Polyploidy-associated paramutation in Arabidopsis is determined by small RNAs, temperature, and structure

Heinrich Bente[1], Andrea M. Foerster[1], Alexis Sarazin[2], Nicole Lettner[1], Ortrun Mittelsten Scheid[1]

[1]Gregor Mendel Institute of Molecular Plant Biology (GMI), Austrian Academy of Sciences, Vienna Biocenter (VBC), 1030 Vienna, Austria

[2]Department of Biology, ETH Zürich, Zurich, Switzerland.

ABSTRACT

Paramutation is a case of non-Mendelian inheritance, long known, but still incompletely understood. It describes the change of gene expression at a paramutable allele by encounter with a paramutagenic allele, causing an alteration that remains even after segregation of both alleles. While discovery, genetic and mechanistic studies of paramutation were made with alleles that change plant pigmentation, paramutation-like phenomena are also known for other traits and in other organisms, and many cases probably went undetected. Paramutation between epialleles, genes with identical sequences but different expression states, argues for an epigenetic mechanism, providing profound evidence that triggered epigenetic changes can be maintained over generations. Here, we use a case of paramutation that affects a transgenic selection reporter gene in tetraploid *Arabidopsis thaliana*, to study the role of small RNAs, epiallele structure and copy number, and environmental influence on degree and timing of paramutation. Our data suggest that different types of small RNA associated with the epialleles, the presence of repeats, and the growth temperature of the epiallelic hybrids contribute to the strength of the allelic interaction. The data further make it plausible why paramutation in this system becomes evident only in the segregating F2 population of tetraploid hybrids. In summary, the results support a model for polyploidy-associated paramutation, with similarities as well as distinctions from other cases of paramutation.

Keywords: Paramutation, epialleles, small RNAs, polyploidy, RNA-directed DNA methylation, Arabidopsis thaliana

INTRODUCTION

Paramutation is the term for a specific interaction between genes, in which one paramutable allele becomes heritably modified by the encounter with a paramutagenic allele. It violates the Mendelian rule of independent segregation, as the paramutated allele is maintained in the new state and even acquires paramutagenicity itself. This stimulated the description as para-(resembling)-mutation or somatic gene conversion by the pioneers working on maize and tomato in the 1950th (Brink, 1958; Hagemann, 1958; Coe, 1959; Szittya et al., 2003). Even earlier genetic experiments (Bateson and Pellew, 1915), curiously with Mendel's main experimental plant pea, are now seen as evidence for paramutation. Numerous cases of paramutation have been reported afterwards in several species including non-vertebrates, animals and plants (reviewed in Gabriel and Hollick, 2015; Pilu, 2015; Ronsseray, 2015), indicating that paramutation is not just a bizarre exception.

Paramutation occurs on an epigenetic basis without changing the DNA sequence of participating alleles and demonstrates that specific, newly acquired epigenetic states can be installed and maintained independent from the trigger of the change. It is likely the best example for true transgenerational inheritance of epigenetic changes that are prompted by a defined setting. Although many principles of epigenetic regulation are now apparent, the molecular basis of paramutation is still not well understood. Genetic screens, mainly in maize, have identified several genes whose function is necessary for establishing and/or maintenance of paramutation, e.g. *"Required to Maintain Repression"* (*RMR*) (Hollick and Chandler, 2001) or *"Mediator Of Paramutation"* (*MOP*) (Dorweiler et al., 2000). The nature of their gene products (reviewed in Giacopelli and Hollick, 2015) suggested that paramutation requires components of the RNA-directed DNA methylation (RdDM) pathway (reviewed in Law and Jacobsen, 2010; Cuerda-Gil and Slotkin, 2016), mainly responsible for silencing repetitive sequences via small RNAs (sRNAs) guiding the RNA-induced silencing complex (RISC) to corresponding genomic loci and installing DNA methylation and heterochromatin. However, the vast majority of RdDM targets does not exert paramutation, suggesting that additional factors, such as *cis*-acting determinants (Hovel et al., 2015) or other *trans*-acting elements are involved. To identify these, additional paramutation systems might help, especially those that differ in penetrance, frequency, and stability of the epigenetic change.

A previously described case of paramutation in *Arabidopsis thaliana* (Mittelsten Scheid et al., 2003) involves epialleles of a transgene that, in an active state, expresses hygromycin phosphotransferase (HPT) conferring resistance (R epiallele) during growth on hygromycin-containing growth medium. *In vitro* culture of protoplasts and regeneration resulted in several tetraploid plants, some hygromycin-resistant, others in which a silent epiallele had occurred spontaneously. This acquired DNA methylation and heterochromatic histone marks (Hetzl et

al., 2007; Baubec et al., 2010), did not transcribe the gene anymore, and rendered plants hygromycin-sensitive (S epiallele). Diploid R and S derivatives were generated by backcrosses, and progeny plants homozygous for either R or S alleles inherited their respective expression state and phenotype over numerous generations, in diploid and tetraploid lines. Molecular analysis confirmed that R and S alleles are completely isogenic, at the same genomic location, inserted in an intergenic region far from transposons, and only differ in their epigenetic state, therefore are true epialleles (Foerster et al., 2011). A mutant screen scoring for reactivation of the silent epiallele resulted in *trans*-acting mutants confirming the importance of DNA and histone methylation (Baubec et al., 2010), and in *cis*-acting mutations with structural variants of the epiallele (RΔ) that led to restored transcription and hygromycin resistance (Foerster et al., 2011).

Crossing diploid R and S plants resulted in 100% hygromycin-resistant F1 plants, and selfed F1 plants gave F2 progeny with close to 75% (3:1) resistant plants, as expected for a dominant trait and confirming independent segregation. Also F1 hybrids from crosses between tetraploid R and S plants were 100% resistant, but their F2 progeny contained significantly less than the expected 97% (35:1) resistant individuals, although with variation between independent tetraploid F2 populations. This argued against a genetic determinant of the lost resistance and suggested epigenetic components interfering with the independence of segregation, similar to the described phenomenon of paramutation (Mittelsten Scheid et al., 2003). However, there were interesting differences to the classical paramutation cases. While the loss of gene expression was visible in maize and tomato, both diploids, already in the F1 hybrid (Patterson and Chandler, 1995), the phenotype indicating paramutation in the R/S interaction was restricted to the F2 generation of tetraploid hybrids only. This late onset of the interaction and its association with polyploidy, the easily quantifiable phenotype, and the toolbox available for the model plant Arabidopsis, stimulated the study presented here.

We investigated the role of the epiallele structure, the connection with transcriptional activity, and the potential for secondary paramutation. We examined the effect of temperature during growth of the F1 hybrids, and we profiled the small RNAs associated with the epialleles in seedlings and flower buds of diploid and tetraploid plants. We show that structure, nature of the small RNAs, environmental conditions, and developmental stage act together to determine the amount of paramutation and present a model that explains the role of copy number in the epiallelic interaction in polyploid plants.

RESULTS

Nature and nomenclature of the epialleles involved in paramutation

The phenotype revealing paramutation in Arabidopsis is based on expression or silencing of the hygromycin phosphotransferase (HPT) gene conferring resistance to the antibiotic hygromycin. A construct with HPT under control of the CaMV 35S promoter was used for PEG-mediated protoplast transformation (Karesch et al., 1991) and became inserted in an intergenic region of chromosome three in the Arabidopsis genome. As often in direct gene transfer, plasmid copies recombine prior to integration, in this case resulting in an insert consisting of a vector fragment upstream of the promoter and the HPT open reading frame, a partially deleted CaMV 35S terminator followed by a second copy of the vector and the CaMV 35S promoter, and approximately 500 bp non-coding non-plant carrier DNA used during the protoplast transformation. This structure has been already described (Mittelsten Scheid et al., 2003; Foerster et al., 2011); a schematic illustration of all used variants (considering the transcription direction opposite to the reference genome) is shown in Fig 1A. Transcribed and silenced epialleles have identical DNA sequence and are referred to as R (resistant) or S (sensitive) in the following. The “empty” insertion site in the wild type will be termed W. Diploid and tetraploid genotypes will be called RR/SS/WW and RRRR/SSSS/WWWW, respectively. An overview about origin and crosses of each epiallele and respective mutants is illustrated in Fig 1B.

Paramutation depends on temperature during growth of tetraploid F1 hybrids

As described, plants homozygous for either R or S inherit the respective hygromycin resistance or sensitivity as stable trait. Hybrids obtained by crossing diploid plants and selfing them resulted in F2 progeny with the expected 3:1 ratio, regardless whether R was combined with S or with the wild type lacking an epiallele (RS or RW). In contrast, paramutation occurs in RRSS hybrids and is apparent by a reduced number of resistant F2 plants, significantly lower than in the control crosses of R with W (RRWW), independent from maternal or paternal contribution of the S allele, but with variation between F2 populations from different F1 parents (Mittelsten Scheid et al. 2003). This variation raised the question whether external factors during the growth of the F1 hybrids would influence the penetrance of paramutation, and we assumed temperature as an important parameter, as its influence on a paramutation-resembling inheritance in Drosophila had been reported (Ciabrelli et al., 2017). Therefore, we grew diploid and tetraploid F1 plants obtained from crosses between R and S (or W) at 10°C, 19°C, or 24°C during their flowering and seed set period (Fig 2A). We germinated F2 seeds on hygromycin selection medium and scored the ratio of resistant seedlings (Fig 2B and 2C). F2 from tetraploid RRSS hybrids grown at 19°C (the standard growth conditions) confirmed the results from earlier experiments by a lower ratio of resistant plantlets than the RRWW

controls. This difference was even more significant in the progeny from parents grown at 24°C, indicating an enhanced interaction between R and S. Surprisingly, progeny from RRSS and RRWW plants grown at 10°C had a similar ratio of resistant to sensitive plants, indicating that the presence of the S epiallele did not lead to reduced activity of R at the lower temperature (Fig 2B). All F2 from diploid hybrids contained approximately 75% of resistant seedlings, excluding any influence of the S epiallele under all tested temperature conditions (Fig 2C). These results confirm that paramutation in this system is associated with tetraploidy and indicate that its penetrance is indeed affected by external factors, exemplified here by temperature.

Epialleles involved in Arabidopsis paramutation have different sRNA signatures

Multiple evidence indicates that *trans*-silencing phenomena, such as paramutation, are associated with small RNAs (sRNA). Indeed, the presence of sRNAs at loci participating in paramutation has been shown for genes in maize (Chandler et al., 2000; Hollick et al., 2000), tomato (Gouil and Baulcombe, 2018), and Arabidopsis (Jiang et al., 2019), including a very limited analysis for the silenced epiallele S of this study (Foerster et al., 2011). To address the possible mechanistic role of sRNA, we generated strand-specific sRNA libraries from total RNA of 14 d-old seedlings of tetraploid and diploid homozygous R and S lines and mapped reads with a length from 18-26 nt to the Arabidopsis TAIR10 reference genome extended *in silico* for the inserted HPT epiallele. Whereas size distribution of sRNAs mapping to the whole genome were similar in all libraries, there was a striking difference regarding the size classes between the epialleles: we found mainly 21 nt sRNAs at the active and mostly 24 nt sRNAs at the silenced epiallele in tetraploid plants (Fig 3A), and a similar profile in diploids (S1A Fig). Both size classes mapped predominantly to the duplicated region (promoter and vector sequence) of the epiallele (Fig 3B, S1B Fig), so that these sRNAs may originate from either copy. Libraries prepared from wild type plants do not contain reads mapping to the epiallele (S2B Fig), confirming the epiallelic origin of the 21/24 nt reads in Fig 3B. Abundance of the 24 nt sRNAs was comparable between diploid and tetraploid S lines, and they were distributed along the whole duplicated promoter/vector region. In contrast, tetraploid R lines had slightly elevated amounts of 21 nt sRNA, especially those mapping to a specific region of the promoter. The HPT-ORF and the non-coding region downstream of the second promoter copy were almost devoid of sRNAs in libraries from all samples (Fig 3B, S1B Fig).

The frequency of the four bases at the 5' of sRNAs correlates with their length and serves as an indicator for its potential association with specific Argonaute (AGO) proteins (Mi et al., 2008). In all lines, sRNAs mapping to the wild type genome confirmed the known preference for either 5'U in 21 nt sRNAs and 5'A for 24 nt sRNAs (S2A Fig). The 5' base distribution of epiallelic sRNAs in both tetraploids (Fig 3C) and diploids (Fig S1C) is coherent: 5' nucleotides

of the 21 nt sRNAs mapping to the R epiallele were biased for 5'U, indicating a preference for AGO1, while half of the 24 nt sRNAs from the S epiallele had an A at the 5' end, suggesting association with either AGO4 (Mi et al., 2008) or AGO3 (Zhang et al., 2016).

sRNA signature and mutant analysis support involvement of RdDM

The relatively high percentage of 5'A in the 24 nt sRNAs from the S epiallele (Fig 3C and S1C Fig) added evidence to the involvement of RNA-directed DNA methylation (RdDM) (Wassenegger et al., 1994; Matzke and Mosher, 2014) in paramutation, as suggested by several mutational studies in maize (Hollick 2016). The limitation of the R/S interaction to tetraploid plants made any forward-directed screen for recessive mutations in this system impractical, but the efficiency of CRISPR-Cas9- based mutagenesis in Arabidopsis (Fauser et al., 2014) allowed to challenge the role of RdDM with a reverse genetic approach. Paramutation between the B-I and B' allele in maize requires functional RDR2 (Stonaker et al., 2009) and several subunits of the plant-specific DNA-dependent RNA polymerases IV and V (Erhard et al., 2009; Sidorenko et al., 2009; Sloan et al., 2014). Therefore, we targeted two corresponding Arabidopsis genes (*RDR2*, At4g11130) and the largest subunit of PolIV (*NRPD1*, At1g63020) in tetraploid R and S plants with Cas9 and suitable sgRNAs and selected lines homozygous for the mutations and absence of the Cas9 gene by genotyping. Loss of function was validated by the absence of endogenous 24 nt small interfering RNAs (S3A Fig). Furthermore, the mutation in *NRPD1* led to depletion of 24 nt sRNAs in SSSS (S3B Fig). Unlike the *NRPD1* mutant *rmr7* in maize that reactivated the silent B' allele (Stonaker et al., 2009), the newly generated Arabidopsis *cas9nrpd1* mutant did not lead to reactivation of the S epiallele. We tested the potential for paramutation by generating F1 hybrids homozygous for the mutation and scoring F2 seedlings for hygromycin resistance, together with progeny from corresponding wild type combinations grown in parallel. The values for mutant progeny of RRSS hybrids were much closer to that of RRWW crosses (Fig 4A) than in wild type background and there were less sensitive seedlings (Fig 4B), indicating reduced establishment of paramutation under impaired RdDM, but no role in the maintenance of the silent state of the S epiallele.

The paramutation effect can be mimicked by ectopic expression of sRNAs homologous to various regions of the epialleles

As the results from the mutant analysis suggested a role of the sRNAs in establishing paramutation, we asked whether sRNAs would be sufficient to silence the otherwise remarkably stable R epiallele, and whether it required allelic positioning between the epialleles. We designed hairpin (hp) constructs that would induce strong sRNA formation targeting three different regions within the duplicated vector/promoter region (hp1/2/3; Fig 1A) and transformed RR and RRRR plants. Seeds from individual T1 plants were separated for

presence or absence of the hairpin construct via seed-specific GFP expression encoded in the vector and grown with or without hygromycin. Under selection, all hairpin variants caused reduced root growth, compared to plants without hairpins (Fig 5A and 5B), indicating silencing of the R allele in diploid and tetraploid lines. The degree of silencing was highest with hp3 homologous to the promoter, but even hp1 and hp2, both matching to the vector sequence further away from the promoter within the epiallele resulted in significant growth inhibition under selection (Fig 5A). To quantify the effect of the hairpins at a later stage and directly at the transcriptional level, we determined HPT expression by qRT-PCR in flower bud samples of diploid and tetraploid T1 plants with and without hairpins selected as above. In diploid plants, all three hairpin variants reduced HPT expression to a similar level (Fig 5C left panel), whereas in tetraploid plants the effect increases with proximity of the hp-matching region to the transcription start site of the HPT gene (Fig 5C right panel). This might reflect different stochiometry: average expression of the hairpin can be expected to be similar across many independent T1 lines, but the processed sRNAs target four rather than only two R copies in tetraploid or diploid lines, respectively.

