THE ROLE OF GENES CO-AMPLIFIED WITH NICASTRIN IN BREAST INVASIVE CARCINOMA

A. SARAJLIĆ[1], A. FILIPOVIĆ[2], V. JANJIĆ[1], R.C. COOMBES[2] and N. PRŽULJ[1]∗

[1] Department of Computing, Imperial College London,
180 Queen's Gate, Huxley, London, SW7 2AZ, UK

[2] Faculty of Medicine, Department of Surgery and Cancer, Imperial College London,
The Hammersmith Hospital, Du Cane Road, London, W12 0NN, UK

∗E-mail: natasha@imperial.ac.uk
Phone: +44-(0)-207-594-1516

1 Abstract

Breast cancer accounts for more than 450,000 deaths per year worldwide. Discovery of novel therapeutic targets that will allow patient-tailored treatment of this disease is an emerging area of scientific interest. Recently, nicastrin (NSCTN) has been identified as one such therapeutic target. Its overexpression is indicative of worse overall survival in the estrogen-receptor negative patient population. In this paper we analyse data from a large invasive breast carcinoma study and confirm nicastrin amplification. In search for genes that are co-amplified with nicastrin, we identify a potential novel breast cancer related amplicon located on chromosome 1. Furthermore, we search for "influential interactors", *i.e.,* genes that interact with a statistically significantly high number of genes which are co-amplified with nicastrin, and confirm their involvement in this female neoplasm. Among influential interactors we find genes which belong to the Core Diseasome (a recently identified therapeutically relevant set of genes which is known to drive disease formation) and propose that they might be important for breast cancer onset, and serve as its novel therapeutic targets. Finally, we identify a pathway that may play a role in nicastrin's amplification process and we experimentally confirm downstream signaling mechanism of nicastrin in breast cancer cells.

Keywords: Breast invasive carcinoma; Nicastrin; Co-amplified genes; Protein-protein interaction network.

2 Introduction

Breast cancer accounts for more than 1,300,000 new cases and 450,000 deaths yearly worldwide [1]. As such, it is the most common female neoplasm. Heterogeneity of breast cancer is reflected in the fact that there are at least seven distinct molecular subtypes of the common estrogen receptor positive disease [2] and six distinct groups of triple negative (estrogen, progesterone and HER2 receptor negative) breast cancers [3]. Such molecular complexity highlights the need for the discovery of novel therapeutic targets that are oncogenic drivers which are overexpressed and/or amplified in individual tumors, allowing for

a precision medicine approach and patient-tailored treatment. Biologically, clinically and functionally relevant targets in cancer often undergo high-level DNA amplification, conferring cancer cells with survival advantage. In solid malignancies, such as breast cancer, some specific regions in the genome harbour amplicons such as the HER2-related amplicon on chromosome 17 [4], as well as the chromosome 1 quantitative anomalies present in 50 – 60% of breast tumors [5].

Discovery and validation of novel therapeutic targets in cancer is a complex process which involves: i) understanding the exact expression patterns of a particular protein target in normal tissues and cancer; ii) evidence of the targets clinical relevance in the disease, as a prognostic or a predictive biomarker; and iii) biological ways and molecular mechanisms in which this target drives cancer. Nicastrin (NCSTN) is one such novel therapeutic target, recently identified and characterized by the group of R.C. Coombes[6]. It is a member of the γ-secretase enzyme complex, which cleaves and activates multiple oncogenic substrates, contributing to cancer development and progression. Its overexpression at the protein level is readily detectable in 47.5% of breast cancers, conferring worse overall survival in the estrogen-receptor negative patient population, whose tumors are by nature deemed more aggressive[6] and it is closely correlated with increased mRNA levels of the target[7]. Furthermore, nicastrin facilitates epithelial to mesenchymal transition in breast cancer cells, a process crucial for metastatic spread and development of treatment resistance. Its genetic depletion successfully attenuates breast tumor growth in *in vivo* xenograft models, with potent depletion of Notch signaling and breast cancer stem cell activity [8]. For the purpose of targeting nicastrin in invasive breast cancer, Filipović *et al.* have developed blocking monoclonal antibodies which confer significant anti-proliferative and anti-metastatic effects *in vitro* and *in vivo*[7]. Clinical administration of anti-nicastrin antibodies would call for careful characterization of patient eligibility for this type of treatment.

Guided by the example of the anti-HER2 antibody therapy, in this study we aim to investigate nicastrin gene amplification status in a large patient cohort available on the cBio Cancer Genome Portal[9], with the notion that detection of nicastrin gene amplification in patients tumors would help stratify those who would be candidates for anti-nicastrin therapy. In this cohort we: 1) confirm nicastrin gene amplification, 2) identify genes that are co-amplified with nicastrin, and 3) examine their potential biological implications in breast cancer based on protein-protein interaction (PPI) network analysis, validated by our experimental data and literature review.

We look into the network of human PPIs to find proteins which interact with a statistically significantly high number of proteins encoded by genes that are co-amplified with nicastrin (we call these *influential interactors*), and we corroborate their relation to breast cancer. The motivation behind the neighborhood analysis comes from earlier successes in extracting new biological knowledge from different biological networks using graph theoretic approaches. It has been shown that proteins which are closer in a PPI network are more likely to perform the same function [10]. In particular, *guilt by association* approach was used to infer functions of functionally unannotated proteins: the direct neighborhoods of proteins in the network were examined in search for the most common functions among annotated direct neighbors [11]. Similarly, the n-neighborhood of proteins [12] and shared neighbors of proteins [13] were used to assign functions to functionally unannotated proteins. When it comes to exploring the connection between topology around protein and its involvement in disease, it has been shown that directly linked proteins in the human PPI network are more likely to cause similar diseases [14, 15]. Goldenberg *et al.* [16] used gene and protein interaction network to identify genes which are important for initiation and progression

of lung cancer. They identified a small set of influential genes by examining genes whose neighbors show high expression change in cancerous tissue versus normal tissue, regardless of their own expression.

