MYC regulation of Glutamine-Proline regulatory axis is key in Luminal B breast cancer

Madeleine L. Craze[a], Hayley Cheung[a], Natasha Jewa[a], Nuno D.M. Coimbra[b], Daniele Soria[c], Rokaya El-Ansari[a], Mohammed A. Aleskandarany[a], Kiu Wai Cheng[a], Maria Diez-Rodriguez[a], Christopher C. Nolan[a], Ian O. Ellis[a,d], Emad A. Rakha[a,d], Andrew R. Green[a]

[a]Academic Pathology, Division of Cancer and Stem Cells, School of Medicine, University of Nottingham, Nottingham City Hospital, Hucknall Road, Nottingham NG5 1PB

[b]Department of Pathology, Instituto Português de Oncologia do Porto FG, Porto, Portugal

[c]Department of Computer Science, University of Westminster, New Cavendish Street, London W1W 6UW

[d]Cellular Pathology, Nottingham University Hospitals NHS Trust, Hucknall Road, Nottingham NG5 1PB

Correspondence:

Dr Andrew R. Green. Academic Pathology, Division of Cancer and Stem Cells, School of Medicine, University of Nottingham, Nottingham City Hospital, Hucknall Road, Nottingham NG5 1PB

Tel: (44) 115 8231407, Email: andrew.green@nottingham.ac.uk

Running title: Glutamine-Proline regulatory axis in breast cancer

Conflict of interest: The authors confirm that they do not have any conflict of interests to declare.

ABSTRACT

Background: Altered cellular metabolism is a hallmark of cancer and some are reliant on Glutamine for sustained proliferation and survival. We hypothesise that the Glutamine-Proline regulatory axis has a key role in Breast cancer (BC) in the highly proliferative classes.

Methods: Glutaminase (GLS), pyrroline-5-carboxylate synthetase (ALDH18A1) and pyrroline-5-carboxylate reductase 1 (PYCR1) were assessed at DNA/mRNA/protein levels in large well-characterised cohorts.

Results: Gain of PYCR1 copy number and high PYCR1 mRNA was associated with luminal B tumours. High ALDH18A1 and high GLS protein expression was observed in the ER+/HER2- high proliferation class (Luminal B) compared with ER+/HER2- low proliferation class (Luminal A) ($p=0.030$ and $p=0.022$ respectively), however this was not observed with mRNA. Cluster analysis of the Glutamine-Proline regulatory axis genes revealed significant associations with molecular subtypes of breast cancer and patient outcome independent of standard clinicopathological parameters ($p=0.012$). High protein expression of the Glutamine-Proline enzymes were all associated with high MYC protein in Luminal B tumours only ($p<0.001$).

Conclusion: We provide comprehensive clinical data indicating that the Glutamine-Proline regulatory axis plays an important role in the aggressive subclass of luminal BC and is therefore a potential therapeutic target.

Key words: metabolism, breast cancer, prognosis, Luminal B, glutamine, proline

1. Introduction

Deregulation of metabolic pathways has been readily accepted as part of the revised hallmarks of cancer where cancer cells are able to regulate their metabolism to provide energy and cellular building blocks required for growth (Hanahan & Weinberg, 2011). Many cancer cells are highly reliant on amino acids for their growth where endogenous synthesis may not provide the rapidly proliferating cells with sufficient nutrients for nuclear biosynthesis. There is also increasing evidence that oncogenes and/or tumour-suppressor genes can reprogram tumour cell metabolism including the direct regulation of the glutamine (Gln)-proline (Pro) regulatory axis by MYC and p53 (Cao *et al*, 2014; Kardos *et al*, 2015; Wise & Thompson, 2010) . This axis is the most important metabolic pathway in tumours after glucose primarily as Gln is used to replenish the tricarboxylic-acid (TCA) cycle and supplies carbon and nitrogen for synthesis of nucleotides, amino acids and glutathione. Indeed, some solid tumours have glutamine dependent cell growth or "glutamine addiction" (Wise & Thompson, 2010).

Gln is a non-essential amino acid synthesised by glutamine synthetase (GS) from glutamate and ammonia. Its utilisation, via reductive carboxylation, is necessary for sustained proliferation/survival and is linked with resistance to certain drugs (Soria *et al*, 2010). A further role for Gln in cancer cell protein translation stems from observations that a master regulator of protein translation, rapamycin complex 1 (mTORC1), which regulates cell growth and protein translation, is also responsive to Gln levels (McShane *et al*, 2005; Phang *et al*, 2015). In breast cancer (BC), high-grade highly proliferative tumours such as the triple negative (TN), have higher levels of glutamate and glutaminase (GLS) together with low levels of Gln than low grade tumours and normal breast epithelium (Gao *et al*, 2009; Kuo *et al*, 2016; Liu *et al*, 2012; Richardson *et al*, 2008; Wise *et al*, 2008). Metabolic profiles of BC show glutaminolysis metabolism as a key pathway discriminating between TN and oestrogen receptor (ER)+ tumours (Cao *et al*, 2014). The CB-839 small-molecule selective inhibitor of GLS which has anti-tumour activity in TNBC cell lines is currently being tested in phase I clinical trials (NCT02071862) (Gao *et al*, 2009).

Gln is converted to glutamate by GLS before entering the TCA cycle as a precursor to α-ketoglutarate, an important energy source which is synthesised by glutamate dehydrogenase. However, glutamate can also be converted into proline via the enzymes pyrroline-5-carboxylate synthetase (ALDH18A1) and pyrroline-5-carboxylate reductase 1 (PYCR1) which subsequently produces NADP used to fuel the breakdown of glucose by glycolysis (Kardos *et al*, 2015) (Figure 1). ALDH18A1 is responsible for reducing glutamate to Gln semialdehyde; a crucial step in the *de novo* biosynthesis of proline (Database, 2016).

With renewed interest in oncometabolism, metabolic enzymes are increasingly targeted to improve therapeutic efficacy and reduce resistance. We therefore hypothesised that the Pro-Gln axis is a key metabolic pathway regulated by MYC in BC particularly as we have recently showed proline dehydrogenase (PRODH) to be down regulated (Kardos *et al*, 2015). This pathway could be used as a potential therapeutic target particularly as the pleotropic MYC has so far proved ineffective.

The aim of this study was to conduct a comprehensive genomic, transcriptomic and proteomic analysis of the Gln-Pro regulatory axis in the molecular subtypes of breast cancer and their associations with MYC. More specifically, we aim to determine GLS, ALDH18A1 and PYCR1 gene copy number and gene expression; together with its protein expression, in large well-characterised annotated cohorts of BC to determine their biological, clinicopathological and prognostic value in the different molecular classes with particular interest in the highly proliferative aggressive subgroups as potential therapeutic targets.

