Supplemental Information

RNA G-Quadruplexes in the model plant species *Arabidopsis thaliana*: Prevalence and possible functional roles

Melissa A. Mullen[1], Kalee J. Olson[1,2], Paul Dallaire[3], François Major[3],

Sarah M. Assmann[2*], and Philip C. Bevilacqua[1*]

[1]Department of Chemistry, [2]Department of Biology,
Pennsylvania State University, University Park, Pennsylvania 16802-5302
[3]Institute for Research in Immunology and Cancer (IRIC), Department of Computer Science and Operations Research, Université de Montréal, PO Box 6128, Downtown Station, Montréal, Québec H3C 3J7, Canada

Contents:

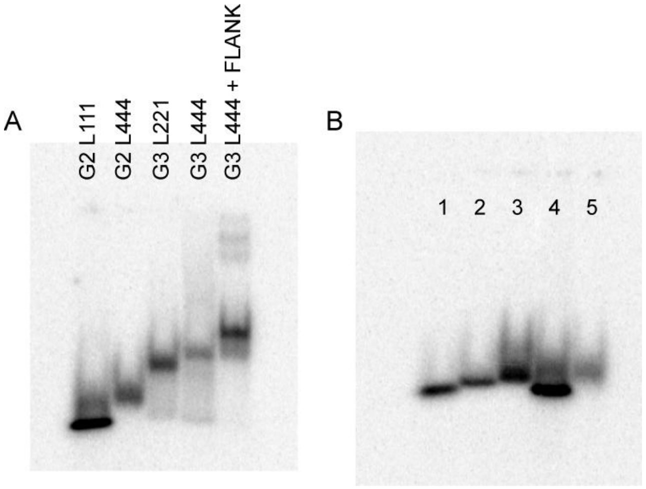

Figure S1. Native gels of GQS RNA oligonucleotides. A. RNA GQS oligonucleotides of interest, renatured identically as for Circular Dichroism samples, and run in 100 mM K^+. B. RNA GQS 'ladder' oligonucleotides. Lane 1: AG$_2$A, Lane 2: A(G$_2$A)$_2$, Lane 3: A(G$_2$A)$_3$, Lane 4: A(G$_2$A)$_4$. Oligonucleotides with increasing numbers of G$_2$A repeats until a full GQS is achieved, as indicated by the faster migration of the oligonucleotide (seen in Lane 4). Lane 5: AG$_2$AGAAG$_2$AG$_2$A. Mutant version of the oligonucleotide in lane 4. Disruption of the G pattern disrupts GQS formation, resulting in slower migration.

