

Additional file 3 of Heat-induced transcriptional activation of LTR retrotransposons in the regenerative flatworm *Macrostomum lignano*

Supplementary Figure Legends

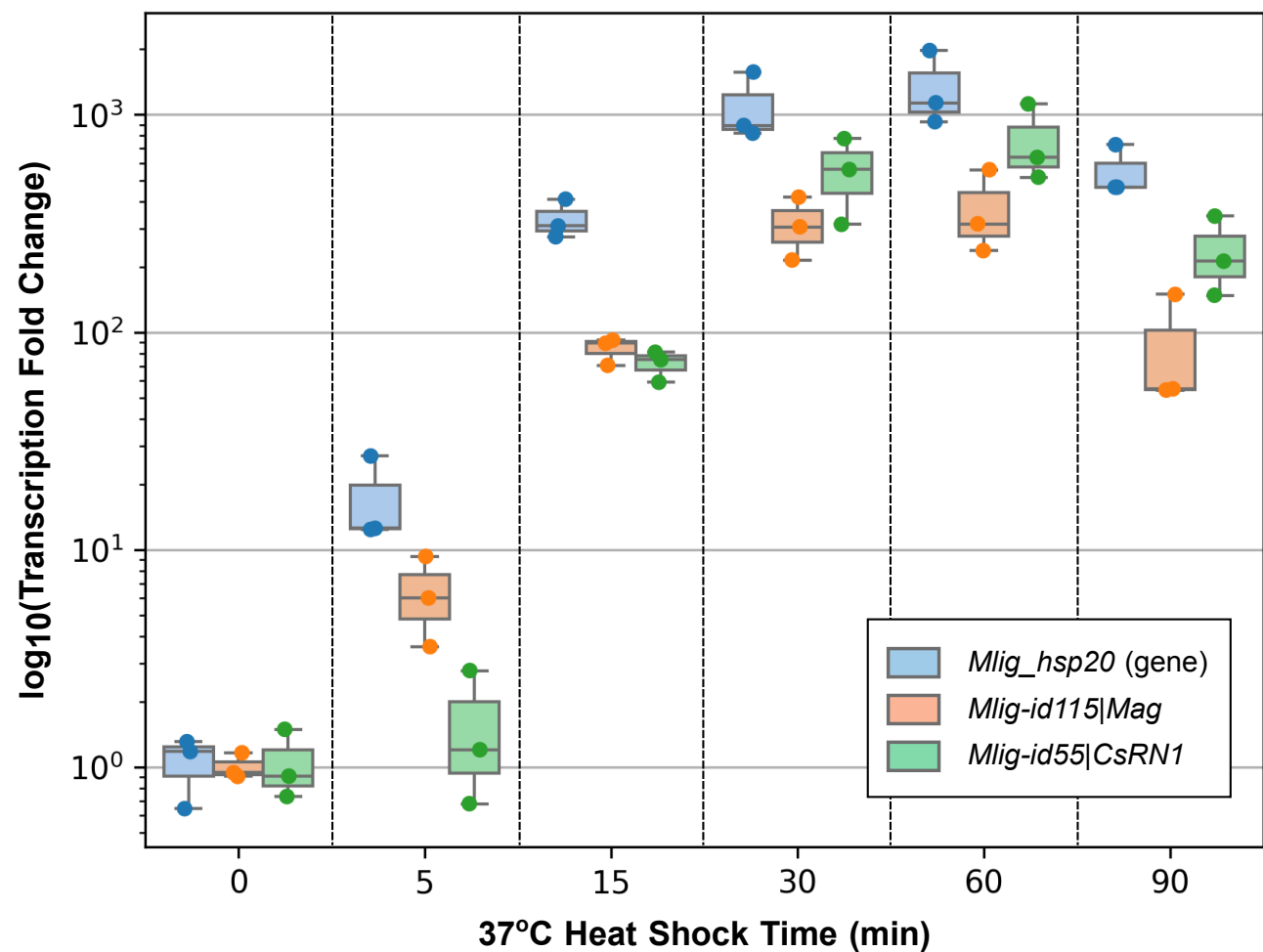
Fig. S1 qRT-PCR analysis of transcriptional activation sensitivity to time of heat stress for two upregulated LTR-RT families. (A) Expression fold changes relative to the control 0 min time point calculated using the $2^{-\Delta\Delta C_t}$ approach. LTR-RT names correspond to those in Figs 1 and 2. The *Mlig_Hsp20* gene was used as a positive control. Expression values between samples were normalized to *Mlig_Gapdh*. (B) Pairwise comparisons between selected time points on log-transformed fold change expression values using ANOVA with post-hoc Tukey test. Adjusted p-values are Benjamini-Hochberg False Discovery Rate (FDR). Significant comparisons (FDR < 0.05) are highlighted in green.

