.e. *GATA3*, *PIK3CA,* and *TP53*) mutated in more than 10% of BC cases (31, 33). Furthermore, specific genetic alterations contradistinguish and are pathognomonic of some BC special types, including: *CDH1* loss of lobular carcinoma; *MYB–NFIB* gene fusion, due to chromosomal translocation t(6;9), of adenoid cystic carcinoma; and *ETV6-NTRK3* rearrangement, due to translocation t(12;15), of secretory carcinoma.

Current therapy for breast carcinoma

The therapeutic approach to BC is often multimodal and contemplates surgery, radiation therapy, hormonal therapy, chemotherapy, and targeted therapy, depending on the extent of the disease and BC intrinsic subtype (Figure 5) (34).

Surgical therapeutical options are partial resection (lumpectomy, segmentectomy, or quadrantectomy) and simple or total mastectomy. Many factors contribute to the selection of the procedure, including tumor stage, breast size, patient wish, availability of reconstructive surgery, and the surgeon's practice (35). Surgical margins must be evaluated and guide decision-making regarding the necessity of further additional surgery and radiation therapy. Positive margins are associated with an increased risk of local recurrence, as much as distant metastasis (36). Postoperative radiotherapy is often employed, especially after conservative surgical procedures, as well as in high-risk postmastectomy patients, to reduce locoregional recurrence (37).

The biopsy of the axilla sentinel lymph node is a standard surgical procedure for axillary staging in BC patients. This practice is associated with less morbidity than axillary lymphadenectomy, that is however required in case of metastatic sentinel lymph node.

Endocrine therapy is the standard treatment for all patients with ER-positive disease (38). At present, it uses either a selective ER modulator (e.g., tamoxifen) or an aromatase inhibitor (e.g., anastrozole) in postmenopausal women, usually in combination with radiotherapy, with or without adjuvant chemotherapy, depending on various parameters.

Trastuzumab is an anti-HER2 monoclonal antibody effective against HER2-positive BCs. This drug is now indicated for patients with early-stage and metastatic disease (39). Furthermore, newer agents against HER2 and its related receptors are in development as alternate or second-line therapies.

Chemotherapy has improved significantly the survival of patients with advanced primary BC, particularly in combination regimens, but is also currently administered to patients with positive axillary nodes, with curative intent (28, 40). Recent studies have shown promising effectiveness of platinum-based chemotherapeutic agents and of PARP (i.e poly-adenosine diphosphate-ribose polymerase) inhibitors in the treatment of BCs with specific DNA repair defects, comprising *BRCA*-mutated and basal-type BCs, but clear evidence of their superiority over conventional regimens is still missing (41-43).

Notably, triple-negative BC (TNBC) patients are not candidates for hormone and anti-HER2 target therapies currently used for other BC subtypes. Although TNBCs are chemosensitive to conventional therapy with anthracycline or anthracycline/taxane and, based on recent evidence, show a specific sensitivity to platinum-based regimens, ionizing radiation, and PARP inhibitors (44, 45). However, unfortunately, TNBC often develops chemoresistance and manifests local recurrence and distant metastases. For this reason, it is urgent the need to develop novel therapeutic approaches for this disease.

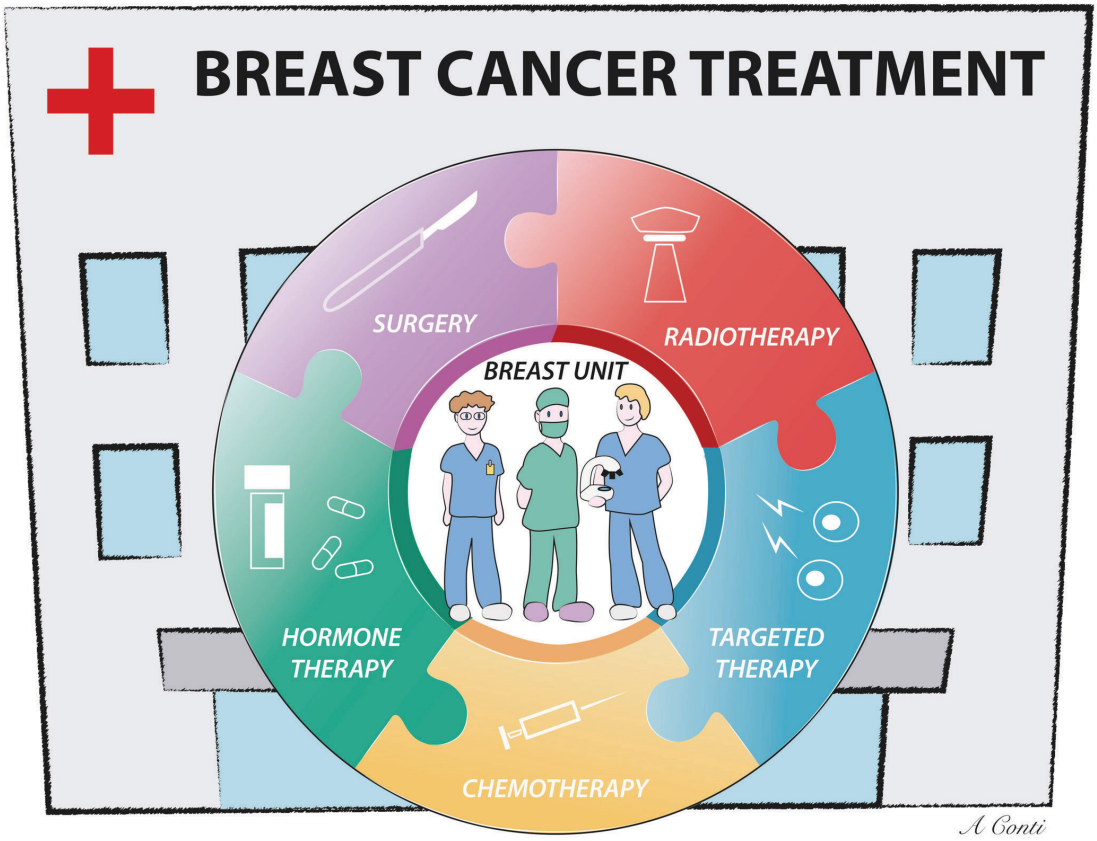

Figure 5. Schematic cartoon of the multimodal treatment for breast cancer patients.

Immunotherapeutics represent an interesting and effective alternative therapeutic option for malignant melanoma and other cancers (46). In particular, the blockade of immune checkpoints, boosting the anticancer immune response by disruption of immune self-tolerance, is emerging as a novel approach in cancer immunotherapy (47). In this regard, the programmed death 1 (PD-1) signaling pathway represents a key immune checkpoint that plays a pivotal role in autoimmunity and peripheral tolerance (48, 49). The activation of PD-1 pathway, inhibiting T cell functions, results significantly abnormal in autoimmune diseases, infections, and cancers. Therefore, PD-1 signaling pathway is currently explored as a target for anti-cancer drugs, and antibodies against PD-1 and PD-L1 have been generated. Encouraging preliminary results have shown antitumor activity of these antibodies in BC, hence, they are currently under investigation in clinical trials for BC patients, specifically with TNBC (50, 51). Nevertheless, predictive factors of response to these antibodies have not been univocally identified (52).

GATA binding protein 3

The GATA family proteins are lineage-specific transcription factors structurally characterized by a central DNA binding domain, composed of two highly conserved zinc fingers, that binds to palindromic GATA motifs and induces chromosome looping, thereby causing global gene expression changes (53). GATA binding protein 3 (GATA3), a member of the GATA family, has been shown to be essential for tissue differentiation of many organs, in particular, the mammary gland and T cells (54-56). Specifically, GATA3 plays a critical role in the morphogenesis of the mammary gland and in the luminal differentiation of breast epithelial cells (54, 56).

Inasmuch as GATA3 expression in human BCs is positively correlated with ER expression and is required for cell cycle progression of ER-positive cell lines, noticeably, GATA3 expression is coregulated with that of ER. In particular, a seminal study by Eeckhoute *et al.* has demonstrated that the direct binding of GATA3 to two-cis regulatory elements within the ERα gene promoter is needed for ERα transcription, allowing the recruitment of RNA polymerase II to ERα promoters (57). Accordingly, GATA3 silencing weakened ERα target genes, including PR. On the other hand, ERα directly activated GATA3 transcription; hence a positive cross-regulatory loop ties GATA3 to ERα (Figure 6A) (58). The transcription coordinated regulation of GATA3 and ERα likely underpins the significant coexpression of these two genes in BC. In addition, the same study has identified GATA3 as a central element of the signal transduction cascade to estradiol in ER-positive BC, together with ERα and FOXA1.

Over the past decade, our understanding of the molecular mechanisms through which estrogen promotes gene transcription cascade and ERα exerts its transcription factor functions has been further broadened. Briefly, estradiol stimulation has been shown to induce the assembly of a “mega transcription factor complex” recruited by ERα dimer at the DNA estrogen-responsive element-containing enhancers (Figure 6B)(59, 60). Specifically, this complex clusters ERα and GATA3 with other transcription factors and co-activators, and the presence of the tripartite enhanceosome composed of the three transcription factors ERα, GATA3, and FOXA1 is necessary for the full estrogen-induced transcriptional activation in BC (59). On the other hand, only the introduction of all these three transcription factors in ERα–negative BC cell lines can reverse their ER-negative status and restore estrogen-dependent biological functions. Finally, GATA3 owns the ability to bind to condensed chromatin and to remodel chromatin accessibility, as such acting as a pioneer factor (53, 61). Therefore, GATA3 shapes the histone modification landscape, changing the chromatin structure and, subsequently, the estrogen-responsive elements accessibility, and the downstream transcription signaling.

Functionally, in BC cell line and murine models, GATA3 expression reduces the tumor-

initiating ability, the epithelial to mesenchymal transition, and the metastatic potential (62). Consequently, GATA3 lack leads to a chemoresistant and mesenchymal phenotype of BC (63, 64). Furthermore, in ER-negative cell lines, GATA3 interacts with wild-type BRCA1, but is incapable to bind to mutant BRCA1 (65). This finding is important since BRCA1-GATA3 disruption could be the biochemical mechanism underlining the aggressive behavior of basal-like BC, which frequently harbors *BRCA1* inactivation.

The importance of GATA3 in breast biology is further emphasized by the fact that *GATA3* is one of the few recurrently mutated genes in BCs, affecting from 12% to 16% of breast carcinomas, across all different subtypes (31, 32). Interestingly, specific *GATA3* mutations cause of a congenital developmental disorder characterized by hypoparathyroidism, sensorineural deafness, and renal insufficiency (so-called HDR syndrome) have been identified. Similarly to HDR syndrome mutations, most of the *GATA3* somatic mutations found in BC occur in the C-terminal zinc finger region, revealing the key role of this region in the normal functionality of the protein (31, 32).

Recently, GATA3 has emerged as a sensitive immunohistochemical marker for breast origin, useful mainly for TNBCs that, by definition, tend to be negative for the other breast specific markers. Furthermore, several studies have investigated GATA3 as a prognostic marker in BC patients, with conflicting findings. On one hand, both GATA3 mutations and protein expression have been associated with better prognosis and improved survival (66-68). On the other hand, GATA3 expression closely parallels ER expression in BCs, so that it does not seem a prognostic factor independent of ER status (69). Hence, whether GATA3 carries independent prognostic information in BC patients yet remains an open question.

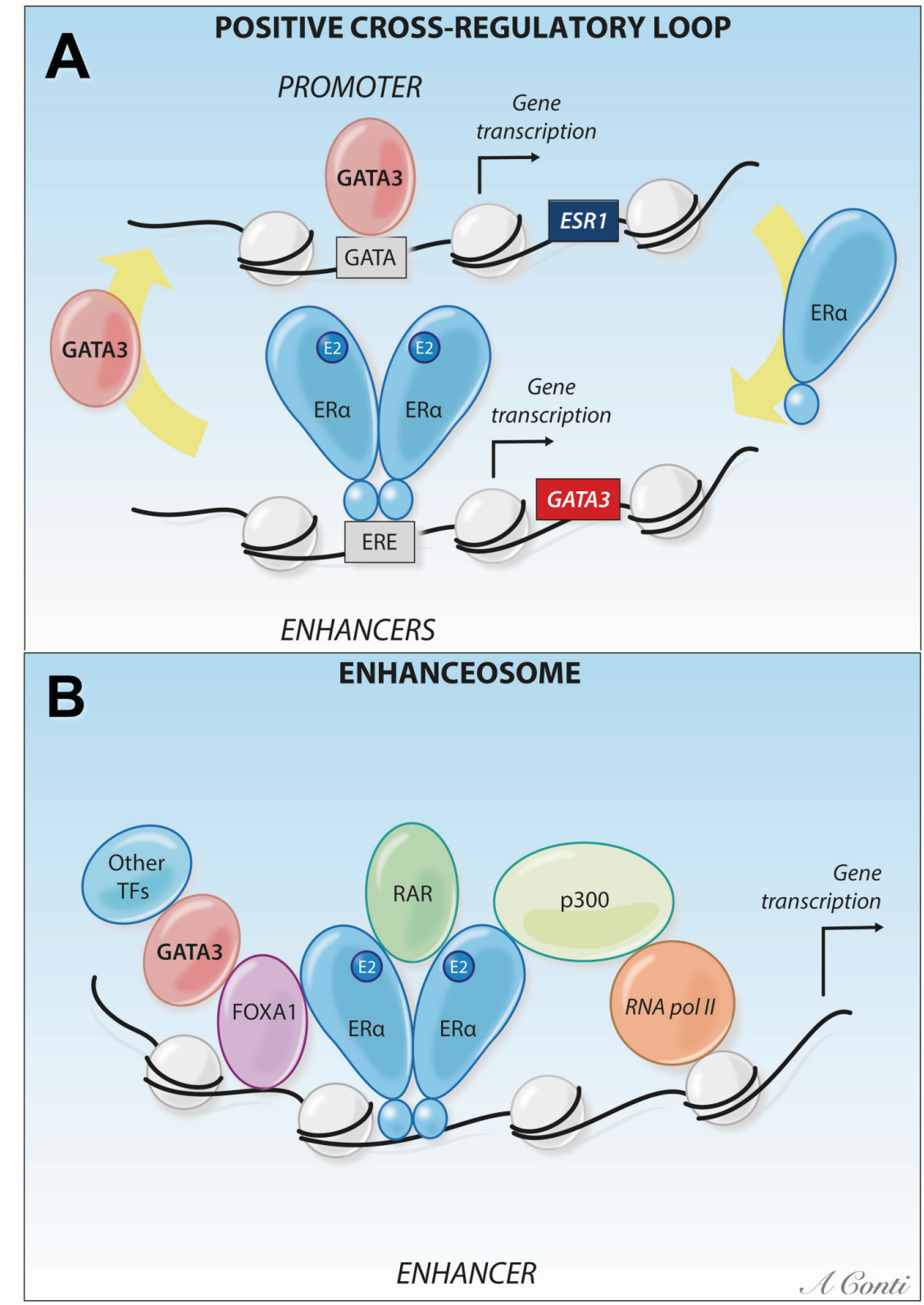

Figure 6. A) GATA3 and ERα expression is coordinated through a positive cross-regulatory feedback loop. GATA3 stimulates *ESR1* transcription by interacting with its binding sites (GATA) in the *ESR1* promoter, vice versa, in the presence of estradiol (E2) ERα stimulates *GATA3* transcription by binding to estrogen-responsive element (ERE) DNA sequences at the enhancer regions of the *GATA3* gene. **B)** E2 stimulation induces the assembly of an enhanceosome, an higher-order protein complex recruited by ERα dimer at the ERE-containing enhancers, where the presence of the three transcription factors ERα, GATA3, and FOXA1 is necessary for the full estrogen-induced transcriptional activation in breast carcinoma.

HYPOTHESIS

We hypothesize that GATA3 status correlates with clinico-pathological characteristics, biological markers, p53 status, and clinical outcome of invasive BC patients. To test the above hypotheses, we propose the following three specific aims.

SPECIFIC AIMS

Aim 1: to perform IHC of GATA3 and the conventional biological markers in a large monoinstitutional series of invasive BCs.

Aim 2: to measure the association of GATA3 IHC expression with biological markers, p53 status, and clinico-pathological characteristics.

Aim 3: to determine the prognostic value of GATA3 IHC expression in BC patients, according to specific BC subgroups.

