Free Radical Biology and Medicine 184 (2022) 170–184

Contents lists available at ScienceDirect

Free Radical Biology and Medicine

journal homepage: www.elsevier.com/locate/freeradbiomed

NRF2 drives an oxidative stress response predictive of breast cancer

Camilla Wolowczyk [a,b,c,*], Ulrike Neckmann [a,b,c], Miriam Ragle Aure [g,l], Martina Hall [h,i], Bjarne Johannessen [d,e], Sen Zhao [d,e], Rolf I. Skotheim [d,e], Sonja B. Andersen [a,b], Rosalie Zwiggelaar [a], Tonje S. Steigedal [f], Ole Christian Lingjærde [j], Kristine Kleivi Sahlberg [g,k], Eivind Almaas [h,i], Geir Bjørkøy [a,b]

[a] Centre of Molecular Inflammation Research and Department of Cancer Research and Molecular Medicine, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway
[b] Department of Biomedical Laboratory Science, Faculty of Natural Sciences, Norwegian University of Science and Technology, Trondheim, Norway
[c] Clinic of Laboratory Medicine, St. Olavs Hospital, Trondheim, Norway
[d] Department of Molecular Oncology, Institute for Cancer Research, Oslo University Hospital-Radiumhospitalet, Oslo, Norway
[e] Norwegian Cancer Genomics Consortium, Oslo, Norway
[f] Department of Clinical and Molecular Medicine, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
[g] Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital, The Norwegian Radium Hospital, Oslo, Norway
[h] Department of Biotechnology and Food Science, Faculty of Natural Sciences, Norwegian University of Science and Technology, Trondheim, Norway
[i] K.G.Jebsen Center for Genetic Epidemiology, Department of Public Health and General Practice, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Norway
[j] Department of Computer Science, University of Oslo, Oslo, Norway
[k] Department of Research, Vestre Viken Hospital Trust, Drammen, Norway
[l] Department of Medical Genetics, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

ARTICLE INFO

Keywords:
4T1 model
66cl4
67NR
NFE2L2
SQSTM1
NQO1
Prognosis

ABSTRACT

Many breast cancer patients are diagnosed with small, well-differentiated, hormone receptor-positive tumors. Risk of relapse is not easily identified in these patients, resulting in overtreatment. To identify metastasis-related gene expression patterns, we compared the transcriptomes of the non-metastatic 67NR and metastatic 66cl4 cell lines from the murine 4T1 mammary tumor model. The transcription factor nuclear factor, erythroid 2-like 2 (NRF2, encoded by *NFE2L2*) was constitutively activated in the metastatic cells and tumors, and correspondingly a subset of established NRF2-regulated genes was also upregulated. Depletion of NRF2 increased basal levels of reactive oxygen species (ROS) and severely reduced ability to form primary tumors and lung metastases. Consistently, a set of NRF2-controlled genes was elevated in breast cancer biopsies. Sixteen of these were combined into a gene expression signature that significantly improves the PAM50 ROR score, and is an independent, strong predictor of prognosis, even in hormone receptor-positive tumors.

1. Introduction

Metastasis is the primary limitation for breast cancer survival and can be detected several years after removal of the primary tumor. The possibilities to prevent or treat metastatic cancer are still limited, thus emphasizing the importance of increasing the knowledge of cancer cells' ability to metastasize. For cells of epithelial origin, metastases result from a series of acquired changes, including the ability to leave the primary tumor, enter into and survive in circulation, extravasate and establish at a secondary site. The acquired metastatic phenotype is driven by genetic changes that result in altered gene expression and protein function. The nature of the acquired changes in gene expression can be analyzed in publicly available data sets, where expression of individual genes or set of genes can be correlated with prognosis. However, these analyses are complicated by inter- and intratumor heterogeneity, different cellular composition of the biopsies and the fact that tumor cells interact with non-transformed cells, including immune cells.

To study the correlation between altered gene expression and

* Corresponding author. Centre of Molecular Inflammation Research and Department of Cancer Research and Molecular Medicine, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway.
E-mail address: camilla.wolowczyk@ntnu.no (C. Wolowczyk).

https://doi.org/10.1016/j.freeradbiomed.2022.03.029
Received 27 January 2022; Received in revised form 29 March 2022; Accepted 30 March 2022
Available online 2 April 2022

Abbreviations

ChEA	ChiP-X enrichment analysis
KEAP1	Kelch-like ECH-associated protein
NRF2	Nuclear factor, erythroid 2-like 2
PCA	principal component analysis
RFS	relapse-free survival
ROR	risk of recurrence
ROS	reactive oxygen species
wTO	weighted Topological Overlap.

metastatic phenotype, animal models reflecting this complexity have been established. The murine 4T1 mammary tumor model is an immunocompetent animal model for breast cancer; consisting of five different cell lines isolated from the same spontaneous BALB/c mouse tumor [1]. The five cell lines differ clearly in their metastatic potential, although they behave similarly in culture and all form primary tumors in the mammary fat pad of syngeneic BALB/c mice [2]. It is also known that cancer cells of epithelial origin during epithelial-to-mesenchymal transition become more stem cell-like and acquire the ability to invade other tissues [3], which is also evident in 66cl4 cells through the increased expression of several stem cell markers [4].

During metastasis, cancer cells experience high levels of oxidative stress, and cancer cells that are able to cope with this challenge can form distant metastasis [5]. Oxidative stress is counteracted by a number of genes regulated by the transcription factor nuclear factor, erythroid-derived 2-like 2 (NRF2) encoded by the *NFE2L2* gene. Somatic mutations causing constitutive activation of NRF2 are frequent in squamous-like and lung cancer subtypes [6,7]. Under normal conditions, NRF2 is ubiquitinated after direct binding to the Kelch-like ECH-associated protein 1 (KEAP1) as a part of the cullin 3 RING ubiquitin ligase complex, and rapidly degraded by the proteasome [8–11]. Upon oxidative stress, NRF2 dissociates from KEAP1, stabilizes, and translocates to the nucleus where it controls the expression of numerous genes encoding enzymatic and non-enzymatic antioxidants or proteins that relieve the damage caused by the insult [12]. The pathway is normally induced in response to raised levels of oxygen radicals, and elevated NRF2 can result from metabolic changes in transformed cells even in the absence of genetic alterations directly affecting NRF2 [13]. Mutual exclusive, somatic mutations in *NFE2L2*, *KEAP1*, or *CUL3*, which activate the NRF2 pathway, have been firmly established [6,7]. These genetic changes also include deletions of exon two of *NFE2L2* [14]. Moreover, the level and activity of NRF2 can be increased by several indirect mechanisms, including SQSTM1-mediated degradation of the negative regulator KEAP1 [15]. Aside from that, SQSTM1 expression is controlled by NRF2, and its induction can set up a positive feedback loop of the pathway [16].

We hypothesized that although somatic mutations in *NFE2L2*, *KEAP1*, and *CUL3* are rarely detected in breast cancer, activation of NRF2 signaling can be important also in this cancer type, as it is a fundamental cytoprotective mechanism and may contribute to the development of an aggressive phenotype. In this study, we compared the non-metastatic 67NR cell line to the metastatic 66cl4 cell line of the 4T1 model. We report that NRF2 is constitutively elevated in the metastatic 66cl4 cells, and that it results in activation of a selective set of NRF2 controlled genes. Depletion of NRF2 abolished 66cl4's ability to form primary tumors and metastases to the lungs. Consistent with a role of NRF2 in metastasis of human breast cancer, the expression of a discrete subset of NRF2 related genes was elevated in breast cancer biopsies, and increased expression correlated with poor prognosis. Combining a specific subset of 16 NRF2-controlled transcripts resulted in a breast cancer-specific gene expression signature that correlated with risk of relapse for breast cancer patients.

2. Results

2.1. Constitutive NRF2 in cultures and tumors formed by the aggressive 66cl4 breast cancer cells

Transcriptome sequencing was conducted on RNA isolated from 67NR and 66cl4 cells grown in culture and as primary tumors as well as macroscopically visible lung metastases of 66cl4-bearing BALB/cJ mice. Principal component analysis (PCA) showed a clear segregation of cell culture versus primary tumor samples and 67NR versus 66cl4 (Fig. 1a). The PCA plot also indicated that 66cl4 lung metastases most closely resembled 66cl4 primary tumors. There was a clear correlation between the relative differences of the cells grown in culture and as primary tumors (Fig. 1b).

These data indicate that the two cell lines originating from the same primary tumor harbor stable gene expression differences, which are present in both culture and tumor.

Of the 23 994 transcripts identified in the sequencing, 1270 genes showed a significantly elevated expression in 66cl4 compared to 67NR cells and primary tumors ($\log2(ExpVal\_66cl4/ExpVal\_67NR > 0.59$; $ExpVal\_66cl4 \geq 1$; p-value ≤ 0.05). On the other side, 1252 genes were expressed higher in 67NR cells and tumors. Enrichment analysis of the genes upregulated in 66cl4 cells and tumors was done using Enrichr [17,18]. ChiP-X enrichment analysis (ChEA) suggested NRF2 (*Nfe2l2)* controlled gene expression as clearly elevated in 66cl4 (Fig. 1c). Moreover, analysis using the KEGG and Reactome databases indicated that cell metabolism was altered in these cells (Fig. 1d, Supplementary Fig. 1). Since NRF2 is known to control transcription of various enzymes in diverse metabolic pathways [19] and hyper-activation of this transcription factor has been linked to tumorigenesis and progression in various cancer types [20], we decided to focus on this pathway.

To discriminate if activation of the oxidative stress response was due to direct genetic alterations, the exomes of 66cl4 and 67NR were sequenced. DNA isolated from BALB/cJ mouse was used as a control. When compared to the mm9 mouse reference genome (C57BL/6J), the majority of gene variants were common between BALB/cJ, 67NR and 66cl4. Of the remaining variants, 18 were common for the two cell lines, whereas 640 and 834 mutations were exclusively found in 67NR and 66cl4, respectively (Fig. 1e). No alterations were found in *Nfe2l2* or *Keap1,* however, 66cl4 harbored a C to G transversion in exon 14 of *Cul3* (c.C1768G; p.H590D). CUL3 is part of the KEAP1-CUL3 E3 ligase complex, which is needed to polyubiquitylate NRF2 and target it for degradation under basal conditions [9,11]. Sanger sequencing showed that 66cl4 is the only among the five cell lines of the 4T1 model that harbors this mutation (Fig. 1f). Consistent with a functional role of this mutation, NRF2 protein levels were clearly elevated whereas the transcript level was in fact lower in 66cl4 compared to 67NR (Fig. 2a and b, Supplementary Fig. 2). However, variable correlation between transcript and the corresponding proteins is known and underlines the fact that NRF2 is more complexly regulated than solely on transcriptional level [21]. In line with constitutive NRF2 signaling, the classical targets NQO1 and HMOX1 were increased at both mRNA and protein level in 66cl4 (Fig. 2a and b, Supplementary Fig. 2). Collectively, these data demonstrate a constitutive activation of the NRF2 pathway in the metastatic 66cl4.

