A CRITICAL ROLE OF YB-1 IN THE GENESIS AND PROGRESSION OF KRAS-MUTATED 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

ABSTRACT

The extensive heterogeneity of spontaneously arising human breast cancers has made it difficult to identify specific mechanisms that determine their malignant properties. We now show that invasive tumours developing in immunodeficient mice from freshly isolated normal human mammary cells transduced with a *KRAS*^{G12D} vector show increased expression of the YB-1 RNA-binding protein within two weeks and is sustained in subsequent tumour passages, thus mimicking advanced human breast cancers with KRAS pathway deregulation. YB-1 is also rapidly upregulated in a new *de novo* model of human ductal carcinoma *in situ* that we show is obtained from similar xenotransplants of *myrAKT1*-transduced primary human mammary cells. Knockdown studies demonstrated that YB-1 is essential for both the initial transforming activity of *KRAS*^{G12D} in primary human mammary cells and the metastatic activity of an established human *KRAS*^{mutant} breast cancer cell line. Accompanying molecular and histological analyses indicate YB-1-mediated activation of a HIF1α response.

INTRODUCTION

Mammalian Y-box binding protein-1 (YB-1) is a member of the family of DNA/RNA binding proteins with an evolutionarily conserved cold-shock domain (CSD). Mammalian CSD proteins are widely expressed and involved in many fundamental processes including DNA repair as well as mRNA transcription, splicing, translation, and stabilization[1,2]. Elevated YB-1 expression correlates with poor patient survival and drug resistance in diverse tumour types, most notably in metastatic tumours[3,4]. In high-risk sarcomas, YB-1 drives metastasis by inducing a hypoxia-inducible factor 1α (HIF1α) response[5], via its ability to bind directly to *HIF1A* mRNA and thereby activate YB-1 translation under hypoxia. YB-1 also induces stress granule formation through translational activation of *G3BP1* mRNAs present in sarcomas and other tumour types[6]. YB-1 also binds to *SNAIL1* and *TWIST* mRNAs to activate their translation in immortalized but nontumorigenic MCF10A human mammary cells overexpressing *HRAS*[7]. This suggests how transformed breast cells would be stimulated to undergo an epithelial-to-mesenchymal transition (EMT), a consequent disruption of their ability to sustain an acinar architecture, and the acquisition of increased invasive and metastatic properties. However, a lack of models of *de novo* breast cancer development from freshly isolated normal human mammary cells has made it difficult to determine if and how YB-1 deregulation may contribute to the initial stages of human mammary cell transformation.

We recently developed an efficient and reproducible method for rapidly generating serially transplantable invasive ductal carcinomas (IDC) by lentiviral transduction of freshly isolated normal human mammary cells with a $KRAS^{G12D}$ cDNA[8]. This approach can be used to transforms two types of mammary cells with innate epidermal growth factor (EGF)-dependent proliferative potential, both of which can be readily isolated from normal human breast tissue.

These two cell types are referred to as basal cells (BCs) and luminal progenitors (LPs), the latter representing a phenotypically and biologically distinct subset of the luminal cell compartment[9]. Although *KRAS* mutations are limited to approximately 4% of all human breast cancers[10], altered *KRAS* expression has a high oncogenic score for many estrogen receptor-positive (ER^+) and ER^- tumours[11]. The tumours produced when either normal BCs or LPs forced to express $KRAS^{G12D}$ are transplanted into immunodeficient mice are highly polyclonal and phenotypically heterogeneous, with a variable content of cells positive for ER, heregulin-2 (HER2), EGFR, Ki67 and cytokeratins (CK) 8/18[8].

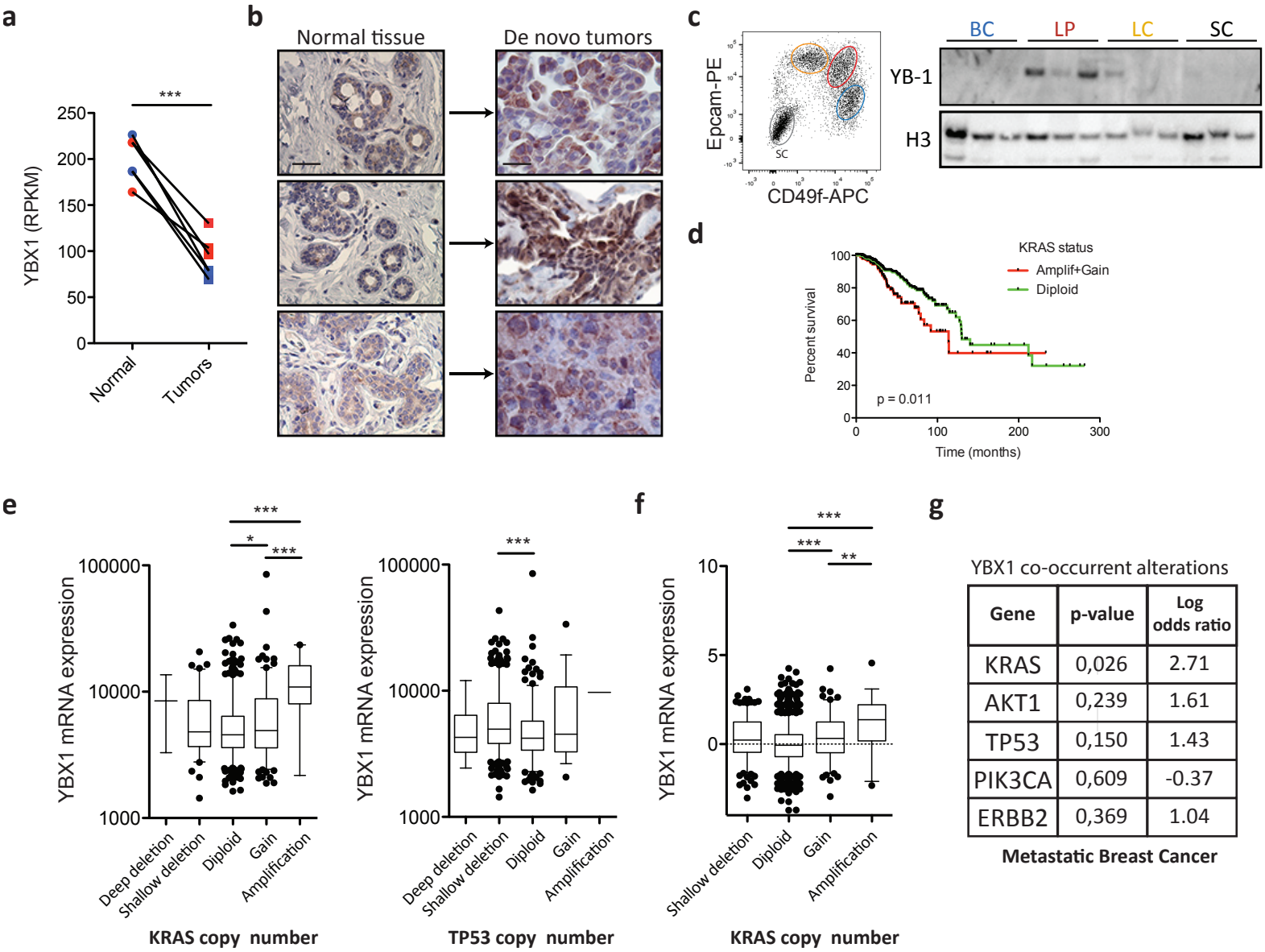
We now show that increased expression of YB-1 constitutes an early and pervasive requirement for human mammary cells expressing an oncogenic form of *KRAS*. This is required for these cells not only to initiate but also to sustain more aggressive malignant properties, even when other mutations have been acquired *a priori*. In addition, we present evidence that mutant KRAS induction of YB-1 expression is associated with the activation of a HIF1α program.

