

Supporting Figures about Breakpoints Accuracy

Qian Feng

11/12/2019

There are six biologically realistic scenarios in the simulation datasets,

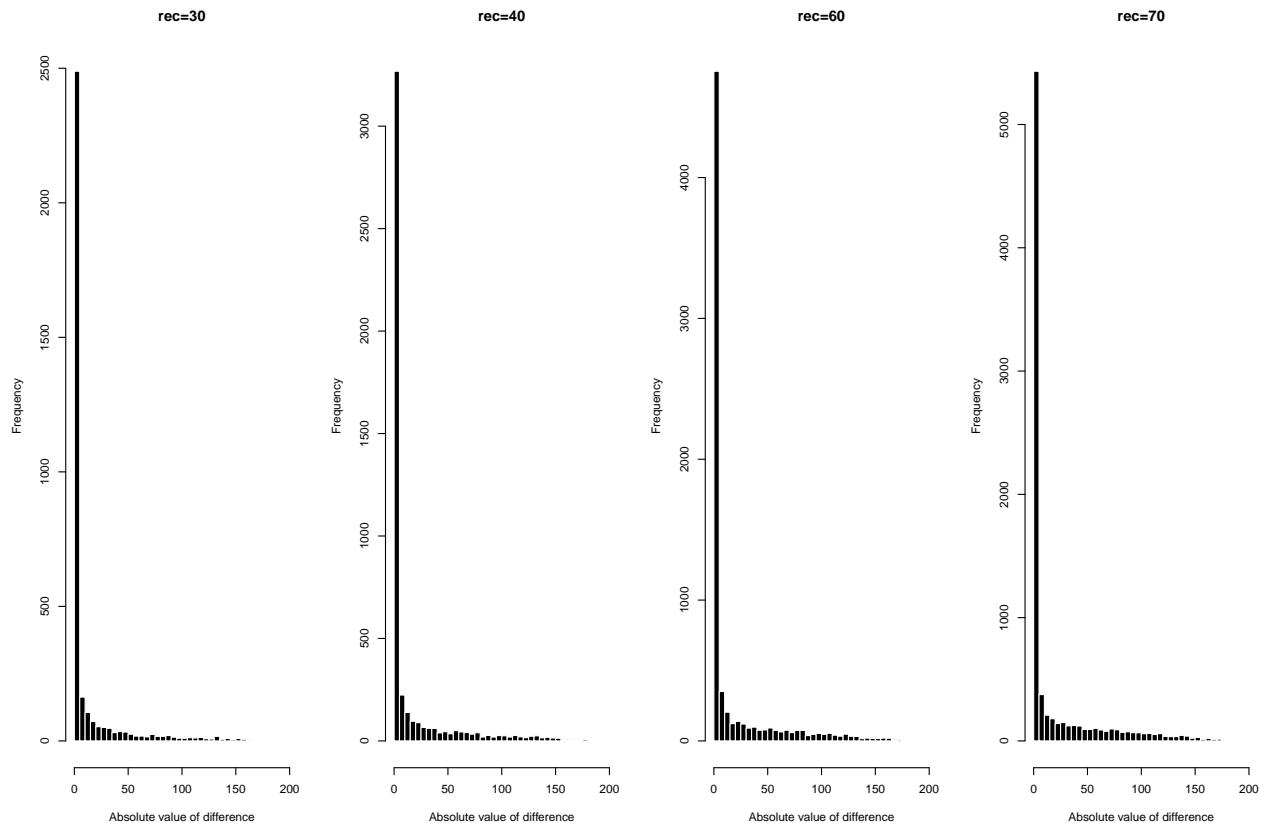
1. different recombination proportions
2. different mutation rates
3. different indel rates
4. different indel sizes
5. different sequence lengths
6. different amino acid models
7. different dataset size (the number of sequences per replicate).

Breakpoints' accuracy evaluation for each scenario is summarized by two plots, one is scatter plot between identified breakpoint and real breakpoint positions, second is the histogram for the absolute values of the difference between identified breakpoint and real breakpoint positions. So there are two plots for each case. Manuscript only offers related results when recombination proportion is equal to 50%. This file provides all the supplementary figures for the manuscript in terms of breakpoints' accuracy evaluation section.

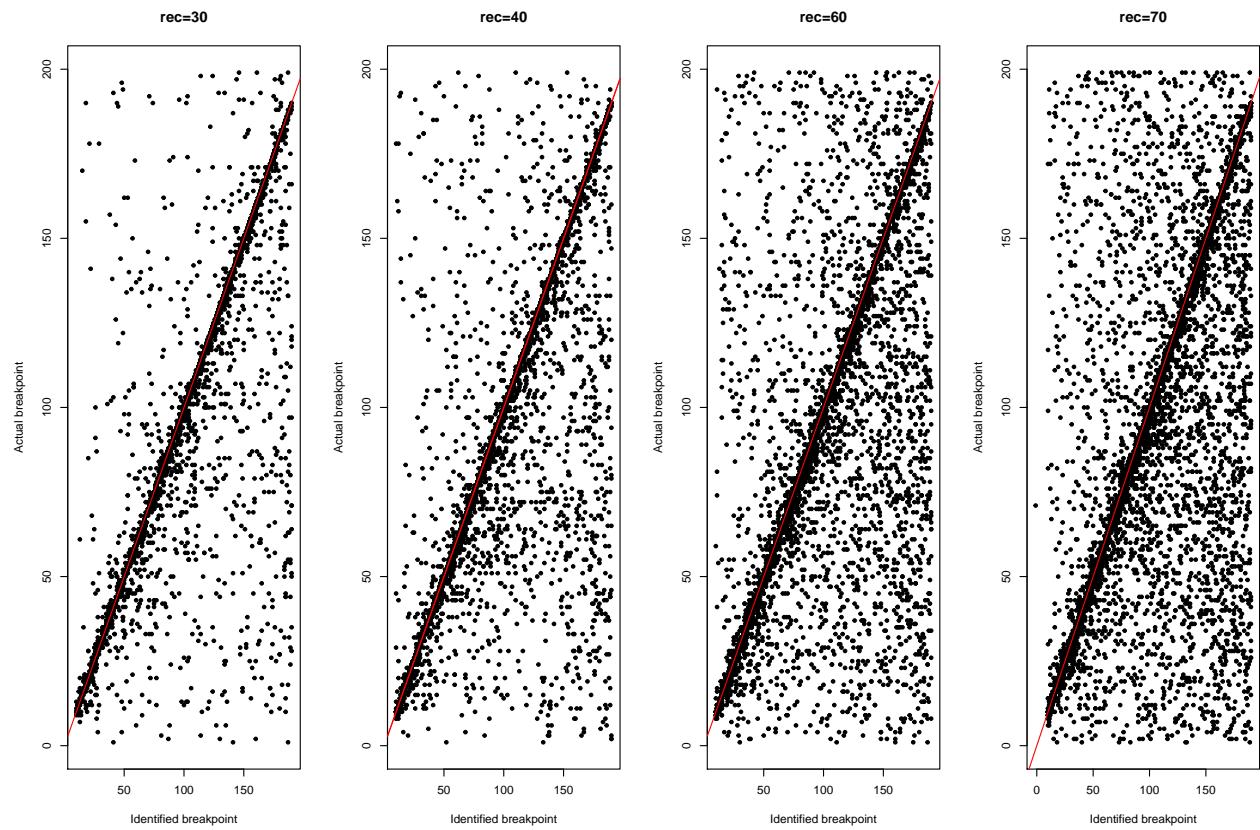
Therefore, all the following figures present scatter plot and histogram for each scenario.