2.2. Constitutive NRF2 activity reduces basal ROS levels, but abrogates responses to additional oxidative stress

In order to sustain aberrant proliferation, cancer cells produce high amounts of ATP, which results in increased ROS and imbalanced redox status [13]. Consistent with activated NRF2 in 66cl4, we found that basal ROS levels were reduced but that these cells were also clearly less able to mobilize endogenous antioxidants in response to the oxidative stress inducing agent hydrogen peroxide and NRF2-inducing factor L-sulforaphane (Fig. 2c – f, and quantification and statistics in

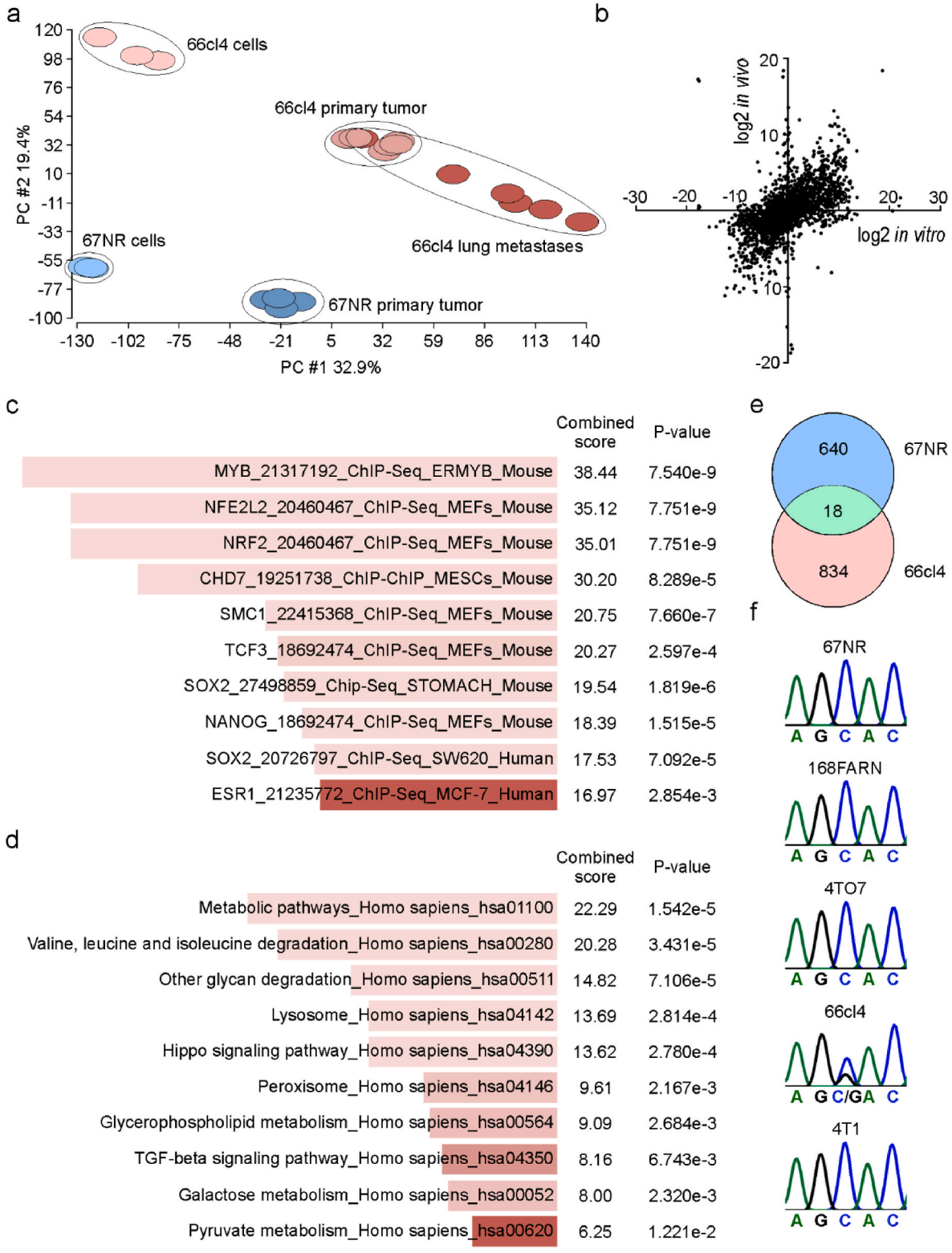
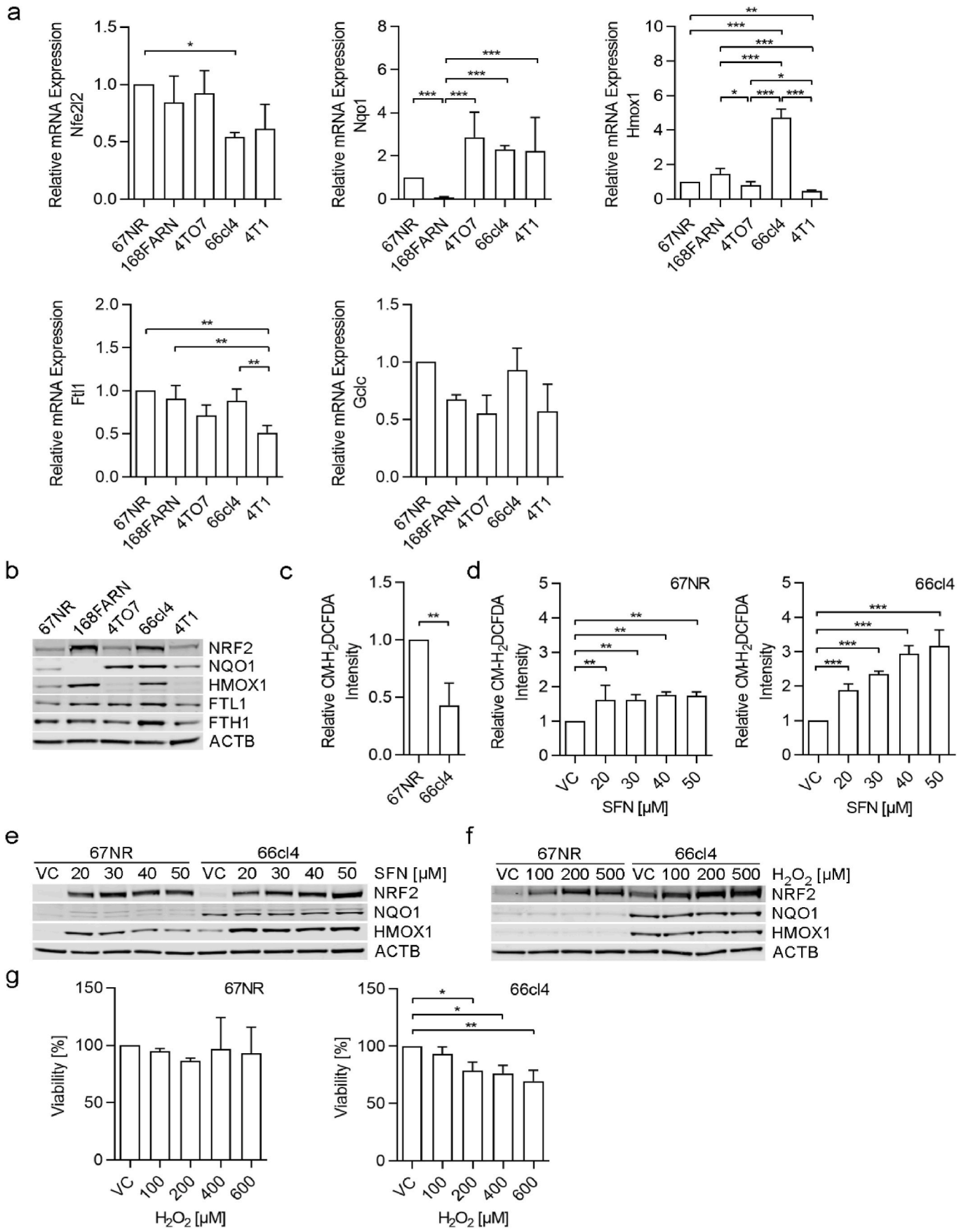


Fig. 1. NRF2 is constitutively activated in 66cl4. (a) Principal components Analysis of 67NR cells and primary tumors and 66cl4 cells, primary tumors and lung metastases. (b) Scatter plot of in vivo versus in vitro differential expression between 66cl4 and 67NR. A positive number indicates higher expression in 66cl4, whereas a negative number indicates higher expression in 67NR. (c, d) Enrichment analysis of genes (1,270) upregulated in 66cl4 cells and primary tumors compared to 67NR cells and primary tumors (log2(ExpVal_66cl4/ExpVal_67NR $>$ 0.59; ExpVal_66cl4 $\geq$ 1; p-value $\leq$ 0.05). Analysis was performed with Enrichr database. Entries were sorted by combined score. (c) ChiP-X enrichment analysis. (d) KEGG cell signaling pathway analysis. (e) Venn-diagram of exonic nonsynonymous mutations detected by exome sequencing in 67NR and 66cl4. (f) The Cul3 DNA sequence in exon 14 from the indicated 4T1 model cell lines determined by Sanger sequencing.



(caption on next page)

Fig. 2. 66cl4 displayed reduced basal ROS levels and impaired ability to adopt to additional oxidative stress. (a) RT-qPCR analysis. Fold changes are relative to Actb and Tbp (n = 3). Data are shown as mean ± SD (ANOVA after log transformation, Tukey's multiple comparisons test). (b) Representative immunoblot of NRF2 (100 μg protein loaded) and NRF2-target genes NQO1, HMOX1, FTL1, and FTH1 (50 μg protein loaded). ACTB (2 μg protein loaded) was used as a loading control (n = 3). (c) Flow-cytometric analysis of basal ROS levels after CM-H2DCFDA staining (n = 4). Data are shown as mean ± SD (one sample *t*-test after log transformation). (d) ROS levels after 4 h' treatment with L-Sulforaphane (SFN) or vehicle (DMSO) (n = 3). Data are shown as mean ± SD (ANOVA after log transformation, Dunnett's multiple comparisons test). (e, f) Representative immunoblots of NRF2, NQO1, and HMOX1 (50 μg protein loaded). ACTB (2 μg protein loaded) was used as a loading control. Cells were treated with (e) L-sulforaphane or vehicle (DMSO) for 4 h and (f) hydrogen peroxide for 30 min. Note that the exposure in (e) is lower due to the great induction of NRF2 expression. (g) MTT analysis of 67NR and 66cl4 cell viability after treatment with hydrogen peroxide (n = 3). Data are shown as mean ± SD (ANOVA after log transformation, Dunnett's multiple comparisons test). VC, vehicle control.

