Supplemental Files

Supplemental Methods

Next-generation sequencing of YAC DNA and bioinformatic analysis

YAC DNA (1) was enzymatically sheared using the Westburg NGS DNA Library Prep Kit. After shearing indexed Illumina TrueSeq adapters were ligated to the ends with the Westburg DNA ligase. Final average insert size of the libraries was 716 bp, determined by a fragment analyzer run (Advanced Analytical, High sensitivity NGS Fragment analysis Kit). QPCR (KAPA library quantification Kit for Illumina) was carried out to determine the precise concentration. Sequencing was carried out on the Illumina Hi-Seq2500 using paired end 125 bp sequencing.

To estimate 5S rRNA gene copy number in each YAC after NGS, we divided the average coverage along the 120 bp transcribed sequence of the 5S rRNA by the average coverage of a region of 1,773 bp in length from the pYAC4 backbone comprising part of ORF2, HIS3 and part of ORF3. We trimmed and mapped reads, and retrieved per-base read depth as described in section “Copy number estimation” of Material and Methods.
T-stretch signatures of each YAC were extracted as described for the sequencing dataset of genomic DNA. T-stretch signatures with at least the same number of reads as for the YAC backbone single copy region comprising ORF2, HIS3 and ORF3 were taken into account.
To detect polymorphisms along the 5S rRNA gene in each YAC library we trimmed our dataset into 50 bp SE-reads, and employed the pipeline described in section “Mapping of polymorphisms in the MAGIC population” of Materials and Methods. Only polymorphisms present at least once for the estimated copies per YAC were plotted.

RNA extraction and RT-qPCR

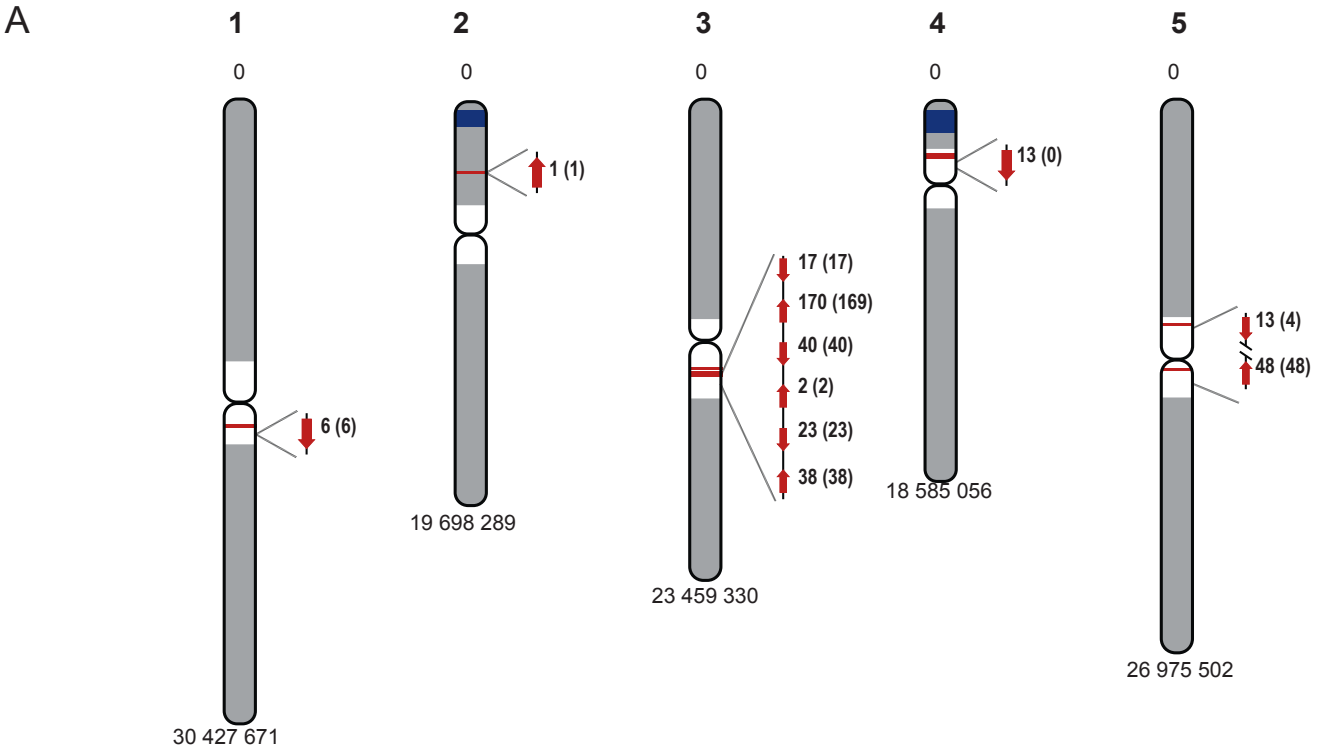
Total RNA from aerial tissue of 2-week old plants grown simultaneously *in vitro* were extracted with Tri-Reagent (Euromedex) according to manufacturer’s instructions, then treated with RQ1 DNase I (Promega) and purified by phenol-chloroform extraction. Purity and concentration of DNase I treated RNA was determined with a Nanodrop.
2 μg of RNA was reverse transcribed using random primers using M-MLV reverse transcriptase (Promega) in a total volume of 25 μL. Controls without reverse transcriptase were performed under the same conditions
Obtained cDNAs were diluted twice in water and 2 μL were used per 10 μl quantitative PCR reaction in the Roche LightCycler® 480 using the LightCycler® 480 SYBR Green I Master kit.

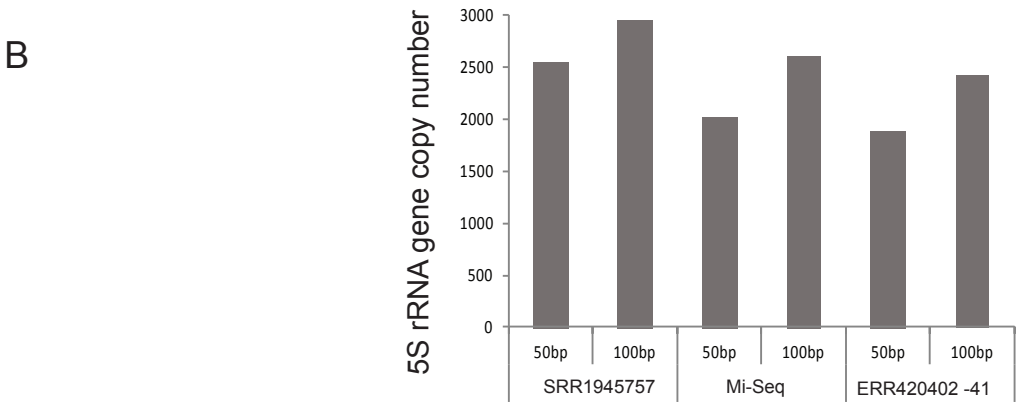
For each biological replicate, a mean of two technical replicates was assessed and qPCR data were analyzed using the ΔΔCt method. The following conditions of amplifications were applied: 3 min at 95°C; 40 cycles of 15 s at 95°C, 15 s at 55°C and 20 s at 72°C. Transcript levels were normalized to 18S rDNA. Primer sequences are indicated in Supplementary Table 6.

