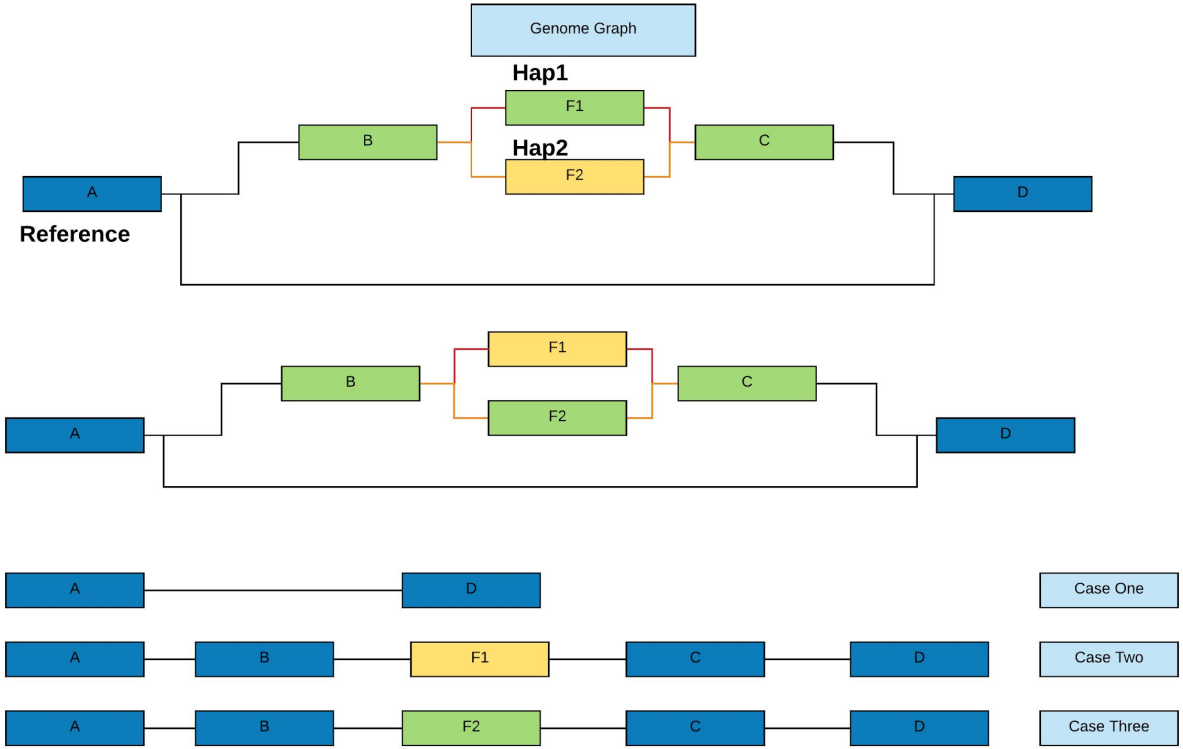
# Graph to Haplotype representation

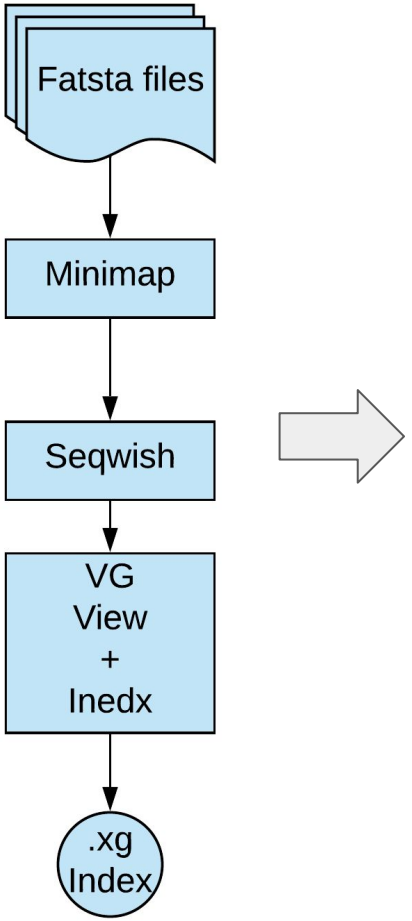



chr21:1-100
chr21:1-50>+H21_1:1-100>+chr21:51-100
chr21:1-50>+H21_1:1-20>+H21_2:1-40>+H21_1:60-100>chr21:51-100

