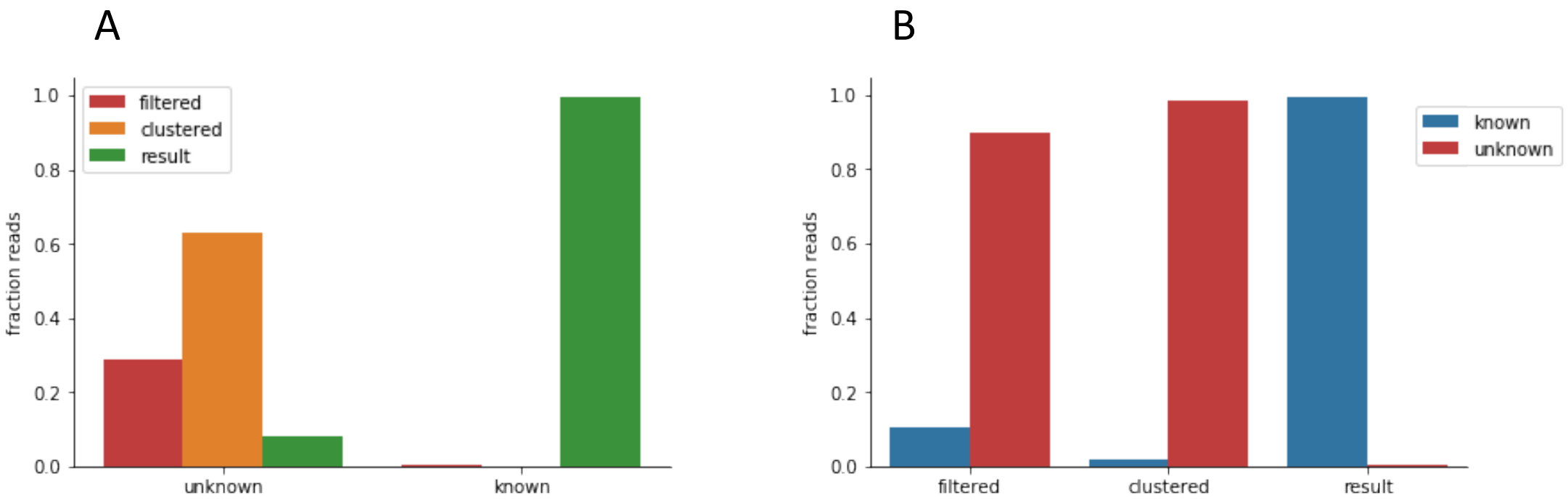
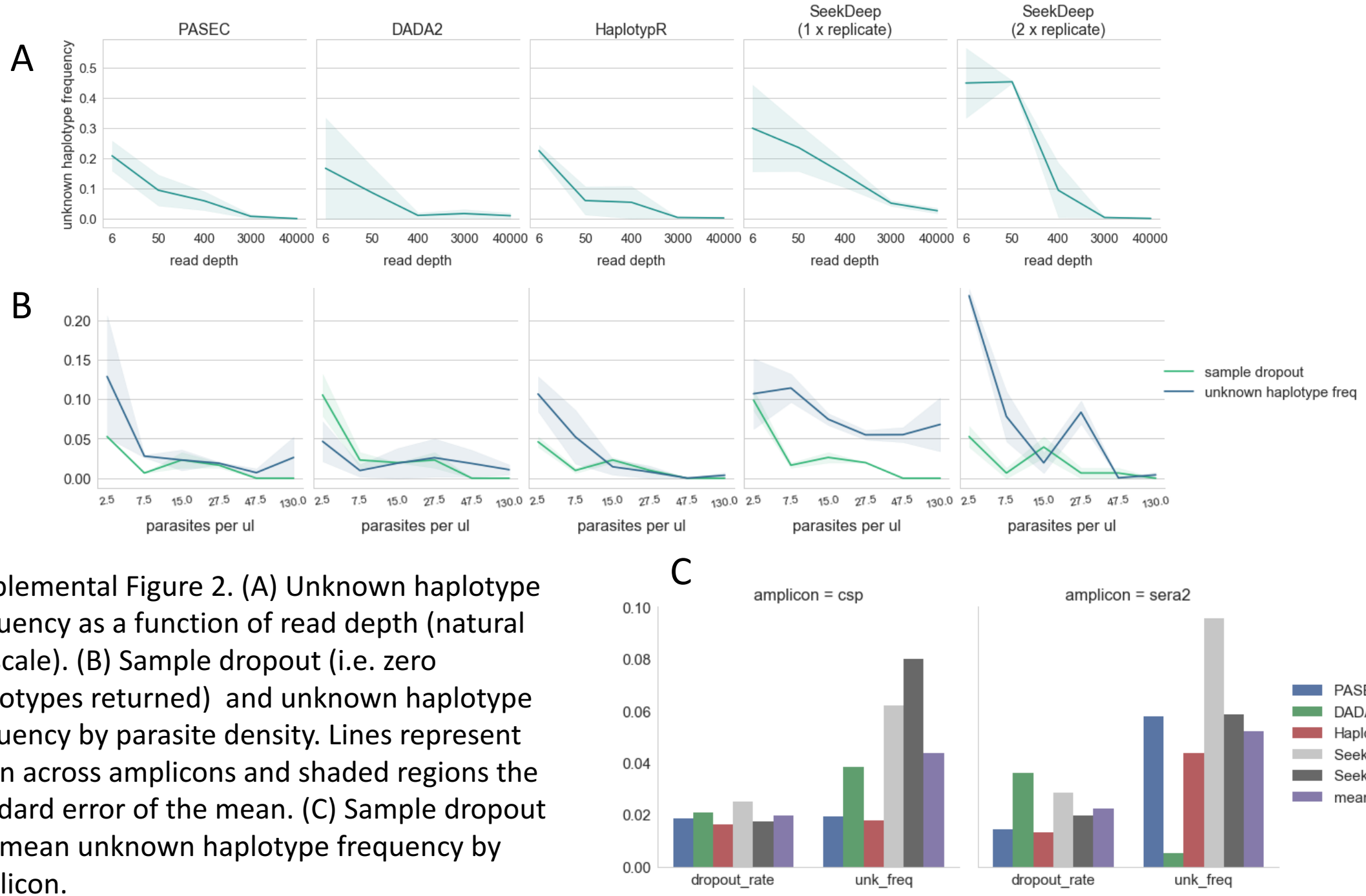
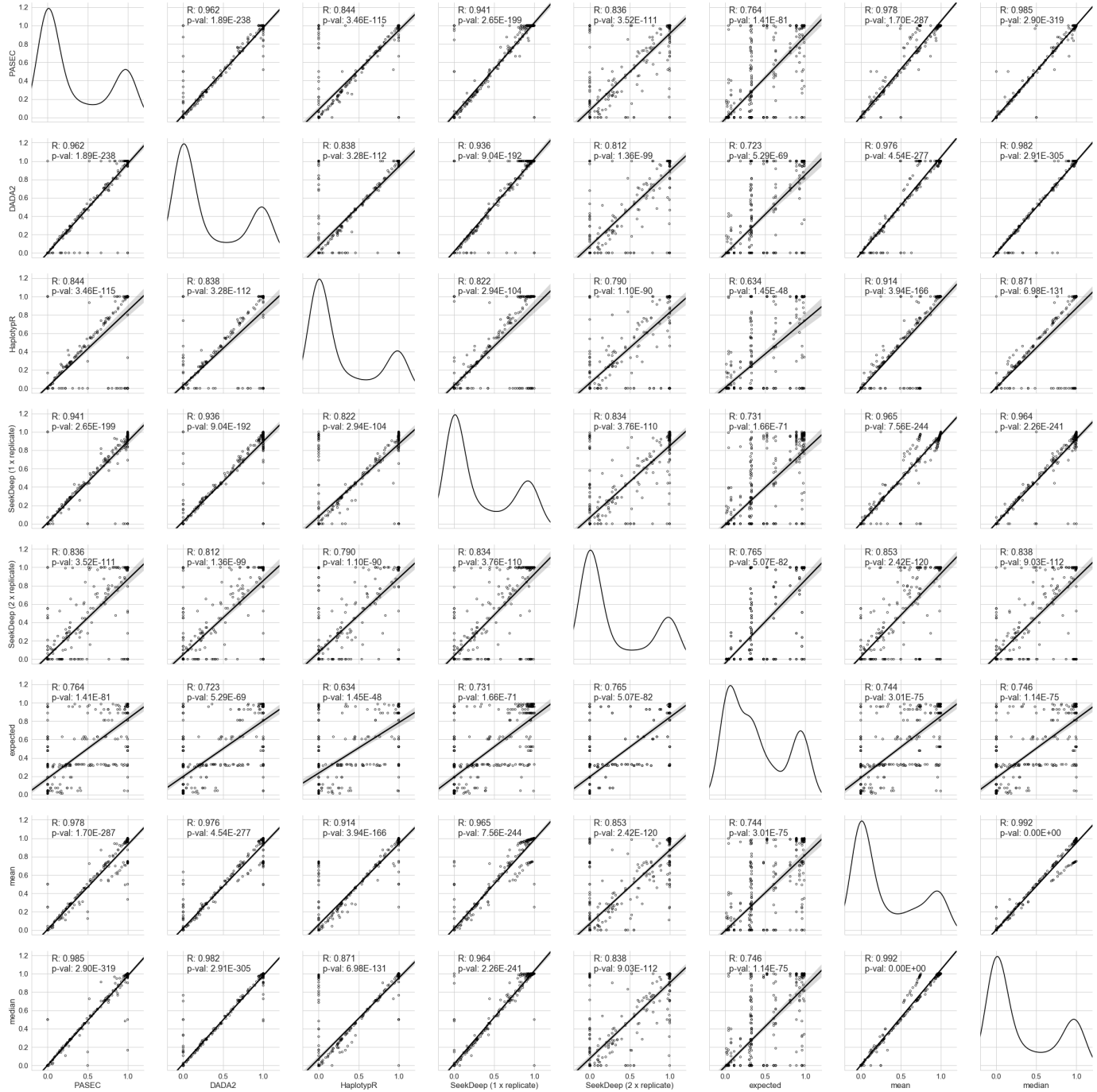


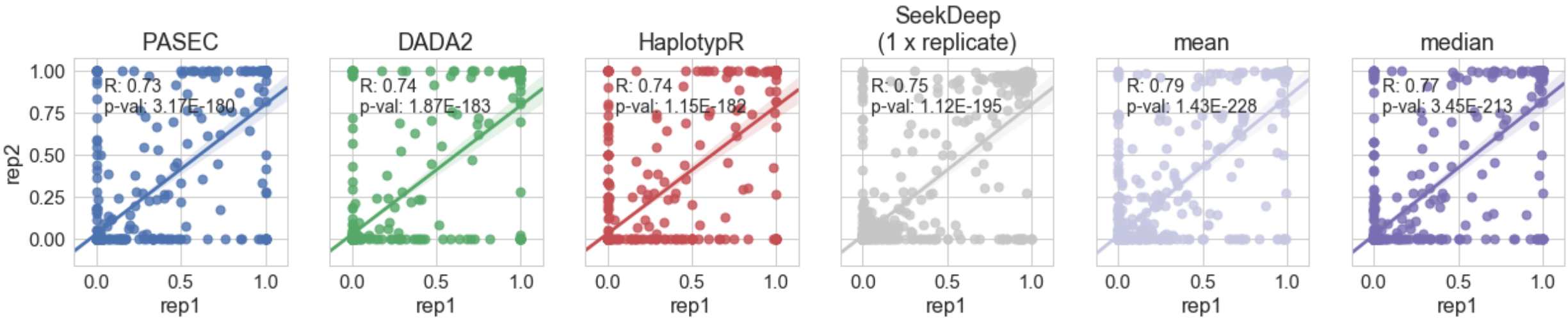
Supplemental Figure 1. PASEC filtering and clustering. (A) Fraction of reads from unknown or known haplotypes that were either filtered, clustered or included in the end results. (B) Fraction of filtered, clustered or resultant reads that were ultimately classified as known or unknown haplotypes.