Fig. S2 LTR-RTs upregulated upon heat stress belong to the *Ty3/Mdg4 Mag*, *CsRN1*, and *Gmr1/Osvaldo* clades. Unrooted maximum likelihood tree of phylogenetic relationships among representatives of four major LTR-RTs groups reconstructed based on amino-acid sequences of their RT protein domains. RT domain sequences of *M. lignano* LTR-RT family consensus were used. Branch support values (aLRT/UFBoot) are indicated above and below to the corresponding cluster. Names of the clusters and LTR-RTs are according to GyDB (<http://gydb.org>). Names of *M. lignano* LTR-RT families are as follows Mlig-id[FAMILY NUMBER] | [NUMBER OF CLUSTERED ELEMENTS].

Fig. S3 KU60 plasmid scheme and its GenBank formatted annotation. An annotation of the LTR sequence with Heat Shock Elements (HSEs) of *Mlig-id115|Mag* is provided before the GenBank annotation. The LTR promoter sequence of the *Mlig-id115|Mag* LTR-RT. HSE motifs are in bold and highlighted in yellow. Putative TATA-box is highlighted in blue.

Supplementary Figure 1

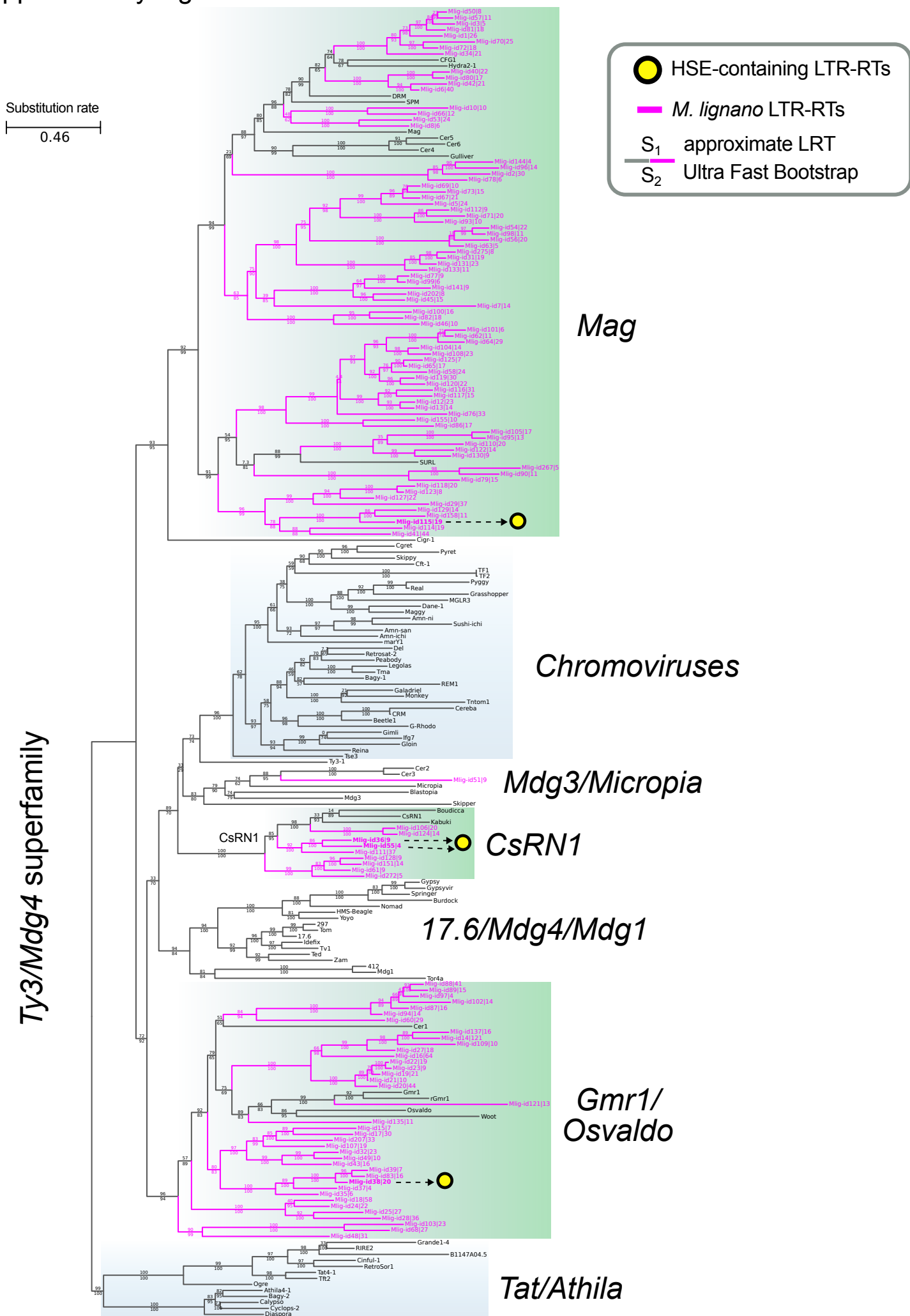
A



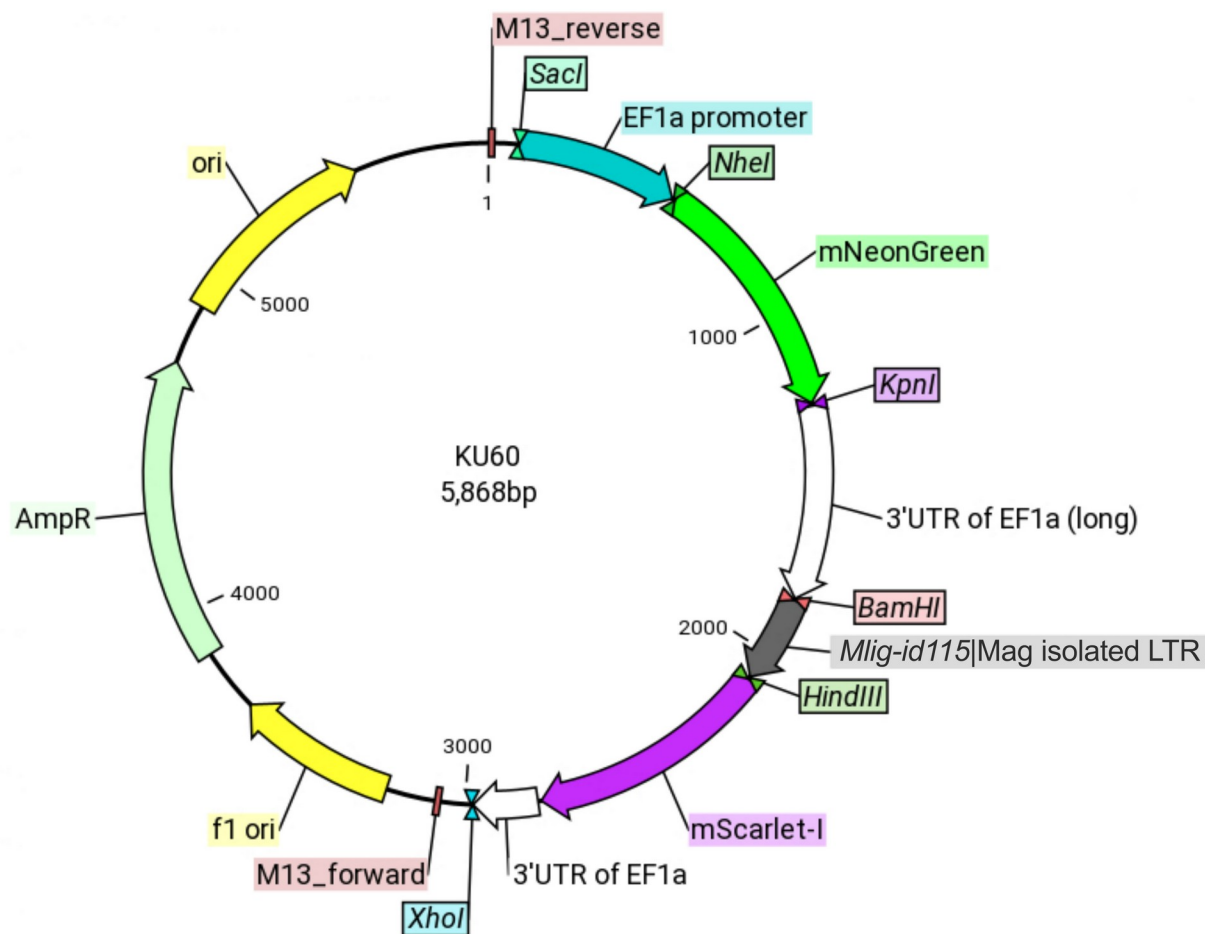
B

HS time (min)		Adjusted p-values (FDR)		
