### Non-Coding Sequences

Qi — Qiagen

SG — SwitchGear Genomics Si — Sigma-Aldrich Ta — TaKaRa

Th - Thermo Scientific

All -35 and -10 bacterial promoter sequences are predicted by BPROM (http://linuxl.softberry.com/berry.phtml) and some are additionally annotated in public sequences.

Name of	# of	Alignment	Size	# Var	Sources*
Variant	0ccur		(bp)	bp	
AmpR_prom-009	967	${\tt CGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGAGT$	105	0	In,No,Or,Pr, etc
AmpR_prom-001	58	CGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATAAT-TTTQAAAAAGGAAGAGT	104	0/1d	MCSG
AmpR prom-006	32	CGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGA-	103	0/2d	Ag
AmpR_prom-003	20		92	0	No, BR, Or, In
AmpR_prom-005		tttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataacct-gataaatgcttcaat <mark>g</mark> atgaaaaagbaagagt	93	1/1i	No
AmpR_prom-008	10	${\tt CGCGGGC}$ ACCCCTATTTGTTTATTTTTCTAAATAC ${\tt A}$ TTCAAATATGTATCCGCTCATGAGACAAT ${\tt A}$ ACCCT ${\tt -GATAAATGCTTCAAT}$ AAT ${\tt -ATT}$ AAAAAG ${\tt GAGAGGC}$	105	1	Or
AmpR_prom-010	5		92	1	MCSG
AmpR_prom-004	2		72	0	Lu
AmpR_prom-011	2	${\tt CGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAAAAAA$	105	1	Ta
AmpR_prom-002	1	${\tt cgcggaacccctatttgtttatttttctaaataca}$ ${\tt ttcaaatatgtatccgctcat}$ ${\tt gagacaat}$ ${\tt accct}$ ${\tt gataaatgcttcaataat}$ - ${\tt attqaaaaa}$ ${\tt gaaaa}$	106	0/1i	Ag
AmpR_prom-007	1	GTTCAAATATGTATCCGCTCAGGGAGACAATAACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGAGT	72	2	No
AmpR_prom-012	1	${\tt CGCGGAACCCCTATTT}$ ${\tt CTTTATTTTTCTAAATACA}$ ${\tt TTCAAATATGTATCCGCTCATGAGACAATACCCT}$ ${\tt GGCGGAACCCTATTT}$ ${\tt GTTAAATGCTTCAATAAT}$ ${\tt ATTGAAAAAGGAAGAGT}$	105	1	BD
		* ************************************			
		-35 bla TEM prom -10 RBS			
	_				
Key ATG	<b>c</b> = ch	ange in nucleotide relative to consensus sequence — non-coding sequence			
AB - AB Vecto	r				
AG — AddGene					
Ag = Agilent					
	BD - BD Biosciences				
BR — Bio-Rad					
Ca - Cambia					
Cl - Clontech					

Ev - Evrogen GB - Geneva Biotech GE - GE Life Sciences iGEM — international Genetically Engineered Machine Foundation IMAGE - I.M.A.G.E. Consortium In - Invitrogen Iv - InvivoGen Lu - Lucigen
MBL - MBL International MCSG - Midwest Center for Structural Genomics Mo - MoBiTec Ne - New England Biolabs No - Novagen OB - Open Biosystems Or - OriGene Ox - Oxford Genetics Pr — Promega

Name of	# of	Alignment
Variant	0ccur	
CMV_enh-003		CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCA
CMV_enh-015	97+4	GACATTGATTATTGCTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCA
CMV_enh-009	96	GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCA
CMV_enh-008	47	GCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCA
CMV_enh-005	23 22	GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCACGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCA
CMV_enh-010 CMV_enh-011	10	CGTTACATACCTTACGTAAATGCCCCGCTGACCCCCA
CMV_enh-016	7	
CMV_enh-002	4	GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGGGTTCCGCGTTAACATAACTTACGGTAAATGGCCCGCTGGCCGCCA
CMV_enh-006	4	CGTTACATAACTTACGGTAAATGGCCCGCCTGACCGCCCA
CMV_enh-013	3	GCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCCCA
CMV_enh-012	2	CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCCCCA
CMV_enh-004	2	CGTTACATAACTTACGGTAAATGGCCCGCCTGACCGCCCA
CMV_enh-007	2	GCATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCA
CMV_pro-007	2	GGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCA
		******* ***********
CMV_enh-003		ACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
CMV_enh-015		ACGACCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
CMV_enh-009		ACGACCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
CMV_enh-008		${\tt ACGACCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC$
CMV_enh-005		ACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATATCACGGTAAACTGCCCACTTGGCAGTACATC
CMV_enh-010		ACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
CMV_enh-011		ACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
CMV_enh-016		ACGACCCCCCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
CMV_enh-002		ACGACCCCCCCCTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
CMV_enh-006		ACGACCCCCCCCTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGATTTACGGTAAACTGCCCACTTGGCAGTACATC
CMV_enh-013 CMV_enh-012		ACGACCCCCGCCATTGACGTCAATAATGACGATAGTTCCCATAGTAACGCCAATAACGGCCATTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCACTTGGCAGTACATC
CMV_enh-012		ACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC ACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
CMV_enh-007		ACGACCCCCGCCATTGGCGTCAATAATGACGTATGTTCCCATAGTAACAGCCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAATAACTGCCCACTTGGCAGTACATC
CMV pro-007		ACGACCCCCCCCCATTGACGTCAATAATGACGTATGTTTCCATTAGCGACCCAATTAGGGACTTTCCATTGACGTCAATGGGTGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
		***************************************
CMV enh-003		AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC-CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTA
CMV enh-015		AAGTGTATCATATGCCAACTGCCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC-CTGGCATTATGCCCAGTACATGACCTTAGGGACTTTCCTACTTGGCAGTACATCTACGTA
CMV enh-009		AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC-CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTA
CMV enh-008		AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC-CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTA
CMV_enh-005		AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC-CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTA
CMV_enh-010		AAGTGTATCATATGCCAAGTGCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC-CTGGCATTATGCCCAGTACATGACCTTAGGGGACTTTCCTACTTGGCAGTACATCTACGTA
CMV_enh-011		AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC-CTGGCATTATGCCCAGTACATGACCTTACGGACTTTCCTACTTGGCAGTACATCTACGTA
CMV_enh-016		AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC_CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTA
CMV_enh-002		AAGTGTATCATATGCCAAGTACGCCCCTATTGACGTCAATGACGGTAAATGGCCCGCTCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTA
CMV_enh-006		AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC-CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTA
CMV_enh-013		AAGTGPATCATATGCCAAGTEGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC-CTGGCATTATGCCCAGTACATCACCTTAEGGGACTTTCCTACTTGGCAGTACATCTACGTA
CMV_enh-012 CMV enh-004		AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC-CTGGCATTATGCCCAGTACATGACCTTATGGGACTTCCTACTTGGCAGTACATCACGTA AAGTGTATCA
CMV_enh-007		AAGTGTATCATATGCCAAGTGCCCCCCTATTGACGTCAATGACGGTAAATGCCCCGC-CTGGCATTATGCCCAGTACATGACCTTA-GGGACTTTCCTACTTGGCAGTACATTACGCCAGTACATGACCTTA-GGGACTTCCTACTTGGCAGTACATTACGCCAGTACATGACCTTA-GGGACTTCCTACTTGGCAGTACATTACGCCAGTACATGACCTTA-GGGACTTCCTACTTGGCAGTACATTACGCCAGTACATGACCTTA-GGGACTTCCTACTTGGCAGTACATTACGCCAGTACATGACCTTA-GGGACTTCCTACTTGGCAGTACATTACGCCAGTACATGACCTTA-GGGACTTCATTACGTA-GGGACTTCATTACGTA-GGGACTTCATTACGTA-GGGACTTCATTACGTA-GGCAGTACATTACGCCAGTACATTACGCCAGTACATGACCTTA-GGGACTTCATTACGTA-GGCAGTACATTACGCCAGTACATGACCTTA-GGGACTTCATTACGTA-GGCAGTACATTACGCCAGTACATTACATTACGCCAGTACATTA
CMV_pro-007		AAGTGTATCATATGCCAAGTACCCCCCTATTGACGTCAATGACGGTAAATGGCCCCG-CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCACTACGTA
0111_p10 007		***************************************
CMV enh-003		TTAGTCATCGCTATTACCATG
CMV_enh-015		TRAGTCATCGCTATTACCATG
CMV_enh-009		TTAGTCATCGCTATTACCATG
CMV_enh-008		TTAGTCATCGCTATTACCATG
CMV_enh-005		TTAGTCATCGCTATTACCATG
CMV_enh-010		TTAGTCATCGCTATTACCATG
CMV_enh-011		TTAGTCATCGCTATTACCATG
CMV_enh-016		TTAGTCATCGCTATTACCATG
CMV_enh-002		TTAGTCATCGCTATTACCATG
CMV_enh-006		TTAGTCATCGCTATTACCATG
CMV_enh-013		TTAGTCATCGCTATTACCATG
CMV_enh-012		TTAGTCATCGCTATTACCATG
CMV_enh-004		TTAGTCATCGCTATTACCATG TTAGTCATCGCTATTACCATG
CMV_enh-007 CMV_pro-007		TTAGTCATCGCTATTACCATG
511V_P10=007		**************************************
CMV enh_001	is reve	erse complement of CMV_enh-003. CMV_enh-001 was merged with CMV_enh-003
		rise complement of CMV_enh-015. CMV_enh-014 was merged with CMV_enh-005 rise complement of CMV_enh-015. CMV_enh-014 was merged with CMV_enh-015
		ally promoter plus enhancer. When split the enhancer segment is unique among the enhancers.
Key A	IGC = ch	mange in nucleotide relative to consensus sequence — non-coding sequence
2		, , ,

Size (bp) 304

380 517 # Var

0 0/1i

2

**bp** 0 Sources\*

Or,Si In,Cl,Or Ag,Mo In,Ne Pr,No,Or

cı,

Qi Cl

No Pr,Cl Cl Lu

Cl Cl

Cl,Ev,Ox,In

•	`	
Name of	# of	Alignment
Variant	0ccur	
CMV pro-004	338+2	
CMV pro-009	114	
CMV pro-008	19	
CMV pro-010	12	
CMV pro-005	10	
CMV pro-006	10	
CMV pro-002	4	
CMV_pro=007	**	GAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGACCGCCCAACGACCCCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCA
	1	GGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCA
CMV_pro-001	1	
CMV_pro-003	1	
CMV_pro-004		
CMV_pro-009		
CMV pro-008		
CMV pro-010		
CMV pro-005		
CMV pro-006		
CMV pro-002		
CMV pro-007		TTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTA
		TIGACGICAAIGGGIGGAGIAITIACGGIAAACIGCCCACIIGGCAGIACAICAAGIGIATCAIAIGCCCACCIAIIGACGICAAIGACGGIAAAIGGCCCGCCIGCAIIA
CMV_pro-001		
CMV_pro-003		
CMV_pro-004		GTGATGCGGTTTTGGCAGTACA <u>T</u> CAATGGGCGTGGATAGCGGTTTTGA
CMV_pro-009		GTGATGCGGTTTTGGCAGTACAGCGGTTTTGA
CMV_pro-008		TGATGCGGTTTTGGCAGTACATCAATGGCGGTTAGCGGTTTTGA
CMV_pro-010		
CMV pro-005		TGATGCGGTTTTGGCAGTACA CAATGGGCGTGGATAGCGGTTTTGA
CMV pro-006		- INDICACIONAL CANTINGA CANTIN
CMV pro-002		GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATACGGTTTTGA
CMV pro-007		$\tt TGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTTGA$
CMV pro-001		TGATGCGGTTTTGGCAGTACATCAATGGCGTTGA
CMV pro-003		GTGATGCGGTTTTGGCAGTACA CAATGGGCGTGGATACCGGTTTGA
CIIV_PIO-003		************************************
CMV pro-004		CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT
		CTCACGGGGATTTCCAAGTCTCACCCCATTGACGTCAATGGGAGTTTGTTT
CMV_pro-009		
CMV_pro-008		CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT
CMV_pro-010		$\tt CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT$
CMV_pro-005		CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT
CMV_pro-006		$\tt CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT$
CMV_pro-002		$\tt CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT$
CMV_pro-007		CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT
CMV pro-001		$\tt CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT$
CMV_pro-003		$\tt CTCACGGGGATTTCCAAGTCTCACCCCATTGACGTCAATGGGAGTTTGTTT$
		***************************************
CMV pro-004		ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT
CMV pro-009		ATGGCCGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT
CMV pro-008		ATGGCCGTAGCCGTACAGGTGGAAGTCTATATAAGCAG
CMV_pro=010		ATGGCCGTAGCCGTGACGTGGGAGTCTATATAGCAGAGCT
CMV_pro-005		ATGGCGTAGGCGTTACGGTGGGAGGTCTATATAAGCAGA
CMV_pro-006		ATGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT
CMV_pro-002		ATGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT
CMV_pro-007		ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT
CMV_pro-001		ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT
CMV_pro-003		ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT
_		******************

Size (bp) 204

# Var

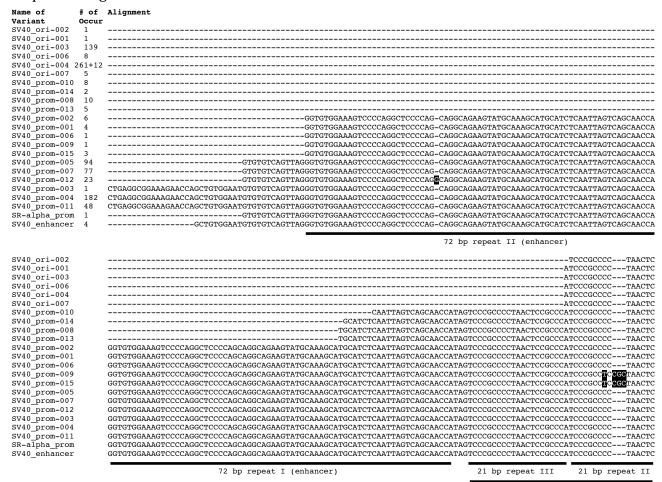
3/8i

Sources\*
C1,In,Ag,Ev
Or,Si,Pr

Or Ox No Cl In Cl Iv Cl

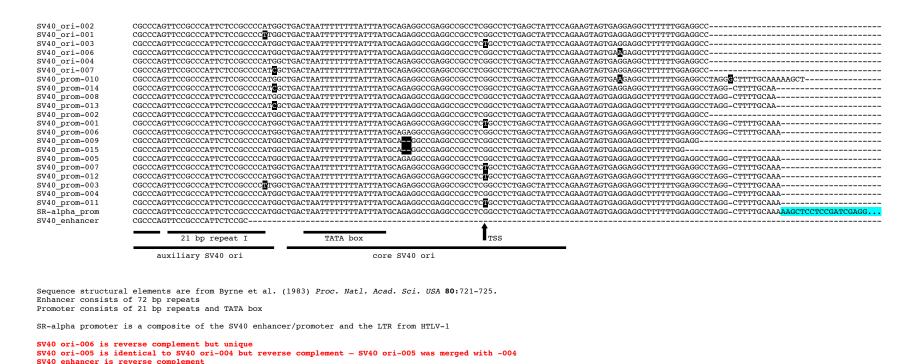
CMV\_pro-007 is actually promoter plus enhancer. When split the promoter segment is identical to CMV\_pro-004.

Key ATGC = change in nucleotide relative to consensus sequence - non-coding sequence



auxiliary SV40 ori

# Var	Sources*
	Or
	Lu
	In, Ne, GE, OB, Lu, etc
	Cl
	Cl,Or,Pr,Ev,Aq, etc
	Pr,Ne
	C1
	Ne
	Or,Cl,Pr
-	Pr
	Si
-	In
-	?
1/2d3i	Si
1	Si
0	Or,Cl
1	In,Th,Cl,GE
1/1i	In,Ne
1	Lu
0	Cl, Aq, Ev, Pr, In, Or
1	Pr,Lu
1	?
0	Pr
	bp 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1



Key ATGC = change in nucleotide relative to consensus sequence — non-coding sequence ATGC = HTLV-1 LTR

Name of	# of	21 company
Name or Variant	# OI Occur	Alignment
fl ori-008	342	ACGCGCCCTGTAGCGGCGCTTTAAGCGCGGGGGGTGTGGTTGGT
f1 ori-007	168	ACGCCCTGTAGCGGCGCGTATAGCGCGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCACGCCCGCTCCTTTCTTCCTTTCTTCCCCCA
f1 ori-009	48	ACGCGCCCTGTAGCGGCGCATTAAGCGCGGGGGGGGGGG
f1 ori-003	34	ACGCGCCCTGTAGCGGCGCATTAAGCGGGGGGTGTGGTGGTTGCTGCGCAGCGTGACCGCTACACTTGCCAGCGCCCCTAGCGCCCCGCTCCTTTCTCTCTTTCCCTTTCCTTTCCCCCA
f1_ori-006	19	${\tt ACGCGCCCTGTAGCGGCGCATTAAGCGCGGGGGGGGGGG$
f1_ori-004	8	ACGCGCCCTGTAGCGCCGCATTAAGCGCGGCGGGTGTGGTTGGT
M13_ori-003		ACGCGCCCTGTAGCGGGCGCATTAAGCGCGGGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCCGCTCCTTTCGCTTTCTTCCCTTTCTCGCCA
f1_ori-011	4	ACGCGCCCTGTAGCGGCGCATTAAGCGCGGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTT
f1_ori-014	4	ACGCGCCCTGTAGCGGCGCATTAAGCGCGGGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCCCTTCCTT
f1_ori-017 f1_ori-005	4 2	ACCCCCCTTAGCGGCCATTAAGCGGCGGCGGTTGGTGGTTACCGCCACGCTCACCGCTACCACGCCCCAGCGCCCTTTCGCTTCTTCTCCCCA
M13 ori-004		ACGGCCCTTTAGGGGGGCATTAAGGGGGGGGTTGGTGGTTAGGGCAGCGTTACGCGTACACTTGCCAGCCCCTTAGGCCCCGCTCCTTTCGCTTTCTCCCCTCCTTTCTCGCA
f1_ori-001	1	ACGGCCCTGTAGCGGCGCATTAAGGCGGGGGGTGTGGTGTTACGGCAGCGTGACCGCTACACTTGCCAGCGCCCTACGGCCCTCTTCTTCTTCCTTC
f1 ori-002	1	ACGCGCCCTGTAGCGGCGCATTAAGCGGGGGGGTGTGGTGGTTGCTGCCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCCGCTCCTTTCGCTTTCTTCCCTGCCA
f1 ori-010	1	ACGCGCCCTGTAGCGGCGCATTAAGCGGGGGGTGTGGTGGTTGCTGCGCAGCGTGACCGCTACACTTGCCAGCGCCCCTAGCGCCCCGCTCCTTTCTCTCTTTCCCTTTCCTTTCCCCCA
f1_ori-012	1	ACGCGCCCTGTAGCGGCGCATTAAGCGCGGGGGTGTGGTTGCTGCTACCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCCGCTCCTTTCGCTTTCTCTCTC
f1_ori-013	1	ACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTTGGT
f1_ori-015	1	ACGCGCCCTGTAGCGGCGCATTAAGCGCGGGGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCCTAGCGCCCCTCCTTTCGCTTTCTTT
f1_ori-016	1	ACGCGCCCTGTAGCGGCGCATTAAGCGCGGGGGGTGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCTTTCTCGCCA
M13_ori-001 M13 ori-002		ACGCGCCCTGTAGCGGCGCATTAAGCGCGGGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTTAGCGCCCCTCCTTTCGCTTTCTCCCTTCCTT
M13_011=002 M13 ori=005		ACCCCCTTAGCGGCGCATTAAGCGCGGCGGTTTGGTGGTTACGCGCACCGTTACCGCTACACTTGCCAGCCCCCTAGCGCCCCGTTCCTTTCGCTTTCTCCCCA
1113_011-003	, -	********* * ******** * ********* * *****
f1 ori-008		$\texttt{CGTTCGCCGGCTTTCQCCGT}^{\textbf{C}} \texttt{AAGC}^{\textbf{T}} \texttt{CTAAATCGGGG}^{\textbf{CT}} \texttt{CCC}^{\textbf{TTTA}} \texttt{GGGTTCCGATTTAG}^{\textbf{T}} \texttt{CCCT}^{\textbf{T}} \texttt{ACGG}^{\textbf{L}} \texttt{ACCTCGACCCCA} \texttt{AAAAACTT}^{\textbf{G}} \texttt{ATTA}^{\textbf{G}} \texttt{GGTGATGGTTCAC}^{\textbf{GTA}} \texttt{GTAGTGGGC}$
f1_ori-007		cgttcgccggctttcqccgt <mark>c</mark> aagc <mark>t</mark> ctaaatcgggg <mark>gct</mark> ccc <mark>ttta</mark> gggttccgatttag <mark>t</mark> gctt <mark>t</mark> acgg¢acctcgaccccaaaaaactt <mark>g</mark> att <mark>a</mark> gggtgatggttcac <mark>gta</mark> gtgggc
f1_ori-009		cgttcgccggctttcqccgt <mark>c</mark> aagc <mark>t</mark> ctaaatcgggg <u>gc</u> tccc <mark>ttta</mark> gggttccgatttag <mark>t</mark> gctt <mark>t</mark> acgg¢acctcgacccca\aaaaactt <mark>g</mark> att <mark>a</mark> gggtgatggttcac <mark>gta</mark> gtgggc
f1_ori-003		cgttcgccggctttcqccgt <mark>c</mark> aagc <mark>t</mark> ctaaatcgggg <mark>ca</mark> tccc <mark>ttta</mark> gggttccgatttag <mark>t</mark> gctt <mark>t</mark> acgg¢acctcgaccccaaaaaactt <mark>g</mark> att <mark>a</mark> gggtgatggtcac <mark>gta</mark> gtgggc
f1_ori-006		cgttcgccggctttcdccgt <mark>c</mark> aagc <mark>t</mark> ctaaatcgggg <mark>gct</mark> ccc <mark>ttta</mark> gggttccgatttag <mark>t</mark> gctt <mark>t</mark> acggfacctcgaccccaaaaaatt <mark>g</mark> att <mark>e</mark> gggtgatggttacgttcac <mark>gta</mark> gtgggc
f1_ori-004		cgttcgccggctttcqccgt <mark>c</mark> aagc <mark>t</mark> ctaaat <b>g</b> gggg <mark>ea</mark> tccg <mark>tttaccc</mark> ttccgattta <mark>g</mark> tgctt <mark>t</mark> acggeacctcgaccccaaaaaactt <mark>g</mark> att <mark>a</mark> ggtgatggtcac <mark>gta</mark> gtgggc
M13_ori-003 f1 ori-011	5	CGTTCGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCCCAAAAAACTTGATTTGGTGATGGTTCACG <mark>TA</mark> GTGGGC CGTTCGCCGGCTTTCCCCGT <mark>C</mark> AAGC <mark>T</mark> CTAAATCGGGG <mark>GCT</mark> CCC <mark>TTTA</mark> GGGTTCCGATTTAGTGCTTTACGGCACCCCAAAAAACTTGATTAGGGTGATGGTTCACG <mark>TA</mark> GTGGGC
f1_ori-014		CGTTCGCCGGCTTTCQCCGTCAAGCCTAAATCGGGGGGTCCCATTTAGGTCCCATTTAGGGCTTTAGGGCACCTCGACCCCAAAAAACTTGATTAGGGTCATGGTCACGTCAGTGGTGGCGCCAACAAAACTTGATTAGGGTCATGGTCACGTAGTGGTGGCGCAACTGAAACTTGATTAGGGTCACGTGTCACGTAGTGGTGGCGCAACTGAAACTTGATTAGGGTCACTGTTCACGTAGTGGTGGACGTCAACTGAACT
f1_ori-017		CGTTCGCCGGCTTTCGCCGTCAAGCTCTAAATCGGGGGCTCCCTTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGC
f1_ori-005		CGTTCGCCGGCTTTCCCCGT <mark>C</mark> AAGC <mark>T</mark> CTAAATCGGGG <mark>GCT</mark> CCC <mark>TTTA</mark> GGGTTCCGATTTAG <mark>T</mark> GCTT <mark>T</mark> ACGGCACCTCGACCCCAAAAAACTT <mark>G</mark> ATT <mark>A</mark> GGGTGATGGTTCAC <mark>GTA</mark> GTGGGC
M13_ori-004	ļ	$\mathtt{CGTTCGCCGGCTTTC} \ \mathtt{CGTC}$ AAGC $\mathbf{TCTAAATCGGGGGCCT}$ CCC $TTTAGGGGTTCCGATTTAGGCCCCTTTACGGCCCCCAAAAAACTTCCCTTCGCTCACCCCAAAAAACCTTCCACTGACCCCCAAAAAACTTCCACTGACCCCCCCAAAAAACTTCCACTGACCCCCAAAAAACTTCCACTGACCCCCAAAAAAACCTTCCACTGACCCCCAAAAAAACCTTCCACTGACCCCCAAAAAAACCTTCCACTGACCCCCAAAAAAACCTTCCACTGACCCCCAAAAAAACCTTCCACCCCCAAAAAAACCTTCCACTGACCCCCAAAAAAACCTTCCACTGACCCCCAAAAAAAA$
f1_ori-001		cgttcgc <mark>e</mark> ggctttcdccgt <mark>c</mark> aagc <mark>t</mark> ctaaatcgggg <mark>gct</mark> ccc <mark>ttta</mark> gggttccgatttag <mark>t</mark> gctt <mark>t</mark> acgg¢acctcgaccccaaaaaactt <mark>g</mark> att <mark>a</mark> gggtgatggtcac <mark>gta</mark> gtgggc
f1_ori-002		CGTTCGC_GGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCCCAAAAAACTTGATTAGGGTGATGGTTCACG <mark>TA</mark> GTGGGC CGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGCC CGTTCGCCGGCTTTCCC_GCTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCCCAAAAA_GCTTGATTAGGGTGATGGTTCACGTA
f1_ori-010		cgttcgccgctttcqc_acc_acc_caraatcgggg <mark>ct</mark> ccc <mark>ttta</mark> gggttccgatttag <mark>t</mark> gctt <mark>t</mark> acggacctcgaccc_caraaa_cttgattaggtgatggttcac <mark>gta</mark> gtgggc
f1_ori-012		cgttcgccggctttcdccgt <mark>c</mark> aagc <mark>t</mark> ctaaatcgggg <mark>gct</mark> ccc <mark>ttta</mark> gggttccgatttag <mark>t</mark> gctt <mark>t</mark> acgglacctcgacc <b>g</b> calaaaactt <mark>g</mark> att <mark>a</mark> ggtgatggtcac <mark>gta</mark> gtgggc
f1_ori-013 f1_ori-015		CGTTCGCCGGCTTTCCCCGT <mark>C</mark> AAGC <mark>T</mark> CTAAATCGGGG <mark>GCT</mark> CCC <mark>TTTA</mark> GGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTT <mark>G</mark> ATTTGGGTGATGGTTCAC <mark>GTA</mark> GTGGGC CGTTCGCCGGCTTTCCCCGT <mark>C</mark> AAGC <mark>T</mark> CTAAATCGGGG <mark>GCT</mark> CCC <u>TTTA</u> GGGTTCCGATTTAG <mark>A</mark> GCTTTACGGCACCTCGACCGCAAAAAACTTGATTTGGGTGATGGTTCACG <mark>TA</mark> GTGGGC
f1_011=013		CGTTCGCCGGCTTTCQCCGTCAAGCCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGTGCC
M13 ori-001		CGTTCGCCGGCTTTCGCCGTCAAGCTCTAAATCGGGGGCTCCCTTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAAACTTGATTTAGGGTAGTTCACGTAGTTGGC
M13 ori-002		CGTTCGCCGGCTTTCCCCGT <mark>C</mark> AAGC <mark>T</mark> CTAAATCGGGG <mark>GCT</mark> CCC <mark>TTTA</mark> GGGTTCCGATTTAGGGCTTTACGGCACCTCGACCGCAAAAAACTT <mark>G</mark> ATT4GGGTGATGGTTCAC <mark>GTA</mark> GTGGGC
M13_ori-005		CGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGG <mark>GCT</mark> CCC <mark>TTTA</mark> GGGTTCCGATTTAG <mark>G</mark> GCTT <mark>T</mark> ACGGCACCTCGACC <mark>G</mark> CAAAAAACTT <mark>G</mark> ATTTGGGTGATGGTTCAC <mark>GTA</mark> GTGGGC CGTTCGCCGGCTTTCCCCGT <mark>C</mark> AAGC <mark>T</mark> CTAAATCGGGG <mark>GCT</mark> CCC <u>TTTA</u> GGGTTCCGATTTAG <mark>A</mark> GCTTTACGGCACCTCGACC <mark>G</mark> CAAAAAACTT <mark>G</mark> ATTTGGGTGATGGTTCAC <mark>GTA</mark> GTGGGC
		****** ***** ****** *** *** *** *** *** ****
		Stem-Loop B T=M13 Stem-Loop C
f1_ori-008		CAMCCOCCUTA DA CA COCCUMUM CA COCCUMUM CA COCUMUM CACCUM COMMUNICA COCCUMUM CACCUM CA
f1_ori-007		CATCGCCC <mark>T</mark> GAT <mark>A</mark> GA <mark>CG</mark> GTTTTTCGCCCTTTGACGTTGGGGTCCACGTT <mark>CTTT</mark> AATAGTGGACTCTTGTTCCAAACTGGAACAACCACACCCTATCTC CATCGCCC <mark>T</mark> GATAGAC <mark>G</mark> GTTTTTCGCCCTTTGACGTTGGGGTCCACGTTCTTTAATAGT
f1 ori-009		CATCGCCCTGATAGACGCTTTTTCGCCCTTTGACGTTGCAGTCCACCTTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTC
f1_ori-003		ር አጥር ርር ርር $-$ ርር ርር ጥጥጥጥ የርር ርር ርጥጥጥር እር ርጥጥር ር እር ርጥር እር ርጥጥ $-$ እጥል ርጥ $$
f1_ori-006		CATCGCCC <mark>T</mark> GAT <mark>A</mark> GA <mark>CG</mark> GTTTTTCGCCCTTTGACGTTGGAGTCCACGTT <mark>CTTT</mark> AATAGTGGACTCCTGTTCTTCCA <mark>AAC</mark> TGGAACAACACCCTATCTC
f1_ori-004		a magaaa <mark>ma ma aa aa ahaa mammadaaa aa magaaaa aa mamma</mark> a ma amaaaaaa aa aa aa aa aa aa aa aa a
M13_ori-003	3	CATCGCCCTGATAGA_GGTTTTTTGCCCTTTGACGTTGGAGTCCTGTTTAATGGT
f1_ori-011		CATCGCCCTGATAGACGGTTTTTPGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGACTCFTGTTCCAGACCTGTACTCAAACTCAACCCTATCTCCATCGCCCTGATAGACGGTTTTTPGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCFTGTTCCAAACTGGAACAALACTCAACCCTATCTC
f1_ori-014		CATCGCCCTGATAGACCGTTTTTPGCCCTTTGACGTTGGGGTCCACGTTGTTTAATAGT
f1_ori-017 f1_ori-005		CATCGCCCTGATBGAGCGTTTTTDGCCCTTGAGGTGCGGTGCTCTTTAATGGT
M13_ori-004	ı	———CATICGOCT <mark>T</mark> GAT <mark>A</mark> GA <mark>CGCTTTTTTTCGCCCTTTTGACGTTCGCACCTTTCTTTTTTTT</mark>
f1 ori-001	•	CAMCCCCCMCAM <mark>ACA CC</mark> CMMMMMCCCCCCMMMCACCCCCACCACCCACCCCAC
f1_ori-002		CATCGCCC <mark>T</mark> GAT <mark>A</mark> GA <mark>CG</mark> GTTTTTCGCCCTTTGACGTTGGAGTCCACGTT <mark>CTTT</mark> AATAGTGGACTCFTGTTCCA <mark>AAC</mark> TGGAACAACACTCAACCCTATCTC
f1_ori-010		categeeetaatagagtttttpgeeetttgaegttggagteeaegttetttaatagt
f1 ori-012		<u></u> categeee <mark>r</mark> gat <mark>a</mark> ga <mark>eg</mark> gttttt bgeeetttgaegttgabgteeaegtt <mark>ettt</mark> aatagt
f1_ori-013		CATGGCCCTGATAGACGGTTTTT GCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGT
f1_ori-015		categeee <mark>t</mark> gat <mark>a</mark> ge <mark>eg</mark> gttttijgeeetttgaggttegagteeaegtttaatagt
f1_ori-016		LATCGCCCTGATAGAGGGTTTTTLGCCCTTTGGACGTTCGAGTTCCACGTTCTTTAATAGT
M13_ori-001		CATCGCCCTGATHGAUGGTTTTTpGCCCTTTGACGTTGGACGCTCTTTAATAGTGACTCTTTTCCAAACTTGGAACAALACTCAACHCTATCTC
M13_ori-002 M13_ori-005		AI-GGCC-GAIRGANGGIIIII-GGCCIIIIGAGIIGGGGGGGGGGGGGGGGGGG
1113_011-003	,	
		Stem-Loop D Stem-Loop E
		Primer RNA (- strand synthesis) Gene II nick site (+ strand initiation)

Size	# Var	Sources*
	-	Boulces.
(bp)	bp	
456	1	Cl,Ag,Or,Ev, etc
456	0	No,In,Cl
429	0	In,Or,Cl
429	4	In,Ne
456	2	BD,Cl,Ta
456	20	In
510	1/1d	7669286
456	1	In
429	2	In,Mo
473	2/17i	Si
456	2	19906724
381	1	Ne
456	2	Pr
456	2	IMAGE
427	0/2d	In
456	2	Pr
459	2/3i	In
456	5	In
459	1/3i	MCSG
456	2	3323803
380	5	Tn

f1 ori-008

f1 ori-007

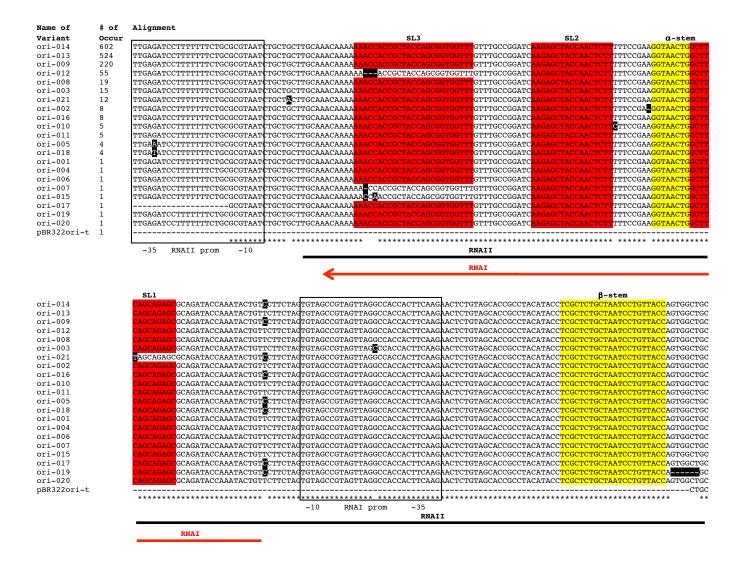
```
f1_ori-009
               f1 ori-003
               f1 ori-006
               {\tt GG}{\tt CTATTCTTTGATTTATAAGGGATTTTGCC}{\tt GATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAT---ATTAACGTTTACAATTT-
               \operatorname{\mathsf{GGT}}CTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGGCTATTGCTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAAT---ATTAACGTTTACAATTT-
f1 ori-004
               M13 ori-003
f1_ori-011
               \mathsf{GGT}^\mathsf{TCTATTCTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAA<math>\mathbf{H}ATTTAACGCGAATTTTAACAAAAT---ATTAACGTTTACAATT-
f1_ori-014
               f1 ori-017
               {\tt GGTCTATTCTTTTGATTTAAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAT---ATTAACGGTTACAATT--
f1_ori-005
               \mathsf{GGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAA<math>\mathsf{GT}
               M13 ori-004
f1 ori-001
               \mathsf{GGT}^\mathsf{TTATTCTTTTGATTTAAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAAT---ATTAACGCTTACAATTT-
f1 ori-002
               \mathsf{GGTCTATTCTTTTGATTTAAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAT---ATTAACGTTTACAATTT-
               f1 ori-010
               \operatorname{\mathsf{GGTCTATTCTTTTGATTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATTGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAAT---ATTAACGCTTACAATTT-
f1_ori-012
f1 ori-013
               \texttt{GG} \underline{\textbf{C}} \texttt{CTATTCTTTGATTTAAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAT---ATTAACGTTTACAATTT-
f1 ori-015
               \mathsf{GGT}^\mathsf{TCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAA<math>\mathsf{TA}ATTTAACGCGAATTTTAACAAAAT---ATTAACGTTTACAATTT-
f1 ori-016
               {\tt GGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAT{\tt CAC}{\tt ATTAACGCTTACAATTT-
M13_ori-001
               M13_ori-002
               M13_ori-005
               GGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGG------
                                     Domain B (AT-rich enhancer of + strand synthesis)
f1 ori-008
f1_ori-007
f1_ori-009
f1 ori-003
f1 ori-006
f1 ori-004
M13_ori-003
               AATATTTGCTTATACAATCTTCCTGTTTTTTGGGGCTTTTCTGATTATCAACCGGGGT
f1_ori-011
f1 ori-014
f1 ori-017
f1 ori-005
M13 ori-004
f1 ori-001
f1_ori-002
f1_ori-010
f1_ori-012
f1 ori-013
f1 ori-015
fl ori-016
M13_ori-001
M13_ori-002
M13 ori-005
Structures are from Dotto et al. (1984) J. Mol. Biol. 172:507-521 and from Baas (1985) Biochimica et Biophysica Acta 825:11-139.
Yellow highlights indicate nucleotides that form bulges or loops in the indicated stem-loop structures formed when the DNA is single-stranded.
fl ori variants -006, -013 and -015 are actually M13 ori as the two origins differ only by an A or T at the indicated position.
```

If the sequences of variants -013 and -017 are correct, these origins are almost certainly compromised or devoid of minus strand and plus strand synthesis, respectively.

