Supplementary information for:

EGFR kinase domain duplication (EGFR-KDD) is a novel oncogenic driver in lung cancer that is clinically responsive to afatinib

Jean-Nicolas Gallant[1,8], Jonathan H. Sheehan[2,7], Timothy M. Shaver[2,8], Mark Bailey[3], Doron Lipson[3], Raghu Chandramohan[4], Monica Red Brewer[5,8], Sally J. York[5,8], Mark G. Kris[9], Jennifer A. Pietenpol[2,8], Marc Ladanyi[4], Vincent A. Miller[3], Siraj M. Ali[3], Jens Meiler[6,7], and Christine M. Lovly[1,5,8,10]

Departments of [1]Cancer Biology, [2]Biochemistry, [5]Medicine, and [6]Chemistry, [7]Center for Structural Biology, and [8]Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN 37232, USA; [3]Foundation Medicine Inc., Cambridge, MA 02141, USA; and the Departments of [4]Pathology and [9]Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA.

[10]Corresponding author:
Vanderbilt-Ingram Cancer Center
2220 Pierce Avenue
777 Preston Research Building
Nashville, TN 37232-6307
Phone 615-936-3457
Fax: 615-343-7602
E-mail: christine.lovly@vanderbilt.edu

Supplementary Methods

Validation of *EGFR*-KDD

Nucleic acids were isolated from formalin-fixed, paraffin-embedded tissue sections using the AllPrep DNA/RNA FFPE Kit (Qiagen). Nucleic acids were isolated from A1235 cell pellets using RNAeasy and DNAeasy kits (Qiagen). RT-/PCR were used to amplify the 'breakpoint' region following exon 25 and preceding exon 18 of *EGFR*-KDD. RT-PCR was performed with the SuperScript III one-step RT-PCR system (Life Technologies) using 100 ng RNA and the following primers:

Ex24-25F: 5'–GTTTTCCCAGTCACGACCTTGTCATTCAGGGGGATG–3'

Ex18-19R: 5'–CAGGAAACAGCTATGACGGGATCCAGAGTCCCTTATACA–3'

PCR was performed with HotStarTaq master mix (Qiagen) using 100 ng DNA and the following primers:

Ex25F: 5'–GTTTTCCCAGTCACGACAGAACCAAGGGGGATTTCAT–3'

Ex18R: 5'–CAGGAAACAGCTATGACTTGGTCTCACAGGACCACTG–3'

RT-/PCR products, which included the *EGFR*-KDD 'breakpoint', were sequenced in both directions using M13 primers, BigDye Terminator chemistry, and 3730 DNA Analyzers (Applied Biosystems). Aside from the *EGFR*-KDD, no mutations or alterations were found in either the sequenced regions of the A1235 cell line or the patient sample (as compared to the *EGFR* CCDS).

Supplementary Figure Legends

Supplementary Figure 1: Sequencing reads of *EGFR-KDD* in index patient with lung adenocarcinoma. (**a**) NGS reads surrounding the genomic breakpoint within the *EGFR*-KDD as detected from the FoundationOne™ assay. (**b**) Paired NGS reads of tumor and matched whole blood from MSK-IMPACT™.

Supplementary Figure 2: cDNA sequence of *EGFR-KDD*.
Sequence of the *EGFR* exons 18-25 kinase domain duplication (*EGFR*-KDD). The first repeat of *EGFR* exons 18–25 is capitalized and highlighted with blue. The second (tandem) repeat of *EGFR* exons 18–25 is capitalized and highlighted with green.

Supplementary Figure 3: Sequencing reads identifying *EGFR-KDD* in a lung adenocarcinoma tumor from The Cancer Genome Atlas. Next-generation sequencing reads from RNA-seq (**a-b**) and whole-genome sequencing (**c-d**) were visualized to confirm the presence and expression of the *EGFR*-KDD in lung adenocarcinoma sample TCGA-49-4512. Paired reads with a pattern of tandem duplication are pictured in green, and breakpoint-spanning (soft-clipped) reads are pictured as colored base pair mismatches. The boundaries of the *EGFR*-KDD are visible at the RNA level in *EGFR* exon 18 (**a**) and exon 25 (**b**) with corresponding intronic DNA breakpoints between exons 17 and 18 (**c**) and exons 25 and 26 (**d**). The data were visualized using the Integrative Genomics Viewer[1].

Supplementary Figure 4: Autophosphorylation of endogenous EGFR-KDD in A1235 cells.
II-18 (EGFR-L858R) and A1235 (EGFR-KDD) cell lines were grown in the presence or absence of serum (FBS), or treated with 1 or 50 ng/mL EGF for 5 minutes after 16 hours of serum starvation, and then lysed for western blot analysis with the indicated antibodies.

Supplementary Figure 5: Colony formation of NR6 cells expressing EGFR variants.

Representative 5x microscopy photographs of NR6 cells transduced with (**a**) pMSCV vector, (**b**) EGFR-WT, (**c**) EGFR-L858R, or (**d**) EGFR-KDD after 15 days of growth in soft agar.

Supplementary Figure 6: Efficacy of EGFR TKIs in endogenous and ectopic models of the EGFR-KDD.

(**a**) 293T cells transfected with EGFR-KDD were treated with increasing doses of erlotinib or afatinib for 2 hours and lysed for western blot analysis. (**b**) A1235 cells, which harbor endogenous EGFR-KDD, were treated with increasing doses of erlotinib, afatinib, or AZD9291 for 2 hours and lysed for western blot analysis. (**c**) A1235 cells were treated with increasing doses of erlotinib, afatinib, or AZD9291 for 72 hours. Cell titer blue assays were performed to assess cell viability. Each point represents quadruple replicates. Data are presented as the percentage of viable cells compared to vehicle control.

Supplementary Table Legends

Supplementary Table 1: Results of MTT curve fitting from Prism.

These data correspond to the graphs shown in Figure 2b.

Supplementary References

1. Robinson, J. T. *et al.* Integrative genomics viewer. *Nature Biotechnology* **29,** 24–26 (2011).