RESULTS

YB-1 protein expression is increased in human breast cancers with mutant KRAS

To investigate the initial effect of $KRAS^{G12D}$-induced transformation of human mammary cells on YB-1 expression, we first examined our previously published RNA-seq data for 3 paired isolates of BCs and LPs obtained by fluorescence-activated cell sorting (FACS) and matching primary tumours derived from each isolate[8]. Despite a consistent decrease in YB-1 transcript levels in the transformants as compared to the normal cells from which the tumours had been generated (Fig. 1a), immunohistochemical (IHC) staining showed consistent strong YB-1 protein expression in most of the tumour cells, mainly in the cytoplasm, although occasional tumour cells also showed

Figure 1



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

evidence of nuclear YB-1 (Fig. 1b). In contrast, similarly stained sections of normal human breast tissue showed YB-1 to be largely restricted to cells in the luminal layer of the gland, primarily in the cytoplasm and at low levels (Figure 1b, left panel). Western blot (WB) analyses of FACS-purified isolates of BCs and LPs, and also of the remaining luminal cells (LCs) of normal human mammary glands confirmed that YB-1 protein levels are highest in normal LPs and barely detectable in BCs or LCs (Fig. 1c).

To determine if upregulated YB-1 protein expression is associated specifically with an oncogenic form of KRAS in patients' breast cancers, we first examined the publicly available data for approximately 800 breast cancers in The Cancer Genome Atlas (TCGA) [12]. This revealed *YBX1* transcripts to be highest in ER^- tumours, and most notably in those that had metastasized (Supplementary Fig. 1a,b). Interestingly, in this dataset, elevated *YBX1* transcripts were positively associated with a gain of function or amplification of the *KRAS* gene, and reduced overall survival (Fig. 1d) as compared to patients whose tumours contained a normal diploid *KRAS* complement (Fig. 1e, left panel). Increased *YBX1* mRNA levels were also seen in tumours with amplified *ERB2, PIK3CA* or *AKT1*, or deletions of *TP53*, compared to diploid breast cancers (Fig. 1e, right panel and Supplementary Fig. 1c). However, elevated *YBX1* expression was not a feature of breast cancers containing *HRAS* or *NRAS* mutations (data not shown). Examination of the larger METABRIC dataset for 2,433 breast cancers [11] confirmed that increased *YBX1* transcripts are associated with *KRAS* amplification (Fig. 1f). In another large study of metastatic breast cancers [13], *YBX1* expression correlated with that of *KRAS*, but showed no association with alterations of *ERBB2, TP53, PIK3CA* or *AKT1* (Fig. 1g).

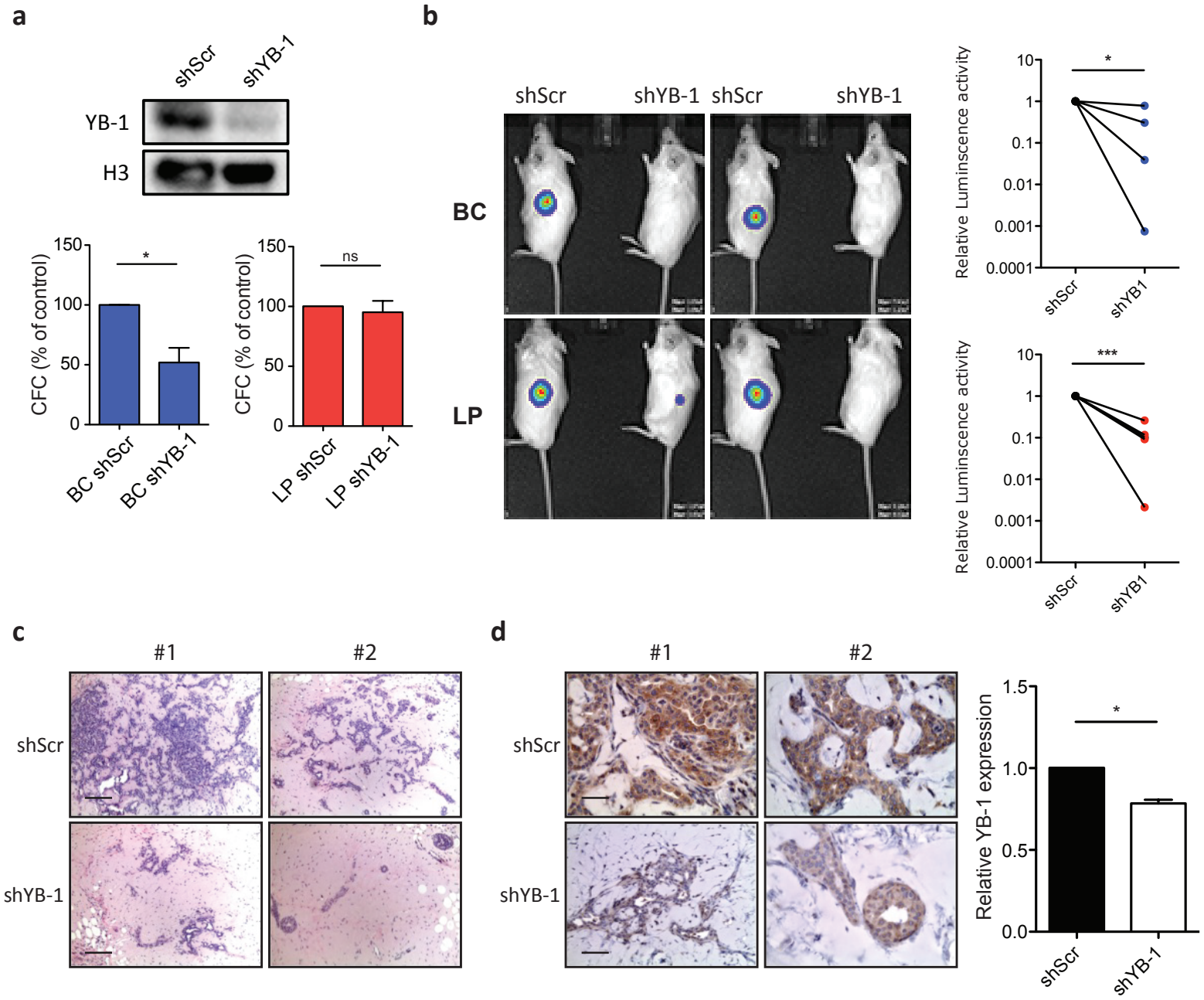
YB-1 is required for tumour initiation by *KRAS*^{G12D}-transduced normal human mammary

cells

To determine whether increased YB-1 expression contributes to the initial acquisition of *in vivo* tumorigenic activity by *KRAS*^{G12D}–transduced human mammary cells, we examined the effect of suppressed YB-1 expression in this model using a shRNA strategy. Accordingly, FACS-purified normal BCs and LPs were first separately transduced with *YBX1* targeting or control lenti-shRNA vectors (Fig. 2a). Then, after 2 days *in vitro* (to allow full expression of the introduced shRNA), the same cells were infected with both a *KRAS*^{G12D}-mCherry and a luciferase (Luc)-YFP vector. Cells were then immediately transplanted subcutaneously in Matrigel plugs into female NOD-*Rag1*$^{-/-}$*IL2Rgc*$^{-/-}$ (NRG) mice ($2x10^4$-$25x10^4$ cells/transplant). By week 2, the level of bioluminescence measured in mice injected with cells that had been co-transduced with *KRAS*^{G12D} and shYB-1 was already much lower (2-1,000-fold) than that evident in the recipients of control cells that had been transduced with *KRAS*^{G12D} and the scrambled shRNA vector. Moreover, this was the case regardless of whether BCs or LPs had been transduced (Fig. 2b). Haematoxylin and eosin (H&E) (Fig. 2c) and human YB-1-immunostained sections (Fig. 2d) of the transplants recovered from these experiments also showed a greatly reduced mammary cell content of the transplants derived from cells transduced with *KRAS*^{G12D} plus the shYB-1 construct as compared to *KRAS*^{G12D} plus the scrambled shYB-1 construct. The numbers of YB-1$^+$ cells in the test transplants were also reduced compared to their matched controls. These experiments thus establish upregulated YB-1 to be an important mediator of the early tumorigenic activity obtained by forced expression of *KRAS*^{G12D} in freshly isolated normal human mammary cells.

