A novel crosstalk between Nrf2 and Smad2/3 bridged by two nuanced Keap1 isoforms

Feilong Chen[1,2,3], Mei Xiao[1,3], Deshuai Lou[4], Qing Wang[1,2,3], Reziyamu Wufur[1,2,3], Shaofan Hu[1,2,3], Zhengwen Zhang[5], Yeqi Wang[1], and Yiguo Zhang[2,3,*]

[1]College of Bioengineering and Graduate School, Chongqing University, No. 174 Shazheng Street, Shapingba District, Chongqing 40044, China

[2]Chongqing University Jiangjin Hospital, School of Medicine, Chongqing University, No. 725 Jiangzhou Avenue, Dingshan Street, Jiangjin District, Chongqing 402262, China

[3]TheLaboratory of Cell Biochemistry and Topogenetic Regulation, College of Bioengineering and Faculty of Medical Sciences, Chongqing University, No. 174 Shazheng Street, Shapingba District, Chongqing 40044, China

[4]Chongqing Key Laboratory of Medicinal Resources in the Three Gorges Reservoir Region, School of Biological and Chemical Engineering, Chongqing University of Education, Chongqing 400067, China

[5]Laboratory of Neuroscience, Institute of Cognitive Neuroscience and School of Pharmacy, University College London, 29-39 Brunswick Square, London WC1N 1AX, England, United Kingdom

*Correspondence: yiguozhang@cqu.edu.cn or eaglezhang64@gmail.com

Abstract

The Keap1-Nrf2 signalling to transcriptionally regulate antioxidant response element (ARE)-driven target genes has been accepted as key redox-sensitive pathway governing a vast variety of cellular stresses during healthy survival and disease development. Herein, we identified two nuanced isoforms α and β of Keap1, arising from its first and another in-frame translation starting codons, respectively. Those common and specific genes monitored by Keap1α and/or Keap1β were unravelled by transcriptomic sequencing of indicated experimental cell lines. Amongst them, an unusual interaction of Keap1 with Smad2/3 was discovered by parsing transcriptome sequencing, protein profiling, and immunoprecipitation data. Further examinations validated that Smad2/3 enable physical interaction with Keap1, as well as its isoforms α and β, by both EDGETSD and DLG motifs in the linker regions between their MH1 and MH2 domains, such that the stability of Smad2/3 and its transcriptional activity are enhanced with the prolonged half-lives and signalling responsiveness from the cytoplasmic to nuclear compartments. Such activation of Smad2/3 by Keap1, Keap1α or Keap1β was contributable to a competitively inhibitory effect of Nrf2. Overall, this discovery presents a novel functional bridge crossing the Keap1-Nrf2 and the TGF-β1-Smad2/3 signalling pathways in healthy growth and development.

Keywords: Keap1; Keap1α; Keap1β; Nrf2; Smad2; Smad3; protein-protein interaction; transcriptional regulation

Introduction

Appropriate cell physiology and homeostasis are necessarily maintained by integrating a vast variety of extrinsic and intrinsic signals transducing towards a series of responsive gene expression profiling in order to meet cellular adaptive requirements for changing environments. Of note, Keap1 (Kelch-like ECH-associated protein 1) towards Nrf2 (nuclear factor E2-related factor 2, called NFE2L2)-regulated genes is a key redox signalling pathway, governing transcriptional expression of a big set of antioxidant and electrophile response elements (AREs/EpREs)-driven genes against many of adverse cellular stresses during normal growth and development, and even diseases[1]. Within the cells, Keap1 exists in the form of homodimers and interacts with the ETGE and DLG motifs of the Neh2 domain in Nrf2 through its C-terminal Kelch/DGR domain[2,3], enabling Nrf2 to be ubiquitynated by the E3 ligase Cullin-3 for its rapid proteasomal degradation[4]. Thereby, it is inferable that under the normal conditions, Keap1 physically binds and sequesters Nrf2 in the cytoplasmic compartments[5,6]. Following stimulation by oxidative stress derived from reactive oxygen species (ROS), Nrf2 is released from Keap1 and then translocated into the nucleus, where it is enabled for transcriptional regulation of AREs/EpREs-driven antioxidant, detoxifying and cytoprotective genes. Such target genes include those encoding haem oxygenase 1 (HO-1), NAD(P)H:quinone oxidoreductase-1 (NQO1), γ-glutamyl cysteine ligase catalytic (GCLC) and modified (GCLM) subunits, and glutathione *S*-transferases (GSTs), among others[7-10].

Between redox-inducible Keap1-Nrf2 pathway and intrinsic developmental signalling cascades, there is existing key signal-integrated interplay that had been interrogated and explored by a few number of scientists over the past decade. The intrinsic developmental signalling cascades are predominantly induced by growth factors (including cytokines such as TGF-β[11]), morphogens (e.g., WNT[12]) mitogens and survival factors (e.g., Hedgehog[13]). However, only considerably less attention has been attracted on such a putative as-yet-unidentified interplay between the TGF-β-Smad and Keap1-Nrf2 signalling pathways, even though a large number of studies have indeed elucidated that the transforming growth factor-

β (TGF-β) superfamily, also including bone morphogenetic proteins (BMPs), play important roles in the embryogenesis, organogenesis and regulation of the whole-body homeostasis by interacting with membrane receptors that transduce information to the nucleus through both Smad-dependent and -independent pathways (including PI3K-AKT and MAPKs). Notably, the TGF-β receptor-regulated Smad2 and Smad3 (collectively called Smad2/3, hereafter) are two important players of the Smad transcription factor family, that mediates this growth factor-stimulated signalling transduction from the cell membrane surface to the nucleus, and thereby regulates critical physio-pathological functions of target genes involved in healthy growth and development, and even disease development. In this signalling pathway, Smad2/3 serve two downstream receptor-activating proteins of TGF-β1 and binding partners for Smad4[14]. All the Smad family proteins, particularly Smad2/3, are evolutionarily homologous in their structures, with two highly conserved domains, i.e., an N-terminal Mad homology domain-1 (MH1, for binding to target genes) and another C-terminal Mad homology domain-2 (MH2, for binding to cognate partners), joined by a linker region containing four co-phosphorylated serine sites for both GSK-3β (glycogen synthase kinase-3β) and MAPKs (including ERKs, JNKs, and p38 mitogen-activated protein kinases)[15,16]. Further studies had also shown that Smad2/3 are closely involved in tumorigenesis, wound healing, immune regulation and the extracellular matrix[17].

Recently, it is elaborately reviewed that many processes that downregulate Nrf2 are *de facto* triggered by TGF-β, with oxidative stress amplifying its signalling[18]. The canonical TGF-β pathway leads to Smad2/3- and Smad4-directed increases in the abundances of Hrd1, Bach1, MafK, and ATF3, all of which have been reported to repress Nrf2 activity. Another increase in NOX4-mediated production of ROS resulting from Smad2/3 activation by TGF-β heightens two non-canonical pathways *via* activation of TAK1 (TGF-β-activated kinase 1) and TRAF6 (TNF receptor-associated factor 6) signalling cascades to MAPKs, and hence further increases c-Jun levels, which antagonise Nrf2, following dimerization of c-Jun with ATF3 or Fos1/Fra1 (to comprise AP-1). Such oxidative stress may also increase suppression of Nrf2 by β-TrCP through facilitating formation of its DSGIS-containing phosphodegron by GSK-3β together with its priming kinases (e.g., JNK and p38 MAPKs). In turn, it is envisaged that downregulation of Nrf2 by TGF-β can also reinforce activation of both canonical and non-canonical TGF-β signalling pathways, if Nrf2-directed antioxidant and detoxification systems will be less able to suppress ROS generated by NOX4. Consequently, the increases in Smad-dependent cancer motility and growth, besides intracellular ROS, resulted from Nrf2 deficiency, as accompanied by concomitant phosphorylation of Smad2/3 in their linker regions and C-terminal ends, induction of Slug, a transcriptional repressor of the cell adhesion E-cadherin[19]. Conversely, over-expression of wild-type Nrf2, but not its dominant-negative mutant, suppressed the transcriptional activity of a TGF-β1-responsive CAGA-directed luciferase reporter gene, whereas knockdown of Nrf2 enhanced the CAGA reporter activity, as well as the expression of endogenous Smad2/3-target genes, which were further identified to be regulated competitively by a nuclear co-immunoprecipitable complex of Nrf2 with Smad2/3/4[19]. This supports the notion that loss of Nrf2 in an oncogenic context-dependent manner also enhances cellular plasticity and motility, at least in part, by the TGF-β1-Smad2/3/4 signalling.

Meanwhile, TGF-β1 is also demonstrated as a potent stimulator of epithelial-to-mesenchymal transition (EMT) by activating the profibrotic genes encoding fibronectin-1 and collagen 1A1 in chronic diseases[20]. The TGFβ1-EMT changes are enhanced by *Nrf2* knockdown, but suppressed in a stable *Keap1*-knockdown model (to pre-activate Nrf2), concomitantly with repression of TGFβ1-stimulated Smad2/3 phosphorylation and transcriptional activity[20]. The Keap1-Nrf2 antioxidant system is further validated as an effective modulator of TGFβ1-stimulated epithelial transition to fibroblastic cells through the Smurf1-Smad7 signalling[20]. Consistently, Nrf2 has been shown to inhibit the TGF-β1-dependent expression of fibrosis markers in a human stellate cell line[21,22], whereas TGF-β1 can also reduce the abundance of Nrf2 in another rat hepatic stellate cell line[21,23]. However, such inter-inhibitory mechanisms between Keap1-Nrf2 antioxidant system and TGF-β1-Smad2/3 signalling have not been elucidated to date.

In the present study, we have discovered an unusual interaction of Keap1 with Smad2/3 by parsing transcriptome sequencing, protein profiling, and co-immunoprecipitation data. Further examinations validated that Smad2/3 enable a

physical interaction with Keap1, including its full-length Keap1α and Keap1β (i.e., Keap1$^{\Delta 1\text{-}31}$)[24], by both EDGETSD and DLG motifs located in the linker region of Smad2 or Smad3, so that they are segregated in the cytoplasmic compartments, but with their protein stability being enhanced by co-expression of Keap1 or its isoforms. Keap1α and Keap1β were also further demonstrated to monitor the TGF-β1-Smad2/3 signalling transduction of the integral signal from the cell surface to the nucleus. Such activation of Smad2/3 by Keap1, Keap1α or Keap1β was contributable to a putative competitively inhibitory effect of Nrf2, which, in turn, acts on the TGF-β1-Smad2/3 signalling. Overall, this discovery presents a novel functional bridge of the redox-sensor Keap1 with its interactors Smad2/3 involved in growth and development.

Results

Identification of two Keap1 isoforms α and β

As shown in Figure 1a, two molecular mass-closed protein isoforms of Keap1 were first found to have been differentially expressed in distinct cell lines. Such two electrophoretic bands of Keap1 were determined in HepG2, MHCC97L and THP-1 cell line, whereas only a single shorter Keap1 protein has been expressed in HL7702, MHCC97H, HCCLM3 and Heap3B cell lines (Fig. 1a). Thereafter, the HepG2 cell line was selectively used for all subsequent studies of two nuanced isoforms of Keap1. Next, distinctive experimental settings were conducted in order to verify whether the protein expression levels of Keap1 were affected by redox agents. Firstly, total lysates of HepG2 cells were incubated *in vitro* for 1 h with a vehicle control (Ctr, distilled water), H_2O_2 (0.5 M), or dithiothreitol[25] (DTT, 0.5 M) (Fig. 1b, *left panel*). Secondly, *in vivo* treatment of HepG2 cells with DTT (1 mM) for different lengths of time (0, 1, 2 or 4 h) and the total lysates were then collected (Fig. 1b, *right panel*). Another similar redox treatment of HepG2 cells that had been transfected with an expression construct for V5-tagged Keap1 was performed (Fig. 1c). These samples were subjected to Western blotting analysis and the results revealed that those two protein bands of Keap1 were all almost unaltered in their molecular sizes by DTT and H_2O_2 (Figs. 1b, 1c and S1a), no matter whether they were endogenously or ectopically expressed in distinct conditions.

To further determine distinct lengths of two independent Keap1 isoforms, a series of expression constructs for N-terminally-truncated Keap1 mutants were co-transfected, together with both Nrf2-expressing and *ARE*-driven luciferase plasmids, into HepG2 cells (Fig. 1, d to f). Western blotting results demonstrated that progressive truncation of Keap1's N-terminal 30 amino acids (aa) enabled its larger isoform (designated Keap1α) to become gradually shortened until this protein disappearance, whereas the mobility of its smaller isoform (called Keap1β) was roughly unaffected by such loss of Keap1's N-terminal 30-aa (Fig. 1d, *top panel*). By contrast, continuous truncation of Keap1's N-terminal 60-aa residues (to yield a Keap1$^{\Delta 2\text{-}60}$ mutant) led to disappearance of Keap1β (besides Keap1α), but it seemed to be replaced by another shorter polypeptide with a faster mobility than that of Keap1β (Fig. 2d, *top panel*). This Keap1$^{\Delta 2\text{-}60}$ mutant, rather than others examined, enabled for significant (but not complete) release from inhibition of Nrf2-mediated *ARE-Luc* activity by wild-type Keap1 (Fig. 2d, *lower panel*). Further examination of internal deletions of Keap1 within its N-terminal 34-aa unravelled that the expression of Keap1β, but not Keap1α, was completely prevented by three mutants Keap1$^{\Delta 31\text{-}34}$, Keap1$^{\Delta 31\text{-}32}$ and Keap1$^{\Delta 32\text{-}33}$ (Figs. 1e & 1f, *top panels*), with partial disinhibition of the Nrf2-mediated *ARE-Luc* activity as compared with the case of wild-type Keap1 (Figs. 1e & 1f, *lower panels*).

