

**The prognostic significance of immune microenvironment in breast ductal
carcinoma *in situ***

Michael S Toss ^{1,2}, Asima Abidi ¹, Dorothea Lesche ^{3,4}, Chitra Joseph ¹, Sakshi Mahale ³,
Hugo Saunders ³, Tanjina Kader ^{3,4}, Islam M Miligy ¹, Andrew R Green ¹, Kylie L Gorringe ^{3,4},
and Emad A Rakha ¹

¹Nottingham Breast Cancer Research Centre, Division of Cancer and Stem Cells, School of Medicine, The University of Nottingham, Nottingham City Hospital, UK. ²Histopathology Department, South Egypt Cancer Institute, Assiut University, Egypt, ³Cancer Genomics Program, Peter MacCallum Cancer Centre, Melbourne, Australia, and ⁴The Sir Peter MacCallum Department of Oncology, University of Melbourne, Parkville, Australia.

Correspondence:

Dr Michael Toss

Department of Histopathology, Nottingham University Hospital NHS Trust,

City Hospital Campus, Hucknall Road, Nottingham, NG5 1PB, UK

Tel: (44) 0115-9691169, Fax: (44) 0115-9627768

Email: msxms8@nottingham.ac.uk

Key words: immune cell markers, PD1, PDL1, TILs, DCIS, prognostic

Running title: Expression of immune cell and checkpoint proteins in DCIS

ABSTRACT

Background: The role of different subtypes of tumour infiltrating lymphocytes (TILs) in breast ductal carcinoma in situ (DCIS) is still poorly defined. This study aimed to assess the prognostic significance of B and T lymphocytes and immune checkpoint proteins expression in DCIS.

Methods: A well characterised DCIS cohort (n=700) with long-term follow-up comprising pure DCIS (n=508) and DCIS mixed with invasive carcinoma (IBC; n=192) were stained immunohistochemically for CD20, CD3, CD4, CD8, FOXP3, PD1 and PDL1. Copy number variation and TP53 mutation status were assessed in a subset of cases (n=58).

Results: CD3+ lymphocytes were the predominant cell subtype in the pure DCIS cohort, while FOXP3 showed the lowest levels. PDL1 expression was mainly seen in the stromal TILs. Higher abundance of TILs subtypes was associated with higher tumour grade, hormone receptor negativity and HER2 positivity. Mutant TP53 variants were associated with higher levels of stromal CD3+, CD4+ and FOXP3+ cells. DCIS coexisting with invasive carcinoma harboured denser stromal infiltrates of all immune cells and checkpoint proteins apart from CD4+ cells. Stromal PD1 was the most differentially expressed protein between DCIS and invasive carcinoma (Z=5.8, p<0.0001). Dense TILs, stromal FOXP3 and PDL1 were poor prognostic factors for DCIS recurrence, while dense TILs was independently associated with poor outcome for all recurrences (HR=7.0; p=0.024), and invasive recurrence (HR=2.1; p=0.029).

Conclusion: Immunosuppressive proteins are potential markers for high risk DCIS and disease progression. Different stromal and intratumoural lymphocyte composition between pure DCIS, DCIS associated with IBC and invasive carcinoma play a potential role in their prognostic significance and related to the underlying genomic instability. Assessment of overall TILs provides a promising tool for evaluation of the DCIS immune microenvironment.

Background

The incidence of breast ductal carcinoma *in situ* (DCIS) increased dramatically after the introduction of mammographic screening and accounts up to 25% of the screen detected breast cancers in the UK with similar rates in the US ^{1,2}. However, management of DCIS remains as a clinical challenge with a considerable proportion of patients are either over- or under-treated due to the lack of a robust prognostic assessment model to stratify patients' risk. Current management decisions rely on conventional clinicopathological parameters as the clinical utility of available gene expression assays and molecular prognostic biomarkers remain to be demonstrated. It is hence important to identify other robust markers for DCIS for behaviour predilection to decide a personalised treatment approach.

Although the role of tumour microenvironment in disease behaviour is undeniable, incorporation of the microenvironmental factors with the clinicopathological and/or molecular signature for DCIS risk assessment is still limited. The tumour immune microenvironment undergoes several changes during carcinogenesis ³. The triggering of an immune response in the form of infiltration of lymphocytes and other inflammatory mediators is a phenomenon known as "cancer immuno-editing" ⁴. It is known that subtype, density and location of tumour infiltrating lymphocytes (TILs) play a role in cancer prognosis. Infiltration of T lymphocytes extends an anti-tumour response through the CD8+ cytotoxic T lymphocytes while an immunosuppressive response occurs through CD4+ FOXP3+ regulatory T cells (Tregs) ^{5,6}. Similarly, CD20+ B lymphocytes are associated with a humoral anti-tumour response ⁷. The role of TILs in invasive breast cancer (IBC) is well validated, with a high TILs density associated with better response to adjuvant or neoadjuvant therapy and favourable outcome especially in triple negative (TNBC) and HER2+ subtypes ⁸⁻¹³. However, there are conflicting data for the role of TILs in DCIS prognosis with some studies claiming an unfavourable role ¹⁴ while others did not find association with the outcome ^{15,16}.

An important component of immuno-editing is the development of an immunosuppressive microenvironment mediated through the activation of the immune checkpoint pathway. The programmed cell death 1 (PD1) protein, a member of the B7/CD28 family of co-stimulatory receptors present mainly on T cells, sends inhibitory signals to T cells to suppress the anti-tumour response ^{4,17}. PD1 functions mainly through its ligand PDL1, which is expressed on a variety of immune cells such as antigen presenting cells and activated T and B lymphocytes and tumour cells ^{18,19}. Revolutionary results have been achieved from clinical trials involving monoclonal antibodies that target PD1 and block the PD1/PDL1 pathway in patients with non-small cell lung cancer, melanoma and renal cancer ²⁰.

In a previous study we have demonstrated the prognostic value of TILs, as assessed on H&E stained slides, in DCIS ¹⁴. Here we hypothesised that various immune cell subpopulations and the immune checkpoint proteins play roles in DCIS behaviour and this can be related to the underlying molecular profile. In this study, we evaluated the expression and the prognostic significance of T and B lymphocytes and immune checkpoint proteins PD1 and PDL1 in a large well characterised DCIS cohort with long-term follow up. We also investigated the relationship between genomic copy number alteration and *TP53* variants (as measures of genomic instability) with immune cell infiltration in a subset of cases.

MATERIALS AND METHODS

Study Cohort

This study was conducted on a well characterised annotated large DCIS cohort, comprising primary pure DCIS cases and DCIS associated with invasive tumour (DCIS-mixed) diagnosed at Nottingham City Hospital, Nottingham, UK, between 1990 to 2012 as described ²¹. The clinicopathological characteristics of the study cohort are summarised in Supplementary Table 1. The patients' demographic data, histopathological parameters, overall TILs density, DCIS management including post-operative radiotherapy (RT) and development of local recurrence

along with patient outcome was available. Tumour grade was classified according to the three-tier system; low, intermediate and high grades ²². Local recurrence free interval (LRFI) was defined as the time interval (in months) between 6 months after the primary DCIS surgery and occurrence of ipsilateral local recurrence either as DCIS or IBC. Recurrence was considered when it occurred in the same quadrant and showed a similar nuclear grade or higher than the primary tumour. In this study, cases with new lesions in other quadrants, or with lower nuclear grade and tumours developing in the contralateral breast were not counted as a recurrence and were censored at time of development of the new event. All recurrence events were checked 3 times by 3 researchers with pathology experience. Moreover, they were histologically reviewed to ensure that the morphology matches that stated in the original histology. We did our best to differentiate true recurrences from new primary tumors with acknowledgement of the limitations of our ability to accurately make such a distinction. We are interested in the behaviour of DCIS under investigation rather than the risk of other events that happened following the primary diagnosis of DCIS, either new primary ipsilateral or contralateral events. We have not included these events in our analysis to avoid the confounding effects of these tumours, which may have different biological characteristics features.

Data on Oestrogen Receptor (ER), Progesterone Receptor (PR), Human Epidermal Growth Factor Receptor 2 (HER2), Ki-67, TP53, CK5/6, and HIF-1 α were available to this cohort as previously described. Briefly, ER and PR positivity were both defined as $\geq 1\%$ nuclear tumour staining ²³. Immunoreactivity of was scored using HercepTest guidelines. Chromogenic in situ Hybridisation (*CISH*) was used to assess *HER2* gene amplification in cases with 2+ score using the HER2 CISH pharmDx™ kit (Dako) ²⁴. Proliferation index was evaluated through Ki-67 antibody immunohistochemical (IHC) staining and defined as high when $\geq 14\%$ of cells showed nuclear positivity ²⁵. The molecular subtypes were classified as: Luminal A (ER+ and/or PR+, HER2- and Ki67 low), Luminal B (ER+ and/or PR+, HER2- and Ki67 high or ER+ and HER2+),

HER2 enriched (HER2+ and ER-) and Triple Negative (ER-, PR- and HER2-). TP53 and CK5/6 proteins were assessed immunohistochemically, and the percentage of positive nuclear and membranous expression were scored, respectively. More than 1% HIF-1 α nuclear immunoreactivity was considered positive ^{26,27}. High TIL density was considered where the average number lymphocytes/duct was 20 or more ¹⁴.

Immunohistochemistry of immune cell markers and checkpoint proteins

Tissue microarrays (TMAs) were prepared from both cohorts using a TMA Grand Master 2.4-UG-EN machine (3D Histech) with 1mm punch sets. The TMA was prepared from all representative areas of heterogeneous morphology and grade, when present. Apart from the TMAs, 30 full face sections were also assessed to reveal the infiltration pattern of various TILs subtypes and immune checkpoint proteins.

Immunohistochemical staining (IHC) was performed on TMA sections using a Novocastra Novolink TM Polymer Detection Systems Kit (Code: RE7280-K, Leica, Biosystems, UK). Primary antibody specific for CD20 (mouse monoclonal, DAKO Clone L26/M0755), CD3 (rabbit monoclonal antibody [SP7], Abcam (ab16669)), CD8 (mouse monoclonal, DAKO Clone C8/144B), CD4 (rabbit monoclonal antibody [EPR6855], Abcam (ab133616)), FOXP3 (mouse monoclonal [236A/E7], Abcam (ab20034)), PD1 (Mouse monoclonal [EH33] Cell signalling (43248S)), PDL1 (Rabbit monoclonal antibody [E1L3N] Cell signalling (13684S)) were used to characterise TILs and immune checkpoint proteins. Supplementary Table 2 summarises the details of the antibodies used for IHC staining. Sections from normal tonsil were used as a positive control while a negative control was generated by omitting the primary antibodies.

Immunoscoreing

Stained slides were digitally scanned using a Nannozoomer (Hamamatsu Photonics), at 20X magnification high resolution images and viewed using the Aperio ImageScope Viewer software version 12.3.3 (developed by Leica Biosystems Imaging). All cases were scored

visually using the digitised images for stromal expression of TILs as previous described ¹⁴, and for intratumour TILs. Briefly, the average number of lymphocytes touching DCIS basement membrane or away from it by one lymphocyte cell thickness was counted. In addition, the average number of intratumour TILs per duct was counted for CD20, CD3, CD4, CD8, FOXP3 and PD1. PDL1 expression in tumour epithelial cells was scored as percent of cells showing membranous and/or cytoplasmic staining, where cases with >1% positive cells were considered positive ¹⁵. The pure DCIS and the in-situ component of mixed cases which clearly defined a demarcated tumour edge were assessed, separately. Average score was used as a final score for cases with multiple cores. All cases were scored independently by two observers (A Abidi and M Toss) blinded to patient histopathological data and clinical outcome. The average score was considered in discrepant cases between the two observers.

DNA Extraction, targeted sequencing library preparation, enrichment, and sequencing

Sections from formalin-fixed paraffin-embedded (FFPE) blocks from a subset of DCIS cases (n=58) were reviewed by two pathologists to select representative sections. The cases were randomly selected based on tissue block availability and adequacy of tumour tissue for DNA extraction. Tumour epithelial cells were needle-microdissected from 10 µm sections (8 – 20 slides) after haematoxylin staining and DNA was isolated using the AllPrep DNA/RNA FFPE kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Targeted sequencing of tumour DNA was performed using an Agilent SureSelect XT Custom Panel (Agilent, Santa Clara, CA, USA) which targeted breast cancer specific genes including *TP53*. Library preparation was performed from at least 80 ng of tumour DNA using the KAPA Hyper system (Agilent). For samples with lower DNA yields available (≤ 10 ng), library preparation was performed with the NEBNext® Ultra™ II DNA Library Prep Kit (New England BioLabs® Inc, Ipswich, MA, USA) and used for low-coverage whole-genome sequencing (LC-WGS) as described previously ²⁸. Illumina Nextseq500 (paired-end 75 bp reads) was used to run

pooled, normalised, indexed libraries according to Illumina protocols. Average sequencing depth was 400X (SureSelect) and 1.3X (LC-WGS).