ATGC = change in nucleotide relative to consensus sequence - non-coding sequence

 ${\tt GGTCTATTCTTTTGATTTAAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAT---ATTAACGCTTACAATTT-$ 

 ${\tt GGTCTATTCTTTTGATTTAAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAAT---ATTAACGTTTACAATTT-$ 



Size	# Var	Sources*
(bp)	bp	
589	1	No,Cl,In,Ca,Ev, etc
589	0	Cl, In, Or, Pr, Ox, etc
589	3	In,Qi,MBL,Ne, etc
585	0/4d	Or,Cl,In,GE
589	2	No,Qi,GE
589	1	Th
589	4	24050148
588	1/1d	Lu
589	4	2659436
589	1	10890530
588	0/1d	Qi,No
589	2	Pr
589	4	15644173
589	1	Or
589	2	IMAGE
589	1	iGEM
588	0/1d	Or
589	2	In
569	3	In
583	2/6d	No
589	1	Cl
353	0	7655517

ori-014

ori-013

ori-009 ori-012

ori-008

ori-003 ori-021

ori-002

ori-016

ori-010

ori-011

ori-005 ori-018

ori-001

ori-004

ori-006

ori-007

ori-015

ori-017

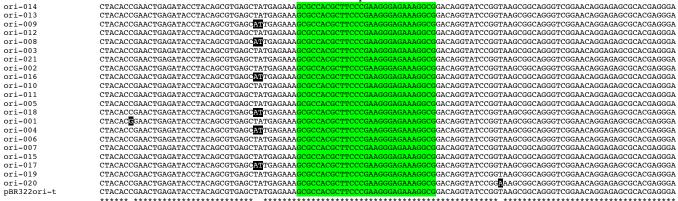
ori-019

ori-020 pBR322ori-t

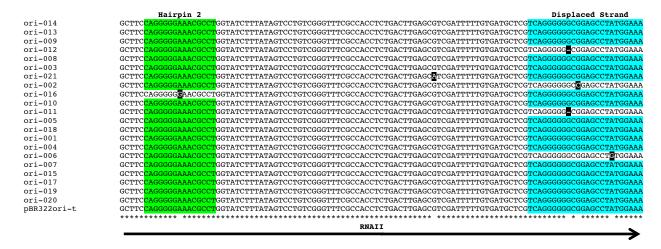
#### V\_S+61

 $\tt TGCCAGTGGCGATAAGTCGTGTCTT{\color{red}{aCCGGGTTGGACTCAAGACGAT}{AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC}$ TGCCAGTGGCGATAAGTCGTGTCTT<mark>ACCGGGTTGGACTCAAGACGAT</mark>AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC TGCCAGTGGCGATAAGTCGTGTCTT<mark>ACCGGGTTGGACTCAAGACGAT</mark>AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC TGCCAGTGGCGATAAGTCGTGTCTT<mark>ACCGGGTTGGACTCAAGACGAT</mark>AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC TGCCAGTGGCGATAAGTCGTGTCTT<mark>ACCGGGTTGGACTCAAGACGAT</mark>AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC TGCCAGTGGCGATAAGTCGTGTCTT<mark>ACCGGGTTGGACTCAAGACGAT</mark>AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC TGCCAGTGGCGATAAGTCGTGTCTT<mark>ACCGGGTTGGACTCAAGACGAT</mark>AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC TGCCAGTGGCGATAAGTCGTGTCTT<mark>ACCGGGTTGGACTCAAGACGAT</mark>AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC TGCCAGTGGCGATAAGTCGTGTCTT<mark>ACCGGGTTGGACTCAAGACGAT</mark>AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC  ${\tt TGCCAGTGGCGATAAGTCGTGTCTT}$  ${\tt TGCCAGTGGCGATAAGTCGTGTCTT}$ TGCCAGTGGCGATAAGTCGTGTCTT<mark>ACCGGGTTGGACTCAAGACGAT</mark>AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC TGCCAGTGGCGATAAGTCGTGTCTT<mark>ACCGGGTTGGACTCAAGACGAT</mark>AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC TGCCAGTGGCGATAAGTCGTGTCTT<mark>ACCGGGTTGGACTCAAGACGAT</mark>AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC  ${\tt TGCCAGTGGCGATAAGTCGTGTCTT}$  ${\tt TGCCAGTGGCGATAAGTCGTGTCTT}$ TGCCAGTGGCGATAAGTCGTGTCTT<mark>ACCGGGTTGGACTCAAGACGAT</mark>AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC  ${ t tgccagtggcgataagtcgtgtcttacceggtttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagctttggagcgaacgac$ 

#### Hairpin 1



RNAII



Structural elements from Selzer et al. (1983) Cell 32:119-129. These ori sequences are ColE1/pMB1 derived origins of replication. RNAII (black arrow) is a transcript that hybridizes with single-stranded DNA in the displaced strand (cyan) region to form a "D-loop" and is processed by RNAse H to form an RNA primer for DNA pol. RNAII is transcribed from a promoter at the extreme 5' end of the sequence.

RNAI is an antisense RNA transcribed from RNAI promoter. RNAI and RNAII form small stem loop structures (SL1-3; red) that can interact and facilitate full sense-antisense hybridization. RNAI-RNAII covers the  $\alpha$ -stem region (yellow; SL1 overlaps), leaving the free region of RNAII to form a large  $\beta$ - $\gamma$  stem loop, which inhibits D-loop formation and replication. Rop protein, expressed from a separate locus not part of ori, stabilizes RNAI-RNAII, preventing replication.

Free RNAII forms an  $\alpha$ - $\beta$  stem loop and small hairpin structures (green) that facilitate D-loop formation and replication.

Plasmids with wild-type ColE1/pMB1 origins and Rop exist at ~15-20 copies/cell. An example of this type of medium copy number plasmid is pBR322.

High copy number (up to 700/cell) ColE1/pMB1 plasmids, such as the pUC and pET series, lack Rop and also carry a pMB1 derivative in which the A at -1 from the start of the RNAI transcript is changed to G (T-C in the above orientation). This mutation moves the transcription start site +3, which removes the "anti-tail" portion of RNAI, which is crucial to nucleate RNAI-RNAII hybrid formation.

The apparent mutations in ori-007, -012 and -015 in SL3 could destabilize stem-loop formation and reduce the ability of RNAI to anneal to RNAII, resulting in higher copy numbers than pBR322.

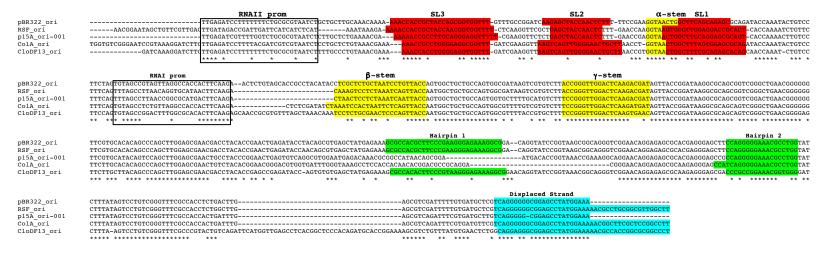
The truncated pBR322 ori is non-functional. The plasmid carrying this (pBINPLUS) uses oriV instead for its replication. pMB1/ColE1 origins can be truncated up to the  $\beta$ -stem, but a strong promoter must be placed in front to make the origin functional.

ori-017 lacks the RNAII promoter. Unless there is a cryptic promoter in SV40 ori immediately 5' to the sequence shown, this ori should be non-functional. BPROM identifies a potential -35 and -10 promoter about 96 bp 5' to ori within SV40 ori.

Key AWGC = change in nucleotide relative to consensus sequence - non-coding sequence

Key

ColE1-Related Promoters using a similar RNA based replication initiation mechanism



Colour scheme for structural elements same as for pMB1/ColE1 ori alignment. From Selzer et al. (1983) Cell 32:119-129.

Plasmids using p15A, ColA, RSF1030 or CloDF13 origins can be co-transformed into cells with plasmids carrying pMB1/ColE1 origins because the RNA sequences are sufficiently different to prevent competition between replication origins. Two plasmid systems are useful for bacterial co-expression of proteins, especially for purification of protein complexes that are difficult to express as individual proteins.

Three plasmid co-expression can be accomplished with a third plasmid that uses an iteron-based replicon. Iterons are repeat sequences to which a DNA-binding protein binds, bending the origin and facilitating its melting and entry of the replication machinery. These plasmids must carry the gene for the DNA-binding protein to express some of these proteins. Alternatively, some E. coli strains are engineered to express an iteron binding protein, and some can do so conditionally to control plasmid copy number.

Examples of iteron origins are ori2/ori5 from F plasmid (uses repE protein), oriV from RK2/IncP- $\alpha$  plasmid (uses trfA protein), oriV from pSC101 plasmid (uses rep101/repA protein), oriV from R6K plasmid (uses pi protein), and oriV from pSa plasmid (uses repA protein).

Some plasmids, such as bacterial artificial chromosomes (BACs) based on the F plasmid ori2/repE replicon, are very low copy number (1-2/cell) to reduce toxic effects of cloned genes or recombination in the insert. These plasmids require partitioning mechanisms (SopA/B/C) to ensure the plasmid is divided between daughter cells. Some plasmids carry two compatible promoters and use conditional expression of iteron proteins to increase copy number when plasmid isolation is required.

The two annotated ori2 variants in the features list are identical but reverse complements oriV-003 and 004 are identical but reverse complements The two annotated pSC101 ori variants are identical but reverse complements

TRES-014

IRES-001

IRES-013

IRES-012

TRES-004

IRES-008

IRES-015

IRES-010

TRES-002

IRES-005

IRES-003

IRES-006

TRES-011

IRES2-001

IRES2-002

IRES2-003

IRES-014

IRES-001

TRES-013

IRES-012

IRES-004

IRES-008

TRES-015

IRES-010

IRES-002

IRES-005

IRES-003

IRES-006 IRES-011

IRES2-001

TRES2-002

IRES2-003

Name of	# of	Alignment
Variant	0ccur	
IRES-014	4	
IRES-001	4	<mark>gtpatttccadcatatgccgtctttggcaatgt</mark> gAggccc
IRES-013	2	GTPATTTTCCACCATATTTGCCAATGTTAGCAATGTGAGGCCC
IRES-012	1	gtpatttccadcatattgccgtcttttgcaatgtgaggccc
IRES-004	2	
IRES-008	1+1	
IRES-015	3	
IRES-010	2	CCCCTCTCCCCCCCCCCC-TAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCGTTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCC
IRES-002	4	${\tt cccctctcccccccc-taacg}$
IRES-005	3	CCCCTCTCCCCCCCCCCC-TAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCGTTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCC
IRES-003	35	CCCCTCTCCCCCCCCCCC-TAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCGTTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCC
IRES-006	1	CCCCTCTCCCCCCCCCC-TAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGFGCGTTTGTCTATATGTFATTTTCCACCATATTTGCCGTCTTTTGGCAATGTGAGGGCCC
IRES-011	2	CCCCTCTCCCCCCCCCCCCCCTTAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCTTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCC
IRES2-001	3	CCCCTCTCCCCCCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGTCTATTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCC
IRES2-002	11	CCCCTCTCCCCCCCCCC-TAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGFGCGTTTGTCTATATGTFATTTTCCACCATATTTGCCGTCTTTTGGCAATGTGAGGGCCC
IRES2-003	4	CCCCTCTCCCCCCCCCCC-TAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCGTTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCC
		Stem-Loop D Stem-Loop E Stem-Loop F

AAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCC AAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCC AAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCC AAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCC AAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCC 9.00AAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCC AAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCC AAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCC AAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCC

Size	# Var	Sources*
(bp)	bp	
464	1/1i	No
463	2	No
499	1/1d	15232106
485	1/1d	No
551	2/1d	Ox
553	0	Or
552	1	Or
569	0	OB
573	1/1d	Cl
573	1d	Cl
574	0	Cl,Or
574	1	Cl
575	1/1i	Cl
588	1/1i	Cl
587	1	Cl
587	0	Cl,Or

	Stem-Loop K A Bulge Stem-Loop J	Stem-Loop L 834	
IRES2-003	TGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAACGTCTAGGCCCCCCCAA		:AACC <mark>ATG</mark>
IRES2-002	TACACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAACGTCTAGGCCCCCCCAA		
IRES2-001	TACACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAACGTCTAGGCCCCCCCAA		
IRES-011	TACACATGCTTTACATGTGTTTAGTCGAGGT <mark>TAAAAAAA</mark> CGTCTAGGCCCCCCCAA		
IRES-006	TACACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAACGTCTAGGCCCCCCCAA		
IRES-003	TGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAACGTCTAGGCCCCCCCAA		
IRES-005	TGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAA CGTCTAGGCCCCCCAA		
IRES-002	TGCACATGCTTTACATGTGTTTAGTCGAGGT <mark>TAAAAAA</mark> CGTCTAGGCCCCCCCAA		
IRES-010	TGCACATGCTTTACATGTGTTTAGTCGAGGT <mark>TAAAAAA</mark> CGTCTAGGCCCCCCCAA		
IRES-015	TGCACATGCTTTACATGTGTTTAGTCGAGGT <mark>TAAAAAAA</mark> CGTCTAGGCCCCCCCAA		
IRES-008	TGCACATGCTTTACATGTGTTTAGTCGAGGT <mark>TAAAAAAA</mark> CGTCTAGGCCCCCCCAA		
IRES-004	TGCACATGCTTTTCATGTGTTTAGTCGAGGTTAAAAAA CGTCTAGGCCCCCCCAA		
IRES-012	TGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAA-CGTCTAGGCCCCCCCAA		
IRES-013	TGCACATGCTTTACATGTGTTTAGTCGAGGT <mark>TAAAAAA</mark> CGTCTAGGCCCCCCCAA		
IRES-001	TGCACATGCTTTACATGTGTTTAGTCGAGGT <mark>TAAAAAAA</mark> CGTCTAGGCCCCCCCAA		
IRES-014	TGCACATGCTTTACATGTGTTTAGTCGAGGT <mark>TAAAAAAA</mark> CGTCTAGGCCCCCCCAA		
	Stem-Loop I	Stem-Loop J	Stem-Loop K
IRES2-003	$\textbf{ACGTTGTGAGTTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGAGTCAAATGGCTCTCTCAAGCGTATTGGAAAGAGAGTCAAATGGCTCTCTCAAGCGTATTGGAAAGAGAGTCAAATGGCTCTCTCAAGCGTATTGGAAAGAGAGAG$	CAACA <mark>AGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGA-</mark> TCT	GATCTGGGGCCTCGG
IRES2-002	${\tt ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGATAGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCTCCTCAAGCGTATTGGAAAGAGTCAAATGGCTCTCTCCTCAAGCGTATTGGAAAGAGAGTCAAATGGCTCTCTCAAGCGTATTGGAAAGAGAGTCAAATGGCTCTCTCCTCAAGCGTATTGGAAAGAGAGTCAAATGGCTCTCTCT$	caaca <mark>a</mark> gggctgaaggatgcccagaaggtaccccattgtatggga-tct	GATCTGGGGCCTCGG
IRES2-001	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTC	caacaagggctgaaggatgcccagaaggtaccccattgtatggga-tct	GATCTGGGGCCTCGG
IRES-011	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTC	caacaagggctgaaggatgcccagaaggtaccccattgtatggga-tct	GATCTGGGGCCTCGG
IRES-006	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTC	caacaagggctgaaggatgcccagaaggtaccccattgtatggga-tct	GATCTGGGGCCTCGG
IRES-003	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATT	caacaagggctgaaggatgcccagaaggtaccccattgtatggga-tct	GATCTGGGGCCTCGG
IRES-005	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTC		
IRES-002	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCACCTCAAGCGTATTC	caacaagggctgaaggatgcccagaaggtaccccattgtatggga-tct	GATCTGGGGCCTCGG
IRES-010	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTC	CAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGA-TCT	GATCTGGGGCCTCGG
IRES-015	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTG	caacaagggctgaaggatgcccagaaggtaccccattgtatggga-tct	GATCTGGGGCCTCGG
IRES-008	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTC	CAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGA-TCT	GATCTGGGGCCTCGG
IRES-004	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCCCCTCAAGCGTATTC		
IRES-012	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCACCTCAAGCGTATTC	caacaagggctgaaggatgcccagaaggtaccccattgtatggga-tct	GATCTGGGGCCTCGG
IRES-013	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCACCTCAAGCGTATTC		
IRES-001	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCACCTCAAGCGTATTC		
IRES-014	ACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCACCTCAAGCGTATTC	CAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGAATCT	GATCTGGGGCCTCGG

#### Red highlighting indicates the IRES borders were mis-annotated. The sequence actually includes the highlighted region.

IRES-009 is derived from foot and mouth disease virus (FMDV) and does not align well with other IRES variants derived from encephalomyocarditis virus (EMCV). However both EMCV and FMDV are picornaviruses and utilize Group III IRES, which initiate translation directly at the START codon without ribosome scanning. This means that the placement of the START codon is important for translation efficiency. The native START codon is highlighted in green. An alternative START codon is highlighted in cyan.

The 5' portion of IRES is dispensable for function up to stem-loop H (not shown), but interference of stem-loop H formation can occur in constructs in which the first cistron is too close. Therefore, it is preferable to include the 5' stem-loops as spacers.

IRES-002, -004, -005, -012 and -013 have wild-type number of A (A6) in the oligo(A) bulge at the K-J border. This bulge is important for interaction with eIF4G. A7 (e.g. in pIRES) has reduced translational activity compared to A6 (Bockhov and Palmenberg. 2006. Biotechniques 41:283-292). Because A6 and A7 variants have slightly different interactions with the translation machinery, they show different activities in different cell types.

The position of the START codon of the downstream ORF is also important for translational efficiency. Translation of the polypeptide from the native START codon at position 834 (native IRES numbering) provides optimal spacing. Including the first few amino acids of native IRES (IRES2 variants) allows the leader polypeptide to interact properly with the ribosome for the most efficient translation. Using the IRES START codon at position 826 reduces translational efficiency due to steric constraints on this leader peptide ribosome interaction. Moreover, placing the start codon too far downstream in a MCS reduces translation efficiency as the ribosome does not scan in Group III IRESs.

Structures are from Kaminski and Jackson (1998) RNA 4:626-638. Invariant stem-loops are not shown.

IRES-007 is identical to IRES-008 but reverse complement - IRES-007 was merged with -008

Key ANGC = change in nucleotide relative to consensus sequence - non-coding sequence

#### Un-Annotated Promoters

Key

Alignments do not include all plasmids carrying the un-annotated promoter

Un-annotated AmpR promoters (colored segments = vector backbone)

#### Name of Plasmid Alignment Or Feature AmpR\_prom-009 CGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGGAAmpR\_prom-001 CGCGGAACCCCTATTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATAAT-TTTGAAAAAAGGAAGAGT cgcggaacccctatttgtttatttttttaaatacattcaaatatgtatccgctcatgagacaataaccct-gataaatgcttcaataat-attqaaaaagaaga AmpR prom-006 AmpR\_prom-008 ${\tt CGCGGGACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATAAT-ATT<math>{\tt CAAAAAAGGAAGAGT}$ $\operatorname{\mathsf{cgcggaacccctatttgtttattttttttaaataca}$ $\operatorname{\mathsf{ttcaaatatgtatccgctcatgagacaataaccct}}$ accct $\operatorname{\mathsf{cgataaatgcttcaataat}}$ a $\operatorname{\mathsf{cattgaaaaa}}$ aagagt AmpR prom-011 AmpR prom-002 $\operatorname{\mathsf{cgcggaacccctatttgtttatttttttaaataca}$ ttcaaatatgtatccgctcatgagacaataaccct $\overline{\mathbf{u}}$ gataaatgcttcaataat-att $\overline{\mathbf{u}}$ aaaaa $\overline{\mathbf{u}}$ gaagact AmpR\_prom-012 cgcggaacccctatttetttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccct-gataaatgcttcaataat-attqaaaaaggaagagt pCMV-Cypridina Luc ACACTGACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGAAT pTK-Cypridina ACACTGACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGAAT ACACTGACCCTATTTGTTTATTTTTCTAAATACATTCAAATAIGTATCGCCTCATGAGCAATAACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAGGAT TGGTCATGACCTATTGTTTATTTTTCTAAATACATTCAAATAIGTATCGCCTCATGAGCAATAACCCT-GATAAATGCTTCAATATGATTGAAAAAGGAGGT TGGTCATGACCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATGATTGAATAAAGGAGGT TGGTCATGACCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATGATTGAATAAAGGAAGGT TGGTCATGACCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATGATTCAAAAAGGAAGAGT TGGTCATGACCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATGATTCAAATAAGGAAGAGGT TGGTCATGACCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATGATTCAATAAAAAGGAAGAGT TCGTCATGACCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT-GATAAATGCTTCAATGATTCAATAAAAAGGAAGAGT pET-43,1a(+) pET-44a(+) pET-45b(+) pET-51b(+) pET-52b(+) AmpR\_prom-005 AmpR\_prom-003 -----tttgtttatttttttaaatacattcaaatatgtatccgctcatgagacaataacct-gataaatgcttcaataat-attqaaaaaggaagagt AmpR prom-010 ------tttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataacct-gataaatgcttcaataat-aatgaaaaagaagg AmpR prom-004 AmpR\_prom-007 -----GTTCAAATATGTATCCGCTCAGGAGAATAACCCT-GATAAATGCTTCAATAAT-ATTQAAAAAGGAAGAGT pCR4-TOPO ${ t tctacggggtctgacgctcagtggaacgaaaact} { t dacgttaagggattttggtca} { t tgagacaat} { t accct-gataaatgcttcaataat-att} { t aaaaaggaagagt}$ pCAS-Guide GCCCGTGTCTCAAAATCTCTGATG<mark>TTACAT</mark>TGCAQAAGATAAAATAA<mark>TATCAT</mark>CATGAGACAATAACCCT-GATAAATGCTTCAATAAT-ATTQAAAAAGGAAGAGT pCMV SPORT raccct-gataaatgcttcaataat-attqaaaaaggaagagt AATTATGTGCTGTGTAGGGATCGCTGGTATCAAATATGTGTGCCCACCCCTGGCATGAGACAA AACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGAGT pCMV SPORT2 AACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGAGT pZL1 pTriEx-1.1 CATGAGACAA AACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGAGT pTriEx-2 CATGA<mark>GACAA'</mark> AACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGAGT pQE-TriSystem 5 catga<mark>gacaat</mark>aaccct-gataaatgcttcaataat-attqaaaaaggaagagt GCGCGTTTATACACATCTTGGGATTTGATTAAAA ATGCAGGAAACGCGC pQE-TriSystem 6 GCGCGTTTATACACATCTTGGGATTTGATTAAAA CATGA<mark>GACAA</mark>: AACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGAGT AACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGAGT pDream2.1\_MCS CATGAGACAA pOE-TriSystem CATGAGACAA! AACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGAGT pQE-TriSys\_His-Str <mark>'GCAG</mark>GAAACGCGCGGACATGA<mark>GACAA</mark>' AACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGAAGAGT pGL4.70-hRluc raaccct-gataaatgcttcaataat-attqaaaaaggacgagt pLightSwitch Prom -GCGGCCTGTCCAATACCTCCCGTACCTTAATATT ACTTACTTATCCTTGAGAGACGTACTAC PAACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGACGAGT pNT-1.1-Nluc AACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGACGAGT pNL2.1-Nluc pGL4.10-luc2 AACCCT-GATAAATGCTTCAATAAT-ATTGAAAAAGGACGAGT Lafmid BA CTTATACAATCTTCCTGTTTTTTGGGGCTTTTTCTGATTATCAACCGGGGTACATATGATTGACAT GCTAGTTTTACGATTACCGTTCATCGATTGAAAAAGGAAGAGT AAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCC AACTTGATTAGGGTGATGGTTCACATTGAAAAAGGAAGAGT pLIC-SGC1 pIB\_V5-His-DEST TCAGTTATTACCCATTGAAAAAGGAAGAGT pIB\_V5-His CATTGAAAAAGGAAGAGT pMIB\_V5-His\_A CATTGAAAAAGGAAGAGT pYC6\_CT CATTGAAAAAGGAAGAGT CATTGAAAAAGGAAGAGT pYES6\_CT pYC2 CT CATTGAAAAAGGAAGAGT pYES2 CT CATTGAAAAAGGAAGAGT pCR2.1-TOPO -ATTGAAAAAGGAAGAGT -35 bla TEM prom

Sequences were trimmed at the maximum length shown or at the junction of an adjacent feature (producing the shorter sequences). The boxed -35 and -10 regions correspond to those of native bla TEM from E. coli Tn3.

pCAS-Guide uses the KanR (aph(3')-Ia) promoter from Tn5 (brown region; -35 and -10 sequences highlighted in yellow).

The violet region for pCMV-SPORT, pCMV-SPORT2 and pZL1 are part of a defunct incA (RepA binding sites) region, but BPROM promoter prediction algorithm suggests ATGTGT forms a viable -35 region with the intact -10.

The pink region for the pTriEx and pQE-TriSystem series are part of baculovirus recombination region (lef2), but BPROM suggests ATGCAG forms a viable -35 region with the intact -10. The light green region for the luciferase vectors appears to be an artificial promoter (there are several synthetic sequences in these plasmids), but BPROM does not predict a -35 and -10. The blue region in Lafmid\_BA corresponds to the 3'end of M13 ori-003 and is predicted by BPROM to form -35 and -10 (yellow highlight).

The orange sequence of pLTC-SGC1 is identical to a portion of f1 origin. The highlighted areas correspond to -35 and -10 sites for RNApol  $\sigma^{70}$  which normally produces an RNA template for DNA polymerase, but has weak promoter activity when the DNA is double-stranded (Higashitani et al. 1997. PNAS USA 94:2909).

The cyan region of the PIB and pMIB series, and pYE series is S. cerevisiae URA3 3' flanking sequence, but is predicted by BPROM to form a viable -35 and -10 (yellow highlight). The AmpR of pCR2.1-TOPO is produced as a bicistronic transcript with NeoR/KanR from the NeoR/KanR promoter.

#### Un-annotated CAT promoters (colored segments = vector backbone)

Name of Plasmid	Alignment
Or Feature	
CAT_prom-001	${ t tgatcgg-qacgtaagagcttccaactttcaccataat}$ ${ t taaataagatcactaccggcgtattttttgagtt}$ ${ t tcagagatttcaacgtaagctaag}$
CAT_prom-002	TGATCGG-QACGTAAGAGGGTCCAACTTTCACCATAAT&AAATAAGATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAA
CAT_prom-003	${ t tgatcgg-dacgtaagaggttccaactttcaccataatgaaataagatcactaccgggcgtattttttgaggttatcgagattttcaggagctaaggaagctaaa$
CAT prom-004	TGATCGG-QACGTAAGAGGTTCCAACTTTCACCATAAT $GAAATAAGATCACTACCGGGCGTATTTTT$ $TGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAA$
pSMART HCKan(KanR)	TGATCGG-QACGTAAGAGGTTCCAACTTTCACCATAAT\$AAATAAGATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAA
pSMART LCKan(KanR)	TGATCGG-QACGTAAGAGGTTCCAACTTTCACCATAAT\$AAATAAGATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTTTCAGGAGCT
pSB1C3(CmR)	${ t t}{ t GATCGG}{ t GACTTAGGGGTTCCACCTTTCACCATAAT { t GAAATAAGATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAA$
pRANGER-BTB-1(AmpR)	ACGTAAGAGGTTCCAACTTTCACCATAATGAAATAAGATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAA
pRANGER-BTB-2 (KanR)	ACGTAAGAGGTTCCAACTTTCACCATAAT\$AAATAAGATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAA
pRANGER-BTB-3 (CmR)	ACGTAAGAGGTTCCAACTTTCACCATAAT\$AAATAAGATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAA
pRANGER-BTB-5 (TpR)	ACGTAAGAGGTTCCAACTTTCACCATAAT&AAATAAGATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAA
pRFP-C-RS(CmR)	ATATGTATCCGCTCATGACTCATGTTTGACAGCTTATCA CGATAAGCTAGCGCCCCTAGCTTAACTTAA
pVP13(CmR)	CCAAGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAAA
pVP16(CmR)	CCAAGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAAA
pVP33K(CmR)	CCAAGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAAA
pVP56K(CmR)	CCAAGCTAGCGTGGCGAGATTTTCAGGAGCTAAGCAAA
pDNR-LIB(CmR)	GGATCTTTCAGGAGCTAAGGAAGCTAAA
pDNR-Dual(CmR)	GGATCTTTCAGGAGCTAAGGAAGCTAAA
	* ******
	-35 <i>CAT</i> prom -10

Sequences were trimmed at the maximum length shown or at the junction of an adjacent feature (producing the shorter sequences). These CAT promoters drive expression of various marker genes indicated in parentheses.

In pSMART\_LCKan, CmR is the only selectable marker. Therefore this sequences must be active.
The cyan portion of pRFP-C-RS is pBR322 backbone but is predicted by BPROM to have -35 and -10 sequences (highlighted in yellow).
The pVP series has two CmR genes - one behind lac UV5 promoter, and the other with the above 5' region, which is not active.

pDNR-LIB/Dual have only CMR as the selectable marker. In the case of pDNR-LIB, CmR is head-to-tail with SacB and is likely produced from a bicistronic transcript with SacB. However, in pDNR-Dual, CMR and SacB are head-to-head. In this case, it seems that the SacB upstream region includes a bidirectional promoter that can drive expression of CMR. Alternatively, the loxP site between CMR and SacB in both plasmids has cryptic promoter activity.

### Un-annotated NeoR/KanR promoters (colored segments = unknown sequence)

Name of Plasmid	Alignment
Or Feature	
pREP4	TCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGA <mark>GG</mark> TCGTTTCGC
pVP56K	TCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCGAGGGGATCAAGATCTGATCAAGAGACAGGATGA <mark>G</mark> GTCGTTTCGC
pVP68K	TCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGA <mark>GG</mark> TCGTTTCGC
pVP65K	TCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCGAGGGGATCAAGATCTGATCAAGAGACAGGATGA <mark>G</mark> GTCGTTTCGC
pVP81K	TCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGA <mark>GG</mark> TCGTTTCGC
pVP33K	TCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCGAGGGGATCAAGATCTGATCAAGAGACAGGATGA <mark>G</mark> GTCGTTTCGC
pSpark_III	TCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAGAGACAGGATGAGGATCGTTTCGC
pCAMBIA5105	TCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTCGC
pZErO-2	TCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTGCCCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTCGC
pSF-pA-PromMCS-Fluc	: GGCAGCGTACCGATCTGTTTAAACCTAGATA <mark>TTGATA</mark> GTCTGATCGGTCAACG <mark>TATAAT</mark> CGAGTCCTAGCTTTTGCAAACATCT_ATCAAGAGACAGGAT <mark>CAGC</mark> AGGA <mark>GGC</mark> TTTCGC
pSF-pA-CMVe-Rluc	GGCAGCGTACCGATCTGTTTAAACCTAGATA <mark>TTGATA</mark> GTCTGATCGGTCAACG <mark>TATAAT</mark> CGAGTCCTAGCTTTTGCAAACATCT_ATCAAGAGACAGGAT <mark>CAGC</mark> AGGA <mark>GGC</mark> TTTCGC
pSF-pA-PromMCS-Rluc	: GGCAGCGTACCGATCTGTTTAAACCTAGATA <mark>TTGATA</mark> GTCTGATCGGTCAACG <mark>TATAAT</mark> CGAGTCCTAGCTTTTGCAAACATCT <mark>ATCAAGAGACAGGAT</mark> CAGCAGGAGGCGGTTTCGC
pSF-PromMCS-Fluc	GGCAGCGTACCGATCTGTTTAAACCTAGATA <mark>TTGATA</mark> GTCTGATCGGTCAACG <mark>TATAAT</mark> CGAGTCCTAGCTTTTGCAAACATCT_ATCAAGAGACAGGAT <mark>CAGC</mark> AGGA <mark>GGC</mark> TTTCGC
pSF-PromMCS-Rluc	GGCAGCGTACCGATCTGTTTAAACCTAGATA <mark>TTGATA</mark> GTCTGATCGGTCAACG <mark>TATAAT</mark> CGAGTCCTAGCTTTTGCAAACATCT_ATCAAGAGACAGGAT <mark>CAGC</mark> AGGA <mark>GGC</mark> TTTCGC
pSF-pA-CMVe-Fluc	GGCAGCGTACCGATCTGTTTAAACCTAGATA <mark>TTGATA</mark> GTCTGATCGGTCAACG <mark>TATAAT</mark> CGAGTCCTAGCTTTTGCAAACATCT_ATCAAGAGACAGGAT <mark>CAGC</mark> AGGA <mark>GGC</mark> TTTCGC
pSF-CMVe-Fluc	GGCAGCGTACCGATCTGTTTAAACCTAGATA <mark>TTGATA</mark> GTCTGATCGGTCAACG <mark>TATAAT</mark> CGAGTCCTAGCTTTTGCAAACATCT_ATCAAGAGACAGGAT <mark>CAGC</mark> AGGA <mark>GGC</mark> TTTCGC
pSF-CMVe-Rluc	GGCAGCGTACCGATCTGTTTAAACCTAGATA <mark>TTGATA</mark> GTCTGATCGGTCAACG <mark>TATAAT</mark> CGAGTCCTAGCTTTTGCAAACATCT_ATCAAGAGACAGGAT <mark>CAGC</mark> AGGA <mark>GGC</mark> TTTCGC
	*** *******
	-35? -10?

The top 9 promoters show identity to the native NeoR/KanR (nptII) promoter from E. coli transposon Tn5.

BPROM did not identify putative -35 and -10 sequences in the native promoter.

The pSF series from Oxford Genetics appear to have a related promoter, but this sequence does not appear to be natural as it generates no hits from homology searches against public databases. Highlighted regions are potential -35 and -10 sequences (BPROM). This promoter is also used to drive AmpR in pSF-COV-Fluc and pSF-CMV-Fluc.

Key

#### Un-annotated KanR promoters (colored segments = plasmid backbone)

Name of Plasmid	Alignment	
Or Feature	<del></del>	_
pShuttle2	ACAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	rcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgtt
pSIREN-Shuttle	ACAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	rcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgtt
pACYC177	ACAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	rcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgtt
pMCSG77	ACAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	rcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgtt
pHSG298	ACAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	rcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgtt
pHSG299	ACAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	
pGreenII_0049	ACAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	rcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgtt
pGreen_0029	ACAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	rcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgtt
pGreen	ACAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	
pKF_18k-2	ACAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	
pKF_19k-2	ACAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	rcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgtt
pOSIP-KT	TGTAACGCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	
pOSIP-KP	TGTAACGCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	
pOSIP-KL	TGTAACGCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	rcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgtt
pOSIP-KO	TGTAACGCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	T
pOSIP-KH	TGTAACGCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	rcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgtt
pOSIP-KC	CTACTTACACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	
pUC57-Kan	GCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	
pCONR201	GCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT	
pMCentr2	GCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAATAATATCAT	
pMCentr3	GCAGCTCTGGCCCGTGTCTCAAAATCTCTGATCTTACATTGCACAAGATAAAATAATATCAT	
pQE-80L	TTCTCGAGGTGAAGACGAAAGGGCCTC <mark>GTGATA</mark> CGCCTATTTTTATAGGT <mark>TAATGT</mark> CATGGT	
pQE-81L	TTCTCGAGGTGAAGACGAAAGGGCCTC <mark>GTGATA</mark> CGCCTATTTTTATAGGT <mark>TAATGT</mark> CATGGT	T
pQE-82L	TTCTCGAGGTGAAGACGAAAGGGCCTC <mark>GTGATA</mark> CGCCTATTTTTATAGGT <mark>TAATGT</mark> CATGG	
pET-9a	TTGATCTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTT <mark>TGGT</mark>	
pET-9b	TTGATCTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTT <mark>TGGT</mark>	
pET-9c	TTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTTGGT	
pET-9d	TTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGT	
pGFP-B-RS		r <mark>catgaacaat</mark> aaaactgtctgcttacataaacagtaatac <mark>aaggggt</mark> gtt
pGFP-V-RS		r <mark>catgaacaat</mark> aaaactgtctgcttacataaacagtaatac <mark>a</mark> aggggtgtt
pCMV6-Entry2		r <mark>catgaacaat</mark> aaaactgtctgcttacataaacagtaatac <mark>a</mark> aggggtgtt
AmpR_prom-009	CGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCT	
		* * * * * * * * * * * * * * * * * * *
	-35 $aph(3')-Ia$ prom $-10$	RBS

These promoters are the natural aph(3')-Ia promoter from E.  $coli\ Tn5$  and all drive expression of KanR (aph(3')-IaI). Highlighting shows putative -35 and -10 sequences (BPROM). The pGFP and pCMV series have a hybrid promoter with AmpR promoter (cyan highlighting) from replacing AmpR with KanR during plasmid construction (TCATGA is a BspHI restriction site).

#### Name of Plasmid Alignment Or Feature pCAMBIA0305.1 pCAMBIA0305.2 pCAMBIA03080 ${\tt GATCACCGCGGTTTCAAAATCGGCTCCGTCGATACTATGTTATACGC.GAACA} {\tt GTGAATTGGAGTTCGTCTTGTTATAAT} {\tt FAGCTTCTTGGGGTATCTTTAAATACTGTAGAAAAGAGGAAATAATAATAA}$ pCAMBIA03090 pEarleyGate 100 pEarleyGate\_101 pEarlevGate 102 pFGC5941 ${\tt GATCACCGCGGTTTCAAAATCGGCTCCGTCGATACTATGTTATACGC.GAACA} {\tt GTGAATTGGAGTTCGTCTTGTTATAAT} {\tt FAGCTTCTTGGGGTATCTTTAAATACTGTAGAAAAGAGGAAATAATAATAA}$ pBI121 ${ t rggtttcaaaatcggctccgtcgatactatgttatacgc.gaaca}$ ${ t rggattcgtcttgttataat racctctttggggtatctttaaatactgtagaaaaagagaagtaaataataa$ pRI\_101-AN pRI\_101-ON CATAATTO pRI\_201-AN CATAATT pRI\_201-ON pRI\_909 CATAATT ${ t T}{ t G}{ t G}{ t G}{ t T}{ t C}{ t G}{ t C}{ t T}{ t C}{ t C}{ t T}{ t C}{ t$ pRI\_910 pBINPLUS CATAATTGTGGTTTCAAAATCGGCTCCGTCGATACTATGTTATACGC.GAACAGTGAATTGGAGTTCGTCTTGTTATAATTAGCTTCTTGGGGTATCTTTAAATACTGTAGAAAAGAGGAAAGAAGAAAATAATAA -35 aphA-3 prom -10

These plasmids all use KanR. These promoters are the natural aphA-3 (KanR-020) promoter from Staphylococcus aureus.

#### Un-annotated SmR promoters

#### Name of Plasmid Alignment Or Feature pDEST20-Pc prom DIDS pDONR223 pZP100 acccadtggacataagcctcgttcggttcgtaagctgtaatgcaagtagcgtaactgccgtcacgcaactggtccagaaccttgaccgaacgcagcggtggtaacggcgcggtggtttcat pZP101 acccadtggacataagcctcgttcggttcgtaagctgtaatgcaagtagcgtaactgccgtcacgcaactggtccagaaccttgaccgaacgcagcggtggtaacggcgcagtggcggttttcat pZP222 acccadtggacataagcctcgttcggttcgtaagctgtaatgcaagtagcgtaactgccgtcacgcaactggtccagaaccttgaccgaacgcagcggtggtaacggcgcagtggcggttttcat pZP221 acccadtggacataagcctcgttcggttcgtaagctgtaatgcaagtagcgtaactgccgtcacgcaactggtccagaaccttgaccgaacgcagcggtggtaacggcgcagtggcggttttcat ${\tt ACCCACTGGACATAAGCCTCGTTCGGTTCGTAAGCTCTAAGCGTAACTGCCGTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGGTTGTTCAT}$ pZP212 pZP211 acccadtggacataagcctcgttcggttcgtaagctgtaatgcaagtagcgtaactgccgtcacgcaactggtccagaaccttgaccgaacgcagcggtggtaacggcggtgtttcat pZP202 acccadtggacataagcctcgttcggttcgtaagctgtaatgcaagtagcgtaactgccgtcacgcaactggtccagaaccttgaccgaacgcagcggtggtaacggcggtgtttcat pZP201 acccadtggacataagcctcgttcggttcgtaagctgtaatgcaagtagcgtaactgccgtcacgcaactggtccagaaccttgaccgaacgcagcggtggtaacggcggtgtttcat pZP200 acccadtggacataagcctcgttcggttcgtaagctgtaatgcaagtagcgtaactgccgtcacgcaactggtccagaaccttgaccgaacgcagcggtggtaacggcgg pCAMBIA1200 ACCCADTGGACATAAGCCTCGTTCGGTTAGTGCAAGTAGCGTAACTGCCGTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGGCGGTTTTCAT pCAMBIA1105.1R acccadtggacataagcctcgttcggttcgtaagctgtaatgcaagtagcgtaactgccgtcacgcaactggtccagaaccttgaccgaacgcagcggtggtaacggcgcagtggcggttttcat pCAMBIA1105.1 acccadtggacataagcctcgttcggttcgtaagctgtaatgcaagtagcgtaactgccgtcacgcaactggtccagaaccttgaccgaacgcagcggtggtaacggcggtgtttcat pCAMBIA0105.1R ACCCADTGGACATAAGCCTCGTTCGGTTCGTAAGCTGTAATGCAAGTAGCGTAACTGCCGTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGGCGGTTTTTCAT pCAMBIA5105 ACCCADTGGACATAAGCCTCGTTCGGTTCGTAAGCTDAATGCAAGTAGCGTAACTGCCGTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGGCGGTTTTTCAT Pc promoter Tn+T1

pDEST20-Pc prom pDONR223 pZP100 pZP101 GGCTTqTTGTTATGACATGTTTTTTTGGGGTACAGTpTATGCCTCGGGCATCCAAGCAGCAGCGCGTTACGCCGTGGGTCGATGTTTATGTGTTATGGAGCAGCAACGATGTTACGCAGCAGGACGATCGCCCTAAAACAAGTTAAACAATC pZP222 pZP221 GGCTTQTTGTTATGACATGTTTTTTTGGGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAGCGCGTTACGCCGTGGGTCGATGTTTATGTGTTATGGAGCAGCAACGATGTTACGCAGCAGGACGATCGCCCTAAAACAAAGTTAAACCATC pZP212 GGCTTQTTGTTATGACATGTTTTTTTGGGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAGCGCGTTACGCCGTGGGTCGATGTTTATGTGTTATGGAGCAGCAACGATGTTACGCAGCAGGACGATCGCCCTAAAACAAAGTTAAACCATC pZP211 pZP202 pZP201 GGCTTOTTGTTATGACATGTTTTTTTGGGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAGTTAAACCATC pZP200 pCAMBIA1200  $GGCTT \\ dTTGTTATGACATGTTTTTTTTGGGGTACAGT \\ \\ CTATGCCTCGGGCATCCAAGCAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTTATGGAGCAGCAGCAGCGAGGGCAGTCGCCCTAAAACAAGTTAAACATC$ pCAMBIA1105.1R GGCTTQTTGTTATGACATGTTTTTTTGGGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAGCGCGTTACGCCGTGGGTCGATGTTTATGTGTTATGGAGCAGCAACGATGTTACGCAGCAGGACGATCGCCCTAAAACAAAGTTAAACCATC pCAMBIA1105.1 GGCTTQTTGTTATGACATGTTTTTTTGGGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAGCGCGTTACGCCGTGGGTCGATGTTTATGTGTTATGGAGCAGCAACGATGTTACGCAGCAGGACGATCGCCCTAAAACAAAGTTAAACCATC GGCTTOTTGTTATGACATGTTTTTTTGGGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAGTTAAACCATC pCAMBTA0105.1R GGCTTOTTGTTATGACATGTTTTTTTGGGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAGTTAAACCATC pCAMBIA5105 \*\*\*\*\*\*\*\* \*\*\*\*\* \* P2 promoter

The native SmR promoter for these plasmids is derived from *E. coli* plasmid R538 and drives expression of the spectinomycin and streptomycin resistance gene (SmR) aadA (aminoglycoside 3"-adenylyltransferase), which is a class 1 integron and includes two promoters to drive expression of gene cassettes that integrate at the att1 site (mutated in these constructs) of the plasmid. The Pc promoter is embedded in the 5' end of a gene, *IntIl*, that encodes the integrase responsible for recombination between the attC containing integron and the att1 site behind the promoter. *IntIl* expression is driven by a promoter, Pint (not shown), that partially overlaps the P2 promoter, but is inverted. With the exception of pIDS, which is missing Pc, these plasmids all have a truncated *IntIl*.

inactivated att1 site

Promoter strength analysis based on Jove et al. (2010) PloS Genet. 6:e1000793. The Pc promoter of pDEST20 and derivatives is a strong version of the promoter called PcS. It has optimal -35 to -10 spacing and two bp changes relative to the other Pc promoters in the alignment, which are weak (PcW). However, the P2 promoter of pDEST20 is probably inactive as it has only 14 bp between the -35 and -10 hexamers.

