



Research paper

A distinct class of antioxidant response elements is consistently activated in tumors with NRF2 mutations



Daniel C. Levings^{a,1}, Xuting Wang^{b,1}, Derek Kohlhas^a, Douglas A. Bell^b, Matthew Slattery^{a,*}

^a Department of Biomedical Sciences, University of Minnesota Medical School, Duluth, MN 55812, USA

^b Environmental Epigenomics and Disease Group, Immunity, Inflammation and Disease Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709, USA

ABSTRACT

NRF2 is a redox-responsive transcription factor that regulates expression of cytoprotective genes via its interaction with DNA sequences known as antioxidant response elements (AREs). NRF2 activity is induced by oxidative stress, but oxidative stress is not the only context in which NRF2 can be activated. Mutations that disrupt the interaction between NRF2 and KEAP1, an inhibitor of NRF2, lead to NRF2 hyperactivation and promote oncogenesis. The mechanisms underlying NRF2's oncogenic properties remain unclear, but likely involve aberrant expression of select NRF2 target genes. We tested this possibility using an integrative genomics approach to get a precise view of the direct NRF2 target genes dysregulated in tumors with NRF2 hyperactivating mutations. This approach revealed a core set of 32 direct NRF2 targets that are consistently upregulated in NRF2 hyperactivated tumors. This set of NRF2 “cancer target genes” includes canonical redox-related NRF2 targets, as well as target genes that have not been previously linked to NRF2 activation. Importantly, NRF2-driven upregulation of this gene set is largely independent of the organ system where the tumor developed. One key distinguishing feature of these NRF2 cancer target genes is that they are regulated by high affinity AREs that fall within genomic regions possessing a ubiquitously permissive chromatin signature. This implies that these NRF2 cancer target genes are responsive to oncogenic NRF2 in most tissues because they lack the regulatory constraints that restrict expression of most other NRF2 target genes. This NRF2 cancer target gene set also serves as a reliable proxy for NRF2 activity, and high NRF2 activity is associated with significant decreases in survival in multiple cancer types. Overall, the pervasive upregulation of these NRF2 cancer targets across multiple cancers, and their association with negative outcomes, suggests that these will be central to dissecting the functional implications of NRF2 hyperactivation in several cancer contexts.

1. Introduction

NRF2 is a Cap-n-Collar (CNC) basic leucine zipper (bZIP) transcription factor encoded by the gene *NFE2L2* (*Nuclear Factor, Erythroid 2 Like 2*; hereafter referred to as *NRF2* for simplicity). NRF2's primary role is cytoprotective; specifically, it controls the transcriptional response to reactive oxygen species (ROS) [1]. These chemically reactive species can come from endogenous or exogenous sources, and have the potential to react with all classes of cellular molecules, including DNA. However, ROS, and H₂O₂ in particular, are also used for a range of physiological cell signaling processes [2]. Accordingly, ROS levels must be kept in check but not eliminated: too little and redox signaling is disrupted, too much and cell damage occurs. Damage induced by ROS is associated with a wide range of chronic diseases, from neurodegenerative disease to cancer [3–6], so the cellular responses that limit these species (i.e. antioxidant pathways) must be finely tuned. NRF2 is a central component of one such pathway, and research over the past decade has demonstrated the importance of proper balance in this pathway (reviewed in [1,7–9]).

NRF2 is a deeply conserved and widely expressed regulator of the response to ROS [1]. NRF2 regulates redox-responsive gene expression via its interaction with *cis*-regulatory sequences commonly referred to as antioxidant response elements (AREs), but also referred to as electrophile response elements (EpREs) or CNC-sMAF-binding elements (CsMBEs) [10–13]; for simplicity we will refer to NRF2's target sequence as the ARE. NRF2 does not interact with AREs alone, but as an obligate heterodimer with one of the three small MAF (sMAF) bZIP proteins (MAF-F, MAF-G, MAF-K) [14,15]. The interaction between NRF2-sMAF complexes and AREs drives induction of a host of cytoprotective genes in response to oxidative stress (excess ROS). The ROS-linked inducibility of this system is a result of NRF2's interaction with the inhibitory protein KEAP1 (Kelch-like ECH Associated Protein 1) [16–19]. KEAP1 can bind to cytoplasmic NRF2 and target it for ubiquitination and proteasomal degradation; this inhibitory process keeps NRF2 activity low under basal (low ROS) conditions. However, KEAP1-mediated inhibition of NRF2 is suppressed as ROS levels increase. KEAP1 contains 27 redox-sensitive cysteine residues that together essentially act as an oxidative stress sensor. ROS modify KEAP1's

* Corresponding author.

E-mail address: msslatter@umn.edu (M. Slattery).

¹ Equal Contribution.

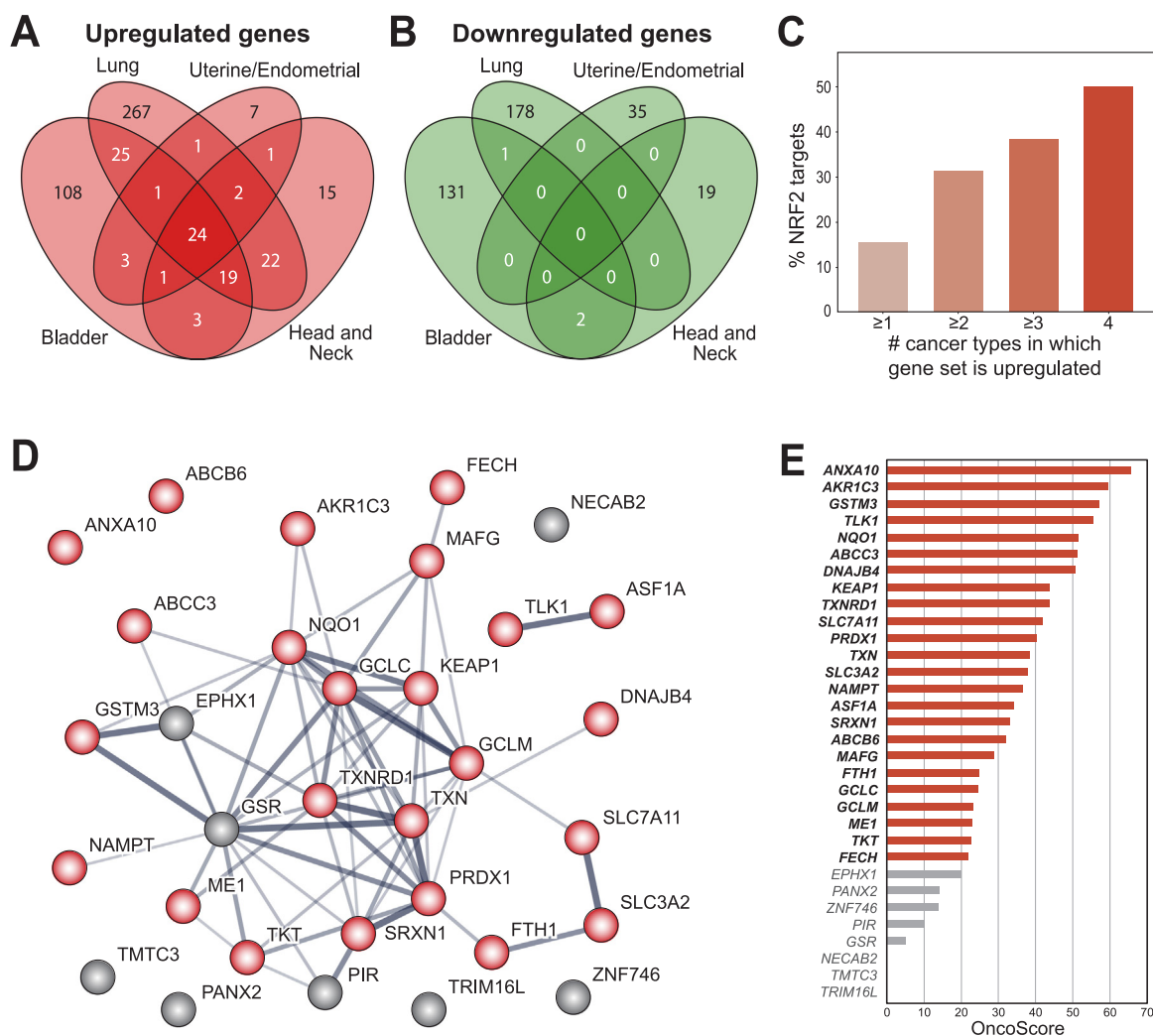


Fig. 1. A subset of NRF2 target genes consistently upregulated in tumors with oncogenic NRF2 mutations. (A) Venn diagram comparing genes called as consistently upregulated with NRF2 hyperactivating mutations in each cancer type, as indicated. Differential gene expression data are from [62]. Out of the 499 genes differentially upregulated: 102 were upregulated in ≥ 2 cancers, 47 were upregulated in ≥ 3 , and 24 were upregulated in all 4 cancers. Cancers listed correspond to the following types from TCGA: Bladder, BLCA; Head and Neck, HNSC; Lung, LUSC; Uterine/Endometrial, UCEC. (B) Same as (A), only for genes consistently downregulated with NRF2 hyperactivating mutations. Out of 366 genes differentially downregulated, only three were shared between more than one cancer type, and none were shared between more than two cancer types. (C) Graph representing the percentage of genes in each overlap category that are called as direct NRF2 targets based on ChIP-seq data. The following fractions of genes were direct NRF2 targets in (x) cancers: (1–4 cancers), 78/499 genes; (2–4), 32/102; (3–4), 18/47; (all 4), 12/24. (D) STRING-based network analysis of the 32 NRF2 targets that are consistently upregulated in at least two cancers. Nodes represent the proteins encoded by the 32 NRF2 target genes, and edge widths are scaled by the strength of evidence supporting an interaction between two nodes. Nodes shaded in red are classified as cancer-associated genes by OncoScore (see (E)). (E) OncoScore [76] values for all 32 NRF2 cancer target genes. At an empirically determined threshold of 21, OncoScore reliably discriminates cancer-associated genes (based on the Cancer Gene Consensus [77]) from non-cancer-associated genes; NRF2 cancer targets with a score above this threshold are highlighted red.

reactive cysteine residues, and this modification prevents KEAP1 from binding newly synthesized NRF2 [20]. Thus, as ROS levels increase, KEAP1 activity decreases, NRF2 is stabilized, and nuclear NRF2 concentration increases. Once in the nucleus, NRF2 can pair with one of the widely expressed SMAD proteins and upregulate ARE-driven target genes [19,21]. The genes targeted by NRF2 in this stress response are varied. Antioxidant genes, proteostasis genes, and intermediary metabolism genes are key NRF2 targets likely to play a central role in helping cells mitigate and recover from oxidative stress [22–25].

The regulatory reach of NRF2 extends far beyond the cytoprotective gene sets described above, however. Notably, NRF2 also directly activates the expression of its own repressor, KEAP1 [24,26]. This negative feedback loop is conserved from fly to human, suggesting tight controls on nuclear NRF2 are an integral part of the regulatory network [24,27]. This NRF2-KEAP1 feedback loop allows for modulation of NRF2-mediated transcriptional activation, presumably ensuring precise

expression of NRF2 target genes across a wide range of redox environments. Although NRF2 activity is cytoprotective [8,28,29], limitation of its activity by KEAP1 is crucial: mutations that impair KEAP1-mediated degradation of NRF2 are associated with tumorigenesis [30–34]. KEAP1 mutations leading to constitutive NRF2 activity were first observed in lung cancer, and cancer-associated NRF2 mutations disrupting KEAP1's inhibition of NRF2 were identified shortly thereafter [30,31,35]. Although these NRF2 hyperactivating mutations are especially prevalent in tumors of the lung, such mutations are also found in a number of other tumor types [34,36–39]. Importantly, coding mutations in NRF2 or KEAP1 are not the only route to constitutive NRF2 activity. For example, in certain tumors the KEAP1 locus is hypermethylated, leading to low KEAP1 expression and high NRF2 activity [40–44], and NRF2 is also indirectly activated via oncogenic KRAS or BRAF signaling [45]. Multiple other proteins that bind KEAP1/NRF2 and/or modulate NRF2 activity have also been identified

Table 1

***NRF2* Cancer Target Genes.** Gene Ontology analysis of the 32 direct *NRF2* target genes, as called by ChIP-seq, that are consistently upregulated in at least two of the four cancer types, as indicated. Differential expression calls are based on data from Araya et al. [62].

Gene	Cancer Types ^a	Redox Balance ^b	Response to Stress/Toxicity ^c
<i>ABCB6</i>	4	–	–
<i>ABCC3</i>	2	–	–
<i>AKR1C3</i>	4	✓	✓
<i>ANXA10</i>	2	–	–
<i>ASF1A</i>	2	–	✓
<i>DNAJB4</i>	2	–	✓
<i>EPHX1</i>	4	–	✓
<i>FECH</i>	4	–	✓
<i>FTH1</i>	2	✓	–
<i>GCLC</i>	3	✓	✓
<i>GCLM</i>	4	✓	✓
<i>GSR</i>	3	✓	✓
<i>GSTM3</i>	2	–	✓
<i>KEAP1</i>	2	–	–
<i>MAFG</i>	2	–	✓
<i>ME1</i>	4	✓	–
<i>NAMPT</i>	3	–	–
<i>NECAB2</i>	2	–	–
<i>NQO1</i>	4	✓	✓
<i>PANX2</i>	4	–	✓
<i>PIR</i>	2	✓	–
<i>PRDX1</i>	3	✓	–
<i>SLC3A2</i>	2	–	–
<i>SLC7A11</i>	3	✓	✓
<i>SRXN1</i>	4	✓	✓
<i>TKT</i>	4	–	–
<i>TLK1</i>	2	–	✓
<i>TMTC3</i>	2	–	–
<i>TRIM16L</i>	4	–	–
<i>TXN</i>	3	✓	✓
<i>TXNRD1</i>	4	✓	✓
<i>ZNF746</i>	2	–	–

^a Number of cancer types in which gene is upregulated (out of 4 total – BLCA, LUSC, UCEC, HNSC).

^b Includes the 'response to oxidative stress' and 'oxidation-reduction process' GO categories (GO:0006979 and GO:0055114).

^c Includes the 'response to stress' and 'response to toxic substance' GO categories (GO:0009636 and GO:0006950).

[46–51]. Considering the diverse molecular changes that can result in hyperactivation of *NRF2*, it is quite plausible that a number of as-yet unidentified mechanisms exist. Ultimately the routes to constitutive *NRF2* activity are varied, and cancer cells make use of the many available options.

Constitutive *NRF2* activity gives cancer cells a selective advantage across a number of organ systems, including the lung, bladder, uterus, and more [34,36–39,52]. The precise mechanisms by which *NRF2* benefits cancer cells are not fully understood, but certainly include chemoresistance, metabolic, and proliferative advantages [25,53–61]. Additionally, because *NRF2* is a transcription factor, many of these advantages are likely due to aberrant expression of its direct target genes. In the work described here, we identify a gene expression signature for oncogenic *NRF2* activity. Specifically, we describe a set of direct *NRF2* target genes consistently upregulated in tumors with oncogenic *NRF2* mutations across multiple organ systems. This functionally important subset of *NRF2* targets includes many of the core cytoprotective genes of the *NRF2*-mediated antioxidant response, and the expression of these *NRF2* “cancer targets” is associated with poor outcome across a number of cancers. Importantly, we determine what mechanistically sets these *NRF2* cancer targets apart from other non-cancer-associated *NRF2* target genes – chiefly, these cancer target loci contain strong AREs in constitutively accessible chromatin, and thus can be turned on in many different cancer-associated tissues.