Collectively, the above mutagenesis analysis suggests that Keap1α is yielded from its first translation codon (UTG, encoding methionine), while Keap1β is likely translated from its internal UTG-starting codon between its N-terminal aa 31-33 (as illustrated in Fig. 1g). To confirm the latter notion, two point mutants of Keap1 at the internal translation start codon encoding Met^{32} into glycine or leucine residues (i.e., Keap1^{M32G} or Keap1^{M32L}, in Fig. 1h) were made. As expected, the results demonstrated that translational expression of Keap1β was totally blocked by those two mutants Keap1^{M32G} and Keap1^{M32L} (Fig. 1h, *upper panel*), but almost no obvious changes in Nrf2-mediated *ARE-Luc* activity as compared to that of its wild-type Keap1 (Fig. 1h, *lower panel*). To gain insight into functional distinctions of Keap1α with Keap1β, the Met^{32} residue was deleted to yield a single protein of Keap1α (i.e., Keap1$^{\Delta M32}$), while another single Keap1β was allowed for translational expression upon deletion of the N-terminal 31-aa residues of Keap1 (i.e., Keap1$^{\Delta 1\text{-}31}$, Fig. 1g). The results indicated that such two single proteins of Keap1α and Keap1β were enabled for their respective independent expression

with differential ability to inhibit Nrf2-mediated *ARE-Luc* activity (Fig, 1I, *upper and lower panels*).

Distinct effects of between Keap1α and Keap1β on Nrf2 and target genes

Clearly, when unstimulated, Keap1 dimers physically binds to Nrf2 and targets this CNC-bZIP protein to the ubiquitin-proteasomal degradation *via* Cullin3[8,26]. After Keap1 is stimulated by oxidative stress, Nrf2 is released from Keap1 and thus translocated into the nucleus before transcriptionally regulating ARE-battery genes, whereas Keap1 is also allowed for another interaction with p62 so to be targeted for autophagy-based degradation[27-29]. Since being the case, a model (in Fig. 2a) is proposed to provide a possible explanation of three putative dimers of Keap1α and/or Keap1β, as involved in the signalling response to redox stress[4-7,30].

The stability of Keap1, Keap1α and Keap1β was determined by their half-lives in HepG2 cells, that been transfected with their respective expression constructs and then treated with CHX (50 μg/ml) for 0 to 48 h (Fig. 2b). The pulse-chase experiments revealed distinctive half-lives of Keap1, Keap1α and Keap1β, which were estimated to be 9.4 h, 6.4 h, or 28.9 h, respectively, following CHX treatment of cells (Fig. 2c). Subsequent examination unravelled that endogenous Nrf2 stability was slightly affected by ectopic expression of Keap1, Keap1α or Keap1β (Fig. 2d), because the half-life of this CNC-bZIP protein was evaluated to be 0.7 h, 0.5 h or 0.6 h after CHX treatment of cells expressing Keap1, Keap1α or Keap1β, respectively (Fig. 2e). Further Western blotting analysis showed that Keap1β appeared to enable for modest release of confined Nrf2, as well as its targets GCLC, GCLM and HO-1, to certain extents, when compared to the inhibitory effects of Keap1 and Keap1α (Fig. 2f, *left panels*). Similar results were also obtained from the cells that had been allowed for co-expression of ectopic Nrf2 with Keap1, Keap1α or Keap1β (Fig. 2f, *right panels*). Moreover, real-time quantitative PCR demonstrated differential inhibition of the mRNA expression levels of Nrf2-target genes *GCLC, GCLM* and *HO-1* by co-expressing Keap1, Keap1α or Keap1β, but with a certain exception of being released by Keap1α or Keap1β, to varying extents (Fig. 2g). Similar effects of Keap1α and Keap1β on ectopic Nrf2-regulated genes were also shown (in Fig. S1b).

Next, confocal imaging illustrated co-localization of V5-tagged Keap1, Keap1α or Keap1β with cortactin (CTTN, an actin-binding cytoskeletal protein, Fig. 2h), as reported previously in the cytoplasm of COS-1 cells[31]. Further subcellular fractionation of distinct genotypic cell lines *Keap1*$^{+/+}$, *Keap1*$^{-/-}$ and Keap1β(*Keap1*$^{\triangle 1\text{-}31}$), along with two stably *Keap1*- or *Keap1α*-expressing cell lines (all derived from HepG2 cells[24]) unravelled that Keap1, Keap1α and Keap1β were recovered primarily in the cytosolic and mitochondrial factions, and only modestly in nuclear fractions (Figs. 2i & S1c). This suggests that they are also likely involved in the nuclear regulatory process, in addition to being responsible for the cytoplasmic and mitochondrial events.

All the inhibitory effects of Keap1 on Nrf2 and its target genes *GCLC, GSR, GCLM, NQO-1* and *HO-1* were significantly attenuated by knockout of *Keap1*$^{-/-}$ (Fig. S1d), whereas Keap1β only retained marginal inhibitory effects on these genes when compared with those controls measured from *Keap1*$^{+/+}$ cells. However, their results for Nrf1 were the opposite. Further examination revealed that, when *Keap1-restored* or *Keap1α-restored*, all those inhibitory effects of Keap1 on Nrf2-mediated downstream genes were recovered from *Keap1*$^{-/-}$, but Nrf1 appeared to be upregulated (Fig. S1e). Of note, Keap1-restored cells had exerted stronger inhibitory effects than those of Keap1α-restored cells. Such distinction in the inhibitory effects of Keap1, Keap1α or Keap1β on Nrf2 and downstream genes were mainly attributable to their tempo-spatial co-localization in the cytoplasmic (and even nuclear) compartments (Figs. 2i & S1c).

Differential gene expression profiles monitored by Keap1α and Keap1β

Distinct genotypic cell lines *Keap1*$^{+/+}$, *Keap1*$^{-/-}$, Keap1β(*Keap1*$^{\triangle 1\text{-}31}$), *Keap1-restored* and *Keap1α-restored*, that had been confirmed by Western blotting (Fig. 3a) and others described previously[24], were subjected to transcriptome sequencing so as to identify differentially regulated gene expression profiles. As shown in a Venn diagram (Fig. 3b), a total of 1592 differentially expressed genes (DEGs) were identified to be the common or unique among four examined cell lines (Table S1). The DEGs were screened according to the following criteria: fold change ≥ 2 or ≤ 0.5, plus false discovery rate (FDR) ≤ 0.001 (when compared with equivalent controls measured from *Keap1*$^{+/+}$ cells). By pairwise comparisons of indicated

four groups (i.e., *Keap1*$^{-/-}$, Keap1β(*Keap1*$^{\triangle 1\text{-}31}$), *Keap1-restored* and *Keap1α-restored*), those common and different regulatory genes in between each two cell lines were selected (Fig. 3c), to verify differential expression genes monitored specifically by Keap1, Keap1α or Keap1β signalling as were subjected to the detailed KEGG pathway and GO enrichment analyses[24].

Consequently, differences in the number of increased or decreased DEGs specifically in each cell line were shown graphically (Fig. S2). In 297 DEGs detected specifically in *Keap1*$^{-/-}$ ells, 167 genes were upregulated, as accompanied by 130 genes downregulated, by loss of Keap1. By contrast, the number of specifically upregulated DEGs was significantly reduced to 55 in *Keap1-restored* cells, with 13 specific DEGs downregulated. Notably, *Keap1α-restored* cells also yielded 143 specific DEGs (i.e., 79 upregulated and 64 downregulated), whereas Keap1β*(Keap1*$^{\triangle 1\text{-}31}$*)* cells had 166 specific DEGs as identified, of which 100 genes were upregulated (Fig. S2, Table S2). These genotypic cell line-specific DEGs possessed by *Keap1*$^{-/-}$*, Keap1-restored, Keap1α-restored* and *Keap1β,* respectively, were subjected to in-depth informatics analysis by means of the top 10 KEGG pathways and GO enrichment methods. The results uncovered that Keap1 and its isoforms α and β were markedly different in terms of the regulated KEGG pathway and GO enrichment analyses (Fig. S3).

The top 20 common and differential genes were selected in each indicated cell line to map these regulatory targets within their functional association networks (Fig. 3d). Based on the transcriptome data, it is inferable that the expression of most of those genes was consistent (Fig. 3d), but only a few of genes (e.g., *CAMK2D, B4GALNT1, KRT4*, *MAP2*) showed to be expressed in inconsistent expression trends (Fig. S4, TableS3). Next, from those three aspects of biological process, cellular component and molecular function, GO enrichment function analysis on DEGs, 1301 gene products were found to associate with (and even bind to) Keap1 (Fig. S5, Table S4). For example, Keap1 was found to bind Smad2. The results of Keap1-putdown protein profiling were subjected to GO enrichment analysis, so that it was shown that Keap1 enabled to bind approximately 2450 proteins detected herein; among these possibly binding proteins was Smad2 (Fig. S6, Table S5). In the transcriptome data, Smad2 was also presented among the total 1592 DEGs (Fig. 3b, Table S1), but it is, to our surprise, it exists as one of the 50 common DEGs for *Keap1*$^{-/-}$ and *Keap1α-restored* cells (Fig. 3b, Table S6). By analysis of peptide and protein profiling data (Tables S7, S8), Smad2 was identified to be represented by its peptide sequence VETPVLPPVLVPR, whereas Smad3 was also identified by the spectral identification (Table S8). According to the BGI gene systemic (Dr Tom) association clustering analysis of Keap1, Smad2 and Smad3 by top 10 KEGG pathways, it was unveiled that they were dominantly concentrated on the TGF-β1 signalling pathway, ubiquitin-mediated proteolysis, fluid shear stress, atherosclerosis, the cell cycle and Wnt signalling pathways (Fig. S7). By the top 10 GO enrichment analysis, their biological functions were also primarily concentrated on the TGF-β1 pathway with related functions (Fig. 3e).

Intriguingly, approximately 51 DEGs were identified by their association with all four cell lines, and their differential expression levels are presented in a hierarchically clustered heatmap (Fig. 3f). Among such association-clustering DEGs, only seven genes, such as *ZFYVE9, BMPR1A, TRIM33, CUL1, UBE2O, HMOX1/HO-1* and *NFE2L2*(*Nrf2)*, were upregulated or downregulated, with each having similar expression trends, in all four cell lines (Fig. 3f). Of striking note, other nine genes, including *SQSTM1/p62, JUN, ACVR1B, BRE2S, UBE2E2, CAND1, SMURF2, ACVR2B* and *Keap1*, were upregulated or downregulated *in Keap1*$^{-/-}$ cells, but their opposite results were obtained from *Keap1-restored, Keap1α-restored* and *Keap1β(Keap1*$^{\triangle 1\text{-}31}$*)* cell lines (Fig. 3f). Further comparison of the latter three cases revealed that the upregulation and/or downregulation of those examined genes by restoration of Keap1 or Keap1α appeared to be largely in similar fashions, whereas 33 of those genes were downregulated by Keap1β (Fig. 3f).

Keap1, Keap1α or Keap1β interacts with two conserved motifs of Smad2/3

Although Smad2 and Smad3 are known to have highly conserved structures and similar biological functions[14-16], it was also, much to our surprise, found by amino acid sequence alignment that two Keap1-binding motifs ETGE and DLG within the Neh2 domain of Nrf2 are represented by additional two homologues EDGETSD and DLQ located within the linker regions of Smad2 and Smad3, respectively (Fig. 4a). By the CABS-dock modelling, it was illustrated that the Keap1-binding peptides could enable it to successfully dock with Smad2 (i.e., YISEDGETSD and NHSLDLQPV) or Smad3

(i.e., YLSEDGETSD and HNNLDLQPV) (Fig. 4b).