SMR expression from pIDS is driven only by P2 which has the optimal spacing (17 bp) between the -35 and -10 hexamers.

Both the PcW and P2 promoters of pDONR223 have optimal spacing.

The pZP and pCAMBIA series have sub-optimal spacing (18 bp) in both the PcS and P2 promoters.

In pDEST20 and derivatives, the Pc promoter is used to express GmR (gentamycin resistance; aacC1, gentamycin acetyltransferase). In pZP1XX series, Pc drives expression of CmR.

#### Un-annotated SmR promoters

### Name of Plasmid Or Feature

pHELLSGATE-sat-aadA1 pHELLSGATE4-sat-aadA1 pHELLSGATE8-sat-aadA1 pHELLSGATE12-sat-aadA1 pSB11-sat-aadA1

#### Name of Plasmid Or Feature

pHELLSGATE 5'sat1 pHELLSGATE4 5'sat1 pHELLSGATE8 5'sat1 pHELLSGATE12\_5'sat1 pSB11\_5'sat1

#### Alignment

 $\texttt{CAAT} \textcolor{red}{\textbf{TGTAC}} \texttt{GCAAAATGTGGCTT} \textcolor{red}{\textbf{TACTCT}} \texttt{CGGCGGCATTGACCTGTTCACG.CAATTCATTCAGGCCGACACCGCTTCGCGGCGGCGTTAATTCAGGAGTTAAACATC}$  $\texttt{CAAT} \textcolor{red}{\textbf{TGTAC}} \texttt{GCAAAATGTGGCTT} \textcolor{red}{\textbf{TACTCT}} \texttt{CGGCGGCATTGACCTGTTCACG.CAATTCATTCAGGCCGACACCGCTTCGCGGCGGCGTTAATTCAGGAGTTAAACATC}$  $\texttt{CAAT} \textcolor{red}{\textbf{TTGTAC}} \texttt{GCAAAATGTGGCTT} \textcolor{red}{\textbf{TACTCT}} \texttt{CGGCGGCATTGACCTGTTCACG.} \textcolor{red}{\textbf{CAATTCATTCAAGCCGACACCGCTTCGCGGCGCGCTTAATTCAGGAGTTAAACATC}$ \*

#### Alignment

 ${\tt AAGTGGCAGCAACCGGATTCGCAAACCTGTCACGCCTTTTGTGCCAAAAGCCGCGCCAGGTTTGCGATCCGCTGTGCCAGGCGTTAGGCGTCAT}$  ${\tt AAGTGGCAGCAACCGGATTCGCAAACCTGTCACGCCTTTTGTGCCAAAAGCCGCGCCAGGTTTGCGATCCGCTGTGCCAGGCGTTAGGCGTCAT}$ A AGTGGC A GC A ACGGATTCGC A A ACCTGTC A CGCCTTTTTGTGC C A A A AGCCGCGCCTTGTGC C A TCGCCTGTGTGC C A TCGCTGTGTGC C A TCGCTGTGTG C CATCATT  ${\tt AAGTGGCAGCAACCGGATTCGCAAACCTGTCACGCCTTTTGTGCCAAAAGCCGCGCCAGGTTTGCGATCCGCTGTGCCAGGCGTTAGGCGTCAT}$  ${\tt AAGTGGCAGCAACCGATTCGCAAACCTGTCACGCCTTTTGTGCCAAAAGCCGCGCCAGGTTTGCGATCCGCTGTGCCAGGCGTTAGGCGTCAT}$ \*

SmR in these plasmids is derived from E. coli Tn7 which contains a class 2 integron including dfrAI (dihydrofolate reductase conferring resistance to trimethoprim), sat (streptothricin acetyltransferase), aadA1 (SmR - spectinomycin and streptomycin resistance, aminoglycoside adenylyltransferase), and ybeA (pseudouridine methyltransferase).

Class 2 integrons have a similar structure to class 1, but the sequences of the Pc promoter (-35 TTTAAT 16 bp TAAAAT -10) and the integrase (IntI2) are different. The Pc promoter of Tn7 drives expression of a polycistronic transcript (dfrA1-sat-aadA1-ybeA). However, neither the pSB11 nor the pHELLSGATE plasmids include the PC promoter from Tn7, which is upstream of dfrA1. These plasmids all include sat, aadA1 and ybeA, but pSB11 includes just a portion of the end of dfrA1, while the pHELLSGATE plasmids start just 5' to sat. Therefore, it is not clear what sequence constitutes a promoter in these constructs.

In the alignments, we show the region covering the 3' end of sat and intergenic region between sat and SMR (top alignment), and the region upstream of sat common to pSB11 and pHELLSGATE (bottom alignment), which has no putative -35 and -10 site. It seems likely that there is a cryptic promoter within these regions or the sat ORF. BPROM predicts a putative promoter at the end of the sat ORF (top alignment; yellow highlight). This region is ~100 bp upstream of the aadA1 START codon.

Kev

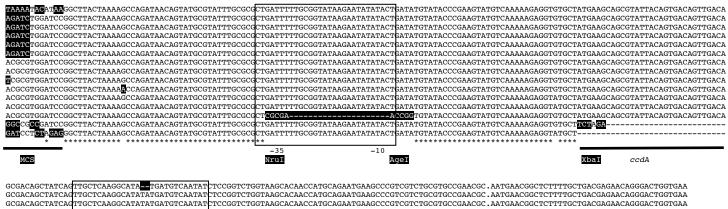
ATGC = change in nucleotide relative to consensus sequence - non-coding sequence

#### Un-annotated ccdB promoters

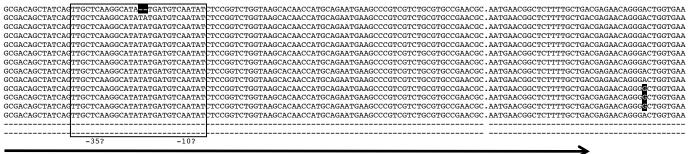
#### Name of Plasmid Or Feature

ccdB\_nat\_prom pDEST26 pLenti4\_V5-DEST pT-Tex-DEST30 pcDNA-DEST40 pANT7-nHA pDEST8 pAd\_BLOCK-iT-DEST pDONR201 pAd\_PL-DEST pET-53-DEST pIB\_V5-His-DEST pMT\_BioEase-DEST pCMV\_SPORT6ccdB pOSIP-CH

#### Alignment



ccdB\_nat\_prom
pDEST26
pLenti4\_V5-DEST
pT-Tex-DEST30
pcDNA-DEST40
pANT7-nHA
pDEST8
pAd\_BLOCK-iT-DEST
pDONR201
pAd\_PL-DEST
pET-53-DEST
pET-53-DEST
pMT\_BioEase-DEST
pCMV\_SPORT6ccdB
pOSIP-CH



ccdA

The native ccdB promoter (ccdB\_nat\_prom) from E. coli F plasmid is shown in the first row. ccdB is transcribed as part of a multi-cistronic transcript along with ccdA (arrow), an inhibitor of ccdB protein. ccdA is inactivated in ccdB plasmids by a two bp insertion, or by complete deletion of the ORF (pCMV SPORT6ccdB and pOSIP-CH).

Restriction enzyme sites used for cloning (MCS in pOSIP-CH) or for construction of the plasmids are shown.

The -10 and -35 sequences are missing from pMT\_BioEase-DEST, but a second potential -35 and -10 is present in the disrupted ccdA ORF (BPROM). Deletion of the natural promoter may allow more efficient transcription from this cryptic promoter, bypassing transcription of the entire ccdA ORF.

## Coding Sequences

Maps	Name of	# of	Alignment	
Appli-01 78 APAGEMENT CONTROL OF THE PROCESS OF THE	Variant	0ccur		
Appl-01 76-72 Appl-02 61-1 Appl-03 76-72 Appl-03 76-72 Appl-03 76-73 Appl-04 76-73 Appl-05 76-73 Appl-06 76-73 Appl-06 76-73 Appl-07 76-73 App				
Augh-21 69 A PAGE AND ATT CONTROL OF THE CONTROL OF				
AmpR-012 4 5-1 APAGEMENT CONTROL OF THE THORSE OF THE THOR			ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTTGCGGCATTTTGCCTTCTTTTTTGCTCACCCAGAAA	CGCTGGTGAAAGTAAAAGATGCTGAAGATCA <u>GT</u> TGGGTGCA
AmpR-012 4 5-1 APAGEMENT CONTROL OF THE THORSE OF THE THOR			ATGAG <mark>C</mark> AT <mark>C</mark> CAACATTT <mark>T</mark> CGTGTCGC <mark>A</mark> CTCATTCCCTT <mark>C</mark> TTTGCGGCATTTTGC <mark>T</mark> CCTGTTTTTTGC <mark>A</mark> CACCC <mark>C</mark> GAAA	CGCTGGTGAAAGTAAAAGATGCTGAAGATCA <mark>AC</mark> TGGGTGCA
APRIL	AmpR-016	45+1	ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTTGCCTTCTTGTTTTTTGCTCACCCAGAAA	CGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCA
APRIL	AmpR-022	28	ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCTGTTTTTTGCTCACCCAGAAA	CGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCA
Augh-19 8 ATAGGATATCAGATTCCGTTTTTGCGCCCTTTTTGCCGCATTTTGCTGTTTTTGCTGCCCGTAAAAACGCTGGTGAAAATAAAAACTCGAGAACCAAAAACCATGGGTGCAAAGAAGACTCGAGAACCAAAAACCATGGGTGCAAAGACAAGACCATTTGCGTTCCCCTTTTTGCCTCCCCTTTTTGCCTGCC	AmpR-010	16	ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCTGTTTTTTGCTCACCCAGAAA	CGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCA
Augh-19 8 ATAGGATATCAGATTCCGTTTTTGCGCCCTTTTTGCCGCATTTTGCTGTTTTTGCTGCCCGTAAAAACGCTGGTGAAAATAAAAACTCGAGAACCAAAAACCATGGGTGCAAAGAAGACTCGAGAACCAAAAACCATGGGTGCAAAGACAAGACCATTTGCGTTCCCCTTTTTGCCTCCCCTTTTTGCCTGCC	AmpR-004	9+3	ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCTTGTTTTTGCTCACCCAGAAA	CGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCA
ABBR-012 8 ARGAGTATCAAGATTCCOTTCCTTTTTTGCCCCCTCACCCAGAAGCTGAAGATCAAGATCCGTGGTGCAAAGTTCAAGTTCCCCTTTTTTGCCGCCATTTTGCCTCCTCTTTTTTTGCCACCCAGAAGCTGCTGAAGTTCAAAGTTCCGTGTGCAAAGTTCAAAGTTCCGTCTTTTTTTGCCCCCAGAAGCTCAAAGTTCCACCCAGAAGTTCAAACTTCCGTCTTTTTTGCTCACCCAGAAGTTCAAACTTCCGTCGTCGTCGTCACCCAGAAGTTCAAACTTCCGTCTTTTTTTGCTCACCCAGAAGTTCAAACTTCCGTCGTCGTCACCCAAAGTTCAAACTTCCGTCTTTTTTGCTCACCCAGAAGTTCAAACTTCCGTCGTCGTCACCCAAAGTTCAAACTTCCGTCTTTTTTGCTCACCCAGAAGCTCTCAAACTTCAAACTTCCGTCGTCACCCAAAGTTCAAACTTCCGTCTTTTTTGCTCACCCAGAAGCTCGTCAAAGTTCAAACTTCCGTCGTCTGTCACCCAGAAGCTCTGTCAAGTTCAAACTTCCGTCGTCTTTTTTGCTCACCCAGAAGCTCGTCAAAGTTCAAAACTTCCGTCGTCGTCTTTTTTGCTCACCCAGAAGCTCGTCAAAGTTCAAAACTTCCGTCGTCTTTTTTGCTCACCCAGAAGCTCGTCAAAGTTCAAAACTTCCGTCGTCTCTTTTTTGCTCACCCAGAAGCTCGTCAAAGTTCAAACTTCCGTCGTCTCTTTTTTGCTCACCCAGAAGCTCGTCAAAGTTCAAACTTCCGTCGTCACCCAGAAGCTCTCAAACTTCAAACTTCCGTCTTTTTTGCTCACCCAGAAGCTCGTCAAAGTTCAAACTTCCGTCGTCACCCAGAAGCTCGTCAAAGTTCAAACTTCCGTCTTTTTTGCTCACCCAGAAGCTCGTCAAAGTTCAAACTTCCGTCGTCACCCAGAAGCTCGTCAAAGTTCAAACTTCCGTCTTTTTTGCTCACCCAGAAGCTCGTCAAAGTTCAAACTTCCGTCAAACTTCAAACTTCCGTCTTTTTTGCTCACCCAGAAACTTCAAACTTCAAACTTCCGTCTTTTTTGCTCACCCAGAAGCTCGTCAAAACTTCAAAACATTCCGTTCTTTTTTTGCTCACCCAGAAGCTCGTCAAAACTTCAAAACATTCCGTTCAAACTTAAAACTTCCAAACTTCAAAACTTCCAAACTTCAAAACTTCCAAACTTCAAAACTTCCAAACTTCAAAACTTCCAAACTTCAAAACTTCCAAAACTTCAAAACTTCCAAAACTTCAAAACTTCCAAAACTTCAAAACTTCAAAACTTCAAAACTTCAAAACTTCAAAACTTCAAAACTTCAAAAAA		8	ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCTTGTTTTTGCTCACCCAGAAA	CGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCA
ABBR-013 5 AGRAFATCHARACTTCCOTTOTTCCCTTTTTTCGCCACCAGAAGCTGAAGTCAAAGCTGCTGAAGTTCAAGCTGCTGAAGTTCAAGCTGCCTAAGTTCAAGCTTCCCTTTTTTTGCCACCAGAAGCTGCTAAGTTCAAGCTGCGTGAAGTTCAAGTT		8		
APRIL				
AmgR-015 5 A ATGASTATCAAGATTCCGTCTTTTTGCGCCCTTATTTGCGCCCCTTATTTGCGGCCACCAGAAGCTGGTGAAGGTCAAAGATCCGTGGTGGAAGGTCAAAGATCCGTGGTGCAAAGTTCCGTTTTTTGCGGCCTTTTTTGCGCCCTTTTTTGCGCCCTTTTTT				
APGA-011				
APGR-002				
Amgin		-		
AmgR=-003   2		2		
Amgr-006				
AmpR-005 1 ATGGGTATCCAGACTTCCGTCTTTTGCCGCCTTTTTGCTCCCGGAAAGCGGTGAAAGCAAAAGCATTGGGTCAAAAGCAAAAGATGAAAGCATTGGGTCAAAGCAAAAGCATTGGGTCAAAAGCAAAAGCATTGGGTCAAAAGCAAAAGCATTGGGTCAAAAGCAAAAGCATTGGGTCAAAAGCAAAAGCATTGGGTCAAAAGCAAAAGCATTGGGTCAAAAGCAAAAGCATTGGGTCAAAAGCAAAAGCATTGGGTCAAAAGCAAAAGCATTGGGTCAAAAGCAAAAGCATTGGGTCAAAAGCAAAAGCATTGGGTCAAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCATTAAAAGCATTGGGTCAAAGCAAAAGCAAAAGCAAAAGCAAAAGCAAGC				
Amgn-012   ATGAGYATTCCACCTTATTCCCTTTTTGCGGCATTTTGCCTCTCTTTTTCTCCCCCAGAAGCGTGGGAAAGTAAAAGTGCTGAAGGTGGGTG				
AmpR-013 1 ATGASTATICAGACTITTCGGGTCATTTGCGGCATTTTGCCTCCTCTTTTTGCTCACCCGAAAGCGTGGGAAAGTAAAAGTGCTGAAAGTACAAGTGGGGGCAAMGR-023 1 ATGASTATICAACTTTCGGGTCATTTTGCGGCATTTTGCGGCCTTTTTGCCTCCCCGGAAAGCGTGGAAAGTAAAAGTGCTGAAAGTACAAGTTGGGGCCAAMGR-023 1 ATGASTATICAACTTTCGGGTCATTTTGCGGCATTTTGCGCCCTCTTTTTGCTCACCCCGAAAGACGTGGAAAGTAAAAGTGCTGAAAGTACAAGTTGGGGCCAAMGR-023 1 ATGASTATICAACATTTCGGTCCCCCTATTTCCCTTTTTTGCGCACTTTTGCGCCCCTATTTCCCTCTTTTTGCGCACTTTGGGCACTTTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAAGTGCTGAAAGTAAAAAGTGCTGAAAGTAAAAGTGCTGAAAGTAAAAAGTGCTGAAAGTAAAAAGTGCTGAAAGTAAAAAGTGCTGAAAGTAAAAAGTGCCTGAAAGTAAAAAAGTGCCTGAAAGTAAAAAAGTGCCTGAAAGTAAAAAGTGCCTGAAAGTAAAAAAGTGCCTGAAAGTAAAAAGTGCCTGAAAGTAAAAAGTGCCTGAAAGTAAAAAAAGTGCCTGAAAGTAAAAAGTGCCTGAAAGTAAAAAGTGCCTGAAAGTAAAAAGTGCCTGAAAGTAAAAAGTGCCTGAAAGTAAAAAAGTGCCTGAAAGTAAAAAGTGCCTGAAAGTAAAAAGTGCCTGAAAGTAAAAAGTGCCTGAAAGTAAAAAGTGCCTGAAAGTAAAAAGTGCCTGAAAGTAAAAGTGCCTGAAAGTTAAAAGTTCCGTGCCCCATAGAAGTTTGCCTTCCTT				
AmpR-012 1 ATGASTATICACACTITICCSTITTICGGGCATTITGGCTCTCTTTTTCCTCACCAGAAAGCTGGGAAAAGTAAAAGTGCTAAAAGTACATGGGGGCTA h1a(N)-001 1		1		
AmpR-013 CGAGTGGGTTACATCGACACTTATTCCCTTTTTTTCCGCCCCGAGAACCTGGTAAAACAAAC				
Dia (N) -001				
bla(N)-002 1 CAGGRAGGTTACATCGATGTCCATTTTTCCCTTTTTTGCGCCTTATTCCTTTTTTGCGCTCCTATTTTTCCCTTCGATGTTTTTTTT		1		
Dia (N) - Oo3		_	CCAGAAA	CGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCA
AmpR-007corr AmpR-007corr AmpR-007corr Ala (M) -001corr Ala (M) -001corr Ala (M) -001corr AmpR-003corr AmpR-003corr AmpR-003corr AmpR-003corr AmpR-003corr AmpR-003corr AmpR-003corr AmpR-003corr AmpR-003 CGAGTGGGTTACATCGGACTGGATTTGCCTTTTTTGGGGCATTTTTGCCTCCCCGAAAACCGTGGTAAAGTAAAAGAGTCGGAGGTCAAAGTACAGTGGGTCA AmpR-013 CGAGTGGGTTACATCGGATCGAATTCCGGTCTCTTTTTTGGGCATTTTTGCCTCCCCGAAAACCGTGGTCAAAGTAAAAGAGTCGGAGAGACCGTTTGCTCAAAGTAAAACAATTCCGGTCCTTTTTGCTACCCCGAAAACCGTGGAAAACCTAAAAGAAAAGAGTCGGAAGAGTCCTAAAGTAAAAGATCCGAACTTTAAAAGTCTGGAACGTCCAAAGTACAAGTTTTGCTCACCCGAAAACCGTGAAAACCTAAAAGAAACCTTTAAAAGTCTGGGTCA AmpR-017 CGAGTGGGTTACATCGAACTGGATTCAACAACGGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAGAACCGTTTTC		1	<mark>ATC</mark> ACCCAGAAA	CGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCA
AmpR-007corr AmpR-007corr AmpR-007corr Ala (M) -001corr Ala (M) -001corr Ala (M) -001corr AmpR-003corr AmpR-003corr AmpR-003corr AmpR-003corr AmpR-003corr AmpR-003corr AmpR-003corr AmpR-003corr AmpR-003 CGAGTGGGTTACATCGGACTGGATTTGCCTTTTTTGGGGCATTTTTGCCTCCCCGAAAACCGTGGTAAAGTAAAAGAGTCGGAGGTCAAAGTACAGTGGGTCA AmpR-013 CGAGTGGGTTACATCGGATCGAATTCCGGTCTCTTTTTTGGGCATTTTTGCCTCCCCGAAAACCGTGGTCAAAGTAAAAGAGTCGGAGAGACCGTTTGCTCAAAGTAAAACAATTCCGGTCCTTTTTGCTACCCCGAAAACCGTGGAAAACCTAAAAGAAAAGAGTCGGAAGAGTCCTAAAGTAAAAGATCCGAACTTTAAAAGTCTGGAACGTCCAAAGTACAAGTTTTGCTCACCCGAAAACCGTGAAAACCTAAAAGAAACCTTTAAAAGTCTGGGTCA AmpR-017 CGAGTGGGTTACATCGAACTGGATTCAACAACGGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAGAACCGTTTTC			CAGAAA	CGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCA
AmpR-019 CGAGTGGGTTACATGGATCTGAATGGGGGTAAGATTTTCGCTCCTGTTTTTGCTCACCCGAAAACCTGGTAAAGAAACAGTTGGGTGCA  Ala (M) - 00 acorr  ATGAGTATTACAGATTTCCGGTGCCCCTTTTTTGCGGCCATTTTGCCTCCTGTTTTTTGCTCACCCGAAAACCTGAAAGTAAAGAGTCGAAGTTGGGTGCA  ATGAGTATTCAACAGTTCCGGTCCCCTTTTTTGCGGCCATTTTGCCTCCTGTTTTTTGCTCACCCGAAAACCTGGTGAAAGTAAAAGAGTTGGGTGCA  ATGAGTATTCAACAGTTCGGACTGGATCTCAACAGTTGGGTCA  ATGAGTATTCAACAGTTCGGACTGGATCTCAACAGGGGTAAGATCCTTGAGAGTTTTTGCCTCCCGAAAACCTGGAAACCTGAAAGTAAAGAGTTGTGGTGCA  ATGAGTATTCAACAGTTGGATCTCAACAGGGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACCGTTTTC				CGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCA
Dia (M) - O01 corr				
Dia (M) - 003 cort   ATGATATTCAACATTTCCTTATTCCCTTATTTCCCTTCTTTTTGCCCCCGAAGAACGTTGGTGAAGTTAAAGATTCATAAGATCATTGGGTCA				
AmpR-013 CGAOTGGGTTACATCGACTGGATCTCAACAGCGGTAAGATCTTGAGAGTTTTCGCCCCCGAAGAACGTTTC——CAATGATGAGCACTTTTAAAGTTCTGCTATOTGGC AmpR-017 CGAOTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTC——CAATGATGAGCACTTTTAAAGTTCTGCTATOTGGC AmpR-021 CGAOTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTC——CAATGATGAGCACTTTTAAAGTTCTGCTATOTGGC AmpR-015 CGAOTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTC——CAATGATGAGCACTTTTAAAGTTCTGCTATOTGGC AmpR-016 CGAOTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTC——CAATGATGAGCACTTTTAAAGTTCTGCTATOTGGC AmpR-010 CGAOTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTC——CAATGATGAGCACTTTTAAAGTTCTGCTATOTGGC AmpR-010 CGAOTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAGAACGTTTC——CAATGATGAGCACTTTTAAAGTTCTGCTATOTGGC AmpR-014 CGAOTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAGAACGTTTC——CAATGATGAGCACTTTTAAAGTTCTGCTATOTGGC AmpR-019 CGAOTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCAAGAACGTTTC——CAATGATGAGCACTTTTAAAGTTCTGCTATOTGGC AmpR-010 CGAOTGGGTTACATCGAACTGGATCCTAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAGAACGTTTC——CAATGATGAGCACTTTTAAAGTTCTGCTATOTGGC AmpR-010 CGAOTGGGTTACATCGAACTGGATCCTAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAGAACCTTTC——CAATGATGAGCACTTTTAAAGTTCTCTCTATOTGGC AmpR-011 CGAOTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAGAACCTTTCC——CAATGATGACCACTTTTAAAGTTCTCTCTATOTGGC AmpR-015 CGAOTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAGAACGTTTC——CAATGATGACCACTTTTAAAGTTCTCTCTATGTGGC AmpR-011 CGAOTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTC——CAATGATGACCACTTTTAAAGTTCTCTCTATGTGGC AmpR-012 CGAOTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTC——CAATGATGACCACTTTTAAAGTTCTCTCTATGTGGC AmpR-001 CGAOTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC——CAATGATGACCACTTTTAAAGTTCTCTCTATGTGGC AmpR-002 CGAOTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC——CAATGATGACCACTTTTAAAGTTCTCTCTATGTGGC AmpR-003 CGAOTGGGTTACAT				
AmpR-013 CGACTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-017 CGAGTGGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-021 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGAGCGTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-016 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGAGCGTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-021 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAGAGTCTTTGAGAGTTTTCGCCCCGAAGAACGTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-010 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-010 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAGAGTCTTTGGCCCGAAGAACGTTTTC — CCAATGATGAGCACTTTTAAAGTTCTCTATGTGGC AmpR-019 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGGCCCGAAGAACGTTTTC — CCAATGATGACACCTTTTAAAGTTCTCTATGTGGC AmpR-019 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGGCCCCAAGAACGTTTTC — CCAATGATGACACCTTTTAAAGTTCTCTATGTGGC AmpR-010 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGGCCCCAAGAACGTTTTC — CCAATGATGACACCTTTTAAAGTTCTCTATGTGGC AmpR-011 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGGCCCCAAGAACGTTTTC — CCAATGATGACACCTTTTAAAGTTCTCTATGTGGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGGCCCCAAGAACGTTTTC — CCAATGATGACACCTTTTAAAGTTCTCTATGTGGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGGCCCCCAAGAACCTTTTC — CCAATGATGACACCTTTTAAAGTTCTCCTTATGGGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGGCCCCCAAGAACCTTTTC — CCAATGATGACACCTTTTAAAGTTCTCCTTTGTGGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGGCCCCCAAGAACCTTTTC — CCAATGATGACACCTTTTAAAGTTCTCCTTTGTGGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGGCCCCCAAGAACCTTTTC — CCAATGATGACACCTTTTAAAGTTCTCCTTTGTGGC AmpR-016 CGAGTGGGTTACATCGAACTGCATCTCAACAGCGGTAAGATCTTTGGCCCCCAAGAACCTTTTC — CCAATGATGACACCTTTTAAAGTTCTCCTTATGTGGC AmpR-012 CGAGTGGGTTACATCGAACTGCATCTAACAGCGGTAAGATCTTTTGGCCCCCAAGAACCTTTTC — CCAATGATGACACCTTTTAAAGTTCTCTTTTGTGGC AmpR-013 CGA				
AmpR-013 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGAGAGTTTTCGCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGAAPR-017 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTGAAGATTTTCGCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGAAPR-018 CGAGTGGGTACATCCAACAGCGGTAAGATCCTTGAAGATTTTCGCCCCGAAGAACGTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGAAPR-018 CGAGTGGGTAAPACCAACTGAACTCCAACAGCGGTAAGATCCTGAAGATTTCGCCCCGAAGAACGTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGAAPR-018 CGAGTGGGTACATCCAACAGCGGTAAGATCCTTGAAGATTTTCGCCCCGAAGAACGTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGAAPR-019 CGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAAGATTTCCGCCCCGAAGAACGTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGAAPR-019 CGAGTGGGTACATCGAACACGGTAAGATCCTTGAAGATTTCGCCCCGAAGAACGTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCAAPR-019 CGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAAGATTTCCGCCCCGAAGAACGTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCAPR-019 CGAGTGGGTACATCGAACACGGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCAPR-019 CGAGTGGGTACATCGAACACGGTAAGATCCTTTGAGAGTTTTCGCCCCGAAGAACGTTTTC CAATGATGACACACTTTTAAAGTTCTGCTATGTGGCARR-019 CGAGTGGGTACATCGAACGACTTGAACGCGGTAAGATCCTTTTTCGCCCCGAAGAACGTTTTC CAATGATGACACACTTTTAAAGTTCTGCTATGTGGCARR-019 CGAGTGGGTACATCAACAGACGGTAAGATCCTTTGAGAGTTTTTCGCCCCGAAGAACGTTTTC CAATGATGACACACTTTTAAAGTTCTGCTATGTGGCARR-019 CGAGTGGGTAACACGACTTTTAAAGTTCTGCTATGTGGCARR-019 CGAGTGGGTAACACGACTTTTAAAGTTCTGCTATGTGGCARR-019 CGACTGGGTAACACGACTTTTAAAGTTCTGCTATGTGGCARR-019 CGACTGGGTAACACGACTTTTAAAGTTCTCATTGTGGCARR-019 CGACTGGGTAACACGACTTTTAAAGTTCTCATGTGGCARR-019 CGACTGGGTAACACGACTTTTAAAGTTCTCATGTGGCAAGAGACTTTTCAACGACGGTAGAACGACTTTTCAACGACGGTAAGATCACACGACTAGAACGACTTTTAAAGTTCTCCATGTGGCARR-019 CGACTGGGTAACACGACTTTAAAGTTCTCAACGACGGTAAGATCCACACGACAAGATCACACACGAGAACACTTTTCGCCCCGAAGAACGTTTTC CAATGATGACACCTTTTAAAGTTCTCATGTGGCARR-019 CGACTGGGTAACACGACTACACGACTAGAACGACTTTTCGCCCCGAAGAACGTTTTC CAATGATGACACCTTTTAAAGTTCTCATTGTGGCAAGAACGTTTTCCCCCCGAAGAACGTTTTC CAATGATGACACCTTTTAAAGTTCTCATTGTGGCAAGAACGTTTTCCCCCCGAAGAACGTTTTC CAATGATGACACCTTTTAAAGTTCTCATTGTGGCA	bla(M)-004c	corr		
AmpR-017 AmpR-021 AmpR-022 AmpR-031 AmpR-032 AmpR-031 AmpR-032 AmpR-033 AmpR-033 AmpR-033 AmpR-034 AmpR-035 AmpR-036 AmpR-036 AmpR-037 AmpR-037 AmpR-037 AmpR-037 AmpR-038 AmpR-038 AmpR-038 AmpR-038 AmpR-039 AmpR-039 AmpR-039 Caafragafracatcgaafracatcgaafracatcgaafractragaafra			** ** ****** ****** ** ****** ******* *	****** *****
AmpR-017 AmpR-021 AmpR-022 AmpR-031 AmpR-032 AmpR-031 AmpR-032 AmpR-033 AmpR-033 AmpR-033 AmpR-034 AmpR-035 AmpR-036 AmpR-036 AmpR-037 AmpR-037 AmpR-037 AmpR-037 AmpR-038 AmpR-038 AmpR-038 AmpR-038 AmpR-039 AmpR-039 AmpR-039 Caafragafracatcgaafracatcgaafracatcgaafractragaafra				
AmpR-011   GAGTGGGTTACATCGAACTGGATCTCAACAGGGGTAAGATCTTTGGCCCCGAAGACGTTTTC-				
AmpR-010 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCGAAAGACGTTTTCC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-014 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTCC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-019 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGAACGTTTTCC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-020 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACCTTTTAAAGTTCTGCTATGTGC AmpR-014 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-011 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-002 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-018 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-003 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTCTCTGCTATGTGCC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTCTCTGCTATGTGCC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTCTCTCTATCTGCCC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCTTTTC — CAATGATGAGCACTTTTAAAGTCTCTCTATCTGCC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCTTTTC — CAATGATG			CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC	CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC
AmpR-010 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCGAAAGACGTTTTCC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-014 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTCC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-019 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGAACGTTTTCC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-020 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACCTTTTAAAGTTCTGCTATGTGC AmpR-014 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-011 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-002 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-018 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-003 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTCTCTGCTATGTGCC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTCTCTGCTATGTGCC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTCTCTCTATCTGCCC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCTTTTC — CAATGATGAGCACTTTTAAAGTCTCTCTATCTGCC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCTTTTC — CAATGATG			CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC	CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC
AmpR-010 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCGAAAGACGTTTTCC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-014 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTCC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-019 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGAACGTTTTCC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-020 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACCTTTTAAAGTTCTGCTATGTGC AmpR-014 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-011 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-002 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-018 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-003 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTCTCTGCTATGTGCC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTCTCTGCTATGTGCC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTCTCTCTATCTGCCC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCTTTTC — CAATGATGAGCACTTTTAAAGTCTCTCTATCTGCC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCTTTTC — CAATGATG			AGAGTGGGCTAMATCGAACTGGATCTCAAMAGCGGCAAGATCCTTGAGMCTTTCCGCCCCGAAGAACGATTCC	Ceatgatgagcacttteaaagtactgctatgtggc
AmpR-010 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCGAAAGACGTTTTCC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-014 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTCC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-019 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGAACGTTTTCC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-020 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACCTTTTAAAGTTCTGCTATGTGC AmpR-014 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAAGACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-011 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-002 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-018 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-003 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTCTCTGCTATGTGCC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTCTCTGCTATGTGCC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCGTTTTC — CAATGATGAGCACTTTTAAAGTCTCTCTATCTGCCC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCTTTTC — CAATGATGAGCACTTTTAAAGTCTCTCTATCTGCC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCCGAAAGACCTTTTC — CAATGATG			CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC	CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC
AmpR-004 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTTCTGCTATCTGCAMPR-019 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC —CAATGATGAGCACTTTTAAAGTTTCTGCTATCTGCCAMPR-014 CCAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTCC ——CAATGATGAGCACTTTTAAAGTTTCTGCTATCTGCCAMPR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTCC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTGCCAMPR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGAATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTGGCAMPR-011 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGAATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTGGCAMPR-011 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTGGCAMPR-018 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTGGCAMPR-018 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTGGCAMPR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTTGGCAMPR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGGCGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTTGGCAMPR-008 CGAGTGGGGTTACATCGAACTGGATCTCAACAGGCGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTTGGCAMPR-006 CGAGTGGGTTACATCTCAACAGGCGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTTGGCAMPR-006 CGAGTGGGTTACATCAACAGGCGTAAGATCCTTCAACAGGCTAAGATCCTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTTGGCAATGAGCACTTTAAAGTTCTCCTAACTGGCAATGAGACTTTTCCCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTGGCAATGAGCACTTTTAAAGTTCTCCTAACTGGCAATGAGACTTTCCAACAGGCGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTGGCAATGGGCACTTTAAAGTTCTCCAACTGGGCTAAGATCCTTAAGAGTCTCCACAGGGGTAAGATCCTTTGAGAGTTTTCCCCCCGAAGAACGTTTTC ——CAATGATGAGCACTTTTAAAGTTCTCCTATCTGGCAATGGGCACTTTTAAAGTTCTCCTA			CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTTC	CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC
AmpR-019 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTAGAGGTTTTCGCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTCCTATGTGC AmpR-020 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTAGAGGTTTTCGCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-014 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTC — — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTC — — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-011 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-018 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-018 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-003 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-005 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-006 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-012 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTCTATGTGCC AmpR-012 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTCCTATGTGCC AmpR-012 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTCCTATGTGCC AmpR-012 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTC — — CAATGATGAGCACTTTTAAAGTTCTCCTATTGGCC bla (M) - 001 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCCCCCCCGAAGAAC			CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC	CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC
AmpR-010 AmpR-017 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTC CGATGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC CGATGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC CGATGAGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC CGATGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC CGATGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC CGATGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC CGATGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC CGATGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AMPR-018 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AMPR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AMPR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AMPR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AMPR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AMPR-012 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AMPR-012 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AMPR-012 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTTGAGAGTTTTCCGCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AMPR-013 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTTGAGAGTTTTCCGCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC  bla(M)-001 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTC CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC  bla(M)-003 CGAGTGGGTTACATCGAACTGGATCCAACAGCGGTAAGATCCT				
AmpR-014 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCTTTGAGCCCCGAAGAACGTTCC				
AmpR-007 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC ————————————————————————————————	AmpR-020		CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTCAGCTGGC	AGGTGGCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC
AmpR-015 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-002 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-018 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-005 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-006 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-012 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC AmpR-023 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC bla(M)-001 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC bla(M)-002 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC bla(M)-003 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC bla(M)-004 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC bla(M)-005 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC bla(M)-006 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCTGCTATGTGC bla(M)-007 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC — CAATGATGAGCACTTTTAAAGTTCT	AmpR-014		CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTCTC	CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC
AmpR-012 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC ————————————————————————————————	AmpR-007		CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC	CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC
AmpR-002 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC	AmpR-015			
AmpR-002 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC	AmpR-011		CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC	CAATGATGAGCACTTTGAAAGTTCTGCTATGTGGC
AmpR-018 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGGTTTTC	AmpR-002		CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC	CAATGATGAGCACTTTTAAAGTTCTGCTATGTGAT
AmpR-003 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC ————————————————————————————————	AmpR-018		CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTC	CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC
AmpR-008 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC ————————————————————————————————			CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC	CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC
AmpR-005 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC				
AmpR-006 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC ————————————————————————————————				
AmpR-009 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC				
AmpR-012 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC				
AmpR-023 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC				
bla(M)-001 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC				
bla(M)-002 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC bla(M)-003 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AMPR-007corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AMPR-009corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC bla(M)-001corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC				
bla(M)-004 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-007corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC AmpR-009corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGCC bla(M)-001corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC				
bla(M)-004 CGAGTGGGTTACATCGAACTGGATCTACAGGGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGCCAMPR-007corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCCCCGAAGAACGTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCCCCGAAGAACGTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCCCCGAAGAACGTTTC				
AmpR-007corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC AmpR-009corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC bla(M)-003corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC bla(M)-004corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC bla(M)-004corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC				
AmpR-009corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC bla(M)-001corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC bla(M)-004corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC		r		
bla(M)-001corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC bla(M)-004corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC bla(M)-004corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCCCCCCGAAGAACGTTTTC				
bla(M)-003corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC bla(M)-004corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC				
bla(M)-004corr CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC				
	DIA(M)-0040	110:		

