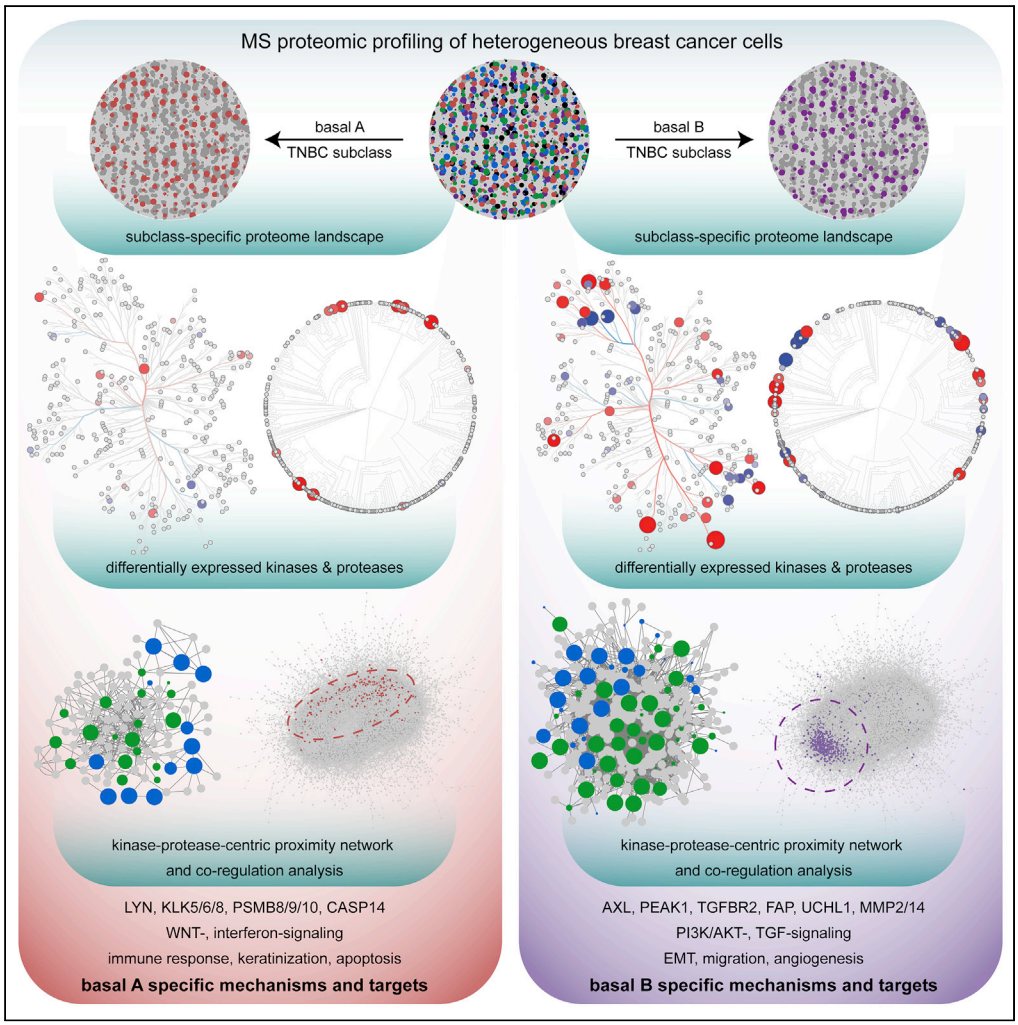
iScience

CellPress

Article

# Comprehensive Proteomic Characterization Reveals Subclass-Specific Molecular Aberrations within Triple-negative Breast Cancer




Max Kosok, Asfa Alli-Shaik, Boon Huat Bay, Jayantha Gunaratne

jayanthag@imcb.a-star.edu.sg

## HIGHLIGHTS

Proteome profiling reveals functionally distinct subclasses within TNBC

Kinases and proteases underlie unique functional signatures among the subclasses

Kinase-protease-centric networks highlight subclass-specific molecular rewiring

Protein association dysregulations reveal TNBC subclass-specific protein targets

## DATA AND CODE AVAILABILITY

PXD017025

Kosok et al., iScience *23*, 100868
February 21, 2020 
https://doi.org/10.1016/j.isci.2020.100868

iScience

CellPress

Article

# Comprehensive Proteomic Characterization Reveals Subclass-Specific Molecular Aberrations within Triple-negative Breast Cancer

Max Kosok,[1,2,3] Asfa Alli-Shaik,[1,3] Boon Huat Bay,[2] and Jayantha Gunaratne[1,2,4,*]

## SUMMARY

**Triple-negative breast cancer (TNBC) is the most aggressive subtype of breast cancer lacking targeted therapies. This is attributed to its high heterogeneity that complicates elucidation of its molecular aberrations. Here, we report identification of specific proteome expression profiles pertaining to two TNBC subclasses, basal A and basal B, through in-depth proteomics analysis of breast cancer cells. We observed that kinases and proteases displayed unique expression patterns within the subclasses. Systematic analyses of protein-protein interaction and co-regulation networks of these kinases and proteases unraveled dysregulated pathways and plausible targets for each TNBC subclass. Among these, we identified kinases AXL, PEAK1, and TGFBR2 and proteases FAP, UCHL1, and MMP2/14 as specific targets for basal B subclass, which represents the more aggressive TNBC cell lines. Our study highlights intricate mechanisms and distinct targets within TNBC and emphasizes that these have to be exploited in a subclass-specific manner rather than a one-for-all TNBC therapy.**

## INTRODUCTION

Triple-negative breast cancer (TNBC) that represents 10%–15% of all breast cancers is the most aggressive breast cancer subtype characterized by high recurrence and poor prognosis (Ovcaricek et al., 2010). Unlike receptor positive subtypes that include estrogen- and/or progesterone-receptor positive ($ER^+/PR^+$) and the human epidermal growth factor receptor 2 positive ($HER2^+$) subtypes, TNBC is a highly heterogeneous breast cancer subtype, making it challenging to understand its underlying disease mechanisms and design targeted treatments (Anders and Carey, 2009; Badve et al., 2011; Bauer et al., 2007). Several research groups have attempted subclassifying TNBC toward the mutual goal of finding targeted therapies. TNBC has been categorized under the basal-like category according to the PAM50 classification (Parker et al., 2009; Perou et al., 2000; Sørlie et al., 2001) and has been further divided into two distinct subclasses as basal A and B based on gene clustering (Neve et al., 2006). Another study based on gene expression analysis of TNBC tumors has identified six subcategories with distinct gene ontologies as two basal-like (BL1 and BL2), two mesenchymal (M and MSL), one immunomodulatory (IM), and one luminal androgen receptor subtype (LAR) (Lehmann et al., 2011). Apart from these, subclassification based on specific marker expression, such as claudin-low subtype characterized by low expression of claudin-3, -4, and -7 (Herschkowitz et al., 2007; Prat et al., 2010) or based on elevated expression of positive markers such as cytokeratins, mainly 5, 6, 13, and 17, have also been reported (Bianchini et al., 2016; Foulkes et al., 2010; Nielsen et al., 2004). The heterogeneous nature of TNBC is also accentuated at the protein level by large proteomic profiling of breast cancer tissues and cell lines (Lapek et al., 2017; Lawrence et al., 2015; Mertins et al., 2016). Although all these studies pointed to existence of subclasses within TNBC, their specific mechanistic differences or pathway dysregulations were not explored in detail.

With no current effective therapy, there has been a huge interest in deconvoluting the oncogenic events that lead to the development of TNBC, and several aberrantly activated pathways have been identified. Among them, TGF-β signaling is linked to aggressive behavior such as TNBC cell migration, early metastasis, as well as invasion and extravasation (Massagué, 2008; Padua et al., 2008; Papageorgis et al., 2010; Xu et al., 2018). Other signaling pathways mediated by neurotrophic tyrosine kinase receptor (TrkB), Erk/NF-κB, PI3K/AKT, and HIF-1α have also been associated with TNBC metastasis (Costa et al., 2018; Delaloge and DeForceville, 2017; Kuo et al., 2017; Neophytou et al., 2018; Ponente et al., 2017; Tsai et al., 2017; Wong et al., 2011). Despite several ongoing trials targeting such aberrant pathways, the clinical outcome has been disappointing with limited benefit in TNBC patients. This underscores the complexity of TNBC

[1]Institute of Molecular and Cell Biology, Agency for Science, Technology and Research, Singapore 138673, Singapore

[2]Department of Anatomy, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117594, Singapore

[3]These authors contributed equally

[4]Lead Contact

*Correspondence: jayanthag@imcb.a-star.edu.sg

https://doi.org/10.1016/j.isci.2020.100868

at the molecular level and emphasizes the need for elucidating TNBC subclass-specific mechanisms and pathways for effective personalized therapies.

Given the distinct genomic alterations and clinicopathological characteristics within different subtypes of TNBC (Bareche et al., 2018), we hypothesized that an in-depth analysis of subclass-specific proteome changes associated with pathway aberrations would reveal molecular vulnerabilities within TNBC, thus identifying novel therapeutic targets. To this end, we carried out comprehensive proteomic profiling of an array of breast cancer cell lines, employing multiplexed quantitative high-resolution mass spectrometry (MS)-based proteomics followed by extensive peptide fractionation to record comprehensive quantitative protein and peptide information across all cell lines. We interrogated the data extensively for network and pathway aberrations, particularly associated with perturbed functional protein classes. Through this strategized and systematic analysis, we identified TNBC inherent proteome-based subclasses, which were also reflected by distinct profiles of kinases and proteases, and their associated functional networks. Finally, using correlation networks, we inferred that dysregulations of kinases and proteases are centric in defining molecular traits unique to each breast cancer subtype and validated kinase AXL as an effective target only in the aggressive TNBC subclass. Altogether, this study provides in-depth proteomic profiles of TNBC cell line panel and reveals subclass-specific mechanisms and potential dysregulated targets, thus offering immense potential for clinical exploration.

## RESULTS

### In-depth Proteomic Profiling

We performed high-resolution MS-based deep proteomics profiling of 18 breast cancer cell lines (12 TNBC, 3 $ER^+/PR^+$, and 3 $HER2^+$) by multiplexed tandem mass tag (TMT) approach that allows reliable and robust quantification of proteins across the different cell lines. Two complimentary offline peptide fractionation methods, namely iso-electric focusing (IEF) and offline basic pH reverse-phase liquid chromatography (bRP-LC), were employed prior to MS analysis to obtain a comprehensive proteome by enhancing peptide coverage. Two biological replicates, accounting for 156 LC-MS runs, were included to ensure reproducibility across the panel. The overall workflow is depicted in Figure 1A. Protein quantification between the biological replicates as well as between the two fractionation approaches was highly correlated (Figure S1A), indicating good reproducibility of our MS dataset. In total, we identified a total of 233,910 peptides derived from 11,709 proteins at a false discovery rate (FDR) of 1%, of which 10,400 were quantified across all 18 cell lines (Figure 1B and Tables S1A and S1B), highlighting consistent quantification with few missing values. By capturing unique peptide clusters, the implementation of the two different fractionation approaches in parallel boosted the overall proteome coverage by 25% compared with the proteome profiles obtained using only either of the two approaches (Figure 1C). With an average of 22 peptides identified per protein, of which 16 peptides on an average were also isoform specific, the improved peptide coverage contributed to reliable protein identification and quantification. To our knowledge, this is the most comprehensive quantified proteomes of breast cancer cells, with the highest average number of unique peptides per protein reported to date. Notably, our dataset also reports the highest numbers for quantified proteins belonging to important functional classes such as transcription factors, kinases, proteases, and phosphatases that have not been captured and profiled extensively thus far by MS methods in the breast cancer landscape (Figure 1B). This high proteome coverage was obtained mainly because of our strategized peptide fractionations that enhanced detection capability of peptides by changing their chemical nature (high pH and isoelectric ambient).

### Proteome-Based Clustering

To evaluate the quality of our MS dataset, we mapped expression levels of known molecular markers to cell lines representing each breast cancer subtype. We observed high expression levels of ESR1 and PGR in all $ER^+/PR^+$ cells, high levels of ERBB2 in $HER2^+$ cells, and no/low expression of all these receptors among the TNBC cell lines, adding confidence to our proteome dataset (Figure 2A upper histogram). Next, we performed unsupervised clustering of the whole proteome expression data that revealed three main clusters with all receptor positive cell lines grouped into one cluster and the TNBC cell lines bifurcating into two distinct clusters (Figures 2A top tree and S1B). The two TNBC clusters corresponded to well-established TNBC classification based on claudin and cytokeratin expression patterns (Figure 2A middle/lower histogram). Our clustering also aligned with PAM gene expression classification (Neve classification) for breast cancer cell lines that is based on 305 genes (Neve et al., 2006). On comparing our proteome-based categories with the TNBCtype classification (Lehmann et al., 2011), we observed that mesenchymal

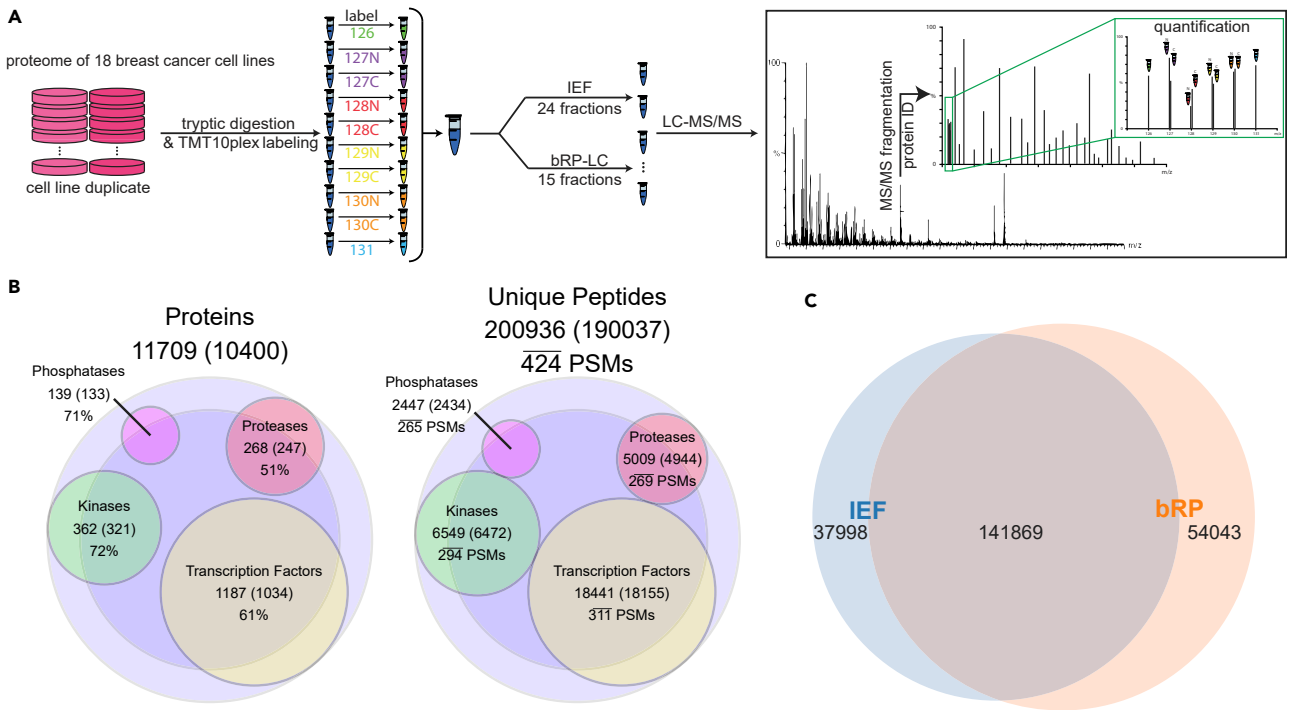



**Figure 1. Comprehensive MS-based Proteomics Analysis of Breast Cancer Cells**

(A) Schematic of the bottom-up proteome profiling pipeline for screening 18 breast cancer cell lines employing TMT multiplexing with two peptide fractionation strategies.

(B) Number of identified (and overall quantified) proteins, unique peptides, and average number of peptide spectrum matches (PSM) in total, along with their distribution in four functional protein classes. Percentage represents completeness of identification of each protein class.

(C) Total number of peptides identified in samples fractionated by IEF or bRP-LC.

(M) and mesenchymal stem-like (MSL) were representative of basal B, whereas basal-like 1 and 2 (BL1 and BL2) resembled basal A subclass (Figure 2A middle panel). The BT20 cell line, which could not be distinctly categorized by the TNBCtype classification, appeared to align with the basal A category in our analysis. A similar grouping of BT20 was also reported by Lawrence et al. (2015) in their label-free proteomics study that assigned them to a cluster resembling the BL1 category, thus validating our assignment. However, in contrast to observations by both our study and that of the TNBCtype classification, the MSL MB-436 cell line was categorized within a group most similar to the BL2 group by Lawrence et al. (2015). Furthermore, our analysis, for the first time, also distinctly assigned the uncategorized TNBC cell line HCC1395 to the basal B subclass based on the proteome pattern. We also compared our proteome profile data with published transcriptome data (Klijn et al., 2015) and noted modest correlation, consistent with previously reported observations (Figure 2A lower panel) (Lapek et al., 2017; Sacco et al., 2016; Zhang et al., 2014). Principal component analysis (PCA) of the whole proteome distinctly demarcated the cell lines into the luminal group with clear differentiation of $ER^+/PR^+$ and the $HER2^+$ groups and the TNBC group showing two subclasses with discrete patterning based on their proteome landscape (Figure 2B).

Next, we sought to ask whether sub-proteomes of major functional protein classes, which are important regulators of cellular signaling, also show similar clustering pattern as the whole proteome. For this, we considered selected functional groups that included kinases, proteases, phosphatases, and transcription factors. PCA of the 321 overall quantified kinases [72% of the human kinome (Manning et al., 2002)] revealed that the expression patterns of only these kinases were sufficient to reflect the same clustering pattern among the different breast cancer cell lines as that obtained using the whole proteome, iterating distinctive kinase expression patterns among the different breast cancer groups (Figure 2C). A similar clustering was also observed for the 247 proteases [51% of human proteases (Rawlings et al., 2018)], with only two cell lines (HCC1806 and MB-231) appearing as outliers based on their protease expression patterns (Figure 2D). However, PCA of 133 phosphatases [71% of human phosphatases (Liberti et al., 2012)] and

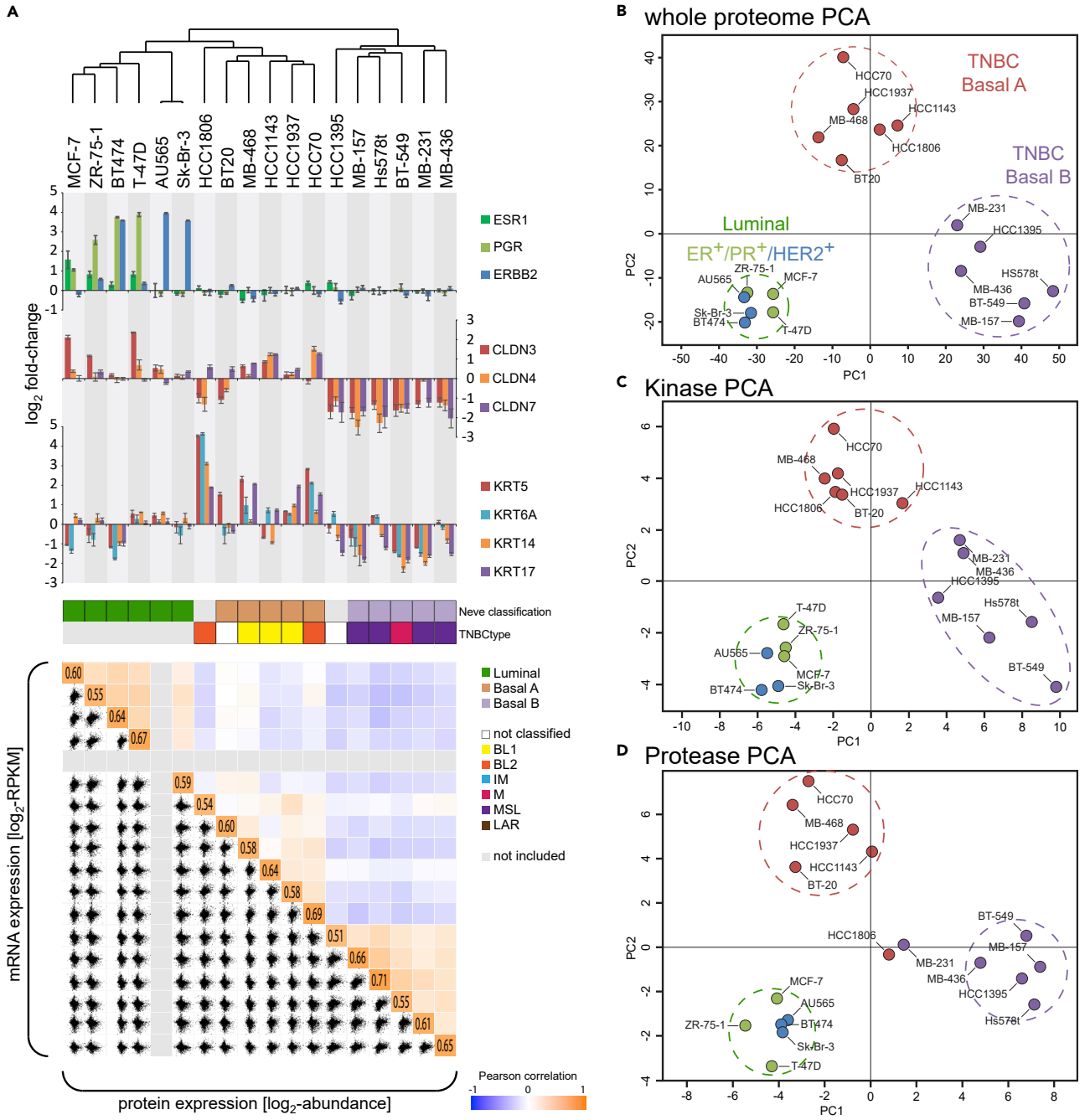



**Figure 2. Proteome Profiles of Breast Cancer Cell Lines**

(A) Clustering of the cell lines based on the whole proteome expression profiles is shown (top). Relative expression of $ER^+/PR^+$ and $HER2^+$ and known claudin and cytokeratin protein markers in breast cancer cell lines. Data are represented as mean ± SEM, n = 4 (upper panel). Comparison to TNBCtype and Neve classification of breast cancer cell lines based on gene expression studies is shown (middle panel). Correlation of protein abundance with mRNA expression is shown (lower panel). The color schema for the respective proteins/categories is shown on the right.

