Saccharomyces cerevisiae-like 1 (SEC14L1) is a prognostic factor in breast cancer associated with lymphovascular invasion

Sonbul SN[1,2], Aleskandarany MA[1,3], Kurozumi S[1], Joseph C[1], Toss MS[1], Mukherjee A[1], Martin S[1], Caldas C[4], Ellis IO[1], Green AR[1], Rakha EA[1,3]

[1]Division of Cancer and Stem Cells, School of Medicine, University of Nottingham, and Nottingham University Hospital NHS Trust, Nottingham City Hospital, Nottingham, UK. [2]Biochemistry Department, Faculty of Sciences, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia. [3]Faculty of Medicine, Minufiya University, Egypt. [4]Addenbrooke's Hospital, Cambridge Breast Unit, Cambridge University Hospital NHS Foundation Trust, Cambridge, UK

Corresponding Author:

Professor Emad Rakha

Department of Histopathology, Division of Cancer and Stem Cells, School of Medicine, The University of Nottingham and Nottingham University Hospitals NHS Trust, Nottingham City Hospital, Nottingham, NG5 1PB, UK.

Email: Emad.Rakha@nottingham.ac.uk

Keywords: Lymphovascular invasion; Prognostic biomarker; Saccharomyces cerevisiae-like 1 (SEC14L1); Invasive breast cancer

Short Title: SEC14L1 expression in breast cancer

Lymphovascular invasion (LVI) is strongly related to breast cancer (BC) metastasis. However, the underlying mechanisms of LVI and its driver molecules in BC remain to be defined. In this study, we explore differential expression of genes in large molecularly characterized and clinically annotated datasets of invasive BC patients (n=8056) coupled with histological review and strict definition for LVI status. The METABRIC series was used to identify genes associated with LVI, as defined using H&E supplemented by immunohistochemistry (IHC), at the genomic/transcriptomic levels. Saccharomyces cerevisiae-like 1 (SEC14L1) was identified as one of the most significant genes associated with LVI. The prognostic significance of *SEC14L1* gene copy number and mRNA expression was further investigated in the METABRIC series and externally validated using the Breast Cancer Gene-Expression Miner v4.0. Protein expression of SEC14L1 was also assessed using IHC in series of early stage BC using tissue microarrays. *SEC14L1* gene copy number gain was significantly associated with high histological grade and poor outcome. *SEC14L1* mRNA expression showed positive association with higher grade, lymph node metastasis and poor outcome. SEC14L1 protein overexpression was significantly associated with LVI (p<0.0001), higher grade (p=0.011), HER2 positivity (p=0.036), and shorter survival (p=0.00075). Our findings specify SEC14L1 as an independent prognostic factor in BC. Its association, at both transcriptome and protein expression levels, with LVI and outcome could imply an important role in tumor progression. A further mechanistic insight into its molecular roles including potential therapeutic utility is warranted.

Although several biomarkers associated with breast cancer progression and response to therapy have been identified, the exact molecular signatures for invasion machineries that lead to metastatic disease remains to be defined. Deciphering the driver genes/proteins that dictate the biological metastatic behavior of breast cancer is essential to understanding cancer progressive mechanisms and opening novel avenues for therapeutic interventions. Lymphovascular invasion (LVI) is an independent prognostic parameter of poor outcome in invasive breast cancer and is major prerequisite for the development of metastasis (1-3). Comprehension of the molecular mechanisms underlying LVI in breast cancer and unveiling its key-players could lead to unique therapeutic targets and improve risk prediction (4-6). However, the sophistication of the molecular mechanisms underlying LVI as part of the invasion-metastasis cascade with involvement of several genes, diverse signaling pathways and interactions of tumor microenvironment in addition to the subjectivity of LVI morphological assessment in clinical samples renders this task highly challenging (6).

Advancement in high-throughput molecular and bioinformatics techniques coupled with thorough pathological assessment could help decipher LVI molecular regulators in breast cancer. Therefore, we have interrogated transcriptomic profiles of the large-scale cohort of invasive breast cancer patients of the METABRIC following morphological and molecular assessment for LVI status to identify LVI differentially expressed genes [7]. Cytosolic factor *Saccharomyces cerevisiae*-like 1 (*SEC14L1*) was identified as one of the most significant genes associated with LVI using genomic data.

The aim of the present study was to evaluate the clinicopathological and prognostic significance of *SEC14L1* copy number alteration (CNA) and mRNA expression in the

large annotated cohort of the METABRIC. Associations of SEC14L1 protein expression with clinicopathological prognostic variables including LVI, intrinsic molecular subtypes, and patient outcome were assessed. Our interpretation of the prognostic impact of SEC14L1 expression in invasive BC patients may unveil new insights on the sophisticated pathways of LVI.

MATERIAL AND METHODS

Analysis of differential gene expression

The Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) cohort [8, 9] is a large genomic and transcriptomic dataset from 1980 primary operable invasive BC female patients. In this cohort, the ER-positive and/or lymph node metastasis negative patients were treated without adjuvant chemotherapy. The ER-negative or lymph node positive patients had adjuvant chemotherapy. No HER2-positive breast cancer patients had trastuzumab therapy. The extracted and purified DNA probes were hybridized to Affymetrix SNP 6.0 arrays (Affymetrix, Santa Clara, CA) at AROS Applied Biotechnology (Aarhus, Denmark). For RNA analysis, the Illumina Totalprep RNA amplification kit and Illumina Human HT-12 v3 Expression BeadChips (Ambion, Warrington, UK) were used as described previously [8].

We interrogated the Nottingham cases from the METABRIC cohort (discovery set), and gene expression levels were compared between LVI positive and LVI negative cases. The LVI status for all patients was evaluated morphologically using hematoxylin and eosin (H&E) staining on full-face tumor sections. Besides, corresponding full-face tumor sections for each patient were stained immunohistochemically (IHC) with CD34 and D2-40 (podoplanin, *PDPN*), [10] to refine the status of LVI. Cases were considered LVI negative by the absence of LVI

in both H&E and IHC stained section. Cases with discordant results or showing lymph node positive but LVI negative were excluded to avoid bias caused by false positive or false negative diagnosis.

A linear model for microarray data and RNA-seq package (LIMMA) method [11] was applied for detecting differentially expressed genes/transcripts between LVI positive and negative groups. This supervised approach of the differential gene expression analysis is compatible with the Affymetrix gene expression data. The Addenbrookes' Hospital cases, within the METABRIC cohort (n=521), with defined LVI status were used to validate the results of the differential gene expression analyses externally. The top differentially expressed genes were ranked based on their p-value of association with LVI. Subsequently, the CNAs of the top differentially expressed genes were determined using the Affymetrix SNP6 Copy Number Inference Pipeline (Cancer Genomics Computation Analysis group of the Broad Institute, USA) [7]. Cytosolic factor Saccharomyces cerevisiae-like 1 (SEC14L1) was identified as one of the highly significant genes associated with LVI. The prognostic impact of SEC14L1 mRNA expression was evaluated in the 1980 cases of METABRIC cohort and externally validated using the Breast Cancer Gene-Expression Miner v4.0 [n=5788] [12].

