fortune Journals

Research Article

JOURNAL OF CANCER SCIENCE AND CLINICAL THERAPEUTICS

ISSN: 2637-5079

Prediction of Intraductal Carcinomas of the Breast that will not evolve to Invasive Carcinomas

Felip Vilardell[1*], Edelmiro Iglesias[2], Carles Canosa[2], Izaskun Urdanibia[3] and Montserrat Martínez[4]

Abstract

Ductal carcinoma in situ (DCIS) of the breast is a malignant neoplastic proliferation of cells with a ductal phenotype, which is confined to the epithelial compartment of ducts and acini. DCIS is regarded as a nonobligate precursor of invasive ductal carcinoma (IDC) of the breast. Currently it is not possible to predict which DCIS will progress to IDC, and so the vast majority of patients with DCIS undergo surgical resection and radiation therapy. DCIS and IDC share most of the genetic alterations, including somatic mutations, genetic expression profiles and genetic copy number alterations. Our aim was to find a genetic mutational profile that could help to predict which DCIS will progress to IDC and which will not. Individuals with a predicted probability of DCIS synchronous with IDC of 0.5 or lower were found to be those with a mutation in 38138689 (POSTN), or no mutation in either 151851109 (KMT2C), 187524525 (FAT1) or 55139771 (PDGFRA).

Keywords: Breast Cancer; Carcinoma in situ; Invasive carcinoma; Surgical Pathology; Gene mutations.

Introduction

DCIS is the most common type of non-invasive breast cancer and forms a group of heterogeneous lesions with malignant potential that includes premalignant and breast preinvasive lesions and to which different treatment options are applied that are usually excessive in most cases. DCIS is a proliferation of neoplastic ductal epithelial cells that are confined to the ductal-lobular system of the mammary gland. DCIS is a non-obligate precursor lesion of invasive breast cancer. It is when the DCIS breaks through and crosses the basement membrane and affects adjacent parenchyma that it becomes a carcinoma with some invasive characteristics. In 1932, Broders [1] categorized DCIS lesions as carcinogenic cells that did not cross the basement membrane of the tissues. Later, Tavassoli [2] and Norris differentiated it from atypical ductal hyperplasia when its size exceeds 2 mm. Before breast cancer early detection programs became widespread, the prevalence of preinvasive lesions such as DCIS was close to 5% at most. In clinically diagnosed cases, their most common form of presentation was in palpable mass form, unilateral discharge of clear or blood-stained fluid from a single milk duct of the nipple, or the association of both. In other cases, nipple eczema appeared to be associated with Paget's disease. After the commencement of breast cancer screening programs, there has been a tendency towards overdiagnosis of DCIS and hence its overtreatment, even though in up to 90% of cases it is clinically undetectable. The most common presentation form, at present, is with the appearance of atypical microcalcifications [3] in the screening

Affiliation:

[1]Department of Pathology and Biomedical Research Institute of Lleida, Hospital Universitari Arnau de Vilanova, Alcalde Rovira Roure 80, 25198, Lleida (Spain)

[2]Department of Surgery and Biomedical Research Institute of Lleida, Hospital Universitari Arnau de Vilanova, Alcalde Rovira Roure 80, 25198, Lleida (Spain)

[3]Oncological Pathology Group, Biomedical Research Institute of Lleida, Alcalde Rovira Roure 80, 25198, Lleida (Spain)

[4]Biostatistics Unit, Biomedical Research Institute of Lleida, Alcalde Rovira Roure 80, 25198, Lleida (Spain)

***Corresponding author:**

Felip Vilardell, Department of Pathology and Biomedical Research Institute of Lleida, Hospital Universitari Arnau de Vilanova, Alcalde Rovira Roure 80, 25198, Lleida (Spain).

Citation: Felip Vilardell, Edelmiro Iglesias, Carles Canosa, Izaskun Urdanibia and Montserrat Martínez. Prediction of Intraductal Carcinomas of the Breast that will not evolve to Invasive Carcinomas. Journal of Cancer Science and Clinical Therapeutics. 8 (2024): 177-188..

Received: June 23, 2024
Accepted: July 07, 2024
Published: July 23, 2024

Vilardell F, et al., J Cancer Sci Clin Ther 2024
DOI:10.26502/jcsct.5079238

mammography and in 30-40% of cases with comedonecrosis. For this reason, the incidence of DCIS has been gradually increasing, reaching age-adjusted rates of around 20-25%, and in some cases up to 30%. This increased diagnosis of DCIS lesions has led in turn to an increase in the treatment of lesions that would not have been detected clinically and which, in many cases, would not have required medical intervention. According to autopsy reviews, the prevalence of asymptomatic DCIS ranges from 7-39% in women over the age of 30 [4]. This leads us to consider that a subgroup of patients are being unnecessarily subjected to surgery, radiation therapy or hormone treatment. This excess or overtreatment, which could have both physical and psychological consequences for patients, could be avoided in a significant percentage of cases. There is no consensus on the factors that are useful in determining which DCIS are most likely to develop into invasive carcinoma. According to some authors, the nuclear grade is not significant [5]. Some studies have considered the possibility of de-escalation in the treatment of DCIS lesions, based on stratification of the potential risk of transformation of the DCIS lesion to an infiltrating lesion with the same molecular alterations [6]. For low-risk lesions, certain treatments could therefore be avoided or, failing that, de-escalation in the treatment applied [7]. Such studies are usually based on the patient's clinical characteristics and a histological study of diagnosed DCIS lesions. Although no classification enjoys widespread acceptance, the two most widely used and with the highest consensus are the Van Nuys prognostic index [8] and the MSKCC (Memorial Sloan Kettering Cancer Center) nomogram [9].

