The SRC-family kinase LYN inhibits tumour growth and activates inflammatory signalling pathways in an *in vivo BRCA1* loss-of-function-associated mammary tumour model

Giusy Tornillo[1], Lauren Warrington[1], Howard Kendrick[1], Trevor Hay[1], Sam Beck[2] and Matthew J. Smalley[1,3]

[1]The European Cancer Stem Cell Research Institute, School of Biosciences, Cardiff University, UK

[2]Independent Anatomic Pathology Ltd, Bath, UK.

[3]Corresponding Author: The European Cancer Stem Cell Research Institute, Hadyn Ellis Building, Maindy Road, Cardiff CF24 4HQ, UK. E-mail: SmalleyMJ@cardiff.ac.uk

Keywords: BRCA1, LYN kinase, CRE-LoxP

Abstract

We have previously demonstrated that the SRC-family kinase LYN is expressed in mammary tumours from a BRCA1 loss-of-function-dependent mouse mammary tumour model, as well as in the cells of origin of such tumours, the mammary luminal epithelial stem/progenitor cells, and in human basal-like/hormone receptor negative breast cancers. We have also shown that suppressing LYN kinase activity in BRCA1-defective cell lines, as well as in *in vitro* cultures of *Brca1*-null mouse mammary tumours, is deleterious to their growth.

Here, we have examined the interaction between LYN kinase and BRCA1 loss-of-function in an *in vivo* mouse mammary tumour model, using conditional knockout *Brca1* and *Lyn* alleles. Comparison of tumour cohorts carrying *Brca1* conditional alleles on a *p53* heterozygous background and also either wild type, heterozygous or homozygous for the conditional *Lyn* allele, demonstrated that homozygous conditional *Lyn* cohorts had a small, but significant, decrease in median survival time compared to the other genotypes. However, this was confounded by the development of multiple tumours in some animals, and continued expression of LYN in some $Lyn^{fl/fl}$ tumours (as well as low levels of LYN in some $Lyn^{wt/wt}$ tumours). We therefore carried out a transcriptomic analysis of $Lyn^{wt/wt}$, $Lyn^{fl/wt}$ and $Lyn^{fl/fl}$ tumours, identifying differentially expressed genes between tumour groups categorised on the basis of supervised (comparison between tumours with high and low *Lyn* levels as measured directly by RNAseq) and unsupervised (Principal Component Analysis-based) methods. There was substantial overlap between tumours determined by the supervised and unsupervised approaches, suggesting levels of *Lyn* expression were, at least in part, responsible for biological differences between the tumours. The analysis also showed that '*Lyn* high' tumours had a significantly slower doubling time than '*Lyn* low' tumours and were enriched in inflammatory pathways and canonical NfkB signalling. Taken together with our previous findings, these results suggest a context-dependency to LYN signalling in mammary epithelial tumour cells, similar to its biology in B-cells.

Introduction

The protein product of the *BRCA1* gene is well-established as a tumour suppressor, principally of breast and ovarian cancer (Fu et al., 2022). Women who inherit one functional and one mutated copy are at an approximately 70% lifetime risk of developing breast cancer and 40% lifetime risk of ovarian cancer (Kuchenbaecker et al., 2017). Tumour formation is associated with loss-of-heterozygosity events which delete the functional copy (Mahdavi et al., 2019) and is accelerated by concomitant inactivation of the TP53 tumour suppressor protein (Kim et al., 2016; Liu et al., 2007; Zhu et al., 2022). However, carriers of *BRCA1* mutations may also show functional haploinsufficiency which increases the risk of overt neoplasia (Lim et al., 2009).

The most well-characterised role of BRCA1 is as a key component of error-free, homologous recombination-dependent repair of double-stranded DNA damage (Foo and Xia, 2022; O'Donovan and Livingston, 2010). It is the defect in this process in BRCA1 loss-of-function associated breast and ovarian cancer that is exploited by the use of PARP inhibitors for therapy (Mateo et al., 2019). However, BRCA1 has a number of additional functions not directly affecting DNA damage repair (although they may influence the process indirectly). These include acting as an E3 ubiquitin ligase, regulating transcription and control of centrosomal replication (Densham and Morris, 2017; Nolan et al., 2017; Yoshida and Miki, 2004).

We have previously demonstrated that the cells of origin of BRCA1-associated mammary tumours, the mammary luminal epithelial estrogen receptor negative stem/progenitor population, express the c-KIT receptor tyrosine kinase at high levels, as well as its downstream pathway member, the SRC-family kinase LYN (Regan et al., 2012; Tornillo et al., 2018). We and others have shown that LYN is overexpressed in human Triple Negative Breast Cancer (TNBC; the breast cancer subtype most strongly associated with BRCA1 loss) (Choi et al., 2010; Croucher et al., 2013; Hochgrafe et al., 2010; Molyneux et al., 2010). We have also shown that LYN is expressed at high levels in mammary tumours from *Brca1* conditional knockout mice (Molyneux et al., 2010). We have directly demonstrated using human breast cancer cell lines *in vitro* and as xenografts, primary cultures from *Brca1*-deleted mouse mammary tumours in two- and three-dimensions, and cultures from human *BRCA1*-null patient-derived breast cancer xenografts, that BRCA1 loss results in LYN activation and activation of downstream pathways, including AKT (Tornillo et al., 2018). BRCA1 loss-dependent activation of LYN resulted in a growth and survival advantage to mammary epithelial cells. We therefore suggested that LYN is an oncogene in the context of BRCA1 loss, and a potential therapeutic target in BRCA1 loss-of-function breast and ovarian

cancers. However, this has not yet been tested in a gold-standard *in vivo* knockout mouse model. Therefore, we obtained a conditional knockout LYN allele and crossed it to our established *BlgCre Brca1*$^{f/f}$ *p53*$^{+/-}$ mouse mammary tumour model. Surprisingly, we found that in this system, *BlgCre Brca1*$^{f/f}$ *p53*$^{+/-}$ mice carrying two conditional knockout *Lyn* alleles had a shorter overall survival than heterozygous or *Lyn* wild type *BlgCre Brca1*$^{f/f}$ *p53*$^{+/-}$ mice, while tumours with low levels of *Lyn* grew faster than tumours with high levels of *Lyn*.

Our study examined *Lyn* transcript and total LYN protein levels. We were not able to directly measure LYN kinase activity in tumours. Nevertheless, in contrast to our initial hypothesis that LYN would act as an oncogene promoting tumour development, the study suggests that LYN acts as a tumour suppressor in this *in vivo* model system. As with B-cells, this argues for a context dependency to LYN function in the mammary epithelium and that tight regulation of its activity is required to maintain 'just right' levels of activity. It also suggests that as a therapeutic target in breast cancer, LYN kinase is likely to present difficulties.

Materials and Methods

See **Supplementary Table 1** for a full list of all primers, antibodies and other reagents. Raw scanned western blots are provided as a **Supplemental Data File**.

Establishment of genetically modified mouse lines

This study was approved by the Cardiff University Animal Welfare and Ethical Review Body and carried out under the authority of appropriate Home Office Personal and Project Licences and with reference to ARRIVE guidelines (Percie du Sert et al., 2020). In particular, animals were monitored regularly and predefined humane endpoints strictly adhered to. Numbers of animals required in each cohort were based on previous experience of requirements for a sufficiently powered tumour cohort study (typically 15 – 20 animals per cohort depending on effect size). Randomisation was not appropriate as animals had to be assigned to cohorts according to their genotype.

The full breeding scheme is illustrated in **Supplementary Figure 1**. Mice carrying the conditional *Brca1* allele on the *p53* heterozygote background as well as a CRE recombinase under the control of the β-lactoglobulin mammary specific promoter (*BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ mice) have been previously described (McCarthy et al., 2007; Molyneux et al., 2010). Mice carrying a tamoxifen-activated CRE ubiquitously expressed from the *Rosa26* locus (*R26C*) (Hameyer et al., 2007) were obtained from Prof. Karen Blyth, CRUK Beatson Centre, Glasgow. Mice carrying the *Lyn floxed exon 4* conditional allele (*Lyn*^{tm1c}) were obtained from the Mary Lyon Centre, MRC Harwell (full nomenclature C57BL/6N-Lyn$^{tm1c(EUCOMM)Hmgu}$/H, derived from an ES cell clone HEPD0704_6_B11).

The full details of all animals used in the study and of all histological samples are provided in **Supplementary Tables 2 and 3**.

Mouse mammary epithelial cell harvest and culture

Mammary epithelial organoids were prepared from fourth mammary fat pads of 10-12 week-old virgin female mice as described (Smalley, 2010). Intramammary lymph nodes were removed prior to tissue collection. Fat pads were finely minced on a McIlwain Tissue Chopper and then digested for 1 hr at 37°C in 3 mg/ml collagenase A/1.5 mg/ml trypsin in serum- and phenol red-free L15 medium (ThermoFisher Scientific) with gentle rotation. Tissue fragments ('organoids') released were incubated for 5 min in Red Blood Cell Lysis buffer (Sigma), washed and then plated for 1 hr at 37°C in DMEM/10%FBS (ThermoFisher) for depletion of fibroblasts by differential attachment.

For three-dimensional (3D) cultures, organoids were incubated with 0.05% trypsin/EDTA for 2 min at 37°C prior to plating onto Growth Factor Reduced phenol red-free Matrigel (Corning) in complete growth medium (DMEM:F12 with 10% Charcoal Stripped FBS, 5 ug/ml insulin, 10 ng/ml cholera toxin and 10 ng/ml epidermal growth factor (Merk-Sigma Aldrich). 4-Hydroxytamoxifen (4OHT) (Merk-Sigma Aldrich) was added at a final concentration 100 nM for 10-12 hours to induce the recombination of the *Lyn*^{fl} allele.

Isolation of Primary Tumour Cells

Primary tumour epithelial cells were obtained using the gentle MACS Dissociator and Mouse Tumor dissociation kit (Miltenyi Biotec, Bisley, Surrey, UK) following the protocol recommended for 'Dissociation of Tough Tumors'. To ensure efficient dissociation, volumes of Enzyme D, Enzyme R and Enzyme A were scaled up according to the size of the tumour piece (100 uL, 50 uL and 12.5 uL respectively per each 0.5 cm^3). The optional red blood cell lysis step was included in the procedure. Resulting cells were plated in complete growth medium in two-dimensional (2D) adherent conditions. Cells at passage 0 were used for all the experiments in this study.

Tumour measurements and doubling times

Tumour width (W) and length (L) were measured using a caliper twice a week by the same person each time to eliminate inter-operator variability. Volume was calculated using the formula ($L \times W^2/2$).

IHC and FFPE sample processing

Mice were euthanised by an approved method when previously established humane endpoints were reached. A full necropsy was performed and any tumour tissue was fixed in 10% neutral buffered formalin for 24h at 4°C before being processed into paraffin blocks according to standard procedures. When a tumour was of sufficient size, a piece (distant from any obvious necrosis) was also snap frozen on dry ice at time of dissection and then stored at -80°C for later RNA/protein extraction. In some cases, pieces of tumour were kept in L15 medium on ice for later isolation and culture of primary tumour cells. Visceral organs (liver, kidneys, spleen, lungs and in some cases heart and stomach, if obvious pathology present) were also fixed in neutral buffered formalin for 24 h and processed into paraffin blocks.

Tissue sections (5 µm) were either stained using Haematoxylin & Eosin (H&E) for histological analysis or used for immunohistochemical staining. For the latter, freshly cut sections were dewaxed and re-hydrated. Sections underwent antigen retrieval in citrate

buffer, pH 6.0 in a pressure cooker for 15 min before incubation with a 3% hydrogen peroxide solution for 20 min and then blocking in 10% goat serum/0.1% Tween-20/TBS for 1 hour. Incubation with primary antibodies **(Supplementary Table 1)** was performed overnight at 4°C. Detection was carried out using the ImmPRESS kit (Vector Labs, Peterborough, UK). Sections were counterstained with haematoxylin and mounted. Images were acquired using a VS200 slide scanner with a 20x objective and visualised using OlyVIA slide viewer software.

Histopathological analysis

Mammary tumour phenotyping was carried out by MJS (who has over ten years' experience of using the four-histotype classification system for mouse mammary tumours) using our previously established criteria based primarily on morphology of H&E stained sections and immunohistochemical staining for ΔNp63. In brief, assessment of metaplasia (either spindle cell or squamous) and the extent of any ΔNp63 staining allows mouse mammary epithelial tumours to be classified as adenosquamous tumours (ASQC; extensive squamous metaplasia and abundant ΔNp63 staining), adenomyoepitheliomas (AME; abundant ΔNp63 staining in a distinct pseudo-basal pattern bordering ΔNp63-negative cells, but little or no metaplasia), metaplastic spindle cell carcinomas (MSCC; extensive or near total spindle cell metaplasia with infrequent nests of epithelial tumour cells; no ΔNp63 staining), adenocarcinomas of no special type (ACNST; little or no metaplasia and little ΔNp63 staining). Histology of other organs was reviewed by MJS with support and advice from SB.

Scoring of Ki67 IHC

For Ki67 IHC quantification, images of five different regions (in one case, six regions) from each section were captured using the OlyVIA software at 10x magnification. Regions were chosen to include areas with the highest level of Ki67 staining for that section, so that the final score represented the highest potential for proliferation, and therefore the most aggressive behaviour, of that tumour. The percentage of positive tumour cells in each image was determined automatically using Cognition Master Professional Ki67 Quantifier (Medline Scientific Limited, Chalgrove, Oxfordshire, UK). Values returned by the program were 'sense-checked' against each image; any obvious errors (e.g. 21-34-03 Field 5; **Supplementary Table 4**) excluded from further analysis.

Scoring of LYN IHC staining by modified histoscore

LYN IHC staining using a rabbit polyclonal antibody (Thermofisher) was quantified by a modified histoscore approach considering strength of staining and area stained. If staining

was visible at x0.2 magnification on the OlyVIA slide viewer images, staining was scored as strength ‘3’. If staining was not visible at x0.2 but was visible at x2, it was scored as strength ‘2’. If staining was not visible at x2 but was visible at x20, it was scored as strength ‘1’. If no staining was visible at x20 it was scored as ‘0’. For the area of tumour stained, scoring was determined as follows: 0, no staining; 1, <10% of tumour cells positive; 2, 10 – 50% of tumour cells positive; 3, >50% of tumour cells positive. These divisions were chosen as easily assessable by eye, without the need for exact counting. The two scores were then multiplied together to give a final value. Note that the antibody used measures total LYN protein, not active protein. Antibodies specific to the phosphorylation site on LYN which indicates activation are not currently available.

Protein isolation and western blot analysis

2D cultured cells were lysed in Laemmli buffer. 3D cultured primary mouse mammary cells were released from Matrigel using the BD cell recovery solution prior to lysis. Protein extracts were separated by SDS-PAGE on 4–15% gradient Mini-PROTEAN TGX Precast Protein Gels (Bio-Rad, Watford, Hertfordshire, UK), transferred to PVDF membranes (IPVH00010, Merck Millipore, Hertfordshire, UK) and immunoblotted with anti-LYN antibodies. GAPDH was used as loading control. Resulting immunocomplexes were detected by HRP-conjugated anti-mouse IgG or anti-rabbit IgG secondary antibodies and enhanced chemiluminescent (ECL) reagents (WBLUF0100, Merck Millipore).

RNA isolation and gene expression analysis

Total RNA was isolated from tumour tissue or 2D-cultured cells using the RNeasy Minikit (QIAGEN) according to the manufacturer's protocol. Trizol (ThermoFisher Scientific) was used for RNA extraction from 3D-cultured primary mammary organoids. Up to 1 ug of RNA was converted into cDNA using the Quantitect Reverse Transcription kit (QIAGEN) following the manufacturer's instructions. Gene expression analysis was carried out using TaqMan Master Mix and Taqman gene expression assays (ThermoFisher Scientific) (**Supplementary Table 1**). Data analysis was carried out using the QuantStudio7 Software. Relative expression levels of target genes were calculated using the ΔΔCt method as described previously (Kendrick et al., 2008).