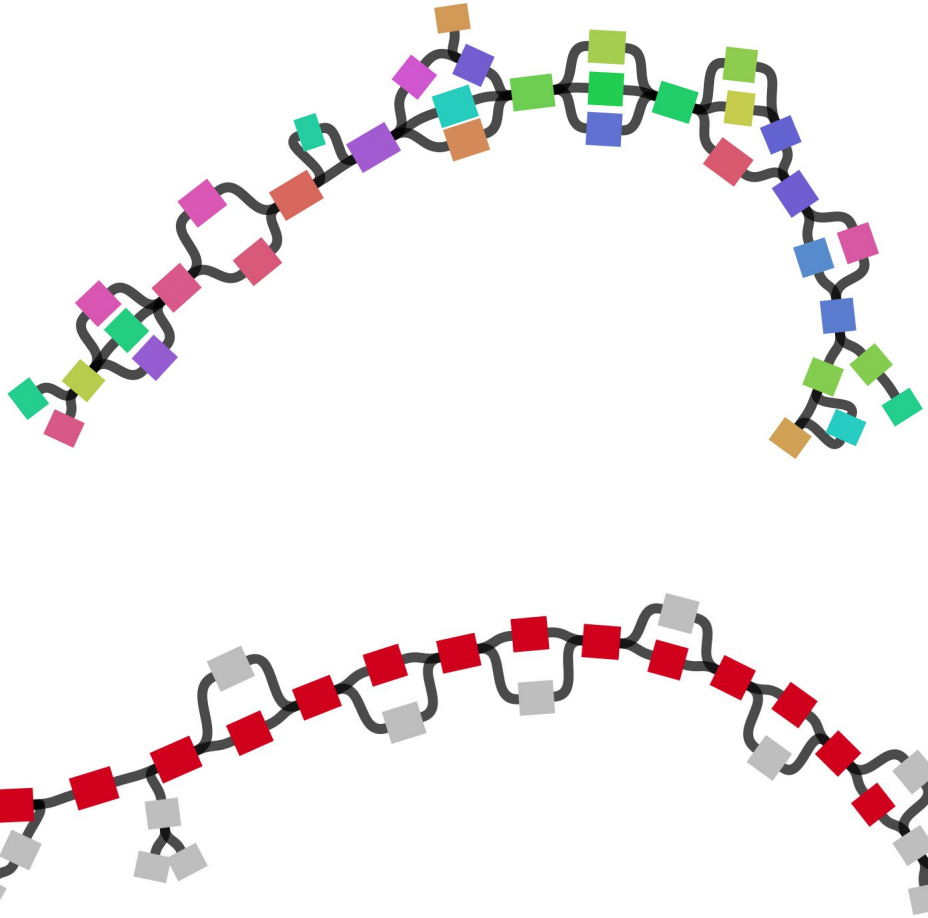
# Rationale

chr21:1-50>+H21_1:1-20>+H21_2:1-40>+H21_1:60-100>chr21:51-100

- Ability to add new haplotype without altering existing reference coordinate
- Reused part of shared sequence among haplotypes (partial overlap, inversion)
- Support patch of reference genome
- Ability to swap out and create custom reference based on preferable haplotype.


Fatsta files
Minimap
Seqwish
VG
View
+
Inedx
.xg
Index





# How to add new haplotype?

- First come first serve, but major haplotype if possible
- No breaking existing haplotype
- Build new Hap based on the closet Hap from MSA
  - Chain file indication junction of haplotype
  - Reuse prefix/suffix of existing haplotype

# When to add new haplotype?

- Sufficient shared allele frequency (>1%) and number of individual (>10k people)
- Significant alteration (X% mismatch/indel/ N bp missing)
    - Caveat: different mapping would have different sensitivity to handle diverged sequence