New silenced epialleles can exert secondary paramutation

The epialleles do not contain any inverted repeat that could form hairpins as a source of sRNAs, but also tandem-oriented repeats can affect paramutation, depending on copy number (Panavas et al., 1999; Hovel et al., 2015). Evidence for a role of the duplicated vector/promoter region in paramutation came from the previously described structural variants (RΔ, derived from the S epiallele by mutagenesis), in which rearrangements downstream of the HPT-ORF led to reactivation of the upstream promoter (Foerster et al., 2011). In RΔ2, the whole transgenic sequence downstream of the defective terminator (including the vector/promoter repeat) was deleted, whereas RΔ4 lacked only the terminal non-coding carrier sequence, maintaining the repeat (Fig 1A). Tetraploid, homozygous derivatives of both lines (RΔ2RΔ2RΔ2RΔ2 and RΔ4RΔ4RΔ4RΔ4, selected after colchicine treatment) were crossed with SSSS or WWWW lines and the F2 progenies screened for hygromycin resistance. Remarkably, only F2 offspring from reciprocal crosses between RΔ4RΔ4RΔ4RΔ4 x SSSS (RΔ4RΔ4SS), but not from RΔ2RΔ2RΔ2RΔ2 x WWWW (RΔ2RΔ2WW) crosses, showed reduced hygromycin resistance as in case of the full length epiallelic crosses RRRR x SSSS (Fig 6A), classifying RΔ2 as a neutral and RΔ4 as still paramutable allele.

The most persuasive evidence for paramutation is the acquired paramutagenicity of a paramutated allele, i.e. the ability to exert secondary paramutation. In the initial paramutation cross RRRR x SSSS, paramutated R alleles do not have any molecular marker that allows distinguishing them from S alleles in the segregating F2 due to their identical sequence. Paramutation of the RΔ4 allele provided this marker necessary to follow the converted allele

by PCR-based screening for the deletion. We crossed RΔ4RΔ4RΔ4RΔ4 with SSSS and selected individual F2 plants that were either homozygous (+/+) or heterozygous (+/-) for the deletion and quantified HPT expression with qRT-PCR. As expected from previous studies and in accordance with the hygromycin resistance assay, RRRR and RΔ4RΔ4RΔ4RΔ4 showed strong expression, whereas SSSS and WWWW did not result in any detectable transcripts. Moreover, only progeny from RΔ4RΔ4WW hybrid parents expressed HPT in comparable amounts (Fig 6B). In contrast, progeny from RΔ4RΔ4SS parents had much less HPT transcript, including those homozygous for the previously active RΔ4 allele (Fig 6B, black star), indicating strong suppression by the presence of the S allele in the previous generation and lasting silencing after segregation. Therefore, in a line with four RΔ4 allele, they all had undergone paramutation, changing it into SΔ4SΔ4SΔ4SΔ4 (Fig 1B), and we used this to assay secondary paramutation. Tetraploid hybrids from crosses between this line and that with the initial active R alleles (Fig 1B) resulted in F2 populations with a similarly reduced hygromycin resistance as progeny from the deletion-activated RΔ4 and the initial silenced S alleles (Fig 6C). This implies that indeed, a once paramutated HPT allele could reduce HPT expression from a naïve R allele in the process of secondary paramutation.

Removal of the downstream duplicate does not reactivate a silent upstream promoter but boosts expression from active epialleles.

RΔ2 and RΔ4 alleles had been identified in a forward genetic screen selecting for reactivation of the S epiallele (Foerster et al., 2011). Both deletions resulted in hygromycin resistance, but RΔ2 lost paramutagenicity while RΔ4 maintained this feature. The different, randomly induced structural changes (Fig 1A) suggested that the removal of the duplicated downstream region in RΔ2, but not in RΔ4, might explain the difference. Thus, we decided to apply CRISPR gene editing to recapitulate the deletion in RΔ2 precisely, eliminating the downstream repeat as well as the non-coding sequence from all alleles. Cas9-encoding transgenes with suitable sgRNAs were transformed into diploid and tetraploid R and S plants but segregated out based on genotyping during the selfing towards homozygosity of the deletions. To distinguish alleles with Cas9-induced deletions from the previously described ones, they are denoted as Rγ and Sγ (Fig 1A). We screened diploid (RγRγ, SγSγ) and tetraploid (RγRγRγRγ, SγSγSγSγ) plants for hygromycin resistance. Despite the targeted deletions in SγSγ and SγSγSγSγ, plants remained completely sensitive and did not have detectable HPT transcripts, in contrast to the deletions after random mutagenesis that were identified by a strict selection screen (Foerster et al., 2011). Asking whether Sγ would still participate in paramutation, we performed reciprocal crosses between tetraploids with full length or deleted R and S alleles and screened F2 segregation under hygromycin selection. Resistance in progeny from RRSγSγ hybrids was higher than for RRSS parents (Fig 7A), indicating reduced paramutagenicity of S without the repeat. Interestingly, the full-length S,

paramutagenic in RRSS, was not anymore paramutagenic in RγRγSS and RγRγSγSγ (Fig 7A and 7B). We therefore scored growth on hygromycin and found longer roots of RγRγRγRγ compared to its full-length progenitor RRRR, confirming more pronounced hygromycin resistance of RγRγRγRγ (Fig 7C and 7D). Northern blots confirmed a shorter transcript of ~1.6 kb according to the deletion with ~3-4x more signal strength than in R plants, in diploids and tetraploids (Fig 7E), in accordance with the enhanced resistance to hygromycin and with the effect of the RΔ alleles (Foerster et al., 2011). sRNA libraries of γ and Δ deletion lines revealed similar profiles along S-derived alleles but differed substantially from full length R, as loss of the duplicated region led to depletion of 21 nt sRNAs (S4 Fig). This suggests that, despite pronounced transcription and stable hygromycin resistance from full length R epialleles, the associated 21 nt RNAs contribute to partial suppression of the upstream promoter, while their absence boosts expression and renders this allele less or not at all sensitive to paramutation by an S epiallele.

DISCUSSION

The term paramutation was created to specify heritable alterations, while distinguishing them from genetic events. Long before mechanistic aspects were investigated, common features of several cases were recognized: (1) a specific combination of alleles, (2) a directed switch of activity for one of the alleles, (3) stability of the switch even after genetic segregation, and (4) the newly acquired paramutagenicity of the converted allele. The epialleles of the HPT transgene used in this study had been shown previously to match condition 1 – 3 ((Mittelsten Scheid et al., 2003; Baubec et al., 2010; Foerster et al., 2011), while the switch from paramutability to paramutagenicity could previously not be documented due the impossibility to discriminate genetically identical epialleles after they were combined in hybrids. The identification of a structural variant that allowed molecular genotyping in a segregating population has now been exploited to also demonstrate secondary paramutation, adding the Arabidopsis system to the series of classical paramutation examples.

Despite the classical features mentioned above, each identified paramutation case has its own particularity. Our system is different from most other described cases, in (1) the lack of apparent paramutation in F1 hybrids, and (2) the occurrence of paramutation only in tetraploid plants. With the model described here, both features are plausible. The potentially strong HPT expression from the upstream Cauliflower Mosaic Virus 35S promoter is weakened in the active epiallele by the formation of the 21 nt sRNA connected with the downstream repeat, but still strong enough to allow sufficient HPT protein production to detoxify hygromycin. Regular F2 resistance segregation after crosses with the “empty” wild type imply that one active allele, even with the repeat, is enough to render a plant resistant, regardless whether it is 1 out of 2 (diploid) or 1 out of 4 (tetraploid). If the cross introduced the inactive epiallele, HPT expression is still maintained high enough to allow resistance as long as the associated 24 nt sRNA are present in limited amounts. The ectopic overexpression of sRNA varieties from the hairpin constructs have demonstrated them as triggers of rapid, efficient, and lasting trans-silencing. However, in R and S lines without hairpin constructs, 21 nt and 24 nt sRNA, respectively, are not detectable on northern blots, have very low read counts in the small RNA libraries and are therefore present in only tiny amounts. In diploid hybrids, both epialleles are present in a 1:1 ratio, which likely does not tilt the balance towards silencing. In tetraploids, 22.2% of the F2 progeny will contain one active and three silent copies, and here the amount of 24 nt sRNAs might traverse the threshold for the trans-silencing effect. The “missing” portion of resistant plants in paramutation cross progeny is usually between 15-30%, compatible with the assumption that especially plants with more silent than active epialleles (never occurring in diploids) are undergoing the switch. A similar higher incidence for paramutation with a triplex genotype for the paramutagenic *sulfurea* allele is apparent in tetraploid tomato hybrids (Berg

and Hagemann, 1979), and paramutation can occur independently at single alleles in subsequent generations (Berg and Hagemann, 1980).

The Arabidopsis/HPT and the tomato/sulf system are the only cases for which a role of tetraploidy in paramutation has been described. Neither for HPT nor for sulf is the silent state of the paramutagenic epiallele as such connected with polyploidy, as both alleles are stable also in diploid background, but the process of paramutation seems to be supported by the 4x karyotype. Genome duplications after formation of polyploids has often been associated with epigenetic changes (Comai et al., 2000; Madlung et al., 2002; Tayale and Parisod, 2013; Ding and Chen, 2018; Wendel et al., 2018) and might trigger the origin of epialleles. However, most studies were performed with allopolyploids, in which combining different genomes modifies chromatin, DNA methylation, or small RNA profiles. In these cases, changes can be due to either increased chromosome numbers, increased genetic diversity, increased gene copy numbers, incompatibility of regulatory components, or a mixture of these parameters. The situation in autopolyploids is simpler, as it excludes genetic determinants. Newly created autopolyploids can contain epigenetic modifications, e.g. switchgrass (Yan et al., 2019) and Arabidopsis (Yu et al., 2010), but effects on gene expression are much smaller than in allopolyploids (Wang et al., 2004) and can occur rather stochastically (Aversano et al., 2013). We did not find prominent differences between diploid and tetraploid lines, beside of divergent HPT expression and small RNA profiles at the epialleles. Therefore, the limitation of paramutation to the tetraploid hybrids is likely not due to the chromosome number as such, or to indirect consequences like larger nuclei (Sas-Nowosielska and Bernas, 2016), cell cycle regulation (Li et al., 2012), or differences in meiosis (Lloyd and Bomblies, 2016). It is more plausible to assume that rather the copy number of the epialleles, or better their ratio, determine the epigenetic change. Copy numbers, not varied by ploidy but by allelic variation of repetitive sequences, are also a major determinant of paramutation in maize (Panavas et al., 1999; Stam et al., 2002; Belele et al., 2013), again mostly in tandem configuration but sometimes far upstream of the coding sequence.

Another paradigm for a dosage effect is transcriptional silencing of a transposon, which occurs only after the copy number of the genomic source has increased beyond a threshold by new insertions (Marí-Ordóñez et al., 2013). A striking similarity with the HPT paramutation is the role of the 24 nt sRNAs that characterize the paramutagenic epiallele and that appear in connection with the switch to transcriptional transposon silencing. 24 nt sRNAs were also reported in very low amounts for the B-I/B' paramutation in maize, depending on the RDR2 ortholog MOP1, but were also found in neutral alleles and therefore not sufficient for paramutation (Barbour et al., 2012). Like in our study, their ectopic expression in larger amounts can mimic paramutation (Arteaga-Vazquez et al., 2010), again supporting the dosage dependency. There is no report for 21 nt sRNAs at paramutable maize alleles, in contrast to

the HPT system, where their presence might be connected to the downstream location of the repeat and its presence in the transcript. However, the repeats are in tandem orientation, so that hairpin formation as precursor for double-stranded RNA is unlikely. The requirement of MOP1 for maize paramutation (Dorweiler et al., 2000) and the reduced paramutation in *rdr2* mutants in our study strongly support a dependence on RNA-dependent RNA polymerase for the creation of the 24 nt small RNAs. As MOP2 and MOP3 are also components of the RdDM based silencing (Sidorenko et al., 2009; Sloan et al., 2014), it is clear that this epigenetic regulatory pathway is a core component of paramutation, and even outside of paramutation siRNA-induced epigenetic changes can persist over several generations (Regulski et al., 2013). However, the interference with RdDM in the Arabidopsis system did not result in hygromycin resistance, in contrast to restored BI and PI expression and regain of pigmentation e.g. in mop1 (Dorweiler et al., 2000). However, this could be due to the exceptionally strong silencing at the S epiallele, which needs simultaneous removal of DNA methylation and repressive histone modifications (Baubec et al., 2010).

The role of the 21 nt sRNAs in the Arabidopsis system remains elusive. They are only present together with the full-length active R allele regardless of ploidy and clearly dependent on the presence of the downstream repeat, as they disappear after its removal in Δ and γ deletion lines. They are usually associated with post-transcriptional silencing (PTGS) (Vaucheret et al., 2001), characterized by 5' preference for U and associated with AGO1 (Mi et al., 2008) and guide mRNA cleavage or translational repression (Bologna and Voinnet, 2014). The much stronger HPT expression in the deletion lines lacking 21 nt sRNAs indicates that, indeed, they could be responsible for partial suppression, but against expectation, the plants remain hygromycin-resistant. In fact, while the role of the 24 nt sRNAs in paramutation is more and more obvious, the stability of the HPT expression despite the presence of the repeat and the 21 nt sRNAs in lines with the R epiallele, for more than 20 years and numerous rounds of seed amplification, is becoming the bigger enigma.

The plant material has been propagated in different locations, over decades, and under a variety of environmental conditions. Among thousands of scored seedlings of the R and S lines, we never observed a spontaneous switch of the epialleles, neither from silent to active, nor from active to silent state. The latter might have gone undetected but certainly remains below 0.1%. However, the temperature dependence of paramutation indicates that after combination of the epialleles in the same plant, external factors can become highly relevant, and they are again supporting the involvement of RNA: the prevention of paramutation at lower temperatures resembles the inhibition of RNA-based virus defense in the cold (Szittya et al., 2003), and higher temperature enhances the production of 24 nt sRNAs (Baev et al., 2014).

In summary, the data presented for the HPT epialleles suggest that occurrence and degree of paramutation are determined by multiple factors. The dependence on functional RdDM, the involvement of repeats, and influence of environmental conditions overlap with features described for other paramutation phenomena. The study revealed an additional role for the copy number of the paramutable allele in the polyploid configuration and determination of paramutation through a delicate balance between transcription on one side and nature and amount of small RNAs on the other. The example of repeated switches between epigenetic states at the same locus, as reflected in the pedigree of the lines, provides evidence that spontaneous changes, targeted triggers, as well as genetic mutation events contribute to heritable stable yet reversible states and thereby to epigenetic diversity. With growing evidence for the existence of epialleles (Minow and Colasanti, 2020), it is tempting to speculate that paramutation between them often goes unnotified and may play an underestimated role in population genetics and evolution.

MATERIAL AND METHODS

Plant material

Arabidopsis thaliana ecotype Zürich (Zh) without (W) or with the transgenic epialleles (R and S) in diploid and tetraploid background were described previously (Mittelsten Scheid et al., 2003). Derivative lines (Δ) with genetic alterations leading to activation of the S epiallele were identified after failed T-DNA insertion (Foerster et al., 2011). RdDM mutants in tetraploid background and targeted deletion of the duplicated region were generated by CRISPR mutagenesis (see below). To generate tetraploid derivatives, 14 d-old diploid seedlings grown on plates were submerged in 0.1% colchicine (w/v) for 2 h, washed extensively in tab water and transferred to soil. Seeds from individual plants were harvested separately and sieved with mesh size 350 µm. Retained larger seeds were more likely to be polyploid and grown on soil to maturity. Ploidy of the progeny was evaluated by flow cytometry, after chopping approx. 100 mg of 14 d-old seedlings with a razor blade in 2 ml Galbraith buffer (Galbraith et al., 1983). Nuclei were filtered through a 30 µm nylon mesh (CellTrics, Sysmex Partec, Görlitz, Germany) and pelleted by 10 min centrifugation at 4000 g and 4°C. The pellet was washed in 1 ml Galbraith buffer and centrifuged again. The pellet was resuspended in 500 µl Galbraith buffer with 5 µl solution containing 0.1 mg/ml 4',6-diamidino-2-phenylindole (DAPI). After 10 min incubation on ice in the dark, samples were profiled for DAPI signal per nucleus on a BD LSRFortessa' (BD Biosciences, Franklin Lakes, New Jersey, US) and compared to that of diploid and tetraploid control samples. An overview about all plant lines in this study is provided in Fig. 1.

Growth conditions

Plants were regularly grown under long day (LD) conditions (16 h light 19°C, 8 h dark 16°C) and 60% relative humidity. For the temperature experiments with F1 hybrids, plants were grown on soil at standard conditions till development of the first inflorescences (approx. 21 days after germination) and then split into three cohorts: one remained in the standard conditions (19°C), the others moved to chambers with 10°C or 24°C under the same LD until seed set. Seeds were harvested, dried, and stratified before use at 4°C for 48 h.

