



- (51) **International Patent Classification:**  
*C12Q 1/68* (2006.01) *A61P 35/04* (2006.01)  
*G01N 33/574* (2006.01)
- (21) **International Application Number:**  
 PCT/US2015/040432
- (22) **International Filing Date:**  
 14 July 2015 (14.07.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
 62/024,639 15 July 2014 (15.07.2014) US
- (71) **Applicant:** SALK INSTITUTE FOR BIOLOGICAL STUDIES [US/US]; 10010 N. Torrey Pine Road, La Jolla, CA 92037-1002 (US).
- (72) **Inventors:** SHAW, Reuben, J.; 10010 N. Torrey Pines Road, La Jolla, CA 92037-1002 (US). GOODWIN, Jonathan, M.; 10010 N. Torrey Pines Road, La Jolla, CA 92037-1002 (US).
- (74) **Agent:** RYBAK, Sheree, Lynn; Klarquist Sparkman, LLP, One World Trade Center, Suite 1600, 121 SW Salmon Street, Portland, OR 97204 (US).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

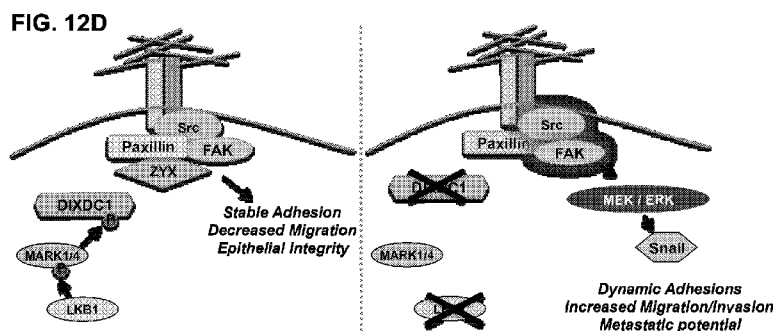
**Declarations under Rule 4.17:**

— of inventorship (Rule 4.17(iv))

**Published:**

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

- (54) **Title:** DETECTING DIXDC1 (DIX DOMAIN-CONTAINING PROTEIN 1) EXPRESSION TO DETERMINE IF A TUMOR WILL RESPOND TO FAK AND SRC KINASE INHIBITORS



- (57) **Abstract:** Provided herein are methods, systems, and kits for analyzing cancer samples (such as those in which FAK is upregulated or activated). Such methods can include detecting DIXDC1 expression and/or phosphorylation to identify cancers that will respond to FAK inhibitors and/or Src inhibitors, and to identify cancers that will likely metastasize.

**DETECTING DIXDC1 (DIX DOMAIN-CONTAINING PROTEIN 1) EXPRESSION TO  
DETERMINE IF A TUMOR WILL RESPOND TO  
FAK AND SRC KINASE INHIBITORS**

5                   **CROSS-REFERENCE TO RELATED APPLICATION**

This application claims priority to U.S. Provisional Application No. 62/024,639, filed July 15, 2014, herein incorporated by reference.

**FIELD**

10           Provided herein are methods, systems, and kits for analyzing cancer samples (such as those in which FAK is upregulated or activated). Such methods can include detecting DIXDC1 expression and/or phosphorylation to identify cancers that will respond to FAK inhibitors and/or Src inhibitors, and to identify cancers that will likely metastasize.

15                   **ACKNOWLEDGMENT OF GOVERNMENT SUPPORT**

This invention was made with government support under R01CA172229-01 and P01CA120964-06A1 awarded by The National Cancer Institute. The government has certain rights in the invention.

20                   **BACKGROUND**

A critical question in cancer biology is the relationship between tumor initiating mutations, including oncogenes and tumor suppressor genes, and the propensity for tumors to metastasize (Hanahan and Weinberg, *Cell* 144, 646-674, 2011). *LKB1/STK11* is the causal gene inactivated in the inherited cancer disorder Peutz-Jeghers Syndrome and is also inactivated in  
25 ~25% of non-small cell lung cancers, making it the fourth most frequent genetic alteration in human lung cancer (Ding et al., *Nature* 455, 1069-1075, 2008). Beyond effects on tumor initiation, loss of *Lkb1* uniquely confers invasive and metastatic behavior in multiple genetically engineered mouse models of cancer when directly compared to other tumor suppressors (e.g. *Trp53*, *Rb*, *p16/Cdkn2a*) (Carretero et al., *Cancer Cell* 17, 547-559, 2010; Contreras et al., *Dis*  
30 *Model Mech* 3, 181-193, 2010; Ji et al., *Nature* 448, 807-810, 2007; Liu et al., *Cancer Cell* 21, 751-764, 2012). The enhanced metastatic potential of *Lkb1*-deficient tumors is also notable for occurring in cancers across multiple lineages, namely endometrium, lung, and melanoma

(Contreras et al., Cancer Res. 68, 759–766, 2008; Ji et al., Nature 448, 807-810, 2007; Liu et al., Cancer Cell 21, 751-764, 2012).

LKB1 is a threonine-specific kinase that functions by directly phosphorylating the activation loop of a family of 14 kinases related to the AMP-activated Protein Kinase (AMPK) (Alessi et al., Annu Rev Biochem 75, 137-163, 2006). Through these 14 downstream kinases, LKB1 also plays highly conserved functions in organismal metabolism and cell polarity in addition to its tumor suppressor function (Jansen et al., Physiol Rev 89, 777-798, 2009). AMPK is a central conserved regulator of metabolism and cell growth, which explains some of the best understood functions of LKB1 seen in different tissues in mammals (Mihaylova and Shaw, Nat Cell Biol 13, 1016-1023, 2011). However, LKB1 is genetically well-established to be a highly conserved regulator of cell polarity as well, stemming from initial findings in *Drosophila* and *C. elegans*, where the LKB1 ortholog was identified as *Partitioning-defective 4 (par-4)* (Jansen et al., Physiol Rev 89, 777-798, 2009). In comparison to AMPK, far less is known about the biological functions and molecular targets of its 12 related kinases activated by LKB1, though one subfamily, the four Microtubule Affinity-Regulating Kinase (MARKs) (or *Partitioning-defective 1/Par1*) genes, control apical-basolateral polarity across eukaryotes and may be primary mediators of LKB1 action in that process (Ollila and Makela, J Mol Cell Biol 3, 330-340, 2011).

Over the past decade, screens and biochemical purifications have uncovered a number of downstream substrates of AMPK and its related kinases that provide molecular mechanisms underlying how LKB1 coordinates cell growth and metabolism (Hardie and Alessi, BMC biology 11, 36, 2013). AMPK regulates cell growth and survival via effects on the mTORC1 and autophagy pathways. AMPK and related LKB1-dependent kinases coordinate cellular and organismal metabolism, triggering acute effects via phosphorylation of metabolic enzymes and glucose transport regulators, and adaptive effects via phosphorylation of transcriptional regulators of metabolism (Mihaylova and Shaw, Nat Cell Biol 13, 1016-1023, 2011).

In spite of these significant advances illuminating how LKB1 loss may lead to metabolic derangement and growth of primary tumors, there are no well-established molecular connections between AMPK or any of its related kinases with effectors known to be involved in the control of metastasis. Moreover, it is not known whether the potent metastatic suppressive activity of LKB1 is due to its effects on AMPK or some of the other 12 related kinases it activates. We therefore set out to decode new targets of LKB1 and its downstream kinases that may play a specific role in limiting metastatic progression. The disclosure discusses the identification of a

single direct substrate of the AMPK-related kinases MARK1 and MARK4 that appears to phenocopy, and mediate, most of the known effects of LKB1 on EMT, migration, and metastatic behavior *in vivo*.

5

## SUMMARY

Based on the discovery that DIX domain-containing protein 1 (DIXDC1) expression and/or phosphorylation is correlated to focal adhesion kinase (FAK) expression, provided herein are methods for analyzing a cancer sample obtained from a subject, for example to determine if the cancer is one that is sensitive (*e.g.*, responsive) to treatment with a FAK inhibitor and/or a Src oncogene inhibitor, to determine if the cancer is one that is likely to metastasize, or both.

Provided herein are methods for analyzing a cancer sample obtained from a subject. In some examples, such methods include contacting the sample with a DIXDC1 protein-specific binding agent (*e.g.*, a DIXDC1-specific antibody or aptamer) or a DIXDC1 nucleic acid probe. The DIXDC1 protein-specific binding agent or a DIXDC1 nucleic acid probe are allowed to interact with the sample (or proteins or nucleic acids isolated from the sample) under conditions that permit binding of the DIXDC1 protein-specific binding agent to DIXDC1 proteins in the sample or that permit binding of the DIXDC1 nucleic acid probe to DIXDC1 nucleic acids (*e.g.*, genomic DNA, mRNA, and the like) in the sample (if DIXDC1 proteins/nucleic acids are present in the sample). After the specific binding agent or probe is allowed to incubate with the sample (or proteins or nucleic acids isolated from the sample), the DIXDC1 proteins or DIXDC1 nucleic acid molecules in the sample are detected, for example using immunohistochemistry, in situ hybridization, or other method available in the art. Such a determination can be qualitative or quantitative. Based on the DIXDC1 proteins (and/or their phosphorylation status, *e.g.*, at Ser 592) or DIXDC1 nucleic acid molecules detected in the sample, it is determined that the cancer is sensitive to a FAK inhibitor and/or a Src inhibitor when DIXDC1 protein expression and/or phosphorylation and/or DIXDC1 nucleic acid molecule expression is reduced in the cancer sample relative to a control or reference value (such as a value or range of values representing DIXDC1 expression expected in a corresponding normal tissue sample, *e.g.*, if cancer is breast cancer, the control can be normal breast tissue). In some examples, the method further includes detecting FAK activation in the sample (*e.g.*, indicated by FAK phosphorylation at Tyrosine 397), wherein increased activation of FAK in combination with decreased DIXDC1 expression indicates that the cancer is sensitive to a FAK inhibitor and/or a Src inhibitor.

In some examples, the method is a method of distinguishing between a subject who is likely to respond to treatment with a FAK inhibitor and/or Src inhibitor from a subject who is not likely to respond to treatment with a FAK inhibitor and/or Src inhibitor. Such methods can also include selecting a subject for treatment with a FAK inhibitor and/or Src inhibitor if the  
5 subject is identified as a subject who is likely to respond to treatment with a FAK inhibitor and/or Src inhibitor. In some examples, the method also includes administering a therapeutically effective amount of the FAK inhibitor and/or Src inhibitor to the subject identified as a subject who is likely to respond to treatment with the FAK inhibitor and/or Src inhibitor.

10 In some examples, the method is a method of predicting the likelihood that the cancer will metastasize. For example, if reduced DIXDC1 protein expression, DIXDC1 protein phosphorylation, or DIXDC1 nucleic acid molecule expression in the cancer sample is detected relative to the control sample, this indicates that the cancer is more likely to metastasize. For example, such a cancer may be at least 20%, at least 50%, at least 75%, or at least 90% more  
15 likely to metastasize within 3 months, within 6 months, within 9 months, or within 1 year.

Methods of treatment are also provided. In some examples such methods include analyzing a cancer sample obtained from a subject according to method as described herein, and administering a therapeutically effective amount of a FAK inhibitor and/or Src inhibitor to the subject identified as a subject who is likely to respond to treatment with the FAK inhibitor  
20 and/or Src inhibitor.

Any of the methods provided herein, such as one or more steps of the disclosed methods, can be performed by a suitably-programmed computer.

Also provided are computer-implemented methods. Such methods can include generating a DIXDC1 protein expression, DIXDC1 protein phosphorylation, and/or nucleic acid  
25 expression score based at least on measured DIXDC1 protein nucleic acid expression within a displayed image depicting a cancer sample detectably labeled with a DIXDC1 specific binding agent or probe, wherein the cancer sample is obtained from a subject. The method can also include outputting a DIXDC1 protein expression, DIXDC1 protein phosphorylation, and/or nucleic acid expression score for the sample. In some examples, the DIXDC1 protein  
30 expression, DIXDC1 protein phosphorylation, and/or nucleic acid expression score is based on the intensity of staining of the sample, such as a score of 0, 1, 2, or 3. In some examples, such a method can include outputting a prognosis for the subject. Such a prognosis can include an

indication as to whether or not the cancer is sensitive to a FAK inhibitor and/or a Src inhibitor, an indication as to the likelihood that the cancer will metastasize, or both.

The disclosure also provides one or more non-transitory computer-readable media comprising computer-executable instructions causing a computing system to perform the methods provided herein.

Also provided herein are systems for analyzing a cancer sample obtained from a subject. Such systems can include a means for measuring a level of DIXDC1 protein, DIXDC1, phosphorylation, or DIXDC1 nucleic acid molecule in the sample (such as a CCD camera), implemented rules for comparing the measured level of DIXDC1 protein or DIXDC1 nucleic acid molecule to a DIXDC1 reference value (such as an appropriate algorithm), and means for implementing the rules (such as an appropriate algorithm), whereby an indication of the likely risk of cancer metastasis and/or sensitivity of the cancer to a FAK inhibitor and/or a Src inhibitor is provided based on the measured level of DIXDC1 protein expression, DIXDC1 protein phosphorylation, and/or nucleic acid expression in the sample.

Kits useful with the disclosed methods are provided. Such kits can include a DIXDC1 specific-binding agent and/or a DIXDC1 nucleic acid probe, and one or more of a pair of primers specific for a DIXDC1 gene sequence; microscope slides; labeled secondary antibodies; and buffers for immunohistochemistry or in situ hybridization. In some examples the kits also include a FAK inhibitor and/or Src inhibitor.

The foregoing and other objects and features of the disclosure will become more apparent from the following detailed description, which proceeds with reference to the accompanying figures.

## BRIEF DESCRIPTION OF THE DRAWINGS

**FIGS. 1A-1I. LKB1 regulates the EMT transcription factor Snail1 in a manner dependent on MARK1 and MARK4, but independent of AMPK.**

FIG. 1A. RNAi knockdown of LKB1 (20  $\mu$ M) was performed across multiple human cancer cell lines of diverse tissue origin. DLD-1 cells are colorectal adenocarcinoma, U2OS and SJSA are osteosarcoma, and MiaPaCa2 and Panc-1 are pancreatic adenocarcinoma cell lines. Knockdown was performed for 72 hours and lysates were analyzed by western blotting with indicated antibodies.

FIG. 1B. H157 human non-small cell lung cancer cells (LKB1-null) were reconstituted with a WT LKB1 (WT) or the kinase-dead K78I mutant (KD). Expression of Snail1 and its target gene Wnt5a were analyzed by western blotting. (Vec) denotes control reconstitution.

FIG. 1C. RNAi knockdown of LKB1 (20  $\mu$ M) for 72 hours in H1299 human lung cancer cells (LKB1-wt).

FIG. 1D. Western blotting for Snail1 and its target gene Wnt5a in *Lkb1*<sup>+/+</sup> and *Lkb1*<sup>fl/fl</sup> MEFs following infection with Ad-Cre.

FIG. 1E. MEFs as described in FIG. 1D were transfected with siRNA oligos (50  $\mu$ M) targeting mouse *Snail1* (Sna). Expression of the Snail1 target Wnt5a was analyzed by western blotting.

FIG. 1F. Primary mouse lung tumors from *Kras*<sup>G12D/+</sup>; *Lkb1*<sup>+/+</sup> or *Kras*<sup>G12D/+</sup>; *Lkb1*<sup>fl/fl</sup> mice immunoblotted with indicated antibodies.

FIG. 1G. LKB1 is the master upstream activating kinase of the AMPK kinase family. siRNA oligos targeting AMPK subfamilies were transfected into U2OS at a final concentration of 20  $\mu$ M for 72 hours. Lysates were analyzed by western blotting with indicated antibodies.

FIG. 1H. RNAi knockdown of individual MARK isoforms (20  $\mu$ M) or AMPK $\alpha$  subunits (20  $\mu$ M final) for 72 hours and immunoblotted with indicated antibodies.

FIG. 1I. Pathway model. LKB1, functioning through the AMPK-related kinases MARK1 and MARK4, specifically represses the levels of the EMT transcription factor Snail1.

**FIGS. 2A-2E. LKB1 suppresses Snail1 mRNA and its subsequent downstream targets including Wnt5a, related to FIG. 1.**

FIG. 2A. LKB1 controls *SNAIL1* mRNA levels in H157 (LKB1-null) human lung cancer cells reconstituted with wild-type LKB1 (WT) or the kinase dead K78I LKB1 (KD) or empty vector (pBABE).  $p < .01$  relative to pBabe control.

FIG. 2B. LKB1 controls *Snail1* mRNA levels in *Lkb1*<sup>fl/fl</sup> MEF treated with Ad-Cre compared to WT. Gene expression was analyzed by RT-PCR.  $p < .01$  relative to *Lkb1*<sup>+/+</sup>.

FIG. 2C. Snail controls Wnt5a/b expression in U2OS. Pooled siRNA oligos (20  $\mu$ M) or two individual siRNA duplexes (20  $\mu$ M each) were transfected into U2OS cells for 72 hours. Lysates were immunoblotted for Wnt5a/b.

FIG. 2D. Validation of Wnt5a/b antibody. Stable overexpression of Snail1 drives expression of its target genes Wnt5a/b, and immunoblot signal for Wnt5a/b is abolished by siRNA transfection of Wnt5a and Wnt5b oligos (5ab) (20  $\mu$ M each). Retro = retroviral transduction.

FIG. 2E. Deconvolution of LKB1 and MARK1 SmartPOOL oligos. Pooled oligos (SP) or individual pool component oligos (20  $\mu$ M final) were transfected into U2OS cells for 72 hours. Lysates were immunoblotted for Snail.

**FIGS. 3A-3I. A bioinformatic and proteomic screen identifies DIXDC1 and other substrates of MARK family kinases, Related to Figure 4.**

FIG. 3A. Peptide phosphorylation selectivity for MARK family members. Sequences of peptide mixtures used to analyze kinase substrate specificity (SEQ ID NOS: 1 to 9, from top to bottom). Each peptide contained one residue (indicated by Z, representing one of the 20 unmodified amino acids, phosphor-Thr or phosphor-Tyr) fixed at one of nine positions relative to the centrally fixed phospho-acceptor (an equal mix of Ser and Thr). Remaining positions (labeled X) were equimolar mixtures of the 17 amino acids excluding Cys, Ser and Thr. Peptide libraries were subjected to a radiolabel kinase assay with the indicated kinase. Radiolabel incorporation was quantified by phosphor imager, and the data were normalized so that the average value at a given position was equal to 1. Heat maps show log transformed data (average value in each position is zero). MARK family members are highly similar in preferred consensus phosphorylation motif.

FIG. 3B. Workflow of screen to identify conserved MARK kinase substrates starting with bioinformatics analysis using optimal consensus phosphorylation motifs (SEQ ID NOS: 10 to 12, from top to bottom). Scansite and Phosphosite databases were mined for proteins containing a preferred sequence. Site conservation across evolution used to increase likelihood of functional phosphorylation sites. This screening approach identified numerous known MARK substrates as well as many novel candidates, including a group with Pfam domain annotation or GO term annotation as playing a role in cytoskeleton. Functionally screening of 25 candidates revealed 8 potential *in vivo* MARK substrates (SEQ ID NOS: 13, 13, 14, 15, 16, 17, 18, and 19, from top to bottom). Sequence alignments and domain structures of the putative MARK substrates illustrating the location of the consensus phosphorylation motif are displayed.

FIG. 3C. GST-14-3-3 binding assay. MARK1 WT or T215A (KD) was co-transfected with FLAG-IRSp53, a known MARK substrate that binds 14-3-3, in HEK 293T cells. Cleared lysates were subjected to pull down with 5  $\mu$ g of recombinant GST or GST-14-3-3 for 2 hr. GST was then captured using GSH-sepharose beads for 1 hr. Immunoprecipitates were immunoblotted for FLAG-IRSp53.

FIG. 3D. GST-14-3-3 binding assay as described in (C) with co-transfection of FLAG-DIXDC1.

FIG. 3E. Deconvolution of DIXDC1 SmartPOOL siRNA oligos. Pooled oligos (SP) and individual pool component oligos (20  $\mu$ M final) were transfected into U2OS cells for 72 hr. Lysates were immunoblotted for Snail protein induction.

FIG. 3F. GST-14-3-3 binding assay described in (C) examining co-transfection of  
5 MARK1 WT with either FLAG-DIXDC1-WT or FLAG-DIXDC1-S592A (S-A).

FIG. 3G. *In vitro*-IP kinase assay using recombinant Active MARK4 as described in FIG. 4D.

FIG. 3H. Transient transfection of myc-tagged WT DIXDC1 cDNA along with “Active LKB1” (pCDNA3-FLAG- WT LKB1 and pCDNA3-STRAD $\alpha$  cDNA constructs) in HEK293T  
10 cells. DIXDC1 was immunoprecipitated and immunoblotted with P-DIXDC1 S592 specific antibody.

FIG. 3I. siRNA knockdown of DIXDC1 (20  $\mu$ M) and SNAIL (50  $\mu$ M) in U2OS cells for 72 hours. Lysates were immunoblotted with indicated antibodies.

**FIGS. 4A-4H. DIXDC1 is a novel substrate of MARK1 that suppresses Snail.**

FIG. 4A. Optimal phosphorylation consensus motifs were determined for all of the  
15 MARK kinases and compared to that of AMPK (Gwinn et al., Mol Cell 30, 214-226, 2008). The consensus motif for MARK1 is shown. *In vitro* kinase assays were performed using an arrayed set of peptide mixtures in which the indicated residue was fixed at the indicated position relative to a central phosphorylation site residue. The level of phosphorylation of each peptide mixture  
20 was quantified and normalized to an average value of 1 at each position to generate the displayed heat maps.

FIG. 4B. siRNA knockdown (20  $\mu$ M) of known and potential MARK substrates implicated in cytoskeletal signaling that were identified in the screen. Lysates were immunoblotted for induction of Snail compared to siRNA knockdown of LKB1 and MARK1  
25 (20  $\mu$ M).

FIG. 4C. DIXDC1 domain structure and conservation of identified S592 residue across vertebrates. SEQ ID NOS: 16, 20, 21 and 22, from top to bottom.

FIG. 4D. *In vitro*-IP kinase assay using recombinant Active MARK1. FLAG-tagged WT or S592A DIXDC1 was expressed and immunoprecipitated from HEK293T cells. Isolated  
30 DIXDC1 was then subjected to an *in vitro* kinase reaction with human MARK1. Reactions were immunoblotted with a phospho-specific antibody raised against the S592 residue. Vec = empty vector.

FIG. 4E. MARK1 WT or KD (kinase-dead) cDNA constructs were expressed in HEK293T cells. DIXDC1 was immunoprecipitated and immunoblotted with P-DIXDC1 S592 antibody.

FIG. 4F. siRNA knockdown for indicated genes (20  $\mu$ M final) was performed in U2OS cells for 72 hrs. Lysates were immunoblotted with indicated antibodies. Key: Control RNAi (*Con*), AMPK $\alpha$ 1 and AMPK $\alpha$ 2 (*A12*), MARK1 (*M1*), MARK2 (*M2*), MARK3 (*M3*), MARK4 (*M4*), MARK1 and MARK4 (*M14*), MARK2 and MARK3 (*M23*).

FIG. 4G. siRNA knockdown for indicated genes (20  $\mu$ M) in U2OS cells for 72 hr. Lysates were immunoblotted with indicated antibodies.

FIG. 4H. Pathway model. DIXDC1 is a novel LKB1-dependent substrate of MARK1 and MARK4 that represses Snail1 through an unknown mechanism.

**FIGS. 5A-5F. DIXDC1 localizes to focal adhesions and regulates their maturation dependent on Ser592 phosphorylation.**

FIG. 5A. Immunolocalization of paxillin and stably expressed FLAG-DIXDC1 by lentiviral transduction in U2OS cells depleted for endogenous DIXDC1 by lentiviral shRNA transduction. Serum-starved cells were placed in suspension for 1 hr in serum-free media followed by plating onto collagen I coated coverslips in serum-free media for 1 hr to stimulate and synchronize focal adhesion formation. Scale bar = 10  $\mu$ m (Left image). Scale bar = 2  $\mu$ m (Zoom image).

FIG. 5B. Immunolocalization of paxillin in U2OS cells stably depleted for DIXDC1 by lentiviral shRNA transduction and plated as described in (A). Representative image shown and total focal adhesion number/cell is quantified at right (n = 50 cells/condition). Scale bar = 10  $\mu$ m. Scale bar = 2  $\mu$ m (Zoom image). \*\* p < .0001 compared to pLKO control. Statistical analysis performed using an unpaired Student's t-test.

FIG. 5C. Colocalization of paxillin and zyxin in U2OS cells stably depleted for DIXDC1 by lentiviral shRNA transduction plated as described in (A). Scale bar = 10  $\mu$ m.

FIG. 5D. Magnified images of boxed regions in (C) showing individual spectra for paxillin and zyxin signal at peripheral adhesions. Scale bar = 2  $\mu$ m.

FIG. 5E. siRNA depletion of indicated genes (20  $\mu$ M) in U2OS cells followed by paxillin and zyxin colocalization. Representative images are shown. Scale bar = 10  $\mu$ m (Left images). Scale bar = 1  $\mu$ m (Zoom image).

FIG. 5F. FLAG-DIXDC1 WT and S592A were transiently expressed by low level lentiviral transduction in U2OS cells. 24 hr post infection, cells were plated as described in (A). Focal adhesions were marked by immunostaining for paxillin. Scale bar = 10  $\mu$ m.

**FIG. 5G. DIXDC1 regulates focal adhesion maturation.** Pearson coefficient of colocalization of Paxillin and Zyxin from cells displayed in Fig. 3C. n = 50 sets of peripheral adhesions as drawn in FIG. 5D.

**FIGS. 6A-6G. Hyperactivation of FAK upon loss of LKB1 or DIXDC1 is responsible for Snail induction.**

FIG. 6A. Immunolocalization of P-FAK Y397 and paxillin in cells transfected with siRNA oligos (20  $\mu$ M) against DIXDC1 or LKB1 and plated as described in Figure 3A. Representative images are shown. Scale bar = 10  $\mu$ m. Scale bar = 2  $\mu$ m (Zoom image).

FIG. 6B. Analysis of focal adhesion signaling upon plating on ECM. U2OS cells depleted of DIXDC1 by siRNA transfection (20  $\mu$ M) were treated as described in FIG. 5A and plated for the indicated timepoint. Lysates immunoblotted with indicated antibodies. Sus = Suspension 1 hr.

FIG. 6C. U2OS cells were infected with lentiviral vectors expressing two separate shRNA duplexes to FAK1 for 24 hr. Immediately following infection, U2OS cells were depleted of DIXDC1 by siRNA transfection (20  $\mu$ M) for 72 h. At 48 hours post siRNA transfection, cells were serum-starved for 24 hours before collection of cell lysates. GFP = shGFP control. Scr = Scramble control.

FIG. 6D. U2OS cells were transfected with siRNA against DIXDC1 (20  $\mu$ M) for 72 hr. 48 hr post RNAi transfection, cells were treated with 1  $\mu$ M PF-573228 FAK inhibitor or DMSO vehicle control for 24 hr in the absence of serum. Lysates were analyzed by immunoblot with the indicated antibodies.

FIG. 6E. siRNA (20  $\mu$ M) knockdown of the indicated genes in U2OS cells for 72 hr. 48 hr post RNAi transfection, cells were treated with 1  $\mu$ M PF-573228 FAK inhibitor or DMSO vehicle control for 24 hr in the absence of serum. Lysates were analyzed by immunoblot with the indicated antibodies.

FIG. 6F. U2OS cells were transfected with siRNA against LKB1 or DIXDC1 (20  $\mu$ M) for 72 hr. 48 hr post RNAi transfection, cells were treated with 20  $\mu$ M U0126 (MEK inhibitor) or DMSO vehicle control for 24 hr in the absence of serum. Lysates were analyzed by immunoblot with the indicated antibodies.

FIG. 6G. Model of proposed LKB1-MARK1-DIXDC1 signaling cascade.

**FIGS. 7A-7D Hyperactivation of FAK upon loss of LKB1 or DIXDC1 is responsible for Snail induction.**

FIG. 7A. U2OS cells were transfected with RNAi oligos (20  $\mu$ M) for 72 hr. Cells were serum-starved overnight and plated as in FIG. 3A. Cells were pretreated for 3 hr with 1  $\mu$ M PF-573228 prior to being placed in suspension. Upon plating onto collagen I coated plates, 1  $\mu$ M PF-573228 was added for the 1 hr duration. Lysates were immunoblotted with the indicated antibodies. Sus = Suspension 1 hr.

FIG. 7B. U2OS cells were transfected with siRNA against DIXDC1 (20  $\mu$ M) for 48 hr. Cells were then serum-starved for 24 hours prior to cell lysis. Endogenous GRB2 was immunoprecipitated using a GRB2 antibody and immunoblotted for FAK and Src binding.

FIG. 7C. U2OS cells were transfected with siRNA oligos (20  $\mu$ M) for 48 hours. Cells were then serum-starved and treated with indicated small molecule inhibitors for 24 hours to distinguish role of MAPK/ERK pathway from P13K/Akt pathway downstream of FAK. U0126 = MEK inhibitor 20  $\mu$ M. Akt Inh V = Akt1/2/3 inhibitor 20  $\mu$ M.

FIG. 7D. U2OS cells were treated as described in (C) with a panel of small molecule inhibitors against canonical downstream effectors of FAK signaling. U0126 = MEK inhibitor 20  $\mu$ M. Wort. = wortmannin 2 $\mu$ M. Torin1 = TOR inhibitor 250nM.

**FIGS. 8A-8E. DIXDC1 suppresses cell migration and invasion in a manner dependent on S592 phosphorylation.**

FIG. 8A. Scratch wound healing assay of U2OS cells depleted of indicated genes by siRNA transfection (20  $\mu$ M) for 72 h. Migration media supplemented with 10 ug/ml Mitomycin C to block cell proliferation. Cells were stained 14 hr post-wounding with Rhodamine-conjugated Phalloidin to demarcate cell borders. Images representative of 3 independent experiments.

FIG. 8B. Stable expression of indicated DIXDC1 cDNA by lentiviral transduction in U2OS cells. Scratch wound healing assays performed as in FIG. 8A. Images from 2 independent experiments shown and representative of 3 experiments.

FIG. 8C. siRNA transfection (20  $\mu$ M) against indicated genes in U2OS cells for 72 h. Scratch wound healing assays performed as in FIG. 8A. Media was supplemented with DMSO vehicle control or 1  $\mu$ M PF-573228 at the time of wounding and incubated for the 14 hr migration period.

FIG. 8D. siRNA knockdown of DIXDC1 (20  $\mu$ M) and Snail1 (50  $\mu$ M) in U2OS cells for 72 h. Scratch wound healing assays performed as in FIG. 5A.

FIG. 8E. Collagen transwell migration assay. Transwell filters were coated with collagen I and U2OS cells depleted for the indicated genes by siRNA transfection were plated onto filters in serum-free media supplemented with 10 ug/ml Mitomycin C for 24 h. Filters were excised and migrated cells were stained with DAPI and counted. \*\*\* =  $p < .0001$  relative to control. Statistical analysis performed using an unpaired Student's t-test.

**FIG. 9. FAK inhibitor blocks increased invasion from DIXDC1 knockdown, Related to Figure 8.** Matrigel transwell invasion assay. siRNA knockdown of DIXDC1 (20  $\mu$ M) was performed in U2OS cells for 72 hr. Cells were then plated into the upper chamber of a transwell migration apparatus in serum-free media supplemented with either DMSO vehicle control or 1  $\mu$ M PF-573228, in addition to 10 ug/ml Mitomycin C. Cell invasion was quantified after 24 hr by crystal violet dye elution. \*\*\*  $p < .0001$  unpaired Student's t-test.

**FIGS. 10A-10G. DIXDC1 suppresses lung colonization of primary mouse lung cancer cell lines in a manner dependent on S592 phosphorylation.**

FIG. 10A. Lysates from non-metastatic ( $T_{\text{nonMet}}$ ) 368T1 and metastatic ( $T_{\text{Met}}$ ) 393T5 *Kras*<sup>G12D</sup>; *p53*<sup>-/-</sup> mouse primary lung tumor cell lines immunoblotted with the indicated antibodies.

FIG. 10B. Stable depletion of Dixdc1 by lentiviral shRNA transduction with two independent shRNA sequences in  $T_{\text{nonMet}}$  368T1 cells. Lysates were immunoblotted with the indicated antibodies.

FIG. 10C. Lung colonization assay.  $T_{\text{nonMet}}$  368T1 cells were stably depleted of endogenous Dixdc1 by lentiviral transduction with two independent shRNA sequences. Cells were then injected intravenously in the lateral tail vein of syngeneic 129/B16 F<sub>1</sub> mice. Lungs were harvested 3 weeks post transplantation and representative hematoxylin & eosin (H&E) - stained sections are shown. Total lung tumor burden was quantified using morphometric Inform software. Scale bar (top image) = 1mm. Scale bar (zoom image) = 100  $\mu$ m. Statistical analysis performed using an unpaired Student's t-test.

FIG. 10D. Rescue of enhanced colonization efficiency by enforced expression of WT, but not S592A mutant hDIXDC1 in  $T_{\text{nonMet}}$  cells. Cells were stably depleted for endogenous Dixdc1 or GFP control by lentiviral shRNA transduction. Knockdown cells were then reconstituted with a WT or S592A hDIXDC1 cDNA using retroviral transduction. Stably selected cells were injected intravenously and colonization efficiency was assessed by H&E staining 3 weeks post transplantation. Total lung tumor burden was quantified using

morphometric Inform software. Scale bar (top image) = 1mm. Scale bar (zoom image) = 100  $\mu$ m. Statistical analysis performed using an unpaired Student's t-test.

FIG. 10E. Suppression of lung colonization by highly metastatic T<sub>Met</sub> 393T5 cells through enforced expression of WT, but not S592A mutant hDIXDC1. T<sub>Met</sub> 393T5 cells, which express low levels of Dixdc1 (FIG. 10A), were transduced with a retroviral vector encoding the cDNA for WT or S592A mutant hDIXDC1. Cells were transduced with a lentiviral vector expressing firefly luciferase and subsequently injected intravenously in the lateral tail vein of syngeneic 129/B16 F<sub>1</sub> mice. Representative H&E sections of mouse lungs at 3 weeks post-transplantation are shown. Scale = 1mm.

FIG. 10F. Quantification of total bioluminescence at 3 weeks (E) compared to 30 min post tumor cell intravenous transplantation. \*\* p < .001 unpaired Student's t-test. N.S. = not significant.

FIG. 10G. At 3 weeks post-transplantation, mice from (E) (n = 5/group) were imaged using IVIS bioluminescence imaging. Signal intensity representing lung tumor burden is shown.

All data are represented as the mean  $\pm$  SEM.

**FIGS. 11A-11H. DIXDC1 controls FAK signaling, Snail, and invasion in T<sub>nonmet</sub> cells similar to U2OS cells.**

FIG. 11A. RT-PCR analysis of a panel of T<sub>nonMet</sub> and T<sub>Met</sub> cell lines for *Dixdc1* mRNA levels. \* p < .01 relative to 802T4.

FIG. 11B. siRNA knockdown (20  $\mu$ M) of indicated genes in T<sub>nonMet</sub> 368T1 cells by siRNA transfection for 72 h. Lysates were immunoblotted with P-Dixdc1 S592 specific antibody. M1 = Mark1. M4 = Mark4.

FIG. 11C. Analysis of proximal focal adhesion signaling in T<sub>nonMet</sub> 368T1 cells. Cells were transfected with siRNA (20  $\mu$ M) against the indicated genes for 72 hr. Serum-starved cells were then placed in suspension for 1 hr in serum-free media, followed by plating onto Fibronectin coated cell culture dishes in the absence of serum for 2 hr to allow for adhesion. Cells were then lysed and immunoblotted with the indicated antibodies. Susp = Suspension 1 hr. FN = Fibronectin.

FIG. 11D. Stable expression of WT-hDIXDC1 in T<sub>Met</sub> cells was achieved by retroviral transduction. Lysates were then immunoblotted with the indicated antibodies.

FIG. 11E. Matrigel transwell invasion assay. T<sub>nonMet</sub> cells stably depleted for Dixdc1 or GFP control used in FIG. 10B were plated in the absence of serum in the upper chamber of a matrigel coated transwell apparatus. Cell invasion was quantified 24 hr post plating using crystal

violet dye elution. \*\*\*  $p < .0001$  relative to shGFP. Statistical analysis performed using an unpaired Student's t-test.

FIG. 11F. MTT growth assay comparing proliferation rates of T<sub>nonMet</sub> shGFP and shDixdc1 cells from FIG. 10B.  $p > 0.9$  using unpaired Student's t-test.

FIG. 11G. Lung colonization assay comparing stable shRNA depletion of Lkb1 to Dixdc1. T<sub>nonMet</sub> 368T1 cells infected with indicated shRNA lentiviral vectors were injected intravenously and lung colonization ability was assessed at 3 weeks post transplantation by H&E staining. Scale = 1mm.

FIG. 11H. Expression levels of DIXDC1 in retrovirally transduced T<sub>Met</sub> 393T5 cells used in FIG. 10F.

**FIGS. 12A-12D. *DIXDC1* is frequently downregulated in human cancer.**

FIG. 12A. Copy number alteration data obtained using the Broad TumorScape platform. Deletion frequency independent of peak regions of *DIXDC1* is shown across tumor types. All results shown reach significance with a Q-value of less than 0.25. Red sidebar denotes tumor subsets where LKB1 mutation/loss-of-function has been shown to promote invasion and metastasis.

FIG. 12B. Expression of *DIXDC1* mRNA in human lung cancer datasets, compared to normal lung. Three independent published datasets were analyzed for *DIXDC1* expression levels. Data is represented in Box-Whisker plots and mRNA expression values are presented on a log<sub>2</sub> scale. Lung cancer datasets were accessed using the Lung Cancer Explorer hosted by UTSW.

FIG. 12C. Correlation of *DIXDC1* expression with patient survival in NSCLC. *DIXDC1* expression was stratified as high vs. low against median expression. Overall survival (OS), time to first progression (FP), and correlation within previously published survival datasets (PPS) were analyzed. Analysis was performed using the kmPlotter ([www.kmplot.com](http://www.kmplot.com)).

FIG. 12D. Model. LKB1, functioning through the downstream kinases MARK1 and MARK4, positively regulates the focal adhesion protein DIXDC1 to promote adhesion maturation/stabilization and suppress cell migration, invasion, and metastatic potential. Upon loss of the LKB1-dependent phosphorylation, or upon downregulation or deletion of the *DIXDC1* gene as occurs in human cancer, focal adhesions become more dynamic and resident kinases FAK/Src activate a signaling cascade through ERK1/2 to induce the EMT transcription factor Snail. Expression of Snail drives genes associated with invasion and migration, notably

members of the noncanonical Wnt family, Wnt5a/b. This identifies a mechanism behind the increased metastatic potential attributed to LKB1-deficient tumors.

**FIGS. 13A-13D. *DIXDC1* is frequently downregulated in human cancer.**

FIG. 13A. Expression of *DIXDC1* mRNA in human lung cancer datasets, compared to normal lung. Two additional published datasets were analyzed for *DIXDC1* expression levels. Data is represented in Box-Whisker plots and mRNA expression values are presented on a log<sub>2</sub> scale. Lung cancer datasets were accessed using the Lung Cancer Explorer hosted by UTSW.

FIG. 13B. Expression levels of *DIXDC1* mRNA across the published dataset GSE19188. Individual log<sub>2</sub> transformed values are expressed showing the high degree of variability of *DIXDC1* mRNA in human lung cancer.

FIG. 13C. Individual representation of raw *DIXDC1* expression values from the published human lung cancer dataset GSE7670. \*\*\*  $p < .0001$  relative to normal lung.

FIG. 13D. *Dixdc1* was identified as one of only 97 genes across the genome that were synergistically regulated by Kras<sup>G12D</sup> mutation and loss of function of p53 (McMurray et al., Nature 453, 1112-1116, 2008) (GSE9199). Stable YAMC cell lines were made with the defined oncogenic perturbations followed by microarray analysis. Raw expression values are displayed for *Dixdc1* across the cell populations. \*  $p < .05$ , \*\*\*  $p < .001$ .

FIG. 14 is a block diagram of an example computing system in which some described methods may be implemented.

## SEQUENCE LISTING

The nucleic and amino acid sequences are shown using standard letter abbreviations for nucleotide bases, and three letter code for amino acids, as defined in 37 C.F.R. 1.822. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand.

SEQ ID NOS: 1-9 are sequences of peptide mixtures used to analyze kinase substrate specificity.

SEQ ID NOS: 10 to 12 show an optimal MARK motif, secondary selections sequence, and additional selections sequences, respectively.

SEQ ID NOS: 13 to 19 show potential *in vivo* MARK substrates.

SEQ ID NOS: 20 to 22 show portions of the *DIXDC1* protein sequence demonstrating conservation of S592 residue across vertebrates.

SEQ ID NOS: 23-24 are qPCR primer sequences used to detect human *SNAI1*.

SEQ ID NOS: 25-26 are qPCR primer sequences used to detect human ACTB.

SEQ ID NOS: 27-28 are qPCR primer sequences used to detect human WNT5A.

SEQ ID NOS: 29-30 are qPCR primer sequences used to detect mouse Wnt5A.

SEQ ID NOS: 31-32 are qPCR primer sequences used to detect mouse Actb.

5 SEQ ID NOS: 33-34 are qPCR primer sequences used to detect mouse Wnt5B.

SEQ ID NOS: 35-36 are qPCR primer sequences used to detect mouse Snail.

SEQ ID NO: 37 is an shRNA for hDIXDC1 3'utr hairpin.

### DETAILED DESCRIPTION

10 The following explanations of terms and methods are provided to better describe the present disclosure and to guide those of ordinary skill in the art in the practice of the present disclosure. The singular forms “a,” “an,” and “the” refer to one or more than one, unless the context clearly dictates otherwise. For example, the term “comprising a cell” includes single or plural cells and is considered equivalent to the phrase “comprising at least one cell.” The term  
15 “or” refers to a single element of stated alternative elements or a combination of two or more elements, unless the context clearly indicates otherwise. As used herein, “comprises” means “includes.” Thus, “comprising A or B,” means “including A, B, or A and B,” without excluding additional elements. Dates of GenBank® Accession Nos. referred to herein are the sequences available at least as early as July 15, 2014. All references, patents and patent applications, and  
20 GenBank® Accession numbers cited herein are incorporated by reference.

Unless explained otherwise, all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this disclosure belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, suitable methods and materials are  
25 described below. The materials, methods, and examples are illustrative only and not intended to be limiting.

In order to facilitate review of the various embodiments of the disclosure, the following explanations of specific terms are provided:

30 **Administration:** To provide or give a subject an agent, such as a FAK inhibitor or Src inhibitor, by any effective route. Exemplary routes of administration include, but are not limited to, oral, injection (such as subcutaneous, intramuscular, intradermal, intraperitoneal,

intravenous, and intratumoral), sublingual, rectal, transdermal, intranasal, vaginal, and inhalation routes.

**Antibody:** Immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, that is, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen (such as a DIXDC1 protein, such as a human DIXDC1 protein). Exemplary antibodies include monoclonal, polyclonal, camelid, and humanized antibodies, such as those that are specific for DIXDC1, such as those that are specific for phosphorylated DIXDC1. In some examples, antibodies can be diagnostic, for example used to detect the presence of a protein such as DIXDC1. In some examples, a DIXDC1 antibody specifically binds to the phosphorylated form of the DIXDC1 protein (such as phosphorylated at S592), but not to a non-phosphorylated DIXDC1.

In some examples, an antibody has a high binding affinity for DIXDC1, such as a binding affinity of at least about  $1 \times 10^{-8}$  M, at least about  $1.5 \times 10^{-8}$ , at least about  $2.0 \times 10^{-8}$ , at least about  $2.5 \times 10^{-8}$ , at least about  $3.0 \times 10^{-8}$ , at least about  $3.5 \times 10^{-8}$ , at least about  $4.0 \times 10^{-8}$ , at least about  $4.5 \times 10^{-8}$ , or at least about  $5.0 \times 10^{-8}$  M. In certain embodiments, an antibody that binds to DIXDC1 has a dissociation constant (Kd) of  $\leq 104$  nM,  $\leq 100$  nM,  $\leq 10$  nM,  $\leq 1$  nM,  $\leq 0.1$  nM,  $\leq 0.01$  nM, or  $\leq 0.001$  nM (*e.g.*,  $10^{-8}$  M or less, *e.g.*, from  $10^{-8}$  M to  $10^{-13}$  M, *e.g.*, from  $10^{-9}$  M to  $10^{-13}$  M). In one embodiment, Kd is measured by a radiolabeled antigen binding assay (RIA) performed with the Fab version of an antibody of interest and its antigen (see, *e.g.*, Chen *et al.*, *J. Mol. Biol.* 293:865-881, 1999). In another example, Kd is measured using surface plasmon resonance assays using a BIACORES-2000 or a BIACORES-3000 (BIAcore, Inc., Piscataway, N.J.) at 25°C with immobilized antigen CM5 chips at about 10 response units (RU). Binding can be measured using a variety of methods standard in the art, including, but not limited to: Western blot, immunoblot, enzyme-linked immunosorbant assay (ELISA), radioimmunoassay (RIA), immunoprecipitation, surface plasmon resonance, chemiluminescence, fluorescent polarization, phosphorescence, immunohistochemical analysis, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, microcytometry, microarray, microscopy, fluorescence activated cell sorting (FACS), and flow cytometry.

A naturally occurring antibody (such as IgG, IgM, IgD) includes four polypeptide chains, two heavy (H) chains and two light (L) chains interconnected by disulfide bonds. As used herein, the term antibody also includes recombinant antibodies produced by expression of a nucleic acid that encodes one or more antibody chains in a cell (for example see U.S. Patent No. 4,745,055; U.S. Patent No. 4,444,487; WO 88/03565; EP 256,654; EP 120,694; EP 125,023;

Faoulkner *et al.*, *Nature* 298:286, 1982; Morrison, *J. Immunol.* 123:793, 1979; Morrison *et al.*, *Ann Rev. Immunol.* 2:239, 1984).

The term antibody also includes an antigen binding fragment of a naturally occurring or recombinant antibody. Specific, non-limiting examples of binding fragments encompassed within the term antibody include Fab, (Fab')<sub>2</sub>, Fv, and single-chain Fv (scFv). Fab is the fragment that contains a monovalent antigen-binding fragment of an antibody molecule produced by digestion of whole antibody with the enzyme papain to yield an intact light chain and a portion of one heavy chain or equivalently by genetic engineering. Fab' is the fragment of an antibody molecule obtained by treating whole antibody with pepsin, followed by reduction, to yield an intact light chain and a portion of the heavy chain; two Fab' fragments are obtained per antibody molecule. (Fab')<sub>2</sub> is the fragment of the antibody obtained by treating whole antibody with the enzyme pepsin without subsequent reduction or equivalently by genetic engineering. F(Ab')<sub>2</sub> is a dimer of two FAb' fragments held together by disulfide bonds. Fv is a genetically engineered fragment containing the variable region of the light chain and the variable region of the heavy chain expressed as two chains. Single chain antibody ("SCA") is a genetically engineered molecule containing the variable region of the light chain, the variable region of the heavy chain, linked by a suitable polypeptide linker as a genetically fused single chain molecule. Methods of making these fragments are routine in the art.