MATERIALS AND METHODS

Patients

A total of 702 consecutive primary invasive BCs were retrieved from the pathological files of our institution and included in this study. All BC patients were diagnosed and surgically treated, from January 1989 to December 1993, at the Surgical Units of the S. Anna University Hospital of Ferrara or of the Ferrara province hospitals. Only cases with available tumor formalin-fixed and paraffin-embedded blocks, from female patients, neither associated with systemic metastasis nor undergone neoadjuvant treatment, were included in this study.

All tumors were categorized according to the WHO classification, the 8th AJCC staging system, and graded according to the Elston-Ellis grading system (Table 2) (1, 70). The molecular classification of BC was based on surrogate definitions by mean of immunohistochemical markers for ER, PR, HER2, and Ki-67, according to the criteria of the 2011 St. Gallen International Breast Cancer Conference (Table 5) (34).

All patients were treated according to our institution guidelines. After the first-line treatment was completed, the patients were re-examined twice a year for the first five years and annually for the following years. Clinical baseline data, including patient age, tumor histology, pathologic stage, grading, and follow-up data (date and site of relapse, last follow-up time, date of death, and cause of death) were retrospectively collected from the pathology files and the Ferrara Cancer Registry. Tissue collection was conformed to the Institutional Research Board regulations of the University-Hospitals of Ferrara. The protocol of this study was approved by the board of the Ministry of the University and Research ('Identification and validation of new markers of metastasizing phenotype of breast cancer', prot. MM06095812_006, 2000).

Tissue Microarray Construction

Tissue microarray (TMA) blocks were built as follows (Figure 7). All available hematoxylin-eosin (HE) stained slides of the selected cases were revaluated by three pathologists, in order to confirm the pathological diagnosis and to select one slide with viable invasive BC to be included in the TMA block. One representative area for each tumor was selected and marked on the HE slide. Corresponding formalin-fixed, paraffin-embedded tissue blocks were retrieved from the hospital archives (donor blocks) and 1 tissue core was extracted from the corresponding donor block marked area by a tissue punch extractor, using the marked HE slide as a guidance. A true-cut needle (4 mm in internal diameter) was used to punch 3-mm spaced holes in the recipient block. A single 4-mm tumor core per donor block was transferred to the recipient TMA block (including 23 different BC samples and 1 reference core of normal liver or lung tissue). The complete BC series was included in a total of 31 TMA blocks. TMA blocks were then

incubated for 15 minutes at 37° C to allow the tumor cores to firmly adhere to the recipient block. Six consecutive 4-µm thick sections were cut and mounted on silanized slides (71).

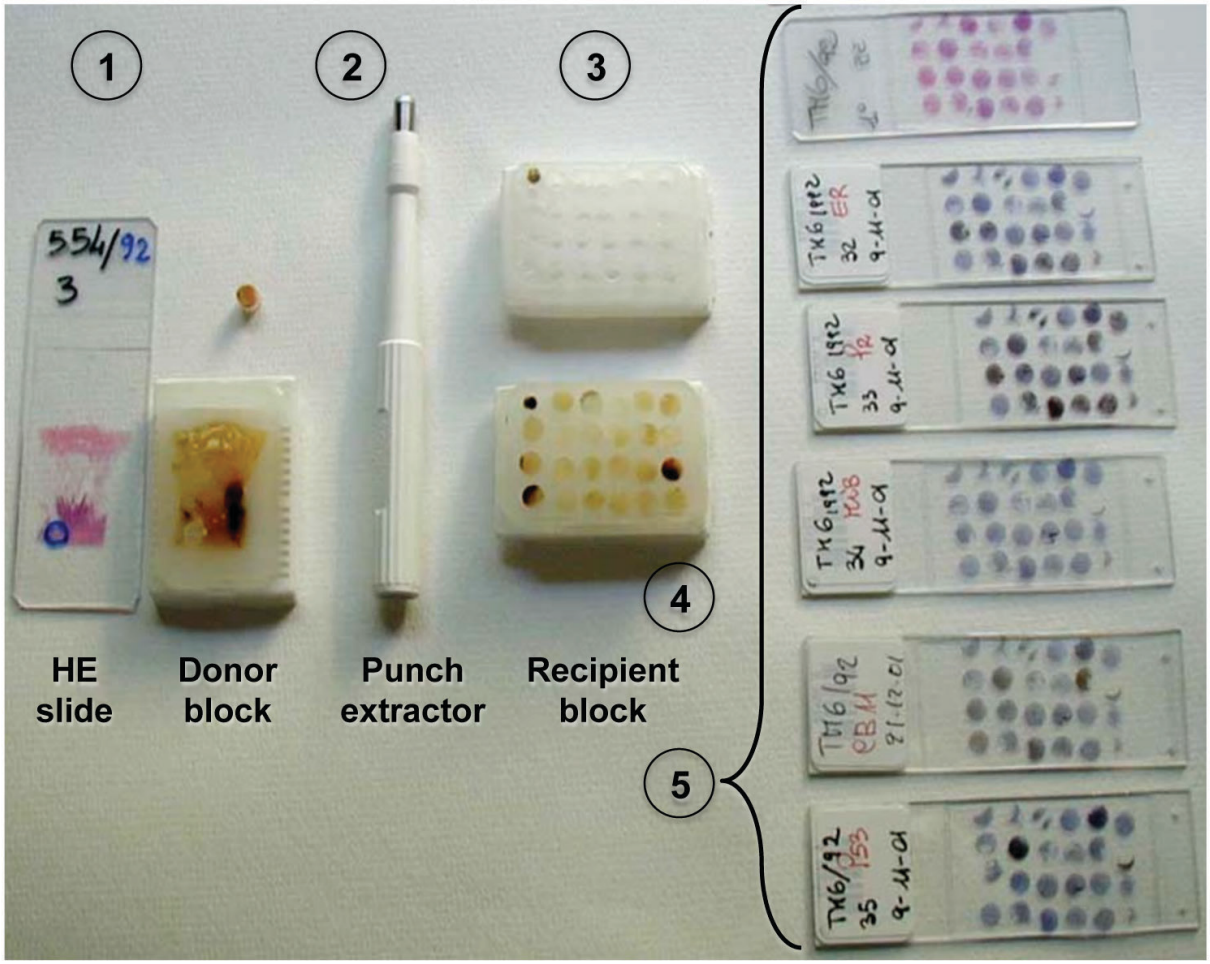

Figure 7 Tissue microarray construction. All hematoxylin-eosin (HE) stained slides were reviewed and one was selected. A representative breast cancer tissue area was marked with a glass marking pencil (1); using the marked HE slide as a guidance, 4 mm diameter cores of tumor tissue were removed from each donor block using a tissue punch extractor (2); recipient block was previously prepared using a true-cut needle (4 mm in internal diameter) to punch 3-mm spaced holes (3); BC tissue cores were then transferred in the recipient block (4); and 6 consecutive 4-µm thick tissue microarray sections were mounted on silanized slides and stained (5).

Immunohistochemical staining

First, one HE-stained slide from the TMA series was prepared and examined to confirm the presence of invasive BC in each core. Then, further TMA sections were stained either manually or by the Ventana NexES automated immunostainer (Ventana Medical Systems/Roche, Tucson, AZ), using the primary antibodies and conditions as reported in Table 6. Briefly, 4-µm thick formalin-fixed, paraffin-embedded sections were deparaffinized in xylene, rehydrated in graded alcohols and incubated in 3% H_2O_2 in methanol for 10 minutes to block endogenous peroxidase activity. Antigen retrieval was carried out by incubating the slides in Tris-EDTA-citrate buffer (pH 7.8) in a microwave

oven, prior to application of primary antibody. Vectastain ABC peroxidase kit (Vector Laboratories, DBA Italia, Segrate, Italy) was applied to reveal antibody binding. The slides were counterstained with hematoxylin for 3 minutes (cat # 790–2208, Ventana Medical Systems/Roche), dehydrated in the graded ethanols and xylene. For each antibody, a positive control slide and a negative control slide, where the primary antibody was replaced with normal serum or isotype-matched antibodies, were included in every staining batch. Endogenous biotin was saturated with a biotin blocking kit (Vector Laboratories).

Table 6: Primary antibodies and conditions used in this study

Antibody	Clone	Vendor	Dilution	Staining location	Threshold value	Stain type
Estrogen receptor alpha	6F11	Ventana Medical Systems Inc.	Prediluted	Nucleus	≥1%	Automated
GATA binding protein 3	HG3-31	Santa Cruz Biotechnology	1:100	Nucleus	≥1%	Manual
HER2/neu	CB11	Cell Marque	Prediluted	Membrane	NA	Automated
Ki-67	Mib1	Biomeda Corp.	1:40	Nucleus	≥20%	Automated
p53	DO7	DBS	Prediluted	Nucleus	0%/≥60%*	Automated
Progesterone receptor	1A6	Ventana Medical Systems Inc.	Prediluted	Nucleus	≥1%	Automated

NA, not applicable.
*intense and diffuse positivity in ≥ 60% of BC cells or complete negativity.

Immunohistochemical scoring

For nuclear stainings, such as ER, GATA3, Ki-67, p53, and PR, tissue cores were scored as the percentage of positive tumor nuclei above the background using a computer-aided image analyzer (Eureka Interface System, Menarini, Firenze, Italy). For each sample, the number of BC positive nuclei per total number of BC nuclei were counted at high magnification (400x) and reported as the percentage of positive cells. Moreover, the intensity score was also recorded for GATA3. In this way, a GATA3 histological score was obtained, given by the product of the percentage of positive cell nuclei (0%-100%) and the four-tier intensity score (0, 1+, 2+, 3+). Therefore, the final histological score ranged from 0 to 300 for each core. Except for p53, scoring results were dichotomized into either negative or positive, using the pre-defined threshold values dictated by the 12th St. Gallen International Breast Cancer Conference and reported in Table 1 (28). Differently, p53 was considered to have either a mutated pattern, when it was completely negative (null pattern) or with at least 60% of BC cell nuclei showing intense positivity (missense pattern), or a wild-type pattern when the tumor showed a variable weak-moderate positivity in 1%-59% of cells (72). The membrane staining HER2 was considered either negative or positive according to conventional guidelines previously described (Table 3) (10, 71).

Statistical Analysis

For correlations with categorical variables, the Spearman's test or the Fisher's exact test, when appropriate, were used. Correlations between continuous and categorical variables were tested using either the Student's t-test or the Wilcoxon's signed rank test when appropriate. The survival curves were estimated using the Kaplan–Meier method, and the log-rank test was used to assess statistical significance. The features found to be significant in univariate analysis were assessed for the multivariate analysis using enter logistic regression model, to evaluate which features were independent. For multivariate analysis, we compared the log–log survival curves and the curves predicted by the Cox model with the observed ones according to the Kaplan-Meier method to check graphically the proportional hazards assumption for all variables. The study time endpoint was evaluated starting at total overall and disease-free survival follow-up (28 years) with a progressive 5-years reduction until 5-years follow-up, then 1-year by 1-year time interval. Then, we used Cox proportional hazards modeling and the likelihood ratio to evaluate survival differences between the different groups (backward parametric statistical Wald method). A setup procedure was used and variables were added to the model if the two-sided significance level was <0.1 in univariate analysis. To control for potential confounding factors, we adjusted HR estimates per age. To evaluate the effect of single variables on patient outcome, the endpoint for disease-free survival analysis was defined as the time from the surgery date to the occurrence of the first adverse event (e.g. local relapse, distant metastasis, contralateral BC, a second tumor and death without evidence of neoplastic disease), while the endpoint for overall survival was considered any death irrespective of cause. Patients without an adverse event were censored at the time of the last follow-up. Multivariate analyses were performed using Cox regression model. Hazard ratios (HRs) with 95% confidence intervals (CIs) were used to quantify the prognostic impact of variables.

A p-value < 0.05 was considered as statistically significant and all tests were two-sided. All statistical analyses were conducted using Stata version 13.0 (StataCorp, College Station, TX, USA) and GraphPad Prism 5 software (GraphPad Software, Inc. CA, USA) was used to plot and to compare data.

RESULTS

Patient clinico-pathological characteristics

Clinico-pathological and immunohistochemical characteristics of the 702 patients included in this study are summarized in Table 7. The median age of the patients at diagnosis was 61 years (range 30-91 years).

Table 7. Clinico-pathological characteristics of our series of breast cancer patients.

Clinico-pathological characteristics	n (%)
Age, total	702
<50 years	166 (23.6)
50-55 years	96 (13.7)
56-70 years	265 (37.7)
> 70 years	175 (24.9)
Grade, total	700
1	135 (19.2)
2	427 (60.8)
3	138 (19.7)
Histotype, total	702
No special type	527 (75.1)
Lobular	109 (15.5)
Other special types	66 (9.4)
pT, total	699
T1	450 (64.4)
T2	236 (33.8)
T3	13 (1.8)
pN, total	702
N0	393 (56.0)
N1	184 (26.2)
N2	72 (10.3)
N3	53 (7.5)
Disease-free survival, failures	
5-year	156 (22.2)
10-year	227 (32.3)
15-year	254 (36.2)
20-year	279 (39.7)
28-year	293 (41.7)
Overall survival, failures	
5-year	119 (17.7)
10-year	237 (33.8)
15-year	291 (41.5)
20-year	375 (53.4)
28-year	432 (61.5)

n, number of cases; total, number of cases for which the data were available; pT, pathological primary tumor status; pN, pathological regional lymph node status.

Among a total of 702 patients included in this study, 513 (73.1%) patients underwent modified radical or radical mastectomy, whereas the remaining 189 (26.9%) underwent partial resection (lumpectomy, segmentectomy or quadrantectomy). Histologically, 527 (75.1%) carcinomas were no special type, 109 (15.5%) lobular, and 66 (9.4%) were other special types, including 24 tubular, 18 mucinous, 7 papillary, 7 medullary, 6 cribriform, 3 apocrine, and 1 micropapillary (Figure 8A).

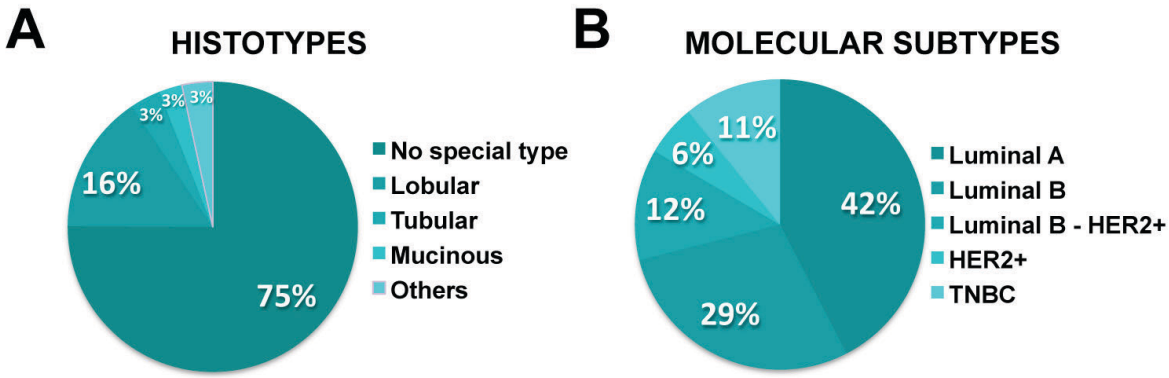

Figure 8. Pie charts representing the distribution by histotypes (**A**) and molecular subtypes (**B**) of our series of breast carcinomas.

A total of 424 (60.4%) patients were treated with adjuvant therapy, including 241 (34.3%) patients treated with only tamoxifen-based endocrine therapy, 134 (18.4%) with 6 cycles of cyclophosphamide, methotrexate, and 5-fluorouracil, while 30 (4.3%) with combined chemotherapy and endocrine therapy. Moreover, 186 (26.5%) patients received loco-regional radiotherapy. 154 (21.9%) patients did not receive adjuvant therapy. Overall, 124 (17.7%) patients had incomplete information regarding their medical treatment.

During a median follow-up of 137 months (range 1-336 months), out of 643 patients with available information, 293 (45.6%) experienced relapse, of which 76 (26.0%) developed local recurrence, 169 (57.7%) distant metastasis, 23 (7.8%) contralateral BC, and 25 (8.5%) had died of BC. The 5-year, 10-year, 15-year, 20-year, and 28-year disease-free survival rates for the entire cohort of patients were 77.8%, 67.7%, 63.8%, 60.3%, and 58.3%, respectively, with a median disease-free survival period of 135 months. The overall survival rates at 5-years, 10-years, 15-years, 20-years, and 28-years for the 702 patients were 83.0%, 66.2%, 58.5%, 46.6%, and 38.5% respectively, with a median overall survival of 182 months.

