Correspondence and requests for materials should be addressed to K.C.B. (Kathlynn. Brown@ UTSouthwestern.edu)

Identification and Characterization of a Suite of Tumor Targeting Peptides for Non-Small Cell Lung Cancer

Michael J. McGuire[1,4], Bethany Powell Gray[1,4], Shunzi Li[1,4], Dorothy Cupka[1,4], Lauren Averett Byers[6], Lei Wu[1,4], Shaghayegh Rezaie[1,4], Ying-Horng Liu[1,4], Naveen Pattisapu[1,4], James Issac[1,4], Tsukasa Oyama[1], Lixia Diao[6], John V. Heymach[6], Xian-Jin Xie[3,4], John D. Minna[1,2,4,5] & Kathlynn C. Brown[1,4]

[1]Departments of Internal Medicine, University of Texas Southwestern Medical Center 5323 Harry Hines Blvd. Dallas, TX 75390-8807, USA, [2]Departments of Pharmacology, University of Texas Southwestern Medical Center 5323 Harry Hines Blvd. Dallas, TX 75390-8807, USA, [3]Departments of Clinical Sciences, University of Texas Southwestern Medical Center 5323 Harry Hines Blvd. Dallas, TX 75390-8807, USA, [4]The Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center 5323 Harry Hines Blvd. Dallas, TX 75390-8807, USA, [5]The Hamon Center for Therapeutic Oncology Research, University of Texas Southwestern Medical Center 5323 Harry Hines Blvd. Dallas, TX 75390-8807, USA, [6]Department of Thoracic and Head and Neck Medical Oncology University of Texas MD Anderson Cancer Center.

Tumor targeting ligands are emerging components in cancer therapies. Widespread use of targeted therapies and molecular imaging is dependent on increasing the number of high affinity, tumor-specific ligands. Towards this goal, we biopanned three phage-displayed peptide libraries on a series of well-defined human non-small cell lung cancer (NSCLC) cell lines, isolating 11 novel peptides. The peptides show distinct binding profiles across 40 NSCLC cell lines and do not bind normal bronchial epithelial cell lines. Binding of specific peptides correlates with onco-genotypes and activation of particular pathways, such as EGFR signaling, suggesting the peptides may serve as surrogate markers. Multimerization of the peptides results in cell binding affinities between 0.0071–40 nM. The peptides home to tumors in vivo and bind to patient tumor samples. This is the first comprehensive biopanning for isolation of high affinity peptidic ligands for a single cancer type and expands the diversity of NSCLC targeting ligands.

I n the United States, 213,000 cases of lung cancer are diagnosed yearly; 60% of these patients die within one year[1]. Non-small cell lung cancer (NSCLC), which accounts for ~80% of lung cancers, is a heterogeneous disease. Even within each histological subtype, there is considerable clinical variability[2]. Personalized therapies tailored to the molecular features of a tumor are anticipated to improve clinical efficacy. To make widespread use of personalized therapies a reality, clinicians must be able to rapidly "classify" clinically relevant subsets, identify treatments effective in that tumor subtype, deliver therapeutics effectively to the tumor, and monitor the tumor phenotype during and after treatment. Tumor targeting ligands are emerging as an important component in customized therapies. Such ligands can be used to help "categorize" tumor cell surface phenotypes and tailor treatment accordingly by attaching appropriate cell-targeting ligands to therapeutic agents. Furthermore, the same ligands can be conjugated to molecular imaging agents to follow expression of the targeted biomarker during treatment.

Most efforts have focused on monoclonal antibodies (MAbs) as targeting agents, resulting in 10 FDA-approved MAb cancer therapies. Despite successes, chemically modifying antibodies for clinical applications is challenging, and production costs are substantial. Additionally, concerns have risen over post-translational modifications, especially glycosylation, which can trigger severe hypersensitivity reactions[3]. Of these 10 MAbs, only 5 different antigens are targeted and 3 are conjugated to therapeutic payload. This limited suite of targets is unlikely to cover the diversity of primary tumors or important subgroups within any tumor type. There remains a need to expand both the types of cell binding ligands used and the diversity of biomarkers they target.

Small, biocompatible, chemically synthesized targeting ligands would expand the potential of molecularly targeted systems. Peptides fulfill these criteria; they are smaller than antibody-based targeting agents, can be synthesized in large quantities, are amenable to regiospecific derivatization and display low toxicity profiles[4]. Peptides can be chemically modified to alter affinity, charge, hydrophobicity, stability, and solubility, allowing for

Table I | NSCLC Binding Peptides

Peptide Name	Cell Line for Selection	Peptide Sequence	Phage Selectivity[1]	Phage Specificity[2]	Half Maximal Binding (nM)[6]	
					Monomer	Tetramer
H1299.1	H1299 (LC)	VSQTMRQTAVPLLWFWTGSL[5]	190	45	1900	42
H1299.2	H1299 (LC)	YAAWPASGAWTGTAPCSAGT[3,5]	83	29	21	0.12
H2009.1	H2009 (AD)	RGDLATLRQLAQEDGVVGVR[5]	300	140	22	0.34
H460.1	H460 (LC)	EAMNSAEQSAAVVQWEKRRI	120	400	>10000	>10000
HCC15.1	HCC15 (SQ)	ATEPRKQYATPRVFWTDAPG	44	34	390	0.0071
HCC15.2	HCC15 (SQ)	FHAVPQSFYTAP[4]	220	73	42	13
A549.1	A549 (AD)	MTVCNASQRQAHAQATAVSL[5]	21	7.8	>10000	>10000
HCC95.1	HCC95 (SQ)	MRGQTGKLPTEHFTDTGVAF	20	69	ND	ND
H1155.1	H1155 (LC)	MTGKAAAPHQEDRHANGLEQ	55	39	ND	ND
H1155.2	H1155 (LC)	MEKLPLSKTGRTVSEGVSPP	61	25	ND	ND
H661.1	H661 (LC)	TNSCRGDWLCDAVPEKARV	26	90	ND	ND

[1]Selectivity is defined as the output phage/input phage normalized to a control phage which display the sequences NQRGTELRSPSVDLNKPGRH (20mer control) or QLGAHNNPRYAP (12mer control).
[2]Specificity is the ratio of the selectivity value for the cancer cell line and a normal bronchial epithelial control cell line, BEAS-2B.
[3]This peptide was isolated from a separately synthesized phage-displayed peptide library[14].
[4]The HCC15.2 is shorter as it was isolated using NEB's 12-mer peptide library.
[5]The isolation of these peptides has been previously reported from our laboratory[14,15]. Half maximal binding was determined using GraphPad Prism using nonlinear curve regression using a normalized response and a variable slope. Binding curves and data analysis are included in the supplementary material.

fine tuning of their *in vivo* biodistribution. A few naturally occurring peptides, such as somatostatin, have been employed for targeting cancer cells[5]. Yet, the number of receptors targeted by naturally occurring ligands is limited. There remains a need to identify peptidic ligands for cell surface cancer biomarkers, even when a naturally occurring ligand is unavailable.

Phage display is useful for identifying peptides that bind cell surface features[6–8]. Affinity selection (biopanning) on intact cells is a versatile approach for identifying ligands, and this approach has yielded targeting peptides[9–15]. Yet, there have not been concerted efforts to generate panels of tumor targeting peptides nor has there been a focus on translating the peptides into high affinity ligands outside of the context of the phage. We report here the isolation and characterization of a panel of peptides that bind all major histologies of NSCLC. By employing multiple phage-displayed peptide libraries and a series of well-characterized NSCLC cell lines, we isolated a suite of high affinity and high specificity NSCLC binding peptides which home to tumors *in vivo*. The peptides show distinct binding profiles across a large sample of NSCLC cell lines and distinguish between immortalized human bronchial epithelial and cancer cell lines. Multimerization of the peptides on a trilysine core results in peptides with cell binding affinities in the picomolar to low nanomolar range. This is the first comprehensive biopanning for the isolation of high affinity peptidic ligands on a panel of cancer cells and demonstrates the feasibility of this approach to expand the diversity of targeting ligands.

Results

Isolation of NSCLC-binding peptides. Instead of purified protein, viable cells were used as bait for peptide selection. This method has several advantages. First, the selection occurs in a more relevant cellular context, preserving the cell surface topography. Second, by changing wash conditions, we can bias the selection towards isolation of internalizing ligands, a critical feature if drug delivery is the desired use of the peptide. Third, the approach requires no prior knowledge of the cellular receptor. To increase the probability of isolating multiple ligands, three *de novo* synthesized peptide libraries were employed. Progression of the selection was followed by sequencing randomly sampled phage clones, and panning was repeated until one or more phage clones emerged as the predominant species. Affinity selection was performed on 8 NSCLC cell lines which include all major subtypes of NSCLC: large cell (LC: H1299, H460, H1155, H6612), adenocarcinoma (AD: H2009, A549) and squamous cell carcinoma (SQ: HCC15, HCC95). This resulted in

isolation of 11 different peptides; each named after the cell line it was selected on followed a number indicating the clone number (Table I). In most cases, the selection converged upon a single sequence between rounds 4–6. The two peptides isolated on H1155 cells were isolated from the same panning experiment; H1155.1 and H1155.2 represented 33% and 25% of the clones respectively; both clones were observed in rounds 4 and 5 but neither dominated. In all selections, consensus sequences were not observed.

The phage clones bind their cognate cell type 20–300-fold better than a control phage; thus cell binding is mediated by the uniquely expressed peptide. Importantly, the phage clones discriminate between their target cancer cell line and an immortalized human bronchial epithelial cell line, BEAS-2B by 8–400-fold (Table I). Negative selections were not performed to remove phage clones that bind both normal and NSCLC cells. It is not clear why the selection isolates peptides with high discriminating power. However, we consistently observe this feature, independent of the cell type[9–15].

