Methods

Reagents

Compounds were obtained from Selleck Chemicals, Tocris Bioscience, Merck Millipore, Sigma Aldrich or as a kind gift from Lead Discovery Center GmbH. They were diluted in DMSO, aliquoted and stored as 10mM stocks at -80°C. Fibroblast Growth Factor (FGFs) proteins were provided by ProSpec, dissolved in water and stored at -20°C. Heparin Solution (0.2%) was purchased from StemCell Technologies and stored at 4°C.

Viability assays and compound activity prediction

Cell lines were plated as triplicates into sterile 96-well plates at 1500 cells/well density as described previously [1]. After 24 hours of incubation, compounds were added at increasing dosages, ranging from 30μM to 0.005 μM together with a separate DMSO control. After 96 hours, relative cell viability was determined by comparing the ATP-content of each well - assessed by CellTiter Glo Assay (Promega, US) - to the content of the DMSO control. Finally, half-maximal growth inhibitory concentrations (GI_{50}) were calculated by the package "ic50" (R programming language) (1).

Apoptosis assays

For analysis of apoptosis, the Annexin V-FITC Apoptosis Detection Kit I (BD Biosciences) was used. Retrovirally transduced (pBabe) NIH3T3 cells were seeded in 6 cm dishes at 30% confluence in cell culture medium containing puromycine (3μg/ml). After 24 hours, supernatant were refreshed and cells were treated with PD173074 (1μM) and DMSO, respectively, for 72 hours. Subsequently, cells were detached by trypsin, washed with cold PBS, incubated with accutase solution (Sigma Aldrich) for one minute, and resuspended in Annexin-V binding buffer (BD Biosciences). Finally, cells were stained with FITC-labeled Annexin V antibody and Propidium Iodide (PI) and incubated in the dark for 20 minutes. Analysis was performed on a FACS Gallios Flow Cytometer

(Beckman Coulter), measuring at least 100,000 events per probe. For calculation of apoptosis, changes from DMSO control to treated samples were evaluated by setting appropriate gate in Kaluza analysis software (Beckman Coulter).

cDNA Transcription

RNA was isolated from cancer cell lines or NIH3T3 cells using TRIZOL reagent (Invitrogen) and cleaned up using the RNeasy MinElute Cleanup Kit (Qiagen) following the manufacturers' protocols. Finally, 1 µg of RNA was transcribed into cDNA using Superscript III reverse transcriptase (Invitrogen, #18064).

FGFR1 Cloning and Site-Directed Mutagenesis

cDNA of H1581 cells (100ng) was used to amplify FGFR1 by attB-overhang primers (**Fig. S30**) and flip it into pDONR.221 using the BP-clonase (Invitrogen). Bacterial transformation of the competent E. coli strain DH5α (Invitrogen) was carried out according to the manufacturer's instructions. Single clones were sequenced from mini-preparation of plasmid DNA using the NucleoSpin Mini Kit (Machery Nagel). For midi-preparation of plasmid DNA we used the NucleoBond Xtra Midi EF Kit (Machery Nagel). pDONR-FGFR1α and β were flipped into the retroviral vector backbones of pBabe-puro, -neo or -hygro gateway (GW) using the LR Clonase Kit (Invitrogen).

For site directed mutagenesis of the pBabe-puro-FGFR1β plasmid we used QuickChange II XL Site-Directed Mutagenesis Kit (Agilent) in order to integrate the following point mutations: V472M, L76T plus V472M, A78L plus V472M, K83E plus V472M, D157N plus V472M, D193N plus V472M and Q195E plus V472M.

Quantitative Real-Time PCR

Quantitative real-time PCR was performed using a 7300 Real-Time PCR System (Applied Biosystems) and Power SYBR Green PCR Master Mix (Applied Biosystems) with primer

pairs (**Fig. S29**) specific for GAPDH (QT01192646, Qiagene) (58°C), MYC (58°C), FGFR1 (56°C), FGFR2 (56°C), FGFR3 (56°C) and FGFR4 (56°C). ΔC_t-values were determined using the 7300 System Software (Applied Biosystems) using GADPH as reference control. Gene expression was calculated by $\Delta\Delta C_t$-method.

