IDENTIFICATION OF A KEY CO-OPERATIVE ROLE OF YB-1 IN THE GENESIS AND METASTASIS OF HUMAN BREAST CANCER

Sylvain Lefort[1], Amal El-Naggar[2,3], Susanna Tan[1,4], Shane Colborne[5], Gian Luca Negri[2], Davide Pellacani[1], Martin Hirst[5-7], Barry Gusterson[8], Gregg B. Morin[5,9], Poul H. Sorensen[2,10*], Connie J. Eaves[1,9*]

*Co-corresponding authors

[1]Terry Fox Laboratory, British Columbia Cancer Agency, Vancouver, BC, Canada

[2]Department of Molecular Oncology, British Columbia Cancer Agency, Vancouver, BC, Canada

[3]Department of Pathology, Faculty of Medicine, Menoufia University, Egypt

[4]Department of Medicine, University of British Columbia, Vancouver, BC, Canada

[5]Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, BC, Canada

[6]Michael Smith Laboratories, University of British Columbia, Vancouver, BC, Canada.

[7]Department of Microbiology and Immunology, University of British Columbia, Vancouver, BC, Canada

[8]Institute of Cancer Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK

[9]Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada

[10]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada

Referenced abstract

Shared mechanisms driving the malignant properties of human breast cancers have been difficult to identify because of their extensive biological and molecular heterogeneity. Here we sought to examine the possibility that mammalian Y-box binding protein-1 (YB-1) might have such a role given its association with poor survival in patients with other types of metastatic tumours[1,2] and recently reported ability in high-risk sarcomas to induce a hypoxia-inducible factor 1α (HIF1α) response[3] and stress granule formation through translational activation of *G3BP1* mRNAs[4]. However, the role of YB-1 in contributing to the malignant properties of human breast cancer cells *in vivo* has not been experimentally investigated. Here we demonstrate a critical co-operative role of YB-1 in mediating the initial as well as later malignant properties and downstream activities displayed by *de novo* transformed normal human mammary cells as well as by an established aggressive breast cancer cell line in transplanted immunodeficient mice.

Examination of the The Cancer Genome Atlas (TCGA)[5] data for breast cancers showed *YBX1* transcripts to be highest in estrogen receptor-negative (ER-) tumours, most notably in those that had metastasized, associated with a gain of function or amplification of the *KRAS* gene, and reduced overall survival (Fig. 1a,b and Supplementary Fig. 1a,b)). In this dataset, increased *YBX1* mRNA levels were also seen in tumours with amplified *ERBB2, PIK3CA* or *AKT1*, or deletions of *TP53*, (Supplementary Fig. 1c), but not in those containing *HRAS* or *NRAS* mutations (data not shown). Examination of the larger METABRIC dataset (2,433 breast cancers)[6] also showed an association of higher levels of *YBX1* transcripts in tumours with *KRAS* amplification (Fig. 1c). In another large study[7], *YBX1* expression was also found to be correlated with that of *KRAS*, and not with changes in expression of *ERBB2, TP53, PIK3CA* or *AKT1* (Fig.

1d). Given the evidence in sarcomas that, under hypoxia, YB-1 induces a HIF1α response via its ability to bind directly to *HIF1A* mRNAs to enhance their translation[3], we then surveyed the same datasets to determine whether the level of HIF1α and its transcriptional targets[8], were similarly altered. This showed that transcripts for *HIF1A* and *CAIX* (carbonic anhydrase 9), as well as for *VEGFA* are all higher in breast cancers that exhibit gain of function or amplified *KRAS* as compared to tumours with diploid *KRAS* (Fig. 1e and Supplementary Fig. 2a). Patients' breast cancers with amplified *KRAS* also showed a strong correlation between *YBX1* and *CAIX* mRNA levels (Fig. 1f). In addition, elevated levels of these transcripts were noted in tumours with a gain of function or amplified *YBX1* gene (Supplementary Fig. 2b). Together, these data suggested that induction of a YB-1/HIF1α-mediated hypoxia response pathway might play a key role in mediating the aggressive malignant properties of human breast cancers driven by deregulated KRAS activity.

To more directly probe the biological significance of this implied association between elevated *YBX1* expression and breast cancer aggressiveness, we first examined the effects of reducing YB-1 protein levels in MDA-MB-231 cells on their *in vivo* tumorigenic properties. MDA-MB-231 cells are an established human breast cancer cell line with a $KRAS^{G13D}$ mutation[9] and elevated *YBX1* expression[10]. Transduction of these cells with lenti-viral vectors encoding short hairpin (sh)*YBX1* constructs reduced their YB-1 protein content by 90% (Fig. 2a, Supplementary Fig. 3a,b) and markedly reduced both the size (Fig. 2b and Supplementary Fig. 3c,d) and YB-1 content (Fig. 2c) of tumours they produced at the site of injection in subcutaneously transplanted immunodeficient mice. Parallel experiments in which mice were injected intravenously with the same test and control cells showed a marked failure of the sh*YBX1*-transduced cells to generate the prevalent metastases evident in multiple organs of the

recipients of the control cells (Fig. 2d,e and Supplementary Fig. 3e-h). Interestingly, however, despite these marked effects of reduced YB-1 levels on the growth of these cells *in vivo*, there was no effect on their proliferative activity *in vitro* (Supplementary Fig. 3i).

RNAseq, proteomic and immunohistochemical (IHC) analyses of cells isolated from 2 of these paired tumour sets (Supplementary Fig. 4a-d) revealed several genes whose encoded RNAs and proteins increased or decreased in concert with reduced levels of YB-1 (Supplementary Fig. 4b-d), including a number that are related to hypoxia (e.g., *PPIF*, *SLC3A2* and *SLC7A1*)[11–13] and are also present at higher levels in breast cancers with increased expression or amplification of *YBX1* (Supplementary Fig. 4e). Immunostaining of tumours generated from sh*YBX1*-transduced MDA-MB-231 cells also showed CAIX levels to be decreased compared to the tumours generated from the control cells (Fig. 2f). Interestingly, under hypoxic conditions *in vitro*, HIF1α expression was also reduced in the sh*YBX1*-transduced MDA-MB-231 cells as compared to their sh*Scr* counterparts (Fig. 2g and Supplementary Fig. 5). Taken together, these findings establish a requisite *in vivo*-specific role of YB-1 in the ability of this advanced human breast cancer cell line to display its local and metastatic tumorigenic activity in an *in vivo* model.

We then asked whether YB-1 expression also plays a key role in the initial stages of human mammary transformation driven by deregulated KRAS activity. For this we turned to a recently developed *de novo* model in which freshly isolated normal human mammary cells transduced with a $KRAS^{G12D}$ vector reproducibly and rapidly generate serially transplantable but non-metastatic invasive ductal carcinomas in immunodeficient mice[14]. IHC analysis of nascent tumours obtained from mice transplanted just 2 weeks previously with $KRAS^{G12D}$-transduced isolates of either normal human mammary basal cells (BCs) or luminal progenitors (LPs) showed strong and prevalent YB-1 expression in the cytoplasm and occasionally in the nucleus

(Supplementary Fig. 6a) at levels already similar to those apparent in the invasive tumours analyzed another 4-6 weeks later (Fig. 3a). These patterns of YB-1 expression contrast sharply with the largely luminal cell-restricted and low levels of YB-1 seen in similarly stained normal human mammary tissue (Fig. 3a) and confirmed by Western Blot (WB) analyses (Supplementary Fig. 6a,b). Interestingly, however, *KRAS*^{G12D}-transduced human mammary cells maintained in 2D culture for 4 days (Supplementary Fig. 6c) or in 3D-Matrigel cultures for up to 15 days (Fig. 3b) showed no evidence of altered YB-1 expression despite production in the latter system of acini with filled lumens and a modified architecture (Fig. 3b). Thus, a rapid increase in YB-1 expression is obtained when *KRAS*^{G12D}-transduced normal human mammary cells are stimulated to proliferate *in vivo* but not *in vitro*.