We also ask whether any of the influential interactors has a key role in the onset of breast cancer by examining their presence in the Core Diseasome [17]. The Core Diseasome is a recently identified topologically and functionally homogeneous "core sub-network" of the human PPI network, which is enriched in disease genes, drug targets, and a small number of genes that have theoretically been proposed to be absolutely required for tumor formation, called "driver genes"[18]. Proteins in this sub-network are postulated to be the key to disease onset and progression and hence should be the primary object of therapeutic intervention. The relation between Core Diseasome and breast cancer has not been examined before, and this paper offers a new perspective on that subject.

3 Methods

3.1 Datasets

We obtain the human PPI network from BioGRID [19] (data downloaded in January 2013). It contains 13,953 proteins and 97,954 interactions. We download the drug-target data from DrugBank[1] (data downloaded in January 2013). To find genes that are co-amplified with nicastrin in breast cancer, we obtain data from breast cancer study from cBio Cancer Genomics Portal [9] (data downloaded in January 2013): "Breast Invasive Carcinoma (TCGA, Nature 2012)" study [20]. The study contains data for 825 tumor cases, out of which we took all genes that were amplified in cases when nicastrin is amplified. Out of the 825 tumor cases in this study, there are 40 where nicastrin is amplified.

We further analyse the following two sets of genes, which we label and define as:

- The "AMPL-stringent" set — this set contains 22 genes, including nicastrin, that are amplified in 95% of "Breast Invasive Carcinoma (TCGA, Nature 2012)" cases in which nicastrin is amplified: ARHGAP30, CD244, CD48, DEDD, F11R, ITLN1, ITLN2, KLHDC9, LY9, NIT1, PFDN2, PVRL4, SLAMF6, SLAMF7, TSTD1, USF1, PEX19, COPA, NCSTN, NHLH1, SUMO1P3, VANGL2. 16 out of these 22 genes are present in our PPI network.

- The "AMPL-extended" set — this set contains 472 genes obtained as follows. For each of the 10,047 amplified genes in the 40 cases when nicastrin is amplified in the study, we ask in how many of the 40 cases it appears. 472 genes appear in at least 19 cases, which is just under 5% of the 10,047 genes. 308 out of these 472 genes are present in our PPI network.

3.2 Functional enrichment and drug targets

We use DAVID database[2] to check enrichment in Gene Ontology (GO) terms (specifically, molecular function, biological processes and cellular component categories of GO) for the two sets of genes that we analyse (AMPL-stringent and AMPL-extended). We choose 0.05 as a cut-off p-value, after applying Benjamini-Hochberg correction for multiple hypothesis testing [21]. Additionally, to check whether a gene is a known drug target, we refer to the DrugBank drug-target dataset (described in Section 3.1).

[1] DrugBank, http://www.drugbank.ca/
[2] http://david.abcc.ncifcrf.gov

3.3 Finding influential interactors of amplified genes

In order to find proteins whose direct neighborhoods are enriched in proteins encoded by genes that are co-amplified with nicastrin, we perform the following computation for AMPL-stringent and AMPL-extended sets separately.

First, we denote the number of proteins in our entire human PPI network as M, and the size of AMPL-stringent (or AMPL-extended) set as K. Then, for each protein in our PPI network, we count: 1) the number of its direct neighbors in the PPI network, and denote it as N; and 2) how many genes from AMPL-stringent (or AMPL-extended) set are among its direct neighbors, and denote that by X. Finally, the p-value, or the probability that X or more amplified genes are found among neighbors of the gene in question, is computed as:

$$p = 1 - \sum_{i=0}^{X-1} \frac{\binom{K}{i}\binom{M-K}{N-i}}{\binom{M}{N}}.$$

We use 0.05 as a cut-off p-value to find genes in the PPI network that have among their neighbors a statistically significantly high number of genes from AMPL-stringent (or AMPL-extended) set. Henceforth, we refer to these genes as *influential interactors*.

3.4 Assesing validity of the neighborhood analyses in an in vivo experimental model system of nicastrin inhibition

In order to explore whether modulation of nicastrin signaling would affect gene expression levels of the genes revealed in the neighbourhood analyses as members of potential pathway associated with nicastrin amplification, we analyse metastatic tumour deposits from the in vivo xenograft model treated with anti-nicastrin monoclonal antibodies. After having received 5 doses of the anti-nicastrin monoclonal antibodty treatment and the control isotype IgG antibody, mice were sacrificed and lungs were harvested for further analyses. mRNA was extracted from lung metastatic deposits and subjected to reverse transcription to produce cDNA and subsequently to RT-qPCR for assessment of annotated gene expression levels.