In addition, some genetic platforms can be used to predict the risk of local recurrence at age 10, as well as recurrences in their invasive form after DCIS surgery [10, 11]. From these recurrences, it is known that approximately 50% will be of another DCIS [12] and will not undergo a transformation to an infiltrating component. In other clinical trials, such as COMET, LORD and LORIS [13], de-escalation in DCIS treatment in low-risk cases has been evaluated by omitting surgery and radiotherapy. There are still no conclusions with a sufficient level of scientific evidence. In view of the above, it is very important to find a methodology that will allow more accurate and reliable reporting to patients diagnosed with DCIS and is more helpful in making the most appropriate therapeutic decisions. The study we conducted seeks to establish a correlation between the different tumor mutations found in lesions with DCIS and the risk of progression to an infiltrating carcinoma. The most frequently mutated genes in breast cancer are *PIK3CA, MYC, CCND1, ERBB2, chr8:ZNF703/FGFR1 locus* and *MAP3K*1 among the oncogenes, and *TP53, PTEN* and *GATA3* among the tumor suppressor genes [14, 15, 16]. DCIS and IDC share most of the genetic alterations, including somatic point mutations, genetic expression profiles and genetic copy number alterations (CNAs) [17, 18]. In fact, the CNAs of DCIS with synchronous IDC are closer to those of IDC than CNAs of pure DCIS [19]. The genetic alterations responsible for the transition from DCIS to IDC are not well known. It is currently not possible to predict which DCIS will progress to IDC, and therefore the vast majority of patients with DCIS are subjected to surgical resection and radiation therapy. Kim SY et al. [19] sought genomic differences between pure DCIS and synchronous DCIS with IDC (S DCIS) comparing fresh samples of pure DCIS obtained from 6 patients with IDC and synchronous DCIS of 5 patients. Although they found a statistically significant difference in the mutational status of 13 genes of the Cancer Gene Census [20] between pure DCIS and DCIS associated with IDC, the study was based on a very limited number of cases and, moreover, the six cases of pure DCIS analyzed were low grade and with a low Ki-67 proliferative index. With the aim of finding a genetic mutational profile that could help to predict which DCIS will progress to IDC, we designed a panel with 40 selected genes including FGFR2, BRCA2, MET, AR, GNAS, NCOA3, PDGFRA, ATM, BCOR, MLL3, NOTCH1, AKT1, ALK, MDM2, MYCL1, MYCN, CDKN2C, GATA3, MAP3K1, NOTCH2, PIK3R1, SMARCA4, FANCE, MKL1, FAT1, DST, TMEM45A, SOX9, MMP8, MMP27, ADAM7, ADAM12, ADAM29, EPHA1, DCLK3, PTPRB, MMP24, MMP25, OMD and POSTN. This selection of genes was based on the results reported by Kim SY et al. [19], together with some genes encoding basement membrane-degrading proteins or involved in epithelial-mesenchymal transition.

Materials & Methods

Tissue samples

Human tissues belonging to 67 patients comprising 37 diagnosed with pure DCIS from year 2011 through 2016 and 30 with extensive DCIS synchronous with IDC were obtained from the pathology service of the University Hospital Arnau de Vilanova (HUAV-L) in Lleida (Spain). The study was approved by the clinical research ethics committee of the HUAV-L (CEIC-2166). With regard to the cases with DCIS and synchronous IDC, cases of patients with extensive peripheral carcinoma were chosen to facilitate macrodissection of the invading and in situ components separately. Pathologic assessment of tumor content was performed by a senior pathologist (PhD) with the use of diagnostic hematoxylin and eosin slides. Clinical and pathologic features of the breast tumor lesions are summarized in Table 1. A total of sixty-seven (67) clinical tumor specimens were analyzed in this study. The specimens consisted of formalin-fixed, paraffin-embedded tissue. Of the 67 clinical specimens, 11 did not pass sample quality control (7 failed real-time PCR DNA quality control; 4 failed quality control after library preparation). The successful tumor specimens included pure DCIS belonging to 28 patients, IDC belonging to 25 patients and synchronous DCIS matched with those 25 IDCs.

Vilardell F, et al., J Cancer Sci Clin Ther 2024
DOI:10.26502/jcsct.5079238

Table 1: Clinicopathologic features of the patients

Case	Tumor	Grade	ER	PR	Ki-67	HER2	DFS	Time DFS	OS	Time OS
1	DCIS	High	285	35	45		0		0	
2	DCIS	High	240	0	25		0		0	
3	DCIS	High	250	35	16		0		0	
4	DCIS	Medium	250	35	16		0		0	
5	DCIS	High	240	10	16		0		0	
6	DCIS	High	0	0	20		0		0	
7	DCIS	High	200	10	20		0		0	
8	DCIS	High	0	0	40		0		0	
9	DCIS	Low	290	290	9		0		0	
10	DCIS	High	0	0	40		0		1	
11	DCIS	High	280	100	12		0		0	
12	DCIS	High	270	270	15		0		0	
13	DCIS	Low	300	110	21		0		0	
14	DCIS	High	230	250	16		0		0	
15	DCIS	High	295	180	20		0		0	
16	DCIS	High	210	140	20		0		0	
17	DCIS	High	180	0	16		0		0	
18	DCIS	High	260	15	35		0		0	
19	DCIS	High	295	295	20		0		0	
20	DCIS	High	240	50	60		0			
21	DCIS	High	300	40	10		0		0	
22	DCIS	High	280	Re0	40		0		0	
23	DCIS	High	290	0	65		0		0	
24	DCIS	High	280	90	40		0		1	
25	DCIS	High	0	0	40		0		0	
26	DCIS	High	300	290	10		0		0	
27	DCIS	High	0	0	25		0		1	
28	DCIS	High	180	0	5		0		0	
1	S DCIS	Medium								
1	IDC	2	280	240	15	1	0		0	
2	S DCIS	High								
2	IDC	2	0	0	60	1	0		0	
3	S DCIS	Medium								
3	IDC	2	295	140	20	2+	0		0	
4	S DCIS	High								
4	IDC	3	0	0	45	3	1	40	1	46
5	S DCIS	Medium								
5	IDC	1	250	115	5	0	0		0	
6	S DCIS	High								
6	IDC	3	0	0	35	1	0		1	144
7	S DCIS	High								
7	IDC	2	0	0	20	1	0		1	112
8	S DCIS	Low								
8	IDC	1	270	260	18	0	0		0	

Vilardell F, et al., J Cancer Sci Clin Ther 2024
DOI:10.26502/jcsct.5079238

9	S DCIS	Low								
9	IDC	1	300	220	14	1	0		0	
10	S DCIS	High								
10	IDC	3	0	0	35	1	0		0	
11	S DCIS	High								
11	IDC	2	290	30	30	1	0		0	
12	S DCIS	Low								
12	IDC	1	290	260	12	1	0		0	
13	S DCIS	Low								
13	IDC	1	300	240	15	1	0		0	
14	S DCIS	High								
14	IDC	1	0	0	20	0	0		0	
15	S DCIS	High								
15	IDC	3	0	0	60	3	0		1	56
16	S DCIS	High								
16	IDC	3	0	0	50	3	1		1	216
17	S DCIS	Medium								
17	IDC	1	250	0	10	0	0		0	
18	S DCIS	Medium								
18	IDC	1	230	70	8	1	0		0	
19	S DCIS	High								
19	IDC	2	160	60	26	2	0		0	
20	S DCIS	Medium								
20	IDC	2	300	230	15	2	0		0	
21	S DCIS	High								
21	IDC	3	295	270	80	3	0		0	
22	S DCIS	High								
22	IDC	3	0	0	27	3	0		0	
23	S DCIS	High								
23	IDC	2	290	110	41	0	0		0	
24	S DCIS	High								
24	IDC	3	0	0	31	3	0		0	
25	S DCIS	Medium								
25	IDC	1	300	230	24	1	0		0	

The first 28 rows show the pure DCIS, and the following rows the 25 patients with matched samples of S DCIS and invasive carcinoma. DCIS: ductal carcinoma in situ, S DCIS: ductal carcinoma in situ synchronous with invasive ductal carcinoma, IDC: invasive ductal carcinoma, ER: estrogen receptor, PR: progesterone receptor, DFS: Diseases free survival, OS: overall survival.