**Contact:** To bring one agent into close proximity to another agent, thereby permitting the agents to interact. For example, a DIXDC1 probe or antibody can be applied to a microscope slide or other surface containing a biological sample (such as a cancer sample), thereby permitting detection of DIXDC1 proteins or nucleic acid molecules in the sample that are specifically recognized by the DIXDC1 antibody or probe, respectively.

**Control:** A sample or standard used for comparison with a test sample, such as a biological sample, *e.g.*, a biological sample obtained from a patient (or plurality of patients) or a cell culture. In some embodiments, the control is a sample obtained from a healthy patient (or plurality of patients) (also referred to herein as a "normal" control), such as a normal sample (*e.g.*, one that does not have cancer cells). In some examples, a normal control is a positive control (*e.g.*, expresses DIXDC1). In some examples, a normal control is a negative control (*e.g.*, does not express detectable DIXDC1). In some embodiments, the control is a historical control or standard value (*i.e.*, a previously tested control sample or group of samples that represent baseline or normal values). In some embodiments the control is a standard value

representing the average value (or average range of values) obtained from a plurality of patient samples.

A control can also be represented by a reference value or range of values representing an amount of activity or expression determined to be representative of a given condition.

5 Reference values can include a range of values, real or relative expected to occur under certain conditions. These values can be compared with experimental values to determine if a given molecule (*e.g.*, DIXDC1) is up-regulated or down-regulated in a particular sample for instance. In one example, a reference value or range of values represents an amount of activity or expression of target (*e.g.*, DIXDC1) nucleic acids or proteins in a sample, such as a sample from  
10 a subject without cancer, non-cancerous tissue as the cancer of the test sample (*e.g.*, if the sample is a breast cancer sample, the control can be a non-cancerous breast tissue sample), or non-cancerous tissue adjacent to the tumor in from the same or other patients. This value can then be used to determine if the tumor sample has one or more mutations in target (*e.g.*, DIXDC1) genes or proteins by comparing this reference value for the target (*e.g.*, DIXDC1)  
15 genes or proteins to the level detected in the test sample.

**Detect:** To determine if an agent is present or absent. In some examples this can further include quantification. For example, use of an antibody specific for a particular protein (*e.g.*, DIXDC1) permits detection of the protein in a sample, such as a sample containing cancer tissue. For example, use of a probe specific for a particular nucleic acid molecule (*e.g.*,  
20 DIXDC1) permits detection of the nucleic acid molecule in a sample, such as a sample containing cancer tissue. In particular examples, an emission signal from a detectable label (such as an increase in the signal if the target is present) is detected. Detection can be in bulk, so that a macroscopic number of molecules can be observed simultaneously. Detection can also include identification of signals from single molecules using microscopy and such techniques as  
25 total internal reflection to reduce background noise.

**Differential Expression:** A nucleic acid sequence is differentially expressed when the amount of one or more of its expression products (*e.g.*, transcript (*e.g.*, mRNA) and/or protein) is higher or lower in one tissue (or cell) type as compared to another tissue (or cell) type. Detecting differential expression can include measuring a change in (*e.g.*, DIXDC1) gene or  
30 protein expression. For example, a gene, the transcript or protein of which is less expressed (or even not expressed) in a cancer tissue (or cells) and more greatly expressed in normal issue (or cells) is differentially expressed.

**DIX domain-containing protein 1 (DIXDC1):** *e.g.*, OMIM 610493. Includes DIXDC1 nucleic acid molecules and proteins, such as DIXDC1 isoform a, b, or c. DIXDC1 is shown herein to be downregulated in some cancers, resulting increased expression of FAK. DIXDC1 sequences are publically available, for example from the GenBank® sequence database (*e.g.*,  
5 Accession Nos. NP\_001033043, NP\_001265471, and NP\_219493 provide exemplary DIXDC1 protein sequences, while Accession Nos. NM\_001037954 and NM\_001278542, and NM\_033425 provide exemplary DIXDC1 nucleic acid sequences). One of ordinary skill in the art can identify additional DIXDC1 nucleic acid and protein sequences, including DIXDC1 variants.

10 **Downregulated or inactivation:** When used in reference to the expression of a molecule, such as a (*e.g.*, DIXDC1) gene or a protein, refers to any process which results in a decrease or elimination in production of a gene product. A gene product can be RNA (such as mRNA, rRNA, tRNA, and structural RNA) or protein. Therefore, downregulation or deactivation includes processes that decrease or even eliminate transcription of a gene or  
15 translation of mRNA and thus decrease the presence of proteins or nucleic acids, such as a target protein or nucleic acid molecule.

Examples of processes that decrease transcription include those that facilitate degradation of a transcription initiation complex, those that decrease transcription initiation rate, those that decrease transcription elongation rate, those that decrease processivity of transcription  
20 and those that increase transcriptional repression. Gene downregulation can include reduction of expression above an existing level. Examples of processes that decrease translation include those that decrease translational initiation, those that decrease translational elongation and those that decrease mRNA stability.

Downregulation includes any detectable decrease in the production of a gene product,  
25 such as a DIXDC1 protein. In certain examples, detectable target protein or nucleic acid expression in a cell (such as a cancer cell) decreases by at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 10-fold, or at least 20-fold (or even not detectable) as compared to a control (such an amount of protein or nucleic acid expression detected in a corresponding normal cell or sample). In one example, a control is a relative amount of expression in a normal  
30 sample (*e.g.*, same tissue that is not cancerous).

**Effective amount or Therapeutically effective amount:** The amount of agent, such as a FAK inhibitor or Src inhibitor, that is an amount sufficient to prevent, treat (including prophylaxis), reduce and/or ameliorate the symptoms and/or underlying causes of any of a

disorder or disease. In one embodiment, an “effective amount” is sufficient to reduce the size, volume, number of metastases, rate of growth, and the like, of a cancer, for example ovarian cancer, thyroid cancer, head and neck cancer, lung cancer, kidney cancer, brain cancer, hepatocellular cancer, pancreatic cancer, colorectal cancer, breast cancer, cervical cancer, prostate cancer, melanoma, neuroblastoma, osteosarcoma, or sarcoma, for example reducing the size or volume of a tumor, reducing metastasis of a tumor, reducing a number of tumor cells in a tumor, reducing the rate of growth of a tumor, and increasing an amount of chemotherapeutic or biologic in the tumor. In one example, a therapeutically effective amount is an amount of a FAK inhibitor or Src inhibitor (or both, alone or in combination with other therapeutic agents) sufficient to reduce symptoms of a cancer (*e.g.*, ovarian cancer, thyroid cancer, head and neck cancer, lung cancer, kidney cancer, brain cancer, hepatocellular cancer, pancreatic cancer, colorectal cancer, breast cancer, cervical cancer, prostate cancer, melanoma, neuroblastoma, osteosarcoma, or sarcoma) for example by at least 10%, at least 20%, at least 50%, at least 70%, or at least 90% (as compared to no administration of the FAK inhibitor and/or Src inhibitor).

**Expression:** The process by which the coded information of a (*e.g.*, DIXDC1) gene is converted into an operational, non-operational, or structural part of a cell, such as the synthesis of a protein. Gene expression can be influenced by external signals (such as a hormone). Expression of a gene also can be regulated anywhere in the pathway from DNA to RNA to protein. Regulation can include controls on transcription, translation, RNA transport and processing, degradation of intermediary molecules such as mRNA, or through activation, inactivation, compartmentalization or degradation of specific protein molecules after they are produced.

The expression of a nucleic acid molecule or protein can be altered relative to a normal (wild type) nucleic acid molecule or protein (such as in a patient not having cancer or a normal tissue). Alterations in gene expression, such as differential expression, include but are not limited to: (1) overexpression (*e.g.*, upregulation); (2) underexpression (*e.g.*, downregulation); or (3) suppression of expression. Alternations in the expression of a nucleic acid molecule can be associated with, and in fact cause, a change in expression of the corresponding protein.

Protein expression can also be altered in some manner to be different from the expression of the protein in a normal (wild type) situation. This includes but is not necessarily limited to: (1) a mutation in the protein such that one or more of the amino acid residues is different; (2) a short deletion or addition of one or a few (such as no more than 10-20) amino acid residues to the sequence of the protein; (3) a longer deletion or addition of amino acid

residues (such as at least 20 residues), such that an entire protein domain or sub-domain is removed or added; (4) expression of an increased amount of the protein compared to a control or standard amount (*e.g.*, upregulation); (5) expression of a decreased amount of the protein compared to a control or standard amount (*e.g.*, downregulation); (6) alteration of the subcellular localization or targeting of the protein; (7) alteration of the temporally regulated expression of the protein (such that the protein is expressed when it normally would not be, or alternatively is not expressed when it normally would be); (8) alteration in stability of a protein through increased longevity in the time that the protein remains localized in a cell; and (9) alteration of the localized (such as organ or tissue specific or subcellular localization) expression of the protein (such that the protein is not expressed where it would normally be expressed or is expressed where it normally would not be expressed), each compared to a control or standard. Controls or standards for comparison to a sample, for the determination of differential expression, include samples believed to be normal (in that they are not altered for the desired characteristic, for example a sample from a subject who does not have cancer) as well as laboratory values, even though possibly arbitrarily set, keeping in mind that such values can vary from laboratory to laboratory.

Laboratory standards and values may be set based on a known or determined population value and can be supplied in the format of a graph or table that permits comparison of measured, experimentally determined values.

**Focal adhesion kinase (FAK):** *e.g.*, OMIM 600758. Includes FAK nucleic acid molecules and proteins. The FAK protein, in humans, is encoded by the PTK2 gene. It is a focal adhesion-associated nonreceptor tyrosine kinase involved in cellular adhesion and spreading. FAK has been implicated in cancer progression, and is a target for anti-cancer agents. FAK has three domains, the amino-N-terminal domain, the central catalytic domain and the carboxy-C-terminal domain. FAK activation relies upon autophosphorylation of the unique Y-397 site that is found in the N-terminal domain. Y-397 binds a number of signaling proteins such as src, PI-3 kinase, and Grb-7. It also binds EGFR, VEGFR, and p53 and other molecules that are critical for carcinogenesis. FAK also activates proteins that promote cell motility and migration, invasion, survival, angiogenesis, lymphangiogenesis, and proliferation, such as paxillin and talin.

FAK is upregulated (*e.g.*, activated) in several types of cancer including ovarian, thyroid, head and neck, lung, kidney, brain, hepatocellular, pancreatic, colorectal, breast, cervical, and prostate cancers, as well as melanoma, neuroblastoma, osteosarcoma, and sarcoma. The

potential role of FAK activation cancer growth and progression has led to the development of several therapeutic agents (called **FAK inhibitors**).

FAK sequences are publically available, for example from GenBank® sequence database (*e.g.*, Accession Nos. NP\_001186578 and NP\_001123881 provide exemplary FAK protein sequences, while Accession Nos. NM\_001199649 and NM\_001130409 provide exemplary FAK nucleic acid sequences). One of ordinary skill in the art can identify additional FAK nucleic acid and protein sequences, including FAK variants.

**Hybridization:** To form base pairs between complementary regions of two strands of DNA, RNA, or between DNA and RNA, thereby forming a duplex molecule. Hybridization conditions resulting in particular degrees of stringency will vary depending upon the nature of the hybridization method and the composition and length of the hybridizing nucleic acid sequences. Generally, the temperature of hybridization and the ionic strength (such as the Na<sup>+</sup> concentration) of the hybridization buffer will determine the stringency of hybridization. The presence of a chemical which decreases hybridization (such as formamide) in the hybridization buffer will also determine the stringency (Sadhu *et al.*, *J. Biosci.* 6:817-821, 1984). Calculations regarding hybridization conditions for attaining particular degrees of stringency are discussed in Sambrook *et al.*, (1989) *Molecular Cloning*, second edition, Cold Spring Harbor Laboratory, Plainview, NY (chapters 9 and 11). Hybridization conditions for ISH are also discussed in Landegent *et al.*, *Hum. Genet.* 77:366-370, 1987; Lichter *et al.*, *Hum. Genet.* 80:224-234, 1988; and Pinkel *et al.*, *Proc. Natl. Acad. Sci. USA* 85:9138-9142, 1988.

**Immunohistochemistry (IHC):** A method of determining the presence, distribution, and/or amount of an antigen (such as a protein) in a sample (such as a tumor sample, for example, a portion or section of tissue) by detecting interaction of the antigen with a specific binding agent, such as an antibody. A sample including an antigen (such as a target antigen) is incubated with an antibody under conditions permitting antibody-antigen binding. Antibody-antigen binding can be detected by means of a detectable label conjugated to the antibody (direct detection) or by means of a detectable label conjugated to a secondary antibody, which is raised against the primary antibody (*e.g.*, indirect detection). Exemplary detectable labels that can be used for IHC include, but are not limited to, radioactive isotopes, fluorochromes (such as fluorescein, fluorescein isothiocyanate, and rhodamine), haptens, enzymes (such as horseradish peroxidase or alkaline phosphatase), and chromogens (such as 3,3'-diaminobenzidine or Fast Red). In some examples, IHC is utilized to detect the presence of or determine the amount of one or more proteins in a sample (such as DIXDC1), for example, a tumor sample.

**Isolated:** An “isolated” biological component (such as a DIXDC1 antibody, protein, or nucleic acid molecule) has been substantially separated, produced apart from, or purified away from other biological components in the cell of the organism in which the component naturally occurs, such as other chromosomal and extrachromosomal DNA and RNA, and proteins.

5 Nucleic acids molecules and proteins which have been “isolated” thus include nucleic acids and proteins purified by standard purification methods. The term also embraces nucleic acid molecules and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acids. A purified or isolated cell, protein, or nucleic acid molecule can be at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at  
10 least 99% pure.

**Label:** An agent capable of detection, for example by spectrophotometry, flow cytometry, or microscopy (such as light microscopy). For example, one or more labels can be attached to an antibody, thereby permitting detection of the target protein (such as DIXDC1), or to a nucleic acid probe, thereby permitting detection of the target nucleic acid (such as  
15 DIXDC1). Exemplary labels include radioactive isotopes, fluorophores, ligands, chemiluminescent agents, haptens, enzymes, and combinations thereof.

**Mammal:** This term includes both human and non-human mammals. Similarly, the term “subject” includes both human and veterinary subjects (such as cats, dogs, cows, and pigs).

**Normal cells or tissue:** Non-tumor, non-malignant cells and tissue.

20 **Pharmaceutically acceptable carriers:** The pharmaceutically acceptable carriers useful in this invention are conventional. *Remington's Pharmaceutical Sciences*, by E. W. Martin, Mack Publishing Co., Easton, PA, 15th Edition (1975), describes compositions and formulations suitable for pharmaceutical delivery of a FAK inhibitor or a Src inhibitor.

In general, the nature of the carrier will depend on the particular mode of administration  
25 being employed. For instance, parenteral formulations usually comprise injectable fluids that include pharmaceutically and physiologically acceptable fluids such as water, physiological saline, balanced salt solutions, aqueous dextrose, glycerol or the like as a vehicle. For solid compositions (*e.g.*, powder, pill, tablet, or capsule forms), conventional non-toxic solid carriers can include, for example, pharmaceutical grades of mannitol, lactose, starch, or magnesium  
30 stearate. In addition to biologically-neutral carriers, pharmaceutical compositions to be administered can contain minor amounts of non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example sodium acetate or sorbitan monolaurate.

**Primer:** An oligonucleotide which hybridizes with a sequence in the target nucleic acid (such as a DIXDC1 nucleic acid sequence) and is capable of acting as a point of initiation of synthesis along a complementary strand of nucleic acid under conditions suitable for such synthesis. A perfect complementarity is not required for the primer extension to occur.

5 However, a primer with perfect complementarity (especially near the 3'-terminus) will be extended more efficiently than a primer with mismatches, especially mismatches at or near the 3'-terminus.

**Probe:** An oligonucleotide which hybridizes with a sequence in the target nucleic acid (such as DIXDC1) and may be detectably labeled. The probe can have modifications, such as a  
10 3'-terminus modification that makes the probe non-extendable by nucleic acid polymerases; and one or more chromophores.

**Quantitative PCR or quantitative RT-PCR:** A nucleic acid amplification reaction wherein the target nucleic acid (such as DIXDC1) is quantitatively detected. QPCR is characterized by a "growth curve" which is a graph of a function, where an independent variable  
15 is the number of amplification cycles and a dependent variable is an amplification-dependent measurable parameter (such as the amount of fluorescence emitted by a specific probe upon hybridization, or upon the hydrolysis of the probe by the nuclease activity of the nucleic acid polymerase) is measured at each cycle of amplification, *see* Holland *et al.*, (1991) *Proc. Natl. Acad. Sci.* 88:7276-7280 and U.S. Patent No. 5,210,015. The amplification-dependent  
20 measurable parameter reflects among other variables, the initial amount of the target nucleic acid. A growth curve is typically characterized by a "cycles to threshold" value or "C<sub>t</sub> value," which is a number of cycles where a predetermined magnitude of the measurable parameter is achieved. A lower or "earlier" C<sub>t</sub> value reflects a greater amount of the input target nucleic acid, while the higher or "later" C<sub>t</sub> value represents a lower amount of the input target nucleic acid.

25 **Proto-oncogene tyrosine-protein kinase Src (Src):** *e.g.*, OMIM 190090. Includes Src nucleic acid molecules and proteins. The Src protein, in humans, is encoded by the SRC gene. It is a non-receptor protein tyrosine kinase. Upregulation of Src has been linked to cancer, such as cancers of the colon, breast and prostate. The potential role of Src activation cancer growth and progression has led to the development of several therapeutic agents (called **Src inhibitors**).

30 Src sequences are publically available, for example from GenBank® sequence database (*e.g.*, Accession Nos. NP\_005408 and AAH11566.1 provide exemplary Src protein sequences, while Accession Nos. NM\_005417 and NM\_198291.2 provide exemplary Src nucleic acid

sequences). One of ordinary skill in the art can identify additional Src nucleic acid and protein sequences, including Src variants.

**Sequence identity of amino acid sequences:** The similarity between amino acid (or nucleotide) sequences is expressed in terms of the similarity between the sequences, otherwise referred to as sequence identity. Sequence identity is frequently measured in terms of percentage identity (or similarity or homology); the higher the percentage, the more similar the two sequences are. Homologs or variants of a polypeptide will possess a relatively high degree of sequence identity when aligned using standard methods.

Methods of alignment of sequences for comparison are well known in the art. Various programs and alignment algorithms are described in: Smith and Waterman, *Adv. Appl. Math.* 2:482, 1981; Needleman and Wunsch, *J. Mol. Biol.* 48:443, 1970; Pearson and Lipman, *Proc. Natl. Acad. Sci. U.S.A.* 85:2444, 1988; Higgins and Sharp, *Gene* 73:237, 1988; Higgins and Sharp, *CABIOS* 5:151, 1989; Corpet *et al.*, *Nucleic Acids Research* 16:10881, 1988; and Pearson and Lipman, *Proc. Natl. Acad. Sci. U.S.A.* 85:2444, 1988. Altschul *et al.*, *Nature Genet.* 6:119, 1994, presents a detailed consideration of sequence alignment methods and homology calculations.

The NCBI Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, *J. Mol. Biol.* 215:403, 1990) is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, MD) and on the internet, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastn and tblastx. A description of how to determine sequence identity using this program is available on the NCBI website on the internet.

Homologs and variants of DIXDC1 protein and coding sequences known in the art and disclosed herein are typically characterized by possession of at least about 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% sequence identity counted over the full length alignment with the amino acid sequence using the NCBI Blast 2.0, gapped blastp set to default parameters. For comparisons of amino acid sequences of greater than about 30 amino acids, the Blast 2 sequences function is employed using the default BLOSUM62 matrix set to default parameters, (gap existence cost of 11, and a per residue gap cost of 1). When aligning short peptides (fewer than around 30 amino acids), the alignment should be performed using the Blast 2 sequences function, employing the PAM30 matrix set to default parameters (open gap 9, extension gap 1 penalties). Proteins with even greater similarity to the reference sequences will show increasing percentage identities when assessed by this method, such as at least 95%, at least 98%, or at least 99% sequence identity. When less than the entire sequence is being compared for

sequence identity, homologs and variants will typically possess at least 80% sequence identity over short windows of 10-20 amino acids, and may possess sequence identities of at least 85% or at least 90% or at least 95% depending on their similarity to the reference sequence. Methods for determining sequence identity over such short windows are available at the NCBI website on the internet. One of skill in the art will appreciate that these sequence identity ranges are provided for guidance only; it is entirely possible that strongly significant homologs could be obtained that fall outside of the ranges provided.

Thus, a DIXDC1 protein can have at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% sequence identity to GenBank® Accession No. NP\_001033043, NP\_001265471, or NP\_219493. Similarly, exemplary DIXDC1 coding sequences in some examples have at least 70%, at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 98%, or at least 99% sequence identity to GenBank® Accession No. NM\_001037954, NM\_001278542, or NM\_033425.

**Sample:** A biological specimen containing genomic DNA, RNA (including mRNA), protein, intact cells (*e.g.*, a tissue sample), or combinations thereof, obtained from a subject. Examples include a specimen containing at least one cancer cell (a cancer sample or cancer tissue sample), for example, a surgical resection specimen, a tissue or tumor biopsy, fine needle aspirate, bronchoalveolar lavage, pleural fluid, breast fluid, sputum, surgical specimen, lymph node, a metastasis, or autopsy material (or proteins or nucleic acids obtained from such samples). In other examples, a sample includes a control sample, such as a non-cancerous cell or tissue sample. In one example the control is a negative control, such as a sample known to not include detectable DIXDC1 protein (such as a lung cancer cell line with defined DIXDC1 deletion). In another example, the control is a positive control, such as a sample known to include detectable DIXDC1 (such as a normal lung or liver tissue sample).

**Specific binding agent:** An agent that binds substantially or preferentially only to a defined target such as a protein, for example a DIXDC1 protein. In some examples, a DIXDC1 specific binding agent specifically binds to the phosphorylated form of DIXDC1 protein, but not to a non-phosphorylated DIXDC1.

A DIXDC1 protein-specific binding agent binds substantially only to DIXDC1 protein, but not to other proteins (such as other proteins routinely found in similar samples). For example, a “DIXDC1 specific binding agent” includes antibodies and other agents, such as aptamers, that bind substantially to a DIXDC1 polypeptide. Antibodies can be monoclonal or polyclonal antibodies that are specific for the polypeptide, as well as immunologically effective

portions (“fragments”) thereof. The determination that a particular agent binds substantially only to a DIXDC1 polypeptide may readily be made by using or adapting routine procedures. One suitable *in vitro* assay makes use of the Western blotting procedure (described in many standard texts, including Harlow and Lane, Using Antibodies: A Laboratory Manual, CSHL, New York, 1999).

**Subject:** Any mammal, such as humans, non-human primates, pigs, sheep, cows, dogs, cats, rodents and the like which is to be the recipient of the particular treatment, such as one having a cancer that can be treated with a FAK inhibitor and/or a Src inhibitor. In two non-limiting examples, a subject is a human subject or a murine subject. In some examples, the subject has a tumor, such as a cancer with increased FAK expression, such as ovarian, thyroid, head and neck, lung, kidney, brain, hepatocellular, pancreatic, colorectal, breast, cervical, and prostate cancers, as well as melanoma, neuroblastoma, osteosarcoma, and sarcoma.

**Under conditions sufficient for:** A phrase that is used to describe any environment that permits the desired activity. An example includes contacting an antibody or a nucleic acid probe with a biological sample sufficient to allow detection of one or more target proteins or nucleic acid molecules (*e.g.*, DIXDC1), respectively, in the sample.

### Overview

The serine/threonine kinase LKB1 is a tumor suppressor whose loss is associated with increased metastatic potential. In an effort to define biochemical signatures of metastasis associated with Lkb1 loss, the inventors discovered that the EMT transcription factor Snail1 was uniquely upregulated upon LKB1-deficiency across a variety of cell types. The ability of LKB1 to suppress Snail1 levels was independent of AMPK, but required the related kinases MARK1 and MARK4. In a screen for substrates of these kinases involved in Snail regulation, the scaffolding protein DIXDC1 was identified. Similar to loss of LKB1, DIXDC1 depletion results in upregulation of Snail1 in a FAK- and ERK-dependent manner, leading to increased cell invasion. MARK1 phosphorylation of DIXDC1 is required for its localization to focal adhesions and ability to suppress metastasis in mice. DIXDC1 is frequently downregulated in human cancers, which correlates with poor survival. The results provided herein define a novel AMPK-independent phosphorylation cascade essential for LKB1-dependent control of metastatic behavior.

It is shown herein that the scaffold protein DIXDC1 is the first direct substrate of LKB1-dependent kinases connected to metastatic progression. Phenotypes observed from gain and loss

of DIXDC1 function also provide a direct molecular basis for the only previously noted biomarkers of Lkb1-deficient metastasis: 1) hyperactivation of FAK/Src family kinase signaling (Carretero et al., *Cancer Cell* 17, 547-559, 2010; Liu et al., *Cancer Cell* 21, 751-764, 2012) and 2) increased expression of some EMT target genes (Carretero et al., *Cancer Cell* 17, 547-559, 2010). This effect is mediated by MARK1/4 dependent activation of DIXDC1, which suppresses the FAK-ERK-Snail1 signaling axis. While MARK1 and MARK4 do not appear redundant in the regulation of DIXDC1, it remains unclear whether this reflects their formation of a heterodimeric complex or alternative possibilities to be pursued in future studies. Phosphorylation of DIXDC1 at Ser592 by MARK1/4 is required for its proper targeting to focal adhesions and required for DIXDC1 to suppress metastatic growth in vivo.

The data provided herein indicates that DIXDC1 is involved in focal adhesion maturation. It was observed that loss of DIXDC1 results in an accumulation of nascent focal adhesions, which are documented to possess enhanced tractile strength in motile cells compared to mature adhesions (Beningo et al., *J Cell Biol* 153, 881-888, 2001). Furthermore, these nascent peripheral adhesions were largely absent for zyxin staining, a protein known to mark mature adhesions as well as decrease adhesion strength and cell migration ability (Sperry et al., *J. Cell. Phys.* 222, 612-624, 2010; Zaidel-Bar et al., *J Cell Sci* 116, 4605-4613, 2003). Zyxin-null cells display increased rates of cell motility as well as hyper-activation of the FAK/Src signaling module (Hoffman et al., *J Cell Biol* 172, 771-782, 2006; Mise et al., *J Biol Chem* 287, 31393-31405, 2012), consistent with the phenotypes we observe when zyxin is localized away from the adhesion plaque in DIXDC1-depleted cells. Thus, in a broad context, factors that regulate the maturation of focal adhesions such as DIXDC1, or even zyxin itself, might be a common target of disruption for cells to gain increased migration and invasion potential.

In addition to the predicted functional inactivation of DIXDC1 in LKB1-deficient tumors from loss of Ser592 phosphorylation, it was observed that deletions at the *DIXDC1* locus were highly selected for in the exact tumor subsets where LKB1 has been characterized as a dominant regulator of metastasis. Attributes that cause these distinct cell types to be particularly sensitive to alteration in the LKB1-DIXDC1 signaling axis remain unknown, but it is of interest to note that both melanoma and lung cancer exhibit little to no latency between initial diagnosis and the development of metastatic disease. Disruption of the function of DIXDC1, through downregulation or loss of LKB1-dependent phosphorylation, could not only result in the prompt mobilization of tumor cells through activation of an EMT-like program, but allow them to “arrive prepared” to immediately initiate metastatic colonization due to the activation of the

FAK-Src signaling axis, shown previously to aid in both the survival and outgrowth of distant micrometastases (Shibue and Weinberg, *Proc Natl Acad Sci U S A* 106, 10290-10295, 2009). This study indicates that certain initiating genetic alterations may provide robust metastatic fitness, alleviating the need for additional genetic lesions to enhance metastatic potential.

5           How does this role for DIXDC1 in suppressing cell migration and invasion fit with the few previously reported functions for DIXDC1? DIXDC1 is best known as a Dvl-binding protein under some conditions (Shiomi et al., *Curr Biol* 13, 73-77, 2003; Wong et al., *J Biol Chem* 279, 39366-39373, 2004), and Dvl has been reported to regulate cell motility and focal adhesion dynamics via affects on FAK and paxillin (Matsumoto et al., *EMBO J* 29, 1192-1204, 10           2010), though the inventors were unable to observe these effects in cells of interest herein. The central involvement of FAK in LKB1-dependent metastatic potential fits well with evidence that hyperactivation of FAK and Src may represent a central biochemical pathway broadly required for metastasis in tumors from many different tissues of origin (Shibue and Weinberg, *Proc Natl Acad Sci U S A* 106, 10290-10295, 2009; Zhang et al., *Cancer Cell* 16, 67-78, 2009)

15           Notably, the activation of DIXDC1 by phosphorylation of Ser592 downstream of LKB1 is mediated by the poorly studied kinases MARK1 and MARK4, and independent of AMPK. DIXDC1 represents one of the first AMPK-independent effectors of LKB1 in human cancer and indicates that the tumor suppressor activity of LKB1 is unlikely to be mediated by AMPK alone in some contexts. The finding that LKB1 actively suppresses the Snail1 EMT transcription 20           factor through a novel signaling pathway downstream of MARK1 and MARK4 indicates that information remains to be decoded for this ancient LKB1 pathway connecting metabolism and tumorigenesis to cell polarity and cytoskeletal control. Very few endogenous suppressors of EMT are known, which may explain why DIXDC1 downregulation would be selected for independent of LKB1 mutation, as its loss promotes Snail-dependent gene expression 25           stimulating invasion and tumor microenvironment remodeling. Analysis of DIXDC1 and its regulation by other pathways that modulate EMT will be of great interest, both in the context of different tumor types and developmental processes.

          In summary, the inventors have been able to attribute much of the suppressive effects of LKB1 on metastasis to a single substrate of some of its downstream AMPK family kinases, 30           which is remarkable given the redundancy amongst the kinases themselves. This study also provides a detailed molecular mechanism for how a single initiating mutation in the LKB1 tumor suppressor not only leads to immediate alterations in mTOR signaling, autophagy, and metabolism, but simultaneously promotes metastatic progression via loss of phosphorylation of

a single serine in a novel downstream effector. Regulation of such seemingly diverse cell processes by LKB1 may underlie its unique potency for tumor initiation and metastatic progression in some tissues, and warrants further investigation into identifying the repertoire of conserved substrates of the LKB1-dependent kinases to reveal additional rate-limiting regulators of cell biology and disease.

### Methods for Analyzing a Cancer Sample, Systems, and Kits

Based on these observations, provided herein are methods for analyzing a cancer sample obtained from a subject, for example to determine if the cancer is one that is sensitive (*e.g.*, responsive) to treatment with a FAK inhibitor and/or a Src inhibitor, to determine if the cancer is one that is likely to metastasize, or both.

Provided herein are methods for analyzing a cancer sample obtained from a subject. Examples of cancers that can be analyzed using the disclosed methods include solid tumors, such as those that may be susceptible (*e.g.*, sensitive) to treatment with a FAK inhibitor and/or Src inhibitor. In some examples, the cancer is a lung cancer (*e.g.*, NSCLC or mesothelioma), ovarian cancer, breast cancer, prostate cancer, pancreatic cancer, head and neck cancer, thyroid cancer, kidney cancer (*e.g.*, RCC), brain, liver cancer (*e.g.*, HCC), colorectal cancer (*e.g.*, colon cancer), cervical cancer, melanoma, neuroblastoma, osteosarcoma, or sarcoma. In some examples the subject from whom the sample is obtained is chemo naïve. The cancer sample can be analyzed using the disclosed methods directly (*e.g.*, as an intact tissue sample or fine needle aspirate, which for example can be fixed), or can be manipulated to extract proteins or nucleic acids, and the resulting extract analyzed (*e.g.*, using an ELISA protein assay or microarray to detect nucleic acids). Exemplary samples include a surgical resection specimen, tissue biopsy (such as a tissue section) or fine needle aspirate. In some examples, the sample is a fixed sample, such as a formalin-fixed, paraffin-embedded (FFPE) sample. Thus, a sample can refer to a tissue or cell sample, as well a sample that includes proteins and/or nucleic acids that have been extracted or isolated from the sample. In some examples, the methods include obtaining the sample from the subject.

In some examples, such methods include contacting the sample with a DIXDC1 protein-specific binding agent (*e.g.*, a DIXDC1-specific antibody, antibody fragment, or aptamer) or a DIXDC1 nucleic acid probe. For example, such contacting can be done manually, or be automated, for example by using an automated tissue stainer (which can also apply additional antibodies or other IHC reagents). Thus, such a step can be controlled by a suitably

programmed computer. The DIXDC1 protein-specific binding agent or a DIXDC1 nucleic acid probe are allowed to interact with the sample (or proteins or nucleic acids isolated from the sample) under conditions that permit binding of the DIXDC1 protein-specific binding agent to DIXDC1 proteins in the sample or that permit binding of the DIXDC1 nucleic acid probe to DIXDC1 nucleic acids (*e.g.*, genomic DNA, mRNA, and the like) in the sample (if DIXDC1 proteins/nucleic acids are present in the sample). After the specific binding agent or probe is allowed to incubate with the sample (or proteins or nucleic acids isolated from the sample), the DIXDC1 proteins or DIXDC1 nucleic acid molecules in the sample are detected or measured, for example using immunohistochemistry, *in situ* hybridization (such as FISH or SISH, or other hybridization methods, such as by using a nucleic acid array), PCR (*e.g.*, quantitative reverse transcription PCR, qRT-PCR, real time PCR, and the like), or other method in the art. In some examples, detecting DIXDC1 proteins (for example to determine if the DIXDC1 protein is present and/or phosphorylated) or DIXDC1 nucleic acid molecules in the sample involves visual inspection (*e.g.*, using light microscopy or fluorescence microscopy), flow cytometry, or image analysis of a corresponding digital image (*e.g.*, of a microscope slide or multi-well array). For example, use of a direct or indirect label on the specific binding agent or probe can be detected. Thus, wherein detecting DIXDC1 proteins or DIXDC1 nucleic acid molecules in the sample can include direct or indirect detection of binding of the DIXDC1 protein-specific binding agent or a DIXDC1 nucleic acid probe to the sample. Such a determination of DIXDC1 expression and/or phosphorylation can be qualitative or quantitative.

In one example, microscopy is used to detect DIXDC1 expression and/or phosphorylation, and can in some examples include detecting DIXDC1 expression and/or phosphorylation in one or more fields of view (FOV) of the sample (such as 1, 2, 3, 4, or 5 FOV). In some examples, the method can include determining an average or median expression of DIXDC1 expression and/or phosphorylation for the sample, for example based on measurements from a values obtained from a plurality of FOVs (or other sources, such as multiple wells on an array, or multiple analysis run on the same tissue (*e.g.*, using different sections of the sample)). In some examples, scoring DIXDC1 expression and/or phosphorylation includes a visual inspection of the total area of the sample, for example that is present on a microscope slide. In some examples, scoring DIXDC1 expression and/or phosphorylation includes an inspection of the total area of the sample, for example using a slide imager. Such a value, which can be called a DIXDC1 expression and/or phosphorylation score, can be inputted into a computer or algorithm. In one example, a sample can be scored based on the staining

intensity observed, such as a scale of 0 (negative), 1 (weak), 2 (moderate), to 3 (strong) for DIXDC1 expression and/or phosphorylation. For example, a sample can be scored negative (0) for DIXDC1 expression and/or phosphorylation if there is an absence of any detectable signal or pale gray/tan signal which is similar to the intensity on the negative control reagent; a sample  
5 can be scored weak (1) if there is by light staining intensity (*e.g.*, light brown) which is more than that seen on the negative control reagent and in the background; a sample can be scored moderate (2) if there is moderate staining intensity (*e.g.*, brown); or a sample can be scored strong (3) intensity if there is dark staining intensity (*e.g.*, brown to black signal intensity). The color detected will depend on the detection system used. In some examples, the score of the  
10 sample for DIXDC1 expression and/or phosphorylation (*e.g.*, 0, 1, 2 or 3) value is inputted into a computer or an algorithm. In some examples, a microscope slide is processed and/or imaged using a slide scanner. Slide scanners are known in the art, and can include those disclosed in U.S. Patent Nos. 8,625,930; 8,609,023, and 8,290,236 (all herein incorporated by reference). Automated scoring of DIXDC1 expression and/or phosphorylation using digital images of the  
15 slides and computer based image analysis can be used, such as those disclosed in U.S. Patent Nos. 8,625,930; 8,537,181; 8,515,683; and 8,428,887 (all herein incorporated by reference).

In another example, DIXDC1 expression is assessed by quantifying the DIXDC1 mRNA in the cancer sample. The method includes isolating RNA from the tumor sample and quantitatively detecting the DIXDC1 mRNA using a specific nucleotide probe via *e.g.*,  
20 quantitative reverse transcription polymerase chain reaction (qRT-PCR) *e.g.*, with TaqMan® probes, or PCR-free systems such as the Invader® assay (Third Wave Technologies), hybridization to immobilized probes, *e.g.*, as a part of a microarray, or any other method of quantifying mRNA that is or will become available. In some examples, relative amount of the DIXDC1 mRNA is determined by comparing the absolute amount of the DIXDC1 mRNA to  
25 that of a housekeeping gene. In yet other examples, the relative amount of the DIXDC1 mRNA in a tumor sample is compared to the relative amount of the DIXDC1 mRNA in a control, *e.g.*, non-tumor sample. The tumor sample is concluded to be DIXDC1 negative (or have reduced DIXDC1 expression, and thus sensitive to an FAK and/or Src inhibitor or likely to metastasize) if the relative amount of the DIXDC1 mRNA is substantially lower than the relative amount of  
30 the DIXDC1 mRNA in a positive control sample. In some examples, the raw or normalized value of DIXDC1 mRNA expression is inputted into a computer or an algorithm.

Based on the DIXDC1 proteins or DIXDC1 nucleic acid molecules detected in the sample, it is determined that the cancer is sensitive to a FAK inhibitor and/or a Src inhibitor

when DIXDC1 protein or DIXDC1 nucleic acid molecule expression (or DIXDC1 phosphorylation, *e.g.*, at Ser 592) is reduced in the cancer sample relative to a control or reference value (such as a value or range of values representing DIXDC1 expression expected in a corresponding tissue normal sample, *i.e.* expression is relative to a normal tissue of the same type as the cancer, *e.g.*, if cancer is breast cancer, compare to normal breast tissue sample).

Thus, in some examples the method can include detecting DIXDC1 proteins or DIXDC1 nucleic acid molecules in a control sample. For example, the control sample can be a positive control sample in which expression and/or phosphorylation of DIXDC1 is expected. Examples of such samples include cell lines from normal non-cancerous tissues such as BJ fibroblasts (*e.g.*, ATCC® Accession No. CRL-2522) or samples from normal liver or lung tissues (*e.g.*, obtained from a human or other mammal). In another example, the sample can be a negative control sample in which expression and/or phosphorylation of DIXDC1 is not expected. Examples of such samples include lung cancer cell lines with defined DIXDC1 deletion or mRNA suppression, *e.g.*, 368T1 TMet cells. The method can then include comparing DIXDC1 protein or DIXDC1 nucleic acid molecule expression (or DIXDC1 and/or phosphorylation) in the control sample to the cancer sample, wherein detection of reduced DIXDC1 protein or DIXDC1 nucleic acid molecule expression (or DIXDC1 phosphorylation, *e.g.*, at Ser 592) in the cancer sample relative to the positive control sample indicates that the cancer is sensitive to a FAK inhibitor and/or Src inhibitor, such as a reduction of at least 25%, at least 50%, at least 75%, at least 90%, at least 95%, or at least 99%. In contrast, detection of similar (such as a variance of no more than 10%, no more than 5% or no more than 1%) DIXDC1 protein or DIXDC1 nucleic acid molecule expression (or DIXDC1 phosphorylation, *e.g.*, at Ser 592) in the cancer sample relative to the negative control sample indicates that the cancer is sensitive to a FAK inhibitor and/or Src inhibitor.

One skilled in the art will appreciate that instead of analyzing a control sample with a test sample, that reference values or ranges can be used. For example, the reference value(s) or range of values can be a number or range of numbers obtained previously from a group of positive control samples in which DIXDC1 expression and/or phosphorylation was detected. Examples of such samples include cell lines from normal non-cancerous tissues such as BJ fibroblasts (*e.g.*, ATCC® Accession No. CRL-2522) or samples from normal liver or lung tissues (*e.g.*, obtained from a human or other mammal). In another example, the reference value(s) or range of values can be a number or range of numbers obtained previously from a group of negative control samples in which expression of DIXDC1 and/or phosphorylation was not detected (or

where the detection was minimal). Examples of such samples include lung cancer cell lines with defined DIXDC1 deletion or mRNA suppression, *e.g.*, 368T1 TMet cells (Monte Winslow, Stanford University). The method can then include comparing DIXDC1 protein or DIXDC1 nucleic acid molecule expression (and/or DIXDC1 phosphorylation) to the reference value sample to the cancer sample, wherein detection of reduced DIXDC1 protein or DIXDC1 nucleic acid molecule expression (or DIXDC1 phosphorylation, *e.g.*, at Ser 592) in the cancer sample relative to the positive reference value indicates that the cancer is sensitive to a FAK inhibitor and/or Src inhibitor, such as a reduction of at least 25%, at least 50%, at least 75%, at least 90%, at least 95%, or at least 99%. In contrast, detection of similar (such as a variance of no more than 10%, no more than 5% or no more than 1%) DIXDC1 protein or DIXDC1 nucleic acid molecule expression (or DIXDC1 phosphorylation, *e.g.*, at Ser 592) in the cancer sample relative to the negative reference value indicates that the cancer is sensitive to a FAK inhibitor and/or Src inhibitor.

In some examples, the method further includes detecting FAK expression in the sample, wherein increased expression of FAK in combination with decreased DIXDC1 expression indicates that the cancer is sensitive to a FAK inhibitor and/or a Src inhibitor.

In some examples, the method is a method of distinguishing between a subject having a cancer that is likely to respond to treatment with a FAK inhibitor and/or Src inhibitor from a subject having a cancer that is not likely to respond to treatment with a FAK inhibitor and/or Src inhibitor. That is, the methods are methods of identifying a subject as having a cancer likely to or predicted to respond to treatment with a FAK inhibitor and/or Src inhibitor. For example, the method can include determining that the subject will benefit from treatment with a FAK inhibitor and/or Src inhibitor if the tumor sample obtained from the subject has decreased DIXDC1 expression and/or phosphorylation (*e.g.*, relative to a positive control or reference value(s)). Such methods can also include selecting a subject for treatment with a FAK inhibitor and/or Src inhibitor if the subject is identified as a subject who is likely to respond to treatment with a FAK inhibitor and/or Src inhibitor. In some examples, the method also includes administering a therapeutically effective amount of the FAK inhibitor and/or Src inhibitor to the subject identified as a subject who is likely to respond to treatment with the FAK inhibitor and/or Src inhibitor. In contrast, if the subject is identified as a subject who is not likely to respond to treatment with a FAK inhibitor and/or Src inhibitor, such subjects are not administered a FAK inhibitor and/or Src inhibitor, but can be assigned for close monitoring or other therapy.

The FAK inhibitor can be an agent that reduces or inhibits expression or activity of FAK, though 100% reduction is not required (for example agents resulting in reductions of at least 95%, at least 90%, at least 75%, or at least 50% can be used). Examples of such agents include inhibitory nucleic acid molecules (*e.g.*, siRNA) as well as small molecules such as any of

5 TAE226 (also known as NVP-226); PF-562,271 (also known as VS-6062)); PF-573,228; PF-04554878 (also known as VS-6063); GSK2256098; PND-1186 (also known as VS-4718); 1,2,4,5-benzenetetraamine tetrahydrochloride; or combinations thereof. The Src inhibitor can be an agent that reduces or inhibits expression or activity of Src, though 100% reduction is not required (for example agents resulting in reductions of at least 95%, at least 90%, at least 75%,

10 or at least 50% can be used). Examples of such agents include inhibitory nucleic acid molecules (*e.g.*, siRNA) as well as small molecules such as any of Dasatinib (N-[2-Chloro-6-methylphenyl]-2-[[6-[4-(2-hydroxyethyl)-1-piperazinyl]-2-methyl-4-pyrimidinyl]amino]-5-thiazolecarboxamide), Saracatinib (AZD0530; N-(5-chlorobenzo[d][1,3]dioxol-4-yl)-7-(2-(4-methylpiperazin-1-yl)ethoxy)-5-(tetrahydro-2H-pyran-4-yloxy)quinazolin-4-amine), Bosutinib

15 (SKI-606), KX2-391 (KX01), quercetin, SU6656, SU6657, PP2, PP1, WHI-P 154; NVP-BHG712, Lavendustin C, ON-01910, SD 1008, Indirubin Derivative E804, ZM-306416, PP121, Src/EGFR inhibitor, Dasatinib  $\beta$ -D-Glucuronide, or combinations thereof.

In some examples, the method is a method of predicting the likelihood that the cancer will metastasize. For example, if reduced DIXDC1 protein expression and/or phosphorylation

20 (*e.g.*, at Ser 592) or DIXDC1 nucleic acid molecule expression in the cancer sample is detected relative to a positive control sample or positive reference value(s), this indicates that the cancer is more likely to metastasize. For example, if DIXDC1 protein expression and/or phosphorylation (*e.g.*, at Ser 592) or DIXDC1 nucleic acid molecule expression in the cancer sample is similar to a negative control sample or negative reference value(s), this indicates that the cancer is less

25 likely to metastasize. Examples of control samples are provided herein. Thus, the method can include comparing DIXDC1 protein or DIXDC1 nucleic acid molecule expression (and/or DIXDC1 phosphorylation) to the reference value sample to the cancer sample. Detection of reduced (*e.g.*, a reduction of at least 25%, at least 50%, at least 75%, at least 90%, at least 95%, or at least 99%) DIXDC1 protein or DIXDC1 nucleic acid molecule expression (or DIXDC1

30 phosphorylation, *e.g.*, at Ser 592) in the cancer sample relative to the positive reference value indicates that the cancer is more likely to metastasize. For example, such a cancer may be at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 75%, at least 80%, at least 90%, or at least 95% more likely to metastasize within 1 month, within 2 months, 3

months, within 4 months, within 5 months, within 6 months, within 9 months, or within 1 year, for example relative to a cancer with normal or increased DIXDC1 expression. In contrast, detection of similar (such as a variance of no more than 10%, no more than 5% or no more than 1%) DIXDC1 protein (or DIXDC1 phosphorylation, *e.g.*, at Ser 592) or DIXDC1 nucleic acid molecule expression in the cancer sample relative to the negative reference value indicates that the cancer is less likely to metastasize. For example, such a cancer may have a likelihood of metastasis of less than 20%, less than 10%, less than 5%, less than 4%, less than 3%, less than 2%, or less than 1% likelihood of metastasis within 1 month, within 2 months, 3 months, within 4 months, within 5 months, within 6 months, within 9 months, or within 1 year, for example relative to a cancer with normal or increased DIXDC1 expression.

Methods of treatment are also provided. In some examples such methods include analyzing a cancer sample obtained from a subject according to method as described herein, and administering a therapeutically effective amount of a FAK inhibitor and/or Src inhibitor to the subject identified as a subject who is likely to respond to treatment with the FAK inhibitor and/or Src inhibitor.