2. METHODS

2.1 *Gln-Pro* enzymes copy number and gene expression

PYCR1, ALDH18A1, GLS and MYC copy number and gene expression were evaluated in a cohort of 1,980 breast cancer samples using the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) cohort (Curtis *et al*, 2012). METABRIC provides data on genomic and transcriptomic profiling of breast cancer using the Affymetrix SNP 6.0 and Illumina HT-12 v3 platforms respectively. In addition, *TP53* mutational profiling was performed. Detailed description of the experimental assays and analytical methods used were described previously (Silwal-Pandit *et al,* 2015). In this cohort, patients with ER+ and/or lymph node negative tumours did not receive adjuvant chemotherapy, whilst those with ER- and/or lymph node positive tumours received adjuvant chemotherapy. Breast Cancer Specific Survival (BCSS) is defined as the time (in months) from the date of primary surgery to the date of BC-related death.

2.2 *Gln-Pro enzymes protein expression*

Immunohistochemistry was conducted for PYCR1, ALDH18A1, and GLS using a large cohort of patients comprising a well-characterised consecutive series of early stage (TNM Stage I-III excluding T3 and T4 tumours) sporadic primary operable invasive BC. Patients (age ≤70 years) were enrolled into the Nottingham Tenovus Primary Breast Carcinoma Series, presented at Nottingham City Hospital between 1989 and 1998 (n=1,837) and managed in accordance to uniform protocol. Patients' clinical history, tumour characteristics, information on therapy and outcomes are prospectively maintained. Outcome data was collected on a prospective basis.

Primary antibody specificity (ALDH18A1 1:250 (HPA012604, Sigma-Aldrich, UK), GLS 1:1000 (EP7212, AbCam, UK) PYCR1 1:250 (HPA047660, Sigma-Aldrich, UK) was validated by western blotting (WB) using MCF7, MDA-MB-231 and SKBR3 human BC cell lines (American Type Culture Collection; Rockville, MD, USA). Proteins were detected using IRDye 800CW and 680RD fluorescent secondary antibodies (1:15000 dilution. 926-32213 and 926-68072, LI-COR Biosciences) and visualised using the Odyssey Fc with Image Studio 4.0 (LI-COR Biosciences). Anti-β-actin primary antibody (Sigma-Aldrich) was used as a loading control

(1:5000). Specific bands were observed at the correct molecular weights for ALDH18A1 (87kDa), GLS (73kDa, 65kDa) and PYCR1 (36kDa). An additional 55kDa band was observed for GLS and a 28kDa band for PYCR1 which represent alternate isoforms of the proteins.

2.3 *Tissue arrays and Immunohistochemistry*

Tumour samples, 0.6mm cores, were arrayed as previously described (Abd El-Rehim *et al*, 2005). Immunohistochemical staining was performed on 4 μm thick sections using the Novolink polymer detection system (Leica Biosystems, RE7150-K), and following the manufacturer's protocol. Primary antibodies for PYCR1, ALDH18A1, and GLS were diluted at 1:50 in Leica antibody diluent (RE7133). Negative (omission of primary antibody) and positive controls were included according to manufacturer's data sheet.

Stained TMA sections were scored using high resolution digital images (NanoZoomer; Hamamatsu Photonics, Welwyn Garden City, UK) at x20 magnification. Staining was assessed using the semi-quantitative, modified histochemical score (H-score) which evaluates both the intensity of staining and the percentage of stained cells resulting in a final score of 0-300 (Dang, 2012). For intensity, a score index of 0, 1, 2 and 3 corresponding to negative, weak, moderate and strong was used and the percentage of positive cells for each intensity was estimated subjectively. Dichotomisation of PYCR1, ALDH18A1, and GLS protein expression was determined using the median H-score. All cores were scored by NJ or HC and a pathologist (MA), blinded from the scores and the clinical data, scored 10% of cores for inter-observer concordance. There was high inter and intra-observer concordance between the scorers (Kappa score ≥0.6).

Immunohistochemical staining and dichotomisation of other biomarkers included in this study were as per previous publications (Green *et al*, 2016). ER and PgR positivity was defined as ≥1% staining. Immunoreactivity of HER2 in TMA cores was scored using standard HercepTest guidelines (Dako). Chromogenic in situ Hybridisation (CISH) was used to quantify HER2 gene amplification in borderline cases using the HER2 FISH pharmDx™ plus HER2 CISH pharmDx™ kit (Dako) and was assessed according to the American Society of Clinical Oncology guidelines. BC molecular subtypes were defined based on the IHC profile as: Luminal A:

ER+/HER2- Low Proliferation (Ki67<10%), Luminal B: ER+/HER2- High Proliferation (Ki67≥10%), HER2-positive class: HER2+ regardless of ER status, Triple Negative: ER-, PgR- and HER2-. Basal phenotype was defined as those tumours expressing cytokeratin (Ck) 5/6, and/or Ck14 and/or Ck17.

2.4 *Cluster analysis*

The partitioning around medoids (PAM) algorithm (also known as k-medoids algorithm) was used to cluster tumours based on gene and protein expression of the Gln-Pro enzymes as previously described (Soria *et al*, 2010). A number of cluster validity indices were used to determine the best number of clusters as the explicit input parameter to the PAM algorithm (Soria *et al*, 2010).

2.5 *Statistical analysis*

Statistical analysis was performed using SPSS 21.0 statistical software (SPSS Inc., Chicago, IL, USA). Univariate and multivariate analyses were performed by chi-squared test, Log rank and Cox regression analysis, respectively. One way ANOVA (Tukey) and Spearman's Correlation coefficient were used for continuous data. Survival curves were analysed by Kaplan-Meier (Kaplan and Meier, 1958). A p-value <0.05 was considered significant. This study complied with reporting recommendations for tumour marker prognostic studies (REMARK) criteria (McShane *et al*, 2005).

2.6 *Ethics*

This study was approved by the Nottingham Research Ethics Committee 2 under the title 'Development of a molecular genetic classification of breast cancer'. All samples from Nottingham used in this study were pseudo-anonymised and collected prior to 2006 and therefore under the Human Tissue Act informed patient consent was not needed. Release of data was also pseudo-anonymised as per Human Tissue Act regulations.

3. Results

3.1 *Gln-Pro regulatory axis expression in breast cancer*

Gain of *PYCR1* copy number primarily occurred in Luminal B tumours with 116/257 (45%) of all gains occurring (Table 1, $p<0.001$). Expression of *PYCR1* mRNA was significantly higher in Luminal B tumours compared with Luminal A tumours, and high expression was also observed with HER2+ and Basal/TNBC subtypes (Figure 2a, $p<0.001$). However there was no association observed between PYCR1 protein expression and molecular classes (Figure 2b, $p=0.118$).

Within the molecular classes, high *ALDH18A1* mRNA was associated with HER2+ tumours although there was no difference observed with mRNA levels between Luminal A and Luminal B tumours (Figure 2c, $p<0.001$). Copy number gain of ALDH18A1 was associated with Basal tumours (Table 1, $p<0.001$). In contrast, high ALDH18A1 protein expression was seen in the ER+/HER2- High proliferation class compared with ER+/HER2- Low proliferation tumours ($p=0.030$) however significance was not observed with all molecular classes (Figure 2d; $p=0.079$).

