

Mapping quantitative trait loci onto a phylogenetic tree

SUPPLEMENT

Karl W. Broman^{*}, Sungjin Kim[†], Saunak Sen[‡],
Cécile Ané^{†,§}, Bret A. Payseur^{**}

^{*}Department of Biostatistics and Medical Informatics, [†]Department of Statistics, [§]Department of Botany, and ^{**}Laboratory of Genetics, University of Wisconsin--Madison, Madison, Wisconsin 53706; [‡]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California 94107

All partitions

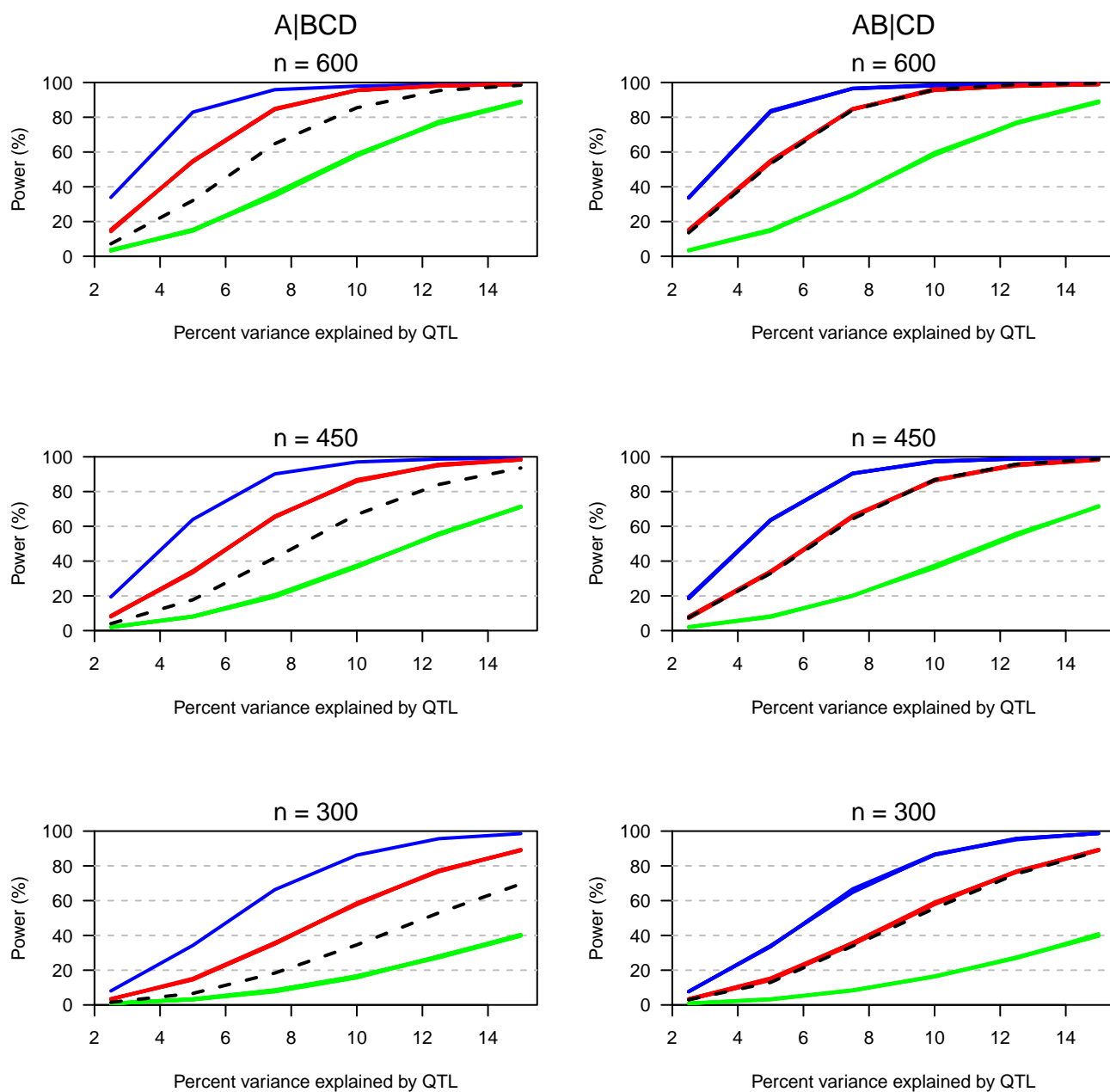


Figure S1: Estimated power in the case of four taxa with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), as a function of the percent phenotypic variance explained by the QTL. The black dashed curves correspond to the use of all six possible crosses. The other curves are for the various choices of a minimal set of three crosses, with the curves in blue, red and green corresponding to cases in which 3, 2 and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering all possible partitions of the taxa.

Tree partitions

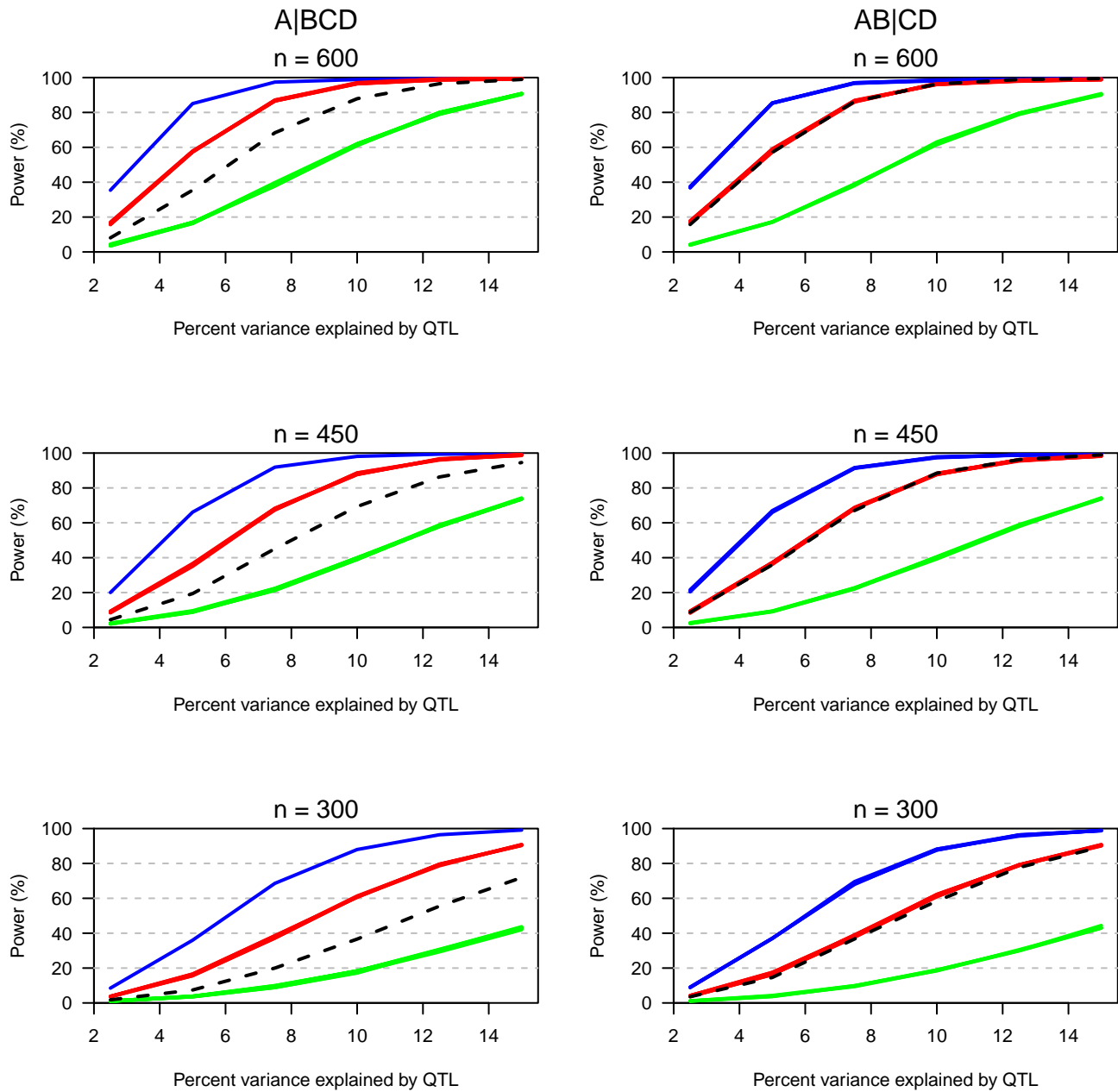


Figure S2: Estimated power in the case of four taxa related as in Figure 1, and with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), as a function of the percent phenotypic variance explained by the QTL. The black dashed curves correspond to the use of all six possible crosses. The other curves are for the various choices of a minimal set of three crosses, with the curves in blue, red and green corresponding to cases in which 3, 2 and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering only the four possible partitions induced by the tree.

