**Materials and Methods**

**Plant material and growth conditions**

The wild-type *Arabidopsis thaliana* accession Col-0 was obtained from Nottingham Arabidopsis Stock Centre. The *vir-1* and VIR complemented (*VIR::GFP-VIR*) lines were provided by Kamil Ruzicka, Brno (*1*). The *hen2-2* (Gabi\_774HO7) mutant was provided by Dominique Gagliardi, Strasbourg. The seeds were sown on MS10 media plates, stratified at 4°C for two days, germinated in a controlled environment at 22°C under 16h light/8h dark conditions, and harvested 14 days after transfer to 22°C.

**RNA isolation**

Following manufacturer’s instructions, total RNA was isolated using RNeasy® Plant Mini kit (QIAGEN) and treated with TURBO™ DNase (Thermo Fisher Scientific). The total RNA concentration was measured using a Qubit 1.0 Fluorometer and Qubit RNA BR Assay Kit (Thermo Fisher Scientific), while the quality and integrity was assessed using the NanoDrop™ 2000 spectrophotometer (Thermo Fisher Scientific) and Agilent 2200 TapeStation System (Agilent).

**mGFP *in vitro* transcription**

mGFP coding sequence was amplified using CloneAmp HiFi PCR Premix (Clontech) and a forward primer containing T7 promoter sequence. PCR product was purified using GeneJET Gel Extraction (Thermo Fisher Scientific) and DNA Cleanup Micro Kit (Thermo Fisher Scientific) as per manufacturer’s instructions. mGFP was *in vitro* transcribed using mMESSAGE mMACHINE™ T7 ULTRA Transcription Kit (Thermo Fisher Scientific) with and without the addition of a cap analog (Anti-Reverse Cap Analog (ARCA)) according to manufacturer’s instructions. mGFP transcripts were treated with TURBOTM DNase, polyA-tailed using *Escherichia coli* Poly(A) Polymerase (E-PAP) and ATP (Thermo Fisher Scientific) and recovered using MEGAclear™ Kit (Thermo Fisher Scientific) as per manufacturer’s instructions. mGFP mRNAs quantity was assessed using a Qubit 1.0 Fluorometer (as described above), while the quality and integrity was checked using the NanoDrop™ 2000 spectrophotometer and denaturing agarose-gel electrophoresis. ARCA-capped and non-capped *in vitro* mGFP mRNAs were used in the library preparation for DRS using nanopores.

**Preparation of libraries for Illumina RNA sequencing**

Illumina RNA sequencing libraries from purified mRNA were prepared and sequenced by the Centre for Genomic Research at University of Liverpool using NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (New England Biolabs® Inc.). Paired-end sequencing with read length of 150bp was carried out on Illumina HiSeq 4000. Illumina RNA libraries from ribosome-depleted RNA were prepared using the Illumina TruSeq Stranded Total RNA with Ribo-Zero Plant kit (Illumina®). Paired-end sequencing with read length of 100bp was carried out on Illumina HiSeq2000 at the Genomic Sequencing Unit of the University of Dundee. ERCC RNA Spike-In mixes (Thermo Fisher Scientific) (*3, 4*) were included in each of the libraries using concentrations advised by the manufacturer.

**Mapping of Illumina RNA sequencing data**

Reads were mapped to the TAIR10 reference genome with Araport11 reference annotation (*17*), using STAR version 2.6.1 (*5*), a maximum multimapping rate of 5, a minimum splice junction overhang of 8 nt (3 nt for junctions in the Araport11 reference), a maximum of 5 mismatches per read, and intron length boundaries of 60 - 10,000 nt.

**Differential gene expression analysis using Illumina RNA sequencing data**

Transcript level counts for Illumina RNA sequencing reads were estimated by pseudoalignment with salmon version 0.11.2 (*6*). Counts were aggregated to gene level using tximport (*7*) and differential expression gene expression analyses for *vir-1* mutantvs wild-type and *vir-1* mutant vs VIR complemented line were conducted in R version 3.5 using edgeR version 3.24.3 (*8*).

**Differentially Expressed region analysis using Illumina RNA sequencing data**

Mapped read pairs originating from forward and reverse strands were separated and coverage tracks were generated using samtools version 1.9 (12). Coverage tracks were then used as input for DERfinder version 1.16.1 (ref). Expressed regions were identified using a minimum coverage of 10 reads, and differential expression between *vir-1* and VIR complemented was conducted using the analyseChr method using 50 permutations.

**Differential Exon Usage analysis using Illumina RNA sequencing data**

Annotated gene models from Araport11 were divided into transcript chunks (i.e. contiguous regions within which each base is present in the same set of transcript models). Read counts for each chunk were generated using bedtools version 2.27.1 (*30*) intersect in count mode. Chunk counts were then processed using DEXseq version XX (ref) to identify differentially expressed chunks between *vir-1* and VIR complemented lines. Chunks were annotated as five prime variation if they included a start site of any transcript, and three prime variation if they contained a termination site. Chunks representing overhangs from alternative donor or acceptor sites were also separately classified. Internal exons were subclassified into cassette exons if they could be wholly contained within any intron.

**Preparation of libraries for Direct RNA sequencing (DRS) using nanopores**

mRNA was isolated from approximately 75 μg of total RNA using the Dynabeads® mRNA purification kit (Thermo Fisher Scientific) following the manufacturer’s instructions. The quality and quantity of mRNA was assessed using the NanoDrop™ 2000 spectrophotometer (Thermo Fisher Scientific). Nanopore libraries were prepared from 1 μg of poly(A)+ RNA combined with 1μl of undiluted ERCC RNA Spike-In mix (Thermo Fisher Scientific) using the nanopore DRS Kit (SQK-RNA001 for wild-type, *vir-1, VIR::GFP-VIR*, and SQK-RNA002 for *hen2-2* mutant; Oxford Nanopore Technologies Ltd.) according to manufacturer’s instructions. Quickly poly(T) adapter was ligated to the mRNA using T4 DNA ligase (New England Biolabs® Inc.) in the Quick Ligase reaction buffer (New England Biolabs® Inc.) for 15 min at room temperature. The first strand cDNA was synthesized by SuperScript III Reverse Transcriptase (Thermo Fisher Scientific) using the oligo(dT) adapter. The RNA-cDNA hybrid was purified using Agencourt RNAClean XP magnetic beads (Beckman Coulter). The sequencing adapter was ligated to the mRNA using T4 DNA ligase (New England Biolabs® Inc.) in the Quick Ligase reaction buffer (New England Biolabs® Inc.) for 15 min at room temperature followed by a second purification using Agencourt beads as described above. Libraries were loaded on R9.4 SpotON Flow Cells (Oxford Nanopore Technologies Ltd.) and sequenced on a 48-hour runtime.