Size	# Var	Sources*
(bp)	bp	
861	0	Most
861	2	In, No, GE, Cl, Ne
861	1	No,GS,MCSG
861	102	Pr,Lu
861	1	Cl,Or,In
861	1	Ev,Qi
861	4	Th,SG
861	8	In
861	2	2659436
873	0/12i	Cl
861	2	7669286
858	0/3đ	Aq
861	5	Lu
861	2	Pr
861	9	In
861	7	Ox
861	1	15782178
861	1	No,Or
861	2	In
861	3	IMAGE
861	1/3d3i	8242750
861	1	In
861	4	MCSG
789	1/6d	12071693
795	0	In
789	1/6d	No
792	0/3đ	In
870	0	
861	1	
861	1	
861	1	
837	0	

AmpR-017 AmpR-001 AmpR-021 AmpR-016 AmpR-022 AmpR-010 AmpR-004 AmpR-019 AmpR-020 AmpR-014 AmpR-007 AmpR-015 AmpR-011 AmpR-002 AmpR-018 AmpR-003 AmpR-008 AmpR-005 AmpR-006 AmpR-009 AmpR-012 AmpR-023 bla(M)-001 bla(M)-002 bla(M)-003 bla(M)-004AmpR-007corr AmpR-009corr bla(M)-001corr bla(M)-003corr bla(M)-004corr GCGGTATTATCCCGTTTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGCGGT<mark>C</mark>TTCTCCCGTATAGACGCCGGGCAAGAGCA<mark>CCTT</mark>GGTCGCCG<mark>T</mark>ATACACTA<mark>C</mark>TCACAAAA<mark>G</mark>GACTTGGTTGAGTACTC<mark>C</mark>CC<mark>C</mark>GTCAC<mark>C</mark>GAAAAGCATCTTACGGATGGCATGACT  $\tt GCGGTATTATCCCGT\_TTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA$  ${\tt GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA$  $\mathbb{R}^{\mathsf{C}}$  CTATTATCCCGTATTGACGCCGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA  $\overline{ ext{GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA$  ${\tt GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA$ GCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA  $\texttt{GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCT\_ACGGATGGCATGACA$ ACAC TATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTA<mark>T</mark>TCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATAC CTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA GCGGTATTATCCCGTTTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA GCGGTATTATCCCGT: ATGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA GCGGTATTATCCCGT: TTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA  ${\tt GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA$ GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA  ${\tt GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA$ 

AmpR-013 AmpR-017 AmpR-001 AmpR-021 AmpR-016 AmpR-022 AmpR-010 AmpR-004 AmpR-019 AmpR-020 AmpR-014 AmpR-007 AmpR-015 AmpR-011 AmpR-002 AmpR-018 AmpR-003 AmpR-008 AmpR-005 AmpR-006 AmpR-009 AmpR-012 AmpR-023 bla(M)-001 bla(M)-002 bla(M)-003 bla(M)-004 AmpR-007corr AmpR-009corr bla(M)-001corr bla(M)-003corr

bla(M)-004corr

GTAAGAGAATT<mark>E</mark>TG<mark>T</mark>AGTGCTGCCAT<mark>T</mark>ACCATGAG<mark>T</mark>GA<mark>T</mark>AACAC<mark>T</mark>GCGGCCAACTTACTTCTGACAACGATCGGAGG<mark>TCCTA</mark>AAGGAGCT<mark>T</mark>ACTGTTTTTTTCATAATATGGGTGATCAT 

AmpR-017 AmpR-001 AmpR-021 AmpR-016 AmpR-022 AmpR-010 AmpR-004 AmpR-019 AmpR-020 AmpR-014 AmpR-007 AmpR-015 AmpR-011 AmpR-002 AmpR-018 AmpR-003 AmpR-008 AmpR-005 AmpR-006 AmpR-009 AmpR-012 AmpR-023 bla(M)-001 bla(M)-002 bla(M)-003 bla(M)-004AmpR-007corr AmpR-009corr bla(M)-001corr bla(M)-003corr

bla(M)-004corr

GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGACGACGATGCCTGGAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA  ${\tt GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTG} {\tt aGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA}$ GT<mark>E</mark>AC<mark>ECCECTTGAECGETGGGAACCEGGAGETGAAE</mark>GAAGCCATACC<mark>E</mark>AACGACGAGCGTGA<mark>E</mark>AACCACGATGCC<mark>E</mark>GTAGCAATGGCEAAACE<mark>E</mark>AAACTAETEACTGCGAACTE  $\tt GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA$  $\mathtt{GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGC{}^{\mathtt{C}}$  $\mathtt{GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGACGACGATGCCTGTAGCAATG @ \mathtt{CAACAACGTTGCGCAAACTATTAACTGGCGAACTA}$  ${ t GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGGGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA$ GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGACGACGACCACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGACGTGACACCACGATGCCTGTAGCAATGGCAACACGTTGCGCAAACTATTAACTGGCGAACTA GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGGTGTGACACCACGATGCCTGTAGCAATG\_CAACAACGTTGCGCAAACTATTAACTGGCGAACTA GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACCTTTGCGTAAACTATTAACTGGCGAACTA ${ t GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGACGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA$ GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGACGACGACGCTGACACCACGATGCCAGGAGCAACGACGACGATGCGCAAACTATTAACTGGCGAACTA GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGGGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA  $\tt GTAACTCGCCTTGATCGTTGGGAACCG\_AGCTGAATGAAGCCATACCAAACGACGGACCGTGACACCACGATGCCTGTAGCAATGGCAACCACGTTGCGCAAACTATTAACTGGCGAACTA$ GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGACGACGATGCCTGGAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTG\_AGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA  $\tt GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA$ GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACACGTTGCGCAAACTATTAACTGGCGAACTA GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGGGTGACACCACGATGCCTGTAGCAATGGCAACGATGCGCAAACTATTAACTGGCGAACTA  $\tt GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGACGACGACGACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA$ GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGGGTGACACCACGATGCCTG AGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA 

AmpR-013 AmpR-017 AmpR-001 AmpR-021 AmpR-016 AmpR-022 AmpR-010 AmpR-004 AmpR-019 AmpR-020 AmpR-014 AmpR-007 AmpR-015 AmpR-011 AmpR-002 AmpR-018 AmpR-003 AmpR-008 AmpR-005 AmpR-006 AmpR-009 AmpR-012 AmpR-023 bla(M)-001 bla(M)-002 bla(M)-003 bla(M)-004 AmpR-007corr AmpR-009corr bla(M)-001corr bla(M)-003corr bla(M)-004corr CTTACTCTAGC<mark>E</mark>TC<mark>E</mark>CG<mark>E</mark>CA<mark>ECTE</mark>ATAGACTGGATGGAGGGCGGA<mark>E</mark>AAAGTEGCAGGACCACTTCT<mark>E</mark>CGCTCGCCCT<mark>C</mark>CCTGGCTGGTTCATTGCTGAEAAATC<mark>E</mark>GG<mark>E</mark>GCCGGT CTTACTCTAGCTTCCCGGCAACAATTAATAGACTCAATGAAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGTTTATTGCTGATAAATCTGGAGCCGGT  $\mathtt{CTTACTCTAGCTTCCCGGCAACA}$  $\tt CTTACTCTAGCTTCCCGGCAACA \overline{A}TTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGTTTATTGCTGATAAATCTGGAGCCGGT$ 

```
GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-017
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-001
                 	ag{2}
AmpR-021
                 GA<mark>RCGEGGO</mark>TCTCGCGG<mark>O</mark>ATCATTGC<mark>E</mark>GCECTGGGGCC<mark>E</mark>GATGGTAAGCCCTC<mark>R</mark>CG<mark>R</mark>ATCGTAGTEATCTACACGACGGGGAGTCAGGC<mark>O</mark>ACTATGGA<mark>O</mark>GAACGAAATAG<mark>E</mark>CAGATCGCT
AmpR-016
                 GAGCGTGGCTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-022
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAGCTATGGATGAACGAAATAGACAGATCGCT
AmpR-010
                 GAGCGTGG<mark>T</mark>TCTCGCGGTATCAT<mark>G</mark>GCAGCGCTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-004
                 GAGCGTGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGG AGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-019
AmpR-020
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-014
AmpR-007
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGAGGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-015
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCG ATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-011
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-002
AmpR-018
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-003
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-008
                 GAGCGTGG_{	ext{T}}^{	ext{T}}TCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
                 GAGCGTGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-005
AmpR-006
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-009
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-012
AmpR-023
                 {\tt GAGCGTGG} \\ \hline {\tt TCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGAGGTCAGCCAACTATGGATGAACGAAATAGACAGATCGCT}
bla(M)-001
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGAGGCAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
bla(M)-002
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
bla(M)-003
                 GAGCGTGG^{T}TCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
bla(M)-004
AmpR-007corr
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
AmpR-009corr
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
bla(M)-001corr
                 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
                 {\tt GAGCGTGG}{\color{red}{\bf T}CTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT}
bla(M)-003corr
bla(M)-004corr
                 GAGCGTGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
                 AmpR-013
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA------
                                                                                        = synonymous
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
AmpR-017
                                                                                        = conservative/semi-conservative amino acid substitution
AmpR-001
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA------
                                                                                        = non-conservative amino acid substitution
AmpR-021
                 GAGAT<mark>C</mark>GGTGCCTCACTGAT<mark>C</mark>AAGCA<mark>C</mark>TGGTAA-----
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
AmpR-016
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
AmpR-022
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
AmpR-010
AmpR-004
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-------
AmpR-019
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA------
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
AmpR-020
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA------
AmpR-014
AmpR-007
                 GAGATAGGTGCCTCACTGATTAAGCATTGG----
                 AmpR-015
AmpR-011
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
AmpR-002
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
AmpR-018
AmpR-003
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
AmpR-008
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
AmpR-005
AmpR-006
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA------
AmpR-009
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAACTG-------
AmpR-012
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA------
AmpR-023
                 bla(M)-001
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
bla(M)-002
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
bla(M)-003
bla(M)-004
                 GAGATAGGTGCCTCACTGATTAAGCATTGG-----
AmpR-007corr
                 GAGATAGGTGCCTCACTGATTAAGCATTGGCCTCAGGATTAG-----
AmpR-009corr
                 bla(M)-001corr
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA-----
                 GAGATAGGTGCCTCACTGATTAAGCATTGGTAA------
bla(M)-003corr
bla(M)-004corr
                 {\tt GAGATAGGTGCCTCACTGATTAAGCATTGGCTGTTATCAACAAGTTTGTACAAAAAAGCTGAACGAGAAACGTAA}
                 .... ......... ....
AmpR gene is bla (β-lactamase) from E. coli
bla(M) lacks the signal peptide
Sequences in bold are incorrectly annotated
AmpR-009 was identified by coordinates that were +3 from the actual start codon and is actually identical to AmpR-016 - -009 merged with -016
AmpR-002 is probably identical to AmpR-004 but the sequence contains an in-frame STOP codon. This is likely a sequence error as AmpR is only bacterial selection marker
for the plasmids carrying AmpR-002 - -002 merged with -004
bla(M)-001 coordinates did not start at the START codon. It is identical to AmpR-001 and has been merged with it
Similarly bla(M)-003 is identical to AmpR-008 and has been merged with it
```

Name of Variant	Alignment of Translated Protein	Size	# Var aa
AmpR-013	MSIOHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDOLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGOEOLGRRIHYSONDLVEYSPVTEKHLTDGMT	<b>(aa)</b> 286	<b>aa</b> 0
AmpR-017	MSIQHFRVALIPFFABFCLPVPAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPPERFFMMSTFKVLLCGAVLSRÜDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	286	2
AmpR-001	MSIQHFRVALIFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFFMMSTFKVLLCGAVLSRIDAGGEOLGRRIHYSONDLVEYSPYTEKHLTDGMT	286	1
AmpR-021	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSONDLVEYSPVTEKHLTDGMT	286	0
AmpR-016	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSR DAGGEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	286	1
AmpR-022	MSIOHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDOLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGOEOLGRRIHYSONDLVEYSPVTEKHLTDGMT	286	1
AmpR-010	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDOLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGOEOLGRRIHYSONDLVEYSPVTEKHLTDGMT	286	0
AmpR-004	MSIOHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDOLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCDTLLSRIDAGGEOLGRRIHYSONDLVEYSPVTEKHLTDGMT	286	6
AmpR-019	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	286	1
AmpR-020	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERF <mark>OVEVA</mark> MMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	290	1/4i
AmpR-014	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERSPMMSTFKVLLCGAVLSR DAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	286	2
AmpR-007corr	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDOLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGOEOLGRRIHYSONDLVEYSPVTEKHLTDGMT	286	0
AmpR-015	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILES RPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	286	3
AmpR-011	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	286	0
AmpR-002	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCDTLLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	166	5/79d
AmpR-018	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	286	0
AmpR-003	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRI YSQNDLVEYSPVTEKHLTDGMT	286	1
AmpR-008	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	286	0
AmpR-005	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSR\DAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	286	2
AmpR-006	$ exttt{MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSR exttt{D}DAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT}$	286	2
AmpR-012	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	286	1
AmpR-023	MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	286	1
bla(M)-002	MMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	264	0
bla(M)-004corr	MMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT	278	0
	***************************************		

 $V \tt RELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAG$ AmpR-013 AmpR-017  ${\tt VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMP\_AMATTLRKLLTGELLTLASROOLIDWMEADKVAGPLLRSALPAGWFIADKSGAG$  ${\tt VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMP\_AMATTLRKLLTGELLTLASROOLIDWMEADKVAGPLLRSALPAGWFIADKSGAG$ AmpR-001 AmpR-021 VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASROOLIDWMEADKVAGPLLRSALPAGWFIADKSGAG AmpR-016  $V {\tt RELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAG$ AmpR-022  $V {\tt RELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAG$ AmpR-010 VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASROOLIDWMEADKVAGPLLRSALPAGWFIADKSGAG VKELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDE<mark>S</mark>DTTMPVAM<mark>E</mark>TTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAG AmpR-004 AmpR-019 VRELCSAAITMSDNTAANLLLTTIGGPKELTAF HNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAG AmpR-020  ${\tt VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAG$ VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASROOLIDWMEADKVAGPLLRSALPAGWFIADKSGAG AmpR-014 AmpR-007corr  $\tt VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQOLIDWMEADKVAGPLLRSALPAGWFIADKSGAG$ AmpR-015  ${\tt VRELCSAAITMSDNTAANLLL}{\tt TIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAG{\tt CLRSALPAGWFIADKSGAG}$ AmpR-011  ${\tt VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAG$ AmpR-002 vrelcsaaitmsdntaanlllttiggpkeltaflhnmgdhvtrldrwepelneaipnde<mark>s</mark>dttmpvan<mark>e</mark>ttlrklltgelltlasr<u>oo</u>lid<mark>-meadkvagpllrsalpagwfiadksgag</mark> AmpR-018  ${\tt VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQDLIDWMEADKVAGPLLRSALPAGWFIADKSGAG$ AmpR-003  $\tt VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAG$ AmpR-008  $V {\tt RELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAG$ VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASROOLIDWMEADKVAGPLLRSALPAGWFIADKSGAG AmpR-005 AmpR-006  $\tt VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMP{\verb|}| AMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAG$ AmpR-009corr  ${\tt VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPV} {\tt MATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAG}$ AmpR-012 VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPKLNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAG AmpR-023 VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMP AMATTLRKLLTGELLTLASROOLIDWMEADKVAGPLLRSALPAGWFIADKSGAG VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASROOLIDWMEADKVAGPLLRSALPAGWFIADKSGAG bla(M)-002 bla(M)-004corr VRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASROOLIDWMEADKVAGPLLRSALPAGWFIADKSGAG 

AmpR-013	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-017	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-001	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-021	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-016	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-022	ERGSRGIIAALGPDGKPSRIVVIYTTGSQAAMDERNRQIAEIGASLIKHW
AmpR-010	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-004	ERGSRGIIAALGPDGK <mark>R</mark> SRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-019	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-020	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-014	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-007corr	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHWPQD
AmpR-015	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-011	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-002	ERGSRGIIAALGPDGK <mark>R</mark> SRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-018	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-003	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-008	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-005	ERGSRGIIAALGPDGK <mark>R</mark> SRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-006	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-009	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-012	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
AmpR-023	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
bla(M)-002	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
bla(M)-004corr	ERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHWLLSTSLYKKAERET
	*********** ********************

Key

```
ATGC = change in nucleotide relative to consensus sequence — synonymous
ATGC = change in nucleotide relative to consensus sequence — conservative/semi-conservative amino acid substitution
ATGC = change in nucleotide relative to consensus sequence — non-conservative amino acid substitution
ATGC = predicted protein sequence following a nonsense mutation
```

CmR-005

CmR-003

CmR-006

CmR-007 CmR-008

CmR-009

CmR-009corr

Name of Variant	# of Occur	Alignment	Size (bp)	# Var bp
CmR-013	48	ATGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG	660	0
CmR-016	39	ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG	660	1
CmR-015	31	$\tt ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG$	660	2
CmR-014	29	ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG	660	1
CmR-010	21	$\tt ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAAGAACATTTTGAGGCATTTCAGTCAG$	660	1
CmR-004	13	ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG	660	2
CmR-011	12	ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG	660	1
CmR-002	4	ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG	660	1
CmR-012	4	$\tt ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAAGAACATTTTGAGGCATTTCAGTCAG$	660	3
CmR-001	3	$\tt ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG$	660	1
CmR-005	2	ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG	660	2
CmR-003	1	ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG	660	2
CmR-006	1	$\tt ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG$	660	3
CmR-007	1	ATGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG	660	2
CmR-008	1	ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG	681	1/1d
CmR-009	1	ATGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTTGAGGCATTTCAGTCAG	654	6d
CmR-009co	orr	ATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG	684	1d
		***************************************		
CmR-013		ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCCGTATGGCAATGAAAGACGGT		
CmR-016		ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCCGTATGGCAATGAAAGACGGT		
CmR-015		ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATT		
CmR-014		ATTAGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCGCCTGATGAATGCTCATCCGGAATTGCTATGGCAATGAAGACACGT		
CmR-010		ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGACTTCCGTATGGCAATGAAAGACGGT		
CmR-004		ATTACGGCCTTTTTTAAAGACGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGA_TTCCGTATGGCAATGAAAGAACACGT		
CmR-011		ATTAGGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCGCCTGATGAATGCTCATCCGGAATTCCGTATGGCAATGAAGAACACGT		
CmR-002		ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCGGCCTGATGAATGCTCATCCGGAATTCCGTATGGCAATGAAAGACGGT		
CmR-012		ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGA <mark>C</mark> TTCCGTATGGCAATGAAAGACGGT		
CmR-001		ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTTCGTATGGCAATGAAAGACGGT		
CmR-005		ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGACTTCCGTATGGCAATGAAAGACGGT		
CmR-003		ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTTCGTATGGCAATGAAAGACGGT		
CmR-006		ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAETTCCGTATGGCAATGAAAGACGGT		
CmR-007		ATTAGGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCGCCTGATGAATGCTCATCCGGAATTCCGTATGGCAATCAAAGACGGT		
CmR-008		ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCGCCTGATGAATGCTATCCGGAATTCCGTATGGCAATGAAAGACGGT		
CmR-009		ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCCCATTCTTGCCCGCCTTGATGAATGCTCATCCGGAATTCCGTATGGCAATGAAGAACACGGT		
CmR-009co	rr	ATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCCGTATGGCAATGAAAGACGGT		
Chin-003CO	,11	11.10.00.00.11.11.11.10.10.11.11.10.10.1		
CmR-013		GAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTTCACACATATAT		
CmR-016		GAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATAT		
CmR-015		GAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATAT		
CmR-014		GAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGAGTTTTCCGGCAGTTTTCTACACCATATAT		
CmR-010		GAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATAT		
CmR-004		GACCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATAT		
CmR-011		GAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGGAAACGTTTTCATGGCTCTGGAGTGAATACCAGACGATTTCCAGCGAGTTTCTACACCATATAT		
CmR-002		GAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGCGCAGTTTCTACACCATATAT		
CmR-012		GAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGGAAACGTTTTCATGGCTCTGGAGTGAATACCAGACGATTTCCAGCGAGTTTCTACACCATATAT		
CmR-001		GAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACGTTTCCATGAGCGTCTGGAGTGAAATACCACGACGATTTCCACCGCATTTCTACACCATTATAT		
5mm-001		CASCISSION CONTROL CON		

 ${\tt GAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACTGTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGCCAGTTTCTACACCATATAT$ 

 ${\tt GAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTTCACACATATAT\\$ 

 ${\tt GAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATAT\\$ 

 Sources\* In,No,Ne In,No,Ne Qi,Ev,Ta Qi,Cl,Lu Or, IMAGE Ca No 16377204 MCSG Ag Or Qi No In In In

```
CmR-013
              {\tt TCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTG
CmR-016
              CmR-015
              TCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTTAAACGTG
CmR-014
              CmR-010
              CmR-004
              CmR-011
              {\tt TCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTG}
CmR-002
              {\tt TCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTG}
CmR-012
              {\tt TCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTG}
CmR-001
              {\tt TCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTG}
CmR-005
              {\tt TCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGT{\tt TCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTG
CmR-003
              CmR-006
CmR-007
              CmR-008
              TCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTTAAACGTG
CmR-009
              TCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTG
CmR-009corr
              {\tt GCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTGTGATGGCTTCCAT
CmR-013
CmR-016
              GCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTGTGTGATGGCTTCCAT
              {\tt GCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGT{\tt GTGATGGCTTCCAT
CmR-015
              {\tt GCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTTGTGATGGCTTCCAT
CmR-014
CmR-010
             CmR-004
              \texttt{GCCAATATGGACAACTTCTTCGCCCCCGTTTTCAC} \texttt{ATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTGTGATGGCTTCCAT
              GCCAATATGGACAACTTCTTCGCCCCCGTTTTCAC<mark>T</mark>ATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTGTGATGGCTTCCAT
CmR-011
              {\tt GCCAATATGGACAACTTCTTCGCCCCGTTTTCAC}^{f c} {\tt ATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTGTGATGGCTTCCAT
CmR-002
              GCCAATATGGACAACTTCTTCGCCCCCGTTTTCAC<mark>L</mark>ATGGGCAAATA<mark>C</mark>TATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGT<mark>O</mark>TGTGATGGCTTCCAT
CmR-012
CmR-001
             {\tt GCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCATCATGCCGTTTGTGATGGCTTCCAT
CmR-005
              {\tt GCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTGTGATGGCTTCCAT
CmR-003
              {\tt GCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTGTGATGGCTTCCAT
              GCCAATATGGACAACTTCTTCGCCCCCGTTTTCAC<mark>T</mark>ATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGT<mark>T</mark>GTGATGGCTTCCAT
CmR-006
CmR-007
              GCCAATATGGACAA
TTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGT
<math>TGTGATGGCTTCCAT
CmR-008
             {\tt GCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTGTGATGGCTTCCAT\\
CmR-009
              GCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTGTGATGGCTTCCAT
              {\tt GCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTGTGATGGCTTCCAT
CmR-009corr
CmR-013
              GTCGGCAGA ATGCTTA ATGA ATTACA AC AGTACTGCGATGA GTGGCAGGCGGGCGTA A ---------------
CmR-016
             CmR-015
              CmR-014
              CmR-010
              CmR-004
CmR-011
              CmR-002
              CmR-012
              CmR-001
             GTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAA------
CmR-005
              CmR-003
              CmR-006
              CmR-007
CmR-008
              GTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGG -GGGGCGTAA<mark>TCTA---GAGGATCCGGCTTACTAA</mark>
CmR-009
             GTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGG GGGG ----
             GTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGG-GGGGCGTAAACGCCGCGTGGATCCGGCTTACTAA
CmR-009corr
              *******************************
CmR gene is cat (chloramphenicol actyltransferase) from E. coli
Sequences in bold are incorrectly annotated
CmR-009 missing STOP codon. Corrected sequence still unique variant
Key
          = change in nucleotide relative to consensus sequence - synonymous
          = change in nucleotide relative to consensus sequence - conservative/semi-conservative amino acid substitution
```

= change in nucleotide relative to consensus sequence - non-conservative amino acid substitution

Key

Name of	Alignment of Translated Protein	Size	# Var
Variant		(aa)	aa
CmR-013	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-016	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-015	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-014	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-010	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-004	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-011	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-002	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-012	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-001	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-005	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPE RMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	1
CmR-003	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-006	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-007	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	219	0
CmR-008	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	226	2
CmR-009corr	MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLHIY	227	1
	***************************************		
CmR-013	${\tt SQDVACYGENLAYFPKGFIENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGA$		
CmR-016	${\tt SQDVACYGENLAYFPKGF1ENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGA$		
CmR-015	SQDVACYGENLAYFPKGF1ENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLA1QVHHAVCDGFHVGRMLNELQQYCDEWQGGA		
CmR-014	SQDVACYGENLAYFPKGF1ENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLA1QVHHAVCDGFHVGRMLNELQQYCDEWQGGA		
CmR-010	${\tt SQDVACYGENLAYFPKGF1ENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGA$		
CmR-004	SQDVACYGENLAYFPKGFIENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGA		
CmR-011	${\tt SQDVACYGENLAYFPKGF1ENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGA$		
CmR-002	SQDVACYGENLAYFPKGFIENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGA		
CmR-012	SQDVACYGENLAYFPKGF1ENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLA1QVHHAVCDGFHVGRMLNELQQYCDEWQGGA		
CmR-001	SQDVACYGENLAYFPKGFIENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGA		
CmR-005	SQDVACYGENLAYFPKGFIENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGA		
CmR-003	${\tt SQDVACYGENLAYFPKGF1ENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGA$		
CmR-006	${\tt SQDVACYGENLAYFPKGF1ENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGA$		
CmR-007	${\tt SQDVACYGENLAYFPKGF1ENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGA$		
CmR-008	$\mathtt{SQDVACYGENLAYFPKGF1ENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQ^GR\mathbf{NLE-DPAY}$		
CmR-009corr	SQDVACYGENLAYFPKGFIENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTÖGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGRKRRVDPAY		

Name of	# of	Alignment
Variant	Occur	
HygR-010	35	ATC A A A A COCACA A CTC A COCAU A CARACTURA DA A A PETECTA DECACA A CTTC A CACACTURA CACACT
		atgaa <mark>c</mark> aagcc <mark>c</mark> gaactcaccgc <mark>r</mark> ac <u>eagc</u> gt <mark>r</mark> galaaltttctcatcgacgagttcgacagtgtgacctgatgcagttctcgacgcgagagaggcgaaga <mark>gagc</mark> cgagcttagcttc
HygR-012	26	ATGAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACTGATGCAGCTCTCGGAGGGCCAAGAATCTCGTGCTTTCAGCTTC
HygR-008	11	ATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGACAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTC
HygR-005	7	AAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTC
HygR-002	4	ATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTC
HygR-011	4	
	_	
HygR-004	3	ATGGGTAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCCAAGAATCTCGTGCTTTCAGCTTC
HygR-007	3	AAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTC
HygR-001	2	ATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTC
hphMX6*	2	ATGGGTAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTC
HygR-003	1	$ATG\frac{GG}{GAAGCCTGA} \stackrel{C}{\operatorname{CT}}ACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTC$
	1	
HygR-006		ATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTC
HygR-009	1	ATGAAAAAGCCTGAACTCACCGCGACGTCTGT <mark>T</mark> GAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCCGAAGAATCTCGTGCTTTCAGCTTC
HygR-013	1	ATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACACGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTC
HygR-013c	orr	ATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTC
		**** ** * **** ** ** ** ** ** *** **** ****
HygR-010		GATGT <mark>GGGGGGTCGGGGTTATGTT</mark> CTGCGGGT <mark>G</mark> AATAGCTGCGC <mark>T</mark> GATGG <mark>C</mark> TTCTACAAAGA <mark>CCG</mark> TA <mark>G</mark> GT <mark>G</mark> TA <mark>C</mark> CG <mark>C</mark> CACTT <mark>GGCGAGCGCTT</mark> CC <mark>G</mark> ATGCC <mark>G</mark> GAAGTG <mark>TTE</mark> GAC
		GATGTRAGGAGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCCTCCCGATTCCGGAACTGCTTCAC
HygR-012		
HygR-008		GATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGAC
HygR-005		GATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGAC
HygR-002		${\tt GATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCTCCCGATTCCGGAAGTGCTTGAC}$
HygR-011		${\tt GATGTAGGAGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCTCCCGATTCCGGAAGTGCTTGACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCTCCCCGATTCCGGAAGTGCTTGACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCTCCCCGATTCCGGAAGTGCTTTGACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCTCCCCGATTCCGGAAGTGCTTTGACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCTCCCCGATTCCGGAAGTGCTTTGACAAAGATCGTTATGTTTATCGGCACTTTTGCATCGGCCGCTCCCCGATTCCGGAAGTGCTTTGACAAAGATCGTTATGTTTATCGGCACTTTTGCATCGGCCGCTCCCGGATTCCGGAAGTGCTTTGACAAAGATCGTTATGTTTATCGGCACTTTTGCATCGGCCGCTCCCGGATTCCGGAAGTGCTTTGACAAAGATCGTTATGTTTATCGGCACTTTTGCATCGGCACTTTCCGGAAGTGCTTTGACAAAGATCGTTATGTTTATCGGCACTTTTGCATCGGCACTTCCCGGAAGTGCTTTGACAAAGATCGTTATGTTTATCGGCACTTTTGCATCGGCACTTCCCGGAAGTGCTTTGACAAAGATCGTTATGTTTATCGGCACTTTTGCATCGGCACTTTCCGGAAGTGCTTTGACAAAGATCGTTATGTTTATCGGCACTTTTGCATCGGCACTTTTGCAAAGATCGTTTATGTTTTATCGGCACTTTTGCAAAGATCGTTTATGTTTTATCGGCACTTTTGCAAAGATCGTTTATGTTTTATCGGCACTTTTGCAAAGATCGTTTATGTTTTATGTTTTTTTT$
HygR-004		GATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATCCGGAAGTGCTTGAC
HygR-007		GATGTAGGAGGGCGTGGATATGTCCTGCGGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCCTCCCGATTCCGGAAGTGCTTGAC
HygR-001		GATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCCTCCCGATTCCGGAAGTGCTTGAC
hphMX6		GATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGAC
HygR-003		GATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGAC
HygR-006		${\tt GATGTAGGAGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCTCCCGATTCCGGAAGTGCTTGACCTTGACCTTGACCTTATGTTTATCGGCACTTTTGCATCGGCCGCTCCCGATTCCGGAAGTGCTTGACCTTGACCTTGACCTTATGTTTATCGGCACTTTTGCATCGGCCGCTCCCGATTCCGGAAGTGCTTGACCTTGACCTTATGTTTATCGGCACTTTTGCATCGGCCGCTCCCGATTCCGGAAGTGCTTGACCTTGACCTTATGTTTATCGGCACTTTTGCATCGGCCGCTCCCGATTCCGGAAGTGCTTGACCTTGACCTTGACTTATGTTTATCGGCACTTTTGCATCGGCACTTCCGGAAGTGCTTGACCTTGACCTTGACCTTATGTTTATCGGCACTTTTGCATCGGCACTTCCGGAAGTGCTTGACCTTGACCTTGACCTTGACCTTGACCTTGACTTTGACTGAC$
HygR-009		GATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCTCCCGATTCCGGAAGTGCTTTGAC
HygR-013		GATGTAGGAGGGCGTGGATATGTCCTGCGGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGAC
HygR-013c	orr	GATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGAC
		**** ** ** ** ** ** ***** ******* ******
HygR-010		AT <mark>GGGGAC</mark> TTCAGCGAGAGCCTGAC <mark>ATAC</mark> TGCATC <mark>ACWACA</mark> CGGCCCAAGGGGTWACTCTCCAAGACCTCCCGGAAACAGAGCTGCCWGCTGTCWTACAGCCWGTTGCCGAAAGACTGCCTGAAACAGACCTGCCGGTGTTCTACAACCGAAGCTGTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAAGCTGTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAAGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAACCGAACTGCCCGCTGTTCTACAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACTGCCCGCTGTTCTACAAACCGAACCTGCCCGCTGTTCTACAAACCGAACCTGCCCGCTGTTCTACAAACCGAACCTGCCCGCTGTTCAAACCGAACCTGCCCGCTGTTCAAACCAAACCAAACCAAACCAAAACCAAACCAAACCAAAA
HygR-012		ATTGGGGAGTTTTAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTG
HygR-008		$\tt ATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTG$
		ATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTCACCAGGGCCATG
HygR-005		
HygR-002		ATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTG
HygR-011		ATTGGGGAATT <mark>T</mark> AGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTG
HygR-004		$\tt ATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTG$
HygR-007		ATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTG
HygR-001		ATTGGGGAGTTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCTGAAACCGAACTGCCGGTGTTCTGCAGCCGGTCGCGGAGGCCATG
hphMX6		ATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTG
HygR-003		ATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTG
HygR-006		ATTGGGGAGTTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAGACCTGCCTG
HygR-009		ATTGGGGAATT <mark>T</mark> AGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTG
HygR-013		ATTGGGGA TTTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTG
HygR-013c	orr	$\mathtt{ATTGGGGA}_{\mathbf{C}}^{\mathbf{C}}$ TTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTG
2 9		** ** ** ** ********** ** ***** * * ** ** ** ** ** *
HygR-010		GATGCTATTGCOGCOGACCTOAGTCATACOAGCGGOTTCGGCCCATTCGGCCCAAGGOATCGGCCACTACACATCGGCGGGATTTCATTTGCGCOATTGCTGATCCCCATGTC
HygR-012		GATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTCATATGCGCGATTGCTGATCCCCATGTG
HygR-008		GATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTG
HygR-005		GATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTG
HygR-002		${\tt GATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGATCGCTGATCCCCATGTGATCGCTGATCCCCATGTGATCCCCATGTGATCATACACTACATGTGCTGATCTCATATGCGCGATTGCTGATCCCCATGTGATCCCCATGTGATCACATGTGCTGATCCCCATGTGATCACATGTGATCACATGTGATCATACATGTGATCATACATGTGATCCCCATGTGATCATGTGATCCCCATGTGATCATACATGTGATCATACATGTGATCATACATGTGATCATACATGTGATCATACATGTGATCATACATGTGATCATATGTGATCATGTGATCCCCATGTGATCATACATGTGATTTCATATGTGATCATGTGATCCCCATGTGATCATACATGTGATCATACATGTGATCATATGTATATGTGATCATATGTGATCATATGTGATCATGTGATCATATGTGATCATATGTGATCATATGTGATCATATGTGATCATATGTGATCATATGTGATCATATGTGATCATATGTGATCATATGTGATCATATGTGATCATATGTGATCATATGTGATCATATGTGATTGAT$
HygR-011		GATCCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTG
		GATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGTTCGGCCCATTCGGACCCAAGGAATCGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCATTCG
HygR-004		
HygR-007		GATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTG
HygR-001		GATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTG
hphMX6		${\tt GATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTG}$
HygR-003		${\tt GATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATTGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTG}$
HygR-006		GATGCGATCGCTGCGGCCGATCTTAGCCGACGACGACGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTG
HygR-009		GATGCAATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTG
HygR-013		GATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTCATATGCGCGATTGCTGATCCCCATGTG
HygR-013c	orr	GATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTG
		**** ** ** ** ** ** ** ** ** ** ** ** *

Size

(**bp)** 1038

1026

1026

1020 1038

1020

1029

1023

1026

1029

1026

1026 1026

1023 1023 # Var

**bp** 222

0/6d

4/6d

0/3d

1/3d

1/3d

13

Sources\*

16200533

16200533

In 10890530

2194165

Pr

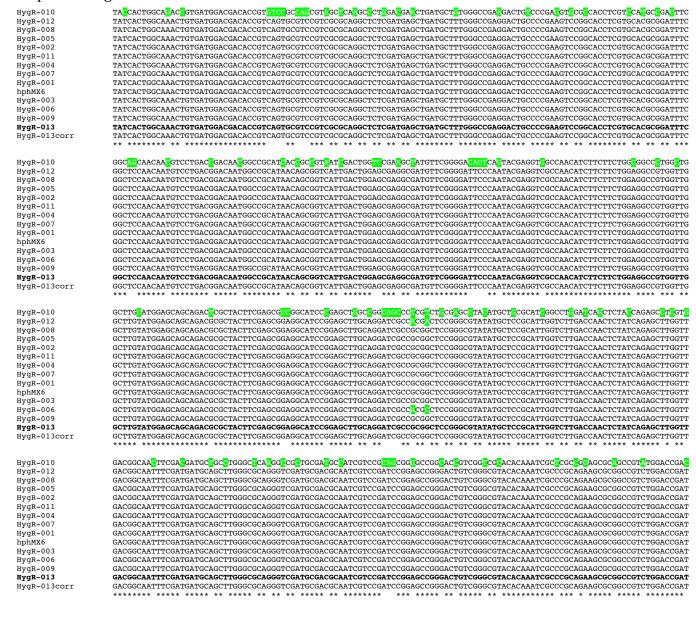
Ca In,Cl,OB

Ag

In

Cl

Or



HygR-013corr

HygR-010	ggctg <mark>ggtggaggtecte</mark> gccga <mark>o</mark> ag <mark>ggo</mark> aaccg <mark>e</mark> cgccag <mark>hacacca</mark> ccg <mark>e</mark> cgc <u>ha</u> agga <mark>echa</mark> ggtcgagtttaa
HygR-012	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCCAAAGAATAG
HygR-008	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAG
HygR-005	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAG
HygR-002	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCG <mark>GATCGG</mark> AG <mark>ATGGGC</mark> GAGGCTAACTGA
HygR-011	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAG
HygR-004	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCCAAAGGAATAA
HygR-007	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCCAAAGGAATAG
HygR-001	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCCAAAGGAATAG
hphMX6	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCCAAAGGAATAA
HygR-003	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAG
HygR-006	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCCAAAGGAATAG
HygR-009	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCCAAAGGAATAG
HygR-013	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAA——————————
HygR-013corr	GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCCAAAGGAA <mark>TAT</mark> CGATAA
	***** ** ** ** ** ***** ** ** ***** **

HygR gene is aph(4)-Ia (aminoglycoside phosphotransferase) from E. coli hphMX6 is identical to HygR-004 except it includes the Ashbya gossypii TEF1 promoter and terminator, which have been trimmed out of the sequence shown Sequences in bold are incorrectly annotated

