PN	AS
WW	w.pnas.org

	www.pnas.org
1 2 3	Supplementary Information for
4	An aphid RNA transcript migrates systemically within plants and is a
5	virulence factor
6	
7	Yazhou Chen ^{1\$} , Archana Singh ^{2\$} , Gemy G. Kaithakottil ² , Thomas C. Mathers ¹ ,
8	Matteo Gravino ¹ , Sam T. Mugford ¹ , Cock van Oosterhout ³ , David Swarbreck ² and
9	Saskia A. Hogenhout ^{1*}
10	
11	¹ Crop Genetics, John Innes Centre, Norwich Research Park, Norwich, NR4 7UH,
12	UK; ² Earlham Institute, and ³ School of Environmental Sciences, University of East
13	Anglia, Norwich Research Park, Norwich, NR4 7UH, UK.
14	^{\$} Y.C and A.S. contributed equally to the work.
15	
16	*To whom correspondence may be addressed:
17	Saskia A. Hogenhout, Department of Crop Genetics, John Innes Centre, Norwich
18	Research Park, Norwich NR4 7UH, United Kingdom
19	Email: <u>saskia.hogenhout@jic.ac.uk</u>
20	ORCID: 0000-0003-1371-5606
21 22 23 24 25 26 27	This PDF file includes: Supplementary file S1 Figures S1 to S19 Tables S1 to S3 SI References
28	Other supplementary materials for this manuscript include the following:
29	Datasets S1-S5

- **1** Supplementary Methods
- 2

3 Plants and growth conditions

4 Seeds of B. rapa, A. thaliana, N. benthamiana were provided by John Innes 5 Centre horticultural department. Solanum tuberosum (Desiree) were purchased from Marks & Spencer (Norwich, UK). Seeds of C. indicum (polar star), H. annuus 6 7 (Helios X Helios Flame), P. vulgaris (NYFB) were purchased from Thompson 8 Morgan (Ipswich, UK). Seeds of P. sativum (JI3253) were provided by Claire 9 Domoney (John Innes Centre, UK). Seeds of Z. mays (Early sunglow corn) were 10 provided by Ian Bedford (John Innes Centre, UK). Seeds were directly sown into 11 soil (peat-based compost). Plants were grown in a controlled environmental room 12 at a constant temperature of 22°C and 70% humidity under a 10 h day /14 h night 13 cycle.

14

15 *M. persicae* transfer to 9 plant species

16 A colony of *M. persicae* clone O that started from a single female was 17 established on *B. rapa* in a growth chamber (14 h light, 10 h dark at constant 20 18 °C, 75% humidity) in 2010. From this founder colony on *B. rapa*, approximately 19 500 asexual individuals were transferred to each of 9 plant species, including B. 20 rapa (used as reference) and A. thaliana, as a representative of another plant 21 species of the family Brassicaceae, and plant species belonging to 4 additional 22 plant families, including N. benthamiana and S. tuberosum (Solanaceae), C. 23 indicum and H. annuus (Asteraceae), P. sativum and P. vulgaris (C. indicum and 24 H. annuus), and the monocot maize (Z. mays) (Poaceae). The aphids were 25 maintained for about 4 generations on these plants, except for *C. indicum* and *P.* 26 vulgaris, on which the aphids were reared for 10 generations. Aphid colonies on all 9 plant species were maintained in the same growth chamber (14 h light, 10 h 27 dark at constant 20 °C, 75% humidity). 28

29 Stable aphid populations on the 9 hosts were analyzed for development time 30 in days between births and emergence of adults, longevities in days between births 31 to deaths, weights in mg per adult and reproduction rates measured by numbers

of nymphs divided by the number of adults on days 7, 9 and 11 for colonies on *B. rapa*, *A. thaliana*, *N. benthamiana* and *S. tuberosum* and on days 7, 9, 11, 13, 15,
 17 and 18 for colonies on *C. indicum*, *H. annuus*, *P. sativum*, *P. vulgaris* and *Z. mays*.

5

6 Transcriptome sequencing of *M. persicae* on 9 hosts

Approximately 100 aphids were harvested from each plant species, at 5 independent biological replicates per plant species, and snap frozen in liquid nitrogen. Aphid samples were ground to powder using 5 mm stainless steel beads (Qiagen, Germany) and a TissueLyser II (Qiagen, Germany), and RNA was extracted using Trizol reagent (Sigma, UK) and the Qiagen RNeasy MinElute Cleanup Kit (Qiagen, Germany), which included and on-column DNase digestion.

13

14 **RNA** sequencing and transcriptome assembly

15 Strand-specific libraries were constructed from mRNAs isolated from 1 ug 16 of total aphid RNA using the poly-A method of the Illumina TruSeq RNA Library 17 Preparation kit (Illumina, US) following the manufacturer's procedures. cDNA was 18 synthesized by 10 cycles of PCR to amplify the fragments. Libraries were then 19 pooled and sequenced on a HiSeq 2000 generating 150-bp paired-end sequences 20 (Earlham Institute, Norwich, UK). A genome-guided transcriptome assembly was 21 generated with RNA-seq data of the 45 libraries of the nine host experiments, five 22 replicates each (data generated herein as described above) and RNA-seg data 23 generated from library LIB1777 (1), as follows. Reads were trimmed for low guality 24 and adapter using Trim Galore! v0.4.0 with default settings (2). Trimmed reads 25 were aligned to *M. persicae* reference genome G006 v1 (1) by HISat2 version-26 2.0.5 (3). The RNA-seq reads of all 46 libraries were merged together in one BAM 27 file (Binary version of Sequence Alignment/Map) using Samtools (v0.1.18) (4) and 28 assembled to create transcript models by StringTie version-1.3.3 (5) guided by the 29 reference genome (6). A consensus assembly was produced using StringTie 30 merge. Transcripts with Fragments Per Kilobase of transcript per Million (FPKM) \leq 31 0.2 were removed from downstream analyses. Details of all transcriptomic libraries

generated for this study are listed in (*SI Appendix*, Table S1). GFF files of the
 transcriptome assembly were submitted to GEO (GSE129667,
 Transcriptome_Assembly_G006_V2.gff).

- 4
- 5

Functional annotation and *M. persicae* IncRNAs identification

The computational workflow for the IncRNAs identification of the M. 6 7 persicae is shown in SI Appendix, Fig. S2A. IncRNA identification was performed 8 on transcript models obtained from the transcriptome assembly described above. 9 Mikado compare (7) was used to identify and subsequently exclude transcripts 10 overlap over 10% of the length of annotated miRNA, tRNA, rRNA and transposons 11 features. To identify putative Open Reading Frame (ORFs), we used 12 TransDecoder (https://transdecoder.github.io) with the default parameters. To 13 further maximize the sensitivity, we scanned all ORFs for homology to curated 14 protein sequence (Arthropods) in the Swiss-Prot database (8), downloaded 15 November 20, 2018 from http://www.uniprot.org/downloads. This was done by 16 BlastP (v2.7.1, evalue 1e-5). HAMMER (v-3.0) was run against the Pfam database 17 (9) with default parameter to search for protein domains. The coding potentials of 18 the remaining transcripts were assessed using CPC2 (10) and those with a coding 19 potential score < 0.5 were selected. To consider transcript as a candidate 20 IncRNA, transcript must be larger than 200nt, not have a hit in the SwissProt, 21 Pfam database, considered non-coding by the CPC2, and not be already 22 classified as another class of functional RNA (rRNA, miRNA, tRNA, 23 transposons). The candidate IncRNA were submitted as a separate gff file 24 (GSE129667, Candidate IncRNA.gff).

Functional annotation for the protein coding genes were generated using annotF v1.02 (https://github.com/EI-CoreBioinformatics/AnnotF) and was submitted to GEO (GSE129667, Functional_annotation.txt.gz). Assessments of whether *M. persicae* genes are expressed in the salivary glands and guts was done by performing blastn against the EST datasets of Ramsey et al., 2007 (28).

30

31 Differential gene expression analysis

1 Differentially expressed transcripts of *M. persicae* colonies on the 9-plant 2 species was determined by comparing transcript expression levels of M. persicae colonies on B. rapa (original host) with those of colonies on the 8 other 3 4 plant species (new hosts) (see experimental design in Fig. 1A) using the 5 DESeq2 package in R (v1.2.10) (12) and transcript count per million (TPM) generated by Kallisto v0.42.3 (https://pachterlab.github.io/kallisto/) (S/ Dataset 6 7 **S1**). Additional filtering was employed in DESeg2 to remove lowly expressed 8 transcripts (mean count < 10) on the basis of normalized counts. Transcripts 9 were considered differentially expressed if they had a p value less than 0.05 after 10 accounting for a 5% FDR according to the Benjamini-Hochberg procedure and if 11 log2Fold change was greater than 1.

12

13 **Co-expression analysis**

14 Weighted gene co-expression network analysis (WGCNA) was used to 15 generate unsigned co-expression networks on nine host swap data (13). Genes 16 with normalized count (TPM) > 5 in at least one sample per plant host were used 17 for the co-expression analysis and clustered into network modules using the 18 topological overlap measure (TOM). Genes were grouped by hierarchical 19 clustering on the basis of dissimilarity of gene connectivity (1-TOM). The co-20 expression clusters were produced by cutreeDynamic in which the minimum size of modules was kept at 20 genes. The modules were randomly colour-labelled. 21 22 An adjacency matrix was built by applying a power function (β) on the Pearson 23 correlation matrix. A β value of 18 was found to be optimal for balancing the scale-24 free property of the co-expression network and the sparsity of connections 25 between genes.

26

27 Manual annotation of Ya genes

A 148 bp nucleotide sequence (*SI Appendix*, Fig. S8) that includes the entire exon 2 of the Ya genes and that was found to be conserved among 23 IncRNA genes in the darkslateblue module was used to search the *M. persicae* reference genome G006 with Blastn (v2.22) using default parameters. Blast hits,

1 with coverage more than 80 percent were converted to GFF and loaded to the 2 Apollo browser (11) along with the annotated gene models (22) and all assembled 3 transcripts herein before the merging step (see above). Gene models and 4 corresponding transcripts that aligned to the 148-bp nucleotide sequence were 5 selected and further curated by manually annotating the 3' ends of each of the transcripts based on the presence of a poly-A tail. The 5' ends were identified 6 7 based on the most conserved sequence among all transcripts combined with 8 existing RT-PCR data for Ya1 (SI Appendix, Fig. S12). The curated annotation of 9 the Ya family enabled more accurate transcript quantifications among its members. 10 We updated the gene models in our previous gff file for the Ya locus and submitted 11 as a separate file (GSE129667, Manual Anno.V3.gff). In addition to *M. persicae*, 12 we also did manual annotation of Ya genes in five other aphid species including A. 13 pisum, M. cerasi, A. glycines, R. padi and D. noxia using publicly available RNA-14 seq data and genome assemblies (SI Dataset S5) and was submitted to GEO 15 (GSE129667).

16

17 Identification *M. persicae* RNAs in plants

18 To assess if aphid translocate transcripts into plants, twenty adult *M. persicae* 19 were caged on rosette leaves of 4-week-old A. thaliana plants for 24 hrs. The 20 caged leaf area with aphids was assigned 'aphid feeding site'. Leaves on plants 21 caged with empty cages were included as controls. Leaf areas covered by the 22 cages were carefully washed three times with deionized water and three time with 23 nuclease-free water. RNAs were isolated from four independent biological 24 replicates of aphid-exposed leaves and non-exposed control leaves and 25 processed for RNA-seq library synthesis and sequenced on the Ilumina HiSeq 26 25000 (Novogene, Beijing, China). Reads were trimmed to remove sequencing 27 genome adapters and aligned to Α. thaliana (TAIR10 database. 28 http://arabidopsis.org) and the *M. persicae* G006 genome (22) with HISAT2 v2.0.5. 29 Reads mapped to the *M. persicae* genome were retrieved and subjected to further 30 filtering by mapping them back to the A. thaliana genome. Reads that did not align 31 to the A. thaliana genome in the last step were considered as unique M. persicae

1 mapping reads. Transcripts with TPM \geq 50 in at least one sample and that were 2 present in at least three samples were selected for further analysis.

3

4 **RT-PCR** analyses to detect systemic migration of aphid transcripts in plants

5 Systemic migration of aphid transcripts was determined by caging a leaf 6 section with aphids and detection of aphid transcripts in the caged area (feeding 7 site), next to the caged area of the same leaf (near-feeding site) and a distal leaf 8 (distal site). See experimental setups shown in Fig. 4A and SI Appendix, Fig. 9 **S14**. Plants exposed to cages without aphids were used as controls. For A. thaliana plants, aphids were caged at the distal halves of the 8th rosette leaf (14) 10 of 4-week-old A. thaliana plants for 24 hrs. The proximal leaf area of the 8th leaf 11 next to the cage near the petiole was assigned near-feeding site and the 5th leaf 12 that is likely phloem-connected with the 8th leaf (14) of A. thaliana plants the distal 13 14 site. Similar setup as for A. thaliana were used for B. rapa, P. sativum and maize 15 plants, except that for maize the near-feeding site was the middle of the leaf next 16 to the caged area and distal site the part of this leaf near the stem (SI Appendix, 17 Fig. S14). Upon the 24 hr period of exposure to the cages with or without aphids, 18 sections of leaves that were caged were immediately separated from the non-19 caged parts with scissors and the cages and aphids removed. Then the remaining 20 parts of the leaves and distal leaves were detached from the plants. Leaf sections 21 detached from plants were cleaned with the brush to remove aphids and visible 22 debris. The leaf tissues were submerged in 5 mL of MilliQ water in a 15 mL tube. 23 The tube was shaken for 30 sec after which the MilliQ water was removed. This 24 was repeated two more times with MilliQ water and three additional times with 25 nuclease-free water. Samples were snap-frozen in liquid nitrogen and storage at -80 °C. 26

Total RNAs were isolated from the leaf tissues by RNeasy Plant Mini Kit (Qiagen, Germany) followed by a DNase treatment with RNase-free DNase I (Thermo Fisher Scientific, US). cDNA was synthesized from 1 µg total RNA at 20 µL reaction volume with poly(A) primers using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, US). The qRT-PCRs reactions were

1 performed on a CFX96 Touch[™] Real-Time PCR Detection System using 2 transcript-specific primers (SI Appendix, Table S3). Each reaction was performed in a 20 µL reaction volume containing 10 µL SYBR Green (Maxima SYBR 3 Green/ROX gPCR Master Mix, Thermo Fisher Scientific, USA), 0.4 µL Rox 4 5 Reference Dye II, 1 µL of each primer (10 µM), 1 µL of 20 µL sample cDNA, and 7.6 µL UltraPure Distilled water (Invitrogen, US). The PCR cycles were: 95 °C for 6 7 10 s, 40 cycles at 95 °C for 20 s, 63 °C for 30 s. The identities of PCR products 8 were verified by Sanger sequencing.

9 To compare Ya1 transcript concentrations in the feeding and distal sites, a 10 standard Ya1 concentration curve was generated (SI Appendix, Fig. S13). For 11 this, the 273-nt Ya1 fragment cloned into the plasmid pBI121 under promoter 35S 12 was used as the PCR template. The concentration curve was generated with a 13 serial of dilution of the plasmid pBI121 35S::Ya1 from the highest concentration 1.44E⁻¹⁰ g/ μ L to the lowest concentration 1.44E⁻¹⁸ g/ μ L and primers (Ya1 primer6 14 and Ya1 primer9, SI Appendix, Table S3) using PCR conditions as described 15 16 above.

17

18 Primer design for amplification of *M. persicae* Ya transcripts

Primers to specifically amplify *M. persicae* Ya transcripts were designed with the PrimerQuest Tool (Integrated DNA Technologies, IA, USA) that predicted five to ten primers for each of the Ya transcripts. The primer pairs were aligned to the sequences of all Ya transcripts and the ones that matched unique sequences of one Ya transcript selected (*SI Appendix*, Table S3). Ya transcripts for which no unique primers were available were excluded from further analyses.