Analysis of SEC14L1 protein expression

The Nottingham BC series which were included in the METABRIC cohort were used in this study. All patients were treated without neoadjuvant treatment. ER, progesterone receptor (PR), HER2 and Ki67 were stained and scored according to guideline recommendations and as previously published (7, 8). The cut-off value of ER and PR were determined as 1% (9). ER-positive and HER2 negative tumors were considered as the luminal class with PR negative and high Ki67 (labelling index

< 10%) cases were determined as luminal B subtype. Clinicopathological data and patients' outcome had been recorded and regularly updated. The recorded outcome data comprised survival status, mean survival in months, recurrence of disease (including distant metastases) and cause of death. BC-specific survival (BCSS) was determined as the time from the day of surgery until the time of death from or with BC in 10 years follow-up.

Antigen binding specificity for rabbit polyclonal anti-SEC14L1 antibody (HPA028703, Sigma Aldrich, UK) was examined by western blotting (WB) at 1:500. A set of different cell lysates was utilized; MCF7, HeLa, MDA-MB-231, Jurkat, human embryonic kidney (HEK) 293, MDA-MB-468 (all obtained from the American Type Culture Collection; Rockville, MD, USA).

Prior to assessment of SEC14L1 IHC expression on tissue microarray (TMA), a subset of full-face tissue sections of breast cancer (n=20) was IHC stained. IHC was performed on (4 μm) TMA sections using Novocastra Novolink Polymer Detection Systems kit (Code: RE7280-K, Leica Biosystems, UK) following the manufacturer's protocol. Briefly, all full-face and TMA sections were incubated on a hot plate at 60˚C for 10 minutes; then deparaffinized in xylene and rehydrated through descending grades of ethanol. SEC14L1 staining was performed utilizing anti-SEC14L1 primary antibody at 1:25 dilution in Leica Antibody Diluent, for 60 minutes. Finally, 3-3' Diaminobenzidine tetrahydrochloride (Novolink DAB substrate buffer plus) was freshly prepared and used as a chromogen. Counterstaining was performed using Meyer's hematoxylin for 6 min. Negative and positive controls (anti-human-β-2-microglobulin; A0072, Dako) were included. The cytoplasmic expression of SEC14L1 was assessed using the percentage of positivity (0-100%) and staining intensity

(Negative: Score 0, Weak: Score 1, Moderate: Score 2, and Strong: Score 3). H-score of SEC14L1 was calculated as previous publication [16].

This work was approved by Nottingham Research Ethics Committee 2 under the title: Development of molecular genetic classification of breast cancer. All tissue samples included in this study were from patients who were consented before inclusion in the study cohort.

Statistical analysis

Statistical analysis was performed using SPSS (IBM SPSS Statistics, Version 22). The relationship between SEC14L1 CNA and mRNA expression was calculated using analysis of variance (ANOVA) test with Bonferroni correction. Differences between two groups were assessed using Mann-Whitney test (non-normal distribution) to determine the associations between SEC14L1 mRNA expression and LVI status. Chi-square test was used to evaluate the relationship between SEC14L1 expression and categorical variables. For dichotomization of the data, X-Tile (X-Tile Bioinformatics Software, Yale University, version 3.6.1) was used. Survival curves were generated by Kaplan-Meier survival analysis with differences in outcome assessed by Log Rank test. Cox's proportional hazard method was performed for multivariate analysis to identify the independent prognostic/predictive factors. The p-value ≤ 0.05 was considered significant.

RESULTS

There was a strong positive correlation between SEC14L1 mRNA expression and SEC14L1 gene copy number gain; higher levels of SEC14L1 mRNA were detected in cases with copy number gain compared to those with neural copy number ($p<0.0001$). SEC14L1 mRNA expression was significantly lower in the copy number

loss group (p<0.0001). Further analysis of SEC14L1 CNA revealed a significant association with histological grade 3 (p<0.0001) and luminal B molecular subtype (p<0.0001) (Table 1). The cut-off value of SEC14L1 mRNA expression was determined at median value. The overexpression of SEC14L1 mRNA was associated with the higher histological grade (p<0.0001), axillary node metastasis (p<0.0001) and the intrinsic molecular subtypes (p<0.0001; Table 1).

In the METABRIC cohort, the survival of patients with SEC14L1 copy number gain was significantly shorter than those of copy number neutral group (p=0.0007). However, no difference of survival was observed between CNA loss and neutral groups (p=0.13) (Figure 1A). High expression of SEC14L1 mRNA conferred a significantly worse prognosis compared to low SEC14L1 mRNA expression (p=0.00049; Figure 1B). External validation of the prognostic power of SEC14L1 mRNA expression in the Breast Cancer Gene-Expression Miner v4.0 revealed that high SEC14L1 mRNA expression was associated with poor prognosis (p<0.05).In the METABRIC cohort, the survival of patients with SEC14L1 copy number gain was significantly shorter than those of copy number neutral group (p=0.0007). However, no difference of survival was observed between CNA loss and neutral groups (p=0.13) (Figure 1A). High expression of SEC14L1 mRNA is relevant with a worse prognosis compared to low SEC14L1 mRNA expression (p=0.00049; Figure 1B). External validation of the prognostic power of SEC14L1 mRNA expression in the Breast Cancer Gene-Expression Miner v4.0 revealed that high SEC14L1 mRNA expression was associated with adverse prognosis (p<0.05).

SEC14L1 protein expression

Evaluation of WB supported a specific binding affinity of the anti-SEC14L1 primary antibody to single protein bands for each cell lysate around the predicted molecular

mass (77 kilo Dalton) of SEC14L1 protein, confirming the specificity of the antibody (Supplementary Figure 1). On full-face tissue sections, strong protein expression of SEC14L1 was detected in the cytoplasm of breast cancer cells (Figure 2). On the other hand, weaker protein expression was observed in the adjacent normal epithelial tissue. Assessment of full-face sections revealed a homogeneous expression pattern for SEC14L1 protein confirming the reliability of TMA to detect its expression (Figure 3). The optimal cut-off value of SEC14L1 protein expression was determined at H score of 80. At this cut-off, 70 % cases showed low/negative cytoplasmic expression whereas positive/high cytoplasmic expression was detected in 30%.

There was an association between SEC14L1 protein expression and LVI status ($p<0.0001$) and other variables of poor prognosis (Table 2). SEC14L1 expression was associated with a higher histological grade ($p=0.011$), HER2 positivity ($p=0.004$), and luminal B subtype ($p=0.006$). Outcome analysis revealed a positive association between SEC14L1 protein expression and shorter survival ($p=0.0008$; Figure 1C). In univariate analysis, SEC14L1 protein expression ($p=0.001$), histological grade ($p<0.0001$), LVI status ($p=0.002$), tumor size ($p=0.005$) and nodal status ($p=0.013$) were all significant predictors of outcome. Multivariate analysis including prognostic variables significant in univariate analysis SEC14L1 protein expression was an independent prognostic variable of survival ($p=0.019$; Table 3).

