

Array-Ready Oligo Set™ for the *Salmonella* Genome Version 1.0

We are pleased to announce the release of our *Salmonella* Genome Array-Ready Oligo Set™ (AROS) containing 5578 70mer probes representing 4451 Open Reading Frames (ORFs) from *Salmonella typhimurium* LT2, 2510 ORFs from *Salmonella typhi*, and 102 ORFs from plasmid pSLT, 2 ORFs from plasmid pSC101, 228 ORFs from pHCM1, 207 ORFs from plasmid R27, and 131 ORFs from plasmid pHCM2. All 5578 probes are designed within predicted ORFs. For probe design we use state-of-the-art methodology and proprietary software. An amino linker is attached to the 5' end of each oligo.

ORF sequence source

1. **NC_003197** *Salmonella typhimurium* LT2, complete genome.
 - 1.1. **NC_003277** *Salmonella typhimurium* LT2 plasmid pSLT;
 - 1.2. **NC_002056** *Salmonella typhimurium* low copy number plasmid pSC101.
2. **NC_003198** *Salmonella typhi* – *Salmonella enterica* subspecies *S. enterica* serovar *Typhi*, complete genome.
 - 2.1. **NC_003384** *Salmonella enterica* subspecies *enterica* serovar *Typhi* plasmid pHCM1;
 - 2.2. **NC_002305** *Salmonella typhi* plasmid R27;
 - 2.3. **NC_003385** *Salmonella enterica* subspecies *enterica* serovar *Typhi* plasmid pHCM2.

GenBank IDs	Release date	Number of ORFs	ORF sequence prefix	Number of ORFs represented by an oligo
NC_003197	11/07/2001	4451	STM	4446
NC_003277	12/09/2002	102	PSLT	102
NC_002056	12/10/2002	2	pSC	2
NC_003198	12/09/2002	4395	STY	2510
NC_003384	12/09/2002	239	HCM1	228
NC_002305	04/23/2002	207	R27	207
NC_003385	12/09/2002	131	HCM2	131

A total of 5578 probes in the *Salmonella* Genome Oligo Set were designed from two strains and five plasmids. Each set of ORF sequences was obtained from the corresponding GenBank entries (NC_003197, NC_003277, NC_002056, NC_003198, NC_003384, NC_002305, NC_003385). Complete *Salmonella* genome GenBank entries were downloaded from the National Center for Biotechnology Information (NCBI) at Entrez Genomes (http://www.ncbi.nlm.nih.gov/PMGifs/Genomes/eub_g.html).

ORF selection for *Salmonella* AROS

The *Salmonella* Genome Oligo Set design is focused on *Salmonella typhimurium* LT2. All 4446 ORFs of the *Salmonella typhimurium* LT2 were included for oligo design. Using BLAST, all ORF sequences from the *Salmonella typhi* strain and five plasmids were aligned to each other as well as ORFs from *Salmonella typhimurium* LT2. If the ORF had less than 85% identity over the length of the ORF to any other ORF of the other strain or plasmid, that ORF was considered as unique and acceptable for probe design.

Probe design and selection rules

Once an ORF has been selected to be included in the set, a probe is selected with an optimal set of parameters. Sufficient numbers of 70mer candidate probes for each ORF are selected using the following criteria for the *Salmonella* Genome Oligo Set.