# Var

0

0

4 2 0

0

0

0

HygR-013 missing STOP codon. Corrected sequence still unique variant

Name of	Alignment of Translated Protein	Size
Variant	Allyment of Hansatet Flotein	(aa)
HygR-010	-MKKPELTATSVEKFLIEKFDSVSDLMOLSEGEESRAFSFDVGGRGYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLODLPETELPAVLOPVAEAM	341
HygR-012	-MKKPELTATSVEKFILTEKFDSVSDLMOLSEGEESRAFSFDVGGRGYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAGGVTLODLFETELRAVLOPVAEAM	345
HygR-008	-MKKPELTATSVEKFLIEKFDSVSDLMQLSEGEGRAFSFDVGGRGYVLRVNSCADGTKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLQDLPETELPAVLQPVAEAM	341
HygR-005		342
HygR-002	-MKPBLIATSVERFILERFDSVSDLMQLSEGESRAFSFDVGGRQVLRVNSCADGFINDTVIRFFSAALPIPEVLDIGESESLTYCISRRAQGVTLQDLPETELPAVLQVVABAM -MKKPBLIATSVERFILERFDSVSDLMQLSEGESRAFSFDVGGRQVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGESESLTYCISRRAQGVTLQDLPETELPAVLQPVABAM	339
HygR-011	- INAMEDITATION OF THE PROPERTY OF THE PROPERT	341
HygR-004	MCKKPELTATSVEKFLIERFDSVSDLMQLSEGESRAFSFDVGGRQVLRVNSCADGFINDTVIRFFASALPIPEVLDIGEFSESITYCISRRAQGVTLQDLPETELPAVLQVAEAM	341
HygR-007	MONRELIATOVER LIERTEN SOLIMUNGESEESRAT STUVGRROY LAWINGLAUG I NEW 1 VERT RAFALET IF DU LIGES SELIT (I SERAÇOV LUÇULE ELETAN LEVALER MEN LA LIERTE LE LIERTE LE LEVALER MEN LA LIERTE LE LIERTE LIERTE LE LIERTE LE LIERTE LE LIERTE LE LIERTE LE LIERTE LE LIERT	341
	-MAKEPLITATSVEKFLIERFDSVSDLMQLSEGESRAFSFDVGGRQVLKVNSCADGFINDTVIRFFASALPIPEVLDIGESESLTYCISRRAQGVTLQDLPETELPAVLQVABAM -MKKPELTATSVEKFLIERFDSVSDLMQLSEGESRAFSFDVGGRQVLKVNSCADGFYKDRYVYRHFASALPIPEVLDIGESESLTYCISRRAQGVTLQDLPETELPAVLQPVABAM	341
HygR-001 hphMX6	-mareliatoverlieridosolimquistotesaa sidvogasivlakvaslalogiidasivikat saaaleilevuldee seeliitoliskaaqoviluguleeleliavulgvabam Mokkepliatovekerlieridosolimquistotesesaa sidvogasivulkvaslalogiikoryvyrheasaaleilevuldidee seelitoliskaaqoviluguleelelavulgvabam	341
	mokkebitatsvekeliekfdsvsblimusbobeska sfivorkni vlkvinslalog i dini vikir sakalpi pevuljubes sobiti liskraqoviluduje eleja. Moskepritatsvekfliekfdsvsblimusbobeska sfivorkni vlkvinska logi vlkvinska alapi pevuljubes sesitvci skraqoviluduje elejavulovabam	
HygR-003		339
HygR-006	-MKKPELTATSVEKFLIERFDSVSDLMQLSEGEESRAFSFDVGGRGYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLQDLPETELPAVLQPVAEAM	341
HygR-009	-MKKPELTATSVEKFLIEKFDSVSDLMQLSEGEESRAFSFDVGGRGYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLQDLPETELPAVLQPVAEAM	341
HygR-013corr	-MKKPELTATSVEKFLIEKFDSVSDLMQLSEGEESRAFSFDVGGRGYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLQDLPETELPAVLQPVAEAM	342
	*** ***********************************	
HygR-010	DAIAAADLSQTSGFGPFGPQGIGQYTTWRDFICAIADPHVYHWQTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVIDWSEAMFGDSQYEVANIFFWRPWL	
HygR-012	DAIAAADLSOTSGFGFFGPGGIGGYTTWRDFICAIADPHYYHWOTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVILWSEAMFGDSGYEVANIFFWRPWL	
HygR-008	DAIAAADLSOTSGFGFFGPGGIGGYTTWRDFICAIADPHYYHWOTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVILWSEAMFGDSGYEVANIFFWRPWL	
HygR-005	DAIAAADLSOTSGFGFGFGGGIGGYTTWRDFICAIADPHYYHWOTVHASVAQALDELMLWAEDCPEVHLVHADFGSNNVLTDNGRITAVIDWSLAH GUNGVEVANIFFWRWL	
HygR-002	DAIAAADLSOTSGFGFFGPGGIGGYTTWRDFICAIADPHYYHWOTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVILWSEAMFGDSGYEVANIFFWRPWL	
HygR-011	DATAAADLSOTSGFGPFGPGGTGQYTTWRDFICAIADPHYYHWOTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVIDWSEAMFGDSQYEVANIFWRPUL	
HygR-004	DAIAAADLSOTSGFGFFGPGGIGGYTTWRDFICAIADPHYYHWOTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVILWSEAMFGDSGYEVANIFFWRPWL	
HygR-007	DAIAAADLSOTSGFGFFGPGGIGGYTTWRDFICAIADPHYYHWOTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVILWSEAMFGDSGYEVANIFFWRPWL	
HygR-001	DAIAAADLSQTSGFGFFGPGGIGQTTTWEDFICAIADFHVYHWQTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVIDWSEAMFGDSQYEVANIFFWRFWL	
hphMX6	DAIAAADLSOTSGFGFFGPGGIGGYTTWRDFICAIADPHYYHWOTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVILWSEAMFGDSGYEVANIFFWRPWL	
HygR-003	DAIAAADLSOTSGFGFGFGGGIGGYTTWRDFICAIADPHYYHWOTVHASVAQALDELMLWAEDCPEVHLVHADFGSNNVLTDNGRITAVIDWSLAH GUNGVEVANIFFWRWL	
HygR-006	DAIAAADLSOTSGFGFFGPGGIGGYTTWRDFICAIADPHYYHWOTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVILWSEAMFGDSGYEVANIFFWRPWL	
HygR-009	DAIAAADLSOTSGFGFFGPGGGGTOTTWRDFICAIADFHVYHWOTVMDDTVSASVAQALDELMLWAEDCFEVRHLVHADFGSNNVLIDNGRITAVIDWSEAMFGDSQYEVANIFFWRWL	
HygR-013corr	DAIAAADLSOTSGFGFFGGGGGTGTTWKDFICAIADFNYHWGTWMDDTVSASVAGALDELMLWAEDCFEVRHLVHADFGSNNVLIDGRITAYIDWSEAWFGDSGTEVARIFFWREWL DAIAAADLSOTSGFGFFGGGGGGTGTWKRDFICAIADFNYHWGTWMDDTVSASVAGALDELMLWAEDCFEVRHLVHADFGSNNVLITDGRITAYIDWSEAWFGDSGYEVARIFFWREWL	
Hygk=013COII	DAIAMADDAJOOTET GTQGIGGIIIRADT LCAIADFOYIAWGIVEDDI VOASVAQADDEDHEMADDE EVARDI GONNYLLDNGALIAYIDWSDAHRGDGELVARIT FARFW ************************************	
HygR-010	${\tt ACMEQOTRYFERRHPELAGSPRLRAYMLRIGLDQLYQSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKEVGRV\\$	
HygR-012	ACMEQOTRYFERRHPELAGSPRLRAYMLRIGLDQLYQSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKK	
HygR-008	ACMEQOTRYFERRHPELAGSPRLRAYMLRIGLDOLYOSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKE	
HygR-005	ACMEQOTRYFERRHPELAGSPRLRAYMLRIGLDQLYQSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKE	
HygR-002	ACMEQQTRYFERRHPELAGSPRLRAYMLRIGLDQLYQSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRP <mark>DR®M</mark> GEAN	
HygR-011	ACMEQOTRYFERRHPELAGSPRLRAYMLRIGLDQLYQSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKE	
HygR-004	ACMEQOTRYFERRHPELAGSPRLRAYMLRIGLDQLYQSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKE	
HygR-007	ACMEQOTRYFERRHPELAGSPRLRAYMLRIGLDOLYOSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKE	
HygR-001	ACMEQOTRYFERRHPELAGSPRLRAYMLRIGLDQLYQSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKE	
hphMX6	ACMEQOTRYFERRHPELAGSPRLRAYMLRIGLDOLYOSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKE	
HygR-003	ACMEQOTRYFERRHPELAGSPRLRAYMLRIGLDQLYQSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKE	
HygR-006	ACMEQOTRYFERRHPELAGSPRLRAYMLRIGLDOLYOSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKE	
HygR-009	ACMEQOTRYFERRHPELAGSPRLRAYMLRIGLDOLYOSLVDGNFPDDAAWAQGRCDAIVRSGAGTVGRTOIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKE	
11-10D 012-000	ACHEQQINITERANIFEDAGO NUMATHIAN TOLOGUI QOLUQAN DOMANAGANCOMI VADGAO I VANI QI TANKOMA WI DOC VEVI DADGANAK I SITA KAKE——	

ACMEQQTRYFERRHPELAGSPRLRAYMLRIGLDQLYQSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKEYR--

#### Name of Variant Occur KanR-009 51 KanR-021 16 KanR-005 KanR-010 KanR-015 KanR-008 5 KanR-019 KanR-014 KanR-001 KanR-002 KanR-003 KanR-004 KanR-006 KanR-007 KanR-011 2 KanR-012 KanR-013 KanR-017

KanR-018 1

KanR-009

KanR-021

KanR-005

KanR-010 KanR-015

KanR-008

KanR-019

KanR-014

KanR-001

KanR-002

KanR-003

KanR-004 KanR-006

KanR-007

KanR-012

KanR-013

KanR-017

KanR-018

KanR-009

KanR-021 KanR-005

KanR-010

KanR-015 KanR-008

KanR-019

KanR-014

KanR-001

KanR-002

KanR-004 KanR-006

KanR-007

KanR-011 KanR-012

KanR-013

KanR-017

KanR-018

### Alignment

 $ATGAGCCATATTCAACGGGAAACGTCTTGCTC{\red{lightighted}{lightighted}{lightightighted}{lightighted}{lightightighted}{lightighte$ ATGAGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT  $A \texttt{TGAGCCATATTCAACGGGAAACGTCTTGCTC} \underline{\textbf{T}} \textbf{AGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT$ ATGAGCCATATTCAACGGGAAACGTCTTGCTCCCCCCTCTAAACTCCAACATGGACGTCTTATATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT  $\textbf{ATGAGCCATATTCAACGGGAAACGTCTTGCTC} \underline{\textbf{A}} \underline{\textbf{AGGCCGCGATTAAATTCCAACATGGA}} \underline{\textbf{TGCTGATTTATATGGGTATAAATTGGCTCGGGCATAATGTCGGGCAATCAGGTGCGACAAT}}$  $ext{ATGAGCCATATTCAACGGGAAACGTCTTGCTC}^{ ext{AGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT$ ATGAGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATATGAGCCATATTCAACGGGAAACGTCTTGCTC<mark>T</mark>AGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT  ${\tt ATGAGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT$  $ATGAGCCATATTCAACGGGAAACGTCTTGCTC{\ref{equation}}{}^{\ref{equation}$ ATGAGCCATATTCAACGGGAAACGT CAGCCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT ATGAGCCATATTCAACGGGAAACGTCTTGCTCAACGGCACAATCAAGTTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT ATGAGCCATATTCAACGGGAAACGT\_\_\_\_\_CGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT ATGAGCCATATTCAACGGGAAACGT-----CGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT  $ATGAGCCATATTCAACGGGAAACGTCTTGCTCGA {\tt GCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT$ ATGAGCCATATTCAACGGGAAACGT-----CGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT ATGAGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATATGAGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT

 $\tt CTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCC$  $\tt CTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGATGATGATGATGATGATGATGATCAGACTAAACTGGCTGACGGAATTTATGCC$  $\tt CTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCC$  $\tt CTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGATGATGATGATGATGATGACTAAACTGGCTGACGGAATTTATGCC$  $\mathtt{CTATCG}^{\mathtt{D}}$   $\mathtt{TTGTATGGGAAGCCCGATGCCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCCTTGCCAATGATGATGATGAGATGAGATGGTC<math>\mathtt{G}^{\mathtt{D}}$   $\mathtt{CT}^{\mathtt{D}}$   $\mathtt{ACTGGCTGACGGA}^{\mathtt{C}}$   $\mathtt{TTTATGCC}$  $\mathtt{CTAT}$  CGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGCCAACTAAACTGGCTGACGGAATTTATGCC  $\tt CTAT \color{red}{\textbf{G}} CAT \color{blue}{\textbf{G}} CAT \color{blue}{\textbf{G}} CAGAGT T \color{blue}{\textbf{G}} CTT \color{blue}{\textbf{G}} CAA \color{blue}{\textbf{G}} CAA \color{blue}{\textbf{G}} CAT \color{blue}{\textbf{G}} CAT \color{blue}{\textbf{G}} CTT \color{blue}{\textbf{G}} CAGA \color{blue}{\textbf{G}} CTT \color{blue}{\textbf{G}} C$  $\mathtt{CTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGCCTAAACTGGCTGACGGAATTTATGCC$  $\tt CTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCC$  $\tt CTATCG{\color{red}{\textbf{C}}}{\textbf{TTGTATGGGAAGCCCGATGGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCC$  $\tt CTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGGTCAGACTAAACTGGCTGACGGAATTTATGCC$  $\mathtt{CTATCG}^{\mathbf{D}}$   $\mathtt{TTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCC$  $\mathtt{CTATCG}^{\mathtt{CTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCC$  $\tt CTATCG\bar{A}TTGTATGGGAAGCCCGATGCGCCAGGTTTTTTTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGATGATCAGACTAAACTGGCTGACGGAATTTATGCC$  $\mathtt{CTATCG}^{\mathtt{C}}$   $\mathtt{TTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGATGATGATGAGATGGACTAAACTGGCTGACGGAATTTATGCC$  $\tt CTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGATGATGATGATGATGATGATGATCAGACTAAACTGGCTGACGGAATTTATGCC$ 

 ${\tt TCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATAT$  $ext{TCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGCAAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAAATAT$ TCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATAT  ${\tt TCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATAT$  $ext{TCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGCAAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAAATAT$  $ext{TCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCAGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATAT$  $ext{TCTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCC}_{ ext{AGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAAATAT}$  $ext{tccgaccatcaagcatttatccgtactcctgatgatgcatggttactcaccactgcgatcccc} aaaaacagcattccaggtattagaagaatatcctgattcaggtgaaaatat$  $ext{TCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATAT$ TCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGG<mark>a</mark>AAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAA<mark>©</mark>AT  ${\tt TCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATAT$  $ext{TCTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATAT$ 

Size	# Var	Sources*
(bp)	bp	
816	1	No,Qi,19906724
810	19/6d	9717241,7747518
816	1	No,Cl,In
816	2	No
816	30	24050148
816	4	10890530
816	3	No
816	6	Lu
816	3	10890530
816	7	Lu,16496398
816	2	19906724
810	2/6d	In,GE
816	2	19906724
810	3/6d	Or
810	3/6d	MCSG
816	2	3327753
810	3/6d	Or
816	3	24057978
816	5	20040575

KanR-009

KanR-021 KanR-005

KanR-010

KanR-015

KanR-008 KanR-019

KanR-014

KanR-001

KanR-002 KanR-003

KanR-004

KanR-006 KanR-007

KanR-011

KanR-012

KanR-013

KanR-017

KanR-018

KanR-009

KanR-021 KanR-005

KanR-010

KanR-015 KanR-008

KanR-019

KanR-014

KanR-001 KanR-002

KanR-003

KanR-004

KanR-007

KanR-011

KanR-012

KanR-013 KanR-017

KanR-018

KanR-021 KanR-005 KanR-010  ${\tt TGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGA} {\tt CGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTT}$ KanR-015  ${\tt TGTTGATGCGCTGGCGGTGTTCCTGCGCCGGTTCGATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGTGTATTTCGTCTCTGGCTCAGGCGCAATCACGGATGAATAACGGTTT$ KanR-008 KanR-019 KanR-014  $\tt TGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAAC {\color{red}\overline{\textbf{G}}} {\color{blue}\textbf{G}} {\color{blue}\textbf{G}$ KanR-001 KanR-002 KanR-003 KanR-004 KanR-006  ${\tt TGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCAGGCGCAATCACGAATAACGGTTT$  $\tt TGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCG\_{\tt GTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTT$ KanR-007 KanR-011 KanR-012 KanR-013  ${ t total}$ KanR-017 KanR-018 

> OPTO PERCONDATE STREET AND A CONTROL OF THE ACCUSAGE AND A CONTROL OF A CONTROL O

> ACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTC ${\tt ACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTC$  ${\tt ACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTC$ ACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCA CTTGATA A CCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGCCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTC

KanR-009	${\tt TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTCTAAACTGCAGATAAAACTGCAGATAAAACTGCAGATAAAACTGCAGATAAAACTGCAGATAAAACTGCAGATAAAACTGCAGATAAAACTGCAGATAAAACTGCAGATAAAACTGCAGATAAAACTGCAGATAAAACTGCAGATAAAACTGCAGAAAAAAAA$
KanR-021	TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA
KanR-005	TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA
KanR-010	${\tt TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA$
KanR-015	${\tt TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA$
KanR-008	${\tt TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA$
KanR-019	${\tt TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA$
KanR-014	$ ext{TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCA} {}^{ ext{TTGATGCTCGATGAGTTTTTCTAA}}$
KanR-001	$ ext{TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA$
KanR-002	$ ext{TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCA ext{TTGATGCTCGATGAGTTTTTCT} ext{CA}$
KanR-003	$ ext{TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA$
KanR-004	${\tt TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA$
KanR-006	${\tt TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA$
KanR-007	TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA
KanR-011	${\tt TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA$
KanR-012	${\tt TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA$
KanR-013	${\tt TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA$
KanR-017	TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA
KanR-018	$ ext{TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCA}{ ext{TTGATGCTCGATGAGTTTTTCTAA}$
	***************************************

KanR is aph(3')-Ia (aminoglycoside 3'-phosphotransferase) from *E. coli Tn5* transposon.

KanR-016 (5 occurrences) is aptII identical to NeoR/KanR but has an additional N-term nos (nopaline synthase) fusion.

KanR-020 (50 occurrences) is aphA-3 from Staphylococcus aureus. There is only one variant of this.

Name of	Alignment of Translated Protein
Variant	
KanR-009	${ t MSHIQRETSC}$ SRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI
KanR-021	MG <mark></mark> KEKT <mark>HV</mark> SRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI
KanR-005	MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI
KanR-010	${ t MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI$
KanR-015	${ t MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI$
KanR-008	MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVAN <mark>U</mark> VTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI
KanR-019	${ t MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI$
KanR-014	${ t MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI$
KanR-001	${ t MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIY} { t CLUBER CONTROLEMENT CONTROLEME$
KanR-002	${ t MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI$
KanR-003	MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIP <mark>R</mark> KTAFQVLEEYPDSGENI
KanR-004	MSHIQRET <mark></mark> SRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI
KanR-006	${ t MSHIQRET} { t SCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI$
KanR-007	MSHIQRET <mark></mark> SRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI
KanR-011	MSHIQRET <mark></mark> SRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI
KanR-012	MSHIQRETSCS PRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI
KanR-013	MSHIQRET <mark></mark> SRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI
KanR-017	MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIP <mark>R</mark> KTAFQVLEEYPDSGENI
KanR-018	MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYGKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGKTAFQVLEEYPDSGENI
	*. :.:* *:*****************************
T 0.00	UNIVERSE AND AN ANALYSIS OF A CONTROL OF A C
KanR-009	${\tt VDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIF} {\tt PSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF}$
KanR-021	VDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF
KanR-005	${\tt VDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF}$
KanR-010	VDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF
KanR-015 KanR-008	${\tt VDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF}$
Kank-008 Kank-019	VDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF VDALAVFLRRLHSIPVCNCPFNSDRVFRLAOAOSRMNNGLVDASDFDDERNGWPVEOVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLOKRLFOKYGIDNPDMNKLOFHLMLDEFF
Kank-019 Kank-014	VDALAYFLERLISIPYCNCFFRSDRYFRLAQAQSKMNRGLVBASDFDDERNGWPVEQWREMHALLFFSPDSVVTHGDFSLDNLIFFSLQRRLFQRYGIDRFDMRKLQFHLMLDEFF VDALAYFLERLISIPYCNCFFRSDRYFRLAQAQSKMNRGLVBASDFDDERNGWPVEQWREMHKLLPFSPDSVVTHGDFSLDNLIFFSLQRRLFQRYGIDRPDMRKLQFHLMLDDFF VDALAYFLERLISIPYCNCFFRSDRYFRLAQAQSKMNRGLVBASDFDDERNGWPVEQWREMHKLLPFSPDSVVTHGDFSLDNLIFFSLQRRLFQRYGIDRPDMRKLQFHLMLDDFF
Kank-014 Kank-001	VDALAVFLEKLISIPVCNCFFNSDRVFRLAQAQSKMNNGLV@ASDFDDERNGWVEQVMEMHALLFFSPDSVVTHGDFSLDNLIFFSLQRRLFQRYGIDNFDMRKLQFHLMLDEFF VDALAVFLERRISIPVCNCFFNSDRVFRLAQAQSKMNNGLVDASDFDDERNGWPVEQVMEMHKLLPFSPDSVVTHGDFSLDNLIFFSLQRRLFQRYGIDNPDMRKLQFHLMLDEFF VDALAVFLERRISIPVCNCFFNSDRVFRLAQAQSKMNNGLVDASDFDDERNGWPVEQVMEMHKLLPFSPDSVVTHGDFSLDNLIFFSLQRRLFQRYGIDNPDMRKLQFHLMLDEFF
Kank-001 Kank-002	VDALAVFLEKLISIPVCNCFFRSDRVFKLAQAQSKMNNGLVBASDFDDERNGWVEQWKEMHKLLFFSPDSVVTHGDFSLDNLIFFSLQKRLFQKYGIDNFDMRKLQFHLMLDEFF VDALAVFLEKRILSIPVCNCFFRSDRVFKLAQAQSKMNNGLVGSSDFDDERNGWVEQWKEMHKLLFFSPDSVVTHGDFSLDNLIFFSLQKRLFQKYGIDNPDMRKLQFHLMLDEFF VDALAVFLEKRLHSIPVCNCFFRSDRVFKLAQAQSKMNNGLVGSSDFDDERNGWPVEQWKEMHKLLFFSPDSVVTHGDFSLDNLIFFSLQKRLFQKYGIDNPDMRKLQFHLMLDEFF
Kank-002 Kank-003	VDALAYFLERLISIPVCNCPFNSDRVFRLAOAGMENNGLV@ADJFDDERNGWPVEGVWAEMHALLFFSPDSVVTHGDFSLDNLIFFSLQRALFQRIGIDRFDMRALDJFHLMLDEFF VDALAYFLERLISIPVCNCPFNSDRVFRLAOAGMENGLUDASDFDDERNGWPVEGVWAEMHALLFFSPDSVVTHGDFSLDNLIFFSLGRALFQRIGIDRHANLDGFF VDALAYFLERLISIPVCNCPFNSDRVFRLAOAGMENGLUDASDFDDERNGWPVEGVWAEMHALLFFSPDSVVTHGDFSLDNLIFFSLGRALFQRIGIDRHANLDGFF
KanR-004	VDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF
KanR-006 KanR-007	VDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF VDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF
KanR-011 KanR-012	VDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF VDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF
Kank-012 Kank-013	VDALAYFLERLHSIPYCNCFTRSDRYFRLAQAQSKMINGLVDASDFDDERNGWPYEQWKEMHKLLFFSPDSVVTHGDFSLDNLIFFSLQRKLFQRYGIDNFDMRKLQFHLMLDEFF VDALAYFLERLHSIPYCNCFTRSGRYFRLAQAQSKMINGLVDASDFDDERNGWPYEQWKHKLLPFSPDSVVTHGDFSLDNLIFFSLQRKLFQRYGIDNFDMRKLQFHLMLDDFF VDALAYFLERLHSIPYCNCFTRSGRYFRLAQAQSKMINGLVDASDFDDERNGWPYEQWKHKLLPFSPDSVVTHGDFSLDNLIFFSLQRKLFQRYGIDNFDMRKLQFHLMLDDFF
KanR-017 KanR-018	VDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF VDALAVFLRRLHSIPVCNCPFN®DRVFRLAOAOSRMNNGLV®ASDFDDERNGWPVEOVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLOKRLFOKYGIDNPDMNKLOFHLMLDEFF
Nauk-U10	VDALAYFLKKLH51PVCNCFFn_DRYFKLAQAQSKMRNGLV_GADDFDDERGGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFPSLQKRLFQxFGLDNPDMRKLQFHLMLDEFF

# Var

6/2d 

2d

2d 2d

1/2d

Size (aa) 



NeoR/KanR-013 NeoR/KanR-002 NeoR/KanR-016 NeoR/KanR=012 NeoR/KanR-017 NeoR/KanR-008 NeoR/KanR-014 NeoR/KanR-022 NeoR/KanR-004 NeoR/KanR-001 NeoR/KanR-010 NeoR/KanR-015 NeoR/KanR-007 NeoR/KanR-009 NeoR/KanR-018 NeoR/KanR-020 NeoR/KanR-023 NeoR/KanR-003 NeoR/KanR-005 NeoR/KanR-006 NeoR/KanR-011 NeoR/KanR-019 NeoR/KanR-021 

NeoR/KanR-013 NeoR/KanR-002 NeoR/KanR-016 NeoR/KanR-012 NeoR/KanR-017 NeoR/KanR-008 NeoR/KanR-014 NeoR/KanR-022 NeoR/KanR-004 NeoR/KanR-001 NeoR/KanR-010 NeoR/KanR-015 NeoR/KanR-007 NeoR/KanR-009 NeoR/KanR-018 NeoR/KanR-020 NeoR/KanR-023 NeoR/KanR-003 NeoR/KanR-005 NeoR/KanR-006 NeoR/KanR=011 NeoR/KanR-019 NeoR/KanR-021

GCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCCCGAGAAAGTATCCATCATGGCTGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGCTGTTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCCCGAGAAAGTATCCATCATTGGCT ${\tt GC}^{\bullet}_{\tt GC}{\tt TGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGCTGCTTGTTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTTATCCATCATGGCT$  ${\tt GCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCT}$ GCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCT ${\tt GCAGC} {\tt \underline{C}TGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGGGAAAGTATCCATCATGGCT$  ${\tt GCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCT}$  ${\tt GCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGCTCATGCTCATCATGGCTATGGCTTGCTCATGTCATCATGGCTTGCTCATGTCATCATGGCTTGCTCATGT$ GCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCCCGAGAAAGTATCCATCATGGCT

NeoR/KanR-013 NeoR/KanR-002 NeoR/KanR-016 NeoR/KanR-012 NeoR/KanR-017 NeoR/KanR-008 NeoR/KanR-014 NeoR/KanR-022 NeoR/KanR-004 NeoR/KanR-001 NeoR/KanR-010 NeoR/KanR-015 NeoR/KanR-007 NeoR/KanR-009 NeoR/KanR-018 NeoR/KanR-020 NeoR/KanR-023 NeoR/KanR-003 NeoR/KanR-005 NeoR/KanR=006 NeoR/KanR-011 NeoR/KanR-019

NeoR/KanR-013 NeoR/KanR-002 NeoR/KanR-016 NeoR/KanR-012 NeoR/KanR-017 NeoR/KanR=008 NeoR/KanR-014 NeoR/KanR-022 NeoR/KanR-004 NeoR/KanR-001 NeoR/KanR-010 NeoR/KanR-015 NeoR/KanR-007 NeoR/KanR-009 NeoR/KanR-018 NeoR/KanR-020 NeoR/KanR-023 NeoR/KanR-003 NeoR/KanR-005 NeoR/KanR-006

NeoR/KanR-011 NeoR/KanR-019

NeoR/KanR-021

NeoR/KanR-021

 $\mathtt{CTGGACGAAGAGCATCAGGGGCTCGCCGACCTGAACTGTTCGCCAGGCTCAAGGCGCGTATGCCCGACGGCGATGATCTCGTCGTGACCCCATGGCGATGCCTGCTTGCCGAATATCATG$ CTGGACGAAGAGCATCAGGGGCTCGCCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGTATGCCCGACGGCGA GATCTCGTCGTGACTACATGCCGAATATCATG  ${\tt CTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTCGTTGCCCATGGCGATGCCTGCTTGCCGAATATCATG$  $\tt CTGGACGAGGGCTCGGGCCCGACCGGACTGTTCGCCAGGCTCAAGGCG\overline{C}_{\hbox{\scriptsize a}} \\ \tt TGCCCGACGGGGGGGGGGGGTCTCGTCGTGCCGAGTGCCTGCTGCCGAATATCATG$  $\mathtt{CTGGACGAAGAGCATCAGGGGCTCGCCGCCGAACTGTTCGCCAGGCTCAAGGCGCGC}$  $\mathtt{CTGGACGAAGAGCATCAGGGGCTCGCCGACCGAACTGTTCGCCAGGCTCAAGGCG}^{\mathtt{G}}\mathtt{CCATGCCCGACGGCGAGGATCTCGTCGTGACT}\mathtt{CATGGCGATGCCTGCTTGCCGAATATCATG}$  $\tt CTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGA \ GATCTCGTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATG$  $\mathtt{CTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCG\overline{\mathtt{CGCATGCCCGACGGCGAGGATCTCGTCGTGACC}}_{\mathtt{CATGGCGATGCCTGCTTGCCGAATATCATG}}$ 

\*\*\*\*\*\*\*\*\*\*

NeoR/KanR-013  $\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTTGGCGGCGAATGGGCT$ NeoR/KanR-002  $\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$  $\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGC\overline{\textbf{G}}\\ GACCGCTATCAGGACATAGCGTTGGCTGCTGATATTGCTGAAGAGCTTTGGCGGCGAATGGGCT$ NeoR/KanR-016 NeoR/KanR-012  $\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAACTTTGGCGGCGAATGGGCT$ NeoR/KanR-017  $\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCG{\ref{GTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$ NeoR/KanR-008  ${\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$ NeoR/KanR-014  $\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGGCCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$ NeoR/KanR-022  ${\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$ NeoR/KanR-004  ${\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$ NeoR/KanR-001 NeoR/KanR-010  $\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$ NeoR/KanR-015  $\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGA^{\overline{\mathsf{C}}}GCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$ NeoR/KanR-007 NeoR/KanR-009 NeoR/KanR-018 NeoR/KanR-020  ${\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$ NeoR/KanR-023  $\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$ NeoR/KanR-003  $\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$  $\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTTGGCGGCGAATGGGCT$ NeoR/KanR=005 NeoR/KanR-006 NeoR/KanR-011  ${\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGT\_TGGCGGATCGGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$ NeoR/KanR-019  $\tt GTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT$ NeoR/KanR=021 GOOD TARGET AND THE ADDITIONS AS A DESCRIPTION OF THE ADDITION OF THE ADDITI

NeoR/KanR-013 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-002 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-016 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR=012 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-017  $GACCGCTTCCT_{1}^{T}GTGCTTTACGGTATCGCCGC_{1}^{T}CCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA$ NeoR/KanR-008 GACCGCTTCCTGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-014 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR=022 NeoR/KanR-004 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-001 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-010 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-015 NeoR/KanR-007 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-009 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-018 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR=020 NeoR/KanR-023 NeoR/KanR-003 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-005 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-006 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-011 NeoR/KanR-019 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA NeoR/KanR-021 GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA

 ${\tt NeoR/KanR \ is \ aph(3')-II \ (aminogly coside \ 3'-phosphotrans ferase) \ aka \ npt II \ (neomy cin \ phosphotrans ferase) \ from \ Th 5 \ transposon}$ 

NeoR/KanR-010 is identical to NeoR/KanR-002 but missing the START codon because it is in frame with hRLuc

Key ATGC = change in nucleotide relative to consensus sequence - synonymous
ATGC = change in nucleotide relative to consensus sequence - conservative/semi-conservative amino acid substitution
ATGC = change in nucleotide relative to consensus sequence - non-conservative amino acid substitution

Name of	Alignment of Translated Protein
Variant	NTRODAY IN GARNA NI WIREY DAUGUN AAMAGAAN MIREY AAAAAN MIREY AAAA MIRA AANA
NeoR/KanR-013 NeoR/KanR-002	MIEQOGLHAGSPAAWVERLFGYDWAQQTIGGSDAAVERLSAQGRPVLFVKTDLSGALNELQDEARKLSWLATTGVPCAAVLDVVTEAGROWLLLGEVPGQDLLSSHLAPAEKVSINA
NeoR/Kank-002 NeoR/Kank-016	MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDLLSSHLAPAEKVSIMAMIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDLLSSHLAPAEKVSIMA
NeoR/KanR-012	HIGODGIHAGSPAAWURRLFGYDWAQOTIGCSDAAVRLSJAGGRPULFVKTDLSGALNELODEARLSHUATIGVCHAVUDVITEAGRDWLLLGEVPGODLLSSHLAPAEKVSINAMIEODGIHAGSPAAWURRLFGYDWAQOTIGCSDAAVRLSAGGRPULFVKTDLSGALNELODEARLSKWLATTGVCHAVLDVVTEAGRDWLLLGEVPGODLLSSHLAPAEKVSINA
NeoR/KanR-017	middoghaospanwerkeginadyiiggdaayrissagaryifyatisganeliybaarismiatigyzgaylivyiibagrumlideyygdlissalafabasima mibooghaospanwerkegynnaootigcsdaayrrisaogrpyifyintdisgalnelodbaarismiatigyzgayldyytbagrumlidbyngodlissalafabaysima
NeoR/KanR-017	midgudelhagspanwerlegidhagviigesdhayrilsaggrpufyatdisghireligidharishibatigvicayuldviibagrumelidevegdelsbihafbarsima 
NeoR/KanR-014	
NeoR/KanR-022	MIEQDGLHAGSPAAWVERLFGYDWAQOTIGCSDAAVFRLSAQGFPULFVXTDLSGALNELQDEARRLSWLIIGVFCAAVLDVVIEAGRUWLLLGEVFGQDLSGALHAFAERVSIMAMIEQDGLHAGSPAAWVERLFGYDWAQOTIGCSDAAVFRLSAQGFPULFVXTDLSGALNELQDEARRLSWLATTGVPCAAVLDVVIEAGRDWLLLGEVFGQDLSGHLAFAERVSIMA
NeoR/KanR-004	
NeoR/KanR-001	SQUOLINGSFARWERLGYDWAQQTIGCSDAVFRLSAQGRPUFFVXTDLSGALNELQDEARRLSHLB11GVFCARVLDVVIEAGRDWLLLGEVPGQDLLSSGALFAERVSINA MGSAIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGRPUFFVXTDLSGALNELQDEARRLSHLATGVPCAAVLDVVTEAGRDWLLLGEVPGQDLLSSHLAFAERVSINA
NeoR/KanR-011	
NeoR/KanR-015	MIEODGLHAGSPAAWVERLFGYDWAQOTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELODEAARLSWLATTGVPCAAVLDVVTEAGROWLLIGEVPGODLLSSHLAPAEKVSIMA
NeoR/KanR-013	HIEQDGLHAGSPAAWVERLFGYDWAQOTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELODEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGODLLSSHLAPAEKVSIMA
NeoR/KanR-009	MIEODGLHAGSPAAWVERLFGYDWAQOTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGODLLSSHLAPAEKVSIMA
NeoR/KanR-018	MIEODGLHAGSPAAWVERLFGYDWAOOTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELODEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGODLLSSHLAPAEKVSIMA
NeoR/KanR-020	MIEODGLHAGSPAAWVERLFGYDWAQOTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDLLSSHLAPAEKVSIMA
NeoR/KanR-023	MGSAIEODGLHAGSPAAWVERLFGYDWAQOTIGCSDAAVFRLSAOGRPVLFVKTDLSGALNELODEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGODLLSSHLAPAEKVSIMA
NeoR/KanR-003	MGSAIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDLLSSHLAPAEKVSIMA
NeoR/KanR-005	MIEODGLHAGSPAAWVERLFGYDWAQOTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGODLLSSHLAPAEKVSIMA
NeoR/KanR-006	M EODGLHAGSPAAWVERLFGYDWAQOTIGCSDAAVFRLSAOGRPVLFVKTDLSGALNELODEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGODLLSSHLAPAEKVSIMA
NeoR/KanR-011	MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDLLSSHLAPAEKVSIMA
NeoR/KanR-019	MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDLLSSHLAPAEKVSIMA
NeoR/KanR-021	MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDLLSSHLAPAEKVSIMA
	***************************************
NeoR/KanR-013	${\tt DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKA{\color{red}SMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVAD} {\tt IAAPDSQRIAFYRLLDEFF}$
NeoR/KanR-013	DANKRIHTLDPATCFFDHOAKHRIERARTKMEAGLVUODLDEENGGLAFAELFARLKARMFDUGELLVYINGDALLFNLINVENGRESGFIUGGREGVADIAAFDSQKIAFIKLDDEFF DANKRIHTLDPATCPFDHOAKHRIERARTKMEAGLVUODDLDEENGGLAPAELFARLKARMFDGGELVVYHGDACLFNLINVENGRESGFIUGGREGVADIAAFDSQKIAFIKLDDEFF DANKRIHTLDPATCPFDHOAKHRIERARTKMEAGLVUODDLDEENGGLAPAELFARLKARMFDGGELVVYHGDACLFNLINVENGRESGFIUGGREGVADIAAFDSQKIAFIKLDDEFF
NeoR/KanR-016	DAMRRIBTIDPATCFFUNGARRIEGARINGEAGUNGUDLDEENGGLAFAELFARLKARMPUGGEDLVVINGDACLFNIIVENGRESGFIDGGROUWIAAFDSQRIAFIRLIDEFF DAMRRIBTIDPATCFFUNGARRIEGARINGEAGUNGUDLDEENGGLAFAELFARLKARMPUGGEDLVVINGDACLFNIIVENGRESGFIDGGROUWIAAFDSQRIAFIRLIDEFF
NeoR/KanR-010	DAMRRIHTLDPATCFFDHQAKHRIERARTREAGLVDQDDLDEEHGGLAPAELFARLKASMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSORIAFYRLDDEFF
NeoR/KanR-017	DAMREHTLDPATCFFDHOAKHRIERARTREAGLVDODDLDEEHGLAPAELFARLKASMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSORIAFYRLLDEFF
NeoR/KanR-008	DAMREHTLDPATCFFDHOAKHRIERARTMEAGLVDODDLDEEHGGLAPAELFARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSORIAFYRLDEFF
NeoR/KanR-014	DAMRRLHTLDPATCPFDHOAKHRIERARTRMEAGLVDODDLDEEHOGLAPAELFARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSORIAFYRLLDEFF
NeoR/KanR-022	DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKARMPDG DLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF
NeoR/KanR-004	DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF
NeoR/KanR-001	DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKARMPDG DLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF
NeoR/KanR-010	DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF
NeoR/KanR-015	$ ext{DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKAS}  ext{MPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF}$
NeoR/KanR-007	DAMRRLHTLDPATCPFDHQAKHRIER RTRMEAGLVDQDDLDEEHQGLAPAELFARLKARMPDGEDLVVT GDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF
NeoR/KanR-009	DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKARMPDG DLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF
NeoR/KanR-018	${ t DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKAR MPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF$
NeoR/KanR-020	${ t DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKA} { t SMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF$
NeoR/KanR-023	DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKARMPDG DLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF
NeoR/KanR-003	${ t DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF$
NeoR/KanR-005	DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKA <mark>S</mark> MPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF
NeoR/KanR-006	${\tt DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF$
NeoR/KanR-011	DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKASMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLG_ALIAAPDSQRIAFYRLLDEFF
NeoR/KanR-019	DAMRRIHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHGGLAPAELFARIKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF
NeoR/KanR-021	DAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDEEHQGLAPAELFARLKA <mark>S</mark> MPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADIAAPDSQRIAFYRLLDEFF
Vov. Amce - ~	hange in nucleatide valative to congarang gaguange, gunanumang
	hange in nucleotide relative to consensus sequence — synonymous hange in nucleotide relative to consensus sequence — conservative/semi-conservative amino acid substitution
	nange in nucleotide relative to consensus sequence — conservative/semi-conservative amino acid substitution hange in nucleotide relative to consensus sequence — non-conservative amino acid substitution
Ardc = C	nange in nacreotive rerative to consensus sequence - non-conservative annino actu substitution