All partitions

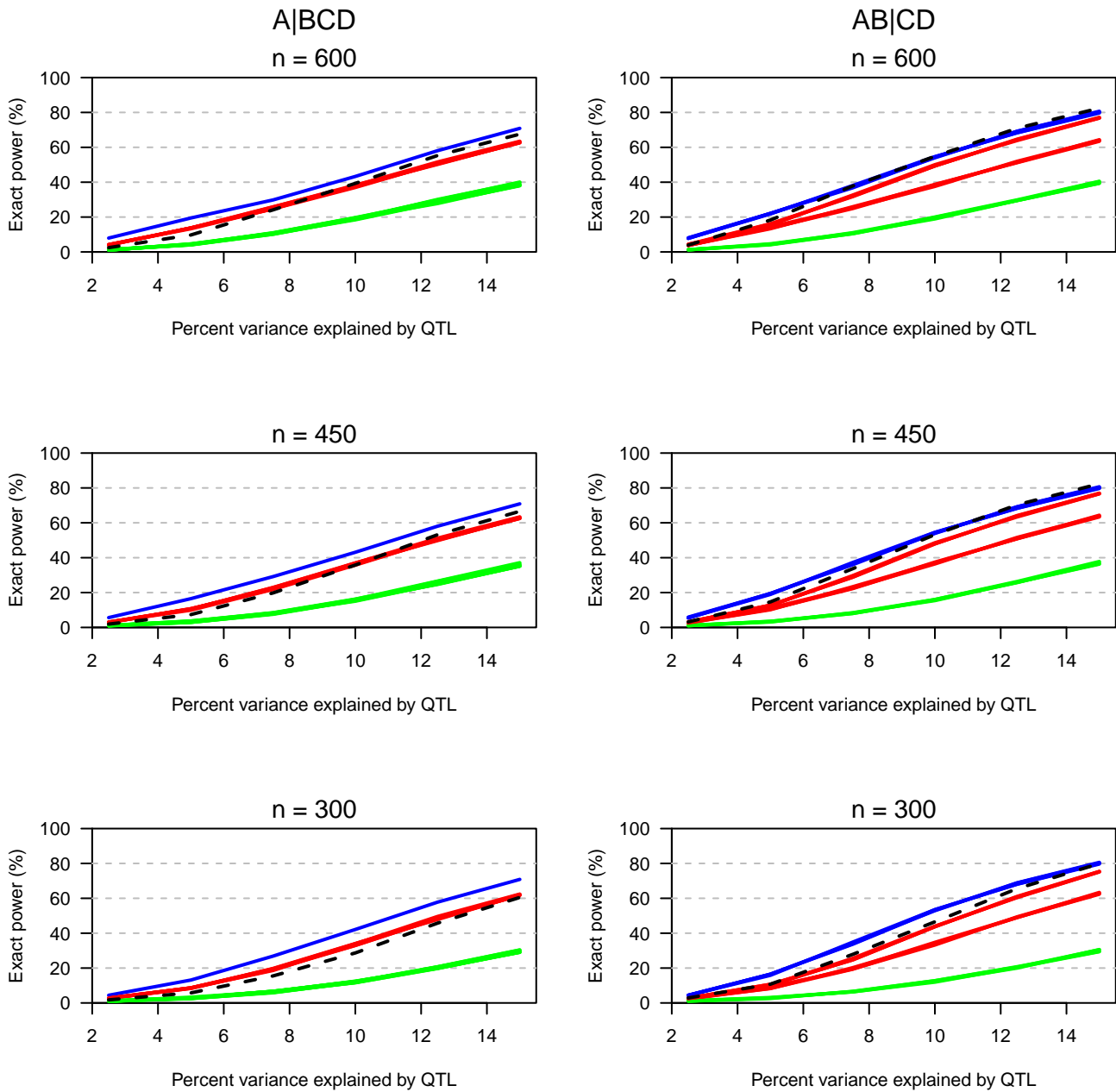


Figure S3: Estimated "exact" power (the chance that a QTL is detected and the credible set of partitions contains only the true partition) in the case of four taxa with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), as a function of the percent phenotypic variance explained by the QTL. The black dashed curves correspond to the use of all six possible crosses. The other curves are for the various choices of a minimal set of three crosses, with the curves in blue, red and green corresponding to cases in which 3, 2 and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering all possible partitions of the taxa.

Tree partitions

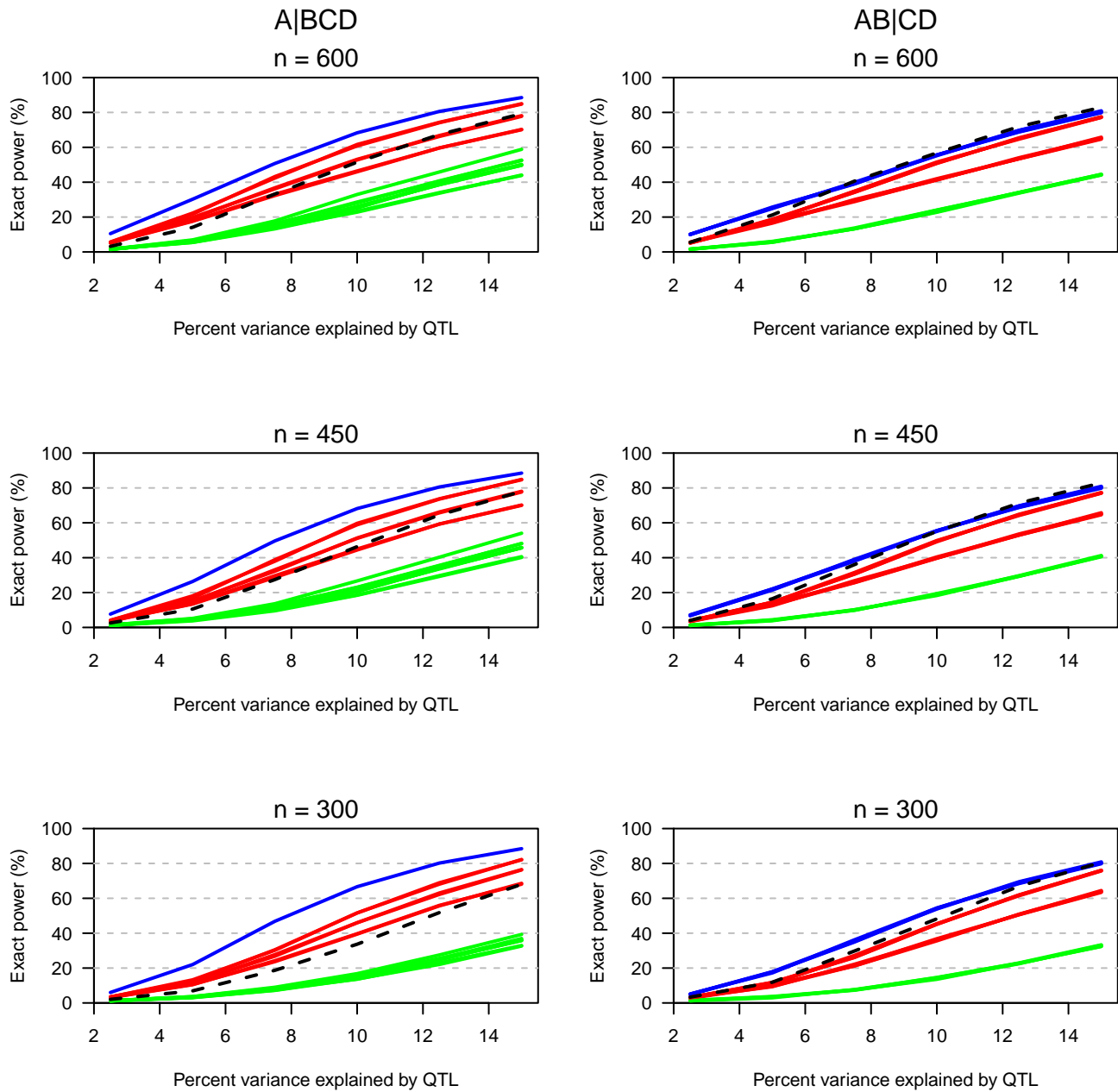


Figure S4: Estimated "exact" power (the chance that a QTL is detected and the credible set of partitions contains only the true partition) in the case of four taxa related as in Figure 1, and with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), as a function of the percent phenotypic variance explained by the QTL. The black dashed curves correspond to the use of all six possible crosses. The other curves are for the various choices of a minimal set of three crosses, with the curves in blue, red and green corresponding to cases in which 3, 2 and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering only the four possible partitions induced by the tree.

