Focal amplification of existing kinase inhibitor targets identified from TCGA data and integrative analysis

Matthew J. Anstett, Paul D. Williams, Peter Wyngaard, and Daniel R. Rhodes

Compendia Bioscience, Ann Arbor, MI.

Abstract

The identification of oncogenic protein kinases in cancer has prompted the development of several novel targeted therapies. In many cases, such therapies have elicited dramatic clinical responses in patients harboring genetic activation of the target kinase. To identify additional opportunities to apply existing targeted therapies, we undertook a systematic DNA copy number analysis of 22 targeted kinases across 7,000 cancer patients from The Cancer Genome Atlas (TCGA) and Oncomine™. Importantly, TCGA studies included normal reference samples, paired gene expression data and in some cases, paired mutation data. To investigate the possibility that the identified target amplifications represent driver amplifications and might thereby confer sensitivity to the appropriate targeted therapy, we performed integrative analysis. First, minimal common region analysis of the genomic regions containing the amplified targets was assessed. Second, DNA amplification vs. mRNA expression correlation analysis was performed to confirm that amplification resulted in significant over-expression. In total, we identified 1073 significant DNA amplifications (>4 copies) of targeted kinases in 862 cancer patients, suggesting that 10-15% of cancer patients harbor a detectable DNA amplification that might indicate benefit from an existing targeted therapy. All 22 targets showed evidence for significant amplification at least once. Our analysis demonstrated that existing kinase inhibitor targets are focally amplified and may be oncogenic drivers in patient sub-populations spanning multiple cancer types and that systematic analysis of TCGA and Oncomine provide the opportunity to ask clinically important questions of large genomic datasets.

Material and Methods

Data Collection & Processing: Copy number variation (CNV) data was analyzed from 50 patient cohorts encompassing 19 general cancer types and nearly 7000 clinical specimens. Data were available in the Oncomine database and were originally gathered from public repositories including GEO and the TCGA portal. Sample data were carefully curated and experimental data were pre-processed either by study authors or using AROMA. Copy number values were median-centered per probe. Circular binary segmentation was performed using the Bioconductor DNACopy package (v1.18). To estimate gene copy number, the resulting segments were mapped to hg18 (NCBI 36.1) RefSeq coordinates (UCSC refGene) as provided by UCSC (UCSC refGene, July 2009, hg18, NCBI 36.1, March 2006).

Analysis: For each of the 22 kinases evaluated and within each cancer type, amplification frequencies were determined. Genes were deemed significantly amplified if the log2 copy number value exceeded 1.0 (i.e., >4 copies). Amplification frequencies were summarized for each gene and each cancer type. To evaluate the likely importance of gene amplification in a given cancer type, minimal common region analysis and copy number vs. expression analysis were performed. The Oncomine DNA Copy Browser was used to assess the genomic regions of amplification. Paired DNA copy number and mRNA expression data were queried from Oncomine and linked on TCGA patient ID.

Conclusions

- 50 genomic datasets
- 7,000 patient tumor specimens
- 22 targets evaluated
- 1073 target amplifications in 862 patients
- 10-15% patients harbor amplifications of kinase targets

Results

Figure 1. Tumors analyzed by cancer type and data source.

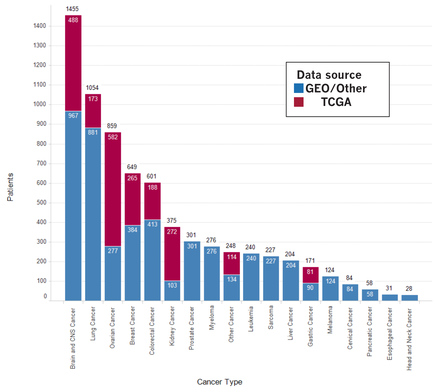

Figure 2. Analysis schematic.

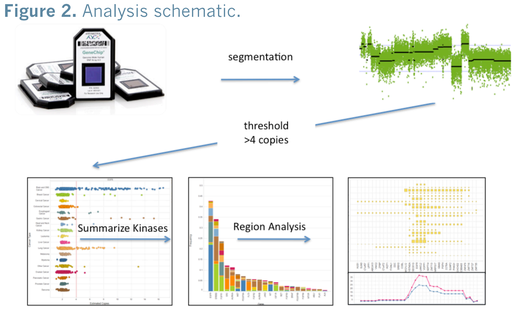

Figure 3. Cumulative amplification frequencies summarized by cancer type (3A) and gene (3B).