25

26 **Detection of aphid transcripts by northern blotting**

For northern blot analyses, 1 ug of total plant and aphid RNA were separated on a 6% denaturing polyacrylamide gel (PAGE) with 1 X TBE buffer (10 X TBE buffer stock, Thermo Fisher Scientific, USA) at 100 V for 60 min. The 273-nt Ya1 fragment was cloned with a SP6 sequence at the 3 the end of Ya1 and used to synthesize a 291-nt Ya1-SP6 RNA of which 100 ng was ran alongside the total

RNAs from aphids and plants on polyacrylamide gels. RNAs were transferred to a nylon membrane (Hybond N, Amersham, UK) by electroblotting at 0.8 A for 2 hrs (BIO-RAD, USA) using 0.5 X TBE buffer at 60 V for 60 min. RNAs were crosslinked to the membranes by a UV cross linker (UVP Inc., CA, USA) using autocrosslink function, twice on the side of the blot exposed to the gel and one time on the other slide of the blot.

7 Ya1 transcript in aphids were detected via hybridization of the northern blots 8 to a biotin-labelled anti-sense sequence of Ya1. Biotin-labelled anti-sense probe 9 of Ya1 was synthesized from anti-sense sequences of Ya1 with the MAXIscript[™] 10 SP6/T7 Transcription Kit (Thermo Fisher Scientific, USA) and Biotin-16-dUTP 11 (Roche, USA). The northern blots were performed according to the manufacturing 12 manual of North2South Chemiluminescent Hybridization and Detection Kit 13 (Thermo Fisher Scientific, US). Briefly, the blot was washed in North2South 14 Hybridization Stringency Wash Buffer at room temperature for one time 20 min, in 15 Wash Buffer containing 2X SSC/0.1% SDS at room temperature for three times 20 16 min, and Stringency Wash Buffer at 65 °C for one time 20 min. The washing was done in 0.2 ml per cm² of blot. The blot was then incubated in the Substrate 17 18 Working Solution containing equal volumes of the Luminol/Enhancer Solution and 19 Stable Peroxide Solution for 5 min. To visualize the hybridization signal, the 20 membrane was exposed to an X-ray film for an appropriate exposure time. 21 Prestain RNA markers, DynaMarker Prestain Marker for RNA High (catalog 22 number DM260, BioDynamics Laboratory Inc. Japan) and DynaMarker® Prestain 23 Marker for Small RNA Plus (catalog number DM253, BioDynamics Laboratory Inc. 24 Japan) were used to estimate the size of RNAs.

Ya1 transcript in plants were detected via hybridization of the northern blots to a 273-nt Ya1 fragment that was labeled using 3000 Ci/mmol of [a-32P] dATP (PerkinElmer Life Sciences. USA) with the Klenow DNA polymerase reaction as per manufacturer's instruction (Megaprime DNA Labeling System, GE Healthcare).The radioactively-labelled PCR probe was denatured at 95°C for 5 min, transferred to ice and then incubated with the blot in 50 ml Hybridization buffer (Sigma-Aldrich) at 42°C for overnight. Washing was done three times at 42°C with

washing buffer (2X SSC/0.1% SDS). The membranes were exposed to storage
 Phosphor Screens (GE Healthcare) and hybridization signals were visualized
 using Typhoon Trio (GE Healthcare) scanner.

4

5 Sequencing of 3' ends of aphid Ya1 transcript

Total RNA was extracted from *M. persicae* aphids reared on *B. rapa* and 3 6 7 ug of aphid total RNAs was added to a 80 µL ligation mixture containing 8 µL T4 8 RNA ligase buffer and 4 µL T4 RNA ligase (Thermo Fisher Scientific, USA), 8 µL 9 ATP (Invitrogen, USA), 8 µL BSA (Invitrogen, USA), and 10 pmol 3' RACE RNA 10 adaptor (SI Appendix, Table S3). RNA ligation was carried out at 16 °C overnight. 11 The ligated RNA was converted into cDNA using oligo sequences complementary 12 to the 3' RACE adaptor with RevertAid First Strand cDNA Synthesis Kit (Thermo 13 Fisher Scientific, USA) (SI Appendix, Table S3). RACE PCRs were performed 14 with the Ya1 forward primer GGACAAGTCCAATCTGC and the adapter primer 15 CAAGCAGAAGACGGCATACGA (SI Appendix, Table S3) in a 50 µL reaction 16 volume containing 0.5 µL Phusion DNA polymerase (NEB), 10 µL 5X Phusion HF 17 buffer, 1 µL 10 mM dNTPs, 1 µL of each primer (10 µM), 1 µL of cDNA sample. The cycle programs were: 98 °C for 30 s, 35 cycles at 98 °C for 10 s, 60 °C for 30 18 19 s, 72 °C for 15 s, final extension 72 °C 10 mins. PCR produces were separated by 20 3% agarose gel. DNA bands were visualized under UV light, then cut to extract 21 DNA using QIAquick Gel Extraction Kit (Qiagen, Germany). DNA were ligated to 22 pGEM-T (Promega, USA) and Sanger sequenced with M13 primer.

23

24 Plasmid construction

To generate pJawohl8-RNAi constructs, a 273 bp of Ya1 was amplified from *M. persicae* cDNA by PCR with specific primers containing additional attB1 and attB2 linkers (*SI Appendix*, Table S3) for cloning with the Gateway system (Invitrogen, USA). The 273-bp Ya1 fragment was introduced into pDONR207 (Invitrogen, USA) plasmid using Gateway BP reaction and transformed into DH5α competent cells (Invitrogen, USA). Subsequent clones were sequenced to verify correct size and sequence of inserts. Via the Gateway LB reaction, inserts were

1 transferred from pDONR207 into the plant transformation vector, pJawohl8-RNAi 2 (kindly provided by I.E. Somssich, Max Planck Institute for Plant Breeding 3 Research, Germany), which is a plasmid that enables the expression of the 4 transgene as a double-stranded hairpin transcript, generating plasmid pJawohl8-5 RNAi Ya1.

Plasmids pBI121 35S::Ya1 and pBI121 35S::Ya1 3TAG that produced Ya1 6 7 and Ya1 3UAG transcripts were constructed as follows. The fragment 8 corresponding to the 273 nt Ya1 transcript was amplified from *M. persicae* cDNA 9 by PCR with specific primers containing *Bam*HI and *SacI* restriction sites in the 10 forward and reverse primers, specifically (SI Appendix, Table S3). The BamHI 11 and Sacl digested PCR fragments were introduced into pBI121 to generate 12 pBI121 35S::Ya1. Plasmid pBI121 35S::Ya1 was used in overlap PCR reactions 13 (15) to generate the Ya1 3TAG construct in which the three ATGs at positions 41, 14 62, and 146 of Ya1 were converted into TAG stop codons. This involved the 15 amplification of fragments 1, 2, 3 and 4 with primer pairs MutateF and ATG41 R, 16 ATG41 F and ATG62 R, ATG62 F and ATG146 R and ATG146 F and 17 MutateR, respectively (see SI Appendix, Table S3 for primer sequences), and 18 subsequent amplification of a 81 bp fragment from fragments 1 and 2 with primers 19 MutateF and ATG62 R, a 230 bp fragment from fragments 3 and 4 with primers 20 ATG62 F and MutateR, and finally the amplification of the 273 nt Ya1 mutant in 21 which three ATGs were mutated to three TAGs by combining the 81 bp and 230 22 bp fragments and primers MutateF and MutateR. The resulting Ya1 3TAGs PCR 23 fragment was digested with BamHI and SacI and introduced into pBI121 to 24 generate pBI121 35S::Ya1 3TAG.

25 The sequences of plasmid inserts were verified by Sanger sequencing using 26 specific primers (SI Appendix, Table S3).

27

28 Generation of transgenic *A. thaliana* plants

29 pJawohl8-RNAi Ya1, pBI121 35S::Ya1, and pBI121 35S::Ya1 3TAG 30 plasmids were introduced into A. tumefaciens strain GV3101 that carried the 31 helper plasmid pMP90RK for subsequent transformation of A. thaliana Col-0 using

1 the floral dip method (16). Transgenic seeds were selected on Murashige and 2 Skoog (MS) medium supplemented with 20 µg/mL phosphinothricin (BASTA) to 3 select dsRNA Ya1 transformants or on MS containing 50 µg/mL kanamycin to selection of 35S::Ya1 and 35S::Ya1 3TAG transformants. F2 seeds were 4 5 germinated on MS medium supplemented with 20 µg/mL BASTA or 50 µg/mL kanamycin for selection. F2 seedlings with 3:1 alive/dead segregation (evidence 6 7 of single insertion) were taken forward to the F3 stage. Seeds from F3 plants were 8 sown on MS with BASTA (for dsRNA Ya1 plants) or on MS with kanamycin (for 9 35S::Ya1 and 35S::Ya1 3TAG) and lines with 100% survival ratio (homozygous) 10 were selected.

11

12 Knock down of aphid transcripts by plant-mediated RNA interference (RNAi)

13 Seeds of the dsRNA Ya1 homozygous lines (expressing dsRNA 14 corresponding to Ya1) were sown on MS medium and, after one week, seedlings 15 were transferred to single pots (8 cm diameter) and transferred to a controlled 16 environmental growth room at temperature 24 °C day/20 °C night under 10 hours 17 of light. One *M. persicae* adult was confined to single 4-weeks-old Arabidopsis 18 lines in sealed experimental cages (15.5 cm diameter and 30 cm height) containing 19 the entire plant. One day later, the adult was removed, five nymphs remained on 20 the plants and become adults in five days. Adults were harvested for gRT-PCR to 21 confirm Ya1 silencing compared to adults on dsGFP plants.

22 For gRT-PCRs, total RNA was isolated from aphids using Trizol reagent 23 (Sigma) and subsequent DNase treatment using an RNase-free DNase I (Thermo 24 Fisher Scientific, US). cDNA was synthesized from 1 µg total RNA with RevertAid 25 First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, US). The qRT-PCRs reactions were performed on CFX96 Touch™ Real-Time PCR Detection System 26 27 using gene-specific primers. Each reaction was performed in a 20 µL reaction 28 volume containing 10 µL SYBR Green (Thermo Fisher Scientific, US), 0.4 µL Rox 29 Reference Dye II, 1 µL of each primer (10 µM), 1 µL of sample cDNA, and 7.6 µL 30 UltraPure Distilled water (Invitrogen, US). The cycle programs were: 95 °C for 10 31 s, 40 cycles at 95 °C for 20 s, 60 °C for 30 s. Relative quantification was calculated

using the comparative 2^{-△}Ct method (17). All data were normalized to the level of
 elongation factor gene (MYZPE13164_G006_v1.0_000087220) from the same
 sample.

4

5 *M. persicae* fecundity assay

6 Seeds of Arabidopsis lines were sown on MS medium and, after one week, 7 seedlings were transferred to single pots (8 cm diameter) and transferred to a 8 controlled environmental growth room at temperature 24 °C day/20 °C night under 9 10 hours of light. One *M. persicae* adult was confined to single 4-weeks-old 10 Arabidopsis lines in sealed experimental cages that contained the entire plant. One 11 day later, the adult was removed, and one nymph remained on the plants. Offspring produced on the 7th, 9th, 11th day of the experiment were scored and 12 13 removed. This experiment was repeated three times to create data from three 14 independent biological replicates with four to six plants per line per replicate.

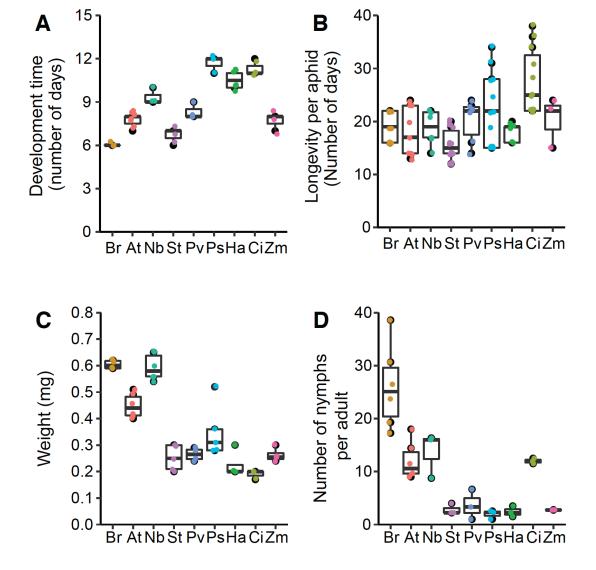
15

16 Statistical analyses

17 All the data analyses were performed in R (v3.5.2). All statistical tests are 18 described in the figure legends.

1 Supplementary Figures and Tables

2



3

4 **Fig. S1.** Performance parameters of stable colonies of *M. persicae* clone O on 9

5 divergent plant species. Colonies on *B. rapa* (Br), *A. thaliana* (At), *N.*

6 benthamiana (Nb), S. tuberosum (St), C. indicum (Ci), H. annuus (Ha), P.

7 sativum (Ps), P. vulgaris (Pv) and Zea mays (Zm), established as shown in Fig.

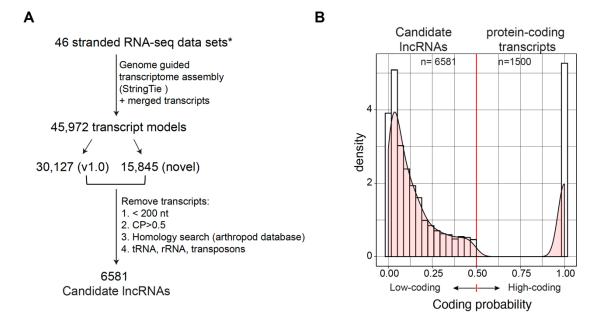
8 **1A**, were analysed for development time (A) in days between birth and

9 emergence of adults (n = 3-7 plants with one aphid each), longevity (B) in days

10 between birth to death (n = 3-15 plants with one aphid each), weight (C) in mg

per adult (n = 3-6 plants with 100 adult aphids each), and number of nymphs per

- 1 adult (D) calculated by dividing the number of nymphs by the number of adults on
- 2 days 7, 9 and 11 (Br, At, Nb and St) or days 7, 9, 11, 13, 15, 17 and 18 (Ci, Ha,
- 3 Ps, Pv and Zm) (n = 3-6 plants seeded with one aphid.



2 Fig. S2. Annotation of *M. persicae* transcript models and candidate IncRNAs. (A) Transcript annotation pipeline. RNA-seq reads from 45 libraries generated from M. 3 persicae on nine divergent hosts (data generated herein, SI Appendix, Fig. S1A) 4 5 and LIB1771 derived from a *M. persicae* colony on *B. rapa* (1) were assembled 6 into transcripts using a genome-guided approach with StringTie (5). Of the 45,972 transcripts in total, 30,127 were annotated previously (1) and 15 845 are novel. 7 8 The 6 581 transcripts that are candidate IncRNAs were identified upon removing 9 transcripts of < 200 nt in size, transcripts with coding potential (CP) >0.5, as 10 determined using the Codina Potential Calculator 2 (CPC2, http://cpc2.cbi.pku.edu.cn/) (10), and that have similarities of deduced protein 11 sequences to known arthropods proteins, house-keeping RNAs (rRNA and tRNA) 12 and transposons. (B) Distribution of CPC2 coding probability scores of 6,581 13 14 candidate IncRNAs identified in Fig. S2A and 1,500 randomly selected transcripts 15 from *M. persicae* protein-coding sequences.