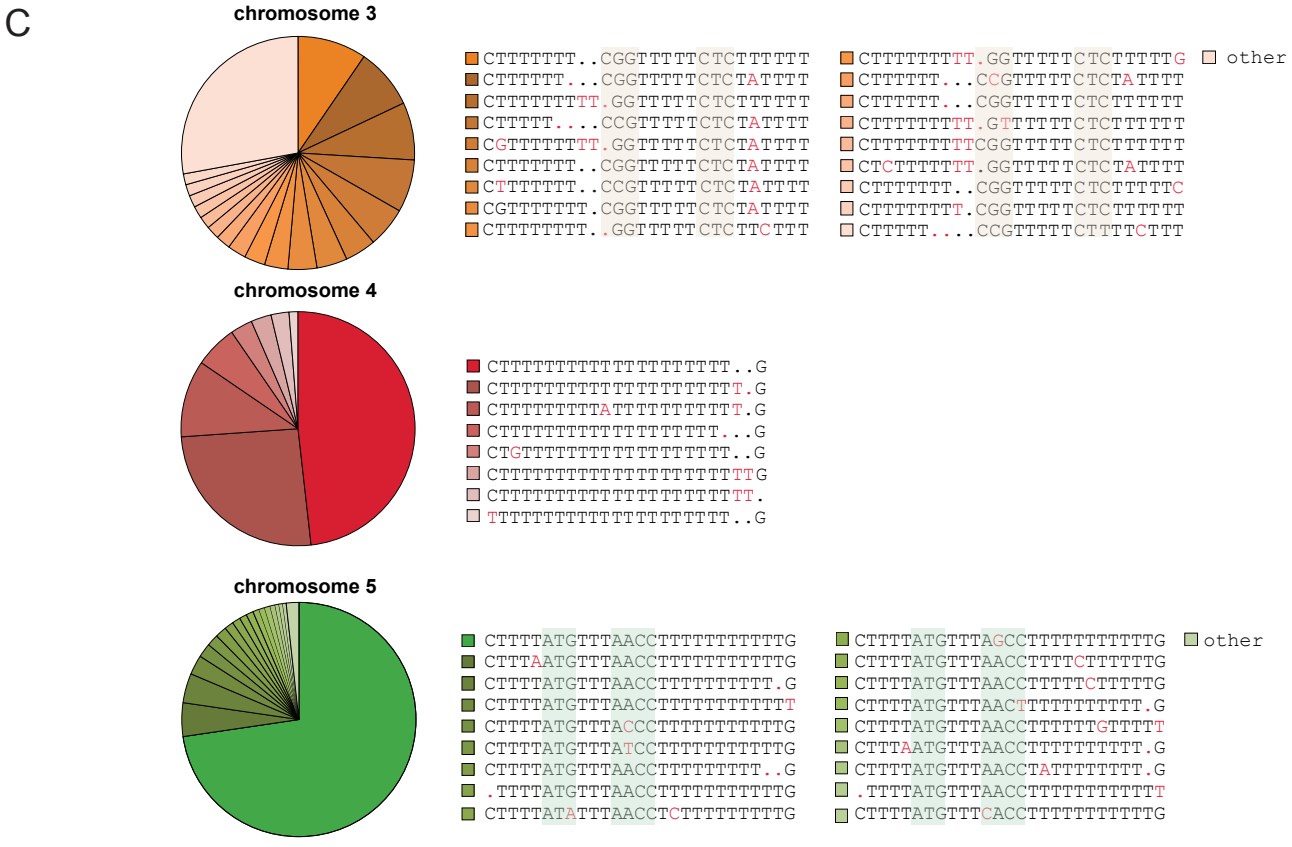
RNA-seq data-mining

To determine expression levels of *Ta3, HXK1* and *ACT2*, their transcript levels in RNA-seq datasets SRR1030234 and SRR1030235 were normalized to *MON1 (At2g28390)* as described previously (2).

Supplemental Fig. 1

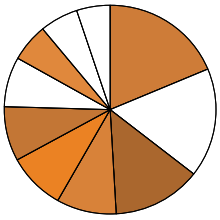






D

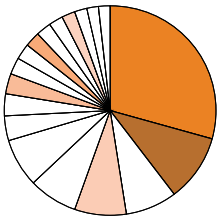
YAC 7E7, CN=92



CGTTTTTTTT.GGTTTTTCTCTATTTT
CTTTTTT...CGGTTTTTCTCTATTTT
CTTTTTTTCTCGGTTTTTCTCTTTTTT
CTTTTTTT..CGGTTTTTCTCTATTTT
CTTTTTTT..CGGTTTTTCTCTTTTTT
CTTTTT....CCGTTTTTCTCTATTTT
CTTTTTTTTT.CGGTTTTTCTCTTTTGT
CTTTTTTT..CCGTTTTTCTCTATTTT
CTTTTTTTTTTCGAGTTTTCTCTTTTTT

CTTTTTTTT..GGTTTTTCTCTATTTT

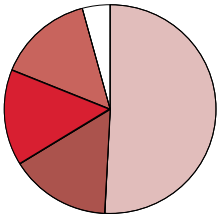
YAC 9A5, CN=195



CTTTTTTT..CGGTTTTTCTCTTTTTT
CTTTTTTTTT.GGTTTTTCTCTTTTTT
CTTTTTTT..CGGGTTTTCTTTTTTTT
CTTTTTTT..CGGTTTTTCTCTTTTTC
CTTTTTT...CGGTTTTTCTCTTTTGT
CTTTTTTTTT.GGTTTTACTCTTTTTT
CTTTTTTT..CAGTTTTTTTTTTTTTT
CTTTTTTT..CTGGTTTTTTTTTTTTT
CTTTTTTTTTCGGTTTTTCTCTTTTTT

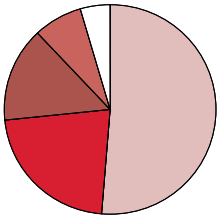
CTTTCTTTT..GGTTTTTCTCTTTTTA
CTTTTTTTT.CGGTTTTTCTCTTTCTT
CTTTTTT...CGGTTTTTCTCTTTTTT
CTTTTTTTTTT.GTTTTTCTCTTTTGT
CTTTTTTTTT.GGTTTTTCTGTTTTGT
CTTTTTTTT.CGGTTTTTCTCTTTTTT
CTTTTTTTTTCGGTTTTTCTCTTTTTC
CCTTTTTT..CGGTTTTTCAATTTTTT
CTTTTTTT..AGGTTTTTCTCTTTTTT

YAC 5C3, CN=52



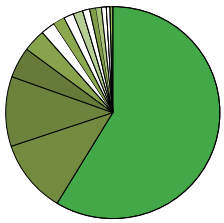
CTTTTTTTTTTTTTTTTTTTTT.
CTTTTTTTTTTTTTTTTTTT...G
CTTTTTTTTTTTTTTTTTTTT.G
CTTTTTTTTTTTTTTTTTTT..G
CTGTTTTTTTTTTTTTTTTTT.

YAC 9D3, CN=40



CTTTTTTTTTTTTTTTTTTTT.
CTTTTTTTTTTTTTTTTTTT..G
CTTTTTTTTTTTTTTTTTTT.G
CTTTTTTTTTTTTTTTTTT...G
CTTTTTTTTTTTTTTTTTT....G

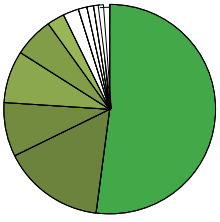
YAC 4E4, CN=279



CTTTTATGTTTAACCTTTTTTTTTTTG
CTTTTATGTTTAACCTTTTTTTTTTT
CTTTTATGTTTAACCTTTTTTTTTT.G
CTTTAATGTTTAACCTTTTTTTTTTG
.TTTTATGTTTAACCTTTTTTTTTTG
CTTTTATATTTAACCTCTTTTTTTTT
CTTTTATATTTAACCTCTTTTTTTTG
CTTTAATGTTTAACCTTTTTTTTTTT
.TTTTATGTTTAACCTTTTTTTTTTT

CTTTAATGTTTAACCTTTTCTTTTTG
CTTTTATGTTTAACCTTTTTTTTT..G
CTTTAATGTTTAACCTTTTTTTTT.G
CTTTTATGTTTAACCTTTTTTTTTTC
CTTTTATGTTTAAC.TTTTTTTTTTG
CTTTTATGTTTAACCTTTTTCTTTTG