2. Results

2.1. A gene expression signature associated with oncogenic *NRF2*

A recent analysis of transcriptome data from The Cancer Genome Atlas (TCGA) identified hundreds of gene expression changes that occur with cancer-associated *NRF2* mutations that disrupt the *NRF2*-KEAP1 interaction interface [62]. This work focused on carcinomas from four tissues – bladder urothelial carcinoma (BLCA); lung squamous cell carcinoma (LUSC); head-neck squamous cell carcinoma (HNSC); uterine corpus endometrial carcinoma (UCEC) – and identified genes significantly up- or downregulated in *NRF2*-mutant tumors as compared to non-*NRF2*-mutant tumors of the same cancer type [62]. Importantly, *NRF2* is classified as a cancer driver gene in all four of these cancers [52]. We used this differential cancer gene expression data to identify genes commonly dysregulated across cancers with *NRF2* hyperactivating mutations.

To explore the tissue specificity of gene expression changes in cancers with mutated *NRF2*, we compared the differential gene expression lists from the four cancers analyzed in [62] to identify shared up- or downregulated gene sets. The overlap pattern for the upregulated gene set (Fig. 1A) was noticeably different from the downregulated gene set. Of the 499 genes upregulated in any of the four cancers, approximately 20% (102 genes) were upregulated in at least two of the four cancers, and ~5% (24 genes) were upregulated in four out of four cancers (Fig. 1A). There was very little overlap of downregulated genes (Fig. 1B). Thus, consistent with *NRF2*'s long-recognized role as a transcriptional activator [15,19], oncogenic *NRF2* is associated with a consistent gene activation signature. Overall, this shared expression profile implies that a common set of gene regulatory changes have taken place in this class of *NRF2*-mutated cancers.

The above results highlight a gene set commonly upregulated with oncogenic *NRF2* in cancers derived from a variety of organ systems, yet the data do not provide information indicating a direct role for *NRF2* in regulating these genes. To identify likely direct targets of *NRF2* within the upregulated gene set, we crossreferenced it with a list of *NRF2* target genes based on our own ChIP-seq data (see Methods) [23,63]. For this comparison we used our data from human lymphoblastoid cell line cultures treated with sulforaphane, a dietary isothiocyanate that activates *NRF2* [23,24,64], because it is a robust dataset that is highly concordant with *NRF2* ChIP-seq data from additional conditions and cell lines [23,24,63]. We calculated the percent overlap between the ChIP-seq derived direct *NRF2* targets and various groupings of the cancer upregulated genes (i.e., those upregulated with *NRF2* mutation in 1–4 cancer types, 2–4 cancer types, etc.). We observed a clear trend that the genes upregulated in multiple cancer types were more likely to be direct *NRF2* targets (Fig. 1C). Indeed, almost a third of the genes upregulated in at least two cancer types (32/102) are direct *NRF2* targets, and fully half of all genes upregulated across all four cancer types (12/24) are such *NRF2* targets. The remainder of this study focuses on the 32 direct *NRF2* target genes consistently upregulated in at least two distinct cancer types; we will subsequently refer to these as the *NRF2* “cancer target genes” (Table 1).

2.2. *NRF2* cancer target genes are functionally interconnected and relevant to oncogenesis

Gene ontology (GO) enrichment analysis [65,66] of the *NRF2* cancer targets revealed that most of these genes play a role in maintaining cellular redox balance and/or responding to toxicants and other stressors (Table 1 and data not shown). However, because GO annotations are not truly comprehensive, a number of the genes not characterized as playing role in redox or stress responses in Table 1 have, in fact, been implicated in these processes. For example, the ATP-binding cassette transporters encoded by *ABCB6* and *ABCC3* both play roles in maintaining redox balance [67–70], as do the proteins encoded by

NAMPT, *SLC3A2*, and *TKT* [71–74]. And, as described multiple times above, *KEAP1* is intimately linked to the NRF2 antioxidant pathway; *KEAP1* is transcriptionally regulated by NRF2 as part of a negative feedback loop [24,26], but this feedback mechanism is broken in cancer cells carrying mutations that disrupt the NRF2-KEAP1 interaction interface [30,35]. However, despite the imperfect nature of ontology annotations, GO analysis still reveals a strong enrichment for genes involved in the cellular response to oxidative stress or other toxic stresses.

STRING-based network analysis [75] provides further evidence that the NRF2 cancer target genes encode proteins that are strongly interconnected at a functional level (Fig. 1D). STRING identifies potential functional associations between proteins based on both physical and indirect interactions. In Fig. 1D, putative interactions are indicated by lines (“edges”) connecting the individual proteins (“nodes”); the thickness of each edge is proportional to the confidence that the interaction is robust and biologically meaningful, given the supporting evidence [75]. The NRF2 cancer target network contains 64 edges. Using a background model based on the human proteome, a network of 32 proteins would be expected to contain 5 edges at random. Using a more conservative background model based on all potential NRF2 targets (as measured by ChIP-seq), a network of 32 proteins would be expected to contain 11 edges. Accordingly, based on the conservative background model, the observed/expected ratio for edges within the NRF2 cancer target genes is 5.8 ($p < 0.0001$). Overall, it is clear that the NRF2 cancer target gene network is significantly interconnected.

The above results do not, however, indicate whether this NRF2 cancer target gene set is relevant to oncogenesis. To address this, we also assessed the cancer relevance of the NRF2 cancer target genes using OncoScore, a text-mining tool that discriminates cancer-associated genes from other genes based on citation frequencies from biomedical literature [76]. We used this tool to move beyond lists of genes associated with cancer based on mutation frequencies alone, which can miss many genes that play important functional roles in oncogenesis. Rather, OncoScore uses automated PubMed queries and text-mining algorithms to identify genes enriched for citations in the cancer research literature. Indeed, 75% (24/32) of the NRF2 cancer target genes are called as cancer associated (Fig. 1E), surpassing an OncoScore threshold that reliably discriminates cancer-associated genes (based on the Cancer Gene Consensus [77]) from genes not associated with cancer [76]. This result, combined with manual analysis of the existing literature (see Discussion) suggests a significant fraction of the NRF2 cancer target genes are likely to play an important role in NRF2’s oncogenic potential.

2.3. NRF2 binds strong AREs at its cancer target gene loci

We next asked whether the NRF2 cancer target genes possess any unique features that could explain why they are more broadly responsive to oncogenic NRF2 than other NRF2 target genes. Because NRF2 is a DNA-binding transcription factor, we focused on characterizing ChIP-seq-derived NRF2 binding sites near the cancer target genes and non-cancer target genes (i.e., NRF2 targets not upregulated in 2+ cancer types). At the ChIP-seq peak-calling threshold used for this study, NRF2 has 3122 genome-wide DNA binding sites. Because some loci contain more than one binding site, this translates into 2016 potential direct NRF2 target genes – 32 cancer target genes and 1984 non-cancer target genes.

To explore the relationship between NRF2 binding at its cancer versus non-cancer target genes, we first focused on strength of NRF2’s ChIP-seq binding signals (fold enrichment over background) at these two classes of target genes. ChIP-seq binding reflects the degree to which a transcription factor is bound at a regulatory DNA region, averaged across a population of cells. This signal is often positively correlated with a transcription factor’s regulatory potential at a given locus, although there are exceptions to this pattern [78,79]. For NRF2,

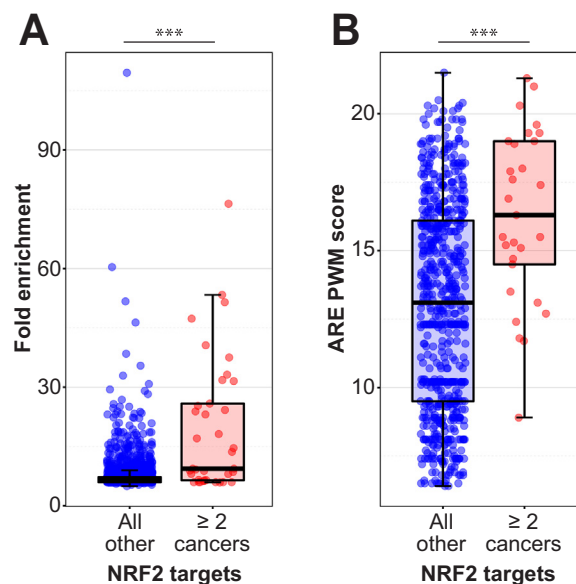


Fig. 2. ChIP enrichment and ARE strength at NRF2 cancer versus non-cancer target genes. (A) ChIP-seq fold enrichment (FE) across the two classes of NRF2 targets. (B) Same setup as (A) only comparing ARE position weight matrix (PWM) scores within ChIP-seq peaks across the two classes of NRF2 targets. For both (A) and (B) boxplots represent the 25th, 50th (median) and 75th percentiles of the data, whiskers show the range of the data minus outliers, and each dot represents a single NRF2 target. (***) $p \leq 0.001$.

we found that its cancer target genes are generally associated with stronger NRF2 binding (median fold enrichment of 9.38; mean 19.64) relative to its non-cancer target genes (median fold enrichment of 6.46; mean 7.35) (Fig. 2A). A similar pattern was observed when we looked at the ARE sequences within NRF2 binding sites near its cancer and non-cancer target genes (Fig. 2B). For this comparison, we used position weight matrix (PWM) scores to quantify how closely a given region’s strongest ARE matched the consensus ARE sequence; the NRF2 cancer target gene AREs were significantly stronger than the AREs at the non-cancer targets (median PWM score of 16.3 versus 13.1). Thus, one difference between the NRF2 cancer and non-cancer targets is that the cancer gene enhancers tend to be regulated by strong ARE sequences that are associated with robust NRF2 binding.

2.4. NRF2 cancer target AREs reside in constitutively permissive chromatin environments

Despite the fact that the 32 cancer target genes tend to be associated with strong NRF2 binding, a number of non-cancer target genes are also associated with strong NRF2 binding (orange box in Fig. 3A). We next focused on this set of robust non-cancer NRF2 targets – the top 200 based on ChIP-seq fold enrichment values – as a reference for comparison with the NRF2 cancer targets. This is a more stringent comparison than the one described in the previous section, but our intention remained the same: to identify features that might explain why the NRF2 cancer targets are more broadly responsive to oncogenic NRF2. This more stringent comparison revealed no significant differences in ChIP-seq enrichment (Fig. 3B) or ARE PWM scores (Fig. 3C) between NRF2’s cancer targets and the top 200 non-cancer targets. Thus, ARE sequence and binding strength alone are not sufficient to explain the broad responsiveness of NRF2’s cancer target genes.

If the NRF2 bound AREs are equivalent at the cancer targets and the strongest non-cancer targets, what other regulatory features at the cancer AREs could allow for their activation across organ systems? Transcription factor interactions with DNA take place in the context of chromatin, and the local chromatin environment at a transcription

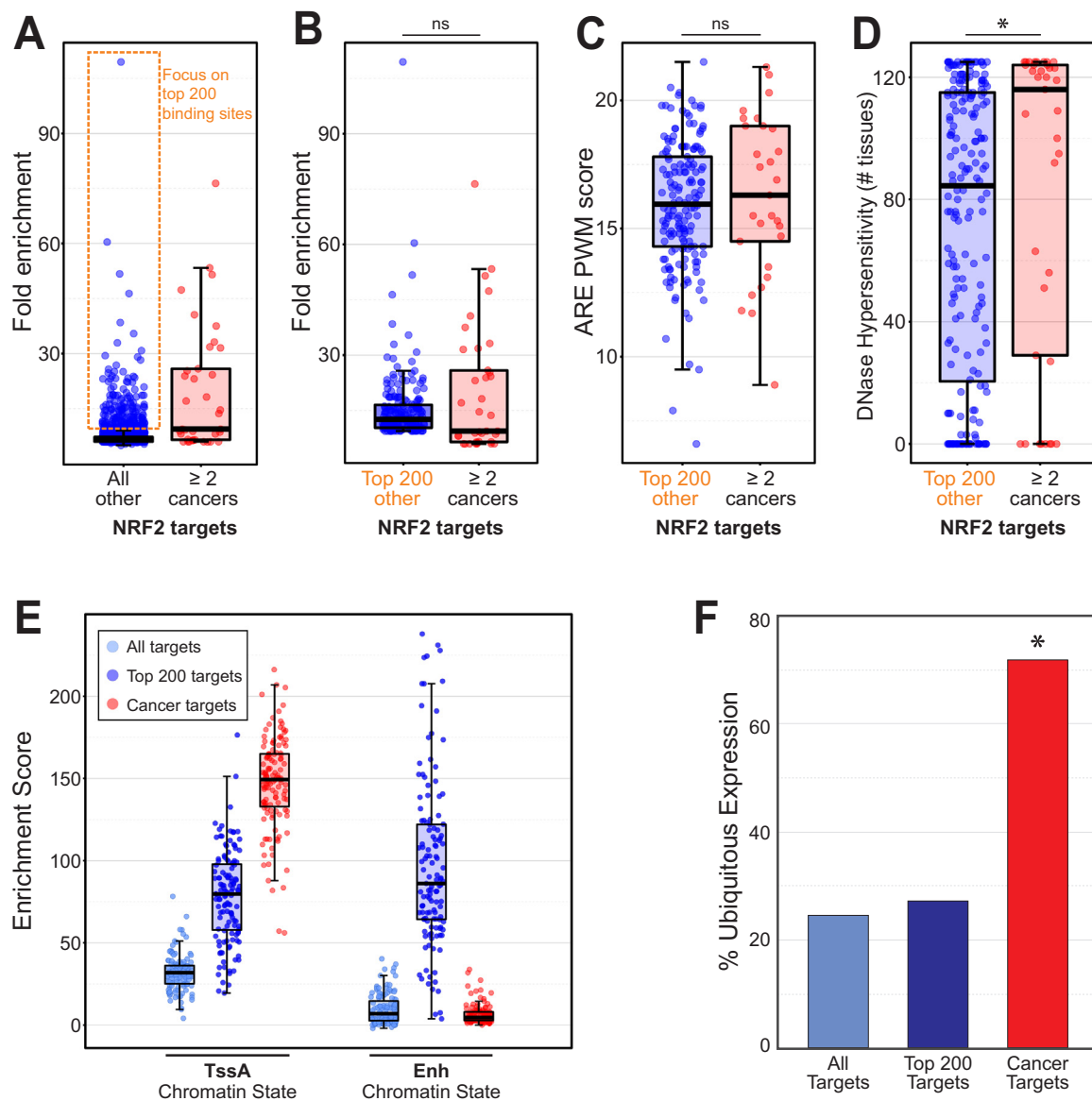


Fig. 3. ChIP enrichment, ARE strength, and chromatin state properties at NRF2 cancer targets and strong non-cancer targets. (A) The top 200 non-cancer NRF2 targets based on ChIP-seq fold enrichment (shown within the red rectangle on the left) were separated from the full set of non-cancer targets and compared to the cancer targets. (B) ChIP-seq fold enrichment (FE) across the top 200 non-cancer targets and NRF2 cancer targets. (C) Same setup as (B) only comparing ARE position weight matrix (PWM) scores within ChIP-seq peaks across the top 200 non-cancer targets and NRF2 cancer targets. (D) A comparison of DNase hypersensitivity patterns across the top 200 non-cancer targets and NRF2 cancer targets. The y-axis indicates the number of cell lines, out of a total of 125, in which the NRF2 binding site overlapped with a DNase I hypersensitive site (data from [85]). (E) NRF2 peaks were aligned to the 15-state chromatin state maps for 127 cell lines and primary cells/tissues from Roadmap Epigenomics using GIGGLE [89,90]. The computed enrichment scores indicate the degree and significance of overlap between the various classes of NRF2 binding sites (all non-cancer, top 200 non-cancer, and cancer targets) and each chromatin state for a given cell/tissue. Each dot represents the enrichment score for an individual cell or tissue type. NRF2 cancer targets predominantly lie in highly active transcription start site regions (TssA; H3K4me3-associated), and non-cancer NRF2 targets tend to lie in enhancer regions (Enh; H3K4me1-associated), so data for those two chromatin states are presented here. Data for all 15 chromatin states are presented in Fig. S3. (F) Percentage of NRF2 target genes in each category that overlap ubiquitously expressed genes as defined by [131]. For (A) through (E), boxplots represent the 25th, 50th (median) and 75th percentiles of the data, and whiskers show the range of the data minus outliers (* $p \leq 0.05$; ns, non-significant).

factor binding site can have a significant effect on binding and regulatory output [79–83]. With this in mind, we next tested whether AREs at the NRF2 cancer target genes were more likely to fall within permissive chromatin environments.