Tissue microarray performance

A total of 31 TMA blocks were built for this study. Overall, of the 702 BC cores arranged in the TMA, an average of 655 (93.3%) cores per antibody were scorable, whereas on average 47 (7.7%) cores per antibody were unscorable, due to tissue loss, unrepresentative tissue, excessive tissue folding, or improper staining per IHC staining (Table 8).

Table 8. immunoistochemical markers of breast cancer patients.

Tissue microarray performance (total=702)	**Scorable core total** (%)
ER	665 (94.7)
PR	663 (94.4)
Ki-67	658 (93.7)
HER2	676 (96.3)
p53	660 (94.0)
GATA3	608 (86.6)
Average	655 (93.3)

Altogether, a total of 3930 TMA cores were suitable for IHC evaluation in this study.

Patient biological factors and molecular subtypes

The summary of the results of the 6 analyzed IHC markers is reported in Table 9 and Figure 9. Overall, ER and PR were expressed in a higher percentage of BC cells compared with Ki-67 and p53 (85%, 57% vs 7%, 13%, respectively). Moreover, the majority of BC cases were positive (≥1%) for ER (81.5%) and PR (75.3%), negative for HER2 (82.1%), showed a low proliferation index measured with Ki-67 (76.0%) and a p53 wild-type IHC pattern (76.2%).

Table 9. Immunohistochemical results of the 6 analyzed IHC markers.

Immunohistochemical results	n (%)
ER median, range	85.2%, 0-98.9%
ER, total Negative (<1%) Positive (≥1%)	665 123 (18.5) 542 (81.5)
PR median, range	57.4%, 0-98.9%
PR, total Negative (<1%) Positive (≥1%)	663 164 (24.7) 499 (75.3)
Ki-67 median, range	7.0%, 0-98.2%
Ki-67, total Negative (<20%) Positive (≥20%)	658 500 (76.0) 158 (24.0)
HER2, total Negative (0-2+) Positive (3+)	676 555 (82.1) 121 (17.9)
p53 median, range	13.0%, 0-98.5%
p53, total Wild-type pattern Mutated pattern	660 503 (76.2) 157 (23.8)
GATA3 score median, range	60, 0-300
GATA3, total Negative (<1%) Positive (≥1%)	608 195 (32.1) 413 (67.9)

n, number of cases; pT, ER, estrogen receptor; PR, progesterone receptor.

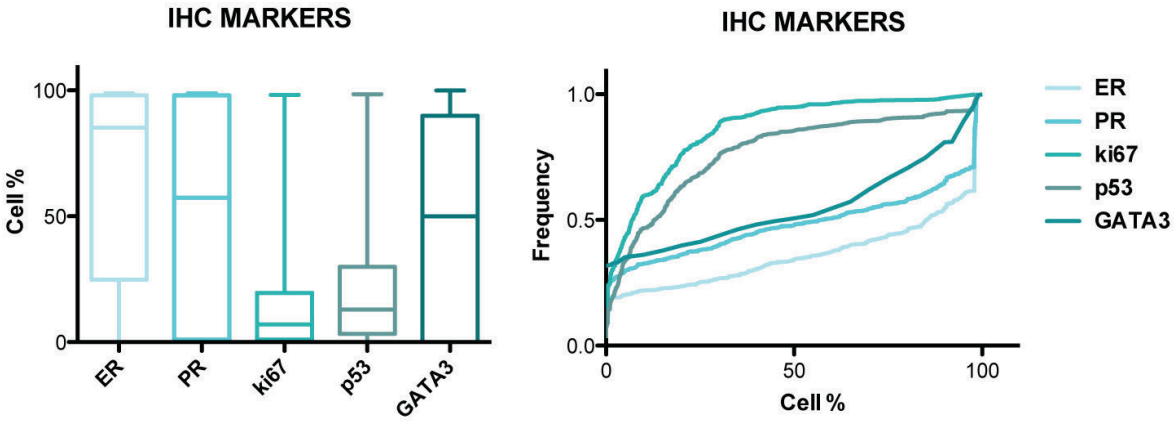

Figure 9 Box plot (left) and cumulative relative frequency charts (right) showing the percentage distribution of immunohistochemical markers estrogen receptors (ER), progesterone receptor (PR), Ki-67 proliferation index, p53 oncosuppressor gene, and GATA3.

Regarding the molecular subtypes, 646 (92.0%) of 702 could be classified based on immunohistochemistry, where all 4 determinant biological markers ER, PR, HER2, and Ki-67 were scored (Table 5). Based on the IHC results, out of 646 BC cases 274 (42.4%) were classified as luminal A, 185 (28.6%) luminal B (HER2-), 80 (12.4%) luminal B-HER2+, 37 (5.7%) HER2-positive, and 70 (10.8%) triple-negative (Figure 8B).

GATA3 expression

Overall, GATA3 was evaluable in 608 (87%) of 702 BC cases and was positive (≥1%) in 413 (68%) cases and negative (<1%) in 195 (32%) cases, with a GATA3 median percentage of 50% (range 0%-100%) and a median histological score of 60 (Figure 10). The remaining 94 (13.4%) TMA cores were not evaluable for GATA3, due to tissue loss, unrepresentative tissue or improper staining.

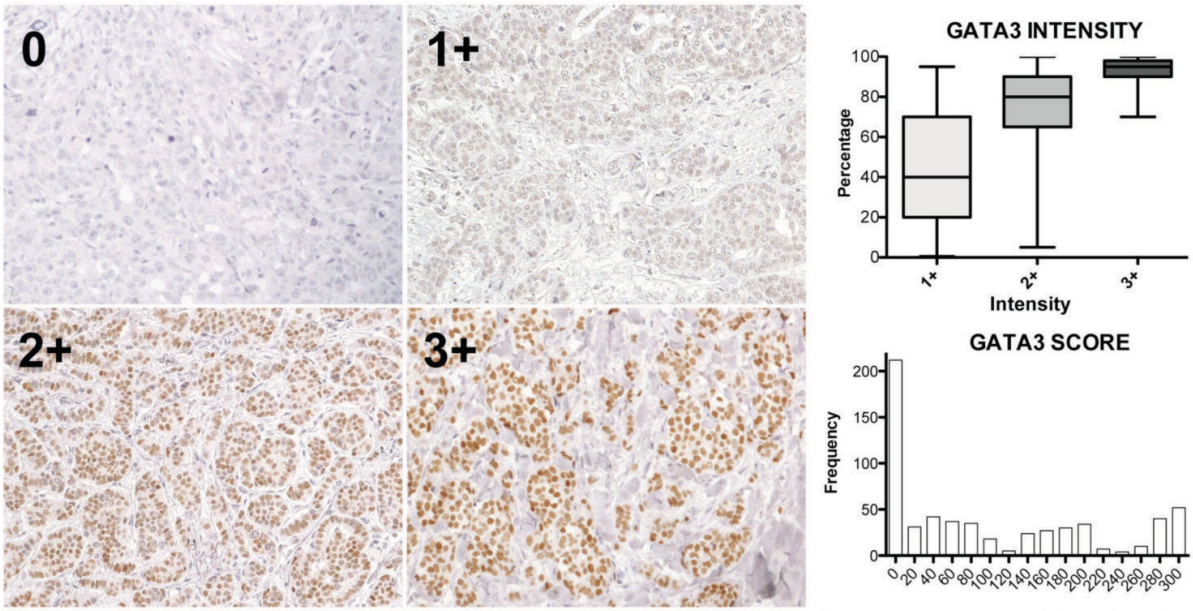

Figure 10. Representative pictures of immunostaining intensity scores of GATA3 in BC tissues (left panel), box-plot diagrams illustrating the results of GATA3 staining of 608 BC samples as positive cell percentage according to intensity scores. The differences between percent positive cells based on intensity were significant (p-value<0.0001, Mann-Whitney U test, two-sided, upper right panel). Histogram of frequency of GATA3 histological score (lower right panel).

Association between GATA3 expression and clinico-pathological features

GATA3 histological score significantly decreased with histological grade (p<0.0001), pT staging (p=0.0463), and stage grouping (p=0.0049) (Figure 11 and Table 10). Therefore, GATA3 expression was higher in BCs with less aggressive clinico-pathological characteristics, such as grade 1 and grade 2, smaller tumor size, and lower stage, than in BC with worse prognosis characteristics. In contrast, GATA3 expression did not significantly differ by age, BC histotype and pN staging.

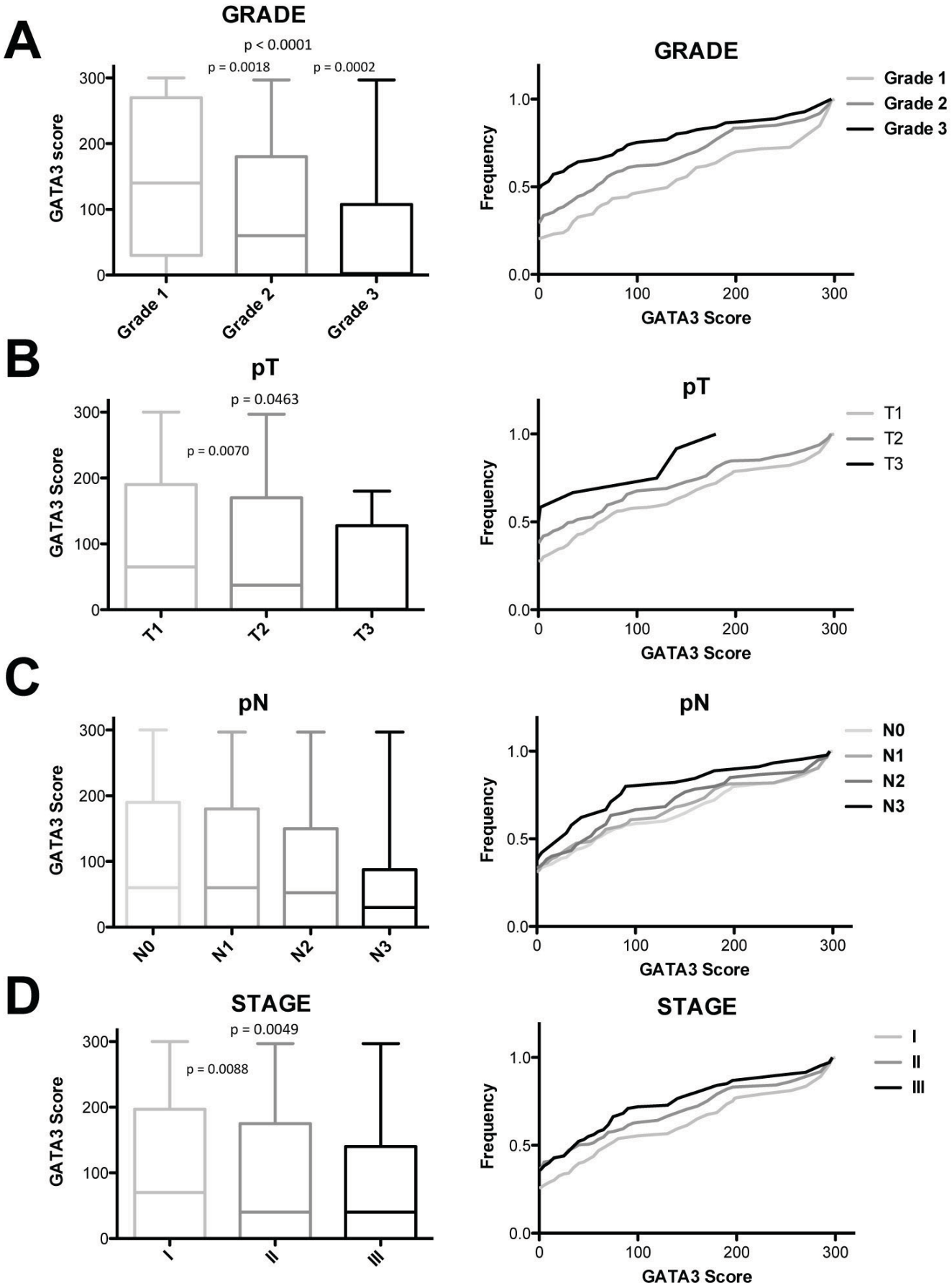

Figure 11. Box plot (left) and cumulative relative frequency charts (right) showing the correlation between GATA3 score and histological grade (**A**), pathological primary tumor size (pT, **B**), pathological lymph node involvement (pN, **C**), and stage grouping (**D**). Significant p-values from two-sided t-test comparisons are reported.

Table 10. Association between GATA3 and clinico-pathological characteristics and biological markers of breast cancer patients.

Clinico-pathological characteristics	n (%)	GATA3 **negative** (<1%)	GATA3 **positive** (≥1%)	p-value
Age, total	608	195	413	
<50 ys	145 (23.8)	40 (27.6)	105 (72.4)	0.3695#
50-55 ys	74 (12.2)	23 (31.1)	51(68.9)	
> 70 ys	389 (64.0)	102 (26.2)	257 (73.8)	
Mean age (range)	61 (30-91)	62 (35-86)	60 (30-91)	0.0559*
Grade, total	607	195	412	
1	113 (18.6)	23 (20.4)	90 (79.6)	**< 0.0001#**
2	368 (60.6)	110 (29.9)	258 (70.1)	
3	126 (20.8)	62 (49.2)	64 (50.8)	
Histologic Type, total	608	195	413	
No special type	463 (76.2)	149 (32.2)	314 (67.8)	0.1587#
Lobular	92 (15.1)	24 (26.1)	68 (73.9)	
Other	53 (8.7)	22 (41.5)	31 (58.5)	
pT, total	605	193	412	
T1	383 (64.4)	106 (27.7)	277 (72.3)	**0.0098#**
T2	210 (33.8)	81 (38.6)	129 (62.4)	
T3	12 (1.8)	6 (50.0)	6 (50.0)	
pN, total	608	195	413	
N0	337 (56.0)	106 (31.5)	231 (68.5)	0.8452#
N1	166 (26.2)	52 (31.3)	114 (68.7)	
N2	60 (10.3)	20 (33.3)	40 (66.7)	
N3	45 (7.5)	17 (37.8)	28 (62.2)	
Stage grouping, total	605	193	412	
I	249 (41.2)	64 (25.7)	185 (74.3)	**0.0233#**
II	249 (41.2)	91 (36.5)	158 (63.5)	
III	107 (17.6)	38 (35.5)	69 (64.5)	
ER, total	591	191	400	
Negative (<1%)	109 (18.5)	83 (76.2)	26 (23.8)	**< 0.0001^**
Positive (≥1%)	482 (81.5)	108 (22.4)	374 (77.6)	
ER mean (range)	63.8% (0-98.9%)	39.6% (0-98.5%)	75.4% (0-98.9%)	**< 0.0001***
PR, total	588	192	396	
Negative (<1%)	143 (24.3)	83 (58.0)	60 (42.0)	**< 0.0001^**
Positive (≥1%)	445 (75.7)	109 (24.5)	336 (75.5)	
PR mean (range)	52.2% (0-98.9%)	34.6% (0-98.8%)	60.8% (0-98.9%)	**< 0.0001***
Ki-67, total	587	192	395	
Negative (<20%)	443 (75.5)	140 (31.6)	303 (68.4)	0.3575^
Positive (≥20%)	144 (24.5)	52 (36.1)	92 (63.9)	
Ki-67 mean (range)	14.3% (0-98.2%)	15.8% (0-98.2%)	13.5%(0-98.0%)	0.4726*
HER2, total	596	193	403	
Negative (0-2+)	487 (81.7)	147 (30.2)	340 (69.8)	**0.0175^**
Positive (3+)	109 (18.3)	46 (42.2)	63 (57.8)	
p53, total	586	185	401	
Wild-type pattern	452 (77.1)	109 (24.1)	343 (75.9)	**< 0.0001^**
Mutated pattern	134 (22.9)	76 (56.7)	58 (43.3)	
p53 mean (range)	23.3% (0-98.5%)	28.1% (0-98.5%)	21.0% (0-98.5%)	0.3928*

n, number of cases; ER, estrogen receptor; PR, progesterone receptor.
*Mann-Whitney test; ^Fisher's test; #Chi-square test.

