

**X-SVLen**

# Input:

**ONT based SV callset using cuteSV**

**Truvari run using default and permissive params**

**hg37**

**TrioCanu assemblies**

# Method:

For each SV defined by chrom, pos, variant:

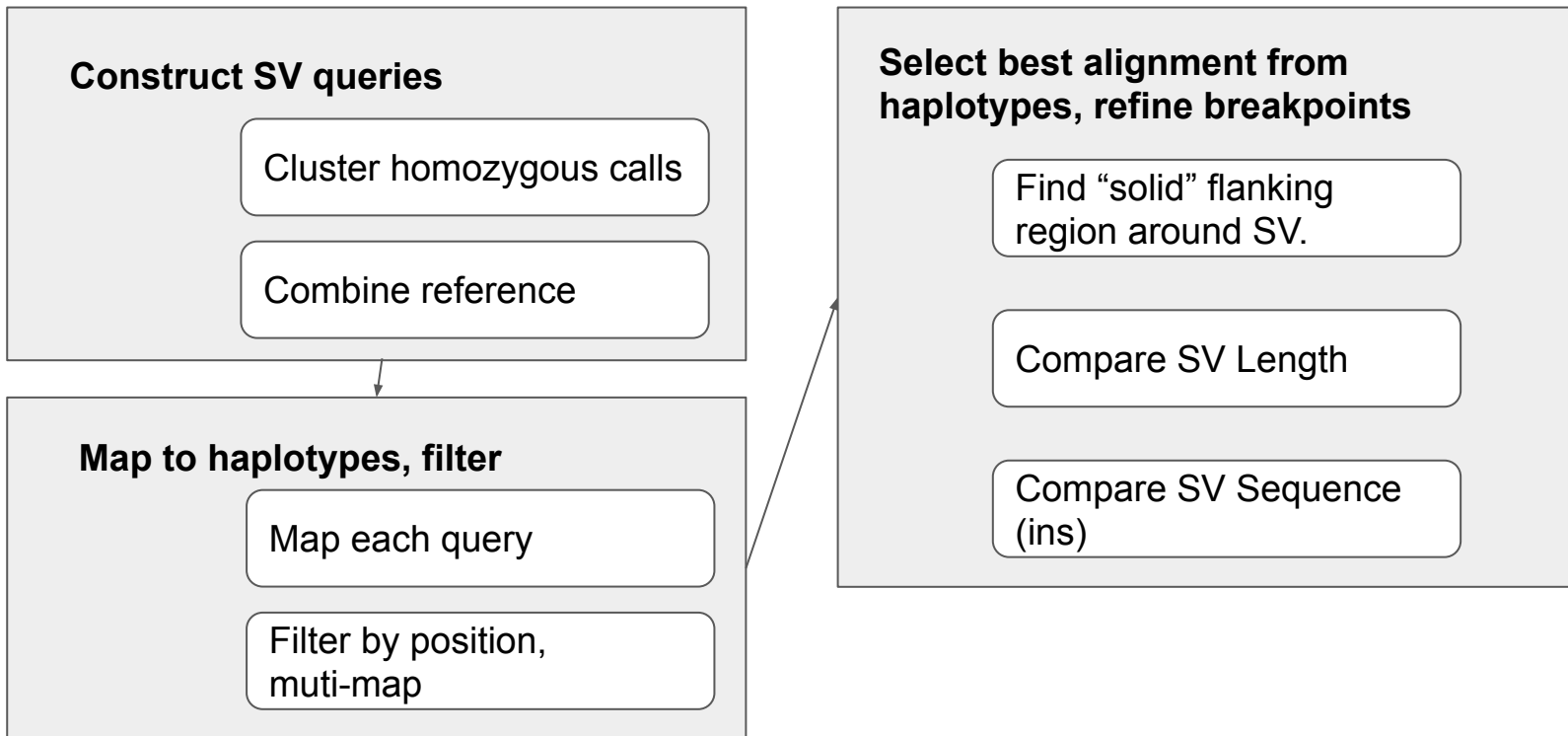
Construct a query as:

`ref[chrom][pos-prefix:pos] + SV + ref[chrom][pos:pos+suffix]`



Map all queries to hap1 and hap2 assemblies.