DNA extraction and quality assessment

Ten sections of 10 µm thickness were cut from formalin-fixed and paraffin-embedded (FFPE) tissue blocks. Areas with a high percentage of neoplastic cells (at least approximately 80%) were scraped using a serial hematoxylin and eosin-stained section as a guide. Paraffin was removed with xylene, and the tissue was washed with 100% ethanol. The FFPE tissue DNA was extracted according to the manufacturer's instructions using the Cobas® DNA Sample Preparation Kit (Ref: 05985536190, Roche Molecular Systems, Inc) and was eluted in a 50 µL volume. The extracted DNA was quantified using the Qubit dsDNA HS assay (Ref: Q32851, Life Technologies). DNA quality was assessed using the Infinium HD FFPE QC Kit (Ref: WG-321-1001, Illumina) according to the manufacturer's instructions; samples with changes in cycle threshold (ΔCt) value of 2 or less (versus a control) were considered acceptable for sequencing.

Vilardell F, et al., J Cancer Sci Clin Ther 2024
DOI:10.26502/jcsct.5079238

Library preparation and quality control

The extracted and quantified FFPE DNA was subsequently amplified in a multiplex PCR using an AmpliSeq Custom Panel from Illumina. The amplicon sizes were 125–175 bps and were especially designed for FFPE material with the DesignStudio (Illumina), a web-based assay design tool. The panel contains probes to generate amplicons from 40 cancer-related genes: FGFR2, BRCA2, MET, AR, GNAS, NCOA3, PDGFRA, ATM, BCOR, MLL3, NOTCH1, AKT1, ALK, MDM2, MYCL1, MYCN, CDKN2C, GATA3, MAP3K1, NOTCH2, PIK3R1, SMARCA4, FANCE, MKL1, FAT1, DST, TMEM45A, SOX9, MMP8, MMP27, ADAM7, ADAM12, ADAM29, EPHA1, DCLK3, PTPRB, MMP24, MMP25, OMD and POSTN. Library preparation used 100 ng of input DNA with AmpliSeq Custom Panel (Ref: 20020495, Illumina), the AmpliSeq Library PLUS for Illumina (Ref: 20019102, Illumina), and the AmpliSeq CD Indexes Set A for Illumina (Ref: 20019105, Illumina) according to the manufacturer's protocol. Barcoded libraries were purified using AMPure Beads XP (Ref: A63881, Beckman Coulter) and amplified. After a second round of purification, the libraries were evaluated on an Agilent 2100 Bioanalyzer using the High Sensitivity DNA Kit (Ref: 5067-4626, Agilent Technologies) and quantified using a Qubit dsDNA HS Assay Kit (Ref: Q32851, Life Technologies) to assess successful enrichment and amplification. In 56 of 60 cases, the DNA library was successfully synthesized from FFPE-derived DNA. Libraries were then normalized and pooled for sequencing according to the manufacturer's protocol. Each pooled library was sequenced on a MiSeq instrument using a 2 × 150 paired-end sequencing design with MiSeq Reagent Kit v3 (600-cycle) (Ref: MS-102-3003, Illumina), achieving an amplicon mean depth of coverage ranging from ×535 to ×2427.

NGS data analysis

Samples tested on the AmpliSeq Custom Panel from Illumina were analyzed by the BaseSpace DNA amplicon v2.1.1 (Illumina). FASTQ files were uploaded to the BaseSpace Sequencing Hub and analyzed by DNA amplicon v2.1.1 (Illumina). Alignment was done against the human reference sequence build GRCh37/Hg19 and variant calling using the default somatic sensitivity parameters. The Bam files were loaded in the Integrative Genomics Viewer (IGV) to visualize variants against the reference genome. Annotated variants were filtered in Excel removing variants outside coding regions (that do not occur in exons) or splice sites, in 3′UTR and 5′UTR regions, synonymous variants and removing variants with minor allele frequency (MAF) >1%. Variant calls with <×100 depth of coverage, or those failing any filtering criteria established by the DNA amplicon pipeline including allele frequency <5%, quality score <30, or significant strand bias, were discarded. The filtered variants were interrogated in different databases that previously identified pathogenic variants. Those databases are ClinVar, Variant Interpreter (Illumina), Ensembl, Combined Annotation Dependent Depletion (CADD), Genome Aggregation Database (gnomAD), Cancer Genome Atlas (TCGA) and Catalog of Somatic Mutations in Cancer (COSMIC). Variants with unreported pathogenicity profile were assessed using various in silico prediction models (SIFT, PolyPhen2, Fathmm-XF, MutationTaster) to identify deleterious/pathogenic mutations. Based on the results of the variant annotations, pathogenic variants of interest and candidate genes were identified.

Statistical analysis

Bivariate analysis included analysis of the number of mutations (in total, by gene and by type of mutation) and the presence of each mutation in patients with pure DCIS compared to those with IDC and S DCIS using Student's t-test or the Mann-Whitney-test to compare quantitative variables with normal distribution or not, respectively. We used the Pearson's χ-squared test for qualitative variables. Additionally, the matched comparison of the IDC and S DCIS samples was performed using the McNemar test for each mutation and the Wilcoxon signed-rank test for the number of mutations (in total, by gene and by type of mutation).

The multivariate analysis to discriminate between pure DCIS and IDC and between pure DCIS and S DCIS with IDC included a preliminary assessment of the importance of each mutation by applying the Boruta algorithm, which allows selecting the subgroup of mutations that are confirmed as important beyond chance. Starting in the subgroup of important mutations for pure DCIS versus IDC, the construction of a conditional inference classification tree based on multiplicity-adjusted p-values by the MonteCarlo test was performed, with a minimum node size of five individuals. The same analysis was applied to the subset of important mutations for pure DCIS versus S DCIS with IDC. The evaluation of the area under the ROC curve (AUC) defined by the estimated probabilities at each node were included for both trees, as well as the sensitivity and specificity of probabilities greater than 0.5.

Finally, two random forest analyses were carried out to estimate the probabilities of each individual from the results of all the trees where they have not participated (out-of-bag prediction), by means of 401 classification trees built from resampling with replacement and using the subgroup of important mutations identified by the Boruta algorithm and with at least five individuals per node. Of the two random forests, the evaluation of the area under the ROC curve defined by the estimated probabilities, as well as the sensitivity and specificity of probabilities greater than 0.5 were included. In addition, the Gardner-Altman plot for the difference in predictions between the two groups in comparison was estimated by bias-corrected and accelerated bootstrap analysis of 5000 resamples.