(B) Principal component analysis (PCA) of the whole proteome distinguished cell lines into the luminal subtype ($HER2^+$ and $ER^+/PR^+$) and TNBC subtype (basal A and basal B).

(C and D) (C) PCA based on only kinases or (D) proteases showed distinct patterns in cell line classifications.

1034 transcription factors [61% of human transcription factors (Ravasi et al., 2010)], which represents the largest protein class, showed no obvious distinctive clustering among the different subclasses as by the kinases or proteases (Figure S1C). These results suggested that kinases and proteases have differential signatures among the different subgroups and potentially play prominent roles in orchestrating unique mechanisms within the groups. Hence, in this study, we focused on a kinase- and protease-centric perspective to reveal subclass-specific aberrations.

### Subgroup-specific Kinase and Protease Expression

Among the 2,440 differentially expressed (DE) proteins (ANOVA p value $\leq$0.05 and two-sample t-test p value $\leq$0.05 for both pairwise comparisons), we observed 87 kinases and 62 proteases (Table S1A) that were exclusively differentially expressed in one group in comparison to the other two groups. Mapping these DE kinases onto the human kinome map (Manning et al., 2002) revealed that a majority of these belonged to tyrosine kinases, CAMK, and AGC kinase families (Figures 3A and S2B). We also compared these DE kinases with targets of approved drugs (www.drugbank.ca; Dec 2019, Figure S2A) and found that a few of them including DDR2, ERBB2, EPHA2, PRKCA, CDK6, and AXL overlapped, indicating that these DE kinases could be potential therapeutic targets, especially for the basal B subclass.

We observed 46, the highest number of DE kinases, specifically within basal B cell lines (Figure 3A right). Some of these kinases have been previously described as promising therapeutic candidates in TNBC. For example, PRKCA and CDK6 have been correlated with poor survival outcomes in TNBC and displayed synergy in impeding tumor growth (Hsu et al., 2014). Targeting the hyperactive CDK4/6 axis in breast cancer has attracted traction as a promising therapeutic (Matutino et al., 2018; Yamamoto et al., 2019), and recently, CDK6 inhibition was proven effective for the MSL subclass of TNBC (Pernas et al., 2018). CAMK1 and DDR2 have been linked to basal-like breast cancer wherein they drive metastasis by triggering EMT (Bergamaschi et al., 2008; Corsa et al., 2016; Zhang et al., 2013). EPHA2 has emerged as a new target impairing tumorigenesis in TNBC, and receptor tyrosine kinase AXL has also been reported to play prominent roles in TNBC tumorigenesis (Leconet et al., 2017; Song et al., 2017). Apart from these targets, our study also highlighted other kinases that are yet to be explored for clinical therapeutic interventions. One among them is PEAK1 that modulates TGF signaling responses toward oncogenic progression in breast cancer (Agajanian et al., 2015a, 2015b). Another kinase, NUAK1, has never been described in breast cancer and opens possibilities for new therapeutic explorations in aggressive TNBC subtypes. Although the roles of FYN in mediating EMT and promoting oncogenesis has been described in many cancers including breast cancer (Chen et al., 2013; Xie et al., 2016), its specificity to TNBC or even particular subclass has not been considered. In contrast to the vast array of kinases modulated in the basal B group, the basal A displayed fewer kinases that changed significantly in abundance compared with both the luminal and basal B groups. Tyrosine kinase LYN of the SRC family of kinases that has been linked to aggressive phenotypes including EMT, migration, and invasion in TNBC and another tyrosine-like kinase MLKL were overexpressed specifically within this group (Figure 3A left) (Choi et al., 2010; Tornillo et al., 2018).

From our data, we found that a majority of basal A-specific DE proteases were upregulated (Figure 3B left). Significant upregulation of three kallikreins (KLK5/6/8), three caspases (CASP2/10/14), and three proteasomal components (PSMB8/9/10) suggests involvement of complex protease machinery in driving basal A-specific tumorigenesis. Interestingly, the proteases PSMB8/9/10, all are components of the immunoproteasome representing the modified proteasome that elicit T-cell-mediated immune response underpinning immunomodulatory roles specifically in the basal A subclass (Rouette et al., 2016). Although some of these proteases have been associated with breast cancer tumorigenesis previously (Haritos et al., 2018; Yang et al., 2015), they have not been described in a subclass-specific manner except for CASP14 that was reported to correlate with tumors expressing high claudin and EGFR, characteristic of basal A subclass (Handa et al., 2017). As compared with both luminal and basal A groups, basal B subclasses showed the highest number of proteases that were altered in abundance. Moreover, in contrast to the luminal and basal A subclasses, the majority of DE proteases belonged to the families of metalloproteases and serine proteases (Figures 3B and S2C). Among the metalloproteases, four proteases belonged to the ADAM and ADAMTS family of proteins, with all showing basal B-specific upregulation except for ADAM15. Serine protease FAP, which was found to be overexpressed by 3.6-fold in basal B specifically, has been shown to increase proliferation, migration, and invasion of cancer cells (Huang et al., 2004; Jia et al., 2014). A known metalloprotease MMP2 responsible for extracellular matrix (ECM) degradation in invasive breast cancer was also particularly upregulated within this subclass (Jezierska and Motyl, 2009;

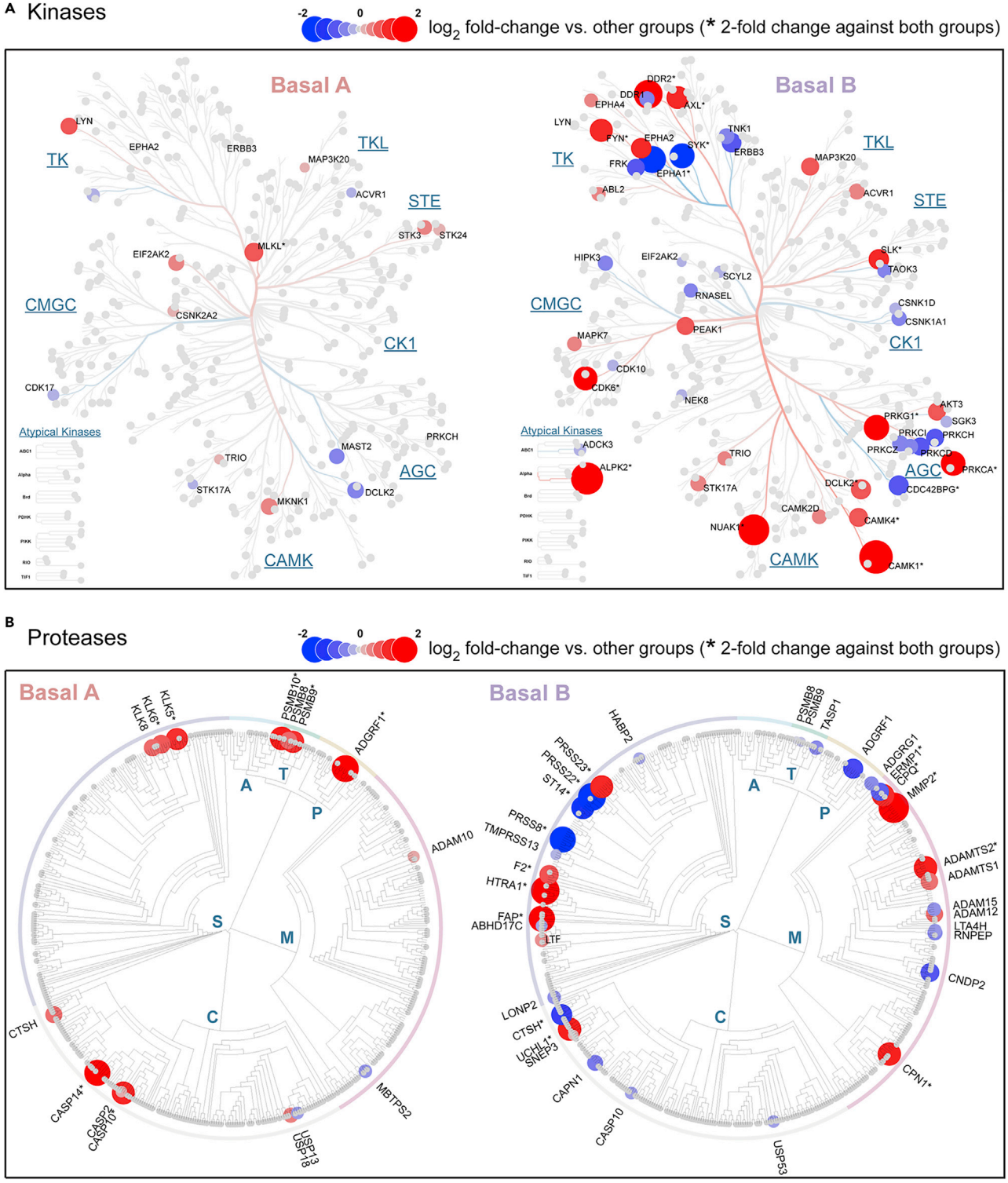



**Figure 3. Kinase and Protease Profiles of TNBC Subclasses**

(A) Differentially expressed kinases ($p \leq 0.05$) specific to either basal A or basal B compared with the other two subclasses mapped onto the human kinome tree. The size and shade of the node represents the median $\log_2$ fold-change of the kinases within each subclass with red and blue representing up- and downregulated kinases, respectively. Those kinases indicated with asterisk are differentially expressed by at least two-fold against both the luminal group

**Figure 3. *Continued***

and the other TNBC subclass. Illustration was created using CORAL online tool (Metz et al., 2018) based on the human kinome tree courtesy Cell Signaling Technology Inc (www.cellsignaling.com).

(B) Differentially expressed proteases ($p \leq 0.05$) specific to either basal A or basal B compared with the other two subclasses mapped according to their protease families. The median $\log_2$ fold-change of the protease within each subclass is represented by the node size and shade. Asterisk denotes those proteases differentially expressed by at least two-fold against both the luminal group and the other TNBC subclass. Red and blue represent up- and downregulated proteases, respectively. The protease families include S – serine, C cysteine, M – metallo, A aspartate, T – threonine proteases, and P – peptidases.

Talvensaari-Mattila et al., 2003). Interestingly, another serine protease HTRA1, which has been widely reported as a tumor suppressor protein with decreased expression being correlated to EMT, proliferation, and angiogenesis (Klose et al., 2018; Lehner et al., 2013; Wang et al., 2012; Xia et al., 2013), was observed to be highly overexpressed only in the basal B subclass. This suggests that there are context-specific roles of HTRA1 that regulate differential mechanisms depending on the breast cancer subclasses. Among all the altered proteases, UCHL1 was the only cysteine protease to be abundantly expressed in the basal B group. Altogether, our study highlights differential expression patterns of kinase and protease families across different breast cancer subtypes and iterates differential mediation of oncogenic responses across the different subtypes.

## Kinase- and Protease-Centric Pathway Aberrations within TNBC Subclasses

Next, we performed expression-based kinase- and protease-centric network analysis in order to understand their involvement in subclass-specific molecular rewiring. For this, all human protein-protein interactions (PPI) associated with the DE kinases and proteases were curated from the Pathway Commons database to identify neighborhood interactors (Cerami et al., 2011). In addition to physical interactions, functional interactions encompassing those that alter phosphorylation, influence expression, and regulate complex formation were also included to dissect functional specifications within each subtype. The network neighborhood was then filtered to retain only those proteins that overlapped with the DE proteins within each subtype to assemble a subtype-specific kinase-protease-centric expression network (Figures 4A and S3A).

On comparing the subclass-specific networks with each other, it was evident that the basal B network displayed a denser topology than the basal A kinase-protease network; partly due to the higher number of DE kinases and proteases identified within basal B subgroup. Despite being a sparse network with only 30 kinases and proteases, the basal A kinase-protease network showed preferential ontology enrichment for regulation of apoptotic processes, cell proliferation, and keratinization. We also identified several DE proteins involved in WNT signaling in the neighborhood network, hinting that the modulated kinases and proteases may have profound influence over WNT signaling regulation in basal A subclass. In addition, the network highlighted kinases including STK24 (MST3), CSNK2A2, and LYN as hub proteins based on their network centrality. Some of these including MST3 and LYN have been identified as pro-proliferative and tumorigenic in TNBC (Cho et al., 2016; Tornillo et al., 2018), and hence implicated as targets, although their specificity to possible subtypes have not been explored. The basal B neighborhood network, on the other hand, revealed a close-knit functional association thus emphasizing on the complex interplay between kinases and proteases in this subtype. In agreement with the aggressive nature of the basal B group, the kinase-protease network showed particular inclination toward processes associated with focal adhesion, extracellular matrix organization, ephrin, and TGF signaling (Canel et al., 2013; Drabsch and ten Dijke, 2011; Lu et al., 2012; Wang, 2011), all of which mediate a more migratory and invasive phenotype. In addition, proteins associated with angiogenic and platelet activation were also enriched in the proximity neighborhood of basal B-specific kinases and proteases. The interaction network spanning a total of 75 kinases and proteases included the before-mentioned kinases AXL, PRKCA, PEAK1, CDK6, and EPHA2 and proteases FAP, MMP2, and UCHL1 in close connectivity with each other, emphasizing an intricate rewiring of kinase-protease networks within the basal B subclass. Some of these kinases and proteases have been shown to correlate with poor prognostic outcomes in breast cancer and are currently being explored as therapeutic targets (Hsu et al., 2014; Leconet et al., 2017; Liu et al., 2019; Song et al., 2017). We also inferred crucial hubs from the neighborhood network, a majority of which play key roles in invasion and metastatic behavior. Specifically, we identified components of TGF signaling including SMAD3, EMT-activator ZEB1, ECM regulators MMP2 and FN1, and transcriptional regulators JUN and SP1 as key proteins in basal B subclass. A cluster assessment based on MCODE topological analysis highlighted several intricately connected proteins, with the top enriched clusters showing functional preference for ECM

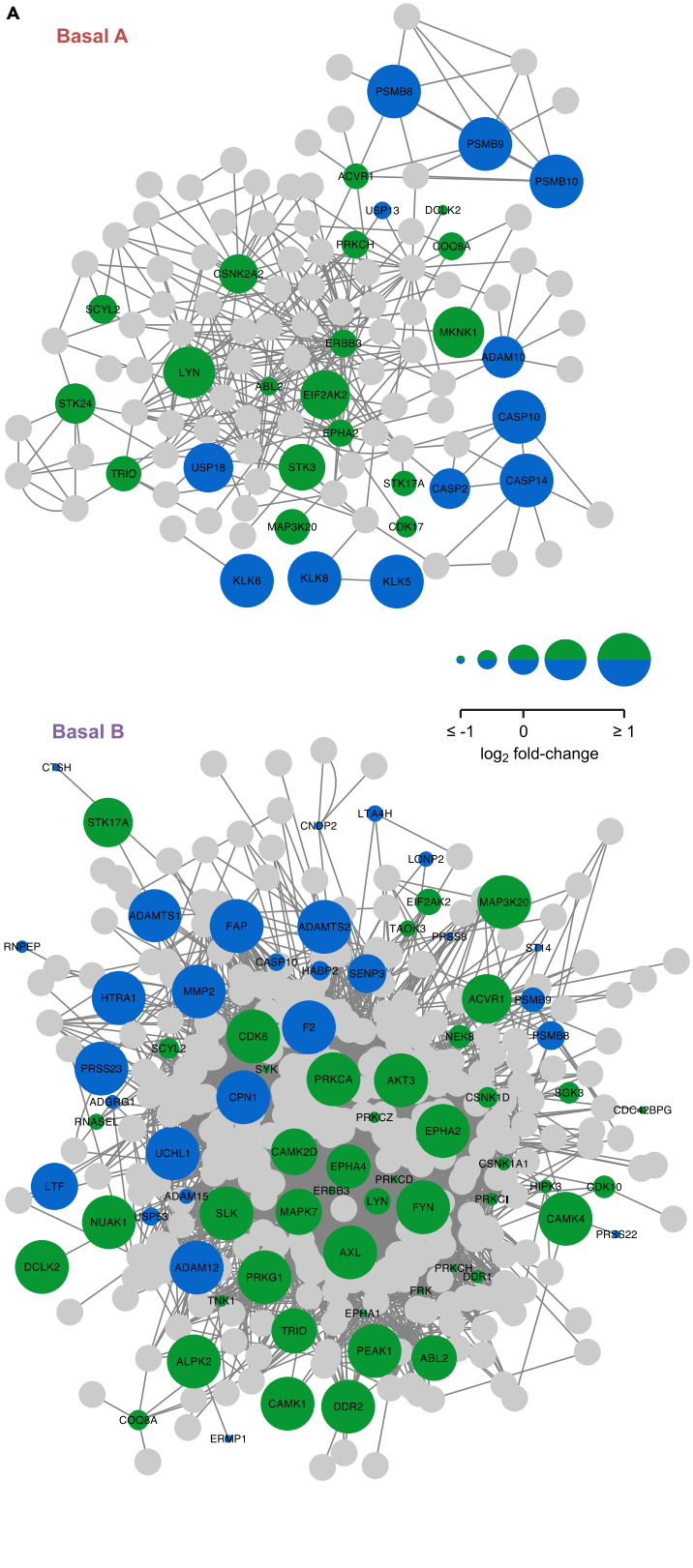

A
Basal A
PSMB8
PSMB9
PSMB10
ACVR1
USP13
DCLK2
PRKCH
COQ8A
CSNK2A2
SCYL2
MKNK1
ERBB3
ADAM10
LYN
ABL2
EIF2AK2
STK24
CASP10
EPHA2
TRIO
USP18
STK3
STK17A
CASP2
CASP14
MAP3K20
CDK17
KLK6
KLK8
KLK5
≤ -1
0
≥ 1
log2 fold-change
Basal B
CTSH
STK17A
CNOT2
LTA4H
LONP2
ADAMTS1
FAP
ADAMTS2
EIF2AK2
MAP3K20
TAOK3
PPP3R1
STK4
RNPEP
CASP10
HABP2
USENP3
HTRA1
MMP2
ACVR1
PSMB9
CDK6
F2
NEK8
PSMB8
SCYL2
PRSS23
SYK
PRKCA
AKT3
CPN1
CSNK1D
SGK3
ADCYAP1
CDC42BPG
RNASEL
PRKCZ
EPHA2
UCHL1
CAMK2D
EPHA4
CSNK1A1
LTF
PRKCD
HIPK3
CDK10
ADAM15
ERBB3
LYN
FYN
PRKCI
NUAK1
USP53
SLK
MAPK7
CAMK4
PRSS22
AXL
DCLK2
ADAM12
PRKG1
PRKCH
DDR1
TNK1
FRK
EPHA1
TRIO
PEAK1
ABL2
ALPK2
CAMK1
DDR2
COQ8A
ERMP1
B
EMT
Coagulation
Interferon gamma Signaling
Estrogen Response
Log2 Median Expression Ratio
Luminal
Basal A
Basal B
**

**Figure 4. Kinase-Protease-Centric Interaction Network**

(A) Proximity network of differentially expressed kinases and proteases in each TNBC subclass. The kinases and proteases are depicted as green and blue nodes, respectively, and the size of the nodes represents the median $\log_2$ fold-change against the other two subclasses, with larger nodes representing increased fold-change and vice versa. The edges represent physical and functional interactions between proteins.

(B) Distribution of protein ratio within each breast cancer subtype for the indicated cancer hallmark gene sets (*, p value $\leq$0.05; **, p value $\leq$0.01; n.s., not significant—using KS test).

remodeling, focal adhesion, inositol phosphate metabolism, HIF-1, TGF, and ephrin signaling within the basal B kinase-protease proximity network (Figure S3B). All these evidences suggest that the basal B kinases and proteases work in concert with each other to rewire signaling pathways particularly involved in cell migration and invasion, thus explaining their aggressive nature. This is strikingly different from the signaling events and processes influenced by basal A-specific kinases and proteases.

With observations of distinct functional biases for the two subclasses based on ontology from the kinase-protease proximity network, we next looked into cancer hallmark gene sets uniquely enriched within each breast cancer subgroup using gene set enrichment analysis (GSEA) to assess oncogenic signatures (Table S2A). As expected, the luminal cell lines showed enrichment of estrogen response process. In addition, the luminal group also showed preferential enrichment of metabolism-associated processes such as oxidative phosphorylation and fatty acid metabolism. Interferon-induced signaling was particularly enhanced in basal A group, and the proliferation-associated p53 signaling pathway was also positively enriched in basal A compared with basal B group. Along with EMT, which emerged as the top hallmark in basal B group, other cancer hallmarks such as hypoxia, inflammatory response, and coagulation also displayed selective enrichment in the basal B group. To ascertain if the kinase and protease signatures within each subtype underlie such differences, we specifically compared the distribution of only proteins within the kinase-protease neighborhood belonging to these cancer gene sets (Figures 4B and S4A). Excitingly, we found that the proximity proteins of kinases and proteases are representative of such functional processes, and accordingly EMT, TGF signaling, TNFa signaling via NF-κB, coagulation, hypoxia, and inflammatory response, all showed elevated response in basal B subclass compared with luminal and basal A groups. Similar functional biases were also observed for basal A subclass for interferon signaling and luminal group for estrogen and metabolic response suggestive of intricate functional roles for kinases and proteases in driving subclass-specific distinctive processes. We also looked into the KEGG pathway database to assess pathway preferences for proteins within the kinase-protease neighborhood and found that positive regulation of TGF, PI3K/AKT, HIF-1, RAP1, and sphingolipid signaling in basal B cell lines could be effectively represented by the abundances of kinase-protease proximity proteins (Figure S4B and Table S2B).

