Operon Biotechnologies, Inc.
2211 Seminole Drive
Huntsville, AL 35805
www.operon.com

Clostridium acetobutylicum Genome Array Ready Oligo Set™ (Version 1.0)

Clostridium acetobutylicum V1.0 AROS is designed using state-of-the-art methodology and proprietary software for detection of alternative splice variants from *C. acetobutylicum*, an organism of high interest for those researching alternative fuels. The *Clostridium acetobutylicum* genome was sequenced by the former Genome Therapeutics Corporation in 2001. Current interest in biofuel research has increased the need for additional tools to support research in this organism.

The annotation of the *C. acetobutylicum* genome predicts 3844 genes which include 3672 protein coding genes and 107 structural RNAs. Of those, 3842 are covered in this AROS. This AROS set also features 24 positive controls and 140 positive stringency controls, 10 SpotReport® Alien® control oligos and 60 corresponding Alien stringency controls and 10 negative controls.

Probe design targets each oligo to be fully contained within each gene and makes this probe set also applicable for CGH (comparative genome hybridization) experiments. The *C. acetobutylicum* AROS V1.0 contains 4231 probes including 70 Stratagene SpotReport™ Alien™ control oligos, 5 Negative controls, 42 Positive and Positive Stringency control oligos. Additionally, 43 Production control oligos used by Operon for internal checks and also useful to researchers for plate layout checks are included free of charge. These AROS control probes supplement with the set to facilitate experimental validation and quality control.

Release date: August, 2007

I. The sequence sources

Clostridium acetobutylicum genome sequence and its associated annotations are from http://www.ncbi.nlm.nih.gov/sites/entrez?db=genome&cmd=Retrieve&dopt=Overview&list_uids=185. More information on the underlying data is available at http://cmr.tigr.org/tigr-scripts/CMR/GenomePage.cgi?org=ntca01

II. Design criteria and selection rules:

The following selection rules were executed in the design.

1) Target length of oligo probe sequences was 70 bases. Exceptions for genes and regions without qualified oligo candidates at this length are routinely made in AROS probe sets in order to maintain tight control of hybridization characteristics to allow for highly stringent experimental conditions. These exceptions are based on the inherent characteristics of the *C. acetobutylicum* genome and not a limitation of the design algorithm.

2) The target T_m was taken to be 60 °C to allow oligos with higher T_m to be shortened to minimize the T_m range 60 ± 5 °C. T_m was calculated using a 'nearest-neighbor' method (Novère, 2001).

3) Probe candidates were processed to shorten those with T_m > 80 although none was allowed to be shorter than 60 bps. The resulting oligos were processed to remove probes representing gene copies and then to replace or concatenate records with cross-hybridization greater than 85% or with more than 20 contiguous bases in common (He *et al*, 2005) as described below.

4) Probe candidates were screened against the low complexity of sequences: (a) The contiguous single nucleotide base repeat or poly (N) tract (≤8 bases); (b) the pre-set normalized simple repeat score.

5) Probe candidates were selected against the potential hairpin structure with stem length of over 9 bases.

6) Probe candidates were screened against the cross-hybridization of the non-self transcripts. The non-self transcript is the transcript not represented by a specific oligo. The selection criteria were used in the selection: (a) the cross-hybridization identity score ≤70%; (b) contiguous match length to any non-self transcript ≤20 bases.

7) Limited exceptions were made if no qualified oligos were available from the selection. Information covering these exceptions is detailed in the *C. acetobutylicum* AROS V1.0 genelist.

8) Probe design targets each oligo to be fully contained within a single exon and makes this probe set also applicable for CGH (comparative genome hybridization) experiments. The probes contained within a single exon are detailed in the

III. Annotation notes:

Oligos were identified by using NCBI BLAST to map each oligo to the sequences in the databases described above. Identifiers are according to the BLAST score (score >97% identity).

The oligos are classified according what genes they represent as shown in Table 1. The design process attempts to find an oligo for every unique combination of transcripts for each gene.

IV. Characteristics of *Clostridium acetobutylicum* AROS V1.0 Control Probes

1) Control Oligos:
    - Negative Control Oligo Probes: Randomly-generated oligo sequences which were selected after filtering against the transcripts.

    - Positive Control Oligo Probes: 24 oligo probes were selected as positive controls.

    - The Positive and Negative controls and the corresponding Hybridization Stringency Control Oligos (see below) are randomly scattered throughout the AROS set to be used to check a variety of potential issues including wash stringency and variation of hybridization stringency in different areas of the array.

    - Hybridization Stringency Control Oligos: Generated based on several positive control oligo sequences with sequence homology of 50% 60% 70%, 80%, 90% and antisense identity to the original oligo probe sequences. They're intended as the standards to calibrate the microarray hybridization conditions and stringency.

    - Stratagene SpotReport™ Alien™ Spike Control Oligos: Licensed from Stratagene (www.stratagene.com). Coupled with the Alien mRNA spikes in Stratagene SpotReport Alien Oligo Array Validation System, they are intended as the internal controls for the normalization and standardization of dye incorporation, microarray hybridization and data analyses.

    - Production Tracking Oligo opHsV04NC000001: It is a randomly-generated oligo sequence with a length of 30 bases. The tracking oligo is randomly positioned in the 384-well and 96-well plates so that each 384-well plate has four, and each 96-well plate has one for the purpose of quality control in production. The CyDye labelled antisense probe for this oligo can be used in a QC assay to ensure the correct placement/order of the probe printing plates on the spotting robot.

References

He, Z., Wu, L., Li, X., Fields, M.W., Zhou, J. (2005). "Empirical Establishment of Oligonucleotide Probe Design Criteria", *Applied and Environmental Microbiology*, **71**:3753-3760.

Novère, N. (2001). "MELTING Computing the melting temperature of a nucleic acid duplex", *Bioinformatics*, 17:1226-1227.