Please cite this article in press as: Addo-Quaye et al., Endogenous siRNA and miRNA Targets Identified by Sequencing of the *Arabidopsis* Degradome, Current Biology (2008), doi:10.1016/j.cub.2008.04.042

# Supplemental Data



# Endogenous siRNA and miRNA Targets Identified by Sequencing of the *Arabidopsis* Degradome

**Charles Addo-Quaye, Tifani W. Eshoo, David P. Bartel, and Michael J. Axtell**

## Supplemental Experimental Procedures

### Library Construction

Libraries 1–3 were generated from 4–6-week-old inflorescence tissue, and library 4 was generated from 10-day-old Col-0 *A. thaliana* seedlings. Total RNA was isolated with TRI Reagent (Sigma) per the recommended protocol followed by polyA$^+$ RNA purification with the Oligotex mRNA Midi Kit (QIAGEN) via the batch protocol. Five micrograms of polyA$^+$ RNA was added to a 5′ DNA-RNA hybrid adaptor containing a 3′ MmeI site at a ratio of 136 pmol/μg of polyA$^+$ RNA to set up the 5′ adaptor ligation. The polyA$^+$ RNA, water, and adaptor were combined and heated at 65°C for 5 min and then transferred to ice for 2 min. T4 RNA ligase buffer and T4 RNA ligase (New England Biolabs) were added for a 0.5 or 1 hr incubation at 37°C. Library 2 underwent a second polyA$^+$ purification after the 5′ adaptor ligation, and libraries 3–4 were passed through MicroSpin S-300 columns (GE Healthcare) to remove unligated adaptor. Library 2 was reverse transcribed with the GeneRacer oligo(dT) primer (Invitrogen GeneRacer) and pool amplified. First-strand cDNA was generated in libraries 3–4 with random hexamers added at a ratio of 83.3 ng/μg mRNA, following the RT-PCR protocol from [S1]. After reverse transcription, libraries 3–4 were eluted in 50 mM NaOH, 5 mM EDTA (pH 8) to degrade remaining RNA. After 1 hr at 65°C, an equal volume of 1M Tris-HCl (pH 7.5) was added to neutralize the reaction. Then, libraries 3–4 were passed through G-25 microspin columns (GE Heathcare) and primer extended. All libraries underwent the MmeI digest with the same conditions with 2U of MmeI being added per μg of library followed by a 1 hr incubation at 37°C. Two oligos were annealed to make a 3′ dsDNA adaptor. One hundred picomoles of this adaptor was added per μg of library along with 400U of T4 DNA Ligase (New England Biolabs) in a 5 hr room-temperature incubation. The library was purified with an 8% (19:1 acryl:bis-acryl) gel and then amplified. The amplified library then underwent a final gel purification prior to sequencing.

### Datasets and Categorizations

Mature *Arabidopsis* miRNAs were all those in miRBase 10.0 [S2] except for those in the miR413-420 and miR426 families. Expanded query sets consisting of expressed 20-22nt small RNAs matching the sense polarity of one or more of the eligible miRNA hairpins described above or 20-22nt expressed small RNAs matching a set of known secondary siRNA-producing loci ([S3, S4]; *TAS1a-c*, *TAS2*, *TAS3a-c*, *TAS4*, *At1g62910*, *At1g63130*, *At1g62930*, *At1g63080*, *At1g63400*, *At1g63150*, *At1g63070*, *At1g63330*, *At1g62590*, *At5g38850*, *At1g12820*, *At5g41610*, *At5g60450*, *At1g06580*, *At1g62860*, *At1g63230*, *At1g64580*, *At5g16640*, *At5g41170*, *At1g62670*, *At1g48410*, *At3g62980*, *At4g03190*, *At3g26810*, and *At3g23690*) were compiled from previously described wild-type small RNA libraries ([S5, S6]; NCBI GEO GSM118372-118375, GSM154361, GSM154370, and GSM154375). Analysis was identical to that of the annotated mature miRNA queries with the exception that fewer control cohorts of randomized queries were used to derived threshold values (Figure S1). Category I targets were those where the sum of the degradome tag abundances from positions nine to eleven was both > 1 and higher than any other tag abundances on the transcript. Category II targets were those where the transcript had four or more matching tags, the sum of the degradome tag abundances from positions nine to eleven was > 1, and this sum was in the top 1/3 compared to any other tag abundances on the transcript. Category III targets were those that were neither Category I nor Category II.

## Supplemental References

S1. Cheng, J., Kapranov, P., Drenkow, J., Dike, S., Brubaker, S., Patel, S., Long, J., Stern, D., Tammana, H., Helt, G., et al. (2005). Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. Science *308*, 1149–1154.

S2. Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A., and Enright, A.J. (2006). miRBase: MicroRNA sequences, targets and gene nomenclature. Nucleic Acids Res. *34*, D140–D144.

S3. Axtell, M.J., Jan, C., Rajagopalan, R., and Bartel, D.P. (2006). A two-hit trigger for siRNA biogenesis in plants. Cell *127*, 565–577.

S4. Howell, M.D., Fahlgren, N., Chapman, E.J., Cumbie, J.S., Sullivan, C.M., Givan, S.A., Kasschau, K.D., and Carrington, J.C. (2007). Genome-wide analysis of the *RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4* pathway in Arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting. Plant Cell *19*, 926–942.

