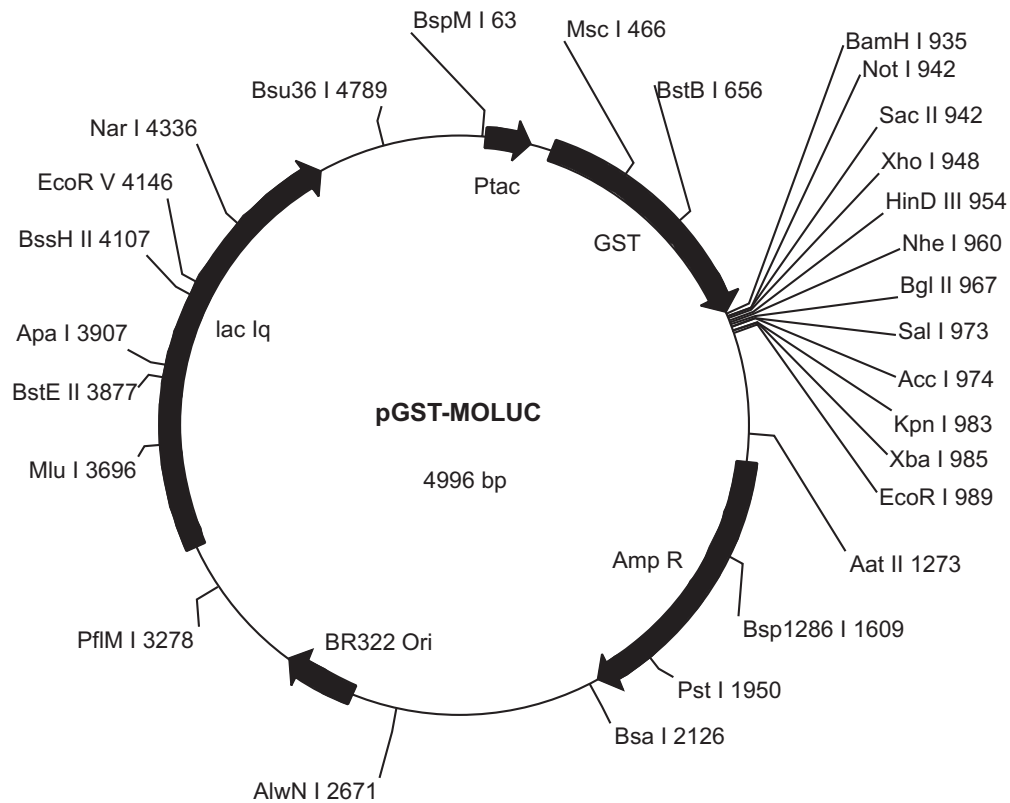


Name of Vector: pGST-MOLUC
Antibiotic Selection: Amp
Created by: Tao Feng
(He Lab @ The University of Chicago)
Date of Creation: September 2001



Linker Sequence

```

901                               921                               941
|                               |                               |
ATCCTCCAAAATCGGATCTGATCGAAGGTCGTGGGATCCGCGGCCGCTCGAGAAGCTTGC
H P P K S D L I E G R G I R G R S R S L

961                               981                               1001
|                               |                               |
TAGCGAGATCTGTCGACGGTACCTCTAGAATTCATCGTGACTGACTGACGATCTGCCTCG
L A R S V D G T S R I H R D -

