

### Array-Ready Oligo Set™ for the Yeast Genome

Our Yeast Genome Oligo Set contains 6388 70mer probes representing 6388 *Saccharomyces cerevisiae* open reading frames (ORFs). An amino linker is attached to the 5' end of each oligo.

#### ORF Sequence Source

All 6388 probes are designed from the set of open reading frames (ORFs) obtained from the *Saccharomyces* Genome Database (SGD) at Stanford University (<http://genomewww.stanford.edu/Saccharomyces/>).

#### Probe Design

For each ORF sequence, a probe is selected with an optimal set of parameters. Large numbers of 70mer candidate probes for each ORF are selected using the following criteria for the Yeast Genome Oligo Set.

- 1) All oligos are within  $73^{\circ}\text{C} \pm 5^{\circ}\text{C}$ . using the following formula:  
 $T_m = 81.5 + 16.6 \times \log[\text{Na}^+] + 41 \times (\#G + \#C)/\text{length} - 500/\text{length}$  where  $[\text{Na}^+] = 0.1 \text{ M}$  and  $\text{length} = \#A + \#C + \#G + \#T$
- 2) Each oligo is within 750 bases from the 3' end of the ORF
- 3) All oligos are scanned for potential hairpins stems
- 4) Using BLAST, each oligo is aligned against all the most current set of SGD ORF sequences. Using the alignment with the candidate oligo versus the highest scoring nonself ORF, a BLAST percent identity score is computed. The highest scoring non-self ORF 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.

One oligo with the minimum cross-hybridization identity is then selected.

A total of 17 oligos are longer or shorter than 70 bases to compensate for high or low GC content in the ORF sequences and one oligo is located more than 750 bases from the 3' end.

#### SUMMARY

Oligo selection criteria	Value	Number of oligos in genome set satisfying these criteria
Length	70mer	6370
Melting temperature	$73^{\circ}\text{C} \pm 5^{\circ}\text{C}$	
Location from 3' end	$\leq 750$	
Length	58–61	13
Length	71–88	4
Location from 3' end	$> 750$	1
<b>Total</b>		<b>6388</b>

A large number of ORFs in the yeast genome have overlapping regions with each other. Using BLAST, all 6388 ORFs were aligned versus each other. The table below summarizes the number of overlaps and the percent identity.

Length of overlapping region between two yeast ORFs	Percent identity of overlap	Number of ORFs >500 bases
> 500 bases	> 95%	240
> 100 bases	> 95%	425
> 50 bases	> 95%	449
> 500 bases	> 90%	291
> 500 bases	> 80%	358

All 6388 oligos were aligned against one another to obtain a percent identity of each probe versus all other non-self probes. The following table summarizes oligo versus oligo percent identity. Again the percent identity is the number of matched bases divided by the length of the oligo.

Percent identity of oligo versus another oligo	Cumulative number of oligos	Cumulative percent of oligos
≤ 30%	2600	40.7%
≤ 40%	4939	77.3%
≤ 70%	6033	94.4%
≤ 90%	6092	95.4%
100%	198	3.1%

#### Design Criteria Distributions

The following illustrations show the distribution of all 6388 oligos for melting temperature, GC content, location from 3' end of ORF sequence, length of longest hairpin stem length, and cross-hybridization identity.

In Figure 5, to generate the cross-hybridization identity graph, each oligo was aligned using BLAST against the most current set of SGD ORFs.