Association between GATA3 expression and biological factors

GATA3 histological score significantly correlated with biological factors ER and PR (*Spearman* r=0.3616, p<0.0001 and r=0.2138, p<0.0001, respectively). Moreover, GATA3 IHC positivity correlated with ER positivity (p<0.0001), PR positivity (p<0.0001), p53 wild-type pattern (p<0.0001), and HER2 negative status (p=0.0175) (Table 10 and Figures 12 and 13). Then, GATA3 score was significantly higher in BCs with wild-type p53 and negative HER2 (Figure 13). No significant association was found between GATA3 and Ki-67 proliferation index.

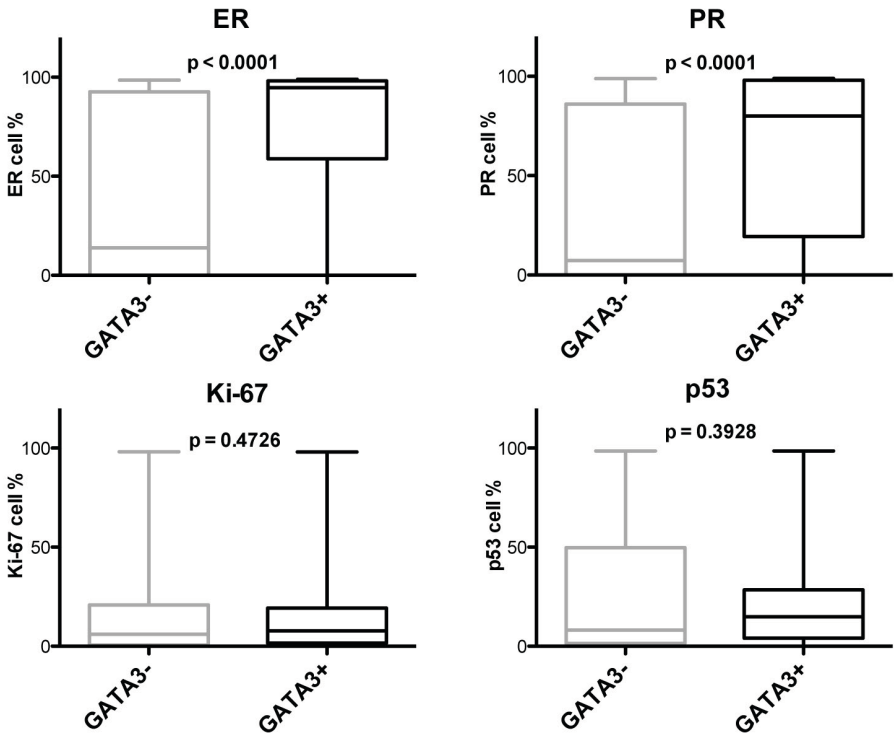

Figure 12. Box-plot charts showing the correlation between GATA3 expression (i.e. - is <1%; + is ≥ 1%) and the biological factors ER, PR, proliferative index Ki-67 and p53. p-values from two-sided t-test comparisons are reported.

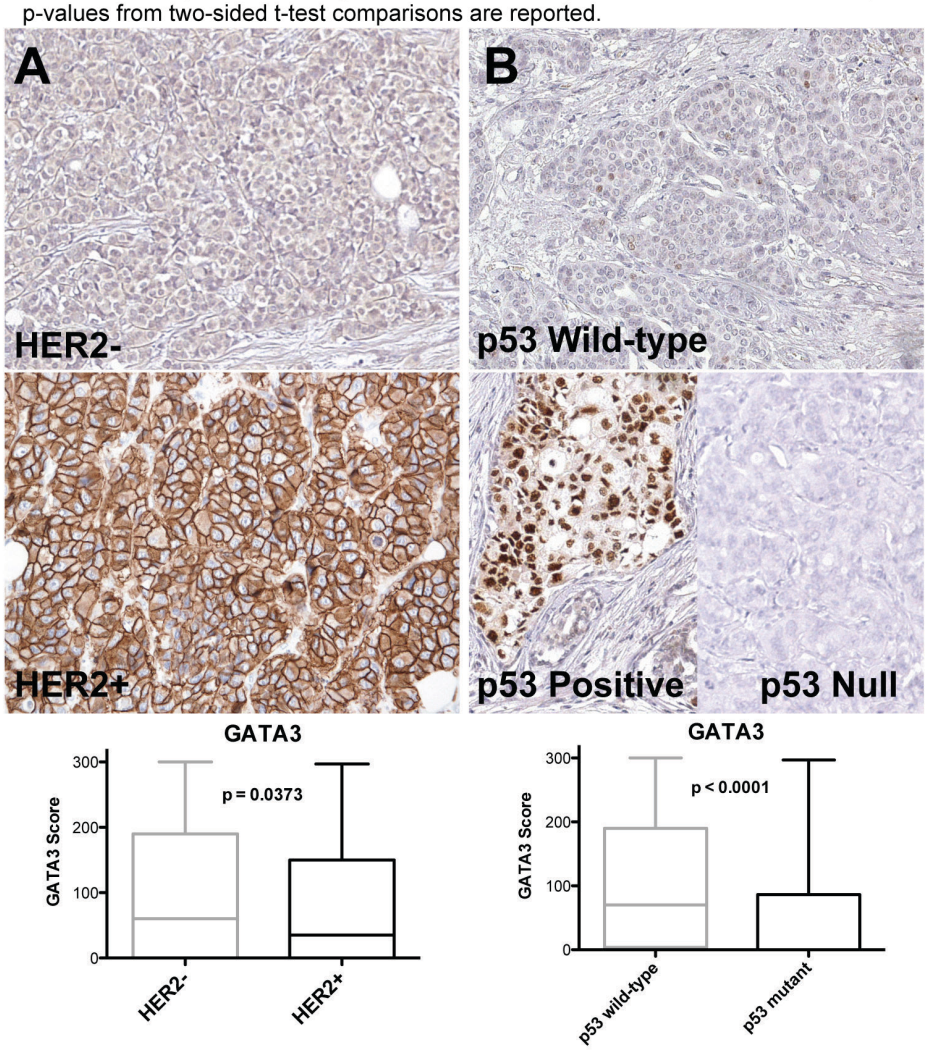

Figure 13. Representative pictures of HER2-negative (-) and HER2-positive (+) BC cases, and box-plot chart of GATA3 histological score according to HER2 status (**A**); representative pictures of p53 immunohistochemical patterns and box-plot chart of GATA3 histological score according to p53 immunohistochemical pattern (**B**). p-values from two-sided t-test comparisons are reported.

Association between GATA3 expression and molecular subtypes

Among the BCs with scored GATA3, 576 cases were subclassified in molecular subtypes according to biological markers by IHC. Overall, GATA3 histological score was significantly correlated with molecular subtypes. Specifically, luminal A, luminal B-HER2-, and luminal B-HER2+ BCs had a significantly higher GATA3 median histological score when compared with HER-enriched and triple-negative BCs (Kruskal-Wallis test, p<0.0001; Table 11 and Figure 14). Moreover, GATA3 positivity correlated with luminal A, luminal B-HER2-, and luminal B-HER2+ subtypes (Chi-square test, p <0.0001).

Table 11. Correlation between GATA3 and molecular subtypes

Molecular subtypes	n (%)	GATA3 score Median (range)	GATA3 negative (<1%), n (%)	GATA3 positive (≥1%), n (%)	p-value
Total	576 (100)	98.2 (0-300)	188 (32.6)	388 (67.4)	
Luminal A	241 (41.8)	123.9 (0-300)	50 (20.7)	191 (79.3)	**< 0.0001^**
Luminal B	167 (29.0)	100.2 (0-297)	45 (26.9)	122 (73.1)	
Luminal B- HER2+	74 (12.9)	107.2 (0-297)	20 (27.0)	54 (73.0)	
HER2-positive	31 (5.4)	18.7 (0-170)	25 (80.6)	6 (20.4)	
Triple-negative	63 (10.9)	23.7 (0-255)	48 (76.2)	15 (23.8)	

^ Chi-square test

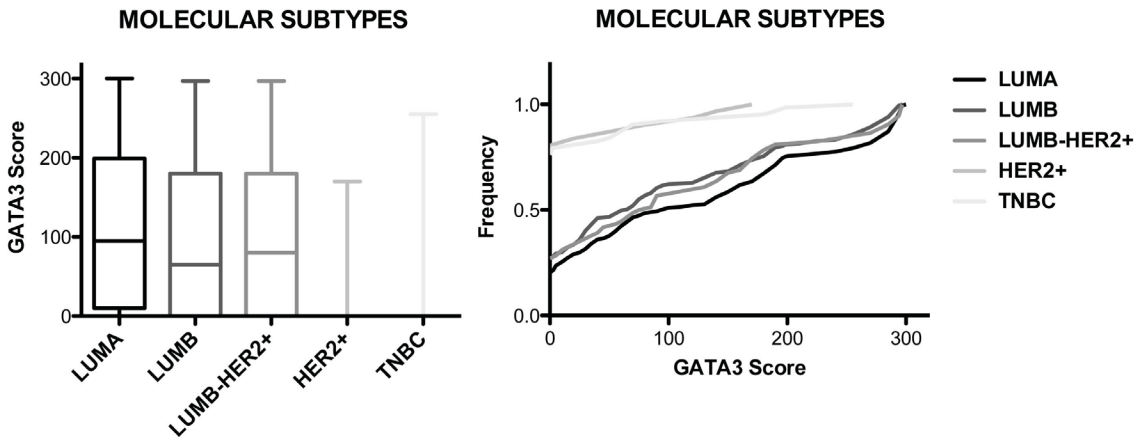

Figure 14. Box-plot and cumulative relative frequency charts show that GATA3 histological score was significantly higher in luminal intrinsic BC subtypes (p<0.0001) when compared to HER2-positive (+) and triple-negative (TNBC) subtypes.

GATA3 expression as a predictor of prognosis

In our patients with a median follow-up of 183 months, GATA3 IHC positivity correlated significantly with a better overall survival (median survival 234 vs.181 months for GATA3+ vs. GATA3-, Figure 15). Specifically, after adjusting for age, the overall hazard ratio for death was 0.70 (95% CI: 0.56 to 0.86, p=0.001) for BC patients with positive GATA3 when compared with negative GATA3 (Figure 15 and Table 12).

In order to explore the clinical relevance of GATA3 positivity, we compared overall survival curves from our BC cohort, categorizing them based on GATA3 expression (Table 13). Interestingly, analysis of grade-, pT- and stage-specific overall survival according to GATA3 IHC showed a significantly better outcome associated with GATA3 positivity mainly among patients with lower grade, and lower stage (Table 13 and Figure 16). Specifically, GATA3 positivity was associated with better prognosis in BC patients with Elston and Ellis grade 1-2 (HR 0.69, p=0.003), pT1-T2 (HR 0.68, p=0.001), pN0 (HR 0.65, p=0.003), stage I-II (HR 0.65, p<0.0001). Moreover, GATA3 positivity maintained the ability to stratify significantly patients with a better prognosis particularly in BC subgroups with positive ER (HR 0.77, p=0.046), positive PR (HR 0.74, p=0.022), Ki-67<20% (HR 0.62, p=0.008), negative HER2 (HR 0.64, p<0.0001) and with p53 wild-type IHC pattern (HR 0.71, p=0.011) (Figure 17). Furthermore, GATA3 expression was associated with a better overall survival of BC patients with luminal B intrinsic subtype (median survival 261 vs. 166 months for GATA3+ vs. GATA3-, HR 0.64, 95% CI: 0.42 to 0.97, p=0.036), but not with the other molecular subtypes (Figure 17). Notably, there was a switch in this trend in patients with positive HER2 (HR 1.06, 95% CI: 0.66 to 1.71, p=0.798) and luminal B-HER2+ intrinsic subtype (HR 1.17, 95% CI: 0.62 to 2.21, p=0.062) where GATA3 positivity was associated with worse OS, but not in a statistical significant way.

The prognostic impact of the other variables analyzed, for example, age, pT stage, pN stage, Stage grouping, was as expected (Table 12). Moreover, a lower risk of failure was associated with ER expression (HR 0.78, CI 95% 0.60 to 1.00, p=0.047), and a higher risk of failure was associated with p53 mutated IHC pattern (HR 1.40, CI 95% 1.12 to 1.75, p =0.004).

Since clinico-pathologic features, including age, pT stage, and pN stage, as well as some biological factors contributed to the outcome of BC patients in the univariate analysis, a multivariate Cox regression analysis was performed to evaluate whether GATA3 positivity can be considered as an independent prognostic factor. For multivariate analysis, we compared the ln–ln survival curves to check graphically the proportional hazards assumption for all clinico-pathological and IHC variables (Figures 18 and 19). The ln-ln survival curves indicated some minor violations of the proportional hazard assumption for pT, pN, stage, and HER2 at the beginning of the follow-up time, that is represented by an initial intersection of the curves.Therefore, we plotted the graphs that compare the survival

Table 12: Kaplan-Meier survival analysis for the clinico-pathological features and biological prognostic factors

Variables	Patients n	Deaths n (%)	KM analysis survival Median, mos	HR (95% CI)	p-value
Age, total	702	432 (61.5)		2.53 (1.95-3.28)	**< 0.0001**
<50 ys	166	65 (39.2)	267		
≥50 ys	536	367 (68.5)	189		
Grade, total	700	431 (61.6)		1.13 (0.89-1.44)	0.311
1-2	562	348 (61.9)	213		
3	138	83 (60.2)	207		
Histologic Type, total	699	430 (61.5)		0.86 (0.74-1.00)	0.053
No special type	527	327 (62.1)	203		
Lobular	109	71 (65.1)	204		
Other	63	32 (50.8)	297		
pT, total	699	431 (61.7)		1.55 (1.30-1.85)	**< 0.0001**
T1	450	251(55.8)	233		
T2	236	171 (72.5)	153		
T3	13	9 (69.2)	66		
pN, total	702	432 (61.5)		0.64 (0.53-0.78)	**< 0.0001**
N0	393	224 (57.0)	242		
N+	309	208 (67.3)	170		
Stage	699	431 (61.7)		1.56 (1.37-1.77)	**< 0.0001**
I	295	161(54.6)	244		
II	277	170 (61.4)	207		
III	127	100 (78.7)	86		
ER, total	665	414 (62.3)		0.78 (0.60-0.99)	**0.047**
Negative (<1%)	123	75 (61.0)	199		
Positive (≥1%)	542	339 (62.6)	213		
PR, total	663	413 (62.3)		0.82 (0.65-1.03)	0.082
Negative (<1%)	164	101 (61.6)	194		
Positive (≥1%)	499	312 (62.5)	214		
Ki-67, total	658	412 (62.6)		1.04 (0.82-1.31)	0.759
Negative (<20%)	500	318 (63.6)	207		
Positive (≥20%)	158	94 (59.5)	229		
HER2, total	676	419 (62.0)		1.27 (0.99-1.63)	0.058
Negative (0-2+)	555	341 (61.4)	215		
Positive (3+)	121	78 (64.5)	195		
p53, total	660	410 (62.1)		1.40 (1.12-1.75)	**0.004**
Wild-type pattern	503	309 (61.4)	217		
Mutated pattern	157	101 (64.3)	190		
GATA3, total	608	379 (62.3)		0.70 (0.56-0.86)	**0.001**
Negative (<1%)	195	136 (69.7)	181		
Positive (≥1%)	413	243 (58.8)	234		

n, number of cases; HR, hazard ratio; CI, confidence interval; ER, estrogen receptor; PR, progesterone receptor.

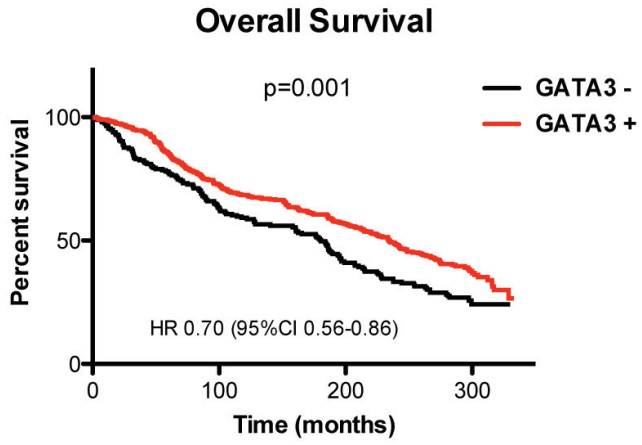

Figure 15. Kaplan–Meier survival curve analysis of GATA3 in breast carcinomas. After adjusting for the patients' age, patients with GATA3 IHC positivity in their tumors had a better overall survival than patients whose tumors were negative GATA3 (p= 0.001, log-rank test). The hazard ratio (HR) for death and the 95% confidence interval (CI) are reported.