Copy number and variant data analysis

Data was processed by an in-house bioinformatics pipeline at the Peter MacCallum Cancer Centre (Melbourne) to detect and filter for high confidence somatic variants. Paired-end sequence reads were aligned to the g1 k v37 hg19 reference genome using BWA. Optical duplicate reads were removed using Picard and the Genome Analysis Tool Kit (GATK) was used to perform local realignment around indels and base quality score recalibration in accordance with the recommended best practice workflow ²⁹. Single nucleotide polymorphism (SNP) and indel variants were called using GATK Unified Genotyper, Platypus ³⁰, and VarScan 2 ³¹. Called variants were annotated using the Ensembl Variant Effect Predictor ³². Somatic mutations were identified by removing previously available germline variant data for hereditary breast and ovarian cancer panel genes and by applying the following filters; canonical transcript; bidirectional read; quality ≥ 100 ; variants identified by at least two variant callers; minor allele frequency (MAF) present at ≤ 0.0001 in ExAc (Version 0.3.1, excluding TCGA data) ³³, GnomAD (Version 2.0), EVS (Version ESP6500SI-V2-SSA137) [Exome Variant Server, available <http://evs.gs.washington.edu/EVS/>], and 1000 Genomes ³⁴ databases. Genomics regions identified by Scheinin et al ³⁵ were removed from analysis as described previously ²⁸.

Copy number profiles were obtained using the SureSelect panel off-target reads with CopywriteR ³⁶ and Control-FREEC for LC-WGS reads ³⁷. A lymphocyte DNA control run (NA12878, Coriell Institute) in the same sequencing batch was used for baseline normalisation. Data were then imported to NEXUS Copy Number™ (software v9.0; BioDiscovery Inc, El Segundo, CA, USA), segmented using a FASST2 segmentation algorithm and visualised using a circular binary segmentation algorithm to determine gains and losses

as described ³⁸. Fraction of the genome altered (FGA) was calculated as the percentage of base pairs with copy number gain or loss. Number of telomeric allelic imbalances (NTAI) was calculated as the count of telomeres (excluding chromosome Y) with a gain or loss ¹⁵. Twenty cases showed deleterious/missense *TP53* variants annotated as described above.

Statistical Analysis

All statistical tests were conducted on IBM SPSS software (version 24.0, Chicago, IL, USA) for Windows. For the dichotomisation of the immune cells and checkpoint protein densities, median values were used as cut off points.

Spearman's Rho test was used to correlate between the expression of different immune cells with each other, overall TILs and with the proliferation marker (Ki67), *TP53* and HIF-1 α . Chi square test was used to determine the association between different markers densities with other clinicopathological parameters. Expression of PDL1 within the tumour cells was not included in the analysis as the percentage of positive cases was very low (2% only). Mann-Whitney U test and Kruskal-Wallis test were used to compare between TILs densities and FGA, NTAI and *TP53* variants. Differential expression of various study markers between pure DCIS and DCIS coexisting with IBC was analysed using the continuous score data and the Mann-Whitney U test. The association with LRFI was assessed in univariate analysis using Kaplan Meier curves and Log Rank Test. Cox regression was carried out for multivariate analysis to adjust for confounding by conventional prognostic factors. Survival analyses were confined to patients treated with breast conserving surgery (BCS) with or without radiotherapy as it was the highest group showing ipsilateral recurrence (>95% of recurrences) and the smaller number of events in patients treated with mastectomy. For each test conducted, a two-tailed *p*-value of <0.05 was considered statistically significant.

RESULTS

Patterns of expression and distribution of various inflammatory cells and immune checkpoints

The assessment of the full-face sections revealed that TILs infiltration was comparable to the TMA sections thus justifying the representability of the TMA sections for immunoscore both in the stromal and intratumoural compartments. Uninformative cores (lost, folded tissue or cores containing <15% tumour cells and/or stroma) were excluded from scoring.

Supplementary Table 3 summarises the distribution of various types of lymphocytes as well as PD1 and PDL1 expression either in the stroma or intratumourally. The predominant lymphocytes in DCIS were CD3+ cells either intratumourally (median 1 cell/duct, range 0-50) or in the stroma (median 3 cells/duct, range 0-100) whereas CD4+ cells were the most frequent subtype of T cells (range 0-60 in the stroma and 0-25 in the intratumoural compartment). Stromal FOXP3+ lymphocytes were the least frequent (median 0 cells/duct, range 0-25) while PD1+ cells were the lowest type intratumourally (median 0 cells/duct, range 0-15). Only 2% of pure DCIS cases showed positive expression for PDL1 within tumour cells. Figure 1 and Supplementary Figure 1 show examples of dense TIL infiltrates in H&E and IHC stained sections with various markers demonstrating the distribution of immune cells and immune checkpoint protein in pure DCIS.

Stromal CD3+, CD8+, CD4+, PDL1+ and PD1+ lymphocytes showed the highest correlation with overall TILs (all $r > 0.4$, $p < 0.0001$). The strongest correlation between various lymphocytes was observed with CD3+ and CD4+ lymphocytes ($r = 0.96$, $p < 0.0001$), followed by CD4+ and CD8+ cells ($r = 0.94$, $p < 0.0001$). Figure 2A shows a correlation matrix that summarises the associations between overall TILs, various immune cell subtypes, PD1, PDL1, hormonal receptors, Ki67, HIF-1 α , CK5/6 and p53 protein.

Moreover, DCIS mixed with synchronous IBC harboured more dense stromal TILs, PD1, PDL1 and various subpopulation of lymphocytes within the stroma than pure DCIS, with the highest difference observed with stromal PD1+ ($Z=5.8$, $p<0.0001$) followed by stromal CD8+ ($Z=5.5$, $p<0.0001$). Only stromal CD4+ cells did not show a significant statistical difference between both groups. In the context of intratumoural immune cells, there was no difference in immune cells infiltrates between pure DCIS and DCIS mixed apart from intratumoural CD20 which was more prevalent in pure DCIS ($Z=2.7$, $p=0.007$) and PD1 which was more dense in DCIS associated with IBC ($Z=2.2$, $p=0.025$). Figure 2B and Supplementary Figure 2 demonstrate the differences in various immune cell markers between both groups.

Association of immune cells subpopulations and immune checkpoint proteins with other clinicopathological parameters in pure DCIS cohort

Higher expression of the majority of markers, either stromal or within the intratumoural compartment, was associated with hormonal receptor negativity. The exceptions were intratumoural CD20+ cells, which were associated with ER and PR positivity, and intratumoural CD3+, CD8+ and CD4+, which showed no association with ER or PR. High expression of all markers within the stroma, and intratumoural expression of PD1 and FOXP3, was associated with higher nuclear grade, HER2 positivity, and higher expression of HIF-1 α . DCIS with smaller size and lower expression of HIF-1 α showed higher expression of intratumoural CD20+ lymphocytes. It was also observed that dense infiltration of PD1, CD3, CD8, CD20 and FOXP3 in the stromal region was significantly associated with the presence of comedo-type necrosis. Supplementary Table 4 (a-d) shows the correlation between various immune cells/checkpoint markers and the clinicopathological parameters in the pure DCIS cohort.

Copy number variation and *TP53* mutations

Higher FGA was associated with higher tumour grade ($p=0.002$), presence of comedo type necrosis ($p=0.0003$), ER negativity ($p=0.009$), HER2 positivity ($p=0.0003$) and presence of at least one deleterious/missense *TP53* variant ($p<0.0001$); (Supplementary Figure 3). Additionally, a higher NTAI was associated with high grade DCIS ($p=0.008$), presence of comedo type necrosis ($p=0.006$), HER2 positivity ($p=0.03$), and with presence of deleterious/missense *TP53* variants ($p=0.002$). The median FGA in pure DCIS with dense TIL was 26% compared to 12% in cases with sparse TILs, however the difference was not statistically significant ($p=0.08$, Figure 3). Similar results were observed with stromal CD3 whereas median FGA in pure DCIS with dense CD3+ cells was 26% compared to 10% in cases with sparse CD3+ cells ($p=0.07$, Figure 3). No association between NTAI and TILs density nor immune cell/checkpoint proteins was observed. However, tumours carrying somatic deleterious/missense *TP53* variants showed higher counts of stromal CD3+ ($p=0.031$), CD4+ ($p=0.04$), and FOXP3+ ($p=0.02$); (Figure 3).

Outcome Analysis

High expression of stromal FOXP3+ and stromal PDL1+ were the sole markers associated with all recurrences in DCIS patients treated with BCS ($HR=2.1$, $95\%CI=1.1-3.7$, $p=0.025$ and $HR=4.4$, $95\%CI=2.4-8.1$, $p<0.0001$, respectively); Figure 4A and 4B. When analysis was restricted to invasive recurrences, only stromal PDL1 was associated with shorter LRFI ($HR=2.4$, $95\%CI=1.1-5.3$, $p=0.032$), Figure 4C. As previous reported¹⁴; overall TILs density was associated with shorter LRFI either for all recurrences ($HR=2.9$, $95\%CI=1.7-5.0$, $p=0.0001$ or invasive recurrences ($HR=2.1$, $95\%CI=1.1-4.3$, $p=0.04$). No significant associations with recurrence were observed for the intratumoural infiltration of any immune cell marker with recurrence (Table 1). Outcome analysis in context of various intrinsic molecular subtypes revealed that high expression of stromal FOXP3 was associated with

shorter LRFI in luminal A subtype ($HR=2.9$, $95\%CI=1.1-8.1$, $p=0.032$) while higher PDL1 expression in stromal TILs was associated with shorter LRFI in luminal A ($HR=7.1$, $95\%CI=2.9-12.2$, $p<0.0001$) and showing a trend of poor prognosis in luminal B ($HR=3.2$, $95\%CI=0.9-8.9$, $p=0.067$) and HER2 enriched ($HR=1.8$, $95\%CI=0.5-7.5$, $p=0.080$) subtypes. Other markers did not show association with outcome in different molecular classes.

In multivariate survival analysis, higher expression of stromal PDL1 was associated with shorter LRFI for all DCIS recurrences (but not for invasive recurrences only) independent of other parameters of DCIS aggressiveness including age at diagnosis, DCIS size, nuclear grade, presence of comedo-type necrosis and margin status ($HR= 2.9$, $95\% CI=1.4-6.1$, $p=0.005$). Interestingly, dense TILs was independently associated with DCIS recurrence either when included in a model comprising FOXP3+ and PDL1+ cells, or in a model where all immune cell markers associated with DCIS recurrence were incorporated with other DCIS clinicopathological parameters. A statistically significant result was obtained for all recurrences or when the analysis was confined to invasive recurrence only. Table 2 A and B summarise the various multivariate models analysed.

Moreover, high TILs density and high stromal PDL1+ expression could categorise high risk DCIS defined as high grade DCIS and larger than >15mm in size into two distinct groups with different outcome; whereas high TILs and stromal PDL1+ cells were associated with shorter LRFI in these patients ($HR=3.2$, $95\%CI=1.1-9.5$, $p=0.039$, and $HR=2.5$, $95\%CI=1.1-6.3$, $p=0.044$, respectively) (Supplementary Figure 4).

DISCUSSION

The importance of the tumour microenvironment dynamics during the neoplastic transformation of breast epithelial tissue is increasingly being recognised and is considered as a hallmark of cancer ^{3,39,40}. However, the effect of infiltrating lymphocytes in various malignancies including DCIS is controversial ^{14-16,41-43} and the role of PD1 and PDL1 proteins in DCIS remains to be defined. Through this study, we aimed to characterise TIL subtypes and their potential prognostic role in DCIS. The dominance of CD3+ cells, including CD4+ and CD8+ subsets, indicates the immunologically active state of these early tumours, which could suggest an active chemokine secretion recruitment mechanism ⁴⁴. Apart from cell-mediated immunity, an active humoral immune response is also indicated with the high expression of CD20+ cells. This result suggests that maturation of B lymphocytes may occur at an early stage of tumour development ⁴⁵. In comparison to pure DCIS, DCIS with synchronous IBC had notably increased levels of CD8+, CD20+, FOXP3+, PD1+ and PDL1+ cells. Since FOXP3 is the single most accurate marker of regulatory T cells (Tregs) ⁴⁶, the high expression of FOXP3+ cells in a proportion of DCIS in addition to high expression of PDL1 suggests an immunosuppressive environment being created simultaneously in early tumour development, in concordance with previous studies ^{47,48,49}. A recent review by Chen *et al* summarised findings from previous work that support our study findings, such as high grade DCIS harbouring more TILs than low grade lesions and being associated with underlying genetic alterations ⁵⁰.