3.4.1 Experimental metastasis model

MDA-MB-231-luc-D3H2LN cells (1 x 106) were injected into the tail vein of female nude mice (n=8 /group). Mice were imaged weekly from dorsal and ventral views for 5 weeks to monitor development of metastasis. Anti-NCSTN monoclonal antibodies and control rat IgG were administered i.v. at 50 mg/kg in 5-daily intervals. At termination, lungs were harvested by dissection. Half of the lung tissue was paraffin embedded and processed for hematoxilin and eosin, as well as Ki67 immunohistochemistry staining. The other half was snap frozen and consequently processed for mRNA extraction and RT-qPCR assessment of the chosen target genes.

3.4.2 RNA isolation and quantitative RT-PCR

Isolation of total RNA was performed using the RNeasy kit (Qiagen, UK). RNA (1g) was reverse transcribed using oligo dT and Transcriptor First Strand cDNA Synthesis Kit (Roche, UK). Quantitative PCR was performed using the SYBR Green PCR mastermix (Applied Biosystems, UK) on a 7900HT

Thermocycler (Applied Biosystems, Warrington, UK). The PCR cycle number that generated the first fluorescence signal above a threshold (threshold cycle, CT; 10 standard deviations above the mean fluorescence generated during the baseline cycles) was determined, and a comparative CT method was then used to measure relative gene expression. Primer sequences used are shown in Table 1.

Gene	Forward primer	Reverse primer
GAPHD	5- TGAAGGTCGGAGTCAACGGATTT-3	5- GCCATGGAATTTGCCATGGGTGG-3
S100A8	5-ATGCCGTCTACAGGGATGAC-3	5-ACGCCCATCTTTATCACCAG-3
S100A9	5-CAGCTGGAACGCAACATAGA-3	5-TCSGCTGCTTGTCTGCATTT-3
CXCL1	5-AGGGAATTCACCCCAAGAAC -3	5-CACCAGTGAGCTTCCTCCTC-3
CXCL2	5-CTCAAGAATGGGCAGAAAGC-3	5-AAACACATTAGGCGCAATCC-3

Table 1. Primer sequences. First column: genes. Second column: forward primers. Third column: reverse primers.

4 Results

For each of the two sets of genes described in Section 3.1 — namely, AMPL-stringent and AMPL-extended — we perform functional enrichment analysis to inspect their role in breast cancer. We also look at their influential interactors (described in Section 3.3) to try to find those which belong to key pathways in breast cancer onset, or which might be causal to the process of amplification. We check if any of the influential interactors are in the Core Diseasome which is postulated to be the key to disease onset and progression.

4.1 AMPL-stringent: a novel amplicon in breast cancer?

We check the cytogenetic location of genes from AMPL-stringent set and find that all of them are mapped to the q arm of chromosome 1, framing a potential novel *amplicon* relevant in breast cancer patients. This finding is in line with previous observations of relevant genes amplified and overexpressed including 6 genes on 1q21 [22]. Additionally, it has already been observed in cancer studies that recurrent gene amplifications may be required for the survival of cancer cells, and as such constitute novel therapeutic targets. Interestingly, more than 50% of genes from the AMPL-stringent set are localized in the cell membrane, which is a property desirable for human drug targets [23].

Influential interactors	Neighboring proteins from AMPL-stringent	p-value
PTPN11	SLAMF6, LY9	3.003878
SH2D1A	CD244, SLAMF6, LY9	$\approx$0
GTF3C1	USF1, NCSTN	0.001520
SMARCA5	USF1, NCSTN	0.002532
TOP2B	USF1, NCSTN	0.000937
FBXO6	COPA, NCSTN	0.012470
XRCC6	NIT1,USF1	0.015468
ELAVL1	PFDN2, NCSTN, PEX19, USF1, NHLH1, F11R	0.009032

Table 2. Influential interactors for AMPL-stringent set. First column: influential interactors (defined in Section 3.3), *i.e.,* genes in the human PPI network that interact with a statistically significantly high number from AMPL-stringent set. Second column: genes from the AMPL-stringent set that are neighbors to the particular influential interactor. Third column: p-value which denotes the probability that an influential interactor would interact with the same or higher number of genes from the AMPL-stringent set purely by chance.

Since neighborhood analysis has previously been used for predicting functions of proteins and establishing their association with diseases, we look for proteins — *i.e.,* influential interactors described in Section 3.3 — whose neighborhoods are enriched in genes from AMPL-stringent (*i.e.,* genes which are co-amplified with nicastrin). Then, based on the guilt by association principle, we hypotesize that influential interactors might be implicated in breast cancer due to a strong presence of amplified genes in their neighborhoods. We identify 8 influential interactors (listed in Table 2) and further inspect their biological functionality:

- The protein encoded by PTPN11 gene is a member of the protein tyrosine phosphatase (PTP) family and we find it in the Core Diseasome. PTPs are known to be signaling molecules that regulate a variety of cellular processes including cell growth [24], cell cycle progression [25], and oncogenic transformation [26].

- SH2D1A has only 9 interacting partners in the human PPI network, 3 of which are present in AMPL-stringent. It encodes a protein that plays a major role in the bidirectional stimulation of T and B cells and can modify signal transduction pathways on activated T, B and NK cells. It is well known that cancer cells avoid immunosurveillance, and the contribution of T-cells, B- cells and NK-cells is significant in the progression of cancer. Signaling of cancer related immune cells and their cross-talk with tumor cells themselves, as well as with tumor-associaed stroma, has an important role in suppressing tumor-specific immunity [27–29]. This is the reason why much effort is being invested into immune therapy of tumors, as well as development of bi-specific monoclonal antibodies that would be directed both at tumor-specific antigens and tumor-related immune cells.

