Supplementary Note

Manuscript by Roman Thomas et al.: “High-throughput oncogene mutation profiling in human cancer”

Supplementary Methods

Validation

We attempted validation for each mutation detected through our high-throughput mutation profiling approach. Mutations were independently validated by conventional Sanger sequencing, by bacterial subcloning followed by Sanger sequencing, by picotiter plate pyrosequencing[1], by independent genotyping of the same sample, by allele-specific PCR or by database search (e.g., frequently used cell lines that have been heavily sequenced). Allele-specific PCR for the BRAF(V600E) mutation was performed in primary melanoma genomic DNA using the 2-DDCt method[2], as described previously.[3] Comparison was made to a LINE-1 reference and normalized to normal human DNA using primers described previously.[3]

Sensitivity was estimated by comparing the results obtained by high-throughput mutation profiling to results obtained by Sanger sequencing and, where available, pyrosequencing-by-synthesis[1] of two DNA sample plates. One of these plates, containing Japanese non-small-cell lung cancer specimens, has previously been sequenced against all coding exons of EGFR[4]; some of the samples have also been analyzed using picotiter plate pyrosequencing.[1] The other plate contained DNA from the NCI60 panel of cancer cell lines plus additional cell lines and has been sequenced against PIK3CA, KRAS, NRAS and BRAF.[5] Of the 39 mutations detected in the

samples from those two DNA sample plates, high-throughput mutation profiling correctly identified 39, yielding a false negative rate of 0%.

We attempted validation of all 437 mutation instances, defined by all samples where high-throughput mutation profiling had called at least one mutation (e.g., one sample can be represented on more than one plate). Of 393 mutation instances that we were able to validate (90%), 349 mutations identified by high-throughput mutation profiling were confirmed, suggesting a specificity of 89%. However, since Sanger sequencing itself is insensitive when applied to stromally admixed tumor samples, we also compared our approach to a highly sensitive pyrosequencing method[1] for the detection of EGFR mutations in lung tumor DNA. Genotyping detected 12/12 mutations identified by pyrosequencing-by-synthesis, including three mutant alleles representing 16%, 12%, and 9% of total alleles, that had been missed by the Sanger method but were detected by single molecule sequencing.[1] Additionally, all of the V600E mutations of BRAF in primary melanoma detected in primary melanoma biopsy specimens were independently confirmed using allele-specific real-time PCR. Out of 15 mutations in that set that we were unable to validate by Sanger sequencing, allele-specific PCR confirmed 13. In all of these was the tumor estimated to constitute only a minute fraction of the biopsy specimen (data not shown). In summary, the accuracy of high-throughput oncogene mutation profiling may exceed that of Sanger sequencing.

Importantly, allele mixing experiments have been performed in the past to test the applicability of mass-spectrometric genotyping for association testing in pooled DNA samples.[6,7] These experiments conclusively revealed a robust identification of underrepresented alleles down to frequencies of 10% and below.[6,7] These results are in line with experiments from our laboratory, where the detection threshold was determined by comparing calls made by mass-spectrometric genotyping to those quantified using picotiter plate pyrosequencing and Sanger sequencing (see above and Winckler and

Meyerson, unpublished results). Also, own allele mixing experiments showed the robust identification of underrepresented alleles to a percentage of 10% and less (data not shown).

References

1. Thomas, R.K. et al. Sensitive mutation detection in heterogeneous cancer specimens by massively parallel picoliter reactor sequencing. *Nat Med* **12**, 852-855 (2006).
2. Livak, K.J. & Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-8 (2001).
3. Garraway, L.A. et al. Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* **436**, 117-22 (2005).
4. Paez, J.G. et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304**, 1497-500 (2004).
5. Solit, D.B. et al. BRAF mutation predicts sensitivity to MEK inhibition. *Nature* **439**, 358-62 (2006).
6. Bansal, A. et al. Association testing by DNA pooling: an effective initial screen. *Proc Natl Acad Sci U S A* **99**, 16871-4 (2002).
7. Werner, M. et al. Large-scale determination of SNP allele frequencies in DNA pools using MALDI-TOF mass spectrometry. *Hum Mutat* **20**, 57-64 (2002).