Increased YB-1 expression in KRAS^{G12D}-induced transformation of human mammary cells is rapid, sustained, and not necessarily associated with their subsequent proliferation

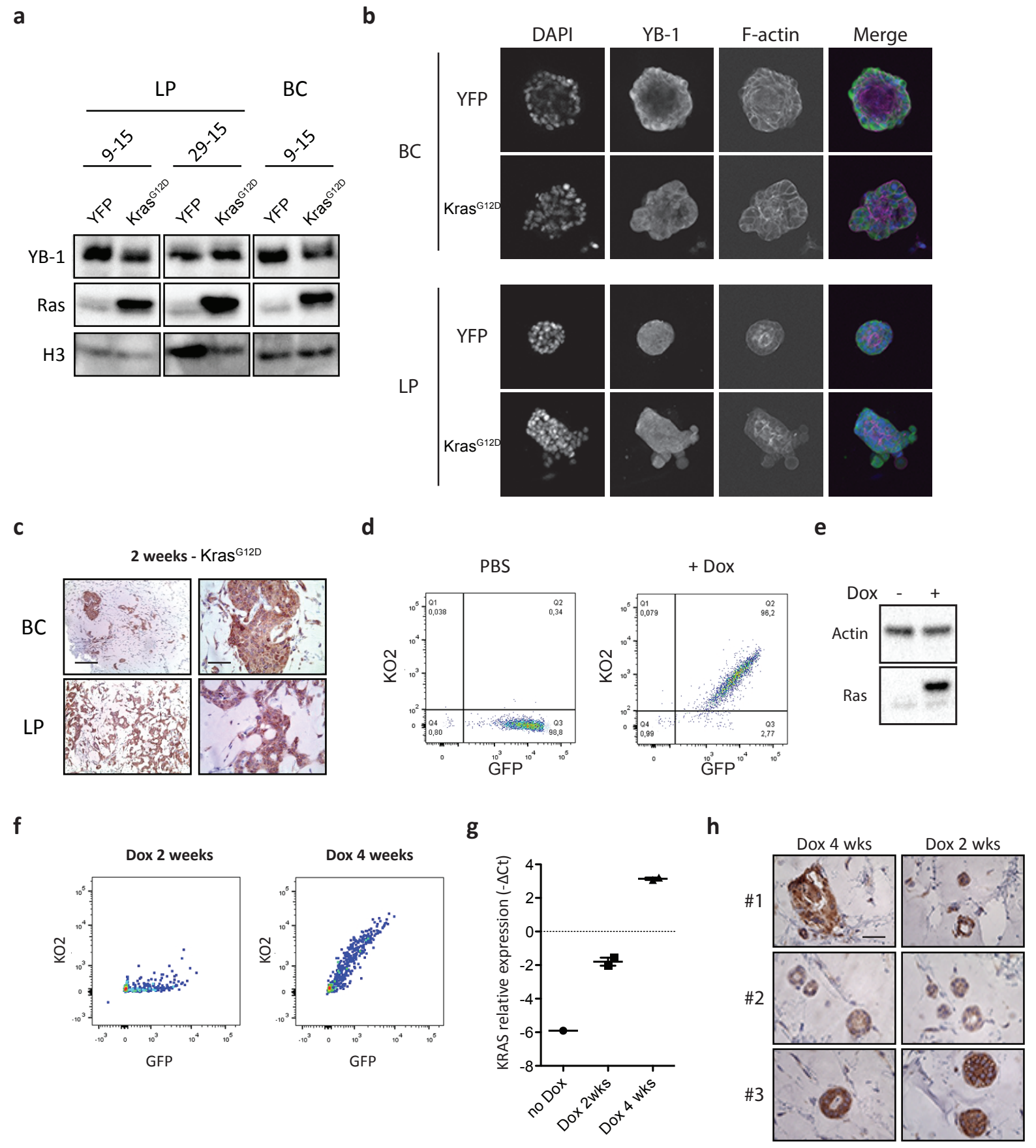
Figure 2


a
shScr
shYB-1
YB-1
H3
CFC (% of control)
150
100
50
0
*
ns
BC shScr
BC shYB-1
LP shScr
LP shYB-1
b
shScr
shYB-1
shScr
shYB-1
BC
LP
Relative Luminscence activity
10
1
0.1
0.01
0.001
0.0001
*
***
shScr
shYB1
c
#1
#2
shScr
shYB-1
d
#1
#2
shScr
shYB-1
Relative YB-1 expression
1.5
1.0
0.5
0.0
*
shScr
shYB-1

To determine if *KRAS*^{G12D} transduction also induces YB-1 expression and evidence of transformation in cells maintained *in vitro*, we transduced freshly isolated normal human mammary cells with *KRAS*^{G12D}, or a control vector, and then assessed YB-1 protein levels in cells maintained in EGF-supplemented 3D-Matrigel cultures. Surprisingly, WB of lysates of these cells obtained 3 days later (Fig. 3a) and also immunofluorescence analyses carried out after 15 days (Fig. 3b) failed to reveal increased YB-1 levels in the *KRAS*^{G12D}-transduced cells, despite their production of acini with filled lumens and a modified architecture (Fig. 3b). In sharp contrast, markedly increased expression of YB-1 was seen in the nascent *de novo* tumours recovered 2 weeks after transplanting the cells into mice (Fig. 3c), as compared to normal breast mammoplasty reduction tissue (Fig. 1b). Indeed, at 2 weeks post-transplantation, the rapidly increased level of YB-1 protein in the *KRAS*^{G12D}-transduced cells was already equivalent to that seen in the invasive tumours analyzed 4-6 weeks later (Fig. 1b and Fig. 3c). Thus, increased YB-1 expression in this *de novo* model of *KRAS*^{G12D}-induced human mammary cell transformation is rapid *in vivo* but not necessarily replicated when the cells are stimulated to proliferate *in vitro*.

We next asked whether the continued presence of KRAS^{G12D} is necessary to maintain the high levels of YB-1 initially obtained in tumours produced from *KRAS*^{G12D}-transduced primary human mammary cells. To address this question, we constructed an inducible *KRAS*^{G12D}-encoding vector that causes the Kusabira Orange (KO) fluorochrome to be co-expressed with *KRAS*^{G12D} under the control of doxycycline (Fig. 3d,e). We then transduced purified human mammary cells with this inducible vector and transplanted the cells subcutaneously with Matrigel plugs into NRG mice ($4x10^4$-$30x10^4$ cells each). The mice were then subdivided into three groups; one group was given regular water for 4 weeks, another was given doxycycline-supplemented water for 4 weeks, and the third one was given doxycycline-supplemented water

Figure 3


a
LP
BC
9-15
29-15
9-15
YFP
KrasG12D
YFP
KrasG12D
YFP
KrasG12D
YB-1
Ras
H3
b
DAPI
YB-1
F-actin
Merge
BC
YFP
KrasG12D
LP
YFP
KrasG12D
c
2 weeks - KrasG12D
BC
LP
d
PBS
+ Dox
KO2
GFP
Q1 0,038
Q2 0,34
Q4 0,80
Q3 98,8
Q1 0,079
Q2 96,2
Q4 0,99
Q3 2,77
e
Dox - +
Actin
Ras
f
Dox 2 weeks
Dox 4 weeks
KO2
GFP
g
KRAS relative expression (-ΔCt)
no Dox
Dox 2wks
Dox 4 wks
h
Dox 4 wks
Dox 2 wks
#1
#2
#3

for the first 2 weeks only, and then switched to regular water. All mice were sacrificed at the end of this 4-week period, and the transplants harvested and single cell suspensions prepared for FACS analysis to determine the number and frequency of KO^+ cells present. Cells isolated from mice receiving doxycycline just for the first 2 weeks of the test schedule, or not at all, displayed minimal KO positivity as compared to mice given doxycycline for the full 4-week period (Fig. 3f; right panel). Quantitative (Q)-PCR analysis showed a 30-fold lower level of RAS transcripts in the minimally KO^+ cells (Fig. 3g). Nevertheless, cells obtained from mice maintained on doxycycline for the first 2 weeks only showed high levels of YB-1 expression similar to those seen in the tumours expressing $KRAS^{G12D}$ for the full 4 weeks (Fig. 3H). Taken together, these experiments suggest that *in vivo*, $KRAS^{G12D}$ can stimulate a critically increased level of YB-1 that can be prolonged even in the absence of a continued expression of $KRAS^{G12D}$.

