

For this example, we will look at data from donor CAP256 34 weeks after infection, from [Doria-Rose et al, J. Virol 2015](#). This is the first time point at which the CAP256-VRC26 antibody lineage was detected. We will use the first 12 members of this lineage (available from [GenBank](#)).

Start by creating a working folder and downloading the data:

```
schrammca@a1-hpcn095:~$ mkdir cap256-week34H
schrammca@a1-hpcn095:~$ cd cap256-week34H/
schrammca@a1-hpcn095:~/cap256-week34H$ fastq-dump SRR2126754
Read 518852 spots for SRR2126754
Written 518852 spots for SRR2126754
```

Now, run module 1 to annotate the V, D, and J genes and collapse duplicate reads:

```
schrammca@a1-hpcn095:~/cap256-week34H$ sonar blast_V --fasta SRR2126754.fastq --locus H --derep --threads 8
```

```
Tue Feb 12 12:10:26 2019 -- SONAR b'' run with command:
/nethome/schrammca/sonar/annotate/1.1-blast_V.py --fasta SRR2126754.fastq --locus H --derep --threads 8
```

```
vsearch v2.8.1_linux_x86_64, 251.8GB RAM, 16 cores
https://github.com/torognes/vsearch
```

```
Reading file /nethome/schrammca/cap256-week34H/work/internal/tempForDerep.fa 100%
204399756 nt in 518852 seqs, min 40, max 1196, avg 394
Dereplicating 100%
Sorting 100%
260705 unique sequences, avg cluster 2.0, median 1, max 767
Writing output file 100%
Writing uc file, first part 100%
Writing uc file, second part 100%
```

```
10122 processed, 10000 good; starting file cap256-week34H_002
20688 processed, 20000 good; starting file cap256-week34H_003
32287 processed, 30000 good; starting file cap256-week34H_004
44519 processed, 40000 good; starting file cap256-week34H_005
56559 processed, 50000 good; starting file cap256-week34H_006
68344 processed, 60000 good; starting file cap256-week34H_007
79953 processed, 70000 good; starting file cap256-week34H_008
91476 processed, 80000 good; starting file cap256-week34H_009
102816 processed, 90000 good; starting file cap256-week34H_010
114164 processed, 100000 good; starting file cap256-week34H_011
125401 processed, 110000 good; starting file cap256-week34H_012
136561 processed, 120000 good; starting file cap256-week34H_013
147653 processed, 130000 good; starting file cap256-week34H_014
158730 processed, 140000 good; starting file cap256-week34H_015
169783 processed, 150000 good; starting file cap256-week34H_016
180827 processed, 160000 good; starting file cap256-week34H_017
191840 processed, 170000 good; starting file cap256-week34H_018
202894 processed, 180000 good; starting file cap256-week34H_019
213898 processed, 190000 good; starting file cap256-week34H_020
224963 processed, 200000 good; starting file cap256-week34H_021
235968 processed, 210000 good; starting file cap256-week34H_022
249256 processed, 220000 good; starting file cap256-week34H_023
TOTAL: 260705 processed, 228881 good
```

```
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_001.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_002.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_003.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_004.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_005.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_006.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_007.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_008.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_009.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_010.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_011.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_012.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_013.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_014.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_015.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_016.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_017.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_018.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_019.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_020.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_021.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_022.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/vgene/cap256-week34H_023.fasta against /nethome/schrammca/sonar/germDB/IgHV.fa...
```

```
schrammca@a1-hpcn095:~/cap256-week34H$ sonar blast_J --threads 8
```

```
Tue Feb 12 12:22:26 2019 -- SONAR b'' run with command:
/nethome/schrammca/sonar/annotate/1.2-blast_J.py --threads 8
```

```
curating 5'end and strand...
10000 done, 9992 good...
20000 done, 19975 good...
30000 done, 29945 good...
40000 done, 39918 good...
50000 done, 49874 good...
60000 done, 59838 good...
70000 done, 69804 good...
80000 done, 79774 good...
90000 done, 89741 good...
100000 done, 99714 good...
110000 done, 109677 good...
120000 done, 119646 good...
130000 done, 129616 good...
140000 done, 139591 good...
150000 done, 149564 good...
160000 done, 159544 good...