Time 1	Time 2	<i>Mlig_hsp20</i>	<i>Mlig-id115 Mag</i>	<i>Mlig-id55 CsRN1</i>
	5	0	< 0.001	0.0013
	15	5	< 0.001	< 0.001
	30	15	0.0143	0.0017
	60	30	0.9798	0.9381
	90	60	0.0872	0.0638
	90	15	0.5174	0.0776

Supplementary Figure 2



Supplementary Figure 3



5' [BamHI]
TGTTGTGTATTGCAAAGGCATGCGATGCGTAGGGGTAGTTCGGGGACAGCGTTTCTAGAACTTCCACCGCGGACTTGTT
TACGCATAGTTCCAGAAAGTTTCCCTCCCCGGTCTTTCTAGAACTTCCCTCCCCGCGTGCATAAATATACGAGCCGCCCT
CCTATTGGCCAAGCCCCCTCTTTTGGTGCCACTGAGCGCCCGGATAATTATCTATCTTGTCTTCTGCGTTCGTTAATTCAAC
AACAGAACAATG[HindIII] 3'

GENE BANK FORMATTED PLASMID ANNOTATION

LOCUS	KU60	5868 bp	DNA	circular	UNA 17-Dec-2024
DEFINITION					
FEATURES	Location/Qualifiers				
primer	2..18 /label="M13_reverse"				
EF1a	90..551 /label="EF1a promoter"				
mNeonGreen	561..1265 /label="mNeonGreen"				
3'UTR	1275..1827 /label="3'UTR of EF1a (long)"				
LTR_promoter	1834..2082 /label="Mlig_LTR-RT_ID206 Mag promoter"				
mScarlet-I	2092..2784 /label="mScarlet-I"				
3'UTR	2794..2977 /label="3'UTR of EF1a"				
primer	complement(3070..3086)				

ori /label="M13_forward"
3227..3682
/label="f1 ori"
AmpR 3865..4725
/label="AmpR"
ori 4896..5484
/label="ori"