# Var

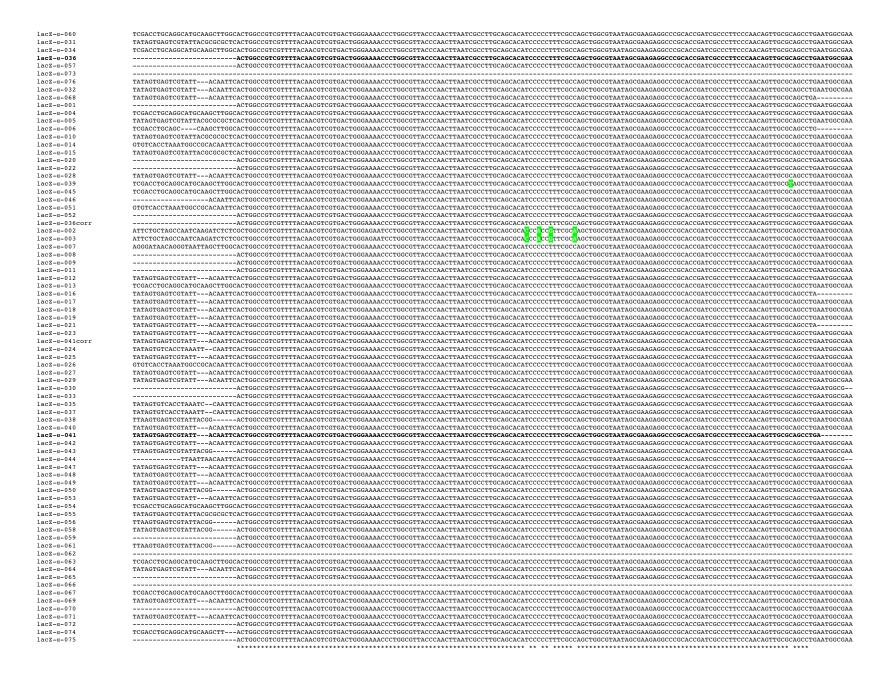
1				
Name of # of	Alignment	Size	# Var	Sources*
Variant Occur		(bp)	bp	
PuroR-002 23		600	1	Cl,AG
	ATGACCGAGTACAAGCCCACGGTGCGCCACCGCGACGACGTCCCC CGCGCGCGCCCCCCCCCC		_	•
PuroR-004 15	ATGACCGAGTACAAGCCCACGGTGCGCCACCCGCCACCCGCGACGACGTCCCC <mark>A</mark> GGGCCGTACGCACCCTCGCCGCCGTTCGCCGACTACCCCCCCC	600	2	Th
PuroR-009 10	ATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCCCCCCGCCGCGTTCCCCGCCACCGCCACCGCCACACCGTCGA	600	3	Cl
PuroR-006 7	ATGACCGAGTACAAGCCCACGGTGCGCCACCCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCCCCCCCC	600	1	Cl
PuroR-007 6	atgaccgagtacaagcc <mark>hace</mark> gtgcgcct <mark>e</mark> gccac <mark>h</mark> cgcga <mark>ngangte</mark> cccggcggcgt <mark>o</mark> cgccac <mark>h</mark> ctegccgccg <mark>t</mark> ttcgccgactaccccgc <mark>nac</mark> cc <mark>e</mark> cacaccgt <mark>gga</mark> cce	600	106	Pr
	ATGACCGAGTACAAGCCTACCCCCT GCCACTCGCGATGATGTCCCCCGCGCCGTTCCGCCGCCGCTTTCGCCGACTACCCCGCTACCCCGCTACACCCGTTCAAGCC			
PuroR-011 6	ATGACCGAGTACAAGCCCACGGTGCGCCACCCGCGACGACGACGTCCCC <mark>A</mark> GGGCCGTACGCACCCTCGCCGCCGTTCGCCGACTACCCCCCCC	600	3	Or,2194165
PuroR-005 2	ATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCAGTTCCCCCGCCGCCGCGCTTCGCCGCCGCCACCGCCACCACCGCCGACACCGTCGATCCA	600	2	Or
PuroR-008 2	ATG <mark>GCC</mark> ACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCACCCTCGCCGCCGCTTCGCCGACTACCCCCGCCACCCCCCCACACCGTCGA <mark>C</mark> CCG	603	6	OB
PuroR-010 2	$\lambda$ TO CONTROL OF THE CONTROL OF TH	600	2	21706014
	ATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCC <mark>A</mark> GGGCCGTACGCCACCCCTCGCCGCCGCTTCGCCGACTACCCCGCCACCACCCGCCACACCGTCGA <mark>T</mark> CCG ATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGACCGTCCCC <mark>A</mark> GGGCCGTACGCCCCCCCGCCGCCGCTCGCCGCCCCCCCCC			
PuroR-003 1	ATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCC12GGGCCACCCCTCGCCGCCGCTTCGCCGCCGCCACCCCCCCC	600	4	Or
PuroR-001 1	ATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCCCCCCGCCGCGTTCGCCGACTACCCCGCCACCGCCACACCGTCGA	597	4	Cl
PuroR-001corr	ATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCACCCTCGCCGCCGCCGCTTCGCCGACTACCCCGCCACCGCCACACCGTCGA <mark>C</mark> CCG	624	1	
	*** ********* ** ******* ** ***** **** ****		=	
PuroR-002	GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGGACGACGACGCCGCGGTGGCGGTTCGGACCACGCCG			
PuroR-004	GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGACGACGACGACGCCGCGGTGGCGGTCTGGACCACGCCG			
PuroR-009	${\tt GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGACGACGACGGCGCGCGGTGGCGGTTCGACCACGCCGGTCGGGCGCGCGGTGGCGGTGGCGGTGGCGCGCGGGTGGCGCGCGGGTGGCGCGCGGGTGGCGCGCGGGTGGCGGC$			
PuroR-006	GACCGCCACATCGAGCGGGCCCGCGGCGGCGGCGCCCCCCCC			
PuroR-007	GACCG <mark>CCACATCGAGCGTGTCACAGAGT</mark> TGCA <mark>CGAGCTC</mark> TTCCT <mark>GACG</mark> CGCGTCGGGCT <mark>G</mark> GACATCGGCAAGGTGTGGGTAGCGGCGCGCGCGCGCGCGC			
PuroR-011	GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGCTCGGGCTCGACATCGGCAAGGTGTGGGTGG			
PuroR-005	GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGACGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCG			
PuroR-008	GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGGCACATCGGCAAGGTGTGGGTCGCGGACGACGGCGCCGGTGGCGGTCTGGACCACGCCCG			
PuroR-010	GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGCTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGACGACGACGGCGCCGC <mark>C</mark> GTGGCGGTCTGGACCACGCCG			
PuroR-003	GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGACGACGACGACGCCGCGGTGGCGGTCTGGACCACGCCG			
PuroR-001	GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGACGACGACGACGCCGCGGTGGCGGTCTGGACCACGCCG			
PuroR-001corr	${\tt GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGGTCGGCGCGGTGGGCGGTCGGACCACGCCGGGTGGGGTCGGGCGGTCGGGGTCGGGCGGTCGGGGTCGGGCGGTGGGGCGGTGGGGCGGGGGG$			
rulok-001col1	uncedetheriedadeuderiedaderietadane (effected edeted edeted edetheriede edeted edeted edeted edeted edeted edet			
	***** ******** ** ** ** ** ** ** ** **			
PuroR-002	GAGAGCGTCGAAGCGGGGGGGGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAACAGATGGAAGGCCTCCTGGCGCCCCACCGCCCAAG			
PuroR-004	GAGAGCGTCGAAGCGGGGGGGGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCACCAGCAACAGATGGAAGGCCTCCTGGCGCCCCACCGCCCAAG			
PuroR-009	${\tt GAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCCGCATGGCCGATTGAGCGGTTCCCGGCTGGCCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGCACCAACAGATGGAAGGCCTCCTGGCGCCCCAACAACAGATGGAAGGCCTCCTGGCGCCCCAACAGATGGAAGGCCTCCTGGCGCCGCACCAACAGATGGAAGGCCTCCTGGCGCCGCACCAACAGATGGAAGGCCTCCTGGCGCCGCACCAACAGATGGAAGGCCTCCTGGCGCCGCACAACAGATGGAAGGCCTCCTGGCCGCACAACAGATGGAAGGCCTCCTGGCCGCCAACAGATGGAAGAA$			
PuroR-006	GAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCGCGTGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCAGCAACAGATGGAAGGCCTCCTGGCGCCCCACCGGCCCAAG			
PuroR-007	GAGAGCGT <mark>TGAE</mark> GCEGGEGCEGTGTTCGCCGAGATCGGCCC <mark>E</mark> CGAATGGCCGAGETGAGCGGEAACCGGCTGGCCGCCCAAATGGAEGGCCTECTTGCECCEAAACGA			
PuroR-011	GAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCAGCAACAGATGGAAGGCCTCCTGGCGCCCCACCGCCCCAAG			
PuroR-005	${\tt GAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCCGCCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCAACAGATGGAAGGCCTCCTGGCGCCCCACCGGCCCAAGGAGGCCACCGGCCCAAGGAGG$			
PuroR-008	GAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCAGCAACAGATGGAAGGCCTCCTGGCGCCCCAAC			
PuroR-010	GAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCGCGTTGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCAGCAACAGATGGAAGGCCTCCTGGCGCCCCACCGCCCAAG			
PuroR-003	GAGAGCGTCGAAGCGGGGGGGGGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAACAGATGGAAGGCCTCCTGGCGCCCCACCGGCCCAAG			
PuroR-001	GAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCAGCAACAGATGGAAGGCCTCCTGGCGCCCCAAC			
PuroR-001corr	${\tt GAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCCGCCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCAACAGATGGAAGGCCTCCTGGCGCCCCACCGGCCCAAGGAGGCCACCGGCCCAAGGAGG$			
Turon-ourcorr	******* ** ** ** ** *********** ** ** *			
PuroR-002	GAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCCGGCGTGCCCTTC			
PuroR-004	GAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCGCGGGGTGCCCGCCTTC			
PuroR-009	GAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCCGAGCGCCCGCC			
PuroR-006	GAGCCGCGTGGTTCCTGGCCACGTCGGCCTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCCCGGGGTGCCCGCCTTC			
PuroR-007	GAGCC <mark>TGCTTTC</mark> CTGGCCAC <mark>TGTT</mark> GGGCTCAGCCCGACCACCAGGGCAAGGG <mark>CT</mark> TGGGCAGCGCCGTCGTG <mark>T</mark> TCCCGGGGTTGAGAGCGGCCGAACCGCCCGGTGTGCCCCCCTTT			
PuroR-011	GAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCGAGCGCGCGGGGTGCCCGCCTTC			
PuroR-005	GAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCGCGC			
PuroR-008	${\tt GAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCGCGTGCGCCCTTC}$			
PuroR-010	GASCCCGCTGGTTCCTGGCCACCCTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCACGCCGTCGTCCTCCCCGGAGTGGGGCCGAGCGCCGCGGGTCCCCGCTTC			
PuroR-003	GAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTGTCCCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCCGCGCGGGTGCCCGCCTTC			
PuroR-001	GAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCGCGGGGTGCCCGCCTTC			
PuroR-001corr	GAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCGGAGCGCGCGGGGTGCCCGCCTTC			
	**** ** **** ***** ** ** ** ** ** ******			
PuroR-002	$\tt CTGGAGACCTCCGCCACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGTCACCGTCACCGCCGACGTCCGAGGACCGCGC-ACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACCGCAGGACCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCACACCAC$			
				_
PuroR-004	CTGGAGACETCCGCGCCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGTCACCGCACGTCGAGGTGCCCGAAGGACCGCGC-ACCTGGTGCATGACCCGCAAGCCCGGTGCCTGA			-
PuroR-009	CTGGAGACCTCCGCGCCCCCCCCCCCCTTCTACGAGCGCCTCGCCTTCACCGCCGCGCCGCGCCGCGAGGTCCCCGAAGGACCGCGCAACCTGCTGCATGACCCGCAAGCCCGCTACCTAC			-
PuroR-006	4 $6$ $6$ $6$ $6$ $6$ $6$ $6$ $6$ $6$ $6$			_
PuroR-007	ct <mark>gar</mark> ac <mark>hac</mark> cscromgraaccticcrttctacgascs <u>s</u> ct <u>s</u> gcttcaccst <u>r</u> accsccsa <u>r</u> gtcgaggtscccga <mark>r</mark> gsccacc <mark>tros</mark> -acctsgtsgratgacrosparaccss			_
	amenda and and and and and and and and and			
PuroR-011	CTGGAGACCTCCGCCCCCCCAACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGCCGCCGACGTCGAG <mark>T</mark> TGCCCGAAGGACCGCGC <mark>G</mark> ACCTGGTGCATGACCCGCAAGCCCGGTGCCTGA			
PuroR-005	CTGGAGACCTCCGCGCCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGCCGCGCACCTGTGCATGACCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGCAAGCCCGGTGCCTGA			
PuroR-008	CTGGAGACCTCCGCGCCCCCCAACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTGCCCGAAGGACCGCGC-ACCTGGTGCATGACCCGCAAGCCCGGTGCCTGA	·		-
PuroR-010	$\tt CTGGAGACCTCCGCGCCCCCCAACCTCCCCTTCTACGAGCGGCTCGCCTTCACCGTCACCGCCGACGTCGAGGGTGCCCGAAGGACCGCGC-ACCTGGTGCATGACCCGCAAGCCCGGTGCCTGAGGAGACCTCCGAGGAGAGACCTCCGCAAGCCCGCAAGCCCGCAAGCCCGGTGCCTGAGAGACCGCGAAGCCCGCAAGCCAAGCCCGCAAGCCCGCAAGCCAAGCCCGCAAGCCCGCAAGCCAAGCCAAGCCAAGCCAAGAAAGCAAGAAAAAA$			
PuroR-003	$\tt CTGGAGACCTCCGCGCCACCTCCTACGAGCGGCTCGGCT$			
PuroR-001	CTGGAGACCTCCGCCCCCCAACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGAGGTGCCCGAAGGACCGCGC-ACCTGGTGCATGACCCGCAAGCCCGGTGCC			
PuroR-001corr	CTGGAGACCTCCGCGCCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGCCGCGC-ACCTGGTGCATGACCCGCAAGCCCGGTGCCACC	GGTGCGGCTTCGF	AGGATAAAATA	A
	** ** ** ** * * ***** ** ****** ** *****			

Name of Variant	Alignment of Translated Protein	Size (aa)	# Var aa
PuroR-002	${\tt M-TEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPESVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPK$	Ì99	0
PuroR-004	M-TEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPESVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPK	199	0
PuroR-009	$ exttt{M-TEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPESVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPK$	199	0
PuroR-006	$ exttt{M-TEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPESVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPK$	199	0
PuroR-007	$ exttt{M-TEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPESVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPK$	199	0
PuroR-011	$ exttt{M-TEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPESVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPK$	199	5
PuroR-005	$ exttt{M-TEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPESVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPK$	199	0
PuroR-008	$\mathbf{M}_{\mathbf{A}}^{\mathbf{A}}$ TEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPESVEAGAVFAEIG $^{\mathbf{C}}$ RMAELSGSRLAAQQQMEGLLAPHRPK	200	2
PuroR-010	M-TEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPESVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPK	199	0
PuroR-003	$ exttt{M-TEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPESVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPK$	199	5
PuroR-001corr	$ exttt{M-TEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPESVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPK$	207	0
	* *************************************		
PuroR-002	EPAWFLATVGVSPDHQGKGLGSAVVLPGVEAAERAGVPAFLETSAPRNLPFYERLGFTVTADVEVPEGPRTWCMTRKPGA		
PuroR-004	EPAWFLATVGVSPDHQGKGLGSAVVLPGVEAAERAGVPAFLETSAPRNLPFYERLGFTVTADVEVPEGPRTWCMTRKPGA		
PuroR-009	EPAWFLATVGVSPDHQGKGLGSAVVLPGVEAAERAGVPAFLETSAPRNLPFYERLGFTVTADVEVPEGPRTWCMTRKPGA		
PuroR-006	EPAWFLATVGVSPDHQGKGLGSAVVLPGVEAAERAGVPAFLETSAPRNLPFYERLGFTVTADVEVPEGPRTWCMTRKPGA		
PuroR-007	EPAWFLATVGVSPDHQGKGLGSAVVLPGVEAAERAGVPAFLETSAPRNLPFYERLGFTVTADVE <mark>VPEGPR</mark> TWCMTRKPGA		
PuroR-011	EPAWFLATVGVSPDHQGKGLGSAVVLPGVEAAERAGVPAFLETSAPRNLPFYERLGFTVTADVE <mark>C</mark> PKD <mark>RA</mark> TWCMTRKPGA		
PuroR-005	EPAWFLATVGVSPDHQGKGLGSAVVLPGVEAAERAGVPAFLETSAPRNLPFYERLGFTVTADVEVPEGPRTWCMTRKPGA		
PuroR-008	EPAWFLATVGVSPDHQGKGLGSAVVLPGVEAAERAGVPAFLETSAPRNLPFYERLGFTVTADVEVPEGPRTWCMTRKPGA		
PuroR-010	${\tt EPAWFLATVGVSPDHQGKGLGSAVVLPGVEAAERAGVPAFLETSAPRNLPFYERLGFTVTADVE} {\tt VPEGPR} {\tt TWCMTRKPGA}$		
PuroR-003	epawflatvgvspdhqgkglgsavvlpgveaaeragvpafletsaprnlpfyerlgftvtadve <mark>c</mark> pkd <mark>ra</mark> twcmtrkpga		
PuroR-001corr	${ t EPAWFLATVGVSPDHQGKGLGSAVVLPGVEAAERAGVPAFLETSAPRNLPFYERLGFTVTADVE\overline{VPEGPRTWCMTRKPGATGAASRIK}$		
	***************************************		

Name of	# of	Alignment Common Restriction Enzyme Sites in MCS	Size	MCS	Sources*
Variant	0ccur	HindIII KpnI XhoI SphI PstI EcoRV Sall XbaI	(bp)		
$1acz-\alpha-060$	24	ATGACCATGATTAC	234	207,328,061,183,195	Ca,7919298
lacZ-α-031		ATGACCATGATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGA	579	115,MCS 2-008	2659436
lacZ-α-034 lacZ-α-036	4	ATGACCATGATTAC	231 <b>285</b>	207,328,061,183,195	Ta,7919298
lacZ-α-057	4	ATGACCATGATTAGCCAAGCTT	324	207,328,061,183,195	3073106
lacZ-α-073	4	ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTACTCGAGCTCAAACATATGAAA	126	327	CB
lacZ-α-076	4		255		CB
$1acz-\alpha-032$	3	ATGACCATGATTACGCCAAGCTCGGAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGA	390	115,MCS 2-008	10890530
lacZ-α-068	3	ATGACCATGATTACGCCAAGCTATTTAGGTGAGACTATAGAATACTCAAGCT-TGCATGCGATACGTATCGTTAACGAT	273		20040575
lacZ-α-001	2	ATGACCATGATTACGCCAAGCTTGTCGACTCTAGA ATGACCATGATTACGTCGACTCTAGA	309 327	207,328,061,183,195	33277753,8566803
lacZ-α-004 lacZ-α-005	2	ATGACCATGATTAC- ATGACTAGATGAGGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAG	527 576	207,183 051,MCS 2-008	Ta,3323803 Aq
lacZ-α-006	2	ATGACCATGATTAC	198	049,117,MCS1-008	Ca
lacZ-α-010	2	ATGACCATGATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGGTACCGGGCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATC	576	223,MCS2,MCS 2-008	Ag,2555794
$lacz-\alpha-014$	2	ATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATC	381	077,3'MCS-2	No,Qi
lacZ-α-015		ATGACCATGATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGA	576	115,MCS 2-008	Ag,2555794
lacZ-α-020 lacZ-α-022	2	ATGACCATGATTACGCCAAGCTTGTCGAATGACCATGATTACGCCAAGCTTGTCGAGATGACCATGATTACGCCAAGCTT	198 348	097	Lu GE
lacz-α-022	2	ATGACCATGATTAGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTT	381	328,061,195	Pr,Ne
lacZ-α-039	2	ATGACCATGATTAC	303	328,061,195	3333305
lacZ-α-045	2	ATGACCATGATTAC	291	328,061,195	3333305
$lacz-\alpha-046$		ATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATC	198	149,261	Ca
lacZ-α-051		ATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATC	366		In
lacZ-α-052 lacZ-α-036c		ATGACCATGATTACGCCAAGCTTGCATGCGTCGACTCTAGA ATGACCATGATTACGCCAAGCTTGCAGCTCTAGAGTCGACTCTAGA	339 339	207,183 207,183	3073106 3073106
lacz-α-0360		ATGAC TATGATTACAACGCCCGGCA	261	207,183	No
lacZ=α=002	1	ATGAT ATGATTACAACGCCCAACTTAATTAACATTAGTGGTGGTGGTGGTGGTGGTCTCCAGGATCCTCTGGGTCTTCAGGAGCGAGTTCTGGCTGG	438	MCS 2-008	No.
	-	ATGACCATGATTACGAATTTAATACGACTCACTATAGGGAATTTAAATTTAATTAA-GGCGCGCCCATGGATGATGTTAACATGATGATATCACGCGTGGCGCCACTAGTGC	504	1100 2 000	Мо
lacZ-α-008	1	ATGACCATGATTACGCCAAGCTTGTCGACTCTAGAGTCGACTCTAGA	351	053	Ne
lacZ-α-009	1	ATGACCATGATTACGCCAAGCTTGTCGACTCTAGA	231	328,061,195	Ta
lacZ-α-011	1	ATGACCATGATTACGCCAAGCTTGTCGACTCTAGA	327	328,061,195	3323803
lacZ-α-012 lacZ-α-013	1	ATGACCATGATTACGCCAAGCTCGGAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGA ATGACCATGATTAC	363 324	115,MCS 2-008 207,183	10890530 2985470
lacz-α-015		ATGACCATGATTAC- ATGACTAGATTACGCCAAGCTATTTAGGTGACGCGTTAGAATACTCAAGCTATGCATC-AAGCTTGGTACCGAGCTC	360	207,183	2985470 In
lacZ-α-017		ATGACCATGATTACGCCAAGCTTGGTACCGAGCTC	387	088	Ag
lacZ-α-018	1	ATGACC TGATTACGCCAAGCTCGAAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCGCGCGCCTGCAGGTGCAGGTCGAC	549	091,3'MCS-2	No
$lacz-\alpha-019$	1	ATGACCATGATTACGCCAAGCTCGAAATTAACCCTCACTAAAGGGAACAAAAGCTGGGTACCGGGCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATC	573	223,MCS2,MCS 2-008	2970625
lacZ-α-021		ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATC-AAGCTTGGTACCGAGCTC	360		In
lacZ-α-023		ATGACCATGATTACGCCAAGCTCGGAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGA ATGACCATGATTACGCCAAGCTCGGAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGA	438	115,MCS 2-008 115,MCS 2-008	10890530
lacZ-α-041c		ATGACCATGATTAGCCAAGCTCTAATAGGACTACATATAGGACAAAAACTGGAGCTCACCGGGTGGGGCCCC	342 369	115,MCS 2-008	10890530 In
lacZ-α-025	1	ATGACCATGATTACGCCAAGCTCGAAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGCCGCTCTAGA	573	115.MCS 2-008	2970625
lacZ-α-026	1	ATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATC	351	077,3'MCS-2	No
$1acz-\alpha-027$		ATGACCATGATTACGCCAAGCTTGGTACCGAGCTC	354		In
lacZ-α-029	1	ATGACCATGATTACGCCAAGCTATTAGGTGACACTATAGAATACTCAAGCTATGCATC-AAGCTTGGTACCGAGCTC	342	107,301,045	In
lacZ-α-030		ATGACCATGATTACGCCAAGCTTGTCGACTCTAGA ATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGA-AAGCTTGCATGCCTGCAGGTCGACTCTAGAGGGAT	324 333	328,061,195 116	6095209 No
lacZ-α-033 lacZ-α-035	1	ATGACCATGATTAGGCCAAGCCT-CTAATACGACT-ACTATAGGGAAAAGCTT	384	116	Pr
lacZ-α-037	1	ATGACCATGATTACGCCAAGCT CTAATACGACTCACTATAGGGA AAGCTGGTAC GCCTGCAG GCCAG GCCTGCAG GCCTGCAG GCC	576	223,MCS2,MCS 2-008	
lacZ-α-038	1	ATGACCATGATTACGCCAAGCTACGATAATACGACTCACTAGTGGGTACCAGAGCTCCCTAGGTTCTAGAACCGGTGACGTCTCCCATGGTG-AAGCTT	282		In
$1acz-\alpha-040$	1	ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGG	384		Pr
lacZ-α-041	_	ATGACCATGATTACGCCAAGCTCGGAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGA ATGACCATGATTACGCCAAGCTCGGAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGA	342	139,MCS 2-008	10890530
lacZ-α-042 lacZ-α-043	1	ATGACCATGATTACGCCAAGGTCGGAATTAACCCTCACTAAAGGGAACCAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGA	576 369	115,MCS 2-008 142.157	7669286
lacZ-α-044	1	ATGACCATGATTAGGCAAGCTGGGGGGCCAAGCTTGCATGCGCATGCGTCGAGGTCGACTCTAGA	351	141,207,183	7655517
lacZ-α-047	1	ATGACCATGATTACGCCAAGCTCGGAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGA	651	115,MCS 2-008	10890530
lacZ-α-048	1	ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGG	369		Pr
lacZ-α-049	1	ATGACCATGATTACGCCAAGCTTGGTACCGAGCTCATGACCATGATGATGATGATGATGATGATGATGATGATGATGATGA	513	328,061,195	
lacZ-α-050	1	ATGACCATGATTACGCCAAGCTACGTAATACGACTCACTATAGGG	372	157	7669286
lacZ-α-053 lacZ-α-054	1	ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATC-AAGCTTGGTACCGAGCTCATGATTACGTACCGAGCTC	369 303	176,5' MCS-004 207,183	7669286 3327753
lacZ-α-055	1	ATGRICATEATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATC	390	184.MCS2.MCS 2-008	IMAGE
lacZ-α-056	1	ATGACCATGATTACGCCAAGCTACGATAATACGACTCACTAGTGGG-GCCCGTGCAATTGAAGCCGGCTGGCGCCAAGCTTCTCTGCAGGATATC	363	209	7669286
lacZ-α-058	1	ATGACCATGATTACGCCAAGCTACGTAATACGACTCACTATAGGG-GCCCGTGCAATTGAAGCCGGCTGGCGCCAAGCTTCTCTGCAGGATATC	363	209	7669286
lacZ-α-059	1	ATGACCATGATTACGCCAAGCTTGTCGACTCTAGA	243	208	3327753
lacZ-α-061	1	ATGACCATGATTACGCCAAGCTACGTAATACGACTCACTAGTGGGTCGACCGCATGCGCATGCGCATGC	366	224	7669286
lacZ-α-062	1	ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGG ATGACCATGATTAC	121 507	229 328,061,195	CB Ne
lacZ-α-064	1	ATGACCATGATTAC	263	328,061,195	CB
lacZ-α-065	1	ATGACCATGATTACGCCAAGCTTGTCGACTCTAGAGCATGCCTGCAGGTCGACTCTAGA	507	207,183	Ne
lacZ-α-066	1	ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATG	115	268	CB
$1acz-\alpha-067$	1	ATG CCATGATTAC	270	328,061,195	Ta
lacZ-α-069	1		254	200 001 1	CB
lacZ-α-070 lacZ-α-071	1	ATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGA-AAGCTTGCATGCCTGCAGGTCGACTCTAGA ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATC-AAGCTTGGTACCGAGCTC	267 606	328,061,195	Th In
lacz-α-071		ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATC-AAGCTTCGTACCGACCTC	477	328,061,195	In Th
lacZ-α-074		ATGACCATGATTACCCC	309	328,061,195	8566803
lacZ-α-075	1	ATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGA-AAGCTTGCATGCAGGCCTCTGCAGTCGACTCGAC	291	097	Th
		*** **			

#### Common Restriction Enzyme Sites in MCS

lacZ-α-060			
lacZ-α-031			GATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTAGATATCAAGCTTATCGATACCGAATTC
lacZ-α-034			GAGCTCGGTACCCGGGGATCCTCT
lacZ-α-036 lacZ-α-057	GGATCC	CCGGGTACCGAGCTCGAATTC	
lacZ-α-073			ACTAGTGGATCCTGAT
lacZ-α-076			GTTTAAACCCATGGGGCCCAATTC
lacZ-α-032			GATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTAGATATCGATTCGAT
lacZ-α-068			
lacZ-α-001 lacZ-α-004			
lacz-α-004 lacz-α-005			GAGCTCGGTACCCGGGGATCCTCTGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCCCGGTAGAGCTCGGTACCCGGGATCCTCT
lacZ-α-006			CCGGGGATC
lacZ-a-010		GAATTC	CTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAATTC
$lacz-\alpha-014$			GTGATATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTC
lacZ-α-015 lacZ-α-020			GATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCCCGGTACCCAATTC
1acz-α-020 1acz-α-022			
lacZ-α-028			
lacZ-α-039		GAATTC	GAGCTCGGTACCCGGGGATCCTCT
lacZ-α-045			GAGCTCGGTACCCGGGGATCCTCT
lacZ-α-046			GTGATATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCC
lacZ-α-051 lacZ-α-052	GGATCC	AGAATTC	GTGATATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTC
lacZ-α-036corr			
lacZ-α-002			GATATCTATATCTCCTTAAAGTTAAACAA
lacZ-a-003	CCGCAAGCTTCGAACGCGTATCGATGGTAC	CCGTCGACGTCCTGCAGGCGCCCTGTACAGAATTC	GGATCCACGAGCTCCCGGGATATC
lacZ-α-007			GGTACCTACGTACTCGAGTCGCGACGTACGTTCGAACAATTGGTTTAAACGCCCGGGCACGTGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTCGGTCC
lacZ-α-008	CTTAATTAAGGATCC	GCCGCGCCCCCGGGTACCGAGCTCGAATTC	
lacZ-α-009 lacZ-α-011	GGATCC	CCGGGTACCGAGCTCGAATTC	
lacZ-α-012			GATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCCCGGTA
lacZ-α-013			GAGCTCGGTACCCGGGGATCCTCT
lacZ-α-016			TTAAGGGCGAATTCTGCAGATATC
lacZ-α-017			TTAAGGGCGAATTCTGCAGATATC
lacZ-α-018			
lacZ-α-019 lacZ-α-021			CTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCTGGAGCTGCAGTTCTGCAGTTCTGCAGATATCTGCAGATATC
lacZ-α-021			
lacZ-a-041corr			GATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTA
lacZ-α-024			
lacZ-α-025			
lacZ-α-026 lacZ-α-027			GTGATATCTGAATTCGTCGACAGCTTCTCGAGCCTAGGCTAGGCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTC TTAAGGGCGAATTCTGCAGATATC
lacZ-α-027			TIANGGCGATATC
lacZ-α-030			
lacZ-α-033			
lacZ-α-035	GGATCC	CCGGGTACCGAGCTCGAATTC	GT
lacZ-α-037			
lacZ-α-038 lacZ-α-040			CTCGAGACCGTACGTGCGCGCGATGCATCCCCCTTAGTCAA
lacZ-α-041			GATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCCCCGGTA
lacZ-α-042			GATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAATTC
lacZ-α-043			CTGCAGGATATCTGGATCCACGAAGCTTCCCATGGTGACGTCACCGGTTCTAGATACCTAGGTGAGCTCTGGTACCCTCTAGTCAA
lacZ-α-044			
	3 am3 am ac		
lacZ-α-047			GATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCCCGGTACCCAATTC
lacZ-α-048		CGGCCGC	ACTAGTGATATCCCGCGGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTC
	GGATCC	ACTAGTAACGGCCGCACTAGTAATTCGCCC	
$lacz-\alpha-048$ $lacz-\alpha-049$ $lacz-\alpha-050$ $lacz-\alpha-053$	GGATCC	ACTAGTAACGGCCGC- ACTAGTAACGGCCGCCAGTGTGGTGGAATTCGCCC SAATGCATCGGGCGCACCGTACGTCTCGAGGAATTC ACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCC	
$\begin{array}{l} 1acZ-\alpha-048 \\ 1acZ-\alpha-049 \\ 1acZ-\alpha-050 \\ 1acZ-\alpha-053 \\ 1acZ-\alpha-054 \end{array}$	GGATCC		
$\begin{array}{l} 1acZ - \alpha - 0.48 \\ 1acZ - \alpha - 0.49 \\ 1acZ - \alpha - 0.50 \\ 1acZ - \alpha - 0.53 \\ 1acZ - \alpha - 0.54 \\ 1acZ - \alpha - 0.55 \\ 1acZ - \alpha - 0.55 \\ \end{array}$		CGGCCGCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCC SAATGCATCGCGGCACCGTACGTCTCGAGGAATTCGCCC	
$\begin{array}{l} 1ac2-\alpha-0.48 \\ 1ac2-\alpha-0.49 \\ 1ac2-\alpha-0.50 \\ 1ac2-\alpha-0.53 \\ 1ac2-\alpha-0.53 \\ 1ac2-\alpha-0.54 \\ 1ac2-\alpha-0.55 \\ 1ac2-\alpha-0.56 \\ \end{array}$	GGATCC	CGGCCGC	
$\begin{aligned} & 1ac2 - \alpha - 048 \\ & 1ac2 - \alpha - 049 \\ & 1ac2 - \alpha - 050 \\ & 1ac2 - \alpha - 053 \\ & 1ac2 - \alpha - 054 \\ & 1ac2 - \alpha - 055 \\ & 1ac2 - \alpha - 056 \\ & 1ac2 - \alpha - 058 \end{aligned}$			
lacZ-u-048 lacZ-u-049 lacZ-u-050 lacZ-u-053 lacZ-u-054 lacZ-u-055 lacZ-u-056 lacZ-u-058 lacZ-u-059		CGGCCGC	
lac2-a-048 lac2-a-049 lac2-a-050 lac2-a-053 lac2-a-054 lac2-a-055 lac2-a-058 lac2-a-058 lac2-a-061 lac2-a-061	GGATCC		
lacZ-a-048 lacZ-a-050 lacZ-a-050 lacZ-a-050 lacZ-a-054 lacZ-a-054 lacZ-a-056 lacZ-a-056 lacZ-a-058 lacZ-a-061 lacZ-a-061 lacZ-a-063		CGCCGC	
lacZ-a-048 lacZ-a-049 lacZ-a-050 lacZ-a-050 lacZ-a-054 lacZ-a-055 lacZ-a-056 lacZ-a-058 lacZ-a-058 lacZ-a-061 lacZ-a-061 lacZ-a-061	GGATCC		
lacZ-a-048 lacZ-a-049 lacZ-a-050 lacZ-a-053 lacZ-a-055 lacZ-a-055 lacZ-a-056 lacZ-a-058 lacZ-a-068 lacZ-a-060 lacZ-a-061 lacZ-a-062 lacZ-a-062 lacZ-a-063 lacZ-a-064 lacZ-a-065	GGATCC	CGCCGC	
lacZ-a-048 lacZ-a-050 lacZ-a-050 lacZ-a-053 lacZ-a-054 lacZ-a-055 lacZ-a-056 lacZ-a-058 lacZ-a-061 lacZ-a-061 lacZ-a-061 lacZ-a-063 lacZ-a-063 lacZ-a-064 lacZ-a-065		CGGCCGC	
lacZ-a-048 lacZ-a-049 lacZ-a-050 lacZ-a-053 lacZ-a-055 lacZ-a-055 lacZ-a-056 lacZ-a-058 lacZ-a-068 lacZ-a-060 lacZ-a-061 lacZ-a-062 lacZ-a-062 lacZ-a-063 lacZ-a-064 lacZ-a-065			
lac2-a-048 lac2-a-049 lac2-a-050 lac2-a-053 lac2-a-054 lac2-a-055 lac2-a-056 lac2-a-058 lac2-a-058 lac2-a-061 lac2-a-061 lac2-a-061 lac2-a-064 lac2-a-066 lac2-a-066 lac2-a-066 lac2-a-066			
lac2-a-048 lac2-a-050 lac2-a-050 lac2-a-053 lac2-a-054 lac2-a-054 lac2-a-056 lac2-a-056 lac2-a-058 lac2-a-061 lac2-a-061 lac2-a-062 lac2-a-063 lac2-a-063 lac2-a-066 lac2-a-066 lac2-a-066 lac2-a-067 lac2-a-067 lac2-a-067	GGATCC	CGGCCGC  ACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCGAATGCATCGCCGCACCGTACGTCTGAGGAATTCGCCCGAATGCATCGCCGCAGTGTGCTGGAGTATCGCCCGGAATCCGAATTCGAATTCGCCCGGAATGCATCGAATTCGCCCATCGAGGGGTACCGAGCTCGAATTCGCCCGCGCGAATTCGCCCGCGCGAATTCGCGCCGCGAATTCCCGGGTACCGAGCTCGAATTCCCGGCTGCCGCGCGAATTCCCGGATTCCCGGCTGCCGAGCTCGAATTCCCGGCTGCGAATTCCCGGCTGCGAATTCCCGGCTGCGAATTCCCGGCTGCGAATTCCCGGCTGCGAATTCCCGGCTGCGAATTCCCGGCTGCGAATTCCCGGCTGCGAATTCCCGGCTGCGAATTCCCGGCTACCGAGCTCGAATTCCCGGCTACCGAGCTCGAATTC	
lac2-a-048 lac2-a-049 lac2-a-050 lac2-a-053 lac2-a-054 lac2-a-055 lac2-a-056 lac2-a-058 lac2-a-058 lac2-a-061 lac2-a-061 lac2-a-061 lac2-a-064 lac2-a-066 lac2-a-066 lac2-a-066 lac2-a-066	GGATCC		



