

This is a preprint of an Article accepted for publication in *The Plant Journal* © 2005
Blackwell Publishing Ltd and The Society for Experimental Biology

Supplemental material

S1 Results from BLAST searches against several ncRNAs databases.

- 1.) Three sequences had perfect matches with bacterial vector sequences¹ (At_oRNA_566, At_oRNA_141, At_oRNA_413). This might be regarded as an error in the assembled *Arabidopsis* genomic sequence.
- 2.) Seven sequences could be assigned to ribosomal RNAs² in *A. thaliana*
Large subunit: At_oRNA_372 on chromosome 2, At_oRNA_567 on chromosome 3, and At_oRNA_568 on chromosome 2.
Small subunit: At_oRNA_092 on chromosome 2, At_oRNA_093 on chromosome 3, At_oRNA_422 on chromosome 2, and At_oRNA_423 on chromosome 3.
- 3.) One sequence (At_oRNA_426) was found to represent a tRNA for asparagine when tested against a set of *A. thaliana* tRNAs³. This finding was confirmed by the program tRNAscan-SE⁴ (Lowe and Eddy, 1997) which is often used to predict tRNA genes in genomic sequences. Sequences At_oRNA_425 and At_oRNA_426 correspond to two different loci for the same transcript, but only At_oRNA_426 is predicted to be a tRNA. Program tRNAscan-SE did not detect a tRNA in At_oRNA_425. Because both sequences are highly similar to the original transcript, it is possible that At_oRNA_425 is a non-functional version of the gene represented by At_oRNA_426. Interestingly, both genomic regions are being transcribed as seen in the “*Arabidopsis* Tiling Array Transcriptome Express Tool” (Yamada *et al.* 2003).

¹ Vector sequences obtained from NCBI <http://www.ncbi.nlm.nih.gov>

² Ribosomal sequences obtained from the European Ribosomal RNA database, <http://oberon.fvms.ugent.be:8080/rRNA/index.html>

³ tRNA sequences obtained from <http://rna.wustl.edu/GtRDB/At/At-seqs.html>

⁴ <http://www.genetics.wustl.edu/eddy/tRNAscan-SE/>

- 4.) Four additional sequences (At_oRNA_134, At_oRNA_486, At_oRNA_542, At_oRNA_555) had regions of high similarity to known non-coding RNAs (ncRNAs) (Erdmann *et al.*, 2000)⁵. Two of them (At_oRNA_134 and At_oRNA_486) produced hits to miRNAs from *A. thaliana*. Mature miRNAs are 20–24nt long. The precursor miRNAs are much larger than the mature forms and are able to fold into an imperfectly base-paired hairpin (Moss and Poethig, 2002). Taken this into account and comparing against the precursor and mature miRNAs reported by Griffiths-Jones (2004), both At_oRNAs could be called *bona fide* miRNAs. Both RNAs align perfectly to the miRNA precursors, and they exhibit a perfect match against the mature miRNAs, i.e. ath-miR168 and ath-miR167, respectively. Ath-miR168 is thought to target the *ARGONAUTE (AGO1)* gene that plays a role in post-transcriptional gene silencing, possibly by acting in a negative-feedback mechanism to regulate the expression of the *AGO1* gene (Griffiths-Jones, 2004). Ath-miR167 targets mRNAs coding for auxin response factors (Griffiths-Jones, 2004).
- 5.) A further RNA (At_oRNA_542) that matched against ncRNAs had a region of 19 nucleotides that is repeated three times in the *Xist* gene, which is responsible of X chromosome inactivation in mammalian female cells. Finally, At_oRNA_555 matched against At_CR20 (*A. thaliana* cytokinin-repressed gene), a ncRNA with high similarity to CR20 in cucumber (Szymanski *et al.*, 2003). This sequence is 610 nucleotides long, 388 nucleotides of which are similar to At_CR20. This region is divided into two sub-regions which have identities greater than 97%. According to BLAST searches both sequences (At_CR20 and At_oRNA_555) reside at the same genomic locus. At_oRNA_555 could be an unprocessed (unspliced) version of At_CR20.

⁵ <http://biobases.ibch.poznan.pl/ncRNA/>

- 6.) Another 18 sequences matched to known small nucleolar RNAs⁶ (snoRNAs; Brown *et al.*, 2003b). Some sequences carried more than one snoRNA (At_oRNA_513, At_oRNA_557, At_oRNA_587; see S2). This observation reflects the fact that some snoRNAs are processed from polycistronic pre-snoRNAs (Leader *et al.*, 1997; Liang-Hu *et al.*, 2001; Barneche *et al.*, 2001; Brown *et al.*, 2003a). In addition, some of the identified sequences exhibited some sequence similarity to snoRNAs from other plant species (At_oRNA_119, At_oRNA_401, At_oRNA_439, and At_oRNA_557). Two sequences (At_oRNA_017, At_oRNA_500) match to the snoRNAs AtsnoR38Y-1 and AtsnoR38Y-2. SnoRNAs with the same name but different numbers are transcribed from different loci. The BLAST result indicates that AtsnoR38Y-1 originates from the locus where At_oRNA_500 is located on chromosome 3. The matches between the snoRNAs and At_oRNA_500 is complete and perfect. On the other hand, AtsnoR38Y-2 would be located in the region where At_oRNA_017 is found. However, sequence At_oRNA_017 overlaps only with a short region of the snoRNA genes just at their ends. One interesting characteristic about the sequences carrying snoRNAs is that in two cases (sequences: At_oRNA_119 and At_oRNA_439) two copies of the snoRNAs were arranged in tandem; one of the copies was a perfect match to the original snoRNAs while the other had several mutations, mainly resulting from small tandem duplications (Barneche *et al.*, 2001).
- 7.) The dataset was compared against ncRNA sequences reported by Marker *et al.* (2002). Four sequences (At_oRNA_119, At_oRNA_485, At_oRNA_557, At_oRNA_587) matched against entries in the database, and all of them were already assigned to known snoRNAs. Finally, the dataset was tested against the sequences of the Small RNA project (<http://cgrb.orst.edu/smallRNA/db/>). Several small RNAs matched sequences annotated as ribosomal RNAs (At_oRNA_092, At_oRNA_093,

⁶ http://bioinf.scri.sari.ac.uk/cgi-bin/plant_snorna/home

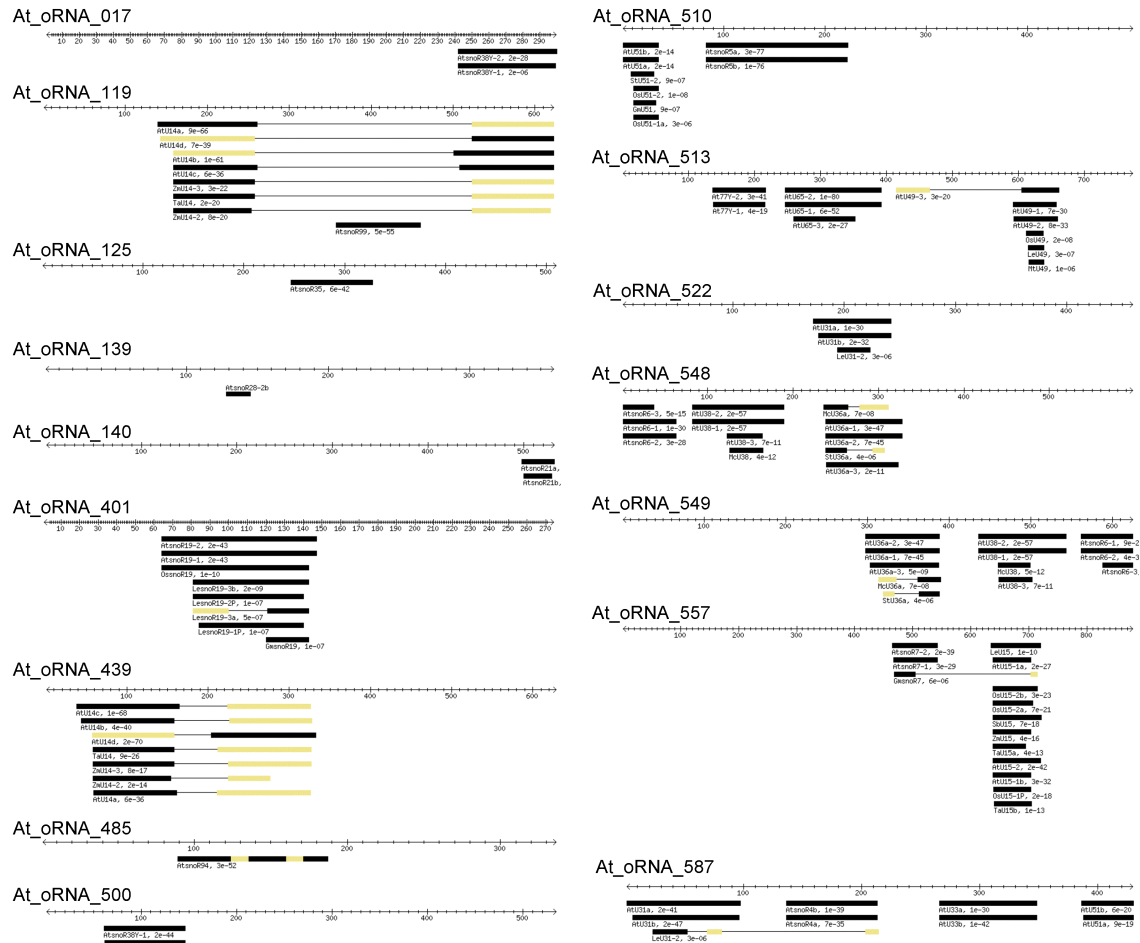
Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_422, At_oRNA_423). Two other RNAs (At_oRNA_134, At_oRNA_486) matched to miRNAs, as mentioned above. One further RNA exhibits several hits against an additional miRNA (miR389), but according to a Rfam search does not appear to be miRNA in itself. Twelve additional orphan transcripts have matches to several non-annotated but experimentally found small RNAs.

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

S2 Some orphan transcripts carry snoRNAs. See text in S1 for explanation.

(Figure can be downloaded in full resolution from <http://atornadb.bio.uni-potsdam.de/S2.php>)



Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

S3 Protein motifs found in orphan transcripts

(The table is available from <http://atornadb.bio.uni-potsdam.de/S3.php>). To find known protein motifs in orphan transcripts, InterProScan (Zdobnov and Apweiler, 2001) was employed. Transcripts were computationally translated in all possible reading frames. Six-frame translated orphan transcripts were scanned against the following databases: PROSITE patterns and profiles (Hulo *et al.*, 2004), PRINTS (Attwood *et al.*, 2003), PFAM (Bateman *et al.*, 2004), PRODOM (Servant *et al.*, 2002), SMART (Letunic *et al.*, 2002) and TIGRFAMs (Haft *et al.*, 2003).

At_oRNA ID	Motif/Domain
At_oRNA_025	InterPro domain IPR000595: Cyclic nucleotide-binding domain
	Pfam domain PF00027: cNMP_binding
	ProfileScan domain PS50042: CNMP_BINDING_3
At_oRNA_027	InterPro domain IPR008539: Arabidopsis thaliana protein of unknown function DUF821
	Pfam domain PF05686: DUF821
At_oRNA_031	InterPro domain IPR001024: Lipoxygenase, LH2 domain
	InterPro domain IPR010916: TONB Box N terminus
	ProfileScan domain PS50095: PLAT
	ScanRegExp domain PS00430: TONB_DEPENDENT_REC_1
At_oRNA_047	InterPro domain IPR000095: PAK-box/P21-Rho-binding
	Pfam domain PF00786: PBD
	ProfileScan domain PS50108: CRIB
At_oRNA_073	InterPro domain IPR000886: Endoplasmic reticulum targeting sequence
	InterPro domain IPR010547: Plant specific mitochondrial import receptor subunit TOM20
	Pfam domain PF06552: TOM20_plant
	ScanRegExp domain PS00014: ER_TARGET
At_oRNA_080	InterPro domain IPR001484: Pyrokinin
	ScanRegExp domain PS00539: PYROKININ
At_oRNA_090	InterPro domain IPR001440: TPR repeat
	ProfileScan domain PS50005: TPR_REPEAT
At_oRNA_111	InterPro domain IPR000313: PWWP domain
	ProfileScan domain PS50812: PWWP
	superfamily domain 63748: Tudor/PWWP
At_oRNA_114	InterPro domain IPR000095: PAK-box/P21-Rho-binding
	Pfam domain PF00786: PBD
	ProfileScan domain PS50108: CRIB
At_oRNA_118	InterPro domain IPR002048: Calcium-binding EF-hand
	ScanRegExp domain PS00018: EF_HAND
At_oRNA_122	HMMTigr domain TIGR00756: PPR
	InterPro domain IPR002885: PPR repeat
	Pfam domain PF01535: PPR
At_oRNA_126	FPrintScan domain PR01415: ANKYRIN
	HMMSmart domain SM00248: ANK

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

	InterPro domain IPR002110: Ankyrin
	Pfam domain PF00023: Ank
	ProfileScan domain PS50088: ANK_REPEAT
	ProfileScan domain PS50297: ANK_REP_REGION
At_oRNA_132	InterPro domain IPR001810: Cyclin-like F-box
	Pfam domain PF00646: F-box
	ProfileScan domain PS50181: FBOX
At_oRNA_136	HMMSmart domain SM00184: RING
	InterPro domain IPR001841: Zn-finger, RING
	Pfam domain PF00097: zf-C3HC4
	ProfileScan domain PS50089: ZF_RING_2
At_oRNA_159	InterPro domain IPR000276: Rhodopsin-like GPCR superfamily
	InterPro domain IPR003656: BED_finger
	InterPro domain IPR007021: Protein of unknown function DUF659
	Pfam domain PF04937: DUF659
	ProfileScan domain PS50808: ZF_BED
	ScanRegExp domain PS00237: G_PROTEIN_RECEP_F1_1
At_oRNA_186	FPrintScan domain PR00259: TMFOUR
	InterPro domain IPR000301: CD9/CD37/CD63 antigen
	Pfam domain PF00335: Tetraspanin
	ProfileScan domain PS50257: TM4_2
At_oRNA_238	InterPro domain IPR008011: Complex 1 LYR protein
	Pfam domain PF05347: Complex1_LYR
At_oRNA_267	InterPro domain IPR003959: AAA ATPase, central region
	Pfam domain PF00004: AAA
At_oRNA_282	HMMSmart domain SM00185: ARM
	InterPro domain IPR000225: Armadillo
	InterPro domain IPR004825: Insulin/IGF/relaxin
	Pfam domain PF00514: Arm
	ProfileScan domain PS50176: ARM_REPEAT
	ScanRegExp domain PS00262: INSULIN
At_oRNA_294	InterPro domain IPR008801: Rapid ALkalinization Factor
	Pfam domain PF05498: RALF
At_oRNA_332	BlastProDom domain PD003928: sp_Q9STV9_ARATH_Q9STV9;
	InterPro domain IPR000845: Purine and other phosphorylases, family 1
	InterPro domain IPR001179: Peptidylprolyl isomerase, FKBP-type
	InterPro domain IPR005829: Sugar transporter superfamily
	Pfam domain PF01048: PNP_UDP_1
	ScanRegExp domain PS00216: SUGAR_TRANSPORT_1
	ScanRegExp domain PS00453: FKBP_PPIASE_1
At_oRNA_333	InterPro domain IPR010682: Plant self-incompatibility response
	Pfam domain PF06876: SCRL
At_oRNA_335	HMMSmart domain SM00185: ARM
	InterPro domain IPR000225: Armadillo
	Pfam domain PF00514: Arm
	ProfileScan domain PS50176: ARM_REPEAT
At_oRNA_353	InterPro domain IPR000490: Glycoside hydrolase, family 17
	Pfam domain PF00332: Glyco_hydro_17
At_oRNA_392	InterPro domain IPR001810: Cyclin-like F-box
	Pfam domain PF00646: F-box

