

## Array-Ready Oligo Set™ for the *Xenopus tropicalis* Genome Version 1.0

We are pleased to announce the release of our *Xenopus tropicalis* Genome Oligo Set Version 1.0. The *Xenopus tropicalis* Genome Oligo Set V1.0 contains 10898 70mer probes representing 10898 transcripts. The transcript sequences were derived from a large scale *Xenopus tropicalis* EST project carried out at the Wellcome Trust/Cancer Research UK Gurdon Institute in Cambridge, and were supplemented by ESTs from NCBI. 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 Sequences Source

A total of 10,878 probes in the *Xenopus tropicalis* Genome Oligo Set V1.0 were designed from transcript sequences obtained from the Wellcome Trust/Cancer Research UK Gurdon Institute in Cambridge.

### 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 *Xenopus tropicalis* Genome Oligo Set.

- 1) All oligo  $T_m$  is within  $78^\circ\text{C} \pm 5^\circ\text{C}$  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 2000 bases from the 3' end of the available transcript sequence.
- 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 *Xenopus tropicalis* 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 transcript is selected with the minimum cross-hybridization identity.

Note that for 197 (1.8%) transcripts 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	70mer	10701
Melting temperature	78°C ± 5°C	
Poly(N)tract length	< 9	
Stem length in potential hairpin	< 9	
Cross-hybridization identity to all other transcripts	≤ 70%	
Contiguous base match to all other transcripts	≤ 20	
Total number of oligos not satisfying one or more of the above criteria		197
Poly(N)tract length	≥ 9	7*
Stem length in potential hairpin	≥ 9	2*
T <sub>m</sub>	< 73 °C or > 83 °C	102*
Cross-hybridization identity to all other transcripts	> 70%	44*
Contiguous base match to all other transcripts	> 20	86*
<b>Total</b>		<b>10898*</b>