Increased YB-1 is a prominent, but not sufficient inducer of human mammary cell transformation

To investigate whether up-regulation of YB-1 occurs in other experimental models of oncogene-induced human mammary cell transformation, we screened several other oncogenes for their ability to induce tumours *de novo* from transplants of similarly transduced primary human cells. For this purpose, we selected cDNAs of genes previously identified as "drivers" from genomic analyses of patients' breast cancers[10]. To simplify the screen, paired pools of BCs and LPs isolated from 3 different normal breast tissue donors were transduced with lenti-Luc-YFP plus one of each of the test cDNAs in a mCherry vector in combination with one of the following: no other vector, plus our lenti-$KRAS^{G12D}$–YFP vector (to look for potential enhancing effects), or plus the lenti-$KRAS^{G12D}$-YFP vector (as a positive control), or plus the lenti-Luc-YFP vector only

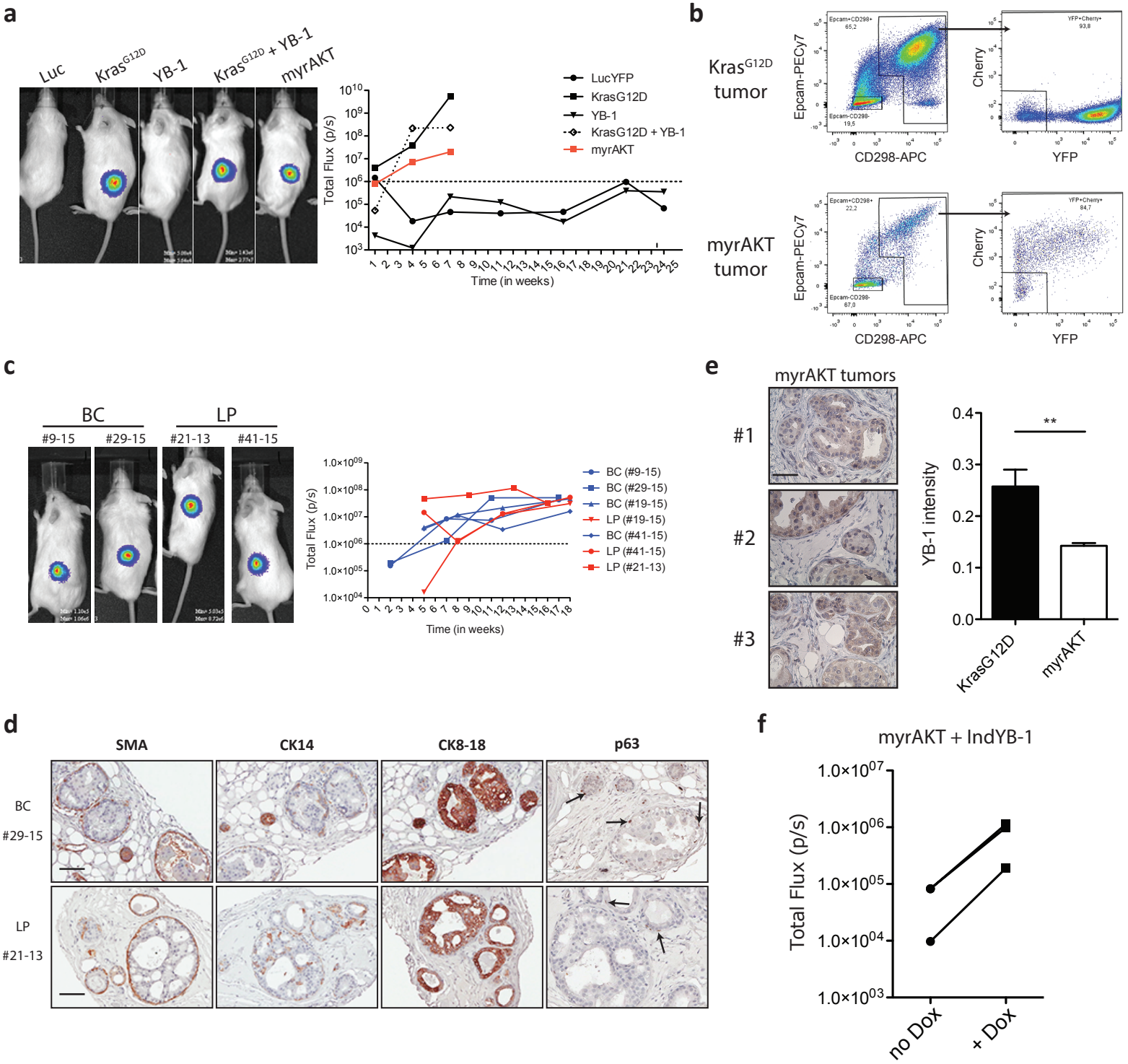
(as a negative control) (Fig. 4a).

Inclusion in this screen of a vector encoding YB-1 showed that forced expression of YB-1 alone was insufficient to induce tumour formation as indicated by a lack of increasing bioluminescence *in vivo*. Forced expression of *YB-1* and *KRAS*^{G12D} in the same cells also did not enhance the growth of the tumours that these cells produced as compared to the cells transduced with *KRAS*^{G12D} only (Fig. 4a). Neither dominant-negative forms of *TP53* nor mutant *PI3K* were able to induce tumorigenesis on their own, as found previously[8], and negative results were also obtained for several other vectors tested including those encoding cDNAs for *EGFR* and *c-MYC*, and shRNAs targeting *PTEN* and *BRCA1* transcripts (data not shown).

However, forced expression of a cDNA encoding myristoylated *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 obtained from *KRAS*^{G12D}-transduced cells (Fig. 4a). Constitutive activity of the myrAKT1 protein is attributed to a 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 [14]. Phenotypic analysis of tumour cells obtained from mice transduced with *myrAKT1* showed they were universally EpCAM^{+}CD298^{+} as well as mCherry^{+}, indicative of an oncogenic role of deregulated AKT1 activity in these cells (Fig. 4b). The modest tumorigenic activity of *myrAKT1* was readily replicated (9/13 tests) using either purified normal BCs or LPs obtained from single donors for transduction (Fig. 4c).

Histological analysis of the tumours produced from the *myrAKT1*-transduced cells appeared morphologically most similar to ductal carcinomas *in situ* (DCIS) with a confined

Figure 4


a
Luc
KrasG12D
YB-1
KrasG12D + YB-1
myrAKT
Total Flux (p/s)
Time (in weeks)
LucYFP
KrasG12D
YB-1
KrasG12D + YB-1
myrAKT
b
KrasG12D tumor
myrAKT tumor
Epcam-PECy7
CD298-APC
Cherry
YFP
c
BC
LP
#9-15
#29-15
#21-13
#41-15
Total Flux (p/s)
Time (in weeks)
BC (#9-15)
BC (#29-15)
BC (#19-15)
LP (#19-15)
BC (#41-15)
LP (#41-15)
LP (#21-13)
e
myrAKT tumors
#1
#2
#3
YB-1 intensity
KrasG12D
myrAKT
**
d
SMA
CK14
CK8-18
p63
BC
#29-15
LP
#21-13
f
myrAKT + IndYB-1
Total Flux (p/s)
no Dox
+ Dox

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 PR^+ cells (Supplementary Fig. 2a). These DCIS-like structures also contained cells with the basal features of smooth muscle actin (SMA) and TP63 mainly in the outer layer (Fig. 4d and Supplementary Fig. 2b), and cells with the luminal features of strong CK14 and CK8/18 positivity more centrally (Fig. 4d and Supplementary Fig. 2b). These structures also showed increased YB-1 protein as compared to normal cells. but less than in the *KRAS*^{G12D}-induced tumours (Fig. 4e and 1b). Notably, tumours from cells co-transduced with the doxycycline inducible KO/YB-1 vector (Supplementary Fig. ^{2c}) plus the lenti-*myrAKT1*-mCherry vector, when treated with doxycycline, showed increased luciferase activity (Fig. 4f) as well as increased YB-1 expression (Supplementary Fig. ^{2d}) compared to mice in the same experiment that were transplanted with the same cells but not treated with doxycycline.