To incorporate cap-dependent ligation of a biotinylated 5’ adapter RNA, the following modifications were introduced into the library preparation protocol. 4 µg of mRNA was de-phosphorylated by Calf Intestinal Alkaline Phosphatase (Thermo Fisher Scientific) and 5’ cap was removed by Cap-Clip™ Acid Pyrophosphatase (Cambio) according to the manufacturer’s instructions. Next, GeneRacer™ RNA oligo biotinylated at the 5’ end was ligated to dephosphorylated, de-capped mRNA using T4 RNA ligase I (New England Biolabs® Inc.) and mRNA was purified using Dynabeads™ MyOne™ Streptavidin C1 beads (Thermo Fisher Scientific) according to the manufacturer’s instructions. mRNA quality and quantity was assessed using the NanoDropTM 2000 spectrophotometer and used for nanopore DRS library preparation as described above.

**Processing of nanopore DRS data**

Reads were basecalled with guppy version 2.3.1 (Oxford Nanopore Technologies) using default RNA parameters and converted from RNA to DNA fastq using seqkit version 0.10.0 (*9*). Reads were aligned to the TAIR10 *Arabidopsis thaliana* genome (*10*) and ERCC RNA spike-in sequences (*3, 4*) using minimap2 version 2.8 (*11*) in spliced mapping mode using a kmer size of 14 and a max intron size of 10,000. SAM and BAM file manipulations were performed using samtools version 1.9 (*12*).

proovread version 2.14.1 (*13*)was used to correct errors in the nanopore DRS read. Each nanopore DRS replicate was split into 200 chunks for parallel processing. Each chunk was corrected using 4 samples of Illumina poly(A) RNAseq data, selected randomly from the 36 Illumina files (6 biological replicates sequenced across 6 lanes). Illumina reads 1 and 2 were merged into fragments using FLASh version 1.2.11 (*14*). Unjoined pairs were discarded. Error correction with proovread was conducted in sampling free mode using a minimum nanopore read length of 50 nt. Corrected reads were then mapped to the reference using minimap2 as described above.

**Error profile analysis using Nanopore DRS data**

Error rate analysis of aligned reads was conducted on ERCC RNA Spike-in mix controls using pysam version 0.15.2 (*15*) for BAM file parsing. Matches, mismatches, insertions and deletions to the reference were extracted from the cs tag (a more informative version of CIGAR string, output by minimap2) and normalised by aligned length of the read. Reference bases and mismatch bases per position were also recorded and used to assess the frequency of each substitution and indel type by reference base.

**Poly (A) length estimation using Nanopore DRS data**

Poly (A) tail length estimations were produced using Nanopolish version 0.11.0 (*16*) and added as tags to bam files using pysam version 0.15.2 (*15*). Per gene length distributions were then produced using the Araport11 annotation, and genes with significant changes in length distribution in the *vir-1* mutant compared to wild-type were identified using a Kolmogorov-Smirnoff test*.* *p*-values were adjusted for multiple testing using Benjamini-Hochberg correction.

**Splicing analysis of Nanopore DRS and Illumina RNAseq data**

Splice junction locations, their flanking sequences, and the read counts supporting them were extracted from Illumina RNA sequencing, Nanopore DRS, and Nanopore error-corrected DRS reads using pysam version 0.14 (*15*), and from Araport11 (*17*) and AtRTD2 (*18*) reference annotations. Splice junctions at the same position, but on opposite strands, were counted independently. Junctions were classified by their most likely snRNP machinery using biopython version 1.71 (*19*), with position weight matrices calculated by (*20*). Position weight matrices were scored against the sequence –3 nt to +10 nt of the donor site, and –14 nt to +3 nt of the acceptor site. Junctions were further categorised as canonical or non-canonical based on the presence or absence of the GT/AG intron motif. For isoform analysis, linked splices from the same read were extracted from full length Nanopore error corrected reads, and counted to create unique sets of splice junctions. Intronless reads were not counted. UpSet plots were generated in Python 3.6 using code adapted from the package upsetplot (*21*).

**3’ end analyses of Nanopore and Helicos DRS reads**

Helicos DRS data were prepared as described in (*22*) and (*23*). Positions with three or more supporting reads were considered to be peaks of Nanopore or Helicos 3’ ends. The distance between each Nanopore peak and the nearest Helicos peak was then determined. 37% of Nanopore peaks occurred at the same position as a Helicos peak, and the standard deviation in distance was 12.5 nt.

To determine the percentage of Nanopore DRS 3’ ends mapping within annotated genic features, transcripts were first flattened into a single record per gene. Exonic annotation was given priority over intronic or intergenic annotation and CDS annotation over UTR. Reads were assigned to genes if they overlapped them by >20% of their aligned length and the annotated feature type of the 3’ end position was determined. Counts were generated both for all reads and for unique positions per gene.

**Over-splitting analysis of Nanopore DRS data**

To identify read pairs which result from over-splitting of signal originating from a single RNA molecule, the sequencing summary files produced by guppy were parsed for sequencing time and channel ID. These were used to identify pairs of consecutively sequenced reads. Genomic locations of reads were parsed from minimap2 mappings and consecutively sequenced reads which aligned adjacently with a genomic distance of between –10 nt and 1000 nt between them were identified. Samples sequenced before or during May 2018 had very low levels of over-splitting (between 0.01 and 0.05% of reads) compared to those sequenced in September 2018 onwards (between 0.8 and 1.5% of reads).

**Analysis of the potential for internal priming in Nanopore DRS data**

To determine if internal priming caused by the RT step can occur in Nanopore data, the location of oligo(A) hexamers within Arabidopsis CDS regions was determined and reads which terminated within a 20 nt window of each hexamer were counted. Of 10,844 CDS oligo(A) runs, 248 (2.29%) had at least one supporting read in one Col-0 Nanopore dataset. Of these, 197 were only supported by one replicate, and only 20 were supported by four or more replicates. 152 (61.29%) occurred in terminal exons, suggesting they may be genuine CPA sites.

**Isoform collapsing of Nanopore DRS data**

Error corrected full length alignments were collapsed into clusters of reads with identical sets of introns. These clusters were then subdivided by 3’ end location by using a Gaussian kernel with sigma of 100 to find local minima between read ends, which were used as cut points to separate clusters. The read with the longest aligned length in each cluster was used as the representative in the figure.

**5’ adapter detection analyses using nanopore DRS data**

To produce positive and negative examples of 5’ adapter containing sequences, 5’ soft-clipped regions were extracted from aligned reads for the Col-0 tissue 2916 dataset, using a custom python script (Supplemental Code?). These soft-clipped sequences were then searched for the presence of the GeneRacer™ adapter sequence using blastn version 2.7.1 (*24*). Two rules were initially applied to filter blastn results: a match of 10 nt or more to the 44 nt adapter, and an E value of less than 100. Reads that failed one or both of these criteria were used as negative training examples. A final rule requiring the match to the adapter sequence to occur directly adjacent to the aligned read was also applied. Reads that passed all three rules were used as the positive training set. When comparing the ratio of positive to negative examples between datasets containing the adapter and those generated from the same tissue without, we found that these three rules gave a signal to noise ratio of >5000 (Table 2).

72,083 positive and 123,739 negative training examples from Col-0 tissue 2916 were collected to train the neural network. We then estimated the amount of raw signal from the 5’ end of the squiggle which was required on average to capture the 5’ adapter. To do this, we used Nanopolish eventalign version 0.11.0 (*25*) to identify the interval in the raw read which corresponded to the mRNA alignment to the reference in the adapter positive examples. Since the adapter can be identified immediately adjacent the alignment in sequence space for these reads, the signal after the event alignment must correspond to signal originating from the adapter. The median length of these signals was 1441 points, and 96% of the signals were less than 3000. Therefore, we used a window size of 3,000 to make predictions.