We next asked whether sustained expression of KRAS^{G12D} is necessary to maintain the high levels of YB-1 initially evident in tumours produced from *KRAS*^{G12D}-transduced primary human mammary cells. To address this question, we first constructed a doxycycline-inducible *KRAS*^{G12D}-Kusabira Orange (KO)-encoding vector (Supplementary Fig. 6d,e), and then used it to transduce purified human mammary cells that we transplanted subcutaneously into mice immediately post-transduction ($4x10^4$-$30x10^4$ cells each). At the end of 4-weeks, cells isolated from the mice that had received doxycycline for either just the first 2 weeks or not at all contained very few KO+ (*KRAS*^{G12D}-expressing) cells as compared to mice given doxycycline for the full 4-week period (Fig. 3c; right panel). Consistent with this expected biological result was the confirmatory finding that the growths obtained with a time-limited exposure to *KRAS*^{G12D} contained a 30-fold lower level of *KRAS* transcripts (Fig. 3d). However, despite the loss of *KRAS* transcripts obtained within 2 weeks after discontinuation of doxycycline, these cells were found to contain elevated *YBX1* transcript levels similar to those evident in the tumours obtained from

cells in which *KRAS*^{G12D} expression was continuously sustained (Fig. 3e). These results indicate that *in vivo*, *KRAS*^{G12D} can stimulate a rapidly increased level of YBX1 transcripts that, even in the absence of a continued expression of *KRAS*^{G12D}, can be prolonged but are insufficient to confer tumorigenic activity.

We then designed experiments using the same *de novo* model to determine whether YB-1 plays an essential role in initiating mutant KRAS-driven acquisition of tumorigenic activity by normal human mammary cells. Accordingly, purified BCs and LPs were first transduced with sh*YBX1* or sh*Scr* lentiviral vectors (Supplementary Fig. 7a). Then after 2 days *in vitro*, the cells were transduced further at high efficiency with both a *KRAS*^{G12D}-mCherry and a luciferase (Luc)-YFP vector prior to subcutaneous transplantation of the cells directly into immunodeficient mice ($2x10^4$-$25x10^4$ cells/transplant). Another 2 weeks later, cells derived from either BCs or LPs that had been co-transduced with *KRAS*^{G12D} and sh*YBX1* produced a 2-1,000-fold lower bioluminescent signal than that obtained from their counterparts transduced with *KRAS*^{G12D} and the sh*Scr* vector (Fig. 4a). H&E-staining also showed the cell content of the test transplants to be greatly reduced compared to the nascent tumours generated from cells transduced with *KRAS*^{G12D} plus the sh*Scr* construct (Supplementary Fig. 7b) and IHC staining confirmed the numbers of YB-1$^+$ cells to also be reduced in the test cell-derived harvests (Supplementary Fig. 7c). Immunostaining of HIF1α (Fig. 4b), CAIX (Fig. 4c), VEGF and CD34 (Supplementary Fig. 7d,e) showed these were also decreased in the same test samples. These experiments further serve to establish that YB-1 is a requisite co-operating mediator of the initial tumorigenic activity and downstream properties acquired by freshly isolated normal human mammary cells when they are forced to express *KRAS*^{G12D} and stimulated to grow *in vivo*.

To investigate whether up-regulation of *YBX1* expression is unique to mutant KRAS-

driven breast cancers or may be relevant in other models of human mammary transformation, we screened several other oncogenes for their ability to induce tumours *de novo* from transplants of transduced primary human cells. For this purpose, we prepared fluorochrome-based lentiviral vectors encoding cDNAs of 7 genes previously identified as potential "drivers" from genomic analyses of patients' breast cancers[15], as well as one encoding YB-1 and then transduced pooled isolates of BCs and LPs with our lenti-Luc-YFP vector plus one of each of the 8 test cDNA vectors either alone or together with a lenti-$KRAS^{G12D}$–YFP vector (to look for potential enhancing effects). As controls, additional aliquots of the same cells were transduced with the lenti-$KRAS^{G12D}$-YFP vector plus the lenti-Luc-YFP vector alone, or with just the lenti-Luc-YFP vector.

Bioluminescent monitoring of all the mice showed transduction of *YBX1* alone was insufficient to induce tumour formation and also did not enhance the growth of $KRAS^{G12D}$-induced tumours when compared to mice injected with cells transduced with $KRAS^{G12D}$ only (Fig. 4d). Dominant-negative forms of TP53 and a mutant version of PI3K also did not induce tumorigenesis on their own, as found previously[14]. Vectors encoding cDNAs for EGFR and MYC, and shRNAs targeting *PTEN* and *BRCA1* transcripts gave similarly negative results (data not shown). In contrast, forced expression of a cDNA encoding a myristoylated form of *AKT1* (*myrAKT1*), even in the absence of $KRAS^{G12D}$, led to significantly increasing luciferase signals over an 8-week period post-transplant, albeit at consistently lower levels than those obtained from $KRAS^{G12D}$-transduced cells (Fig. 4d). Constitutive activity of the myrAKT1 protein is attributed to the removal of the wild type AKT1 pleckstrin homology domain and addition of an engineered SRC myristoylation signal sequence that targets the protein to the cell membrane[16]. Phenotypic analysis of tumour cells obtained from mice transduced with *myrAKT1* (only)

showed they were universally EpCAM^{+}CD298^{+} as well as mCherry+, indicative of a weak oncogenic effect of deregulated AKT1 activity in these cells (Supplementary Fig. 8a). Additional experiments using purified normal BCs or LPs from single donors showed this modest tumorigenic activity of *myrAKT1* was readily replicated (9/13 tests, Supplementary Fig. 8b).

Histological analysis of tumours produced from the *myrAKT1*-transduced cells appeared morphologically similar to ductal carcinomas *in situ* (DCIS), with a confined organization of the cells in duct-like structures with extensive luminal filling, a low frequency of ER+ and/or Ki67+ cells, and an absence of progesterone receptor-positive (PR+) cells (Supplementary Fig. 8c). The outer layer of cells in these structures displayed the basal features of smooth muscle actin (SMA) and TP63 (Fig. 4e and Supplementary Fig. 8d), with cells displaying the luminal features of strong CK14 and CK8/18 positivity more centrally (Fig. 4e and Supplementary Fig. 8d), consistent with a DCIS morphology. The cells in these structures also showed increased levels of YB-1 protein detected by IHC staining as compared to normal cells, albeit less than in the *KRAS*^{G12D}-induced tumours (Fig. 4f). However, tumours generated from cells that had been co-transduced with a doxycycline-inducible lenti-*YBX1* vector (Supplementary Fig. 8e) as well as -*myrAKT1* and then treated with doxycycline, showed increased luciferase activity consistent with an increased growth response (Supplementary Fig. 8f), as well as the expected increase in YB-1 expression (Supplementary Fig. 8g) when compared to mice in the same experiment that were transplanted with the same cells in mice not treated with doxycycline. Thus, even in an experimental stage of early breast cancer development obtainable with forced expression of *myrAKT1,* a moderate increase in YB-1 expression was evident, and further expression of YB-1 in this model enhanced the abnormal growth *in vivo* of *myrAKT1*-transduced normal human mammary cells.

