## Draft assembly sequence

- 1. Check each word of size *k* (kmer) in filter
- 2. Check k kmer subset  $(S_k)$  for absence
  - 3. Permutate 3'-end base
    - Check k kmer\* subset  $(S_{k\_alt})$  for presence

Edited draft assembly sequence

If  $S_k^- \ge k/x$ :

If  $S_{k\_alt}^+ \ge k/y$ : apply change to sequence, resume 3

- Insert 3'-end positions
   Check k kmer subset presence
   If S<sub>k alt</sub> <sup>+</sup> ≥ k/y : apply change, resume 1
  - 5. Delete 3'-end positions Check k kmer subset presence If  $S_{k \ a/t}^+ \ge k/y$ : apply change, resume 1

haploid or diploid DNA source

Sequence reads



\*kmers with alternate 3'end base (k alt)