



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

## **Operon *Streptococcus mitis* Genome Array Ready Oligo Set™ (Version 1.0)**

---

*Streptococcus mitis* Genome AROS (version 1.0) was designed based on *S. mitis* strain B6 genome, which was assembled and annotated by the scientists at the Laboratory of Professor Regine Hakenbeck. *S. mitis* AROS, consisting of 2,489 oligonucleotide probes, represent all gene sequences as well as a selected group of intergenic sequences from the B6 genome. In addition, *S. mitis* AROS is supplemented with a set of positive controls, randomly-generated negative controls, stringency controls, alien spike controls and tracking control for experimental validation and quality control.

Release date: April, 2006

### **I. The sequence sources**

*S. mitis* genome sequence and its associated annotations are from the Laboratory of Professor Regine Hakenbeck, Department of Microbiology and Nanobio-Center, University of Kaiserslautern, Germany, 67663. The data will be later submitted to the data repositories (GenBank and EMBL) for public access. Additional sequences used in the AROS design are accessible at the National Center for Biotechnology Information ([www.ncbi.nih.gov](http://www.ncbi.nih.gov)).

### **II. The design criteria and selection rules:**

The following selection rules were executed in the design.

- 1) The intergenic regions at or over 200 bases were included for the oligo design.
- 2) The length of oligo sequences was set at 70 bases with the exceptions for genes and regions without qualified oligo candidates.
- 3) The melting temperatures ( $T_m$ ) of  $74 \pm 5^\circ\text{C}$  was used for oligo screening.  $T_m$  was calculated with the formula:  $T_m = 81.5 + 16.6 * \log[\text{Na}^+] + 41 * \text{GC}\% - 500/\text{length}$ , where  $[\text{Na}^+] = 0.1 \text{ M}$  and  $\text{length} = \#A + \#C + \#G + \#T$ .
- 4) The oligo candidates were pre-selected with 40 bases away from both 5'- and 3'-end of

transcripts as well as a maximal distance of less than 1000 bases from the 3'-end.

- 5) The oligo candidates were screened against the low complexity of sequences: (a) The contiguous single nucleotide base repeat or poly (N) tract ( $\leq 8$  bases); (b) the pre-set normalized simple repeat score.
- 6) The oligo candidates were selected against the potential hairpin structure with stem length of over 9 bases.
- 7) The oligo candidates were screened against the cross-hybridization of the non-self transcripts as well as the selected intergenic sequences. 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  $\leq 70\%$ ; (b) contiguous match length to any non-self transcript  $\leq 20$  bases.
- 8) Exceptions were made if no qualified oligos were available from the selection.

### **III. The specifications and characteristics of *Streptococcus mitis* AROS (version 1.0)**

The AROS encompasses both the gene sequences and the intergenic sequences from the genomes of *S. mitis* strain B6. The AROS are divided into four components.

- 1) Gene-specific oligos: 2,020 oligos cover all gene-specific sequences.
- 2) Intergenic oligos: 464 oligos represent the selected intergenic regions of B6 genome.
- 3) Misc. oligos: 5 oligos represent rup1, rup2, rup3, box AB and box BC elements.
- 4) Control oligos:
  - Negative control oligos: They're randomly-generated oligo sequences, and were selected after filtering against the transcripts and intergenic sequences for non cross-hybridization property.
  - Eukaryotic control oligos: Four oligos (H300006234, H300013722, H300022861, opHsV04TC000032) from human AROS (version 4.0) are included in this AROS as eukaryotic controls for human ACTB, STAT1, ISGF3G and GAPDH genes.
  - Positive control oligos: Ten oligos were selected from this AROS as the positive controls. They represent the following genes: Shikimate 5-dehydrogenase, Glucose-6-phosphate 1-dehydrogenase, RNA polymerase, delta subunit (RNAP delta factor), Glucose kinase, 30S ribosomal protein S3, signal peptidase I, Xanthine phosphoribosyltransferase, D-ala ligase, Cell division protein FtsZ, glutamate dehydrogenase, acetyl-CoA carboxylase beta subunit.
  - Stratagene alien spike control oligos: They are licensed from Stratagene ([www.stratagene.com](http://www.stratagene.com)). In coupling 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.

