

hg19.fa, genes

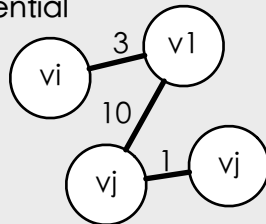
Extract exon sequence of targeted genes from reference genome; this step can be skipped if such sequences already exist.

Input  
exon.fa R1.fq R2.fq

Index exon.fa by  
k-mer hash table

Construct Breakend  
Associated Graph (BAG)

In  $BAG=(V,E)$ , any node  $v_i$  indicates a gene and edge  $e_{ij}$  indicates potential fusion between  $v_i$  and  $v_j$ . For any  $e_{ij}$  there is a integer  $w(e_{ij})$  indicates the number of reads that support fusion between  $v_i$  and  $v_j$ .

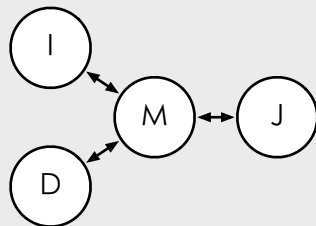


Identify fusion  
junction by  
fitting alignment

Construct fusion transcript

Realign reads against  
fused transcript

For any edge  $e_{ij}$  with  $w(e_{ij})$  greater than 4, reads that support this edge will be aligned to the string concatenated by  $v_i$  and  $v_j$  using modified fitting algorithm and junction site will be identified



Score fusion

Output  
fusion.out