We first focused on DNA accessibility. Nuclear DNA is associated with nucleosomes, and this interaction can act as an impediment, making regulatory DNA elements inaccessible to transcription factors. Nucleosome positioning is a regulated process that can constrain or allow transcription factor access to DNA in a cell-type specific manner, and it can be monitored genome-wide using DNase-seq (DNase I

hypersensitive site sequencing) [84,85]. DNase-seq is based on the fact that nucleosome-associated inaccessible DNA is protected from cleavage by DNase I, whereas more accessible DNA presents as DNase hypersensitive sites (DHS). The ENCODE (Encyclopedia of DNA Elements) consortium has generated genome-wide DHS maps for 125 human cell lines and primary cells [85]; we used these data to test whether NRF2 binding sites at its cancer target loci are more broadly accessible than its non-cancer-associated binding sites. Indeed, in comparison to the NRF2 cancer targets, the full list of non-cancer-associated NRF2 bound regions had a significantly higher fraction of targets with a DHS score of

0, meaning these NRF2 binding regions are inaccessible in most cell types (Supplemental Fig. 1B). Importantly, this pattern also held true when comparing cancer-associated NRF2 bound regions to the top 200 non-cancer-associated NRF2 bound regions – the cancer-associated regions had a significantly higher median DHS score than the top 200 non-cancer-associated NRF2 bound regions (Fig. 3D). In fact, the small number of cancer-associated NRF2 bound regions that are not broadly accessible are either: (1) weaker secondary binding events at a locus with a separate strong, accessible NRF2 binding region, or (2) fall within repetitive DNA elements (SINEs or LINEs), which can be associated with nonconsensus transcription factor binding and/or present with sequencing mappability issues that can lead to false positive or negative results in various -seq assays (Supplemental Fig. 2) [86–88]. Thus, with a few explainable exceptions, these results support a model in which NRF2's cancer target genes are activated in multiple organ systems because strong AREs located in constitutively accessible regulatory DNA regions drive their expression.

A comparison of NRF2 binding sites with chromatin state data from the Roadmap Epigenomics Consortium provides further support for the model proposed above. The Roadmap project used ChIP-seq data for a set of five histone modifications profiled across 127 diverse tissues and cell types (including > 70 primary cell or tissue types and > 25 primary cultures) to generate a high resolution view of cell-specific chromatin states across the human genome [89]. In total, 15 chromatin states were defined in the chromatin state model we used. To determine if there is a correlation between chromatin state and NRF2 binding at its cancer-associated and non-cancer-associated targets, we used GIGGLE [90] to calculate the significance of overlap between the different classes of NRF2 binding and the 127 cell- and tissue-specific chromatin state models (Fig. 3E and Supplemental Fig. 3). In general, NRF2 binding occurs in chromatin states associated with *cis*-regulatory DNA across all cell types, particularly in: (1) TssA, which is characterized by high levels of histone H3 lysine 4 trimethylation (H3K4me3); (2) TssAFlnk, which is characterized by a combination of H3K4me3 and histone H3 lysine 4 monomethylation (H3K4me1); and (3) Enh, which is characterized by high levels of H3K4me1. However, NRF2's chromatin state preferences are not uniform across its cancer-associated and non-cancer-associated binding sites (Fig. 3E and Supplemental Fig. 3). Cancer-associated NRF2 binding is strongly enriched in the TssA (H3K4me3) chromatin state across all 127 cell types, whereas the non-cancer-associated NRF2 binding is far less TssA-enriched across cell types (Fig. 3E). On the other hand, non-cancer-associated NRF2 binding is uniquely enriched in the Enh (H3K4me1) chromatin state (Fig. 3E). Importantly, although both H3K4me1 and H3K4me3 mark regulatory DNA regions, H3K4me1 does not indicate whether an enhancer or promoter is active; only H3K4me3 is positively correlated with *active* promoter or enhancer regions [91–93], and regions classified as TssA have more permissive *cis*-regulatory activity than Enh regions based on massively parallel reporter assays [94]. Thus, cancer-associated and non-cancer-associated NRF2 binding sites correlate with distinct chromatin landscapes, with the cancer-associated NRF2 binding sites more strongly linked to ubiquitously permissive chromatin regions.

Multiple results described thus far indicate that the regulatory environments at NRF2's cancer target gene loci are more permissive than at NRF2's other target genes. Consistent with this, we found that the NRF2 cancer target genes are also more likely to be broadly expressed than the non-cancer targets (Fig. 3F).

A comparison of four classical NRF2 target genes highlights the key difference between its cancer target genes and its other targets. *NQO1*, *GCLC*, *GCLM*, and *HMOX1* (also known as *HO-1*) are all well-established, highly responsive NRF2 target genes [1], however, only *NQO1*, *GCLC*, and *GCLM* are called as NRF2 cancer targets (see Table 1). Indeed, a closer look at expression of these four classic NRF2 targets in tumors with no *NFE2L2* mutation versus tumors with an *NFE2L2* mutation reveals that *HMOX1* is not as consistent in its activation by *NFE2L2* mutation (Fig. 4A). While *NQO1*, *GCLC*, and *GCLM* induction is

strongly significant in almost all cancer contexts, *HMOX1* induction is robust (Wilcoxon p-value < 0.01) only in LUSC. *HMOX1* induction is nominally significant in HNSC (Wilcoxon p-value < 0.05), and is not significant in UCEC or BLCA. The dominant ARE sequences within the two high confidence NRF2 binding sites at *HMOX1* are just as strong as the AREs at the other three genes, so this discrepancy in *HMOX1* expression cannot be explained by ARE sequence (Fig. 4B). Similarly, ChIP enrichment signals at the two NRF2 binding sites near *HMOX1* are as strong as, or stronger than, binding at *NQO1*, *GCLC*, or *GCLM* (Fig. 4C). Thus, ARE sequence and NRF2 binding strength are not sufficient to explain the less consistent induction of *HMOX1* by oncogenic NRF2.

A far different pattern emerges when we look at the dominant chromatin states at the NRF2 binding sites near *NQO1*, *GCLC*, *GCLM*, and *HMOX1*. As described above, the Roadmap Epigenomics project has generated high-resolution maps of chromatin states across 127 human tissues and cell types [89]. We used these data to identify the dominant chromatin state at each of these NRF2-targeted AREs across all 127 tissues/cells (Fig. 4D). NRF2 binding at the three cancer-associated genes (*NQO1*, *GCLC*, or *GCLM*) falls in regions classified as TssA or TssA-flanking (permissive, H3K4me3-associated) chromatin states in essentially all 127 cell/tissue types profiled (Fig. 4D). The NRF2 binding sites near *HMOX1*, however, fall within regions with substantial variation in chromatin state across the 127 cell/tissue types (Fig. 4D). For example, the *HMOX1* enhancer with the most robust NRF2 binding (labeled *HMOX1*^{Enh1} in Fig. 4) is classified as TssA or TssA-flanking in < 50% (55/127) of the Roadmap-profiled samples, and is classified as Enh (less permissive, H3K4me1-associated) in > 30% (42/107) samples. In fact, the *HMOX1*^{Enh1} region is even classified as quiescent chromatin in 11 samples. Therefore, these data support a model where *NQO1*, *GCLC*, and *GCLM* are broadly responsive to oncogenic NRF2 because their AREs fall within ubiquitously permissive chromatin; *HMOX1* is less broadly responsive because its AREs fall within chromatin environments that are only permissive in select cell or tissue types.

Taken together, these analyses suggest that oncogenic NRF2 is able to upregulate its cancer target genes in a variety of cell types because these genes contain fewer cell-specific repressive inputs, particularly in the chromatin environment surrounding their ARE-containing regulatory DNA regions.

2.5. High expression of NRF2 cancer target genes is associated with NRF2 pathway variation

Since the NRF2 cancer target genes are consistently upregulated by oncogenic NRF2, and their AREs fall within permissive chromatin in most cell types, we reasoned that they could serve as proxies for NRF2 activation status beyond the four tumor types addressed above. This is an important point because identifying tumors with hyperactive NRF2 is not as simple as looking for mutations in NRF2 – there are many routes to NRF2 activation [37]. Thus, we used RNA-seq data from TCGA to infer NRF2 activity across > 9000 tumors based on mean expression of the 32 NRF2 cancer target genes (Fig. 5). The tumor populations of at least ten different cancer types show a clear positively skewed or bimodal distribution, indicative of a significant tumor subset with higher than average inferred NRF2 activity. Cancer types with this profile include the four discussed so far (LUSC, HNSC, BLCA, UCEC), as well as liver hepatocellular carcinoma (LIHC), esophageal carcinoma (ESCA), lung adenocarcinoma (LUAD), kidney renal papillary cell carcinoma (KIRP), cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), and kidney renal clear cell carcinoma (KIRC). In addition, although not as dramatic, additional cancer types also have distributions that are positively skewed (e.g. breast invasive carcinoma, or BRCA). These patterns of inferred NRF2 activity imply that NRF2 hyperactivation is common in a range of carcinomas and adenocarcinomas.

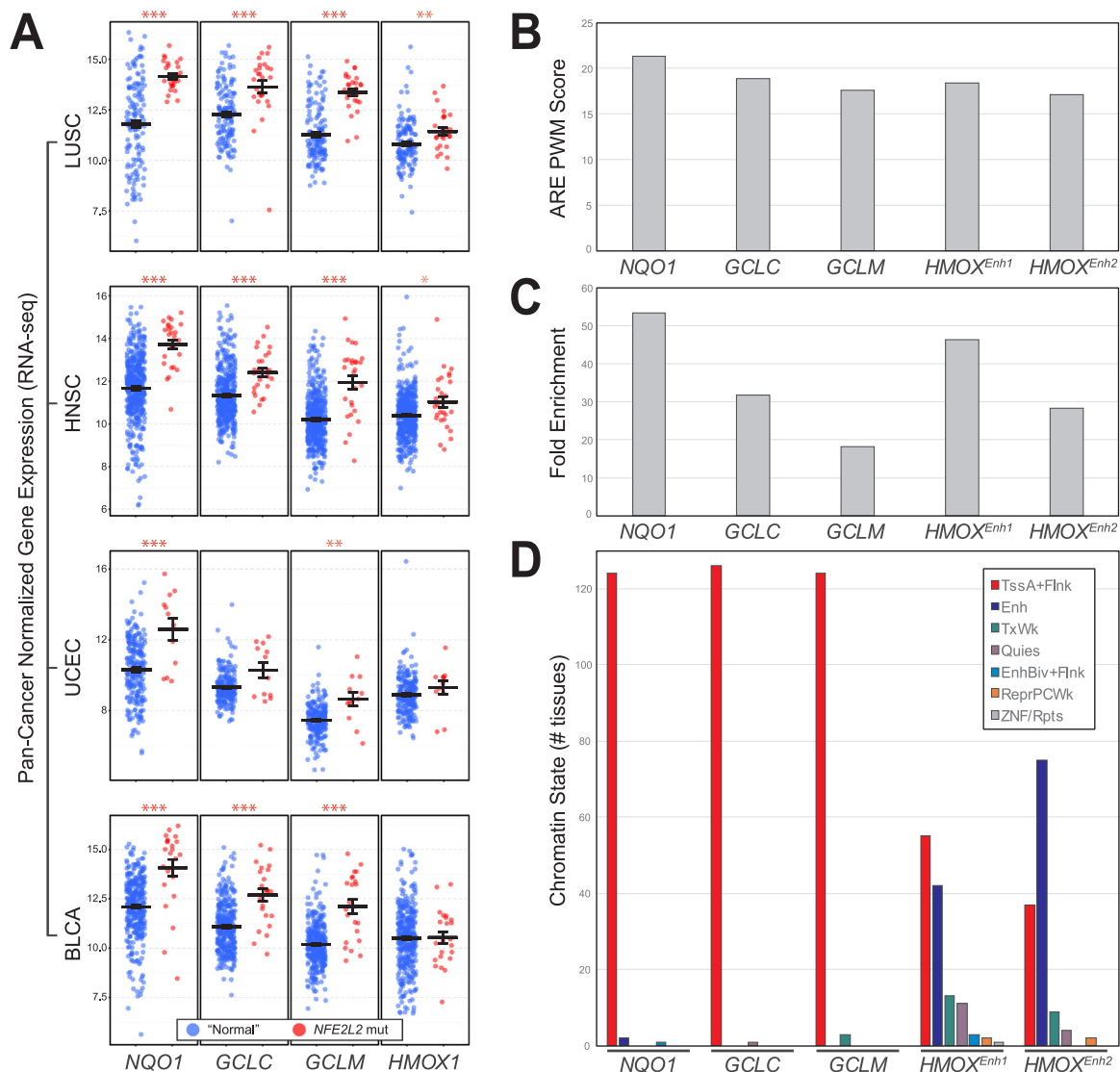


Fig. 4. Gene expression, NRF2 binding, and chromatin state properties at *NQO1*, *GCLC*, *GCLM*, and *HMOX1*. (A) Gene expression values (RNA-seq) from TCGA for three NRF2 cancer target genes (*NQO1*/*GCLC*/*GCLM*) and a strong, non-cancer target gene (*HMOX1*) in the four cancers described in Fig. 1. For each cancer, tumors with mutated *NRF2/NFE2L2* were separated from tumors with normal *NRF2/NFE2L2*; each dot represents a single tumor. Horizontal lines represent the mean (\pm SE) expression for each gene. (***) $p \leq 0.001$; ** $p \leq 0.01$; * $p \leq 0.05$ (B) ChIP-seq fold enrichment (FE) across NRF2-bound enhancers at *NQO1*, *GCLC*, *GCLM*, and *HMOX1*. *HMOX1* is associated with two NRF2-bound enhancers, represented here as *HMOX1^{Enh1}* and *HMOX1^{Enh2}*. (C) Same setup as (B) only comparing ARE position weight matrix (PWM) scores at *NQO1*, *GCLC*, *GCLM*, and *HMOX1*. (D) As in Fig. 4E, NRF2 peaks were aligned to the 15-state chromatin state maps for 127 cell lines and primary cells/tissues from Roadmap Epigenomics using GIGGLE [89,90]. NRF2 peaks were assigned a single chromatin state (based on largest genomic overlap) for each cell/tissue, and these assignments were tallied for all 127 cells/tissues. The y-axis indicates the number of cell/tissues, out of a total of 127, in which the NRF2 binding site overlapped with the indicated chromatin state (each bar represents a chromatin state, as indicated in graph legend).