Vilardell F, et al., J Cancer Sci Clin Ther 2024
DOI:10.26502/jcsct.5079238

Results

Of the 40 genes included in the customized panel, mutations were found in 28. More specifically, mutations were found in MYCL, NOTCH2, ALK, TMEM45A, PDGFRA, ADAM29, FAT1, MAP3K1, PIK3R1, FANCE, DST, MET, EPHA1, KMT2C, NOTCH1, ADAM12, MMP27, MMP8, ATM, PTPRB, BRCA2, POSTN, AKT1, SMARCA4, NCOA3, MKL1, BCOR and AR (Table 2, Figure 1). The number of mutated genes per patient ranged from 11 to 22 among patients with pure DCIS, and from 13 to 20 mutated genes per patient in patients with IDC and S DCIS. The total number of genetic mutations ranged from 27 to 57 in pure DCIS, from 28 to 58 in IDC, and from 31 to 58 in S DCIS. Mutations considered pathogenic in the Cancer Gene Census were found in 12 of the 28 above mentioned genes. Table 3 shows their frequency by kind of lesion. The multivariate analysis to discriminate between pure DCIS and IDC and between pure DCIS and S DCIS with IDC began with the Boruta algorithm to assess the importance of each mutation. The discriminative mutations between IDC and DCIS were the missense variant 151851109 in lysine N-methyltransferase 2C (KMT2C) gene (G➔A transition, not reported at Ensembl data base), the missense variant 187524525 in protocadherin FAT1 gene (C➔T transition, pVal3719Met) and the missense variant 38138689 in periostin POSTN gene (C ➔ T, pVal814Met) (Figure 2). The mutation 151851109 was present in 5/28 (17.9%) patients with pure DCIS, 13/25 (52%) patients with S DCIS and 12/25 (48%) patients with IDC. The conditional inference tree for the discrimination between IDC and DCIS using these three important mutations (Figure 4a) showed that the presence of the 187524525 mutation in FAT1 was predictive of IDC with a probability of 1 (5/5 patients in node 7). In absence of this mutation, the presence of the 38138689 mutation in POSTN gene predicts 0 probability of having IDC (0/6 patients in node 3). In absence of the two former mutations, the presence of the 151851109 mutation in KMT2C gene was predictive of IDC with a probability of 0.7 (8/11 patients with IDC in node 5). In the absence of this last mutation, the probability of IDC is 0.387 (12/31 patients in node 6). By means of random forests, 401 trees were estimated from resampled datasets and employing the three confirmed important mutations 151851109, 187524525 and 38138689. The results are shown in Figure 4b and Figure 4c. With probabilities higher than 0.5 for IDC, we obtained a sensitivity of 52% (13/25) and a specificity of 89.3% (25/28). The discriminative mutations between pure DCIS and S DCIS associated with IDC (Figure 3) were the previous three plus the missense variant 55139771 in platelet-derived growth factor receptor alpha (PDGFRA) gene (T➔C transition, COSM5008347). This mutation was present in 2/28 (7.14%) patients with pure DCIS, 7/25 (28%) patients with S DCIS and 7/25 (28%) patients with IDC. The conditional inference tree for the discrimination between pure DCIS and S DCIS using the four important mutations (Figure 4d) showed that the presence of the 151851109 mutation in KMT2C was predictive of S DCIS with a probability of 0.722 (13/18 patients in node 2). In absence of this mutation, the presence of the 55139771 mutation in PDGFRA gene predicts a 0.714 probability of having S DCIS (5/7 patients in node 4). In absence of the two former mutations, the probability of S DCIS is 0.25 (7/28 patients in node 5). With probabilities higher than 0.5 for S DCIS, we obtained a sensitivity of 72% (18/25) and a specificity of 75% (21/28). By means of random forests, 401 trees were estimated from resampled datasets and employing the four confirmed important mutations 151851109, 187524525, 38138689 and 55139771. Figure 4e and Figure 4f show the sorted predicted probabilities of S DCIS for each patient in this study. Among the 25 patients with S DCIS, 20 were correctly predicted by the random forest analysis. With probabilities higher than 0.5 for S DCIS, we obtained a sensitivity of 80% (20/25) and a specificity of 82.8% (23/28). More specifically, individuals with a predicted probability of having a DCIS synchronous with IDC higher than 0.5, are those with at least a mutation in 151851109 (KMT2C gene), 187524525 (FAT1 gene) or 55139771 (PDGFRA gene) and no mutation in 38138689 (POSTN gene). Inversely, individuals with a predicted probability of DCIS S of 0.5 or lower are those with a mutation in 38138689 or no mutation in either 151851109, 187524525 or 55139771.

Table 2: Total genes with mutations found, number of mutations and variant effect

	Consequence						
Gene	**Missense**	**Frameshift variant**	**Inframe deletion**	**Stop codon**	**Splicing**	**Number of mutations**	**%**
MYCL	1					1	0.92
NOTCH2	3					3	2.75
ALK	4				2	6	5.5
TMEM45A	1					1	0.92
PDGFRA	1					1	0.92
ADAM29			1			1	0.92
FAT1	17					17	15.6
MAP3K1	3					3	2.75

Vilardell F, et al., J Cancer Sci Clin Ther 2024
DOI:10.26502/jcsct.5079238

PIK3R1	1					1	0.92
FANCE	1					1	0.92
DST	8				1	9	8.26
MET	1					1	0.92
EPHA1	2					2	1.83
KMT2C	13			3		16	14.68
NOTCH1	2				2	4	3.67
ADAM12	2					2	1.83
MMP27	5					5	4.59
MMP8	3					3	2.75
ATM	4				2	6	5.5
PTPRB	7				2	9	8.26
BRCA2	4				1	5	4.59
POSTN	1					1	0.92
AKT1	1					1	0.92
SMARCA4	2					2	1.83
NCOA3	2		1		1	4	3.67
MKL1	1					1	0.92
BCOR	1				1	2	1.83
AR			1			1	0.92
Total	**91**	**0**	**3**	**3**	**12**	**109**	

Table 3: Frequency of pathogenic mutations

Gene	Position	% T	% P	% S	% I
NOTCH2	120478125	1.81	0	3.84	3.84
TMEM45A	100295801	83.63	79.31	70.37	81.48
FAT1	187517873	14.54	10.34	7.41	18.52
	187524525	9.09	0	19.23	19.23
	187542755	21.81	17.24	25.93	29.63
DST	56335937	3.63	3.45	3.7	3.7
	56373536	7.27	6.9	7.41	7.41
MET	116340086	3.63	6.9	0	0
KMT2C	151927016	34.54	31.03	38.46	34.61
	151927021	98.18	95.55	100	100
	151927025	87.27	86.21	88.46	88.46
	151945007	96.39	89.66	100	100
	151962168	15.66	17.24	14.81	14.81
	151962265	71.08	68.97	74.07	70.37
	151970856	95.2	93.1	100	92.59
	151970859	61.45	65.52	62.96	55.56
	151970877	85.54	75.86	96.3	85.19
NOTCH1	139396319	71.08	72.41	74.07	66.67

Vilardell F, et al., J Cancer Sci Clin Ther 2024
DOI:10.26502/jcsct.5079238

MMP27	102576382	7.23	6.9	7.41	7.41
PTPRB	70925864	14.46	13.79	14.81	14.81
	70954547	3.63	0	7.69	7.69
	70965009	6.02	3.45	7.41	7.41
	70984019	37.35	44.83	33.33	33.33
POSTN	38138689	10.9	20.69	0	0
AKT1	105246551	4.82	3.45	7.41	3.7
NCOA3	46256424	18.07	10.34	22.22	22.22
	46264888	27.71	27.59	29.63	25.93

T= Total, P= pure DCIS, S= synchronous DCIS, I= IDC.