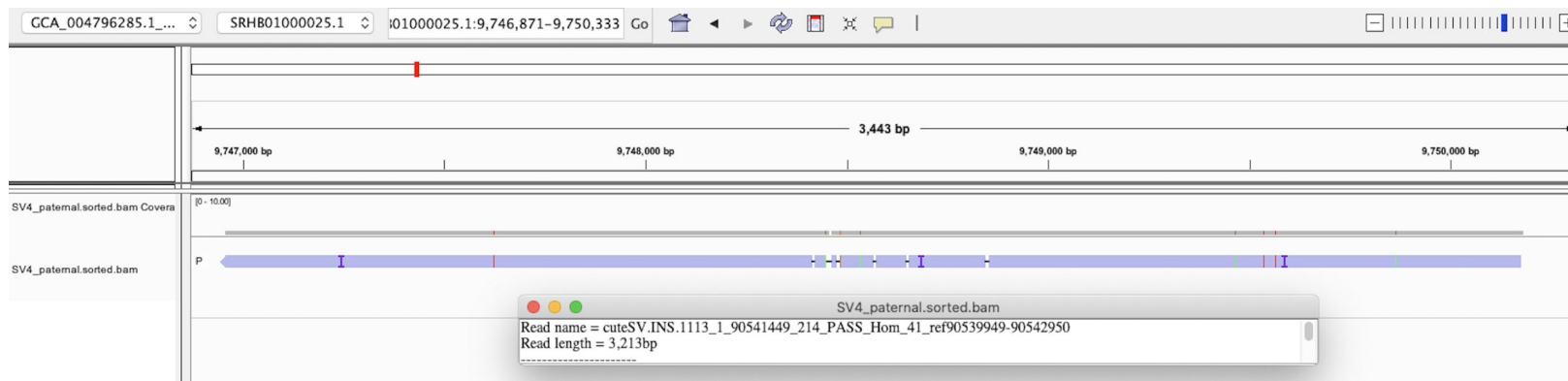
# Workflow



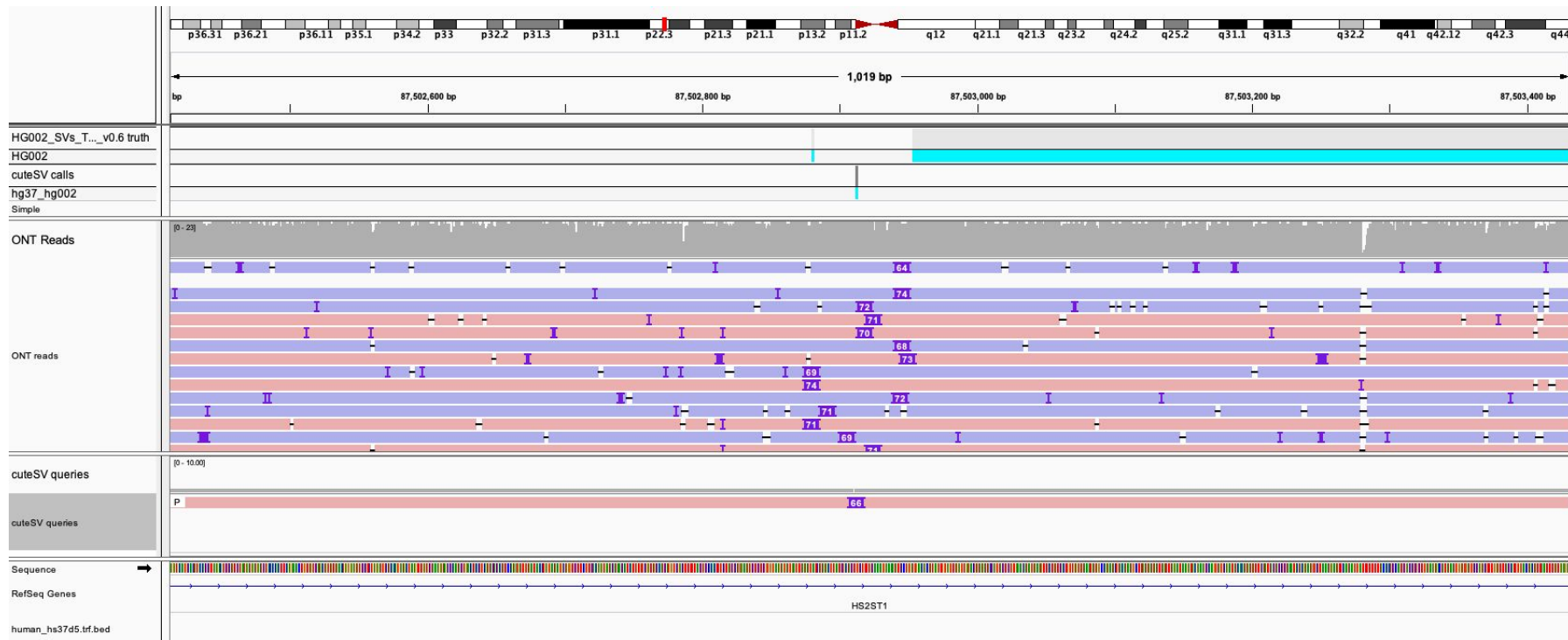
# Assembly calls that overlap contigs.