170000 done, 169524 good...
180000 done, 179496 good...
190000 done, 189475 good...
200000 done, 199447 good...
210000 done, 209432 good...
220000 done, 219392 good...
228881 done, 228244 good...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_001.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_002.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_003.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_004.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_005.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_006.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_007.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_008.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_009.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_010.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_011.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_012.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_013.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_014.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_015.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_016.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_017.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_018.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_019.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_020.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_021.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_022.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_023.fasta against /nethome/schrammca/sonar/germDB/IgHJ.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_001.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_002.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_003.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_004.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_005.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_006.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_007.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_008.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
```

```
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_009.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_010.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_011.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_012.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_013.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_014.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_015.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_016.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_017.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_018.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_019.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_020.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_021.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_022.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_023.fasta against /nethome/schrammca/sonar/germDB/IgHC_CH1.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_001.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_002.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_003.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_004.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_005.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_006.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_007.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_008.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_009.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_010.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_011.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_012.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_013.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_014.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_015.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_016.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_017.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_018.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_019.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_020.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_021.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_022.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
Starting blast of /nethome/schrammca/cap256-week34H/work/annotate/jgene/cap256-week34H_023.fasta against /nethome/schrammca/sonar/germDB/IgHD.fa...
```

```
schrammca@a1-hpcn095:~/cap256-week34H$ sonar finalize
```

```
Tue Feb 12 12:26:25 2019 -- SONAR b'' run with command:
/nethome/schrammca/sonar/annotate/1.3-finalize_assignments.py
```

```
loading sequence info from /nethome/schrammca/sonar/germDB/IgHV.fa...
loading sequence info from /nethome/schrammca/sonar/germDB/IgHV.fa...
curating junction and 3' end...
/sysapps/cluster/software/BioPython/1.73-foss-2016b-Python-3.7.1/lib/python3.7/site-packages/Bio/Seq.py:2609: BiopythonWarning: Partial codon, len(sequence) not a multiple of three, E
10000 done, found 8676; 8778 good...
20000 done, found 13936; 16348 good...
30000 done, found 29750; 23128 good...
40000 done, found 39434; 29927 good...
50000 done, found 49118; 36755 good...
60000 done, found 58820; 43626 good...
70000 done, found 68524; 50423 good...
80000 done, found 78243; 57011 good...
90000 done, found 87940; 63554 good...
100000 done, found 97642; 69949 good...
110000 done, found 107344; 76227 good...
120000 done, found 117029; 82437 good...
130000 done, found 126759; 88461 good...
140000 done, found 136502; 94281 good...
150000 done, found 146229; 99865 good...
160000 done, found 155963; 105351 good...
170000 done, found 165705; 110443 good...
180000 done, found 175467; 115338 good...
190000 done, found 185240; 119846 good...
200000 done, found 195000; 123931 good...
210000 done, found 204784; 127522 good...
220000 done, found 214435; 133630 good...
228881 done, found 223022; 139448 good...

Total raw reads: 260702
Correct Length: 228881
V assigned: 228790
J assigned: 223022
CDR3 assigned: 221892
In-frame junction: 192304
No indels: 142657
Continuous ORF with no stop codons: 139448
```

```
schrammca@a1-hpcn095:~/cap256-week34H$ sonar cluster_Sequences --id .97 --min2 2
```

```
Tue Feb 12 12:36:51 2019 -- SONAR b'' run with command:
/nethome/schrammca/sonar/annotate/1.4-cluster_sequences.py --id .97 --min2 2
```

```
vsearch v2.8.1 linux x86_64, 251.8GB RAM, 16 cores
https://github.com/torognes/vsearch
```

```
Reading file temp.fa 100%
51161824 nt in 139448 seqs, min 127, max 445, avg 367
Dereplicating 100%
Sorting 100%
90756 unique sequences, avg cluster 4.2, median 1, max 4215
Writing output file 100%
Writing uc file, first part 100%
Writing uc file, second part 100%
```

```
vsearch v2.8.1 linux x86_64, 251.8GB RAM, 16 cores
https://github.com/torognes/vsearch
```

```
Reading file temp dedup.fa 100%
33294745 nt in 90756 seqs, min 127, max 445, avg 367
Masking 100%
Sorting by abundance 100%
Counting k-mers 100%
Clustering 100%
Sorting clusters 100%
Writing clusters 100%
Clusters: 42826 Size min 1, max 7156, avg 2.1
Singletons: 29330, 32.3% of seqs, 68.5% of clusters
```

Let's look at different ways to divide the bulk data into lineages:

```
schrammca@a1-hpcn095:~/cap256-week34H$ sonar id-div -a CAP256-VRC26.01-12H.fa -t 8
```

```
Tue Feb 12 12:58:11 2019 -- SONAR b'' run with command:
/nethome/schrammca/sonar/lineage/2.1-calculate_id-div.py -a CAP256-VRC26.01-12H.fa -t 8
```

```
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000000.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000001.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000002.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000003.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000004.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000005.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000006.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000007.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000008.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000009.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000010.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000011.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000012.fa...