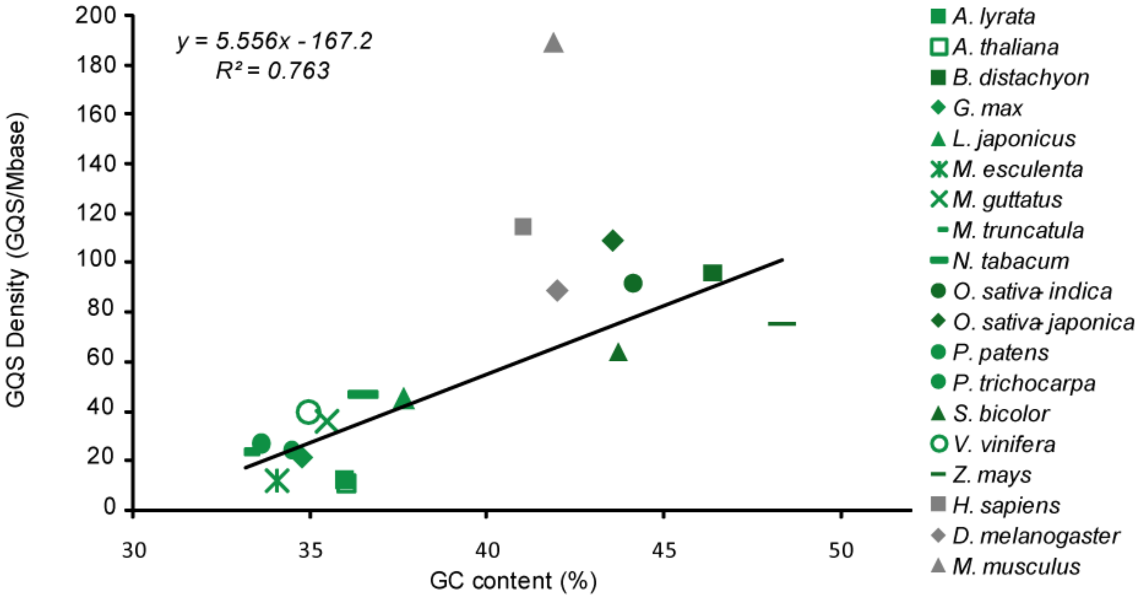

Figure S2. G$_3$L$_{1-7}$ GQS density as a function of GC content. 15 plant species and 3 non-plant eukaryotes were analyzed. All plant species are in green and fit to a linear regression curve. Non-plant eukaryotes are shown in gray. GC content was determined as described in materials and methods.

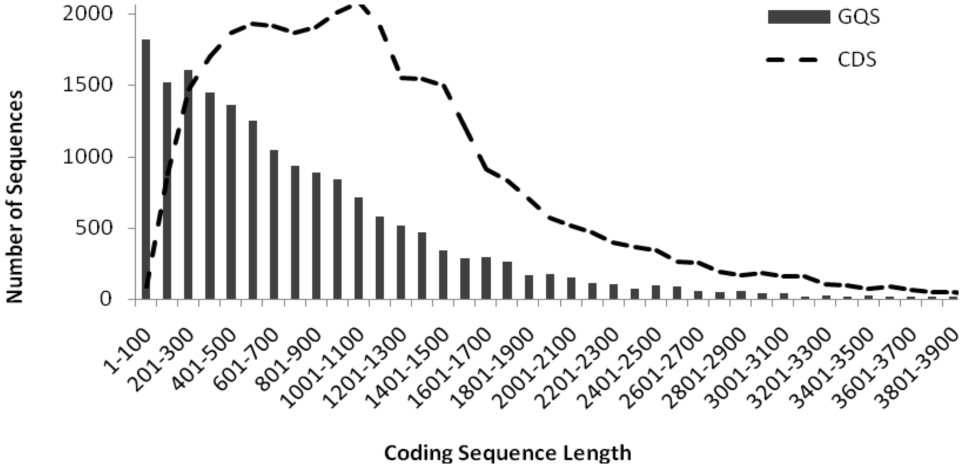

Figure S3. Distribution of G_2L_{1-4} GQS in the coding sequence. The number of GQS starting positions that fall within each 100 nt segment of CDS is shown above with bars. The plot is cropped at a CDS length of 4,000; it incorporates 17,497 GQS out of a total of 17,619 GQS (>99%) that are found in the RNA CDS. The dashed black line represents the distribution of the length of all coding sequences from *Arabidopsis*.

Table S1. Genome sequence information

Species[a]	**Common Name**[b]	**Assembly**[c]	**Release Date**[d]	**Server**[e]
Arabidopsis thaliana (57)	mouseear cress	TAIR9	06/09	TAIR
Arabidopsis lyrata	lyrate rockcress	Araly1	12/08	JGI[f]
Brachypodium distachyon (58)	a wild grass	Brachy1.0	01/09	Ensembl
Drosophila melanogaster (59)	fruit fly	BDGP, 5	3/06	Ensembl
Glycine max (60)	soybean	v 1.0	12/08	JGI
Lotus japonicus (61)	a model legume	v 1.0	05/09	KDRI
Manihot esculenta	cassava	v. 1.0		JGI[f]
Medicago truncatula (62,63)	barrelclover	v 3.0	12/09	MSGC
Mimulus guttatus	seep monkeyflower	v 1.0	01/10	JGI[f]
Mus musculus (64)	mouse	NCBI m37	04/07	Ensembl
Nicotiana tabacum (65)	tobacco		11/08	PNGG
Oryza sativa – indica (66,67)	rice – indica	2005-01-BGI	01/05	Ensembl
Oryza sativa – japonica (68,69)	rice - japonica	MSU6	01/09	Ensembl
Populus trichocarpa (70)	black cottonwood	JGI2.0	12/04	Ensembl
Pyscomitrella patens (71)	moss	v 1.1	03/07	JGI
Sorghum bicolor (72)	sorghum	Sbi1	12/07	Ensembl
Vitis vinifera (73)	wine grape	IGGP_12x	12/07	Ensembl
Zea mays (74)	corn	AGP1 4a.53	03/09	MaizeSeq

Genome sequence information for all downloaded sequences. Provided for each are the [a]official species name, including subspecies for *O. sativa*, [b]accepted common name, [c]genome assembly [d]release date for the assembly, and [e]server where the genomes were obtained. [f]These sequence data were produced by the US Department of Energy Joint Genome Institute (http://www.jgi.doe.gov) in collaboration with the user community.