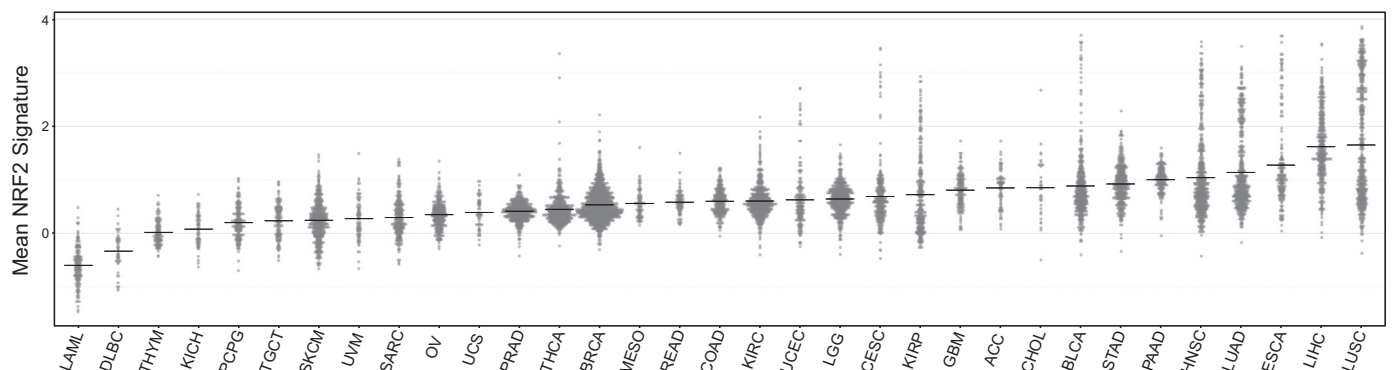


Fig. 5. Inferred NRF2 activity across cancer types. Inferred NRF2 activity based on mean transcript expression of the 32 NRF2 cancer target genes across > 9000 tumors profiled by The Cancer Genome Atlas. Each dot represents an individual tumor, and black lines represent the mean value for each cancer type.

The data presented in Fig. 5 are mutation agnostic – NRF2 activity is inferred based on gene expression data alone – so it is not clear which mutations are driving high NRF2 activity. Thus, although the NRF2 cancer target genes were identified based on changes associated with NRF2 mutation in a limited number of tumors, our inferred activity data might identify additional genetic alterations driving high NRF2 activity. To test this possibility, we used a receiver operating characteristic (ROC) based approach to ask which cancer driver mutations [52] are enriched in tumors with high NRF2 activity. First, for each cancer type a mean inferred NRF2 activity value was determined. Cancer-specific mean inferred NRF2 activity values were then used to normalize the inferred NRF2 activity for each tumor across all cancers, thus removing cancer specific differences in basal NRF2 activity (see Supplemental Fig. 4). Next, all tumors were ranked in decreasing order of inferred NRF2 activity. We screened all 439 cancer driver genes [52] for positive association with inferred NRF2 activity by calculating the area under a ROC curve (AUC) (see Supplemental Fig. 4). To calculate the AUC for a given driver gene, ranked tumors were called ‘positive’ if they had a mutation in the gene, and ‘negative’ if they had no mutation in the gene. For this analysis, AUC values closer to 1 correspond to enrichment (high NRF2 activity is a strong classifier for tumors with a mutation in given gene), 0.5 corresponds to no enrichment (random relationship between NRF2 activity and mutation), and values < 0.5 correspond to inverse relationships (low NRF2 activity identifies tumors with a mutation in given driver gene). This approach essentially asks whether inferred NRF2 activity alone is a strong classifier for a given mutation. We limited our analysis to driver genes that were mutated in at least 50 of the 7679 tumors with both RNA-seq data and non-silent mutation data. Using this method, the top three genes with mutations associated with high NRF2 activity were *NFE2L2/NRF2* (AUC = 0.843), *KEAP1* (AUC = 0.826), and *CUL3* (AUC = 0.609) (Fig. 6A, solid lines). Importantly, in addition to mutations, *NFE2L2/NRF2* amplification (copy number gain) and *KEAP1* or *CUL3* deletion (copy number loss) are also associated with high inferred NRF2 activity (Fig. 6A, dashed lines). Together, mutations in these three genes can account for many of the “high NRF2” tumors in multiple cancer types (Fig. 6B). Ultimately, considering it captures multiple mutations and copy number variations known or expected to alter NRF2 activity, these results suggest this gene set is a reliable NRF2 pathway signature.

2.6. High expression of NRF2 cancer target genes is associated with poor cancer prognosis

The work described in the previous sections implies that the NRF2 cancer target genes (1) are potentially important mediators of NRF2-driven oncogenesis, and (2) could serve as proxies for NRF2 activation status in a tumor. Thus, considering the functional relevance and biomarker potential of this gene set, we next asked whether inferred NRF2 activity, based on expression of the NRF2 cancer target genes, is associated with significant differences in survival across 29 cancers profiled by TCGA. First, for each cancer type, we used expression data to group tumor samples according to the expression levels of the 32 NRF2 cancer target genes. Using hierarchical clustering, we divided the tumors for each cancer type into “normal NRF2” and “high NRF2” groups, and compared the long-term survival outcomes for the patients from which the tumors were obtained (Fig. 7 and Supplemental Fig. 5). A total of 10 cancers allowed for clear clustering of tumors into groups of high and normal NRF2 cancer target gene expression. This included the four cancer types that were used to generate the NRF2 signature (BLCA, LUSC, HNSC, and UCEC), as well as kidney renal papillary cell carcinoma (KIRP), liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), esophageal carcinoma (ESCA), cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), and testicular germ cell tumors (TGCT). Within this overall group of 10, there was a distinct trend toward high NRF2 activity corresponding to decreased survival (Fig. 7 and Supplemental Fig. 5A), and four cancer types –

BLCA, HNSC, KIRP, and LIHC – exhibited a significantly lower overall survival in the high NRF2 group. No cancers showed a significant increase in overall survival with high NRF2 activity.

Although clustering-based separation of tumors of a given cancer type into high and normal NRF2 groups is ideal for these analyses, this was not possible for more than half of the TCGA profiled cancers. Therefore, to further verify our clustering-based survival analysis, we used a simple thresholding method to separate the tumors for each cancer type profiled by TCGA. Briefly, for each cancer type, we first ordered the tumors based on average NRF2 cancer target gene expression, and then used a 90th percentile cut-off to split them into normal (< 90th percentile) and high (\geq 90th percentile) NRF2 activity. We then looked for differences in overall survival between the high and low NRF2 tumor sets. Despite the weaknesses inherent in this thresholding approach – high NRF2 tumors will be placed in low NRF2 category in cancers with many (> 10%) high NRF2 tumors – this analysis was consistent with the trends in the hierarchical clustering-based analysis. Again, high expression of the NRF2 cancer target genes is primarily associated with decreased overall survival (Supplemental Fig. 6), and 11 cancer types showed a significant decrease in overall survival in the high NRF2 group using this method. Three cancer types – KIRP, BLCA, and LIHC – were associated with significant decreases in overall survival using both methods, and the clustering-based results for these three cancer types are shown in Fig. 7. Overall, this work implies that, at least for certain cancers, high expression of the NRF2 cancer target genes is associated with significantly shortened overall survival times.

3. Discussion

Dysregulation of transcription factor function is a common occurrence in cancer [52,95–97]. In order to understand the role of cancer-associated transcription factor dysregulation, one must have an accurate view of the factor’s target genes, as these encode the effector proteins that ultimately drive oncogenic outcomes. Here, we focused on the transcription factor NRF2, which is often hyperactivated in cancer via mutations that disrupt the NRF2-KEAP1 interaction interface [1,7,30,35]. In fact, based on these mutations, *NFE2L2* (which encodes NRF2) is considered a cancer driver gene in several cancer types, including carcinomas of the lung, bladder, head/neck, and uterus/endometrium; and *KEAP1* is classified as a cancer driver in both lung carcinoma and adenocarcinoma [34,38,52,98–100]. To gain insight into the core oncogenic NRF2 regulatory network, we used an integrative genomics approach that combined our own ChIP-seq data with TCGA gene expression data to identify direct NRF2 target genes consistently upregulated by oncogenic NRF2 across multiple cancer contexts. Below we discuss what makes these NRF2 cancer target genes unique from a gene regulation standpoint, as well as the implications of this gene set for understanding NRF2’s role in oncogenesis.

Like most transcription factors, NRF2 binds to thousands of genomic regions based on ChIP-seq experiments. At the relatively stringent peak calling threshold used in this study, NRF2 has approximately 3100 genome-wide binding sites, and these binding sites map to approximately 2100 unique genes (based on the nearest transcription start site to a peak). Thus, NRF2 could potentially alter expression at a considerable number of loci. However, when we compared these potential NRF2 targets to gene expression changes associated with oncogenic NRF2 mutation in carcinomas of the lung, bladder, head/neck, or uterus/endometrium [62], only 32 direct target genes were upregulated by NRF2 in at least two cancer types. These 32 genes tend to be robust NRF2 targets, in that NRF2 binding at these loci is strong based on ChIP-seq fold enrichment, and these binding events are associated with sequences that are strong matches to the ARE consensus. Yet many other NRF2 target genes (e.g. *HMOX1*) share these features but fail to be consistently upregulated with oncogenic NRF2 mutations; thus, ARE sequence and NRF2 binding strength alone do not mechanistically

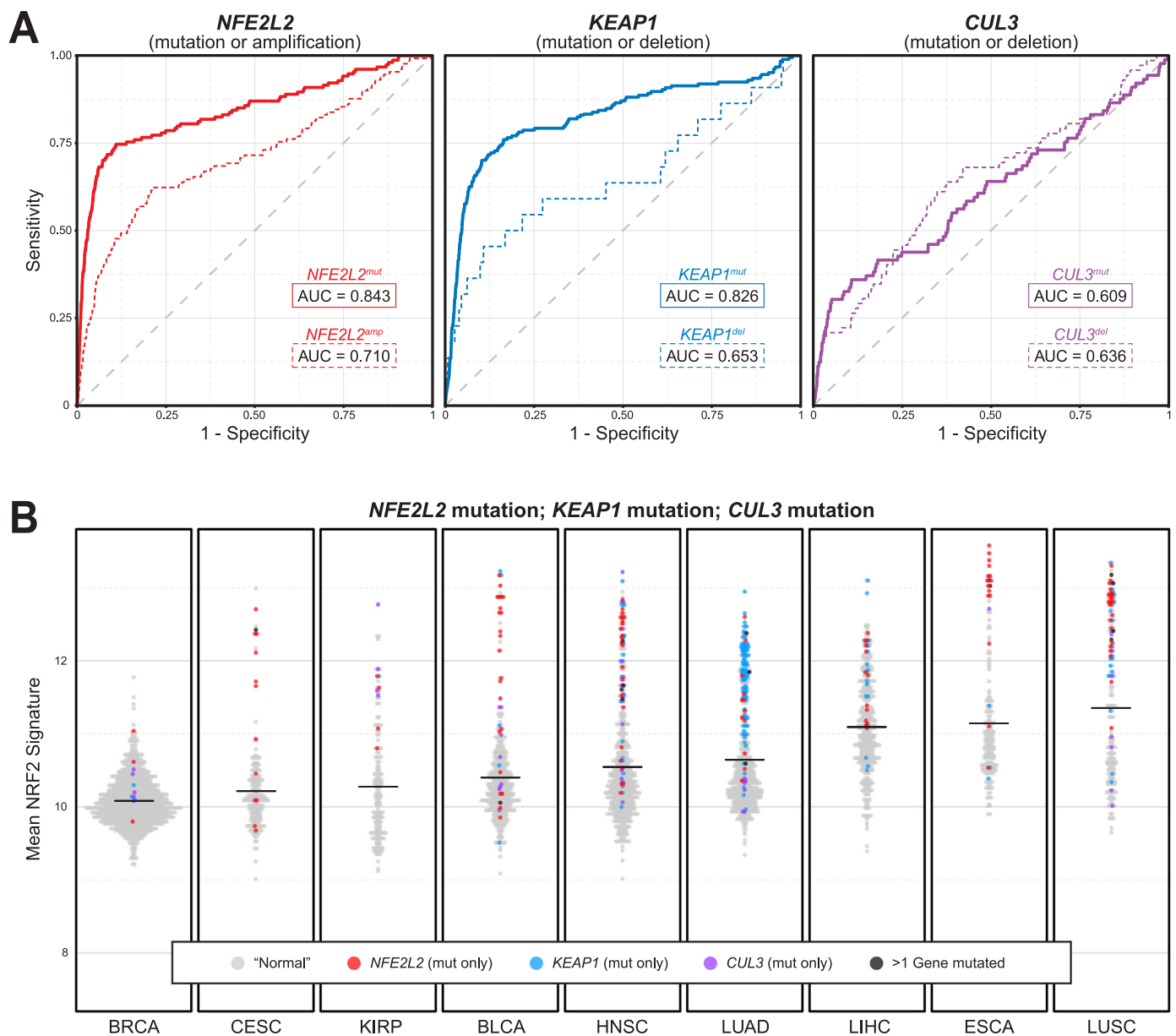


Fig. 6. Inferred NRF2 activity and mutation or copy number variation in the NRF2-KEAP1-CUL3 axis. (A) Receiver operating characteristic (ROC) curves for the classification of mutation or copy number variation using pan-cancer inferred NRF2 activity. Left panel: *NFE2L2/NFE2L2* non-silent mutation (solid red line) or amplification (dashed red line). Center panel: *KEAP1* non-silent mutation (solid blue line) or deletion (dashed blue line). Right panel: *CUL3* non-silent mutation (solid purple line) or deletion (dashed purple line). Area under the curve (AUC) values are also represented for each mutation/amplification/deletion. (B) Inferred NRF2 activity for a subset of cancer types is presented in the same manner as in Fig. 5. Tumors with mutations in *NFE2L2/NRF2* (red), *KEAP1* (blue), or *CUL3* (purple) are highlighted for the indicated cancer types.

differentiate NRF2-targeted cancer AREs from its non-cancer AREs. Instead, what differentiates the cancer AREs from the strong, non-cancer AREs is the chromatin environment in which the AREs are located. Cancer AREs are located in regions of the genome that are characterized by high DNA accessibility and permissive chromatin signatures in the vast majority of cell and tissue types tested, whereas equivalent non-cancer AREs fall in regions with more complex chromatin environments that vary by tissue. It is not yet clear which regulatory factors establish and maintain the unique, broadly permissive chromatin environment at the cancer AREs. Nor is it clear which factors regulate the restrictive chromatin environment at strong, non-cancer AREs, although BACH1 is one candidate. BACH1 is a repressive transcription factor that can also partner with small MAF proteins to bind ARE sequences, and it plays a role in context-specific repression of *HMOX1* [101]. Importantly, BACH1 can also recruit the chromatin

remodeling factor CHD8 and the DNA methyltransferase DNMT3B to DNA, so it has the potential to create a repressive chromatin environment at ARE sequences [102]. However, it is also likely that transcription factors targeting non-ARE sequences can modulate NRF2 activity at many of its binding sites, so there may be distinct mechanisms limiting the activity of individual non-cancer AREs. Nevertheless, regardless of the mechanism(s) restricting non-cancer ARE activity, our work supports a model in which cancer AREs are unique: at the cancer AREs, NRF2 can bind and elicit a transcriptional response in almost all cellular contexts because these AREs are not subject to the same tissue-specific regulatory constraints as most other AREs.

