Aus dem Institut für Pathologie
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Expression statuses of FGFR2, TOX3 and LSP1 proteins in human breast cancer tissues and cell lines and the associations with clinical and pathological characteristics

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Juan Luo

aus Sichuan, China

Datum der Promotion: 09.09.2016

My Family

CONTENTS

Abstract

In recent years, several genome-wide association studies (GWAS) have identified novel breast cancer susceptibility loci which locate within or near some known genes, such as *FGFR2*, *TOX3* and *LSP1*. However, the biological roles of these plausible candidate genes in breast cancer still remain unclear. By using immunohistochemistry/immunocytochemistry, we detected the expression statuses of FGFR2, TOX3 and LSP1 in 110 invasive breast cancers (Inv-BCs, including 39 familial and 19 triple-negative breast cancers), 39 benign breast lesions and 13 human breast cell lines (5 non-tumorous and 8 cancerous cell lines). Histologically, a mixed intracellular localization of FGFR2 was observed in both malignant and benign breast epithelial cells, including one breast cancer cell line (T47D). Statistically, the expression and high-level expression of FGFR2 were detected in 75.2% and 22.9% of Inv-BCs, respectively, while no high-level expression was observed in benign lesions. Furthermore, negative correlation of FGFR2 expression with tumor grade and obviously positive correlations with ER and PR expressions were confirmed. Entirely negative for FGFR2 staining was even observed in 89.5% (17/19) of triple-negative breast cancers. Higher expression of FGFR2 in invasive lobular carcinoma (ILC) than in invasive ductal carcinoma (IDC) and medullary carcinoma (MEC) was further revealed. No associations of FGFR2 expression with other clinical and pathological characteristics of Inv-BC, including family history, were demonstrated. TOX3 staining was shown in nuclei of all kinds of observed epithelial and mesenchymal cells, including 13 breast cell lines. A lower expression of TOX3 in familial Inv-BCs than in sporadic ones and in cases positive for malignant tumor history than negative ones was further implied. No staining of LSP1 was detected in any breast epithelia or cell lines except for 2 Inv-BCs showing ambiguous staining.

In conclusion, FGFR2 is expressed at varying levels and intracellular localizations in both malignant and benign breast tissues and can be highly expressed or lose expression in Inv-BCs. Its expression is positively correlated with the expressions of ER and PR and negatively correlated with tumor grade. Nuclear expression of TOX3 was detected in all kinds of cells

observed in our study, and seemingly lower expression was shown in familial Inv-BCs and cases with malignant tumor history in other organs/tissues. But replicated and functional studies are still needed to clarify these tentative findings. No convincing expression of LSP1 occurs in either malignant or benign breast epithelial cells. LSP1 may not play direct roles in benign or malignant breast epithelial cells.

Keywords

breast cancer, breast cell lines, FGFR2 (fibroblast growth factor receptor 2), TOX3 (TOX high mobility group box family member 3), LSP1 (lymphocyte-specific protein 1), immunohistochemistry, immunocytochemistry

Zusammenfassung

In den letzten Jahren wurden mittels genomweiten Assoziationsstudien (GWAS) mit hereditären, nicht BRCA1- oder BRCA2-abhängigen Mammakarzinomproben neue Loci für ein erhöhtes Brustkrebsrisiko identifiziert. Diese liegen z.B. innerhalb oder in der Nähe der Gene für FGFR2, TOX3 und LSP1. Eine mögliche biologische Rolle dieser Kandidatengene in der Karzinogenese von Mammakarzinomen ist jedoch unklar.

Zur weiteren Charakterisierung einer möglichen Bedeutung dieser Faktoren beim Mammakarzinom wurde die Expression von FGFR2, TOX3 und LSP1 mittels Immunhistochemie in Gewebeproben von 110 invasiven Mammakarzinomen (Inv-BCs), 39 benignen Mammagewebeproben und 13 humanen Mammazellinien (8 Karzinomlinien, 5 nichttumoröse Linien) untersucht.

Hierbei zeigte sich eine gemischte intrazelluläre Lokalisierung von FGFR2 in Mammakarzinomgewebe, normalen Brustepithelzellen und in einer Mammakarzinomzelllinie (T47D). In 75,2% der Inv-BCs zeigte sich eine starke -, in 22,9% eine moderate Expression von FGFR2. Im benignen Mammagewebe fand sich keine Überexpression. Auβerdem konnte eine negative Korrelation der FGFR2-Expression mit dem Tumor-Grading und eine deutlich positive Korrelationen mit der Expression vom Östrogenrezeptor (ER) und Progesteronrezeptor (PR) nachgewiesen werden. Ferner zeigte sich eine höhere FGFR2 Expression im invasiven lobulären Karzinom (ILC) als im invasiven duktalen Karzinom (IDC) und MEC Mukoepidermoidkarzinom (MEC). Assoziationen der FGFR2-Expression mit anderen klinisch-pathologischen Parametern wurden nicht gefunden.

TOX3 konnte in unterschiedlichen ephithelialen und mesenchymalen Zellen einschlieβlich der 13 Mammazelllinien detektiert werden. Zudem fand sich eine geringere Expression von TOX3 bei familiären als bei sporadischen invasiven Mammakarzinomen.

Eine Färbung von LSP1 konnte lediglich in zwei invasiven Mammakarzinomen (jedoch nicht eindeutige Färbung) nachgewiesen werden.

Zusammenfassend kann festgestellt werden, dass aufgrund der höheren Expression von FGFR2 in den malignen Gewebeproben, FGFR2 tatsächlich eine Rolle bei der Entwicklung des

malignen Phänotyps bei diesen Tumoren spielen könnte. Die eher ubiquitäre Expression von TOX3 und die sehr geringe Expression von LSP1 sprechen dagegen gegen eine biologische Bedeutung dieser Faktoren beim Mammakarzinom.

Schlüsselwörter:

Brustkrebs, Mammakarzinom, FGFR2 (Fibroblastenwachstumsfaktor-Rezeptor 2), TOX3 (Mitglied der TOX High-Mobility-Group Box Familie 3), LSP1 (Lymphozyten-spezifisches Protein 1), Immunhistochemie

Abbreviations

1-Oct	organic cation transporter 1
ASCO/CAP	American Society of Clinical Oncology/College of American Pathologists
AR	androgen receptor
ATCC	American Type Culture Collection
BC	breast cancer
BCL-2	B-cell lymphoma 2
BLAST	Basic Local Alignment Search Tool
bp	base pair
BRCA1	breast cancer 1, early onset *or* breast cancer type 1 susceptibility protein
BRCA2	breast cancer 2, early onset *or* breast cancer type 2 susceptibility protein
C/EBP β	CCAAT/enhancer-binding protein β
Ca^{2+}	calcium
CAGF9	CAG trinucleotide repeat-containing gene F9
cAMP	cyclic adenosine monophosphate
CK	cytokeratin
CO_2	carbon dioxide
CRIBC	invasive cribriform breast cancer
DNA	deoxyribonucleic acid
DCIS	ductal carcinoma in situ
EGFR	epidermal growth factor receptor
EMA	epithelial membrane antigen
ER	estrogen receptor
FCS	fetal calf serum
FGF	fibroblast growth factor
FGFR2	fibroblast growth factor receptor 2

FISH	Fluorescence In Situ Hybridization
GWAS	genome wide-association studies
HE	haematoxyline-eosine
hEGF	human epidermal growth factor
HER2 (c-erbB2)	human epidermal growth factor receptor 2
HMG-box	high mobility group box
HPSGs	heparan sulphate proteoglycans
HRP	horseradish peroxidase
HRT	hormone replacement therapy
H score	histochemical score
ICC	immunocytochemistry
IDC	invasive ductal carcinoma
Ig	immunoglobulin
IHC	immunohistochemistry
ILC	invasive lobular carcinoma
Inv-BC	invasive breast cancer
IQR	interquartile range
L	liter
LD	linkage disequilibrium
LSP1	lymphocyte-specific protein 1
MAF	minor allele frequency
MAP3K1	mitogen-activated protein kinase kinase kinase 1
MEC	medullary carcinoma
min	minute
mRNA	messenger ribonucleic acid
MUC	mucinous carcinoma
NEC	neuroendocrine carcinoma
NLS	nuclear location signal
nM	nanomole per liter

No.	number
OC	ovarian cancer
OR	odds ratio
PBS	phosphate buffered saline
PDGFR	platelet-derived growth factor
pH	potential of hydrogen
PI3K	phosphoinositide 3-kinases
PLCγ	Phospholipase C-γ
PR	progesterone receptor
PTEN	phosphatase and tensin homolog
rpm	revolutions per minute
rs	reference SNP
RT	room temperature
RT-PCR	reverse transcription-polymerase chain reaction
SMA	smooth muscle actin
SNP	single nucleotide polymorphism
STK11/LKB1	serine/threonine kinase 11 /liver kinase B1
TBS	tris-buffered saline
TBST	tris-buffered saline Tween-20
TGFα	transforming growth factor α
TK	tyrosine kinase
TM	transmembrane
TNBC	triple-negative breast cancer
TNC	triple-negative cell line
TNRC9	trinucleotide repeat containing 9
TOX	thymocyte selected-association HMG-box
TOX3	TOX HMG-box family member 3
TP53	tumor protein p53
v/v	volume per volume

VEGF	vascular endothelial growth factor
W	watt
WHO	World Health Organization

1 INTRODUCTION

1.1 Genetic susceptibility to breast cancer

Breast cancer is the most frequently diagnosed cancer and the leading cause of cancer death in women worldwide, accounting for 23% (1.38 million) of the total new cancer cases and 14% (458,400) of the total cancer deaths in 2008 [1]. Although the major contributions to breast cancer risk are explained by environmental factors and hormonal and reproductive factors, such as alcohol consumption, age at menarche and menopause, parity and hormone replacement therapy (HRT), there is also good evidence for a genetic component [2-6]. Female individuals who have one or more first-degree relatives affected by breast cancer have approximately a 1.5- to 3-fold increased risk for developing breast cancer and the risk ratio varies with the age of individuals and the number of affected first-degree relatives [7,8]. Studies focusing on monozygotic twins and correlations of environmental factors with familial risk have suggested that most of the excess familial risk is likely to be explained by genetic factors rather than shared environments [3,9]. To date, different approaches, such as genetic linkage study, candidate gene-association study and genome-wide association study (GWAS), have been used to identify susceptibility loci contributing to breast cancer. These genetic variants associated with breast cancer risk can be classified as high-penetrance mutations that are rare in the population but associated with very high risk (relative risk of carriers versus non-carriers of 5 to >20); moderate-penetrance variants which are associated with moderate increased risk; and low-penetrance polymorphisms which are common but associated with small increases in breast cancer risk (relative risk <1.5) [9,10].

1.1.1 High-penetrance mutations

Genetic linkage studies conducted in the 1990s led to identify two tumor suppressor genes, *BRCA1* and *BRCA2*, whose mutations conferred a high risk of breast cancer [9,10]. *BRCA1* was initially localized to chromosome 17q21 by a genetic linkage study based on 23 families with 146 cases of breast cancer in 1990 [11] and was subsequently cloned in 1994 [12]. *BRCA2* was mapped to chromosome 13q12-13 in 1994 [13] and cloned in the following two years [14,15]

Genetic heterogeneity and penetrance analysis in 237 families, each of which contained at least 4 breast cancer cases, indicated that breast cancer was linked to *BRCA1*, *BRCA2* and other genes in estimated 52%, 32% and 16% of families, respectively. And the majority (81%) of breast-ovarian cancer families was due to *BRCA1* while the majority (76%) of families with male and female breast cancer was due to *BRCA2*. However, that the largest proportion (67%) of families due to other genes was found only in families with 4 or 5 cases of female breast cancer [16]. Pathologically, *BRCA1*-associated breast cancer differs from *BRCA2*-associated and non-*BRCA*-associated tumors. Although most of *BRCA1*- and *BRCA2*-associated tumors are invasive ductal carcinoma (IDC), *BRCA1*-associated ones are more likely to be diagnosed as atypical medullary carcinoma and poorly differentiated (grade 3, G3) tumors showing less tubule formation, higher mitotic count and more pleomorphism [17-19]. Particularly, breast cancer in *BRCA1* mutation carriers tends to be triple-negative (ER-, PR-, without HER2 overexpression) and shows a basal-like phenotype (expressing basal/myoepithelial markers, such as CK5/6, CK14, SMA and EGFR) [18].

Other high-penetrance mutations have been identified as part of cancer syndromes. Germline mutations in the *TP53* gene (localized at chromosome 17p13 and encoding the p53 protein) predispose to a rare spectrum of malignant tumors known as the Li-Fraumeni syndrome (LFS) and Li-Fraumeni-like syndrome (LFL), including soft tissue sarcomas, brain tumors, osteosarcoma, adrenocortical carcinoma, premenopausal breast cancer and other tumors [20-22]. LFS was initially proposed in 1969 and classically defined as a proband with a sarcoma at <45 years with a first-degree relative at <45 years with any cancer, plus an additional first- or second-degree relative in the same lineage with any cancer at <45 years or a sarcoma at any age. Birch definition and Eeles definition are two definitions of LFL, which are based on more extensive types of tumors and wider age-ranges of onset in related family members than LFS [21,22]. Another term, the *PTEN* hamartoma tumor syndromes (PHTS), has been used to describe a collection of several rare clinical syndromes, including Cowden syndrome (CS), Bannayan-Riley-Ruvalcaba syndrome (BRRS), Proteus syndrome (PS), and a Proteus-like syndrome (PSL). PHTS is characterized by germline mutations of the *PTEN* tumor suppressor gene, which is localized at chromosome 10q23.3 and encodes a major lipid phosphatase [23]. Age-related penetrance estimated in 368 patients with PHTS revealed the highest lifetime risk

(85.2%) for female invasive breast cancer (Inv-BC) compared with several other cancers and melanoma [24]. Peutz-Jeghers syndrome (PJS) is a rare but dominantly inherited condition characterized by benign intestinal hamartomatous polyps, mucocutaneous pigmentation and increased risk of cancers, including cancers of gastrointestinal tract, breast, testis and ovary [25]. The relative risk of non-cutaneous cancer was 9.9 in all patients, and 50.5 for gastrointestinal cancer and 20.3 for gynecologic and breast cancer in female patients [26]. Germline mutations of *STK11/LKB1* gene, localized at 19p13.3 and encoding a serine/threonine kinase, have been identified as causal gene mutations for PJS [25,27].

In spite of the high risks, the high-penetrance mutations mentioned above are rare in the population and are estimated to account for a relatively small percentage (about 20%-25%) of the familial risk of breast cancer [9,10].

1.1.2 Moderate-penetrance variants

A combination of family-based and population-based approaches has identified some relatively uncommon variants associated with modestly increased risk for breast cancer. These include variants in *CHEK2*, *ATM*, *BRIP1* and *PALB2* genes [9,10]. The *CHEK2* gene is located at chromosome 22q12.1 and encodes the checkpoint kinase 2 (CHEK2) protein which functions within the cellular network that responds to DNA damage and protects genomic integrity [28]. Germline mutations, especially c.1100delC, in *CHEK2* are associated with breast cancer in different populations with a combined odds ratio (OR) of 2.77 [29]. The *ATM* (for 'ataxia-telangiectasia mutated') gene is located at 11q22.3 and encodes a protein which belongs to the family of PI3K-related protein kinases (PIKK) and plays a central role in the complex processes, involving *TP53*, *BRCA1* and *CHEK2*, that repair DNA double-strand breaks. Mutations in the *ATM* gene have been shown in 0.5%-1% of Western populations and account for an autosomal recessive condition called *ataxia-telangiectasia*. The relative risk of *ATM*-mutation for breast cancer has been estimated to be 2.23 to 4.9 [30]. The *BRIP1* (BRCA1-interacting protein 1) gene maps to chromosome 17q22 near the *BRCA1* locus and encodes a DNA helicase that interacts with BRCA1 protein and then contributes to its DNA repair function [31]. Biallelic *BRIP1* mutations have been shown to cause Fanconi anemia

complementation group J (FA-J), and the relative risk of breast cancer associated with truncating mutations in *BRIP1* has been estimated to be 2.0 [32]. Fanconi anemia is a genetically heterogeneous recessive condition that includes 13 subtypes, 12 of which have been attributed to distinct genes [33] . Similar to *BRIP1*, biallelic mutations of *PALB2* (for 'partner and localizer of BRCA*2*'), a gene which is located at 16p12 and encodes a protein interacting with BRCA2, cause Fanconi anemia subtype FA-N and these monoallelic mutations have been estimated to be associated with an approximately 2-fold increased risk of female breast cancer [9,33].

As inactivating mutations in each of these modest risk genes are rare, with less than 1% of the population being heterozygote, the contribution of these genes to the overall familial risk of breast cancer is estimated to be less than 3% [9,10,33].

1.1.3 Low-penetrance polymorphisms

As shown above, although important high-penetrance genes, such as *BRCA1* and *BRCA2*, and some other moderate-penetrance variants underlying the genetic susceptibility to breast cancer have been identified during the past two decades, mutations of these genes are rare and can only explain a small fraction of the familial aggregation of this disease. The failure to identify further high-penetrance loci has led to the hypothesis that the susceptibility to this common cancer is in a polygenic model, and the residual risk is due to combined effects of a large number of more common but lower-penetrance variants [9,10,34].

1.1.3.1 Candidate gene-associated studies

With the progress in technology, some powerful approaches have been offered to identify low-penetrance variants. Early genetic association studies have been widely used and generally focused on limited numbers of polymorphisms in candidate genes that were suspected to be important in carcinogenesis [34-36]. However, this kind of individual studies with insufficient sample-size have some drawbacks, such as increased likelihood of false positive, lack of ability to detect moderate relative risks and low replicability [36]. With these drawbacks, only very limited common susceptibility alleles to breast cancer have been confirmed or replicated. The most convincing association among these variants has been confined to the nonsynonymous

variant D302H (rs1045485) in the caspase 8 gene (*CASP8*), which is located at human chromosome 2q33 and encodes one of the initiator caspases that transduce apoptotic signals from the death receptors on the cell surface [9,35-38]. The minor allele of *CASP8* D302H has been demonstrated to be associated with a reduced risk for breast cancer and shows a minor allele frequency (MAF) of 0.13-0.29 and a per-allele OR of 0.87-0.89 in different studies [35,38,39], and a four-SNP haplotype, including the D302H locus, has been identified with significant association with breast cancer (per-allele OR of 1.30) [39]. Another common variant with weaker evidence for an association with breast cancer risk is transforming growth factor β (*TGFB1*) L10P (rs1982073), which was indicated to be associated with PR- rather than PR+ tumors with an overall per-allele OR of 1.08 [35]. However, this association still needs further confirmation, and these two variants are estimated to account for approximately 0.3% and 0.2% of the excess familial risk of breast cancer in populations of European ancestry, respectively [35].

1.1.3.2 GWAS and related studies

Recently, technological advances have provided platforms, which allow hundreds of thousands of SNPs to be genotyped simultaneously, for analyzing risk alleles without prior knowledge of the position or function of certain genes [10,40]. To date, by adopting this agnostic approach, several GWAS have identified about 24 novel breast cancer susceptibility loci within regions containing genes or no known genes (Table 1.1) [40-52]. Among these loci, *FGFR2* (fibroblast growth factor receptor 2) and *TOX3* (TOX HMG-box family member 3, also known as *TNRC9* or *CAGF9*) loci are two replicable ones showing the largest effect sizes and most significant associations [40-42,45,48,49].

Table 1.1 Summary of 24 novel susceptibility loci to breast cancer identified by recent GWAS

Locus	SNP ID	Plausible gene	MAF[a]	per-allele OR	*P*-trend	Ref.[d]
1p11.2	rs11249433	LD block (*FCGR1B*, *NOTCH2*)	0.39[b]	1.16[c]	6.74×10^{-10}	[45]
2q35	rs13387042	no kown (nearest: *TNP1*, *IGFBP5*, *IGFBP2*)	0.497	1.2	1.3×10^{-13}	[41]
3p24	rs4973768	LD block (*SLC4A7*, *NEK10*)	0.46	1.11	4.1×10^{-23}	[44]
5p12	rs4415084	LD block (*FGF10*, *MRPS30*)	0.396[b]	1.16	6.4×10^{-10}	[43]

5p15	rs10069690	*TERT* (intron 4)	0.26	1.18	1.0×10^{-10}	[50]
5q11.2	rs889312	LD block (*MAP3K1*, *MGC33648*, *MIER3*)	0.28	1.13	7×10^{-20}	[40]
6q22.33	rs2180341	LD block (*ECHDC1*, *RNF146*)	0.211	1.41	2.9×10^{-8}	[52]
6q25.1	rs2046210	LD block (*ESR1*,*C6orf97*)	0.35	1.29	2.0×10^{-15}	[46]
8q24	rs13281615	no known	0.40	1.08	5×10^{-12}	[40]
9p21	rs1011970	LD block (*CDKN2A*, *CDKN2B*; *CDKN2BAS*)	0.17	1.09	2.5×10^{-8}	[48]
9q31.2	rs865686	no known (nearest: *KLF4*, *RAD23B*, *ACTL7A*)	0.39	0.89	1.75×10^{-10}	[49]
10p15	rs2380205	LD block (*ANKRD16*, *FBXO18*)	0.43	0.94	4.6×10^{-7}	[48]
10q21	rs10995190	*ZNF365* (intron 4)	0.15	0.86	5.1×10^{-15}	[48]
10q22	rs704010	LD block (*ZMIZ1*)	0.39	1.07	3.7×10^{-9}	[48]
10q26.13	rs2981582	*FGFR2*(intron 2)	0.38	1.26	2×10^{-76}	[40]
10q26.13	rs1219648	*FGFR2*(intron 2)	0.39	1.20[c]	1.1×10^{-10}	[42]
11p15.5	rs3817198	*LSP1*(intron 10)	0.30	1.07	3×10^{-9}	[40]
11q13	rs614367	no known (nearest: *MYEOV*, *CCND1*, *ORAOV1*, *FGF19*, *FGF4*, *FGF3*)	0.15	1.15	3.2×10^{-15}	[48]
12p11	rs10771399	LD block (*PTHLH*)	0.12	0.85	2.7×10^{-35}	[51]
12q24	rs1292011	no known (nearest: *MAPKAPK5*, *TBX3*)	0.41	0.92	4.3×10^{-19}	[51]
14q24.1	rs999737	*RAD51L1* (intron 12)	0.76[b]	0.94[c]	1.74×10^{-7}	[45]
16q12.1	rs3803662	LD block (*TOX3*, *LOC643714*)	0.25	1.2	10^{-36}	[40]
16q12.1	rs3803662	LD block (*TOX3*, *LOC643714*)	0.269	1.28	5.9×10^{-19}	[41]
17q23.2	rs6504950	LD block (*STXBP4* (intron 1), *COX11*, *TOM1L1*)	0.27	0.95	1.4×10^{-8}	[44]
19p13	rs8170	LD block (*ANKLE1*, *C19orf62*, *ABHD8*)	0.17	1.26	2.3×10^{-9}	[47]
19p13	rs2363956	LD block (*ANKLE1*, *C19orf62*, *ABHD8*)	0.52	0.84	5.5×10^{-9}	[47]
21q21	rs2823093	no known (nearest: *NRIP1*)	0.27	0.94	1.1×10^{-12}	[51]

[a]: minor allele frequency in controls (mainly in European ancestry); [b]: risk allele frequency;

[c]: heterozygous OR; [d]: reference

Further analyses indicated that the risk association of these common genetic variants that predispose to breast cancer may also be modified by clinical and pathological characteristics, particularly by ER, PR statuses, family history and *BRCA1/2* mutation status. For example, while

variants at the 5p15, 6q25.1 and 19p13 loci showed stronger risk association with ER- and triple-negative tumors, most of the other risk loci, such as *FGFR2*, *TOX3*, 8q24, and 1p11.2, showed stronger risk association with ER+ than ER- ones [41,43-47,50,51,53]. The remaining susceptibility loci, such as 10p15 and 10q22, showed similar association with both ER+ and ER- tumors [48]. When the analysis was narrowed in German familial breast cancer patients, significantly higher ORs of *FGFR2* and *TOX3* variants in high-risk breast cancers (≥3 breast cancer carriers in each family) were demonstrated than unselected cohort, and a putative novel susceptibility variant within *LSP1* gene seemed more predominant in high-risk groups [54]. Similarly, rs614367 at 11q13 showed a consistently stronger association with a positive family history at both GWAS stages [48]. Additional studies provided evidence that SNPs in *FGFR2*, *LSP1* and *MAP3K1* were associated with breast cancer risk in *BRCA2* mutation carriers and the SNPs in *TOX3* and 2q35 were associated with breast cancer risk in both *BRCA1* and *BRCA2* mutation carriers [55,56], and the association of one SNP in *TOX3* with familial *BRCA2* mutation carriers was replicated in another study [57]. The susceptibility variants, rs8170 and rs2363956, at 19p13, which were initially identified from *BRCA1* mutation carriers, seemed to be specifically associated with *BRCA1* but not *BRCA2* mutation carriers [47]. When ER status was further taken into consideration, the SNPs in *FGFR2* and *LSP1* showed further associations in ER+ and ER- *BRCA1* mutation carriers, respectively. However, the associations of the SNPs in *MAP3K1* and 2q35 in *BRCA1* mutation carriers could not be replicated [58].

As shown in Table 1.1, these relatively common susceptibility variants (MAF of >0.1) identified by GWAS are estimated to contribute to ~9% of familial risk of breast cancer, and the per-allele ORs for the risk alleles are usually <1.5-fold [51].

While guidelines for genetic testing of *BRCA1/2* and several models for determining carrier probabilities and cancer risks of individuals with a family history of breast cancer have been available, no risk prediction models or practical criteria concerning these polymorphisms have been established and validated. Thus, currently it is not the right time to introduce these low-penetrance variants into routine medical care. However, it is still possible to integrate the screening of these polymorphisms into risk and therapeutic assessments in the future for

BRCA1/2 mutation carriers and non-carriers with family history, as these variants can act as genetic modifiers in these family members [59,60].

Except for the needs to establish and validate the potential risk-prediction models of these susceptibility polymorphisms, to date, the mechanisms underlying these statistical significances still remain to be elucidated. Limited functional studies indicated that the expression of *FGFR2* might be up-regulated in breast cancer tissues by the risk allele through altering binding affinity for transcription factors Oct-1/Runx2 and C/EBP β, whereas the effect of risk allele on the expression of FGFR2 was opposite in normal breast tissues [61,62]. Although no difference of *TOX3* mRNA expression was found between normal and cancerous breast tissues in one research [41], the risk alleles of rs3803662 and rs12443621 at *TOX3* locus showed significant associations with lower mRNA expression of *TOX3* in breast cancer tissues in a dose-dependent manner in another study [63]. However, none of the risk alleles at another 6 loci, including *FGFR2*, *LSP1*, *MAP3K1* and *8q24*, was observed a significant correlation with tumorous mRNA expression of the nearest genes in the latter research [63].