Crossings

All crosses were made after manual emasculation of immature flower buds and pollination by transfer of mature pollen on the stigma of the prepared recipient. Contaminations were avoided by using sterilized tools and optical control under a binocular. All crosses were made in reciprocal orientation to exclude potential parent-of-origin effects.

Hygromycin resistance assay

Seeds were aliquoted in 2 ml tubes and surface-sterilized 10 min in a closed box in chlorine gas emerging after mixing 50 ml 14% sodium hypochlorite and 10 ml 37% chloric acid solution. Seeds were sown on plates containing germination medium (GM, https://www.oeaw.ac.at/gmi/research/research-groups/ortrun-mittelsten-scheid/resources/) without or with 20 µg/ml hygromycin B (Calbiochem) and grown for 14 d under long day (LD) conditions (16 h light, 8 h dark 21°C). Plates were incubated horizontally or vertically in case of root growth assays. Resistance was evaluated by counting green plants with extended leaves and well-developed roots. Resistance ratios were calculated with reference to the number of all germinated seeds. Root length of plants grown on vertical plates was determined with Fiji ((Version 1.51e) (Schindelin et al., 2012) and the corresponding SmartRoot plugin (Version 4.21) (Lobet et al., 2011) relative to a ruler and size of the plate.

DNA and RNA extraction

For genotyping, DNA was extracted from one young leaf each in a 2 ml tube containing approx. 10 glass beads and 500 µl extraction buffer (200 mM Tris pH 7.5, 250 mM NaCl, 25 mM EDTA) and ground in an MM400 homogenizer (Retsch, Düsseldorf, Germany) for 1 min at 30 Hz. Samples were centrifuged 1 min at 16000 g and 450 µl supernatant each was transferred into a 1.5 ml tube containing 450 µl isopropanol and 50 µl sodium acetate pH 5.2. After >1 h at -20°C, precipitates were collected by 20 min centrifugation at 16000 g, washed once in 75% EtOH and air-dried. Pellets were dissolved in 70 µl TE-buffer and stored at 4°C until usage.

For PCR analysis and bisulfite sequencing, DNA was prepared from 0.1 – 1 g plant material ground in liquid nitrogen and suspended in 0.3 – 3 ml CTAB elution buffer (A4150; PanReac AppliChem, Darmstadt, Germany, supplied with 1% polyvinyl pyrrolidone (PVP) and 1% β-mecaptoethanol (β-ME)). After adding RNase A (EN0531, Thermo Fisher Scientific, Waltham,

Massachusetts, US), samples were shaken 2 h at 65°C followed by phenol-chloroform extraction as described (Clarke, 2009). DNA was precipitated as above and dissolved in an appropriate amount of TE buffer. Concentrations were determined on a Nanodrop ND1000 (Thermo Fisher Scientific, Waltham, Massachusetts, US). DNA samples were stored at -20°C.

RNA was extracted from either 100 mg 14 d-old seedlings or 30 mg 35 d-old flower buds with TRI-Reagent (T9424, Sigma Aldrich, St. Louis, Missouri, US) according to the manufacturer's protocol. Total RNA was eluted in 50 µl RNase-free ddH2O and the concentration determined as for DNA. RNA integrity was analyzed by running aliquots on a 1.5% agarose-TAE gel, before treating 2.5 µg with DNaseI (EN0525, Thermo Fisher Scientific, Waltham, Massachusetts, US) according to the manual. DNAseI was heat-inactivated at 65°C for 10 min in the presence of 10 mM EDTA. DNA-free total RNA (RNA+) was precipitated by adding 1/10 volume sodium acetate pH 5.2 and 3 volumes 100% cold EtOH for > 2 h at -20°C. RNA was pelleted and washed once in 700 µl cold 75% EtOH. Air-dried pellets were dissolved in 40 µl RNase-free ddH2O and stored at -70°C until usage.

Quantitative RT-PCR

cDNA was synthesized from total RNA with RevertAid H Minus Reverse Transcriptase (EP0451; Thermo Fisher Scientific, Waltham, Massachusetts, US), with Random Hexamer primers (S0142; Thermo Fisher Scientific, Waltham, Massachusetts, US) according to the manufacturer's recommendation, but extending cDNA synthesis at 42°C for 90 min in the presence of 20 U RiboLock RNase Inhibitor (EO0381; Thermo Scientific). Prior to quantitative real-time PCR (qRT-PCR), absence of genomic DNA contaminations was controlled by 40 cycles PCR without reverse transcriptase (RT) at 60°C annealing and 1 min elongation, or with RT and primers spanning the intron of the UBC28 gene (At1g64230), in 30 cycles with 20 sec elongation. qRT-PCR was performed with FastStart Essential DNA Green Master (Roche; Rotkreuz, Switzerland) with approx. 5 ng of cDNA on a LightCycler96 system (Roche) according to the manufacturer's instructions. Amplification was performed in a two-step protocol with pre-incubation at 95°C for 10 min and 45 cycles alternating between 95°C 10 sec/60°C 30 sec. A final melting cycle at 97°C preceded melting curve analysis with the ΔΔCt method (Livak and Schmittgen, 2001) relative to the housekeeping genes *SAND* (At2g28390) or *EIF4A1* (At3g13920). The following primers were used: fw 5' GGGTAAATAGCTGCGCCGATGGTT and rv 5' CACGGCGGGAGATGCAATAGGTC (expression of HPT); fw 5' AACTCTATGCAGCATTTGATCCACT and rv 5' TGATTGCATATCTTTATCGCCATC (qRT-PCR normalization to *SAND*/At2g28390); fw 5' ATCCAAGTTGGTGTGTTCTCC and rv 5' GAGTGTCTCGAGCTTCCACTC (qRT-PCR normalization to *EIF4a*/At3g13920). fw 5' TCCAGAAGGATCCTCCAACTTCCTGCAGT and

rv 5' ATGGTTACGAGAAAGACACCGCCTGAATA (spanning a small intron in *UBC28*/At1g64230 to exclude contamination of cDNA with genomic DNA).

Northern blotting

For northern blotting of small RNAs, 10 to 30 µg total RNA were dissolved in 27 µl 50% deionized formamide, denatured at 65°C for 10 min and chilled on ice 5 min before loading on a 17.5% denaturing urea polyacrylamide TBE gel (dPAGE, UreaGel System, National Diagnostics, Atlanta, Georgia, US) in a 10-well 1.5 mm gel (Mini-PROTEAN; BioRad, Hercules, California, US). Gels were run 90 min at 200 V in 0.5 x TBE buffer.

RNA gels were blotted onto Hybond NX membranes (Amersham, UK) by capillary blotting with 20 x SSC overnight (DNA and mRNA gels) or by electroblotting (BioRad) for 1 h at 300 mA in 0.5 x TBE (sRNA gels). Membranes were washed once in 2 x SSC, air-dried and crosslinked in a UV Stratalinker 2400 (Stratagene, La Jolla, California, US) in Autocrosslink mode (DNA and mRNA blots) or chemically crosslinked (sRNA blots) for 90 min at 60°C in the dark with N-(3-dimethylaminopropyl)-N′-ethylcarbodiimide (EDC; E7750; Sigma Aldrich, St. Louis, Missouri, US) as described in (Pall and Hamilton, 2008). To visualize size markers and to confirm equal loading, crosslinked membranes (DNA and mRNA blots) were stained for 5 min at RT with methylene blue staining solution (0.04% methylene blue [w/v] in 0.5 M sodium acetate pH 5.2) and washed in ddH2O. Methylene blue stain was removed by incubation at RT for 20 min in 2 x SSC with 1% SDS on a shaker and then air-dried.

Probes for nucleic acid blots were either made from oligonucleotides, generated by PCR amplification of template DNA of the corresponding region, or obtained from plasmids via restriction digest. Twenty-five mM oligonucleotide were end-labeled with T4 Polynucleotide Kinase (PNK; EK0031; Thermo Scientific) according to the manual with 25 µCi [γ-32P]ATP (SRP-501; Hartmann Analytic; Braunschweig; Germany) for 45 min at 37°C. Radioactive labeled oligonucleotide probes were purified with Illustra MicroSpin G-25 micro columns (GE Healthcare, Chalfont St Giles, UK). Amplicons or restrictions fragments were gel-purified (D4002; Zymo, Irvine, California, US) and diluted to 25 ng/µl. Twenty-five ng were labeled with 50 µCi [α-32P]dCTP (SRP-305; Hartmann Analytic; Braunschweig; Germany) with the Amersham Rediprime II Random Prime Labeling System (RPN1633; GE Healthcare, Chalfont St Giles, UK) and purified on Illustra G-50 ProbeQuant micro columns (GE Healthcare, Chalfont St Giles, UK). Prior to hybridization, the longer, double-stranded probes were denatured for 5 min at 99°C.

Blot were hybridized as described (Church and Gilbert, 1984) in an hybridization oven overnight at 42°C for oligo-probes or 60°C for longer probes. Membranes were washed twice for 20 min each in 2 x SSC + 2% SDS at 50°C for oligo-probes or 65°C for longer probes, followed by 15 min in 2 x SSC +0.2% SDS at the respective temperatures. Membranes were

wrapped in Saran foil and exposed to phosphoscreens (GE Healthcare, Chalfont St Giles, UK) for up to 3 days. Screens were scanned on a phosphoimager (Typhoon FLA 9500, GE Healthcare, Chalfont St Giles, UK).

Preparation and sequencing of sRNAs

Sequencing libraries were prepared from either 500 ng total RNA or 3 µl TraPR-enriched (Grentzinger et al., 2020) sRNA with the NEBNext Multiplex Small RNA Library Prep Set for Illumina (E7300; NEB, Ipswich, Massachusetts) according to the manufacturer's protocol. Libraries were amplified in 14 cycles and cleaned up with AMPure XP beads (Beckman Coulter, Brea, California, US) prior to size selection on 6% polyacrylamide gels. cDNA libraries from approx. 140 bp to 150 bp (miRNAs/siRNAs including adapters) were gel-eluted as described in the manufacturer's protocol, EtOH-precipitated at -20°C overnight and eluted in 12 µl TE buffer. Size range of the libraries was confirmed on a Fragment Analyzer (Agilent formerly Advanced Analytical, Santa Clara, California, US) with the HS NGS Fragment Kit (1-6000 bp, DNF-474, Agilent, formerly Advanced Analytical, Santa Clara, California, US). sRNAs were sequenced on a HiSeq 2500' system in single end 50 bp mode (Illumina, San Diego, Santa Clara, California, US) by the VBCF NGS facility.

Plasmid construction and plant transformation

Cloning for subsequent Sanger sequencing and subcloning steps was made with the CloneJET PCR cloning kit (K1231, Thermo Fisher Scientific, Waltham, Massachusetts, US). For plant transformation with the recloned epiallele, we used a pSUN vector containing an oleosin promoter upstream of GFP (Shimada et al., 2010). For Cas9-expressing vectors, the ubiquitin4-2 promoter (PcUbi4-2) of *Petroselinum crispum* in the binary vector pDE-Cas9 (Fauser et al., 2014) was replaced with the egg cell-specific promoter EC1.2p (Wang et al., 2015) and the seed-specific GFP marker described before inserted to generate the pDEECO vector backbone. Two specific sgRNAs each were expressed from an Arabidopsis U6-26 promoter and processed via the tRNA-based multiplex system (Xie et al., 2015). For ectopic silencing constructs, the PPT cassette of pDE-Cas9 containing a minimal 35S promoter sequence was removed, and the SpCas9 coding sequence was replaced by the PDK/catalase intron sequence with flanking multiple cloning sites from the pHELLSGATE12 vector (Helliwell and Waterhouse, 2003), resulting in the pDEPPi backbone. Sequences matching different regions of the epiallele were inserted in sense and antisense direction to result in hairpin-forming transcripts. All vectors were transformed into E. coli strains DH5α. Plasmids were prepared with Zyppy Plasmid Miniprep Kit (D4020, Zymo, Irvine, California, US) and eluted in 50 µl ddH2O. Plasmids were controlled by Sanger sequencing before being transformed into electrocompetent Agrobacterium tumefaciens strain GV3101 pMP90. Arabidopsis plants were transformed with the floral dip method (Clough and Bent, 1998) and transformants selected

among seeds of the T0 plants based on the GFP signal in seeds as described (Shimada et al., 2010).

CRISPR mutagenesis

sgRNAs were designed with Chopchop (Montague et al., 2014; Labun et al., 2016; Labun et al., 2019) and ranked based on their efficiency score and suitable target sites. Prior to cloning the vectors for deletion mutagenesis, we tested several sgRNA candidates for efficient cleavage at the target site in an *in vitro* assay, incubating reconstituted sgRNA-Cas9 ribonucleoprotein with amplicons containing the target site (Bente et al., 2020). For genotyping potential mutant candidates, the *in vitro* assay was modified by replacing the wild type amplicon with that of pooled genomic DNA from candidate plants (Bente et al., 2020). PCR products were analyzed by electrophoresis, and the analysis repeated for individual plants in positive pools. Amplicons around the mutated site were further analyzed by Sanger sequencing to determine the type of mutation.

sRNA mapping and profiling

General sequence analysis, alignment, *in silico* cloning, and primer design were made with the CLC Main workbench (Version 7.7.2; Qiagen Bioinformatics, Venlo, Netherlands). For the analyses of sRNA sequencing, files in .bam format were first converted to .fastq files using BEDtools bamtofastq (Version 2.29.2) (Quinlan and Hall, 2010). Adapters were trimmed and reads were size-selected to the length of 18 nt – 26 nt by deploying cutadapt (Version 1.9.1; https://cutadapt.readthedocs.io/en/v1.9.1/), using the options -a AGATCGGA-g CGACGATC-m 18 -M 26 -discard-untrimmed for sRNA reads. Read quality and size distribution was controlled for mRNA libraries and sRNA libraries using FastQC (Version 0.11.5; http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Sequencing reads were aligned to the TAIR10 genome (Berardini et al., 2015) including an extra contig containing the sequence of the transgenic epiallele plus upstream and downstream flanking regions. sRNA libraries were aligned with Bowtie2 (Langmead and Salzberg, 2012), deploying the options -p 14 -k 500 -no-unal. Mapped files were sorted and indexed using Samtools (Version 1.3.1) (Li et al., 2009). In depth sRNA sequence analysis was performed by custom scripts provided by Dr. Alexis Sarazin (ETH Zürich, Switzerland). sRNA data were normalized to the library size as reads per million mapped reads (RPM).

Statistical analyses

Statistical analyses were performed either with Chi-square goodness-of-fit tests on total values or one factorial ANOVA on values with normal distribution, like normalized percentages and measurements. Statistical comparisons of two groups were performed by Student's T-test. Root length of plants grown on vertical plates was determined with Fiji (Version 1.51e)

(Schindelin et al., 2012) and the corresponding SmartRoot plugin (Version 4.21) (Lobet et al., 2011).

ACKNOWLEDGEMENTS

We are grateful for helpful discussions within our group and colleagues on campus, especially Arturo Marí-Ordóñez, Peter Schlögelhofer, and Michael Nodine. We are also grateful to Helene Fasching and Emi Miyakoda for help with resistance assays and genotyping. We thank the campus facilities, especially Next Generation Sequencing and Plant Facilities from the VBCF and MolBiol Service and BioOptics from the IMP/IMBA/GMI facilities for their excellent support. We further appreciate the financial support by the Austrian Science Fund (FWF DK1238) to the Doctoral Program “Chromosome Dynamics”.

REFERENCES

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics **31,** 166-169.

Arteaga-Vazquez, M., Sidorenko, L., Rabanal, F.A., Shrivistava, R., Nobuta, K., Green, P.J., Meyers, B.C., and Chandler, V.L. (2010). RNA-mediated trans-communication can establish paramutation at the b1 locus in maize. Proc Natl Acad Sci U S A **107,** 12986-12991.

Aversano, R., Caruso, I., Aronne, G., De Micco, V., Scognamiglio, N., and Carputo, D. (2013). Stochastic changes affect Solanum wild species following autopolyploidization. J Exp Bot **64,** 625-635.

Baev, V., Milev, I., Naydenov, M., Vachev, T., Apostolova, E., Mehterov, N., Gozmanva, M., Minkov, G., Sablok, G., and Yahubyan, G. (2014). Insight into small RNA abundance and expression in high- and low-temperature stress response using deep sequencing in Arabidopsis. Plant Physiol Biochem **84,** 105-114.