DISCUSSION

In the present research, we have tried to decipher the molecular mechanism underlying LVI in BC. High throughput molecular techniques coupled with bioinformatics and strict definition of LVI status were utilized to identify genes

associated with LVI that can potentially be used as therapeutic targets. Following internal and external validation, SEC14L1 was identified as one of the top differentially expressed genes associated with LVI and also with prognosis. Further examination of SEC14L1 at protein expression level confirmed its association with LVI and indicated its independent prognostic value in BC.

SEC14L1 gene is located within a discrete region of 17q25 that frequently shows copy number alterations in BC [17]. A proximate locus, the17q23 locus is well documented in BC, and the key feature of this amplified locus is its oncogenic activity in BC both in vitro and *in vivo* [18]. In this study, there was an association not only between SEC14L1 CNA and mRNA expression but also between CNA and mRNA expression and LVI status. In a recent investigation of prostatic carcinoma, a prognostic 12-gene signature associated with criteria of aggressive clinical outcome included SEC14L1, also confirmed by real-time quantitative RT-PCR and immunohistochemistry assays [19]. SEC14L1 cytoplasmic protein overexpression was frequently found in the Transmembrane Protease, Serine 2 (TMPRSS2)/ Ets-related gene (ERG) fusion-positive prostate cancer, and the clinical and prognostic power of SEC14L1 strongly depends on this fusion status in prostate cancer [21]. SEC14 like proteins have been implicated in hepatocellular carcinomas [20].

TMPRSS2/ERG fusion has been revealed to be associated with several molecular alterations including deletion of the phosphatase and tensin homolog gene (PTEN) [22]. Several previous studies suggested that PTEN loss is strongly correlated with transcriptional instability and was associated with poor outcome of HER2 positive BC [23, 24]. TMPRSS2/ERG fusion was also thought to be correlated with androgen receptor activity [25], transcriptional stability [26], and stem cell maintenance [27] in cancer. These mechanisms have essential roles in cell proliferation, adhesion,

invasion and migration. Thus, SEC14L1 may play a pivotal role in the regulation of cell growth/tissue development and cell adhesion, underlying LVI and metastasis in breast cancer.

Moreover, this study has revealed that SEC14L1 protein expression had a significant relationship with HER2 status and high histological grade. We previously found that definite LVI was significantly correlated with HER2 positivity [1] and tumor microenvironment plays a crucial role in the HER2 signaling pathway, invasion and metastasis including the development of LVI [28, 29]. A previous study suggested that SEC14L1 might play an important role in trafficking proteins on cellar membrane [30]. SEC14L1 belongs to SEC14 cytosolic factor family that plays a role in the intracellular transport system [31] and innate immunity [32]. SEC14L1 overexpression may be responsible for enhanced intra-tumor signaling pathways including HER2 and influence the tumor microenvironment to promote tumor growth, promoting LVI and metastasis in BC. Although SEC14L1 was associated with higher histological grade, the association with outcome was independent of grade. In this study, SEC14L1 was associated with lymph node metastasis at mRNA levels, and its expression at mRNA and protein levels was associated with outcome independent of other prognostic variables.

In conclusion, this study revealed and confirmed that SEC14L1 expression is a significant prognostic factor of BC. Over-expression of SEC14L1 plays a crucial role in the development of LVI, BC progression and metastasis. Further functional assessment of SEC14L1 to determine its therapeutic value in BC is warranted.

REFERENCES

1. Rakha, E.A., et al., *The prognostic significance of lymphovascular invasion in invasive breast carcinoma.* Cancer, 2012. **118**(15): 3670-3680.
2. Zhang, S., et al., *The relationship of lymphatic vessel density, lymphovascular invasion, and lymph node metastasis in breast cancer: a systematic review and meta-analysis.* Oncotarget, 2017. **8**(2): 2863-2873.
3. Liu, Y.L., et al., *Lymphovascular invasion is an independent predictor of survival in breast cancer after neoadjuvant chemotherapy.* Breast Cancer Res Treat, 2016. **157**(3): 555-564.
4. Mohammed, R.A., et al., *Objective assessment of lymphatic and blood vascular invasion in lymph node-negative breast carcinoma: findings from a large case series with long-term follow-up.* J Pathol, 2011. **223**(3): 358-365.
5. Khwaja, S.S., et al., *Lymphovascular space invasion and lack of downstaging after neoadjuvant chemotherapy are strong predictors of adverse outcome in young women with locally advanced breast cancer.* Cancer Med, 2016. **5**(2): 230-238.
6. Aleskandarany, M.A., et al., Molecular Mechanisms Underlying Lymphovascular Invasion in Invasive Breast Cancer. Pathobiology, 2015. **82**(3-4): 113-123.
7. Aleskandarany MA, Sonbul S, Surridge R, Mukherjee A, Caldas C, Diez-Rodriguez M et al. Rho-GTPase activating-protein 18: a biomarker associated with good prognosis in invasive breast cancer. Br J Cancer 2017; 117: 1176–1184.
8. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ et al. The genomic and transcriptomic architecture of 2,000 breast tumors reveals novel subgroups. Nature 2012; 486: 346–352.
9. Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat Commun 2016; 10: 11479.
10. Mohammed RA, Martin SG, Mahmmod AM, Macmillan RD, Green AR, Paish EC et al. Objective assessment of lymphatic and blood vascular invasion in lymph node-negative breast carcinoma: findings from a large case series with long-term follow-up. J Pathol 2011; 223: 358–365.
11. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43(7): e47.
12. Jézéquel P, Campone M, Gouraud W, Charbonnel C, Leux C, Ricolleau G et al. bc-GenExMiner: an easy-to-use online platform for gene prognostic analyses in breast cancer. Breast Cancer Res Treat 2012; 131: 765–775.
13. Rakha, E.A., et al., Updated UK Recommendations for HER2 assessment in breast cancer. J Clin Pathol, 2015. 68(2): p. 93-9.
14. Rakha, E.A., et al., Prognostic stratification of oestrogen receptor-positive HER2-negative lymph node-negative class of breast cancer. Histopathology, 2017. 70(4): p. 622-631.
15. Hammond, M.E., et al., American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in