- GTF3C1 is required for RNA polymerase III-mediated transcription[30].

- SMARCA5 is involved in transcriptional regulation [31]

- TOP2B gene encodes an enzyme which controls topological states of DNA during transcription [32]
- FBXO6 is part of the of a family of human F-box proteins[33] which are critical for the controlled degradation of cellular regulatory proteins[34]. FBXO6 has previously been mentioned as a cancer marker[35].
- XRCC6 is involved in the DNA double-strand break repair pathway, which is importnat for genomic stability and cancer prevention. Genetic variations of this gene are associated with increased risk of cancer, including breast cancer[36].
- ELAVL1 is involved in a variety of biological processes and has been previously linked to cancer as a highly expressed gene[37, 38]. It has a high potential for diagnosing cancer, its prognosis, and therapy[39].

Influential interactors PTPN11 and SH2D1A have common neighbors among the following genes from the AMPL-stringent set: CD244, SLAMF6 and LY9 (see Table 2). These three genes are members of the signaling lymphocyte activation molecule (SLAM) family. This family forms a subset of the larger CD2 cell-surface receptor immunoglobulin superfamily which is involved in cellular activation [40]. This indicates that influential interactors PTPN11 and SH2D1A might, through these three genes, play a role in activation of amplification process. Whether this holds true is a subject of future research. On the other hand, influential interactors GTF3C1, SMARCA5 and TOP2B have two common amplified neighbors: NSCTS and a transcription factor USF1. Having in mind importance of GTF3C1, SMARCA5 and TOP2B in the transcription process, as described above, we suggest exploring their role in co-amplification with nicastrin through regulation of transcription.

4.2 Involvement of AMPL-extended in breast cancer

In order to assess the biological significance of the genes in the AMPL-extended set, we check its enrichment in GO terms and find it to be enriched in functions which represent features closely related to cancer development and progression: epithelial cell differentiation (p-value $= 5.5 \times 10^{-23}$), cell differentiation (p-value $= 1.9 \times 10^{-3}$) and immune response (p-value $= 4.6 \times 10^{-2}$). Also, 117 of AMPL-extended genes are localized on the plasma membrane (p-value $= 5.4 \times 10^{-3}$). Recall that membrane localization is a desirable property of drug target genes [23]. Out of 117, the following 23 genes are already known drug targets: CD1D, CD1A, FCGR3A, SELP, FCGR3B, CD247, S100A12, FCER1G, FCER1A, CCT3, CHRNB2, NTRK1, SV2A, S100A7, HSD17B7, NPR1, FCGR2A, CD55, IL6R, F5, SELE, KCNN3, and SLC19A2.

Analogously to the analysis done for the AMPL-stringent genes, we check the human PPI network for influential interactors of the AMPL-extended gene set and identify 226 of them. We find that 6 of them (listed in Table 3) are present in the Core Diseasome, which highlights their potential importance in the onset of breast cancer. We verify the involvement of these 6 genes in breast cancer through literature:

- SHC1 gene couples activated growth factor receptors to signaling pathways and has already been related to breast cancer [41].
- ABL1, SP1 and PTPN11 are involved in cell growth and were shown to be implicated in breast cancer [42, 43] .

- MDM2 is involved in tumor protein p53 pathway [44] and was also linked to breast cancer [45]. Its over-expression may result in excessive inactivation of tumor protein p53, diminishing its tumor suppressor function.
- LYN transmits signals from cell surface receptors and, among other functions, plays an important role in the regulation of response to growth factors [46]. It was also identified as a mediator of invasion and a possible new therapeutic target, with particular relevance to clinically aggressive basal-like breast cancer [47].

Influential interactors	Neighboring proteins from AMPL-extended	p-value
ABL1	SHC1, PSMD4, NTRK1, SPTA1, SHE, MDM4	0.029954
SP1	ARNT, ZBTB7B, TPM3, NTRK1, POGZ, IFI16, DCAF6, MEF2D, ATF6, POU2F1	0.012628
MDM2	RPS27, PBX1, MDM4, PYHIN1, HIST2H2BE, PSMD4, GORAB, S100A4, S100A6, S100A1, S100A2	0.018274
PTPN11	MPZL1, SELE, FCGR2B, CD84, LY9, SLAMF6, FCRL3, SLAMF1	0.000445
SHC1	HSPA6, DDR2, FCGR2B, HAX1, NTRK1, SPTA1, PIK3C2B, S100A9, S100A8, S100A7, DUSP23, MAPKAPK2	0.001060
LYN	MUC1, FCGR2A,FASLG, ADAM15, SHC1	0.028959

Table 3. Influential interactors for AMPL-extended set that are in the Core Diseasome. First column: influential interactors (defined in Section 3.3), *i.e.,* genes in the human PPI network that interact with a statistically significantly high number from AMPL-extended set that are in the Core Diseasome. Second column: genes from the AMPL-extended set that are neighbors to the particular influential interactor. Third column: p-value which denotes the probability that an influential interactor would interact with the same or higher number of genes from the AMPL-extended set purely by chance.