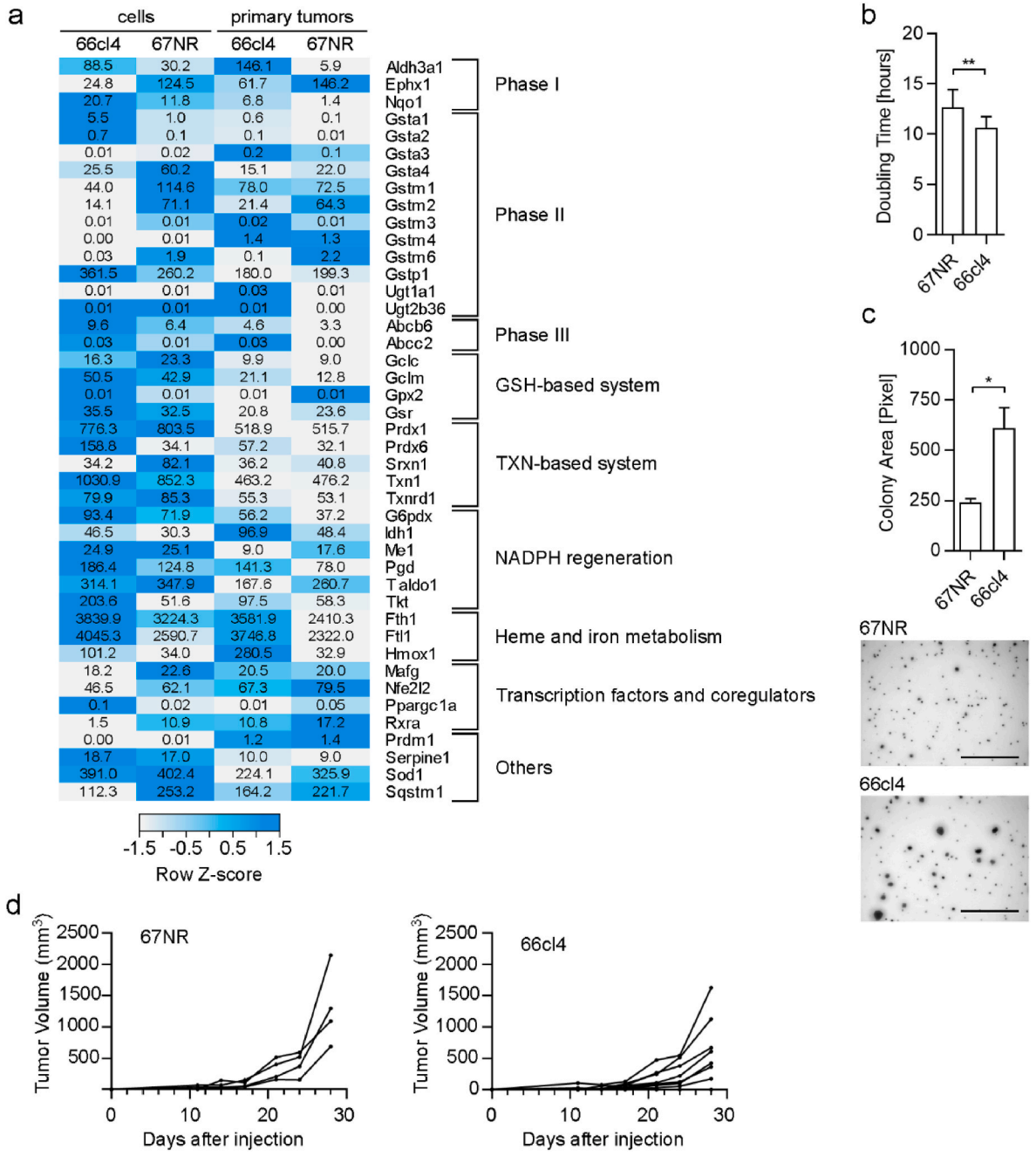


Fig. 3. A subset of NRF2 genes is elevated in the metastatic 66cl4 cells but the primary tumor growth is not different. (a) Heat-map of the expression patterns of 43 NRF2-target genes from expression sequencing data of 67NR and 66cl4 cells and primary tumors. (b) Doubling time of 67NR and 66cl4 in culture (n = 5). Data are shown as mean ± SD (Student's t-test after log transformation). (c) Soft-agar assay. Colony area was measured in pixels. Z-stack of representative image of colonies. Scale bar: 1 mm (n = 3). Data are shown as mean ± SEM (Student's t-test after log transformation). (d) Growth curve of primary tumors after injection of 67NR (4 mice) and 66cl4 (8 mice) into the fat-pad of BALB/cJ mice.

Supplementary Fig. 3). In line with the reduced ability to adopt to additional oxidative stress, hydrogen peroxide supplementation led to lower viability in 66cl4 cells as measured by reduced metabolic activity (Fig. 2g).

2.3. Constitutive NRF2 activation coincides with metabolic reprogramming

The transcriptome analysis revealed that a subset of well-known NRF2 target genes was upregulated in 66cl4 (Fig. 3a). We then asked if constitutive NRF2 activation correlated with growth rate in culture, the ability to form colonies in soft agar or growth rate of the primary tumor. Compared to 67NR cells, the 66cl4 cells multiplied marginally faster in culture and formed larger colonies in soft agar (Fig. 3b and c). However, there was no significant difference in the growth of the primary tumors formed by the two cell lines. In fact, tumors formed by 67NR cells displayed a tendency to grow faster than 66cl4 tumors (Fig. 3d). NRF2 regulates genes involved in glucose and glutamine metabolism [19] and pathway enrichment analysis revealed alterations in cell metabolism in 66cl4 (Fig. 1d and Supplementary Fig. 1). In addition, GO cellular component analysis showed that mitochondrial components were upregulated in 66cl4 (Supplementary Fig. 4a). Analysis of cellular bioenergetics using a Seahorse XF Analyzer demonstrated that 66cl4 had a higher glycolytic flux (Supplementary Figs. 4b–e). The glycolytic capacity did not differ, and the glycolytic reserve was largest in 67NR. Moreover, the basal mitochondrial respiration and ATP production were similar. The maximum respiration was highest in 67NR and accordingly, the spare respiratory capacity was significantly higher in these cells. Together, these results indicate that upregulation of mitochondrial components in 66cl4 is a compensatory mechanism that does not result in increased mitochondrial activity.

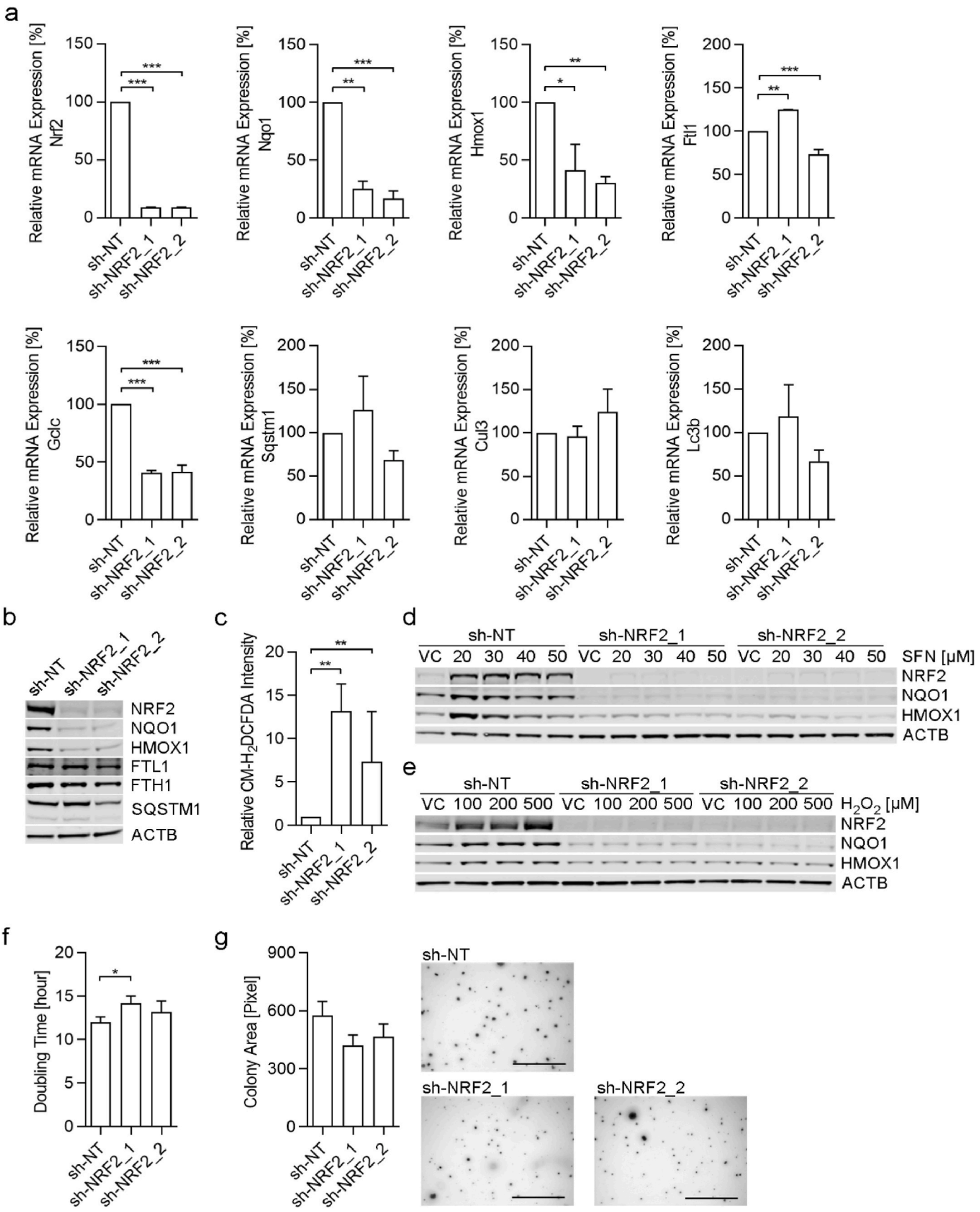
2.4. NRF2 depletion impairs primary tumor growth and metastasis

To evaluate the role of NRF2 in the metastasis model, the protein was depleted by shRNA-mediated knockdown in 66cl4. Knockdown efficiency was assessed both at transcript and protein level and was significantly reduced by approximately 90% in two independent clones (Fig. 4a and b). Depletion of NRF2 coincided with up to 90% reduction in mRNA levels of *Nqo1*, *Hmox1* and *Gclc,* and protein levels of NQO1 and HMOX1. However, mRNA and protein expression of FTH1, FTL1 and SQSTM1 were not changed by NRF2 downregulation. As NRF2 independent controls, mRNA expression of *Cul3* and *Lc3b* were unaffected. As expected, NRF2 depletion led to a clear increase in basal ROS and loss of NRF2, NQO1, and HMOX1 induction in response to oxidative stress (Fig. 4c–e). We also observed a slightly reduced growth rate and reduction in the ability to form colonies in soft agar (Fig. 4f and g). Moreover, NRF2 depleted clones displayed reduced efficiency of glycolysis and significantly increased mitochondrial respiration (Supplementary Fig. 5). Together, these data indicate that interfering with the constitutive NRF2 expression restored energy production via oxidative phosphorylation. To evaluate the role of constitutive NRF2 for the tumor formation, both control cells and NRF2 depleted cells were injected into the fat pad of female BALB/c mice. Strikingly, down regulation of NRF2 resulted in a clear loss in the ability to form primary tumors (Fig. 5a and b, and Supplementary Fig. 6). The clear difference in ability to form primary tumors precluded a direct comparison of metastatic ability. Instead, control cells and NRF2 depleted cells were injected into the tail vein of female BALB/c mice and lung metastasis measured as total lung weight after 19 days (Fig. 5c). Injection of the control cells resulted in swollen lungs with numerous metastatic foci. In contrast, the lungs from the mice injected with the NRF2 depleted cells had normal size and appearance. Together, these results suggest that constitutively active NRF2 is functionally important for an aggressive phenotype of breast cancer cells.

2.5. The NRF2 target gene signature shows prognostic value in ER negative breast tumors

Somatic mutations in *NFE2L2*, *KEAP1*, and *CUL3*, resulting in activation of NRF2 signaling, have been described in various cancer types, particularly in squamous-like cancers [6,7]. Searching the TCGA Pan-Cancer Atlas in the online tool cBioPortal [22,23] for gene alterations in *NFE2L2*, *KEAP1* and *CUL3*, we saw that alterations in these three genes were found in 33% of samples from lung squamous cell carcinoma patients, 17% of samples from head and neck squamous cell carcinoma and only 4% of samples from breast invasive carcinoma patients (Fig. 6a). Although these specific genetic alterations are rare in breast cancer, the pathway may also be activated by several indirect mechanisms, including competitive binding of SQSTM1, DPP3, PALB2, or AMER1 to KEAP1 or direct binding of CDKN1A, which all lead to stabilization of the NRF2 protein [15,24–27]. Since metastases are the major cause of breast cancer-related deaths, we hypothesized that if any particular gene controlled by NRF2 should be important for the ability to metastasize, its mRNA expression level should predict poor prognosis in breast cancer patients. We selected 46 established NRF2-target genes (Supplementary Table 1) and analyzed whether high mRNA expression of any of these correlates with poor prognosis in breast cancer patients using KM plotter [28,29] and BreastMark [30]. These analyses revealed that high mRNA expression of 10 out of the 46 NRF2 target genes, correlated with reduced relapse free survival (RFS) in both databases (Supplementary Table 2). Furthermore, by using the Ualcan cohort database [31] we found that 17 NRF2-target genes were higher expressed in breast tumor tissue than normal tissue (Supplementary Fig. 7). Accordingly, of the 46 NRF2-target genes, only six correlated with poor prognosis and were also upregulated only in tumor compared to normal tissue (*NQO1*, *SERPINE1*, *SRXN1*, *TALDO1*, *TXN*, and *TXNRD1*) (Supplementary Fig. 8a – f)). By combining these six genes into a gene signature, the prognostic value increased substantially compared to each transcript alone (Fig. 6b). Since NRF2 levels can also be elevated by other proteins, we hypothesized that mRNA expression of such proteins may improve the prognostic value of this NRF2 gene signature. Indeed, including the transcripts for *SQSTM1*, *PALB2*, and *AMER1* to the NRF2 gene signature resulted in a nine-gene signature with slightly increased hazard ratio, and reduced p-value (Fig. 6c). Of note, the same NRF2 gene expression signature displayed no prognostic value in gastric or lung cancer patients (Supplementary Figs. 8g and h). Together, these findings demonstrate that NRF2-gene signatures are predictive in a cancer type specific manner and that the importance of individual NRF2 target genes for breast cancer aggressiveness differ from important to neutral.