breast cancer (unabridged version). Arch Pathol Lab Med, 2010. 134(7): p. e48-72.
16. McCarty KS Jr, Miller LS, Cox EB, Konrath J, McCarty KS Sr. Estrogen receptor analyses. Correlation of biochemical and immunohistochemical methods using monoclonal antireceptor antibodies. Arch Pathol Lab Med 1985; 109: 716–721.
17. Kalikin LM, Bugeaud EM, Palmbos PL, Lyons RH Jr, Petty EM (2001) Genomic characterization of human SEC14L1 splice variants within a 17q25 candidate tumor suppressor gene region and identification of an unrelated embedded expressed sequence tag. Mamm Genome;12: 925-929.
18. Sinclair, C.S., et al., The 17q23 Amplicon and Breast Cancer. Breast Cancer Research and Treatment, 2003. 78(3): p. 313-322.
19. Agell, L., et al., A 12-gene expression signature is associated with aggressive histological in prostate cancer: SEC14L1 and TCEB1 genes are potential markers of progression. Am J Pathol, 2012. 181(5): p. 1585-94.
20. Zhao S, Xu C, Qian H, Lv L, Ji C, Chen C et al (2008) Cellular retinaldehyde-binding protein-like (CRALBPL), a novel human Sec14p-like gene that is upregulated in human hepatocellular carcinomas, may be used as a marker for human hepatocellular carcinomas. DNA Cell Biol; 27: 159-163.
21. Burdelski, C., et al., Saccharomyces cerevisiae-like 1 overexpression is frequent in prostate cancer and has markedly different effects in Ets-related gene fusion-positive and fusion-negative cancers. Hum Pathol, 2015. 46(4): p. 514-23.
22. Carver BS, Tran J, Gopalan A, Chen Z, Shaikh S, Carracedo A, Alimonti A, Nardella C, Varmeh S, Scardino PT, Cordon-Cardo C, Gerald W, Pandolfi PP. Aberrant ERG expression cooperates with loss of PTEN to promote cancer progression in the prostate. Nat Genet. 2009. 41(5):619-24.
23. Puc J, Keniry M, Li HS, Pandita TK, Choudhury AD, Memeo L, Mansukhani M, Murty VV, Gaciong Z, Meek SE, Piwnica-Worms H, Hibshoosh H, Parsons R. Lack of PTEN sequesters CHK1 and initiates genetic instability. Cancer Cell. 2005. 7(2):193-204.
24. Stern HM, Gardner H, Burzykowski T, Elatre W, O'Brien C, Lackner MR, Pestano GA5, Santiago A4, Villalobos I4, Eiermann W6, Pienkowski T7, Martin M, Robert N, Crown J, Nuciforo P, Bee V, Mackey J, Slamon DJ, Press MF. PTEN Loss Is Associated with Worse Outcome in HER2-Amplified Breast Cancer Patients but Is Not Associated with Trastuzumab Resistance. Clin Cancer Res. 2015. 1;21(9):2065-74.
25. Yu J, Yu J, Mani RS, Cao Q, Brenner CJ, Cao X, Wang X, Wu L, Li J, Hu M, Gong Y, Cheng H, Laxman B, Vellaichamy A, Shankar S, Li Y, Dhanasekaran SM, Morey R, Barrette T, Lonigro RJ, Tomlins SA, Varambally S, Qin ZS, Chinnaiyan AM. An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. Cancer Cell. 2010; 17(5):443-54.
26. Shah AV, Birdsey GM, Randi AM. Regulation of endothelial homeostasis, vascular development and angiogenesis by the transcription factor ERG. Vascul Pharmacol. 2016;86:3-13.

27. Polson ES, Lewis JL, Celik H, Mann VM, Stower MJ, Simms MS, Rodrigues G, Collins AT, Maitland NJ. Monoallelic expression of TMPRSS2/ERG in prostate cancer stem cells. Nat Commun. 2013;4:1623.
28. Banin-Hirata BK, de Oliveira CEC, Losi-Guembarovski R, Ozawa PMM, Vitiello GAF, de Almeida FC, Derossi DR, André ND, Watanabe MAE. (2017) The prognostic value of regulatory T cells infiltration in HER2-enriched breast cancer microenvironment. Int Rev Immunol. 19:1-7.
29. Mohammed RA1, Ellis IO, Lee AH, Martin SG. Vascular invasion in breast cancer; an overview of recent prognostic developments and molecular pathophysiological mechanisms. Histopathology. 2009; 55(1):1-9.
30. Mousley CJ, Tyeryar KR, Vincent-Pope P, Bankaitis VA (2007) The Sec14-superfamily and the regulatory interface between phospholipid metabolism and membrane trafficking. Biochim Biophys Acta; 1771: 727-736.
31. Ribeiro, F.M., et al., SEC14-like protein 1 interacts with cholinergic transporters. Neurochem Int, 2007. 50(2): p. 356-64.
32. Li, M.T., et al., Negative regulation of RIG-I-mediated innate antiviral signaling by SEC14L1. J Virol, 2013. 87(18): p. 10037-46.

Acknowledgements

Cell lysates were obtained courtesy of Prof. Stewart Martin (Faculty of Medicine & Health Sciences, The University of Nottingham). We would also like to show our gratitude to Dr. Oscar Rueda (CRUK Cambridge Research Institute, University of Cambridge, Cambridge, UK) for providing us with the clinical, genomic, and transcriptomic profiles of the METABRIC patients in this study. We are also immensely thankful to Dr. Elena Provenzano (Lead Breast Histopathologist, Addenbrooke's Hospital, Cambridge Breast Unit, Cambridge University Hospital NHS Foundation Trust, Cambridge, UK) and to Dr. Abhik Mukherjee for their validation of LVI status in the patients of Addenbrookes' Hospital cases. We thank the Breast Cancer Research Trust and the University of Nottingham for funding this project and the Nottingham Health Science Biobank and Breast Cancer Now Tissue Bank for the provision of tissue samples.

Author Contributions

Rakha EA, Green AR, Caldas C, Martin S, and Ellis IO conceived of the study, contributed to study design and provided samples and data; Sonbul SN, Aleskandarany MA, Mukherjee A, Toss MS, and Green AR collected data; Sonbul SN carried out experiments; Sonbul SN, Joseph C, and Kurozumi S analysed and interpreted data and generated the figures and tables; all authors contributed to drafting and reviewing the manuscript and approved the submitted and final version.

Conflict of interest

None of the authors have any conflict of interest to declare.

Figure Legends

Figure 1: Kaplan-Meier survival plots showing the association between *SEC14L1* copy numbers (A), mRNA expression (B) and SEC14L1 protein expression (C) and outcome.

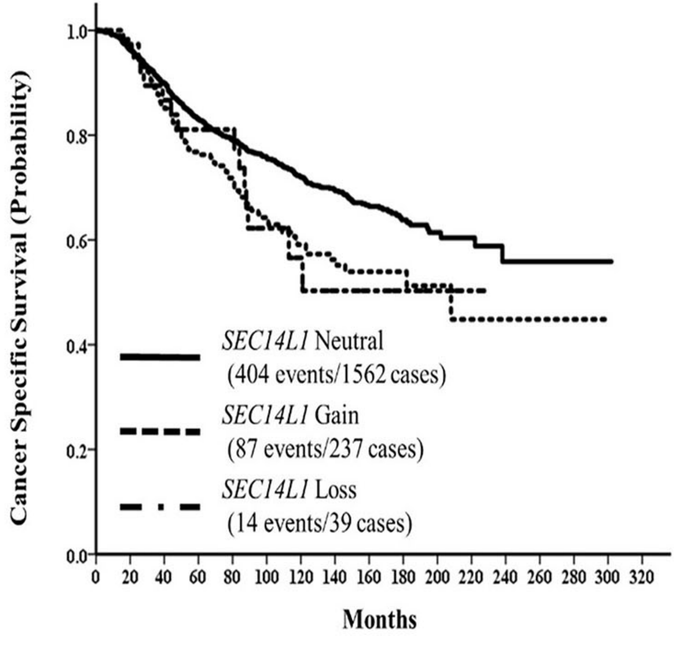
Figure 2: SEC14L1 protein expression in normal and breast cancer tissue.