Confocal imaging of immunocytochemistry showed that the red fluorescent signals representing V5-tagged Keap1, Keap1α or Keap1β were superposed with flag-Smad2/3 exhibiting green fluorescent signals primarily in the cytoplasm, but modestly superposed images of Keap1β with Smad2 were also presented in the nucleus (Figs. 4c, 4d). The putative interaction of Keap1, Keap1α or Keap1β with Smad2/3 were further corroborated by a series of co-immunoprecipitation assays, revealing that flag-Smad2/3 enabled for immunoprecipitation of V5-tagged Keap1, Keap1α or Keap1β (Figs. 4e, 4f) and *vice versa* (i.e., V5-tagged Keap1, Keap1α or Keap1β also enabled Smad2 or Smad3 to put down, Fig. S8). Further mutagenesis mapping of putative Keap1-binding EDGETSD and DLQ motifs in the linker region of Smad2/3 revealed that the resulting mutants caused their protein expression levels to be obviously decreased (Fig. 4g), particularly in the cases of Smad2$^{\triangle 228\text{-}264}$ and Smad3$^{\triangle 82\text{-}117}$. The latter two mutants lacked both the EDGETSD and DLQ motifs of Smad2/3 so that their immuneprecipitates were not put down by V5-tagged Keap1 (Fig. 4h), although the expression abundance of Keap1 was unaffected or even slightly enhanced (Fig. 4g). Collectively, these suggest that the stability of Smad2/3 may not only be modulated by their EDGETSD/DLQ-encompassing linker regions, and also monitored by their interacting Keap1.

Distinct impacts of Keap1, Keap1α and Keap1β on Smad2/3 and cognate genes

To further verify whether the protein stability of Smad2/3 was influenced by Keap1, Keap1α or Keap1β, a series of the CHX-pulse chase experiments were carried out (Figs. 4i to 4m and S9). Firstly, the stability of endogenous Smad2/3 was determined by measuring their half-lives that were estimated to be 1.4 h or 1.1 h, respectively, after CHX treatment of cells (Figs. 4i, 4j), while the half-life was evaluated to be 2.1 h (Fig. 4i, 4j). Secondly, it is very interesting to note that forced expression of Keap1 Keap1α or Keap1β caused the endogenous Smad2/3 half-lives to be, to different extents, prolonged (Fig. 4k). As it is indeed, the half-life of endogenous Smad2 was prolonged respectively by ectopic Keap1 to 3.1 h, Keap1α to 2.2 h, and Keap1β to 1.6 h (Fig. 4l), while the endogenous Smad3 half-life was also prolonged by ectopic Keap1 to 3.4 h, Keap1α to 1.7 h, and Keap1β to 1.3 h, respectively, after CHX treatment (Fig. 4m). Thirdly, the exogenous Smad2/3 half-lives were also prolonged by ectopic Keap1 and its isoforms α and β to be much longer than those of their endogenous proteins (Fig. S9). Just so ectopic Smad2 half-life was indeed extended by over-expressing Keap1, Keap1α or Keap1β to reach 4.3 h, 3.1 h, or 2.0 h, respectively (Figs. S9a, S9b), whereas ectopic Smad3 half-life was also extended by over-expressing Keap1, Keap1α or Keap1β to reach 4.0 h, 2.7 h or 1.7 h, respectively, after treatment of cells with CHX (Figs. S9c, S9d). Taken together, these demonstrate that the stability of Smad2/3 is enhanced distinctively by Keap1, as well as its isoforms α and β, because the latter three can differentially promote Smad2's half-life more strongly than that of Smad3.

Next, distinct Keap1-based genotypic cell lines were employed to explore the effects of Keap1, Keap1α or Keap1β on basal expression of Smad2/3-target genes. When compared with *Keap1*$^{+/+}$ cells, Smad2, Smad3, Smad4, TGF-β1 and E2F4 (the latter three had been identified as downstream genes of Smad2/3[32]) were down-expressed in *Keap1*$^{-/-}$ cells, but conversely, they were all up-expressed in Keap1β(*Keap1*$^{\triangle 1\text{-}31}$) cells (Fig. 5a). Similarly downregulated or upregulated changes in their mRNA expression levels of these examined genes were observed in *Keap1*$^{-/-}$ or Keap1β(*Keap1*$^{\triangle 1\text{-}31}$) cell lines, respectively (Fig. 5b). Upon the restoration of Keap1 or Keap1α, all those examined protein abundances and their mRNA expression levels were strikingly promoted, to different extents, in *Keap1-restored* and *Keap1α-restored* cell lines (Figs. 5c, 5d). The results also demonstrated the function of restored Keap1 appeared to be significantly stronger than that of its α isoform. Further *E2F4-Luc* assays revealed that its transcriptional expression was markedly upregulated by Smad2 or Smad3, but also manifested with the highest expression levels enhanced by Keap1 and the lowest expression levels elicited by Keap1β (Fig. 5e).

To further determined effects of Keap1 and its isoforms (α and β) on endogenous and exogenous Smad2/3, HepG2 cells were co-transfected with expression constructs for Keap1, Keap1α or Keap1β alone or plus Smad2 or Smad3 (Fig. S10). As anticipated, the results showed that distinct extents of increases in the exogenous and endogenous expression levels of Smad2, Smad3, Smad4, TGF-β1 and E2F4 were promoted by Keap1, Keap1α or Keap1β (Fig. S10, a to c). Among

them, their promotion of Keap1 was the strongest, followed by Keap1α, and Keap1β exerted the weakest or least effects. Next, to gain an insight into the effect of Keap1 on Smad2/3, both *Nrf1α*$^{-/-}$ (with diminished Keap1 abundance[33]) and *Nrf2*$^{-/-}$ (with enhanced Keap1 expression) cell lines were employed herein. The results unravelled that Keap1, Smad2, Smad3, Smad4, TGF-β1 and E2F4 were all decreased, at their protein abundances, by loss of *Nrf1α* (Fig. 5f, *left panel*), although their mRNA expression levels were only less or not altered (*right panel*). By contrast, they were all upregulated due to Keap1 activation by knockout of *Nrf2* (Fig. 5f, *left and right panels*). Collectively, these demonstrate that Keap1 contributes to positive regulation of Smad2/3 and cognate genes.

To further examine the redox-responsive effects of Keap1, Keap1α or Keap1β on the expression of Smad2/3 and cognate genes, distinct genotypic cell lines (*Keap1*$^{+/+}$, *Keap1*$^{-/-}$, *Keap1-restored*, *Keap1α-restored* and *Keap1β*(*Keap1*$^{\triangle 1\text{-}31}$)) were treated for 0 to 24 h with DTT (1 mM, as a strong reducing agent[25]) or 50 μM of *tert*-butylhydroquinone (tBHQ, as a pro-oxidative stressor and also widely-used Nrf2 activator[34]). As excepted, all the mRNA and protein express levels of Keap1, Nrf2, Nrf1, Smad2, Smad3, Smad4, TGF-β1 and E2F4 were significantly augmented in redox responsive to DTT (Fig. S11) or tBHQ (Fig. S12). Such redox-stimulated increases of Nrf1, Nrf2 and E2F4 were higher in *Keap1*$^{-/-}$ cells, whereas stimulated increases of Smad2, Smad3, Smad4, and TGF-β1 were observed in *Keap1β*(*Keap1*$^{\triangle 1\text{-}31}$) cells, when compared with their equivalent controls of *Keap1*$^{+/+}$ cells (Figs. S11 & S12). Conversely, restoration of Keap1 or Keap1α enabled redox-stimulatory increases of Nrf2 and Nrf1 (with an exception of Keap1α stimulated by DTT) to be reduced, but stimulated increases of Smad2, Smad3, Smad4, TGF-β1 and E2F4 to be promoted, differentially in *Keap1-restored* or *Keap1α-restored* cell lines. These indicate that redox-stimulated expression of Smad2/3 are also monitored by Keap1-indepdenpdent pathway, besides Keap1-dependent redox signalling. Overall, Keap1, together with its isoforms α and β, provides a novel functional crosstalk between the relevant redox signalling (e.g., mediated by Nrf2 and/or Nrf1) and its interactors Smad2/3 involved in the growth and development.

Discussion

In this study, we have identified that two Keap1 isoforms α and β are differentially expressed in distinct cell lines and also exerted discrepant effects on Nrf2 protein stability and transcriptional activity to regulate *ARE*-driven target genes (Fig. 5g). This finding just provides a novel explanation of why Nrf2, as a versatile chameleon-like regulator of antioxidant, detoxification, cytoprotective, and other genes in the selective responses to repressing or promoting cancer[35], must have to be highly tightly regulated by a finely-tuned control system (involving Keap1 and its isoforms α and β). Upon stimulation of cells by oxidative stress, chemopreventive compounds can block the activity of Keap1-Cul3-Rbx1 ubiquitin ligase[36], such that Nrf2 does not only circumvent the ubiquitin-mediated proteasomal degradation and is also precisely activated in order to properly regulate transcriptional expression of its downstream target genes. Thereby, it is inferable that the protein expression abundance of Nrf2 and its turnover after its functional action are all accurately fined tuned, to certain proper extents, by alternative translation of Keap1 into two isoforms α and β and by differential functioning of the ubiquitin-mediated proteasomal degradation system triggered by three distinct Keap1 dimers (αα, αβ, ββ). This is also modulated by its naturally-occurring dominant-negative mutant Keap1$^{\triangle C}$ lacking most essential portion of the C-terminally Nrf2-interacting Kelch/DGR domain[37].

Considering the strict specificity of the relationship between Keap1 and Nrf2[38-40], it is hypothesized that the Keap1 isoforms α and β can also exert their intrinsic inhibitory effects on Nrf2 by physical interaction with the latter ETGE and DLG motifs in order to target this CNC-bZIP protein to the ubiquitin-mediated proteasomal degradation pathway. In fact, Keap1α is the full-length prototypical Keap1, whereas Keap1β arises from the secondary internal translation start codon to yield an N-terminally 31aa-truncated mutant Keap1$^{\triangle 1\text{-}31}$. Such nuanced two Keap1 isoforms α and β have disparate half-lives, which are determined to be significantly shortened (from 9.4 h to 6.4 h) or prolonged (from 9.4 h to 28.9 h), respectively. It should also be noted that 9.4 h is the half-life of ectopic Keap1, but that of endogenous Keap1 is only 2.1 h). From these, taken together with our confocal imaging and subcellular fractionation data, it is thus reasoned that the three putative functional dimers of Keap1 (αα, αβ, and ββ) are manifested with its interactors (e.g., Nrf2, p62, Smad2/3,

and others[41], and that they enable to distribute tempo-spatially in distinct subcellular (i.e., cytosolic, mitochondrial and nuclear) locations and hence exert distinctive regulatory effects and relevant biological functions. Yet, it warrants further study of the mechanistic distinction in the subcellular compartmentation of Keap1α and Keap1β to be tethered.

As expected, several lines of experimental evidence have been provided revealing that Keap1, Keap1α and Keap1β have exerted similar, but yet different, inhibitory effects on those examined Nrf2-target genes, although they all enable to physically interact with this CNC-bZIP factor[24]. Such differential expression genes (DEGs) monitored by Keap1, Keap1α and Keap1β are scrutinized by the transcriptome sequencing of distinct Keap1-based genotypic cell lines (i.e., *Keap1*$^{+/+}$, *Keap1*$^{-/-}$, *Keap1-restored*, *Keap1α-restored* and *Keap1β*(*Keap1*$^{\triangle 1\text{-}31}$)). Then, similarities and differences between Keap1, Keap1α and Keap1β in governing those gene expression profiles were further parsed by detailed KEGG pathway and GO enrichment analyses. Of striking note, a surprising association of Keap1 with Smad2 was, for the first time, found herein, and further validated by mass spectrometry analysis of Keap1-putdown interactome, implying there exists a competitive relevance of the Keap1-Nrf2 signalling with the TGFβ-Smad2/3 pathways (Fig. 5g). For instance, Nrf2 can inhibit the TGF-β1 expression of in a human stellate cell line[22], but conversely TGF-β1 can also enable downregulation of Nrf2 in another rat hepatic stellate cell line[23].

Subsequently, further experimental evidence has also been herein provided substantiating that Keap1, as well as its isoforms α and β, enables for a physical interaction with the two highly conserved EDGETSD and DLQ motifs of Smad2/3 (respectively compared with the ETGE and DLG motifs in Nrf2 enabling it to bind a dimeric Keap1[42,43]), which are located within the central linker regions of Smad2/3 between the N-terminal MH1 (binding target genes) and C-terminal MH2 (forming a functional heterodimer with Smad4, as illustrated in Figs. 5h). Notably, such an interact of Keap1, Keap1α or Keap1β with Smad2 or Smad3 was corroborated by their co-immunoprecipitation and co-localization assays. These collective data demonstrate that the redox sensor Keap1 is likely involved in cross-talking with the Smad2/3-mediated signalling (Fig. 5g), although this pathway is triggered predominantly by the TGF-β family of growth factors[44]. In this signalling cascades, TGF-β1 stimulates activation of Smad2/3 and in turn, Smad2/3 had also been shown to be central to the expression of both TGF-β1 and Smad4 (as an co-binding partner of Smad2/3 recruited their cognate target genes)[45], so as to comprise a positive feedback regulatory circuit.