Binding profiles of phage clones across a panel of cell lines. Phage clone binding was determined for 40 different NSCLC cell lines (Figure 1A, Supplemental Table S-I). The clones are not specific for histopathological class, indicating they recognize molecular features not resolved by pathology. Additionally, binding is not patient specific; ten of the phage clones bind to another cell line derived from a different patient. Cell lines H2073 and H1993, which represent a primary tumor and corresponding metastasis from the same patient, respectively, have identical peptide binding profiles within experimental error. Although this represents only one paired sample, it suggests that the peptides may be useful for detection and therapy of metastases. Importantly, this panel of 11 phage displayed-peptides binds 85% of the NSCLC cell lines tested.

Binding of the phage clones for cancer cell lines from other organs was determined (Figure 1B, Supplemental Table S-II). Eight phage displayed-peptides have significant affinity for other cancer cell lines. Thus, molecular features recognized by the peptides are shared between tumor types, and these peptides may have utility in cancers beyond NSCLC.

The peptides have affinities comparable to antibodies. Most peptides isolated from phage displayed libraries have affinities unsuitable for clinical uses when synthesized as monomers. On the phage, peptides are displayed on the pIII coat protein in 5 copies at the tip of the filamentous particle. As such, peptides selected may bind the cell surface in a multivalent manner. Tetramerization of peptides using a
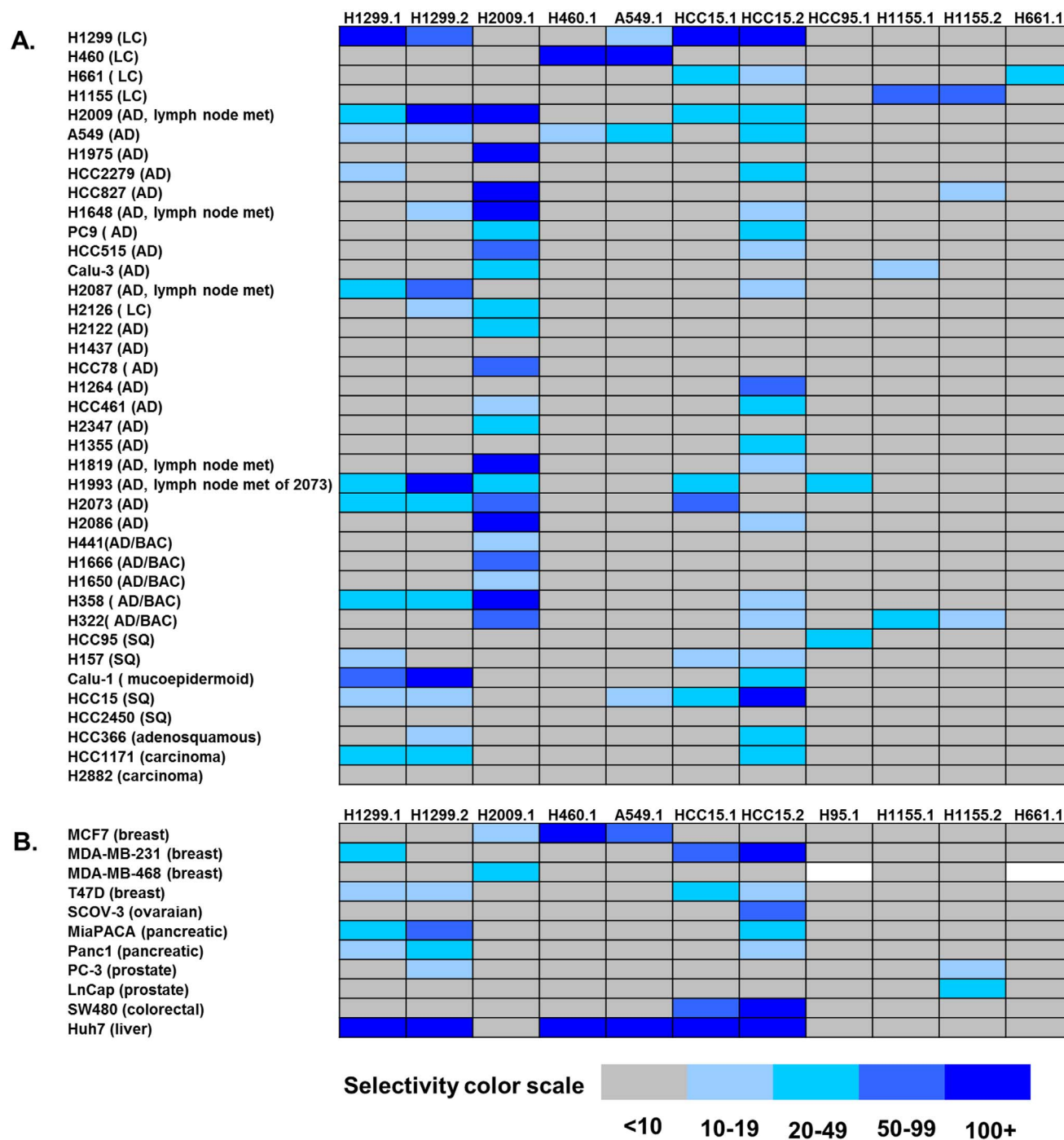
Figure 1 | Binding profile of targeting phage clones for a panel of cancer cell lines. The selectivity for each phage clone on the cell lines is color coded as a heat map. Selectivity is defined as the output phage/input phage normalized to a control phage. White indicates that the selectivity value was not determined. Panel A: NSCLC cell lines. Panel B: Cancer cell lines derived from other organ sites.

trilysine core maintains the valency and orientation found on the phage particles (Figure 2). The linker between the trilysine core and the peptide is a thioether bond, not a native amide bond, allowing for convergent synthesis. Modifications are placed on the trilysine core; as such the peptide sequence is not disturbed, peptide affinity is not altered, and single modifications can be made. Each peptide contains a polyethylene glycol linker (PEG_{11}) to increase solubility of and prevent peptide aggregation.

The most promising peptides as defined by the broadest binding of the phage clone across the panel of cell lines were chosen for further study; H1299.1, H1299.2, H2009.1, H460.1, HCC15.1, and HCC15.2, were synthesized as monomers and tetramers. These peptides in

combination bind to 75% of the NSCLC lines and 90% of the other tumor lines tested.

Half-maximal binding affinities of synthetic peptides were determined (Table I). Except for the HCC15.2 peptide, tetramerization increases the apparent affinity of the peptides for their target cells by >45-fold compared to the monomer. The non-additive increase in affinity likely results from multivalent binding on the cell surface, i.e., the avidity effect. Tetrameric HCC15.2 peptide may not exhibit multivalent binding on the cell surface due to the arrangement of its receptor. The half-maximal binding affinities range from low nanomolar to low picomolar, rivaling affinities of MAbs. Surprisingly, synthetic H460.1 peptide was unable to block phage binding in either
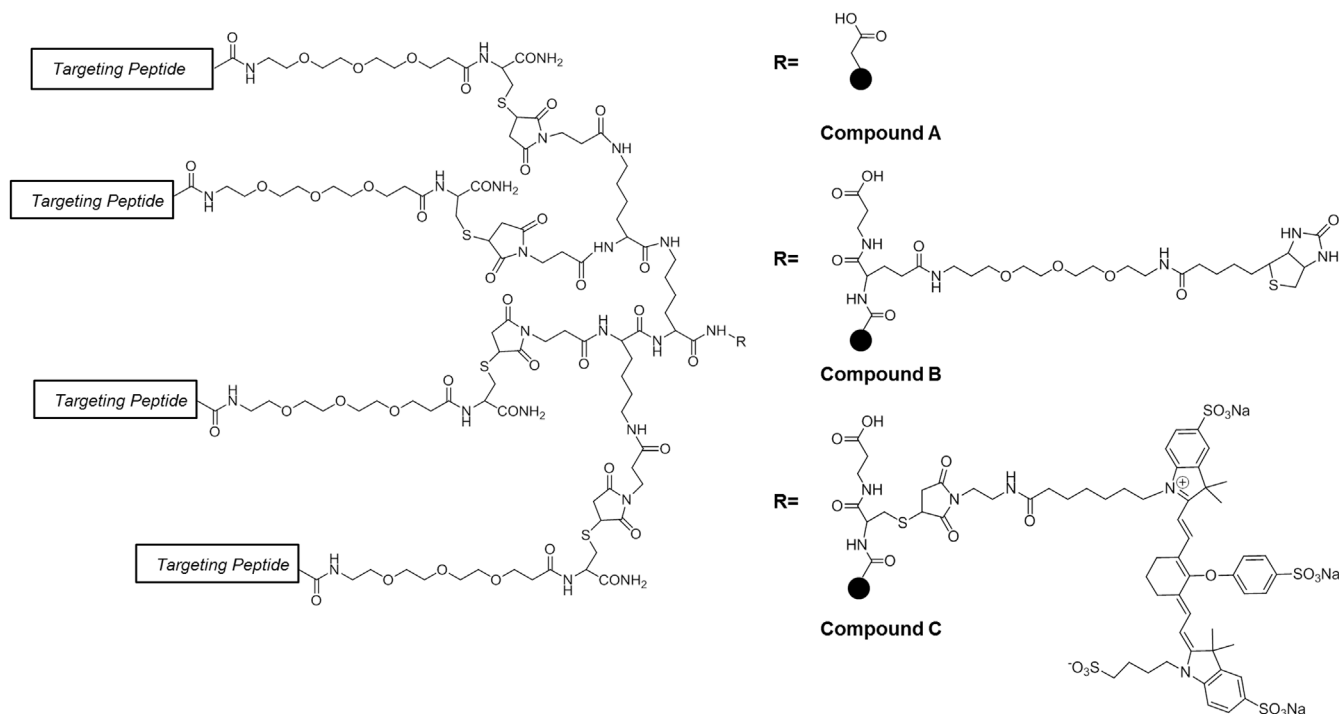
Figure 2 | Structure of tetrameric peptides. Peptides were synthesized on different trilysine cores appropriate for the application: Compound A: β-alanine core, employed for peptide affinity assays; Compound B: biotinylated core, employed for microscopy and flow cytometry; Compound C: 800 CW NIR dye labeled core, employed for animal imaging.

valency, even at 10 μM. H460.1 peptide binding may be dependent on its presentation on pIII coat protein of the phage particle. Alternatively, free H460.1 peptide may bind but is unable to saturate binding.