RNA Sequencing and Analysis

Samples for RNAseq analysis underwent an on-column DNase I digestion step for genomic DNA removal prior to further processing. Total RNA quality and quantity was assessed using Agilent 4200 TapeStation and hsRNA or RNA ScreenTapes (Agilent Technologies). mRNA was isolated from 50ng of total RNA (RIN value >7) using the NEBNext® Poly(A) mRNA

magnetic isolation module, (New England BioLabs, NEB, #E7490) and the sequencing libraries were prepared using the NEB® Ultra™ II Directional RNA Library Prep Kit for Illumina® (NEB, #E7760). The sequencing libraries were prepared following Chapter 1 of the NEB® Ultra™ II Directional RNA Library Prep Kit for Illumina® (New England BioLabs, NEB) protocol. The steps included mRNA isolation, fragmentation and priming, first strand cDNA synthesis, second strand cDNA synthesis, adenylation of 3' ends, adapter ligation (1:80 dilution) and PCR amplification (14-cycles). Libraries were validated using the Agilent 4200 TapeStation and hsD1000 ScreenTapes (Agilent Technologies) to ascertain the insert size, and the Qubit® (Thermo Fisher Scientific) was used to perform fluorometric quantitation. The manufacturer's instructions were followed except for the replacement of SPRIselect Beads or NEBNext Sample Purification Beads by AMPure XP beads (Beckman Coulter®) in purification steps. The validated libraries were normalized to 4nM, pooled together and the pool sequenced on an S1 (200 cycle) flow cell using a 2x100bp PE dual index format on the Illumina® NovaSeq6000 sequencing system according to the manufacturer's instructions. Samples were sequenced to a read depth of at least 35 Million prior to quality trimming with fastp (Chen et al., 2018). Quality trimmed reads were mapped to GRCm38 using STAR (v2.5.1b) (Dobin and Gingeras, 2015) with read multimapping filter set to 1 and gtf Gencode GRCm38 vM17. Exon and gene counts were calculated with featureCounts (v1.5.1) (Liao et al., 2014). Differential gene expression was calculated using SARtools using the DESeq2 package (Varet et al., 2016). RNAseq data have been deposited at Gene Expression Omnibus (GEO) with the accession number GSE222075.

Gene Set Enrichment Analysis (GSEA)

For GSEA, significantly differentially expressed genes from the different tumour groups (adjusted p value <0.05; log2 fold change <0.5 or >2.0) were uploaded to g:Profiler (https://biit.cs.ut.ee/gprofiler/gost) and queried against the *Mus musculus* database with a 'term size limit' of 1000 but otherwise using default options (with the exception of one group, the differentially expressed genes overlapping between the 'PCA 1/2' and '*Lyn* low' tumour groups, for which the 'term size limit' was allowed default values, otherwise no results were returned). Results were downloaded as a CSV file (default options). Gene Ontology Bioprocess grouping into functional categories was carried out manually.

CIBERSORTx analysis (https://cibersortx.stanford.edu/) (Newman et al., 2019) was carried out using the LM22 signature matrix file for 22 immune cell types and was run in 'absolute mode' so that relative differences between the proportions of immune cell types would be maintained.

Statistics

All statistical analysis was carried out in GraphPad Prism. Survival curves were analysed by Log Rank test. For all other experiments, normally distributed data were analysed by ANOVA and/or t-tests where appropriate. Non-parametric data were analysed by Kruskel-Wallis and/or Mann-Whitney where appropriate. A p value of <0.05 was taken as significant. For analysis of difference in distribution of categorical variables, Chi2 test (two groups) or Chi2 test for trend (more than two groups) was used.

Results

The *Lyn*$^{fl(ex4)}$ allele is efficiently recombined *ex vivo* resulting in loss of *Lyn* expression

We had previously assessed the relationship between LYN and (*BRCA1*-associated) mammary tumourigenesis using, among other approaches, shRNA knockdown with two independent shRNA sequences. We had controlled for off-target effects where appropriate by rescuing any observed phenotype by re-expressing a *Lyn* cDNA which was resistant to the effects of the knockdown (Tornillo et al., 2018). Our findings in these studies suggested to us that functional LYN kinase was required for survival of BRCA1-null mammary epithelial cells. However, the role of LYN in *BRCA1*-associated mammary tumourigenesis had not been investigated using the gold standard of conditional *in vivo* knockouts. Therefore, we obtained a mouse line carrying a conditional (*floxed* exon 4) *Lyn* allele (the *Lyn*^{tm1c} allele; hereafter *Lyn*^{fl}) from the Mary Lyon Centre, Harwell **(Figure 1A)**. Full details of these mice and the breeding strategies in which they were used to generate experimental cohorts are provided in the **Materials and Methods** and **Supplementary Figure 1**.

We first used *Lyn*^{fl} mice to establish a cohort of animals in which expression of a tamoxifen-inducible CRE recombinase was driven from the ubiquitously active *Rosa26* promoter (hereafter *R26C*). This cohort also included a conditional *floxed Brca1* allele (*Brca1*$^{fl(ex22-24)}$; hereafter *Brca1*^{fl}) (McCarthy et al., 2007; Molyneux et al., 2010). To test the recombination of the *Lyn*^{fl} allele, mammary epithelial cells were harvested from *R26C Brca1*$^{fl/wt}$ *Lyn*$^{wt/wt}$, *R26C Brca1*$^{fl/wt}$ *Lyn*$^{fl/wt}$ or *R26C Brca1*$^{fl/wt}$ *Lyn*$^{fl/fl}$ mice and cultured as three-dimensional organoids according to our previous protocols (Tornillo et al., 2018). After one day, cultures were treated with 100 nM 4OHT or vehicle. After overnight incubation, cultures were washed to remove 4OHT and then cultured for a further 72 hours prior to lysis for isolation of either RNA or protein. qrtPCR analysis (**Figure 1B**) demonstrated no difference in expression of *Lyn exon 4* between vehicle-treated *R26C Brca1*$^{fl/wt}$ *Lyn*$^{wt/wt}$, *R26C Brca1*$^{fl/wt}$ *Lyn*$^{fl/wt}$ or *R26C Brca1*$^{fl/wt}$ *Lyn*$^{fl/fl}$ cells, or between vehicle and 4OHT-treated cells from *R26C Brca1*$^{fl/wt}$ *Lyn*$^{wt/wt}$ mice. However, there was a significant reduction in *Lyn exon 4* levels in 4OHT-treated compared to vehicle-treated *R26C Brca1*$^{fl/wt}$ *Lyn*$^{fl/wt}$ cells, while *Lyn exon 4* was undetectable in 4OHT-treated cells from *R26C Brca1*$^{fl/wt}$ *Lyn*$^{fl/fl}$ mice. Western blot analysis of protein extracts from these cultures (**Figure 1C**) confirmed these results.

To provide further evidence that deletion of *Lyn exon 4* results in a complete loss of LYN protein rather than, for example, generating a truncated protein which may have dominant negative effects, protein lysates from *R26C Brca1*$^{fl/wt}$ *Lyn*$^{wt/wt}$ and *R26C Brca1*$^{fl/wt}$ *Lyn*$^{fl/fl}$ cells, treated with either vehicle or 4OHT, were analysed by western blot using three different anti-LYN antibodies, one mouse monoclonal (Abcam ab1890) and two rabbit polyclonal (Abcam

ab32398 and Thermofisher PA5-81925) in parallel (**Figure 1D**). In all other respects, the analysis was run identically, with identical amounts of protein loaded in all lanes. All three antibodies showed a substantial reduction in the amount of LYN protein detected in 4OHT-treated *R26C Brca1*$^{fl/wt}$ *Lyn*$^{fl/fl}$ cells, compared to the other samples, and indeed in the mouse monoclonal and Abcam polyclonal samples LYN was undetectable. The mouse monoclonal antibody generated a number of non-specific bands only observed with the other reagents in very over-exposed blots (see **Supplemental Data File**). The Abcam polyclonal gave only a faint single band at the expected size which disappeared in 4OHT-treatment of *Lyn*^{fl} cells. The Thermofisher rabbit polyclonal gave a strong signal of the expected size which was substantially reduced in the 4OHT-treated *R26C Brca1*$^{fl/wt}$ *Lyn*$^{fl/fl}$ sample, although a faint band was still detectable. The lower band visible with the Thermofisher polyclonal was not an artefact associated with the *Lyn*^{fl} allele, as it was visible in extracts from both *Lyn*^{wt} and *Lyn*^{fl} cells. It was specific to *Lyn*, as it disappeared upon 4OHT-treatment of *Lyn*^{fl} cells. It was not, however, the LYNB isoform (Tornillo et al., 2018), as the two LYN isoforms, LYNA and LYNB, are not fully resolved in the gradient gels used here (Tornillo et al., 2018). We suggest that the lower band is an endogenous product of caspase cleavage of LYN; a ΔNLYN variant has been previously described (Marchetti et al., 2009). Therefore, the *Lyn*^{fl} allele is recombined by CRE recombinase and as a result is unable to generate protein.

A *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/fl}$ mouse cohort has reduced overall survival but few other differences compared to cohorts with wild type *Lyn* alleles

Next, three cohorts of mice in which CRE expression was driven by the *Beta-lactoglobulin* promoter (*BlgCre*) were established. All three were homozygous for *floxed Brca1* alleles and also germline heterozygous for *p53* (*BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$), replicating the lines we have previously used (McCarthy et al., 2007; Molyneux et al., 2010). One cohort was wild type for *Lyn* (*BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{wt/wt}$ n=21), in the second animals were heterozygous for the conditional *Lyn* allele (*BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/wt}$ n=21), in the third they were homozygous for the conditional *Lyn* allele (*BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/fl}$ n=21). Mice were aged until defined humane endpoints were reached, at which point animals were euthanised and underwent a full necropsy. Where mice developed mammary tumours, these were regularly measured to determine tumour doubling times prior to the point at which euthanasia was necessary. Full details of all cohort animals and their pathology is provided in **Supplementary Tables 2 and 3**.

We hypothesised that as LYN kinase activity was required for survival of cells that had lost BRCA1 activity (Tornillo et al., 2018), introducing the conditional *Lyn*$^{fl(ex4)}$ allele into the *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ background would result in a significant increase in overall survival

(i.e. age at which mice had to be euthanised for any reason) and also in mammary tumour-specific survival (i.e. the age at which mice had to be euthanised specifically as a result of the size of a mammary tumour). In contrast to our hypothesis, however, overall survival for *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/fl}$ mice was slightly, but significantly, shorter than for *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/wt}$ mice (median survival 350 days vs 365 days, respectively), although not significantly different to *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{wt/wt}$ mice (median survival 366 days) (**Figure 2A**). There was no significant difference in mammary tumour-specific survival between the cohorts (**Figure 2B**).

The majority of mice in all cohorts were euthanised because of the growth of the mammary tumours (**Figure 2C**). Other reasons for euthanasia included scratches, vestibular syndrome or poor body condition score, with a number of examples of non-mammary neoplasia found upon necropsy, as well as reactive hyperplasia of the spleen. In seven cases, neoplastic epithelial deposits were observed in the lungs of animals carrying mammary tumours (**Supplementary Figure 2A,B**); in one of these cases, both the lung deposit and the primary tumour had a squamous histology, consistent with metastatic spread of the primary tumour (**Supplementary Figure 2B**). Other neoplastic lesions included osteosarcoma (two cases) and haemangiosarcoma (one case) (**Supplementary Figure 2C-F**). Histological analysis of enlarged spleens suggested this was largely reactive, a response to either wounds, infection or necrosis of a mammary tumour, but in some cases the histology was consistent with histiocytic sarcoma (two cases; not shown) or lymphoma (clonal analysis of B- and/or T-cell receptor rearrangement to confirm lymphoma was not carried out as this was not the primary focus of this study). Florid extra-medullary haematopoiesis was also noted in some cases. There was no visible difference in LYN staining in the white pulp of the spleen in *Lyn*$^{fl/fl}$ animals compared to others, with an expected decrease in LYN staining in proliferating B-cell germinal centres (**Supplementary Figure 3**).

There were no significant differences in histotype of mammary tumours (the majority of which were adenocarcinomas of no special type) between the cohorts (**Supplementary Figure 4 and 5A**). Many animals developed more than one tumour, but there was no significant difference between the cohorts in terms of numbers of tumours developed by each animal (**Supplementary Figure 5B**). The growth of every mammary tumour that was palpable while an animal was alive was measured daily until a humane endpoint was reached and that animal needed to be euthanised. This enabled doubling times for every tumour with three or more measurements to be established. There was no significant difference in doubling times of mammary tumours between the cohorts (**Figure 2D**). This was confirmed by Ki67 staining of sections from mammary tumours across the cohorts

(considering only adenocarcinomas to eliminate different tumour histotypes as a potential confounding factor) (**Figure 2E; Supplementary Table 4**). Within each cohort there was, however, considerable heterogeneity in tumour doubling times and when animals which developed more than one tumour were considered individually, it was apparent that even in a single animal, doubling times of tumours could vary widely (**Figure 2F**).

Cohort genotype only partly predicts LYN expression in tumours

Variation in behaviour of tumours within mice in a cohort, and indeed in multiple tumours in a single mouse, could result from partial *floxed* allele recombination in *Lyn*$^{fl/wt}$ or *Lyn*$^{fl/fl}$ mice, or from suppression of LYN expression by other mechanisms in *Lyn*$^{wt/wt}$ mice. If such variation existed, it would confound any analysis of the role of LYN in *Brca1*-dependent mammary tumourigenesis based solely on genotypes of the cohort.

Therefore, to directly assess LYN expression levels in tumours from the three cohorts, thirteen *Lyn*$^{wt/wt}$ tumours, twenty-one *Lyn*$^{fl/wt}$ tumours and twenty *Lyn*$^{fl/fl}$ tumours, all adenocarcinomas (no special type) were stained for LYN protein (**Figure 3**). LYN staining was assessed both qualitatively and semi-quantitatively using a histoscore approach based on strength of staining and the area of the tumour stained (see **Materials and Methods**, **Supplementary Figure 5C, Supplementary Table 4**). Scoring was carried out blinded to genotype; once scored, tumours were unblinded and analysed.

Staining patterns fell into three types: sheets and nests of epithelial-like tumour cells showing membrane staining (**Figure 3A**); single cells scattered throughout the tumour which were typcially cells with pseudopodia and an appearance suggest a motile phenotype (**Figure 3B**, white arrowheads); cells with the appearance of tumour-associated fibroblasts (**Figure 3B,** black arrows).

Histoscore quantitation (a combined measure of both the proportion of the tumour stained for the marker and of the intensity of staining) of LYN staining was carried out for the neoplastic epithelial-like tumour cells. LYN staining of these varied significantly with tumour genotype (Kruskal-Wallis test, P=0.0063) (**Figure 3C**). There was a significant reduction in LYN staining in the cells of *Lyn*$^{fl/fl}$ tumours compared to *Lyn*$^{wt/wt}$ tumours (Mann-Whitney test, P=0.002) and *Lyn*$^{fl/wt}$ tumours (Mann-Whitney, P=0.0344). However, some *Lyn*$^{fl/fl}$ tumours clearly retained strong LYN staining while some *Lyn*$^{wt/wt}$ tumours showed very little or no staining. LYN staining in *Lyn*$^{fl/wt}$ tumours was reduced compared to *Lyn*$^{wt/wt}$ but the difference was not statistically significant.

While these results showed that, overall, there was a correlation between LYN staining of a tumour and the genotype of the animal the tumour came from, they also highlighted the variability in staining between tumours of the same genotype. This suggested that genotype could not be fully relied upon to predict LYN expression in any one individual tumour. To understand more objectively the relationship between tumour genotype and LYN expression in tumour cells, in the absence of cells from the tumour microenvironment which may also express LYN, we isolated live cells from 17 tumours (six *Lyn*$^{wt/wt}$, five *Lyn*$^{fl/wt}$ and six *Lyn*$^{fl/fl}$) and cultured them in conditions optimised for epithelial tumour cell primary culture before harvesting DNA, RNA and protein for analysis.

