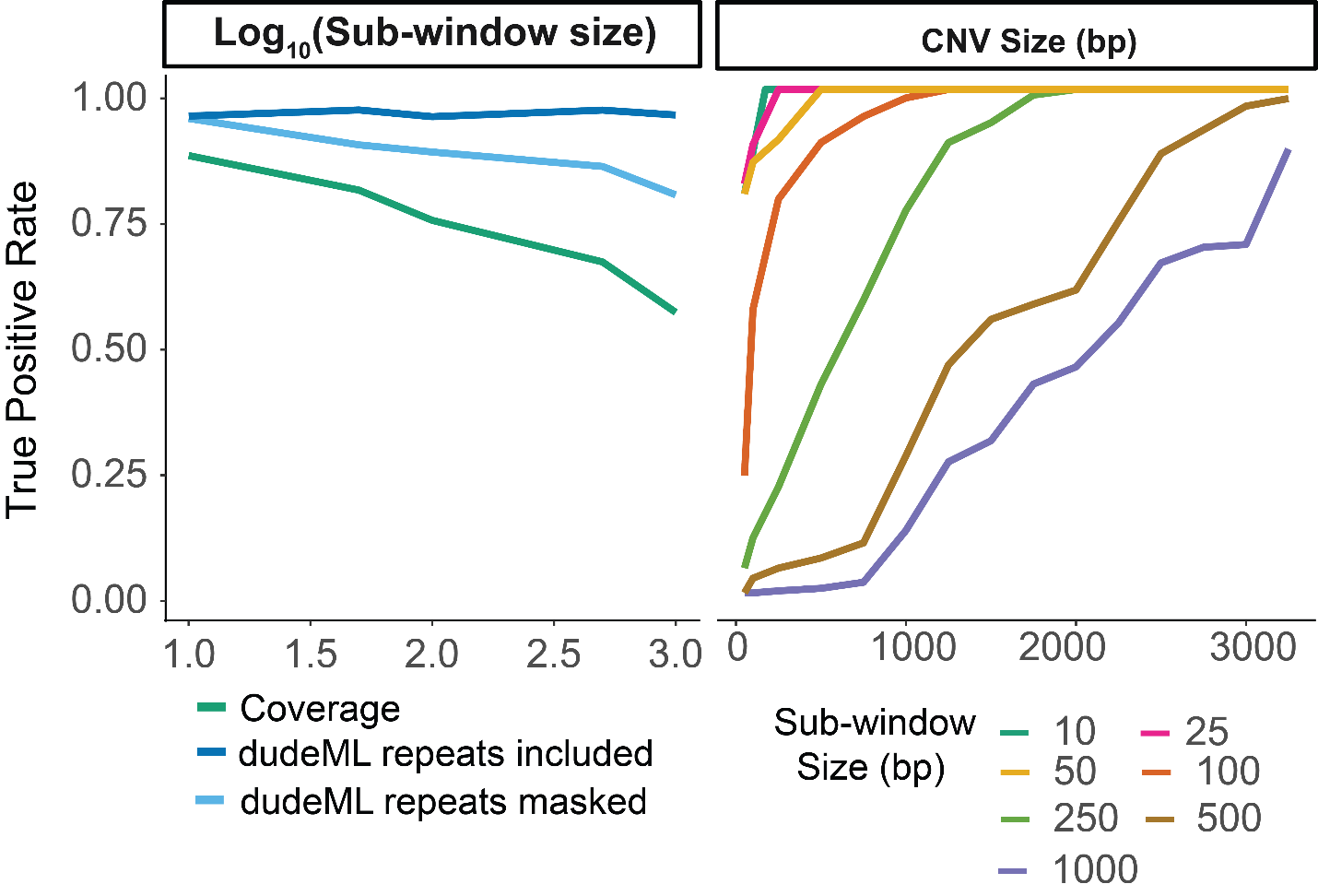
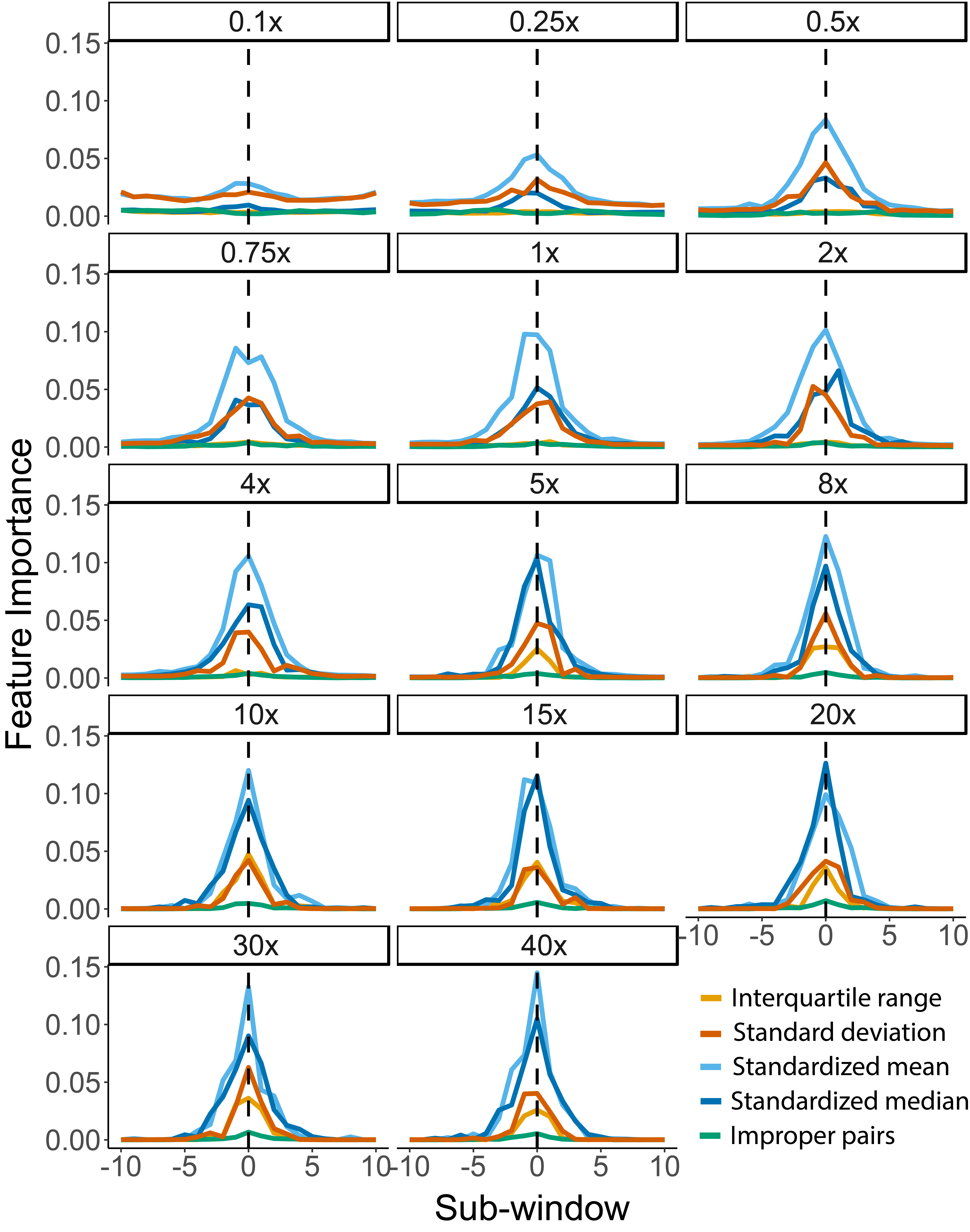
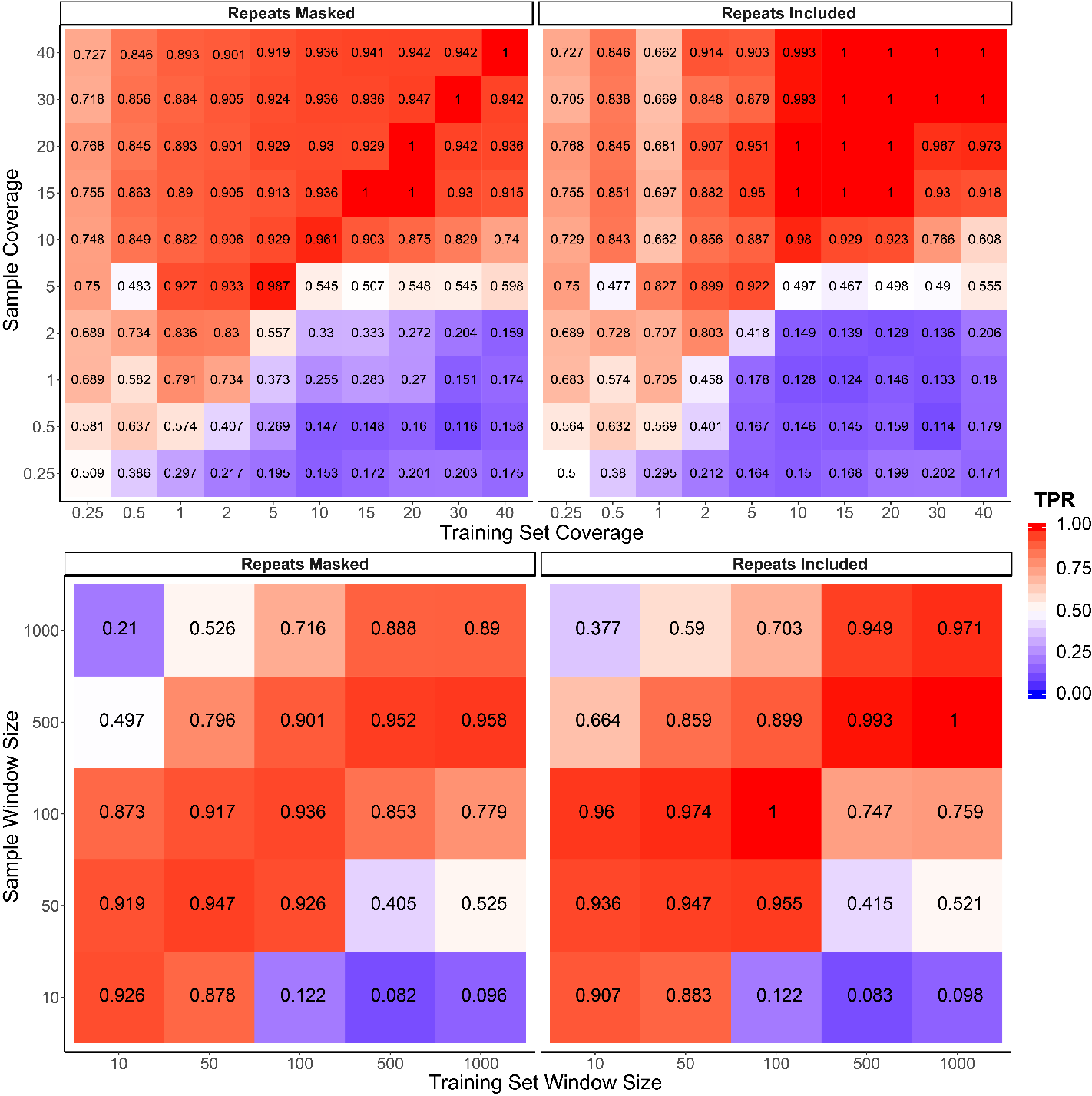
**Figure S1:** Effect of sub-window size on detection of CNVs for dudeML and coverage. In dudeML as window size increases, detection rate decreases due to smaller CNVs being averaged over and missed, this is seen in the second plot where smaller sub-windows have a higher true-positive rate for smaller CNVs.