High copy number gain of *GLS* (Table 1), GLS mRNA (Figure 2e) and protein (Figure 2f) were all significantly associated with Basal/TNBC subtypes ($p<0.001$). Lower *GLS* mRNA expression was seen in Luminal B tumours compared with Luminal A and the other molecular subtypes (Figure 2e; $p<0.001$). In contrast higher protein expression of GLS was seen in ER+/HER2- High proliferation compared with those classified as ER+/HER2- Low proliferation (Figure 2f; $p=0.022$), where highest expression was seen in the TNBC tumours ($p<0.001$).

3.2 *Gln-Pro regulatory axis confers poor prognosis in Luminal B breast cancer*

Breast tumours were further clustered based on the comparison of several indices comparing the gene expression of the Gln-Pro enzymes. These were characterised as follows: Cluster 1 ($GLS^-/ALDH18A1^+/PYCR1^+$), Cluster 2 ($GLS^+/ALDH18A1^+/PYCR1^+$) and Cluster 3 ($GLS^+/ALDH18A1^-/PYCR1^-$) (Figure 3a). The clusters were significantly associated with molecular subtypes of breast cancer where Cluster 1 were predominately Luminal B

and HER2+ tumours, Cluster 2 were associated with Basal/HER2+ tumours and Cluster 3 were primarily Luminal A and Normal subtypes (Table 2). With respect to patient outcome, Cluster 1 tumours had the worst survival compared with Cluster 2 which had moderate outcome and Cluster 3 which showed the best survival (Figure 3b; $p<0.001$). In Cox Regression, Gln-Pro PAM Clusters remained independent of the standard clinicopathological parameters in predicting patient survival (Table 3; $p=0.012$).

We looked to replicate the clusters derived from the mRNA expression using the protein expression of the Gln-Pro enyzmes (using the median H-scores) which similarly showed that Cluster 1 ($GLS^-/ALDH18A1^+/PYCR1^+$) was strongly associated with ER+/HER2- High Proliferation tumours, whereas Cluster 2 ($GLS^+/ALDH18A1^+/PYCR1^+$) and Cluster 3 ($GLS^+/ALDH18A1^-/PYCR1^-$) were associated with ER+/HER2- Low Proliferation and TNBC/HER2+ tumours respectively (Table 2; $p=0.001$). In terms of patient outcome, Cluster 1 tumours showed the worst survival compared with Clusters 2 and 3 (Figure 3c; $p=0.0003$). In Cox Regression analysis, the Glu-Pro Clusters remained independent of tumour grade, lymph node stage, tumour size, ER and HER2 status (Table 3; $p=0.001$).

3.5 *c-MYC is associated with high Gln-Pro enzymes in Luminal B tumours*

MYC mRNA was negatively correlated with *PCYR1* ($p=0.01$) and *ALDH18A1* ($p<0.001$) in all breast tumours, but not *GLS* (Table 4). In specific subtypes, *MYC* was positively correlated with *PYCR1* mRNA in Luminal B tumours ($p=0.006$) and negatively correlated in Luminal A tumours (Table 4; $p<0.001$). The only other correlations between *MYC* and the Gln-Pro genes was a negative correlation with *ALDH18A1* ($p<0.001$) in Luminal A ($p<0.001$) and HER2+ tumours ($p=0.002$).

In terms of protein expression, MYC protein was positively associated with the individual protein expression of GLS and ALDH18A1 in all breast cancers, but was only associated with all three Gln-Pro enzymes in ER+/HER2- High Proliferation tumours ($p<0.001$, Table 4).

4. Discussion

ER+/luminal tumours, which comprise the majority of BC (55-80%), remain a heterogeneous group in terms of molecular biology and patients' outcome. Despite the significant benefit of hormone therapy, a proportion of patients with luminal BC develop recurrences and die of their disease. There is therefore a clear need for improved understanding of the biology of the luminal class of BC, with subsequent translation into more effective methods of prognostic and predictive stratification of this most common form of BC.

Metabolic reprogramming in cancer including BC provides a vital role in the provision of supplementary elements including nutrients and energy which are essential for cellular growth. Tumour cells can become reliant on glutamine metabolism and become "addicted" to this amino acid for sustained proliferation/survival. Additionally, proline is often used as a precursor for other amino acids; therefore it is important in the synthesis of new proteins, as well as mediating redox signaling and protecting cells from oxidative stress (Kuo *et al*, 2016; Phang *et al*, 2015).

It has also been shown that there is a major metabolic shift towards *de novo* proline synthesis in metastatic breast tumours (Richardson *et al*, 2008). High levels of MYC are required to maintain this glutaminolytic phenotype which not only sees an increase in Gln cellular uptake but also regulates Gln-Pro enzymes (Gao *et al*, 2009; Wise *et al*, 2008). MYC promotes the conversion of Gln to Pro by nearly tenfold by upregulating ALDH18A1 (Hu *et al*, 2008; Liu *et al*, 2012).

Evidence of glutamine dependence in TNBC has previously been established (Cao *et al*, 2014; Gross *et al*, 2014; Hanahan & Weinberg, 2011; Lukey *et al*, 2016) but studies that address the prognostic significance of the key Gln-Pro enzymes in BC and their potential influence on Gln metabolism in the other molecular subtypes remains limited particularly in luminal BC. We have therefore investigated the expression of PYCR1, ALDH18A1 and GLS at the genomic, transcriptomic and proteomic level, and the impact of MYC on their expression, utilising a

large number of breast tumours in order to better understand the potential role of this regulatory axis in BC and its molecular subtypes, particularly in the luminal tumours.

Overexpression of ALDH18A1 increases proline levels and lowers reactive oxygen species, as well as increasing cell survival. The increase in proline has been linked to protect cells against hydrogen peroxide induced cell death, as well as carcinogenic stressors; therefore promoting tumour growth and proliferation (Krishnan *et al*, 2008). In line with this, we show that high ALDH18A1 confers a poor prognosis in Luminal B tumours. Inhibition of ALDH18A1 in melanoma significantly disrupts proline synthesis limiting cellular metabolism and decreasing cell viability, tumour growth and protein synthesis (Kardos *et al*, 2015; Liu *et al*, 2012). Knockdown or inhibition of ALDH18A1 in luminal B tumours could therefore potentially lead to a benefit in therapeutic targeting of this enzyme.

PYCR1 is frequently overexpressed in many cancer types (Kuo *et al*, 2016); in particular breast, pancreatic and ovarian cancer (Natarajan *et al*, 2012; Phang *et al*, 2015). In meta-analysis, *PYCR1* mRNA is highly expressed in the more aggressive BC subtypes, in addition to its protein in a small number of BC cases (Ding *et al*, 2017). Whilst we show similar results for mRNA in the METABRIC dataset, we were unable to confirm the translation into similarly high protein expression in TNBC and HER2+ BC nor its association with patient outcome. However, we do show that Luminal B have a high copy number gain of *PYCR1* and consequently supports a role in this poor prognostic group of tumours. In prostate cancer, knockdown of *PYCR1* results in the inhibition of cell proliferation via cell cycle arrest and enhanced apoptosis (Zeng *et al*, 2017) although the effect on breast cancer cells, particularly luminal B tumours remains to be determined.