- 1) All oligos are within 78°C ±5°C using the following formula:
$$T_m = 81.5 + 16.6 \times \log[\text{Na}^+] + 41 \times (\#G + \#C)/\text{length} - 500/\text{length}$$
where $[\text{Na}^+] = 0.1 \text{ M}$ and $\text{length} = \#A + \#C + \#G + \#T$
- 2) Each oligo is selected anywhere within the ORF but does not occur within 70 bases from the 3' end of the ORF. This design allows the researcher to use either random priming or ORF-specific priming. Location from the 3' end is defined as the coordinate of the first base on the 5' end of the oligo from the 3' end of the ORF. All coordinates start from 1.
- 3) An oligo cannot have a contiguous single nucleotide base repeat or poly (N) tract longer than 8 bases.
- 4) An oligo cannot have a potential hairpin structure with a stem length longer than 8 bases.
- 5) A normalized score is assigned to each oligo based on the number of repeats. Oligos with more repeats having a normalized score greater than a certain threshold are filtered out.
- 6) Using BLAST, each oligo was aligned against all ORFs represented two *Salmonella* genomes and five plasmids. This BLAST percent identity is also referred to as cross-hybridization identity of the non-self gene. Each oligo selected has less than or equal to 70% identity to all other ORFs within the same genome or plasmid. For the *Salmonella typhi*, 2510 oligos are common (100% identical) with oligos

designed for *Salmonella typhimurium* LT2 strain. BLAST alignment results were used for final selection of unique oligos within each strain or plasmid and common oligos for two strains and five plasmids. The highest scoring non-self gene is defined as the sequence that yields the most matched bases in an alignment.

This cross-hybridization percent identity score is dependent on the size of the sequence database used to BLAST against, oligo sequence, and use of no-gap alignment method.

Once oligo candidates have been selected satisfying all the selection rules mentioned above, each oligo is ranked based on BLAST percent identity as computed in Step 6. One final oligo for each gene is selected with the minimum cross-hybridization identity.

Oligo selection criteria for <i>Salmonella</i> AROS	Value	Number of oligos in genome set satisfying these criteria
Length	70mer	5578
Melting temperature	78°C±5°C	
Location from 3' end	>70	
Poly(N) tract length	≤8	
Stem length in potential hairpin	≤8	
Cross-hybridization identity to all other genes	≤70%	

By request, cross-hybridization identity score and other score values of each oligo will be provided to the customer.

The following illustrations show distribution of melting temperature (T_m), GC content, location from 3' end of the ORF, and length of longest stem for all 5578 oligos.