Semi-quantitative PCR analysis of the *Lyn*^{fl} and *Lyn*^{wt} alleles demonstrated that the abundance of the *Lyn*^{wt} allele in primary cultures of *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/wt}$ tumours was approximately half that in cultures of *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{wt/wt}$ tumours, as expected (**Figure 3D,E**). However, there was also no significant difference overall in the abundance of the *Lyn*^{fl} allele between *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/wt}$ and *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/fl}$ tumour cells (**Figure 3D,F**), and it is clear that while in some tumours the *Lyn*^{fl} allele had recombined effectively, in others it remained intact (**Figure 3D**). Assessment of *Lyn* expression by probe-based real time rtPCR using a probe targeting exon 4, demonstrated that there was a significant reduction in *Lyn* expression in *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/fl}$ tumour cells compared to wild-type cells, but again in some individual tumours *Lyn* exon 4 expression was comparable to that seen in *Lyn*$^{wt/wt}$ tumours (**Figure 3G**). *Brca1* expression was, as expected very low in tumours from all three lines relative to normal mammary epithelial cells (**Figure 3H**). Finally, assessment of LYN protein levels in these cells by western blot showed that *Lyn*$^{fl/wt}$ tumour cells had significantly less protein than *Lyn*$^{wt/wt}$ tumour cells, but again while some *Lyn*$^{fl/fl}$ tumour cells had low LYN protein levels, others had levels of LYN comparable to the wild type cultures (**Figure 3I**).

Transcriptional analysis of tumours demonstrates tumours with high *Lyn* expression have a slower doubling time

Analysis of LYN protein levels in tumours, either by staining of tumour sections or analysis of *in vitro* cultures of tumour cells, had demonstrated that while overall there was a correlation between cohort genotype and LYN expression, there were a number of cases in which tumours from a *Lyn* wildtype mouse had very low or undetectable levels of LYN, while tumours from mice homozygous for the *Lyn* flox allele could actually have high levels of LYN expression. This may have confounded the analysis, in particular, of tumour doubling times and their relationship to *Lyn* expression. An alternative approach was needed to assess

differences in the biology of the tumours from the cohorts we had established, and to determine whether or not such differences were correlated with *Lyn* expression.

Therefore, we carried out an RNAseq analysis of tumour pieces from 39 tumours – 12 from *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{wt/wt}$ mice, 14 from *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/wt}$ mice and 13 from *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/fl}$ mice. All were adenocarcinomas of no special type, to eliminate histological variation as a confounding factor. The sample details are provided in **Supplementary Table 5**. The data were analysed in three ways, two of which were 'supervised' and one which was 'unsupervised'. First, Differentially Expressed Genes (DEGs) significantly (<0.05 adjusted P-value; ≤0.05 or ≥2.0 log2 fold change) differentially expressed between tumours from *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{wt/wt}$ and *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/fl}$ mice were identified (supervised analysis on the basis of genotype). Second, the normalised expression values for *Lyn* were used to rank the whole 39-sample tumour set from the strongest *Lyn* expressing to the weakest *Lyn* expressing tumour. Then, the 13 tumours most strongly expressing *Lyn* ('*Lyn* high' group) were compared to the 13 tumours with the weakest *Lyn* expression ('*Lyn* low' group) to identify DEGs (supervised analysis on the basis of *Lyn* expression). Finally, a Principal Component Analysis (PCA) analysis was carried out on the complete normalised data set of 39 tumours to identify any groups of tumours which could be distinguished from each other on the basis of transcriptional profiles in an unbiased manner. Significant DEGs were identified between the PCA groups (unsupervised analysis). The raw and normalised data for these comparisons are provided in **Supplementary Tables 6 – 8**. Significant DEGs are provided in **Supplementary Table 9**, Gene Set Enrichment Analysis (GSEA) using g:Profiler in **Supplementary Table 10** and a summary of enriched Gene Ontogeny Bioprocess and KEGG pathways in **Supplementary Table 11**.

With the *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{wt/wt}$ and *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/fl}$ comparison, only 93 significant DEGs were identified, 12 upregulated in *Lyn*$^{fl/fl}$ tumours relative to *Lyn*$^{wt/wt}$ tumours and 81 downregulated in *Lyn*$^{fl/fl}$ tumours relative to *Lyn*$^{wt/wt}$ tumours (**Figure 4A**; **Supplementary Table 9**). This emphasised our previous findings that animal genotype was not necessarily a good surrogate for *Lyn* expression or differences in tumour biology. Indeed, there was no difference in *Lyn* expression, as defined by normalised RNAseq *Lyn* counts, between the cohorts (**Figure 4B**).

Next, we examined *Lyn* expression in all tumours as determined by the RNAseq data (ignoring genotypes) and compared this to LYN staining (**Supplementary Table 4**). Normalised *Lyn* counts were elevated in tumours with a LYN Histoscore of 1/2/3 compared

to tumours with a score of 0, and elevated further in 4/6/9 LYN Histoscore tumours compared to 1/2/3 scored tumours (P<0.05; ANOVA; **Figure 4C**). Consistent with this, when the tumours defined by normalised *Lyn* counts in the RNAseq as '*Lyn* high' (top 13 most strongly *Lyn* expression tumours) and '*Lyn* low' tumours (13 tumours with the weakest *Lyn* expression) (**Supplementary Table 5**) were compared, the Histoscore of the *Lyn* high tumours was significantly greater than that of the *Lyn* low tumours (P<0.01; Mann-Whitney; **Figure 4D**). There were, however, outliers showing that in some tumours there was not a direct correlation between *Lyn* expression by RNAseq and LYN IHC staining. Nevertheless, taking the tumour set as a whole, *Lyn* expression (by RNAseq) reflected levels of LYN protein in the neoplastic epithelial tumour cells (by Histoscore), even though we had not purified neoplastic cells away from non-neoplastic cells within the tumours for analysis. This suggests that the main contributor to RNAseq signatures were the neoplastic tumour cells.

We next compared the *in vivo* doubling time of the tumours defined as *Lyn* high and *Lyn* low (**Figure 4E**). *Lyn* high tumours had a significantly longer doubling time (P<0.05; t-test), suggesting that in general they grew more slowly than *Lyn* low tumours and that they formed a distinct biological group.

There were 1655 significant DEGs upregulated in *Lyn* high relative to *Lyn* low tumours and 100 significant DEGs downregulated in *Lyn* high relative to *Lyn* low tumours (equivalent to 100 DEGs significantly upregulated in *Lyn* low relative to *Lyn* high tumours; **Figure 4F**; **Supplementary Table 9**). This number of DEGs, compared to the number identified when comparing by genotype, showed that categorising tumours by *Lyn* expression was better at defining sets of tumours with biologically meaningful differences than categorising tumours by the genotype of the cohort from which they were derived.

Finally, we used PCA analysis on the normalised RNAseq expression values for the whole tumour set to identify in an unbiased manner groups of tumours with similar gene expression patterns. This analysis initially suggested the tumours could be split into either four (PCA groups 1, 2, 3 and 4; **Figure 5A**) or two groups (combined groups 1/2 and 3/4). The normalised *Lyn* counts, LYN Histoscore and *in vivo* tumour doubling time were compared across either the four PCA groups individually or combined into two groups (**Figure 5** and **Supplementary Figure 5D - F**). Groups 3 and 4 had significantly elevated *Lyn* counts and a significantly higher LYN Histoscore than groups 1 and 2 (**Supplementary Figure 5D,E**); the differences were more marked when comparing groups 3 and 4 combined as a single group against groups 1/2 (**Figure 5B,C**). The combined group 3/4 had a significantly slower *in vivo* doubling time than group 1/2 (**Figure 5D**) but these differences were not significant in the

four-group analysis (**Supplementary Figure 5F**). As the group 3 and 4 tumours appeared to behave similarly to each other, and the group 1 and 2 tumours also behaved similarly, the differences between the groups were more marked in the two-group analysis, the difference in doubling time suggested a real biological difference between groups 3/4 and 1/2, and the two-group approach would enable greater numbers of tumours to be compared in each group, we concentrated on the two-group approach and identified significant DEGs from PCA group 3/4 compared to PCA group 1/2. There were 835 significant DEGs upregulated in PCA group 3/4 relative to PCA group 1/2 and 2437 significant DEGs downregulated in PCA group 3/4 relative to PCA group 1/2 (equivalent to 2437 DEGs significantly upregulated in PCA group 1/2 relative to PCA group 3/4; **Figure 5E**).

Considering that the PCA group 3/4 tumours had elevated *Lyn* expression (**Figure 5B,C**), we next assessed the overlap between the sets of significant DEGs from the *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{wt/wt}$ vs. *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/fl}$ comparison, the *Lyn* high vs. *Lyn* low tumour comparison and the PCA group 3/4 vs. PCA group 1/2 comparison, using Venny (https://bioinfogp.cnb.csic.es/tools/venny/) (**Figure 5F**). There was very little overlap between the *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{wt/wt}$ vs. *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/fl}$ DEGs and the other comparisons. However, 340 genes were identified as significantly differentially expressed in both the *Lyn* high vs. *Lyn* low and the PCA group 3/4 vs. PCA group 1/2 comparison. Of these, 282 were upregulated in both *Lyn* high and PCA group 3/4 tumours while 57 were upregulated in both *Lyn* low and PCA group 1/2 tumours. Only one gene was elevated in *Lyn* high and PCA group 1/2 tumours (**Supplementary Table 10**).

The distribution of the *Lyn* high and *Lyn* low tumours within the PCA groups was consistent with these results. Of the PCA group 3/4 tumours, six were also in the *Lyn* high tumour group while four were tumours with intermediate *Lyn* expression. There were no *Lyn* low tumours in PCA group 3/4. PCA group 1/2 included all the *Lyn* low tumours, nine intermediate *Lyn* expression tumours and seven belonging to the *Lyn* high group. Notably, the *Lyn* high tumours in PCA group 3/4 were six of the seven tumours with the highest rank for *Lyn* expression by RNAseq, the one exception being one of the *Lyn* high tumours in PCA group 1/2 which ranked second in *Lyn* expression by RNAseq but had been scored 0 for LYN expression by IHC (**Supplementary Table 5**).

***Lyn* high and PCA group 3/4 tumours are enriched in inflammatory signalling pathways while PCA group 1/2 tumours are enriched in morphogenesis and cancer-associated signalling pathways**

We next undertook Gene Set Enrichment Annotation (GSEA) of the DEGs from the *Lyn* high versus *Lyn* low and the PCA group comparisons using g:Profiler. Genes significantly differentially expressed in the two comparisons were annotated separately and then overlaps between the annotations assessed. Full details are provided in **Supplementary Table 10**. For ease of interpretation, we concentrated on understanding differentially enriched Gene Ontology Bioprocess terms (GO BP terms) and KEGG pathways. GO BP terms were grouped by functional categories to facilitate this. The GO BP and KEGG analysis is summarised in **Supplementary Table 11** and **Figure 6.**

Lyn high tumours were enriched for 48 GO BP terms and 17 KEGG pathways. *Lyn* low tumours were enriched for 3 GO BP terms but no KEGG pathways. PCA group 1/2 tumours were enriched for 477 GO BP terms and 14 KEGG pathways while PCA group 3/4 tumours were enriched for 48 terms and 17 KEGG pathways (**Supplementary Table 11**). The overlaps in GO BP and KEGG pathways between the tumour groups was striking and reflected the overlap seen in the DEGs. The list of enriched GO BP and KEGG pathways in the *Lyn* high and PCA group 3/4 tumours was identical. Seven of these GO Bioprocesses were also enriched in PCA group 1/2, however, the majority (469 out of 477) of GO BP terms and all KEGG terms enriched in PCA group 1/2 were not found in the other groups (**Figure 6A, B**).

GO BP terms were categorised into functional classes (**Supplementary Table 11**) and the proportions of enriched terms from each functional class assessed for the tumour groups (**Figure 6C**). Unsurprisingly, given that the terms enriched in the *Lyn* high and PCA group 3/4 tumours was identical, the classification of GO BP terms in these groups was identical. For both of these sets, the most numerous classification of enriched GO BP terms was inflammation and immune response (40 terms, 83.3%), followed by cell signalling (4 terms, 8.3%), cell adhesion and ECM (2 terms, 4.2%) and neurodevelopment and function (2 terms, 4.2%). For the *Lyn* low tumours, the three enriched GO BP terms were all classified as associated with neurodevelopment and function (100%). The 477 GO BP terms enriched in PCA group 1/2 could be classified into 17 different functional classes. The four largest of these (to which >10% of the 477 GO BP terms were assigned) were morphogenesis, remodelling and healing (128 terms, 26.8%), cellular homeostasis and metabolism (92 terms, 19.3%), cell signalling (53 terms, 11.1%) and neurodevelopment and function (48 terms, 10.1%). In addition to the 'cell signalling' class, six GO BP terms (1.3%) were separately categorised as being WNT-signalling associated. Also of interest were three GO BP terms enriched in PCA group 1/2 (0.6%) associated with mammary gland development.

Closer examination of the GO BP terms associated with cell signalling in the *Lyn* high/PCA group 3/4 tumours and the PCA group 1/2 tumours (**Supplementary Table 11**) suggested that ERK1/2 pathways were active in the *Lyn* high/PCA group 3/4 tumours. In the PCA group 1/2 tumours, ERK1/2 was being inhibited, while there was both positive and negative regulation of MAPK pathways. PI3K-associated pathways were also being regulated in PCA group 1/2 tumours. Activation of these pathways in this group of tumours appeared to be associated with TGFβ and VEGF signalling, while smoothened and WNT signalling were also active.

We next examined the enriched KEGG pathways. Consistent with the GO BP analysis, and the overlap of the annotations between the tumour groups, the *Lyn* high/PCA group 3/4 tumours were enriched for KEGG pathways including 'Cytokine-cytokine receptor interaction', 'NF-kappa B signaling', 'Chemokine signaling', 'TNF signaling pathway' and 'Apoptosis'. In contrast, the PCA group 1/2 tumours were enriched for KEGG pathways including 'Breast cancer', 'Wnt signaling', 'Notch signaling', 'PI3K-Akt signaling' and 'Pathways in cancer' (**Supplementary Table 11**).

No significant differences in immune cell infiltrate between tumour groups

The overlap in GO Bioprocess annotations and KEGG pathways between the *Lyn* high and PCA group 3/4 tumours suggests that high *Lyn* expression levels, and pathways associated with such high expression levels, are a fundamental driver of the biology of the PCA group 3/4 tumours. Many of these pathways have likely originally been identified in immune cells (LYN itself is best known as a regulator of B-cells) and therefore have been annotated with names associated with inflammation and immune cells. However, it cannot be assumed that these tumours have a significantly higher immune cell infiltrate that than *Lyn* low or PCA group 1/2 tumours simply from the names of the pathways, although this might be the case. Therefore, to assess whether *Lyn* high and PCA group 3/4 tumours had elevated immune cell infiltration or whether the GO BP terms/KEGG pathways with which they were enriched reflected a fundamental difference in intrinsic tumour cell biology, we used CIBERSORTx (Steen et al., 2020). This enabled us to estimate relative abundance of immune cell subsets based on RNAseq expression profiles, in the *Lyn* high compared to the *Lyn* low tumour groups, and PCA group 3/4 compared to the PCA group 1/2 tumours.

Other than lower levels of resting NK cells in the PCA group 3/4 tumours group compared to the PCA group 1/2 tumours, there were no significant differences in immune cell subsets between the tumour groups, as estimated by CIBERSORTx (**Figure 7; Supplementary Table 12**). Therefore, the enrichment in biological pathways associated with inflammation

and immune function likely reflects fundamental differences in the *Lyn* high/PCA group 3/4 tumour cells themselves rather than changes in immune cell infiltrate. The enrichment of biological functions associated with NfkB signalling (NF-kappa B signaling pathway, KEGG:04064; cytokine-mediated signaling pathway, GO:0019221; inflammatory response, GO:0006954; adaptive immune response, GO:0002250; regulation of immune response, GO:0050776) in particular suggested a role for inflammatory signalling in these tumours.