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_402	HMMSmart domain SM00487: DEXDc
	HMMSmart domain SM00490: HELICc
	InterPro domain IPR000330: SNF2 related domain
	InterPro domain IPR001410: DEAD/DEAH box helicase
	InterPro domain IPR001650: Helicase, C-terminal
	InterPro domain IPR004825: Insulin/IGF/relaxin
	Pfam domain PF00176: SNF2_N
	Pfam domain PF00271: Helicase_C
At_oRNA_409	ScanRegExp domain PS00262: INSULIN
	InterPro domain IPR001810: Cyclin-like F-box
At_oRNA_411	Pfam domain PF00646: F-box
	InterPro domain IPR000504: RNA-binding region RNP-1
	InterPro domain IPR003441: No apical meristem (NAM) protein
	Pfam domain PF02365: NAM
At_oRNA_427	ScanRegExp domain PS00030: RRM_RNP_1
	FPrintScan domain PR00239: RHODOPSNTAIL
	FPrintScan domain PR01217: PRICHEXTENSIN
At_oRNA_441	InterPro domain IPR006030: Molluscan rhodopsin C-terminal tail
	HMMSmart domain SM00184: RING
	InterPro domain IPR001841: Zn-finger, RING
	Pfam domain PF00097: zf-C3HC4
At_oRNA_447	ProfileScan domain PS50089: ZF_RING_2
	FPrintScan domain PR00194: TROPOMYOSIN
	InterPro domain IPR000533: Tropomyosin
At_oRNA_449	InterPro domain IPR006867: Protein of unknown function DUF632
	InterPro domain IPR006868: Protein of unknown function DUF630
	Pfam domain PF04782: DUF632
	Pfam domain PF04783: DUF630
At_oRNA_459	HMMSmart domain SM00256: FBOX
	InterPro domain IPR001810: Cyclin-like F-box
	Pfam domain PF00646: F-box
	ProfileScan domain PS50181: FBOX
At_oRNA_460	HMMSmart domain SM00185: ARM
	InterPro domain IPR000225: Armadillo
	InterPro domain IPR001007: von Willebrand factor, type C
	InterPro domain IPR001450: 4Fe-4S ferredoxin, iron-sulfur binding domain
	ScanRegExp domain PS00198: 4FE4S_FERREDOXIN
	ScanRegExp domain PS01208: VWFC_1
At_oRNA_462	HMMSmart domain SM00185: ARM
	InterPro domain IPR000225: Armadillo
	InterPro domain IPR001007: von Willebrand factor, type C
	InterPro domain IPR001450: 4Fe-4S ferredoxin, iron-sulfur binding domain
	ScanRegExp domain PS00198: 4FE4S_FERREDOXIN
	ScanRegExp domain PS01208: VWFC_1
At_oRNA_469	InterPro domain IPR000307: Ribosomal protein S16
	ScanRegExp domain PS00732: RIBOSOMAL_S16
At_oRNA_473	BlastProDom domain PD000001: sp_Q9LMN2_ARATH_Q9LMN2;
	HMMSmart domain SM00219: TyrKc
	HMMSmart domain SM00220: S_TKc
	InterPro domain IPR000719: Protein kinase

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

	InterPro domain IPR001245: Tyrosine protein kinase
	InterPro domain IPR002290: Serine/threonine protein kinase
	InterPro domain IPR006056: YjgF-like protein
	InterPro domain IPR010632: Protein of unknown function DUF1221
	Pfam domain PF00069: Pkinase
	Pfam domain PF06760: DUF1221
	ProfileScan domain PS50011: PROTEIN_KINASE_DOM
	ScanRegExp domain PS01094: UPF0076
At_oRNA_515	InterPro domain IPR000315: Zn-finger, B-box
	InterPro domain IPR006734: Protein of unknown function DUF597
	Pfam domain PF04640: DUF597
	ProfileScan domain PS50119: ZF_BBOX
At_oRNA_525	HMMTigr domain TIGR01589: A_thal_3526
	InterPro domain IPR006476: Uncharacterized plant-specific domain 01589
At_oRNA_526	HMMTigr domain TIGR01589: A_thal_3526
	InterPro domain IPR006476: Uncharacterized plant-specific domain 01589
At_oRNA_534	InterPro domain IPR001810: Cyclin-like F-box
	Pfam domain PF00646: F-box
At_oRNA_541	InterPro domain IPR002110: Ankyrin
	Pfam domain PF00023: Ank
	ProfileScan domain PS50088: ANK_REPEAT
	ProfileScan domain PS50297: ANK_REP_REGION

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

S4 Representation of orphan sequences on the Affymetrix chip arrays (Tables also available from <http://atornadb.bio.uni-potsdam.de/S4.php>).

From the 591 orphan sequences identified in this study, 573 are not represented by the Affymetrix GeneChip AtGenome1.

Matches against GeneChip AtGenome1 are indicated.

At_oRNA_ID	Chip AtGenome1											
	12842_s_at	12846_s_at	13333_at	14125_at	15382_at	16066_at	18049_at	18060_i_at	18061_r_at	18673_at	20033_at	20223_at
At_oRNA_028												1
At_oRNA_092	5											
At_oRNA_093	6											
At_oRNA_137					4							
At_oRNA_396**		4										
At_oRNA_397**		4										
At_oRNA_398**		4										
At_oRNA_401*							9	13				
At_oRNA_438						1						
At_oRNA_473							12					
At_oRNA_493				7								
At_oRNA_516			1									
At_oRNA_531											16	
At_oRNA_551**		26										
At_oRNA_552**		29										
At_oRNA_553**		14										
At_oRNA_554**		14										
At_oRNA_555*									4			

*Sequences At_oRNA_401 and At_oRNA_555 correspond to ncRNAs. They were annotated accordingly by Affymetrix.

- Sequence At_oRNA_555 matches perfectly to 4 probes in the probe set 18673_at, which is annotated as AtCR20-1 (Liu *et al.*, 2003), *Arabidopsis thaliana* cytokinin-repressed gene, which was first described in cucumber, named CR20 (Teramoto *et al.*, 1996).
- Sequence At_oRNA_401 matches to 9 probes in the probe set 18060_i_at, and to 13 probes in the probe set 18061_r_at, both of which were annotated by Affymetrix as U25 snoRNA. However, according to Brown *et al.* (2003b) this sequence corresponds to AtsnoR19-1.

**Seven sequences match to several probes of the probe set 12846_s_at, which is annotated as an intergenic region in an rRNA operon.

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

12842_s_at Nulear ribosomal ITS-1, 5.8S ribosomal RNA	18049_at anti-sense transcript, AKL kinase-like gene
12846_s_at DNA for rRNA intergenic region	18060_i_at mRNA for U25 snoRNA
13333_at At2g46770	18061_r_at mRNA for U25 snoRNA
14125_at	18673_at AtCR20-1 mRNA
15382_at No annotation available	20033_at
16066_at At3g46100	20223_at At4g23230,At4g23220

From the 591 orphan sequences identified in this study, 586 are not represented by the Affymetrix GeneChip ATH1.

Matches against GeneChip ATH1 are indicated.

At_oRNA_ID	chip ATH1				
	248126_at (At5g54760)	257329_at	258833_at (At3g07280)	264494_at (At1g27470)	266714_at (At2g46770)
At_oRNA_176		2			
At_oRNA_228	1				
At_oRNA_405				3	
At_oRNA_516					2
At_oRNA_579			2		

Of the 5 sequences identified on the ATH1 chip, none of them matched to more than 3 probes.

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

S5 Summary information for orphan transcripts in At oRNA DB. This information can be downloaded as a plain-text file from http://atornadb.bio.uni-potsdam.de/At_oRNA_Summary-04042005.txt

At oRNA ID	Chromosome	Rigth end	Left end	Strand
At_oRNA_001	5	26517102	26517432	+
At_oRNA_002	5	26519328	26519588	-
At_oRNA_003	5	20718314	20718349	-
At_oRNA_004	3	15983625	15983849	+
At_oRNA_005	3	1139193	1139426	+
At_oRNA_006	5	8072682	8072920	-
At_oRNA_007	4	14558989	14559188	+
At_oRNA_008	3	7830932	7831100	+
At_oRNA_009	1	27913591	27913715	+
At_oRNA_010	2	7971783	7971988	-
At_oRNA_011	5	19640896	19641416	-
At_oRNA_012	3	7471149	7471499	+
At_oRNA_013	2	14914456	14914630	-
At_oRNA_014	1	5192325	5192614	-
At_oRNA_015	1	5191886	5192194	+
At_oRNA_016	5	23699648	23699862	+
At_oRNA_017	5	23699865	23700163	-
At_oRNA_018	3	11502106	11503352	-
At_oRNA_019	4	13294992	13295421	+
At_oRNA_020	1	23980207	23980657	-
At_oRNA_021	1	2492681	2493178	-
At_oRNA_022	2	3322439	3322682	+
At_oRNA_023	5	3995397	3995632	+
At_oRNA_024	3	7873803	7874050	-
At_oRNA_025	2	8658885	8659308	-
At_oRNA_026	3	21674742	21675013	-
At_oRNA_027	3	22692123	22692892	+
At_oRNA_028	4	12152715	12152992	-
At_oRNA_029	5	18739011	18739572	+
At_oRNA_030	5	3932758	3933204	+
At_oRNA_031	5	26047579	26047851	+
At_oRNA_032	2	3447782	3448104	-
At_oRNA_033	3	9418791	9419221	+
At_oRNA_034	4	18574508	18574943	+
At_oRNA_035	4	4311279	4311941	-
At_oRNA_036	1	25140872	25141165	+
At_oRNA_037	3	7492166	7492608	-
At_oRNA_038	3	2140114	2140556	+
At_oRNA_039	1	27874938	27875385	+
At_oRNA_040	1	23334213	23334441	-
At_oRNA_041	1	2581043	2581271	+
At_oRNA_042	5	17988008	17988676	-
At_oRNA_043	5	7667774	7668212	+
At_oRNA_044	1	19500348	19501349	-

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_045	5	498631	498882	-
At_oRNA_046	3	8312678	8313081	-
At_oRNA_047	1	22823821	22824662	-
At_oRNA_048	2	13125425	13125885	-
At_oRNA_049	2	19306562	19307000	+
At_oRNA_050	1	19606818	19607233	+
At_oRNA_051	5	26581191	26581445	+
At_oRNA_052	1	29314273	29314757	-
At_oRNA_053	4	11368	11743	+
At_oRNA_054	3	6038866	6039315	-
At_oRNA_055	5	607850	608254	-
At_oRNA_056	5	16887317	16888115	-
At_oRNA_057	1	20082750	20083230	+
At_oRNA_058	4	8347315	8347752	-
At_oRNA_059	5	5462594	5463019	-
At_oRNA_060	3	19320849	19321278	+
At_oRNA_061	5	25540473	25540831	-
At_oRNA_062	4	17340441	17340835	-
At_oRNA_063	5	12102130	12102517	+
At_oRNA_064	4	1382439	1382913	+
At_oRNA_065	5	22821177	22821600	-
At_oRNA_066	1	28057991	28058307	-
At_oRNA_067	4	329579	329994	+
At_oRNA_068	3	2296156	2296818	+
At_oRNA_069	4	15659921	15660229	-
At_oRNA_070	1	2014005	2014204	+
At_oRNA_071	3	727118	727502	-
At_oRNA_072	2	10682350	10682740	+
At_oRNA_073	5	16416696	16417465	-
At_oRNA_074	2	19076231	19076640	-
At_oRNA_075	1	23700239	23700561	+
At_oRNA_076	2	15112807	15113167	-
At_oRNA_077	3	4358187	4358537	+
At_oRNA_078	5	24272124	24272561	+
At_oRNA_079	1	21178438	21178818	+
At_oRNA_080	1	3264815	3265183	-
At_oRNA_081	2	5057580	5057977	+
At_oRNA_082	2	6671885	6672338	-
At_oRNA_083	1	19509427	19509831	-
At_oRNA_084	1	12914410	12914782	-
At_oRNA_085	5	6032817	6033223	-
At_oRNA_086	1	5895047	5895442	-
At_oRNA_087	1	19855416	19855795	-
At_oRNA_088	3	17206044	17206265	-
At_oRNA_089	1	11138511	11138946	-
At_oRNA_090	5	3204200	3204709	-
At_oRNA_091	1	5031543	5031774	-
At_oRNA_092	2	5287	5702	-
At_oRNA_093	3	14210244	14210659	-
At_oRNA_094	3	852949	853685	+
At_oRNA_095	4	8985489	8985886	-

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_096	3	22301405	22302281	-
At_oRNA_097	2	15066423	15066826	+
At_oRNA_098	4	9115782	9116191	-
At_oRNA_099	4	16480941	16481318	+
At_oRNA_100	1	23620931	23621430	-
At_oRNA_101	1	24353404	24353911	-
At_oRNA_102	2	17369707	17370175	+
At_oRNA_103	5	6324028	6324519	+
At_oRNA_104	5	4337946	4338098	+
At_oRNA_105	1	2212304	2212545	-
At_oRNA_106	1	162575	162757	-
At_oRNA_107	5	22348909	22349039	-
At_oRNA_108	4	14998302	14998657	-
At_oRNA_109	3	9419299	9419941	-
At_oRNA_110	1	21218460	21219057	-
At_oRNA_111	3	7489243	7489713	+
At_oRNA_112	4	10038139	10038263	+
At_oRNA_113	3	8311805	8312201	+
At_oRNA_114	1	22823363	22824110	+
At_oRNA_115	2	4312558	4313026	-
At_oRNA_116	5	21455375	21455817	-
At_oRNA_117	4	12147	13346	-
At_oRNA_118	1	26897441	26898464	-
At_oRNA_119	4	18304746	18305367	+
At_oRNA_120	4	17338699	17339311	+
At_oRNA_121	4	334156	334487	-
At_oRNA_122	2	5594683	5595347	-
At_oRNA_123	3	2297088	2297746	-
At_oRNA_124	5	10238556	10239212	+
At_oRNA_125	1	2381647	2382582	+
At_oRNA_126	5	4591744	4593764	+
At_oRNA_127	5	24074634	24075295	+
At_oRNA_128	2	18477846	18478352	-
At_oRNA_129	3	4358512	4359014	-
At_oRNA_130	3	9476588	9476796	+
At_oRNA_131	2	3400085	3400728	-
At_oRNA_132	3	9921700	9922435	+
At_oRNA_133	3	17202654	17206138	+
At_oRNA_134	1	11137347	11137963	+
At_oRNA_135	5	3201520	3201644	+
At_oRNA_136	5	14160335	14160463	-
At_oRNA_137	1	23038134	23038564	-
At_oRNA_138	5	4528586	4529055	+
At_oRNA_139	5	18875179	18875739	+
At_oRNA_140	2	15036290	15036820	-
At_oRNA_141	3	13748218	13748264	+
At_oRNA_142	2	8876735	8877351	-
At_oRNA_143	4	2422618	2422848	+
At_oRNA_144	4	2327009	2327483	-
At_oRNA_145	5	20986483	20986730	-
At_oRNA_146	3	19495409	19495767	+