16

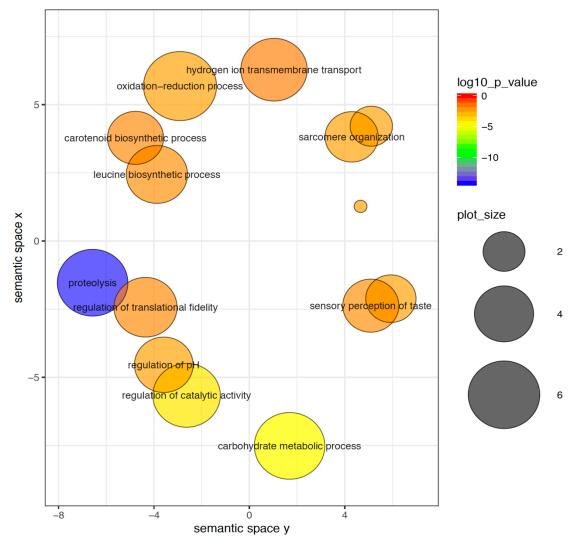
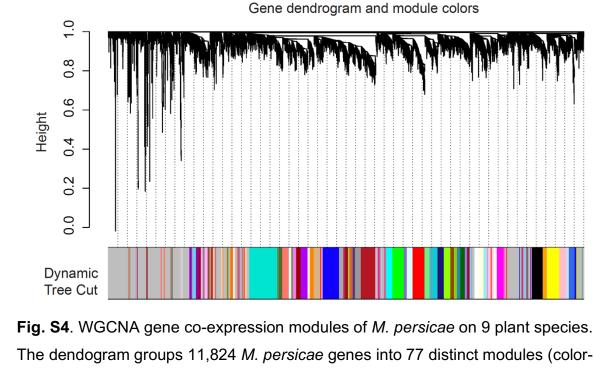




Fig. S3. Biological process GO terms that are enriched among DE genes of M. 2 3 persicae on 9 plants species. Graphs was generated from data shown in Fig. 1B. 4 Enriched GO terms are represented by circles and are clustered according to 5 semantic similarities to other GO terms in the gene ontology. Distance between 6 the circles indicates the semantic similarity between the corresponding GO terms. 7 Circle size is proportional to the frequency of the GO term, whereas colour 8 indicates the $log_{10}(p \text{ value})$ for the enrichment calculated using GO-seq (red 9 higher, blue lower). Only GO-terms having p value and padj < 0.05 are shown. Data were summarized and visualized by REVIGO online tool (http://revigo.irb.hr/) 10 11 (18).



4 coded) using hierarchical clustering. Dynamic cut tree was used to identify 5 modules, dividing modules at significant branch points in the dendrogram. The x-6 axis shows the width of each module that is defined by the number of genes. The 7 y-axis corresponds to distance determined by the extent of topological overlap (1-8 TOM). Genes not assigned to a module are labelled in grey.

1 2

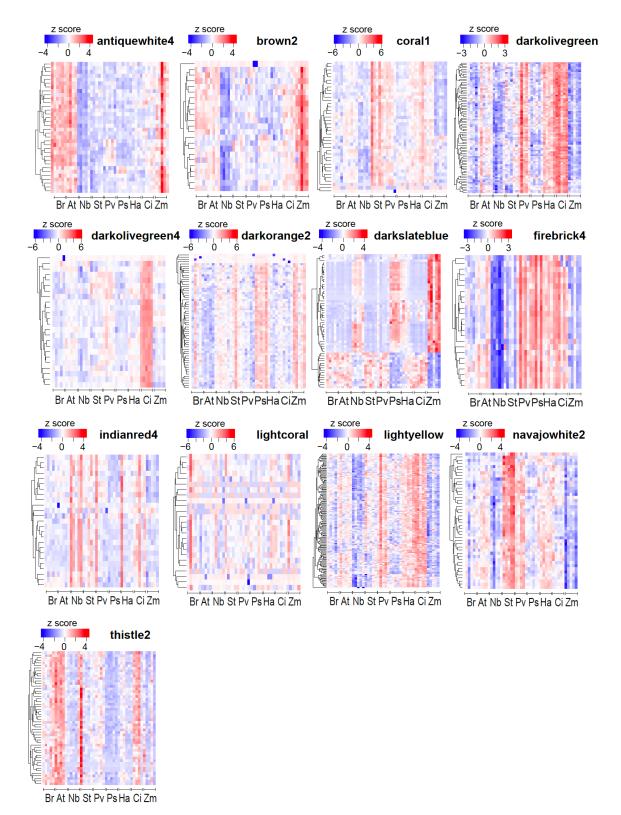
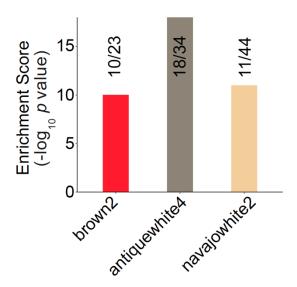


Fig. S5. Heatmaps of 13 modules enriched for DE genes of *M. persicae* colonies
on 9 plant species. The colour codes of the modules are indicated above the

- 1 heatmaps. Hierarchical clustering was done based on plant species as follows: Br,
- 2 B. rapa; At, A. thaliana; Nb, N. benthamiana; St, S. tuberosum; Ci, C. indicum; Ha,
- 3 H. annuus; Ps, P. sativum; Pv, P. vulgaris; Zm, Z. mays. Rows are scaled based
- 4 log transformed TPM values, which are shown using a z-score as indicated.
- 5

Modules enriched for genes expressed in guts



1

Fig. S6. Modules enriched for genes expressed in *M. persicae* guts. *M. persicae*DE genes were searched against a gut EST dataset (28). The x/y numbers above
the bars indicate the number of genes in the enriched category (x) and the total
number of genes in the module (y).

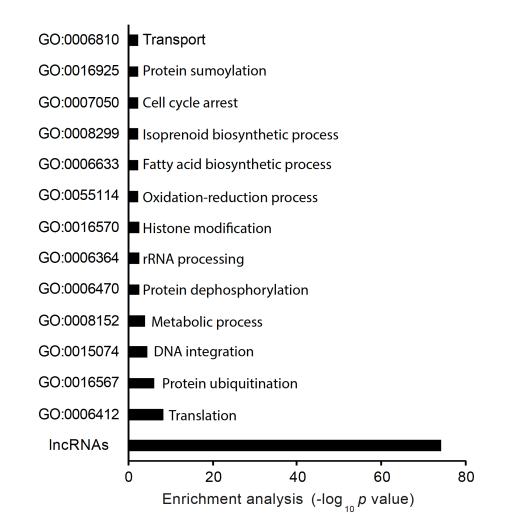
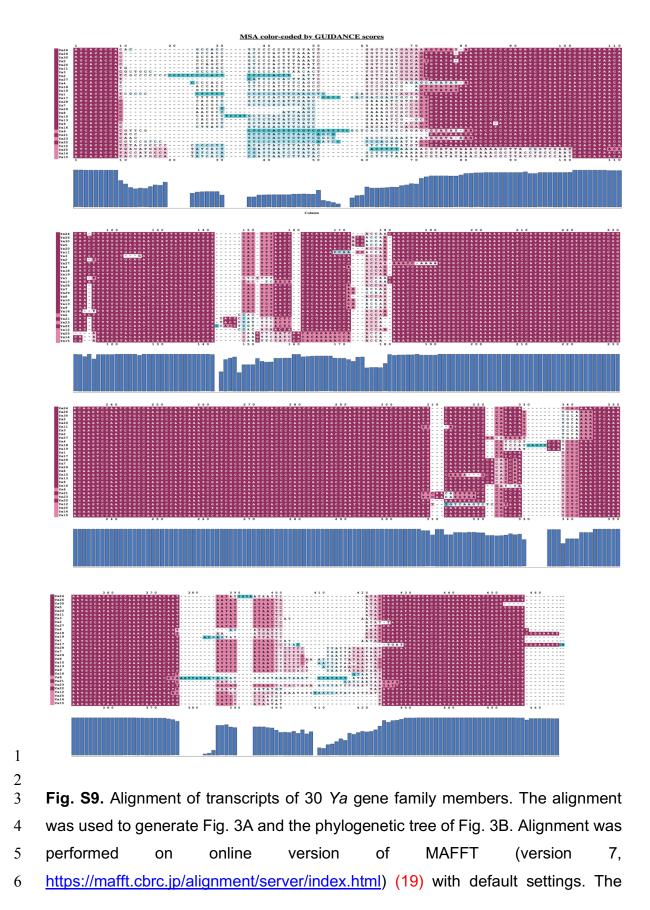


Fig. S7. Enrichment analysis of aphid transcripts found in the feeding sites. Gene functional annotations were used. Enrichment analysis was calculated with Fisher's Exact Test by comparing the ratios of genes that belong to a GO category versus the total number of genes between the feeding site and the whole aphid.

>Ya1_exon2_148nt TTTCTCTTTTAAACCTAAAAAACCAACCAACAAATCAAAAATGGGCGCTGAAAAGGTATCCATGAACA TCGTTGTCGTCGGACAAGTCCAATCTGCCAAGGCCATCAAGACCGTCGCCAAGGCCTCCCCAGCCACC AACCAATTGATG

1

- 2 Fig. S8. Nucleotide sequence of Ya1 exon 2 used to identify Ya family members
- 3 in *M. persicae* and other aphid species.





multiple alignment was subjected to GUIDANCE2 (<u>http://guidance.tau.ac.il/ver2/</u>)
(20) to compute the residue-wise confidence scores and extract well-aligned
residues. Nucleotides are highlighted (colour scale between red and turquoise) to
indicate confidence of alignment (red is high confidence and turquoise is low). The
bottom bar graph displays the scores of aligned nucleotides, as determined by
GUIDANCE2.

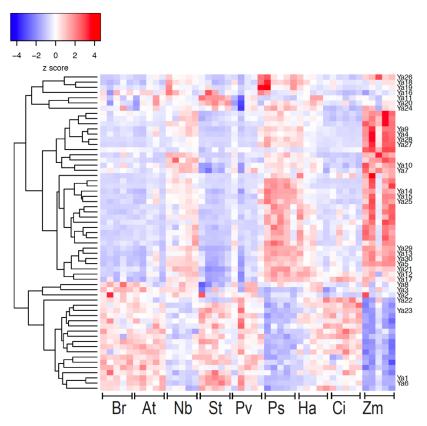


Fig. S10. Heatmap of darkslateblue module after manual annotation of the Ya
family. WGCNA analysis was repeated with the corrected set of Ya genes.
Heatmap was generated using log-transformed TPM values of all genes in the
module. Br, *B. rapa*; At, *A. thaliana*; Nb, *N. benthamiana*; St, *S. tuberosum*; Ci, *C. indicum*; Ha, *H. annuus*; Ps, *P. sativum*; Pv, *P. vulgaris*; Zm, *Z. mays.* Rows are
scaled and represented as z-score.

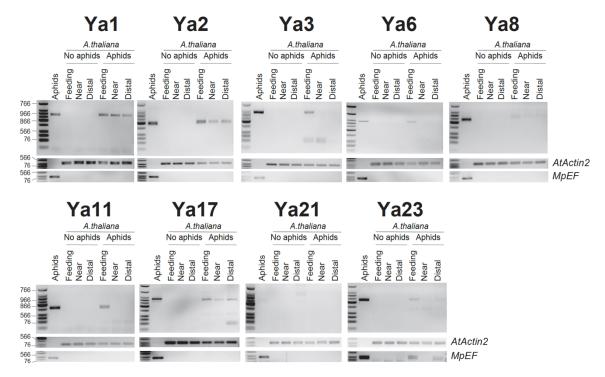
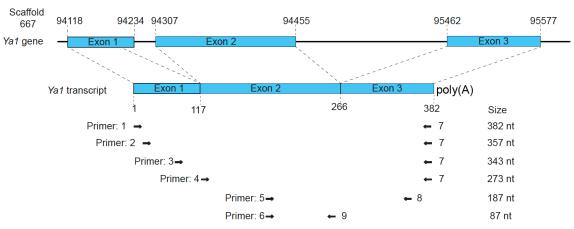




Fig. S11. RT-PCR experiments showing translocation of 9 *Ya* transcripts into plants and systemic migration of *Ya1*, *Ya2*, and *Ya17* transcripts. Experimental setup is shown in Fig. 4A. The *Ya* transcripts were amplified by RT-PCR using gene-specific primers (*SI Appendix*, Table S3) and PCR products were separated by 3% agarose gel. The bands corresponding to the PCR products were extracted from the gel and sequenced directly by forward and/or reverse primers to verify the identity of the *Ya* sequence.





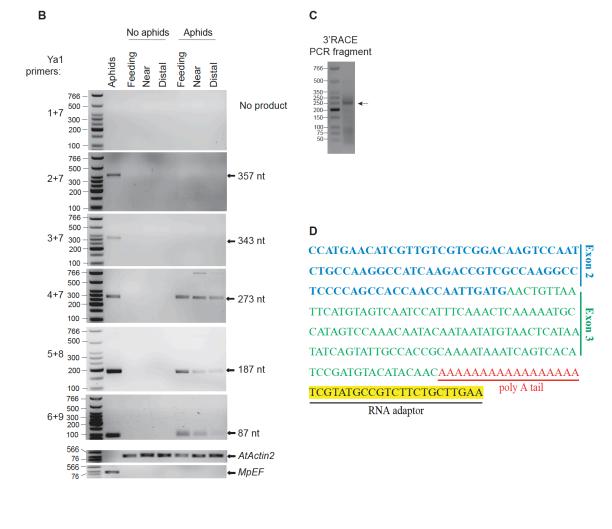
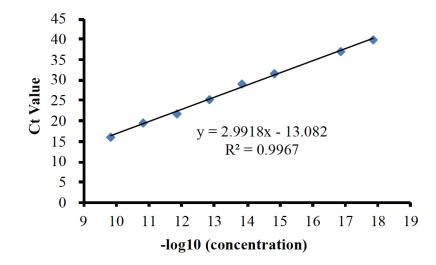


Fig. S12. Identification of *Ya1* transcript sizes in aphids and plants. (A) Transcript
 model of the *Ya1* gene. The three exons of the *Ya1* gene that form the transcript

1 are indicated as blue rectangles. The numbers above the Ya gene indicate the first 2 and last nucleotide of each exon on scaffold 667 of the *M. persicae* genome assembly. The locations of the Ya1 primers 1-9 (SI Appendix, Table S2) used to 3 4 assess the presence of Ya1 transcripts in aphids and in plants by RT-PCR are 5 indicated as black arrows. (B) A 357-nt Ya1 transcript was detected in aphids and not in plants, whereas the largest Ya1 transcript detected in plants at aphid feeding 6 7 sites and distal leaves is 273-nt (primers 4 and 7). Experimental setup is shown in 8 Fig. 4A. RT-PCR products generated with primers shown in A were separated on 9 3% agarose gels. Fragment sizes are indicated at right of the gels and correspond 10 with the expected sizes as shown in A. Primers 1 and 7 did not generate a product 11 in aphids nor in plants. (C) Results of a 3' RACE experiment. The 3' RACE PCR 12 products were separated a 3% agarose gel. The band isolated from the gel and 13 sequenced is shown with an arrow at right of the gel. (D) Sequence of the 3' RACE 14 PCR fragment of the band shown in C. The sequence is identical to the sequences 15 of the end of exon 2 (letters in blue font) and the entire exon 3 (letters in green 16 font) of the annotated Ya1 gene in the M. persicae genome. The Ya1 transcript 17 has a poly(A) tail (letters in red font). The RNA adaptor sequence used in 3' RACE 18 protocol is indicated in yellow.



2 Fig. S13. Standard concentration curve of Ya1 and cycles of PCR amplification.

The 273-nt Ya1 fragment in the plasmid pBI121 was used as the PCR template. QPCR were performed on a serial of dilution of the plasmid pBI121_35S::Ya1 from the highest concentration $1.44E^{-10}$ g/ µL to the lowest concentration $1.44E^{-18}$ g/ µL.

6 x axis are minus log 10 transferred concentrations.