1.2 *FGFR2*

1.2.1 Basic properties of *FGFR2* gene and FGFR2 protein

Among the plausible genes listed in Table 1.1, *FGFR2* gene, located at 10q26.13, consists of 21 exons and encodes a tyrosine kinase transmembrane receptor belonging to the FGFR family which contains 4 highly conserved members, FGFR1, FGFR2, FGFR3 and FGFR4 [64]. Structurally, all the four FGFRs are similar and each has up to 3 immunoglobulin-like (Ig-like) extracellular domains (IgI-IgIII), a transmembrane segment (TM) and an intracellular tyrosine kinase domain (TK). While the IgI domain and the acid box which is an acidic serine-rich sequence in the linker between IgI and IgII of each FGFR are proposed to have a role in receptor auto-inhibition, the IgII and IgIII fragments are necessary and sufficient for specific ligand (fibroblast growth factor, FGF) binding (Figures 1.1a and 1.1b). The human FGF family has at least 18 members, FGF1-10 and FGF16-23, which are secreted glycoproteins and classically considered to be paracrine factors except that FGF19, FGF21 and FGF23 have been shown to

function in an endocrine manner [65]. The FGF ligands exert their diverse functions by binding and activating FGFRs in an HPSGs (for 'heparan sulphate proteoglycans')-dependent and overlapping pattern and through several intracellular signaling pathways [66]. The alternative splicing at the extracellular ligand-binding site of FGFR1-3 leads to production of FGFR1b-3b and FGFR1c-3c isoforms which show distinct FGF binding specificities. Thus, it enhances the complexity and functional diversity of this FGF-FGFR system [65,67] (Figure 1.1b**)**. Moreover, these receptor isoforms are often cell lineage-specific. In the case of FGFR2, FGFR2 IIIb (FGFR2b) isoform is predominantly expressed in epithelial cells and shows high affinity for FGF1, FGF3, FGF7, FGF10 and FGF22, while FGFR2 IIIc (FGFR2c) is preferentially expressed in mesenchymal cells and shows affinity for FGF1, FGF2, FGF4, FGF6, FGF9, FGF16 and FGF20 [64,67].

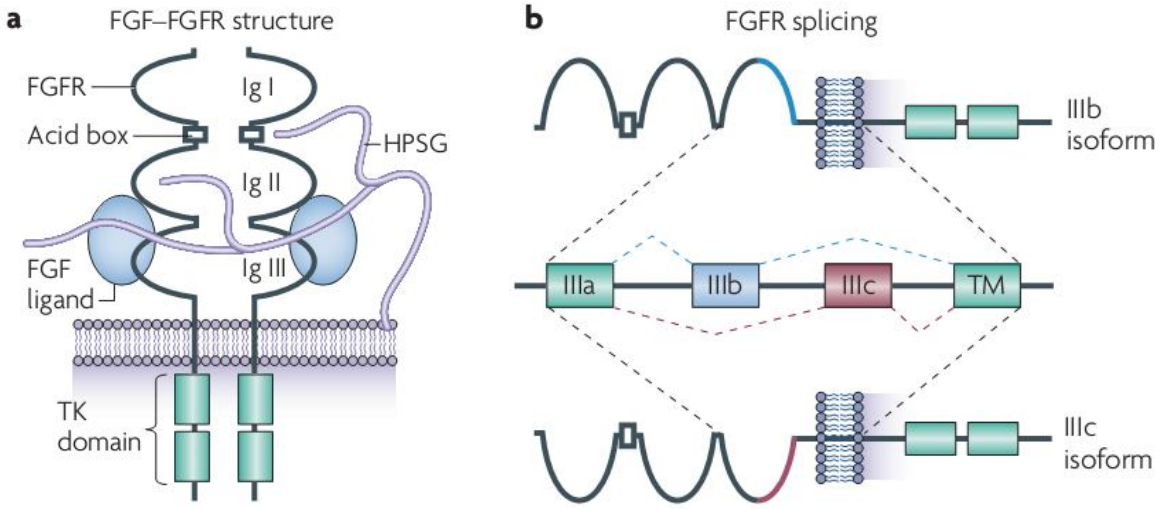

Figure 1.1 Schematic diagrams of a functional FGF-FGFR unit and alternative isoforms of FGFR1-3
a: A functional FGF-FGFR unit consists of two 1:1:1 FGF-FGFR-HPSG complexes juxtaposed in a symmetrical dimer. HPSGs facilitate FGF-FGFR dimerization by simultaneously binding to both FGF and FGFR, thereby promoting and stabilizing protein-protein contacts in the dimer [65,68]. **b.** Each of the monomer of FGFR1-3 has two alternative isoforms (IIIb and IIIc) which own two different carboxyl-terminal (C-terminal) half of the third Ig-like (IgIII) domains due to alternative splicing and thus manifest different ligand binding specificities and affinities [64,68]. **(The figures are quoted from Turner *et al.* [68])**

1.2.2 Physiological and pathophysiological roles of FGFR2

Generally, the FGFs-FGFR2 system participates in a variety of cellular processes during several organ systems' embryogenesis, adult tissue homeostasis, and carcinogenesis.

During mouse embryogenesis, while Fgfr2IIIc is expressed mainly in tissues of mesenchymal origin, Fgfr2IIIb is expressed in epithelium of many ectodermal and endodermal organs, including the mammary gland, and its specific ligands, Fgf10 and Fgf7, are usually expressed predominantly in mesenchyme adjacent to the epithelia expressing Fgfr2IIIb. Germline *Fgfr2*-knockout mice died shortly after birth because of multi-organ abnormalities, such as agenesis of lungs, limbs, mammary glands and so on. Unlike other branching organs, mammary branching morphogenesis may be divided into embryonic, adolescent and adult stages, each of which is differentially regulated by endocrine and local regulators. Selective abrogation of *Fgfr2IIIb* isoform (*Fgfr2IIIb-/-*) has revealed its crucial roles in instructive mesenchymal-epithelial signaling in several organ systems during mouse development and the mice lacking *Fgfr2IIIb* in particular showed a complete absence of mammary glands. Moreover, conditional ablation of FGFR2 leads to a severe delay in adolescent ductal development and the epithelia without FGFR2 are eliminated from the ducts that do develop (reviewed in [64,67,69,70]).

Pathophysiologically, some germline mutations in *FGFR2* cause several congenital, non-cancerous, skeletal disorders. Moreover, except for the association of SNPs at *FGFR2* loci with breast cancer risk mentioned above, somatic mutations in *FGFR2* have been demonstrated in endometrial, ovarian, breast, lung and gastric cancers. Missense mutations of *FGFR2* around the IgIII domain alter the ligand specificity and induce oncogenic FGFR2 activation, while C-terminally truncated FGFR2 induces constitutive activation of FGFR2 signaling cascades in a ligand-independent manner. Additionally, gene amplification and overexpression of *FGFR2* has also been demonstrated in 10% of human gastric cancer and ~1% of breast cancer. Research on FGFR2-targeted therapeutics for cancers is ongoing (reviewed in [40,64,68,69]). One recent study established a lapatinib-resistant cell line (UACC812/LR) in vitro from a HER2-positive parent breast cancer cell line and detected an amplification of FGFR2 gene, but a reduction of HER2 in this drug-resistant cell line. After treated with a small-molecule inhibitor of FGFRs, PD173074, the IC_{50} was 10,000 times lower in UACC812/LR than the parent cells. These results indicated a switch of addiction from the HER2 to the FGFR2 pathway enabled cancer cells to become resistant to HER2-targeted therapy and suggested that FGFR-targeted therapy might become a promising salvage strategy after lapatinib failure in patients with HER2-positive breast

cancer [71].

1.3 *TOX3*

Human *TOX3* gene was first identified in a screen for transcripts containing long CAG trinucleotide repeats which are associated with a number of neurodegenerative disorders of the human brain [72]. Later, murine LOC244579(*Tox3*) together with *Tox(Tox1),* LOC241768*(Tox2)* and LCP1(langerhans cell protein 1, *Tox4)* were identified as the TOX HMG-box subfamily by using BLAST program searches [73]. The human homologues of the latter 3 genes are *TOX* (*TOX1*, KIAA0808, at chromosome 8q21.1), *TOX2* (C20ORF100, at chromosome 20q13.12) and *TOX4* (KIAA0737, at chromosome 14q11.2), respectively. The 4 members of this TOX HMG-box subfamily, belonging to the HMG-box family which is one of three classes of HMG proteins, similarly have an amino-terminal (N-terminal) domain, a nuclear location signal (NLS), a C-terminal domain and share a common HMG-box domain next to NLS and likely fall into the sequence-independent rather than sequence-dependent category of HMG-box proteins. Each given member of this subfamily is highly conservative between murine and human and shows approximately 20% to 30% of identity outside of the highly similar NLS/HMG-box region among these 4 members [73].

Up to now, very limited functional assays concerning this TOX HMG-box subfamily have indicated that TOX (thymocyte selection-associated HMG-box) is abundantly expressed in the thymus and participates in the regulation of T-cell selection; the function of TOX2 in human is not characterized yet, while a rat ortholog of this gene is primarily expressed and functions in the hypothalamo-pitutary-gonadal axis of reproduction; TOX4 has been demonstrated to recognize DNA adducts specifically generated by platinum-based anticancer drugs and the LCP1(*Tox4*) gene is most highly expressed in testis, but the specific function of this protein is not yet known [73-75]. A few available studies [76,77] suggested that TOX3 was predominantly expressed in the brain and involved in mediating Ca^{2+}-dependent transcription in neuronal cells through interacting with both cAMP-response-element-binding protein (CREB) and CREB-binding protein (CBP). And this interaction induced the CREB-responsive *BCL-2* promoter and protected neuronal cells from cell death. Besides, TOX3 could also interact with CITED1 (for

'CBP/p300-interacting transactivator with glutamic-acid- and aspartic-acid-rich C-terminal domain 1') which enhances transcription mediated by diverse transcription factors (including ERs) and increase the estrogen-response element (ERE)-dependent transcription partly through EREs in the complement C3 promoter (Figure 1.2). One additional study which aimed to identify genes relevant to bone metastasis in breast cancer revealed higher expression of *TOX3* and other genes than non-bone relapsed ones[78]. Another recently published study focusing on epigenetic regulation of TOX HMG-box subfamily in lung and breast cancers has discovered a novel aberrantly hypermethylated CpG island within the *TOX2* promoter in 43% of breast cancer and 5% of lung cancer cases, whereas TOX3 was more frequently methylated in lung (58%) than in breast tumors (30%) and TOX4 was unmethylated in all samples and showed the highest expression in normal lung [75].

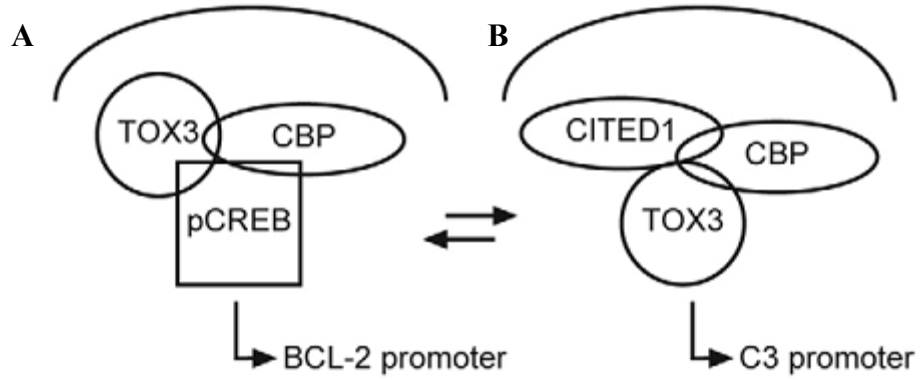

Figure 1.2 Schematic diagrams of proposed interacting and functioning mechanisms of TOX3 with CITED1 and CREB **A:** TOX3 was proposed to interact with pCREB (phosphorylated CREB)-CBP and then mediate the active transcription of *BCL-2* promoter. **B:** TOX3 was indicated to interact with CITED1 (but not CITED2, 3 or 4) and maybe together with an additional protein CBP to mediate an ERE-dependent transcription of C3 promoter in a ligand- or ER- independent way. **(The figures are quoted from Dittmer *et al.* [77])**

1.4 *LSP1*

The human *LSP1* gene is located at chromosome 11p15. Initially, this gene was identified from a mouse B lymphoma cell line and then isolated from human T cell lines [79-81]. Human and mouse LSP1 proteins are highly conserved showing 85% of identity in the C-terminal half and 53% of identity in the N-terminal half. The highly conservative basic C-terminal half of LSP1 contains the F-actin-binding site(s) and several serine/threonine residues that are identified as potential phosphorylation sites by serine/threonine kinase such as MAPKAPK2(MK2) or protein

kinase C (PKC), while the acidic N-terminal half contains two putative Ca^{2+}-binding motifs [82](Figure 1.3).

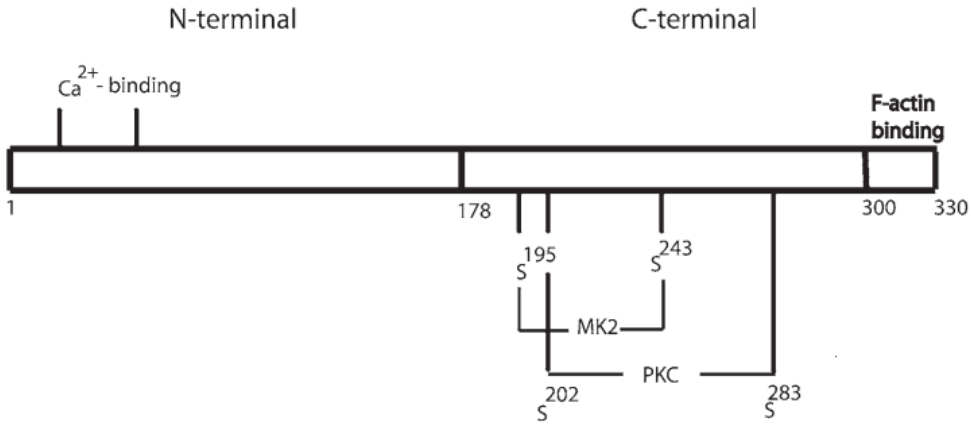

Figure 1.3 Schematic presentation of mouse LSP1 protein Human LSP1, consisting of 339 amino acids (a.a), and mouse LSP1, consisting of 330 a.a., are highly conserved. The putative Ca2+-binding sites locate at the N-terminal region (1-178), while the PKC, MK2 phosphorylation sites and the F-actin-binding region locate in the C-terminal region (178-330). **(Quoted from Jongstra-Bilen *et al.* [82])**

Although the variants at LSP1 loci were shown to be associated with increased breast cancer risk in some GWAS, previous studies have shown that LSP1 is restrictedly expressed in normal leukocytes, B-cell and some T-cell lymphomas and leukemias, localizing at the cytoplasmic face of the plasma membrane of the cells, but not in non-hematopoietic cells (including HT29 colon cancer and MCF7 breast cancer cell lines), tissues (including breast tissue) or tumors (including carcinomas) [79-81,83-85]. However, the studied cancer cell lines or cancer tissues with LSP1 are usually in a small sample-size. Additional studies also reported a nuclear localization of LSP1 in endothelial cells [86,87].

1.5 Working hypothesis

As mentioned above, the risk alleles of the *FGFR2*, *TOX3* and *LSP1* susceptibility loci were common in population (the risk allele frequency ranged approximately from 20% to 45%), their associations with breast cancer risk could be modified at least by ER and PR statuses, triple negative phenotype, family history and *BRCA1/2* germline mutation status, and the risk alleles may alter the expression of these genes. So we speculated that: (i) The FGFR2, TOX3 and LSP1 proteins might be expressed in breast cancer with varied expression levels. (ii) Their expressions might be associated with ER and PR statuses, family history and *BRCA1/2* germline

mutation status, and be different between TNBCs and non-TNBCs; (iii). There might be other clinical and pathological characteristics that can alter the expressions of these three proteins.

To date, little is known about the functions of TOX3 in breast caner. The restricted expression pattern of LSP1 still needs further investigation in relatively large sample size studies of breast cancer. Although a lot of studies of FGFR2 have been performed, its roles in breast cancer seem very complex and not fully clarified. Moreover, several recent studies reported a nuclear intracellular localization of FGFR2 in breast cancer that was not observed in previous research [88-91]. In order to clarify our speculations listed above and to confirm the possibility of a nuclear localization of FGFR2 in breast cancer, our current study detected the protein-level expressions of FGFR2, TOX3 and LSP1 by immunocytochemistry (ICC) in 5 non-tumorous and 8 cancerous breast cell lines and by immunohistochemistry (IHC) in 110 Inv-BCs (especially including 39 familial tumors with 16 ones among them having known *BRCA1/2* germline mutation statuses [92,93]), 27 benign breast tumors and 12 non-tumor breast lesions. Further statistical analyses were performed to investigate the associations of the expressions of these proteins with 15 clinical and pathological characteristics of Inv-BC.

2 MATERIALS AND METHODS

2.1 Materials

2.1.1 Equipments and Instruments

Equipment/Instrument	Manufacturer
BenchMark XT IHC/ISH Slide Stainer	Ventana Medical System, Inc.
Captair Filter 804N	Captair
Centrifuge GS-6KR	Beckman Coulter GmbH
CO_2 Incubator HeraCell 240	Heraeus Instruments GmbH
Lab Precision Balance BL1500S	Sartorius AG
Lab Water Purification Systems Milli-Ro/Milli-Q Plus	Millipore Corporation
LaminAir HB 2472 (Laminar flow workbench)	Heraeus Instruments GmbH
Light microscope DMRXA	Leica
Magnetic Stirrer RET Basic	IKA® Werke GmbH & Co.KG
Memmert Incubator Model 200	Memmert GmbH + Co.KG
Microwave Crisper 1L, 2.4L	Komax Industrial Co., Limited
Nikon ECLIPSE E200 Light Microscope	Nikon Instruments
Olympus Inverted Microscope IMT-2	Olympus Optical Co., LTD.
Panasonic Microwave Oven NN-3356	Panasonic Corporation
pH-Meter CG840	Schott-Geräte GmbH
Pipetboy acu	Integra Bioscience GmbH
Pipettes (10µl, 20µl, 100µl, 200µl, 1000µl)	Eppendorf AG
Steam Sterilizer, Varioklav Type 300/400/500 EP-Z	Heraeus Instruments GmbH
Water bath 1002	GFL

2.1.2 Consumables

Consumable	Manufacturer
SuperFrost® Plus Object Slides	R. Langenbrinck
Cover Glasses	R. Langenbrinck
PAP PEN	The Binding Site
Cell Culture Ware and Petri Dishes	BD Falcon
Pipette Tips	Eppendorf AG
Glass Ware	Schott-Geräter GmbH

2.1.3 Chemicals and Liquids

Chemical/Liquid	Manufacturer
Agarose Ultra Pure	Invitrogen GmbH
Cholera toxin	Sigma Co. LLC
Citric acid monohydrate	Merck KGaA
D-PBS(10×)	Invitrogen GmbH
Dimethyl sulphoxide (DMSO)	Sigma-Aldrich Co. LLC
β-Estradiol	Sigma Co. LLC
Ethanol 100 Vol. -% (MEK)	Herbeta Arzneimittel
Ethylene diamine tetraacetic acid, disodium salt dehydrate (EDTA)	Merck KGaA
Fetal calf serum (FCS)	Biochrom AG
Insulin	Sigma Co. LLC
Human epidermal growth factor (hEGF)	Sigma Co. LLC
Hydrocortisone	Sigma Co. LLC
L-Glutamine (200 nM)	Lonza Group Ltd.
Mayer's Haematoxylin	Dr. K. Hollborn & Söhne
Phosphate Buffered Formalin Solution	J.T. Baker

Polyethylene glycol sorbitan monolaurate (Tween 20)	SERVA Electrophoresis GmbH
Sodium chloride (NaCl)	Merck KGaA
Sodium Hydroxide (NaOH)	Merck KGaA
Tris(hydroxymethyl)-aminomethane (Tris-base)	Merck KGaA
Tris(hydroxymethyl)-aminomethane-hydrochlorid (Tris-HCl)	Merck KGaA
Tri-sodium citrate dihydrate	Merck KGaA
Trypan blue	Sigma-Aldrich Chemie GmbH
Tween-20	Merck KGaA
Vitro-Clud®	R. Langerbrink
Xylene	J.T. Baker

2.1.4 Antibodies

Antibody	**Catalog No.**	**Manufacturer**
Rabbit anti-human ER (monoclonal, Clone SP1)	RM-9101-S	Thermo Scientific
Mouse anti-human PR (monoclonal, Clone PgR 636)	M3569	Dako
Rabbit anti-human c-erbB-2 (HER2, polyclonal)	A0485	Dako
Mouse anti-human FGFR2 (monoclonal)	ab58201	abcam
Mouse anti-human LSP1(monoclonal, Clone 16/LSP-1)	610734	BD Biosciences
Rabbit anti-TOX3 (TOX3_Center, polyclonal)	AP4814c	ABGENT
Goat anti-mouse Immunoglobulins/HRP	P0447	Dako
Goat anti-rabbit Immunoglobulins/HRP	P0448	Dako

2.1.5 Commercial Solutions and Kits

Solution/Kit	**Catalog No.**	**Manufacturer**
Antibody Diluent Solution	00-3218	Invitrogen
DAB Detection Kit	K3468	Dako

Normal Goat Serum	PCN5000	Invitrogen
Peroxidase-Blocking Solution	S2023	Dako
Rabbit Primary Antibody Isotype Control	08-6199	Invitrogen
*ultra*View Universal DAB Detection Kit	760-500	Ventana
Universal Negative Control- Cocktail of mouse IgG_1, IgG_{2a}, IgG_{2b}, IgG_3 and IgM	N1698	Dako

2.1.6 Solutions and Buffers

10×TBS (pH7.4)	**1000ml**
Tris-base	9g
Tris-HCl	68.5g
NaCl	87.8g
dH_2O	up to 1000ml
Adjusting the pH to 7.4.	
Stored at room temperature (RT).	

10×TBST(pH7.4, 0.05% Tween20)	**1000ml**
10×TBS (pH7.4)	1000ml
Tween 20	5ml
Stored at RT.	

10×Citrate buffer (pH6.0)	**1000ml**
Citric acid	3.78g
Tri-Natriumcitrate-Dihydrate	24.21g
dH_2O	up to 1000ml
Adjust the pH to 6.0	
Stored at RT or 4℃ for longer storage.	

10×Tris-EDTA buffer (pH9.0, 0.05%Tween 20)	**1000ml**
Tris	12.1g
EDTA	3.7g
dH_2O	up to 1000ml
Adjusting the pH to 9.0 if necessary.	
Tween 20	5ml
Stored at RT or 4 ℃ for longer storage.	

2.1.7 Cell Culture Media

Medium	**Manufacturer**
MEGM (mammary epithelial growth medium)	Clonetics
DMEM (Dulbecco's modified eagle medium)	Lonza Group Ltd.
DMEM/Ham's F12 (1:1)	Biochrom AG
VLE RPMI1640	Biochrom AG

2.1.8 Software

Software	**Manufacturer**
DISKUS software (version 4.80.3505)	HILGERS Technisches Buero
SPSS PASW Statistics Software (version 18.0.0)	IBM/SPSS

2.1.9 Background of human breast cell lines included in this study

Except for the cell line, MDA-MB-453 was kindly provided by Prof. Denkert (Institute of Pathology at Charité University Hospital), all the other 12 cell lines included in this study were kindly provided by Dr. M. Theile (Department of Tumor Genetics, Max Delbrück Center for Molecular Medicine (MDC), Berlin, Germany). The background of the cell lines was collected from American Type Culture Collection (ATCC) or from publications (Table 2.1).

Table 2.1 Background of 13 human breast cell lines included in this study

Cell line	Type	Derivation	Properties	Ref.
hTERT-HME1	non-tumor	53y, mammoplasty	Expressing exogenous telomerase gene and not undergoing growth arrest; pan-CK+/MUC1+	ATCC
MCF 10A	non-tumor	36y, fibrocystic disease, non-tumorigenic	Showing no signs of terminal differentiation or senescence; not myoepithelial cells; positive for epithelial sialomucins, CKs, milk fat globule antigen, MFC-Breast, MC-5	ATCC
MCF 12A	non-tumor	60y, spontaneous immortalization, non-tumorigenic	Exhibiting typical luminal epithelial morphology; producing epithelial mucin and sialomucin, milk fat globule antigen	ATCC
184A1	non-tumor	21y, Chemically transformed	Being immortal but not malignant	ATCC
HBL100	non-tumor	27y, exfoliated milk cells 3 days after delivery	Showing several transformed characteristics	[94]
BT20	cancer	74y, primary: breast carcinoma	ER-, but do express an ER mRNA that has deletion of exon 5	ATCC
BRC230	cancer	79y, primary: infiltrating ductal carcinoma of breast	Primary tumor showed ER-/PR/-; showing predominant nucleoli	[95-97]
MCF7	cancer	69y, metastatic: pleural effusion	ER+, differentiated carcinoma; capability of forming domes	ATCC
MDA-MB-231	cancer	51y, metastatic: pericardial effusion	EGF+, TGFα+	ATCC
MDA-MB-453	cancer	48y, metastatic: pleural effusion	Overexpressing FGFRs	ATCC
T47D	cancer	54y, metastatic: pleural effusion	Differentiated, ER+/PR+/AR+/calcitonin+	ATCC
CAL51	cancer	44y, metastatic: after radiotherapy and chemotherapy, pleural effusion	ER-, CK+, EMA+; heterogeneous shapes: small or larger, more rounded	[98]
MDA-MB-435	cancer	31y, metastatic: pleural effusion	Previously derived from ductal carcinoma but then was suspected for contaminated with melanoma M14.	ATCC, [99]

2.1.10 Tissue samples

As the risk associations of the *FGFR2*, *TOX3* and *LSP1* susceptibility loci could be altered by family history and the gene expressions might be changed by these low-penetrance but high-frequency variants, we speculated that the expression of these three proteins might also be altered by family history. So we included one familial breast cancer group (Group 1) and one family-history-unselected breast cancer group (Group 2) in our study.

2.1.10.1 Group 1: familial breast cancer cases collected from 1981 to 1996

Thirty-one tumor blocks of 30 confirmed familial Inv-BC cases (at least one first-degree relative suffered from breast cancer (BC) and/or ovarian cancer (OC)) from 23 independent families, from 1981 to 1996, were selected from the collection of the Department of Tumor Genetics, MDC, Berlin, Germany. Details are listed in Table 2.2. Among them, 2 cases were male and the others were female. The median age at which the first BC was diagnosed in the selected patient (s) in each family was 48.0 (from 28 to 71) years and 7 (23.3%) cases were less than 40 years. Five cases were *BRCA1*, 4 were *BRCA2*, and 7 were non-*BRCA1/2* germline mutation carriers which were determined by previous studies [92,93]. According to WHO breast tumor classification (2003), this group contained 23 invasive ductal carcinoma, NOS (IDC), 3 invasive lobular carcinoma (ILC), 1 mucinous carcinoma (MUC), 1 invasive cribriform breast cancer (CRIBC), 1 mixed carcinoma (IDC+MUC), and 1 multifocal cancer which was regarded as two separate tumors in the following statistical analysis because of the different histological types (IDC and ILC) at two separate locations.