The peptides mediate cellular uptake. To determine cellular localization of the peptides upon binding, biotinylated tetrameric peptides were conjugated to streptavidin coated fluorescent quantum dots (SAQdot) and incubated with H2009, H1299, or H460 cells. The C-terminal 5 amino acids were removed from H1299.2 peptide for these experiments. This region of the peptide does not contribute to binding as determined by alanine scanning and the 15-mer has the same affinity as the parental peptide (Supplemental Figure S1). Deleting this region removes the cysteine residue which must remain protected to avoid intra- and inter-branch aggregation. Similarly, a fully functional 10-mer form of H2009.1 peptide was employed[16].

Microscopy reveals that the peptides are internalized and show homogenous staining within the cell populations (Figure 3A). In all cases, cell specificity observed with each phage clone is maintained; no uptake is observed when the H2009.1 peptide-conjugate is incubated with the non-binding H1299 cell line but is clearly observed in H2009 cells. H1299.2, HCC15.1, and HCC15.2 peptides mediate cellular uptake in H1299 cells and H2009 cells. No binding was observed on H460 cells (data not shown). Prominent perinuclear punctate staining is observed. Unfortunately, poor solubility of H1299.1 hampered further studies with this peptide. Yet the H1299.1 and H1299.2 phage clones have a similar binding profile on the panel of cell lines; removal of the H1299.1 peptide does not limit the number of cell types targeted by the suite of peptides.

Although the H460.1 peptide is unable to compete with its cognate phage clone, specific binding to MCF7 cells is observed with H460.1 peptide-conjugated to SAQdots (Figure 3B). However, the binding pattern is distinctly different from the other peptides. The H460.1 peptide-Qdot is observed in patches localized at the membrane, and most cells have one region of peptide-Qdot staining which is

excluded from cell-cell contacts. A similar staining pattern is observed when the H460.1 phage clone binds to MCF7 cells (Supplemental Figure S2). Phage is detected with anti-phage antibodies on non-permeabilized cells, indicating the peptide remains bound to the outer cell membrane. This is surprising because both the selection process and microscopy experiments include low pH washes, designed to remove surface bound phage or peptide. Staining is also observed in the extracellular space. This is not observed when H460.1 peptide or phage clone is incubated with non-binding cell lines suggesting that the MCF7 cells secrete or shed the biological target. Binding to an extracellular target may explain the inability of the free peptide to block H460.1 phage binding.

Binding profiles of the free tetrameric peptides across a panel of cell lines. We determined binding of the isolated phage clones on a panel of cell lines in order to select peptides for further study. However, selectivity is based on a ratio and does not provide absolute values of phage binding. Additionally, there is no guarantee that the peptides have the same specificity as the parental phage clones. For these reasons, peptide binding was determined by flow cytometry using labeled tetrameric peptides (Figure 4A–B). This assay allows us to quantify relative uptake in different cell lines. In general, the peptide binding results mirror those observed with the phage clones. We observe concordance in binding of the H2009.1 tetrameric peptide with the H2009.1 phage clones. The H1299.2, HCC15.1, and HCC15.2 peptides demonstrate 80% concordance with phage clone binding but in several cases, binding is positive for the free peptide but not for the phage clone. However, binding of the phage clone but not the corresponding peptide is not observed. The difference in binding between the phage clone and free peptide likely stems from two factors. First, phage clone binding may be underrepresented due to loss of phage viability. Second, the molar concentration of the free peptide is much greater than that of the phage clone. For example, 1×10^8 phage/mL corresponds to 0.83 pM of peptide, assuming 5 peptides per phage clone.
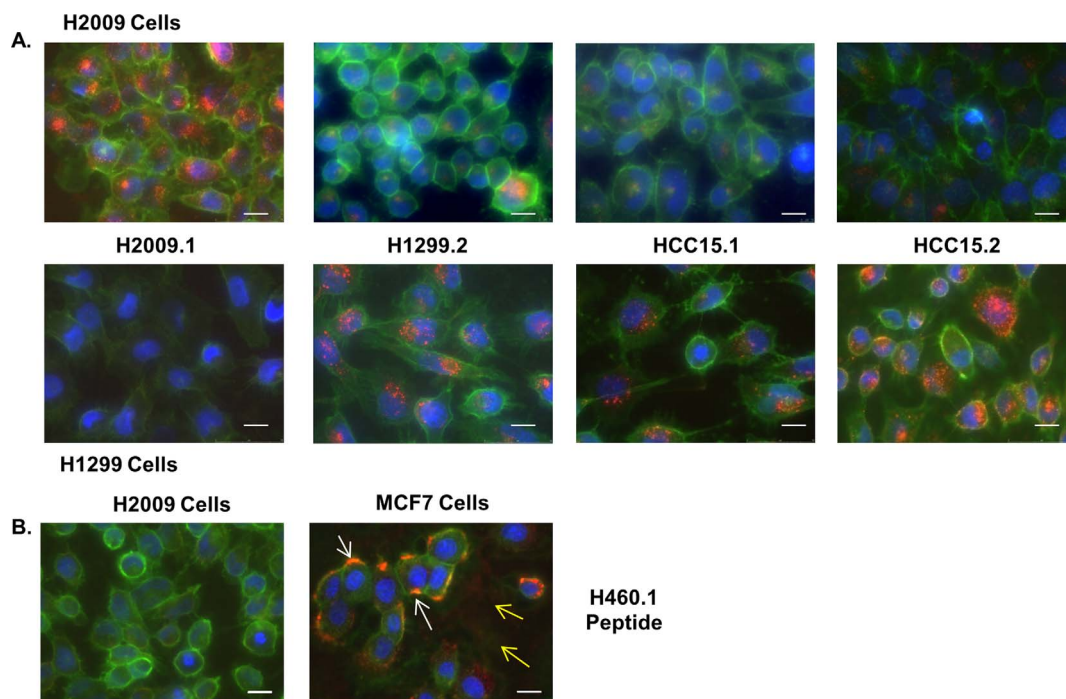
4

Figure 3 | The free tetrameric peptides bind to and mediate internalization. (A). Cells were incubated with 10 nM of the SAQdot605 conjugated to the tetrameric peptide via a streptavidin-biotin interaction. Images were obtained on a Leica DM5500 Fluorescence Microscope at 400× total magnification. The scale bar represents 20 microns. The top panels are H2009 cells incubated with the indicated peptides while the bottom panels are H1299 cells. The red color represents the peptide-Qdot conjugates; blue is DAPI staining of the nuclei; the green is the cell surface stained with wheat germ agglutinin. The peptides do not bind the non-targeted cell line. In all cases, the micrographs for each peptide/cell pair are obtained at the same gain. (B). H460.1-SAQdot605 was incubated with H2009 cells (left panel) and MCF7 cells (right panel). MCF7 cells have a high selectivity for the peptide and are more suitable for microscopy experiments than H460 cells due to their larger size and higher cytoplasm/nucleus ratio. The target cells have localized patches of peptide binding that are excluded from cell-cell contacts as indicated by white arrows. Hazy staining is seen in the extracellular space for the MCF7 cells (indicated by yellow arrows) but not the H2009 cells signifying binding to shed or secreted proteins.

The free peptides maintain their specificity for cancer cells compared to normal control cells. Binding was determined on a non-transformed but immortalized human bronchial epithelial cells (HBEC3) and the BEAS-2B cell line[17]. There was clear discrimination between cancer cell lines and the normal control cell lines (Figure 4A–B). Additionally, we synthesized scrambled control peptides for each of the 4 peptides. No binding was observed for the scrambled peptides on any of the cell lines. Most cell lines fall into one of three main clusters: cluster one which binds all four peptides, cluster two which binds H1299.2, HCC15.1, and HCC15.2, and a third cluster which do not bind any of the peptides. Interestingly, the H1299.2, HCC15.1, and HCC 15.2 bind as a triplet. Except for 4 cell lines (cluster 4), binding of the H2009.1 peptide is accompanied by binding of the triplet. However, binding of the triplet does not necessarily dictate binding of the H2009.1 peptide.

