

Table S4.187 putative miRNA hairpins in ESTs.

miRNA family	genus	ascencion number	begin	end	strand
	156 Brassica	CD835236	102	122	+
	156 Glycine	AW756919	150	170	+
	156 Glycine	BE210632	207	227	+
	156 Glycine	BE807821	163	183	+
	156 Glycine	BG650023	237	257	+
	156 Helianthus	BQ916415	146	166	+
	156 Lotus	BP046017	414	433	-
	156 Medicago	AW696064	195	215	+
	156 Nicotiana	BP129308	301	320	+
	156 Oryza	AU091537	10	29	+
	156 Oryza	CB643501	259	279	+
	156 Saccharum	CA294779	200	220	+
	156 Saccharum	CA294779	499	519	+
	156 Sesamum	BU667823	42	62	+
	156 Solanum	BF187371	137	157	+
	156 Solanum	BI432985	53	73	+
	156 Sorghum	CN132665	306	326	+
	156 Vitis	CF213216	613	632	-
	156 Zea	CF035522	274	294	+
	156 Zea	CF035885	87	107	+
	156 Zea	CF036467	272	292	+
	156 Zea	CF036852	224	244	+
	156 Zea	CF039276	299	319	+
	156 Zea	CF059465	221	241	+
159/JAW	Glycine	BE475558	364	383	+
159/JAW	Glycine	BG237979	294	313	+
159/JAW	Glycine	BM893181	275	295	+
159/JAW	Glycine	BQ453148	190	210	+
159/JAW	Glycine	BQ630503	315	334	+
159/JAW	Hordeum	BJ448559	376	395	+
159/JAW	Liriodendron	CK760683	210	229	+
159/JAW	Medicago	AW691937	476	496	+
159/JAW	Oryza	CF303003	369	389	+
159/JAW	Oryza	CR289947	402	421	+
159/JAW	Pennisetum	CD725199	431	451	+
159/JAW	Physcomitrella	BJ191979	443	462	+
159/JAW	Saccharum	CA079434	213	233	+
159/JAW	Saccharum	CA148454	337	357	+
159/JAW	Saccharum	CA203816	443	463	+
159/JAW	Saccharum	CA229394	138	157	-
159/JAW	Saccharum	CA229394	293	313	+
159/JAW	Saccharum	CA238586	430	450	+
159/JAW	Schedonorus	CK801264	363	383	+
159/JAW	Sorghum	CD204047	464	484	+
159/JAW	Sorghum	CD213290	448	468	+
159/JAW	Sorghum	CD224004	449	469	+
159/JAW	Triticum	CA483944	96	115	-
159/JAW	Triticum	CA484819	460	480	+
159/JAW	Triticum	CA731881	411	431	+
159/JAW	Vitis	CF210361	463	483	+
159/JAW	Vitis	CF983804	125	145	+

159/JAW	Vitis	CN007551	485	505	-
159/JAW	Zea	BM338067	314	334	-
159/JAW	Zea	BM338067	472	491	+
160	Glycine	CA801322	23	43	+
160	Oryza	CF327534	402	422	+
160	Oryza	CF327976	254	274	+
160	Triticum	BF293809	259	279	+
160	Zea	CD445121	285	305	+
162	Lupinus	BG149136	274	294	+
162	Medicago	BF003769	221	241	+
162	Oryza	CA764295	425	445	+
162	Vitis	CF516290	368	388	+
164	Populus	BU869001	16	35	+
164	Populus	CK113235	159	178	+
164	Triticum	CA704421	10	30	+
166	Glycine	BG316028	238	258	+
166	Glycine	BI893541	125	145	+
166	Glycine	BI972515	306	326	+
166	Glycine	BM308126	440	460	+
166	Glycine	CA819984	241	261	+
166	Hedyotis	CB086732	224	244	+
166	Hordeum	BQ760548	137	156	+
166	Ipomoea	BJ553847	310	329	+
166	Medicago	AI737566	133	152	+
166	Medicago	AJ502524	319	338	+
166	Medicago	AW685461	220	240	+
166	Medicago	AW694053	371	391	+
166	Oryza	CA760464	442	462	-
166	Sorghum	CN126049	157	177	+
166	Zea	CK369135	163	183	+
167	Glycine	BG509097	227	247	+
167	Glycine	BI095235	405	425	+
167	Oryza	BX928793	277	296	+
167	Oryza	CF309326	102	122	+
167	Phaseolus	CA916400	135	155	+
167	Saccharum	CA284319	123	143	+
167	Saccharum	CA284394	629	649	-
167	Saccharum	CA287325	136	156	+
167	Zea	CF630597	365	385	-
168	Arabidopsis	H77158	42	62	+
168	Betula	CD271355	99	119	+
168	Glycine	AW424354	20	40	+
168	Hedyotis	CB076866	168	188	+
168	Lycopersicon	BE461110	16	36	+
168	Lycopersicon	BF097936	16	36	+
168	Oryza	CA756101	118	138	+
168	Populus	BU809183	109	129	+
168	Populus	BU886509	70	90	+
168	Saccharum	CA123434	184	204	+
168	Saccharum	CF576659	193	213	+
168	Solanum	CK243936	10	30	+
168	Solanum	CK246126	12	32	+
168	Solanum	CK254942	85	105	+

168 Sorghum	CD205059	182	202	+
168 Vitis	CF604588	52	72	+
168 Zea	BG842528	97	117	+
168 Zea	BU099153	62	82	+
168 Zea	CA826777	62	82	+
169 Glycine	AW596073	26	46	+
169 Glycine	BF598910	85	105	+
169 Glycine	CA953278	90	110	+
169 Oryza	BQ906497	169	189	+
169 Oryza	BX899553	198	217	+
169 Oryza	CF196420	107	127	-
169 Oryza	CF280503	141	161	+
169 Oryza	CF305246	175	195	-
169 Oryza	CF310317	155	174	-
169 Oryza	CF336482	692	673	-
169 Populus	BU862460	89	109	+
169 Populus	BU865420	164	184	+
169 Triticum	BJ225371	502	521	-
169 Zea	CD960569	293	313	-
171 Arabidopsis	BX838271	125	144	+
171 Glycine	CA937914	265	285	+
171 Hedyotis	CB087210	482	502	+
171 Hordeum	CA009309	394	414	+
171 Oryza	CA756196	210	230	+
171 Oryza	CF329292	308	328	+
171 Triticum	BJ275219	359	378	-
171 Triticum	CD910903	165	185	+
171 Zea	CD440809	221	241	+
171 Zea	CD441617	530	550	-
172 Citrus	CF506714	404	424	+
172 Glycine	BI320499	384	404	+
172 Glycine	BU084569	200	220	+
172 Lycopersicon	AI484737	448	468	+
172 Solanum	BQ114970	492	512	-
393 Oryza	CA765497	55	75	+
393 Oryza	CB639509	137	157	+
393 Populus	CF231897	95	115	+
394 Glycine	AW099182	37	56	+
394 Glycine	BG653184	24	43	+
394 Robinia	BI642560	222	241	+
395 Glycine	AW596801	100	120	+
395 Oryza	CA760441	112	132	-
395 Oryza	CA760441	255	275	-
395 Oryza	CA760441	398	418	-
395 Oryza	CA764701	188	208	+
395 Oryza	CA764701	348	368	+
395 Oryza	CA764701	489	509	+
395 Oryza	CA764701	627	647	+
395 Triticum	CK193704	347	366	-
395 Triticum	CK194045	375	395	+
396 Brassica	BQ704984	486	506	-
396 Glycine	BG405077	196	216	+
396 Glycine	CA784869	176	196	+

396	Mesembryanthemum	BM301359	134	154	+
396	Mesembryanthemum	BM302146	159	179	+
396	Mesembryanthemum	BM302151	159	179	+
396	Mesembryanthemum	CA833245	158	178	-
396	Oryza	CA765692	209	229	+
396	Populus	BU883864	195	215	+
396	Populus	CK113079	195	215	+
396	Prunus	BU039965	39	59	+
396	Prunus	BU047426	189	209	+
396	Saccharum	CA240723	81	101	+
396	Solanum	BI178715	123	143	+
396	Zea	BM350675	507	527	-
396	Zea	CF349199	83	103	+
397	Hordeum	BG415888	248	267	-
397	Hordeum	BU966899	296	315	-
398	Citrus	C24241	168	187	+
398	Glycine	BM732696	138	158	+
398	Glycine	CA937628	127	147	+
398	Glycine	CB063312	77	97	+
398	Helianthus	CD849864	175	195	-
398	Lactuca	BQ988187	98	118	+
398	Lactuca	BU005500	95	115	+
398	Lotus	CB829453	136	156	+
398	Medicago	AW584548	130	150	+
398	Medicago	BE943500	89	109	+
398	Nicotiana	CK289462	84	104	+
398	Oryza	CB673340	388	408	+
398	Zea	CF064240	330	350	-
399	Medicago	AJ502674	100	119	+
399	Populus	BI125288	135	155	+