Table 2.2 Tumor tissue samples of familial breast cancer included in this study (n=39)

Fam. ID	Pat. Number[a]	Pat. ID[b]	Age (y)[c]	Subtype[d]	pT	pN	pM	G	*BRCA1/2*	Mutation[e]
Fam.30	8 BC	303	46	IDC	1	+	n.a.	2	*BRCA1*	Exon2, 185delAG
Fam.31	5 BC(4 males)	300(male)	49	IDC	1	+	n.a.	2	*BRCA2*	Exon10, 2041insA
Fam.52	3 BC	200	63	IDC	1	+	n.a.	2	n.a.	
		300	52	MUC	2	-	n.a.	2	n.a	
Fam.58	3 BC	301	70	IDC	1	-	n.a.	1	n.a.	
Fam.59	3 BC	300	55	ILC+IDC[f]	2	-	n.a.	1	n.a.	
Fam.60	10 BC,1 OC	302	55	ILC[h]	3	-	n.a.	2	*BRCA2*	del8894bp

		401	39	IDC	2	n.a.	n.a.	2	*BRCA2*	del8894bp
		416	28	IDC	2	+	0	3	*BRCA2*	del8894bp
Fam.81	5 BC(4 males)	202(male)	70	IDC	4	+	0	3	*BRCA1*	Exon23, 5448T>G
		305	35	IDC	1	-	n.a.	3	*BRCA1*	Exon23, 5448T>G
Fam.92	4 BC	302	54	IDC	1	n.a.	n.a.	2	no	
Fam.102	3 BC	302	46	IDC	1	+	n.a.	2	n.a.	
Fam.103	3 BC	300	43	ILC[h]	1	+	n.a.	1	n.a.	
Fam.105	2 BC,1 OC	401	34	IDC	1	+	n.a.	2	*BRCA1*	Exon11-A, 962del 4bp
Fam.1351	3 BC	300	37	IDC	1	-	n.a.	1	n.a.	
Fam.1551	3 BC	204	45	IDC	2	+	n.a.	2	n.a.	
Fam.1688	4 BC, 1 OC	209	71	IDC	2	+	n.a.	2	no	
Fam.2141	2 BC	203	47	IDC[h]	2	+	n.a.	2	n.a.	
		204	46	IDC	1	+	n.a.	1	n.a.	
Fam.2750	2 BC	300	47	(IDC+MUC)[g]	2	+	n.a.	1	*BRCA1*	Exon20, 5382insC
Fam.2927	2 BC	202	50	CRIBC	n.a.	n.a.	n.a.	1	n.a.	
Fam.2953	4 BC	202	54	ILC	1	-	n.a.	1	no	
Fam.3040	5 BC	200	60	IDC	2	+	0	2	n.a. (*BRCA1*)[i]	Exon20, 5382insC[i]
Fam.3641	3 BC	300	32	IDC	n.a.	n.a.	n.a.	3	n.a.	
Fam.3665	6 BC	300	49	IDC	2	n.a.	n.a.	2	n.a. (*BRCA2*)[i]	Exon11.3, 3036del4bp[i]
Fam.3692	3 BC	200	39	IDC[h]	n.a.	n.a.	n.a.	3	no	
Fam.3784	4 BC	200	67	IDC	2	-	n.a.	2	no	
		202	68	IDC	1	+	n.a.	2	no	
		300	46	IDC	3	+	n.a.	2	no	
Fam.C1	4 BC	Case57	66	IDC	1	-	0	1	no	
Fam.C2	2 BC	Case63	60	IDC	2	+	0	2	n.a.	
Fam.C3	4 BC	Case80	66	IDC	3	+	0	3	n.a.	
Fam.C4	2 BC(1 male)	Case89(male)	74	ILC	3	+	n.a.	2	n.a.	
Fam.C5	2 BC	Case94	69	ILC	2	+	0	2	n.a.	
Fam.C6	3 BC	Case100	69	MEC	2	+	0	3	n.a.	
Fam.C7	2 BC	Case117	65	NEC	4	n.a.	0	2	n.a.	
Fam.C8	2 BC	Case120	52	IDC[f]	2	+	0	2	n.a.	

Note: 1. Fam.C1 to Fam. C8 were from Group 2, others were from Group 1.

2. Tissue blocks of 4 cases could only be available in the relapsed tumors but the ages of carrying the first BC were included in final analysis.

3. Abbreviations: Fam.: family; Pat.: patient; BC: breast cancer; OC: ovarian cancer; IDC: invasive ductal carcinoma; ILC: invasive lobular carcinoma; MUC: mucinous carcinoma; CRIBC: invasive cribriform breast cancer; NEC: neuroendocrine carcinoma; n.a.: not available.

4. [a]: total number of BC/OC patients in one family; [b]: detected patient in this study; [c]: age at which the first BC

was diagnosed; [d]: histological subtype; [e]: germline mutation; [f]: multifocal or bifocal tumors; [g]: mixed tumor; [h]: relapsed lesion; [i]: with mutation carriers in the family but without mutation information of the patient included in this study.

2.1.10.2 Group 2: routine cases of breast cancer and benign lesions selected from 2003 to 2007

Seventy-eight Inv-BC (female, including 1 male) without neo-adjuvant chemotherapy or radiotherapy, 11 breast intraductal papilloma (female), 16 breast fibroadenoma (female) and 12 non-tumorous breast lesion (female, most of them were diagnosed as fibrocystic breast disease with hyperplasia) cases, from 2003 to 2007, were collected from the Institute of Pathology at Charité University Hospital in Berlin, Germany. The clinical data of Inv-BC cases, including age, family history, menopause status, HRT before diagnosis and history of suffering from malignant tumors in other organs/tissues (malignant history), were collected from the clinical archives and most of the information was available in 63 (80.8%) cases. Particularly, 8 cases having first-degree family history were included into the familial group (Table 2.2) and 48 cases showing no first-degree family history were regarded as sporadic cases. The pathological data, including multicentric/multifocal BC, bilateral BC, pTNM, tumor grade (G), histological subtype, ER, PR and HER2 statuses, were collected from the histopathological reports. One bilateral case was regarded as two separate tumors in the following analysis because it showed two different histological subtypes (IDC and CRIBC) at different sides. IHC staining of ER and PR was rescored according to Allred's scoring system [100]. IHC staining of HER2 was already scored in the reports according to HerceptTest^{TM} (Dako) scoring system and rescored according to the American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) guidelines [101].

2.1.10.3 Tissue microarray (TMA) samples

In order to verify the antibodies used in this study before the final detection on the samples included in Group 1 and 2, we first pre-tested these antibodies by using a TMA sample collection which contained 22 Inv-BC, 2 normal breast, 2 DCIS, 2 hysteromyoma, and 1 kidney cases from 1984 to 2008 and prepared by the Group of Prof. Denkert (Institute of Pathology at Charité University Hospital). However, because the pathological and clinical data were not available for

all the cases, these results were not included in our final analyses except to mention it when LSP1 was presented in the sections of ***Results*** and ***Discussion***.

2.2 Methods

2.2.1 Cell culture and cell block preparation

2.2.1.1 Cell culture

All the 13 cell lines were grown in cell-culture flasks containing proper media listed below (Table 2.3) and incubated in the incubator with 5% CO_2 and 95% humidity at 37℃. Once the cells reached confluence after 3- to 4-day cultivation, the medium was removed and the cells were washed by adding 1×PBS, and then digested by incubating in pre-warmed 0.05% trypsin/0.02% EDTA solution at 37℃ until all cells detached from the flask surface (checked under the inverted microscope). Subsequently, new medium was added to inactive the trypsin. The trypsinized cell suspension was then removed from the flask and transferred to a 15ml conical flask. After an aliquot for counting on a hemacytometer was removed from the conical flask, the rest cell suspension was spun in a centrifuge for 5min at 1000 rpm. Finally, after the cell count was calculated, cells were plated into new flasks at 2.5×10^5 per flask with fresh medium.

Table 2.3 Cell-culture medium for each cell line

Cell line	Type	Medium
hTERT-HME1	non-tumor	MEGM
MCF 10A	non-tumor	DMEM/F12(1:1) + 10%FCS + 2mM Glutamine + Hydrocortisone(0.5mg/ml) + hEGF(20ng/ml) + insulin(10μg/ml) + Cholera toxin(100ng/ml)
MCF 12A	non-tumor	DMEM/F12(1:1) + 10%FCS + 2mM Glutamine + Hydrocortisone(0.5mg/ml) + hEGF(20ng/ml) + insulin(10μg/ml) + Cholera toxin(100ng/ml)

184A1	non-tumor	DMEM/F12(1:1) + 10%FCS + 2mM Glutamine + Hydrocortisone(0.5mg/ml) + hEGF(20ng/ml) + insulin(10µg/ml) + Cholera toxin(100ng/ml)
HBL100	non-tumor	RPMI1640 + 10%FCS
BT20	cancer	RPMI1640 + 20%FCS + insulin
BRC230	cancer	DMEM/F12(1:1) + 10%FCS + 2mM Glutamine + Hydrocortisone(0.5mg/ml) + hEGF(20ng/ml) + insulin(10µg/ml) + Cholera toxin(100ng/ml)
MCF7	cancer	DMEM + 10%FCS+Estradiol(0.02mM)
MDA-MB-231	cancer	RPMI1640 + 10%FCS
MDA-MB-453	cancer	RPMI1640 + 10%FCS
T47D	cancer	RPMI1640 + 10%FCS + Estradiol(0.02mM)
CAL51	cancer	DMEM + 10%FCS
MDA-MB-435	cancer	RPMI1640 + 10%FCS

2.2.1.2 Cell block preparation

When growing to confluence, 5 to 8 plates of cells were digested slightly by 0.1% trypsin-EDTA in PBS. Subsequently, fresh medium was added to stop the enzyme reaction. After being centrifuged in a 50ml conical tube for 5min at 1000 rpm, the supernatant was discarded and the cell pallet was washed once in 1×PBS. Then the cell pallet was fixed in 10% (v/v) phosphate buffered formalin solution for 1 to 2 hours at RT. At the same time during fixation, 2% agarose in 1×PBS was prepared by boiling in a microwave oven and kept warm in a water bath or thermomixer at 70℃ until usage. When the fixation was finished, the formalin solution containing cells was centrifuged and washed once in 1×PBS. After the supernatant was discarded, the fixed cells were resuspended in 200µl to 350µl prewarmed 2% agarose solution and dropped by a pipet tip onto a pre-cooled water-resistant paper sheet (from the parafilm). After becoming solid, the cell-drop was transferred into an embedding cassette for dehydration (performed by an automatic dehydration machine sequentially in 80% ethanol for 20min ×1 time, 100% ethanol for 20min ×2 times and xylene for 20min ×2 times) and finally embedded in paraffin.

2.2.2 HE staining and IHC/ICC

Cell and tissue blocks were sectioned into 2μm sections prior to mounting onto glass slides. HE staining was performed by routine HE-staining group in the Institute of Pathology with an HE autostainer. IHC/ICC of ER, PR and HER2 for cases from Group 1 and 13 cell lines were performed by IHC Lab in the same institute using IHC autostainer. ER (1:25), PR (1:50) and c-erbB-2 (1:300) primary antibodies and ultraView Universal DAB Detection Kit were applied for autostaining.

IHC/ICC of FGFR2, TOX3 and LSP1 in all tissues and cell lines were performed manually. Briefly, deparaffinized and rehydrated sections were put into microwave pre-boiled (800W×7min) antigen retrieval buffer (citrate buffer pH6.0 for FGFR2 and LSP1, EDTA-Tris pH9.0 for TOX3) and heating was continued at 250W×15min. Then the samples were blocked with Peroxidase Blocking Solution and 10% normal goat serum (diluted by 1×TBS) and then incubated with anti-FGFR2 (1:200), anti-TOX3 (1:100) or anti-LSP1 (1:600) primary antibody for 1 hour at RT. Goat anti-mouse Immunoglobulins/HRP (1:100) or goat anti-rabbit Immunoglobulins/HRP (1:150) secondary antibody was used according to the primary antibody for 30min at RT. The immunoreaction was finally visualized by DAB Detection Kit. Nuclei were counterstained with Mayer's hematoxylin. Normal breast, brain and thymus tissues were stained simultaneously as positive control for FGFR2, TOX3 and LSP1, respectively. Mouse or rabbit primary antibody isotype control was used instead of corresponding primary antibody as negative control.

2.2.3 Scoring of IHC/ICC

IHC of ER, PR, HER2, FGFR2, TOX3 and LSP1 in Inv-BC, benign tumors, non-tumor breast conditions (non-tumor) and ICC of cell lines were assessed by two pathologists. The intensity of staining was recorded as 0 (no staining), 1 (weak, light brown), 2 (moderate, brown) or 3 (strong, dark brown). The percentage of positive cells was recorded as 0% to 100%. Final scores of ER and PR were calculated according to Allred's scoring system [100]. In detail, the portion of positive cells was scored as 0 (no staining), 1 (>0 to 1/100), 2 (>1/100 to 1/10), 3 (>1/10 to 1/3), 4 (>1/3 to 2/3) and 5 (>2/3 to 1), and the intensity of staining was scored as 0 (negative), 1 (weak), 2 (intermediate) and 3 (strong). Then these two scores were added to a total score and the total score of ≥3 was finally regarded as positive. Assessment of HER2 was done by IHC

categories 0, 1+, 2+ and 3+ according to Dako's HerceptTest^{TM} interpretation and the ASCO/CAP guidelines [101] and recorded here as Hercept score. Finally, cases with Hercept score of 3+ (uniform intense membrane staining of >30% of invasive tumor cells) or gene amplification confirmed by FISH (from the pathological reports) were regarded as HER2 overexpression. The expression of FGFR2, TOX3 and LSP1 was estimated as histochemical score (H score) [102,103], in which the percentage of positive cells staining in each intensity category (0, no staining; 1, weak; 2, moderate; or 3, strong) was derived, multiplied by its intensity and summed (range, 0 to 300). H score ≤ 10, 10 < H score ≤100, 100 < H score ≤ 200, 200 < H score ≤ 300 were finally regarded as negative (-), weak positive (+), moderate positive (++) and strong positive (+++), respectively. Cases with tissue loss, including 2 in FGFR2, 3 in TOX3 and 3 in LSP1 staining, or without available clinical or pathological data were not included into final analyses.

2.2.4 Statistical analysis

During statistical analyzing, expression proportions of FGFR2 and TOX3 in different characteristic subsets were compared by Fisher's exact test. Comparisons of expression level (reflected by H score) were performed using nonparametric tests. In detail, Mann-Whitney *U* test was chosen for 2-independent sample comparison, Kruskal-Wallis and Nemenyi test were chosen for multi-independent sample comparison. Correlations of ER and PR Allred's score with the H scores of FGFR2 and TOX3 were estimated by using Spearman's ranked correlation methods. All the statistical analyses were carried out using SPSS PASW Statistics Software (version 18.0.0, USA). Significance was defined as a *P*-value of <0.05 (two-sided).

3 RESULTS

3.1 Expression statuses of ER, PR, HER2, FGFR2, TOX3 and LSP1 in 13 breast cell lines

Breast cancer is a heterogenous disease. ER, PR and HER2 expression statuses are three important profiles of this disease for clinical diagnosis, prognosis and therapy. As research has demonstrated that cell lines also exhibit pronounced genomic heterogeneity and biological characteristics as with primary tumors [104], we first detected the expression pattern of ER, PR and HER2 in 13 breast cell lines with ICC (Table 3.1). Consistent with earlier studies [95,104-108], expression of both ER and PR was detected in MCF7 and T47D breast cancer cell lines (Figure 3.1). None of the other 10 non-tumorous or cancerous breast cell lines was detectable for ER or PR. Although previous reports showed no amplification/overexpression of *ERBB*2/HER2 in BT20 [104-107], we observed a strong positive staining of the complete membrane in >30% cells (Hercept score of 3+ [101]). Amplification of *ERBB2* was observed in MDA-MB-453 but without overexpression in previous studies [104-107], however, we saw a predominant plasma membrane staining in >30% cells (Hercept score of 3+) (Figure 3.2). Therefore, according to the expression patterns of ER, PR and HER2 detected in the current study, 4 breast cancer cell lines (BRC230, MDA-MB-231, CAL51 and MDA-MB-435) and 5 non-tumor breast cell lines were regarded as triple-negative beast cell lines (TNC) and 2 (BT20 and MDA-MB-453) were regarded as HER2 overexpression cell lines for the following analysis.

As shown in Figure 3.3, positive staining of FGFR2 was only observed in the cytoplasm of about 80% T47D cells and negative in the other 12 cell lines, although <1% positive cells could be observed in MDA-MB-453, CAL51 and MDA-MB-435 (Figure 3.3 A to E). By contrast, TOX3 was detected in all the 13 breast cell lines with a clear nuclear and sometimes also nucleolar localization and with moderate to strong intensity (Figure 3.3 F to J). Although statistical analysis revealed no significance, the expression level of TOX3 (H score) seemed lower in cancer cell lines than in non-tumor ones (Figure 3.4 A). No association of the TOX3 expression with ER, PR or HER2 status was determined (Figure 3.4 B to E). No positive staining for LSP1 was detected in any cell line, while it was strong positive in thymocytes (as positive control)

(Figure 3.3 K to O).

The *BRCA1* mutation status [109] and genotypes of variants of rs2981582 at *FGFR2*, rs3803662 at *TOX3* and rs3817198 at *LSP1* loci [106] had been detected in 5 of the 8 breast cancer cell lines (Table 3.1). Because all the 5 detected cell lines harbored wild type of *BRCA1*, no comparison could be done. The sole cell line positive for FGFR2 (T47D) showed wild type of *BRCA1* and homozygote of major allele for the *FGFR2* risk locus (rs2981582). Homozygote of major allele at rs3803662 presented in 3 cell lines with moderate expression level of TOX3 (H score of 135 to 180); heterozygote presented in 1 cell line with a TOX3 H score of 180; homozygote of minor allele presented in 1 cell line with a TOX3 H score of 135 (Figure 3.4 F). In spite of the small sample size used to perform statistical comparison, this result could not exclude the possibility that the expression of TOX3 might be different between different variants at rs3803662. Irrespective of the kind of allele that was harboring at rs3817198, no any expression of LSP1 was detected.

Table 3.1 Expression statuses of ER, PR, HER2, FGFR2, TOX3 and LSP1 in 13 human breast cell lines

Cell line	*BRCA1*[a]	Type	ER	PR	HER2[b]	TNC	FGFR2[c]	TOX3[c]	LSP1[c]
hTERT-HME1	n.a.	non-tumor	-	-	1+	yes	-	+	-
MCF10A	n.a.	non-tumor	-	-	2+	yes	-	+	-
MCF12A	n.a.	non-tumor	-	-	2+	yes	-	+	-
184A1	n.a.	non-tumor	-	-	2+	yes	-	+	-
HBL100	n.a.	non-tumor	-	-	1+	yes	-	+	-
BT20	wild	cancer	-	-	3 +	no	-/Maj	+/Maj	-/Het
BRC230	n.a.	cancer	-	-	2+	yes	-	+	-
MCF7	wild	cancer	+	+	2+	no	-/Maj	+/Het	-/Maj
MDA-MB-231	wild	cancer	-	-	2+	yes	-/Het	+/Maj	-/Het
MDA-MB-453	wild	cancer	-	-	3+	no	-/Min	+/Maj	-/Maj
T47D	wild	cancer	+	+	2+	no	+/Maj	+/Min	-/Min
CAL51	n.a.	cancer	-	-	2+	yes	-	+	-
MDA-MB-435	n.a.	cancer	-	-	2+	yes	-	+	-

[a]: *BRCA1* mutation status [109]: wild: wild type, without mutation;

b: Hercept score of HER2 for ICC;

c: Genotyping of rs2981582 in *FGFR2*, rs3803662 near *TOX3* and rs3817198 in *LSP1* was reported in reference [106]. Maj: Homozygote of major allele; Min: Homozygote of minor allele; Het: Heterozygote;

+: positve for ICC; -: negative for ICC; n.a.: not available; yes: TNC (triple-negative cell line); no: non-TNC.

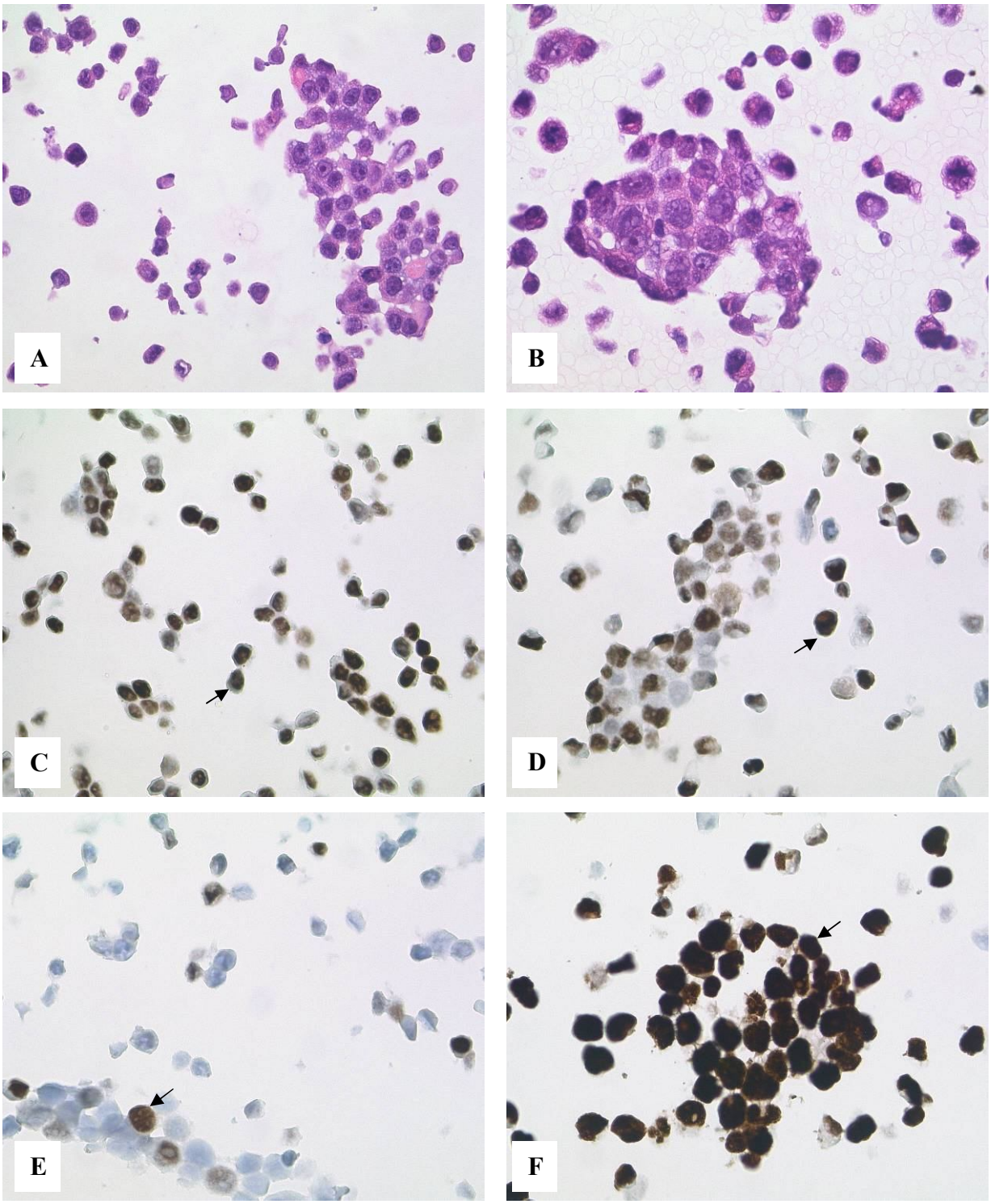

Figure 3.1 Representative HE staining and immunostaining of ER and PR in breast cancer cell lines (HE or DAB×400) **A , B:** HE staining of MCF7 (**A**) and T47D (**B**); **C, D:** Nuclear immunostaining of ER (↑) in MCF7 (Allred's score of 7, **C**) and T47D (Allred's score of 7, **D**); **E, F:** Nuclear immunostaining of PR (↑) in MCF7 (Allred's score of 5, **E**) and T47D (Allred's score of 8, **F**).

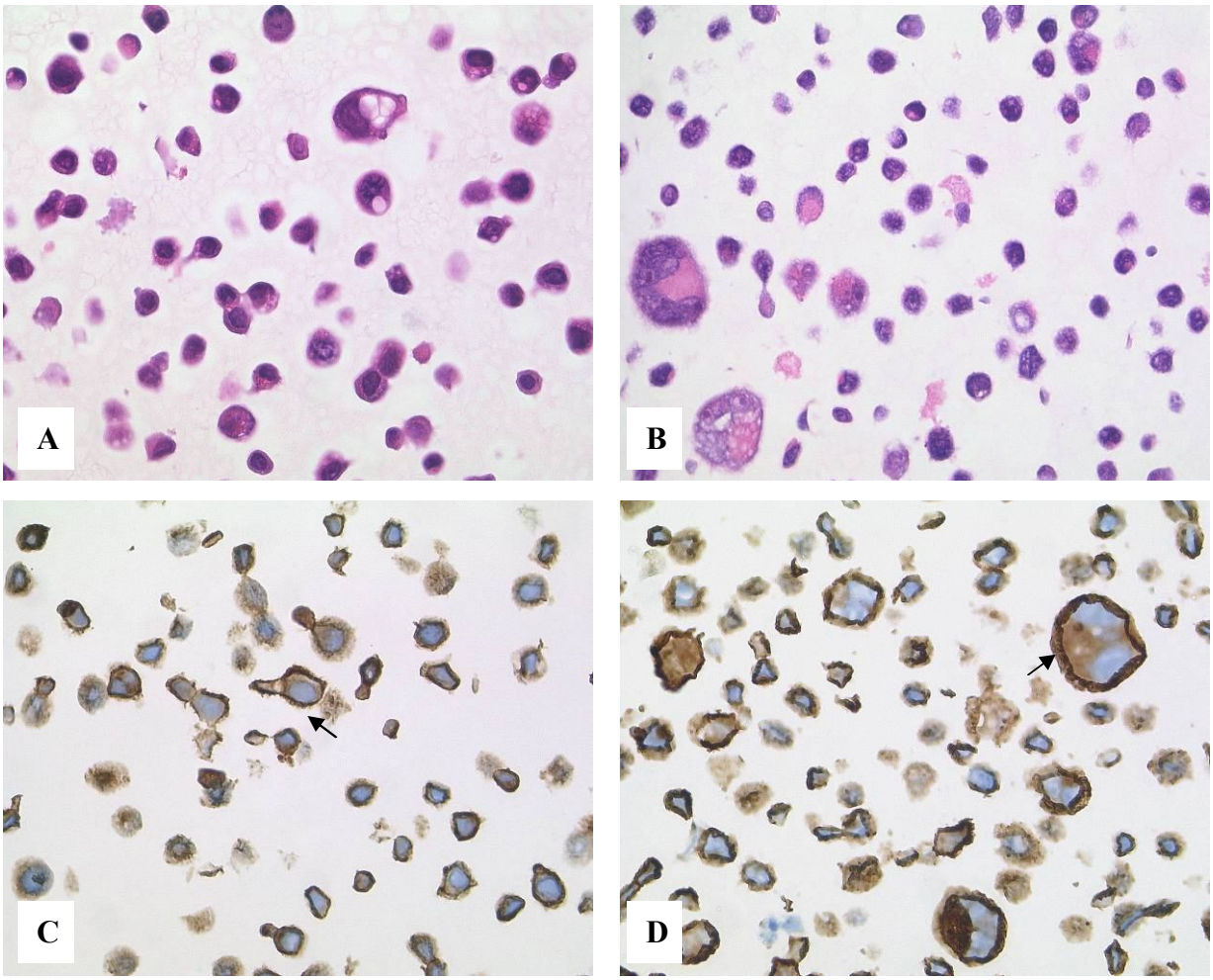

Figure 3.2 Representative HE staining and immunostaining of HER2 in breast cancer cell lines (HE or DAB×400) **A , B:** HE staining of BT20 (**A**) and MDA-MB-453 (**B**); **C, D:** Strong and complete plasma membrane staining of HER2 (↑) in >30% cells of BT20 (**C**) and MDA-MB-453 (**D**), especially predominant in the latter.