ORIGIN

```
1 ACAGGAAACA GCTATGACCA TGATTACGCC AAGCTATTTA GGTGACACTA TAGAATACTC
61 AAGCTATGCA TCCAACGCGT TGGGAGCTCC GTCTTCTGTT CAGTGTTACA AGTTTTGATG
121 TAAAAAACAA TATTGGCCTA TAACATTTTC ATATATCCTC TCTACATTTT CATATATTTT
181 AAAAAAGTATT TAAAAACAT CTAAAAAGTA TTCCATTTTA TTATTAAGTG AAAAACTAAC
241 GTAAAATTAC AATATATCTC AAAACTTTCT AAAGTCGTAT ATTTTCCGGC ATCGTTAATT
301 TTAACGACAA TCGTGAAAGT TCGTGAACAA GTTTGCTGTT CCTCACCTAA ATTTGAATTA
361 CTATCGCTCT TGTCTGAATG GATTTCTTAC TGCAGATATC CTGAAGTAAG TCTTTCAATT
421 TGTGAATTGT AAGTAAAGTA TTTTTTATTA CTGCTATCAT ATTTGTTGAA TTTGTTCCAA
481 ATTACTATTT CTTGCTTTTA ATTCAAGTTC TTTTATTTAT TCCTAAACTC AGGTAGTTTA
541 ACAACAGCAT TATGGCTAGC GTGAGCAAGG GCGAGGAGGA CAACATGGCC AGCCTGCCGG
601 CCACCCACGA GCTGCACATC TTCGGCAGCA TCAACGGCGT GGAATTCGAC ATGGTGGGCC
661 AGGGCACCGG CAACCCGAAC GACGGCTACG AGGAGCTGAA CCTGAAGAGC ACCAAGGGCG
721 ACCTTCAGTT CAGCCCGTGG ATTCTGGTGC CGCACATCGG CTACGGCTTC CACCAGTACC
781 TGCCGTACCC GGACGGCATG AGCCCGTTCC AGGCCGCTAT GGTGGACGGC AGCGGCTACC
841 AGGTGCACCG CACCATGCAG TTCGAGGACG GCGCCAGCCT GACCGTGAAC TACCGCTACA
901 CCTACGAGGG CAGCCACATC AAGGGCGAGG CCCAGGTGAA GGGCACCGGC TTCCCGGCCG
961 ACGGCCCGGT GATGACCAAC AGCCTGACCG CCGCCGACTG GTGCCGCAGC AAGAAGACCT
1021 ACCCGAACGA CAAGACCATC ATCAGCACCT TCAAGTGGAG CTACACCACC GGCAACGGCA
1081 AGCGCTACCG CAGCACCGCC CGCACACCT ACACCTTCGC CAAGCCGATG GCCGCCAACT
1141 ACCTGAAGAA CCAGCCGATG TACGTGTTCC GCAAGACCGA GCTGAAGCAC AGCAAGACCG
1201 AGCTGAACCT CAAGGAGTGG CAGAAGGCCT TCACCGACGT GATGGGCATG GACGAGCTGT
1261 ACAAGGGTAC CTGAAATGAC CCGCACCTGT GTAATTCTGG AGGCATTGCA GTTAAGTAA
1321 CTATTTGGAA TGCTCGATTG TCGATCAATA AAGAATGCGA AGTTAAAGTT GCGAGACTTG
1381 TTTGTTTGGA AATTCCTGTC TCGATTTGTT TAATTTCTTT GTGTTTCAAT CTATTTGTTG
1441 CTAGTATTTT CTGACTAAAG AGCGACGTAA AGCCCGAACT TCACAAATGT ATGGGCACAA
1501 GTATAGGCTG CATAATGAAG TAGCCAGTTA ATTGCAACGA ATTAACAGTT CATTTATTCA
1561 TATTACAAGG GGCGGGAAC TAAAAAGTTG GAGGCTGTAC TTTTAAATAA CATTGCTTCA
1621 AGCATTACTT GAAGCAGTAT TTACCACAGA CTAAGTGAAT GCAAAGGAAC ACTAACTCGT