Near matches to *Arabdiopsis* and *Oryza* miRNAs (19/20 nt matches) were found in non-human, non-mouse ESTs in the April 5, 2004 release of dbEST from NCBI. Matches with potential miRNA-like hairpins were identified with MIRcheck. The 5' and 3' ends of the miRNAs are inferred from the ends of the *Arabdiopsis* or *Oryza* homolog. Because many plant miRNAs have heterogeneity at either the 5' or 3' end, the ends of the sequences listed should be considered to be approximations. Hairpin length is defined as the minimal sequence length containing the miRNA, miRNA*, and intervening sequence

miRNA sequence	hairpin ϵ hairpin length	hairpin sequence
UGACAGAAGAGAGUGAGCACA	5prime	81 CAUAGCAACUGACAGA/
UUGACAGAAGAUAGAGAGCAC	5prime	112 UUAAGGUUGUUGACAG
UUGACAGAAGAUAGAGAGCAC	5prime	85 GUGAUGCUGUUGACAG
UGACAGAAGAGAGUGAGCACA	5prime	83 GAGAGAGGCUGACAGA
UGACAGAAGAGAGAGAGCACA	5prime	83 AUCUCAUGUUGACAGA.
UUGACAGAAGAUAGAGAGCAC	5prime	86 UGAUGGAUGUUGACAG
UUGACAGAAGAGAGAGAGCAC	5prime	86 UUCAUGCAUGUUGACA
UUGACAGAAGAUAGAGGGCAC	5prime	111 GUAAGGUUGUUGACAG
UUGACAGAAGAUAGAGAGCAC	5prime	131 UGUGAGAUUGUUGACA
UGACAGAAGAGAGUGAGCACA	5prime	87 UGGGAGNUCUGACAGA
UGACAGAAGAGAGUGAGCACA	5prime	86 GCGAGAUUGUUGACAG
UGACAGAAGAGAGUGAGCACA	5prime	98 GGUGGAGGCUGACAGA/
UGACAGAAGAGAGUGAGCACA	5prime	84 UUUGAAGGUUUGACAG
UUGACAGAAGAGAGAGAGCAC	5prime	83 AUUAAUUUGUUGACAG.
UUGACAGAAGAUAGAGAGCAC	5prime	91 UGAUAAUUGUUGACAG
UGACAGAAGAGAGUGAGCACA	5prime	86 AAUCAAGACUGACAGA/
UGACAGAAGAGAGUGAGCACA	5prime	84 CUUGAGAGAUUGACAG
UGACAGAAGAGAGAGAGCAUG	5prime	206 UGCCUCACAAUGACAG.
UGACAGAAGAGAGUGGGCACA	5prime	99 AGGUGAAAGCUGACAG
UGACAGAAGAGAGUGAGCACA	5prime	84 UUGAAGGUUUGACAGA
UGACAGAAGAGAGUGAGCACA	5prime	99 AGGUGAAAGCUGACAG
UGACAGAAGAGAGUGAGCACA	5prime	84 UUGAAGGUUUGACAGA
UGACAGAAGAGAGUGAGCACA	5prime	84 UUUGAAGGUUUGACAG
UGACAGAAGAGAGUGAGCACA	5prime	84 UCGAGAGAUUGACAGA
UUGGACUGAAGGGAGCUCCC	3prime	169 GCGACGGUAAGAGAGC
UUGGACUGAAAGGAGCUCCU	3prime	182 AAGAGAGUGAAGGAGC
UUUGGAUUGAAGGGAGCUCUA	3prime	175 AUUAUGAAGUGGAGCU
AUUGGAGUGAAGGGAGCUCCA	3prime	166 AAACCCAACUUGGAGU/
UUGGACUGAAGGGAGCUCCC	3prime	168 CUAAGGUAAGAGAGC
UUUGGAUUGAAGGGAGCUCUG	3prime	174 GUUUGGAGGUGGAGCI
UUGGACUGAAGGGAGCUCCC	3prime	167 GUUAUGGACUAAGGAG
UUUGGAUUGAAGGGAGCUCUA	3prime	168 UUAAGGGGUGGAGCL
UUUGGAUUGAAGGGAGCUCUG	3prime	252 UUGUGGACGUUGAGCL
UUUGGAUUGAAGGGAGCUCUG	3prime	249 UUGUGGACGUUGAGCL
UUUGGAUUGAAGGGAGCUCUG	3prime	226 GAUUGGAAGCGGAGCL
UUGGACUGAAGGGAGCUCCA	3prime	169 ACCUUGAUUGUGGAGC
UUUGGAUUGAAGGGAGCUCUG	3prime	226 UUUGAAGCGGAGCUCC
UUUGGAUUGAAGGGAGCUCUG	3prime	226 GAUUUGAAGCGGAGCU
UUUGGAUUGAAGGGAGCUCUG	3prime	227 GAUUUGAAGCGGAGCU
UUGGAUCGAAGGGAGCUCUU	3prime	176 GGAAAGAGAGAGGAGC
CUUGGAUUGAAGGGAGCUCCU	3prime	176 AAGUGAUCGAAGAGCU
UUUGGAUUGAAAGGAGCUCUU	3prime	225 GAUUUGAAGCGGAGCU
UUUGGAUUGAAGGGAGCUCUG	3prime	175 GUUUUGAGGUGGAGCI
UUUGGAUUGAAGGGAGCUCUG	3prime	222 GAUUCGAAGCGGAGCU
UUUGGAUUGAAGGGAGCUCUG	3prime	222 GAUUCGAAGCGGAGCL
UUUGGAUUGAAGGGAGCUCUG	3prime	222 GAUUCGAAGCGGAGCL
UUGGACUGAAGGGAGCUCCC	3prime	196 AGUUUGAGGGAGCUC
UUUGGAUUGAAGGGAGCUCUG	3prime	252 UUGUGGACGUUGAGCL
UUUGGAUUGAAGGGAGCUCUG	3prime	173 GUUUAGAGGUGGAGCL
UUUGGAUUGAAGGGAGCUCUA	3prime	172 GUUUUGGAGUGGAGCL
UUUGGACUGAAGGGAGCUCCU	3prime	87 CUGCAGAAAUGGGGGL

UUUGGAUUGAAGGGAGCUCUA	3prime
CUUGGAUUGAAGGGAGCUCCU	3prime
UUGGAUCGAAGGGAGCUCUU	3prime
UGCCUGGCUCCCUGUAUGCCA	5prime
UGCCUGGCUCCCUGUAUGCCA	5prime
UGCCUGGCUCCCUGUAUGCCA	5prime
UGCCUGGCUCCCUGUAUGCCA	5prime
UGCCUGGCUCCCUGUAUGCCA	5prime
UGCCUGGCUCCCUGUAUGCCA	5prime
UCGAUAAACCUCUGCAUCCAG	3prime
UCGAUAAACCUCUGCAUCCAG	3prime
UCGAUAAAGCCUCUGCAUCCAG	3prime
UCGAUAAACCUCUGCAUCCAG	3prime
UGGAGAAGCAGGGCACGUGCA	5prime
UGGAGAAGCAGGGCACGUGCA	5prime
UGGAGAAGCAGGGCACGUGCA	5prime
UCGGACCAGGCUUCAUUCCCU	3prime
UCGGACCAGGCUUCAUUCCCC	3prime
UCGGACCAGGCUUCAUUCCCG	3prime
UCGGACCAGGCUUCAUUCCCC	3prime
UCGGACCAGGCUUCAUUCCCC	3prime
UCGGACCAGGCUUCAUUCCCC	3prime
UCGGACCAGGCUUCAUUCCCC	3prime
UCGGACCAGGCUUCAUUCCUC	3prime
UCGGACCAGGCUUCAUUCCCC	3prime
UCGGACCAGGCUUCAUUCCCC	3prime
UCGGACCAGGCUUCAUUCCCC	3prime
UCGGACCAGGCUUCAUUCCUC	3prime
UCGGACCAGGCUUCAUCCCCU	3prime
UCGGACCAGGCUUCAUUCCCC	3prime
UCGGACCAGGCUUCAUUCCCC	3prime
UGAAGCUGCCAGCAUGAUCUG	5prime
UGAAGCUGCCAGCAUGAUCUA	5prime
UGAAGCUGCCAGCAUGAUCUG	5prime
UGAAGCUGCCAGCAUGAUCUG	5prime
UGAAGCUGCCAGCAUGAUCUU	5prime
UGAAGCUGCCAGCAUGAUCUG	5prime
UGAAGCUGCCAGCAUGAUCUG	5prime
UGAAGCUGCCAGCAUGAUCUG	5prime
UGAAGCUGCCAGCAUGAUCUG	5prime
UCGCUUGGUGCAGGUCGGGAA	5prime
UCGCUUGGUGCAGGUCGGGAA	5prime
UCGCUUGGUGCAGGUCGGGAA	5prime
UCGCUUGGUGCAGGUCGGGAC	5prime
UCGCUUGGUGCAGGUCGGGAC	5prime
UCGCUUGGUGCAGAU CGGGAC	5prime
UCGCUUGGUGCAGGUCGGGAA	5prime
UCGCUUGGUGCAGGUCGGGAA	5prime
UCGCUUGGUGCAGAU CGGGAC	5prime
UCGCUUGGUGCAGAU CGGGAC	5prime
UCGCUUGGUGCAGGUCGGGAC	5prime
UCGCUUGGUGCAGGUCGGGAC	5prime

