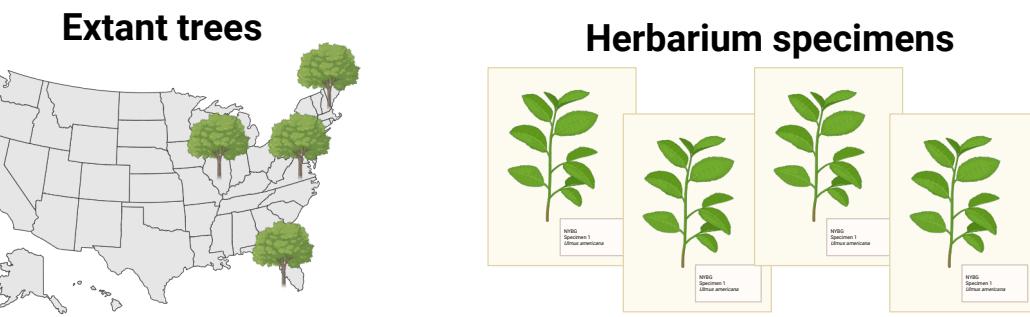
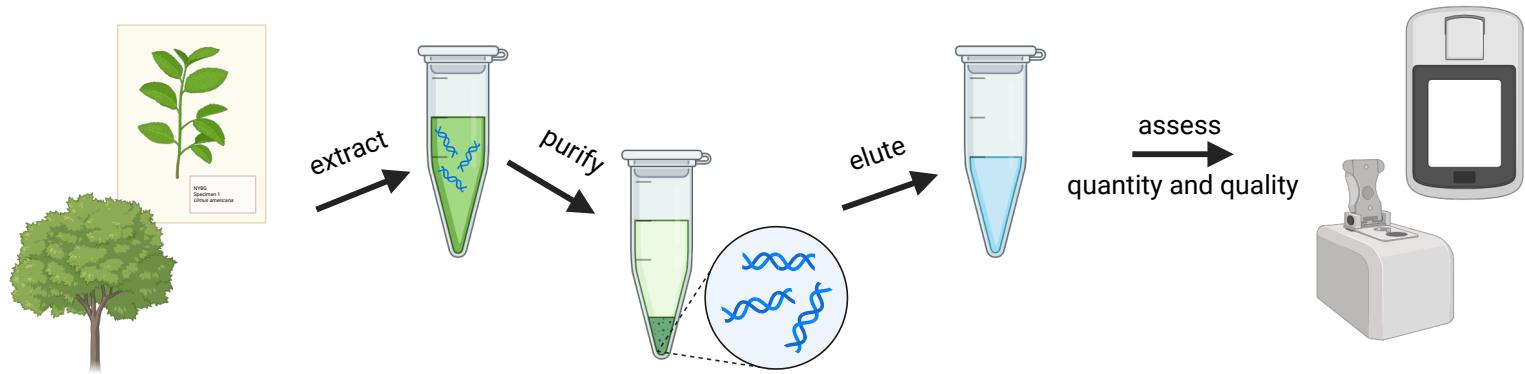