The model was trained in python 3.6 using Keras version 2.2.4 with Tensorflow version 1.10.0 backend (*26, 27*). A ResNet style architecture was used (*28*), composed of 8 residual blocks containing two convolutional layers of kernel size 5 and a shortcut convolution with kernel size 1. Down-sampling using maximum pooling layers with a stride of 2 was used between each residual block. A penultimate densely connected layer of size 16 was used, with training dropout of 0.5. Input signals were standardized by median absolute deviation scaling across the whole read before the final 3,000 points were taken, and the negative samples were augmented by addition of random internal signals from reads and pure gaussian, multi-gaussian, and perlin noise signals. The whole dataset was also augmented on the fly during training by the addition of gaussian noise with standard deviation of 0.1. Models were trained for a maximum of 100 epochs (batch size of 128, 100 batches per epoch, positive and negative examples sampled in a 1:1 ratio) using the RMSprop optimiser with an initial learning rate of 0.001, which was reduced by a factor of 10 after three epochs with no reduction in validation loss. Early stopping was used after five epochs with no reduction in validation loss. Model performance was evaluated using five-fold cross validation and by testing on independently generated datasets from Col-0 tissue 2918, produced with and without the adapter ligation protocol (Fig S3B,C).

For evaluation of the reduction in 3’ bias of adapter ligated datasets, we used the Araport11 exon annotations to produce per base coverage for each gene in the Col-0 tissue 2918 dataset. Coverage was generated separately for reads predicted to contain adapters and those that did not. Leading zeros at the very 5’ and 3’ of genes were assumed to be caused by mis-annotation of UTRs and so were trimmed. The quartile coefficient of variation (interquartile range / median) was then used as a robust measure of variation in coverage across each gene. To orthogonally validate the 5’ ends of adapter ligated reads, the distance to TSSs identified by EST cloning in the Plant Promoter DB3 dataset (*29*) was calculated using bedtools version 2.27.1. The amount of 5’ end sequence which is rescued when 5’ adapters are used was estimated by identifying the largest peak in 5’ end locations per gene in the absence of adapter, and then measured how this peak shifted in reads predicted to contain adapters.

**Differential error site analysis using nanopore DRS data**

To detect sites of Virilizer dependent m6A RNA modifications, we developed scripts to test changes in per base error profiles of aligned reads. Pileup columns for each position with >10 reads coverage were generated using pysam (*15*) and reads in each column were categorised as either A, C, G, T or indel. The relative proportions of each category were counted. Counts from replicates of the same condition were aggregated and a 2x5 contingency table was produced for each base comparing ­*vir-1* and VIR complemented lines. A G-test was performed to identify bases with significantly altered error profiles. For bases with *p-*value less than 0.05, G-tests for homogeneity between replicates of the same condition were then performed. Multiple testing correction was carried out using the Benjamini-Hochberg method. Sites with false discovery rate (FDR) indels. The log2 fold change in mismatch to match ratio between VIR complemented and *vir-1* was calculated using the Haldane correction for zero counts. Bases which had a log fold change > 1 were considered to have a reduced error rate in the *vir-1* mutant.

To identify motifs enriched at sites with reduced error rate, error sites were increased in size by 5nt in each direction and overlapping sites were merged using bedtools version 2.27.1 (*30*). Sequences corresponding to these sites were extracted from the TAIR10 reference and overrepresented motifs were detected in the sequences using MEME version 5.0.2 (*31*). The presence of these motifs at error sites was then detected using FIMO version 5.0.2 (*32*), with an FDR threshold of 0.1.

**Differential gene expression analysis using nanopore DRS data**

Gene level counts were produced for each Nanopore DRS replicate using featureCounts version 1.6.3 (*33*) in long-read mode with strand specific counting. Differential expression analysis between VIR complemented and *vir-1* lines was then performed in R version 3.5 using edgeR version 3.24.3 (*8*).

**Identification of alternative 3’ end positions and chimeric RNA using nanopore DRS data**

Genes with differential 3’ end usage were identified by producing 3’ profiles of reads which overlapped with each annotated gene locus by >20%, relative to the annotated start site of the gene. These profiles were then compared between VIR complemented and *vir-1* using a Kolmogorov Smirmov test to identify changes. Multiple testing correction was performed using the Benjamini Hochberg method. To approximately identify the direction and distance of the change, the normalised single base level histograms of the Col-0 profile was subtracted from that of the mutant profile, and the minimum and maximum points in the difference profile were found. These represent the site of most reduced and increased relative usage, respectively. Results were filtered for FDR < 0.05 and absolute change of site > 13 nt.

To identify genes with significant increases in chimeric reads in the *vir-1* mutant, we used the Araport11 annotation (*17*) to identify error corrected reads which overlapped with multiple adjacent gene loci (chimeric reads) and those that originated from a single locus (non-chimeric reads). Chimeric reads were considered to originate from the most upstream gene with which they overlapped. We pooled reads from replicates of a condition and used 50 bootstrapped 75% samples without replacement to estimate the ratio of chimeric to non-chimeric reads at each gene in that condition. This was then used to estimate the change in chimeric to non-chimeric ratio in the *vir-1* mutant compared to wild-type. Loci which had more than one chimeric read in *vir-1* and demonstrated at least a two-fold increase in chimeric read ratio in more than 50% of bootstraps were considered to be sites of increased chimeric RNA formation in *vir-1* mutant.

**miCLIP**

Total RNA for miCLIP was isolated from 7.5 mg of two-week old Arabidopsis Col-0 seedlings as previously described (*34*). mRNA was isolated from ~1 mg of total RNA using oligo(dT) and streptavidin paramagnetic beads (PolyATtract® mRNA Isolation Systems, Promega) as per manufacturer’s instructions. miCLIP was carried out using 15 µg of mRNA as described in (*35*) using antibody against N6-methyladenosine (#202 003 Synaptic Systems) with minor modifications. No-antibody controls were processed throughout the experiment. 4-12% Bis-Tris gel electrophoresis was run at 180V for 50 minutes, while transfer of the RNA-antibody complexes was carried out at 30V for 60 minutes. Following the transfer, membranes were exposed to the Medical X-Ray Film Blue (Agfa) at -80⁰C overnight. Reverse transcription was carried out using barcoded RT primers - RT41, RT48, RT49 and RT50 (see Table…) (Integrated DNA Technologies). After reverse transcription, one cDNA fraction corresponding to 70-200 nt was purified from the 6% TBE-urea gel (Thermo Fisher Scientific). After final PCR, all libraries were pooled together and purified using Agencourt Ampure XP magnetic beads (Beckman Coulter) and eluted in nuclease-free water. Paired-end sequencing with read length of 100 bp was carried out on Illumina MiSeq v2 at Edinburgh Genomics of the University of Edinburgh. Input sample libraries were prepared using NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (New England Biolabs® Inc.) and sequenced on Illumina HiSeq2000 at the Tayside Centre for Genomics analysis of the University of Dundee with pair-end read length of 75bp.