Section 1: Different recombination proportions

histograms

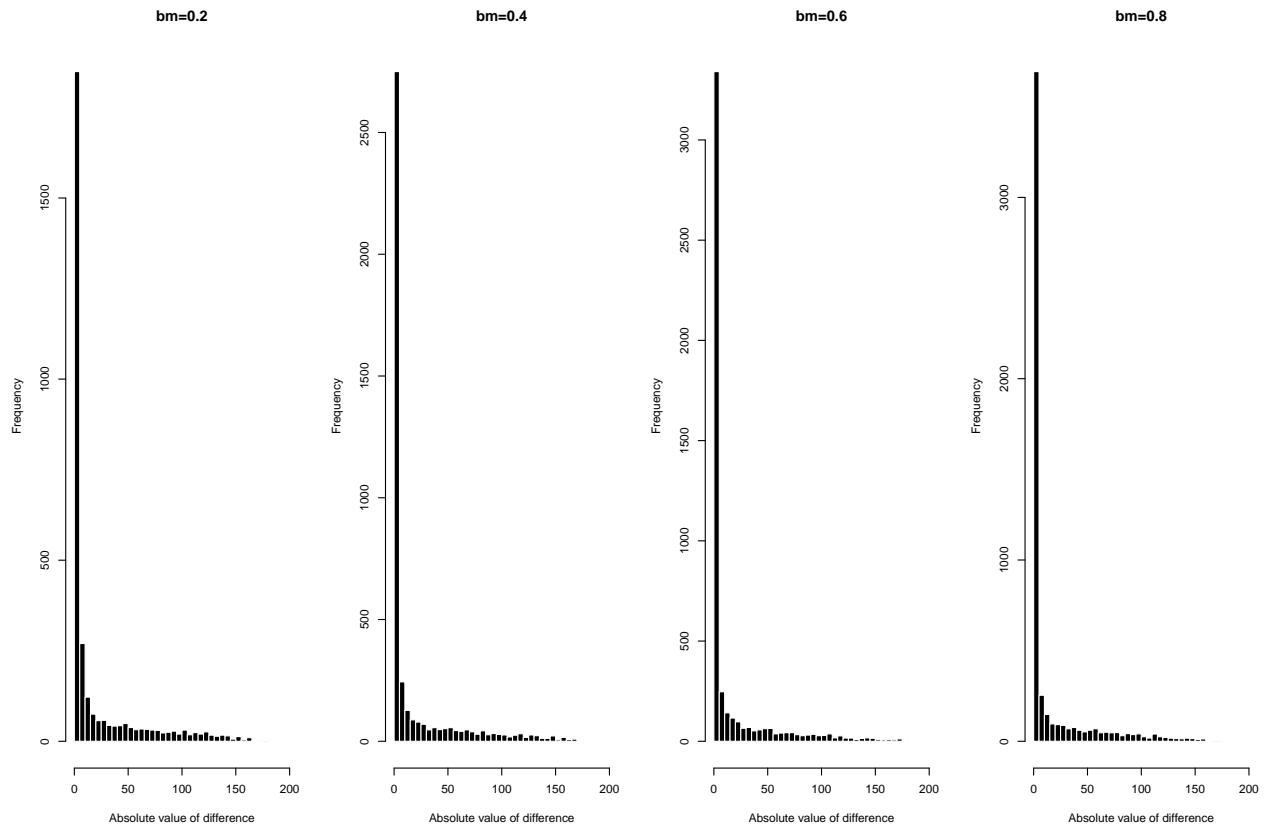


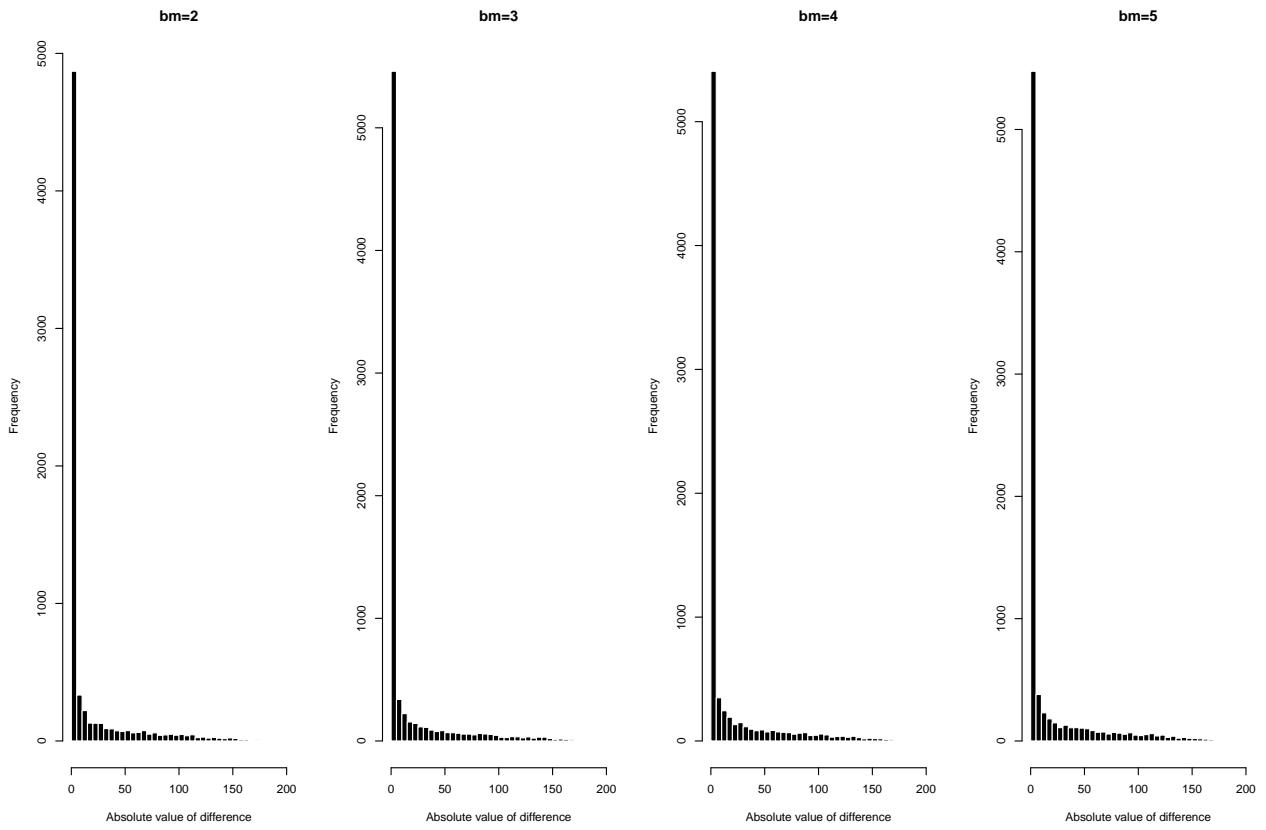
scatter plots for the actual_bkp vs identified_bkp



