

Supporting Information Figures

Generation of unequal nuclear genotype proportions in *Rhizophagus irregularis* progeny causes allelic imbalance in gene transcription

Chanz Robbins*, Joaquim Cruz Corella*, Consolée Aletti, Réjane Seiler, Ivan D. Mateus, Soon-Jae Lee, Frédéric G. Masclaux, Ian R. Sanders

* These authors contributed equally to this work

Accepted 26th May 2021

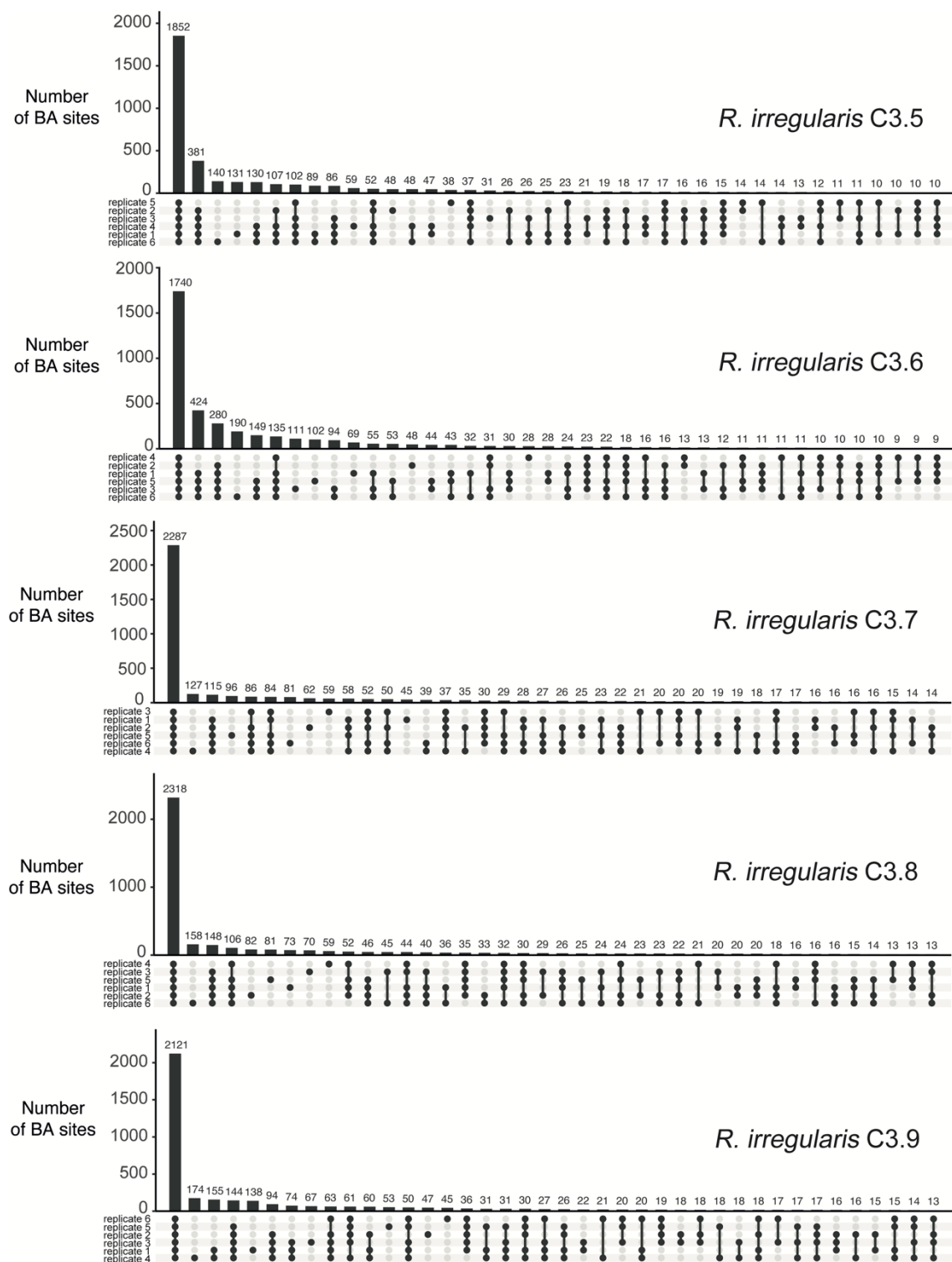


Figure S1: Reproducibility of bi-allelic (BA) sites in the ddRAD-sequencing data among the six replicates in each of the 5 dikaryon Single spores sibling lines (SSSLs) of *R. irregularis* C3. The y-axis represents the number of

BA sites commonly detected among the replicates being compared. The replicates considered for comparison are labelled at the left and shown along the horizontal axis by darkened circles indicating they were included in the comparison. Isolates are listed from top to bottom and indicated at the far right.