S5. Kasschau, K.D., Fahlgren, N., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., and Carrington, J.C. (2007). Genome-wide profiling and analysis of *Arabidopsis* siRNAs. PLoS Biol. *5*, e57.

S6. Rajagopalan, R., Vaucheret, H., Trejo, J., and Bartel, D.P. (2006). A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. Genes Dev. *20*, 3407–3425.

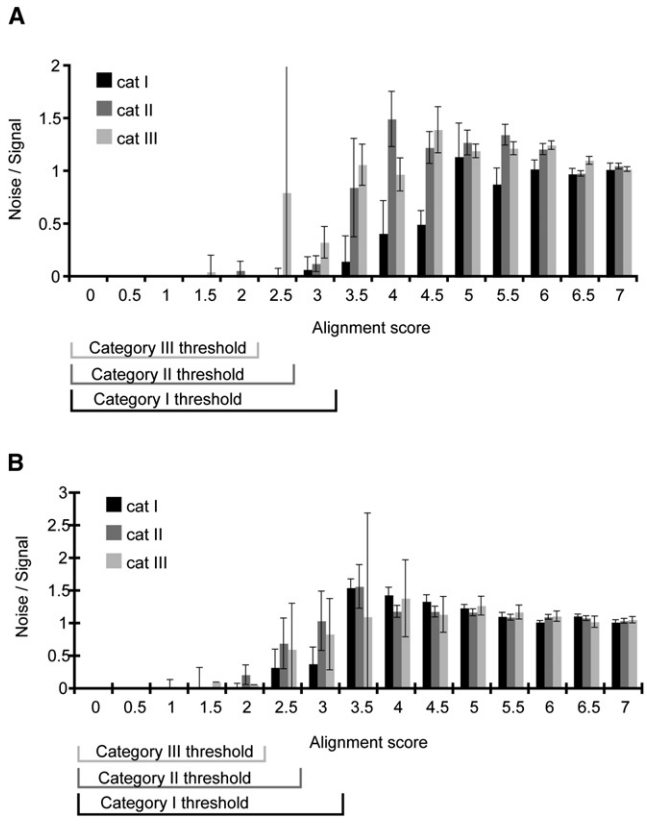


Figure S1. Signal-to-Noise Analyses of Expanded Query Sets

(A) Histogram displaying mean ratio of targets found with ten cohorts of randomly permuted queries to the number found with a data set comprising all expressed 20–22-mers from known *Arabidopsis* miRNA hairpins at different alignment scores. Error bars represent one standard deviation. Alignment score thresholds are indicated for category I, II, and III targets.

(B) As in (A) for a data set of expressed 20–22-mers from known secondary siRNA producing loci by using nine cohorts of randomly permuted queries.

Please cite this article in press as: Addo-Quaye et al., Endogenous siRNA and miRNA Targets Identified by Sequencing of the *Arabidopsis* Degradome, Current Biology (2008), doi:10.1016/j.cub.2008.04.042



Table S1. *Arabidopsis* Degradome Libraries

| Library | Tissue | RT[a] | Second Strand[b] | Sequencing Method | Tags[c] | Structural RNAs[c,d] | Transcript-Matched[c,e] | Fraction of Transcriptome Covered (%)[f] |
|---|---|---|---|---|---|---|---|---|
| 1 | Inflorescence | dT | Pool-amplified | 454 | 93,670 (15,613) | 245 (90) | 12,403 (11,682) | 16 |
| 2 | Inflorescence | dT | Pool-amplified | Solexa | 1,371,358 (763,960) | 21,449 (1,594) | 889,157 (430,622) | 64 |
| 3 | Inflorescence | Random | Primer extension | Solexa | 7,179,831 (851,938) | 1,868,877 (13,420) | 2,234,831 (345,288) | 65 |
| 4 | Seedlings | Random | Primer extension | Solexa | 5,648,121 (660,829) | 1,119,944 (10,389) | 1,956,177 (223,138) | 59 |
| Union[g] | | | | | 14,292,980 (2,025,860) | 3,010,515 (15,651) | 5,092,568 (853,902) | 73 |

[a] Priming method for reverse transcription.
[b] Second-strand-synthesis method.
[c] Number of reads, with number of unique sequences in parentheses.
[d] Tags matching mature rRNAs, tRNAs, snRNAs, or snoRNAs—these were removed from further analysis.
[e] Tags matching the sense polarity of one or more annotated mRNAs.
[f] Percent of annotated *Arabidopsis* mRNAs with at least one unambiguous degradome tag.
[g] Nonredundant summary of all four libraries. All tags were deposited at NCBI GEO (GSE11007).

Table S2. Oligos for RLM 5′-RACE

| Gene | Location | Oligo Sequence (5′–3′) |
|---|---|---|
| AT5G39610 | Outer | GTACCGGACGAATCACGACCGTCGAAA |
| AT5G39610 | Inner | CCGTCGAAAGGTTTGCCGGAGAATTGA |
| AT1G10120 | Outer | AGGATTCAGCCCGAGTGTAGGAGTGTT |
| AT1G10120 | Inner | GGTGATCTTGTTGCATCCGGGAACAAG |