All partitions

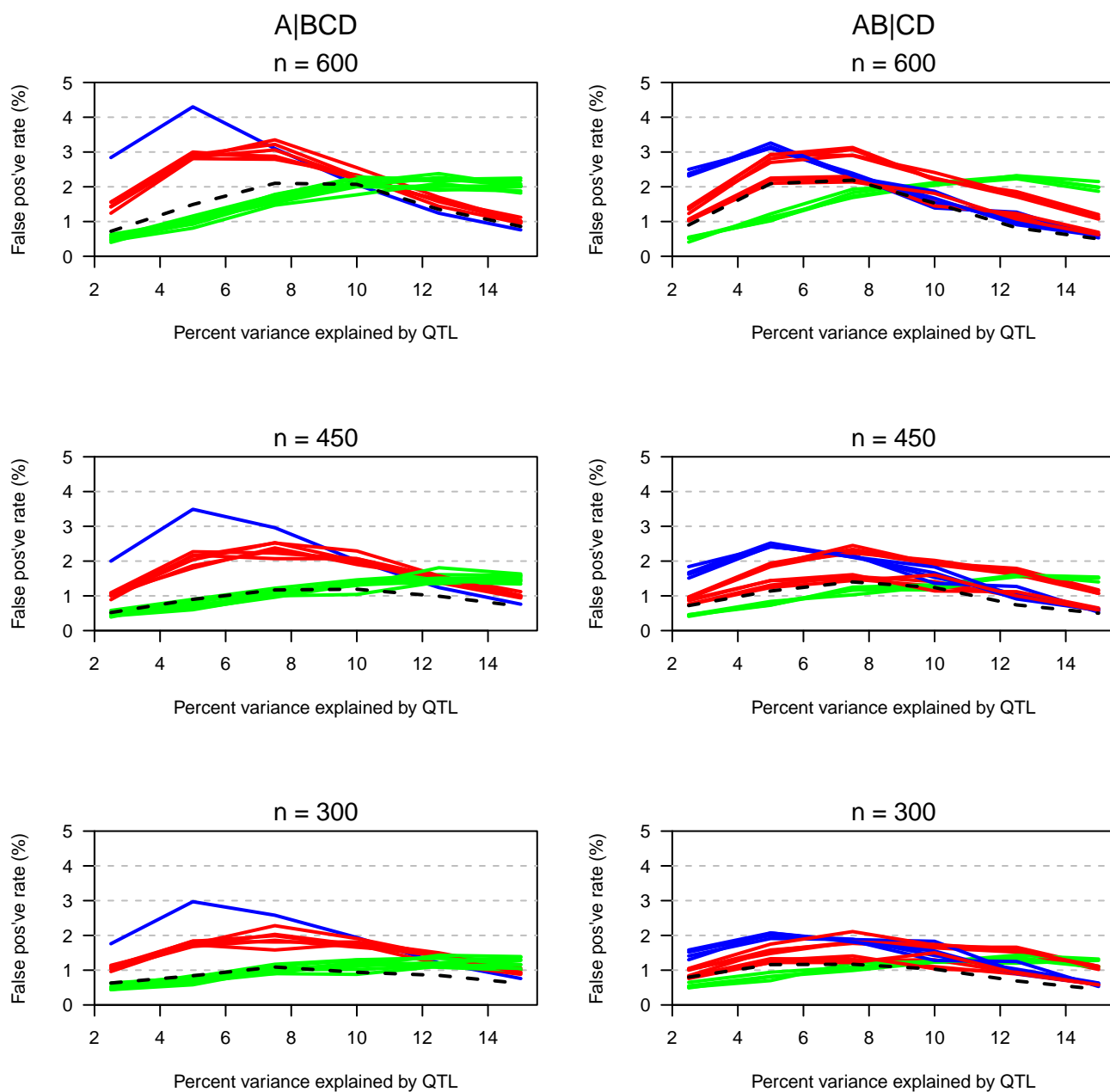


Figure S5: Estimated false positive rate in the case of four taxa with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), as a function of the percent phenotypic variance explained by the QTL. The black dashed curves correspond to the use of all six possible crosses. The other curves are for the various choices of a minimal set of three crosses, with the curves in blue, red and green corresponding to cases in which 3, 2 and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering all possible partitions of the taxa.

Tree partitions

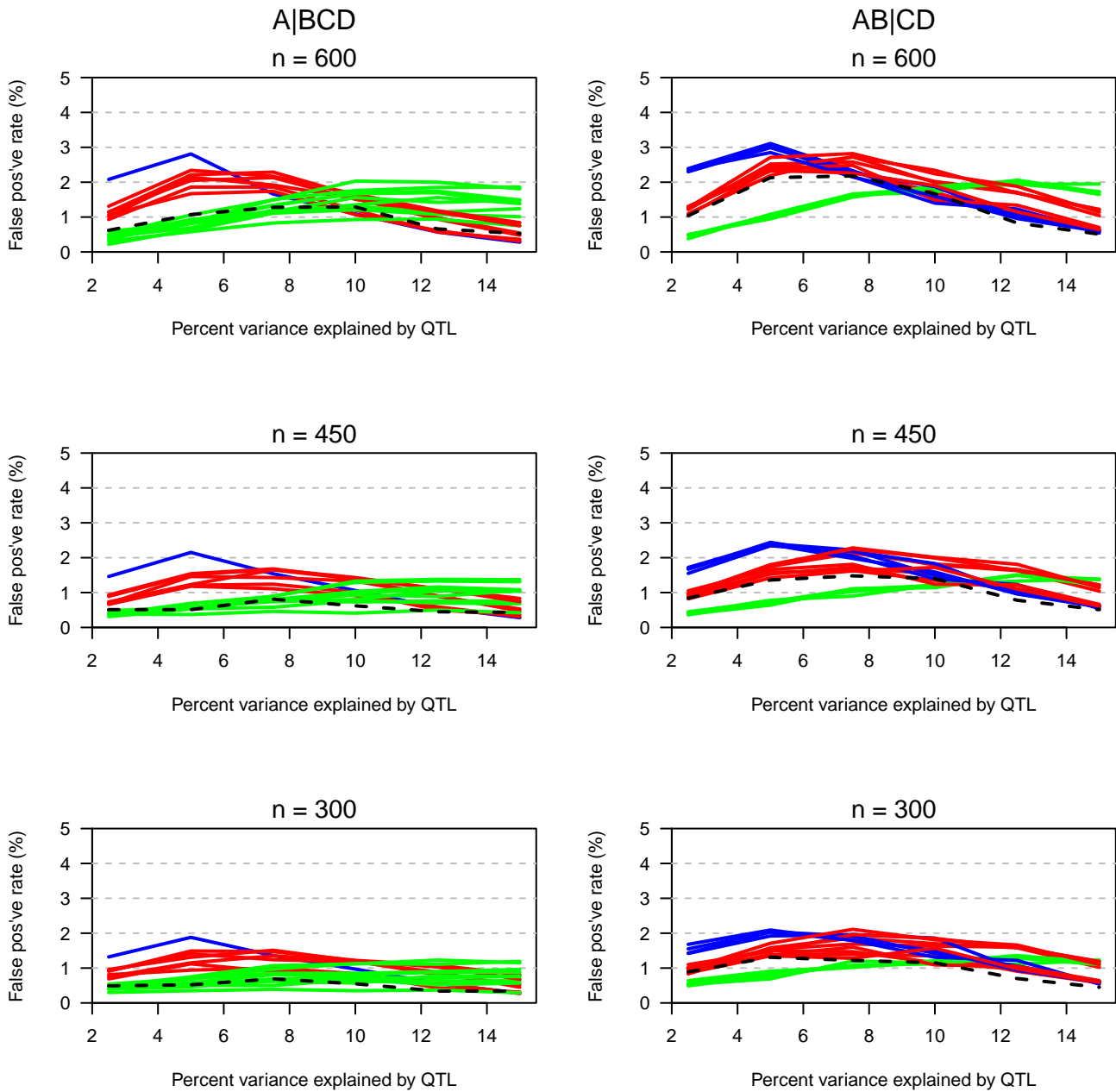


Figure S6: Estimated false positive rate in the case of four taxa related as in Figure 1, and with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), as a function of the percent phenotypic variance explained by the QTL. The black dashed curves correspond to the use of all six possible crosses. The other curves are for the various choices of a minimal set of three crosses, with the curves in blue, red and green corresponding to cases in which 3, 2 and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering only the four possible partitions induced by the tree.