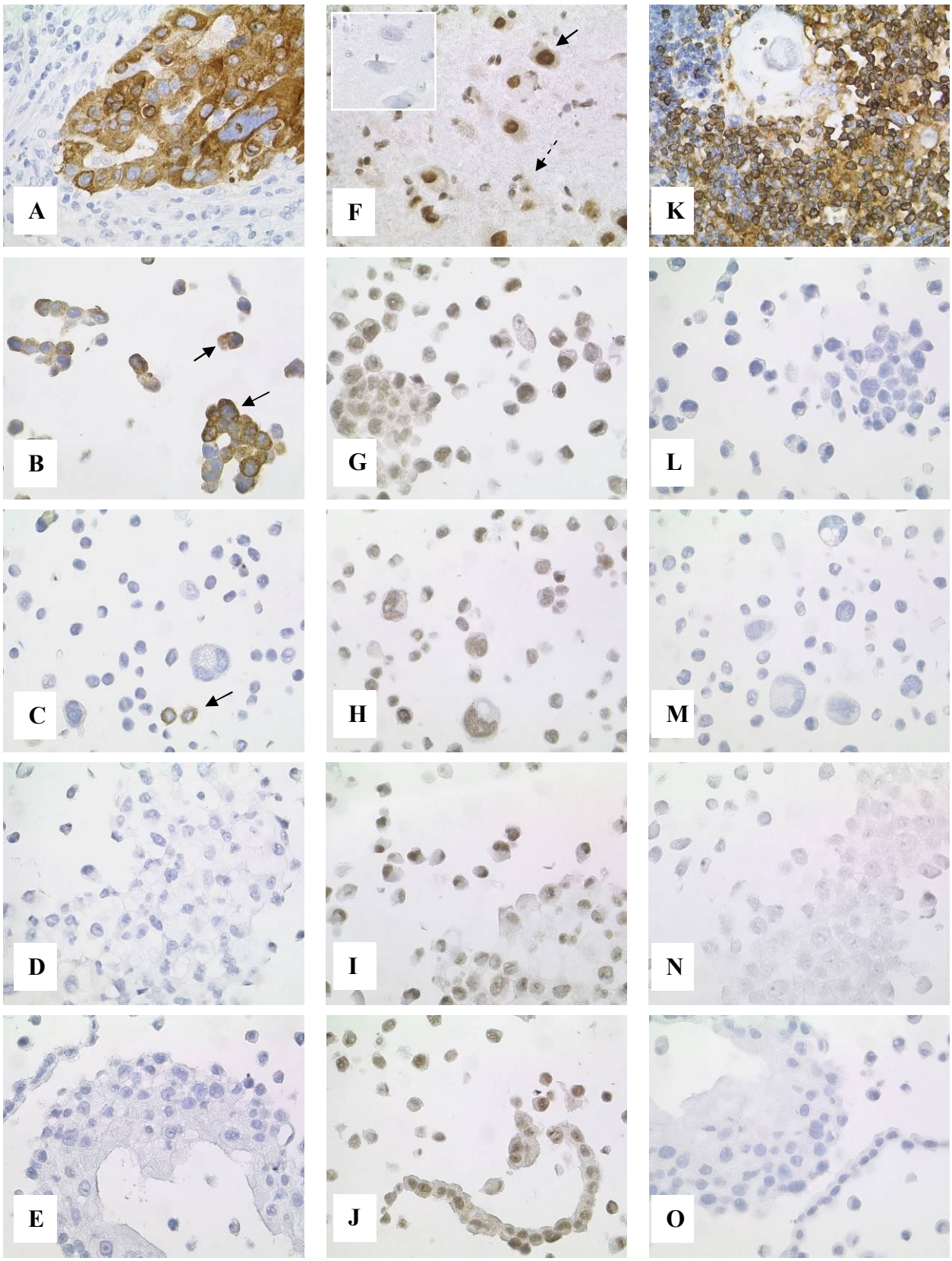

Figure 3.3 Representative immunostaining of FGFR2, TOX3 and LSP1 in breast cancer and non-tumor cell lines (DAB×400) **A** to **E:** FGFR2 **A:** Positive control; **B:** moderate staining of FGFR2 in the cytoplasm of T47D cells (↑); **C:** <1% cells showed positive in MDA-MB-453 (↑); **D:** no staining in MDA-MB-231; **E:** no staining in non-tumor cell line (MCF12A). **F** to **J:** TOX3 When strong and moderate positive staining of

TOX3 was presented in both neuron cells (**F**, ↑) and gliocytes (**F**, dashed ↑) as positive control, all the 13 breast cancer and non-tumor cell lines showed nuclear staining for TOX3 (The small square area in **F** was shown as isotype control for TOX3). **K** to **O:** LSP1 While thymus tissue showed strong positive for LSP1 (**K**), no staining was detected in any cell line. (**B, G, L:** T47D, non-TNC; **C, H, M:** MDA-MB-453 (HER2-overexp.); **D, I, N:** MDA-MB-231 (TNC)**; E, J, O:** MCF12A (non-tumor, TNC)).

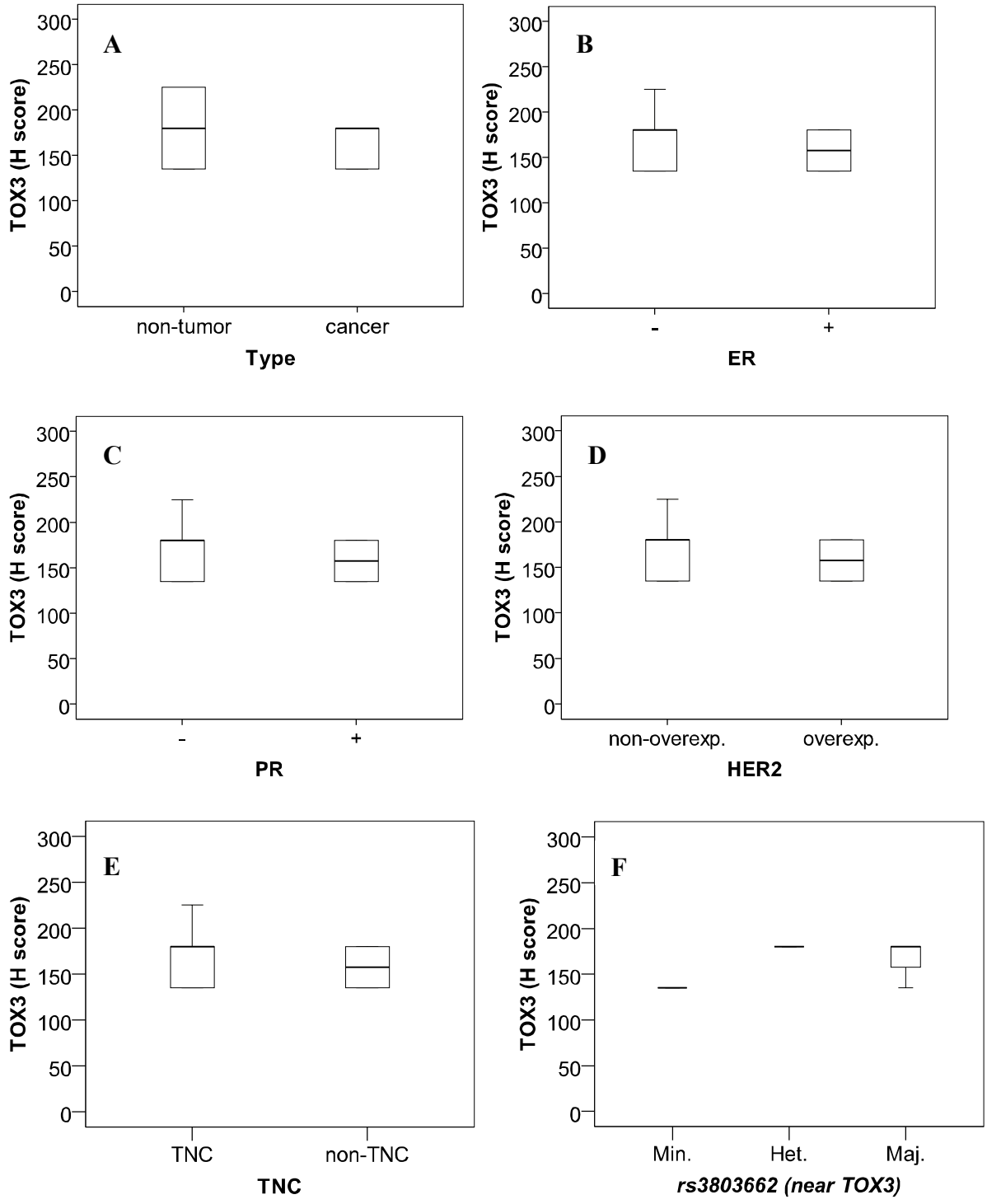

Figure 3.4 Expression of TOX3 in 5 non-tumorous and 8 cancerous breast cell lines and the associations

with ER, PR and HER2 statuses and the SNP near *TOX3* locus (median and IQR of H score, Mann-Whitney *U* test) The median (min, max) of TOX3 H score was 163.13(135, 180) in breast cancer cell lines and 180.00(135, 225) in non-tumorous breast cell lines. No significant difference was confirmed between breast cancer and non-tumorous cell lines by Mann-Whitney *U* test (P=0.474; **A**). When ER, PR, HER2 statuses and TNC were considered, no difference was revealed (P=0.592, 0.592, 0.592 and 0.402, respectively; **B**, **C**, **D**, **E**). When variants at rs3803662 were considered, no statistical analysis could be done because the sample size was too small (**F**).

3.2 Potential differences in several characteristics of Inv-BC samples between Group 1 and Group 2

As our Inv-BC tissue samples in Group 1 were from a 1st-degree family history confirmed collection during 1981-1996 and Inv-BC cases in Group 2 were from a family history unselected collection from 2003-2007, potential heterogeneity of tissue treatments and tumor characteristics between these two groups might influence the subsequent comparison. Therefore, we first compared several important features, which would be combined subsequent for comparison analyses of Group 1 with Group 2. The distant metastasis (pM) could not be compared because only 2 cases provided information about pM in Group 1.The results showed no significant differences in pT, lymph node involvement (pN), tumor grade (G), histological subtype, ER or PR statuses, HER2 overexpression of Inv-BC or TNBC phenotype between these two groups. The only difference was that the age at first diagnosis of breast cancer in Group 1 was younger than that in Group 2 (with the median age of 48.0 years and 66.50 years, respectively; P=0.002 with Fisher's exact test and P=0.000 with Mann Whitney *U* test) (Table 3.2).

Based on these statistical results, we could combine Inv-BC cases in Group 1 and Group 2 into a third Inv-BC sample group which contained 110 Inv-BCs (including at least 39 familial Inv-BC cases and 48 sporadic ones) for following analyses. Overall, this combined group included 17.8% (19/107) of pT3 and pT4, 49.5% (51/103) of LN involved (pN1~N3, LN+), 3.2% (2/62) of distant metastasized (pM1), 27.3% (30/110) of G3, 80% (88/110) of ER+, 72.2% (78/108) of PR+, 7.4% (8/108) of HER2 overexpressed (HER2-overexp.) tumors and 17.3% (19/110) were defined as TNBC.

Table 3.2 Comparison of several characteristics of Inv-BC between Group1 and Group2 (Fisher's exact test)

Characteristic	N1	Group 1 (n,%)	N2	Group 2 (n,%)	N	Total (n, %)	*P*-value[a]
Age of first BC	30		78		108		*0.002**
median(min, max) years		48.0 (28, 71)		66.5 (33, 85)		64.5 (28, 85)	
<40 years		7 (23.3)		2 (2.6)		9 (8.3)	
≥40 years		23 (76.7)		76 (97.4)		99 (91.7)	
pT	31		79		110		0.389
pT 1,2		25 (80.6)		63 (79.7)		88 (80)	
pT 3,4		3 (9.7)		16 (20.3)		19 (17.3)	
unknown		3 (9.7)		0		3 (2.7)	
LN involvement (pN)	31		79		110		0.112
-		9 (29.0)		43 (54.4)		52 (47.3)	
+		16 (51.6)		35 (44.3)		51 (46.4)	
unknown		6 (19.4)		1 (1.3)		7 (6.4)	
Grade	31		79		110		0.152
1,2		26 (83.9)		54 (68.4)		80 (72.7)	
3		5 (16.1)		25 (31.6)		30 (27.3)	
Histological subtype	31		79		110		0.645
IDC		24 (77.4)		46 (58.2)		70 (63.6)	
ILC		4 (12.9)		12 (15.2)		16 (14.5)	
MEC		0 (0.0)		4 (5.1)		4 (3.6)	
MUC		1 (3.2)		5 (6.3)		6 (5.5)	
NEC		0 (0.0)		5 (6.3)		5 (4.5)	
CRIBC		1 (3.2)		3 (3.8)		4 (3.6)	
Others		1 (3.2)		4 (5.1)		5 (4.5)	
ER	31		79		110		1.000
-		6 (19.4)		16 (20.3)		22 (20.0)	
+		25 (80.6)		63 (79.7)		88 (80.0)	
PR	31		79		110		0.637
-		9 (29.0)		21 (26.6)		30 (27.3)	
+		20 (64.5)		58 (73.4)		78 (70.9)	
unknown		2 (6.5)		0		2 (1.8)	
HER2 overexpression	31		79		110		0.439
no		29 (93.5)		71 (89.9)		100 (90.9)	
yes		1 (3.2)		7 (8.9)		8 (7.3)	
unknown		1 (3.2)		1 (1.3)		2 (1.8)	
TNBC	31		79		110		0.781
no		25 (80.6)		66 (83.5)		91 (82.7)	
yes		6 (19.4)		13 (16.5)		19 (17.3)	

[a]: The *P*-value was calculated excluding tumors without information.

*: with statistical significance

3.3 General expression patterns of FGFR2, TOX3 and LSP1 in Inv-BC, intraductal papilloma, fibroadenoma and non-tumorous breast lesions

3.3.1 Expression pattern of FGFR2

Although FGFR2 is a transmembrane protein, several recent articles have reported a nuclear localization of FGFR2 in normal and cancerous breast tissues [88-91]. However, from the results of cell lines shown above, we did not detect a nuclear localization of FGFR2. For this reason, we first examined the expression pattern of FGFR2 in 109 cancerous and 38 benign breast tissues (excluding 3 samples with tissue-losing). Histologically, in both Inv-BCs and benign breast lesions, a cytoplasmic staining pattern of FGFR2 was observed in nearly all positive cases and obviously surrounded the nucleus (perinuclear pattern) in many cases. Staining on the plasma membrane could also be seen in some cases, although it seemed more like a localization at the intracellular side of the cell membrane. No certain nuclear staining was observed. Either, no staining was detected in myoepithelial or non-epithelial cells either (Figure 3.5). Numerically, positive expression of FGFR2 (+ to +++) was observed in 75.2% (82/109) Inv-BC, 63.6% (7/11) intraductal papilloma, 66.7% (10/15) fibroadenoma and 73.5% (9/12) non-tumor breast tissues (P=0.722, Table 3.3). The proportion of high-level expression (strong positive, +++) was determined in 22.9% of Inv-BCs but 0% of the three benign lesions (P=0.013). However, possibly different expression level of FGFR2 could not be further figured out within these 4 lesions by non-parameter comparison, although intraductal papilloma seemed generally to show the lowest expression level (with H score median of 20, Figure 3.6).

Table 3.3 Expression and high-level expression of FGFR2 in Inv-BC and benign breast lesions (Fisher's exact test)

Lesion	N	FGFR2 (expression, n)		*P*-value	FGFR2 (high-level, n)		*P*-value
		negative	positive (%)		low	high (%)	
	147			0.722			*0.013**
Inv-BC		27	82 (75.2)		84	25 (22.9)	
Intraductal papilloma		4	7 (63.6)		11	0 (0.0)	
Fibroadenoma		5	10 (66.7)		15	0 (0.0)	
Non-tumorous lesions		3	9 (75.0)		12	0 (0.0)	

*: with statistical significance

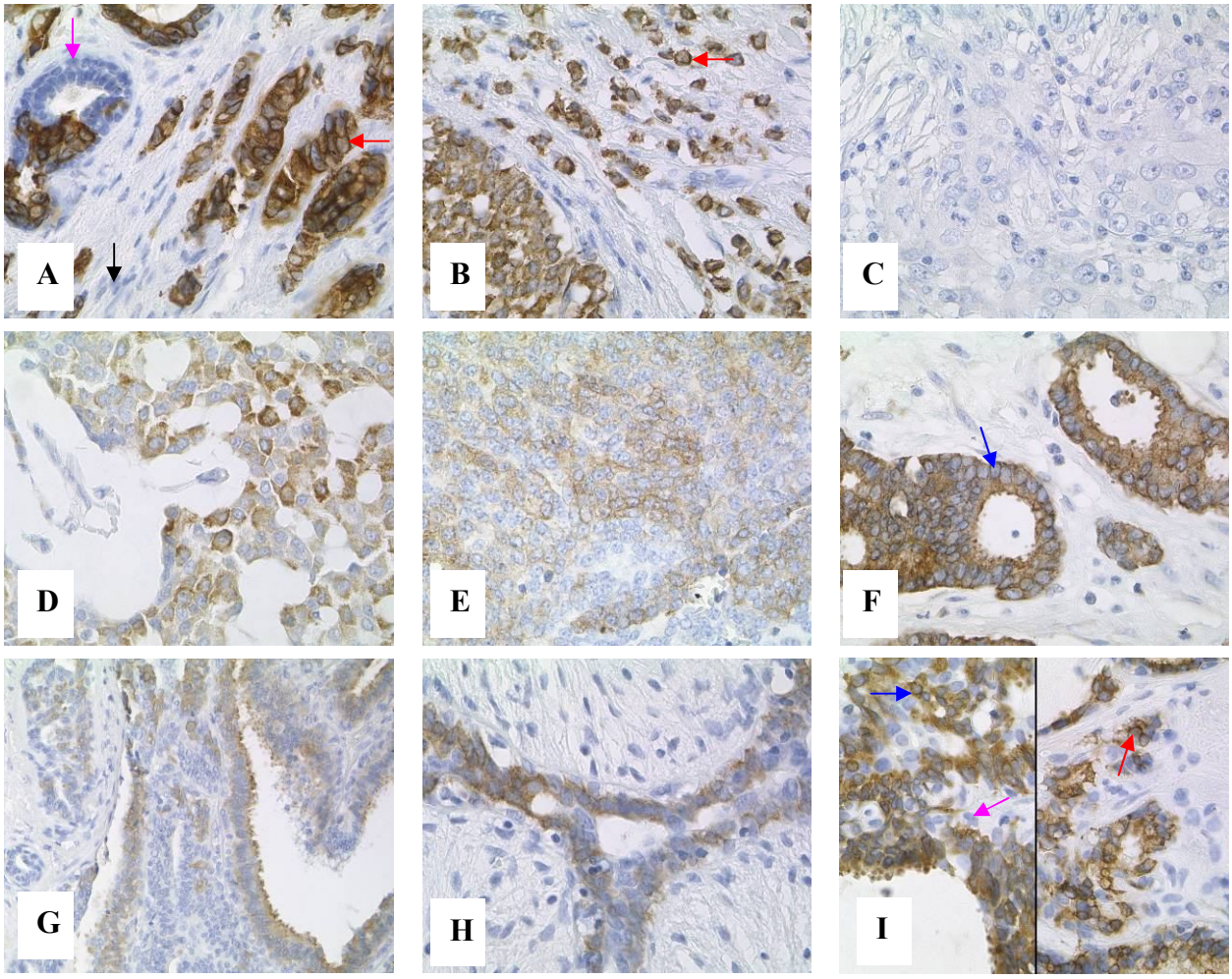

Figure 3.5 Representative immunostaining of FGFR2 in Inv-BC and benign breast lesions (DAB ×400)

A to **E:** Immunostaining of FGFR2 was observed in most kinds of Inv-BCs at varying levels, except for 4 cases of MEC, which were totally unstained for FGFR2. **A:** IDC; **B:** ILC; **C:** MEC; **D:** MUC; **E:** NEC; **F:** CRIBC; **G:** intraductal papilloma (×200); **H:** fibroadenoma; **I:** non-tumorous breast lesion (from 2 cases). Cytoplasmic and plasma membrane's (red ↑) and perinuclear (blue ↑) staining was observed in cancerous and benign ductal/lobular epithelial cells but not in myoepithelial (pink ↑) or mesenchymal (black ↑) cells.

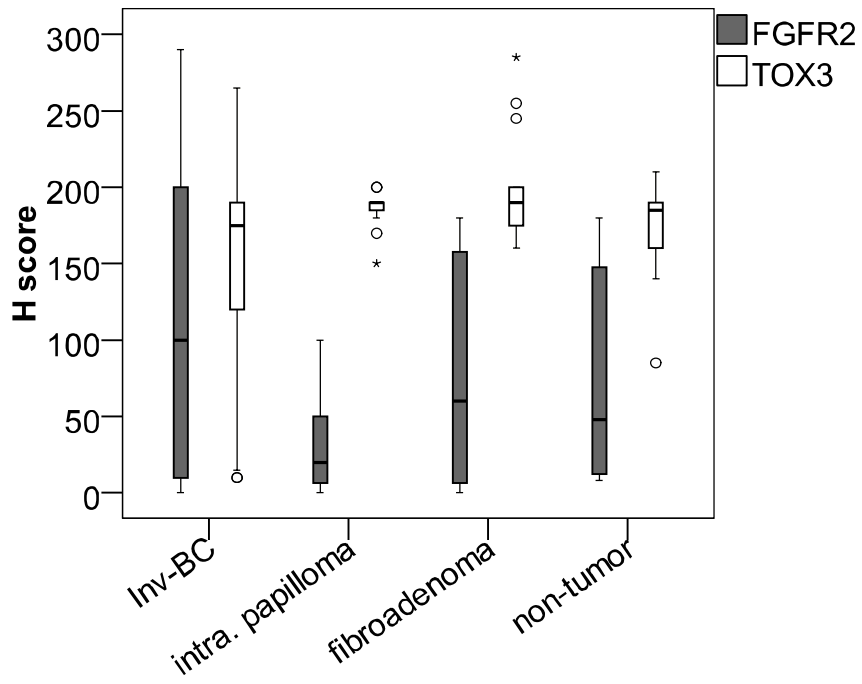

Figure 3.6 Expression levels of FGFR2 and TOX3 in Inv-BC and benign breast lesions (median and IQR of H score, Kruskal-Wallis and Nemenyi tests) When only Inv-BC and the other two kinds of benign breast tumors were compared, expression of FGFR2 in intraductal papilloma (shown as intra. papilloma) was lower than in Inv-BC but with a marginally significant *P*-value (*P*=0.05). When non-tumorous lesions (shown as non-tumor) were also included in the comparison, no significance was further confirmed. Significantly lower expression of TOX3 in Inv-BC than in fibroadenoma was determined (*P*=0.04). ("○" and "*": extreme values)

3.3.2 Expression pattern of TOX3

TOX3 has been reported as a nuclear protein [77]. However, no visualized data has been available in publications until now. By using a commercial polyclonal antibody, positive immunostaining of TOX3 was observed in the nuclei of nearly all cases in all kinds of tissues. While strong and moderate staining of TOX3 was respectively presented in neuron cells and gliocytes in the normal human brain tissues (used as a positive control), weak to strong staining, mainly moderate, was observed in both benign ductal/lobular epithelial cells, myoepithelial cells and also benign/malignant tumorous epithelial cells (Figure 3.7). It seemed that the staining in those epithelial cells with apocrine metaplasia was always somewhat more predominant. Mesenchymal fibroblasts, endothelial and smooth muscle cells of the vessels and also eccrine glands, sebaceous glands and skeletal muscle cells surrounding the tumor lesions in several cases were also moderately positive for TOX3. Numerically, positive expression of TOX3 (+ to +++) in epithelial cells was determined in 98.1% of Inv-BC and in 100% of other 3 lesions. The only

two so-called negative Inv-BCs (one from Group1, ILC; another one from Group2, neuroendocrine carcinoma (NEC), without available information about family history), showed very faint and ambiguous staining in about 10% tumor cells and therefore were regarded as negative. As seemingly weaker staining in Inv-BCs was noticed, we paid special attention to the weak expression of TOX3 and a overall difference was revealed within these 4 lesions (*P*=0.035, Table 3.4). Furthermore, Kruskal-Wallis test of H score also showed significance with a *P*-value of 0.013 within these 4 lesions and lower expression in Inv-BC than in fibroadenoma was determined (*P*=0.04, Figure 3.6). However, no difference of high-level (+++) expression was observed within these lesions.

Table 3.4 Weak and high-level expressions of TOX3 in Inv-BC and benign breast lesions

(Fisher's exact test)

Lesion	N	TOX3 (weak, n)		*P*-value	TOX3 (high-level, n)		*P*-value
		weak (%)	non-weak		low	high (%)	
	146			*0.035**			0.587
Inv-BC		24(22.4)	83		95	12 (11.2)	
Intraductal papilloma		0(0.0)	11		11	0 (0.0)	
Fibroadenoma		0(0.0)	16		13	3 (18.8)	
Non-tumor disease		1(8.3)	11		11	1 (8.3)	

*: with statistical significance

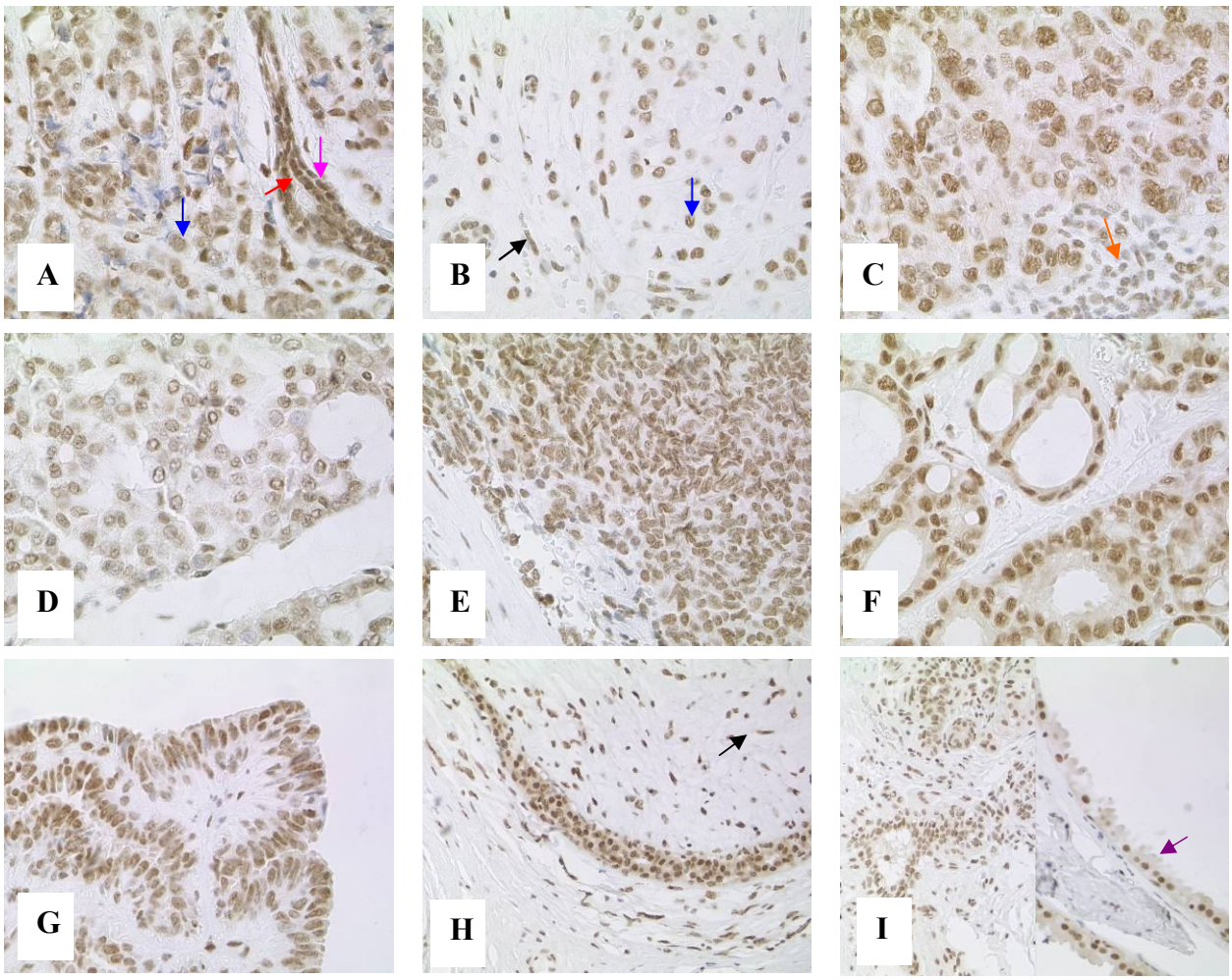

Figure 3.7 Representative immunostaining of TOX3 in Inv-BC and benign breast lesions (DAB ×400)
A to **E:** Clear nuclear immunostaining of TOX3 was observed in all kinds of Inv-BC and benign breast lesions. **A:** IDC; **B:** ILC; **C:** MEC; **D:** MUC; **E:** NEC; **F:** CRIBC; **G:** intraductal papilloma; **H:** fibroadenoma (×200); **I:** non-tumorous breast lesion (×200, from 2 cases). Tumor cells were marked with a blue arrow (↑); normal ductal epithelial cells were marked with a red arrow (↑); the pink arrow (↑) highlighted myoepithelial cells; the black arrow (↑) marked out the stromal fibroblasts and fibrocytes; the orange arrow (↑) represented lymphocyte; the purple arrow (↑) pointed out the epithelial cells with apocrine metaplasia.

