Supplemental materials.

To see if one could use this approach to identify viral isolates, we conducted a set of studies in which the identities of the viral strains were unknown to those performing the experiments and data analysis. Four anonymous strains of human rhinoviruses were obtained from our collaborators. Before constructing their sequence motif maps, the nearly full length dsDNA was generated by reverse transcription followed by long range PCR with a pair of universal PCR primers for all strains. PCR forward primer is an 18bp oligo (5'TCCTCCGGCCCCTGAATG), which is conserved for all the strains and the reversed primer (5'Cy3ATTCTAGAGGCCGAGGCGGCCGACATG-d$(T)_{30}$) was labeled with Cy3 to indicate the orientation of the DNA molecules. In these experiments, the nicking endonuclease Nb.Bsm I (GCATTC) was used and the nicking sites were labeled with Tamra-ddCTP. After labeling, the samples were purified with S400 spin column (Amersham, UK) to remove the labeled PCR primers. The predicted maps for all four strains are shown in Figure S1. Figure S2 shows some typical false color two-channel composite images of the stretched DNA contours (YOYO in blue) and the Nb.Bsm I sites (Tamra-ddCTP in green). On average, the labeling efficiency is about 50%. The DNA molecules that stretched over 2.0 microns in length with one end green label and at least one green internal label were chosen to construct sequence motif maps shown in Figure S3. About 100-200 DNA molecules were used. Clearly, three peaks at 0kb, 1.6 kb and 3.69 kb are observed in sample 1. The two internal peaks of the map match well with the predicted nicking sites of HRV-15 virus at 1.57kb and 3.62kb respectively. Only one internal peak was identified in sample 2 at 3.52 kb, which is consistent with the map of HRV28. Sample 3 was confirmed to be HRV36 virus. Interestingly, the HRV15 and HRV36 would not be distinguished if no end label was used to indicate the orientation of the DNA molecules. Sample 4 is a little bit problematic since some of DNA molecules have two end labels, which makes it impossible to discern the orientation of the DNA molecules. The map of sample 4 was generated disregard of the end label. Two pairs of mirrored peaks were observed at 0,6.7kb and 1.53kb, 5.27kb, which correspond to the predicted map of HRV73 virus. The minimal distance of 1.6kb was detected between two closest labels, indicating the resolution of 1.6 kb can be achieved.

Figure S1. Sequence motif maps (Nb.Bsm I) of human rhinovirus 15, 28, 36 and 73.