Virus production

HEK293T cells were plated on 6 cm dishes in DMEM + 10% FCS and incubated over night at 37 °C. The next day, cells were transfected with retroviral plasmids. For this, 12µl TransIT-LT1 (Mirus) was added drop-wise to 400 µl OptiMem medium (Invitrogen). In a separate tube, 4 µg of pBabe expression plasmid was mixed with 4 µg of pCL-eco or pCL-ampo packaging plasmid in 400 µl OptiMem medium. After 5 minutes of incubation, both tubes were mixed carefully and incubated at room temperature for 20 minutes. Subsequently, this mixture was added to HEK293T cells. The next day, medium was removed and changed to DMEM + 30% FCS. After 24 hours and 48 hours, supernatants were collected and centrifuged at 200 g for 5 minutes, filtered, aliquoted and stored at -80 °C.

Similarly, replication-incompetent lentivirus was produced from pLKO.1-puro vector containing a short hairpin RNA (shRNA) specific for the respective target gene. For this HEK293T cells were co-transfected with Δ8.9, pMGD2 and pLKO.1 vector as described previously [1]. Viral titers were determined by transduction of NIH3T3 cells (ATCC) at increasing virus dilutions. Hereby the virus titer was calibrated equally for all samples.

RNAi and stable transduction

Cancer cell lines were transduced by lentiviral supernatants at equal titers in the presence of polybrene (10µg/ml) for 24 hours and selected by puromycine (1 - 3µg/ml). Relative cell survival was calculated as ratio to the empty-vector construct (Addgene). Knockdown efficacy was validated by immunoblotting. The following target sequences

were used for MYC and FGFR2, respectively:

CCTGAGACAGATCAGCAACAA (*shMYC*),

Stable cDNA expression

Cancer cell lines and NIH3T3 cells were transduced by retroviral supernatants in the presence of polybrene (10µg/ml) for 24 hours and selected by puromycine (3 µg/ml), G418 (800 µg/ml) or hygromycine (400 µg/ml), respectively, for 2-3 weeks. NIH3T3 cells transduced with FGFR1 were incubated with 15ng/ml FGF2 and 2µg/ml heparin (StemCell). Expression of the respective cDNA was confirmed by immuoblotting or quantitative real-time PCR. Finally, cells were expanded and frozen in liquid nitrogen for long-term storage.

Soft-Agar Assay

All soft-agar experiments were performed as triplicate in 96-well plates. For this, we re-suspended cell pellets in growth media containing 10% FCS and 0.6% agar and plated 1000 cells in 50 µl per well. Subsequently, 50 µl of solidified growth medium (10% FCS; 1.0% agar) was added on top. After 3-4 weeks of incubation at 37°C and 5% CO_2, colonies were analyzed by Scanalyzer imaging system (LemnaTec) or counted by hand.

Immunoblotting

Cells were washed with cold PBS and lysed in lysis buffer (Cell Signaling) supplemented with protease (Roche) and phosphatase inhibitor (Calbiochem) cocktails. After 20 minutes of incubation on ice, lysates were were centrifuged at 18,000 g for 25 minutes. Protein concentration in supernatants was measured using BCA Protein Assay (ThermoScientific). Equivalent amounts of protein (30–60µg) were denatured and separated on 4–12% SDS-PAGE gels and after blotting on nitrocellulose membranes (Amersham Hybond-C Extra). The following antibodies were used for immunoblotting:

β-actin (MP Bioscience); phospho-FGFR (Tyr^{653}, Tyr^{654}), phospho-FRS2 (Tyr^{196}), phospho-AKT (Ser^{473}), AKT, phospho-ERK, and ERK, c-myc (Cell Signaling Technology); total FGFR1 (Epitomics / Abcam); caspase-3 (Cell Signaling 9662S); cyto-chrom C (BD Pharmingen, mouse); cyclin D1 (Santa Cruz); conjugated antibodies to rabbit and mouse (Millipore).