Cytoplasmic SEC14L1 expression was overexpressed in breast cancer cells compared to the expression in epithelial cells of normal duct lobular units (Black arrow: invasive carcinoma and white arrow: normal mammary gland).

Figure 3: Immunohistochemistry expression of SEC14L1 in tissue microarray images.

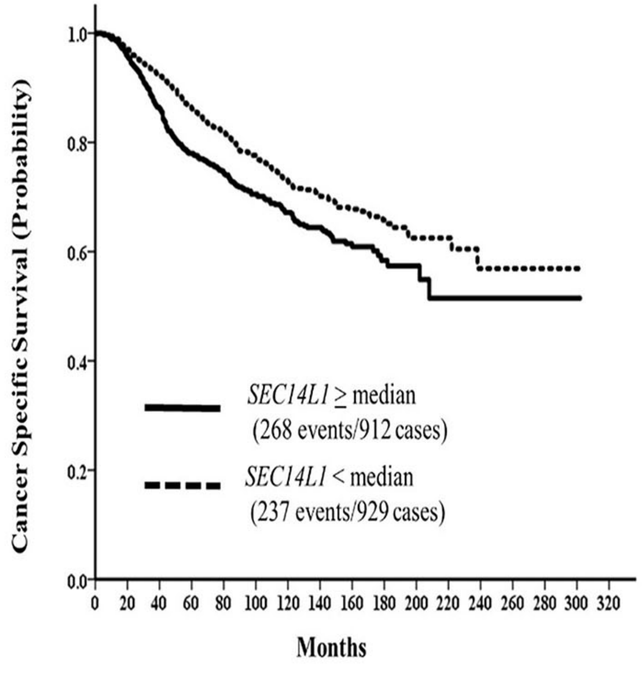
SEC14L1 expression of the cytoplasm in cancer cells was distributed as follows; a) No staining, b) weak staining, c) moderate staining, and d) strong staining.

A)



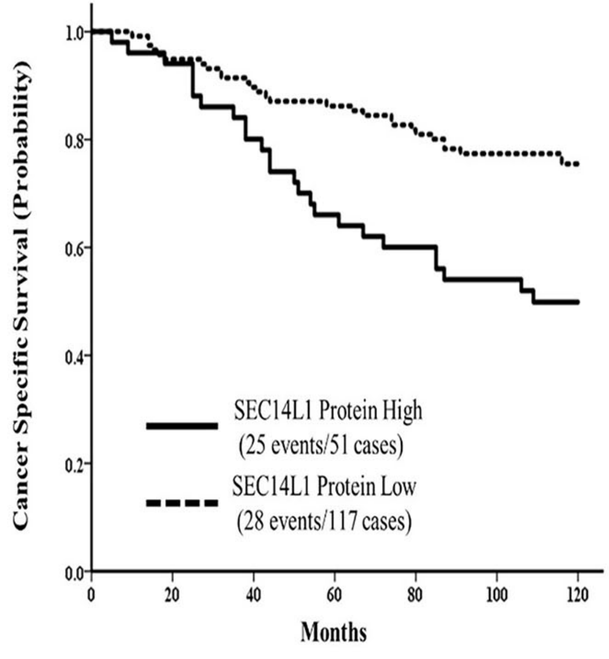
Copy numbers of *SEC14L1:*
Gain vs. Neutral: Hazard Ratio: 11.64, p= 0.00065
Loss vs. Neutral: Hazard Ratio: 2.28, p= 0.13

B)

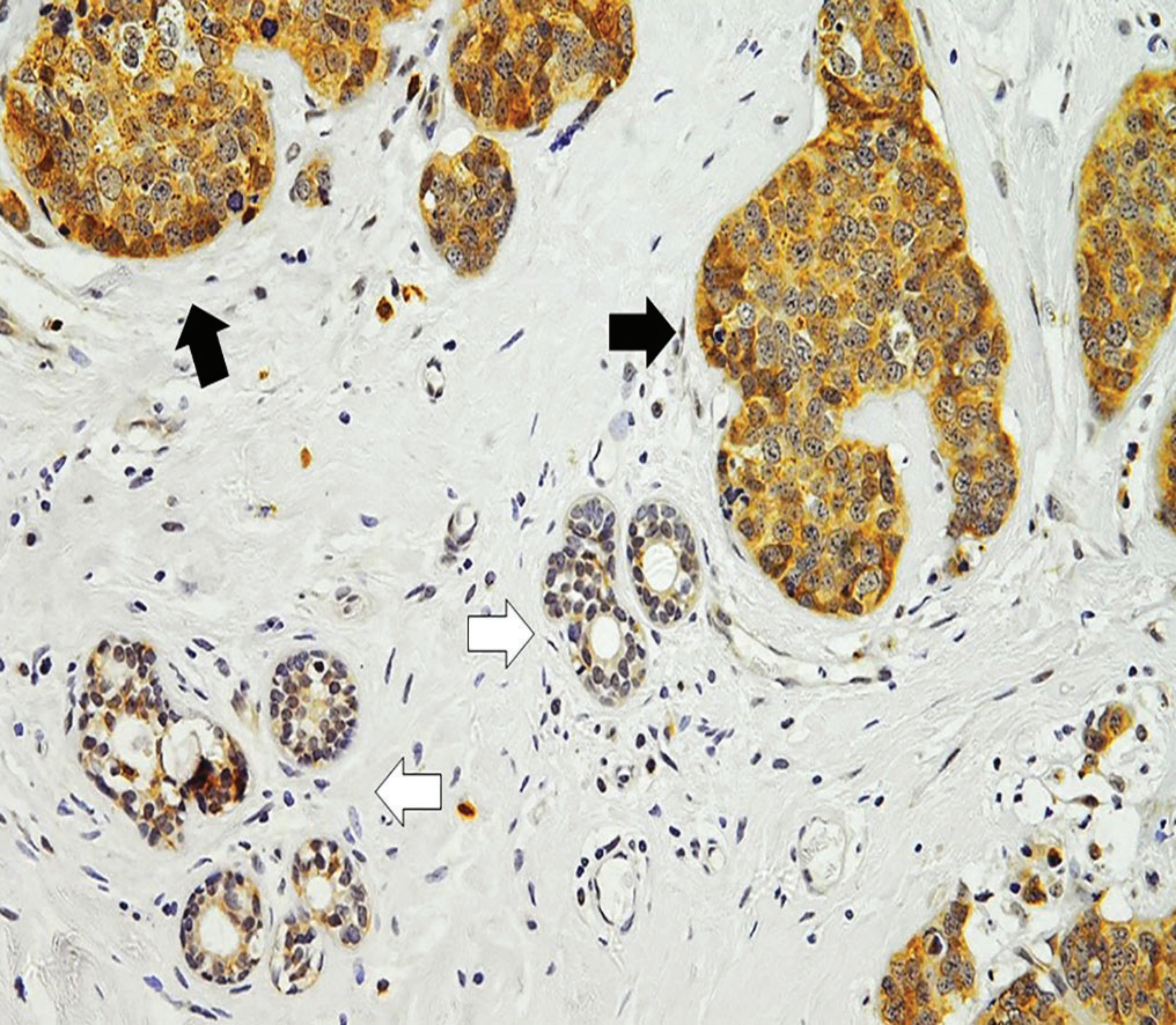


mRNA expression of *SEC14L1:*
≥ median vs. < median: Hazard Ratio: 12.17, p= 0.00049