YAC 6A1, CN=137



CTTTTATGTTTAACCTTTTTTTTTTG
CTTTTATGTTTAACCTTTTTTTTT.G
CTTTTATGTTTAACCTTTTTTTTTTG
CTTTTATATTTAACCTCTTTTTTTTG
CTTTTATGTTTAACCTTTTTTTT..G
CTTTTATGTTTAACCTTTTTCTTTTG
CTTTTATGTTTAAC.TTTTTTTTTTG
CTTTTATATTTAACCTCTTTCTTTTG
CTTTTATGTGTAACCTTTTTTTTTTG

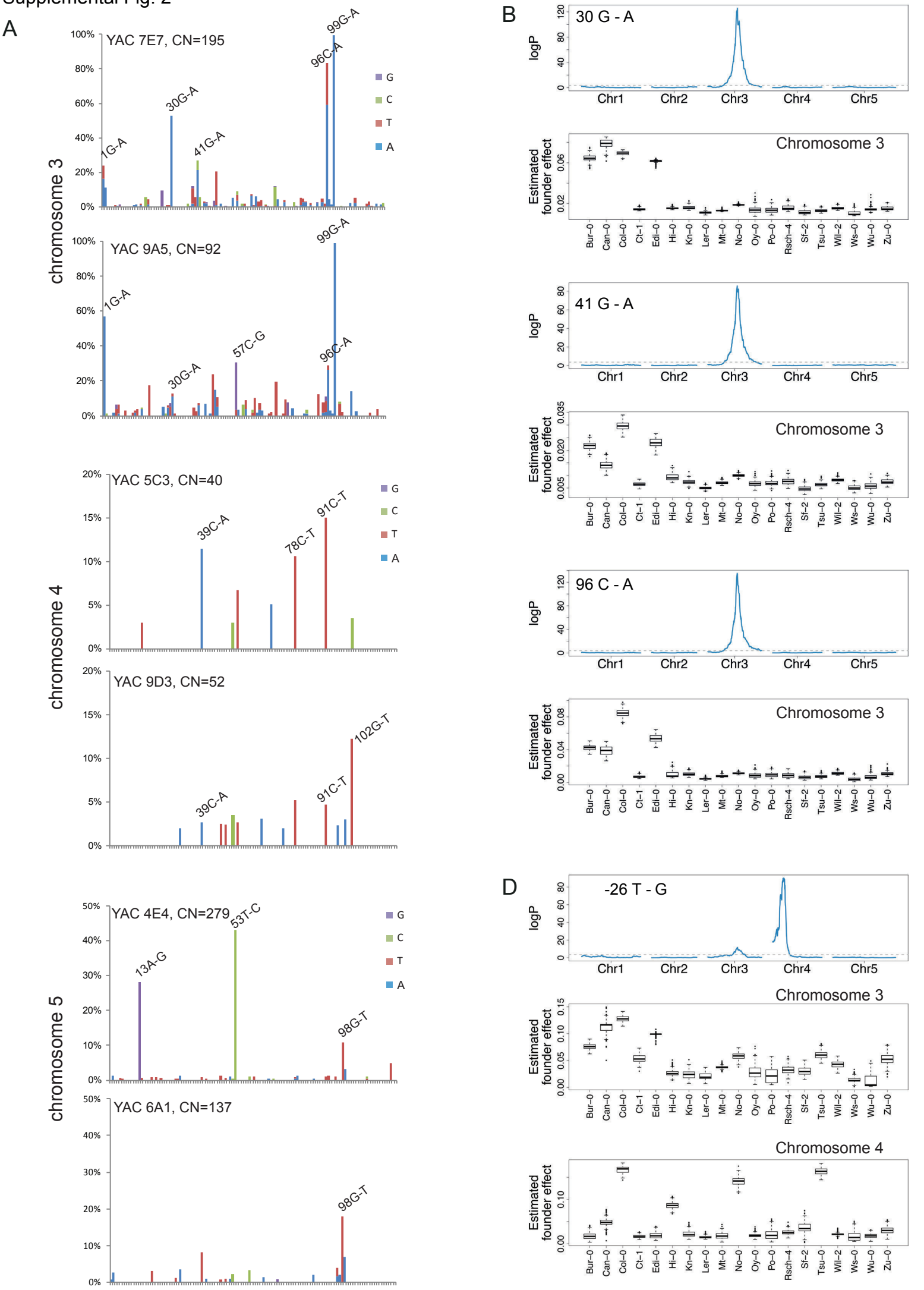
CTTTTATGTTTAACCTTTATTTTTT.G
CTTTTTTGTTTAACCTTTTTTTTT.G
ATTTTATGTTTAACCTTTTTTTTTTG

Supplemental Figures

Supplemental Figure 1: 5S rDNA organization in the pseudomolecule, influence of read length on 5S rDNA copy number determination and T-stretch signatures in Col-0.

(A) Schematic representation of the 5 Arabidopsis chromosomes showing the number and orientation of sequences with homology to the 5S rRNA consensus sequence as present in the Tair10 Col-0 genome assembly. The total number of copies is given including the number of copies bearing polymorphisms compared to the consensus sequence of the 120 bp long transcribed sequence (in brackets).
(B) 5S rRNA gene copy number was estimated in three independent next-generation-sequencing datasets for Col-0 (SRR1945757, Mi-Seq and ERR420402 - 41), with reads length set at 50 or 100bp, by dividing the average coverage along the transcribed sequence of the 5S rRNA gene by the average coverage along 18 single copy genes (see Materials and Methods). While sequence coverage along the genes used for normalization augments roughly proportionate to read length, coverage along the 5S rRNA gene raises disproportionally. As a consequence, the longer the read length, the higher the estimate of 5S rRNA gene copies.
(C) Identified T-stretch signatures and their relative percentages per chromosome in the Mi-Seq dataset of the Col-0 genome.
(D) T-stretch signatures identified in 5S rRNA genes comprised in several YACs positioned on chromosome 3 (7E7 and 9A5), chromosome 4 (5C3 and 9D3) or chromosome 5 (4E4 and 6A1). Color code is the same as in panel C. Note that most of the T-stretches determined in the Mi-Seq dataset are also found in the 5S rRNA genes comprised in the YACs. Differences between T-stretches in the Mi-Seq dataset and in the YACs are shown in white. We hypothesize that these differences, which include insertion or deletion of Ts or 1 to 2 polymorphisms can be due to replication defects through longer T-stretches in yeast or to differences between Col-0 lab strains.