Importantly, we have also found that the protein stability of Smad2/3 and its transcription activity are markedly enhanced to different extents by Keap1, Keap1α or Keap1β, although they enable both Nrf2 stability and *trans*-activity to be significantly reduced. This notion is supported by evidence revealing that the half-lives of endogenous Smad2 and Smad3 proteins are determined to be 1.4 h or 1.1 h, respectively, after CHX treatment, but ectopic Smad2 half-life is prolonged to 3.1 h, 2.2 h or 1.6 h, while ectopic Smad3 half-life is also extended to 2.4 h, 1.7 h or 1.3 h, by ectopically-expressing Keap1, Keap1α or Keap1β, respectively. Such opposing effects of Keap1, Keap1α or Keap1β on Smad2/3 and Nrf2 demonstrate a significant different mechanism accounting for Smad2/3 turnover from the Nrf2-controlling mechanism, albeit this remains to be further elucidated. Anyways, it is inferable that putative conformation changes in the Smad2/3's functional complexes are likely to result from physical interaction of dimeric Keap1 with highly conserved EDGETSD and DLQ motifs in their central linker regions. This is due to the facts that this linker region includes a transactivation domain[46], enables homo- or hetero-oligomerization and provides critical phosphorylation sites by GSK3β, MAPKs, CDKs and calcium-calmodulin dependent kinases, such that it can exert key regulatory roles in governing Smad2/3 activity and function[47,48]. For example, the linker region phosphorylation cannot only regulate the inhibition of the nuclear translocation of Smad2/3, and also influences their protein degradation by the ubiquitin-proteasomal system[11,49]. This is also further evidenced by our experiments showing that the linker region is essential for the protein stability of Smad2/3.

In addition to its interactors Smad2/3, Keap1-putdown interactome analysis also revealed that it may act as a key player in the process to sequentially add and remove ubiquitin by dynamic formation of distinctive functional complexes with ubiquitin conjugating enzymes (e.g., Cullin-1, 2, 3 , 4B, 5, ATG3, UBE3A, and UBE3C) and deubiquitinating enzymes (e.g., USP5, 7, 10, 12, 14, 15, 17L, 22, 24, 47 and UCHL5), as well as with cullin-associated NEDD8-dissociated protein 1

(CAND1, as an inhibitor promoting dissociation of substrate receptor components from the Cullin RING ligases[50]). For this, it inferable that the selective dynamic formation of such opposite functional complexes is much likely to depend on distinctive tempo-spatial contexts of between Smad2/3 and Nrf2. Thereby, based on our experimental data, its inferable that after inhibiting the expression of Nrf2, Keap1 activates Smad2/3 and promotes the expression of TGF-β1, Smad4 and relevant targets such as E2F4. This is further evidenced by our data obtained from *Nrf1α*$^{-/-}$ (with diminished Keap1 abundance) and *Nrf2*$^{-/-}$ (with enhanced Keap1 expression levels). In the meantime, it is intriguing to note that a table of Smad2/3-target genes were also activated, upon redox stimulation by DTT and tBHQ (as an Nrf2-inducer), implying a coordinated mechanism also existing between Nrf2 and Smad2/3.

In summary, we have herein discovered two nuanced Keap1 isoforms α and β that arise from alternative translation starting sites within its mRNA transcripts and also interact with Smad2/3, besides Nrf2. Their similarities and differences in monitoring relevant gene expression profiles have been interrogated by transcriptome sequencing. Subsequent series of evidence unravelled distinct and even opposing effects of Keap1, as well as its isoforms α and β, on Nrf1 and Smad2/3 at their protein stability and transcriptional activity to mediate distinct sets of target genes. Overall, this study presents a novel functional bridge of Keap1, Keap1α or Keap1β crossing both the redox-responsive Nrf2 and the developmental TGFβ1-Smad2/3 signalling pathways. Thereby, this discovery also provides a novel helpful understanding of the 'double-edged sword' effects of Keap1-Nrf2 or TGFβ1-Smad2/3 signalling on paradoxically suppressing or promoting cancer and other diseases.

Materials and Methods

Chemicals and antibodies

All chemicals were of the best quality commercially available. Cycloheximide (CHX, 50 μg/ml) was purchased from Sigma–Aldrich (St Louis, MO, USA), whereas DTT and tBHQ were from MedChemExpress (MCE). The antibody against Nrf1 proteins was saved in our own lab (as previously described[51]). Anti-V5 and Anti-Flag monoclonal antibodies were from Invitrogen, with all other primary antibodies from Abcam. β-actin antibody, as well as the secondary antibodies, were from ZSGB-BIO (Beijing, China).

Cell lines and transfection

Four distinct genotypic cell lines (*Kepa1*$^{-/-}$, *Keap1-Restored, Keap1α-Restored* and *Keap1β(Keap1*$^{\triangle 1\text{-}31}$*)*) were created from wild-type HepG2 cells (i.e., *Keap1*$^{+/+}$) described elsewhere[24]. They were all allowed to grow in DMEM supplemented with 10% (v/v) foetal bovine serum (FBS), 5 mM glutamine, and 100 units/mL penicillin and streptomycin at 37 °C in an incubator with 5% CO_2. The cells were transfected for 8 h with indicated plasmids in Lipofectamine 3000 (Invitrogen) and then allowed for 24-h recovery from transfection by changing a fresh medium before subsequent experiments.

Expression Constructs

Those expression constructs for human Keap1, Nrf2, CTTN, Smad2 and Smad3 were made by cloning each of their full-length cDNA sequences into a pcDNA3 or 3xFlag vector. The N-terminal 32nd methionine of the full-length Keap1 was mutated (or deleted) to yield only Keap1α-expressing plasmids, whereas the first N-terminal 31 amino acids of the full-length Keap1 were deleted to yield the Keap1β plasmid. The other plasmids specifically for the genome editing of Keap1, and Smad2/3 mutants were made and identified (as shown in Figs. 1d-h & 4g). In addition, *E2F4*-promoter luciferase reporter was also created herein. All plasmids were confirmed by sequencing and validated by relevant experiments.

Construction of *Keap1*$^{-/-}$, *Keap1β*(*Keap1*$^{\triangle 1\text{-}31}$), *Keap1-Restored* and *Keap1α-Restored* cell lines

Both Keap1 (sgRNA-Keap1: 5'-TATGAGCCAGAGCGGGATG-3') and Keap1α (sgRNA-Keap1α: 5'-AGGCCTAGCGGGGCTGGG GC-3') genomes had been created *via* CRISPR/Cas9 and identified[24]. Then, each of those Cas9 plasmids for Keap1 and Keap1α were introduced into HepG2 cells to construct *Keap1*$^{-/-}$ and *Keap1β(Keap1*$^{\triangle 1\text{-}31}$*)* cell lines (Fig. S1c). Thereafter,

the full-length cDNA sequences of the Keap1 plasmid and Keap1α plasmid (F: CGCGGATCCATGCAGCCAGATCCCAGG, R: AAATATGCGGCCGCTCATAGCCTCCTCTCCACACT) were subcloned into the pLVX-IRES-ZsGreen-puro vector for lentivirus packaging, which were verified by gene sequencing (Fig. S1d). These two viruses were prepared in 293T cells with three plasmids (2.52 μg pMD2G, 8.34 μg psPAX2, and 10.5 μg pLVX-mcmv-ZsGreen-puro pKeap1/pKeap1α) according to a previously-described method[52]. Either of Keap1 and Keap1α viruses was infected into *Keap1*$^{-/-}$cells, and then screened in puromycin-containing media to select the positively-infected cells. The Keap1- or Keap1α-restored cells were verified by Western blotting.

RNA isolation and quantitative real-time PCR (qPCR)

Experimental cells were subjected to total RNA isolation (using an RNA Simple Total RNA Kit, DP419, Tiangen Biotech Co., Ltd., Beijing, China). The first strand of cDNA was obtained by adding total RNA (1.5 μg) in a reverse transcriptase reaction (using a Revert Aid First Strand cDNA Synthesis Kit, from Thermo), and served as the template for quantitative PCR (using GoTaq qPCR Master Mix, A6002, Promega, USA). Then, each pair of those reciprocal primers (listed in Table 1) was also added to the indicated qPCR, which was carried out under the following conditions: 95 °C for 5 min, followed by 40 cycles of 15 s at 95 °C and 30 s at 60 °C. The mRNA expression level of β-actin was used as the optimal internal standard control.

Table 1. Primer pairs used for the RT–qPCR analysis

ID	Name	Forward sequences (5'-3')	Reverse sequences (5'-3')
60	β-Actin	TGGCATCCACGAAACTACCTT	CTTCTGCATCCTGTCGGCAAT
9817	Keap1	CGTGGCTGTCCTCAATCGTCT	ATTGCTGTGATCATTCGCCACT
4780	Nrf2	GAGAGCCCAGTCTTCATTGCTA	CCGTCTAAATCAACAGGGGCTA
4776	Nrf1	TTTGAAGCCCACCAAGACCGAA	CTGCCTCTTCCTGTACACTGACC
2729	GCLC	TCAATGGGAAGGAAGGTGTGTT	TCAATGGGAAGGAAGGTGTGTT
2936	GSR	CACGAGTGATCCCAAGCCC	CAATGTAACCTGCACCAACAATG
1728	NQO1	AAGAAGAAAGGATGGGAGGTGG	GAACAGACTCGGCAGGATACTG
2730	GCLM	TCAATGGGAAGGAAGGTGTGTT	CGCTTGAATGTCAGGAATGCTT
3162	HO-1	CAGAGCCTGGAAGACACCCTAA	AAACCACCCCAACCCTGCTAT
4087	Smad2	TACTCTCCAATGTTAACCGAA	TATGTAGTATAAGCGCACTCC
4088	Smad3	TCCCCGAAAACACTAACTTCC	TCCATCTTCACTCAGGTAGCCA
4089	Smad4	AAATATTGTCAGTATGCGTTT	TACTTGATGGAGCATTACTCT
7040	TGF-β1	TGTATTTAAGGACACCCGTGCC	AATGACACAGAGATCCGCAG
1874	E2F4	CTCCCCAAAGAGCTGTCAGA	GCTCATGCACTCTCGTGT

Luciferase Reporter Assay

Equal numbers (1.5 × 10^5) of experimental cells were seeded in each well of 12-well plates. After reaching 70-80% confluence, the cells were transfected with luciferase plasmids alone or plus other expression plasmids in the Lipo3000 reagent, in which the *pRL-TK* plasmid served as an internal control for transfection efficiency. The luciferase activity was determined by the Dual-Lumi™ double luciferase reporter gene assay (RG088M, Beyotime).

Western Blotting analysis

After experimental cells were rinsed three times with PBS and harvested in a total protein extraction buffer, the total lysates were denatured immediately for 10 min at 100 °C, and then the eluted proteins were subjected to separation by SDS–PAGE. The indicated protein abundances were visualized by Western blotting with different primary and secondary antibodies.

Subcellular fractionation

Experimental cells (6×10^5) were allowed for growth in 6-cm cell culture plates for 48 h and subjected to isolation of the cytoplasmic and nuclear fractions by using relevant separation kit (N-3408-200ML, Sigma). During this procedure, the extracted cytoplasmic portion was further centrifuged again at 20,000 g for 10 min to obtain precipitated mitochondrial fraction, while the supernatant was viewed as cytosolic fraction. All these subcellular fractions were further evaluated by Western blotting with specific antibodies against indicated proteins and relevant compartmental markers

Transcriptome sequencing and protein profiling

Experimental cells were subjected to total RNA isolation using an RNA Simple Total RNA Kit, followed by transcriptome sequencing (Beijing Genomics Institute (BGI), Shenzhen, China) to obtain the relevant data from an Illumina HiSeq 2000 sequencing system (Illumina, San Diego, CA). All differentially expressed genes (DEGs) were identified. The standard fold change was ≥2 or ≤0.5, with FDR (false discovery rate) ≤0.001 determined by the Poisson distribution model method (PossionDis). The total proteins of COS-1 cells, that had been transfected with expression constructs for Keap1, Keap1α or Keap1β, were extracted and put down by Keap1-specific antibodies, followed by mass spectrometry analysis (BGI) to obtain the protein profiling data.

The CABS-dock structural modelling

The CABS-dock web server (http://biocomp.chem.uw.edu.pl/CABSdock/) was here used for the flexible protein-peptide docking, which enables a full flexibility of the peptide structure with large-scale flexibility of the protein during the search for their binding site. The two putative KEAP1-binding peptides within SMAD2 (YISEDGETSD and NHSLDLQPV) or SMAD3 (YLSEDGETSD and HNNLDLQPV) were selected for docking into KEAP1 (PDB code: 1ZGK) within default settings. The ten top ranked CABS-dock models were analysed.

