

### Array-Ready Oligo Set™ for the *Listeria* Genome

Our *Listeria* Genome Oligo Set contains 2857 70mer probes representing 2857 ORFs from *Listeria monocytogenes* strain EGD. This set of 2857 probes also represents 877 ORFs from *Listeria innocua* Clip11262. For our probe design we use state-of-the-art methods and proprietary software.

#### ORF sequence and selection

##### Primary source

The 2857 probes for the *Listeria* Genome Oligo Set are designed from the National Center of Biotechnology Information (NCBI) RefSeq NC\_003210 for *Listeria monocytogenes* strain EGD (updated March 2002; source: [ftp://ftp.ncbi.nih.gov/genomes/Bacteria/Listeria\\_monocytogenes/NC\\_003210.gb](ftp://ftp.ncbi.nih.gov/genomes/Bacteria/Listeria_monocytogenes/NC_003210.gb)). More information is available at the following website: <http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/framik?db=Genome&gi=204>.

##### Secondary source

A second set of 3043 ORFs was obtained for *Listeria innocua* Clip11262 from the NCBI Refseqs NC\_003212 and NC\_003383 (updated in March 2002; source: [ftp://ftp.ncbi.nih.gov/genomes/Bacteria/Listeria\\_innocua](ftp://ftp.ncbi.nih.gov/genomes/Bacteria/Listeria_innocua)). More information is available at the following website: <http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/framik?db=Genome&gi=205>.

All oligos are designed directly from the primary set but are selected such that a large number of *L. innocua* ORFs are also represented. The final set of 2857 oligos in this set also represents 877 *L. innocua* ORFs. The gene list contains information for oligos and their corresponding *L. monocytogenes* and *L. innocua* ORFs.

#### Probe Design and Selection Rules

A probe is selected with an optimal set of parameters. Large numbers of 70mer candidate probes for each *L. monocytogenes* ORF are selected using the following criteria.

1) An oligo is within  $72^{\circ}\text{C} \pm 5^{\circ}\text{C}$  using the following formula:

$$T_m = 81.5 + 16.6 \times \log[\text{Na}^+] + 41 \times (\#G + \#C) / \text{length} - 500 / \text{length} \text{ where } [\text{Na}^+] = 0.1 \text{ M and length} = \#A + \#C + \#G + \#T$$

2) To accommodate both random and gene-specific priming, the oligo is located more than 50 bases away from both the 5' and 3' end of the ORF. In other words, the 5' end of the oligo is more than 50 bases from the 5' end of the ORF and the 3' end of the oligo is at least 50 bases from the 3' end of the ORF.

3) An oligo cannot have a contiguous single nucleotide base repeat or poly (N) tract longer than 7 bases.

4) An oligo cannot have a potential hairpin structure with a stem length longer than 7 bases.

5) Each oligo has less than or equal to 70% identity to all other ORFs. Each oligo, designed from the primary set of *L. monocytogenes* ORFs, is aligned against the primary set of all ORF sequences in the *Listeria monocytogenes* genome. Using the alignment with the candidate oligo versus the highest scoring non-self sequence, a BLAST percent identity score is computed. The highest scoring nonself sequence is defined as the sequence that yields the most matched bases in an alignment. This BLAST percent identity is also referred to as cross-hybridization identity of the oligo.

6) Each oligo of any length cannot have greater than 20 contiguous bases common to any other gene.

7) Using BLAST, each oligo designed from the primary set of *L. monocytogenes* ORFs is aligned against the secondary set of *Listeria innocua* ORFs. This is done so that oligos designed from the primary set can also represent a large fraction of ORFs in the secondary set.

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 5. One final oligo for each gene is selected with the minimum cross-hybridization identity.

Note that for 27 (0.9%) ORFs that did not yield oligos satisfying all the above criteria, certain rules were relaxed.