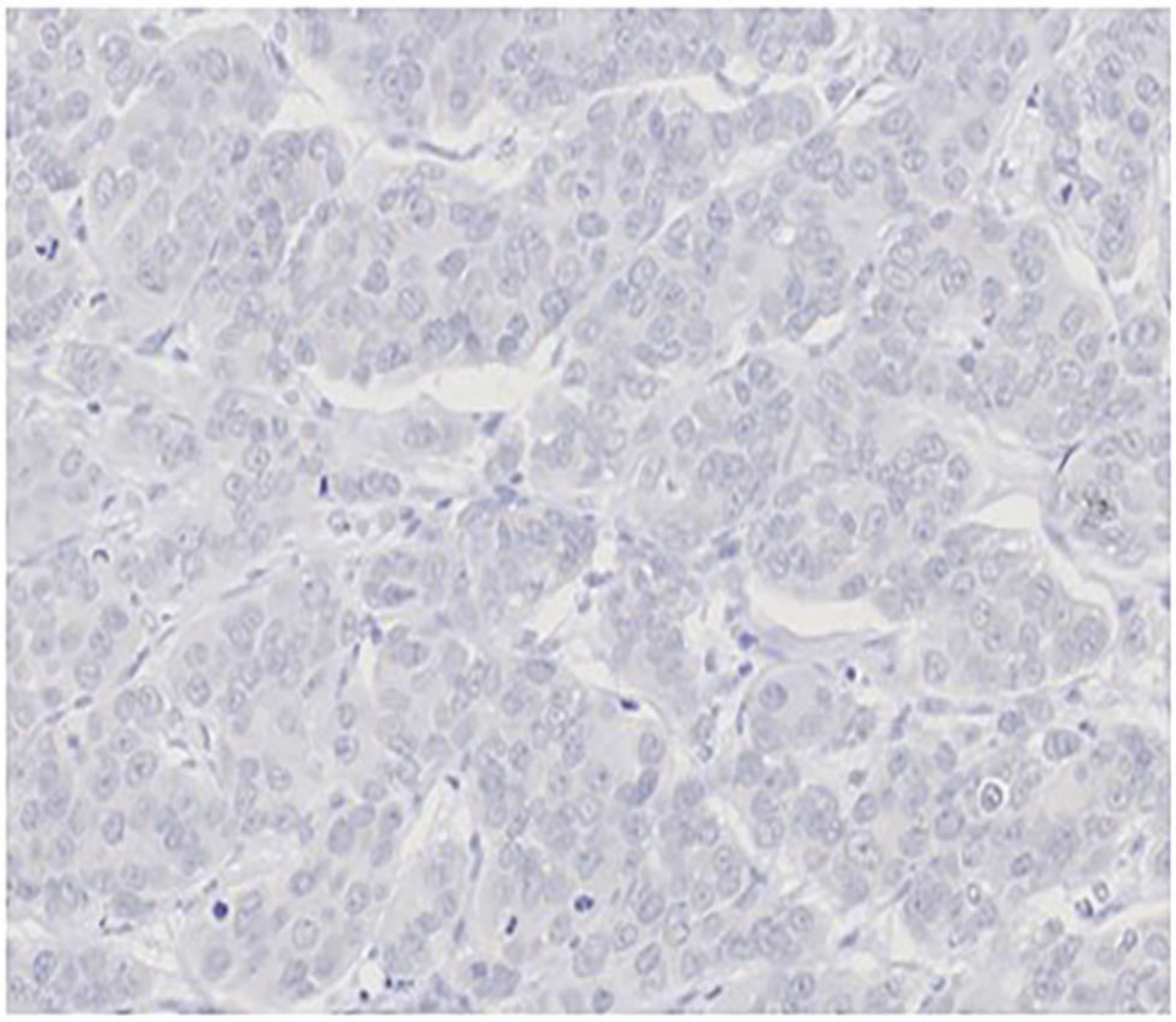
C)



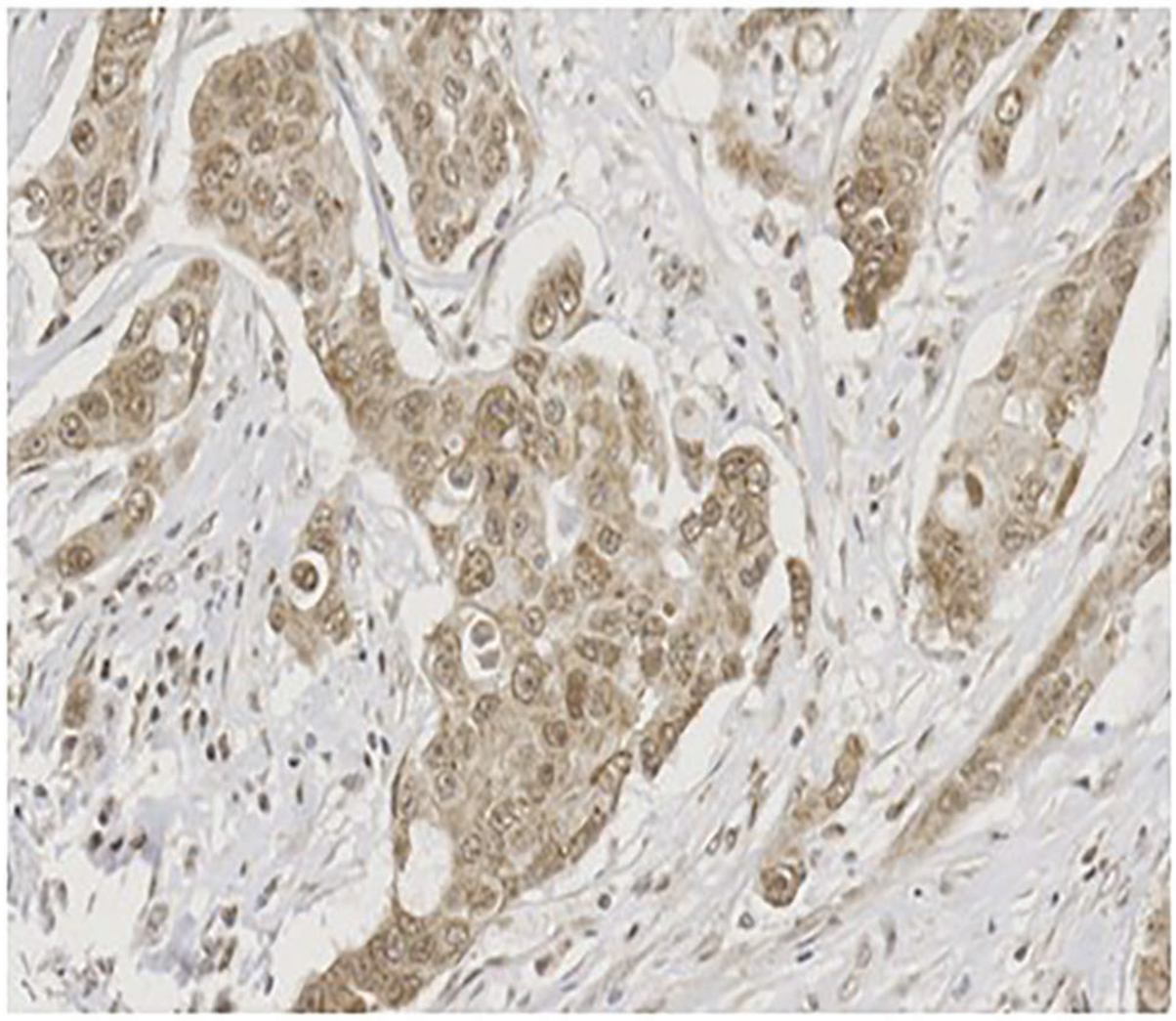
Protein expression of SEC14L1
≥ H-socre 90 vs. < H-score 90: Hazard Ratio: 11.37, p= 0.00075