Co-immunoprecipitation analysis

COS-1 cells (4.0×10^5), that were grown on 6-well cell culture plates for 24 h, were transfected with different expression plasmids for Keap1-V5, Keap1α-V5, Keap1β-V5, Smad2-Flag or Smad3-Flag, and then allowed for 24-h recovery in the normal medium. The cells were lysed in RIPA buffer (C1053, PPLYGEN) containing protease and phosphatase inhibitors (Solarbio, P6730). The total lysates were centrifuged to collect the supernatants, followed by co-immunoprecipitation with antibodies against V5 or FLAG epitopes and the BeyoMag™ Protein A+G beads[53] (Beyotime. P2108-1ML). The eluted proteins were determined by Western blotting with anti-V5 and anti-FLAG antibodies (Invitrogen) and visualized by the horseradish peroxidase-conjugated secondary antibodies (ZSGB-BIO, ZB-2305), along with being developed by enhanced chemiluminescence reagents[54] (Thermo, iBright 750).

Immunocytochemistry and confocal microscopy

COS-1 cells (2.5×10^5) were grown for 24 h on 6-well cell culture plates (with the glass coverslips of 1 cm^2 c being stored flat) and co-transfected for 8 h with expression constructs for Smad2/3-Flag plus Keap1-V5, Keap1α-V5 or Keap1β-V5. The cells were allowed for 24-h recovery, then washed in PBS and fixed with paraformaldehyde fixative (Servicebio) for 30 min at room temperature. After the fixed cells were rinsed three times with PBS, they were permeabilized with 0.2% Triton X-100 for 20 min, before immunocytochemical staining. The samples were further rinsed again three times with PBS, blocked with 1% BSA for 60 min, incubated overnight with each of indicated primary antibody and then re-incubated for 4 h with the secondary Alexa Fluor–conjugated secondary antibody (ZSGB-BIO). The antibodies-stained coverslips were mounted with DAPI (Beyotime, C1005), and subjected to confocal imaging by using an IN Cell Analyser ZeissLSM900 cellular imaging system[31].

Statistical analysis

The statistical significances of differential gene expression levels measured by real-time qPCR and luciferase activity were determined using Student's *t*-test or two-way ANOVA. All the relevant data were obtained from at least three independent experiments, and shown as a fold change (mean ± SEM or ±SD) with significant differences being calculated by the value of $p < 0.01$ when compared with controls). The transcriptome sequencing data were also subjected to statistical analysis described previously[52].

Author Contributions: Both F.C. and M.X. performed the experiments with help of R.W. and S.H., collected all the relevant data, and wrote a draft of this manuscript with most figures and supplemental information. D.L. modelled the CABS-dock structural interaction of Keap1 with Smad2/3. Q.W. helped to analyse the big data from transcriptome sequencing and protein profiling. Y.W. and Z.Z. provided invaluable and critical discussion. Z.Z. also edited and corrected this paper in standard English. And, Y.Z. designed and supervised this study, analyzed all the data, helped to prepare all figures with cartoons, rewrote and revised the paper. All coauthors have read and agreed to the published version of the manuscript.

Acknowledgments: We are greatly thankful to all the other present and past members of Zhang's laboratory for giving critical discussion and invaluable help with this work. This study was funded by the National Natural Science Foundation of China (NSFC, with two projects 81872336 and 82073079) awarded to Prof. Yiguo Zhang (at Chongqing University). This is also, at part, supported by the Initiative Foundation of Jiangjin Hospital affiliated to Chongqing University (2022qdjfxm001).

Conflicts of Interest: The authors declare no conflict of interest.

References

1 Zhou, X. L., Zhu, C. Y., Wu, Z. G., Guo, X. & Zou, W. The oncoprotein HBXIP competitively binds KEAP1 to activate NRF2 and enhance breast cancer cell growth and metastasis. *Oncogene* **38**, 4028-4046, doi:10.1038/s41388-019-0698-5 (2019).

2 Panieri, E., Telkoparan-Akillilar, P., Suzen, S. & Saso, L. The NRF2/KEAP1 Axis in the Regulation of Tumor Metabolism: Mechanisms and Therapeutic Perspectives. *Biomolecules* **10**, doi:10.3390/biom10050791 (2020).

3 Tian, W., Rojo de la Vega, M., Schmidlin, C. J., Ooi, A. & Zhang, D. D. Kelch-like ECH-associated protein 1 (KEAP1) differentially regulates nuclear factor erythroid-2-related factors 1 and 2 (NRF1 and NRF2). *J Biol Chem* **293**, 2029-2040, doi:10.1074/jbc.RA117.000428 (2018).

4 Kobayashi, A. *et al.* Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate proteasomal degradation of Nrf2. *Mol Cell Biol* **24**, 7130-7139, doi:10.1128/MCB.24.16.7130-7139.2004 (2004).

5 Liu, Y. *et al.* TRIM25 promotes the cell survival and growth of hepatocellular carcinoma through targeting Keap1-Nrf2 pathway. *Nat Commun* **11**, 348, doi:10.1038/s41467-019-14190-2 (2020).

6 Leinonen, H. M., Kansanen, E., Polonen, P., Heinaniemi, M. & Levonen, A. L. Role of the Keap1-Nrf2 pathway in cancer. *Adv Cancer Res* **122**, 281-320, doi:10.1016/B978-0-12-420117-0.00008-6 (2014).

7 Vriend, J. & Reiter, R. J. The Keap1-Nrf2-antioxidant response element pathway: a review of its regulation by melatonin and the proteasome. *Mol Cell Endocrinol* **401**, 213-220, doi:10.1016/j.mce.2014.12.013 (2015).

8 Zheng, Y. H. *et al.* A novel Keap1 inhibitor iKeap1 activates Nrf2 signaling and ameliorates hydrogen peroxide-induced oxidative injury and apoptosis in osteoblasts. *Cell Death Dis* **12**, 679, doi:10.1038/s41419-021-03962-8 (2021).

9 Schmidlin, C. J., Dodson, M. B., Madhavan, L. & Zhang, D. D. Redox regulation by NRF2 in aging and disease. *Free Radic Biol Med* **134**, 702-707, doi:10.1016/j.freeradbiomed.2019.01.016 (2019).

10 Zhang, H., Davies, K. J. A. & Forman, H. J. Oxidative stress response and Nrf2 signaling in aging. *Free Radic Biol Med* **88**, 314-336, doi:10.1016/j.freeradbiomed.2015.05.036 (2015).

11 Derynck, R. & Budi, E. H. Specificity, versatility, and control of TGF-beta family signaling. *Sci Signal* **12**, doi:10.1126/scisignal.aav5183 (2019).

12 Nusse, R. & Clevers, H. Wnt/beta-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* **169**, 985-999, doi:10.1016/j.cell.2017.05.016 (2017).

13 Cai, C., Thorne, J. & Grabel, L. Hedgehog serves as a mitogen and survival factor during embryonic stem cell neurogenesis. *Stem Cells* **26**, 1097-1108, doi:10.1634/stemcells.2007-0684 (2008).

14 Derynck, R., Zhang, Y. & Feng, X. H. Smads: transcriptional activators of TGF-beta responses. *Cell* **95**, 737-740, doi:10.1016/s0092-8674(00)81696-7 (1998).

15 Brown, K. A., Pietenpol, J. A. & Moses, H. L. A tale of two proteins: differential roles and regulation of Smad2 and Smad3 in TGF-beta signaling. *J Cell Biochem* **101**, 9-33, doi:10.1002/jcb.21255 (2007).

16 Fleming, N. I. *et al.* SMAD2, SMAD3 and SMAD4 mutations in colorectal cancer. *Cancer Res* **73**, 725-735, doi:10.1158/0008-5472.CAN-12-2706 (2013).

17 Bottinger, E. P. & Bitzer, M. TGF-beta signaling in renal disease. *J Am Soc Nephrol* **13**, 2600-2610, doi:10.1097/01.asn. 0000033611. 79556.ae (2002).

18 Bathish, B., Robertson, H., Dillon, J. F., Dinkova-Kostova, A. T. & Hayes, J. D. Nonalcoholic steatohepatitis and mechanisms by which it is ameliorated by activation of the CNC-bZIP transcription factor Nrf2. *Free Radic Biol Med* **188**, 221-261, doi:10.1016/j.freeradbiomed.2022.06.226 (2022).

19 Rachakonda, G. *et al.* Increased cell migration and plasticity in Nrf2-deficient cancer cell lines. *Oncogene* **29**, 3703-3714, doi:10.1038/onc.2010.118 (2010).

20 Ryoo, I. G., Ha, H. & Kwak, M. K. Inhibitory role of the KEAP1-NRF2 pathway in TGFbeta1-stimulated renal epithelial transition to fibroblastic cells: a modulatory effect on SMAD signaling. *PLoS One* **9**, e93265, doi:10.1371/journal.pone.0093265 (2014).

21 Prestigiacomo, V. & Suter-Dick, L. Nrf2 protects stellate cells from Smad-dependent cell activation. *PLoS One* **13**, e0201044, doi:10.1371/journal.pone.0201044 (2018).

22 Oh, C. J. *et al.* Sulforaphane attenuates hepatic fibrosis via NF-E2-related factor 2-mediated inhibition of transforming growth factor-beta/Smad signaling. *Free Radic Biol Med* **52**, 671-682, doi:10.1016/j.freeradbiomed.2011.11.012 (2012).

23 Yang, J. J. *et al.* MicroRNA-200a controls Nrf2 activation by target Keap1 in hepatic stellate cell proliferation and fibrosis. *Cell Signal* **26**, 2381-2389, doi:10.1016/j.cellsig.2014.07.016 (2014).

24 Chen, F. *et al.* Different Inhibition of Nrf2 by Two Keap1 Isoforms alpha and beta to Shape Malignant Behaviour of Human Hepatocellular Carcinoma. *Int J Mol Sci* **23**, doi:10.3390/ijms231810342 (2022).

25 Wang, Y. *et al.* Toxicity of Dithiothreitol (DTT) to Drosophila melanogaster. *Toxicol Rep* **8**, 124-130, doi:10.1016/j.toxrep. 2020.12.014 (2021).

26 Keum, Y. S. & Choi, B. Y. Molecular and chemical regulation of the Keap1-Nrf2 signaling pathway. *Molecules* **19**, 10074-10089, doi:10.3390/molecules190710074 (2014).

27 Lee, D. H. *et al.* SQSTM1/p62 activates NFE2L2/NRF2 via ULK1-mediated autophagic KEAP1 degradation and protects mouse liver from lipotoxicity. *Autophagy* **16**, 1949-1973, doi:10.1080/15548627.2020.1712108 (2020).

28 Filomeni, G., De Zio, D. & Cecconi, F. Oxidative stress and autophagy: the clash between damage and metabolic needs. *Cell Death Differ* **22**, 377-388, doi:10.1038/cdd.2014.150 (2015).

29 Ichimura, Y. *et al.* Phosphorylation of p62 activates the Keap1-Nrf2 pathway during selective autophagy. *Mol Cell* **51**, 618-631, doi:10.1016/j.molcel.2013.08.003 (2013).

30 Kong, L. *et al.* Sitagliptin activates the p62-Keap1-Nrf2 signalling pathway to alleviate oxidative stress and excessive autophagy in severe acute pancreatitis-related acute lung injury. *Cell Death Dis* **12**, 928, doi:10.1038/s41419-021-04227-0 (2021).

31 Ito, A. *et al.* The subcellular localization and activity of cortactin is regulated by acetylation and interaction with Keap1. *Sci Signal* **8**, ra120, doi:10.1126/scisignal.aad0667 (2015).

32 Chen, C. R., Kang, Y., Siegel, P. M. & Massague, J. E2F4/5 and p107 as Smad cofactors linking the TGFbeta receptor to c-myc repression. *Cell* **110**, 19-32, doi:10.1016/s0092-8674(02)00801-2 (2002).

33 Zhu, Y. P. *et al.* Unification of Opposites between Two Antioxidant Transcription Factors Nrf1 and Nrf2 in Mediating Distinct Cellular Responses to the Endoplasmic Reticulum Stressor Tunicamycin. *Antioxidants (Basel)* **9**, doi:10.3390/antiox9010004

(2019).

34 Wufuer, R., Fan, Z., Liu, K. & Zhang, Y. Differential Yet Integral Contributions of Nrf1 and Nrf2 in the Human HepG2 Cells on Antioxidant Cytoprotective Response against Tert-Butylhydroquinone as a Pro-Oxidative Stressor. *Antioxidants (Basel)* **10**, doi:10.3390/antiox10101610 (2021).

35 Rojo de la Vega, M., Chapman, E. & Zhang, D. D. NRF2 and the Hallmarks of Cancer. *Cancer Cell* **34**, 21-43, doi:10.1016/j.ccell.2018.03.022 (2018).

36 Wang, J. *et al.* Nestin regulates cellular redox homeostasis in lung cancer through the Keap1-Nrf2 feedback loop. *Nat Commun* **10**, 5043, doi:10.1038/s41467-019-12925-9 (2019).

37 Qiu, L., Wang, M., Zhu, Y., Xiang, Y. & Zhang, Y. A Naturally-Occurring Dominant-Negative Inhibitor of Keap1 Competitively against Its Negative Regulation of Nrf2. *Int J Mol Sci* **19**, doi:10.3390/ijms19082150 (2018).