Figure 1. Melting Temperature

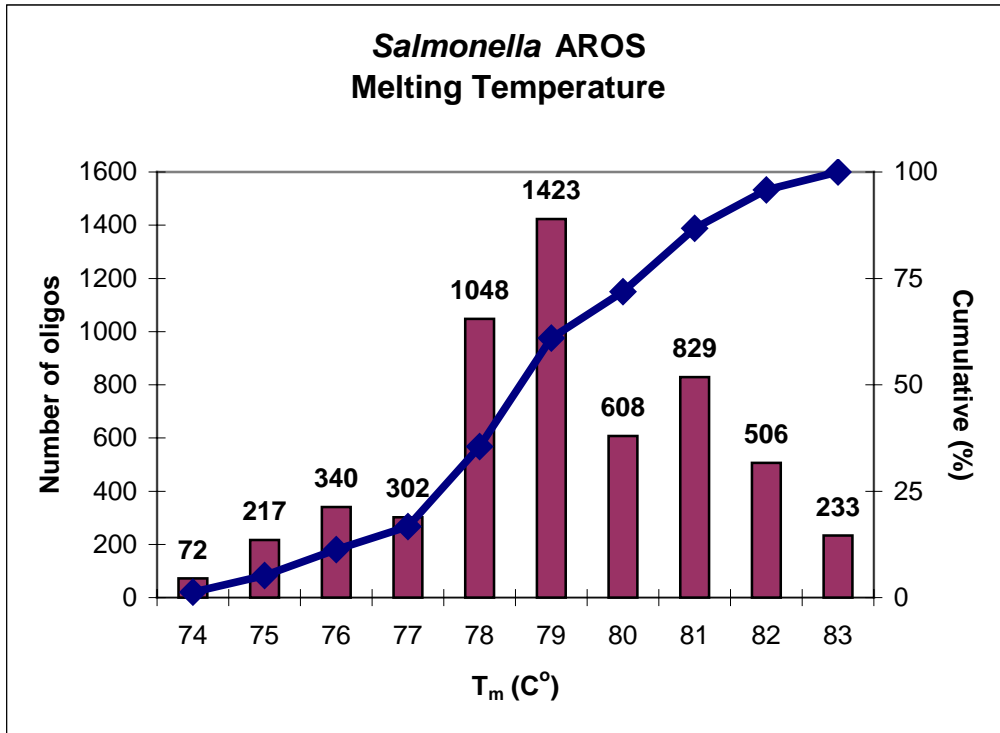


Figure 2. GC Content

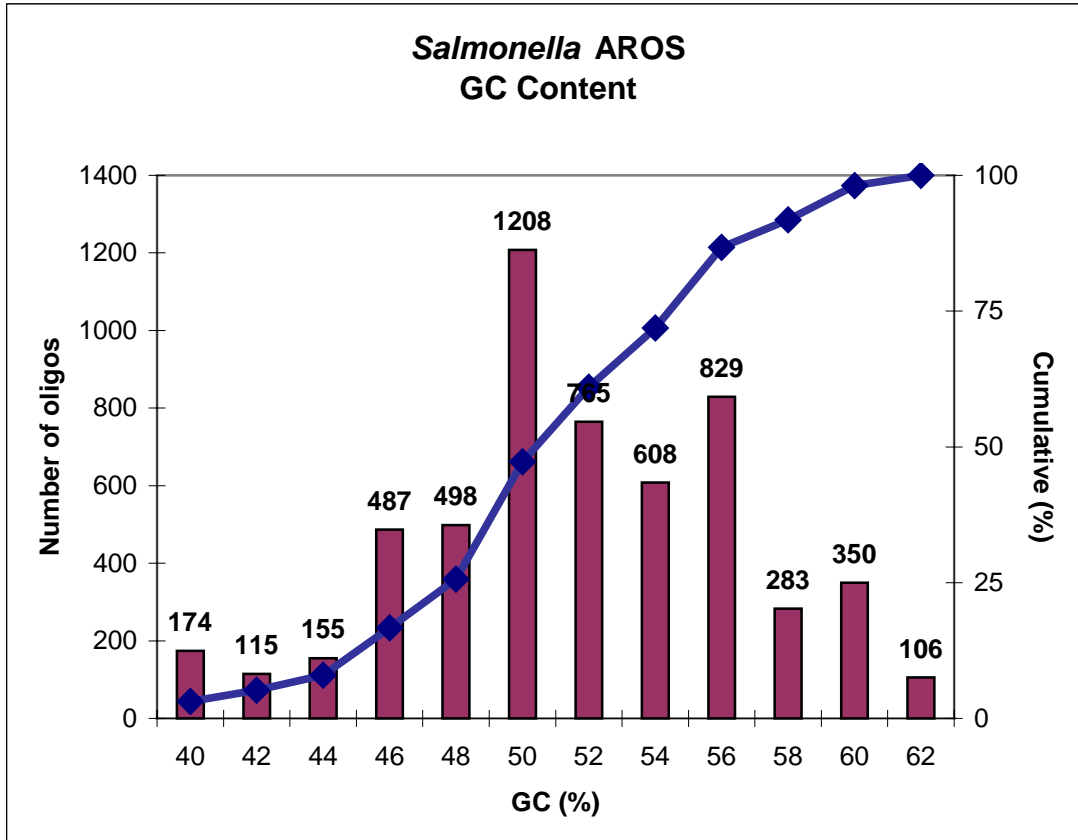


Figure 3. Location from 3' End

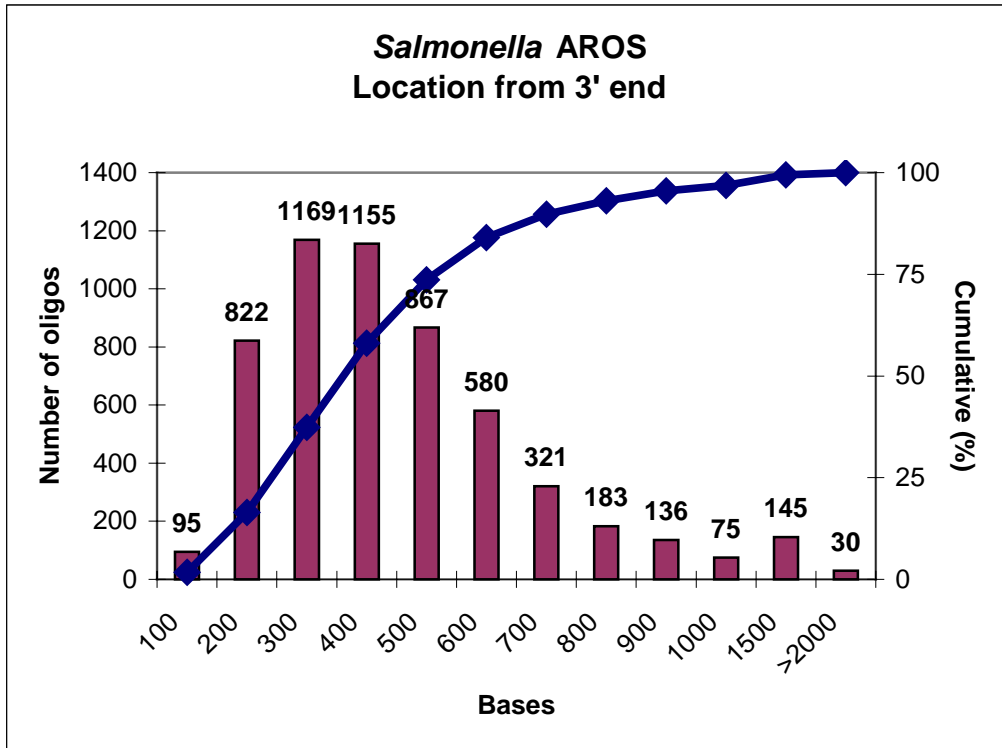


Figure 4. Length of the Longest Hairpin Stem

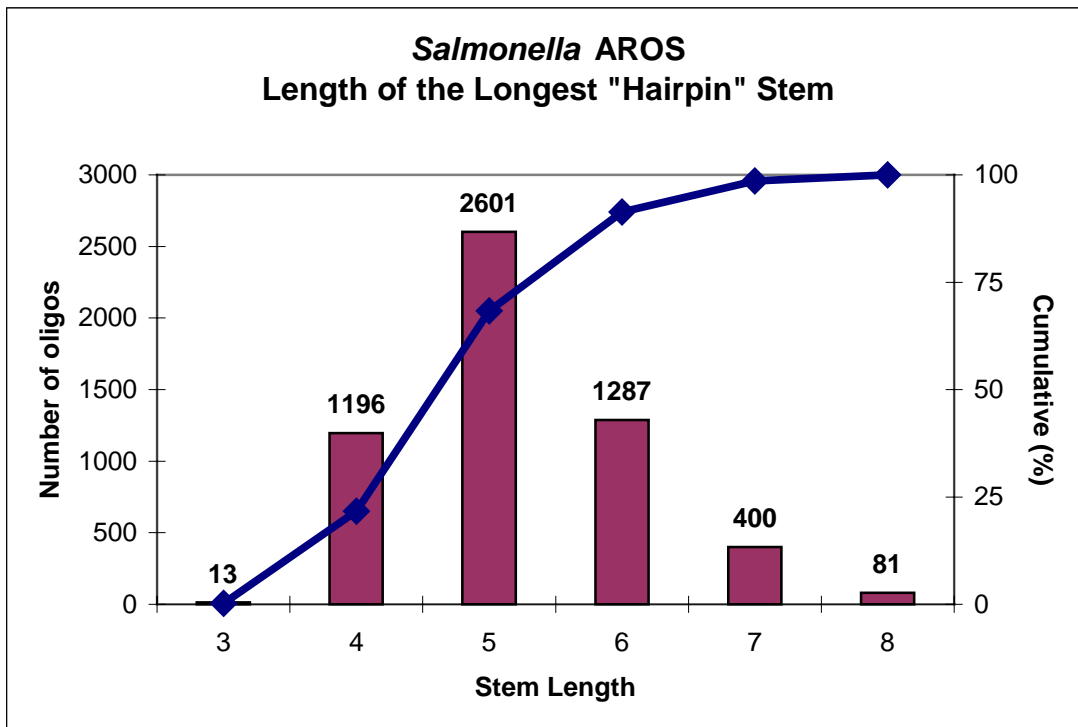
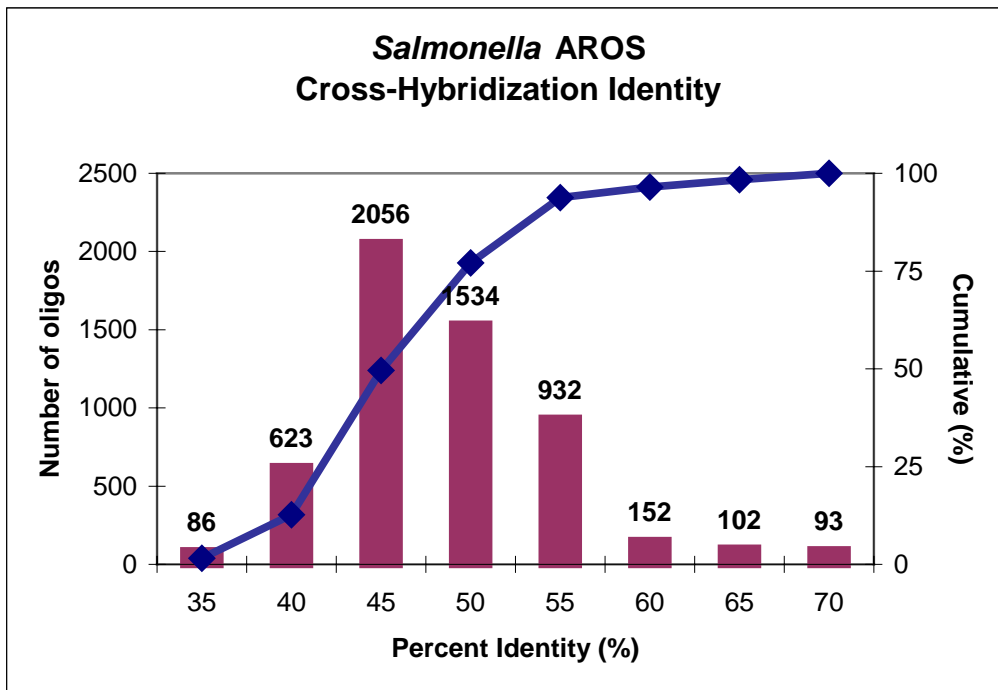


Figure 5. Cross-Hybridization Identity for *Salmonella* AROS



Quality check of probe design specifications

Once the final oligo has been selected to represent a gene or ORF, each oligo undergoes design specifications quality control where we use an independent method to confirm that all oligos have met the specified design specifications. The table below summarizes data from our quality check for probe design specifications for all 5578 oligos in the set.

Probe design specification	Expected value	Verified range	Number of oligos
Melting temperature (C°)	78°C \pm 5°C	73.0–83.0	5578
Poly(N) tract length	3–8	3–8	5578
Hairpin stem length (basepairs)	3–8	3–8	5578
Cross-hybridization identity (%)	34–70	34 –70	5578