Supplemental Fig. 2



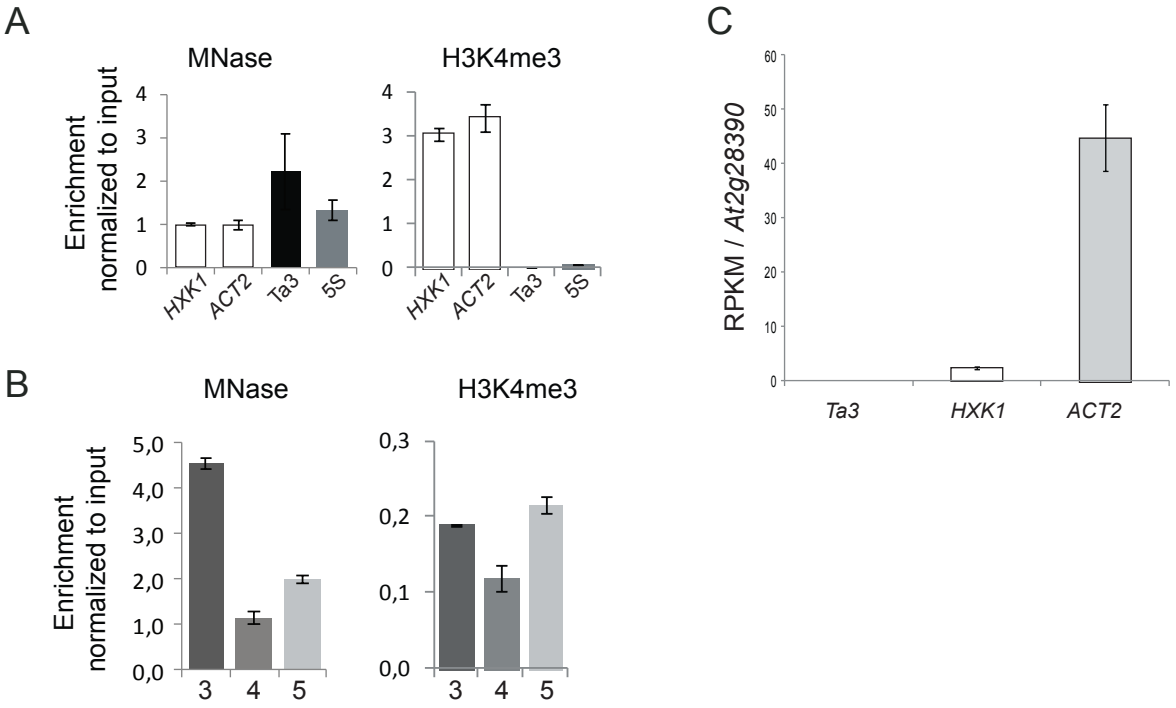
(C)

```
Chr3        AGATGCGATCATACCAGCACTAATGCACCGGATCCCATCAGAACTCCGCA
Chr4        GGATGCGATCATACCAGCACTAATGCACCGGATCCCATCAGAACTCCGCA
Chr5        GGATGCGATCATACCAGCACTAATGCACCGGATCCCATCAGAACTCCGCA
Campell     GGATGCGATCATACCAGCACTAATGCACCGGATCCCATCAGAACTCCGCA

Chr3        GTTAAGCGTGCTTGGGCGAGAGTAGTACTAGGATGGGTGACCTCCAGGAA
Chr4        GTTAAGCGTGCTTGGGCGAGAGTAGTACTAGGATGGGTGACCTCCCGGGA
Chr5        GTTAAGCGTGCTTGGGCGAGAGTAGTACTAGGATGGGTGACCTCCCGGGA
Campell     GTTAAGCGTTCTTGGGCGAGAGTAGTACTAGGATGGGTGACCTCCCGGGA

Chr3        AGTCCTCGTGTTGCATCCCTCTTTTTTTCGGTTTTTCTCTTTTTTTTTTG
Chr4        AGTCCTCGTGTTGCATCCCTCTTTTTTT-TTTT------TTTTTTTTGG
Chr5        AGTCCTCGTGTTGCATCCCTCTTT-TAT-GTTTAACCTTTTTTTTTTGG
Campell     AGTCCTCGTGTTGCATCCCTCTTTATAT-GTTTAACCTTTTTTTTTTGG

Chr3        TTAAAAATGTATGACTCTATAACTTTTAGACCGTGAGGCCAAACTTGGCA
Chr4        TTAAAACTTTATGACTCTATAACTTCTATACCGTGAGGCCAAACTTGGCA
Chr5        TTAAAACTTTATGACTCCATAACTTTTAGACCGTGAGGCCAAACTTGGCA
Campell     TTAAAACTTTATGACTCCATAACTTTTAGACCGTGAG-CCAAACTTGGCA

Chr3        TGTGATACCTTTTCGGAAAGCCCAAAGAGAGCTCTCCGATGAATGAGGAA
Chr4        TGTGATACCTTTTCGGAAAGCCCAAAGAGAGCCCTCCGACGAAAGAAGCA
Chr5        TGTGATACCTTTTCGGAAAGCCCAAAGACAGCCCTCCGACGAAAGAAGCA
Campell     TGTGATACCTTTTCGGAAAGCCCAAAGACAGCCCTCCGACGAAAGAAGCA

Chr3        GGAAAATGGAATTCTTCTATTGTTTTTTTTTCTACTCCAAATTTTGACCT
Chr4        AGACAATGGAACTTTTCCATTGACTTTTTGTCGACTCCAAATTTTGACCT
Chr5        GGACAATGGAACTTTTCCATTGACTTTTTGTCGACCCCAAATTTTGACCT
Campell     GGACAA-----CTTTTCCATTGACTTTTTGTCGACCCCAAATTTTGACCT

Chr3        TAATGTACTTTTTCGGGCCTTTTCGTGATCTTTGCTATATTACGGGGCGA
Chr4        TTAAGTACTTTTTCGGGCATTTTCGTGATTTGTGCTATATTACGGACCCA
Chr5        TTAAGTACTTTTTCGGGCATTTTCGTGATTTGGGCTATATTACGGACCCA
Campell     TTAAGTACTTTTTCGGGCATTTTCGTGATTTGGGCTATATTACGGACCCA

Chr3        AAATTATATGTTCGGACATTGTTTTCGAATATTTTGCATGTATCAAAGCT
Chr4        AAATTACTTTTTCAAGCATTGTTTTCGAATATTTTTCATGCATCAAAGCT
Chr5        AAATTACTTGTTCAAGCATTGTTTTCGAAT-TTTTTCATGCATCAAAGCT
Campell     AAATTACTTGTTCAAGCATTGTTTTCGAAT-TTTTTCATGCATCAAAGCT

Chr3        CTTATAAACAAGATAGGGTATTCCTACATAGCTGGTGGGACCCACGGCAA
Chr4        CGTTAAGACTAGATGGGGTATCCCTACATGGCGGGTGGGACCCACGGCGA
Chr5        CGTTAAGACTAGATGGGGTATCCCTACATAGCGGGTGGGACCCACGGCGA
Campell     CGTTAAGACTAGATGGGGGATCCCTACATAGCGGGTGGGACCCACGGCGA

Chr3        ATGGATCATAAAGTCTTCAAAAAAGAATAGATACGATTGCATTGCATGTA
Chr4        ACGGTTCATCAAGACTTAAAAAAAGAATAGATACGATTGCATTGCATATA
Chr5        ATGGTTCATCAAGTCTTCAAAAAAGAATATATACGATTGCATTGCATATA
Campell     ATGGTTCATCAACTCTTCAAAAAAGAATATATACGATTGCATTGCATATA

Chr3        GTAAC
Chr4        CTAAC
Chr5        CTAAC
Campell     CTAAC
```

Supplemental Figure 2: Polymorphisms in Col-0 5S rRNA gene sequences

(A) Linkage mapping of the abundance of single nucleotide polymorphisms at positions 30, 41 and 96 estimated by NGS in 393 individuals of the MAGIC population with their corresponding boxplots of the estimated founder ecotype effect by multiple imputation using R/happy (3, 4) at the major quantitative trait locus (QTL) in each scan.
(B) Single nucleotide polymorphisms compared to the consensus 120 bp sequence identified in the 5S rRNA genes comprised in YAC 7E7 and 9A5 (chromosome 3), 5C3 and 9D3 (chromosomes 4) as well as 4E4 and 6A1 (chromosome 5).
(C) Alignment of the established consensus sequences for rRNA genes of chromosome 3, 4 and 5 derived from the Col-0 Mi-seq dataset with the reference sequence originally determined by Campell et al (5). Internal promoter sequences are colored in green, the TATA-box in blue.
(D) Linkage mapping of the abundance of the T to G single nucleotide polymorphisms at position -26 estimated by NGS in 393 individuals of the MAGIC population with the corresponding boxplots of the estimated founder ecotype effect by multiple imputation using R/happy (3, 4) at the major quantitative trait locus (QTL) in chromosome 3 and chromosome 4.