Using a network systems biology approach, we identified genes with expression patterns most strongly correlated with the six initial NRF2 target genes of the first gene expression signature in breast cancer tissue from TCGA, by using the weighted Topological Overlap (wTO) approach [32]. Considering the ego-centric networks of each target gene with the 25 strongest associated links, we found that 32 genes were connected to more than one target gene (Table 1). Comparing the ranking of these genes to the corresponding ego-centric network for the controls, we found that all except two of these genes were specific to patients with breast cancer (Supplementary Table 3). Subsequently, these 30 specific transcripts were individually tested for association between elevated level and poor prognosis using both the KM plotter and BreastMark web tools. Based on this analysis, we identified seven transcripts that directly correlated with the NRF2 signature (Supplementary Fig. 9). Interestingly, these seven transcripts connect to the same three target genes, TXN, TXNRD1 and TALDO1 (Table 1). Elevated levels of the transcripts correlated with poor prognosis and adding them to the existing nine-gene signature resulted in a 16-transcript signature with even stronger prognostic value (Fig. 6d).



(caption on next page)

Fig. 4. NRF2 depletion in 66cl4 slows proliferation and limit anchorage-independent growth. (a) RT-qPCR analysis. Fold changes are relative to Actb and Tbp (n = 3). Data are shown as mean ± SD (ANOVA after log transformation, Dunnett's multiple comparisons test). (b) Representative immunoblot of NRF2 (100 μg protein loaded), NQO1, HMOX1, FTL1, FTH1 and SQSTM1 (50 μg protein loaded). ACTB was used as a loading control (2 μg protein loaded). (c) Flow-cytometric analysis of basal ROS levels after CM-H2DCFDA staining (n = 3). Data are shown as mean ± SD (ANOVA after log transformation, Dunnett's multiple comparisons test). (d, e) Representative immunoblots of NRF2, NQO1, and HMOX1 (50 μg protein loaded). ACTB (2 μg protein loaded) was used as a loading control. Cells were treated with (d) L-sulforaphane (SFN) (20 μM – 50 μM) for 4 h or vehicle (DMSO) and (e) hydrogen peroxide (H2O2) (100 μM – 500 μM) for 30 min. (f) Doubling time of 66cl4 sh-NT control and sh-Nrf2 knockdowns in culture (n = 5). Data are shown as mean ± SD (ANOVA after log transformation, Dunnett's multiple comparisons test). (g) Soft-agar assay. Colony area was measured in pixels. Z-stack of representative image of colonies. Scale bar: 1 mm (n = 5), Data are shown as mean ± SEM (ANOVA after log transformation, Dunnett's multiple comparisons test). VC, vehicle control.

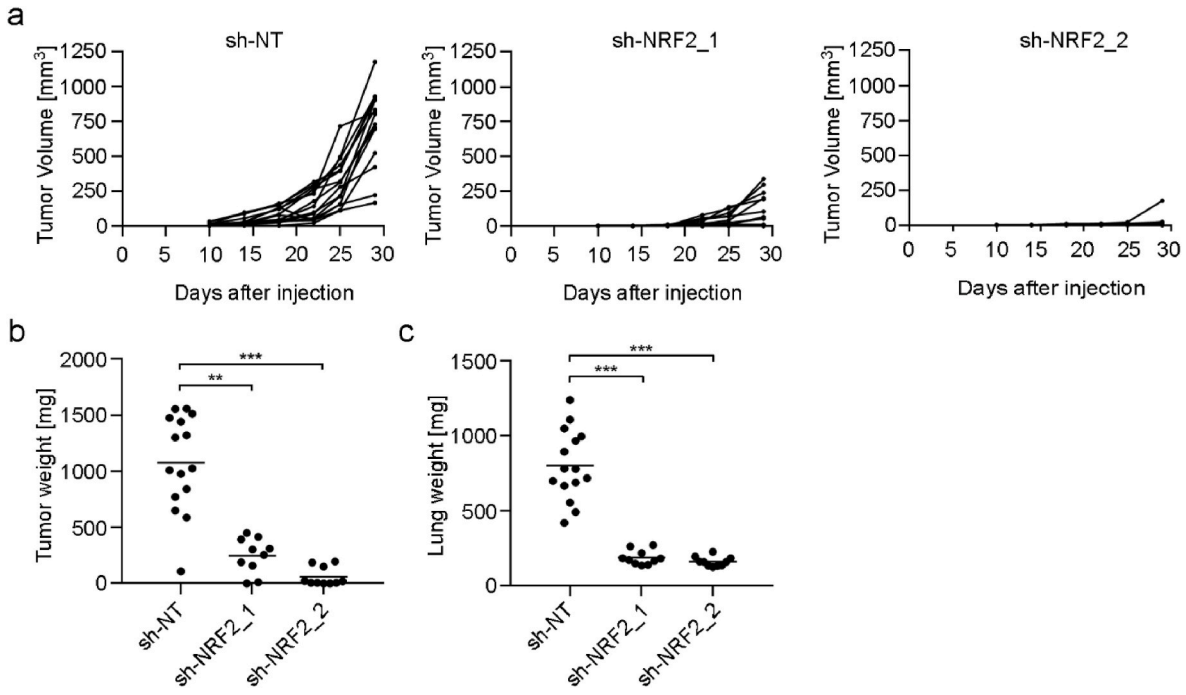


Fig. 5. NRF2 depletion abolishes primary and secondary tumor growth. (a) Growth curve of primary tumors after injection of 66cl4 sh-NT (15 mice), sh-Nrf2_1 (10 mice) and sh-Nrf2_2 (10 mice). (b) Tumor weight recorded at end point (n = 10–15). Data are shown as a scatter dot plot with a line at mean value (ANOVA after log transformation, Dunnett's multiple comparisons test). (c) Lung weight recorded at end point (n = 10–15). Data are shown as scatter dot plot with a line at mean value (ANOVA after log transformation, Dunnett's comparisons test).

2.6. The 16-transcript, NRF2 related gene expression signature is a strong and independent predictor of risk of recurrence compared to the PAM50 gene expression signature

For breast cancer patients with hormone receptor-positive, HER2 negative tumors, it is currently challenging to separate those that are unlikely to relapse from those with a higher probability for advanced disease and that will benefit from chemotherapy. Several expression signatures have been developed the last years, to help in the stratification of patients [33–35]. The FDA approved PAM50 is a classifier based on the expression of 50 genes that divides breast cancer into five subgroups; luminal A, luminal B, HER2-enriched, basal-like and normal-like [35]. Together with the PAM50 classifier, a risk of recurrence (ROR) score was developed based on the correlation to the different subtypes. Together, the PAM50 subtypes and ROR scores have an impact on the assessment of late recurrence and is clinically approved to aid in clinical decision-making. As the 16-gene NRF2 signature was identified totally independently of the transcripts included in the PAM50 gene set, and since only one gene is in common (*MELK*), we asked if the NRF2 signature identified here could increase the prediction of outcomes based on the ROR score. First, we assessed the signature score in tumor biopsies of the METABRIC cohort (n = 1904) [36]. Considering the PAM50 molecular subtypes we found that the signature level was significantly different between the subtypes (Fig. 7a) (ANOVA p-value < 0.01). Furthermore, the signature was overall higher in ER negative compared to ER positive tumors as well as in HER2 positive compared to HER2 negative, although there was a diverse spectrum within each group (Fig. 7b and c). We then used univariate Cox regression to assess the prognostic value of the signature alone using death from breast cancer as endpoint. We found that the signature was significantly prognostic alone (Table 2). Importantly, this was also true only within the hormone receptor-positive, HER2 negative tumors (Table 2). From this, we concluded that the NRF2 related signature represents an independent prognostic predictor of breast cancer. Finally, to investigate any added prognostic value of using the signature together with ROR we used ANOVA testing to perform a deviance analysis to compare regression models of PAM50 ROR alone or PAM50 ROR together with the signature. These comparisons revealed that the NRF2 related signature significantly improved the prognostic performance of PAM50 ROR (p < 2e-06). Importantly, this held true also for the hormone receptor-positive, HER2 negative tumors, suggesting an improved prognostication of this clinically challenging subgroup of patients.

3. Discussion

Transcriptome analyses revealed a constitutive activation of NRF2 in the metastatic 66cl4 cell line both in culture and in established tumors. Somatic mutations in *NFE2L2*, *KEAP1*, or *CUL3* causing activation of NRF2, have frequently been described in lung- and other squamous-like cancers [6,7,14]. In breast cancer, however, such alterations are less