3.3.3 Expression pattern of LSP1

Although the expression of LSP1 has been shown to be restricted in hematopoietic cells but not in epithelial cells, the possible risk association of the SNP at *LSP1* locus with breast cancer might indicate a special role of LSP1 in breast cancer. Moreover, we did observed one IDC case (54y; G3; pT_1N_0) in 22 Inv-BCs included in the TMA sample collection, which was used for pre-test, that showed moderate cytoplasmic staining for LSP1 in 20% to 60% tumor cells (Figure 3.8). So we further detected the expression of LSP1 in the 147 Inv-BC and benign breast tissues

(two samples with tissue-losing were excluded). While thymocytes (as inter positive control) in normal thymus and leukocytes scattering in all types of breast tissues (as inner positive control) presented strong staining in cytoplasm and intracellular side of cell membrane for LSP1, only 2 Inv-BC cases were found to present ambiguous light brown staining (Figure 3.8). One of them was diagnosed as ILC (59y; G2; $pT_1N_0M_0$; ER+, PR+, HER2 3+) which showed weak cell membrane staining in <10% of scattered tumor cells. The other one was diagnosed as MEC (68y; $pT_1N_0M_0$; TNBC) which showed weak cell membrane staining and cytoplasmic faint staining in about 15% of tumor cells. IHC of these two cases were repeated once, similar staining patterns were obtained. Interestingly, the tumor cells of all these three "positive" cases had large nuclei and predominant nucleoli but without predominant infiltration of inflammatory cells into the epithelial cell sheets, which is a typical appearance in lymphoepithelial carcinoma, especially in the "positive"-case included in the TMA. In all the other tumorous or non-tumorous breast tissues, no positive staining was observed in any of the epithelial cells. Therefore, no further statistical analysis was done upon the expression of LSP1.

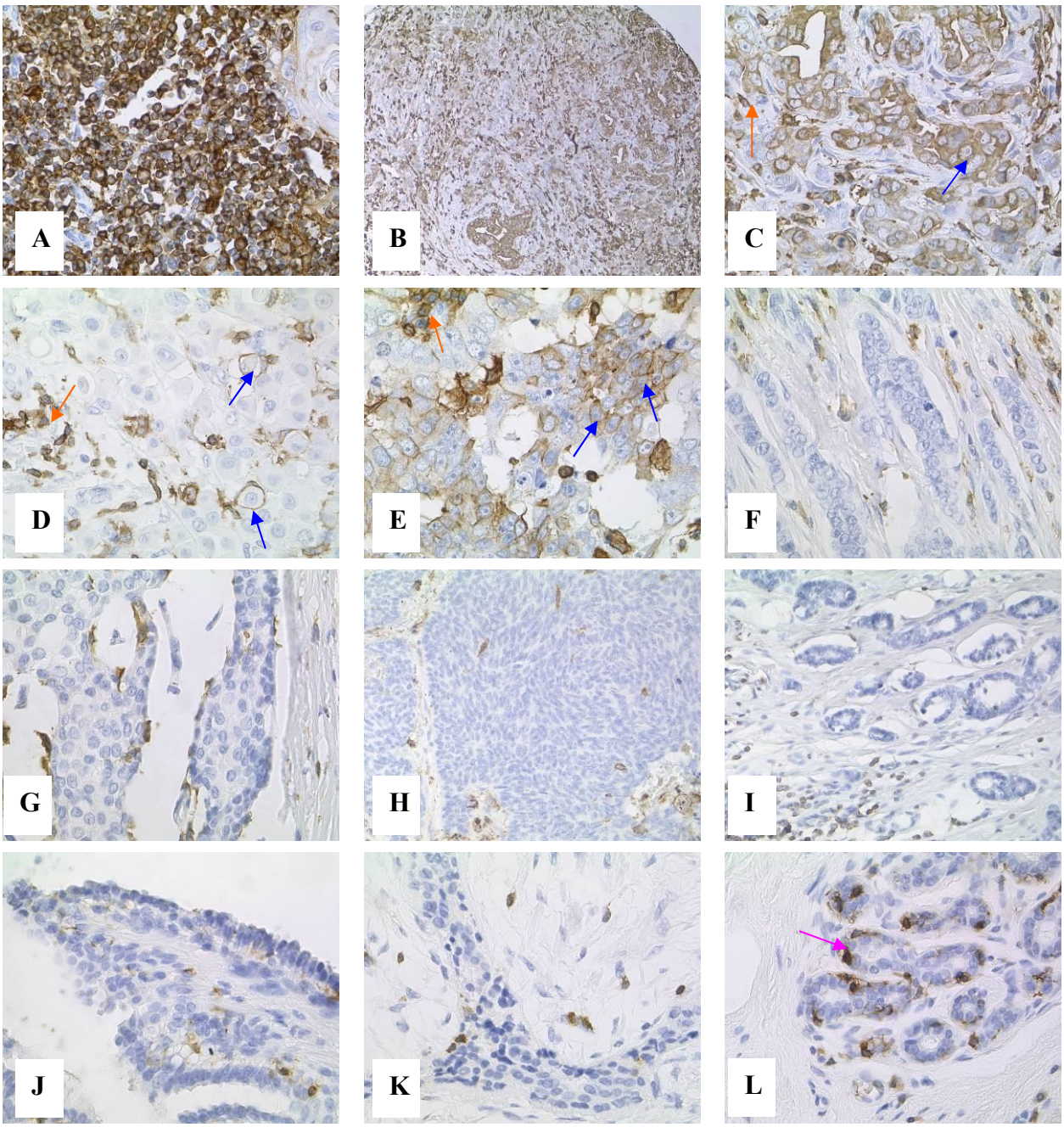

Figure 3.8 Representative immunostaining of LSP1 in Inv-BC and benign breast lesions (DAB ×400)

A: Strong staining for LSP1 on the inner side of plasma membrane and in the cytoplasm of normal thymocytes (positive control). **B** (×100) and **C** (×400)**:** One case of IDC in the TMA collection showed cytoplasmic staining for the current LSP1 antibody. One ILC (**D**) and one MEC (**E**) in Group2 presented ambiguous staining for LSP1. **F:** IDC; **G:** MUC; **H:** NEC (×200); **I:** CRIBC (×200); **J:** intraductal papilloma; **K:** fibroadenoma; **L:** non-tumorous breast lesion. "Positive" tumor cells were marked with a blue arrow (↑); the orange arrow (↑) highlighted lymphocytes which were regarded as inner positive control; the pink arrow (↑) pointed out the lymphocytes infiltrating in the lobular epithelia.

3.4 Associations of FGFR2 expression in Inv-BC with clinical and pathological characteristics

When clinical and pathological characteristics of Inv-BC were examined, higher positive proportion of FGFR2 in subsets of elder age (≥40 years), lower tumor grade (G1, 2), ER+ and PR+, as well as in non-TNBCs (Table 3.5) was revealed. As shown above, different age distributions between Group 1 and Group 2 might influence the combined analysis of FGFR2 expression with patients' age, replicated analyses were executed separately in these two groups. However, no significance was confirmed in either group (P=0.053 in Group 1, P=0.430 in Group2). Thus, we could not confirm that expression of FGFR2 was associated with age. A higher high-level expression proportion of FGFR2 (with 200 < H score ≤ 300) was demonstrated in ER+ and non-TNBC than ER- (P=0.023) and TNBC (P=0.006) subsets (Table 3.5). Nonparametric tests of H score further demonstrated a significantly lower expression level of FGFR2 in G3 tumors and a predominantly lower level in ER-, PR- and TNBC ones (Figure 3.9 and Figure 3.10 A, B, D). In fact, FGFR2 was entirely undetectable in 89.5% (17/19) of TNBCs, and the other 2 cases only showed very weak expression with an H score of <50. Because PR is well-known as an ER-dependent protein, further fine comparisons within 4 different ER/PR phenotypes indicated that the ER+ might be the key event for FGFR2 expression, because the expression levels were similar in ER+/PR+ and ER+/PR- subsets (Figure 3.10 E). According to these findings, negative correlation of FGFR2 expression (- to +++) with tumor grade (G1 to G3) (r_s= -0.269, P=0.005) and positive correlations of FGFR2 H score with ER (r_s= 0.513, P=0.000) and PR (r_s= 0.414, P=0.000) Allred's scores were further defined by Spearman's rank correlation analysis. Besides, as shown in Figure 3.10 E (↑), we still noticed that 1/21 ER-/PR- case showed high-level-expression for FGFR2, while nearly all the others in this subset lacked FGFR2 expression. When examining the primary data, we found that this case showed HER2 overexpression. However, a comparison of the two subsets with different expression status of HER2 did not reveal an association between the expression of FGFR2 and HER2 overexpression (Table 3.5 and Figure 3.10 C). Nor was more significant difference between subsets defined by family history, menopausal status, HRT history, malignant tumor history, bilateral, multicentric/multifocal, pT, LN involvement or distant metastasis observed.

Within different histological subtypes, FGFR2 was expressed in 70% (49/109) of IDC, 100 % of (16/16) ILC, 100% (5/5) of MUC, 80% (4/5) of NEC and 100% (4/4) of CRIBC cases but it was entirely negative in 100% (4/4) of MEC ones which were TNBCs (Table 3.5, Figure 3.11). Kruskal-Wallis and further Nemenyi tests demonstrated a significantly higher expression level of FGFR2 in ILC than in IDC and in MEC cases (Figure 3.11).

Table 3.5 Associations of FGFR2 expression in Inv-BC with clinical and pathological characteristics (Fisher's exact test)

Characteristic	N	FGFR2 (expression, n)		*P*-value	FGFR2 (high-level, n)		*P*-value
		negative	positive (%)		low	high (%)	
Group	109			1.000			0.801
1		8	23 (74.2)		23	8 (25.8)	
2		19	59 (75.6)		61	17 (21.8)	
Age	109			*0.040**			0.681
<40y		5	4 (44.4)		8	1 (11.1)	
≥40y		22	78 (78.0)		76	24 (24.0)	
Family history	86			0.626			0.799
-		14	33 (70.2)		36	11(23.4)	
+		9	30 (76.9)		31	8(20.5)	
Menopause$^{\Delta}$	76			0.452			0.690
Pre- (or <50y)		4	7 (63.6)		8	3 (27.3)	
Post- (or ≥50y)		15	50 (76.9)		52	13 (20.0)	
HRT$^{\Delta}$	58			1.000			0.433
-		11	33 (75.0)		37	7 (15.9)	
+		4	10 (71.4)		10	4 (28.6)	
Malignant tumor history$^{\Delta\blacktriangle}$	54			0.418			0.667
-		13	32 (71.1)		37	8 (17.8)	
+		1	8 (88.9)		7	2 (22.2)	
Bilateral$^{\Delta}$	78			0.671			0.362
no		18	52 (74.3)		56	14 (20.0)	
yes		1	7 (87.5)		5	3 (37.5)	
Multicentric/multifocal$^{\Delta}$	78			0.188			1.000
no		19	51 (72.9)		55	15 (21.4)	
yes		0	8 (100.0)		6	2 (25.0)	
pT	106			0.231			1.000
pT1, 2		23	64 (73.6)		66	21 (24.1)	
pT3, 4		2	17 (89.5)		15	4 (21.1)	
LN involvement	102			1.000			0.816
-		13	38 (74.5)		38	13 (25.5)	
+		12	39 (76.5)		40	11 (21.6)	

pM$^{\Delta}$		62			1.000		0.378
	M0		16	44(73.3)		48	12(20.0)	
	M1		0	2(100.0)		1	1(50.0)	
Tumor grade		109			*0.000**		0.203
	G1, 2		12	67 (84.8)		58	21 (26.6)	
	G3		15	15 (50.0)		26	4 (13.3)	
ER		109			*0.000**		*0.023**
	-		19	3 (13.6)		21	1 (4.5)	
	+		8	79 (90.8)		63	24 (27.6)	
PR		107			*0.000**		0.447
	-		20	10 (33.3)		25	5 (16.7)	
	+		7	70 (90.9)		58	19 (24.7)	
HER2 overexpression		107			1.000		0.680
	no		25	74 (74.7)		76	23 (23.2)	
	yes		2	6 (75.0)		7	1 (12.5)	
TNBC		109			*0.000**		*0.006**
	no		10	80 (88.9)		65	25 (27.8)	
	yes		17	2 (10.5)		19	0 (0.0)	
BRCA1/2 mutation		16			0.263		0.263
	non-*BRCA1/2*		2	5(71.4)		4	3(42.9)	
	BRCA1		3	2(40.0)		5	0(0.0)	
	BRCA2		0	4(100.0)		2	2(50.0)	

$^{\Delta}$: Information on menopausal status, HRT, malignant tumor history, bilateral BC, multicentric/multifocal BC and pM was only available in some cases in Group 2. Some cases in Group 2 had undergone hysterectomy before natural menopause and no information about inner hormone level was available, so we defined age at ≥50 years as postmenopausal status in such cases.

▲: Nine cases in Group 2 had suffered from malignant tumors of other organs or tissues: 3 cases had colon caner, 1 had endometrial cancer, 1 had skin basal cell carcinoma, 1 had subaxillary non-Hodgkin's lymphoma (NHL), 1 had CIN III, 1 had papillary thyroid carcinoma, and 1 had paranasal squamous cell carcinoma.

*: with statistical significance

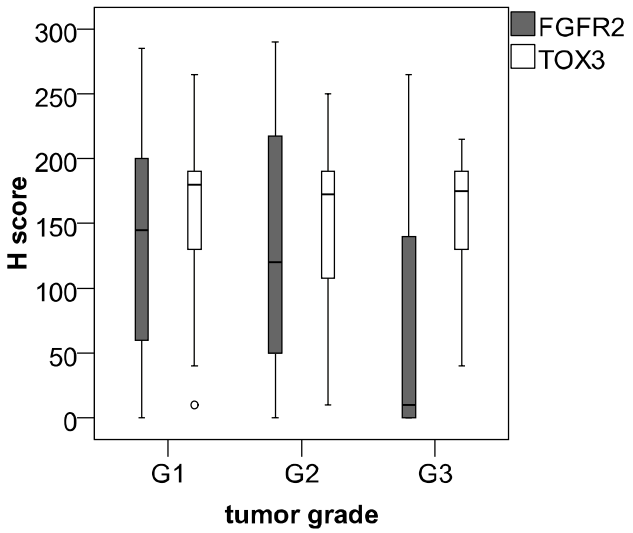

Figure 3.9 Expression levels of FGFR2 and TOX3 in Inv-BC with different tumor grades (median and

IQR of H score, Kruskal-Wallis and Nemenyi tests) Statistical analyses demonstrated a significantly lower expression of FGFR2 in G3 Inv-BCs than in G1 and G2 ones (*P*=0.003). No different expression of TOX3 was revealed within subsets of different tumor grades. ("○": extreme values)

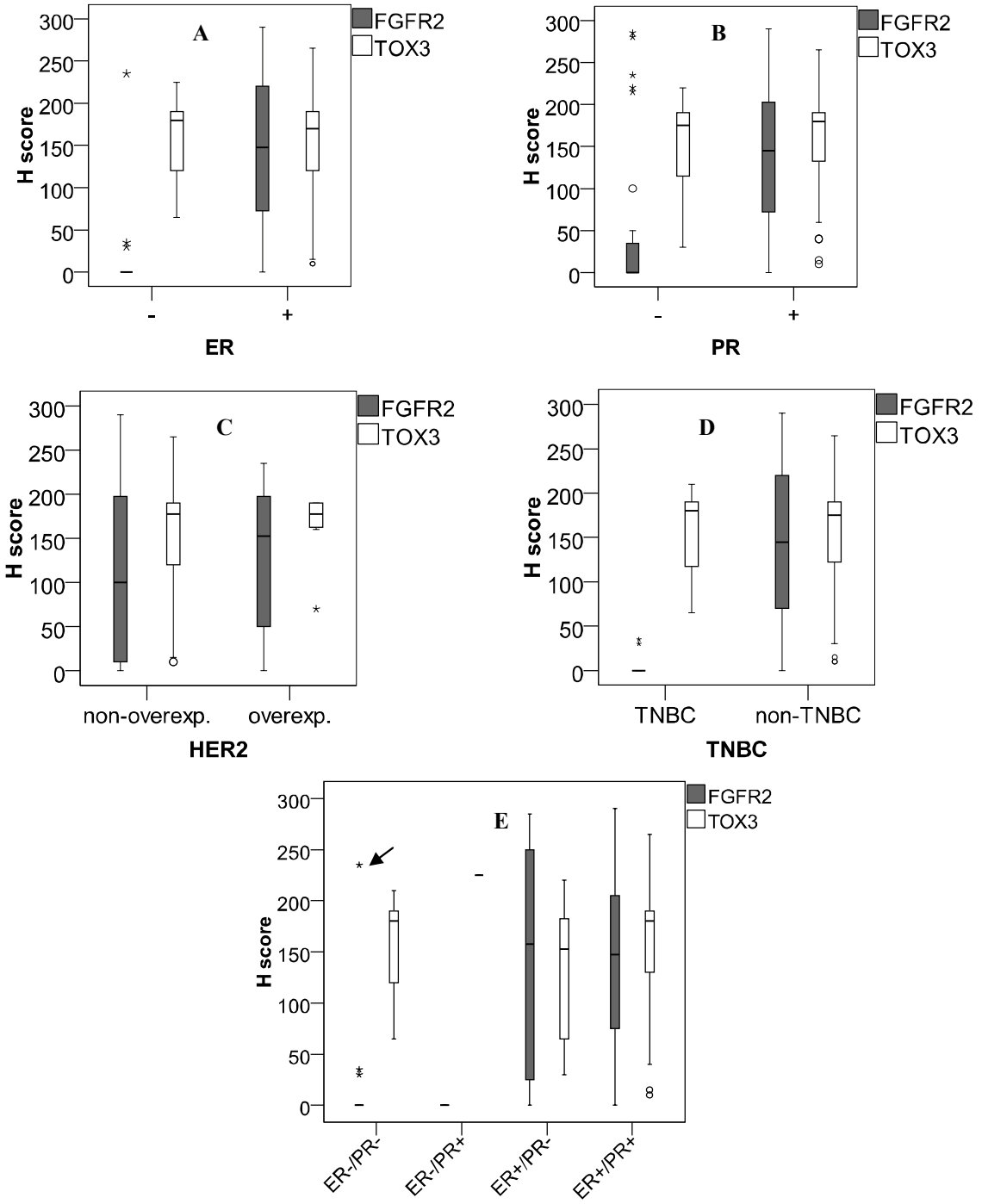

Figure 3.10 Expression levels of FGFR2 and TOX3 in Inv-BCs with different ER, PR and HER2 statuses (median and IQR of H score, Mann-Whitney test or Kruskal-Wallis and Nemenyi tests) **A, B, D:** Predominantly lower expression of FGFR2 was shown in ER-, PR- and TNBC tumors (all with *P*=0.000). **E:** Further, greatly different expression of FGFR2 between ER-/PR- subset and ER+/PR+ (*P*=0.00) or ER+/PR- (*P*=0.00) subsets was confirmed. One case with ER-/PR- and HER2 overexpressed still showed strong expression of FGFR2 (↑), but no difference was detected between HER2 over- and non-overexpression tumors (**C**). No statistical difference was revealed when TOX3 expression was concerned. ("○" and "*": extreme values)

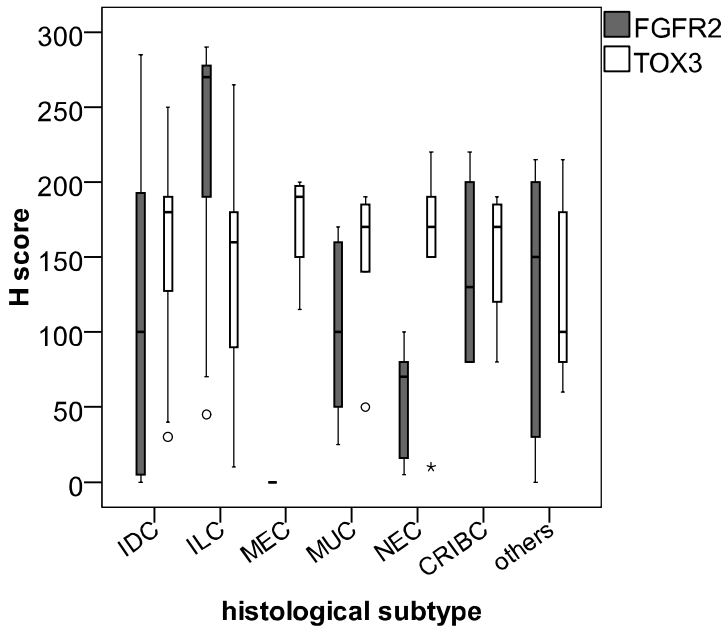

Figure 3.11 Expression levels of FGFR2 and TOX3 in different histological subtypes of Inv-BC (median and IQR of H score, Kruskal-Wallis and Nemenyi tests) Significantly lower expression of FGFR2 in MEC (P=0.01) and IDC (P=0.02) than in ILC was confirmed. Seemingly lower expression of FGFR2 in NEC (only 5 cases) could not be confirmed by these statistical analyses. No different expression of TOX3 was revealed within these subtypes. ("○" and "*": extreme values)

3.5 Associations of TOX3 expression in Inv-BC with clinical and pathological characteristics

Although nearly all Inv-BC tissues showed positive for TOX3, and no difference in TOX3 high-level-expression was revealed between subsets in any of the characteristics, Fisher's exact test indicated a lower expression proportion in familial than sporadic cases (P=0.002, Table 3.6). Mann-Whitney U tests additionally confirmed a weaker expression level of TOX3 in familial Inv-BCs than sporadic ones. Because the samples in Group 1 were obtained from old blocks, and we found that tumor-adjacent normal breast tissues also showed obviously lower expression level of TOX3 in Group 1 than in Group 2 (P=0.000 with Mann-Whitney test), we further compared the expression of TOX3 only in Group 2 with family history. As shown in Figure 3.12, the significance could not be confirmed although the H score in familial tumors seemed lower than in sporadic ones. There were very few familial cases in Group 2, which could cause a false negative. Therefore, more familial cases with equal quality are still needed to clarify the possibility of lower expression of TOX3 in familial Inv-BC. However, a significantly lower

expression level of TOX3 was revealed in cases positive for malignant tumor history than negative ones in Group 2 (Figure 3.13). No additional difference was observed with respect to other clinical and pathological characteristics (Table 3.6, Figure 3.9 to Figure 3.11).

Table 3.6 Associations of TOX3 expression in Inv-BC with clinical and pathological characteristics (Fisher's exact test)

Characteristics	N	TOX3 (weak, n)		*P* value	TOX3 (high-level, n)		*P* value
		weak (%)	non-weak		low	high(%)	
Group	107			*0.001**			0.742
1		14 (54.2)	17		27	4 (12.9)	
2		10 13.2)	66		68	8 (10.5)	
Age	107			0.112			1.000
<40y		4 (44.4)	5		8	1 (11.1)	
≥40y		20 (20.4)	78		87	11 (11.2)	
Family history	84			*0.002**			1.000
-		5 (11.1)	40		40	5 (11.1)	
+		16 (41.0)	23		35	4 (10.3)	
Menopause$^{\Delta}$	74			1.000			0.238
Pre- (or <50y)		1 (10.0)	9		8	2 (20.0)	
Post- (or ≥50y)		9 (14.1)	55		59	5 (7.8)	
HRT$^{\Delta}$	56			1.000	37	5 (11.9)	1.000
-		5 (11.9)	37		13	1 (7.1)	
+		2 (14.3)	12				
Malignant tumor history$^{\Delta\blacktriangle}$	52			0.057			0.574
-		3 (7.0)	40		37	6 (14.0)	
+		3 (33.3)	6			0 (0.0)	
Bilateral$^{\Delta}$	76			1.000			1.000
no		9 (13.2)	59		61	7 (10.3)	
yes		1 (12.5)	7		7	1 (12.5)	
Multicentric/multifocal$^{\Delta}$	76			0.587			1.000
no		10 (14.7)	58		61	7 (10.3)	
yes		0 (0.0)	8		7	1 (12.5)	
pT	104			0.527			0.432
pT1, 2		17 (19.8)	69		77	9 (10.5)	
pT3, 4		5 (27.8)	13		15	3 (16.7)	
LN involvement	100			0.329			0.758
-		8 (16.3)	41		43	6 (12.2)	
+		13 (25.5)	38		46	5 (9.8)	
pM$^{\Delta}$	60			1.000			0.192
M0		7 (12.1)	51		53	5 (8.6)	
M1		0 (0.0)	2		1	1 (50.0)	

Tumor grade	107			0.603			1.000
G1, 2		19 (24.4)	59		69	9 (11.5)	
G3		5 (17.2)	24		26	3 (10.3)	
ER	107			0.776			0.708
-		4 (18.2)	18		19	3 (13.6)	
+		20 (23.5)	65		76	9 (10.6)	
PR	105			0.603			1.000
-		7 (24.1)	22		26	3 (10.3)	
+		15 (19.7)	61		67	9 (11.8)	
HER2 overexpression	105			0.682			0.592
no		22 (22.7)	75		85	12 (12.4)	
yes		1 (12.5)	7		8	0 (0.0)	
TNBC	107			1.000			1.000
no		20 (22.7)	68		78	10 (11.4)	
yes		4 (21.1)	15		17	2 (10.5)	
BRCA1/2 mutation	16			1.000			0.250
non-*BRCA1/2*		4(57.1)	3		7	0(0.0)	
BRCA1		3(60.0)	2		5	0(0.0)	
BRCA2		2(50.0)	2		3	1(25.0)	

[Δ]: Information on menopausal status, HRT, malignant tumor history, bilateral BC, multicentric/multifocal BC and pM was only available in some cases in Group 2. Some cases in Group 2 had undergone hysterectomy before natural menopause and no information about inner hormone level was available, so we defined age at ≥50 years as postmenopausal status in such cases.

[▲]: Nine cases in Group 2 had suffered from malignant tumors of other organs or tissues: 3 cases had colon caner, 1 had endometrial cancer, 1 had skin basal cell carcinoma, 1 had subaxillary non-Hodgkin's lymphoma (NHL), 1 had CIN III, 1 had papillary thyroid carcinoma, and 1 had paranasal squamous cell carcinoma.