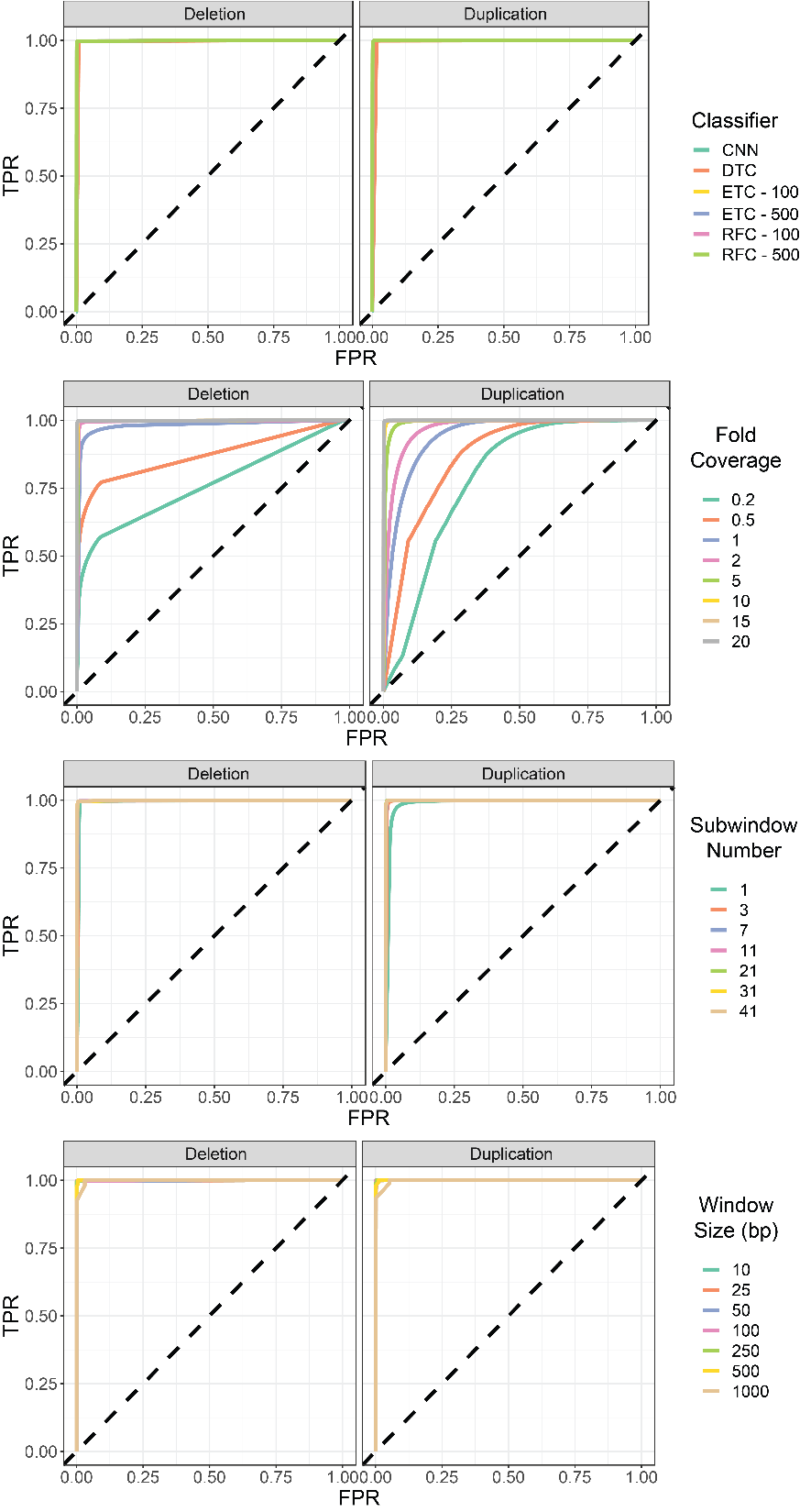
**Figure S2:** Feature importance for detection of CNVs as coverage decreases. As coverage decreases, the focal standardized median of each window decreases in importance and the standardized mean and standard deviation of surrounding windows becomes more important.



