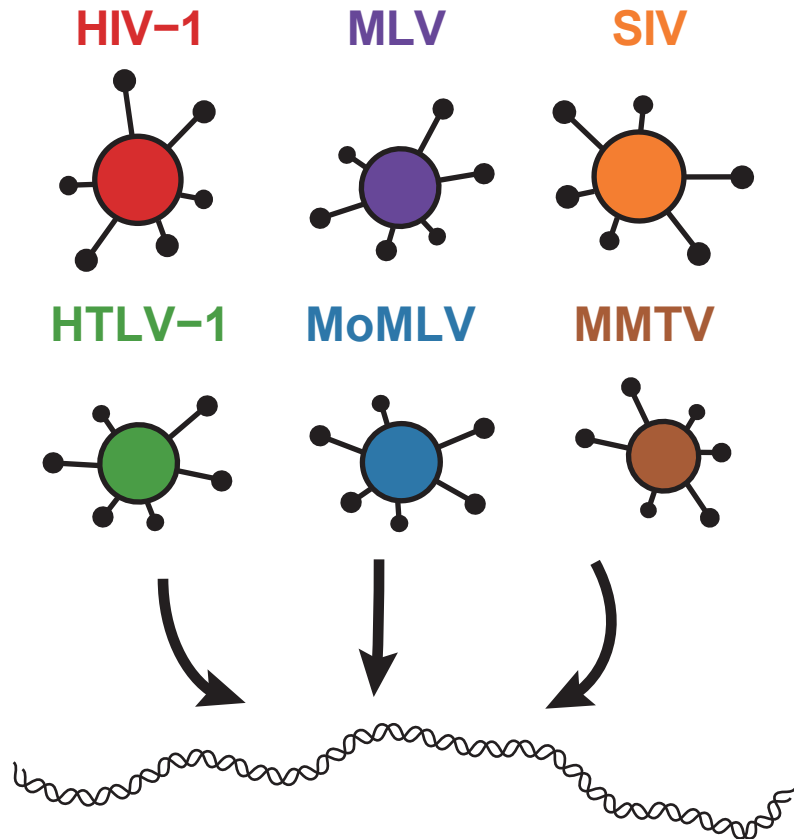


# Effects of biochemical (DNA motif) and biophysical (DNA shape) DNA properties on the integration specificity across retroviruses

What determine the integration specificity in host genome?



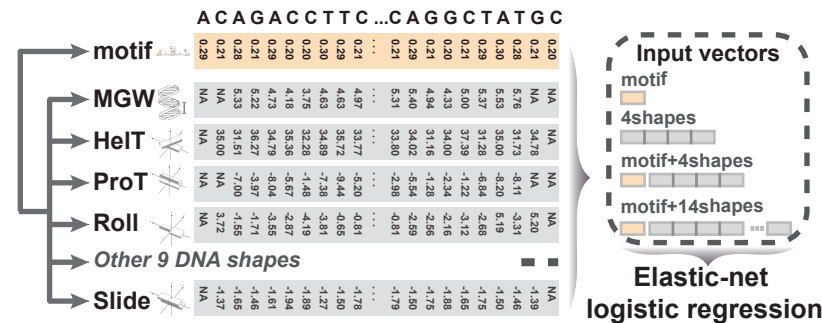
## QRIS framework

(Quantify the Retrovirus Integration Specificity)

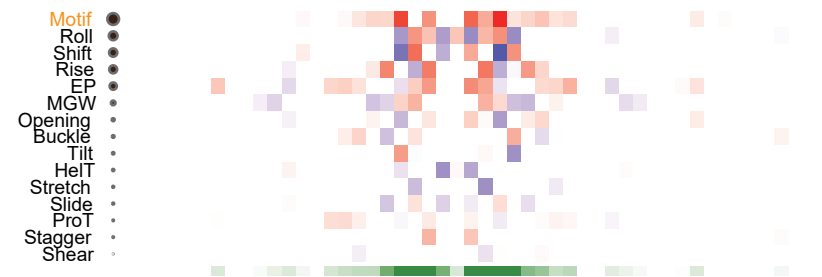
**Step 1: Train the machine learning model**

Integration sites (n=10,000)  
Random sites (n=10,000)

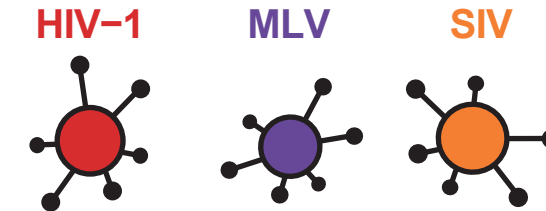
Encode DNA motif & DNA shape information  
surrounding a 51bp region for each site



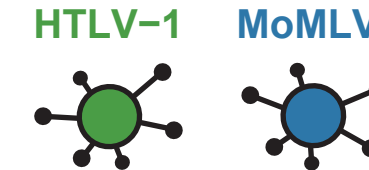
**Step 2: Extract the coefficient**



① StrongFavor retrovirus



② WeakFavor retrovirus



③ StrongShape retrovirus



**Conclusions:**

DNA motif and DNA shape cooperatively have strong or weak effects on the integration specificity of StrongFavor or WeakFavor retrovirus;

DNA shape independently and strongly affects the integration specificity of StrongShape retrovirus.