Levels and activity of GLS are significantly increased in TNBC and HER2+ breast cancer cell lines due to their propensity for glutamine dependence (Gross *et al*, 2014; Lukey *et al*, 2016; Wang *et al*, 2010) which we here confirm *in vivo*. However, we further demonstrate that GLS predicts a better outcome in HER2+ tumours and

selective GLS inhibitors, such as CB-839 which blocks glutamine consumption and reduces subsequent glutamate-derived metabolic intermediates, is likely to have limited use in these patients (Gross *et al*, 2014).

Protein analysis also revealed that PYCR1 and ALDH18A1 are significantly associated with the MYC oncogene. This coincides with a plethora of studies which have shown that the MYC oncogene reprograms cell metabolism and causes an upregulation of genes involved in cancer (Li & Simon, 2013; Liu *et al*, 2012; Phang *et al*, 2013). Multiple studies have also supported the findings in this study, as the *MYC* oncogene has resulted in an increased expression of PYCR1 and ALDH18A1 to drive proline production; alongside fuelling the glutamine and MYC addiction that cancer cells exhibit (Li & Simon, 2013; Phang *et al*, 2015; Wise *et al*, 2008).

It is thought that the MYC gene interacts with these enzymes by amplifying the existing transcriptional signalling of the gene leading to an increase in mRNA. MYC then drives the translation of mRNA in order to increase protein levels of the enzymes (Li & Simon, 2013; Natarajan *et al*, 2012). Interestingly, studies by Wei Liu et al have shown that downregulating the MYC oncogene with siRNAs not only reduced the conversion of glutamine to proline but also resulted in reduced mRNA and protein expression of ALDH18A1 and PYCR1 (Liu *et al*, 2015; Liu *et al*, 2012). Silencing MYC results in a reduction of cell proliferation and induction of apoptosis; as well as decreasing levels of PYCR1 and ALDH18A1 directly (Dang, 2012; Liu *et al*, 2012) and targeting all three enzymes led to the substantial reduction in proline production (Liu *et al*, 2012).

Our observational findings of the Gln-Pro enzymes and their differing associations with MYC within the sub-groups of breast cancer further support the many functions of MYC within glutaminolysis. We have previously shown that MYC is potentially driving glucose metabolism in ER-negative tumours and translational function in ER-positive tumours (Green *et al*, 2016), and we further show in this study that MYC is likely driving the conversion of Gln to Pro within Luminal B tumours. Proline metabolism has a functional role in redox regulation and the conversion of Gln to Pro could be one mechanism where this balance can be achieved, especially in the

highly proliferative Luminal B tumours which show increased metabolic rates resulting in the accumulation of ROS (Benassi *et al*, 2006).

In this study, we have shown that the Gln-Pro enzymes are highly expressed in a subset of ER+ tumours that have high proliferation, i.e. Luminal B tumours, and are related with poor patient outcome in this group, suggesting that expression of these enzymes are regulated and driven by alternative mechanisms within the different molecular subtypes of breast cancer and hence the pathways utilised for cell metabolism will also differ It was also interesting to observe the difference in gene and protein expression of these enzymes within the ER+/HER2- low proliferation tumours and TNBC/HER2+ tumours. This could have arisen due to the definitions used between mRNA (PAM50) and protein (ER plus Ki67 expression). It might also reflect the differences in transcription and translational modifications.

The high expression of the Gln-Pro regulatory axis in Luminal B tumours is perhaps not unsurprising as they will have heavier demands of nutrients and energy essential for cell survival and proliferation compared with Luminal A tumours and it is known that tumours will alter their metabolic profiles to meet the needs for their growth and proliferation. This is supported by a study carried out by Kim et al in which differential expression patterns of glutamine metabolism-related proteins were identified according to the molecular subtype of breast cancer, with HER2+ tumours displaying highest glutamine metabolic activity and higher MYC amplification and Luminal B tumours displaying higher glutamine metabolic activity than Luminal A tumours (Kim *et al*, 2013).
We further show that within the highly proliferative luminal B tumours, high expression of the Gln-Pro enzymes are all significantly associated with MYC suggesting that it is the driving force behind the metabolic status for this subclass of breast cancer.

Therefore, we believe that continued refinement in understanding of the biological diversity of BC, particularly the Luminal B class, with linked development of classification strategies suitable for routine clinical use are essential to achieve a personalised approach to BC management. Further investigation of Gln metabolism using a

variety of *in vitro* functional assays are therefore essential to assess its the potential therapeutic value in the highly proliferative aggressive subclass of ER+ Luminal B BC.

Acknowledgments

We thank the Nottingham Health Science Biobank and Breast Cancer Now Tissue Bank for the provision of tissue samples. We thank the University of Nottingham (Nottingham Life Cycle 6 and Cancer Research Priority Area) for funding.

Author contributions

ARG **conceived and designed study**

MLC, HC, NJ, NDMC, KWC, MAA, RE, MDR, CCN, IOE, EAR, ARG **carried out experiments and collected data**

MLC, HC, NJ, KWC, DS, ARG **analysed data**

All authors were involved in writing the paper and had final approval of the submitted and published versions.

Titles and Legends to Figures

Figure 1. Schmatic representation of the enzymes involved in the glutamine regulatory axis

Figure 2. Gln-Pro enzyme mRNA expression in breast cancer molecular subtypes

Figure 3. Gln-Pro clusters in breast cancer. Boxplots for a) mRNA b) protein. Patient outcome c) mRNA d) protein

References

Abd El-Rehim DM, Ball G, Pinder SE, Rakha E, Paish C, Robertson JF, Macmillan D, Blamey RW, Ellis IO (2005) High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. *International journal of cancer* **116**(3)**:** 340-50

Benassi B, Fanciulli M, Fiorentino F, Porrello A, Chiorino G, Loda M, Zupi G, Biroccio A (2006) c-Myc phosphorylation is required for cellular response to oxidative stress. *Molecular cell* **21**(4)**:** 509-19

Cao MD, Lamichhane S, Lundgren S, Bofin A, Fjosne H, Giskeodegard GF, Bathen TF (2014) Metabolic characterization of triple negative breast cancer. *BMC cancer* **14:** 941

Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, Graf S, Ha G, Haffari G, Bashashati A, Russell R, McKinney S, Group M, Langerod A, Green A, Provenzano E, Wishart G, Pinder S, Watson P, Markowetz F, Murphy L, Ellis I, Purushotham A, Borresen-Dale AL, Brenton JD, Tavare S, Caldas C, Aparicio S (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**(7403)**:** 346-52

Dang CV (2012) MYC on the path to cancer. *Cell* **149**(1)**:** 22-35

Database GHG (2016) ALDH18A1 Gene - GeneCards | P5CS Protein | P5CS Antibody

Ding J, Kuo ML, Su L, Xue L, Luh F, Zhang H, Wang J, Lin TG, Zhang K, Chu P, Zheng S, Liu X, Yen Y (2017) Human mitochondrial pyrroline-5-carboxylate reductase 1 promotes invasiveness and impacts survival in breast cancers. *Carcinogenesis*