168 GGUUUUGGAGUGAGCU(
178 AGGCGAUCGAAGAGCU
178 ACAGACAGAGAGGAGC
80 AUGUGUAUGUGCCUGG
130 AAAGGGGAUAUGCCUG
82 CUUGAGAGCGUGCCUC
87 UGAUAUGAUGUGCCUG
81 GUGUCGUGUGUGCCUC
80 GUGAAGUCACUGGAAG
87 AAGUUCGUCACUGGAU
114 GGGUGAUGCCUGGGC(
85 UUGAAGUCACUGGAUG
137 UGAGCAAGAUGGAGAA
137 UGAGCAAGAUGGAGAA
152 CGCGCGAGGUGGAGAA
136 CACGUCUUGAGGGGAA
107 UGUCUUUUGAGGGGAA
156 UGGGGUUGAUGGGGAAL
86 AGGUGUUGAGGGGAAL
128 UGAGGUUGAGAGGAAL
165 UAUCUUUUGAGGGGAA
78 AUGGUUGUCGAGGGGA
169 GGAGUUUGAGGGGGAL
92 GUUAGGUUGAGAGGA
96 CUUAUUUGAGGGGAAU
97 UAUUUUUUGAGGGGAA
173 AGGAGGUGUUUGGAAL
97 UUAGGUUAAGGGGAUU
108 UUGCUCUGAGUGGA
87 CUUACUUUGAGGGGAA
64 UGCAGCAGUUGAAGCL
82 AAGGAAAAAGUGAAGCI
90 CAUUAAGGAGCUGAAGC
100 ACAAGUUGGUGAAGCU
69 ACUACCAGUUGAAGCU
149 ACAAGUUGGUGAAGCU
149 CACAAGUUGGUGAAGC
139 ACAAGUUGGUGAAGCU
80 CUCUAGUAGCUGAAGC
104 GGCUCGGAUUCGCUUC
166 GUCUCUAAUUCGCUUG
88 GUCUCUAAUUCGCUUG
140 GUCUCCGAUUCGCUUG
145 GCCUCUAAUUCGCUUG
145 GCCUCUAAUUCGCUUG
67 GCCUCGGGCUCGCUUC
118 GUCUCUAAUUCGCUUG
140 GUCUCUGAUUCGCUUG
65 GCCUCGGGCUCGCUUC
65 GCCUCAGGCUCGCUUG
133 GCCUCUCAUUCGCUUG
134 GCCUCUCAUUCGCUUG
133 GCCUCUCAUUCGCUUG

UCGCUUGGUGCAGAU CGGGAC	5prime
UCGCUUGGUGCAGGUC GGGAA	5prime
UCGCUUGGUGCAGAU CGGGAC	5prime
UCGCUUGGUGCAGAU CGGGAC	5prime
UCGCUUGGUGCAGAU CGGGAC	5prime
CAGCCAAGGAUGACU UGCCGG	5prime
CAGCCAAGGAUGACU UGCCGG	5prime
CAGCCAAGGAUGACU UGCCGG	5prime
CAGCCAAGGAUGACU UGCCGA	5prime
CAGCCAAGGAUGACU UGCCGG	5prime
UAGCCAAGGAGACUG CCUAUG	5prime
CAGCCAAGGAUGACU UGCCGA	5prime
UAGCCAAGGAGACUG CCCAUG	5prime
UAGCCAAGGAGACUG CCUAUC	5prime
UAGCCAAGAAUGGCU UGCCUA	5prime
CAGCCAAGGAUGACU UGCCGA	5prime
CAGCCAAGGAUGAUU UGCCGA	5prime
UAGCCAAGGAUGAUU UGCCUG	5prime
UAGCCAAGGACAGACU UGCCG	5prime
UGAUUGAGCCGCGCC AAUAUC	3prime
UCAUUGAGCCGUGCC AAUAUC	3prime
CGAUUGAGCCGUGCC AAUAUC	3prime
UGAUUGAGCCGUGCC AAUAUC	3prime
UGAUUGAGCCGUGCC AAUAUC	3prime
UGAUUGAGCCGUGCC AAUAUC	3prime
CGAUUGAGCCGUGCC AAUAUC	3prime
UGAUUGAGCCGUGCC AAUAUC	3prime
UGAUUGAGCCGUGCC AAUAUC	3prime
UGAUUGAGCCGUGCC AAUAUC	3prime
AGAAUCUUGAUGAUG CUGCAA	3prime
AGAAUCUUGAUGAUG CUGCAU	3prime
AGAAUCUUGAUGAUG CUGCAU	3prime
AGAAUCUUGAUGAUG CUGCAU	3prime
AGAAUCUUGAUGAUG CUGCAU	3prime
UCCAAAGGGAUCGCAU UGAUC	5prime
UCCAAAGGGAUCGCAU UGAUC	5prime
UCCAAAGGGAUCGCAU UGAUC	5prime
UUGGCAUUCUGUCC ACCUCC	5prime
UUGGCAUUCUGUCC ACCUCC	5prime
UUGGCAUUCUGUCC ACCUCC	5prime
UGAAGUGUUUGGGGGA ACUCU	3prime
GUGAAGUGUUUGGGGGA ACUC	3prime
GUGAAGUGUUUGGGGGA ACUC	3prime
AUGAAGUGUUUGGAGGA ACUC	3prime
GUGAAGUGCUUGGGGGA ACUC	3prime
GUGAAGUGUUUGGGGGA ACUC	3prime
GUGAAGUGUUUGGAGGA ACUC	3prime
GUGAAGUGUUUGGGGGA ACUC	3prime
GUGAAGUGUUUGGAGGG ACUC	3prime
AUGAAGUGUUUGGGGGA ACUC	3prime
UUCCACAGCUUUCUUG AACUU	5prime
UUCCACAGCUUUCUUG AACUU	5prime
UUCCACAGCUUUCUUG AACUG	5prime

67 GCCUCGGGCUCGCUU(
110 GUCUCUAAUUCGCUUG
65 GUCUCGGGCUCGCUU(
65 GCCUCGGGCUCGCUU(
65 GCCUCGGGCUCGCUU(
111 AGUGUAGUGCAGCCAA
103 AGUGUAGUGCAGCCAA
84 AGAGUGAUGCAGCCAA
95 GCCAUGGUGCAGCCAA
92 AACGGGAUGCAGCCAA
88 CGACUCAGACUAGCCA
91 GCCAUGGUGCAGCCAA
88 CCACUCAGGCUAGCCA
105 CCACUCAGGCUAGCCA
107 UUGCCUCUGUUAGCCA
101 AGAUUGAUGCAGCCAA
84 AGUUUGAUGCAGCCAA
99 CCAUCUUCGAUAGCCA
93 GAACUAGGUGUAGCCA
92 GAGUCCCUUUGAUUU
79 GUUCAACGGGAUUAUG
79 CAUAAACGAGAUCUUG
78 GGUCACUAUGAUGUUG
107 GGGAGAGUGCGAUGUL
99 GGUAGCUAUGAUGUUG
80 AGUGAACGCGUAUUG
94 UGGUCACUAUGAUGUU
113 GAGAGAGUGCGAUGUL
113 AGAGAGUGCGAUGUUG
107 UGCUCGCUGUAGCAGC
114 UGUUUUGCGGAUGUAGC
120 UUAUUUGCGGAUGUAG
106 UGUUUUGCAUAUGUGGC
93 UGCUUGCUAGUGCAGC
95 GGGGAAGCAUCCAAAG
95 GUGGAGGACUCCAAAG
89 UUGGAGUGUCCAAAG
65 CAGAGUUUCUUGGCAU
79 CAGAGUUUAUUGGCAU
88 CAGAGUUUCUUGGCAU
69 GUUGUCUCUUGGAGUL
68 UAUUAUCGUGAGUUCC
68 CAUUGUCGUGAGUUCC
69 UAUUACUAUGAGUUCU
122 UUGUCCACUGGAGUUC
66 AGUCCCUAGGAGUUCC
69 UAUUAUCAAGAGUUCU
68 UAUUGUCGUGAGUUCC
62 GUCUUAGCAUGGGUUC
66 UCUUACCAUGGGUUCC
117 GGUUAUAUUUUUCCAC
87 GUCAUGCUUUUCCACA
101 UUUGUAUUCUCCACA