One implication of the NRF2 cancer target gene set's potential broad responsiveness to NRF2 is that it will serve as a reliable proxy for NRF2 activation status in many cellular contexts. We therefore used this expression signature to infer NRF2 activity across thousands of TCGA-

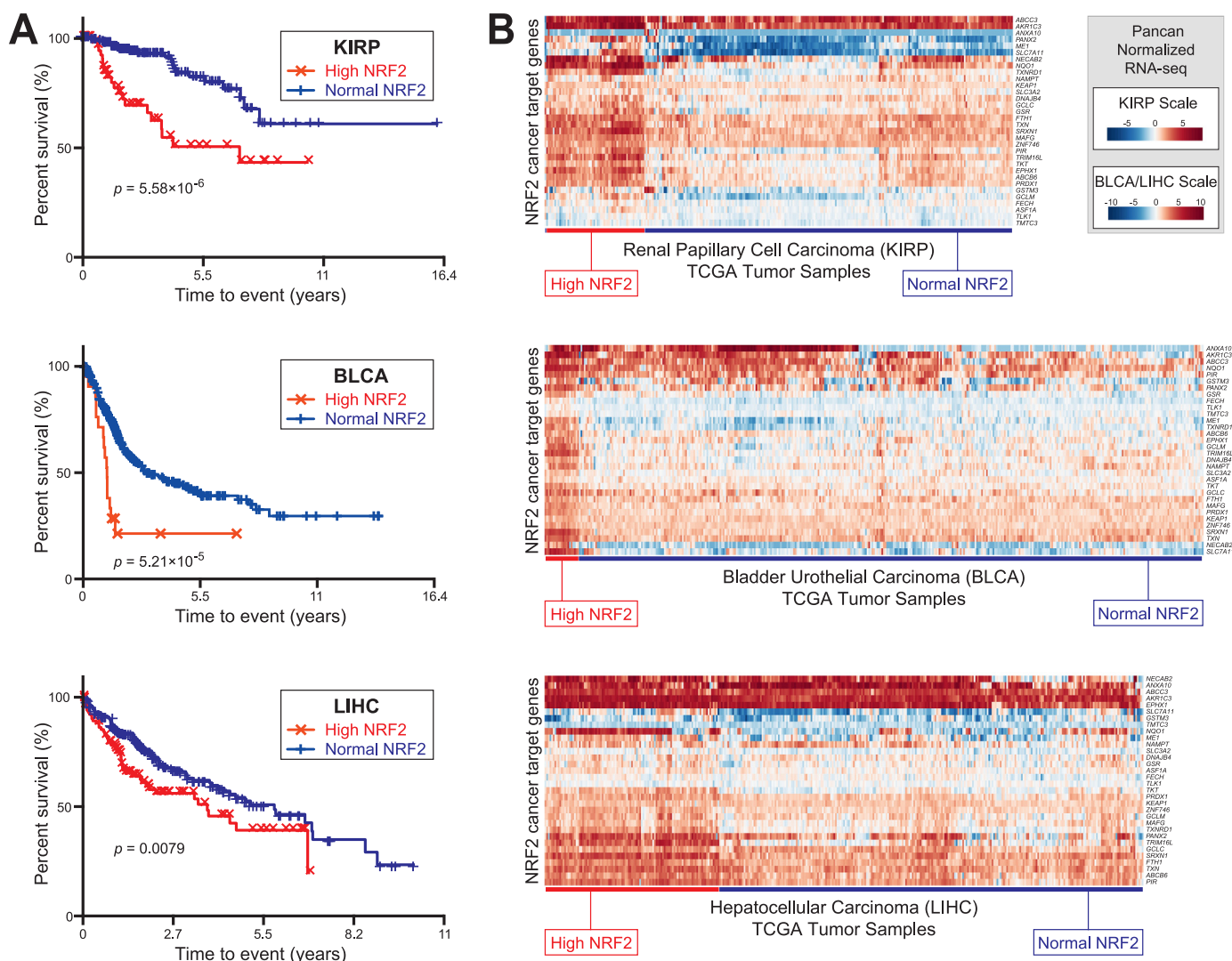


Fig. 7. Survival outcomes associated with high inferred NRF2 activity. (A) Kaplan-Meier survival analysis for three cancer types – KIRP (top), BLCA (middle), and LIHC (bottom) – where high expression of NRF2 target genes is associated with decreased overall survival. (B) Hierarchical clustering of KIRP (top), BLCA (middle), and LIHC (bottom) tumors based on expression of NRF2 cancer target genes. Clusters were annotated as “High NRF2” and “Normal NRF2” groups as indicated, and these classifications were used for survival analysis in (A).

profiled tumors, and test whether high NRF2 tumors are associated with the expected NRF2 pathway mutations. Indeed, the top three mutated genes identified were *NFE2L2*, *KEAP1*, and *CUL3* – all central players in the NRF2 signaling pathway. As discussed previously, mutations that disrupt NRF2-KEAP1 interaction lead to constitutive NRF2 activity [1,7]. *CUL3* encodes the ubiquitin ligase that KEAP1 uses to target NRF2 for proteasomal degradation [103–105]. Mutations in *CUL3* have previously been linked to NRF2 activation in papillary renal cell carcinoma [106], and our analysis confirms this, although we also see *CUL3* mutation associated with high NRF2 activity in breast, esophageal, and head/neck cancer (see Fig. 6B). Importantly, we also found copy number variation in the NRF2-KEAP1-CUL3 axis is linked to high NRF2 activity: *NRF2* copy number gain (amplification) and *KEAP1* or *CUL3* copy number loss (deletion) are all associated with high NRF2 activity. This is consistent with the prevailing model where gain of function variation in *NRF2* and loss of function variation in *KEAP1* or *CUL3* drive high NRF2 activity, and highlights the potential broad oncogenic impact of copy number variation across the NRF2-KEAP1-CUL3 axis.

The NRF2 cancer target genes are also useful for exploring the consequences of NRF2 hyperactivation. The genetics described above demonstrate that our inferred NRF2 activity is mutation-agnostic – a

range of mutations or copy number variants that increase NRF2 activity are captured – so it allows for a more inclusive look at the clinical implications of hyperactive NRF2. In general, when comparing high NRF2 tumors (top 10%) to all other tumors of a given cancer type, high inferred NRF2 activity is associated with poor survival, with significant decreases in overall survival in 11 cancer types. Further, ten cancer types have expression profiles that result in unambiguous clustering of tumors into groups of high and normal NRF2 cancer target gene expression, thus bypassing the need for an arbitrary threshold defining “high” NRF2 activity. High NRF2 is associated with significant decreases in overall survival in four of these cancers, and three of the four (KIRP, BLCA, and LIHC) were also deemed significant in the threshold-based comparisons. Thus, NRF2 hyperactivation is associated with decreased overall survival in multiple cancer types, and this relationship is especially evident for papillary renal cell carcinoma (KIRP), bladder carcinoma (BLCA), and hepatocellular carcinoma (LIHC).

Although NRF2's role in oncogenesis has attracted significant attention, it is still not clear exactly how NRF2 provides cells with an oncogenic advantage. Notably, several of the NRF2 cancer target genes from this study have been functionally linked to metabolic, proliferative, and chemoresistance advantages in cancer. Many of the NRF2

targets described here directly or indirectly increase nucleophilic tone, which would help cancer cells compensate for the increased ROS levels associated with rapid proliferation [107,108]. The antioxidants glutathione and thioredoxin are essential for cancer initiation and progression [109], and key regulators of these two antioxidant pools are NRF2 cancer target genes. Glutathione levels are largely dependent on the actions of *GCLC* and *GCLM*, which encode the catalytic and modifier subunits of glutamate-cysteine ligase, *SLC7A11* and *SLC3A2*, which encode the two subunits of the system x_c^- cystine/glutamate antiporter, and *GSR*, which codes for glutathione reductase. Thioredoxin is encoded by the gene *TXN*, and reduced thioredoxin is regenerated by thioredoxin reductase, which is encoded by *TXNRD1*. Moreover, the NRF2 cancer targets *TKT* (encodes transketolase) and *ME1* (encodes malic enzyme 1) are important for generation of the reducing agent NADPH that are also key regulators of oncogenic metabolic reprogramming [110–112], and the metabolic enzymes and transporters encoded by *ABCB6*, *ABCC3*, *AKR1C3*, *EPHX1*, *GSTM3*, and *NQO1* have been implicated in chemoresistance and/or radiation resistance [70,113–117]. In addition, multiple NRF2 cancer targets have cell cycle-related functions that could contribute to NRF2's oncogenic potential. For example, the antioxidant enzyme peroxiredoxin 1 (encoded by *PRDX1*) selectively protects telomeres from ROS-mediated damage during S and G2 phases of the cell cycle, which facilitates telomere elongation and delays cellular aging [118]. And *TLK1* and *ASF1A* encode a kinase and histone chaperone, respectively, that are important for chromatin assembly during DNA replication and DNA repair [119–122]. Taken together, the core cancer target genes described here offer a multitude of functions that could explain the oncogenic effects of constitutive NRF2 activity.

In summary, our results demonstrate that this mechanistically justified, cell-type-independent readout of NRF2 activity is a useful tool for dissecting the pathways driving NRF2 activation and for exploring the implications of NRF2 activation in cancer. The NRF2 cancer gene set will also be a valuable resource for asking more mechanistic questions about how NRF2 influences clinical outcomes beyond overall survival (chemoresistance, metastasis, etc.) and for exploring how additional genetic or environmental variables interact with NRF2 status in various cancer types. And, importantly, now that “high NRF2” tumors are readily identifiable, future work must move beyond this ubiquitously responsive gene set and focus on identifying tissue-specific NRF2 target genes that might further modify NRF2's malignant potential, especially in cancer types where activation of this pathway is associated with poor prognosis.

4. Materials and methods

4.1. Data sources

A table summarizing data sources for all analyses in this study is provided in the [Supplemental Materials \(Table S1\)](#), and further analysis details are provided in specific subsections below.

4.2. Differential expression analysis, ChIP-seq, and identification of NRF2 cancer target genes

We obtained the full list of genes differentially expressed in NRF2 mutant tumors from Araya et al., [Supplementary Table 11](#) (RNA-seq Differential Expression) [62]. We then extracted all data corresponding to genes carrying an “NFE2L2” tag in the *Code: Cancer_Region_Gene* column. This data was loaded into the R environment, where we split the differentially expressed genes by their logFC (log fold change) value into separate upregulated and downregulated sets. We then generated separate lists of up- or downregulated genes for each cancer type from TCGA (BLCA, HNSC, LUSC, and UCEC) [123,124]. Finally, we calculated the extent of overlap between each of the groups (shown in [Fig. 1](#)), using the *match* function and generated the final Venn diagrams

using the R package *VennDiagram* [125].

Direct NRF2 targets were identified based on reprocessing of our previously published ChIP-seq data from lymphoblastoid cell lines treated with sulforaphane – the sequencing data for this ChIP-seq experiment can be accessed under GEO Accession Number [GSE37589](#) [23,63]. ChIP peaks and fold-enrichment values were called using MACS2, with merged “NRF2 ChIP SFN” sample replicates from [GSE37589](#) treated as the “ChIP-seq” file, and merged “IgG ChIP SFN treated” and “input SFN treated” samples from [GSE37589](#) treated as the “Control” file. Peaks were called using a *q*-value cutoff of 1×10^{-2} , but to focus on the most robust NRF2 peaks, we implemented a stringent cutoff and only used peaks with a *q*-value $\leq 1 \times 10^{-5}$ for the remainder of our analyses ([Table S2](#)). ChIP-seq peaks were assigned to target genes as previously described [23,63], and the strongest putative ARE in each peak was identified using PWM calculations based on a set of 57 published AREs [126].

To identify direct NRF2 cancer targets, we cross-referenced the list of upregulated genes identified in the differential expression analysis with the list of NRF2 targets obtained by ChIP-seq. We made the following four groups of differentially upregulated genes based on the previously described differential expression overlap analysis: 1) those found in any of the four NRF2-associated cancers, 2) those found in at least 2 cancers, 3) those found in ≥ 3 , and 4) those found in all 4 cancers. We then used the *match* function in R to determine how many of these genes were found in the set of > 3000 high-confidence ($-\log_{10}(q) \geq 5$) NRF2 targets. The 32 direct NRF2 target genes that were upregulated in at least two cancer types were used for subsequent analyses.

4.3. Integration with DNase hypersensitivity and chromatin state data

To investigate the DNA accessibility of NRF2 target regions in chromatin across many tissue/cell types, we looked at DNase hypersensitivity (DHS) data. We obtained this data from the ENCODE project at <http://hgdownload.soe.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeRegDnaseClustered/> (file name: *wgEncodeRegDnaseClusteredV3.bed.gz*) [127]. This data includes both genomic regions of DHS sites, as well as the number of cell lines each DHS site is shared across (the *name* column in this dataset, described here: https://genome.ucsc.edu/cgi-bin/hgTables?db=hg19&hgta_group=regulation&hgta_track=wgEncodeRegDnaseClustered&hgta_table=wgEncodeRegDnaseClusteredV3&hgta_doSchema=describe+table+schema). We then used the *intersect* function of the BEDTools software suite [128] to find regions of overlap between high-confidence NRF2 targets and the DNase hypersensitivity sites, and combined the ChIP-seq data with this DHS overlap data. In R, we then extracted only DHS sites that shared at least 50 bp of overlap with an NRF2 target site. We finalized the DHS data by assigning any NRF2 target with at least 50 bp overlap a DHS score corresponding to the number of cell types the DHS signal was detected across in the ENCODE data (the *name* score), and any target with no DHS overlap or an overlap less than 50 bp a DHS score of 0.

To ‘score’ the amount of overlap for each NRF2 target with the chromatin states present in the Roadmap Epigenomics 5-mark, 15-state chromatin state model (<http://egg2.wustl.edu/-roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final>), we used the Python branch of the GIGGLE genomics search engine (<https://github.com/ryanlayer/giggle>, [90]). Briefly, we used an index, built across the 15 chromatin states and 127 tissue and cell types found in the Roadmap model, to query each NRF2 target interval (from ChIP-seq MACS2 peaks). We then calculated the fraction of overlap of the target sequence with each of the 15 chromatin states per cell/tissue type, and assigned a score of 1 to the chromatin state with the largest fraction of overlap and a 0 to all other (minor or zero-overlap) states. Finally, we calculated the sum score for each state across all cells/tissues and recorded this. Thus, each NRF2 target had a single score/value for each chromatin state, which represents the

number of cell types in Roadmap for which that chromatin state was the ‘dominant’ signal (Fig. 4D).