```

pGST-MOLUC Sequence

AGCTTATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGAAGCTGTG 60
GTATGGCTGTGCAGGTCGTAAATCACTGCATAATTCGTGTCGCTCAAGGCGCACTCCCGT 120
TCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCCTGGCAAATATTCGTGAAATGAGC 180
TGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTTGTGAGCGGATAACAATTTCA 240
CACAGGAAACAGTATTCATGTCCCTATACTAGGTTATTGGAAAAATTAAGGGCCTTGTGC 300
AACCCACTCGACTTCTTTTTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGC 360
GCGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCAATC 420
TTCTTATTTATATTTGATGGTGATGTTAAATTAACACAGTCTATGGCCATCATAACGTTATA 480
TAGCTGACAAGCACAAACATGTTGGGTGGTGTGCCAAAAGAGCGTGCAGAGATTTCAATGC 540
TTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTCGAGAATTGCATATAGTAAAGACT 600
TTGAAACTCTCAAAGTTGATTTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTTCGAAG 660
ATCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGT 720
TGTATGACGCTCTTGATGTTGTTTTATAACATGGACCCAATGTGCCTGGATGCGTTCCCAA 780
AATTAGTTTGTFTTTAAAAAACGTATTGAAGCTATCCACAAAATTGATAAGTACTTGAAAAT 840
CCAGCAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACC 900
ATCCTCCAAAATCGGATCTGATCGAAGGTCGTGGGATCCgcggccgctcgagaagcttgc 960
tagcgagatctgtcgacggtacctctagAATTCATCGTACTGACTGACGATCTGCCTCG 1020
CGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCCACAG 1080
CTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTG 1140
GCGGGTGTGCGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATAATTCCTT 1200
GAAGACGAAAGGGCCTCGTGATACGCCATTTTTTATAGGTTAATGTCATGATAATAATGG 1260
TTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTAT 1320
TTTTCTAAAATACATTCAAAATATGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTC 1380
AATAATATTGAAAAAGGAAGAGTATGAGTATTC AACATTTCCGTGTCGCCCTTATTCCCT 1440
TTTTTGCGGCATTTTGCCTTCTGTTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAG 1500
ATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTA 1560
AGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTTCCAATGATGAGCACTTTTTAAAGTTC 1620
TGCTATGTGGCGCGGTATTATCCCGTGTGACGCCGGGCAAGAGCAACTCGGTGCGCCGCA 1680
TACACTATTTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGG 1740
ATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGG 1800

CCAAC TTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACA 1860
TGGGGGATCATGTA ACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAA 1920
ACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAA ACTATTTAA 1980
CTGGCGAACTACTTACTCTAGCTTCCC GGCAACAATTAATAGACTGGATGGAGGCGGATA 2040
AAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAAT 2100
CTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGC 2160
CCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATA 2220
GACAGATCGCTGAGATAGGTGCCCTCACTGATTAAGCATTTGGTAACTGTCAGACCAAGTTT 2280
ACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTTAATTTAAAAAGGATCTAGGTGA 2340
AGATCCTTTTTTGATAATCTCATGACCAAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAG 2400
CGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAA 2460
TCTGCTGCTTGCAAAACAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAG 2520
AGCTACCAACTCTTTTTCCGAAGGTA ACTGGCTTCAGCAGAGCGCAGATACCAAATACTG 2580
TCCTTCTAGTGTAGCCGTAGTTAGGCCACC ACTTCAAGAACTCTGTAGCACCGCTACAT 2640
ACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTA 2700
CCGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGTCCGGCTGAACGGGGG 2760
GTTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGC 2820
GTGAGCTATGAGAAAAGCGCCACGCTTCCC GAAGGGAGAAAAGGCGGACAGGTATCCGGTAA 2880
GCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAAACGCTGGTATC 2940
TTTTATAGTCTGTGCGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTTGTGATGCTCGT 3000
CAGGGGGGCGGAGCCTATGAAAAACGCCAGCAACCGGCCCTTTTTACGGTTCTTGCCCT 3060
TTTGCTGGCCTTTTGCTCACATGTTCTTTCC TGCGTTATCCCCTGATTCTGTGGATAACC 3120
GTATTACCGCTTTGAGTGAGCTGATACCGCTCGCCCGACCCGAACGACCGAGCGCAGCG 3180
AGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGT 3240
GCGGTATTTACACCCGCATAAAATTCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTA 3300
TGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAAACCAGTAACGT 3360
TATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCC GCGTGGTGAACC 3420
AGGCCAGCCACGTTTTCTGCGAAAAACGCGGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGA 3480
ATTACATTTCCCAACCGCTGGCACAACA ACTGGCGGGCAAACAGTCGTTGCTGATTGGCG 3540
TTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTTGTCGCGGCGATTAAATCTC 3600
GCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTGCGATGGTAGAACGAAGCGGCGTCGAAG 3660

CCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACT 3720
ATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGACTAATGTTCCGGCGT 3780
TATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTCTCCCATGAAGACG 3840
GTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAG 3900
CGGGCCCATTAAGTTCTGTCTCGGCGCTCTGCGTCTGGCTGGCTGGCATAAAATATCTCA 3960
CTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTT 4020
TTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCCTGCGATGCTGGTTGCCA 4080
ACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTG 4140
CGGATATCTCGGTAGTGGGATACGACGATAACCGAAGACAGCTCATGTTATATCCCGCCGT 4200
TAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGC 4260
AACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAA 4320
GAAAAACCACCCTGGCGCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCAT 4380
TAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAAGCGGGCAGTGAGCGCAACGCAATT 4440
AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGT 4500
ATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGAT 4560
TACGGATTCACTGGCCGTCGTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCA 4620
ACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCG 4680
CACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTT 4740
TCCGGCACCAGAAGCGGTGCCGAAAAGCTGGCTGGAGTGCATCTTCTGAGGCCGATAC 4800
TGTCGTGTCCTCCCTCAAACCTGGCAGATGCACGGTTACGATGCGCCATCTACACCAACGT 4860
AACCTATCCCATTACGGTCAATCCGCCGTTTGTTCACGAGAAATCCGACGGGTTGTTA 4920
CTCGCTCACATTTAATGTTGATGAAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTT 4980
TGATGGCGTTGGAATT 4996

pGST-MOLUC Restriction Sites

Unique enzymes in pGST-MOLUC:

BspM I	ACCTGC 10/14	63	Ban I	(7)	979	2239	3616	4335
Msc I	TGG CCA	466	Ban II	(1)	4465	4745	4757	
BsiC I	TT`CG,AA	656	Bbe I	(1)	3907			
BstB I	TT`CG,AA	656	Bbs I	(3)	1209	3842	4181	
BamH I	G`GATC,C	935	Bbv I	(11)	55	1071	1168	1960
Sac II	CC,GC`GG	942			2149	2751	3170	3188
Not I	GC`GGCC,GC	942	Bbv II	(3)	1208	3841	4180	
Eag I	C`GGCC,G	942	Bcl I	(2)	693	3710		
Xma III	C`GGCC,G	942	Bcn I	(7)	1067	1102	1657	2008
Ava I	C`YCGR,G	948			2704	3317	4126	
PaeR7 I	C`TCGA,G	948	Bfa I	(6)	271	961	986	1999
Xho I	C`TCGA,G	948			2334	2587		
HinD III	A`AGCT,T	954	Bgl I	(2)	2074	4717		
Nhe I	G`CTAG,C	960	Bgl II	(1)	967			
Bgl II	A`GATC,T	967	Bpm I	(4)	2123	3534	4023	4794
Sal I	G`TCGA,C	973	Bsa I	(1)	2126			
Acc I	GT`MK,AC	974	BsaA I	(1)	1174			
Acc65 I	G`GTAC,C	979	BsaH I	(5)	36	1270	1652	3653
Asp718	G`GTAC,C	979			4336			
Kpn I	G,GTAC`C	983	BsaJ I	(6)	939	2920	4331	4470
Xba I	T`CTAG,A	985			4607	4897		
EcoR I	G`AATT,C	989	BsaW I	(4)	1896	2727	2874	4015
BsaA I	YAC GTR	1174	Bsg I	(5)	-3	91	545	3547
Aat II	G,ACGT`C	1273			3747			
Bsp1286 I	G,DGCH`C	1609	BsiC I	(1)	656			
Pst I	C,TGCA`G	1950	BsiE I	(6)	945	1674	1823	2746
Bsa I	GGTCTC 7/11	2126			3170	4688		
Ahd I	GACNN,N`NNGTC	2192	BsiHKA I	(5)	23	1524	1609	2770
AlwN I	CAG,NNN`CTG	2671			3680			
PflM I	CCAN,NNN`NTGG	3278	BsmA I	(7)	1064	1351	2125	3392
Mlu I	A`CGCG,T	3696			3797	3923	4310	
BstE II	G`GTNAC,C	3877	BsmB I	(2)	1063	4311		
Bsp120 I	G`GGCC,C	3903	BsmF I	(2)	246	4794		
Apa I	G,GGCC`C	3907	BsoF I	(30)	44	942	945	1060
Ban II	G,RCYC`C	3907			1157	1448	1677	1772
BssH II	G`CGCG,C	4107			1799	1949	2138	2466
EcoR V	GAT ATC	4146			2672	2675	2740	2883
Hpa I	GTT AAC	4202			3038	3156	3159	3177
Kas I	G`GCGC,C	4335			3464	3587	3651	3671
Nar I	GG`CG,CC	4336			3758	4129	4258	4387
Ehe I	GGC GCC	4337			4636	4709		
Bbe I	G,GCGC`C	4339	Bsp120 I	(1)	3903			
Bsu36 I	CC`TNA,GG	4789	Bsp1286 I	(1)	1609			
Number of enzymes = 43			BspH I	(3)	1247	1352	2360	
			BspM I	(1)	63			
			Bsr I	(20)	1168	1546	1718	1985
					2028	2146	2552	2666
					2679	3353	3514	3552
					3616	3853	4005	4315
					4414	4575	4601	4823
			BsrB I	(6)	226	550	947	1350
					3151	4521		
			BsrD I	(4)	1958	2132	3743	4109
			BssH II	(1)	4107			
			BssS I	(3)	1216	1523	2907	
			BstB I	(1)	656			
			BstE II	(1)	3877			
			BstN I	(11)	766	2921	2934	3055
					3421	3736	4276	4333
					4472	4609	4736	
			BstU I	(27)	362	941	1021	1023
					1126	1301	1633	2126
					2456	3037	3296	3410
					3447	3497	3568	3586
					3602	3691	3698	3846
					3892	3927	4109	4133
					4364	4366	4970	
			BstX I	(3)	3498	3627	3750	
			BstY I	(9)	915	935	967	1545
					1562	2330	2342	2428
					2439			
			Bsu36 I	(1)	4789			
			Cac8 I	(30)	524	631	874	878
					958	962	1946	2079
					2470	3030	3067	3153
					3426	3516	3566	3621
					3762	3903	3942	3946
					4109	4228	4255	4391
					4420	4655	4659	4679
					4770	4950		
			Cfr10 I	(2)	2107	3382		
			Csp6 I	(4)	831	980	1710	3842
			Dde I	(8)	625	1265	1691	2231
					2397	2806	4268	4789
			Dpn I	(25)	662	695	917	922
					937	969	1012	1511
					1547	1564	1822	1868
					1886	2227	2332	2344
					2422	2430	2441	2516
					3608	3712	4085	4687
					4783			
			DpnII	(25)	660	693	915	920
					935	967	1010	1509
					1545	1562	1820	1866
					1884	2225	2330	2342
					2420	2428	2439	2514
					3606	3710	4083	4685
					4781			
			Dra I	(5)	686	795	1614	2306