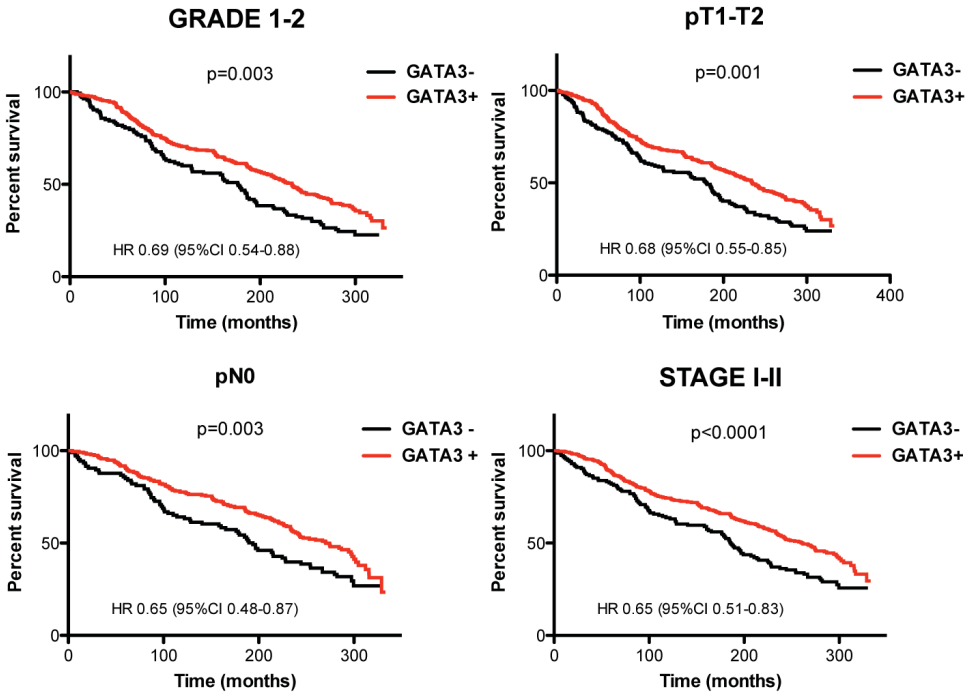

Figure 16. Kaplan–Meier overall survival curves of subgroups of breast cancer patients according to IHC expression of GATA3. After adjusting for the patients' age, GATA3 IHC positivity is associated with a significant better overall survival in breast carcinoma patients with histological grade 1 and 2, pT1-T2, pN0 and Stage I and II. The hazard ratio (HR) for death and the 95% confidence interval (CI) of univariate Cox regression analyses are reported.

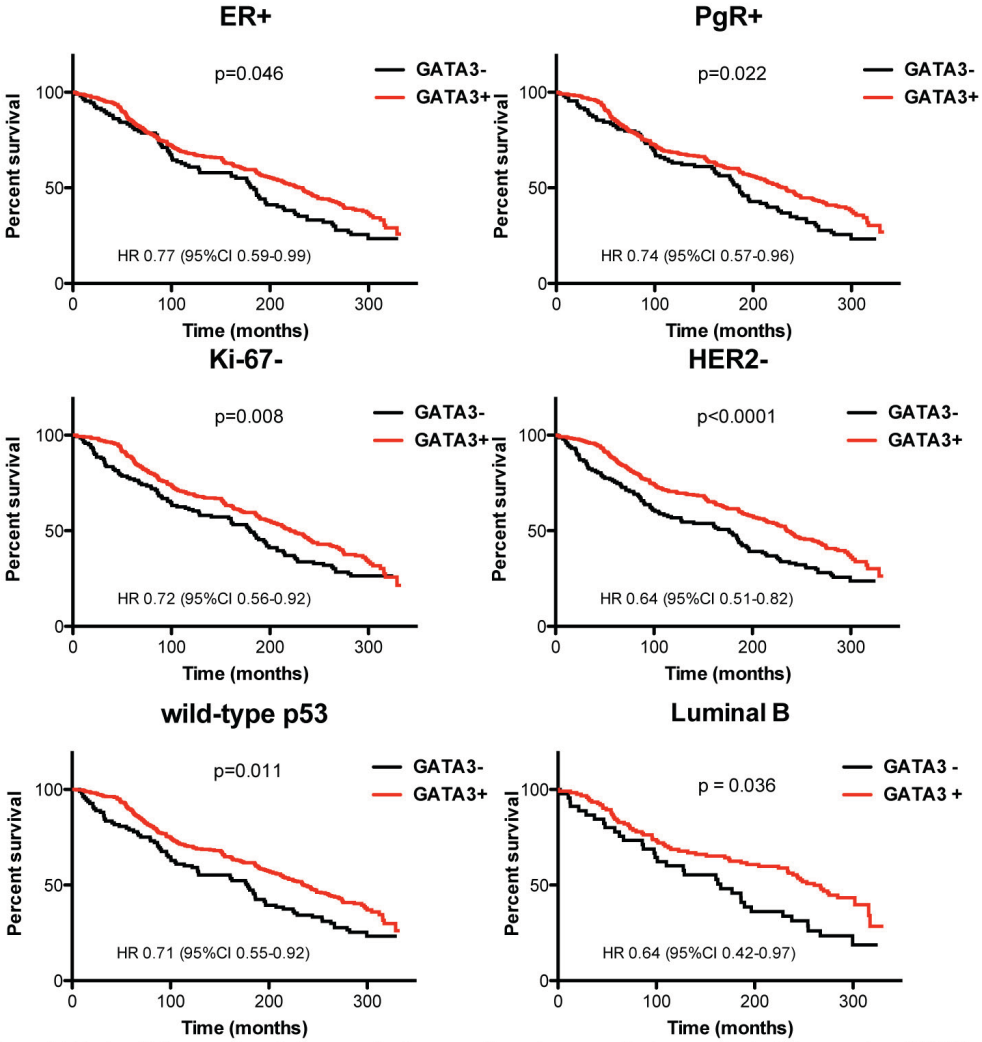

Figure 17. Kaplan–Meier overall survival curves of subgroups of breast cancer patients according to IHC expression of GATA3. After adjusting for the patients' age, GATA3 IHC positivity is associated with a significant better overall survival in breast carcinoma patients with positive (+) estrogen receptors (ER) and progesterone receptor (PR), negative HER2 and Ki-67 (i.e. <20%), and p53 wild-type IHC pattern and Luminal B intrinsic subtype. The hazard ratio (HR) for death and the 95% confidence interval (CI) of univariate Cox regression analyses are reported.

Table 13: Prognostic value of GATA3 positivity according to clinico-pathological features and biological prognostic factors

Variables	Univariate Cox regression analysis		
	Hazard ratio	95% Confidence Interval	p-value
Age			
<50 ys	0.92	0.53-1.60	0.767
≥50 ys	0.66	0.52-0.83	**<0.0001**
Histologic Grade			
1-2	0.69	0.54-0.88	**0.003**
3	0.74	0.47-1.16	0.189
Histologic Type			
No special type	0.71	0.56-0.91	**0.006**
Lobular	0.69	0.39-1.22	0.197
Other	0.35	0.15-0.85	0.019
pT			
T1-T2	0.68	0.55-0.85	**0.001**
T3	1.85	0.41-8.41	0.425
pN			
N0	0.65	0.48-0.87	**0.003**
N+	0.76	0.56-1.03	0.078
Stage			
I-II	0.65	0.51-0.83	**<0.0001**
III	0.96	0.61-1.49	0.840
ER			
Negative (<1%)	0.55	0.30-1.00	0.051
Positive (≥1%)	0.77	0.59-0.99	**0.046**
PR			
Negative (<1%)	0.72	0.47-1.10	0.125
Positive (≥1%)	0.74	0.57-0.96	**0.022**
Ki-67			
Negative (<20%)	0.72	0.56-0.92	**0.008**
Positive (≥20%)	0.67	0.44-1.02	0.065
HER2			
Negative (0-2+)	0.64	0.51-0.82	**<0.0001**
Positive (3+)	1.06	0.66-1.71	0.798
p53			
Wild-type pattern	0.71	0.55-0.92	**0.011**
Mutated pattern	0.93	0.61-1.41	0.725

n, number of cases; ER, estrogen receptor; PR, progesterone receptor.

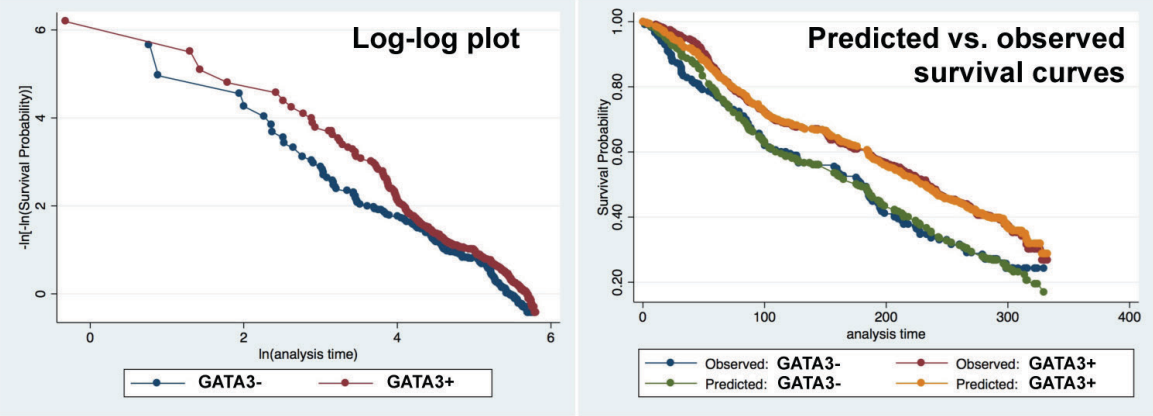

Figure 18. To check the proportional hazards assumption of the Cox-model, we compared the log-log plots (left panel) and the predicted (Cox Proportional Hazard Model) vs. observed (Kaplan-Meier) survival curves (right panel) according to GATA3 expression. Log-log plots are parallel, while predicted vs. observed survival curves are overlapping. Both graphs demonstrate that the assumption has not been violated for GATA3.

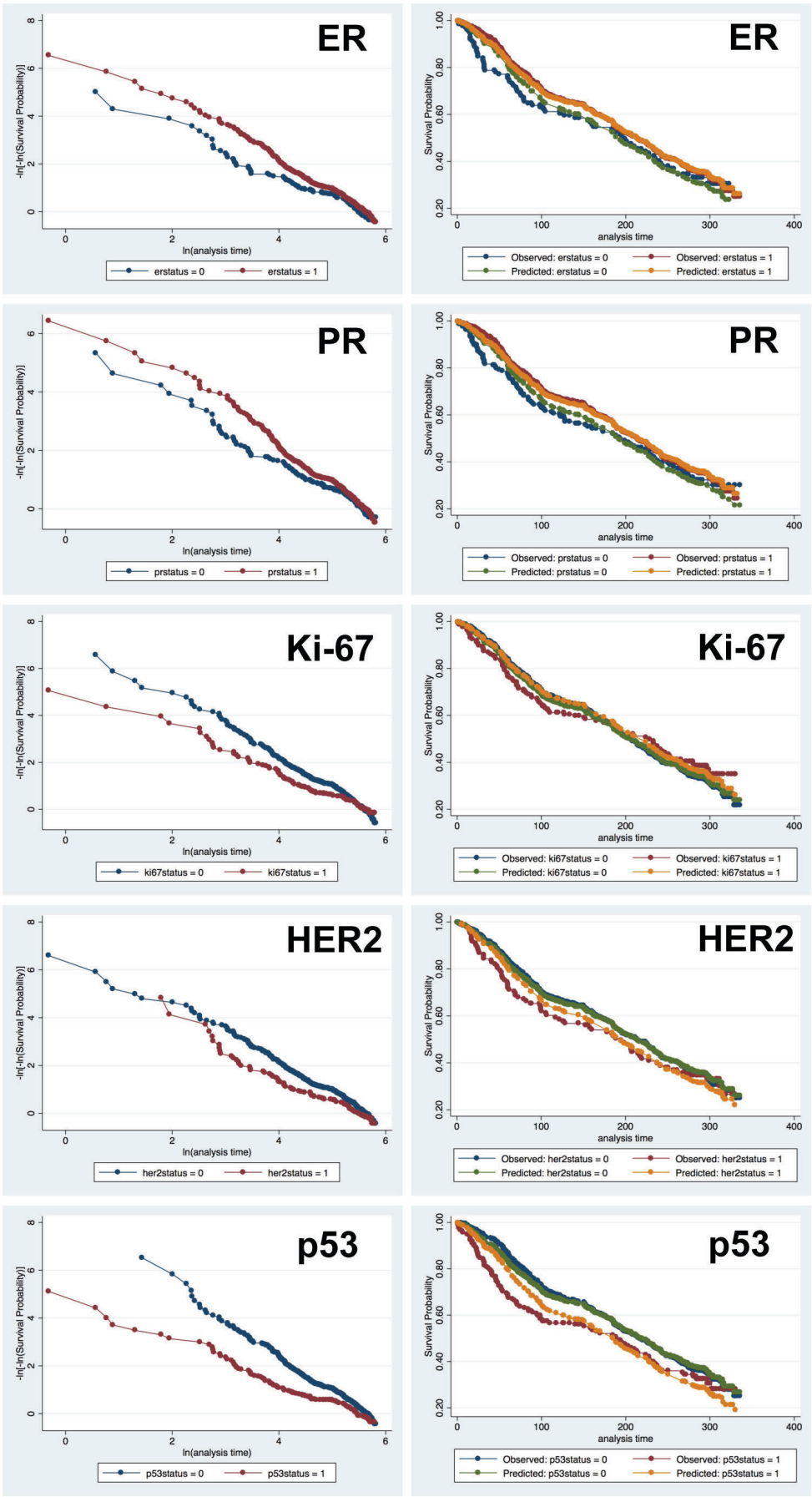

Figure 19. The log-log plots (left panels) and the predicted (Cox Proportional Hazard Model) vs. observed (Kaplan-Meier) survival curves (right panels) show that for all these biological markers the hazard ratio could be considered constant only in the first 48 months

curves predicted by the Cox model with the observed ones according to the Kaplan-Meier method. According to the log-log plot and predicted versus observed survival curves, GATA3 is a time-independent variable, as a consequence, we could apply the Cox proportional hazards modeling (Figure 18). After adjusting for all clinico-pathological and IHC variables, GATA3 functioned as an independent prognostic factor for BC patients (Table 14).

Table 14: Multivariate Cox proportional hazards model for the overall survival

Multivariate Cox regression analysis					
Variables	**Hazard ratio**	**Standard Error**	**Z coefficient**	**p-value**	**95% CI**
Age	1.0578	0.0053	11.1900	0.0000	1.0474-1.0682
GATA3	0.7705	0.0916	-2.1900	0.0280	0.6104-0.9726

CI, confidence interval.

Differently, based on the log-log plots and predicted vs observed survival curves, for grade, ER, PR, HER2, Ki-67, and p53 the Cox model does not predict correctly the survival during the whole time period, but just during the first 48 months after the surgical treatment, the period for which the HR is constant in time (Figure 19). As a consequence, we applied univariate Cox regression analysis in this limited time period and found that all covariates had a significant impact on short-term overall survival (Table 15).

Table 15: Univariate Cox regression analysis for the overall survival at 48 months follow-up

Variables	**Univariate Cox regression analysis**		
	Hazard ratio	**95% Confidence Interval**	**p-value**
Age	1.03	1.01-1.05	0.002
Histologic Grade	3.00	1.95-4.63	<0.0001
pT	2.91	1.89-4.49	<0.0001
pN	2.53	1.62-3.94	<0.0001
Stage	3.63	2.08-6.34	<0.0001
ER	0.40	0.25-0.63	<0.0001
PR	0.43	0.28-0.68	<0.0001
Ki-67	1.66	1.03-2.67	0.037
HER2	1.81	1.11-2.95	0.017
p53	3.57	2.31-5.51	<0.0001
GATA3	0.37	0.24-0.59	<0.0001

ER, estrogen receptor; PR, progesterone receptor.