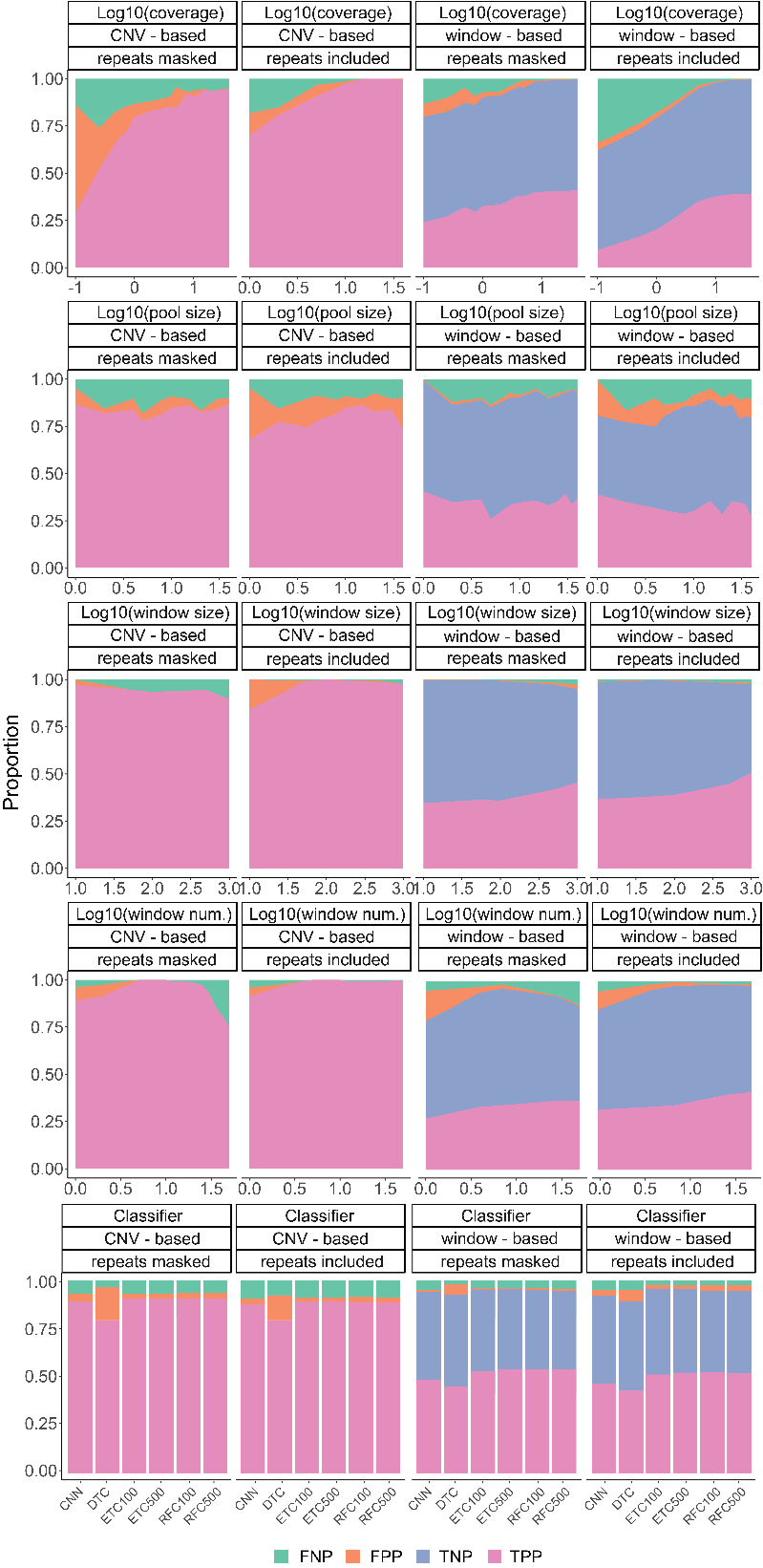
**Figure S3:** True-Positive rates (TPR) of mis-specified training sets across different fold-coverage samples and classifiers, and different window sizes in samples and classifiers.