UUCCACAGCUUUCUUGAACUG	5prime	181 UUUGUAUUCUUCCACA
UUCCACAGCUUUCUUGAACUG	5prime	182 UUUGUAUUCUUCCACA
UUCCUCAGCUUUCUUGAACUG	5prime	183 UUUGUAUUCUUCCUCA
UCCACAGCUUUCUUGAACUU	5prime	98 CGCCAUUUUUUCCAC
UUCCACAGCUUUCUUGAACUG	5prime	134 UUUGUGAUCUUCCACA
UUCCACAGCUUUCUUGAACUG	5prime	114 UUUGUAUUCUUCCACA
UUCCACAGCUUUCUUGAACUG	5prime	114 UUUGUAUUCUUCCACA
UUCCACAGCUUUCUUGAACUG	5prime	113 UUUGUAUUCUUCCACA
UUCCACAGCUUUCUUGAACUG	5prime	113 UUUGUAUUCUUCCACA
UUCCACAGCUUUCUUGAACUG	5prime	95 UUUGUGAUCUUCCACA
UUCCACAGCUUUCUUGAACUU	5prime	103 GUCAUGCUUUUCCACA
UUCCACAGCUUUCUUGAACUG	5prime	97 CUUUGUGAUCUUCCAC
UUCCACAGCUUUCUUGAACUG	5prime	110 CCUGCCAUCUUCCACA
CGUUGAGUGCAGCGUUGAUG	5prime	87 GCAGAGGUGCCGUUG/
CGUUGAGUGCAGCGUUGAUG	5prime	87 GCAGAGGUGCCGUUG/
UGUGUUCUCAGGUCACCCCUU	3prime	133 GUGAACCCCAGAGGAG
UGUGUUCUCAGGUCGCCCCUG	3prime	98 AGGAAUUCUACAGGGU
UGUGUUCUCAGGUCACCCCUU	3prime	96 GUUUUUCUCAGAGGAG
UGUGUUCUCAGGUCACCCCUU	3prime	93 CUCGGAGGAGUGAAUC
CGUGUUCUCAGGUCGCCCCUG	3prime	66 GACAUCCAACAGGUGU
UGUGUUCUCAGGUCGCCCCUG	3prime	73 AGGCAUCCAACAGGUG
UGUGUUCUCAGGUCGCCCCUG	3prime	73 AGGCAUCCAACAGGUG
UGUGUUCUCAGGUCACCCCUU	3prime	113 GUCUAUCUCAAGGAG
UGUGUUCUCAGGUCACCCCUU	3prime	88 UUCUAUCUCAGAGGAG
UGUGUUCUCAGGUCGCCCCUG	3prime	77 UGAUGUUCUACAGGGU
UGUGUUCUCAGGUCGCCCCUG	3prime	89 AAGUGUUCAACAGGGG
UGUGUUCUCAGGUCGCCCCUG	3prime	68 GAGUUCCUACAGGGGC
UGUGUUCUCAGGUCGCCCCCG	3prime	108 CAGUUCCGGCGGGGGC
UGCCAAAGGAGAGUUGCCCUG	3prime	74 AAAUCAGCUAUAGGGC
UGCCAAAGGGAGAAUUGCCCUG	3prime	82 UAAAGAAUAACAGGGC

nn-

3

3

3.

ascencion numbers with redundant foldbacks

AGAGAGUGAGCACACAAAAGUAAUCUGCAUAUACUGCAUUUGCUUCUCUUGCGUGCUCACUGC
;AAGAUAGAGAGCACAGAUGAUGAUUGCAUAUUUAUAAUAUAGCAGGGAACUCAUGAUGA
;AAGAUAGAGAGCACUGAUGAUGAAUUGCAUGAAAGGGAAUGGCAUCUCACUCCUUUGUGCUC
\AGAGAGUGAGCACAUAGCUAGUGGUUUUUGUAUGAGGGCAUACAAUUGCGGGUGCGUGCUCAC
AGAGAGAGAGCACAAACCCGGGAUUGGCUAAAGGAGUCUUUGCCUUUGUUGGGAGUGUGCCCL
;AAGAUAGAGAGCACAGAGAAGCAUGAAUUGCACAUAAGAUUUGCAAUUCACUCCUUCGUGCUC
,GAAGAGAGAGAGCACAAACCCAGGAUUGGUGAAAGAGAGUCUUUGCUUUUUGUUGGGAGUGUGC
;AAGAUAGAGGGGCACUAAGGAUGAUAUGCAUACACAUAUAUACAACAUGGAGGAGGAGCUU/
\GAAGAUAGAGAGCACAGAUGAUGAUGAUGUCUGCUAAAUCUGGGAUUGGAGAGGGGCACUGAA
\AGAGAGUGAGCACACACGGUGCUUUUCUAGCAUGCAAGAGCCNAUGCUGGGAGCUGUGCGUI
CB647907,CB653304,CB677501,CB677777

\AGAGAGUGAGCACACAUGGUGCCUUUCUUGCAUGAUGAACGAUCGAGAGGUUCAUGCUCGA/
;AAGAGAGUGAGCACACACGGUGGUUUUCUJANCAUGAGUGCCAUGCUGGGAGCUGUGCGUGC/
AAGAGAGAGAGCACAGCCCGCCAUUGACAAAGAGGUCUUUGCCUUUUGUGGGAUUGUGCUCU
BQ119413

AGAGAGUGAGCACACGCAGUCGAAUUGUAUAAACAUUUUAUACA AUUGUCAUUUGCGUGUGCUC
;AAGAGAGUGAGCACACGGCGCGACGAACGGCAUAAUAUGUAUGUCGUCCUCGCCGCGUGCUC
AAGAGAGAGAGCAUGCUGGUGGGAAAACAAUUAACAACCUUUGCUCAUUCUGAUCUGGAAAUGCI
;AAGAGAGUGGGGCACACAUGGUGCCUUUCUUGCAUGAUGUAUGAUCGAGAGAGUUCAUGCUC
CF036182,CF036471,CF036982,CF037245

;AAGAGAGUGAGCACACAUGGUGCCUUUCUUGCAUGAUGUAUGAUCGAGAGAGUUCAUGCUCG
\AGAGAGUGAGCACACACGGUGGUUUUCUJACCAUGAGUGUCAUGCUAGGAGCUGUGCGUGCUC/
;AAGAGAGUGAGCACACACGGUGGUUCCUJACCAUGAGUGUCAUGCUAGGAGCUGUGCGUGC/
\AGAGAGUGAGCACACGGCGCGACGAACGACAUAUGCAUGUAUGCCGUCCUCGCCGCGUGCUC/
;UUUCUUCAGUCCACUUAUGGGUGACAUAAGAUUUCAAUUAAGCUGCCGACUCAUUCAUCCAA/
;UCCCCUCAGCCCAUUCAUUGGAGAUAAACGAAAGAUUGGGUUGCUGAAUUAACUGCUAGCUCAC,
ICCUUGAAGUCCAAUUGAGGAUCUUAACUGGGUGAAUUGAGCUGCUUAGCUAUGGAUCCACAG
UCCUGCACUCCAAGUCUGAAAGGAUAUGAUGGUAAACCUCUACUGCUAGUUCAUGGAUACCL
BQ630517

BJ456281

;CUCUCUUCAGUCCAGUCCAAGAUAGUAUUAAAGCCAAUCUCCGCUGCUGACUCGUUGGCUCAU
JUCCUUUAGUCCAAAUUGGAUCUUGCUAUGUUGAUAGAGCUGCUUAGCUAUGGGUCCCUCA/
BI265403

JCCUUUCGGUCCAAAAAGGGGUGUUGCUGUGGGUCGAUUGAGCUGCUGGGUCAUGGAUCCCC
JCCUAUCAUUCCAAUGAAAGGUUGUUUUUGUGGGUUGGUACAGCUGCUCGUUCAUGGUUCCC/
;UCCGUUUUCGGUCCAAUAGUGGCUGCGACGGAAGGUGGUCCCGCUGCCGAUACACGUCC
CA236922,CA261537

JCCUAUCAUUCCAAUGAAGGGCCGUUCUGAAGGGUUGUUCGCGUCUCGUUCAUGGUUCCCA/
JCCUAUCAUUCCAAUGAAGGGCCGUUCUGAAGGGUGGUUCCGCGUCUCGUUCAUGGUUCCCA/
;UCCCUUCAUCCAAGCACGAGGGAAAGAUUGGUGGGUUAUCUCCCGGGUCAUGCACACC
|CCCUUCGAUCCAUCCAGGAGGGGAAGUGGUCGGUUGCAGCUGCCGGUUAUGGAUGCCUC/
JCCUAUCAUUCCAAUGAAGGGCCGUUCUGAAGGGUGGUUCCGCGUCUCGUUCAUGGUUCCCA/
CK801265

JCCUAUCAUUCCAAUGAAGGGCCCUUUUCAUGGGUGGUUCCGCGUCUCGUUCAUGGUUCCCA/
CD213304

CD224964,CF433018

\CUUCAGUCCACUCAUGGGAGGUAGCGGGGAUUGAACGAGCUGCCGACUCAUUCACUCGAGC/
JCCUUUCGGUCCAAAAAGGGGUGUUGCUGUGGGUCGAUUGAGCUGCUGGGUCAUGGAUCCCC
JCCUAUCAUUCCAAUGAAGGGUCUACCGGAAGGGUUGUGCAGCUGCUCGUUCAUGGUUCCC,
JCCUUGAAGUCCAAUAGAGGGUCUUAACUGGGUAGAUUGAGCUGCUGACUUAUGGAUCCACAC/
JUCCUUUGCAGCCCAAACAACUCCAUCGCUGAAGAAGAUGAUGAACUUCAUGCUCUUGUUU