#### SUMMARY

Oligo selection criteria	Value	Number of oligos in genome set satisfying these criteria
Length Melting temperature Location from 5' and 3' end Poly(N)tract length Stem length in potential hairpin Cross-hybridization identity to all other ORFs Contiguous base match to all other ORFs	70mer 72°C ± 5°C > 50 ≤ 7 ≤ 7 ≤ 70% ≤ 20	2830
Total number of oligos not satisfying one or more of the above criteria		27
Location from 5' and 3' end	≤ 50	19*
Cross-hybridization identity to all other ORFs	>70%	8*
Contiguous base match to all other ORFs	>20	8*
<b>Total</b>		<b>2857</b>

\*Out of 27 probes.

As mentioned above, each *L. monocytogenes* oligo is aligned against all *L. innocua* ORFs. The final oligos for each *L. monocytogenes* ORF were selected to meet the above criteria and to cover a large number of the *L. innocua* ORFs. The following table shows the number of *L. innocua* ORFs covered by a *L. monocytogenes* oligo.

Cumulative crosshybridization identity with a <i>L. monocytogenes</i> oligo (%)	Number of <i>L. innocua</i> ORFs represented by an oligo	Fraction of <i>L. innocua</i> represented by an oligo
>90%	2127	70.0%
>95%	967	31.8%
>97%	877	28.8%

The following illustrations show the distribution of all 2857 oligos for melting temperature, GC content, location from 3' end of gene sequence, length of maximum stem length, and cross-hybridization identity.

