

# ***Organism Identification Using a Genome Sequence-Independent Universal Microarray Probe Set***

## ***Experimental Results Description***

### **Oligos Sequence**

The oligos sequence is extracted from the Chip Probe file by the DOCUMENTOR software and each experiment using the same type of chip has the same order of oligos sequences and can be compared.

### **Experiment Name and Oligos Number**

Each experiment is named for the individual chip used and the sample hybridized to it. For example, 6-24B Sub is the experiment name for the sample *Bacillus subtilis* hybridized on chip 6-24B. A list of experiment names and their corresponding sample types are below.

The oligos number is an arbitrary name given to each specific sequence feature on the chip to distinguish them from the others. OligoPM, OligoSS and OligoRS are automatically determined by the chip design software as a internal control of chip quality. OligoPM is a perfect match to the control sample. OligoSS is a control for a Scrambled-Sequence oligo, and OligoRS is a control for a ReSequence oligo.

### **Rejected Mean**

This is the average of the pixel intensity of each feature on the chip, with the individual outliers two standard deviations from the mean of the feature rejected. This statistical normalization is done in the DOCUMENTOR program. Another normalization is performed in the GeneSpring evaluation.

## Data analysis

These results were imported into GeneSpring for analysis. The internal replicates were averaged and the data with values less than 2 standard deviations above no-hybridization controls were set to an arbitrary value, 0.01. The data was normalized by standard means, with the measurements normalized to the median of all the measurements in the sample. Then each feature is normalized to the median of its measurements in all the samples. The Global Error Model was used to calculate a threshold. Probes that did not meet the threshold were excluded from further analysis. The replicate microarrays were then compared and probes that showed poor reproducibility between replicates were filtered and excluded from further analysis. The filtered data is clustered first by feature and then by sample using the standard correlation method.

## Experiment Names:

6-24B Sub	slide 6-24B <i>Bacillus subtilis</i> (ATCC # 6633)
6-24H Sub	slide 6-24H <i>Bacillus subtilis</i> (ATCC # 6633)
4-29O Sub	slide 4-29O <i>Bacillus subtilis</i> (ATCC # 6633)
4-29F Strep	slide 4-29F <i>Streptococcus pneumoniae</i> (ATCC # BAA-334D)
4-29K Strep	slide 4-29K <i>Streptococcus pneumoniae</i> (ATCC # BAA-334D)
6-24O Human	slide 6-24O <i>Homo sapiens</i> (Boehringer Mannheim # 1691112).
6-24O BA	slide 6-24O <i>Bacillus anthracis</i>
6-17H BA	slide 6-17H <i>Bacillus anthracis</i>
7-1S YP	slide 7-1S <i>Yersinia pestis</i>
6-11I YP	slide 6-11I <i>Yersinia pestis</i>