Starting work on /nethome/schrammca/cap256-week34H/work/lineage/align/align000013.fa...
```

Note that this next script generates an interactive plot allowing you to select the island corresponding to the lineage of interest.

```
schrammca@cogsworth:/nethome/schrammca/cap256-week34H$ sonar get_island output/tables/cap256-week34H_goodVJ_unique_id-div.tab --mab CAP256-VRC26.01 --mab CAP256-VRC26.08
```

```
Loading required package: misc3d
```

```
Left click to accept, right click to re-select region
Left click to accept, right click to re-select region
[1] "Saved 129 transcript IDs to output/tables/islandSeqs"
There were 50 or more warnings (use warnings() to see the first 50)
```

Another option is a phylogenetic based analysis:

```
schrammca@cogsworth:~/nethome/schrammca/cap256-week34H$ sonar intradonor --n CAP256-VRC26.01-12H.fa --v IGHV3-30*18 --threads 8 -f

loading sequence info from CAP256-VRC26.01-12H.fa...
loading sequence info from /nethome/schrammca/sonar/germDB/IgHV.fa...

Tue Feb 19 16:47:54 2019 -- SONAR b'v2.0.0-133-g091cbcl-dirty' run with command:
/nethome/schrammca/sonar/lineage/2.3-intradonor_analysis.py --n CAP256-VRC26.01-12H.fa --v 'IGHV3-30*18' --threads 8 -f

16:47:54 - Starting a new analysis from scratch...
loading reads from /nethome/schrammca/cap256-week34H/output/sequences/nucleotide/cap256-week34H_goodVJ_unique.fa assigned to IGHV3-30...
2808 loaded...
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00001.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00002.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00003.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00004.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00005.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00006.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00007.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00008.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00009.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00010.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00011.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00012.fa
Found 26 reads in subtree #1. Total saved so far: 27 / 249
Found 6 reads in subtree #2. Total saved so far: 34 / 499
Found 25 reads in subtree #3. Total saved so far: 60 / 749
Found 13 reads in subtree #4. Total saved so far: 74 / 999
Found 228 reads in subtree #5. Total saved so far: 303 / 1249
Found 9 reads in subtree #6. Total saved so far: 313 / 1499
Found 221 reads in subtree #7. Total saved so far: 535 / 1749
Found 4 reads in subtree #8. Total saved so far: 540 / 1999
Found 2 reads in subtree #9. Total saved so far: 543 / 2249
Found 5 reads in subtree #10. Total saved so far: 549 / 2499
Found 16 reads in subtree #11. Total saved so far: 566 / 2749
Found 4 reads in subtree #12. Total saved so far: 571 / 2807
16:48:19 - Finished processing round 1: 571 reads, 20.33% of input
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00001.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00002.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00003.fa
Found 220 reads in subtree #1. Total saved so far: 221 / 249
Found 13 reads in subtree #2. Total saved so far: 235 / 499
Found 61 reads in subtree #3. Total saved so far: 297 / 570
16:48:30 - Finished processing round 2: 297 reads, 52.01% of input
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00001.fa
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00002.fa
Found 135 reads in subtree #1. Total saved so far: 136 / 249
Found 16 reads in subtree #2. Total saved so far: 153 / 296
16:48:38 - Finished processing round 3: 153 reads, 51.52% of input
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00001.fa
Found 61 reads in subtree #1. Total saved so far: 62 / 152
16:48:44 - Finished processing round 4: 62 reads, 40.52% of input
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00001.fa
Found 49 reads in subtree #1. Total saved so far: 50 / 61
16:48:46 - Finished processing round 5: 50 reads, 80.65% of input
Building NJ tree from /nethome/schrammca/cap256-week34H/work/lineage/NJ00001.fa
Found 48 reads in subtree #1. Total saved so far: 49 / 49
/nethome/schrammca/.local/lib/python3.5/site-packages/Bio/Seq.py:2309: BiopythonWarning: Partial codon, len(sequence) not a multiple of three. Explicitly trim the sequence or add trailing
After round #6: tree has converged with 49 reads!