Barbour, J.E., Liao, I.T., Stonaker, J.L., Lim, J.P., Lee, C.C., Parkinson, S.E., Kermicle, J., Simon, S.A., Meyers, B.C., Williams-Carrier, R., Barkan, A., and Hollick, J.B. (2012). required to maintain repression2 is a novel protein that facilitates locus-specific paramutation in maize. Plant Cell **24,** 1761-1775.

Bateson, W., and Pellew, C. (1915). On the genetics of "rogues" among culinary peas (Pisum sativum). Journal of Genetics **5,** 13-36.

Baubec, T., Dinh, H.Q., Pecinka, A., Rakic, B., Rozhon, W., Wohlrab, B., von Haeseler, A., and Mittelsten Scheid, O. (2010). Cooperation of multiple chromatin modifications can generate unanticipated stability of epigenetic States in Arabidopsis. Plant Cell **22,** 34-47.

Belele, C.L., Sidorenko, L., Stam, M., Bader, R., Arteaga-Vazquez, M.A., and Chandler, V.L. (2013). Specific tandem repeats are sufficient for paramutation-induced trans-generational silencing. PLoS Genetics **9**.

Bente, H., Mittelsten Scheid, O., and Donà, M. (2020). Versatile in vitro assay to recognize Cas9-induved mutations. Plant Direct.

Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., and Huala, E. (2015). The Arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. Genesis **53,** 474-485.

Berg, W., and Hagemann, R. (1979). Paramutation am *sulfurea* Locus von *Lycopersicon esculentum* Mill. VIII. Der Dosiseffekt in Tetraploiden. . Biologisches Zentralblatt **98,** 307-314.

660 **Berg, W., and Hagemann, R.** (1980). Paramutation aom *sulfurea* Locus von *Lycopersicon esculentum*
661 Mill. IX. Schrittweise Paramutation in Tetraploiden. Biologisches Zentralblatt **99,** 25-36.
662 **Bologna, N.G., and Voinnet, O.** (2014). The diversity, biogenesis, and activities of endogenous silencing
663 small RNAs in Arabidopsis. Annu Rev Plant Biol **65,** 473-503.
664 **Brink, R.A.** (1958). Paramutation at the R locus in maize. Cold Spring Harb Symp Quant Biol **23,** 379-
665 391.
666 **Chandler, V.L., Eggleston, W.B., and Dorweiler, J.E.** (2000). Paramutation in maize. Plant Mol Biol **43,**
667 121-145.
668 **Cheng, C.Y., Krishnakumar, V., Chan, A.P., Thibaud-Nissen, F., Schobel, S., and Town, C.D.** (2017).
669 Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. Plant
670 Journal **89,** 789-804.
671 **Church, G.M., and Gilbert, W.** (1984). Genomic sequencing. Proceedings of the National Academy of
672 Sciences of the United States of America-Biological Sciences **81,** 1991-1995.
673 **Ciabrelli, F., Comoglio, F., Fellous, S., Bonev, B., Ninova, M., Szabo, Q., Xuereb, A., Klopp, C., Aravin,**
674 **A., Paro, R., Bantignies, F., and Cavalli, G.** (2017). Stable Polycomb-dependent
675 transgenerational inheritance of chromatin states in Drosophila. Nat Genet **49,** 876-886.
676 **Clarke, J.D.** (2009). Cetyltrimethyl ammonium bromide (CTAB) DNA miniprep for plant DNA isolation.
677 Cold Spring Harb Protoc **2009**.
678 **Clough, S.J., and Bent, A.F.** (1998). Floral dip: a simplified method for Agrobacterium-mediated
679 transformation of Arabidopsis thaliana. Plant Journal **16,** 735-743.
680 **Coe, E.H.** (1959). A regular and continuing conversion-type phenomenon at the B-locus in maize.
681 Proceedings of the National Academy of Sciences of the United States of America **45,** 828-832.
682 **Comai, L., Tyagi, A.P., Winter, K., Holmes-Davis, R., Reynolds, S.H., Stevens, Y., and Byers, B.** (2000).
683 Phenotypic instability and rapid gene silencing in newly formed arabidopsis allotetraploids.
684 Plant Cell **12,** 1551-1568.
685 **Cuerda-Gil, D., and Slotkin, R.K.** (2016). Non-canonical RNA-directed DNA methylation. Nat Plants **2,**
686 16163.
687 **Ding, M., and Chen, Z.J.** (2018). Epigenetic perspectives on the evolution and domestication of
688 polyploid plant and crops. Current Opinions in Genetics and Development **42,** 37-48.
689 **Dorweiler, J.E., Carey, C.C., Kubo, K.M., Hollick, J.B., Kermicle, J.L., and Chandler, V.L.** (2000).
690 mediator of paramutation1 is required for establishment and maintenance of paramutation at
691 multiple maize loci. Plant Cell **12,** 2101-2118.
692 **Erhard, K.F., Jr., Stonaker, J.L., Parkinson, S.E., Lim, J.P., Hale, C.J., and Hollick, J.B.** (2009). RNA
693 polymerase IV functions in paramutation in Zea mays. Science **323,** 1201-1205.
694 **Fauser, F., Schiml, S., and Puchta, H.** (2014). Both CRISPR/Cas-based nucleases and nickases can be
695 used efficiently for genome engineering in Arabidopsis thaliana. Plant Journal **79,** 348-359.
696 **Foerster, A.M., Dinh, H.Q., Sedman, L., Wohlrab, B., and Mittelsten Scheid, O.** (2011). Genetic
697 rearrangements can modify chromatin features at epialleles. Plos Genetics **7,** e1002331.
698 **Gabriel, J.M., and Hollick, J.B.** (2015). Paramutation in maize and related behaviors in metazoans.
699 Semin Cell Dev Biol **44,** 11-21.
700 **Galbraith, D.W., Harkins, K.R., Maddox, J.M., Ayres, N.M., Sharma, D.P., and Firoozabady, E.** (1983).
701 Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science **220,** 1049-1051.
702 **Giacopelli, B.J., and Hollick, J.B.** (2015). Trans-Homolog Interactions Facilitating Paramutation in
703 Maize. Plant Physiol **168,** 1226-1236.
704 **Gouil, Q., and Baulcombe, D.C.** (2018). Paramutation-like features of multiple natural epialleles in
705 tomato. BMC Genomics **19,** 203.
706 **Grentzinger, T., Oberlin, S., Schott, G., Handler, D., Svozil, J., Barragan-Borrero, V., Humbert, A.,**
707 **Duharcourt, S., Brennecke, J., and Voinnet, O.** (2020). A universal method for the rapid
708 isolation of all known classes of functional silencing small RNAs. Nucleic Acids Res **48,** e79.
709 **Hagemann, R.** (1958). Somatische Konversion bei *Lycopersicon esculentum* Mill. Zeitschrift für
710 Vererbungslehre **89,** 587-613.

Helliwell, C., and Waterhouse, P. (2003). Constructs and methods for high-throughput gene silencing in plants. Methods **30,** 289-295.

Hetzl, J., Foerster, A.M., Raidl, G., and Mittelsten Scheid, O. (2007). CyMATE: a new tool for methylation analysis of plant genornic DNA after bisulphite sequencing. Plant Journal **51,** 526-536.

Hollick, J.B., and Chandler, V.L. (2001). Genetic factors required to maintain repression of a paramutagenic maize pl1 allele. Genetics **157,** 369-378.

Hollick, J.B., Patterson, G.I., Asmundsson, I.M., and Chandler, V.L. (2000). Paramutation alters regulatory control of the maize pl locus. Genetics **154,** 1827-1838.

Hovel, I., Pearson, N.A., and Stam, M. (2015). Cis-acting determinants of paramutation. Semin Cell Dev Biol **44,** 22-32.

Jiang, W., Li, Z., Yao, X., Zheng, B., Shen, W.H., and Dong, A. (2019). jaw-1D: a gain-of-function mutation responsive to paramutation-like induction of epigenetic silencing. J Exp Bot **70,** 459-468.

Karesch, H., Bilang, R., Mittelsten Scheid, O., and Potrykus, I. (1991). Direct gene transfer to protoplasts of Arabidopsis thaliana. Plant Cell Reports **9,** 571-574.

Labun, K., Montague, T.G., Gagnon, J.A., Thyme, S.B., and Valen, E. (2016). CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. Nucleic Acids Research **44,** W272-W276.

Labun, K., Montague, T.G., Krause, M., Torres Cleuren, Y.N., Tjeldnes, H., and Valen, E. (2019). CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res **47,** W171-w174.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods **9,** 357-U354.

Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet **11,** 204-220.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics **25,** 2078-2079.

Li, X., Yu, E., Fan, C., Zhang, C., Fu, T., and Zhou, Y. (2012). Developmental, cytological and transcriptional analysis of autotetraploid Arabidopsis. Planta **236,** 579-596.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. Methods **25,** 402-408.

Lloyd, A., and Bomblies, K. (2016). Meiosis in autopolyploid and allopolyploid Arabidopsis. Curr Opin Plant Biol **30,** 116-122.

Lobet, G., Pagès, L., and Draye, X. (2011). A novel image-analysis toolbox enabling quantitative analysis of root system architecture. Plant Physiol **157,** 29-39.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology **15**.

Madlung, A., Masuelli, R.W., Watson, B., Reynolds, S.H., Davison, J., and Comai, L. (2002). Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic Arabidopsis allotetraploids. Plant Physiol **129,** 733-746.

Marí-Ordóñez, A., Marchais, A., Etcheverry, M., Martin, A., Colot, V., and Voinnet, O. (2013). Reconstructing de novo silencing of an active plant retrotransposon. Nature Genetics **published online**.

Matzke, M.A., and Mosher, R.A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat Rev Genet **15,** 394-408.

McDermaid, A., Monier, B., Zhao, J., Liu, B., and Ma, Q. (2019). Interpretation of differential gene expression results of RNA-seq data: review and integration. Brief Bioinform **20,** 2044-2054.

Mi, S.J., Cai, T., Hu, Y.G., Chen, Y., Hodges, E., Ni, F.R., Wu, L., Li, S., Zhou, H., Long, C.Z., Chen, S., Hannon, G.J., and Qi, Y.J. (2008). Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5 ' terminal nucleotide. Cell **133,** 116-127.

763 **Minow, M.A.A., and Colasanti, J.** (2020). Does variable epigenetic inheritance fuel plant evolution?
764 Genome **63,** 253-262.
765 **Mittelsten Scheid, O., Afsar, K., and Paszkowski, J.** (2003). Formation of stable epialleles and their
766 paramutation-like interaction in tetraploid Arabidopsis thaliana. Nat Genet **34,** 450-454.
767 **Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M., and Valen, E.** (2014). CHOPCHOP: a
768 CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Research **42,** W401-W407.
769 **Pall, G.S., and Hamilton, A.J.** (2008). Improved northern blot method for enhanced detection of small
770 RNA. Nat Protoc **3,** 1077-1084.
771 **Panavas, T., Weir, J., and Walker, E.L.** (1999). The structure and paramutagenicity of the R-marbled
772 haplotype of Zea mays. Genetics **153,** 979-991.
773 **Patterson, G.I., and Chandler, V.L.** (1995). Paramutation in maize and related allelic interactions. Curr
774 Top Microbiol Immunol **197,** 121-141.
775 **Pilu, R.** (2015). Paramutation phenomena in plants. Semin Cell Dev Biol **44,** 2-10.
776 **Quinlan, A.R., and Hall, I.M.** (2010). BEDTools: a flexible suite of utilities for comparing genomic
777 features. Bioinformatics **26,** 841-842.
778 **Regulski, M., Lu, Z., Kendall, J., Donoghue, M.T., Reinders, J., Llaca, V., Deschamps, S., Smith, A., Levy,**
779 **D., McCombie, W.R., Tingey, S., Rafalski, A., Hicks, J., Ware, D., and Martienssen, R.A.** (2013).
780 The maize methylome influences mRNA splice sites and reveals widespread paramutation-like
781 switches guided by small RNA. Genome Res **23,** 1651-1662.
782 **Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov,**
783 **J.P.** (2011). Integrative genomics viewer. Nat Biotechnol **29,** 24-26.
784 **Ronsseray, S.** (2015). Paramutation phenomena in non-vertebrate animals. Semin Cell Dev Biol **44,** 39-
785 46.
786 **Sas-Nowosielska, H., and Bernas, T.** (2016). Spatial relationship between chromosomal domains in
787 diploid and autotetraploid Arabidopsis thaliana nuclei. Nucleus-Phila **7,** 216-231.
788 **Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S.,**
789 **Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein, V., Eliceiri, K.,**
790 **Tomancak, P., and Cardona, A.** (2012). Fiji: an open-source platform for biological-image
791 analysis. Nat Methods **9,** 676-682.
792 **Shimada, T.L., Shimada, T., and Hara-Nishimura, I.** (2010). A rapid and non-destructive screenable
793 marker, FAST, for identifying transformed seeds of Arabidopsis thaliana. Plant J **61,** 519-528.
794 **Sidorenko, L., Dorweiler, J.E., Cigan, A.M., Arteaga-Vazquez, M., Vyas, M., Kermicle, J., Jurcin, D.,**
795 **Brzeski, J., Cai, Y., and Chandler, V.L.** (2009). A dominant mutation in mediator of
796 paramutation2, one of three second-largest subunits of a plant-specific RNA polymerase,
797 disrupts multiple siRNA silencing processes. PLoS Genet **5,** e1000725.
798 **Sloan, A.E., Sidorenko, L., and McGinnis, K.M.** (2014). Diverse gene-silencing mechanisms with distinct
799 requirements for RNA polymerase subunits in Zea mays. Genetics **198,** 1031-1042.
800 **Stam, M., Belele, C., Dorweiler, J.E., and Chandler, V.L.** (2002). Differential chromatin structure within
801 a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. Genes
802 Dev **16,** 1906-1918.
803 **Stonaker, J.L., Lim, J.P., Erhard, K.F., Jr., and Hollick, J.B.** (2009). Diversity of Pol IV function is defined
804 by mutations at the maize rmr7 locus. PLoS Genet **5,** e1000706.
805 **Szittya, G., Silhavy, D., Molnar, A., Havelda, Z., Lovas, A., Lakatos, L., Banfalvi, Z., and Burgyan, J.**
806 (2003). Low temperature inhibits RNA silencing-mediated defence by the control of siRNA
807 generation. Embo J **22,** 633-640.
808 **Tayale, A., and Parisod, C.** (2013). Natural pathways to polyploidy in plants and consequences for
809 genome reorganization. Cytogenetic and Genome Research **140,** 79-96.
810 **Trapnell, C., Pachter, L., and Salzberg, S.L.** (2009). TopHat: discovering splice junctions with RNA-Seq.
811 Bioinformatics **25,** 1105-1111.
812 **Vaucheret, H., Beclin, C., and Fagard, M.** (2001). Post-transcriptional gene silencing in plants. Journal
813 of Cell Science **114,** 3083-3091.

Wang, J.L., Tian, L., Madlung, A., Lee, H.S., Chen, M., Lee, J.J., Watson, B., Kagochi, T., Comai, L., and Chen, Z.J. (2004). Stochastic and epigenetic changes of gene expression in arabidopsis polyploids. Genetics **167,** 1961-1973.

Wang, Z.P., Xing, H.L., Dong, L., Zhang, H.Y., Han, C.Y., Wang, X.C., and Chen, Q.J. (2015). Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome Biology **16**.

Wassenegger, M., Heimes, S., Riedel, L., and Sanger, H.L. (1994). RNA-directed de novo methylation of genomic sequences in plants. Cell **76,** 567-576.

Wendel, J.F., Lisch, D., Hu, G.Q., and Mason, A.S. (2018). The long and short of doubling down: polyploidy, epigenetics, and the temporal dynamics of genome fractionation. Current Opinions in Genetics and Development **49,** 1-7.

Xie, K.B., Minkenberg, B., and Yang, Y.N. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proceedings of the National Academy of Sciences of the United States of America **112,** 3570-3575.

Yan, H., Bombarely, A., Xu, B., Wu, B., Frazier, T.P., Zhang, X., Chen, J., Chen, P., Sun, M., Feng, G., Wang, C., Cui, C., Li, Q., Zhao, B., and Huang, L. (2019). Autopolyploidization in switchgrass alters phenotype and flowering time via epigenetic and transcription regulation. Journal of Experimental Botany **70,** 5673-5686.

Yu, Z., Haberer, G., Matthes, M., Rattei, T., Mayer, K.F., Gierl, A., and Torres-Ruiz, R.A. (2010). Impact of natural genetic variation on the transcriptome of autotetraploid Arabidopsis thaliana. Proc Natl Acad Sci U S A **107,** 17809-17814.