## Dysregulated Kinase and Protease Associations within TNBC Subclasses

We expanded our analysis to investigate differences in the protein co-regulation networks between the subclasses. For this, protein correlations were obtained between all protein pairs of the 10,400 quantified proteins, and a high-confidence correlation network was obtained after filtering for significant associations. The positive correlation network included a total of 92,643 co-regulated protein pairs, and the negative correlation network comprised of 61,061 associating pairs, both corresponding to a minimum Spearman's rank correlation coefficient of ~0.76 (Table S3). To infer cell-line-specific network aberration, we specifically identified protein pairs displaying significant deviations from correlation in each cell line by performing outlier analysis on all protein pairs in the significant protein-association networks (Lapek et al., 2017) (Table S4). The median dysregulated associations based on positively and negatively correlating protein pairs across all cell lines were 1,444 and 860, respectively. On mapping these dysregulated protein associations, we observed differential patterning of proteins with deviations that aligned with the breast cancer subtypes, particularly for those inferred from positive correlations. This suggested that protein dysregulations are more likely to be conserved across different subtypes, and investigating such dysregulations can offer insights on subtype-specific pathways and mechanisms. Thus to identify dysregulated proteins specifically enriched within each breast cancer subtype we carried out enrichment analysis using hypergeometric testing (Figures 5A and S5A and Tables S5A and S5B). This analysis highlighted that the basal B subclass included the most number of dysregulated proteins (672), followed by luminal subtype with 455 and basal A subclass with 171 dysregulated proteins. Through ontology-based enrichment analysis of these dysregulated proteins, we observed subtype-specific dysregulated pathways (Figure 5A and Table S6). For instance, proteins mainly involved in metabolism and mitochondrial translation were dysregulated in

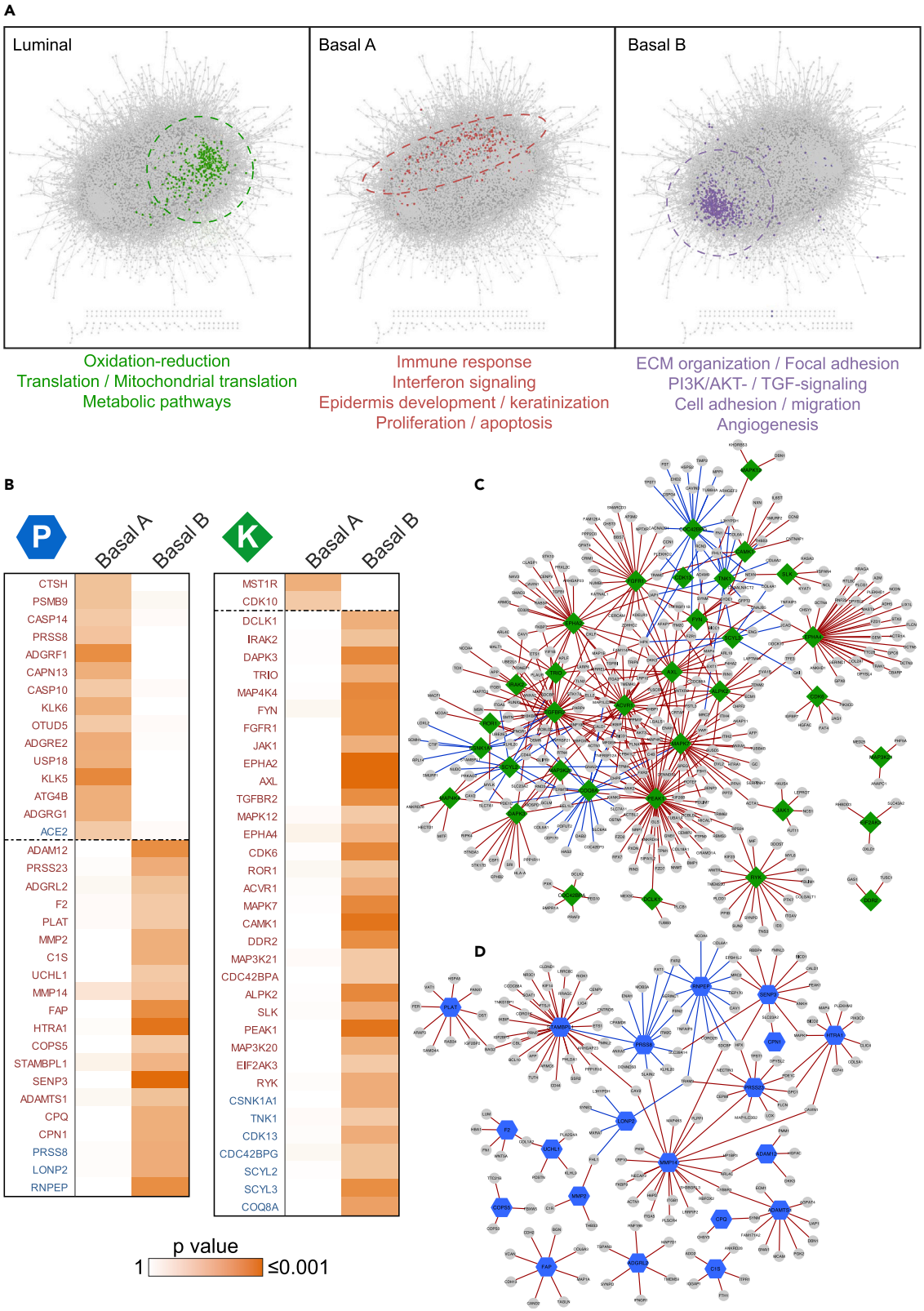

A
Luminal
Basal A
Basal B
Oxidation-reduction
Translation / Mitochondrial translation
Metabolic pathways
Immune response
Interferon signaling
Epidermis development / keratinization
Proliferation / apoptosis
ECM organization / Focal adhesion
PI3K/AKT- / TGF-signaling
Cell adhesion / migration
Angiogenesis
B
P
Basal A
Basal B
CTSH
PSMB9
CASP14
PRSS8
ADGRF1
CAPN13
CASP10
KLK6
OTUD5
ADGRE2
USP18
KLK5
ATG4B
ADGRG1
ACE2
ADAM12
PRSS23
ADGRL2
F2
PLAT
MMP2
C1S
UCHL1
MMP14
FAP
HTRA1
COPS5
STAMBPL1
SENP3
ADAMTS1
CPQ
CPN1
PRSS8
LONP2
RNPEP
K
Basal A
Basal B
MST1R
CDK10
DCLK1
IRAK2
DAPK3
TRIO
MAP4K4
FYN
FGFR1
JAK1
EPHA2
AXL
TGFBR2
MAPK12
EPHA4
CDK6
ROR1
ACVR1
MAPK7
CAMK1
DDR2
MAP3K21
CDC42BPA
ALPK2
SLK
PEAK1
MAP3K20
EIF2AK3
RYK
CSNK1A1
TNK1
CDK13
CDC42BPG
SCYL2
SCYL3
COQ8A
p value
1
≤0.001
C
D

**Figure 5. Dysregulated Proteins in Breast Cancer Subclasses**

(A) Proteins with dysregulated associations in each breast cancer subclass mapped onto the background protein-protein association network derived from co-regulation analysis of protein abundances from all breast cancer cell lines. The network includes only positive correlations (for negative correlations see Figure S5A). The corresponding functional annotation of the dysregulated proteins for each subclass is indicated below.

(B) Dysregulated kinases and proteases specifically enriched within each TNBC subclass are depicted as heatmap based on -$\log_{10}$ transformed p values derived from hypergeometric testing. Kinases or proteases derived from dysregulations of positive associations and negative associations are depicted as red and blue, respectively. K denotes kinases and P denotes proteases.

(C and D) (C) Perturbed association network of kinases and (D) proteases specifically dysregulated in basal B subclass. Red edges denote perturbations inferred from positive associations and blue edges represent perturbations inferred from negative associations.

the luminal subgroup. In basal A, proteins involved in epidermis development/keratinization, innate immune response, interferon signaling, and those regulating proliferation and apoptosis showed dysregulation, whereas for basal B, those associated with ECM organization, migration, angiogenesis, PI3K/AKT, and TGF signaling pathway were dysregulated. Thus the dysregulations inferred captured functional distinctions between the two subclasses. In several instances, the associated proteins also displayed consistent pairwise dysregulation within subclasses, with basal B displaying the most number of deviated protein pairs (Figure S5B and Tables S5C and S5D). For example, protein associations involving EPHA2, TGFBI, DKK3, AXL, SERPINE1, and several collagens were dysregulated in a majority of the basal B cell lines. In contrast, the basal A cell lines revealed far less number of conserved protein pairs, with associations involving SLPI, KLK5, and keratins showing perturbations in many cell lines in a basal A-specific manner.

We posit that kinases or proteases displaying association perturbations with several other proteins in a subtype-specific manner reflect their activity modulation in that particular subtype. Interrogating the dysregulated kinases and proteases among different breast cancer types, we found that the basal B group had the highest representation of perturbed kinases and proteases (Figures 5B and S6 and Tables S5A and S5B). Among them, AXL, DDR2, PEAK1, EPHA2, and FYN were also specifically overexpressed in basal B stressing on possible key roles played by these kinases in the aggressive basal B subclass. In addition, our analysis also identified several other kinases that were not significantly differently expressed but whose associations have been dysregulated in a basal B-specific manner (Figures 5B and 5C). For example, receptor kinases FGFR1 and TGFBR2 were predicted to be dysregulated only in basal B subclass. FGFR1 is regarded as a prognostic factor in TNBC, and TGF-β pathway components have been considered promising targets for impairing tumorigenesis in TNBC (Bhola et al., 2013; Cheng et al., 2015; Jamdade et al., 2015). Associated kinases that regulate these pathways including the neurotrophic tyrosine kinase ROR1 that potentiates FGFR signaling in basal breast cancer, and activin signaling receptor-like kinase AVCR1 that activates SMAD pathway were also predicted to be dysregulated in a basal B-specific manner (Bashir et al., 2015; Pandey et al., 2019; Ramachandran et al., 2018) A member of the IRAK family of kinase, IRAK1, which mediates toll-like receptor (TLR) signaling, has been particularly proposed as a therapeutic target driving TNBC metastasis (Wee et al., 2015). Our dysregulation analysis predicted another member IRAK2, which is crucial in late-phase TLR signaling, to be modulated in basal B. It is interesting to note that many of the dysregulated kinases including TGFBR2, PEAK1, ACVR1, EPHA2, EPHA4, and AXL shared perturbed associations with multiple proteins within the basal B subclass, suggesting prominent roles for these kinases in driving oncogenic aberrations (Figure 5C). Apart from these kinases inferred from positively correlated dysregulated associations, we also identified others that displayed deviations in their negative correlation especially in the basal B group. Among them, the non-receptor tyrosine kinase TNK1 (KOS1) has been reported as a tumor suppressor in breast cancer that downregulates RAS signaling (Hoare et al., 2008). Another transcriptional kinase, CDK13, however, is associated with aggressive features in breast cancer and is regarded as a potential target for TNBC (Quereda et al., 2019). In comparison to the extensive modulation of kinase association network in basal B group, only two kinases were identified to be dysregulated in basal A class. The dysregulated MST1R kinase, also known as RON receptor tyrosine kinase, is frequently overexpressed in TNBC, and anti-RON targeted therapy has been advocated based on promising outcomes from preclinical TNBC models (Suthe et al., 2018).

Apart from kinases, we also observed several proteases displaying subtype specificity (Figures 5B, 5D, and S6 and Tables S5A and S5B). Strikingly, in contrast to kinase perturbations, protease association dysregulations were more pronounced in basal A subgroup. Proteases such as caspases 10 and 14, kallikreins 5 and 6, and calpain 13, all displayed perturbations with multiple proteins among the basal A cell lines, implicating dysregulation of a wide array of protease families in the basal A subclass. Some of these

including KLK5 and CASP14 are prognostic markers for aggressive TNBC (Handa et al., 2017; Yang et al., 2015). In the basal B subclass, we predicted multiple metalloproteases (MMP2 and MMP14), ADAM family proteases (ADAM12 and ADAMTS1), and deubiquitinases (UCHL1 and STAMBPL1) to be specifically dysregulated. Although MMP2 protease was found to be both dysregulated and also overexpressed in basal B group, we predicted MMP14 to be particularly modulated in basal B subclass based on our dysregulation analysis. Of note, ADAM12 has been reported to promote cancer stemness in claudin-low breast cancer (Duhachek-Muggy et al., 2017), and this aligns with our prediction that ADAM12 is dysregulated specifically in the basal B subclass that expresses low claudin. Apart from these, a class of serine proteases encompassing PRSS23 and PRSS8 were also found to have dysregulated associations in the basal B group. PRSS8, also known as prostasin, has been reported to inhibit breast cancer invasiveness (Chen and Chai, 2002), and our analysis predicted negative association perturbations.

### AXL Regulates Subclass-specific Migration and Invasion

Our comprehensive protein-protein network analysis revealed several TNBC subclass-specific potential therapeutic targets including kinases and proteases. We validated some of these to highlight protein expression differences within the three breast cancer groups by Western blotting in multiple representative cell lines (Figure S7).

Multiple evidences suggest dysregulation of AXL, and its associations specifically in basal B subclass, and validation experiments confirmed AXL overexpression exclusively in basal B subclass as against the other groups (Figures 3A, 5C & 6A, S3B, and S7A). Hence, we hypothesized that basal B cells should be more sensitive to AXL inhibitors than basal A cells. To test this, we treated seven TNBC cell lines representing basal A (BT20, MB468, and HCC70) and basal B (BT-549, Hs578t, MB-157, and MB-231) subclass with AXL-specific inhibitor R428 that inhibits the autophosphorylation of AXL required for its activity (Korshunov, 2012; Sadahiro et al., 2018). Using an antibody against phosphorylated AXL at tyrosine 779 (pAXL-Tyr$^{779}$), we confirmed effective inhibition of AXL activity upon treatment with R428 inhibitor in basal B cell lines (Figure 6B). Next, we investigated whether inhibition of AXL suppresses the aggressive behavior of TNBC cell lines by assessing the migration and invasion response of the selected TNBC cell lines to AXL inhibitor R428 using modified wound-healing assays. The reduction in AXL activity with inhibitor treatment was accompanied with a significant decrease of invasion for all basal B cell lines, with Hs578t showing the highest sensitivity (~80% reduction in invasion) (Figure 6C). As expected, the basal A cell lines showed no significant changes in their invasive potential in response to AXL inhibitor (Figure 6C). The invasive feature of basal B cell lines was specifically responsive to AXL inhibition as compared with the migratory potential that displayed cell-line-dependent responses (Figure S7B). We also performed the functional assays using transient AXL knockdown in Hs578 T cell line, and the results showed significant decrease in invasion, whereas migration showed no notable effects (Figures 6D and 6E). Altogether, this suggested that AXL kinase and its associated mechanisms mediate aggressive behavior specifically in basal B subclass and hence is an effective target for basal B but not basal A type tumors.

## DISCUSSION

Deconvolution of TNBC complexity is necessary to resolve molecular mechanisms and devise effective targeted therapies for this highly heterogeneous and aggressive breast cancer subtype. In this study, we aimed to explore patterns of proteome profiles and pertinent pathway dysregulations within TNBC by employing advanced proteomics MS in combination with methodical computational approaches. The coupling of proteomic MS with two peptide fractionation methods boosted the overall quality of our proteome data with higher peptide and proteome coverage along with robust quantification accuracy as revealed by good concordance between the biological replicates. This is exemplified by the identification of regulatory proteins belonging to various functional classes such as kinases, transcription factors, and proteases, which are generally not extensively profiled by conventional MS-based proteomics profiling and hence are often underrepresented. With consistent high coverage of proteome across the panel, we obtained the most comprehensive quantitative proteome profile of breast cancer cell lines reported to date. Unsupervised clustering of the unprecedented high-quality proteome revealed two major subclasses within TNBC, basal A and B, which aligned with previously reported transcriptome-based classification. Moreover, for the first time, we observed distinct expression profile patterning of kinases and proteases that were able to segregate the two TNBC subclasses, stressing on key roles played by these protein classes within TNBC.

Although similar TNBC subclasses based on proteomics differences have been reported previously (Lapek et al., 2017; Lawrence et al., 2015), our observation of exclusive subclass-specific kinase and protease

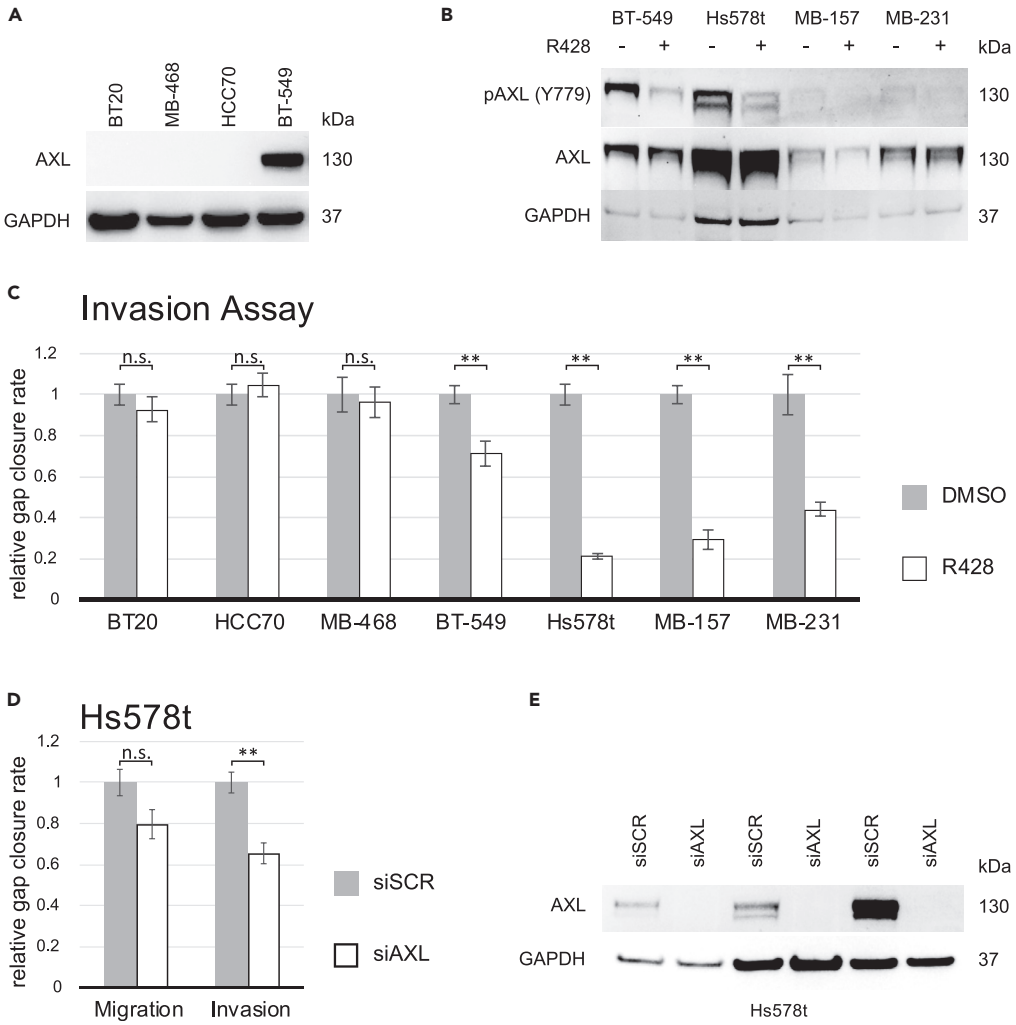



**Figure 6. Subclass-specific Regulation of AXL Kinase**

(A) Western blot analysis of AXL expression in selected basal A and B cell lines (only a section of Figure S7A is shown).

(B) Western blot analysis of AXL and pAXL-Tyr$^{779}$ in selected basal B cell lines with and without AXL inhibitor (R428) treatment.

(C) Invasive potential represented as relative gap closure rates between R428 and DMSO-treated cells. Data shown as mean ± SEM, n = 3. *, p value ≤0.05; **, p value ≤0.01; n.s., not significant (using t-test).

(D) Migratory and invasive potential represented as relative gap closure rates for Hs578t cell line treated with siRNA against AXL (siAXL) or scramble (siSCR) as control. Data shown as mean ± SEM, n = 3. *, p value ≤0.05; **, p value ≤0.01; n.s., not significant (using t-test).

(E) Western blot analysis of AXL expression in Hs578T cell line treated with siRNAs.

signatures offered direct phenotypic perspectives on the different TNBC subclasses. In fact, some of the observed basal B-specific kinases and proteases have been previously shown to be linked with migratory and invasive potential of tumors, supporting a more aggressive feature of basal B cell lines. We observed several kinases including PEAK1, ACVR1, and NUAK1 specifically within the basal B subclass, and it is worthwhile to assess their roles in mediating aggressive oncogenic features for exploring newer targets for this subclass. We also found AXL with increased abundance specifically within the basal B subclass. AXL has been previously shown to be associated with TNBC tumorigenesis, particularly in those cancers cells displaying mesenchymal features (Bottai et al., 2016). AXL activates PI3K/AKT and JAK/STAT pathways, resulting in increased cell survival and reduction of immune response, and also mediates EMT and increases invasion (Gay et al., 2017; Goyette et al., 2018; Wilson et al., 2014). With its prominent oncogenic roles, AXL has been proposed as a therapeutic target for TNBC and also for other cancers, and one of its selective inhibitors R428 (BGB324; Bemcentinib) is under clinical trials for acute myeloid leukemia

(AML), non-small cell lung cancer (NSCLC) (NCT02488408/NCT02424617) as well as for advanced TNBC and adenocarcinoma of the lung in combination with pembrolizumab (Yule et al., 2018). Despite several studies showing that AXL mediates TNBC tumorigenesis (Holland et al., 2010; Wilson et al., 2014), the possibility of subclass-specific functions or therapeutic implications has not been explored thus far. In fact, in our study, we demonstrated that inhibition of AXL suppresses the migratory and invasive behavior specifically in the basal B subclass and hence, can be an effective target for tumors belonging to this subclass. In comparison to kinases, proteases are far less studied, their repertoire of targets is not well understood, and hence, only a few available specific inhibitors have been exploited in the clinic. With their profound roles in various functional processes, particularly in mediating migration and invasion during tumor progression, there has been a leap in the number of studies focusing on elucidating the mechanistic roles and defining their target substrates. It is therefore not surprising that except for CASP14, none of our observed DE proteases have been reported in a subclass-specific manner, even though some of them including FAP, MMP2, UHCL1, and kallikreins have been reported to be involved in breast cancer tumorigenesis in general (Haritos et al., 2018; Huang et al., 2004; Liu et al., 2019; Talvensaari-Mattila et al., 2003; Yang et al., 2015).