We observed a significantly decreased recurrence-free survival in patients with a higher stromal infiltration of FOXP3. This suppressive environment appears to be particularly important for differentiating outcome in luminal A tumours, in which both FOXP3+ and PDL1+ were associated with poor outcome. FOXP3 was also described as an independent negative prognostic marker in IBC ⁵¹. Higher infiltration of FOXP3+ cells also predicts poor overall survival in ovarian carcinoma where release of the chemokine CCL22 in the tumour

microenvironment by tumour associated macrophages facilitates the recruitment of FOXP3+ Tregs and hence, the suppression of cytotoxic T cells ⁵². Another mechanism by which FOXP3 aids in tumour recurrence could be by binding to the upstream region of transcription start site of chemokines such as CCR7 and CXCR4.

The crucial role of the PD1/PDL1 pathway in the attenuation of T cell mediated immunity is well established ⁵³. We observed that high expression of stromal PDL1, but not PD1, was associated with a shorter recurrence free interval. High expression of PDL1 is associated with a poorer outcome in pancreatic cancer, renal cell carcinoma and breast cancer ⁵⁴⁻⁵⁶. Similarly, high expression of PD1 is associated with poorer outcome in renal cell carcinoma and breast cancer ^{57,58}. This association can be explained by the attenuation of the cytotoxic function of T lymphocytes in the presence of active PD1/PDL1 interactions thus aiding the tumour in its escape mechanisms ^{17,59}. Another mechanism that might be at play is the upregulation of the PI3K pathway due to loss of PTEN or mutation of *PIK3CA* which ultimately leads to upregulation of PDL1 ⁶⁰. However, we saw no association of *PIK3CA* mutations with PDL1 positivity.

Despite the associations of an immunosuppressive microenvironment in DCIS with recurrence, in our multivariate analysis, overall TIL density was the strongest independent predictor for invasive recurrence. Since we did not find any significant associations of dense intratumoural infiltration of any immune cell with recurrence, we consider that assessment of stromal infiltration of different subtypes of TILs is a better method of TILs evaluation for outcome analysis in DCIS ¹⁴. It is similarly recommended to assess stromal TILs in IBC, but in IBC TILs are strongly associated with better outcome, particularly in TNBC ³⁸. The contradictory association of TILs as a good prognostic marker in IBC and as a poor prognostic marker in DCIS could support a hypothesis that direct contact between invasive tumour cells and immune cells leads to destruction of tumour cells and hence better prognosis, while in DCIS, a dense inflammatory reaction could provide a way for destruction of the surrounding

myoepithelial layer ⁶¹ and basement membrane. This facilitation of tumour cell infiltration to the surrounding stroma ⁶² would be enhanced by the immunosuppressive proteins produced by the tumour cells to escape destruction by host immunity and guarantee maintenance of the invasive tumour cells ⁴⁹.

An alternative (and not necessarily exclusive) explanation could be that a heightened immune response in DCIS and its association with poor prognosis reflects an underlying aggressive breast tumour biology. Our findings regarding the significant association of the stromal immune markers with higher grade, HER2 positivity and hormone receptor negativity has been reported previously in DCIS ⁶³ and IBC ³⁸. The ability to recruit more inflammatory cells through production of neoantigens and chemokines, as well as their ability to modulate the immunosuppressive microenvironment to evade the host immune response, combine to develop the immune resistance shown by these tumours and lead to still further aggressive behaviour ⁶⁴. ER negative DCIS are more proliferative and genomically unstable, hence stimulate a stronger attack from the immune cells. The association of dense infiltration in HER2+ tumours and TNBC can be best interpreted as a pro-tumourigenic function of TILs especially due to higher infiltration of Tregs ^{10,63}. Our findings regarding the association of PD1 and PDL1 in TILs with higher tumour grade, were supported by the findings of Muenst *et al.* ⁵⁸. This association with higher grade emphasises the role of PD1/PDL1 pathway adopted by DCIS to evade anti-tumour immunity and spread aggressively. Higher expression of PDL1 in TNBC has also been reported in other studies ^{60,65}. The association of more proliferative tumours such as HER2 enriched, TNBC and basal like tumours, with high expression of PDL1 could be due to the increased immunogenicity of these tumours and increased presence of neoantigens.

All solid tumours elicit an inflammatory response that is critical for the recruitment of immune cells, tumour cell proliferation, survival, and angiogenesis. Tumours initiate these responses through several mechanisms including hypoxia and derived factors ⁶⁶. Hypoxia was reported

to shape the immunosuppressive microenvironment of the tumour ⁶⁷. The positive correlation between HIF1a and FOXP3 in our study supports such a hypothesis and shows the crosstalk between tumour microenvironment components in DCIS progression. This finding is supported other studies that showed tight interaction between hypoxia and the tumour inflammatory process ^{66,68}.

DCIS harbouring dense inflammatory cells was reported to have more copy number alteration ¹⁵ and higher Oncotype DX score ⁶⁹ than those with low TILs density, which reflect the underlying biological association between TILs density and DCIS behaviour. As expected, in our study high risk DCIS (high grade, comedo type necrosis, hormonal receptor negativity and HER2 positivity) showed a higher degree of genomic instability represented by an increased burden chromosomal gains and losses and the presence of deleterious/missense *TP53* mutations ⁷⁰. In general, genomic instability leading to high risk DCIS lesions might promote T cell activation (CD3 and CD4) and recruitment of regulatory T cells (FOXP3) as suggested by our data. These findings are supported by Hendry *et al*, who concluded that tumour immune-editing and evasion of the immune system could be a key step in progression from DCIS into IBC ¹⁵. Our data lacked sufficient numbers of cases analysed for mutation and CNVs to evaluate the interaction of these events with the immune context in terms of patient outcome.

Our findings could refine high risk DCIS into two distinct groups based on TILs density and stromal PDL1 expression, where high risk cases with dense TILs or higher stromal PDL1+ cells behaved more aggressively and had higher recurrence rate than high risk DCIS with lower TILs or stromal PDL1+ cells. This observation reflects the role of immune microenvironment in DCIS risk stratification and management guidance. Implementation of TILs evaluation in clinical practice could avoid further management, i.e. unnecessary postoperative radiotherapy in a subgroup of patients.

In our study, evaluation of TILs in the routine H&E sections was the only inflammatory related tool that showed independent association with DCIS recurrence, especially with invasive recurrence. Taken together, assessment of TILs in H&E sections could provide a robust prognostic data with no need for further specific immune cell markers assessment.

Conclusion

This study highlights the importance of the immune microenvironment playing a role in DCIS behaviour. Subsets of TILs, immune checkpoint proteins and their distribution should be carefully analysed for patient risk assessment. Since patient risk stratification has been difficult due to conflicting evidence regarding TILs affecting DCIS prognosis, this study provides a more uniform inference with the higher infiltration of FOXP3+ and PDL1 immune cells being associated with poor outcome. Further mechanistic studies to clarify the crosstalk between DCIS tumour cells and the immune microenvironment are warranted.

Limitation of the study

This study was conducted on TMA sections which may underestimate TILs infiltration although we did compare the distribution between TMA cores and full-face sections which showed comparability. In addition, due to the small number of recurrence events in patients treated with RT or when the cohort was split by molecular class or risk groups (high vs. low risk DCIS), further studies of TILs in context of these limitations are recommended. Also, the role of hormonal therapy is overlooked here as none of the study patients received such therapy. Data from genomic assessment was only available for a small subset of tumours limiting statistical power and the capability to conduct sub-group analysis (e.g. stratification by receptor status or grade), which might have weakened associations between TIL subtypes and genetic features.

Additional information

Acknowledgements

This research was supported and funded by the Egyptian Ministry of Higher Education and Scientific Research. We thank the Nottingham Health Science Biobank and Breast Cancer Now Tissue Bank for the provision of tissue samples

We also thank the Peter MacCallum Cancer Centre Bioinformatics and Molecular Genomics core facilities, which were supported by the Australian Cancer Research Foundation, as well as Dr Jia-Min Pang for the histopathological assessment and Dr Siobhan Hughes for technical assistance. This research was supported by the Peter MacCallum Cancer Foundation and by the National Breast Cancer Foundations (IIRS-18-051).

Authors' contributions

Conception and design: Michael Toss, Asima Abidi, Chitra Josef, Andrew R Green and Emad A Rakha

Cohort Identification, Collection and assembly of data: Michael Toss, Islam M Miligy, and Kylie L Gorringe

Staining and scoring analysis: Michael Toss, Asima Abidi

Copy number and molecular analysis: Dorothea Lesche, Sakshi Mahale, Hugo Saunders, Tanjina Kader and Kylie L Gorringe.

Data analysis and interpretation: Michael S Toss, Asima Abidi, Chitra Josef, Andrew R Green and Emad A Rakha

Manuscript writing: Michael S Toss, Asima Abidi, Dorothea Lesche, Chitra Joseph, Sakshi Mahale, Hugo Saunders, Tanjina Kader, Islam M Miligy, Andrew R Green, Kylie L Gorringe, and Emad A Rakha.

All authors contributed to revise and approve the final version of the manuscript

Ethical approval and consent to participate

This work obtained ethics approval by the North West – Greater Manchester Central Research Ethics Committee under the title; Nottingham Health Science Biobank (NHSB), reference number 15/NW/0685. All patients included in this study were consented to participate in the study and to use their materials in research. All samples from Nottingham used in this study were pseudo-anonymised and stored in compliance with the UK Human Tissue Act.

Availability of data and materials

The authors confirm the data that has been used in this work is available on reasonable request.

Competing interest

The authors declare no conflict of interest.

Funding

This research was supported and funded by the Egyptian Ministry of Higher Education and Scientific Research.

References

1. Ernster VL, Ballard-Barbash R, Barlow WE, Zheng Y, Weaver DL, Cutter G *et al.* Detection of ductal carcinoma in situ in women undergoing screening mammography. *J Natl Cancer Inst.* 2002; **94**: 1546-1554; doi 10.1093/jnci/94.20.1546.
2. Groen EJ, Elshof LE, Visser LL, Th EJ, Winter-warnars HAO, Lips EH *et al.* Finding the balance between over- and under-treatment of ductal carcinoma in situ (DCIS). *The Breast* 2017; **31**: 274-283; doi 10.1016/j.breast.2016.09.001.
3. Artacho-Cordón A, Artacho-Cordón F, Ríos-Arrabal S, Calvente I, Núñez MI. Tumor microenvironment and breast cancer progression: A complex scenario. *Cancer Biol Ther.* 2012; **13**: 4-35; doi 10.4161/cbt.13.1.18869.
4. Schreiber RD, Old LJ, Smyth MJ. Cancer Immunoediting: Integrating Immunity's Roles in Cancer Suppression and Promotion. *Science* 2011; **331**: 1565-1570; doi 10.1126/science.1203486.
5. Emens LA. Breast cancer immunobiology driving immunotherapy: vaccines and immune checkpoint blockade. *Expert Rev Anticancer Ther.* 2012; **12**:1597-611;doi: 10.1586/era.12.147.
6. Cruz-merino LD, Barco-sánchez A, Carrasco FH, Fernández EN, Benítez AV, Molina JB *et al.* New Insights into the Role of the Immune Microenvironment in Breast Carcinoma. *Clin Dev Immunol* 2013;2013:785317; doi: 10.1155/2013/785317.
7. Nelson BH. CD20+ B cells: the other tumor-infiltrating lymphocytes. *J Immunol.* 2010 ;185:4977-82; doi: 10.4049/jimmunol.1001323.
8. Kotoula V, Chatzopoulos K, Lakis S, Alexopoulou Z. Tumors with high-density tumor infiltrating lymphocytes constitute a favorable entity in breast cancer : a pooled analysis of four prospective adjuvant trials. *Oncotarget.* 2016 26:5074-87; doi: 10.18632/oncotarget.6231.
9. García-martínez E, Gil GL, Benito AC, González-billalabeitia E, Angeles M, Conesa V *et al.* Tumor-infiltrating immune cell profiles and their change after neoadjuvant chemotherapy predict response and prognosis of breast cancer. *Breast Cancer Res* 2014: 1-17; doi 10.1186/s13058-014-0488-5.
10. Loi S, Michiels S, Salgado R, Sirtaine N, Jose V, Fumagalli D *et al.* Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: results from the FinHER trial. *Ann Oncol* 2014; **25**: 1544-1550; doi 10.1093/annonc/mdu112.
11. Denkert C, von Minckwitz G, Brase JC, Sinn BV, Gade S, Kronenwett R *et al.* Tumor-infiltrating lymphocytes and response to neoadjuvant chemotherapy with or without carboplatin in human epidermal growth factor receptor 2-positive and triple-negative primary breast cancers. *J Clinl Oncol* 2015; **33**: 983-991; doi 10.1200/JCO.2014.58.1967.
12. Adams S, Gray RJ, Demaria S, Goldstein L, Perez EA, Shulman LN *et al.* Prognostic Value of Tumor-Infiltrating Lymphocytes in Triple-Negative Breast Cancers From Two Phase III Randomized Adjuvant Breast Cancer Trials: ECOG 2197 and ECOG 1199. *J Clin Oncol* 2014; **32**(27): 2959-2966; doi 10.1200/jco.2013.55.0491.
13. Ibrahim EM, Al-Foheidi ME, Al-Mansour MM, Kazkaz GA. The prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancer: a meta-analysis. *Breast Cancer Res Treat* 2014; **148**: 467-476; doi 10.1007/s10549-014-3185-2.
14. Toss MS, Miligy I, Al-Kawaz A, Alsleem M, Khout H, Rida PC *et al.* Prognostic significance of tumor-infiltrating lymphocytes in ductal carcinoma in situ of the breast. *Mod Pathol* 2018; **31**: 1226-1236; doi 10.1038/s41379-018-0040-8.