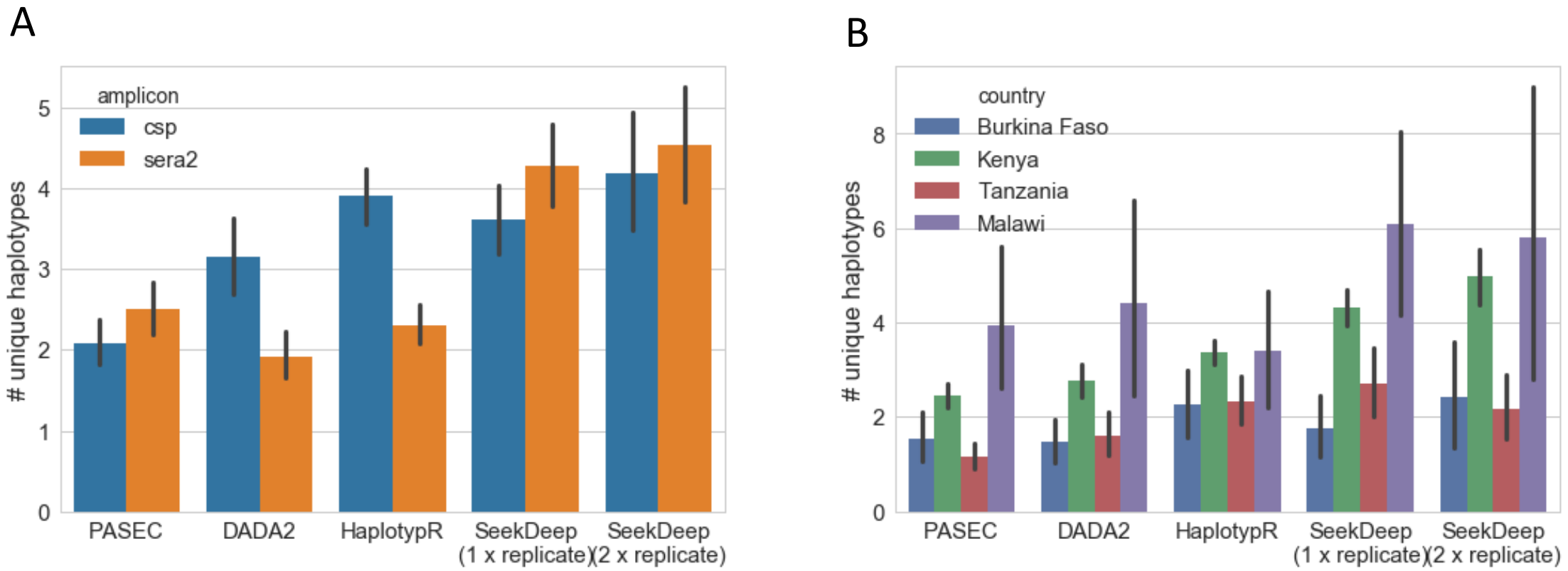
Supplemental Figure 2. (A) Unknown haplotype frequency as a function of read depth (natural log scale). (B) Sample dropout (i.e. zero haplotypes returned) and unknown haplotype frequency by parasite density. Lines represent mean across amplicons and shaded regions the standard error of the mean. (C) Sample dropout and mean unknown haplotype frequency by amplicon.