The following enzymes do not cut in pGST-MOLUC:

Afl II	Age I	Asc I	Avr II	Blp I
BsaB I	BseR I	BsiW I	Bsm I	BspM II
BsrG I	Bst1107 I	Cla I	Dra III	Eco47 III
Eco72 I	Esp I	Fse I	Mun I	Nae I
Nco I	Nde I	NgoM I	Nru I	Nsi I
Pme I	Pml I	PpuM I	PspA I	Rsr II
Sac I	Sfi I	Sma I	SnaB I	Spe I
Sph I	Spl I	Srf I	Stu I	Sty I

pGST-MOLUC: sites sorted by name:

Aat II	(1)	1273						
Acc I	(1)	974						
Acc65 I	(1)	979						
Aci I	(59)	226	550	939	941			
		945	1093	1132	1142			
		1184	1301	1348	1447			
		1556	1633	1677	1798			
		1844	2035	2126	2488			
		2497	2632	2742	2863			
		2882	3009	3037	3128			
		3149	3156	3199	3216			
		3242	3255	3296	3408			
		3447	3463	3472	3495			
		3514	3586	3650	3670			
		3673	3724	3901	3985			
		4141	4195	4251	4279			
		4352	4362	4418	4521			
		4679	4755	4884				
Afl III	(3)	497	3080	3696				
Aha II	(5)	36	1270	1652	3653			
		4336						
Ahd I	(1)	2192						
Alu I	(27)	3	56	180	484			
		633	811	956	1062			
		1081	1839	1902	2002			
		2523	2780	2826	2916			
		3142	3477	3757	4181			
		4296	4389	4453	4548			
		4657	4768	4948				
Alw I	(11)	922	931	942	1552			
		1558	1873	2337	2338			
		2435	2435	2521				
AlwN I	(1)	2671						
Apa I	(1)	3907						
ApaL I	(4)	19	1520	2766	3676			
Apo I	(3)	989	3261	3971				
Ase I	(4)	191	2017	4381	4440			
Asp718	(1)	979						
Ava I	(1)	948						
Ava II	(4)	753	1828	2050	4248			
BamH I	(1)	935						