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_147	5	17447894	17448474	-
At_oRNA_148	3	10348340	10348811	-
At_oRNA_149	5	22595944	22596578	-
At_oRNA_150	5	9902804	9903491	+
At_oRNA_151	3	8171597	8171915	-
At_oRNA_152	1	18034461	18034759	+
At_oRNA_153	5	1640658	1640843	-
At_oRNA_154	5	4956713	4957588	+
At_oRNA_155	1	30369372	30369668	+
At_oRNA_156	4	11891924	11892072	+
At_oRNA_157	5	3303547	3304695	-
At_oRNA_158	5	9136268	9136616	+
At_oRNA_159	1	30011315	30011757	-
At_oRNA_160	2	229463	229826	+
At_oRNA_161	4	9041941	9043454	+
At_oRNA_162	3	11139410	11139816	+
At_oRNA_163	1	28941154	28941614	-
At_oRNA_164	1	28790029	28790446	+
At_oRNA_165	5	20922077	20922210	-
At_oRNA_166	4	14034661	14035090	+
At_oRNA_167	1	24916402	24917206	-
At_oRNA_168	5	15649259	15649584	-
At_oRNA_169	3	22097457	22097852	+
At_oRNA_170	1	21133710	21133954	+
At_oRNA_171	3	5095516	5095817	-
At_oRNA_172	2	18395519	18396298	+
At_oRNA_173	2	18398541	18398858	+
At_oRNA_174	2	3431194	3431577	-
At_oRNA_175	1	5084348	5084672	+
At_oRNA_176	2	3352987	3353414	-
At_oRNA_177	2	3444920	3445252	+
At_oRNA_178	4	15897595	15898014	-
At_oRNA_179	1	558347	558730	-
At_oRNA_180	1	556668	557055	+
At_oRNA_181	2	3374766	3375067	-
At_oRNA_182	2	3508934	3509234	-
At_oRNA_183	2	3370993	3371307	+
At_oRNA_184	2	3505161	3505475	+
At_oRNA_185	1	8542435	8543090	+
At_oRNA_186	2	10144341	10144527	-
At_oRNA_187	5	9560345	9560434	+
At_oRNA_188	5	25742838	25743182	+
At_oRNA_189	5	15923195	15923454	-
At_oRNA_190	3	2601935	2602162	-
At_oRNA_191	5	16726894	16727137	-
At_oRNA_192	2	3409920	3410247	+
At_oRNA_193	3	22416944	22417279	-
At_oRNA_194	2	3253154	3253494	+
At_oRNA_195	5	25229960	25230301	-
At_oRNA_196	1	17870477	17870738	+
At_oRNA_197	2	3504037	3504441	+

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_198	2	3369869	3370274	+
At_oRNA_199	2	3370385	3370790	-
At_oRNA_200	2	3504553	3504958	-
At_oRNA_201	5	26384810	26385099	-
At_oRNA_202	4	17377405	17377625	-
At_oRNA_203	4	12987127	12987422	-
At_oRNA_204	3	5895323	5895710	-
At_oRNA_205	1	24129042	24129132	-
At_oRNA_206	1	25890468	25892687	+
At_oRNA_207	1	2561625	2562101	+
At_oRNA_208	1	22667628	22668176	+
At_oRNA_209	1	29349719	29350139	-
At_oRNA_210	5	25876537	25877134	-
At_oRNA_211	5	25488605	25489166	+
At_oRNA_212	5	26090857	26091345	+
At_oRNA_213	3	4576159	4576677	+
At_oRNA_214	5	25997774	25998131	-
At_oRNA_215	1	2388032	2388134	-
At_oRNA_216	1	1523924	1524432	-
At_oRNA_217	3	2636665	2636727	-
At_oRNA_218	1	4318825	4318983	-
At_oRNA_219	1	4061364	4061449	+
At_oRNA_220	3	1108846	1109080	+
At_oRNA_221	1	2564590	2564680	+
At_oRNA_222	3	4374260	4374695	-
At_oRNA_223	4	15750705	15750862	+
At_oRNA_224	1	12601579	12601805	-
At_oRNA_225	1	8543019	8543403	-
At_oRNA_226	5	14551839	14551903	+
At_oRNA_227	1	29190781	29190862	+
At_oRNA_228	1	9175599	9177543	-
At_oRNA_229	2	17648979	17649409	-
At_oRNA_230	2	10897449	10897667	-
At_oRNA_231	3	5564426	5564554	-
At_oRNA_232	1	22080270	22080638	+
At_oRNA_233	5	7794146	7795013	-
At_oRNA_234	1	8896785	8897085	-
At_oRNA_235	4	9420938	9420991	-
At_oRNA_236	3	9185858	9186141	-
At_oRNA_237	4	11128379	11128882	-
At_oRNA_238	1	28549642	28550228	+
At_oRNA_239	3	21441815	21441978	+
At_oRNA_240	2	17657552	17658089	-
At_oRNA_241	2	7506313	7506358	-
At_oRNA_242	2	3286932	3287321	-
At_oRNA_243	2	3243368	3243757	+
At_oRNA_244	3	23309794	23310291	-
At_oRNA_245	2	3429390	3429737	+
At_oRNA_246	1	24108366	24108858	-
At_oRNA_247	1	21619870	21620452	-
At_oRNA_248	1	21653629	21655553	+

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_249	1	21675779	21676423	+
At_oRNA_250	1	21870257	21871851	+
At_oRNA_251	4	18120707	18120953	-
At_oRNA_252	2	3277109	3277520	-
At_oRNA_253	3	5158459	5158686	+
At_oRNA_254	3	17526071	17526190	-
At_oRNA_255	3	9106003	9106133	+
At_oRNA_256	3	9109947	9110077	+
At_oRNA_257	1	3597921	3598068	-
At_oRNA_258	5	6573625	6574146	+
At_oRNA_259	1	12330654	12330763	+
At_oRNA_260	5	6235047	6235195	+
At_oRNA_261	5	18856978	18857305	-
At_oRNA_262	1	1990760	1991812	+
At_oRNA_263	4	17447502	17447904	+
At_oRNA_264	2	12971482	12971743	+
At_oRNA_265	1	8922972	8923317	-
At_oRNA_266	1	1991759	1992161	-
At_oRNA_267	5	21451667	21452082	+
At_oRNA_268	1	27864801	27866307	+
At_oRNA_269	1	11472070	11472775	-
At_oRNA_270	1	2443779	2445118	+
At_oRNA_271	4	7887430	7893684	-
At_oRNA_272	5	19296798	19296848	-
At_oRNA_273	3	4368931	4368992	+
At_oRNA_274	5	6727517	6727927	+
At_oRNA_275	3	7815501	7815884	-
At_oRNA_276	4	1516504	1516853	+
At_oRNA_277	1	28424791	28425137	+
At_oRNA_278	1	24895594	24895709	+
At_oRNA_279	1	29192410	29192692	-
At_oRNA_280	1	18751958	18752095	+
At_oRNA_281	2	788038	788492	+
At_oRNA_282	2	9416629	9417000	+
At_oRNA_283	5	22254057	22254543	-
At_oRNA_284	4	7843740	7844283	+
At_oRNA_285	4	16020752	16020832	-
At_oRNA_286	3	22180554	22180794	-
At_oRNA_287	4	18159544	18159892	-
At_oRNA_288	4	5724019	5724129	+
At_oRNA_289	5	9470176	9470340	-
At_oRNA_290	2	3425662	3426072	-
At_oRNA_291	5	83220	83563	+
At_oRNA_292	5	3072947	3073335	+
At_oRNA_293	5	22597118	22597688	+
At_oRNA_294	4	7624678	7624944	+
At_oRNA_295	2	9556612	9557137	+
At_oRNA_296	1	4730380	4730896	+
At_oRNA_297	4	912998	913265	-
At_oRNA_298	3	21988224	21988366	-
At_oRNA_299	5	15847329	15847449	+

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_300	1	30286253	30286520	+
At_oRNA_301	3	525819	526216	-
At_oRNA_302	5	16845963	16846314	-
At_oRNA_303	5	25247165	25247544	+
At_oRNA_304	1	29760621	29761014	-
At_oRNA_305	2	13972864	13972989	+
At_oRNA_306	2	3416692	3417096	-
At_oRNA_307	4	7484698	7484801	+
At_oRNA_308	5	9671511	9671720	+
At_oRNA_309	5	14670382	14671284	-
At_oRNA_310	3	16799899	16800258	+
At_oRNA_311	5	18002440	18002700	+
At_oRNA_312	3	5123659	5123896	+
At_oRNA_313	5	22389269	22390252	+
At_oRNA_314	3	9916352	9916437	-
At_oRNA_315	3	5248559	5248879	+
At_oRNA_316	5	17978430	17979360	-
At_oRNA_317	2	19356076	19356571	-
At_oRNA_318	3	4368757	4368886	-
At_oRNA_319	1	20651397	20651804	-
At_oRNA_320	5	20054325	20054450	+
At_oRNA_321	4	5338954	5339059	-
At_oRNA_322	2	13759530	13759709	+
At_oRNA_323	1	2056575	2056972	+
At_oRNA_324	5	1690598	1690808	+
At_oRNA_325	2	13567477	13567785	+
At_oRNA_326	2	193215	193812	+
At_oRNA_327	2	18350853	18350979	-
At_oRNA_328	1	6669364	6669536	-
At_oRNA_329	3	10103471	10103624	-
At_oRNA_330	2	8857394	8857975	-
At_oRNA_331	4	11611305	11611732	-
At_oRNA_332	4	14274491	14275292	+
At_oRNA_333	1	22465781	22466207	-
At_oRNA_334	5	6552928	6553259	-
At_oRNA_335	2	9413344	9414577	+
At_oRNA_336	3	4704555	4705529	-
At_oRNA_337	4	18296617	18296757	+
At_oRNA_338	4	12577655	12577877	+
At_oRNA_339	2	8837172	8837582	+
At_oRNA_340	5	9355984	9356506	+
At_oRNA_341	2	16391487	16391719	-
At_oRNA_342	5	13927050	13927403	-
At_oRNA_343	5	16643459	16643552	+
At_oRNA_344	1	24057761	24057857	+
At_oRNA_345	5	10239441	10239984	-
At_oRNA_346	1	30221758	30221983	-
At_oRNA_347	4	17283109	17283187	+
At_oRNA_348	3	4154936	4155567	+
At_oRNA_349	2	6988735	6989183	-
At_oRNA_350	2	3262243	3262773	-

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_351	2	17533954	17534164	+
At_oRNA_352	1	5890651	5891130	+
At_oRNA_353	4	15143387	15143467	+
At_oRNA_354	1	20587832	20588231	+
At_oRNA_355	4	12354592	12354971	-
At_oRNA_356	4	17709675	17709759	-
At_oRNA_357	1	28028632	28028709	+
At_oRNA_358	1	13197519	13197954	-
At_oRNA_359	4	14415077	14415492	-
At_oRNA_360	5	16082227	16082293	+
At_oRNA_361	2	3465404	3465945	-
At_oRNA_362	3	20280043	20280340	+
At_oRNA_363	4	17248973	17249429	-
At_oRNA_364	4	2327921	2328250	-
At_oRNA_365	4	17945660	17946051	+
At_oRNA_366	3	17355651	17356032	+
At_oRNA_367	3	22820521	22820807	+
At_oRNA_368	5	8797450	8797879	+
At_oRNA_369	3	20615197	20615458	+
At_oRNA_370	4	329199	330448	-
At_oRNA_371	1	927629	928045	-
At_oRNA_372	2	3385613	3385919	-
At_oRNA_373	1	11929453	11929631	-
At_oRNA_374	2	6898134	6898399	-
At_oRNA_375	3	9860906	9861256	-
At_oRNA_376	4	4310954	4311346	+
At_oRNA_377	1	2187347	2187726	+
At_oRNA_378	1	28026795	28027208	+
At_oRNA_379	5	2559783	2559970	-
At_oRNA_380	2	432296	432595	+
At_oRNA_381	3	18399441	18399595	-
At_oRNA_382	1	28177511	28177964	+
At_oRNA_383	2	10683096	10683411	-
At_oRNA_384	2	11995411	11995727	+
At_oRNA_385	5	26516569	26516930	-
At_oRNA_386	5	17396243	17396452	-
At_oRNA_387	5	18922528	18922780	+
At_oRNA_388	1	16377935	16378196	+
At_oRNA_389	4	12923101	12923354	-
At_oRNA_390	1	28941575	28941835	-
At_oRNA_391	1	28789554	28789805	+
At_oRNA_392	4	6454403	6454672	-
At_oRNA_393	1	26777190	26777441	-
At_oRNA_394	4	9244112	9244356	+
At_oRNA_395	2	3415079	3415359	-
At_oRNA_396	2	10057	10372	-
At_oRNA_397	3	14204689	14204930	-
At_oRNA_398	3	14215098	14215339	-
At_oRNA_399	3	9857282	9857529	-
At_oRNA_400	5	16461890	16462168	-
At_oRNA_401	3	17456304	17456577	+

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_402	1	1620575	1620821	-
At_oRNA_403	1	13457231	13457488	-
At_oRNA_404	3	7441842	7442521	+
At_oRNA_405	1	9537780	9538206	+
At_oRNA_406	5	19684101	19684550	+
At_oRNA_407	3	20978720	20979144	+
At_oRNA_408	5	6730130	6730410	+
At_oRNA_409	3	16011867	16012325	+
At_oRNA_410	1	888322	888813	+
At_oRNA_411	5	2894720	2895422	-
At_oRNA_412	2	3437728	3438367	+
At_oRNA_413	3	13748406	13748796	-
At_oRNA_414	5	6160180	6160449	+
At_oRNA_415	1	28933786	28934103	+
At_oRNA_416	1	11453628	11454014	+
At_oRNA_417	5	1852606	1852988	-
At_oRNA_418	3	1141239	1141491	+
At_oRNA_419	5	4373243	4373679	+
At_oRNA_420	5	4373781	4374200	-
At_oRNA_421	1	17508802	17509192	-
At_oRNA_422	2	3736	4725	+
At_oRNA_423	3	14208693	14209682	+
At_oRNA_424	2	3364129	3364536	+
At_oRNA_425	2	3186718	3187039	-
At_oRNA_426	2	3496108	3496506	+
At_oRNA_427	1	23054139	23054648	+
At_oRNA_428	4	8070185	8070638	+
At_oRNA_429	3	19667467	19667870	+
At_oRNA_430	2	3333281	3334123	-
At_oRNA_431	4	10835003	10835712	-
At_oRNA_432	3	19210761	19211200	+
At_oRNA_433	4	15779309	15779702	-
At_oRNA_434	1	23987721	23988464	-
At_oRNA_435	2	10660650	10661148	+
At_oRNA_436	3	13360179	13360669	+
At_oRNA_437	2	804174	804616	+
At_oRNA_438	5	841071	841861	+
At_oRNA_439	4	18305372	18305999	-
At_oRNA_440	1	19886092	19886551	+
At_oRNA_441	5	21549799	21550376	+
At_oRNA_442	5	17978450	17979030	-
At_oRNA_443	5	17978490	17978742	-
At_oRNA_444	4	7785682	7786085	-
At_oRNA_445	1	24363543	24363724	-
At_oRNA_446	1	4675926	4676470	-
At_oRNA_447	2	13695702	13698133	-
At_oRNA_448	3	22581881	22582356	+
At_oRNA_449	1	19488025	19488559	+
At_oRNA_450	5	9944817	9945337	-
At_oRNA_451	3	9476846	9477195	+
At_oRNA_452	3	4698571	4699406	+