We were surprised to see the same binding patterns for the H1299.2, HCC15.1, and HCC15.2 peptides. There is a linear correlation between cell binding values of each of the three peptides against each other (Supplemental Figure S3). By contrast, there is no correlation of H2009.1 binding compared to H1299.2. These data suggests that the H1299.2, HCC15.1, and HCC15.2 peptides are binding the same receptor or receptors that are co-regulated. The ability of each peptide to block uptake of another was determined (Supplemental Table S-IV). A 10-fold excess of HCC15.1 peptide is able to block uptake of dye-labeled HCC15.1, HCC 15.2, and H1299.2. However, reciprocity does not exist; HCC15.2 blocks itself as expected but not HCC15.1 and H1299.2. Similarly, a 10-fold excess H1299.2 peptide is unable to completely inhibit HCC15.1 uptake but blocks both itself and HCC15.2. It is possible that these data reflect the relative Kd values of each peptides for a single site on

the same receptor, however until the receptors are identified, it is ambiguous whether the peptides bind the same target. It is important to remember that we are measuring a relative amount of peptide uptake, which is governed by affinity of the peptide for its receptor, the rate of internalization, and receptor recycling. The factors that influence the amount of peptide uptake are complicated.

Correlation of Peptide Binding to the Onco-Genotype and Onco-Phenotype of the Cell Lines. Although the receptors are unknown for the majority of the peptides, their binding could serve as surrogate markers for onco-genotypes or activation of particular pathway. As an exploratory study, we determined if peptide binding correlated with mutational status of genes known to play a role in NSCLC and/or protein expression and activation (Supplemental Table S-V). No significant correlation was observed with binding and mutations in EGFR, BRAF, KRAS, NRAS, LKB1, PTEN, PI3K, p53 except for a negative correlation between H1299.1, HCC15.2, or HCC15.2 binding and presence of mutated CDKN2A and PIK3CA (P = 0.0104 and 0.0033 respectively, Fisher exact test, Supplemental Table S-VI).

Reverse phase protein arrays (RPPA) were then used to measure differences in protein expression between 200 total and phosphoproteins between cell lines with distinct peptide binding profiles[18,19]. Ten proteins have significantly different level in cells that bind H2009.1 compared to cells that do not bind H2009.1 (t test, P < 0.01, Supplemental Table S-VII). Importantly, this protein profile includes higher expression of proteins associated with EGFR activation. For example, levels EGFR_pY1173, Her2_pY1248, Met_pTyr1234/1235, STAT5_pY694, and STAT6_pY641 are increased >2-fold in the H2009.1 binding cells (Figure 5A). E-cadherin and beta-catenin
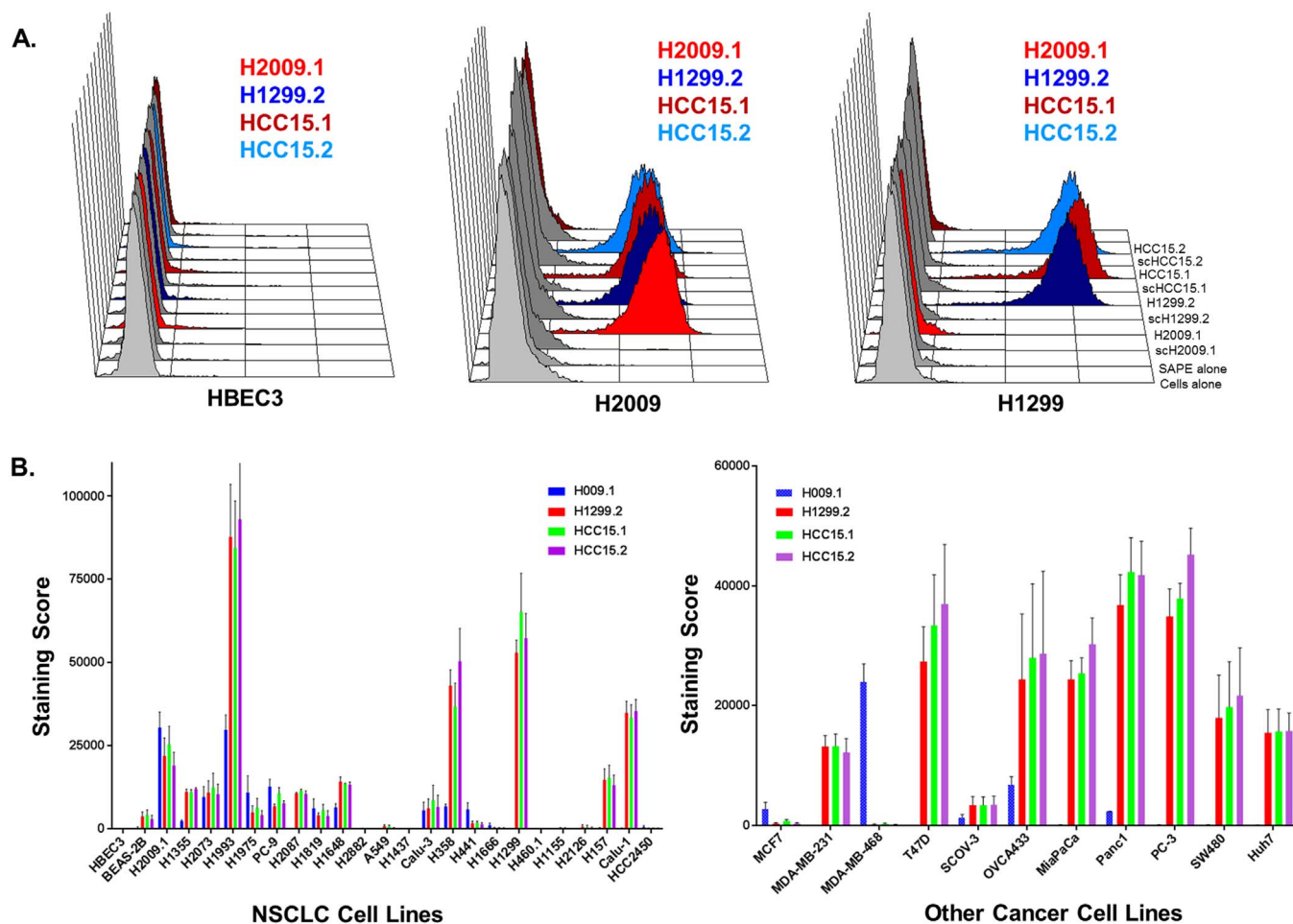
5

Figure 4 | The tetrameric peptides display discrete binding profiles across a panel NCSCL cell lines as well as other cancer types. (A). Peptide binding was determined by flow cytometry. Representative histograms are shown for HBEC3, H2009, and H1299 cells. Nonspecific uptake and fluorescence controls include cells incubated without SAPE, unconjugated SAPE, and a scrambled sequence control peptide conjugated to SAPE. In each case a corresponding scrambled peptide is shown on the line before the targeting peptide. (B). Binding profiles of the free peptides for different cell lines. Cells were considered to be positively stained if their fluorescence intensity was above the first decade. To compare staining across the panel of cells, a staining score was defined as the percentage of positively stained gated live, single cells multiplied by the mean fluorescence intensity of that population. The average and SEM of three experiments are shown. Staining score averages below 5000 are considered negative for significant uptake of the individual peptide conjugate.

expression is significantly higher in cell lines that bind H2009.1 peptide ($P = 0.00105$ and 0.00546) while N-Cadherin expression is reduced ($P = 0.00185$), suggestive of an epithelial phenotype. The cellular receptor for the H2009.1 has been previously identified as $\alpha v \beta 6$. Consistent with this observation, $\alpha v \beta 6$ gene expression has recently been suggested to be an epithelial marker in NSCLC[18]. Similarly strong correlations with protein expression were not observed with H1299.2, HCC15.1, and HCC15.2 binding although cells that bind these peptides have 2-fold greater expression of p16 ($P = 0.00949$, Supplemental Figure S4, Supplemental Table S-VII). XIAP and caspase-7 expression is decreased in expression relative to the cells that do not bind the triplet ($P = 0.0160$ and 0.0231, respectively). The lack of other proteins identified as markers for H1299.2, HCC15.1, and HCC15.2 binding could be either because they are involved in pathways not well represented on the RPPA or because they bind to a range of NSCLC cell lines with heterogeneous profiles.

The NSCLC cell lines have been previously categorized as having epithelial or mesenchymal phenotypes[18]. In cluster 1, 6 of the 7 cell lines have epithelial phenotypes while all five NSCLC cell lines in cluster 2 are mesenchymal. Both cluster 1 and cluster 2 bind the H1299.2, HCC15.1, and HCC15.2 triplet, differing only by binding of the H2009.1 peptide. Interestingly, group 3 is equally split between

the two phenotypes; a lack of $\alpha v \beta 6$-expression does not dictate a mesenchymal phenotype. Consistent with these observations, E-cadherin expression is increased in cluster 1 compared to cluster 2 while cell lines in cluster 3 have a wide spread of E-cadherin levels (Figure 5B, $P = 0.0035$, ANOVA). The inverse is observed with N-cadherin although with less statistical significance. This data suggest that combinatorial patterns of peptide binding can provide cellular information.