15. Pruneri G, Lazzeroni M, Bagnardi V, Tiburzio GB, Rotmensz N, DeCensi A *et al.* The prevalence and clinical relevance of tumor-infiltrating lymphocytes (TILs) in ductal carcinoma in situ of the breast. *Ann Oncol* 2017; **28**: 321-328; doi 10.1093/annonc/mdw623.
16. Van Bockstal M, Lambein K, Gevaert O, De Wever O, Praet M, Cocquyt V *et al.* Stromal architecture and periductal decorin are potential prognostic markers for ipsilateral locoregional recurrence in ductal carcinoma in situ of the breast. *Histopathology* 2013; **63**: 520-533; doi 10.1111/his.12188.
17. Freeman GJ, Long AJ, Iwai Y, Iwai Y, Bourque K, Bourque K, Chernova T, Chernova T, Nishimura H, Nishimura H, Fitz LJ *et al.* Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. *J Exp Med*. 2000, 192:1027-34; doi 10.1084/jem.192.7.1027
18. Keir ME, Liang SC, Guleria I, Latchman YE, Qipo A, Albacker LA *et al.* Tissue expression of PD-L1 mediates peripheral T cell tolerance. *J Exp Med* 2006; **203**: 883-895; doi 10.1084/jem.20051776.
19. Wang X, Teng F, Kong L, Yu J. PD-L1 expression in human cancers and its association with clinical outcomes. *Oncotargets Ther* 2016; **9**: 5023-5039; doi 10.2147/OTT.S105862.
20. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF *et al.* Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med* 2012; **366**: 2443-2454; doi 10.1056/NEJMoa1200690.
21. Miligy IM, Gorringer KL, Toss MS, Al-Kawaz AA, Simpson P, Diez-Rodriguez M *et al.* Thioredoxin-interacting protein is an independent risk stratifier for breast ductal carcinoma in situ. *Mod Pathol* 2018; **31**: 1807-1815; doi 10.1038/s41379-018-0086-7.
22. Lester SC, Bose S, Chen YY, Connolly JL, de Baca ME, Fitzgibbons PL *et al.* Protocol for the examination of specimens from patients with ductal carcinoma in situ of the breast. *Arch Pathol Lab Med* 2009; **133**: 15-25; doi 10.1043/1543-2165-133.1.15.
23. Hammond ME, Hayes DF, Wolff AC, Mangu PB, Temin S. American society of clinical oncology/college of american pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Oncol Pract*. 2010; **6**: 195-197; doi 10.1200/jop.777003.
24. Rakha EA, Pinder SE, Bartlett JM, Ibrahim M, Starczynski J, Carder PJ *et al.* Updated UK Recommendations for HER2 assessment in breast cancer. *J Clin Pathol* 2015; **68**: 93-99; doi 10.1136/jclinpath-2014-202571.
25. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol* 2011; **22**: 1736-1747; doi 10.1093/annonc/mdr304.
26. Bos R, van Diest PJ, de Jong JS, van der Groep P, van der Valk P, van der Wall E. Hypoxia-inducible factor-1alpha is associated with angiogenesis, and expression of bFGF, PDGF-BB, and EGFR in invasive breast cancer. *Histopathology* 2005; **46**: 31-36; doi 10.1111/j.1365-2559.2005.02045.x.
27. van der Groep P, van Diest PJ, Smolders YH, Ausems MG, van der Luijt RB, Menko FH *et al.* HIF-1alpha overexpression in ductal carcinoma in situ of the breast in BRCA1 and BRCA2 mutation carriers. *PLoS one* 2013; **8**: e56055; doi 10.1371/journal.pone.0056055.
28. Kader T, Goode DL, Wong SQ, Connaughton J, Rowley SM, Devereux L *et al.* Copy number analysis by low coverage whole genome sequencing using ultra low-input DNA from formalin-fixed paraffin embedded tumor tissue. *Genome Med* 2016; **8**: 121-121; doi 10.1186/s13073-016-0375-z.
29. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 2013; **43**: 11-33; doi 10.1002/0471250953.bi1110s43.

30. Rimmer A, Phan H, Mathieson I, Iqbal Z, Twigg SRF, Wilkie AOM *et al.* Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet* 2014; **46**: 912-918; doi 10.1038/ng.3036.
31. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012; **22**: 568-576; doi 10.1101/gr.129684.111.
32. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A *et al.* The Ensembl Variant Effect Predictor. *Genome Bio* 2016; **17**: 122; doi 10.1186/s13059-016-0974-4.
33. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016; **536**: 285-291; doi 10.1038/nature19057.
34. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO *et al.* A global reference for human genetic variation. *Nature* 2015; **526**: 68-74; doi 10.1038/nature15393.
35. Scheinin I, Sie D, Bengtsson H, van de Wiel MA, Olshen AB, van Thuijl HF *et al.* DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res* 2014; **24**: 2022-2032; doi 10.1101/gr.175141.114.
36. Kuilman T, Velds A, Kemper K, Ranzani M, Bombardelli L, Hoogstraat M *et al.* Copywriter: DNA copy number detection from off-target sequence data. *Genome Biol.* 2015; **16**: 49; doi 10.1186/s13059-015-0617-1.
37. Boeva V, Popova T, Bleakley K, Chiche P, Cappo J, Schleiermacher G *et al.* Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* 2012; **28**: 423-425; doi 10.1093/bioinformatics/btr670.
38. Rathore AS, Kumar S, Konwar R, Makker A, Negi MP, Goel MM. CD3+, CD4+ & CD8+ tumour infiltrating lymphocytes (TILs) are predictors of favourable survival outcome in infiltrating ductal carcinoma of breast. *Indian J Med Res* 2014; **140**: 361-369;
39. Campbell MJ, Baehner F, O'Meara T, Ojukwu E, Han B, Mukhtar R *et al.* Characterizing the immune microenvironment in high-risk ductal carcinoma in situ of the breast. *Breast Cancer Res Treat* 2017; **161**: 17-28; doi 10.1007/s10549-016-4036-0.
40. Doebar SC, de Monye C, Stoop H, Rothbarth J, Willemsen SP, van Deurzen CH. Ductal carcinoma in situ diagnosed by breast needle biopsy: Predictors of invasion in the excision specimen. *Breast* 2016; **27**: 15-21; doi 10.1016/j.breast.2016.02.014.
41. Kuroda H, Tamaru J-i, Sakamoto G, Ohnisi K, Itoyama S. Immunophenotype of lymphocytic infiltration in medullary carcinoma of the breast. *Virchows Archiv* 2005; **446**: 10-14; doi 10.1007/s00428-004-1143-9.
42. Iseki Y, Shibutani M, Maeda K, Nagahara H. A new method for evaluating tumor- infiltrating lymphocytes (TILs) in colorectal cancer using hematoxylin and eosin (H-E) - stained tumor sections. *PLoS One.* 2018: 1-12; doi: 10.1371/journal.pone.0192744.
43. Matkowski R, Gisterek I, Halon A, Lacko A, Szewczyk K, Staszek U *et al.* The prognostic role of tumor-infiltrating CD4 and CD8 T lymphocytes in breast cancer. *Anticancer Res* 2009; **29**: 2445-2451;
44. Kristensen VN, Vaske CJ, Ursini-Siegel J, Van Loo P, Nordgard SH, Sachidanandam R *et al.* Integrated molecular profiles of invasive breast tumors and ductal carcinoma in situ (DCIS) reveal differential vascular and interleukin signaling. *Proc Natl Acad Sci U S A* 2012; **109**: 2802-2807; doi 10.1073/pnas.1108781108.
45. Miligy I, Mohan P, Gaber A, Aleskandarany MA, Nolan CC, Diez-rodriguez M *et al.* Prognostic significance of tumour infiltrating B lymphocytes in breast ductal carcinoma in situ. *Histopathology* 2017: 1-11; doi 10.1111/his.13217.

46. Mahmoud SMA, Paish EC, Lee AHS, Ellis IO, Green AR. An evaluation of the clinical significance of FOXP3 + infiltrating cells in human breast cancer. 2011: 99-108; doi 10.1007/s10549-010-0987-8.
47. Bates GJ, Fox SB, Han C, Leek RD, Garcia JF, Harris AL *et al.* Quantification of Regulatory T Cells Enables the Identification of High-Risk Breast Cancer Patients and Those at Risk of Late Relapse. *J Clin Oncol* 2018; **24**: 5373-5380; doi 10.1200/JCO.2006.05.9584.
48. Lal A, Chan L, Devries S, Chin K, Scott GK, Benz CC *et al.* FOXP3-positive regulatory T lymphocytes and epithelial FOXP3 expression in synchronous normal , ductal carcinoma in situ , and invasive cancer of the breast. *Breast Cancer Res Treat.* 2013: 381-390; doi 10.1007/s10549-013-2556-4.
49. Gil Del Alcazar CR, Huh SJ, Ekram MB, Trinh A, Liu LL, Beca F *et al.* Immune Escape in Breast Cancer During In Situ to Invasive Carcinoma Transition. *Cancer Discov* 2017; **7**: 1098-1115; doi 10.1158/2159-8290.cd-17-0222.
50. Chen XY, Yeong J, Thike AA, Bay BH, Tan PH. Prognostic role of immune infiltrates in breast ductal carcinoma in situ. *Breast Cancer Res Treat* 2019; **177**: 17-27; doi 10.1007/s10549-019-05272-2.
51. Merlo A, Casalini P, Carcangiu ML, Malventano C, Triulzi T, Me S. FOXP3 Expression and Overall Survival in Breast Cancer. *J Clin Oncol* 2018; **27**; doi 10.1200/JCO.2008.17.9036.
52. Curiel TJ, Coukos G, Zou L, Alvarez X, Cheng P, Mottram P *et al.* Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. *Nature Med* 2004; **10**: 942-949; doi 10.1038/nm1093.
53. Dermani FK, Samadi P, Rahmani G, Kohlan AK, Najafi R. PD-1/PD-L1 immune checkpoint: Potential target for cancer therapy. *J Cell Physiol* 2019; **234**: 1313-1325; doi 10.1002/jcp.27172.
54. Nomi T, Sho M, Akahori T, Hamada K, Kubo A, Kanehiro H *et al.* Clinical Significance and Therapeutic Potential of the Programmed Death-1 Ligand/Programmed Death-1 Pathway in Human Pancreatic Cancer. *Clin Cancer Res* 2007; **13**: 2151; doi 10.1158/1078-0432.CCR-06-2746.
55. Thompson RH, Gillett Md, Cheville JC, Lohse CM, Dong H, Webster WS, *et al.* Costimulatory B7-H1 in renal cell carcinoma patients: Indicator of tumor aggressiveness and potential therapeutic target. *Proc Natl Acad Sci.* 2004, 101:17174-9; doi 10.1073/pnas.0406351101
56. Trella E, Droeser RA, Muenst S, Schaerli AR, Gao F, Muraro MG *et al.* Expression of programmed death ligand 1 (PD-L1) is associated with poor prognosis in human breast cancer. *Breast Cancer Res Treat* 2014; **1**; doi 10.1007/s10549-014-2988-5.
57. Thompson RH, Dong H, Lohse CM, Leibovich BC, Blute ML, Cheville JC *et al.* PD-1 is expressed by tumor-infiltrating immune cells and is associated with poor outcome for patients with renal cell carcinoma. *Clin Cancer Res.* 2007 15:1757-61, doi 10.1158/1078-0432.CCR-06-2599
58. Muenst S, Soysal SD, Gao F, Obermann EC, Oertli D, EGillanders W. The presence of programmed death 1 (PD-1)-positive tumor-infiltrating lymphocytes is associated with poor prognosis in human breast cancer. *Breast Cancer Res Treat.* 2013 ;139:667-76; doi: 10.1007/s10549-013-2581-3.
59. Schalper KA, Velcheti V, Carvajal D, Wimberly H, Brown J, Puztai L *et al.* In Situ Tumor PD-L1 mRNA Expression Is Associated with Increased TILs and Better Outcome in Breast Carcinomas. *Clin Cancer Res* 2014: 2773-2783; doi 10.1158/1078-0432.CCR-13-2702.
60. Mittendorf EA, Philips AV, Meric-Bernstam F, Qiao N, Wu Y, Harrington S *et al.* PD-L1 Expression in Triple-Negative Breast Cancer. *Cancer Immunol Res* 2014; **2**: 361-370; doi 10.1158/2326-6066.CIR-13-0127.
61. Man YG, Sang QX. The significance of focal myoepithelial cell layer disruptions in human breast tumor invasion: a paradigm shift from the "protease-centered" hypothesis. *Exp Cell Res* 2004; **301**: 103-118; doi 10.1016/j.yexcr.2004.08.037.
62. Man YG, Stojadinovic A, Mason J, Avital I, Bilchik A, Bruecher B *et al.* Tumor-infiltrating immune cells promoting tumor invasion and metastasis: existing theories. *J Cancer* 2013; **4**: 84-95; doi 10.7150/jca.5482.