The analysis of subclass-specific kinase-protease-centric signatures and their associated networks represented known functional signatures of each breast cancer type and also revealed several molecular hubs, altered pathways, and signaling crosstalk that may signify intricate molecular rewiring within each type. For the basal B class, the kinase-protease-centric network and signatures displayed functional inclination to a more aggressive phenotype with potential crosstalk between signaling cascades. We found that HIF-1 signaling and hypoxia-driven signatures were more pronounced in basal B as compared with basal A subclass. Hypoxia has been shown to be linked to EMT and cancer progression in many solid tumors (Muz et al., 2015), and its major regulator transcription factor HIF-1α is known to be associated with advanced disease stage and poor clinical outcome in breast cancers (Liu et al., 2015; Yamamoto et al., 2008). Recent studies demonstrated that HIF-1α is particularly overexpressed in TNBC tumors as compared with other subtypes and favors invasive nature (Chen et al., 2014; Ponente et al., 2017). From the basal B kinase-protease network, we inferred that HIF-1α lies within the proximity neighborhood of protein kinase C α (PKCα encoded by PRKCA). PKC isoforms including PKCα have been reported to play crucial roles in regulating the stability, accumulation, and transactivation of HIF-1α during hypoxia (Datta et al., 2004; Lee et al., 2007). PKCα has also been implicated in TNBC aggressiveness, and PKCα inhibitors have been advocated as promising therapeutic agents (Humphries et al., 2014; Lin et al., 2017). With the observed basal B-specific enhanced expression of PKCα, we postulate a robust activation and maintenance of HIF-1 signaling particularly within the basal B TNBC subgroup. Further, HIF-1α stabilization through deubiquitination by UCHL1, a protease we found to be overexpressed within basal B, has also been reported to increase HIF-1 signaling, leading to angiogenesis and distant metastasis (Goto et al., 2015). Taken together, these data suggest HIF-1 signaling to be specifically hyperactivated within the basal B group. The network also revealed transcription factor JUN as one of the key functional hubs within the proximity neighborhood of several altered kinases and proteases including MAPK7, DDR2, EPHA2, and FAP in the basal B subclass. Among breast cancer types, overexpression and activation of JUN transcriptional cascade has been reported especially in TNBC wherein it aggravates proliferation and migratory potential along with triggering a stem cell phenotype (Xie et al., 2017). A potential crosstalk between JUN signaling and EPHA2 receptor is known to promote phenotypic alterations associated with tumorigenesis and acquiring stem-cell-like properties in cancers (Song et al., 2014). With EPHA2 abundantly expressed in basal B, our results support sustained activation of EPHA2 via JUN, which may be conducive for promoting aggressive features within basal B subclass. In addition to pinpointing signaling hubs, we also identified EMT inducer ZEB1 as a key molecular pivot within the basal B kinase-protease network supporting a more invasive and migratory signature for this aggressive subclass (Zhang et al., 2015). Such evidences from our analyses thus add validity to the relevance of unique kinase and protease signatures among different breast cancer subtypes and reinforce their profound influence over key functional processes.

Although the analysis based on differentially regulated proteins, especially kinases and proteases, provided insights on possible functional rewiring of basal A and basal B TNBC subclasses, signaling pathways and networks underlie intricate involvement of several proteins beyond those that are only differentially expressed. There exists a high degree of protein co-regulation and association in various cellular processes, and any imbalance affecting the stoichiometry between proteins or within functional complexes often leads to signaling aberrations making them vulnerable to transformations. We

hypothesized that inherent protein-association dysregulations exist within each subtype of breast cancer, and mining for such functional differences at the systemic level will identify subtype-specific molecular aberrations. Indeed, we observed several protein associations that were consistently dysregulated in particular groups of breast cancer subtypes. Although not all dysregulated proteins identified through this analysis were differentially expressed in a subtype-specific manner, the functional differences among the breast cancer subgroups were well represented. The basal B subclass displayed the highest number of dysregulated proteins involved in perturbed associations. Several of these intricately associated proteins are known to play vital roles in invasion and metastatic processes. For example, the specific prediction of MMP14 in the more aggressive basal B group rather than the basal A group is interesting, as MMP14 is implicated in aggressive phenotypes of breast cancer and is involved in mediating EMT, thus contributing to invasion and metastasis (Jiang et al., 2006). We also predicted TGFBR2 and associated TGF signaling to be specifically modulated in the basal B subclass, and indeed TGF-β pathway is well documented in mediating oncogenesis in the mesenchymal-like and claudin-low subgroups of TNBC (Wahdan-Alaswad et al., 2016). Besides triggering EMT mechanisms, TGF-β pathway also enhances cell proliferation and reduces apoptosis (Massagué, 2008; Xu et al., 2018). Although both TGFBR1 and TGFBR2 heterodimerize during transcriptional activation via SMAD signaling (Principe et al., 2014), only TGFBR1 inhibitors are predominant in clinical trials for effectively blocking the pathway. Given the pro-proliferative role of TGFBR2 to activate MAPK and PI3K/AKT pathways independently of TGFBR1 (Bakin et al., 2000; Janda et al., 2002), and the fact that we predict TGFBR2 perturbation within basal B subclass, it is prudent to inhibit TGFBR2 for effective abolishment of downstream oncogenic effects pertaining to this pathway within aggressive tumors belonging to basal B subclass. Indeed, several of the kinases that we observed to vary in abundance within the basal B subclass also have been reported to engage in signaling crosstalk with TGFBR2 and their associated signaling. For instance, PEAK1 has been reported as a molecular switch altering TGFBR2 responses from canonical tumor suppressing to non-canonical tumor promoting, thus driving EMT and invasion (Agajanian et al., 2015a, 2015b). For protease FAP, which was both overexpressed and predicted to be dysregulated within basal B, we observed perturbed associations with several EMT-associated proteins such as BGN, CDH2, TAGLN, COL6A3, and VCAN (Chen et al., 2019; Hu et al., 2014; Huang et al., 2018; Mrozik et al., 2018; Sheng et al., 2006). Such dysregulated associations of FAP strongly argue for an important functional role specific to basal B cells; however, little is known about its actual *in vivo* targets and tumorigenic mechanisms (Hamson et al., 2014). All these assert that protein associations and co-regulations are critical determinants in defining cellular mechanisms and functional alterations. Although some of the dysregulated kinases and proteases including AXL, EPHA2, MMP2/14, and FAP have previously been shown to be possible targets for TNBC, they have not been studied in a subclass-specific manner. While some of these are currently being explored for TNBC therapy, our analyses suggest that targeting these proteins could prove more effective in a particular subclass rather than in TNBC in a broader context. In addition to these, we also unraveled several other novel kinases and proteases that have the potential to be exploited as TNBC subclass-specific druggable targets.

Although the molecular heterogeneity of TNBC is well documented, success with regard to clinical intervention has been disappointing. Despite several studies reporting varied expression patterns of genes and proteins within TNBC, a thorough investigation from a therapeutic perspective to unravel the complexity has been lacking. Our systemic and systematic workflow, with emphasis on protein association dysregulations, opens up new avenues for understanding molecular perturbations at the subtype level and also extracts subclass-specific therapeutic targets for strategized clinical applications. The candidates identified in our study are now at secondary validation stage, where tumor screening for target verification and *in vivo* biological studies should be performed. At the same time, establishment of robust subclass-specific biomarkers is mandatory for patient stratification for successful targeted treatments. Here, our analysis has been carried out focusing only on two major TNBC subclasses. Nevertheless, we postulate that there could be more than two subclasses within TNBC with different functional signatures. For this, deep proteome profiling of all available TNBC cells as well as tumors, followed by systematic analysis herein reported is required. Altogether, our study uncovers molecular mechanisms within TNBC subclasses and thus holds potential to enhance applications of personalized medicine for TNBC.

## Limitations of the Study

This study revealed distinct proteome-based subclasses within TNBC that are functionally discrete and highlighted mechanisms and potential targets unique to the subclasses opening new avenues for further functional explorations. However, certain limitations and challenges remain to be addressed. First,

the repertoire of TNBC cell lines profiled by this study is limited and may not represent the complex heterogeneity underlying TNBC. Extensive proteomic screening of a wide array of breast cancer cell lines that represents the entirety of TNBC complexity is likely to identify additional subclasses or better resolve existing subclasses with distinctive functional signatures. These subclasses may refine our current understanding of transcriptome-derived TNBC clusters and offer functional insights on inherent molecular specificities within TNBC. An additional drawback of this study is the use of cell lines for proteomic profiling and validation. Existing cell line models are homogeneous and lack components of the tumor such as infiltrating cells, macrophages, blood vessels, and ECM that represent the tumor microenvironment and hence are not perfect in mimicking actual tumor conditions. Functional interplay between tumors and its environment is crucial in defining oncogenic progression, and such complexities are not considered in the current cell-line-based proteome profiling. Yet, our study shows feasibility of unraveling TNBC complexity using proteomic approaches and extension of such investigations to profile tumors derived from large patient cohorts will benefit from patient stratification based on functional subclasses that hold potential for improved TNBC-targeted therapies. With emphasis laid on subclass-specific roles of kinases and proteases based on our proteomic characterization, more efforts are needed in investigating their oncogenic roles within TNBC subclasses to explore their clinical utility for personalized medicine. This is especially important for targets that are currently under clinical trial where knowledge on subclass specificity will favor successful clinical outcomes.

## METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

## DATA AND CODE AVAILABILITY

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PXD017025.

## SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.100868.

## ACKNOWLEDGMENTS

We thank Hannah L.F. Swa and Suat Peng Neo for advice on cell line maintenance and Siok Ghee Ler for technical assistance with mass spectrometric sample analysis. This work is supported by Agency for Science, Technology and Research (A*STAR), Singapore. M.K. is supported by a SINGA (Singapore International Graduate Award) fellowship.

## AUTHOR CONTRIBUTIONS

M.K., A.A.S., and J.G. designed the experiments and analysis. M.K. executed the proteomic mass spectrometry and biological validation experiments. M.K. and A.A.S. performed proteomic data interpretation and computational analyses. M.K. and A.A.S. generated all the figures and wrote the manuscript. M.K., A.A.S., and J.G. critically interpreted and evaluated the data. J.G. supervised all experiments and data analyses, directed the project, and wrote the manuscript. B.H.B. evaluated and revised the manuscript. All authors read and approved the final version.

## DECLARATION OF INTERESTS

The authors declare that they have no competing interests.

Received: March 5, 2019
Revised: December 30, 2019
Accepted: January 20, 2020
Published: February 21, 2020

## REFERENCES

Agajanian, M., Campeau, A., Hoover, M., Hou, A., Brambilla, D., Kim, S.L., Klemke, R.L., and Kelber, J.A. (2015a). PEAK1 acts as a molecular switch to regulate context-dependent TGFbeta responses in breast cancer. PLoS One *10*, e0135748.

Agajanian, M., Runa, F., and Kelber, J.A. (2015b). Identification of a PEAK1/ZEB1 signaling axis during TGFbeta/fibronectin-induced EMT in

breast cancer. Biochem.Biophys. Res. Commun. *465*, 606–612.

Anders, C.K., and Carey, L.A. (2009). Biology, metastatic patterns, and treatment of patients with triple-negative breast cancer. Clin.Breast Cancer *9*, S73–S81.

Badve, S., Dabbs, D.J., Schnitt, S.J., Baehner, F.L., Decker, T., Eusebi, V., Fox, S.B., Ichihara, S., Jacquemier, J., Lakhani, S.R., et al. (2011). Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists. Mod.Pathol. *24*, 157–167.

Bakin, A.V., Tomlinson, A.K., Bhowmick, N.A., Moses, H.L., and Arteaga, C.L. (2000). Phosphatidylinositol 3-kinase function is required for transforming growth factor beta-mediated epithelial to mesenchymal transition and cell migration. J. Biol. Chem. *275*, 36803–36810.

Bareche, Y., Venet, D., Ignatiadis, M., Aftimos, P., Piccart, M., Rothe, F., and Sotiriou, C. (2018). Unravelling triple-negative breast cancer molecular heterogeneity using an integrative multiomic analysis. Ann. Oncol. *29*, 895–902.

Bashir, M., Damineni, S., Mukherjee, G., and Kondaiah, P. (2015). Activin-A signaling promotes epithelial-mesenchymal transition, invasion, and metastatic growth of breast cancer. NPJ Breast Cancer *1*, 15007.

Bauer, K.R., Brown, M., Cress, R.D., Parise, C.A., and Caggiano, V. (2007). Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype. Cancer *109*, 1721–1728.

Bergamaschi, A., Kim, Y.H., Kwei, K.A., La Choi, Y., Bocanegra, M., Langerod, A., Han, W., Noh, D.Y., Huntsman, D.G., Jeffrey, S.S., et al. (2008). CAMK1D amplification implicated in epithelial-mesenchymal transition in basal-like breast cancer. Mol. Oncol. *2*, 327–339.

Bhola, N.E., Balko, J.M., Dugger, T.C., Kuba, M.G., Sánchez, V., Sanders, M., Stanford, J., Cook, R.S., and Arteaga, C.L. (2013). TGF-β inhibition enhances chemotherapy action against triple-negative breast cancer. J. Clin. Invest. *123*, 1348–1358.

Bianchini, G., Balko, J.M., Mayer, I.A., Sanders, M.E., and Gianni, L. (2016). Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. Nat. Rev. Clin. Oncol. *13*, 674–690.

Bottai, G., Raschioni, C., Székely, B., Tommaso, L.D., Szász, A.M., Losurdo, A., Győrffy, B., Ács, B., Torrisi, R., Karachaliou, N., et al. (2016). AXL-associated tumor inflammation as a poor prognostic signature in chemotherapy-treated triple-negative breast cancer patients. NPJ Breast Cancer *2*, 16033.

Canel, M., Serrels, A., Frame, M.C., and Brunton, V.G. (2013). E-cadherin-integrin crosstalk in cancer invasion and metastasis. J. Cell Sci. *126*, 393–401.

Cerami, E.G., Gross, B.E., Demir, E., Rodchenkov, I., Babur, O., Anwar, N., Schultz, N., Bader, G.D., and Sander, C. (2011). Pathway commons, a web resource for biological pathway data. Nucleic Acids Res. *39*, D685–D690.

Chen, H.Y., Yang, Y.M., Stevens, B.M., and Noble, M. (2013). Inhibition of redox/Fyn/c-Cbl pathway function by Cdc42 controls tumour initiation capacity and tamoxifen sensitivity in basal-like breast cancer cells. EMBO Mol. Med. *5*, 723–736.

Chen, L.M., and Chai, K.X. (2002). Prostasin serine protease inhibits breast cancer invasiveness and is transcriptionally regulated by promoter DNA methylation. Int. J. Cancer *97*, 323–329.

Chen, X., Iliopoulos, D., Zhang, Q., Tang, Q., Greenblatt, M.B., Hatziapostolou, M., Lim, E., Tam, W.L., Ni, M., Chen, Y., et al. (2014). XBP1 promotes triple-negative breast cancer by controlling the HIF1alpha pathway. Nature *508*, 103–107.

Chen, Z., He, S., Zhan, Y., He, A., Fang, D., Gong, Y., Li, X., and Zhou, L. (2019). TGF-beta-induced transgelin promotes bladder cancer metastasis by regulating epithelial-mesenchymal transition and invadopodia formation. EBioMedicine *47*, 208–220.

Cheng, C.L., Thike, A.A., Tan, S.Y., Chua, P.J., Bay, B.H., and Tan, P.H. (2015). Expression of FGFR1 is an independent prognostic factor in triple-negative breast cancer. Breast Cancer Res. Treat. *151*, 99–111.

Cho, C.Y., Lee, K.T., Chen, W.C., Wang, C.Y., Chang, Y.S., Huang, H.L., Hsu, H.P., Yen, M.C., Lai, M.Z., and Lai, M.D. (2016). MST3 promotes proliferation and tumorigenicity through the VAV2/Rac1 signal axis in breast cancer. Oncotarget *7*, 14586–14604.

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.

Corsa, C.A.S., Brenot, A., Grither, W.R., Van Hove, S., Loza, A.J., Zhang, K., Ponik, S.M., Liu, Y., DeNardo, D.G., Eliceiri, K.W., et al. (2016). The action of discoidin domain receptor 2 in basal tumor cells and stromal cancer-associated fibroblasts is critical for breast cancer metastasis. Cell Rep. *15*, 2510–2523.

Costa, R.L.B., Han, H.S., and Gradishar, W.J. (2018). Targeting the PI3K/AKT/mTOR pathway in triple-negative breast cancer: a review. Breast Cancer Res. Treat. *169*, 397–406.

Datta, K., Li, J., Bhattacharya, R., Gasparian, L., Wang, E., and Mukhopadhyay, D. (2004). Protein kinase C zeta transactivates hypoxia-inducible factor alpha by promoting its association with p300 in renal cancer. Cancer Res. *64*, 456–462.

Delaloge, S., and DeForceville, L. (2017). Targeting PI3K/AKT pathway in triple-negative breast cancer. Lancet Oncol. *18*, 1293–1294.

Drabsch, Y., and ten Dijke, P. (2011). TGF-beta signaling in breast cancer cell invasion and bone metastasis. J. Mammary Gland Biol. Neoplasia *16*, 97–108.

Duhachek-Muggy, S., Qi, Y., Wise, R., Alyahya, L., Li, H., Hodge, J., and Zolkiewska, A. (2017). Metalloprotease-disintegrin ADAM12 actively promotes the stem cell-like phenotype in claudin-low breast cancer. Mol. Cancer *16*, 32.

Foulkes, W.D., Smith, I.E., and Reis-Filho, J.S. (2010). Triple-negative breast cancer. New Engl. J. Med. *363*, 1938–1948.

Gay, C.M., Balaji, K., and Byers, L.A. (2017). Giving AXL the axe: targeting AXL in human malignancy. Br. J. Cancer *116*, 415–423.

Goto, Y., Zeng, L., Yeom, C.J., Zhu, Y., Morinibu, A., Shinomiya, K., Kobayashi, M., Hirota, K., Itasaka, S., Yoshimura, M., et al. (2015). UCHL1 provides diagnostic and antimetastatic strategies due to its deubiquitinating effect on HIF-1α. Nat. Commun. *6*, 1–13.

Goyette, M.-A., Duhamel, S., Aubert, L., Pelletier, A., Savage, P., Thibault, M.-P., Johnson, R.M., Carmeliet, P., Basik, M., Gaboury, L., et al. (2018). The receptor tyrosine kinase AXL is required at multiple steps of the metastatic cascade during HER2-positive breast cancer progression. Cell Rep. *23*, 1476–1490.

Hamson, E.J., Keane, F.M., Tholen, S., Schilling, O., and Gorrell, M.D. (2014). Understanding fibroblast activation protein (FAP): substrates, activities, expression and targeting for cancer therapy. Proteomics Clin. Appl. *8*, 454–463.

Handa, T., Katayama, A., Yokobori, T., Yamane, A., Horiguchi, J., Kawabata-Iwakawa, R., Rokudai, S., Bao, P., Gombodorj, N., Altan, B., et al. (2017). Caspase14 expression is associated with triple negative phenotypes and cancer stem cell marker expression in breast cancer patients. J. Surg. Oncol. *116*, 706–715.

Haritos, C., Michaelidou, K., Mavridis, K., Missitzis, I., Ardavanis, A., Griniatsos, J., and Scorilas, A. (2018). Kallikrein-related peptidase 6 (KLK6) expression differentiates tumor subtypes and predicts clinical outcome in breast cancer patients. Clin. Exp. Med. *18*, 203–213.

Herschkowitz, J.I., Simin, K., Weigman, V.J., Mikaelian, I., Usary, J., Hu, Z., Rasmussen, K.E., Jones, L.P., Assefnia, S., Chandrasekharan, S., et al. (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol. *8*, R76.

Hoare, S., Hoare, K., Reinhard, M.K., Lee, Y.J., Oh, S.P., and May, W.S., Jr. (2008). Tnk1/Kos1 knockout mice develop spontaneous tumors. Cancer Res. *68*, 8723–8732.

Holland, S.J., Pan, A., Franci, C., Hu, Y., Chang, B., Li, W., Duan, M., Torneros, A., Yu, J., Heckrodt, T.J., et al. (2010). R428, a selective small molecule inhibitor of Axl kinase, blocks tumor spread and prolongs survival in models of metastatic breast cancer. Cancer Res. *70*, 1544–1554.

Hsu, Y.H., Yao, J., Chan, L.C., Wu, T.J., Hsu, J.L., Fang, Y.F., Wei, Y., Wu, Y., Huang, W.C., Liu, C.L., et al. (2014). Definition of PKC-alpha, CDK6, and MET as therapeutic targets in triple-negative breast cancer. Cancer Res. *74*, 4822–4835.

Hu, L., Duan, Y.T., Li, J.F., Su, L.P., Yan, M., Zhu, Z.G., Liu, B.Y., and Yang, Q.M. (2014). Biglycan enhances gastric cancer invasion by activating FAK signaling pathway. Oncotarget *5*, 1885–1896.

Huang, Y., Li, G., Wang, K., Mu, Z., Xie, Q., Qu, H., Lv, H., and Hu, B. (2018). Collagen type VI alpha 3 chain promotes epithelial-mesenchymal transition in bladder cancer cells via transforming

growth factor beta (TGF-beta)/Smad pathway. Med. Sci. Monitor *24*, 5346–5354.

Huang, Y., Wang, S., and Kelly, T. (2004). Seprase promotes rapid tumor growth and increased microvessel density in a mouse model of human breast cancer. Cancer Res. *64*, 2712–2716.

Humphries, B., Wang, Z., Oom, A.L., Fisher, T., Tan, D., Cui, Y., Jiang, Y., and Yang, C. (2014). MicroRNA-200b targets protein kinase Calpha and suppresses triple-negative breast cancer metastasis. Carcinogenesis *35*, 2254–2263.

Jamdade, V.S., Sethi, N., Mundhe, N.A., Kumar, P., Lahkar, M., and Sinha, N. (2015). Therapeutic targets of triple-negative breast cancer: a review. Br. J. Pharmacol. *172*, 4228–4237.

Janda, E., Lehmann, K., Killisch, I., Jechlinger, M., Herzig, M., Downward, J., Beug, H., and Grünert, S. (2002). Ras and TGFβ cooperatively regulate epithelial cell plasticity and metastasis. J. Cell Biol. *156*, 299–314.