*: with statistical significance

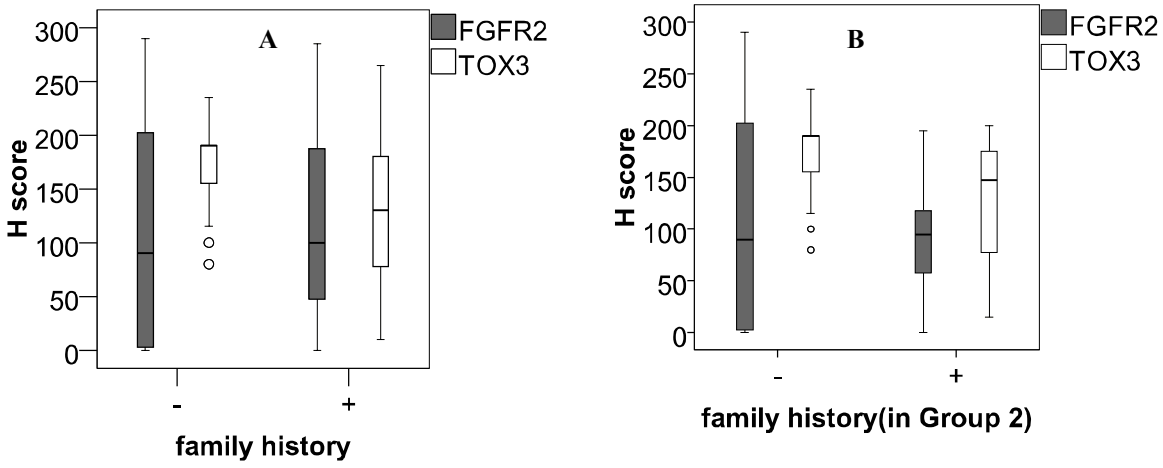

Figure 3.12 Expression levels of FGFR2 and TOX3 in Inv-BCs with different statuses of family and malignant tumor history (median and QR of H score, Mann-Whitney *U* test) Significantly lower expression of TOX3 was observed in familial Inv-BCs than sporadic ones when all Inv-BCs in Group 1 and Group 2 were conbined (P=0.001, **A**). However, when comparison was only done in cases in Group 2, such significance could not be repeated (P=0.068, **B**). No differences were seen when expression of FGFR2 was

analyzed. ("○": extreme values)

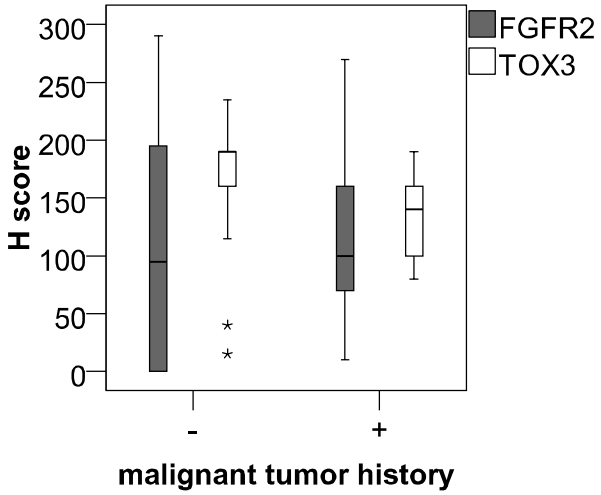

Figure 3.13 Expression levels of FGFR2 and TOX3 in Inv-BCs with different statuses of malignant tumor history (only including cases in Group 2, median and QR of H score, Mann-Whitney *U* test) Significantly lower expression of TOX3 was observed in cases positive for malignant tumor history than in those negative for (P=0.004) ("*": extreme values)

3.6 Expression statuses of FGFR2 and TOX3 in Inv-BC tissues of *BRCA1/2* germline mutation carriers

Because the breast cancer risk associations of *FGFR2* and *TOX3* loci might also be altered by *BRCA1/2* mutation status except for family history and ER status, the expression of FGFR2 and TOX3 was also compared within *BRCA1/2* germline mutation and non-mutation carriers. As shown in Table 3.7, there were 5 cases with *BRCA1* and 4 cases with *BRCA2* germline mutations, and 7 cases without *BRCA1/2* germline mutation in Group 1. By Fisher's exact (Tables 3.5 and 3.6) and Kruskal-Wallis test (Figure 3.14), no significant difference in FGFR2 or TOX3 expression was demonstrated. However, seemingly lower expression of FGFR2 was shown in *BRCA1* mutation carriers than in the other two subsets. Interestingly, 3/5 *BRCA1*, 0/4 *BRCA2* and 1/7 non-*BRCA1/2* mutation carriers were TNBCs. Moreover, these 3 *BRCA1* carriers showing TNBC totally lacked FGFR2 expression, while none of the 4 *BRCA2* carriers were negative for FGFR2, and 2/4 of them even showing high-level expression for FGFR2 (Table 3.7).

These findings were in agreement with earlier studies that *BRCA1* associated tumors are more likely to be triple negative and basal-like breast cancer, while *BRCA2* associated ones are more likely to be luminal breast cancer which shows ER+ and/or PR+. And *BRCA1*-related

tumors are more likely to lack FGFR2 expression, while *BRCA2*- and non-mutation tumors are more likely to present diverse expression of FGFR2 [110].

Table 3.7 FGFR2 and TOX3 expression levels in cases with *BRCA1/2* germline mutation information (n=16)

Fam. ID	Detected Pat. ID	Histological subtype	G	*BRCA1/2*	ER	PR	HER2	TNBC	FGFR2 (H score)	TOX3 (H score)
Fam.30	303	IDC	2	*BRCA1*	-	-	1+	yes	0	150
Fam.31	300(male)	IDC	2	*BRCA2*	+	-	2+	no	50	220
Fam.60	302	ILC	2	*BRCA2*	+	+	1+	no	285	80
	401	IDC	2	*BRCA2*	+	+	1+	no	235	180
	416	IDC	3	*BRCA2*	+	+	2+	no	150	80
Fam.81	202(male)	IDC	3	*BRCA1*	+	+	1+	no	70	100
	305	IDC	3	*BRCA1*	-	-	1+	yes	0	100
Fam.92	302	IDC	2	no	+	+	2+	no	100	40
Fam.105	401	IDC	2	*BRCA1*	-	-	2+	yes	0	180
Fam.1688	209	IDC	2	no	-	-	1+	yes	0	75
Fam.2750	300	(IDC+MUC)	1	*BRCA1*	+	+	1+	no	200	60
Fam.2953	202	ILC	1	no	+	+	1+	no	270	180
Fam.3692	200	IDC	3	no	+	+	2+	no	10	40
Fam.3784	200	IDC	2	no	+	+	n.a.	no	220	150
	202	IDC	2	no	+	+	1+	no	235	90
	300	IDC	2	no	+	+	2+	no	170	180

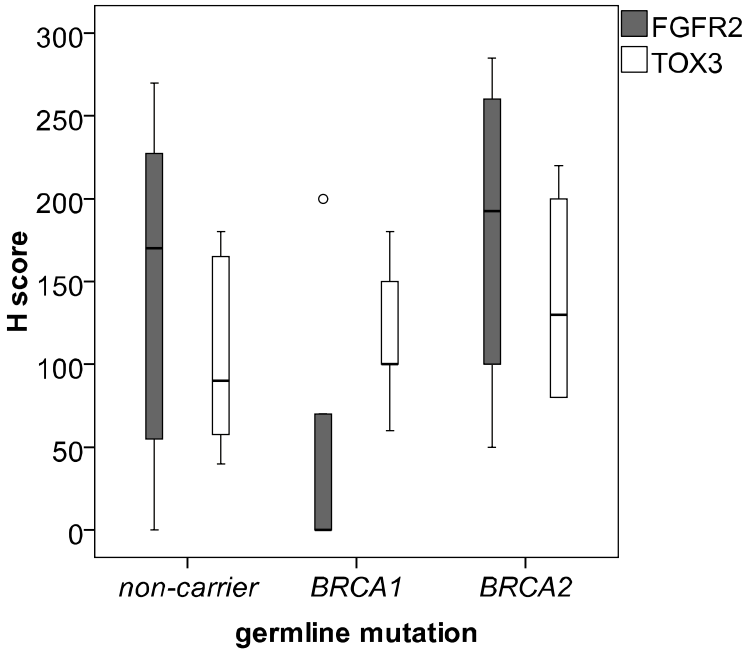

Figure 3.14 Expression levels of FGFR2 and TOX3 in different *BRCA1/2* germline mutation carriers (median and IQR of H score, Kruskal-Wallis test) Although no significant difference was demonstrated by Kruskal-Wallis test (*P*=0.129), expression level of FGFR2 in *BRCA1* mutation carriers seemed much lower than in the other two subsets. Expression of TOX3 showed no obvious difference within these 3 subsets (*P*=0.635). ("○": extreme value)

4 DISCUSSION

In the present study, we detected the expression statuses of FGFR2, TOX3 and LSP1 proteins by ICC/IHC in 5 non-tumor and 8 human cancerous breast cell lines, 110 Inv-BCs, 11 intraductal papillomas, 16 fibroadenomas and 12 non-tumor breast conditions. Further, the associations of the expression of these proteins in Inv-BC with 15 clinical and pathological characteristics including age, menopausal status, HRT, family history, history of malignant tumors of other organs or tissues, multicentric/multifocal and bilateral tumors, pTNM, ER, PR and HER2 statuses and histological subtypes, were statistically analyzed.

4.1 Expression status of FGFR2 in human breast cancer

4.1.1 Distribution and intracellular localization of FGFR2 in breast lesions

By using a commercial monoclonal antibody which is raised against recombinant fragment corresponding to amino acid residues (aa) 621-724 of human FGFR2 (intracellular TK domain), we observed a mixed cytoplasmic, perinuclear and plasma membrane immunostaining pattern of FGFR2 in malignant and benign breast tumor cells, as well as in non-tumorous ductal/lobular breast epithelial cells, but not in myoepithelial cells or mesenchymal tissues. As mentioned in the introduction, FGFR2 has at least two alternative isoforms (FGFR2 IIIb and FGFR2 IIIc) which have two different C-terminal half of the IgIII domains (within the extracellular domain) due to alternative usage of exon 9 or 10 and thus manifest different ligand binding specificities and affinities in a cell lineage-specific manner. Generally, FGFR2 IIIb is predominantly expressed in epithelial cells, while FGFR2 IIIc is preferentially expressed in mesenchymal cells [64,67,111]. In human breast cancer, both isoforms are expressed and FGFR2 IIIc was demonstrated to express with a correlation with epithelial-to-mesenchymal transition (EMT) [111]. The distribution pattern observed here seems to resemble the distribution of FGFR2 IIIb rather than that of FGFR2 IIIc. However, according to the recognizable site, this antibody might not distinguish these two isoforms. Except for the alternative splicing isoforms of IgIII domain, there are still at least 3 C-terminal splicing variants of human FGFR2 IIIb, which were designated as

C1 (822aa), C2 (788aa) and C3 (769aa), having been identified [111,112]. Additional variants having different or deleted signal peptides, IgI, acid box, IgII and TM domains due to different mRNA splicing further add complexity to the presenting forms of FGFR2 [113]. However, according to the amino acid sequences, the differences of these variants are also beyond the recognizable site of this antibody. Despite a nuclear localization of FGFR2 was reported by using one or several commercial polyclonal antibodies recognizing the C-terminal cytoplasmic domain of FGFR2 in human normal and cancerous breast, mouse breast cancer, and also human non-small-cell-lung cancer tissues (NSCL) [88-91,114], we did not observe a certain nuclear localization in any sample examined by using the current monoclonal antibody. Earlier studies using one rabbit polyclonal antibody raised against the 805-821aa (C-terminal) of human FGFR2 or a self-produced rabbit antibody raised against the IgII domain of FGFR2 [110,115] revealed a cytoplasmic staining pattern of FGFR2 in cancerous or normal breast cells of tissue samples and a mixed cytoplasmic and perinuclear, but not a nuclear staining pattern in BC cell lines either [116]. Interestingly, different variants of another FGFR family member, FGFR3, showing different cellular localization, have also been reported. When four different variants of FGFR3 were transfected into COS-7 cells, one variant missing both the N-terminal signal peptide and the TM domains just showed a nuclear localization, while the other 3 variants showed a cytoplasmic staining pattern [116]. Moreover, the current antibody recognizing site (621-724 aa) of FGFR2 contains 3 major tyrosine sites of phosphorylation [117]. Thus, the potential causes underlying the different findings within these antibodies may be explained by the following possibilities: (i) Different antibodies targeting different domains may recognize different variants or functional (phosphorylated) statuses of FGFR2; (ii) Lower sensitivity of the monoclonal antibody we used here or lower specificity of the polyclonal antibodies used in those studies.

Because of the complexity of alternative splicing, comprehensive comparisons with antibodies targeting different domains of FGFR2, detection of the mRNA variants and phosphorylated statuses of this protein is still needed to clarify this apparent inconsistency.

4.1.2 Variing expression levels of FGFR2 in benign and malignant breast cell lines and tissues

In 5 non-tumor and 8 malignant human breast cell lines, we first determined by ICC that the breast cancer cell lines T47D and MCF7 are ER+/PR+; MDA-MB-453 and BT20 are HER2 overexpressed; BRC230, MDA-MB-231, CAL51 and MDA-MB-435, as well as all the 5 non-tumorous breast cell lines are triple-negative phenotypes. By using this antibody targeting the cytoplasmic TK domain, we only observed a moderate cytoplasmic staining of FGFR2 in the T47D breast cancer cell line but not in any other malignant or non-tumorous breast cell lines. Previous studies have shown that both T47D and MCF7 express *FGFR2* mRNA and also FGFR2 protein [116,118,119]. Here we only detected a positive expression of FGFR2 in T47D but not in MCF7. In a large-scale real-time quantitative RT-PCR study, the mRNA expression level of *FGFR2* was 3-fold higher in T47D than in MCF7 and no difference occurred in MCF7 with or without E2 (17β-estradiol) treatment [120]. In the same study, similar to our observations, MDA-MB-231 and MDA-MB-435 were undetectable for *FGFR2* mRNA [120]. The expression of FGFR2 has also been reported to be expressed in 184A1, MCF10A and BT-20 cell lines but controversially in HBL100 and MDA-MB-453 [116,118,119,121,122]. Therefore, compared with our results, except for the possibility that the antibody we chose here recognizes a certain isoform or functional status of FGFR2 which presents in T47D but not in the other breast cell lines, it's also possible that the expression level of FGFR2 in other cell lines is too low to be detectable by the present ICC detecting system.

The mRNA expression of *FGFR2* has been reported in 89% to 95% of human breast cancers and nearly in 100% of non-malignant breast tissues at varying levels but without significant difference between these two kinds of breast tissues [119,121]. By using the antibody we chose here, positive expression of FGFR2 with varied levels was detected in most cases (63.6% to 75.2%) of Inv-BCs, intraductal papillomas, fibroadenomas and non-tumor breast conditions (most cases were fibrocystic breast disease with hyperplasia). Although no significance could be pointed out by non-parameter comparisons, the seemingly lowest expression level was shown in intraductal papilloma. And high-level expression (H score of >200) of this protein was observed in 22.9% of Inv-BCs but not in any of the other three benign lesions. One study using a polyclonal anti-FGFR2 antibody detected a positive cytoplasmic expression of FGFR2 in only 17% of breast cancers [110]. However, in that study, a different scoring system (Allred's scoring

system [100]) was adopted and a score of 7 or above was regarded as positive, which is more equal to the high-level expression cut-off in the present study. Thus, our results generally agree with the previous findings and also confirm the notion that FGFR2 plays a role in both benign and malignant breast epithelial lesions.

4.1.3 Significant associations of the expression of FGFR2 with tumor grade, ER and PR statuses and non-triple negative phenotype of Inv-BC

When clinical and pathological features of breast cancer were considered, we observed an expression of FGFR2 in 90.8% of ER+ and 90.9% of PR+ Inv-BCs versus in 13.6% of ER- and 33.3% of PR- ones. Further statistical analyses revealed a positive correlation with ER and PR expression levels. Moreover, the staining of FGFR2 in 89.5% (17/19) of TNBCs was even entirely undetectable, while the other 2 cases only showed weak positive for FGFR2. In the large-scale real-time quantitative RT-PCR study mentioned above, 51/517 genes, including *PGR* (the gene encoding PR) and *FGFR2*, showed >3-fold up-regulated expression of mRNA in the ERα-positive breast tumor pool compared with the ERα-negative pool [120]. Similar results had been reported in another immunohistochemical study in which FGFR2 positive expression was shown in 6% of *BRCA1*-mutation (being predominantly ER- and having a basal-like phenotype) and 30% of *BRCA2*-mutation (characterized by ER+ and a luminal phenotype) carriers and positive associations of FGFR2 with ER and PR expressions were revealed [110]. Studies on breast cancer cell lines also demonstrated a much greater expression of FGFR2 in ER+ than in ER- cell lines [123]. In the 13 breast cell lines included in this study, FGFR2 expression was only detected in one ER+/PR+ breast cancer cell line (T47D) but not in other ER-/PR- ones. As PR is well-known as an ER-dependent protein and FGFR2 was expressed in both ER+/PR+ and ER+/PR- subsets but not in ER-/PR- cases, the key correlation might mainly lie in ER and FGFR2. Besides, while 89.5% of TNBCs were completely undetectable for FGFR2, no association of FGFR2 expression with HER2 overexpression was observed. The potential mechanisms underlying these results may at least include the following possibilities: (i) The expression of FGFR2 in breast cancers might be partly dependent on, or regulated by, the expression of ER, or they may locate at the downstream of one common signaling passway. While the increased risk in breast cancer conferred by the *FGFR2* allele is predominant for ER+

subset and no significant increase in risk for ER- subset, a functional study on this locus has shown evidence to support the notion that the Oct-1/Runx2 binding site is the dominant determinant of different expressions between the common and minor haplotypes of *FGFR2* locus [61]. Additionally, a genome-wide analysis of ER-binding sites has revealed that Oct-1/Runx2 may cooperate to increase gene expression [124]. Unfortunately, as far as the genotypes at rs2981592 (*FGFR2* locus) in 5 breast cancer cell lines were concerned, the only one (T47D) positive for FGFR2 shows homozygote of major allele, and the other 4 cell lines showing homozygote of minor allele or heterozygote of major and minor allele were undetectable for FGFR2. Therefore, no more proof could be given by this study to demonstrate the possible association of FGFR2 expression with the genotypes at *FGFR2* locus. Despite the fact that Spearman's ranked correlation test demonstrated significantly positive correlations of FGFR2 expression with ER and PR expressions, the r_s only reaches 0.513 and 0.414, respectively. In other words, there might be some other factors, beyond ER and PR, regulating the expression of FGFR2. It has also been shown that FGFR2 IIIb plays a role in the forming and branching of embryonic mammary glands without the presence of ERα, ERβ, PR, or the receptors for growth hormone (GH) and prolactin [67,70]. In breast cancers, similar mechanisms might also exist. (ii) The lack of expression of FGFR2 might be a common feature of most TNBCs, although TNBC has been demonstrated to be a genetically heterogeneous subset of breast cancer. Amplification and overexpression of *FGFR2* have been found in 4% TNBCs [125], however, most (86%) of the basal-like breast cancer (comprising the major part of TNBCs) were negative for FGFR2 when examined by IHC [110]. In our study, we didn't find any TNBC with overexpression of FGFR2. This may be because of the relatively smaller sample-size (only 19 cases of TNBC were determined here). (iii). As discussed above, the antibody used here may recognize a special functional status or isoform of FGFR2 which differed from those showing a nuclear localization. Thus, the positive correlation with ER/PR status may only represent one status or isoform of FGFR2. Nuclear localization of FGFR2 in breast cancer was not demonstrated to be associated with ER or PR status, or the genotype of rs2981582 at *FGFR2* locus [90].

In the current 16 cases with known *BRCA1/2* germline mutation status, although without significance, 4/5 (80%) of *BRCA1*-associated cases were negative or only weak positive for

FGFR2, while 3/4 (75%) of *BRCA2* and 4/7 (57%) of non-*BRCA1/2* mutation carriers were moderate or strong positive for FGFR2. Combined with the previous report [110], we agree that a higher expression of FGFR2 is in *BRCA2*-associated than *BRCA1*-associated breast cancers.

Additionally, a negative correlation of the expression of FGFR2 with tumor grade was observed. G3 tumors are known to be less differentiated and more likely to show lower expression for ER and PR and more likely to be TNBCs. As regards histological types of breast cancer, higher expression of FGFR2 in ILC than in IDC and MEC was revealed. This is possibly also due to the different expression statuses of ER in these three histological types, because ILC is usually positive for ER and MEC usually presents triple-negative profile. However, a recent study [126] which compared the expression of several growth factor receptors between male and female breast cancers did not reveal any significantly different expression of FGFR2 within IDC (121 males and 211 females), ILC (3 males and 25 females) and other histological types (9 males and 30 females), regardless of the gender. Due to the limited number of cases of ILC (16 cases) and MEC (4 cases) included in our study, further investigation with larger sample size is still needed to clarify the possible differences.

Despite the fact that family history could alter the association of breast cancer with SNPs at *FGFR2* locus, we did not demonstrate any association of family history with FGFR2 expression. Either, no further associations were observed when the other characteristics were concerned, except that a doubtfully higher positive proportion of FGFR2 was seen in elder ages (≥ 40y) which might more likely be due to the different ER statuses between younger and elder groups.

4.2 Expression status of TOX3 in human breast cancer

4.2.1 Expression profiling of TOX3 in benign and malignant breast tissues and cell lines

Similar to the other three members of the TOX HGM-box subfamily, TOX3 protein contains three separable domains: an N-terminal domain with a NLS, the highly conserved HMG-box and a C-terminal polyglutamine stretch [77]. By using a commercial polyclonal antibody which is generated from rabbits immunized with synthetic peptide between 221~250 aa from the central

region of human TOX3 (close to the HMG-box region), we observed a wide cell lineage distribution of TOX3 in the nuclei of breast ductal/lobular epithelial cells, myoepithelial cells, surrounding mesenchymal cells, eccrine glands, sebaceous glands, skeletal muscle cells, as well as all the 13 non-tumorous and malignant breast cell lines. This cellular localization pattern is consistent with the previous description of it as a nuclear protein [77], and its wide distribution may imply a basic role in cells.

Although all the 13 breast cell lines were detectable for TOX3 protein, a seemingly lower expression level in breast cancer cell lines than in non-tumor cell lines was implied but without statistical significance. No significant difference was observed when ER, PR and HER2 statuses were taken into consideration. As the SNPs at TOX3 locus have been demonstrated to be associated with breast cancer risk [40,41], we tried to examine the expression of TOX3 in breast cancer cell lines available for this information. In all the five breast cancer cell lines with detected SNP status of rs3803662 in earlier reports, TOX3 showed moderate positivity. The lowest H score of TOX3 was seen in the cell line harboring the homozygote of minor allele and one cell line harboring homozygote of major allele, while other 2 cell lines with major alleles and the one with heterozygote alleles presented similar scores. Since the sample-size was too small, the expression level could not be compared by statistical analysis. Then this study can not indicate whether the SNP status at rs3803662 may alter the expression of TOX3 protein.

Trying to determine the role of TOX3 in breast cancer, we further compared the expression of TOX3 in malignant and benign breast epithelial cells in tissue samples. Statistical analysis revealed that weak positive pattern was easier to find in Inv-BC than fibroadenoma, intraductal papilloma and non-tumor lesions. Significant difference was observed between Inv-BC and fibroadenoma.

4.2.2 Possibly lower expression of TOX3 in familial Inv-BCs and cases with malignant tumor history

As regards 15 clinical and pathological factors of Inv-BC, lower expression level of TOX3 was observed in familial Inv-BCs than sporadic ones and in Inv-BC cases positive for other organ's

malignant tumor history than negative ones. However, since most of the familial samples included in the present study were collected from 1981 to 1996 while all the sporadic ones were from 2003 to 2007, we cannot rule out the possibility that potentially different tissue sample qualities may influence the IHC results of TOX3, although no such phenomena were noticed when other proteins, including ER, PR, HER2, FGFR2 and LSP1, were observed in the same samples. In spite of this, we still cannot rule out the possibility that a lower expression of TOX3 may be associated with family history. More samples of equal quality are still needed to clarify this possibility.

To our knowledge, the present study is the first attempt to investigate the expression of TOX3 focusing on the histological localization, and our results imply a potential basic role of TOX3 in cancerous and non-cancerous breast tissues. However, quantification analysis of human TOX3 by quantitative real-time PCR in another study indicated that *TOX3* was expressed most prominently in the central nervous system and ileum but not in normal mammary tissues, and mainly in epithelial cells but not in endothelial or mesenchymal cells [77]. In view of the contradictions between these two studies, more work is still needed to clarify the underlying causes, such as the specificity of the antibody used here, consistency between the mRNA and protein level expressions of TOX3 and its exact functions in different cells and tissues.

4.3 No definite expression of LSP1 in benign or malignant human breast epithelial cells

As mentioned above, LSP1 has been reported to be restrictedly expressed in normal and malignant hematopoietic cells. Additional expression of LSP1 in mouse endothelial cells was demonstrated and was indicated to regulate neutrophil transendothelial migration [87]. Our IHC/ICC results also replicated a typical expression of LSP1 in normal thymocytes and in leukocytes presented in the stromal of benign and malignant breast tissues. With regard to breast epithelial cells, we first observed one IDC case presenting moderate cytoplasmic staining for LSP1 in tumor cells in a TMA sample collection. Additional 2 Inv-BCs further showed ambiguous weak positive in <15% of tumor cells. In order to exclude an artificial positive, the immunostaining of LSP1 in all the three cases had been repeated and similar staining pattern had

been observed. Another explanation is passive epithelial acquisition of LSP1 from the surrounding tumour-infiltrating lymphocytes [127]. However, no predominant lymphocytes infiltrating into the tumor cell sheets were observed in the 3 positive cases. Thus, aberrant expression might be a possible explanation. Such aberrant expression of lymphocytic antigen CD5 is well known in thymic carcinoma, as well as in malignant mesothelioma, gastric adenocarcinoma, endometrial carcinoma, and so on [127,128]. CD45 positive carcinomas have also been reported, although very rarely [127]. Despite of the possibility of occasional aberrant LSP1 expression in breast cancerous epithelial cells, it's still too early to conclude that LSP1 can play a role directly in breast cancerous epithelial cells. Moreover, we did not detect a positive expression of LSP1 in 13 benign or malignant breast cell lines, including 5 breast cancer cell lines showing all the three types of alleles at rs3817198 in the intron 10 of *LSP1* gene. Thus, it is still not likely that the risk variants at *LSP1* loci may exert a direct effect on the expression of LSP1 in breast cancer cells.

5 CONCLUSIONS

According to the working hypothesis and the current results, we draw the following conclusions: (i) FGFR2 expresses at varing levels in both benign and malignant breast epithelial cells. However, the nuclear intracellular localization of FGFR2 cannot be replicated by using the current commercial monoclonal antibody. (ii) FGFR2 can be highly expressed in Inv-BCs but usually not in benign lesions. (iii) The expression of FGFR2 in Inv-BC is positively correlated with the expressions of ER and PR and negatively correlated with tumor grade and usually lose expression in TNBCs. In spite of a small sample size, we still reckon that the *BRCA1/2* germline mutation status alters the expression of FGFR2. (iv) TOX3 may play a basic role in breast, as nuclear expression of TOX3 was detected in all kinds of breast cells in the current study. (v) The expression of TOX3 may be down-regulated in familial Inv-BCs and cases with malignant tumor history in other organs/tissues. *BRA1/2* germline mutation status of Inv-BC may not alter the expression levels of TOX3. However repeated and functional studies are still needed to clarify these tentative findings. (vi) LSP1 is usually not expressed in tumorous or non-tumorous breast epithelial cells, including those in familial Inv-BCs and *BRCA1/2* mutation carriers. It seems that LSP1 does not play a direct role in breast epithelial cells.

6 REFERENCES

1 Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA: A Cancer Journal for Clinicians 2011;61:69-90.

2 Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. Collaborative Group on Hormonal Factors in Breast Cancer. Lancet 1997;350:1047-59.

3 Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skytthe A, Hemminki K. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med 2000;343:78-85.

4 Peto J, Mack TM. High constant incidence in twins and other relatives of women with breast cancer. Nat Genet 2000;26:411-4.

5 Brewster A, Helzlsouer K. Breast cancer epidemiology, prevention, and early detection. Curr Opin Oncol 2001;13:420-5.

6 Collaborative Group on Hormonal Factors in Breast Cancer. Alcohol, tobacco and breast cancer – collaborative reanalysis of individual data from 53 epidemiological studies, including 58515 women with breast cancer and 95067 women without the disease. British Journal of Cancer 2002;87:1234-45.

7 Thompson WD. Genetic epidemiology of breast cancer. Cancer 1994;74:279-87.

8 Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. Lancet 2001;358:1389-99.

9 Mavaddat N, Antoniou AC, Easton DF, Garcia-Closas M. Genetic susceptibility to

breast cancer. Molecular Oncology 2010;4:174-91.