In summary, we show here that the elevated expression of *YBX1* characteristic of patients' breast cancers with elevated KRAS activity, an activated hypoxic program and a poor prognosis, is reproduced in several experimental models of human breast cancer. These include the established MDA-MB-231 human breast cancer cell line that has a $KRAS^{G13D}$ mutation[9] and highly penetrant metastatic as well as tumorigenic activity in transplanted immunodeficient hosts. In addition we show that upregulation of YB-1 protein is a rapidly acquired and sustained feature of experimentally derived primary human mammary cell transformants created *de novo* by lenti-viral transduction of $KRAS^{G12D}$. We also describe a new, somewhat less aggressive model of *de novo* normal human mammary cell transformation that results from lenti-viral deregulated AKT activity and also displays an elevated expression of YB-1. Taken together, these findings indicate a role of YB-1 both at the point of initial transformation of normal human mammary cells, as well as in more advanced derivatives that display metastatic activity and/or continued tumorigenic activity in serially transplanted immunodeficient hosts.

Previous experiments in which YB-1 was overexpressed in the immortalized but non-tumorigenic MCF10A cell line transformed with HRAS showed that high YB-1 expression contributes to the disruption of mammary cell architecture and promotes an EMT[17]. However a survey of changes in published breast cancer datasets did not show any evidence of increased *YBX1* expression in patient breast cancers with amplified HRAS or NRAS. Here we confirmed the loss of a normal acinar architecture in 3D cultures of $KRAS^{G12D}$-transduced primary human mammary cells, but noted that there were no accompanying effects on *TWIST1, SLUG* or *ZEB2* expression in patients' breast cancers with gain of function mutations or amplified *KRAS* (data not shown). It is also interesting that forced expression of $KRAS^{G12D}$ in primary human mammary cells did not modulate *YBX1* transcript levels in cells maintained in normoxic conditions *in vitro*,

in contrast to the rapid increases observed *in vivo*. This difference may reflect the creation of a hypoxic environment in nascent tumours forming *in vivo*, as shown by the high levels of HIF1α and CAIX in these populations, since the *in vitro* experiments performed under hypoxic conditions suggested this could allow YB-1 to modulate HIF1α expression.

High expression of HIF1α target genes is a shared feature of triple-negative breast cancers with correlated levels of expression of *HIF1A,* but not *HIF2A*[18]. Thus, our results showing that suppression of YB-1 expression leads to decreased expression of HIF1α and CAIX in both *de novo* and advanced tumours, combined with overexpression of HIF1α and HIF target genes in patients' breast cancers with amplified KRAS, is consistent with a model in which activated KRAS activates YB-1 levels which, in turn, promote an elevated HIF1α response *in vivo*. Further investigations will be required to determine whether YB-1 directly binds to the *HIF1A* 5'-UTR in human breast cancer cells to enhance the acute synthesis of HIF1α as previously shown in sarcoma cells[3].

These findings reveal a key role of YB-1 in initiating as well as sustaining the growth of transformed human mammary cells with accompanying activation of a HIF1α response. YB-1 may therefore represent a relevant target for therapeutic intervention in breast cancer. At the same time, the present findings suggest that targeting HIF1α or its downstream effectors may offer alternative and potentially more tractable clinical targets. The *de novo* tumour systems we describe here should provide new and robust preclinical models to elucidate these key disease mechanisms not otherwise readily accessible using other approaches.

ACKNOWLEDGMENTS

The authors thank D. Wilkinson, G. Edin, M. Hale, and A. Li for excellent technical support; and to Drs. E. Bovill, J. Boyle, S. Bristol, P. Gdalevitch, A. Seal, J. Sproul, and N. van Laeken for access to discarded reduction human mammoplasty tissue. This work was supported by grants to CJE from the Canadian Cancer Society Research Institute, the Cancer Research Society, and the Cancer Institute for Health Research (CIHR grant CRP-154482), to, PHS from the Terry Fox Research Institute (Team Grant 1021) and the CIHR (Foundation 143280. ST holds a CIHR Banting and Best Studentship and A.E-N a Michael Smith Foundation for Health Research trainee award (# 17159).

AUTHOR CONTRIBUTIONS

S.L., P.H.S. and C.J.E. conceptualized this project and drafted the manuscript. S.L., A.E-N., S.T., and S.C. performed experiments. S.C and G.L.N. carried out the computational and bioinformatics analysis of the proteomics and RNA-seq data. S.L., A.E-N., S.T., S.C., G.L.N., M.H., B.G., G.B.M., P.H.S. and C.J.E. analyzed and interpreted the data.

Conflict of Interest Disclosures: The authors have no conflicts of interest.

METHODS

Cells and cultures

Normal human reduction mammoplasty discard tissue was collected with informed consent, according to protocols approved by the University of British Columbia Research Ethics Board. Organoid-rich pellets were then isolated and viably cryopreserved[19]. As required, thawed organoids were rinsed with 2% fetal bovine serum (FBS, from STEMCELL Technologies) in Hank's Balanced Salt Solution (HF), and the cells then dissociated in 2.5 mg/ml trypsin with 1 mM EDTA and 5 mg/ml dispase (STEMCELL Technologies) with 100 μg/ml DNaseI (Sigma) and washing of the cells with HF between each step. The resulting cell suspension was filtered through a 40 μm mesh and BCs then isolated by FACS according to their $CD45^{-}CD31^{-}EpCAM^{lo}CD49f^{+}$ phenotype, LPs according to their $CD45^{-}CD31^{-}EpCAM^{hi}CD49f^{+}$ phenotype, LCs according to their $CD45^{-}CD31^{-}EpCAM^{hi}CD49f^{-}$ phenotype and stromal cells (SCs) according to their $CD45^{-}CD31^{-}EpCAM^{-}CD49f^{-}$ phenotype using well established protocols and reagents [20]. Following FACS, cells were transduced or cultured in SF7 media supplemented with 5% FBS. MCF10A cells (obtained from J Brugge, Harvard University, Cambridge, MA) were maintained in phenol-free DMEM/F12 nutrient mix supplemented with 5% horse serum, 10 mg/ml insulin, 0.5 mg/ml hydrocortisone, 100 ng/ml cholera toxin, 20 ng/ml EGF (all Sigma), and 1% penicillin/streptomycin (Life Technologies). 3D assays of human mammary cells were performed by culturing the cells in the presence of irradiated 3T3 fibroblasts for 8, 10 or 14 days in Matrigel (Corning) SF7 media supplemented with 5% FBS as previously described [21]. MDA-MB-231 cells were obtained from S. Dunn (Child and Family Research Institute, Vancouver, BC) and maintained in DMEM with 10% FBS. Their identity was confirmed by DNA sequencing, including detection of the $KRAS^{G13D}$ allele[9].

Transduction and transfection

Primary cells were transduced with lenti-viral vectors prepared and used as previously described[14]. For stable inhibition of YBX1 transcripts, lentiviral vectors encoding sh*YBX1* (sc-38634-V, Santa Cruz, #1307/TRCN0000315307, #1309/TRCN0000315309, and #948/TRCN0000007948, Sigma).or sh*Scr* (sc-108080, Santa Cruz or shScr (SHC202, Sigma) were purchased or constructed.

Xenografts

Female immunodeficient NOD-Rag1$^{-/-}$IL2Rγc$^{-/-}$ (NRG) mice were bred and housed in the animal facility at the British Columbia Cancer Research Centre under SPF conditions. Surgery was performed on 5- to 10-week-old mice. All experimental procedures were approved by the University of British Columbia Animal Care Committee.