Jezierska, A., and Motyl, T. (2009). Matrix metalloproteinase-2 involvement in breast cancer progression: a mini-review. Med. Sci. Monitor *15*, RA32–40.

Jia, J., Martin, T.A., Ye, L., and Jiang, W.G. (2014). FAP-α (Fibroblast activation protein-α) is involved in the control of human breast cancer cell line growth and motility via the FAK pathway. BMC Cell Biol. *15*, 16.

Jiang, W.G., Davies, G., Martin, T.A., Parr, C., Watkins, G., Mason, M.D., and Mansel, R.E. (2006). Expression of membrane type-1 matrix metalloproteinase, MT1-MMP in human breast cancer and its impact on invasiveness of breast cancer cells. Int. J. Mol. Med. *17*, 583–590.

Klijn, C., Durinck, S., Stawiski, E.W., Haverty, P.M., Jiang, Z., Liu, H., Degenhardt, J., Mayba, O., Gnad, F., Liu, J., et al. (2015). A comprehensive transcriptional portrait of human cancer cell lines. Nat. Biotechnol. *33*, 306–312.

Klose, R., Adam, M.G., Weis, E.-M., Moll, I., Wüstehube-Lausch, J., Tetzlaff, F., Oka, C., Ehrmann, M., and Fischer, A. (2018). Inactivation of the serine protease HTRA1 inhibits tumor growth by deregulating angiogenesis. Oncogene *37*, 4260.

Korshunov, V.A. (2012). Axl-dependent signaling: a clinical update. Clin. Sci. *122*, 361–368.

Kuo, Y.-H., Chiang, E.-P.I., Chao, C.-Y., Rodriguez, R.L., Chou, P.-Y., Tsai, S.-Y., Pai, M.-H., and Tang, F.-Y. (2017). Dual inhibition of key proliferation signaling pathways in triple-negative breast cancer cells by a novel derivative of Taiwanin A. Mol. Cancer Ther. *16*, 480–493.

Lapek, J.D., Greninger, P., Morris, R., Amzallag, A., Pruteanu-Malinici, I., Benes, C.H., and Haas, W. (2017). Detection of dysregulated protein-association networks by high-throughput proteomics predicts cancer vulnerabilities. Nat. Biotechnol. *35*, 983.

Lawrence, R.T., Perez, E.M., Hernández, D., Miller, C.P., Haas, K.M., Irie, H.Y., Lee, S.-I., Blau, C.A., and Villén, J. (2015). The proteomic landscape of triple-negative breast cancer. Cell Rep. *11*, 630–644.

Leconet, W., Chentouf, M., du Manoir, S., Chevalier, C., Sirvent, A., Aït-Arsa, I., Busson, M., Jarlier, M., Radosevic-Robin, N., Theillet, C., et al. (2017). Therapeutic activity of anti-AXL antibody against triple-negative breast cancer patient-derived xenografts and metastasis. Clin.Cancer Res. *23*, 2806–2816.

Lee, J.W., Park, J.A., Kim, S.H., Seo, J.H., Lim, K.J., Jeong, J.W., Jeong, C.H., Chun, K.H., Lee, S.K., Kwon, Y.G., et al. (2007). Protein kinase C-delta regulates the stability of hypoxia-inducible factor-1 alpha under hypoxia. Cancer Sci. *98*, 1476–1481.

Lehmann, B.D., Bauer, J.A., Chen, X., Sanders, M.E., Chakravarthy, A.B., Shyr, Y., and Pietenpol, J.A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J. Clin. Invest. *121*, 2750–2767.

Lehner, A., Magdolen, V., Schuster, T., Kotzsch, M., Kiechle, M., Meindl, A., Sweep, F.C.G.J., Span, P.N., and Gross, E. (2013). Downregulation of serine protease HTRA1 is associated with poor survival in breast cancer. PLoS One *8*, e60359.

Liberti, S., Sacco, F., Calderone, A., Perfetto, L., Iannuccelli, M., Panni, S., Santonico, E., Palma, A., Nardozza Aurelio, P., Castagnoli, L., et al. (2012). HuPho: the human phosphatase portal. FEBS J. *280*, 379–387.

Lin, W., Huang, J., Yuan, Z., Feng, S., Xie, Y., and Ma, W. (2017). Protein kinase C inhibitor chelerythrine selectively inhibits proliferation of triple-negative breast cancer cells. Sci. Rep. *7*, 2022.

Liu, S., Gonzalez Prieto, R., Zhang, M., Geurink, P.P., Kooij, R., Iyengar, P.V., van Dinther, M., Bos, E., Zhang, X., Le Devedec, S.E., et al. (2019). Deubiquitinase activity profiling identifies UCHL1 as a candidate oncoprotein that promotes TGFbeta-induced breast cancer metastasis. Clin. Cancer Res. https://doi.org/10.1158/1078-0432.CCR-19-1373.

Liu, Z.J., Semenza, G.L., and Zhang, H.F. (2015). Hypoxia-inducible factor 1 and breast cancer metastasis. J. Zhejiang Univ. Sci. B *16*, 32–43.

Lu, P., Weaver, V.M., and Werb, Z. (2012). The extracellular matrix: a dynamic niche in cancer progression. J. Cell Biol. *196*, 395–406.

Manning, G., Whyte, D.B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. Science *298*, 1912–1934.

Massagué, J. (2008). TGFβ in cancer. Cell *134*, 215–230.

Matutino, A., Amaro, C., and Verma, S. (2018). CDK4/6 inhibitors in breast cancer: beyond hormone receptor-positive HER2-negative disease. Ther. Adv. Med. Oncol. *10*, 1758835918818346, https://doi.org/10.1177/1758835918818346.

Mertins, P., Mani, D.R., Ruggles, K.V., Gillette, M.A., Clauser, K.R., Wang, P., Wang, X., Qiao, J.W., Cao, S., Petralia, F., et al. (2016). Proteogenomics connects somatic mutations to signalling in breast cancer. Nature *534*, 55–62.

Metz, K.S., Deoudes, E.M., Berginski, M.E., Jimenez-Ruiz, I., Aksoy, B.A., Hammerbacher, J., Gomez, S.M., and Phanstiel, D.H. (2018). Coral: clear and customizable visualization of human kinome data. Cell Syst. *7*, 347–350.e1.

Mrozik, K.M., Blaschuk, O.W., Cheong, C.M., Zannettino, A.C.W., and Vandyke, K. (2018). N-cadherin in cancer metastasis, its emerging role in haematological malignancies and potential as a therapeutic target in cancer. BMC Cancer *18*, 939.

Muz, B., de la Puente, P., Azab, F., and Azab, A.K. (2015). The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. Hypoxia *3*, 83–92.

Neophytou, C., Boutsikos, P., and Papageorgis, P. (2018). Molecular mechanisms and emerging therapeutic targets of triple-negative breast cancer metastasis. Front. Oncol. *8*, 31.

Neve, R.M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F.L., Fevr, T., Clark, L., Bayani, N., Coppe, J.-P., Tong, F., et al. (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell *10*, 515–527.

Nielsen, T.O., Hsu, F.D., Jensen, K., Cheang, M., Karaca, G., Hu, Z., Hernandez-Boussard, T., Livasy, C., Cowan, D., Dressler, L., et al. (2004). Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin.Cancer Res. *10*, 5367–5374.

Ovcaricek, T., Frkovic, S.G., Matos, E., Mozina, B., and Borstnar, S. (2010). Triple negative breast cancer – prognostic factors and survival. Radiol.Oncol. *45*, 46–52.

Padua, D., Zhang, X.H.F., Wang, Q., Nadal, C., Gerald, W.L., Gomis, R.R., and Massagué, J. (2008). TGFβ primes breast tumors for lung metastasis seeding through angiopoietin-like 4. Cell *133*, 66–77.

Pandey, G., Borcherding, N., Kolb, R., Kluz, P., Li, W., Sugg, S., Zhang, J., Lai, D.A., and Zhang, W. (2019). ROR1 potentiates FGFR signaling in basal-like breast cancer. Cancers *11*, 718.

Papageorgis, P., Lambert, A.W., Ozturk, S., Gao, F., Pan, H., Manne, U., Alekseyev, Y., Thiagalingam, A., Abdolmaleky, H.M., Lenburg, M., et al. (2010). Smad signaling is required for the maintenance of epigenetic gene silencing during breast cancer progression. Cancer Res. *70*, 968–978.

Parker, J.S., Mullins, M., Cheang, M.C.U., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X., Hu, Z., et al. (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. J. Clin. Oncol. *27*, 1160–1167.

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. *47*, D442–D450.

Pernas, S., Tolaney, S.M., Winer, E.P., and Goel, S. (2018). CDK4/6 inhibition in breast cancer: current practice and future directions. Ther. Adv. Med. Oncol. *10*, 1758835918786451, https://doi.org/10.1177/1758835918786451.

Perou, C.M., Sørlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross,

D.T., Johnsen, H., Akslen, L.A., et al. (2000). Molecular portraits of human breast tumours. Nature *406*, 747–752.

Ponente, M., Campanini, L., Cuttano, R., Piunti, A., Delledonne, G.A., Coltella, N., Valsecchi, R., Villa, A., Cavallaro, U., Pattini, L., et al. (2017). PML promotes metastasis of triple-negative breast cancer through transcriptional regulation of HIF1A target genes. JCI Insight *2*, e87380.

Prat, A., Parker, J.S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J.I., He, X., and Perou, C.M. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res. *12*, R68.

Principe, D.R., Doll, J.A., Bauer, J., Jung, B., Munshi, H.G., Bartholin, L., Pasche, B., Lee, C., and Grippo, P.J. (2014). TGF-β: duality of function between tumorprevention and carcinogenesis. J. Natl. Cancer Inst. *106*, djt369.

Quereda, V., Bayle, S., Vena, F., Frydman, S.M., Monastyrskyi, A., Roush, W.R., and Duckett, D.R. (2019). Therapeutic targeting of CDK12/CDK13 in triple-negative breast cancer. Cancer Cell *36*, 545–558.e7.

Ramachandran, A., Vizan, P., Das, D., Chakravarty, P., Vogt, J., Rogers, K.W., Muller, P., Hinck, A.P., Sapkota, G.P., and Hill, C.S. (2018). TGF-beta uses a novel mode of receptor activation to phosphorylate SMAD1/5 and induce epithelial-to-mesenchymal transition. eLife *7*, e31756.

Ravasi, T., Suzuki, H., Cannistraci, C.V., Katayama, S., Bajic, V.B., Tan, K., Akalin, A., Schmeier, S., Kanamori-Katayama, M., Bertin, N., et al. (2010). An atlas of combinatorial transcriptional regulation in mouse and man. Cell *140*, 744–752.

Rawlings, N.D., Barrett, A.J., Thomas, P.D., Huang, X., Bateman, A., and Finn, R.D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res. *46*, D624–D632.

Rouette, A., Trofimov, A., Haberl, D., Boucher, G., Lavallee, V.P., D'Angelo, G., Hebert, J., Sauvageau, G., Lemieux, S., and Perreault, C. (2016). Expression of immunoproteasome genes is regulated by cell-intrinsic and -extrinsic factors in human cancers. Sci. Rep. *6*, 34019.

Sacco, F., Silvestri, A., Posca, D., Pirro, S., Gherardini, P.F., Castagnoli, L., Mann, M., and Cesareni, G. (2016). Deep proteomics of breast cancer cells reveals that metformin rewires signaling networks away from a pro-growth state. Cell Syst. *2*, 159–171.

Sadahiro, H., Kang, K.-D., Gibson, J.T., Minata, M., Yu, H., Shi, J., Chhipa, R., Chen, Z., Lu, S., Simoni, Y., et al. (2018). Activation of the receptor tyrosine kinase AXL regulates the immune microenvironment in glioblastoma. Cancer Res. *78*, 3002–3013.

Sheng, W., Wang, G., La Pierre, D.P., Wen, J., Deng, Z., Wong, C.K., Lee, D.Y., and Yang, B.B. (2006). Versican mediates mesenchymal-epithelial transition. Mol. Biol. Cell *17*, 2009–2020.

Song, W., Hwang, Y., Youngblood, V.M., Cook, R.S., Balko, J.M., Chen, J., and Brantley-Sieders, D.M. (2017). Targeting EphA2 impairs cell cycle progression and growth of basal-like/triple-negative breast cancers. Oncogene *36*, 5620–5630.

Song, W., Ma, Y., Wang, J., Brantley-Sieders, D., and Chen, J. (2014). JNK signaling mediates EPHA2-dependent tumor cell proliferation, motility, and cancer stem cell-like properties in non-small cell lung cancer. Cancer Res. *74*, 2444–2454.

Sørlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., et al. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc. Natl. Acad. Sci. U S A *98*, 10869–10874.

Suthe, S.R., Yao, H.P., Weng, T.H., Hu, C.Y., Feng, L., Wu, Z.G., and Wang, M.H. (2018). RON receptor tyrosine kinase as a therapeutic target for eradication of triple-negative breast cancer: efficacy of anti-RON ADC Zt/g4-MMAE. Mol. Cancer Ther. *17*, 2654–2664.

Talvensaari-Mattila, A., Pääkkö, P., and Turpeenniemi-Hujanen, T. (2003). Matrix metalloproteinase-2 (MMP-2) is associated with survival in breast carcinoma. Br. J. Cancer *89*, 1270–1275.

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.e10.

Tsai, Y.-F., Tseng, L.-M., Hsu, C.-Y., Yang, M.-H., Chiu, J.-H., and Shyr, Y.-M. (2017). Brain-derived neurotrophic factor (BDNF) -TrKB signaling modulates cancer-endothelial cells interaction and affects the outcomes of triple negative breast cancer. PLoS One *12*, e0178173.

Wahdan-Alaswad, R., Harrell, J.C., Fan, Z., Edgerton, S.M., Liu, B., and Thor, A.D. (2016). Metformin attenuates transforming growth factor beta (TGF-β) mediated oncogenesis in mesenchymal stem-like/claudin-low triple negative breast cancer. Cell Cycle *15*, 1046–1059.

Wang, B. (2011). Cancer cells exploit the Eph-ephrin system to promote invasion and metastasis: tales of unwitting partners. Sci. Signal. *4*, pe28.

Wang, N., Eckert, K.A., Zomorrodi, A.R., Xin, P., Pan, W., Shearer, D.A., Weisz, J., Maranus, C.D., and Clawson, G.A. (2012). Down-regulation of HtrA1 activates the epithelial-mesenchymal transition and ATM DNA damage response pathways. PLoS One *7*, e39446.

Wee, Z.N., Yatim, S.M., Kohlbauer, V.K., Feng, M., Goh, J.Y., Bao, Y., Lee, P.L., Zhang, S., Wang, P.P., Lim, E., et al. (2015). IRAK1 is a therapeutic target that drives breast cancer metastasis and resistance to paclitaxel. Nat. Commun. *6*, 8746.

Wilson, C., Ye, X., Pham, T., Lin, E., Chan, S., McNamara, E., Neve, R.M., Belmont, L., Koeppen, H., Yauch, R.L., et al. (2014). AXL inhibition sensitizes mesenchymal cancer cells to antimitotic drugs. Cancer Res. *74*, 5878–5890.

Wong, C.C.-L., Gilkes, D.M., Zhang, H., Chen, J., Wei, H., Chaturvedi, P., Fraley, S.I., Wong, C.-M., Khoo, U.-S., Ng, I.O.-L., et al. (2011). Hypoxia-inducible factor 1 is a master regulator of breast cancer metastatic niche formation. Proc. Natl. Acad. Sci. U S A *108*, 16369–16374.

Xia, J., Wang, F., Wang, L., and Fan, Q. (2013). Elevated serine protease HtrA1 inhibits cell proliferation, reduces invasion, and induces apoptosis in esophageal squamous cell carcinoma by blocking the nuclear factor-κB signaling pathway. Tumor Biol. *34*, 317–328.

Xie, X., Kaoud, T.S., Edupuganti, R., Zhang, T., Kogawa, T., Zhao, Y., Chauhan, G.B., Giannoukos, D.N., Qi, Y., Tripathy, D., et al. (2017). c-Jun N-terminal kinase promotes stem cell phenotype in triple-negative breast cancer through upregulation of Notch1 via activation of c-Jun. Oncogene *36*, 2599–2608.

Xie, Y.G., Yu, Y., Hou, L.K., Wang, X., Zhang, B., and Cao, X.C. (2016). FYN promotes breast cancer progression through epithelial-mesenchymal transition. Oncol. Rep. *36*, 1000–1006.

Xu, X., Zhang, L., He, X., Zhang, P., Sun, C., Xu, X., Lu, Y., and Li, F. (2018). TGF-beta plays a vital role in triple-negative breast cancer (TNBC) drug-resistance through regulating stemness, EMT and apoptosis. Biochem.Biophys.Res. Commun. *502*, 160–165.

Yamamoto, T., Kanaya, N., Somlo, G., and Chen, S. (2019). Synergistic anti-cancer activity of CDK4/6 inhibitor palbociclib and dual mTOR kinase inhibitor MLN0128 in pRb-expressing ER-negative breast cancer. Breast Cancer Res. Treat. *174*, 615–625.

Yamamoto, Y., Ibusuki, M., Okumura, Y., Kawasoe, T., Kai, K., Iyama, K., and Iwase, H. (2008). Hypoxia-inducible factor 1alpha is closely linked to an aggressive phenotype in breast cancer. Breast Cancer Res. Treat. *110*, 465–475.

Yang, F., Li, J.Y., Yin, Q.N., Yang, K., Dong, S.N., Bai, L.J., Liu, P., and Tong, X.W. (2015). Human kallikrein 5 as a novel prognostic biomarker for triple-negative breast cancer: tissue expression analysis and relationship with disease course. Genet.Mol. Res. *14*, 9655–9666.

Yule, M., Davidsen, K., Bloe, M., Hodneland, L., Engelsen, A., Lie, M., Bougnaud, S., D'Mello, S., Aguilera, K., Ahmed, L., et al. (2018). Combination of bemcentinib (BGB324): a first-in-class selective oral AXL inhibitor, with pembrolizumab in patients with triple negative breast cancer and adenocarcinoma of the lung. J. Clin. Oncol. *36*, TPS43.

Zhang, B., Wang, J., Wang, X., Zhu, J., Liu, Q., Shi, Z., Chambers, M.C., Zimmerman, L.J., Shaddox, K.F., Kim, S., et al. (2014). Proteogenomic characterization of human colon and rectal cancer. Nature *513*, 382–387.

Zhang, K., Corsa, C.A., Ponik, S.M., Prior, J.L., Piwnica-Worms, D., Eliceiri, K.W., Keely, P.J., and Longmore, G.D. (2013). The collagen receptor discoidin domain receptor 2 stabilizes SNAIL1 to facilitate breast cancer metastasis. Nat. Cell Biol. *15*, 677–687.

Zhang, P., Sun, Y., and Ma, L. (2015). ZEB1: at the crossroads of epithelial-mesenchymal transition, metastasis and therapy resistance. Cell Cycle *14*, 481–487.

**iScience, Volume *23***

## Supplemental Information

# Comprehensive Proteomic Characterization Reveals Subclass-Specific Molecular Aberrations within Triple-negative Breast Cancer