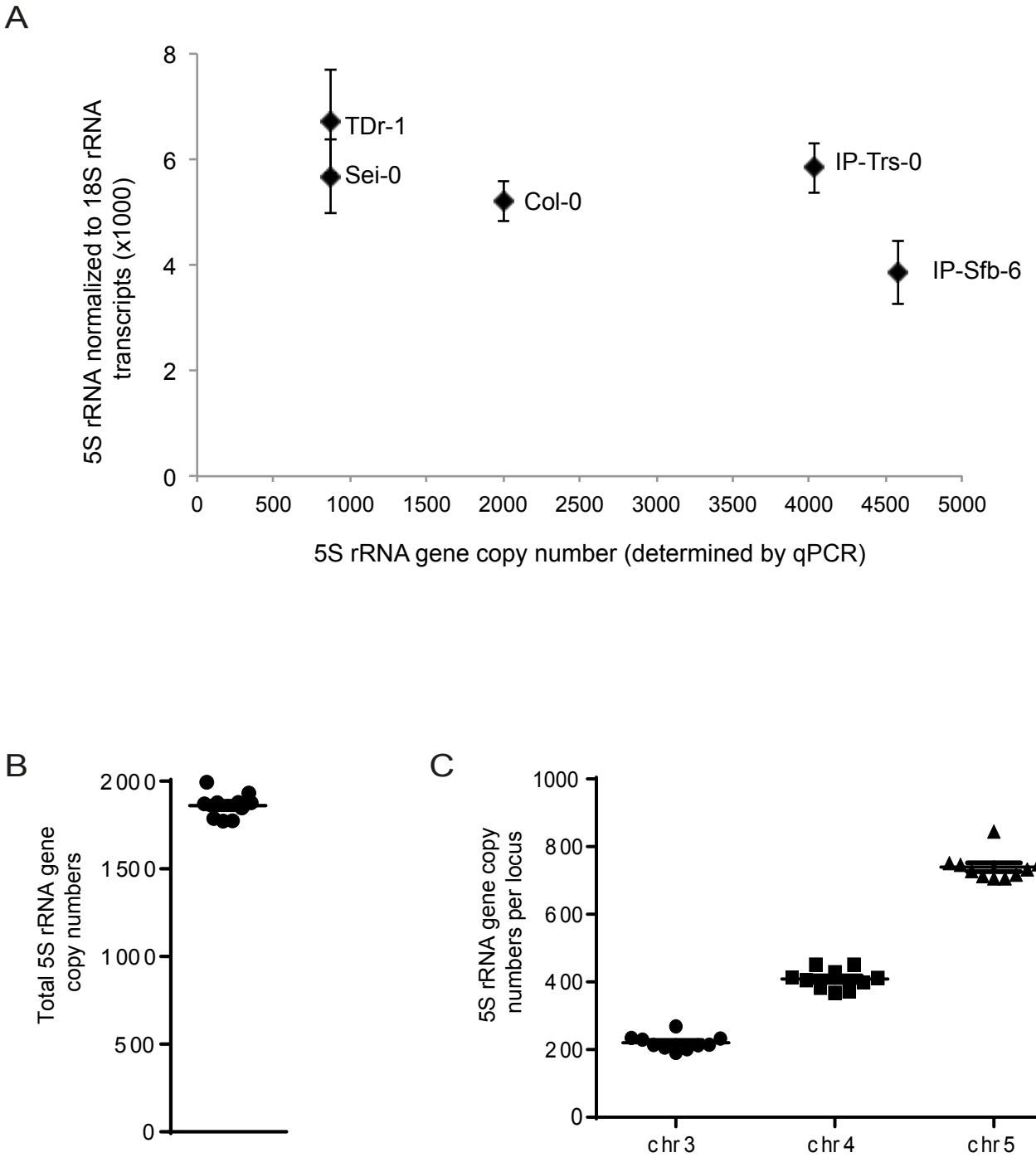
Supplemental Fig. 3



Supplemental Figure 3: Nucleosomal occupancy and enrichment in H3K4me3 at 5S rDNA and relative expression levels of *Ta3*, *HXK1* and *ACT2*

(A-B) Nucleosomal occupancy (MNase) and enrichment in H3K4me3 at (A) the transcriptionally active genes *HEXOKINASE1* (*HXK1*, *At4g29130*) and *ACTIN2* (*ACT2*, *At3g18780*), the retrotransposon *Ta3* and the 120 bp transcribed sequence of a 5S rRNA gene or (B) at the three different 5S rDNA loci. Enrichment was calculated by normalizing to the mean of three input datasets. Error bar corresponds to SEM for H3K4me3 for which two datasets were available.
(C) *Ta3*, *HXK1* and *ACT2* transcript levels normalized to *MON1* (*At2g28390*) from two RNA-seq datasets (6) of *A. thaliana* shoot tissue.

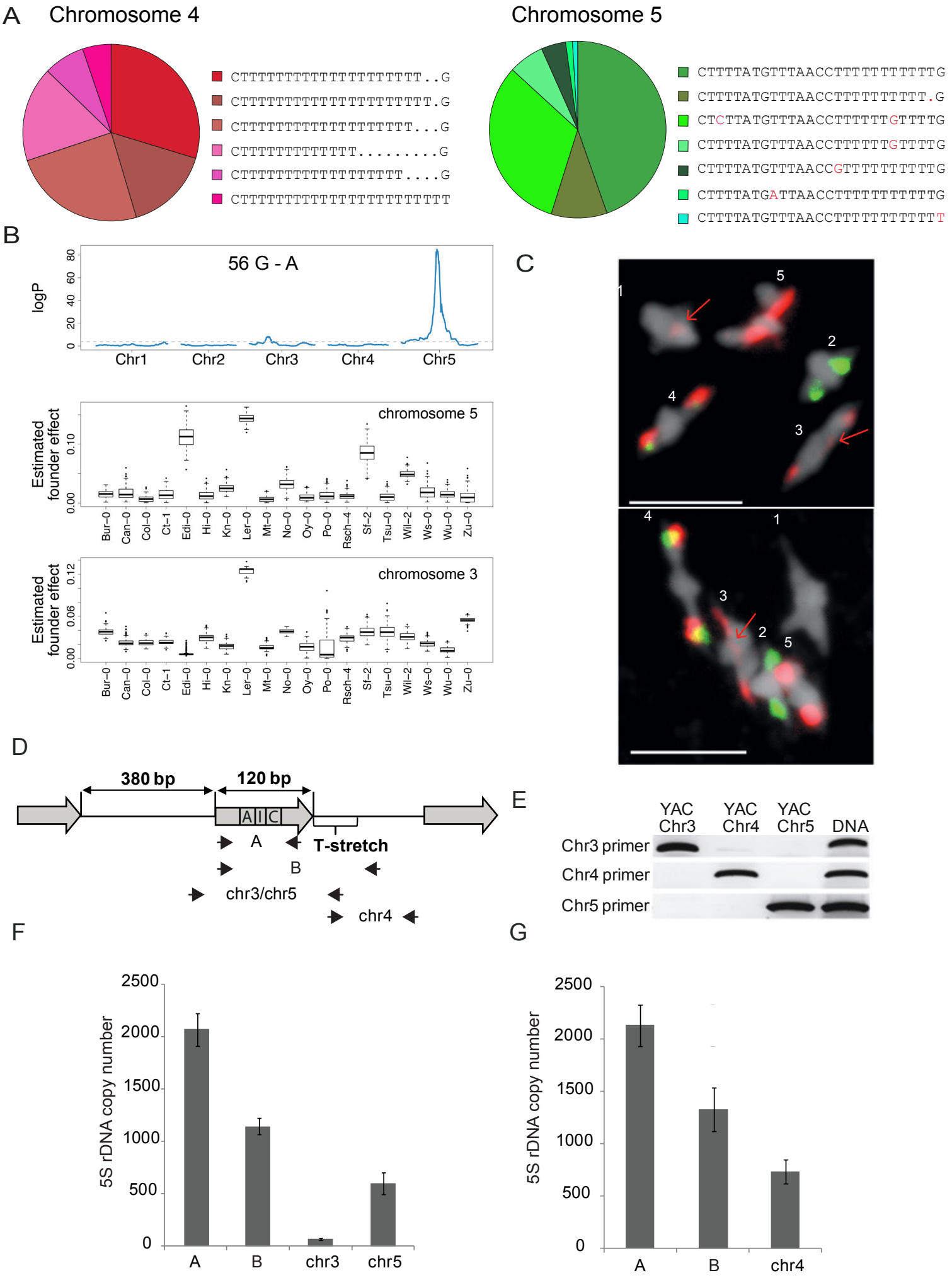
Supplemental Fig. 4



Supplemental Figure 4: 5S rRNA expression in different ecotypes and 5S rRNA gene copy number in Col-0 single seed descendants