The peptides bind to human NSCLC samples. Our data suggest that the peptides bind to NSCLC-specific biomarkers. However, these experiments were performed with cell lines. To determine if peptides bind to human NSCLC tumors, we probed a lung cancer tumor microarray with the H1299.2, HCC15.1 and HCC 15.2 peptides. We previously determined $\alpha v \beta 6$ to be expressed in human NSCLC. As such, we did not explore binding of the peptide on this TMA. The distribution by histopathological subclasses included 37 squamous, 5 large cell, 8 adenocarcinoma, 7 BAC, and 2 carcinosarcomas. Of 59 samples, 24% were scored positive for binding of HCC15.2, 19% of H1299.2, and 14% of the HCC15.1 peptide. Peptide binding is observed for adenocarcinoma (4 samples), large cell (2 samples) and squamous carcinoma (8 samples). A representative tumor which was scored positive and another tumor which scored
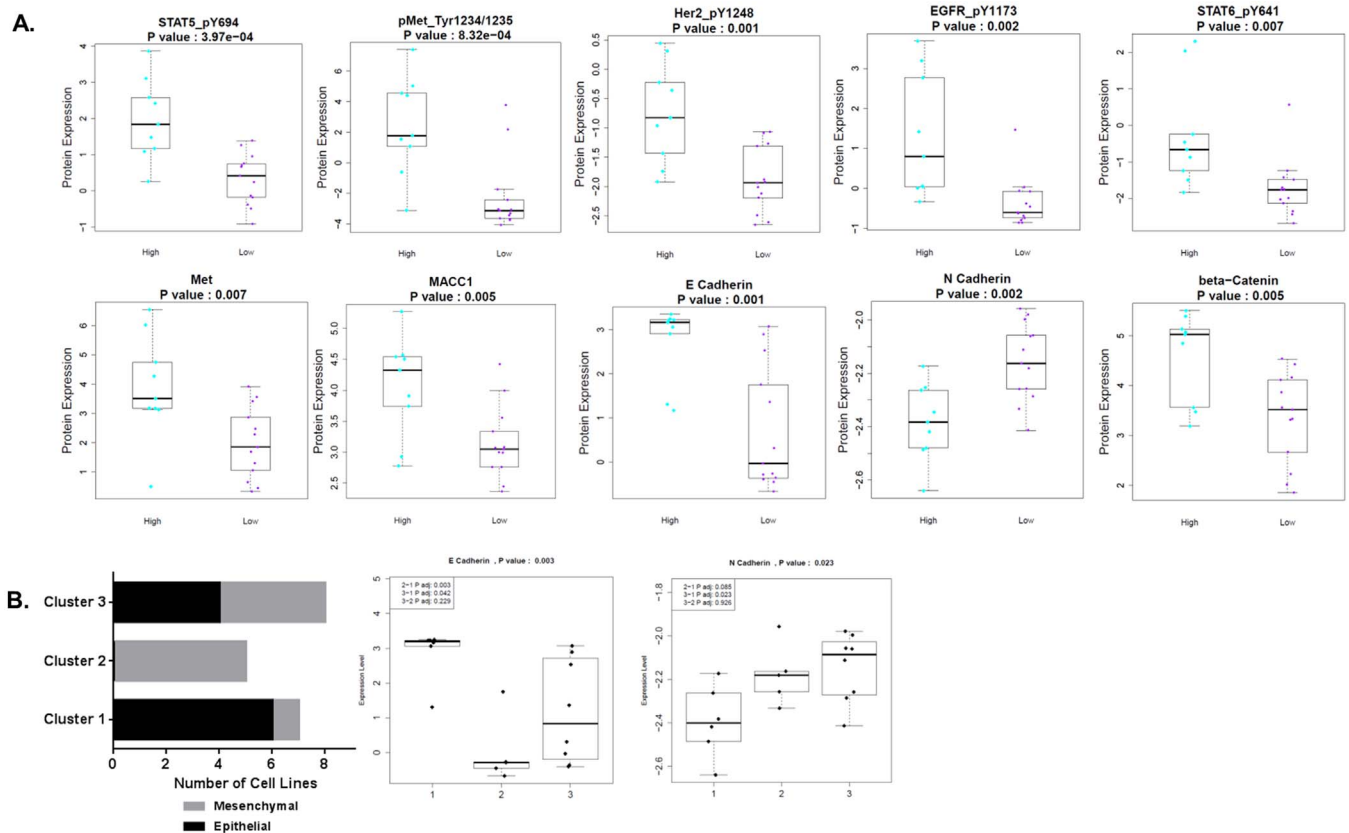
Figure 5 | Peptide binding correlates with cellular phenotypes. Cells were dichotomized into high and low binders using a cell binding score as 5000 as a benchmark and compared to expression levels of 200 proteins analyzed by RPPA. (A). Ten proteins have differential expression of >2-fold between cell lines which are H2009.1 binders and non-binders (t-test, P value <0.01). (B). Epithelial or mesenchymal phenotype of the cell lines correlates with peptide binding (Chi-square analysis, P = 0.0138). Cluster 1 binds H2009.1, H1299.2, HCC15.1, and HCC15.2 peptides. Cluster 2 binds H1299.2, HCC15.1, and HCC15.2 peptides. Cluster 3 binds none of the 4 peptides. Distribution of E-cadherin and N-cadherin expression level amongst each cluster is shown (Anova analysis).

as negative are shown in Figure 6. Again, most positive samples bound the triplet set of peptides. Corresponding normal lung tissue was negative. These data indicate that the peptides bind tumor-associated markers relevant to human cancer.

The peptides home to NSCLC xenograft tumors in vivo. NIR dye-labeled H1299.2, H2009.1, H460.1, HCC15.1, and HCC15.2 tetrameric peptides were injected intravenously into mice bearing H2009 and H460 tumors or H1299 and H460 tumors. Signal from the targeted tumor compared to other tissues was determined for each animal (Figure 7A–B, Supplemental Table S-VIII). Labeled-H2009.1 accumulates in the H2009 tumor with a 120-fold preference over the H460 tumor. A 180-fold greater signal is observed in the H2009 tumor compared to the lung; thus, the peptide discriminates between the tumor and normal lung tissue in vivo. Primary clearance occurs via the kidneys and to a much lesser extent in the liver. Negligible fluorescence was observed in the heart or spleen. While we do see animal to animal variation, in all cases the tumor specificity is maintained (Figure 7B). For example, the ratio of tumor to lung varies from 8–180 (average 75 ± 34), but in no case do we see significant accumulation of the H2009.1 peptide in the lung. By comparison, the average tumor to lung ratio for the non-targeted H460 tumor is 2.5 ± 1.6.

Similar results are observed for the H460.1 and HCC15.1 conjugates, although the ratio of peptide uptake in the target tumor vs. other tissues is lower than that observed with the H2009.1 peptide. H460.1-dye conjugate is observed in the H460 tumor preferentially (8-fold H460 tumor/H1299 tumor and 15-fold H460 tumor/lung for animal

shown, average for all animals 27 and 14 ± 6.5, respectively). HCC15.1 shows a 5-fold preference for the H1299 tumor compared to the lungs and maintains its specificity for H1299 tumors compared to the control H460 tumor (average for all animals 4.8 ± 1.8 and 3.3 ± 0.97). Relative liver accumulation of HCC15.1 compared to the target tumor is greater than observed with other peptides. It remains to be determined if this is from clearance or an overlap in receptor profiles.

No tumor uptake was originally observed for the H1299.2 and HCC15.2 peptides. Yet acetylation of the amino-terminus rescues the peptides' in vivo homing capabilities, suggesting the amino-termini are rapidly cleaved in vivo and this region is critical for binding. Acetylation of H1299.2 results in a peptide with specificity for the H1299 tumor compared to the H460 tumor and the lungs (5-fold and 7-fold respectively for animal shown, average for all animals 3.5 ± 0.80 and 3.3 ± 1.8). This peptide displays negligible uptake in the liver. The HCC15.2 peptide is not as specific as the other peptides. It accumulates in both H1299 and H460 tumors and shows almost equal uptake in the lung. Due to the lack of specificity, in vivo experiments with this peptide were not continued. Loss of specificity may be due to the acetylation or instability of peptide in vivo. Alternatively, the peptide may not have the affinity necessary to drive tumor accumulation. Finally, we cannot rule out the possibility that the cellular receptor for HCC15.2 is ubiquitously expressed.

