Generate all substrings of length k from the reads ACAGC TCCTG GTCTC AGCGC CTCTT GGTCG CACAG TTCCT GGTCT CAGCG CCTCT TGGTC CCACA TGGTC TCAGC TTGGT TTGTT CCCAC GCTTC CTGGT CTCAG TTCCT GTTGG k-mers (k=5) GCCCA CGCTT CCTCA GCTGG CTTCC TGTTG CGCCC GCGCT TGCTG TCTTG CCCTC GCTTC TTGTT CGTAG CCGCC AGCGC GCCCT TCGTA ACCGC GTCGT ACCGCCCACAGCGCTTCCTGCTGGTCTCTTGTTG CGCCCTCAGCGCTTCCTCTTGTTGGTCGTAG b Generate the De Bruijn graph Sequencing error or SNP > CCCAC > CCACA > CACAG > ACAGC ACCGC - CCGCC - CGCCC CAGCG → AGCGC → GCGCT → CGCTT → GCTTC GCCCT -> CCCTCA -> CTCAG -> TCAGC Deletion or intron CTGCT - TGCTG - GCTGG - CTGGT - TGGTC - GGTCT - GTCTC - TCTCT CTTCC - TTCCT TCCTC - CCTCT ctctt \rightarrow (cttg) \rightarrow (cttg)c Collapse the De Bruijn graph TCCTGCTGGTCTCT GCCCACAGC CTCTTGTTGGTCGTAG CAGCGCTTCCT ACCGCCC TCCTCT GCCCTCAGC d Traverse the graph

ACCGCCCACAGCGCTTCCTGCTGGTCTCTTGTTGGTCGTAG

Assembled isoforms