Normal human mammary cells expressing *myrAKT1* thus provide a new genetic model of *de novo* DCIS formation and the cells produced also show a moderate increase in YB-1 expression. In contrast, overexpression of YB-1 on its own is insufficient to induce evidence of transformation of primary human mammary cells.

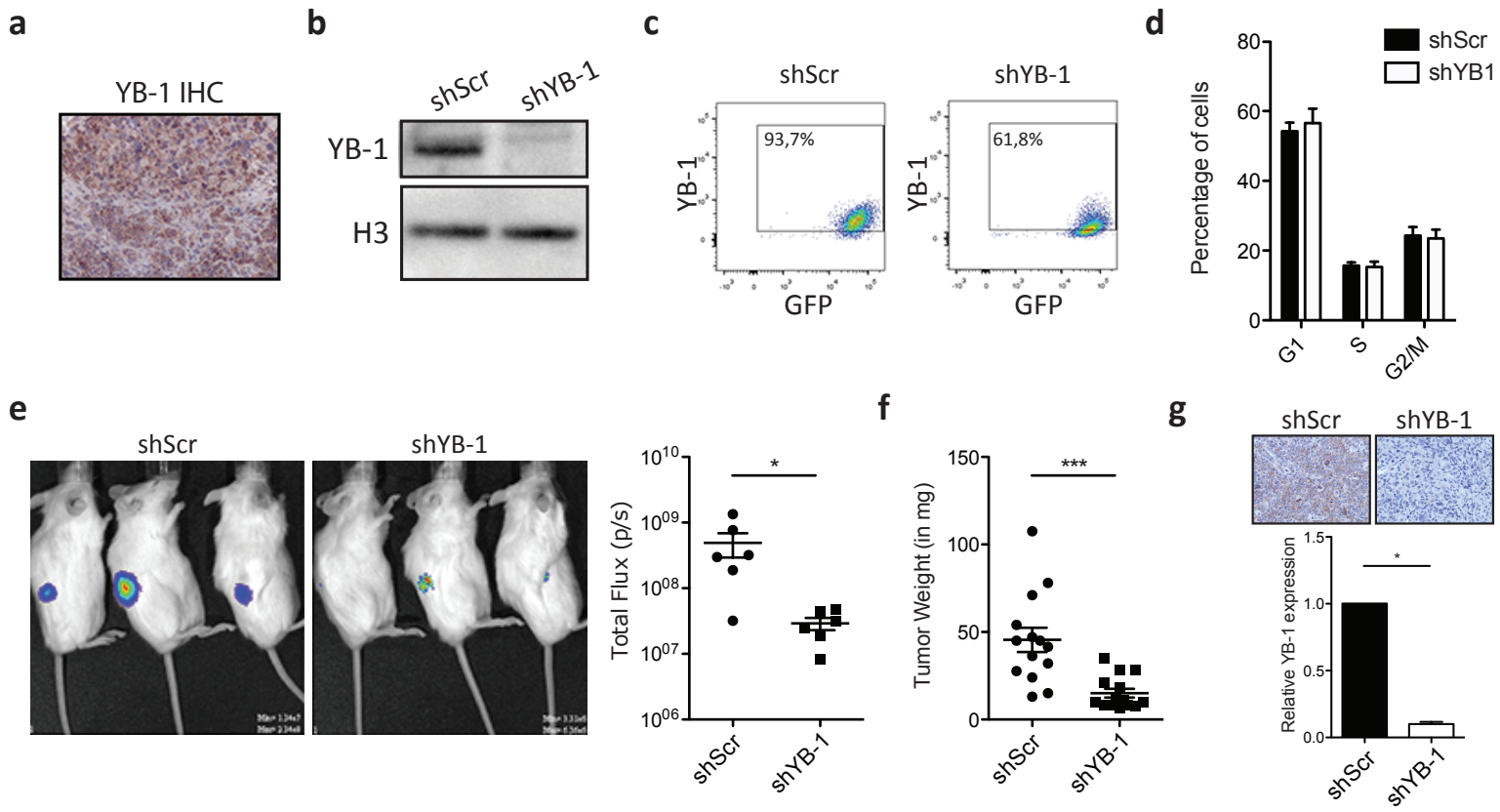
YB-1 expression is elevated in tumours produced from an established tumorigenic breast cancer cell line and is required for its metastatic ability

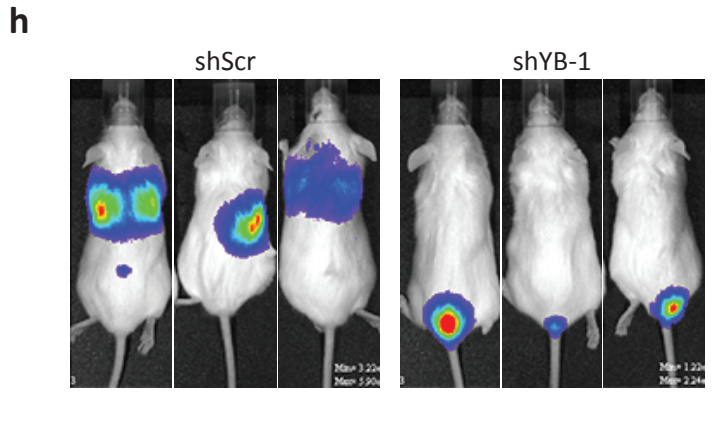
We next designed experiments to determine whether YB-1 expression might also play a critical role in more advanced breast cancers with deregulated KRAS activity. MDA-MB-231 cells was chosen as a model as it is a well-established human breast cancer cell line with a mutant

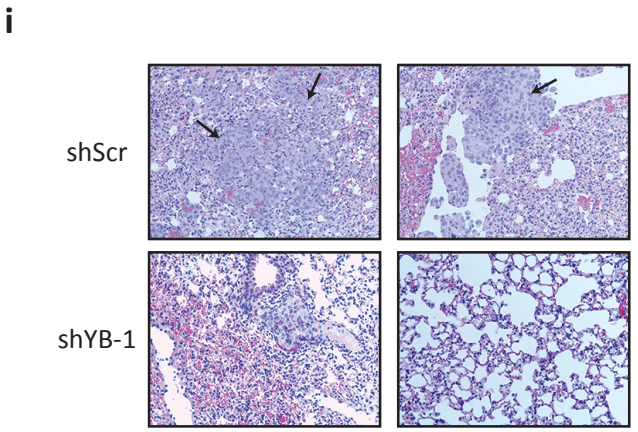
KRAS^{G13D} gene[15] that we confirmed here together with their previously reported high levels of YB-1[7] (Fig. 5a-i and Supplementary Fig. 3). WB analysis of MDA-MB-231 cells stably transduced with a shScrambled (shScr), or a shYB-1 vector, showed that YB-1 expression was specifically reduced by approximately 90% in the shYB-1-transduced MDA-MB-231 cells (Fig. 5b). Single-cell analysis showed that YB-1 expression in the shYB-1-transduced cells was highly variable, with some cells containing readily detectable levels of YB-1 (Fig. 5c). Interestingly, *in vitro* cell proliferation parameters were similar in control and YB-1 knockdown cells (Fig. 5d). In contrast, subcutaneous injection of low doses shYB-1-transduced MDA-MB-231 cells (1,000 cells/mouse) resulted in marked reductions of luciferase activity (Fig. 5e), tumour weights (Fig. 5f), and YB-1 levels (Fig. 5g) compared to the transplants of the control-transduced cells.