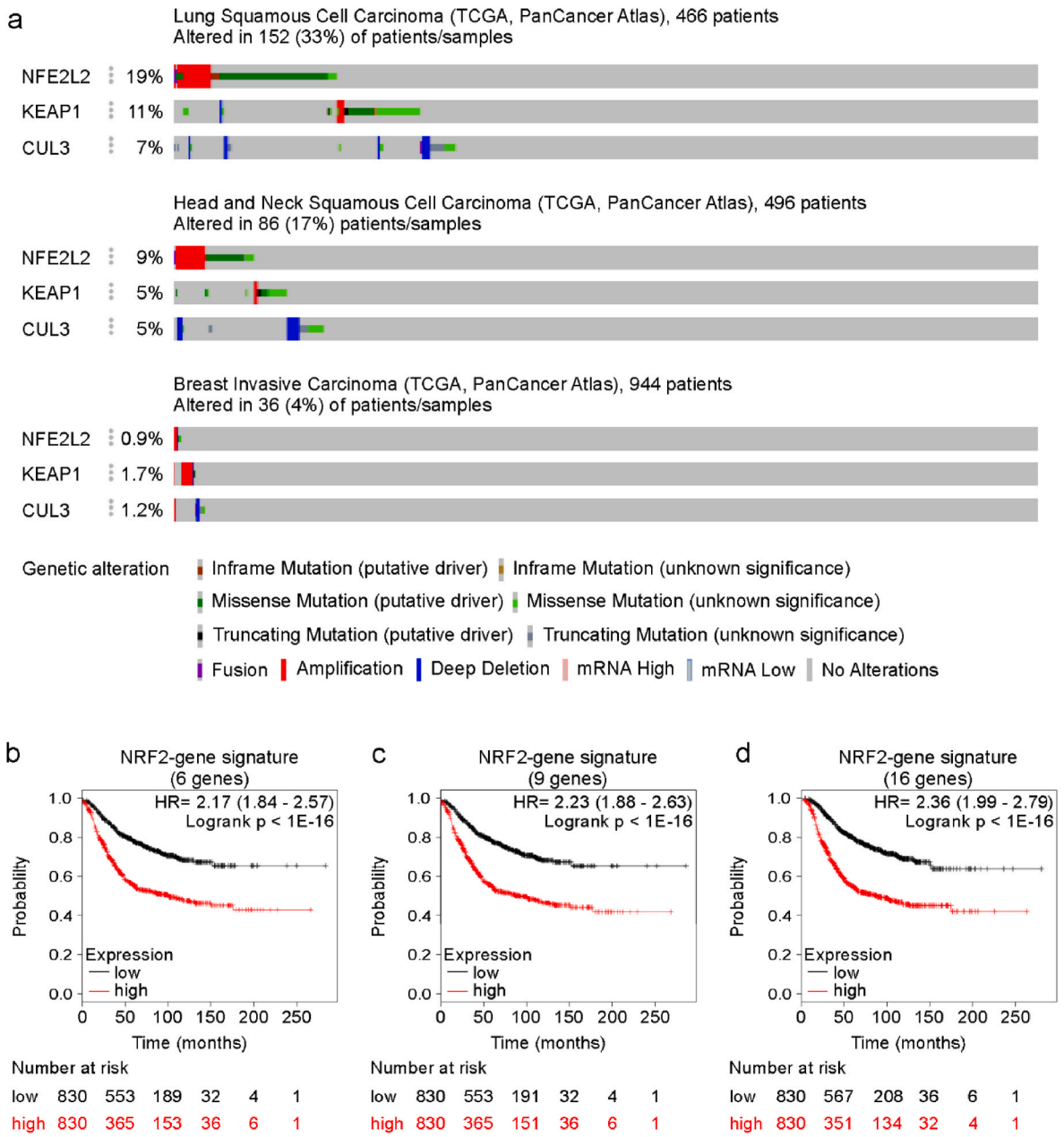


Fig. 6. Breast cancer specific NRF2-gene signature predicts relapse-free survival. (a) Gene expression and mutation status of NFE2L2, KEAP1 and CUL3 in biopsies from lung squamous cell carcinoma (n = 466), head and neck squamous cell carcinoma (n = 496) and breast invasive carcinoma (n = 944) from the TCGA PanCancer Atlas analyzed using the cBioPortal database. Each line represents all patients in the cohort and the lines are aligned to illustrate genetic alterations in each individual patient. (b–d) Analysis of relationships between gene expression and relapse-free survival (RFS) in breast cancer patients using the online tool KM plotter. High and low expression were defined as above and below median. (b) Relationship between mean expression of the six NRF2-target genes. (c) Relationship between mean expression of the 9-gene expression signature and RFS. (d) Relationship between mean expression of the 16-gene expression signature and RFS. HR, hazard ratio.

frequent. Exome sequencing uncovered a missense mutation in *Cul3* (c. C1768G) in 66cl4 that results in the substitution of histidine with aspartic acid. Sanger sequencing showed that 66cl4 is the only of the five cell lines of the 4T1 model that carries this mutation. In this conserved region of *Cul3*, somatic missense mutations leading to the exchange of a positively charged amino acid with an uncharged or negatively charged amino acid have also been found in human cancers [37]. We therefore speculate that these mutations encode loss of function or dominant-negative CUL3 variants, which cause the upregulation of NRF2 protein levels in these breast cancer cells.

The constitutively raised NRF2 level in 66cl4 coincided with reduced basal ROS levels, which was reverted by depletion of NRF2. Down-regulation of NRF2 protein levels severely reduced 66cl4's ability to form primary tumors and lung metastases. Interestingly, 66cl4 showed a lower capacity to cope with additional oxidative stress than 67NR. As a consequence of the constant upregulation of redox adaption mechanisms, cancer cells are more sensitive to further exogenous ROS insults than normal cells and this difference in sensitivity might provide a therapeutic window in certain types of cancers such as ovarian, lung, breast, liver, and esophageal cancer [13,38–41].

NRF2 driven oxidative stress responses and autophagy are closely intertwined. In autophagy defective livers (Atg5 −/−), there is a persistent liver inflammation ultimately leading to liver cancer [42]. In these livers, the autophagy receptor SQSTM1 (also known as p62) accumulates to high levels due to lack of lysosomal degradation [42]. The high levels of SQSTM1 sequester KEAP1 and cause accumulation of

Table 1
Top 25 gene transcripts with the highest wTO-score that correlates the best with the expression of each of the initial 6 NRF2 controlled genes in breast cancer biopsies. The colored cells are transcripts that predict poor prognosis when expressed above median using BreastMark and KM plotter. Transcripts in bold indicate gene products that appear more than one time in the table.

NQO1	SRXN1	TXN	TXNRD1	TALDO1	SERPINE1
MLPH 0.0505	TBC1D20 0.0317	**BUB1** 0.1014	**TPX2** 0.0678	**TPX2** 0.0642	FAP 0.08
FOXA1 0.0483	CSNK2A1 0.0232	**TPX2** 0.0993	**CENPA** 0.0674	**BUB1** 0.0614	GLT8D2 0.0768
SPDEF 0.0398	C20orf116 0.0195	**CENPA** 0.0983	**BUB1** 0.0659	**KIF2C** 0.0607	THBS2 0.0758
SRrp35 0.0393	PSMF1 0.0162	**KIF2C** 0.0981	**CCNB2** 0.0653	**CENPA** 0.0603	VCAN 0.0758
AGR2 0.0372	NSFL1C 0.0154	**CDCA8** 0.0937	**KIF2C** 0.0639	**CCNB2** 0.059	FSTL1 0.0743
FOXC1 0.0357	FAM110A 0.0147	**CCNB2** 0.0927	**CCNA2** 0.0638	**UBE2C** 0.0586	DCN 0.0742
B3GNT5 0.0356	CSNK2A1P 0.0115	**TTK** 0.0923	**UBE2C** 0.0625	**CDCA8** 0.058	COL5A1 0.0734
PRR15 0.0339	SFRP1 0.0115	**DLGAP5** 0.0922	**CDCA8** 0.0624	**CDCA5** 0.0567	COL6A3 0.0726
C10orf38 0.0338	CCDC82 0.0112	MCM10 0.0892	**CEP55** 0.0614	**DLGAP5** 0.0555	COL5A2 0.0723
XBP1 0.0329	**ROPN1** 0.011	**CEP55** 0.0888	**CDCA5** 0.061	**C15orf42** 0.0548	COL6A2 0.0717
ROPN1 0.0311	**FOXC1** 0.0109	**MELK** 0.0888	**DEPDC1B** 0.0606	**KIF4A** 0.0546	LUM 0.0704
CA12 0.0303	**C10orf38** 0.0102	**UBE2C** 0.0884	**MCM10** 0.0601	**CEP55** 0.0545	SPARC 0.0703
GATA3 0.0293	MRPS26 0.01	**CDCA5** 0.0873	**DLGAP5** 0.0596	**CCNA2** 0.0545	COL1A2 0.07
AR 0.0293	**SRrp35** 0.0098	**CCNA2** 0.0864	**FOXM1** 0.058	**MCM10** 0.054	FBN1 0.0697
ART3 0.0292	ATRN 0.0096	**C15orf42** 0.0842	**C15orf42** 0.0574	**FOXM1** 0.0524	SERPINF1 0.0688
ESR1 0.0289	**B3GNT5** 0.0095	**DEPDC1B** 0.084	**CENPE** 0.0569	**TTK** 0.0523	ADAM12 0.0679
FOXL1 0.0288	**RGMA** 0.0095	**FOXM1** 0.0832	**KIF4A** 0.0568	**MELK** 0.0523	CCDC80 0.0674
TBC1D9 0.0285	C20orf29 0.0094	**KIF4A** 0.0811	**TTK** 0.0568	**DEPDC1B** 0.0518	PCOLCE 0.0672
BCL11A 0.0284	MIA 0.0093	**CDC20** 0.0809	**PTTG3** 0.0567	**CENPE** 0.0512	C2orf32 0.0667
PPP1R14C 0.028	C17orf28 0.0093	**BUB1B** 0.0794	PTTG1 0.0563	**CDC20** 0.0504	DACT1 0.0659
UGT8 0.0277	**ROPN1B** 0.0092	**CENPE** 0.0792	FAM83D 0.0556	**BUB1B** 0.0495	BNC2 0.0657
SLC44A4 0.0275	ZNF343 0.0091	NCAPH 0.0785	**MELK** 0.0551	**PTTG3** 0.0489	MMP2 0.0655
AGR3 0.0273	**BCL11A** 0.009	EXO1 0.0784	PTTG2 0.0549	**CDCA3** 0.0487	SPON1 0.0651
ROPN1B 0.0263	**MLPH** 0.0088	ASPM 0.0783	CDC45L 0.0548	CENPI 0.0482	HTRA1 0.0643
RGMA 0.0261	STK35 0.0088	**PTTG3** 0.0781	**CDCA3** 0.0532	KIF20A 0.0481	AEBP1 0.0642

NRF2 [15,43]. In this setting, NRF2 also induce the expression of SQSTM1 to form a positive feedback loop [16]. Despite constitutive activation of NRF2 in 66cl4, SQSTM1 mRNA and proteins levels were similar in 66cl4 and 67NR. We also found no consistent downregulation of SQSTM1 protein levels in NRF2 depleted cells, even though NQO1 and HMOX1 were clearly reduced. These results demonstrate that there are cell specific differences in the NRF2 controlled gene regulation and that other regulators of SQSTM1 could be more important in these cells. Reciprocally, the ability of SQSTM1 to regulate NRF2 through sequestering KEAP1 is context dependent and change by post-translational modifications and alternative splicing (in mice) of SQSTM1 [44,45]. However, elevated expression of SQSTM1 might still be important for aggressive breast cancer development since there is a correlation between elevated SQSTM1 transcript levels in tumor biopsies and poor prognosis using KM plotter (HR = 1.35, p = 8.7 logrank P = 8.7e-08, n = 3951).

The clinical relevance of our findings is highlighted by the observed correlation between elevated score of the 16-transcript signature in RNA