63. Thompson E, Taube JM, Elwood H, Sharma R, Meeker A, Warzecha HN *et al.* The immune microenvironment of breast ductal carcinoma in situ. *Mod pathol* 2016; **29**: 249-258; doi 10.1038/modpathol.2015.158.
64. Ma X-J, Dahiya S, Richardson E, Erlander M, Sgroi DC. Gene expression profiling of the tumor microenvironment during breast cancer progression. *Breast Cancer Res* 2009; **11**: R7; doi 10.1186/bcr2222.
65. Sabatier R, Finetti P, Mamessier E, Adelaide J, Chaffanet M, Ali HR *et al.* Prognostic and predictive value of PDL1 expression in breast cancer. *Oncotarget* 2015; **6**: 5449-5464; doi 10.18632/oncotarget.3216.
66. Triner D, Shah YM. Hypoxia-inducible factors: a central link between inflammation and cancer. *J Clin Invest* 2016; **126**: 3689-3698; doi 10.1172/jci84430.
67. Chouaib S, Umansky V, Kieda C. The role of hypoxia in shaping the recruitment of proangiogenic and immunosuppressive cells in the tumor microenvironment. *Contemp Oncol (Pozn)* 2018; **22**: 7-13; doi 10.5114/wo.2018.73874.
68. Sitkovsky M, Lukashev D. Regulation of immune cells by local-tissue oxygen tension: HIF1 alpha and adenosine receptors. *Nat Rev Immunol* 2005; **5**: 712-721; doi 10.1038/nri1685.
69. Knopfmacher A, Fox J, Lo Y, Shapiro N, Fineberg S. Correlation of histopathologic features of ductal carcinoma in situ of the breast with the oncotype DX DCIS score. *Mod Pathol* 2015; **28**: 1167-1173; doi 10.1038/modpathol.2015.79.
70. Kasthuber ER, Lowe SW. Putting p53 in Context. *Cell* 2017; **170**: 1062-1078; doi 10.1016/j.cell.2017.08.028.

Figure legends

Figure 1: Photomicrograph for immune cells and checkpoint proteins in DCIS; A) Dense Overall TILs infiltration in H&E stained sections (20x). Dense stromal and intratumoral infiltration of various immune cell markers, B) CD3, C) CD4, D) CD8, E) CD20, F) FOXP3, G) PD1 and H) PDL1. Expression of PDL1 within the tumor epithelial cells is shown in (I). Green arrows highlight the intratumoral lymphocytes. Inset: higher magnification (40x) showing stromal FOXP3 positive cells.

Figure 2: A) Correlation Matrix showing the correlation between TILs subtypes, overall TILs, and other associated biomarkers. (s; stromal, i; intratumoral). Blue colour refers to positive correlation, while the red colour reflects negative correlations. The intensity of the colour proportionate to the correlation coefficient. **B)** Tornado (Butterfly) plot showing the differential expression level of various lymphocytes between pure DCIS and DCIS mixed with synchronous invasive breast cancer (IBC).

Figure 3: Violin plots showing the association between *TP53* variants and (A) stromal CD3+ cells, (B) stromal CD4+ cells, (C) stromal FOXP3+ cells and (D) stromal PDL1+ TILs. The lower figures show the association between fraction of genome altered and (E) overall TILs density and (F) stromal CD3+ cells. The central red dots represent the mean, boxes represent the interquartile range, central line represent the median and whiskers shows the 95% confidence interval.

Figure 4: Kaplan Meier Curves showing; Higher stromal FOXP3+ cells was associated with poor outcome for all DCIS recurrences (A), Dense stromal PDL1 expression was associated with shorter local recurrence free interval (LRFI) either for all recurrences (B), or invasive recurrences (C). All analyses were carried out on patients treated with breast conserving surgery with or without radiotherapy.

Table 1: Univariate survival analysis between overall TILs and various immune cell markers with the local recurrence free interval in DCIS patients treated with breast conserving surgery

Marker	All recurrences				Invasive recurrences			
	Hazard ratio	95% Confidence interval		p-value	Hazard ratio	95% Confidence interval		p-value
		Lower	Upper			Lower	Upper	
Overall TILs	2.9	1.7	5.0	0.0001	2.1	1.1	4.3	0.040
sCD20	0.9	0.5	1.7	0.793	0.9	0.4	1.9	0.770
sCD3	1.1	0.6	1.9	0.877	0.8	0.4	1.7	0.563
sCD8	0.9	0.5	1.8	0.992	0.5	0.2	1.2	0.109
sCD4	1.5	0.9	2.8	0.150	1.2	0.5	2.5	0.708
sFOXP3	2.1	1.1	3.7	0.025	1.1	0.4	2.6	0.895
sPD1	1.3	0.7	2.3	0.412	0.7	0.3	1.6	0.347
sPDL-1	4.4	2.4	8.1	<0.0001	2.4	1.1	5.3	0.032
iCD20	1.2	0.7	2.2	0.520	1.1	0.5	2.3	0.937
iCD3	1.2	0.7	2.2	0.522	0.9	0.5	2.1	0.979
iCD8	1.1	0.6	1.9	0.839	0.7	0.3	1.7	0.529
iCD4	1.2	0.7	2.1	0.572	1.1	0.5	2.2	0.985
iFOXP3	1.6	0.8	3.3	0.180	0.7	0.2	2.5	0.617
iPD1	0.9	0.4	2.1	0.766	0.7	0.2	2.5	0.652

s: Stromal, i: Intratumoral.

Significant p values are in bold

Table 2: Multivariate survival analysis (Cox regression model) of variables (with and without immune cell markers) predicting outcome in terms of ipsilateral local all recurrences (A) and invasive recurrences (B) in patients treated by breast conserving surgery in pure DCIS.

A) All recurrences

Conventional clinicopathological parameters associated with high risk DCIS				
Parameters	Hazard ratio (HR)	95% confidence interval (CI)		p-value
		Lower	Upper	
Patient Age	0.4	0.2	0.8	0.006
DCIS presentation*	1.5	0.9	2.4	0.111
DCIS size	1.5	1.1	2.1	0.040
DCIS nuclear Grade	1.9	1.3	2.7	0.001
Comedo necrosis	0.6	0.4	0.9	0.049
Margin status	0.8	0.7	0.9	0.004
TILs density and other clinicopathological parameters				
Patient Age	0.4	0.1	0.8	0.023
DCIS presentation	1.5	0.9	3.8	0.395
DCIS size	0.8	0.4	1.6	0.498
DCIS nuclear Grade	1.2	0.6	2.4	0.662
Comedo necrosis	0.6	0.3	1.3	0.196
Margin status	0.9	0.9	1.1	0.637
Overall TILs density	7.0	2.8	1.2	0.024
Dense stromal FOXP3+ cells	1.7	0.7	0.3	0.453
Dense stromal PDL1+ cells	6.8	3.1	1.4	0.004
Overall TILs density with stromal FOXP3+ and stromal PDL1+ inflammatory cells				
Overall TILs density	3.5	1.5	8.2	0.003
Dense stromal FOXP3+ cells	0.8	0.4	1.7	0.554
Dense stromal PDL1+ cells	2.9	1.4	6.1	0.005

B) Invasive recurrence

Conventional clinicopathological parameters associated with high risk DCIS				
Parameters	Hazard ratio (HR)	95% confidence interval (CI)		p-value
		Lower	Upper	
Patient Age	0.5	0.2	1.3	0.156
DCIS presentation	1.4	0.8	2.6	0.245
DCIS size	1.8	1.1	2.4	0.013
DCIS nuclear Grade	1.9	1.1	3.0	0.013
Comedo necrosis	0.7	0.4	1.3	0.274
Margin status	0.9	0.8	1.1	0.075
TILs density and other clinicopathological parameters				
Patient Age	0.7	0.2	2.4	0.527
DCIS presentation	2.2	0.6	7.5	0.206
DCIS size	1.7	0.6	4.7	0.301
DCIS nuclear Grade	2.6	0.8	8.2	0.111
Comedo necrosis	0.7	0.2	2.0	0.497
Margin status	0.9	0.9	1.1	0.877
Overall TILs density	2.1	1.2	5.5	0.029
Dense stromal FOXP3+ cells	0.3	0.1	1.4	0.143
Dense stromal PDL1+ cells	1.7	0.6	5.1	0.307
Overall TILs density with stromal FOXP3+ and stromal PDL1+ inflammatory cells				
Overall TILs density	2.8	1.1	7.9	0.045
Dense stromal FOXP3+ cells	0.4	0.1	1.5	0.176
Dense stromal PDL1+ cells	1.7	0.6	4.8	0.318

Significant p values are in bold

* DCIS presentation refers to the method of DCIS diagnosis either through the mammography screening program or symptomatic presentation.