4.3 A sub-network of NCSTN-co-amplified genes reveals breast cancer cells' propagation inhibiting pathway

In order to inspect weather the amplified genes from AMPL-stringent and AMPL-extended sets form pathways which might play a role in breast cancer, or specifically the amplification process, we check whether these genes form relevant connected components in the human PPI network. We find that the sub-network formed from 16 AMPL-stringent genes present in our PPI network does not form a connected component. However, when we look at the sub-network formed from a larger set of 308 AMPL-extended genes present in our PPI network, we find that it forms 17 connected components: the largest connected component has 16 genes; the second largest component consists of 5 genes; and all remaining components have 4 or less genes. This means that proteins which are co-amplified with nicastrin are mostly disconnected and scattered around the PPI network. Surprisingly, adding the 6 influential interactors of AMPL-extended set which are present in the Core Diseasome (SHC1, PTPN11,

LYN, ABL1, MDM2 and SP1; see Section 4.2) to the sub-network, results in merging of several of these connected components. Note that influential interactor SHC1 was already present in the sub-network as an amplified gene from AMPL-extended.

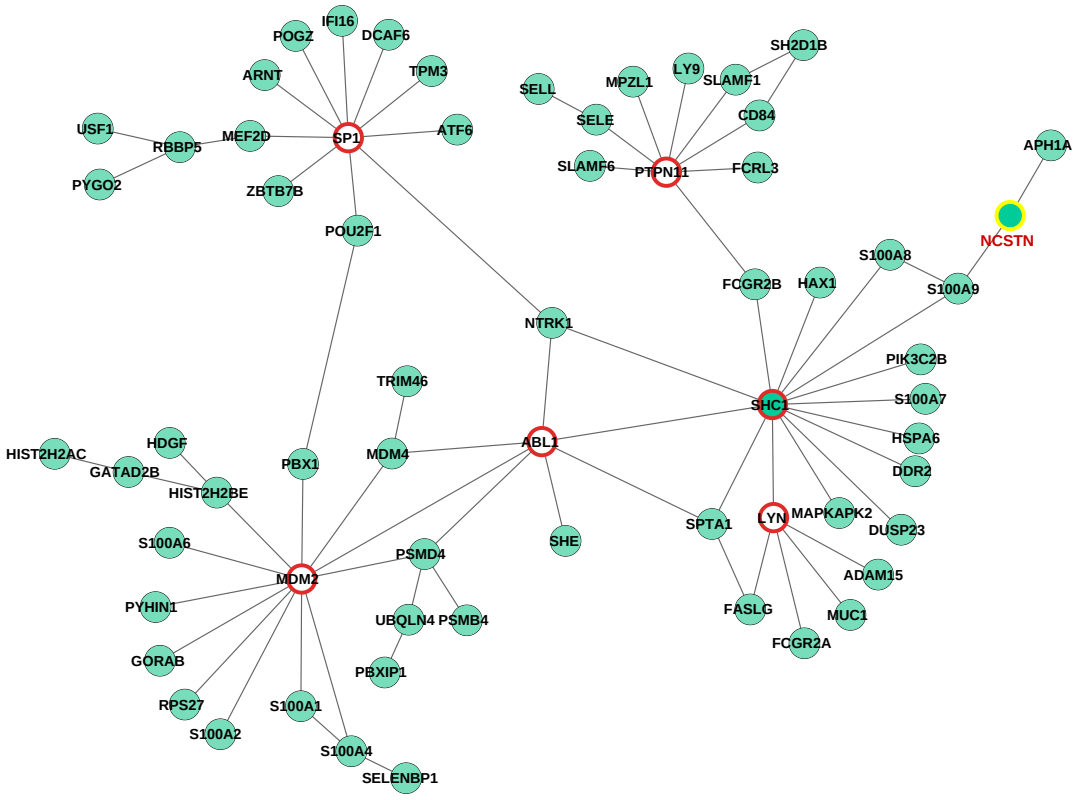

Figure 1. The largest connected component in the human PPI sub-network formed from AMPL-extended set of genes (shown in green) and their influential interactors which are in the Core Diseasome (bordered in red). SHC1 gene is also an amplified gene. Nicastrin (NCSTN) is bordered in yellow.

The largest connected component of the above-described merged sub-network consists of 64 genes (shown in Figure 1). Among these 64 genes, the most highly connected are the 6 added Core Diseasome genes, with SHC1 being the only highly connected one which is simultaneously co-amplified with nicastrin (*i.e.,* present in AMPL-extended). Even more, SHC1 is the only AMPL-extended gene in the entire PPI network which belongs to the Core Diseasome. This leads to a hypothesis that SHC1 might drive the process of amplification in cases when nicastrin is amplified. Hence, we explore pathways which may facilitate the influence of SHC1 on nicastrin amplification.

Nicastrin shares two interacting partners with SHC1 — amplified genes S100A8 and S100A9. Proteins from the S100 family are localized in the cytoplasm and the nucleus of a cell, and play a role in cell cycle progression and differentiation processes. Their coding genes cluster on chromosome 1q21 [48] close to both NCSTN and SHC1. Little has experimentally been confirmed in terms of direct binding partners of

nicastrin *in vivo*. Co-immunoprecipitation experiments show that nicastrin, a member of the γ-secretase complex, interacts with other members of the γ-secretase complex, and proposes transient interactions between nicastrin's substrate-binding DYIGS domain (located on the extracellular portion of the protein) and substrates of the γ-secretase [49]. Our sub-network analysis suggests that proteins S100A8 and S100A9, as interacting partners of nicastrin, might be involved in a pathway that plays a role in the amplification process. The members of the S100 protein family are commonly up-regulated in tumors which is associated with cancer progression [50]. Cellular processes, that these proteins are involved in, range from proliferation, chemoattraction of immunomodulatory cells, control of cellular actin-dynamics and tubular-dynamics, etc. S100A8 and S100A9 are implicated in the metastatic process facilitating the homing of tumor cells to pre-metastatic sites, particularly within the lung parenchyma. They also increase the motility of circulating cancer cells, deemed the culprits of disease relapse.