(A) 5S rRNA expression normalized to 18S rRNA in 5 Arabidopsis ecotypes with different 5S rRNA gene copy numbers. Each data point depicts mean transcript levels of at least 5 biological replicates pooled from 2 independent experiments. Error bar represents SEM. (B-C) Total 5S rRNA gene copy number (B) and copies with the chromosome 3, 4 or 5 specific T-stretch signatures (C) as determined in Illumina datasets for 10 individual single seed descendants in the 30th generation of the mutation accumulation (MA) lines.

Supplemental Fig. 5



Supplemental Figure 5: 5S rDNA polymorphisms and variations in the L. *erecta* ecotype and in *dcl* mutants and characterization of locus-specific PCR primer sets

(A) Identified T-stretch signatures and their relative percentages per chromosome in the L*er* genome. The first T-stretches in the list are identical to those found in Col-0 and appear in identical color code. T-stretches specific for chromosome 4 and 5 in L*er* are in pink and light green respectively.

(B) (Top) Linkage mapping of the abundance of G to A single nucleotide polymorphism at position 56 estimated by NGS in 393 individuals of the MAGIC population. (Middle) Boxplot of the estimated founder accession effect by multiple imputation using R/happy at the major QTL in chromosome 5. (Bottom) Similar to middle panel, but for the major QTL in chromosome 3.

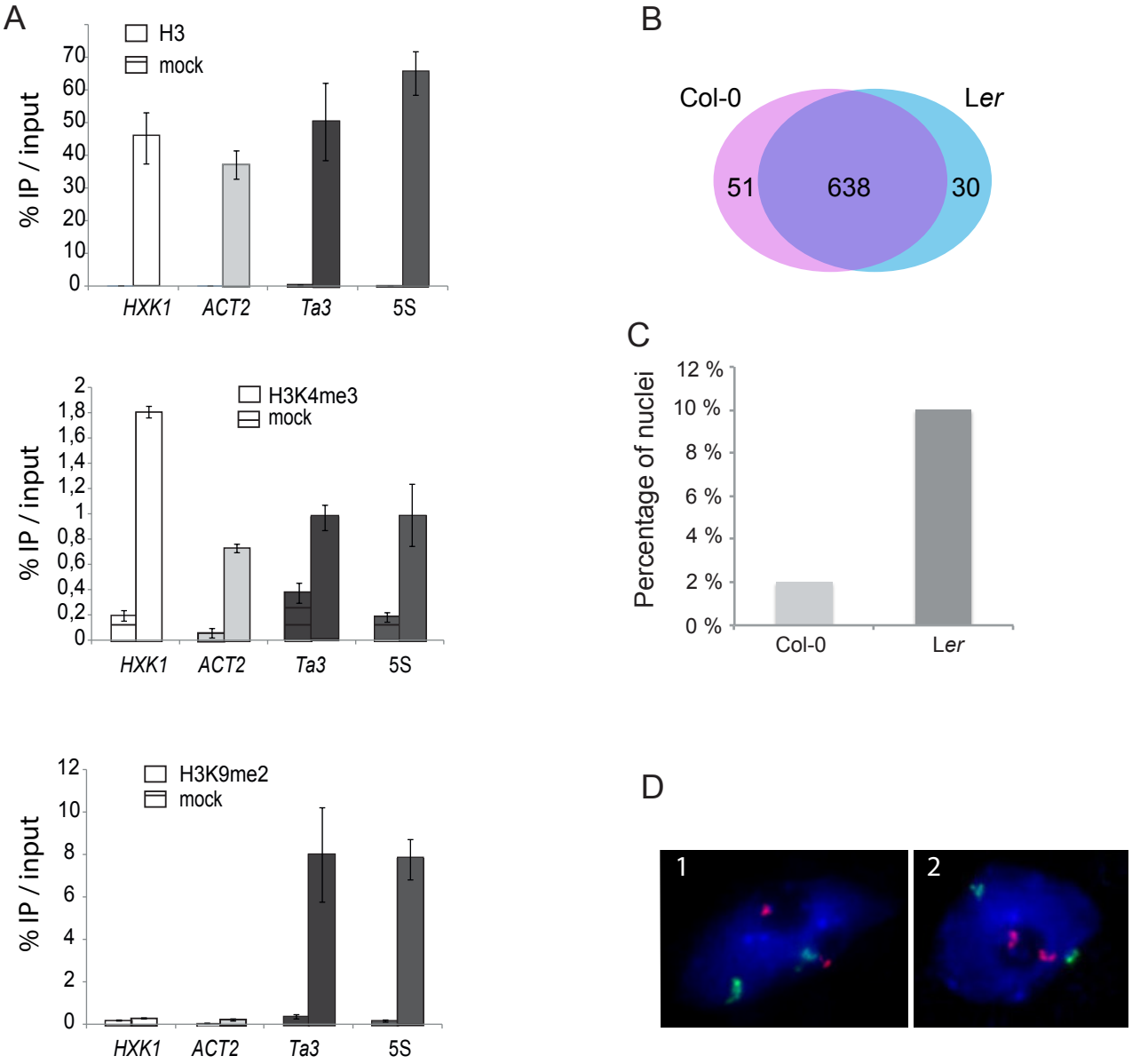
(C) FISH analysis using 5S (red) and 45S (green) probes on metaphase chromosomes of *dcl2-2* and *dcl3-1* mutants. Red arrows indicate additional 5S loci on chromosome 1 and 3 in *dcl2-2* and on chromosome 3 in *dcl3-1* mutant plants.

(D) Schematic representation of two 5S rRNA gene units and positions of the different primer sets used.

(E) Specificity of the chromosome-specific primer pairs was confirmed by PCR on DNA from Yeast Artificial Chromosomes containing either the loci of chromosomes 3, 4 or 5 and genomic DNA as control.

(F, G) Number of 5S rDNA copies quantified by qPCR using different primer pair combinations. Mean values of 5S rDNA copy numbers determined from three biological replicates and normalized to two single copy genes (*HXK1*, *At4g29130* and *UEV1C*, *At2g3660*). Different qPCR conditions have been used for the chromosome 3 and 5 (F) versus chromosome 4 specific primers (G) (see Material and Methods). Note that primer sets that cover a larger region and are not confined to the highly conserved transcribed sequence underestimate 5S rRNA gene copy numbers.

Supplemental Fig. 6



Supplemental Figure 6: 5S rDNA loci organization in the nucleus and ChIP-qPCR analysis

(A) H3 occupancy and enrichment in H3K4me3 and H3K9me2 normalized to input, including IP levels in the no-antibody fraction (mock) for Col-0 at two single copy genes (*HXK1*, *At4g29130* and *UEV1C*, *At2g3660*), the retrotransposon *Ta3* and the 5S rDNA transcribed sequence globally confirming enrichments found in ChIP-seq datasets.
(B) Venn diagram showing common and specific SNPs in the 5S-120 bp RNA-seq datasets from Col-0 and L*er*.
(C) Percentage of nuclei, for which 5S rDNA loci with chromosome 4 and chromosome 5 specific T-stretch signatures cluster at the same chromocenter in Col-0 or L*er* nuclei (150~200 nuclei by genotype).
(D) Cotyledon interphase nuclei stained by FISH with locus-specific LNA/DNA mixmer probes. 5S rRNA gene copies from chromosome 4 are labeled in red and those from chromosome 5 in green. (1) Example for a nucleus for which the 5S rDNA loci from chromosome 4 and 5 are clustered at the same chromocenter. (2) Example for a nucleus in which are 5S rDNA loci are found at distinct chromocenters.