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_453	1	765503	766153	-
At_oRNA_454	1	23628334	23628832	+
At_oRNA_455	4	2504914	2505525	-
At_oRNA_456	5	8311056	8311988	-
At_oRNA_457	4	7892388	7893685	-
At_oRNA_458	4	7842994	7843736	+
At_oRNA_459	4	1611743	1612616	+
At_oRNA_460	2	9418666	9419191	+
At_oRNA_461	3	971684	972159	+
At_oRNA_462	2	9419261	9419872	-
At_oRNA_463	5	2854958	2855581	-
At_oRNA_464	4	4311483	4311953	-
At_oRNA_465	4	16761075	16761461	+
At_oRNA_466	5	9285135	9285553	+
At_oRNA_467	4	3939298	3940737	+
At_oRNA_468	2	4312032	4312596	+
At_oRNA_469	4	18220562	18221107	+
At_oRNA_470	4	2506831	2507416	+
At_oRNA_471	4	14045973	14046666	-
At_oRNA_472	1	3544992	3545458	+
At_oRNA_473	1	165736	166290	-
At_oRNA_474	3	23259105	23260072	+
At_oRNA_475	3	2140475	2141492	-
At_oRNA_476	3	2140922	2141529	-
At_oRNA_477	1	3749549	3749964	+
At_oRNA_478	5	10632270	10632799	+
At_oRNA_479	1	13479519	13479606	+
At_oRNA_480	1	23038862	23039564	-
At_oRNA_481	1	24915790	24916595	+
At_oRNA_482	4	11730147	11730418	-
At_oRNA_483	5	26217255	26217899	+
At_oRNA_484	1	20581102	20581948	-
At_oRNA_485	3	11155392	11155727	+
At_oRNA_486	4	10578623	10579191	+
At_oRNA_487	4	2507435	2507913	-
At_oRNA_488	1	2030008	2031384	+
At_oRNA_489	4	10978983	10979471	-
At_oRNA_490	4	7843214	7843616	+
At_oRNA_491	2	19198321	19198848	-
At_oRNA_492	3	2287177	2287550	-
At_oRNA_493	2	2901309	2901732	+
At_oRNA_494	1	27890088	27890718	+
At_oRNA_495	1	16419408	16420248	-
At_oRNA_496	3	1848968	1850602	+
At_oRNA_497	1	13282385	13283213	+
At_oRNA_498	3	20143865	20144728	-
At_oRNA_499	3	10148381	10148966	+
At_oRNA_500	3	17456888	17457421	-
At_oRNA_501	5	6569197	6570874	+
At_oRNA_502	2	3378344	3378914	-
At_oRNA_503	3	14318605	14319279	+

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_504	4	1961140	1961890	+
At_oRNA_505	3	19498105	19499533	+
At_oRNA_506	3	22001257	22001764	-
At_oRNA_507	1	12002324	12002857	-
At_oRNA_508	2	19617544	19618034	+
At_oRNA_509	1	25815504	25815981	+
At_oRNA_510	1	10300786	10301288	-
At_oRNA_511	2	2023293	2023992	+
At_oRNA_512	2	16346979	16347358	+
At_oRNA_513	2	2184427	2185198	-
At_oRNA_514	5	4020935	4021536	-
At_oRNA_515	2	353457	355522	-
At_oRNA_516	2	19227018	19227678	+
At_oRNA_517	3	17265429	17266147	+
At_oRNA_518	5	896769	897849	-
At_oRNA_519	2	17648863	17649884	+
At_oRNA_520	2	17648990	17650053	-
At_oRNA_521	2	17232599	17233040	+
At_oRNA_522	1	10308561	10309778	+
At_oRNA_523	1	25242163	25242741	+
At_oRNA_524	2	17697627	17698096	+
At_oRNA_525	1	3530506	3531755	+
At_oRNA_526	1	3530469	3531690	+
At_oRNA_527	3	6762136	6762705	-
At_oRNA_528	1	25873359	25873838	-
At_oRNA_529	4	9604528	9604983	+
At_oRNA_530	1	25124899	25125560	+
At_oRNA_531	2	11077711	11078522	-
At_oRNA_532	1	2407605	2408049	+
At_oRNA_533	1	12935841	12936386	-
At_oRNA_534	3	16012531	16013059	-
At_oRNA_535	5	2313130	2313563	-
At_oRNA_536	3	21132326	21132856	+
At_oRNA_537	3	16353203	16353999	-
At_oRNA_538	5	2940536	2941257	-
At_oRNA_539	1	19985090	19985617	+
At_oRNA_540	3	4136969	4137644	-
At_oRNA_541	3	1191532	1192642	-
At_oRNA_542	1	18552873	18553791	+
At_oRNA_543	4	11294272	11295440	+
At_oRNA_544	2	18577650	18579035	-
At_oRNA_545	2	2087230	2087999	+
At_oRNA_546	2	2087147	2088125	-
At_oRNA_547	1	21426126	21427008	+
At_oRNA_548	1	875280	875876	-
At_oRNA_549	1	1008511	1009134	+
At_oRNA_550	4	2452882	2453564	+
At_oRNA_551	3	14202696	14203794	-
At_oRNA_552	3	14205361	14206710	-
At_oRNA_553	3	14215258	14215698	-
At_oRNA_554	3	14204849	14205289	-

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_555	4	17282126	17282735	-
At_oRNA_556	1	11457146	11459168	+
At_oRNA_557	3	18903747	18904624	-
At_oRNA_558	5	21738958	21740063	+
At_oRNA_559	3	17196439	17197110	+
At_oRNA_560	4	9715787	9716356	+
At_oRNA_561	3	790187	790648	-
At_oRNA_562	1	17255410	17256342	+
At_oRNA_563	4	2581840	2582276	-
At_oRNA_564	5	16432630	16433057	-
At_oRNA_565	3	23470202	23470721	-
At_oRNA_566	2	3624510	3624718	-
At_oRNA_567	3	14211107	14211782	+
At_oRNA_568	2	6150	6825	+
At_oRNA_569	1	2656689	2657060	-
At_oRNA_570	1	3744919	3745653	+
At_oRNA_571	4	425951	426652	-
At_oRNA_572	2	13453685	13453794	-
At_oRNA_573	1	7766108	7766518	+
At_oRNA_574	3	1848893	1849302	+
At_oRNA_575	1	9538738	9539033	-
At_oRNA_576	1	25061143	25061924	-
At_oRNA_577	1	28132938	28133434	+
At_oRNA_578	2	6399301	6400093	+
At_oRNA_579	3	2317411	2318060	+
At_oRNA_580	3	5055389	5055895	-
At_oRNA_581	3	9666802	9667585	+
At_oRNA_582	3	22849077	22850045	-
At_oRNA_583	4	7846242	7846741	-
At_oRNA_584	5	17980557	17981079	+
At_oRNA_585	5	19089100	19089231	+
At_oRNA_586	3	12561836	12562506	-
At_oRNA_587	1	10300226	10300776	+
At_oRNA_588	2	13453124	13453826	+
At_oRNA_589	4	4866914	4868339	-
At_oRNA_590	5	10632246	10632752	+
At_oRNA_591	5	19640714	19641435	-

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

S6 ESTs per orphan transcript. This information can be downloaded as a plain-text file

from http://atornadb.bio.uni-potsdam.de/At_oRNA_DB_ESTs -04042005.txt

At oRNA ID	ESTs (GenBank Identifier)
At_oRNA_001	1015440
At_oRNA_002	1015441
At_oRNA_003	10277179
At_oRNA_004	1033257
At_oRNA_005	1033276
At_oRNA_006	1082057
At_oRNA_007	1100922
At_oRNA_008	1103332
At_oRNA_009	1130576
At_oRNA_010	1154983
At_oRNA_011	1159090
At_oRNA_012	1159607
At_oRNA_013	1184224
At_oRNA_014	1215862
At_oRNA_015	1215863
At_oRNA_016	1215912
At_oRNA_017	1215913
At_oRNA_018	1216944
At_oRNA_019	1217382
At_oRNA_020	1268915
At_oRNA_021	13382529
At_oRNA_022	13382583
At_oRNA_023	13382637
At_oRNA_024	13382678
At_oRNA_025	13382732
At_oRNA_026	13382903
At_oRNA_027	14580138
At_oRNA_028	14580203
At_oRNA_029	1520594
At_oRNA_030	1565764
At_oRNA_031	16784
At_oRNA_032	16936
At_oRNA_033	19740655
At_oRNA_034	19740731
At_oRNA_035	19740763
At_oRNA_036	19741097
At_oRNA_037	19741266
At_oRNA_038	19741274
At_oRNA_039	19741338
At_oRNA_040	19741376
At_oRNA_041	19741513
At_oRNA_042	19741590
At_oRNA_043	19741782
At_oRNA_044	19741805
At_oRNA_045	19741805
At_oRNA_046	19741902

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_047	19741965
At_oRNA_048	19742054
At_oRNA_049	19742059
At_oRNA_050	19742150
At_oRNA_051	19742559
At_oRNA_052	19742624
At_oRNA_053	19742730
At_oRNA_054	19742760
At_oRNA_055	19742800
At_oRNA_056	19742809
At_oRNA_057	19742820
At_oRNA_058	19742873
At_oRNA_059	19742892
At_oRNA_060	19742904
At_oRNA_061	19742927
At_oRNA_062	19742943
At_oRNA_063	19742952
At_oRNA_064	19743011
At_oRNA_065	19743230
At_oRNA_066	19743302
At_oRNA_067	19743408
At_oRNA_068	19743419
At_oRNA_069	19743437
At_oRNA_070	19743466
At_oRNA_071	19743506
At_oRNA_072	19796875
At_oRNA_073	19796968
At_oRNA_074	19797122
At_oRNA_075	19797854
At_oRNA_076	19797999
At_oRNA_077	19798068
At_oRNA_078	19798442
At_oRNA_079	19798666
At_oRNA_080	19798753
At_oRNA_081	19799009
At_oRNA_082	19799186
At_oRNA_083	19799249
At_oRNA_084	19799290
At_oRNA_085	19799291
At_oRNA_086	19799392
At_oRNA_087	19799510
At_oRNA_088	19799826
At_oRNA_089	19799988
At_oRNA_090	19800003
At_oRNA_091	19803945
At_oRNA_092	19806444
At_oRNA_093	19806444
At_oRNA_094	19828052
At_oRNA_095	19835898
At_oRNA_096	19839638
At_oRNA_097	19840034

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_098	19842202
At_oRNA_099	19847392
At_oRNA_100	19859666
At_oRNA_101	19862060
At_oRNA_102	19867921
At_oRNA_103	19867921
At_oRNA_104	19870344
At_oRNA_105	19871413
At_oRNA_106	19871451
At_oRNA_107	19871906
At_oRNA_108	19873486
At_oRNA_109	19874507
At_oRNA_110	19874777 8684154
At_oRNA_111	19875017
At_oRNA_112	19875500
At_oRNA_113	19875521
At_oRNA_114	19875574
At_oRNA_115	19875895
At_oRNA_116	19875996
At_oRNA_117	19876232
At_oRNA_118	19876391
At_oRNA_119	19876425
At_oRNA_120	19876428
At_oRNA_121	19876848
At_oRNA_122	19876854
At_oRNA_123	19876859
At_oRNA_124	19876888
At_oRNA_125	19876913
At_oRNA_126	19876944
At_oRNA_127	19877207
At_oRNA_128	19877276
At_oRNA_129	19877610
At_oRNA_130	19877839
At_oRNA_131	19878285
At_oRNA_132	19878697
At_oRNA_133	19878941
At_oRNA_134	19879089
At_oRNA_135	19879102
At_oRNA_136	19879191
At_oRNA_137	2047697
At_oRNA_138	2047770
At_oRNA_139	2048913
At_oRNA_140	2062902
At_oRNA_141	23301862
At_oRNA_142	23302811
At_oRNA_143	23303534
At_oRNA_144	2393399
At_oRNA_145	2412930
At_oRNA_146	2413184
At_oRNA_147	2413405
At_oRNA_148	2413451

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_149	2413861
At_oRNA_150	2446159
At_oRNA_151	2596168
At_oRNA_152	2597098
At_oRNA_153	2704872
At_oRNA_154	2722977
At_oRNA_155	2733600
At_oRNA_156	2733883
At_oRNA_157	2748261
At_oRNA_158	2748489
At_oRNA_159	2748970
At_oRNA_160	2749143
At_oRNA_161	2749496
At_oRNA_162	2763995
At_oRNA_163	2764018
At_oRNA_164	2764018
At_oRNA_165	28197627
At_oRNA_166	315316
At_oRNA_167	315953
At_oRNA_168	3236403
At_oRNA_169	3868978
At_oRNA_170	397199
At_oRNA_171	397262
At_oRNA_172	402952
At_oRNA_173	402952
At_oRNA_174	402976
At_oRNA_175	406619
At_oRNA_176	406804
At_oRNA_177	406815
At_oRNA_178	406818
At_oRNA_179	406821
At_oRNA_180	406822
At_oRNA_181	406845
At_oRNA_182	406845
At_oRNA_183	406846
At_oRNA_184	406846
At_oRNA_185	4210214
At_oRNA_186	4210226
At_oRNA_187	4714038
At_oRNA_188	493273
At_oRNA_189	496824
At_oRNA_190	506982
At_oRNA_191	509099
At_oRNA_192	510775
At_oRNA_193	510792
At_oRNA_194	510812
At_oRNA_195	511165
At_oRNA_196	520971
At_oRNA_197	550269
At_oRNA_198	550269
At_oRNA_199	550269 550270

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_200	550269 550270
At_oRNA_201	557607
At_oRNA_202	563841
At_oRNA_203	572566
At_oRNA_204	5842120
At_oRNA_205	5842279
At_oRNA_206	5842356
At_oRNA_207	5842364
At_oRNA_208	5842378
At_oRNA_209	5842507
At_oRNA_210	5842507
At_oRNA_211	5842507
At_oRNA_212	5842507
At_oRNA_213	5842579
At_oRNA_214	5842893
At_oRNA_215	5842908
At_oRNA_216	5842952
At_oRNA_217	5843096
At_oRNA_218	5843126
At_oRNA_219	5843286
At_oRNA_220	5843850
At_oRNA_221	5843938
At_oRNA_222	5844072
At_oRNA_223	5844165
At_oRNA_224	5844252
At_oRNA_225	5844270
At_oRNA_226	5844508
At_oRNA_227	5844538
At_oRNA_228	5844553
At_oRNA_229	5844949
At_oRNA_230	5845068
At_oRNA_231	5845921
At_oRNA_232	5845961
At_oRNA_233	5846147
At_oRNA_234	5846269
At_oRNA_235	5846437
At_oRNA_236	5846462
At_oRNA_237	5846466
At_oRNA_238	5846573
At_oRNA_239	5846714
At_oRNA_240	5851225
At_oRNA_241	5851231
At_oRNA_242	5851300
At_oRNA_243	5851300
At_oRNA_244	5851450
At_oRNA_245	599631
At_oRNA_246	623023
At_oRNA_247	6536645
At_oRNA_248	6536646
At_oRNA_249	6536647
At_oRNA_250	6536648