To generate primary tumours, enzymatically dissociated human mammary cell suspensions were prepared, transduced and transplanted into mice subcutaneously in 50% (v/v) Matrigel [14]. To measure tumour bioluminescence from the co-transduced luciferase cDNA, mice were injected intraperitoneally with 150 mg/kg body weight of d-luciferin (Promega) and 10 minutes later the mice were imaged using a Xenogen IVIS Lumina system with Living Image version 3.0 software (Caliper Life Sciences). To prepare single cell suspensions from excised tumours, the tissue was minced with a scalpel, incubated at 37 °C in DMEM/F12 media supplemented with 5% FBS and 300 U/ml collagenase and 100 U/ml hyaluronidase for 1 to 2 hours with periodic vortexing, washed with HF, and treated with 2.5 mg/ml trypsin with 1 mM EDTA and 5 mg/ml dispase with 100 µg/ml DNaseI. Human cells were sorted after staining with

anti-human specific antibodies directed against EpCAM and CD298 (Biolegend) with simultaneous depletion of mouse cells stained with anti-mouse-specific antibodies directed against CD45 and CD31 (Biolegend).

IHC staining

Pieces of tumours obtained from mice or normal breast were fixed in 10% buffered formalin (Fisher), washed in 70% ethanol and embedded in paraffin. Sections of paraffin-embedded tissue (3 mm) were first treated with Target Retrieval solution (DAKO) and then a cytomation serum-free protein block (DAKO) followed by staining with specific antibodies recognizing human YB-1 (#HPA040304, Sigma), ER (SP1; 1/50; Thermofisher; RM9101), PR (SP2; 1/50; Neomarker; 9102), Ki67 (SP6; 1/50; Thermofisher; RM9106), CK14 (Novocastra/Leica; 1/50; NCL-L-LL02), CK8/18 (Novocastra/Leica; 1/50; NCL-L-5D3), p63 (4A4; 1/50; Gentex; GTX23239), SMA (1A4; 1/100; Dako; MO851), HIF1a (Novus Biologicals; 1/150; NB100-134), CAIX (R&D Systems; 1/200; AF2188), CD34 (Abcam; 1/400; Ab81289). A secondary mouse or rabbit antibody conjugated to horseradish peroxidase and treatment with 3,3′-diaminobenzidine (DAB, DAKO) was used to obtain a positive brown staining. Negative IgG controls were performed on normal reduction mammoplasty tissue.

Quantitative analysis of IHC samples was conducted using the colour deconvolution plugin which implements stain separation and the ImmunoRatio plugin for ImageJ software (developed at the National Institutes of Health, USA, and available at http://rsb.info.nih.gov/ij/). Student's t-test was used for data analysis, unless indicated otherwise.

Plasmids

Inducible *KRAS*^{G12D}- and YB-1- encoding vector was derived from pINDUCER21 backbone [22] by replacing the attR1-ORF-attR2 cassette with a Kras-2A-KO2 fragment.

WB and densitometry analysis

After the required treatment, cells were washed with cold PBS and incubated for 15 minutes at 4°C with RIPA lysis buffer (30 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10% glycerol, 1% Triton X-100 (Sigma) supplemented with a 1 mM NaF, 1 mM NaVO3 and 1 mM PMSF (all Sigma). Cells extracts were centrifuged at 13,000 g for 10 minutes at 4°C. The protein concentration of the supernatant fraction was determined using the Bio-Rad Bradford Protein Assay Kit according to the manufacturer's instructions. For each sample, an equal amount of total protein was diluted in sample buffer (Invitrogen) and boiled for 5 minutes. Samples were loaded onto precast NuPAGE 4-12% polyacrylamide gels (Invitrogen). After electrophoresis, proteins were transferred to a PVDF transfer membrane. Membranes were then blotted overnight at 4°C with appropriate primary antibodies, such as anti-ACTIN (Santa Cruz, sc-1615, 1/10,000), anti-H3 (Cell Signaling Technology, 12648, 1/10,000), anti-RAS (Cell Signaling Technologies, 3339, 1/1,000), anti-YB-1 (Cell Signaling Technology, 4202, 1/1,000) and anti-HIF1a (Cayman; 10006421; 1/1000). Specific binding of antibodies was detected using appropriate secondary antibodies conjugated to horseradish peroxidase, and visualized with SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermofisher) on a ChemiDoc Gel Imaging system (Bio-rad). Densitometric analyses of immunoblots were performed using ImageJ.

RNAseq data

Copy number alterations and Z-score normalized RNAseq expression values (V2 RSEM) were

obtained from cBioPortal[23], from TCGA[5], METABRIC[6] and Mutational profiles of metastatic breast cancers[7] datasets. Paired-end reads were generated from MDA-MB-231 shScr or shYBX1 cells or derived tumors, on an Illumina HiSeq2500 sequencer. Read sequences were aligned to the hg19 human reference using the BWA-SW algorithm[24] to generate binary alignment/map (BAM) files. Transcript counts were obtained with the summarizeOverlaps function from GenomicAlignments package[25]. Differential expression analysis was performed with DESeq2 package[26].

Proteomic data

Tissues were thawed and lysed in 100 µL lysis buffer containing 500 mM Tris-HCL pH 8, 2% SDS (w/v), 1% NP-40 (v/v), 1% Triton X100 (v/v), 0.5 mM EDTA, 50 mM NaCl, 10 mM Tri(2-carboxyethyl)phosphine (TCEP) and 40 mM chloroacetamide (CAA). The proteins were then denatured by heating at 95°C for 90 minutes with shaking at 1,100 rpm before incubation at room temperature for 90 minutes in the dark to allow reduction and alkylation of disulfide bonds by TCEP and CAA respectively. SP3 beads[27,28] were added and the tissues were sonicated in a Bioruptor Pico (Diagenode) for 10 cycles (30 seconds ON, 30 seconds OFF). The samples were purified and prepared for trypsin digestion using the SP3 method[28]. Tryptic peptides from each sample were individually labeled with TMT 10-plex labels (Thermo Scientific), pooled, and fractionated into 12 fractions by high pH RP-HPLC, desalted, and then analyzed using an Easy-nLC1000 liquid chromatograph (LC) (Thermo Scientific) coupled to a Orbitrap Fusion Tribrid mass spectrometry (MS) (Thermo Scientific) operating in MS3 mode. The offline peptide fractionation and LC-MS conditions are as described[28]. The raw MS data were searched using Proteome Discoverer (version 2.1.1.21) using the embedded Sequest HT algorithm against a

combined UniProt Human proteome database with a list of common contaminants appended (24,624 total sequences). Sequest HT parameters were specified as: trypsin enzyme, allowance for 2 missed cleavages, minimum peptide length of 6, precursor mass tolerance of 20 ppm, and a fragment mass tolerance of 0.6. Dynamic modifications allowed were oxidation of methionine residues, and TMT at lysine residues and peptide N-termini. Carbamidomethylation of cysteine residues was set as a static modification. Peptide spectral match (PSM) error rates were determined using the target-decoy strategy coupled to Percolator modeling of positive and false matches[29,30]. Data were filtered at the PSM-level to control for false discoveries using a q-value cutoff of 0.05 as determined by Percolator. Contaminant and decoy proteins were removed from all datasets prior to downstream analysis. Statistical analysis of differential protein expression was performed at the peptide level using a modified version of the PECA function that is appropriate for input of log-transformed data[31]. PECA uses Limma[32] to generate a linear model for estimating fold changes and standard errors prior to empirical Bayes smoothing. Median t-statistics of the assigned peptides were used to calculate false-discovery rate-adjusted *p*-values determined from the beta distribution, as described previously[31].