**Processing for miCLIP sequencing data**

miCLIP data were assessed for quality using FastQC version 0.11.8 (*36*) and MultiQC version 1.7 (*37*). Only the forward read was used for analysis as the miCLIP site is located at the 5’ position of the forward read. 3’ adapter and poly(A) sequences were trimmed using cutadapt version 1.18 (*38*) and Unique Molecular Identifiers were extracted from the 5’ of the reads using UMI-tools version 0.5.5 (*39*). Immunoprecipitation and no antibody controls were demultiplexed and multiplexing barcodes were trimmed using seqkit version 0.10.0 (*9*). Reads were mapped to the TAIR10 reference genome with Araport11 reference annotation (*17*), using STAR version 2.6.1 (*5*), a maximum multimapping rate of 5, a minimum splice junction overhang of 8 nt (3 nt for junctions in the Araport11 reference), a maximum of 5 mismatches per read, and intron length boundaries of 60 - 10,000 nt. SAM and BAM file manipulations were performed using samtools version 1.9 (*12*). Removal of PCR duplicates was then performed using UMI-tools in directional model (*39*). miCLIP 5’ coverage and matched input 5’ coverage tracks were generated using bedtools version 2.27.1 (*30*) and these were used to call miCLIP peaks at single nucleotide resolution with Piranha version 1.2.1 (*40*) with relaxed *p*-value thresholds of 0.5. Reproducible peaks across pairwise combinations of the three replicates were identified by Irreproducible Discovery Rate (IDR) analysis using the python package idr version 2.0.3 using an IDR threshold of 0.05 (*41*). The final set of peaks was identified by pooling the three replicates and re-analysing using Piranha, then ranking the peaks by FDR and selecting the top N peaks, where N was the smallest number of reproducible peaks discovered by pairwise comparisons of the three replicates.

**RT-PCR and RT-qPCR**

Total RNA was reverse transcribed using SuperScript™ III polymerase or SuperScript™ IV VILO™ Master Mix (Thermo Fisher Scientific) according to the manufacturer’s protocol. For RT-PCR, the reactions were performed using the Advantage 2 Polymerase Mix (Clontech) using primers listed in Table S3. Next, the PCR products were purified from the gel using GeneJET Gel Extraction and DNA Cleanup Micro Kit (Thermo Fisher Scientific) and cloned into the pGEM T-Easy vector (Promega) (according to the manufacturer’s instruction) and sequenced. For RT-qPCR, the reactions were carried out using SYBR Green I (Quiagen) mix with primers listed in Table S3 following manufacturer’s instructions.

**Clock phenotype analysis**

Clock phenotype experiments were performed as described in (*42*). Briefly, plants were grown in 12‐h light/12‐h dark cycles at 22oC and 80 μmol m−2 sec−1 light for 9 days. Next, delayed fluorescence measurements were recorded every hour for 6 days at constant temperature (22oC) and under constant light (20 uM red and 20uM blue mix) . FFT non-linear Least Squares fitting to estimate period length was conducted using Biodare (*43*).

**m6A LC-MS analysis**

The m6A content analysis using Liquid Chromatography - Mass Spectroscopy (LC-MS) was performed as described in (*44*). Chromatography was carried out by the FingerPrints Proteomics facility at the University of Dundee.

**Data Availability**

Sequencing datasets described in this study have been deposited at the European Nucleotide Archive (ENA): Study, PRJEB3993; accession no, ERP003245.

1. K. Růžička *et al.*, Identification of factors required for m6A mRNA methylation in Arabidopsis reveals a role for the conserved E3 ubiquitin ligase HAKAI. *New Phytologist* **215**, 157-172 (2017).

2. J. Kim, Y. Kim, M. Yeom, J.-H. Kim, H. G. Nam, FIONA1 Is Essential for Regulating Period Length in the <em>Arabidopsis</em> Circadian Clock. *The Plant Cell* **20**, 307-319 (2008).

3. H. Lee, P. S. Pine, J. McDaniel, M. Salit, B. Oliver, External RNA Controls Consortium Beta Version Update. *Journal of genomics* **4**, 19-22 (2016).

4. L. Jiang *et al.*, Synthetic spike-in standards for RNA-seq experiments. *Genome research* **21**, 1543-1551 (2011).

5. A. Dobin *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* **29**, 15-21 (2013).

6. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nature methods* **14**, 417 (2017).

7. C. Soneson, M. Love, M. Robinson, Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences [version 2; peer review: 2 approved]. *F1000Research* **4**, (2016).

8. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* **26**, 139-140 (2010).

9. W. Shen, S. Le, Y. Li, F. Hu, SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLOS ONE* **11**, e0163962 (2016).

10. I. The Arabidopsis Genome, Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* **408**, 796-815 (2000).

11. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100 (2018).

12. G. P. D. P. Subgroup *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).

13. T. Hackl, R. Hedrich, J. Schultz, F. Förster, proovread: large-scale high-accuracy PacBio correction through iterative short read consensus. *Bioinformatics (Oxford, England)* **30**, 3004-3011 (2014).

14. T. Magoč, S. L. Salzberg, FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics (Oxford, England)* **27**, 2957-2963 (2011).

15. A. Heger, Belgrad, T. G., Goodson, M., & Jacobs, K., pysam: Python interface for the SAM/BAM sequence alignment and mapping format., (2014).

16. R. E. Workman *et al.*, Nanopore native RNA sequencing of a human poly(A) transcriptome. *bioRxiv*, 459529 (2018).

17. C.-Y. Cheng *et al.*, Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *The Plant Journal* **89**, 789-804 (2017).

18. R. Zhang *et al.*, A high quality Arabidopsis transcriptome for accurate transcript-level analysis of alternative splicing. *Nucleic acids research* **45**, 5061-5073 (2017).

19. A. Dalke *et al.*, Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422-1423 (2009).

20. N. Sheth *et al.*, Comprehensive splice-site analysis using comparative genomics. *Nucleic acids research* **34**, 3955-3967 (2006).

21. pyUpSet.

22. A. Sherstnev *et al.*, Direct sequencing of Arabidopsis thaliana RNA reveals patterns of cleavage and polyadenylation. *Nat Struct Mol Biol* **19**, 845-852 (2012).

23. N. J. Schurch *et al.*, Improved Annotation of 3′ Untranslated Regions and Complex Loci by Combination of Strand-Specific Direct RNA Sequencing, RNA-Seq and ESTs. *PLOS ONE* **9**, e94270 (2014).

24. C. Camacho *et al.*, BLAST+: architecture and applications. *BMC bioinformatics* **10**, 421-421 (2009).

25. N. J. Loman, J. Quick, J. T. Simpson, A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nature methods* **12**, 733 (2015).

26. F. Chollet, Keras. (2018).

27. A. A. Martín Abadi, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mane, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viegas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, Xiaoqiang Zheng, TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. (2016).

28. X. Z. Kaiming He, Shaoqing Ren, Jian Sun, Deep Residual Learning for Image Recognition. (2015).

29. A. Hieno *et al.*, ppdb: plant promoter database version 3.0. *Nucleic acids research* **42**, D1188-D1192 (2014).

30. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)* **26**, 841-842 (2010).