DCUUGAGUCCAAUAGAGGUCUACUGGGUAGAUGAGCUGCUGACUUAUGGAUCCACAGCCCI
JCCCUUCGAUCCAAUCCAGGAGGGGAAGUGGUCGGUUGCAGCUGCCGGUUCAUGGAUACCUCI
:UCCCUUCAAUCCAAGCACGAGGGAAAGAUGAUGGCGGGCUCaucUCCCGGGUCAUGCACACC
3CUCUUGUAUGCCAUUUGUAGAGCUCaucGAAGCAUCAUAGACCUUUGUGGAUGGCGUAUG/
CF327535
CF330305
3GCUCCUGUAUGCCACUCAUCCAGAGCAACACCUUUUGCAAUAAGGUUGCCUGCGAUGGAUC
3GCUCCUGUAUGCCACACAUGUAGCCAACCCGUGGCGUGAUUGGAUGCUGUGGGUGGCGUC
3CAGCGGUUUUAUCGAUCUCUUCUGAAUUUGAUUAACACAAACCAUGAAUCGAUCGAUAAACCL
IGCAGCGGUUCAUCGAUCUGUUCUGAAUUUGUUUGUCUGUAACAAACAUGAAUCGGUC
3CAGUGGUUUUAUCGAUCUCUUCUCCUGCCUUGUGCUGCUCUGAUCGAUGCCCGUGCUGAUUCL
3CAGCGGUUCAUCGAUCUCUUCUGAAAUUGUUGUGAAAAAAGCAGAUCAAGAAUCGGUCGAU,
,GCAGGGCACGUGCAUUAACUCAUGCAGACAGAGUGAGAGAGACAUUUCUUGCUGGAGUU
,GCAGGGCACGUGCAUUAACUCAUGCAUACAGAGUGAGAGAGACAUUUCUUGCUCGAGUU
AGCAGGGCACGUGCAUCCAUUUCAGCUCGGCAUUCUCCGGCGUCCGGCCGGCCGGCUGCCG(
\UGCAGUGUGGUCCAAGGAGAUGAUUAUACAUACCAUACUCAUAUCUUCACAAACUCAUC/
\UGUUGUCUGGCUCGAGGACCCUUCUUCAUUCUUGAUUCUUGUGUAGACUACUAUGCUUGUGGU
JGUUGUUUGGCUCGAGGUAACUAUGCAUGGUCUUAUUUUUGUUCAUUCUUGAAGCUUUAAUUI
JGUUGGCUGGCUCGAGGCUUUUCAAGAGGAGGUUCUCACUGGCAAGAACUAUAAGGCUUCG/
CD391899,CD405934
\UGUUGUCUGGUGCGAGGCCACCAACUAGAUCCAUGGAUCCUUCUUUAUAUAUUAUACAGAU
AAUGACNCCGGGUCCNAAAGAGAGACNCUCGCAUGGCGUGCGGUGGUGCGUUUCGGACCAG
JGUUGGCUGGCUCGAUGCACUUAUCUAUCUUCUCAAACAAUGCGACAUUAACAUACAU,
BQ255147
JGUUGGCUGGCUCGAGGCUUUUCAGUUUCACAAAGGAAGUUCAGUCUUAUUUGUAUGAACUAI
\UGUUGUCUGGCUCGAGGACGCUUUCUUCUGAUCUAAUGCAAUUUGUGGUCAUGGAUUGU/
JGAGGUUUUGUCCAAGAUCAUCACACAUCAUGUUUCUUCUUCUCCUUAUAAUUCUUGAUUA,
JGUUGUCUGGUUCAAGGUCUCCACAUUGUGCAAAAUGUUCAUUCAUGGAGGCACAGGAUGCUL
\UGUUGUCUGGUUCAAGGUCUCGCUUGUGAUUUAAAGGAUGAUUUUGUGCAUGCGUAAUUUUUAI
\UGUUGUCUGGCUCGAGGUGCAGAAACAUGCAGAUUCUCAUCGGUCUAGGUUCUUGUCGAUCU(
BM892909,BU548271,CD411229
UGCCAGCAUGAUCUAGCUUUGGUUAGUGGGAGCGAGAUAGUGCUAACCCUCACUAGGUCAUG
:UGCCAGCAUGAUCUGAUGAGUGCUUAUUAGGUGAGGGCAGAAUUGACUGCCAAAACAAAGAU
CF322848
GCCAGCAUGAUCUUAACUUCUCCUCACUUGGUUGAGGAGAGAUCAUGAUCUUGGCGAGUUUCA
IGCCAGCAUGAUCUGAUGGUGGUUAUAUAUGAAUAUAUGAUGUCUUUACCUCUGAUCUCUCCCU
:UGCCAGCAUGAUCUGAUGGUGGUUAUAUAUGAAUAUAUGAUGUCUUUACCUCUGAUCUCUCCC
CF571260
:UGCCAGCAUGAUCUGAGGUGUCCACAGCAUAUAUAUGGAAGCAGCUAGCGAUCAGAUCAUGC
3GUGCAGGUCGGGAACCAAUUCGGCUGACACAGCCUCGUGACUUUUAAACCUUUAUUGGUUU(
3GUGCAGGUCGGGAACUGCUUCGCUUUUGCCCUUAGAACACGCAUACAUGUUUGAGAGUACUC
BE659175,BE661028
3GUGCAGGUCGGGAACUGCUUCUUCACUCACGGAAAGAAAAUAUGUGCCUUUACAGAGAAGC
BE461111
3GUGCAGGUCGGGACCUCAUUCGGCGGCGCCGGGAUAUAUGCCGGACGAACGACGGCGGUGL
CB629782,CB655478,CB662543,CB667935
BU811772,BU871538,BU888204,BU889659,BU893331
3GUGCAGGUCGGGAACUGAUUCGGCGAUUUUGAUUGCCAGCUGGCUGGACAUGACUGGUUGUL
CA156827,CA173899
3GUGCAGAUCCGGGACCCGCCGCCCGGCCGACGGGACGGAUCCCGCCUUGCACCAAGUGAAUC
3GUGCAGGUCGGGACCUACCUCGCCGGCACAAUGGCGGUAGCUGACGGCGACGCCAGCGUAC
CK249947
3GUGCAGGUCGGGACCUACCUCGCCGGCAACAAUGGCGGUAGCUGACGGCGACGGCGAGCUUA

CD205823,CD209338,CD209437,CD210663,CD213423,CN133168,CN145907,CD207772

;GUGCAGGUCGGGAACCGACUUCGCCGCUCCGGCAGCGCCGGAGGCACGCGGCGGCCUACG/
BG842539

3GUGCAGAUCCGGGACCCGCCGCCCGGCCGACGGGACGGAUCCCGCCUUGCACCAAGUGAAUC
3GUGCAGAUCCGGGACCCGCCGCCCGGCCGACGGGACGGAUCCCGCCUUGCACCAAGCGAAUC
,GGAUGACUUGCCGGCAUUAGCCAAGUGAAUGAGCAUCAUAUAUAUAUAUAUAUAUAUAUAUA
BM524615,BM520766

,GGAUGACUUGCCGGCGUUAUUUAUUUGCUCUAUGUUCUAUGCUCACCGGUUUCCUUGCCGGCAAC/
,GGAUGACUUGCCGAUCGAUCGAUCUAUCUAUGAAGCUAAGCUAGCUGGCCAUUGGAUCCAUC
,GGAUGACUUGCCGGCUCUCCUGUAUUUGGGGAAUCUCAGCUUUGCUGAAGCGCCUUGGAGUU
CF316749