All this suggests that oncogenic signaling through nicastrin may feed into the pro-metastatic functions of S100A8 and S100A9 in breast cancer. Our hypothesis, therefore, is that attenuating nicastrin signaling will impinge on the S100A8/9 signaling axis, providing a mechanism of inhibiting metastatic engraftment at the secondary site and propagation of cancer cells therein. Given the fact that growth of metastatic disease, still inadequately controlled by current therapies, inevitably results in patients death, identifying targets that control this process is of paramount importance. In support of our hypothesis, performing experiments as described in Section 3.4 , we show that anti-nicastrin monoclonal antibodies used for treatment of a breast cancer metastatic model to the lungs, potently inhibit S100A8/9, as well as their downstream effectors CXCL1 and CXCL2 (Figure 1).

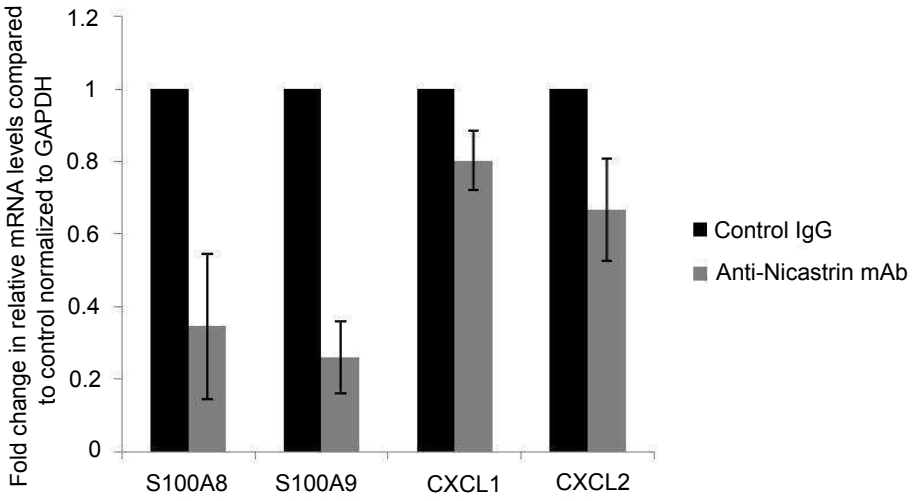

Figure 2. Relative fold change of S100A8, S100A9, CXCL1 and CXCL2 gene expression levels in the excised lung metastasis of animals treated with the rat IgG control and anti-nicastrin monoclonal antibody, in the in vivo metastatic model of MDA-MB-231 cells inoculated through the tail vein of BalbC nude mice. Gene expression levels in n = 4 animals per group were analysed and compared.

Based on the above validation of the role of two interacting partners of SHC1 in breast cancer, the hub properties of all 6 Core Diseasome genes (SHC1, PTPN11, LYN, ABL1, MDM2 and SP1), and the fact that those 6 genes could play a role in activation of the amplification process (discussed in Section 4.2),

we propose to further experimentally explore the involvement of the remaining 5 Core Diseasome genes in breast cancer.

5 Conclusion

Many studies using comparative genomic hybridization have highlighted sites of DNA amplification in cancer that might harbor novel oncogenes. Here, we show that nicastrin is amplified in a subset of breast cancers, and we identify a set of 22 genes co-amplified with nicastrin in at least 95% of cases in a large patient cohort available on cBio Cancer Genomics Portal. We confirm that all 22 genes from the set are located on the chromosome 1, proposing existence of a novel amplicon in breast cancer tissues. Using functional siRNA screening approach, we plan to assess the relative contribution of each of the genes pertaining to this amplicon towards breast cancer cell proliferation and invasion. The relevance of nicastrin in these oncogenic processes has already been validated in multiple *in vitro* and *in vivo* experimental approaches.

We further explore a set of top 5% co-amplified genes and find that statistically significant number of them is found in the plasma membrane. Among them we find genes that are already known to be drug targets. Also, a statistically significant number of amplified genes is involved in biological processes closely related to cancer development and progression.

To study the key proteins involved in breast cancer, we refer to the Core Diseasome — a recently proposed therapeutically relevant set of genes known to trigger disease onset. We find Core Diseasome genes which interact with a statistically significantly high number of genes co-amplified with nicastrin, and show that they have already been linked to breast cancer, indicating that they might play a key role in the onset of this disease. We identify SHC1 as the only gene co-amplified with nicastrin which is present in the Core Diseasome. SHC1 is involved in the highest number of interactions among genes co-amplified with nicastrin, which suggests its high influence in nicastrin co-amplification. We identify a link between nicastrin and SHC1 which involves S100A8 and S100A9 genes, and postulate that S100A8 and S100A9 genes, as interacting partners of nicastrin, may be implicated in a pathway significant for breast invasive carcinoma progression. In support of this hypothesis we experimentally show that anti-nicastrin antibodies inhibit S100A8 and S100A9 genes and their downstream effectors CXCL1 and CXCL2, not only confirming our bioinformatics analysis, but also identifying a downstream signaling mechanism of nicastrin in breast cancer cells.