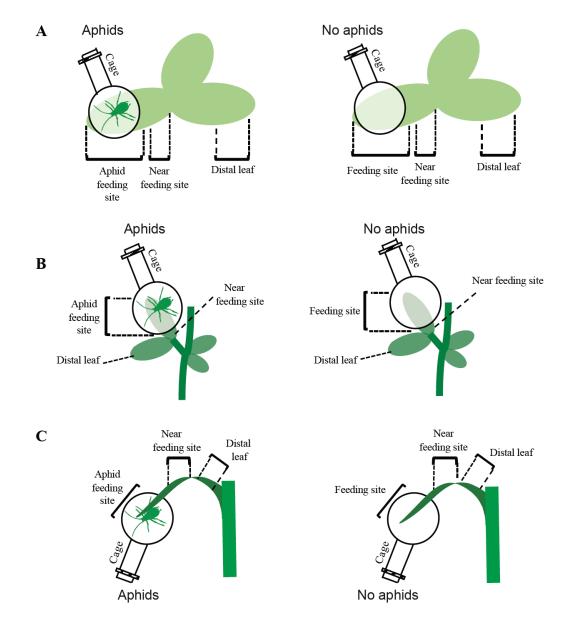


Fig. S14. Schematic overview of experimental setups to investigate systemic migration of aphid Ya1 transcripts in *Brassica rapa*, *Pisum sativum* and *Zea mays*.
(A) *Brassica rapa*. (B) *Pisum sativum*. (C) *Zea mays*. Aphid feeding site is the caged parts of the leaf. Distal sites include near-feeding sites that are located just externally of the caged leaf areas and more distal locations that are different leaves for *B. rapa* and *P. sativum* and further away of the caged area on the same leaf of *Z. mays*.

A 1 TTTCTCTTTTAAACCTAAAAAACCAACCAACAAATCAAAAATGGGCGCTGAAAAGGTATCCATGAACATCGTTGTCGTCGGA 82 83 CAAGTCCAATCTGCCAAGGCCATCAAGACCGTCGCCAAGGCCTCCCCAGCCACCAACTAATTGATGAACTGTTAATTCATGT 164 165 AGTCAATCCATTTCAAACTCAAAAATGCCATAGTCCAAACAATACAATAACAATAATATGTAACTCATAATATCAGTATTGCCACCG2 246 247 AAAATAAATCAGTCACATCCGATGTAC 273

в		Forward primer	Reserve primer
	Ya1 sequence	e 60 CCATGAACATCGTTGTCGGCCAAGGCCCAAGGCCATCAAGACCGTCGCCAAGGCCTCC	COAGCCACCAACCAATTGATG 148
а	Feeding site	Manna Malan Malan Man Man Man Man Man Man Man Man Man M	mmmm
A.thaliana	Near- feeding site	mannahamandahamanahaman	mmmmm
	Distal leaf	mana man and a man and a man and a man	And Markan
за	Feeding site	kan mar and a star and the second sec	Marka Marka
B. rapa	Near- feeding site	my hand hand hand had marked by hand hand hand hand hand hand hand hand	Mm Mm Mm Mar Mar
	Distal leaf	manna han han han han han han han han han	mmmmmm
m.	Feeding site	montheman man and man a	Mar
P. sativum	Near- feeding site	man	manam
	Distal leaf	monorman and an and and an an an an and and	mmmmmm
S	Feeding site ∭	month and	mmmmm
Z. mays	Near- feeding site	month and	mmmmm
	Distal leaf	monthalimentalime	Multi Marine M

Fig. S15. Sequences of RT-PCR products of *M. persicae* Ya1 transcripts amplified from plants. The Ya1 transcripts were amplified with Ya1 forward and reverse primers from feeding sites, near-feeding sites, and distal sites of *A. thaliana*, *B. rapa*, *P. sativum* and *Z. mays* as per experimental setups shown in Fig. 4A and Suppl. Fig. S13. The sequences shown were identical to the corresponding region within the Ya1 gene.