lacZ-α-060	TGCTAG
lacZ-α-031	TGGCGCGACGCCCTGTAGCGGCGCATTAACCGCGGGGGTGTGGTGTTACGCGCAGCGTGACCGCTACACCTTGCCAGCCCCTCTTTCGCTTCTTCCCTTCCTT
lacz-α-034	TGA
	TGGGGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATA
lacZ-α-036	
lacZ-α-057	TGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAG
lacZ-α-073	
lacZ-α-076	TGGACGCGCCCTGTAGCGGCGCATTAA
lacZ-α-032	TGGCGCGAAATTGTAAACGTTAATGTAATCGATACATGA
lacZ-α-068	
lacZ-α-001	TGCGATTTATTCAACAAAGCCGCCGTCCCGTCAAGTCAGCGTAATGCTCTGCCAGTGTTACAACCAATTCAGA
lacZ-α-004	TGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATACGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAA
lacZ-α-005	TGGGACGCCCCTGTAGCGGCGCATTAAGCCGGCGGGTGTGGTGATACCCGCAGCGTGACCGTACACTTGCCAGCGCCCTAGCGCCCGCTCTTTTCCTTTCTTCTCTTTCTCCCACGTTCGCCGCTTACCCCGCAAGCTCTAAATCGGGGGCTCCTTTAGGGTTC
lacZ-α-006	
	TGGGACGCGCCCTGTASCGGCGCTTTAASCGGCGGGTGGTGGTGGTTACCGCAGCGCTACCTTCCAGCCCCCTCCTTTCCTTTCCTTCC
lacZ-α-010	
lacZ-α-014	TGGAAATTGTAA
lacZ-α-015	${\tt TGGG} \\ \textbf{AGCGCCCTGTAGCGGCGCGTTGGCGGCGGGTGTGGTGTTACCCGCGAGCGTACCGCTACCTGCCAGCCCCTAGCGCCCCGTTCCCTTTCCTTTCCTTTCCTTCC$
lacZ-α-020	TGGCGCTAA
lacZ-α-022	TGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAG
lacZ-α-028	TGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAG
lacZ-α-039	TGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGGCATAGATCCGGTACGTCGTTAA
lacZ-α-045	TGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATAGGGTAA
lacZ=α=046	TGGAAPTTCTAA
lacZ-α-051	TGGAANTGTAA
lacZ-α-051	TIGGGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATATCGCTGGGCCATTCTCATGAAGAATATCTTGAATTTATTGTCATATTACTAG
	TGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGTATTTACACCCGATATATCGCTGGGCCATTCTATGAGAATATCTTGATTTATTGTATTATACTAG
lacZ-α-036corr	
lacZ-α-002	TGGAAATTGTAA
lacZ-α-003	TGGANATTGTAA
lacZ-α-007	TGGGAAATTGTAAACGTTAATATTTTGTTAATATTTTGTTAA
lacZ-α-008	TGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAG
lacZ-α-009	TGA
lacZ-α-011	TGGCGCCTGATGCGGTATTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATACGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAA
lacZ-α-012	TGGAAATTGTAA
lacZ-α-013	TGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAG
lacZ-α-016	
lacZ-α-017	TGGACGCGCCCTGTAGCGGCGCATTAA
lacZ-α-018	TGGAGATCCAATTTTAA
lacZ-α-019	TGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGGGTGTGGTGGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCCTCTTTCCTTTCTTCCTTC
lacZ-α-021	
lacz-q-023	TGGCGCGA AATTGTAAACGTTAATGTTATCGTACCCCTACTCCAAAAATGTCAAAGATACAGTCTCAGAAGAGCCAAAGGGCTATTGA
lacZ-α-023	TGGCGCGABATTGTAAACGTTAATGTTATCGTACCCCTACTCCAAAAATGTCAAGGTCTAGAGGTCCAAAGGCCCTATGA- TGGCGCGGABATTGTABATCGTBATCGTTATCGTACCCCTACTCCAAAAATGTCAAGAGTCCAAGGGCCTATGA- TGGCGCGGABATTGTABATCGTBATCTTGTACTCCTACTACTCAAAAATGTCAAAAGTCCTAAGAGGCCTATTGA- TGGCGCGGABATTGTABATCGTBATCTTGTACCCTTACTCAAAATGTCAAAATGTCAAAATGTCAAAAGGCCTATTGA- TGGCGCGGABATTGTABATCGTBATCTTGTACCCTTACTCAAAATGTCAAAAATGTCAAAAATGTCAAAAATGTCAAAATGTCAAAATGTCAAAATGTCAAAATGTCAAAATGTCAAAATGTCAAAAATGTCAAAAATGTCAAAATGTCAAAATGTCAAAATGTCAAAATGTCAAAATGTCAAAATGTCAAAATGTCAAAAATGTCAAAAATGTCAAAAATGTCAAAATGTCAAAAATGTCAAAATGTCAAAAATGTCAAAATGTCAAAAATGTCAAAAATGTCAAAAATGTCAAAAATGTCAAAAAAATGTCAAAAATGTCAAAAATGTCAAAAAAAA
lacZ-a-041corr	TGGCGCGAAATTGTAAACGTTAATCGTACCCCTACTCCAAAAATGTCAAAGATACAGATCCAAAGGCCAAAGGGCCTATTGA
lacZ-α-04lcorr lacZ-α-024	TGGCGCGAAATTGTAAACGTTAATGTTATCGTACCCCTACTAAAAATGTCAAAGATACAGTCCAGAAGACCAAAGGGCTATTGA
$\begin{array}{l} 1acZ-\alpha-041corr \\ 1acZ-\alpha-024 \\ 1acZ-\alpha-025 \end{array}$	TGGCGCGANATTGTAAACGTTAATGTTATCGTACCCCTACTAAAAATGTCAAAGATACAGTCCAGAAGACCAAAGGGCTATTGA
$\begin{array}{l} {\tt lacZ-\alpha-041corr} \\ {\tt lacZ-\alpha-024} \\ {\tt lacZ-\alpha-025} \\ {\tt lacZ-\alpha-026} \end{array}$	top:top:top:top:top:top:top:top:top:top:
$\begin{array}{l} \texttt{lacZ} - \alpha - \texttt{041corr} \\ \texttt{lacZ} - \alpha - \texttt{024} \\ \texttt{lacZ} - \alpha - \texttt{025} \\ \texttt{lacZ} - \alpha - \texttt{026} \\ \texttt{lacZ} - \alpha - \texttt{027} \end{array}$	TGGGCGARATTOTRACGTTANTGTTATCGTACCCTACTARARTGTARAGATACAGTCTCAGARGACCARAGGGCTATTGA- TGGACGCGCCCTGTAGCGGGGCATTRA- TGGARATTOTRA- TGGARATTOTRA- TGGCGCGCCCTGTAGCGGGGCATTRA
$\begin{array}{l} \text{lac} z - \alpha - 0.41 \text{corr} \\ \text{lac} z - \alpha - 0.24 \\ \text{lac} z - \alpha - 0.25 \\ \text{lac} z - \alpha - 0.26 \\ \text{lac} z - \alpha - 0.27 \\ \text{lac} z - \alpha - 0.27 \\ \text{lac} z - \alpha - 0.29 \end{array}$	TGGCGCGANATGTAACGTTAATGTTATCGTACCCCTACCAAAAATGTCAAAGATACAGTCCAGAAGACCAAAGGGCTATGA
$\begin{array}{l} lac2-\alpha-041corr \\ lac2-\alpha-024 \\ lac2-\alpha-025 \\ lac2-\alpha-026 \\ lac2-\alpha-026 \\ lac2-\alpha-027 \\ lac2-\alpha-029 \\ lac2-\alpha-030 \\ \end{array}$	TGGGCGAAATTGTAAACGTTAATGTTATCGTACCCCTACTCAAAAATGTCAAAGATACAGTCCGAAGACCAAAGGGCTATTGA- TGGACGCGCCCTGTACGGGCGCATTAA- TGGAAATTGTAA TGGAAATTGTAA TGGAAATTGTAA TGGACGCGCCCTTAACGGGGGGGTTAGGGCGGGGGGGGGG
$\begin{array}{l} {\rm lac2} - \alpha - 0.41 {\rm corr} \\ {\rm lac2} - \alpha - 0.24 \\ {\rm lac2} - \alpha - 0.25 \\ {\rm lac2} - \alpha - 0.26 \\ {\rm lac2} - \alpha - 0.26 \\ {\rm lac2} - \alpha - 0.27 \\ {\rm lac2} - \alpha - 0.29 \\ {\rm lac2} - \alpha - 0.30 \\ {\rm lac2} - \alpha - 0.33 \\ \end{array}$	TGGG-CGARATTOTAAACGTTAATGTTATCGTACCCTACTAAAATGTCAAAGATACAGTCCAGAGACCAAAGGGCTATTGA- TGGACGCGCCCTGTAGCGGGCATTAA- TGGAAATTGTAA- TGGAAATTGTAA- TGGAGCGCCCTGTAGCGGCGCATTAA- TGGACGCCCCTGTAGCGGCGCATTAACGGGGGGGGGG
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-029 lacZ-u-033 lacZ-u-035	TGGCGCGANATTGTANACGTTANTGTTATCGTACCCCTACCAAAAATGTCAAAGATACAGTCCAGAAGACCAAAGGGCTATTGA
lacZ-u-04lcorr lacZ-u-024 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-027 lacZ-u-030 lacZ-u-033 lacZ-u-033	TGGG-CGARATTOTRANCGTTANTGTTATCGTACCCTACTACAAAATGTCAAAGATACAGTCCAGAAGACCAAAGGGCTATTGA- TGGACGCGCCTGTAGCGGGCGATTAA- TGGAAATTGTAA- TGGAAATTGTAA- TGGGACGCGCCTGTAGCGGGCGATTAA- TGGGACGCGCCTGTAGCGGGCGATTAA- TGGGACGCGCCTGTAGCGGGCGATTAACGGGGGGGGGG
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-029 lacZ-u-033 lacZ-u-033 lacZ-u-035	TGGCGCGANATTOTANCGTTANTGTTACCGTACCCTACTAAAATGTCAAAGATACAGTCCAGAGACCAAAGGGCTATTGA—  TGGACCGCCCCTTACCGCGCCATTAA—  TGGAAATTGTAAA  TGGAAATTGTAAA  TGGACGGCCCTGTACGGCGCATTAA—  TGGGAAATTGTAAA  TGGGACGGCCTGTACGGCGCATTAA—  TGGGACGGCCTGTACGGCGCATTAA—  TGGGACGGCCTGTACGGCGCATTAA—  TGGGACGGCCTGTACGGCGCATTAA  TGGGACGGCCTGTACGGCGCATTAA  TGGGACGGCCTGTACGGCGCATTAACGGCACGTGCGCCGCTACACTTGCCACGCCCCTACCGCCCCTTTCCTTTCTCCTTCCT
lacZ-a-04lcorr lacZ-a-024 lacZ-a-025 lacZ-a-025 lacZ-a-027 lacZ-a-027 lacZ-a-030 lacZ-a-030 lacZ-a-033 lacZ-a-037 lacZ-a-037	TGGGCGAAATTGTAAACGTTAATGTTATCGTACCCCTACCAAAAATGCAAAGATACAGTCCAGAAGACCAAAGGGCTATGA— TGGACGCCCCTGTAGCGGGCGATTAA— TGGGAGGGCCATTAACGTGCGGGGGGTATAACGGGGGGGGGG
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-029 lacZ-u-033 lacZ-u-033 lacZ-u-035	TGG-CCCTTATACGGGGCATTAA- TGGACGCCCTGTAGCGGGCATTAA- TGGACGCCCTGTAGCGGGCATTAA- TGGACGCCCTGTAGCGGGCATTAA- TGGACGCCCTGTAGCGGGGCATTAA- TGGACGCCCTGTAGCGGGGCATTAA- TGGACGCCCTGTAGCGGGGCATTAA- TGGACGCCCTGTAGCGGGGATTAACGCGGGGGGGGGGG
lacZ-a-04lcorr lacZ-a-024 lacZ-a-025 lacZ-a-025 lacZ-a-027 lacZ-a-027 lacZ-a-030 lacZ-a-030 lacZ-a-033 lacZ-a-037 lacZ-a-037	GGGGGAAATTGTAAACGTTAATGTTATCGTACCCCTACCAAAAATGTCAAAGATACAGTCCAGAAGACCAAAGGGCTATGA TGGACGCCCCTGTAGCGGCGCATTAA- TGGCAGGGCCCTTAACGGGGCGATTAACGGGGGGGGGG
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-033 lacZ-u-033 lacZ-u-035 lacZ-u-036 lacZ-u-041	GGGCGANATTOTAACGTTAATCTTATCGTACCCTACTAAAAATGTCAAAGATACAGTCCAGAAGACCAAAGGGCTATTGA— TGGACGCGCCCTGTAGCGGCGCATTAA— TGGAAATTGTAA- TGGAAATTGTAA- TGGGAAATTGTAACGCGCCGTTAAGGGGGGGGGGGGG
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-029 lacZ-u-033 lacZ-u-033 lacZ-u-035 lacZ-u-036 lacZ-u-0404 lacZ-u-041 lacZ-u-041	TGGCGCGANATTOTAAACGTTAATCTTATCGTACCCTACAAAATGTCAAAGATACAGTCCAGAAGACCAAAGGGCTATTGA—  TGGACGCGCCCTGTAGCGGCGATTAA—  TGGAAATTGTAAA  TGGAAATTGTAAA  TGGGAAATTGTAACGTTAACGGGGGGTGGTGGTGGTGGTACGGCAGCGTACACTTGCCAGCCCCTAGGCCCCTATCCTTTCTCCTTCCT
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-033 lacZ-u-033 lacZ-u-035 lacZ-u-040 lacZ-u-040 lacZ-u-040	GGGCGARATTOTAAACGTTAATGTTATCGTACCCCTACCAAAAATGTCAAAGATACGTCAGAAGACCAAAAGGGCTATGA— TGGACGGCCCCTGTAGCGGCGCATTAA— TGGCACGCCCTGTAGCGGCGCATTAAGCGCGGGGGGGGGG
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-029 lacZ-u-033 lacZ-u-033 lacZ-u-035 lacZ-u-037 lacZ-u-041 lacZ-u-044 lacZ-u-044 lacZ-u-044 lacZ-u-044	TGGACGCCCTGTAGCGGCATTAA- TGGAATTGTAACGTTAACGTTAACGCGCAGGGGGGGGGG
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-030 lacZ-u-031 lacZ-u-037 lacZ-u-037 lacZ-u-040 lacZ-u-040 lacZ-u-040 lacZ-u-044 lacZ-u-044 lacZ-u-044	GGGGGAAATTGTAAACGTTAATGTTATCGTACCCCTACCAAAAATGCAAAGATCAAGATCAAGATCAAGACCCAAAGGCCATTGCAGGCCCTAGGGCCCTTTCCTTTCTCCTTCCT
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-033 lacZ-u-033 lacZ-u-037 lacZ-u-040 lacZ-u-040 lacZ-u-041 lacZ-u-044 lacZ-u-044 lacZ-u-044	GGGCGANATTOTANACGTTANTCTTATCGTACCCTACTACAAAATGTCAAAGATACAGTCCAGAAGACCAAAGGGCTATTGA— TGGAGCGGCCCTGTAGCGGGGCATTAAGCGGGGGGGTGGTGGTGTACGGGAGGTGACCGTACACTTGCAGCCCCTAGGGCCCCTTTCGTTTCTCCTTCCT
lacZ-u-04lcorr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-033 lacZ-u-033 lacZ-u-035 lacZ-u-040	GGGCGANATTOTANACGTTANTCTTATCGTACCCTACTACAAAATGTCAAAGATACAGTCCAGAAGACCAAAGGGCTATTGA— TGGAGCGGCCCTGTAGCGGGGCATTAAGCGGGGGGGTGGTGGTGTACGGGAGGTGACCGTACACTTGCAGCCCCTAGGGCCCCTTTCGTTTCTCCTTCCT
lac2-a-04lcorr lac2-a-024 lac2-a-025 lac2-a-026 lac2-a-027 lac2-a-027 lac2-a-030 lac2-a-030 lac2-a-035 lac2-a-035 lac2-a-036 lac2-a-040 lac2-a-040 lac2-a-044 lac2-a-044 lac2-a-044 lac2-a-044 lac2-a-044 lac2-a-044 lac2-a-044 lac2-a-044 lac2-a-044 lac2-a-045 lac2-a-055	TGGGGGCAAATTGTAAACGTTAATGTTATCGTACCCCTACCAAAAATGTCAAAGATACGTCAGAAGACCAAAAGGCCATTGGGCCCTAGGCCCTAGGCCCTTTCCTTTCTCCTTCCT
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-030 lacZ-u-031 lacZ-u-037 lacZ-u-037 lacZ-u-040 lacZ-u-041 lacZ-u-044 lacZ-u-044 lacZ-u-044 lacZ-u-048 lacZ-u-047 lacZ-u-048 lacZ-u-047 lacZ-u-047 lacZ-u-047 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-050	GGGCGANATTOTANACGTTANTOTTATCGTACCCTACAAAATGCAAAGATCAAGATCAAGATCAAGACCAAAAGGCCATAGACCCTATGCACCCCTGTAGCGGCGCGCTTTCCTTCTCGCCACGTTCCCTTCCTT
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-033 lacZ-u-033 lacZ-u-037 lacZ-u-041 lacZ-u-041 lacZ-u-044 lacZ-u-044 lacZ-u-044 lacZ-u-046 lacZ-u-047 lacZ-u-048 lacZ-u-049 lacZ-u-050 lacZ-u-050 lacZ-u-050	GGGGGGAAATTGTAAACGTTAATGTTATCGTACCCCTACTCCAAAAATGTCAGAAGACCAAAAGGCCATTGGCGCCCTTTCGCTTCTTCCCTCCTTTCTCCCACGTTCGCCGGCTTTCCCTTAGGGGCCCCTTTAGGGTCCTTTGGCAAGTCCAAAATGTGAAATTGGGGAATTAAATTGGGAAAATTGTAAAATTGGAAAAACGTTAAATTGGGGATGGCGACGCGACCTTAGGGCCCCTTTGGCAGGCCCGCCC
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-030 lacZ-u-031 lacZ-u-037 lacZ-u-037 lacZ-u-040 lacZ-u-041 lacZ-u-044 lacZ-u-044 lacZ-u-044 lacZ-u-048 lacZ-u-047 lacZ-u-048	GGGCGANATTOTANACGTTANTOTTATCGTACCCTACAAAATGCAAAGATCAAGATCAAGATCAAGACCAAAAGGCCATAGACCCTATGCACCCCTGTAGCGGCGCGCTTTCCTTCTCGCCACGTTCCCTTCCTT
lacZ-u-04lcorr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-033 lacZ-u-033 lacZ-u-037 lacZ-u-037 lacZ-u-040 lacZ-u-041 lacZ-u-041 lacZ-u-044 lacZ-u-045 lacZ-u-045 lacZ-u-054 lacZ-u-055 lacZ-u-055 lacZ-u-055 lacZ-u-056	GGGCGCAAATTGTAAAGGTTAATGTTATCGTACCCTCACTCCAAAAATGTCAAAGATACAGTCTCAGAAAGCCAAAGGGCTATTGA— TGGACCCGCCCTTGTAGCGGGCCATTAA— TGGACCCGCCCTTGTAGCGGGCCATTAA— TGGACCCGCCCTTGTAGCGGGCCATTAA— TGGACCCGCCCTTGTAGCGGCGCATTAA— TGGACCCGCCCTTGTAGCGGCGCATTAA— TGGACCCGCCTTGTAGCGGCGCATTAA— TGGACCCGCCCTTGTAGCGGCGCATTAA— TGGACCCGCCCTTGTAGCGGCGCATTAA— TGGACCCGCCTTGTAGCGGCGCATTAACGCGCAGGGTAGCCCTACACTTCCCACCTCCCTAGCGCCCTCTTTTCTCCTTCCT
lacZ-a-04lcorr lacZ-a-024 lacZ-a-025 lacZ-a-025 lacZ-a-026 lacZ-a-027 lacZ-a-027 lacZ-a-030 lacZ-a-030 lacZ-a-031 lacZ-a-037 lacZ-a-037 lacZ-a-040 lacZ-a-041 lacZ-a-044 lacZ-a-044 lacZ-a-046 lacZ-a-047 lacZ-a-047 lacZ-a-056 lacZ-a-055	TGGC-AGCGCCTTGTAGGGGGCATTAA TGG-AGCGCCCTTGAGGGGGCATTAA TGG-AGCGCCCTTGAGGGGGCATTAA TGG-AGCGCCCTTGAGGGGGGCATTAA TGG-AGCGCCCTTGAGGGGGGGGATAAGGGGGGGGGGGGG
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-033 lacZ-u-033 lacZ-u-037 lacZ-u-040 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-056 lacZ-u-056 lacZ-u-058 lacZ-u-058 lacZ-u-059	GGGCGCAAATTGTAAAGGTTAATGTTATCGTACCCTCACTCCAAAAATGTCAAAGATACAGTCTCAGAAAGCCAAAGGGCTATTGA— TGGACCCGCCCTTGTAGCGGGCCATTAA— TGGACCCGCCCTTGTAGCGGGCCATTAA— TGGACCCGCCCTTGTAGCGGGCCATTAA— TGGACCCGCCCTTGTAGCGGCGCATTAA— TGGACCCGCCCTTGTAGCGGCGCATTAA— TGGACCCGCCTTGTAGCGGCGCATTAA— TGGACCCGCCCTTGTAGCGGCGCATTAA— TGGACCCGCCCTTGTAGCGGCGCATTAA— TGGACCCGCCTTGTAGCGGCGCATTAACGCGCAGGGTAGCCCTACACTTCCCACCTCCCTAGCGCCCTCTTTTCTCCTTCCT
lac2-a-04lcorr lac2-a-024 lac2-a-025 lac2-a-026 lac2-a-027 lac2-a-027 lac2-a-030 lac2-a-030 lac2-a-035 lac2-a-035 lac2-a-036 lac2-a-040 lac2-a-040 lac2-a-044 lac2-a-044 lac2-a-044 lac2-a-046 lac2-a-055 lac2-a-056 lac2-a-055 lac2-a-055 lac2-a-055 lac2-a-056 lac2-a-059 lac2-a-059 lac2-a-059 lac2-a-059 lac2-a-061	TGGC-ACGCCATTACGTACCCCTATCCAAAAATTTCAACCCCAACTCCAACACCCAACCCCTAACCCCAACCCCTAACCCCCACCCCCC
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-030 lacZ-u-031 lacZ-u-037 lacZ-u-037 lacZ-u-040 lacZ-u-040 lacZ-u-040 lacZ-u-040 lacZ-u-040 lacZ-u-041 lacZ-u-040 lacZ-u-040 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-055 lacZ-u-056 lacZ-u-056 lacZ-u-059 lacZ-u-059 lacZ-u-061 lacZ-u-063	TGGC-CANTITOTANCOTTANTOTATCOTACCCTACTACAGAACAGATCCAGAAGATCCAGAAGACCATAGGCCCTATGCCTTCTTCTCCTTCTTCCCACGTTGCGCCCTTAAGGCCCCACGTTAAACGGGGCCCCTTAAGGGGCCCCTTAAGGCCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGCCCTTAAGGGCCCTTAAGGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGCGCCCCTTAAGCGCCCCCTTAAGCGCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCTTAAGCCCCCCTTAAGCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCCC
lacZ-u-04lcorr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-033 lacZ-u-033 lacZ-u-037 lacZ-u-037 lacZ-u-040 lacZ-u-041 lacZ-u-041 lacZ-u-044 lacZ-u-045 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-050 lacZ-u-056 lacZ-u-056 lacZ-u-058 lacZ-u-058 lacZ-u-058 lacZ-u-058 lacZ-u-061 lacZ-u-062 lacZ-u-061 lacZ-u-062 lacZ-u-063	GGG-CARGGCCCTATAGGGGCCATTAA- TGGG-CACGGCCCTAGGGGCCATTAA- TGGGGCCTAGGGGCCATTAA- TGGGCCCTAGGGGCCATTAA- TGGGCCCTAGGGGCCATTAA- TGGGCCCTAGGGGCCATTAA- TGGGCCCTAGGGGCCATTAA- TGGGCCCTAGGGGCCATTAA- TGGGCCCTAGGGGCCATTAA- TGGGCCCTAGGGGCCATTAA- TGGGCCCTAGGGGCCATTAA- TGGGCCCTAGGGGCCATTAA- TGGGCCACGGCCCTAGGGGCCATTAA- TGGGCCACGGCCCTAGGGGCCATTAA- TGGGCCACGGCCCTAGGGGCCATTAA- TGGGCCACGGCCCTAGGGGCCATTAA- TGGGCCACGGCCCTAGGGGCCATTAA- TGGGCCACGGCCCTAGGGGCCATTAA- TGGGCCACGGCCCTAGGGGCCATTAA- TGGGCCACGGCCCTAGGGGCCATTAA- TGGGCCACGGCCCTAGGGGCCATTAA- TGGCCACGGCCCTAGGGGCCCATTAA- TGGCCACGGCCCTAGGGGCCATTAA- TGGCCCACGGCCCTAGGGGCCCATTAA- TGGCCCACGGCCCTAGGGGCCCATTAA- TGGCCCCACGGCCCTAGGGGCCCATTAA- TGGCCCCACGGCCCTAGGGGCCCATTAA- TGGCCCACGGCCCTAGGGGCCCATTAA- TGGCCCACGGCCCTAGGGGCCCATTAA- TGGCCCACGGCCCTAGGGGCCCATTAA- TGGCCCACGGCCCTAGGGGCCCATTAA- TGGCCCTAGGGCCCTAGGGGCCCATTAA- TGGCCCTAGGGGCCCTAGGGGCCCATTAA- TGGCCCACGGCCCTAGGGGCCCATTAA- TGGCCACTGCCTAGGGGCCCATTAA- TGGCCACTGCCTTGGTAA- TGGCCACTGCCCTTGGAGGCCCATTAA- TGGCCCTAGGGCCCTAGGGGCCCATTAA- TGGCCACTGCCCTTGGAA- TGGCCACTGCCCTTGGAA- TGGCCACTGCCCTTGGAA- TGGCCACTGCCCTTGGAA- TGGCCACTGCCCTTGGAA- TGGCCACTGCCCTTGGAA- TGGCCACTGCCCTTGAGGCGCCATTAA- TGGCCCTTGCCTTGGAA- TGGCCACTGCCCTTGGAA- TGGCCACTGCCCTTGGAA- TGGCCACTGCCCTTGGAA- TGGCCACTGCCCTTGGAA- TGGCCACTGCCCCTTGAGGGCCCATTAA- TGGCCCTTGCTAGGGCCCATTAA- TGGCCCTTGCTAGA- TGGCCCTTGCTAGGCCCCTTGGAA- TGGCCCTTGCTAGGGCCCATTAA- TGGCCCTTGCTAGA- TGGCCTTGCTAGGCCCCTTGGAA- TGGCCCTTGCTAGGGCCCATTAA- TGGCCCTTGCTAGGCCCCTTGCAGCGCCCATTAA- TGGCCCTTGCCTTGCTAGA- TGGCCCTTGCTAGGCCCCTTGCAGCGCCATTAA- TGGCCCTTGCCCTTGCTAGA- TGGCCCTTGCCCCTTGCAGCGCCCATTAA- TGGCCCTTGCCCCTTGCAGCGCCCATTAA- TGGCCCTTGCCCTTGCAGCGCCCATTAA- TGGCCCTTGCCCCTTGCAGCGCCCATTAA- TGGCCCTTGCCCCTTGCAGCGCCCATTAA- TGGCCCTTGCCCCTTGCAGCGCCCATTAA- TGGCCCTTGCCCCTTGCCCCCCCATTAA- TGGCCCTTCCCCCTCAGCGCCCATTAA- TGGCCCTTGCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
lacZ-a-04lcorr lacZ-a-024 lacZ-a-025 lacZ-a-025 lacZ-a-026 lacZ-a-027 lacZ-a-030 lacZ-a-030 lacZ-a-033 lacZ-a-037 lacZ-a-037 lacZ-a-040 lacZ-a-041 lacZ-a-044 lacZ-a-044 lacZ-a-044 lacZ-a-045 lacZ-a-055 lacZ-a-056 lacZ-a-056 lacZ-a-056 lacZ-a-056 lacZ-a-056 lacZ-a-056 lacZ-a-056 lacZ-a-066 lacZ-a-066 lacZ-a-066 lacZ-a-066 lacZ-a-066 lacZ-a-066	TGGC-CANTITOTANCOTTANTOTATCOTACCCTACTACAGAACAGATCCAGAAGATCCAGAAGACCATAGGCCCTATGCCTTCTTCTCCTTCTTCCCACGTTGCGCCCTTAAGGCCCCACGTTAAACGGGGCCCCTTAAGGGGCCCCTTAAGGCCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGGCCCTTAAGGGCCCTTAAGGGCCCTTAAGGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGGGCCCCTTAAGCGCCCCTTAAGCGCCCCCTTAAGCGCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCCCCCCTTAAGCGCCCCCTTAAGCGCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCTTAAGCCCCCCTTAAGCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCTTAAGCCCCCCCC
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-027 lacZ-u-033 lacZ-u-033 lacZ-u-033 lacZ-u-037 lacZ-u-040 lacZ-u-041 lacZ-u-041 lacZ-u-044 lacZ-u-044 lacZ-u-045 lacZ-u-045 lacZ-u-055 lacZ-u-056 lacZ-u-056 lacZ-u-056 lacZ-u-058 lacZ-u-056 lacZ-u-058 lacZ-u-056 lacZ-u-058 lacZ-u-058 lacZ-u-061 lacZ-u-063 lacZ-u-066	TGGG-ACCCCCTGTACGCCGCATTA. TGG-ACCCCCCTGTACGCGCGATTA. TGG-ACCCCCTGTACGCGCGATTAA. TGG-ACCCCCTGTACGCGCGATTAA. TGG-ACCCCCTGTACGCGCGATTAA. TGG-ACCCCCTGTACGCGCGATTAA. TGGG-ACCCCCTGTACGCGCGATTAA. TGGGCCCCGTGTACGCGCGATTAA. TGGGCCCCTGTACGCGCGATTAA. TGGGCCCCTGTACGCGCGATTAA. TGGCCCCTGTACGCGCGATTAA. TGGCCCTTGTACGCGCGATTAA. TGGCCCTTGTACGCGCGATTAA. TGGCCCTTGTACGCGCGATTAA. TGGCCCTTGTACGCGCGATTAA. TGGCCCTTGTACGCGCGATTAA. TGGCCCTTGTACGCGCCTTGTACGCGCGATTAA. TGGCCCTTGTACGCGCCTTGTACGCGCGATTAA. TGGCCCTTGTACGCGCCTTGTACGCGCGATTAA. TGGCCCTTGTACGCGCCTTGTACGCGCGATTAA. TGGCCCTTGTACGCGCCTTGTACGCGCGATTAA. TGGCCCTTGTACGCGCCTTGTACGCGCGATTAA. TGGCCCTTGTACGCGCCTTGTACGCGCGATTAA. TGGCCTTCCTTGTAA. TGGCCCTTCCTTGTAA. TGGCCCTTCCTTCGTAA. TGCCCTTCCTTCGTAA. TGCCCTTCCTTCGTAA. TGCCCTTCCTTCGTAA. TGCCCTTCCTTCGTAA. TGCCCTTCCTTCGTAA. TGCCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTT
lacZ-a-04lcorr lacZ-a-024 lacZ-a-025 lacZ-a-025 lacZ-a-026 lacZ-a-027 lacZ-a-027 lacZ-a-030 lacZ-a-030 lacZ-a-031 lacZ-a-037 lacZ-a-037 lacZ-a-038 lacZ-a-040 lacZ-a-041 lacZ-a-044 lacZ-a-044 lacZ-a-047 lacZ-a-054 lacZ-a-055 lacZ-a-056 lacZ-a-056 lacZ-a-056 lacZ-a-056 lacZ-a-056 lacZ-a-066 lacZ-a-066 lacZ-a-066 lacZ-a-066 lacZ-a-066 lacZ-a-066 lacZ-a-066 lacZ-a-066	TSGGGGAAATTGTAAACGTAATGTATCTGACCCCTACTCCAAAATGTCAAAGACCAAAGGCTATTCACTGGGGCCGCTCTTCCCTTTCTCCCCACGTTGCCGCGTTTCCCGCAACCTCAAACCGGGGCCCCTTTAGGGTTCCCGCAACCTCCCTTAGGGTTCCCGCAACCTCCAAACCTCCAAACCGGGCCCCTTTAGGGTTCCCGCCAACCTCCCTTTCCCCCAACCTCCCAACCTCCCAACCTCCCTTTCGCCAACCTCCCTTTCCCCCAACCTCCCAACCTCCCAACCTCCCTTTAGGGTTCCCGCCAACCTCCCTTTAGGGTTCCCGCCAACCTCCCTTTAGGGTTCCCGCCAACCTCCCTTTAGGGTTCCCGCCAACCTCCCTTTAGGGTTCCCGCCAACCTCCCTTTAGGGTTCCCGCCAACCTCCCAACCTCCAACCTCCCTTTAGGGTTCCCGCCAACCTCCCAACCTCCAACCTCCAACCTCCAACCTCCAACCTCCAACCTCCAACCTCCCCAACCTCCCCAACCTCCCAACCCCCAACCCCCAACCCCCAACCCCCAACCCCCAACCCC
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-030 lacZ-u-031 lacZ-u-037 lacZ-u-037 lacZ-u-041 lacZ-u-041 lacZ-u-044 lacZ-u-044 lacZ-u-045 lacZ-u-056 lacZ-u-056 lacZ-u-056 lacZ-u-056 lacZ-u-056 lacZ-u-056 lacZ-u-056 lacZ-u-056 lacZ-u-066 lacZ-u-066 lacZ-u-066 lacZ-u-066 lacZ-u-066 lacZ-u-066 lacZ-u-066	TGGGGCAANTTGTAAACGTTAATGTTATCGTCACCTTACCCAAAAATCTCAAAAATCACCTACAAGGCCATTGA.  TGGGACCGCCCCTGTAGGGGGCATTAA.  TGGGACCGCCCTGTAGGGGGCATTAAA  TGGGACCGCCCTGTAGGGGGCATTAAATCTGCGGAAGCGTAACGTTACCACCTTACCTTCCTGCTTCTTCCTCCTTCTTCCGCACGTTCCCGGAAACTCAAATCGGGGCCCCTTTAGGGTTACCACCTTAGGGTTCCCGGAAATCTTACCATTACGGGAAACCCAATTAATCTGCGAAACTCGCAAGATCTCAAACCGGAAATCTGAAACCGGAAACCCAATTAACCGGGAACCGAAATCTGCAAGACCCCAAAACCCAAATTAATCTGCGCAAACCCAAATTAACCACAAATTAACCACCAAATTAACCACC
lac2-a-041corr lac2-a-024 lac2-a-025 lac2-a-026 lac2-a-027 lac2-a-027 lac2-a-030 lac2-a-031 lac2-a-035 lac2-a-035 lac2-a-036 lac2-a-040 lac2-a-041 lac2-a-044 lac2-a-044 lac2-a-044 lac2-a-055 lac2-a-056 lac2-a-056 lac2-a-055 lac2-a-056 lac2-a-059 lac2-a-066	TSGGCCAATTTTTATACTTAACGTTATCTTACCTCACCCCATTAL  GGACCGCCCTGATACGCCCATTAL  GGACCGCCCTGATACGCCCATTAL  GGACCGCCCTGATACGCCCATTAL  GGACCGCCCTGATACGCCCCATTACGCCCATTACGCCACACCCATACCCCATACCCCATACCCCCTTACGCCCCTTCCCTTCCCTTCCCCCCCC
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-030 lacZ-u-031 lacZ-u-037 lacZ-u-037 lacZ-u-041 lacZ-u-041 lacZ-u-042 lacZ-u-044 lacZ-u-045 lacZ-u-045 lacZ-u-046 lacZ-u-047 lacZ-u-047 lacZ-u-048 lacZ-u-049 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-066 lacZ-u-066 lacZ-u-066 lacZ-u-066 lacZ-u-067	TGGG-CANTTOTAMCOTTANTOTTACCCTACCCCATTA  GG
lac2-a-04lcorr lac2-a-024 lac2-a-025 lac2-a-026 lac2-a-027 lac2-a-027 lac2-a-030 lac2-a-030 lac2-a-035 lac2-a-035 lac2-a-038 lac2-a-040 lac2-a-041 lac2-a-044 lac2-a-044 lac2-a-044 lac2-a-045 lac2-a-055 lac2-a-056 lac2-a-056 lac2-a-056 lac2-a-056 lac2-a-056 lac2-a-056 lac2-a-056 lac2-a-056 lac2-a-066 lac2-a-066 lac2-a-066 lac2-a-066 lac2-a-066 lac2-a-069 lac2-a-070 lac2-a-071 lac2-a-071 lac2-a-071	TRGGGCGAATTTATATGTTATACGTTATGTTATCGTTACCCCTATACAAAAATGCCAAAAAGTCCAAAAAGTCCAAAAAGTCCAAAAAGCCCTAGCAAGGCCCATTACACCCCCCCTTACGGCCCCGTTACACCCCCCCTTACGGCCCCGTTACACCCCCCCTTACGGCCCCGTTACACCCCCCCTTACGGCCCCGTTACACCCCCCCTTACGGCCCCGTTACACCCCCCCTTACGGCCCCGTTACACCCCCCTTACGGCCCCTTACGGCCCCCTTACGGCCCCTTACGGCCCCTTACGGCCCCTTACGGCCCCCTTACGGCCCCTTACGGCCCCCTTACGGCCCCCTTACGGCCCCCTTACGGCCCCCTTACGGCCCCCTTACGGCCCCTTACGGCCCCCCTTACGGCCCCCCTTACGGCCCCCCCC
lacZ-u-041corr lacZ-u-024 lacZ-u-025 lacZ-u-025 lacZ-u-026 lacZ-u-027 lacZ-u-030 lacZ-u-030 lacZ-u-031 lacZ-u-037 lacZ-u-037 lacZ-u-041 lacZ-u-041 lacZ-u-042 lacZ-u-044 lacZ-u-045 lacZ-u-045 lacZ-u-046 lacZ-u-047 lacZ-u-047 lacZ-u-048 lacZ-u-049 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-059 lacZ-u-066 lacZ-u-066 lacZ-u-066 lacZ-u-066 lacZ-u-067	TGGG-CANTTOTAMCOTTANTOTTACCCTCCCCTCACAMATTCTCACAGACCCACACCCTCACAGCCCCTACGCCCCCTTCCCCTTCCCTTCCCTTCCCTTCCCTCCC

Sequences in yellow correspond to the f1 origin of replication, which is in frame with  $lacz-\alpha$  in these cases.

 $1acZ-\alpha-016$  and  $1acZ-\alpha-021$  are missing STOP codons because these variants are in frame fusions with ccdB. -016 identical to -053. -021 identical to -029 Sequences in bold are incorrectly annotated

 $1acZ-\alpha-041$  sequence has a single bp deletion that generates a frame-shift and premature STOP codon. Very likely sequence error corrected above. -041 merged with -023  $1acZ-\alpha-036$  is annotated without a STOP codon. Sequence actually extends another 54 bp and is identical to  $1acZ-\alpha-052$ . -036 merged with -052.

Name of Variant	Alignment	Plasmids
$ \begin{array}{l} 1acZ-\alpha-040 \\ 1acZ-\alpha-062 \\ 1acZ-\alpha-064 \end{array} $	ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCGAATTCACTAGT- ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCGAATTCACTAGTA	pGEM-T pSpark Done (linearized) pSpark Done (linearized)
$lacz-\alpha-048$ $lacz-\alpha-066$	ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCCCATATGGTCGACCTGCAGGCGGCCGCACTAGT- ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCACTAGTA	pGEM-T Easy pSpark II (linearized) pSpark II (linearized)
$1acz-\alpha-069$ $1acz-\alpha-073$ $1acz-\alpha-076$	ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTACTCGAGCTCAAACATATGAAAGTCGACCTGCAGGCGGCCGCACTAGTG	pSpark III, IV and V (linearized) pSpark III, IV and V (linearized)
lacZ-α-040	GATATCGAATTCCCGCGGCCGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTT	
$ \begin{array}{l} 1acZ-\alpha-062\\ 1acZ-\alpha-064\\ 1acZ-\alpha-048 \end{array} $	GAT	
lacZ-α-066 lacZ-α-069 lacZ-α-073	GAT	
lacZ-α-076	ATCAGGCCTGATCCGGAGAATTCGTTTAAAC-CCATGGGGGCCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTT	
lacZ-α-040	${\tt TACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAG} GACGCGCCCTGTAGCGGCGCATTAA$	
$ \begin{array}{l} 1acZ-\alpha-062\\ 1acZ-\alpha-064\\ 1acZ-\alpha-048 \end{array} $	TACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGACGCGCCCTGTAGCGGCGCATTAA TACAACGTCGTGACTGGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGACGCGCCCTGTAGCGGCGCATTAA	
lacZ-α-066 lacZ-α-069 lacZ-α-073	TACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGACGCGCCCTGTAGCGGCGCATTAA	
lacZ-α-076	${\tt TACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGACGCGCCCTGTAGCGGCGCATTAACCCAGCTGCGCGCATTAACCGAACACTCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGACGCGCCCTGTAGCGGCGCATTAACCCAGCTGCGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGACGCGCCCTGTAGCGGCGCATTAACCGAACACTCAACTC$	
Combined lacZ-α Fr	ragments from Recircularized pSpark Series	
Name of Variant	Alignment Common Restriction Enzyme Sites in MCS	
$\begin{array}{l} 1acZ - \alpha - 040 \\ 1acZ - \alpha - 062 + 064 \\ 1acZ - \alpha - 048 \\ 1acZ - \alpha - 066 + 069 \\ 1acZ - \alpha - 073 + 076 \end{array}$	Ser Net Sair PSt Not EcoRI Spei ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCGAATTCACTA ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCGAATTCACTA ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCACTA ATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCACTA ATGACCATGATTACGCCAAGCTATTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGCGGAGCTCCCATATGGTCGACCTGCAGGCGGCCGCACTA ATGACCATGATTACGCCAAGCTATTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTACTCGACCTGCATAGAAACTTACGCAGGCGGCCGCACTA ATGACCATGATTACGCCAAGCTATTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTACTCGACTGCATAGAAACTTACGACAGGCGGCCGCACTA ATGACCATGATTACGCCAAGCTATTAGGTGACACTATAGAATACTCAAGCTATCCAACGCGTTACTCCAATGCATATGCAATACTAGAATACTCAAGCTATCAAGCTATCAAGCTATCCAACGCGTTACTCGAATACTATGCATCCAACGCGTTACTCGACTGCAAGCCTGCCAGGCGGCCGCACTA ATGACCATGATTACGCCAAGCTATTAGGATACTCAAGCTATCCAACGCGTTACTCCAACGCGTTACAATACTAAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATCAACGCTTACTCCAATACTAAACTATAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATTAGAATACTCAAGCTATAGAATACTCAAGCTATTAGAATACTCAAGCTATAGAATACTCAAGCTATTAGAATACTAGAATACTCAAGCTATAGAATACTCAAGCTATAGAATACTCAAGCTATAGAATACTCAAGCTATAGAATACTCAAGCTATAGAATACTCAAGCTATAGAATACTCAAGCTATAGAATACTCAAGCTATAGAATACTCAAGCTATAGAATACTCAAGCTATAGAATACTCAAGCATATTAGAATACTAGAATACTAGAATACTCAAGCTATAGAATACTAGAATACTAGAATACTAGAATACTAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAGAATACTAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACTAAGAATACT	
$\begin{array}{l} {\rm lac}{\rm Z}\!-\!\alpha\!-\!040 \\ {\rm lac}{\rm Z}\!-\!\alpha\!-\!062\!+\!064 \\ {\rm lac}{\rm Z}\!-\!\alpha\!-\!064 \\ {\rm lac}{\rm Z}\!-\!\alpha\!-\!066\!+\!069 \\ {\rm lac}{\rm Z}\!-\!\alpha\!-\!073\!+\!076 \end{array}$	ECORV SACII ECORI SACII NOTI Sph1  GTGATATC	
$\begin{array}{l} \texttt{lacZ} - \alpha - 040 \\ \texttt{lacZ} - \alpha - 062 + 064 \\ \texttt{lacZ} - \alpha - 048 \\ \texttt{lacZ} - \alpha - 066 + 069 \\ \texttt{lacZ} - \alpha - 073 + 076 \end{array}$	TTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCCCCTGTAGCGGCGCATTAA TTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCCCCTGTAGCGGCGCATTAA TTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACTCCCAACAGTTGCGCAGCCTGAATGGCAATGGACGCCCCTGTAGCGGCGCATTAA TTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCCCCTGTAGCGGCGCATTAA TTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCCCCTGTAGCGCGCCATTAA ********************************	

 $1acz-\alpha-062$ ,  $1acz-\alpha-066$  and  $1acz-\alpha-073$  are linearized plasmids and have added A at end possibly because vector is like Topo with 3' A overhang to make PCR cloning more efficient. A is removed in alignment. Plasmid comes linearized by EcoRV.