Gao P, Tchernyshyov I, Chang TC, Lee YS, Kita K, Ochi T, Zeller KI, De Marzo AM, Van Eyk JE, Mendell JT, Dang CV (2009) c-Myc suppression of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism. *Nature* **458**(7239)**:** 762-5

Green AR, Aleskandarany MA, Agarwal D, Elsheikh S, Nolan CC, Diez-Rodriguez M, Macmillan RD, Ball GR, Caldas C, Madhusudan S, Ellis IO, Rakha EA (2016) MYC functions are specific in biological subtypes of breast cancer and confers resistance to endocrine therapy in luminal tumours. *British journal of cancer* **114**(8)**:** 917-28

Gross MI, Demo SD, Dennison JB, Chen L, Chernov-Rogan T, Goyal B, Janes JR, Laidig GJ, Lewis ER, Li J, Mackinnon AL, Parlati F, Rodriguez ML, Shwonek PJ, Sjogren EB, Stanton TF, Wang T, Yang J, Zhao F, Bennett MK

(2014) Antitumor activity of the glutaminase inhibitor CB-839 in triple-negative breast cancer. *Molecular cancer therapeutics* **13**(4)**:** 890-901

Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. *Cell* **144**(5)**:** 646-74

Hu CA, Khalil S, Zhaorigetu S, Liu Z, Tyler M, Wan G, Valle D (2008) Human Delta1-pyrroline-5-carboxylate synthase: function and regulation. *Amino acids* **35**(4)**:** 665-72

Kardos GR, Wastyk HC, Robertson GP (2015) Disruption of Proline Synthesis in Melanoma Inhibits Protein Production Mediated by the GCN2 Pathway. *Molecular cancer research : MCR* **13**(10)**:** 1408-20

Kim S, Kim DH, Jung WH, Koo JS (2013) Expression of glutamine metabolism-related proteins according to molecular subtype of breast cancer. *Endocrine-related cancer* **20**(3)**:** 339-48

Krishnan N, Dickman MB, Becker DF (2008) Proline modulates the intracellular redox environment and protects mammalian cells against oxidative stress. *Free radical biology & medicine* **44**(4)**:** 671-81

Kuo ML, Lee MB, Tang M, den Besten W, Hu S, Sweredoski MJ, Hess S, Chou CM, Changou CA, Su M, Jia W, Su L, Yen Y (2016) PYCR1 and PYCR2 Interact and Collaborate with RRM2B to Protect Cells from Overt Oxidative Stress. *Scientific reports* **6:** 18846

Li B, Simon MC (2013) Molecular Pathways: Targeting MYC-induced metabolic reprogramming and oncogenic stress in cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **19**(21)**:** 5835-41

Liu W, Hancock CN, Fischer JW, Harman M, Phang JM (2015) Proline biosynthesis augments tumor cell growth and aerobic glycolysis: involvement of pyridine nucleotides. *Scientific reports* **5:** 17206

Liu W, Le A, Hancock C, Lane AN, Dang CV, Fan TW, Phang JM (2012) Reprogramming of proline and glutamine metabolism contributes to the proliferative and metabolic responses regulated by oncogenic transcription factor c-MYC. *Proceedings of the National Academy of Sciences of the United States of America* **109**(23)**:** 8983-8

Lukey MJ, Greene KS, Erickson JW, Wilson KF, Cerione RA (2016) The oncogenic transcription factor c-Jun regulates glutaminase expression and sensitizes cells to glutaminase-targeted therapy. *Nature communications* **7:** 11321

McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, Statistics Subcommittee of the NCIEWGoCD (2005) REporting recommendations for tumour MARKer prognostic studies (REMARK). *British journal of cancer* **93**(4)**:** 387-91

Natarajan SK, Zhu W, Liang X, Zhang L, Demers AJ, Zimmerman MC, Simpson MA, Becker DF (2012) Proline dehydrogenase is essential for proline protection against hydrogen peroxide induced cell death. *Free Radic Biol Med* **53**(5)**:** 1181-91

Phang JM, Liu W, Hancock C (2013) Bridging epigenetics and metabolism: role of non-essential amino acids. *Epigenetics* **8**(3)**:** 231-6

Phang JM, Liu W, Hancock CN, Fischer JW (2015) Proline metabolism and cancer: emerging links to glutamine and collagen. *Current opinion in clinical nutrition and metabolic care* **18**(1)**:** 71-7

Richardson AD, Yang C, Osterman A, Smith JW (2008) Central carbon metabolism in the progression of mammary carcinoma. *Breast cancer research and treatment* **110**(2)**:** 297-307

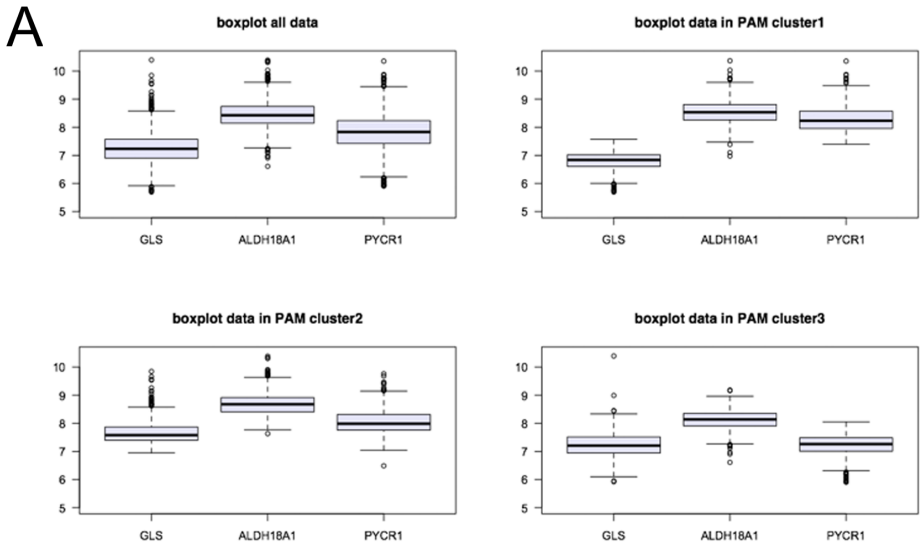
Soria D, Garibaldi JM, Ambrogi F, Green AR, Powe D, Rakha E, Macmillan RD, Blamey RW, Ball G, Lisboa PJ, Etchells TA, Boracchi P, Biganzoli E, Ellis IO (2010) A methodology to identify consensus classes from clustering algorithms applied to immunohistochemical data from breast cancer patients. *Computers in biology and medicine* **40**(3)**:** 318-30