Always sanity check your results: In this case, a quick scan of the CDR3 sequences shows that we have failed to pick up any lineage
members using CAP256-VRC26.01 and CAP256-VRC26.08 as our anchors in this analysis!
```

We can also do unseeded lineage analysis:

```
schrammca@cogsworth:~/nethome/schrammca/cap256-week34H$ sonar groups -v 'IGHV3-30*18' -j 'IGHJ3*01' -t 8

Fri Feb 22 09:25:26 2019 -- SONAR b'v2.0.0-133-g091cbcl-dirty' run with command:
/nethome/schrammca/sonar/lineage/2.4-cluster_into_groups.py -v 'IGHV3-30*18' -j 'IGHJ3*01' -t 8

Processing chunk #1...
Processing chunk #3...
Processing chunk #5...
Processing chunk #4...
Processing chunk #6...
Processing chunk #2...
minseqlength 15: 3 sequences discarded.
minseqlength 15: 1 sequence discarded.
minseqlength 15: 1 sequence discarded.
minseqlength 15: 1 sequence discarded.
```

Let's make this into a longitudinal analysis, by adding data from two other time points: 48 weeks after infection (SRR1057705) and 59 weeks after infection (SRR1057707). I've downloaded each of those to their own directories and processed them as above. We're going to use the islands identified by 2.1/2.1 in each case.

```
schrammca@cogsworth:~/nethome/schrammca/cap256-week34H$ sonar getfasta -l output/tables/islandSeqs.txt -f output/sequences/nucleotide/cap256-week34H_goodVJ_unique.fa -o output/sequence

Fri Feb 22 13:36:39 2019 -- SONAR b'v2.0.0-133-g091cbcl-dirty' run with command:
/nethome/schrammca/sonar/utilities/getFastaFromList.py -l output/tables/islandSeqs.txt -f output/sequences/nucleotide/cap256-week34H_goodVJ_unique.fa -o output/sequences/nucleotide/c

schrammca@cogsworth:~/nethome/schrammca/cap256-week34H$ cd ..
schrammca@cogsworth:~/nethome/schrammca$ mkdir cap256-longitudinal
schrammca@cogsworth:~/nethome/schrammca$ cd cap256-longitudinal/
schrammca@cogsworth:~/nethome/schrammca/cap256-longitudinal$ sonar merge_time --seqs ../cap256-week34H/output/sequences/nucleotide/cap256-week34H_islandSeqs.fa --labels w34 --seqs ../c

Tue Feb 19 17:15:23 2019 -- SONAR v2.0.0-133-g091cbcl-dirty run with command:
/nethome/schrammca/sonar/phylogeny/3.1-merge_timepoints.pl --seqs ../cap256-week34H/output/sequences/nucleotide/cap256-week34H_islandSeqs.fa --labels w34 --seqs ../cap256-week48H/out

/nethome/schrammca/sonar/third-party/vsearch -derep_fulllength work/phylo/all_seqs.fa -uc work/phylo/uc

vsearch v2.8.1_linux_x86_64, 31.1GB RAM, 8 cores
https://github.com/torognes/vsearch

Reading file work/phylo/all_seqs.fa 100%
128379 nt in 297 seqs, min 391, max 436, avg 432
Dereplicating 100%
Sorting 100%
297 unique sequences, avg cluster 1.0, median 1, max 1
Writing uc file, first part 100%
Writing uc file, second part 100%
```

Finally, let's construct a tree. For best results, I've added the VH3-30 germline sequence and our 12 native mAbs to the collected sequences from all 3 time points, aligned them, and manually validated the alignment. These steps can be automated using the -v and --natives flags to 3.2, but the automatic alignments get really messed up if there are a lot of insertions or deletions. This step takes several hours, given the amount of data (~300 sequences); use the --quick argument to skip slower optimizations.