Discussion

The SRC-family kinase LYN is most highly expressed in haematopoietic cells but is also expressed in a wide variety of other tissues, including epithelial tissues (reviewed in (Brian and Freedman, 2021)). It is a downstream target of c-KIT signalling in luminal epithelial stem/progenitor cells of the mammary gland (Regan et al., 2012; Tornillo et al., 2018) and also expressed in breast cancers, particularly TNBC (Choi et al., 2010; Hochgrafe et al., 2010; Molyneux et al., 2010).

LYN kinase is best known for its role as both a positive and negative regulator of myeloid and B-cell development and differentiation (Brian and Freedman, 2021). The positive functions of LYN are context dependent and in positive signalling loss of LYN may be compensated for by other SRC-family kinases. In contrast, LYN appears to be absolutely required for negative regulation of B-cell proliferation (Brian and Freedman, 2021; Xu et al., 2005). Notably, both *Lyn* knockout mice and LYN constitutively over-expressing mice develop lethal autoimmune kidney disease, although of distinctly different pathologies (Hibbs et al., 2002; Hibbs et al., 1995).

LYN has two splice isoforms (LYN^{FL}/LYN p56/LYNA and $LYN^{\Delta 25\text{-}45}$/LYN p53/LYNB) (Brian and Freedman, 2021; Tornillo et al., 2018; Xu et al., 2005) and a ΔNLYN caspase-cleaved variant which affects NFkB signalling has also been described (Marchetti et al., 2009). These splice isoforms are still not fully understood, although it is known that LYNA regulates a signalling checkpoint in macrophages (Freedman et al., 2015) and re-expression of either LYNA or LYNB in *Lyn* knockout mice restores B-cell developmental defects but neither rescues the autoimmune phenotype on its own (Brian et al., 2022). Interestingly, the autoimmune pathology is different in mice lacking only LYNA or LYNB and, surprisingly, in LYNA-only deficient mice is only seen in females (Brian et al., 2022).

Previous studies on the mammary epithelium and breast cancer, including our own, have highlighted LYN as a positive regulator of cell growth and survival (Choi et al., 2010; Tornillo et al., 2018). We have shown LYNA specifically regulates cell invasion and migration in TNBC cell lines *in vitro* but both LYNA and LYNB enhance breast cancer cell line survival, suggesting it is an oncogene (Tornillo et al., 2018). The tyrosine residue (Y32) in the LYNA-specific N-terminal region is a target of EGFR kinase activity and, once phosphorylated, results in LYN-mediated activation of the MCM7 DNA replication licencing factor (Huang et al., 2013). LYN activity has been reported to promote EMT through Vav-Rac1-PAK1-mediated control of SNAI protein localisation and stability in multiple cancer types, including

breast cancer (Thaper et al., 2017). A role for LYN in regulation of p53 has also been described. LYN is reported to directly interact with p53 and prevent its nuclear export, suppressing MDM-2 mediated p53 degradation and enhancing p53-dependent apoptosis (Ren et al., 2002).

However, LYN has both positive and negative effects on signal pathway activation, cell proliferation and differentiation in haematopoietic cells, which are context dependent (Brian and Freedman, 2021; Xu et al., 2005), and indeed here we have shown that in an *in vivo* model, tumours with high levels of *Lyn* expression grew more slowly than tumours with low expression levels. Therefore, in this *in vivo* context *Lyn* is acting as a tumour suppressor gene. Unfortunately, in our model we were not able to distinguish between the role of LYNA and LYNB (or indeed ΔNLYN) or its activation or kinase activity. Our results support a model in which LYN activity and function in mammary tumorigenesis is also context dependent, promoting proliferation and invasion in *in vitro* systems and in xenografts, but inhibiting growth in the setting of an autochthonic tumour model. The 'contexts' in these cases may include two-dimensional or three-dimensional growth patterns, interaction between neoplastic cells, non-neoplastic cells (such as immune cells or tumour-associated fibroblasts) and the stroma, as well as the spectrum of additional mutations within the tumour. Overall, these findings suggest context dependency of activity and function may be a general property of this kinase.

Unsupervised principal component analysis of RNAseq data from the tumours from our cohorts showed that they fell into two main transcriptional profile groups, PCA group 1/2 and PCA group 3/4. Importantly, these had substantial overlap with the transcriptional profiles of a supervised division of the tumours into one group of thirteen tumours with the lowest levels of *Lyn* expression by RNAseq and a second group of thirteen tumours with the highest levels of *Lyn* expression. Gene Ontology Bioprocess and KEGG pathway analysis demonstrated pathways associated with inflammation, cytokine signalling, immune cell regulation and in particular NFkB signalling were enriched in both the *Lyn* high and the PCA group 3/4 tumours. The PCA group 1/2 tumours were enriched in processes associated with morphogenesis and development, WNT, Notch and PI3K signalling.

The PCA group 1/2 and PCA group 3/4 tumours were clearly characterised by distinct biological processes, which (from the CIBERSORTx analysis) were a result of more than simply a difference in immune cell infiltrate. The *Lyn* low tumours were a subset of the PCA group 1/2 tumours, while the *Lyn* high tumours were split between PCA group 3/4 and PCA group 1/2. Overall, however, PCA group 3/4 had significantly higher *Lyn* gene expression by

RNAseq and LYN protein by Histoscore than PCA group 1/2, and six of the seven most strongly expressing *Lyn* high tumours were in PCA group 3/4, the exception being one tumour with high *Lyn* counts by RNAseq but no LYN staining by IHC. Taken together with the findings of the GO BP and KEGG pathway analysis, this suggests that processes regulated by LYN are a major driver of the fundamental biology of PCA group 3/4 and its difference from PCA group 1/2.

The enrichment of the *Lyn* high/PCA group 3/4 tumours for, in particular, processes and pathways associated with NFkB signalling is consistent with previous studies which demonstrate that LYN is a regulator of this pathway. The caspase-cleaved ΔNLYN form (cleaved after Asp18) increased STAT-3 activation but suppressed activation of NfkB pathways in a transgenic mouse model of ΔNLYN overexpression (Marchetti et al., 2009). These mice developed a psoriasis-like inflammatory syndrome in juveniles, which was transient for less severe cases. Interestingly, this syndrome could be rescued by crossing the mice with TNFR1 knockout mice (Marchetti et al., 2009).

MEK and IKKα, which activates NFkB signalling, were shown to be downstream targets of LYN in an analysis of the kinomes of imatinib-sensitive and resistant leukaemia cells (Cooper et al., 2013). LYN is required for TLR2-mediated NFkB activation though PI3K signalling in HEK293 cells engineered to express TLR2 (Toubiana et al., 2015) while TLR4-mediated NFkB activation in macrophages as a response to LPS treatment is associated with LYN activation and subsequent activation of downstream MAP pathways (Li et al., 2022). Furthermore, *Lyn*$^{-/-}$ mast cells show less MAPK and IkB phosphorylation and less NFkB nuclear translocation after TLR4 triggering than wild type cells (Avila et al., 2012).

The link between LYN and NFkB signalling in immunity is therefore supported by a number of lines of evidence. However, there is also support for a role of NFkB activation in BRCA1 loss-of-function-associated breast cancers. *BRCA1*-deficient mammary epithelial progenitors show sustained NFkB activation, suggested to be a consequence of replication-associated DNA damage (Sau et al., 2016). This NFkB activation was proposed to be the mechanism underlying hormone-independent growth of *BRCA1*-deficient luminal progenitors in colony formation assays (Lim et al., 2009; Sau et al., 2016). In contrast, a subset of *BRCA1*-mutant breast cancers were reported to show increased NFkB activity correlating with good prognosis (Buckley et al., 2016). This study demonstrated NFkB activity in a subset of *BRCA1*-mutant breast cancer cell lines *in vitro*, and identified a gene signature associated with that activity which identified tumours with an ‘NFkB-on’ signature from TNBC datasets.

These tumours had good prognosis and were associated with an increase in numbers of CD8+ cytotoxic T-cells, suggested to create an ‘anti-tumour microenvironment’ (Buckley et al., 2016). These contrasting studies suggest the link between BRCA1 and NFkB signalling is context-dependent, including differences depending on whether cell intrinsic or cell extrinsic processes are operating. We suggest that the context-dependent link between *BRCA1* breast tumours and NFkB activation is LYN kinase.

Our manuscript has limitations. It is clear that RNA expression does not necessarily reflect protein levels and protein levels do not necessarily reflect active protein levels. Classifying tumours on the basis of RNAseq counts for *Lyn* may not, therefore, actually reflect LYN activity in the tumour, and LYN staining may not reflect active LYN protein. Unfortunately, antibodies are not currently available for immunohistochemistry which are specific for active pLYN. Those that are available stain the phosphorylated active site of all SRC-family kinases. Flow cytometry to purify neoplastic tumours cells for analysis by, for example, western blot, is also problematic. Antibodies are not available which can be used to specifically mark all neoplastic epithelial-origin cells as opposed to non-neoplastic epithelium, or indeed other components of the tumour. However, there was broad agreement in most (although not all) tumours between RNAseq counts for *Lyn* and LYN histoscore, and it is clear that *Lyn* high tumours were biologically different to *Lyn* low tumours.

In the absence of a robust approach to purifying tumour cells, we opted to carry out RNAseq analysis from pieces of tumour which likely contained mixed populations of cells. Such pieces were taken from tumour regions away from obvious necrosis, but without any other selection criteria. This has the advantage of ensuring that sensitive RNA expression patterns are not altered during cell purification protocols but it is clearly not representative of a pure single cell type. It is, in effect, a snapshot of the biology of the tumour as a whole, rather than of a single cell type within in. Having identified the NFkB pathway as active in the *Lyn* high tumours, future studies should use approaches such as spatial transcriptomics to test the assumption that it is the epithelial tumour cells driving this biology.

As flow cytometric analysis of the tumours for immune cell subsets was not available to us, we used CIBERSORTx as a proxy to estimate immune cell populations in the tumours. This approach has previously proved to be robust (Steen et al., 2020).

Overall, our study supports a role for LYN kinase in the biology of BRCA1, and in *BRCA1*-mutation associated mammary cancer, particularly in the regulation of NFkB signalling. However, our findings also emphasise that context dependency of LYN likely extends

beyond immune cells and is a general property of this kinase. As a therapeutic target in breast cancer, therefore, LYN kinase is likely to present difficulties.

Acknowledgements

This work was supported by Breast Cancer Now (Grant Ref no: 2018NovPR1211). The RNAseq work was supported by Breast Cancer Research Aid. We acknowledge our colleagues at Wales Gene Park for their insight and expertise that assisted this research, for their technical and bioinformatic support in generating and analysing the NGS data. Wales Gene Park is an infrastructure support group with funding from Welsh Government through Health and Care Research Wales.

The authors would like to thank the Freedman Lab and also Catherine Hogan for their comments on the manuscript.

The authors declare they have no conflicts of interest.

Contributions

Conceptualisation MJS, GT; Investigation GT, LW, HK, TH, MJS; Formal analysis GT, LW, HK, TH, SB, MJS; Writing – Original Draft, MJS, GT; Supervision, MJS; Project Administration, MJS; Funding Acquisition, MJS, GT.

References

Avila, M., Martinez-Juarez, A., Ibarra-Sanchez, A., and Gonzalez-Espinosa, C. (2012). Lyn kinase controls TLR4-dependent IKK and MAPK activation modulating the activity of TRAF-6/TAK-1 protein complex in mast cells. Innate Immun *18*, 648-660.

Brian, B.F., and Freedman, T.S. (2021). The Src-family Kinase Lyn in Immunoreceptor Signaling. Endocrinology *162*.

Brian, B.F.t., Sauer, M.L., Greene, J.T., Senevirathne, S.E., Lindstedt, A.J., Funk, O.L., Ruis, B.L., Ramirez, L.A., Auger, J.L., Swanson, W.L.*, et al.* (2022). A dominant function of LynB kinase in preventing autoimmunity. Sci Adv *8*, eabj5227.

Buckley, N.E., Haddock, P., De Matos Simoes, R., Parkes, E., Irwin, G., Emmert-Streib, F., McQuaid, S., Kennedy, R., and Mullan, P. (2016). A BRCA1 deficient, NFkappaB driven immune signal predicts good outcome in triple negative breast cancer. Oncotarget *7*, 19884-19896.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884-i890.

Choi, Y.L., Bocanegra, M., Kwon, M.J., Shin, Y.K., Nam, S.J., Yang, J.H., Kao, J., Godwin, A.K., and Pollack, J.R. (2010). LYN is a mediator of epithelial-mesenchymal transition and a target of dasatinib in breast cancer. Cancer Res *70*, 2296-2306.

Cooper, M.J., Cox, N.J., Zimmerman, E.I., Dewar, B.J., Duncan, J.S., Whittle, M.C., Nguyen, T.A., Jones, L.S., Ghose Roy, S., Smalley, D.M.*, et al.* (2013). Application of multiplexed kinase inhibitor beads to study kinome adaptations in drug-resistant leukemia. PLoS One *8*, e66755.

Croucher, D.R., Hochgrafe, F., Zhang, L., Liu, L., Lyons, R.J., Rickwood, D., Tactacan, C.M., Browne, B.C., Ali, N., Chan, H.*, et al.* (2013). Involvement of Lyn and the atypical kinase SgK269/PEAK1 in a basal breast cancer signaling pathway. Cancer Res *73*, 1969-1980.

Densham, R.M., and Morris, J.R. (2017). The BRCA1 Ubiquitin ligase function sets a new trend for remodelling in DNA repair. Nucleus *8*, 116-125.

Dobin, A., and Gingeras, T.R. (2015). Mapping RNA-seq Reads with STAR. Curr Protoc Bioinformatics *51*, 11 14 11-11 14 19.

Foo, T.K., and Xia, B. (2022). BRCA1-Dependent and Independent Recruitment of PALB2-BRCA2-RAD51 in the DNA Damage Response and Cancer. Cancer Res *82*, 3191-3197.

Freedman, T.S., Tan, Y.X., Skrzypczynska, K.M., Manz, B.N., Sjaastad, F.V., Goodridge, H.S., Lowell, C.A., and Weiss, A. (2015). LynA regulates an inflammation-sensitive signaling checkpoint in macrophages. Elife *4*.

Fu, X., Tan, W., Song, Q., Pei, H., and Li, J. (2022). BRCA1 and Breast Cancer: Molecular Mechanisms and Therapeutic Strategies. Front Cell Dev Biol *10*, 813457.

Hameyer, D., Loonstra, A., Eshkind, L., Schmitt, S., Antunes, C., Groen, A., Bindels, E., Jonkers, J., Krimpenfort, P., Meuwissen, R.*, et al.* (2007). Toxicity of ligand-dependent Cre recombinases and generation of a conditional Cre deleter mouse allowing mosaic recombination in peripheral tissues. Physiol Genomics *31*, 32-41.

Hibbs, M.L., Harder, K.W., Armes, J., Kountouri, N., Quilici, C., Casagranda, F., Dunn, A.R., and Tarlinton, D.M. (2002). Sustained activation of Lyn tyrosine kinase in vivo leads to autoimmunity. J Exp Med *196*, 1593-1604.

Hibbs, M.L., Tarlinton, D.M., Armes, J., Grail, D., Hodgson, G., Maglitto, R., Stacker, S.A., and Dunn, A.R. (1995). Multiple defects in the immune system of Lyn-deficient mice, culminating in autoimmune disease. Cell *83*, 301-311.

Hochgrafe, F., Zhang, L., O'Toole, S.A., Browne, B.C., Pinese, M., Porta Cubas, A., Lehrbach, G.M., Croucher, D.R., Rickwood, D., Boulghourjian, A.*, et al.* (2010). Tyrosine phosphorylation profiling reveals the signaling network characteristics of Basal breast cancer cells. Cancer Res *70*, 9391-9401.

Huang, T.H., Huo, L., Wang, Y.N., Xia, W., Wei, Y., Chang, S.S., Chang, W.C., Fang, Y.F., Chen, C.T., Lang, J.Y.*, et al.* (2013). Epidermal growth factor receptor potentiates MCM7-mediated DNA replication through tyrosine phosphorylation of Lyn kinase in human cancers. Cancer Cell *23*, 796-810.