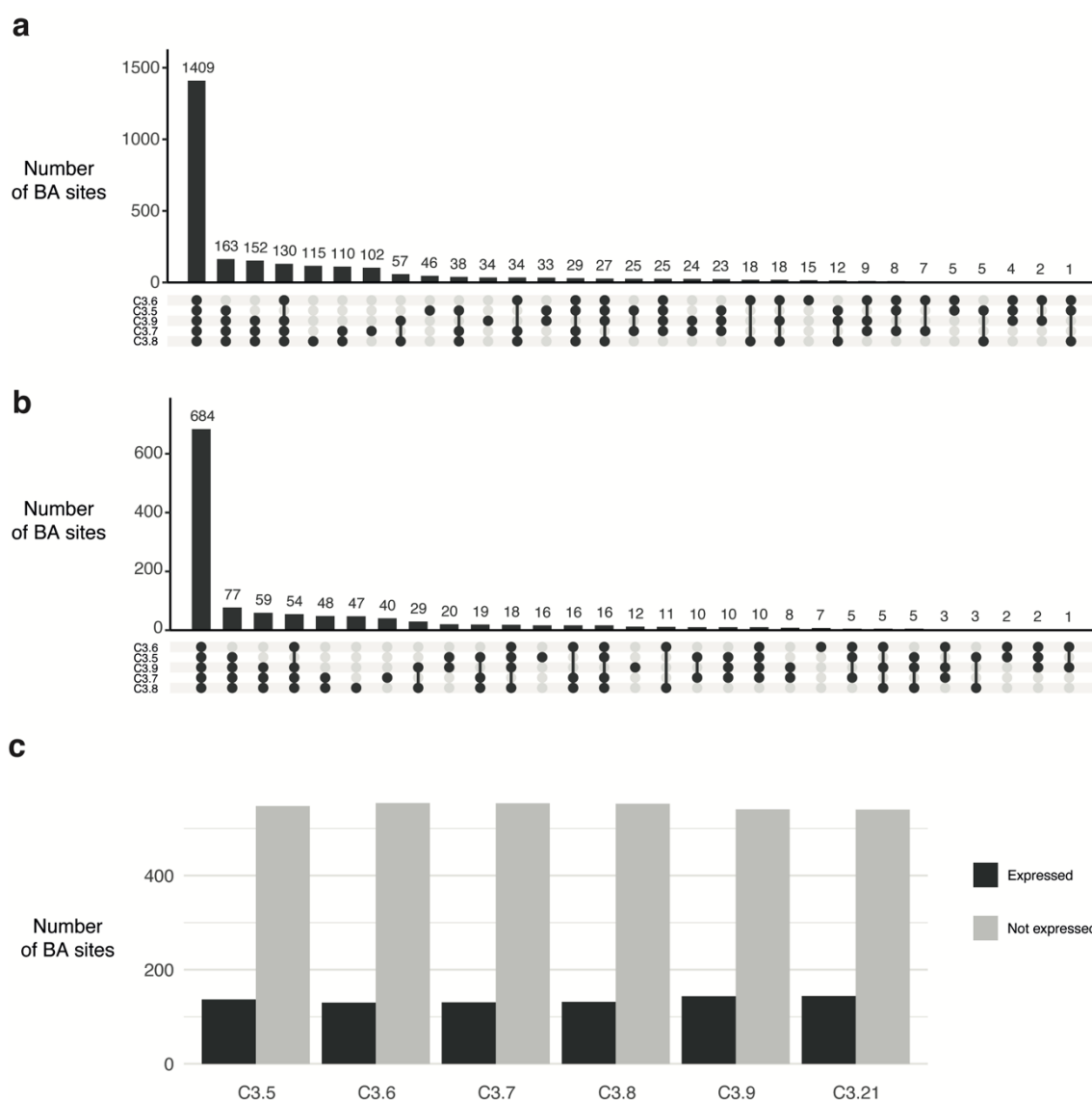


Figure S2: Consistency of bi-allelic (BA) sites in ddRADseq data and their transcription among single spore sibling lines (SSSLs) of *R. irregularis* C3.