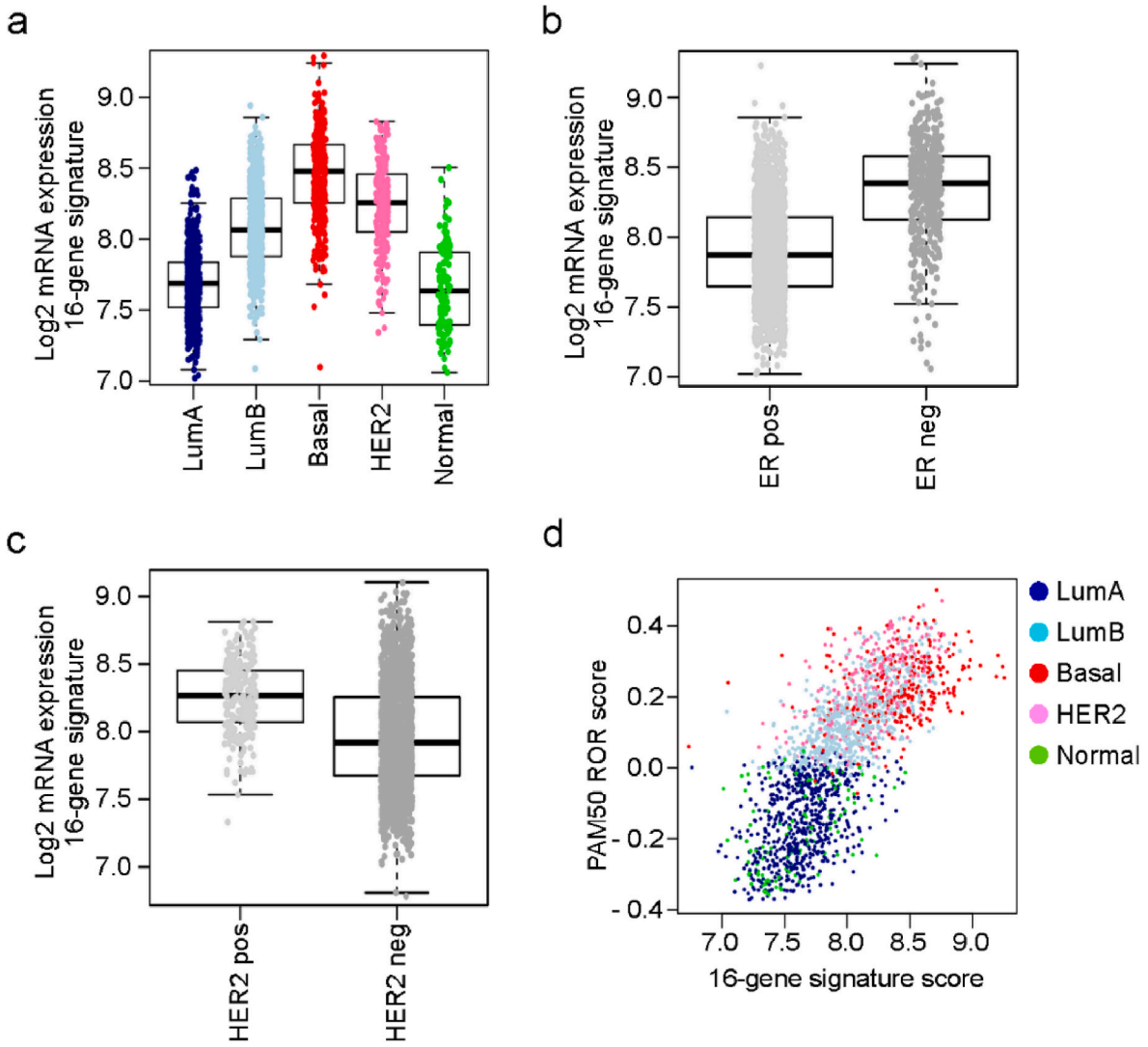


Fig. 7. The 16-transcript, NRF2 related gene expression signature is a strong predictor of risk of recurrence compared to the PAM50 gene expression signature. (a) 16-gene expression signature level considering the PAM50 molecular subtypes in tumor biopsies of the METABRIC cohort (n = 1904, p < 1E-273). (b) The 16-gene expression signature level in ER positive versus ER negative tumors (p < 6E-83). (c) The 16-gene expression signature level in HER2 positive versus HER2 negative tumors (p < 6E-37). (d) Correlation between PAM50 ROR and the 16-gene expression signature (correlation score = 0.76).

Table 2
Prognostic value of the 16-gene expression signature and PAM50 ROR with death from breast cancer as endpoint. The univariate Cox regression analysis was performed using gene expression of breast cancer biopsies from METABRIC (all samples n = 1904, hormone receptor-positive, HER2 negative samples n = 944) SE; standard error, HR; hazard ratio.

	All samples (n = 1904)				Hormone receptor-positive, HER2 negative samples (n = 944)			
	Coefficient	SE	HR	p-value	Coefficient	SE	HR	p-value
NRF2 signature	0.2592555	0.02444357	1.295965	7.97E-26	0.3236054	0.03801978	1.382102	9.70E-17
PAM50 ROR	0.1823023	0.01918524	1.199977	1.46E-22	0.210953	0.0292933	1.234854	4.07E-13

from tumor biopsies and unfavorable prognosis of breast cancer patients. A striking observation was that the signature was prognostic irrespective of subtype. The majority of breast cancer patients are diagnosed with a small, low-grade, hormone receptor-positive tumor. This patient group is challenging to treat since only few will relapse after surgery and radiotherapy. Subsequent to this treatment, the question is whether these patients should proceed with long-term chemotherapy. Such additional treatment further reduces the low risk of relapse but also give side effects reducing quality of life for many of these patients. A risk of recurrence score based on a combination of the FDA approved PAM50 gene expression signature and clinical features of the tumor is already in use to aid in the stratification of breast cancer patients. We found 16 transcripts that report on NRF2 activity and that when elevated correlate with reduced RFS in breast cancer patients. Combined, these 16 transcripts have a greater prognostic value than the level of either transcript alone. The prognostic value was evident even for patients with low-grade, hormone receptor-positive breast tumors. Thus, the NRF2-gene signature could potentially be used to identify patients that will most likely not experience metastatic relapse. The 16-transcript signature was also an independent prognostic predictor compared to PAM50 ROR, and that significantly improved PAM50 ROR when combined. We hope that such a diagnostic approach may contribute to reduce the number of

patients that develop short-term and long-term adverse effects of anti-cancer therapies.

4. Materials and methods

4.1. Cell culture and generation of stable cell lines

67NR and 66cl4 cells were obtained from Barbara Ann Karmanos Cancer Institute. 168FARN, 4TO7 and 4T1 were kindly provided by Dr. Fred Miller (Wayne State University, Detroit, MI). All cell lines were cultured in DMEM (Lonza, BioWhittaker, BE12- 604F) supplemented with 10% fetal calf serum (FCS, Thermo Fischer Scientific, #10270-106), 2 mM L-Glutamine (Lonza Group, De-17-605E) and 50 U/ml penicillin-streptomycin (Thermo Fischer Scientific, Gibco, #15070-063). Cells were incubated at 37 °C with 5% CO_2.

ShRNA-NRF2 knockdowns and respective controls were generated by viral transduction (Sigma Aldrich: TRCN0000054658, SHC216V-1 EA). 6 h after seeding, 66cl4 cells were infected with lentiviral particles (MOI 0.1) in medium containing hexadimethrine bromide (8 μg/ml). After 24 h, cells were split 1:17. Starting 48 h after infection, cells were selected with puromycin (3.25 μg/ml) for 1 week. The selection medium was replaced every 2 – 3 days. Single cell colonies were picked using cloning cylinders and tested for NRF2 expression levels.

4.2. Orthotopic mouse tumors and in vivo lung colonization assay

Experiments were approved by the National Animal Research Authorities and carried out according to the European Convention for the Protection of Vertebrates used for Scientific Purposes (FOTS ID 4536 and FOTS 10049). For all experiments, female BALB/cJRj mice (8–11 weeks old, Janvier Labs, France) were used.

For orthotopic tumors, mice were anaesthetized and injected with 5×10^5 viable cells into the fourth mammary fat pad. When palpable, tumor size was measured twice weekly using electronic calipers. Tumor volume was calculated: $V_T = (\text{length x width}^2)/2$. All mice belonging to the same experiment were sacrificed on the same day, unless otherwise stated. Mice were sacrificed after three to four weeks as indicated in the results. Weight of primary tumors and lungs was recorded. For the in vivo lung colonization assay female mice were anaesthetized and injected with 5×10^5 cells/100 μl PBS in the lateral tail vein. Mice were monitored daily and sacrificed two to three weeks after injections as indicated in the results. Entire lungs were weighed.

4.3. Transcriptome analysis

As described previously, RNA was isolated from three passages of 67NR and 66cl4 in culture, four and seven primary tumors of 67NR and 66cl4, respectively. Detailed information about sample preparation and data analysis can be found in Ref. [4]. The transcriptome data obtained by sequencing mRNA isolated from cells and primary breast tumors of 67NR and 66cl4 is accessible from NCBI (SUB6422687).

4.4. Genomic DNA isolation

DNA from 67NR, 66cl4, and blood of BALB/c mice intended for exon sequencing was isolated using the QIAGEN Blood & Cell Culture DNA Kit (Qiagen, #13323). Everything was done according to the QIAGEN Genomic DNA Handbook.

DNA used for Sanger sequencing was isolated according to the following protocol. Parts of primary tumors and entire lungs from in vivo experiments with 66cl4 NRF2 knockdowns were homogenized in lysis buffer (w/o SDS and digestion enzymes) using 1,4 mm ceramic beads from Precellys (2 × 40 s homogenization and 2 min break in between cycles). Homogenized tissue samples and cell pellets from tumor cell lines were digested overnight at 55 °C in lysis buffer (50 mM Tris, HCl pH 8.0; 100 mM NaCl; 100 mM EDTA, pH 8.0; 0,5% SDS; 100 μg/ml proteinase K [Roche, #03 115 828 001); 100 μg/ml RNaseA [Qiagen, #19101]). Genomic DNA (gDNA) was isolated using super saturated salt solution NaCl and ethanol precipitation. The DNA pellet was dissolved in T1/10E buffer. Purity and quantity were measured by a Nanodrop microvolume spectrophotometer.

4.5. Exome sequencing

Exome sequencing libraries were prepared from 1 μg gDNA using SureSelectXT target enrichment system for Illumina paired-end sequencing libraries (Agilent Technologies, Santa Clara, CA, USA), according to the manufacturer's instructions. Briefly, the DNA was fragmented using the Covaris M220 system (Covaris, Woburn, MA, USA). The DNA fragments (app 200 bp) were end-repaired using T4 DNA polymerase, Klenow DNA polymerase and T4 polynucleotide kinase (PNK), followed by purification using AMPure XP beads (Beckman Coulter, Brea, CA, USA). An A-base was ligated to the blunt ends of the DNA fragments using the Klenow DNA polymerase and the sample was purified using AMPure XP beads. Indexing specific adapters for sequencing were ligated to the DNA fragments, followed by purification using AMPure XP beads. The adapter-ligated libraries were amplified for six PCR cycles, followed by a second purification using AMPure XP beads. The quality of the enriched libraries was evaluated using the 2100 Bioanalyzer and a DNA 1000 kit (Agilent, Santa Clara, CA, USA) and qPCR. Exon capture was performed from 1000 ng of each sequencing library using the SureSelectXT SureSelect Mouse All Exon Kit (Agilent Technologies, Santa Clara, CA, USA). Briefly, the fragments in the library were hybridized to capture probes (20 h at 65C), unhybridized material was washed away and the captured fragments were amplified for ten PCR cycles, followed by purification using AMPure XP beads. The quality of the enriched libraries was evaluated using the 2100 Bioanalyzer and a High-Sensitivity DNA-kit (Agilent Technologies, Santa Clara, CA, USA). The adapter-ligated fragments were quantified by qPCR using the KAPA SYBR FAST library quantification kit for Illumina Genome Analyzer (KAPA Biosystems, Woburn, MA, USA). A 20 pM solution of the sequencing libraries was subjected to cluster generation on a HiSeq2500 rapid ruin mode flowcell by the cBot instrument (Illumina, Inc., San Diego, CA, USA). Paired-end sequencing was performed for 2 × 100 cycles on a HiSeq2500 instrument (Illumina, Inc. San Diego, CA, USA). Everything was done according to the manufacturer's instructions. Base calling was done on the HiSeq instrument by RTA 1.17.21.3. FASTQ sequence files were generated using CASAVA 1.8.2 (Illumina, Inc. San Diego, CA, USA). Raw FASTQ files from three replicates were combined and aligned to the mm9 genome reference by BWA (http://bio-bwa.sourceforge.net/), version 0.6.2, for each of 66c14, 67NR, and blood, respectively. Sequence Alignment Map (SAM) files were converted to Binary Alignment Map (BAM) files by Picard (https://broadinstitute.github.io/picard/), version 1.102 and sorted by Samtools (http://samtools.sourceforge.net/), version 0.1.18. The BAM files were subsequently preprocessed by the GATK pipeline. Single nucleotide variants were called by Mutect (https://www.nature.com/articles/nbt.2514), version 1.1.7, and short insertions and deletions (indels) were called by Strelka (https://www.ncbi.nlm.nih.gov/pubmed/22581179), version 1.0.14. SNVs and indels were annotated by Annovar (http://annovar.openbioinformatics.org/en/latest/), version 2013May09, and exonic, nonsynonymous mutations were kept for further analysis.

