

Array-Ready Oligo Set™ for the Grape (*Vitis vinifera*) Genome Version 1.0

We are pleased to announce the release of our Array-Ready Oligo Set™ for the *Vitis vinifera* genome containing 14,562 70mer probes representing 14,562 transcripts from The Institute for Genomic Research (TIGR) Grape Gene Index (VvGI), release 3. 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.

Transcript Sequence Source

A total of 14,562 probes in the Grape Genome Oligo Set were designed from transcript sequences obtained from TIGR Grape Gene Index (VvGI), release 3 (August 13, 2003). The complete grape gene index was downloaded from the TIGR website (<http://www.tigr.org/tdb/tgi/vvgi/index.shtml>).

Probe Design and Selection Rules

Once a transcript 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 transcript are selected using the following criteria for the Grape Genome Oligo Set.

- 1) All oligos are within $78 \pm 5^\circ\text{C}$ using the following formula:
 $T_m = 81.5 + 16.6 \times \log[\text{Na}^+] + 41 \times (\#G + \#C)/\text{length} - 500/\text{length}$ where $[\text{Na}^+] = 0.1 \text{ M}$ and $\text{length} = \#A + \#C + \#G + \#T$
- 2) Each oligo is designed within 1000 bases from the 3' end of the available gene sequence. Coordinate starts from 1.
- 3) An oligo cannot have a contiguous single nucleotide base repeat or poly (N) tract longer than 9 bases.
- 4) An oligo cannot have a potential hairpin structure with a stem length longer than 9 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 the Basic Local Alignment Search Tool (BLAST), each oligo was aligned against all transcripts, representing the *Vitis vinifera* genome. 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 transcripts. BLAST alignment results were used for final selection of unique oligos within the genome. 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.
- 7) Each oligo cannot have greater than 20 contiguous bases common to any non-self transcript.

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	Value	Number of oligos in genome set satisfying these criteria (out of 14,562)
Length	70mer	14,240
Melting temperature	78 ± 5°C	
Location from 3' end	≤1000	
Poly(N) tract length	≤9	
Stem length in potential hairpin	≤9	
Contiguous bases common to any non-self transcript	≤20	
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.

Please note that for 322 transcripts that did not yield oligos satisfying all the above criteria certain rules were relaxed. For those genes, one or more of the following criteria may apply:

- T_m less than 73°C or greater than 83°C;
- Probe location greater than 1000 bases from 3' end of the gene or transcript;
- Greater than 70% cross-hybridization percent identity;
- Contiguous base match greater than 20 bases to another gene.

The following illustrations show distribution of melting temperature (T_m), GC content, location from 3' end of the transcript, length of longest stem, and cross-hybridization identity for all 14,562 oligos.