Supplemental Tables

Chromosome	Source		5S copy number	Total 5S copy number
Chromosome 3	BAC	F21A14	25	250
	BAC	F1C23	10	
	BAC	T15B3	1	
	BAC	F4M19	22	
	BAC	T28G19	133	
	BAC	T25F15	10	
	BAC	T18B3	38	
	BAC	F23H6	11	
Chromosome 4	BAC	F13J19	15	15
Chromosome 5	BAC	T26N4	9	108
	BAC	T3P01	1	
	BAC	T32B03	47	
	BAC	T25B21	51	
				373

Chromosome 3	YAC	7E7	92	287
	YAC	9A5	195	
Chromosome 4	YAC	5C3	52	92
	YAC	9D3	40	
Chromosome 5	YAC	4E4	279	416
	YAC	6A1	137	
				795

Supplemental Table 1: 5S rRNA genes in available BAC and YACs

Number of sequences with homology to 5S rRNA genes on chromosomes 3, 4 and 5 identified by Blast search of available sequences from Bacterial Artificial Chromosomes (BACs) and copy number estimation of 5S rRNA gene sequences comprised in the 6 different YACs analyzed by NGS.

Col-0	**Determination of copy number**		
	SRR1945757	The 1001 Genomes Consortium, 2016	Hi-seq
	ERR420402 - ERR420412	Hagmann et al., 2015	
	PRJNA369183	this study	Mi-seq
L*er*	**Calculation of loci ratios**		
	ERR031544	Gan et al, 2011	Hi-seq
	SRR352145	Lu et al, 2011	
	DRR003391	Abe et al, 2015	
	Determination of polymorphisms		
	SRR3166543	Zapata et al, 2016	Hi-seq
	SRR3157034		
	Determination of T-stretchs		
	DRR003391	Abe et al, 2015	Hi-seq

Supplemental Table 2: Genome sequencing datasets for Columbia-0 (Col-0) and Landsberg *erecta* (L*er*) used in this study.

ChIP-seq datasets	H3K4me3	SRR1964977	Brusslan et al, 2015
		SRR1964979	
	H3K36me3	SRR5364427	Wollmann et al, 2017
		SRR5364428	
	H3K9me2	SRR3087128	Bewick et al, 2016
		SRR1999292	Stroud et al, 2014
	H3	SRR1999291	
		SRR1005403	
	Input H3.1	SRR394081	Stroud et al, 2012
	Input H3.3	SRR394080	
	H3.3	SRR394078	
		SRR6207346	Lu et al, 2018
	MNase	SRR957780	Li et al, 2014
		SRR5298544	Lorković et al, 2017
	H2A.W.6	SRR5298545	
		SRR988544	Yelagandula et al, 2014
	Input	SRR988541	
RNA-seq datasets for quantification of *ACT2, HXK1* and *Ta3* expression	Shoot tissue from 16 days old plants	SRR1030234	Zhang et al., 2015
		SRR1030235	
RNA-Seq dataset for 5S-120 transcripts	2 day-old Col-0 and L*er* plantlets	PRJNA378941	this study
RNA-Seq dataset for 5S-210 transcripts	2 day-old Col-0 and L*er* plantlets	PRJNA369190	this study

Supplemental Table 3: ChIP-Seq and RNA-Seq datasets used in this paper.

Polymorphisms	5S-120 transcripts		DNA		
	%	STDEV	chr3 (%)	chr4 (%)	chr5 (%)
78 C-T	0.98	+/- 0,089	2.98	0.36	0.15
49 C-T	0.45	+/- 0,009	2.17	0.12	0.51
105 C-T	0.42	+/- 0,017	0.08	0.73	0.60
18 C-T	0.42	+/- 0,023	0.49	0.05	0.07
96 C-T	0.39	+/- 0,054	2.44	1.67	0.09
62 T-C	0.26	+/- 0,008	0.35	0.29	0.06
15 C-T	0.25	+/- 0,009	0.27	0.10	0.04
94 C-T	0.24	+/- 0,007	10.02	0.69	2.46
25 G-A	0.21	+/- 0,020	2.99	0.04	0.03
35 C-T	0.21	+/- 0,016	0.64	0.07	0.21

Supplemental Table 4: Polymorphisms in the Col-0 5S-120 rRNA pool compared to the abundance in DNA

Frequency of the 10 most abundant polymorphisms in the 5S-120 rRNA dataset compared to the prevalence of the respective polymorphism in the Mi-Seq reads corresponding to the three different 5S rRNA gene loci in Col-0. Color-code corresponds to abundance in the RNA-Seq dataset (red > blue). The most abundant polymorphisms in the RNA-seq datasets are not necessarily found most frequently at the DNA level.

Polymorphisms	5S-120 transcripts				
	Col-0		L*er*		Ratio: L*er*/Col
	%	STDEV	%	STDEV	
78 C-T	0.98	0.089	0.02	0.002	0.02
49 C-T	0.45	0.009	0.31	0.022	0.69
105 C-T	0.42	0.004	0.05	0.004	0.12
18 C-T	0.42	0.103	0.03	0.001	0.06
96 C-T	0.39	0.054	0.65	0.077	1.64
62 T-C	0.26	0.008	1.69	0.036	6.47
15 C-T	0.25	0.009	0.89	0.016	3.50
94 C-T	0.24	0.007	0.14	0.004	0.57
25 G-A	0.21	0.020	0.10	0.004	0.49
35 C-T	0.21	0.016	0.03	0.002	0.16
21 T-G	0.19	0.005	0.20	0.010	1.01
45 T-G	0.19	0.004	0.20	0.013	1.05
19 A-T	0.10	0.001	0.25	0.012	2.41
93 T-A	0.08	0.001	0.39	0.012	4.81
61 C-T	0.04	0.001	0.32	0.007	8.95
59 T-C	0.03	0.002	0.70	0.027	20.37
95 C-T	0.03	0.002	0.43	0.011	12.86
69 A-T	0.03	0.002	0.29	0.008	11.09
77 A-T	0.02	0.001	0.23	0.008	13.54
40 A-T	0.01	0.000	0.28	0.013	21.43

Supplemental Table 5: Comparison of polymorphisms in the 5S-120 rRNA pool between Col-0 and L*er*.

Percentages (≥ 0.2% in at least one of the two ecotypes) of different single nucleotide polymorphisms in the 5S-120 RNA pools from Col-0 and L*er*. Shown are mean percentage of 3 biological replicates and the corresponding standard deviations. The most abundant SNPs in the Col-0 ecotype are not necessarily frequent in the 5S rRNA pool of L*er*.

