Hybridization Efficiency Analysis Of Probes Targeting 16S rRNA Genes Using The Affymetrix Genechip Format.

T. Z. DeSantis[1], K. D. Hansen[2], E. L. Brodie[1], Y. M. Piceno[1], J. Bullard[2], P. Hu[1], **G. L. Andersen**[1];
[1]Lawrence Berkeley National Laboratory, Berkeley, CA,
[2]University of California, Berkeley, Berkeley, CA.

Background: Detection of diverse 16S rRNA gene types in complex mixtures can be achieved using arrays of probes targeting specific sequences in 16S rRNA genes. Whereas probes for expression arrays are designed to leverage the diversity among various genes in one genome, 16S probes rely upon the diversity of the same gene found in many genomes. Also, expression arrays are validated by their accurate estimation of changes in analyte concentration, but 16S arrays are expected to provide definitive present-absent scoring of each prokaryotic taxa. The degree of uniqueness of a probe for a particular target species or other defined operational taxonomic unit will dictate its reliability but has yet to be quantified for prediction of hybridization accuracy. **Methods:** To obtain these metrics, amplicons of the16S rRNA gene from *Francisella tularensis* were fragmented, labeled and isothermally hybridized to replicate Affymetrix custom arrays containing 491,069 unique 25mer probes with various degrees of probe-target complimentarity, melting temperature, and secondary structure potential. Hybrid abundance at each probe location was determined by florescence intensity. **Results:** As expected, probes exactly complimentary to the target but with various sequence composition produced intensities ranging over 3 orders of magnitude yet replicate probes on the same array produced a coefficient of variation under 20%. Although mismatching probes were able to capture target sequence, a general decrease in intensity was observed with probes divergent from the target and the effect could be attenuated by masking probes with high melting temperatures. **Conclusion:** The data collected allows the development of a probabilistic model that aids in predicting the confidence that a probe's response is due to the presence of the corresponding target in solution.