Figure 1. Melting Temperature

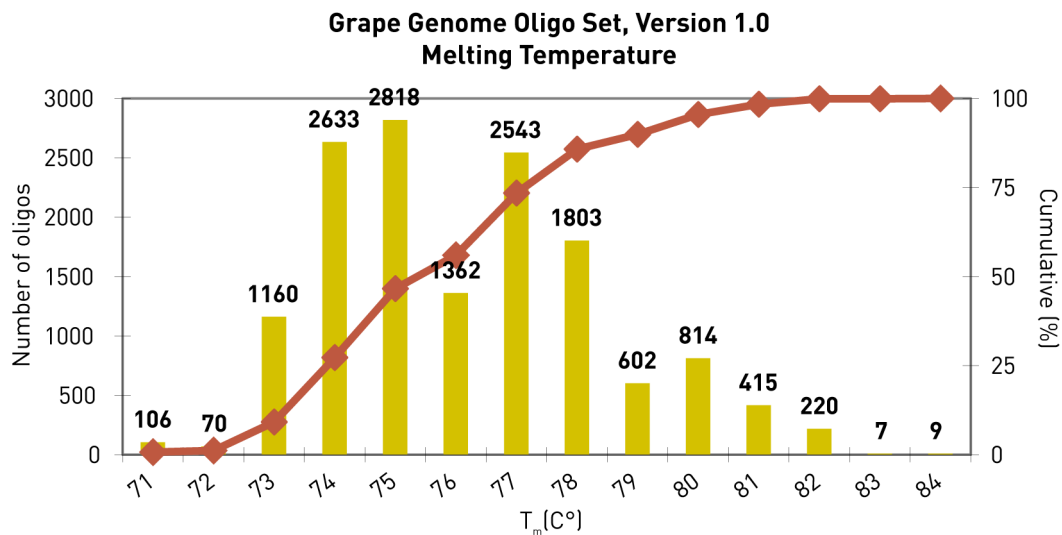


Figure 2. GC Content

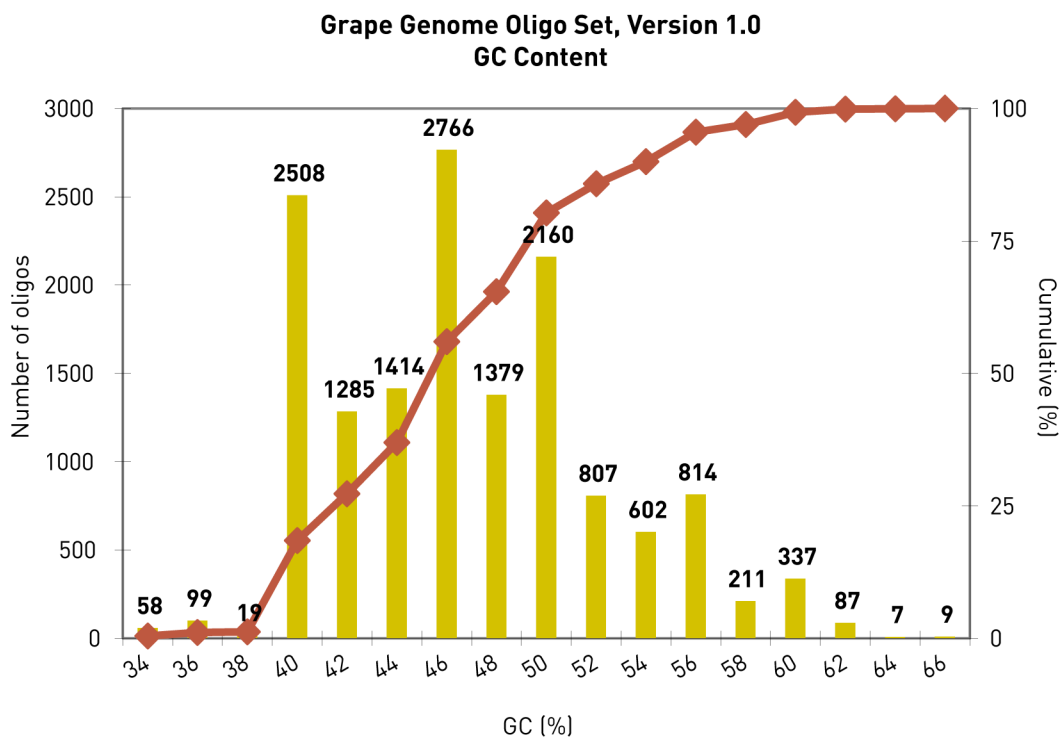


Figure 3. Location from 3' End

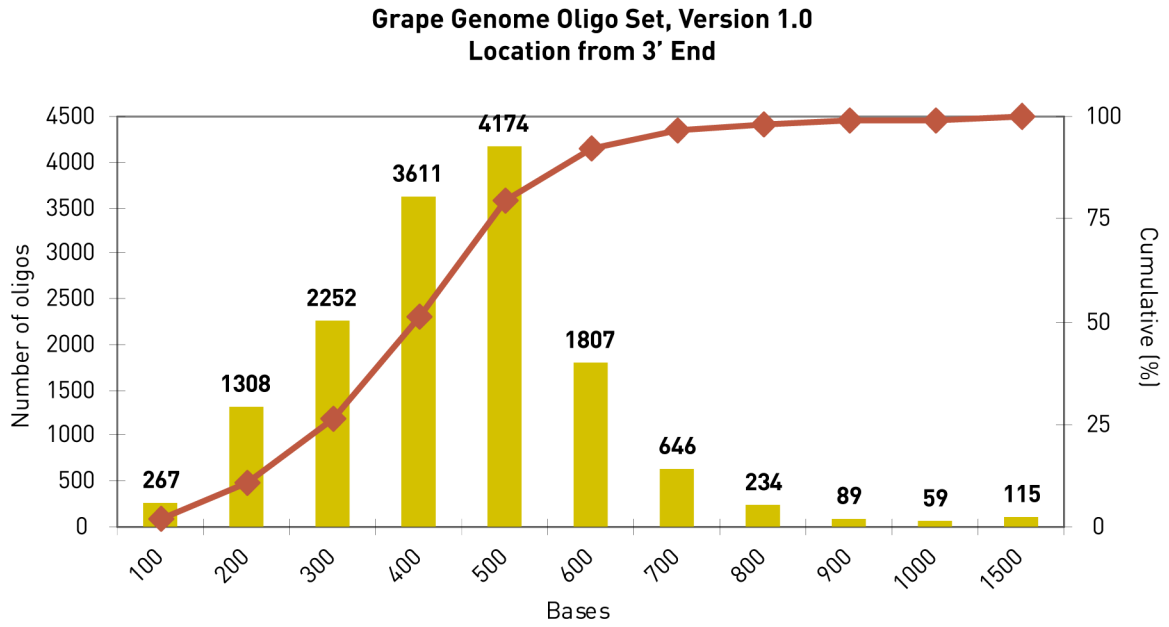


Figure 4. Length of the Longest Hairpin Stem

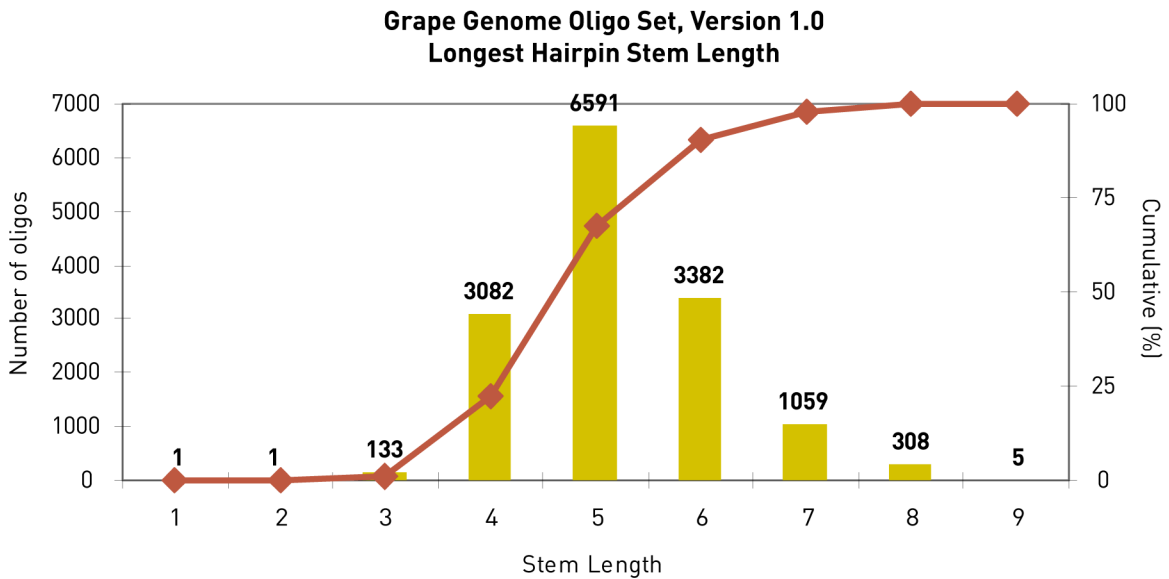
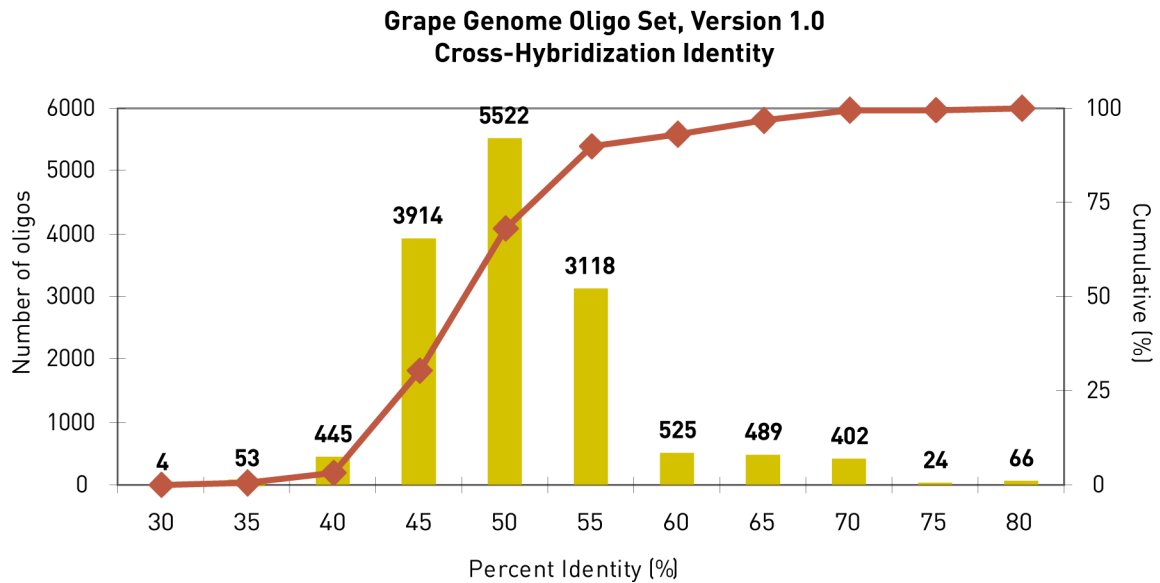


Figure 5. Cross-Hybridization Identity



Quality Check of Probe Design Specifications

Once the final oligo has been selected to represent a gene or transcript, 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 14,562 oligos in the set.

Probe design specification	Expected value	Verified range	Number of oligos
Melting temperature (C°)	78 ± 5 °C	73.0-83.0	14,440
Poly(N) tract length	1-9	2-9	14,562
Hairpin stem length (basepairs)	1-9	1-9	14,562
Contiguous bases common to any non-self transcript	2-20	12-20	14,240
Cross-hybridization identity (%)	10-70	25-70	14,472