(a) Reproducibility of BA sites in the ddRAD-sequencing data among 5 SSSLs of *R. irregularis* C3. SSSLs being compared are indicated with darkened circles and the number of BA sites is indicated on the y-axis. **(b)** Reproducibility of BA sites within coding sequences in the ddRAD-sequencing data among the 5 SSSLs of *R. irregularis* C3. SSSLs being compared are indicated with darkened circles and the number of BA genes is indicated on the y-axis. **(c)** Evidence of expression of

the 684 consistent BA sites from the ddRAD-sequencing data in the RNA-sequencing data.

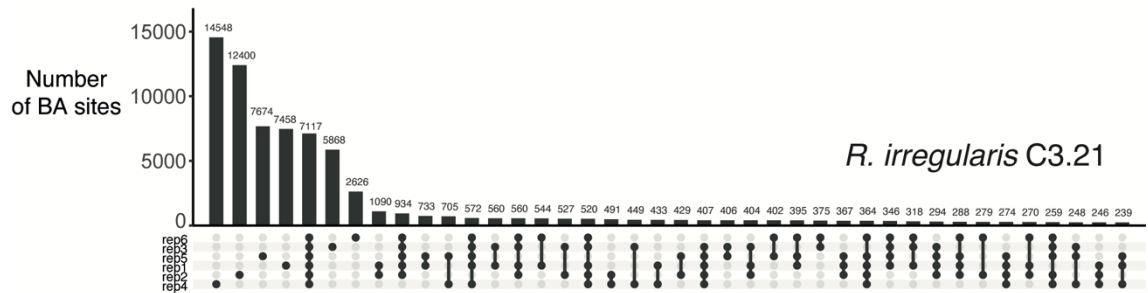
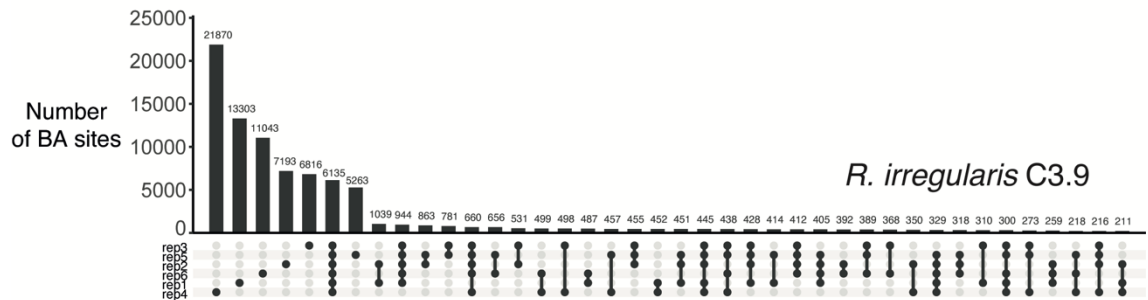
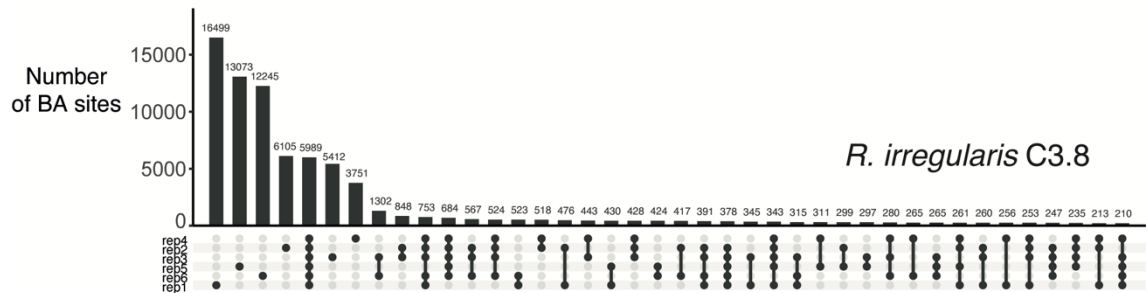
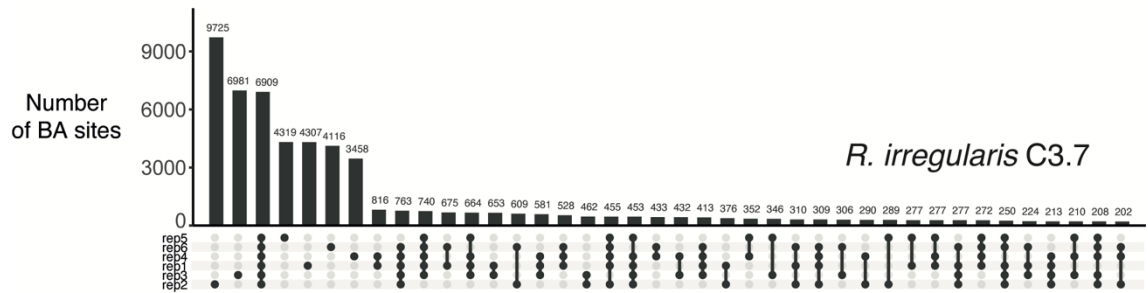
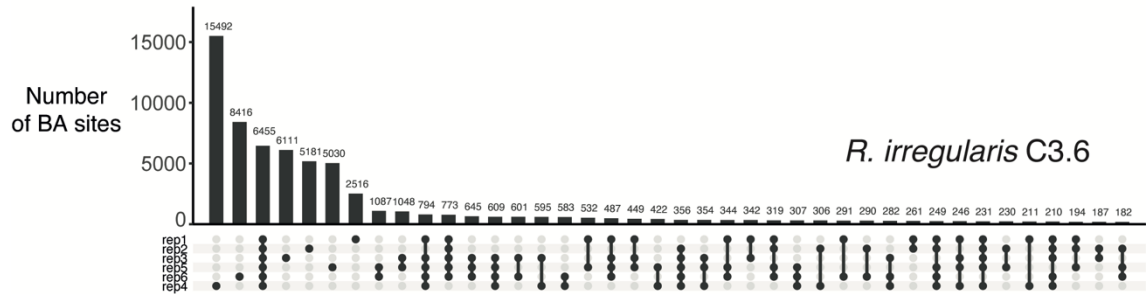
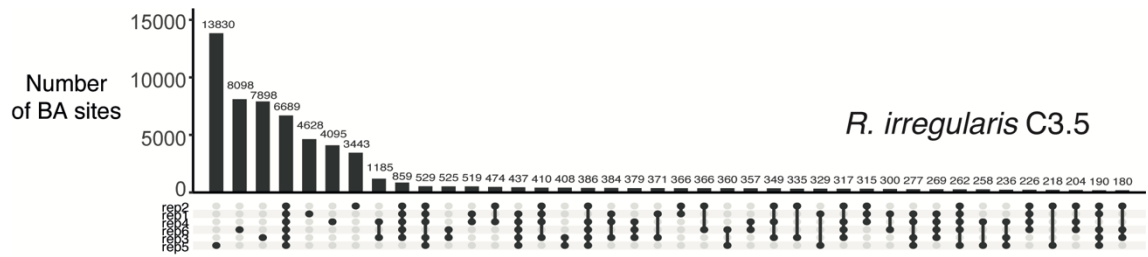