Step 1a) Sample collection

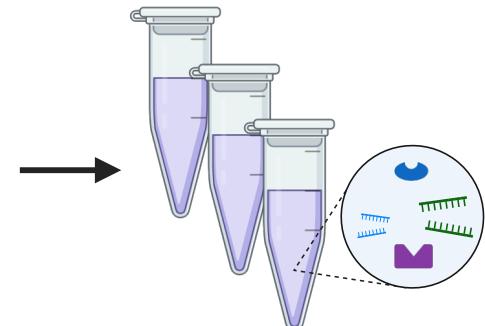


Step 2a) DNA extraction with modified CTAB protocol

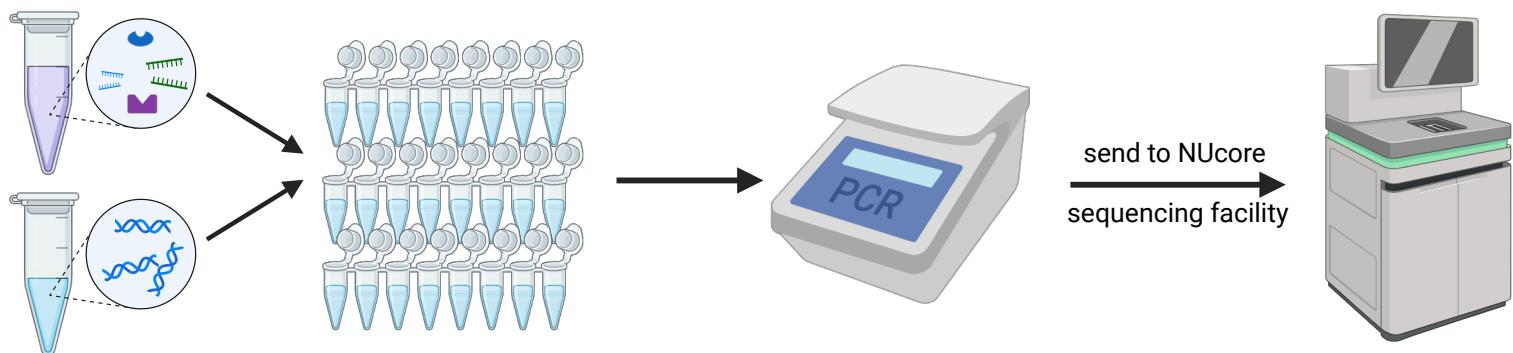


Step 3a) Create custom amplicon panel

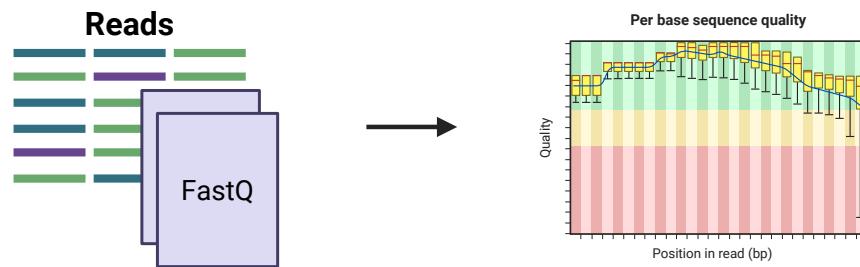
```
>elm gene 1  
ATGCATTGGCGCGTACTAGCTGCTAACG  
GCGCGTACTAGCTGCTAACGCTTAGCAT  
>elm gene 2  
CGTACTAGCTGCTAACGCTTAGCATCTA  
AGCATTAGCATCTATACTAGCTGCTAACG  
>disease gene 1  
AAGCATTAGCATCTAATGCATTGGCGCG  
TACTAGCTGCTAACGCTACTAGCTGTAG
```



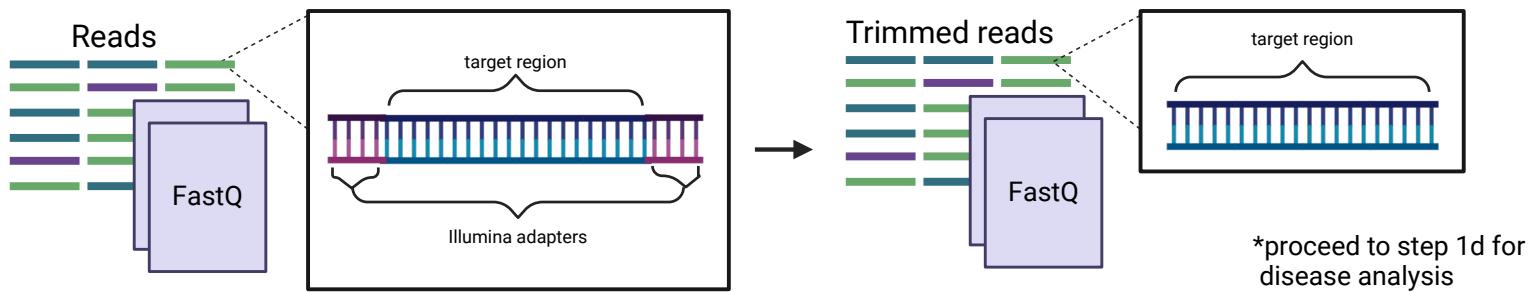
Step 4a) Library preparation and sequencing



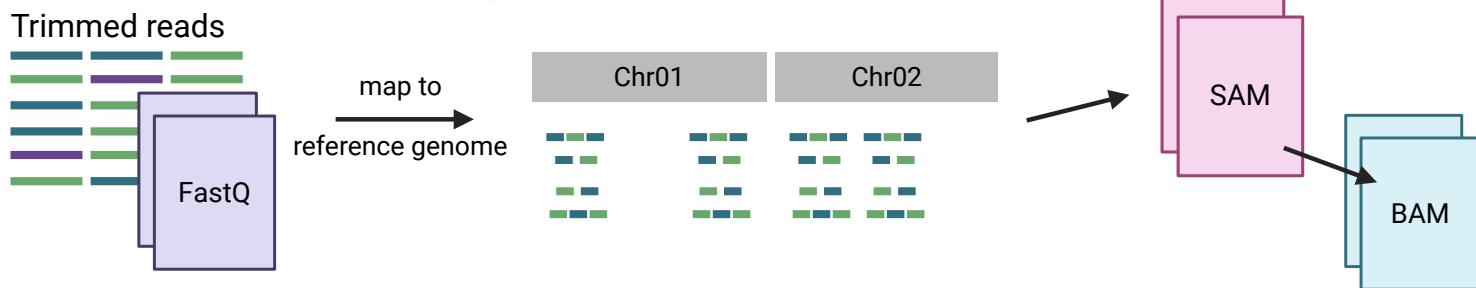
Step 1b) Quality control with FastQC and MultiQC