,GGAUGACUUGCCGAUCGAUCUAUCUAUGAAGCUAAGCUAGCUGGCCAUUGGAUCCAUCUCA
,AGGAGACUGCCCAUGAACCGACUUAAGGAUCAUUAUUUGCUAAUCCUUCAGGGAGGACAC/
,AGGAGACUGCCUUAUCUCGGCUCUAUCUGAACCGAGCCAGGCCACACUGAUCAUUGGCACUGCAUC
,AGAAUGGCUUGCCUUAUCUCCACUAUUUGGUUCAUCACUGGAACCCACUUGGGGUUCUCCGAL
,GGAUGACUUGCCGACGACUCGCUUUUGCUUUUGCUUCCAUAUAUAGGCAUAAUCAAGAAG/
,GGAUGAUUUUGCCGACGACUCAGUUUUUGCUUCCAUAUGGUAGGAGAGAUGAAGAGGUUGGC/
,AGGAUGAUUUUGCCUGUGAAACUCCCUUGGCAGCCGAGCUCUCUGCCACAGAGAGCGGCGUCC
,AGGACAGACUUGCCGACUGAGUUCUAAGCCUCAGCAGCAAGCUGAGACGCCUCCAGGUUCCA
IGGCCUGGUUCACUCAGAUUUUACCUGACCACACACGUAGAUUAACAUAUUUCUCUCUAGAUU/
;GUCCGGUUCAAUAAGAAAGCAAUGCUCAAAAUGUUUUUGGGUCCUGUUUUUUCAUUGAGCCG
GUGCGGUUCAAUGGUAACGGUUGUGGUCGUAACAUAAGACCCCAUUUUUUCGAUUGAGCCG
;GCUCGACUCACUCAGACCACGCCGGAGGGAGCCAUCUCGCGCGGCGGUUCUGAUUGAGCCG
JGGCAUGGUUCAAUCAAACCGGGCAAACUUAUGCACUAGCUAAGCAAGAUGCAGGGAUCUGCA
;GCUCGGCUCACUCAGACGGCAUUGGCGUGAUGCAAAGCAUGCAUGCGUGCUUGCUAGCUCA/
;GUGCGGUUCAAUCAGAGAGCUGGCGCCCCAGGAGGCAAGGGGUUCCUCCCUUCGAUUGAGC
JGGCUCGACUCACUCAGACCACGCCUGCCGGCCGGCCGUAGCCAUGCAUCUGCAUGCGGUGG
JGGCAUGGCUCAAUCAAACUCGCCGGCCGCGGGUGGCUUAGCUUAUAUAUUCUGCGUUUUUGA
;GCAUGGCUCAAUCAAACUCGCCGGCCGCGGGUGGCUUAGCUUAUAUAUUCUGCGCGUUCGAU
;GUCCUCAAGAUUCACAUCACAGUCUAAAGGCAAAGCAGCAAUUUUUUCUUAUUUUUGCUUGC/
;AUCAUCAAGAUUCACAUGCAAUAAGGUGGGUGGGACUAUGAUGCAAUCCAAGUGCUCUGC
;CAUCAUCAAGAUUCACAUGCAAAGCGCAGGUGGUGGGUGGGACUUGAUGCAAUCUAAGUGCUC
;AUAUAUCAAGAUUCACGUGAAAAGUUGCAAUAUUGGUUAUAUAUAUUGAUGAAAUAUUGGCUGG/
;ACCAUCAAGAUUCACAUAGAAAUAUUGGACUAUGAAAUAUAGCCCAUUUUUUGAAUACA/
;GGAUCGCAUUGAUCCUUCAUUCGCUUCGCUUCGCUUCCAUGGCGGUCGUCGUCUACAAGCAG/
;GGAUCGCAUUGAUCUGGCUAGCUAUCUCGAUCGAUCGCCUCAUCGAUCGACGACGACGUGC/
;GGAUCGCAUUGAUCUAAUGACUUCGAUGUCUAUAUGAUGUUAUUGUUUAGUCAUUUCAUUG
JUCUGUCCACCUCACUUCUUGGCCCUAUCUACGUACUCGGAGGUGGAUAUACUGCCAAUAGA
BU082875

JUCUGUCCACCUCACUUCUUCAUCCUCCUGUUUCUGUUAUUGGAUCUCUCUACCUAAUAUGUG
JCCUCUGAACGCUUCAUGUGAUUGGCUAGUUAUAGGCCUUGAUGAAGUGUUUGGGGGAACU
;CUUCAAGCACUUCACAUGGCCCUAUUUCAAUGUCUAAUAUGUGAAGUGUUUGGGGGAACUCU
;CUUCAAGCACUUCACGUGGCACUAUUUCAAUGCGUACCGUGUGAAGUGUUUGGGGGAACUCU
CUUUAAGCACUUCAUACGACACCAUUAUUUGUUAAGGGUUGUUAUGAAGUGUUUGGAGGAACUC/
;UCCUCAAUCCACUUCAGUAGAUAGCUAUGGCUAGGCCUCAUUGCAUUGCACUGUUACAUAAC/
;UUUCAAGCACUUUACGACACACCGUAUUAGAGAUUGUCGUGAAGUGUUUGGGGGAACUCUUA
CUUUAAGCACUUCAUACGACACCAUUAUUUAUAGGGUUGUUGUGAAGUGUUUGGAGGAACUCI
;AUUCAAGCACUUCACGUGGCACUAUCUCAAUGCCUACUAUGUGAAGUGUUUGGGGGAACUCU
;CUUACAAGCACUUCACUAGGCAUUGAAUUGCCAAUGUGAAGUGUUUGGAGGGACUCUUAUG
;UUGCAAGCACUUCAUAGGCAUUAUUUGAGAUGCCACUAUGAAGUGUUUGGGGGAACUCUUG
;AGCUUUCUUGAACUUUUUUUCAUUUCCCUUAUUUUASAGCGAAAUUAAUAACUAAAAUCI
BM521827
CA853579

.GCUUUCUUGAACUGCAUCUUCCCUAAUUUUUUUAUCUUUUGUUUUUCCUCUCUCAGAUCCAAI
BM302147,BM302148,BM302149
.GCUUUCUUGAACUGCAUCUUCCCUAAUUUUUUUAUCUUUUGUUUUUCCUCUCUCAGAUCCAAI
AGCUUUCUUGAACUUUCCCAAUGAUGGUUUGUUUCUCACUAGAAAGAAAAAAAAGAAGAAAAI
.GCUUUCUUGAACUGCACGCAUGAUGAAUAAUCCCUUUGGUUAAUUGUGAUCUGGUCUCUGAG
.GCUUUCUUGAACUGCACCUAUUAGAUUUUAGUUGAUGUUGUUGUGCGAUUUGCCAUGACCAU
.GCUUUCUUGAACUGCACCUAUUAGAUUUUAGUUGAUGUUGUUGUGCGAUUUGCCAUGACCAU
BU046219
.GCUUUCUUGAACUGCAUCCAUGAGAUCGAUCGAUCUUUGCAUGUGAUGCUGCAGUCACUCAC
.GCUUUCUUGAACUGCAUCUCUAAGAGGAGCAGCUCGAAGCCUCGAACUCUACCUGCAUGAGC
BQ509712
;AGCUUUCUUGAACUGCAUCUUUCAGAGGAGCGGCAGUUUCAACUCCUCCACCCGCAUCAGCA
.GCUUUCUUGAACUGCAUCAUGCAUGCAGCAGGCUGUGCUGUGGACCUGAUCGAGUUUCAUL
BQ461815,BU966762,CA014459,BM097863,BQ665274,BU966899
AGUGCAGCGUUGAUGAACC GUCCGGCCAUGGNCCGUCCGCCUCCACCGAGGCCGGAGCGGUL
;UGAACCUGAGAACAGAGGGUGGCGUUGGCUUAAAUUUGAUUUGCAUUGCUGCUGCCUGCUC(
ICGUCCUGAGACCACAUGAAACAGAUUCAAAAUACAAGCAUAAUUGCUUGUGACCUUUUGUGAC
;UGGAUCUGAGAACACAAGGCUGGUUUGCACUGCUAUAAUUGAUCGAUUGGUAAUAAGGUGAA
;UGAGAACACAAGGCUGGUUUGCACUGCUAUUAUCAUCUAAUUGGUAAUAAGGUGAAUUUACUUUG
IGACAUGAGAACACAUGAUGAUAUACUUAGAUUUUUCACGUGUUCUCAGGUCGCCCCUGCUG,
;CGACAUGGGAACACAUGUUAAAUGUGCAACAAAUCACAUUCCGCCAUGUGUUCUCAGGUCGC(
;CGACAUGGCAACACAUGUUAAAUGUGCAACAAAUCACAUUCCGCCAUGUGUUCUCAGGUCGC(
;UGAGCCUGAGAACACAAGCUGAAUUGGUUUGAAUUGCCAUAUCACAUACUGAUUUCUGGUAAU/
BG585080
JCGACAUGAGAGCACAUGAAGCUAUCAUGGUUGUCUAUGUUUAUCCAACUCAUGUGUUCUCAGG
;CAACCUGAGAUCACAUUUUGUCAUUUUUUCUUUAGUUGUUGAGUCUGGUUCAUAUCUCACUAI
;GAGCUGGGAACACACGGUGAUGAGGCGGUCUGGUCUUUCGUGUGUUCUCAGGUCGCCCCUC
CGGACUGGGAACACAUGGGAAUGAGAUGAGAUCAUUGCUCGGUCGUGCUGGCCUGGGCCGUC
AJ502806,AJ503124,AJ502406,AW328877
CK094874

UCUUUCUGUCAGAUUCUAGU
AUUGUGCAUCUUACUCCUUUGUGCUCUCUAUACUUCUGUCAUCACCUUCAG
UCUAGUCUUCUGUCAUCAUCCUUCU
CUUCUCUAUCUGUCAGCUUCCAU
JCUCUCCUCUGUCAUCAUCAU
CUCUAUGCUUCUGUCAUUACCUAUUA
UCUCCCUUCUUCUGUCAUCAUCAUGA
AAUUGCAUUUCAUUUCCUUUGUGCUCUCUAGACUUCUGUCAUCACCUCAUC
UCAAUUAAACUGCAGAGAAUAAAAAGCAUCUCAAUUCAUUUGUGCUCUCUAUGCUUCCGUCAUCAC
GCUCACUCUCUAUCUGUCAGCCCGUUA