Drd I	(2)	2325								2351	2424	3213	3330
Dsa I	(2)	1115	2978							3677	3846	4185	4682
Eae I	(5)	939	4897							4777			
Eag I	(1)	464	942	1799	4370			Mlu I	(1)	3696			
Ear I	(5)	4573						Mme I	(5)	2688	2872	3289	4931
Eco57 I	(2)	942						Mnl I	(19)	4971			
EcoN I	(2)	338	1393	3197	3314					913	992	1026	1056
EcoO109 I	(2)	4666								1224	1819	2025	2171
EcoR I	(1)	1525	2539							2252	2652	2903	2976
EcoR II	(11)	269	1831							3186	3557	4041	4364
EcoR V	(1)	291	1212					Msc I	(1)	4667	4784	4822	
Ehe I	(1)	989						Mse I	(23)	466			
Fnu4H I	(30)	764	2919	2932	3053					191	287	446	451
		3419	3734	4274	4331					685	794	1241	1613
		4470	4607	4734						1978	2017	2252	2305
		4146								2319	2324	2376	3593
		4337								3716	3910	4201	4381
		44	942	945	1060					4440	4624	4933	
		1157	1448	1677	1772			Msl I	(5)	353	1421	1780	3748
		1799	1949	2138	2466					4036			
		2672	2675	2740	2883			Msp I	(19)	1066	1100	1655	1897
		3038	3156	3159	3177					2007	2074	2108	2512
		3464	3587	3651	3671					2702	2728	2875	3316
		3758	4129	4258	4387					3383	3775	4016	4124
		4636	4709							4493	4743	4761	
Fok I	(10)	692	781	887	1108			MspAl I	(9)	941	1132	1556	2497
		1753	2040	2221	3742					2742	3726	4296	4389
		3751	4627							4657			
Fsp I	(2)	1969	4707					Nar I	(1)	4336			
Gdi II	(5)	941	943	1800	4369			Nci I	(7)	1066	1101	1656	2007
		4572								2703	3316	4125	
Gsu I	(4)	2122	3535	4022	4793			Nhe I	(1)	960			
Hae I	(7)	466	861	2606	3058			Nla III	(22)	261	501	700	720
Hae II	(6)	3069	3424	4963						753	860	1059	1164
Hae III	(21)	2840	3210	3315	4096					1251	1356	1749	1785
		4339	4730							1863	1873	2364	3084
		293	466	861	944					3308	3835	4014	4035
		1214	1801	2068	2148					4187	4559		
		2606	3040	3058	3069			Nla IV	(16)	755	937	981	1305
		3424	3559	3905	4274					1895	2106	2147	2241
		4372	4575	4677	4794					3013	3052	3618	3905
		4963								4337	4467	4747	4759
Hga I	(11)	26	735	1116	1659			Not I	(1)	942			
		2391	2969	3643	3688			Nsp7524 I	(3)	497	1055	3080	
		3917	3923	4975				NspB II	(9)	941	1132	1556	2497
HgiA I	(5)	23	1524	1609	2770					2742	3726	4296	4389
		3680								4657			
HgiE II	(3)	2491	3287	4746				NspH I	(3)	501	1059	3084	
Hha I	(34)	112	141	362	1023			PaeR7 I	(1)	948			
		1126	1156	1301	1633			Pal I	(21)	293	466	861	944
		1970	2063	2456	2565					1214	1801	2068	2148
		2739	2839	2906	3176					2606	3040	3058	3069
		3209	3314	3570	3604					3424	3559	3905	4274
		3693	3894	3927	4095					4372	4575	4677	4794
		4102	4109	4111	4133					4963			
		4338	4366	4431	4708			PflM I	(1)	3278			
		4729	4844					Ple I	(4)	2201	3189	3332	4128
HinC II	(4)	185	975	1650	4202			Psp1406 I	(3)	1590	1963	3358	
Hind II	(4)	185	975	1650	4202			Pst I	(1)	1950			
HinD III	(1)	954						Pvu I	(2)	1823	4688		
Hinf I	(9)	2193	2710	3106	3181			Pvu II	(3)	4296	4389	4657	
		3324	4120	4375	4566			Rsa I	(4)	832	981	1711	3843
		4904						Sac II	(1)	942			
HinI I	(5)	36	1270	1652	3653			Sal I	(1)	973			
		4336						Sap I	(2)	338	3197		
HinP I	(34)	110	139	360	1021			Sau3A I	(25)	660	693	915	920
		1124	1154	1299	1631					935	967	1010	1509
		1968	2061	2454	2563					1545	1562	1820	1866
		2737	2837	2904	3174					1884	2225	2330	2342
		3207	3312	3568	3602					2420	2428	2439	2514
		3691	3892	3925	4093					3606	3710	4083	4685
		4100	4107	4109	4131					4781			
		4336	4364	4429	4706			Sau96 I	(13)	291	753	1212	1828
		4727	4842							2050	2067	2146	3558
Hpa I	(1)	4202								3903	3904	4248	4272
Hpa II	(19)	1066	1100	1655	1897					4676			
		2007	2074	2108	2512			Sca I	(2)	832	1711		
		2702	2728	2875	3316			ScrF I	(18)	766	1066	1101	1656
		3383	3775	4016	4124					2007	2703	2921	2934
		4493	4743	4761						3055	3316	3421	3736
Hph I	(15)	381	450	702	1041					4125	4276	4333	4472
		1050	1466	1499	1707					4609	4736		
		2121	2348	3350	3425			Sec I	(6)	939	2920	4331	4470
		3872	4292	4325						4607	4897		
Kas I	(1)	4335						SfaN I	(14)	760	1087	1492	1739
Kpn I	(1)	983								1932	2984	3204	3242
Mae I	(6)	271	961	986	1999					3730	3869	4059	4060
		2334	2587							4817	4830		
Mae II	(12)	474	801	883	1173			Sfc I	(4)	1946	2624	2815	4953
		1270	1590	1963	2379			Ssp I	(2)	167	1387		
		3358	3431	4589	4858			Taq I	(12)	8	309	577	656
Mae III	(19)	699	997	1074	1169					923	949	974	1538
		1531	1719	1872	1930					2982	3275	3634	3656
		2261	2544	2660	2723			Tfi I	(4)	3106	4375	4566	4904
		3354	3877	4593	4613			Tsp45 I	(7)	997	1074	1169	1719
		4833	4859	4917						1930	3877	4593	
Mbo I	(25)	660	693	915	920			Tth111 I	(2)	1167	3522		
		935	967	1010	1509			Tth111 II	(7)	504	530	2458	2488
		1545	1562	1820	1866					2497	3534	4227	
		1884	2225	2330	2342			Vsp I	(4)	191	2017	4381	4440
		2420	2428	2439	2514			Xba I	(1)	985			
		3606	3710	4083	4685			Xcm I	(3)	3552	4068	4086	
		4781						Xho I	(1)	948			
Mbo II	(17)	342	354	413	669			Xho II	(9)	915	935	967	1545
		1213	1409	1518	1596					1562	2330	2342	2428