1681 TCTGTAAGTC CTTAAATTTG ATGTCTTCTG TTCAATTTCA AACAATAACT TGGAATCTCT
1741 GGGTAAATTT CTGAATCATA CAATTGTCAG TATCAGTACA TGCTAACAACT TGAGTGTACA
1801 TGCCAACTGT GGGCTAGTAT CAGCGAAGGA TCCTGTTGTG TATTGCAAAG GCATGCGATG
1861 CGTAGGGGTA GTTCGGGGAC AGCGTTTCTA GAAACTTTCA CCGCGGACTT GTTTACGCAT
1921 AGTTCCAGAA GTTTCCTTCC CCGGCTTTTC TAGAACTTTT CTTCCCCGCG TGCATAAATA
1981 TACGACCGCG CCTCCTATTG GCCAAGCCCC TCTTTTGGTG CCACTGAGCG CCCGGATAAT
2041 TATCTATCTT GTTCTGCGT TCGTTAATTC AACAACAGAA CAATGAAGCT TGTGAGCAAG
2101 GGCGAGGCCG TGATCAAGGA GTTCATGCGC TTCAAGGTGC ACATGGAGGG CAGCATGAAC
2161 GGCCACGAGT TCGAGATCGA GGGCGAGGGC GAGGGCCGCC CGTACGAGGG CACCCAGACC
2221 GCCAAGCTGA AGGTGACCAA GGGCGGCCCG CTGCCGTTCA GCTGGGACAT CCTGAGCCCG
2281 CAGTTCATGT ACGGCAGCCG CGCCTTCATC AAGCACCCGG CCGACATCCC GGACTACTAC
2341 AAGCAGAGCT TCCCGGAGGG CTTCAGTGG GAGCGCGTGA TGAACCTCGA GGACGGCGGC
2401 GCCGTGACCG TGACCCAGGA CACCAGCCTG GAGGACGGCA CCCTGATCTA CAAGGTGAAG
2461 CTGCGCGGCA CCAACTTCCC GCCGGACGGC CCGGTGATGC AGAAGAAGAC TATGGGCTGG
2521 GAGGCCAGCA CCGAGCGCCT GTACCCGGAG GACGGCGTGC TGAAGGGCGA CATCAAGATG
2581 GCCCTGCGCC TGAAGGACGG CGGTGCTGCT CTGGCCGACT TCAAGACCAC CTACAAGGCC
2641 AAGAAGCCGG TGCAGATGCC GGGCGCCTAC AACGTGGACC GCAAGCTGGA CATCACCAGC
2701 CACAACGAGG ACTACACCGT GGTGGAGCAG TACGAGCGCA GCGAGGGCCG CCACAGCACC
2761 GGCGGCATGG ACGAGCTGTA CAAGTCTAGA TGAGCAATGA CCCGCACCTG TGTAATTCTG
2821 GAGGCATTGC AGTTAACTGA ACTATTTGGA ATGCTCGATT GTCGATCAAT AAAGAATGCG
2881 AAGTTAAAGT TGCGAGACTT GTTTGTTTGG AAATTCCTGT CTCGATTTGT TTAATTTCTT
2941 TGTGTTTCAA TCTATTTGTT GCTAGTATTT CCTGACTCTC GAGATAATCG AATTCCCCGC
3001 GCCGCATGG CCGCCGGGAG CATGCGACGT CGGGCCCAAT TCGCCCTATA GTGAGTCGTA
3061 TTACAATTCA CTGGCCGTG TTTTACAACG TCGTGAATGG GAAAACCTTG GCGTACCGTA
3121 ACTTAATCGC CTTGCAGCAC ATCCCCCTTT CGCCAGCTGG CGTAATAGCG AAGAGGCCCG
3181 CACCGATCGC CCTTCCCAAC AGTTGCGCAG CCTGAATGGC GAATGGACGC GCCCTGTAGC
3241 GGCGCATTA GCGCGGCGGG TGTGGTGGTT ACGCGCAGCG TGACCGCTAC ACTTGCCAGC
3301 GCCCTAGCGC CCGCTCCTTT CGCTTTCTTC CCTTCCTTTC TCGCCACGTT CGCCGGCTTT
```