Kendrick, H., Regan, J.L., Magnay, F.A., Grigoriadis, A., Mitsopoulos, C., Zvelebil, M., and Smalley, M.J. (2008). Transcriptome analysis of mammary epithelial subpopulations identifies novel determinants of lineage commitment and cell fate. BMC Genomics *9*, 591.

Kim, M.C., Choi, J.E., Lee, S.J., and Bae, Y.K. (2016). Coexistent Loss of the Expressions of BRCA1 and p53 Predicts Poor Prognosis in Triple-Negative Breast Cancer. Ann Surg Oncol *23*, 3524-3530.

Kuchenbaecker, K.B., Hopper, J.L., Barnes, D.R., Phillips, K.A., Mooij, T.M., Roos-Blom, M.J., Jervis, S., van Leeuwen, F.E., Milne, R.L., Andrieu, N.*, et al.* (2017). Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. JAMA *317*, 2402-2416.

Li, X., Xu, M., Shen, J., Li, Y., Lin, S., Zhu, M., Pang, Q., Tan, X., and Tang, J. (2022). Sorafenib inhibits LPS-induced inflammation by regulating Lyn-MAPK-NF-kB/AP-1 pathway and TLR4 expression. Cell Death Discov *8*, 281.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.

Lim, E., Vaillant, F., Wu, D., Forrest, N.C., Pal, B., Hart, A.H., Asselin-Labat, M.L., Gyorki, D.E., Ward, T., Partanen, A.*, et al.* (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat Med *15*, 907-913.

Liu, X., Holstege, H., van der Gulden, H., Treur-Mulder, M., Zevenhoven, J., Velds, A., Kerkhoven, R.M., van Vliet, M.H., Wessels, L.F., Peterse, J.L.*, et al.* (2007). Somatic loss of BRCA1 and p53 in mice induces mammary tumors with features of human BRCA1-mutated basal-like breast cancer. Proc Natl Acad Sci U S A *104*, 12111-12116.

Mahdavi, M., Nassiri, M., Kooshyar, M.M., Vakili-Azghandi, M., Avan, A., Sandry, R., Pillai, S., Lam, A.K., and Gopalan, V. (2019). Hereditary breast cancer; Genetic penetrance and current status with BRCA. J Cell Physiol *234*, 5741-5750.

Marchetti, S., Gamas, P., Belhacene, N., Grosso, S., Pradelli, L.A., Colosetti, P., Johansen, C., Iversen, L., Deckert, M., Luciano, F.*, et al.* (2009). The caspase-cleaved form of LYN mediates a psoriasis-like inflammatory syndrome in mice. EMBO J *28*, 2449-2460.

Mateo, J., Lord, C.J., Serra, V., Tutt, A., Balmana, J., Castroviejo-Bermejo, M., Cruz, C., Oaknin, A., Kaye, S.B., and de Bono, J.S. (2019). A decade of clinical development of PARP inhibitors in perspective. Ann Oncol *30*, 1437-1447.

McCarthy, A., Savage, K., Gabriel, A., Naceur, C., Reis-Filho, J.S., and Ashworth, A. (2007). A mouse model of basal-like breast carcinoma with metaplastic elements. J Pathol *211*, 389-398.

Molyneux, G., Geyer, F.C., Magnay, F.A., McCarthy, A., Kendrick, H., Natrajan, R., Mackay, A., Grigoriadis, A., Tutt, A., Ashworth, A.*, et al.* (2010). BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. Cell Stem Cell *7*, 403-417.

Newman, A.M., Steen, C.B., Liu, C.L., Gentles, A.J., Chaudhuri, A.A., Scherer, F., Khodadoust, M.S., Esfahani, M.S., Luca, B.A., Steiner, D.*, et al.* (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. Nature Biotechnology *37*, 773-782.

Nolan, E., Lindeman, G.J., and Visvader, J.E. (2017). Out-RANKing BRCA1 in Mutation Carriers. Cancer Res *77*, 595-600.

O'Donovan, P.J., and Livingston, D.M. (2010). BRCA1 and BRCA2: breast/ovarian cancer susceptibility gene products and participants in DNA double-strand break repair. Carcinogenesis *31*, 961-967.

Percie du Sert, N., Hurst, V., Ahluwalia, A., Alam, S., Avey, M.T., Baker, M., Browne, W.J., Clark, A., Cuthill, I.C., Dirnagl, U.*, et al.* (2020). The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. Br J Pharmacol *177*, 3617-3624.

Regan, J.L., Kendrick, H., Magnay, F.A., Vafaizadeh, V., Groner, B., and Smalley, M.J. (2012). c-Kit is required for growth and survival of the cells of origin of Brca1-mutation-associated breast cancer. Oncogene *31*, 869-883.

Ren, X., Cao, C., Zhu, L., Yoshida, K., Kharbanda, S., Weichselbaum, R., and Kufe, D. (2002). Lyn tyrosine kinase inhibits nuclear export of the p53 tumor suppressor. Cancer Biol Ther *1*, 703-708.

Sau, A., Lau, R., Cabrita, M.A., Nolan, E., Crooks, P.A., Visvader, J.E., and Pratt, M.A. (2016). Persistent Activation of NF-kappaB in BRCA1-Deficient Mammary Progenitors Drives Aberrant Proliferation and Accumulation of DNA Damage. Cell Stem Cell *19*, 52-65.

Smalley, M.J. (2010). Isolation, culture and analysis of mouse mammary epithelial cells. Methods Mol Biol *633*, 139-170.

Steen, C.B., Liu, C.L., Alizadeh, A.A., and Newman, A.M. (2020). Profiling Cell Type Abundance and Expression in Bulk Tissues with CIBERSORTx. Methods Mol Biol *2117*, 135-157.

Thaper, D., Vahid, S., Nip, K.M., Moskalev, I., Shan, X., Frees, S., Roberts, M.E., Ketola, K., Harder, K.W., Gregory-Evans, C.*, et al.* (2017). Targeting Lyn regulates Snail family shuttling and inhibits metastasis. Oncogene *36*, 3964-3975.

Tornillo, G., Knowlson, C., Kendrick, H., Cooke, J., Mirza, H., Aurrekoetxea-Rodriguez, I., Vivanco, M.D.M., Buckley, N.E., Grigoriadis, A., and Smalley, M.J. (2018). Dual Mechanisms

of LYN Kinase Dysregulation Drive Aggressive Behavior in Breast Cancer Cells. Cell Rep *25*, 3674-3692 e3610.

Toubiana, J., Rossi, A.L., Belaidouni, N., Grimaldi, D., Pene, F., Chafey, P., Comba, B., Camoin, L., Bismuth, G., Claessens, Y.E.*, et al.* (2015). Src-family-tyrosine kinase Lyn is critical for TLR2-mediated NF-kappaB activation through the PI 3-kinase signaling pathway. Innate Immun *21*, 685-697.

Varet, H., Brillet-Gueguen, L., Coppee, J.Y., and Dillies, M.A. (2016). SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. PLoS One *11*, e0157022.

Xu, Y., Harder, K.W., Huntington, N.D., Hibbs, M.L., and Tarlinton, D.M. (2005). Lyn tyrosine kinase: accentuating the positive and the negative. Immunity *22*, 9-18.

Yoshida, K., and Miki, Y. (2004). Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. Cancer Sci *95*, 866-871.

Zhu, L., Pan, J.N., Qian, Z., Ye, W.W., Wang, X.J., and Cao, W.M. (2022). High chromosome instability identified by low-pass whole-genome sequencing assay is associated with TP53 copy loss and worse prognosis in BRCA1 germline mutation breast cancer. Breast Cancer *29*, 103-113.

Figure Legends

Figure 1: The *Lyn^{fl}* allele is efficiently recombined *ex vivo* to deplete LYN protein. A) Schematic of the *Lyn^{tm1c}* (*Lyn$^{fl(ex4)}$* or *Lyn^{fl}*) allele as supplied by the Mary Lyon Centre, MRC Harwell. Location of *loxP* sites flanking exon 4 indicated by red triangles. The single *Frt* site (green) is a remnant of the targeting strategy used to generate the allele. **B)** Quantitative real-time rtPCR analysis of *Lyn* levels using an exon 4-specific probe in cultured primary mouse mammary epithelial cells from *R26C Brca1$^{fl/+}$ Lyn$^{wt/wt}$*, *R26C Brca1$^{fl/+}$ Lyn$^{fl/wt}$* and *R26C Brca1$^{fl/+}$ Lyn$^{fl/fl}$* mice, treated with either vehicle control or tamoxifen. Results expressed as mean relative exon 4 levels ± 95% confidence intervals compared to to vehicle-treated *R26C Brca1$^{fl/+}$ Lyn$^{fl/wt}$* cells. **C)** Western blot for LYN expression in protein extracts from the cell cultures analysed in (B) (Thermofisher polyclonal antibody). **D)** Western blot analysis of LYN expression in *R26C Brca1$^{fl/+}$ Lyn$^{wt/wt}$* and *R26C Brca1$^{fl/+}$ Lyn$^{fl/fl}$* primary mammary epithelial cells, with or without tamoxifen treatment, using three different commercial anti-LYN antibodies. Each antibody was tested on identical loading of the same protein extracts.

Figure 2: *Lyn$^{fl/fl}$* mice have decreased overall survival in the *BlgCre Brca1$^{fl/fl}$ p53$^{+/-}$* model and have tumours with highly heterogeneous growth characteristics. A) Kaplan-Meier curve of overall survival of *BlgCre Brca1$^{fl/fl}$ p53$^{+/-}$ Lyn$^{wt/wt}$*, *BlgCre Brca1$^{fl/fl}$ p53$^{+/-}$ Lyn$^{fl/wt}$* and *BlgCre Brca1$^{fl/fl}$ p53$^{+/-}$ Lyn$^{fl/fl}$* cohorts. *Lyn$^{fl/fl}$* mice have a significantly shorter survival than *Lyn$^{fl/wt}$* mice ($P<0.05$; Log Rank test). **B)** Kaplan-Meier curve of survival of *BlgCre Brca1$^{fl/fl}$ p53$^{+/-}$ Lyn$^{wt/wt}$*, *BlgCre Brca1$^{fl/fl}$ p53$^{+/-}$ Lyn$^{fl/wt}$* and *BlgCre Brca1$^{fl/fl}$ p53$^{+/-}$ Lyn$^{fl/fl}$* cohorts considering only mice euthanased as a result of a mammary tumour reaching a specified endpoint. There are no significant differences. **C)** Reason for euthanasia in *BlgCre Brca1$^{fl/fl}$ p53$^{+/-}$ Lyn$^{wt/wt}$*, *BlgCre Brca1$^{fl/fl}$ p53$^{+/-}$ Lyn$^{fl/wt}$* and *BlgCre Brca1$^{fl/fl}$ p53$^{+/-}$ Lyn$^{fl/fl}$* cohorts. No significant differences (Chi^2 test for trend). **D)** Doubling times for individual tumours (days) in each cohort. The value for each tumour is plotted with the mean ± SD. No significant differences (one-way ANOVA across cohorts and also t-tests comparing each cohort to the others). **E)** Ki67 percentage positivity in sections of mammary adenocarcinomas from each cohort. The value for each tumour is plotted with the mean ± SD. No significant differences (Kruskel-Wallis across cohorts and Mann-Whitney tests comparing each cohort to the others). **F)** Tumour volume doubling times (days) by animal showing only animals from each cohort with more than one tumour measured.

Figure 3: LYN expression in tumours is heterogeneous but decreased overall in *BlgCre Brca1$^{fl/fl}$ p53$^{+/-}$ Lyn$^{fl/fl}$* mice. A) LYN staining pattern in epithelial-like tumour cells. Bar = 20 μm. **B)** LYN staining in single cells (white arrowheads) and cancer-associated

fibroblast-like cells (black arrows). Bar = 20 μm. **C)** 'Histoscore' quantitation of LYN staining in epithelial-like tumour cells, estimating strength of staining (0, 1, 2, 3) and area of tumour stained (0, no staining; 1, <10% of tumour cells positive; 2, 10 – 50% of tumour cells positive; 3, >50% of tumour cells positive) and then multiplying the two scores together to give a final value. The *Lyn*$^{fl/fl}$ cohort has significantly less staining than the *Lyn*$^{wt/wt}$ or *Lyn*$^{fl/wt}$ cohort (**P<0.01, *P<0.05, Mann-Whitney test). **D)** Semi-quantitative PCR analysis of *Lyn*$^{fl(ex4)}$ and *Lyn*^{wt} alleles in genomic DNA isolated from primary cultures of *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{wt/wt}$ (n=6), *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/wt}$ (n=5) and *BlgCre Brca1*$^{fl/fl}$ *p53*$^{+/-}$ *Lyn*$^{fl/fl}$ (n=6) tumour cells. Analysis of the *Apc* locus was included as a control to enable relative quantitation. Each PCR reaction included the primers for all three alleles. **E)** Quantitation of relative abundance of *Lyn*^{wt} alleles in (D), considering only *Lyn*$^{wt/wt}$ and *Lyn*$^{fl/wt}$ cultures (as no *Lyn*^{wt} alleles are present in *Lyn*$^{fl/fl}$ cells). Data presented as abundance in each sample relative to the *Apc* band in that sample. The mean abundance ± SD of each group is indicated. There are 50% fewer (P<0.01, Mann-Whitney test) *Lyn*^{wt} alleles in the *Lyn*$^{fl/wt}$ compared to the *Lyn*$^{wt/wt}$ cultures, as expected. **F)** Quantitation of relative abundance of *Lyn*$^{fl(ex4)}$ alleles in (D), considering only *Lyn*$^{fl/wt}$ and *Lyn*$^{fl/fl}$ cultures (as no *Lyn*^{fl} alleles are present in *Lyn*$^{wt/wt}$ cells). Data presented as abundance in each sample relative to the *Apc* band in that sample. The mean abundance ± SD of each group is indicated. No significant difference between the samples, but the heterogeneity of the samples reflects the clear differences in band intensities seen in (D). **G)** Quantitative real-time rtPCR analysis of *Lyn exon 4* expression in primary tumour cell cultures. Data presented as expression relative to the mean value for the *Lyn*$^{wt/wt}$ cells (*P<0.05, Mann-Whitney test). **H)** Quantitative real-time rtPCR analysis of *Brca1* expression in primary tumour cell cultures. Data presented as expression relative to *Brca1* levels in lysates of freshly isolated normal mouse mammary epithelial cells for *Brca1* (no significant differences, Mann-Whitney tests). **I)** Relative expression of LYN protein as determined by Western blot analysis, quantified relative to standard loading controls and normalised to one *Lyn*$^{wt/wt}$ tumour sample (see **Supplementary Data File** for raw blots).