****

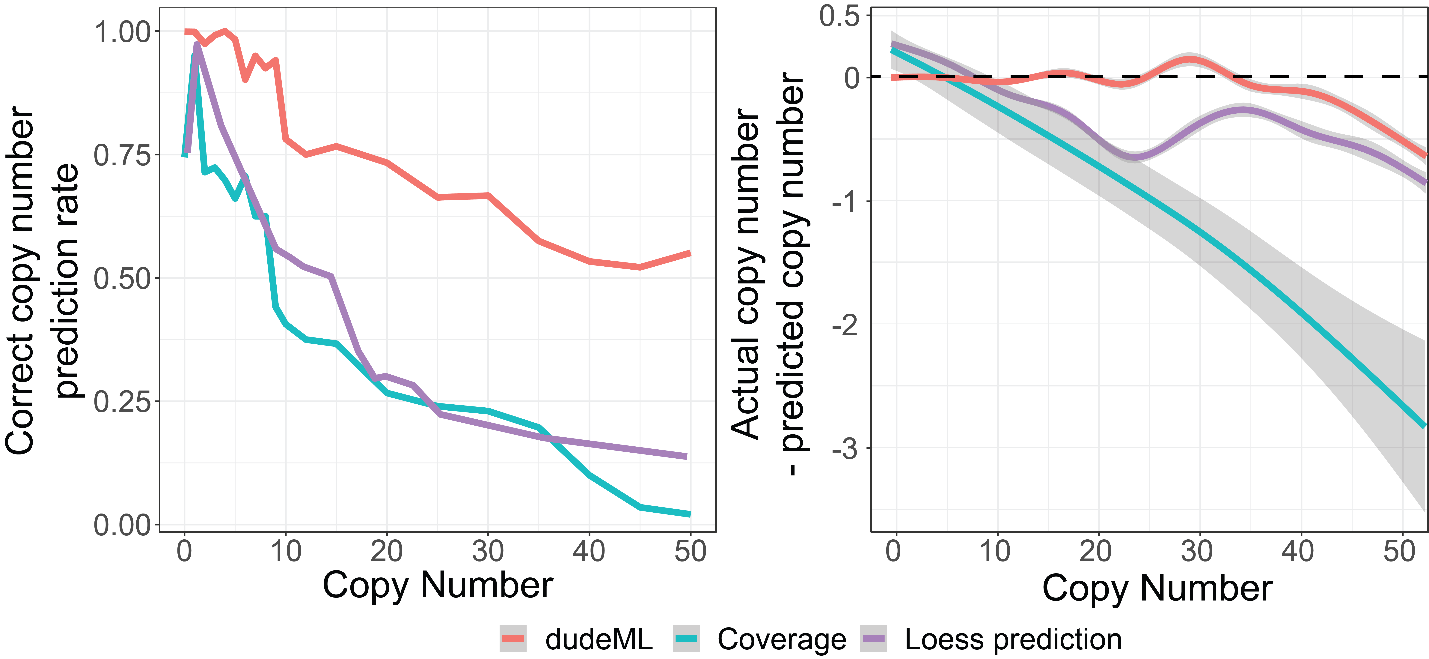
**Figure S4:** Receiver operating characteristic (ROC) curves for correctly detecting duplications and deletions across different classifiers, sample coverages, sub-window numbers and window-sizes (denoted by line color). Classifiers used as follows: convolutional neural network (CNN), decision tree classifier (DTC), extra trees classifier with 100 estimators (ETC100), extra trees classifier with 500 estimators (ETC500), random forest classifier with 100 estimators (RFC100), random forest classifier with 500 estimators (RFC500).