Figure 1

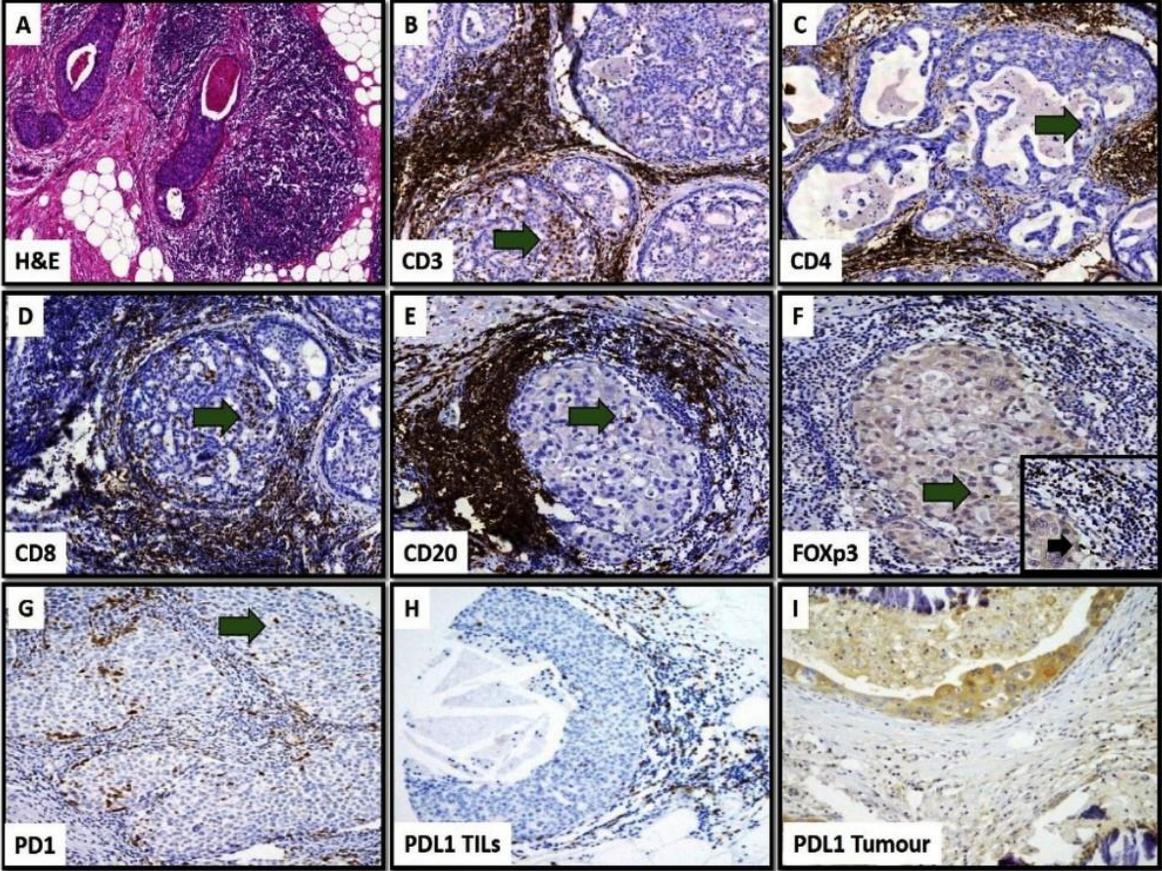


Figure 2

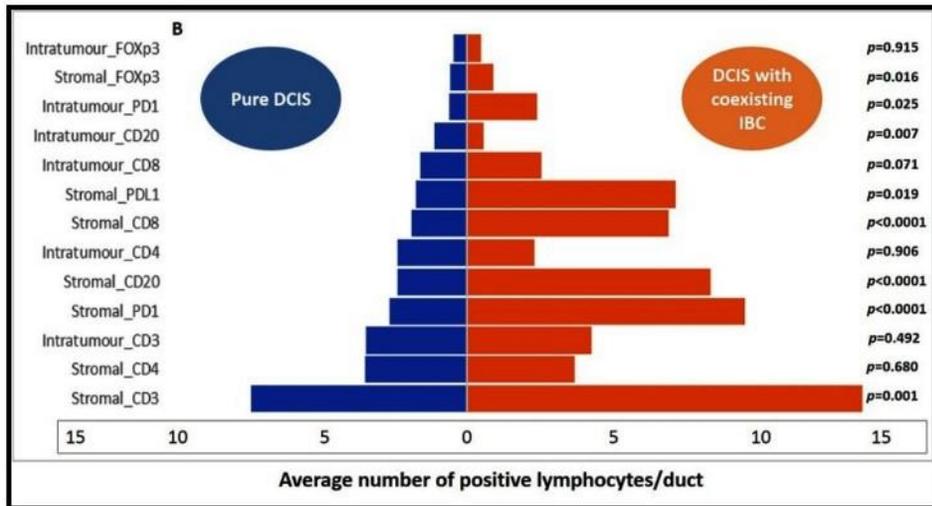
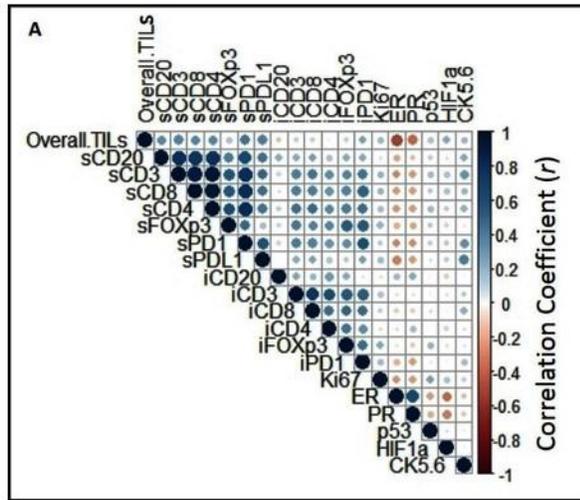


Figure 3

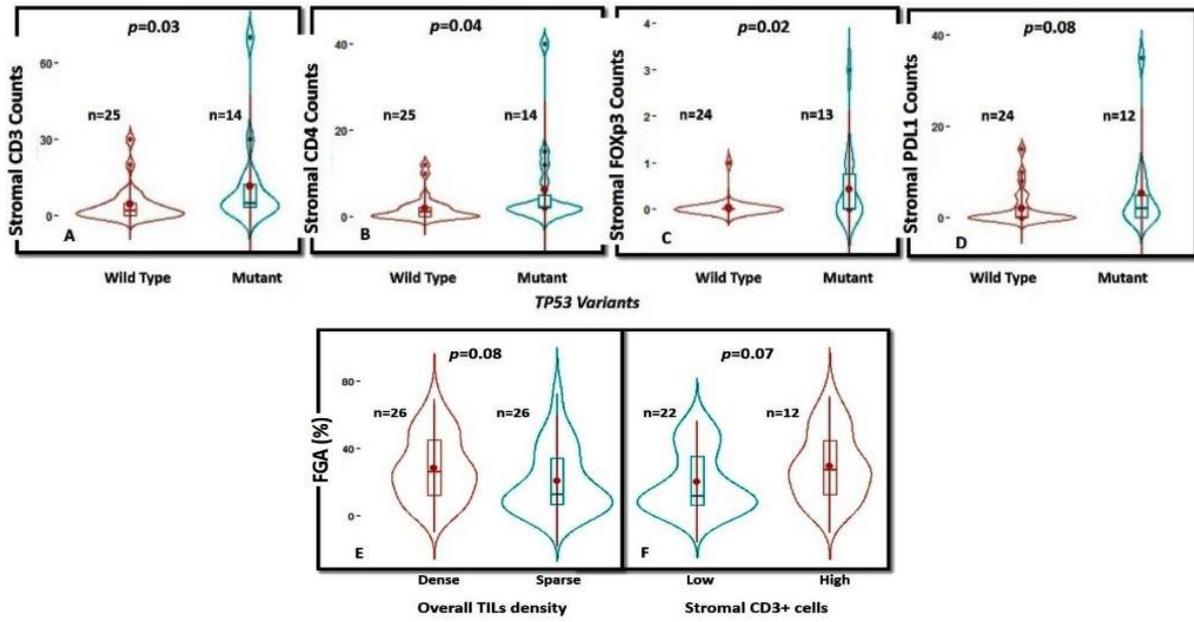
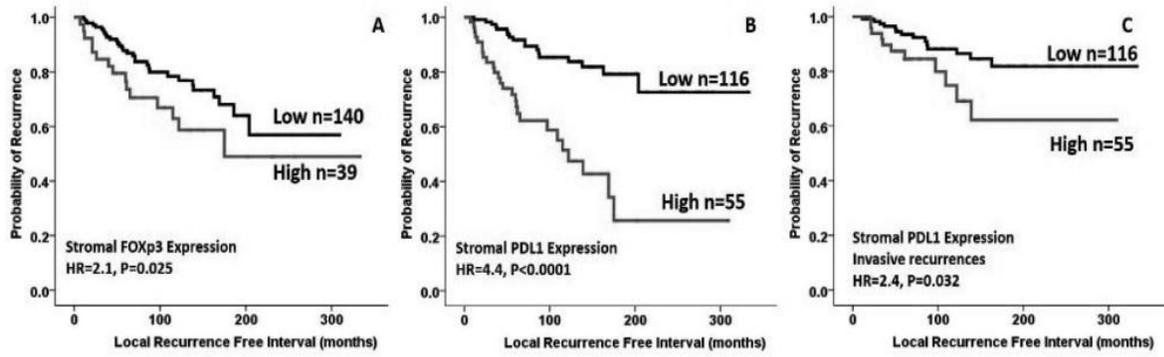


Figure 4



Supplementary Table 1: Clinicopathological characteristics of the study cohort

Parameters	Whole Pure DCIS cohort (n= 780) (n/%)	Pure DCIS cohort included in this study* (n=413) (n/%)	Whole DCIS/IBC cohort** (n=239) (n/%)	DCIS/IBC cohort included in this study (n=154) (n/%)
Age				
≤50 years	172 (22)	113 (27)	108 (45)	73 (47)
>50 years	608 (78)	300 (73)	131 (55)	81 (53)
Presentation				
Screening	430 (55)	186 (45)	NA	NA
Symptomatic	350 (45)	227 (55)		
DCIS Size				
≤20mm	392 (51)	168 (41)	88 (35)	53 (34)
>20mm	381 (49)	245 (59)	155 (65)	101 (66)
Nuclear Grade				
Low	104 (13)	54 (13)	12 (5)	8 (5)
Moderate	194 (25)	107 (26)	60 (25)	38 (25)
High	482 (62)	252 (61)	167 (70)	108 (70)
Comedo necrosis***				
Yes	497 (64)	273 (66)	175 (73)	111 (72)
No	283 (36)	140 (34)	64 (27)	43 (28)
Management				
Breast conserving surgery	376 (48)	181 (44)	NA	NA
Mastectomy	404 (52)	232 (56)		
Estrogen receptor				
Negative	148 (25)	102 (21)	NA	NA
Positive	439 (75)	260 (79)		
Progesterone receptor				
Negative	248 (42)	166 (45)	NA	NA
Positive	336 (58)	200 (55)		
HER2 status				
Negative	516 (80)	288 (75)	NA	NA
Positive	130 (20)	96 (25)		
Molecular classes				
Luminal A	248 (52)	159 (49)	NA	NA
Luminal B	88 (19)	66 (20)		
HER2 Enriched	64 (14)	53 (16)		
Triple Negative	73 (15)	47 (15)		

*Refers to the number of cases included in the TMA used in this study as the number differs from the original cohort. The actual number of cases with informative TMA cores included in each study varies due to loss of some cores during sectioning and tissue processing prior to staining and was reduced with more markers stained.

**DCIS component in the mixed cases

***Comedo necrosis was defined as presence of central acellular necrosis with nuclear debris that can be seen using low power examination. Tiny / punctate necrosis or focal apoptosis was not considered as comedo necrosis.

NA: No data available for the DCIS component in these cases, as the molecular classification and treatment were related to the invasive component not the DCIS component

Supplementary Table 2: Different antibodies used and their staining protocols

Antibody	Clone	Source	Species	Antigen retrieval	Dilution	Incubation Time
CD3	SP7	Abcam	Rabbit monoclonal	Microwave in citrate PH6	1:300	24h
CD8	C8/144B	Dako	Mouse monoclonal	Microwave in citrate PH6	1:300	60min
FOXP3	236A/E7	Abcam	Mouse monoclonal	Microwave in citrate PH6	1:300	24h
CD20	L26	Dako	Mouse monoclonal	Microwave in citrate PH6	1:300	30min
CD4	EPR6855	Abcam	Rabbit monoclonal	Microwave in citrate PH6	1:500	24h
PD1	EH33	Cell signalling	Mouse monoclonal	Microwave in citrate PH6	1:75	24h
PDL1	E1L3N	Cell signalling	Rabbit monoclonal	EDTA PH9, water bath heating at 100c	1:25	24h

Supplementary Table 3: Distribution of various TILs in the pure DCIS cohort

	sCD20	iCD20	sCD3	iCD3	sCD8	iCD8	sCD4	iCD4	sFOXP3	iFOXP3	sPD1	iPD1	sPDL1	tPDL1
Median	0	0	3	1	0	0	0	0	0	0	0	0	0	0
Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Maximum	60	40	100	50	30	20	60	25	25	10	40	15	35	100
% Pure DCIS with high expression	35	26	46	50	41	37	44	41	20	16	39	14	31	2

Numbers represent average number of cells / duct except PDL1 in the tumour which represent the percentage of positive cells. s; stromal, i; intratumoural and t; tumour epithelial cells

