



**Figure S2** The average nucleotide frequencies flanking the insertion sites were determined. All Hermes insertions were aligned and the nucleotide patterns were analyzed by sequence logo created by a program written in Visual Basic based on an existing algorithm (Schneider and Stephens, 1990). The positions of the eight nucleotides at the target sites that are duplicated during integration are indicated by TSD. The nucleotide preferences for the *in vivo* integration sites are shown in (A) and the *in vitro* integration sites are shown in (B).

Schneider, T. D., and Stephens, R. M., 1990 Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* 18:6097-6100.