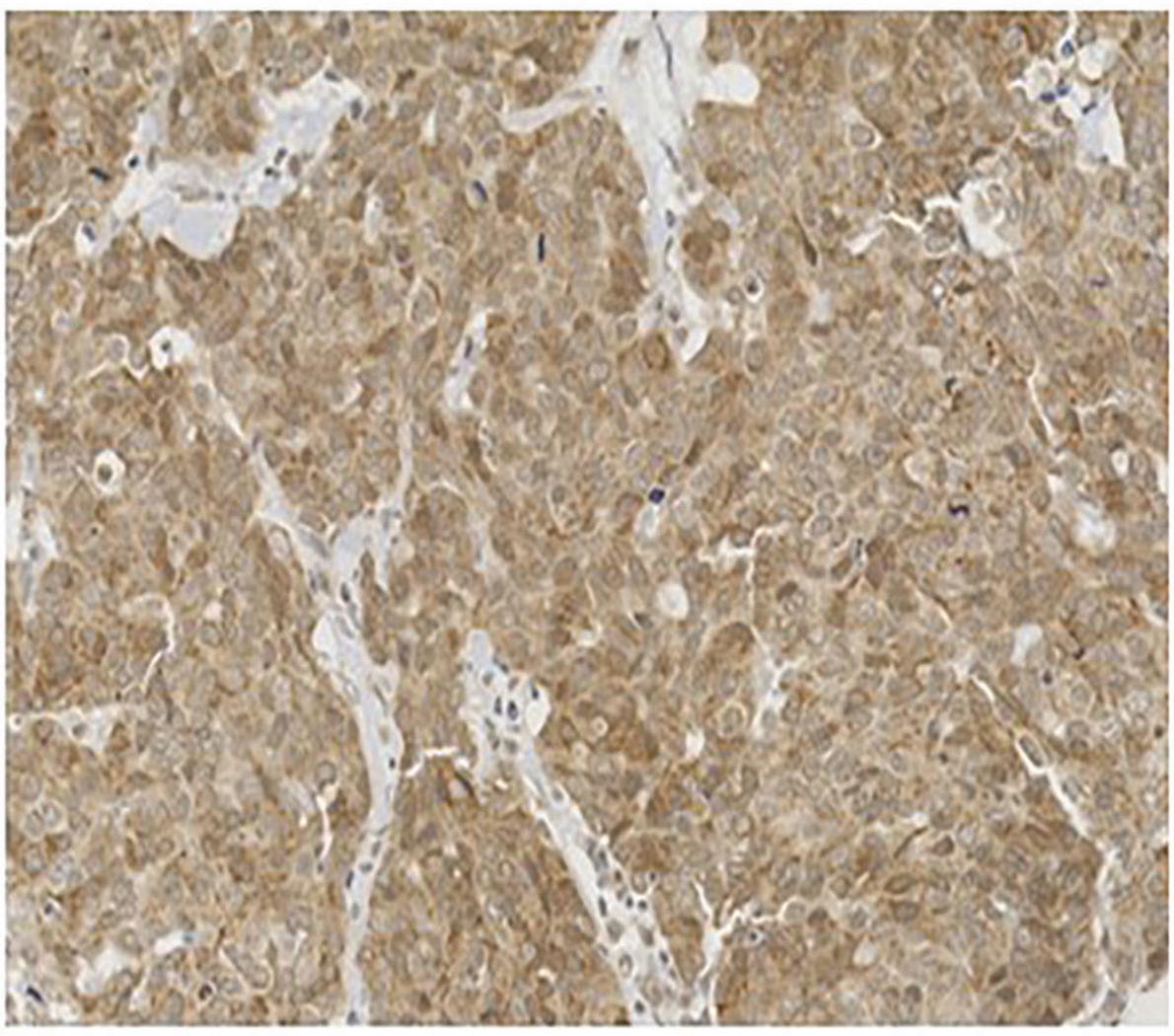
A)



B)



C)



D)

