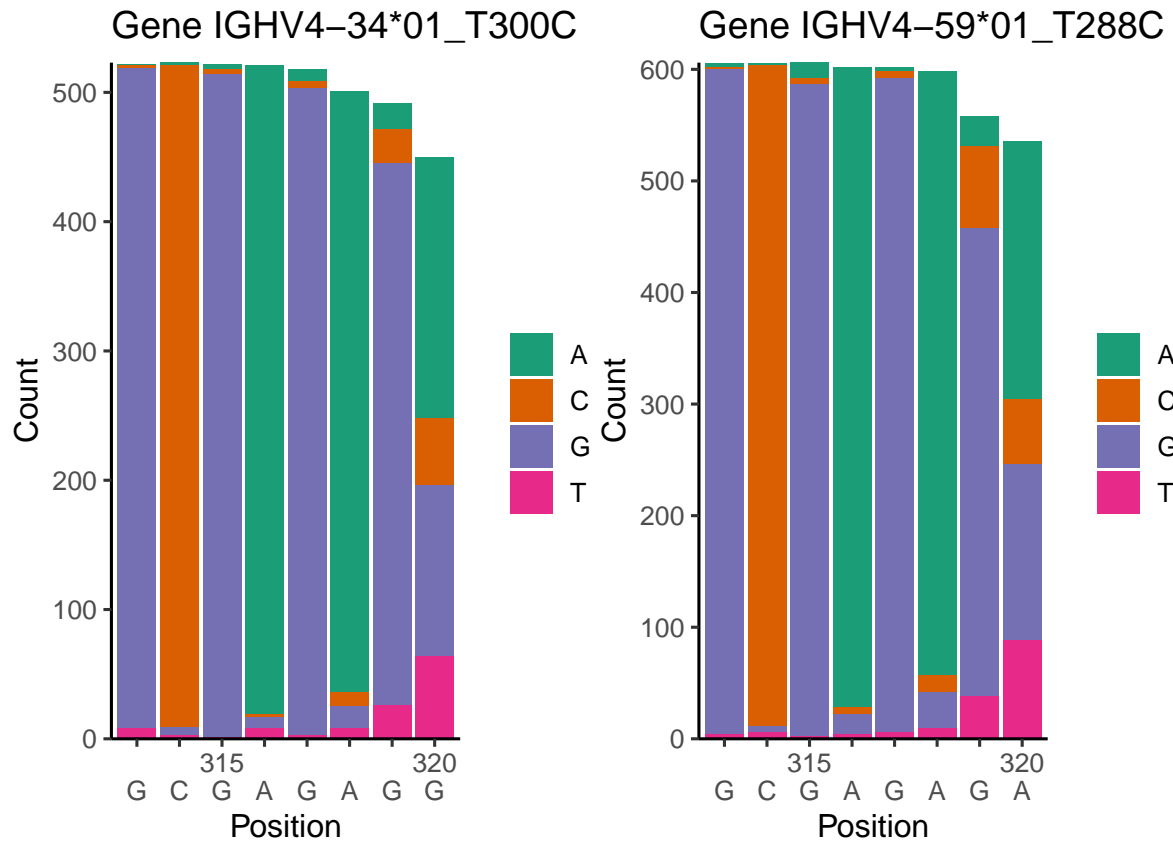# OGRDBstats Report

Contents

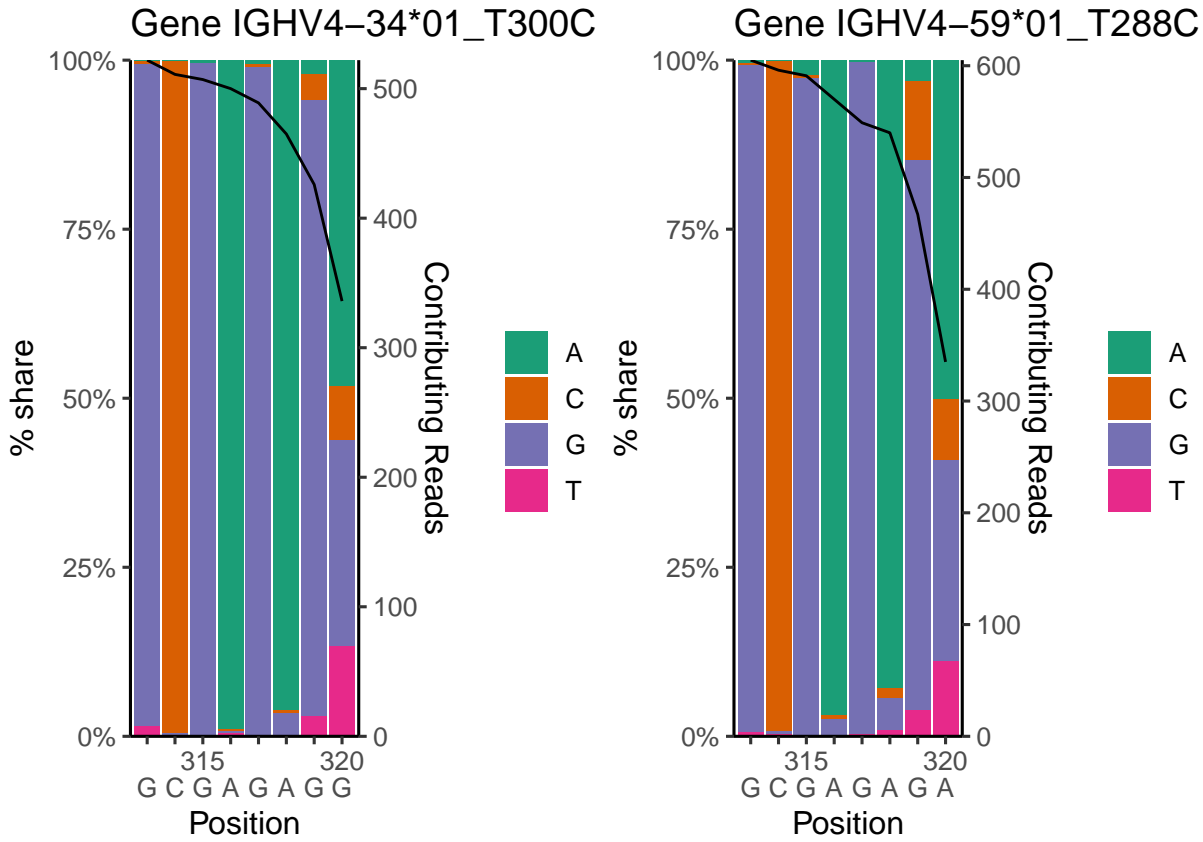# 1 Novel sequence analysis

1.1 End-nucleotide composition