Acknowledgements

This work was supported by the European Research Council (ERC) Starting Independent Researcher Grant 278212, the National Science Foundation (NSF) Cyber-Enabled Discovery and Innovation (CDI) OIA-1028394, the Serbian Ministry of Education and Science Project III44006, and ARRS project J1-5454.

Disclosures

Disclosures: none.

References

1. American Cancer Society.(2011) Global cancer facts and figures 2nd edition. Atlanta: American Cancer Society

2. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, Graf S, Ha G, Haffari G, Bashashati A, Russell R, McKinney S, Langerod A, Green A, Provenzano E, Wishart G, Pinder S, Watson P, Markowetz F, Murphy L, Ellis I, Purushotham A, Borresen-Dale AL, Brenton JD, Tavare S, Caldas C, Aparicio S (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 486(7403):346–352

3. Lehmann B, Bauer J, Chen X, Sanders M, Chakravarthy A, Shyr Y, Pietenpol J (2011) Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. The Journal of Clinical Investigation 121(7):2750–2767

4. Sahlberg KK, Hongisto V, Edgren H, Makela R, Hellstrom K, Due EU, Vollan HKM, Sahlberg N, Wolf M, Brresen-Dale AL, Perl M, Kallioniemi O (2013) The {HER2} amplicon includes several genes required for the growth and survival of {HER2} positive breast cancer cells. Molecular Oncology 7(3):392 – 401

5. Orsetti B, Nugoli M, Cervera N, Lasorsa L, Chuchana P, Roug C, Ursule L, Nguyen C, Bibeau F, Rodriguez C, Theillet C (2006) Genetic profiling of chromosome 1 in breast cancer: mapping of regions of gains and losses and identification of candidate genes on 1q. British Journal of Cancer 95(10):1439–1447

6. Filipović A, Gronau J, Green A, Wang J, Vallath S, Shao D, Rasul S, Ellis I, Yage E, Sturge J, Coombes R (2011) Biological and clinical implications of nicastrin expression in invasive breast cancer. Breast Cancer Research and Treatment 125(1):43–53

7. Filipović A, Lombardo Y, Fronato M, Abrahams J, Aboagy E, Nguyen Q, Borda d' Aqua B, Ridley A, Green A, Rahka E, Ellis I, Recchi C, Pržulj N, Sarajlić A, Alattia J, Fraering P, Mahendra D, Coombes R (2013) Monoclonal antibody targeting of nicastrin inhibits growth and metastasis of invasive breast cancer cells. Submitted

8. Lombardo Y, Filipovic A, Molyneux G, Periyasamy M, Giamas G, Hu Y, Trivedi PS, Wang J, Yague E, Michel L, Coombes RC (2012) Nicastrin regulates breast cancer stem cell properties and tumor growth in vitro and in vivo. Proceedings of the National Academy of Science 109:16,558–16,563

9. Cerami E, Gao J, Dogrusoz U (2012) The cbio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. Cancer Discovery (2):401–404

10. Sharan R, Ulitsky I, Shamir R (2007) Network-based prediction of protein function. Molecular Systems Biology 3

11. Schwikowski B, Uetz P (2000) A network of protein-protein interactions in yeast. Nature Biotechnology 18:1257–1261

12. Chua HN, Sung WK, Wong L (2006) Exploiting indirect neighbours and topological weight to predict protein function from protein-protein interactions. Bioinformatics 22(13):1623–1630

13. Samanta MP, Liang S (2003) Predicting protein functions from redundancies in large-scale protein interaction networks. Proceedings of The National Academy of Sciences USA 100:12,579–12,583

14. Ideker T, Sharan R (2008) Protein networks in disease. Genome Research 18:644–652

15. Aragues R, Sander C, Oliva B (2008) Predicting cancer involvement of genes from heterogeneous data. BMC Bioinformatics 9(172):172

16. Goldenberg A, Mostafavi S, Quon G, Boutros PC, Morris QD (2011) Unsupervised detection of genes of influence in lung cancer using biological networks. Bioinformatics 27(22):3166–3172

17. Janjić V, Pržulj N (2012) The core diseasome. Molecular Biosystems 8(10):2614–2625

18. Ashworth A, Lord CJ, Reis-Filho JS (2011) Genetic interactions in cancer progression and treatment. Cell 145(1):30–38

19. Chatr-aryamontri A, Breitkreutz BJ, Heinicke S, Boucher L, Winter A, Stark C, Nixon J, Ramage L, Kolas N, O'Donnell L, Reguly T, Breitkreutz A, Sellam A, Chen D, Chang C, Rust J, Livstone M, Oughtred R, Dolinski K, Tyers M (2013) The biogrid interaction database. Nucleic Acids Research 41(D1):D816–D823

20. Collaboration (2012) Comprehensive molecular portraits of human breast tumours. Nature 490(7418):61–70

21. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological) 57(1):pp. 289–300

22. Basik M, Caplen NJ, Kallioniemi OP, Mousses S (2003) Amplified genes as therapeutic targets in cancer. {TARGETS} 2(4):147 – 153

23. Bakheet TM, Doig AJ (2008) Properties and identification of human protein drug targets. Bioinformatics 25(4):451–457

24. Thompson L, Jiang J, Madamanchi N, Runge M, Patterson C (2001) Ptp-epsilon, a tyrosine phosphatase expressed in endothelium, negatively regulates endothelial cell proliferation. American journal of physiology Heart and circulatory physiology 281(1):H396–403