Table S2. Unique gene models with a GQS in the RNA

GQS Motif	Genic		Coding		5’ UTR		3’ UTR		Intron	
	Gene models	*Loci6*	*Gene models*	*Loci*	*Gene models*	*Loci*	*Gene models*	*Loci*	*Gene models*	*Loci*
$G_{3+} L_{1-7}$	215	177	166	142	10	10	14	11	26	19
$G_{3+} L_{1-3}$	43	38	28	26	4	4	4	4	7	5
$G_{2+} L_{1-4}$	13,156	10,382	11,816	9,699	371	310	588	473	782	630
$G_{2+} L_{1-2}$	4,026	3,185	3,610	2,932	119	100	123	102	211	165
$G_{2+} L_{1}$	2,428	1,913	2,190	1,781	70	60	62	54	114	183

Number of unique gene models and loci with at least one GQS. Genic, Coding, 5’ UTR, 3’ UTR and Intron are defined in Table 4. Quadparser search parameters were set to include G-patterns, which will be found in RNA, and exclude C-patterns. In Arabidopsis, there are 39,640 gene models and 33,518 loci, of which 27,379 are protein coding. The non-coding loci correspond to rRNAs, tRNAs, and other annotated ncRNAs (47). Note that GQS from gene models were used for Tables 1-5.

Table S4. Functional analysis of genes with at least one G_2L_{1-4} GQS present in the RNA

	GO ID[a]	GO Cat[b]	GO term[c]	GQS genes[d]	All genes[e]	% GQS genes[f]	p-value[g]
overrep	0003824	MF	catalytic activity	2894[h]	6393[i]	45%	9E-65
	0006468	BP	protein aa phosphorylation	506	798	63%	1E-53
	0016310	BP	phosphorylation	528	852	62%	1E-51
	0016740	MF	transferase activity	1118	2176	51%	2E-49
	0006796	BP	phosphate metabolic process	543	894	61%	2E-29
	0006793	BP	phosphorus metabolic process	543	895	61%	3E-49
	0016301	MF	kinase activity	661	1151	57%	2E-48
	0043687	BP	post-translational protein mod.	579	985	59%	2E-46
	0016772	MF	transferase act, phos. groups	708	1284	55%	2E-43
	0006464	BP	protein modification process	612	1107	55%	1E-37
	0043412	BP	biopolymer modification	639	1229	52%	7E-29
	0005524	MF	ATP binding	337	607	56%	2E-20
	0032559	MF	adenyl ribonucleotide binding	339	613	55%	3E-20
	0030554	MF	adenyl nucleotide binding	351	642	55%	7E-20
	0005478	MF	transporter activity	504	993	51%	1E-19
	0032501	BP	multicellular organismal process	462	906	51%	2E-18
	0017111	MF	nucleoside-triphosphatase act.	237	408	58%	2E-17
	0000166	MF	nucleotide binding	490	980	50%	2E-17
	0048856	BP	anatomical structure dev.	359	681	53%	5E-17
	0032502	BP	developmental process	493	992	50%	6E-17
	0007275	BP	multicellular organismal dev.	441	871	51%	7E-17
	0015646	MF	transmembrane transporter act.	366	701	52%	2E-16
	0032555	MF	purine ribonucleotide binding	387	751	52%	3E-16
	0032443	MF	ribonucleotide binding	387	751	52%	3E-16
	0016020	CC	membrane	1011	2266	45%	3E-16
	0017076	MF	purine nucleotide binding	399	781	51%	4E-16
	0016817	MF	hydrolase act, on acid anhydrides	244	432	57%	5E-16
	0016818	MF	hydrolase act, acid anhyd. wit h P	243	431	56%	7E-16
	0016462	MF	pyrophosphatase activity	242	429	56%	7E-16
	0022804	MF	active transmem. transporter act.	241	427	56%	7E-16
	0022414	BP	reproductive process	275	502	55%	8E-16
	0043227	CC	membrane-bound organelle	1929	4637	42%	1E-15
	0050876	BP	reproduction	277	509	54%	2E-15
	0043231	CC	intracell. mem-bound organelle	1924	4630	42%	2E-15
	0016887	MF	ATPase activity	183	312	59%	4E-14
	0043226	CC	organelle	2092	5111	41%	1E-13
	0043229	CC	intracellular organelle	2091	5110	41%	1E-13
	0042623	MF	ATPase activity, coupled	134	214	63%	3E-13
	0044424	CC	intracellular part	2331	5754	41%	3E-13
	0005622	CC	intracellular	2477	6146	40%	3E-13
	0016787	MF	hydrolase activity	975	2231	44%	7E-13
	0004386	MF	helicase activity	88	126	70%	1E-12
	0044464	CC	cell part	4378	11331	39%	3E-12
	0005623	CC	cell	4378	11331	39%	3E-12
	0003006	BP	reproductive dev. process	226	420	54%	6E-12
	0048608	BP	reproductive structure dev.	226	420	54%	6E-12
	0051234	BP	establishment of localization	588	1279	46%	6E-12
	0022892	MF	substrate-specific transporter act.	328	657	50%	1E-11
	0051179	BP	localization	592	1294	46%	1E-11
	0006810	BP	transport	583	1273	46%	2E-11
	0048731	BP	system development	190	348	55%	9E-11