Since YB-1 is a known driver of sarcoma dissemination[5] and MDA-MB-231 cells have known metastatic activity, we also injected mice intravenously with stably shScr and shYB-1-transduced MDA-MB-231 cells (Fig. 5h-i and Supplementary Fig. 3a,b) or cells transfected with 2 different YB-1 targeting siRNAs (siYB-1#2 and siYB-1#6, Supplementary Fig. 3c-e). Control MDA-MB-231 cells metastasized as expected into different organs including the lung (Fig. 5h-i and Supplementary Fig. 3d,e). Strikingly, however, MDA-MB-231 cells with down-regulated YB-1 expression showed greatly reduced dissemination and the only cells detected were those retained at the base of the tail adjacent to the site of injection (Fig. 5h-i and Supplementary Fig. 3d,e). Interestingly, the few small lung metastasis found in mice injected with MDA-MB-231 cells transfected with siYB-1#2- displayed high YB-1 levels (Supplementary Fig. 3e), reflecting their possible "escape" from YB-1 inactivation (likely due either to a failed or inadequate suppression of YB-1), further underscoring the importance of YB-1 in contributing to the ability of these cells to disseminate.

Figure 5


a
YB-1 IHC
b
shScr
shYB-1
YB-1
H3
c
shScr
shYB-1
93,7%
61,8%
YB-1
GFP
d
shScr
shYB1
Percentage of cells
80
60
40
20
0
G1
S
G2/M
e
shScr
shYB-1
Total Flux (p/s)
10^{10}
10^{09}
10^{08}
10^{07}
10^{06}
*
f
Tumor Weight (in mg)
150
100
50
0
***
g
Relative YB-1 expression
1.5
1.0
0.5
0.0
*


h
shScr
shYB-1

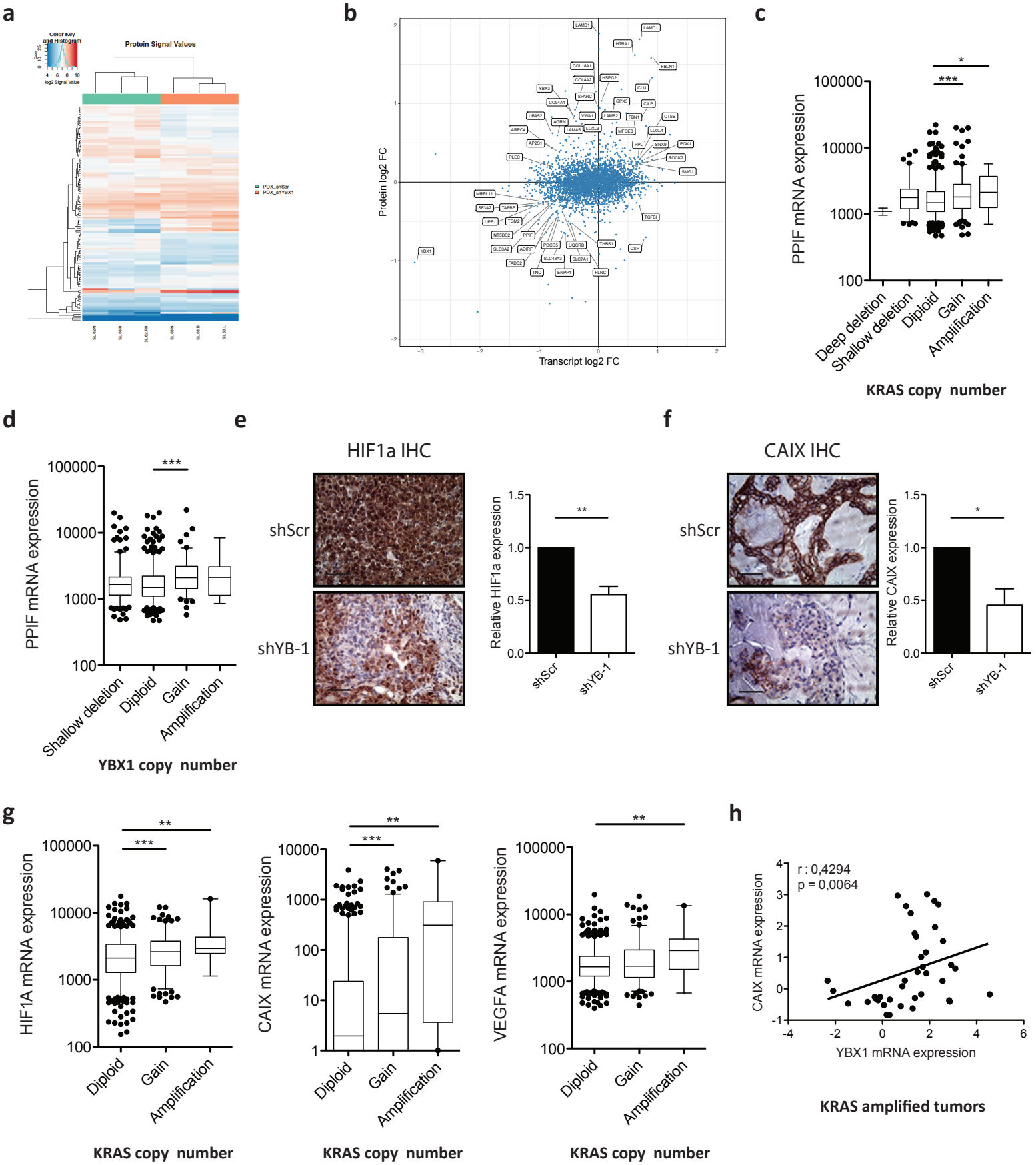

i
shScr
shYB-1

KRAS activation leads to an increased HIF1α response through YB-1

To decipher how $KRAS^{G12D}$ and YB-1 cooperate to promote tumour progression, we performed RNAseq and proteomic analyses of tumours generated from MDA-MB-231 cells stably expressing shYB-1 or a shScr- construct (Fig. 6a and Supplementary Fig. 4a,b). From these analyses, we identified several candidate target genes whose encoded RNAs and proteins increased or decreased in concert with YB-1 inhibition (Fig. 6b and Supplementary Fig. 4a,b). This revealed many proteins encoded by genes whose expression was similarly altered. Several of these (e.g., *PPIF*, *SLC3A2* and *SLC7A1*) are related to hypoxia[16–18] and also found to be increased in patients' breast cancers with gain of function mutations or amplification of *KRAS* as compared to tumours with diploid *KRAS* (Fig. 6c and Supplementary Fig. S4B), as well as in others with increased expression or amplified *YBX1* (Fig. 6d and Supplementary Fig. 4c). Notably, expression of PPIF, SLC3A2 and SLC7A1 was not further modulated in stably shYB-1-transduced MDA-MB-231 cells maintained in 2D cultures (Supplementary Fig. 4a,b), and these cells also showed no alterations in cell cycle progression (Fig. 5d).