RT-PCR

Total RNA was extracted from cryopreserved tumour samples or cultured cells using the Total RNA Isolation Micro kit (Agilent) and cDNA then synthesized using SuperScript VILO cDNA synthesis kit (Life Technologies). RT-PCR was performed using a SYBR Green master mix (Applied Biosystems) and samples run in triplicate with custom-designed primers.

Statistical analyses

Values are expressed as mean ± SEM, unless otherwise specified. Significance was evaluated using Student's t-test, unless otherwise specified. $*P<0.05$, $**P<0.01$, ns = not significant.

REFERENCES

1. Lovett, D. H., Cheng, S., Cape, L., Pollock, A. S. & Mertens, P. R. YB-1 alters MT1-MMP trafficking and stimulates MCF-7 breast tumor invasion and metastasis. *Biochem. Biophys. Res. Commun.* **398**, 482–488 (2010).
2. Wu, Y. *et al.* Strong YB-1 expression is associated with liver metastasis progression and predicts shorter disease-free survival in advanced gastric cancer. *J Surg Oncol* **105**, 724–730 (2012).
3. El-Naggar, A. M. *et al.* Translational Activation of HIF1α by YB-1 Promotes Sarcoma Metastasis. *Cancer Cell* **27**, 682–697 (2015).
4. Somasekharan, S. P. *et al.* YB-1 regulates stress granule formation and tumor progression by translationally activating G3BP1. *J. Cell Biol.* **208**, 913–929 (2015).
5. Ciriello, G. *et al.* Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* **163**, 506–519 (2015).
6. Pereira, B. *et al.* The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun* **7**, 11479 (2016).
7. Lefebvre, C. *et al.* Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. *PLoS Med.* **13**, e1002201 (2016).
8. Wykoff, C. C. *et al.* Hypoxia-inducible expression of tumor-associated carbonic anhydrases. *Cancer Res.* **60**, 7075–7083 (2000).
9. Kozma, S. C. *et al.* The human c-Kirsten ras gene is activated by a novel mutation in codon 13 in the breast carcinoma cell line MDA-MB231. *Nucleic Acids Res.* **15**, 5963–5971 (1987).
10. Lim, J. P. *et al.* YBX1 gene silencing inhibits migratory and invasive potential via CORO1C in breast cancer in vitro. *BMC Cancer* **17**, 201 (2017).

11. Kucharzewska, P., Christianson, H. C. & Belting, M. Global Profiling of Metabolic Adaptation to Hypoxic Stress in Human Glioblastoma Cells. *PLOS ONE* **10**, e0116740 (2015).
12. Park, J. S., Pasupulati, R., Feldkamp, T., Roeser, N. F. & Weinberg, J. M. Cyclophilin D and the mitochondrial permeability transition in kidney proximal tubules after hypoxic and ischemic injury. *Am. J. Physiol. Renal Physiol.* **301**, F134-150 (2011).
13. Cui, H., Chen, B., Chicoine, L. G. & Nelin, L. D. Overexpression of cationic amino acid transporter-1 increases nitric oxide production in hypoxic human pulmonary microvascular endothelial cells. *Clin. Exp. Pharmacol. Physiol.* **38**, 796–803 (2011).
14. Nguyen, L. V. *et al.* Barcoding reveals complex clonal dynamics of de novo transformed human mammary cells. *Nature* **528**, 267–271 (2015).
15. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
16. Coticchia, C. M., Revankar, C. M., Deb, T. B., Dickson, R. B. & Johnson, M. D. Calmodulin modulates Akt activity in human breast cancer cell lines. *Breast Cancer Res Treat* **115**, 545–560 (2009).
17. Evdokimova, V. *et al.* Translational activation of snail1 and other developmentally regulated transcription factors by YB-1 promotes an epithelial-mesenchymal transition. *Cancer Cell* **15**, 402–415 (2009).
18. Mahara, S. *et al.* HIFI-α activation underlies a functional switch in the paradoxical role of Ezh2/PRC2 in breast cancer. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E3735-3744 (2016).
19. Eirew, P. *et al.* A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. *Nat. Med.* **14**, 1384–1389 (2008).

20. Kannan, N. *et al.* The luminal progenitor compartment of the normal human mammary gland constitutes a unique site of telomere dysfunction. *Stem Cell Reports* **1**, 28–37 (2013).

21. Kannan, N. *et al.* Glutathione-dependent and -independent oxidative stress-control mechanisms distinguish normal human mammary epithelial cell subsets. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7789–7794 (2014).

22. Meerbrey, K. L. *et al.* The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 3665–3670 (2011).

23. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**, 401–404 (2012).

24. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).

25. Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).

26. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

27. Hughes, C. S. *et al.* Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* **10**, 757 (2014).

28. Moggridge, S., Sorensen, P. H., Morin, G. B. & Hughes, C. S. Extending the Compatibility of the SP3 Paramagnetic Bead Processing Approach for Proteomics. *J. Proteome Res.* **17**, 1730–1740 (2018).

29. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S. & MacCoss, M. J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* **4**, 923–925 (2007).

30. Spivak, M., Weston, J., Bottou, L., Käll, L. & Noble, W. S. Improvements to the percolator algorithm for Peptide identification from shotgun proteomics data sets. *J. Proteome Res.* **8**, 3737–3745 (2009).

31. Suomi, T., Corthals, G. L., Nevalainen, O. S. & Elo, L. L. Using Peptide-Level Proteomics Data for Detecting Differentially Expressed Proteins. *J. Proteome Res.* **14**, 4564–4570 (2015).

32. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3**, Article3 (2004).

FIGURE LEGENDS

Fig. 1. Patients' *KRAS*-amplified tumours express high levels of *YBX1*. (**a**) Kaplan-Meier curves of overall survival (OS) for the TCGA cohort, with respect to *KRAS* copy number (N=206 for tumours with amplified *KRAS* or a gain of function *KRAS* gene, and N=522 for tumours with diploid *KRAS*). (**b-c**) *YBX1* mRNA levels compared to *KRAS* copy number status (right panel) in invasive breast carcinoma samples in the TCGA dataset (**b**)**,** or in the METABRIC dataset (**c**). Values for *YBX1* are shown as RPKMs. (**d**) Representative table of *YBX1* alteration co-occurrence in metastatic breast cancer. (**e**) *HIF1A* (left panel) *CAIX* (middle panel) and *VEGFA* (right panel) mRNA expression according to *KRAS* copy number status in invasive breast carcinomas in TCGA dataset. Values for *HIF1A, CAIX and VEGFA* are shown as RPKMs. (**f**) Scatter plot of *YBX1* and *CAIX* mRNA expression in amplified-*KRAS* invasive breast carcinomas.