**Max Kosok, Asfa Alli-Shaik, Boon Huat Bay, and Jayantha Gunaratne**

# Supplemental Figures

Fig. S1

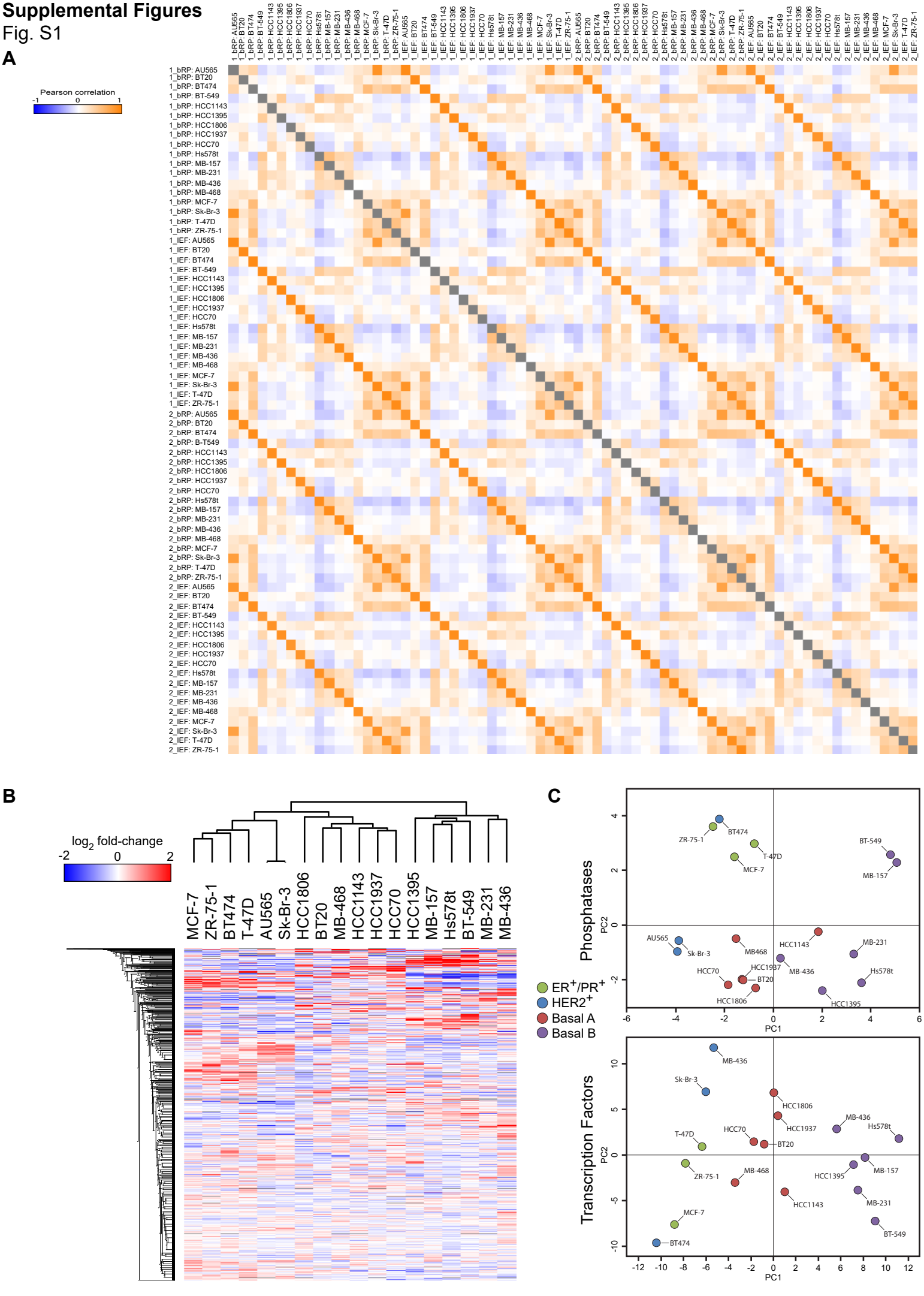
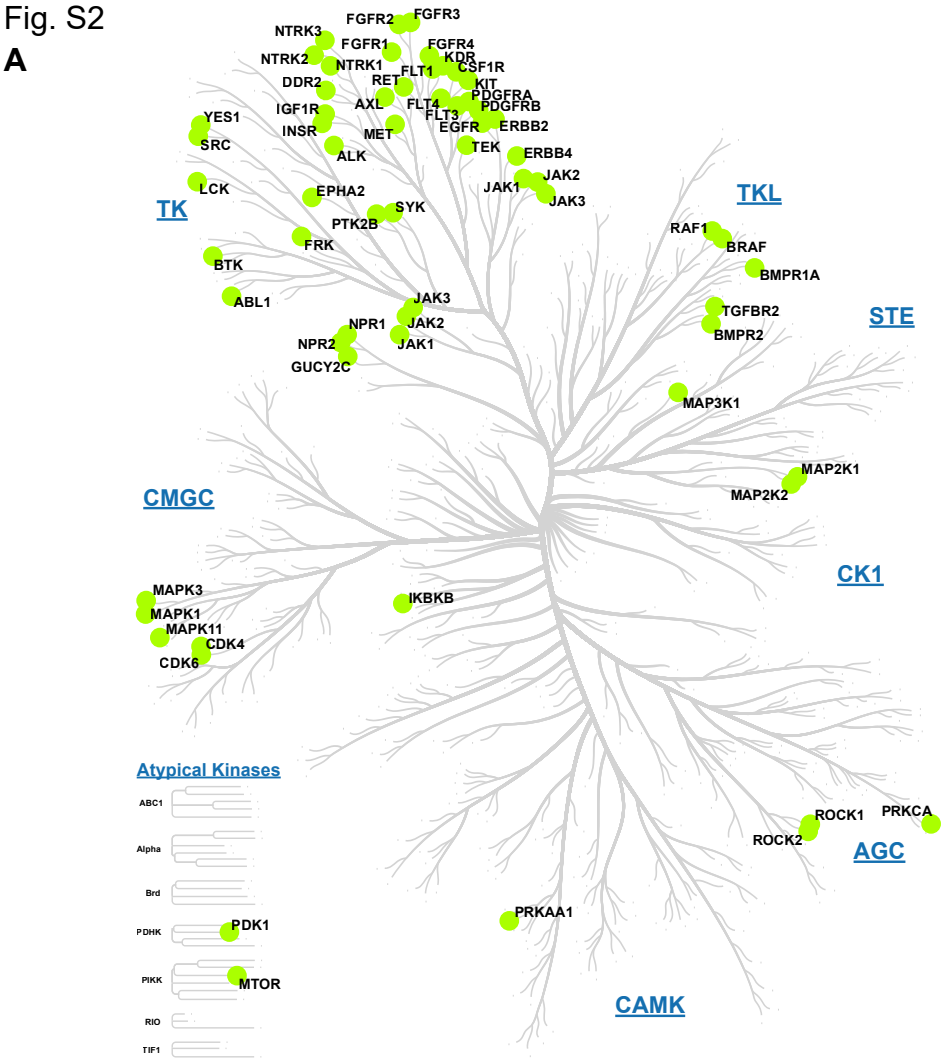
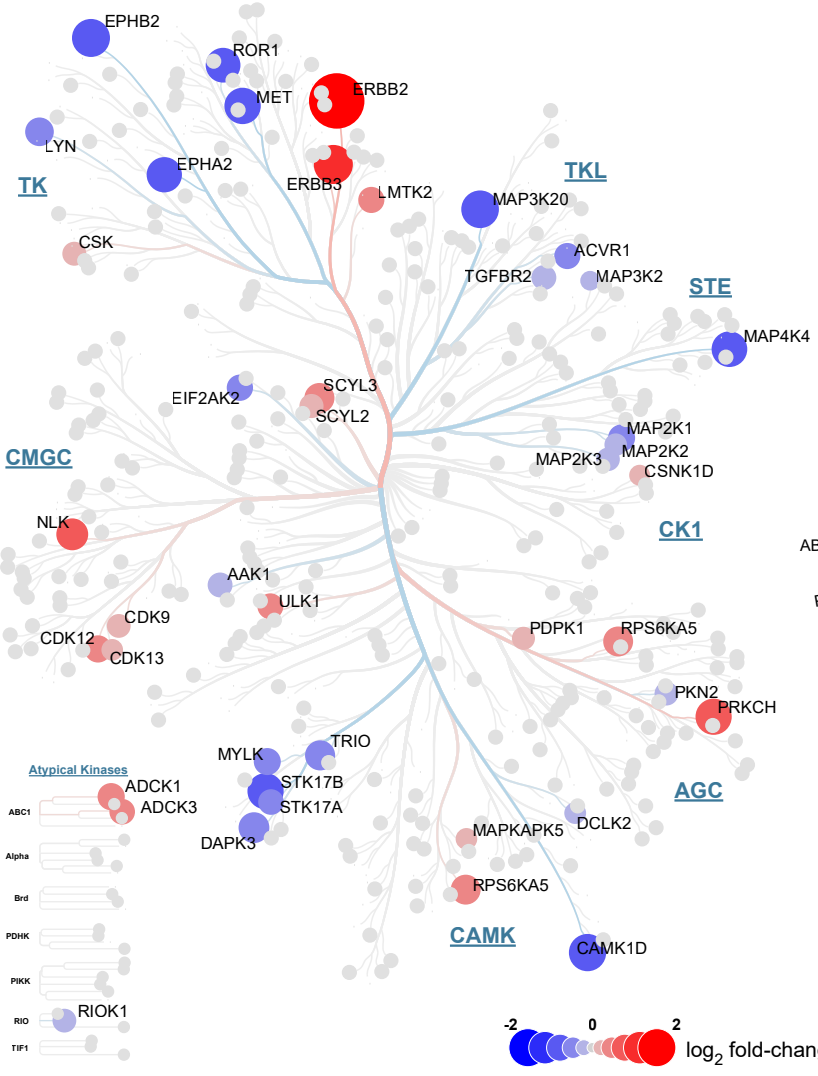
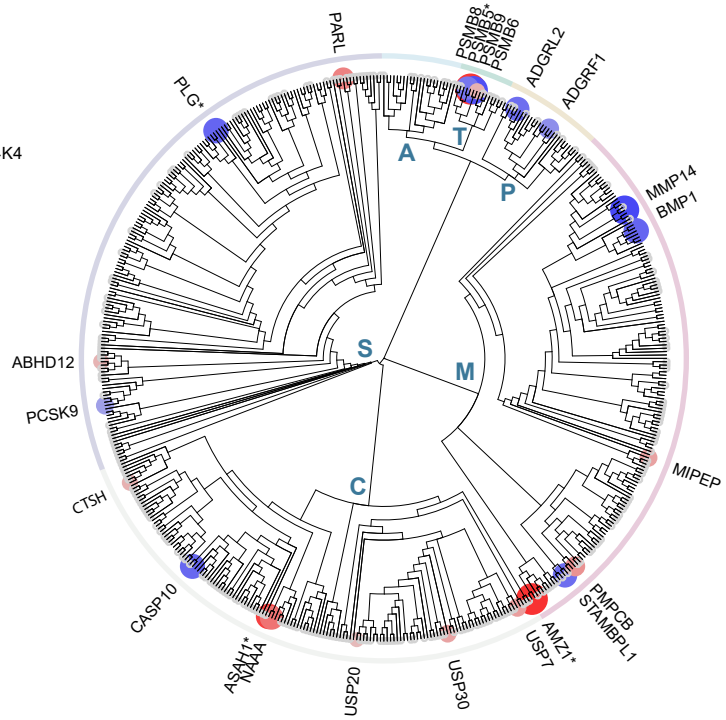


***Figure S1. Proteome profiles of breast cancer cell line panel.*** Related to Figures 1 & 2, and Table S1. (**A**) Pearson correlation of $log_2$ fold-change plotted as heat map for all biological replicates and between the two different peptide fractionation methods, IEF and bRP-LC. (**B**) Unsupervised whole proteome-based clustering of the breast cancer cell line panel. Normalized median $log_2$ fold-change of proteins is represented as heat map. (**C**) Principal component analysis for phosphatases or transcription factors showed no distinct distribution pattern compared to kinases (Figure 2C) and proteases (Figure 2D). ER/$PR^+$ and $HER2^+$ represent luminal subtype, basal A and basal B represent TNBC subtype.

Fig. S2


A
FGFR2
FGFR3
NTRK3
FGFR1
FGFR4
NTRK2
KDR
NTRK1
CSF1R
FLT1
RET
KIT
DDR2
PDGFRA
AXL
FLT4
PDGFRB
IGF1R
FLT3
EGFR
ERBB2
INSR
MET
YES1
TEK
SRC
ERBB4
ALK
JAK2
LCK
JAK1
JAK3
EPHA2
TK
SYK
PTK2B
TKL
FRK
RAF1
BTK
BRAF
BMPR1A
ABL1
JAK3
TGFBR2
STE
NPR1
JAK2
BMPR2
NPR2
JAK1
GUCY2C
MAP3K1
MAP2K1
CMGC
MAP2K2
CK1
MAPK3
IKBKB
MAPK1
MAPK11
CDK4
CDK6
Atypical Kinases
ABC1
Alpha
Brd
ROCK1
PRKCA
ROCK2
AGC
PDHK
PDK1
PRKAA1
PIKK
MTOR
CAMK
RIO
TIF1


B Kinases
EPHB2
ROR1
MET
ERBB2
LYN
EPHA2
TK
ERBB3
LMTK2
TKL
MAP3K20
CSK
ACVR1
TGFBR2
MAP3K2
STE
MAP4K4
EIF2AK2
SCYL3
SCYL2
MAP2K1
MAP2K2
MAP2K3
CSNK1D
CMGC
NLK
CK1
AAK1
ULK1
CDK9
CDK12
CDK13
PDPK1
RPS6KA5
PKN2
PRKCH
TRIO
MYLK
Atypical Kinases
ADCK1
STK17B
AGC
ADCK3
STK17A
DAPK3
MAPKAPK5
DCLK2
RPS6KA5
CAMK
CAMK1D
RIOK1


C Proteases
PARL
PSMB8*
PSMB5*
PSMB9
PSMB6
ADGRL2
ADGRF1
PLG*
A
T
P
MMP14
BMP1
S
M
ABHD12
PCSK9
MIPEP
C
CTSH
CASP10
PMPCB
STAMBPL1
AMZ1*
USP7
ASAH1*
NAAA
USP20
USP30

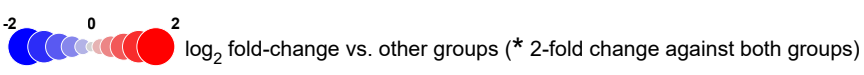

-2
0
2
log2 fold-change vs. other groups (* 2-fold change against both groups)

***Figure S2. Differentially expressed kinase and protease profiles.*** Related to Figure 3. (**A**) Key targets of approved drugs from www.drugbank.co (Dec 2019) mapped onto the human kinome tree. Illustration was created using CORAL online tool (Metz et al., 2018) based on the human kinome tree courtesy Cell Signaling Technology, Inc. (www.cellsignaling.com). (**B**) Differentially expressed kinases for the luminal breast cancer subtype as compared to both basal A and basal B subclasses mapped onto the human kinome tree. The node size and shade represents the median $\log_2$ fold-change of the kinases in the luminal subclass. Red represents upregulated kinases and blue denotes downregulated kinases in luminal as against the other two TNBC groups. Those kinases indicated with asterisk are differentially expressed by at least two-fold in luminal as against the other two TNBC subclasses. Illustration was created using CORAL online tool (Metz et al., 2018) based on the human kinome tree courtesy Cell Signaling Technology, Inc. (www.cellsignaling.com). (**C**) Differentially expressed proteases for the luminal breast cancer subtype as compared to both basal A and basal B subclasses mapped according to their protease families. The node size and shade represents the median $\log_2$ fold-change of the protease within the luminal subclass. Red represents upregulated and blue denotes downregulated proteases in luminal as against the other two TNBC groups. Nodes are scaled to the sizes represented in the scale bar in Figure S2B. Those proteases indicated with asterisk are differentially expressed by at least two-fold in luminal group as against the other two TNBC subclasses. The protease families include S – serine, C – cysteine, M – metallo, A – aspartate, T – threonine proteases, and P – peptidases.

Fig. S3

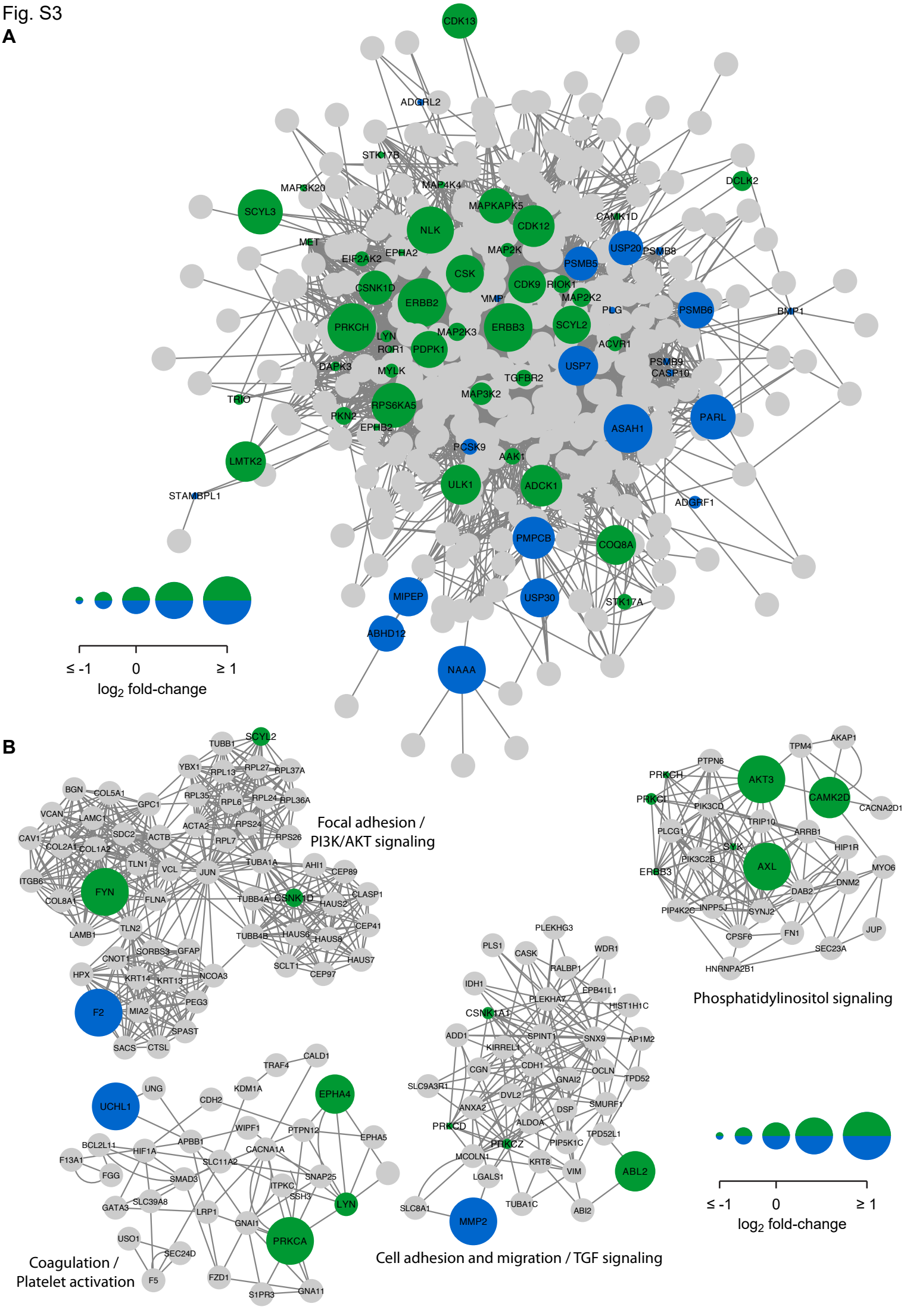

A
CDK13
ADGRL2
STK17B
MAP3K20
MAP4K4
MAPKAPK5
DCLK2
SCYL3
NLK
CDK12
CAMK1D
MET
EIF2AK2
EPHA2
MAP2K
USP20
PSMB8
PSMB5
CSK
CDK9
RIOK1
CSNK1D
ERBB2
MMP
MAP2K2
PLG
PSMB6
BMP1
PRKCH
ERBB3
SCYL2
LYN
MAP2K3
ACVR1
ROR1
PDPK1
USP7
PSMB9
CASP10
DAPK3
MYLK
TGFBR2
TNIK
MAP3K2
RPS6KA5
PKN2
ASAH1
PARL
EPHB2
PCSK9
AAK1
LMTK2
ULK1
ADCK1
STAMBPL1
ADGRF1
PMPCB
COQ8A
MIPEP
USP30
STK17A
ABHD12
NAAA
≤ -1
0
≥ 1
log2 fold-change
B
SCYL2
TUBB1
YBX1
RPL13
RPL27
RPL37A
BGN
COL5A1
GPC1
RPL35
RPL6
RPL24
RPL36A
VCAN
LAMC1
ACTA2
RPS24
CAV1
SDC2
ACTB
RPL7
RPS26
COL2A1
COL1A2
TLN1
VCL
JUN
TUBA1A
AHI1
CEP89
ITGB6
FYN
COL8A1
FLNA
TUBB4A
CSNK1D
HAUS2
CLASP1
CEP41
LAMB1
TLN2
TUBB4B
HAUS6
HAUS8
SORBS3
GFAP
CNOT1
SCLT1
HAUS7
CEP97
HPX
KRT14
KRT13
NCOA3
PEG3
F2
MIA2
SPAST
SACS
CTSL
Focal adhesion /
PI3K/AKT signaling
TPM4
AKAP1
PTPN6
PRKCH
AKT3
CAMK2D
PRKCI
CACNA2D1
PIK3CD
TRIP10
PLCG1
ARRB1
SYK
HIP1R
PIK3C2B
AXL
ERBB3
MYO6
DNM2
DAB2
PIP4K2C
INPP5J
SYNJ2
CPSF6
FN1
JUP
SEC23A
HNRNPA2B1
Phosphatidylinositol signaling
PLEKHG3
PLS1
CASK
WDR1
RALBP1
IDH1
EPB41L1
PLEKHA7
HIST1H1C
CSNK1A1
ADD1
SPINT1
SNX9
AP1M2
KIRREL1
CGN
CDH1
OCLN
TPD52
SLC9A3R1
GNAI2
DVL2
SMURF1
ANXA2
DSP
PRKCD
ALDOA
TPD52L1
PRKCZ
PIP5K1C
MCOLN1
KRT8
VIM
ABL2
LGALS1
SLC8A1
MMP2
TUBA1C
ABI2
Cell adhesion and migration / TGF signaling
CALD1
TRAF4
UNG
KDM1A
EPHA4
UCHL1
CDH2
WIPF1
PTPN12
EPHA5
APBB1
BCL2L11
HIF1A
CACNA1A
F13A1
SLC11A2
FGG
SMAD3
ITPKC
SNAP25
SSH3
LYN
GATA3
SLC39A8
LRP1
GNAI1
PRKCA
USO1
SEC24D
F5
FZD1
S1PR3
GNA11
Coagulation /
Platelet activation
≤ -1
0
≥ 1
log2 fold-change

***Figure S3. Kinase-protease-centric interaction network.*** Related to Figure 4. (**A**) Proximity network of differentially expressed kinases and proteases in the luminal subtype. The kinases are depicted as green nodes and proteases as blue nodes, and the size of the nodes represents the median $\log_2$ fold-change against the other two TNBC subclasses, with larger nodes representing increased fold-change and vice versa. The edges represent physical and functional interactions between proteins. (**B**) The top densely connected clusters in the basal B kinase-protease proximity network identified through MCODE graph clustering algorithm is shown. The sizes of the colored nodes correspond to the median $\log_2$ fold-change of the basal B subclass against both the luminal and the basal A subclass. The kinases and proteases are depicted as green and blue nodes, respectively, and the edges represent physical and functional interactions between proteins.