25. Carr B, Wang Z, Kar S (2002) K vitamins, ptp antagonism, and cell growth arrest. Journal of Cellular Physiology 193(3)

26. Kikawa K, Vidale D, Van Etten R, Kinch M (2002) Regulation of the epha2 kinase by the low molecular weight tyrosine phosphatase induces transformation. The journal of biological chemistry 277(42):39,274–9

27. Ha TY (2009) The role of regulatory t cells in cancer. Immune Netw 9:209–235

28. Balkwill F, Montfort A, Capasso M (2013) B regulatory cells in cancer. Trends in Immunology 34(4):169 – 173

29. Zamai L, Ponti C, Mirandola P, Gobbi G, Papa S, Galeotti L, Cocco L, Vitale M (2007) Nk cells and cancer. The Journal of Immunology 178(7):4011–6

30. Ray M, Ruan J, Zhang W (2008) Variations in the transcriptome of alzheimer's disease reveal molecular networks involved in cardiovascular diseases. Genome Biology 9(10):R148

31. Aihara T, Miyoshi Y, Koyama K, Suzuki M, Takahashi E, Monden M, Nakamura Y (1998) Cloning and mapping of smarca5 encoding hsnf2h, a novel human homologue of drosophila iswi. Cytogenetic and Genome Research 81(3-4):191–193

32. Champoux JJ (2001) Dna topoisomerases: structure, function, and mechanism. Annual review of biochemistry 70(1):369–413

33. Cenciarelli C, Chiaur D, Guardavaccaro D, Parks W, Vidal M, Pagano M (1999) Identification of a family of human f-box proteins. Current Biology 9(20):1177 – S3

34. Patton E, Willems AR, Tyers M (1998) Combinatorial control in ubiquitin-dependent proteolysis: don't skp the f-box hypothesis. Trends in genetics 14(6):236 – 243

35. Clarke MF, Liu R (2013) Compositions and methods for treating and diagnosing cancer

36. Ouimet M, Cassart P, Larivire M, Kritikou E, Simard J, Sinnett D (2012) Functional analysis of promoter variants in ku70 and their role in cancer susceptibility. Genes Chromosomes Cancer 51(11):1007–1

37. Chu IM, Hengst L, Slingerland JM (2008) The cdk inhibitor p27 in human cancer: prognostic potential and relevance to anticancer therapy. Nat Rev Cancer 8(4):253–267

38. Abdelmohsen K, Gorospe M (2010) Posttranscriptional regulation of cancer traits by hur. Wiley Interdisciplinary Reviews - RNA 1(2):214–229

39. Kazarian M, Laird-Offringa IA (2011) Small-cell lung cancer-associated autoantibodies: potential applications to cancer diagnosis, early detection, and therapy. Molecular Cancer 10(1):33

40. Tangye SG, Phillips JH, Lanier LL (2000) The cd2-subset of the ig superfamily of cell surface molecules: receptor-ligand pairs expressed by nk cells and other immune cells. Seminars in Immunology 12(2):149–57

41. Wagner K, Hemminki K, Grzybowska E, Klaes R, Butkiewicz D, Pamula J, Pekala W, Zientek H, Mielzynska D, Siwinska E, Försti A (2004) The insulin-like growth factor-1 pathway mediator genes: Shc1 met300val shows a protective effect in breast cancer. Carcinogenesis 25(12):2473–2478

42. Wang X (2012) Abstract 2150: The role of shp2 in her2+ breast cancer. Proceedings of the 103rd Annual Meeting of the American Association for Cancer Research 72(8):1538–7445

43. Srinivasan D, Sims JT, Plattner R (2008) Aggressive breast cancer cells are dependent on activated abl kinases for proliferation, anchorage-independent growth and survival. Oncogene 27(8):1095–1105

44. Mieyal J, Gallogly M, Qanungo S, Sabens E, Shelton M (2008) Molecular mechanisms and clinical implications of reversible protein s-glutathionylation. Antioxid Redox Signal 10(11):1941–88

45. McCann AH, Kirley A, Carney DN, Corbally N, Magee HM, Keating G, Dervan PA (1995) Amplification of the mdm2 gene in human breast cancer and its association with mdm2 and p53 protein status. British Journal of Cancer 71(5):981–985

46. Hibbs ML, Harder KW (2006) The duplicitous nature of the Lyn tyrosine kinase in growth factor signaling. Growth Factors 24(2):137–149

47. Choi YL, Bocanegra M, Kwon MJ, Shin YK, Nam SJ, Yang JH, Kao J, Godwin AK, Pollack JR (2010) Lyn is a mediator of epithelial-mesenchymal transition and a target of dasatinib in breast cancer. Cancer Research 70(6):2296–2306

48. Donato R, Cannon B, Sorci G, Riuzzi F, Hsu K, Weber D, CL G (2013) Functions of s100 proteins. Antioxid Redox Signal 13(1):24–57

49. Hayashi I, Takatori S, Urano Y, Miyake Y, Takagi J, Sakata-Yanagimoto M, Iwanari H, Osawa S, Morohashi Y, Li T, Wong PC, Chiba S, Kodama T, Hamakubo T, Tomita T, Iwatsubo T (2012) Neutralization of the gamma-secretase activity by monoclonal antibody against extracellular domain of nicastrin. Oncogene 31:787–798

50. Salama I, Malone P, Mihaimeed F, Jones J (2008) A review of the {S100} proteins in cancer. European Journal of Surgical Oncology (EJSO) 34(4):357 – 364