Afterward, multivariate analysis on this time period showed that GATA3 IHC positivity was an independent favorable predictive factor for improved overall survival in BC patients, reducing the hazard ratio by 53% (Table 16). In addition, p53 mutation IHC pattern was an independent risk factor for reduced overall survival, increasing the hazard ratio by 2.7-fold (Table 16).

Table 16: Multivariate Cox proportional hazards model of overall survival at 48 months

Multivariate Cox regression analysis					
Variables	**Hazard ratio**	**Standard Error**	**Z coefficient**	**p-value**	**95% CI**
Age	1.030964	0.0101604	3.09	0.002	1.011241-1.051072
GATA3	0.5269395	0.1341376	-2.52	0.012	0.3199478-0.8678456
p53	2.725966	0.6961342	3.93	<0.0001	1.652524-4.496692

CI, confidence interval.

In a second time, we analyzed the effects of GATA3 and both all clinico-pathological variables and IHC markers on disease-free survival of our BC patient cohort. Neither IHC markers nor molecular classification were significantly associated with disease-free survival, based on the Kaplan-Meier analysis during the entire period. On the other hand, the following clinico-pathological features were found to be associated with it: histologic grade (HR 1.27, CI 95% 1.06-1.52, p<0.0001), tumor size (HR 1.67, CI 95% 1.36-2.05, p<0.0001), lymph node status (HR 1.89, CI 95% 1.50-2.38, p<0.0001), and staging (HR 1.74, CI 95% 1.49-2.04, p<0.0001).

However, GATA3 positivity correlated with a significantly better disease-free survival at 48 months (HR 0.63, CI 95% 0.44-0.93, p=0.001), as well as the other prognostic factors that we have analyzed using the Kaplan-Meier survival analysis (Table 17 and Figure 20).

Table 17: Kaplan-Meier disease-free survival analysis for the clinico-pathological features and biological prognostic factors at 48 months

Variables	**KM analysis survival**	
	HR (95% CI)	**p-value**
Age, total <50 ys ≥50 ys	0.93 (0.63-1.38)	0.723
Grade, total 1-2 3	2.16 (1.63-2.86)	**< 0.0001**
Histologic Type, total No special type Lobular Other	0.64 (0.46-0.90)	**0.010**
pT, total T1 T2 T3	1.96 (1.47-2.62)	**< 0.0001**
pN, total N0 N+	2.34 (1.65-3.32)	**< 0.0001**
Stage I II III	1.94 (1.56-2.43)	**< 0.0001**
ER, total Negative (<1%) Positive (≥1%)	0.51 (0.34-0.75)	**0.001**
PR, total Negative (<1%) Positive (≥1%)	0.60 (0.41-0.87)	**0.007**
Ki-67, total Negative (<20%) Positive (≥20%)	1.58 (1.08-2.31)	**0.018**
HER2, total Negative (0-2+) Positive (3+)	1.67 (1.12-2.49)	**0.012**
p53, total Wild-type pattern Mutated pattern	2.72 (1.90-3.89)	**< 0.0001**
GATA3, total Negative (<1%) Positive (≥1%)	0.63 (0.44-0.93)	**0.001**

CI, confidence interval; ER, estrogen receptor; PR, progesterone receptor.

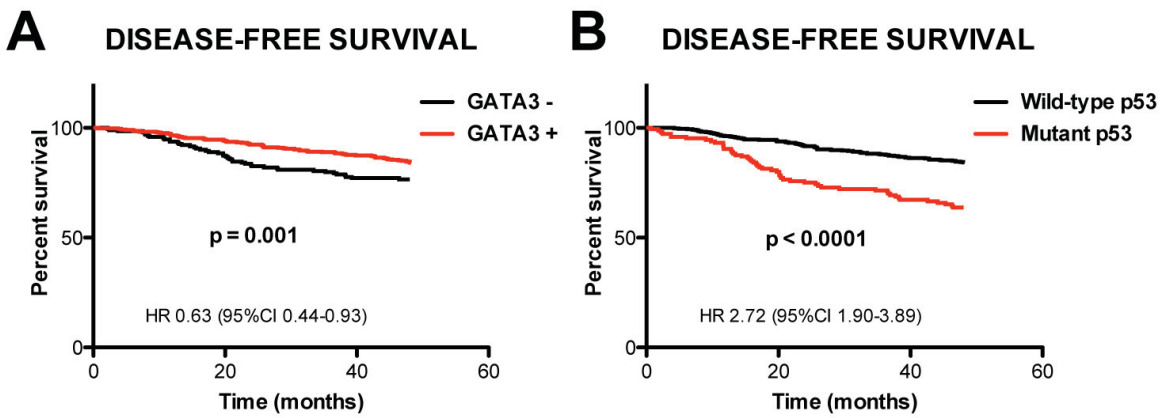

Figure 20. Kaplan–Meier disease-free survival curves of breast cancer patients according to IHC expression of GATA3 and p53. Both GATA3 IHC positivity (A) and p53 mutant IHC pattern (B) are significantly associated with a better disease-free survival in breast carcinoma patients. Hazard ratio (HR) for disease relapse, 95% confidence interval (CI) and p-value of Kaplan-Meier survival analyses are reported.

Table 18: Prognostic value of GATA3 positivity on disease-free survival at 48 months, according to clinico-pathological features and biological prognostic factors

Variables	Univariate Cox regression analysis		
	Hazard ratio	95% Confidence Interval	p-value
Age			
<50 ys	1.35	0.55-3.35	0.513
≥50 ys	0.51	0.34-0.80	**0.003**
Histologic Grade			
1-2	0.62	0.38-1.01	0.055
3	0.97	0.51-1.80	0.910
Histologic Type			
No special type	0.64	0.43-0.96	**0.032**
Lobular	0.69	0.21-2.30	0.548
Other	NA	NA	NA
pT			
T1	0.78	0.44-1.39	0.398
T2	0.62	0.36-1.06	0.081
T3	0.51	0.09-3.07	0.463
pN			
N0	0.75	0.40-1.42	0.377
N+	0.57	0.36-0.92	**0.021**
Stage			
I	1.07	0.46-2.50	0.880
II	0.61	0.36-1.06	0.081
III	0.51	0.85-3.07	0.463
ER			
Negative (<1%)	0.45	0.16-1.29	0.136
Positive (≥1%)	0.87	0.51-1.49	0.618
PR			
Negative (<1%)	0.44	0.21-0.94	**0.034**
Positive (≥1%)	0.89	0.52-1.52	0.672
Ki-67			
Negative (<20%)	0.61	0.38-0.97	**0.038**
Positive (≥20%)	0.76	0.37-1.53	0.436
HER2			
Negative (0-2+)	0.56	0.36-0.87	**0.010**
Positive (3+)	1.42	0.63-3.18	0.399
p53			
Wild-type pattern	0.61	0.35-1.03	0.065
Mutated pattern	1.47	0.81-2.66	0.202

ER, estrogen receptor; PR, progesterone receptor; NA, not available.

In this case, the benefit on disease-free survival of GATA3 positivity was not maintained in most BC subgroups according to the other variables. In particular, GATA3 preserved a prognostication value only on the following BC subgroups: no special type, lymph node metastatic, PR-negative, HER2-negative, and low Ki-67 (<20%) (Table 18).

We evaluated the ln–ln survival curves to check graphically the proportional hazards assumption for all clinico-pathological and IHC variables at 48 months. The log-log plots indicated some minor violations of the proportional hazard assumption for most of the variables (not shown). Then, we plotted the graphs that compare the survival curves predicted by the Cox model with the observed ones according to the Kaplan-Meier method. According to the log-log plot and predicted versus observed survival curves, GATA3 can be considered time-independent at 48 months of disease-free survival, as such, we could apply the Cox proportional hazards modeling (Figure 21).

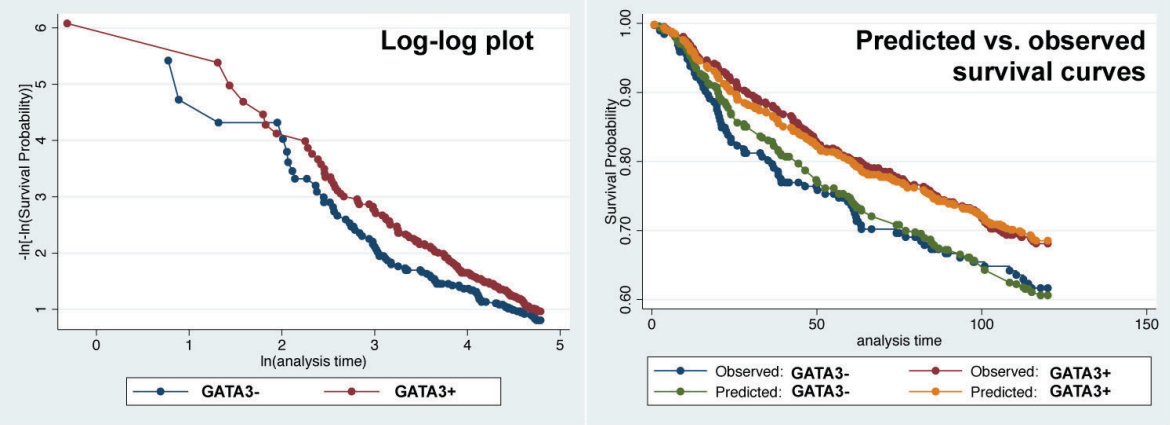

Figure 21. To check the proportional hazards assumption of the Cox-model, we compared the log-log plots and the predicted (Cox Proportional Hazard Model) vs. observed (Kaplan-Meier) survival curves according to GATA3 expression. Log-log plots show a minor violation of the proportional hazards assumption (left panel), while predicted vs. observed survival curves are closed (right panel). Both graphs demonstrate that the assumption has not been violated for GATA3 for disease-free survival at 48 months.

After adjusting for all clinico-pathological and IHC variables, GATA3 does not result as an independent prognostic factor for disease-free survival in our BC patients (Table 19). The only significant predictive variable entering into the multivariate model was p53 (HR 2.43, CI 95% 1.66-3.04, p<0.0001).

Table 19: Multivariate Cox proportional hazards model of disease-free survival at 48 months

Multivariate Cox regression analysis					
Variables	**Hazard ratio**	**Standard Error**	**Z coefficient**	**p-value**	**95% CI**
p53	2.427183	0.4676534	4.6	<0.0001	1.663791-3.54084
Ki-67	1.329254	0.2681619	1.41	0.158	0.8951331-1.973914

CI, confidence interval.

DISCUSSION

Physiologically, GATA3 is a transcriptional activator determinant for the specification and maturation of the breast luminal cells (56). Accordingly, in normal mammary tissue GATA3 is expressed exclusively in lobular and ductal epithelial cells, but not in myoepithelial cells (56, 73). In murine BC models, GATA3 inhibits the transition from an epithelial to mesenchymal phenotype and the development of metastases (54, 74). GATA3 is an ER-related gene and, as such, has been associated with improved outcome in BC patients (67, 68, 75). However, conclusive data demonstrating GATA3 as a prognostic factor independent from ER status are still missing (69).

By using TMA technology, in this study, we have analyzed GATA3 IHC in a consecutive monoinstitutional large cohort of 702 BCs in order to assess its association with clinico-pathological and biological characteristics and to define its role as an independent prognostic factor. In our study, we found that 68% (413/608) of BCs expressed GATA3 by IHC. Previous studies reported that GATA3 IHC expression ranged between 31% and 92% in BCs (69, 76, 77). The apparent discrepancy in GATA3 expression between these studies may reflect both the methodology used and the different cutoff values assigned to define GATA3 IHC positivity. In fact, whether the tumor tissue tested is a whole section or TMA cores may have an impact on the results, and in the latter case, also the number of cores per tumor and the core size will modulate the results (78). Other technical issues may contribute to this difference such as the use of different GATA3 antibodies and their sensitivity (79). Interestingly, it has been shown that the sensitivity of GATA3 different antibodies varies between BC subgroups. Furthermore, the threshold definition for GATA3 IHC positivity is critical for all subsequent quantification and varied arbitrarily between previous studies, ranging from 1% to 30%. In our study, we used relatively large cores and the positive GATA3 IHC cutoff was set at 1% (to parallel cutoff percentages of ER and PR) obtaining high positive rate (28). Finally, the assessment of positive IHC can be subjective, and operator-dependent, to limit this variability and to increase score accuracy we applied an automated scoring method (80).

Between BC histotypes, lobular BCs demonstrated the slightly highest GATA3 positive rate (74%), whereas NST BCs were GATA3 IHC positive in 68% of cases. Coherently, previous studies reported the highest frequency of GATA3 IHC positivity in lobular BC when compared to NST BC (73, 77, 81). Moreover, we found that GATA3 positivity is strongly associated ($p<0.0001$) with BC low-intermediated histologic grades, similarly to previous studies, while GATA3 expression is reduced in high-grade BCs (76, 82, 83).

In addition to the morphological characterization, in 2001 a molecular classification of BC has been proposed and progressively has taken over during the past decade (25-27). Currently, five intrinsic molecular subtypes of BC are recognized: luminal A, luminal B-

HER2-, luminal B-HER2+, HER2-positive and triple-negative (84). Surrogates based on IHC have been proposed with a discussable accuracy and partial overlapping to molecular subtypes according to gene expression (26-28, 85).

In our series, GATA3 positivity significantly varied between different molecular subtypes. Luminal A and luminal B showed the highest GATA3 levels and positivity percentages (79% and 73%, respectively). However, GATA3 was also expressed in a relevant percentage of the other molecular subtypes, namely in 20% of HER2-positive BCs and 24% of TNBC. In four previous studies GATA3 IHC positivity in luminal A, luminal B, and HER2-positive subtypes have been reported to range from 54-100%, 27-100%, and 8%-86%, respectively (76, 83, 86, 87). Many more studies have specifically investigated the expression of GATA3 in TNBC for the relevance that it might have as an ancillary diagnostic test. They reported the following rates of GATA3 positivity in TNBCs: 6% (83), 15% (76), 20% (87), 25% (88), 35% (82), 40% (89), 44% (79), 46% (86, 90), 48% (91), 61% (92), and 66% (79), depending on the antibody clone used and cutoff settled. Therefore, together our GATA3 results are roughly in line with literature reports, whilst in the lower ranges, possibly as a consequence of the use of only one TMA core per tumor, that could cause an underestimate of the IHC staining, particularly in case the marker has a heterogeneous expression (78).

Again, we found an excellent correlation ($p<0.0001$) between hormone receptor positivity and GATA3 expression. In fact, in our study, 78% of ER-positive and 76% of PR-positive BCs were also GATA3-positive and, vice versa, among GATA3-positive cases 93% were also ER-positive and 85% were PR-positive. These results are consistent with all previous studies that identified this association in BC tissues (67, 93, 94) and coherent with the integral role of GATA3 into ER signaling pathway revealed by both cell line studies and microarray data analyses (26, 94, 95). As a matter of fact, it has been demonstrated a positive cross-regulated feedback loop between GATA3 and ER, which, through a mutual stimulation of expression between these two genes, is important for maintenance of ER signaling pathway and, likely, sustains this robust coexpression of GATA3 and hormone receptors (57, 58).

Moreover, we found a significant association between GATA3 positivity and HER2-negativity. Notably, this association is implicit in the GATA3 distribution between BC molecular subtypes and also considering the reciprocal associations with hormone expression. Nevertheless, our results are consistent with the majority of the previous studies that have investigated this relationship, but not all (69, 76, 82, 93, 96-99). To the best of our knowledge, yet no direct relationship at a molecular level between GATA3 and HER2 has been identified.

We have found no association between Ki-67 labeling index and GATA3-expression as

well as other investigators (98, 99). Conversely, two previous studies found a significantly lower Ki-67 labeling index in GATA3-positive BCs, but both have analyzed only selected BC subtypes, i.e. hormone receptor-positive/HER2-negative and TNBC, respectively (88, 100).

In addition to usual biological prognostic factors, we evaluated p53 IHC. *TP53* is the most frequently mutated gene in human cancers and, as such, it has been extensively studied (101, 102). *TP53* acts as a tumor suppressor gene, which may be disrupted by either missense mutation, that is a point mutation that, changing a single nucleotide, results in a changed codon that specifies a different amino acid, or null mutation, that is any other mutation causing a truncated protein, usually nonfunctional. Also, in BC *TP53* is the most commonly mutated gene, given that *TP53* mutations occur in 30% of BCs (31, 33). Importantly, *TP53* mutation has been strongly associated with worse outcome in BC patients (103, 104).