All partitions

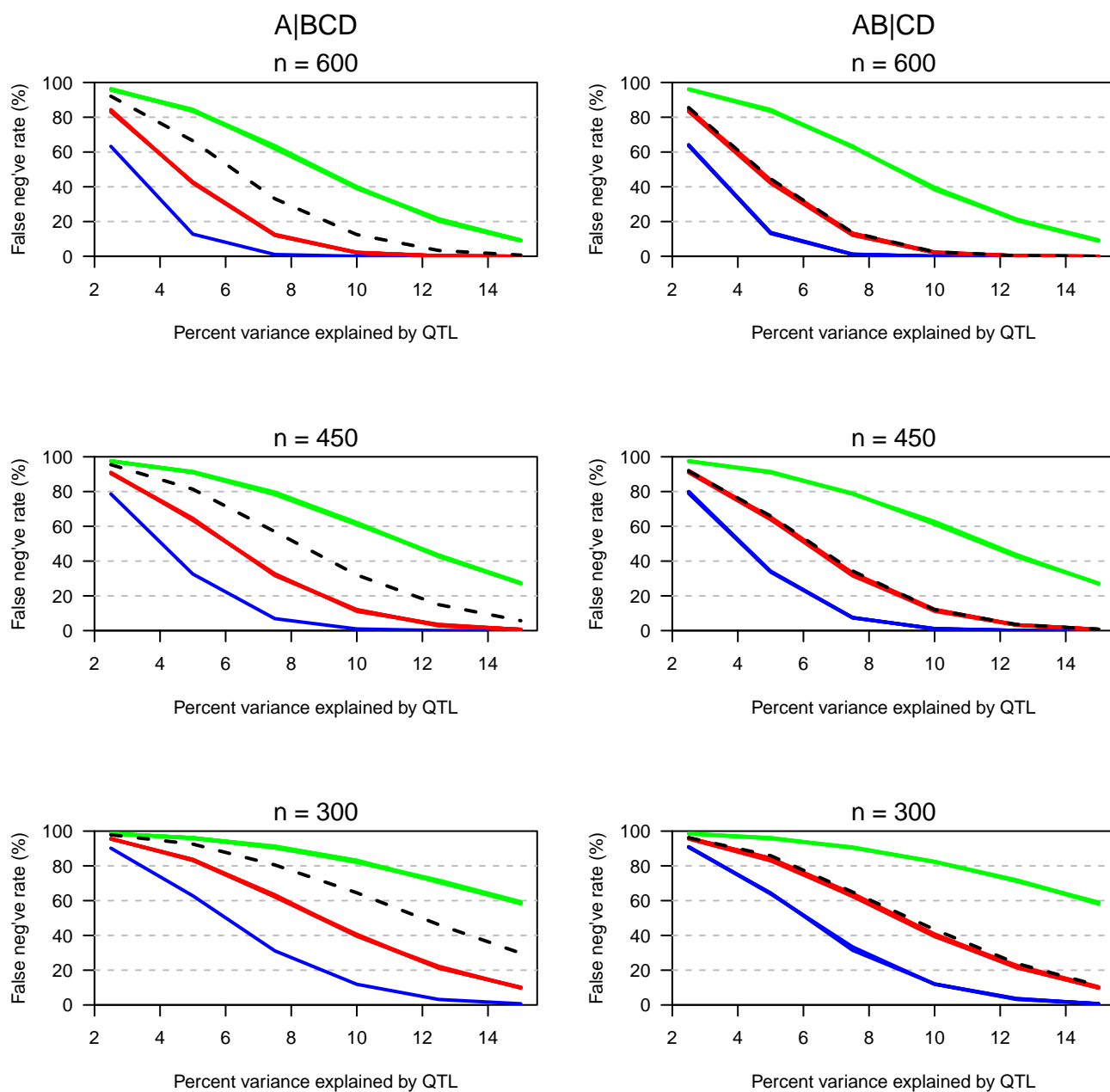


Figure S7: Estimated false negative rate in the case of four taxa with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), as a function of the percent phenotypic variance explained by the QTL. The black dashed curves correspond to the use of all six possible crosses. The other curves are for the various choices of a minimal set of three crosses, with the curves in blue, red and green corresponding to cases in which 3, 2 and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering all possible partitions of the taxa.

Tree partitions

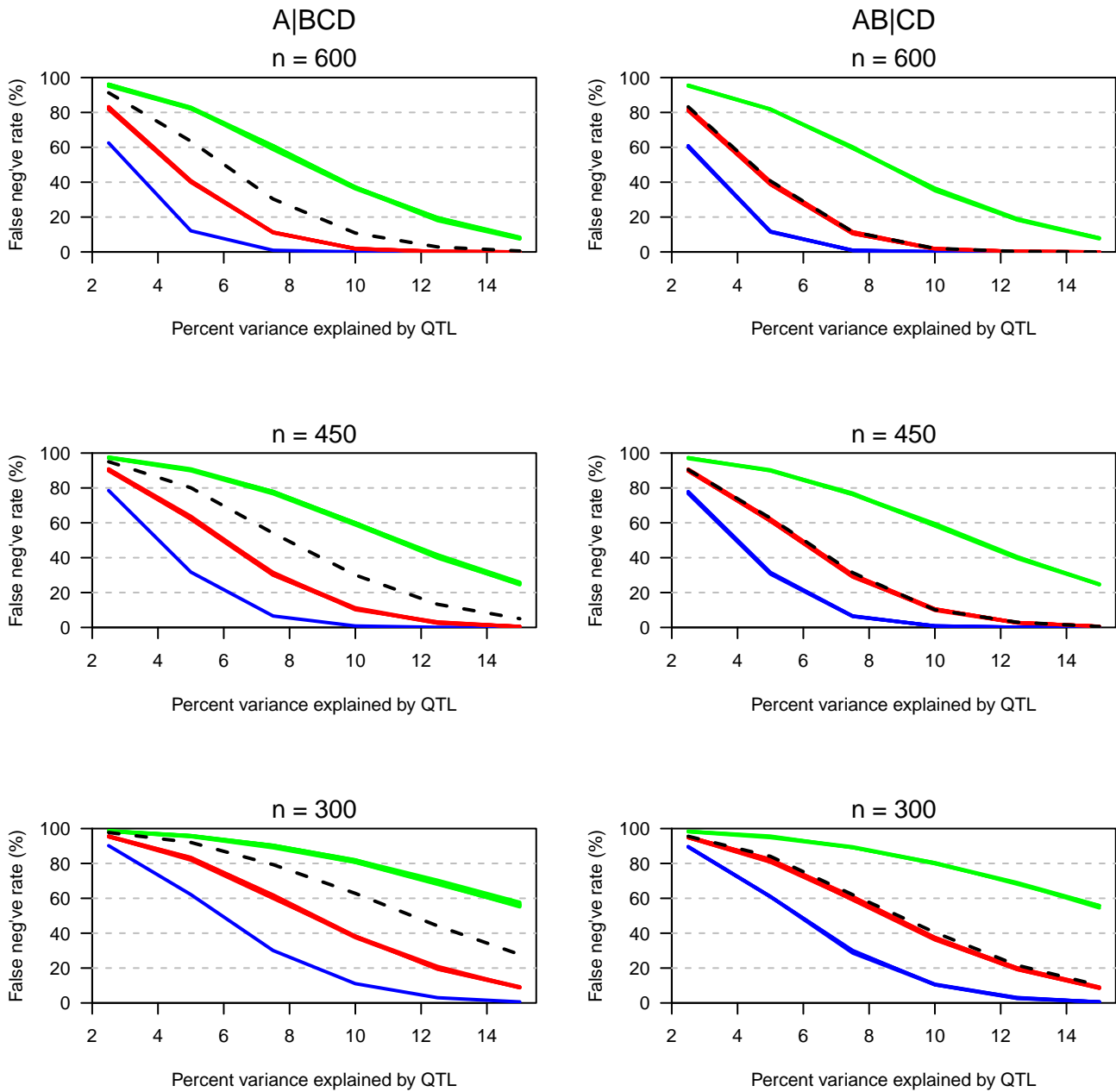


Figure S8: Estimated false negative rate in the case of four taxa related as in Figure 1, and with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), as a function of the percent phenotypic variance explained by the QTL. The black dashed curves correspond to the use of all six possible crosses. The other curves are for the various choices of a minimal set of three crosses, with the curves in blue, red and green corresponding to cases in which 3, 2 and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering only the four possible partitions induced by the tree.