4.6. Sanger sequencing

PCR was performed with KOD Xtreme Hot Start DNA polymerase (Novagen, 71975-3). The 50 μl reactions contained 25 μl 2x Xtreme Buffer, 10 μL 2 mM dNTPs, 1.5 μl 10 μM forward primer (ACTGTGTAGGGAAATGGGGAC) and 1.5 μl 10 μM reverse primer (CCTACTGGTATTCAAATGGCATCTC) and 100 ng gDNA. The cycling conditions of the touchdown PCR were 94 °C for 120 s, 16 cycles of 98 °C for 10 s, 62

°C (minus 0.5 °C/cycle) for 15 s, 68 °C for 20 s, 24 cycles of 98 °C for 10 s, 55 °C for 15 s, 68 °C for 20 s and 68 °C for 120 s. The size and purity of the PCR product was examined on a 2% agarose gel.

ExoSAP-IT (Thermo Fisher Scientific, 78200.200.UL) was used for PCR cleanup. Sanger sequencing was done with the Cul3 reverse primer by GATC Biotech AG, Germany.

4.7. Quantitative real-time PCR

Total RNA from cells in culture was extracted using RNeasy Mini Kit (Qiagen, 74104). RNA concentration and quality were measured by Nanodrop. cDNA was synthesized from 500 ng total RNA by QuantiTect Reverse Transcription Kit (Qiagen, 205310). cDNA was diluted 1:5. Quantitative real-time PCR was performed in parallel 25 μl reactions containing 12.5 μl 2X QuantiTect SYBR Green PCR master mix (Qiagen, 204141) and 2.5 μl 10X QuantiTect Primer Assay. Qiagen QuantiTect Primer Assays used in this study: Mm_Nfe2l2_1_SG (QT00095270), Mm_Nqo1_1_SG (QT00094367), Mm_Hmox1_1_SG (QT00159915), Mm_Ftl1_2_SG (QT02526006), Mm_Gclc_1_SG (QT00130543), Mm_Sqstm1_1_SG (QT00127855), Mm_Cul3_1_SG (QT00108010), Mm_Map1lc3b_2_SG (QT01750322), Mm_Actb_2_SG (QT01136772), Mm_Tbp_1_SG (QT00198443). RT-PCR was performed on the StepOne plus system (Applied Biosystems, Foster City, CA, USA) using the following cycling conditions: 95 °C for 15 min, 40 cycles of 94 °C for 15 s, 55 °C for 30 s and 72 °C for 30 s. Relative gene expression levels were calculated with the $2^{\wedge(-\text{delta delta CT})}$ method. Transcripts were normalized to *Actb* and *Tbp*.

4.8. Immunoblotting

Cells were harvested in urea lysis buffer containing 8 M urea (Merck Millipore, 1084870500), 0.5% (v/v) Triton X-100 (Sigma, T8787), 100 mM DTT (Sigma, 646563), Complete® protease inhibitor (Roche, 1187350001) and phosphatase inhibitor cocktail II (Sigma, P5726) and III (Sigma, P0044). When indicated the cells were pretreated for 4 h with L-Sulforaphane dissolved in DMSO (20 μM – 50 μM, Sigma Aldrich, S6317) or for 30 min with hydrogen peroxide (100 μM – 500 μM, Sigma Aldrich, H1009). Protein concentration was measured using BioRad protein assay (Bio-Rad, 500-0006). Equal amounts of proteins were run on Invitrogen NuPAGE Bis-Tris protein gels (10%, 12% or 3–12%), transferred onto nitrocellulose membranes and probed with antibodies against NRF2 (Cell signaling, 12721, 1:1000), NQO1 (Abcam, ab34173, 1:1000), HMOX1 (Enzo, ADI-OSA-110, 1:1000), FTL1 (ThermoFisher Scientific PA5-27357, 1:1000), FTH1 (Abcam, ab65080, 1:1000), SQSTM1/p62 (Progen, GP62-C, 1:1000) or ACTB (Abcam, ab6276, 1:10 000). Proteins of interest were detected with near-infrared fluorescent (IRDye) secondary antibodies (Li-Cor Biosciences, 926–32211, 926–32411, 926–68070, 1:10 000) and imaged with the Odyssey Near Infrared scanner (Li-Cor Biosciences, Lincoln, Nebraska, USA). Images were analyzed using Image Studio v3.1.

4.9. Detection of reactive oxygen species

ROS levels determined by flow cytometric measurement of CM-H2DCFDA (ThermoFisher Scientific, C6827). When indicated the cells were pretreated for 4 h with L-Sulforaphane dissolved in DMSO (20 μM – 50 μM, Sigma Aldrich, S6317) or vehicle (DMSO). The cells were incubated at 37 °C and 5% CO_2 with 5 μM CM-H_2DCFDA for 30 min before intracellular ROS was determined. The experiments were performed in triplicates and the data represent mean intensity of 10 000 cells per well ± SD.

4.10. MTT

67NR and 66cl4 were seeded in 96 well plates (6000 cells/well) one day prior stimulation. The cells were stimulated with H_2O_2 (100 μM – 600 μM, Sigma Aldrich, H1009) for 30 min and the medium was replaced by new growth medium (200 μl). After 16 h incubation at 37 °C and 5% CO_2, 20 μl MTT (5 mg/ml) were added to the each well and incubated at 37 °C, 5% CO_2 for 4 h. Next, 150 μl of supernatant were discarded, 100 μl acidic isopropanol (40 mM HCl) added and the plate incubated for 1 h while shaking. To dissolve the MTT formazan crystals, samples were thoroughly pipetted up and down. Absorbance was measured at 570 nm with a Biorad iMark Microplate™ reader. Per condition 6 parallels were measured.

4.11. Cell proliferation assay

12 000 cells were seeded in 24 well plates and counted at around 24 h, 36 h, 48 h, 60 h and 72 h. Measurements were done in triplicates and a single plate was prepared for each time point. Each of the triplicates was counted three times with the Beckman Coulter Z2 Coulter Particle Count and Size analyzer. An average cell number per well was calculated for each cell line and time point ($cells_{avg}$). The natural logarithm (ln(X)) of each $cells_{avg}$ was calculated and plotted against time (hours). Growth rate (μ max [h-1]) was determined and doubling time (tD [h]) calculated: ln(2)/μ max.

4.12. XF96 Seahorse measurements of key parameters of glycolytic and mitochondrial function

67NR (13 000 cells/well), 66cl4 (20 000 c/w) and 66cl4 clones (sh-NT and sh-NRF2 knockdowns) (20 000 c/w) were seeded in XF96 polystyrene cell culture microplates that had been pre-treated with 0.01% poly-L-lysine. The cells were incubated overnight, reaching around 95% confluency the following day. At least 1 h prior to measurements, the cells were washed and incubated at 37 °C in a CO_2-free incubator in serum-free XF Assay Medium Modified DMEM (Seahorse Biosciences, Part #102352-000) supplemented with either 6 mM glutamine (glycolysis stress test) or a combination of 10 mM glucose, 10 mM pyruvate and 6 mM glutamine (mito stress test).

Glycolysis stress test: Basal extracellular acidification rate (ECAR) was measured in an XF96 Extracellular Flux Analyzer (Seahorse Bioscience, North Billerica, MA) and then the ECAR was followed after injections of a saturating concentration of glucose (10 mM), oligomycin (1 μM) and finally 2-deoxy-glucose (50 mM).

Mito stress test: Basal oxygen consumption rate (OCR) was measured, and then the OCR was followed after injections of oligomycin (1 μM), carbonylcyanide p-trifluoromethoxyphenylhydrazone (FCCP; 1.5 μM) and the combination of antimycin A (AA; 1 μM) and rotenone (1 μM).

After the analysis, most of the medium in the wells was removed (ca. 40 μl left), and 16% paraformaldehyde (PFA) was added to a final concentration of 4%. The cells were fixed at RT for ca. 15 min. For normalization, cell nuclei were stained with Draq5 (Cell Signalling Technology, 4084 L) by carefully removing the PFA and then adding 50 μl Draq5 (diluted 1:1000 in PBS) and incubating in the dark at RT for 0.5 h. The wells were scanned at 680 nm using an Odyssey NIR scanner (Li-Cor Biosciences).

All experiments were performed 3 times, each time using 14–20 wells per cell line. All chemicals used in the glycolysis and mito stress tests were from Sigma-Aldrich.

4.13. Soft-agar assay

2x DMEM growth medium was prepared by dissolving 10 g DMEM low glucose (Sigma Aldrich, D2902), 3.7 g sodium bicarbonate, ($NaHCO_3$), and 3.5 g glucose in Milli-Q water. After sterile filtration 20% fetal calf serum (FCS, Thermo Fischer Scientific, #10270-106), 4 mM L-Glutamine (Lonza Group, De-17-605E) and 100 U/ml penicillin-streptomycin (ThermoFischer Scientific, Gibco, #15070-063) were added. PH of the final 2x DMEM was 7.4. A bottom layer of a 1:1 mix of

0.75% SeaKem® LE Agarose (LONZA, 50004) and 2x DMEM was prepared and added to each well in 6 well plates. The top layer containing a 1:1 mix of 0.36% SeaKem® LE Agarose and 2x DMEM with 10 000 cells was added to the solid bottom layer. The cells were fed twice a week with 1x DMEM and incubated at 37 °C and 5% CO_2 for 12 days. Plates were stained with 0.04% crystal violet in 2% ethanol/PBS for 1 h and washed 6 × 60 min with 1x PBS while rocking. Z-stack images consisting of nine images were taken with the EVOS FL Auto Cell Imaging System (Invitrogen, Carlsbad, California, USA). For each cell line, 5 to 6 wells, containing 25 beacons each, were analyzed. Z-stacks were merged and analyzed with CellProfiler 2.2.0.

4.14. Survival, enrichment and gene expression analysis

KM plotter and Breastmark – BreastMark (glados.ucd.ie/BreastMark) [30] and KM plotter (http://kmplot.com/analysis/) (database version: 2014) [28,29] are online custom tools for examining prognostic markers in BC subtypes, which utilize data from multiple cDNA microarray experiments. Analysis of disease-free/relapse-free survival (DFS/RFS) of 46 NRF2-target genes was analyzed in breast cancer as a whole and intrinsic subtypes (lymph node positive/negative, Luminal A, Luminal B, Her2 positive, basal, and estrogen receptor [ER] positive/negative).

Enrichr - Enrichment analysis, including Chip-X enrichment analysis (ChEA), KEGG cell signaling pathway analysis, Reactome pathway analysis and GO cellular component analysis, were done with the online tool Enrichr (http://amp.pharm.mssm.edu/Enrichr/) [17,18].

Ualcan - mRNA expression levels of 61 NRF2-target genes were determined in breast invasive carcinoma tissue and normal tissue (TCGA dataset) using Ualcan (http://ualcan.path.uab.edu/index.html), which is interactive web resource for analyzing cancer transcriptome data [31].

cBioPortal - cBioPortal (http://www.cbioportal.org/) is an open access database that allows visualization and analysis of large-scale cancer genomics data sets [22,23]. Our analyses utilize the OncoPrints visualization to identify genomic alterations, including somatic mutations, mRNA expression and amplifications across a set of cases. This visualization shows the genes as rows, while individual cases are shown as columns. For this analysis we used the TCGA Provisional data set for invasive breast carcinoma, and selected mutations, putative copy-number alterations and mRNA expression as genomic profiles.

4.15. Statistical analysis of the signature

The signature score was assessed in the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) cohort [36] using microarray measurements of tumor mRNA expression and clinical information downloaded from the cBioPortal on December 1st, 2017, http://www.cbioportal.org/study?id=brca_metabric#clinical [22,23]. Statistical analyses were performed using R [46]. For comparisons of two groups, the Student's *t*-test was used and for comparison of more than two groups (PAM50) an Analysis of Variance Model (ANOVA) was used. A p-value < 0.05 was considered statistically significant.

To assess the prognostic value of the signature univariate Cox regression from the survival library in R was used with death from breast cancer as endpoint.