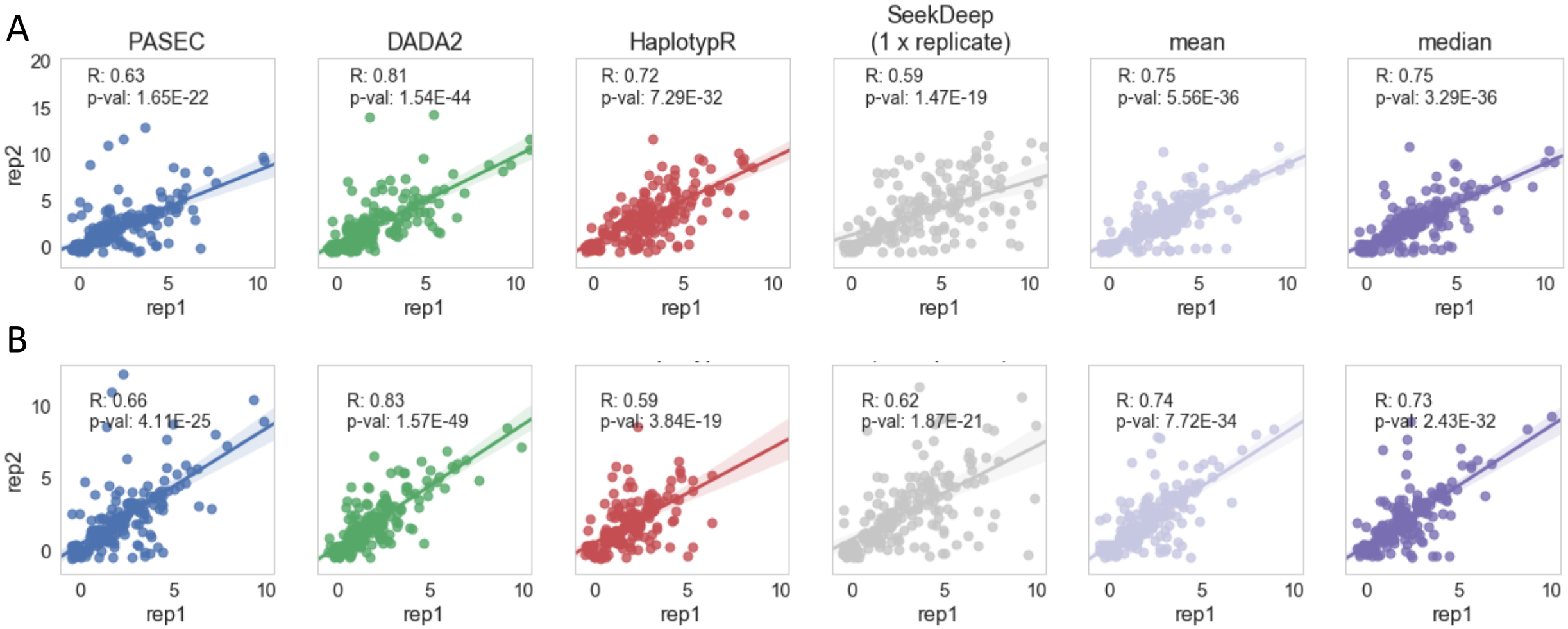
Supplemental Figure 5. Concordance between observed haplotype frequencies for all tools, as well as the mean and median of those observed frequencies and the expected frequencies. Pearson's R and p-values are displayed in the upper left of each of the scatter plots. Distributions are plotted on the diagonal.