0048513	BP	organ development	190	348	55%	9E-11
0015290	MF	secondary active transmembrane transporter act.	126	212	60%	3E-10
0022891	MF	substrate-specific transmembrane transporter act.	271	537	51%	3E-10
0008026	MF	ATP-dependent helicase activity	56	76	74%	2E-9
0009857	BP	recog or rejection of self pollen	29	31	94%	2E-9
0008151	BP	cellular process	2642	6706	39%	3E-9
0009791	BP	post-embryonic development	152	275	55%	4E-9
0008037	BP	cell recognition	29	32	91%	1E-8
0015075	MF	ion transmem. transporter act.	201	392	51%	3E-8
0050222	MF	protein kinase activity	225	448	50%	3E-8
005169	BP	response to stimulus	854	2016	42%	6E-8
0005737	CC	cytoplasm	1671	4157	40%	9E-8
0009875	BP	pollen-pistil interaction	29	34	85%	3E-7
0044444	CC	cytoplasmic part	1542	3831	40%	3E-7
0016773	MF	phosphotransferase act., alcohol group as acceptor	262	546	48%	4E-7
0008152	BP	metabolic process	2446	6261	39%	8E-7
0009790	BP	embryonic development	145	277	52%	1E-6
0003674	MF	molecular function	8070	21958	37%	2E-6
0016932	MF	transferase act., transferring glycosyl groups	195	395	50%	2E-6
0006950	BP	response to stress	526	1210	44%	2E-6
0015293	MF	symporter activity	61	96	64%	2E-6
0003677	MF	DNA binding	717	1699	42%	3E-6
0050791	BP	regulation of biological process	891	2149	42%	3E-6
0048519	BP	negative reg. of biological process	80	136	59%	3E-6
0015294	MF	solute:cation symporter activity	60	95	63%	4E-6
0005634	CC	nucleus	642	1510	43%	4E-6
0065007	BP	biological regulation	955	2320	41%	4E-6
0048316	BP	seed development	144	280	51%	5E-6
0005402	MF	cation:sugar symporter activity	49	74	66%	6E-6
0005403	MF	sugar:hydrogen symporter act.	49	74	66%	6E-6
0015295	MF	solute:hydrogen symporter act.	49	74	66%	6E-6
0046527	MF	glucosyltransferase activity	55	86	64%	6E-6
0051119	MF	sugar transmem. transporter act.	54	85	64%	6E-6
0009793	BP	embryonic development ending in seed dormancy	133	256	52%	6E-6
0051706	BP	multi-organism process	180	367	49%	1E-5
0005515	MF	protein binding	778	1873	42%	1E-5
0009628	BP	response to abiotic stimulus	279	607	46%	1E-5
0048523	BP	negative reg. of cellular process	59	96	62%	2E-5
0015144	MF	carbohydrate transmembrane transporter activity	58	94	62	2E-5
0008324	MF	cation transmem. transporter act.	151	302	50%	2E-5
0005200	CC	struc. constituent of cytoskeleton	22	26	85%	2E-5
0044425	CC	membrane part	406	926	44%	2E-5
0006952	BP	defense response	233	49	47%	3E-5
0048409	BP	flower development	68	116	59%	3E-5
0051244	BP	regulation of cellular process	830	2018	41%	3E-5
0016820	MF	hydrolase activity, acting on acid anhydrides…	72	125	58%	3E-5
0009536	CC	plastid	800	1941	41%	3E-5
0044238	BP	primary metabolic process	2011	5159	39%	4E-5