Zhang, Z., Liu, X., Guo, X., Wang, X.J., and Zhang, X. (2016). Arabidopsis AGO3 predominantly recruits 24-nt small RNAs to regulate epigenetic silencing. Nat Plants **2,** 16049.

FIGURE LEGENDS

Figure 1. Overview over structure of epialleles and origin of plant material.

(A) Location of the epialleles on chromosome three, their structural variants, and the location of primers and probes. HPT: ORF of hygromycin phosphotransferase; ΔT: incomplete CaMV 35S terminator; P1, P2: CaMV 35S promoter; V1, V2: vector plasmid sequence; NC: non-coding sequence; S: silent epiallele rendering plants sensitive (red); R: active epiallele rendering plants resistant (blue) to hygromycin; Δ, γ: deletion derivatives. (B) Pedigree of lines. S, R, Δ, γ as in (A), W: wild type allele without insert (black); 2 letters: diploids; 4 letters: tetraploids; x: crossing; ⊗ selfing; colchicine: generation of tetraploids, * converted by paramutation. Links to other figures refer to the use of the material in different context.

Figure 2. The degree of paramutation in tetraploids depends on growth temperature in F1. (A) Experimental setup: Diploid or tetraploid plants with R epialleles were crossed with those containing S or W, and F1 hybrids from reciprocal crossings were grown at 19°C for three weeks before transfer to either 10°C, 19°C, or 24°C until seed maturity. F2 seeds were germinated at 19°C on GM plates containing 20 mg/L hygromycin B and resistance ratios determined after 14 days for tetraploids (B) and diploids (C). Data for reciprocal crosses were combined, as no parent-of-origin difference was observed. N = number of tested seedlings in each group. Bars represent the mean from two biological with three technical replicates each (B) and one biological with three technical replicates (C), with 100 plated seeds (B) each, or 50 plated seeds (C). Error bars indicate standard deviation. Dashed lines represent the expected segregation for tetraploid (B, 35:1 ≙ 97.2%) and diploid F2 populations (C, 3:1 ≙ 75%). F1 growth temperature is indicated by colour: 10°C (blue), 19°C (green), 24°C (red); light colours: F2 of R/S hybrids, dark colours: F2 of R/W hybrids. Statistical analysis was performed by Student's T-test ($p \leq 0.05$); p-values of relevant comparisons are indicated.

Figure 3. Epialleles are associated with distinct classes of sRNAs. (A) Proportion of sRNA length from all reads mapped to the epialleles in 14 day-old seedlings with tetraploid active (RRRR) or silenced (SSSS) epialleles. See S1 Fig for the corresponding data from diploids. (B) Coverage plots of 18 – 26 nt sRNA along the RRRR or SSSS epialleles. (C) Proportion of 5' prime nucleotides of epiallele-specific sRNAs in RRRR (left) and SSSS (right) plotted by size. See S1 Fig for the corresponding data for reads mapped to the genome.

Figure 4. Lack of paramutation in the background of RdDM mutants. (A) F2 seeds obtained by selfing tetraploid F1 hybrids were germinated on GM plates containing 20 mg/L hygromycin B and resistance ratios determined after 14 days. RRWW control (white), paramutation test hybrid RRSS in wild type (grey), *nrpd1* mutant (blue) or *rdr2* mutant (red). Data for reciprocal crosses were combined. N = number of tested seedlings in each group. The dashed line represents the expected F2 segregation for tetraploids (97.2%), statistical

analysis is based on a summed Chi-square goodness-of-fit test with indicated values. (B) Representative images for resistance assays quantified in (A). Orange asterisks indicate sensitive seedlings, scale bar = 3 cm.

Figure 5. Ectopic RNAi expression reduces HPT expression to different extent in diploids and tetraploids. (A) Three constructs with constitutive expression of hairpin-forming palindromes homologous to either the vector region of the R epiallele (hp1, green; hp2, grey) or the promoter region (hp3, blue) were inserted in ectopic positions, and their transcripts are processed into sRNAs. (A) Box plots of root length (measured with SmartRoot software) as indicator of resistance in the presence of 20 mg/L hygromycin in seven d-old T1 seedlings grown from seeds of T0 plants with or without the hairpin construct (sorted by seed-specific GFP fluorescence provided by the construct and grown on the same plate). White diamonds indicate the mean of each group. Statistical analysis was made with one factorial ANOVA ($p \leq 0.05$), p-values of relevant comparisons are indicated. N = number of tested seedlings in each group. (B) Representative images for root growth in lines without (upper panel) and with (lower panel) the hairpin construct, scale bar = 1 cm. (C) Relative expression of HPT in flower bud tissue of hairpin-containing T1 plants in tetraploids (left panel) and diploids (right panel), compared to sibling plants without hairpins. Measurements were done in three biological and three technical replicates. Statistical analysis was performed by doing Student's T-test (* = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$).

Figure 6. Paramutation depends on epiallele structure, and paramutated epialleles can exert secondary paramutation. (A) Tetraploid F2 seeds obtained by selfing F1 hybrids of crosses between S/W/R and RΔ2 (light grey) or RΔ4 (dark grey) were germinated on GM plates containing 20 mg/L hygromycin B and resistance ratios determined after 14 days. Data for reciprocal crosses were combined. N = number of tested seedlings in each group. The dashed line represents the expected F2 segregation for tetraploids (97.2%), statistical analysis is based on a summed Chi-square goodness-of-fit test with indicated values. (B) RΔ4 deletion alleles were combined with either W or S partners (all tetraploid). Individual plants in F2 progeny were genotyped for RΔ4, sorted into groups heterozygous (+/-) or homozygous (+/+) for the deletion and subjected to HPT expression analysis relative to homozygous RRRR. Bars indicate the mean of two biological and three technical replicates. Error bars indicate standard deviation. Statistical analysis was performed with Student's T-test (* = $p \leq 0.05$;), p values for relevant comparisons are indicated. (C) Plants from primary paramutation crosses, homozygous for RΔ4 alleles (excluding the presence of the old S allele) and only residual HPT expression (black star in B, now called SΔ4SΔ4SΔ4SΔ4, Fig. 1) were crossed to plants with the active, full length R alleles and analysed for hygromycin as described before in F2, in parallel to controls. The dashed line represents expected F2 segregation for tetraploid

populations at 97.2%; statistical analysis is based on a summed Chi-square goodness-of-fit test with indicated values. N = number of tested seedlings in each group.

Figure 7. Deletion of the downstream repeated region leads to higher expression at active epialleles and reduces paramutation. (A) Hygromycin resistance assay with 14 d-old F2 plants after crossing parents with full length or truncated epialleles. Data from reciprocal crosses were combined, and statistical analysis was performed by Student's T-test ($p \leq 0.05$); p-values of relevant comparisons are indicated. N = number of tested seedlings in each group. (B) Representative image for resistance assays in (A), scale bar = 3 cm. (C) Box plots of root length (measured with SmartRoot software) as indicator of resistance in the presence of 20 mg/L hygromycin in seven d-old plants containing full length or truncated active epialleles. White diamonds indicate the mean of each group. Statistical analysis was performed with Student's T-test with indicated p-value. N = number of tested seedlings in each group. (D) Representative images for root growth assays in (C), scale bar = 3 cm. (E) Northern blots of HPT transcripts in independent lines of tetraploid (left) and diploid (right) plants with full length or truncated active epialleles. Location of the probe is illustrated in Fig. 1.

Figure 8. Features of polyploidy-dependent paramutation. (A) The expressed epiallele (blue) confers hygromycin resistance, although the downstream repeat and the 21 nt sRNA limit transcription from the upstream promoter to some extent. The silent epiallele (red) is associated with 24 nt sRNAs and exerts strong silencing *in trans*, involving temperature-dependent RdDM. (B) Distribution of epigenotypes in F2 populations. Paramutation occurs in a dosage-dependent way and may be restricted to hybrids with more silent than active copies. This combination does not occur in diploids and explains the limitation of paramutation to tetraploids.

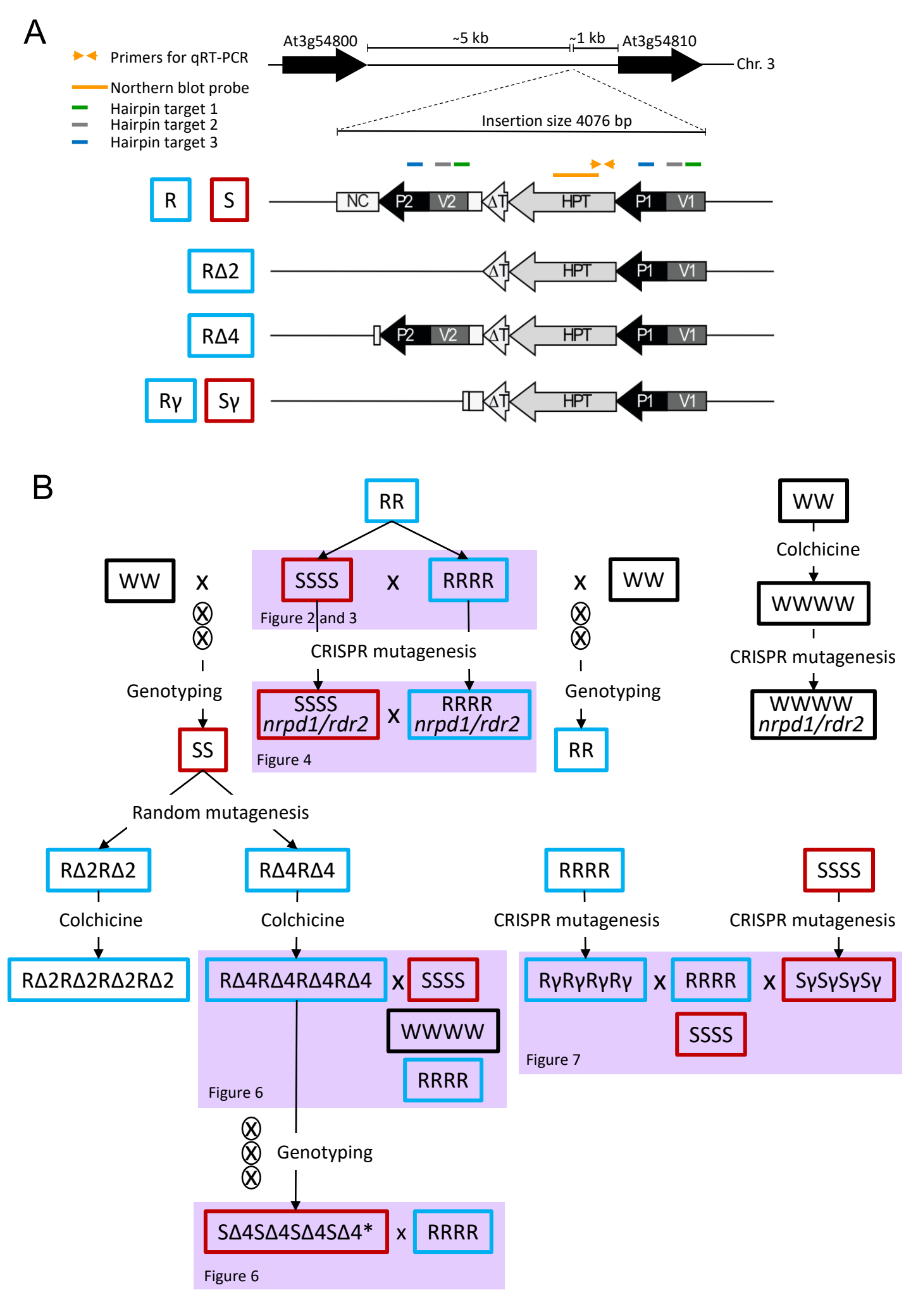
SUPPORTING FIGURE LEGENDS

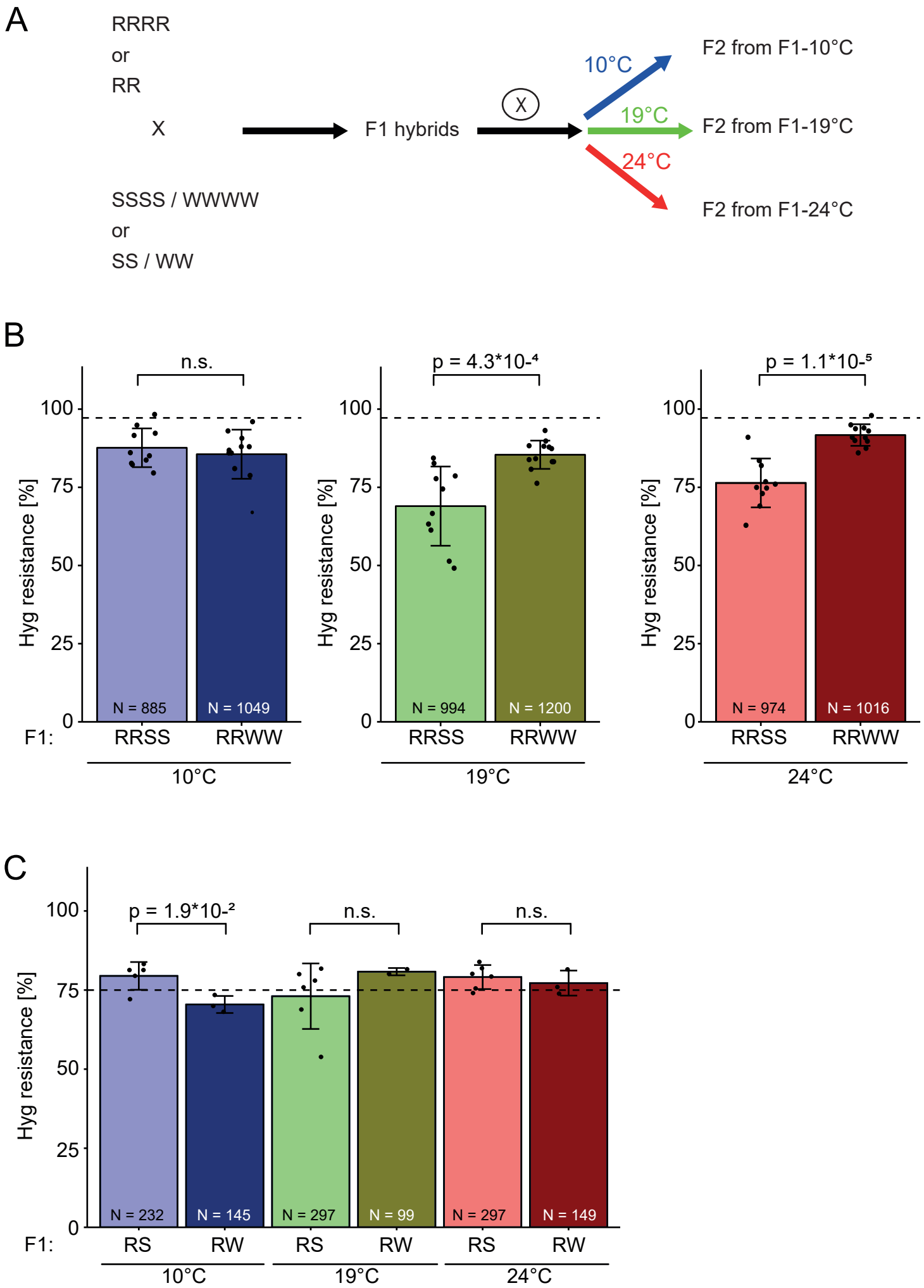
S1 Figure. Epialleles associated with distinct classes of sRNAs in diploids. (A) Proportion of sRNA length from all mapped reads in 14 day-old seedlings with diploid active (RR) or silenced (SS) epialleles. (B) Coverage plots of 18 - 26 nt sRNA along the RR or SS epialleles. (C) Proportion of 5' prime nucleotides of epiallele-specific sRNAs in RR (left) and SS (right) plotted by size.