Any of the methods provided herein, such as one or more steps of the disclosed methods, can be performed by a suitably-programmed computer. In some examples, the detected DIXDC1 protein expression, DIXDC1 phosphorylation, and/or DIXDC1 nucleic acid molecule expression in the cancer sample and/or the control sample(s) are inputted into a computer or an algorithm. For example, a score calculated for each of DIXDC1 protein expression, DIXDC1 phosphorylation, and/or DIXDC1 nucleic acid molecule expression in the cancer sample and/or the control sample(s) can be inputted into a computer or an algorithm. The computer or algorithm can then generate an output, for example in a visual or audible output (such as a print-out), thereby analyzing the sample. In some examples, the output is stored, for example on a computer readable medium. Such an output can be related to whether or not the cancer is one that is sensitive to an FAK and/or Src inhibitor, whether the tumor is likely to metastasize, or combinations thereof. For example, the output can be a visual or audible “YES”, “NO”, indicating that the tumor is or is not likely to metastasize or be sensitive to an FAK inhibitor and/or Src inhibitor, or provide an indication that a FAK inhibitor and/or Src inhibitor should or should not be administered to the subject.

Also provided are computer-implemented methods for determining whether a cancer is likely to respond to a FAK inhibitor and/or Src inhibitor and/or likely to metastasize. Such methods can include generating a DIXDC1 protein or nucleic acid expression (or

phosphorylation) score or value (such as a staining intensity score of 0, 1, 2 or 3, or a raw or normalized value obtained from the detection/measuring method used) based at least on measured DIXDC1 protein expression, DIXDC1 phosphorylation, and/or DIXDC1 nucleic acid expression within a displayed image depicting a cancer sample detectably labeled with a DIXDC1 specific binding agent or probe, wherein the cancer sample is obtained from a subject. The method can also include outputting (such as a visual or audible output) a DIXDC1 protein expression, phosphorylation, or nucleic acid expression score for the sample. In some examples, the DIXDC1 expression or phosphorylation score is based on the intensity of staining of the sample, such as a score of 0, 1, 2, or 3. In some examples, such a method can include outputting a prognosis for the subject by the computer, for example to a user or an algorithm. Such a prognosis can include an indication as to whether or not the cancer is sensitive to a FAK inhibitor and/or a Src inhibitor, an indication as to the likelihood that the cancer will metastasize, or both. For example, the output can be a visual or audible “YES”, “NO”, indicating that the tumor is or is not likely to metastasize or be sensitive to an FAK inhibitor and/or Src inhibitor, or provide an indication that a FAK inhibitor and/or Src inhibitor should or should not be administered to the subject.

The disclosure also provides one or more non-transitory computer-readable media comprising computer-executable instructions causing a computing system to perform the methods provided herein.

Also provided herein are systems for analyzing a cancer sample obtained from a subject. Such systems can include a means for measuring a level of DIXDC1 protein expression and/or phosphorylation and/or DIXDC1 nucleic acid molecule expression in the test sample or in a control sample (such as a DIXDC1 -specific antibody or probe, a CCD camera, a spectrophotometer, a computer screen, an appropriately programmed computer, a microscope, or combinations thereof). In some examples, such means include a light microscope, automated tissue or slide stainer, computer, or combinations thereof. In some examples, the system includes means for imputing test or control/reference scores or values for DIXDC1 protein expression and/or phosphorylation and/or DIXDC1 nucleic acid molecule expression into an algorithm, such as a keyboard or computer program. In some examples, the DIXDC1 protein expression and/or phosphorylation and/or DIXDC1 nucleic acid molecule expression reference values are stored values or stored digital images. In some examples, the DIXDC1 protein expression and/or phosphorylation and/or DIXDC1 nucleic acid molecule expression reference values are a level of DIXDC1 protein expression and/or phosphorylation and/or DIXDC1

nucleic acid molecule expression measured from a control sample by said means for measuring. In some examples, the system includes implemented rules for comparing the measured level of DIXDC1 protein expression and/or phosphorylation or DIXDC1 nucleic acid expression to a DIXDC1 reference value (such as an appropriate algorithm), and means for implementing the rules (such as an appropriate programmed computer or algorithm), whereby an indication of the likely risk of cancer metastasis and/or sensitivity of the cancer to a FAK inhibitor and/or a Src inhibitor is provided based on the measured level of DIXDC1 protein or DIXDC1 nucleic acid molecule in the sample.

Kits useful with the disclosed methods are provided. Such kits can include a DIXDC1 specific-binding agent (such as a DIXDC1 specific-antibody or fragment thereof, or a DIXDC1 specific-aptamer) and/or a DIXDC1 nucleic acid probe (wherein the DIXDC1 specific-binding agent and/or the probe may be labeled), and one or more of a pair of primers specific for a DIXDC1 gene sequence; microscope slides; labeled secondary antibodies; materials for collecting and/or storing a fine needle aspirate or surgical specimen, nucleoside triphosphates, one or more DNA polymerases, and buffers for immunohistochemistry or nucleic acid hybridization (such as *in situ* hybridization). In one example a kits includes oligonucleotide primers and one or more probes specific for the control gene, *e.g.*, a “house-keeping” gene. Such reagents can be in separate vials or containers. In some examples the kits also include one or more vials containing FAK inhibitor and/or Src inhibitor. In some examples the kit includes reference values for FAK expression expected if FAK is up or down regulated.

#### **A. Detection of DIXDC1**

The cancer sample obtained from a subject is analyzed to determine if it contains DIXDC1, such as detectable levels of DIXDC1 protein and/or nucleic acid molecules. Thus, the sample can be analyzed to detect or measure the presence of DIXDC1 protein or nucleic acid (*e.g.*, genomic DNA or mRNA) in the sample. In some examples, the sample is analyzed to determine if the DIXDC1 protein is phosphorylated, for example at Ser 592 (for example using a DIXDC1 phospho-specific antibody). Such measurements can be a qualitative, quantitative, or semi-quantitative measurement. In particular embodiments, the disclosed methods utilize qualitative measurement of the presence of DIXDC1 protein, DIXDC1 phosphorylated protein or DIXDC1 nucleic acid in cancer cells in the sample. Expression of DIXDC1 can be assessed by any method known to one of ordinary skill in the art. In particular embodiments, DIXDC1 is detected at the protein level (for example, presence and/or amount of one or more proteins). In

other embodiments, DIXDC1 is detected at the nucleic acid level (for example, presence and/or amount of genomic DNA, RNA (such as mRNA) and/or presence or absence of a gene copy number variation).

### 1. Protein Expression

5 In some examples, DIXDC1 is about 200 to 683 amino acids (aa) in length, such as 210 aa to 683 aa, 215 aa to 683 aa, 218 aa to 683 aa, or 218 aa, 472 aa or 683 aa in length, such as GenBank® Accession No. NP\_001033043, NP\_001265471, or NP\_219493. Protein expression can include measuring or detecting the DIXDC1 protein, and/or measuring or determining whether DIXDC1 protein, if present, is phosphorylated.

10 In some examples, the DIXDC1 specific binding agent is an antibody, such as a polyclonal or monoclonal antibody, or fragment thereof. Such a DIXDC1-specific binding agent, such as an antibody or aptamer, can in some examples be used to distinguish between phosphorylated and non-phosphorylated forms of DIXDC1. Thus, in some examples the DIXDC1 antibody only bind with high affinity to DIXDC1 when it is phosphorylated (*e.g.*, at  
15 Ser592) (for example produces detectable signal if only if phosphorylated DIXDC1 is present). If desired, the DIXDC1 antibody can include a detectable label to permit detection and in some cases quantification of the DIXDC1 protein/antibody complex. In other examples, the DIXDC1 antibody is detected with an appropriate labeled secondary antibody. In additional examples, the DIXDC1 antibody is detected with an appropriate labeled tertiary antibody.

20 Antibodies specific for DIXDC1 can be used for detection and quantitation of DIXDC1 protein by one of a number of immunoassay methods that are well known in the art, such as those presented in Harlow and Lane (*Antibodies, A Laboratory Manual*, CSHL, New York, 1988). Methods of constructing such antibodies are known in the art. In addition, such antibodies are commercially available.

25 Exemplary commercially available antibodies include DIXDC1 antibodies available from Santa Cruz Biotechnology, Santa Cruz, CA (*e.g.*, catalog numbers sc-109529, sc-292126, and sc-377160, and Abcam, Cambridge, MA (*e.g.*, catalog numbers ab177517 and ab139306). Other examples are provided in Example 1 below.

30 A person of ordinary skill in the art will appreciate that other DIXDC1 antibodies can be used in the methods provided herein, including those now available or developed in the future. For example, methods of preparing antibodies against a specific target protein are well known in the art, including methods of making antibodies specifically phosphorylated at a particular position. A DIXDC1 protein or a fragment or conservative variant thereof can be used to

produce antibodies which are immunoreactive or specifically bind to an epitope of the DIXDC1 protein. Polyclonal antibodies, antibodies which consist essentially of pooled monoclonal antibodies with different epitopic specificities, as well as distinct monoclonal antibody preparations are included. The preparation of polyclonal antibodies is well known to those skilled in the art. See, for example, Green *et al.*, “Production of Polyclonal Antisera,” in: *Immunochemical Protocols*, pages 1-5, Manson, ed., Humana Press, 1992; Coligan *et al.*, “Production of Polyclonal Antisera in Rabbits, Rats, Mice and Hamsters,” in: *Current Protocols in Immunology*, section 2.4.1, 1992. The preparation of monoclonal antibodies likewise is conventional (see, for example, Kohler & Milstein, *Nature* 256:495, 1975; Coligan *et al.*, sections 2.5.1-2.6.7; and Harlow *et al.* in: *Antibodies: a Laboratory Manual*, page 726, Cold Spring Harbor Pub., 1988).

Any standard immunoassay format (such as ELISA, Western blot, immunoprecipitation, or radioimmunoassay) can be used to measure protein levels. Thus, in one example, polypeptide levels of DIXDC1 in a cancer sample can readily be evaluated using these methods.

Immunohistochemical techniques (such as IHC) can also be utilized for DIXDC1 detection and quantification. General guidance regarding such techniques can be found in Bancroft and Stevens (*Theory and Practice of Histological Techniques*, Churchill Livingstone, 1982) and Ausubel *et al.* (*Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1998). In some examples an average, mean, or median amount of a protein (such as DIXDC1) is determined in a sample, for example in one or more microscope fields of view (FOV) of an IHC sample or in a plurality of sections from the sample. In a particular example, the median DIXDC1 level is determined in a cancer sample from a subject using IHC.

IHC can determine the presence or distribution of an antigen (such as a protein) in a sample (such as a cancer sample, for example, a portion or section of tissue including DIXDC1-expressing cancer cells or tissue) by detecting interaction of the antigen with a specific binding agent, such as an antibody or aptamer. A sample including an antigen (such as DIXDC1) is incubated with a DIXDC1-specific antibody or aptamer under conditions permitting antibody-antigen or aptamer-antigen binding. Antibody-antigen binding can be detected by means of a detectable label conjugated to the antibody/aptamer (direct detection) or by means of a detectable label conjugated to a secondary antibody, which is raised against the primary antibody (*e.g.*, indirect detection). In other examples of indirect detection, antibody-antigen binding is detected by means of a detectable label conjugated to a tertiary antibody which is capable of binding to a secondary antibody (*e.g.*, is raised against the secondary antibody or is

raised against a molecule conjugated to the secondary antibody, such as a hapten). Exemplary detectable labels that can be used for IHC include, but are not limited to, radioactive isotopes, fluorochromes (such as fluorescein, fluorescein isothiocyanate, and rhodamine), haptens, enzymes (such as horseradish peroxidase or alkaline phosphatase), and chromogens (such as 3,3'-diaminobenzidine (DAB) or Fast Red). In some examples, detection of antigen-antibody binding also includes signal amplification (such as tyramide signal amplification or related methods). The signal amplification method may include methods described in U.S. Pat. Publ. No. 2012/0171668.

In some examples, a sample is obtained from a subject (such as a tumor sample that is known or suspected of being sensitive to treatment with an FAK inhibitor or Src inhibitor), and processed for IHC. For example, the sample can be fixed and embedded, for example with formalin and paraffin. The sample can then be mounted on a support, such as a glass microscope slide. For example, the sample can be sliced into a series of thin sections (for example, using a microtome), and the sections mounted onto a microscope slide. In some examples, a single slide includes multiple tissue sections from the same cancer sample or sections from the same cancer sample can be placed on different slides. Different sections of the cancer sample can then be individually labeled with different antibodies, for example an anti-DIXDC1 antibody and a negative control antibody (for example, an antibody that does not specifically bind to an endogenous antigen in the sample). That is, one section can be labeled with DIXDC1 antibody and another section can be labeled with a negative control antibody (such as an antibody that binds to a target that does not occur endogenously in the sample). In some examples, a separate slide from the same subject is stained with H&E (such as an adjacent or serial section from the same tumor sample). In some examples, additional proteins of interest can be detected in the same or additional tissue samples by labeling with further antibodies (for example other tumor markers, such as p53). In some examples, an automated slide or tissue stainer (such as VENTANA BENCHMARK instruments, for example BenchMark XT or BenchMark GX instruments) can be used to stain and process the slides.

In some examples, detecting DIXDC1 protein in the sample includes indirect detection of binding of the DIXDC1 antibody to the sample (for example, the DIXDC1 (primary) antibody is not detectably labeled). For example, the sample is contacted with a DIXDC1 antibody under conditions sufficient for the DIXDC1 antibody to bind to DIXDC1 protein in the sample. The sample is then contacted with a secondary antibody that can specifically bind to the DIXDC1 antibody (such as an anti-rabbit antibody, if the DIXDC1 antibody is a rabbit antibody or an

anti-mouse antibody, if the DIXDC1 antibody is a mouse antibody) under conditions sufficient for the secondary antibody to bind to the DIXDC1 antibody. The secondary antibody can be detectably labeled. The detectable label can be conjugated to the secondary antibody. In some examples, the detectable label conjugated to the secondary antibody can be directly detected (such as a fluorescent label, or an enzyme, which can produce a detectable reaction product in the presence of suitable substrate). In other examples, the secondary antibody is conjugated to one or more haptens (such as fluorescein, dinitrophenyl, biotin, or 3-hydroxyquinoxaline-2-carboxylic acid (HQ)). The sample is then contacted with a tertiary antibody that can specifically bind the hapten-conjugated secondary antibody (for example, an anti-hapten antibody, such as an anti-HQ antibody) under conditions sufficient for the tertiary antibody to bind to the hapten. In some examples, the tertiary antibody is conjugated to a detectable label, such as an enzyme (for example, horseradish peroxidase (HRP) or alkaline phosphatase (AP)). The sample is then contacted with one or more reagents that produce a detectable reaction product in the presence of the enzyme. In some examples, the sample is contacted with an HRP substrate (such as hydrogen peroxide) and a chromogen (such as DAB) that produces a visually detectable product in the presence of HRP. In some examples, detecting DIXDC1 protein in the sample is carried out using light microscopy, fluorescence microscopy, or flow cytometry.

In particular embodiments, detecting DIXDC1 protein in the sample includes indirect detection including signal amplification. In some examples, signal amplification allows unequivocal detection of DIXDC1-positive specimens which may exhibit only weak staining without signal amplification. Signal amplification methods for IHC are known to one of ordinary skill in the art. In some examples, signal amplification includes CAtalyzed Reporter Deposition (CARD), also known as Tyramide Signal Amplification (TSA™). In one variation of this method an enzyme-conjugated secondary antibody (such as an HRP-conjugated secondary antibody) binds to the primary antibody. Next a substrate of biotinylated tyramide (tyramine is 4-(2- aminoethyl)phenol) is used, which presumably becomes a free radical when interacting with the HRP enzyme. The phenolic radical then reacts quickly with the surrounding material, thus depositing or fixing biotin in the vicinity. This process is repeated by providing more substrate (biotinylated tyramide) and building up more localized biotin. Finally, the “amplified” biotin deposit is detected with streptavidin attached to a fluorescent molecule. Alternatively, the amplified biotin deposit can be detected with avidin-peroxidase complex, which is then contacted with DAB to produce a brown color.

In other examples, signal amplification includes contacting the sample with hydrogen peroxide and a tyramide-HQ conjugate after contacting the sample with an HRP-conjugated tertiary antibody under conditions sufficient for depositing HQ at or near the site of the primary antibody bound to the sample. The sample is then contacted with an enzyme-conjugated  
5 antibody (such as an HRP- or AP-conjugated antibody) that specifically binds to HQ. In some examples, this enzyme-conjugated antibody is the same as the HRP-conjugated tertiary antibody. In other examples, the enzyme-conjugated antibody is a different antibody than the HRP-conjugated tertiary antibody. The sample is then contacted with one or more reagents that produce a detectable reaction product in the presence of the enzyme. In some examples, the  
10 sample is contacted with an HRP substrate (such as hydrogen peroxide) and a chromogen (such as DAB) that produces a visually detectable product in the presence of HRP.

Quantitation of DIXDC1 proteins can be achieved by immunoassay. The amount of the proteins can be assessed in the tumor and optionally in the adjacent non-tumor tissue or in tissue from cancer-free subjects. Quantitative spectroscopic methods, such as SELDI, can be used to  
15 analyze protein expression in a sample (such as tumor tissue, non-cancerous tissue, and tissue from a cancer-free subject). In one example, surface-enhanced laser desorption-ionization time-of-flight (SELDI-TOF) mass spectrometry is used to detect protein expression, for example by using the ProteinChip™ (Ciphergen Biosystems, Palo Alto, CA). Such methods are well known in the art (for example see U.S. Pat. No. 5,719,060; U.S. Pat. No. 6,897,072; and U.S. Pat. No.  
20 6,881,586). SELDI is a solid phase method for desorption in which the analyte is presented to the energy stream on a surface that enhances analyte capture or desorption.

## 2. Nucleic Acid Expression

DIXDC1 gene expression can also be evaluated by detecting nucleic acids (including mRNA or genomic DNA). Thus, the disclosed methods can include detecting the presence or  
25 amount of one or more DIXDC1 nucleic acid molecules.

Methods of detecting nucleic acids include methods based on hybridization analysis of polynucleotides, methods based on sequencing of polynucleotides, and proteomics-based methods. In some examples, mRNA expression in a sample is quantified using Northern blotting or *in situ* hybridization (Parker & Barnes, *Methods in Molecular Biology* 106:247-283,  
30 1999); RNase protection assays (Hod, *Biotechniques* 13:852-4, 1992); and PCR-based methods, such as reverse transcription polymerase chain reaction (RT-PCR) (Weis *et al.*, *Trends in Genetics* 8:263-4, 1992). Alternatively, antibodies can be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-

protein duplexes. Representative methods for sequencing-based gene expression analysis include Serial Analysis of Gene Expression (SAGE), and gene expression analysis by massively parallel signature sequencing (MPSS).

In some examples, gene expression is identified or confirmed using the microarray technique. Thus, expression can be measured in either fresh or paraffin-embedded tumor tissue, or nucleic acids isolated from such tissue, using microarray technology. In this method, nucleic acid sequences of interest (*e.g.*, DIXDC1 cDNAs and oligonucleotides) are plated, or arrayed, on a microchip substrate. The arrayed sequences are then hybridized with isolated nucleic acids (such as cDNA or mRNA) from cells or tissues of interest (such as from a cancer sample). The presence of detectable hybridization between the DIXDC1-specific nucleic acids on the array, and the nucleic acids from the sample, indicate the presence of DIXDC1 in the sample.

*In situ* hybridization (ISH), such as FISH and SISH, is another method for detecting and comparing expression or copy number of DIXDC1. ISH applies and extrapolates the technology of nucleic acid hybridization to the single cell level, and, in combination with the art of cytochemistry, immunocytochemistry and immunohistochemistry, permits the maintenance of morphology and the identification of cellular markers to be maintained and identified, and allows the localization of sequences to specific cells within populations, such as tissues and blood samples. ISH is a type of hybridization that uses a complementary nucleic acid to localize one or more specific nucleic acid sequences in a portion or section of tissue (*in situ*), or, if the tissue is small enough, in the entire tissue (whole mount ISH). RNA ISH can be used to assay expression patterns in a tissue, such as the expression of DIXDC1.

Sample cells or tissues can be treated to increase their permeability to allow a probe (such as a DIXDC1 probe), to enter the cells. The probe is added to the treated cells, allowed to hybridize at pertinent temperature, and excess probe is washed away. A complementary probe is labeled so that the probe's location and quantity in the tissue can be determined, for example, using autoradiography, fluorescence microscopy or immunoassay. The sample may be any sample as herein described, such as a cancer sample.

In some examples an average, mean, or median amount of a nucleic acid (such as DIXDC1) is determined in a sample, for example in one or more microscope fields of view (FOV) of an ISH sample or in a plurality of sections from the sample. In a particular example, the median DIXDC1 level is determined in a cancer sample from a subject using ISH.

### C. Exemplary Samples

In some examples, the disclosed methods include the step of obtaining a sample, preparing the sample for analysis (for example fixing the sample, contacting it with DIXDC1 antibodies or probes), or both. Methods of obtaining a biological sample, such as a surgical resection specimen, from a subject are known in the art. For example, methods of obtaining tissue and cells, such as tissue or cells from the lung, ovary, breast, prostate, pancreas, head and neck cancer, thyroid, kidney, brain, liver, colon, rectum, cervix, skin, bone or brain, are routine. For example, a sample from a tumor that contains cellular material, can be obtained by surgical excision of all or part of the tumor, by collecting a fine needle aspirate from the tumor, as well as other methods known in the art. In some examples, the sample is obtained from a subject having or suspected to have lung cancer (*e.g.*, NSCLC or mesothelioma), ovarian cancer, breast cancer, prostate cancer, pancreatic cancer, head and neck cancer, thyroid cancer, kidney cancer (*e.g.*, RCC), brain, liver cancer (*e.g.*, HCC), colorectal cancer (*e.g.*, colon cancer), cervical cancer, melanoma, neuroblastoma, osteosarcoma, or sarcoma. In particular examples, the sample obtained from the subject includes tumor cells, such as at least a portion of a tumor. In some examples, the sample from the subject also includes normal (*e.g.*, non-tumor) tissue or cells. In some examples, the tumor sample is placed in 10% neutral buffered formalin upon removal from the subject.

The sample can be fresh, frozen, or fixed. In some examples, samples are processed post-collection by fixation and in some examples are wax- (*e.g.*, paraffin-) embedded. Fixatives for mounted cell and tissue preparations are well known in the art and include, without limitation, formalin fixative, 95% alcoholic Bouin's fixative; 95% alcohol fixative; B5 fixative, Bouin's fixative, Karnovsky's fixative (glutaraldehyde), Hartman's fixative, Hollande's fixative, Orth's solution (dichromate fixative), and Zenker's fixative (see, *e.g.*, Carson, *Histotechnology: A Self-Instructional Text*, Chicago:ASCP Press, 1997). DIXDC1 staining intensity may vary depending on the particular fixatives used (such as 95% alcohol, AFA, B5, or Prefer). In particular examples, the sample is fixed in neutral buffered formalin (such as 10% neutral buffered formalin) or zinc formalin. In some examples, the sample is fixed for at least about 6 hours (for example, about 6-48 hours, 12-24 hours or about 6, 12, 16, 18, 24, 36, or 48 hours). In additional examples, the sample is placed in fixative within about 6 hours of collection (for example, within about 15 minutes, 30 minutes, 1, 2, 3, 4, 5, or 6 hours).

In some examples, the sample can be a fixed, wax-embedded tissue sample, such as a fixed, wax-embedded tissue sample including tumor cells. In some examples, the sample is a

tissue section including tumor cells labeled with a primary antibody specific for DIXDC1, which may be labeled directly or indirectly (*e.g.*, with a labeled secondary antibody), which in some examples is further stained with H&E (*e.g.*, using an adjacent or serial section from the same sample). In some examples, the sample is a tissue section including tumor cells labeled with a nucleic acid probe specific for DIXDC1, which may be labeled directly or indirectly, which in some examples is further stained with H&E (*e.g.*, using an adjacent or serial section from the same sample).

In some examples the tissue or cells obtained from the subject is treated to extract or isolate proteins and/or nucleic acids from the sample. Such methods are routine. For example, isolated proteins can be analyzed using antibodies, for example by Western blotting, or by mass spec. Isolated nucleic acids can be analyzed using nucleic acid primers and/or probes, for example by PCR, Northern or Southern blotting, ISH (such as FISH or SISH) (for example of whole chromosomes), and by nucleic acid based arrays (*e.g.*, applying the nucleic acids to an array containing a plurality of target probes, such as a nucleic acid probe that can specifically hybridize to DIXDC1, or the reverse, wherein the sample nucleic acids are applied to an array, which is then contacted with target probes, such as a nucleic acid probe that can specifically hybridize to DIXDC1).

In some examples, the sample (or a fraction thereof) is present on a solid support. Solid supports bear the biological sample and permit the convenient detection of components (*e.g.*, proteins or nucleic acids) in the sample. Exemplary supports or substrates include microscope slides (*e.g.*, glass microscope slides or plastic microscope slides), coverslips (*e.g.*, glass coverslips or plastic coverslips), tissue culture dishes, multi-well plates, membranes (*e.g.*, nitrocellulose or polyvinylidene fluoride (PVDF)) or BIACORE™ chips.

#### **D. FAK and Src Inhibitors**

The disclosed diagnostic methods can be used in companion diagnostic assays with FAK and/or Src inhibitors. That is, the disclosed methods can be used to identify cancers that will be sensitive to a FAK and/or Src inhibitor. For example, cancers with decreased DIXDC1 expression or decreased phosphorylation (*e.g.*, at Ser 592), are cancers which are predicted to be responsive to FAK and/or Src inhibitors. In addition, cancers with decreased DIXDC1 expression or decreased phosphorylation, are cancers which are predicted to metastasize.

FAK inhibitors include a family of compounds that can be used to reduce (*e.g.*, down regulate) FAK activity, for example by reducing FAK nucleic acid and/or FAK expression.

Examples of such compounds include inhibitory nucleic acid molecules and small molecule inhibitors, such as those that specifically block FAK kinase activity. Examples include, but are not limited to: TAE226 (blocks the ATP binding site and inhibits FAK phosphorylation at both Y397 and Y861); PF-562,271 (blocks the ATP binding site of FAK and of protein-rich tyrosine kinase 2), PF-573,228, PF-04554878 (also known as VS-6063, defactinib), GSK2256098; molecules that target the Y397 site that is highly specific to FAK such as 1,2,4,5-benzenetetraamine tetrahydrochloride (Y15) (which in some examples is combined with gemcitabine). Other exemplary FAK inhibitors include but are not limited to, those shown below in Table 1, and VS-4718 from Verastem (Cambridge, MA). The structure of PND-1186 (VS-4718) is

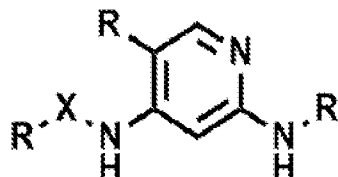


Table 1: FAK inhibitors (from Dunn *et al.*, *Anticancer Agents Med Chem.* Dec 2010; 10(10):722–734)

| Inhibitor Name          | Chemical Name                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Structure                     |          |
|-------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------|----------|
| PF 573,228 (PF-228)     | 6-(4-(3-(methylsulfonyl)benzylamino)-5-(trifluoromethyl)pyrimidin-2-ylamino)-5,4-dihydroquinolin-2(1H)-one                                                                                                                                                                                                                                                                                                                                                                                                                |                               | Pfizer   |
| PF562, 271-26 (PF-271)  | N-Methyl-N-(3-{{[2-(2-oxo-2,3-dihydro-1H-indol-5-ylamino)-5-trifluoromethylpyrimidin-4-ylamino]-methyl}-pyridin-2-yl)-methanesulfonamide                                                                                                                                                                                                                                                                                                                                                                                  |                               | Pfizer   |
| NVP-226                 | (2-[5-Chloro-2-[2-methoxy-4-(4-morpholinyl)phenylamino]pyrimidin-4-ylamino]-N-methylbenzamide                                                                                                                                                                                                                                                                                                                                                                                                                             |                               | Novartis |
| Y15 <sup>*</sup>        | 1,2,4,5-Benzenetetraamine tetrahydrochloride                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | <br><br>· 4HH <sub>2</sub> Cl |          |
| PND-1186 <sup>***</sup> | 2,4-diamino-pyridine-based;<br>X=bond or<br>(C1-C3)alkyl comprising<br>0-1 heteroatom selected<br>from the group consisting<br>of N, O, S(O) and S(O) <sub>2</sub> ,<br>the (C1-C3)alkyl is<br>substituted with 0-1 hydroxy,<br>halo, (C1-C3)alkoxy,<br>(C1-C3)alkylamino or<br>(C1-C3)2dialkylamino groups.<br>R1 and R2 are 5-12 membered<br>Mono-, bi- or polycyclic,<br>aromatic or partially aromatic rings.<br>R3 is a trifluoromethyl,<br>halo, nitro or cyano; salt, tautomer,<br>solvate, hydrate, or a prodrug. |                               |          |

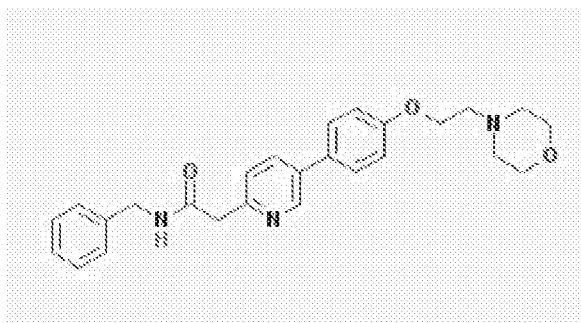
<sup>\*</sup>Reported by Golubovskaya *et al.*, *J. Med. Chem.*, V 51, p7405-7416, 2008, ref. in (15).

<sup>\*\*\*</sup>Reported by Tanjoni *et al.*, *Cancer Biol. Ther.*, 2010, V 15, 2010, ref in (17).

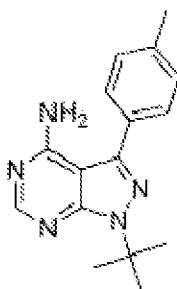
Proto-oncogene tyrosine-protein kinase Src also known as proto-oncogene c-Src or simply c-Src is a non-receptor protein tyrosine kinase protein that in humans is encoded by the SRC gene (OMIM: 190090).

Src inhibitors include a family of compounds that can be used to reduce (*e.g.*, down regulate) Src activity, for example by reducing Src nucleic acid and/or Src expression. Examples of such compounds include inhibitory nucleic acid molecules and small molecule inhibitors, such as those that specifically block Src kinase activity. Examples include, but are not limited to: Dasatinib (N-[2-Chloro-6-methylphenyl]-2-[[6-[4-(2-hydroxyethyl)-1-piperazinyl]-2-methyl-4-pyrimidinyl]amino]-5-thiazolecarboxamide), Saracatinib (AZD0530; N-(5-chlorobenzo[d][1,3]dioxol-4-yl)-7-(2-(4-methylpiperazin-1-yl)ethoxy)-5-(tetrahydro-2H-pyran-4-yloxy)quinazolin-4-amine), Bosutinib (SKI-606), KX2-391 (KX01), quercetin, SU6656, SU6657, PP2, PP1, WHI-P 154; NVP-BHG712, Lavendustin C, ON-01910, SD 1008, Indirubin Derivative E804, ZM-306416, PP121, Src/EGFR inhibitor (sequence derived from v-Src 137-157, Santa Cruz Biotechnology Catalog # sc-3050), and Dasatinib  $\beta$ -D-Glucuronide.

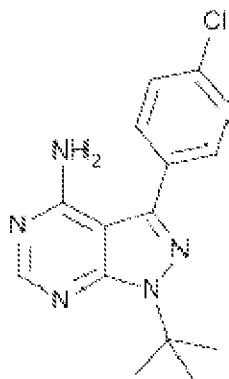
The structure of some of these compounds is shown below



KX2-391



PP1



PP2

In one example, a FAK or Src inhibitor is an inhibitory nucleic acid molecule, such as an antisense oligonucleotide, a siRNA, a microRNA (miRNA), a shRNA or a ribozyme. Such molecules can be used to decrease or eliminate FAK or Src gene expression. Any type of antisense compound that specifically targets and regulates expression of a FAK or Src nucleic acid is contemplated for use. An antisense compound is one which specifically hybridizes with and modulates expression of a target nucleic acid molecule (such as FAK or Src). These compounds can be introduced as single-stranded, double-stranded, circular, branched or hairpin compounds and can contain structural elements such as internal or terminal bulges or loops. Double-stranded antisense compounds can be two strands hybridized to form double-stranded compounds or a single strand with sufficient self-complementarity to allow for hybridization and formation of a fully or partially double-stranded compound. In some examples, an antisense FAK or Src oligonucleotide is a single stranded antisense compound, such that when the antisense oligonucleotide hybridizes to a FAK or Src mRNA, the duplex is recognized by RNaseH, resulting in cleavage of the mRNA. In other examples, a miRNA is a single-stranded RNA molecule of about 21-23 nucleotides that is at least partially complementary to an mRNA molecule that regulates gene expression through an RNAi pathway. In further examples, a shRNA is an RNA oligonucleotide that forms a tight hairpin, which is cleaved into siRNA. siRNA molecules are generally about 20-25 nucleotides in length and may have a two nucleotide overhang on the 3' ends, or may be blunt ended. Generally, one strand of a siRNA is at least partially complementary to a target nucleic acid. Antisense compounds specifically targeting a FAK or Src gene can be prepared by designing compounds that are complementary to a FAK or Src nucleotide sequence, such as a mRNA sequence. FAK or Src antisense compounds need not be 100% complementary to the FAK or Src nucleic acid molecule to

specifically hybridize and regulate expression of FAK or Src. For example, the antisense compound, or antisense strand of the compound if a double-stranded compound, can be at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99% or 100% complementary to a FAK or Src nucleic acid sequence. Methods of screening antisense compounds for specificity are well known (see, for example, U.S. Publication No. 2003-0228689). In addition, methods of designing, preparing and using inhibitory nucleic acid molecules are within the abilities of one of skill in the art. Furthermore, sequences for FAK or Src are publicly available.

## **E. Cancers**

Examples of cancers that can be analyzed using the disclosed methods include those having or suspected of having increased FAK and/or Src expression, and or decreased DIXDC1 expression and/or phosphorylation. In one example, the cancer is a solid tumor, such as a lung cancer (*e.g.*, NSCLC or mesothelioma), ovarian cancer, breast cancer, prostate cancer, pancreatic cancer, head and neck cancer, thyroid cancer, kidney cancer (*e.g.*, RCC), brain, liver cancer (*e.g.*, HCC), colorectal cancer (*e.g.*, colon cancer), cervical cancer, melanoma, neuroblastoma, osteosarcoma, or sarcoma.

## **F. Methods of Treatment**

The disclosed methods can further include identifying and/or selecting subjects for treatment with a FAK inhibitor and/or Src inhibitor. Such inhibitors can be used alone, or in combination with other chemotherapy or biotherapy. For example, if the disclosed methods indicate that the cancer is one that is likely sensitive to FAK inhibitor and/or Src inhibitor, the subject with that cancer can be selected for treatment with FAK inhibitor and/or Src inhibitor, while if the disclosed methods indicate that the cancer is not likely sensitive to FAK inhibitor and/or Src inhibitor, then the subject can be selected to abstain from such treatment. Additionally, the disclosed methods can further include administering one or more chemotherapies or biotherapies to the subject receiving the FAK inhibitor and/or Src inhibitor.

### **1. Exemplary chemotherapies and biologic therapies**

The FAK inhibitor and/or Src inhibitor administered to a subject (such as the therapies described herein) can be administered in combination with other chemotherapies and biotherapies (such as sequentially, or concurrently, or simultaneously). Such chemotherapies and biotherapies include therapeutic agents that when administered in therapeutically effective

amounts induce the desired response (*e.g.*, treatment of a cancer, for example by reducing the size or volume of the tumor, or reducing the size, volume or number of metastases, reducing the rate of growth of the tumor).

Chemotherapies and bio-therapies that can be used in combination with FAK and/or Src inhibitors, can include anti-neoplastic chemotherapeutic agents, antibiotics, alkylating agents and antioxidants, kinase inhibitors, and other agents such as antibodies. Particular examples of additional chemotherapeutic agents that can be used include alkylating agents, such as nitrogen mustards (for example, chlorambucil, chlormethine, cyclophosphamide, ifosfamide, and melphalan), nitrosoureas (for example, carmustine, fotemustine, lomustine, and streptozocin), platinum compounds (for example, carboplatin, cisplatin, oxaliplatin, and BBR3464), busulfan, dacarbazine, mechlorethamine, procarbazine, temozolomide, thiotepea, and uramustine; folic acid (for example, methotrexate, pemetrexed, and raltitrexed), purine (for example, cladribine, clofarabine, fludarabine, mercaptopurine, and tioguanine), pyrimidine (for example, capecitabine), cytarabine, fluorouracil, and gemcitabine; plant alkaloids, such as podophyllum (for example, etoposide, and teniposide); microtubule binding agents (such as paclitaxel, docetaxel, vinblastine, vindesine, vinorelbine (navelbine) vincristine, the epothilones, colchicine, dolastatin 15, nocodazole, podophyllotoxin, rhizoxin, and derivatives and analogs thereof), DNA intercalators or cross-linkers (such as cisplatin, carboplatin, oxaliplatin, mitomycins, such as mitomycin C, bleomycin, chlorambucil, cyclophosphamide, and derivatives and analogs thereof), DNA synthesis inhibitors (such as methotrexate, 5-fluoro-5'-deoxyuridine, 5-fluorouracil and analogs thereof); anthracycline family members (for example, daunorubicin, doxorubicin, epirubicin, idarubicin, mitoxantrone, and valrubicin); antimetabolites, such as cytotoxic/antitumor antibiotics, bleomycin, rifampicin, hydroxyurea, and mitomycin; topoisomerase inhibitors, such as topotecan and irinotecan; monoclonal antibodies, such as alemtuzumab, bevacizumab, cetuximab, gemtuzumab, rituximab, panitumumab, pertuzumab, and trastuzumab; photosensitizers, such as aminolevulinic acid, methyl aminolevulinate, porfimer sodium, and verteporfin, enzymes, enzyme inhibitors (such as camptothecin, etoposide, formestane, trichostatin and derivatives and analogs thereof), kinase inhibitors (such as imatinib, gefitinib, and erlotinib), gene regulators (such as raloxifene, 5-azacytidine, 5-aza-2'-deoxycytidine, tamoxifen, 4-hydroxytamoxifen, mifepristone and derivatives and analogs thereof); and other agents, such as alitretinoin, altretamine, amsacrine, anagrelide, arsenic trioxide, asparaginase, axitinib, bexarotene, bevacizumab, bortezomib, celecoxib, denileukin

diftitox, estramustine, hydroxycarbamide, lapatinib, pazopanib, pentostatin, masoprocol, mitotane, pegaspargase, tamoxifen, sorafenib, sunitinib, vemurafinib, vandetanib, and tretinoin..

Specific examples of chemotherapies and bio-therapies that can be used with the FAK inhibitor and/or Src inhibitor include but are not limited to one or more of the following: 5-fluorouracil (*e.g.*, Adrucil®, Efudex®, Fluoroplex®), Avastin® (bevacizumab), Camptosar® (Irinotecan Hydrochloride), capecitabine (*e.g.*, Xeloda®), oxaliplatin (*e.g.*, Eloxatin®), Erbitux® (cetuximab), leucovorin calcium, regorafenib, Stivarga® (Regorafenib), Vectibix® (Panitumumab), Wellcovorin® (Leucovorin Calcium), and Zaltrap® (Ziv-Aflibercept).

Examples of drug combinations include but are not limited to: GSK2256098 and trametinib; VS-6063 and paclitaxel; Dasatinib and Erlotinib; Dasatinib and Bevacizumab; as well as Dasatinib and Dacarbazine.

In one example, a FAK inhibitor and/or Src inhibitor, alone or in combination with one or more chemotherapies or bio-therapies, increases killing of cancer cells (or reduces their viability). Such killing need not result in 100% reduction of cancer cells; for example a therapy that results in reduction in the number of viable cancer cells by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 75%, at least 90%, or at least 95% (for example as compared to no treatment with the therapy) can be used in the methods provided herein. For example, the therapy can reduce the growth of cancer cells by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 75%, at least 90%, or at least 95% (for example as compared to no therapy).

In one example, a FAK inhibitor and/or Src inhibitor, alone or in combination with one or more chemotherapies or bio-therapies, decreases FAK and/or Src expression or activity. Such inhibition need not result in 100% reduction of FAK and/or Src expression or activity; for example chemotherapy that reduces FAK and/or Src expression or activity by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 75%, at least 90%, or at least 95% (for example as compared to no treatment with therapy) can be used in the methods provided herein. For example, the therapy may interfere with gene expression (transcription, processing, translation, post-translational modification).

In one example, a cancer bio-therapy includes or consists of an antibody, such as a humanized antibody. Such antibodies can be polyclonal, monoclonal, or chimeric antibodies. Examples of antibodies that can be used in combination with a FAK inhibitor and/or Src inhibitor include Cetuximab, Gemtuzumab, Ibritumomab tiuxetan, Nivolumab, Panitumumab,

Rituximab, Tositumomab and Trastuzumab). Methods of making antibodies specific for a particular target is routine. In some example, the therapeutic antibody is conjugated to a toxin.

## 2. Administration of Therapeutic Agents

5 In some examples, the disclosed methods include providing a therapeutically effective amount of one or more FAK and/or Src inhibitors, alone or in combination with therapeutically effective amounts of other chemotherapies or bio-therapies, to a subject having a cancer with decreased DIXDC1 expression and/or non-phosphorylated DIXDC1 (*e.g.*, lack of phosphorylation at Ser592). Methods and therapeutic dosages of such agents and treatments are  
10 known to those of ordinary skill in the art, and for example, can be determined by a skilled clinician. In some examples, the disclosed methods further include providing surgery and/or radiation therapy to the subject in combination with the FAK and/or Src inhibitors, alone or in combination with other chemotherapies or bio-therapies, (for example, sequentially, substantially simultaneously, or simultaneously). Administration can be accomplished by single  
15 or multiple doses. Methods and therapeutic dosages of such agents and treatments are known to those skilled in the art, and can be determined by a skilled clinician. The dose required will vary from subject to subject depending on the species, age, weight and general condition of the subject, the particular therapeutic agent being used and its mode of administration.

Therapeutic agents, including FAK and/or Src inhibitors, alone or in combination with  
20 other chemotherapies or bio-therapies, can be administered to a subject in need of treatment using any suitable means known in the art. Methods of administration include, but are not limited to, intradermal, transdermal, intramuscular, intraperitoneal, parenteral, intravenous, subcutaneous, vaginal, rectal, intranasal, inhalation, oral, or by gene gun. Intranasal administration refers to delivery of the compositions into the nose and nasal passages through  
25 one or both of the nares and can include delivery by a spraying mechanism or droplet mechanism, or through aerosolization of the therapeutic agent.

Administration of the therapeutic agents, including FAK and/or Src inhibitors, alone or in combination with other chemotherapies or bio-therapies, by inhalant can be through the nose or mouth via delivery by spraying or droplet mechanisms. Delivery can be directly to any area  
30 of the respiratory system via intubation. Parenteral administration is generally achieved by injection. Injectables can be prepared in conventional forms, either as liquid solutions or suspensions, solid forms suitable for solution or suspension in liquid prior to injection, or as

emulsions. Injection solutions and suspensions can be prepared from sterile powders, granules, and tablets. Administration can be systemic or local.

Therapeutic agents, including FAK and/or Src inhibitors, alone or in combination with other chemotherapies or bio-therapies, can be administered in any suitable manner, for example with pharmaceutically acceptable carriers. Pharmaceutically acceptable carriers are determined in part by the particular composition being administered, as well as by the particular method used to administer the composition. Accordingly, there is a wide variety of suitable formulations of pharmaceutical compositions of the present disclosure. The pharmaceutically acceptable carriers (vehicles) useful in this disclosure are conventional. *Remington's Pharmaceutical Sciences*, by E. W. Martin, Mack Publishing Co., Easton, PA, 15th Edition (1975), describes compositions and formulations suitable for pharmaceutical delivery of one or more therapeutic agents

Preparations for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's, or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers (such as those based on Ringer's dextrose), and the like. Preservatives and other additives may also be present such as, for example, antimicrobials, anti-oxidants, chelating agents, and inert gases and the like.

Formulations for topical administration can include ointments, lotions, creams, gels, drops, suppositories, sprays, liquids and powders. Conventional pharmaceutical carriers, aqueous, powder or oily bases, thickeners and the like may be necessary or desirable.

Therapeutic agents, including FAK and/or Src inhibitors, alone or in combination with other chemotherapies or bio-therapies, for oral administration include powders or granules, suspensions or solutions in water or non-aqueous media, capsules, sachets, or tablets. Thickeners, flavorings, diluents, emulsifiers, dispersing aids or binders may be desirable.

Therapeutic agents, including FAK and/or Src inhibitors, alone or in combination with other chemotherapies or bio-therapies, can be administered as a pharmaceutically acceptable acid- or base-addition salt, formed by reaction with inorganic acids such as hydrochloric acid, hydrobromic acid, perchloric acid, nitric acid, thiocyanic acid, sulfuric acid, and phosphoric acid, and organic acids such as formic acid, acetic acid, propionic acid, glycolic acid, lactic acid,

pyruvic acid, oxalic acid, malonic acid, succinic acid, maleic acid, and fumaric acid, or by reaction with an inorganic base such as sodium hydroxide, ammonium hydroxide, potassium hydroxide, and organic bases such as mono-, di-, trialkyl and aryl amines and substituted ethanolamines.

### 5                   3. Exemplary Doses

Methods and therapeutic dosages of such agents are known to those skilled in the art, and can be determined by a skilled clinician. Other therapeutic agents, for example anti-tumor agents, that may or may not fall under one or more of the classifications above, also are suitable for administration in combination with the described specific binding agents. Selection and  
10 therapeutic dosages of such agents are known to those skilled in the art, and can be determined by a skilled clinician.

In some examples, the dose of a FAK or Src inhibitory nucleic acid (such as an antisense molecule, siRNA, shRNA, or miRNA) is about 1 mg to about 1000 mg, about 10 mg to about 500 mg, or about 50 mg to about 100 mg. In some examples, the dose of antisense compound is  
15 about 1 mg, about 10 mg, about 50 mg, about 100 mg, about 250 mg, about 500 mg or about 1000 mg. In some embodiments, the dose of an inhibitory nucleic acid is about 1.0 mg/kg to about 100 mg/kg, or about 5.0 mg/kg to about 500 mg/kg, about 10 mg/kg to about 100 mg/kg, or about 25 to about 50 mg/kg. In some examples, the dose of a inhibitory nucleic acid is about 1.0 mg/kg, about 5 mg/kg, about 10 mg/kg, about 12.5 mg/kg, about 15 mg/kg, about 20 mg/kg,  
20 about 25 mg/kg, about 30 mg/kg, about 35 mg/kg, about 40 mg/kg, about 45 mg/kg, about 50 mg/kg, about 60 mg/kg, about 70 mg/kg, about 80 mg/kg or about 100 mg/kg. It will be appreciated that these dosages are examples only, and an appropriate dose can be determined by one of ordinary skill in the art using only routine experimentation.

In some embodiments, the dose of an inhibitory FAK or Src antibody or antibody  
25 conjugate is about 1 mg/kg to about 25 mg/kg, such as about 2 mg/kg to about 15 mg/kg, about 2 mg/kg to about 10 mg/kg, or about 2 mg/kg to about 8 mg/kg. In some examples, the dose of antibody is about 1 mg/kg, about 2 mg/kg, about 4 mg/kg, about 5 mg/kg, about 6 mg/kg, about 8 mg/kg, about 10 mg/kg, about 15 mg/kg, about 20 mg/kg, or about 25 mg/kg. In other embodiments, the dose of antibody is about 50 mg/m<sup>2</sup> to about 500 mg/m<sup>2</sup>, such as about 50  
30 mg/m<sup>2</sup> to about 400 mg/m<sup>2</sup>, about 100 mg/m<sup>2</sup> to about 400 mg/m<sup>2</sup>, or about 250 mg/m<sup>2</sup> to about 400 mg/m<sup>2</sup>. In some examples, the dose is about 50 mg/m<sup>2</sup>, about 100 mg/m<sup>2</sup>, about 150 mg/m<sup>2</sup>, about 200 mg/m<sup>2</sup>, about 250 mg/m<sup>2</sup>, about 300 mg/m<sup>2</sup>, about 400 mg/m<sup>2</sup>, or about 500 mg/m<sup>2</sup>.