Figure 1: Number and type of mutations per gene, in the 28 genes with found mutations.

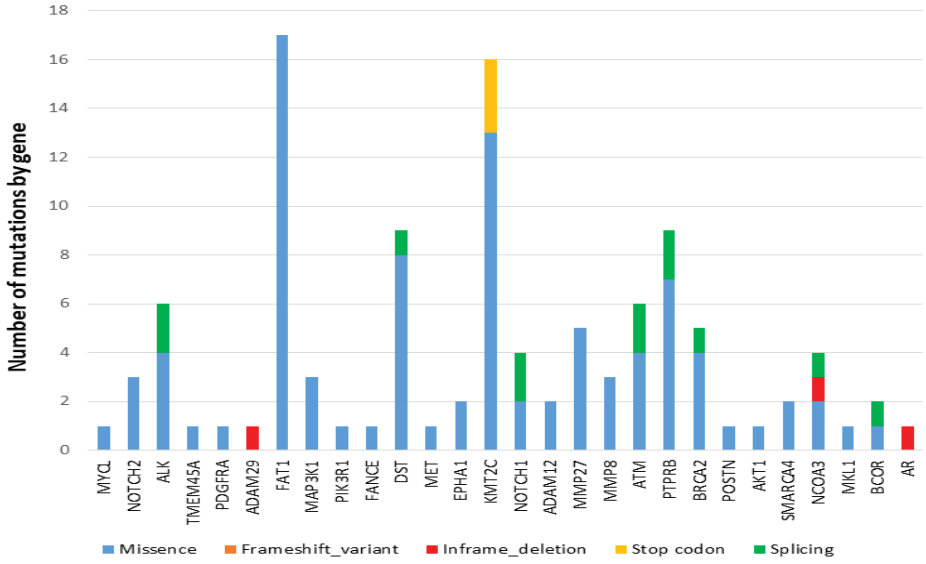
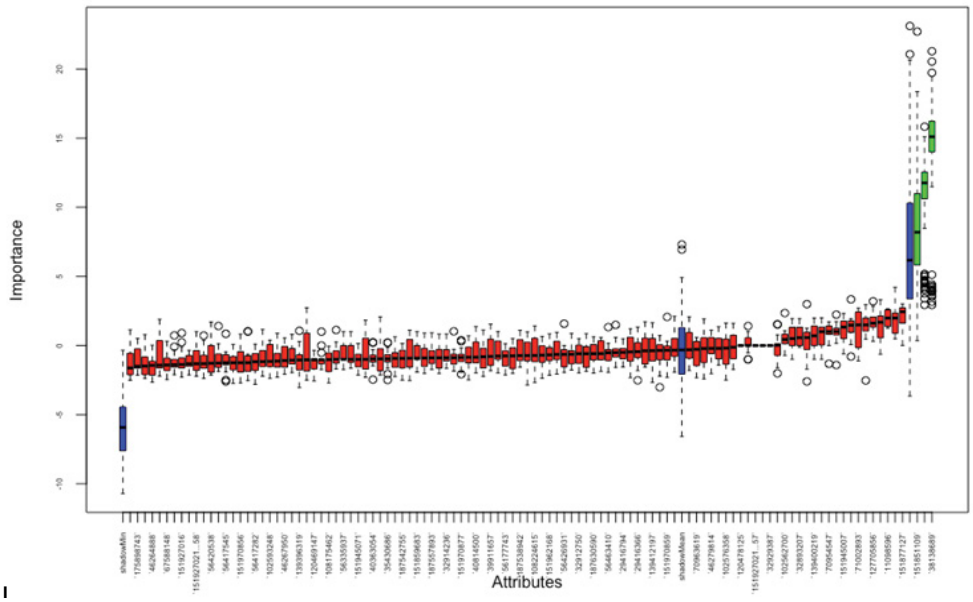


Figure 2: The discriminatory mutations between IDC and pure DCIS are '151851109' (KMT2C gene, 'benign' variant), '187524525' (FAT1 gene, 'possibly pathogenic' variant) and '38138689' (POSTN gene, 'contradictory' variant)



Citation: Felip Vilardell, Edelmiro Iglesias, Carles Canosa, Izaskun Urdanibia and Montserrat Martínez. Prediction of Intraductal Carcinomas of the Breast that will not evolve to Invasive Carcinomas. Journal of Cancer Science and Clinical Therapeutics. 8 (2024): 177-188.

Vilardell F, et al., J Cancer Sci Clin Ther 2024
DOI:10.26502/jcsct.5079238

Unadjusted estimated OR for the selected mutations (IDC vs. DCIS)

	DCIS	IDC	OR	p.ratio
	N=28	***N=25***		
151851109	5 (17.9%)	12 (48.0%)	4.07 [1.20;15.7]	0.023
187524525	0 (0.00%)	5 (20.0%)	. [.;.]	.
38138689	6 (21.4%)	0 (0.00%)	. [.;.]	.