Discussion

We have expanded the panel of NSCLC targeting ligands with features that will facilitate their use: high affinity, specificity for cancer vs. normal cells, ease of chemical synthesis, and the ability to home to
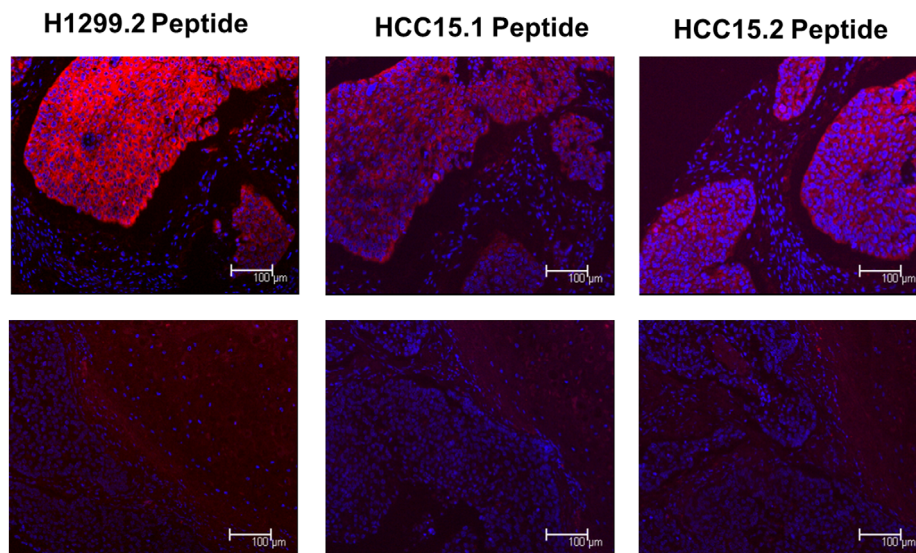
Figure 6 | Peptide-dye conjugates bind to human NSCLC tumors. Triplicate Imgenex arrays containing 59 human NSCLC tumor samples were stained with 10 nM Peptide-SAQdot 605 and DAPI. A representative squamous cell lung cancer tumor sample which was scored as positive for peptide binding is shown along the top row. All three panels in a row are from the same tumor sample. By comparison, a different squamous carcinoma sample scored as negative for peptide binding is seen on the bottom row. Magnification of 100× shown and the scale bar represents 100 microns.
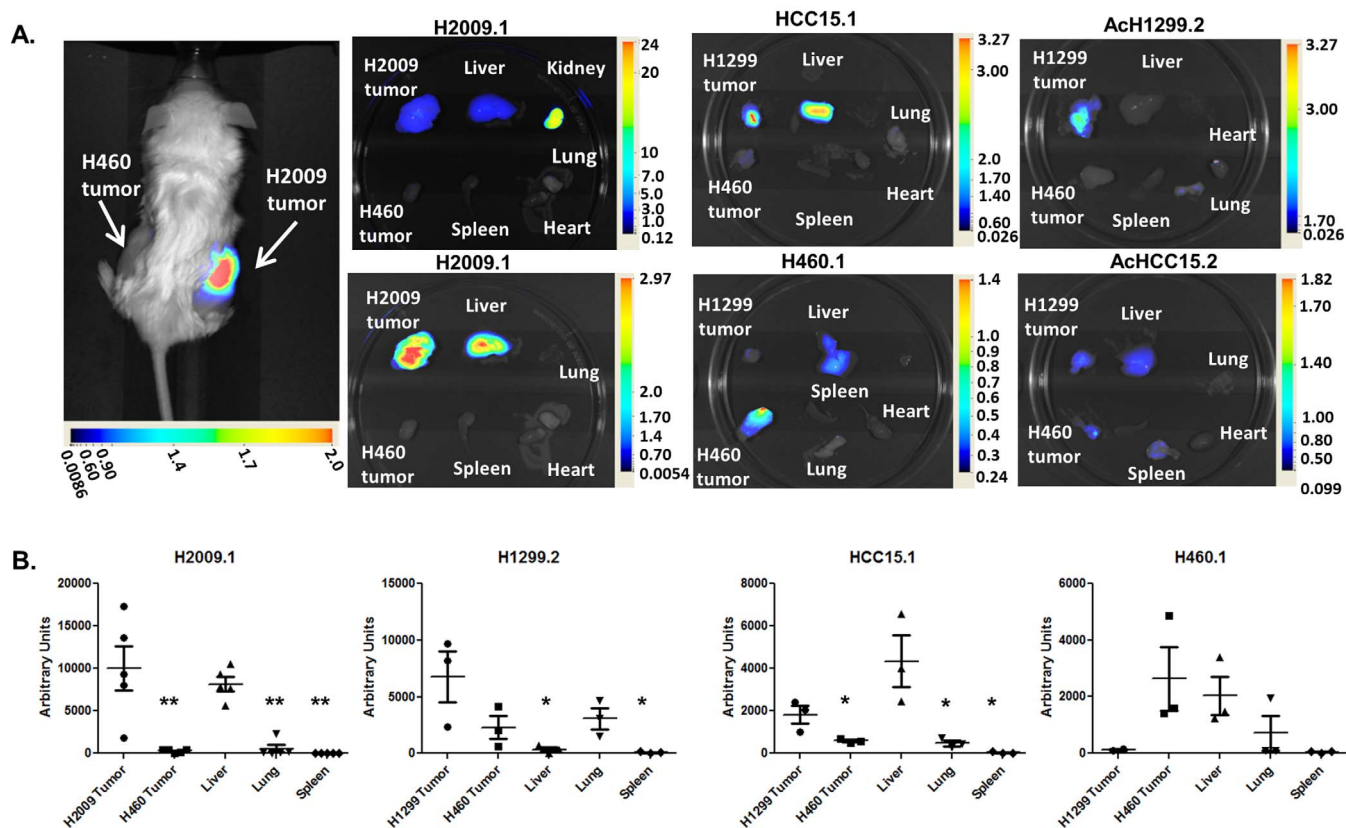
Figure 7 | The peptides home to specific tumors *in vivo*. (A). NOD/SCID mice bearing either dual H2009 and H460 xenografts or dual H1299 and H460 xenografts were injected via tail vein with 30–53 μg of one near-infrared dye labeled tetrameric peptide. The H2009.1 peptide was injected into mice bearing dual H2009 and H460 tumors while the H460.1, HCC15.1, acetylated HCC15.2, and acetylated H1299.2 peptides were injected individually into mice bearing dual H1299 and H460 tumors. At 24 hours after peptide injection, each mouse was imaged for dye fluorescence using a LI-COR Pearl Impulse. Immediately afterwards, the mice were sacrificed for *ex vivo* imaging of organs (heart, liver, spleen, kidneys, lung, and tumors). *Ex vivo* imaging of the tumors and organs revealed significant clearance through kidney for all peptides tested. With the kidney removed from the plate, imaging identifies accumulation in the expected tumor type with additional clearance in the liver. The intensity scale has been optimized for each animal experiment. Pseudocolor scales are shown to the right of each panel. (B). The distribution of peptide accumulation in individual animals is shown. Groups were compared to the target tumor vs. other tissue using an unpaired t-test. ** P 0.001–0.01, * $P < 0.01$–0.05.

tumors *in vivo*. Five of these ligands are functional as free peptides with affinities that rival those of MAbs. Importantly, the peptides bind to the cancer cells with a 40–2000-fold specificity compared to a normal human bronchial epithelial cell line. Of note, these peptides do not stimulate growth or affect cell viability (Supplemental Figure S5). Thus these peptides do not activate growth signaling nor are they expected to be toxic *in vivo*.

The peptides show discrete binding patterns on a panel of cell lines; greater than 80% of the cancer cell lines tested bind to at least 1 peptide. By contrast, HER-2, the target of clinically used MAb trastuzumab (Herceptin®), is expressed in 25% of breast cancers[20]. Binding of the peptides extends to other cancer types as well. In sum, this panel of peptides may cover a large fraction of NSCLC cases encountered in the clinic. While we focused on 5 of the peptides, the remaining 6 may also prove useful.

Two other NSCLC-binding peptides have been isolated from phage displayed peptide libraries by other labs[21,22]; neither show sequence similarity to the peptides identified here. As their binding profile has yet to be determined on a large cell line panel, they may have different binding profiles that supplement our collection of NSCLC-binding ligands. A large scale biopanning has been performed on the NCI-60 panel of cancer cell lines[23]. The library design and selection protocol differ from that used here. Specifically, large scale sequencing was performed at round three of selection to determine presence of tripeptide consensus sequences that represent broad specificity. Additionally, the peptides were not tested for affinity, cellular specificity, or uptake. Of the 38 tripeptide motifs identified, only the GVR motif is observed in our collection of peptides. However, GVR is at the C-terminus of the H2009.1 peptide in a region known not to be involved in cellular binding[24].

Using phage clones in clinical applications is not feasible; peptides must be functional outside the context of the phage. Poor to moderate affinities of isolated peptides has limited the number of clinically viable peptides obtained from phage-displayed libraries. We have overcome this problem by multimerization of the cell-targeting peptides on a trilysine core. The synergistic boost in affinity indicates that these scaffolds take advantage of multimeric binding on the cell surface. The peptides maintain their binding specificity. Although selected for binding *in vitro*, the peptides are able to home to tumors in animals with minimal uptake in other organs. Thus, we have generated peptides with affinities and specificities similar to antibodies but with 1/10 the molecular mass, combining the favorable characteristics of both MAbs and peptides.

The specificity of the peptides indicates that they bind distinct features on the cell surface. The underlying cellular biology affects expression, localization, modifications, and activity of cellular receptors. A major hurdle will be determining the cellular receptor for each peptide. In a recent survey, cellular receptors have been identified for only 15% of peptides identified by biopanning[7]. Biochemical approaches for target identification are complicated due to difficulties in dealing with membrane proteins, especially low abundance receptors. Additionally, specificity of the peptides may arise not from total protein expression but from a rearrangement of receptors to the cell surface. This information can be lost upon the preparation of membrane protein for affinity purification. Furthermore, it is possible that the cellular target is not a proteinacous receptor but a phospholipid or carbohydrate. The cellular receptor for H2009.1, αvβ6, was identified by sequence similarity with a ligand for this integrin[24]. BLAST analysis for the remaining 10 peptides found no significant matches. Additionally, RPPA analysis did not yield any candidate receptors. While receptor identification will provide important biological information, the peptides can be used without knowledge of the receptor.

The similar binding profile of H1299.2, HCC15.1, and HCC15.2 peptides suggests that they are binding the same receptor or receptors that are co-regulated. It is surprising that they would bind the same receptor site as they share no significant sequence similarity. It is possible that the peptides bind the same receptor but at different binding sites. Alternatively, the peptides may bind receptors that are co-regulated. Unfortunately, the peptide blocking experiments do not conclusively discriminate between these scenarios. However, even if they do bind the same receptor, the *in vivo* biodistribution data support the importance of identifying multiple ligands from different peptide libraries. Case-in-point, the H1299.2 and HCC15.1 peptides have similar tumor uptake but H1299.2 has significantly less liver accumulation.

Surrogate markers for tumor phenotypes or genotypes can be of clinical use. We explored the possibility that peptide binding profiles could classify cells accordingly. Although preliminary, clear patterns emerge. Cell lines which bind H2009.1 display a protein phosphorylation pattern consistent with EGFR activation and downstream signaling via STAT5. Met protein expression is higher in the H2009.1 binding cells as is MACC-1, a key regulator of Met expression. Up-regulation of Met protein and increased Met phosphorylation is consistent with known interaction of the Met/EGFR signaling pathways. However, Met phosphorylation is dramatically increased compared to total Met expression and expression of EGFR, HER2, and STAT5 are similar for both group; increased phosphorylation of these proteins represents increased signaling, not just protein overexpression. Of note, only 2 of the cell lines have activating EGFR mutations. It remains to be determined if αvβ6 plays a role in EGFR signaling or is simply a read-out. None-the-less, this suggests that the H2009.1 peptide may be useful in detecting activation of EGFR *in vitro* and *in vivo*.