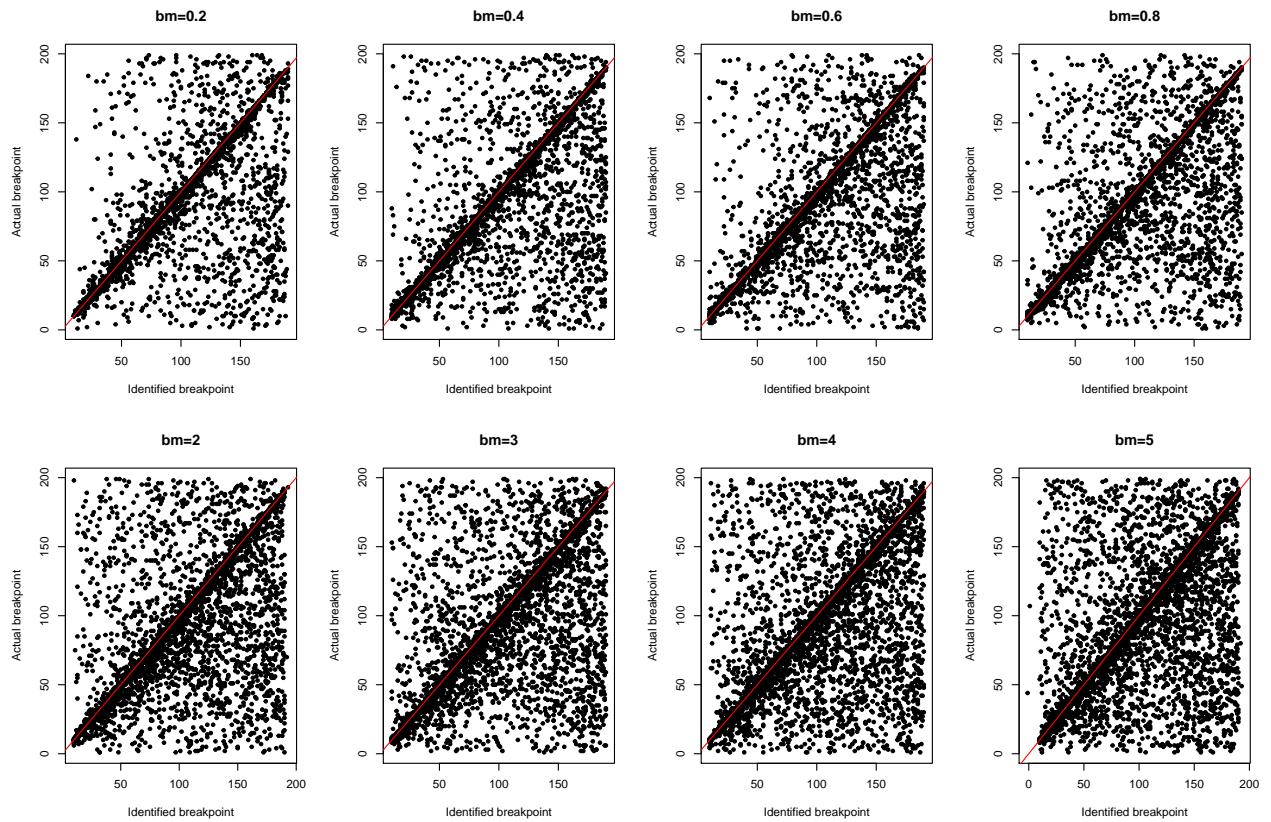
Section 2 : Different mutation rates

histograms



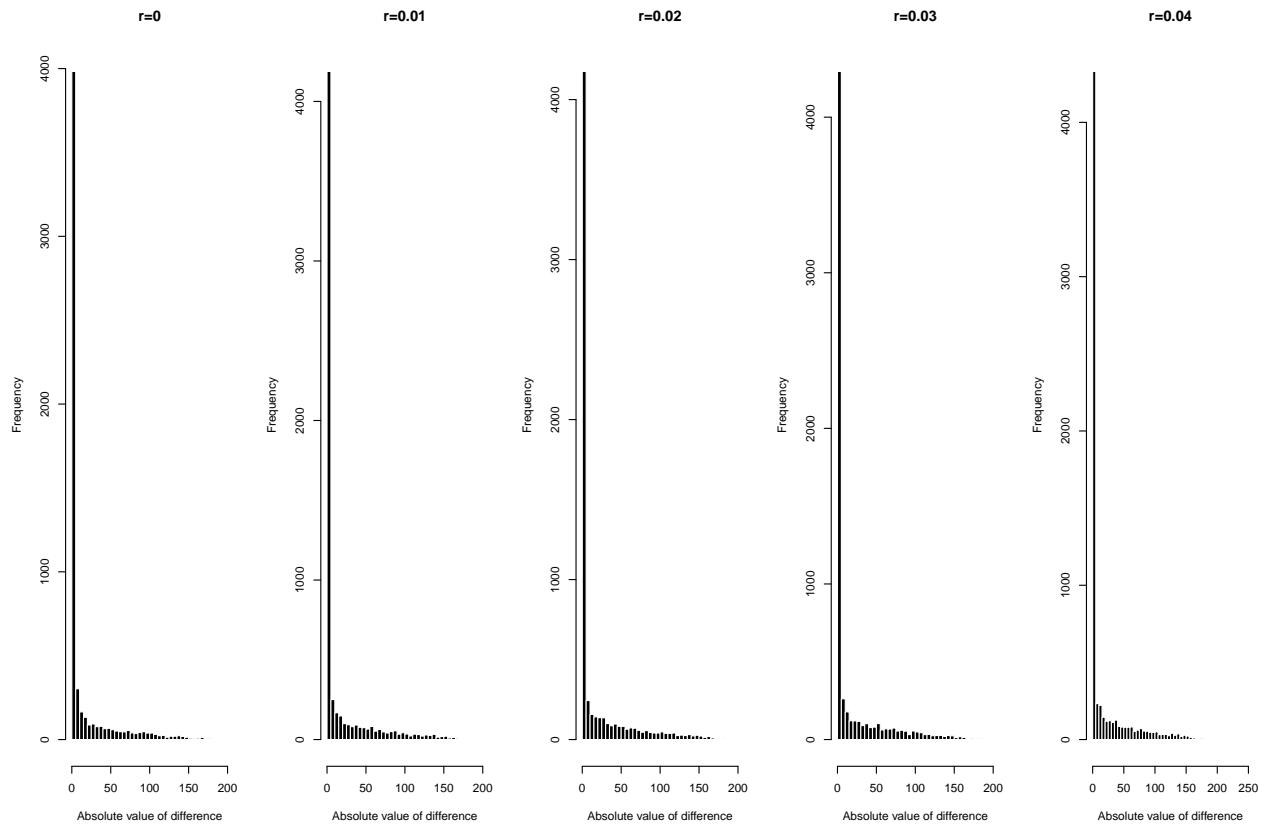


scatter plots for the actual_bkp vs identified_bkp