31. T. L. Bailey *et al.*, MEME Suite: tools for motif discovery and searching. *Nucleic Acids Research* **37**, W202-W208 (2009).

32. C. E. Grant, T. L. Bailey, W. S. Noble, FIMO: scanning for occurrences of a given motif. *Bioinformatics (Oxford, England)* **27**, 1017-1018 (2011).

33. Y. Liao, G. K. Smyth, W. Shi, The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic acids research* **41**, e108-e108 (2013).

34. V. Quesada, R. Macknight, C. Dean, G. G. Simpson, Autoregulation of FCA pre-mRNA processing controls Arabidopsis flowering time. *The EMBO Journal* **22**, 3142-3152 (2003).

35. A. V. Grozhik, B. Linder, A. O. Olarerin-George, S. R. Jaffrey, Mapping m(6)A at Individual-Nucleotide Resolution Using Crosslinking and Immunoprecipitation (miCLIP). *Methods in molecular biology (Clifton, N.J.)* **1562**, 55-78 (2017).

36. S. Andrews, FastQC: a quality control tool for high throughput sequence data., (2010).

37. P. Ewels, M. Magnusson, M. Käller, S. Lundin, MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047-3048 (2016).

38. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17**, 3 (2011).

39. T. Smith, A. Heger, I. Sudbery, UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome research* **27**, 491-499 (2017).

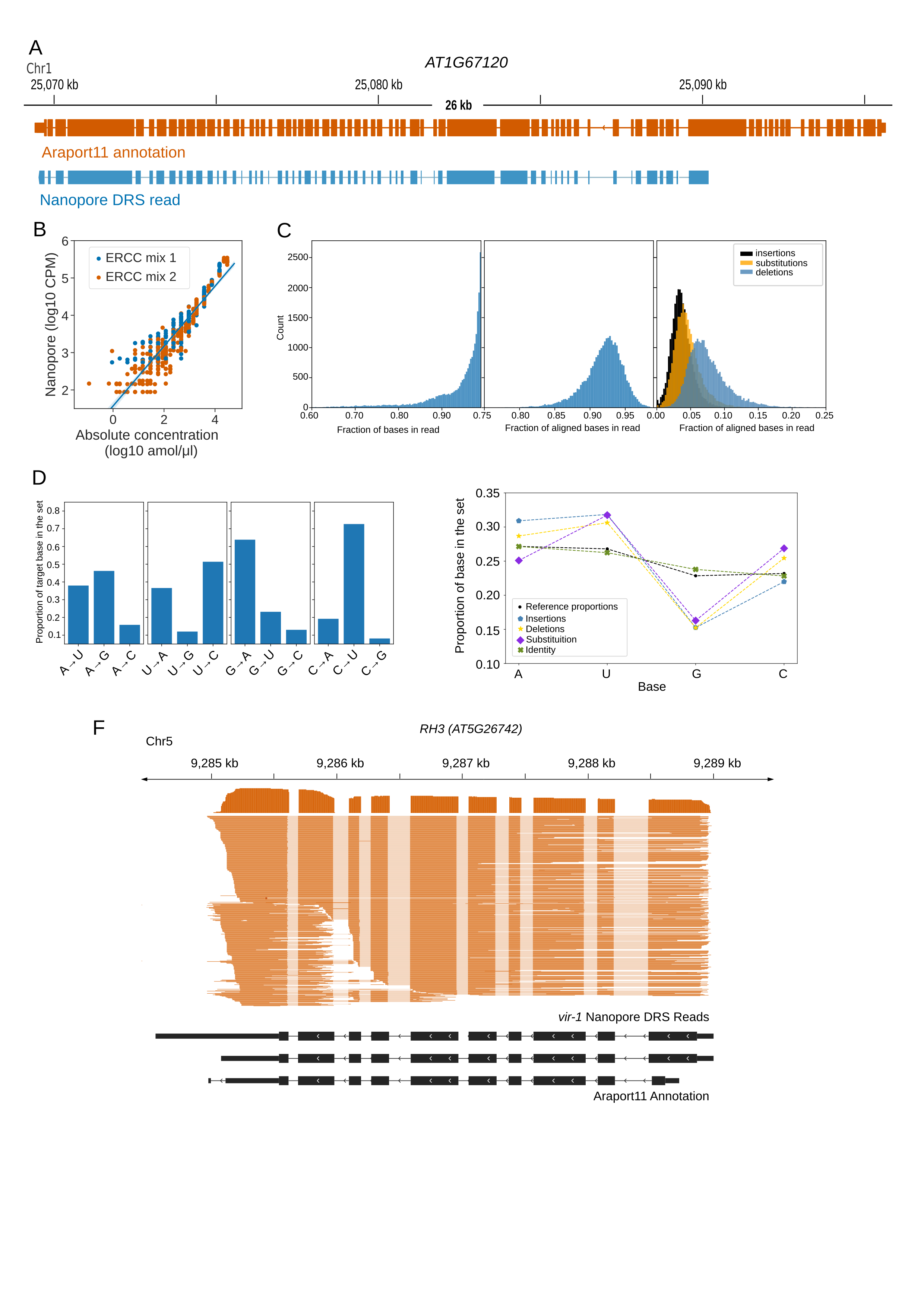
40. P. J. Uren *et al.*, Site identification in high-throughput RNA-protein interaction data. *Bioinformatics (Oxford, England)* **28**, 3013-3020 (2012).

41. Q. Li, J. B. Brown, H. Huang, P. J. Bickel, Measuring reproducibility of high-throughput experiments. *Ann. Appl. Stat.* **5**, 1752-1779 (2011).

42. P. D. Gould *et al.*, Delayed fluorescence as a universal tool for the measurement of circadian rhythms in higher plants. *The Plant Journal* **58**, 893-901 (2009).