All partitions

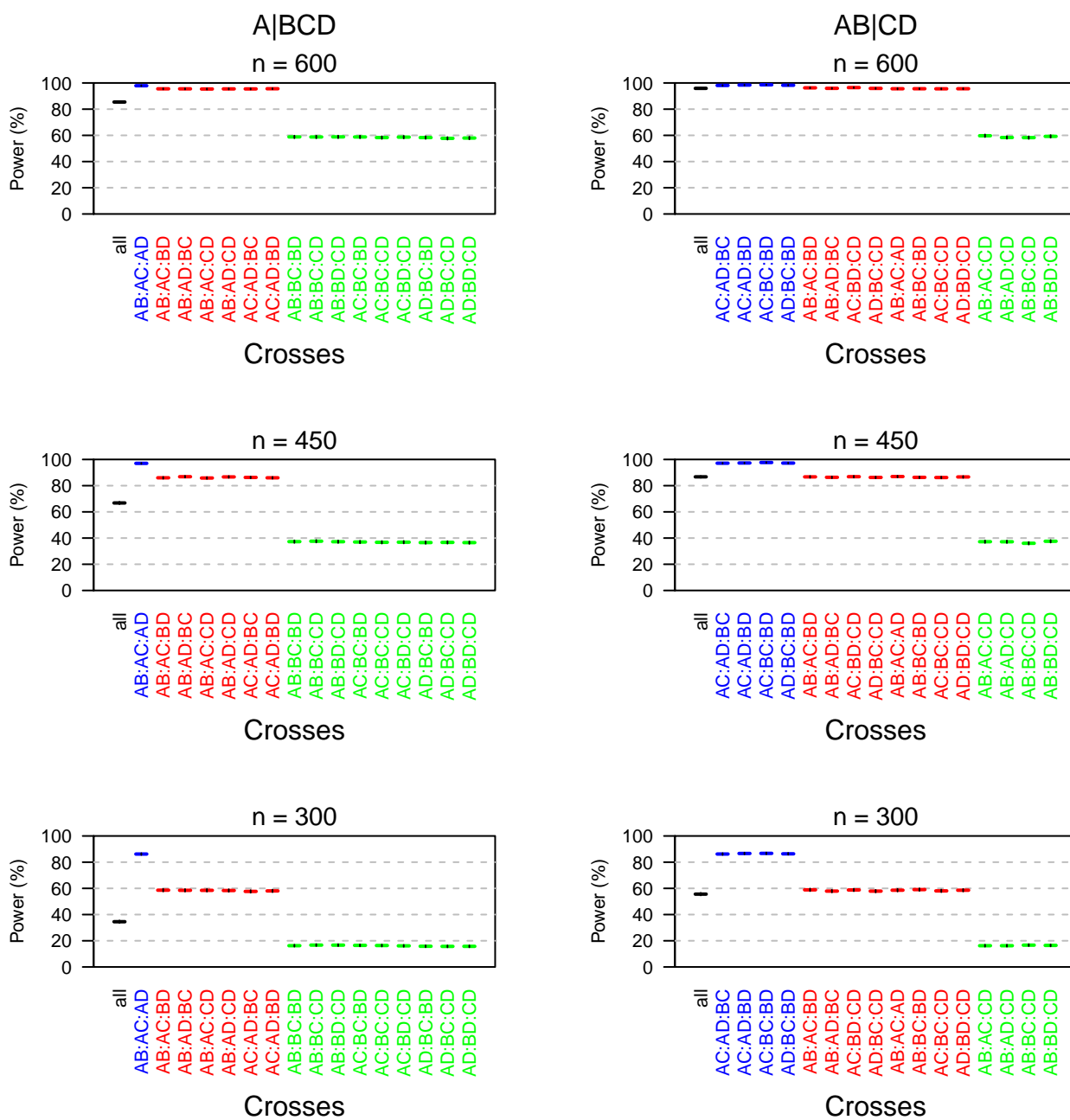


Figure S9: Detailed results on the estimated power for individual choices of crosses, in the case of four taxa with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), and with the QTL being responsible for 10% of the phenotypic variance in crosses in which it is segregating. Blue, red and green correspond to cases in which 3, 2, and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering all possible partitions of the taxa. The black vertical line segments indicate 95% confidence intervals.

Tree partitions

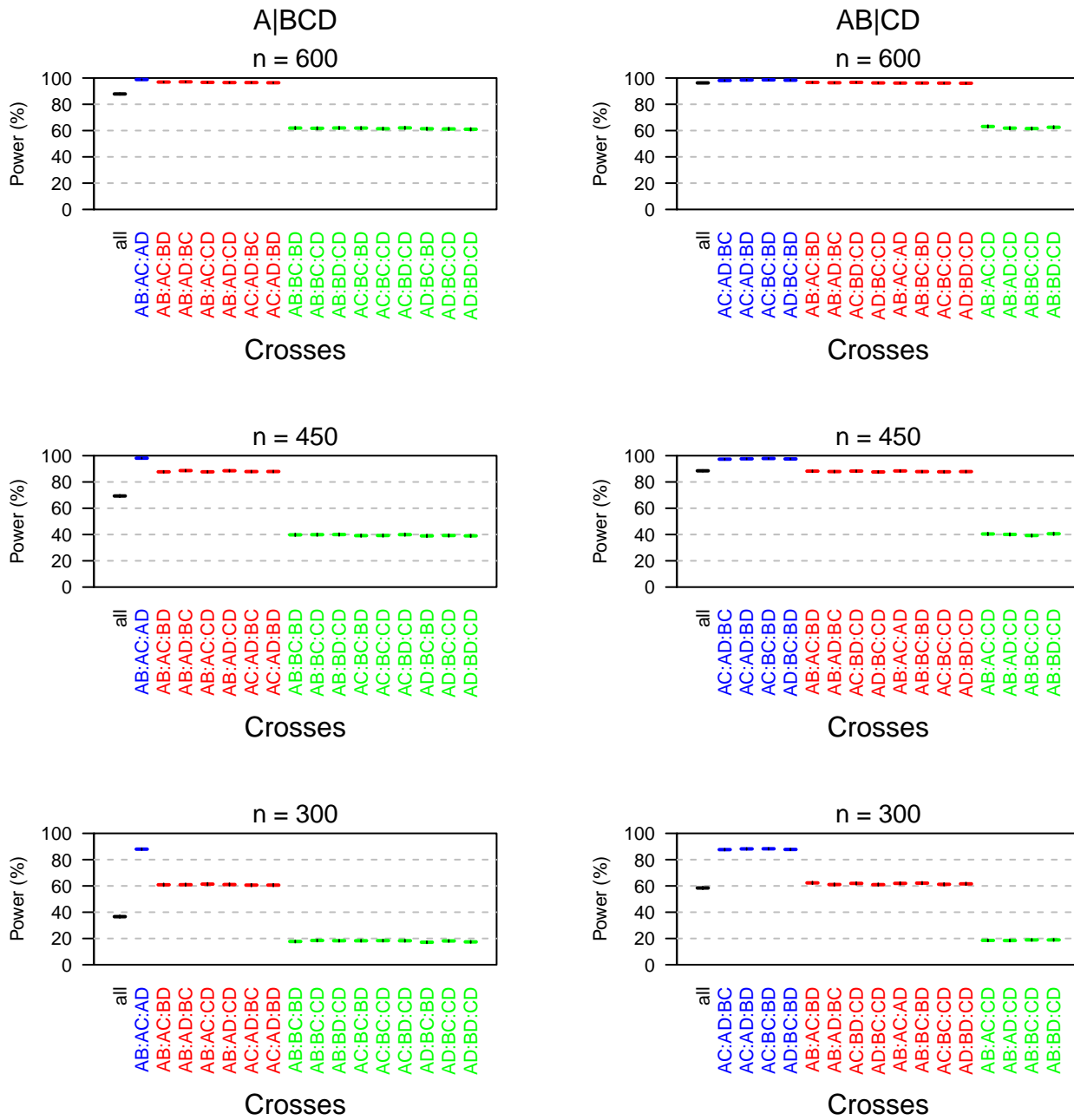


Figure S10: Detailed results on the estimated power for individual choices of crosses, in the case of four taxa related as in Figure 1, with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), and with the QTL being responsible for 10% of the phenotypic variance in crosses in which it is segregating. Blue, red and green correspond to cases in which 3, 2, and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering only the four possible partitions induced by the tree. The black vertical line segments indicate 95% confidence intervals.