Figure 1. Melting Temperature

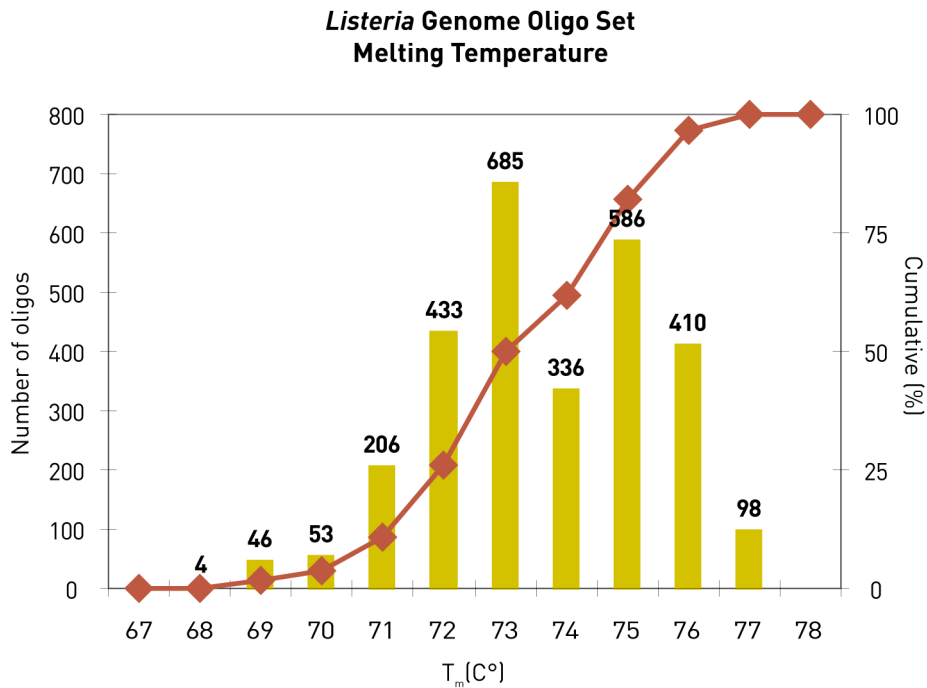


Figure 2. GC Content

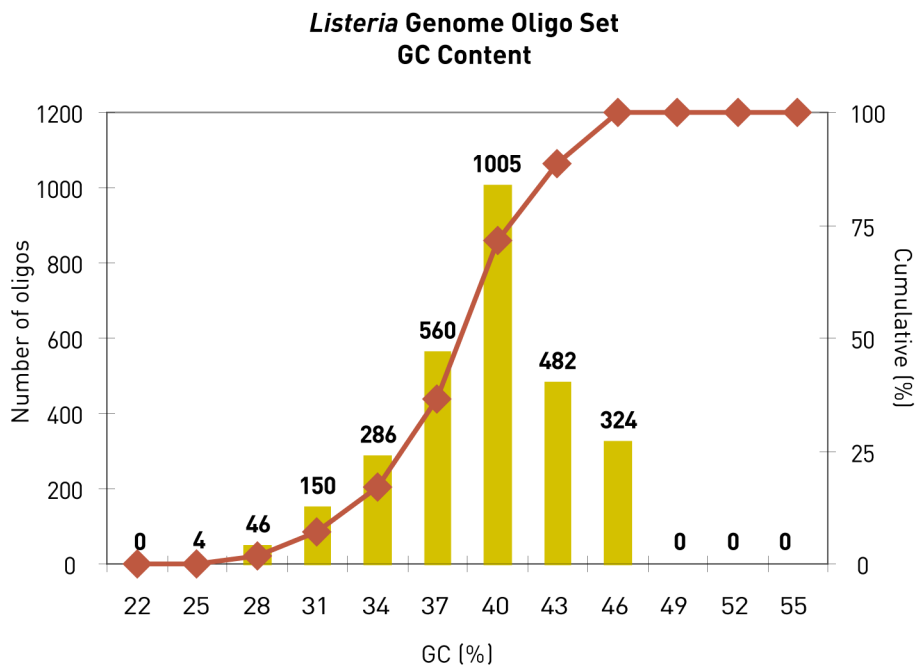


Figure 3. Location from 3' End

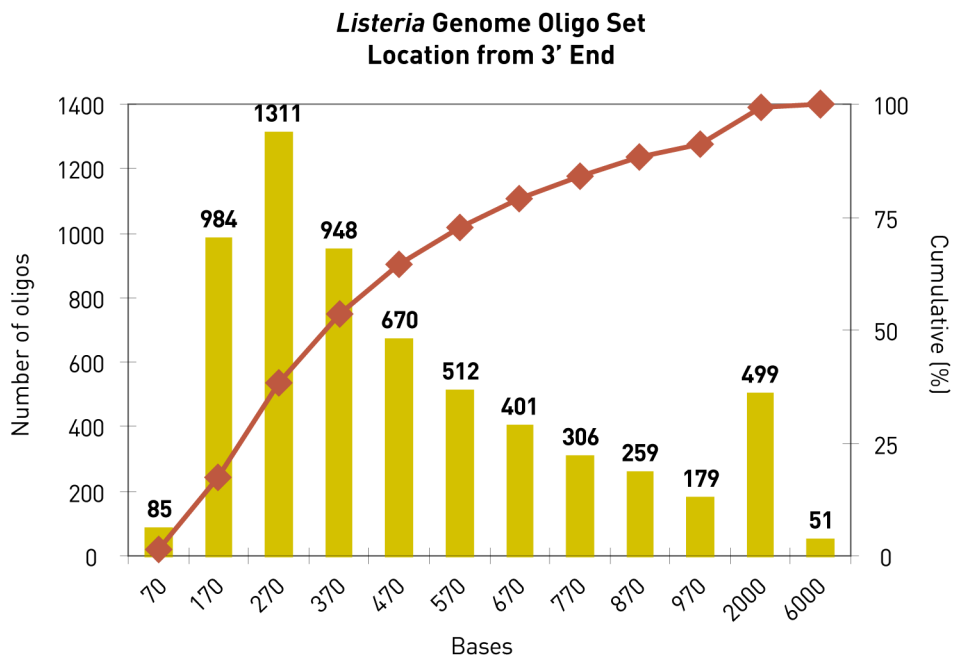


Figure 4. Length of the Longest Hairpin Stem

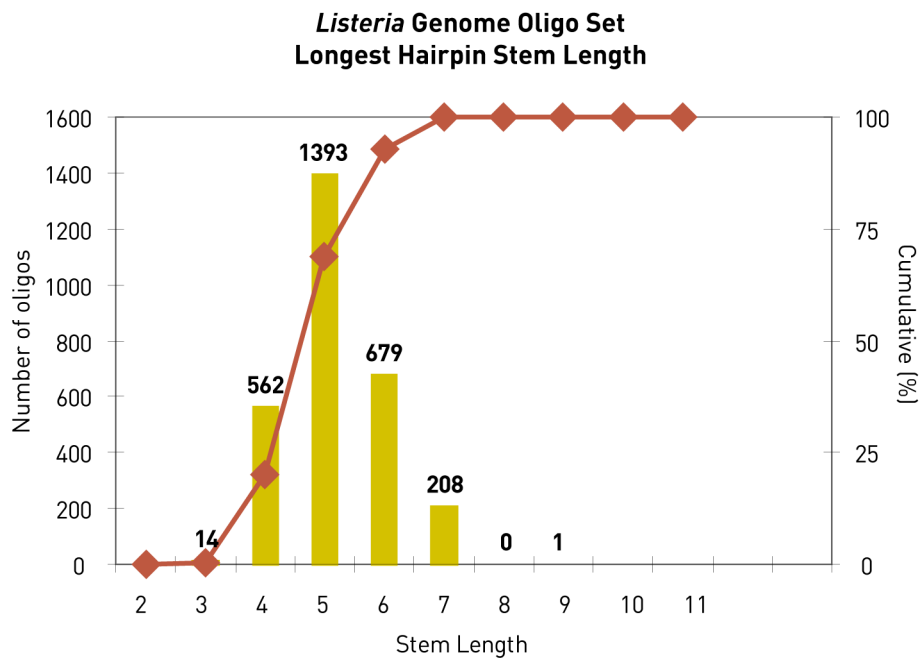
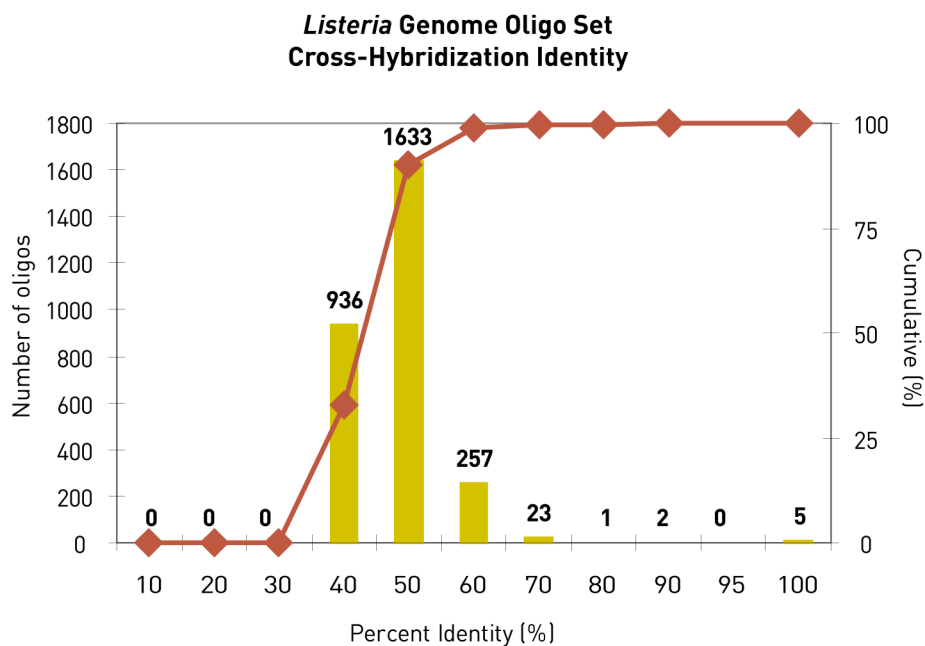


Figure 5. Cross-Hybridization Identity



#### Quality Check of Probe Design Specifications

Once the final oligo has been selected to represent a gene, 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 2857 oligos in the set.

Probe design specification	Expected value	Verified range	Number of oligos
Melting temperature (C°)	72°C ± 5°C	67.3– 76.5	2857
Cross-hybridization identity (%)	≤70	31–70	2849
Cross-hybridization identity (%)	71–100	77–100	8