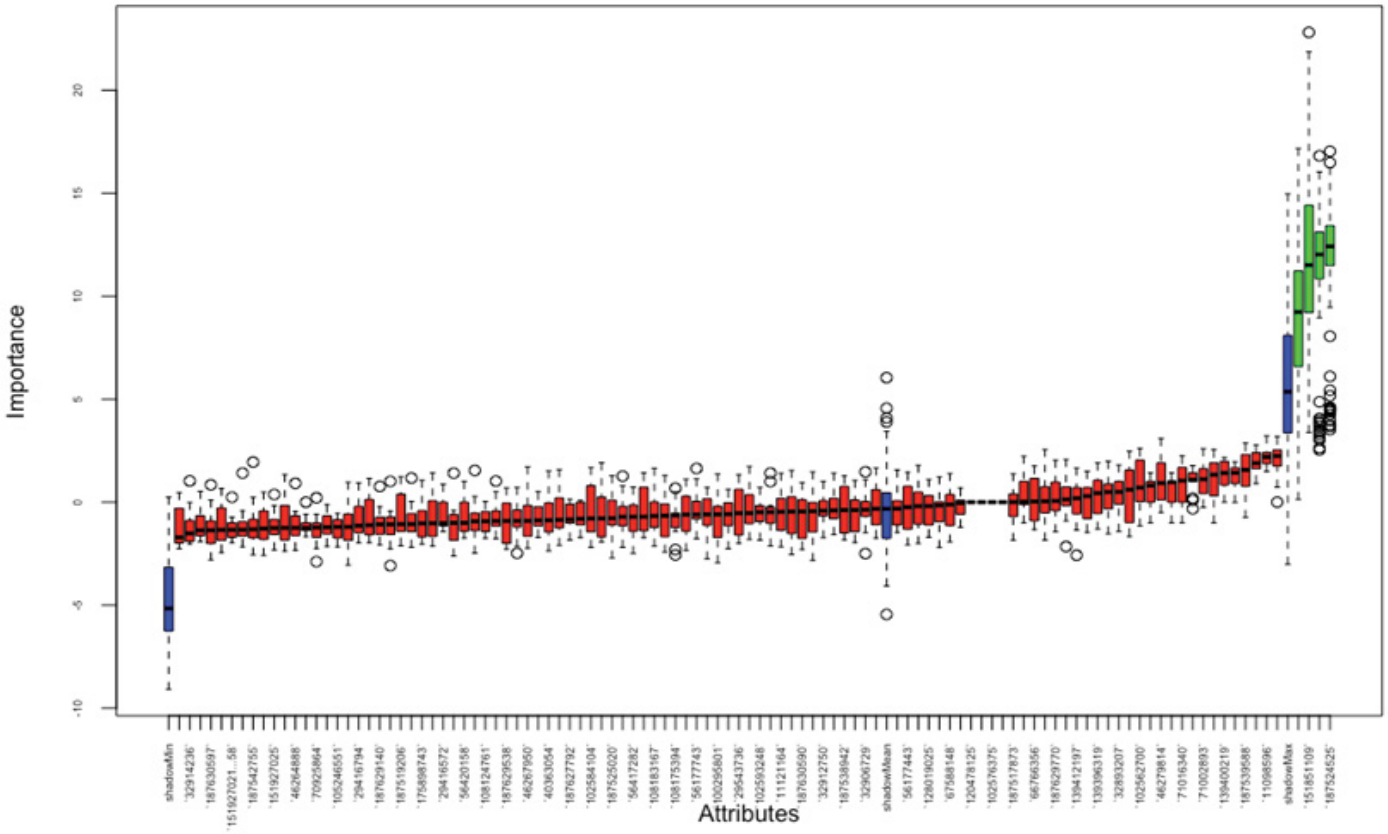


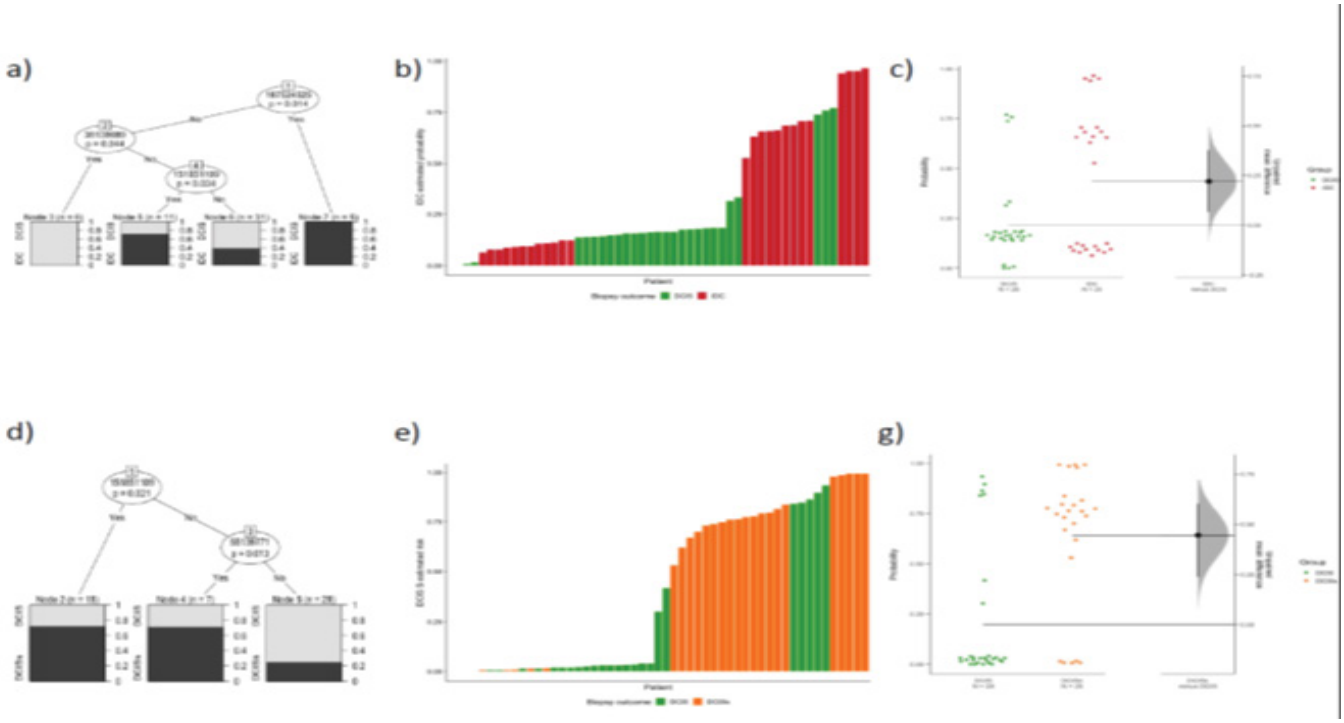
Figure 3: The discriminatory mutations between DCIS S and pure DCIS are the three reported in Fig. 2 ('151851109', '187524525' and '38138689') and an additional mutation '55139771' (PDGFRA gene, 'benign'variant)

Unadjusted estimated OR for the selected mutations (DCIS S vs. DCIS)

	DCIS	DCIS S	OR	p.ratio
	N=28	***N=25***		
151851109	5 (17.9%)	13 (52.0%)	4.75 [1.41;18.4]	0.011
187524525	0 (0.00%)	5 (20.0%)	. [.;.]	.
38138689	6 (21.4%)	0 (0.00%)	. [.;.]	.
55139771	2 (7.14%)	7 (28.0%)	4.69 [0.97;37.9]	0.056

Vilardell F, et al., J Cancer Sci Clin Ther 2024
DOI:10.26502/jcsct.5079238

Figure 4: a) IDC vs. DCIS, conditional inference tree.