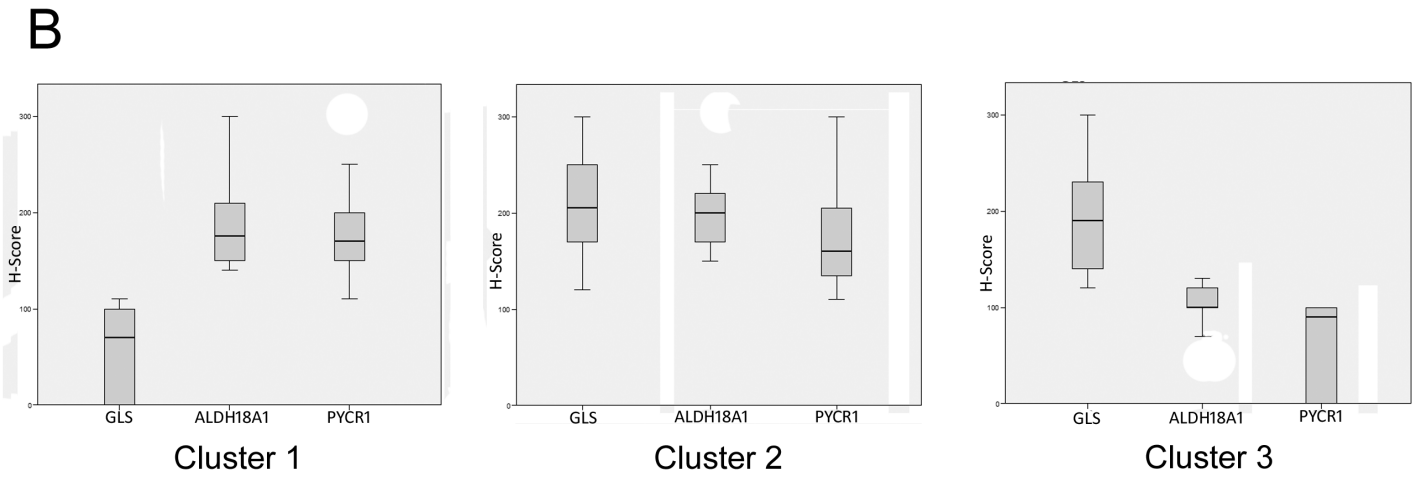
Wang JB, Erickson JW, Fuji R, Ramachandran S, Gao P, Dinavahi R, Wilson KF, Ambrosio AL, Dias SM, Dang CV, Cerione RA (2010) Targeting mitochondrial glutaminase activity inhibits oncogenic transformation. *Cancer cell* **18**(3)**:** 207-19

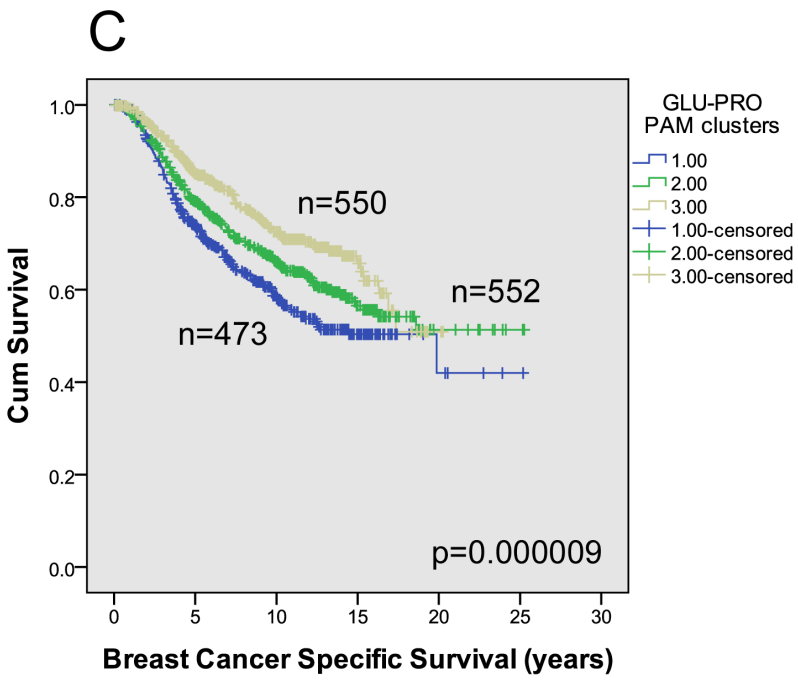
Wise DR, DeBerardinis RJ, Mancuso A, Sayed N, Zhang XY, Pfeiffer HK, Nissim I, Daikhin E, Yudkoff M, McMahon SB, Thompson CB (2008) Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction. *Proceedings of the National Academy of Sciences of the United States of America* **105**(48)**:** 18782-7

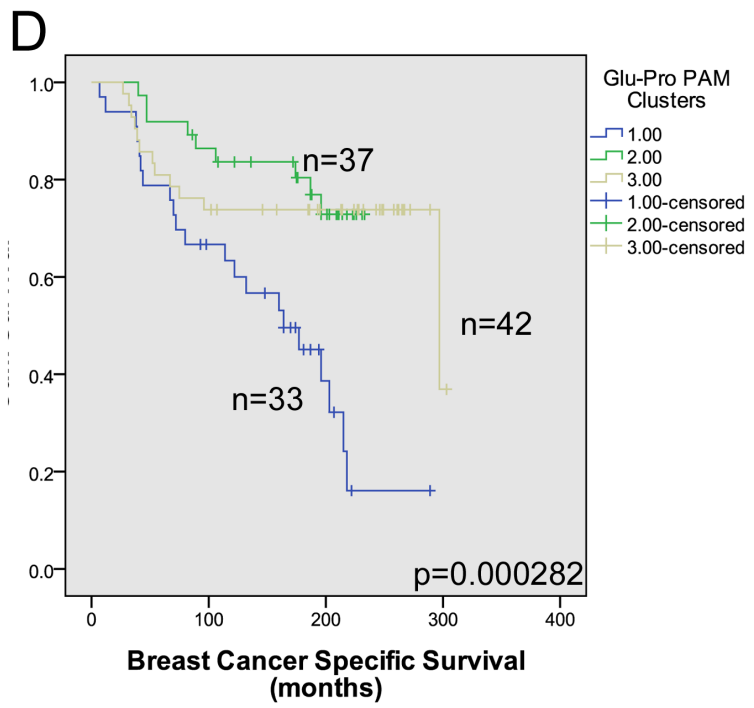
Wise DR, Thompson CB (2010) Glutamine addiction: a new therapeutic target in cancer. *Trends in biochemical sciences* **35**(8)**:** 427-33