8

33S::Yal lineline_8-842 line_9-941tttcttttaaacctaaa tttcttttaaacctaaa aaccaaccaacaacaaa aaccaaccaacaacaaa aaccaaccaacaacaaa atgggegtgaaaggtatecatgaacatgttgegte edgacatgtgtgee ggacaagtccaatctgccaa aggacagtccaatctgccaa aaggacgtgaaaggtatecatgaacatgttgegte edgacatgtgtgee ggacaagtccaatctgccaa aggacagtccaatctgccaa tagggegtgaaaggtatecatgaacatgttgegte edgacatgtgtgee ggacaagtccaatctgccaa taggacgtgaaaggtatecatgaacatgttgegte edgacagtccaatctgccaa taggacagtccaatctgccaa tagggegtgaaaggtatecatgaacatgttgegte edgacagtccaatctgccaa taggacagtccaatctgccaa tagggegtgaaaggtatecatgaacatgttgegte edgacagtccaatctgccaa taggacagtccaatctgccaa tagggegtgaaaggtatecatgaacatgttgegte edgacagtccaatctgccaa taggacagtccaatctgccaa tagggegtgaaaggtatecatgaacatgttgegte edgacagtccaatctgccaa taggacagtcgaaaggtatecatgaacatgttgegte edgacagtccaatctgccaa taggacagtcgaaaggtatecatgaacatgttgegte edgacagtccaatctgccaa taggacagtgaaaggtatecatgaacatgttgegte edgacagtccaatctgccaa taggacagtcgaaaggtatecatgaacatgttgegte edgacagtccaatctgccaa taggacagtcgaaaggtatecatgaacatgttgegte edgacagtcgacatgttgegte ggacaagtccaatctgccaa taggegetgaaaggtatecatgaacatgttgegte edgacatggttgegte ggacaagtccaatctgccaa taggegetgaaaggtatecatgaacatgttgegte edgacatggttgegte ggacaagtccaatctgccaa taggegetgaaaggtatecatgaacatgttgegte edgacatggttgegte ggacaagtccaatctgccaa taggegetgaaaggtatecatgaacatgttgegte edgacatggttgegte edgacatggttgegte ggacaagtccaatcgcacac taggegetgaaaggtatecatgaacatgttgegte edgacatggttgegte edgacatggttgegte taggegetgaaaggtatecatgaacatgttgegte edgacatggttgegte edgacatggttgegte edgacatggttgegte taggegetgaaaggtatecatgaacatggtgegetgaacggt	Ya1 reference		Froward primer
Ine 35S:YalleeIntercentination 1 and 2	sequences	Ya1 273 nt	tttetettttaaasetaaaa aaccaaccaacaaatcaaaa atgggegetgaaaaaggtate catgaacategttgtegte ggacaagtecaatetgecaa
358: YallinsImeIttertitiaaaretaaaaccaacaacaacaacaacaacatgaggetgaaaggtatcatgaacatgttytotoggaaagtecaattotgeaa358: YallinsIme_1-141Ittertitiaaaretaaaaccaacaacaacaacaatagggetgaaaggtatctagaacategttytotoggaaagtecaattotgeaa358: Yalline_1-142Ittertitiaaaretaaaaccaacaacaacaacaatagggetgaaaggtatctagaacategttytotoggaaagtecaattotgeaa358: Yalline_1-142Ittertititaaretaaaaccaacaacaacaacaatagggetgaaaggtatctagaacategttytotoggaaagtecaattotgeaa358: Yalline_1-142Ittertititaaaretaaaaccaacaacaacaacaatagggetgaaaggtatctagaacategttytogtoggacagtecaattogeaa34T6s_mulline_1-442IttertititaaaretaaaaccaacaacaacaacaatagggetgaaaggtatctagaacategttytogtoggacagtecaattogeaaYal refereerYal_273_ntggccatcaagacegtogeaaggettececagcaccaacaattgatgaacgttatcatgaactacattataactgaacgtecattita358: YallineIme_0-001ggccatcaagacegtogeaaggettececagcaccaacattgatgaacgttatcatgaactecattataactaaaaatgecatgt358: YallineIme_0-001ggccatcaagacegtogeaaggettececagcaccaacattgatgaacgttatatcatgaactacattataactaaaaatgecatgt358: YallineIme_0-001ggccatcaagacegtogeaaggettececagcaccaacattgatgaacgttatatcatgaactacattataactaaaaatgecatgt358: YallineIme_0-001ggccatcaagacegtogeaaggettececagcaccaacaattatgaacgttatatcatgaactacattataactaaaaatgecatgt3416s_mullineIme_0-001ggccatcaagacegtogeaaggettececagcaccaa<		line 8-8#1	tttetettttaaacetaaaa aaccaaccaaatcaaaa atgggegetgaaaaggtate catgaacategttgtegte ggacaagtecaatetgecaa
line_9:91itterrittaarriaaa arccarcaaraatacaaa atgggggtgaaaaggtat catgaacagtitgtegt ggacagtcaattgrealine_9:92itterritterrittaarriaaa arccarcaaraatcaaa atgggggtgaaaaggtat catgaacagtitgtegt ggacagtcaattgrea358:Yal_line_1:11itterritterrittaarriaaa arccarcaaraatcaaa tagggggtgaaaggtat catgaacatgtitgtegt ggacagtcaattgrea358:Yal_line_1:12itterritterrittaarriaaa arccarcaaraatcaaa tagggggtgaaaggtat catgaacatgtitgtegt ggacagtcaattgrea3ATGs_muliesline_1:12itterritterrittaarriaaa arccarcaaraatcaaa tagggggtgaaaggtat catgaacatgtitgtegt ggacagtcaattgrea3ATGs_muliesline_1:12itterritterrittaarriaaa arccarcaaraatcaaa tagggggtgaaaggtat catgaacatgtitgtegt ggacagtcaattgreayal referenceYal_273_ntggccatcaagacgtogca aggecteccagcaccaa caatgatgaacgtitat catgtagtcaatcattea aactcaaaatgccatgtyan sequenceYal_273_ntggccatcaagacgtogca aggecteccagcaccaa caatgatgaacgtitat catgtagtcaatcattea aactcaaaatgccatgtine_9:91ggccatcaagacgtogca aggecteccagcaccaa caatgatgaacgtitat catgtagtcaatcattea aactcaaaatgccatgtine_9:92ggccatcaagacgtogca aggecteccagcaccaa caatgatgaacgtitat catgtagtcaatcattea aactcaaaatgccatgtine_9:91ggccatcaagacgtogca aggecteccagcaccaa caatgatgaacgtitat catgtagtcaatcattea aactcaaaatgccatgtine_9:92ggccatcaagacgtogca aggecteccagcaccaa caatgatgaacgtitat catgtagtcaatcattea aactcaaaatgccatgtine_9:92ggccatcaagacgtogca aggecteccagcaccaa caatgatgaacgtitat catgtagtcaatcattea aactcaaaatgccatgtine_9:92ggccatcaagacgtogca aggecteccagcaccaa caatgatgaacgtitat catgtagtcaatcattea aactcaaaatgccatgtine_9:92ggccatcaagacgtogca aggecteccagcaccaa caatgtagaacgtitatine_9:92ggccatcaagac			tttetettttaaacetaaaa aaccaaccaaccaatcaaaa atgggegetgaaaaggtate catgaacategttgtegte ggacaagtecaatetgecaa
Ine 35S::Yal 34TGs_multineIntelCititiancetama ine_1-142IntelCititiancetama ittelCititiancetama accanceancanateana taggegetgamangtate claganetegttaget ggecangtecantetgee taggegetgamangtate claganetegttaget ggecangtecantetgee ggecangtecantetgee accanceancanateana titelcititiancetama anceanceancanateana taggegetgamangtate claganetegttaget ggecangtecantetgee accanceancanateana taggegetgamangtate claganetegttaget ggecangtecantetgee accanceancanateana taggegetgamangtate claganetegttaget ggecangtecantetgee accanceancanateana taggegetgamangtate claganetegttaget ggecangtecantetgee accanceancanateana taggegetgamangtate claganetegttaget ggecangtecantetgee accanceancanateana taggegetgamangtate claganetegttaget ggecangtecantetgee accanceancanateana taggegetgamangtate claganetegttaget claganetegttaget ggecangtecantetgee accanceancanateana taggegetgamangtate claganetegttaget claganetegttaget claganetegttaget claganetegttaget ggecangtecantegee accanceancanceana taggegetgamangtate claganetegttaget cl	35S::Ya1 lines		tttetettttaaacetaaaa aaccaaccaacaaatcaaaa atggggggtgaaaaaggtate catgaacategttgtegte ggacaagtecaatetgecaa
33S::Yalline_1-142tttcttttaaacctaaaaaccaacaacaacaaatcaaataggggctgaaaggtatoctagacatottgtegtoggacaagtccaatctgacaa33S::Yalline_3-311tttctttttaaacctaaaaaccaacaacaaatcaaatagggcgtgaaaggtatoctagacactotttgtogtoggacaagtccaatctgacaa3ATG: mutlineline_4-441tttctttttaaacctaaaaaccaaccaacaatcaaatagggcgtgaaaggtatoctagacactotttgtogtoggacaagtccaatctgacaa1ine_4-442tttctttttaaacctaaaaaccaaccaacaatcaaatagggcgtgaaaggtatoctagacactotttgtogtoggacaagtccaatctgacaa3argemereeral_273_ntggccatcaagacgtgccaaggcctcccagccaccaaccaattgtgaactgttattcattgatgcaatccattcaaaccaacaaatgccatgt3SS::Yal lineline_8-891ggccatcaagacgtgccaaggcctcccagccaccaaccaattgatgaactgttattcattgatgcaatccattcaaaccaacaaatgccatgt3SS::Yal lineline_9-992ggccatcaagacgtgccaaggcctcccagccaccaaccaattgatgaactgttattcattgatgcaatccattcaaaccaaatgccatgt3ATG: mutlineline_1-142ggccatcaagacgtgccaaggcctcccagccaccaaccaattgatgaactgttattcatgatgcaatccattcaaaccaaaatgccatgt3ATG: mutlineline_1-142ggccatcaagacgtgccaaggcctcccagccaccaaccaattgatgaactgttattcatgatgcaatccattcaaaccaaatgccatgt3ATG: mutlineline_1-142ggccatcaagacgtgccaaggcctcccagccaccaaccaattgatgaactgttattcatgatgcaatccattcaaaccaaatgccatgt3ATG: mutlineline_1-142ggccatcaagacgtgccaaggcctcccagccaccaaccaattgatgaactgttattcatgatgtcaatccattcaaaccaaatgccatgt3ATG: mutline </td <td></td> <td>line 9-9#2</td> <td>tttetettttaaacetaaaa aaccaaccaacaaatcaaaa atgggegetgaaaaggtate catgaacategttgtegte ggacaagtecaatetgecaa</td>		line 9-9#2	tttetettttaaacetaaaa aaccaaccaacaaatcaaaa atgggegetgaaaaggtate catgaacategttgtegte ggacaagtecaatetgecaa
35::Yal atterline 3-342ttetettttaaneetaaa attetettttaaneetaaa aaccaaccaacaatcaaa taggegetgaaaggtate etagaacatgttgege ggacagtecaategecaa taggegetgaaaggtate etagacatgttgege ggacagtecaategecaa taggegetgaaaggtate etagacatgttgege ggacagtecaategecaa taggegetgaaaggtate etagacatgttgege ggacagtecaategecaa taggegetgaaaggtate etagacatgttgege ggacagtecaategeca aaccaacaagacgtegea aggeeteccageecaace aggeeteccageecaace taggeeteccageecaace taggeetgaaaggtate etagacatgttatt tagtgetgatatecatte aatgagteatecatte aatecaaaatgeetagt aatecaaaatgeetge aggeeteccageecaace aggeeteccageecaace aggeeteccageecaace caatgatgaaetgtaat tagtgagtegtate tagtgateatecatte aatecaaaatgeetagt aatecaaatgeetgea aggeeteccageecaeac caatgatgaaetgtaat tagtgategtaate tagtgateatecatte aatecaaaatgeetagt aatecaaatgeetagt aatecaaatgeetagt aatecaaatgeetagt aatecaaatgeetagt aatecaaatgeetgea aggeeteccageeceae caatgagaetgtaat tagtgatgaaetgtaat tagtgateatecatte aatecaaatgeetagt aatecaaatgeetagt aatecaaatgeetage aggeeteccageeceaee caatgagaetgtaat tagtgaetgtaate tagtgateatecatte aatecaaatgeetgea aggeeteccageeceaee caatgagaetgtagaatgtate tagtgaetgtaate tagtgaetgtaateetgtaat tagtgaetgtaateetgtaateetgtaa aatecaaatgeetgea aggeeteccageeceaee caatgagaetgtagaaetgtaat caatgagetgtaateetgtaateetgtaa aatecaaatgeetgea aggeeteccageeceaee caatgagaetgtagaetgtaateetgtaateetgtae aatgagaetgtateetgtagaetgtaateetgtae aatecaaaatgeetgtag aatecaaatgeetgee aggeetec		line_1-1#1	tttetettttaaacetaaaa aaccaaccaaatcaaaa tagggegetgaaaaggtate etagaacategttgtegte ggacaagtecaatetgecaa
ATGs_muTimesline_1-332tttetettttaaacetaaatacggggegtgaaaggtatectagaacategttgege ggacaagtecaatetegcaa tagggegetgaaaggtateYal reference sequencesYal_273_ntggccatcaagacegtggcca aggcctccaagacegtggcca aggcctcccagccaccaaccattgatgaactgttaattcatgtagtcaatecattca aaccaaccacaaatecaaa38S::Yal imesline_4-442ggccatcaagacegtggcca aggcctcccaagccaccaaccattgatgaactgttaattcatgtagtcaatecattca aactcaaatagtccatagt38S::Yal imesline_8-842 ine_9-991ggccatcaagacegtggcca aggcctcccaagccaccaaccattgatgaactgttaattcatgtagtcaatecattca aactcaaaatgccatagt3ATGs_mutimesline_1-141 ine_3-992ggccatcaagacegtggcca aggcctccccagccaccaaccattgatgaactgttaattcatgtagtcaatecattca aactcaaaatgccatagt3ATGs_mutimesline_1-141 ine_3-992ggccatcaagacegtggcca aggcctccccagccaccaaccattgatgaactgttaattcatgtagtcaatecattca aactcaaaatgccatagt3ATGs_mutimesline_1-141 ine_3-932 ingccatcaagacegtggcca aggcctcccaagcccaccaccattgatgaactgttaattcatgtagtcaatccattca aactcaaaatgccatagt aggcctcccaagcccaccac3ATGs_mutimesline_3-932 ingccatcaagacegtggcca aggcctcccaagcccaccaccattgagactgttaattcatgtagtcaatccattca aactcaaaatgccatagt aggcctcccaagcccaccacYal reference sequencesreference ine_9-991caacaatacaataatgr aactcaaatatgr aggcctcccaagcccaccaccaattgatgaactgttaattcatgtagtcaatccattca aactcaaaatgccatagt aggcctcccagcccaccacYal reference sequencesreference sequencescaacaatacaatatgr aactcaaatatgr aactcaatatatgr acctcaatatcagtatgcaacatcagtacgtgcca aggcctcc		line_1-1#2	tttetettttaaacetaaaa aaccaaccaaatcaaaa tagggcgctgaaaaggtate ctagaacatcgttgtcgte ggacaagtecaatetgecaa
line_4-441tttcttttaaactaaa accaacaaataaa taggggctgaaaggtat ctagacatgttgtc ggacagtccattgcaa taggggctgaaaggtat ctagacatgttgtc ggacagtcgatattgcaat taggggctgaaaggtat ctagacatgttgtc ggacagtccattgcaa taggggctgaaaggtat ctagacatgttgtc ggacagtcgatattgcaat taggggctgaaaggtat ctagacatgttgtcat cagtagtcaatcgtta aattgagacatgttat cagtagtcaatcatta aattcaaaatgccatagt line_9-94233TGs_mutlineline_8-841 ggccatcagacgtgcca ggccatcagacgtgca ggccatcagacgtgca aggcctccagccacaca aggcctccagccacaca castgatgaacgttatt castgatgaacgttatt catgtagtcatcattta aattcaaaatgccatagt line_9-942ggccatcagacgtgca ggccatcagacgtgca aggcctccagccacaca ggccatcagacgtgca aggcctccagccacaca castgatgaacgttatt castgatgaacgttatt catgtagtcatcattta aattcaaaatgccatagt line_3-942ggccatcagacgtgca ggccatcagacgtgca aggcctccagccacaca ggcctccagccacaca castgatgaacgttatt castgatgaatgttatt catgtagtcatcattta aattcaaaatgccatagt line_3-342 ggccatcagacgtgca ggccatcagacgtgca aggcctccaggccacaac ggccatcagacgtgca aggcctcccagccacaca aggcctcccagccacaca castgatgaacgttatt castgatgaatgttatt catgtagtcatcattta aattcaaaatgccatagt line_4-442Yal referenc xequencesYal referenc y line_8-841Yal referenc xequencesYal referenc y line_8-841Yal referenc xequencesYal referenc y line_8-841Yal referenc xequencesYal referenc y line_9-942Yal referenc xequencesYal referenc y line_9-941Aggcctcagacgtgcca caacaatacatatagt accaacaatacatatagt acccaacaatacatatgt acccaacaacatacatatgt acccaacaacatacatatgt acccaacaatacatatgt acccaacaatacatatgt acccaacaatacatatgt acccaacaacatacatatagt ccacegcaaataatcagt cacatcegatgtac ccacegcaa	35S::Ya1_	line_3-3#1	tttetettttaaaeetaaaa aaccaaccaaatcaaaa tagggegetgaaaaggtate etagaacategttgtegte ggacaagtecaatetgecaa
line_4-442tttetettttaaaectaaa aaccaacaaataaa tagggeetgaaaaggtate etgaacatgttgtegte ggacaagtecaatetgeaaYal refereer sequencesYal_273_ntggcatcaagacegtegea aggeeteceagecacaaa caatgatgaactgttaatt catgtagteaatecattea aacteaaaatgeeatagt line_8-841 ggcatcaagacegtegea aggeeteceagecacaaa caatgatgaactgttaatt catgtagteaatecattea aacteaaaatgeeatagt line_9-942355::Yal linesline_8-842 ggcatcaagacegtegea aggeeteceeagecacaa aggeeteceagecacaaa caatgatgaactgttaatt catgtagteaatecattea aacteaaaatgeeatagt line_9-9423ATGs_mut lineline_8-841 ggcatcaagacegtegea aggeeteceeagecacaa aggeeteceagecacaaa caatgatgaactgttaatt catgtagteaatecattea aacteaaaatgeeatagt line_1-142 ggceatcaagacegtegea aggeeteceeagecacaa ggeeteceagecacaaa caatgatgaactgttaatt catgtagteaatecattea aacteaaaatgeeatagt line_3-341 ggeeatcaagacegtegea aggeeteceeagecacaa ggeeatecaagacegtegea aggeeteceeagecacaa caatgtagaactgttaatt catgtagteaatecattea aacteaaaatgeeatagt line_3-342 ggeeatecaagacegtegea aggeeteceeagecacaa ggeeatecaagacegtegea aggeeteceeagecacaa caatgtagaactgttaatt catgtagteaateeattea aacteaaaatgeeatagt line_4-442 ggeeatecaagacegtegea aggeeteceeagecacaa ggeeteceageeacaaa caatgtagaactgttaatt catgtagteaateeattea aacteaaaatgeeatagt line_4-442 ggeeatecaagacegtegea aggeeteceeagecacaa caatgtagaactgttaatt catgtagteateeattea aacteaaaaatgeeatagt caaecaataeatatagt acteataatagtatg caaecaataeatatagt acteatatategtat caaecaateaatategt acteatatategtate caaecaateaatategt acteatatategtateg caaecaateaatategt acteatatategtateg caaecaateaatatagt acteatatategtateg caaecaateaatategt acteatatategtateg caaecaateaatategt acteatatategtateg caaecaateaatategt acteatatategtateg caaecaateaatategt acteatatategtateg caaecaateaatategt acteatatategtateg caaecaateaatategt acteatatategtateg caaecaateaatategt actea	3ATGs_mut lines	line_3-3#2	tttetettttaaaeetaaaa aaccaaccaaatcaaaa tagggegetgaaaaggtate etagaacategttgtegte ggacaagtecaatetgecaa
Yal reference sequencesYal_273_ntggccatcaagaccgtcgcca aggcctccaagaccgtcgcca aggcctcccagccaccacccaattgatgaactgttaatt caattgatgaactgttaatt catgtagtcaatccatttca caattgatgaactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagt inne_9-9#2ggccatcaagaccgtcgcca aggcctcccagccaccacc aggcctcccagccaccacc caattgatgaactgttaatt caattgatgaactgttaatt catgtagtcaatccatttca aactcaaaatgccatagt gtgcatcaagaccgtcgcca aggcctcccagccaccacc aggcctcccagccaccacc caattgatgaactgttaatt catgtagtcaatccatttca aactcaaaatgccatagt line_9-9#2ggccatcaagaccgtcgca aggcctcccagccaccacc aggcctcccagccaccacc caattgatgaactgttaatt catgtagtcaatccatttca aactcaaaatgccatagt line_3-9#1 ggccatcaagaccgtcgcca aggcctcccagccaccacc aggcctcccagccaccacc aggcctcccagccaccacc caattgatgaactgttaatt catgtagtcaatccatttca aactcaaaatgccatagt line_3-3#1 ggccatcaagaccgtcgcca aggcctcccagccaccacc aggcctcccagccaccacc aggcctcccagccaccacc caattgatgaactgttaatt catgtagtcaatccatttca aactcaaaatgccatagt line_3-3#2 ggccatcaagaccgtcgcca aggcctcccagccaccacc aggcctcccagccaccacc aggcctcccagccaccacc caattgatgaactgttaatt catgtagtcaatccatttca aactcaaaatgccatagt line_4-4#2 ggccatcaagaccgtcgcca aggcctccagccaccacc aggcctccagcaccaccacc caattgatgaactgttaatt caatgtagtcaatccatttca aactcaaaatgccatagt gccatcaagaccgtcgcca aggcctccagccaccaccaccaccaccaccaccaccaccaccacca		line_4-4#1	<u>tttetettttaaacet</u> aaaa aaccaaccaacaaatcaaaa tagggegetgaaaaggtate etagaacategttgtegte ggacaagtecaatetgeeaa
sequencesYal_273_ntggccatcaagacgtgcocaaggcctcccagccaccacacaattgatgacgtgtaatcattgatgacatcattaatactcaaaatgccatagt35S::Yallinesline_8-841ggccatcaagacgtgcocaaggcctcccagccaccacacaattgatgacgtgtaatcattgatgacatcattaatactcaaaatgccatagt35S::Yallinesline_9-941ggccatcaagacgtgcocaaggcctcccagccaccacacaattgatgacgtgtaatcatgtatgacatcattaatactcaaaatgccatagt34TGs_mutlineline_1-141ggccatcaagacgtgcocaaggcctcccagccaccacacaattgatgacgttaatcatgtatgacatcattatactcaaaatgccatagt34TGs_mutlineline_3-341ggccatcaagacgtgcocaaggcctcccagccaccacacaattgatgaactgttaatcatgtatgcaatccatttaactcaaaatgccatagt11ne_4-442ggccatcaagacgtgcocaaggcctccccagccaccacacaattgatgaactgttaatcatgtatgcaatccatttaactcaaaatgccatagt11ne_4-442ggccatcaagacgtgcocaaggcctccccagccaccacacaattgatgaactgttaatcatgtatgcaatccatttaactcaaaatgccatagt11ne_4-442ggccatcaagacgtgcocaaggcctccccagccaccacacaattgatgaactgttaatcatgtatgcaatccatttaactcaaaatgccatagt11ne_8-841ggccatcaagacgtgcocaaggcctccccagccaccacacaattgatgaactgttaatcatgtatgcaatccatttaactcaaaatgccatagt355::Yallineline_8-841caaccaatacatatagtactcaatatcagtatgcaacgccaatagt aggcatcagatgtcaattgatgaatgtatacatgtatgcatcgattaa366sequencesYal_273_ntcaaccaatacatatagtactcaatatagtgtcaacgcaatagtagtcaatgcatggtaga3765_mutlineline_9-942caaccaatacatatagtactcaatatagtgt		line_4-4#2	tttetettttaaacetaaaa aaccaaccaaatcaaaa tagggcgctgaaaaggtate etagaacategttgtegte ggacaagtecaatetgecaa
sequencesYal_273_ntggccatcaagacgtgcocaaggcctcccagccaccacacaattgatgacgtgtaatcattgatgacatcattaatactcaaaatgccatagt35S::Yallinesline_8-841ggccatcaagacgtgcocaaggcctcccagccaccacacaattgatgacgtgtaatcattgatgacatcattaatactcaaaatgccatagt35S::Yallinesline_9-941ggccatcaagacgtgcocaaggcctcccagccaccacacaattgatgacgtgtaatcatgtatgacatcattaatactcaaaatgccatagt34TGs_mutlineline_1-141ggccatcaagacgtgcocaaggcctcccagccaccacacaattgatgacgttaatcatgtatgacatcattatactcaaaatgccatagt34TGs_mutlineline_3-341ggccatcaagacgtgcocaaggcctcccagccaccacacaattgatgaactgttaatcatgtatgcaatccatttaactcaaaatgccatagt11ne_4-442ggccatcaagacgtgcocaaggcctccccagccaccacacaattgatgaactgttaatcatgtatgcaatccatttaactcaaaatgccatagt11ne_4-442ggccatcaagacgtgcocaaggcctccccagccaccacacaattgatgaactgttaatcatgtatgcaatccatttaactcaaaatgccatagt11ne_4-442ggccatcaagacgtgcocaaggcctccccagccaccacacaattgatgaactgttaatcatgtatgcaatccatttaactcaaaatgccatagt11ne_8-841ggccatcaagacgtgcocaaggcctccccagccaccacacaattgatgaactgttaatcatgtatgcaatccatttaactcaaaatgccatagt355::Yallineline_8-841caaccaatacatatagtactcaatatcagtatgcaacgccaatagt aggcatcagatgtcaattgatgaatgtatacatgtatgcatcgattaa366sequencesYal_273_ntcaaccaatacatatagtactcaatatagtgtcaacgcaatagtagtcaatgcatggtaga3765_mutlineline_9-942caaccaatacatatagtactcaatatagtgt			
Ine_8-8#1ggccatcaagaccgtcgccaaggcctcccagccacacacaattgatgaactgttaattcattgatgcaatcatttaaaactcaaaatgccatagt385::Yal linesline_9-8#2ggccatcaagaccgtcgccaaggcctcccagccacacacaattgatgaactgttaattcattgatgcaatccatttaaaactcaaaaatgccatagt3ATGs_mutlineline_1-1#1ggccatcaagaccgtcgccaaggcctcccagccacacacaattgatgaactgttaattcattgatgcaatcattaataactcaaaaatgccatagt3ATGs_mutlineline_1-1#2ggccatcaagaccgtcgccaaggcctcccagccacacacaattgatgaactgttaattcatgtatgcaatcattaaaactcaaaatgccatagt3ATGs_mutlineline_1-1#2ggccatcaagaccgtcgccaaggcctcccagccacacacaattgatggaactgttaattcatgtatgcaatcattaaaactcaaaatgccatagt3ATGs_mutlineline_1-1#2ggccatcaagaccgtcgccaaggcctcccagccacacacaattgatggaactgttaattcatgtatgcaatcattaaaactcaaaatgccatagt3ATGs_mutlineline_1-3#1ggccatcaagaccgtcgccaaggcctcccagccacacacaattgatggaactgttaattcatgtatgcaatcattaaaactcaaaaatgccatagt3ATGs_mutlinefine_8-8#1caaacaatacaataatgtaactcaaaatatgtaatattagtcaattgatggaactgttaattcatgtatgcaatcattaaaactcaaaaatgccatagt3ATGs_mutlinefine_8-8#1caaacaatacaataatgtaactcaaatatgtatattattagtattaatcatgtatgcaatcgattaattaattagtataatcaattagtaactcaaaatgccatagt3SS:Yal line_9-9#1caaacaatacaataatgtaactcaaatattagtattattattagtattagtccacgccgaaataaattagtccactccgatgtac3ATGs_mutlineline_1-1#1caaacaatacaataatgtaactcaaatatagtgtccacgccgaaataaatagtcacatcggatgtact3SS:Yal line_9-9#2 <td>Ya1 reference</td> <td></td> <td></td>	Ya1 reference		
35S::Yal line:line_8-842 line_9-941ggccatcaagaccgtcgcca aggccatcaagaccgtcgca aggccatccaagaccgtcgca aggccatccaagaccgtcgca aggccatccaagaccgtcgca aggccatccaagaccgtcgca aggccatccaagaccgtcgca aggccatccaagaccgtcgca aggccatccaagaccgtcgca aggccatccaagaccgtcgca aggccatccaagaccgtcgca aggccatccaagaccgtcgca aggccatccaagaccgtcgca aggccatccaagaccgtcgca aggccatccaagaccgtcgca aggcctcccagccacaac caattgtagaactgttatt caattgtagactgttatt catgtagtcaatccattta aattgaactgttatt catgtagtcaatccattta aattcaaaattgcatagt tatt catgtagtcaatccattta aattcaaaattgcatagt tatt catgtagtcaatccattta aattcaaaattgt actcaatattagt actcaatattagt actcaatattagt actcaatattagt ccaccgcaaataa ccaattgatgactgttatt ccaattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgatcatcaattat ccaccgcaaataa attgatgactgttatt caattgatgactgttatt caattgatgactgatcattagt caattgatgactgttatt caattgatgactgatcattagt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgactgttatt caattgatgatc	sequences		ggccatcaagaccgtcgcca aggcctccccagccaccaac caattgatgaactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagt
JaneImeggccatcaagaccgtcgccaggccatcaagaccgtcgccaggccatcaagaccgtcgacagaccatcgacgaccgaccgaccgaccaacgaccgacgaccgtcgacagaccgacgaccgacgaccgaccgacaagacgtcgacagaccgtcgacgaccg			ggccatcaagaccgtcgcca aggcctccccagccaccaac caattg <mark>atg</mark> aactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagtc
line_9-9#2ggccatcaagaccgtcgccaaggcctcccagccacacacaattgatgaactgttaattcattgatgcatcatttaaaactcaaaatgccatagt3ATGs_mutimeline_1-1#1ggccatcaagaccgtcgccaaggcctcccagccacacacaattgatgaactgttaattcatgtatgcaatcatttaaaactcaaaatgccatagt3ATGs_mutimeline_1-1#2ggccatcaagaccgtcgccaaggcctcccagccacacacaattgtagaactgttaattcatgtagtcaatccatttaaaactcaaaaatgccatagt1ine_3-31ggccatcaagaccgtcgccaaggcctcccagccacacacaattgtagaactgttaattcatgtagtcaatccatttaaaactcaaaaatgccatagt1ine_4-441ggccatcaagaccgtcgccaaggcctcccagccacacacaattgtagaactgttaattcatgtagtcaatccatttaaaactcaaaaatgccatagt1ine_4-442ggccatcaagaccgtgcgcaaggcctcccagccacacacaattgtagaactgttaattcatgtagtcaatccatttaaaactcaaaatgccatagt3SS:Yal line_line_8-841caaacaatacaataatgtaactcataatatggtagtagtccacgcaaataaatagtcacatccgatgtac3SS:Yal line_9-941caaacaatacaataatgtaactcataatatggtccacgcaaataaatagtaccatccgatgtac1ine_9-942caaacaatacaataatgtaactcataatatggtccacgcaaataaatagtaccatccgatgtac1ine_1-141caaacaatacaataatgtaactcataatatggtccacgcaaataaatagtaccatccgatgtac1ine_1-142caaacaatacaataatgtaactcataatatggtccacgcaaataaatagtacatccatagtagt3ATGs_mutimeline_1-141caaacaatacaataatgtacctatatatagtagtccacgcaaataaatagtacatccatagtagt3ATGs_mutimeline_1-141caaacaatacaataatgtacctatatatagtagtccacgccgaaataaatagtacatccagtgtc <td< td=""><td>35S::Ya1 lines</td><td></td><td>ggccatcaagaccgtcgcca aggcctccccagccaccaac caattg<mark>atg</mark>aactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagtc</td></td<>	35S::Ya1 lines		ggccatcaagaccgtcgcca aggcctccccagccaccaac caattg <mark>atg</mark> aactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagtc
Jangeline_1-1#1ggccatcaagaccgtcgccaaggcctcccagccaccaccaattgtagaactgttaattcatgtagtcatccatttcaaactcaaaatgccatagtJATGs_mullimeline_3-3#1ggccatcaagaccgtcgccaaggcctcccagccaccaccaattgtagaactgttaattcatgtagtcatccatttcaaactcaaaatgccatagtJATGs_mullimeline_3-3#2ggccatcaagaccgtcgccaaggcctcccagccaccaccaattgtagaactgttaattcatgtagtcaatccatttcaaactcaaaaatgccatagtJine_3-3#2ggccatcaagaccgtcgccaaggcctccccagccaccaccaattgtagaactgttaattcatgtagtcaatccatttcaaactcaaaaatgccatagtJine_4-4#1ggccatcaagaccgtcgccaaggcctccccagccaccaccaattgtagaactgttaattcatgtagtcaatccatttcaaactcaaaaatgccatagtggccatcaagaccgtcgccaaggcctccccagccaccaccaattgtagaactgttaattcatgtagtcaatccatttcaaactcaaaaatgccatagtggccatcaagaccgtcgccaaggcctccccagccaccaccaattgtagaactgttaattcatgtagtcaatccatttcaaactcaaaaatgccatagtggccatcaagaccgtcgccaaggcctccccagccaccaccaattgtagaactgttaattcatgtagtcaatccatttcaaactcaaaatgccatagtggccatcaagaccgtcgccaaggcctccccagccaccaccaattgtagaactgttaattcatgtagtcaatccatttcaaactcaaaatgccatagtggccatcaagaccgtcgccaaggcctccccagccaccaccaattgtagaactgttaattcatgtagtcatccatttcaaactcaaaatgccatagtggccatcaagaccgtcgccaaggcctcccagccaccaccaattgtagaactgttaattcatgtagtcatccatttcaaactcaaaatgccatagtggccatcaagaccgtcgccaaggcctcccagccaccaccaattgtagaactgttaattcatgtagtcatccatttcaaactcaaaatgcatgtagtggccatcagaccgtcgccaggccatcaagtagtgtcaccagtagtccacgc		line_9-9#1	ggccatcaagaccgtcgcca aggcctccccagccaccaac caattg <mark>atg</mark> aactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagtc
ATGs_mutlineline_1-1#2ggccatcaagaccgtcgccaaggcctcccagccaccaccaattgtagaactgttaattcatgtagtactactatttaaaactcaaaatgccatagt3ATGs_mutline1ine_3-3#1ggccatcaagaccgtcgccaaggcctcccagccaccaccaattgtagaactgttaattcatgtagtactaccatttaaaactcaaaaatgccatagt1ine_4-4#1ggccatcaagaccgtcgccaaggcctcccagccaccaccaattgtagaactgttaattcatgtagtactaccatttaaaactcaaaaatgccatagt1ine_4-4#1ggccatcaagaccgtcgccaaggcctcccagccaccaccaattgtagaactgttaattcatgtagtactaccatttaaaactcaaaaatgccatagt1ine_4-4#1ggccatcaagaccgtcgccaaggcctcccagccaccaccaattgtagaactgttaattcatgtagtactaccatttaaaactcaaaaatgccatagt1ine_4-4#1ggccatcaagaccgtcgccaaggcctcccagccaccaccaattgtagaactgttaattcatgtagtcaatccatttaaaactcaaaaatgccatagtsequencesYal_273_ntcaaacaatacaataatgtaactcataattagtagtccaccgcaaaataaatagtccactgcagtgtac3SS:Yal line1ine_8-8#1caaacaatacaataatgtaactcataatagtagtccaccgcaaataaatagtccactgcagtgtac1ine_9-9#2caaacaatacaataatgtaactcataatagtagtccacdgcaaataaatagtccactgcagtgtac1ine_1-1#1caaacaatacaataatgtaactcataatatgtagtccacgcaaataaatagtccactggagtgac3ATGs_mutline1ine_3-3#1caaacaatacaataatgtaactcataatatgtagtccacgcaaataaatagtccactggagtgac1ine_3-3#2caaacaatacaataatgtaactcaatatatgtagtccacgcaaataaatagtccactggagtgac1ine_3-3#2caaacaatacaataatagtaactcaatatatgtagtccactggagtgacine1ine_3-4#1 </td <td></td> <td>line_9-9#2</td> <td>ggccatcaagaccgtcgcca aggcctccccagccaccaac caattg<mark>atg</mark>aactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagtc</td>		line_9-9#2	ggccatcaagaccgtcgcca aggcctccccagccaccaac caattg <mark>atg</mark> aactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagtc
3ATGs_mut line_3-341 ggccatcaagaccgtcgcca aggcctcccagccacaca cattgtagaactgttaatt catgtagtcaatcatttca aactcaaaatgccatagt 1ine_3-342 ggccatcaagaccgtcgcca aggcctccccagccacaca caattgtagaactgttaatt catgtagtcaatccatttca aactcaaaatgccatagt 1ine_4-441 ggccatcaagaccgtcgcca aggcctcccagccacaca caattgtagaactgttaatt catgtagtcatccatttca aactcaaaaatgccatagt Yal reference Yal_273_nt caaacaatacaataatgt aactcataatatgg cactgcagt cacagaccgtcgcca cacagcagtagtagt cacagtcgatgtagt 3SS:Yal line_8-841 caaacaatacaataatgt aactcataatatgg tagtagtagt cacatcgatgtagt cacaatcgatgtagt 1ine_8-842 caaacaatacaataatgt aactcataatatgg cacagcgagt cacagacgtgcgc caacgacagtgcgc 3SS:Yal line_8 1ine_9-941 caaacaatacaataatgt aactcataatatgg tagtcattagt cacatcggatgtac 1ine_9-942 caaacaatacaataatgt aactcataatatgg cacaggcaaataaatagg cacagggga 3ATGs_mut line_1-141 caaacaatacaataatgt aactcataatatgg cacaggcaaataaatagg cacaggatgtac 3ATGs_mut line_3-341 caaacaatacaataatgt aactcataatatggt cacagcaaataaataggt cacatcgatgtac 1ine_3-342 caaacaatacaataatgt aactcataatatggt cacagcaaataaataggt cacatcgatgtac 3ATGs_mut line_4-442 caaacaatacaataatgt aactcataatatggt cacagcaaataaataggt cacatcgatgtac 1ine_1-141 caaacaatacaataatagt aactcataataggt cacagcaaataaatagt cacatcgatgtac 1ine_3-342 caaacaa		line 1-1#1	ggccatcaagaccgtcgcca aggcctccccagccaccaac caattgtagaactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagtc
Site June 3-342 ggccatcaagaccgtcgcca aggcctcccagccaccac cattgtagaactgttaatt catgtagtcaatcatttaa aactcaaaatgccatagt Jine 4-441 ggccatcaagaccgtcgcca aggcctccccagccaccac cattgtagaactgttaatt catgtagtcaatcatttaa aactcaaaaatgccatagt Yal reference ggccatcaagaccgtcgcca aggcctccccagccaccac cattgtagaactgttaatt catgtagtcaatccatttaa aactcaaaaatgccatagt SS::Yal lines Yal 273_nt caaacaatacaataatgt aactcataatatogtattg ccaccgcaaataa_atcagt cacatccgatgtac Jine 8-841 caaacaatacaataatgt aactcataatatogtattg ccaccgcaaataa_atcagt cacatccgatgtac Jine 9-941 caaacaatacaataatgt aactcataatatogtattg ccaccgcaaataa_atcagt cacatccgatgtac Jine 9-942 caaacaatacaataatgt aactcataatatggtattg ccaccgcaaataa_atcagt cacatccgatgtac Jine 1-141 caaacaatacaataatgt aactcataatatogtattg ccaccgcaaataa_atcagt cacatccgatgtac Jine 1-142 caaacaatacaataatgt aactcataatatogtattg ccaccgcaaataa_atcagt cacatccgatgtac Jine 1-141 caaacaatacaataatgt aactcataatatogtattg ccaccgcaaataa_atcagt cacatccgatgtac Jine 3-341 caaacaatacaataatgt aactcaataatatgt ccaccgcaaataaatagt cacatcgatgtac Jine 3-342 caaacaatacaataatgt aactcaataatatgt ccaccgcaaataaatagt cacatcgatgtac Jine 3-343 caaacaatacaataatgt aactcaataatatgt ccaccgcaaataaatagt cacatcgatgtac Jine 3-342 caaacaatacaataatagt aactcaataatatgt ccaccgcaaataaatagt cacatcgatgtac Jine 3-342 caaacaatacaa		line_1-1#2	ggccatcaagaccgtcgcca $aggcctccccagccaccaac$ $caattgtagaactgttaatt$ $catgtagtcaatccatttca$ $aactcaaaaatgccatagtcaatgtcaatgtcaatccatttca$
Ine_3-342 ggccatcaagaccgtcgcca aggcctccccagccacaca caattgtagaactgttaat catgtagtcaatccatttca aactcaaaatgccatagt Ine_4-441 ggccatcaagaccgtcgcca aggcctccccagccacaca cattgtagaactgttaat catgtagtcaatccatttca aactcaaaatgccatagt Jine_4-442 ggccatcaagaccgtcgcca aggcctccccagccacaca cattgtagaactgttaat catgtagtcaatccatttca aactcaaaatgccatagt Yal reference sequences Yal reference saccaatacaataatgt aactcataatatggt acctatattcagtatt Ine_8-841 caaacaatacaataatgt aactcataatatggt cacagccaaataaatagt cacatccgatgtac Ine_9-941 caaacaatacaataatgt aactcataatatggt acctatatatagt cacatccgatgtac Ine_9-942 caaacaatacaataatgt aactcataatatggt cacagcaaataaatagt cacatccgatgtac Ine_1-141 caaacaatacaataatgt aactcataatatggtag caaacaatacaataatgt aactcataatatgg cacaggcaaataaatagt cacatccgatgtac ine_1-141 SSS:Yal line_1-141 caaacaatacaataatgt aactcataatatgg cacagcaaataaatagt cacatccgatgtac Ine_1-141 caaacaatacaataatgt aactcataatatgt cacagcaaataaatagt cacatccgatgtac Ine_3-341 caaacaatacaataatgt aactcataatatgatg cacagcaaataaatagt cacatcgatgtac Ine_3-342 caaacaatacaataatgt aactcataatatgatg cacagcaaataaatagt cacatcgatgtac Ine_3-342 caaacaatacaataatgt aactcataatatgatg cacagcaaataaatagt cacatcgatgtac Ine_3-342 caaacaatacaataatgt aactcataatatggtag Cacagcaa	3ATGs mut lines	line_3-3#1	ggccatcaagaccgtcgcca aggcctccccagccaccaac caattg <mark>tag</mark> aactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagtc
line_4-442 ggccatcaagaccgtcgcca aggcctccccagccaccac caattgtagaactgttaatt catgtagtcaatcatttca aactcaaaatgccatagt Yal referenc sequences Yal_273_nt caaacaatcaataatgt aactcataatatgt acccatactagtatt cacatccgatgtac line_8-841 caaacaatcaataatgt aactcataatatgt gccacgcaaataa <u>atcagt cacatccgatgtac</u> 35S::Yal lines line_9-941 caaacaatcaataatgt aactcataatatggt gccacgcaaataa <u>atcagt cacatccgatgtac</u> line_1-9-942 caaacaatcaataatgt aactcataatatggt gccacgcaaataa <u>atcagt cacatccgatgtac</u> line_1-141 caaacaatcaataatgt aactcataatatggt gccacgcaaataa <u>atcagt cacatccgatgtac</u> line_3-341 caaacaatcaataatgt aactcataatatggt cacagcaaataa <u>atcagt cacatccgatgtac</u> line_3-342 caaacaatcaataatgt aactcataatatggt cacagcaaataa <u>atcagt cacatccgatgtac</u> line_3-342 caaacaatcaataatgt aactcataatatgg cacagcaaataa <u>atcgt cacatccgatgtac</u> line_3-342 caaacaatcaataatgt aactcataatatgg cacagcaaataa <u>atcgt cacatccgatgtac</u> line_3-342 caaacaatcaataatgt aactcataatatgg cacagcaaataa <u>atcgt cacatccgatgtac</u> line_4-441 caaacaatcaataatgt aactcataatatgg cacagcaaataa <u>atagt cacatccgatgtac</u> line_4-442 caaacaatcaatatagt aactcataatatggt cacagcaaataa <u>atcgt cacatccgatgtac</u>	-	line_3-3#2	ggccatcaagaccgtcgcca aggcctccccagccaccaac caattg <mark>tag</mark> aactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagtc
Yal reference sequences Yal 273 nt caacaatacaatatatgt aactcataatatogt ccacogcaaataa <u>atcaqt cacatcoqatgtac</u> 1ine_8-8#1 caaccaatacaatatatgt aactcataatatogtattg ccacogcaaataa <u>atcaqt cacatcoqatgtac</u> 35S::Yal lines line_9-9#1 caaccaatacaatatatgt aactcataatatogtatg ccacogcaaataa <u>atcaqt cacatcoqatgtac</u> ine_9-9#2 caaccaatacaatatatgt aactcataatatogtattg ccacogcaaataa <u>atcaqt cacatcoqatgtac</u> line_1-1#1 caaccaatacaatatatgt aactcataatatogtattg ccacogcaaataa <u>atcaqt cacatcoqatgtac</u> line_1-1#2 caaccaatacaatatatgt aactcataatatogtattg ccacogcaaataa <u>atcaqt cacatcoqatgtac</u> line_3-3#1 caaccaatacaataatgt aactcataatatogtattg ccacogcaaataa <u>atcqt cacatcoqatgtac</u> line_3-3#2 caaaccaatacaataatgt aactcataatatogtattg ccacogcaaataa <u>atcqt cacatcoqatgtac</u> line_4-4#2 caaaccaatacaataatgt aactcataatatogtattg ccacogcaaataa <u>atcqt cacatcoqatgtac</u>		line_4-4#1	ggccatcaagaccgtcgcca aggcctccccagccaccaac caattgtagaactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagtc
sequencesYa1_273_htcaacaatacaatatgt aactaatatogt aactaatatogtattg cacacqgatataatacagt cacatccqatgtacass:Ya1ineline_8-841caacaatacaatatgt aactaatatogt aactaatatogtatgcacacqataatatgt cacatccqatgtacass:Ya1ineline_8-842caacaatacaatatgt aactaatatogtatgt cacagtcgatataatagt cacatccqatgtacine_9-941caacaatacaatatgt aactaatatogtatgccacqgaaataatagt cacatccqatgtacine_9-942caacaatacaatatgt aactaatatagtatgccacqgaaataatagt cacatccqatgtacine_1-142caacaatacaatatgt aactaatatagt gccacqgaaataatgt cacatccqatgtacine_3-341caacaatacaatatgt aactaatatagt gccacqgaaataattgt cacatccqatgtacine_3-342caacaatacaatatgt aactaatatagt aactaatatagt gccacqgaaataattgt cacatccqatgtacine_4-442caacaatacaatatgt aactaatatagt aactaatatagt cacatcqatgtacccacqgaaataattgt cacatcqatgtacine_4-442caacaatacaatatgt aactaatatagt aactaatatagt gccacqgaaataattgt cacatcqatgtacine_4-442caacaatacaatatgt aactaatatcgatgtagtccacqgaaataatagt cacatcqatgtacine_4-442caacaatacaatatgt aactaatatagt aactaatatcggtagtccacqgaaataatagt cacatcqatgtacine_4-442caacaatacaatatgt aactaatatcggtagtccacqgaaataatagt cacatcqatgtacine_4-442caacaatacaatatgt aactaatatagt aactaatatcggtagtccacqgaaataaatagt accatcqatggtacine_4-442caacaatacaatatgt aactaatatcggtagtccacqgaaataaatagt accatcqatggtacine_4-442caacaatacaatatgt aactaatatagt aactaatatcggtagtccacggaaataaatggt accatcqagtgacine_4-442caacaatacaatatgt aactaatatcggtagtccacgcaaataaataggt accatcqagtgacine_4-442caacaataca		line_4-4#2	ggccatcaagaccgtcgcca aggcctccccagccaccaac caattgtagaactgttaatt catgtagtcaatccatttca aactcaaaaatgccatagtc
sequencesYa1_273_htcaacaatacaatatgt aactaatatogt aactaatatogtattg cacacqgatataatacagt cacatccqatgtacass:Ya1ineline_8-841caacaatacaatatgt aactaatatogt aactaatatogtatgcacacqataatatgt cacatccqatgtacass:Ya1ineline_8-842caacaatacaatatgt aactaatatogtatgt cacagtcgatataatagt cacatccqatgtacine_9-941caacaatacaatatgt aactaatatogtatgccacqgaaataatagt cacatccqatgtacine_9-942caacaatacaatatgt aactaatatagtatgccacqgaaataatagt cacatccqatgtacine_1-142caacaatacaatatgt aactaatatagt gccacqgaaataatgt cacatccqatgtacine_3-341caacaatacaatatgt aactaatatagt gccacqgaaataattgt cacatccqatgtacine_3-342caacaatacaatatgt aactaatatagt aactaatatagt gccacqgaaataattgt cacatccqatgtacine_4-442caacaatacaatatgt aactaatatagt aactaatatagt cacatcqatgtacccacqgaaataattgt cacatcqatgtacine_4-442caacaatacaatatgt aactaatatagt aactaatatagt gccacqgaaataattgt cacatcqatgtacine_4-442caacaatacaatatgt aactaatatcgatgtagtccacqgaaataatagt cacatcqatgtacine_4-442caacaatacaatatgt aactaatatagt aactaatatcggtagtccacqgaaataatagt cacatcqatgtacine_4-442caacaatacaatatgt aactaatatcggtagtccacqgaaataatagt cacatcqatgtacine_4-442caacaatacaatatgt aactaatatagt aactaatatcggtagtccacqgaaataaatagt accatcqatggtacine_4-442caacaatacaatatgt aactaatatcggtagtccacqgaaataaatagt accatcqatggtacine_4-442caacaatacaatatgt aactaatatagt aactaatatcggtagtccacggaaataaatggt accatcqagtgacine_4-442caacaatacaatatgt aactaatatcggtagtccacgcaaataaataggt accatcqagtgacine_4-442caacaataca			
1ine 35S:Yalline1ine 6-841caacaatacaataatgt aactcataatacgt actcataatacgt g cacacgaaaataa actcatgt cacatccaqtqtac caacacatacaatatgt aactcataatactagt g cacacgaaaataa t cacaccaataatgt g cacacgaaataa t cacaccaataatgt g cacacgaaataa 		Val 273 nt	
35S::Yal lineline_8-8#2 ine_9-9#1caacaatacaatadtgt aactcataatacgtatgt cacaccgataataa caacgacaataaatagt aactcataatacgtatgt cacaccgataataatagt caacgacaataaatagt aactcataatacgtatgt caacgacaataaatagt cacatcgatgtac1ine_9-9#2caacaatacaataatgt aactcataatacgtatg caacgacaataatagt aactcataatacgtatgt cacagtcgatgtac caacgacaataagtgt cacatcgatgtac1ine_1-1#1caacaatacaataatgt aactcataatacgtatg caacgacaataaattgt aactcataatacgtatg cacaggacaataaattgt cacatcgatgtac3ATGs_mutline1ine_1-1#2 ine_3-3#1caacaatacaataatgt aactcataatacgattg cacagcaaataaattgt cacatcgatgtac caacgacaataaattgt aactcataatacgattg cacaggacaataaattgt cacatcgatgtac cacaggacaataaattgt cacatcgatgtac cacaggacaataaattgt cacatcgatgtac caacgacaataattgt aactcataatatgt g cacaggacaataaattgt cacatcgatgtac cacaggacaataaattgt cacatcgatgtac cacaggacaataaattgt cacatcgatgtac cacaggacaataaattgt cacatcgatgtac cacaggacaataaatagt cacatcgatgtac cacaggacaataaattgt cacatcgatgtac cacaggacaataaatgt cacatcgatgtac cacaggacaataaatgt cacatcgatgtac cacaggacaataaatagt cacatcgatgtac cacaggacaaataatagt cacatcgatgtac cacaggacaaataaatagt cacatcgatgtac cacaggacaaataaatagt cacatcgatgtac cacaggacaaataaatagt cacatcgatgtac cacaggacaaataatgt cacatcgatgtac cacaggacaaataaatagt cacatcgatgtac cacaggacaaataaatagt cacatcg	sequences		
line_9-9#1 caacaatacaataataf a aactaatacagtatf accacagtata accagt acagt cacatcegatgtac line_9-9#2 caacaatacaataataf aactaatatagtaf acctaatatacgattf accacagtaataatagt accatcegatgtac line_1-1#1 caacaatacaataatgt aactaatatagtaataatatagtagt accacagaaataaatagt accatcegatgtac JATGs_mut line_3-3#1 caacaatacaataatgt aactaatatagtagtagt accatacagtagt accatcegatgtagt accatcegatgtagt aactaatatagt aactaatatagt accataatatagt accatcagatagt accatcegatgtagt accatcegatgtagt accatcegatgtagt accatcegatgtagt accatcegatgtagt accatcegatgtagt accatcegatgtagt accatcegatgtagt accatcegatgtagt accatcegatatagt accatcegatgtagt accatcegatgtagtagtagtagtagtagtagtagtagtagtagtag			
line_9-9#2 caacaatacaatatgt aactcataatacgt ccacgcaaaataa <u>atcagt cacatccgatgtac</u> line_1-1#1 caacaatacaatatgt aactcataatatgt gcacgatgtgt ccacgcaaaataa <u>atcagt cacatccgatgtac</u> line_1-1#2 caacaatacaatatgt aactcataatatggtagt ccacgcaaaataa <u>atcagt cacatccgatgtac</u> ine_3-3#1 caacaatacaataatgt aactcataatatggtagt ccacgcaaaataa <u>atcgt cacatccgatgtac</u> line_3-3#2 caacaatacaataatgt aactcataatatggtagt ccacgcaaataa <u>atcgt cacatccgatgtac</u> line_4-4#1 caacaatacaataatgt aactcataatatcggtagt caacgatacaataatgt aactcataatatggtagt ccacgcaaataa <u>atcgt cacatccgatgtac</u> line_4-4#2 caacaatacaataatgt aactcataatatcggtagt cacgcgaaaataa <u>atcgt cacatccgatgtac</u> caacgaatacaataatgt aactcataatatggtagt cacatgtagtg cacacgaatacaataatgt aactcataatatcggtagt cacacgcgaaataa <u>atcgt cacatccgatgtac</u> line_4-4#2 caacaatacaataatgt aactcataatatcggtagt	358::Yal lines		
Jine_1-1#1 caaacaataaatatgt aactoataatacqgtatg ccaccgcaaaataa <u>atcqt cacatccgatgtac</u> JATGs_mutline line_1-1#2 caaacaatacaataatgt aactoataatacqgtatg ccaccgcaaaataa <u>atcqt cacatccgatgtac</u> JATGs_mutline line_3-3#1 caaacaatacaataatgt aactoataatatggtatg ccaccgcaaaataa <u>atcqt cacatccgatgtac</u> line_3-3#2 caaacaatacaataatgt aactoataatatggtaatg ccaccgcaaaataa <u>atcqt cacatccgatgtac</u> line_4-4#1 caaacaatacaataatgt aactoataatatcggtatg ccaccgcaaaataa <u>atcqt cacatccgatgtac</u> line_4-4#2 caaacaatacaataatgt aactoataatatcggtatg ccaccgcaaaataa <u>atcqt cacatccgatgtac</u>		—	
JATGs_mutline_1-1#2 caaacaataaatatgt aactcataatacgtatg ccaccgcaaaataa <u>atcqt cacatccgatgtac</u> JATGs_mutline_1-1#2 caaacaatacaataatgt aactcataatacgtatg ccaccgcaaaataa <u>atcqt cacatccgatgtac</u> JATGs_mutline_3-3#1 caaacaatacaataatgt aactcataatatgt g ccaccgcaaaataa <u>atcqt cacatccgatgtac</u> line_3-3#2 caaacaatacaataatgt aactcataatatcgatgtg ccaccgcaaaataa <u>atcqt cacatccgatgtac</u> line_4-4#1 caaacaatacaataatgt aactcataatatcggtatg ccaccgcaaaataa <u>atcagt cacatccgatgtac</u> line_4-4#2 caaacaatacaataatgt aactcataatacggtatg		-	
3ATGs_mutline 1ine_3-3#1 caaacaataaatatgt aactcataatacgtattg ccaccgcaaaataa <u>atctgt ccactccqatgtac</u> 1ine_3-3#2 caaacaatacaataatgt aactcataatattgtattg ccaccgcaaaataa <u>atctgt ccactccqatgtac</u> 1ine_4-4#1 caaacaatacaataatgt aactcataatatcgtagtagt ccaccgcaaaataa <u>atcgt ccactccqatgtac</u> 1ine_4-4#2 caaacaatacaataatgt aactcataatatcgtagt			
SAIGs_mutime line_3-3#2 caaacaataaatatgt aactcataatatcagtattg ccaacaataaatagt line_4-4#1 caaacaataaatatgt aactcataatatcagtattg ccaacagt ccaacaatacaataatatgt ccaacaatacaataatatgt ccaacaatacaataatatgt ccaacagt ccaacaatacaataatatgt ccaacaatacaataatatgt ccaacagt ccaacaatacaataatatgt ccaacaatacaataatatgt ccaacaatacaataatatgt ccaacagt ccaacag			
line_4-4#1 caaacaatacaataaatatgt aactcataatatcagtattg ccaccgcaaaataa atcagt cacaatccgtatgtac line_4-4#2 caaacaatacaataatatgt aactcataatatcagtattg ccaccgcaaaataa atcagt ccactccgatgtac	3ATGs_mut lines		
line_4-4#2 caaacaatacaataatatgt aactcataatatcagtattg ccaccgcaaaataa <u>atcagt cccatccgatgtac</u>			
Reverse primer		TTH6 4-4#2	
			Reverse primer