38 Wakabayashi, N. *et al.* Keap1-null mutation leads to postnatal lethality due to constitutive Nrf2 activation. *Nat Genet* **35**, 238-245, doi:10.1038/ng1248 (2003).

39 Mitsuishi, Y. *et al.* Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming. *Cancer Cell* **22**, 66-79, doi:10.1016/j.ccr.2012.05.016 (2012).

40 Suzuki, T. *et al.* Hyperactivation of Nrf2 in early tubular development induces nephrogenic diabetes insipidus. *Nat Commun* **8**, 14577, doi:10.1038/ncomms14577 (2017).

41 Hayes, J. D. & Dinkova-Kostova, A. T. Oncogene-Stimulated Congestion at the KEAP1 Stress Signaling Hub Allows Bypass of NRF2 and Induction of NRF2-Target Genes that Promote Tumor Survival. *Cancer Cell* **32**, 539-541, doi:10.1016/j.ccell.2017.10.009 (2017).

42 Tong, K. I. *et al.* Keap1 recruits Neh2 through binding to ETGE and DLG motifs: characterization of the two-site molecular recognition model. *Mol Cell Biol* **26**, 2887-2900, doi:10.1128/MCB.26.8.2887-2900.2006 (2006).

43 Kobayashi, A. *et al.* Oxidative and electrophilic stresses activate Nrf2 through inhibition of ubiquitination activity of Keap1. *Mol Cell Biol* **26**, 221-229, doi:10.1128/MCB.26.1.221-229.2006 (2006).

44 Hata, A., Lo, R. S., Wotton, D., Lagna, G. & Massague, J. Mutations increasing autoinhibition inactivate tumour suppressors Smad2 and Smad4. *Nature* **388**, 82-87, doi:10.1038/40424 (1997).

45 Yu, J. S. *et al.* PI3K/mTORC2 regulates TGF-beta/Activin signalling by modulating Smad2/3 activity via linker phosphorylation. *Nat Commun* **6**, 7212, doi:10.1038/ncomms8212 (2015).

46 Wang, G., Long, J., Matsuura, I., He, D. & Liu, F. The Smad3 linker region contains a transcriptional activation domain. *Biochem J* **386**, 29-34, doi:10.1042/BJ20041820 (2005).

47 Kamato, D. *et al.* Transforming growth factor-beta signalling: role and consequences of Smad linker region phosphorylation. *Cell Signal* **25**, 2017-2024, doi:10.1016/j.cellsig.2013.06.001 (2013).

48 Vasilaki, E. *et al.* Novel regulation of Smad3 oligomerization and DNA binding by its linker domain. *Biochemistry* **48**, 8366-8378, doi:10.1021/bi9005489 (2009).

49 Lonn, P., Moren, A., Raja, E., Dahl, M. & Moustakas, A. Regulating the stability of TGFbeta receptors and Smads. *Cell Res* **19**, 21-35, doi:10.1038/cr.2008.308 (2009).

50 Nakashima, K. *et al.* Cullin-associated NEDD8-dissociated protein 1, a novel interactor of rabphilin-3A, deubiquitylates rabphilin-3A and regulates arginine vasopressin secretion in PC12 cells. *Endocr J* **65**, 325-334, doi:10.1507/endocrj.EJ17-0399 (2018).

51 Zhang, Y. & Hayes, J. D. Identification of topological determinants in the N-terminal domain of transcription factor Nrf1 that control its orientation in the endoplasmic reticulum membrane. *Biochem J* **430**, 497-510, doi:10.1042/BJ20100471 (2010).

52 Wang, M. *et al.* TCF11 Has a Potent Tumor-Repressing Effect Than Its Prototypic Nrf1alpha by Definition of Both Similar Yet Different Regulatory Profiles, With a Striking Disparity From Nrf2. *Front Oncol* **11**, 707032, doi:10.3389/fonc.2021.707032 (2021).

53 Bae, S. H. *et al.* Sestrins activate Nrf2 by promoting p62-dependent autophagic degradation of Keap1 and prevent oxidative liver damage. *Cell Metab* **17**, 73-84, doi:10.1016/j.cmet.2012.12.002 (2013).

54 Park, J. S. *et al.* Dual roles of ULK1 (unc-51 like autophagy activating kinase 1) in cytoprotection against lipotoxicity. *Autophagy* **16**, 86-105, doi:10.1080/15548627.2019.1598751 (2020).

Figure legends

Figure 1

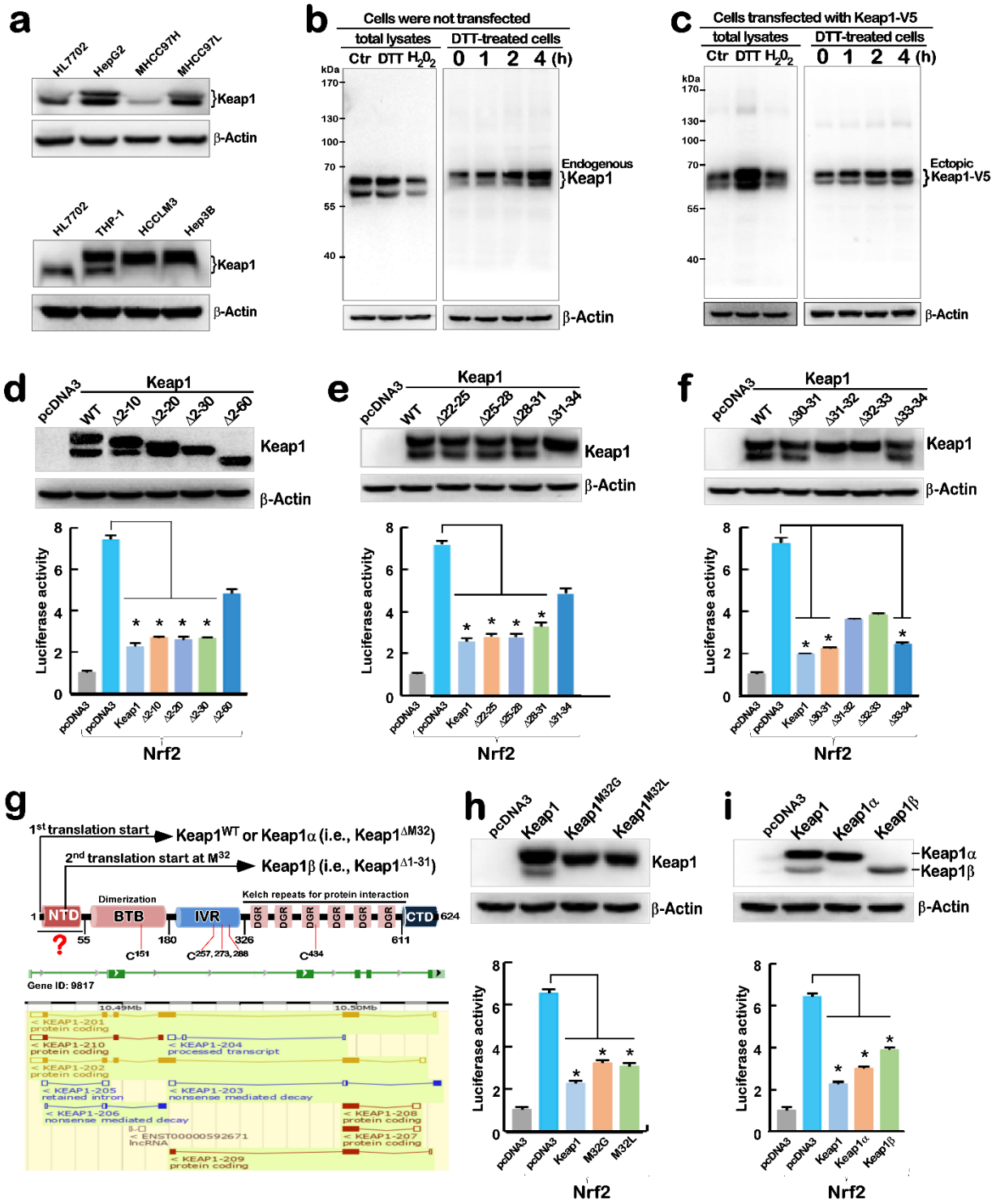

Fig. 1. Discovery of Keap1 isoforms α and β.

a. Differential expression of Keap1 isoforms in distinct cell lines examined by Western blotting with its specific antibody.

b. The existence of endogenous Keap1 isoforms are unaffected by redox agents. Total lysates of HepG2 cells were treated (*in vitro*) with control water, DTT (0.5 M) or H_2O_2 (0.5 M), along with the lysates of HepG2 cells that had or had not been treated (*in vivo*)

with 0.1 mM DTT for 0 to 4 h, were separated by SDS-PAGE and visualized by Western blotting with Keap1-specific antibody.

c. The existence of ectopically-expressing Keap1 isoforms are unaffected by redox agents. Total lysates of HepG2 cells that had been transfected with an expression construct for Keap1-V5 and treated with 0.1 mM DTT for 0 to 4 h, along with the Keap1-expressing lysates being *in vitro* treated with control water, DTT (0.5 M) or H_2O_2 (0.5 M), were subjected to Western blotting analysis.

d-f. Keap1 mutagenesis mapping. HepG2 cells that had been co-transfected with expression constructs for Keap1-V5 and its mutants, along with the Nrf2-expressing plasmid, *ARE-Luc* and *pRL-TK* reporters, were subjected to Western blotting analysis (*upper panels*) and the Dual-Lumi™ luciferase assays as shown graphically (*lower panels*). There data were representative of at least three independent experiments (n = 3 x 3), and significant decreases (*, $p < 0.01$) were determined relative to the positive controls (of Nrf2 with empty pcDNA3 instead of Keap1).

g. Schematic representation of Keap1 and its isoforms yielded at its mRNA and protein levels. Among them, Keap1α and Keap1β are yielded from alternative translation starting codons at Met^{1st} and Met^{32nd}, respectively. All other isoforms of Keap1 were searched from the e!ensembl database (https://asia.ensembl.org/index.html)

h-i. Identification of Keap1 isoforms α and β. HepG2 cells that had been co-transfected with expression constructs for Keap1-V5 and its mutants (at the Met^{32nd} position), along with the Nrf2-expressing plasmid, *ARE-Luc* and *pRL-TK* reporters, were subjected to Western blotting analysis (*upper panels*) and the Dual-Lumi™ luciferase assays as shown graphically (*lower panels*). These data were representative of at least three independent experiments (n = 3 x 3), and significant decreases (*, $p < 0.01$) were determined relative to the positive controls (of Nrf2 with empty pcDNA3 instead of Keap1).

Figure 2

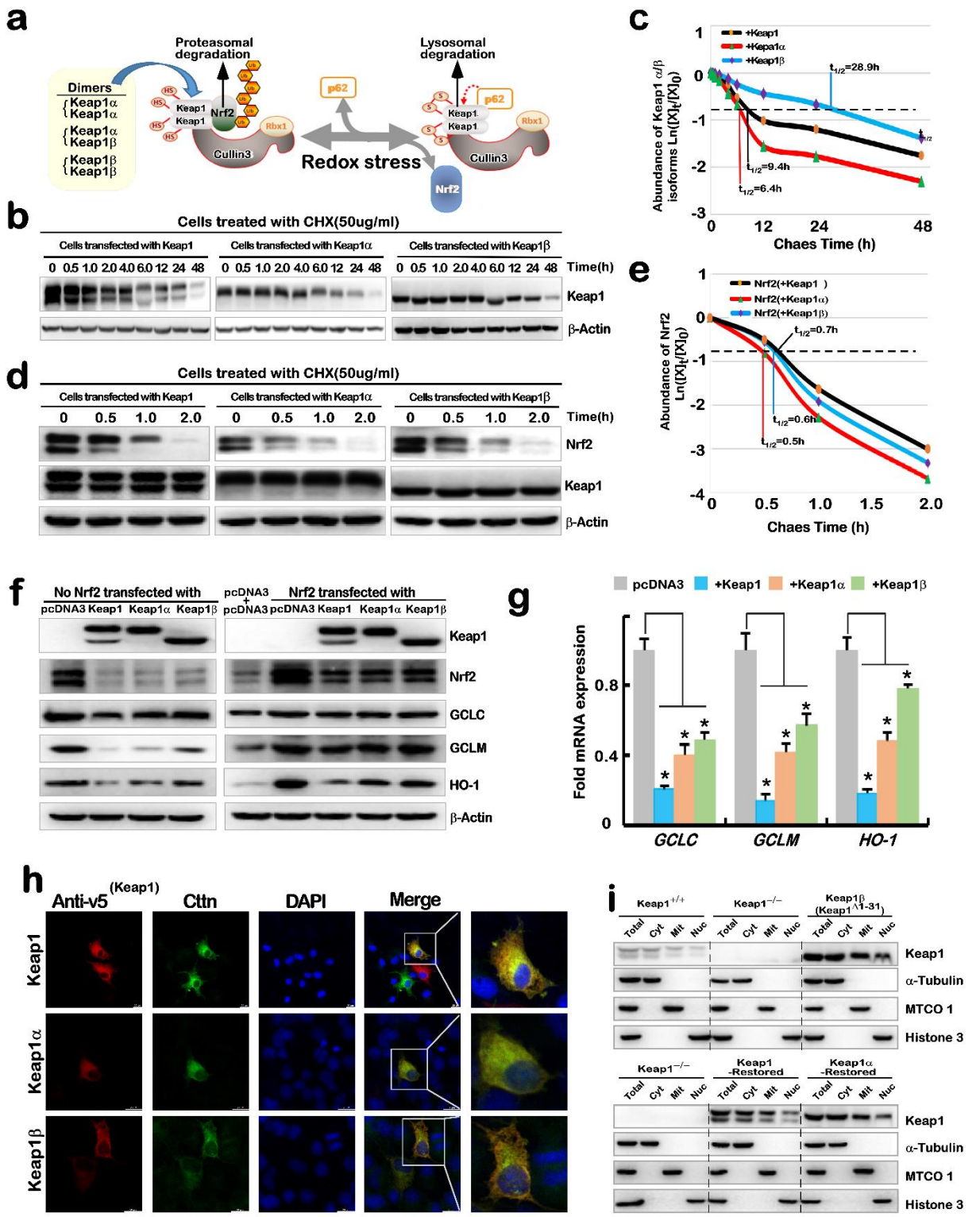

Fig. 2. Distinct inhibitory effects of Keap1 and its isoforms (α and β) on Nrf2 stability and *trans*-activity.