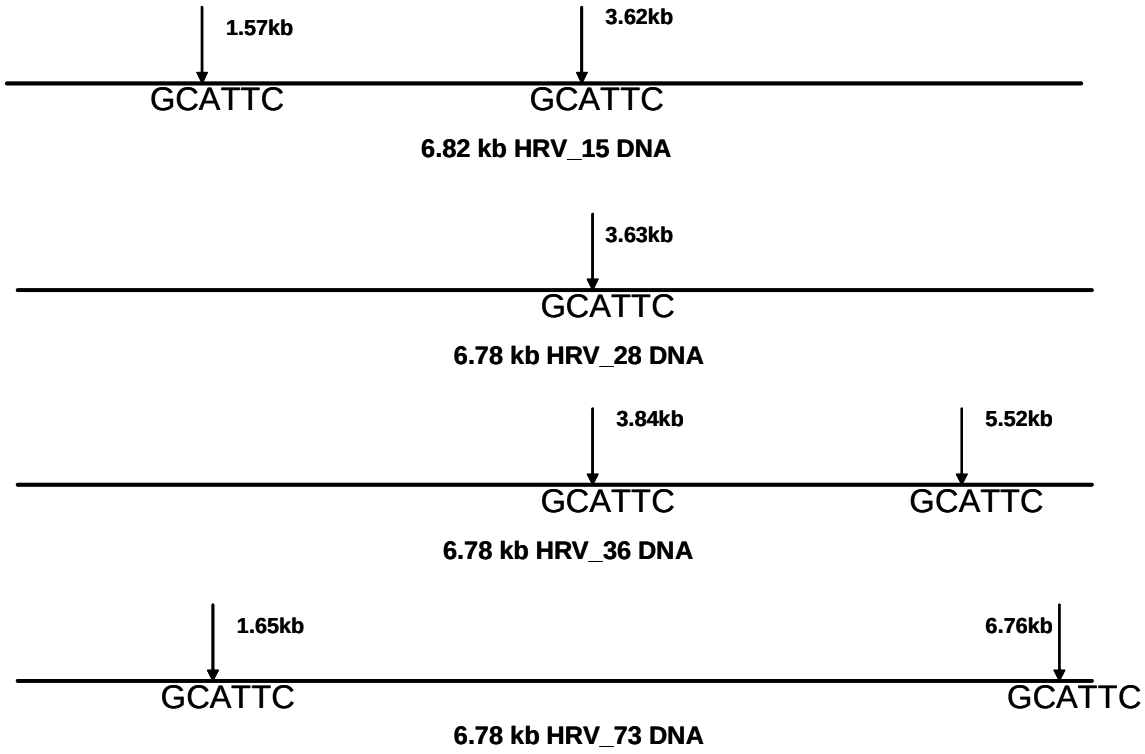

Figure S1. The predicted Nb.Bsm I (GCATTC) map of human rhinovirus 15, 28, 36, and 73.

Figure S2. The intensity scaled composite two-false-color images.

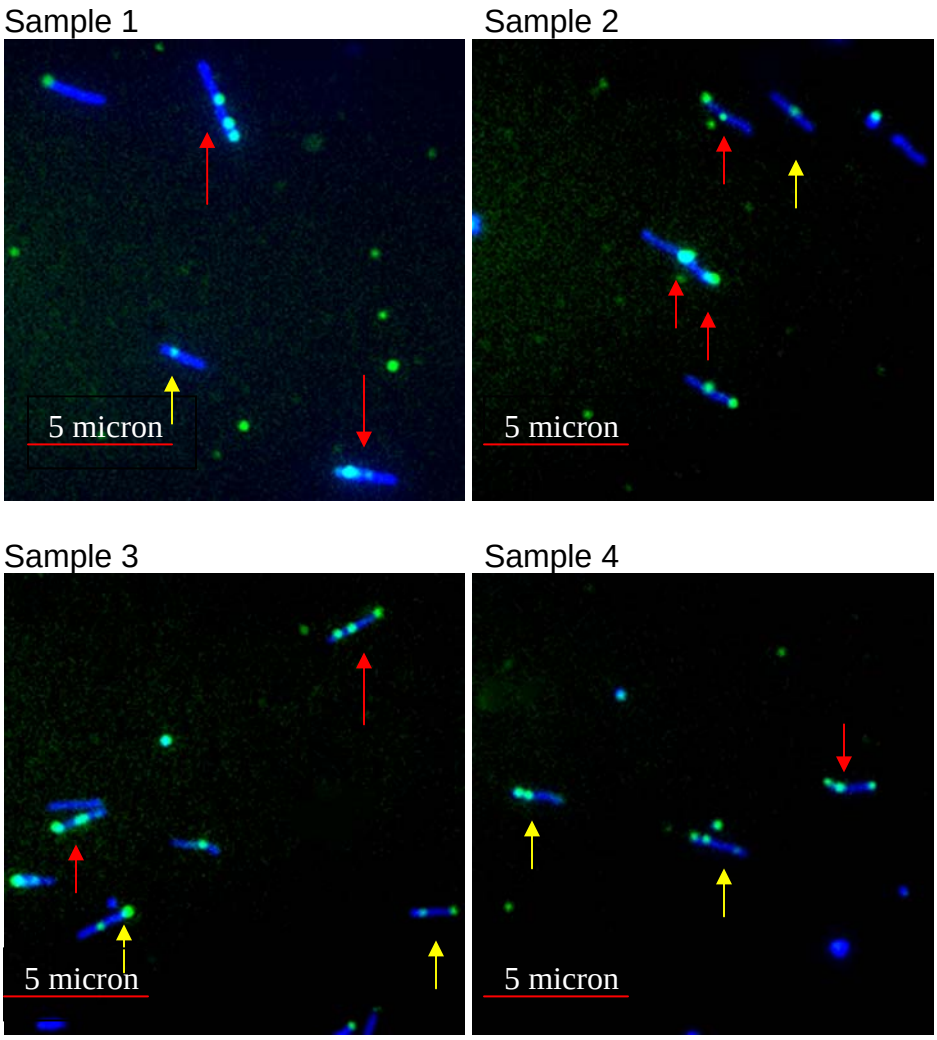

Figure S2. The intensity scaled composite two-false-color images. In the large field of the intensity scaled composite image, numerous molecules are found. Some molecules were fully labeled (red arrows) and some were partially labeled (yellow arrows).