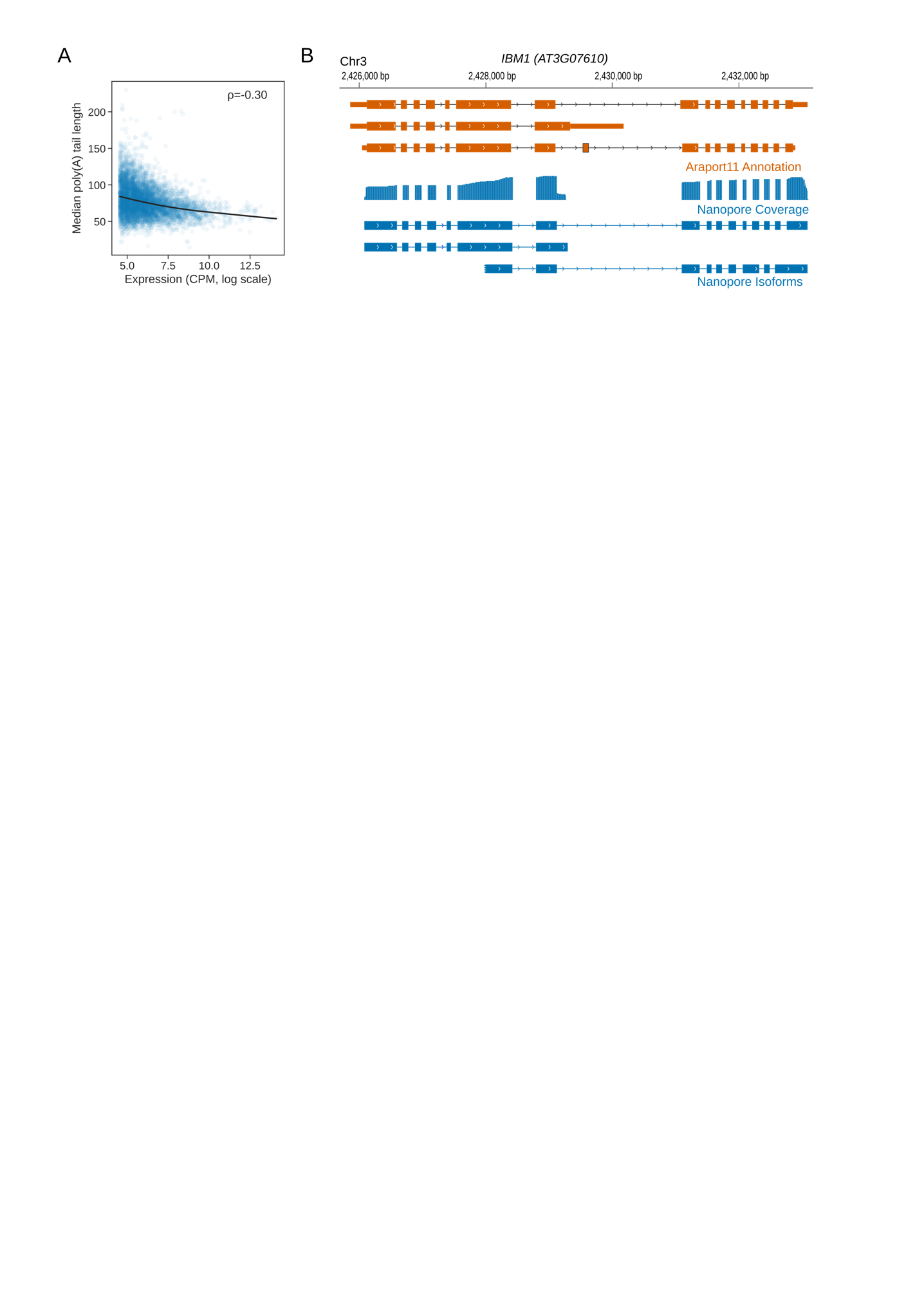
43. A. Moore, T. Zielinski, A. J. Millar, in *Plant Circadian Networks: Methods and Protocols,* D. Staiger, Ed. (Springer New York, New York, NY, 2014), pp. 13-44.

44. H. Huang *et al.*, Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nature Cell Biology* **20**, 285-295 (2018).

****

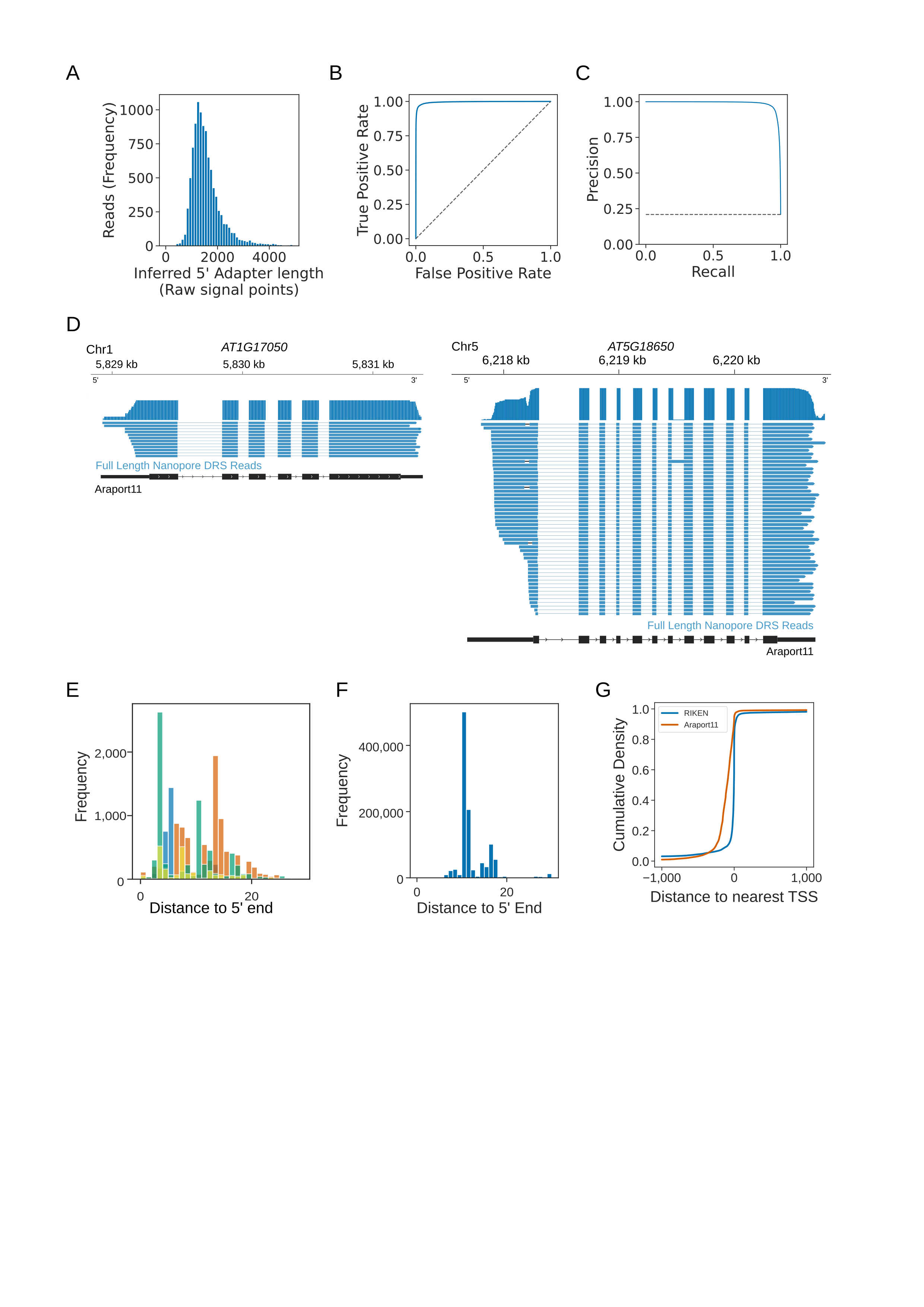
**Figure S1. Properties of the Nanopore DRS sequencing data.**