Fig. 2. Suppressed YBX1 expression impairs generation of local and metastatic tumour formation by an aggressive breast cancer cell line. (**a**) Western blots of MDA-MB-231 cells transduced with sh*YBX1(#1307)*, sh*YBX1(#1309)* or shY*BX1(#948)* or a control sh*Scr* vector. (**b**) Representative pictures of bioluminescence signals in mice injected subcutaneously with MDA-MB-231 cells transduced with sh*YBX1(#1307)*, sh*YBX1(#1309)*, sh*YBX1(#948)* or sh*Scr*. Dot plots shows the bioluminescence measured in these tumours 34 days post-transplant. (**c**) Representative views of YB-1 levels in the IHC-stained *in vivo* progeny of sh*YBX1*- or sh*Scr*-transduced MDA-MB-231 cells. Bar graph shows quantification of YB-1 expression. (**d**) Representative pictures of bioluminescence signals obtained in mice 34 days after being injected

intravenously with MDA-MB-231 cells transduced with sh*YBX1 or* sh*Scr* vectors. Dot plot shows the measured bioluminescence in the progeny of these cells generated *in vivo*. (**e**) Representative photomicrographs of H&E- or YB-1-stained sections of lungs of mice injected intravenously with MDA-MB-231 cells transduced with sh*YBX1* or sh*Scr* vectors. Dot plot shows the number of metastasis per mouse (2 lungs). (**f**) Representative images of CAIX-stained sections of tumours produced from sh*YBX1*- or sh*Scr*-transduced MDA-MB-231 cells. Scale bar, 100 μm. Bar graph shows measured levels of CAIX (staining intensity). (**g**) Western blots showing HIF1α and YB-1 levels (relative to GAPDH) in sh*YBX1* (#1307)- and sh*Scr*-transduced MDA-MB-231 cells maintained for up to 24 hours in 1% O_2.

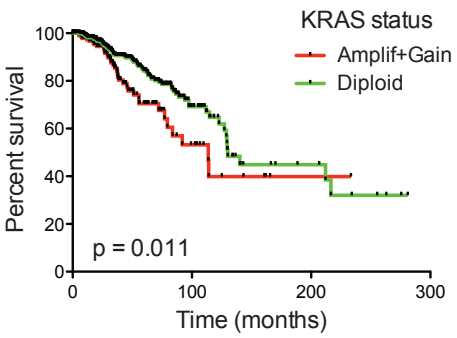
Fig. 3. *KRAS^{G12D}* upregulates YB-1 in *de novo KRAS^{G12D}*-transformed normal human mammary cells. (**a**) Representative views of YB-1 immunostaining of normal human mammary tissue (left) and 8-week tumours derived from *KRAS^{G12D}*-transduced normal mammary cells isolated from the same matching 3 different donors (right). Scale bar, 50 μm. (**b**) Representative immunofluorescence images of control and *KRAS^{G12D}*-transduced BCs (top) and LPs (bottom) assessed 15 days post-transduction, and cultured in 3D in Matrigel. Staining was performed using an anti-YB-1 antibody, Phalloidin (F-ACTIN) and DAPI (DNA). (**c**) Representative FACS profile of a 4 week-old graft of inducible *KRAS^{G12D}*-transduced human BCs obtained from mice maintained on doxycyline-supplemented water (Dox) for 2 or 4 weeks post-transplant. (**d**) *KRAS* mRNA levels measured in 4 week-old grafts of inducible *KRAS^{G12D}*-transduced cells obtained from mice maintained on doxycyline-supplemented water for 0, 2, or 4 weeks, as shown. (**e**) Representative views of YB-1 immunostaining of 4 week-old tumours derived from inducible *KRAS^{G12D}*-transduced cells in mice maintained on doxycyline-supplemented water for 2 or 4

weeks (N=3 donors). Scale bar, 50 μm.

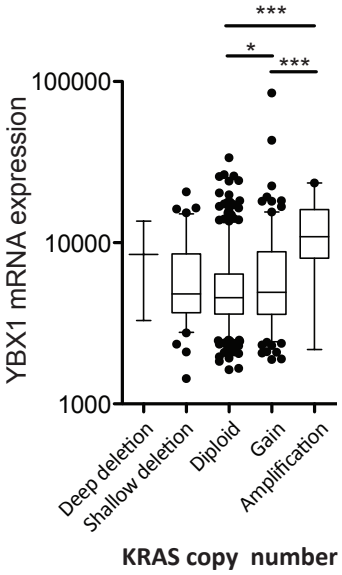
Fig. 4. Increased YB-1 is necessary but neither sufficient nor restricted to *KRAS^{G12D}*-induced mammary tumorigenesis. (**a**) Representative pictures of bioluminescence signals in mice injected subcutaneously with *KRAS^{G12D}*+sh*YBX1*- or *KRAS^{G12D}*+sh*Scr*-transduced primary cells 2 weeks earlier. Dot plot showing bioluminescence activity of tumours derived from BCs (blue) and LPs (red). (**b-c**) Representative images of HIF1α (**b**)- and CAIX (**c**)- stained sections from different BC- or LP-derived tumours arising from *KRAS^{G12D}*+sh*YBX1*- or *KRAS^{G12D}*+sh*Scr*-transduced cells. Scale bar, 100 μm. Bar graphs shows quantification of HIF1α (**b**)- and CAIX (**c**)- intensities in tumours derived from *KRAS^{G12D}*+sh*YBX1*- or *KRAS^{G12D}*+sh*Scr* -transduced cells. N = 8. (**d**) Representative photos of bioluminescence signals measured in mice injected subcutaneously 7 weeks earlier with human mammary cells transduced with Luc-YFP alone or in combination with *KRAS^{G12D}*, *YBX1*, *KRAS^{G12D}*+*YBX1*, or *myrAKT1*. BCs and LPs from 3 donors were pooled before transduction. Graph plot shows changes in bioluminescence activity over time. (**e**) Representative images of sections of *myrAKT1*-derived tumours initiated from either BCs or LPs and stained for SMA, CK14, CK8/18 and p63. Scale bar, 100 μm. (**f**) Representative views of YB-1 immunostaining of 18-week tumours generated from normal mammary cells from 3 different donors (#1-3) transduced with *myrAKT1*. Bar graph shows a comparison of YB-1 staining intensity in tumours derived from *KRAS^{G12D}*- or *myrAKT1*-transduced cells. N = 10 (*KRAS^{G12D}*) and 6 (*myrAKT1*) tumours.

Figure 1

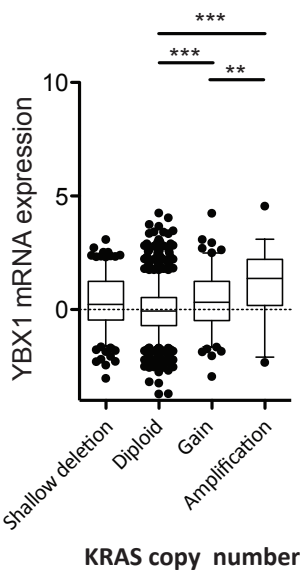
a



b



c



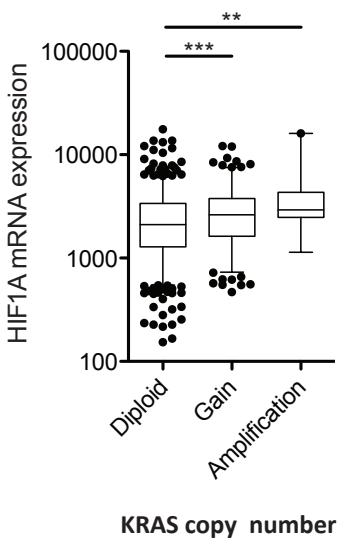
d

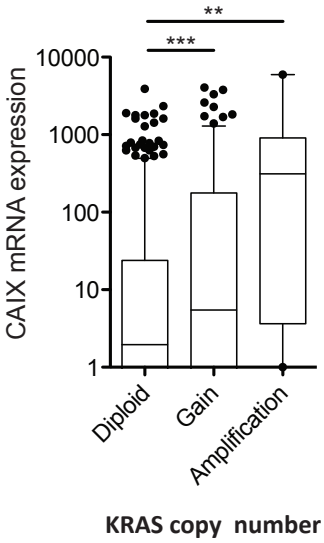
YBX1 co-occurrent alterations

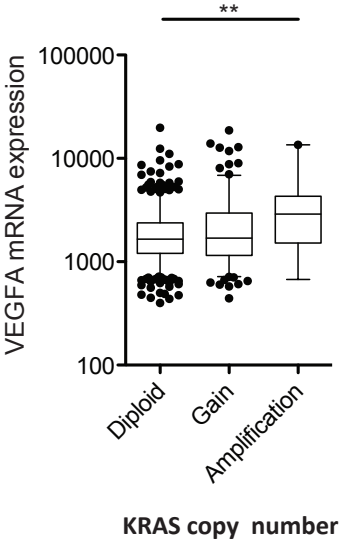
Gene	p-value	Log odds ratio
KRAS	0,026	2.71
AKT1	0,239	1.61
TP53	0,150	1.43
PIK3CA	0,609	-0.37
ERBB2	0,369	1.04