4.4. Inferred NRF2 activity and receiver operator characteristic analysis

To assess the usefulness, or accuracy, of our 32 target NRF2 activity signature in classifying NRF2-related mutations or variants, we employed a receiver operator characteristic (ROC) curve approach on TCGA gene expression and mutation (or copy number variation) data. First, we downloaded three sets of data for all samples in TCGA using XenaPython: 1) pan-cancer normalized gene expression values from the UCSC TOIL recompute, 2) all non-silent mutation data for these samples for *NFE2L2*, *KEAP1* and *CUL3*, and 3) all ordinal copy-number variation (CNV) scores for these samples for *NFE2L2*, *KEAP1* and *CUL3*. The CNV data we used had been assigned a score using a simple thresholding approach and the GISTIC2 algorithm [129], with the following categories: $-2 \approx$ homozygous loss, $-1 \approx$ heterozygous loss, $0 \approx$ normal diploid, $+1 \approx$ low-level copy number gain, $+2 \approx$ high-level amplification. We then extracted a set of 7679 samples that had gene expression values for all 32 targets, as well as mutation and CNV data for the three genes.

We then ‘binned’ each sample according to which of the 33 cancer types in TCGA it corresponded to and calculated its NRF2 signature score: the average of all gene expression values for the 32 NRF2 cancer targets, in each sample (dots in Fig. 5). Next, we calculated the mean NRF2 signature across all samples, as well as all samples within each cancer separately (black lines in Fig. 5; see also Supplemental Fig. 4), as it is expected that different tissues will require different basal levels of NRF2 activity. We used these mean ‘pan-cancer’ and ‘cancer-specific’ NRF2 signature scores to control for tissue-specific NRF2 activity; specifically, we normalized each sample by subtracting the difference between the pan-cancer and cancer-specific score from the sample’s individual score. For instance, if a LUSC sample had: NRF2 signature score = 12 and cancer-specific score = 11.4, and the pan-cancer score = 10.5, the normalized NRF2 signature score for that sample would be $12 - (11.4 - 10.5) = 11.1$.

4.5. Inferred NRF2 activity and survival analysis

Gene expression and survival data for the four NRF2-associated cancers (BLCA, HNSC, LUSC, and UCEC) were obtained using the Xena Browser tool from UCSC Xena at <http://xena.ucsc.edu/>. Briefly, we selected the TCGA cohort associated with each cancer – for instance ‘TCGA Bladder Cancer (BLCA)’ – and then added the following data types: *gene expression RNAseq (polyA + IlluminaHiSeq pancan normalized)*, phenotype: *_TIME_TO_EVENT (overall survival in days)*, phenotype: *_EVENT (overall survival indicator, 1 = death 0 = censor)*, and phenotype: *sample type*. We downloaded the generated information in a comma-separated table, and then used R to extract only samples with: 1) a *sample type* of ‘Primary Tumor’, ‘Metastatic’ or ‘Recurrent Tumor’ (excluding ‘Normal Tissue’), and 2) data for all data types (gene expression and survival). This ‘pruned’ dataset was then used for the remaining analyses. Gene expression and survival data for all other TCGA-associated cancers were obtained using the XenaPython tool (<http://xena.ucsc.edu/xena-python-api/>) with the same filtering criteria as above.

With the TCGA data retrieved using Xena, we carried out agglomerative hierarchical clustering of the tumor sample gene expression data (for the 32 cancer target genes) with the *hclust* function of the *stats* package in R. ‘Euclidean’ distance and the ‘average’ clustering method were employed. The heatmaps in Fig. 7 and Supplemental Fig. 5 were generated using the same clustering parameters with the *heatmap* function [130] (*heatmap* package). Tumor samples (and the patients they were derived from) were assigned to groups based on these clustering results; specifically, the samples were split at the highest branch in the dendrogram produced by the clustering algorithm, into two groups. The samples in the group with the highest average expression of

the cancer target genes were dubbed “high NRF2”, the others were labeled “normal NRF2”.

For the thresholding results shown in Supplemental Fig. 6, we used the same TCGA data as for cluster analysis and calculated the NRF2 signature score from the 32 targets (as above) for each sample. We then computed the value for the 90th percentile of this NRF2 signature score across all samples. Any sample with an NRF2 signature score above this 90th percentile was assigned to the high NRF2 group, and any at or below this value was assigned normal NRF2. Four cancer types (DLBC, KICH, PCPG, and TGCT) were omitted from the thresholding-based analysis in Supplemental Fig. 6 because hazard ratios and *p*-values could not be computed due to no deaths in one of the groups; these cancers all have good long-term outcomes, with both groups (normal or high NRF2) having a 9 + year survival probability greater than 70%.

For survival analysis, the survival data (*_EVENT* and *_TIME_TO_EVENT*) were then loaded into either R or Graphpad Prism for analysis. Samples were assigned to “high NRF2” or “normal NRF2” groups either according to hierarchical clustering, or as part of the top 10% based on their NRF2 signature score. Hazard ratios (HR), confidence intervals (CI), and log-rank (Mantel-Cox) *p*-values were generated using the R *survival* package, with a *Surv()* object and fit to a *coxph()* proportional hazards regression model (default settings for both). HR and CI values were \log_2 transformed for presentation in Supplemental Figs. 5 and 6. Kaplan-Meier plots were generated using the *Survival* function in Prism.

4.6. Statistics

For comparisons of NRF2 targets by fold enrichment, PWM score and DHS score, *p*-values were obtained using the non-parametric Mann-Whitney-Wilcoxon (alternatively known as Wilcoxon rank sum) test in R, with the *wilcox.test* function set to unpaired, two-sided and using continuity correction.

Acknowledgements

The authors wish to acknowledge Sarah Lacher for helpful discussions. We are also indebted to the researchers who worked to generate and process the data cited. This work was supported by funding from the National Institute of General Medical Sciences, National Institutes of Health, USA (R35-GM-119553 to M.S.) and the Intramural Research Program of the National Institute of Environmental Health Sciences, National Institutes of Health, USA [Z01-ES-100475 and Z01-ES-46008 (D.A.B)].

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.redox.2018.07.026](https://doi.org/10.1016/j.redox.2018.07.026).

References

- [1] G.P. Sykiotis, D. Bohmann, Stress-activated cap'n/collar transcription factors in aging and human disease, *Sci. Signal.* 3 (2010) re3.
- [2] K.M. Holmstrom, T. Finkel, Cellular mechanisms and physiological consequences of redox-dependent signalling, *Nat. Rev. Mol. Cell Biol.* 15 (2014) 411–421.
- [3] G. Waris, H. Ahsan, Reactive oxygen species: role in the development of cancer and various chronic conditions, *J. Carcinog.* 5 (2006) 14.
- [4] G. Barja, Free radicals and aging, *Trends Neurosci.* 27 (2004) 595–600.
- [5] M. Jimenez-Del-Rio, C. Velez-Pardo, The bad, the good, and the ugly about oxidative stress, *Oxid. Med. Cell. Longev.* 2012 (2012) 163913.
- [6] B. Halliwell, Biochemistry of oxidative stress, *Biochem. Soc. Trans.* 35 (2007) 1147–1150.
- [7] H.M. Leinonen, E. Kansanen, P. Polonen, M. Heinaniemi, A.L. Leinonen, Dysregulation of the Keap1-Nrf2 pathway in cancer, *Biochem. Soc. Trans.* 43 (2015) 645–649.
- [8] H. Zhang, K.J. Davies, H.J. Forman, Oxidative stress response and Nrf2 signaling in aging, *Free Radic. Biol. Med.* (2015).
- [9] T. Suzuki, M. Yamamoto, Stress-sensing mechanisms and the physiological roles of the Keap1-Nrf2 system during cellular stress, *J. Biol. Chem.* 292 (2017)

- 16817–16824.
- [10] R.S. Friling, A. Bensimon, Y. Tichauer, V. Daniel, Xenobiotic-inducible expression of murine glutathione S-transferase Ya subunit gene is controlled by an electrophile-responsive element, *Proc. Natl. Acad. Sci. USA* 87 (1990) 6258–6262.
- [11] A. Otsuki, M. Suzuki, F. Katsuoaka, K. Tsuchida, H. Suda, M. Morita, R. Shimizu, M. Yamamoto, Unique cisrome defined as CsMBE is strictly required for Nrf2-SMaf heterodimer function in cytoprotection, *Free Radic. Biol. Med.* 91 (2016) 45–57.
- [12] T.H. Rushmore, M.R. Morton, C.B. Pickett, The antioxidant responsive element. activation by oxidative stress and identification of the DNA consensus sequence required for functional activity, *J. Biol. Chem.* 266 (1991) 11632–11639.
- [13] K. Itoh, T. Chiba, S. Takahashi, T. Ishii, K. Igarashi, Y. Katoh, T. Oyake, N. Hayashi, K. Satoh, I. Hatayama, M. Yamamoto, Y. Nabeshima, An Nrf2/small Maf heterodimer mediates the induction of phase II detoxifying enzyme genes through antioxidant response elements, *Biochem. Biophys. Res. Commun.* 236 (1997) 313–322.
- [14] M.B. Kannan, V. Solovieva, V. Blank, The small MAF transcription factors MAFF, MAFK and MAFK: current knowledge and perspectives, *Biochim. Et. Biophys. Acta* 1823 (2012) 1841–1846.
- [15] T. Toki, J. Itoh, J. Kitazawa, K. Arai, K. Hatakeyama, J. Akasaka, K. Igarashi, N. Nomura, M. Yokoyama, M. Yamamoto, E. Ito, Human small Maf proteins form heterodimers with CNC family transcription factors and recognize the NF-E2 motif, *Oncogene* 14 (1997) 1901–1910.
- [16] L.M. Zipper, R.T. Mulcahy, The Keap1 BTB/POZ dimerization function is required to sequester Nrf2 in cytoplasm, *J. Biol. Chem.* 277 (2002) 36544–36552.
- [17] N. Wakabayashi, K. Itoh, J. Wakabayashi, H. Motohashi, S. Noda, S. Takahashi, S. Imakado, T. Kotsuji, F. Otsuka, D.R. Roop, T. Harada, J.D. Engel, M. Yamamoto, Keap1-null mutation leads to postnatal lethality due to constitutive Nrf2 activation, *Nat. Genet.* 35 (2003) 238–245.
- [18] M. McMahon, K. Itoh, M. Yamamoto, J.D. Hayes, Keap1-dependent proteasomal degradation of transcription factor Nrf2 contributes to the negative regulation of antioxidant response element-driven gene expression, *J. Biol. Chem.* 278 (2003) 21592–21600.
- [19] K. Itoh, N. Wakabayashi, Y. Katoh, T. Ishii, K. Igarashi, J.D. Engel, M. Yamamoto, Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain, *Genes Dev.* 13 (1999) 76–86.
- [20] E. Kansanen, S.M. Kuosmanen, H. Leinonen, A.L. Levenon, The Keap1-Nrf2 pathway: mechanisms of activation and dysregulation in cancer, *Redox Biol.* 1 (2013) 45–49.
- [21] A.T. Dinkova-Kostova, W.D. Holtzclaw, R.N. Cole, K. Itoh, N. Wakabayashi, Y. Katoh, M. Yamamoto, P. Talalay, Direct evidence that sulfhydryl groups of Keap1 are the sensors regulating induction of phase 2 enzymes that protect against carcinogens and oxidants, *Proc. Natl. Acad. Sci. USA* 99 (2002) 11908–11913.
- [22] D. Malhotra, E. Portales-Casamar, A. Singh, S. Srivastava, D. Arenillas, C. Happel, C. Shyr, N. Wakabayashi, T.W. Kensler, W.W. Wasserman, S. Biswal, Global mapping of binding sites for Nrf2 identifies novel targets in cell survival response through ChIP-Seq profiling and network analysis, *Nucleic Acids Res.* 38 (2010) 5718–5734.
- [23] B.N. Chorley, M.R. Campbell, X. Wang, M. Karaca, D. Sambandan, F. Bangura, P. Xue, J. Pi, S.R. Kleeberger, D.A. Bell, Identification of novel NRF2-regulated genes by ChIP-Seq: influence on retinoid X receptor alpha, *Nucleic Acids Res.* 40 (2012) 7416–7429.
- [24] S.E. Lacher, J.S. Lee, X. Wang, M.R. Campbell, D.A. Bell, M. Slattery, Beyond antioxidant genes in the ancient Nrf2 regulatory network, *Free Radic. Biol. Med.* 88 (2015) 452–465.
- [25] J.D. Hayes, A.T. Dinkova-Kostova, The Nrf2 regulatory network provides an interface between redox and intermediary metabolism, *Trends Biochem. Sci.* 39 (2014) 199–218.
- [26] O.H. Lee, A.K. Jain, V. Papusha, A.K. Jaiswal, An auto-regulatory loop between stress sensors Inrf2 and Nrf2 controls their cellular abundance, *J. Biol. Chem.* 282 (2007) 36412–36420.
- [27] G.P. Sykiotis, D. Bohmann, Keap1/Nrf2 signaling regulates oxidative stress tolerance and lifespan in *Drosophila*, *Dev. Cell* 14 (2008) 76–85.
- [28] D.H. Oh, D. Rigas, A. Cho, J.Y. Chan, Deficiency in the nuclear-related factor erythroid 2 transcription factor (Nrf1) leads to genetic instability, *FEBS J.* 279 (2012) 4121–4130.
- [29] L. Gan, J.A. Johnson, Oxidative damage and the Nrf2-ARE pathway in neurodegenerative diseases, *Biochim. Et. Biophys. Acta* 1842 (2014) 1208–1218.
- [30] A. Singh, V. Misra, R.K. Thimmulappa, H. Lee, S. Ames, M.O. Hoque, J.G. Herman, S.B. Baylín, D. Sidransky, E. Gabrielson, M.V. Brock, S. Biswal, Dysfunctional KEAP1-NRF2 interaction in non-small-cell lung cancer, *PLoS Med.* 3 (2006) e420.
- [31] T. Shibata, T. Ohta, K.I. Tong, A. Kokubu, R. Odogawa, K. Tsuta, H. Asamura, M. Yamamoto, S. Hirohashi, Cancer related mutations in NRF2 impair its recognition by Keap1-Cul3 E3 ligase and promote malignancy, *Proc. Natl. Acad. Sci. USA* 105 (2008) 13568–13573.
- [32] P.A. Konstantinopoulos, D. Spentzos, E. Fountzilas, N. Francoeur, S. Sanisetty, A.P. Grammatikos, J.L. Hecht, S.A. Cannistra, Keap1 mutations and Nrf2 pathway activation in epithelial ovarian cancer, *Cancer Res.* 71 (2011) 5081–5089.
- [33] Y.R. Kim, J.E. Oh, M.S. Kim, M.R. Kang, S.W. Park, J.Y. Han, H.S. Eom, N.J. Yoo, S.H. Lee, Oncogenic NRF2 mutations in squamous cell carcinomas of oesophagus and skin, *J. Pathol.* 220 (2010) 446–451.
- [34] N. Cancer Genome Atlas Research, Comprehensive genomic characterization of squamous cell lung cancers, *Nature* 489 (2012) 519–525.
- [35] B. Padmanabhan, K.I. Tong, T. Ohta, Y. Nakamura, M. Scharlock, M. Ohtsui, M.I. Kang, A. Kobayashi, S. Yokoyama, M. Yamamoto, Structural basis for defects of Keap1 activity provoked by its point mutations in lung cancer, *Mol. Cell* 21 (2006) 689–700.
- [36] I. Ganan-Gomez, Y. Wei, H. Yang, M.C. Boyano-Adanez, G. Garcia-Manero, Oncogenic functions of the transcription factor Nrf2, *Free Radic. Biol. Med.* 65 (2013) 750–764.
- [37] S. Menegon, A. Columbano, S. Giordano, The Dual Roles of NRF2 in Cancer, *Trends Mol. Med.* 22 (2016) 578–593.
- [38] N. Cancer Genome Atlas, Comprehensive genomic characterization of head and neck squamous cell carcinomas, *Nature* 517 (2015) 576–582.
- [39] W.J. Gibson, E.A. Hoivik, M.K. Halle, A. Taylor-Weiner, A.D. Cherniack, A. Berg, F. Holst, T.I. Zack, H.M. Werner, K.M. Staby, M. Rosenberg, I.M. Stefansson, K. Kusunmano, A. Chevalier, K.K. Mauland, J. Trovik, C. Krakstad, M. Giannakis, E. Hodis, K. Woie, L. Bjorge, O.K. Vintermyr, J.A. Wala, M.S. Lawrence, G. Getz, S.L. Carter, R. Beroukhi, H.B. Salvesen, The genomic landscape and evolution of endometrial carcinoma progression and abdominopelvic metastasis, *Nat. Genet.* 48 (2016) 848–855.
- [40] D. Wang, Y. Ma, X. Yang, X. Xu, Y. Zhao, Z. Zhu, X. Wang, H. Deng, C. Li, F. Gao, J. Tong, K. Yamanaka, Y. An, Hypermethylation of the Keap1 gene inactivates its function, promotes Nrf2 nuclear accumulation, and is involved in arsenite-induced human keratinocyte transformation, *Free Radic. Biol. Med.* 89 (2015) 209–219.
- [41] V.D. Martinez, E.A. Vucic, L.A. Pikor, K.L. Thu, R. Hubaux, W.L. Lam, Frequent concerted genetic mechanisms disrupt multiple components of the NRF2 inhibitor KEAP1/CUL3/RBX1 E3-ubiquitin ligase complex in thyroid cancer, *Mol. Cancer* 12 (2013) 124.
- [42] V.D. Martinez, E.A. Vucic, K.L. Thu, L.A. Pikor, S. Lam, W.L. Lam, Disruption of KEAP1/CUL3/RBX1 E3-ubiquitin ligase complex components by multiple genetic mechanisms: Association with poor prognosis in head and neck cancer, *Head. neck* 37 (2015) 727–734.
- [43] P. Zhang, A. Singh, S. Yegnasubramanian, D. Esopi, P. Kombairaju, M. Bodas, H. Wu, S.G. Bova, S. Biswal, Loss of Kelch-like ECH-associated protein 1 function in prostate cancer cells causes chemoresistance and radioresistance and promotes tumor growth, *Mol. Cancer Ther.* 9 (2010) 336–346.
- [44] R. Wang, J. An, F. Ji, H. Jiao, H. Sun, D. Zhou, Hypermethylation of the Keap1 gene in human lung cancer cell lines and lung cancer tissues, *Biochem. Biophys. Res. Commun.* 373 (2008) 151–154.
- [45] G.M. DeNicola, F.A. Karreth, T.J. Humpton, A. Gopinathan, C. Wei, K. Frese, D. Mangal, K.H. Yu, C.J. Yeo, E.S. Calhoun, F. Scrimieri, J.M. Winter, R.H. Hruban, C. Iacobuzio-Donahue, S.E. Kern, I.A. Blair, D.A. Tuveson, Oncogene-induced NRF2 transcription promotes ROS detoxification and tumorigenesis, *Nature* 475 (2011) 106–109.
- [46] M. Komatsu, H. Kurokawa, S. Waguri, K. Taguchi, A. Kobayashi, Y. Ichimura, Y.S. Sou, I. Ueno, A. Sakamoto, K.I. Tong, M. Kim, Y. Nishito, S. Iemura, T. Natsume, T. Ueno, E. Kominami, H. Motohashi, K. Tanaka, M. Yamamoto, The selective autophagy substrate p62 activates the stress responsive transcription factor Nrf2 through inactivation of Keap1, *Nat. Cell Biol.* 12 (2010) 213–223.
- [47] D. Papp, K. Lenti, D. Modos, D. Fazekas, Z. Dul, D. Turei, L. Foldvari-Nagy, R. Nussinov, P. Csermely, T. Korcsmaros, The NRF2-related interactome and regulome contain multifunctional proteins and fine-tuned autoregulatory loops, *FEBS Lett.* 586 (2012) 1795–1802.
- [48] D. Turei, D. Papp, D. Fazekas, L. Foldvari-Nagy, D. Modos, K. Lenti, P. Csermely, T. Korcsmaros, NRF2-ome: an integrated web resource to discover protein interaction and regulatory networks of NRF2, *Oxid. Med. Cell. Longev.* 2013 (2013) 737591.
- [49] J.D. Wardyn, A.H. Ponsford, C.M. Sanderson, Dissecting molecular cross-talk between Nrf2 and NF-kappaB response pathways, *Biochem. Soc. Trans.* 43 (2015) 621–626.
- [50] M.M. Alam, K. Okazaki, L.T.T. Nguyen, N. Ota, H. Kitamura, S. Murakami, H. Shima, K. Igarashi, H. Sekine, H. Motohashi, Glucocorticoid receptor signaling represses the antioxidant response by inhibiting histone acetylation mediated by the transcriptional activator NRF2, *J. Biol. Chem.* 292 (2017) 7519–7530.
- [51] A. Cuadrado, G. Manda, A. Hassan, M.J. Alcaraz, C. Barbas, A. Daiber, P. Ghezzi, R. Leon, M.G. Lopez, B. Oliva, M. Pajares, A.I. Rojo, N. Robledinos-Anton, A.M. Valverde, E. Guney, H. Schmidt, Transcription factor NRF2 as a therapeutic target for chronic diseases: a systems medicine approach, *Pharmacol. Rev.* 70 (2018) 348–383.
- [52] A. Gonzalez-Perez, C. Perez-Llamas, J. Deu-Pons, D. Tamborero, M.P. Schroeder, A. Jene-Sanz, A. Santos, N. Lopez-Bigas, IntOGen-mutations identifies cancer drivers across tumor types, *Nat. Methods* 10 (2013) 1081–1082.
- [53] Y. Mitsuishi, K. Taguchi, Y. Kawatani, T. Shibata, T. Nukiwa, H. Aburatani, M. Yamamoto, H. Motohashi, Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming, *Cancer Cell* 22 (2012) 66–79.
- [54] D.V. Chartoumpakis, N. Wakabayashi, T.W. Kensler, Keap1/Nrf2 pathway in the frontiers of cancer and non-cancer cell metabolism, *Biochem. Soc. Trans.* 43 (2015) 639–644.
- [55] S. Murakami, H. Motohashi, Roles of Nrf2 in cell proliferation and differentiation, *Free Radic. Biol. Med.* 88 (2015) 168–178.
- [56] I.I.C. Chio, S.M. Jafarnejad, M. Ponz-Sarvis, Y. Park, K. Rivera, W. Palm, J. Wilson, V. Sangar, Y. Hao, D. Ohlund, K. Wright, D. Filippini, E.J. Lee, B. Da Silva, C. Schoepfer, J.E. Wilkinson, J.M. Buscaglia, G.M. DeNicola, H. Tiriach, M. Hammell, H.C. Crawford, E.E. Schmidt, C.B. Thompson, D.J. Pappin, N. Sonenberg, D.A. Tuveson, NRF2 promotes tumor maintenance by modulating mRNA translation in pancreatic cancer, *Cell* 166 (2016), pp. 963–976.
- [57] H. Kitamura, H. Motohashi, NRF2 addiction in cancer cells, *Cancer Sci.* 109 (2018) 900–911.
- [58] S.E. Lacher, M. Slattery, Gene regulatory effects of disease-associated variation in the NRF2 network, *Curr. Opin. Toxicol.* 1 (2016) 71–79.