```
schrammca@cogsworth:~/nethome/schrammca/cap256-longitudinal$ sonar igphyml -i output/sequences/nucleotide/cap256-longitudinal-collected_aligned.fa -f

Thu Feb 21 10:22:01 2019 -- SONAR b'v2.0.0-133-g091cbcl-dirty' run with command:
/nethome/schrammca/sonar/phylogeny/3.2-run_IgPhyML.py -i output/sequences/nucleotide/cap256-longitudinal-collected_aligned.fa -f

Looks like a fasta file..

.....
oooooooooooooooooooooo CURRENT SETTINGS ooooooooooooooooooooooo
.....

. Sequence filename:      infile
. Number of taxa:        310
. Sequence length:       150
. Data type:             Codon
. Alphabet size:         61
. Genetic code:          Standard
. Sequence format:       Interleaved
. Number of data sets:    1
```


. (2755 sec) [-23571.7441] [Topology] [depth= 1]
. (2812 sec) [-23571.3238] [Topology] [depth= 3]
. (2814 sec) [-23571.30331] [Topology] [depth= 1]
. (2824 sec) [-23569.4621] [Topology] [depth= 2]
. (2825 sec) [-23566.5676] [Topology] [depth= 1]
. (2839 sec) [-23566.5476] [Topology] [depth= 1]
. (2849 sec) [-23565.1503] [Topology] [depth= 1]
. (2851 sec) [-23565.0637] [Topology] [depth= 3]
. (2861 sec) [-23563.6548] [Topology] [depth= 1]
. (2888 sec) [-23562.6874] [Topology] [depth= 2]
. (2931 sec) [-23562.4116] [Topology] [depth= 1]
. (2941 sec) [-23561.5402] [Topology] [depth= 1]
. (2947 sec) [-23560.8791] [Topology] [depth= 2]
. (2950 sec) [-23559.9582] [Topology] [depth= 2]
. (2951 sec) [-23559.0408] [Topology] [depth= 1]
. (2952 sec) [-23556.4009] [Topology] [depth= 2]
. (2955 sec) [-23555.9560] [Topology] [depth= 1]
. (2957 sec) [-23553.9391] [Topology] [depth= 1]
. (2968 sec) [-23553.3608] [Topology] [depth= 1]
. (2984 sec) [-23553.2902] [Topology] [depth= 1]
. (3025 sec) [-23549.8803] [Topology] [depth= 1]
. (3025 sec) [-23541.6453] [Topology] [depth= 2]
. (3034 sec) [-23537.5951] [Topology] [depth= 1]
. (3035 sec) [-23535.8387] [Topology] [depth= 4]
. (3063 sec) [-23530.5067] [Topology] [depth= 1]
. (3065 sec) [-23529.6718] [Topology] [depth= 1]
. (3085 sec) [-23529.6357] [Topology] [depth= 2]
. (3087 sec) [-23529.0320] [Topology] [depth= 1]
. (3092 sec) [-23527.9303] [Topology] [depth= 1]
. (3121 sec) [-23527.6649] [Topology] [depth= 1]
. (3123 sec) [-23520.9261] [Topology] [depth= 1]
. (3145 sec) [-23518.7966] [Topology] [depth= 3]
. (3157 sec) [-23514.8861] [Topology] [depth= 4]
. (3191 sec) [-23511.1951] [Topology] [depth= 1]
. (3319 sec) [-23505.2378] [Topology] [depth= 4]
. (3320 sec) [-23501.8110] [Topology] [depth= 1]
. (3322 sec) [-23500.7986] [Topology] [depth= 1]
. (3332 sec) [-23500.7526] [Topology] [depth= 1]
. (3341 sec) [-23500.0348] [Topology] [depth= 1]
. (3443 sec) [-23499.8840] [Topology] [depth= 1]
. (3444 sec) [-23498.2899] [Topology] [depth= 2]
. (3481 sec) [-23490.6298] [Topology] [depth= 6]
. (3482 sec) [-23489.1381] [Topology] [depth= 1]