Immunohistochemistry

Tissues were fixed in 4% PBS-buffered formalin and embedded in paraffin (FFPE). Immunohistochemistry was performed as described previously [2] on 3 µm slides with specific antibodies for pFGFR1 (Abnova, Y154) and MYC (Abcam). Staining intensities were individually evaluated by 3 independent observers, using a 4-tier scoring system. The areas of highest staining intensity was scored. Examples nuclear MYC and cytoplasmic and membranous pFGFR1 staining are shown in Figure S13. Statistical analysis was performed using a Fisher's exact test.

Computational Analysis

In total, we analyzed segmented copy number data of a collection of 306 primary squamous cell lung cancer samples (CLCGP data set, manuscript submitted). All scripts applied were performed using R Statistical Computing Environment as well as Java and C Language Compilers. Calculation was performed using a 2.8 GHz Intel Core i7 Processor with 8GB DDR3 Memory on Mac OS X Version 10.7.5 operating system. Further validation of the procedures briefly described will be provided elsewhere (manuscript in preparation).

Macro Lesion Purified CN Index (Smoothening):

For calculation of purified copy number indices, segmented inferred copy numbers were normalized by locally expected reference copy number data. The latter were computed for each sample separately by a segment-wise model, particularly taking its individual

macro lesions into account. For further analysis, copy indices were segmented and analyzed by the Genomic Identification of Significant Targets in Cancer (GISTIC) algorithm (GenePattern Platform, Broad Institute, [3]).

Hierarchical Clustering:

In order to cluster copy number data, which are continuously recorded over the human genome (Hg18), we had to segment the latter **first**. For this, we derived the segmentation boarders from Affymetrix Probe Localization Files, so that segment length were adapted to aberration frequencies in cancer. This resulted in 213 distinct genomic localizations. **Secondly**, we derived the segment wise copy numbers from macro lesion purified copy number indices. **Thirdly**, 64 segments, meeting a predefined copy number variance threshold, as well as 69% of samples, which showed at least one distinct copy number aberration, were selected to derive a 64 x 210 copy number matrix. This was used as input for pairwise complete-linkage column-wise hierarchical clustering (GenePattern Platform, Broad Institute, [3]), which was performed using Pearson correlation metrics (**Step 4**). Finally, clusters were defined by significance as to t-distribution as well as positive correlation, and sorted by their amplification maximum subsequently (**Step 5**).

Average ranks indices:

For each recurrent amplicon in squamous cell lung cancer, we calculated raw inferred copy numbers directly from SNP 6.0 array data. For each sample, we detected the borders of the amplicon and ordered the included genes by copy number in decreasing order. From this, we derived a relative gene-wise rank index (range: 0-1). An amplicon was labeled as centered, if the average index of its target gene exceeded 0.6. Finally, the fraction of centered amplified samples was calculated for each recurrent amplicon separately.

Chromosomal Breakpoint Analysis:

A distance metrics capturing genomic coherence was derived from segmented copy

number data to decipher genomic regions, prone to chromosomal breakage. Clusters of coherence, centered on loci of maximum genomic consistency, were demarcated if distance metrics exceeded a predefined threshold.

References

[1] Sos ML, Michel K, Zander T, Weiss J, Frommolt P, Peifer M, et al. Predicting drug susceptibility of non-small cell lung cancers based on genetic lesions. J Clin Invest 2009; 119: 1727-1740.

[2] Heukamp LC, Schröder D, Plassmann D, Homann J, Büttner R. Marked Clinical and Histologic Improvement in a Patient with Type-1 Gaucher's Disease Following Long-Term Glucocerebroside Substitution A Case Report and Review of Current Diagnosis and Management. Pathol Res Pract 2003; 199: 159–163.

[3] Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. Nat Genet 2006; 38: 500–1.

Supplementary Figure Legends

Figure S1 *Recurrent amplified regions in SQLC based on purified CLCGP copy number*

Table of 206 genes from 16 recurrent amplified regions in SQLC, based on copy number index data.

Figure S2 *Amplicon focality on copy number aberrations in SQLC*

a) Recurrence of copy number aberrations (GISTIC algorithm) in the TCGA dataset. It is shown the raw segmented inferred copy number data of SQLC tumor samples (left panel, red) and in macro lesion purified CN index (right panel, green). **b)** Representative screenshots of segmented copy number data before (left) and after (right) processing with FAPP algorithm. Genomic regions (15Mbp) containing the *FGFR1* (8p12) gene are displayed (n=132, 56 samples (42%) contain *FGFR1*, cut-of 3.0). Samples were sorted by the genomic coordinate of the highest inferred copy number value. Position of *FGFR1* is highlighted in green.