1.2 Per-nucleotide consensus where previous nucleotides match the consensus
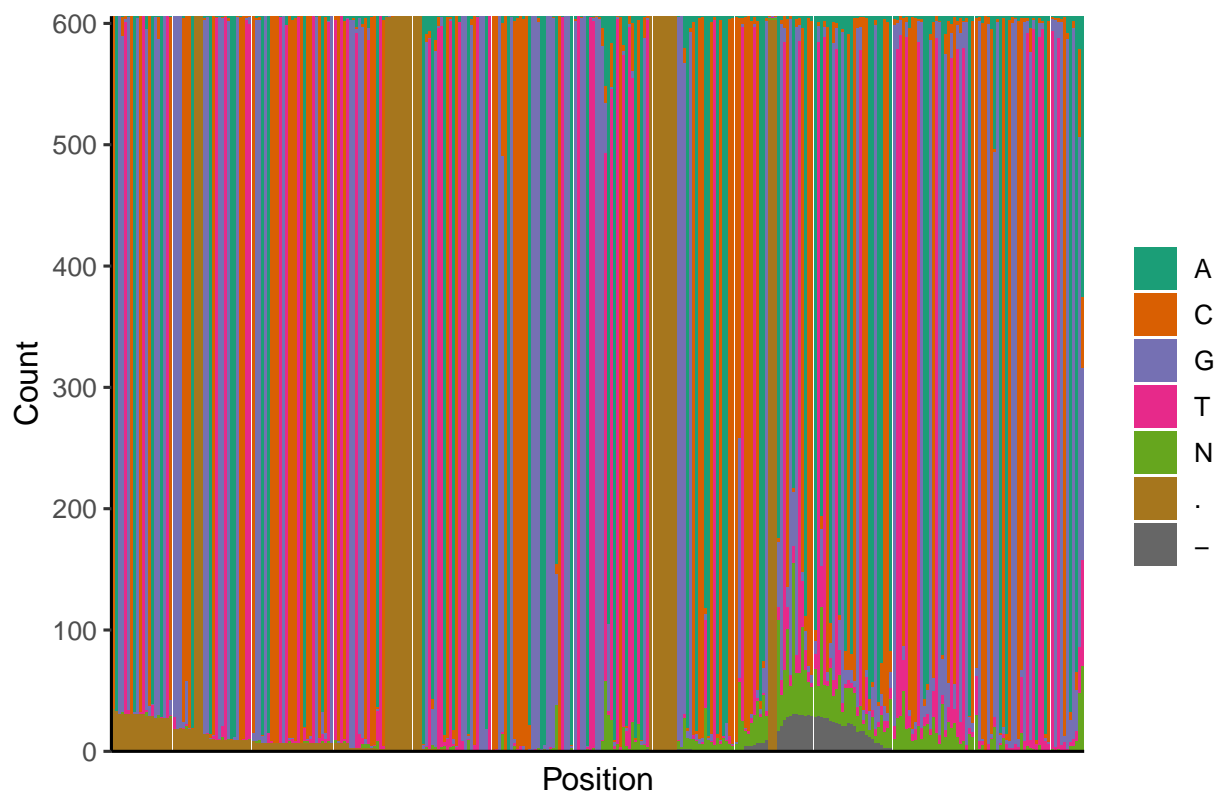

Gene IGHV4−34*01_T300C