We also found a striking correlation with an epithelial or mesenchymal phenotype and the peptide binding profile. Cell lines in cluster 1, which binds the 4 peptide panel (H2009.1, H1299.2, HCC15.1, and HCC15.2) are almost exclusively epithelial while cluster 2 is comprised primarily of cells denoted as having a mesenchymal phenotype. Binding of the H1299.2, HCC15.1, and HCC15.2 triplet is not sufficient to predict a mesenchymal phenotype; the cells must also be αvβ6 negative. While all cells that bind H2009.1 and therefore express αvβ6 have an epithelial phenotype (except H2073), lack of αvβ6 expression does not demand a mesenchymal phenotype. The role of αvβ6 in EMT is conflicting[25,26]. While expression of αvβ6 has been found to be a marker for an aggressive phenotype and poor patient outcome, other reports suggest it is a tumor suppressor[26–29]. In fact, we previously reported it to be a marker of EMT in early stage NSCLC[30]. As αvβ6 activates latent-TGF-β, it is likely to have a similar duality[31–33]. Further large scale studies on tumor samples need to be performed to determine if the peptide binding profile has similar predictive ability.

These peptides add to the growing toolbox of tumor targeting reagents. The high specificities and affinities suggest that these peptides have potential in a clinical setting, much in the same way antibodies have been used. Importantly, the versatility of the peptide platform allows for easy incorporation into a variety of drug delivery and imaging formats. These peptidic ligands have potential to complement currently used antibody targeting ligands and may expand the number of clinical cases that can be treated with a targeted therapy.

Methods

laboratory animals. Imaging experiments were carried out with a minimum of 3 animals per group per experimental condition. Errors are represented as standard error measurements (SEM) unless explicitly stated otherwise.

Isolation of Cell-Specific Peptides. Selection of peptides was performed as described[10]. Briefly, cells were seeded in a 12 well plate and allowed to reach ~90% confluency. Complete media was removed and replaced with media without serum to clear receptors. After 2 hours, media was removed and replaced with 10–100 phage library equivalents in 1 ml PBS supplemented with 10 mM $MgCl_2$, 0.5 mM $CaCl_2$ (PBS^+) and 0.1% bovine serum albumin (PBS^+/BSA). The mixture also contained 0.1 mM chloroquine and complete protease inhibitor cocktail without EDTA (Roche). Cells were incubated with the phage cocktail for 1 hour after which they were washed 4 times with 1 ml PBS^+/BSA and 2 times with 0.1 M HCl-glycine pH 2.2 in 0.9% NaCl. Cells were incubated 30 minutes in 30 mM Tris-HCl, pH 8.0, at 4°C and disrupted by freeze-thaw. Phage contained in the cell lysates were amplified in *E. coli*. Three peptide libraries were employed: a 20-mer peptide library of 10^8 different peptide members previously described[35], a second 20-mer library generated in-house[14] and a commercially available 12-mer library with 10^{10} peptide sequences (Ph.D.-12™, New England Biolabs). Starting at round three, 24 individual phage clones were sequenced by colony PCR[12]. The panning was repeated until sequence analysis revealed that one or more phage clones had emerged as the predominant species in the output phage pool. Of note, negative selections on normal control cells were not performed.

To compare the selected peptides using the same assay format (i.e., colony forming units), the peptide sequences obtained from the 12mer library was placed in our standard phage vector. Briefly, fAFF1 vector was cut with BstXI and ligated in-frame to annealed synthetic oligonucleotides encoding the sequence of the HCC15.2 peptide or the 12mer control peptide. The constructs were transformed into electrocompetent K-91 *E. coli*.

Selectivity Calculations. Phage binding assays were performed as described using 1×10^8 phage particles of a single clone[10]. Phage were incubated with the cell type of interest in the presence of 0.1 mM chloroquine and $1\times$ protease inhibitor in PBS^+/BSA. After 10 minute incubation at 37°C, the cells were washed and lysed as described. The quantity of cell associated phage was determined by titering serial dilutions. Selectivity is defined as the ratio of the uptake of the targeting phage to the total phage input versus that of a control phage[14,15]. Control phage displays the sequence NQRGTELRSPSVDLNKPGRH (20mer control) or QLGAHNNPRYAP (12mer control). Both peptides display the same affinity as the whole library and an empty phage containing no peptide. Experiments were repeated a minimum of 3 times. Errors are represented as standard error measurements (SEM).

Peptide Synthesis. Peptide synthesis was performed on a Symphony Synthesizer (Rainin Instruments) by Fmoc solid-phase peptide synthesis on Rink Amide AM resin (substitution level 0.71 mmol/g). Peptides were purified by reverse phase HPLC employing a C18 column. Maleimido tetrameric cores were synthesized by sequential coupling (Fmoc)₂Lys on Fmoc-β-Ala-CLEAR™ Acid Resin, (substitution level 0.40 mmol/g) and purified by reverse phase HPLC[16]. Tetrameric peptides were synthesized by coupling the monomeric peptide to the activated multimeric core (8 : 1 mole ratio of the peptide: maleimide core) for 2 hours in Ar-purged PBS, pH 7.4, containing 0.01 M EDTA. Excess monomeric peptide was removed by reverse phase HPLC. Peptides were characterized by MALDI mass spectrometry. Peptides are ≥95% pure as assessed by HPLC.

For near infrared imaging (NIR) imaging, the tetrameric peptides possessing a cysteine before the trilysine branch (168 nmol) and IRDye 800 CW Maleimide (0.1 mg, 84 nmol) were dissolved in 1.5 mL PBS, pH 7.4, containing 0.01 M EDTA and stirred for 2 hours. The products were purified by RP-HPLC and characterized by MALDI MS.

Peptide Blocking Experiments. Peptide stock solutions were prepared in PBS, pH 7.4. Concentrations were confirmed by Edman sequencing or by absorbance at 280 nm if the peptide contained a tyrosine or tryptophan. The ability of the peptides to block their cognate phage was determined as described, varying the peptide concentration from 1 pM to 1 µM[10,16]. Peptide stock solutions were diluted in PBS^+/BSA containing 10^8 phage particles, chloroquine (0.1 mM) and $1 \times$ protease inhibitor and incubated with cells for 10 min. Cells were washed, lysed, and titered as described. Peptide blocking is defined as the output phage to input phage ratio in the presence of the peptide compared to the ratio without added peptide. Half-maximal binding affinities were determined by nonlinear regression analysis using GraphPad Prism (Supplemental Table S-III).

Peptide Binding Experiments. Peptide conjugates with Steptavidin-phycoerythrin (1 : 1 ratio, SAPE: Peptide, BD Bioscience) were incubated with cells at a final peptide concentration of 10 nM for 1 hour at 37°C. Peptide-SAPE was removed and the cells were washed 3 times in PBS, 2 times in acid wash pH 2.2, and removed from the well in PBS containing 10 mM EDTA. Peptide-mediated SAPE uptake in cells was determined by flow cytometry on a Cell Lab Quanta (Beckman Coulter Inc.). For each set of experiments the instrument settings were adjusted to a normalized value based on fluorescence of H1299 cells accumulating 1299.2-SAPE or H2009 cells accumulating 2009.1-SAPE. Nonspecific uptake and fluorescence controls included H1299 or H2009 cells incubated without SAPE, unconjugated SAPE, or a scrambled sequence control peptide conjugated to SAPE. For each sample, 10,000 events were

counted. Collected data was evaluated using WinMDI v2.9 flow analysis software. Single live cells were gated by electronic volume and side scatter measurement. Gated cells were considered to be positively stained if their fluorescence intensity was above the first decade. In order to compare staining across the panel of cells, we generated a staining score for each sample. The staining score is defined as the percentage of positively stained gated live, single cells multiplied by the mean fluorescence intensity of that population. Peptide-SAPE uptake was assessed for each cell line in three separate assays. The staining scores were averaged. Staining score averages below 5000 are considered negative for significant uptake of the individual peptide conjugate.

Fluorescent Microscopy. Streptavidin coated Qdots605 (SAQdot605, 20 nM, Life Technologies) and tetrameric peptides (100 nM) were incubated in PBS for 30 minutes. After conjugation, the mixture was diluted with an equal volume of RPMI containing 5% bovine serum to saturate the remaining biotin binding sites. Cells were seeded onto poly-D-lysine-coated, 8-well chamber slides (BD Biosciences) using 10,000 cells per well. After adhering overnight, growth medium was removed and replaced with the peptide-SAQdot solution (final concentration: 10 nM SAQdots and 50 nM peptide). Cells were treated for 10 minutes at 37°C, washed with PBS^+/BSA, rinsed twice with 0.5 ml/well glycine-HCl and fixed with PBS-buffered formalin for 10 minutes. The cell membrane was labeled with wheat germ agglutinin coupled to AlexaFluor-488. DAPI was employed for nuclear visualization.

Reverse Phase Protein Arrays (RPPA). The RPPA were previously generated and analyzed for protein expressions[18,19].