AGCUAUGCGUGCUCACUUCUCUCUCUGUCAGCCGUUAGA
UCACUCUCUAUCUGUCAGCCACUCAUC
JCUUGCUUCUGCCAACGACCGUCU

ACUUCUCAUUCUGUCAGCUCUCUA
ACUUCUCUUUCGGUCAGCCUCUUCUG
UUGUAAGCGGCAUUCUCUUGGAUUGUAAUCUGAAUUCUGCCUCUAUCAUCAACCUGCCCACAAACG/
AAGCUAUGCGUGCUCACUUCUCUCUCUGUCAGCCAUUAGAA

AAGCUAUGCGUGCUCACUUCUCUCUCUGUCAGCCAUUAGAA
CACUCUCUAUCUGUCAACCACUCAU
UCACUCUCUAUCUGUCAGCCACUCAUC
ACUUCUCUUUCUGUCAGCCUCUUUC
AUGCUGAGUGAAAGCGAAGAAAGAUACUCAGCAAUAGUGAAUGAUGCGGGAGACAAUUGAUUCI
ACAUUCAUUCUAACAUAAGUAUUCAAUUAGGGUAAUUAUUGUGUGAAUGAAGCGGGAGUAUAUAGUAL
UUCUACCCAUCAUAAGUGCUUUUGUGGUAGUCUUGUGGCCUCCAUAUCUGGGGAGCUUCAUUUG
JCUGACUUCUUAACAACAUGCGUUCGAAGUCAAGGGUUGCAUGCCCUGGGAGAUGAGUUUACCUU

IGAACUCAUCCAACGGCUAGGAUUCGGAUGUGUUUUUGAUCCAACGAUGCGGGAGCCGUGUUUGGU
ACUCUACCCAUCUUGUUCUUUGUGGUAGUUUUUGUGGCCUCCAUAUCUAGGGAGCCUUUACCCUUU

UUAGACUACUCCAUGUUCAUCAUUCAGCUCGAGAUCUGAAAGAAACUACUCCAUUUUUACUAAUA/
ACUAUCCUAUCUCAUGGCUGGAGGUUUAUCUGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGUGGUCGGI
GGGUUCUUUAUCGGGGGCAGGGCCCCGAUACGGUAUCCGAACGUUUGUCCCGGGAACUGGUCGAC

CUAUCCUAUCUCAUCAUGUAUGUGUGUAUGUAUUUUUCGAGAGGGAGGAGAGGAGCUAGACUCUCAL
CUAUCCUAUCUCAUCAUGUAUGUGUGUAUGUAUUUUUCGAGAGGGAGGAGAGGAGCUAGACUCUC/
CAUGCAAGUGCAGGUGAGCAUUAGUCAUUGCUGCACCAGAGAGGCAUCCAUGAACCGGCAGCUGC
UCUGGUGCAGCAAUGACUAAUGCUCACCUGCACUUGCAUGGGUGUGCAUGACCCGGGAGAUGAAC/
CUAUCCUAUCUCAUCAUGUAUGUGUGUAUGUAUUUUUCGAGAGGGAGGAGAGGAGCUAGACUCUCAC

CUAUCCUAUCUCAUCAUGUAUCUGUGUAUGUACUCUAGAGGGCCGGAGAAGAGAUUCAUGUGGUCC

ACACAGUAGAUUGAGACUAGUCCAGGGCAUACCAGUAUGUUACAAUUGUACUGUGCGAAUGAGC/
UUAGCCUACUCCAUGUUCAUCAUUCAGCUCGAGAUCUGAAAGAAACUACUCCAUUUUUACUAAUA/
ACUAUCCUAUCUCAUAGAAAACGAGGAGAGAGGCCUGUGGUUUGCAUGACCGAGGAGCCGCUUCC
GCCCUAUCCCGUCAUUGGGGGGCAUUGGAUAGGCUUGUGGCCUUGCAUAUCUCAGGAGCUGCAUUAI
UGGACUGAAGGGAGCUCCUAGUUCUUCUC

3GCGUGCACGGAUCCAAGCAUAUCGAACCC
3CAAGGAGCCAAGCAUGCAUAACAG
JCUGCAUCCAGCGCUCACUU
;GAUAAACCUCUGCAUCCAGCGCUCACUUU
JUGAUAAUAUACAACGCAGGAUUCGAUCGAUAAGCCUCUGCAUCCAGAUCUCACUU
AAACCUCUGCAUCCAGCGUUCACUC
IAUGACUCUUACCUACAAUAGAUUUUGUUGGCCUUCAGCGAGUUAGUUCUUCAUGUGCCUGUCUUCC(
IAUGACUCUUACCUACUUAUAGAUUUUGUUGGCCUCAGUGAGUUAGUUCUUCAUGUGCCUGUCUUCC(
CGGCCUUGCCUGGCUGGGUAGUGCGUCGCUCGAUCCGGCCGUGCGCCGGCGGCCGCCCUUGCAI
AGAUUCAGAUGUAAUAAUAAUUUGUAAUAUAUGCAUAUUAUGUCUCCUCGGACCAGGCUUCAUUC(C
CAAGGAAUACAUAGUGUUGUCGGACCAGGCUUCAUUC(CCCCCCAAUUAUAI
UAUGGGUUUCGAUCUCUUUGAUCCCUUGAAACAAAGAAAGCUUUAAGGUUGGAUUUUGAGGCUUU
GACCAGGCUUCAUUC(CCCCUCAAAAAU

JAAGGCUUCGACCAGGCUUCAUUCCCCUCAAAAUCU
AAAGUAUUCUCGGACCAGGCUUCAUUCCCCCAUUUAU
AUUUUAUGUUCAUACAUCUAUAUAUAUAUCUAUAUAUAUAUCAAGUCAUUACAUGCAUGGGAUGAGA
JGGUGAUCUCGGACCAGGCUUCAUCCCUUUAACCAGG
UUCUUGAAUCUAUGAGAUCUCGGACCAGGCUUCAUUCCCCUCAGCAUAG
CGGACCAGGCUUCAUUCCCCUCAAGUGGAG

CCUAGUUGU
GACUGUCACGGAUCCAUGAAUCCAGGAUGAGGGGAGGGAAGAAAGAGGGAUAAUGAGCAUCAGGU
UGACUGUCACCGAUCCAUGAAUCCAGGAUGAGGGGAGGGAAGAAAGAGGGAUAAUGAGCAUCAGGU

AACGUUUUUUAUUUUUAUAAUUUUUUGAACGACGUUUCUGGUGAAUUUAAGUUCAUGCCUUGCAUCA
 JUAAUUCUACUAAAGCUGUCACCGACGGAUAGAUGUUUGAUUAGCGGCGAAAUUUGGGUCCUGCCU

;GGAGCC
 ;CGGUAAAAACUAAUUUUUUACAUGUUGUCUGUGGCGUAGUUUGGGUCCCGCCUUGCAUCAAACUGA/
 ;CCGGUAAAAACUUUUUUUUUUUUACAUGUCUGUGGCGUAGUUUGGGUCCCGCCUUGCAUCAAACUGA/