Supplementary Table 4 (a-d): Correlation between immune cell markers, immune checkpoints with the clinicopathological parameters in pure DCIS cohort

a) Parameters	Stromal CD20 +ve cells		χ^2 (p-value)	Intratumour CD20 +ve cells		χ^2 (p-value)	Stromal CD3 +ve cells		χ^2 (p-value)	Intratumour CD3 +ve cells		χ^2 (p-value)
	Low (N=268) N. (%)	High (n=145) N. (%)		Low (N=305) N. (%)	High (n=108) N. (%)		Low (N=207) N. (%)	High (n=203) N. (%)		Low (N=207) N. (%)	High (n=203) N. (%)	
Age ≤50 years >50 years	66 (25) 202 (75)	47 (32) 98 (68)	2.9 (0.090)	79 (26) 226 (74)	34 (31) 74 (69)	1.2 (0.264)	55 (25) 165 (75)	56 (30) 134 (70)	1.1 (0.309)	52 (25) 155 (75)	59 (29) 144 (71)	0.8 (0.369)
Presentation Screening Symptomatic	126 (47) 142 (53)	60 (41) 85 (59)	1.2 (0.272)	138 (45) 167 (55)	48 (44) 60 (56)	0.021 (0.886)	101 (50) 119 (50)	84 (44) 106 (56)	0.1 (0.730)	92 (44) 115 (56)	93 (46) 110 (54)	0.1 (0.781)
DCIS Size ≤20mm >20mm	109 (41) 158 (59)	59 (41) 85 (59)	0.1 (0.977)	116 (38) 187 (62)	52 (48) 56 (52)	3.2 (0.073)	92 (42) 128 (58)	75 (40) 113 (60)	0.2 (0.694)	74 (36) 132 (64)	93 (46) 109 (54)	4.3 (0.038)
Nuclear Grade Low Moderate High	43 (15) 81 (28) 168 (57)	9 (6) 27 (19) 85 (75)	19.1 (<0.0001)	36 (12) 81 (27) 188 (61)	16 (15) 27 (25) 65 (60)	0.676 (0.713)	41 (18) 71 (33) 108 (49)	10 (5) 35 (18) 145 (77)	34.5 (<0.0001)	26 (13) 57 (28) 124 (59)	25 (12) 49 (24) 129 (64)	0.7 (0.711)
Comedo necrosis Yes No	163 (61) 105 (39)	110 (76) 35 (24)	9.5 (0.002)	207 (68) 98 (32)	66 (61) 42 (39)	1.6 (0.202)	132 (60) 88 (40)	139 (73) 51 (27)	7.9 (0.005)	140 (68) 67 (32)	131 (65) 72 (35)	0.4 (0.507)
Estrogen receptor Negative Positive	45 (19) 188 (81)	57 (44) 72 (56)	25.4 (<0.0001)	83 (32) 180 (68)	19 (19) 80 (81)	5.4 (0.020)	28 (15) 160 (85)	73 (42) 99 (58)	33.8 (<0.0001)	53 (30) 124 (70)	48 (26) 135 (74)	0.6 (0.433)
Progesterone receptor Negative Positive	87 (37) 151 (63)	79 (62) 49 (38)	21.3 (<0.0001)	132 (49) 135 (51)	34 (34) 65 (66)	6.6 (0.010)	62 (32) 131 (68)	103 (61) 67 (39)	29.5 (<0.0001)	86 (48) 94 (52)	79 (43) 104 (57)	0.8 (0.378)
HER2 status * Negative Positive	202 (80) 49 (20)	86 (65) 47 (35)	11.6 (0.001)	205 (73) 76 (27)	83 (81) 20 (19)	2.3 (0.126)	170 (83) 34 (17)	115 (65) 62 (35)	16.9 (<0.0001)	144 (75) 49 (25)	141 (75) 47 (25)	0.1 (0.930)
Molecular classes Luminal A Luminal B HER2 Enriched Triple Negative	119 (58) 42 (20) 22 (11) 23 (11)	40 (34) 24 (20) 31 (26) 24 (20)	24.1 (<0.0001)	106 (46) 45 (19) 43 (19) 38 (16)	53 (57) 21 (23) 10 (12) 9 (10)	6.6 (0.086)	101 (62) 32 (20) 14 (9) 14 (9)	57 (35) 34 (21) 39 (24) 32 (20)	31.1 (<0.0001)	71 (46) 31 (20) 27 (18) 25 (16)	87 (52) 35 (21) 26 (15) 21 (12)	1.5 (0.674)
HIF1a expression Low High	207 (83) 44 (17)	90 (63) 53 (37)	18.7 (<0.0001)	207 (72) 80 (28)	90 (84) 17 (16)	6.3 (0.014)	171 (84) 33 (16)	126 (66) 64 (34)	16.8 (<0.0001)	144 (75) 49 (25)	151 (76) 48 (24)	0.1 (0.771)

b) Parameters	Stromal CD4 +ve cells		χ^2 (p-value)	Intratumour CD4 +ve cells		χ^2 (p-value)	Stromal CD8 +ve cells		χ^2 (p-value)	Intratumour CD8 +ve cells		χ^2 (p-value)
	Low (N=225) N. (%)	High (N=178) N. (%)		Low (N=233) N. (%)	High (N=170) N. (%)		Low (N=238) N. (%)	High (n=164) N. (%)		Low (N=253) N. (%)	High (n=149) N. (%)	
Age ≤50 years >50 years	55 (24) 170 (76)	57 (32) 121 (68)	2.8 (0.092)	65 (28) 168 (72)	47 (28) 123 (72)	0.1 (0.956)	54 (23) 184 (77)	57 (35) 107 (65)	6.6 (0.010)	65 (26) 188 (74)	46 (31) 103 (69)	1.3 (0.262)
Presentation Screening Symptomatic	106 (47) 119 (53)	75 (42) 103 (58)	0.9 (0.363)	103 (44) 130 (56)	78 (46) 92 (54)	0.1 (0.738)	114 (48) 124 (52)	66 (40) 98 (60)	2.3 (0.129)	107 (42) 146 (58)	73 (49) 76 (51)	1.7 (0.192)
DCIS Size ≤20mm >20mm	88 (39) 137 (61)	75 (43) 101 (57)	0.6 (0.424)	83 (36) 150 (64)	80 (48) 88 (52)	5.8 (0.016)	94 (40) 144 (60)	67 (41) 95 (59)	0.1 (0.773)	92 (36) 161 (64)	69 (47) 78 (53)	4.3 (0.038)
Nuclear Grade Low Moderate High	40 (18) 75 (34) 110 (48)	11 (6) 29 (16) 138 (78)	36.9 (<0.0001)	29 (12) 63 (27) 141 (61)	22 (13) 41 (24) 107 (63)	0.4 (0.803)	43 (19) 71 (29) 123 (52)	7 (4) 35 (21) 123 (75)	26.0 (<0.0001)	37 (15) 72 (29) 144 (56)	13 (9) 34 (23) 102 (68)	5.8 (0.055)
Comedo necrosis Yes No	143 (63) 82 (37)	125 (70) 53 (30)	2.2 (0.134)	165 (71) 68 (29)	103 (61) 67 (39)	4.6 (0.032)	148 (63) 90 (37)	119 (73) 45 (27)	4.3 (0.037)	168 (66) 85 (34)	99 (66) 50 (34)	0.1 (0.993)
Estrogen receptor Negative Positive	28 (14) 168 (86)	73 (46) 86 (54)	44.8 (<0.0001)	61 (31) 138 (69)	40 (26) 116 (74)	1.1 (0.299)	38 (19) 173 (81)	62 (43) 82 (57)	25.1 (<0.0001)	54 (25) 165 (75)	46 (34) 90 (66)	3.5 (0.062)
Progesterone receptor Negative Positive	66 (33) 132 (67)	96 (60) 64 (40)	25.0 (<0.0001)	101 (49) 106 (51)	61 (40) 90 (60)	2.5 (0.115)	77 (36) 139 (64)	85 (59) 59 (41)	19.0 (<0.0001)	93 (41) 132 (59)	69 (51) 66 (49)	3.3 (0.071)
HER2 status * Negative Positive	175 (83) 36 (17)	107 (65) 58 (35)	14.3 (<0.0001)	159 (73) 60 (27)	123 (78) 34 (22)	1.6 (0.205)	184 (82) 40 (18)	97 (64) 54 (36)	15.4 (0.001)	182 (77) 54 (23)	99 (71) 40 (29)	1.6 (0.203)
Molecular classes Luminal A Luminal B HER2 Enriched Triple Negative	110 (65) 33 (19) 13 (8) 15 (8)	47 (31) 32 (21) 40 (27) 31 (21)	44.4 (<0.0001)	80 (45) 38 (22) 32 (18) 27 (15)	77 (53) 27 (19) 21 (15) 19 (13)	2.2 (0.527)	120 (63) 31 (16) 18 (10) 21 (11)	37 (28) 34 (26) 35 (27) 24 (19)	40.3 (<0.0001)	102 (54) 35 (18) 29 (15) 24 (13)	55 (42) 30 (23) 24 (19) 21 (16)	4.1 (0.260)
HIF1a expression Low High	180 (85) 33 (15)	112 (64) 63 (36)	24.1 (<0.0001)	168 (76) 53 (24)	124 (74) 43 (26)	0.2 (0.690)	196 (85) 32 (15)	97 (61) 62 (39)	29.1 (<0.0001)	188 (78) 52 (22)	105 (71) 42 (29)	2.4 (0.124)

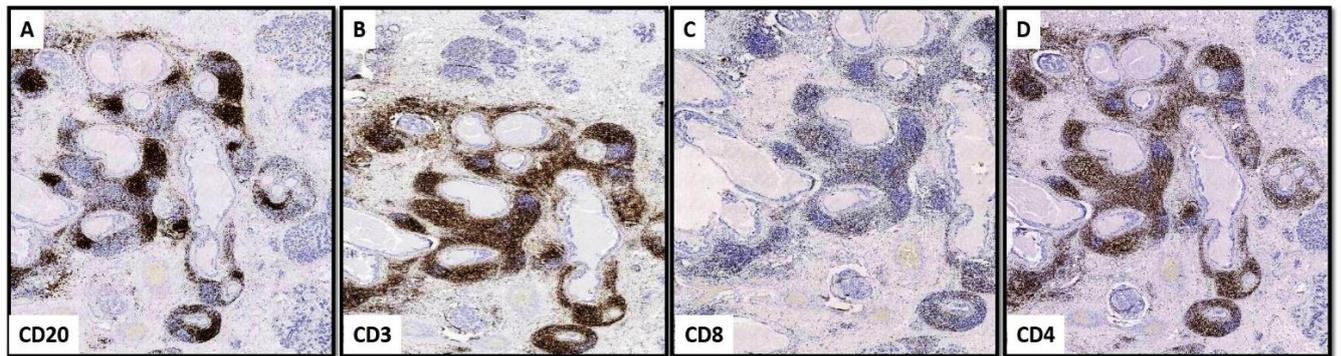
c) Parameters	Stromal FOXP3 +ve cells		χ^2 (p-value)	Intratumour FOXP3 +ve cells		χ^2 (p-value)
	Low (N=327) N. (%)	High (N=79) N. (%)		Low (N=330) N. (%)	High (N=66) N. (%)	
Age ≤50 years >50 years	91 (28) 236 (72)	20 (25) 59 (75)	0.2 (0.681)	91 (27) 239 (73)	20 (30) 46 (70)	0.3 (0.593)
Presentation Screening Symptomatic	143 (44) 180 (56)	39 (49) 40 (51)	0.7 (0.405)	156 (46) 180 (54)	26 (39) 40 (61)	1.1 (0.294)
DCIS Size ≤20mm >20mm	132 (41) 191 (59)	31 (40) 46 (60)	0.1 (0.931)	134 (40) 201 (60)	29 (45) 36 (55)	0.5 (0.488)
Nuclear Grade Low Moderate High	48 (15) 90 (28) 185 (57)	1 (1) 13 (17) 65 (82)	19.9 (<0.0001)	47 (14) 92 (27) 197 (59)	2 (3) 11 (16) 53 (81)	12.1 (0.002)
Comedo necrosis Yes No	204 (63) 119 (37)	64 (19) 15 (81)	9.4 (0.002)	218 (65) 118 (35)	50 (76) 16 (24)	2.9 (0.087)
Estrogen receptor Negative Positive	67 (24) 214 (76)	34 (53) 39 (47)	15.1 (<0.0001)	78 (27) 216 (73)	23 (38) 37 (62)	3.4 (0.049)
Progesterone receptor Negative Positive	114 (40) 173 (60)	49 (69) 22 (31)	19.8 (<0.0001)	127 (42) 172 (53)	36 (61) 23 (39)	6.8 (0.009)
HER2 status * Negative Positive	240 (80) 64 (20)	40 (58) 31 (42)	13.7 (<0.0001)	244 (78) 70 (22)	36 (59) 25 (41)	9.4 (0.002)
Molecular classes Luminal A Luminal B HER2 Enriched Triple Negative	140 (55) 48 (20) 31 (12) 35 (13)	17 (25) 17 (25) 22 (33) 11 (17)	24.6 (<0.0001)	139 (53) 46 (17) 39 (15) 37 (15)	17 (29) 18 (31) 14 (24) 9 (16)	12.3 (0.006)
HIF1a expression Low High	251 (81) 58 (19)	39 (50) 39 (50)	32.7 (<0.0001)	256 (80) 66 (20)	34 (52) 31 (48)	21.3 (<0.0001)