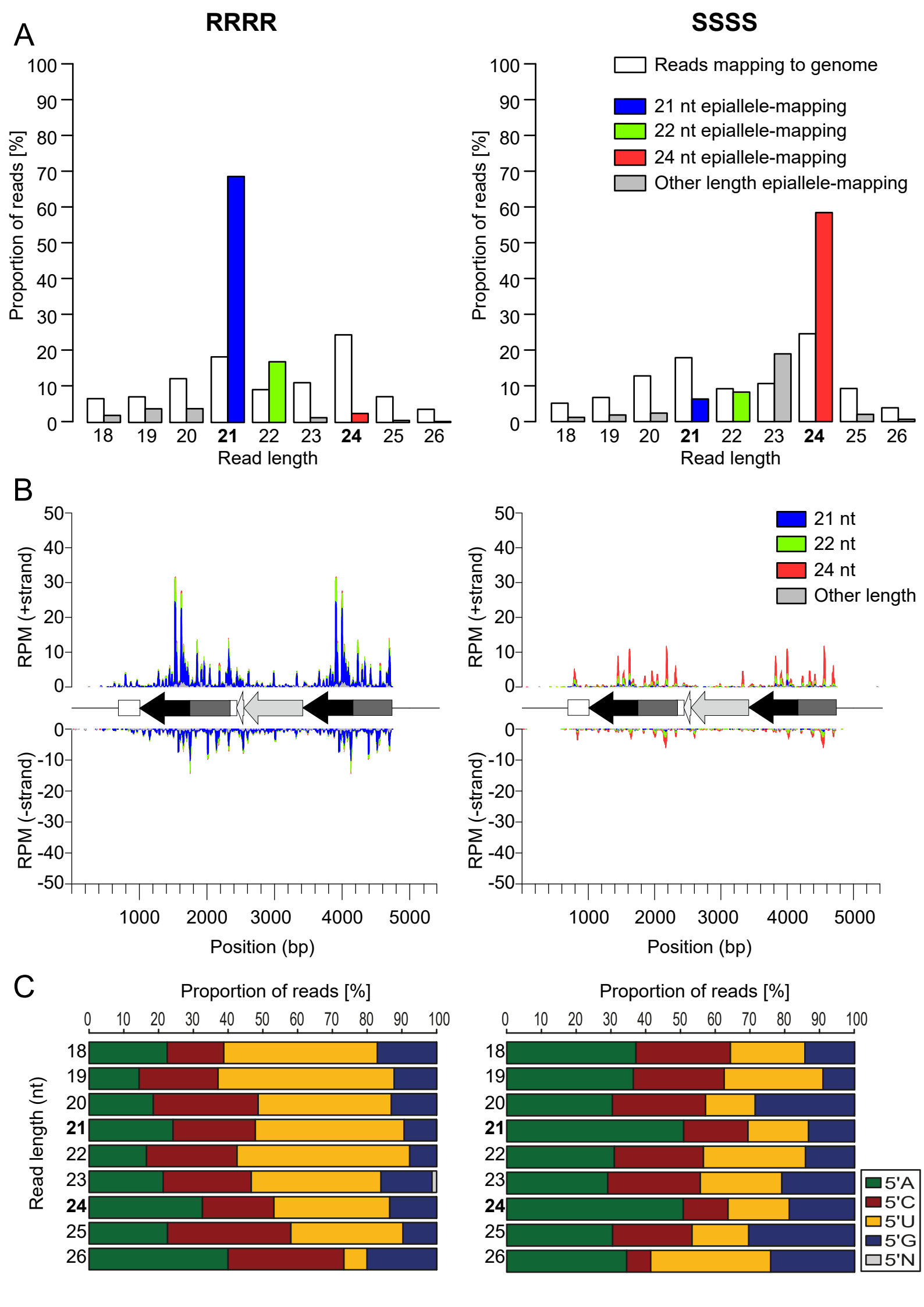
S2 Figure. Equal sRNA distribution genome-wide and no sRNAs mapping to the epiallele sequence in wild type. (A) Proportions of 5' nucleotides in different size classes in sRNA libraries from tetraploids (left) and diploids (right) in lines with active epialleles (top), silent epialleles (middle), and wildtype lines (bottom). (B) RNA libraries from tetraploid (WWWW, left) or diploid (WW, right) plants were attempted to map to the full length sequence of the epiallele.

S3 Figure. Effects of the Cas9-generated PolIV mutant on sRNA profiles. (A) Strong reduction of 24 nt sRNAs in a homozygous loss-of-function mutation in the *NRPD1* gene (encoding the largest subunit of PolIV) generated by CRISPR in the background of the tetraploid silent epiallele; including (B) those mapping to the epiallele.

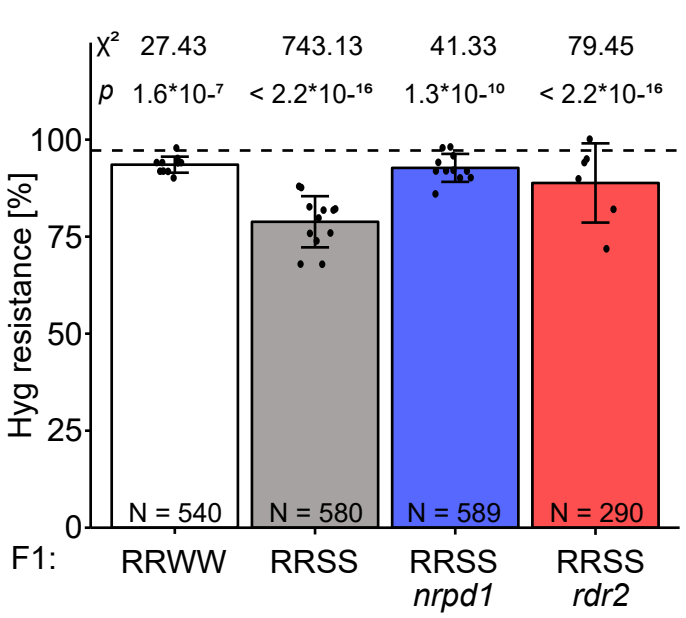
S4 Figure. sRNA profiles along the epialleles in F1 hybrids and in deletion mutants. (A) Profile from flower bud material in a pair of reciprocal paramutation test hybrids. (B – F) Profiles of tetraploid (left) and diploid (right) lines with the indicated alleles; (B) silent full length allele; (C) active allele missing downstream repeat after random mutagenesis and hygromycin screen; (D) CRISPR-generated deletion of the downstream repeat in the background of the silent epiallele; (E) active full length allele; (F) CRISPR-generated deletion of the downstream repeat in the background of the active epiallele. All values are reads per million mapped reads (RPM), and the y-axes are all set to the same maximum of ± 50 RPM.


A
Primers for qRT-PCR
Northern blot probe
Hairpin target 1
Hairpin target 2
Hairpin target 3
At3g54800
~5 kb
~1 kb
At3g54810
Chr. 3
Insertion size 4076 bp
R
S
NC
P2
V2
ΔT
HPT
P1
V1
RΔ2
RΔ4
Rγ
Sγ
B
RR
SSSS
x
RRRR
Figure 2 and 3
CRISPR mutagenesis
SSSS nrpd1/rdr2
x
RRRR nrpd1/rdr2
Figure 4
WW
x
Genotyping
SS
Random mutagenesis
RΔ2RΔ2
Colchicine
RΔ2RΔ2RΔ2RΔ2
RΔ4RΔ4
Colchicine
RΔ4RΔ4RΔ4RΔ4
x
SSSS
WWWW
RRRR
Figure 6
Genotyping
SΔ4SΔ4SΔ4SΔ4*
x
RRRR
Figure 6
x
WW
Genotyping
RR
WW
Colchicine
WWWW
CRISPR mutagenesis
WWWW nrpd1/rdr2
RRRR
CRISPR mutagenesis
RγRγRγRγ
x
RRRR
x
SγSγSγSγ
SSSS
Figure 7
SSSS
CRISPR mutagenesis

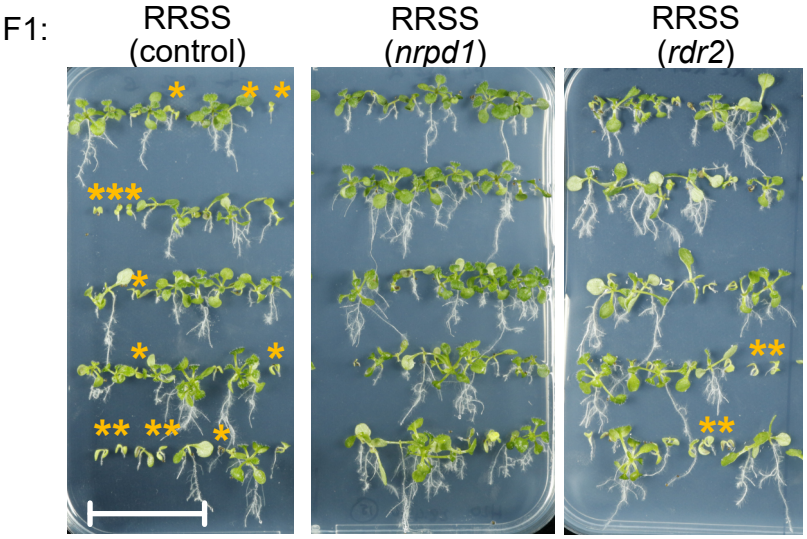

A
RRRR
or
RR
X
SSSS / WWWW
or
SS / WW
F1 hybrids
X
10°C
19°C
24°C
F2 from F1-10°C
F2 from F1-19°C
F2 from F1-24°C
B
n.s.
p = 4.3*10-⁴
p = 1.1*10-⁵
Hyg resistance [%]
100
75
50
25
0
N = 885
N = 1049
N = 994
N = 1200
N = 974
N = 1016
F1:
RRSS
RRWW
10°C
19°C
24°C
C
p = 1.9*10-²
n.s.
n.s.
N = 232
N = 145
N = 297
N = 99
N = 297
N = 149
RS
RW

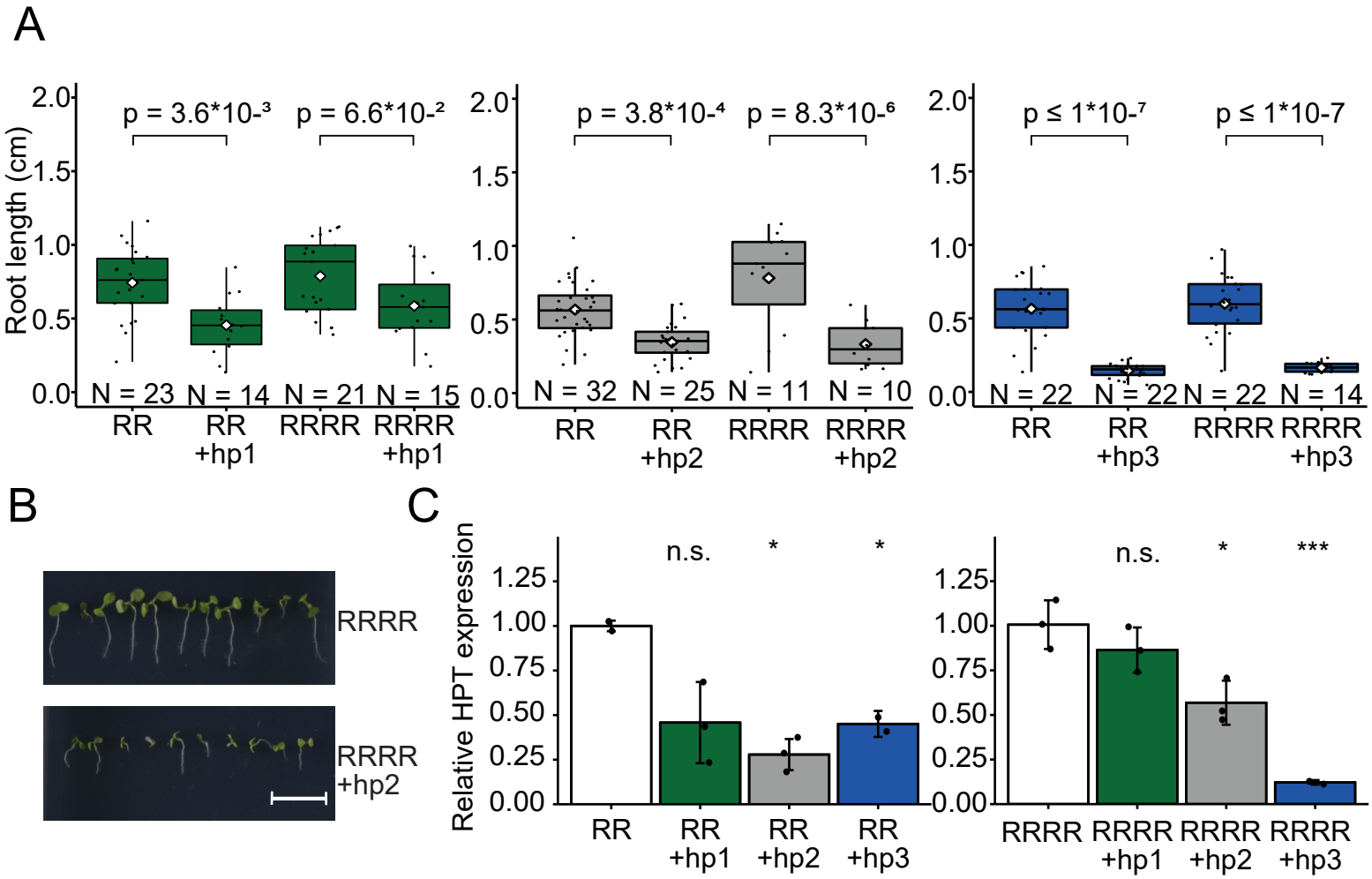

A
RRRR
SSSS
Reads mapping to genome
21 nt epiallele-mapping
22 nt epiallele-mapping
24 nt epiallele-mapping
Other length epiallele-mapping
Proportion of reads [%]
Read length
B
21 nt
22 nt
24 nt
Other length
RPM (+strand)
RPM (-strand)
Position (bp)
C
Proportion of reads [%]
Read length (nt)
5'A
5'C
5'U
5'G
5'N

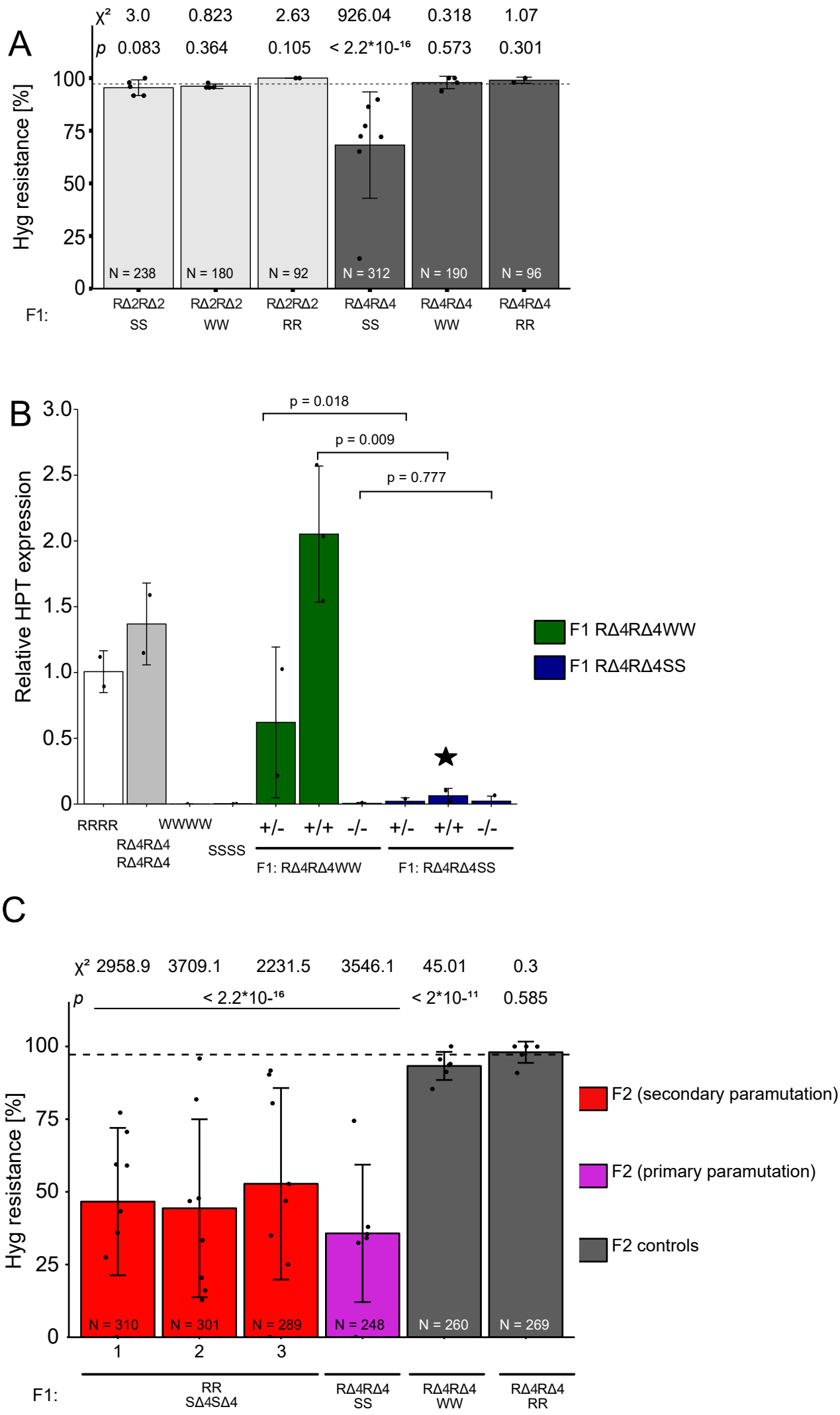
A

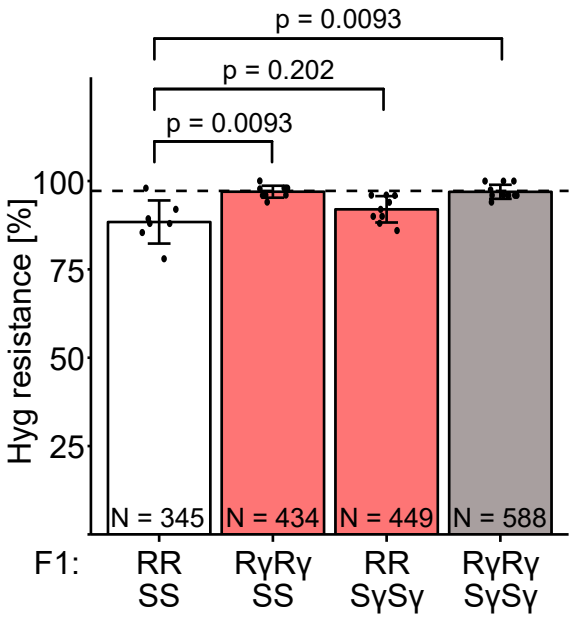

χ^2 27.43 743.13 41.33 79.45
p $1.6*10^{-7}$ $< 2.2*10^{-16}$ $1.3*10^{-10}$ $< 2.2*10^{-16}$
Hyg resistance [%]
100
75
50
25
0
N = 540
N = 580
N = 589
N = 290
F1:
RRWW
RRSS
RRSS *nrpd1*
RRSS *rdr2*