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_251	671783
At_oRNA_252	671805
At_oRNA_253	683518
At_oRNA_254	7610309
At_oRNA_255	7610453
At_oRNA_256	7610453
At_oRNA_257	7610472
At_oRNA_258	7610479
At_oRNA_259	7610784
At_oRNA_260	7610784
At_oRNA_261	7610913
At_oRNA_262	7611466
At_oRNA_263	7611865
At_oRNA_264	7611927
At_oRNA_265	7611947
At_oRNA_266	7612817
At_oRNA_267	7612839
At_oRNA_268	8333168
At_oRNA_269	8333326
At_oRNA_270	8334040
At_oRNA_271	8334560
At_oRNA_272	8678120
At_oRNA_273	8680752
At_oRNA_274	8682798
At_oRNA_275	8682877
At_oRNA_276	8683067
At_oRNA_277	8683775
At_oRNA_278	8683778
At_oRNA_279	8684162
At_oRNA_280	8687056
At_oRNA_281	8687636
At_oRNA_282	8687737
At_oRNA_283	8688550
At_oRNA_284	8688994
At_oRNA_285	8689054
At_oRNA_286	8689164
At_oRNA_287	8689414
At_oRNA_288	8689818
At_oRNA_289	8689869
At_oRNA_290	8690356
At_oRNA_291	8690514
At_oRNA_292	8690549
At_oRNA_293	8691834
At_oRNA_294	8692583
At_oRNA_295	8692684
At_oRNA_296	8692759
At_oRNA_297	8692848
At_oRNA_298	8693246
At_oRNA_299	8693246
At_oRNA_300	8693292
At_oRNA_301	8693699

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_302	8693699
At_oRNA_303	8693873
At_oRNA_304	8694092
At_oRNA_305	8694375
At_oRNA_306	8694639
At_oRNA_307	8694729
At_oRNA_308	8695760
At_oRNA_309	8696141
At_oRNA_310	8696817
At_oRNA_311	8696868
At_oRNA_312	8696873
At_oRNA_313	8697049
At_oRNA_314	8697356
At_oRNA_315	8697357
At_oRNA_316	8697526
At_oRNA_317	8715908
At_oRNA_318	8717115
At_oRNA_319	8717415
At_oRNA_320	8717626
At_oRNA_321	8717743
At_oRNA_322	8718485
At_oRNA_323	8718506
At_oRNA_324	8718587
At_oRNA_325	8718702
At_oRNA_326	8718740
At_oRNA_327	8719126
At_oRNA_328	8719781
At_oRNA_329	8720096
At_oRNA_330	8720437
At_oRNA_331	8721541
At_oRNA_332	8722460
At_oRNA_333	8724604
At_oRNA_334	8726214
At_oRNA_335	8726230
At_oRNA_336	8726335
At_oRNA_337	8726357
At_oRNA_338	8727547
At_oRNA_339	8728044
At_oRNA_340	8728082
At_oRNA_341	8728339
At_oRNA_342	8728861
At_oRNA_343	8728990
At_oRNA_344	8729283
At_oRNA_345	8729497
At_oRNA_346	8729759
At_oRNA_347	8729815
At_oRNA_348	8729870
At_oRNA_349	8730424
At_oRNA_350	8730975
At_oRNA_351	8733236
At_oRNA_352	8734364

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_353	8734543
At_oRNA_354	8735027
At_oRNA_355	8735690
At_oRNA_356	8735931
At_oRNA_357	8736152
At_oRNA_358	8736268
At_oRNA_359	8736948
At_oRNA_360	8737413
At_oRNA_361	8737467
At_oRNA_362	8738331
At_oRNA_363	895720
At_oRNA_364	930842
At_oRNA_365	937703
At_oRNA_366	9779283
At_oRNA_367	9779304
At_oRNA_368	9779373
At_oRNA_369	9779632
At_oRNA_370	9779638
At_oRNA_371	9779955
At_oRNA_372	9779996
At_oRNA_373	9780082
At_oRNA_374	9781491
At_oRNA_375	9781839
At_oRNA_376	9781856
At_oRNA_377	9781950
At_oRNA_378	9782160
At_oRNA_379	9782390
At_oRNA_380	9782437
At_oRNA_381	9782637
At_oRNA_382	9782699
At_oRNA_383	9782700
At_oRNA_384	9782745
At_oRNA_385	9783031
At_oRNA_386	9783221
At_oRNA_387	9783725
At_oRNA_388	9783872
At_oRNA_389	9783936
At_oRNA_390	9784288
At_oRNA_391	9784288
At_oRNA_392	9785329
At_oRNA_393	9785614
At_oRNA_394	9785793
At_oRNA_395	9785899
At_oRNA_396	9783494 9786653
At_oRNA_397	9783494 9786653
At_oRNA_398	9783494 9786653
At_oRNA_399	9787061
At_oRNA_400	9788338
At_oRNA_401	9788368
At_oRNA_402	9788526
At_oRNA_403	9788581

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_404	19801981 19857043
At_oRNA_405	19743358 19830555
At_oRNA_406	19743023 19799406
At_oRNA_407	19742891 2413021 2580978
At_oRNA_408	19741812 8689194
At_oRNA_409	19741799 2393664
At_oRNA_410	19741739 2161009
At_oRNA_411	13382569 13382799 19867766
At_oRNA_412	10277136 2762744 5851277 8681781
At_oRNA_413	19806747 9784159
At_oRNA_414	9783483 9786658
At_oRNA_415	9782134 9782189
At_oRNA_416	9781290 9785038
At_oRNA_417	9779362 9779363
At_oRNA_418	19743251 9779067
At_oRNA_419	19742889 9779035
At_oRNA_420	9779033 9779034 9784446 9788652
At_oRNA_421	9778551 9778552
At_oRNA_422	23303047 23303893 26019573
At_oRNA_423	23303047 23303893 26019573
At_oRNA_424	9778335 9778483
At_oRNA_425	9778135 9788590
At_oRNA_426	9778135 9788590
At_oRNA_427	16910 16911 8738587
At_oRNA_428	19799034 8735250
At_oRNA_429	19796940 8732578 8737237
At_oRNA_430	5851261 8690161 8729866
At_oRNA_431	8726431 8734995
At_oRNA_432	19798789 8718014 8719116
At_oRNA_433	2759179 8717420
At_oRNA_434	8717367 8717561
At_oRNA_435	8716789 8730037
At_oRNA_436	8716789 8730037
At_oRNA_437	19826107 8715811
At_oRNA_438	19743549 2758369 8715557
At_oRNA_439	19742940 403274 8702189
At_oRNA_440	19740822 2722062 2722146 8701557
At_oRNA_441	2048777 2412987 2412990 2576649 8697922
At_oRNA_442	5844061
At_oRNA_443	28197487 28197489 28197490
At_oRNA_444	5844266 8697455
At_oRNA_445	2412942 8696432
At_oRNA_446	5842615 8696372 8733350 8734790
At_oRNA_447	19865036 8690437 8717325 8726193 8726307 8726331
At_oRNA_448	2393167 8690164
At_oRNA_449	19875941 8690156

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_450	8688993 8689011
At_oRNA_451	8685364 9785753
At_oRNA_452	8682382 8729338 8734812
At_oRNA_453	19740972 8682166
At_oRNA_454	8682014 8717018 8730110
At_oRNA_455	8681766 8717622 8718261 8719663 8719944
At_oRNA_456	8678749 8719870
At_oRNA_457	8334147
At_oRNA_458	8334147
At_oRNA_459	7611257 7613156 8723634
At_oRNA_460	7612126 8696992
At_oRNA_461	19741384 7611394 8682408 8689816
At_oRNA_462	7610606 8720197
At_oRNA_463	2722740 5846379
At_oRNA_464	2393315
At_oRNA_465	5844716 8737354
At_oRNA_466	19857712
At_oRNA_467	5843570 8715777
At_oRNA_468	19742343 5843196
At_oRNA_469	10277200 19799917 5843190 7610612 8736758 8738013
At_oRNA_470	19799255 2749194 5843148
At_oRNA_471	5843056 8692051
At_oRNA_472	19799318 19860648 2722956 5842949
At_oRNA_473	19743543 5840629
At_oRNA_474	2413835 5839501 8677996 8720061
At_oRNA_475	5839468
At_oRNA_476	19875024
At_oRNA_477	2763551 5840302 5841321
At_oRNA_478	19876018 8696833 9779887
At_oRNA_479	5840492 937818
At_oRNA_480	2759848 8682509
At_oRNA_481	19860601 502132
At_oRNA_482	2748157 2762780
At_oRNA_483	19742887 2748146 8717638
At_oRNA_484	1159403 19861145 19874167 5841680 9778428
At_oRNA_485	19799806 19878925
At_oRNA_486	1054409 5839323
At_oRNA_487	19878426 2733845
At_oRNA_488	5840683 906064 9785012
At_oRNA_489	19741509 2722921 8717831
At_oRNA_490	2722079 935066
At_oRNA_491	19799452 19878605
At_oRNA_492	8693244 930679
At_oRNA_493	19799426 19878580 397253
At_oRNA_494	19877040 2445844
At_oRNA_495	19799041 19878496

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_496	19797106
At_oRNA_497	1217065 19799227 19878401 5842341 8681087 8681458 8685606 8724411 8728282 8736416
At_oRNA_498	1217065 19799227 19878401 5842341 8681087 8681458 8685606 8724411 8728282 8736416
At_oRNA_499	8696027 895736 895737
At_oRNA_500	19799171 19878352 8716903
At_oRNA_501	19799096 19878279
At_oRNA_502	416463 493415
At_oRNA_503	19798977 19878174
At_oRNA_504	19798977 19878174
At_oRNA_505	19798954 19878151
At_oRNA_506	19742756 19876257 398624 406613 435180 435181
At_oRNA_507	19798818 19878031
At_oRNA_508	19798817 19878030
At_oRNA_509	19798785 19877997
At_oRNA_510	19798581 19877879 8716034
At_oRNA_511	19798436 19877826
At_oRNA_512	19798405 19877809
At_oRNA_513	19798259 19877730
At_oRNA_514	19798245 19877720 8717905 8718044
At_oRNA_515	19798237 19877714
At_oRNA_516	19797341 19877249 8694119
At_oRNA_517	19797341 19877249 8694119
At_oRNA_518	19742804 19876302
At_oRNA_519	19798211
At_oRNA_520	19876219
At_oRNA_521	19742520 19876038
At_oRNA_522	19742497 19876017
At_oRNA_523	19742451 19875983
At_oRNA_524	19742324 19875880
At_oRNA_525	19799499 19878648
At_oRNA_526	19742311 19875869
At_oRNA_527	19742277 19875841
At_oRNA_528	19742133 19875704
At_oRNA_529	19741996 19875602
At_oRNA_530	19741958 19875567
At_oRNA_531	1103355 1520870 19741937 19875549 2596503 5839479 8333313
At_oRNA_532	19741917 19875533 2412861
At_oRNA_533	19741863 19875493
At_oRNA_534	19875446 502134
At_oRNA_535	19741742 19875422
At_oRNA_536	19741727 19875410
At_oRNA_537	19741723 19875406 2412795

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_538	19741550 19875263 8691667
At_oRNA_539	19741461 19875186
At_oRNA_540	19741355 19800046 19875094 19879137 2722532 2757418 5839922
At_oRNA_541	19741262 19875013 8689131 8718596
At_oRNA_542	19741233 19874989 5844370
At_oRNA_543	19741155 19874917
At_oRNA_544	19741092 19874869
At_oRNA_545	5851223
At_oRNA_546	19740665 19874517
At_oRNA_547	19740639 19874491
At_oRNA_548	19872980 19876601
At_oRNA_549	19872980 19876601
At_oRNA_550	19871256 19872354
At_oRNA_551	19837785 19838375 19843670 19871048
At_oRNA_552	19837785 19838375 19843670 19871048
At_oRNA_553	2748432
At_oRNA_554	2748432
At_oRNA_555	19869433 5842822
At_oRNA_556	19868744 28251363
At_oRNA_557	19825847 19867718
At_oRNA_558	19823150 19867051
At_oRNA_559	19803895 19866058 2048551 2758514
At_oRNA_560	19865749 19870519
At_oRNA_561	19803280 19865567
At_oRNA_562	19801290 19863947 2749266 757548 757567 937320
At_oRNA_563	19841931 19844939
At_oRNA_564	19836100 19846974
At_oRNA_565	19830351 2739594 2747409 3449910
At_oRNA_566	493284 506983
At_oRNA_567	1053739 1269534 1327809 2047375 2048122 23302014 2412786 2412957 2723192 2747696 2763478 8689187 8726147 9787859
At_oRNA_568	1053739 1269534 1327809 2047375 2048122 23302014 2412786 2412957 2723192 2747696 2763478 8689187 8726147 9787859
At_oRNA_569	19743515 8717680
At_oRNA_570	19741560 19875270 8333229
At_oRNA_571	19741560 19875270 8333229
At_oRNA_572	19798900
At_oRNA_573	19877136 2413262 2722754
At_oRNA_574	19877136 2413262 2722754

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_575	9784471 9784738 9788681
At_oRNA_576	19741166 19874926 5843900 895707 902613
At_oRNA_577	7610422 8678279
At_oRNA_578	1054453 1158388 1520866 19799023 19802321 19803758 19804616 19824250 19829159 19864764 2757247 2762571 558159 5846171 8332760 8691821 8693518 8714779 8714801 8715867 8723501 906969 935042 937922 9785365 9786101 9786964
At_oRNA_579	1327901 5843560
At_oRNA_580	19825956 2722250 8719172
At_oRNA_581	1517303 19799797 19878917 2763617 937891
At_oRNA_582	2062939 8729312 8732740
At_oRNA_583	2062816 5844268 8682642
At_oRNA_584	19742343 5843196 8693699
At_oRNA_585	8726233 8726252 8726353
At_oRNA_586	11170749 1159033 8691145 8693313 8694505 8697334
At_oRNA_587	2759966 8724383
At_oRNA_588	1159090 19740713 19740880 19798900 19860060 19873970 19878108 2722157 2739556 5842322
At_oRNA_589	19741128 5844424 8333680 8694359 8727208 8734168 9782044
At_oRNA_590	2762715 930491
At_oRNA_591	19740713 19740880 19798900 19860060 19873970 19878108 2722157 2739556 5842322

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

S7 Interesting cases about some orphan transcripts that overlap annotated genes

There are 17 oRNAs that still have a partial overlap with known genes. Fifteen of them overlap in less than 50 bp with known transcriptional units. The remaining two oRNAs which overlap in more than 50 bp with known transcriptional units might be interpreted as parts of these. However, at closer inspection, this possible interpretation is questioned. The region where At_oRNA_133 is located overlaps completely with the region containing the annotated gene At3g46670 (http://atornadb.bio.uni-potsdam.de/cgi-bin/gbrowse/arab?name=At_oRNA_133). However, At3g46670 appears to be located in an intron of At_oRNA_133. A similar case is found for At_oRNA_556 with the annotated gene At1g31900 (http://atornadb.bio.uni-potsdam.de/cgi-bin/gbrowse/arab?name=At_oRNA_556).

At_oRNA_519 has an overlap of 4 bp with At2g42360 (http://atornadb.bio.uni-potsdam.de/cgi-bin/gbrowse/arab?name=At_oRNA_519), still we have proposed it to be an independent transcriptional unit for which we have predicted a protein-coding region. It is also interesting that in the region where At_oRNA_519 maps, At_oRNA_229 and At_oRNA_520 also map, but with different spliced alignments, which generates different processed transcripts.