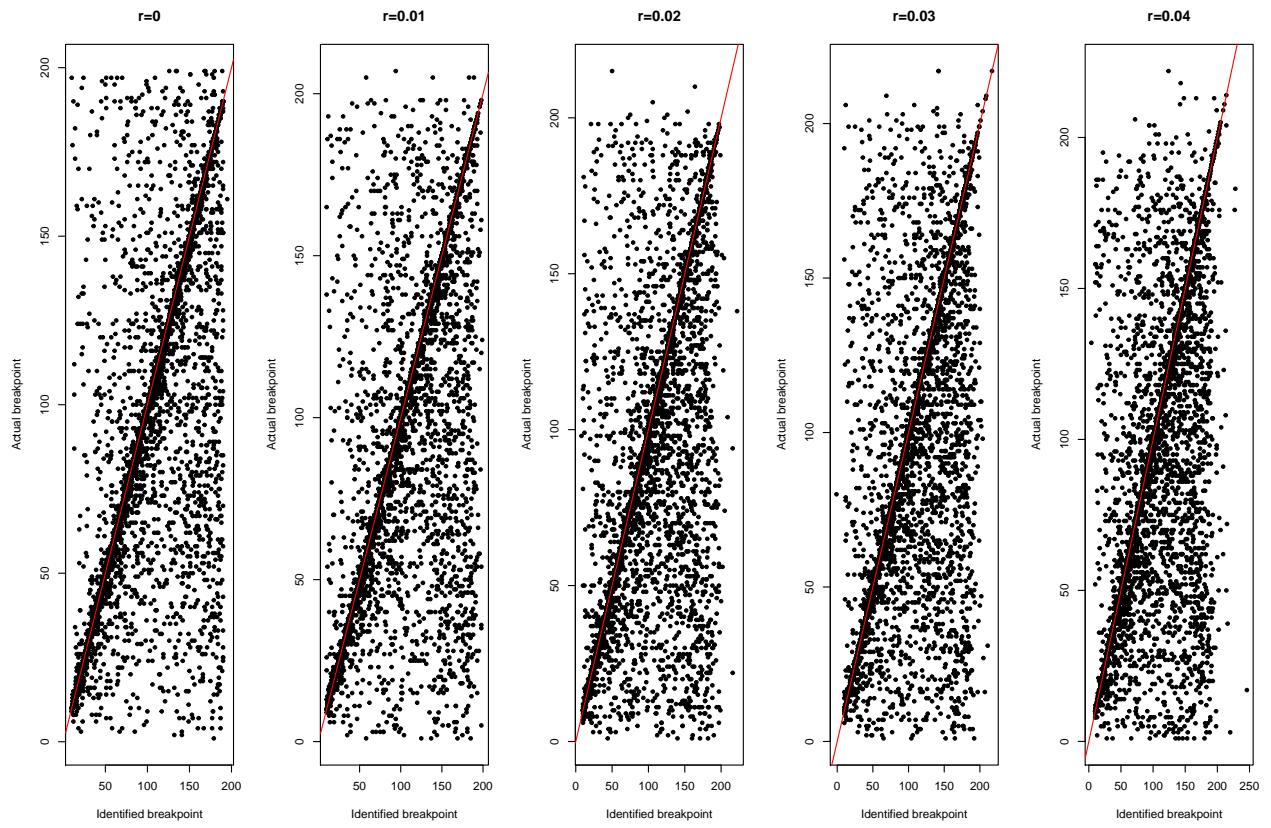


Section 3: Different indel rates

histograms

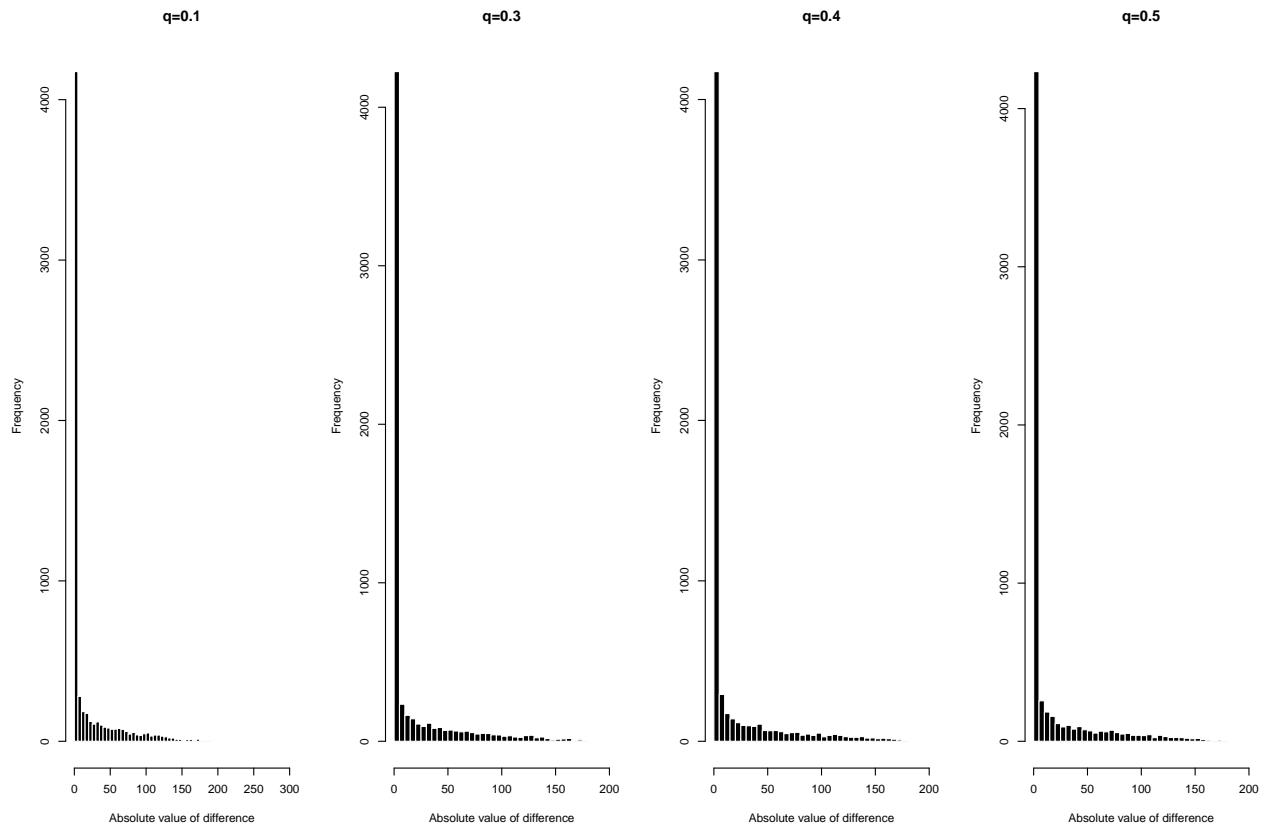