****

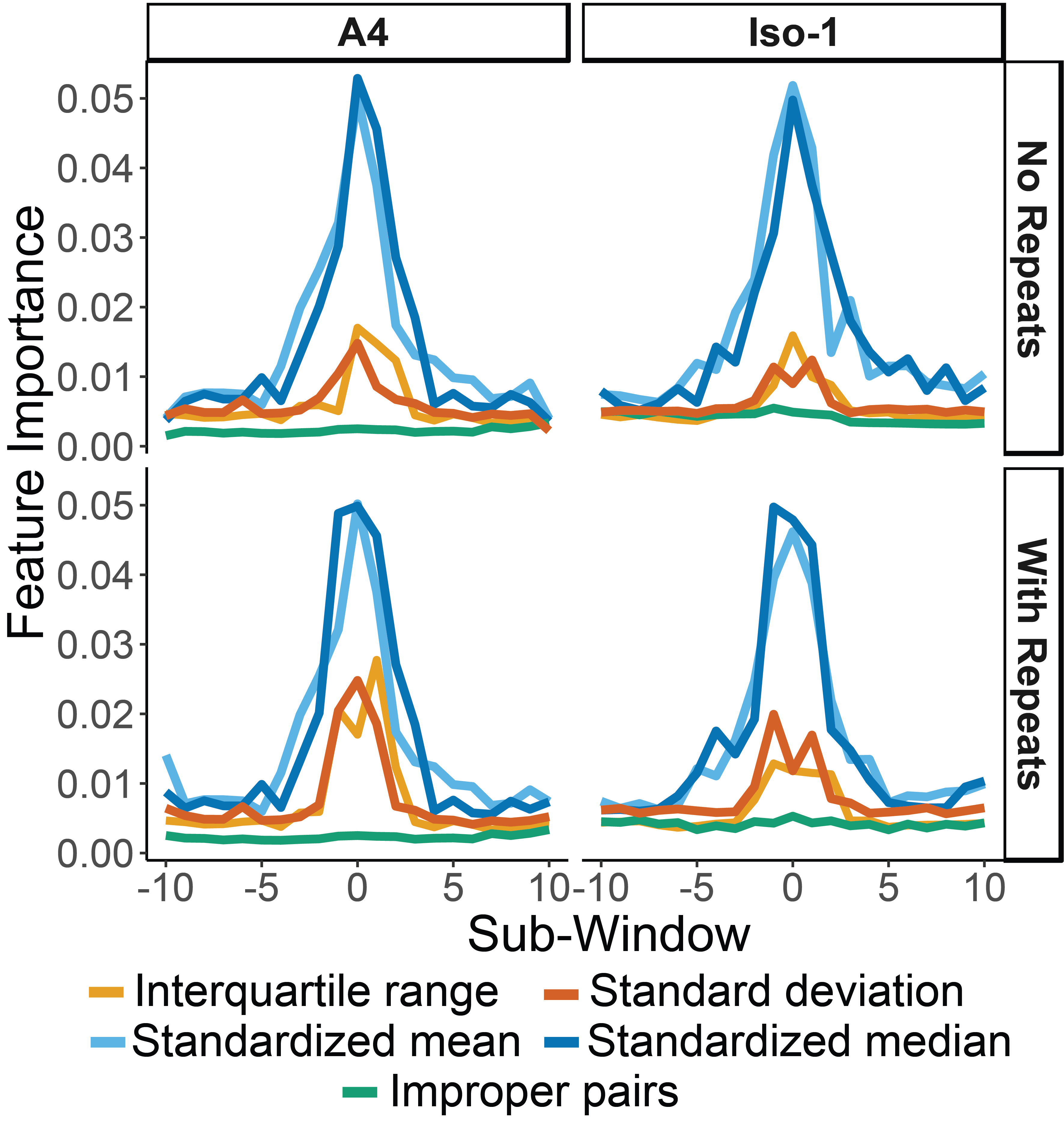
**Figure S5:** Proportion of windows correctly classified and CNVs properly detected or missed given changing parameters, including different fold-coverage, window size, pool size, and number of windows). If parameter is not variable, it is set as follows: 1 individual, 20-fold coverage, 11 sub-windows, 50bp windows, random forest classifier (100 estimators). TPP = true-positive proportion, TNP = true-negative proportion, FPP = false-positive proportion, FNP = false-negative proportion.