These findings suggested a link between YB-1 expression and an altered hypoxia response in breast tumours harboring oncogenic *KRAS* mutations or deregulated *KRAS* signaling. Previously, YB-1 was found to be a primary regulator of HIF1α protein expression in sarcoma cells[5]. We therefore asked whether HIF1α expression and that of its direct transcriptional target, *CAIX* (carbonic anhydrase 9)[19], were also altered in tumours produced in mice transplanted with stably shScr- versus shYB-1-transduced normal mammary (Fig. 6e-f) or MDA-MB-231 cells (Supplementary Fig. 5a). Immunostaining of both HIF1α (Fig. 6e) and CAIX (Fig. 6f and Supplementary Fig. 5a) showed both were decreased in tumours generated from shYB-1-MDA-

Figure 6


a
Color Key and Histogram
Protein Signal Values
log2 Signal Value
PDX_shScr
PDX_shYBX1
b
Protein log2 FC
Transcript log2 FC
c
PPIF mRNA expression
Deep deletion
Shallow deletion
Diploid
Gain
Amplification
KRAS copy number
d
PPIF mRNA expression
Shallow deletion
Diploid
Gain
Amplification
YBX1 copy number
e
HIF1a IHC
shScr
shYB-1
Relative HIF1a expression
f
CAIX IHC
shScr
shYB-1
Relative CAIX expression
g
HIF1A mRNA expression
CAIX mRNA expression
VEGFA mRNA expression
Diploid
Gain
Amplification
KRAS copy number
h
r : 0,4294
p = 0,0064
CAIX mRNA expression
YBX1 mRNA expression
KRAS amplified tumors

MB-231 cells compared to those generated from control cells. In addition, MDA-MB-231-shYB-1 cells displayed reduced HIF1α compared to their shScr counterparts when cultured under hypoxic conditions (Supplementary Fig. 5b). Interestingly, levels of *HIF1A* and *CAIX* as well as *VEGFA*, a known HIF1α transcriptional target, were also increased in patients' breast cancers with gain of function or amplified *KRAS* compared to tumours with diploid *KRAS* (Fig. 6g and Supplementary Fig. 5c), and were likewise increased in tumours with gain of function or amplified *YBX1* (Supplementary Fig. 5d). Finally, patient s' breast cancers with amplified *KRAS* show a strong correlation between *YBX1* and *CAIX* mRNA levels (Fig. 6h). Together, these data point to an activated YB-1/HIF1α-mediated hypoxia response pathway in $KRAS^{G12D}$-driven breast cancers (Fig. 1d).

DISCUSSION

In this study, we show that elevated expression of YB-1 plays a requisite role in enabling primary human mammary cells to acquire a number of distinct malignant properties, both during their initial transformation and at later stages of advanced disease. Evidence for a key early role of YB-1 was revealed here in experiments with $KRAS^{G12D}$-transduced cells also carrying a shYB-1 construct. This impaired the ability of the $KRAS^{G12D}$-transduced cells to produce invasive ductal carcinomas previously shown to be otherwise consistently obtained within 6-8 weeks in immunodeficient mice[8]. We now also report the ability of another oncogene, *myrAKT1,* to produce an earlier DCIS-like stage of human breast cancer that is similarly YB-1-dependent. Finally, we show that the established MDA-MB-231 breast cancer cell line remains similarly dependent on an elevated expression of YB-1 protein to display their metastatic potential. The importance of these findings is underscored by analysis of published data for patients' breast

cancers showing elevated YB-1 expression in breast cancers that have both an amplified *KRAS* genotype and a poor prognosis. These tumours also show an activated HIF1α response in the presence of high YB-1 expression, in common with the tumours produced in our experimental models.

Previous experiments in which YB-1 was overexpressed in the immortalized but non-tumorigenic MCF10A cell line transformed with H-RAS showed that high YB-1 expression contributes to the disruption of mammary cell architecture and promotes an EMT[7]. However a survey of changes in published breast cancer datasets did not show any evidence of increased YB-1 expression in patient breast cancers with amplified HRAS or NRAS (data not shown). 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 to note that forced expression of $KRAS^{G12D}$ in primary human mammary cells did not modulate YB-1 levels in cells maintained in normoxic conditions *in vitro*, in contrast to the rapid increases in YB-1 levels observed *in vivo*. This difference may reflect the rapid creation of a hypoxic environment in nascent tumours forming *in vivo*, as shown by their high staining of HIF1α and CAIX, whereas the results of our *in vitro* experiments performed under hypoxic conditions suggest that YB-1 might modulate HIF1α in such an environment.

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* [20]. Thus, our results showing that YB-1 inhibition 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 in turn activates YB-1 expression, which in turn promotes an elevated HIF1α response *in vivo*. Further investigations will be required to determine whether YB-1 directly binds to the *HIF1A* 5'-UTR to enhance the acute synthesis of HIF1α as shown in sarcoma cells [5].

In summary, our findings reveal novel mechanistic features by which $KRAS^{G12D}$ activates YB-1 to both initiate and disseminate 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. Alternatively, targeting HIF1α itself or its downstream effectors may offer more tractable clinical targeting approaches. 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 AME 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 wrote the manuscript. S.L., A.E-N., S.T., and S.C. performed the experiments. S.C and G.L.N. performed 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[21]. 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 a well established protocol and reagents [9]. 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 [22]. 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[15].

Transduction and transfection

Primary cells were transduced with lentiviral vectors prepared and used as previously described[8]. For transient inhibition of YB-1, primary human mammary cells were transfected with siYB-1 (siRNA1, 5′-UGACACCAAGGAAGAUGUA-3′; siRNA 2, 5′-GUGAGAGUGGGGAAAAGAA-3′, from GE Healthcare), using RNAiMAX following the manufacturer's protocol (Thermofisher). For stable inhibition, shYB-1 (sc-38634-V, Santa Cruz) or shScr (sc-108080) lentiviral particles were used.

Xenografts

NRG mice were bred and housed in the animal facility at the British Columbia Cancer Research Centre. 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 subcutaneously with 50% (v/v) Matrigel into mice[8]. To measure tumour bioluminescence from expressed luciferase, 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 cell suspensions from 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).

Immunohistochemical (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). 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 $KRAS^{G12D}$-encoding vector was derived from pINDUCER21 backbone [23] by replacing the attR1-ORF-attR2 cassette with a Kras-2A-KO2 fragment.

Western blot 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), and anti-YB-1 (Cell Signaling Technology, 4202, 1/1,000). 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

RNAseq data from matched normal and *de novo* tumours were derived from Nguyen et al[8] and expressed as RPKM values (reads per kilobase per million mapped reads). P-values were calculated using a paired t-test. Copy number alterations and Z-score normalized RNAseq

expression values (V2 RSEM) were obtained from cBioPortal[24], from TCGA[12], METABRIC[11] and Mutational profiles of metastatic breast cancers[13] datasets. Paired-end reads were generated on an Illumina HiSeq2500 sequencer. Read sequences were aligned to the hg19 human reference using the BWA-SW algorithm[25] to generate binary alignment/map (BAM) files. Transcript counts were obtained with the summarizeOverlaps function from GenomicAlignments package[26]. Differential expression analysis was performed with DESeq2 package[27].

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[28,29] 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[29]. 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[29]. 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[30,31]. 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[32]. PECA uses Limma[33] 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[32].

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. Evdokimova, V. M. & Ovchinnikov, L. P. Translational regulation by Y-box transcription factor: involvement of the major mRNA-associated protein, p50. *Int. J. Biochem. Cell Biol.* **31,** 139–149 (1999).
2. Kohno, K., Izumi, H., Uchiumi, T., Ashizuka, M. & Kuwano, M. The pleiotropic functions of the Y-box-binding protein, YB-1. *Bioessays* **25,** 691–698 (2003).
3. 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).
4. 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).
5. El-Naggar, A. M. *et al.* Translational Activation of HIF1α by YB-1 Promotes Sarcoma Metastasis. *Cancer Cell* **27,** 682–697 (2015).
6. 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).
7. 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).
8. Nguyen, L. V. *et al.* Barcoding reveals complex clonal dynamics of de novo transformed human mammary cells. *Nature* **528,** 267–271 (2015).
9. 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).

10. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490,** 61–70 (2012).
11. Pereira, B. *et al.* The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun* **7,** 11479 (2016).
12. Ciriello, G. *et al.* Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* **163,** 506–519 (2015).
13. Lefebvre, C. *et al.* Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. *PLoS Med.* **13,** e1002201 (2016).
14. 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).
15. 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).
16. Kucharzewska, P., Christianson, H. C. & Belting, M. Global Profiling of Metabolic Adaptation to Hypoxic Stress in Human Glioblastoma Cells. *PLOS ONE* **10,** e0116740 (2015).
17. 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).
18. 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).