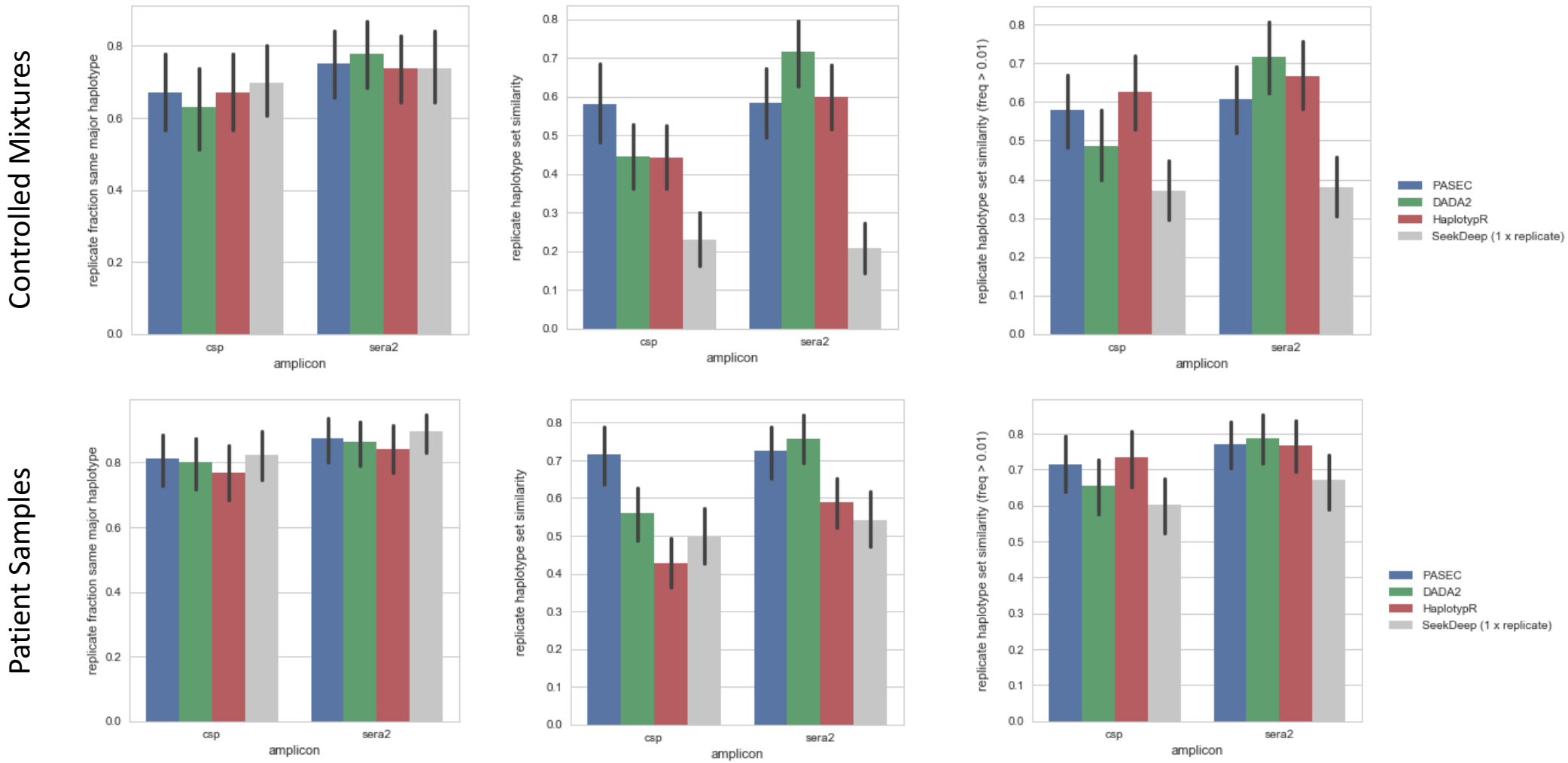
Supplemental Figure 6. Concordance between observed haplotype frequencies of controlled mixture technical replicates, for each tool as well as mean and median frequencies across tools.