Gene IGHV4−59*01_T288C

1.3 Whole-sequence composition of each assigned read

Gene IGHV4–34*01_T300C
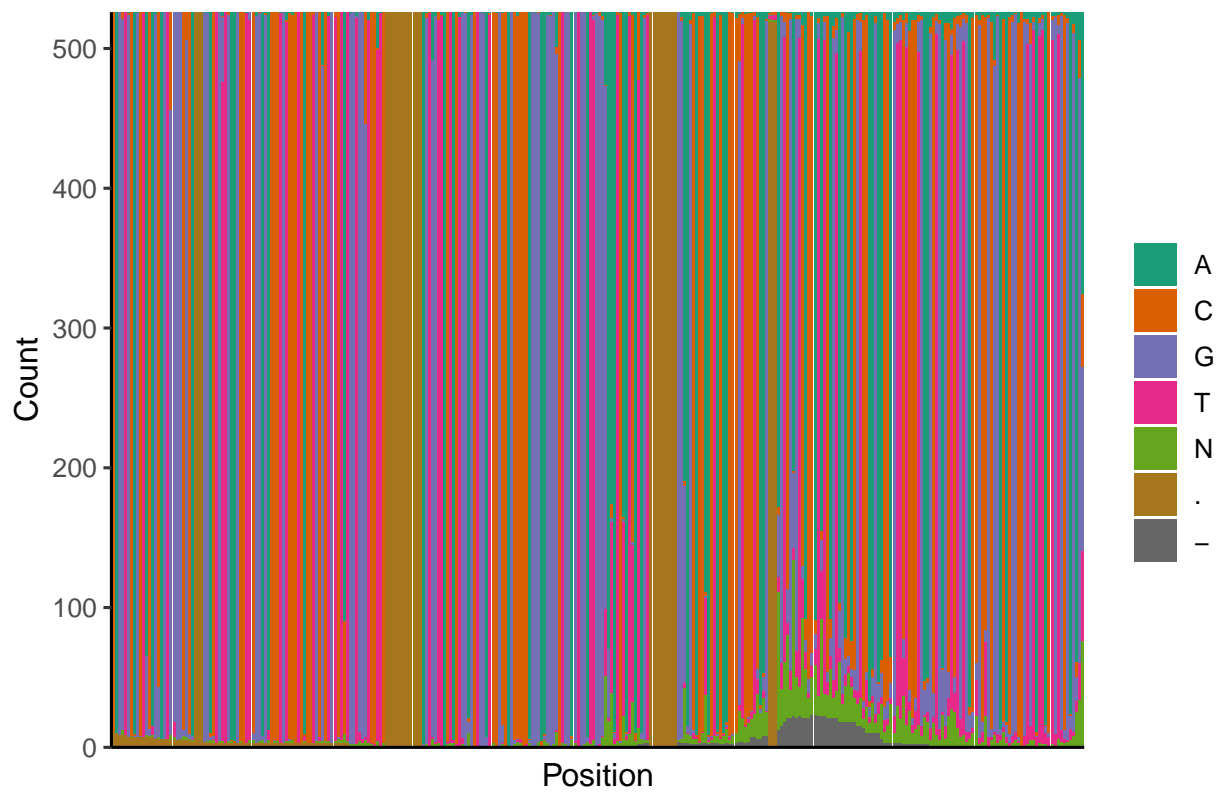
Gene IGHV4–59*01_T288C