- Hybridization stringency control oligos: They were generated based on the four oligo sequences with sequence homology of 70%, 80%, 90% identity to the original oligo sequences. They're intended as the standards to calibrate the microarray hybridization conditions and stringency. The four oligos represent the following genes: acetyl-CoA carboxylase beta subunit, 50S ribosomal protein L20, Topoisomerase IV subunit A, UDP-glucose 4-epimerase. In addition, the same type of stringency controls (70%, 80%, 90% identity) were generated for Stratagene SpotReport Alien oligo #1, #2, #3 and #4.
- Production tracking oligo: It is a randomly-generated oligo sequence with a length of 30 bases. It was filtered against the cross-hybridizations with the gene sequences used in the AROS design. 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.

Figure 1. Tm

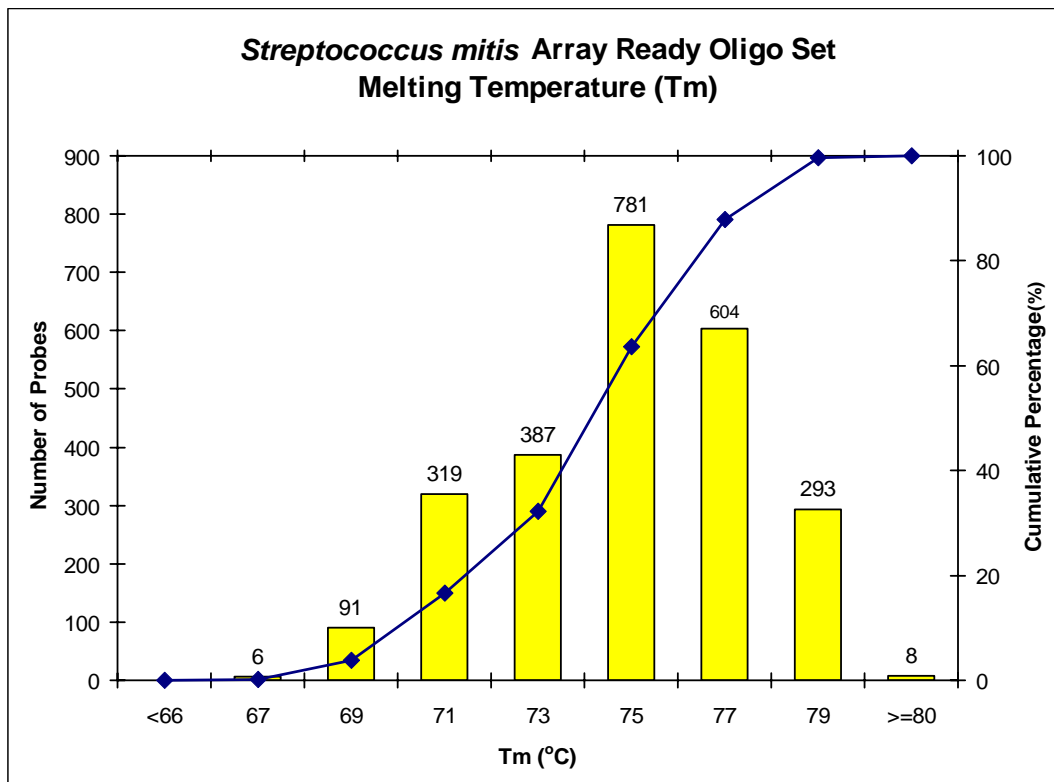


Figure 2. GC content

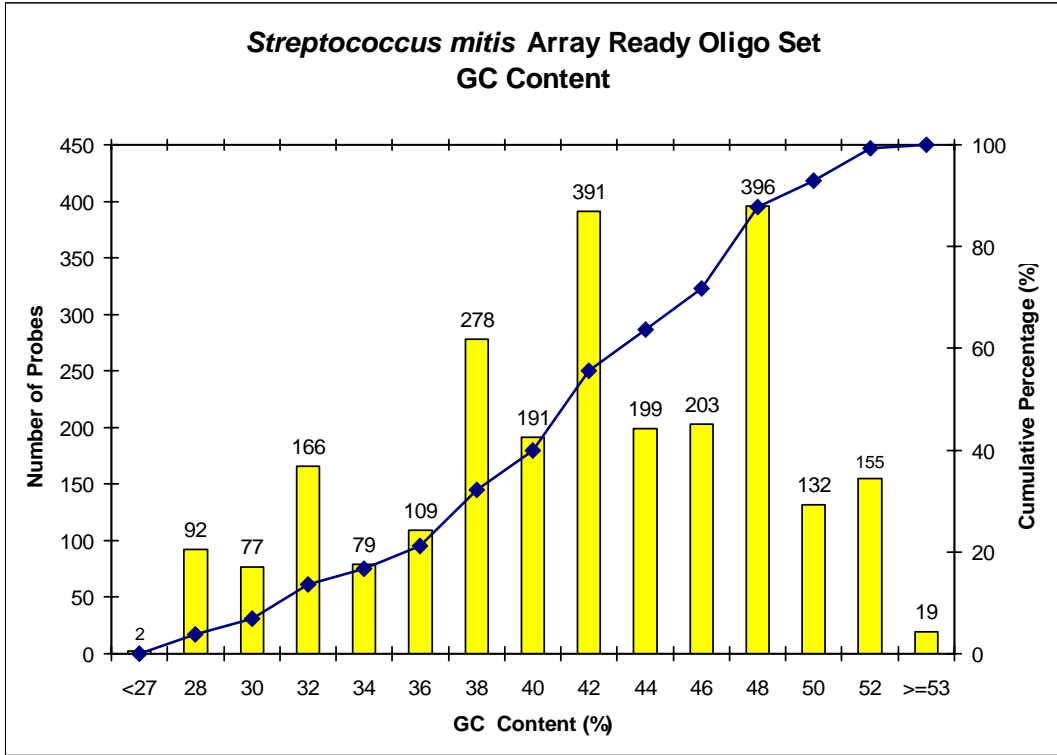


Figure 3. Distance to the 3'-end of transcripts

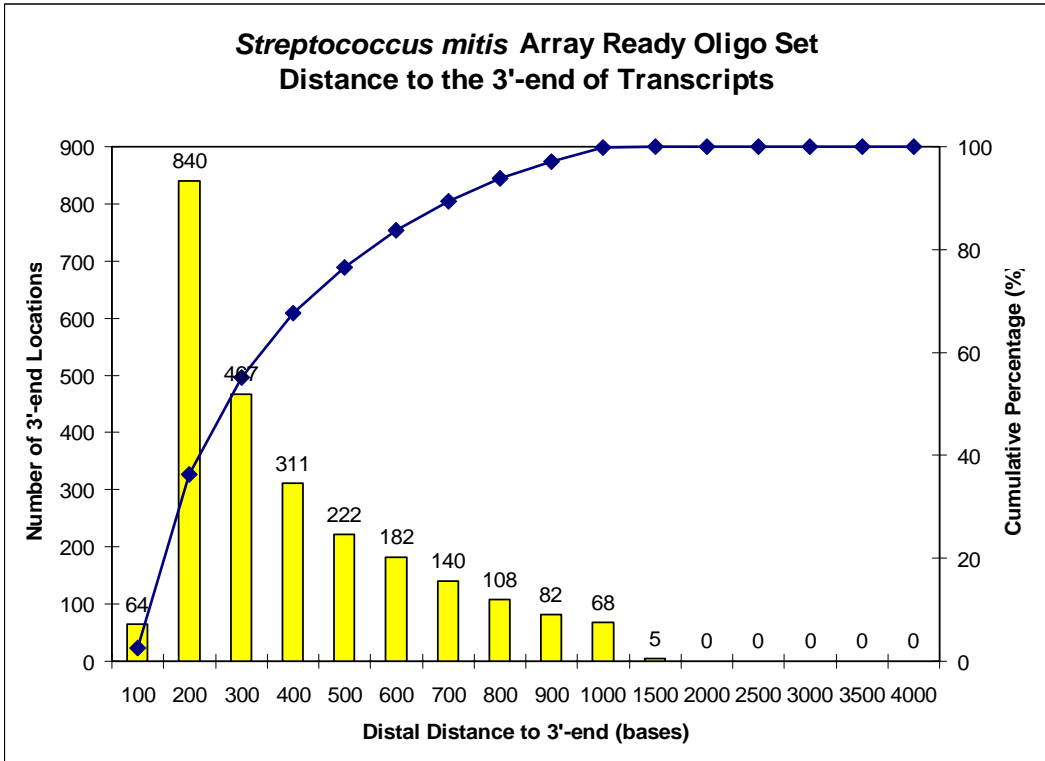


Figure 4. Hairpin stem length

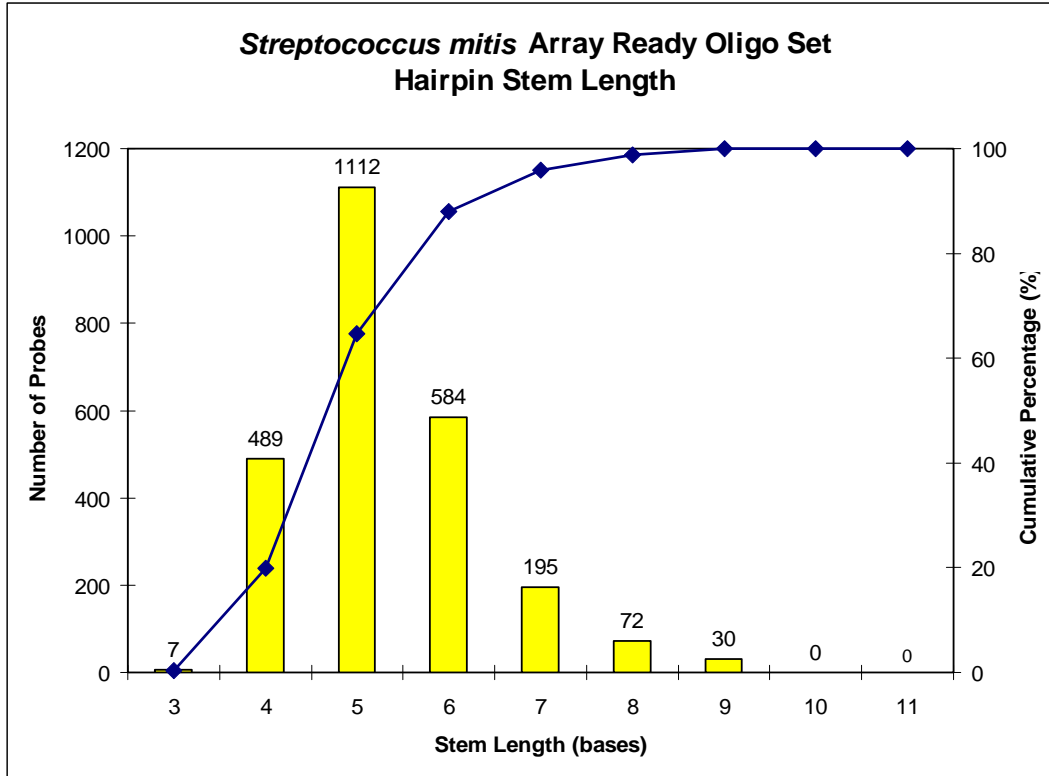


Figure 5. Cross-hybridization to non-self transcripts