Tumor Microarray Binding. Human lung cancer (IMH-305) and matching normal tissue (IMH-340) microarrays were purchased from Imgenex Inc. Each microarray contained 59 independent tissue samples. The distribution of the lung cancer by histopathological subclasses included 37 squamous, 5 large cell, 8 adenocarcinoma, 7 BAC, and 2 carcinosarcomas. Paraffin was removed using EZ-DeWax solution (Biogenex). Slides were rinsed and incubated in a 65°C bath of 10 mM Citrate-Citric acid, pH 6.0 for 10 min. Following washes with PBS, slides were blocked with PBS + 2.5% BSA. After antigen retrieval, microarrays were treated with 10 nM peptide conjugated 1 : 1 with SAQdot605 in 50% PBS 50% RPMI with 2.5% bovine serum. Microarrays were incubated with the peptide-Qdot solution for 30 minutes. Slides were washed with PBS and samples were covered in Vectashield + DAPI (Vector Laboratories). Images were collected from three areas of each microarray tissue sample using a Leica fluorescence microscope and LAS-AF software. Peptide-mediated Qdot staining was identified in each sample by comparing the red Qdot channel with background intrinsic fluorescence of each sample using the green fluorescence channel. The scoring of positive staining was determined by five independent scientists.

In Vivo Targeting. Animal protocols were approved by the Institutional Animal Care and Use Committee at UT Southwestern. NOD/SCID mice were injected with either H2009 or H1299 cells in the right flank and with H460 cells in the left flank, so that each mouse bore either dual H2009 and H460 xenografts or dual H1299 and H460 xenografts. When tumors reached ~100 mm³ mice were injected via tail vein with 30–53 µg of NIR dye labeled tetrameric peptide, based upon the molecular weight of the peptide (see Table S-VIII for exact amounts per mouse) and imaged using a LI-COR Pearl Impulse 24 hours post injection. Immediately afterwards, the mice were sacrificed for *ex vivo* imaging of organs.

Statistical Analysis. Statistical analysis was conducted using SAS 9.1.3 Service Pack 4. Selectivity values were dichotomized as Low (<5000) and High (≥5000) based on flow cytometry analysis. Association between mutation status and the dichotomized selectivity value was examined using Fisher's exact test. In vivo images were analyzed using GraphPad Prism using an unpaired t test. All the reported p-values are two-sided. A p-value of ≤0.05 is considered statistically significant.

1. Siegel, R. *et al.* Cancer treatment and survivorship statistics, 2012. *CA. Cancer J. Clin.* **62**, 220–241 (2012).
2. Brambilla, E. & Gazdar, A. Pathogenesis of lung cancer signaling pathways: Roadmap for therapies. *Eur. Respir. J.* **33**, 1485–1497 (2009).
3. Chung, C. H. *et al.* Cetuximab-induced anaphylaxis and ige specific for galactose-α-1,3-galactose. *N. Engl. J. Med.* **358**, 1109–1117 (2008).
4. Bray, B. L. Large-scale manufacture of peptide therapeutics by chemical synthesis. *Nat. Rev. Drug Discov.* **2**, 587–593 (2003).
5. Nagy, A. & Schally, A. V. Targeting cytotoxic conjugates of somatostatin, luteinizing hormone-releasing hormone and bombesin to cancers expressing their receptors: A "smarter" therapy. *Curr. Pharm. Des.* **11**, 1167–1180 (2005).
6. Aina, O. H. *et al.* From combinatorial chemistry to cancer targeting peptides. *Mol. Pharm.* **4**, 631–651 (2007).
7. Brown, K. C. Peptidic tumor targeting agents: The road from phage display peptide selections to clinical applications. *Curr. Pharm. Des.* **16**, 1040–1054 (2010).
8. Gray, B. P. & Brown, K. C. Combinatorial peptide libraries: Mining for cell-binding peptides. *Chem. Rev.* **114**, 1020–1081 (2014).
9. De, J. *et al.* Isolation of a *Mycoplasma*-specific binding peptide from an unbiased phage-displayed peptide library. *Mol. Biosyst.* **1**, 149–157 (2005).

10. Mcguire, M. J., Li, S. & Brown, K. C. (2009) Biopanning of phage displayed peptide libraries for the isolation of cell-specific ligand, In *Biosensors and Biodetection Methods and Protocols* (Walker, J. M., Ed.), pp 291–321. Humana Press, New York.

11. Mcguire, M. J., Samli, K. N., Chang, Y. & Brown, K. C. Novel ligands for cancer diagnosis: Selection of peptide ligands for identification and isolation of B-cell lymphomas. *Exp. Hematol.* **34**, 443–452 (2006).

12. Mcguire, M. J., Samli, K. N., Johnston, S. A. & Brown, K. C. In vitro selection of a peptide with high selectivity for cardiomyocytes in vivo. *J. Mol. Biol.* **342**, 171–182 (2004).

13. Mcguire, M. J. et al. A library selected Langerhans cell-targeting peptide enhances an immune response. *DNA Cell Biol.* **23**, 742–752 (2004).

14. Oyama, T. et al. Isolation of multiple cell-binding ligands from different phage displayed-peptide libraries. *Biosens. Bioelectron.* **21**, 1867–1875 (2006).

15. Oyama, T. et al. Isolation of lung tumor specific peptides from a random peptide library: Generation of diagnostic and cell-targeting reagents. *Cancer Lett.* **202**, 219–230 (2003).

16. Li, S. et al. Synthesis and characterization of a high affinity $\alpha_v\beta_6$-specific ligand for in vitro and in vivo applications *Mol. Cancer Ther.* **8**, 1239–1249 (2009).

17. Ramirez, R. D. et al. Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. *Cancer Res.* **64**, 9027–9034 (2004).

18. Byers, L. A. et al. An epithelial–mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. *Clin. Cancer Res.* **19**, 279–290 (2013).

19. Byers, L. A. et al. Reciprocal regulation of c-Src and STAT3 in non-small cell lung cancer. *Clin. Cancer Res.* **15**, 6852–6861 (2009).

20. Slamon, D. J. et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* **244**, 707–712 (1989).

21. Chang, D.-K., Lin, C.-T., Wu, C.-H. & Wu, H.-C. A novel peptide enhances therapeutic efficacy of liposomal anti-cancer drugs in mice models of human lung cancer. *PLoS ONE* **4**, e4171 (2009).

22. Zang, L. et al. Screening and identification of a peptide specifically targeted to NCI-H1299 from a phage display peptide library. *Cancer Lett.* **281**, 64–70 (2009).

23. Kolonin, M. G. et al. Ligand-directed surface profiling of human cancer cells with combinatorial peptide libraries. *Cancer Res.* **66**, 34–40 (2006).

24. Elayadi, A. N. et al. A peptide selected by biopanning identifies the integrin $\alpha_v\beta_6$ as a prognostic biomarker for non-small cell lung cancer. *Cancer Res.* **67**, 5889–5895 (2007).

25. Ramos, D. M., Dang, D. & Sadler, S. The role of the integrin alphav beta6 in regulating the epithelial to mesenchymal transition in oral cancer. *Anticancer Res.* **29**, 125–130 (2009).

26. Bates, R. C. Transcriptional activation of integrin $\beta6$ during the epithelial-mesenchymal transition defines a novel prognostic indicator of aggressive colon cancer. *J. Clin. Invest.* **115**, 339–347 (2005).

27. Ahmed, N. et al. $\alpha_v\beta_6$ Integrin- a marker for the malignant potential of epithelial ovarian cancer. *J. Histochem. Cytochem.* **50**, 1371–1379 (2002).

28. Hazelbag, S. et al. Overexpression of the $\alpha_v\beta_6$ integrin in cervical squamous cell carcinoma is a prognostic factor for decreased survival. *J. Pathol.* **212**, 316–324 (2007).

29. Yang, G.-Y. et al. Integrin alpha$_v$beta$_6$ mediates the potential for colon cancer cells to colonize in and metastasize to the liver. *Cancer Sci.* **99**, 879–887 (2008).

30. Prudkin, L. et al. Epithelial-to-mesenchymal transition in the development and progression of adenocarcinoma and squamous carcinoma of the lung. *Mod. Pathol.* **22**, 668–678 (2009).

31. Bierie, B. & Moses, H. L. Tumour microenvironment: TGFβ: The molecular Jekyll and Hyde of cancer. *Nat. Rev. Cancer* **6**, 506–520 (2006).

32. Annes, J., Chen, Y., Munger, J. & Rikfin, D. Integrin αvβ6-mediated activation of latent TGF-β requires the latenr TGF-β binding protein-1. *J. Cell Biol.* **165**, 723–734 (2004).

33. Worthington, J., Klementowicz, J. & Travis, M. TGFβ: A sleeping giant awoken by integrins. *Trends Biochem. Sci.* **36**, 47–54 (2011).

34. Phelps, R. M. et al. NCI-Navy medical oncology branch cell line database. *J. Cell. Biochem.* **63**, 32–91 (1996).

35. Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. Peptides on phage: A vast library of peptides for identifying ligands. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 6378–6382 (1990).

Acknowledgments

Author contributions

M.J.M. performed the peptide binding studies by flow cytometry, fluorescent microscopy experiments, and peptide selections, and had significant input on experimental design and data analysis. B.P.G. performed the in vivo biodistribution experiments, peptide synthesis and assisted with the fluorescent microscopy. S.L. and D.C. performed peptide synthesis and the peptide affinity assays. Y.H.L., L.W., N.P. and T.O. performed peptide selections and determined the phage clone selectivity across the cell line panel. S.R. and J.I. provided general lab assistance and participated in the phage clone selectivity experiments. L.A.B. and J.V.D. provided RPPA and performed the analysis of the arrays. L.D. and X.X. performed statistical analysis. J.D.M. provided cell lines along with their genetic characterization. K.C.B. designed experiments, supervised the work, analyzed data, provided financial support and wrote the paper.

Additional information