Figure S3 *Hierarchical clustering of CLCGP SNP data*

Pairwise complete linkage hierarchical clustering as to Pearson correlation of 210 SQLC samples (Columns, 69%) along 64 aberrant regions (Rows, 30%). The hierarchical tree illustrates cluster significance, as derived from Student's t-distribution. Chromosomal mapping (right) was derived from Hg18 annotation.

Figure S4 *Centrality frequencies for a common choice of smoothing parameters*

Copy number aberrations of amplified samples were classified as central, if the average rank index of the respective target gene exceeded 0.6. Amplification cut

off of was 2.5 and 3.0, respectively. The fraction of centered amplicons was compared for the five most significant amplifications in SQLC genome.

Figure S5 *Genotypic cell line annotation*

Genomic aberrations [Seidel et al., revised] were plotted against histotypes of 148 lung cancer cell lines (CCLE database). Lines indicate histology-specific alterations (orange = adenocarcinoma (AD), green = large cell carcinoma (Large), red = small cell lung cancer (SCLC), blue = squamous cell lung cancer (SQLC)). Cell lines included in this study are highlighted (dark green = adenocarcinoma, pink = large-cell carcinoma, turquoise = small-cell lung cancer, yellow = squamous-cell lung cancer)

Figure S6 *Receptor signaling activation H520 [FGFR1^{amp}] and HCC15 [NRAS$^{mut.}$/FGFR1^{del}]*

Cell lines were starved for 24 hours and stimulated by a collection of 6 FGFs (1 ng/ml) and heparin (10 µg/ml) for 20 minutes. Additionally, the FGFR inhibitor PD173074 (1 µM) was added 40 minutes before stimulation by aFGF and bFGF. Subsequently, phosphorylation of MAPK, AKT and the FGFR1 signaling adapter protein FRS2α as well as MYC expression were assessed by immunoblotting.

Figure S7 *Experimental draft for ligand dependency*

Ligand-Dependent: Endogenous receptors binds FGF, dimerize and activate downstream signaling. Exogenous receptor with gatekeeper mutation (FGFR1, V472M) takes over signaling during FGFR-Inhibitor treatment. Exogenous double mutated receptor (FGFR1, R161Q, V472M) cannot take over signaling during therapy (lag of FGF binding). Reconstitution of FGFR1 signaling through FGF stimulation.

Ligand-Independent: Endogenous receptors dimerize and activate downstream signaling. Exogenous receptor with gatekeeper mutation (FGFR1, V472M) takes over signaling during FGFR-Inhibitor treatment. Exogenous double mutated receptor (FGFR1, R161Q and V472M) still signals during therapy (lag of FGF binding).

Figure S8 *Whole transcriptome analysis shows alternative FGFR1 splicing for amplified tumor samples and high MYC expression in FGFR1 centered samples*

a) Exonwise coverage of whole transcriptome sequenced cell lines (H1581, H520, DMS114 and SBC-7)(left) and six patient samples (S00062, S00186, S00473, S01225, S01265 and S01472) (right) are plotted. They are mapped to translated exons of *FGFR1*. Coverage is highly affected by GC content a collection. Low coverage of exon 8 (IIIb) suggests dominance of mesenchymal splice variants, whereas small read density on exon 2 indicates dominance of IIIc-β variants. **b)** Amplified samples were classified as central, if the average FGFR1 rank index exceeded 0.6. Average FPKM-normalized (Whole Transcriptome Sequencing) values of *MYC* expression were compared by t-test. Error bars display standard derivation within the respective groups.

Figure S9 *Validation of gene expression by quantitative real time PCR*

Color intensities indicate expression levels in transduced NIH3T3 cells as assessed by quantitative real time PCR specific for FGFR1 or the genes annotated below. For each gene, three independent primer pairs were used.