Figure 4: Tumour molecular profiles correlate with *Lyn* expression but not with tumour cohort. A) Volcano plot of significantly differentially expressed genes comparing *Lyn*$^{wt/wt}$ and *Lyn*$^{fl/fl}$ tumours. Genes with an adjusted P value of <0.05 and a fold change of ≤0.5 or ≥2 were considered significant. **B)** Normalised RNAseq *Lyn* counts in tumours from each cohort. Mean values ± SD are indicated. No significant differences (ANOVA). **C)** Normalised RNAseq *Lyn* counts in tumours with different Histoscore grading (0, no LYN staining; 1/2/3, low LYN staining or strong staining but only in a small region; 4/6/9, moderate to strong LYN staining). Increased *Lyn* counts are associated with increasing

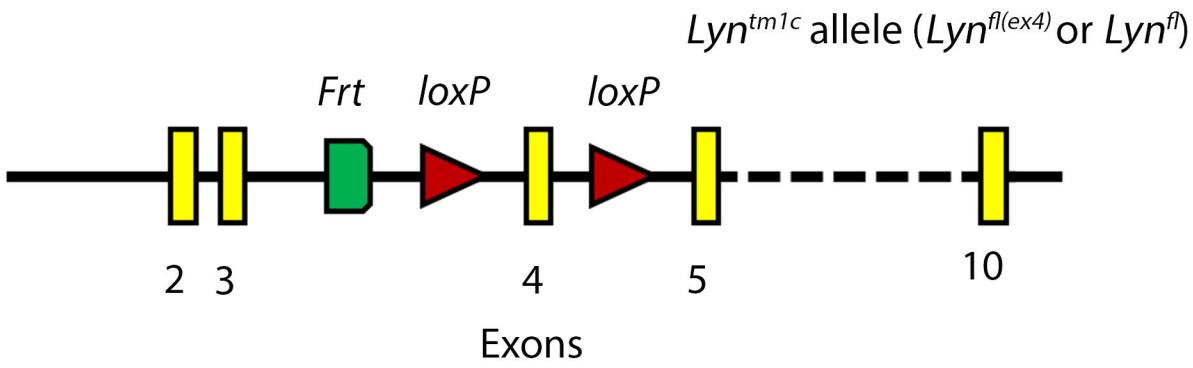
histoscore (ANOVA, *P<0.05). **D)** Histoscore of top tertile (*Lyn* high) vs bottom tertile (*Lyn* low) *Lyn* expressing tumours by RNAseq. Mean values ± SD are indicated. **P<0.01 (Mann-Whitney). **E)** *In vivo* doubling time (days) of top tertile (*Lyn* high) vs bottom tertile (*Lyn* low) *Lyn* expressing tumours by RNAseq. Mean values ± SD are indicated. *P<0.05 (t-test). **F)** Volcano plot of significantly differentially expressed genes comparing top tertile (*Lyn* high) vs bottom tertile (*Lyn* low) *Lyn* expressing tumours defined by RNAseq. Genes with an adjusted P value of <0.05 and a fold change of ≤0.5 or ≥2 were considered significant.

Figure 5: Principal Component Analysis (PCA) identifies two main groups of tumours which partially overlap with *Lyn* high and *Lyn* low tumours. **A)** PCA plot of tumours analysed by RNAseq showing an unbiased assessment of tumour gene expression differences and similarities. PC1 divides the tumours into groups 1/2 and 3/4, PC2 further divides the tumours to give four groups; however, the majority of the differences between the tumour groups was generated by PC1. **B)** Normalised RNAseq *Lyn* counts in tumours from PCA groups 1/2 and 3/4. Mean values ± SD are indicated. *P<0.05 (t-test). **C)** Histoscore of tumours from PCA groups 1/2 and 3/4. Mean values ± SD are indicated. **P<0.01 (Mann-Whitney). **D)** *In vivo* doubling time (days) of tumours from PCA groups 1/2 and 3/4. Mean values ± SD are indicated. *P<0.05 (t-test). **E)** Volcano plot of significantly differentially expressed genes comparing tumours from PCA groups 1/2 and 3/4. Genes with an adjusted P value of <0.05 and a fold change of ≤0.5 or ≥2 were considered significant. **F)** Venn diagram showing overlap between differentially expressed genes identified when comparing $Lyn^{wt/wt}$ vs $Lyn^{fl/fl}$ tumours, *Lyn* high vs *Lyn* low tumours and PCA group 1/2 vs 3/4 tumours. Note 340 genes overlap between the latter two groups but there is very little overlap with the $Lyn^{wt/wt}$ vs $Lyn^{fl/fl}$ tumour data.

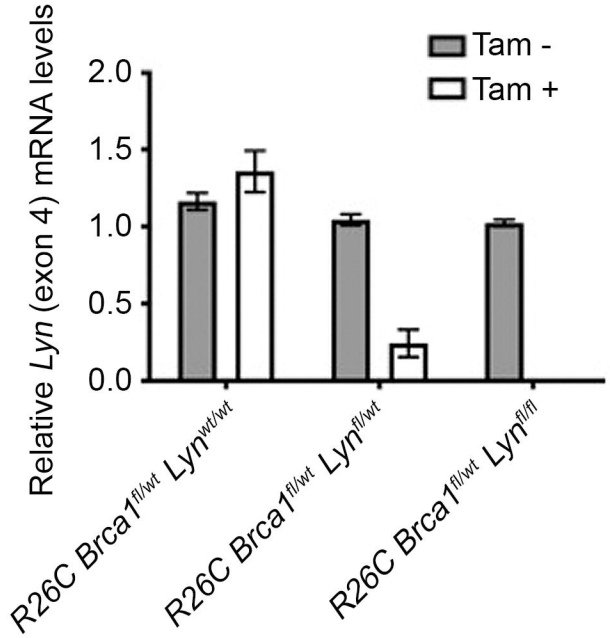
Figure 6: The *Lyn* high and PCA group 3/4 tumours are enriched for identical biological functions. A) Venn diagram analysis of overlap between GO Bioprocess terms enriched in the DEGs from *Lyn* high, *Lyn* low, PCA group 1/2 and PCA group 3/4 tumours. **B)** Venn diagram analysis of overlap between KEGG pathways enriched in the DEGs from *Lyn* high, *Lyn* low, PCA group 1/2 and PCA group 3/4 tumours. **C)** Distribution of enriched GO Bioprocess terms within functional categories for the *Lyn* high, *Lyn* low, PCA group 1/2 and PCA group 3/4 tumours. The number of enriched GO BP terms identified in the DEGs of each tumour group is indicated above each bar. Each bar is divided up according to the percentage of enriched GO BP terms falling into each functional category (indicated by the colour key).

Figure 7: Immune cell complexity is not altered between tumour groups with different levels of *Lyn* expression. CIBERsortx estimates of immune cell abundance in the *Lyn* low vs *Lyn* high **(A)** and PCA group 3/4 vs PCA group 1/2 tumours **(B)** based on normalised RNAseq expression data. The only significant difference is in resting NK cells in the PCA tumour group comparison (t-test). Estimates are provided using the absolute abundance (arbitrary units) approach (Newman et al., 2019) so that the proportions of each different immune cell type are comparable.

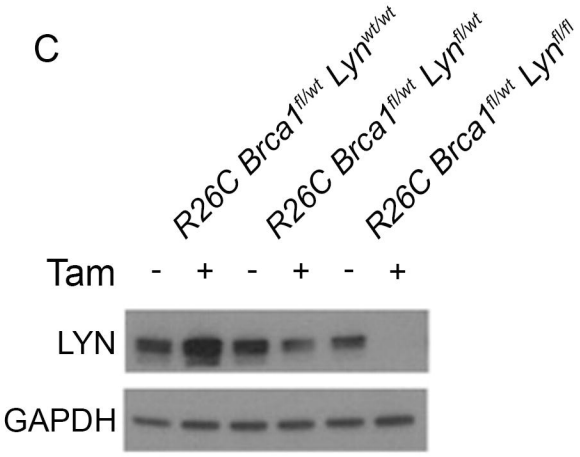
A


Lyntm1c allele (Lynfl(ex4) or Lynfl)
Frt
loxP
loxP
2
3
4
5
10
Exons

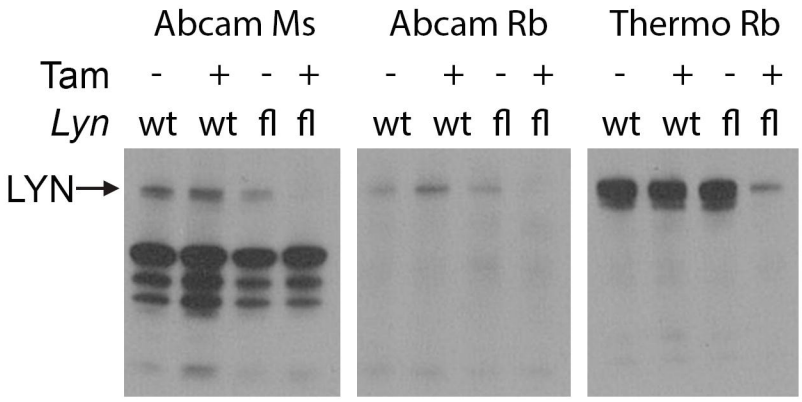
B

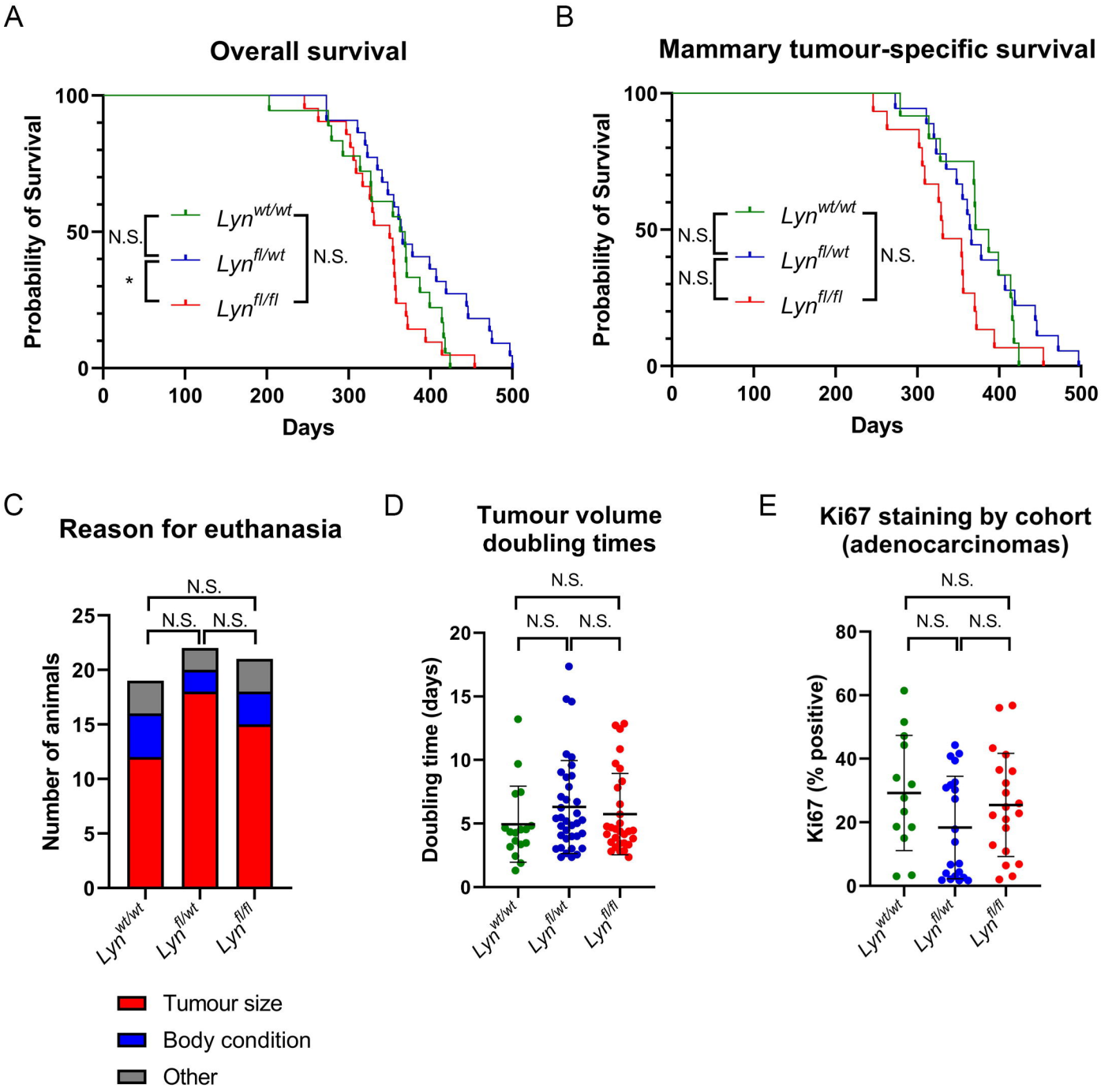

Relative Lyn (exon 4) mRNA levels
2.0
1.5
1.0
0.5
0.0
Tam -
Tam +
R26C Brca1fl/wt Lynwt/wt
R26C Brca1fl/wt Lynfl/wt
R26C Brca1fl/wt Lynfl/fl

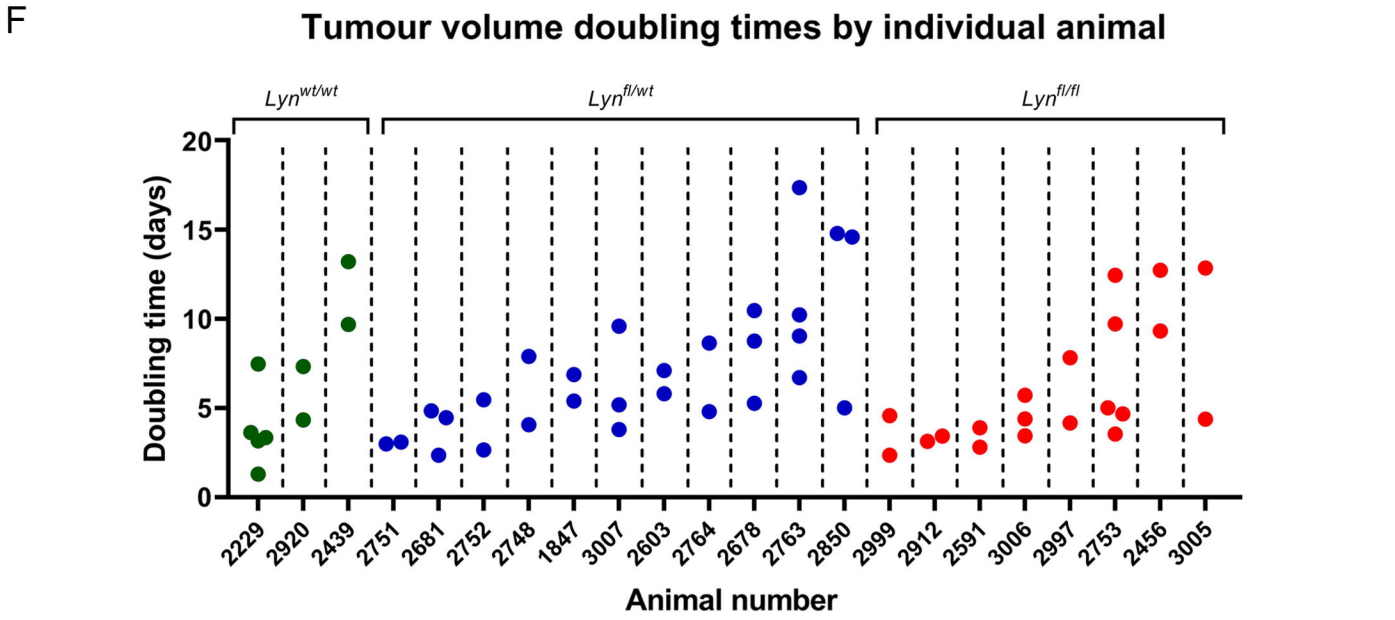
C

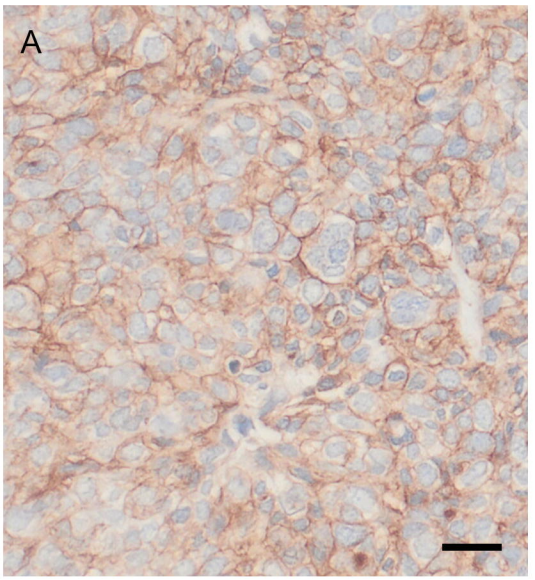

R26C Brca1fl/wt Lynwt/wt
R26C Brca1fl/wt Lynfl/wt
R26C Brca1fl/wt Lynfl/fl
Tam - + - + - +
LYN
GAPDH