All partitions

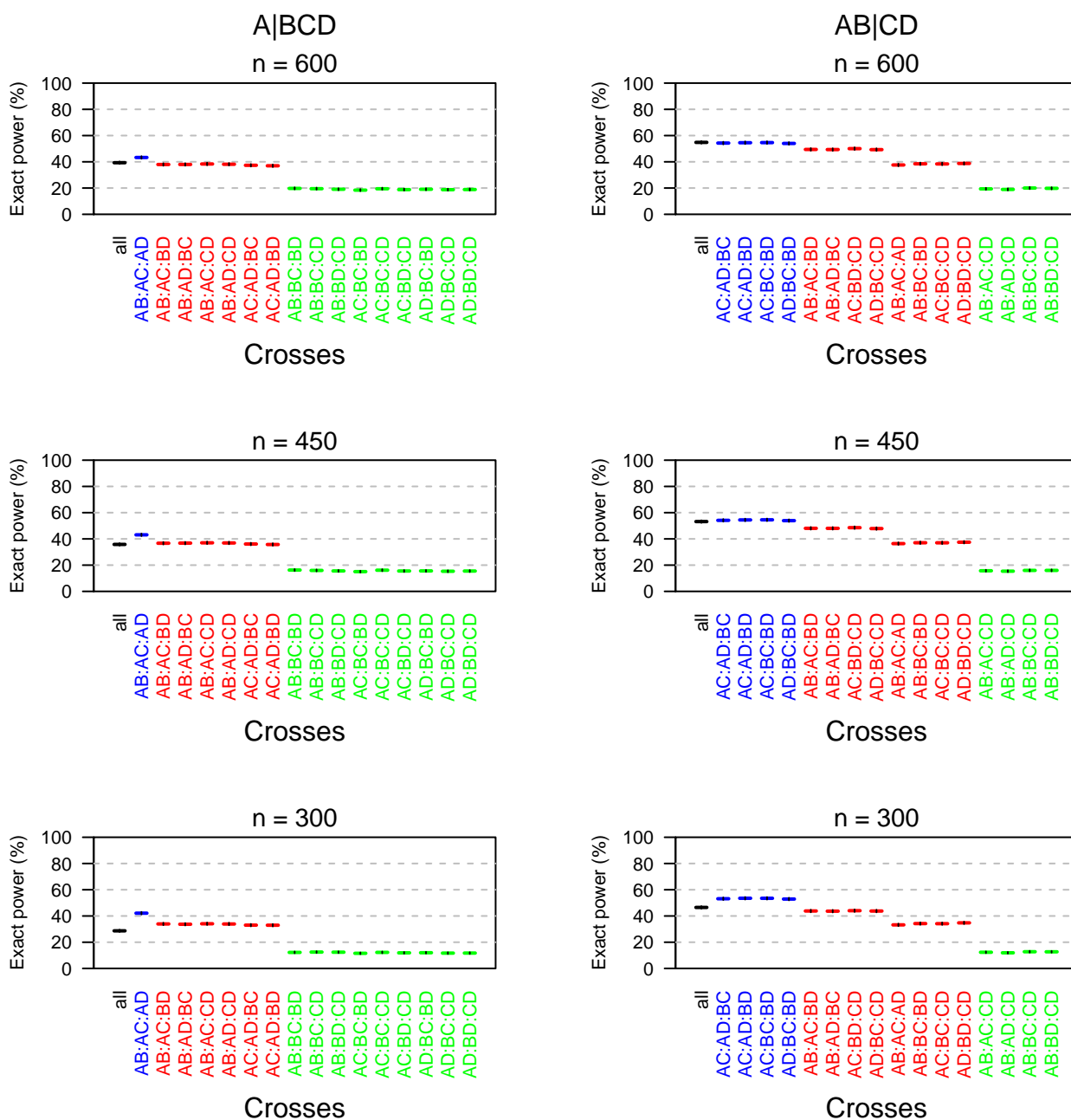


Figure S11: Detailed results on the estimated "exact" power (the chance that a QTL is detected and the credible set of partitions contains only the true partition) for individual choices of crosses, in the case of four taxa with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), and with the QTL being responsible for 10% of the phenotypic variance in crosses in which it is segregating. Blue, red and green correspond to cases in which 3, 2, and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering all possible partitions of the taxa. The black vertical line segments indicate 95% confidence intervals.

Tree partitions

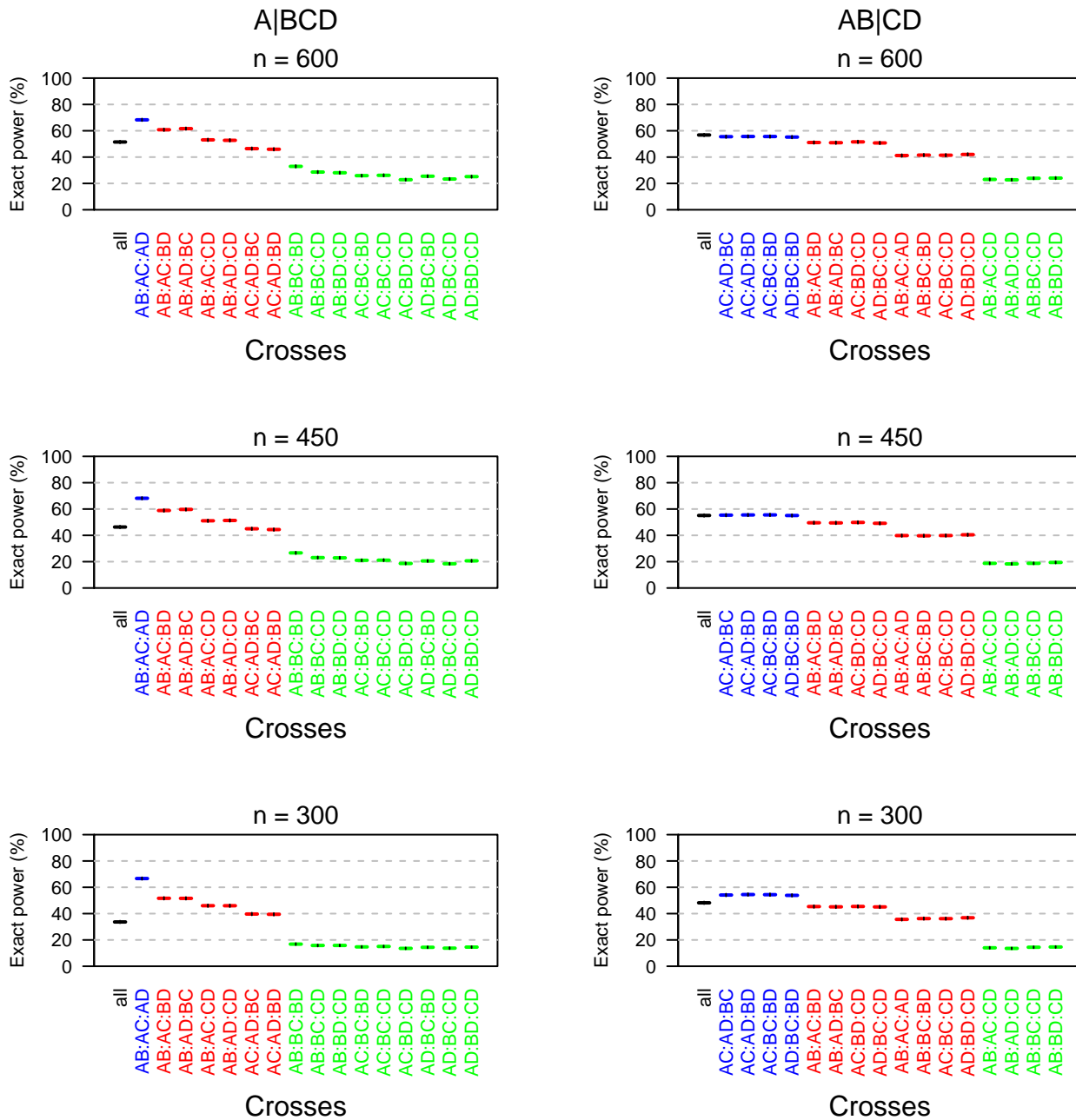


Figure S12: Detailed results on the estimated "exact" power (the chance that a QTL is detected and the credible set of partitions contains only the true partition) for individual choices of crosses, in the case of four taxa related as in Figure 1, with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), and with the QTL being responsible for 10% of the phenotypic variance in crosses in which it is segregating. Blue, red and green correspond to cases in which 3, 2, and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering only the four possible partitions induced by the tree. The black vertical line segments indicate 95% confidence intervals.

All partitions

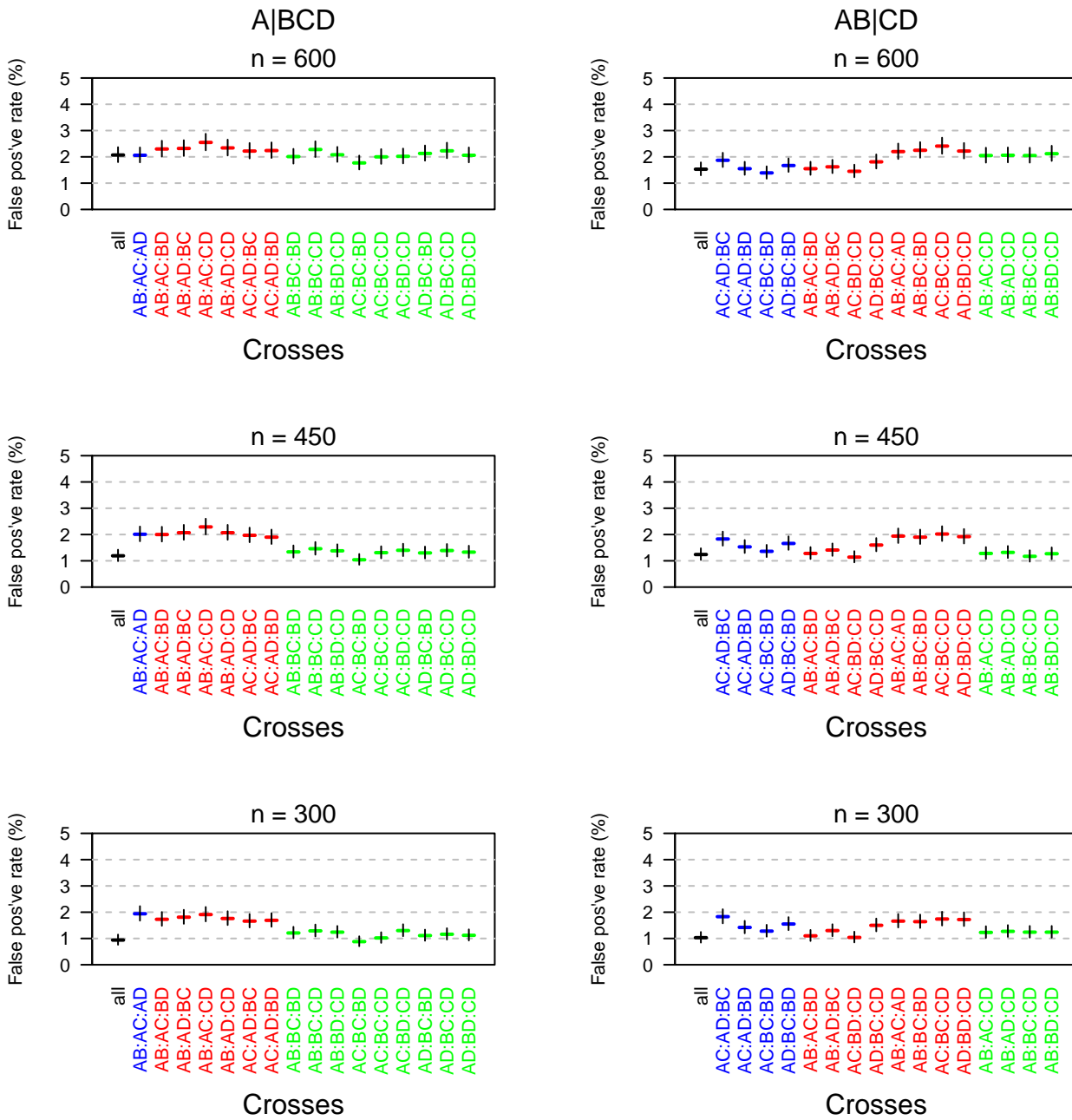


Figure S13: Detailed results on the estimated false positive rate for individual choices of crosses, in the case of four taxa with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), and with the QTL being responsible for 10% of the phenotypic variance in crosses in which it is segregating. Blue, red and green correspond to cases in which 3, 2, and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering all possible partitions of the taxa. The black vertical line segments indicate 95% confidence intervals.