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

S8 List of orphan transcripts overlapping with known transcriptional units

Orphan transcript	Overlapping with
At_oRNA_017	At5g58595
At_oRNA_090	At5g10200
At_oRNA_092	At2g01010
At_oRNA_093	At3g41768
At_oRNA_119	At4g39361
At_oRNA_119	At4g39363
At_oRNA_125	At1g07702
At_oRNA_131	At2g07712
At_oRNA_132	At3g26920
At_oRNA_133	At3g46670
At_oRNA_140	At2g35744
At_oRNA_143	At4g04750
At_oRNA_144	At4g04610
At_oRNA_234	At1g25360
At_oRNA_314	At3g26910
At_oRNA_344	At1g64760
At_oRNA_353	At4g31140
At_oRNA_401	At3g47342
At_oRNA_426	At2g07764
At_oRNA_439	At4g39364
At_oRNA_439	At4g39366
At_oRNA_466	At5g26742
At_oRNA_500	At3g47348
At_oRNA_505	At3g52540
At_oRNA_508	At2g47900
At_oRNA_513	At2g05765
At_oRNA_519	At2g42360
At_oRNA_529	At4g17090
At_oRNA_531	At2g25950
At_oRNA_548	At1g03502
At_oRNA_548	At1g03506
At_oRNA_549	At1g03935
At_oRNA_556	At1g31900
At_oRNA_557	At3g50825

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

S9 Annotated genes closest to orphan transcripts

Annotated genes	Orphan transcript	Chr	Distance to left gene (bp)	Distance to right gene (bp)
At5g66320-At5g66330	At_oRNA_001	5	2043	195
At5g66330-At5g66340	At_oRNA_002	5	146	139
At5g50870-At5g50880	At_oRNA_003	5	11	242
At3g44280-At3g44290	At_oRNA_004	3	1541	46
At3g04290-At3g04300	At_oRNA_005	3	2895	797
At5g23910-At5g23920	At_oRNA_006	5	71	306
At4g29720-At4g29730	At_oRNA_007	4	3815	72
At3g22180-At3g22190	At_oRNA_008	3	111	567
At1g74200-At1g74210	At_oRNA_009	1	571	263
At2g18320-At2g18330	At_oRNA_010	2	7169	750
At5g48410-At5g48420	At_oRNA_011	5	244	2171
At3g21260-At3g21270	At_oRNA_012	3	6048	3185
At2g35460-At2g35470	At_oRNA_013	2	872	1374
At1g15090-At1g15100	At_oRNA_014	1	2569	806
At1g15090-At1g15100	At_oRNA_015	1	2130	1226
At5g58590-At5g58595	At_oRNA_016	5	628	243
At3g29642-At3g29643	At_oRNA_018	3	277	3789
At4g26250-At4g26260	At_oRNA_019	4	3898	2526
At1g64560-At1g64570	At_oRNA_020	1	2366	1873
At1g08030-At1g08035	At_oRNA_021	1	761	1632
At2g07784-At2g07755	At_oRNA_022	2	3807	5984
At5g12340-At5g12350	At_oRNA_023	5	1785	175
At3g22260-At3g22270	At_oRNA_024	3	230	177
At2g20040-At2g20050	At_oRNA_025	2	430	245
At3g58580-At3g58590	At_oRNA_026	3	67	2045
At3g61270-At3g61280	At_oRNA_027	3	768	74
At4g23210-At4g23220	At_oRNA_028	4	1285	974
At5g46180-At5g46190	At_oRNA_029	5	512	773
At5g12160-At5g12170	At_oRNA_030	5	667	298
At5g65140-At5g65160	At_oRNA_031	5	8148	567
At2g07724-At2g07725	At_oRNA_032	2	1067	294
At3g25790-At3g25800	At_oRNA_033	3	2550	4613
At4g40060-At4g40070	At_oRNA_034	4	1429	1691
At4g07507-At4g07510	At_oRNA_035	4	610	267
At1g67190-At1g67200	At_oRNA_036	1	2777	2762
At3g21290-At3g21300	At_oRNA_037	3	3224	836
At3g06770-At3g06780	At_oRNA_038	3	2879	2937
At1g74110-At1g74120	At_oRNA_039	1	2908	198
At1g62975-At1g62980	At_oRNA_040	1	25	899
At1g08210-At1g08220	At_oRNA_041	1	376	299
At5g44570-At5g44575	At_oRNA_042	5	5166	2391
At5g22920-At5g22930	At_oRNA_043	5	505	19
At1g52340-At1g52350	At_oRNA_044	1	5030	738
At5g02350-At5g02360	At_oRNA_045	5	1366	1741
At3g23250-At3g23260	At_oRNA_046	3	1844	259

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At1g61790-At1g61800	At_oRNA_047	1	4725	3442
At2g30780-At2g30790	At_oRNA_048	2	288	238
At2g46970-At2g46980	At_oRNA_049	2	1666	1935
At1g52615-At1g52620	At_oRNA_050	1	1267	245
At5g66550-At5g66560	At_oRNA_051	5	2786	19
At1g77950-At1g77960	At_oRNA_052	1	520	770
At4g00020-At4g00030	At_oRNA_053	4	912	1783
At3g17660-At3g17670	At_oRNA_054	3	438	330
At5g02690-At5g02700	At_oRNA_055	5	241	1124
At5g42210-At5g42220	At_oRNA_056	5	1090	1466
At1g53785-At1g53790	At_oRNA_057	1	79	131
At4g14540-At4g14550	At_oRNA_058	4	2100	65
At5g16640-At5g16650	At_oRNA_059	5	45	163
At3g52050-At3g52060	At_oRNA_060	3	290	231
At5g63770-At5g63780	At_oRNA_061	5	347	239
At4g36790-At4g36800	At_oRNA_062	4	2064	119
At5g32460-At5g32470	At_oRNA_063	5	751	40
At4g03110-At4g03115	At_oRNA_064	4	3036	452
At5g56310-At5g56320	At_oRNA_065	5	36	4266
At1g74670-At1g74680	At_oRNA_066	1	180	4541
At4g00760-At4g00770	At_oRNA_067	4	530	1162
At3g07210-At3g07220	At_oRNA_068	3	697	1073
At4g32440-At4g32450	At_oRNA_069	4	166	868
At1g06570-At1g06580	At_oRNA_070	1	460	234
At3g03140-At3g03150	At_oRNA_071	3	17	58
At2g25090-At2g25100	At_oRNA_072	2	2617	3866
At5g40910-At5g40920	At_oRNA_073	5	338	416
At2g46450-At2g46455	At_oRNA_074	2	793	61
At1g63840-At1g63850	At_oRNA_075	1	5768	63
At2g35950-At2g35960	At_oRNA_076	2	162	851
At3g13400-At3g13410	At_oRNA_077	3	751	3351
At5g60250-At5g60260	At_oRNA_078	5	57	98
At1g56510-At1g56520	At_oRNA_079	1	2512	460
At1g10000-At1g10010	At_oRNA_080	1	24	794
At2g12460-At2g12470	At_oRNA_081	2	2916	11719
At2g15310-At2g15320	At_oRNA_082	2	9301	1059
At1g52360-At1g52370	At_oRNA_083	1	360	319
At1g35220-At1g35230	At_oRNA_084	1	515	2344
At5g18240-At5g18250	At_oRNA_085	5	2012	482
At1g17230-At1g17240	At_oRNA_086	1	198	966
At1g53230-At1g53240	At_oRNA_087	1	22	2487
At3g46670-At3g46680	At_oRNA_088	3	661	37
At1g31170-At1g31175	At_oRNA_089	1	4124	1583
At1g14650-At1g14660	At_oRNA_091	1	755	192
At3g03550-At3g03560	At_oRNA_094	3	1351	290
At4g15800-At4g15802	At_oRNA_095	4	75	736
At3g60310-At3g60320	At_oRNA_096	3	50	579
At2g35850-At2g35860	At_oRNA_097	2	2348	15
At4g16110-At4g16120	At_oRNA_098	4	105	121
At4g34470-At4g34480	At_oRNA_099	4	456	565
At1g63680-At1g63690	At_oRNA_100	1	166	610

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At1g65485-At1g65490	At_oRNA_101	1	2476	4522
At2g41640-At2g41650	At_oRNA_102	2	93	809
At5g18930-At5g18940	At_oRNA_103	5	10810	1268
At5g13490-At5g13500	At_oRNA_104	5	314	59
At1g07200-At1g07210	At_oRNA_105	1	1460	2689
At1g01440-At1g01450	At_oRNA_106	1	355	1347
At5g55040-At5g55050	At_oRNA_107	5	922	5887
At4g30790-At4g30800	At_oRNA_108	4	510	2515
At3g25790-At3g25800	At_oRNA_109	3	3058	3893
At1g56610-At1g56620	At_oRNA_110	1	232	4425
At3g21290-At3g21300	At_oRNA_111	3	301	3731
At4g18110-At4g18120	At_oRNA_112	4	408	626
At3g23250-At3g23260	At_oRNA_113	3	971	1139
At1g61790-At1g61800	At_oRNA_114	1	4267	3994
At2g10930-At2g10940	At_oRNA_115	2	1508	4467
At5g52880-At5g52890	At_oRNA_116	5	4436	6243
At4g00020-At4g00030	At_oRNA_117	4	1691	180
At1g71360-At1g71370	At_oRNA_118	1	185	2202
At4g36790-At4g36800	At_oRNA_120	4	322	1643
At4g00770-At4g00780	At_oRNA_121	4	678	245
At2g13420-At2g13430	At_oRNA_122	2	38	1786
At3g07210-At3g07220	At_oRNA_123	3	1629	145
At5g28260-At5g28263	At_oRNA_124	5	2545	2128
At5g14220-At5g14230	At_oRNA_126	5	4339	45
At5g59700-At5g59710	At_oRNA_127	5	1982	264
At2g44800-At2g44810	At_oRNA_128	2	2219	7775
At3g13400-At3g13410	At_oRNA_129	3	1076	2874
At3g25882-At3g25890	At_oRNA_130	3	4461	492
At1g31170-At1g31175	At_oRNA_134	1	2960	2566
At5g10190-At5g10200	At_oRNA_135	5	155	734
At5g36000-At5g36005	At_oRNA_136	5	1212	780
At1g62305-At1g62310	At_oRNA_137	1	5158	1139
At5g14030-At5g14035	At_oRNA_138	5	174	1420
At5g46500-At5g46510	At_oRNA_139	5	164	1938
At3g33076-At3g33069	At_oRNA_141	3	43	65894
At2g20590-At2g20595	At_oRNA_142	2	397	1280
At5g51620-At5g51630	At_oRNA_145	5	1004	116
At3g52527-At3g52530	At_oRNA_146	3	5979	77
At5g43400-At5g43405	At_oRNA_147	5	2446	1398
At3g27883-At3g27886	At_oRNA_148	3	1754	1112
At5g55780-At5g55790	At_oRNA_149	5	49	1407
At5g27885-At5g27890	At_oRNA_150	5	6023	152
At3g22990-At3g23000	At_oRNA_151	3	5244	688
At1g48740-At1g48750	At_oRNA_152	1	4832	4875
At5g05550-At5g05560	At_oRNA_153	5	50	609
At5g15260-At5g15270	At_oRNA_154	5	1777	751
At1g80800-At1g80810	At_oRNA_155	1	169	405
At4g22580-At4g22590	At_oRNA_156	4	1209	1486
At5g10490-At5g10500	At_oRNA_157	5	414	636
At5g26140-At5g26150	At_oRNA_158	5	4628	847
At1g79740-At1g79750	At_oRNA_159	1	192	461

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At2g01505-At2g01510	At_oRNA_160	2	226	924
At4g15950-At4g15955	At_oRNA_161	4	42	272
At3g29170-At3g29175	At_oRNA_162	3	323	3940
At1g76990-At1g77000	At_oRNA_163	1	720	4111
At1g76700-At1g76705	At_oRNA_164	1	1886	897
At5g51460-At5g51470	At_oRNA_165	5	311	2302
At4g28365-At4g28370	At_oRNA_166	4	814	525
At1g66780-At1g66790	At_oRNA_167	1	1253	4468
At5g39040-At5g39050	At_oRNA_168	5	2227	2230
At3g59770-At3g59780	At_oRNA_169	3	330	51
At1g56420-At1g56430	At_oRNA_170	1	5810	6654
At3g15120-At3g15130	At_oRNA_171	3	26	1342
At2g44540-At2g44550	At_oRNA_172	2	332	158
At2g44550-At2g44560	At_oRNA_173	2	322	131
At2g07718-At2g07719	At_oRNA_174	2	3992	1214
At1g14750-At1g14760	At_oRNA_175	1	1832	10
At2g07696-At2g07697	At_oRNA_176	2	1203	596
At2g07756-At2g07724	At_oRNA_177	2	488	1138
At4g32920-At4g32930	At_oRNA_178	4	1585	825
At1g02620-At1g02630	At_oRNA_179	1	360	2276
At1g02610-At1g02620	At_oRNA_180	1	813	36
At2g07702-At2g07703	At_oRNA_181	2	1613	445
At2g07766-At2g07680	At_oRNA_182	2	5328	6724
At2g07793-At2g07701	At_oRNA_183	2	1555	521
At2g07766-At2g07680	At_oRNA_184	2	1555	10483
At1g24145-At1g24150	At_oRNA_185	1	706	6416
At2g23810-At2g23820	At_oRNA_186	2	133	3128
At5g27150-At5g27160	At_oRNA_187	5	2828	2164
At5g64320-At5g64330	At_oRNA_188	5	160	1380
At5g39740-At5g39750	At_oRNA_189	5	506	648
At3g08560-At3g08570	At_oRNA_190	3	21	101
At5g41760-At5g41765	At_oRNA_191	5	181	6642
At2g07714-At2g07715	At_oRNA_192	2	1416	929
At3g60610-At3g60620	At_oRNA_193	3	920	101
At2g07671-At2g07779	At_oRNA_194	2	1043	1279
At5g62770-At5g62780	At_oRNA_195	5	1126	310
At1g48340-At1g48350	At_oRNA_196	1	101	137
At2g07766-At2g07680	At_oRNA_197	2	431	11517
At2g07793-At2g07701	At_oRNA_198	2	431	1554
At2g07793-At2g07701	At_oRNA_199	2	947	1038
At2g07766-At2g07680	At_oRNA_200	2	947	11000
At5g65920-At5g65925	At_oRNA_201	5	1585	1171
At4g36880-At4g36890	At_oRNA_202	4	1184	1547
At4g25410-At4g25420	At_oRNA_203	4	433	3471
At3g17250-At3g17260	At_oRNA_204	3	870	370
At1g64940-At1g64950	At_oRNA_205	1	54	2091
At1g68860-At1g68870	At_oRNA_206	1	195	56
At1g08160-At1g08170	At_oRNA_207	1	1224	836
At1g61420-At1g61430	At_oRNA_208	1	324	157
At1g78030-At1g78040	At_oRNA_209	1	1000	592
At5g64685-At5g64690	At_oRNA_210	5	1933	3060