Figure S10 *NIH3T3-FGFR1α and HEK298T cells form tumors in nude mice*

a) Nude mice, engrafted with retrovirally transduced NIH3T3 cells, received BGJ398 (15 mg/kg, q.d.) and 5% glucose, respectively, upon formation of palpable tumors. Volumes of tumors formed by NIH3T3-FGFR1α cells (left). Representative immunohistochemical MYC stains (x40 magnification) of subcutaneous mouse tumors after 14-day FGFR inhibitory therapy (right bottom) and vehicle (right top). **b)** HEK cells were retrovirally (pBabe) transduced by FGFR1α and selected by 800 µg/ml G418, subsequently. Protein expression of transduced cells was analyzed by immunoblotting (left). Nude mice were subcutaneously engrafted with HEK-FGFR1α cells and tumor size was measured using an external caliper every second day (right). **c)** *FGFR1*-amplified H520 cells were treated with PD173074 (1µM, 24 hours). Expression levels of MYC, Cyclin D1 and Actin as well as ERK phosphorylation were monitored by immunoblotting.

Figure S11 *Quantification of MYC stains in murine tumor samples*

Tumors explanted from mice were examined for MYC expression by immunohistochemistry prior (black) and post (white) therapy. Average fractions of cells, which display positive MYC stains in their nucleus, are shown as bar plot. At least 1,000 tumor cells in 10 independent fields were counted for each sample. Significant (*) and strongly significant (**) differences are marked by asterisks.

Figure S12 *Apoptosis is accompanied by cytochrome c release and breakdown of mitochondrial potential*

a) H1581 and H358 cells were treated with PD173074 (1µM, 72 hours). For mitochondria precipitation preparation cells were trypsinised and centrifuged (5 min., 300g). 1 volume of Hep-Buffer (20mM Hepes, pH=7.5, 10mM KCl, 1,5

mM $MgCl_2$, 1mM EDTA, 10µM cytochalasin B, 1mM DTT, protease inhibitor) was added and incubated for 15 minutes on ice. Cells were punched 15-20 times with 26G needle and centrifuged for 5 min., 600g. Supernatant was centrifuged again for 10 min., 600g. Supernatant was than centrifuged for 15 min., 10000g and supernatant was loaded on gel. Cytochrome c release and expression levels of MYC and caspase 3 were monitored by immunoblotting. **b)** A panel of seven cell lines was treated with PD173074 (1 µM) and DMSO for 24, 48, 72 and 96 hours, respectively. Breakdown of mitochondrial potential was examined by comparison of JC-1 stains flow cytometrically.

Figure S13 IHC-*Scoring of FGFR phosphorylation and MYC expression exhibits enriched FGFR phosphorylation and variant MYC expression*

a) Occurrence of FGFR phosphorylation and high MYC expression in $FGFR1^{amp}$ and $FGFR^{non\text{-}amp}$ patient cohorts, respectively. Inter-dependency of MYC expression, FGFR1 amplification and FGFR phosphorylation was assessed by Fisher's Exact Test. **b)** Phospho-FGFR (top) and MYC (bottom) IHC stains were scored from 0 to 3. A representative sample is shown for each score.

Figure S14 *Focality of the 8p12 amplicon as assessed by Deep Cap Analysis Gene*

a) Expression (CAGE) Sequencing. DNA was extracted from the formalin-fixed tumor sample and cloned into a CAGE sequencing library. Genes of the 8p12 amplicon were enriched in the CAGE chip design. Copy number was inferred from gene coverage and mapped to genomic positions of the Hg18 annotation. **b)** Pathological examination of a tumor biopsy of the pazopanib responder before therapy. After diagnosis SQLC histology (top left), the sample was scored (degrees 0-3) by FGFR1 FISH (top left middle), phospho-FGFR1 IHC (top right middle) as well as nuclear staining of MYC IHC (top right). Dual color FISH was

performed with FGFR1 (green) and CEN8 (red, centromere) probes in order to derive a normalized copy number ratio for FGFR1 amplification. Baseline computer tomographic (CT) scan with tumor in the left lung (bottom left); CT after 4 weeks (bottom middle) and 8 weeks (bottom right) of pazopanib, showing tumor regression with cavitation. Target lesions for evaluation of tumor response are highlighted by red arrows.