Name	Sequence	Purpose
A-For	GGATGCGATCATACCAG	**qPCR on DNA**
A-Rev	GGGAGGTCACCCATCCTAGT	
B-For	GGATGCGATCATACCAG	
B-Rev	CGAAAAGGTATCACATGCC	
chr3-For	GACCCAYGGCGAATGGTT	
chr3-Rev	ACCAAAAAAAAAAAGGTAAACAT	
chr4-For	AAAGTACWTWWAGGTCAAAATTTGGRG	
chr4-Rev	TCCCTCTTTTTTTTTTTTTTTTTTTTG	
chr5-For	GGACCCAGGACAAATGGA	
chr5-Rev	AARACAAAAAAAAAGAGAAAAACCG	
HXK1-For	AGGAGCTCGTCTCTCTGCTG	
HXK1-Rev	GCTCAAACAATCCACCATCC	
UEV1C-For	GGTGACTGAAATGTGAATTTGC	
UEV1C-Rev	ATGCAGCCATCTCCTTCTTC	
5S-210 bp-For	GGATGCGATCATACCAG	
5S-210 bp-Rev	CGAAAAGGTATCACATGCC	
18S qRT-PCR-For	CGTAGTTGAACCTTGGGATG	**RT-qPCR to quantify 5S-120 transcripts**
18S qRT-PCR Rev	CACGACCCGGCCAATTA	
5S qRT-PCR For	GGATGCGATCATACCAG	
5S qRT-PCR Rev	GGGAGGTCACCCATCCTAGT	
RT-210 for RNA-seq	AGTTATGGAGTCATAAAGTTTTAAC	**RT reaction for RNA-seq**
5S-210 RNA-Seq-For	GGATGCGATCATACCAG	
5S-210 RNA-Seq-Rev1	AGTTATGGAGTCATAAAGTTTTAAC	
5S-210 RNA-Seq-Rev2	AGTTATAGAGTCATAAAGTTTTAAC	
RT 5S-120 RNA-seq	CCTTGGCACCCGAGAATTCCAAGGGATGCAACACG	

Supplemental Table 6: Primer sequences used in this study.

Supplemental References

1. Cloix,C., Tutois,S., Yukawa,Y., Mathieu,O., Cuvillier,C., Espagnol,M.C., Picard,G. and Tourmente,S. (2002) Analysis of the 5S RNA pool in Arabidopsis thaliana: RNAs are heterogeneous and only two of the genomic 5S loci produce mature 5S RNA. *Genome Res.*, **12**, 132–144.
2. Poulet,A., Duc,C., Voisin,M., Desset,S., Tutois,S., Vanrobays,E., Benoit,M., Evans,D.E., Probst,A. V and Tatout,C. (2017) The LINC complex contributes to heterochromatin organisation and transcriptional gene silencing in plants. *J. Cell Sci.*, **130**, 590–601.
3. Mott,R., Talbot,C.J., Turri,M.G., Collins, a C. and Flint,J. (2000) A method for fine mapping quantitative trait loci in outbred animal stocks. *Proc. Natl. Acad. Sci. U. S. A.*, **97**, 12649–12654.
4. Kover,P.X., Valdar,W., Trakalo,J., Scarcelli,N., Ehrenreich,I.M., Purugganan,M.D., Durrant,C. and Mott,R. (2009) A multiparent advanced generation inter-cross to fine-map quantitative traits in Arabidopsis thaliana. *PLoS Genet.*, **5**.
5. Campell,B.R., Song,Y., Posch,T.E., Cullis,C.A. and Town,C.D. (1992) Sequence and organization of 5S ribosomal RNA-encoding genes of Arabidopsis thaliana. *Gene*, **112**, 225–228.
6. Zhang,X., Zhu,Y., Liu,X., Hong,X., Xu,Y., Zhu,P., Shen,Y., Wu,H., Ji,Y., Wen,X., *et al.* (2015) Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in Arabidopsis. *Science*, **348**, 120–3.
7. The 1001 Genomes Consortium,X. (2016) 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. *Cell*, **166**, 481–491.
8. Hagmann,J., Becker,C., Müller,J., Stegle,O., Meyer,R.C., Wang,G., Schneeberger,K., Fitz,J., Altmann,T., Bergelson,J., *et al.* (2015) Century-scale Methylome Stability in a Recently Diverged Arabidopsis thaliana Lineage. *PLoS Genet.*, **11**.
9. Gan,X., Stegle,O., Behr,J., Steffen,J.G., Drewe,P., Hildebrand,K.L., Lyngsoe,R., Schultheiss,S.J., Osborne,E.J., Sreedharan,V.T., *et al.* (2011) Multiple reference genomes and transcriptomes for Arabidopsis thaliana. *Nature*, **477**, 419–23.
10. Lu,P., Han,X., Qi,J., Yang,J., Wijeratne,A.J., Li,T. and Ma,H. (2012) Analysis of Arabidopsis genome-wide variations before and after meiosis and meiotic recombination by resequencing Landsberg erecta and all four products of a single meiosis. *Genome Res.*, **22**, 508–18.
11. Abe,M., Kaya,H., Watanabe-Taneda,A., Shibuta,M., Yamaguchi,A., Sakamoto,T., Kurata,T., Ausín,I., Araki,T. and Alonso-Blanco,C. (2015) FE, a phloem-specific Myb-related protein, promotes flowering through transcriptional activation of FLOWERING LOCUS T and FLOWERING LOCUS T INTERACTING PROTEIN 1. *Plant J.*, **83**, 1059–68.
12. Zapata,L., Ding,J., Willing,E.-M., Hartwig,B., Bezdan,D., Jiao,W.-B., Patel,V., Velikkakam James,G., Koornneef,M., Ossowski,S., *et al.* (2016) Chromosome-level assembly of Arabidopsis thaliana Ler reveals the extent of translocation and inversion polymorphisms. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, E4052-60.
13. Brusslan,J.A., Bonora,G., Rus-Canterbury,A.M., Tariq,F., Jaroszewicz,A. and Pellegrini,M. (2015) A Genome-Wide Chronological Study of Gene Expression and Two Histone Modifications, H3K4me3 and H3K9ac, during Developmental Leaf Senescence. *Plant Physiol.*, **168**, 1246–61.
14. Wollmann,H., Stroud,H., Yelagandula,R., Tarutani,Y., Jiang,D., Jing,L., Jamge,B., Takeuchi,H., Holec,S., Nie,X., *et al.* (2017) The histone H3 variant H3.3 regulates gene body DNA methylation in Arabidopsis thaliana. *Genome Biol.*, **18**, 94.
15. Bewick,A.J., Ji,L., Niederhuth,C.E., Willing,E.-M., Hofmeister,B.T., Shi,X., Wang,L., Lu,Z., Rohr,N.A., Hartwig,B., *et al.* (2016) On the origin and evolutionary consequences of gene body DNA methylation. *Proc. Natl. Acad. Sci.*, **113**, 9111–9116.
16. Stroud,H., Do,T., Du,J., Zhong,X., Feng,S., Johnson,L., Patel,D.J. and Jacobsen,S.E. (2014) Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. *Nat. Struct. Mol. Biol.*, **21**, 64–72.

17. Stroud,H., Hale,C.J., Feng,S., Caro,E., Jacob,Y., Michaels,S.D. and Jacobsen,S.E. (2012) DNA methyltransferases are required to induce heterochromatic re-replication in Arabidopsis. *PLoS Genet.*, **8**, e1002808.
18. Lu,L., Chen,X., Qian,S. and Zhong,X. (2018) The plant-specific histone residue Phe41 is important for genome-wide H3.1 distribution. *Nat. Commun.*, **9**, 630.
19. Li,G., Liu,S., Wang,J., He,J., Huang,H., Zhang,Y. and Xu,L. (2014) ISWI proteins participate in the genome-wide nucleosome distribution in Arabidopsis. *Plant J.*, **78**, 706–14.
20. Lorković,Z.J., Park,C., Goiser,M., Jiang,D., Kurzbauer,M.-T., Schlögelhofer,P. and Berger,F. (2017) Compartmentalization of DNA Damage Response between Heterochromatin and Euchromatin Is Mediated by Distinct H2A Histone Variants. *Curr. Biol.*, **27**, 1192–1199.
21. Yelagandula,R., Stroud,H., Holec,S., Zhou,K., Feng,S., Zhong,X., Muthurajan,U.M., Nie,X., Kawashima,T., Groth,M., *et al.* (2014) The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in Arabidopsis. *Cell*, **158**, 98–109.
22. Zhang,X., Zhu,Y., Liu,X., Hong,X., Xu,Y., Zhu,P., Shen,Y., Wu,H., Ji,Y., Wen,X., *et al.* (2015) Plant biology. Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in Arabidopsis. *Science*, **348**, 120–3.