	0043492	MF	ATPase activity, coupled to movement of substances	71	124	58%	5E-5
	0042626	MF	ATPase activity, coupled to transmembrane movement	71	124	58%	5E-5
	0044237	BP	cellular metabolic process	1994	5120	39%	5E-5
	0044459	CC	plasma membrane part	60	101	59%	6E-5
	0030246	MF	carbohydrate binding	63	108	58%	7E-5
	0006855	BP	multidrug transport	35	51	69%	8E-5
	0008238	MF	exopeptidase activity	49	79	62%	9E-5
	0050973	BP	reg. of developmental process	90	168	54%	9E-5
underrep	0000496	MF	base pairing	0	631	0%	<1E-99
	0000498	MF	base pairing with RNA	0	631	0%	<1E-99
	0000499	MF	base pairing with mRNA	0	631	0%	<1E-99
	0060090	MF	molecular adaptor activity	0	631	0%	<1E-99
	0030533	MF	triplet codon-AA adaptor activity	0	631	0%	<1E-99
	0006455	BP	translational elongation	18	662	0%	4E-98
	0043284	BP	biopolymer biosynthetic process	103	874	12%	6E-59
	0006412	BP	translation	174	1129	15%	4E-51
	0010467	BP	gene expression	293	1487	20%	1E-41
	0009059	BP	macromolecule biosynth. process	280	1399	20%	7E-39
	0003723	MF	RNA binding	179	983	18%	7E-30
	0044249	MF	cellular biosynthetic process	429	1719	25%	3E-22
	0009058	BP	biosynthetic process	602	2149	28%	2E-14
	0000154	BP	rRNA modification	1	70	0.01%	3E-9
	0048046	CC	apoplast	8	102	8%	1E-8
	0005732	CC	small nucleolar ribonucleoprotein complex	8	94	9%	2E-7
	0005576	CC	extracellular region	28	172	16%	2E-6

Provided are [a]overrepresented and underrepresented gene ontology (GO) ID numbers, [b]GO categories (Cat), and [c]GO term for gene products encoded by transcripts with at least one $G_{2+}L_{1-4}$ GQS. Included are [d]the number of genes (CDS, 5'UTR, 3'UTR, and introns) with a GQS that are annotated for the listed GO term, and [e]the total number of genes in Arabidopsis with the listed GO term. Also included are [f]the percentage of genes with GQS with a given GO term and [g]the appropriate p-value, as determined using the BiNGO program. [h]The total number of GO-annotated genes with a GQS in G_2L_{1-4} is 9,097. [i]The total number of GO-annotated genes in *A. thaliana* is 25,179. Table is sorted in order of increasing p-value. Some GO terms are sub-categories of others.

Table S5. Functional analysis of genes with at least one G_2L_{1-2} GQS present in the RNA