Fig. S4

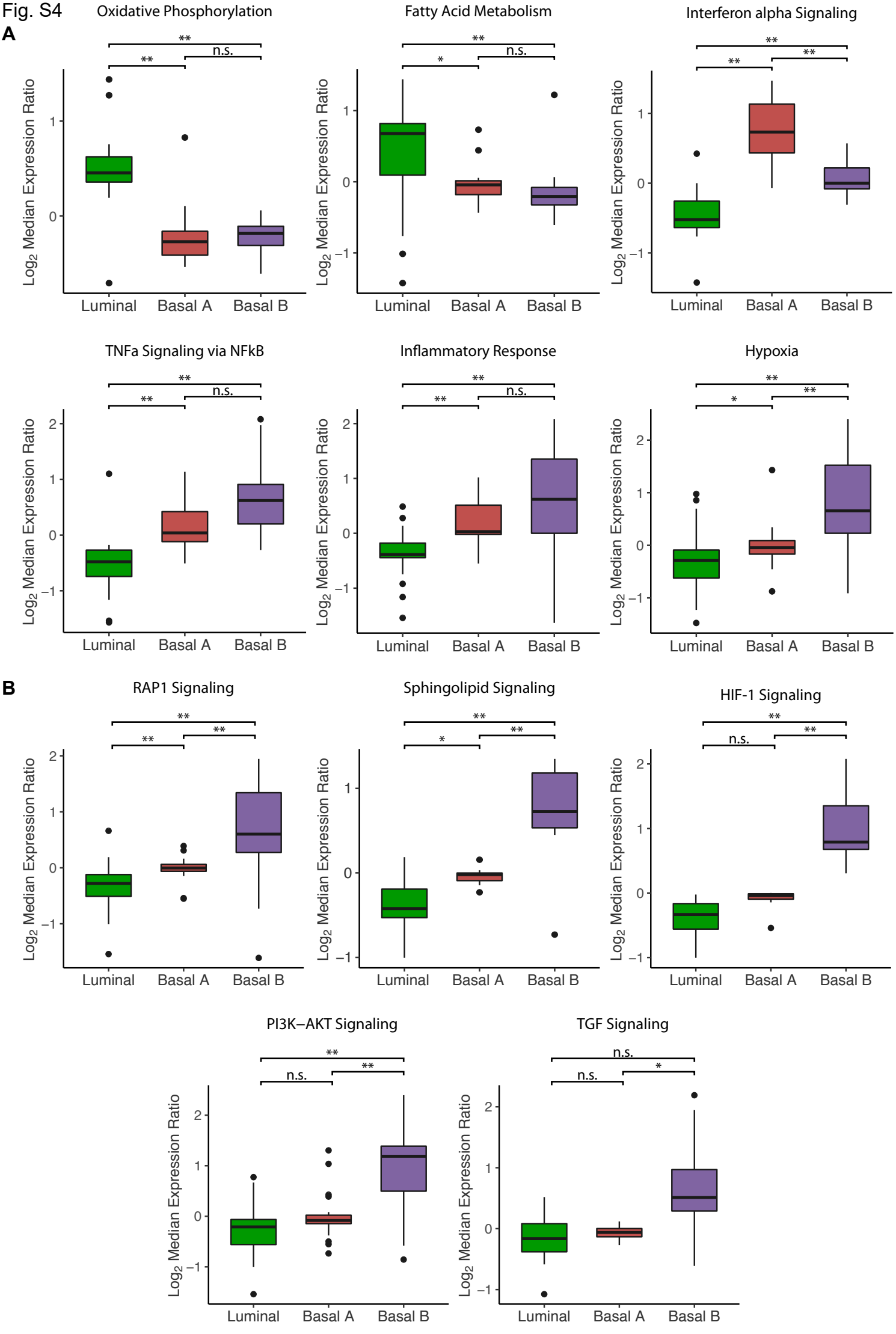

A
Oxidative Phosphorylation
Fatty Acid Metabolism
Interferon alpha Signaling
TNFa Signaling via NFkB
Inflammatory Response
Hypoxia
B
RAP1 Signaling
Sphingolipid Signaling
HIF-1 Signaling
PI3K–AKT Signaling
TGF Signaling
Log2 Median Expression Ratio
Luminal
Basal A
Basal B
**
*
n.s.

***Figure S4. Functional enrichment among breast cancer subgroups.*** Related to Figure 4 and Table S2. (**A**) Distribution of protein ratio within each breast cancer subgroups for the indicated cancer hallmark gene sets (*, p value ≤0.05; **, p value ≤0.01; n.s., not significant – using KS test). (**B**) Distribution of protein ratio within each breast cancer subtype for the indicated KEGG signaling pathways (*, p value ≤0.05; **, p value ≤0.01; n.s., not significant – using KS test).

Fig. S5

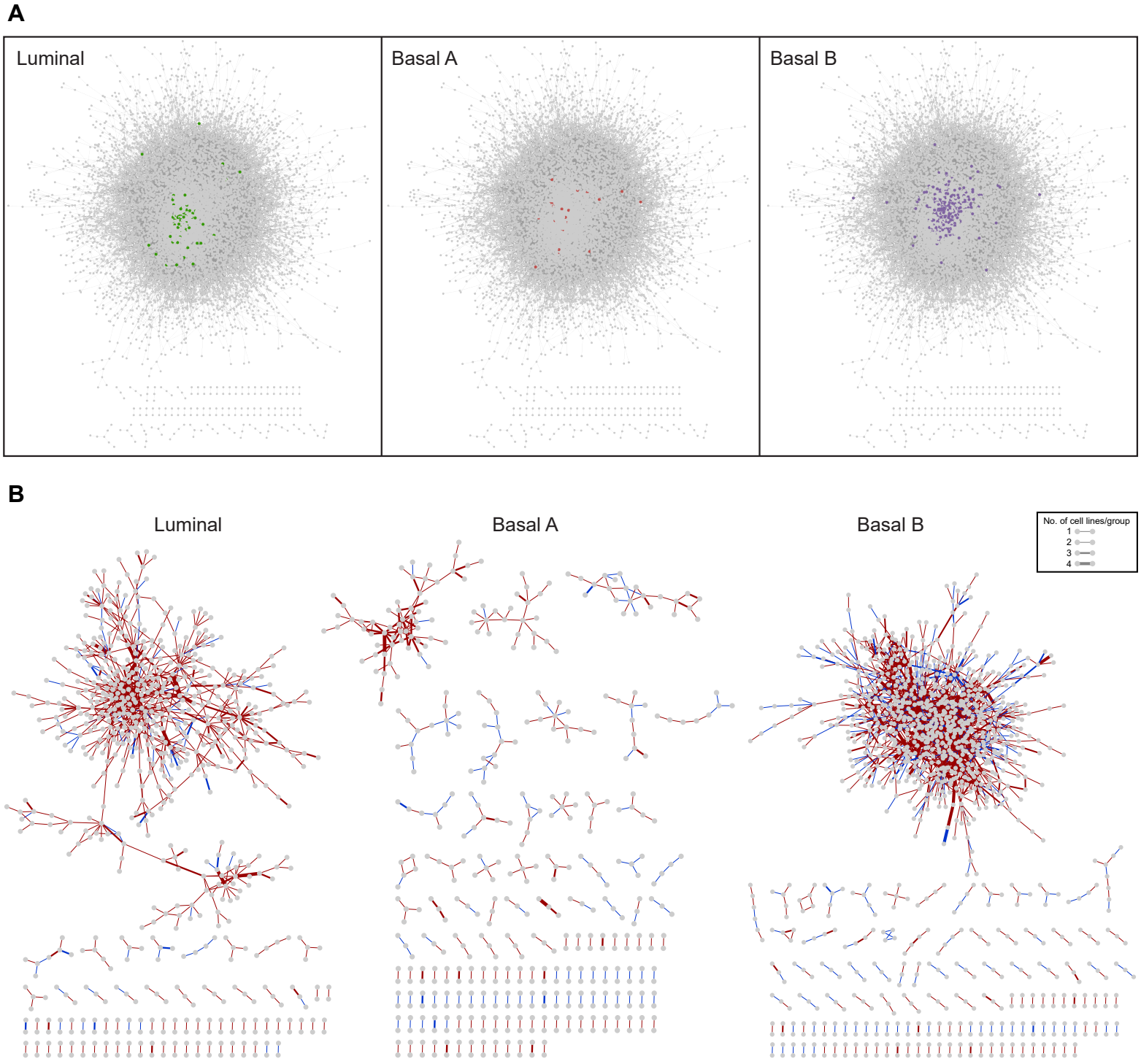



***Figure S5. Protein association dysregulations in different breast cancer subgroups.*** Related to Figure 5, and Tables S3, S4 and S5. (**A**) Proteins with dysregulated associations in each breast cancer subgroup shown as colored nodes mapped onto the background protein-protein association network derived from co-regulation analysis of protein abundances from all breast cancer cell lines. The network includes only negative correlations. (**B**) Pairs of proteins that are consistently dysregulated among the different breast cancer groups are shown. The red and blue edges indicate positive and negative protein co-regulations, respectively. The size of the edges corresponds to the number of cell lines within a particular group displaying dysregulations in the indicated protein-protein association.

Fig. S6

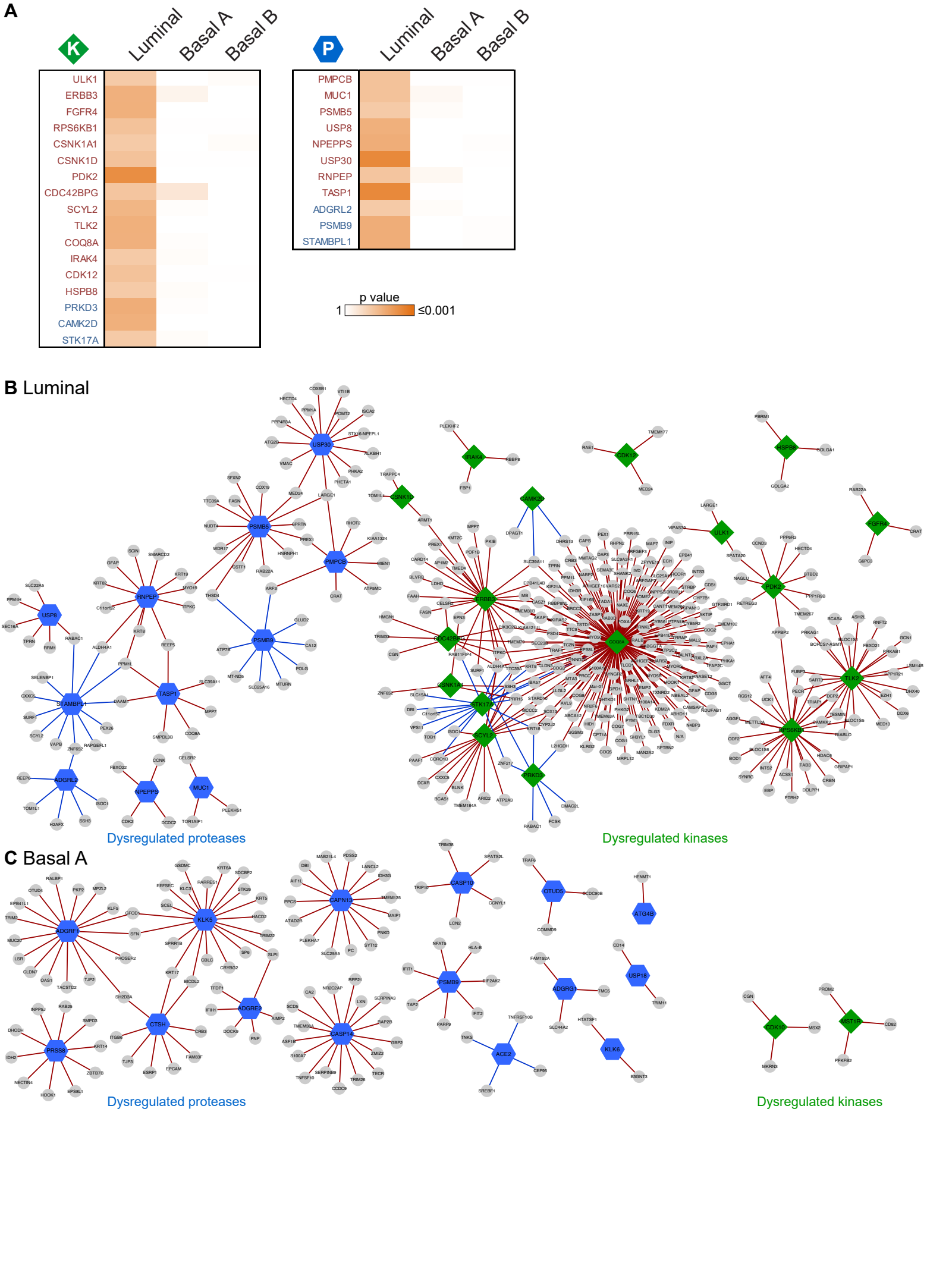
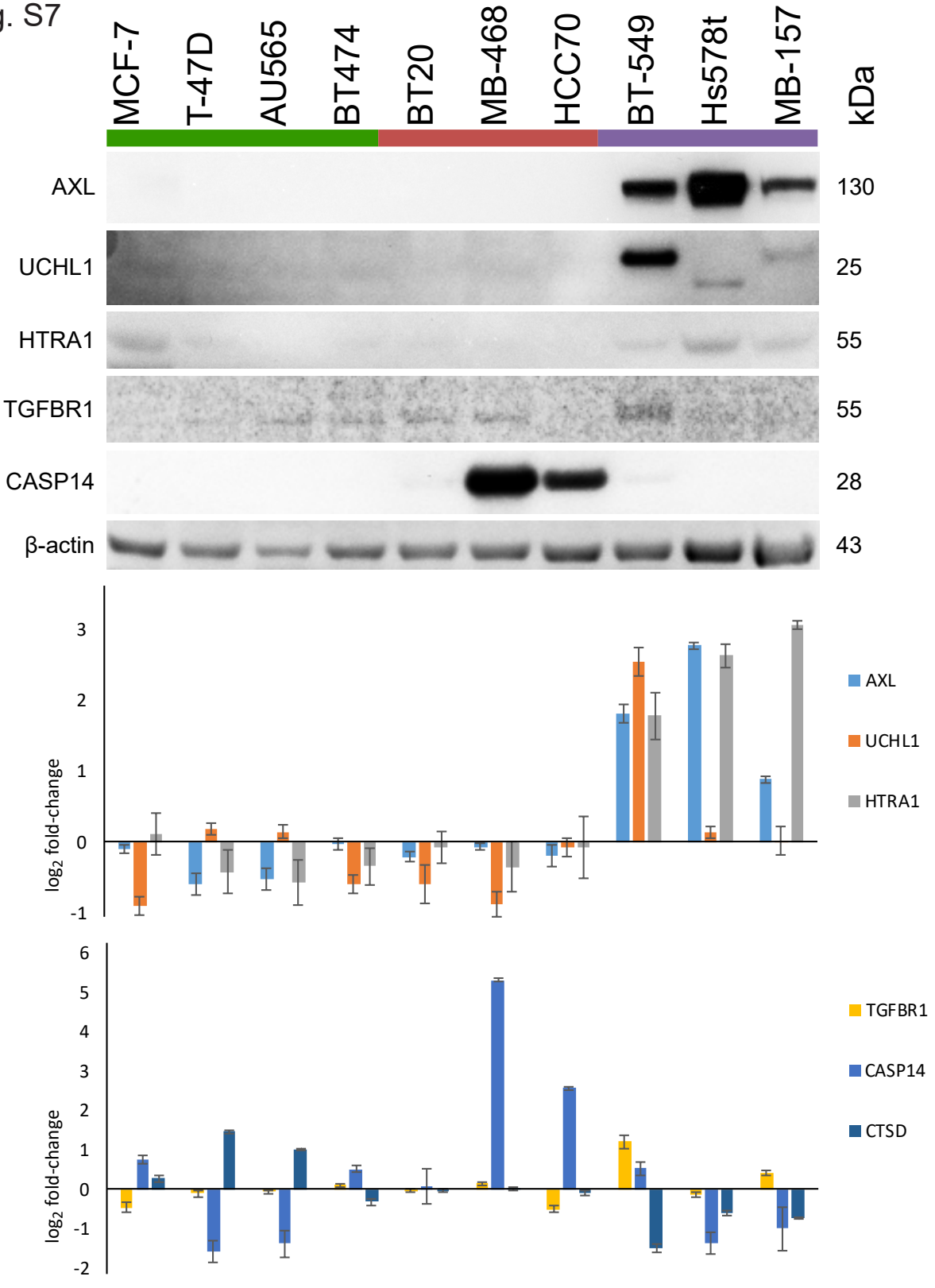

A
K
Luminal
Basal A
Basal B
ULK1
ERBB3
FGFR4
RPS6KB1
CSNK1A1
CSNK1D
PDK2
CDC42BPG
SCYL2
TLK2
COQ8A
IRAK4
CDK12
HSPB8
PRKD3
CAMK2D
STK17A
P
Luminal
Basal A
Basal B
PMPCB
MUC1
PSMB5
USP8
NPEPPS
USP30
RNPEP
TASP1
ADGRL2
STAMBPL1
p value
1
≤0.001
B Luminal
Dysregulated proteases
Dysregulated kinases
C Basal A
Dysregulated proteases
Dysregulated kinases

***Figure S6. Subtype-specific kinase and protease dysregulations.*** Related to Figure 5, and Table S5. (**A**) Dysregulated kinases and proteases specifically enriched in the luminal subtype are depicted as a heat map based on negative log10 transformed p values derived from hypergeometric testing. Kinases or proteases derived from dysregulations of positive associations and negative associations are indicated in red and blue, respectively. K denotes kinases and P denotes proteases. (**B**) Perturbed association network of kinases and proteases in the luminal subtype. Red edges denote perturbations inferred from positive associations and blue edges represent perturbations inferred from negative associations. (**C**) Perturbed association network of kinases and proteases in the basal A subclass. Red edges denote perturbations inferred from positive associations and blue edges represent perturbations inferred from negative associations.

Fig. S7

**A**


MCF-7
T-47D
AU565
BT474
BT20
MB-468
HCC70
BT-549
Hs578t
MB-157
kDa
AXL
UCHL1
HTRA1
TGFBR1
CASP14
β-actin
130
25
55
55
28
43
log2 fold-change
AXL
UCHL1
HTRA1
TGFBR1
CASP14
CTSD

**B**

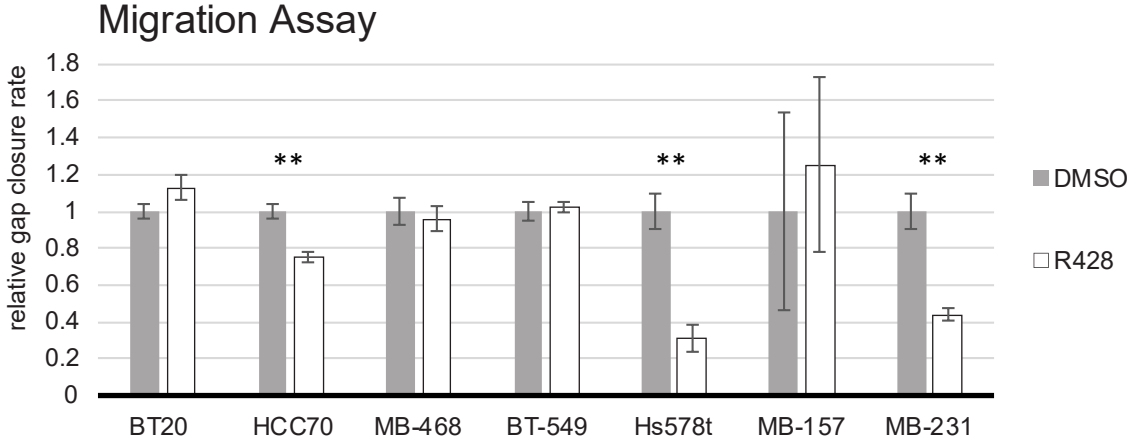

Migration Assay
relative gap closure rate
**
BT20
HCC70
MB-468
BT-549
Hs578t
MB-157
MB-231
DMSO
R428

***Figure S7. Western-blot validation and migration assay.*** Related to Figure 6. (**A**) Expression differences of selected kinases and proteases in representative cell lines for the luminal, basal A and basal B subclass. Bar-chart represents quantified fold-change by MS. (**B**) Migratory potential represented as relative gap-closure rate between R428 and DMSO treated cell lines. Data are represented as mean ± SEM, n=3. (*, p value ≤0.05; **, p value ≤0.01; n.s., not significant – using t-test).