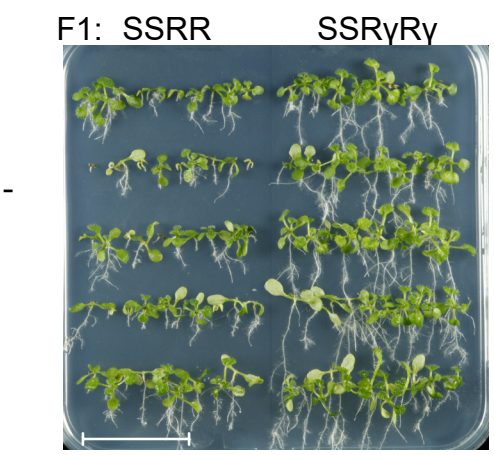
B

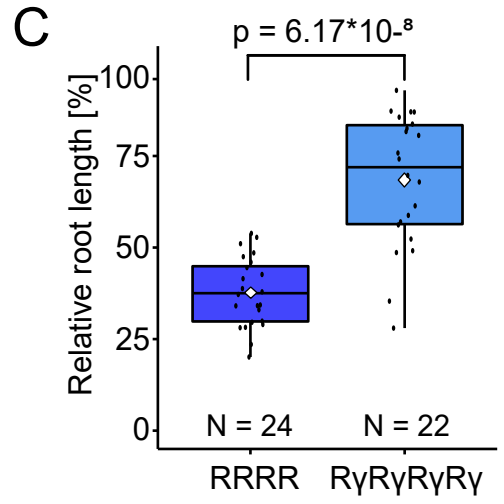

F1:
RRSS (control)
RRSS (*nrpd1*)
RRSS (*rdr2*)

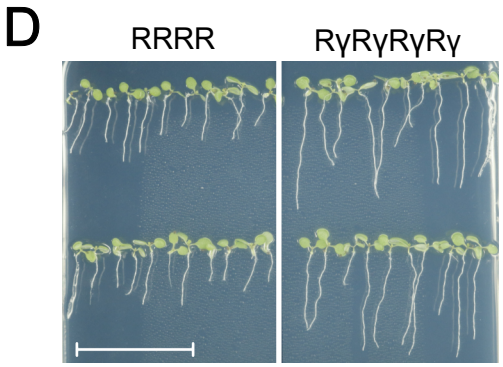

A
Root length (cm)
p = 3.6*10-3
p = 6.6*10-2
N = 23
N = 14
N = 21
N = 15
RR
RR +hp1
RRRR
RRRR +hp1
p = 3.8*10-4
p = 8.3*10-6
N = 32
N = 25
N = 11
N = 10
RR
RR +hp2
RRRR
RRRR +hp2
p ≤ 1*10-7
p ≤ 1*10-7
N = 22
N = 22
N = 22
N = 14
RR
RR +hp3
RRRR
RRRR +hp3
B
RRRR
RRRR +hp2
C
Relative HPT expression
n.s.
*
*
n.s.
*
***
RR
RR +hp1
RR +hp2
RR +hp3
RRRR
RRRR +hp1
RRRR +hp2
RRRR +hp3

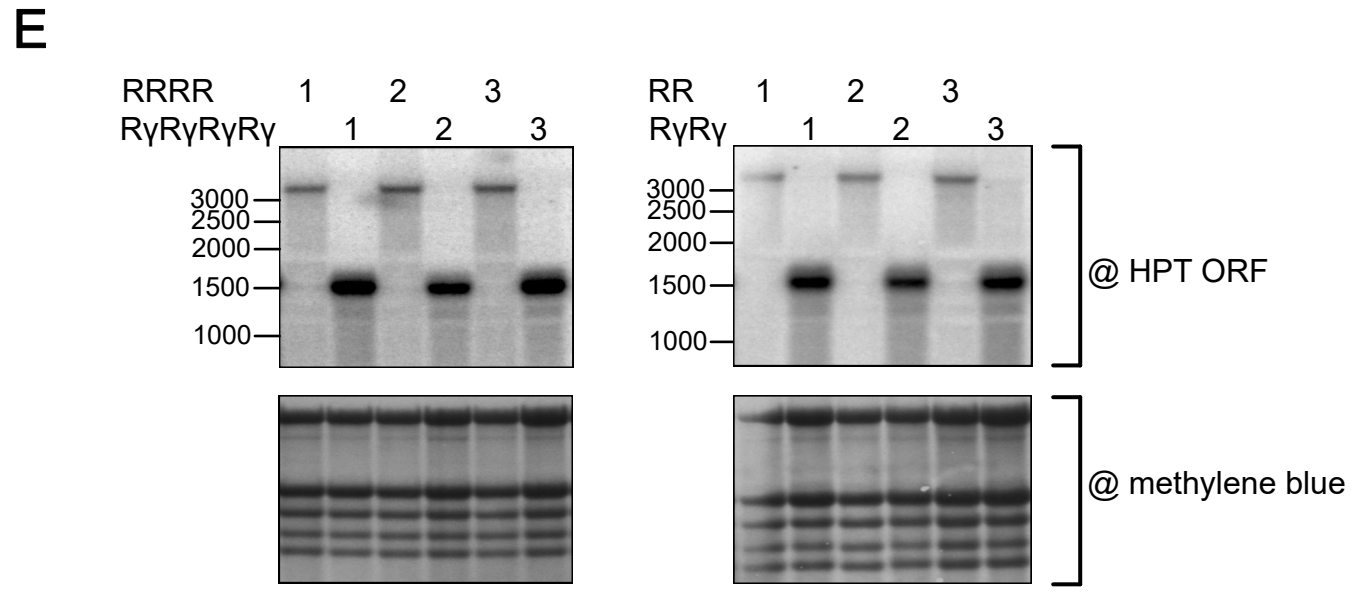

A
χ² 3.0 0.823 2.63 926.04 0.318 1.07
p 0.083 0.364 0.105 < 2.2*10-16 0.573 0.301
Hyg resistance [%]
100
75
50
25
0
N = 238
N = 180
N = 92
N = 312
N = 190
N = 96
F1:
RΔ2RΔ2 SS
RΔ2RΔ2 WW
RΔ2RΔ2 RR
RΔ4RΔ4 SS
RΔ4RΔ4 WW
RΔ4RΔ4 RR
B
Relative HPT expression
3.0
2.5
2.0
1.5
1.0
0.5
0
p = 0.018
p = 0.009
p = 0.777
RRRR
RΔ4RΔ4 RΔ4RΔ4
WWWW
SSSS
+/- +/+ -/-
F1: RΔ4RΔ4WW
+/- +/+ -/-
F1: RΔ4RΔ4SS
F1 RΔ4RΔ4WW
F1 RΔ4RΔ4SS
C
χ² 2958.9 3709.1 2231.5 3546.1 45.01 0.3
p < 2.2*10-16 < 2*10-11 0.585
Hyg resistance [%]
100
75
50
25
0
N = 310
N = 301
N = 289
N = 248
N = 260
N = 269
1
2
3
F1:
RR SΔ4SΔ4
RΔ4RΔ4 SS
RΔ4RΔ4 WW
RΔ4RΔ4 RR
F2 (secondary paramutation)
F2 (primary paramutation)
F2 controls


A
p = 0.0093
p = 0.202
p = 0.0093
Hyg resistance [%]
100
75
50
25
N = 345
N = 434
N = 449
N = 588
F1:
RR SS
RγRγ SS
RR SγSγ
RγRγ SγSγ

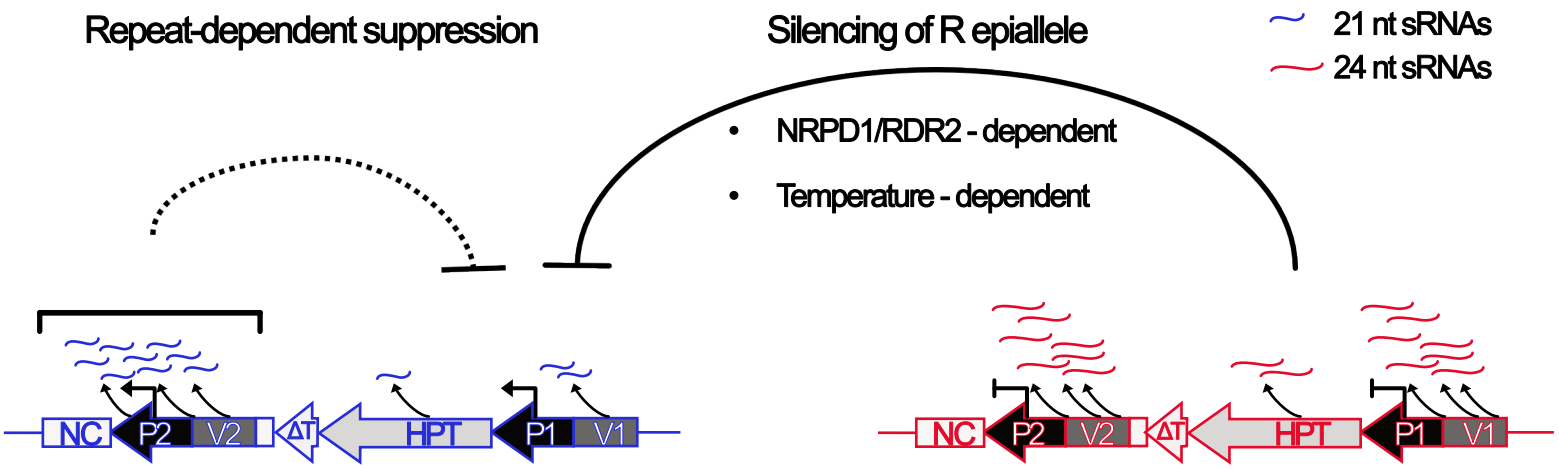

B
F1: SSRR
SSRγRγ


C
p = 6.17*10-8
Relative root length [%]
100
75
50
25
0
N = 24
N = 22
RRRR
RγRγRγRγ

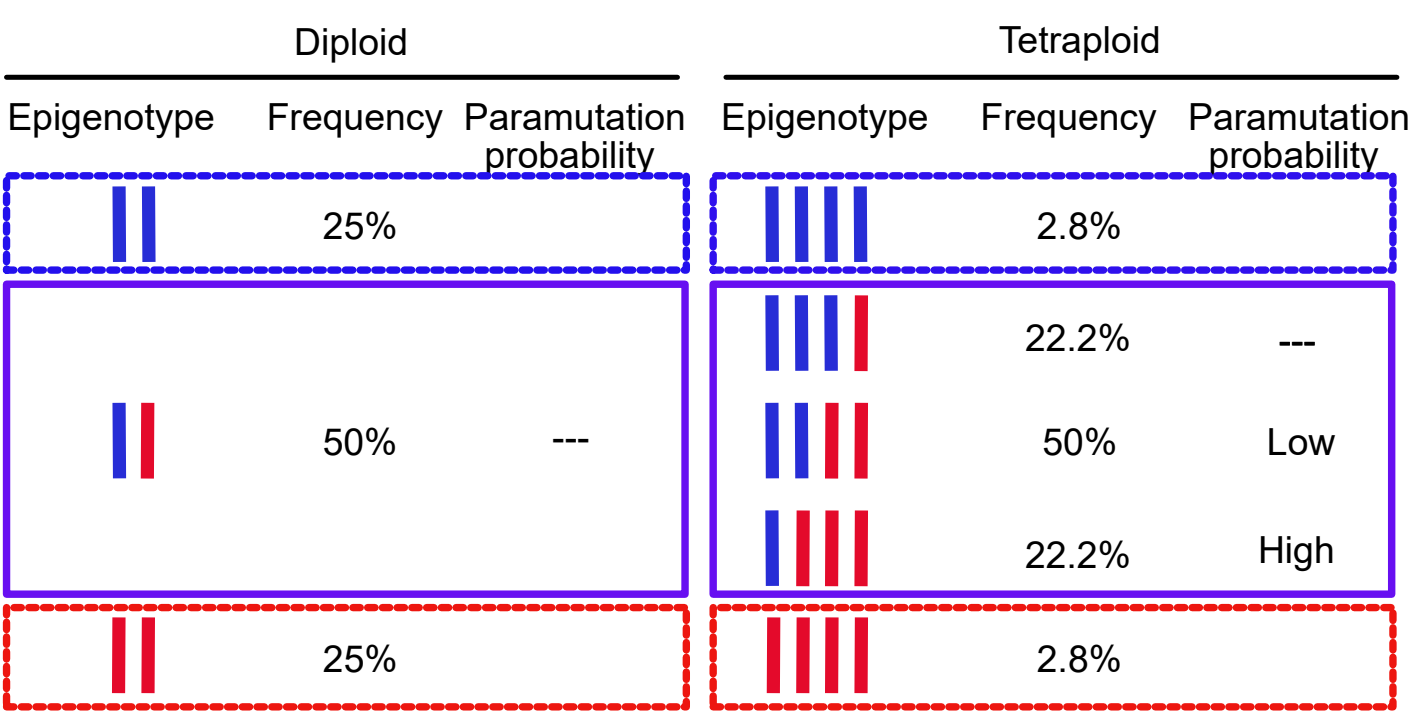

D
RRRR
RγRγRγRγ


E
RRRR 1 2 3
RγRγRγRγ 1 2 3
3000
2500
2000
1500
1000
RR 1 2 3
RγRγ 1 2 3
3000
2500
2000
1500
1000
@ HPT ORF
@ methylene blue