	GO ID[a]	GO Cat[b]	GO term[c]	GQS genes[d]	All genes[e]	% GQS genes[f]	p-value[g]
overrep	0003677	MF	DNA binding	297[h]	1699[i]	17%	1E-12
	0030528	BP	transcription regulator activity	247	1441	17%	2E-9
	0019219	BP	reg. of nucleobase, nucleoside, & nucleic acid metabolic process	202	1153	18%	2E-8
	0050791	BP	regulation of biological process	335	2149	16%	3E-8
	0003700	BP	transcription factor activity	208	1207	17%	3E-8
	0009889	BP	regulation of biosynthetic process	202	1164	17%	3E-8
	0045449	BP	regulation of transcription	198	1139	17%	3E-8
	0004386	MF	helicase activity	41	126	33%	3E-8
	0031323	BP	reg. of cellular metabolic process	208	1216	17%	3E-8
	0010556	BP	reg. of macromolecule biosynthetic process	199	1152	17%	3E-8
	0051244	BP	regulation of cellular process	315	2018	16%	3E-8
	0010468	BP	regulation of gene expression	209	1232	17%	5E-8
	0043227	CC	membrane-bounded organelle	641	4636	14%	5E-8
	0060255	BP	regulation of macromolecule metabolic process	200	1171	17%	6E-8
	0043231	CC	intracell mem-bounded organelle	639	4629	14%	6E-8
	0019222	BP	regulation of metabolic process	209	1238	17%	6E-8
	0065007	BP	biological regulation	350	2320	15%	1E-7
	0032501	BP	multicellular organismal process	159	906	18%	5E-7
	0008026	MF	ATP-dependent helicase activity	28	76	37%	5E-7
	0005622	CC	intracellular	809	6145	13%	1E-6
	0043229	CC	intracellular organelle	685	5109	13%	1E-6
	0043226	CC	organelle	685	5110	13%	1E-6
	0016887	MF	ATPase activity	69	312	22%	2E-6
	0007275	BP	multicellular organismal dev	151	871	17%	2E-6
	0005634	CC	nucleus	236	1510	16%	3E-6
	0044424	CC	intracellular part	752	5753	13%	1E-5
	0051234	BP	establishment of localization	202	1279	16%	1E-5
	0006810	BP	transport	201	1273	16%	2E-5
	0048856	BP	anatomical structure dev.	120	681	18%	2E-5
	0051179	BP	localization	203	1294	16%	2E-5
	0042623	MF	ATPase activity, coupled	50	214	23%	2E-5
	0003676	MF	nucleic acid binding	428	3078	14%	2E-5
	0017111	MF	nucleoside-triphosphate activity	80	408	20%	3E-5
	0005488	MF	binding	881	6895	13%	3E-5
	0032502	BP	developmental process	161	992	16%	4E-5
	0001681	MF	hydrolase act, acting on p-containing acid anhydrides	2	431	0.5%	6E-5
	0016817	MF	hydrolase act, acting on acid anhydrides	82	432	19%	7E-5
	0016462	MF	pyrophosphatase activity	81	429	19%	9E-5
underrep	0000496	MF	base pairing	0	631	0%	1E-30
	0000498	MF	base pairing with RNA	0	631	0%	1E-30
	0000499	MF	base pairing with mRNA	0	631	0%	1E-30
	0030533	MF	triplet codon-AA adaptor activity	0	631	0%	1E-30
	0060090	MF	molecular adaptor activity	0	631	0%	1E-30
	0006455	BP	translational elongation	13	662	2%	5E-17

0006416	BP	translation	53	1129	5%	3E-12
0043284	BP	biopolymer biosynthetic process	38	874	4%	2E-10
0010467	BP	gene expression	89	1487	6%	6E-10
0044249	BP	cellular biosynthetic process	109	1719	6%	6E-10
0009059	BP	macromolecule biosynth. process	82	1399	6%	7E-10
0012505	CC	endomembrane system	316	3859	8%	9E-9
0009058	BP	biosynthetic process	154	2149	7%	2E-8

Provided are [a]overrepresented and underrepresented gene ontology (GO) ID numbers, [b]GO categories (Cat), and [c]GO term for gene products encoded by transcripts with at least one $G_{2+}L_{1-4}$ GQS. Included are [d]the number of genes (CDS, 5'UTR, 3'UTR, and introns) with a GQS that are annotated for the listed GO term, and [e]the total number of genes in Arabidopsis with the listed GO term. Also included are [f]the percentage of genes with GQS with a given GO term and [g]the appropriate p-value, as determined using the BiNGO program. [h]The total number of GO-annotated genes with a GQS in G_2L_{1-4} is 2,811. [i]The total number of GO-annotated genes in *A. thaliana* is 25,179. Table is sorted in order of increasing p-value. Some GO terms are sub-categories of others.

Table S6. Functional analysis of genes with at least one G_2L_1 GQS present in the RNA