1.4 Final three nucleotides: frequency of each observed triplet

IGHV4−34*01_T300C− Final 3 nucleotides as a triplet

IGHV4−59*01_T288C− Fina
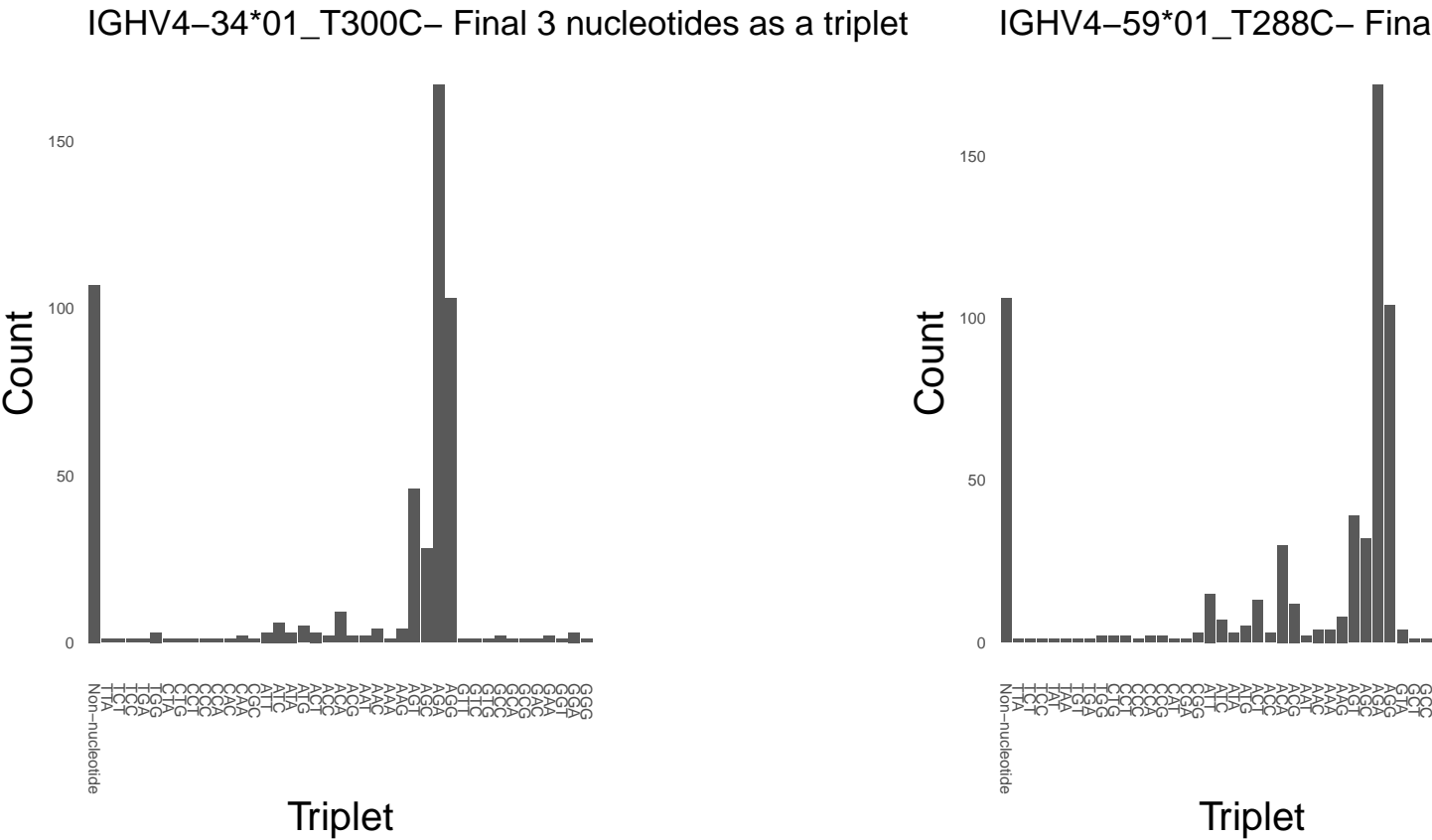
1.5 CDR3 length distribution, in assignments to novel alleles