Figure S3: Reproducibility of bi-allelic (BA) sites in the RNA-sequencing data among the six replicates in each of the 6 dikaryon single spore sibling lines (SSSLs) of *R. irregularis* C3. The y-axis represents the number of BA sites commonly detected among replicates being compared. The replicates considered for comparison are labelled at the left and shown along the horizontal axis by darkened circles indicating they were included in the comparison. Isolates are listed from top to bottom and indicated at the far right.

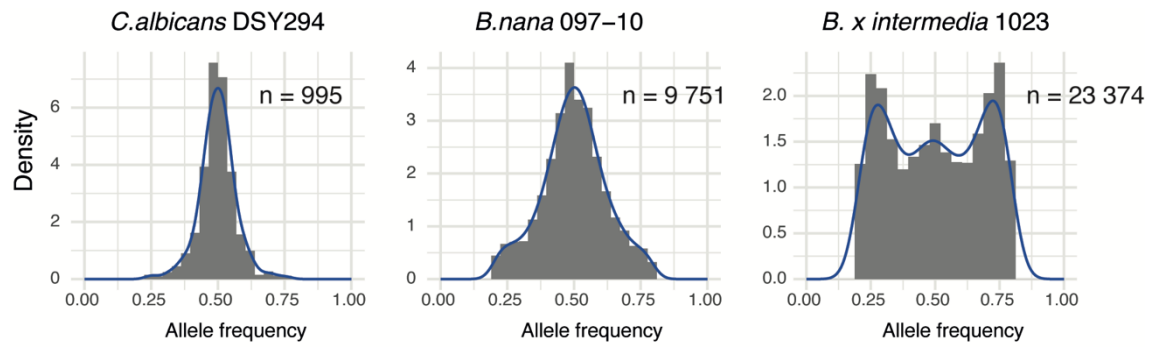
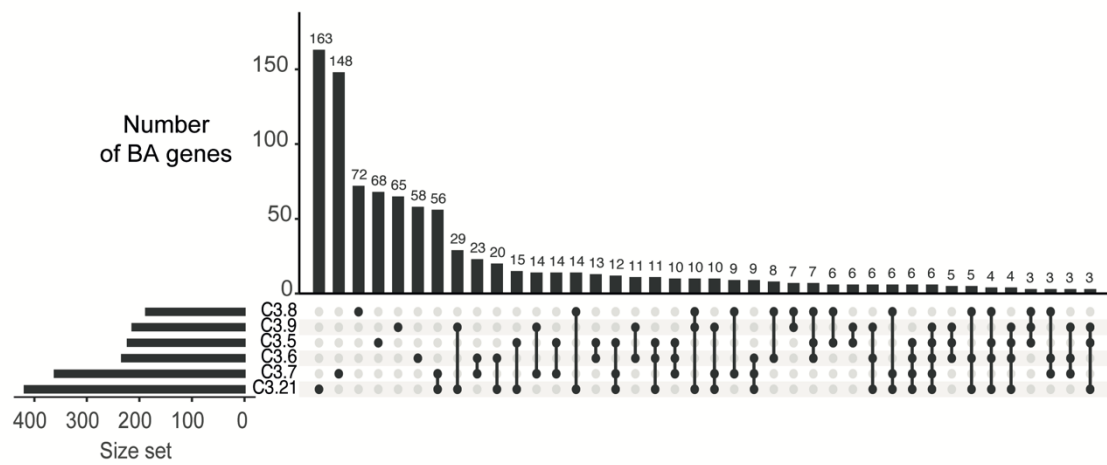


Figure S4: Allele frequency distributions of diploid (*C. albicans* and *B. nana*) and tetraploid (*B. x intermedia*) species derived from ddRAD-seq data. The number of bi-allelic sites analyzed for each species is indicated as (n). Diploid species are expected to exhibit one peak (0.5), and tetraploid species are expected to exhibit three peaks (0.25, 0.5 and 0.75).

a



b

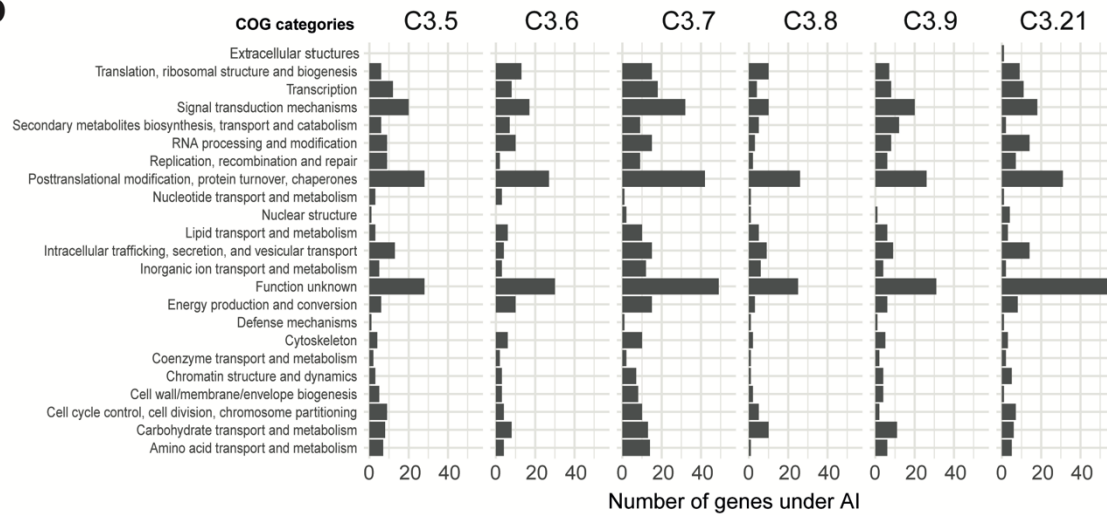


Figure S5: Genes with allele transcription under allelic imbalance. (a) Reproducibility of genes under allelic imbalance among the six *R. irregularis* SSSLs. **(b)** Cluster of orthologous (COG) categories of the genes under allelic imbalance for each *R. irregularis* SSSL.