10 Garcia-Closas M, Chanock S. Genetic susceptibility loci for breast cancer by estrogen receptor status. Clin Cancer Res 2008;14:8000-9.

11 Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, King MC. Linkage of early-onset familial breast cancer to chromosome 17q21. Science 1990;250:1684-9.

12 Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 1994;266:66-71.

13 Wooster R, Neuhausen SL, Mangion J, Quirk Y, Ford D, Collins N, Nguyen K, Seal S, Tran T, Averill D, et al. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. Science 1994;265:2088-90.

14 Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G. Identification of the breast cancer susceptibility gene BRCA2. Nature 1995;378:789-92.

15 Tavtigian SV, Simard J, Rommens J, Couch F, Shattuck-Eidens D, Neuhausen S, Merajver S, Thorlacius S, Offit K, Stoppa-Lyonnet D, Belanger C, Bell R, Berry S, Bogden R, Chen Q, Davis T, Dumont M, Frye C, Hattier T, Jammulapati S, Janecki T, Jiang P, Kehrer R, Leblanc JF, Mitchell JT, McArthur-Morrison J, Nguyen K, Peng Y, Samson C, Schroeder M, Snyder SC, Steele L, Stringfellow M, Stroup C, Swedlund B, Swense J, Teng D, Thomas A, Tran T, Tranchant M, Weaver-Feldhaus J, Wong AK, Shizuya H, Eyfjord JE, Cannon-Albright L, Labrie F, Skolnick MH, Weber B, Kamb A, Goldgar DE. The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. Nat Genet 1996;12:333-7.

16 Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, Bishop DT, Weber B, Lenoir G, Chang-Claude J, Sobol H, Teare MD, Struewing J, Arason A, Scherneck S, Peto J, Rebbeck TR, Tonin P, Neuhausen S, Barkardottir R, Eyfjord J, Lynch H, Ponder BA, Gayther SA, Zelada-Hedman M, 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.

17 Breast Cancer Linkage Consortium. Pathology of familial breast cancer: differences between breast cancers in carriers of BRCA1 or BRCA2 mutations and sporadic cases. Breast Cancer Linkage Consortium. Lancet 1997;349:1505-10.

18 Da Silva L, Lakhani SR. Pathology of hereditary breast cancer. Mod Pathol 2010;23 Suppl 2:S46-51.

19 van der Groep P, van der Wall E, van Diest PJ. Pathology of hereditary breast cancer. Cell Oncol (Dordr) 2011;34:71-88.

20 Malkin D, Li FP, Strong LC, Fraumeni JF, Jr., Nelson CE, Kim DH, Kassel J, Gryka MA, Bischoff FZ, Tainsky MA, et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 1990;250:1233-8.

21 Olivier M, Goldgar DE, Sodha N, Ohgaki H, Kleihues P, Hainaut P, Eeles RA. Li-Fraumeni and related syndromes: correlation between tumor type, family structure, and TP53 genotype. Cancer Res 2003;63:6643-50.

22 Varley JM. Germline TP53 mutations and Li-Fraumeni syndrome. Hum Mutat 2003;21:313-20.

23 Eng C. PTEN: one gene, many syndromes. Hum Mutat 2003;22:183-98.

24 Tan MH, Mester JL, Ngeow J, Rybicki LA, Orloff MS, Eng C. Lifetime cancer risks in individuals with germline PTEN mutations. Clin Cancer Res 2012;18:400-7.

25 Hemminki A, Markie D, Tomlinson I, Avizienyte E, Roth S, Loukola A, Bignell G, Warren W, Aminoff M, Hoglund P, Jarvinen H, Kristo P, Pelin K, Ridanpaa M, Salovaara R, Toro T, Bodmer W, Olschwang S, Olsen AS, Stratton MR, de la Chapelle A, Aaltonen LA. A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. Nature 1998;391:184-7.

26 Boardman LA, Thibodeau SN, Schaid DJ, Lindor NM, McDonnell SK, Burgart LJ,

Ahlquist DA, Podratz KC, Pittelkow M, Hartmann LC. Increased risk for cancer in patients with the Peutz-Jeghers syndrome. Ann Intern Med 1998;128:896-9.

27 Jenne DE, Reimann H, Nezu J, Friedel W, Loff S, Jeschke R, Muller O, Back W, Zimmer M. Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. Nat Genet 1998;18:38-43.

28 Nevanlinna H, Bartek J. The CHEK2 gene and inherited breast cancer susceptibility. Oncogene 2006;25:5912-9.

29 Desrichard A, Bidet Y, Uhrhammer N, Bignon YJ. CHEK2 contribution to hereditary breast cancer in non-BRCA families. Breast Cancer Res 2011;13:R119.

30 Ahmed M, Rahman N. ATM and breast cancer susceptibility. Oncogene 2006;25:5906-11.

31 De Nicolo A, Tancredi M, Lombardi G, Flemma CC, Barbuti S, Di Cristofano C, Sobhian B, Bevilacqua G, Drapkin R, Caligo MA. A novel breast cancer-associated BRIP1 (FANCJ/BACH1) germ-line mutation impairs protein stability and function. Clin Cancer Res 2008;14:4672-80.

32 Seal S, Thompson D, Renwick A, Elliott A, Kelly P, Barfoot R, Chagtai T, Jayatilake H, Ahmed M, Spanova K, North B, McGuffog L, Evans DG, Eccles D, Easton DF, Stratton MR, Rahman N. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. Nat Genet 2006;38:1239-41.

33 Rahman N, Seal S, Thompson D, Kelly P, Renwick A, Elliott A, Reid S, Spanova K, Barfoot R, Chagtai T, Jayatilake H, McGuffog L, Hanks S, Evans DG, Eccles D, Easton DF, Stratton MR. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. Nat Genet 2007;39:165-7.

34 Varghese JS, Easton DF. Genome-wide association studies in common cancers--what have we learnt? Curr Opin Genet Dev 2010;20:201-9.

35 Cox A, Dunning AM, Garcia-Closas M, Balasubramanian S, Reed MW, Pooley KA, Scollen S, Baynes C, Ponder BA, Chanock S, Lissowska J, Brinton L, Peplonska B, Southey MC, Hopper JL, McCredie MR, Giles GG, Fletcher O, Johnson N, dos Santos Silva I, Gibson L, Bojesen SE, Nordestgaard BG, Axelsson CK, Torres D, Hamann U, Justenhoven C, Brauch H, Chang-Claude J, Kropp S, Risch A, Wang-Gohrke S, Schurmann P, Bogdanova N, Dork T, Fagerholm R, Aaltonen K, Blomqvist C, Nevanlinna H, Seal S, Renwick A, Stratton MR, Rahman N, Sangrajrang S, Hughes D, Odefrey F, Brennan P, Spurdle AB, Chenevix-Trench G, Beesley J, Mannermaa A, Hartikainen J, Kataja V, Kosma VM, Couch FJ, Olson JE, Goode EL, Broeks A, Schmidt MK, Hogervorst FB, Van't Veer LJ, Kang D, Yoo KY, Noh DY, Ahn SH, Wedren S, Hall P, Low YL, Liu J, Milne RL, Ribas G, Gonzalez-Neira A, Benitez J, Sigurdson AJ, Stredrick DL, Alexander BH, Struewing JP, Pharoah PD, Easton DF. A common coding variant in CASP8 is associated with breast cancer risk. Nat Genet 2007;39:352-8.

36 Breast Cancer Association Consortium. Commonly studied single-nucleotide polymorphisms and breast cancer: results from the Breast Cancer Association Consortium. J Natl Cancer Inst 2006;98:1382-96.

37 Yin M, Yan J, Wei S, Wei Q. CASP8 polymorphisms contribute to cancer susceptibility: evidence from a meta-analysis of 23 publications with 55 individual studies. Carcinogenesis 2010;31:850-7.

38 Zhang B, Beeghly-Fadiel A, Long J, Zheng W. Genetic variants associated with breast-cancer risk: comprehensive research synopsis, meta-analysis, and epidemiological evidence. Lancet Oncol 2011;12:477-88.

39 Shephard ND, Abo R, Rigas SH, Frank B, Lin WY, Brock IW, Shippen A, Balasubramanian SP, Reed MW, Bartram CR, Meindl A, Schmutzler RK, Engel C, Burwinkel B, Cannon-Albright LA, Allen-Brady K, Camp NJ, Cox A. A breast cancer risk haplotype in the caspase-8 gene. Cancer Res 2009;69:2724-8.

40 Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, Healey CS, Bowman R, Meyer KB, Haiman CA, Kolonel LK, Henderson BE, Le Marchand L, Brennan P, Sangrajrang S, Gaborieau V, Odefrey F, Shen CY, Wu PE, Wang HC, Eccles D, Evans DG, Peto J, Fletcher O,

Johnson N, Seal S, Stratton MR, Rahman N, Chenevix-Trench G, Bojesen SE, Nordestgaard BG, Axelsson CK, Garcia-Closas M, Brinton L, Chanock S, Lissowska J, Peplonska B, Nevanlinna H, Fagerholm R, Eerola H, Kang D, Yoo KY, Noh DY, Ahn SH, Hunter DJ, Hankinson SE, Cox DG, Hall P, Wedren S, Liu J, Low YL, Bogdanova N, Schurmann P, Dork T, Tollenaar RA, Jacobi CE, Devilee P, Klijn JG, Sigurdson AJ, Doody MM, Alexander BH, Zhang J, Cox A, Brock IW, MacPherson G, Reed MW, Couch FJ, Goode EL, Olson JE, Meijers-Heijboer H, van den Ouweland A, Uitterlinden A, Rivadeneira F, Milne RL, Ribas G, Gonzalez-Neira A, Benitez J, Hopper JL, McCredie M, Southey M, Giles GG, Schroen C, Justenhoven C, Brauch H, Hamann U, Ko YD, Spurdle AB, Beesley J, Chen X, Mannermaa A, Kosma VM, Kataja V, Hartikainen J, Day NE, Cox DR, Ponder BA. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 2007;447:1087-93.

41 Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, Masson G, Jakobsdottir M, Thorlacius S, Helgason A, Aben KK, Strobbe LJ, Albers-Akkers MT, Swinkels DW, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Godino J, Garcia-Prats MD, Polo E, Tres A, Mouy M, Saemundsdottir J, Backman VM, Gudmundsson L, Kristjansson K, Bergthorsson JT, Kostic J, Frigge ML, Geller F, Gudbjartsson D, Sigurdsson H, Jonsdottir T, Hrafnkelsson J, Johannsson J, Sveinsson T, Myrdal G, Grimsson HN, Jonsson T, von Holst S, Werelius B, Margolin S, Lindblom A, Mayordomo JI, Haiman CA, Kiemeney LA, Johannsson OT, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 2007;39:865-9.

42 Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A, Wang J, Yu K, Chatterjee N, Orr N, Willett WC, Colditz GA, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Hayes RB, Tucker M, Gerhard DS, Fraumeni JF, Jr., Hoover RN, Thomas G, Chanock SJ. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 2007;39:870-4.

43 Stacey SN, Manolescu A, Sulem P, Thorlacius S, Gudjonsson SA, Jonsson GF, Jakobsdottir M, Bergthorsson JT, Gudmundsson J, Aben KK, Strobbe LJ, Swinkels DW, van Engelenburg KC, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Saez B, Lambea J, Godino J, Polo E, Tres A, Picelli S, Rantala J, Margolin S, Jonsson T, Sigurdsson H,

Jonsdottir T, Hrafnkelsson J, Johannsson J, Sveinsson T, Myrdal G, Grimsson HN, Sveinsdottir SG, Alexiusdottir K, Saemundsdottir J, Sigurdsson A, Kostic J, Gudmundsson L, Kristjansson K, Masson G, Fackenthal JD, Adebamowo C, Ogundiran T, Olopade OI, Haiman CA, Lindblom A, Mayordomo JI, Kiemeney LA, Gulcher JR, Rafnar T, Thorsteinsdottir U, Johannsson OT, Kong A, Stefansson K. Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 2008;40:703-6.

44 Ahmed S, Thomas G, Ghoussaini M, Healey CS, Humphreys MK, Platte R, Morrison J, Maranian M, Pooley KA, Luben R, Eccles D, Evans DG, Fletcher O, Johnson N, dos Santos Silva I, Peto J, Stratton MR, Rahman N, Jacobs K, Prentice R, Anderson GL, Rajkovic A, Curb JD, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Diver WR, Bojesen S, Nordestgaard BG, Flyger H, Dork T, Schurmann P, Hillemanns P, Karstens JH, Bogdanova NV, Antonenkova NN, Zalutsky IV, Bermisheva M, Fedorova S, Khusnutdinova E, Kang D, Yoo KY, Noh DY, Ahn SH, Devilee P, van Asperen CJ, Tollenaar RA, Seynaeve C, Garcia-Closas M, Lissowska J, Brinton L, Peplonska B, Nevanlinna H, Heikkinen T, Aittomaki K, Blomqvist C, Hopper JL, Southey MC, Smith L, Spurdle AB, Schmidt MK, Broeks A, van Hien RR, Cornelissen S, Milne RL, Ribas G, Gonzalez-Neira A, Benitez J, Schmutzler RK, Burwinkel B, Bartram CR, Meindl A, Brauch H, Justenhoven C, Hamann U, Chang-Claude J, Hein R, Wang-Gohrke S, Lindblom A, Margolin S, Mannermaa A, Kosma VM, Kataja V, Olson JE, Wang X, Fredericksen Z, Giles GG, Severi G, Baglietto L, English DR, Hankinson SE, Cox DG, Kraft P, Vatten LJ, Hveem K, Kumle M, Sigurdson A, Doody M, Bhatti P, Alexander BH, Hooning MJ, van den Ouweland AM, Oldenburg RA, Schutte M, Hall P, Czene K, Liu J, Li Y, Cox A, Elliott G, Brock I, Reed MW, Shen CY, Yu JC, Hsu GC, Chen ST, Anton-Culver H, Ziogas A, Andrulis IL, Knight JA, Beesley J, Goode EL, Couch F, Chenevix-Trench G, Hoover RN, Ponder BA, Hunter DJ, Pharoah PD, Dunning AM, Chanock SJ, Easton DF. Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. Nat Genet 2009;41:585-90.

45 Thomas G, Jacobs KB, Kraft P, Yeager M, Wacholder S, Cox DG, Hankinson SE, Hutchinson A, Wang Z, Yu K, Chatterjee N, Garcia-Closas M, Gonzalez-Bosquet J, Prokunina-Olsson L, Orr N, Willett WC, Colditz GA, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Diver R, Prentice R, Jackson R, Kooperberg C, Chlebowski R, Lissowska J, Peplonska B, Brinton LA, Sigurdson A, Doody M, Bhatti P, Alexander BH, Buring J, Lee IM, Vatten LJ, Hveem K, Kumle M, Hayes RB, Tucker M, Gerhard DS, Fraumeni JF, Jr., Hoover RN, Chanock SJ, Hunter DJ. A multistage genome-wide

association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). Nat Genet 2009;41:579-84.

46 Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, Wen W, Levy S, Deming SL, Haines JL, Gu K, Fair AM, Cai Q, Lu W, Shu XO. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nat Genet 2009;41:324-8.

47 Antoniou AC, Wang X, Fredericksen ZS, McGuffog L, Tarrell R, Sinilnikova OM, Healey S, Morrison J, Kartsonaki C, Lesnick T, Ghoussaini M, Barrowdale D, Peock S, Cook M, Oliver C, Frost D, Eccles D, Evans DG, Eeles R, Izatt L, Chu C, Douglas F, Paterson J, Stoppa-Lyonnet D, Houdayer C, Mazoyer S, Giraud S, Lasset C, Remenieras A, Caron O, Hardouin A, Berthet P, Hogervorst FB, Rookus MA, Jager A, van den Ouweland A, Hoogerbrugge N, van der Luijt RB, Meijers-Heijboer H, Gomez Garcia EB, Devilee P, Vreeswijk MP, Lubinski J, Jakubowska A, Gronwald J, Huzarski T, Byrski T, Gorski B, Cybulski C, Spurdle AB, Holland H, Goldgar DE, John EM, Hopper JL, Southey M, Buys SS, Daly MB, Terry MB, Schmutzler RK, Wappenschmidt B, Engel C, Meindl A, Preisler-Adams S, Arnold N, Niederacher D, Sutter C, Domchek SM, Nathanson KL, Rebbeck T, Blum JL, Piedmonte M, Rodriguez GC, Wakeley K, Boggess JF, Basil J, Blank SV, Friedman E, Kaufman B, Laitman Y, Milgrom R, Andrulis IL, Glendon G, Ozcelik H, Kirchhoff T, Vijai J, Gaudet MM, Altshuler D, Guiducci C, Loman N, Harbst K, Rantala J, Ehrencrona H, Gerdes AM, Thomassen M, Sunde L, Peterlongo P, Manoukian S, Bonanni B, Viel A, Radice P, Caldes T, de la Hoya M, Singer CF, Fink-Retter A, Greene MH, Mai PL, Loud JT, Guidugli L, Lindor NM, Hansen TV, Nielsen FC, Blanco I, Lazaro C, Garber J, Ramus SJ, Gayther SA, Phelan C, Narod S, Szabo CI, Benitez J, Osorio A, Nevanlinna H, Heikkinen T, Caligo MA, Beattie MS, Hamann U, Godwin AK, Montagna M, Casella C, Neuhausen SL, Karlan BY, Tung N, Toland AE, Weitzel J, Olopade O, Simard J, Soucy P, Rubinstein WS, Arason A, Rennert G, Martin NG, Montgomery GW, Chang-Claude J, Flesch-Janys D, Brauch H, Severi G, Baglietto L, Cox A, Cross SS, Miron P, Gerty SM, Tapper W, Yannoukakos D, Fountzilas G, Fasching PA, Beckmann MW, Dos Santos Silva I, Peto J, Lambrechts D, Paridaens R, Rudiger T, Forsti A, Winqvist R, Pylkas K, Diasio RB, Lee AM, Eckel-Passow J, Vachon C, Blows F, Driver K, Dunning A, Pharoah PP, Offit K, Pankratz VS, Hakonarson H, Chenevix-Trench G, Easton DF, Couch FJ. A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. Nat Genet 2010;42:885-92.

48 Turnbull C, Ahmed S, Morrison J, Pernet D, Renwick A, Maranian M, Seal S, Ghoussaini M, Hines S, Healey CS, Hughes D, Warren-Perry M, Tapper W, Eccles D, Evans DG, Hooning M, Schutte M, van den Ouweland A, Houlston R, Ross G, Langford C, Pharoah PD, Stratton MR, Dunning AM, Rahman N, Easton DF. Genome-wide association study identifies five new breast cancer susceptibility loci. Nat Genet 2010;42:504-7.

49 Fletcher O, Johnson N, Orr N, Hosking FJ, Gibson LJ, Walker K, Zelenika D, Gut I, Heath S, Palles C, Coupland B, Broderick P, Schoemaker M, Jones M, Williamson J, Chilcott-Burns S, Tomczyk K, Simpson G, Jacobs KB, Chanock SJ, Hunter DJ, Tomlinson IP, Swerdlow A, Ashworth A, Ross G, dos Santos Silva I, Lathrop M, Houlston RS, Peto J. Novel breast cancer susceptibility locus at 9q31.2: results of a genome-wide association study. J Natl Cancer Inst 2011;103:425-35.

50 Haiman CA, Chen GK, Vachon CM, Canzian F, Dunning A, Millikan RC, Wang X, Ademuyiwa F, Ahmed S, Ambrosone CB, Baglietto L, Balleine R, Bandera EV, Beckmann MW, Berg CD, Bernstein L, Blomqvist C, Blot WJ, Brauch H, Buring JE, Carey LA, Carpenter JE, Chang-Claude J, Chanock SJ, Chasman DI, Clarke CL, Cox A, Cross SS, Deming SL, Diasio RB, Dimopoulos AM, Driver WR, Dunnebier T, Durcan L, Eccles D, Edlund CK, Ekici AB, Fasching PA, Feigelson HS, Flesch-Janys D, Fostira F, Forsti A, Fountzilas G, Gerty SM, Giles GG, Godwin AK, Goodfellow P, Graham N, Greco D, Hamann U, Hankinson SE, Hartmann A, Hein R, Heinz J, Holbrook A, Hoover RN, Hu JJ, Hunter DJ, Ingles SA, Irwanto A, Ivanovich J, John EM, Johnson N, Jukkola-Vuorinen A, Kaaks R, Ko YD, Kolonel LN, Konstantopoulou I, Kosma VM, Kulkarni S, Lambrechts D, Lee AM, Marchand LL, Lesnick T, Liu J, Lindstrom S, Mannermaa A, Margolin S, Martin NG, Miron P, Montgomery GW, Nevanlinna H, Nickels S, Nyante S, Olswold C, Palmer J, Pathak H, Pectasides D, Perou CM, Peto J, Pharoah PD, Pooler LC, Press MF, Pylkas K, Rebbeck TR, Rodriguez-Gil JL, Rosenberg L, Ross E, Rudiger T, Silva Idos S, Sawyer E, Schmidt MK, Schulz-Wendtland R, Schumacher F, Severi G, Sheng X, Signorello LB, Sinn HP, Stevens KN, Southey MC, Tapper WJ, Tomlinson I, Hogervorst FB, Wauters E, Weaver J, Wildiers H, Winqvist R, Van Den Berg D, Wan P, Xia LY, Yannoukakos D, Zheng W, Ziegler RG, Siddiq A, Slager SL, Stram DO, Easton D, Kraft P, Henderson BE, Couch FJ. A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. Nat Genet 2011;43:1210-4.

51 Ghoussaini M, Fletcher O, Michailidou K, Turnbull C, Schmidt MK, Dicks E,

Dennis J, Wang Q, Humphreys MK, Luccarini C, Baynes C, Conroy D, Maranian M, Ahmed S, Driver K, Johnson N, Orr N, dos Santos Silva I, Waisfisz Q, Meijers-Heijboer H, Uitterlinden AG, Rivadeneira F, Hall P, Czene K, Irwanto A, Liu J, Nevanlinna H, Aittomaki K, Blomqvist C, Meindl A, Schmutzler RK, Muller-Myhsok B, Lichtner P, Chang-Claude J, Hein R, Nickels S, Flesch-Janys D, Tsimiklis H, Makalic E, Schmidt D, Bui M, Hopper JL, Apicella C, Park DJ, Southey M, Hunter DJ, Chanock SJ, Broeks A, Verhoef S, Hogervorst FB, Fasching PA, Lux MP, Beckmann MW, Ekici AB, Sawyer E, Tomlinson I, Kerin M, Marme F, Schneeweiss A, Sohn C, Burwinkel B, Guenel P, Truong T, Cordina-Duverger E, Menegaux F, Bojesen SE, Nordestgaard BG, Nielsen SF, Flyger H, Milne RL, Alonso MR, Gonzalez-Neira A, Benitez J, Anton-Culver H, Ziogas A, Bernstein L, Dur CC, Brenner H, Muller H, Arndt V, Stegmaier C, Justenhoven C, Brauch H, Bruning T, Wang-Gohrke S, Eilber U, Dork T, Schurmann P, Bremer M, Hillemanns P, Bogdanova NV, Antonenkova NN, Rogov YI, Karstens JH, Bermisheva M, Prokofieva D, Khusnutdinova E, Lindblom A, Margolin S, Mannermaa A, Kataja V, Kosma VM, Hartikainen JM, Lambrechts D, Yesilyurt BT, Floris G, Leunen K, Manoukian S, Bonanni B, Fortuzzi S, Peterlongo P, Couch FJ, Wang X, Stevens K, Lee A, Giles GG, Baglietto L, Severi G, McLean C, Alnaes GG, Kristensen V, Borrensen-Dale AL, John EM, Miron A, Winqvist R, Pylkas K, Jukkola-Vuorinen A, Kauppila S, Andrulis IL, Glendon G, Mulligan AM, Devilee P, van Asperen CJ, Tollenaar RA, Seynaeve C, Figueroa JD, Garcia-Closas M, Brinton L, Lissowska J, Hooning MJ, Hollestelle A, Oldenburg RA, van den Ouweland AM, Cox A, Reed MW, Shah M, Jakubowska A, Lubinski J, Jaworska K, Durda K, Jones M, Schoemaker M, Ashworth A, Swerdlow A, Beesley J, Chen X, Muir KR, Lophatananon A, Rattanamongkongul S, Chaiwerawattana A, Kang D, Yoo KY, Noh DY, Shen CY, Yu JC, Wu PE, Hsiung CN, Perkins A, Swann R, Velentzis L, Eccles DM, Tapper WJ, Gerty SM, Graham NJ, Ponder BA, Chenevix-Trench G, Pharoah PD, Lathrop M, Dunning AM, Rahman N, Peto J, Easton DF. Genome-wide association analysis identifies three new breast cancer susceptibility loci. Nat Genet 2012;44:312-8.

52 Gold B, Kirchhoff T, Stefanov S, Lautenberger J, Viale A, Garber J, Friedman E, Narod S, Olshen AB, Gregersen P, Kosarin K, Olsh A, Bergeron J, Ellis NA, Klein RJ, Clark AG, Norton L, Dean M, Boyd J, Offit K. Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. Proc Natl Acad Sci U S A 2008;105:4340-5.

53 Garcia-Closas M, Hall P, Nevanlinna H, Pooley K, Morrison J, Richesson DA, Bojesen SE, Nordestgaard BG, Axelsson CK, Arias JI, Milne RL, Ribas G, Gonzalez-Neira A,

Benitez J, Zamora P, Brauch H, Justenhoven C, Hamann U, Ko YD, Bruening T, Haas S, Dork T, Schurmann P, Hillemanns P, Bogdanova N, Bremer M, Karstens JH, Fagerholm R, Aaltonen K, Aittomaki K, von Smitten K, Blomqvist C, Mannermaa A, Uusitupa M, Eskelinen M, Tengstrom M, Kosma VM, Kataja V, Chenevix-Trench G, Spurdle AB, Beesley J, Chen X, Devilee P, van Asperen CJ, Jacobi CE, Tollenaar RA, Huijts PE, Klijn JG, Chang-Claude J, Kropp S, Slanger T, Flesch-Janys D, Mutschelknauss E, Salazar R, Wang-Gohrke S, Couch F, Goode EL, Olson JE, Vachon C, Fredericksen ZS, Giles GG, Baglietto L, Severi G, Hopper JL, English DR, Southey MC, Haiman CA, Henderson BE, Kolonel LN, Le Marchand L, Stram DO, Hunter DJ, Hankinson SE, Cox DG, Tamimi R, Kraft P, Sherman ME, Chanock SJ, Lissowska J, Brinton LA, Peplonska B, Hooning MJ, Meijers-Heijboer H, Collee JM, van den Ouweland A, Uitterlinden AG, Liu J, Lin LY, Yuqing L, Humphreys K, Czene K, Cox A, Balasubramanian SP, Cross SS, Reed MW, Blows F, Driver K, Dunning A, Tyrer J, Ponder BA, Sangrajrang S, Brennan P, McKay J, Odefrey F, Gabrieau V, Sigurdson A, Doody M, Struewing JP, Alexander B, Easton DF, Pharoah PD. Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. PLoS Genet 2008;4:e1000054.

54 Hemminki K, Muller-Myhsok B, Lichtner P, Engel C, Chen B, Burwinkel B, Forsti A, Sutter C, Wappenschmidt B, Hellebrand H, Illig T, Arnold N, Niederacher D, Dworniczak B, Deissler H, Kast K, Gadzicki D, Meitinger T, Wichmann HE, Kiechle M, Bartram CR, Schmutzler RK, Meindl A. Low-risk variants FGFR2, TNRC9 and LSP1 in German familial breast cancer patients. Int J Cancer 2010;126:2858-62.