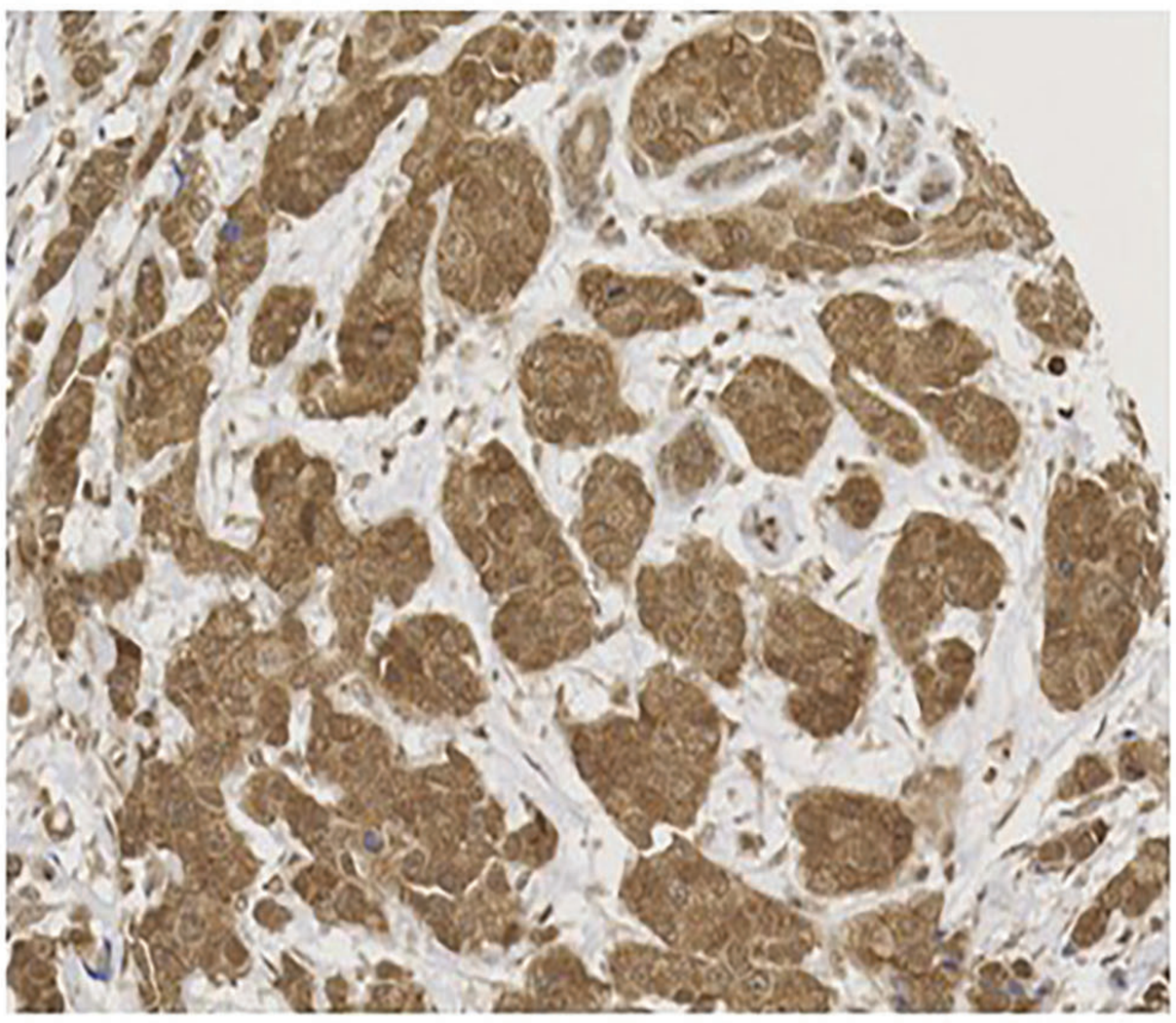


Table 1: Association between *SEC14L1* copy number alterations and *SEC14L1* mRNA expression and clinicopathological parameters in the METABRIC cohort of invasive breast cancer (n=1980)

Factors		Expression of *SEC14L1* (copy number alterations)					Expression of *SEC4L1* (mRNA)			
		Loss	Neutral	Gain	Total	*p*-value	> Median	≤ Median	Total	*p*-value
Tumour size	≥ 2cm	30 (2.2%)	1114 (83.3%)	193 (14.4%)	1337	0.063	657 (49.1%)	681 (50.9%)	1338	0.10
	< 2cm	13 (2.1%)	536 (87.3%)	65 (10.6%)	614		327 (53.1%)	289 (46.9%)	616	
Nodal status	Positive	20 (2.1%)	787 (84.1%)	129 (13.8%)	936	0.76	413 (44.0%)	525 (56.0%)	938	<0.0001
	Negative	23 (2.2%)	880 (85.1%)	131 (12.7%)	1034		575 (55.6%)	460 (44.4%)	1035	
Histological grade	Grade 3	22 (2.3%)	764 (80.3%)	165 (17.4%)	951	<0.0001	355 (37.3%)	597 (62.7%)	952	<0.0001
	Grade 1, 2	21 (2.2%)	834 (88.9%)	83 (8.8%)	938		589 (62.7%)	351 (37.3%)	940	
Molecular subtypes	Luminal A	14 (2.0%)	651 (90.8%)	52 (7.3%)	717	<0.0001	481 (67.0%)	237 (33.0%)	718	<0.0001
	Luminal B	15 (3.1%)	358 (73.4%)	115 (23.6%)	488		252 (51.6%)	236 (48.4%)	488	
	HER2 enriched	5 (2.1%)	201 (84.1%)	33 (13.8%)	239		78 (32.5%)	162 (67.5%)	240	
	Basal-like	5 (1.5%)	275 (83.8%)	48 (14.6%)	328		74 (22.5%)	255 (77.5%)	329	
	Normal-like	4 (2.0%)	183 (92.0%)	12 (6.0%)	199		101 (50.8%)	98 (49.2%)	199	

Table 2: Association between SEC14L1 protein expression and the clinicopathological factors

Factors		Expression of SEC14L1			
		Low	High	Total	*p*-value
Tumor size	≥ 2cm	68 (68.7%)	31 (31.3%)	99	0.75
	< 2cm	49 (71.0%)	20 (29.0%)	69	
Nodal status	Positive	55 (67.1%)	27 (32.9%)	82	0.48
	Negative	62 (72.1%)	24 (27.9%)	86	
Histological grade	Grade 3	53 (60.9%)	34 (39.1%)	87	0.011
	Grade 1, 2	64 (79.0%)	17 (21.0%)	81	
Lymphovascular invasion	Positive	43 (51.8%)	40 (48.2%)	83	<0.0001
	Negative	74 (88.1%)	10 (11.9%)	84	
Estrogen receptor	Positive	85 (69.1%)	38 (30.9%)	123	0.80
	Negative	32 (71.1%)	13 (28.9%)	45	
Progesterone receptor	Positive	66 (70.2%)	28 (29.8%)	94	0.96
	Negative	44 (69.8%)	19 (30.2%)	63	
HER2	Positive	14 (53.8%)	12 (46.2%)	26	0.036
	Negative	98 (74.2%)	34 (25.8%)	132	
Molecular subtypes	Luminal A-like	30 (88.2%)	4 (11.8%)	34	0.006
	Luminal B-like	39 (65.0%)	21 (35.0%)	60	
	HER2-positive	14 (53.8%)	12 (46.2%)	26	
	Triple negative	23 (85.2%)	4 (14.8%)	27	
Abbreviations: ER, Estrogen receptor; PgR, Progesterone receptor					

Table 3: Univariate and multivariate survival analysis of clinicopathological factors and SEC14L1 protein expression in breast cancer

Factors		Univariate analysis			Multivariate analysis		
		Hazard Ratio	95% CI	*p*-value	Hazard Ratio	95% CI	*p*-value
SEC14L1	Low	Reference			Reference		
	High	2.5	1.43-4.21	0.001	2.0	1.12-3.54	0.019
Lymphovascular invasion	Negative	Reference			Reference		
	Positive	2.5	1.41-4.47	0.002	1.5	0.79-2.74	0.22
Tumor size	< 2cm	Reference			Reference		
	≥ 2cm	2.4	1.30-4.43	0.005	1.8	0.96-3.47	0.067
Nodal status	Negative	Reference			Reference		
	Positive	2.0	1.16-3.53	0.013	1.4	0.76-2.47	0.29
Histological grade	Grade1-2	Reference			Reference		
	Grade 3	3.7	1.96-6.87	<0.0001	2.7	1.41-5.15	0.003