3361	CCCCGTCAAG	CTCTAAATCG	GGGGCTCCCT	TTAGGGTTCC	GATTTAGTGC	TTTACGGCAC
3421	CTCGACCCCA	AAAAACTTGA	TTAGGGTGAT	GGTTCACGTA	GTGGGCCATC	GCCCTGATAG
3481	ACGGTTTTTC	GCCCTTTGAC	GTTGGAGTCC	ACGTTCTTTA	ATAGTGGACT	CTTGTTCCAA
3541	ACTGGAACAA	CACTCAACCC	TATCTCGGTC	TATTCTTTTG	ATTTATAAGG	GATTTTGCCG
3601	ATTTTCGGCCT	ATTGGTTAAA	AAATGAGCTG	ATTTAACAAA	AATTTAACGC	GAATTTTAAC
3661	AAAAATATTAA	CGCTTACAAT	TTCCCTGATGC	GGTATTTTCT	CCTTACGCAT	CTGTGCGGTA
3721	TTTCACACCG	CATCAGGTGG	CACTTTTTCGG	GGAAATGTGC	GCGGAACCCC	TATTTGTTTA
3781	TTTTTCTAAA	TACATTCAAA	TATGTATCCG	CTCATGAGAC	AATAACCCTG	ATAAATGCTT
3841	CAATAATATT	GAAAAAGGAA	GAGTATGAGT	ATTCAACATT	TCCGTGTGCG	CCTTATTCCC
3901	TTTTTTTGCGG	CATTTTGCCCT	TCCTGTTTTT	GCTCACCCAG	AAACGCTGGT	GAAAGTAAAA
3961	GATGCTGAAG	ATCAGTTGGG	TGCACGAGTG	GGTTACATCG	AACTGGATCT	CAACAGCGGT
4021	AAGATCCTTG	AGAGTTTTTCG	CCCCGAAGAA	CGTTTTTCAA	TGATGAGCAC	TTTTAAAGTT
4081	CTGCTATGTG	GCGCGGTATT	ATCCCGTATT	GACGCCGGGC	AAGAGCAACT	CGGTCGCCGC
4141	ATACACTATT	CTCAGAATGA	CTTGGTTGAG	TACTCACCAG	TCACAGAAAA	GCATCTTACG
4201	GATGGCATGA	CAGTAAGAGA	ATTATGCAGT	GCTGCCATAA	CCATGAGTGA	TAACACTGCG
4261	GCCAACTTAC	TTCTGACAAC	GATCGGAGGA	CCGAAGGAGC	TAACCGCTTT	TTTGCACAAC
4321	ATGGGGGATC	ATGTAACTCG	CCTTGATCGT	TGGGAACCGG	AGCTGAATGA	AGCCATACCA
4381	AACGACGAGC	GTGACACCAC	GATGCCTGTA	GCAATGGCAA	CAACGTTGCG	CAAACATTATTA
4441	ACTGGCGAAC	TACTTACTCT	AGCTTCCCGG	CAACAATTAA	TAGACTGGAT	GGAGGCGGAT
4501	AAAGTTGCAG	GACCACTTCT	GCGCTCGGCC	CTTCCGGCTG	GCTGGTTTAT	TGCTGATAAA
4561	TCTGGAGCCG	GTGAGCGTGG	GTCTCGCGGT	ATCATTGCAG	CACTGGGGCC	AGATGGTAAG