Xma III (1) 942
 Site usage in pGST-MOLUC:

Aat II	G,ACGT`C	1	Acc I	GT`MK,AC	1
Acc65 I	G`GTAC,C	1	Aci I	C`CG,C	59
Afl II	C`TTAA,G	-	Afl III	A`CRYG,T	3
Age I	A`CCGG,T	-	Aha II	GR`CG,YC	5
Ahd I	GACNN,N`NNGTC	1	Alu I	AG CT	27
Alw I	GGATC 8/9	11	AlwN I	CAG,NNN`CTG	1
Apa I	G,GGCC`C	1	Apal I	G`TGCA,C	4
Apo I	R`AATT,Y	3	Asc I	GG`CGCG,CC	-
Ase I	AT`TA,AT	4	Asp718	G`GTAC,C	1
Ava I	C`YCGR,G	1	Ava II	G`GWC,C	4
Avr II	C`CTAG,G	-	BamH I	G`GATC,C	1
Ban I	G`GYRC,C	7	Ban II	G,RCY`C	1
Bbe I	G,CGC`C	1	Bbs I	GAAGAC 8/12	3
Bbv I	GCAGC 13/17	11	Bbv II	GAAGAC 7/11	3
Bcl I	T`GATC,A	2	Bcn I	CC,S`GG	7
Bfa I	C`TA,G	6	Bgl I	GCCN,NNN`NGGC	2
Bgl II	A`GATC,T	1	Blp I	GC`TNA,GC	-
Bpm I	CTGGAG 22/20	4	Bsa I	GGTCTC 7/11	1
BsaA I	YAC GTR	1	BsaB I	GATNN NNATC	-
BsaH I	GR`CG,YC	5	BsaJ I	C`CNNG,G	6
BsaW I	W`CCGG,W	4	BseR I	GAGGAG 16/14	-
Bsg I	GTGCAG 22/20	5	BsiC I	TT`CG,AA	1
BsiE I	CG,RY`CG	6	BsiHKA I	G,WGCW`C	5
BsiW I	C`GTAC,G	-	Bsm I	GAATG,C 7	-
BsmA I	GTCTC`/9	7	BsmB I	CGTCTC 7/11	2
BsmF I	GGGAC 15/19	2	BsoF I	GC`N,GC	30
Bsp120 I	G`GGCC,C	1	Bsp1286 I	G,DGCH`C	1
BspH I	T`CATG,A	3	BspM I	ACCTGC 10/14	1
BspM II	T`CCGG,A	-	Bsr I	ACT,GG`	20
BsrB I	GAG CGG	6	BsrD I	GCAATG, 8	4
BsrG I	T`GTAC,A	-	BssH II	G`CGCG,C	1
BssS I	C`TCGT,G	3	Bst1107 I	GTA TAC	-
BstB I	TT`CG,AA	1	BstE II	G`GTNAC,C	1
BstN I	CC`W,GG	11	BstU I	CG CG	27
BstX I	CCAN,NNNN`NTGG	3	BstY I	R`GATC,Y	9
Bsu36 I	CC`TNA,GG	1	Cac8 I	GCN NGC	30
Cfr10 I	R`CCGG,Y	2	Cla I	AT`CG,AT	-
Csp6 I	G`TA,C	4	Dde I	C`TNA,G	8
Dpn I	GA TC	25	DpnII	`GATC,	25
Dra I	TTT AAA	5	Dra III	CAC,NNN`GTG	-
Drd I	GACNN,MN`NNGTC	2	Dsa I	C`CRYG,G	2
Eae I	Y`GGCC,R	5	Eag I	C`GGCC,G	1
Ear I	CTCTTC 7/10	5	Eco47 III	AGC GCT	-
Eco57 I	CTGAAG 21/19	2	Eco72 I	CAC GTG	-
EcoN I	CCTNN`N,NNAGG	2	Eco0109 I	RG`GNC,CY	2
EcoR I	G`AATT,C	1	EcoR II	`CCWGG,	11
EcoR V	GAT ATC	1	Ehe I	GGC GCC	1
Esp I	GC`TNA,GC	-	Fnu4H I	GC`N,GC	30