1. Nanopore DRS identified 13kb transcript generated from *AT1G67120* gene that includes 58 exons. An RNA isoform present in the *Araport11* annotation is shown in orange and an RNA isoform identified using nanopore DRS – in blue.
2. Synthetic RNA spike-ins are detected in a quantitative manner. Absolute concentrations of RNA spike-ins are plotted against counts per million reads in log10 scale. ERCC mix 1 is shown in blue, and ERCC mix 2 – in orange.
3. Overview of the sequencing and alignment characteristics of Nanopore DRS data for ERCC spike-ins. Distribution of the length fraction of each sequenced read that aligns to the ERCC spike-in reference is shown on the left; distribution of fraction of identity that matches between the sequence of the read and the ERCC reference for the aligned portion of each read – in the centre; distributions of the occurrence of insertions (black), substitutions (orange) and deletions (blue) as a proportion of the number of aligned bases in each read – on the right.
4. Substitution preference for each nucleotide (A, U, G, C, left-to-right). When substituted, guanidine is replaced with adenosine in more than 63% of its substitutions, while cytosine is replaced by uridine 73% of the time. Conversely uridine is rarely replaced with guanidine (12%) and adenine is rarely substituted with cytosine (16%).
5. Nucleotide representation within the ERCC spike-in reference sequences (black dots) compared with the nucleotide representation within four categories from the Nanopore DRS reads; identity matches between the sequence of the read and the ERCC reference (green crosses), insertions (blue pentagons), deletions (yellow stars) and substitutions (purple diamonds). Guanidine (G) is under-represented, and uridine (U) is over-represented, for all three categories of error (insertion, deletion and substitution (relative to the reference nucleotide distribution). Cytosine (C) is over-represented in the set of deletions and substitutions. Adenine (A) is over-represented for insertions and deletions and, under-represented in the set of substitutions.
6. Signals originating from the RH3 transcripts are susceptible to systematic over-splitting around exons 7-9, resulting in reads with apparently novel 5’ or 3’ positions. This appears to only occur in datasets collected after May 2018 (Table 1) and may be the result of an update to the MinKNOW software.

****

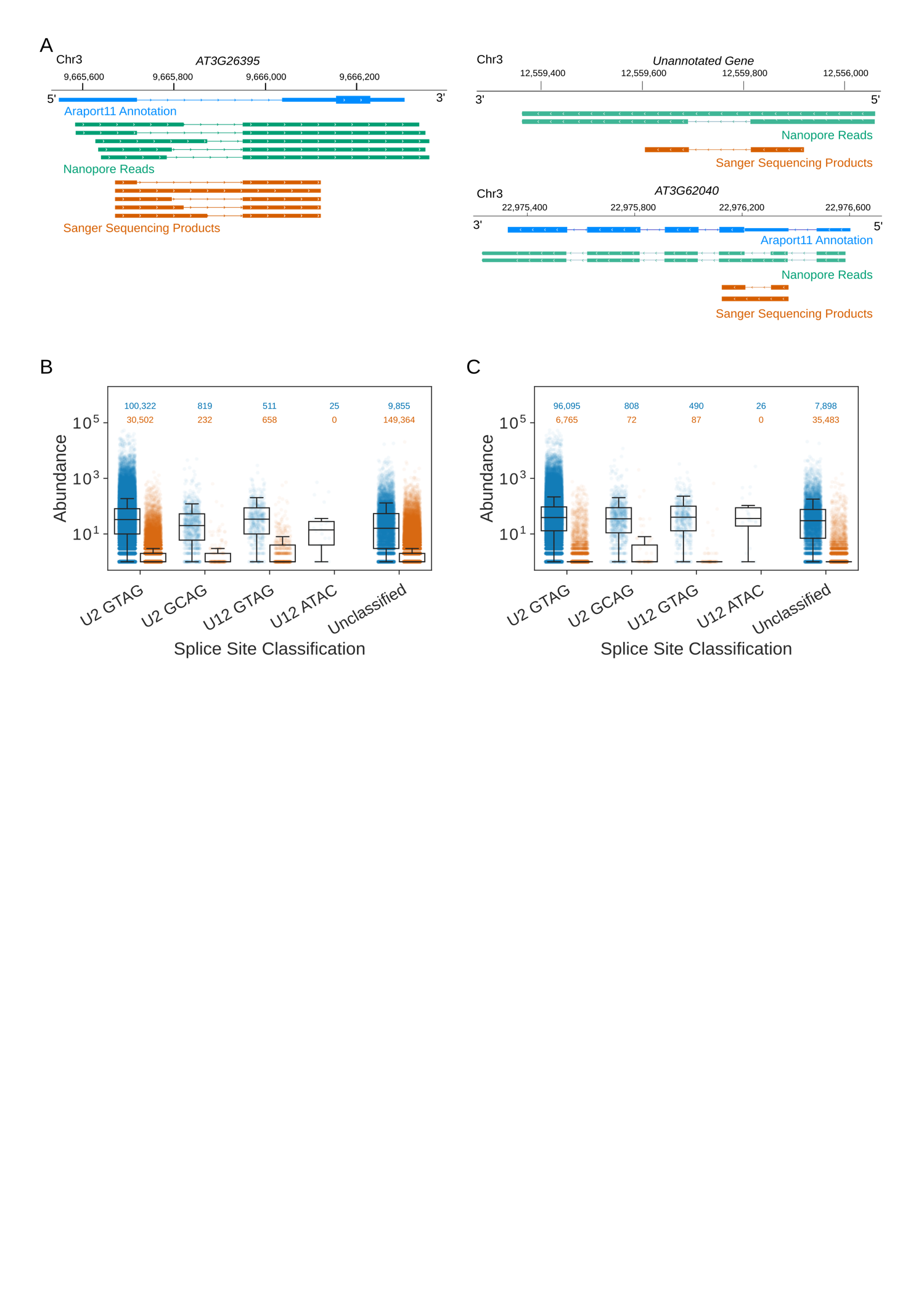
**Figure S2. 3’ end processing information provided by Nanopore DRS.**

1. Poly(A) tail length of RNA negatively correlates with the expression of the gene. Expression in log scale of counts per million (CPM) obtained from Nanopore DRS data is plotted against median poly(A) tail length.
2. Nanopore DRS identified new 3’ polyadenylation sites in RNAs transcribed from *IBM1 (AT3G07610)* gene. Blue track shows coverage of Nanopore DRS reads. Isoforms found in Araport11 annotation are shown in orange and those detected by Nanopore DRS are in blue.