Figure 1. Melting Temperature

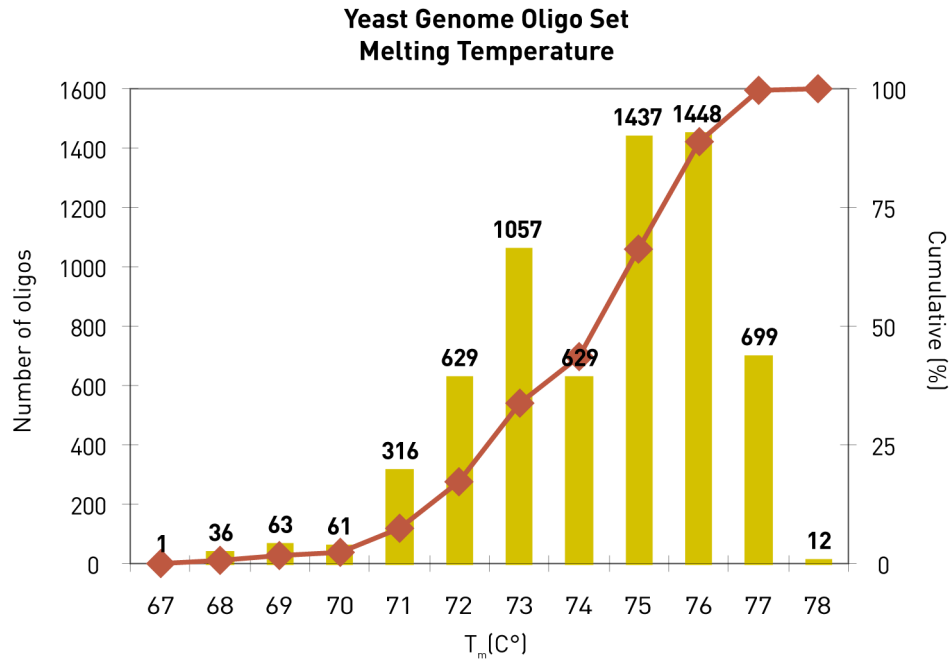


Figure 2. GC Content

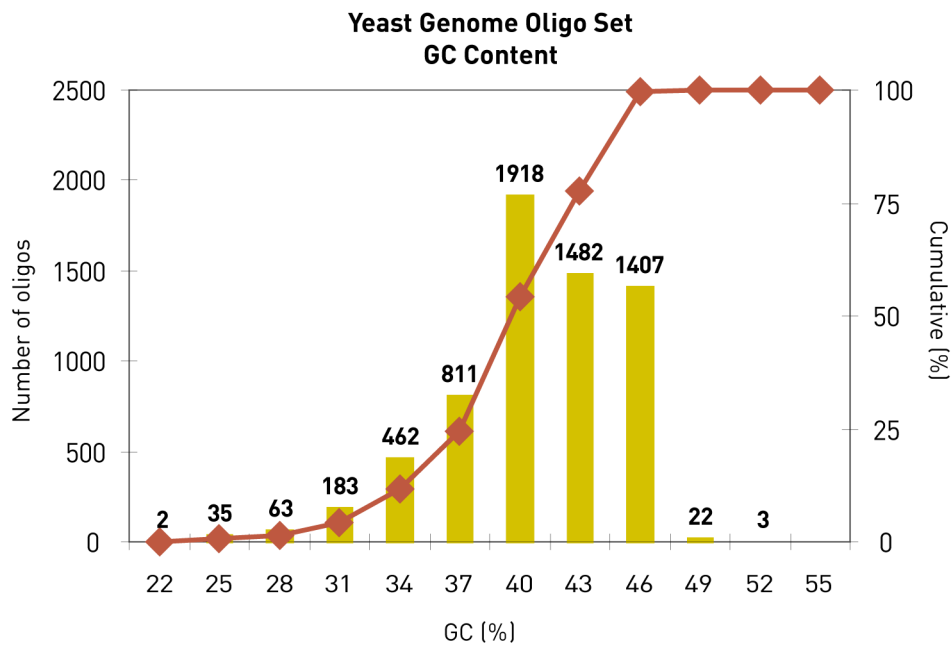


Figure 3. Location from 3' End