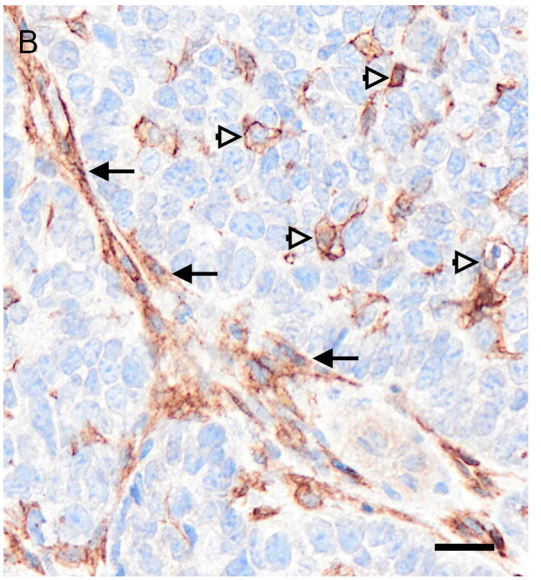
D

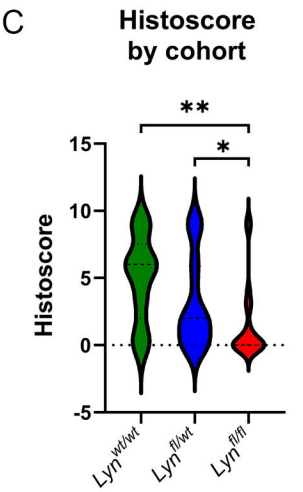

Abcam Ms
Abcam Rb
Thermo Rb
Tam - + - + - + - + - + - +
Lyn wt wt fl fl wt wt fl fl wt wt fl fl
LYN

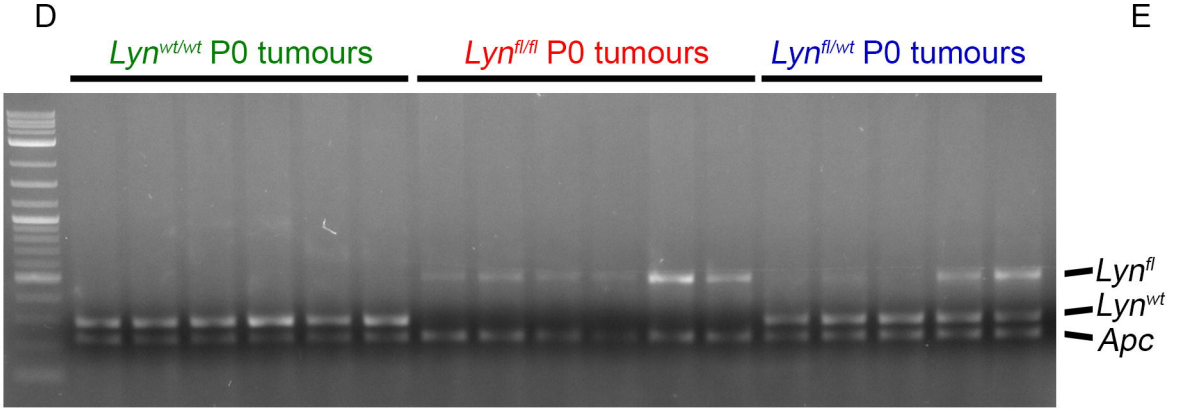

A
Overall survival
Probability of Survival
Days
N.S.
*
N.S.
Lyn wt/wt
Lyn fl/wt
Lyn fl/fl
B
Mammary tumour-specific survival
Probability of Survival
Days
N.S.
N.S.
N.S.
C
Reason for euthanasia
Number of animals
N.S.
N.S.
N.S.
Tumour size
Body condition
Other
D
Tumour volume doubling times
Doubling time (days)
E
Ki67 staining by cohort (adenocarcinomas)
Ki67 (% positive)

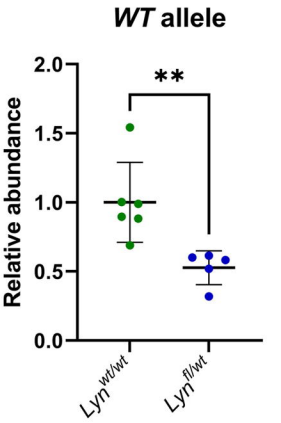

F
Tumour volume doubling times by individual animal
Lyn wt/wt
Lyn fl/wt
Lyn fl/fl
Doubling time (days)
2229
2920
2439
2751
2681
2752
2748
1847
3007
2603
2764
2678
2763
2850
2999
2912
2591
3006
2997
2753
2456
3005
Animal number

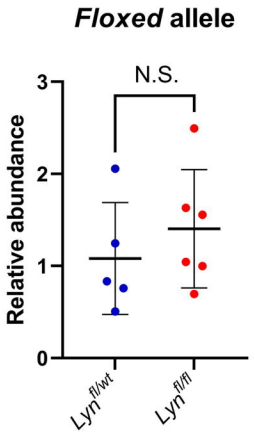

A

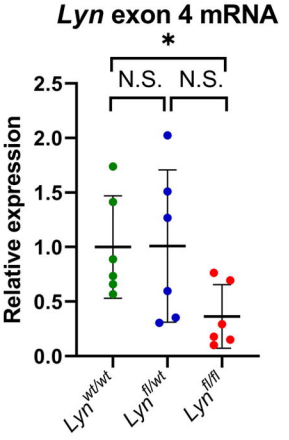

B

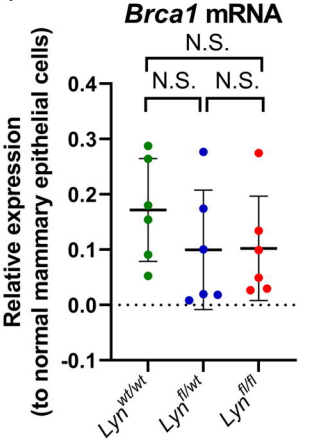

C
Histoscore by cohort
**
*
Histoscore
15
10
5
0
-5
Lyn wt/wt
Lyn fl/wt
Lyn fl/fl

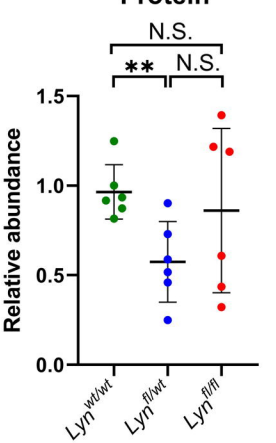

D
Lyn wt/wt P0 tumours
Lyn fl/fl P0 tumours
Lyn fl/wt P0 tumours
Lyn fl
Lyn wt
Apc

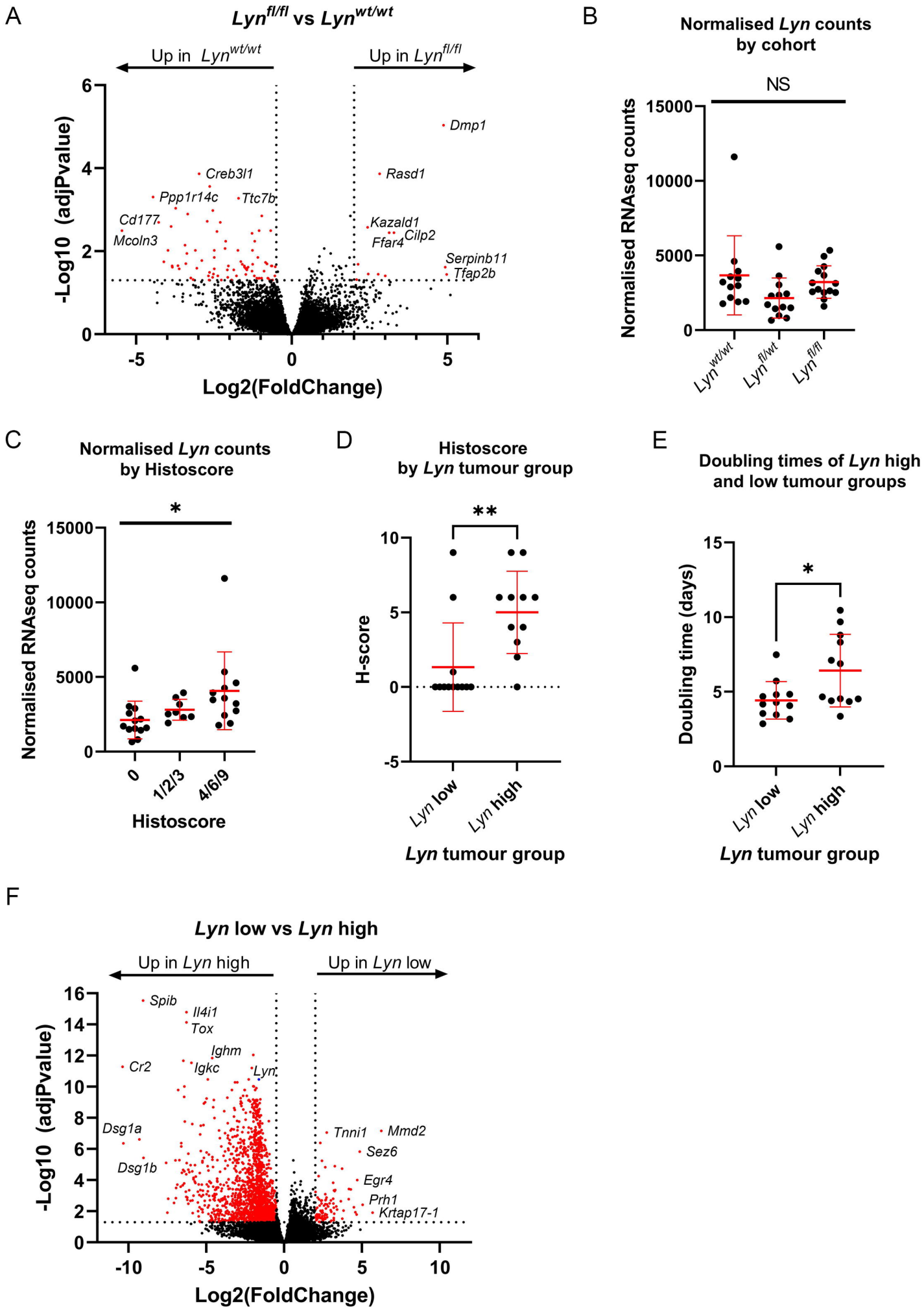

E
WT allele
**
Relative abundance
2.0
1.5
1.0
0.5
0.0
Lyn wt/wt
Lyn fl/wt

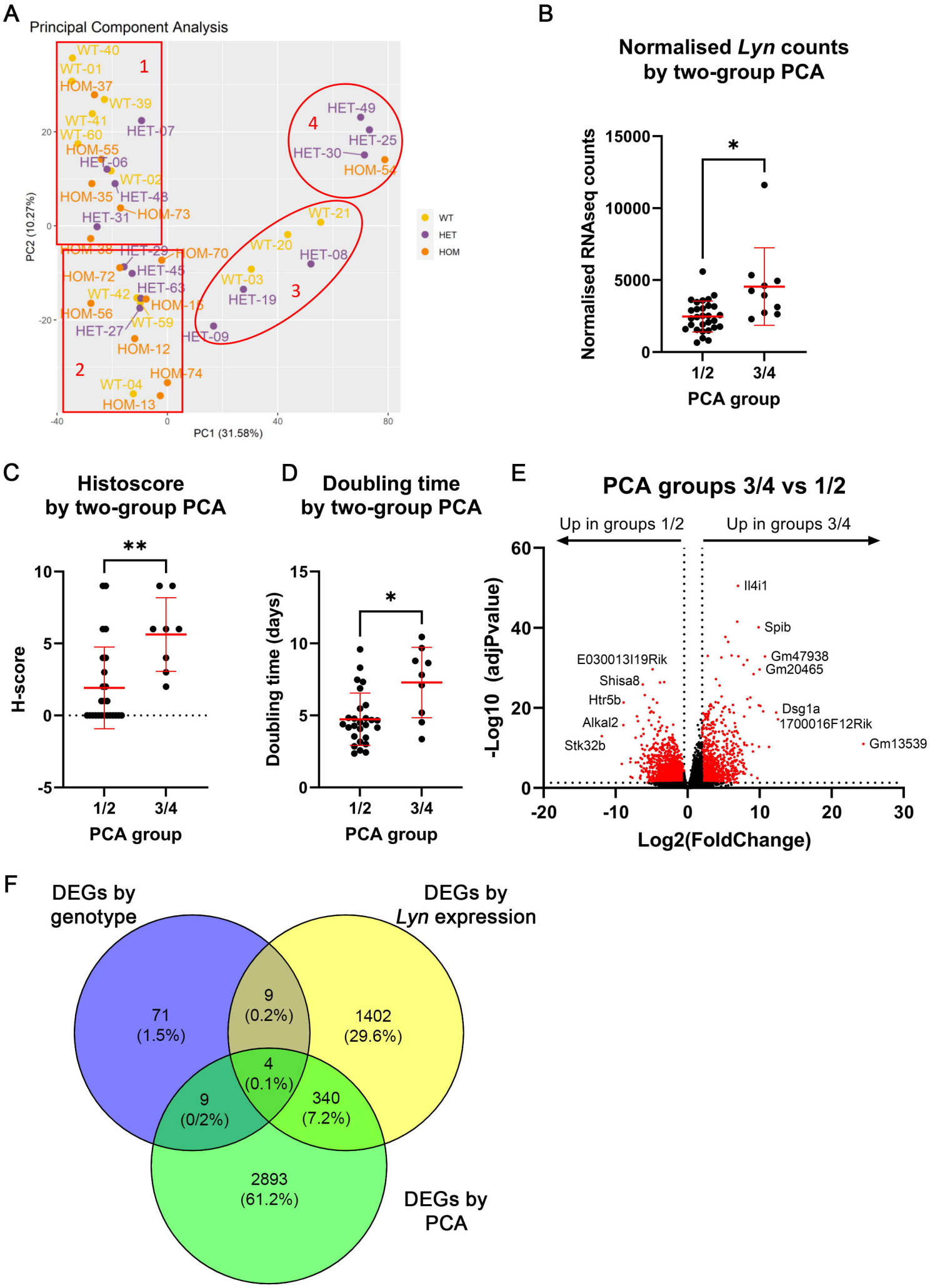

F
Floxed allele
N.S.
Relative abundance
3
2
1
0
Lyn fl/wt
Lyn fl/fl


G
Lyn exon 4 mRNA
*
N.S.
N.S.
Relative expression
2.5
2.0
1.5
1.0
0.5
0.0
Lyn wt/wt
Lyn fl/wt
Lyn fl/fl

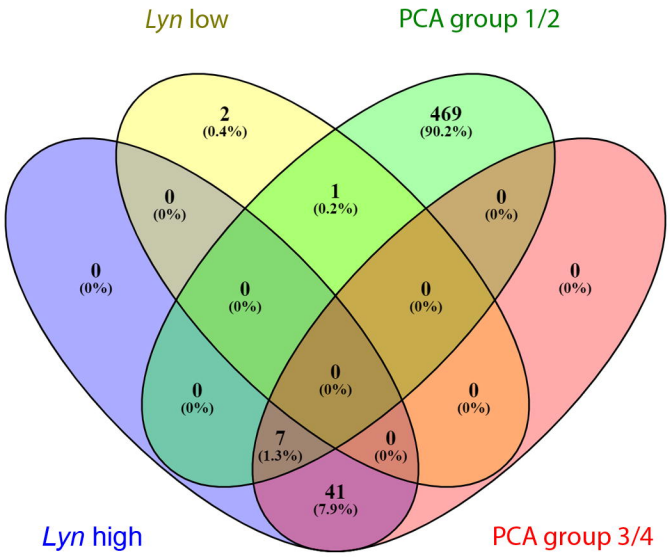

H
Brca1 mRNA
N.S.
N.S.
N.S.
Relative expression
(to normal mammary epithelial cells)
0.4
0.3
0.2
0.1
0.0
-0.1
Lyn wt/wt
Lyn fl/wt
Lyn fl/fl