To investigate any added prognostic value of using the signature together with ROR we used ANOVA testing to compare regression models of PAM50 ROR alone or PAM50 ROR together with the signature.

4.16. Weighted topological overlap network analysis

The network analysis was performed on RNA-sequenced data on 17 214 genes from 421 cases and 51 controls, downloaded from The Cancer Genome Atlas (TCGA; The Cancer Genome Atlas research network: http://cancergenome.nih.gov. The data was already normalized within samples to the median gene expression, and genes with missing values were removed prior to the analysis, resulting in 421 cases and 61 controls with 16 750 gene expressions.

Applying the standard procedure of Weighted Gene Co-Expression Network Analysis (WGCNA) [47], we generated correlation-based networks for both cases and controls, where nodes correspond to genes and a link between a pair of genes represents the interaction between the genes. Briefly, the WGCNA approach starts by creating correlation adjacency matrices for the gene expressions. Using a power function, strong correlations accentuate, while weak correlations are reduced [47,48]. These transformed correlations are used to calculate the weighted Topological Overlap (wTO),

$$\omega_{ij} = \frac{a_{ij} + \sum_{u \neq i,j} a_{iu} a_{uj}}{\min\{\sum_u a_{iu},\ \sum_u a_{ju}\} + 1 - a_{ij}},$$

where $a_{ij} = |cor_{ij}|^{\beta}$ is the absolute Pearson correlation between gene i and j raised to a power $\beta = 5$. Accounting also for the correlation of the surrounding nodes, the wTO has been shown to be a robust and biological meaningful measurement [32,47–49]. Letting the wTO-values represent links in the network, we focused on the target genes by creating ego-centric networks for the target genes, selecting only the 25 strongest links for each target gene.

4.17. Statistics

Statistical analyses were performed in GraphPad Prism 7 and 9. Values are expressed as mean ± standard deviation (SD) or standard error of the mean (SEM) if not otherwise stated. Details about statistical analyses are specified in the figure legends. P value < 0.05 was considered statistically significant and is labeled with *), $p < 0.01$ is labeled with **) and $p < 0.001$ is labeled with ***).

Author contributions

CW and UN designed and performed the study and experiments, analyzed and interpreted data and prepared the manuscript. MRA, MH, SBA, RZ and OCL performed experiments and analyzed data. BJ, SZ and RIS analyzed data. TSS, KSS and EA provided insightful comments on the findings through the project and reviewed the manuscript. GB designed and supervised the study, interpreted data and prepared the manuscript.

Funding

This work was supported by grants from the Norwegian Cancer Society (project number #6846671 and #419654), the Research Council of Norway (Centres of Excellence funding program, project number 223255/F50), the Central Norway Regional Health Authority (project number #90181900) and a PhD-grant to CW from Norwegian Women's Public Health Association.

Declaration of competing interest

The authors declare that they have no competing interests.

Acknowledgements

We would like to thank the Cellular and Molecular Imaging Core Facility (CMIC), Faculty of Medicine, NTNU for access to instruments and support. The library preparation and sequencing analysis (RNA Seq and Exome Seq) were provided in close collaboration with the Genomics Core Facility (GCF), Norwegian University of Science and Technology (NTNU). GCF is funded by the Faculty of Medicine and Health Sciences at NTNU and Central Norway Regional Health Authority. We also thank the Comparative medicine Core Facility (CoMed), NTNU, for the help regarding the mice experiments. CoMed is funded by the Faculty of

Medicine and Health Sciences at NTNU and Central Norway Regional Health Authority. Parts of the results shown here are based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.freeradbiomed.2022.03.029.

References

[1] F.R. Miller, B.E. Miller, G.H. Heppner, Characterization of metastatic heterogeneity among subpopulations of a single mouse mammary tumor: heterogeneity in phenotypic stability, Invasion Metastasis 3 (1) (1983) 22–31.
[2] C.J. Aslakson, F.R. Miller, Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor, Cancer Res. 52 (6) (1992) 1399–1405.
[3] E. Batlle, H. Clevers, Cancer stem cells revisited, Nat. Med. 23 (10) (2017) 1124–1134.
[4] U. Neckmann, et al., GREM1 is associated with metastasis and predicts poor prognosis in ER-negative breast cancer patients, Cell Commun. Signal. 17 (1) (2019) 140.
[5] E. Piskounova, et al., Oxidative stress inhibits distant metastasis by human melanoma cells, Nature 527 (7577) (2015) 186–191.
[6] Cancer Genome Atlas Research Network, Comprehensive genomic characterization of squamous cell lung cancers, Nature 489 (7417) (2012) 519–525.
[7] K.A. Hoadley, et al., Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin, Cell 158 (4) (2014) 929–944.
[8] Y. Watai, et al., Subcellular localization and cytoplasmic complex status of endogenous Keap1, Gene Cell. 12 (10) (2007) 1163–1178.
[9] S.B. Cullinan, et al., The Keap1-BTB protein is an adaptor that bridges Nrf2 to a Cul3-based E3 ligase: oxidative stress sensing by a Cul3-Keap1 ligase, Mol. Cell Biol. 24 (19) (2004) 8477–8486.
[10] A. Kobayashi, et al., Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate proteasomal degradation of Nrf2, Mol. Cell Biol. 24 (16) (2004) 7130–7139.
[11] D.D. Zhang, et al., Keap1 is a redox-regulated substrate adaptor protein for a Cul3-dependent ubiquitin ligase complex, Mol. Cell Biol. 24 (24) (2004) 10941–10953.
[12] L.E. Tebay, et al., Mechanisms of activation of the transcription factor Nrf2 by redox stressors, nutrient cues, and energy status and the pathways through which it attenuates degenerative disease, Free Radic. Biol. Med. 88 (Pt B) (2015) 108–146.
[13] C. Gorrini, I.S. Harris, T.W. Mak, Modulation of oxidative stress as an anticancer strategy, Nat. Rev. Drug Discov. 12 (12) (2013) 931–947.
[14] L.D. Goldstein, et al., Recurrent loss of NFE2L2 exon 2 is a mechanism for Nrf2 pathway activation in human cancers, Cell Rep. 16 (10) (2016) 2605–2617.
[15] M. Komatsu, et al., The selective autophagy substrate p62 activates the stress responsive transcription factor Nrf2 through inactivation of Keap1, Nat. Cell Biol. 12 (3) (2010) 213–223.
[16] A. Jain, et al., p62/SQSTM1 is a target gene for transcription factor NRF2 and creates a positive feedback loop by inducing antioxidant response element-driven gene transcription, J. Biol. Chem. 285 (29) (2010) 22576–22591.
[17] E.Y. Chen, et al., Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool, BMC Bioinf. 14 (2013) 128.
[18] M.V. Kuleshov, et al., Enrichr: a comprehensive gene set enrichment analysis web server 2016 update, Nucleic Acids Res. 44 (W1) (2016) W90–W97.
[19] M. Rojo de la Vega, E. Chapman, D.D. Zhang, NRF2 and the Hallmarks of cancer, Cancer Cell 34 (1) (2018) 21–43.
[20] E. White, Deconvoluting the context-dependent role for autophagy in cancer, Nat. Rev. Cancer 12 (6) (2012) 401–410.
[21] M. Gry, et al., Correlations between RNA and protein expression profiles in 23 human cell lines, BMC Genom. 10 (2009) 365.
[22] J. Gao, et al., Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal, Sci. Signal. 6 (269) (2013) pl1.
[23] E. Cerami, et al., The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data, Cancer Discov. 2 (5) (2012) 401–404.
[24] B.E. Hast, et al., Proteomic analysis of ubiquitin ligase KEAP1 reveals associated proteins that inhibit NRF2 ubiquitination, Cancer Res. 73 (7) (2013) 2199–2210.
[25] W. Chen, et al., Direct interaction between Nrf2 and p21(Cip1/WAF1) upregulates the Nrf2-mediated antioxidant response, Mol. Cell 34 (6) (2009) 663–673.
[26] J. Ma, et al., PALB2 interacts with KEAP1 to promote NRF2 nuclear accumulation and function, Mol. Cell Biol. 32 (8) (2012) 1506–1517.
[27] N.D. Camp, et al., Wilms tumor gene on X chromosome (WTX) inhibits degradation of NRF2 protein through competitive binding to KEAP1 protein, J. Biol. Chem. 287 (9) (2012) 6539–6550.
[28] A. Lanczky, et al., miRpower: a web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients, Breast Cancer Res. Treat. 160 (3) (2016) 439–446.
[29] B. Gyorffy, et al., An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients, Breast Cancer Res. Treat. 123 (3) (2010) 725–731.
[30] S.F. Madden, et al., BreastMark: an integrated approach to mining publicly available transcriptomic datasets relating to breast cancer outcome, Breast Cancer Res. 15 (4) (2013) R52.
[31] D.S. Chandrashekar, et al., UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses, Neoplasia 19 (8) (2017) 649–658.
[32] B. Zhang, S. Horvath, A general framework for weighted gene co-expression network analysis, Stat. Appl. Genet. Mol. Biol. 4 (2005) Article17.
[33] L.J. van 't Veer, et al., Gene expression profiling predicts clinical outcome of breast cancer, Nature 415 (6871) (2002) 530–536.
[34] S. Paik, et al., A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer, N. Engl. J. Med. 351 (27) (2004) 2817–2826.
[35] J.S. Parker, et al., Supervised risk predictor of breast cancer based on intrinsic subtypes, J. Clin. Oncol. 27 (8) (2009) 1160–1167.
[36] C. Curtis, et al., The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups, Nature 486 (7403) (2012) 346–352.
[37] J.G. Tate, et al., COSMIC: the catalogue of somatic mutations in cancer, Nucleic Acids Res. 47 (D1) (2019) D941–d947.
[38] Q. Kong, K.O. Lillehei, Antioxidant inhibitors for cancer therapy, Med. Hypotheses 51 (5) (1998) 405–409.
[39] Q. Kong, J.A. Beel, K.O. Lillehei, A threshold concept for cancer therapy, Med. Hypotheses 55 (1) (2000) 29–35.
[40] K. Taguchi, M. Yamamoto, The KEAP1-NRF2 system as a molecular target of cancer treatment, Cancers 13 (1) (2020).
[41] L. Baird, M. Yamamoto, NRF2-Dependent bioactivation of mitomycin C as a novel strategy to target KEAP1-NRF2 pathway activation in human cancer, Mol. Cell Biol. 41 (2) (2021).
[42] M. Komatsu, et al., Homeostatic levels of p62 control cytoplasmic inclusion body formation in autophagy-deficient mice, Cell 131 (6) (2007) 1149–1163.
[43] A. Lau, et al., A noncanonical mechanism of Nrf2 activation by autophagy deficiency: direct interaction between Keap1 and p62, Mol. Cell Biol. 30 (13) (2010) 3275–3285.
[44] Y. Ichimura, et al., Phosphorylation of p62 activates the Keap1-Nrf2 pathway during selective autophagy, Mol. Cell 51 (5) (2013) 618–631.
[45] S. Kageyama, et al., Negative regulation of the Keap1-Nrf2 pathway by a p62/Sqstm1 splicing variant, Mol. Cell Biol. 38 (7) (2018), https://doi.org/10.1128/MCB.00642-17.
[46] R. Team, A language and environment for statistical computing, Computing 1 (2006).
[47] A. Voigt, E. Almaas, Assessment of weighted topological overlap (wTO) to improve fidelity of gene co-expression networks, BMC Bioinf. 20 (1) (2019) 58.
[48] M.C. Oldham, et al., Functional organization of the transcriptome in human brain, Nat. Neurosci. 11 (11) (2008) 1271–1282.
[49] E. Ravasz, et al., Hierarchical organization of modularity in metabolic networks, Science 297 (5586) (2002) 1551–1555.