. (3533 sec) [-23488.8863] [Topology] [depth= 1]
. (3750 sec) [-23485.8338] [Topology] [depth= 1]
. (3763 sec) [-23479.3990] [Topology] [depth= 1]
. (3773 sec) [-23476.1941] [Topology] [depth= 1]
. (3853 sec) [-23475.6361] [Topology] [depth= 1]
. (3890 sec) [-23474.6127] [Topology] [depth= 2]
. (3893 sec) [-23474.5891] [Topology] [depth= 1]
. (3925 sec) [-23474.5578] [Topology] [depth= 1]
. (3932 sec) [-23474.3298] [Topology] [depth= 2]
. (3961 sec) [-23460.9169] [Topology] [depth= 3]
. (3998 sec) [-23460.8214] [Topology] [depth= 1]
. (4088 sec) [-23460.2845] [Topology] [depth= 1]
. (4122 sec) [-23458.9262] [Topology] [depth= 9]
. (4123 sec) [-23456.7143] [Topology] [depth= 1]
. (4218 sec) [-23455.2201] [Topology] [depth= 1]
. (4243 sec) [-23453.9452] [Topology] [depth= 6]
. (4289 sec) [-23453.9350] [Topology] [depth= 1]
. (4289 sec) [-23449.7930] [Topology] [depth= 1]
. (4291 sec) [-23447.8912] [Topology] [depth= 1]
. (4387 sec) [-23443.2407] [Topology] [depth= 4]
. (4433 sec) [-23443.1314] [Topology] [depth= 1]
. (4458 sec) [-23443.1044] [Topology] [depth= 1]
. (4476 sec) [-23443.0916] [Topology] [depth= 2]
. (4689 sec) [-23440.7088] [Topology] [depth= 1]
. (4891 sec) [-23440.6041] [Topology] [depth= 1]
. (4925 sec) [-23440.5917] [Topology] [depth= 1]
. (4976 sec) [-23438.2854] [Topology] [depth= 5]
. (4978 sec) [-23436.1784] [Topology] [depth= 5]
. (5158 sec) [-23435.3434] [Topology] [depth= 3]
. (5159 sec) [-23428.9329] [Topology] [depth= 2]
. (5161 sec) [-23425.8845] [Topology] [depth= 1]
. (5162 sec) [-23424.7007] [Topology] [depth= 2]
. (5181 sec) [-23421.5607] [Topology] [depth= 2]
. (5268 sec) [-23399.2787] [Branch lengths] [
. (5269 sec) [-23398.5293] [ts/tv ratio] [1.85]
. (5269 sec) [-23398.5293] [dn/ds ratio] [0.73]
. (5545 sec) [-23396.0457] [Topology] [depth= 1]
. (5629 sec) [-23392.0615] [Topology] [depth= 2]
. (5630 sec) [-23381.9841] [Topology] [depth= 1]
. (5676 sec) [-23380.9014] [Topology] [depth= 5]
. (5704 sec) [-23379.2825] [Topology] [depth= 1]
. (5739 sec) [-23376.9250] [Topology] [depth= 1]
. (5918 sec) [-23375.8368] [Topology] [depth= 2]
. (5919 sec) [-23371.7194] [Topology] [depth= 1]
. (5926 sec) [-23368.8729] [Topology] [depth= 1]
. (6020 sec) [-23368.7517] [Topology] [depth= 1]
. (6140 sec) [-23367.7729] [Topology] [depth= 1]
. (6146 sec) [-23367.6458] [Topology] [depth= 1]
. (6211 sec) [-23367.5722] [Topology] [depth= 1]
. (6600 sec) [-23366.3941] [Topology] [depth= 3]
. (6649 sec) [-23366.3728] [Topology] [depth= 1]
. (6679 sec) [-23366.3390] [Topology] [depth= 1]
. (6720 sec) [-23366.3229] [Topology] [depth= 1]
. (6775 sec) [-23364.7175] [Topology] [depth= 1]
. (6840 sec) [-23363.8130] [Topology] [depth= 3]
. (6843 sec) [-23363.7865] [Topology] [depth= 1]
. (6892 sec) [-23360.2146] [Topology] [depth= 3]
. (6893 sec) [-23360.2034] [Topology] [depth= 1]
. (6901 sec) [-23359.6265] [Topology] [depth= 4]