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At5g63620-At5g63630	At_oRNA_211	5	2945	657
At5g65230-At5g65240	At_oRNA_212	5	3978	860
At3g13880-At3g13890	At_oRNA_213	3	1725	73
At5g65040-At5g65050	At_oRNA_214	5	2130	1369
At1g07710-At1g07720	At_oRNA_215	1	48	2517
At1g05240-At1g05250	At_oRNA_216	1	1262	1167
At3g08670-At3g08680	At_oRNA_217	3	122	870
At1g12670-At1g12680	At_oRNA_218	1	1191	900
At1g12015-At1g12020	At_oRNA_219	1	2789	1010
At3g04210-At3g04220	At_oRNA_220	3	769	44
At1g08170-At1g08180	At_oRNA_221	1	920	57
At3g13430-At3g13435	At_oRNA_222	3	5551	739
At4g32640-At4g32650	At_oRNA_223	4	154	471
At1g34460-At1g34470	At_oRNA_224	1	1503	2472
At1g24145-At1g24150	At_oRNA_225	1	1290	6103
At5g36870-At5g36880	At_oRNA_226	5	678	26
At1g77655-At1g77660	At_oRNA_227	1	2714	104
At1g26550-At1g26560	At_oRNA_228	1	2672	821
At2g42360-At2g42370	At_oRNA_229	2	112	1002
At2g25580-At2g25590	At_oRNA_230	2	421	705
At3g16390-At3g16400	At_oRNA_231	3	63	1796
At1g59970-At1g59980	At_oRNA_232	1	1755	3815
At5g23150-At5g23160	At_oRNA_233	5	1653	398
At4g16745-At4g16750	At_oRNA_235	4	6804	151
At3g25220-At3g25225	At_oRNA_236	3	8	65
At4g20740-At4g20760	At_oRNA_237	4	33	143
At1g76060-At1g76070	At_oRNA_238	1	897	1047
At3g57870-At3g57880	At_oRNA_239	3	659	14
At2g42380-At2g42390	At_oRNA_240	2	1528	241
At2g17240-At2g17250	At_oRNA_241	2	24	200
At2g07678-At2g07781	At_oRNA_242	2	5100	1247
At2g07773-At2g07747	At_oRNA_243	2	1711	796
At3g63050-At3g63060	At_oRNA_244	3	429	973
At2g07718-At2g07719	At_oRNA_245	2	2188	3054
At1g64870-At1g64880	At_oRNA_246	1	871	446
At1g58270-At1g58280	At_oRNA_247	1	2087	3543
At1g58330-At1g58340	At_oRNA_248	1	7929	1135
At1g58350-At1g58360	At_oRNA_249	1	4684	3777
At1g59520-At1g59530	At_oRNA_250	1	401	75
At4g38810-At4g38820	At_oRNA_251	4	1635	583
At2g07746-At2g07677	At_oRNA_252	2	600	852
At3g15320-At3g15330	At_oRNA_253	3	3380	78
At3g47520-At3g47530	At_oRNA_254	3	44	2176
At3g24927-At3g24954	At_oRNA_255	3	2006	30
At3g24982-At3g25010	At_oRNA_256	3	1002	32
At1g10790-At1g10800	At_oRNA_257	1	72	90
At5g19480-At5g19490	At_oRNA_258	5	693	2624
At1g33930-At1g33940	At_oRNA_259	1	3589	140
At5g18690-At5g18700	At_oRNA_260	5	918	193
At5g46450-At5g46460	At_oRNA_261	5	204	226
At1g06500-At1g06510	At_oRNA_262	1	35	571

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At4g37020-At4g37030	At_oRNA_263	4	107	4239
At2g30420-At2g30430	At_oRNA_264	2	2489	3509
At1g25420-At1g25430	At_oRNA_265	1	4753	4401
At1g06500-At1g06510	At_oRNA_266	1	1034	222
At5g52880-At5g52890	At_oRNA_267	5	728	9978
At1g74080-At1g74090	At_oRNA_268	1	3605	262
At1g31940-At1g31950	At_oRNA_269	1	749	3011
At1g07900-At1g07910	At_oRNA_270	1	150	1590
At4g13570-At4g13580	At_oRNA_271	4	1769	6685
At5g47520-At5g47530	At_oRNA_272	5	1204	1752
At3g13430-At3g13435	At_oRNA_273	3	222	6442
At5g19890-At5g19900	At_oRNA_274	5	1591	234
At3g22150-At3g22160	At_oRNA_275	3	3	1979
At4g03420-At4g03430	At_oRNA_276	4	2475	278
At1g75680-At1g75690	At_oRNA_277	1	1600	633
At1g66730-At1g66740	At_oRNA_278	1	107	293
At1g77660-At1g77670	At_oRNA_279	1	15	1102
At1g50610-At1g50620	At_oRNA_280	1	3564	109
At2g02780-At2g02790	At_oRNA_281	2	3392	214
At2g22122-At2g22125	At_oRNA_282	2	5740	2929
At5g54730-At5g54740	At_oRNA_283	5	26	1040
At4g13490-At4g13500	At_oRNA_284	4	3459	12479
At4g33210-At4g33220	At_oRNA_285	4	51	1610
At3g60030-At3g60040	At_oRNA_286	3	4	6119
At4g38950-At4g38960	At_oRNA_287	4	1088	1473
At4g08910-At4g08920	At_oRNA_288	4	10637	1
At5g26910-At5g26920	At_oRNA_289	5	649	6328
At2g07717-At2g07718	At_oRNA_290	2	1487	187
At5g01200-At5g01210	At_oRNA_291	5	4677	909
At5g09860-At5g09870	At_oRNA_292	5	2012	21
At5g55780-At5g55790	At_oRNA_293	5	1223	297
At4g13070-At4g13080	At_oRNA_294	4	1345	1664
At2g22480-At2g22490	At_oRNA_295	2	968	3745
At1g13790-At1g13800	At_oRNA_296	1	153	156
At4g02070-At4g02075	At_oRNA_297	4	67	153
At3g59450-At3g59460	At_oRNA_298	3	2323	123
At5g39535-At5g39533	At_oRNA_299	5	3003	1768
At1g80530-At1g80540	At_oRNA_300	1	1036	9
At3g02510-At3g02520	At_oRNA_301	3	1195	227
At5g42090-At5g42100	At_oRNA_302	5	280	64
At5g62840-At5g62850	At_oRNA_303	5	255	359
At1g79100-At1g79110	At_oRNA_304	1	1339	3064
At2g32920-At2g32930	At_oRNA_305	2	329	77
At2g07716-At2g07717	At_oRNA_306	2	1856	1129
At4g12700-At4g12710	At_oRNA_307	4	295	32
At5g27380-At5g27390	At_oRNA_308	5	652	127
At5g37070-At5g37080	At_oRNA_309	5	1004	5078
At3g45720-At3g45730	At_oRNA_310	3	1968	706
At5g44580-At5g44590	At_oRNA_311	5	4805	2117
At3g15210-At3g15220	At_oRNA_312	3	1089	2708
At5g55120-At5g55130	At_oRNA_313	5	175	153

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At3g15510-At3g15520	At_oRNA_315	3	3169	56
At5g44565-At5g44570	At_oRNA_316	5	8069	2305
At2g47120-At2g47130	At_oRNA_317	2	767	34
At3g13430-At3g13435	At_oRNA_318	3	48	6548
At1g55325-At1g55330	At_oRNA_319	1	511	328
At5g49420-At5g49430	At_oRNA_320	5	928	113
At4g08405-At4g08410	At_oRNA_321	4	7499	23
At2g32390-At2g32400	At_oRNA_322	2	529	32
At1g06700-At1g06710	At_oRNA_323	1	1324	305
At5g05657-At5g05660	At_oRNA_324	5	847	53
At2g31890-At2g31900	At_oRNA_325	2	632	51
At2g01430-At2g01440	At_oRNA_326	2	2846	280
At2g44450-At2g44460	At_oRNA_327	2	32	2596
At1g19290-At1g19300	At_oRNA_328	1	409	1591
At3g27325-At3g27327	At_oRNA_329	3	146	4351
At2g20560-At2g20570	At_oRNA_330	2	342	4463
At4g21880-At4g21890	At_oRNA_331	4	84	122
At4g28930-At4g28940	At_oRNA_332	4	2832	518
At1g60987-At1g60990	At_oRNA_333	1	1220	228
At5g19420-At5g19430	At_oRNA_334	5	77	567
At2g22122-At2g22125	At_oRNA_335	2	2455	5352
At3g14170-At3g14180	At_oRNA_336	3	6558	1590
At4g39345-At4g39350	At_oRNA_337	4	2771	145
At4g24250-At4g24260	At_oRNA_338	4	61	3
At2g20470-At2g20480	At_oRNA_339	2	593	75
At5g26630-At5g26617	At_oRNA_340	5	4509	3958
At2g39230-At2g39240	At_oRNA_341	2	158	144
At5g35738-At5g35740	At_oRNA_342	5	131	244
At5g41570-At5g41580	At_oRNA_343	5	253	12
At5g28260-At5g28263	At_oRNA_345	5	3430	1356
At1g80370-At1g80380	At_oRNA_346	1	4	173
At4g36640-At4g36650	At_oRNA_347	4	3935	100
At3g12980-At3g12990	At_oRNA_348	3	33	675
At2g16050-At2g16060	At_oRNA_349	2	699	519
At2g07672-At2g07673	At_oRNA_350	2	4181	3801
At2g41980-At2g41990	At_oRNA_351	2	2183	170
At1g17220-At1g17230	At_oRNA_352	1	240	186
At1g55170-At1g55180	At_oRNA_354	1	2189	492
At4g23710-At4g23720	At_oRNA_355	4	2957	3745
At4g37690-At4g37700	At_oRNA_356	4	79	506
At1g74590-At1g74600	At_oRNA_357	1	244	104
At1g35660-At1g35663	At_oRNA_358	1	256	1255
At4g29230-At4g29240	At_oRNA_359	4	114	3118
At5g40160-At5g40170	At_oRNA_360	5	668	25
At2g07758-At2g07731	At_oRNA_361	2	3255	1250
At3g54750-At3g54760	At_oRNA_362	3	316	296
At4g36550-At4g36560	At_oRNA_363	4	1387	3232
At4g04610-At4g04620	At_oRNA_364	4	909	243
At4g38300-At4g38310	At_oRNA_365	4	174	2424
At3g47090-At3g47100	At_oRNA_366	3	20	522
At3g61630-At3g61640	At_oRNA_367	3	3021	208

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At5g25350-At5g25360	At_oRNA_368	5	434	1833
At3g55550-At3g55560	At_oRNA_369	3	2144	136
At4g00760-At4g00770	At_oRNA_370	4	150	708
At1g03720-At1g03730	At_oRNA_371	1	1255	2782
At2g07794-At2g07707	At_oRNA_372	2	2138	369
At1g32920-At1g32930	At_oRNA_373	1	171	2195
At2g15820-At2g15830	At_oRNA_374	2	645	2800
At3g26790-At3g26800	At_oRNA_375	3	3679	2153
At4g07507-At4g07510	At_oRNA_376	4	285	862
At1g07120-At1g07130	At_oRNA_377	1	767	48
At1g74580-At1g74590	At_oRNA_378	1	65	79
At5g07980-At5g07990	At_oRNA_379	5	5112	424
At2g01930-At2g01940	At_oRNA_380	2	3499	73
At3g49600-At3g49610	At_oRNA_381	3	1327	1802
At1g75020-At1g75030	At_oRNA_382	1	413	87
At2g25090-At2g25100	At_oRNA_383	2	3363	3195
At2g28120-At2g28130	At_oRNA_384	2	595	25
At5g66320-At5g66330	At_oRNA_385	5	1510	697
At5g43300-At5g43310	At_oRNA_386	5	145	366
At5g46580-At5g46590	At_oRNA_387	5	5655	125
At1g43387-At1g43415	At_oRNA_388	1	973	3217
At4g25220-At4g25230	At_oRNA_389	4	205	574
At1g76990-At1g77000	At_oRNA_390	1	1141	3890
At1g76700-At1g76705	At_oRNA_391	1	1411	1538
At4g10420-At4g10430	At_oRNA_392	4	197	26
At1g71000-At1g71010	At_oRNA_393	1	3416	9059
At4g16350-At4g16360	At_oRNA_394	4	425	859
At2g07716-At2g07717	At_oRNA_395	2	243	2866
At2g01020-At2g01022	At_oRNA_396	2	4111	105
At3g41345-At3g41768	At_oRNA_397	3	21541	3732
At3g41979-At3g42050	At_oRNA_398	3	4195	24142
At3g26790-At3g26800	At_oRNA_399	3	55	5880
At5g41080-At5g41090	At_oRNA_400	5	676	823
At1g05480-At1g05490	At_oRNA_402	1	14	177
At1g36060-At1g36070	At_oRNA_403	1	324	10634
At3g21210-At3g21215	At_oRNA_404	3	1009	103
At1g27460-At1g27470	At_oRNA_405	1	231	1893
At5g48520-At5g48530	At_oRNA_406	5	464	1245
At3g56590-At3g56600	At_oRNA_407	3	65	930
At5g19900-At5g19910	At_oRNA_408	5	82	851
At3g44325-At3g44330	At_oRNA_409	3	957	5359
At1g03550-At1g03560	At_oRNA_410	1	365	1350
At5g09330-At5g09340	At_oRNA_411	5	10	1253
At2g07721-At2g07722	At_oRNA_412	2	2736	371
At3g33076-At3g33069	At_oRNA_413	3	231	65362
At5g18540-At5g18550	At_oRNA_414	5	5742	67
At1g76980-At1g76990	At_oRNA_415	1	1649	2952
At1g31885-At1g31900	At_oRNA_416	1	1662	3842
At5g06120-At5g06125	At_oRNA_417	5	3	314
At3g04300-At3g04310	At_oRNA_418	3	201	96
At5g13580-At5g13590	At_oRNA_419	5	38	1041

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At5g13580-At5g13590	At_oRNA_420	5	576	520
At1g47610-At1g47620	At_oRNA_421	1	397	1363
At2g07698-At2g07699	At_oRNA_424	2	4	3611
At2g07590-At2g07600	At_oRNA_425	2	1123	1353
At1g62330-At1g62340	At_oRNA_427	1	280	18
At4g13965-At4g13970	At_oRNA_428	4	2512	52
At3g53000-At3g53010	At_oRNA_429	3	408	409
At2g07792-At2g07689	At_oRNA_430	2	524	1982
At4g19990-At4g20000	At_oRNA_431	4	36	2607
At3g51760-At3g51770	At_oRNA_432	3	376	106
At4g32700-At4g32710	At_oRNA_433	4	257	550
At1g64570-At1g64580	At_oRNA_434	1	132	276
At2g25040-At2g25050	At_oRNA_435	2	1669	38
At3g32397-At3g32400	At_oRNA_436	3	743	38
At2g02810-At2g02820	At_oRNA_437	2	738	1397
At5g03400-At5g03415	At_oRNA_438	5	1869	637
At1g53300-At1g53310	At_oRNA_440	1	48	1143
At5g53110-At5g53120	At_oRNA_441	5	2828	1403
At5g44565-At5g44570	At_oRNA_442	5	8089	2635
At5g44565-At5g44570	At_oRNA_443	5	8129	2923
At4g13390-At4g13400	At_oRNA_444	4	359	1732
At1g65510-At1g65520	At_oRNA_445	1	10	980
At1g13635-At1g13640	At_oRNA_446	1	77	553
At2g32240-At2g32250	At_oRNA_447	2	300	2048
At3g60990-At3g61010	At_oRNA_448	3	932	110
At1g52315-At1g52320	At_oRNA_449	1	1681	670
At5g27920-At5g27925	At_oRNA_450	5	34	4352
At3g25882-At3g25890	At_oRNA_451	3	4719	93
At3g14170-At3g14180	At_oRNA_452	3	574	7713
At1g03160-At1g03170	At_oRNA_453	1	899	3478
At1g63690-At1g63700	At_oRNA_454	1	2274	38
At4g04920-At4g04925	At_oRNA_455	4	115	635
At5g24350-At5g24355	At_oRNA_456	5	157	2919
At4g13570-At4g13580	At_oRNA_457	4	6727	6684
At4g13490-At4g13500	At_oRNA_458	4	2713	13026
At4g03630-At4g03640	At_oRNA_459	4	1707	2195
At2g22122-At2g22125	At_oRNA_460	2	7777	738
At3g03800-At3g03810	At_oRNA_461	3	216	37
At2g22122-At2g22125	At_oRNA_462	2	8372	57
At5g08750-At5g08770	At_oRNA_463	5	461	57
At4g07507-At4g07510	At_oRNA_464	4	814	255
At4g35230-At4g35240	At_oRNA_465	4	2978	22
At4g06697-At4g06699	At_oRNA_467	4	1056	1616
At2g10930-At2g10940	At_oRNA_468	2	982	4897
At4g39100-At4g39110	At_oRNA_469	4	285	1369
At4g04925-At4g04930	At_oRNA_470	4	46	1714
At4g28395-At4g28400	At_oRNA_471	4	1074	1693
At1g10680-At1g10690	At_oRNA_472	1	1208	4534
At1g01450-At1g01460	At_oRNA_473	1	218	2432
At3g62890-At3g62900	At_oRNA_474	3	92	83
At3g06770-At3g06780	At_oRNA_475	3	3240	2001