- [59] H. Peng, H. Wang, P. Xue, Y. Hou, J. Dong, T. Zhou, W. Qu, S. Peng, J. Li, P.L. Carmichael, B. Nelson, R. Clewell, Q. Zhang, M.E. Andersen, J. Pi, Suppression of NRF2-ARE activity sensitizes chemotherapeutic agent-induced cytotoxicity in human acute monocytic leukemia cells, *Toxicology Appl. Pharmacol.* 292 (2016) 1–7.
- [60] L. Zhan, H. Zhang, Q. Zhang, C.G. Woods, Y. Chen, P. Xue, J. Dong, E.J. Tokar, Y. Xu, Y. Hou, J. Fu, K. Yarborough, A. Wang, W. Qu, M.P. Waalkes, M.E. Andersen, J. Pi, Regulatory role of KEAP1 and NRF2 in PPARgamma expression and chemoresistance in human non-small-cell lung carcinoma cells, *Free Radic. Biol. Med.* 53 (2012) 758–768.
- [61] J. Zhu, H. Wang, F. Chen, J. Fu, Y. Xu, Y. Hou, H.H. Kou, C. Zhai, M.B. Nelson, Q. Zhang, M.E. Andersen, J. Pi, An overview of chemical inhibitors of the Nrf2-ARE signaling pathway and their potential applications in cancer therapy, *Free Radic. Biol. Med.* 99 (2016) 544–556.
- [62] C.L. Araya, C. Cenik, J.A. Reuter, G. Kiss, V.S. Pande, M.P. Snyder, W.J. Greenleaf, Identification of significantly mutated regions across cancer types highlights a rich landscape of functional molecular alterations, *Nat. Genet.* 48 (2016) 117–125.
- [63] X. Wang, M.R. Campbell, S.E. Lacher, H.Y. Cho, M. Wan, C.L. Crowl, B.N. Chorley, G.L. Bond, S.R. Kleeberger, M. Slattery, D.A. Bell, A polymorphic antioxidant response element links NRF2/sMAF binding to Enhanced MAPT expression and reduced risk of parkinsonian disorders, *Cell Rep.* (2016).
- [64] W. Hur, N.S. Gray, Small molecule modulators of antioxidant response pathway, *Curr. Opin. Chem. Biol.* 15 (2011) 162–173.
- [65] E.Y. Chen, C.M. Tan, Y. Kou, Q. Duan, Z. Wang, G.V. Meirelles, N.R. Clark, A. Ma'ayan, Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool, *BMC Bioinforma.* 14 (2013) 128.
- [66] M.V. Kuleshov, M.R. Jones, A.D. Rouillard, N.F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S.L. Jenkins, K.M. Jagodnik, A. Lachmann, M.G. McDermott, C.D. Monteiro, G.W. Gunderen, A. Ma'ayan, Enrichr: a comprehensive gene set enrichment analysis web server 2016 update (W90-97), *Nucleic Acids Res.* 44 (2016) (W90-97).
- [67] J. Lynch, Y. Fukuda, P. Krishnamurthy, G. Du, J.D. Schuetz, Cell survival under stress is enhanced by a mitochondrial ATP-binding cassette transporter that regulates hemoproteins, *Cancer Res.* 69 (2009) 5560–5567.
- [68] T.A. Schaedler, B. Faust, C.A. Shintre, E.P. Carpenter, V. Srinivasan, H.W. van Veen, J. Balk, Structures and functions of mitochondrial ABC transporters, *Biochem. Soc. Trans.* 43 (2015) 943–951.
- [69] Z. Yu, C. Zhang, H. Wang, J. Xing, H. Gong, E. Yu, W. Zhang, X. Zhang, G. Cao, C. Fu, Multidrug resistance-associated protein 3 confers resistance to chemoradiotherapy for rectal cancer by regulating reactive oxygen species and caspase-3-dependent apoptotic pathway, *Cancer Lett.* 353 (2014) 182–193.
- [70] M.J. Canet, M.D. Merrell, B.G. Harder, J.M. Maher, T. Wu, A.J. Lickteig, J.P. Jackson, D.D. Zhang, M. Yamamoto, N.J. Cherrington, Identification of a functional antioxidant response element within the eighth intron of the human ABCG3 gene, *Drug Metab. Dispos.: Biol. fate Chem.* 43 (2015) 93–99.
- [71] S.M. Hong, C.W. Park, S.W. Kim, Y.J. Nam, J.H. Yu, J.H. Shin, C.H. Yun, S.H. Im, K.T. Kim, Y.C. Sung, K.Y. Choi, NAMPT suppresses glucose deprivation-induced oxidative stress by increasing NADPH levels in breast cancer, *Oncogene* 35 (2016) 3544–3554.
- [72] D. Cerna, H. Li, S. Flaherty, N. Takebe, C.N. Coleman, S.S. Yoo, Inhibition of nicotinamide phosphoribosyltransferase (NAMPT) activity by small molecule GMX1778 regulates reactive oxygen species (ROS)-mediated cytotoxicity in a p53- and nicotinic acid phosphoribosyltransferase1 (NAPRT1)-dependent manner, *J. Biol. Chem.* 287 (2012) 22408–22417.
- [73] Z. Moore, G. Chakrabarti, X. Luo, A. Ali, Z. Hu, F.J. Fattah, R. Vemireddy, R.J. DeBerardinis, R.A. Brekken, D.A. Boothman, NAMPT inhibition sensitizes pancreatic adenocarcinoma cells to tumor-selective, PAR-independent metabolic catastrophe and cell death induced by beta-lapachone, *Cell Death Dis.* 6 (2015) e1599.
- [74] I.M. Xu, R.K. Lai, S.H. Lin, A.P. Tse, D.K. Chiu, H.Y. Koh, C.T. Law, C.M. Wong, Z. Cai, C.C. Wong, I.O. Ng, Transketolase counteracts oxidative stress to drive cancer development, *Proc. Natl. Acad. Sci. USA* 113 (2016) E725–E734.
- [75] D. Szklarczyk, J.H. Morris, H. Cook, M. Kuhn, S. Wyder, M. Simonovic, A. Santos, N.T. Doncheva, A. Roth, P. Bork, L.J. Jensen, C. von Mering, The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible, *Nucleic Acids Res.* 45 (2017) D362–D368.
- [76] R. Piazza, D. Ramazzotti, R. Spinelli, A. Pirola, L. De Sano, P. Ferrari, V. Magistri, N. Cordani, N. Sharma, C. Gambacorti-Passerini, OncoScore: a novel, Internet-based tool to assess the oncogenic potential of genes, *Sci. Rep.* 7 (2017) 46290.
- [77] P.A. Futreal, L. Coin, M. Marshall, T. Down, T. Hubbard, R. Wooster, N. Rahman, M.R. Stratton, A census of human cancer genes, *Nat. Rev. Cancer* 4 (2004) 177–183.
- [78] W.W. Fisher, J.J. Li, A.S. Hammonds, J.B. Brown, B.D. Pfeiffer, R. Weiszmann, S. MacArthur, S. Thomas, J.A. Stamatoyannopoulos, M.B. Eisen, P.J. Bickel, M.D. Biggin, S.E. Celniker, DNA regions bound at low occupancy by transcription factors do not drive patterned reporter gene expression in *Drosophila*, *Proc. Natl. Acad. Sci. USA* 109 (2012) 21330–21335.
- [79] M. Slattery, T. Zhou, L. Yang, A.C. Dantas Machado, R. Gordan, R. Rohs, Absence of a simple code: how transcription factors read the genome, *Trends Biochem. Sci.* 39 (2014) 381–399.
- [80] X.Y. Li, S. Thomas, P.J. Sabo, M.B. Eisen, J.A. Stamatoyannopoulos, M.D. Biggin, The role of chromatin accessibility in directing the widespread, overlapping patterns of *Drosophila* transcription factor binding, *Genome Biol.* 12 (2011) R34.
- [81] J. Ernst, M. Kellis, Interplay between chromatin state, regulator binding, and regulatory motifs in six human cell types, *Genome Res.* 23 (2013) 1142–1154.
- [82] M. Slattery, R. Voutev, L. Ma, N. Negre, K.P. White, R.S. Mann, Divergent transcriptional regulatory logic at the intersection of tissue growth and developmental patterning, *PLoS Genet.* 9 (2013) e1003753.
- [83] B. Xin, R. Rohs, Relationship between histone modifications and transcription factor binding is protein family specific, *Genome Res.* (2018).
- [84] P.J. Sabo, M.S. Kuehn, R. Thurman, B.E. Johnson, E.M. Johnson, H. Cao, M. Yu, E. Rosenzweig, J. Goldy, A. Haydock, M. Weaver, A. Shafer, K. Lee, F. Neri, R. Humbert, M.A. Singer, T.A. Richmond, M.O. Dorschner, M. McArthur, M. Hawrylycz, R.D. Green, P.A. Navas, W.S. Noble, J.A. Stamatoyannopoulos, Genome-scale mapping of DNase I sensitivity in vivo using tiling DNA microarrays, *Nat. Methods* 3 (2006) 511–518.
- [85] R.E. Thurman, E. Rynes, R. Humbert, J. Vierstra, M.T. Maurano, E. Haugen, N.C. Sheffield, A.B. Stergachis, H. Wang, B. Vernot, K. Garg, S. John, R. Sandstrom, D. Bates, L. Boatman, T.K. Canfield, M. Diegel, D. Dunn, A.K. Ebersol, T. Frum, E. Giste, A.K. Johnson, E.M. Johnson, T. Kutayavin, B. Lajoie, B.K. Lee, K. Lee, D. London, D. Lotakis, S. Neph, F. Neri, E.D. Nguyen, H. Qu, A.P. Reynolds, V. Roach, A. Safi, M.E. Sanchez, A. Sanyal, A. Shafer, J.M. Simon, L. Song, S. Vong, M. Weaver, Y. Yan, Z. Zhang, B. Lenhard, M. Tewari, M.O. Dorschner, R.S. Hansen, P.A. Navas, G. Stamatoyannopoulos, V.R. Iyer, J.D. Lieb, S.R. Sunyaev, J.M. Akey, P.J. Sabo, R. Kaul, T.S. Furey, J. Dekker, G.E. Crawford, J.A. Stamatoyannopoulos, The accessible chromatin landscape of the human genome, *Nature* 489 (2012) 75–82.
- [86] D. Chung, P.F. Kuan, B. Li, R. Sanalkumar, K. Liang, E.H. Bresnick, C. Dewey, S. Keles, Discovering transcription factor binding sites in highly repetitive regions of genomes with multi-read analysis of ChIP-Seq data, *PLoS Comput. Biol.* 7 (2011) e1002111.
- [87] A. Afek, H. Cohen, S. Barber-Zucker, R. Gordan, D.B. Lukatsky, Nonconsensus protein binding to repetitive DNA sequence elements significantly affects eukaryotic genomes, *PLoS Comput. Biol.* 11 (2015) e1004429.
- [88] P. Deininger, Alu elements: know the SINES, *Genome Biol.* 12 (2011) 236.
- [89] A. Khandaj, W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kuderapour, Z. Zhang, J. Wang, M.J. Ziller, V. Amin, J.W. Whitaker, M.D. Schultz, L.D. Ward, A. Sarkar, G. Quon, R.S. Sandstrom, M.L. Eaton, Y.C. Wu, A.R. Pfennig, X. Wang, M. Claussnitzer, Y. Liu, C. Coarfa, R.A. Harris, N. Shores, C.B. Epstein, E. Gjoneska, D. Leung, W. Xie, R.D. Hawkins, R. Lister, C. Hong, P. Gascard, A.J. Mungall, R. Moore, E. Chuah, A. Tam, T.K. Canfield, R.S. Hansen, R. Kaul, P.J. Sabo, M.S. Bansal, A. Carles, J.R. Dixon, K.H. Farh, S. Feizi, R. Karlic, A.R. Kim, A. Kulkarni, D. Li, R. Lowdon, G. Elliott, T.R. Mercer, S.J. Neph, V. Onuchic, P. Polak, N. Rajagopal, P. Ray, R.C. Sallari, K.T. Siebenthal, N.A. Sinnott-Armstrong, M. Stevens, R.E. Thurman, J. Wu, B. Zhang, X. Zhou, A.E. Beaudet, L.A. Boyer, P.L. De Jager, P.J. Farnham, S.J. Fisher, D. Haussler, S.J. Jones, W. Li, M.A. Marra, M.T. McManus, S. Sunyaev, J.A. Thomson, T.D. Tlsty, L.H. Tsai, W. Wang, R.A. Waterland, M.Q. Zhang, L.H. Chadwick, B.E. Bernstein, J.F. Costello, J.R. Ecker, M. Hirst, A. Meissner, A. Milosavljevic, B. Ren, J.A. Stamatoyannopoulos, T. Wang, M. Kellis, Integrative analysis of 111 reference human epigenomes, *Nature* 518 (2015) 317–330.
- [90] R.M. Layer, B.S. Pedersen, T. DiSera, G.T. Marth, J. Gertz, A.R. Quinlan, GIGGLE: a search engine for large-scale integrated genome analysis, *Nat. Methods* 15 (2018) 123–126.
- [91] A. Pekowska, T. Benoukraf, J. Zacarias-Cabeza, M. Belhocine, F. Koch, H. Holota, J. Imbert, J.C. Andrau, P. Ferrier, S. Spicuglia, H3K4 tri-methylation provides an epigenetic signature of active enhancers, *EMBO J.* 30 (2011) 4198–4210.
- [92] L.J. Core, A.L. Martins, C.G. Danko, C.T. Waters, A. Siepel, J.T. Lis, Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers, *Nat. Genet.* 46 (2014) 1311–1320.
- [93] R. Andersson, A. Sandelin, C.G. Danko, A unified architecture of transcriptional regulatory elements, *Trends Genet.: TIG* 31 (2015) 426–433.
- [94] X. Wang, L. He, S. Goggin, A. Saadat, L. Wang, M. Claussnitzer, M. Kellis, High-resolution genome-wide functional dissection of transcriptional regulatory regions in human, *bioRxiv* 193136 (2017).
- [95] J.E. Darnell Jr., Transcription factors as targets for cancer therapy, *Nat. Rev. Cancer* 2 (2002) 740–749.
- [96] A.S. Bhagwat, C.R. Vakoc, Targeting transcription factors in cancer, *Trends Cancer* 1 (2015) 53–65.
- [97] J.E. Bradner, D. Hnisz, R.A. Young, Transcription Addiction in Cancer, *Cell* 168 (2017) 629–643.
- [98] N. Cancer Genome Atlas Research, C. Kandoth, N. Schultz, A.D. Cherniack, R. Akbani, Y. Liu, H. Shen, A.G. Robertson, I. Pashtan, R. Shen, C.C. Benz, C. Yau, P.W. Laird, L. Ding, W. Zhang, G.B. Mills, R. Kucherlapati, E.R. Mardis, D.A. Levine, Integrated genomic characterization of endometrial carcinoma, *Nature* 497 (2013) 67–73.
- [99] N. Cancer Genome Atlas Research, Comprehensive molecular characterization of urothelial bladder carcinoma, *Nature* 507 (2014) 315–322.
- [100] N. Cancer Genome Atlas Research, Comprehensive molecular profiling of lung adenocarcinoma, *Nature* 511 (2014) 543–550.
- [101] S.J. Chapple, T.P. Keeley, D. Mastronicola, M. Arno, G. Vizcay-Barrena, R. Fleck, R.C.M. Siow, G.E. Mann, Bach1 differentially regulates distinct Nrf2-dependent genes in human venous and coronary artery endothelial cells adapted to physiological oxygen levels, *Free Radic. Biol. Med.* 92 (2016) 152–162.
- [102] M. Fang, J. Ou, L. Hutchinson, M.R. Green, The BRAF oncoprotein functions through the transcriptional repressor MAFG to mediate the CpG Island Methylator phenotype, *Mol. Cell* 55 (2014) 904–915.
- [103] A. Kobayashi, M.I. Kang, H. Okawa, M. Ohtsuji, Y. Zenke, T. Chiba, K. Igarashi, M. Yamamoto, Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate proteasomal degradation of Nrf2, *Mol. Cell. Biol.* 24 (2004) 7130–7139.