Tree partitions

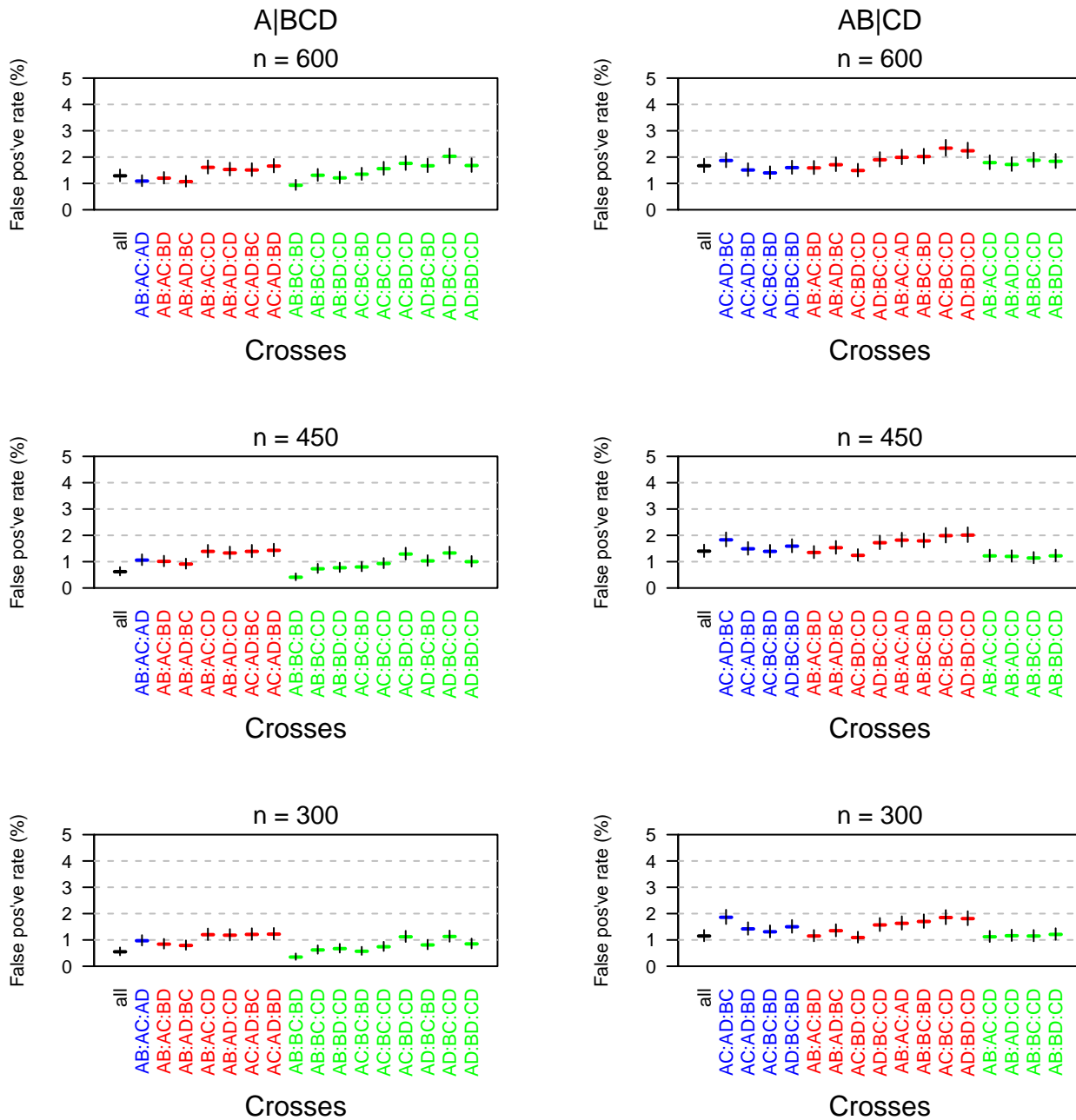


Figure S14: Detailed results on the estimated false positive rate for individual choices of crosses, in the case of four taxa related as in Figure 1, with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), and with the QTL being responsible for 10% of the phenotypic variance in crosses in which it is segregating. Blue, red and green correspond to cases in which 3, 2, and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering only the four possible partitions induced by the tree. The black vertical line segments indicate 95% confidence intervals.

All partitions

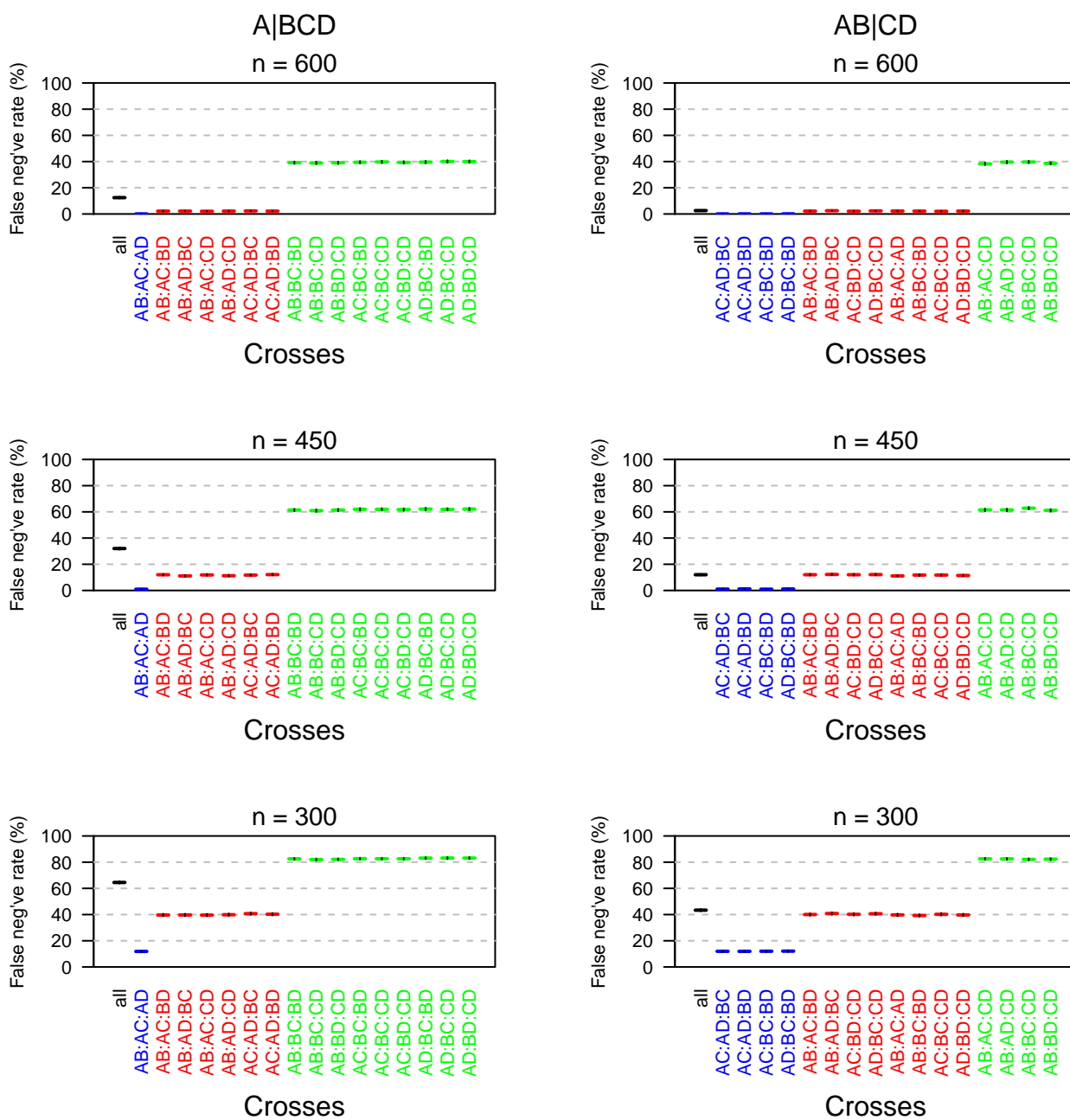


Figure S15: Detailed results on the estimated false negative rate for individual choices of crosses, in the case of four taxa with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), and with the QTL being responsible for 10% of the phenotypic variance in crosses in which it is segregating. Blue, red and green correspond to cases in which 3, 2, and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering all possible partitions of the taxa. The black vertical line segments indicate 95% confidence intervals.