It will be appreciated that these dosages are examples only, and an appropriate dose can be determined by one of ordinary skill in the art using only routine experimentation.

In some embodiments, the dose of an inhibitory FAK or Src small molecule is at least 1 mg/kg, at least 10 mg/kg, at least 50 mg/kg, at least 100 mg/kg, at least 500 mg/kg, or at least 1000 mg/kg, such as about 1 mg/kg to about 1000 mg/kg, about 10 mg/kg to about 200 mg/kg, about 10 mg/kg to about 100 mg/kg, or about 100 mg/kg to about 500 mg/kg. In some examples, the dose of an inhibitory FAK or Src small molecule is about 1 mg/kg, about 2 mg/kg, about 4 mg/kg, about 5 mg/kg, about 6 mg/kg, about 8 mg/kg, about 10 mg/kg, about 15 mg/kg, about 20 mg/kg, or about 25 mg/kg. In other embodiments, the dose of an inhibitory FAK or Src small molecule is at least 1 mg/m<sup>2</sup>, at least 10 mg/m<sup>2</sup>, at least 50 mg/m<sup>2</sup>, at least 100 mg/m<sup>2</sup>, or at least 500 mg/m<sup>2</sup>, such as about 50 mg/m<sup>2</sup> to about 500 mg/m<sup>2</sup>, such as about 50 mg/m<sup>2</sup> to about 400 mg/m<sup>2</sup>, about 100 mg/m<sup>2</sup> to about 400 mg/m<sup>2</sup>, or about 250 mg/m<sup>2</sup> to about 400 mg/m<sup>2</sup>. In some examples, the dose is about 50 mg/m<sup>2</sup>, about 100 mg/m<sup>2</sup>, about 150 mg/m<sup>2</sup>, about 200 mg/m<sup>2</sup>, about 250 mg/m<sup>2</sup>, about 300 mg/m<sup>2</sup>, about 400 mg/m<sup>2</sup>, or about 500 mg/m<sup>2</sup>.

It will be appreciated that these dosages are examples only, and an appropriate dose can be determined by one of ordinary skill in the art.

In one example, the FAK inhibitor is GSK2256098 and is administered orally at a dose of 250 mg. For example, GSK2256098 250 mg can be administered 30 minutes after a light meal with approximately 240 milliliter of water. In some example it is combined with 0.5 mg oral trametinib under fasting conditions two hours after a meal. In one example the FAK inhibitor is Y15 and is administered by injection (*e.g.*, ip) at a dose of 125 mg/kg. In another example, the FAK inhibitor is PF00562271 (also known as PF-562,271 and VS-6062) and is administered orally twice a day at a dose of 125 mg with food. In another example, the Src inhibitor is dasatinib and is administered orally at a dose of 70 mg twice a day.

### Exemplary Computer Systems

FIG. 14 illustrates a generalized example of a suitable computing system (100) in which the disclosed methods may be implemented. The computing system (100) is not intended to suggest any limitation as to scope of use or functionality, as embodiments of the methods may be implemented in diverse general-purpose or special-purpose computing systems. Thus, the computing system can be any of a variety of types of computing system (*e.g.*, desktop computer, laptop computer, tablet or slate computer, smartphone, etc.).

With reference to FIG. 14, the computing system (100) includes one or more processing units (110, 115) and memory (120, 125). In FIG. 14, this most basic configuration (130) is included within a dashed line. The processing units (110, 115) execute computer-executable instructions. A processing unit can be a general-purpose central processing unit (CPU), processor in an application-specific integrated circuit (ASIC) or any other type of processor. In a multi-processing system, multiple processing units execute computer-executable instructions to increase processing power. For example, FIG. 14 shows a central processing unit (110) as well as a graphics processing unit or co-processing unit (115). The tangible memory (120, 125) may be volatile memory (*e.g.*, registers, cache, RAM), non-volatile memory (*e.g.*, ROM, EEPROM, flash memory, etc.), or some combination of the two, accessible by the processing unit(s). The memory (120, 125) stores software (180) implementing one or more methods for determining if a cancer is sensitive to a FAK inhibitor and/or a Src inhibitor, or whether the cancer is likely to metastasize, in the form of computer-executable instructions suitable for execution by the processing unit(s).

A computing system may have additional features. For example, the computing system (100) includes storage (140), one or more input devices (150), one or more output devices (160), and one or more communication connections (170). An interconnection mechanism (not shown) such as a bus, controller, or network interconnects the components of the computing system (100). Typically, operating system software (not shown) provides an operating environment for other software executing in the computing system (100), and coordinates activities of the components of the computing system (100). The other software can include common applications (*e.g.*, email applications, calendars, word processors and other productivity software, Web browsers, messaging applications).

The tangible storage (140) may be removable or non-removable, and includes magnetic disks, magnetic tapes or cassettes, CD-ROMs, DVDs, thumb drives, or any other medium which can be used to store information in a non-transitory way and which can be accessed within the computing system (100). The storage (140) stores instructions for the software (180) implementing one or more methods for determining if a cancer is sensitive to a FAK inhibitor and/or a Src inhibitor, or whether the cancer is likely to metastasize.

The input device(s) (150) may include one or more audio input devices (*e.g.*, a microphone adapted to capture audio or similar device that accepts audio input in analog or digital form). The input device(s) (150) may also include a touch input device such as a keyboard, mouse, pen, or trackball, a touchscreen, a scanning device, or another device that

provides input to the computing system (100). For determining if a cancer is sensitive to a FAK inhibitor and/or a Src inhibitor, or whether the cancer is likely to metastasize the input device(s) (150) may include a reader adapted to receive values that quantify expression levels of DIXDC1 proteins, DIXDC1 protein phosphorylation, and/or DIXDC1 nucleic acid expression. The input device(s) (150) may further include a CD-ROM or CD-RW that reads values into the computing system (100). The output device(s) (160) may include one or more audio output devices (*e.g.*, one or more speakers). The output device(s) (160) may also include a display, touchscreen, printer, CD-writer, or another device that provides output from the computing system (100). For determining if a cancer is sensitive to a FAK inhibitor and/or a Src inhibitor, or whether the cancer is likely to metastasize, the output device(s) (160) may indicate results of the prediction/prognosis visually on a screen or printed report, audibly, or in some other way.

The communication connection(s) (170) enable communication over a communication medium to another computing entity. The communication medium conveys information such as computer-executable instructions, values that quantify protein expression level, clinical information or variables, or other data in a modulated data signal. A modulated data signal is a signal that has one or more of its characteristics set or changed in such a manner as to encode information in the signal. By way of example, and not limitation, communication media can use an electrical, optical, RF, or other carrier.

The methods can be described in the general context of computer-readable media. Computer-readable media are any available tangible media that can be accessed within a computing environment. By way of example, and not limitation, with the computing system (100), computer-readable media include memory (120, 125), storage (140), and combinations of any of the above.

The innovations can be described in the general context of computer-executable instructions, such as those included in program modules, being executed in a computing system on a target real or virtual processor. Generally, program modules include routines, programs, libraries, objects, classes, components, data structures, etc. that perform particular tasks or implement particular abstract data types. The functionality of the program modules may be combined or split between program modules as desired in various embodiments. Computer-executable instructions for program modules may be executed within a local or distributed computing system. The computer-executable instructions can be executable code that results from compilation into instructions directly executable on a processor, interpreted code such as a script, byte codes or other intermediate representation of software instructions, or logic or

formulas adapted for execution in a spreadsheet or other tool. The computer-executable instructions can be specified in any of various computer languages, such as C, C++, Java, Perl, JavaScript, Adobe Flash, or any other suitable language.

The terms “system” and “device” are used interchangeably herein. Unless the context clearly indicates otherwise, neither term implies any limitation on a type of computing system or computing device. In general, a computing system or device can be local or distributed, and can include any combination of special-purpose hardware and/or general-purpose hardware with software implementing the functionality described herein.

### 1. Exemplary Assay Outputs

Following the measurement of the DIXDC1 expression and/or phosphorylation level in the cancer sample (and in some example a control sample), the assay results, findings, diagnoses, predictions and/or treatment recommendations are typically recorded and communicated to technicians, physicians and/or patients, for example. In certain embodiments, computers will be used to communicate such information to interested parties, such as, patients and/or the attending physicians. Thus, in some examples the methods include producing or generating a report, such as a report with an icon, for example an icon indicating the patient’s risk of metastasis and/or whether the cancer will respond to an FAK inhibitor and/or Src inhibitor, in a graphical or pictorial manner. Based on the measurement, a FAK inhibitor and/or Src inhibitor administered to a subject, if the output indicates that the cancer will be sensitive to a FAK inhibitor and/or Src inhibitor.

In one embodiment, a diagnosis, prognosis, prediction and/or treatment recommendation based on DIXDC1 expression and/or phosphorylation in a cancer sample from a test subject is communicated to interested parties as soon as possible after the assay is completed and the diagnosis and/or prediction is generated. The results and/or related information may be communicated to the subject by the subject's treating physician. Alternatively, the results may be communicated directly to interested parties by any means of communication, including writing, such as by providing a written report, electronic forms of communication, such as email, or telephone. In particular examples, the output is an icon, which communicates the patient’s risk of metastasis in a graphical or pictorial manner. Thus, if the patient has a low risk of metastasis, such as a risk of recurrence of less than 20% (for example, less than 20%, 10%, 5%, 4%, 3%, 2%, or 1%) the icon could be a positive indicator (including, but not limited to a plus sign, check mark, or a thumbs-up graphic). If the patient has a high risk of metastasis, such as a risk of 20% or more (for example, 25%, 30%, 35%, 40%, 50%, 75%, or more) the icon

could be a negative indicator (including, but not limited to a minus sign, X mark, or a thumbs-down graphic). In particular examples, the output is an icon, which communicates whether the cancer is responsive to a FAK and/or Src inhibitor in a graphical or pictorial manner. Thus, if the cancer is predicted to respond to the FAK and/or Src inhibitor the icon could be a positive indicator (including, but not limited to a plus sign, check mark, or a thumbs-up graphic). If the cancer is predicted to be one that will not respond to the FAK and/or Src inhibitor the icon could be a negative indicator (including, but not limited to a minus sign, X mark, or a thumbs-down graphic).

Communication may be facilitated by use of a suitably programmed computer, such as in case of email communications. In certain embodiments, the communication containing results of a diagnostic test and/or conclusions drawn from and/or treatment recommendations based on the test, may be generated and delivered automatically to interested parties using a combination of computer hardware and software which will be familiar to artisans skilled in telecommunications. One example of a healthcare-oriented communications system is described in U.S. Pat. No. 6,283,761; however, the present disclosure is not limited to methods which utilize this particular communications system.

In certain embodiments of the methods of the disclosure, all or some of the method steps, including the assaying of samples, prognosis of cancer (*e.g.*, likelihood of metastasis), determining whether the cancer is sensitive/responsive to a FAK and/or Src inhibitor), and communicating of assay results or prognosis, may be carried out in diverse (*e.g.*, foreign) jurisdictions.

### **Computer-Readable Media**

Any of the computer-readable media herein can be non-transitory (*e.g.*, memory, magnetic storage, optical storage, or the like).

Any of the storing actions described herein can be implemented by storing in one or more computer-readable media (*e.g.*, computer-readable storage media or other tangible media).

Any of the things described as stored can be stored in one or more computer-readable media (*e.g.*, computer-readable storage media or other tangible media).

Any of the methods described herein can be implemented by computer-executable instructions in (*e.g.*, encoded on) one or more computer-readable media (*e.g.*, computer-readable storage media or other tangible media). Such instructions can cause a computer to perform the

method. The technologies described herein can be implemented in a variety of programming languages.

Any of the steps of the methods described herein can be implemented by computer-executable instructions stored in one or more computer-readable storage devices (*e.g.*, memory, magnetic storage, optical storage, or the like). Such instructions can cause a computer to perform the method.

The present disclosure is illustrated by the following non-limiting Examples.

## Example 1

### Materials and Methods

This example provides the materials and methods used to generate the results obtained in the subsequent examples.

#### Plasmid Constructs

The cDNA for human DIXDC1 was cloned from a cDNA library prepared from U2OS cells and sequence verified to match the sequence of the long isoform of human DIXDC1 (NM\_001037954.2). DIXDC1 cDNA was subcloned into pENTR4-FLAG (Addgene #17423) or pENTR4-myc (edited from pENTR4-FLAG) to create FLAG or myc tagged ENTR clones, respectively. ENTR clones for human LKB1 and LYK5 (STRADalpha) were obtained from Invitrogen (IOH21169, IOH45129 respectively). Human Snail cDNA (Addgene #16218) and Slug cDNA (Addgene #31698) were PCR amplified and subcloned to pDONR221 with BP Clonase (Invitrogen). Site-directed mutagenesis was performed using QuikChange II XL (Stratagene) according to the manufacturer's instructions. All ENTR/DONR clones were sequence verified to ensure no additional mutations. To create mammalian expression vectors, ENTR clones were recombined into DEST vectors using LR clonase (Invitrogen). DEST vectors used in this study include: pQCXIB (Addgene #17400), pQCXIN (Addgene #17395), pLentiCMV/TO DEST (Addgene #17293), pcDNA3 N-term myc DEST, pcDNA3 N-term FLAG DEST, pBabe-Hygro DEST. pEBG2T-hMARK1 WT and T215A(KD) were a gift from Dario Alessi (Dundee, UK).

## Antibodies

Cell Signaling Technology antibodies used: Wnt5a/b (#2530), Snail (#3879), Slug (#9585), ZEB1 (#3396), LKB1 (#3047), P-ACC S79 (#3661), Total ACC (#4190), Axin2 (#2151),  
 5 MARK2 (#9118), MARK3 (#9311), AMPKalpha (#2532), P-ULK1 S555 (#5869), Nuak1 (#4458), SIK2 (#6919), GST (#2622), myc-tag (#2272), P-Src family Y416 (#2113), Src (#2109), P-Paxillin Y118 (#2541), Pathscan I for P-ERK1/2 and P-Akt S473 (#5301), Total ERK1/2 (#4695), P-S6K (#9234), HA-tag (#3724), P-MEK1/2 (#9154), P-p90RSK S380 (#11989), P-FAK Y925 (#3284) and Phospho-DIXDC1 S592. Epitomics antibodies used:  
 10 Phospho-FAK Y397 (#2211-1), Phospho-FAK Y576/577 (#2183-1), Total FAK1 (#2146-1), Zyxin (#3586-1). BD Transduction Labs antibodies used: Paxillin (P13520). Sigma antibodies used: Actin (A5441), Flag polyclonal (F7425). Protein Tech antibodies used: MARK1 (21552-1-AP), MARK2 (15492-1-AP). Millipore antibodies used: ZEB2 (ABT332), MARK4 (07-699). Abcam antibodies used Twist (ab50887), IRSp53 (ab15697). CLASP2 antibody was from Santa  
 15 Cruz Biotechnology (sc-98440). DIXDC1 total antibody was from R&D Systems (AF5599).

## Cell lines

U2OS, H157, H1299, SJSA, Panc-1, MiaPaca-2, and 293T were from ATCC. Littermate-derived *Lkb1*<sup>+/+</sup> and *Lkb1*<sup>-/-</sup> MEFs were described previously (Shaw et al., Proc Natl Acad Sci U  
 20 S A 101, 3329-3335, 2004). T<sub>nonMet</sub> and T<sub>Met</sub> primary mouse lung cancer cell have been described previously (Winslow et al., Nature 473, 101-104, 2011). Specific lines used include 368T1, 394T4, 802T4, 565T2, and 393T5. All cells were maintained in DMEM (Mediatech, Manassas, VA) supplemented with 10% fetal bovine serum (Hyclone, Thermo Scientific) and cultured at 37°C in 10% CO<sub>2</sub>.

25

## Transfection and Virus Production

For transient expression of proteins and packaging of virus, HEK293T cells were transfected with DNA or short hairpin RNA (shRNA) plasmids using Lipofectamine 2000 (Invitrogen, Carlsbad, CA) following the manufacturer's protocol. Lentiviral shRNA transduction and  
 30 retroviral gene expression was performed as described previously (Gwinn et al., Mol Cell 30, 214-226, 2008). Briefly, for retroviral infection, the pQCXIN or pQCXIB constructs were transfected along with the pCL-Ampho packaging plasmid into growing, low-passage HEK293T cells. Virus containing supernatants were collected 48 hours after transfection, centrifuged, and

syringe-filtered to eliminate residual HEK293T cells. Viral transduction was performed for 16 hours in the presence of polybrene (8 ug/ml). Selective antibiotic was applied 24 hours after transduction. For lentiviral shRNA knockdown, hairpin sequences were either used as supplied in the pLKO vector backbone (TRC Collection, Sigma) or subcloned into the pENTR/pSUPER+ or pENTR/pTER+ entry cassettes (Addgene #17338 or Addgene #17453, respectively). shRNA sequences driven by the H1 promoter were then recombined in the pLentiX2 series of viral vectors (Addgene #17296 and #17390). The shRNA-containing vectors were transfected into HEK293T cells with lentiviral packaging plasmids vsvg, GAG/pol, and REV using Lipofectamine 2000. Viruses were collected 48 hours after transfection, and target cells were infected for 16 hours to achieve depletion of endogenous proteins.

### Animal Studies

Tumor lysates from *Kras*<sup>G12D</sup>;*Lkb1*<sup>+/+</sup> and *Kras*<sup>G12D</sup>;*Lkb1*<sup>-/-</sup> primary lung tumors were generated from *Lox-Stop-Lox Kras*<sup>G12D</sup> (Jackson et al., Cancer Res. 65, 10280–10288, 2005) and *Lkb1*<sup>fllox/fllox</sup> mice (Shaw et al., Science 310, 1642–1646, 2005) as previously described (Shackelford et al., Cancer Cell 23, 143–158, 2012). Mice used for syngeneic mouse transplantation assays were 129/B16 F<sub>1</sub> hybrid mice (Jackson Laboratories). Mice were 7-9 weeks old at time of injections. For lung colonization assays, 7 x 10<sup>4</sup> cells resuspended in 200 µl PBS were injected through the lateral tail vein. Mice were sacrificed 3 weeks after injection and lungs were harvested for tissue processing.

### RNA extraction and qPCR analysis

Total RNA was extracted using Trizol (Invitrogen) according to the manufacturer's protocol. RNA was treated with DNase I (New England Biolabs, MA). cDNA amplification was performed using the High Capacity Reverse Transcription Kit (Applied Biosystems). qPCR primers were designed using Primer 3 MIT software and sequences were chosen that spanned exon junctions. Sequences are listed below 5' to 3'.

#### Human

|                      |                          |                 |
|----------------------|--------------------------|-----------------|
| <b><i>hSNAI1</i></b> | F-ggccctggctgctacaaggc   | (SEQ ID NO: 23) |
|                      | R-ctcgagggtcagcggggaca   | (SEQ ID NO: 24) |
| <b><i>hACTB</i></b>  | F-cgggaaatcggtgcgtgacatt | (SEQ ID NO: 25) |

|    |                      |                         |                 |
|----|----------------------|-------------------------|-----------------|
|    |                      | R-tgatctccttctgcatcctgt | (SEQ ID NO: 26) |
|    | <b><i>hWNT5A</i></b> | F-tggctttggccatatttttc  | (SEQ ID NO: 27) |
|    |                      | R-ccgatgtactgcatgtggtc  | (SEQ ID NO: 28) |
| 5  | <u><i>Mouse</i></u>  |                         |                 |
|    | <b><i>mWnt5A</i></b> | F-ctggcaggcctttctcaagg  | (SEQ ID NO: 29) |
|    |                      | R-gtctctcggctgcctatttg  | (SEQ ID NO: 30) |
| 10 | <b><i>mActb</i></b>  | F-tgttaccaactgggacgaca  | (SEQ ID NO: 31) |
|    |                      | R-tctcagctgtggtggtgaag  | (SEQ ID NO: 32) |
|    | <b><i>mWnt5B</i></b> | F-tcagagagtgccaacaccag  | (SEQ ID NO: 33) |
|    |                      | R-agccgtactccacgttgtct  | (SEQ ID NO: 34) |
| 15 |                      |                         |                 |
|    | <b><i>mSnai1</i></b> | F-cttgtgtctgcacgacctgt  | (SEQ ID NO: 35) |
|    |                      | R-cttcacatccgagtgggttt  | (SEQ ID NO: 36) |

## 20 **shRNA sequences**

The pLKO-based shRNA plasmids were obtained from the TRC Collection (Sigma-Aldrich, St. Louis, MO). Gene names and TRC numbers are listed below.

|    |               |                                                            |    |
|----|---------------|------------------------------------------------------------|----|
|    | hDIXDC1       | TRCN0000134516                                             | X2 |
|    | mDixdc1 (#57) | TRCN0000097257                                             | X2 |
| 25 | mDixdc1 (#58) | TRCN0000097258                                             | X2 |
|    | hMARK1        | TRCN0000006332                                             | X2 |
|    | hMARK4        | TRCN0000007158                                             |    |
|    | mLkb1         | TRCN0000024148                                             | X2 |
|    | hPTK2 #10     | TRCN0000196310                                             |    |
| 30 | hPTK2 #37     | David Schlaepfer Lab (Lim et al., Mol Cell 29, 9-22, 2008) |    |

“X2” designation above indicates that hairpin sequences were ordered from IDT DNA Services (San Diego, CA) and subsequently subcloned into pENTR/pTER+ or pENTR/pSuper+ according to Eric Campeau Lab cloning protocol ([ericcampeau.com/manuals.html](http://ericcampeau.com/manuals.html)). E NTR-

shRNA constructs were recombined into the pLentiX2 series of plasmids also obtained from the Eric Campeau lab through Addgene (Campeau et al., PLoS One 4, e6529, 2009). shRNA targeting GFP was also obtained from Eric Campeau lab through Addgene (shEGFP#1 sequence Addgene #17470).

5

Custom shRNA were designed using Block-IT RNAi Designer (Invitrogen). Custom sequence for hDIXDC1 3'utr hairpin listed below:

shDIXDC1 Human 3'utr

10 5'-GCATCATTCCTGTGTGTTAGTGTGCTGTCCTAACACACAGGAATGATGCTTTTT  
(SEQ ID NO: 37)

### Reagents and Chemicals

|                  |                    |               |                        |
|------------------|--------------------|---------------|------------------------|
| PF-573228        | FAK Inhibitor      | 1 $\mu$ M     | Tocris 3239            |
| Collagen I (rat) | ECM protein        | 50 $\mu$ g/ml | BD Biosciences 354236  |
| Mitomycin C      | Growth arrest      | 10 $\mu$ g/ml | Sigma M4287            |
| U0216            | MEK Inhibitor      | 20 $\mu$ M    | CST 9903               |
| Torin1           | TOR inhibitor      | 250 nM        | Dr. David Sabatini Lab |
| Akt Inh V        | Akt1/2/3 Inhibitor | 20 $\mu$ M    | EMD/Millipore 124012   |
| Wortmannin       | PI3K Inhibitor     | 2 $\mu$ M     | Sigma W1628            |
| Y27632           | ROCK Inhibitor     | 10 $\mu$ M    | Tocris 1254            |

### 15 Migration and Invasion Assays

For migration analysis, scratch wounds were made to a confluent monolayer to stimulate direction cell migration. Briefly, cells were plated to confluency in 35mm plates and then scratched with P10 pipette tips. Monolayers were washed 4 times with PBS and then fresh media was added. Cells were allowed to migrate to close the wound in the presence of

20 Mitomycin C (10  $\mu$ g/ml) for 12-16 hrs. Cells were then fixed with 4% paraformaldehyde/PBS for 10 min at room temperature and subsequently stained with rhodamine-labelled Phalloidin (Invitrogen/Molecular Probes; 1:200) to distinguish cell boundaries (Simpson et al., Nat Cell Biol 10, 1027-1038, 2008).

For invasion assays, cells were plated into matrigel-coated transwell inserts and invasion assays were carried out according to the manufacturer's protocol. Briefly,  $10^5$  cells were plated in serum free media in the upper chambers of equilibrated matrigel-coated transwell inserts. Inserts were then placed into wells containing media with 10% FBS as a chemoattractant. Cells were then incubated for 24-48 hours to allow invasion. Inserts were then fixed and stained with crystal violet stain solution (0.5% crystal violet, 0.5% formal saline, 145 mM NaCl, 50% EtOH) and cells remaining in the upper chamber were removed with a cotton swab. Crystal violet dye was eluted from invaded cells on the bottom of the insert in 33% acetic acid and absorbance was quantified using a microplate reader with a 562nm ref 650nm absorbance. In FIG. 8B, membranes were first cleared of cells remaining in the upper chamber, then excised from the transwell insert and stained with DAPI. Number of migrated cells/field was counted for 8 fields/membrane.

### Immunofluorescence

Cells were plated on collagen-coated glass coverslips for the indicated timepoints and then fixed in 4% paraformaldehyde/PBS for 10 minutes at room temperature followed by permeabilization in 0.2% Triton X-100/PBS for 10 minutes. Cells were then blocked with 3% BSA/PBST and subsequently incubated with primary antibodies diluted in blocking buffer at 4°C overnight. Primary antibodies used: mouse anti-Paxillin (BD Biosciences; 1:250), rabbit anti-Paxillin N-term (Epitomics; 1:2000), M2 flag (Sigma-Aldrich; 1:2000), rabbit anti-Zyxin (Epitomics; 1:1000), rabbit anti-P-FAK Y397 (Invitrogen; 1:100). F-Actin was stained using rhodamine labeled Phalloidin (Invitrogen/Molecular Probes; 1:200). Secondary antibodies, donkey anti-mouse AlexaFluor 488 and donkey anti-rabbit AlexaFluor 568, were used at 1:2000 diluted in blocking buffer. Cells were counterstained with DAPI and mounted using Fluoromount G (Southern Biotech, Birmingham, AL).

All confocal microscopy was performed on an LSM 710 spectral confocal microscope mounted on an inverted Axio Observer Z1 frame (Carl Zeiss, Jena, Germany) as previously described (Egan et al., *Science* 331, 456-461, 2011; Mihaylova et al., *Cell* 145, 607-621, 2011).

Colocalization analysis was performed using ImageJ with the PSC colocalization plugin (French et al., *Nature protocols* 3, 619-628, 2008). Briefly, region of interests were created around peripheral adhesions of immunostained cells (4-6 regions/cell). Pearson's coefficient of colocalization was then determined using the PSC plugin. 10-20 cells were analyzed per

condition and box-whisker plots were generated representing the Pearson's coefficients across each sample.

### ***In vitro* Kinase Assay**

5 Recombinant active MARK1 and MARK4 were purchased from Sigma (M8447 and SRP5046, respectively) and kinase assays were carried out according to manufacturer's protocol. Briefly, WT or S592A myc-hDIXDC1 was immunoprecipitated from transiently transfected HEK293T cells. Immunoprecipitates were washed three times in IP buffer followed by two washes in kinase buffer (25mM MOPS pH 7.2, 12.5mM glycerol-2-phosphate, 25 mM MgCl<sub>2</sub>, 5 mM  
10 EGTA, and 2mM EDTA). Immunoprecipitates were then incubated with or without 15 µl of active MARK1 or MARK4 under recommended reaction conditions in 100 µl final volume, using cold ATP. Reactions were then boiled and run out on SDS-PAGE gel. Phosphorylation of DIXDC1 was detected with a P-DIXDC1 S592 specific antibody (CST).

### **15 Peptide library screening**

Peptide mixtures (50 µM) were incubated 2 hours at 30 °C in multiwell plates in the presence of the indicated kinase in 50 mM HEPES, pH 7.4, 25 mM MgCl<sub>2</sub>, 0.25 mM DTT, 12.5 mM β-glycerophosphate, 5 mM EGTA, 2 mM EDTA, 0.1% Tween 20, and 50 µM ATP (0.03 µCi/ml). Aliquots of each reaction were transferred to streptavidin-coated membrane (Promega), which  
20 was quenched, washed and dried as described previously (Hutti et al., Nat Methods 1, 27-29, 2004). Membranes were exposed to a phosphor imager screen to quantify radiolabel incorporation. Heat maps were generated using Microsoft Excel.

### **MARK substrate screen**

25 To identify substrates of the MARK kinases, the optimal substrate motif for all four MARK family members was determined using arrayed positional scanning peptide libraries. First, the human proteome was bioinformatically queried to identify proteins containing candidate phosphorylation sites matching the optimal MARK substrate motif using the ScanSite (scansite3.mit.edu) and ProSite (prosite.expasy.org) databases. Proteins with consensus motif  
30 matches were further queried using the PhosphoSite (www.phosphosite.org) database for existing evidence of *in vivo* phosphorylation. Site conservation was examined across evolution using UniProt (www.uniprot.org) and NCBI PubMed Protein databases. Additional stringency was achieved by re-performing the search on other eukaryotic genomes allowed us to narrow the

initial ~1000 candidate human substrates to ~300 highly conserved candidate substrate sites. Given that control of the cytoskeleton is a major conserved function of the MARK/Par1 kinase family, 25 candidate MARK substrates bearing conserved phosphorylation motifs that contained a Pfam domain indicative of a role in cytoskeleton or were classified by gene ontology (GO) as having a role in cytoskeletal signaling were examined.

To test of whether these 25 candidates may be regulated by the MARK kinases, two approaches were used. Substrate phosphorylation was examined by exogenous gain and loss of function of MARK1 kinase activity either indirectly through regulated binding of the substrate to the phospho-dependent binding scaffold protein 14-3-3, or directly by reactivity with phospho-motif antibodies generated against peptides overlapping with the optimal MARK substrate motif (CST #5759, #9606, #9601, and #4381). In the first approach, the 25 candidates cDNAs were cloned into an epitope-tagged mammalian expression construct and then were co-expressed in HEK293T cells with wild-type or kinase-dead MARK1. Cells were lysed in IP buffer, followed by addition of 5 µg of GST or GST-14-3-3 to cell extracts and incubation at 4 degrees for 2 hr with gentle rocking. Complexes were precipitated with the addition of glutathione sepharose at 4°C for 1 hr. This assay reveals direct substrates that associate specifically with GST-14-3-3 only when the wild-type, but not kinase-dead, AMPKR is co-expressed. As half the known AMPKR substrates do not interact with 14-3-3 upon phosphorylation, our second approach was to examine whether exogenous MARK1 activity would regulate recognition of the substrate by phospho-motif antibodies. Here candidate substrate cDNAs in an epitope-tagged mammalian expression construct are co-expressed in HEK293T cells with wild-type or kinase-dead MARK1, then immunoprecipitated with the tag antibody and immunoblotted with the phospho-motif antibody. Of note, both of these approaches are based on co-overexpression of both the candidate substrate and the MARK kinases, so positive results here indicate that the phosphorylation event is possible in cells, but does reflect whether the endogenous candidate substrate is targeted by the endogenous AMPKR kinase when each are expressed at their endogenous levels and localized appropriately, which requires additional experimentation.

### **Cell Lysis, Immunoblotting, and Immunoprecipitation**

Cells were lysed at indicated timepoints in lysis/IP buffer (20 mM Tris (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1mM EGTA, 1% Triton X-100, 2.5 mM pyrophosphate, 50 mM NaF, 5 mM β-glycero-phosphate, 50 nM calyculin A, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 10 mM PMSF, 4 mg/ml leupeptin, 4

mg/ml pepstatin, 4 mg/ml aprotinin). Total protein was normalized using BCA protein kit (Pierce Protein, Rockford, IL) and lysates were resolved on SDS-PAGE gel.

Immunoprecipitations were performed on equilibrated protein lysates. Briefly, primary antibody  
5 was added to cleared cell lysates and incubated at 4°C on a rocking platform for 2 hours. Antibody complexes were then captured with the addition of Protein G-Sepharose beads (Invitrogen, Carlsbad, CA) for 1 hour. Immunoprecipitates were then washed 3 times in lysis buffer and samples were boiled and run on SDS-PAGE gels. Primary antibodies used for immunoprecipitation: mouse anti-FLAG M2 (Sigma), or mouse anti-myc 9E10 ascites fluid  
10 (from Dr. Tony Hunter).

### **Recombinant GST-14-3-3 binding assay**

Recombinant GST or GST-14-3-3 were produced in *E. coli* as previously described (Yaffe et al., Cell 91, 961-971, 1997) then purified on glutathione sepharose and eluted with free glutathione.  
15 MARK1 WT or MARK1 T215A along with the indicated FLAG-tagged substrate of interest was transfected into HEK 293T for 18 hr using Lipofectamine 2000. Cells were lysed in IP buffer, followed by addition of 5 µg of GST or GST-14-3-3 to cell extracts and incubation at 4°C for 2 hr with gentle rocking. Complexes were precipitated with the addition of glutathione sepharose at 4°C for 1 hr.

### **RNAi transfection and oligos**

RNAi mediated gene depletion was carried out by reverse transfection of siRNA oligos using RNAiMAX (Invitrogen) according to manufacturer's protocol. siRNA duplexes were used at 20 nM unless otherwise noted. Negative universal control was from Invitrogen (Medium GC  
25 content). ON-TARGET SmartPOOL RNAi against *hLKB1*, *hPRKAA1*, *hPRKAA2*, *hMARK1*, *hMARK2*, *hMARK3*, *hMARK4*, *hNUAK1*, *hNUAK2*, *hSIK1*, *hSIK2*, *hAPC*, *hDIXDC1*, *hBAIAP2*, *hCLASP1*, *hCLASP2*, *hKIF13B*, *hPARD3*, *hMAP4*, *mDixdc1*, *hCK1epsilon*, *hWNT5A*, *hWNT5B*, *hSNAI1*, and *mSnai1* were obtained from Dharmacon Thermo Scientific. Cells were harvested at 72 hours post siRNA transfection or further manipulated in functional assays.

### **Adhesion and replating experiments**

To analyze activity of tyrosine kinases present at focal adhesions, cells were stimulated by plating onto extracellular matrix coated dishes as previously described (Bernard-Trifilo et al.,

Curr Prot Cell Bio *Chapter 14; Unit 14.7.*, 2006). Briefly, cells were serum-starved the night previous to plating experiments. Cell culture dishes were coated with rat tail Collagen I (BD) in sterile water for 2 hours at 37°C. Collagen solution was aspirated off and dishes were allowed to dry for 30 min in a sterile hood. Serum-free DMEM was then added to coated plates and  
5 allowed to equilibrate for 30 min at 37°C. Meanwhile, cells were trypsinized, centrifuged, and resuspended at equal cell density in serum-free DMEM. Cells were kept in suspension for 1 hr to synchronize focal adhesion formation. After 1 hr, cells were plated onto equilibrated collagen-coated dishes and incubated for the indicated timepoints.

- 10 To inhibit focal adhesion kinase (FAK) activity using the chemical inhibitor PF-573228 (FIG. 7A), cells were exposed to the inhibitor (1  $\mu$ M) for 3 hours prior to cell suspension and also replated in the context of the inhibitor (1  $\mu$ M) for 1 hour on collagen-coated plates.

### **Mouse Bioluminescent Imaging**

- 15 Bioluminescent imaging was performed on mice using an IVIS Kinetic 200 from Caliper Life sciences. Mice were intraperitoneal (i.p.) injected with 150 mg/kg D-luciferin (Caliper Life Science Hopkinton, MA), anesthetized with isoflurane and dorsal and ventral images were then captured 10 minutes post luciferin injection. Bioluminescence was quantified by normalizing relative photon counts using Living Image 3.2 software (Caliper Life Science Hopkinton, MA).  
20 Primary mouse lung tumor cell lines were transduced in vitro to express firefly luciferase using a lentiviral vector (pLenti-pgk-LUC BLAST). Imaging was performed 30 minutes post intravenous transplantation, and then again at 3 weeks post transplantation before sacrifice. Total change in BLI was determined by normalizing signal at 3 weeks to signal at 30 minutes for each mouse.

25

### **Histology and Quantitation of lung tumor burden using Nuance/Inform software**

- Mouse tissues were perfusion fixed in 10% neutral buffered formalin overnight, processed routinely and embedded in paraffin. Four  $\mu$ M thick sections from formalin fixed paraffin embedded (FFPE) lung tumors were cut and stained with hematoxylin and eosin (H&E). Lung  
30 tumor burden was determined by histological analysis of tumor area. H&E stained lung sections were imaged using a Zeiss Axio Imager.M2 and the Nuance FX multispectral imaging system (Cambridge Research and Instrumentation, Cambridge, MA). Whole cross sections of lungs were imaged at 1.25X magnification. Images were converted to digital image cubes and spectral

libraries were made for Hematoxylin and Eosin (H&E). H&E stained whole lung sections were imaged and spectrally unmixed then pseudo-colored and the pixel counts were quantified using Nuance v1.0.2 and Inform v1.0.0 image analysis software (Cambridge Research and Instrumentation, Cambridge, MA). Analysis of tumor area in whole lung sections was  
5 determined by calculating the total pixel count for the red pseudo-colored tumors in lobes 1-5 of each mouse. Normal tissue was pseudo-colored green and blank space was colored blue. The mean tumor burden for each treatment group was calculated by averaging the total pixel count for each mouse in the treatment group.

## 10 Bioinformatics

In order to identify candidate MARK substrates, the optimal MARK motif was queried in the ScanSite ([scansite3.mit.edu](http://scansite3.mit.edu)) and ProSite ([prosite.expasy.org](http://prosite.expasy.org)) databases. Proteins with consensus motif matches were further queried using the PhosphoSite ([www.phosphosite.org](http://www.phosphosite.org)) database for existing evidence of *in vivo* phosphorylation. Site conservation was examined  
15 across evolution using UniProt ([www.uniprot.org](http://www.uniprot.org)) and NCBI PubMed Protein databases.

mRNA expression analysis of *DIXDC1* in human lung cancer datasets was performed using Lung Cancer Explorer (UTSW QBRC/CCBSR; [qbrc.swmed.edu/lce/](http://qbrc.swmed.edu/lce/)). Studies represented are referenced in the main text. Graphs in FIGS. 12D-12F were generated through curation of  
20 publicly available datasets on the NCBI GEO Omnibus using the GEO2R viewer. Datasets analyzed in this study include GSE7670, GSE19188, GSE9199. Profile graphs were obtained using the Affymetrix gene probe-set ID number for *DIXDC1* on the given array platform. Mapping of *DIXDC1* copy number alterations was done using TumorScape platform encompassing the TCGA databases (Broad Institute; [www.broadinstitute.org/tcga](http://www.broadinstitute.org/tcga)).

25 Correlation of *DIXDC1* expression levels to patient survival in non-small cell lung cancer was done using the KMplotter ([kmplot.com](http://kmplot.com); (Gyorffy et al., PLoS One 8, e82241, 2013). Patients were split by median *DIXDC1* expression, and analysis was performed across all molecular subtypes irrespective of grade, stage, or prior treatment regimen. OS = Overall Survival. FP =  
30 First Progression (Median Follow-up). PPS = Previously published survival.

## Example 2

### **Snail is upregulated from loss of LKB1 or its downstream kinases MARK1 and MARK4, independent of AMPK**

5           Given the findings that LKB1-deficiency promotes aggressive metastatic behavior of diverse tumor types (Contreras et al., Cancer Res. 68, 759–766, 2008; Ji et al., Nature 448, 807–810, 2007; Liu et al., Cancer Cell 21, 751–764, 2012), critical downstream effectors of LKB1 that contribute to its control of metastatic behavior were identified. One hallmark of the metastatic cascade is the requirement for spatially restrained epithelial cells to acquire the ability  
10   to invade local tissue, intravasate, and extravasate at distant sites. Experimental evidence indicates that this cascade can be mediated by the induction of an epithelial to mesenchymal transition (EMT), a transcriptional program orchestrated by re-expression of conserved transcription factors critical in embryonic development (Chaffer and Weinberg, Science 331, 1559–1564, 2011). It was determined whether the expression of any EMT transcription factors  
15   was repressed by LKB1 function by examining the impact of LKB1 knockdown in a panel of LKB1-proficient tumor cell lines.

          Across cell lines from diverse tumor types, the zinc-finger transcription factor Snail1 was consistently upregulated when LKB1 was suppressed, in contrast to expression of Snail2 (Slug), ZEB1, ZEB2, or Twist (FIG. 1A). In human NSCLC cell lines bearing LKB1 inactivation,  
20   expression of wild-type but not kinase-dead LKB1 reduced Snail1 levels (FIG. 1B) and LKB1 knockdown in a LKB1-proficient NSCLC cell line increased Snail1 levels (FIG. 1C). Surprisingly, the ability of LKB1 to suppress Snail1 was not confined to tumor cells or cells bearing oncogenic signals as *Lkb1*<sup>-/-</sup> mouse embryonic fibroblasts (MEFs) exhibited Snail1 upregulation compared to littermate *Lkb1*<sup>+/+</sup> MEFs (FIG. 1D). Examining the mechanism of  
25   Snail regulation in genetic contexts, LKB1 was necessary and sufficient for suppression of Snail at the mRNA level (FIGS. 2A and 2B).

          A core function of Snail family transcription factors is to create a cellular state favorable to cell migration and invasion (Thiery et al., Cell 139, 871–890, 2009). Snail mediates this through repression of the E-cadherin promoter, but also through transcriptional induction  
30   (Rembold et al., Genes Dev 28, 167–181, 2014) of mRNAs for extracellular matrix components, metalloproteinases, and numerous secreted growth factors, including the noncanonical Wnt ligands, Wnt5a and Wnt5b (Moreno-Bueno et al., Cancer Res 66, 9543–9556, 2006; Ren et al., Genes Cells 16, 304–315, 2011). As *Wnt5a* was previously documented as a highly upregulated

mRNA in *Lkb1*-deficient gastrointestinal polyps (Lai et al., J Pathol 223, 584-592, 2011), the relationship between Wnt5a and Wnt5b levels and Snail1 levels across cell types was examined. It was observed that Snail1 was necessary (FIG. 2C) and sufficient (FIG. 2D) for induction of Wnt5a and Wnt5b in U2OS cells. Similarly, Wnt5a/Wnt5b levels paralleled Snail protein levels in various cell types when LKB1 was silenced (FIGS. 1B, 1C, 1D, and 2B). Moreover, elevated Wnt5a/Wnt5b in *Lkb1*<sup>-/-</sup> MEFs was attenuated by knockdown of Snail1 (FIG. 1E). Collectively these results indicate that Snail is necessary and sufficient for Wnt5a/5b expression in LKB1-deficient contexts, indicating that Wnt5a/5b levels can serve as biomarkers of Snail activity.

Snail expression was higher in lysates from lung tumors isolated from *Kras*<sup>LSL-G12D/+</sup> *Lkb1*<sup>L/L</sup> mice than in tumors from *Kras*<sup>LSL-G12D/+</sup> *Lkb1*<sup>+/+</sup> mice and Wnt5a/b levels paralleled elevated Snail levels here (FIG. 1F). Given the propensity of *Lkb1*-deficiency to enhance metastatic potential in *Kras*<sup>LSL-G12D/+</sup> lung tumors (Ji et al., Nature 448, 807-810, 2007), the upregulation of Snail in this context makes it a potential molecular component of how LKB1 controls *bona fide* metastasis. The molecular mechanisms by which LKB1 controls Snail levels across cell types was therefore determined.

Because LKB1 can activate multiple AMPK-related kinases (“AMPKRs”), which downstream kinases controlled Snail levels were identified. For screening purposes, U2OS cells were utilized as a human cell system in which LKB1 signaling is fully intact, but can be readily suppressed by RNAi-mediated silencing of LKB1. As previously observed (FIG. 1A), LKB1 depletion in U2OS cells resulted in elevated Snail levels, yet surprisingly combined knockdown of the two genes encoding the AMPK catalytic subunits (AMPKα1 and AMPKα2) had no effect, even though phosphorylation of the AMPK substrate ACC was fully suppressed (FIG. 1G). In contrast, knockdown of the four member MARK/Par-1 subfamily of LKB1-dependent kinases resulted in Snail induction to an equal or greater extent than LKB1 knockdown (FIG. 1G). Knockdown of each of the MARK family members individually revealed that MARK1 (also known as Par1c) and MARK4 (Par1d) were most critical to suppression of Snail in these cells (FIG. 1H), while the related kinases MARK2 and MARK3 had no effect on Snail levels. Deconvolution of the RNAi pools for LKB1 and MARK1 revealed that multiple independent siRNA duplexes against each target resulted in Snail induction, indicating that our observations are unlikely to be due to off-target silencing of unintended genes (FIG. 2E).

### Example 3

#### **DIXDC1 is a novel substrate of MARK1 and MARK that suppresses Snail levels**

This example describes methods used to further dissect the mechanism by which MARK1 and MARK4 regulate Snail levels, as very little is known about these two kinases (FIG. 1D). Previously, a biochemical screen was used to identify novel direct substrates of AMPK based on the determination of an optimal substrate motif phosphorylated by AMPK family kinases using arrayed positional scanning peptide libraries (FIG. 3A) (Egan et al., *Science* 331, 456-461, 2011; Gwinn et al., *Mol Cell* 30, 214-226, 2008; Mihaylova et al., *Cell* 145, 607-621, 2011; Turk et al., *Nat Protoc* 1, 375-379, 2006). To identify substrates of the MARKs, the optimal substrate motif was determined for all four MARK kinases using the same method, and nearly identical profiles were observed (FIGS. 3B and 4A). The optimal *in vitro* peptide phosphorylation sequence for the MARK kinases ("MARK motif") was also quite similar to a previously described AMPK motif (Gwinn et al., *Mol Cell* 30, 214-226, 2008) (FIG. 4A). This common consensus sequence includes strong preferences for aliphatic residues at positions -5 and +4 relative to the phosphorylation site, as well as for a basic residue at position -3. The similarity of these optimal sequences is consistent with previous studies demonstrating that phosphorylation sites induced by distinct members of the AMPKR family are very similar in their sequence specificity, and corresponds well with a survey of well-studied AMPK and MARK substrate phosphorylation sites (Shackelford and Shaw, *Nat Rev Cancer* 9, 563-575, 2009).

The human proteome was bioinformatically scanned to identify proteins containing candidate phosphorylation sites matching the optimal MARK substrate motif. Re-performing the search on other eukaryotic genomes allowed us to narrow the initial ~1000 candidate human substrates to ~300 highly conserved candidate substrate sites (FIG. 3B). Given that control of the cytoskeleton is a major conserved function of the MARK/Par1 kinase family, the 25 candidate MARK substrates bearing conserved phosphorylation motifs that also contained a Pfam domain indicative of a role in cytoskeleton or were classified by gene ontology (GO) as having a role in cytoskeletal signaling were examined. As a first test of whether these 25 candidates may be regulated by the MARK kinases, it was determined whether exogenous gain and loss of function of MARK1 kinase activity could regulate the phosphorylation of these proteins, as measured indirectly through their binding to the phospho-dependent binding scaffold protein 14-3-3 or recognition by a phospho-motif antibodies generated against the

optimal AMPK family substrate motif (Gwinn et al., Mol Cell 30, 214-226, 2008). Many of the known substrates of AMPK, the MARKs, and their related kinases inducibly bind to 14-3-3 upon their phosphorylation by these kinases (Shackelford and Shaw, Nat Rev Cancer 9, 563-575, 2009). This screening approach revealed a handful of previously reported MARK  
5 substrates - MAP4, Par3, KIF13B, IRSp53 (see FIGS. 3B and 3C), as well as four novel targets whose phosphorylation was controlled in a MARK-dependent manner (CLASP1, CLASP2, APC, DIXDC1), (FIGS. 3B and 3D).