Figure 1. Melting Temperature

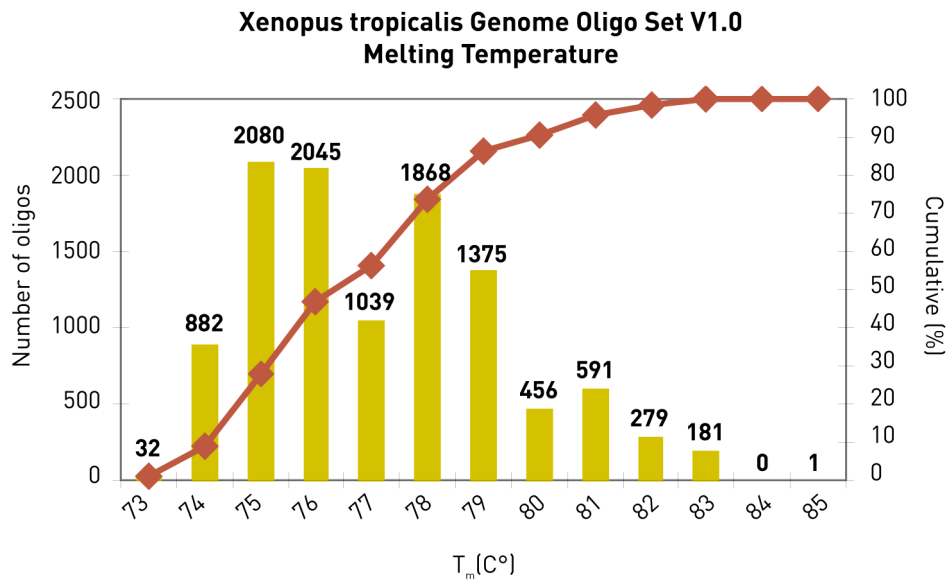


Figure 2. GC Content

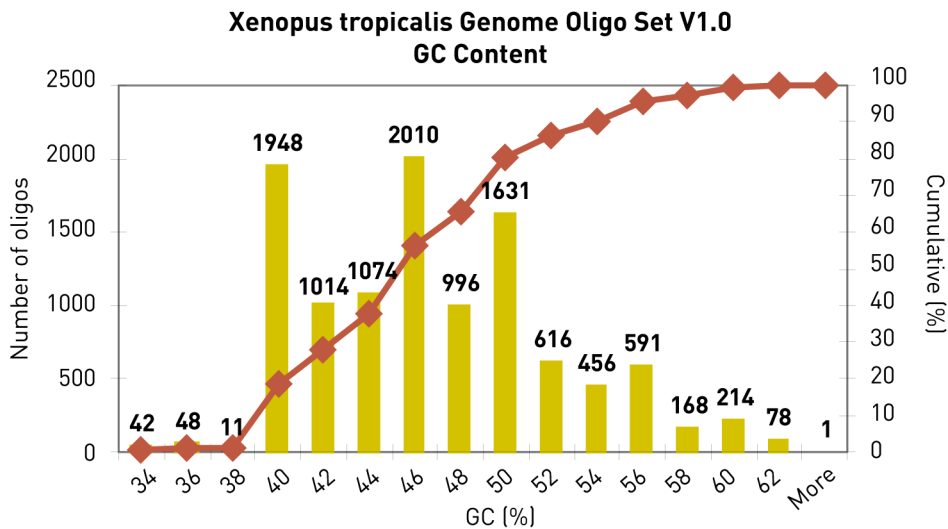


Figure 3. Location from 3' End

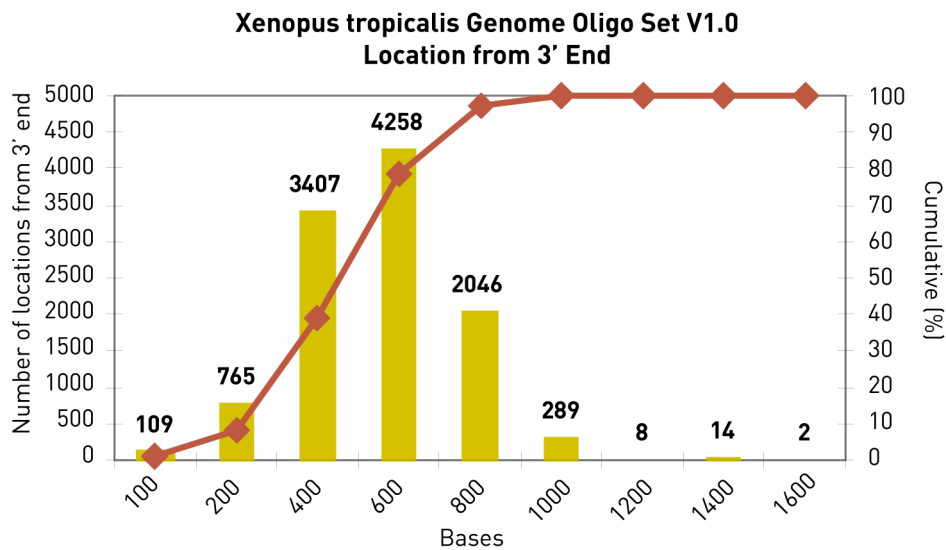


Figure 4. Length of the Longest Hairpin Stem

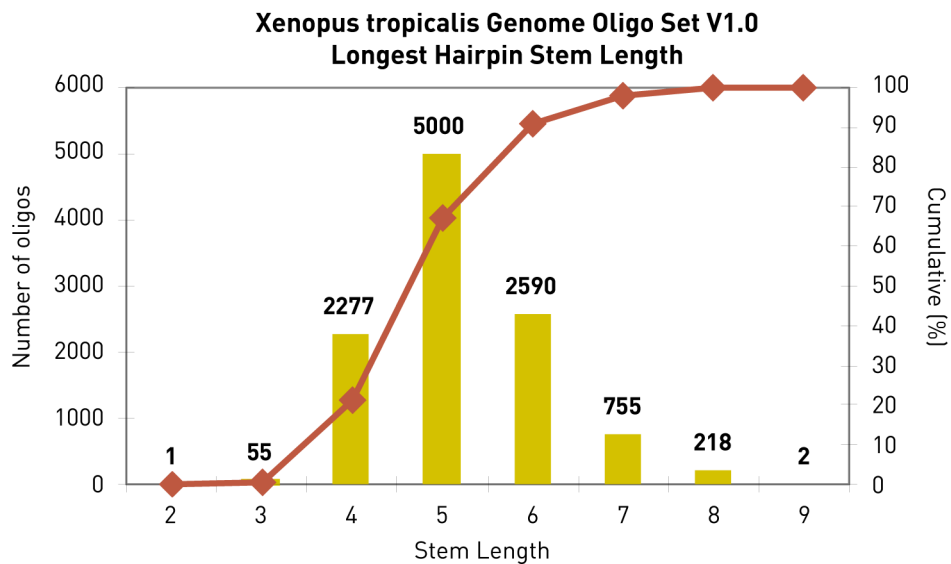
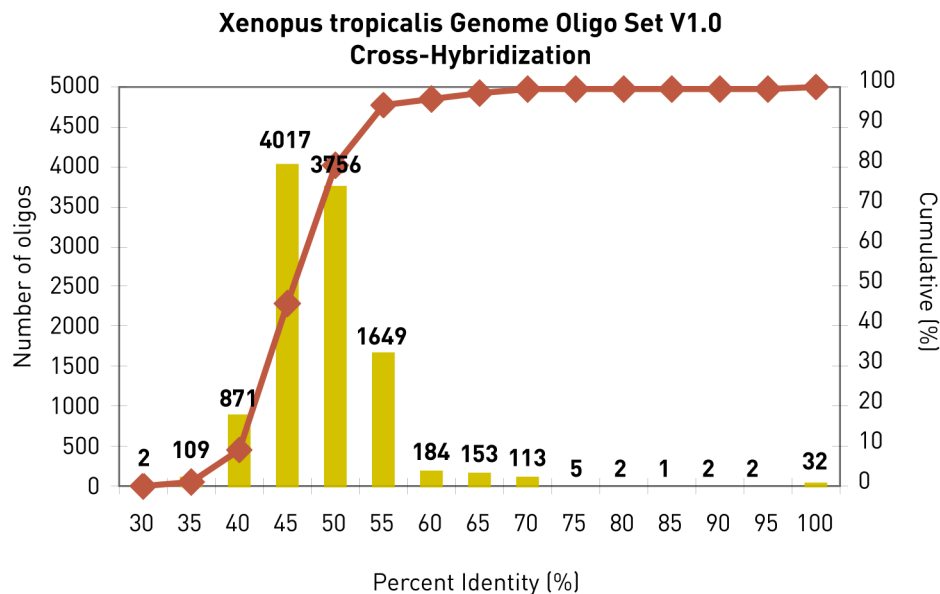


Figure 5. Cross-Hybridization Identity



Quality check of probe design specifications Once the final oligo has been selected to represent a 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 10898 oligos in the set.

Probe design specification	Expected value	Verified range	Number of oligos
Melting temperature (C°)	78°C ± 5°C	71.23–84.11	10898
Cross-hybridization identity [%]	≤ 70	27–70	10854
Cross-hybridization identity [%]	71–100	71–100	44

#### Acknowledgements

The *Xenopus tropicalis* Genome Oligo Set V1.0 design was based on sequence data provided by researchers at the Wellcome Trust/Cancer Research UK Gurdon Institute in Cambridge. The main contributors were Dr Enrique Amaya, Dr Andrew Chalmers, Dr Michael J. Gilchrist, Dr Nancy Papalopulu and Professor James C. Smith. Sequences were derived from a large scale *Xenopus tropicalis* EST project carried out in collaboration with the Wellcome Trust Sanger Institute, and were supplemented by ESTs from NCBI, and, by kind permission, from Nicolas Pollet at IBAIC, Université Paris-Sud, Orsay. Transcript identification was performed with computational support from the Cambridge-Cranfield High Performance Computing Facility.