a. Schematic diagram of distinct Keap1 dimers (αα, αβ and ββ) and its interactor Nrf2. In normal (or reductive) conditions, Keap1 interacts with Nrf2 to target this CNC-bZIP factor to the ubiquitin-proteasomal degradation system, but the exposure to oxidative stress, Keap1 disassociates from Nrf2, but instead interacts with p62 targeting to the lysosomal autophagic pathway.

b-e. Distinct stability of Keap1 and isoforms α and β, with distinct effects on Nrf2 stability. HepG2 cells that had been transfected with expression constructs for Keap1, Keap1α or Keap1β alone (***b***) or plus Nrf2(***d***), were treated with 50 μg/ml of cycloheximide (CHX) for indicated lengths of time in the pulse chase experiments, followed by visualization by Western blotting. The intensity of anti-Keap1 or -Nrf2 immunoblots was quantified by the Quantity One 4.5.2 software (Bio-Rad, CA, USA). The resulting data of Keap1 (***c***) or Nrf2 (***e***) abundances with distinct half-lives were representative of at least three independent experiments, as they are shown graphically, after being calculated by a formula of $Ln([A]_t /[A]_0$, in which $[A]_t$ indicated a fold change (mean ± SD) in each of those examined protein expression levels at different times relative to the corresponding controls measured at 0 h (i.e., $[A]_0$).

f. Distinct effects of ectopic Keap1 Keap1α or Keap1β on Nrf2-targets. HepG2 cells were transfected with expression constructs for Keap1, Keap1α or Keap1β alone (*left panels*) or plus Nrf2 (*left panels*) and then subjected to Western blotting analysis of Nrf2 (including endogenous and exogenous proteins) and its downstream targets.

g. Differential inhibitory effects of ectopic Keap1, Keap1α or Keap1β on Nrf2-target genes. HepG2 cells were transfected with expression constructs for Keap1, Keap1α or Keap1β alone or plus Nrf2 (Fig. S1b) and then subjected to real-time qPCR analysis of Nrf2 and its target genes. The data were representative of at least three independent experiments (n = 3 x 3), and significant decreases (*, $p < 0.01$) were determined relative to the corresponding controls (of Nrf2 with empty pcDNA3 instead of Keap1).

h. Subcellular localization of Keap1, Keap1α or Keap1β with CTTN. COS-1 cells were co-transfected with an expression construct for V5-tagged Keap1, Keap1α or Keap1β, together with a CTTN-Flag plasmid, and then subjected to immunocytochemistry with anti-V5 and -Flag and fluorescent-labelled secondary antibody, followed by DNA-staining with DAPI. The representative images were obtained by confocal microscopy. Scale bars =20 μm.

i. Subcellular fractionation of Keap1, Keap1α or Keap1β from distinct genotypic cell lines. the cytosolic, mitochondria and nuclear fractions were isolated from five examined cell lines, according to the manufacturer's instruction, and then subjected to Western blotting analysis of Keap1 and its isoforms α and β. Of note, distinct subcellular compartment-specific markers were represented by α-Tubulin, MTCO1 and Histone3, respectively.

Figure 3

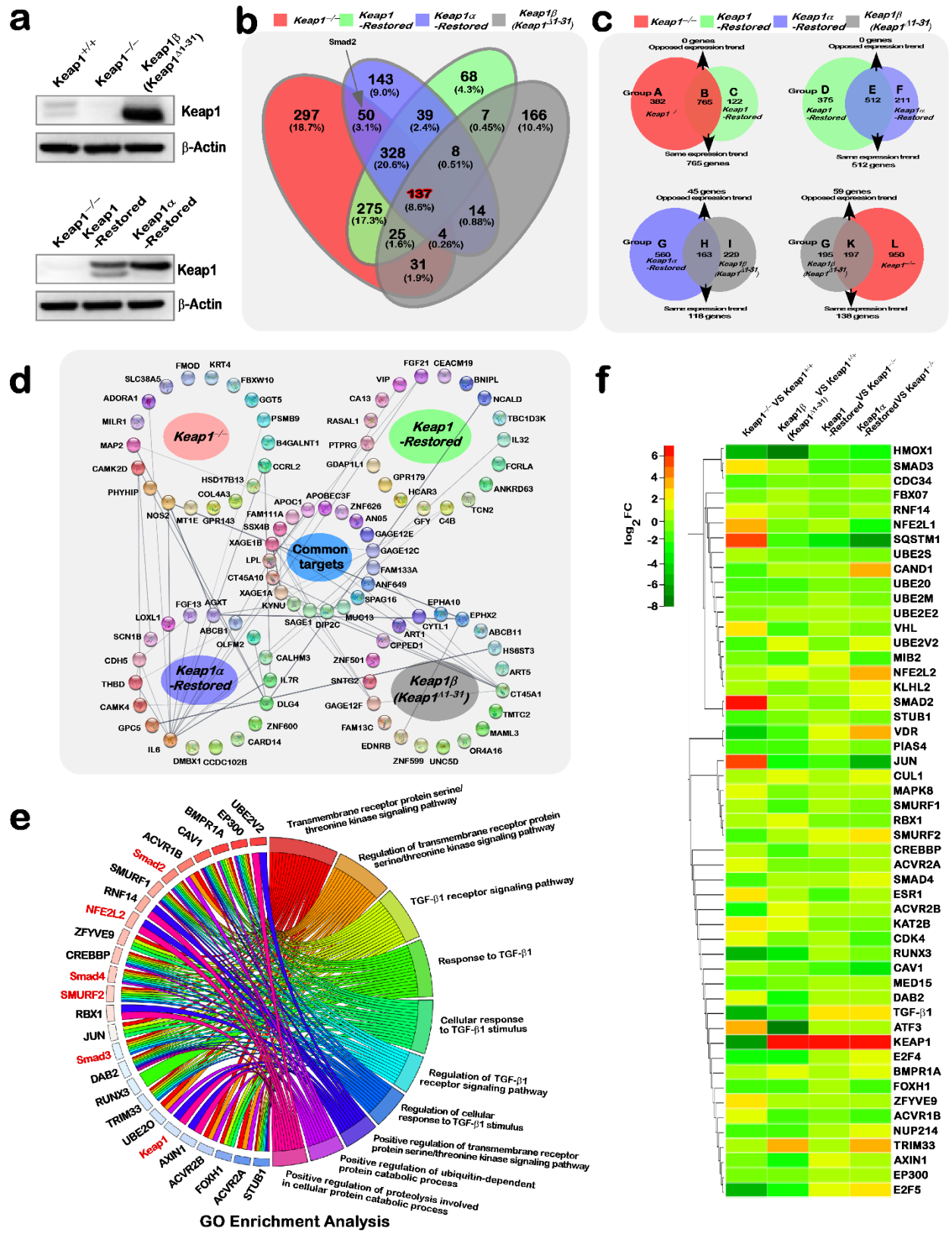

Fig. 3. Differential expression genes monitored in distinct Keap1-based genotypic cell lines.

a. Distinct Keap1-based genotypic cell lines were confirmed by Western blotting (herein) and genome sequencing[24].

b. The Venn chart displays the common or unique DEGs among those cell lines compared to *Keap1*$^{+/+}$, in which Smad2 was arrowed in an intersecting region of *Keap1*$^{-/-}$ with *Keap1α-restored.*

c. The DEGs between every two indicated cell lines were grouped with their subsequent functional annotation in the Venn diagram.

d. Functional protein association networks governed by Keap1 or not. Only the top 20 targets were shown in each subnetwork that were unique and common to each cell line.

e. The GO enrichment analysis. Those genes associated with Keap1, Smad2 and Smad3 were denoted for the top 10 GO enrichment functions. Their relevant genes were linked to the same coloured enrichment function.

f. The heatmap with hierarchical clusters of 51 DEGs shared in all four distinct cells lines. As indicated, *Keap1*$^{-/-}$ and *Keap1β (Keap1*$^{\triangle 1\text{-}31}$*)* were compared with *Keap1*$^{+/+}$, whereas *Keap1-restored* and *Keap1α-restored* were compared with *Keap1*$^{-/-}$. Distinct nodes in the heatmap are represented by the coloured bars showing their values of log2 (fold change). Upregulation was shown in red and yellow, while downregulation was deciphered in green.

Figure 4

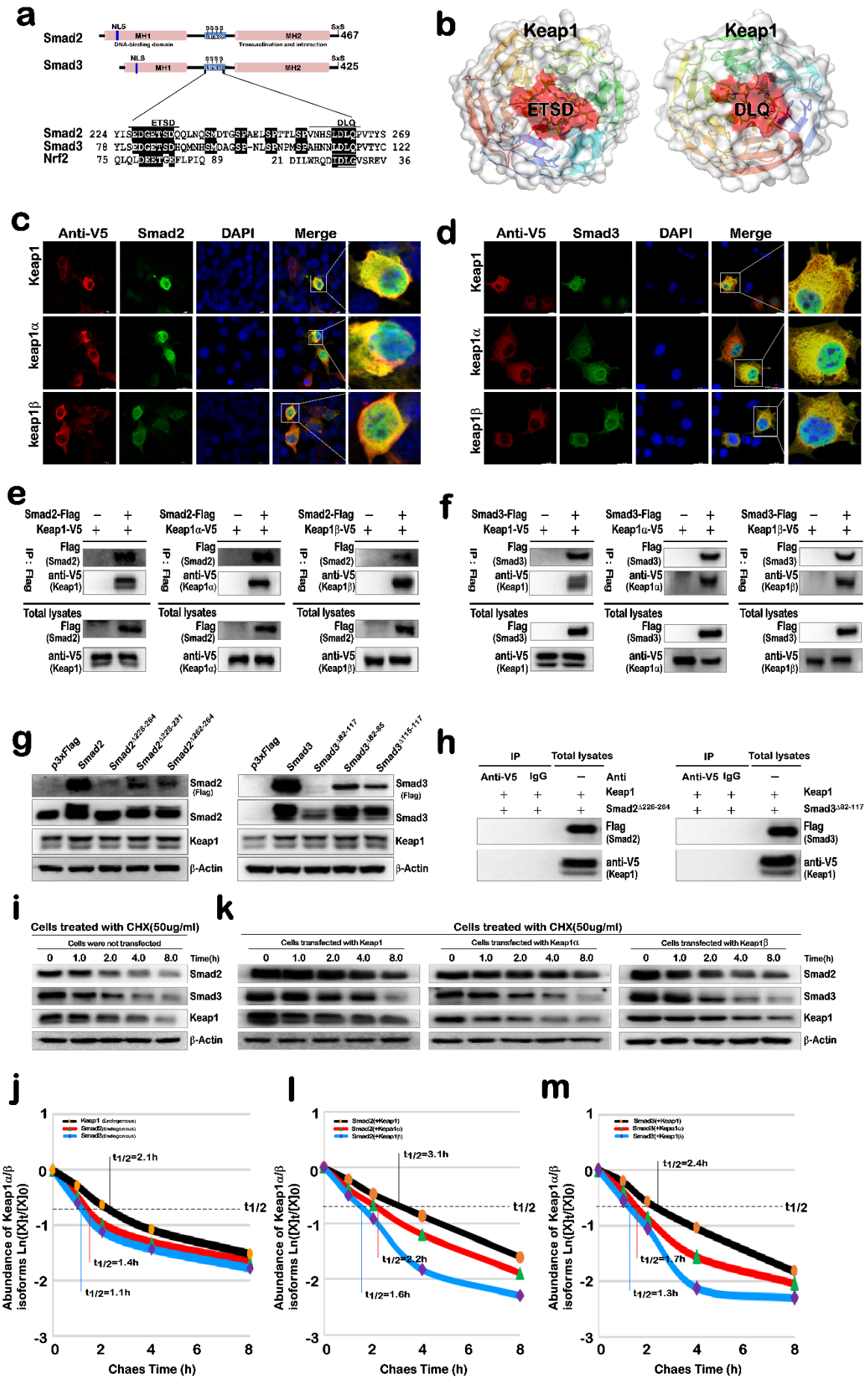
a
Smad2
Smad3
NLS
MH1
MH2
DNA-binding domain
Transactivation and interaction
467
425
ETSD
DLQ
Smad2 224 YISEDGETSDQQLNQSMDTGSPAELSPTTLSLVNHSLDLQPVTYS 269
Smad3 78 YLSEDGETSDHQMNHSMDAGSE-NLSPNPMSPAHNNLDLQPVTYC 122
Nrf2 75 QLQLDEETGEFLPIQ 89 21 DILWRQDIDLGVSREV 36
b
Keap1
ETSD
DLQ
c
Anti-V5
Smad2
DAPI
Merge
Keap1
keap1α
keap1β
d
Anti-V5
Smad3
DAPI
Merge
Keap1
keap1α
keap1β
e
Smad2-Flag
Keap1-V5
Keap1α-V5
Keap1β-V5
IP : Flag
Flag (Smad2)
anti-V5 (Keap1)
anti-V5 (Keap1α)
anti-V5 (Keap1β)
Total lysates
f
Smad3-Flag
Keap1-V5
Keap1α-V5
Keap1β-V5
Flag (Smad3)
g
p3xFlag
Smad2
Smad2Δ228-264
Smad2Δ228-231
Smad2Δ262-264
Smad3
Smad3Δ82-117
Smad3Δ82-85
Smad3Δ115-117
Smad2 (Flag)
Smad3 (Flag)
Keap1
β-Actin
h
IP
Total lysates
Anti-V5
IgG
Anti
Keap1
Smad2Δ228-264
Smad3Δ82-117
Flag (Smad2)
Flag (Smad3)
anti-V5 (Keap1)
i
Cells treated with CHX(50ug/ml)
Cells were not transfected
Time(h)
0
1.0
2.0
4.0
8.0
Smad2
Smad3
Keap1
β-Actin
k
Cells treated with CHX(50ug/ml)
Cells transfected with Keap1
Cells transfected with Keap1α
Cells transfected with Keap1β
j
l
m
Abundance of Keap1α/β isoforms Ln([X]t/[X]0)
Chaes Time (h)
t1/2
t1/2=2.1h
t1/2=1.4h
t1/2=1.1h
t1/2=3.1h
t1/2=2.2h
t1/2=1.6h
t1/2=2.4h
t1/2=1.7h
t1/2=1.3h