Input mutations are: '151851109', '187524525' and '38138689'. All important mutations are used to end with four IDC probabilities: 0 (0/6) in node 3, 0.727 (8/11) in node 5, 0.387 (12/31) in node 6 and 1 (5/5) in node 7. With probabilities higher than 0.5 for IDC, we obtained a sensitivity of 52% (13/25) and a specificity of 89.3% (25/28). AUC = 0.769, 95%CI [0.66, 0.865] for the predicted probabilities. b) Random forests IDC vs. DCIS. 401 trees are estimated from resampled datasets and employing the confirmed mutations: '151851109', '187524525' and '38138689'. This figure shows the sorted predicted probabilities of IDC for each patient in the study, with red or green identifying the biopsy outcome, IDC or DCIS, respectively. c) The Gardner-Altman two group estimation plot shows the predicted probabilities of IDC and its unpaired mean difference (IDC - DCIS). Mean difference and its 95% confidence interval are displayed as a point estimate and vertical bar respectively, using the sample density distribution calculated from a bias-corrected and accelerated bootstrap analysis from 5000 resamples. d) DCIS S vs.DCIS, conditional inference tree. Input mutations are: '151851109', '187524525', '38138689' and '55139771'. Only '151851109' and '55139771' are used, with estimated probabilities of DCIS S of: 0.722 (13/18) in node 2, 0.714 (5/7) in node 4 and 0.25 (7/28) in node 5. With probabilities higher than 0.5 for DCIS S, we obtained a sensitivity of 72% (18/25) and a specificity of 75% (21/28). AUC = 0.736, 95%CI [0.601, 0.846] for the predicted probabilities. e) Random forest DCIS vs. S DCIS. 401 trees are estimated from resampled datasets and employing the confirmed mutations: '151851109', '187524525', '38138689' and '55139771'. This figure shows the sorted predicted probabilities of S DCIS for each patient in the study, with orange or green identifying the biopsy outcome, S DCIS or DCIS, respectively. f) The Gardner-Altman two group estimation plots show the predicted probabilities of DCIS S and its unpaired mean difference (DCIS S - DCIS). Mean difference and its 95% confidence interval are displayed as a point estimate and vertical bar respectively, using the sample density distribution calculated from a bias-corrected and accelerated bootstrap analysis from 5000 resamples.

Discussion

DCIS precedes IDC, and genetically DCIS and IDC share somatic mutations and CNAs. There are mainly two groups of theories that explain the progression from DCIS to IDC: those that consider invasiveness as a property acquired by the neoplastic cells through genetic alterations, and those that consider that it is changes of genetic expression in the peritumoral stroma that make tumor progression possible [21]. Currently, there is no molecular or histological tests that allows to predict the progression of DCIS to IDC. We designed, considering the first group of theories, a pilot study that would allow us to predict which DCIS will progress to IDC and which will not, based on the analysis of samples embedded in paraffin, in order to avoid overtreatment of DCIS with low risk of progression to invasive cancer. With this objective, we retrospectively selected samples of pure DCIS corresponding to 37 patients diagnosed between 2011 and 2016 and who, after five years of follow-up since their diagnosis, have not recurred or progressed to IDC. We compared the somatic gene mutation profile of this group of patients with that of 30 patients with extensive DCIS associated with IDC. The main limitation of this study, designed in a surgical pathology department, is the use of samples embedded in paraffin. This does not allow analysis of differential gene expression or CNA profiles, which are largely the genetic alterations responsible for the progression from DCIS to IDC. The use of paraffin-embedded samples only allows us to perform a comparative analysis of the profile of somatic genetic mutations in pure DCIS, IDC and DCIS associated with IDC. However, the study design with paraffinized samples has the advantage of facilitating, in the cases with DCIS associated with IDC, the selection of the areas corresponding

Vilardell F, et al., J Cancer Sci Clin Ther 2024
DOI:10.26502/jcsct.5079238

to the two tumor components, as well as their dissection. Most of the genetic alterations of breast cancer have already been acquired at the DCIS stage and there appear to be few somatic genetic mutations that can lead the transition from DCIS to IDC. Some of these differential mutations between DCIS and IDC were found by Kim SH et al. [19], and were those genes differentially mutated in synchronous DCIS-IDC but not in pure DCIS that we chose for our study. In order to find a genetic mutational profile predictor of indolent DCIS, it was ordered to make a customized panel with 40 chosen genes, comprising the differentially mutated genes reported in the work by Young Kim S et al. [19], plus some stromal genes differentially expressed at the preinvasive-to-invasive transition point. In our study, mutations were found in MYCL, NOTCH2, ALK, TMEM45A, PDGFRA, ADAM29, FAT1, MAP3K1, PIK3R1, FANCE, DST, MET, EPHA1, KMT2C, NOTCH1, ADAM12, MMP27, MMP8, ATM, PTPRB, BRCA2, POSTN, AKT1, SMARCA4, NCOA3, MKL1, BCOR and AR. Eight of these 28 genes belong to the group of 13 genes from the Cancer Gene Census that were found to be differentially mutated in synchronous DCIS but not in pure DCIS [19]. More specifically, these genes are BRCA, ATM, KMT2C, NOTCH1, PDGFRA, SMARCA4, FANCE and BCOR. Since we are looking for mutational profiles in DCIS that can predict low risk of progression to invasive carcinoma, we decided to take into account not only the mutations reported as pathogenic in the Cancer Gene Census, but also those previously considered benign. We found that FAT1 gene also might be differentially mutated in IDC and S DCIS but not in pure DCIS. FAT1 encodes a 500-600 kDa cadherin protein and can work as tumor suppressor or oncogene in different contexts. Its increased expression in all forms of neoplastic transformation indicates an important role in carcinogenesis, as reported by Kaewpila N et al. [23]. Surprisingly, POSTN, a gene that encodes a protein implicated in the epithelial-mesenchymal transition, probably only harbor mutations in patients with pure DCIS but not in patients with IDC or S DCIS. Mutations in POSTN gene could lead to a loss of function of the periostin protein and, therefore, a low mesenchymal and invasive state.

In the current work, we found that the variants 151851109 in KMT2C gene, 187524525 in FAT1 gene and 38138689 in POSTN gene, could discriminate between IDC and DCIS. After statistical analysis by means of random forests, we obtained a sensitivity of 52% (13/25) and a specificity of 89.3% (25/28) to predict IDC with a probability higher than 0.5. However, the three mutations mentioned above, along with the mutation 55139771 in PDGFRA gene, show a high ability to discriminate between S DCIS and pure DCIS, which is actually an indirect way of predicting the presence of infiltrating carcinoma next to the in situ component, and may be a way to predict the progression from DCIS to IDC. More specifically, according to our work, individuals with a predicted probability of having a DCIS synchronous with IDC higher than 0.5, are those with at least a mutation in 151851109 (KMT2C gene), 187524525 (FAT1 gene) or 55139771 (PDGFRA gene) and no mutation in 38138689 (POSTN gene). Given our aim of predicting the risk of evolution to invasive carcinoma in patients with DCIS, individuals with a predicted probability of DCIS S of 0.5 or lower would be those with a mutation in 38138689 or no mutation in either 151851109, 187524525 or 55139771. Finding this mutational profile can help identify DCIS with a low risk of progression to IDC, but to give robustness to these results it may be necessary to carry out new analyzes with larger sample sizes, as well as prospective and much more complex studies from fresh tissue that allow the performing of differential gene expression analyses between DCIS and IDC. We believe that aggressive cancer treatment could be avoided in patients diagnosed with breast carcinoma in situ in whom this profile of genetic mutation is found, although clinical trials and follow-up of these patients will be needed to verify that they do not develop infiltrating carcinoma despite application of a conservative treatment.