4621	CCCTCCCGTA	TCGTAGTTAT	CTACACGACG	GGGAGTCAGG	CAACTATGGA	TGAACGAAAT
4681	AGACAGATCG	CTGAGATAGG	TGCCTCACTG	ATTAAGCATT	GGTAACTGTC	AGACCAAGTT
4741	TACTCATATA	TACTTTAGAT	TGATTTAAAA	CTTCATTTTT	AATTTAAAAA	GATTTAGGTG
4801	AAGATCCTTT	TTGATAATCT	CATGACCAAA	ATCCCTTAAC	GTGAGTTTTT	GTTCCACTGA
4861	GCGTCAGACC	CCGTAGAAAA	GATCAAAGGA	TCTTCTTGAG	ATCCTTTTTT	TCTGCGCGTA
4921	ATCTGCTGCT	TGCAAACAAA	AAAACCACCG	CTACCAGCGG	TGGTTTGTTT	GCCGGATCAA
4981	GAGCTACCAA	CTCTTTTTTCC	GAAGGTAACT	GGCTTCAGCA	GAGCGCAGAT	ACCAAATACT
5041	GTTCTTCTAG	TGTAGCCGTA	GTTAGGCCAC	CACTTCAAGA	ACTCTGTAGC	ACCGCCTACA
5101	TACCTCGCTC	TGCTAATCCT	GTTACCAGTG	GCTGCTGCCA	GTGGCGATAA	GTCGTGTCTT
5161	ACCGGGTTGG	ACTCAAGACG	ATAGTTACCG	GATAAGGCGC	AGCGGTCGGG	CTGAACGGGG
5221	GGTTCGTGCA	CACAGCCAG	CTTGAGCGA	ACGACCTACA	CCGAACGAG	ATACCTACAG
5281	CGTGAGCTAT	GAGAAAGCGC	CACGCTTCCC	GAAGGGAGAA	AGGCGGACAG	GTATCCGGTA
5341	AGCGGCAGGG	TCGGAACAGG	AGAGCGCACG	AGGGAGCTTC	CAGGGGGAAA	CGCCTGGTAT
5401	CTTTATAGTC	CTGTCGGGTT	TCGCCACCTC	TGACTTGAGC	GTCGATTTTT	GTGATGCTCG
5461	TCAGGGGGGC	GGAGCCTATG	GAAAAACGCC	AGCAACGCGG	CCTTTTTTACG	GTTCTTGCC
5521	TTTTTGCTGGC	CTTTTGCTCA	CATGTTCTTT	CCTGCGTTAT	CCCCTGATTC	TGTGGATAAC
5581	CGTATTACCG	CCTTTGAGTG	AGCTGATACC	GCTCGCCGCA	GCCGAACGAC	CGAGCGCAGC
5641	GAGTCAGTGA	GCGAGGAAGC	GGAAGAGCGC	CCAATACGCA	AACCGCCTCT	CCCCGCGCGT
5701	TGGCCGATTC	ATTAATGCAG	CTGGCACGAC	AGGTTTCCCG	ACTGGAAAGC	GGGCAGTGAG
5761	CGCAACGCAA	TTAATGTGAG	TTAGCTCACT	CATTAGGCAC	CCCAGGCTTT	ACACTTTATG
5821	CTTCCGGCTC	GTATGTTGTG	TGGAATTGTG	AGCGGATAAC	AATTTTAC	

//