Fig. 4. Keap1 and its isoforms α and β interact with Smad2/3 to enhance their protein stability.

a. Schematic representation of structural domains of Smad2 (NM_005901.6) and Smad3 (NM_005902.4), together with their amino acid sequences within the linker regions being aligned with the known Keap1-binding motifs of Nrf2 (NM_006164.5).

b. Structural modelling of interaction of Smad2/3 with Keap1 by using the CABS-dock method. The putative Keap1-binding peptides within Smad2 (YISEDGETSD and NHSLDLQPV) or Smad3 (YLSEDGETSD and HNNLDLQPV) were selected for docking into the KEAP1 (PDB code: 1ZGK) with default settings.

c-d. Subcellular co-localization of Smad2/3 with Keap1, Keap1α or Keap1β. COS-1 cells that had been co-transfected with expression constructs for V5-tagged Keap1, Keap1α or Keap1β, together with Smad2-Flag (*c*) or Smad3-Flag (*d*), were subjected to imaging of immunocytochemistry with antibodies against V5 or Flag epitopes, along with the fluorescent secondary antibodies, followed by being stained with DAPI. The immuno-fluorescent images were acquired by confocal microscopy. Scale bars =20 μm

e-f. Co-immunoprecipitation of Smad2 (*e*) or Smad3 (*f*) to Keap1, as well as its isoforms α and β. Total lysates of COS-1 cells that had been co-transfected with expression constructs for Smad2 or Smad3, along with V5-tagged Keap1, Keap1α or Keap1β, were subjected to co-immunoprecipitation with (herein) and antibody (Fig. S8), followed by Western blotting with anti-V5 or -Flag antibodies, respectively.

g. No effects of Smad2/3 mutants on the abundance of Keap1. Total lysates of HepG2 cells were transfected with expression constructs for Smad2/3-Flag or their mutants and then subjected to determination by Western blotting with antibodies against Keap1, Smad2/3 and their C-terminal Flags.

h. No immunoprecipitates of Smad2$^{\triangle 228\text{-}264}$-Flag or Smad3$^{\triangle 82\text{-}117}$-Flag were put down by V5-tagged Keap1. Total lysates of COS-1 cells that had been co-transfected with expression constructs for V5-tagged Keap1 plus either Smad2$^{\triangle 228\text{-}264}$-Flag or Smad3$^{\triangle 82\text{-}117}$-Flag, were subjected to co-immunoprecipitation by V5-tagg antibody, followed by Western blotting with Flag or V5 antibodies.

i-J. The CHX-pulse chase experiments to determine the half-lives of endogenous Smad2/3 and Keap1 expressed in HepG2 cells that had been treated with 50 g/mL CHX for 0 to 8 h, followed by visualization by Western blotting with their specific antibodies. Then the intensity of the immunoblotts representing Keap1, Smad2 or Smad3 was quantified by the Quantity One 4.5.2 software (Bio-Rad, Hercules, CA, USA). The data were representative of at least three independent experiments, and are shown graphically, after being calculated by a formula of $Ln([A]_t/[A]_0$, in which $[A]_t$ indicated a fold change (mean ± SD) in each of those examined protein expression levels at different times relative to the corresponding controls measured at 0 h (i.e., $[A]_0$).

k-m. Differential enhancement of Smad2/3 stability by Keap1, Keap1α and Keap1β. HepG2 cells were transfected with expression constructs for Keap1, Keap1α and Keap1β and treated with 50 g/mL CHX for 0 to 8 h (*k*), and then subjected to the pulse chase experimental analysis of endogenous Smad2/3 stability, which was estimated by their half-lives of Smad2 (*l*) and Smad3 (*m*). The data were representative of at least three independent experiments.

Figure 5

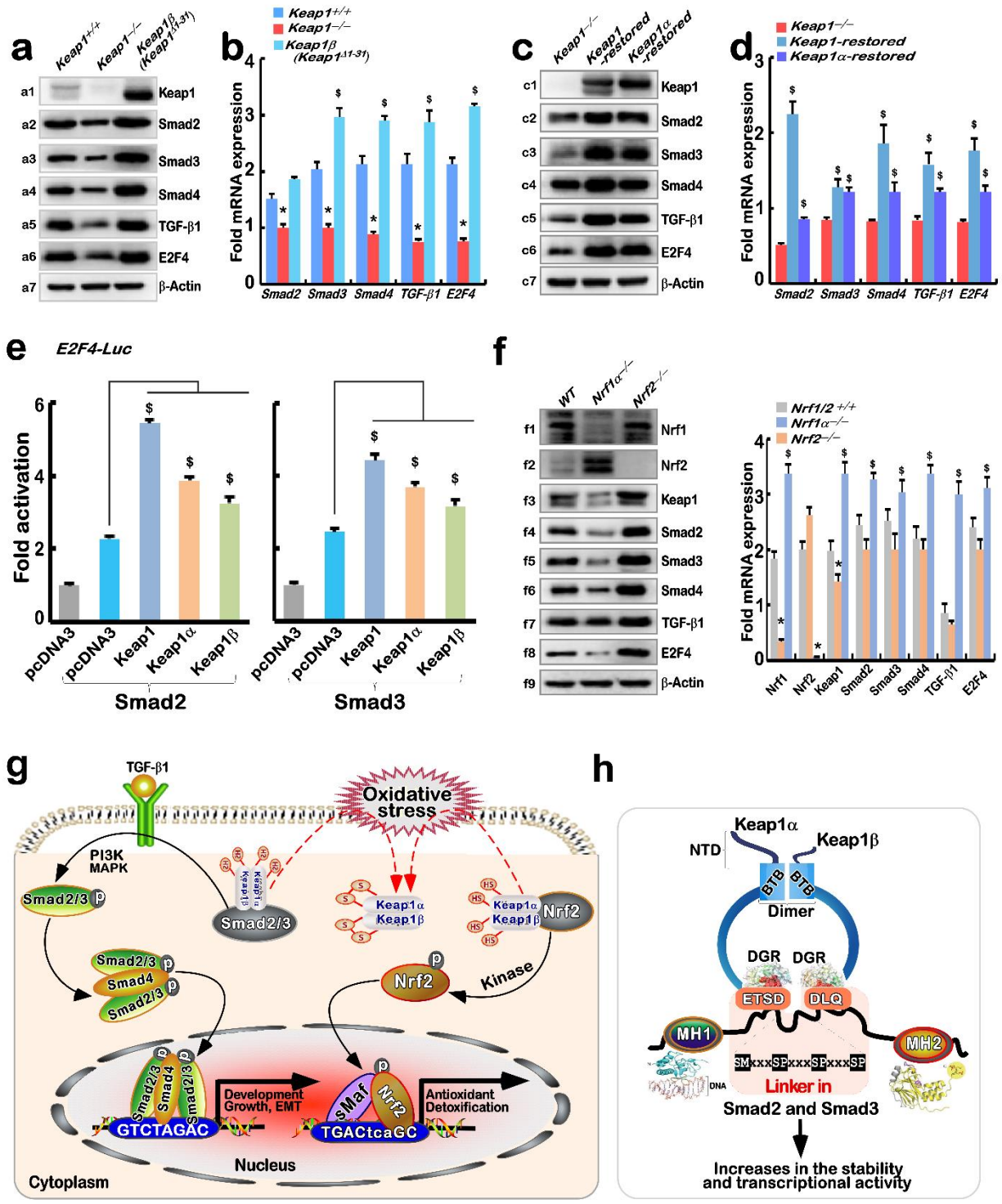

Fig 5. Upregulation of Smad2/3-mediated genes by Keap1 and its isoforms α and β.

a-d. Distinct effects of Keap1, Keap1α and Keap1β on Smad2/3 and related cognate genes**.** These protein and mRNA expression levels in *Keap1*$^{+/+}$, *Keap1*$^{-/-}$ and *Keap1β(Keap1*$^{\Delta 1\text{-}31}$*)* cell lines were determined by Western blotting (*a*) and quantitative real-time PCR (*b*). Similar experiments were carried out to determine the relevant protein and mRNA abundances (*c* and *d*) in *Keap1*$^{-/-}$, *Keap1-restored* and *Keap1α-restored* cell lines. The data were representative of at least three independent experiments (n = 3 x 3), and significant decreases (*, $p < 0.01$) and significant increases ($, $p < 0.01$) were determined relative to the corresponding controls.

e. The effects of Keap1 and its isoforms (α and β) on Smad2/3-mediated reporter activity. HepG2 cells that had been co-transfected with expression constructs for Smad2 or Smad3, plus Keap1, Keap1α, Keap1β or empty pcDNA3 vector, together with the *E2F4-Luc* reporter gene and *pRL-TK* control, were subjected to the Dual-Lumi™ reporter assay. The Smad2/3-mediated reporter activity was calculated from at least three independent experiments (n = 3 x 3). Significant increases ($, $p < 0.01$) were also determined relative to the corresponding controls.

f. Regulation of Smad2/3 and cognate genes by altered Keap1 in *Nrf2*$^{-/-}$ and *N rf1*α$^{-/-}$ cell lines when compared to wild-type (*WT*) HepG2 cells. The protein and mRNA expression levels Keap1, Smad2/3 and cognate genes in the examined three cell lines were determined by Western blotting with each specific antibody (*left panels*) and quantitative real-time PCR (*right panel*). The data were representative of at least three independent experiments (n = 3 x 3), and significant decreases (*, $p < 0.01$) and significant increases ($, $p < 0.01$) were determined relative to the corresponding controls.

g-h. A novel model was proposed to give a better explanation of the newly-identified functional cross-talk of Keap1, as its isoforms α and β, with its interactors Nrf2 and Smad2/3 (g). Of note, the stability of Smad2/3 and their transcriptional activity to mediate target gene expression are enhanced by resulting from a physical interaction of Keap1 with the highly conserved EDGETSD and DLQ motifs within the linker region of Smad2/3 (h), although Nrf2 is negatively regulated by Keap1. Overall, this study presents a novel functional bridge of Keap1, Keap1α or Keap1β crossing both the redox-responsive Nrf2 and the developmental TGFβ1-Smad2/3 signalling pathways. This is further supported by our experimental evidence revealing that the phosphorylated Smad2 activation occurs *via* MAPKs and PI-3 kinase signalling pathways in response to TGF-β1 and UVA (as oxidative stress) (Fig. S13).