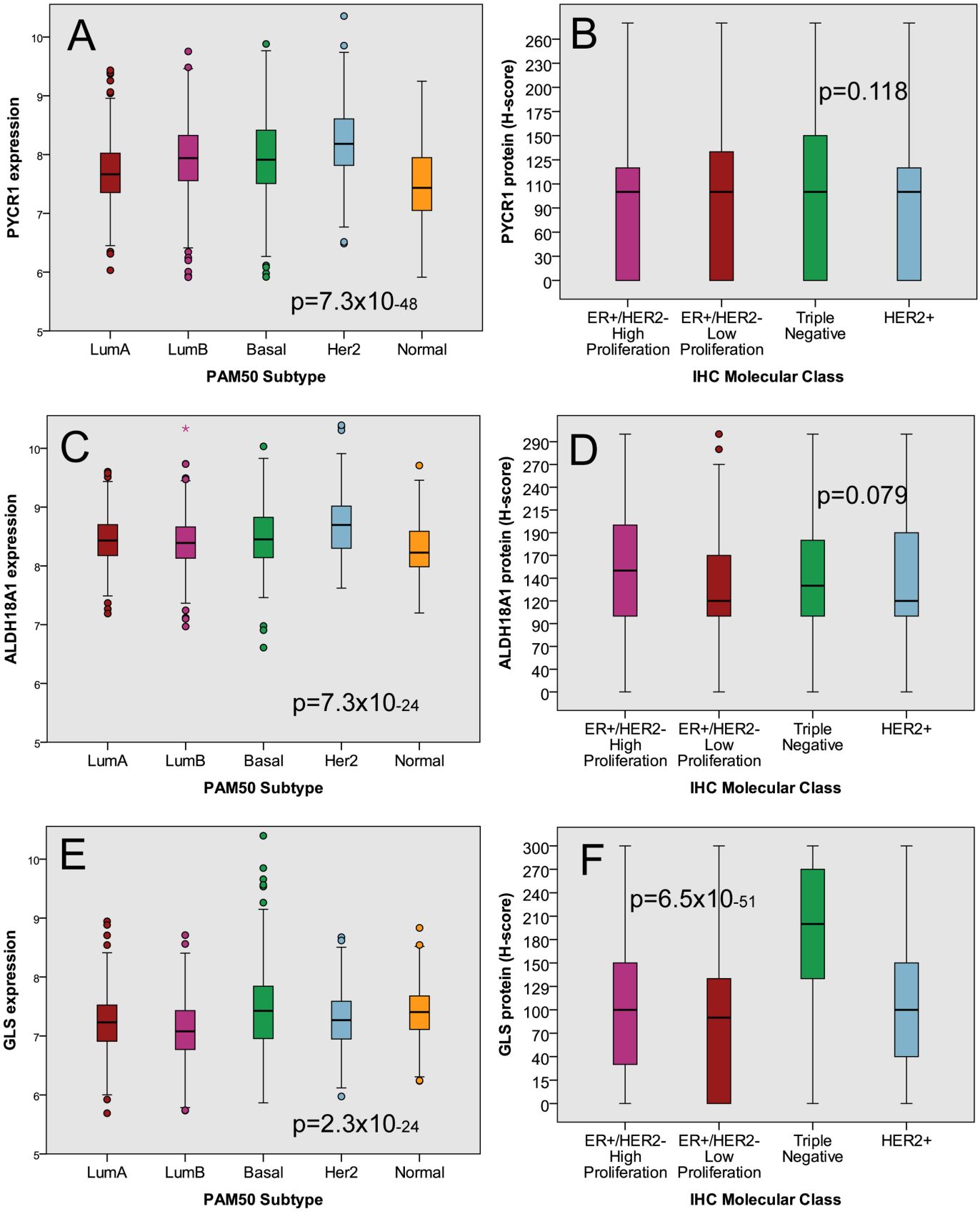
Zeng T, Zhu L, Liao M, Zhuo W, Yang S, Wu W, Wang D (2017) Knockdown of PYCR1 inhibits cell proliferation and colony formation via cell cycle arrest and apoptosis in prostate cancer. *Medical oncology* **34**(2)**:** 27

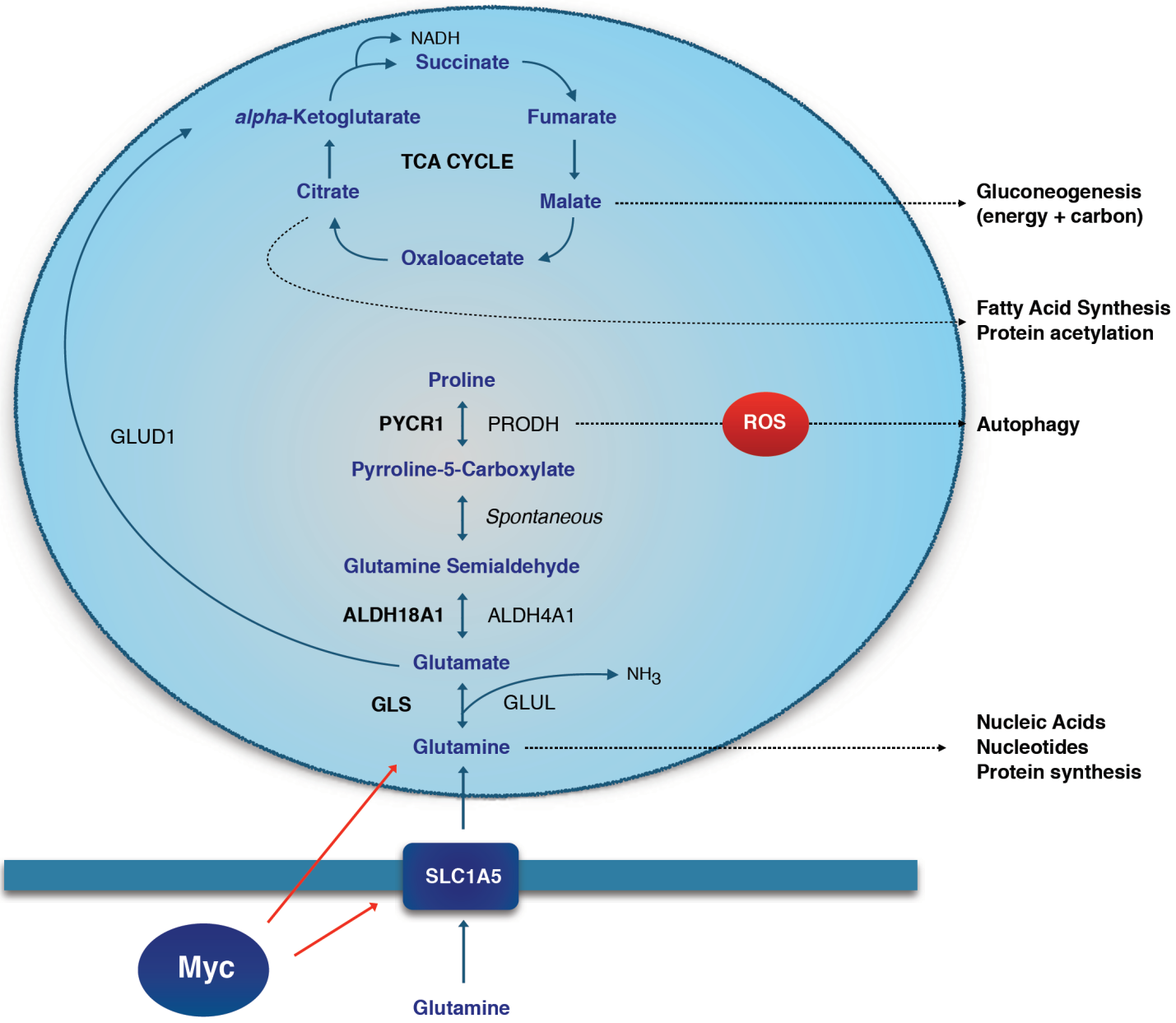

A
boxplot all data
boxplot data in PAM cluster1
boxplot data in PAM cluster2
boxplot data in PAM cluster3
GLS
ALDH18A1
PYCR1