. (6933 sec) [-23359.6055] [Topology] [depth= 1]
. (6935 sec) [-23356.6344] [Topology] [depth= 1]
. (6939 sec) [-23356.0085] [Topology] [depth= 1]
. (6963 sec) [-23354.6301] [Topology] [depth= 2]
. (6974 sec) [-23354.6128] [Topology] [depth= 1]
. (6975 sec) [-23352.7784] [Topology] [depth= 1]
. (7016 sec) [-23352.7080] [Topology] [depth= 1]
. (7017 sec) [-23352.4778] [Topology] [depth= 1]
. (7089 sec) [-23351.4518] [Topology] [depth= 4]
. (7108 sec) [-23349.4092] [Topology] [depth= 2]
. (7153 sec) [-23349.2799] [Topology] [depth= 2]
. (7184 sec) [-23348.6036] [Topology] [depth= 1]
. (7281 sec) [-23348.5844] [Topology] [depth= 1]
. (7540 sec) [-23348.3061] [Topology] [depth= 5]
. (7721 sec) [-23343.1643] [Topology] [depth= 2]
. (7734 sec) [-23340.1753] [Topology] [depth= 2]
. (7782 sec) [-23334.2233] [Topology] [depth= 8]
. (7828 sec) [-23334.1388] [Topology] [depth= 1]
. (7947 sec) [-23333.1277] [Topology] [depth= 2]
. (8051 sec) [-23332.9740] [Topology] [depth= 1]
. (8137 sec) [-23331.5764] [Topology] [depth= 2]
. (8409 sec) [-23331.5567] [Topology] [depth= 1]
. (8677 sec) [-23326.7949] [Branch lengths] [
. (8678 sec) [-23326.7819] [ts/tv ratio] [1.85]
. (8678 sec) [-23326.7819] [dn/ds ratio] [0.73]
. (10804 sec) [-23326.7018] [Topology] [depth= 2]
. (11676 sec) [-23325.2561] [Topology] [depth= 3]
. (11677 sec) [-23323.3582] [Topology] [depth= 4]
. (11734 sec) [-23323.3446] [Topology] [depth= 1]
. (11750 sec) [-23321.1661] [Topology] [depth= 2]
. (11768 sec) [-23320.2728] [Topology] [depth= 1]
. (11980 sec) [-23320.2546] [Topology] [depth= 1]
. (12234 sec) [-23320.1525] [Topology] [depth= 2]
. (17582 sec) [-23320.1317] [Topology] [depth= 1]
. (17852 sec) [-23320.0502] [Branch lengths] [
. (17854 sec) [-23320.0455] [ts/tv ratio] [1.85]
. (17854 sec) [-23320.0455] [dn/ds ratio] [0.73]

```
. (24488 sec) [ -23320.0444] [Branch lengths ]
. (24488 sec) [ -23320.0444] [Topology      ]
. (24506 sec) [ -23320.0436] [Topology      ]
. (24514 sec) [ -23320.0436] [ts/tv ratio   ] [1.85 ]
. (24514 sec) [ -23320.0436] [dn/ds ratio   ] [0.73 ]

. Checking for NNIs, optimizing five branches...

. Log-likelihood of the current tree: -23320.04
Ratio test? 0

. Printing the most likely tree in file:
.... 'infile_igphyml_tree.txt_gy94'

. Time used 6h50m12s

oooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooo

. Loading hotspot tables..
/home/schrammca/igphyml/IgPhyML/src/motifs/HTABLE_WRC_2
/home/schrammca/igphyml/IgPhyML/src/motifs/HTABLE_GYW_0
/home/schrammca/igphyml/IgPhyML/src/motifs/HTABLE_WA_1
/home/schrammca/igphyml/IgPhyML/src/motifs/HTABLE_TW_0
/home/schrammca/igphyml/IgPhyML/src/motifs/HTABLE_SYC_2
/home/schrammca/igphyml/IgPhyML/src/motifs/HTABLE_GRS_0

Looks like a fasta file..

***** rest of output truncated *****
```

Use `sonar tree` to create a figure-ready plot of the tree!