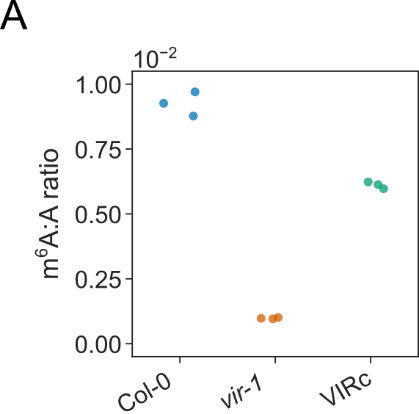
**Figure S3. 5’ end processing information provided by Nanopore DRS with cap-dependent ligation of 5’ adapter.**

1. Histogram showing the distribution of adapter length in Nanopore raw current signal, inferred from alignment of mRNA sequence to signal using nanopolish eventalign. The median signal length was 1441 points, and 96% of adapter signals were 3000 points or less.
2. Out-of-bag Receiver Operator Curve showing performance of trained convolutional neural network at detecting 5’ adapters using 3000 points of signal. Curve was generated using 5-fold cross validation.
3. Precision recall curve showing performance of trained neural network.
4. Alternative transcription start sites were identified using Nanopore DRS with cap-dependent ligation of 5’ end adapter at *AT1G17050* and *AT5G18650* genes. Blue track shows coverage of Nanopore DRS reads. Isoforms found in Araport11 annotation are shown in black and those detected by Nanopore DRS with cap-dependent liagtion of 5’ adapter are in blue.
5. Reads mapping to ERCC spike-in RNAs lack approximately 11 nt of sequence at the 5’ end. Histogram shows distance to the 5’ end for ERCC RNA spike-ins reads (each shown in different colours).
6. Reads mapping to *in vitro* transcribed GFP lack approximately nn nt of sequence at the 5’ end. Histogram shows distance to the 5’ end for *in vitro* transcribed GFP.
7. Araport11 annotation overestimates the length of 5’ UTRs. Cumulative distribution function shows the distance to the nearest TSS identified from full length transcripts cloned as part of the RIKEN RAFL project (blue) and Araport11 annotation (orange).

****

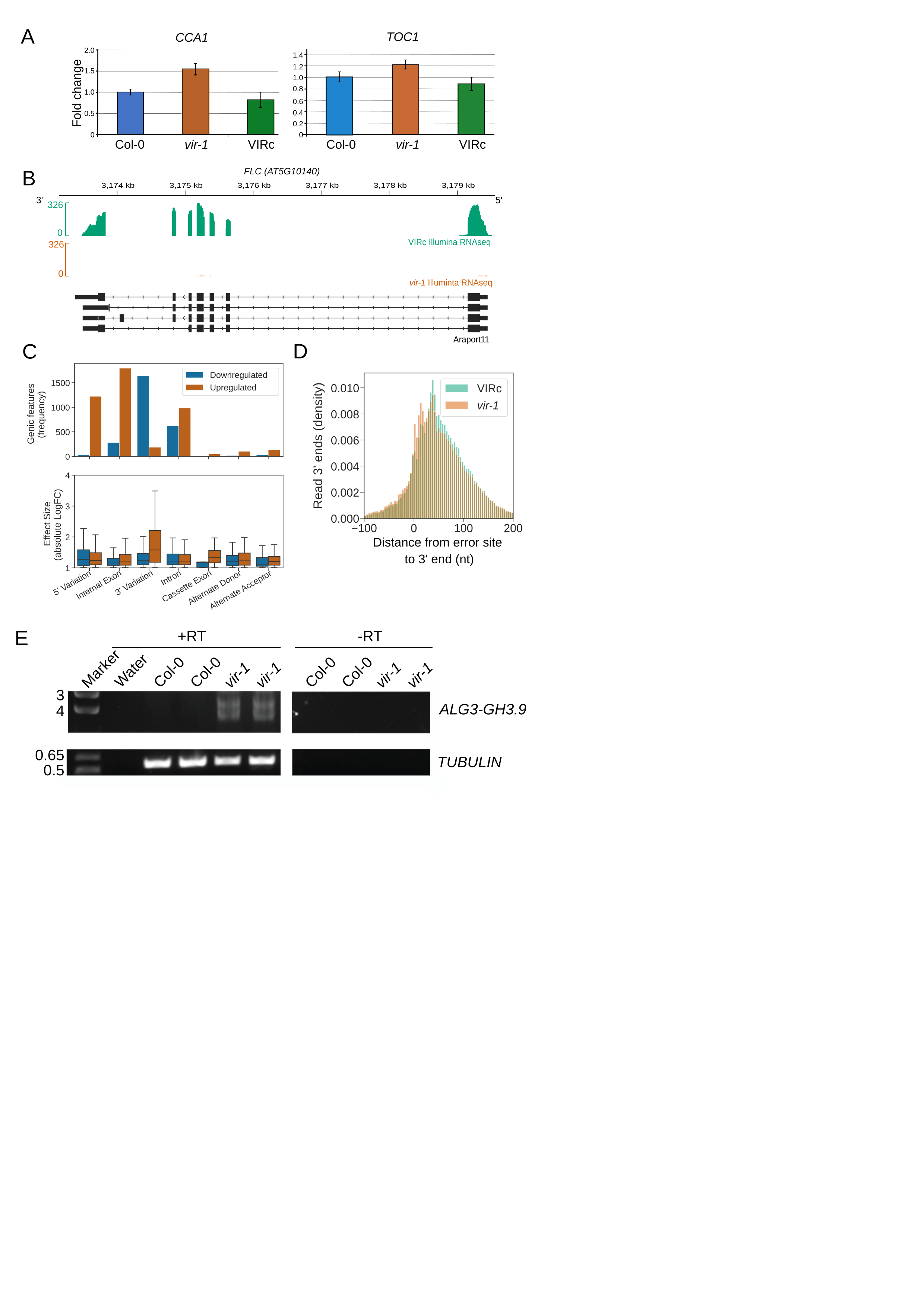
**Figure S4. New details on splicing identified using Nanopore DRS.**

1. Some of the novel splicing events identified by Nanopore DRS can be validated using RT-PCR. RNA isoforms present in Araport11 annotation are shown in blue; RNA isoforms found using Nanopore DRS – in green; Sanger sequencing products obtained using RT-PCR – in orange.
2. Classification of annotated (blue) and novel (orange) splice sites found non-error corrected Nanopore DRS data. Counts are plotted in log10 scale and the exact numbers are indicated above each category in blue for annoted events and in orange for novel events.
3. Classification of annotated (blue) and novel (orange) splice sites found error corrected Nanopore DRS data. Counts are plotted in log10 scale and the exact numbers are indicated above each category in blue for annoted events and in orange for novel events.

****

**Figure S5. VIR complemented line shows increase in m6A methylation compared to *vir-1* mutant.**

1. Col-0 ratio of methylated adenosine at the 6th position to adenosine obtained using LC-MS analysis is shown in blue; *vir-1* mutant ratio – in orange; VIR complemented line ratio – in green.

****

**Figure S6. Changes in circadian clock and generation of chimeric RNAs are observed in the mutant with reduced m6A modification.**

1. Expression of *CCA1* and *TOC1* genes, encoding regulators of circadian rhythm in Arabidopsis, is increased in *vir-1* mutant. Blue bar shows gene expression measured by RT-qPCR in Col-0; orange – in *vir-1*; green – in VIR complemented line.
2. Expression of *Flowering Locus C* (*FLC*) gene is decreased in the *vir-1* mutant. Gene tracks show gene expression from Illumina RNAseq data for VIR complemented line (green) and *vir-1* mutant (orange).
3. Splicing is moderately disrupted in *vir-1* mutant. Less retained (blue) and more retained (orange) introns are shown in various categories of splice sites.
4. A shift in use of more proximal polyadenylation site is observed in *vir-1* mutant. Histogram shows distance from error site to 3’ end in *vir-1* mutant (orange) and VIR complemented line (green).
5. *ALG3-GH3.9* chimeric RNAs are generated in *vir-1* mutant. RT-PCR gel shows formation of chimeric RNAs in the ­*vir-1* mutant compared to Col-0.