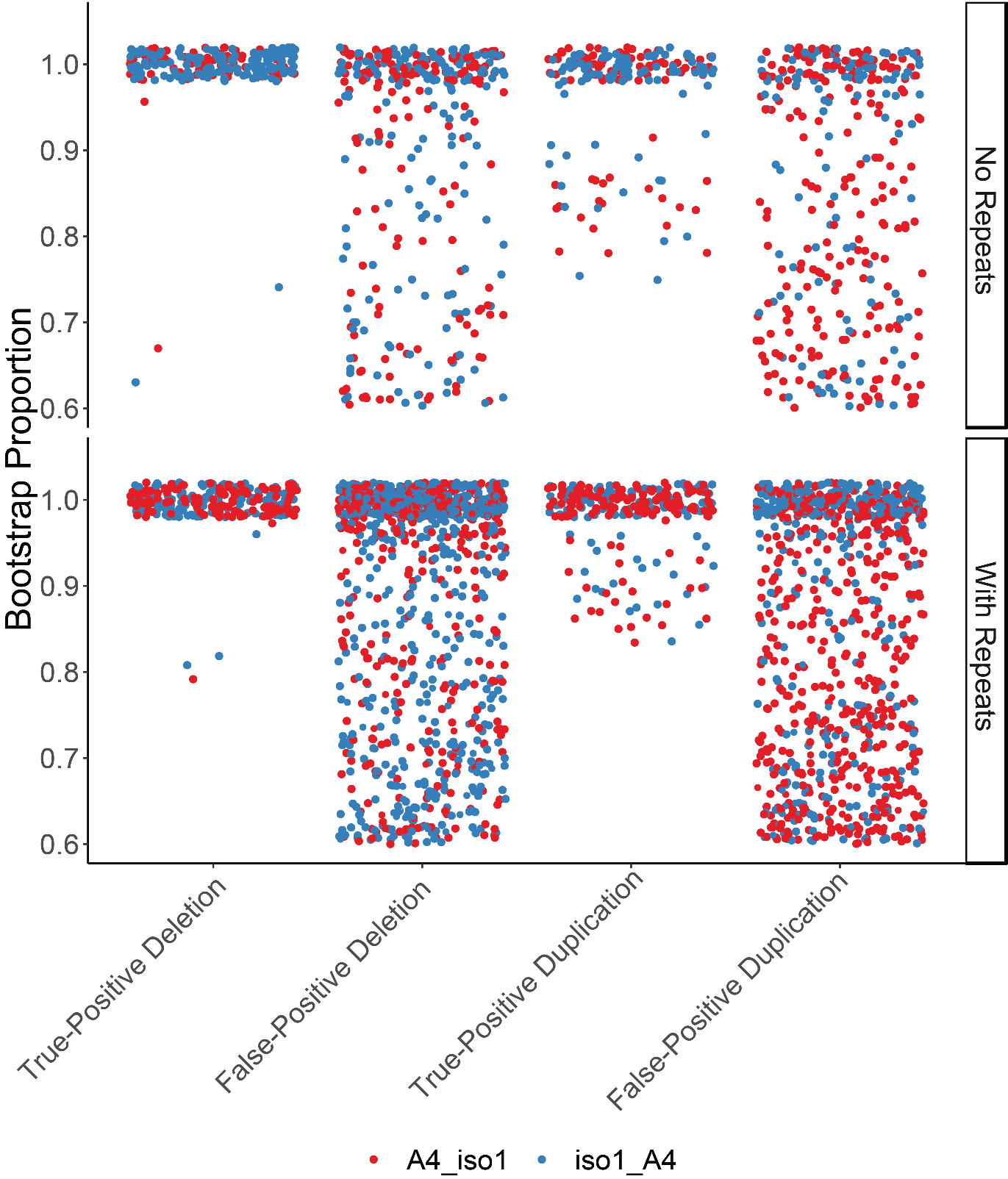
**Figure S6:** The success rate of correct copy number prediction for dudeML, for pure coverage and for the predicted values based on a loess smoothed regression. We also examined the difference between actual copy number and predicted copy number (with correct copy number shown as 0) for dudeML, pure coverage (shown as the median window copy number divided by the median copy number of the total chromosome) and the predicted copy number from the loess smoother regression.