First, we found no association between GATA3 expression and p53 by IHC percentage, similarly to Hosoda *et al.* and Jacquemier *et al.* (98, 99). As opposite, two different studies have identified higher p53 expression in GATA3-negative BCs (96, 97). Relevantly, it has been shown that diffuse, intense IHC positivity for p53 associates with a missense mutation, while complete p53 immunonegativity is consistent with null mutation (due to frameshift, splicing junction, and nonsense mutation) (67, 99). By contrast, rare, patchy, and weak p53 staining corresponds to wild-type *TP53* (67, 99). Therefore, differently from previous studies, besides considering p53 as a continuous variable, we categorized p53 immunostaining in patterns associated with mutated and wild-type *TP53* status (72, 105, 106). In this way, we found a positive association between GATA3 positivity and p53 wild-type pattern. This novel finding is reasonably expected since *GATA3* and *TP53* mutations have been shown to be mutually exclusive in BCs (31).

Looking at clinico-pathological parameters, based on our findings, GATA3 positivity is inversely associated with the most important clinical prognostic factors of BC outcome, namely tumor size and lymph node metastasis. Most of the previous studies that have examined GATA3 IHC expression in relation with tumor size and/or presence of lymph node metastasis have not found any association (76, 83, 93, 97, 100, 107), however four studies analogously reported inverse association between GATA3 and tumor size (67-69, 100) and only one reported low GATA3 expression associated with lymph node metastasis (67).

Altogether, our findings associate, univocally and clearly, GATA3 expression to favorable clinical, biological and pathological features of BC, coherently with and supporting previous observations.

To date, there is still undisguised controversy regarding the ability of GATA3 to predict the long-term prognosis in BC patients. Specifically, although high GATA3 gene expression

level has been convincingly associated with a better outcome of BC patients, independently of other clinico-pathological features (67, 108-111), the prognostic value of GATA3 protein level by IHC remains inconclusive (67-69, 82, 112). We found that GATA3 IHC expression was significantly associated with improved overall survival in our entire BC cohort. Moreover, we found that this association was also maintained in individual BC subgroups, particularly those at lower risk, such as ER-positive, PR-positive, HER2-negative, with lower proliferation (Ki-67<20%), wild-type p53, lower grade, smaller size (pT1-T2), without lymph node metastasis (pN0), at lower stage. To our knowledge, this is the first study that systematically has evaluated specifically the prognostic impact in so many BC subgroups based on all recognized prognostic factors. Furthermore, among intrinsic molecular subtypes, GATA3 expression was associated with a better overall survival only in luminal B.

It is well-known that different intrinsic molecular subtypes show distinct clinical evolution. Characteristically, luminal B BC has a poorer outcome compared with luminal A, but similar to HER2-positive and TNBC (25). Furthermore, luminal B exhibit a delayed aggressive behavior with the highest mortality rate after 5-8 years, when compared to HER2-positive and TNBC that progress rapidly with the highest mortality during the first two years after the diagnosis (113). In theory, late recurrence may be fostered by a withdrawal of endocrine therapy or may simply reflect its intrinsic aggressiveness and natural evolution. Of note, luminal B subtype has a distinctive pattern of metastasis dissemination, with a predilection for bone followed at a discreet distance by the lung. This is in contrast with luminal A BC, for which bone is also the principal site of metastasis, although it shows a low frequency of metastasis to other sites (114). Noticeable, there is an unmet need to identify patients with luminal B cancers who may benefit from adjuvant chemotherapy, particularly among those without lymph node metastasis. Therefore, given our results, GATA3 should be explored as a potential biomarker and could demonstrate an immediate, clinically relevant application in this contest and for this specific clinical issue.

Later, our multivariate analysis has demonstrated GATA3 as a strong, independent favorable prognostic factor for overall survival (OS) in BC patients, both in long-term and short-term outcome. On the other hand, disease-free survival (DFS) analysis could not identify a significant association between GATA3 and BC relapse at long-term follow-up, but, only in univariate analysis, our results have shown a predictive value of GATA3 in DFS at 48 months. Overall, in our BC cohort, GATA3 correlated effectively with improved OS, and demonstrated the strongest predictive power in multivariate analysis (HR 0.77, p=0.028), whereas in DFS it has not proved an independent prognostic effect.

A discrete number of studies have investigated the prognostic value of GATA3 by IHC in continuous series of BCs. The first study by Mehra *et al.* has analyzed GATA3 IHC in a

cohort of 139 consecutive invasive BCs arranged in TMA blocks and has found that BC patients with low GATA3 expression had a significantly shorter OS and DFS (67). The prognostic value of GATA3 IHC was demonstrated as independent from other prognostic variables by Cox multivariate analysis of DFS, but it was not shown for OS. Furthermore, low GATA3 was associated with DFS in patient subgroups without lymph node metastasis or with ER-positive BC. Successively, Voduc *et al.* could not confirm GATA3 IHC as an independent prognostic predictor in over 3100 consecutive cases of BCs as well as in ER-positive subgroup (69). Analogously, in the same year, Ciocca *et al.* found a non-significant modest protective effect of GATA3 positivity in OS and DFS of 166 consecutive BC patients (112). No association between GATA3 expression and outcome was found by Albergaria *et al.* in 249 consecutive BC patients (76). Thereafter, Yoon *et al.* could not only confirm the association of GATA3 IHC positivity with better BC-related survival in a consecutive series of 242 BCs, and in the ER-positive subgroup, but also found it in low-grade groups (68). Two studies have investigated in particular over 200 hormone receptor-positive and HER2-negative BCs, respectively; both found GATA3 significantly associated with better prognosis in univariate analysis, but not in multivariate analysis (98, 100). Then, McCleskey *et al*. obtained analogous results by analyzing 62 advanced BC patients (115). Another study has specifically evaluated the GATA3 prognostic impact on 516 BC patients treated with systemic therapy, i.e. chemotherapy and/or tamoxifen (82). GATA3 status could not reveal significant differences in OS among these patients, but GATA3 positivity affected negatively the OS of ER-negative patients (who do not receive tamoxifen), only in univariate analysis. Finally, two studies investigated the GATA3 prognostic effect in patients treated with neoadjuvant chemotherapy (116, 117). Both showed that GATA3 negativity was significantly associated with a more likely complete pathological response (116, 117). In summary, altogether these studies, that have interrogated GATA3 as a prognostic factor in BC patients, have found that even though GATA3 is associated significantly with a favorable prognosis, especially a prolonged DFS, it may not represent an independent prognostic factor, since in the majority of the studies GATA3 could not demonstrate its predictor value in multivariate analysis (118). Therefore, GATA3 has been shown so far as a critical biomarker associated with improved survival, but this association can depend on the association of GATA3 expression with different variables of good prognosis. Instead, our data have demonstrated a strong correlation of GATA3 positivity with improved prognosis both in OS and DFS, but an independent predictive value only for OS.

Ours is the second largest study that has analyzed GATA3 by IHC, taking advantage of a TMA approach and a monoinstitutional consecutive cohort of BCs. By comparison, we obtained many similar results to previous reports, including the association of GATA3 expression with ER-positivity, PR-positivity, HER2-negativity, luminal subtypes, low grade,

and smaller size, but also we found novel associations. Specifically, in our series GATA3 was associated with lower stage, lack of lymph node metastasis, and p53 wild-type IHC pattern, also all these are undoubtedly markers of favorable outcome. Moreover, our results demonstrated incontestably GATA3 as a prognostic factor in all BCs, but, more interestingly, it retained its predictive potential in BC subgroups with less aggressive characteristics, such as low grade, lower stage, ER positivity, PR positivity, HER2 negativity, and possibly wild-type TP53 BC subgroups. In addition, among molecular subtypes taken singularly, GATA3 negativity is associated with worse prognosis only in luminal B BC patients. More generally, our results emphasize the possibility that GATA3 plays a different prognostic role in various BC subtypes; this should be kept in mind for further studies, in order to better clarify the additional prognostic information supplied by GATA3 for therapeutic decision-making, especially in luminal B BCs.

Many differences among the studies that have analyzed GATA3 by IHC, including ours, may reflect differences in the IHC technique, analysis (especially cutoff definition), study design, or simply types of analyzed BCs, therefore a standardization of the methodology is needed. In particular, regarding the cutoff value to be used, we strongly encourage to use 1%, since it is easier and more reproducible.

Despite p53 was not our principal focus, by applying a different IHC evaluation method, we found interesting data, beyond a correlation between GATA3 positivity and p53 wild-type IHC pattern. Nevertheless, albeit whether *TP53* mutation is really inversely correlated with GATA3 positivity has to be confirmed in further studies, however, our results in terms of survival correlate mutant p53 pattern with an exceedingly worse prognosis at 48 months (HR 2.73 and 2.43 for OS and DFS, respectively). These findings are quite coherent with previous studies on the clinical value of *TP53* mutation in BC (103). For this reason, our results further bolster p53 IHC as a prognostic test in BC tissues.

In conclusion, our study, building on previous observations, provides new evidence of the prognostic value of GATA3 in BC, inasmuch it correlates GATA3 negativity to worse prognosis, especially in less aggressive BC subgroups. Indeed, GATA3 IHC could uncover BC patients with worse clinical outcome in low-risk categories, which would potentially benefit from additional tailored treatment. Our data, therefore, support the possible clinical utility of incorporating GATA3 IHC analysis into routine practice, as an adjunct to standard IHC panel, in order to further risk stratify BC patients. Since GATA3 IHC is currently used in routine diagnostic practice as a surrogate marker for breast and urothelial origin of carcinomas of unknown primary, this should be straightforward; however, standardized methods with univocal, reproducible cutoff are needed. Still, further controlled studies are required before applying GATA3 as a prognostic biomarker.

REFERENCES

1. Lakhani SR EI, Schnitt SJ, Tan PH, van de Vijver MJ. WHO Classification of Tumours of the Breast. Fourth ed. Lyon: International Agency for Research on Cancer; 2012.
2. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136:E359-86.
3. Lacey JV, Jr., Kreimer AR, Buys SS, Marcus PM, Chang SC, Leitzmann MF, et al. Breast cancer epidemiology according to recognized breast cancer risk factors in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial Cohort. BMC cancer. 2009;9:84.
4. Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, et al. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. Am J Hum Genet. 1998;62:676-89.
5. Easton DF. How many more breast cancer predisposition genes are there? Breast cancer research : BCR. 1999;1:14-7.
6. Consensus Conference on the classification of ductal carcinoma in situ. The Consensus Conference Committee. Cancer. 1997;80:1798-802.
7. Wellings SR, Jensen HM, Marcum RG. An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. Journal of the National Cancer Institute. 1975;55:231-73.
8. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 1991;19:403-10.
9. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25:118-45.
10. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2013;31:3997-4013.
11. Caleffi M, Teague MW, Jensen RA, Vnencak-Jones CL, Dupont WD, Parl FF. p53 gene mutations and steroid receptor status in breast cancer. Clinicopathologic correlations and prognostic assessment. Cancer. 1994;73:2147-56.
12. van Agthoven T, Timmermans M, Foekens JA, Dorssers LC, Henzen-Logmans SC. Differential expression of estrogen, progesterone, and epidermal growth factor receptors in normal, benign, and malignant human breast tissues using dual staining immunohistochemistry. Am J Pathol. 1994;144:1238-46.
13. Lal P, Tan LK, Chen B. Correlation of HER-2 status with estrogen and progesterone receptors and histologic features in 3,655 invasive breast carcinomas. Am J Clin Pathol. 2005;123:541-6.
14. Baselga J. Why the epidermal growth factor receptor? The rationale for cancer therapy. The oncologist. 2002;7 Suppl 4:2-8.
15. Cobleigh MA, Vogel CL, Tripathy D, Robert NJ, Scholl S, Fehrenbacher L, et al. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1999;17:2639-48.
16. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science. 1987;235:177-82.
17. Joensuu H, Kellokumpu-Lehtinen PL, Bono P, Alanko T, Kataja V, Asola R, et al. Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. The New England journal

of medicine. 2006;354:809-20.
18. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. The New England journal of medicine. 2005;353:1659-72.
19. Baselga J, Cortes J, Kim SB, Im SA, Hegg R, Im YH, et al. Pertuzumab plus trastuzumab plus docetaxel for metastatic breast cancer. The New England journal of medicine. 2012;366:109-19.
20. Burstein HJ, Harris LN, Marcom PK, Lambert-Falls R, Havlin K, Overmoyer B, et al. Trastuzumab and vinorelbine as first-line therapy for HER2-overexpressing metastatic breast cancer: multicenter phase II trial with clinical outcomes, analysis of serum tumor markers as predictive factors, and cardiac surveillance algorithm. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2003;21:2889-95.
21. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. The New England journal of medicine. 2001;344:783-92.
22. Wang S, Saboorian MH, Frenkel E, Hynan L, Gokaslan ST, Ashfaq R. Laboratory assessment of the status of Her-2/neu protein and oncogene in breast cancer specimens: comparison of immunohistochemistry assay with fluorescence in situ hybridisation assays. J Clin Pathol. 2000;53:374-81.
23. Rhodes A, Jasani B, Anderson E, Dodson AR, Balaton AJ. Evaluation of HER-2/neu immunohistochemical assay sensitivity and scoring on formalin-fixed and paraffin-processed cell lines and breast tumors: a comparative study involving results from laboratories in 21 countries. Am J Clin Pathol. 2002;118:408-17.
24. http://www.fda.gov/Drugs/default. U.S. Food and Drug Administration: Drugs.
25. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406:747-52.
26. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;98:10869-74.
27. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. Journal of the National Cancer Institute. 2009;101:736-50.
28. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thurlimann B, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO. 2013;24:2206-23.
29. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol. 2010;11:55-65.
30. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, et al. Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. The New England journal of medicine. 2015;373:2005-14.
31. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490:61-70.
32. Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nature communications. 2016;7:11479.
33. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature. 2012;486:400-4.
34. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ, et al. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Annals of oncology : official j

ournal of the European Society for Medical Oncology / ESMO. 2011;22:1736-47.
35. Lee MC, Rogers K, Griffith K, Diehl KA, Breslin TM, Cimmino VM, et al. Determinants of breast conservation rates: reasons for mastectomy at a comprehensive cancer center. Breast J. 2009;15:34-40.
36. Goldstein NS, Kestin L, Vicini F. Factors associated with ipsilateral breast failure and distant metastases in patients with invasive breast carcinoma treated with breast-conserving therapy. A clinicopathologic study of 607 neoplasms from 583 patients. Am J Clin Pathol. 2003;120:500-27.
37. Clarke M, Collins R, Darby S, Davies C, Elphinstone P, Evans V, et al. Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005;366:2087-106.
38. Goldhirsch A, Wood WC, Gelber RD, Coates AS, Thurlimann B, Senn HJ. Meeting highlights: updated international expert consensus on the primary therapy of early breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2003;21:3357-65.
39. Brufsky A. Trastuzumab-based therapy for patients with HER2-positive breast cancer: from early scientific development to foundation of care. Am J Clin Oncol. 2010;33:186-95.
40. Early Breast Cancer Trialists' Collaborative G. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005;365:1687-717.
41. Cadoo KA, Morris PG, Cowell EP, Patil S, Hudis CA, McArthur HL. Adjuvant Chemotherapy and Trastuzumab Is Safe and Effective in Older Women With Small, Node-Negative, HER2-Positive Early-Stage Breast Cancer. Clin Breast Cancer. 2016;16:487-93.
42. Kriege M, Seynaeve C, Meijers-Heijboer H, Collee JM, Menke-Pluymers MB, Bartels CC, et al. Sensitivity to first-line chemotherapy for metastatic breast cancer in BRCA1 and BRCA2 mutation carriers. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2009;27:3764-71.
43. Tutt A, Robson M, Garber JE, Domchek SM, Audeh MW, Weitzel JN, et al. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. Lancet. 2010;376:235-44.
44. Villarreal-Garza C, Khalaf D, Bouganim N, Clemons M, Pena-Curiel O, Baez-Revueltas B, et al. Platinum-based chemotherapy in triple-negative advanced breast cancer. Breast cancer research and treatment. 2014;146:567-72.
45. Kaufman B, Shapira-Frommer R, Schmutzler RK, Audeh MW, Friedlander M, Balmana J, et al. Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33:244-50.
46. Robert C, Thomas L, Bondarenko I, O'Day S, Weber J, Garbe C, et al. Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. The New England journal of medicine. 2011;364:2517-26.
47. Emens LA, Ascierto PA, Darcy PK, Demaria S, Eggermont AMM, Redmond WL, et al. Cancer immunotherapy: Opportunities and challenges in the rapidly evolving clinical landscape. Eur J Cancer. 2017;81:116-29.
48. Li Y, Li F, Jiang F, Lv X, Zhang R, Lu A, et al. A Mini-Review for Cancer Immunotherapy: Molecular Understanding of PD-1/PD-L1 Pathway & Translational Blockade of Immune Checkpoints. Int J Mol Sci. 2016;17.
49. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer. 2012;12:252-64.
50. Nanda R, Chow LQ, Dees EC, Berger R, Gupta S, Geva R, et al. Pembrolizumab in Patients With Advanced Triple-Negative Breast Cancer: Phase Ib KEYNOTE-012 Study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2016;34:2460-7.
51. Dirix LY, Takacs I, Jerusalem G, Nikolinakos P, Arkenau HT, Forero-Torres A, et al. Avelumab, an anti-PD-L1 antibody, in patients with locally advanced or metastatic breast cancer: a phase 1b JAVELIN Solid Tumor study. Breast cancer research and treatment. 2017.