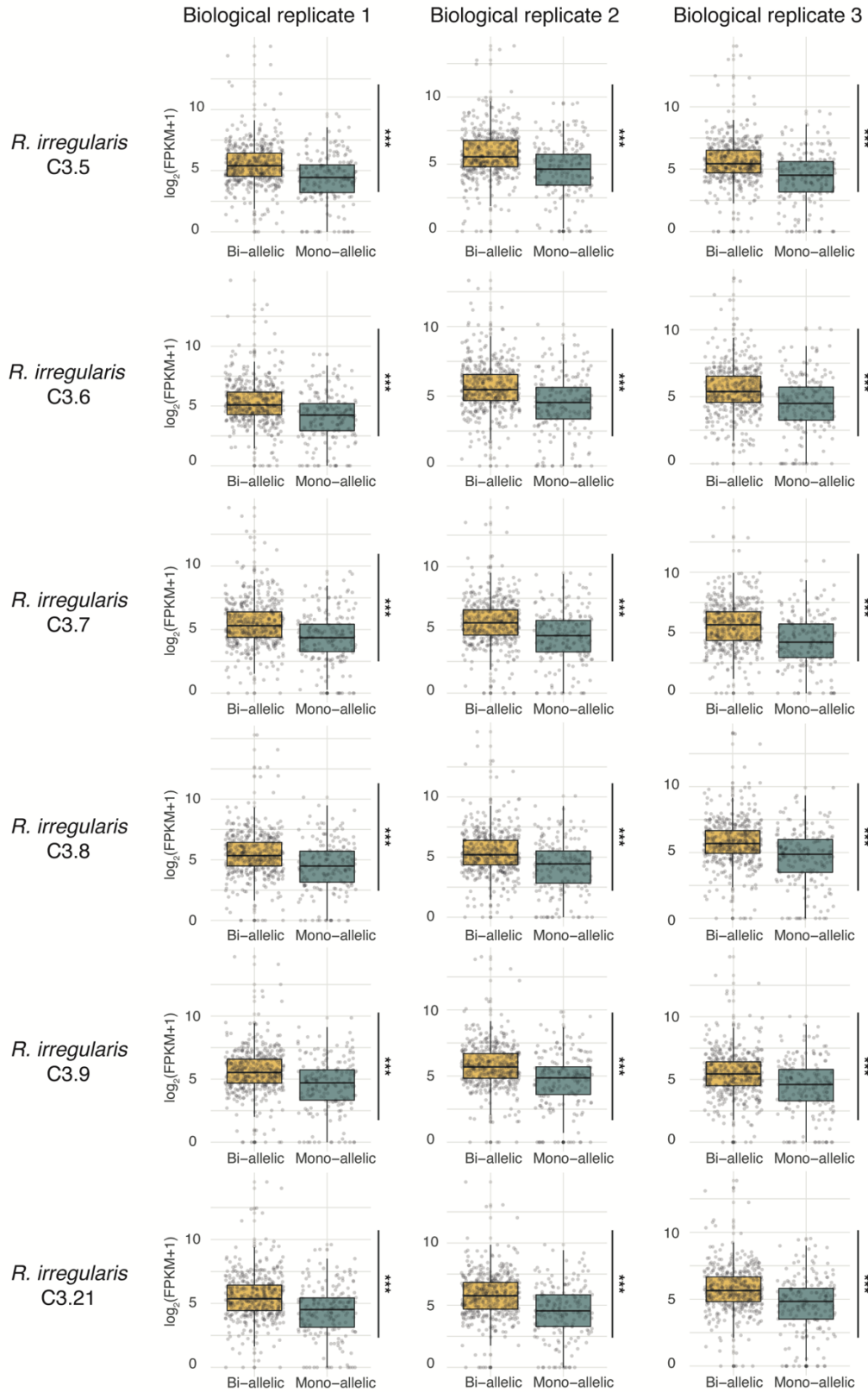


Figure S6: Reproducibility of mono-allelic and bi-allelic expression among three biological replicates of each of the 6 dikaryon single spores sibling

lines (SSSLs) of *R. irregularis* C3. Comparison of the expression levels after a $\log_2(\text{FPKM}+1)$ transformation, between the 187 reproducible genes exhibiting mono-allelic expression and 187 reproducible genes exhibiting bi-allelic expression, for each biological replicate of each *R. irregularis* SSSL. The three asterisks (***) indicate a statistically significant difference between the medians (Mann-Whitney U-test, with a significance threshold of $p < 0.05$). The horizontal line represents the median, the box represents in the interquartile range and the vertical lines represent the maximum and minimum values.