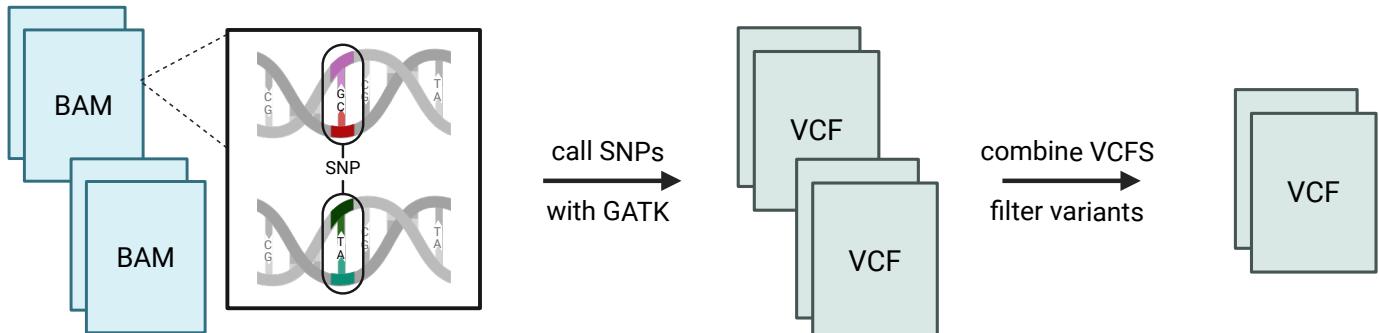
Step 2b) Trimming with Trimmomatic



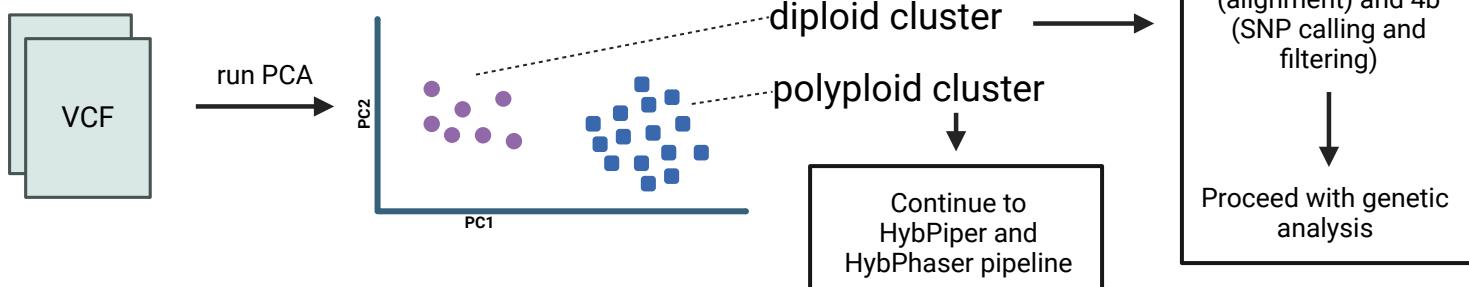
Step 3b) Align with to reference with BWA



Step 4b) Call SNPs with GATK

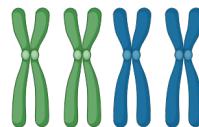


Step 5b) Identify cytotypes



Step 0) Allopolyploid subgenome identification

Subgenome A:
diploid *U. americana*



Subgenome B:
unknown (likely *U. rubra*,
used *U. glabra* as reference)