It was determined whether knockdown of the four known and four novel MARK substrates could increase, or decrease, basal Snail expression levels in U2OS cells. Notably, loss  
10 of the novel target DIXDC1 demonstrated the greatest effect, triggering an upregulation of Snail levels similar to that induced by knockdown of MARK1 or LKB1 (FIG. 4B). Three distinct siRNA oligos targeting DIXDC1 all resulted in upregulation of Snail demonstrating this was not an off-target effect of the siRNA pool used (FIG. 3E).

DIXDC1 (also known as CCD1) is a poorly-studied scaffolding protein composed of  
15 three protein-protein interaction domains: an actin-binding calponin homology (CH) domain, a coiled-coil (CC) domain, and a dishevelled and axin (DIX) oligomerization domain (FIG. 4C). Notably, the candidate MARK phosphorylation site in DIXDC1, Ser592, is highly conserved, and lies just at the start of the DIX domain (FIG. 4C). Given that wild-type but not kinase-dead MARK1 induced binding of DIXDC1 to 14-3-3, it was determined whether the putative  
20 MARK1 phosphorylation site Ser592 was required for the observed association between DIXDC1 and 14-3-3. The ability of wild-type MARK1 to promote 14-3-3 binding of wild-type versus a non-phosphorylatable Ser592Ala mutant was determined. Mutation of Ser592 completely abolished the ability of wild-type MARK1 to induce DIXDC1 to bind 14-3-3 (FIG. 3F).

Based on this observation, a phospho-specific antibody was developed against Ser592.  
25 To establish the specificity of the antibody, it was used to immunoblot purified DIXDC1 that had been incubated with active recombinant MARK1 protein in an *in vitro* kinase assay (FIG. 4D). Active MARK1 induced reactivity of WT DIXDC1 with the Phospho-Ser592 antibody, and this was not observed in the DIXDC1 S592A non-phosphorylatable mutant. Similar results  
30 were observed with MARK4, demonstrating that both of these MARK kinases can phosphorylate DIXDC1 Ser592 *in vitro* (FIG. 3E). Using the purified phosphor-specific antibody, it was observed that overexpression of WT but not kinase inactive (KD) MARK1 increased co-expressed DIXDC1 Ser592 phosphorylation in HEK293T cells, as did co-

expression of LKB1 and its activating subunit STRAD $\alpha$  with DIXDC1 (FIG. 3H), confirming exogenously expressed DIXDC1 can serve as a target for LKB1- and MARK1-dependent signals in cells.

To verify the endogenous kinase(s) responsible for phosphorylation of endogenous DIXDC1 at Ser592, LKB1, AMPK, and individual MARK family kinases were knocked down. Silencing of LKB1, MARK1, or MARK4 resulted in loss of endogenous Ser592 DIXDC1 phosphorylation, while knockdown of AMPK or MARK2 or MARK3 had no effect (FIG. 4F). The regulation of DIXDC1 Ser592 phosphorylation by MARK1 and MARK4 paralleled the observations regarding which AMPK family kinases suppress Snail levels (FIG. 1H). Further examination confirmed that levels of phospho-Ser592 DIXDC1 were inversely correlated with the level of Snail expression (FIG. 4G). These data indicate that the ability of DIXDC1 to suppress Snail levels may require LKB1- and MARK1/4- dependent phosphorylation of DIXDC1 Ser592. Consistent with the observed effects on Snail, DIXDC1 knockdown also resulted in increased Wnt5a/b, an effect that required Snail (FIG. 3I) as previously noted in the context of LKB1-deficiency (FIG. 1E).

#### Example 4

##### **DIXDC1 localizes to focal adhesions in a Serine 592-dependent manner and regulates their maturation**

To investigate how DIXDC1 controls Snail levels, its subcellular localization was examined. Indirect immunofluorescence on U2OS cell lines in which endogenous DIXDC1 was suppressed and stably replaced with a FLAG-tagged DIXDC1, revealed that DIXDC1 localized primarily to focal adhesions (FIG. 5A), consistent with the only study to report the localization of endogenous DIXDC1 (Wang et al., *Biochem Biophys Res Commun* 347, 22-30, 2006). DIXDC1 localized just distal to the focal adhesion resident protein paxillin, in a region where the focal adhesion complexes link to the intracellular actin cytoskeleton (Wehrle-Haller, *Curr Opin Cell Biol* 24, 116-124, 2012).

To determine if DIXDC1 affects focal adhesion dynamics, the morphology of focal adhesions was examined in U2OS knocked down for DIXDC1 using the marker paxillin. At one hour post-plating, a striking contrast was observed between mature, actin-linked adhesion plaques in control shRNA cells and an abundance of nascent, immature focal contacts in DIXDC1 shRNA cells (FIGS. 5B and 5C). The change in focal adhesion morphology within 1 hour after plating in DIXDC1 knockdown cells led us to examine whether there was a defect in

adhesion maturation. Immunolocalization of the focal adhesion protein zyxin, which localizes preferentially to mature adhesions and can be used to distinguish the age of a focal complex, was performed (Zaidel-Bar et al., J Cell Sci 116, 4605-4613, 2003). In control U2OS cells, zyxin was perfectly colocalized with paxillin within 1 hour of adhesion onto collagen matrix. In contrast, zyxin was largely absent from focal contacts in *DIXDC1* knockdown cells, indicative of a defect in focal adhesion maturation (FIGS. 5C, 5D and 5G). Collectively, these results indicate for the first time that *DIXDC1* controls the dynamics of focal adhesions. Knockdown for LKB1 or MARK1 in U2OS cells phenocopied the effect of *DIXDC1* knockdown on focal adhesions as observed by zyxin and paxillin co-staining 1 hour after plating (FIG. 5E).

To directly examine the effect of MARK1 phosphorylation of Ser592 on *DIXDC1* action at focal adhesions, stable cell lines expressing FLAG- *DIXDC1*<sup>WT</sup> or FLAG-*DIXDC1*<sup>S592A</sup> were compared to cells with FLAG-*DIXDC1*<sup>WT</sup> but with shRNAs to MARK1 and MARK4. Non-phosphorylatable FLAG-*DIXDC1*<sup>S592A</sup> was no longer capable of localizing to focal adhesions like FLAG-*DIXDC1*<sup>WT</sup> (FIG. 5F). FLAG-*DIXDC1*<sup>WT</sup> was similarly mislocalized away from focal adhesions by knockdown for MARK1 and MARK4 (FIG. 5F, bottom image), collectively indicating that Ser592 phosphorylation is required for proper localization of *DIXDC1* to focal adhesions.

## Example 5

### Hyperactivation of FAK from loss of LKB1 or *DIXDC1* function is responsible for elevated Snail

As FAK is normally recruited to nascent focal adhesions concurrently with paxillin (Kuo et al., Nat Cell Biol 13, 383-393, 2011), and FAK was previously reported to be hyperactivated in LKB1-deficient lung tumors (Carretero et al., Cancer Cell 17, 547-559, 2010), the impact of *DIXDC1* silencing on FAK was determined. Activated Tyr397-phosphorylated FAK colocalized with paxillin 1 hour after plating, in both nascent adhesions in the *DIXDC1* or LKB1 knockdown treated cells, as well as the larger more mature focal adhesions formed in the control cells (FIG. 6A). Notably, nascent adhesions were previously observed to be highly enriched for phospho-tyrosine (Zaidel-Bar et al., J Cell Sci 116, 4605-4613, 2003). Consistent with this, cells bearing *DIXDC1* knockdown exhibited a greater peak and prolonged activation of FAK upon plating (FIG. 6B). The increase in FAK activation from *DIXDC1* knockdown also resulted in elevated Src signaling, as defined by Src activation loop Tyr416 phosphorylation or Src-dependent phosphorylation of paxillin at Tyr118, which were both attenuated in cells

expressing two distinct FAK shRNAs (FIG. 6C) or by treatment of DIXDC1 RNAi bearing cells with the small molecule FAK inhibitor PF-573228 (Slack-Davis et al., J Biol Chem 282, 14845-14852, 2007) (FIG. 7A).

Moreover, it was observed that two FAK shRNAs or the FAK inhibitor PF-573228 attenuated Snail1 induction from knockdown of DIXDC1 (FIGS. 6C and 6D), MARK1, or LKB1 (FIG. 6E), consistent with a study reporting that FAK regulates Snail1 levels in MEFs (Li et al., J Cell Biol 195, 729-738, 2011). The extent that Snail levels were suppressed from the two FAK shRNAs was proportional to their ability to reduce FAK protein levels (FIG. 6C).

In DIXDC1 knockdown cells, FAK also exhibited increased phosphorylation at Tyr925, which mediates binding to Grb2 and activation of the Ras-MEK-ERK pathway. Consistent with the increased P-Y925, immunoprecipitation of endogenous GRB2 revealed more endogenous FAK and Src co-immunoprecipitating with GRB2 in the DIXDC1 knockdown cells (FIG. 7B). Correspondingly, higher levels of activating phosphorylation events were observed on MEK, ERK1/2, and the ERK substrate RSK Ser380 in DIXDC1 knockdown cells, all of which were blocked by treatment with the FAK inhibitor PF-573228 (FIG. 6D).

To further examine what pathways downstream of FAK were involved in Snail upregulation, cells were treated with small molecule inhibitors against MEK, PI3K, Akt, or mTOR, for which only the MEK inhibitor U0126 caused complete loss of Snail protein, similar to the extent from the FAK inhibitor itself (FIGS. 6F, 7C, and 7D). This observation is consistent with previous studies reporting multiple mechanisms of ERK-dependent increases in Snail, including both mRNA induction (Barbera et al., Oncogene 23, 7345-7354, 2004; Hsu et al., Oncogene 32, 4436-4447, 2013; Strippoli et al., Dis Model Mech 1, 264-274, 2008; Toettcher et al., Cell 155, 1422-1434, 2013) and direct ERK phosphorylation sites in Snail suppressing protein turnover (Zhang et al., Nat Cell Biol 15, 677-687, 2013). Collectively, these findings indicate that when LKB1, MARK1, or DIXDC1 function is compromised, FAK becomes hyperactivated resulting in MEK/ERK-dependent induction of Snail (FIG. 6G).

## Example 6

### DIXDC1 suppresses cell migration and invasion in a Ser592-dependent manner

To examine how DIXDC1 loss impacts cell behavior, cell migration and invasion was examined.

Scratch assays were used to examine directed migration, in the presence of mitomycin C to prevent proliferation during the experiment. RNAi silencing of LKB1 or DIXDC1 resulted in

increased wound healing, which was mimicked by stable overexpression of non-phosphorylatable DIXDC1<sup>S592A</sup>, but not DIXDC1<sup>WT</sup> (FIGS. 5A, 5B). This finding indicates that in this context, DIXDC1<sup>S592A</sup> behaves as a dominant negative allele to suppress the activity of the endogenous DIXDC1. Increased migration upon DIXDC1 knockdown was attenuated by the small molecule FAK inhibitor PF-573228 (FIGS. 8C and 9) or Snail RNAi (FIG. 8D).

To determine if DIXDC1 similarly impacted invasive behavior, cells were placed in a collagen-coated Boyden chamber and allowed to invade for 24 hours prior to fixation. Cells with stable LKB1 or DIXDC1 knockdown had a ~1.8-fold increase in cell invasion in that time window (FIG. 8E). The enhanced collagen invasion of cells bearing DIXDC1 knockdown was also blocked by incubation with the small molecule FAK inhibitor PF-573228 (FIG. 9).

### Example 7

#### Dixdc1 dictates metastatic behavior in non-small cell lung cancer lines

Given the role of *Lkb1* in the suppression of lung tumor metastasis in mouse *Kras*-dependent NSCLC models (Carretero et al., Cancer Cell 17, 547-559, 2010; Ji et al., Nature 448, 807-810, 2007), the function of *Dixdc1* in lung tumor cell lines with defined metastatic characteristics was determined. A panel of *Kras*<sup>G12D</sup>, *p53*<sup>-/-</sup> mutant lung adenocarcinoma cell lines derived from nonmetastatic and metastatic primary mouse lung tumors (*T*<sub>nonMet</sub> and *T*<sub>Met</sub> respectively) that have been demonstrated to retain the metastatic capacity of their tumor of origin (Li et al., Genes Dev 27, 1557-1567, 2013; Winslow et al., Nature 473, 101-104, 2011) were used.

Notably, while *Lkb1* protein levels were unaffected, *Dixdc1* protein and mRNA levels were lower in the *T*<sub>Met</sub> lines compared to the *T*<sub>nonMet</sub> lines (FIGS. 10A, 11A), whereas Snail and FAK Phospho-Y397 levels were higher in the metastatic setting. Thus *Dixdc1* reduction and the previously observed associated biochemical signature of DIXDC1 suppression correlated with metastatic potential in these cell lines. Furthermore, stable knockdown of *Dixdc1* in *T*<sub>nonMet</sub> cells resulted in hyperactivation of FAK signaling and increased Snail levels to an extent similar to that found in the *T*<sub>Met</sub> cells (FIG. 10B), an effect observed with two distinct effective *Dixdc1* shRNAs.

To further investigate *Dixdc1* function in these murine lung adenocarcinoma cell lines, it was first confirmed that *Lkb1*, *Mark1*, and *Mark4* regulate endogenous *Dixdc1* Ser592 phosphorylation in the *T*<sub>nonMet</sub> cells (FIG. 11B). Also similar to what was observed in U2OS cells, acute RNAi to *Mark1* or *Dixdc1* resulted in hyperactivation of FAK upon plating (FIG.

11C). Conversely, stable enforced expression of Dixdc1 into the T<sub>Met</sub> cells suppressed FAK signaling (P-FAK, P-Src, P-Paxillin) (FIG. 11D). Consistent with regulation of the FAK-Snail signature, Dixdc1 knockdown in T<sub>nonMet</sub> cells resulted in increased invasion (FIG. 11E).

To directly examine the impact of loss of Dixdc1 on metastatic potential *in vivo*, Dixdc1 was stably knocked down with two independent shRNAs in T<sub>nonMet</sub> cells and a lung colonization assay was performed in syngeneic mice, an assay previously shown to mirror metastatic ability of this cell system (Winslow et al., Nature 473, 101-104, 2011). Dixdc1 knockdown dramatically increases the lesion number and total tumor burden (FIG. 10C), without any differences in the growth rate of these cells in culture (FIG. 11F). Further characterization revealed the increase in tumor burden in T<sub>nonMet</sub> cells observed from shRNAs against Dixdc1 was comparable to the extent of lesions induced by a previously validated shRNA to Lkb1 (FIG. 11G). Importantly, the increased tumor burden in the T<sub>nonMet</sub> cell lines bearing Dixdc1 shRNA was fully rescued by stable re-expression of a wild-type but not S592A hDIXDC1 cDNA (FIG. 10D). Moreover, hDIXDC1 cDNAs were expressed in T<sub>Met</sub> cells, which lack appreciable Dixdc1 (FIGS. 10A and 11H), and observed a potent suppression of experimental metastasis by wild-type but not S592A hDIXDC1 (FIGS. 6E-G), providing crucial proof of the importance of this single phosphorylation site in governing metastasis *in vivo*. Quantification using luciferase imaging confirmed the decrease in total tumor burden (FIGS. 10F and 10G). The suppression of metastasis by overexpression of wild-type but not S592A Dixdc1 also demonstrates that the T<sub>Met</sub> cells are capable of receiving Lkb1-dependent signals to Dixdc1, but simply lack sufficient Dixdc1 expression as compared to the nonmetastatic cells.

## Example 8

### DIXDC1 is frequently lost in human cancers

Though the T<sub>Met</sub> cell lines retain Lkb1, given their loss of *Dixdc1* mRNA expression and potent effects of Dixdc1 on tumor burden in this setting, it was determined whether DIXDC1 is deleted, mutated, or downregulated in human tumors. Copy number variations were determined for human DIXDC1, which lies in a broad tumor suppressor region on chromosome 11q23.1. GISTIC analysis reveals that DIXDC1 is significantly deleted in 15 out of 26 cancer subsets across the entire Tumorscape/TCGA dataset of 8177 tumors. Notably, DIXDC1 is most significantly deleted independent of a peak region in metastatic cutaneous melanoma, lung cancer, and cervical squamous cell carcinoma, of which melanoma and lung cancer represent well-defined settings where loss of Lkb1 in GEMM models modulates metastasis and FAK/Src

signaling (Liu et al., *Cancer Cell* 21, 751-764, 2012) (FIG. 12A). In support of this analysis, using a new algorithm (CONEXIC) that combines gene expression and copy number alterations to identify driver events in melanoma, suppression of DIXDC1 was identified as one of the top 30 drivers in this tumor type (Akavia et al., *Cell* 143, 1005-1017, 2010). Importantly, DIXDC1 mRNA expression correlated better with global gene expression changes than any of the other 16 genes in a focal deletion region of 11q23.1 or the focal deletion of 11q23.1 itself (Akavia et al., *Cell* 143, 1005-1017, 2010), reinforcing that alterations in *DIXDC1* mRNA levels are a critical measure of its functionality beyond the subset of tumors bearing LKB1 mutations where DIXDC1 function would be abrogated from loss of Ser592 phosphorylation.

Analysis of DIXDC1 mRNA expression in human NSCLC cancer datasets revealed its frequent downregulation (FIGS. 12B, 13A-13C). Interestingly, *Dixdc1* was previously identified as one of only 97 genes genomewide that was cooperatively suppressed at the mRNA level by Kras and p53 mutation in a pattern correlating with the malignant phenotype (McMurray et al., *Nature* 453, 1112-1116, 2008) (FIG. 13D). Genes identified in the study proved relevant across a panel of human tumor cell lines, containing a plethora of mutations in addition to the defined Kras and p53, indicating that genes present in the dataset may be relevant for tumorigenesis across cell contexts.

To further explore DIXDC1 expression in human cancer, DIXDC1 levels were examined as a function of patient outcome in an annotated set of NSCLC tumors. Mining a recent meta-analysis of published lung cancer microarray datasets that identified biomarkers related to survival (Gyorffy et al., *PLoS One* 8, e82241, 2013), it was observed that low levels of DIXDC1 are correlated with decreased survival, whether one considers overall survival, follow-up survival, or progression-free survival (FIG. 12C). As survival in such cohorts is largely due to metastasis, the correlation of lowered DIXDC1 mRNA and poor survival suggests a lower DIXDC1-increased metastatic potential correlation.

In view of the many possible embodiments to which the principles of the disclosure may be applied, it should be recognized that the illustrated embodiments are only examples of the disclosure and should not be taken as limiting the scope of the invention. Rather, the scope of the disclosure is defined by the following claims. We therefore claim as our invention all that comes within the scope and spirit of these claims.

We claim:

1. A method for analyzing a cancer sample obtained from a subject, comprising:  
contacting the sample with a DIX domain-containing protein 1 (DIXDC1) protein-specific binding agent or a DIXDC1 nucleic acid probe;  
5 detecting DIXDC1 proteins or DIXDC1 nucleic acid molecules in the sample; and  
determining that the cancer is sensitive to a focal adhesion kinase (FAK) inhibitor and/or a Src inhibitor when DIXDC1 protein expression, DIXDC1 protein phosphorylation, and/or DIXDC1 nucleic acid molecule expression is reduced in the sample relative to a control or  
10 reference value, thereby analyzing the cancer sample.
2. The method of claim 1, further comprising:  
detecting DIXDC1 proteins or DIXDC1 nucleic acid molecules in a positive or negative control sample; and  
15 comparing DIXDC1 protein expression, DIXDC1 protein phosphorylation, and/or DIXDC1 nucleic acid molecule expression in the control sample to the cancer sample, wherein detection of reduced DIXDC1 protein expression, DIXDC1 protein phosphorylation, and/or DIXDC1 nucleic acid molecule expression in the cancer sample relative to the positive control sample indicates that the cancer is sensitive to a FAK inhibitor and/or Src inhibitor, and wherein  
20 detection of similar DIXDC1 protein expression, DIXDC1 protein phosphorylation, and/or DIXDC1 nucleic acid molecule expression in the cancer sample relative to the negative control sample indicates that the cancer is sensitive to a FAK inhibitor and/or Src inhibitor.
3. The method of claim 1 or 2, further comprising:  
25 inputting the detected DIXDC1 protein expression, DIXDC1 protein phosphorylation, and/or DIXDC1 nucleic acid molecule expression in the cancer sample and/or the control sample into a computer; and  
generating an output from the computer, thereby analyzing the sample.
- 30 4. The method of any of claims 1 to 3, wherein the DIXDC1 protein-specific binding agent comprises a DIXDC1 antibody or fragment thereof.

5. The method of any of claims 1 to 4, wherein the sample comprises a surgical resection specimen, tissue biopsy or fine needle aspirate.
6. The method of claims 1 to 5, wherein detecting DIXDC1 proteins comprises detecting  
5 DIXDC1 protein phosphorylation at Ser 592.
7. The method of any of claims 1 to 6, wherein contacting the sample with the DIXDC1 protein-specific binding agent is performed with an automated tissue stainer.
- 10 8. The method of any of claims 1 to 7, wherein detecting DIXDC1 proteins or DIXDC1 nucleic acid molecules in the sample comprises visual inspection or image analysis of a corresponding digital image.
9. The method of claim 8, wherein the visual inspection is performed utilizing light  
15 microscopy or fluorescence microscopy.
10. The method of any one of claims 1 to 9, wherein detecting DIXDC1 proteins or DIXDC1 nucleic acid molecules in the sample comprises direct or indirect detection of binding of the DIXDC1 protein-specific binding agent or a DIXDC1 nucleic acid probe to the sample.  
20
11. The method of any of claims 1 to 10, further comprising obtaining the sample.
12. The method of any of claims 1 to 11, wherein the method is a method of distinguishing  
25 between a subject who is likely to respond to treatment with a FAK inhibitor and/or Src inhibitor from a subject who is not likely to respond to treatment with a FAK inhibitor and/or Src inhibitor.
13. The method of claim 12, further comprising selecting the subject for treatment with the FAK inhibitor and/or Src inhibitor if the subject is identified as a subject who is likely to  
30 respond to treatment with a FAK inhibitor and/or Src inhibitor.

14. The method of claim 11 or 12, further comprising administering a therapeutically effective amount of the FAK inhibitor and/or Src inhibitor to the subject identified as a subject who is likely to respond to treatment with the FAK inhibitor and/or Src inhibitor.

15. The method of any of claims 11 to 14, wherein the FAK inhibitor comprises TAE226; PF-562,271; PF-573,228; PF-04554878; GSK2256098; PND-1186; 1,2,4,5-benzenetetraamine tetrahydrochloride; or combinations thereof.

16. The method of any of claims 11 to 14, wherein the Src inhibitor comprises Dasatinib (N-[2-Chloro-6-methylphenyl]-2-[[6-[4-(2-hydroxyethyl)-1-piperazinyl]-2-methyl-4-pyrimidinyl]amino]-5-thiazolecarboxamide), Saracatinib (AZD0530; N-(5-chlorobenzo[d][1,3]dioxol-4-yl)-7-(2-(4-methylpiperazin-1-yl)ethoxy)-5-(tetrahydro-2H-pyran-4-yloxy)quinazolin-4-amine), Bosutinib (SKI-606), KX2-391 (KX01), quercetin, SU6656, SU6657, PP2, PP1, WHI-P 154; NVP-BHG712, Lavendustin C, ON-01910, SD 1008, Indirubin Derivative E804, ZM-306416, PP121, Src/EGFR inhibitor, Dasatinib  $\beta$ -D-Glucuronide, or combinations thereof.

17. The method of any of claims 1 to 16, wherein the subject is a chemo-naïve subject.

18. The method of any of claims 1 to 17, wherein the method is a method of predicting the likelihood that the cancer will metastasize, wherein detection of reduced DIXDC1 protein expression, DIXDC1 protein phosphorylation, and/or DIXDC1 nucleic acid molecule expression in the cancer sample relative to the control or reference value indicates that the cancer is more likely to metastasize.

19. The method of any of claims 1 to 18, wherein the control is a normal tissue of the same type as the cancer sample and the reference value is an amount of DIXDC1 protein expression, DIXDC1 protein phosphorylation, and/or DIXDC1 nucleic acid molecule expression expected in a normal tissue of the same type as the cancer sample.

20. A method of treatment comprising:

analyzing a cancer sample obtained from a subject according to method of any of claims 1 to 18; and

administering a therapeutically effective amount of the FAK inhibitor and/or Src inhibitor to the subject identified as a subject who is likely to respond to treatment with the FAK inhibitor and/or Src inhibitor.

5 21. The method of any of claims 1 to 20, wherein one or more steps are performed by a suitably-programmed computer.

22. A computer-implemented method, comprising:  
generating a DIXDC1 protein expression or protein phosphorylation score based at least  
10 on measured DIXDC1 protein expression or protein phosphorylation within a displayed image depicting a cancer sample detectably labeled with a DIXDC1 specific binding agent, wherein the cancer sample is obtained from a subject; and  
outputting a DIXDC1 protein expression or protein phosphorylation score for the sample.

15 23. A computer-implemented method, comprising:  
generating a DIXDC1 nucleic acid expression score based at least on measured DIXDC1 nucleic acid expression within a displayed image depicting a cancer sample, or nucleic acids from the cancer sample, detectably labeled with a DIXDC1 nucleic acid probe, wherein the  
20 cancer sample is obtained from a subject; and  
outputting a DIXDC1 nucleic acid expression score for the sample.

24. The computer-implemented method of claim 22 or 23, further comprising outputting a prognosis for the subject, wherein the prognosis comprises:  
25 an indication as to whether or not the cancer is sensitive to a FAK inhibitor and/or a Src inhibitor;  
an indication as to the likelihood that the cancer will metastasize; or  
combinations thereof.

30 25. One or more non-transitory computer-readable media comprising computer-executable instructions causing a computing system to perform the method of any of claims 1 to 24.

26. A system for analyzing a cancer sample obtained from a subject, comprising:

means for measuring a level of DIXDC1 protein or DIXDC1 nucleic acid molecule in the sample;

5 implemented rules for comparing the measured level of DIXDC1 protein or DIXDC1 nucleic acid molecule to a DIXDC1 reference value; and

means for implementing the rules, whereby an indication of the likely risk of cancer metastasis and/or sensitivity of the cancer to a FAK inhibitor and/or a Src inhibitor is provided based on the measured level of DIXDC1 protein, DIXDC1 phosphorylation, or DIXDC1 nucleic  
10 acid molecule in the sample.

27. A kit comprising:

a DIXDC1 specific-binding agent and/or a DIXDC1 nucleic acid probe, and one or more of:

15 a pair of primers specific for a DIXDC1 gene sequence;  
microscope slides;  
labeled secondary antibodies; and  
buffers for immunohistochemistry or in situ hybridization.

20 28. The kit of claim 27, wherein the DIXDC1 protein specific binding agent comprises a DIXDC1 antibody.

29. The kit of claim 27 or 28, wherein the kit further comprises a FAK inhibitor and/or Src inhibitor.

25

**FIG. 1A**

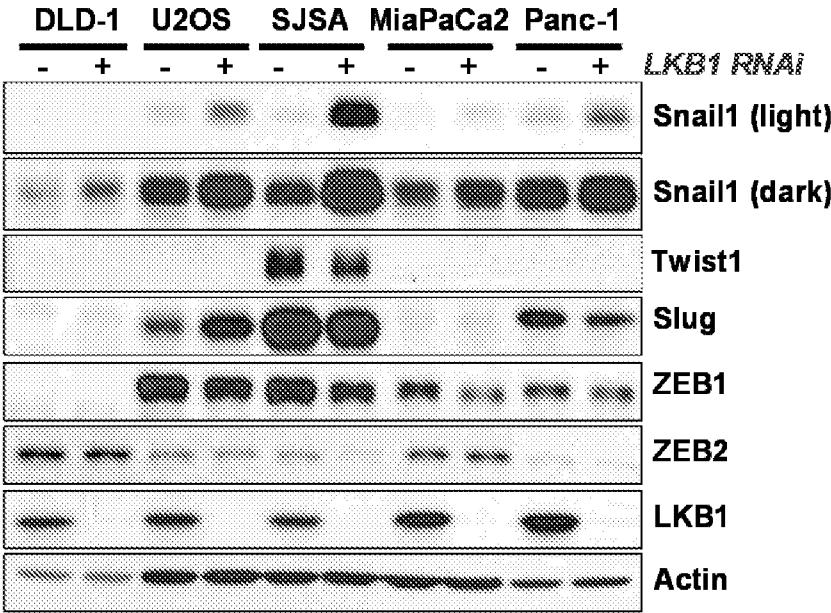


**FIG. 1I**

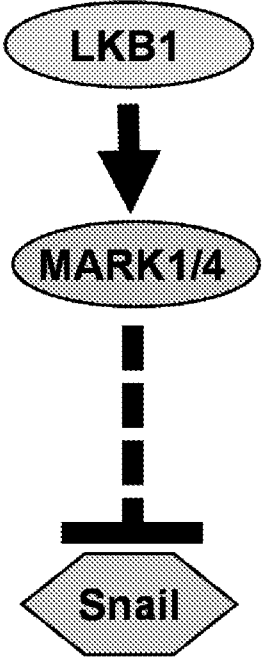
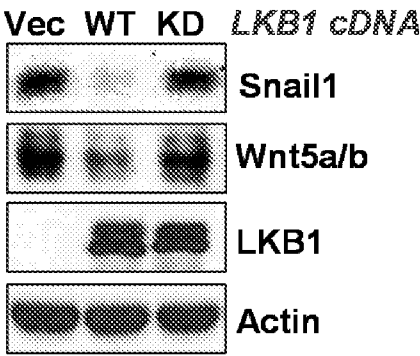
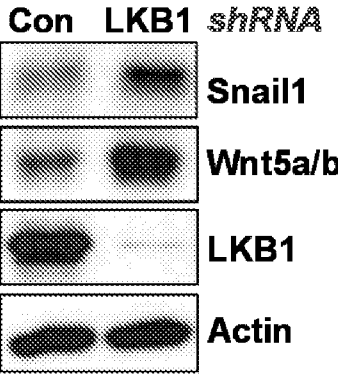


**FIG. 1B**



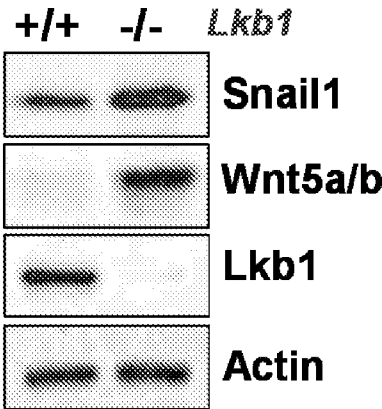
*H157 (LKB1<sup>-/-</sup>)*

**FIG. 1C**



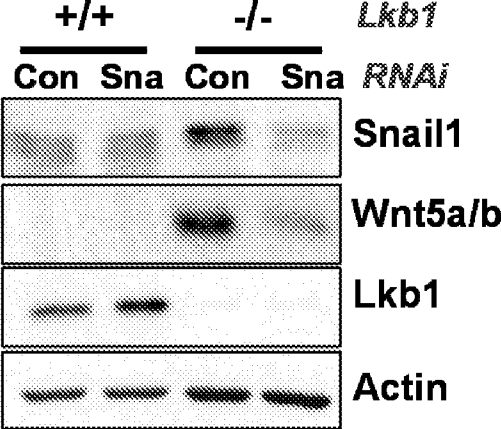
*H1299 (LKB1<sup>+/+</sup>)*

**FIG. 1D**



*MEFs*

**FIG. 1E**



*MEFs*

FIG. 1F

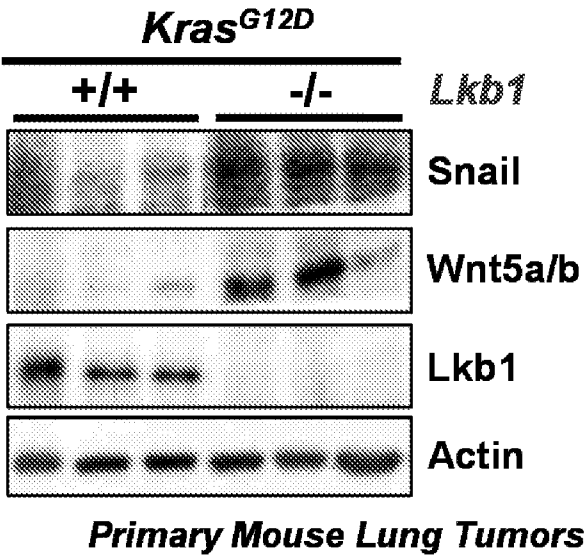


FIG. 1G

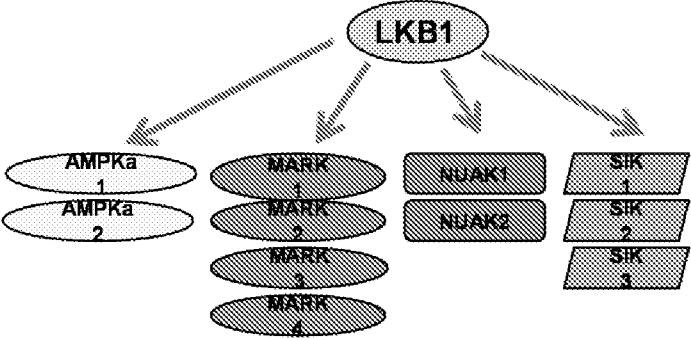
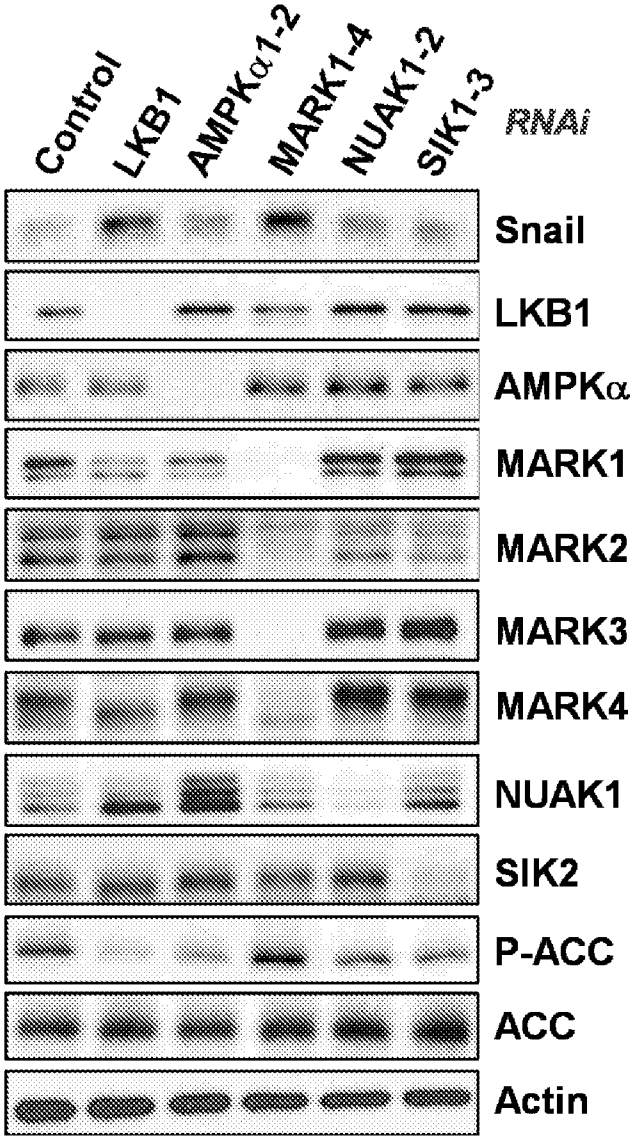
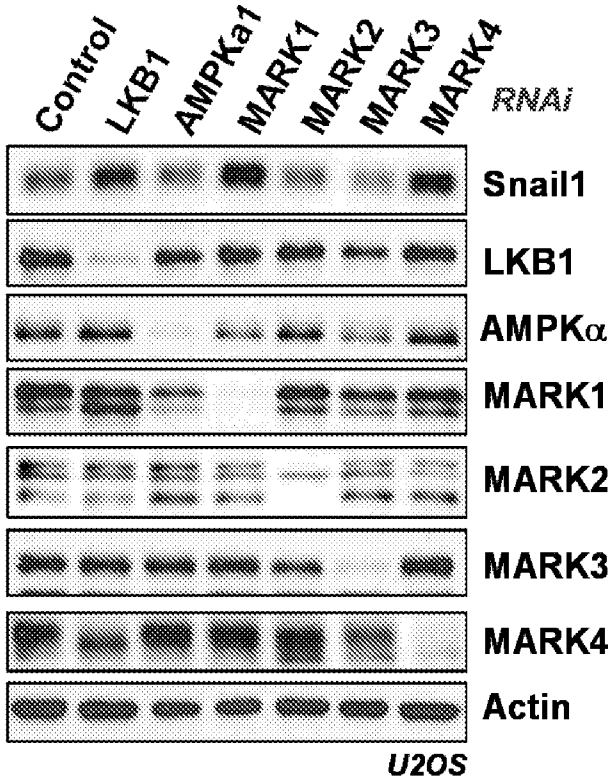


FIG. 1H



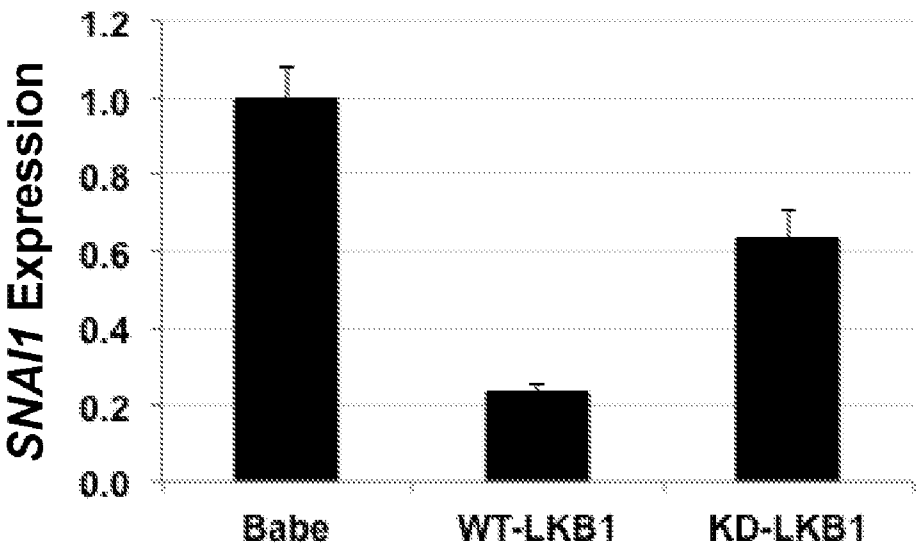


FIG. 2A

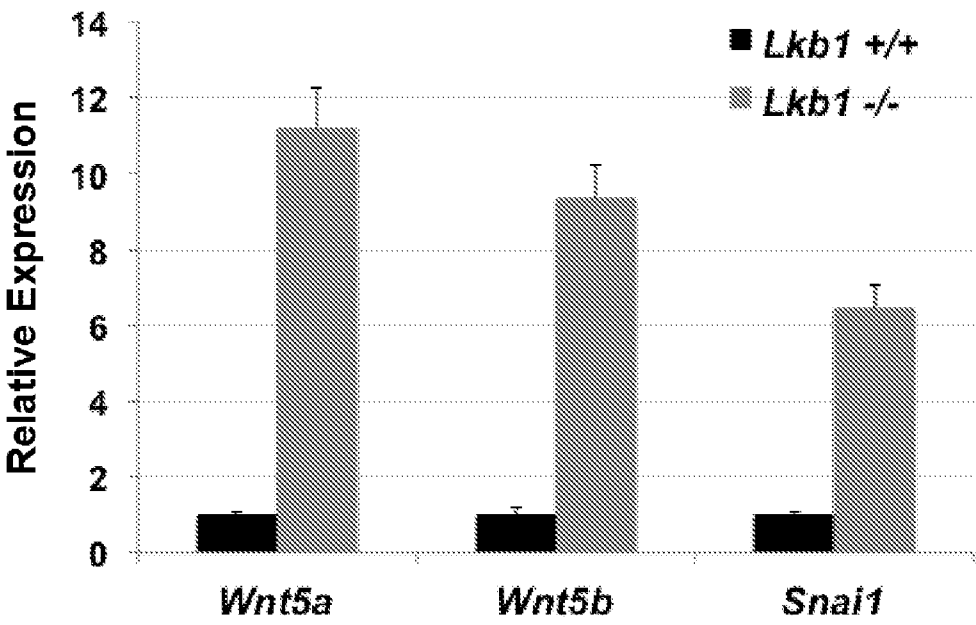
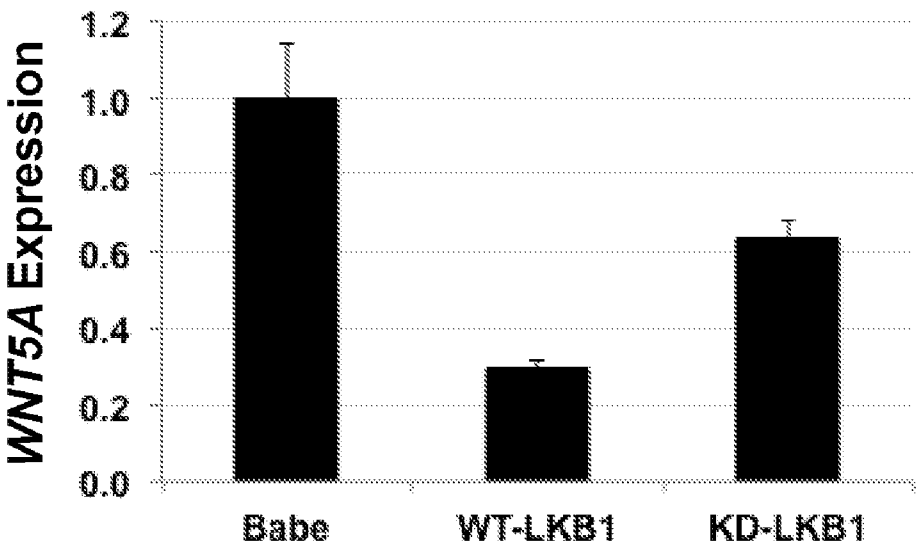


FIG. 2B

FIG. 2C

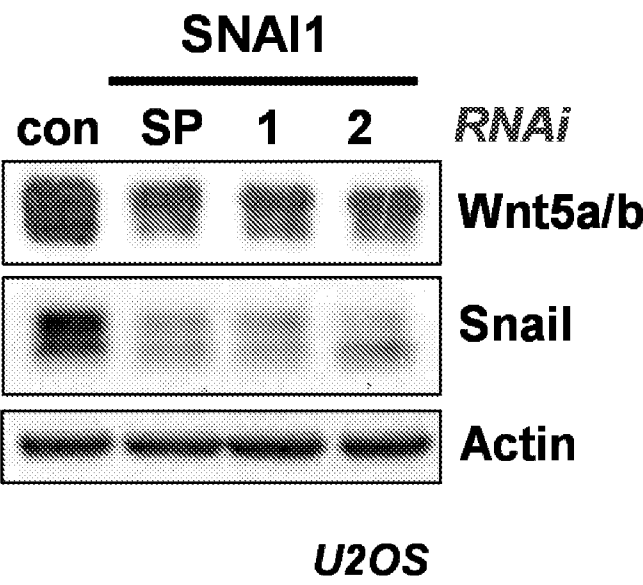


FIG. 2D

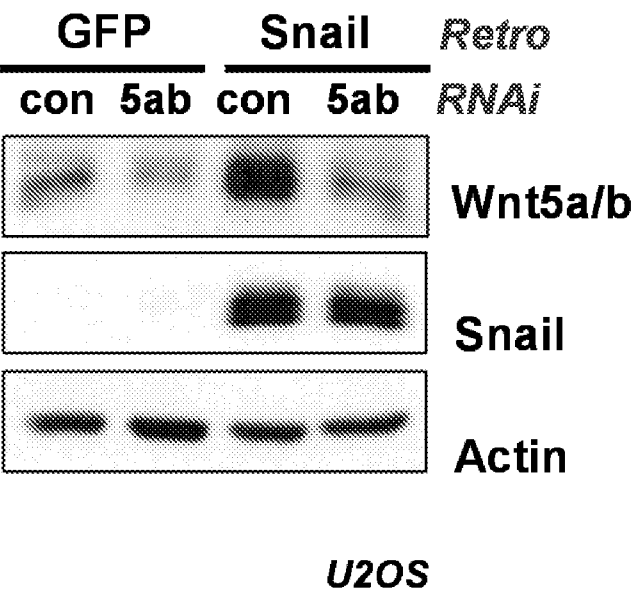


FIG. 2E

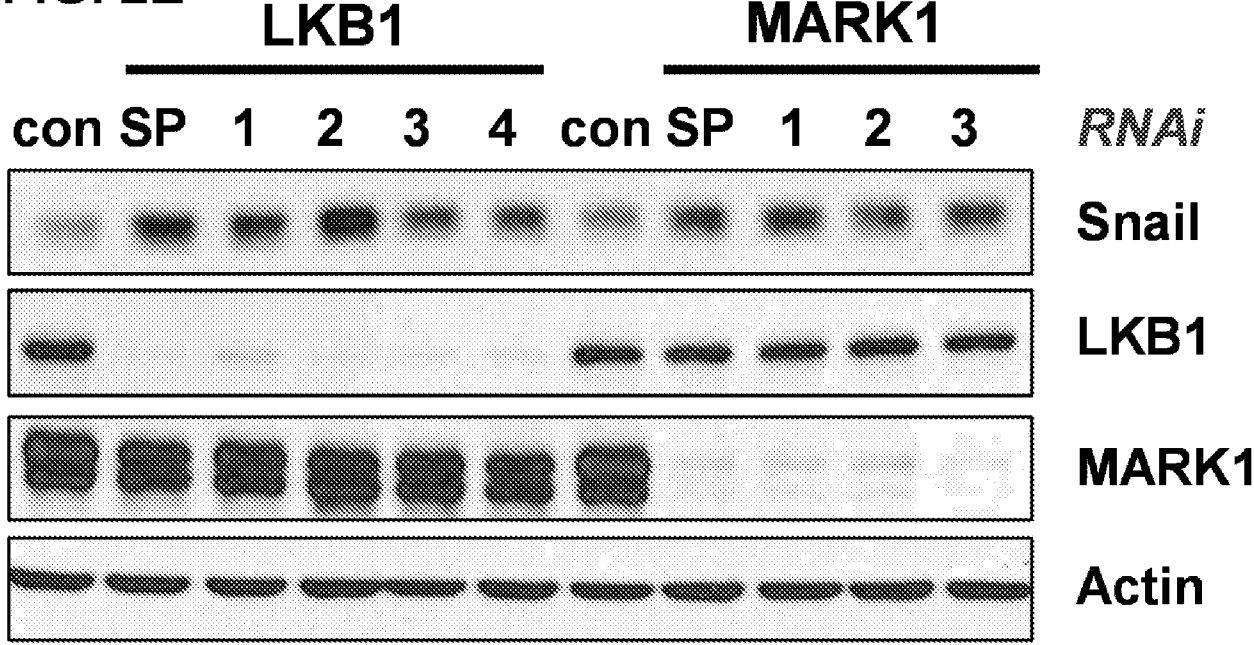
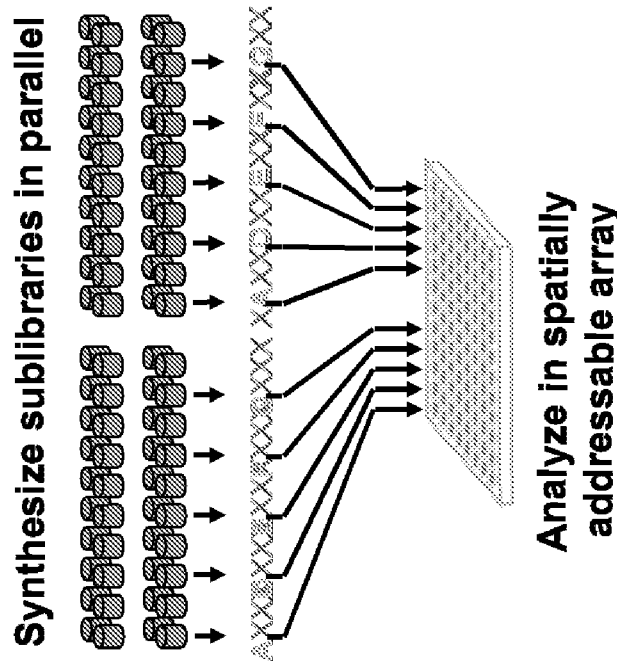