scatter plots for the actual_bkp vs identified_bkp

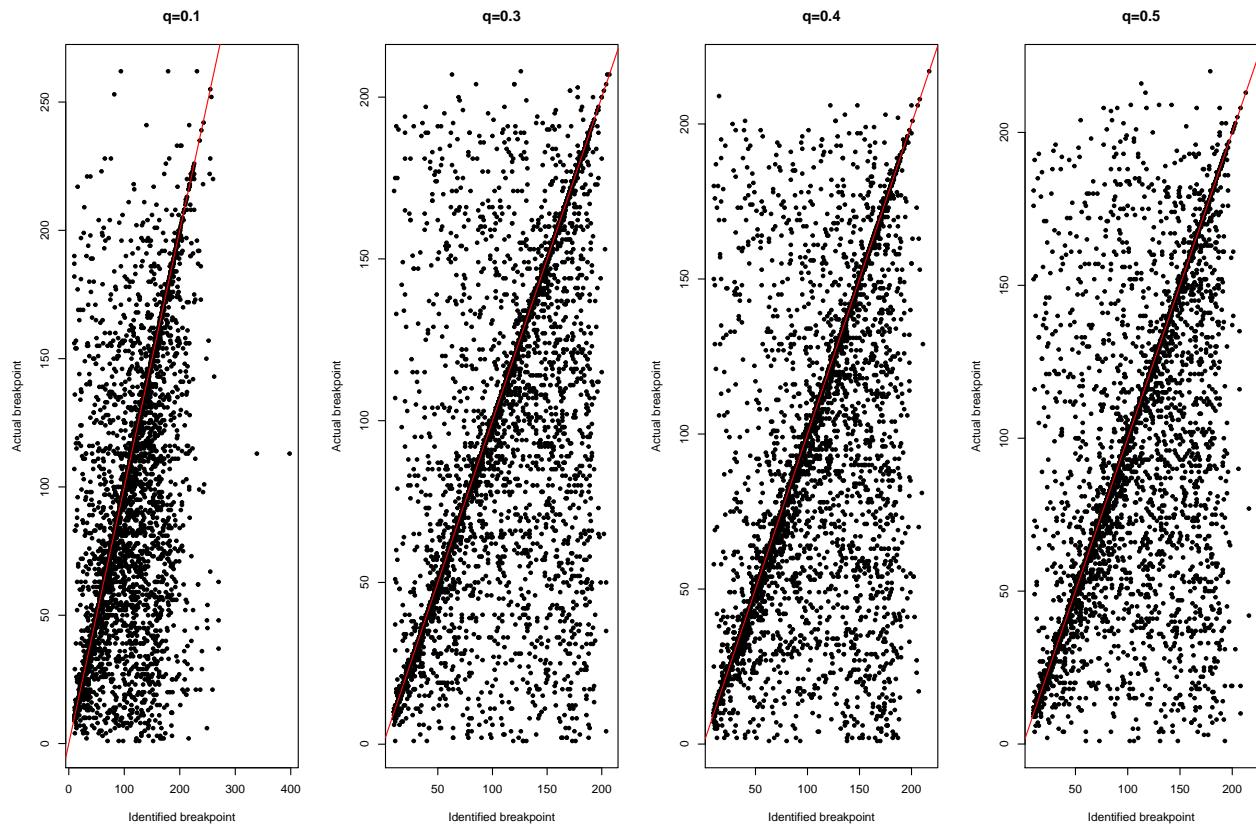


Section 4: Different indel sizes

histograms

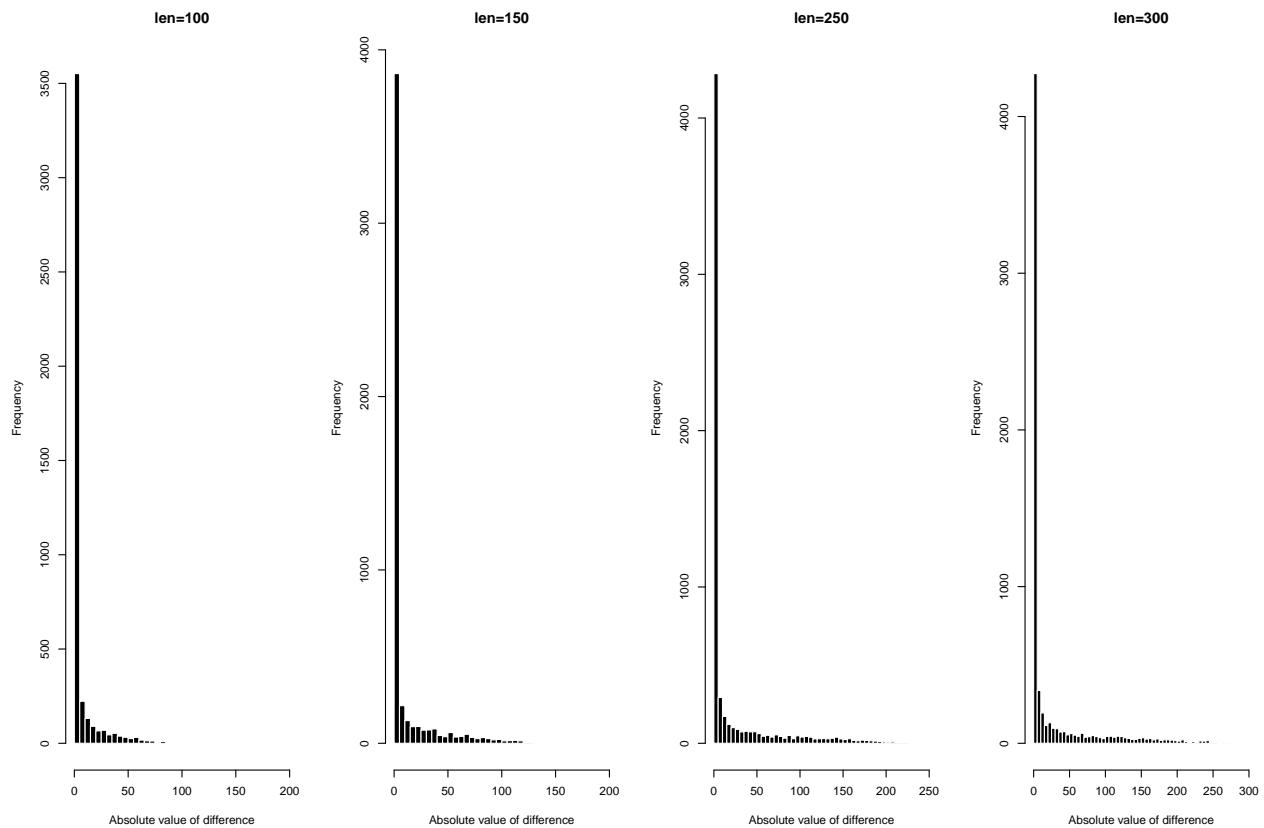