- 2 **Fig. S16.** Sequences of RT-PCR products of *Ya1* and *Ya1_3ATGs_mut*
- 3 transcripts amplified from transgenic *A. thaliana* plants (Figure 5D). Nucleotides
- 4 underlined lowercase are sequences of forward and reverse primer. ATG and
- 5 TAG (mutated ATG) are in red.

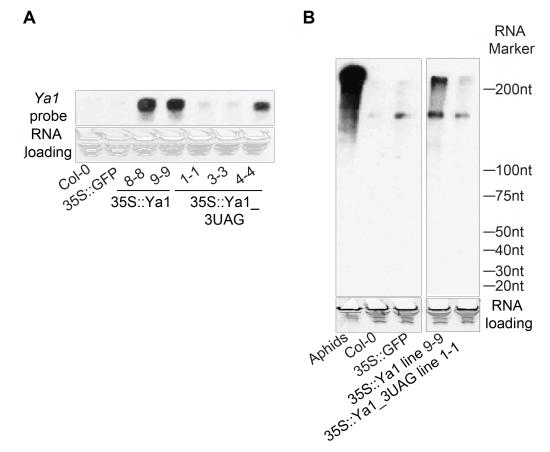


Fig. S17. Northern blot to confirm expression of Ya1 and Ya1 3UAG expressed in transgenic A. thaliana plants. (A) Northern blot hybridizations with a Ya1 probe to detect Ya1 and Ya1 3UAG expressed in transgenic A. thaliana plants. (B) Full length of Ya1 and Ya1 3UAG were detected in the transgenic plants. Prestain RNA markers (DynaMarker) were used to estimate the size of RNAs.

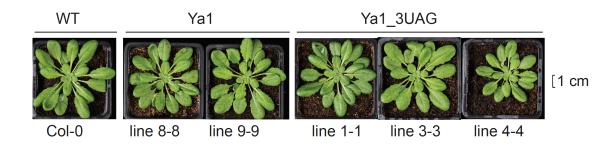


Fig. S18. Comparison of phenotypes of Ya1 transgenic A. thaliana plants.
Transgenic A. thaliana (Col-0) plants that express Ya1 and Ya1_3UAG under
control of the 35S promoter did not show obvious morphological differences
compared to the non-transformed wild type Col-0 plants.