FIG. 3A



Y-A-Z-X-X-X-X-S/T-X-X-X-A-G-K-K(bioin)  
Y-A-X-Z-X-X-X-S/T-X-X-X-A-G-K-K(bioin)  
Y-A-X-X-Z-X-X-S/T-X-X-X-A-G-K-K(bioin)  
Y-A-X-X-X-Z-X-S/T-X-X-X-A-G-K-K(bioin)  
Y-A-X-X-X-X-Z-S/T-X-X-X-A-G-K-K(bioin)  
Y-A-X-X-X-X-X-S/T-Z-X-X-X-A-G-K-K(bioin)  
Y-A-X-X-X-X-X-S/T-X-Z-X-X-A-G-K-K(bioin)  
Y-A-X-X-X-X-X-S/T-X-X-Z-X-A-G-K-K(bioin)  
Y-A-X-X-X-X-X-S/T-X-X-X-Z-A-G-K-K(bioin)

-5 -4 -3 -2 -1 0 +1 +2 +3 +4

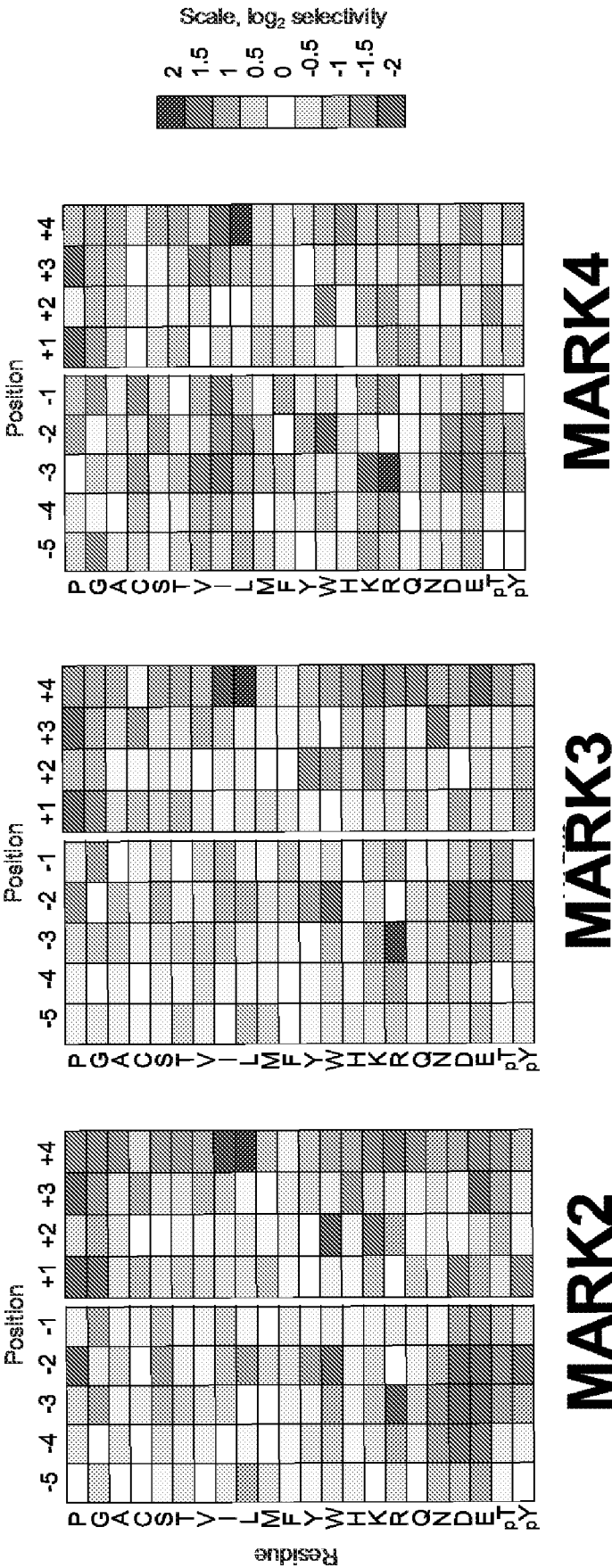


**FIG. 3B****Optimal MARK Motif**

Secondary selections

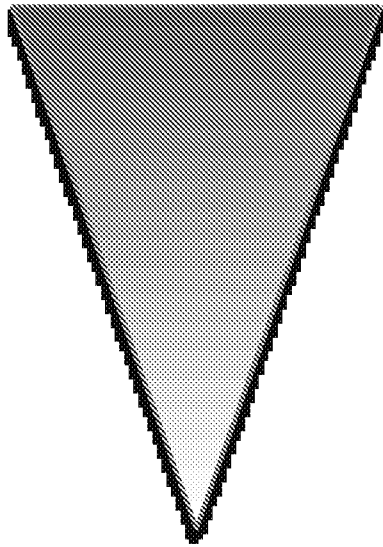
Additional selections

-5 -4 -3 -2 -1 0 +1 +2 +3 +4

L R P S G S x W N L

M K W P S x x D I

S x Y x S x x H M



**1000 human candidate MARK substrates  
based on presence of consensus sequence**

**300 conserved across upper eukaryotes**

**Functionally screened 25 annotated by  
“cytoskeletal control” GO terms**

**MARK substrates identified**

|                 |         |
|-----------------|---------|
| CLASP1          | Ser600  |
| CLASP2          | Ser603  |
| APC             | Ser2835 |
| KIF13B          | Ser1381 |
| DIXDC1          | Ser592  |
| IRSp53 (BAIAP2) | Ser366  |
| PAR3a           | Ser889  |
| MAP4            | Ser1000 |

L Q R S R S D I D V

L Q R S R S D I D V

S P R H S G S Y L

S R S I S S P N V

L P H S Q S S P T V

L P S S S M A A G

L K S S S L E S L

V S K K V S Y S H

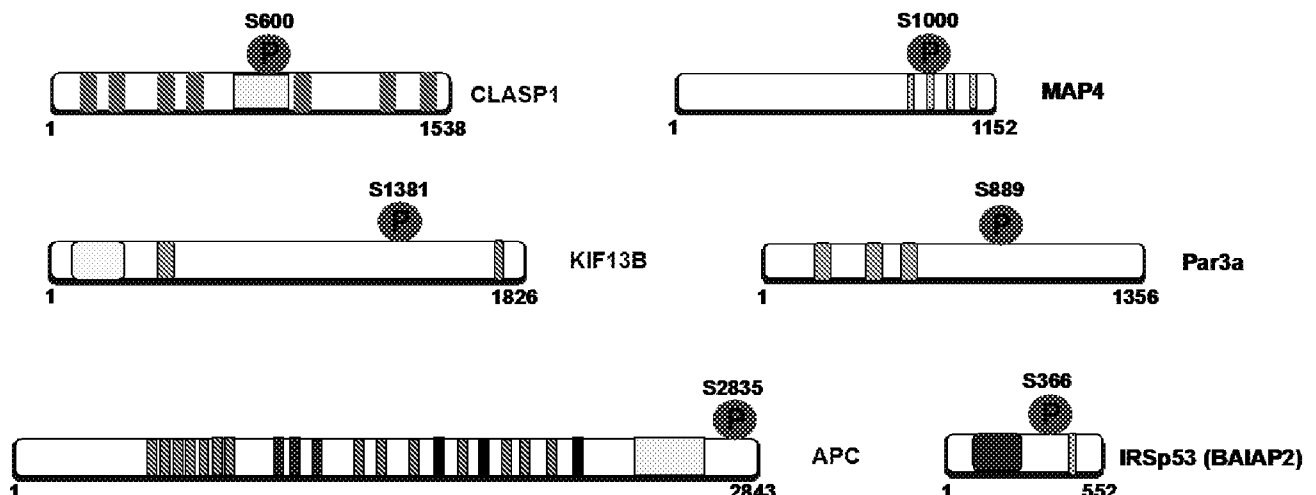


**FIG. 3C**

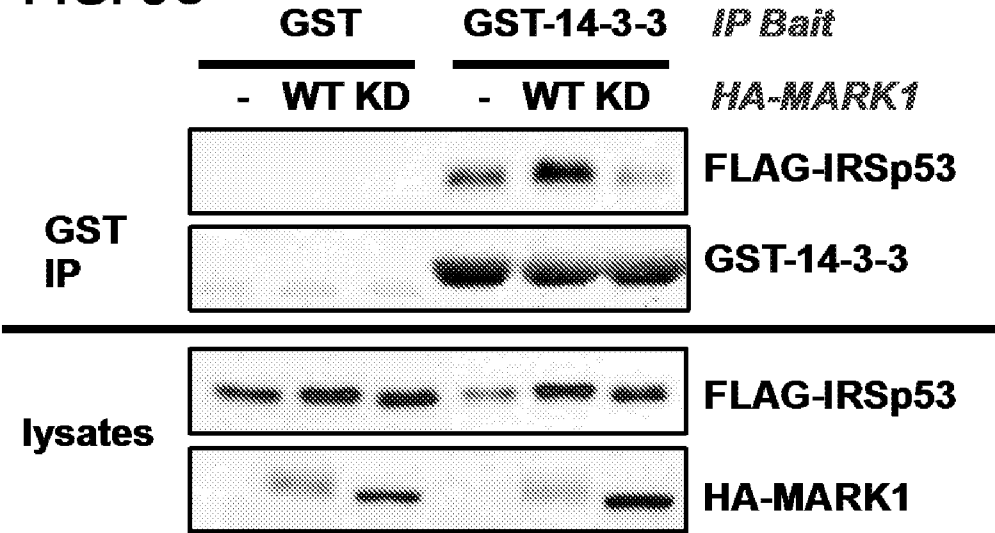


**FIG. 3D**

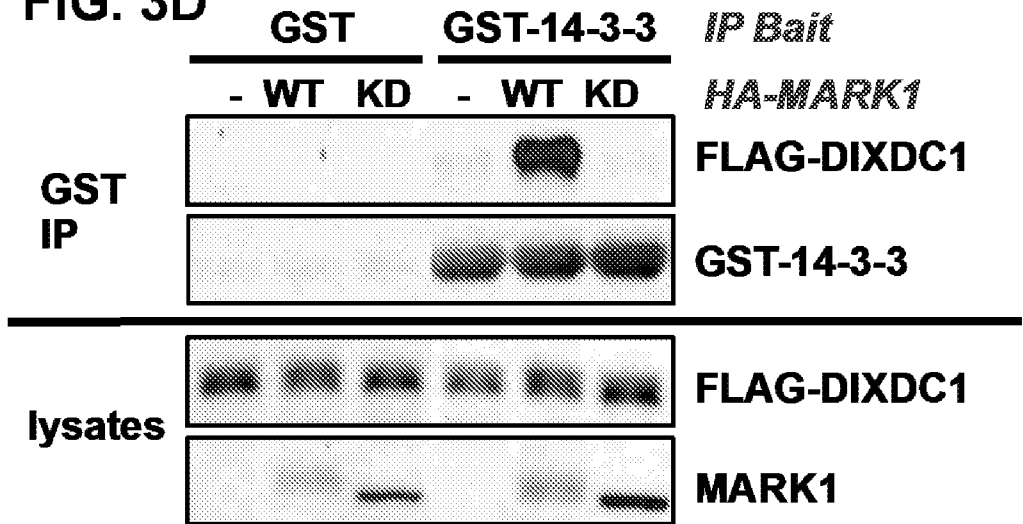


**FIG. 3E**

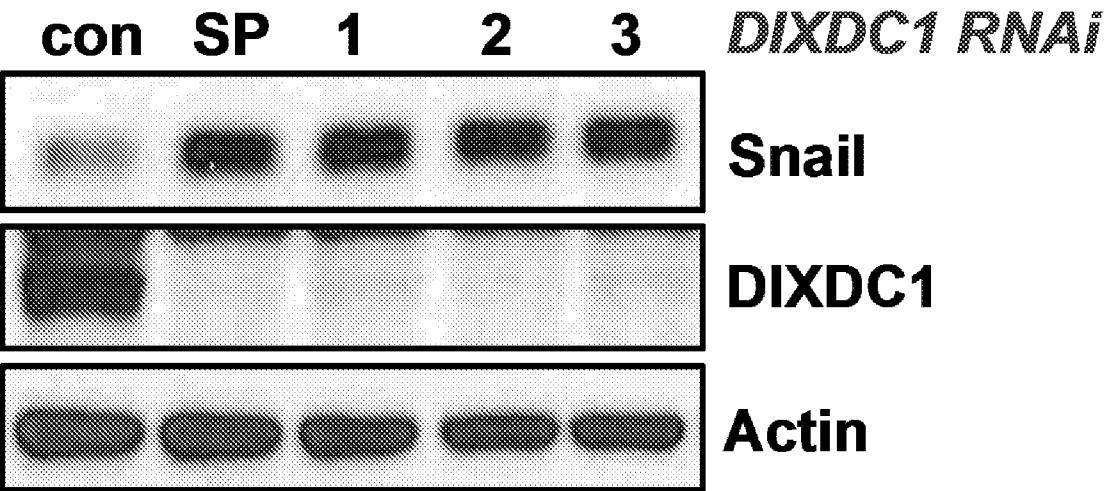


**FIG. 3F**

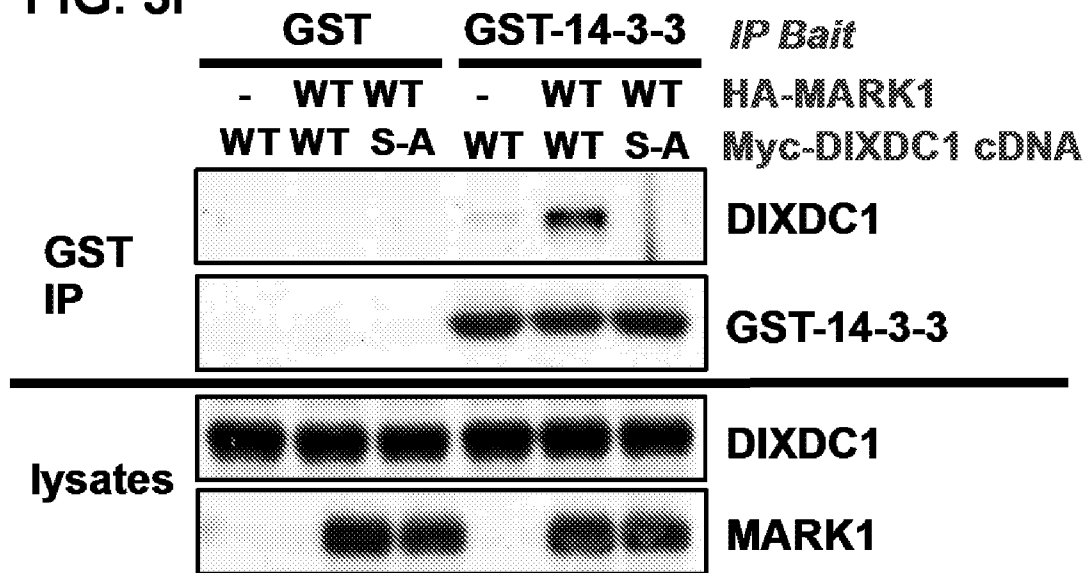


**FIG. 3G**

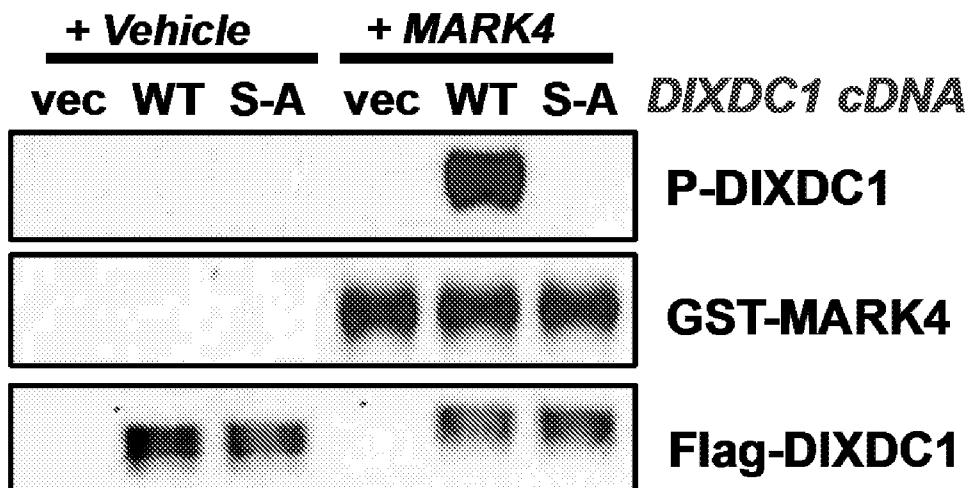


**FIG. 3H**

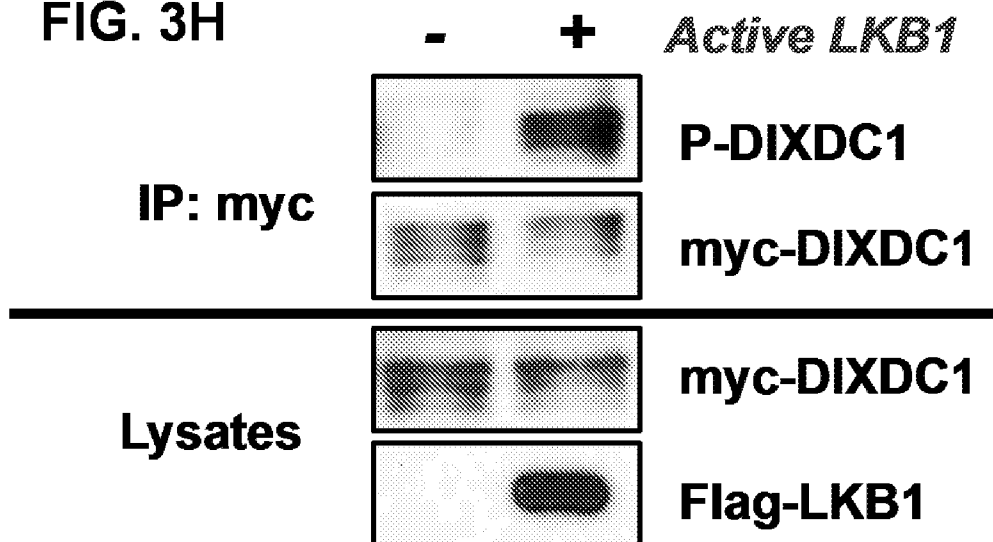


FIG. 3I

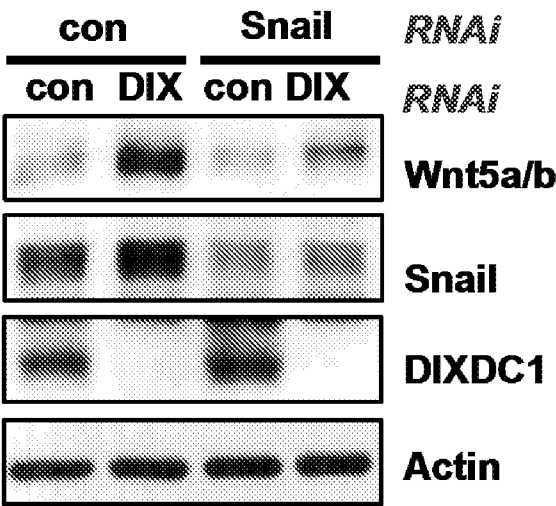


FIG. 4B

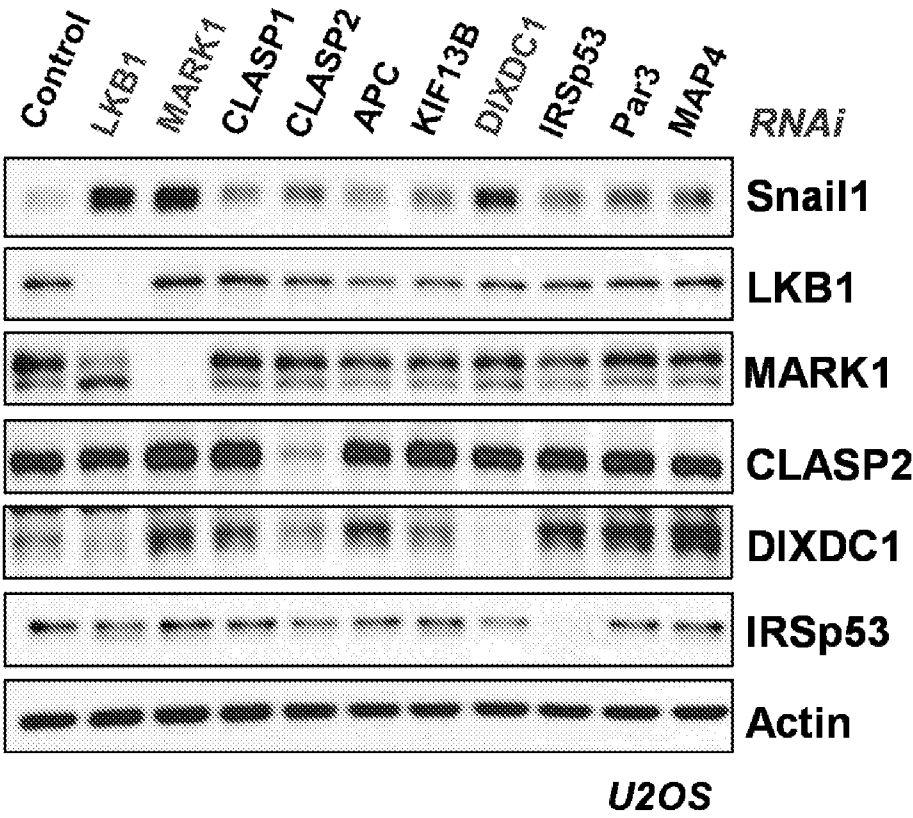


FIG. 4C

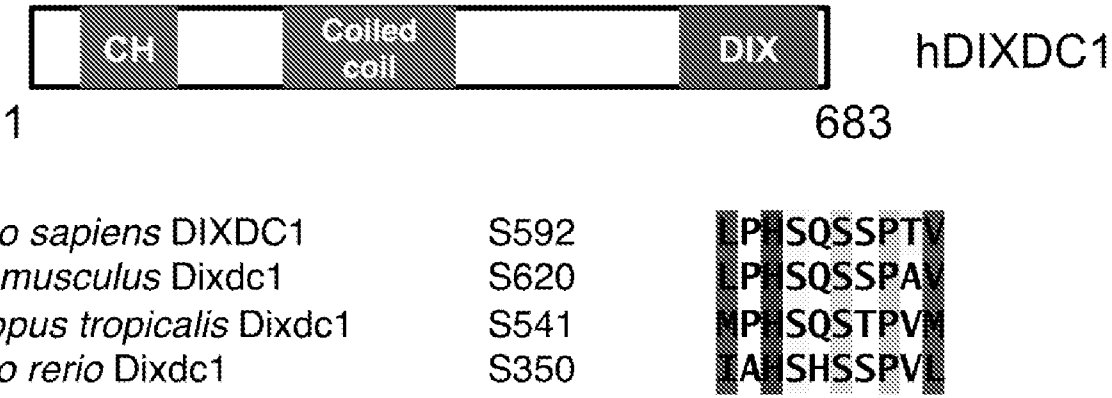


FIG. 4A

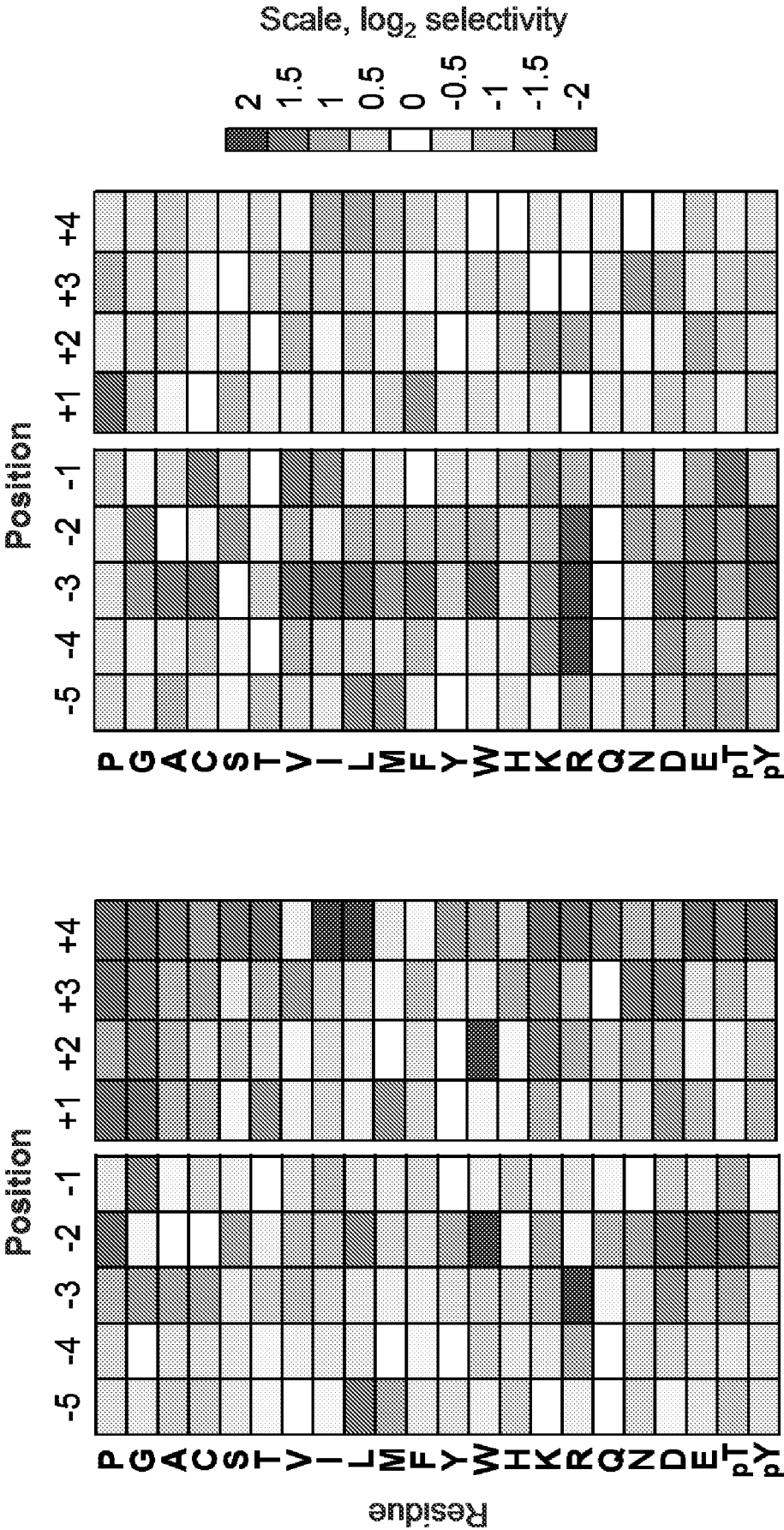


FIG. 4D

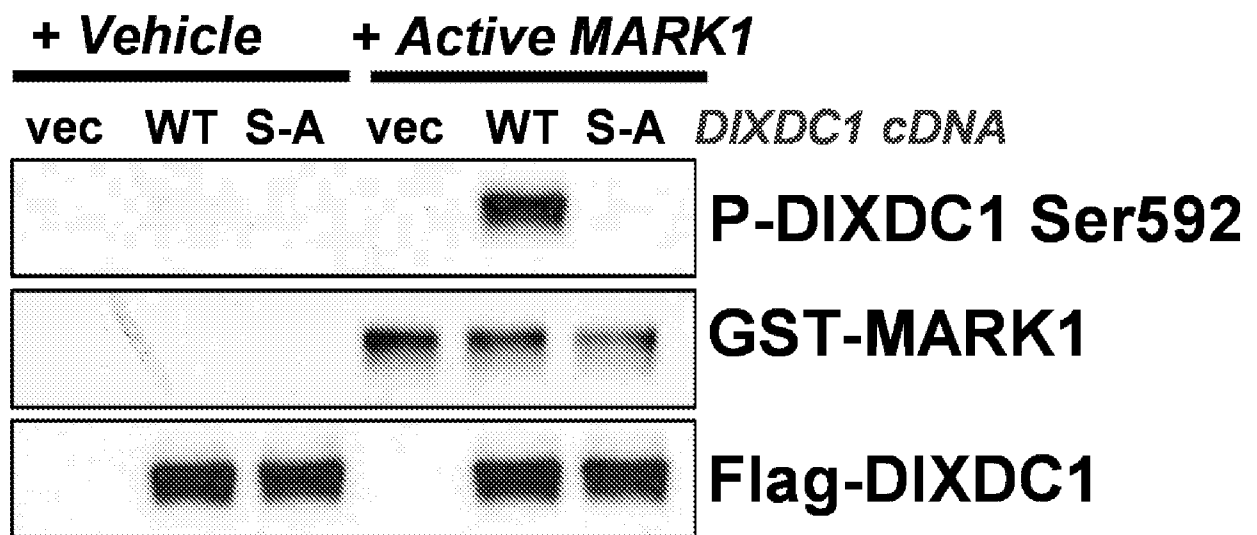
*In Vitro Kinase Assay*

FIG. 4E



**FIG. 4F**

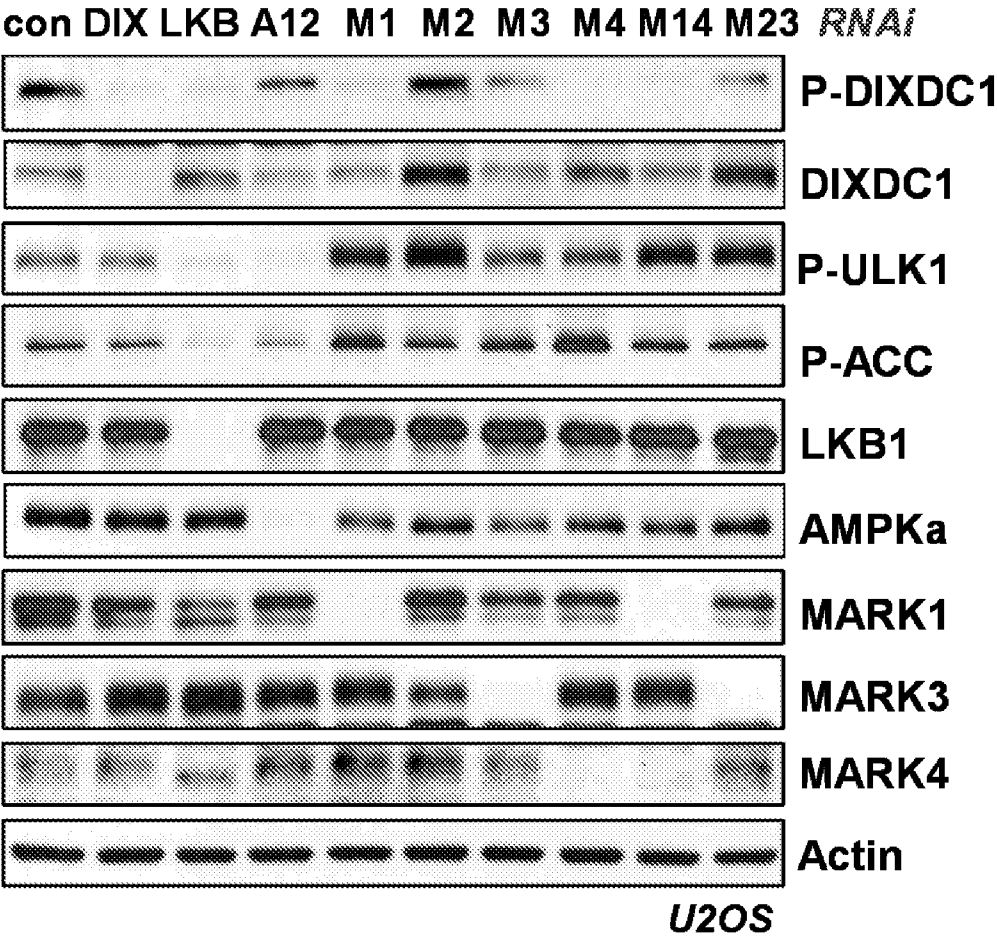


**FIG. 4G**

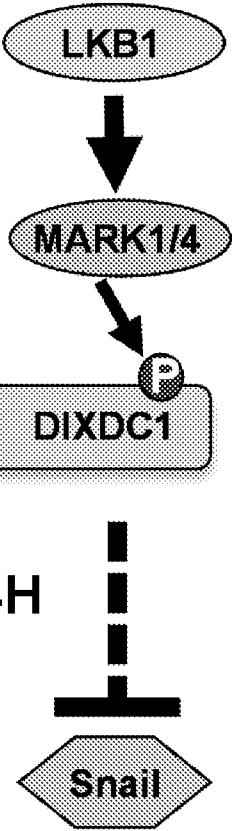
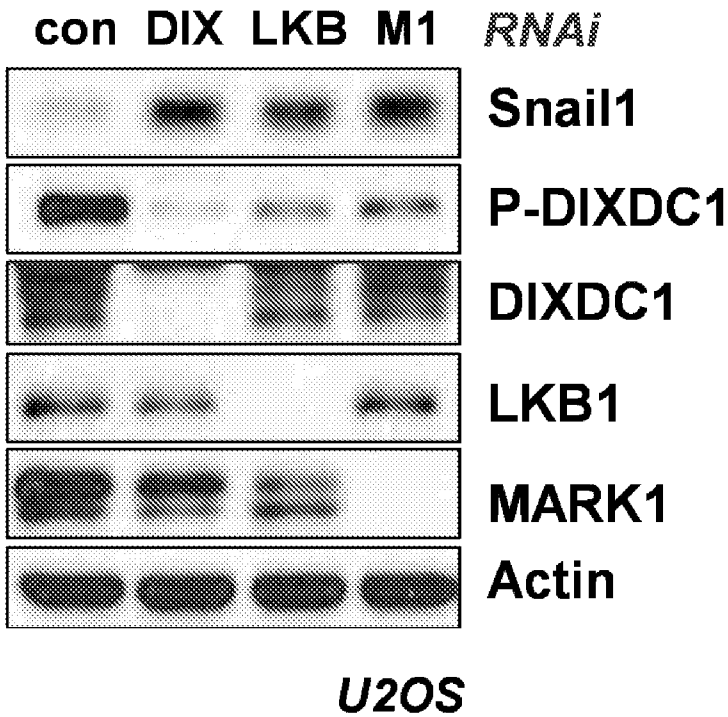


**FIG. 4H**

FIG. 5A

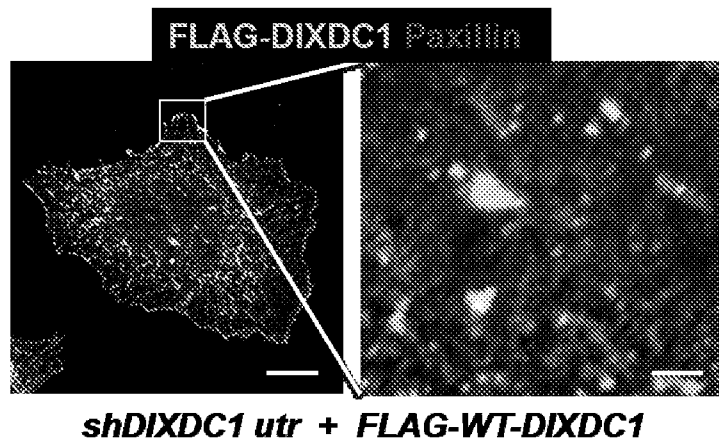


FIG. 5B

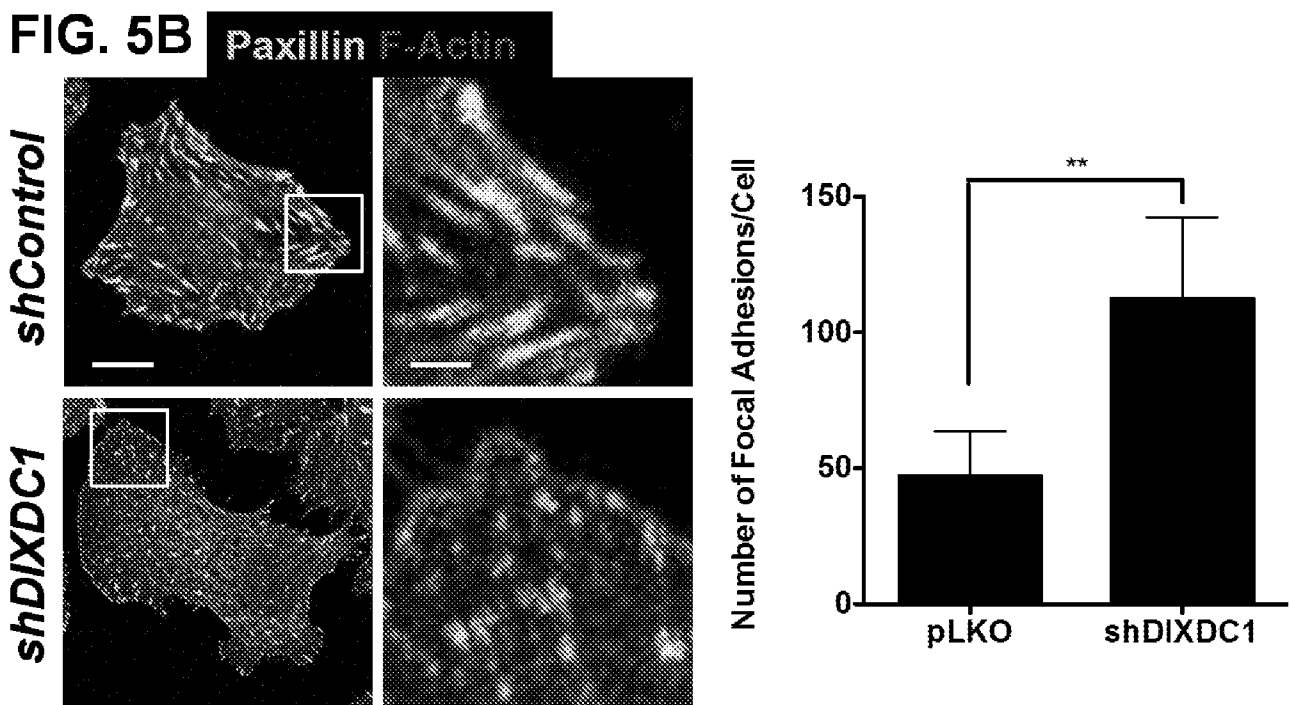


FIG. 5C

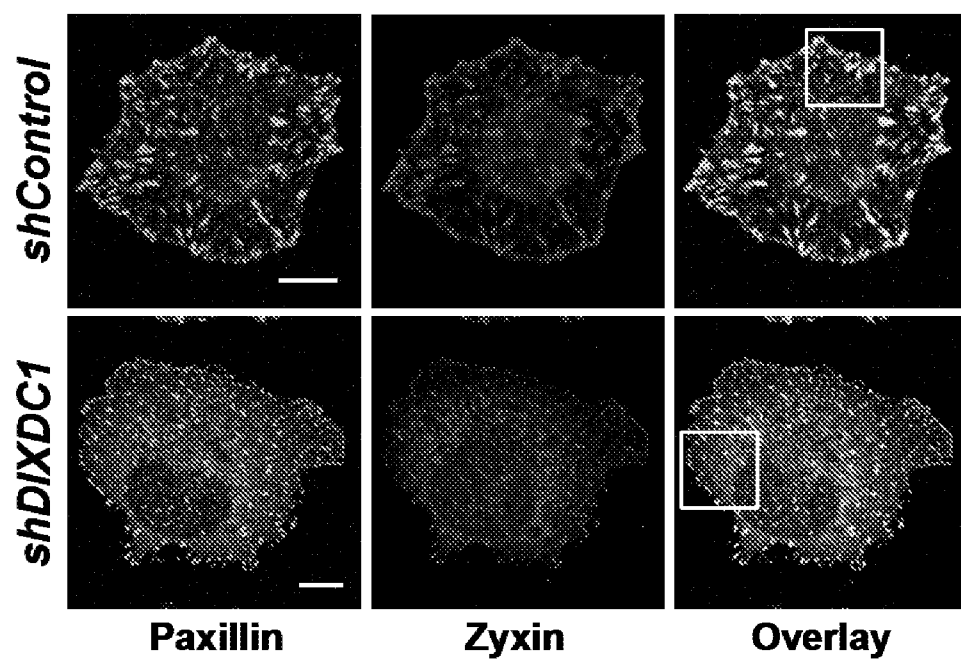
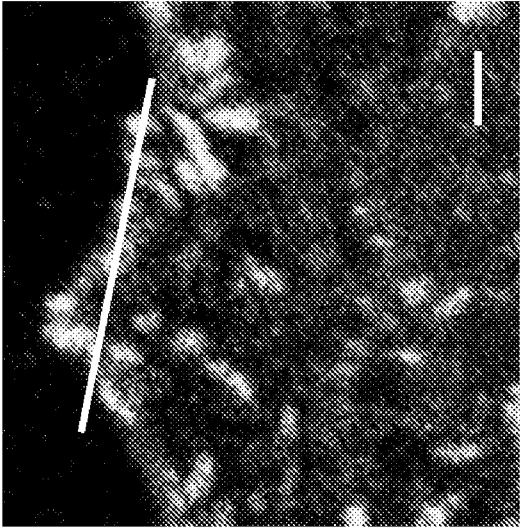
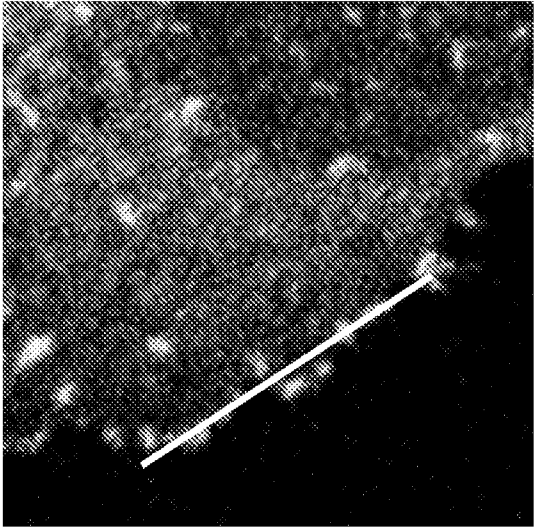


**FIG. 5D**



***shControl***



***shDIXDC1***

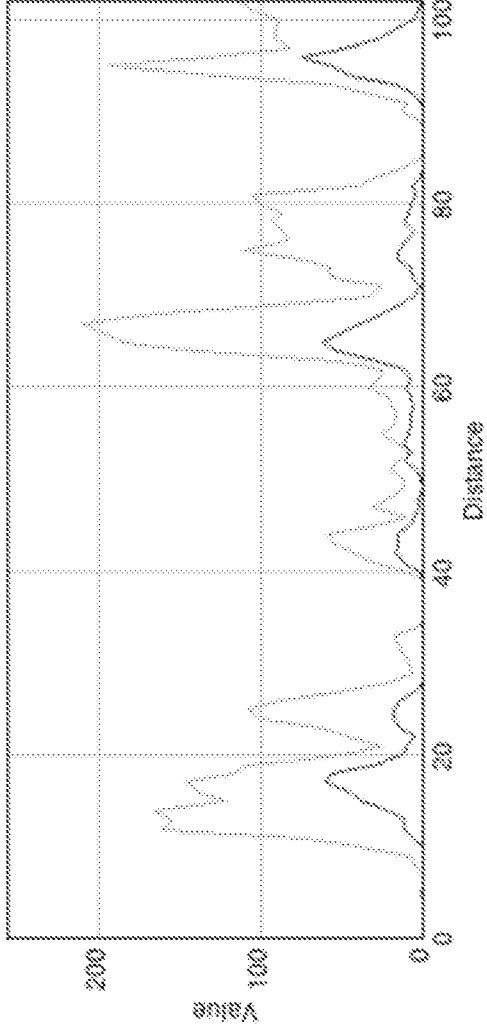
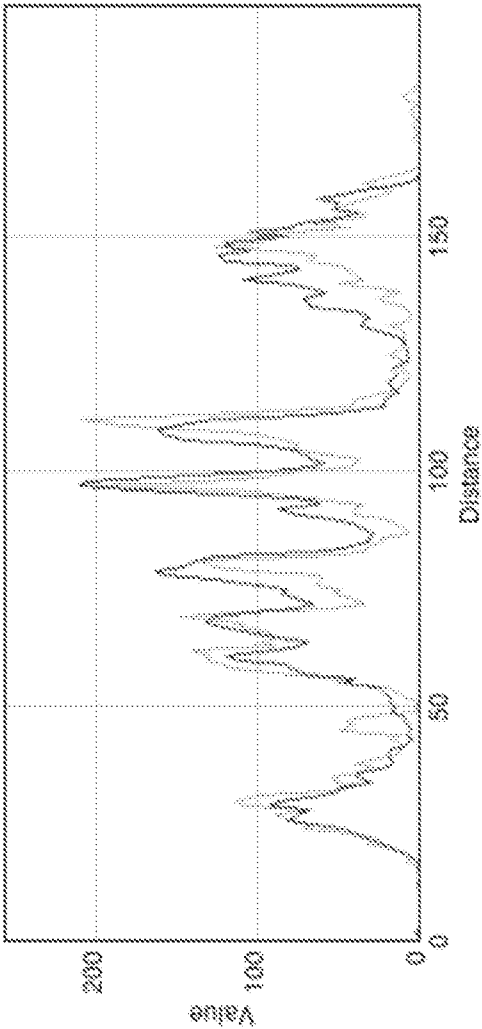