scatter plots for the actual_bkp vs identified_bkp

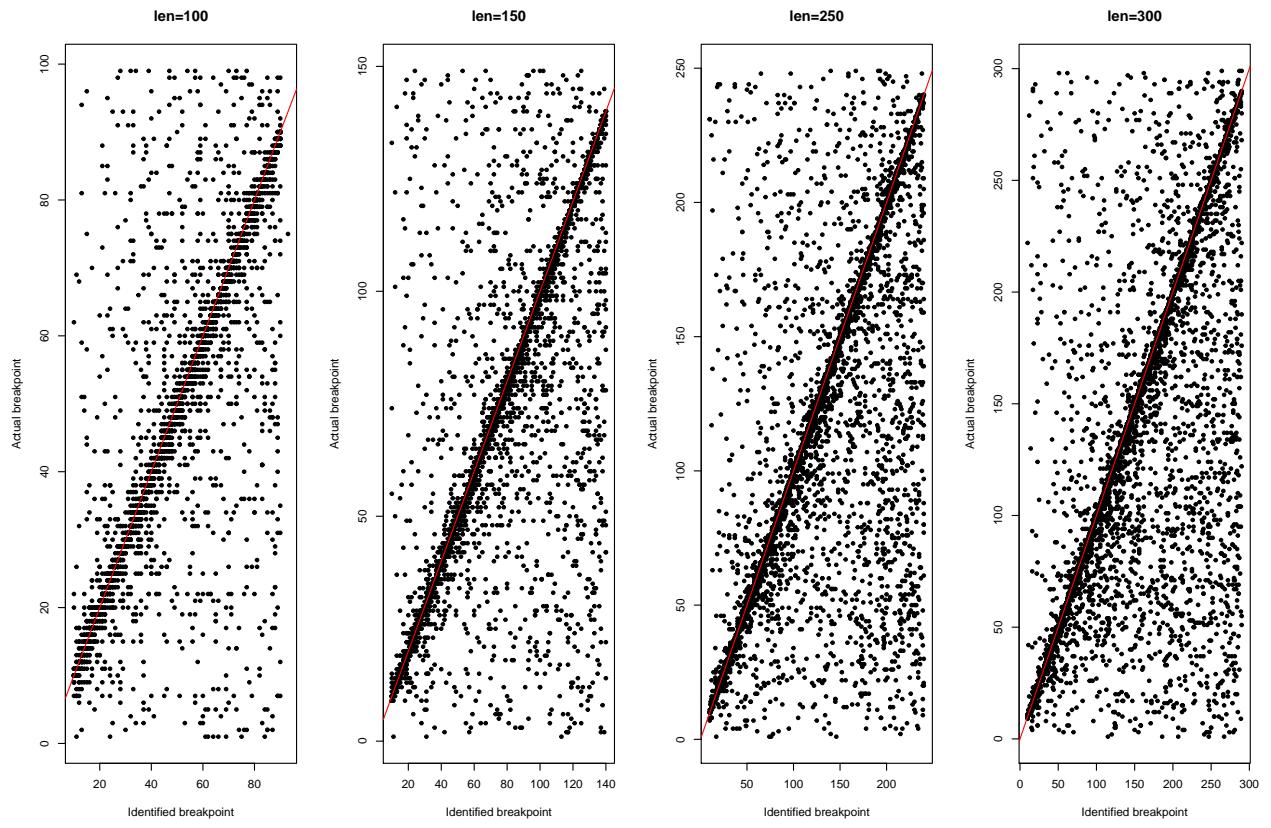


Section 5 : Different sequence lengths

histograms

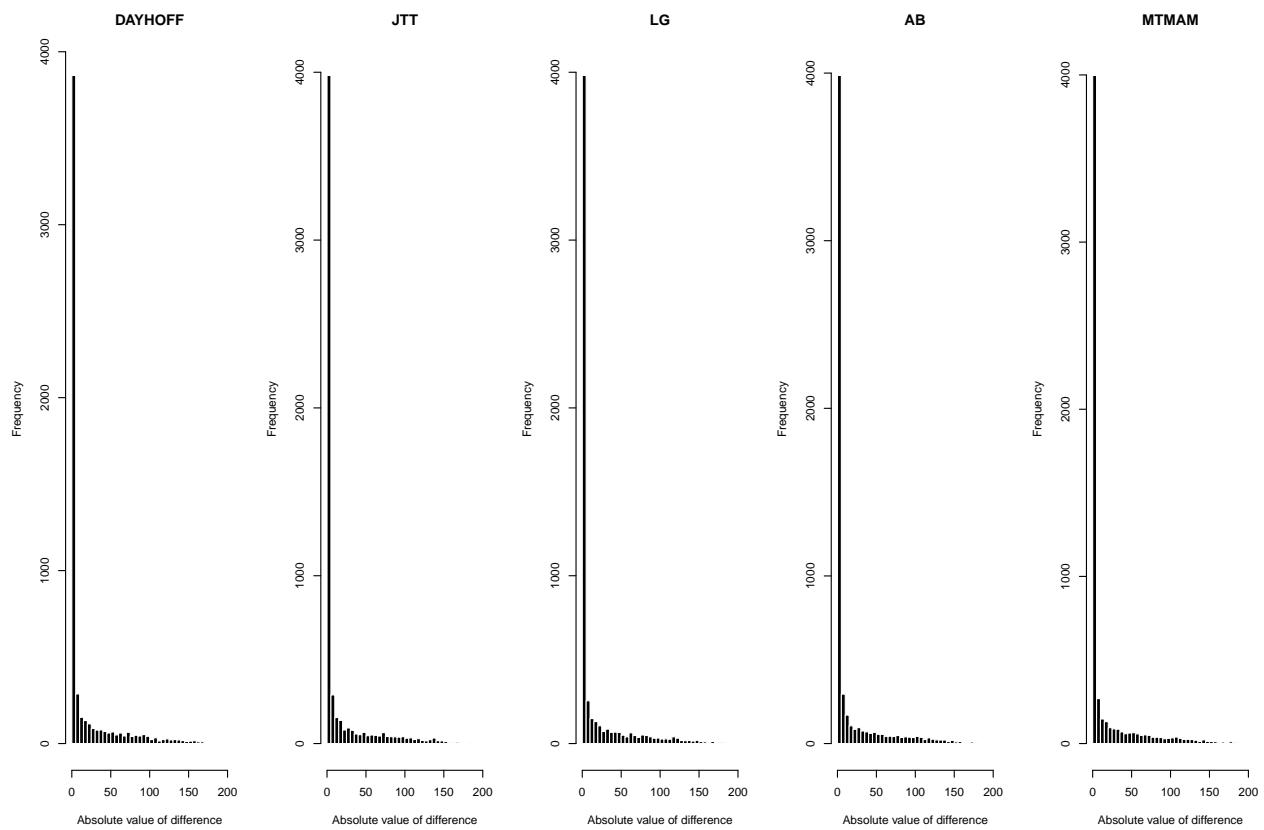