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At3g06770-At3g06780	At_oRNA_476	3	3687	1964
At1g11180-At1g11190	At_oRNA_477	1	1027	101
At5g28625-At5g28627	At_oRNA_478	5	1820	823
At1g36070-At1g36080	At_oRNA_479	1	7630	4748
At1g62305-At1g62310	At_oRNA_480	1	5886	139
At1g66780-At1g66790	At_oRNA_481	1	641	5079
At4g22140-At4g22150	At_oRNA_482	4	755	735
At5g65550-At5g65560	At_oRNA_483	5	201	338
At1g55150-At1g55160	At_oRNA_484	1	124	135
At3g29180-At3g29185	At_oRNA_485	3	2594	438
At4g19390-At4g19400	At_oRNA_486	4	2183	1646
At4g04925-At4g04930	At_oRNA_487	4	650	1217
At1g06630-At1g06640	At_oRNA_488	1	415	959
At4g20320-At4g20330	At_oRNA_489	4	174	3056
At4g13490-At4g13500	At_oRNA_490	4	2933	13146
At2g46700-At2g46710	At_oRNA_491	2	4672	156
At3g07190-At3g07195	At_oRNA_492	3	15	1414
At2g06990-At2g07000	At_oRNA_493	2	233	134
At1g74150-At1g74160	At_oRNA_494	1	2930	78
At1g43600-At1g43610	At_oRNA_495	1	945	6769
At3g06120-At3g06130	At_oRNA_496	3	731	2296
At1g35780-At1g35790	At_oRNA_497	1	912	13188
At3g54360-At3g54380	At_oRNA_498	3	1032	143
At3g27410-At3g27420	At_oRNA_499	3	1501	857
At5g19470-At5g19480	At_oRNA_501	5	389	955
At2g07703-At2g07704	At_oRNA_502	2	2468	244
At3g42140-At3g42150	At_oRNA_503	3	4600	1020
At4g04070-At4g04080	At_oRNA_504	4	6533	1493
At3g59510-At3g59520	At_oRNA_506	3	567	1038
At1g33100-At1g33110	At_oRNA_507	1	1035	2091
At1g68730-At1g68740	At_oRNA_509	1	712	130
At1g29410-At1g29420	At_oRNA_510	1	794	180
At2g05510-At2g05520	At_oRNA_511	2	6546	2166
At2g39170-At2g39180	At_oRNA_512	2	3690	3886
At5g12400-At5g12410	At_oRNA_514	5	24	279
At2g01810-At2g01820	At_oRNA_515	2	3505	1914
At2g46760-At2g46770	At_oRNA_516	2	4781	118
At3g46840-At3g46850	At_oRNA_517	3	330	1175
At5g03550-At5g03555	At_oRNA_518	5	624	390
At2g42360-At2g42370	At_oRNA_520	2	123	358
At2g41310-At2g41330	At_oRNA_521	2	1093	1075
At1g29460-At1g29470	At_oRNA_522	1	431	452
At1g67360-At1g67370	At_oRNA_523	1	405	38
At2g42480-At2g42490	At_oRNA_524	2	697	400
At1g10650-At1g10660	At_oRNA_525	1	4169	673
At1g10650-At1g10660	At_oRNA_526	1	4132	738
At3g19500-At3g19510	At_oRNA_527	3	1190	64
At1g68820-At1g68830	At_oRNA_528	1	710	2280
At1g67140-At1g67150	At_oRNA_530	1	3863	690
At1g07760-At1g07770	At_oRNA_532	1	702	156
At1g35250-At1g35260	At_oRNA_533	1	1383	440

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At3g44325-At3g44330	At_oRNA_534	3	1621	4625
At5g07320-At5g07330	At_oRNA_535	5	1046	2417
At3g57060-At3g57070	At_oRNA_536	3	295	2046
At3g44796-At3g44800	At_oRNA_537	3	2197	320
At5g09440-At5g09450	At_oRNA_538	5	1073	555
At1g53540-At1g53550	At_oRNA_539	1	314	1461
At3g12960-At3g12970	At_oRNA_540	3	134	3413
At3g04470-At3g04480	At_oRNA_541	3	307	993
At1g50050-At1g50060	At_oRNA_542	1	158	1062
At4g21190-At4g21200	At_oRNA_543	4	252	7320
At2g45020-At2g45030	At_oRNA_544	2	591	450
At2g05630-At2g05635	At_oRNA_545	2	2350	1481
At2g05630-At2g05635	At_oRNA_546	2	2267	1355
At1g57830-At1g57840	At_oRNA_547	1	1657	1455
At4g04840-At4g04850	At_oRNA_550	4	1294	229
At3g41345-At3g41768	At_oRNA_551	3	19548	4868
At3g41345-At3g41768	At_oRNA_552	3	22213	1952
At3g41979-At3g42050	At_oRNA_553	3	4355	23783
At3g41345-At3g41768	At_oRNA_554	3	21701	3373
At4g36640-At4g36650	At_oRNA_555	4	2952	552
At5g53480-At5g53487	At_oRNA_558	5	3949	276
At3g46640-At3g46650	At_oRNA_559	3	989	235
At4g17390-At4g17410	At_oRNA_560	4	162	869
At3g03340-At3g03350	At_oRNA_561	3	110	1843
At1g46552-At1g46624	At_oRNA_562	1	517	1857
At4g05040-At4g05050	At_oRNA_563	4	44	5725
At5g40980-At5g40990	At_oRNA_564	5	181	3090
At3g63530-END	At_oRNA_565	3	767	84
At2g07981-At2g08986	At_oRNA_566	2	6796	422
At3g41979-At3g42050	At_oRNA_567	3	204	27699
At2g01020-At2g01022	At_oRNA_568	2	204	3652
At1g08420-At1g08430	At_oRNA_569	1	127	1659
At1g11170-At1g11180	At_oRNA_570	1	243	120
At4g00985-At4g00990	At_oRNA_571	4	1162	291
At2g31580-At2g31590	At_oRNA_572	2	756	111
At1g22030-At1g22040	At_oRNA_573	1	5562	1394
At3g06120-At3g06130	At_oRNA_574	3	656	3596
At1g27460-At1g27470	At_oRNA_575	1	1189	1066
At1g67100-At1g67110	At_oRNA_576	1	2635	3469
At1g74870-At1g74880	At_oRNA_577	1	1001	105
At2g14870-At2g14880	At_oRNA_578	2	741	320
At3g07270-At3g07280	At_oRNA_579	3	216	867
At3g15010-At3g15020	At_oRNA_580	3	439	187
At3g26390-At3g26400	At_oRNA_581	3	1614	267
At3g61700-At3g61710	At_oRNA_582	3	146	254
At4g13490-At4g13500	At_oRNA_583	4	5961	10021
At5g44565-At5g44570	At_oRNA_584	5	10196	586
At5g47000-At5g47010	At_oRNA_585	5	1539	4
At3g30834-At3g30836	At_oRNA_586	3	1861	3400
At1g29410-At1g29420	At_oRNA_587	1	234	692
At2g31580-At2g31590	At_oRNA_588	2	195	79

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At4g08032-At4g08033	At_oRNA_589	4	2800	8421
At5g28625-At5g28627	At_oRNA_590	5	1796	870
At5g48410-At5g48420	At_oRNA_591	5	62	2152

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

S10 Conserved orphan transcripts in rice (*Oryza sativa*)

Highly conserved orphan transcript.

Orphan transcripts were blasted against the rice genome sequence. A hit was considered highly conserved if the e-value was $<10^{-10}$ and an alignment occurred over at least 40% of the length of the orphan transcript.

Orphan transcript
At_oRNA_092
At_oRNA_093
At_oRNA_141
At_oRNA_290
At_oRNA_372
At_oRNA_422
At_oRNA_423
At_oRNA_426
At_oRNA_430

Weakly conserved orphan transcripts.

A hit was considered weakly conserved when the e-value was $<10^{-4}$, and at least 5% of the orphan transcript aligned to any sequence of the rice genome. The length aligned was always $<40\%$

Orphan transcript
At_oRNA_021
At_oRNA_025
At_oRNA_051
At_oRNA_117
At_oRNA_119
At_oRNA_131
At_oRNA_134
At_oRNA_181
At_oRNA_182
At_oRNA_197
At_oRNA_198
At_oRNA_266
At_oRNA_289
At_oRNA_309
At_oRNA_313
At_oRNA_316
At_oRNA_401
At_oRNA_409
At_oRNA_425
At_oRNA_427
At_oRNA_439
At_oRNA_442
At_oRNA_443
At_oRNA_460
At_oRNA_510
At_oRNA_513
At_oRNA_540

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

At_oRNA_548
At_oRNA_549
At_oRNA_557
At_oRNA_567
At_oRNA_568

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

References

- Attwood, T.K., Bradley, P., Flower, D.R., Gaulton, A., Maudling, N., Mitchell, A.L., Moulton, G., Nordle, A., Paine, K., Taylor, P., Uddin, A. and Zygouri, C. (2003) PRINTS and its automatic supplement, prePRINTS. *Nucl. Acids Res.*, **31**, 400-402.
- Barneche, F., Gaspin, C., Guyot, R. and Echeverria, M. (2001) Identification of 66 Box C/D snoRNAs in *Arabidopsis thaliana*: Extensive Gene Duplication Generated Multiple isoforms Predicting New Ribosomal RNA 2'-O-Methylation Sites. *J. Mol. Biol.* **311**, 57-73.
- Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L., Studholme, D.J., Yeats, C. and Eddy, S.R. (2004) The Pfam Protein Families Database *Nucl. Acids Res.*, **32**, D138-D141.
- Brown, J.S.W., Echeverria, M. and Qu, L.H. (2003a) Plan snoRNAs: Functional Evolution And New Modes of Gene Expression. *Trends Plant Sci.*, **8**, 42-49.
- Brown, J.S.W., Echeverria, M., Qu, L.H., Lowe, T.M., Bachellerie, J.P., Huttenhofer, A., Kastenmayer, J.P., Green, P.J., Shaw, P. and Marshall, D.F. (2003b) Plant snoRNA database. *Nucl. Acids Res.*, **31**, 432-435.
- Erdmann, V.A., Szymanski, M., Hochberg, A., Groot, N. and Barciszewski, J. (2000) Non-coding, mRNA like RNAs Database Y2K. *Nucl. Acids Res.*, **28**, 197-200.
- Griffiths-Jones, S. (2004) The microRNA Registry. *Nucl. Acids Res.*, **32**, D109-D111.
- Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A. and Eddy, S.R. (2003) Rfam: An RNA Family database. *Nucl. Acids Res.*, **31**, 439-441.
- Haft, D.H., Selengut, J.D. and White, O. (2003) The TIGRFAMs database of protein families. *Nucl. Acids Res.*, **31**, 371-373.
- Hulo, N., Sigrist, C.J.A., Le Saux, V., Langendijk-Genevaux, P.S., Bordoli, L., Gattiker, A., De Castro, E., Bucher, P. and Bairoch, A. (2004) Recent improvements to the PROSITE database *Nucl. Acids Res.*, **32**, D134-D137.
- Leader, D.J., Clark, G.P., Watters, J., Beven, A.F., Shaw, P.J. and Brown, J.W.S. (1997) Clusters of multiple different nucleolar RNA genes in plants are expressed as and processed from polycistronic pre-snoRNAs. *EMBO J.*, **16**, 5742-5751.
- Letunic, I., Goodstadt, L., Dickens, N.J., Doerks, T., Schultz, J., Mott, R., Ciccarelli, F., Copley, R.R., Ponting, C.P. and Bork, P. (2002) Recent improvements to the SMART domain-based sequence annotation resource. *Nucl. Acids Res.*, **30**, 242-244.
- Liang-Hu, Q., Qing, M., Hui, Z. and Yue-Qin, C. (2001) Identification of 10 novel snoRNA gene clusters from *Arabidopsis thaliana*. *Nucl. Acids Res.*, **29**, 1623-1630.

Orphan transcripts in *Arabidopsis thaliana*: identification of several hundred previously unrecognized genes.

Liu, G., Loraine, A.E., Shigeta, R., Cline, M., Cheng, J., Valmeekam, V., Sun, S., Kulp, D. and Siani-Rose, M.A. (2003) NetAffx: Affymetrix probesets and annotations. *Nucl. Acids Res.*, **31**, 82-86.

Lowe, T.M. and Eddy, S.R. (1997) tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucl. Acids Res.*, **25**, 955-964.

Marker, C., Zemann, A., Terhorst, T., Kiefmann, M., Kastenmayer, J.P., Green, P., Bachellerie, J.P., Brosius, J. and Hüttenhofer, A. (2002) Experimental RNomics: identification of 140 candidates for small non-messenger RNAs in the plant *Arabidopsis thaliana*. *Curr. Biol.*, **12**, 2002-2013.

Moss, E.G. and Poethig, R.S. (2002) MicroRNAs: Something New Under the Sun. *Curr. Biol.* **12**, R688-R690.

Servant, F., Bru, C., Carrère, S., Courcelle, E., Gouzy, J., Peyruc, D. and Kahn, D. (2002) ProDom: Automated clustering of homologous domains. *Brief Bioinform.*, **3**, 246-251.

Szymanski, M., Erdmann, V.A. and Barciszewski, J. (2003) Noncoding regulatory RNAs database. *Nucl. Acids Res.* **31**, 429-431.

Teramoto, H., Toyama, T., Takeba, G. and Tsuji, H. (1996). Noncoding RNA for CR20, a cytokinin-repressed gene of cucumber. *Plant Mol. Biol.*, **32**, 797-808.

Zdobnov, E.M. and Apweiler, R. (2001) InterProScan - an integration platform for the signature-recognition methods in InterPro. *Bioinformatics*, **17**, 847-848.