**FIG. 5F**

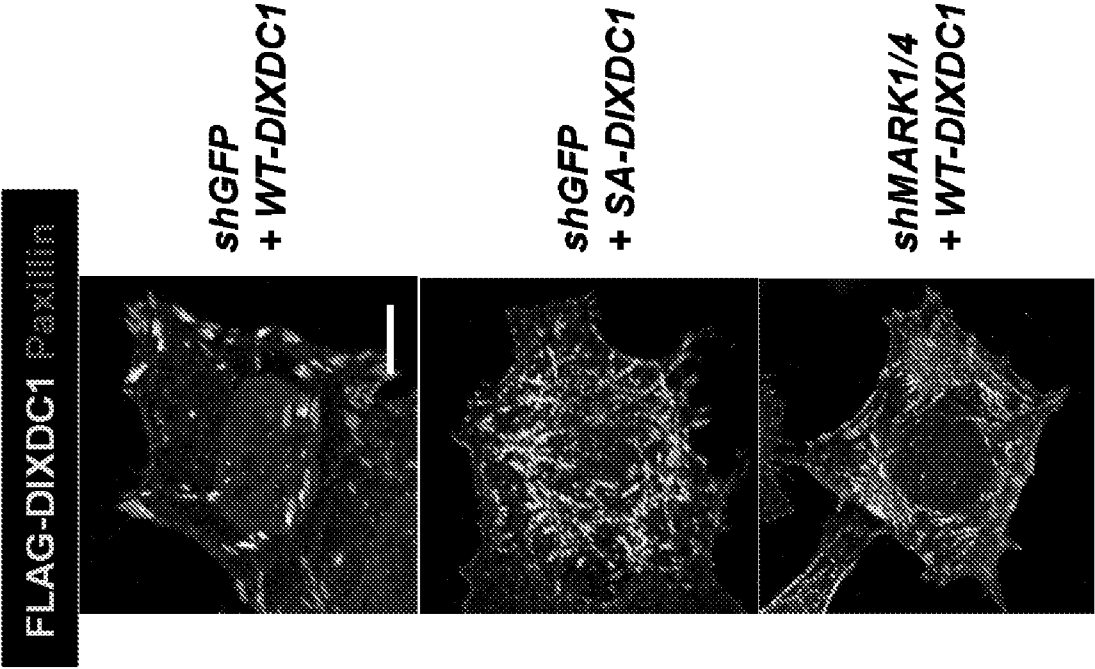


**FIG. 5E**

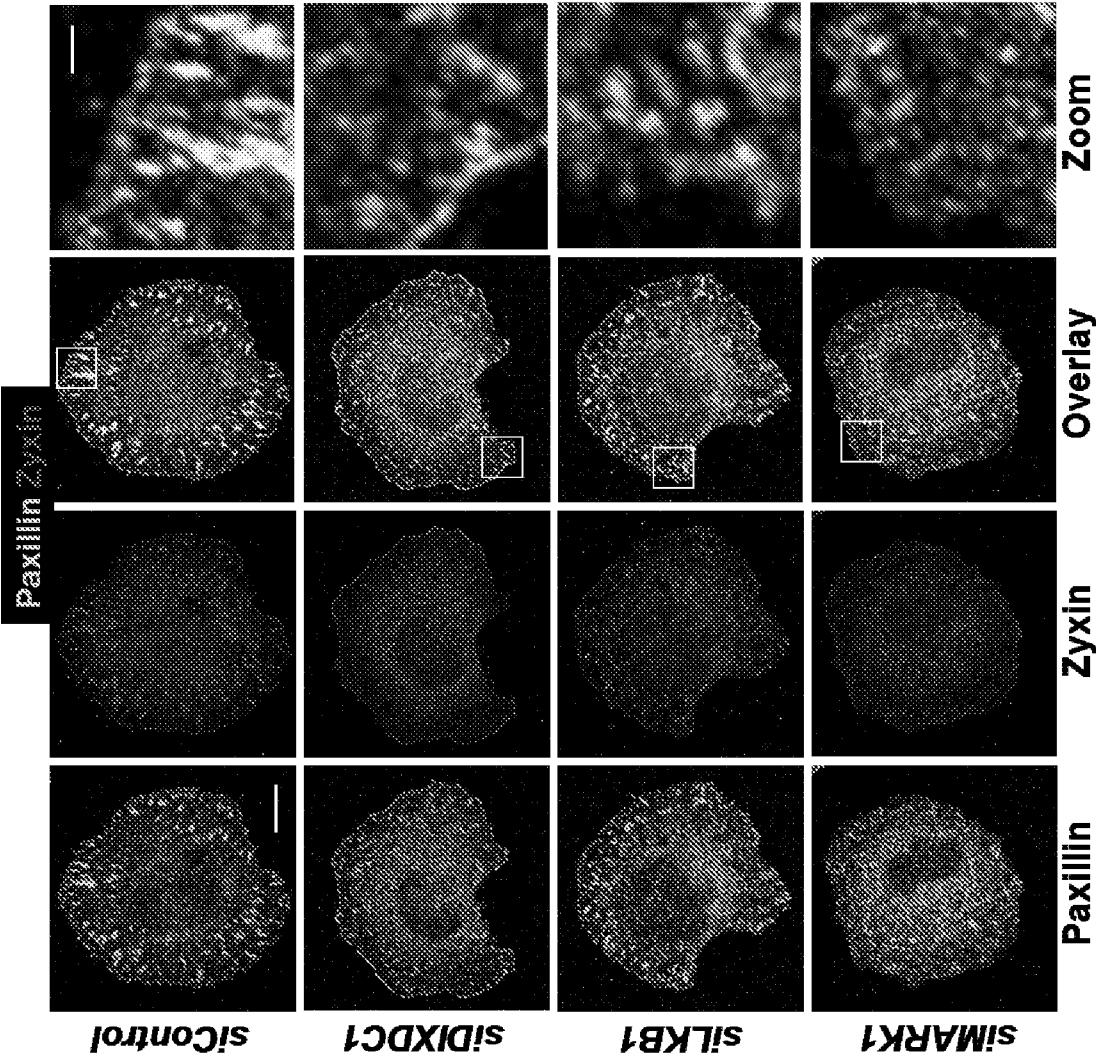


FIG. 5G

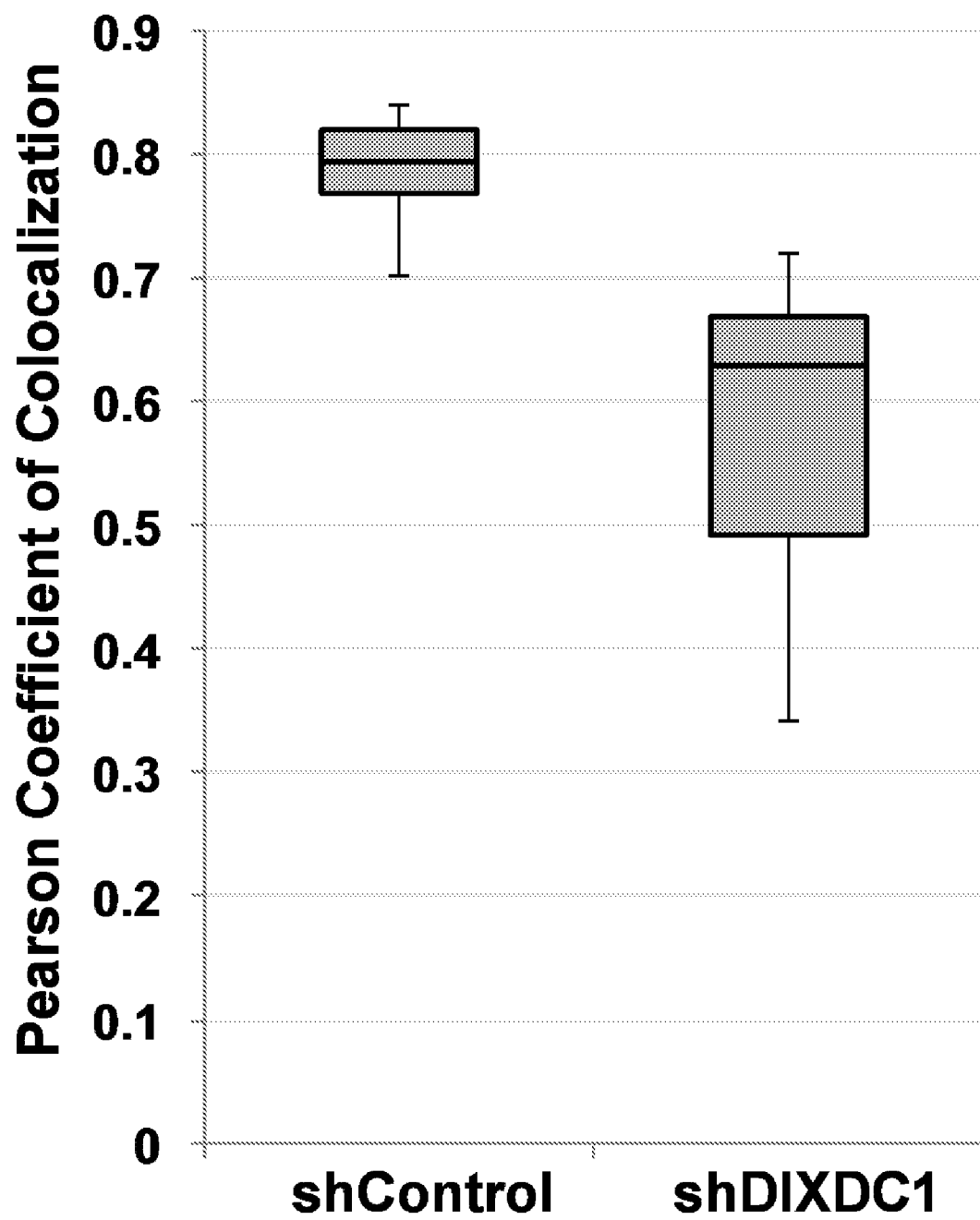


**FIG. 6A**

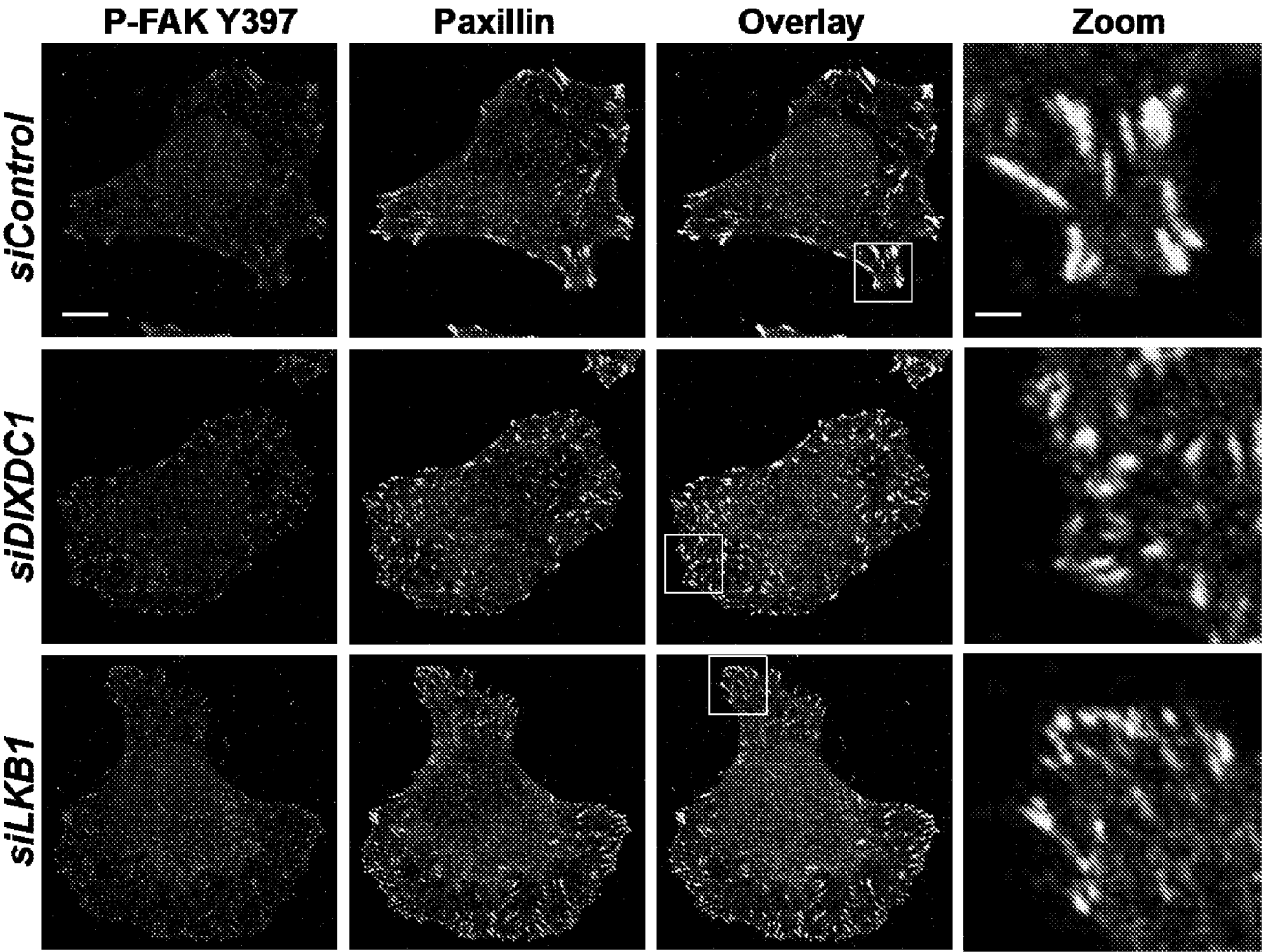


**FIG. 6B**

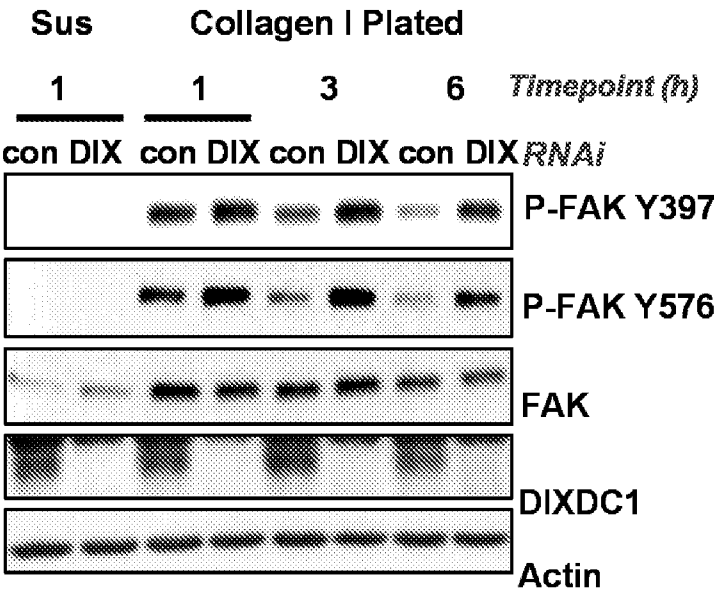


**FIG. 6F**

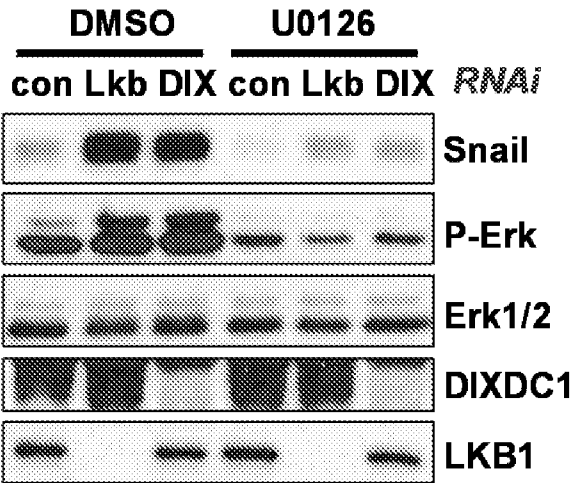
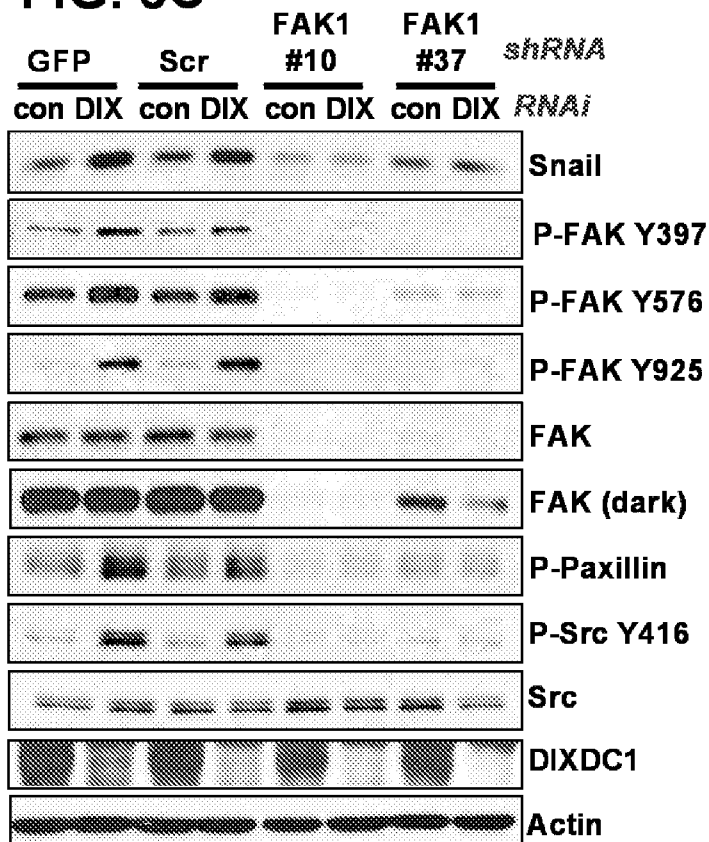
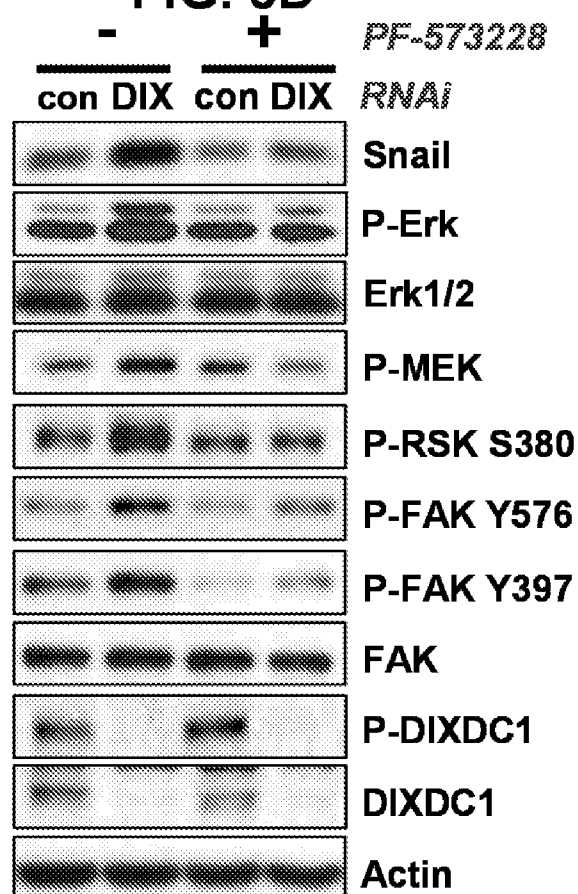
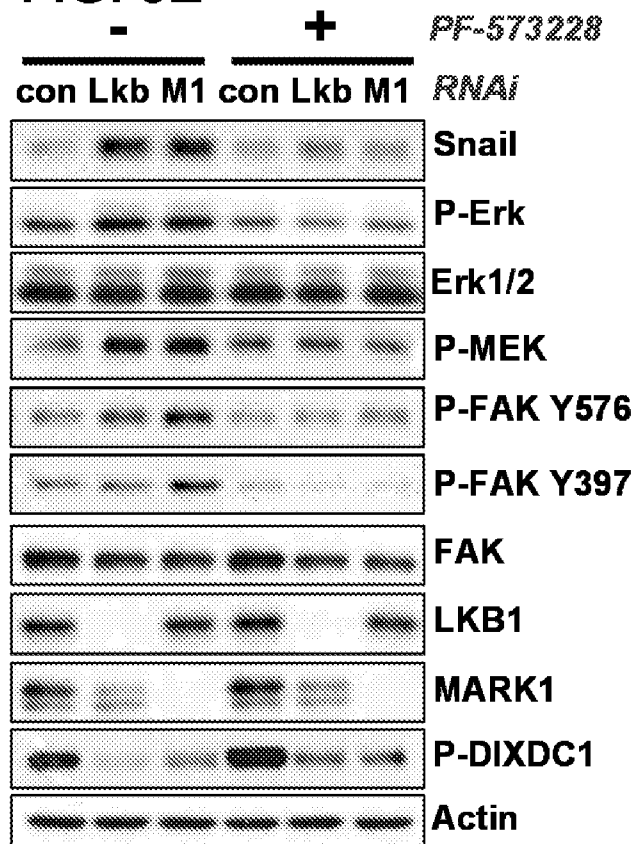
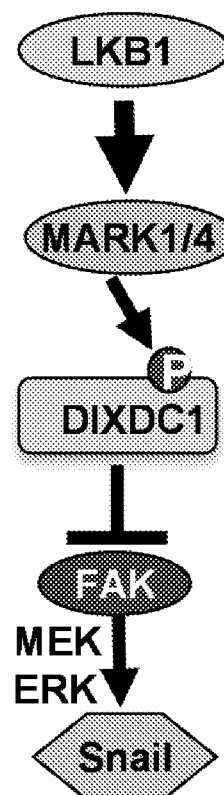


**FIG. 6C****FIG. 6D****FIG. 6E****FIG. 6G**

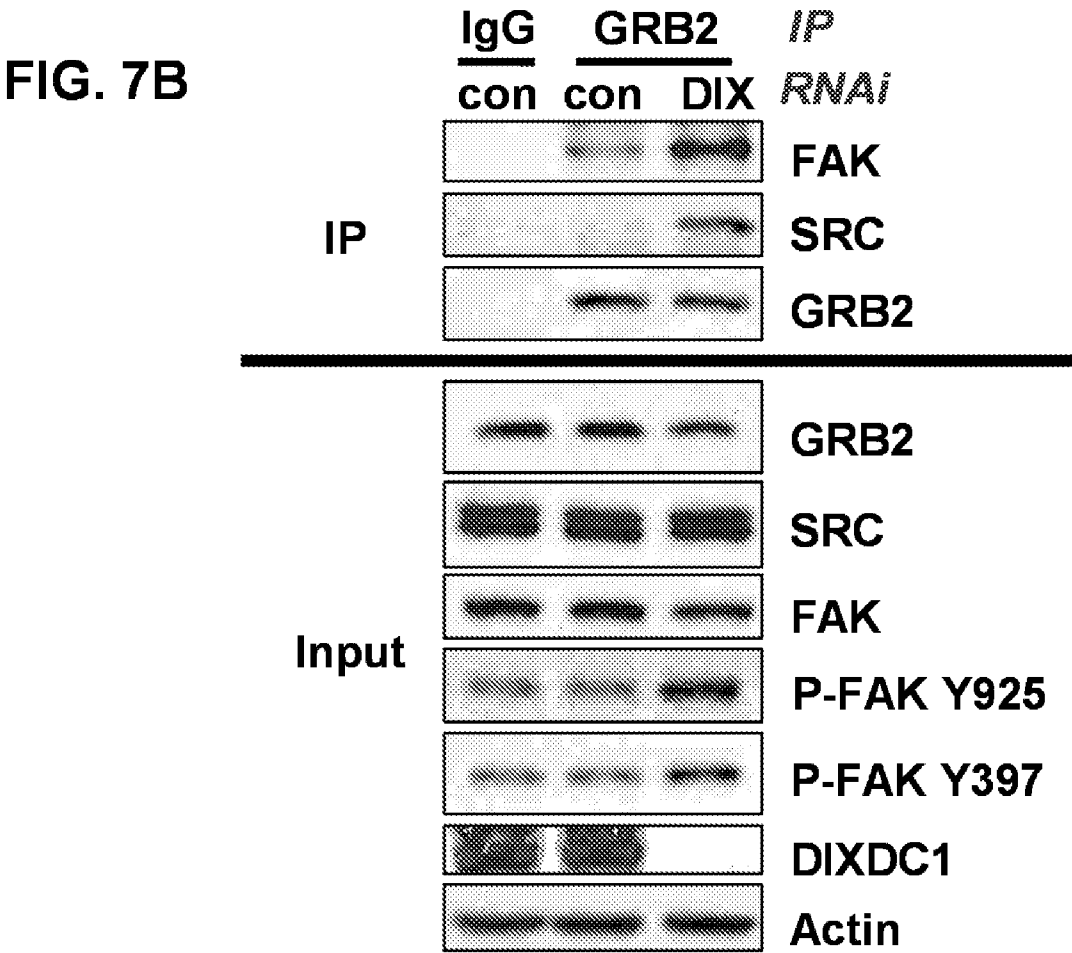
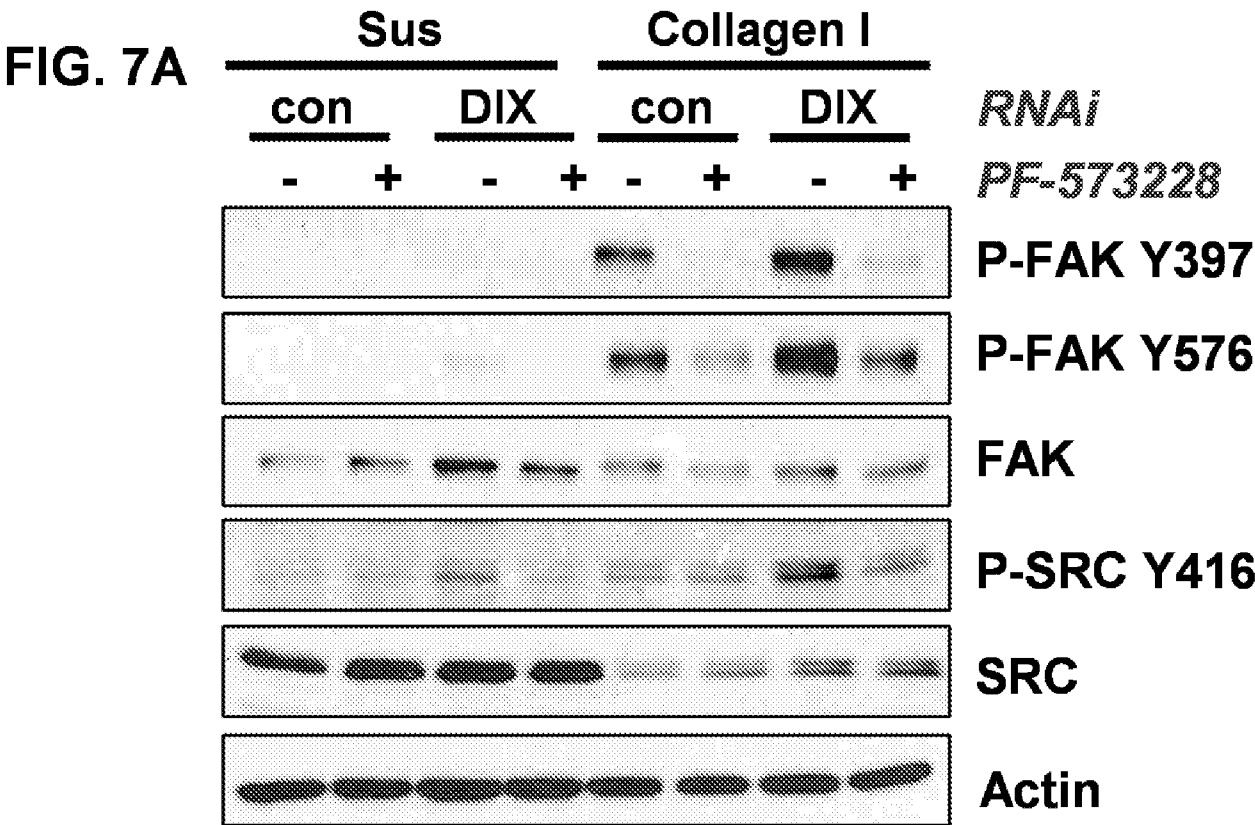


FIG. 7C

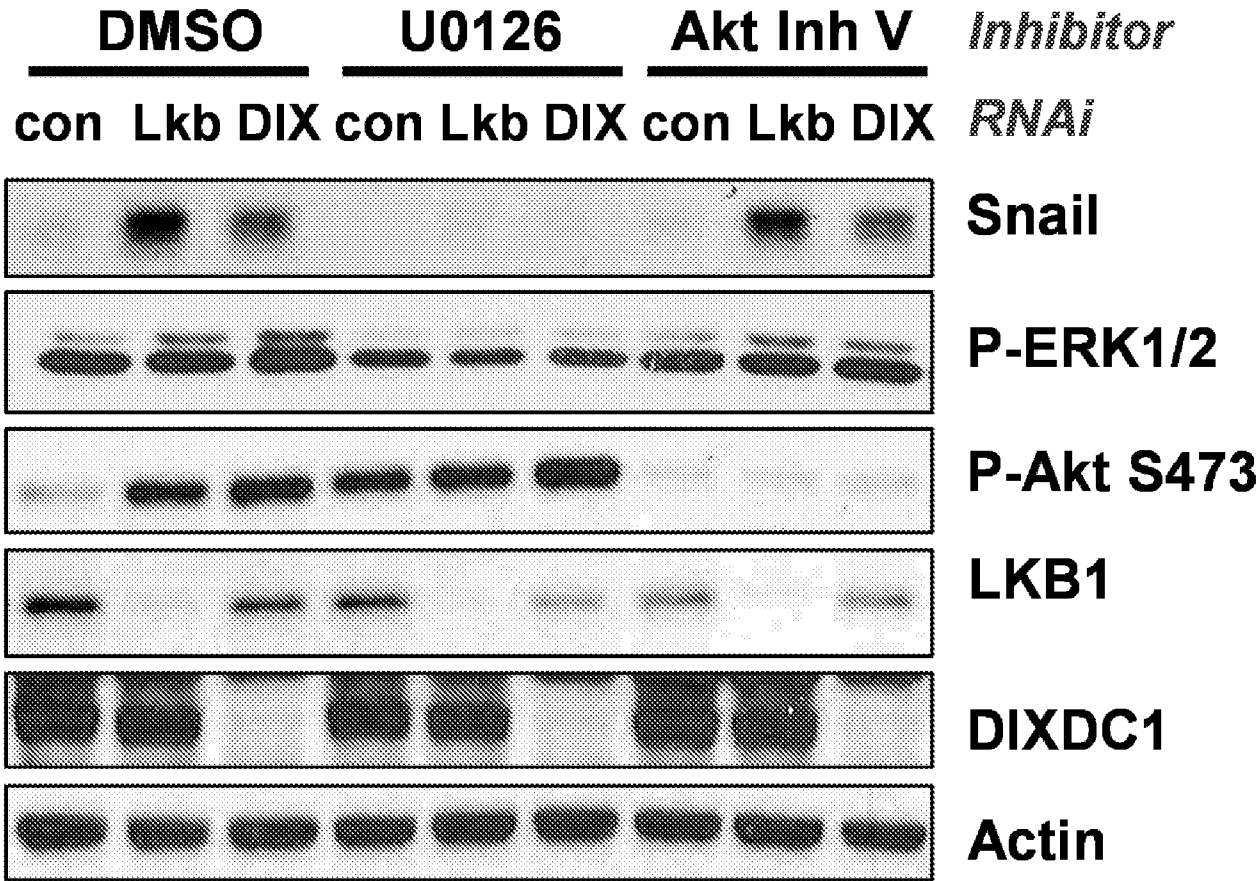


FIG. 7D

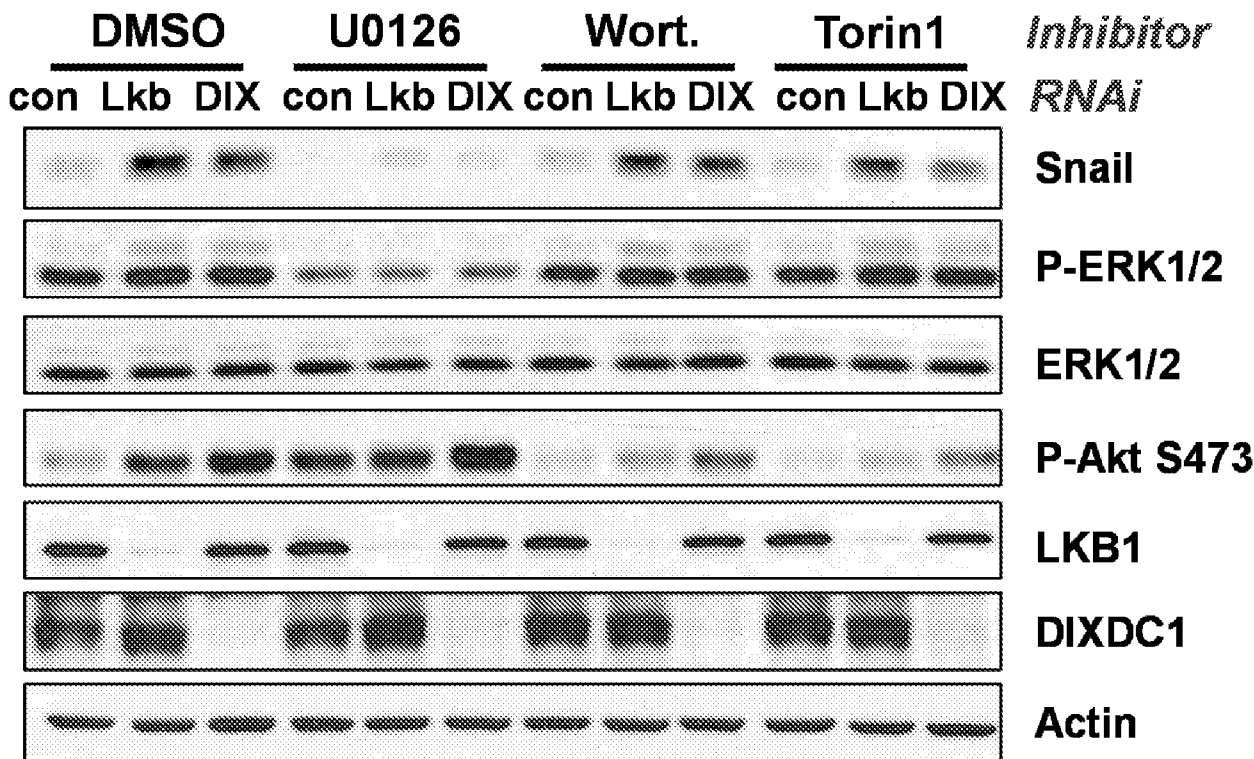


**FIG. 8A**

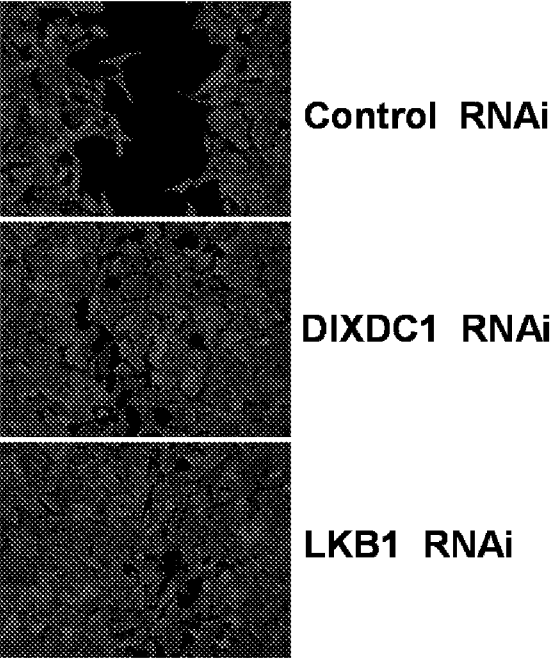


**FIG. 8C**

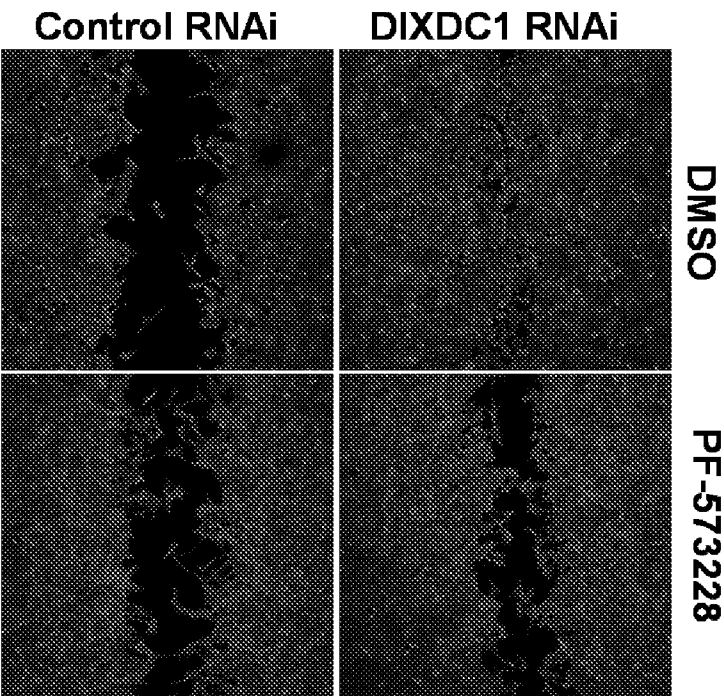


**FIG. 8B**

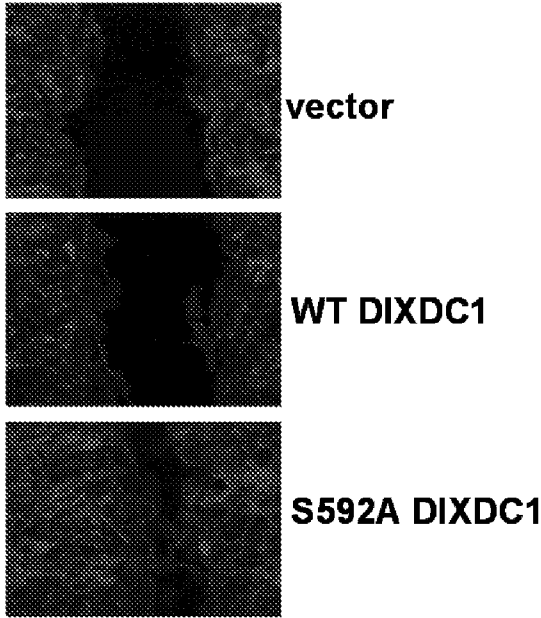


**FIG. 8D**

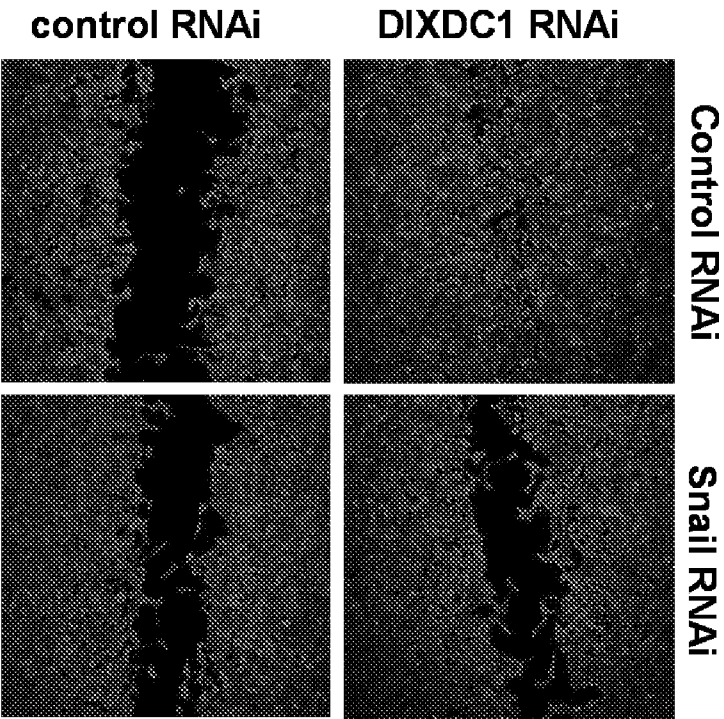


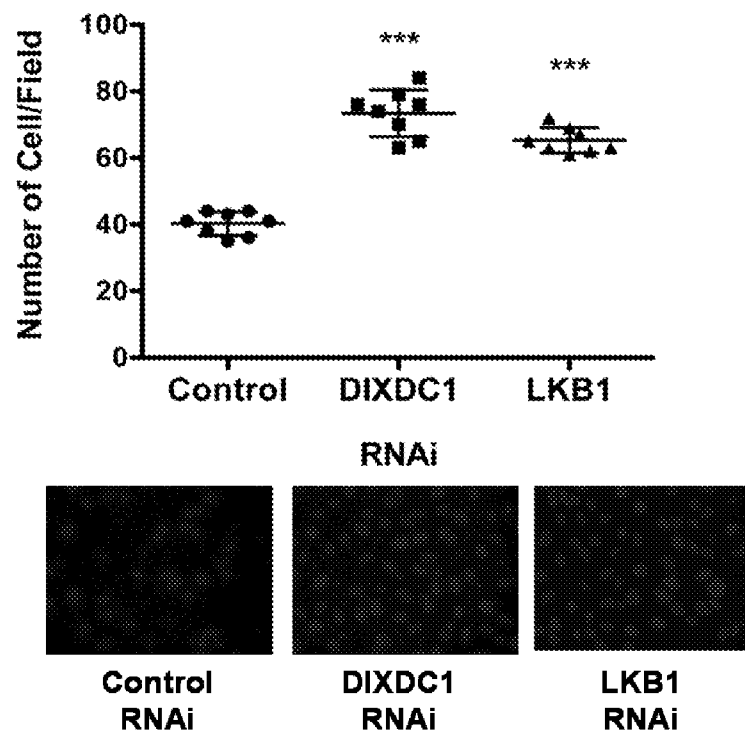
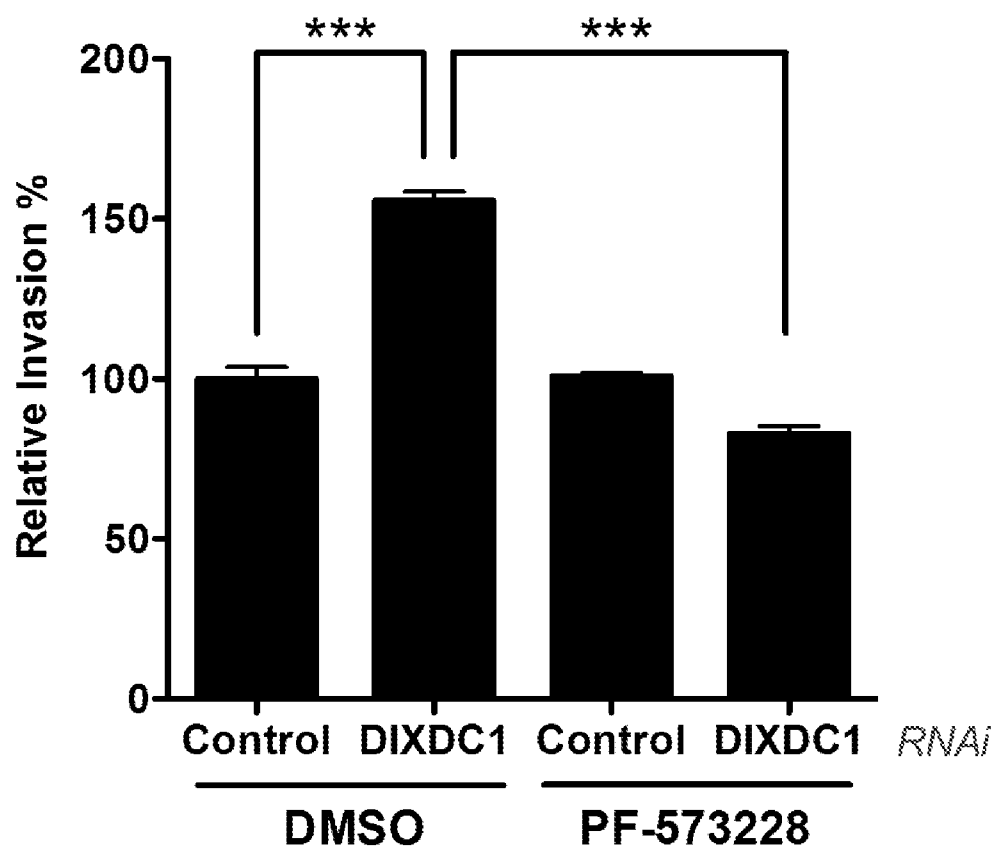
**FIG. 8E***Collagen I Transwell Invasion Assay***FIG. 9**

**FIG. 10A**

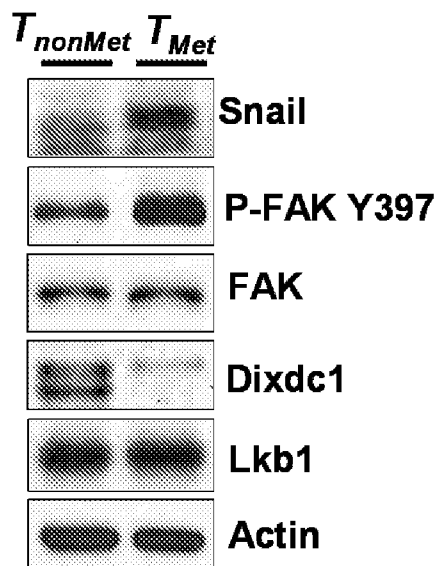
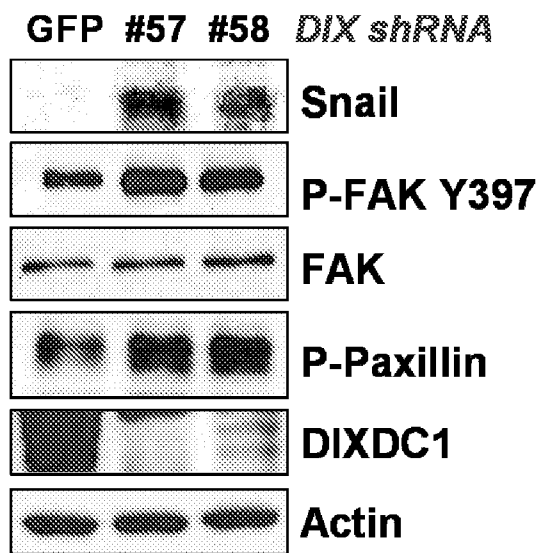


**FIG. 10B**



$T_{nonMet}$

**FIG. 10C**

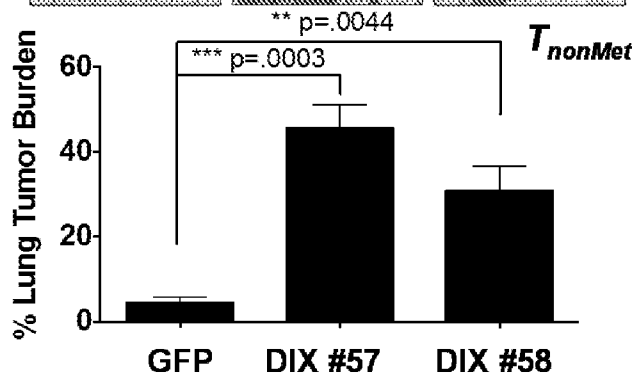
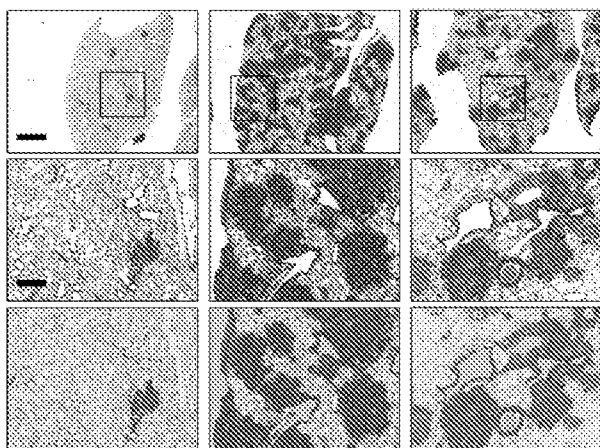