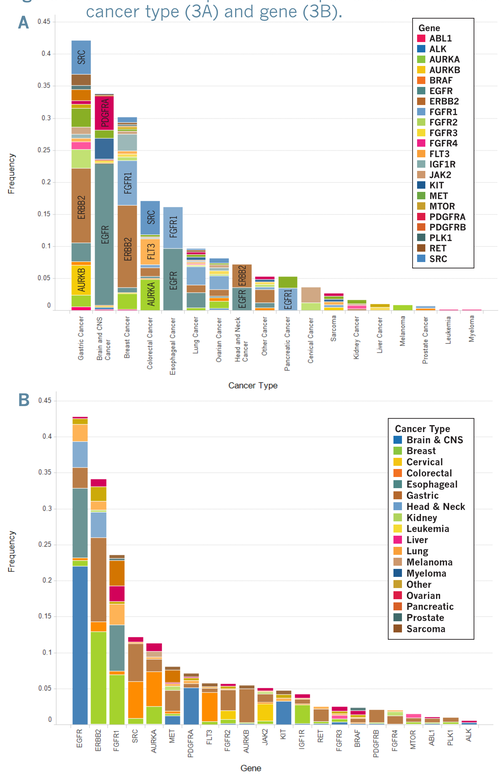

Figure 4. EGFR amplifications in cancer.

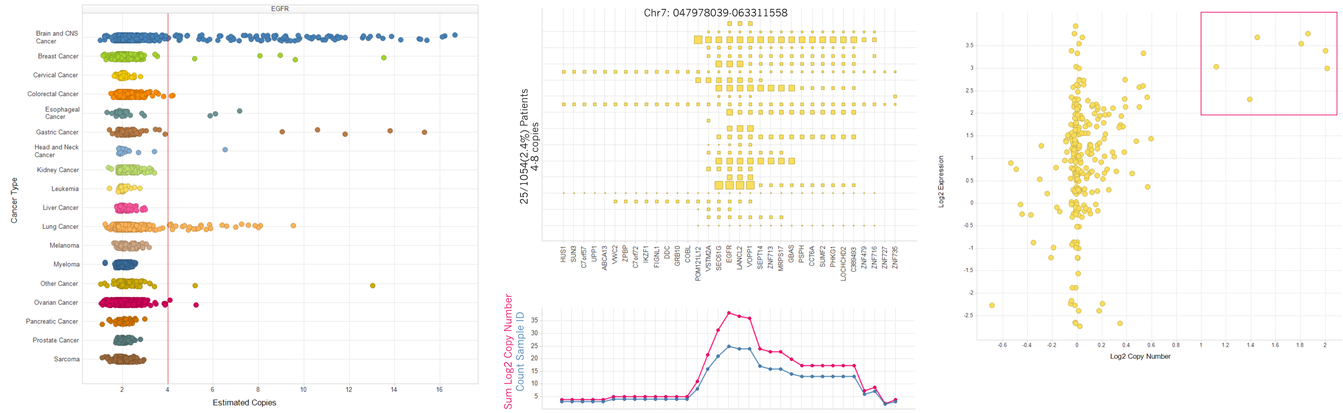

Figure 5. FGFR1 amplifications in cancer.

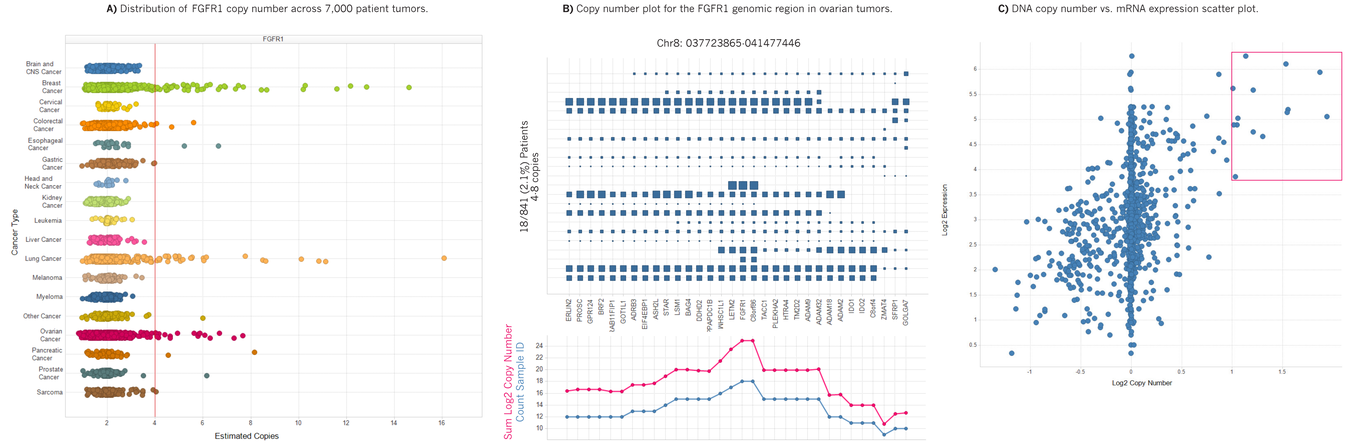

Figure 6. IGF1R amplifications in cancer.