Fok I	GGATG 14/18	10	Fse I	GG,CCGG`CC	-
Fsp I	TGC CCA	2	Gdi II	`YGGC,CG	5
Gsu I	CTGGAG 21/19	4	Hae I	WGG CCW	7
Hae II	R,CGC`Y	6	Hae III	GG CC	21
Hga I	GACGC 9/14	11	HgiA I	G,WGCW`C	5
HgiE II	ACCNNNNNNGGT -1/133	-	Hha I	G,CG`C	34
HinC II	GTY RAC	4	Hind II	GTY RAC	4
HinD III	A`AGCT,T	1	Hinf I	G`ANT,C	9
HinI I	GR`CG,YC	5	HinP I	G`CG,C	34
Hpa I	GTT AAC	1	Hpa II	C`CG,G	19
Hph I	GGTGA 12/11	15	Kas I	G`GCGC,C	1
Kpn I	G,GTAC`C	1	Mae I	C`TA,G	6
Mae II	A`CG,T	12	Mae III	`GTNAC,	19
Mbo I	`GATC,	25	Mbo II	GAAGA 12/11	17
Mlu I	A`CGCG,T	1	Mme I	TCCRAC 25/23	5
Mnl I	CCTC 10/10	19	Msc I	TGG CCA	1
Mse I	T`TA,A	23	Msl I	CAYNN NNRTG	5
Msp I	C`CG,G	19	MspAl I	CMG CKG	9
Mun I	C`AATT,G	-	Nae I	GCC GGC	-
Nar I	GG`CG,CC	1	Nci I	CC`S,GG	7
Nco I	C`CATG,G	-	Nde I	CA`TA,TG	-
NgoM I	G`CCGG,C	1	Nhe I	G`CTAG,C	1
Nla III	,CATG`	22	Nla IV	GGN NCC	16
Not I	GC`GGCC,GC	1	Nru I	TCG CGA	-
Nsi I	A,TGCA`T	-	Nsp7524 I	R`CATG,Y	3
NspB II	CMG CKG	9	NspH I	R,CATG`Y	3
PaeR7 I	C`TCGA,G	1	Pal I	GG CC	21
Pf1M I	CCAN,NNN`NTGG	1	Ple I	GAGTC 9/10	4
Pme I	CTTT AAAC	-	Pml I	CAC GTG	-
PpuM I	RG`GWC,CY	-	Psp1406 I	AA`CG,TT	3
PspA I	C`CCGG,G	-	Pst I	C,TGCA`G	1
Pvu I	CG,AT`CG	2	Pvu II	CAG CTG	3
Rea I	GT AC	4	Rsr II	CG`GWC,CG	-
Sac I	G,AGCT`C	-	Sac II	CC,GC`GG	1
Sal I	G`TCGA,C	1	Sap I	GCTCTTC 8/11	2
Sau3A I	`GATC,	25	Sau96 I	G`GNC,C	13
Sca I	AGT ACT	2	ScrF I	CC`N,GG	18
Sec I	C`CNNG,G	6	SfaN I	GCATC 9/13	14
Sfc I	C`TRYA,G	4	Sfi I	GGCCN,NNN`NGGCC	-
Sma I	CCC GGG	-	SnaB I	TAC GTA	-
Spe I	A`CTAG,T	-	Sph I	G,CATG`C	-
SpI I	C`GTAC,G	-	Srf I	GCCC GGGC	-
Ssp I	AAT ATT	2	Stu I	AGG CCT	-
Sty I	C`CWG,G	-	Taq I	T`CG,A	12
Tfi I	G`AWT,C	4	Tsp45 I	`GTSAC,	7
Tth111 I	GACN`N,NGTC	2	Tth111 II	CAARCA 16/14	7
Vsp I	AT`TA,AT	4	Xba I	T`CTAG,A	1
Xca I	GTA TAC	-	Xcm I	CCANNNN,N`NNNNTGG3	-
Xho I	C`TCGA,G	1	Xho II	R`GATC,Y	9
Xma I	C`CCGG,G	-	Xma III	C`GGCC,G	1