Supplemental Figure 7. Number of unique haplotypes called without applying a frequency threshold for each tool for 95 patient blood spot samples from 4 countries, stratified by (A) amplicon and (B) country.

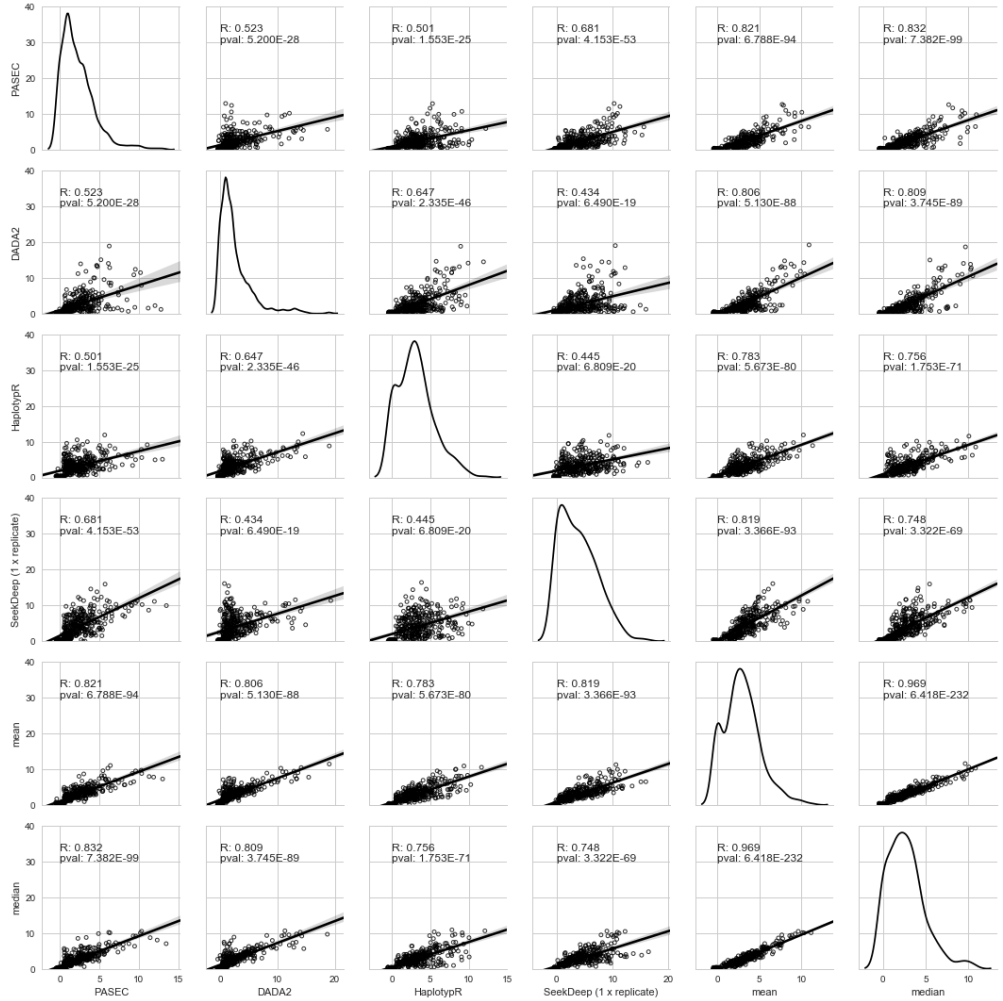


Supplemental Figure 8. (A) Concordance between number of unique haplotypes observed between technical replicates of 95 patient samples with no minimum haplotype frequency. (B) With minimum haplotype frequency of 0.01.