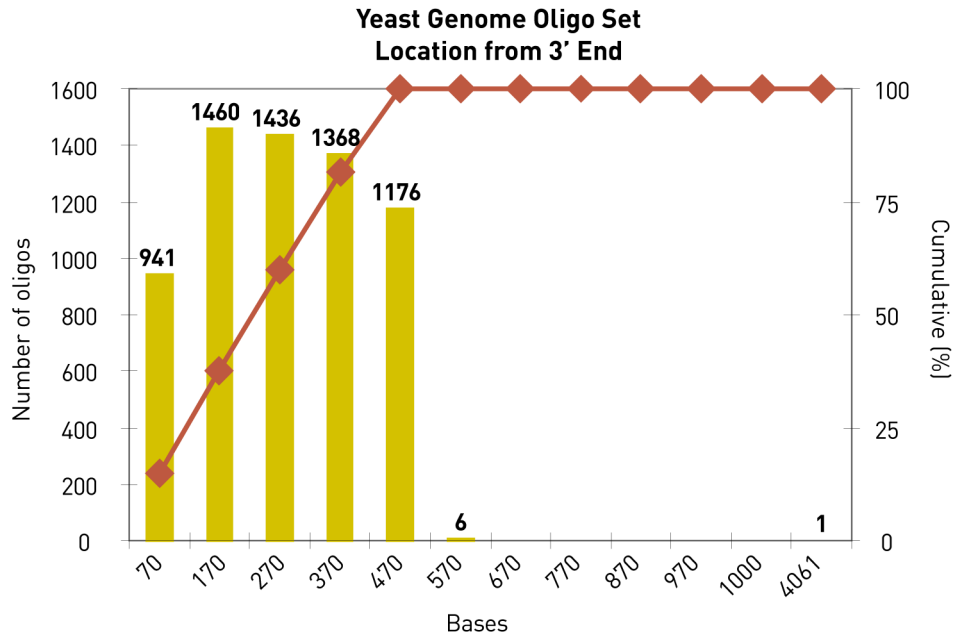


Figure 4. Length of the Longest Hairpin Stem

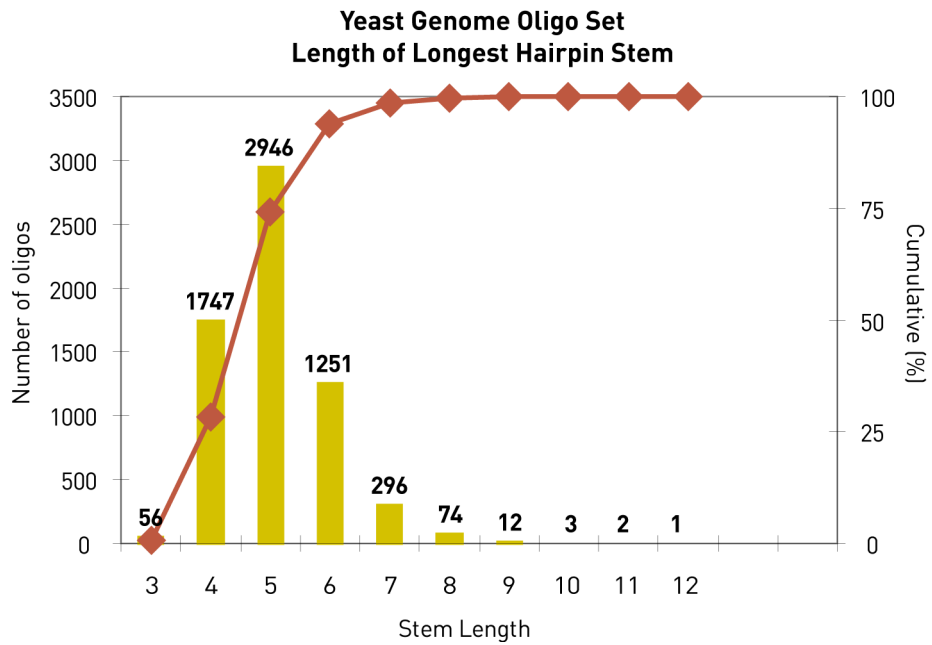


Figure 5. Cross-Hybridization Identity