AUUGGUUGCUGAGCGAAUUCCGAUCCCGCCUUGCAUCAACUGAAUCGGAGAC

GGAGCC

GGAGCC

IAUAUAUGACUCAUGUUCUUGUCGGCAAGUUGGCCUUGGCUAUUUGGACU

GUUGUGUUUGGCUAUGUUUUGCU

UAUCAAUUGGCAAGUUGUUCUUGGCUACAUCUUGGC

IAGCCGGCAAGUCUGUCCUUGGCUACACCUAGCU

AUUGGCAAGUUGUUCUUGGCUACAUCUUGGC

AGGCAAGUCAUCCUUGGCUAUCAGAGAUAA

AUUCAGAAGAGCACAAUAGGCAAGUCAUCCUUGGCUACCAGAGGCAG

JGGUGGAUGAAUAUGGAAGAUGGUGAGCCUUCAUGGCUAAGAGAGUGAU

AGAUGAAUCCGUUGGCAGGUUGUUCUUGGCUACAUUUUUCU

AGGUUUUCCUUGGCUACAUUUUCCU

GGCGGUUCCAUGGGCAAGUCACCCUGGGCUACCCGAAGUAC

GGAGCCGGCAAGUCAUCCUUGGCUGCAUCCCGUUC

AUCUGAUUGAGCCGCGCCAAUAUCUCAGUACUCU

UGCCAAUAUCACGAACCAC

UGCCAAUAUCACGUACUUA

UGCCAAUAUCUUAGUGCUC

UAUGGUUUUGUUGGUCUGAUUGAGCCGUGCCAAUAUCACAAGCUUGC

CUUGUGUUUCUGAUUGAGCCGUGCCAAUAUCUUAGUGCUC

CGUGCCAAUAUCACGCGGUUU

UGGCUCUGAUUGAGCCGUGCCAAUAUCUCAGUGCUCU

UCGAGGUGCGGGCGCAGUGUUUAAUUGAUUGAGCCGUGCCAAUAUCACAACCUUCU

CGAGGUGCGGGCGCAGUGUUUAAUUGAUUGAGCCGUGCCAAUAUCACAACCUUC

CUUGGUUUUUGUCAGUGAGAAUCUUGAUGAUGCUGCAACGGCGAUUA

CAAUCCAUCGGUCUUUUUGAUGUGAGAAUCUUGAUGAUGCUGCAUCAGCCAUAA

UGCCAGCCAAGCCAUAGGUCUUUUGGAACUGAGAAUCUUGAUGAUGCUGCAUCAGCCAUAAA

CUAUUUGAAACUCACGAGAAUCUUGAUGAUGCUGCAUCAGCAAUAA

UGAGAAUCUUGAUGAUGCUGCAUUGGCCAAAUU

CUUGACGGAUCAUGCGAUCCUUUUGGAGGCUUCCUC

UGAUCGAUCAGUGCAAUCCCUUUGGAAUUUCCUC

GAUCAUGCGAUCCCUUAGGAAUUUCCAG

GCUGU

GAGGUGGGCAUACUGCCAACAGAGCUGU

CUUAGGUUCAAC

IUGGUAUCG

JUGGCAUCC

UCAGUGCCA

UGUGAUCAUGGGGCCAAAAGCUAGCUAUGUACAGUGAAGUGCUUGGGGGAACUCCAGUUGACAC

AGUGUCG

UCAGUGUCA

JCGGUAUCA

GC AU

UGAUG

UCUAACAUUUAACACUCUASAAAAAAAAAGCUCAASAAAGCUGUGGGAAAACAUGACA

UUUUUUCUUUGGUAAAUACUUUAUAAUUUAUCAAAAAAAAAAAAAAGAUGGAUUUUUGGGAAUUGAAAGAI

UUUUUUCUUUGGUAAAUACUUUAUAAUUUAUCAAAAAAAAAAAAAAGAUGGAUUUUUGGGAAUUGAAAG,
GAACCGGAAAGUUCAAGAAAGCUGUGGAAAAGCAUGGCA
;AGAUCGUAGCUAGACUCGAUCGGUUGCAUUGGCAUCAGAGAGAGCAGUUCAAUAAAGCUGUGGGA,
IAUGACAUUGUAUUCAUUUUUGCUGCGGUUCAUAAAGCUGUGGGAAGAUACAAA
IAUGACAUUGUAUUCAUUUUUGCUGCGGUUCAUAAAGCUGUGGGAAGAUACAAA

:UCACUCUCUCUCUAACUGGCUUGCGGGUCAUAAAGCUGUGGGAAGAUACAGA
:AGGUGCAGUUCAUAAAGCUGUGGGAACUGCAGA

,GGUGCAUGCAGUUCAAUAAAGCUGUGGGAAACUGGAAAG
JGAUCCAAGCAAGCAAGAGGGCAGUUCAUAAAGCUGUGGGAAAUUGCAGA

JCACCGGCGCUGCACGCAAUGACGCCUCUGC
CUGCCAUACAUAUAAUCAAAGAUUGUAUGUGGCCAACGCACCUUGUGUUCUCAGGUCACCCCUU
CUCAGUUCAUGUGUUCUCAGGUCGCCCCUGCUGAACUUUU
\UUUACUUUGUGUUCUCAGGUCACCCCUUUGAGCCAACC
;UGUUCUCAGGUCACCCCUUUGAGCCAACC
AAUUUU
CCCUGCAGGGUUUUU
CCCUGCAGGGUUUUU
AGGCUUUAUGUUGCUAUUUUUUUGUGUUCUCAGGUCACCCCUUUGAGCUGACC

;UCGCCCCUGCUGAAUUUUC
UGUGUUCUCAGGUCGCCCCUGUCGAAUUAUU
3CCGGGACUC
CGGCGCGCGUUGAUCUUGCAUGUGUUCUCAGGUCGCCCCCGGAGGGCCUU

CUUCACC

AGUUCCUUCAGCUGAGUGCCUUUCCGGCUUGAGCCUUCUGCAUGAUCAGCUGAGUUCUUUCUGCG

UUAAGUGUCCUGUACUUGGACUGAAGGGAGCUCCCUUUUUCUUUU
JCUAUUUGCAACCCUCUUUCUCUGUGCUUGGACUGAAAGGAGCUCCUUCUUUUUCUG
;CCUUUAUAGUAUUAACCUUCUUUGGAUUGAAGGGAGCUCUACACCCUUCU
IGAUCUUUUGGUUUGGAGUGAAGGGAGCUCCAGAGGGUAUUC

IUCUGUCUGUCUCGUCUUGGACUGAAGGGAGCUCCCUUCUGUCCAC
IAGUUUAAUCUUUCUUUGGAUUGAAGGGAGCUCUACAUCUUGCU

GUAUGUGUGUAGAUAGGAAAUGAUGGAGUACUCGUUGGGAUAGGCUUAUGGCUUGCAUGCCCCA
UCUGGUGUUGGCUUGAGAUAGGCUUGUGGCUUGCAUGACCGAGGAGCUGCACCGUCCCUUGCUG
;CUUCCGCCCGGCGUCUCUUGGACUGAAGGGAGCUCCACUCUUGGCGU

JGGGGGUCGUCUUGAGAUACGCUUGUGGUUUGCAUGACCGAUGAGCUGCACCGUCCCUUGCUG
ACGGUGGUCGUCUUUGAGAUAGGCUUGUGGUUUGCAUGACCGAGGAGCUGCACCGUCCCUUGCU
:AACCGACCACUUCCCUCCUGGAUUGGAUCGAAGGGAGCUCUUCGAUCACUUU
CCACCAUCAUCUUUCCCUUGGCUUGGAUUGAAGGGAGCUCCUCUCUCUUUC
;GGUGGUCGUCUUUGAGAUAGGCUUGUGGUUUGCAUGACCGAAGAGCUGCACCGUCCCUUGCUG

;UCAGUCUUUGAGAUAGGCUUGUGGUUUGCAUGACCGAGGAGCUGCACCGUCCCUUGCUGGCCG

GAAUGCAGCGGGAGAUUGUUCUCUCUUUCCUCCUCCAUGCUUGGACUGAAGGGAGCUCCCUCAUC
GUAUGUGUGUAGAUAGGAAAUGAUGGAGUACUCGUUGGUUGGGAUAGGCUUAUGGCUUGCAUGCC
;AUCCUCGCUGACCGCUGUUUGGAUUGAAGGGAGCUCUGCAUCUUGAU
UCCAAGCUUAGAUCUUGUUUGGAUUGAAGGGAGCUCUACACCUCUCU

AGUUUAGAUCUUGUUUGGAUUGAAGGGAGCUCUACACCUCUCU
AGCCCGCCAUAUCUUCUCCUCGUGCUUGGAUUGAAGGGAGCUCUCUCUGUCUG
AGAACCGACCACUUCUCCUCCUGGAUUGGAUCGAAGGGAGCUCUUCGAUCGCCUU

CCAUCAUGAUCA
CCAUCAUGAUCA
UGCAUGUGCCUUUCUUCUCCACCGUGCACA
CUUCAAUUACAG

CUCGGACCAGGCUUCAUUCUCCGUAAACCUUA

CGUUAGUGUCGUCGGACCAGGCUUCAUCCCCCCCCAAUUGUUU
CUUGGAGUGAUGGCAUCGGACCAGGCUUCAUUCUCCCAAACAC

UAUUGAGAGUGAUCUCGGACCAGGCUUCAUUCUAC

CAUGCUGUAGUUUCAUCUGCUGGU
JCAUGCUGUAGUUUCAUCUGCUGGUG

JAGGGCGAAUUGGAUCCCGCCUUGCAUCAACUGAAUCGGAGAC

ACUGAAUUGGAUGA

UGCAUCAACUGAAUUGGAGAC

AACUGAAUCGGAGGC

AUAGGAGAC

AUAGGAGAC

UGGAAAUUUUAAUUUGUGGGUUGCGGUUCAUAAAGCUGUGGGAAGAUACAAA
AUGGAAAUUUUAAUUUGUGGGUUGCGGUUCAUAAAGCUGUGGGAAGAUACAAAU
AAUUGCAGA

UGGGAUUUCA

CCUUUCAUUGUGUCCUG

3GAGCUGCAUCAACCCUACAUGGACCCUCUUUGGAUUGAAGGGAGCUCUGCAUCUUUGG
;GCCGCUCUUUGGAUUGAAGGGAGCUCUGCAUCUUGAU

GCCGCUCUUUGGAUUGAAGGGAGCUCUGCAUCCUGAU
GGCCGCUCUUUGGAUUGAAGGGAGCUCUGCAUCCUGAA

GCCGCUCUUUGGAUUGAAAGGAGCUCUUGCAUCUGAU

CUCUUUGGAUUGAAGGGAGCUCUGCAUCCUGAU

JCUCA
CCAGGAGCUGCAUCAACCCUACAUGGACCCUCUUUGGAUUGAAGGGAGCUCUGCAUCUUUGG