- [104] S.B. Cullinan, J.D. Gordan, J. Jin, J.W. Harper, J.A. Diehl, 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 (2004) 8477–8486.
- [105] D.D. Zhang, S.C. Lo, J.V. Cross, D.J. Templeton, M. Hannink, Keap1 is a redox-regulated substrate adaptor protein for a Cul3-dependent ubiquitin ligase complex, *Mol. Cell. Biol.* 24 (2004) 10941–10953.
- [106] A. Ooi, K. Dykema, A. Ansari, D. Petillo, J. Snider, R. Kahnoski, J. Anema, D. Craig, J. Carpten, B.T. Teh, K.A. Furge, CUL3 and NRF2 mutations confer an NRF2 activation phenotype in a sporadic form of papillary renal cell carcinoma, *Cancer Res.* 73 (2013) 2044–2051.
- [107] V. Sosa, T. Moline, R. Somoza, R. Paciucci, H. Kondoh, L.L. ME, Oxidative stress and cancer: an overview, *Ageing Res. Rev.* 12 (2013) 376–390.
- [108] H.J. Forman, K.J. Davies, F. Ursini, How do nutritional antioxidants really work: nucleophilic tone and para-hormesis versus free radical scavenging in vivo, *Free Radic. Biol. Med.* 66 (2014) 24–35.
- [109] I.S. Harris, A.E. Treloar, S. Inoue, M. Sasaki, C. Gorrini, K.C. Lee, K.Y. Yung, D. Brenner, C.B. Knobbe-Thomsen, M.A. Cox, A. Elia, T. Berger, D.W. Cescon, A. Adeoye, A. Brustle, S.D. Molyneux, J.M. Mason, W.Y. Li, K. Yamamoto, A. Wakeham, H.K. Berman, R. Khokha, S.J. Done, T.J. Kavanagh, C.W. Lam, T.W. Mak, Glutathione and thioredoxin antioxidant pathways synergize to drive cancer initiation and progression, *Cancer Cell* (2015).
- [110] C.W. Tseng, W.H. Kuo, S.H. Chan, H.L. Chan, K.J. Chang, L.H. Wang, Transketolase regulates the metabolic switch to control breast cancer cell metastasis via the alpha-ketoglutarate signaling pathway, *Cancer Res.* (2018).
- [111] Y.X. Lu, H.Q. Ju, Z.X. Liu, D.L. Chen, Y. Wang, Q. Zhao, Q.N. Wu, Z.L. Zeng, H.B. Qiu, P.S. Hu, Z.Q. Wang, D.S. Zhang, F. Wang, R.H. Xu, ME1 regulates NADPH homeostasis to promote gastric cancer growth and metastasis, *Cancer Res.* 78 (2018) 1972–1985.
- [112] S. Murai, A. Ando, S. Ebara, M. Hirayama, Y. Satomi, T. Hara, Inhibition of malic enzyme 1 disrupts cellular metabolism and leads to vulnerability in cancer cells in glucose-restricted conditions, *Oncogenesis* 6 (2017) e329.
- [113] D.M. Townsend, K.D. Tew, The role of glutathione-S-transferase in anti-cancer drug resistance, *Oncogene* 22 (2003) 7369–7375.
- [114] M. Decker, M. Arand, A. Cronin, Mammalian epoxide hydrolases in xenobiotic metabolism and signalling, *Arch. Toxicol.* 83 (2009) 297–318.
- [115] K. Minami, Y. Kamijo, Y. Nishizawa, S. Tabata, F. Horikuchi, M. Yamamoto, K. Kawahara, Y. Shinsato, T. Tachiwada, Z.S. Chen, K. Tsujikawa, M. Nakagawa, N. Seki, S. Akiyama, K. Arima, Y. Takeda, T. Furukawa, Expression of ABCB6 is related to resistance to 5-FU, SN-38 and vincristine, *Anticancer Res.* 34 (2014) 4767–4773.
- [116] W. Xiong, J. Zhao, H. Yu, X. Li, S. Sun, Y. Li, Q. Xia, C. Zhang, Q. He, X. Gao, L. Zhang, D. Zhou, Elevated expression of AKR1C3 increases resistance of cancer cells to ionizing radiation via modulation of oxidative stress, *PLoS One* 9 (2014) e111911.
- [117] M. Fiorillo, F. Sotgia, D. Sisci, A.R. Cappello, M.P. Lisanti, Mitochondrial "power" drives tamoxifen resistance: NQO1 and GCLC are new therapeutic targets in breast cancer, *Oncotarget* 8 (2017) 20309–20327.
- [118] E. Aeby, W. Ahmed, S. Redon, V. Simanis, J. Lingner, Peroxiredoxin 1 protects telomeres from oxidative damage and preserves telomeric DNA for extension by telomerase, *Cell Rep.* 17 (2016) 3107–3114.
- [119] H.H. Sillje, E.A. Nigg, Identification of human Asf1 chromatin assembly factors as substrates of Tousled-like kinases, *Curr. Biol.* 11 (2001) 1068–1073.
- [120] J.A. Mello, H.H. Sillje, D.M. Roche, D.B. Kirschner, E.A. Nigg, G. Almouzni, Human Asf1 and CAF-1 interact and synergize in a repair-coupled nucleosome assembly pathway, *EMBO Rep.* 3 (2002) 329–334.
- [121] S.P. Sen, A. De Benedetti, TLK1B promotes repair of UV-damaged DNA through chromatin remodeling by Asf1, *BMC Mol. Biol.* 7 (2006) 37.
- [122] I.M. Klimovskaia, C. Young, C.B. Stromme, P. Menard, Z. Jasencakova, J. Mejlvang, K. Ask, M. Ploug, M.L. Nielsen, O.N. Jensen, A. Groth, Tousled-like kinases phosphorylate Asf1 to promote histone supply during DNA replication, *Nat. Commun.* 5 (2014) 3394.
- [123] R.C.R. Team, A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2017.
- [124] H.B. Wickham, J. Bryan, readxl: Read Excel Files. (2017).
- [125] H. Chen, VennDiagram: generate high-resolution Venn Euler plots (2016).
- [126] X. Wang, D.J. Tomso, B.N. Chorley, H.Y. Cho, V.G. Cheung, S.R. Kleeberger, D.A. Bell, Identification of polymorphic antioxidant response elements in the human genome, *Human. Mol. Genet.* 16 (2007) 1188–1200.
- [127] E.P. Consortium, An integrated encyclopedia of DNA elements in the human genome, *Nature* 489 (2012) 57–74.
- [128] A.R. Quinlan, I.M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features, *Bioinformatics* 26 (2010) 841–842.
- [129] C.H. Mermel, S.E. Schumacher, B. Hill, M.L. Meyerson, R. Beroukhi, G. Getz, GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers, *Genome Biol.* 12 (2011) R41.
- [130] R. Kolde, *Pheatmap: Pretty heatmaps*. 2015.
- [131] D. Ramskold, E.T. Wang, C.B. Burge, R. Sandberg, An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data, *PLoS Comput. Biol.* 5 (2009) e1000598.