

Tissue Microarrays

Michael D. Radmacher
Richard Simon

mdradmac@helix.nih.gov
Rsimon@nih.gov

Statistical Methods

For specimens on tissue microarrays, the average nuclear intensity (ANI) was measured as an ordinal variable, taking integer values between 0 (weakest staining) and 3 (strongest). Each tumor was represented on the array by two cores; the agreement of ANI scoring for these was high, as measured by a weighted kappa statistic. The Wilcoxon-Mann-Whitney test was used to test for differences between *BRCA1* mutation positive and *BRCA2* mutation positive tissues (for these tests, we used the mean ANI score of the two cores from a tumor instead of including both values in the analysis). Reported p-values are two-sided and exact.

Results

Tests of Differential Expression Between *BRCA1*s and *BRCA2*s

| Gene | cDNA Array ¹ | Tissue Array ² |
|------------------|-------------------------|---------------------------|
| <i>cyclin D1</i> | 0.0004 | 0.0001 |
| <i>MEK-1</i> | 0.054 | 0.23 |
| <i>ERBB-2</i> | 0.0027 | 0.31 |
| <i>keratin 8</i> | 0.0001 | 0.18 |

¹ p-value from two-sample t-test

² p-value from Wilcoxon-Mann-Whitney test

Discussion

We used a high-density tissue microarray to validate the gene expression findings within a larger number of tumors. As expected, IHC results from tissue microarrays did not absolutely correlate with every gene identified by cDNA microarrays as separating sample sets. Explanations for the concordance between tissue microarray and cDNA microarray data for some genes (e.g. *cyclin D1* and *MEK-1*) but not for others (e.g., *ERBB-2*) are most likely due to deficiencies of the IHC staining reaction. First, with the

highly variable and dynamic nature of gene expression and protein translation as well as varying half-lives of individual mRNA and protein molecules, a one-to-one relationship does not always exist between every transcript and its protein levels. Second, the nature of the IHC reaction is usually semi-quantitative and has a poor dynamic range, whereas cDNA microarrays produce linear information on gene expression. This was clearly evident in the comparison of *ERBB-2* data between IHC and cDNA microarrays where IHC staining for *ERBB-2* was titrated to be positive only in the cases with very high-level overexpression (confirmed by FISH to be caused by gene amplification). In contrast, cDNA microarrays detected measurable *ERBB-2* expression in virtually all cases. Third, differences in tissue fixation may cause non-linear staining and deficient staining reactions by IHC. Finally, the expression of thousands of genes were measured by the cDNA microarray experiments. Because of the “regression-to-the-mean” phenomenon exacerbated by the large number of genes being examined, reliably identifying the specific genes which account for the difference in expression patterns among the mutation groups is more difficult than reliably establishing that the patterns differ. This in itself may account for why a few specific genes that are found on the consensus gene list from the cDNA array data do not appear to discriminate on the tissue array. Accordingly, we recognize that our result attempting to directly correlate cDNA and tissue microarrays should be interpreted with caution, as we have studied only a small tumor sample set, a limited number of antibodies, and we have used an institution-matched tissue microarray.