19. Wykoff, C. C. *et al.* Hypoxia-inducible expression of tumor-associated carbonic anhydrases. *Cancer Res.* **60,** 7075–7083 (2000).

20. 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).

21. 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).

22. 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).

23. 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).

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

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

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

27. 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).

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

29. 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).
30. 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).
31. 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).
32. 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).
33. 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 ***KRAS^{G12D}*-transformed primary isolates of human mammary cells and patients' *KRAS*-amplified tumours express high levels of YB-1.** (**a**) RNAseq data from normal human mammary cell subsets versus *de novo* tumours (BC or BC-derived tumours in blue; LP or LP-derived tumours in red). Values for *YBX1* are shown as RPKMs. N = 3 donors. P-values are from paired t-test. (**b**) Representative views of YB-1 immunostaining of normal human mammary tissue (left) and 8-week tumours derived from *KRAS^{G12D}*-transduced mammary cells isolated from the same normal donors (right). N = 3 donors. Scale bar, 50 µm. (**c**) Western blots showing YB-1 levels (relative to H3) in human normal BCs, LPs, LCs and SCs. N = 3 donors. Subsets were sorted according to their surface EPCAM and CD49f levels (top panel). (**d**) 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*). (**e-f**) *YBX1* mRNA levels compared to *KRAS* mRNA levels (**e**, left panel) or *TP53* copy number status (**e**, right panel) in invasive breast carcinoma samples in the TCGA dataset**,** and *YBX1* mRNA levels compared to *KRAS* mRNA levels in the METABRIC dataset (**f**). Values for *YBX1* are shown as RPKMs. (**g**) Representative table of *YBX1* alteration co-occurrence in metastatic breast cancer.

Fig. 2 **YB-1 inactivation impairs initial tumour formation by *KRAS^{G12D}*-transduced normal human mammary cells.** (**a**) Western blot (top) showing YB-1 expression in cells expanded *in vitro* from isolated BCs and LPs transduced with *siScr, siYB-1#2* or *siYB-1#6*. CFC frequencies (bottom) in BCs and LPs transduced with *siScr, siYB-1#2* or *siYB-1#6*. Values shown are expressed as a percentage of the number of CFCs detected in the paired *siScr*-transfected cells. N

= 3 donors. (**b**) Representative pictures of bioluminescence signals in mice injected subcutaneously with *KRAS*^{G12D} + shScr- or shYB-1-transduced primary cells 2 weeks earlier. Dot plot showing bioluminescence activity of tumours derived from BCs (blue) and LPs (red). (**c-d**) Representative images of H&E (**c**)- and YB-1 (**d**)-stained sections from different BC- or LP-derived tumours arising from *KRAS*^{G12D} + shScr- or shYB-1-transduced cells. Scale bar, 200 µm (**c**) or 100 µm (**d**). Bar graph (**d**) shows quantification of YB-1 intensity in tumours derived from *KRAS*^{G12D} + shScr or shYB-1 cells. N = 8.

Fig. 3 **YB-1 expression remains high after reversal of *KRAS* expression to basal levels.** (**a**) Western blots showing YB-1 and RAS levels (relative to H3) in control and *KRAS*^{G12D}-transduced human BCs and LPs assessed 3 days post-transduction. N = 3 donors. (**b**) Representative photomicrographs 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 and DAPI. (**c**) Representative images of YB-1 immunostaining of 2 week-old xenografts of *KRAS*^{G12D}-transduced primary cells. Scale bar, 200 µm (left) or 100 µm (right). (**d**) Representative FACS profile of human mammary MCF10A cells stably expressing an inducible KRAS-2A-KO2 construct after being maintained in the presence or absence of doxycycline. (**e**) Western blots showing RAS levels (relative to ACTIN) in the same cells as in (**d**). (**f**) 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. (**g**) *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. (**h**) 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 ***De novo* formation of DCIS-like tumours leads to intermediate levels of YB-1.** (**a**) Representative photos of bioluminescence signals measured in mice injected subcutaneously 7 weeks earlier with Luc-YFP alone or in combination with *KRAS*^{G12D}-, YB-1-, *KRAS*^{G12D}+YB-1-, *myrAKT1*-transduced human mammary cells. BCs and LPs from 3 donors were pooled before transduction. Graph plot shows changes in bioluminescence activity over time. (**b**) Representative FACS plots of human (CD298/EPCAM)$^{+}$ and mCherry (myrAKT1)$^{+}$ or YFP (KRAS^{G12D})$^{+}$ cells present in dissociated tumours generated from human mammary cells transduced with *KRAS*^{G12D} or *myrAKT1*. (**c**) Representative photos of bioluminescence signals in mice injected subcutaneously 5 or 7 weeks earlier with Luc-YFP and *myrAKT1*-transduced human mammary cells. Graph plot shows bioluminescence activity from tumours derived from BCs (blue) and LPs (red). N = 5 donors. (**d**) Representative images of SMA-, CK14-, CK8-18- and p63-stained sections of *myrAKT1*-derived tumours initiated from either BCs or LPs. Scale bar, 100 μm. (**e**) Representative views of YB-1 immunostaining of 18-week primary *myrAKT1*-derived tumours generated from normal mammary cells from 3 different donors (#1-3). Bar graph shows a comparison of YB-1 staining intensity in *KRAS*^{G12D}- or *myrAKT1*-derived tumours. N = 10 (*KRAS*^{G12D}) or 6 (*myrAKT1*) tumours. (**f**) Dot plot shows the bioluminescence measured in mice injected subcutaneously with *myrAKT1*+inducible YB-1-transduced human mammary cells and given water with or without doxycyline. N = 3 donors.

Fig. 5 **YB-1 inactivation impairs tumour formation and dissemination of advanced tumours.** (**a-c**) YB-1 expression by IHC (**a**) Western blot (**b**) or FACS analysis (**c**) of MDA-MB-231 cells. (**d**) Cell cycle analysis of MDA-MB-231 cells. (**e-g**) Tumorigenesis. (**e**)

Representative pictures of bioluminescence signals in mice injected subcutaneously with MDA-MB-231 shScr or shYB-1 cells. Dot plot shows the measured bioluminescence in these tumours 26 days post-transplant. (**f**) Weights of the tumours shown in (**e**). (**g**) Representative views of YB-1 immunostaining of tumours derived from MDA-MB-231 shScr or shYB-1 cells. Bar graph shows quantification of YB-1 intensity. (**h-i**) Effect of YB-1 suppression on tumour dissemination. (**h**) Representative pictures 45 days post-transplant of bioluminescence signals measured in mice injected intravenously with MDA-MB-231 shScr or shYB-1 cells. (**i**) Representative H&E-stained photomicrographs of MDA-MB-231 shScr- or shYB-1-derived tumours present in the lungs of intravenously injected mice.

Fig. 6 **KRAS amplification leads to an enhanced HIF1α response through YB-1.** (**a**) Hierarchical clustering of proteomics data obtained on tumour cells generated from shScr- or shYB-1-transduced MDA-MB-231 cells. (**b**) Correlation plot between transcripts and proteins identified by RNAseq and proteomic analysis of tumours derived from shScr- and shYB-1-transduced MDA-MB-231 cells. (**c-d**) *PPIF* mRNA expression according to *KRAS* (**c**) or *YBX1* (**d**) copy number status in samples of invasive breast carcinoma. Values for *PPIF* are shown as RPKMs. (**e-f**) Representative images of HIF1α (**e**)- and CAIX (**f**)-stained sections from different BC- or LP-derived tumours arising from $KRAS^{G12D}$ + shScr- or shYB-1-transduced cells. Scale bar, 100 μm. Bar graph shows quantification of HIF1α (**e**) and CAIX (**f**) staining intensity in primary shScr or shYB-1 tumours. N = 8. (**g**) *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. (**h**) Scatter plot of *YBX1* and *CAIX* mRNA expression in amplified-*KRAS* invasive breast carcinomas.