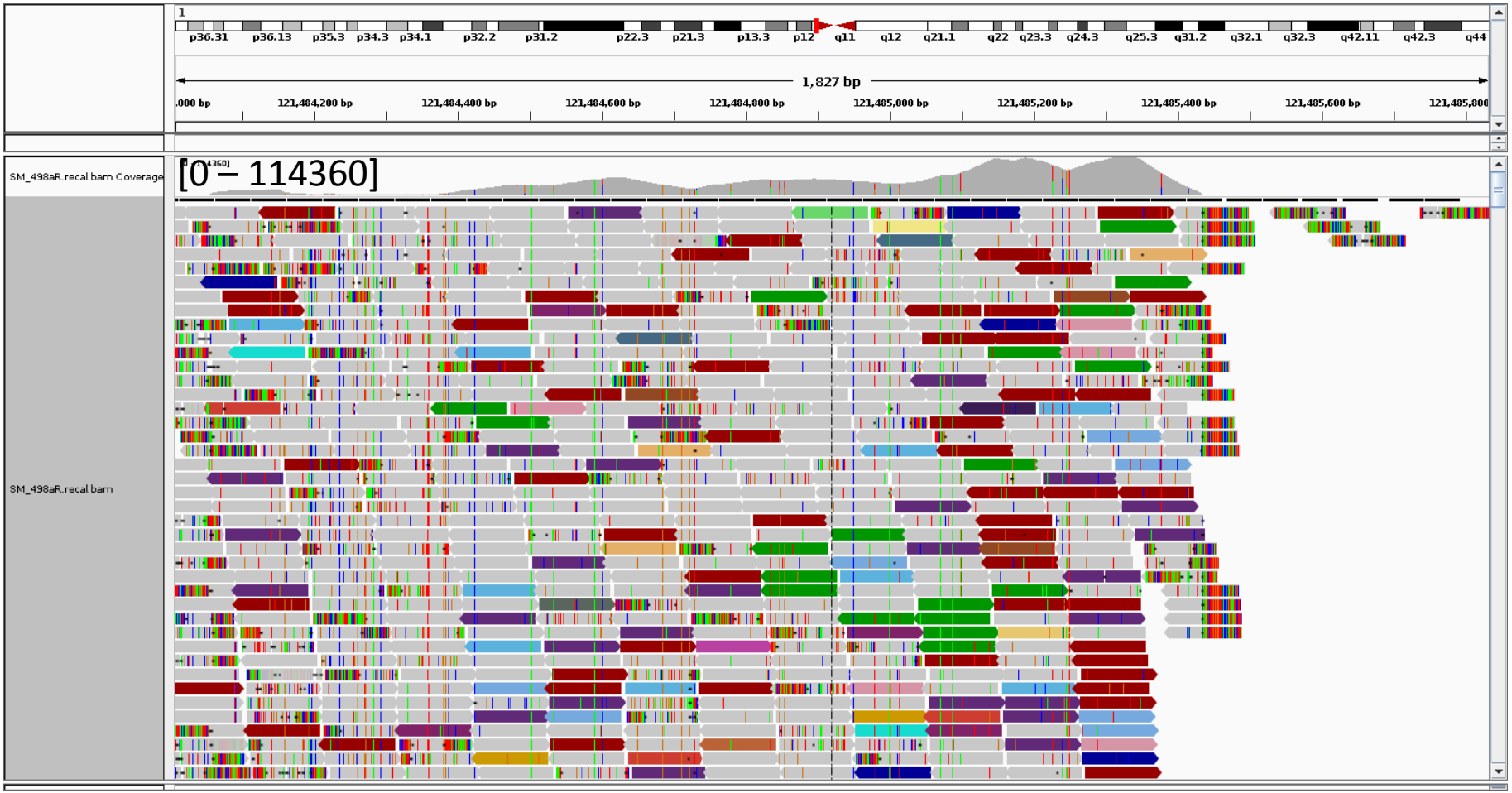
FILTERING SVS

Lots of noise!



Sources of noise – what filter should we use?

- Repeat regions
- High-depth regions
- Poor quality mapping
- Mobile elements
- Bacterial genome insertion
- Viral genome insertion
- Poor quality reference (telomere and centromere)

Other common filters

- Read depth
- Reads supporting both sides of the break
- Concomitant copy-number change

BRASS – Breakpoint by assembly

- Supporting read > 4
- Remove read groups overlapping:
  - repeats
  - high GC content
  - high-depth regions
  - known viral insertion sites
  - known bacterial insertion sites
  - telomeric and centromeric regions
- Require events to have
  - Concomitant copy-number change
  - Assembly support

LOOK AT YOUR BRASS OUTPUT

EXERCISE 3