52. Muenst S, Soysal SD, Gao F, Obermann EC, Oertli D, Gillanders WE. The presence of programmed death 1 (PD-1)-positive tumor-infiltrating lymphocytes is associated with poor prognosis in human breast cancer. Breast cancer research and treatment. 2013;139:667-76.
53. Chen Y, Bates DL, Dey R, Chen PH, Machado AC, Laird-Offringa IA, et al. DNA binding by GATA transcription factor suggests mechanisms of DNA looping and long-range gene regulation. Cell reports. 2012;2:1197-206.
54. Asselin-Labat ML, Sutherland KD, Barker H, Thomas R, Shackleton M, Forrest NC, et al. Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. Nat Cell Biol. 2007;9:201-9.
55. Ting CN, Olson MC, Barton KP, Leiden JM. Transcription factor GATA-3 is required for development of the T-cell lineage. Nature. 1996;384:474-8.
56. Kouros-Mehr H, Slorach EM, Sternlicht MD, Werb Z. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. Cell. 2006;127:1041-55.
57. Eeckhoute J, Keeton EK, Lupien M, Krum SA, Carroll JS, Brown M. Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer. Cancer research. 2007;67:6477-83.
58. Marconett CN, Sundar SN, Poindexter KM, Stueve TR, Bjeldanes LF, Firestone GL. Indole-3-carbinol triggers aryl hydrocarbon receptor-dependent estrogen receptor (ER)alpha protein degradation in breast cancer cells disrupting an ERalpha-GATA3 transcriptional cross-regulatory loop. Mol Biol Cell. 2010;21:1166-77.
59. Kong SL, Li G, Loh SL, Sung WK, Liu ET. Cellular reprogramming by the conjoint action of ERalpha, FOXA1, and GATA3 to a ligand-inducible growth state. Mol Syst Biol. 2011;7:526.
60. Liu Z, Merkurjev D, Yang F, Li W, Oh S, Friedman MJ, et al. Enhancer activation requires trans-recruitment of a mega transcription factor complex. Cell. 2014;159:358-73.
61. Theodorou V, Stark R, Menon S, Carroll JS. GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. Genome Res. 2013;23:12-22.
62. Usary J, Llaca V, Karaca G, Presswala S, Karaca M, He X, et al. Mutation of GATA3 in human breast tumors. Oncogene. 2004;23:7669-78.
63. Dydensborg AB, Rose AA, Wilson BJ, Grote D, Paquet M, Giguere V, et al. GATA3 inhibits breast cancer growth and pulmonary breast cancer metastasis. Oncogene. 2009;28:2634-42.
64. Yan W, Cao QJ, Arenas RB, Bentley B, Shao R. GATA3 inhibits breast cancer metastasis through the reversal of epithelial-mesenchymal transition. J Biol Chem. 2010;285:14042-51.
65. Tkocz D, Crawford NT, Buckley NE, Berry FB, Kennedy RD, Gorski JJ, et al. BRCA1 and GATA3 corepress FOXC1 to inhibit the pathogenesis of basal-like breast cancers. Oncogene. 2012;31:3667-78.
66. Jiang YZ, Yu KD, Zuo WJ, Peng WT, Shao ZM. GATA3 mutations define a unique subtype of luminal-like breast cancer with improved survival. Cancer. 2014;120:1329-37.
67. Mehra R, Varambally S, Ding L, Shen R, Sabel MS, Ghosh D, et al. Identification of GATA3 as a breast cancer prognostic marker by global gene expression meta-analysis. Cancer research. 2005;65:11259-64.
68. Yoon NK, Maresh EL, Shen D, Elshimali Y, Apple S, Horvath S, et al. Higher levels of GATA3 predict better survival in women with breast cancer. Human pathology. 2010;41:1794-801.
69. Voduc D, Cheang M, Nielsen T. GATA-3 expression in breast cancer has a strong association with estrogen receptor but lacks independent prognostic value. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2008;17:365-73.
70. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 2002;41:154-61.
71. Querzoli P, Coradini D, Pedriali M, Boracchi P, Ambrogi F, Raimondi E, et al. An immunohistochemically positive E-cadherin status is not always predictive for a good prognosis in human breast cancer. British journal of cancer. 2010;103:1835-9.
72. Kuhn E, Kurman RJ, Vang R, Sehdev AS, Han G, Soslow R, et al. TP53 mutations in serous

tubal intraepithelial carcinoma and concurrent pelvic high-grade serous carcinoma--evidence supporting the clonal relationship of the two lesions. The Journal of pathology. 2012;226:421-6.
73. Miettinen M, McCue PA, Sarlomo-Rikala M, Rys J, Czapiewski P, Wazny K, et al. GATA3: a multispecific but potentially useful marker in surgical pathology: a systematic analysis of 2500 epithelial and nonepithelial tumors. Am J Surg Pathol. 2014;38:13-22.
74. Kouros-Mehr H, Kim JW, Bechis SK, Werb Z. GATA-3 and the regulation of the mammary luminal cell fate. Curr Opin Cell Biol. 2008;20:164-70.
75. Parikh P, Palazzo JP, Rose LJ, Daskalakis C, Weigel RJ. GATA-3 expression as a predictor of hormone response in breast cancer. Journal of the American College of Surgeons. 2005;200:705-10.
76. Albergaria A, Paredes J, Sousa B, Milanezi F, Carneiro V, Bastos J, et al. Expression of FOXA1 and GATA-3 in breast cancer: the prognostic significance in hormone receptor-negative tumours. Breast cancer research : BCR. 2009;11:R40.
77. Wendroth SM, Mentrikoski MJ, Wick MR. GATA3 expression in morphologic subtypes of breast carcinoma: a comparison with gross cystic disease fluid protein 15 and mammaglobin. Ann Diagn Pathol. 2015;19:6-9.
78. Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. Journal of the National Cancer Institute. 2011;103:1656-64.
79. Krings G, Nystrom M, Mehdi I, Vohra P, Chen YY. Diagnostic utility and sensitivities of GATA3 antibodies in triple-negative breast cancer. Human pathology. 2014;45:2225-32.
80. Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. Journal of the National Cancer Institute. 1994;86:829-35.
81. Ellis CL, Chang AG, Cimino-Mathews A, Argani P, Youssef RF, Kapur P, et al. GATA-3 immunohistochemistry in the differential diagnosis of adenocarcinoma of the urinary bladder. Am J Surg Pathol. 2013;37:1756-60.
82. Gulbahce HE, Sweeney C, Surowiecka M, Knapp D, Varghese L, Blair CK. Significance of GATA-3 expression in outcomes of patients with breast cancer who received systemic chemotherapy and/or hormonal therapy and clinicopathologic features of GATA-3-positive tumors. Human pathology. 2013;44:2427-31.
83. Cakir A, Isik Gonul I, Ekinci O, Cetin B, Benekli M, Uluoglu O. GATA3 expression and its relationship with clinicopathological parameters in invasive breast carcinomas. Pathol Res Pract. 2017;213:227-34.
84. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. Mol Oncol. 2011;5:5-23.
85. Cianfrocca M, Gradishar W. New molecular classifications of breast cancer. CA Cancer J Clin. 2009;59:303-13.
86. Cimino-Mathews A, Subhawong AP, Illei PB, Sharma R, Halushka MK, Vang R, et al. GATA3 expression in breast carcinoma: utility in triple-negative, sarcomatoid, and metastatic carcinomas. Human pathology. 2013;44:1341-9.
87. Shaoxian T, Baohua Y, Xiaoli X, Yufan C, Xiaoyu T, Hongfen L, et al. Characterisation of GATA3 expression in invasive breast cancer: differences in histological subtypes and immunohistochemically defined molecular subtypes. J Clin Pathol. 2017;70:926-34.
88. Kim S, Moon BI, Lim W, Park S, Cho MS, Sung SH. Expression patterns of GATA3 and the androgen receptor are strongly correlated in patients with triple-negative breast cancer. Human pathology. 2016;55:190-5.
89. Huo L, Gong Y, Guo M, Gilcrease MZ, Wu Y, Zhang H, et al. GATA-binding protein 3 enhances the utility of gross cystic disease fluid protein-15 and mammaglobin A in triple-negative breast cancer by immunohistochemistry. Histopathology. 2015;67:245-54.
90. Dang DN, Raj G, Sarode V, Molberg KH, Vadlamudi RK, Peng Y. Significantly increased PELP1 protein expression in primary and metastatic triple-negative breast carcinoma: comparison with GATA3 expression and PELP1's potential role in triple-negative breast carcinoma. Human pathology. 2015;46:1829-35.

91. Byrne DJ, Deb S, Takano EA, Fox SB. GATA3 expression in triple-negative breast cancers. Histopathology. 2017;71:63-71.
92. Deftereos G, Sanguino Ramirez AM, Silverman JF, Krishnamurti U. GATA3 immunohistochemistry expression in histologic subtypes of primary breast carcinoma and metastatic breast carcinoma cytology. Am J Surg Pathol. 2015;39:1282-9.
93. Gonzalez RS, Wang J, Kraus T, Sullivan H, Adams AL, Cohen C. GATA-3 expression in male and female breast cancers: comparison of clinicopathologic parameters and prognostic relevance. Human pathology. 2013;44:1065-70.
94. Hoch RV, Thompson DA, Baker RJ, Weigel RJ. GATA-3 is expressed in association with estrogen receptor in breast cancer. Int J Cancer. 1999;84:122-8.
95. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. The New England journal of medicine. 2002;347:1999-2009.
96. Min KW, Kim DH, Do SI, Chae SW, Kim K, Sohn JH, et al. Expression Pattern of Smad4/GATA3 as a Predictor of Survival in Invasive Ductal Carcinoma of the Breast. Pathobiology. 2017;84:130-8.
97. Min KW, Kim DH, Do SI, Chae SW, Kim K, Sohn JH, et al. Negative association between GATA3 and fascin could predict relapse-free and overall survival in patients with breast cancer. Virchows Arch. 2016;468:409-16.
98. Hosoda M, Yamamoto M, Nakano K, Hatanaka KC, Takakuwa E, Hatanaka Y, et al. Differential expression of progesterone receptor, FOXA1, GATA3, and p53 between pre- and postmenopausal women with estrogen receptor-positive breast cancer. Breast cancer research and treatment. 2014;144:249-61.
99. Jacquemier J, Charafe-Jauffret E, Monville F, Esterni B, Extra JM, Houvenaeghel G, et al. Association of GATA3, P53, Ki67 status and vascular peritumoral invasion are strongly prognostic in luminal breast cancer. Breast cancer research : BCR. 2009;11:R23.
100. Hisamatsu Y, Tokunaga E, Yamashita N, Akiyoshi S, Okada S, Nakashima Y, et al. Impact of GATA-3 and FOXA1 expression in patients with hormone receptor-positive/HER2-negative breast cancer. Breast Cancer. 2015;22:520-8.
101. Goh AM, Coffill CR, Lane DP. The role of mutant p53 in human cancer. The Journal of pathology. 2011;223:116-26.
102. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med. 2004;10:789-99.
103. Olivier M, Langerod A, Carrieri P, Bergh J, Klaar S, Eyfjord J, et al. The clinical value of somatic TP53 gene mutations in 1,794 patients with breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2006;12:1157-67.
104. Simao TA, Ribeiro FS, Amorim LM, Albano RM, Andrada-Serpa MJ, Cardoso LE, et al. TP53 mutations in breast cancer tumors of patients from Rio de Janeiro, Brazil: association with risk factors and tumor characteristics. Int J Cancer. 2002;101:69-73.
105. Alsner J, Jensen V, Kyndi M, Offersen BV, Vu P, Borresen-Dale AL, et al. A comparison between p53 accumulation determined by immunohistochemistry and TP53 mutations as prognostic variables in tumours from breast cancer patients. Acta Oncol. 2008;47:600-7.
106. Kim JY, Park K, Jung HH, Lee E, Cho EY, Lee KH, et al. Association between Mutation and Expression of TP53 as a Potential Prognostic Marker of Triple-Negative Breast Cancer. Cancer Res Treat. 2016;48:1338-50.
107. Ademuyiwa FO, Thorat MA, Jain RK, Nakshatri H, Badve S. Expression of Forkhead-box protein A1, a marker of luminal A type breast cancer, parallels low Oncotype DX 21-gene recurrence scores. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2010;23:270-5.
108. Jenssen TK, Kuo WP, Stokke T, Hovig E. Associations between gene expressions in breast cancer and patient survival. Hum Genet. 2002;111:411-20.
109. Bertucci F, Houlgatte R, Benziane A, Granjeaud S, Adelaide J, Tagett R, et al. Gene expression profiling of primary breast carcinomas using arrays of candidate genes. Hum Mol Genet. 2000;9:2981-91.

110. West M, Blanchette C, Dressman H, Huang E, Ishida S, Spang R, et al. Predicting the clinical status of human breast cancer by using gene expression profiles. Proc Natl Acad Sci U S A. 2001;98:11462-7.
111. Liu J, Prager-van der Smissen WJ, Look MP, Sieuwerts AM, Smid M, Meijer-van Gelder ME, et al. GATA3 mRNA expression, but not mutation, associates with longer progression-free survival in ER-positive breast cancer patients treated with first-line tamoxifen for recurrent disease. Cancer Lett. 2016;376:104-9.
112. Ciocca V, Daskalakis C, Ciocca RM, Ruiz-Orrico A, Palazzo JP. The significance of GATA3 expression in breast cancer: a 10-year follow-up study. Human pathology. 2009;40:489-95.
113. Langerod A, Zhao H, Borgan O, Nesland JM, Bukholm IR, Ikdahl T, et al. TP53 mutation status and gene expression profiles are powerful prognostic markers of breast cancer. Breast cancer research : BCR. 2007;9:R30.
114. Smid M, Wang Y, Zhang Y, Sieuwerts AM, Yu J, Klijn JG, et al. Subtypes of breast cancer show preferential site of relapse. Cancer research. 2008;68:3108-14.
115. McCleskey BC, Penedo TL, Zhang K, Hameed O, Siegal GP, Wei S. GATA3 expression in advanced breast cancer: prognostic value and organ-specific relapse. Am J Clin Pathol. 2015;144:756-63.
116. Chen Y, Chen C, Yang B, Xu Q, Wu F, Liu F, et al. Estrogen receptor-related genes as an important panel of predictors for breast cancer response to neoadjuvant chemotherapy. Cancer Lett. 2011;302:63-8.
117. Tominaga N, Naoi Y, Shimazu K, Nakayama T, Maruyama N, Shimomura A, et al. Clinicopathological analysis of GATA3-positive breast cancers with special reference to response to neoadjuvant chemotherapy. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO. 2012;23:3051-7.
118. Guo Y, Yu P, Liu Z, Maimaiti Y, Chen C, Zhang Y, et al. Prognostic and clinicopathological value of GATA binding protein 3 in breast cancer: A systematic review and meta-analysis. PLoS One. 2017;12:e0174843.

ACKNOWLEDGEMENTS

I thank Dr. Massimo Pedriali for the precious help, Mr. Eros Magri for technical assistance, Dr. Paolo Giorgi Rossi for the biostatistical guidance, and Dr. Antonia Conti for the wonderful figures. Above all, I wish to gratefully and sincerely thank my advisors, Prof. Paola Secchiero e Prof. Patrizia Querzoli, for their invaluable advice and excellent guidance.