****

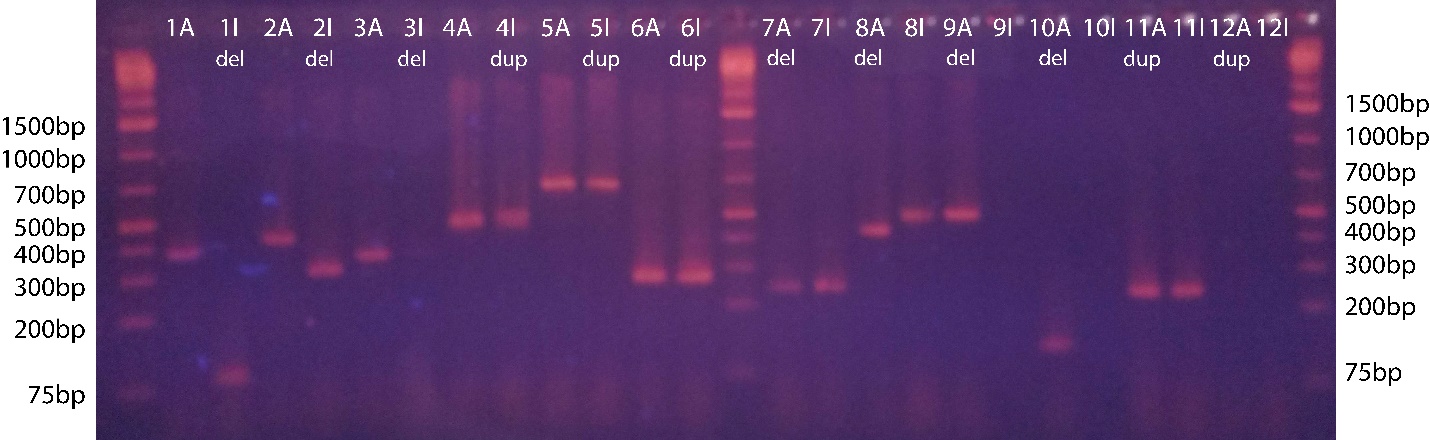
**Figure S7:** Feature importance for real CNVs identified in Iso-1 and A4 strains relative to the alternate reference genomes.



**Figure S8:** The proportion of bootstraps for each detected CNV in Figure 2, separated by if they are a false-positive, true-positive, duplication or deletion, and if the training set contained repeats or not.

****

**Figure S9:** Gel electrophoresis image of PCR products from primers designed around putative CNVs missed in the previous survey, numbered as Supplementary Data 2. Deletions are shown as products shorter than expected, while duplications should be longer or show laddering. Products are ordered showing A4 (A) as the left of the pair, while iso-1 (I) is on the right.



**Data S1:** Screenshots of the integrated genomics viewer for a subset of called duplications and deletions in A4 data mapped to Iso-1 reference genome and vice versa (compared to the data mapped to its own reference). These CNVs were called as false-positives due to their absence in the previous survey. Coverage and reads with supplementary alignments supporting their existence.

**Data S2:** Primer Sequences of a subset of putative duplications and deletions described in Supplementary Figure 8 and Supplementary Data 1.

**Data S3:** Number of true-positive (TPR), false-positive (FPR) and false-negative (TNR) deletion and duplication calls for each dataset (Iso-1 and A4 mapped to either A4 or Iso-1 genomes), given different parameters, such as repeats included in the genome or repeats masked and ignored. In cases with repeat masked genomes, CNVs which would be masked and are therefore uncallable have been removed.