Step 1) Assemble allopolyploid reads with HybPiper

Reads

Allopolyploid samples +
subgenome A + subgenome B references

Map to target file

Gene 1

Gene 2



Step 2) Generate subgenome consensus sequences with HybPhaser I

Subgenome A

Gene 1

Gene 2

Subgenome B

Gene 1

Gene 2

Step 3) Subgenome Association and Phasing with HybPhaser II & III

Input

Subgenome References

On-Target Reads

Reference Mapping

map to subgenome consensus sequences



Subgenome Association HybPhaser II

	subgenome A	subgenome B
tree01	0.46	0.25
tree02	0.32	0.14

Phased Reads HybPhaser III

Subgenome A Subgenome B

Step 4) Assemble phased reads with HybPiper

Phased reads -
subgenome A

Map to subgenome consensus sequences

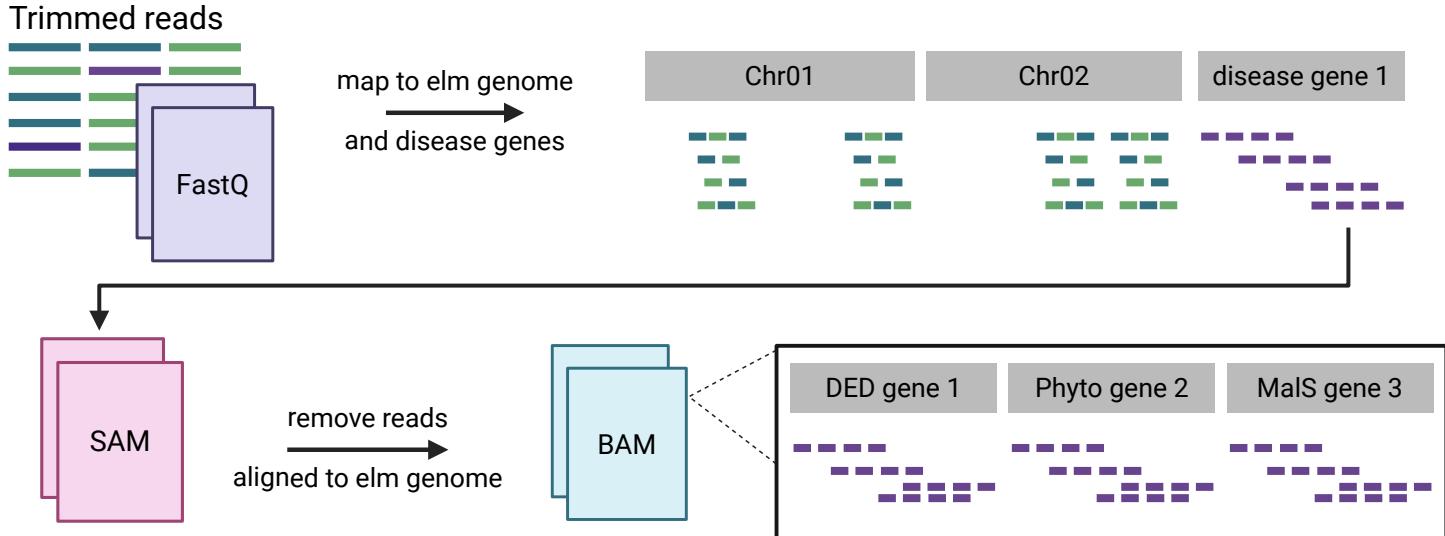


Phased reads -
subgenome B

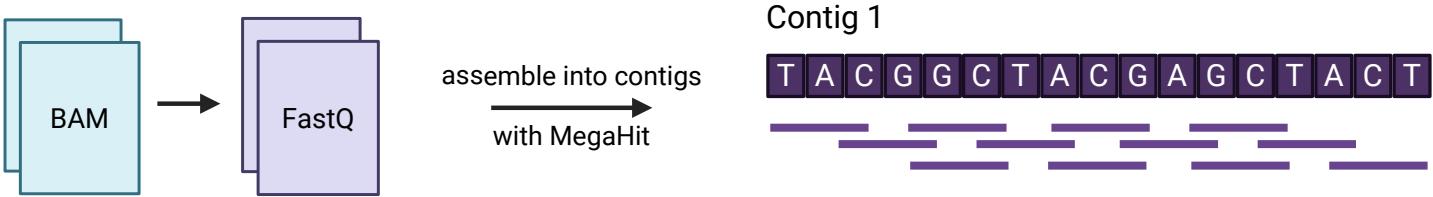
SNP calling and VCF
filtering of individual
subgenomes

Combine VCF files and
proceed with genetic
analysis

Step 1d) Align to reference with BWA



Step 2d) Assemble reads into contigs with MegaHit



Step 3d) Compare contigs to reference database

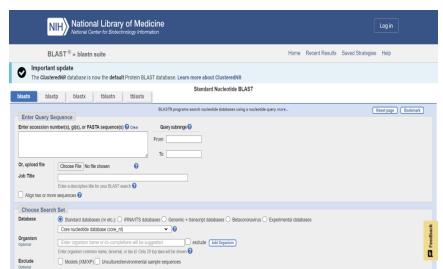
Contig 1

```
T A C G G C T A C G A G C T A C T
```

Contig 2

```
A T G T C G T C A A T A G C G C A
```

Blast against
NCBI Nucleotide Database



Step 4d) Identify low and high confidence infections

Low confidence

An individual who has 1 or more Blast hits matching the **family** of the causal agent.

High confidence

An individual who has 1 or more Blast hits matching the **genus** of the causal agent.

tree01	DED	Phyto	MalS
Low confidence	4	0	0
High confidence	1	0	0

tree02	DED	Phyto	MalS
Low confidence	0	2	0
High confidence	0	0	0

Map of Infected Trees