B
H-Score
GLS
ALDH18A1
PYCR1
Cluster 1
Cluster 2
Cluster 3


C
Cum Survival
n=550
n=552
n=473
p=0.000009
Breast Cancer Specific Survival (years)
GLU-PRO PAM clusters
1.00
2.00
3.00
1.00-censored
2.00-censored
3.00-censored


D
n=37
n=42
n=33
p=0.000282
Breast Cancer Specific Survival (months)
Glu-Pro PAM Clusters
1.00
2.00
3.00
1.00-censored
2.00-censored
3.00-censored


A
PYCR1 expression
p=7.3x10-48
LumA
LumB
Basal
Her2
Normal
PAM50 Subtype
B
PYCR1 protein (H-score)
p=0.118
ER+/HER2- High Proliferation
ER+/HER2- Low Proliferation
Triple Negative
HER2+
IHC Molecular Class
C
ALDH18A1 expression
p=7.3x10-24
PAM50 Subtype
D
ALDH18A1 protein (H-score)
p=0.079
IHC Molecular Class
E
GLS expression
p=2.3x10-24
PAM50 Subtype
F
GLS protein (H-score)
p=6.5x10-51
IHC Molecular Class


NADH
Succinate
alpha-Ketoglutarate
Fumarate
TCA CYCLE
Citrate
Malate
Oxaloacetate
Gluconeogenesis
(energy + carbon)
Fatty Acid Synthesis
Protein acetylation
Proline
PYCR1
PRODH
ROS
Autophagy
GLUD1
Pyrroline-5-Carboxylate
Spontaneous
Glutamine Semialdehyde
ALDH18A1
ALDH4A1
Glutamate
NH3
GLS
GLUL
Glutamine
Nucleic Acids
Nucleotides
Protein synthesis
SLC1A5
Myc
Glutamine

	PYCR1			ALDH18A1			GLS		
DNA Copy number aberrations	**None**	**Gain**	χ^2 **(p-value)**	**None**	**Loss**	χ^2 **(p-value)**	**None**	**Gain**	χ^2 **(p-value)**
PAM50									
Luminal A	649 (92.3)	54 (7.7)		704 (98.9)	8 (1.1)		708 (99.2)	6 (0.8)	
Luminal B	354 (75.3)	116 (24.7)		466 (96.1)	19 (3.9)		470 (98.3)	8 (1.7)	
Basal	272 (83.4)	54 (16.6)	89.3 (**$1.9x10^{-18}$**)	308 (94.2)	19 (5.8)	22.4 (**0.0002**)	303 (92.7)	24 (7.3)	49.5 (**$4.7x10^{-10}$**)
HER2	211 (89.8)	24 (10.2)		231 (97.9)	5 (2.1)		229 (95.8)	10 (4.2)	
Normal	184 (95.3)	9 (4.7)		196 (98.5)	3 (1.5)		196 (100)	0	

Table 1. Gln-Pro enzymes and breast cancer molecular subtypes

	Cluster 1 n (%)	Cluster 2 n (%)	Cluster 3 n (%)	χ^2 (p-value)
	$GLS^-/ALDH18A1^+/PYCR1^+$	**$GLS^+/ALDH18A1^+/PYCR1^+$**	**$GLS^+/ALDH18A1^-/PYCR1^-$**	
mRNA				
Luminal A	174 (24.2)	242 (33.7)	302 (42.1)	177.5 (**$3.7x10^{-34}$**)
Luminal B	214 (43.9)	133 (27.3)	141 (28.9)	
Basal	91 (27.7)	148 (45.0)	90 (27.4)	
HER2	95 (39.6)	110 (45.8)	35 (14.6)	
Normal	23 (11.6)	60 (30.2)	116 (58.3)	
Protein				
ER+/HER2- Low Proliferation	8 (34.8)	10 (43.5)	5 (21.7)	23.8 (**0.001**)
ER+/HER2- High Proliferation	19 (48.7)	8 (20.5)	12 (30.8)	
Triple Negative	1 (2.8)	13 (36.1)	22 (61.1)	
HER2+	4 (28.6)	4 (28.6)	6 (42.9)	

Table 2. Gln-Pro Clusters in breast cancer molecular subtypes

Table 3. Gln-Pro Clusters and patient outcome

Parameter	mRNA Hazard ratio (95% CI)	p-value	Protein Hazard ratio (95% CI)	p-value
Grade	1.24 (1.05-1.47)	**0.012**	1.22 (0.71-2.09)	0.466
Lymph Node Stage	1.83 (1.62-2.06)	**$3.9x10^{-23}$**	2.23 (1.24-4.03)	**0.008**
Size	1.58 (1.26-1.97)	**0.00008**	0.89 (0.46-1.72)	0.731
ER	1.12 (1.01-1.25)	**0.035**	1.07 (0.48-2.38)	0.878
HER2	0.84 (0.75-0.95)	**0.004**	1.79 (0.74-4.36)	0.198
Gln-Pro Clusters	0.86 (0.76-0.97)	**0.012**	0.46 (0.29-0.75)	**0.001**

	MYC		
	PYCR1	**ALDH18A1**	**GLS**
mRNA			
All cases	-0.058 (**0.010**)	-0.221 (**$2.3x10^{-23}$**)	0.038 (0.092)
PAM50			
Luminal A	-0.133 (**0.0003**)	-0.339 (**$9.8x10^{-21}$**)	-0.017 (0.659)
Luminal B	0.124 (**0.006**)	-0.070 (0.123)	0.052 (0.248)
Basal	0.094 (0.088)	0.023 (0.677)	-0.072 (0.193)
HER2	0.030 (0.646)	-0.201 (**0.002**)	0.026 (0.284)
Normal	-0.335 (**0.000001**)	-0.395 (**$7.5x10^{-9}$**)	-0.148 (**0.037**)
Protein			
All cases	-0.011 (0.772)	0.394 (**$1.8x10^{-33}$**)	0.148 (**0.00002**)
IHC molecular classes			
ER+/HER2- Low Proliferation	0.166 (**0.048**)	0.253 (**0.001**)	0.124 (0.111)
ER+/HER2- High Proliferation	0.156 (**0.013**)	0.221 (**0.0001**)	0.145 (**0.015**)
Triple Negative	0.149 (0.074)	0.378 (**$3.9x10^{-7}$**)	0.292 (**0.0003**)
HER2+	-0.060 (0.551)	0.492 (**$2.8x10^{-8}$**)	0.094 (0.329)

Table 4. Gln-Pro enzymes and MYC expression in breast cancer molecular subtypes