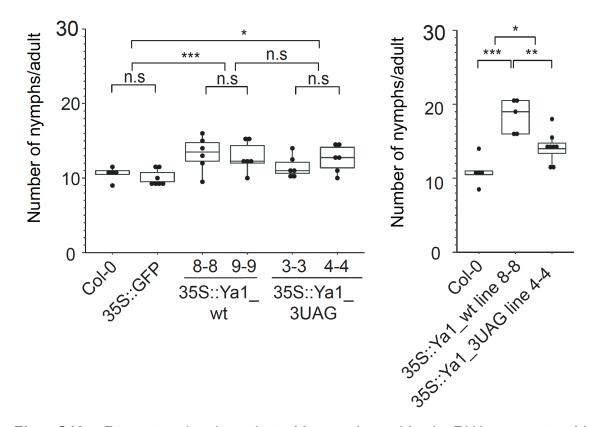


Fig. S19. Repeats showing that *M. persicae* Ya IncRNA promotes *M. persicae* colonization on *A. thaliana*. Stable expression of Ya1_wt and Ya1_3UAG promotes *M. persicae* reproduction on plants. Each data point (black dot) represents number of nymphs produced by an adult female aphid per plant. Box plots show distribution of data points collected from n = 5-8 female aphids per *A. thaliana* line. *p < 0.05, ** p < 0.01, ***p < 0.001, ANOVA followed by a Tukey– Kramer post-hoc test.

9

1

1 **Table S1.** Transcriptome assembly statistics. RNA-seq reads were derived from

2 samples LIB1777 (1) and colonies on 5 independent plants (1-5) of *Arabidopsis*

3 thaliana (At); Br, Brassica rapa (Br), Nicotiana benthamiana (Nb), Solanum

- 4 tuberosum (St), Chrysanthemum indicum (Ci), Helianthus annuus (Ha), Pisum
- 5 sativum (Ps), Phaseolus vulgaris (Pv) and Zea mays (Zm). The reads were
- 6 mapped to the *M. persicae* G006 genome assembly (1).
- ⁷ ^a Total transcripts were assembled from mapped reads.
- 8 ^{b,c} Transcripts that correspond to genes identified previously in Mathers et al.
- 9 **2017 (1)**.
- ¹⁰ ^d Novel transcript as identified in the transcriptome assembly reported herein.
- 11

	Total	Mapped	Percentage		Transcripts				
Samples	reads	reads	of mapped reads	Total ^a	Known ^ь	Genes⁰	Novel ^d		
LIB1777	182928264	117579937	64.28	87444	18750	11332	68694		
At1	35027735	31609619	90.24						
At2	30317626	27641464	91.17						
At3	32270976	29367540	91	61983	17781	10486	44202		
At4	33889504	30765104	90.78						
At5	29871796	27369312	91.62						
Ps1	33997244	31321585	92.13						
Ps2	33572880	30992391	92.31						
Ps3	30706147	28148247	91.67	60443	17737	10528	42706		
Ps4	38851016	35655100	91.77						
Ps5	23736763	20424890	86.05						
Br1	36105440	33390772	92.48						
Br2	27368493	23985300	87.64						
Br3	28484604	26110873	91.67	58293	17750	10489	40543		
Br4	27486534	25035464	91.08						
Br5	30044024	27584912	91.81						
Ci1	31213071	28395530	90.97						
Ci2	26059223	23430480	89.91						
Ci3	27469301	24621304	89.63	65917	17856	10621	48061		
Ci4	32534564	29376935	90.29						
Ci5	36510146	33342272	91.32						
Zm1	23898955	21330499	89.25						
Zm2	32663254	30362632	92.96						
Zm3	38833395	32998965	84.98	61157	17886	10610	43271		
Zm4	36223785	33199732	91.65						
Zm5	37985100	34900276	91.88						
Nb1	26749263	23973979	89.62						
Nb2	28476921	25844725	90.76						
Nb3	30481687	28348224	93	57533	17567	10444	39966		
Nb4	31075574	28936577	92.12						
Nb5	25055052	22819867	91.08						

	Total	Mapped	Percentage	Transcripts			
Samples	reads	reads	of mapped reads	Total ^a	Known ^b	Genes⁰	Novel ^d
Ps1	22202831	19882967	89.55				
Ps2	26907680	24899984	92.54				
Ps3	30923478	28678037	92.74	56431	17590	10488	38841
Ps4	37835896	34845885	92.10				
Ps5	30745946	28130700	91.49				
St1	22647808	19265819	85.07				
St2	26239515	23358283	89.02				
St3	29668113	27521716	92.77	60662	17731	10537	42931
St4	32534760	30241012	92.95				
St5	32905595	30331494	92.18				
Ha1	34475345	31538679	91.48				
Ha2	39135872	36179328	92.45				
Ha3	32762153	29883530	91.21	58152	17744	10490	40408
Ha4	20135006	18438242	91.57				
Ha5	24340322	22404095	92.05				
	т		45972	30127	18529	15845	

Table S2. Statistics of analyses of RNA-seq data retrieved from aphid-exposed (feeding sites) and non-exposed (control) leaves. RNA-seq reads derived from four biological replicates per treatment were mapped to both *A. thaliana* Col-0 (TAIR10) and *M. persicae* (Mp) genomes (1). *Reads mapping to Mp were realigned to At to find uniquely mapping Mp reads. The unique reads were then assigned to the transcripts obtained from the transcriptome assembly v2. ^ represents transcripts at TPM \geq 50 in at least one biological replicate and presence in at least three replicates.

		Total	Deede	Deede	Deede Dee	Deede	Deede		Mp transcripts (TPM \geq		
	Biological Replicate	reads (million)	Reads mapped to At (million)	Reads mapped to At (%)	Reads mapped to Mp	Reads mapped to Mp (%)	Reads mapped to At	Unique Mp reads*	Number of transcripts	Number of low- coding potential transcripts	
	1	28	25	90.61	18007	0.0631	5	18002	2928	170	
Aphid	2	21	19	90.02	242628	1.1091	0	242628	2599	280	
feeding	3	24	22	90.39	3847	0.0153	2	3845	1837	79	
sites	4	24	21	88.11	657024	2.6969	3	657021	3154	336	
									3186^	201^	
No	1	31	28	90.53	35	0.0001	23	12	0	0	
aphid	2	22	20	91.08	9	0.0000	5	4	0	0	
control	3	28	26	90.79	5	0.0000	5	0	0	0	
sites	4	21	19	90.55	17	0.0001	14	3	0	0	

Table S3. Primers used in this study.

Primers for qRT-PCR for RNAi

Primer names	Sequence (5'> 3')
Ya1 F	ATCAAGACCGTCGCCAAG
Ya1 R	TTGCGGTGGCAATACTGATA

Primers for qRT-PCR for detection in plants

•	
Primer names	Sequence (5'> 3')
Ya1 primer1	ATCACGCCATCGCCCCTCTC
Ya1 primer2	AAGCTGCAGGCCATCAGT
Ya1 primer3	AGTAGTGTATATAAAGGAGTCC
Ya1 primer4	CGATTCCGGCATAACACAT
Ya1 primer5	GTTTTCTCTTTTAAACCT
Ya1 primer6	CCATGAACATCGTTGTCGTC
Ya1 primer7	ATGTGACTGATTTATTTTGC
Ya1 primer8	GCGGTGGCAATACTGATATTAT
Ya1 primer9	CATCAATTGGTTGGTGGCT
Ya1 Forward	CCATGAACATCGTTGTCGTC
Ya1 Reverse	GCGGTGGCAATACTGATATTAT
Ya2 Forward	ACAAGTCCAATCTGCCAAGG
Ya2 Reverse	TGCGGTGTCAATATAATATG
Ya3F Forward	CTTGATTCCGGCATAACACATC
Ya3 Reverse	AAATGACTGATTTATTTTGC
Ya6 Forward	CATAACACATCTTTTTGTG
Ya6 Reverse	AAAATCGTGGTTTATTTTGC
Ya8 Forward	CCATGAACATCGTTGTCGTC
Ya8 Reverse	AAATGACTGATTTATTTTGT
Ya11 Forward	CCACCAAAAACCAAAAATGG
Ya11 Reverse	AAGTGTCTGATTTATTTTGC
Ya17 Forward	GGCATAACACATATTTTTGT
Ya17 Reverse	TGAATATCGGATGTGGCTGTT
Ya23 Forward	CTTCTGTCGGCCGCCTCGAATT
Ya23 Reverse	AGAAGACTGATTTATTTTGCAA

Primers for cloning

Underlined lower case nucleotides indicate restriction enzyme sequences Primer names Sequence (5'> 3')

Ya1_F (273 nt)	ТТТСТСТТТТАААССТААААА
Ya1_R (273 nt)	GTACATCGGATGTGACTGATT
attB1_Ya1	acaagtttgtacaaaaaagcaggctagtTTTCTCTTTTAAACCTAAAAAA
attB2_Ya1	accactttgtacaagaaagctgggt GTACATCGGATGTGACTGATT
Ya1_BamHI	cgggggactctagaggatccagtTTTCTCTTTTAAACCTAA
Ya1_Sacl	aaattcgagctcGTACATCGGATGTGACTGATTT
MutateF	cgggggactctagaggatccagtTTTCTCTTTTAAACCTAA
ATG41_F	TTTAAACCTAAAAAACCAACCAACAAATCAAAATAGGGCGCTGAAAAGGTAT
ATG41_R	ATACCTTTTCAGCGCCCTATTTTGATTTGTTGGTTGGTTTTTTAGGTTTAAA
ATG62_F	GGCGCTGAAAAGGTATCCTAGAACATCGTTGTCGTCGG
ATG62_R	CCGACGACAACGATGTTCTAGGATACCTTTTCAGCGCC
ATG146_F	CCCCAGCCACCAACCAATTGTAGAACTGTTAATTCATGTAGT
ATG146_R	ACTACATGAATTAACAGTTCTACAATTGGTTGGTGGCTGGGG
MutateR	aaattcgagctcGTACATCGGATGTGACTGATTT

Probes for northern blot hybridizations

Sequence (5'> 3') Probe Probe of AtU6 CTCGATTTATGCGTGTCATCCTTGC

Adapter and primers for 3'RACE

Names

Names	Sequence (5'> 3')
Ya1 Forward primer	GGACAAGTCCAATCTGC
3' RNA Adapter (5' Phosphorylation and 3' Inverted dT)	P-UCGUAUGCCGUCUUCUGCUUGUidT
oligo sequence complementary to 3'RNA adapter	CAAGCAGAAGACGGCATACGA

Datasets

Dataset S1. List of differentially expressed transcripts of *M. persicae* colonies on 8 divergent plant species compared to *M. persicae* on *Arabidopsis thaliana* (At); Br, *Brassica rapa* (Br), *Nicotiana benthamiana* (Nb), *Solanum tuberosum* (St), *Chrysanthemum indicum* (Ci), *Helianthus annuus* (Ha), *Pisum sativum* (Ps), *Phaseolus vulgaris* (Pv) and *Zea mays* (Zm). 1, differentially expressed (DE) transcript; 0, not DE. Transcript ID refers to the transcriptome assembly *M. persicae* G006 v2 (GSE129669) generated herein.

Dataset S2. Results of WGCNA analyses with 13 modules enriched in differentially expressed genes (Tab A) and the 64 other co-expression modules (Tab B). Genes of tandem repeats in the scaffold are highlighted in yellow or green and numbers of tandemly repeated groups identified in each module are summarized in the table at right. ^ indicates whether genes are differentially expressed (DE), encode candidate IncRNAs, and are expressed in salivar glands and guts of *M. persicae*.

Dataset S3. List of *M. persicae* candidate IncRNAs (Tab A) and other transcripts (Tab B) found *in A. thaliana* leaves exposed to *M. persicae* (feeding sites).

Dataset S4. Characteristics of the 30 Ya gene family members upon manual annotation. Gene IDs correspond to version 3 GFF generated after manual correction.

Dataset S5. Annotation of *Ya* gene in five aphid species in addition to *M. persicae*. Accession numbers and versions of genome assemblies and RNASeq data used for annotations of the *Ya* genes (Tab A); Genomic locations, strand information and annotation of *Ya* genes (Tab B).

SI References

- T. C. Mathers, Y. Chen, G. Kaithakottil, F. Legeai, S. T. Mugford, P. Baa-Puyoulet, A. Bretaudeau, B. Clavijo, S. Colella, O. Collin, T. Dalmay, T. Derrien, H. Feng, T. Gabaldón, A. Jordan, I. Julca, G. J. Kettles, K. Kowitwanich, D. Lavenier, P. Lenzi, S. Lopez-Gomollon, D. Loska, D. Mapleson, F. Maumus, S. Moxon, D. R. Price, A. Sugio, M. van Munster, M. Uzest, D. Waite, G. Jander, D. Tagu, A. C. Wilson, C. van Oosterhout, D. Swarbreck, S. A. Hogenhout, Rapid transcriptional plasticity of duplicated gene clusters enables a clonally reproducing aphid to colonise diverse plant species. *Genome Biol.* 18, 27 (2017).
- F. Krueger, Trim Galore: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, with some extra functionality for MspI-digested RRBS-type (Reduced Representation Bisufite-Seq) libraries. URL http://www. bioinformatics. babraham. ac. uk/projects/trim galore/. (2012).
- M. Pertea, D. Kim, G. M. Pertea, J. T. Leek, S. L. Salzberg, Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650-1667 (2016).
- H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R .Durbin, 1000 Genome Project Data Processing Subgrou, The sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079 (2009).
- M. Pertea, G. M. Pertea, C. M. Antonescu, T. C. Chang, J. T. Mendell, S. L. Salzberg, StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290-295 (2015).
- A. Roberts, H. Pimentel, C. Trapnell, L. Pachter, Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics* 27, 2325-2329 (2011).
- L. Venturini, S. Caim, G. G. Kaithakottil, D. L. Mapleson, D. Swarbreck, Leveraging multiple transcriptome assembly methods for improved gene structure annotation. *Gigascience* 7 (2018).

- 8. A. Bairoch, R. Apweiler, The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic. Acids Res.* **28**, 45-48 (2000).
- R. D. Finn, P. Coggill, R. Y. Eberhardt, S. R. Eddy, J. Mistry, A. L. Mitchell, S. C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, G. A. Salazar, J. Tate, A. Bateman, The Pfam protein families database: towards a more sustainable future. *Nucleic. Acids Res.* 44, D279-285 (2016).
- Y. J. Kang., D. C. Yang, L. Kong 1, M. Hou, Y. Q. Meng, L. Wei, G. Gao, CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic. Acids Res.* 45, W12-W16 (2017).
- S. E. Lewis, S. M. Searle, N. Harris, M. Gibson, V. Lyer, J. Richter, C. Wiel, L. Bayraktaroglu, E. Birney, M. A. Crosby, J. S. Kaminker, B. B. Matthews, S. E. Prochnik, C. D. Smithy, J. L. Tupy, G. M. Rubin, S. Misra, C. J. Mungall, M. E. Clamp, Apollo: a sequence annotation editor. *Genome Biol.* 3, RESEARCH0082 (2002).
- 12. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
- P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559 (2008).
- J. S. Ramsey, A. C. Wilson, M. de Vos, Q. Sun, C. Tamborindeguy, A. Winfield, G. Malloch, D. M. Smith, B. Fenton, S. M. Gray, G. Jander, Genomic resources for *Myzus persicae*: EST sequencing, SNP identification, and microarray design. *BMC Genomics* 8: 423 (2007).
- 15. M. D. Nelson, D. H. Fitch, Overlap extension PCR: an efficient method for transgene construction. *Methods Mol. Biol.* **772**, 459-470 (2011).
- N .Bechtold, J. Ellis, G. Pelletier, In planta Agrobacterium-mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C. R. Acad. Sci. Ser. III Sci. Vie. Life Sci.* **316**,1194–1199 (1993).
- 17. T. D. Schmittgen, K. J. Livak, Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* **3**, 1101-1108 (2008).
- 18. F. Supek, M. Bosnjak, N. Skunca, T. Smuc, REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, e21800 (2011).

- K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772-780 (2013).
- 20. I. Sela, H. Ashkenazy, K. Katoh, T. Pupko, GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucleic. Acids Res.* **43**, W7-14 (2015).