Figure S3. The constructed sequence motif maps.

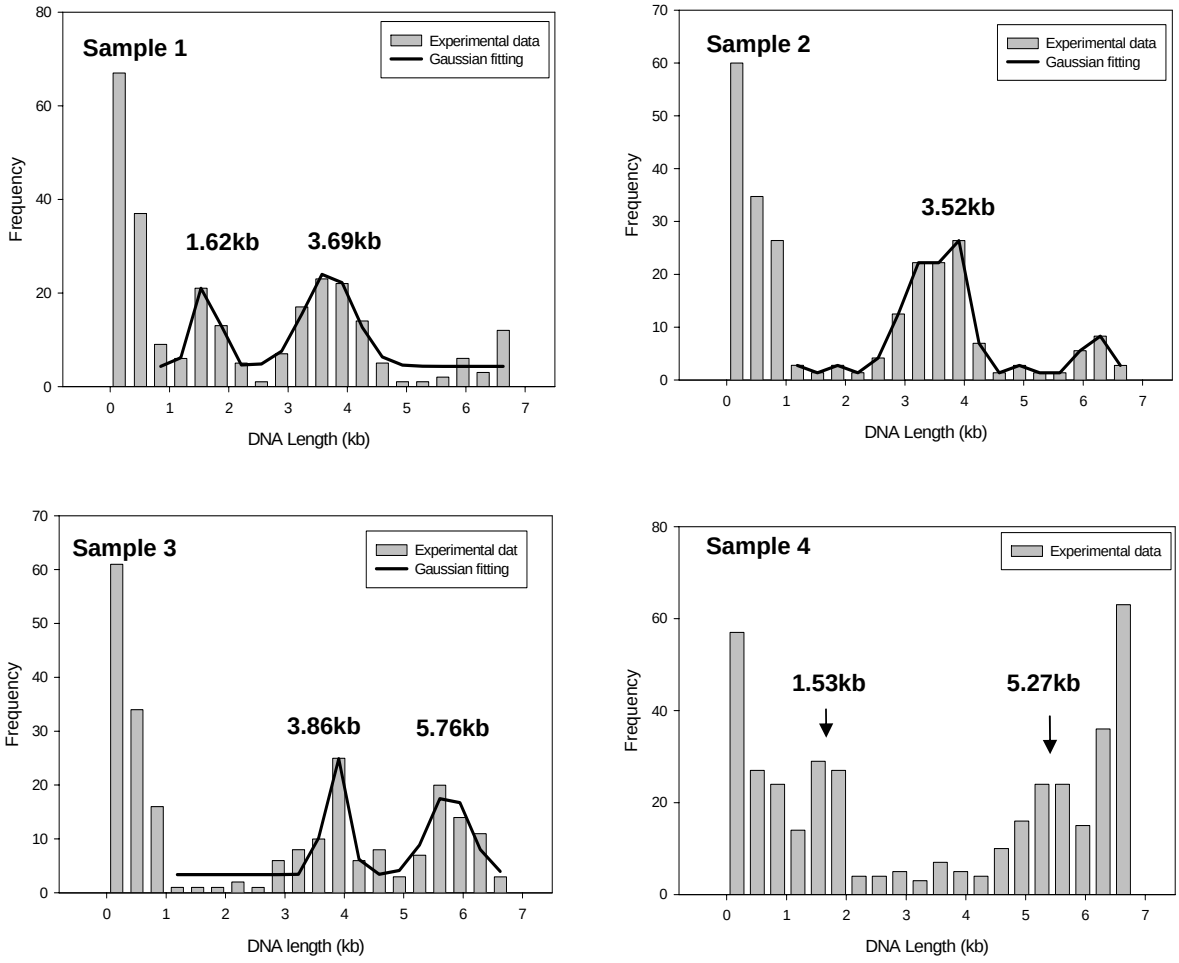

Figure S3. The constructed sequence motif maps. The sequence motif maps in the graph were obtained by analyzing 100-200 DNA molecules. The solid line is the Gaussian curve fitting and the peaks correspond well to the predicted locations of the sequence motif.