Data Availability Statement

The datasets generated and/or analysed during the current study are available in the Zenodo repository, at link:

https://zenodo.org/record/6828480#.YtGpS2gzYuU
DOI: 10.5281/zenodo.6828479

Funding

The source of funding was Ms. Francisca Rafales, an individual patron unrelated to the pharmaceutical industry or biomedical sciences.

Authors' contributions

FV and EI conceived the study. CC, FV and IU wrote the article. IU performed the experiments. MM and IU performed the computational analyses. FV collected the specimens and performed the clinical review. All authors have read and approved the manuscript for publication.

Declarations of interest: None

Acknowledgement

Ms. Francisca Rafales, whom we thank for her financial support

Ethical approval

The Healthcare Ethics Committee of Hospital Universitari Arnau de Vilanova de Lleida, of the Territorial Management of Lleida-GSS, at the meeting of December 3, 2019, minutes 14/2019, reported favorably on research project B00044, entitled "Prediction, by Comparative Mutational Analysis, of the Risk of Progression from Intraductal Carcinoma to

Vilardell F, et al., J Cancer Sci Clin Ther 2024
DOI:10.26502/jcsct.5079238

Infiltrating Ductal Carcinoma of the Breast. Pilot Project (B00044)", code CEIC-2166, with doctors E. Iglesias and F. Vilardell as co-investigators, and considered: The procedure for obtaining informed consent from the subjects participating in the study is adequate. All the experiments were performed in accordance with Declaration of Helsinki regulations. Informed consent was obtained from all subjects and/or their legal guardian(s).

References

1. Broders A C. Carcinoma in situ contrasted with benign penetrating epithelium. JAMA 99 (1932): 1670-1674.
2. Tavassoli FA, Norris HJ. A comparison of the results of long-term follow-up foratypical intraductal hyperplasia and intraductal hyperplasia of the breast. Cancer 65 (1990): 518-529.
3. Tavassoli FA. Lobular and ductal intraephitelial neoplasia, Pathologe 2 (2008): 107-111.
4. Welch HG, Black WC: Using autopsy series to estimate the disease "reservoir" for ductal carcinoma in situ of the breast: how much more breast cancer can we find? Ann Inter Med 127 (1997): 1023-1028.
5. Bijker N, J L Peterse, L Duchateau, et al. Risk factors for recurrence and metastasis after breast conserving therapy for ductal carcinoma in situ: analysis of European Organization for Research and Treatment of Cancer Trial 10853. J Clin Oncol 19 (2001): 2263-2271.
6. Vincent-Salomon A, Carlo L, Nadège G, et al. Integrated genomic and transcriptomic analysis of ductal carcinoma in situ of the breast. Clin Cancer Res 14 (2008): 1956-1965.
7. Van Seijen M, Esther H L, Alastair M T, et al. Ductal carcinoma in situ: to treat or not to treat, that is the question. Br J Cancer 121 (2019): 285-292.
8. Silverstein MJ. The University of Southern California/ Van Nuys prognostic index for ductal carcinoma in situ of the breast. Am J Surg 186 (2003): 337-343.
9. Rudloff U, Jack L M, Goldberg J I, et al. Nomogram for predicting the risk of local recurrence after breast-conserving sugery for ductal carcinoma in situ. J Clin Oncol 28 (2010): 3762-3769.
10. Solin LJ, Gray R, Baehner F L, et al. A multigene expression assay to predict local recurrence risk for ductal carcinoma in situ of the breast. J Natl Cancer Inst 105 (2013): 701-710.
11. Rakovitch E, Mozes S N, Hanna W, et al. Multigene Expression Assay and Benefit of Radiotherapy After Breast Conservation in Ductal Carcinoma in Situ. J Natl Cancer Inst 109 (2017): djw256.
12. Solin LJ. Management of Ductal Carcinoma In Situ (DCIS) of the Breast: Present Approaches and Future Directions. Curr Oncol Rep 21 (2019): 33.
13. Solin LJ, J Kurtz, A Fourquet, et al. Fifteen-year results of breast conserving surgery and definitive breast irradiation for the treatment of ductal carcinoma in situ of the breast. J Clin Oncol 14 (1996): 754-763.
14. Tsuda H. Gene and chromosomal alterations in sporadic breast cancer: correlation with histopathological features and implications for genesis and progression. Breast Cancer 16 (2009): 186-201.
15. Ma XJ, Salunga R, Tuggle J T, et al. Gene expression profiles of human breast cancer progression. Proc Natl Acad Sci U S A 100 (2003): 5974-5979.
16. Cancer Genome Atlas Network: Comprehensive molecular portraits of human breast tumours. Nature 490 (2021): 61-70.
17. Aubele M, A Mattis, H Zitzelsberger, et al. Extensive ductal carcinoma in situ with small foci of invasive ductal carcinoma: evidence of genetic resemblance by CGH. Int J Cancer 85 (2000): 82-86.
18. Petridis C, Brook M N, Shah V, et al. Genetic predisposition to ductal carcinoma in situ of the breast. Breast Cancer Res 18 (2016): 22.
19. Kim SY, Jung S H, Kim M S, et al. Genomic differences between pure ductal carcinoma in situ and synchronous ductal carcinoma in situ with invasive breast cancer. Oncotarget 6 (2015): 7597-7607.
20. Futreal PA, Coin L, Marshall M, et al. A census of human cancer genes. Nat Rev Cancer 4 (2004): 177-183.
21. Banerji S, Cibulskis K, Claudia R E, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. Nature 486 (2012): 405-409.
22. Cowell CF, Weigelt B, Sakr R A, et al. Progression from ductal carcinoma in situ to invasive breast cancer: Revisited. Mol Oncol 7 (2013): 859-869.
23. Knudsen ES, Ertel A, Davicioni E, et al. Progression of ductal carcinoma in situ to invasive breast cancer is associated with gene expression programs of EMT and myoepithelia. Breast Cancer Res Treat (2021): 1009-1024.
24. Kwaepila N, Burns G, Leong AS. Immunohistological localization of human FAT1 (hFAT1) protein in 326 breast cancers. Does this adhesion molecule have a role in pathogenesis? Pathology 38 (2006): 125-131.