Tree partitions

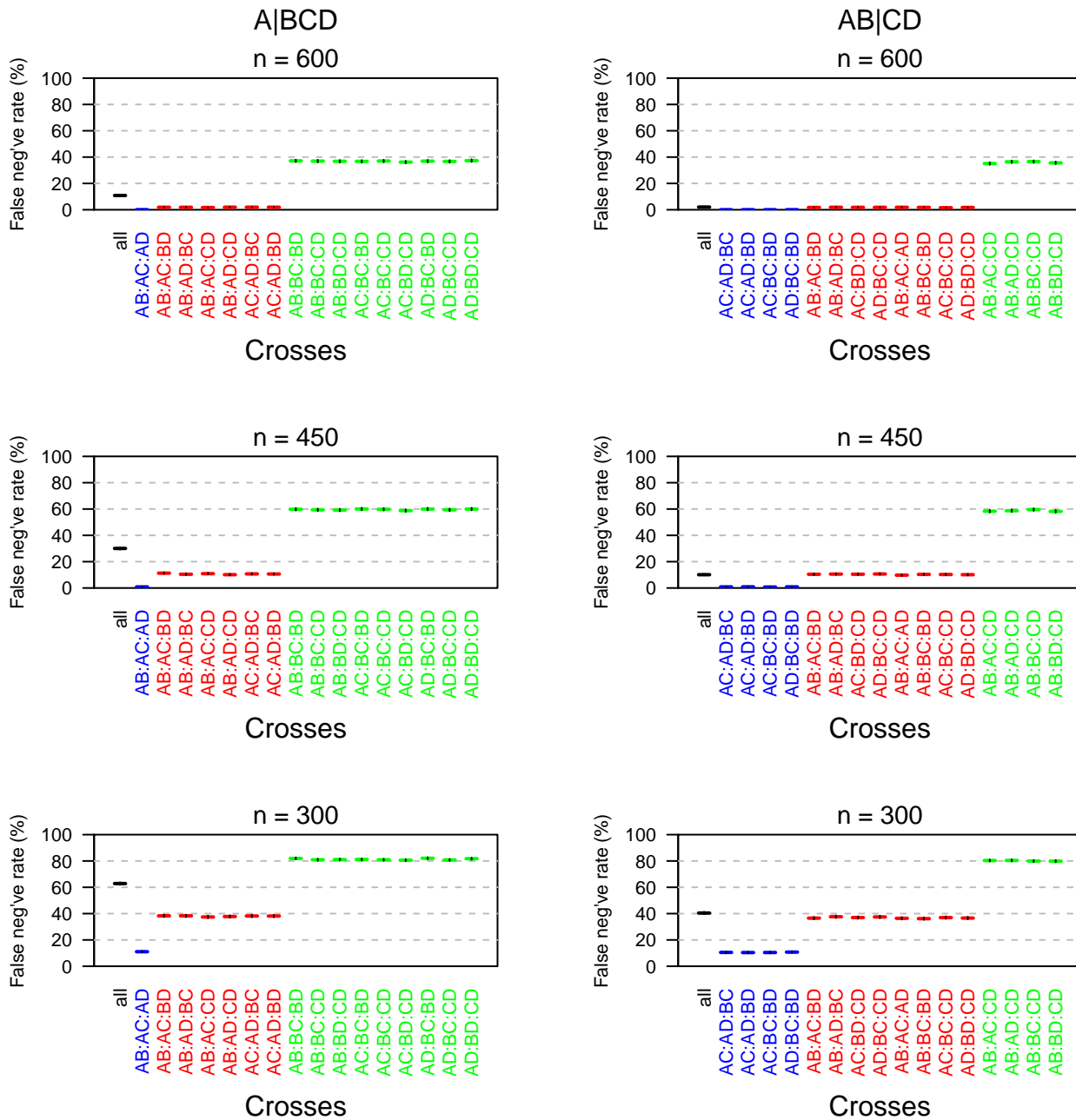


Figure S16: Detailed results on the estimated false negative rate for individual choices of crosses, in the case of four taxa related as in Figure 1, with a total sample size of either 300 (bottom panels) 450 (middle panels) or 600 (top panels), and with the QTL being responsible for 10% of the phenotypic variance in crosses in which it is segregating. Blue, red and green correspond to cases in which 3, 2, and 1 of the crosses are segregating the QTL, respectively. The results are based on 10,000 simulation replicates, with analyses considering only the four possible partitions induced by the tree. The black vertical line segments indicate 95% confidence intervals.

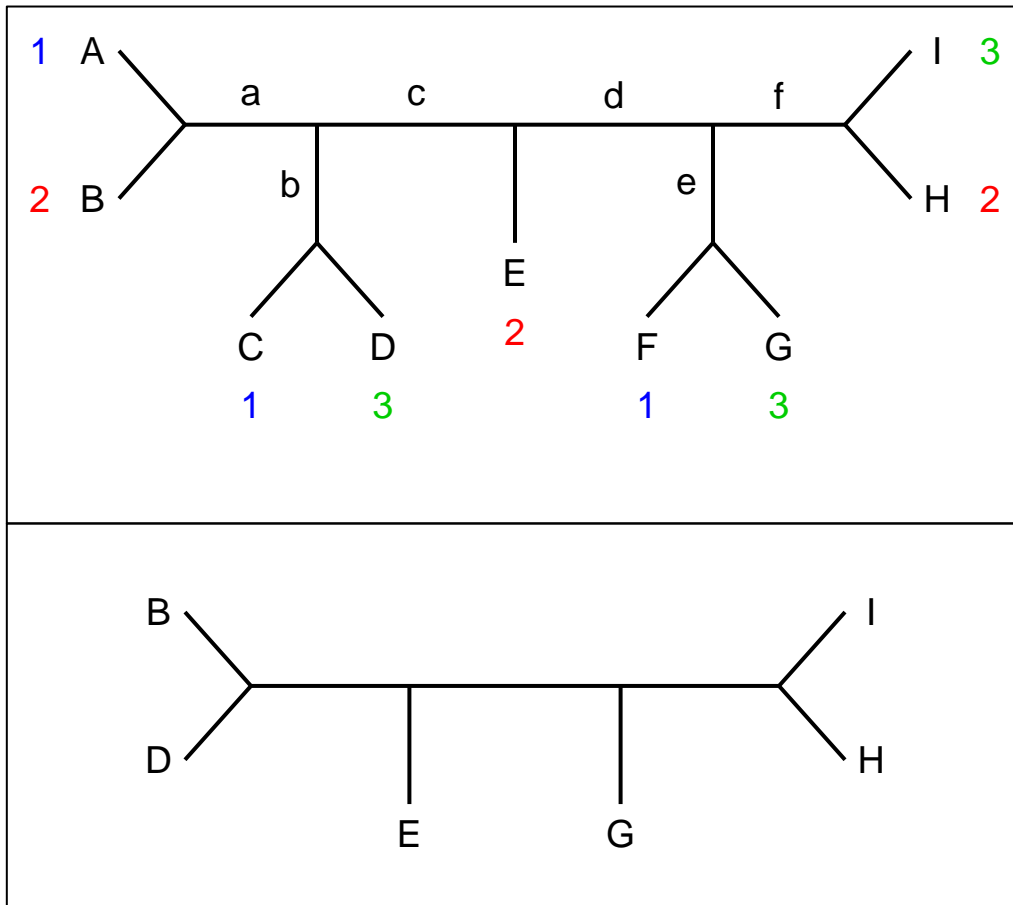


Figure S17: Illustration of the construction of a minimal set of crosses to distinguish the partitions induced by a tree. In the upper panel, the nine taxa (indicated with upper-case letters) are divided into three groups (indicated with numbers) by the method described in the Appendix. Internal edges are indicated with lower-case letters. The lower panel contains the subtree of six taxa defined by excluding the first group of three taxa (A, C, and F). The method described in the Appendix identifies the set $A \times C$, $C \times F$, $B \times E$, $E \times H$, $D \times G$, and $G \times I$ as six crosses sufficient to distinguish all 15 partitions induced by the tree.

Table S1: Estimated 5% genome-wide significance thresholds, based on 10,000 simulation replicates, for a single intercross. We assumed an autosomal genome modeled after the mouse, with genetic markers at a 10 cM spacing.

sample size	threshold
50	3.80
75	3.72
100	3.65

Table S2: Estimated 5% genome-wide significance thresholds for the maximum LOD score across partitions in the case of four taxa, based on 10,000 simulation replicates. We assumed an autosomal genome modeled after the mouse, with genetic markers at a 10 cM spacing.

total sample size	All partitions		Tree partitions	
	all crosses	minimal crosses	all crosses	minimal crosses
300	4.56	4.48	4.43	4.33
450	4.51	4.47	4.36	4.33
600	4.49	4.44	4.32	4.29