SVType SVLen	INS		DEL	
	<50	>=50	<50	>=50
number of these calls that overlap assembly contigs	16054	17481	30494	6475

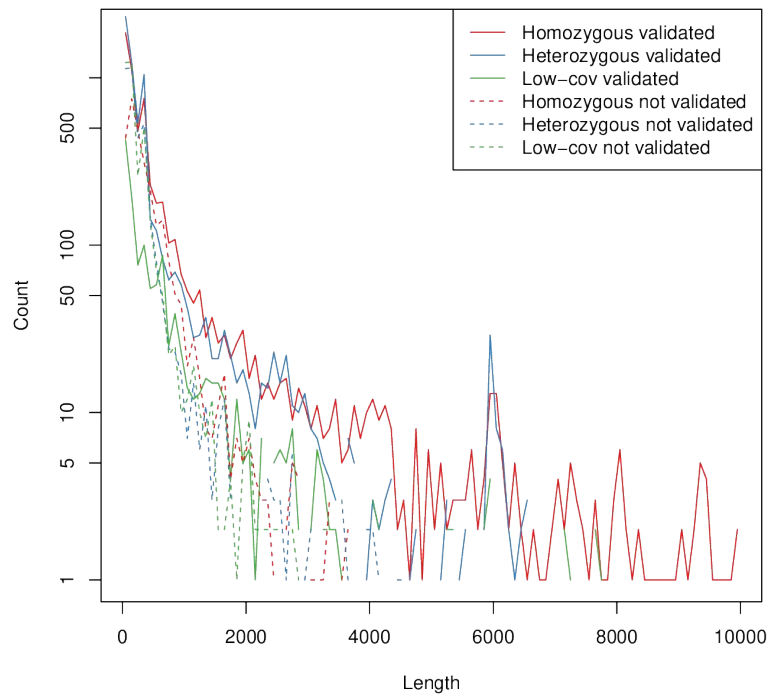
# Mapped query sequence



# Example difficult to validate call.



# Annotation of true positives





# Comparison against Truvari

SVType SVLen	INS		DEL	
	<50	>=50	<50	>=50
number of truvari TP/FP/no-call	3351 FP	2835 FP	880 FP	1571 FP
	7200 TP	4095 TP	1925 TP	2099 TP
	15974	80	30196	298
	not_processed	not_processed	not_processed	not_processed
number of assembly TP/FP/no-call	5220 FP	8979 FP	23202 FP	3444 FP
	10834 TP	8502 TP	7292 TP	3031 TP

# Thank you!!

Rocio Esteban

Debapriya Sinha

Chunxiao Liao

And all the organizers of the 2020 eHackathon YOU ARE AWESOME