Metastatic Breast Cancer

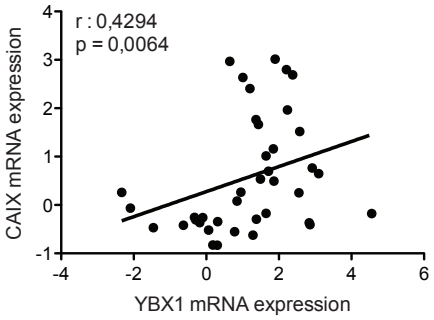
e







f



KRAS amplified tumors

Figure 2

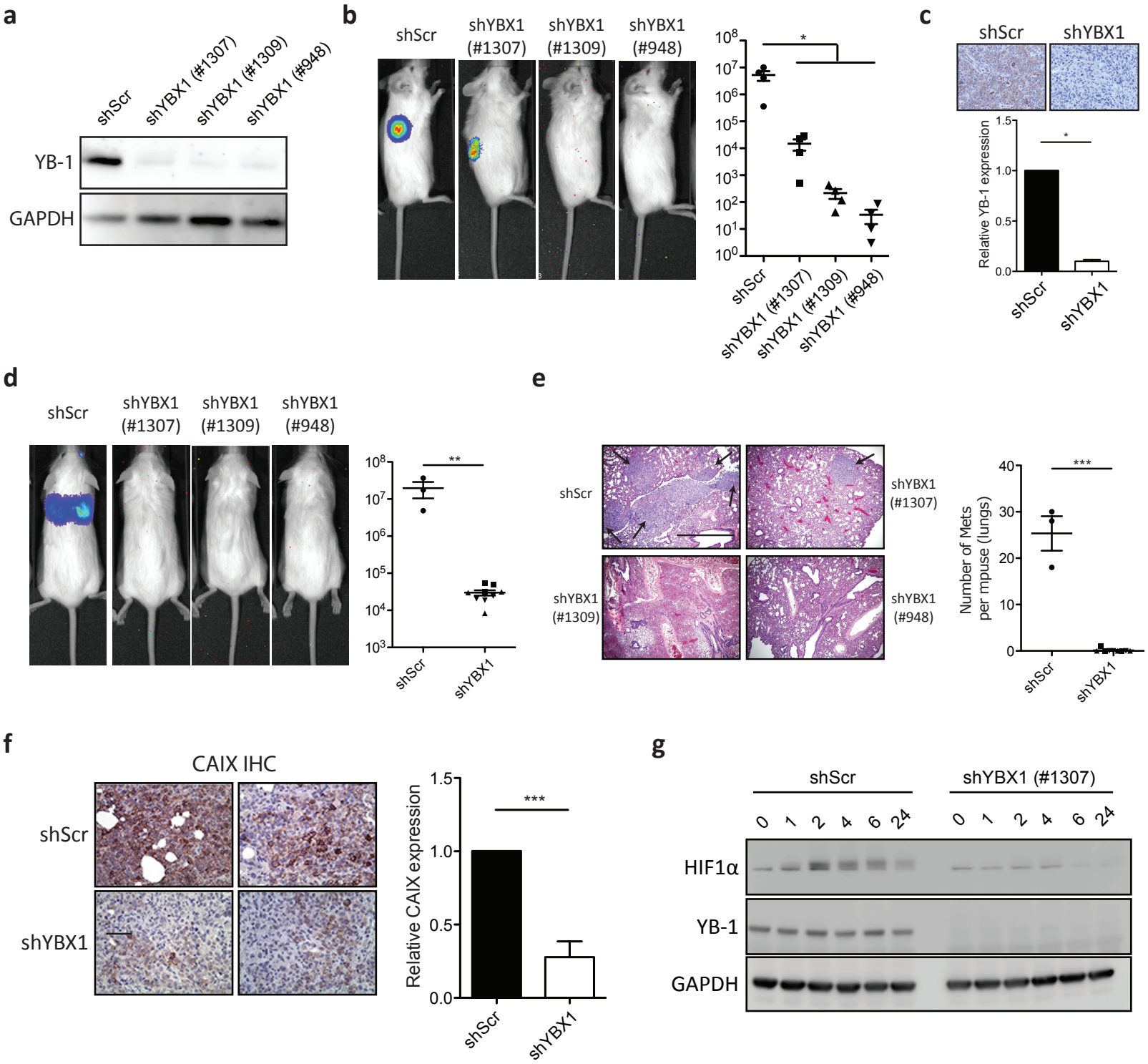

a
shScr
shYBX1 (#1307)
shYBX1 (#1309)
shYBX1 (#948)
YB-1
GAPDH
b
shScr
shYBX1 (#1307)
shYBX1 (#1309)
shYBX1 (#948)
*
c
shScr
shYBX1
Relative YB-1 expression
*
d
shScr
shYBX1 (#1307)
shYBX1 (#1309)
shYBX1 (#948)
**
e
shScr
shYBX1 (#1307)
shYBX1 (#1309)
shYBX1 (#948)
Number of Mets per mpuse (lungs)
***
f
CAIX IHC
shScr
shYBX1
Relative CAIX expression
***
g
shScr
shYBX1 (#1307)
0 1 2 4 6 24
HIF1α
YB-1
GAPDH