I
Protein
N.S.
**
N.S.
Relative abundance
1.5
1.0
0.5
0.0
Lyn wt/wt
Lyn fl/wt
Lyn fl/fl


A
Lyn fl/fl vs Lyn wt/wt
Up in Lyn wt/wt
Up in Lyn fl/fl
-Log10 (adjPvalue)
Log2(FoldChange)
Dmp1
Creb3l1
Rasd1
Ppp1r14c
Ttc7b
Cd177
Mcoln3
Kazald1
Ffar4
Cilp2
Serpinb11
Tfap2b
B
Normalised Lyn counts by cohort
NS
Normalised RNAseq counts
Lyn wt/wt
Lyn fl/wt
Lyn fl/fl
C
Normalised Lyn counts by Histoscore
*
Normalised RNAseq counts
0
1/2/3
4/6/9
Histoscore
D
Histoscore by Lyn tumour group
**
H-score
Lyn low
Lyn high
Lyn tumour group
E
Doubling times of Lyn high and low tumour groups
*
Doubling time (days)
Lyn low
Lyn high
Lyn tumour group
F
Lyn low vs Lyn high
Up in Lyn high
Up in Lyn low
-Log10 (adjPvalue)
Log2(FoldChange)
Spib
Il4i1
Tox
Ighm
Cr2
Igkc
Lyn
Dsg1a
Dsg1b
Tnni1
Mmd2
Sez6
Egr4
Prh1
Krtap17-1

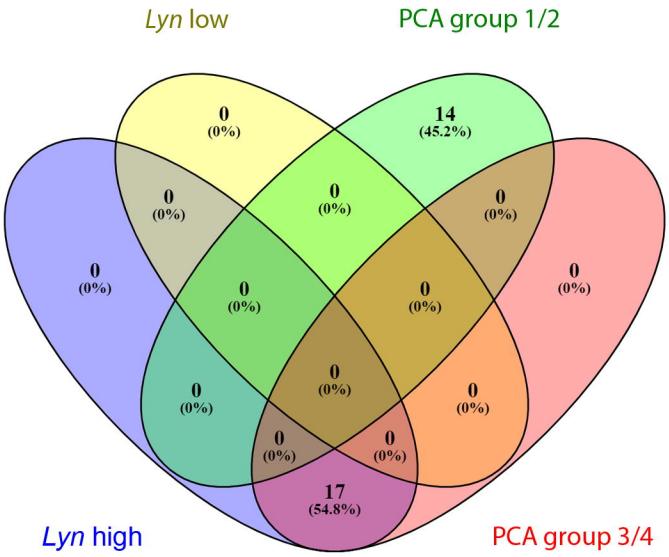

A
Principal Component Analysis
PC2 (10.27%)
PC1 (31.58%)
WT
HET
HOM
1
2
3
4
B
Normalised Lyn counts
by two-group PCA
Normalised RNAseq counts
PCA group
1/2
3/4
C
Histoscore
by two-group PCA
H-score
PCA group
D
Doubling time
by two-group PCA
Doubling time (days)
PCA group
E
PCA groups 3/4 vs 1/2
Up in groups 1/2
Up in groups 3/4
-Log10 (adjPvalue)
Log2(FoldChange)
Il4i1
Spib
Gm47938
Gm20465
Dsg1a
1700016F12Rik
Gm13539
E030013I19Rik
Shisa8
Htr5b
Alkal2
Stk32b
F
DEGs by genotype
DEGs by Lyn expression
DEGs by PCA
71 (1.5%)
9 (0.2%)
1402 (29.6%)
4 (0.1%)
9 (0/2%)
340 (7.2%)
2893 (61.2%)

A



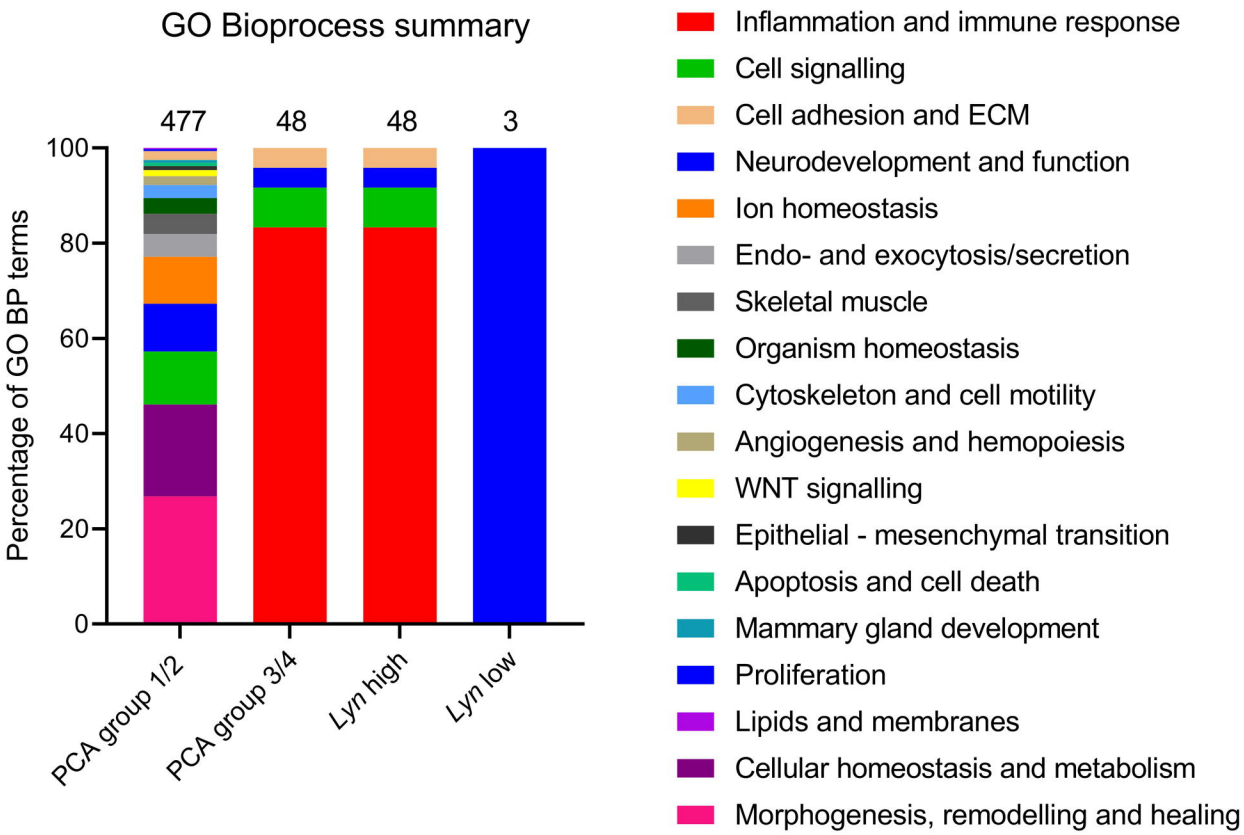
GO Bioprocess overlaps

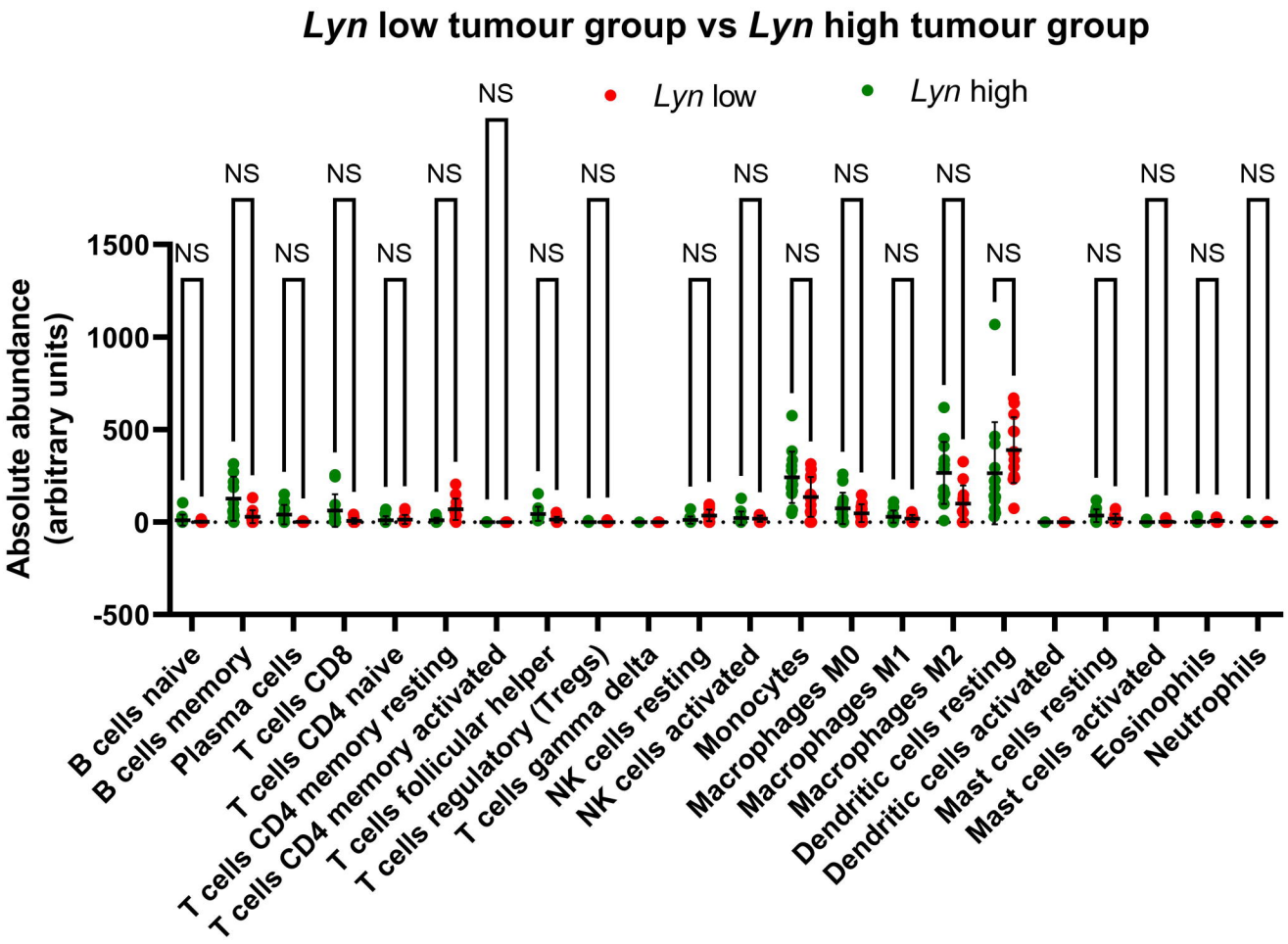
B

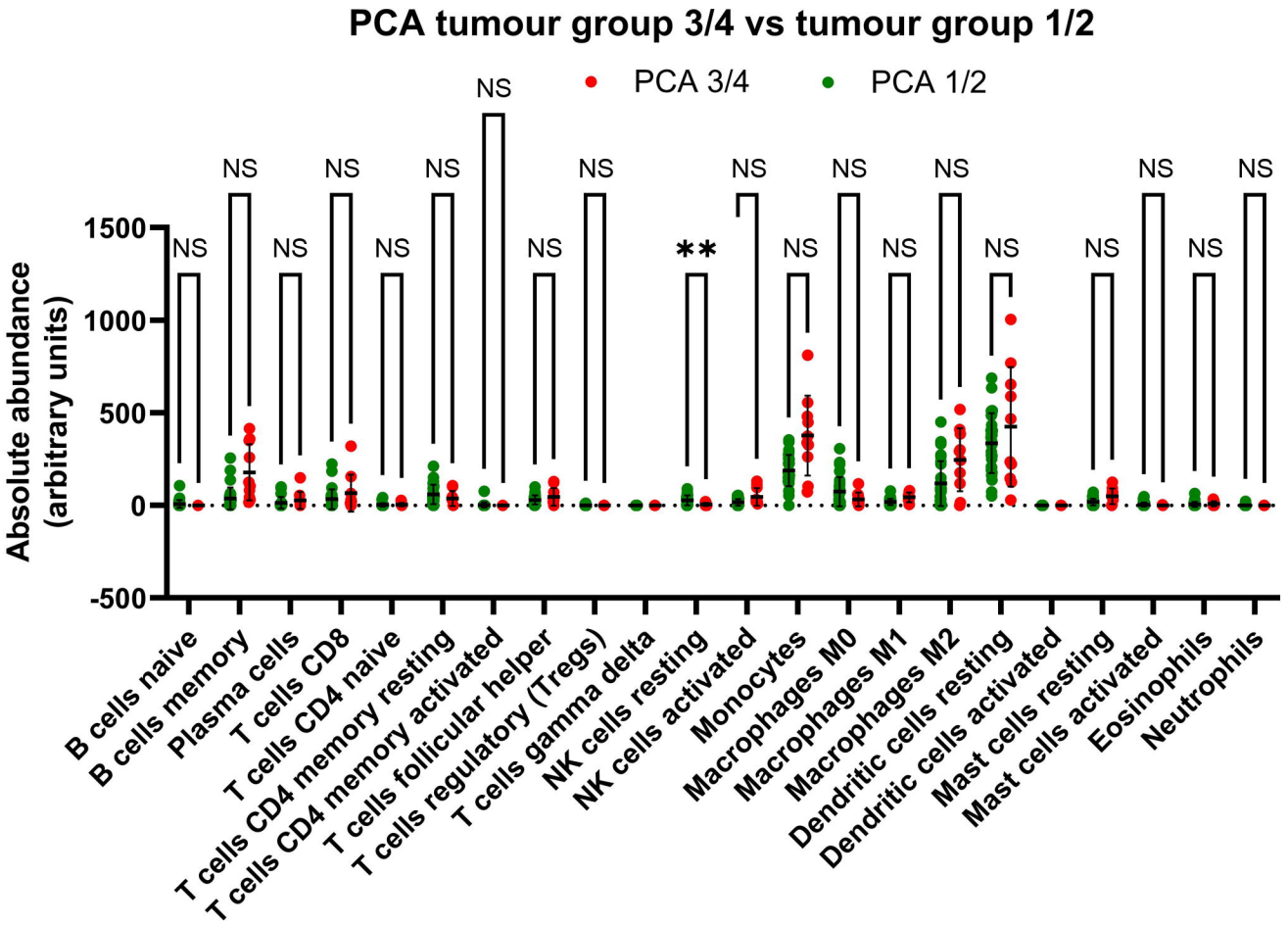


KEGG pathway overlaps

C




Lyn low tumour group vs *Lyn* high tumour group
Lyn low
Lyn high
Absolute abundance (arbitrary units)
1500
1000
500
0
-500
NS
B cells naive
B cells memory
Plasma cells
T cells CD8
T cells CD4 naive
T cells CD4 memory resting
T cells CD4 memory activated
T cells follicular helper
T cells regulatory (Tregs)
T cells gamma delta
NK cells resting
NK cells activated
Monocytes
Macrophages M0
Macrophages M1
Macrophages M2
Dendritic cells resting
Dendritic cells activated
Mast cells resting
Mast cells activated
Eosinophils
Neutrophils


PCA tumour group 3/4 vs tumour group 1/2
PCA 3/4
PCA 1/2
Absolute abundance (arbitrary units)
1500
1000
500
0
-500
NS
**
B cells naive
B cells memory
Plasma cells
T cells CD8
T cells CD4 naive
T cells CD4 memory resting
T cells CD4 memory activated
T cells follicular helper
T cells regulatory (Tregs)
T cells gamma delta
NK cells resting
NK cells activated
Monocytes
Macrophages M0
Macrophages M1
Macrophages M2
Dendritic cells resting
Dendritic cells activated
Mast cells resting
Mast cells activated
Eosinophils
Neutrophils