	GO ID[a]	GO Cat[b]	GO term[c]	GQS genes[d]	All genes[e]	% GQS genes[f]	p-value[g]
overrep	0030528	BP	transcription regulator activity	191[h]	1441[i]	13%	8E-17
	0003677	MF	DNA binding	214	1699	13%	9E-17
	0003700	MF	transcription factor activity	161	1207	13%	2E-14
	0019219	BP	reg. of nucleobase, nucleoside, & nucleic acid metabolic process	150	1153	13%	1E-12
	0009889	BP	regulation of biosynthetic process	151	1164	13%	1E-12
	0045449	BP	regulation of transcription	148	1139	13%	2E-12
	0031323	BP	reg. of cellular metabolic process	154	1216	13%	4E-12
	0010556	BP	reg. of macromolecule biosynthetic process	148	1152	13%	4E-12
	0019222	BP	regulation of metabolic process	155	1238	13%	6E-12
	0003676	MF	nucleic acid binding	311	3078	10%	6E-12
	0060255	BP	regulation of macromolecule metabolic process	148	1171	13%	9E-12
	0010468	BP	regulation of gene expression	152	1232	12%	3E-11
	0050791	BP	regulation of biological process	230	2149	11%	7E-11
	0051244	BP	regulation of cellular process	218	2018	11%	1E-10
	0065007	BP	biological regulation	236	2320	10%	5E-9
	0005488	MF	binding	579	6895	8%	2E-8
	0005634	CC	nucleus	151	1510	10%	6E-5
underrep	0060528	MF	transcription regulator activity	191	1441	13%	8E-17
	0003677	MF	DNA binding	214	1699	13%	9E-17
	0003700	MF	transcription factor activity	161	1207	13%	1E-14
	0019219	BP	regulation of nucleobase… and nucleic acid metabolic process	150	1153	13%	1E-12
	0009889	BP	regulation of biosynthetic process	151	1164	13%	1E-12
	0045449	BP	regulation of transcription	148	1139	13%	2E-12
	0031323	BP	reg. of cellular metabolic process	154	1216	13%	3E-12
	0010556	BP	reg. of macromolecule biosynthetic process	148	1152	13%	3E-12
	0019222	BP	regulation of metabolic process	155	1238	13%	6E-12
	0003676	MF	nucleic acid binding	311	3078	10%	6E-12
	0060255	BP	regulation of macromolecule metabolic process	148	1171	13%	9E-12
	0010468	BP	regulation of gene expression	152	1232	12%	3E-11
	0050791	BP	regulation of biological process	230	2149	11%	7E-11
	0051244	BP	regulation of cellular process	218	2018	11%	1E-10
	0065007	BP	biological regulation	236	2320	10%	5E-9
	0005488	MF	binding	59	6895	8%	2E-8
	0006534	CC	nucleus	151	1510	10%	6E-5

Provided are [a]overrepresented and underrepresented gene ontology (GO) ID numbers, [b]GO categories (Cat), and [c]GO term for gene products encoded by transcripts with at least one $G_{2+}L_{1-4}$ GQS. Included are [d]the number of genes (CDS, 5'UTR, 3'UTR, and introns) with a GQS that are annotated for the listed GO term, and [e]the total number of genes in Arabidopsis with the listed GO term. Also included are [f]the percentage of genes with GQS with a given GO term and [g]the appropriate p-value, as determined using the BiNGO program. [h]The total number of GO-annotated genes with a GQS in G_2L_{1-4} is 1,697. [i]The total number of GO-annotated genes in *A. thaliana* is 25,179. Table is sorted in order of increasing p-value. Some GO terms are sub-categories of others.

Table S7. Functional analysis of genes with at least one G_3L_{1-7} GQS present in the RNA

	GO ID[a]	GO Cat[b]	GO term[c]	GQS genes[d]	All genes[e]	% GQS genes[f]	p-value[g]
overrep	0005507	MF	copper ion binding	9[h]	109[i]	8%	8.8E-4
	0009790	BP	embryonic development	11	277	4%	3.3E-2
	0007275	BP	multicellular organismal dev.	20	871	2%	8.8E-2
underrep	0009059	BP	macromolecule biosynth process	0	1399	0	3.7E-3
	0009058	BP	biosynthetic process	3	2149	0.1%	3.7E-3
	0010467	BP	gene expression	1	1487	0.1%	1.2E-2
	0044249	BP	cellular biosynthetic process	2	1719	0.1%	1.2E-2
	0006416	BP	translation	0	1129	0	1.9E-2

Provided are [a]overrepresented and underrepresented gene ontology (GO) ID numbers, [b]GO categories (Cat), and [c]GO term for gene products encoded by transcripts with at least one $G_{2+}L_{1-4}$ GQS. Included are [d]the number of genes (CDS, 5'UTR, 3'UTR, and introns) with a GQS that are annotated for the listed GO term, and [e]the total number of genes in Arabidopsis with the listed GO term. Also included are [f]the percentage of genes with GQS with a given GO term and [g]the appropriate p-value, as determined using the BiNGO program. [h]The total number of GO-annotated genes with a GQS in G_2L_{1-4} is 246. [i]The total number of GO-annotated genes in *A. thaliana* is 25,179. Table is sorted in order of increasing p-value. Some GO terms are sub-categories of others.