

**Assembly:** GCA\_024586455.2\_APGP\_CSIRO\_Bneo\_wtdbg2-racon-allhic-juicebox.fasta\_v2\_genomic

5' motif: GCGGT 3' motif: ATCGC Internal max. Poly-T: 3nt Trailing-T: 8nt PSE: 0.84

Bac\_neo2 | CM044975.1:10331670-10332109 (-) | 440 nt | lncRNA:noe consensus e-value: 1.2e+00

5' motif: GCGGT 3' motif: ATCGC Internal max. Poly-T: 4nt Trailing-T: 6nt PSE: 0.90

**Bac\_neo3 | CM044975.1:10077181-10077616 (+) | 436 nt | lncRNA: noe consensus e-value: 3.2e-03**

5' motif: GCGGT 3' motif: ATCGC Internal max. Poly-T: 4nt Trailing-T: 6nt PSE: 0.90

**Bac\_neo4 | CM044977.1:56095089-56095385 (-) | 297 nt | IncRNA:noe consensus e-value: 2.0e-08**

5' motif: GCGGT 3' motif: ATCGC Internal max. Poly-T: 4nt Trailing-T: 6nt PSE: 0.93

GGCGGGTT CAGAT CCTCCCA GTTGAAAA GGTGGT CACAT CCTCA ATTGTG ACAAAT GTATAC TTTGAA TAAAG CAAACAT AAAAAAAAAA GCATAC GCATAC GCATATAC AACA AAAAAA CGAAAA GCAAAAAAAT GAATCA ACAAACAAG  
AATCTAT GGTGAAAAC TGTAAAAA CAAC TTCAAT GATCAC TTTCTAT CCTTAA GAA GCGATT TGAAGG TTAAG AAAAAA TTC TCTCAAT TTTTCT GTCG AAGGGT TAATT GATAA GATAC TGTG GAGGT CAAT GTGAT CGCCT TTTT

	Bac_neo1	Bac_neo2	Bac_neo3	Bac_neo4
Bac_neo1	-	30	30	21
Bac_neo2		-	89	41
Bac_neo3			-	40
Bac_neo4				-

Pairwise sequence identity matrix (%) showing similarity between lncRNA sequences.

[illegible]

```

U3      -----ttatactcttttttctttccctttgtctttcatttctgcgAATTCTCAACTATTTTATGtgccctgttattgaaataccctgcgttggcctttgaaaaatt-----
U3      -----gtatdctcttttttctttcatttctttcatttctccgAATTCTCAACTATTTTATGCGtatattttdttdaaataccctgcgttggcctttgaaaaatt-----
U3      -----atactgcttttttctttcatttctttcatttctgcgAATTCTCAACTACTATTTTgcgtatatttggtaaataccctgcgttggcctttgaaaaatt-----
U3      -----ttatactcttttttctttcatttctttcatttctccgAATTCTCAACTATTTTATGtgcccttttattgaaataccctgcgttggcctttgaaaaatt-----
U3      -----gaagtatactgcttttttctttcatttctttcatttctgcgAATTCTCAACTATTTTATGtgacccttttgtgaaataccctgcgttggcctttgaaaaatt-----
U4      -----gcaaggcatggttgtgtcagcgcacaacgtactgacttgcAATTCTCATATGGTTAATAaatatgtgcaatgaatagttttttgttggcctttttaaatt-----
U4      -----ttcactdccaatcacctttccaaattddccattaccaactACTTCTCAACCAACTCATAAatdadaaattttaaataotccatottoddaadcatcttt-----
U4atac  -----tcggataataatcttttctgtaccgcgggttcagtgaaatgcAATTCTCAATGATTTGTgaaacgcattatagaatttgaattgtttgcacaact-----
U5      -----caggcagttgacgtctttgttgaatatcatcttgcctgtACTTCAACCCCTTCATGgcacgtgtatactaatagcaaatgtatgcgtacataaag-----
U5      -----cctgcctaagdtatctttaatttttctttcacttttaaaAATTCTCAACCATTTTATAtatttctacactcfatdctttctctttcttttgcacaaa-----
U5      -----catagttcaagcataatttcacatgtgtcttatatctttacAATTCTCAACCCTTTTAAAGttattcctagcttataaacttcagtatctgctttgtacag-----
U12     -----cttcattttatactttactgcactaacaacacccctctGATTCTCAACCTGTTTAAAtacaacactgggtaaatagcatgttgatgtctgctgtataatt-----

```

Upstream promoter alignment showing conserved regions. Sequences include Pol III promoters (noeCR34335, Arthropod\_7SK, 7SL, RNase\_MRP, RNaseP\_nuc, U6, U6atac, tRNAsec, snoRNAMe18SA1806) and Pol II promoters (U1, U2, U3, U4, U4atac, U5, U7, U8, U11, U12, OrCD1). Conserved positions are highlighted in the alignment.