d) Parameters	Stromal PD1 +ve cells		χ^2 (p-value)	Intratumour PD1 +ve cells		χ^2 (p-value)	Stromal PDL1 +ve cells		χ^2 (p-value)
	Low (N=241) N. (%)	High (n=152) N. (%)		Low (N=338) N. (%)	High (n=55) N. (%)		Low (N=264) N. (%)	High (n=119) N. (%)	
Age ≤50 years >50 years	63 (26) 178 (74)	47 (31) 105 (69)	0.9 (0.328)	95 (28) 243 (72)	15 (27) 40 (73)	0.1 (0.898)	65 (25) 199 (75)	40 (34) 79 (66)	3.3 (0.068)
Presentation Screening Symptomatic	114 (47) 127 (53)	64 (42) 88 (58)	0.8 (0.357)	154 (46) 184 (54)	24 (44) 31 (56)	0.1 (0.790)	128 (49) 136 (51)	45 (38) 74 (62)	3.8 (0.052)
DCIS Size ≤20mm >20mm	97 (41) 144 (59)	59 (39) 91 (61)	0.6 (0.806)	132 (39) 206 (61)	24 (45) 29 (55)	0.742 (0.389)	114 (43) 150 (57)	39 (33) 79 (67)	3.8 (0.062)
Nuclear Grade Low Moderate High	40 (17) 83 (35) 118 (48)	9 (6) 22 (15) 121 (79)	37.7 (<0.0001)	48 (14) 95 (28) 195 (58)	1 (2) 10 (18) 44 (80)	11.4 (0.003)	42 (16) 79 (30) 143 (54)	7 (6) 22 (19) 90 (75)	16.7 (<0.0001)
Comedo necrosis Yes No	146 (60) 95 (40)	117 (77) 35 (23)	11.7 (0.001)	223 (66) 115 (34)	40 (73) 15 (27)	0.9 (0.324)	169 (64) 95 (36)	86 (72) 33 (28)	2.5 (0.113)
Estrogen receptor Negative Positive	32 (15) 176 (85)	66 (46) 76 (54)	41.6 (<0.0001)	76 (25) 222 (75)	22 (42) 30 (58)	6.2 (0.013)	46 (20) 187 (80)	53 (49) 55 (51)	30.8 (<0.0001)
Progesterone receptor Negative Positive	66 (30) 149 (70)	95 (67) 46 (33)	47.8 (<0.0001)	125 (41) 180 (59)	36 (71) 15 (29)	15.5 (<0.0001)	82 (35) 152 (65)	74 (68) 35 (32)	32.4 (<0.0001)
HER2 status * Negative Positive	187 (84) 37 (16)	85 (60) 57 (40)	12.9 (<0.0001)	243 (77) 71 (23)	29 (56) 23 (44)	10.9 (0.001)	199 (82) 44 (18)	64 (58) 47 (42)	23.4 (<0.0001)
Molecular classes Luminal A Luminal B HER2 Enriched Triple Negative	117 (64) 34 (19) 14 (7) 18 (10)	39 (29) 30 (23) 40 (30) 24 (18)	47.4 (<0.0001)	137 (51) 55 (21) 39 (15) 35 (13)	19 (38) 9 (18) 15 (30) 7 (14)	7.5 (0.057)	123 (60) 37 (18) 23 (11) 21 (11)	25 (25) 25 (25) 30 (29) 22 (21)	38.4 (<0.0001)
HIF1a expression Low High	191 (85) 34 (15)	91 (61) 59 (39)	28.1 (<0.0001)	251 (78) 70 (22)	31 (57) 23 (43)	10.7 (0.001)	208 (85) 37 (15)	65 (55) 53 (45)	37.9 (<0.0001)

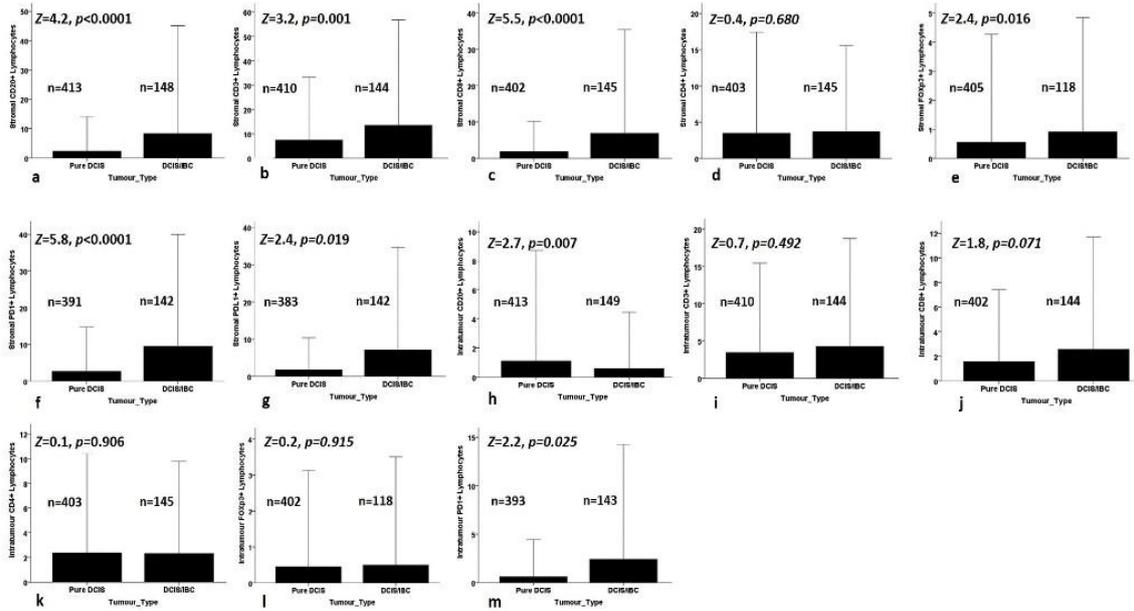
Significant p values are in bold.

DCIS; ductal carcinoma in situ, HER2; Human epidermal growth factor receptor 2, HIF1A; hypoxia inducible factor 1 alpha.

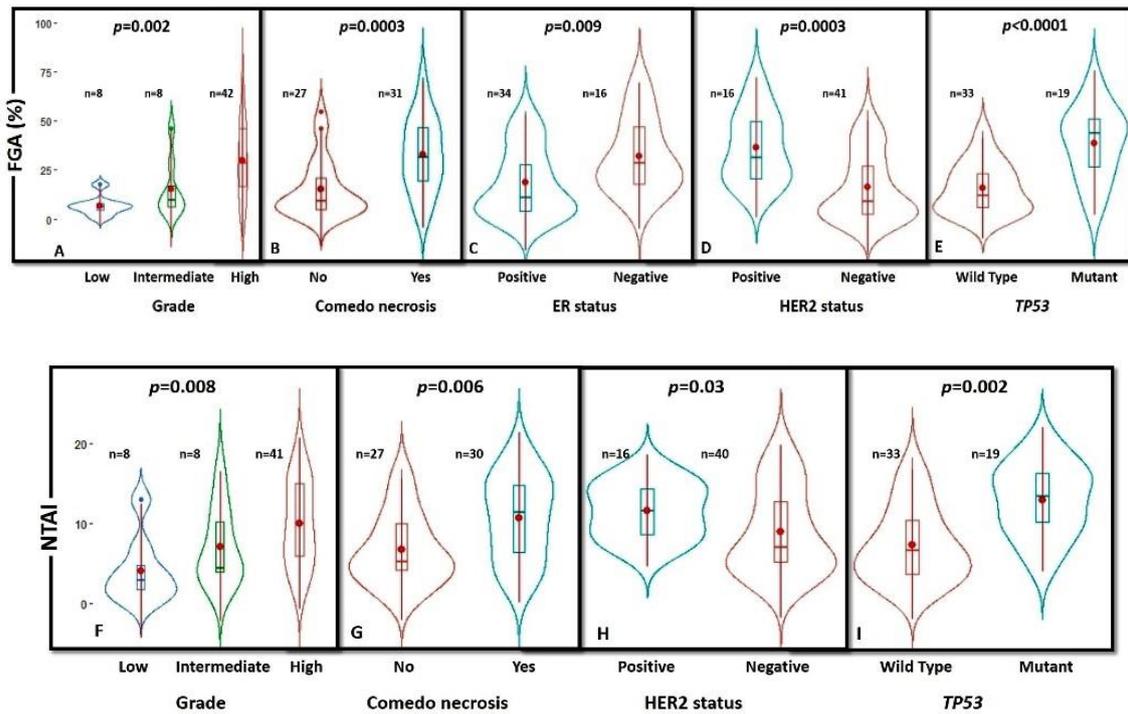
Supplementary Figures



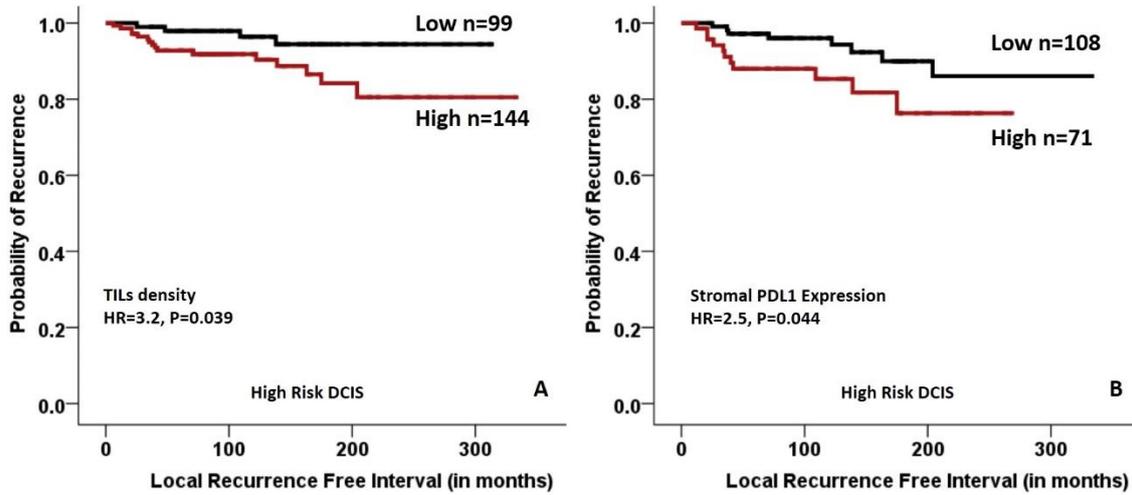
Supplementary Figure 1: Low power view (4x), showing distribution of CD20, CD3, CD4 and CD8 in a DCIS case. It is noticed that CD20 and CD3 are the most predominant subtypes. Also, CD4+ cells are more predominant T lymphocytes' subtype than CD8+ cells.



Supplementary Figure 2: Bar charts showing the difference between expression of various immune cells and checkpoint proteins in pure DCIS and DCIS coexisting with invasive disease; a) stromal CD20+, b) stromal CD3+, c) stromal CD8+, d) stromal CD4+, e) stromal FOXP3+, f) stromal PD1+, g) stromal PDL1+, h) intratumoural CD20+, i) intratumoural CD3+, j) intratumoural CD8+, k) intratumoural CD4+, l) intratumoural FOXP3+ and m) intratumoural PD1+ lymphocytes. Error bars represent +2 standard deviation.



Supplementary Figure 3: Violin plots showing the association between various clinicopathological parameters, TP53 mutation with (A-E) the fraction of genomic altered (FGA) and with (F-I) number of telomeric allelic imbalances (NTAI). The central red dots represent the mean, boxes represent the interquartile range, central line represent the median and whiskers shows the 95% confidence interval.



Supplementary Figure 4: Kaplan Meier Curves showing; A) Higher TILs density is associated with shorter local recurrence free interval in high risk DCIS defined as high grade DCIS with size >1.5cm. B) High stromal PDL1 expression is associated with shorter local recurrence free interval (LRFI) in high risk DCIS.