scatter plots for the actual_bkp vs identified_bkp

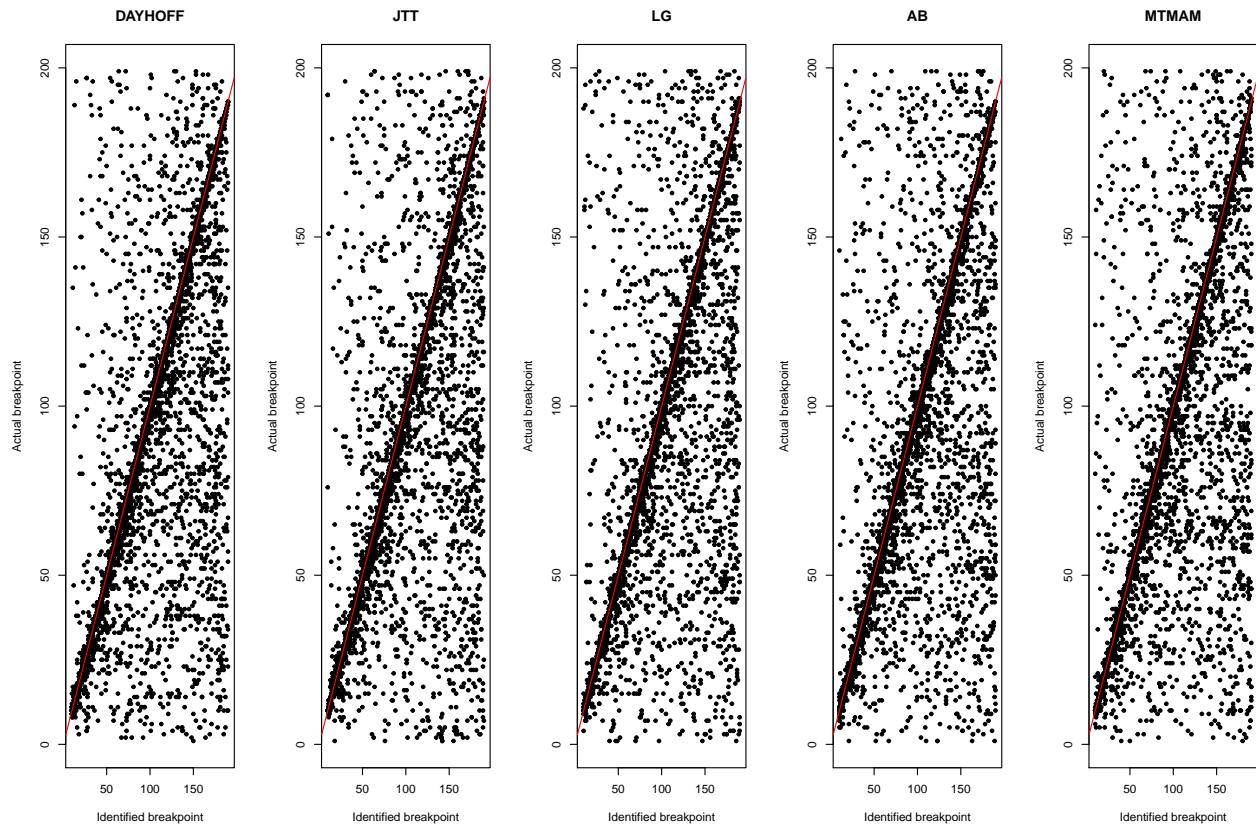


Section 6 : Different AA models

histograms

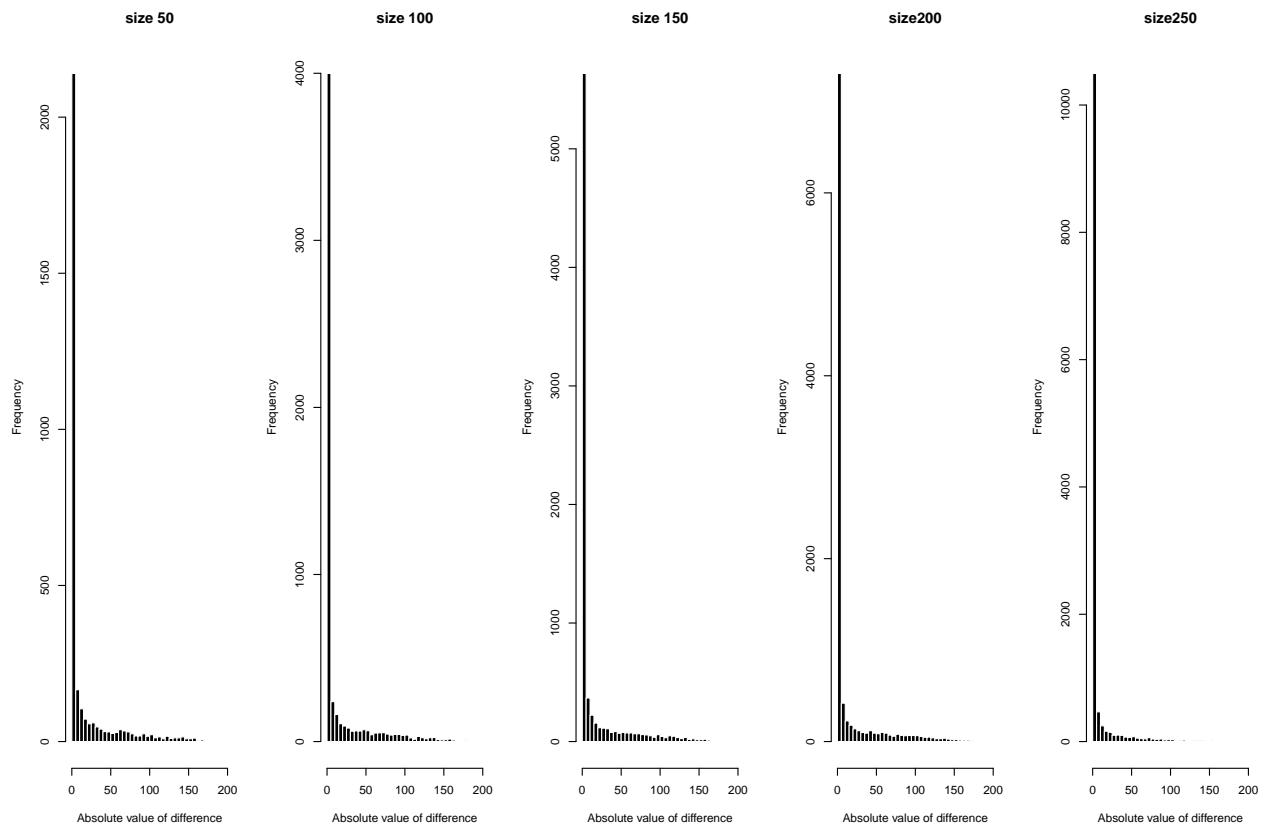


scatter plots for the actual_bkp vs identified_bkp



Section 7 : Different dataset size

histograms



scatter plots for the actual_bkp vs identified_bkp