**FIG. 10F**

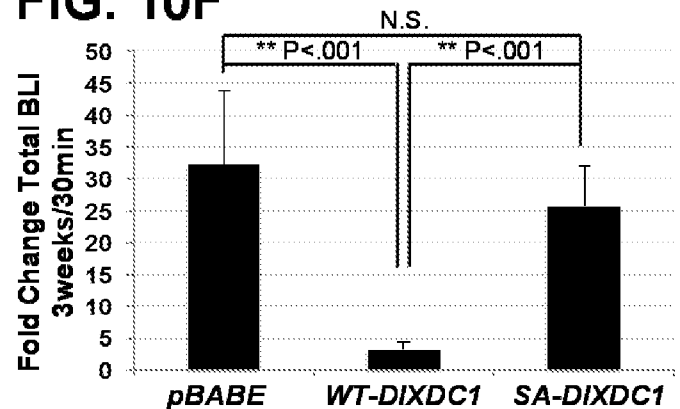


**FIG. 10G**

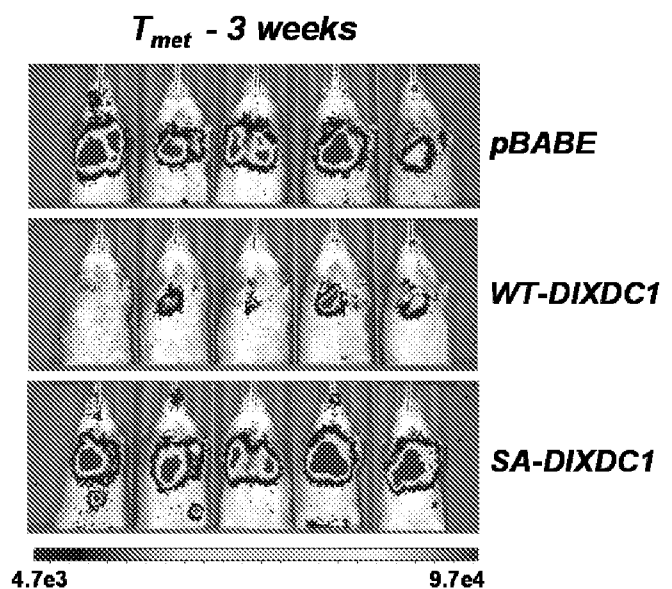


**FIG. 10D**

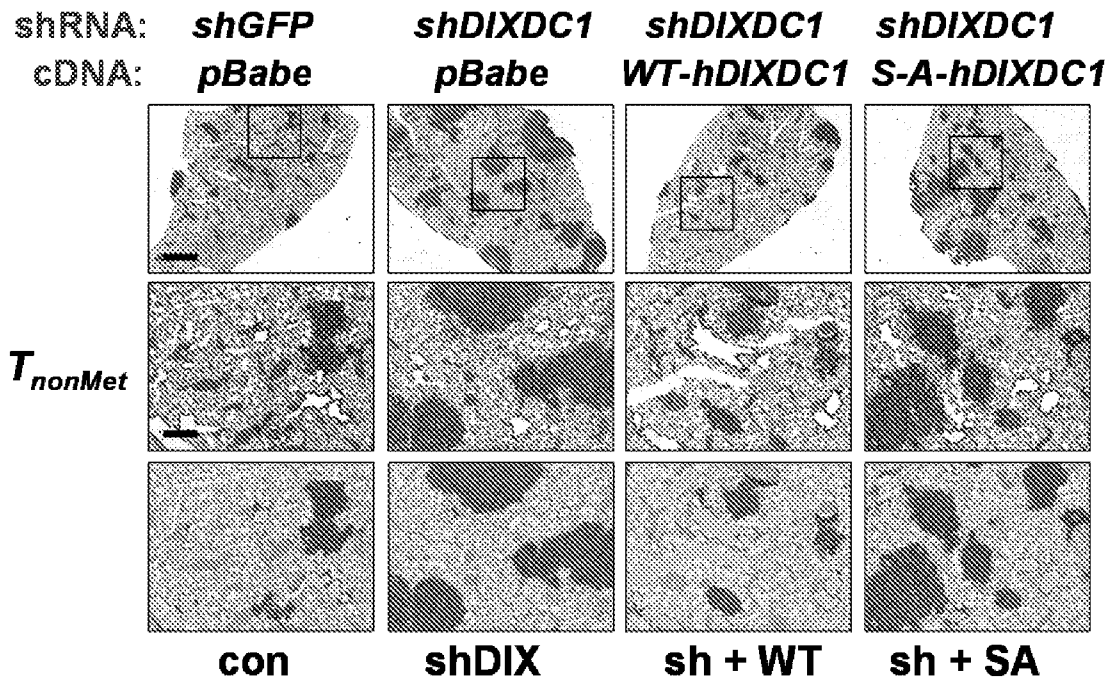


**FIG. 10E**

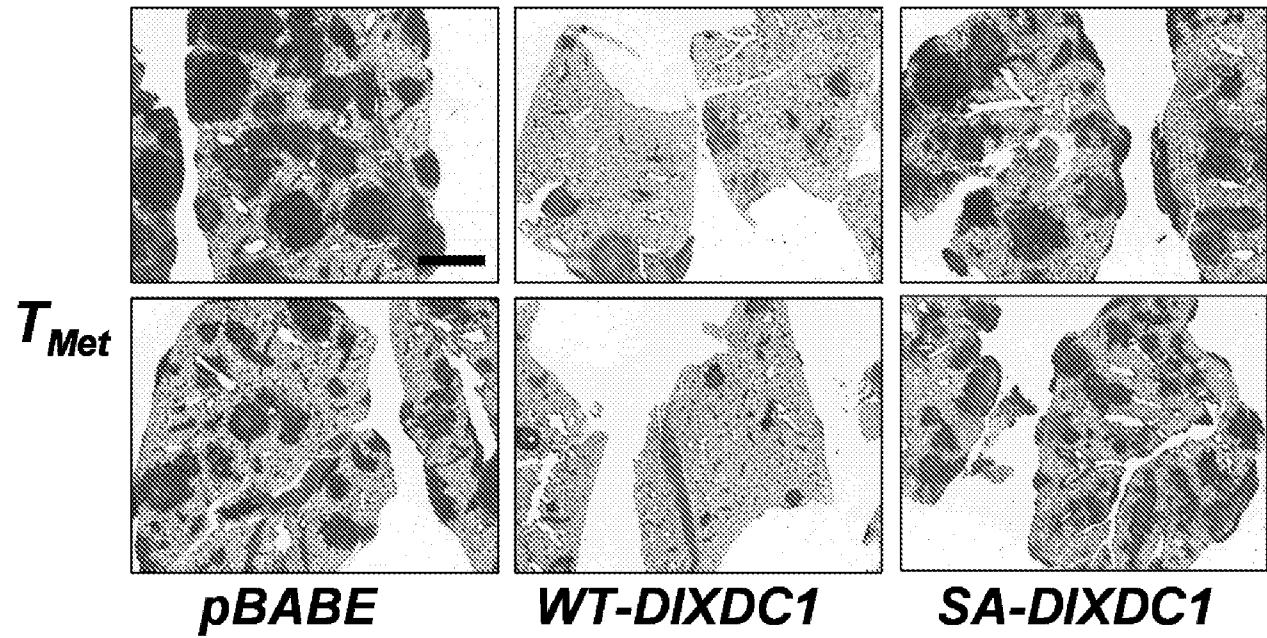


**FIG. 11A**

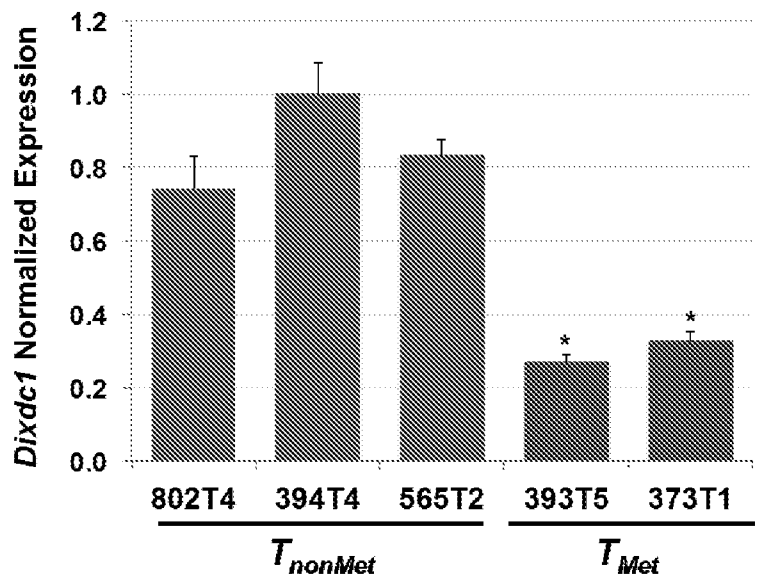


**FIG. 11B**

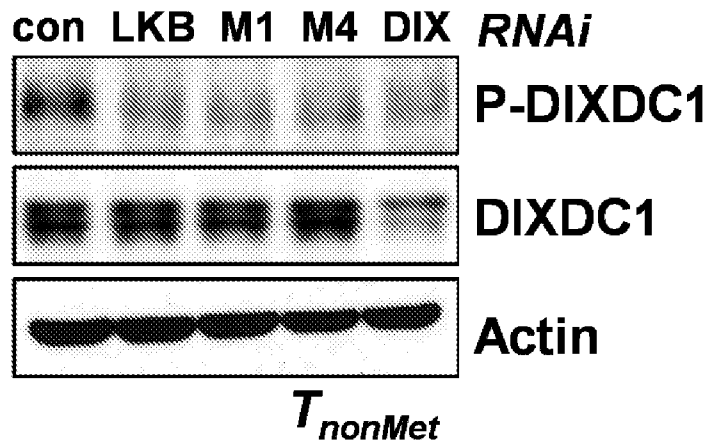


**FIG. 11C**

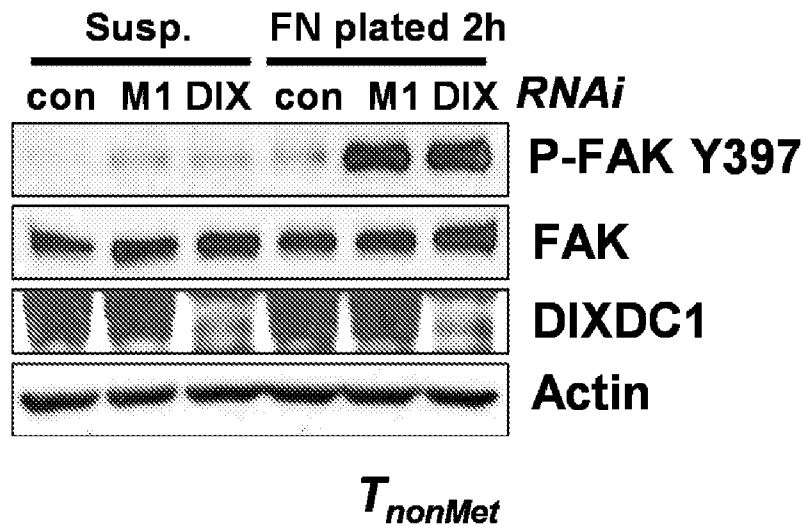


**FIG. 11D**

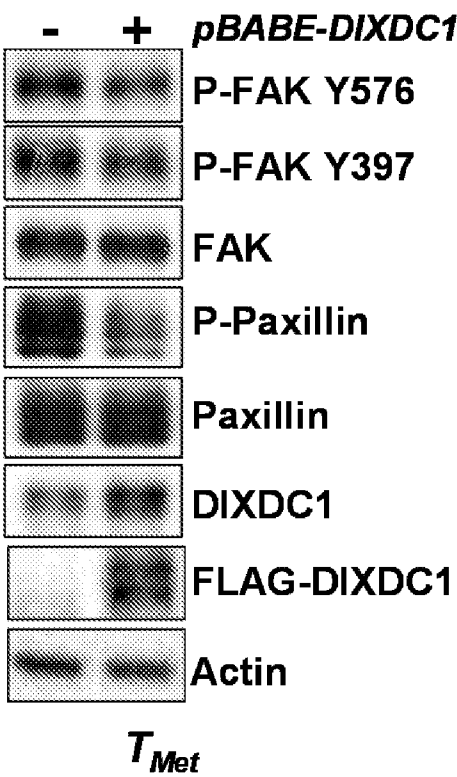


FIG. 11E

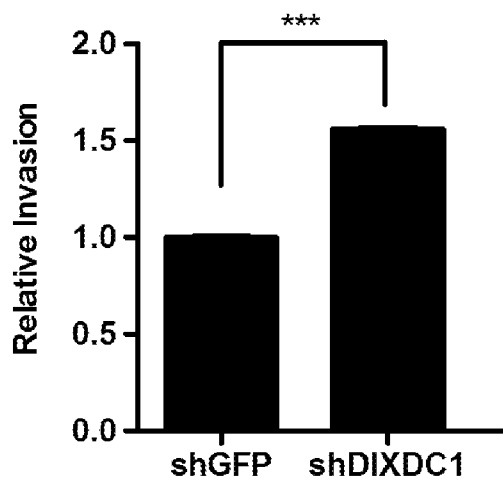


FIG. 11F

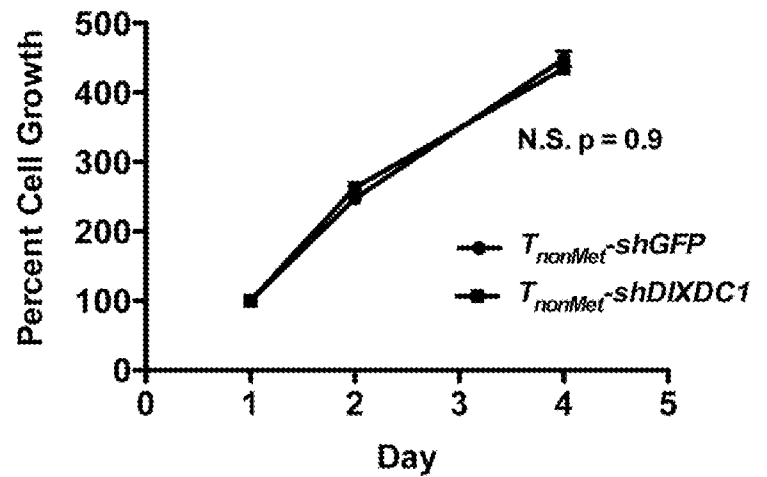


FIG. 11G

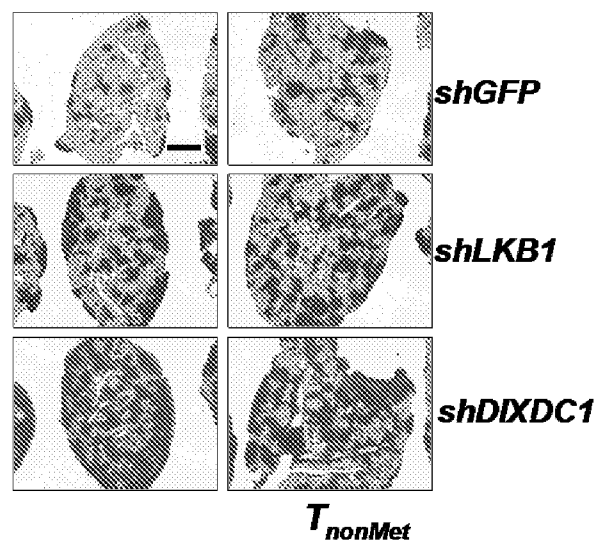


FIG. 11H

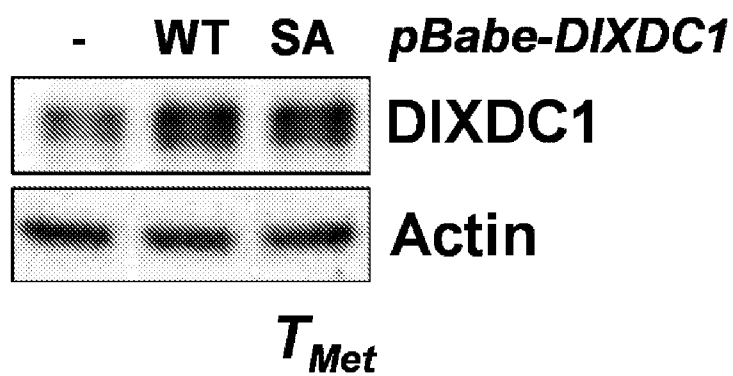


FIG. 12A

| Cancer Subset                         | Deletion Frequency | Minimal common boundary (Chr11) | Q-value  |
|---------------------------------------|--------------------|---------------------------------|----------|
| Cutaneous melanoma (metastatic)       | 55.48%             | 110,933,269 - 114,585,676       | 3.94E-17 |
| Cervical squamous cell carcinoma      | 51.61%             | 106,828,892 - 111,795,202       | 2.95E-11 |
| Lung cancers                          | 29.50%             | 111,718,432 - 111,914,311       | 8.16E-04 |
| Lung adenocarcinoma                   | 23.53%             | 111,714,281 - 111,907,596       | 1.58E-03 |
| Sarcoma                               | 42.67%             | 111,536,882 - 112,092,921       | 1.91E-03 |
| Uterine corpus endometrioid carcinoma | 16.37%             | 109,716,576 - 112,344,381       | 1.18E-02 |

FIG. 12B

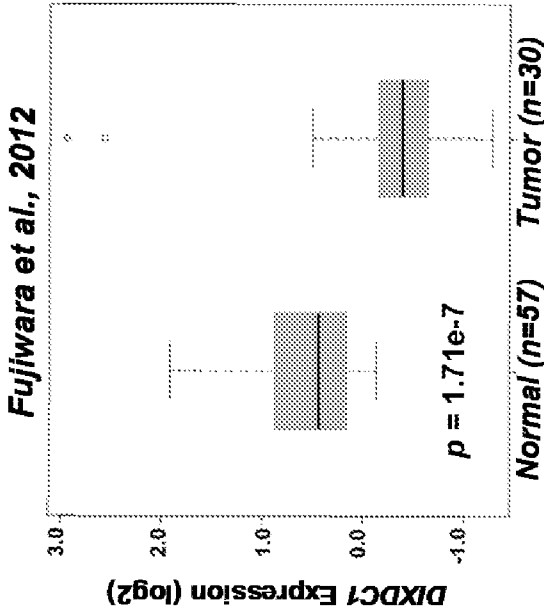
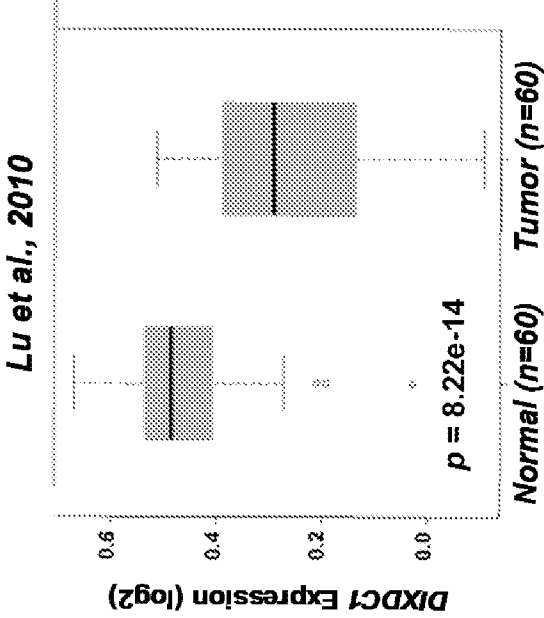
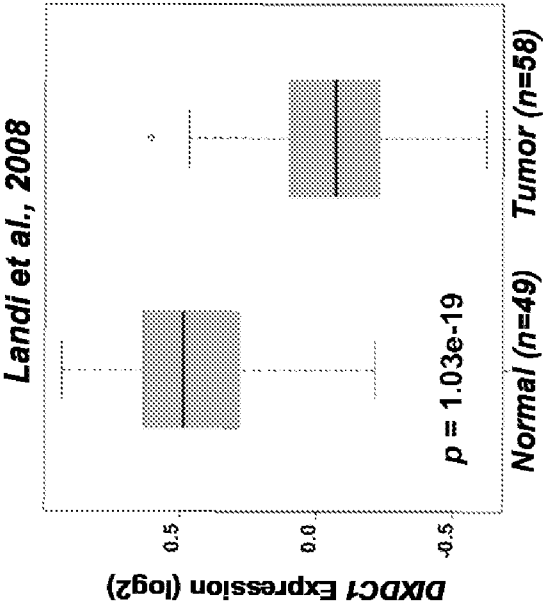


FIG. 12C

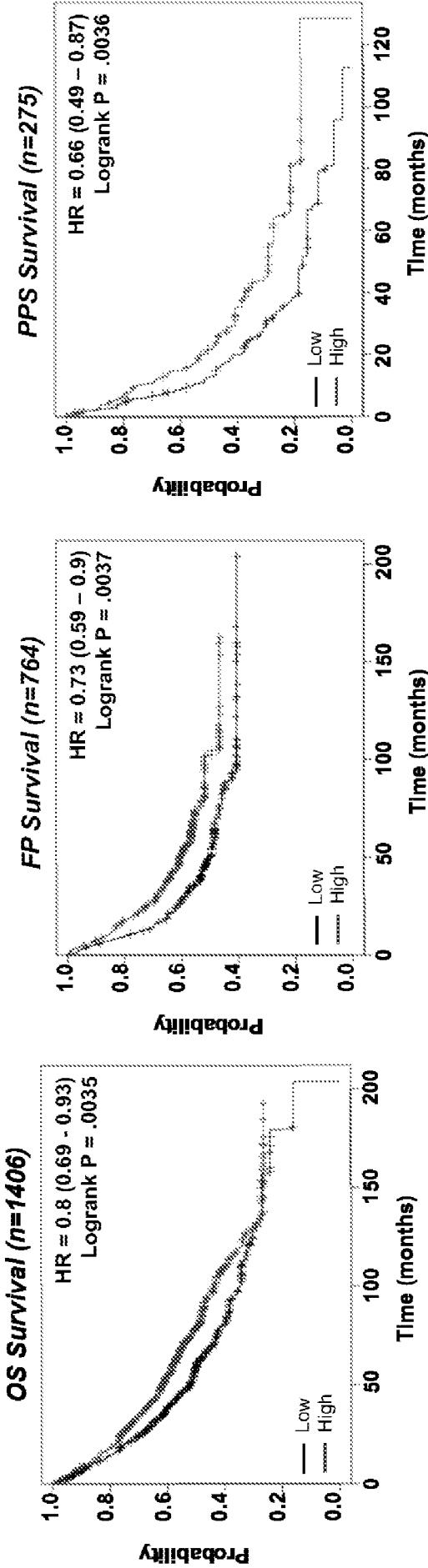


FIG. 12D

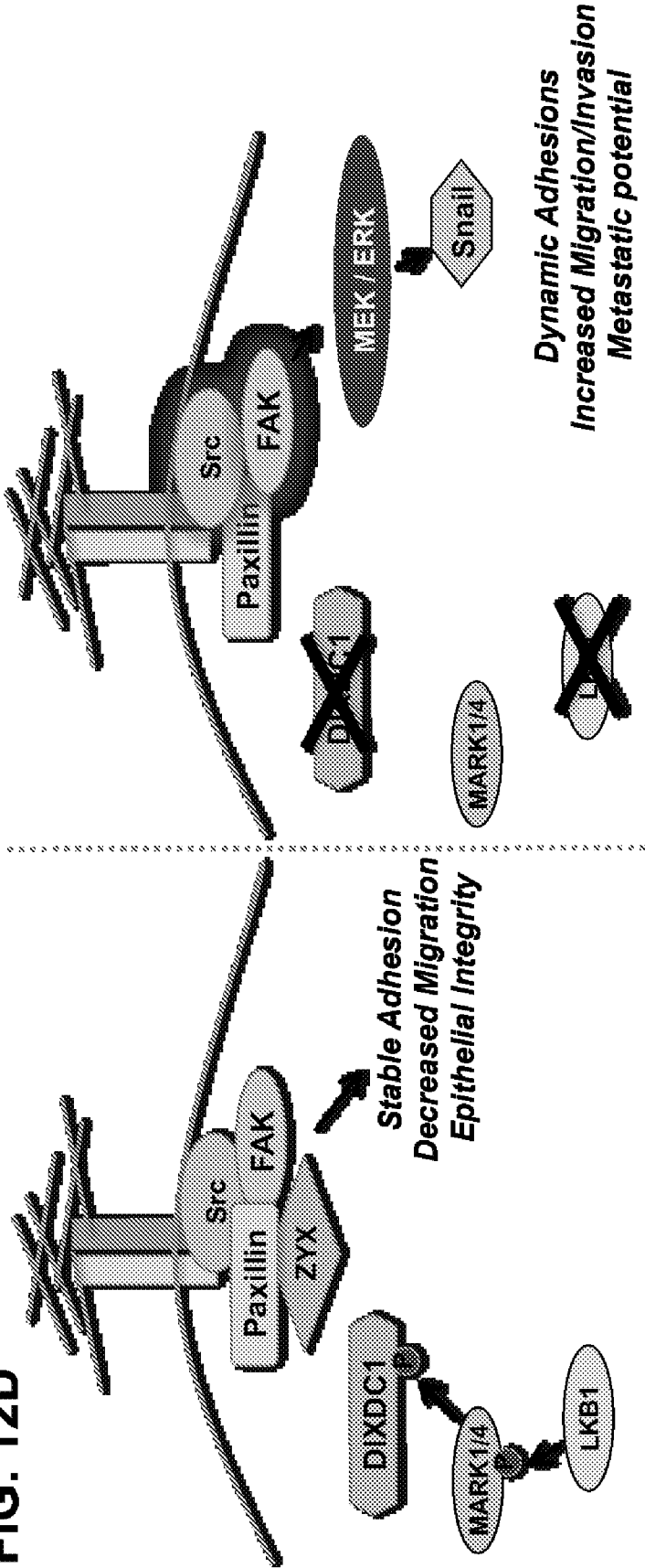
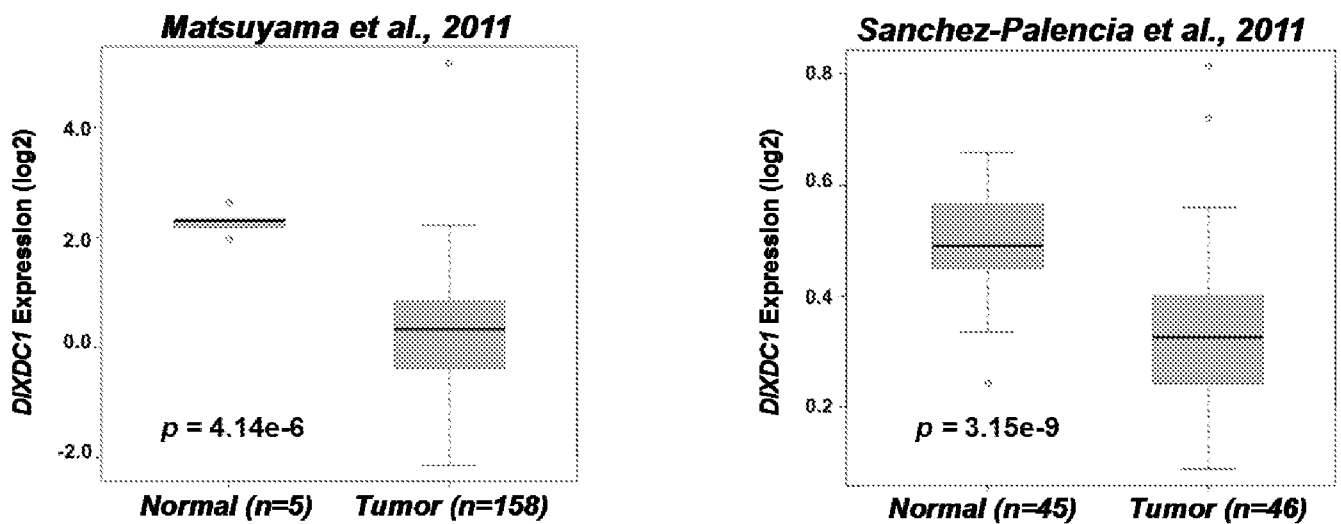
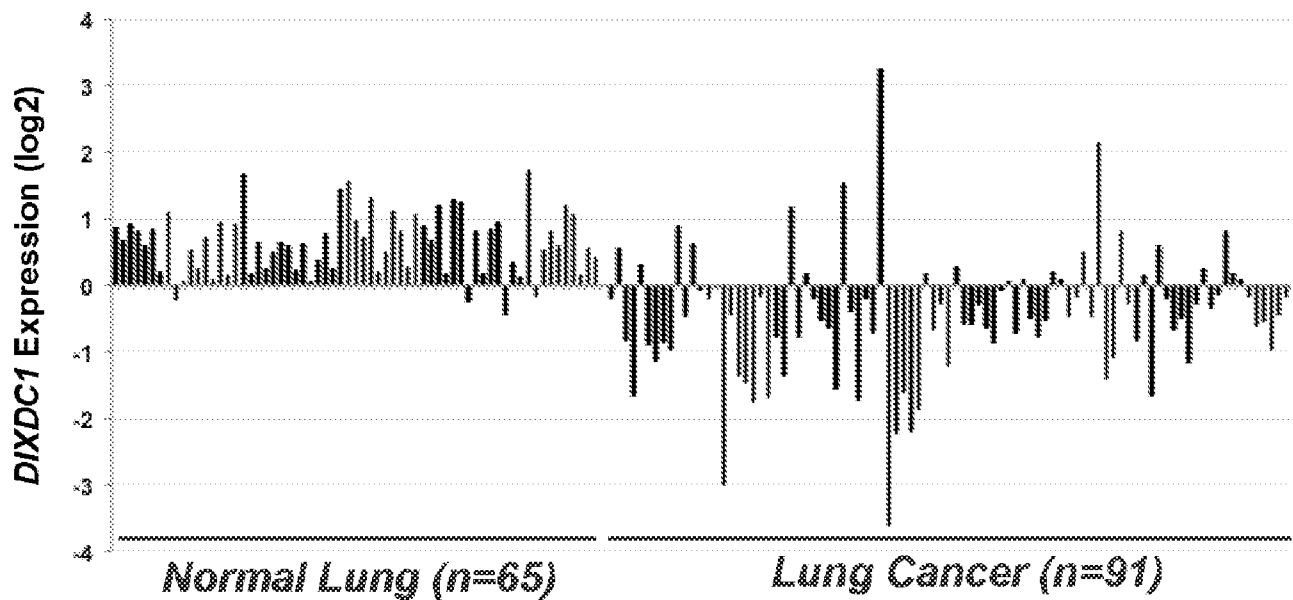
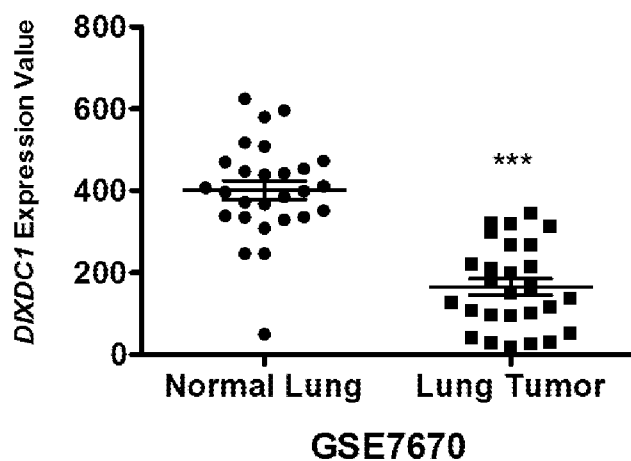
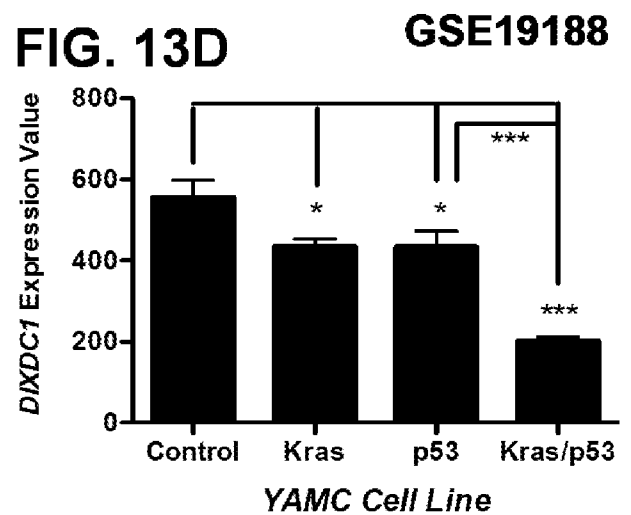
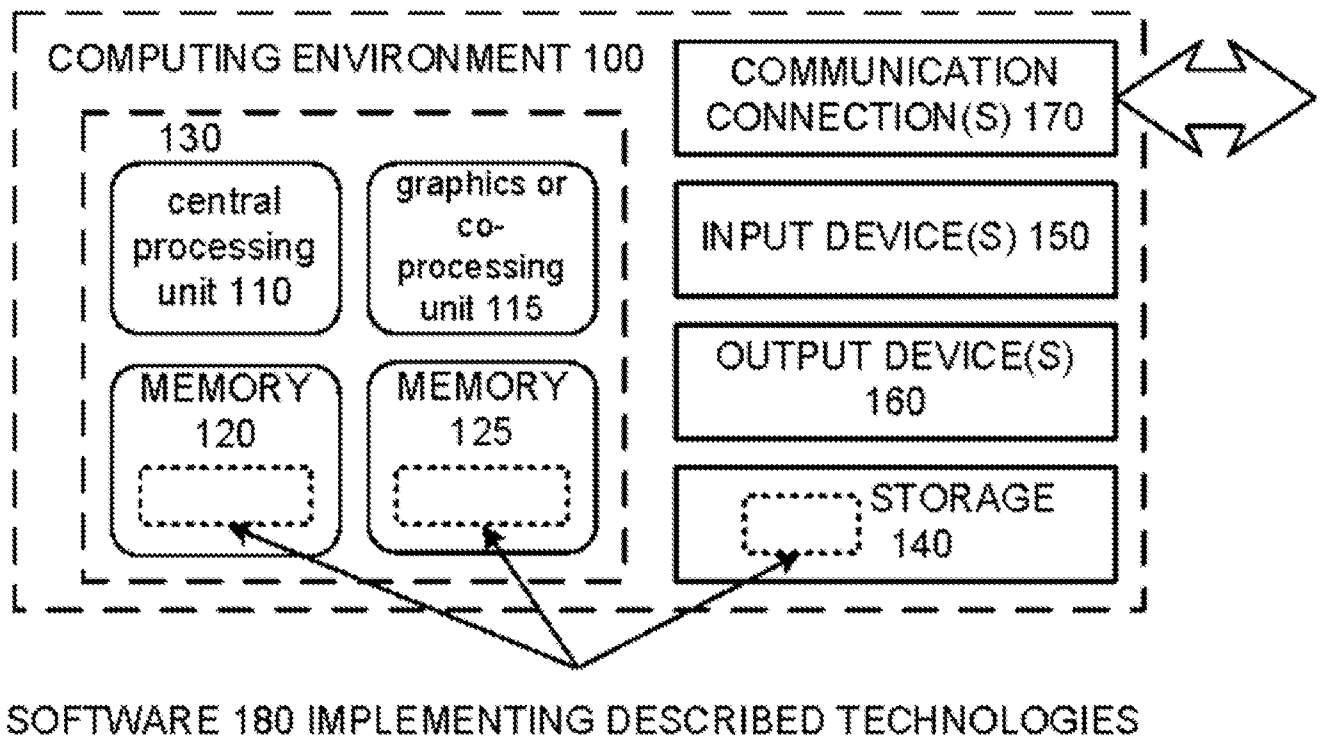


**FIG. 13A****FIG. 13B****FIG. 13C****FIG. 13D**

GSE9199

**FIG. 14**

## INTERNATIONAL SEARCH REPORT

International application No.  
**PCT/US2015/040432**

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                  |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>A. CLASSIFICATION OF SUBJECT MATTER</b><br><b>C12Q 1/68 (2006.01) G01N 33/574 (2006.01) A61P 35/04 (2006.01)</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                  |
| According to International Patent Classification (IPC) or to both national classification and IPC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                  |
| <b>B. FIELDS SEARCHED</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                  |
| Minimum documentation searched (classification system followed by classification symbols)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                  |
| Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                  |
| Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)<br>Databases: EPODOC, WPIAP, TXTE, MEDLINE, CAPLUS, BIOSIS, EMBASE. Search for keywords: Cancer, tumour, carcinoma, sarcoma, leukaemia, metastasis, DIXDC1, coiled coil protein DIX1, DIXIN, KIAA1735 and like terms.<br>Databases: THE LENS, AUSPAT, PUBMED, GOOGLE SCHOLAR. Searched for applicant and inventor names.                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                  |
| <b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                  |
| Category*                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Relevant to claim No.                                                                                                                            |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Documents are listed in the continuation of Box C                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                  |
| <input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C <input checked="" type="checkbox"/> See patent family annex                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                  |
| * Special categories of cited documents:<br>"A" document defining the general state of the art which is not considered to be of particular relevance<br>"E" earlier application or patent but published on or after the international filing date<br>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)<br>"O" document referring to an oral disclosure, use, exhibition or other means<br>"P" document published prior to the international filing date but later than the priority date claimed | "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention<br>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone<br>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art<br>"&" document member of the same patent family |                                                                                                                                                  |
| Date of the actual completion of the international search<br>29 September 2015                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | Date of mailing of the international search report<br>29 September 2015                                                                          |
| <b>Name and mailing address of the ISA/AU</b><br>AUSTRALIAN PATENT OFFICE<br>PO BOX 200, WODEN ACT 2606, AUSTRALIA<br>Email address: pct@ipaustalia.gov.au                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | <b>Authorised officer</b><br>Kelly Hitchens<br>AUSTRALIAN PATENT OFFICE<br>(ISO 9001 Quality Certified Service)<br>Telephone No. +61 2 6222 3604 |

| INTERNATIONAL SEARCH REPORT                           |                                                                                                                                                                                                                                                                 | International application No. |
|-------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------|
| C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT |                                                                                                                                                                                                                                                                 | PCT/US2015/040432             |
| Category*                                             | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                                                              | Relevant to claim No.         |
| X<br>A                                                | US 2010/0184125 A1 (HUANG ET AL.) 22 July 2010<br>0025-8, 0072-3.<br>Abstract, Table 2, Claim 1.                                                                                                                                                                | 27-28<br>1-26                 |
| X<br>A                                                | US 2011/0320392 A1 (BLACK ET AL.) 29 December 2011<br>0029-33, 0113, 0117.<br>Examples 1 and 4, Tables 1 and 3, Claims 61 and 59.                                                                                                                               | 27-28<br>1-26                 |
| X<br>A                                                | US 2010/0247528 A1 (HUNTER ET AL.) 30 September 2010<br>0040, 0062-7, 0082, Claim 16.<br>Tables 2 and 7, Claims 16-30, 0088, 0040, 0062.                                                                                                                        | 27-28<br>1-26                 |
| A                                                     | WO 2011/017106 A1 (THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK) 10 February 2011<br>Abstract, Table V, 069.                                                                                                                                     | 1-26                          |
| A                                                     | WO 2008/082730 A2 (NOVARTIS AG) 10 July 2008<br>Table VIII, Claims 1 and 11, Example 1, Page 298 line 17 - Page 304 line 6.                                                                                                                                     | 1-26                          |
| A                                                     | US 2013/0331294 A1 (FOX CHASE CANCER CENTER) 12 December 2013<br>Claims 31-33.                                                                                                                                                                                  | 1-26                          |
| X                                                     | GOODWIN, J., ET AL., 'Abstract A59: DIXDC1 links the LKB1 tumor suppressor to Snail, noncanonical Wnt signaling, and metastasis', Cancer Research, 2013, vol. 73, no. 3, Supplement A59.<br>Abstract                                                            | 1-29                          |
| A                                                     | XU, Z., ET AL., 'DIXDC1 increases the invasion and migration ability of non-small-cell lung cancer cells via the PI3K-AKT/AP-1 pathway', Molecular Carcinogenesis, 2014, vol. 53, no. 11, pages 917-25.<br>Figures 1-4, Table 2, Page 925 column 1 paragraph 3. | 1-26                          |
| A                                                     | UCHINO, K., ET AL., 'Therapeutic effects of microRNA-582-5p and -3p on the inhibition of bladder cancer progression', Molecular Therapies, 2013, vol. 21, no. 3, pages 610-9.<br>Figures 4, 5 and S4A, Tables S4 and S5.                                        | 1-26                          |
| A                                                     | WO 2014/030602 A1 (NATIONAL CANCER CENTER) 27 February 2014<br>& US2015/0197747A1 used for English translation<br>Examples 6-11, Figures 7-11.                                                                                                                  | 1-26                          |
| A                                                     | WANG, L., ET AL., 'DIXDC1 targets p21 and cyclin D1 via PI3K pathway activation to promote colon cancer cell proliferation', Cancer Science, 2009, vol. 100, no. 10, pages 1801-8.<br>Figures 1-3 and 5.                                                        | 1-26                          |
| A                                                     | WANG, L., ET AL., 'Wnt signaling stabilizes the DIXDC1 protein through decreased ubiquitin-dependent degradation', Cancer Science, 2010, vol. 101, no. 3, pages 700-6.<br>Figures 1-6.                                                                          | 1-26                          |
| Form PCT/ISA/210 (fifth sheet) (July 2009)            |                                                                                                                                                                                                                                                                 |                               |

| INTERNATIONAL SEARCH REPORT                                                                                                                                                                                                                                           |                   | International application No. |                  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|-------------------------------|------------------|
| Information on patent family members                                                                                                                                                                                                                                  |                   | PCT/US2015/040432             |                  |
| This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information. |                   |                               |                  |
| Patent Document/s Cited in Search Report                                                                                                                                                                                                                              |                   | Patent Family Member/s        |                  |
| Publication Number                                                                                                                                                                                                                                                    | Publication Date  | Publication Number            | Publication Date |
| US 2010/0184125 A1                                                                                                                                                                                                                                                    | 22 July 2010      | US 2010184125 A1              | 22 Jul 2010      |
|                                                                                                                                                                                                                                                                       |                   | US 8492328 B2                 | 23 Jul 2013      |
|                                                                                                                                                                                                                                                                       |                   | EP 2145021 A2                 | 20 Jan 2010      |
|                                                                                                                                                                                                                                                                       |                   | EP 2559771 A2                 | 20 Feb 2013      |
|                                                                                                                                                                                                                                                                       |                   | WO 2008144345 A2              | 27 Nov 2008      |
| US 2011/0320392 A1                                                                                                                                                                                                                                                    | 29 December 2011  | US 2011320392 A1              | 29 Dec 2011      |
|                                                                                                                                                                                                                                                                       |                   | GB 2480980 A                  | 07 Dec 2011      |
|                                                                                                                                                                                                                                                                       |                   | WO 2010101916 A1              | 10 Sep 2010      |
| US 2010/0247528 A1                                                                                                                                                                                                                                                    | 30 September 2010 | US 2010247528 A1              | 30 Sep 2010      |
|                                                                                                                                                                                                                                                                       |                   | WO 2009032915 A2              | 12 Mar 2009      |
| WO 2011/017106 A1                                                                                                                                                                                                                                                     | 10 February 2011  | WO 2011017106 A1              | 10 Feb 2011      |
| WO 2008/082730 A2                                                                                                                                                                                                                                                     | 10 July 2008      | WO 2008082730 A2              | 10 Jul 2008      |
|                                                                                                                                                                                                                                                                       |                   | AU 2007340265 A1              | 10 Jul 2008      |
|                                                                                                                                                                                                                                                                       |                   | AU 2007340265 B2              | 26 Jul 2012      |
|                                                                                                                                                                                                                                                                       |                   | BR PI0716944 A2               | 17 Sep 2013      |
|                                                                                                                                                                                                                                                                       |                   | CA 2662508 A1                 | 10 Jul 2008      |
|                                                                                                                                                                                                                                                                       |                   | CN 101541977 A                | 23 Sep 2009      |
|                                                                                                                                                                                                                                                                       |                   | EP 2074226 A2                 | 01 Jul 2009      |
|                                                                                                                                                                                                                                                                       |                   | JP 2010504530 A               | 12 Feb 2010      |
|                                                                                                                                                                                                                                                                       |                   | JP 2013128487 A               | 04 Jul 2013      |
|                                                                                                                                                                                                                                                                       |                   | KR 20090071603 A              | 01 Jul 2009      |
|                                                                                                                                                                                                                                                                       |                   | MX 2009002936 A               | 01 Apr 2009      |
|                                                                                                                                                                                                                                                                       |                   | RU 2009114745 A               | 27 Oct 2010      |
|                                                                                                                                                                                                                                                                       |                   | SG 174826 A1                  | 28 Oct 2011      |
|                                                                                                                                                                                                                                                                       |                   | US 2010004253 A1              | 07 Jan 2010      |
| US 2013/0331294 A1                                                                                                                                                                                                                                                    | 12 December 2013  | US 2013331294 A1              | 12 Dec 2013      |
|                                                                                                                                                                                                                                                                       |                   | US 2010239656 A1              | 23 Sep 2010      |
|                                                                                                                                                                                                                                                                       |                   | WO 2009062199 A1              | 14 May 2009      |

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

Form PCT/ISA/210 (Family Annex)(July 2009)

|                                                                                                                                                                                                                                                                       |                         |                                                           |                         |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|-----------------------------------------------------------|-------------------------|
| <b>INTERNATIONAL SEARCH REPORT</b><br>Information on patent family members                                                                                                                                                                                            |                         | International application No.<br><b>PCT/US2015/040432</b> |                         |
| This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information. |                         |                                                           |                         |
| <b>Patent Document/s Cited in Search Report</b>                                                                                                                                                                                                                       |                         | <b>Patent Family Member/s</b>                             |                         |
| <b>Publication Number</b>                                                                                                                                                                                                                                             | <b>Publication Date</b> | <b>Publication Number</b>                                 | <b>Publication Date</b> |
| WO 2014/030602 A1                                                                                                                                                                                                                                                     | 27 February 2014        | WO 2014030602 A1                                          | 27 Feb 2014             |
|                                                                                                                                                                                                                                                                       |                         | CA 2879410 A1                                             | 27 Feb 2014             |
|                                                                                                                                                                                                                                                                       |                         | EP 2886122 A1                                             | 24 Jun 2015             |
|                                                                                                                                                                                                                                                                       |                         | KR 20150027302 A                                          | 11 Mar 2015             |
|                                                                                                                                                                                                                                                                       |                         | US 2015197747 A1                                          | 16 Jul 2015             |
| <b>End of Annex</b>                                                                                                                                                                                                                                                   |                         |                                                           |                         |
| <div> <p>Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.</p> <p>Form PCT/ISA/210 (Family Annex)(July 2009)</p> </div>                                                                       |                         |                                                           |                         |