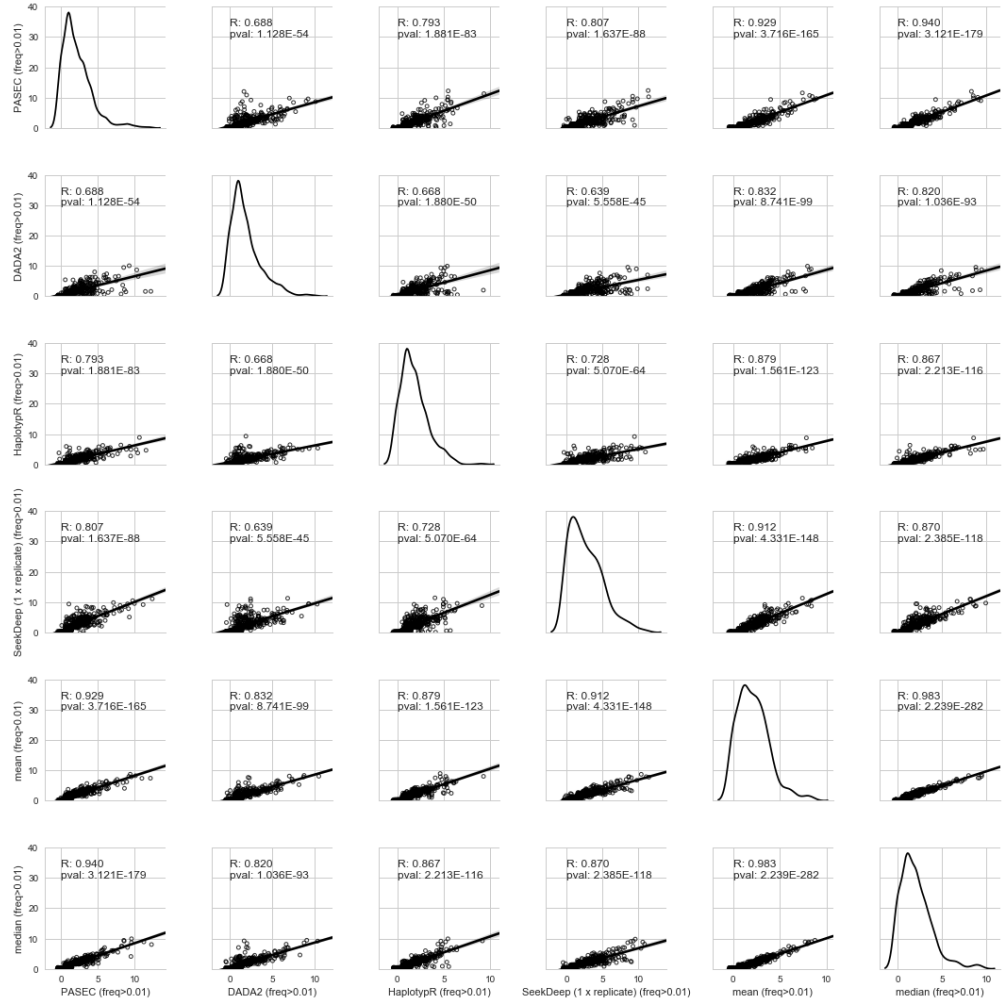


Supplemental Figure 9. Concordance of major haplotypes and haplotype sets between technical replicates for each tool. Results for controlled mixtures are shown across the top row, with results from patient samples shown across the bottom. The first column shows the fraction of major haplotypes that agree between technical replicates. The second and final columns show Jaccard set similarity of haplotype sets called between replicates, with the final column showing results after filtering haplotypes with frequency less than 0.01.

A



B



Supplemental Figure 10. Correlations between the number of unique haplotypes from each 95 patient blood spot sample, as returned by each tool along with the mean and median. Pearson's R and its p-value are displayed in the upper left of each of the scatter plots. Distribution of each set of predicted MOIs are plotted on the diagonal. (A) No minimum haplotype frequency. (B) Minimum haplotype frequency of 0.01.