Figure 3

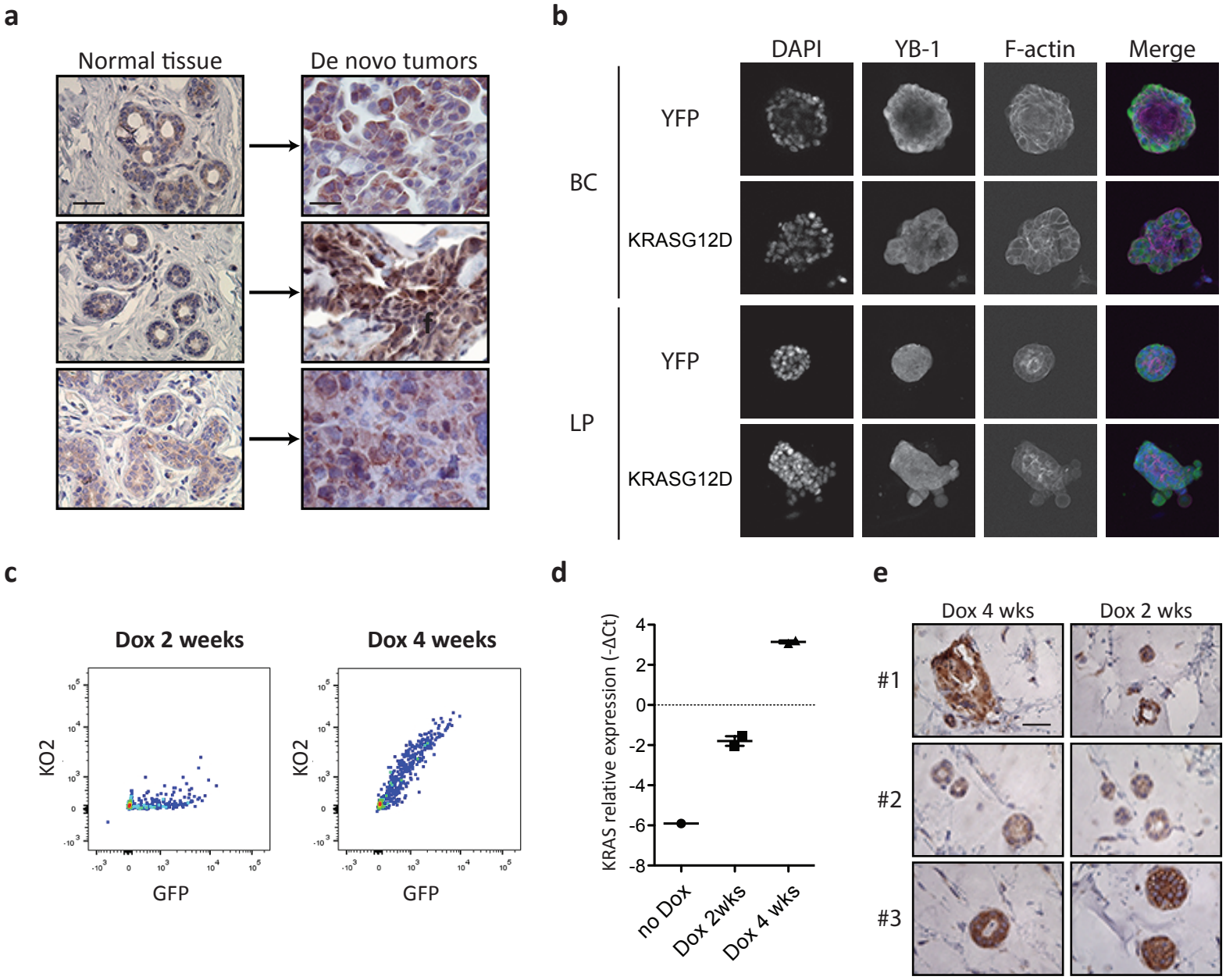
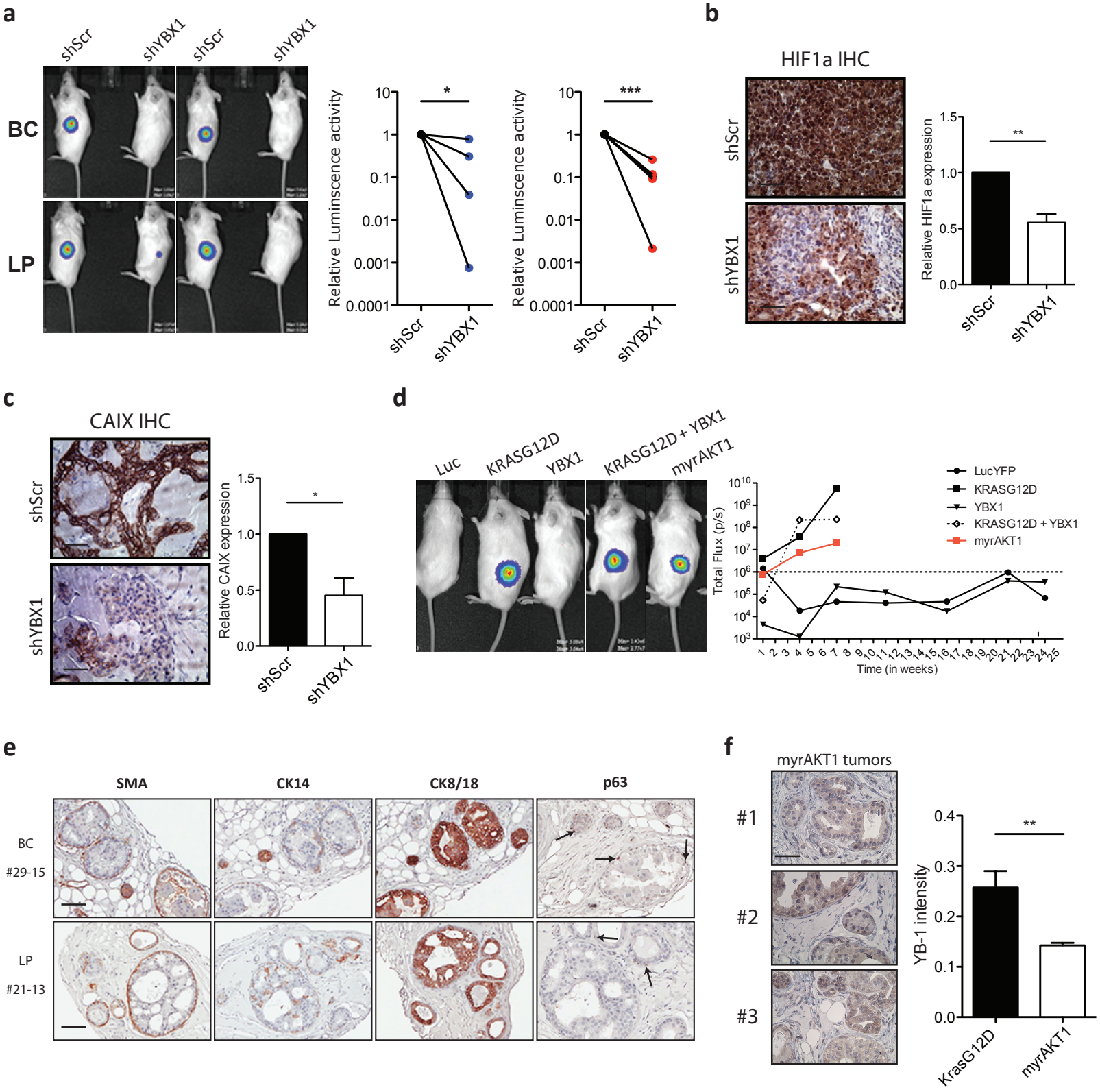

a
Normal tissue
De novo tumors
b
DAPI
YB-1
F-actin
Merge
BC
YFP
KRASG12D
LP
YFP
KRASG12D
c
Dox 2 weeks
Dox 4 weeks
KO2
GFP
d
KRAS relative expression (-ΔCt)
4
2
0
-2
-4
-6
-8
no Dox
Dox 2wks
Dox 4 wks
e
Dox 4 wks
Dox 2 wks
#1
#2
#3

Figure 4


a
shScr
shYBX1
shScr
shYBX1
BC
LP
Relative Luminscence activity
10
1
0.1
0.01
0.001
0.0001
*
***
shScr
shYBX1
b
HIF1a IHC
shScr
shYBX1
Relative HIF1a expression
1.5
1.0
0.5
0.0
**
c
CAIX IHC
Relative CAIX expression
*
d
Luc
KRASG12D
YBX1
KRASG12D + YBX1
myrAKT1
Total Flux (p/s)
Time (in weeks)
LucYFP
KRASG12D
YBX1
KRASG12D + YBX1
myrAKT1
e
SMA
CK14
CK8/18
p63
BC
#29-15
LP
#21-13
f
myrAKT1 tumors
#1
#2
#3
YB-1 intensity
**
KrasG12D
myrAKT1