55 Antoniou AC, Spurdle AB, Sinilnikova OM, Healey S, Pooley KA, Schmutzler RK, Versmold B, Engel C, Meindl A, Arnold N, Hofmann W, Sutter C, Niederacher D, Deissler H, Caldes T, Kampjarvi K, Nevanlinna H, Simard J, Beesley J, Chen X, Neuhausen SL, Rebbeck TR, Wagner T, Lynch HT, Isaacs C, Weitzel J, Ganz PA, Daly MB, Tomlinson G, Olopade OI, Blum JL, Couch FJ, Peterlongo P, Manoukian S, Barile M, Radice P, Szabo CI, Pereira LH, Greene MH, Rennert G, Lejbkowicz F, Barnett-Griness O, Andrulis IL, Ozcelik H, Gerdes AM, Caligo MA, Laitman Y, Kaufman B, Milgrom R, Friedman E, Domchek SM, Nathanson KL, Osorio A, Llort G, Milne RL, Benitez J, Hamann U, Hogervorst FB, Manders P, Ligtenberg MJ, van den Ouweland AM, Peock S, Cook M, Platte R, Evans DG, Eeles R, Pichert G, Chu C, Eccles D, Davidson R, Douglas F, Godwin AK, Barjhoux L, Mazoyer S, Sobol H, Bourdon V, Eisinger F, Chompret A, Capoulade C, Bressac-de Paillerets B, Lenoir GM, Gauthier-Villars M, Houdayer C, Stoppa-Lyonnet D, Chenevix-Trench G, Easton DF. Common breast

cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers. Am J Hum Genet 2008;82:937-48.

56 Antoniou AC, Sinilnikova OM, McGuffog L, Healey S, Nevanlinna H, Heikkinen T, Simard J, Spurdle AB, Beesley J, Chen X, Neuhausen SL, Ding YC, Couch FJ, Wang X, Fredericksen Z, Peterlongo P, Peissel B, Bonanni B, Viel A, Bernard L, Radice P, Szabo CI, Foretova L, Zikan M, Claes K, Greene MH, Mai PL, Rennert G, Lejbkowicz F, Andrulis IL, Ozcelik H, Glendon G, Gerdes AM, Thomassen M, Sunde L, Caligo MA, Laitman Y, Kontorovich T, Cohen S, Kaufman B, Dagan E, Baruch RG, Friedman E, Harbst K, Barbany-Bustinza G, Rantala J, Ehrencrona H, Karlsson P, Domchek SM, Nathanson KL, Osorio A, Blanco I, Lasa A, Benitez J, Hamann U, Hogervorst FB, Rookus MA, Collee JM, Devilee P, Ligtenberg MJ, van der Luijt RB, Aalfs CM, Waisfisz Q, Wijnen J, van Roozendaal CE, Peock S, Cook M, Frost D, Oliver C, Platte R, Evans DG, Lalloo F, Eeles R, Izatt L, Davidson R, Chu C, Eccles D, Cole T, Hodgson S, Godwin AK, Stoppa-Lyonnet D, Buecher B, Leone M, Bressac-de Paillerets B, Remenieras A, Caron O, Lenoir GM, Sevenet N, Longy M, Ferrer SF, Prieur F, Goldgar D, Miron A, John EM, Buys SS, Daly MB, Hopper JL, Terry MB, Yassin Y, Singer C, Gschwantler-Kaulich D, Staudigl C, Hansen TO, Barkardottir RB, Kirchhoff T, Pal P, Kosarin K, Offit K, Piedmonte M, Rodriguez GC, Wakeley K, Boggess JF, Basil J, Schwartz PE, Blank SV, Toland AE, Montagna M, Casella C, Imyanitov EN, Allavena A, Schmutzler RK, Versmold B, Engel C, Meindl A, Ditsch N, Arnold N, Niederacher D, Deissler H, Fiebig B, Suttner C, Schonbuchner I, Gadzicki D, Caldes T, de la Hoya M, Pooley KA, Easton DF, Chenevix-Trench G. Common variants in LSP1, 2q35 and 8q24 and breast cancer risk for BRCA1 and BRCA2 mutation carriers. Hum Mol Genet 2009;18:4442-56.

57 Latif A, Hadfield KD, Roberts SA, Shenton A, Lalloo F, Black GC, Howell A, Evans DG, Newman WG. Breast cancer susceptibility variants alter risks in familial disease. J Med Genet 2010;47:126-31.

58 Mulligan AM, Couch FJ, Barrowdale D, Domchek SM, Eccles D, Nevanlinna H, Ramus SJ, Robson M, Sherman M, Spurdle AB, Wappenschmidt B, Lee A, McGuffog L, Healey S, Sinilnikova OM, Janavicius R, Hansen T, Nielsen FC, Ejlertsen B, Osorio A, Munoz-Repeto I, Duran M, Godino J, Pertesi M, Benitez J, Peterlongo P, Manoukian S, Peissel B, Zaffaroni D, Cattaneo E, Bonanni B, Viel A, Pasini B, Papi L, Ottini L, Savarese A, Bernard L, Radice P, Hamann U, Verheus M, Meijers-Heijboer HE, Wijnen J, Gomez Garcia EB, Nelen MR, Kets CM,

Seynaeve C, Tilanus-Linthorst MM, van der Luijt RB, van Os T, Rookus M, Frost D, Jones JL, Evans DG, Lalloo F, Eeles R, Izatt L, Adlard J, Davidson R, Cook J, Donaldson A, Dorkins H, Gregory H, Eason J, Houghton C, Barwell J, Side LE, McCann E, Murray A, Peock S, Godwin AK, Schmutzler RK, Rhiem K, Engel C, Meindl A, Ruehl I, Arnold N, Niederacher D, Sutter C, Deissler H, Gadzicki D, Kast K, Preisler-Adams S, Varon-Mateeva R, Schoenbuchner I, Fiebig B, Heinritz W, Schafer D, Gevensleben H, Caux-Moncoutier V, Fassy-Colcombet M, Cornelis F, Mazoyer S, Leone M, Boutry-Kryza N, Hardouin A, Berthet P, Muller D, Fricker JP, Mortemousque I, Pujol P, Coupier I, Lebrun M, Kientz C, Longy M, Sevenet N, Stoppa-Lyonnet D, Isaacs C, Caldes T, de la Hoya M, Heikkinen T, Aittomaki K, Blanco I, Lazaro C, Barkardottir RB, Soucy P, Dumont M, Simard J, Montagna M, Tognazzo S, D'Andrea E, Fox S, Yan M, Rebbeck T, Olopade O, Weitzel JN, Lynch HT, Ganz PA, Tomlinson GE, Wang X, Fredericksen Z, Pankratz VS, Lindor NM, Szabo C, Offit K, Sakr R, Gaudet M, Bhatia J, Kauff N, Singer CF, Tea MK, Gschwantler-Kaulich D, Fink-Retter A, Mai PL, Greene MH, Imyanitov E, O'Malley FP, Ozcelik H, Glendon G, Toland AE, Gerdes AM, Thomassen M, Kruse TA, Jensen UB, Skytte AB, Caligo MA, Soller M, Henriksson K, Wachenfeldt v A, Arver B, Stenmark-Askmalm M, Karlsson P, Ding YC, Neuhausen SL, Beattie M, Pharoah PD, Moysich KB, Nathanson KL, Karlan BY, Gross J, John EM, Daly MB, Buys SM, Southey MC, Hopper JL, Terry MB, Chung W, Miron AF, Goldgar D, Chenevix-Trench G, Easton DF, Andrulis IL, Antoniou AC. Common breast cancer susceptibility alleles are associated with tumour subtypes in BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of BRCA1/2. Breast Cancer Res 2011;13:R110.

59 Ripperger T, Gadzicki D, Meindl A, Schlegelberger B. Breast cancer susceptibility: current knowledge and implications for genetic counselling. Eur J Hum Genet 2009;17:722-31.

60 Meindl A. Identification of Novel Susceptibility Genes for Breast Cancer - Genome-Wide Association Studies or Evaluation of Candidate Genes? Breast Care (Basel) 2009;4:93-9.

61 Meyer KB, Maia AT, O'Reilly M, Teschendorff AE, Chin SF, Caldas C, Ponder BA. Allele-specific up-regulation of FGFR2 increases susceptibility to breast cancer. PLoS Biol 2008;6:e108.

62 Sun C, Olopade OI, Di Rienzo A. rs2981582 is associated with FGFR2 expression

in normal breast. Cancer Genet Cytogenet 2010;197:193-4.

63 Riaz M, Berns EM, Sieuwerts AM, Ruigrok-Ritstier K, de Weerd V, Groenewoud A, Uitterlinden AG, Look MP, Klijn JG, Sleijfer S, Foekens JA, Martens JW. Correlation of breast cancer susceptibility loci with patient characteristics, metastasis-free survival, and mRNA expression of the nearest genes. Breast Cancer Res Treat 2011.

64 Katoh M. Cancer genomics and genetics of FGFR2 (Review). Int J Oncol 2008;33:233-7.

65 Beenken A, Mohammadi M. The FGF family: biology, pathophysiology and therapy. Nat Rev Drug Discov 2009;8:235-53.

66 Haugsten EM, Wiedlocha A, Olsnes S, Wesche J. Roles of fibroblast growth factor receptors in carcinogenesis. Mol Cancer Res 2010;8:1439-52.

67 Dillon C, Spencer-Dene B, Dickson C. A crucial role for fibroblast growth factor signaling in embryonic mammary gland development. J Mammary Gland Biol Neoplasia 2004;9:207-15.

68 Turner N, Grose R. Fibroblast growth factor signalling: from development to cancer. Nat Rev Cancer 2010;10:116-29.

69 Katoh M. FGFR2 abnormalities underlie a spectrum of bone, skin, and cancer pathologies. J Invest Dermatol 2009;129:1861-7.

70 Sternlicht MD. Key stages in mammary gland development: the cues that regulate ductal branching morphogenesis. Breast Cancer Res 2006;8:201.

71 Azuma K, Tsurutani J, Sakai K, Kaneda H, Fujisaka Y, Takeda M, Watatani M, Arao T, Satoh T, Okamoto I, Kurata T, Nishio K, Nakagawa K. Switching addictions between HER2 and FGFR2 in HER2-positive breast tumor cells: FGFR2 as a potential target for salvage after lapatinib failure. Biochem Biophys Res Commun 2011;407:219-24.

72 Margolis RL, Abraham MR, Gatchell SB, Li SH, Kidwai AS, Breschel TS, Stine OC, Callahan C, McInnis MG, Ross CA. cDNAs with long CAG trinucleotide repeats from human brain. Hum Genet 1997;100:114-22.

73 O'Flaherty E, Kaye J. TOX defines a conserved subfamily of HMG-box proteins. BMC Genomics 2003;4:13.

74 Wilkinson B, Chen JY, Han P, Rufner KM, Goularte OD, Kaye J. TOX: an HMG box protein implicated in the regulation of thymocyte selection. Nat Immunol 2002;3:272-80.

75 Tessema M, Yingling CM, Grimes MJ, Thomas CL, Liu Y, Leng S, Joste N, Belinsky SA. Differential epigenetic regulation of TOX subfamily high mobility group box genes in lung and breast cancers. PLoS One 2012;7:e34850.

76 Yuan SH, Qiu Z, Ghosh A. TOX3 regulates calcium-dependent transcription in neurons. Proc Natl Acad Sci U S A 2009;106:2909-14.

77 Dittmer S, Kovacs Z, Yuan SH, Siszler G, Kogl M, Summer H, Geerts A, Golz S, Shioda T, Methner A. TOX3 is a neuronal survival factor that induces transcription depending on the presence of CITED1 or phosphorylated CREB in the transcriptionally active complex. J Cell Sci 2011;124:252-60.

78 Smid M, Wang Y, Klijn JG, Sieuwerts AM, Zhang Y, Atkins D, Martens JW, Foekens JA. Genes associated with breast cancer metastatic to bone. J Clin Oncol 2006;24:2261-7.

79 Jongstra J, Tidmarsh GF, Jongstra-Bilen J, Davis MM. A new lymphocyte-specific gene which encodes a putative Ca2+-binding protein is not expressed in transformed T lymphocyte lines. J Immunol 1988;141:3999-4004.

80 Jongstra-Bilen J, Young AJ, Chong R, Jongstra J. Human and mouse LSP1 genes code for highly conserved phosphoproteins. J Immunol 1990;144:1104-10.

81 Kadiyala RK, McIntyre BW, Krensky AM. Molecular cloning and characterization

of WP34, a phosphorylated human lymphocyte differentiation and activation antigen. Eur J Immunol 1990;20:2417-23.

82 Jongstra-Bilen J, Jongstra J. Leukocyte-specific protein 1 (LSP1): a regulator of leukocyte emigration in inflammation. Immunol Res 2006;35:65-74.

83 Li Y, Guerrero A, Howard TH. The actin-binding protein, lymphocyte-specific protein 1, is expressed in human leukocytes and human myeloid and lymphoid cell lines. J Immunol 1995;155:3563-9.

84 Pulford K, Jones M, Banham AH, Haralambieva E, Mason DY. Lymphocyte-specific protein 1: a specific marker of human leucocytes. Immunology 1999;96:262-71.

85 Marafioti T, Jabri L, Pulford K, Brousset P, Mason DY, Delsol G. Leucocyte-specific protein (LSP1) in malignant lymphoma and Hodgkin's disease. Br J Haematol 2003;120:671-8.

86 Liu L, Cara DC, Kaur J, Raharjo E, Mullaly SC, Jongstra-Bilen J, Jongstra J, Kubes P. LSP1 is an endothelial gatekeeper of leukocyte transendothelial migration. J Exp Med 2005;201:409-18.

87 Petri B, Kaur J, Long EM, Li H, Parsons SA, Butz S, Phillipson M, Vestweber D, Patel KD, Robbins SM, Kubes P. Endothelial LSP1 is involved in endothelial dome formation, minimizing vascular permeability changes during neutrophil transmigration in vivo. Blood 2011;117:942-52.

88 Giulianelli S, Cerliani JP, Lamb CA, Fabris VT, Bottino MC, Gorostiaga MA, Novaro V, Gongora A, Baldi A, Molinolo A, Lanari C. Carcinoma-associated fibroblasts activate progesterone receptors and induce hormone independent mammary tumor growth: A role for the FGF-2/FGFR-2 axis. Int J Cancer 2008;123:2518-31.

89 Cerliani JP, Guillardoy T, Giulianelli S, Vaque JP, Gutkind JS, Vanzulli SI, Martins R, Zeitlin E, Lamb CA, Lanari C. Interaction between FGFR-2, STAT5, and progesterone

receptors in breast cancer. Cancer Res 2011;71:3720-31.

90 Martin AJ, Grant A, Ashfield AM, Palmer CN, Baker L, Quinlan PR, Purdie CA, Thompson AM, Jordan LB, Berg JN. FGFR2 protein expression in breast cancer: nuclear localisation and correlation with patient genotype. BMC Res Notes 2011;4:72.

91 Sun S, Jiang Y, Zhang G, Song H, Zhang X, Zhang Y, Liang X, Sun Q, Pang D. Increased expression of fibroblastic growth factor receptor 2 is correlated with poor prognosis in patients with breast cancer. J Surg Oncol 2011;105:773-9.

92 Jandrig B, Grade K, Seitz S, Waindzoch B, Muller M, Bender E, Nothnagel A, Rohde K, Schlag PM, Kath R, Hoffken K, Scherneck S. BRCA1 mutations in German breast-cancer families. Int J Cancer 1996;68:188-92.

93 Seitz S, Rohde K, Bender E, Nothnagel A, Kolble K, Schlag PM, Scherneck S. Strong indication for a breast cancer susceptibility gene on chromosome 8p12-p22: linkage analysis in German breast cancer families. Oncogene 1997;14:741-3.

94 Gaffney EV. A cell line (HBL-100) established from human breast milk. Cell Tissue Res 1982;227:563-8.

95 Amadori D, Bertoni L, Flamigni A, Savini S, De Giovanni C, Casanova S, De Paola F, Amadori A, Giulotto E, Zoli W. Establishment and characterization of a new cell line from primary human breast carcinoma. Breast Cancer Res Treat 1993;28:251-60.

96 Savini S, Zoli W, Nanni O, Volpi A, Frassineti GL, Magni E, Flamigni A, Amadori A, Amadori D. In vitro potentiation by lonidamine of the cytotoxic effect of adriamycin on primary and established breast cancer cell lines. Breast Cancer Res Treat 1992;24:27-34.

97 Zoli W, Ricotti L, Barzanti F, Dal Susino M, Frassineti GL, Milandri C, Casadei Giunchi D, Amadori D. Schedule-dependent interaction of doxorubicin, paclitaxel and gemcitabine in human breast cancer cell lines. Int J Cancer 1999;80:413-6.

98 Gioanni J, Le Francois D, Zanghellini E, Mazeau C, Ettore F, Lambert JC,

Schneider M, Dutrillaux B. Establishment and characterisation of a new tumorigenic cell line with a normal karyotype derived from a human breast adenocarcinoma. Br J Cancer 1990;62:8-13.

99 Rae JM, Creighton CJ, Meck JM, Haddad BR, Johnson MD. MDA-MB-435 cells are derived from M14 melanoma cells--a loss for breast cancer, but a boon for melanoma research. Breast Cancer Res Treat 2007;104:13-9.

100 Allred DC, Harvey JM, Berardo M, Clark GM. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. Mod Pathol 1998;11:155-68.

101 Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, Pegram MD, Perez EA, Press MF, Rhodes A, Sturgeon C, Taube SE, Tubbs R, Vance GH, van de Vijver M, Wheeler TM, Hayes DF. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol 2007;25:118-45.

102 McCarty KS, Jr., Miller LS, Cox EB, Konrath J, McCarty KS, Sr. Estrogen receptor analyses. Correlation of biochemical and immunohistochemical methods using monoclonal antireceptor antibodies. Arch Pathol Lab Med 1985;109:716-21.

103 Cohen DA, Dabbs DJ, Cooper KL, Amin M, Jones TE, Jones MW, Chivukula M, Trucco GA, Bhargava R. Interobserver agreement among pathologists for semiquantitative hormone receptor scoring in breast carcinoma. Am J Clin Pathol 2012;138:796-802.

104 Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F, Speed T, Spellman PT, DeVries S, Lapuk A, Wang NJ, Kuo WL, Stilwell JL, Pinkel D, Albertson DG, Waldman FM, McCormick F, Dickson RB, Johnson MD, Lippman M, Ethier S, Gazdar A, Gray JW. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 2006;10:515-27.

105 Kao J, Salari K, Bocanegra M, Choi YL, Girard L, Gandhi J, Kwei KA, Hernandez-Boussard T, Wang P, Gazdar AF, Minna JD, Pollack JR. Molecular profiling of breast

cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. PLoS One 2009;4:e6146.

106 Riaz M, Elstrodt F, Hollestelle A, Dehghan A, Klijn JG, Schutte M. Low-risk susceptibility alleles in 40 human breast cancer cell lines. BMC Cancer 2009;9:236.

107 Subik K, Lee JF, Baxter L, Strzepek T, Costello D, Crowley P, Xing L, Hung MC, Bonfiglio T, Hicks DG, Tang P. The Expression Patterns of ER, PR, HER2, CK5/6, EGFR, Ki-67 and AR by Immunohistochemical Analysis in Breast Cancer Cell Lines. Breast Cancer (Auckl) 2010;4:35-41.

108 Manning DL, Archibald LH, Ow KT. Cloning of estrogen-responsive messenger RNAs in the T-47D human breast cancer cell line. Cancer Res 1990;50:4098-104.

109 Elstrodt F, Hollestelle A, Nagel JH, Gorin M, Wasielewski M, van den Ouweland A, Merajver SD, Ethier SP, Schutte M. BRCA1 mutation analysis of 41 human breast cancer cell lines reveals three new deleterious mutants. Cancer Res 2006;66:41-5.

110 Bane AL, Pinnaduwage D, Colby S, Reedijk M, Egan SE, Bull SB, O'Malley FP, Andrulis IL. Expression profiling of familial breast cancers demonstrates higher expression of FGFR2 in BRCA2-associated tumors. Breast Cancer Res Treat 2009;117:183-91.

111 Cha JY, Lambert QT, Reuther GW, Der CJ. Involvement of fibroblast growth factor receptor 2 isoform switching in mammary oncogenesis. Mol Cancer Res 2008;6:435-45.

112 Itoh H, Hattori Y, Sakamoto H, Ishii H, Kishi T, Sasaki H, Yoshida T, Koono M, Sugimura T, Terada M. Preferential alternative splicing in cancer generates a K-sam messenger RNA with higher transforming activity. Cancer Res 1994;54:3237-41.

113 Tannheimer SL, Rehemtulla A, Ethier SP. Characterization of fibroblast growth factor receptor 2 overexpression in the human breast cancer cell line SUM-52PE. Breast Cancer Res 2000;2:311-20.

114 Behrens C, Lin HY, Lee JJ, Raso MG, Hong WK, Wistuba, II, Lotan R.

Immunohistochemical expression of basic fibroblast growth factor and fibroblast growth factor receptors 1 and 2 in the pathogenesis of lung cancer. Clin Cancer Res 2008;14:6014-22.

115 Hughes SE. Differential expression of the fibroblast growth factor receptor (FGFR) multigene family in normal human adult tissues. J Histochem Cytochem 1997;45:1005-19.

116 Johnston CL, Cox HC, Gomm JJ, Coombes RC. Fibroblast growth factor receptors (FGFRs) localize in different cellular compartments. A splice variant of FGFR-3 localizes to the nucleus. J Biol Chem 1995;270:30643-50.

117 Ahmad I, Iwata T, Leung HY. Mechanisms of FGFR-mediated carcinogenesis. Biochim Biophys Acta 2012;1823:850-60.

118 Penault-Llorca F, Bertucci F, Adelaide J, Parc P, Coulier F, Jacquemier J, Birnbaum D, deLapeyriere O. Expression of FGF and FGF receptor genes in human breast cancer. Int J Cancer 1995;61:170-6.

119 Bansal GS, Cox HC, Marsh S, Gomm JJ, Yiangou C, Luqmani Y, Coombes RC, Johnston CL. Expression of keratinocyte growth factor and its receptor in human breast cancer. Br J Cancer 1997;75:1567-74.

120 Tozlu S, Girault I, Vacher S, Vendrell J, Andrieu C, Spyratos F, Cohen P, Lidereau R, Bieche I. Identification of novel genes that co-cluster with estrogen receptor alpha in breast tumor biopsy specimens, using a large-scale real-time reverse transcription-PCR approach. Endocr Relat Cancer 2006;13:1109-20.

121 Luqmani YA, Graham M, Coombes RC. Expression of basic fibroblast growth factor, FGFR1 and FGFR2 in normal and malignant human breast, and comparison with other normal tissues. Br J Cancer 1992;66:273-80.

122 Luqmani YA, Bansal GS, Mortimer C, Buluwela L, Coombes RC. Expression of FGFR2 BEK and K-SAM mRNA variants in normal and malignant human breast. Eur J Cancer 1996;32A:518-24.

123 Zang XP, Pento JT. Keratinocyte growth factor-induced motility of breast cancer cells. Clin Exp Metastasis 2000;18:573-80.

124 Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, Brodsky AS, Keeton EK, Fertuck KC, Hall GF, Wang Q, Bekiranov S, Sementchenko V, Fox EA, Silver PA, Gingeras TR, Liu XS, Brown M. Genome-wide analysis of estrogen receptor binding sites. Nat Genet 2006;38:1289-97.

125 Turner N, Lambros MB, Horlings HM, Pearson A, Sharpe R, Natrajan R, Geyer FC, van Kouwenhove M, Kreike B, Mackay A, Ashworth A, van de Vijver MJ, Reis-Filho JS. Integrative molecular profiling of triple negative breast cancers identifies amplicon drivers and potential therapeutic targets. Oncogene 2010;29:2013-23.

126 Vermeulen JF, Kornegoor R, van der Wall E, van der Groep P, van Diest PJ. Differential expression of growth factor receptors and membrane-bound tumor markers for imaging in male and female breast cancer. PLoS One 2013;8:e53353.

127 Ngo N, Patel K, Isaacson PG, Naresh KN. Leucocyte common antigen (CD45) and CD5 positivity in an "undifferentiated" carcinoma: a potential diagnostic pitfall. J Clin Pathol 2007;60:936-8.

128 Schirosi L, Nannini N, Nicoli D, Cavazza A, Valli R, Buti S, Garagnani L, Sartori G, Calabrese F, Marchetti A, Buttitta F, Felicioni L, Migaldi M, Rea F, Di Chiara F, Mengoli MC, Rossi G. Activating c-KIT mutations in a subset of thymic carcinoma and response to different c-KIT inhibitors. Ann Oncol 2012.

Affidavit

"I, [Juan Luo] certify under penalty of perjury by my own signature that I have submitted the thesis on the topic [Expression statuses of FGFR2, TOX3 and LSP1 proteins in human breast cancer tissues and cell lines and the associations with clinical and pathological characteristics] I wrote this thesis independently and without assistance from third parties, I used no other aids than the listed sources and resources.

All points based literally or in spirit on publications or presentations of other authors are, as such, in proper citations (see "uniform requirements for manuscripts (URM)" the ICMJE www.icmje.org) indicated. The sections on methodology (in particular practical work, laboratory requirements, statistical processing) and results (in particular images, graphics and tables) correspond to the URM (s.o) and are answered by me. My interest in any publications to this dissertation corresponds to those that are specified in the following joint declaration with the responsible person and supervisor. All publications resulting from this thesis and which I am author correspond to the URM (see above) and I am solely responsible.

The importance of this affidavit and the criminal consequences of a false affidavit (section 156,161 of the Criminal Code) are known to me and I understand the rights and responsibilities stated therein.

Date Signature

Declaration of any eventual publications

[Name of the doctoral candidates] had the following share in the following publications:

Publication 1: [Authors], [titles], [magazine], [year of publication]
Contribution in detail (please briefly explain):

Publication 2: [Authors], [titles], [magazine], [year of publication]
Contribution in detail (please briefly explain):

Publication 3: [Authors], [titles], [magazine], [year of publication]
Contribution in detail (please briefly explain):

Signature, date and stamp of the supervising University teacher

_____

Signature of the doctoral candidate

_____

My curriculum vitae does not appear in the electronic version of my paper for reasons of data protection.

Acknowledgments

I would like to express my gratitude to my supervisor Prof. Dr. Dr. Hermann Lage for offering me the opportunity and support to carry out this experimental and clinical research on breast cancer, and for his continuous guidance and encouragement throughout my study. Also, his critical reading of my manuscript, remarkable insights and thoughtful comments are highly appreciated.

My further acknowledgements go to Mrs. Manuela Pacyna-Gengelbach who helped me with cell culture and cell block preparation and other technical support, and to Dr. med. Alaa who helped me prepare for the experiment.

I would also like to show my special gratitude to the Immunohistochemistry Lab in the Institute of Pathology for their skilled work in sectioning a plenty of slides and staining ER, PR and HER2 for me, and to Prof. Dr. med. Frank Heppner and PD Dr. med. Arend Koch and their colleagues in the Institute of Neuropathology for supplying me with high quality brain samples for control. I also sincerely thank OA Dr. med. Ruza Arsenic for her professional work in scoring the immunostaining of my abundant samples. Additionally, I show my sincere appreciation to Mr. Torsch and Mr. John for their great help in photographing of a plenty of samples.

Moreover, I would like to thank Prof. Dr. med. Manfred Dietel and Dr. rer. nat. Alexandra Stege, without their help, I could not get the chance and obtain the clinical tissue samples and data for my research. Also, a big “Thank you!” goes to Frau Hermann and her colleagues in the Archive.

I especially thank all other colleagues and students in AG Lage for being such a pleasant group and letting me have such a comfortable and informative time! Further, my hearty thankfulness goes to Mrs. Pamela Glowacki and Mrs. Zhonghua Helmke in the International Cooperation Office at Charité and all my Chinese friends who gave me a lot of help in my work and my living in Berlin.

Finally I own my sincere appreciation to my parents, my brother, my sister-in-law, my niece and especially my boyfriend, now my husband and my son’s father, for their love and encouragement throughout my whole study in Berlin. I also thank my pre-hospital in China, Zhejiang Cancer Hospital, for its financial support.