## Transparent Methods

### *Key Resource Table*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| AXL | R and D systems | Cat# AF154, RRID:AB_354852 |
| Phospho-AXL (pAXL-Tyr$^{779}$) | R and D systems | Cat# AF2228, RRID:AB_2062560 |
| GAPDH | Santa Cruz | Cat# sc-32233, RRID: AB_627679 |
| TGFBR1 | Abcam | Cat# ab31013, RRID:AB_778352 |
| CASP14 | Abcam | Cat# ab174847 |
| UCHL1 | Cell Signaling | Cat# 13179, RRID:AB_2798141 |
| HTRA1 | Abcam | Cat# ab38611, RRID:AB_733053 |
| β-actin | Santa Cruz | Cat# sc-47778, RRID:AB_626632 |
| Chemicals, Peptides, and Recombinant Proteins | | |
| R428, AXL kinase inhibitor, 10mM in 1ml DMSO | MedChem Express | Cat# HY-15150 |
| Collagen-I, Rat Tail | gibco | Cat# A10483-01 |
| Sequencing Grade Modified Trypsin | Promega | Cat# V5111 |
| LysC (Lysyl Endopeptidase) | Wako | Cat# 125-05061 |
| TMT10plex Label Reagent Set | Thermo Scientific | Cat# 90110, LOT# SC239886 |
| Immobiline® Drystrip pH 3-10 NL, 24 cm | GE Healthcare | Cat# 17-6002-45 |
| Critical Commercial Assays | | |
| Pierce 660nm Protein Assay Reagent | Thermo Scientific | Cat# 22660 |
| Culture-Inserts 3 Well for self-insertion | ibidi | Cat# 80369 |
| EZ-PCR Mycoplasma Detection Kit | Biological Industries | Cat# 20-700-20 |
| Deposited Data | | |
| The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository. | This Study | http://www.ebi.ac.uk/pride<br>Project Accession: PXD017025 |
| Experimental Models: Cell Lines | | |
| AU565 | ATCC | Cat# CRL-2351, RRID:CVCL_1074 |
| BT20 | ATCC | Cat# HTB-19, RRID:CVCL_0178 |
| BT474 | ATCC | Cat# CRL-7913, RRID:CVCL_0179 |
| BT-549 | ATCC | Cat# HTB-122, RRID:CVCL_1092 |
| HCC1143 | ATCC | Cat# CRL-2321, RRID:CVCL_1245 |
| HCC1395 | ATCC | Cat# CRL-2324, RRID:CVCL_1249 |
| HCC1806 | ATCC | Cat# CRL-2335, RRID:CVCL_1258 |
| HCC1937 | ATCC | Cat# CRL-2336, RRID:CVCL_0290 |
| HCC70 | ATCC | Cat# CRL-2315, RRID:CVCL_1270 |
| Hs578t | ATCC | Cat# HTB-126, RRID:CVCL_0332 |
| MB-157 (MDA-MB-157) | ATCC | Cat# HTB-24, RRID:CVCL_0618 |
| MB-231 (MDA-MB-231) | ATCC | Cat# HTB-26, RRID:CVCL_0062 |
| MB-436 (MDA-MB-436) | ATCC | Cat# HTB-130, RRID:CVCL_0623 |
| MB-468 (MDA-MB-468) | ATCC | Cat# HTB-132, RRID:CVCL_0419 |
| MCF-7 | ATCC | Cat# HTB-22, RRID:CVCL_0031 |

| | | |
|---|---|---|
| Sk-Br-3 | ATCC | Cat# HTB-30, RRID:CVCL_0033 |
| T-47D | ATCC | Cat# T-47D, RRID:CVCL_0553 |
| ZR-75-1 | ATCC | Cat# CRL-1500, RRID:CVCL_0588 |
| Software and Algorithms | | |
| Proteome Discover 2.2 | Thermo Scientific | RRID:SCR_014477 |
| Perseus Version 1.6.6.0 | (Cox and Mann, 2012; Tyanova et al., 2016) | https://maxquant.net/perseus, RRID:SCR_015753 |
| Image Lab Version 5.2 build 14 | Bio Rad | http://www.bio-rad.com/en-sg/sku/1709690-image-lab-software?ID=1709690, Cat# 1709690, RRID:SCR_014210 |
| iBright Analysis Software | Thermo Scientific | RRID:SCR_017632 |
| ImageJ 1.51w | Wayne Rasband National Institute of Health, USA | http://imagej.nih.gov/ij, RRID:SCR_003070 |
| Scratch wound assay (ImageJ Macro) | John Lim (IMB Microscopy Unit, A*STAR) (Bigliardi et al., 2015) | https://www.a-star.edu.sg/imb/Tech-Platforms/AMP-Light-Microscopy/Image-analysis/Scratch-wound-assay |
| MEGA X Version 10.0.2 | (Kumar et al., 2018) | www.megasoftware.net |
| R statistical environment Version x64 3.4.0 | (R Core Team, 2014) | http://www.r-project.org, RRID:SCR_001905 |
| Gene set enrichment analysis (GSEA) Version 3.0 | (Subramanian et al., 2005) | http://www.broadinstitute.org/gsea, RRID:SCR_003199 |
| Cytoscape Version 3.6.0 | (Shannon et al., 2003) | https://cytoscape.org |
| Enrichr | (Kuleshov et al., 2016) | http://amp.pharm.mssm.edu/Enrichr, RRID:SCR_001575 |
| DAVID Version 6.8 | (Huang da et al., 2009) | http://david.abcc.ncifcrf.gov, RRID:SCR_001881 |
| Others | | |
| UniProtKB Human (2019-06) | UniProt Consortium | http://www.uniprot.org/help/uniprotkb, RRID:SCR_004426 |
| Kincat Hsap 08.02; Update Dec 07 | (Manning et al., 2002) | www.kinase.com |
| HuPho: human phosphatase portal (2018-03-19_07-16) | (Liberti et al., 2012) | http://hupho.uniroma2.it/index.php |
| MEROPS DB release 12.0 (2018-03-19) | (Rawlings et al., 2018) | http://merops.sanger.ac.uk, RRID:SCR_007777 |
| The FANTOM consortium and RIKEN Omics Science Center | (Ravasi et al., 2010) | http://fantom.gsc.riken.jp, RRID:SCR_002678 |
| Pathway Commons release 12 (2019-09-18) | (Cerami et al., 2011) | https://www.pathwaycommons.org/ |
| Cancer Gene Census release v90 (2019-09-05) | (Sondka et al., 2018) | http://cancer.sanger.ac.uk/census |

***Contact for reagent and resource sharing***

Further information and requests for resources and reagents should be directed to the Lead Contact, Jayantha Gunaratne (jayanthag@ imcb.a-star.edu.sg).

***Cell line models***

AU565, BT20, BT474, BT-549, SK-BR-3, HCC1143, HCC1395, HCC1806, HCC1937, HCC70, Hs578t, MCF-7, MDA-MB-157 (MB-157), MDA-MB-231 (MB-231), MDA-MB436 (MB-436), MDA-MB-468 (MB-468), T-47D and ZR-75-1, were derived from female breast tumors. Cell lines were either freshly obtained from ATCC (pre-authenticated) or authentication confirmed by ATCC using STR Profiling prior to proteome profiling. All cell lines were regularly tested for mycoplasma contaminations using EZ-PCR Mycoplasma Detection Kit. BT474, Hs578t, MCF-7, MB-157, MB-231, and MB-468 were cultured in DMEM supplemented with 10% fetal bovine serum (FBS), 1%PenStrep and 1% L-glutamine. BT20 was cultured in MEM supplemented with 10% FBS, 1% PenStrep, 2% Sodium Bicarbonate, 1% Non-Essential Amino Acids and 1% Sodium Pyruvate. Sk-Br-3 was cultured in McCoy's 5a Medium supplemented with 10% FBS and 1%PenStrep. The remaining cell lines were cultured in RPMI1640 supplemented with 10% FBS and 1% PenStrep. All cell lines were cultured at 37 °C with 5% $CO_2$.

***Cell lysis and sample preparation for mass spectrometry***

For harvesting, cell lines were grown to 90-100% confluency and rinsed with cold PBS before removal from cell culture dish by trypsinizing. Cells were pelleted by centrifugation at 1500 rpm for 7 minutes at 4 °C followed by washing with cold PBS and re-pelleting. Subsequently cell pellets were snap-frozen on dry-ice and stored at -80 °C.

Cells were lysed by addition of approximately 3x pellet volume lysis buffer (6 M urea, 2 M thiourea and 100 mM ammonium bicarbonate) and repeated passing of the suspension through a 1 ml pipette tip. Undissolved debris was pelleted by centrifugation at 15000 rpm for 20 minutes, and the supernatant was transferred to a new low-bind tube followed by addition of 1 µl benzonase and incubated for 15 minutes at room temperature, before protein concentration was determined using Pierce 660 nm Protein Assay Reagent as per manufacturer's instructions. For pooled sample control, equal protein amounts from each cell lysate were combined and used for all replicates. 150 µg of proteins per sample were subsequently reduced by addition of 100 mM dithiothreitol to a final concentration of 5 mM and incubated for 30 minutes. The reduced samples were subjected to alkylation by addition of 100 mM iodacetamide to a final concentration of 10 mM and incubated further for 30 minutes in the dark. Proteins were then digested by addition of LysC (Lysyl Endopeptidase, Wako) in a protein:enzyme ratio of 1:100 and overnight incubation at 37 °C. Subsequently, urea concentration in the samples was reduced to 1 M by addition of 50 mM ammonium bicarbonate before protein digestion using Sequencing Grade Modified Trypsin (Promega) at a 1:50 enzyme:protein ratio for 8 hours at 37 °C. After digestion, the samples were desalted by C18 solid-phase extraction and peptide concentration was determined using Nanodrop (Thermo Scientific) followed by drying under vacuum.

***TMT10plex labeling***

The TMT10plex Label reagents (Thermo Scientific) were first equilibrated to room temperature and suspended using 41 µl anhydrous acetonitrile. 100 µg of digested dried peptides were re-suspended in 40 µl of 100 mM triethyl ammonium bicarbonate (TEAB) before being added to the appropriate TMT label reagent in random

order. After gentle mixing of the TMT label and sample, the labeling reaction was incubated for 1 hour at room temperature, followed by addition of 8 µl quenching reagent (5% Hydroxylamine in 100 mM TEAB) for 15 minutes to quench the reaction. Subsequently, labeled samples were combined, briefly vortexed, and split into two equal parts for fractionation using IEF and bRP-LC. Samples were stored at -80 °C until further use.

***Sample fractionation by basic pH reversed-phase liquid chromatography (bRP-LC)***

Peptide sample volume was reduced to 20 µl under vacuum followed by resuspension in 20 mM ammonium formate pH 10 to final volume of 110 µl. The sample was then fractionated using a 3.0x150mm XBridge C18 3.5 µm column on a Shimadzu Nexera XR HPLC system with fraction collector. Fractionation was accomplished by applying a 45 minute gradient from 4 to 72% acetonitrile in 20 mM ammonium formate pH 10 with a flowrate of 0.5 ml/min. A total of 48 fractions were consecutively combined into 15, followed by freeze-drying under vacuum prior to MS analysis.

***Sample fractionation by off-gel isoelectric focussing (IEF)***

Peptide samples were fractionated using Immobiline Drystrip pH 3-10 NL, 24 cm on an Agilent 3100 OFFGEL Fractionator (Agilent, G3100AA) according to manufacturer's protocol. 24 fractions were collected and desalted using self-packed C18 stage tips, followed by freeze-drying under vacuum prior to MS analysis.

***Liquid Chromatography – Mass Spectrometry (LC-MS)***

Dried samples were reconstituted in 0.1% formic acid, followed by loading of 2 µg total peptides per fraction onto a trap column of C18 Acclaim PepMap 100 of 5µm, 100 Å, 100 µm I.D. x 2 cm and an analytical column of PepMap RSLC C18, 2um, 100 Å, 75µm I.D. x 50 cm using Easy nLC 1000 (Thermo Scientific). The LC solvent A comprised of 0.1 M formic acid in 2% acetonitrile and LC solvent B comprised of 0.1 M formic acid in 95% acetonitrile. Tryptic peptides were separated using a 3 hour gradient from 8–38% solvent B with a flowrate of 200 nl/min and subsequently analyzed using Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Scientific) operating in positive ion mode. MS data was recorded in data-dependent acquisition mode, for acquiring full scan MS spectra (m/z 310 – 1510) with a resolution of R=120,000 at an AGC target of 4e5, and a maximum injection time of 50 ms. Data dependent mode was set to cycle time for 3 seconds. Each MS2 scan was sequentially isolated to an AGC target value of 1e5 with resolution of 60,000 and maximum injection time of 105 ms and fragmented using HCD collision energy of 34% at MS2 level with fixed first mass at 120 and isolation window of 1.2 m/z.

***Protein/peptide identification and quantification***

Mass spectrometry raw data was analyzed using Proteome Discoverer 2.2 (Thermo Scientific) and proteins were identified using the integrated Sequest HT search engine. Search was performed against human database (2019-06) for fully tryptic peptides with a maximum of two missed cleavage sites. Methionine oxidation and N-terminal acetylation were stated as dynamic modifications, and carbamidomethylation at cysteine and TMT-modifications at any peptide N-terminus and lysine were included as static modifications in the search parameters. Threshold for peptide and protein validation was set to 0.01 false discovery rate (FDR). Quantification was based on MS2 TMT reporter ion abundance with a co-isolation threshold of 50 and an average reporter S/N threshold of 10. Protein abundance was normalized to equal total protein for each TMT-channel.

***Proteome expression data analysis***

A pooled sample from all cell lysates was labeled with the 126 TMT-reporter within each TMT10plex set. To normalize for batch variation between TMT sets, median protein abundance of all 126-channels was calculated and subsequently protein abundance of each TMT set was multiplied by a corresponding correction factor, followed by $log_2$-ratio calculation to median pooled protein abundance.

Pearson's correlation was calculated for the protein expression ratios of all replicates. Following good correlation, median $log_2$-ratios were calculated for all replicates and finally, all protein expression ratios were median-normalized across all cell lines.

The cell lines were categorized into three defined groups as 'luminal', 'basal A' and 'basal B' following PCA and clustering analysis. Only those proteins with more than 50% values in each of the defined groups were considered for subsequent statistical testing using one-way ANOVA. Significant differentially expressed (DE) proteins with p value ≤ 0.05 were filtered and subjected to pairwise comparisons using two-sample Student's t-tests. Significant proteins in each group were identified as those with p value ≤ 0.05 in all the pairwise t-tests.

Correlation analysis, ANOVA and Student's t-test analysis, unsupervised clustering and PCA were performed using Perseus 1.6.6.0 (Tyanova et al., 2016)

***Functional protein class mapping***

Kinase information of human non-pseudogenes was extracted from www.kinase.com (Kincat Hsap 08.02; Update Dec 07) (Manning et al., 2002) and the corresponding Uniprot accessions were obtained by Uniprot ID mapping tool. Missing values and duplicates were cross-checked and validated at www.kinasenet.ca using Entrez gene name. This resulted in 506 kinases with distinct Uniprot accessions. DE kinases were subsequently mapped onto the human kinome tree (Manning et al., 2002) using the web-based Coral tool (Metz et al., 2018).

Uniprot accessions for 197 human non-pseudogene phosphatases were extracted from the human phosphatase portal (2018-03-19_07-16) (Liberti et al., 2012).

Protease information for human non-pseudogenes was extracted from MEROPS DB release 12.0 (2018-03-19) (Rawlings et al., 2018) and the MEROPS IDs were mapped to their corresponding Uniprot accessions using Uniprot ID mapping tool, resulting in a total of 529 unique proteases. Similar to the human kinome tree, a phylogenetic tree was created for protease families. Amino acid sequences of proteases families were aligned using ClustalW algorithm in MEGA-X (Kumar et al., 2018) followed by creation of maximum likelihood tree using Poisson model. Subsequently DE proteases were mapped onto the constructed tree.

Information for human non-pseudogene transcription factors was extracted from the supplementary data curated by The FANTOM consortium and RIKEN Omics Science Center (Ravasi et al., 2010). Gene IDs were mapped to Uniprot accession numbers using the Uniprot ID mapping tool, resulting in 1961 transcription factor IDs.

***Expression based protein-protein network analysis***

Kinase- and protease-centric networks were explored using interactions curated in the Pathway Commons database release 12 (2019-09-18) that integrates biological interaction and pathway information from multiple databases (Cerami et al., 2011). Neighborhood interactions for all the kinases and proteases that were significantly altered in each of the breast cancer subtype as against the other groups were first derived by searching against the interaction database. Apart from physical interactions, functional interactions such as those that control phosphorylation, regulate protein expression and influence complex formation were also included. In this manner, for all the DE kinases and proteases, a proximity network encompassing signaling, regulatory and metabolic networks were derived. The network neighborhood for each kinase and protease was then filtered to retain those proteins that overlapped with the DE proteins in that particular subtype. The proximity networks for all the kinases and proteases were merged and any possible interactions among the retained DE proteins were also considered. The kinase- and protease-proximity networks for each subtype were then integrated and visualized in Cytoscape (Shannon et al., 2003). The topological parameters of the network were analyzed for centrality and degree distribution. Finally, cluster assessment was performed using graph clustering algorithm MCODE to identify densely connected regions based on network topology (Bader and Hogue, 2003).

***Enrichment analysis***

The DE proteins including the kinases and proteases for each breast cancer subtype were analyzed for enrichment of cancer hallmark gene sets from the Molecular Signature Database (MSigDB) collections (Liberzon et al., 2015). The enrichment was performed as implemented in gene set enrichment analysis (GSEA) at FDR ≤ 25% (Subramanian et al., 2005). For comparison of gene set distributions across different breast cancer subtypes, only those proteins within the enriched gene sets that were also part of the kinase-protease neighborhood network were considered. The significance of protein abundance distribution of each gene set across the different breast cancer groups was assessed using Kolmogorov–Smirnov (KS) test, and those with p value ≤ 0.05 were considered significant. In addition, gene ontology and pathway enrichment by KEGG pathways was assessed using DAVID (Huang da et al., 2009), and significantly enriched ontology terms and signaling pathways (p value ≤ 0.05) were derived. Statistical testing of distributions of protein abundances within each signaling pathway across the subtypes were assessed using KS test as described above.

***Correlation-based dysregulated association analysis***

The quantification profiles of proteome from all cell lines were correlated using Spearman's rank correlation. For this analysis, only those proteins that were quantified across all cell lines (10,400 proteins) were considered. The correlation coefficients and p values for all correlating pairs were calculated and adjusted for multiple testing using Benjamini-Hochberg correction. The correlating pairs were filtered for adjusted p value ≤ 0.1, which retained those with an absolute correlation coefficient of ~0.76 from both positively and negatively correlated associations. Unweighted correlation networks were constructed using these highly correlated protein-protein associations and specific deviations of individual cell lines from such correlations were assessed by calculating the Mahalanobis distance followed by Grubbs's test for outliers (p value ≤ 0.05). The above analysis that revealed dysregulated protein associations was adapted from the 'interactome mapping by high-throughput quantitative proteome analysis' (IMAHP) method (Lapek et al., 2017). Following the observation that the dysregulated proteins from each cell line displayed discrete patterns according to their subtype on mapping to the background positive correlation network, hypergeometric test was performed to assess the enrichment of

dysregulated proteins within each breast cancer subtype (p value ≤ 0.1). Known cancer drivers among the dysregulated proteins were identified by mapping against genes curated in COSMIC Cancer Gene Census (CGC) (Sondka et al., 2018). All the above analyses were done in R statistical environment (R Core Team, 2014). In addition, enrichment analysis was also carried out to identify protein associations/ pairs of proteins that were consistently dysregulated in each subtype. Ontology and pathway assessment of the dysregulated proteins for KEGG and Reactome pathways were carried out using DAVID (Huang da et al., 2009) and Enrichr (Kuleshov et al., 2016). Only those terms with at least 5 proteins and p value ≤ 0.05 were considered significant and are shown in Table S6. Dysregulated kinases and proteases that were enriched within each subclass based on assessment of the positive or negative correlation networks were visualized as heat maps, and their corresponding co-regulated associations were assembled into a network for visualization in Cytoscape.

***Western Blotting***

Primary antibodies against human AXL and human phospho AXL (pAXL-Tyr$^{779}$) were purchased from R and D systems. Antibody against human GAPDH was purchased from Santa Cruz biotechnology. Equal amount of protein lysate from selected cell lines were separated on a 4-12% Bis-Tris gel, followed by transfer onto a PVDF membrane. Membranes were subsequently blocked using 5% skim milk in tris-buffered saline and tween20 (TBST), followed by incubation with primary and HRP-conjugated secondary antibody. Bands were visualized using HRP Substrate and medical X-ray film or iBright Western Blot Imaging System (Thermo Scientific, USA).

***Invasion/migration wound-healing assay***

AXL kinase inhibitor R428 was purchased from MedChem Express. 24 well plates were pre-treated with 50µg/ml rat-tail collagen-I (gibco) in 20 mM acetic acid for invasion assays or 20 mM acetic acid for migration assays. After pre-treatment, Culture-Inserts 3 Well for self-insertion (ibidi) was inserted in each of the 24 wells. 70 µl cell suspension at 3 x $10^5$ cells/ml in growth media was added to each well of the cell culture insert, followed by incubation at 37 °C and 5% $CO_2$. After 24 hours of primary incubation, growth media was replaced with FBS-free growth media for approximately 16 hours. After serum starvation, media was replaced by full growth media with addition of 0.75 µM AXL inhibitor R428 or DMSO and further incubated for another 2 hours. After pre-incubation with R428 or DMSO, cell culture inserts were removed and attached cells were washed twice with cold PBS. For invasion assay, cells were covered with 200 µl of 1 mg/ml cold rat tail collagen-I in growth media (+0.75 µM R428 or +DMSO). For migration assays, cells were covered with 200 µl cold growth media (+0.75 µM R428 or +DMSO). After 20 minutes incubation at 37 °C and 5% $CO_2$ for collagen polymerization, additional 800 µl growth media (+0.75 µM R428 or +DMSO) was added to each well. Time-lapse images of the created gap by cell culture inserts were acquired every 30 minutes for an average of 24 hours using a Zeiss LSM800 Inverted Confocal with environmental chamber (37°C, 5% $CO_2$).

***Experimental validation analysis***

Western blot images were either analyzed using iBright Analysis Software or scanned images of blotted protein bands on X-ray film were analyzed using Image Lab TM software (Bio Rad) to calculate relative abundance of proteins between different samples.

For wound healing assays, the area of the gap for each time-lapse image was determined using 'Scratch wound assay' ImageJ macro (Bigliardi et al., 2015). Gap closure rate of wound healing assays was calculated by regression for linear part of the gap closure.

## References


Bader, G.D., and Hogue, C.W. (2003). An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics *4*, 2.

Bigliardi, P.L., Neumann, C., Teo, Y.L., Pant, A., and Bigliardi-Qi, M. (2015). Activation of the δ-opioid receptor promotes cutaneous wound healing by affecting keratinocyte intercellular adhesion and migration. Br J Pharmacol *172*, 501-514.

Cerami, E.G., Gross, B.E., Demir, E., Rodchenkov, I., Babur, O., Anwar, N., Schultz, N., Bader, G.D., and Sander, C. (2011). Pathway Commons, a web resource for biological pathway data. Nucl Acids Res *39*, D685-690.

Cox, J., and Mann, M. (2012). 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. BMC Bioinformatics *13*, S12.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols *4*, 44-57.

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A.*, et al.* (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucl Acids Res *44*, W90-97.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular Biology and Evolution *35*, 1547-1549.

Lapek, J.D., Greninger, P., Morris, R., Amzallag, A., Pruteanu-Malinici, I., Benes, C.H., and Haas, W. (2017). Detection of dysregulated protein-association networks by high-throughput proteomics predicts cancer vulnerabilities. Nature Biotechnology *35*, 983.

Liberti, S., Sacco, F., Calderone, A., Perfetto, L., Iannuccelli, M., Panni, S., Santonico, E., Palma, A., Nardozza Aurelio, P., Castagnoli, L.*, et al.* (2012). HuPho: the human phosphatase portal. The FEBS Journal *280*, 379-387.

Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. cels *1*, 417-425.

Manning, G., Whyte, D.B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The Protein Kinase Complement of the Human Genome. Science *298*, 1912-1934.

Metz, K.S., Deoudes, E.M., Berginski, M.E., Jimenez-Ruiz, I., Aksoy, B.A., Hammerbacher, J., Gomez, S.M., and Phanstiel, D.H. (2018). Coral: Clear and Customizable Visualization of Human Kinome Data. cels *7*, 347-350.e341.

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M.*, et al.* (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucl Acids Res *47*, D442-D450.

R Core Team (2014). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

Ravasi, T., Suzuki, H., Cannistraci, C.V., Katayama, S., Bajic, V.B., Tan, K., Akalin, A., Schmeier, S., Kanamori-Katayama, M., Bertin, N.*, et al.* (2010). An atlas of combinatorial transcriptional regulation in mouse and man. Cell *140*, 744-752.

Rawlings, N.D., Barrett, A.J., Thomas, P.D., Huang, X., Bateman, A., and Finn, R.D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucl Acids Res *46*, D624-D632.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res *13*, 2498-2504.

Sondka, Z., Bamford, S., Cole, C.G., Ward, S.A., Dunham, I., and Forbes, S.A. (2018). The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nat Rev Cancer *18*, 696-705.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America *102*, 15545-15550.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature Methods *13*, 731-740.