 $\texttt{lacZ}-\alpha-\texttt{064, lacZ}-\alpha-\texttt{069} \text{ and } \texttt{lacZ}-\alpha-\texttt{076 missing A} \text{ at start to complete EcoRV site of linearization (blunt) added in alignment.}$ 

Name of Variant		Size
variant lacZ-α-041		(aa) 113
lacZ-α-041 corr		141
lacZ-α-023		145
lacZ-α-032	MTMITPSSELTLTKGNKSWSSTAVAAALELVDPPGCRNSISSLSIPSTSRGGP	129
lacZ-α-047		216
$lacz-\alpha-012$		120
lacZ-α-015		191
lacZ-α-031		192
lacZ-α-005		191
lacZ-α-042		191
lacZ-α-019		190
lacZ-α-037		191
lacZ-α-010	MTMITPSAQLTLTKGNKSWVPGPPSRSTVSISLISNSCSPGD	191
lacZ-α-055		129
lacZ-α-038		94
lacZ-α-028 lacZ-α-017		126 128
lacz-α-017		170
lacZ-α-008		116
lacZ-α-030		107
lacZ-α-065	MTMITPPRVPSSNSLAVVLQRRDWENPGVTQLNRLAAHPP	168
lacZ-α-009	MTMITPSLHACRLAVVLQRRDWENPGVTQLNRLAAHPP	76
lacZ-α-057	MTMITPPRVPSSNSLAVVLQRRDWENPGVTQLRLAAHPP	107
lacZ-α-036		95
lacZ-a-036corr/052		112
lacZ-α-011 lacZ-α-001		384 108
1acz-α-001 1acz-α-072		158
lacZ-α-070		88
lacZ-α-035		127
lacZ-α-033	MTMITPSSNTTHYRESLHACRLAVVLQRRDWENPGVTQLNRLAAHPP	110
lacZ-α-075		96
lacZ-α-022	MTMITPVPSSNSLAVVLQRRDWENPGVTQLNRLAAHPP	115
lacZ-α-059		80
lacZ-α-016 lacZ-α-029		120 114
lacz-α-029		120
lacZ-α-021		117
lacZ-α-071	MTMITPSYLGDTIEYSSYASSLVPSSDPLVTAASVLEFCRYPSHWRPLEH	201
lacZ-α-040	MTMITPSYLGDTIEYSSYASNALGALPYGRPAGGREFTSDIEF-PRPPWRPGACDV	127
lacZ-α-048	MTMITPSYLGDTIEYSSYASNALGALPYGRPAGGRTSDIPR-PWRPGACDV	122
lacZ-α-068		90
lacZ-α-044 lacZ-α-002		116 86
1acz-α-002 1acz-α-003		145
lacZ-α-006		66
lacZ-α-063		168
lacZ-α-004	MTMITDPLESTCRHASLALAVVLQRRDWENPGVTQLNRLAAHPP	108
lacZ-α-045	MTMITDPLESTCRHASLALAVVLQRRDWENPGVTQLNRLAHPP	96
lacZ-α-039		100
lacZ-α-013		107
lacZ-α-034 lacZ-α-067		76 89
lacZ=α=054		100
lacZ-α-074		102
lacZ-α-060		77
lacZ-α-007	MTMITNLIRLTIGNLNLIKARPWIDVNMHDITRGATSASRSGRPTGGPYGPAAANSSSVPTYSSRDVRSNNWFKRPGTWDPLESTCRHASFGPPRDNRVISLALAVVLQRRDWENPGVTQLNRLAAHPP	167
lacZ-a-024		122
lacZ-α-018		182
lacZ-α-014		126
lacZ-α-026 lacZ-α-051		126
lacZ-α-046		116
lacz-α-043	MTMITPSYVIRLTSGQIFECIARTVRLEEFLQDIWIHEASHGDVTGSRY-LGELWYPLVKALSESYY-GLAVVLQRRDWENPGVTQLNRLAAHPP	123
lacZ-α-050	MTMITPSYVIRLTIGQIFECIARTVRLEEF	123
lacZ-α-056	MTMITPSYVIRLTSGARAIEAGWRQASLQDIWIHEFASFGRDASPD-VQACVDPLVKALSESYYGLAVVLQRRDWENPGVTQLNRLAAHPP	120
lacZ-α-058	MTMITPSYVIRLTIGARAIEAGWRQASGLAVVLQRRDWENPGVTQLRRLAAHPP	120
lacZ-α-053		122
lacZ-α-061		121
lacZ-α-020 lacZ-α-062+064		65 61
lacz-α-062+064		56
lacZ-α-073+076	MINITESTIGDTEYSSANALELKHKWOLQA-AALUPDTERPERSFKPMGAGFAL-	60

Key

s o quo i i i i	8
lacz-α-041	LSPAGVIAKPAPIALPNSCAA
lacZ-α-041 corr	FASWRNSEEARTDRPSOOLRSLNGEWREIVNVNVIVPLLOKCORYSLRRPKGY
lacZ-α-023	FASWRNSEEARTDRPSOOLRSINGEWREIVNVNVIVPLLOKCORYSIRRPKGY
lacZ-α-032	FASWRNSEEARTDRPSQQLRSLNGEWREIVNVNVIDT
lacZ-α-047	FASWRNSEEARTDRPSQQLRSLNGEWREIVNVNVNVTPQYILPRSLIPGIGNPEAREVAAFRAFSFSKKKIYKTICCGRPPGCGQRRWRSTVGNRLRLSTGGAGARSALSTSQGRPIIDIYIHN
lacZ-α-012	FASWRNSEEARTDRPSQQLRSLNGEWKL
lacZ-α-015	FASWRNSEEARTDRPSQQLRSLNGEW-DAPCSGALSAAGVVVTRSVTATLASALAPAPFAFFPSFLATFAGFPRQALNRGLPLGFRFSALRHLDPKKLD
lacZ-α-031	FASWRNSEEARTDRPSQQLRSLNGEWRDAPCSGALSAAGVVVTRSVTATLASALAPAPFAFFPSFLATFAGFPRQALNRGLPLGFRFSALRHLDPKKLD
lacZ-α-005	FASWRNSEEARTDRPSQQLRSLNGEW-DAPCSGALSAAGVVVTRSVTATLASALAPAPFAFFPSFLATFAGFPRQALNRGLPLGFRFSALRHLDPKKLD
lacZ-α-025	FASWRNSEEARTDRPSQQLRSLNGEW-DAPCSGALSAAGVVVTRSVTATLASALAPAPFAFFPSFLATFAGFPRQALNRGLPLGFRFSALRHLDPKKLD
lacZ-α-042	FASWRNSEEARTDRPSQQLRSLNGEWRDAPCSGALSAAGVVVTRSVTATLASALAPAPFAFFPSFLATFAGFPRQALNRGLPLGFRFSALRHLDPKKLD
lacZ-α-019	FASWRNSEEARTDRPSQQLRSLNGEW-DAPCSGALSAAGVVVTRSVTATLASALAPAPFAFFPSFLATFAGFPRQALNRGLPLGFRFSALRHLDPKKLD
lacZ-α-037	FASWRNSEEARTDRPSQQLRSLNGEWRDAPCSGALSAAGVVVTRSVTATLASALAPAPFAFFPSFLATFAGFPRQALNRGLPLGFRFSALRHLDPKKLD
lacZ-α-010	FASWRNSEEARTDRPSQQLRSLNGEW-DAPCSGALSAAGVVVTRSVTATLASALAPAPFAFFPSFLATFAGFPRQALNRGLPLGFRFSALRHLDPKKLD
lacZ-α-055	FASWRNSEEARTDRPSQQLRSLNGEWKL
lacZ-α-038	FASWRNSEEARTDRPSQQLRSL
lacZ-α-028	FASWRNSEEARTDRPSQQLRSLNGEWRLMRYFLLTHLCGISHRIWCTLSTICSDAA
lacZ-α-017	FASWRNSEEARTDRPSQQLRSLNGEWRSNF
lacZ-α-049	FASWRNSEEARTDRPSQQLRSLNGEWRDAPCSGALSAAGVVVTRSVTATLASALAPAPFAFFPSFLATFAGFPRQALNRGLPLGFRFSALRHLDPKKLD
lacZ-α-008	FASWRNSEEARTDRPSQQLRSLNGEWRLMRYFLLTHLCGISHRIWCTLSTICSDAA
lacZ-α-030	FASWRNSEEARTDRPSQQLRSLNGARSPRFLPFLSRHVRRLSPSSSKSGAPFRVPI
lacZ-α-065	FASWRNSEEARTDRPSQQLRSLNGEWRFAWFPAPEAVPESWLECDLPEADTVVVPSNWQMHGYDAPIYTNVTYPITVNPPFVPTENPTGCYSLTFNVDESWLQEGQTRIIFDGVPIG
lacZ-α-009	FASWRNSEEARTDRPSQQLRSLNGE
lacZ-α-057	FASWRNSEEARTDRPSQQLRSLNGEWRLMRYFLLTHLCGISHRIWCTLSTICSDAA
lacZ-α-036	FASWRNSEEARTDRPSQQLRSLNGEWRLMRYFLLTHLCGISHRI
lacZ-α-036corr/052	FASWRNSEEARTDRPSQQLRSLNGEWRLMRYFLLTHLCGISHRIYRWAILMKNILNLLSYY
lacZ-α-011	FASWRNSEEARTDRPSQQLRSLNGEWRLMRYFLLTHLCGISHRIRQSNHSTRPVAAH
lacZ-α-001	FASWRNSEBARTDRPSQQLRSLNGE DLFNKAAVPSSQRNALPVLQPINQF
lacZ-α-072	FASWRNSEEARTDRPSQQLRSLNGEW-DAPCSGALSAAGVVVTRSVTATLASALAPAPFAFFPSFLATFAGFPRQALNRGLPLGFRFSALRHLDPKKLD
lacZ-α-070	FASWRNSEEARTDRPSQQLRSLNGEWKL
lacZ-α-035	FASWRNSEEARTDRPSQQLRSLNGEWRLMRYFLLTHLCGISHRIWCTLSTICSDAA
lacZ-α-033	FASWRNSEEARTDRPSQQLRSLNGEWEIVNVNILLKFALNFC
lacZ-α-075	FASWRNSEEARTDRPSQQLRSLNGEWKL
lacZ-α-022	FASWRNSEEARTDRPSQQLRSLNGEWRLMRYFLLTHLCGISHRIWCTLSTICSDAA
lacZ-α-059	FASWRNSEEARTDRPSQQLRSLNGE
lacZ-α-016	FASWRNSEEARTDRPSQQLRSLFASWRNSEEARTDRPSQQLRSL
lacZ-α-029 lacZ-α-021	FASWRNSEBARTDRPSQURSLFASWRNSEBARTDRPSQURSL
lacz-α-021	FASWRISEARTURFSQUKSI.—
lacz-α-027 lacz-α-071	
lacz-α-0/1 lacz-α-040	FASWENSEBARTDRPSQLESLNGEW-DAPCSGALSAAGVVVTRSVTATLASALAPAFFAFFPSFLATFAGFPRQALNRGLPLGFRFRALRHLDRKKLDLGDGSRSGFSP
1acz-α-040 1acz-α-048	FASWKNSEARTURFSQQLKSLINGEWIKYVAAH
1acz-α-048 1acz-α-068	FASWKNSEARTURFSQQLKSLINGEWIKYVAAI
lacZ-α-044	FASWRASELARI DRIFSQULKS
1acz-α-044 1acz-α-002	FASWKRSEARTURFSQUKSLINGENS FREIFESKRIKESSSSSSSAFFKVI- FASWKRSEARTURFSQUKSLINGENEL- FASWKRSEARTURFSQUKSLINGENEL-
lacz-u-002	FASWRADERAN DAY SYULABURGANIL- FASWRADERAN DAY SYULABURGAN DAY SYU
lacz-u-005	FASWRISDEART DRF SQUIRSL
lacz-u-000	FASWRASEEARTIDESQUESLAGEWEFAWFPAPEAVPESWLECDLPBADTVVVPSNWQMHGYDAPIYTNVTYPITVNPPFVPTENPTGCYSLTFNVDESWLQBGQTRIIFDGVPIG
lacZ=α=003	FASWINSEEARTDPSOOLRSINGEWRIMTYFLITHLGGISHRIROSNHSTRPVAAH
lacZ-α-045	FASWINSEEARTDRPSQQLRSLNGEWRLMRYFLLTHLCGISHRIG
lacZ=α=039	FASWRNSEEARTDRPSQQLRSINGEWRLMRYFLLTHLCGISHRIDPVRR
lacZ=α=013	FASWRNSEEARTDRPSQQLRSLNGEWRLMRYFLLTHLCGISHRIWCTLSTICSDAA
lacZ-α-034	FASWRNSEEARTDRPSQQLRSLNGE
lacZ-α-067	FASWRNSEEARTDRPSQQLRSLNGEWRLRNHYYHDINL
lacZ-α-054	FASWRNSEEARTDRPSOOLRSLNGE DLFNKAAVPSSORNALPVLOPINOF
lacZ-α-074	FASWRNSEEARTDRPSOOLRSINGE DIFNKAAVPSSORNALPVLOPINOF
lacZ-α-060	FASWRNSEEARTDRPSQQLRSLNGE
lacZ-α-007	FASWRNSEEARTDRPSQQLRSLNGEWEIVNVNILLIFC
lacZ-α-024	FASWRNSEEARTDRPSQQLRSLNGEWTRPVAAH
lacZ-α-018	FASWRNSEEARTDRPSQQLRSLNGEWEIVNVNILLKFALNFC
lacZ-α-014	FASWENSEE ARTIDEPSOOLESINGEWKI.
lacZ-α-026	FASWRNSEEARTDRPSOOLRSINGEWKI
lacZ-α-051	FASWRNSEEARTDRPSQQLRSLNGEWKL
lacZ-α-046	FASWRNSEEARTDRPSQQLRSLNGEWKL
lacZ-α-043	FASWRNSEEARTDRPSQQLRSLNGEWRFAW
lacZ-α-050	FASWRNSEEARTDRPSQQLRSLNGEWRFAW
lacZ-α-056	FASWRNSEEARTDRPSQQLRSLNGEWRFAW
lacZ-α-058	FASWRNSEEARTDRPSQQLRSLNGEWRFAW
lacZ-α-053	FASWRNSEEARTDRPSOOLRSLNGEWRFAW
lacZ-α-061	FASWRNSEEARTDRPSQQLRSLNGEWRFAW
lacZ-α-020	FASWRNSEEARTDRPSQQLRSLNGEWR
lacZ-a-062+064	
lacZ-a-066+069	
lacZ-α-073+076	
	************

Name of	# of	Alignment
Variant	Occur	•
MBP-004	10	AAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGAT
MBP-010 MBP-001	10 7	AAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGAT
MBP-001 MBP-002	2	AAAAT COACOAAGGTAAACTGGTAATCTGGATTAACGGCGAT
MBP-009	2	
MBP-003	1	ATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGAT
MBP-005	1	AAA@GGAAGGTAAACTGGTAATCTGGATTAACGGCGAT
MBP-006	1	AAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGAT
MBP-007	1	ATGAAAATAAAACAGGTGCACGCATCCTCGCATTATCCGCATTAACGACGATGATGTTTTCCGCCTCGGCTCTCGCCAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGAT
MBP-008	1	AAAATCGAGAAGGTAAATCTGGATTAATCGGCGAT
MBP-004		A A A G G C TATA A C G G T C T G A G T A G A A A G T T C G G A A T A A A G T C C G T T A A C T G G A A G A T A C T G G C A A C T G G C A C T G C G C A C T G C G C A C T G C G C A C T G C G C C A C T G C C C C G C C C C C C C C C C C C
MBP-010		A A A G G C TATA A C G G T C T G A G T C G G T A A G A A T T C G G A A G A T T A A G T C C C T T G G C A T C T G G A G G A A T T C C C A C G G T T C C G C A C T G C C T G G C A C T G C C C C C C C C C C C C C C C C C
MBP-001		AAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGC
MBP-002		AAAGGCTATAACGGTCTGGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGC
MBP-009 MBP-003		AAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACAGCAAAAGGCTATAAACGGTCTCGCTGAAGTCACACAGGTTGCGGCAACTAGCCAACTGGCAACTGGCCAACTGGCCAACTGGCCAACTGGCCAACTGGCCAACTGGCCAACTGGCCAACTGGCCAACTGGCAACTGGCAACTGGCCAACTGGCAACTGGCAACTGGCAACTGGCAACTGGCCAACTGGCAACTGAACTGAACTGAACTGAACTGGAAAACTGAAAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAAAACTGAACTGAAAAAACAACTAAAACTGAACTAAAAACTGAACTGAAAAAAACAACTAAAAAAAA
MBP-005		AAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGGAATTCGGGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCACGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGG
MBP-006		AAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAAAATTCCCACAGGTTGCGGCAACTGGC
MBP-007		A A A G G C TATA A C G G T C T C G C T G A G T A G A A A T T C C G G A A T T A A G T C C C T T G G C A T C T G G A G G A A T T C C C A C G T T C C C C C T G G C T C C C C C C
MBP-008		AAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAAATTCCCACAGGTTGCGGCAACTGGC
		***************************************
MBP-004		${\tt GATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGGTTACCTGGCTGAAATCACCCCGGACAAAGCTGAATCACCCCGGACAAACCTGTACGCTCAATCAGCTCAGACAATCACCCCGGACAAACCAAACCTGTACACCCCGGACAAACCAAACCAAACCAGACAAACCAAACCAAACCAAACCAAACCAAACCAAACCAAACCAAACCAAACCAAACCAAACAAACAAACAAACAAAA$
MBP-004 MBP-010		GATGGCCCTGGACATTATCTTCTGGGCACACGACGCCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCGTTTACCTG
MBP-001		GATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGGAATCACCCCGGACAAAGCGTTCCAGGACAAGCGTTTACCTGTTTACCTGG
MBP-002		GATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGG
MBP-009		${\tt GATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGG}$
MBP-003		GATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGG
MBP-005		GATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGG
MBP-006 MBP-007		GATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGG <mark>CAGC</mark> AGCGTTCCAGGACAAGCTGTATCCGTTTACCTGG GATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACATCTGGCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGG
MBP-007		GATGGCCCTGACATTATCTTCTGGGCACACGACGCCTTTGGTGGCTACGCTCAATCTGGCTGTGGCTGGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCGTTTACCTG
1151 000		***************************************
MBP-004		GATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTTGGGAAGAGATCCCGGCG
MBP-010 MBP-001		GATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCG GATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCG
MBP-001		GATGCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATGCGTGTTGAAGCGTTATTGCGTGATTTATAACAAAGATCTGCTGCCGAACCGGCAAAAACCTGGGAAGAGAGAG
MBP-009		GATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGCAACCCGCCAAAAACCTGGGAAGAGATCCCGGCG
MBP-003		${\tt GATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCAAAAACCTGGGAAGAGATCCCGGCGCAAAAACCTGGGAAGAGATCCCGGCGAACCAAAAACCTGGGAAGAGATCCCGGCGAACAAAAACCTGGGAAGAGATCCCGGCGAACAAAAAACCTGGGAAGAGAAAAACCTGGGAAGAGATCCCGGCGAAAAAAAA$
MBP-005		GATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGCTCCCGGCG
MBP-006		GATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCG
MBP-007		GATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATTATATACGAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAAGAGCTCCGGGG
MBP-008		GATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGGTCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCG
MBP-004		$\tt CTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGC$
MBP-010		$\tt CTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGC$
MBP-001		CTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCCGTTCAAGTATGAAAACGGC
MBP-002 MBP-009		CTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGC CTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCCTTCAAGTATGAAAACGGC
MBP-003		CTGGATAAAGAACTGAAAGCGAAAGCGTAAGGACGCCCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCCCTGATTGCTGCTGCAGGAGGTTATGCGTTCAAGTATCAAGTATGAAAAACGGC
MBP-005		CTGGATAAAGAACCGAAAGCGAAAGGTAAGACGCCCTGAACTTCAACCTGCAACAACCGTTCACCTGGCCCCTGATTGCTGCTGACGGGGGTTATTGCATTCAAAAACGGC
MBP-006		CTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATG <mark>E</mark> A <mark>GGA</mark> GGC
MBP-007		$\tt CTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATG\overline{A}A\overline{A}\overline{A}\overline{A}\overline{C}GGC$
MBP-008		$\tt CTGGATAAAGGAAGCGAAAGGTAAGGCGGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCCGTGATGGATG$
		***
MBP-004		AAGTACGACATTAAAGACGTGGGCGTGATAACGCTGGCGGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAAC
MBP-010		AAGTACGACATTAAAAGACGTGGGCGTGGATAACGCTGGCGCAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAAC
MBP-001		${\tt AAGTACGACATTAAAGACGTGGGCTGGATAACGCTGGCGCAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAAC$
MBP-002		AAGTACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAAC
MBP-009		AAGTACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAAC
MBP-003		AAGTACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTTCCTGGTTGACCTGATTAAAAACAAAC
MBP-005 MBP-006		AAGTACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAAC
MBP-007		AAGTACGACATTAAAAGACCTGGGCGTGGATAAACGCTGGCCGCAAAGCGGGTCTGACCTTCCTGGTTGACCTAATTAAAAACAAAC
MBP-008		${\tt AAGTACGACATTAAAGACGTGGGCGTGATAACGCTGGCGCAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAAC$
		***************************************

# Var

bp 3d 1/3d 4/3d  Sources\*

MBP-004 MBP-010 MBP-001 MBP-002 MBP-009 MBP-003 MBP-005 MBP-006 MBP-007 MBP-008	GCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA GCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA GCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA GCTGCCTTTAATAAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACCACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA GCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACCACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA GCTGCCTTTAATAAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA GCTGCCTTTAATAAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA GCTGCCTTTAATAAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA GCTGCCTTTAATAAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA GCTGCCTTTAATAAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA GCTGCCTTTAATAAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA GCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA CCTGCCTTTAATAAAGGCGAACACGGATGACCATCAACGGCCCGTGGGCATGGCCAACATCGACACCAAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCA ********************************
MBP-004	${\tt TCCAAACCGTTCGTTGGCGTGGCGGAGGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGA$
MBP-010	${\tt TCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTTGCCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGA$
MBP-001	${\tt TCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGA$
MBP-002	TCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGA
MBP-009	TCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGA
MBP-003	TCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAAAAGAGTTCCTCGAAAACTATCTGCTGACTGA
MBP-005 MBP-006	TCCAAACCGTTCGTTGGCGGCGCGGCGCAGGTATTAACGCCCGCC
MBP-007	TCCAAACCGTTCGTTGGCGTCCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGCCAACAAAGAGTTCCTCGAAAACTATCTGCTGACTGA
MBP-008	TCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCEAAAGAGTTCCTCGAAAAACTATCTGCTGACGATGAAGGTCTGGAAGCGGTTAAT
	***************************************
MBP-004	${\tt AAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACCATGGAAAAGCGCCCAGAAAGGTGAAATCATGCCGAACATCCACCATGCAACATCCCACCATGCAACATCCCACCATGCAACATCCACCATCCACCATCCACCATCCACCATCCACCA$
MBP-010	AAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACCATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCG
MBP-001	AAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCAC <mark>T</mark> ATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCG
MBP-002	AAAGACAAACCGCTGGGTGCCGTAAGGTCTTAACAGGAAGAGTTGGGGAAAGATCCACGTATTGCCGCAC <mark>T</mark> ATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCG
MBP-009 MBP-003	AAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCAC <mark>L</mark> ATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCG AAAGACAAACCGCTGGGTGCCGGTAGCGCTGAAGTCTTACGAGGAAAGGTTGGCGAAAGATCCACGTATTGCCGCCACLATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCG
MBP-005	ANAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCACCACTGGAAAACGCCCAGAAAGGTCAAATCATCCCGACACATCCAGAAGGTGAAATCATCCCGAACATCAAAACCGCCAGAAAACGTCAAATCATCCCGACACATCCAGAAAACGCCCAGAAAGGTGAAATCATCCCGACACATCCAGAAAACGTCAAAACGTCAAAATCATCCCGACACATCAAAACGTCAAAAACGTCAAAATCATCCCGACACATCAAAAACGTCAAAAACGTCAAAATCATCCCGACACATCAAAAACGTCAAAAACGTCAAAATCATCCCGAAAAACGTCAAAATCATCCCGAAAAACGTCAAAATCATCCCGAAAAACGTCAAAATCATCCCGAAAAACGTCAAAATCATCCCGAAAAACGTCAAAATCATCCCGAAAAACGTCAAAATCATCCCGAAAAACGTCAAAATCATCCCGAAAAACGTCAAAATCATCCCGAAAAACGTCAAAATCATCCCGAAAATCATCCCGAAAAACGTCAAAATCATCCCGAAAATCATCCCGAAAAACGTCAAAAACATCCAAAAACATCAAAAAAAA
MBP-006	AAAGCAAACCGCTGGGTGCCCTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAACATCCACGTATTGCCGCACCATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCG
MBP-007	AAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGG GAAAGATCC-CGTATTGCCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCG
MBP-008	AAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAAGATCCACGTATTGCCGCCACCATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCG
	***************************************
MBP-004	${\tt CAGATGTCCGCTTTCTGGTATGCCGTGCTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACT}$
MBP-010	CAGATGTCCGCTTTCTGGTATGCCGTGCGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCCTGAAAGACGCCGCAGACT
MBP-001 MBP-002	CAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCGGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACT CAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACT
MBP-002 MBP-009	CAGATOTICCCTTTCTGGTATGCCGTGCTACTGCGGTGATCAACGCCGCAGCGGTCGTCAGACTGTCGATGAACGACCCTGAAAGACGCCGCAGACT CAGATOTCCGCTTTCTGGTATGCCGTGCTACTGCGGTGATCAACCCGCCAGCGGTCGTCAGACTGTCGATGAACCCCTGAAAGACGCCGCAGACT CAGATOTCCGCTTTCTGGTATGCCGTGATCATGCGGTGATCAACCCGCCAGCGGTCGTCAGACTGTCGATGAACCCTGAAAGACGCGCAGACT
MBP-003	CAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACCCCTGAAAGACCCCTGAAAGACTGCACTGACTG
MBP-005	CAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACT
MBP-006	CAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGAT
MBP-007	${\tt CAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACT}$
MBP-008	CAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGAT
	********
	al to MBP-004 but missing START codon for C-terminal in-frame protein tagging al to MBP-003 but missing START codon for C-terminal in-frame protein tagging
Key ATGC = c	hange in nucleotide relative to consensus sequence — synonymous
	nange in nucleotide relative to consensus sequence — synonymous hange in nucleotide relative to consensus sequence — conservative/semi-conservative amino acid substitution
ATGC = c	hange in nucleotide relative to consensus sequence — non-conservative amino acid substitution

Name of Variant	Alignment of Translated Protein
MBP-004	MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPOVAATGDGPDIIFWAHDRFGGYAOSGLLAEITPDKAFODKLYPFTW
MBP-010	KIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPOVAATGDGPDIIFWAHDRFGGYAOSGLLAEITPDKAFODKLYPFTW
MBP-001	KIBEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPOVAATGDGPDIIFWAHDRFGGYAOSGLLAEITPDKAFODKLYPFTW
MBP-002	MGKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPOVAATGDGPDIIFWAHDRFGGYAOSGLLAEITPDKAFODKLYPFTW
MBP-009	MEAEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPOVAATGDGPDIIFWAHDRFGGYAOSGLLAEITPDKAFODKLYPFTW
MBP-003	MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFFOVAATGDGPDIIFWAHDRFGGYAOSGLLAEITPDKAFODKLYPFTW
MBP-005	KBEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPOVAATGDGPDIIFWAHDRFGGYAOSGLLAEITPDKAFODKLYPFTW
MBP-006	KIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPOVAATGDGPDIIFWAHDRFGGYAOSGLLAEITPAAAFODKLYPFTW
MBP-007	MKIKTGARILALSALTTMMFSASALAKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPOVAATGDGPDIIFWAHDRFGGYAOSGLLAEITPDKAFODKLYPFTW
MBP-008	KIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTW
	***************************************
MBP-004	${\tt DAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAE$
MBP-010	${\tt DAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAE$
MBP-001	${\tt DAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAE$
MBP-002	${\tt DAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAE$
MBP-009	${\tt DAVRYNGKLIAYPIAVEALSLIYNKOLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAE\\$
MBP-003	DAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAE
MBP-005	DAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAE
MBP-006	DAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKY <mark>AL</mark> GKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAE
MBP-007 MBP-008	DAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAE DAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAE
MBP-006	DAVKINGALIATPIAVEALSLIINADLLENPPKINSEIPALDAELAANOASALMENLYDPITIWPLIAADGIIAFNIENGATDIADVGVDNAGAAAGLITLVULINNAMMADIDISIAE ************************************
MBP-004 MBP-010	AAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQT AAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQT
MBP-010	AAFNKGETAMTINGEWAWSNIDTSKVNIGGTULPTEKGGYLEAVNKDKPLGAVALKSYEEELAKDFRIAATMENAGKGEIMPRIPUMSAFWIAVKTAVINAASGRGTUDEALKDAGT AAFNKGETAMTINGPKAWSNIDTSKVNIGGTULPTEKGOPLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAOKGEIMPRIPOMSAFWYAVRTAVINAASGRGTUDDEALKDAOT AAFNKGETAMTINGPKAWSNIDTSKVNIGGTULPTEKGOPLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAOKGEIMPRIPOMSAFWYAVRTAVINAASGRGTUDDEALKDAOT
MBP-001	AAF NACETAMII NGEWAWSNIDISAV NICUTUPIT ROGY LEAVNKDKPLGAVALKSIEBELAADFRIAATHENAQAGEIHMII YUMSAF WIAVKIAV INAASGRUVDEALKDAQI AAFNKGETAMII NGPWAWSNIDISKVNYGUTULPTFKGOP LEAVNKDKPLGAVALKSYEBELAKDPRIAATHENAOKGEIHMI POMSAFWAVRTAVINAASGRUTUDEALKDAOT
MBP-002 MBP-009	aaf nacetamiingwawsnidisavnicvituffiacyLeavnkdreigavalasiebelaadriaatenagagelimtiiqusafwiavatavinaasgrivdealadagi AafnacetamiingwawsnidiskunggvitulptikgopLeavnkdkelgavalasyebelakdriaatmenagagelimtiiqusafwavatavinaasgrotvobalkdagt
MBP-003	AAPINGETAMTINGWAMONIDISKVINGVITUPTEKGOLEAVINDKPLGAVALKSISEELLAKDPRIAATHENAGKGEIMPNIPGMSHWIAVAITAVINAASGGGTUDEALKAAQI AAPINGETAMTINGPWAWONIDISKVINGGTULPTEKGOLEAVINDKPLGAVALKSISEELLAKDPRIAATHENAGKGEIMPNIPGMSAFWAVATRAVINAASGGGTUDEALKAAQT
MBP-005	AAPINGETAMTINGWAMONIDISKVNYGYTULPITKGGFLEAVNKDKPLGAVALKSISEELEKDPRIAATHENAGKGEIMVAITGUSAFWIAVAITAVINAASGGGTUDEALKDAGT AAPINGETAMTINGGWAWONIDISKVNYGYTULPITKGGFLEAVNKDKPLGAVALKSISEELEKDPRIAATHENAGKGEIMVNIPOMSAFWAVATRAVINAASGGGTUDEALKDAGT
MBP-006	AAPNKEETAMTINGPWAWSNIDTSAVNYGYTULPTFKGOPLEAVNKDKPLGAVALKSYEEELAKDPRIAATHENAOKGEIHPNIPOMSAFWYAVRTAVINAASGROTVD
MBP-007	AAPNKGETAMTINGPWAWSNIDTSKVNYGYTVLPTFKGODLEAVNKDKPLGAVALKSYEEEL KDPRIAATMENAOKGEIMPNIPOMSAFWYAVRTAVINAASGROTVDEALKDAOT
MBP-008	AAPNKGETAMTINGPWAWSNIDTSKVNYGGTVLPTFKGODLEAVNKDKPLGAVALKSYEEELAKDPRIAATHENQOKGEIMPNIPOMSAFWYAVRTAVINAASGROTVD
IIDI -000	***************************************
Key	AWCC = change in nucleotide relative to consensus sequence — synonymous
,	Aucc = change in nucleotide relative to consensus sequence = Synonymous and substitution
	Argo = change in nucleotide relative to consensus sequence — non-conservative amino acid substitution

Name of Variant	# of Occur	Alignment			Size (bp)	# Var bp	Sources*
GST-002	44	ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTT.GCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGA	CCATCCTCCAA	A	654	0	GE, No, BD
GST-004	3	${\tt ATG}^{\tt GCA}_{\tt TCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTT.}_{\tt GCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGAAGCCACGTTTGGTGGTGGTGGCGAAGCCACGTTTGGTGGTGGTGGCGAAGCCACGTTTGGTGGTGGTGGCGAAGCCACGTTTGGTGGTGGTGGCGAAGCCACGTTTGGTGGTGGTGGCGAAGCCACGTTTGGTGGTGGTGGCGAAGCCACGTTTGGTGGTGGTGGCGAAGCCACGTTTGGTGGTGGTGGCGAAGCCACGTTTGGTGGTGGTGGCGAAGCCACGTTTGGTGGTGGTGGCGAAGCCACGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG$	CCATCCTCCAA	A	657	0/3i	No
GST-005	2	TCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTT.GCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGA	CCATCCTCCAA	A	651	0/3d	19906724
GST-001	2	ATGGCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTT.GCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGA	CCATCCTCCAA	A	654	1	No
GST-003	1	ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTT.GCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGA			657	0/3i	#
GST-006	1	ATGACCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTT,GCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGAGCCACGTTTGGTGGTGGTGGCGAGCCACGTTTGGTGGTGGTGGCGAGCCACGTTTGGTGGTGGTGGCGAGCCACGTTTGGTGGTGGTGGCGAGCCACGTTTGGTGGTGGTGGCGAGCCACGTTTGGTGGTGGTGGCGAGCCACGTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGTGGCGAGCCACGTTTGCAGGCCTTTGCAGGGCTTGCAGGCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCAGGCCAGGCCAGGCAGGCCAGGCAGGCAGGCCAGGAGG	CCATCCTCCAA	A	654	1	AB
GST-007	1	CCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTT.GCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGA			648	6	15232106
GST-008	1	ATGCANTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTT.GCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGA			657	3	15782178
	_	*** ***********************************	******	*		-	
Name of		Alignment of Translated Protein	Size	Tag	# Var		
Variant			(aa)	Pos	aa		
GST-002		-MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYI.KKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK	218	N/C1	0		
GST-004		M <mark>E</mark> SPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYI.KKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK	219	N	1i		
GST-005		SPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYI.KKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK	217	$N^2$	1d		
GST-001		-Mapilgywkikglvoptrllleyleekyeehlyerdegdkwrnkkfelglefpnlpyyi.kkrieaipoidkylksskyiawplogwoatfgggdhppk	218	N	1		
GST-003		-MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYI.KKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK	218	#	0		
GST-006		-MSPILGYWKIKGLVOPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYI.KKRIEAIPOIDKYLKSSKYIAWPLOGWQATFGGGDHPPK	218	N	0		
GST-007		PILGYWKIKGLVOPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYI.KKRIEAIPOIDKYLKSSKYIAWPLOGWQATFGGGDHPPK	216	С	2d		
GST-008		MESPILGYWKIKGLVOPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYI.KKRIEAIPOIDKYLKSSKYIAWPLOGWOATFGGGDHPPK	219	N	1i		
		*****************************					

- # from feature file. Not found in any vector in the plasmid library.
  1 followed by STOP codon after additional in-frame residues or followed by another tag.
- 2 preceded either by another tag or a START codon and additional in-frame residues.

GST-005 is identical to GST-002 but missing START codon for C-terminal in-frame protein tagging

Name of Variant	#of Occur	Alignment	Size (bp)	-	Var bp	Express Host	CAIHost/Non-Host <sup>1</sup>	Sources*
HA-008	10	TACCCATACGATGTTCCAGATTACGCT	27		5	Mammal-Plant	0.783/0.508	Or,Cl
HA-007	5	TATCCGTACGACGTACCAGACTACGCA	27		3	Mammal	0.463/0.510	C1
HA-003	4	TA <mark>T</mark> CCGTA <mark>T</mark> GA <mark>T</mark> GTGCC <mark>G</mark> GACTACGC <mark>A</mark>	27		5	E. coli	0.424/0.571	MCSG
HA-009	3	TA <mark>T</mark> CC <mark>T</mark> TA <mark>T</mark> GACGTGCC <mark>T</mark> GACTA <mark>T</mark> GCC	27		5	Mammal-S. cer.	0.941-0.071/0.155	Cl
HA-002	2	TATCCGTACGACGTACCGGACTACGCA	27		4	E. coli	0.642/0.395	MCSG
HA-004	2	TACCCATACGATGTTCCGGATTACGCT	27		6	Mammal	0.668/0.640	Mo
HA-006	2	TACCCATACGACGTACCAGATTACGCT	27		4	S. cer.	0.182/0.478	Cl
HA-001	1	TA <mark>T</mark> CCGTA <mark>T</mark> GACGTGCC <mark>C</mark> GACTA <mark>T</mark> GCC	27		4	Mammal	0.784/0.140	Lu
HA-005	1	TACCC <mark>TTAT</mark> GA <mark>T</mark> GTGCCAGA <mark>T</mark> TA <mark>T</mark> GCC	27		5	S. cer.	0.108/0.153	8242750
HA-010	1	TACCCCTACGACGTGCC GACTACGCC	27		1	Mammal	1.000/0.096	15232106
		** ** ** ** ** ** ** **						

1 Non-Host is E.coli for Mammal, Plant and S. cerevisiae. Non-Host is Mammal (Mouse) for E. coli. Carbone et al. (2003) reference set and Eyre-Walker (1996) equation.

Key

ATGC = change in nucleotide relative to consensus sequence — synonymous

ATTGC = change in nucleotide relative to consensus sequence - synonymous

ATTGC = change in nucleotide relative to consensus sequence - conservative/semi-conservative amino acid substitution

ATTGC = change in nucleotide relative to consensus sequence - non-conservative amino acid substitution

c = change in nucleotide relative to consensus sequence — non-conservative amino acid substitution