pGST-MOLUC GST Coding Region

```
1          21          41
|          |          |
AGCTTATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGAAGCTGTG

61          81          101
|          |          |
GTATGGCTGTGCAGGTCGTAAATCACTGCATAATTCGTGTCGCTCAAGGCGCACTCCCGT

121         141         161
|          |          |
TCTGGATAATGTTTTTGGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGC

181         201         221
|          |          |
TGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCA

241         261         281
|          |          |
CACAGGAAACAGTATTCATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGC
          M S P I L G Y W K I K G L V

301         321         341
|          |          |
AACCCACTCGACTTCTTTTTGGAATATCTTGAAGAAAAATATGAAGAGCATTGTATGAGC
Q P T R L L L E Y L E E K Y E E H L Y E

361         381         401
|          |          |
GCGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATC
R D E G D K W R N K K F E L G L E F P N

421         441         461
|          |          |
TTCCTTATTATATTGATGGTGATGTTAAATTAACACAGTCTATGGCCATCATAACGTTATA
L P Y Y I D G D V K L T Q S M A I I R Y

481         501         521
|          |          |
TAGCTGACAAGCACAAACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGC
I A D K H N M L G G C P K E R A E I S M

541         561         581
|          |          |
TTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTCGAGAATTGCATATAGTAAAGACT
L E G A V L D I R Y G V S R I A Y S K D

601         621         641
|          |          |
TTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAG
F E T L K V D F L S K L P E M L K M F E
```