A

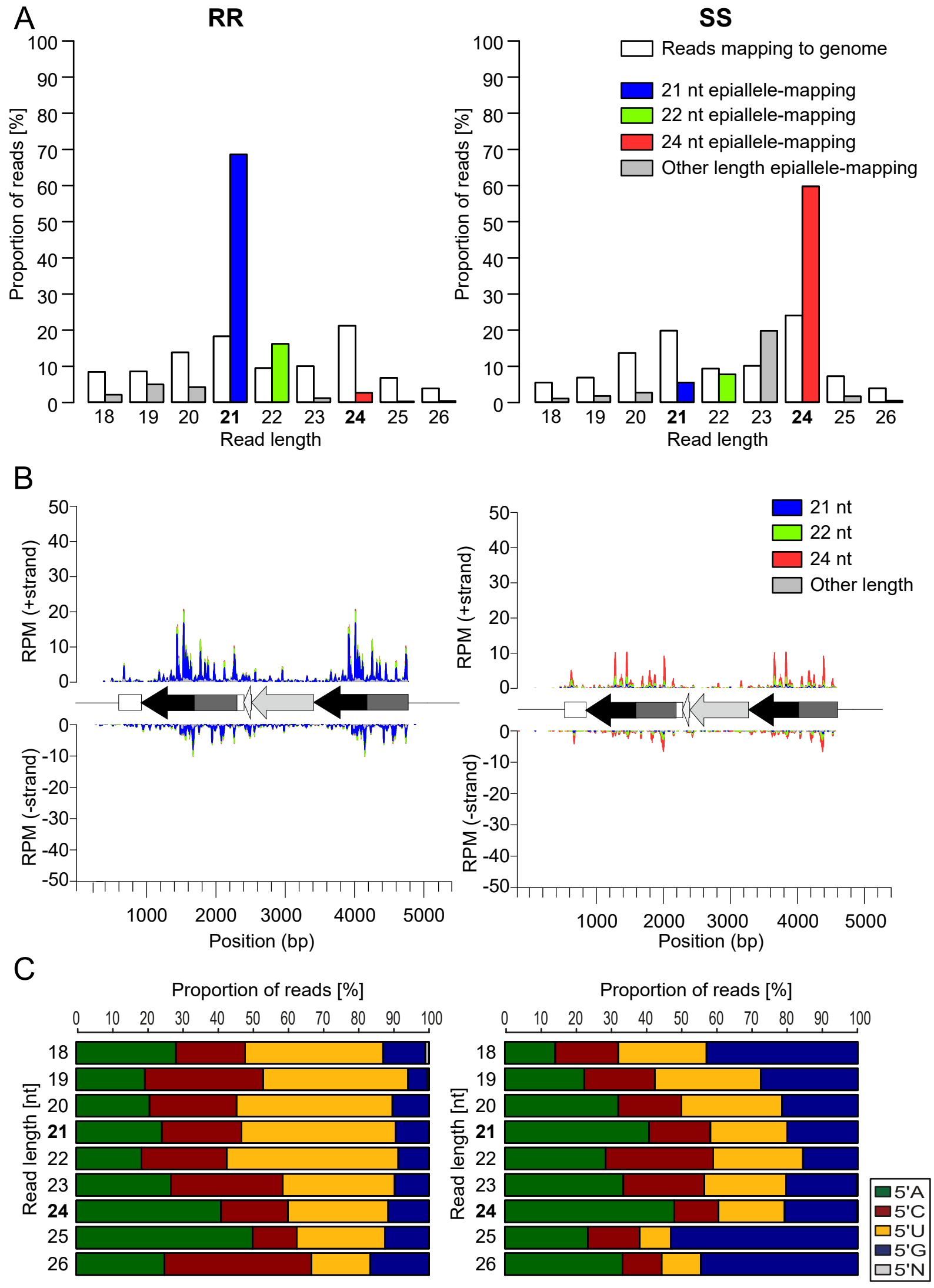

Repeat-dependent suppression
Silencing of R epiallele
21 nt sRNAs
24 nt sRNAs
NRPD1/RDR2 - dependent
Temperature - dependent
NC
P2
V2
ΔT
HPT
P1
V1
NC
P2
V2
ΔT
HPT
P1
V1

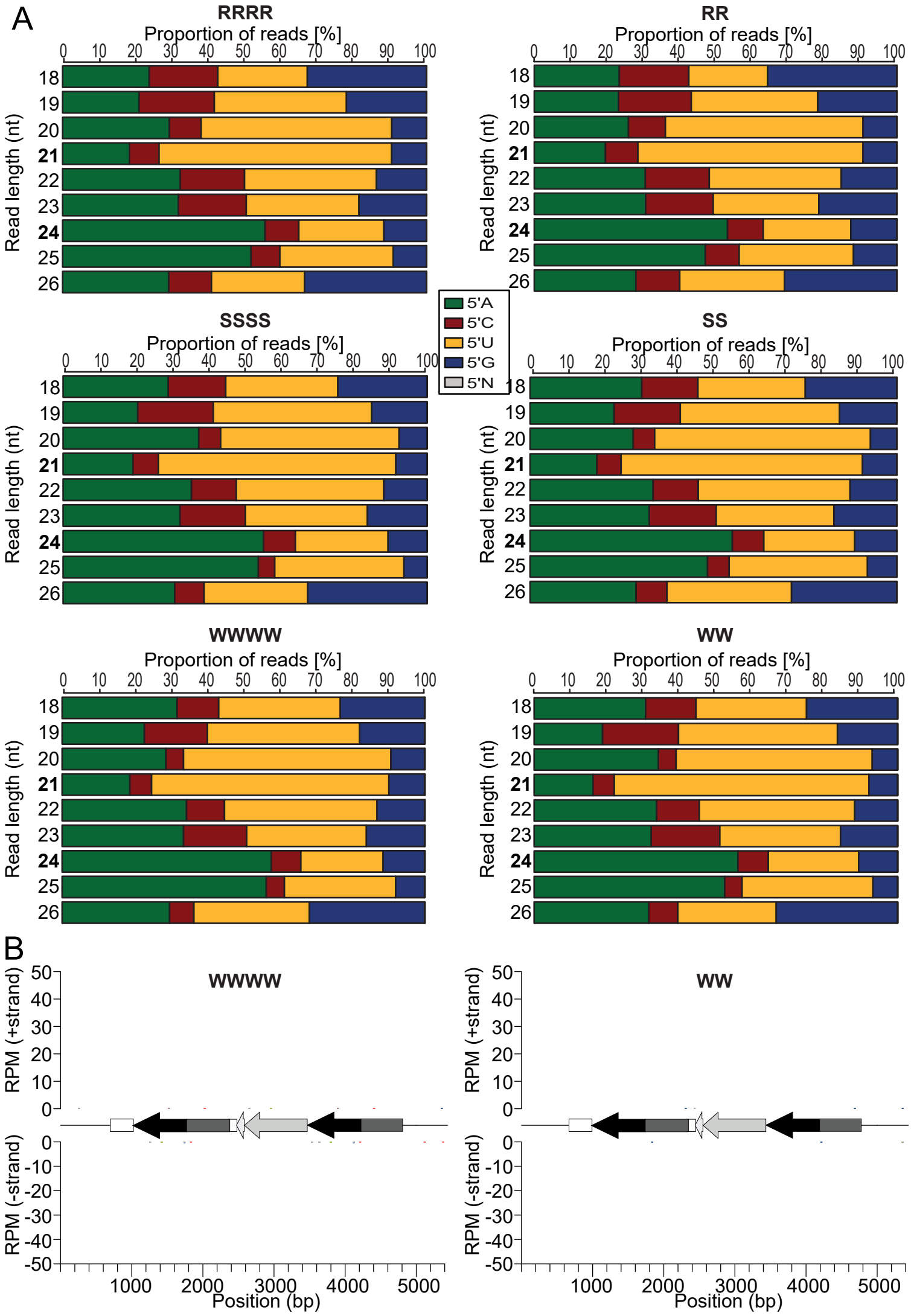
B

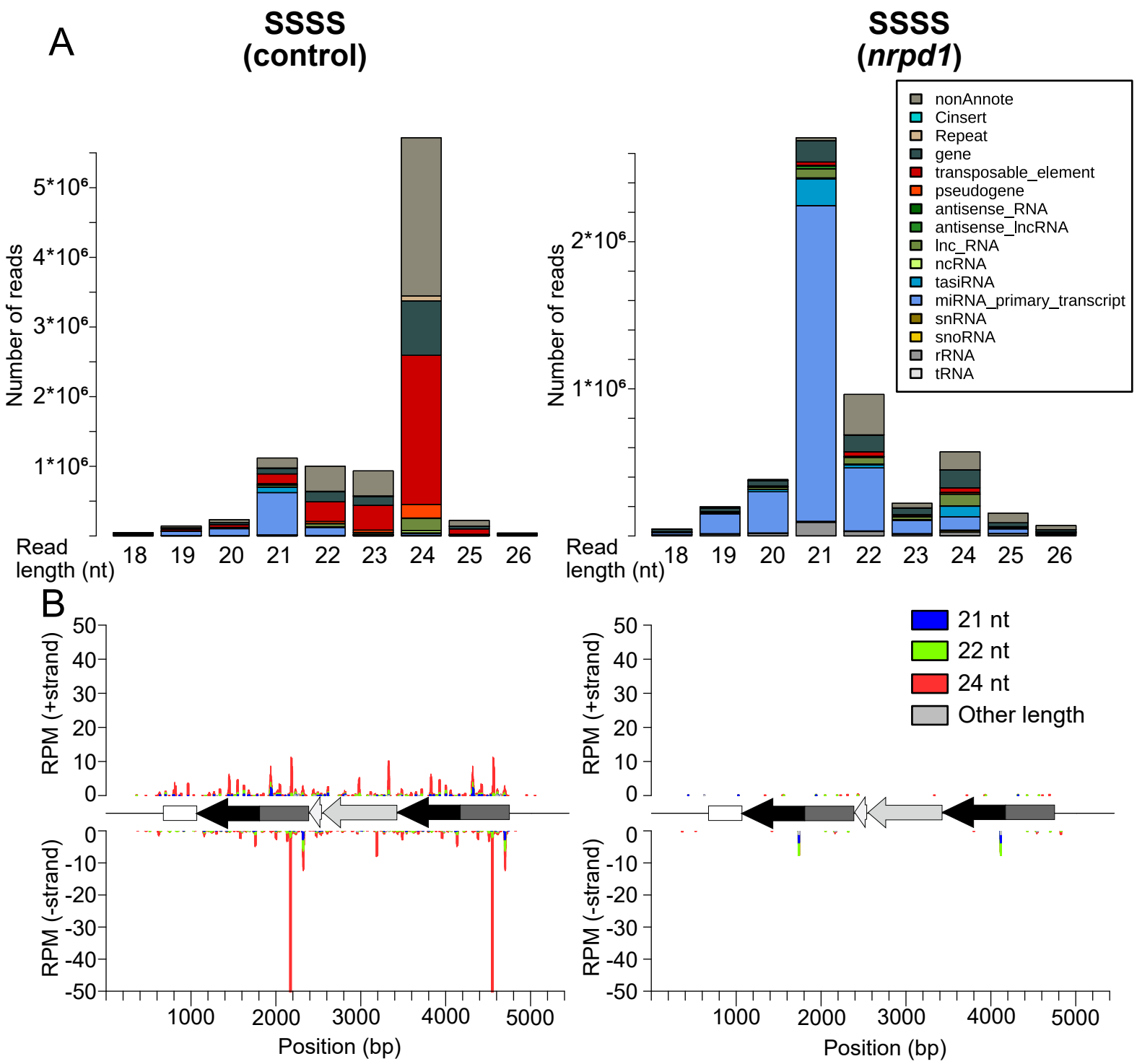

Diploid
Epigenotype
Frequency
Paramutation probability
25%
50%

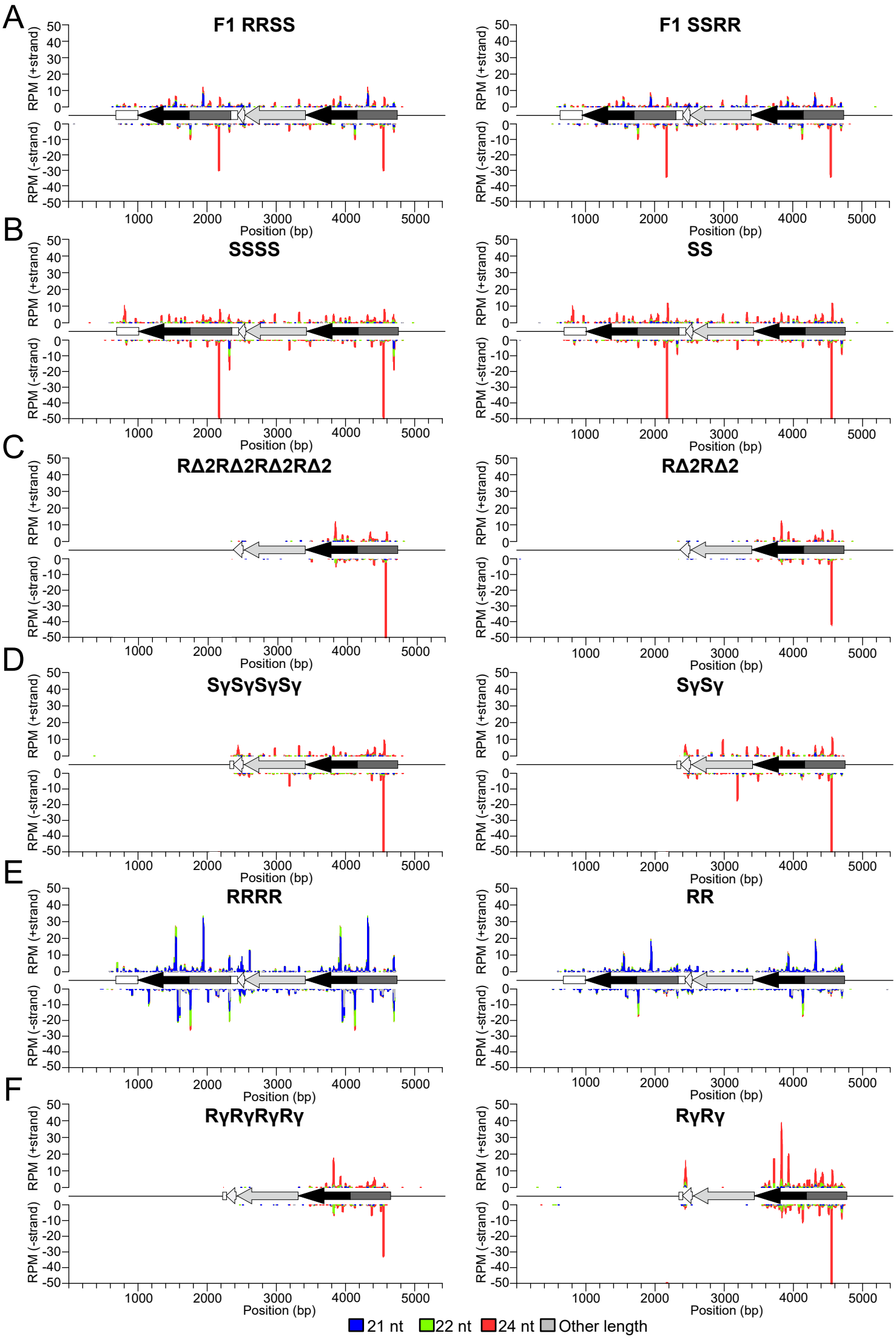
25%
Tetraploid
Epigenotype
Frequency
Paramutation probability
2.8%
22.2%

50%
Low
22.2%
High
2.8%


A
RR
SS
Reads mapping to genome
21 nt epiallele-mapping
22 nt epiallele-mapping
24 nt epiallele-mapping
Other length epiallele-mapping
Proportion of reads [%]
Read length
B
21 nt
22 nt
24 nt
Other length
RPM (+strand)
RPM (-strand)
Position (bp)
C
Proportion of reads [%]
Read length [nt]
5'A
5'C
5'U
5'G
5'N


A
RRRR
RR
SSSS
SS
WWWW
WW
Proportion of reads [%]
0 10 20 30 40 50 60 70 80 90 100
Read length (nt)
18 19 20 21 22 23 24 25 26
5'A
5'C
5'U
5'G
5'N
B
RPM (+strand)
RPM (-strand)
50 40 30 20 10 0
0 -10 -20 -30 -40 -50
1000 2000 3000 4000 5000
Position (bp)


A
SSSS
(control)
SSSS
(nrpd1)
Number of reads
5*10⁶
4*10⁶
3*10⁶
2*10⁶
1*10⁶
Read
length (nt)
18
19
20
21
22
23
24
25
26
nonAnnote
Cinsert
Repeat
gene
transposable_element
pseudogene
antisense_RNA
antisense_lncRNA
lnc_RNA
ncRNA
tasiRNA
miRNA_primary_transcript
snRNA
snoRNA
rRNA
tRNA
B
RPM (+strand)
RPM (-strand)
21 nt
22 nt
24 nt
Other length
Position (bp)
1000
2000
3000
4000
5000


A
F1 RRSS
F1 SSRR
B
SSSS
SS
C
RΔ2RΔ2RΔ2RΔ2
RΔ2RΔ2
D
SγSγSγSγ
SγSγ
E
RRRR
RR
F
RγRγRγRγ
RγRγ
RPM (+strand)
RPM (-strand)
Position (bp)
21 nt
22 nt
24 nt
Other length