a

CR   SV   HV1   HV2

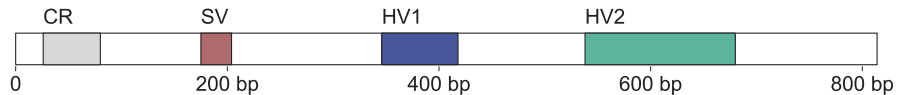0   200 bp   400 bp   600 bp   800 bp

b

(1) Calculate k-mer distances between *opa* for each variable region
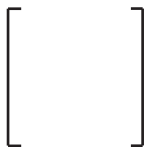
Hypervariable 1 sequence example

Distance matrix

caacacaaaaaaggtgaac...
caacaca
 aacacaa
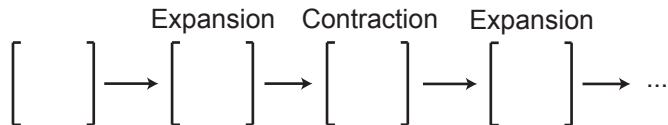 acacaaa
   cacaaaa
    acaaaaa
     caaaaaa
      ...

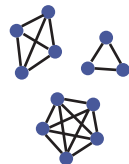→ Find # matching k-mers between sequences →

(2) Create sequence clusters using network clustering

Expansion   Contraction   Expansion

→ ... → Network clustering

c

occurrences

Semivariable cluster type

occurrences

Hypervariable 1 cluster type

occurrences

Hypervariable 2 cluster type