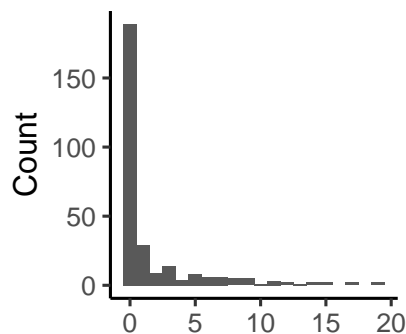### IGHV4−34*01_T300C

IGHV4−59*01_T288C
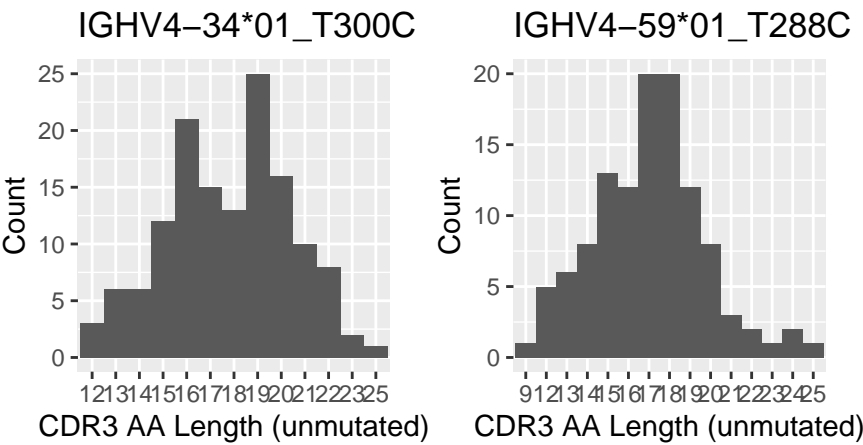
2 Variation from germline, in assignments to each allele

IGHV1−2*02

798 sequences assigned
154 (19.3%) exact matches, in which:
151 unique CDR3
8 unique J

IGHV1−3*01

375 sequences assigned
100 (26.7%) exact matches, in which:
82 unique CDR3
7 unique J

IGHV1−8*01

469 sequences assigned
78 (16.6%) exact matches, in which:
78 unique CDR3
8 unique J

IGHV1−18*01

1569 sequences assigned
400 (25.5%) exact matches, in which:
395 unique CDR3
9 unique J

IGHV1−24*01

403 sequences assigned
111 (27.5%) exact matches, in which:
107 unique CDR3
8 unique J

IGHV1−46*01

520 sequences assigned
100 (19.2%) exact matches, in which:
95 unique CDR3
7 unique J

IGHV1−58*01

24 sequences assigned
7 (29.2%) exact matches, in which:
7 unique CDR3
3 unique J

IGHV1−69*04

439 sequences assigned
141 (32.1%) exact matches, in which:
139 unique CDR3
9 unique J

IGHV1−69*06

1219 sequences assigned
459 (37.7%) exact matches, in which:
455 unique CDR3
11 unique J

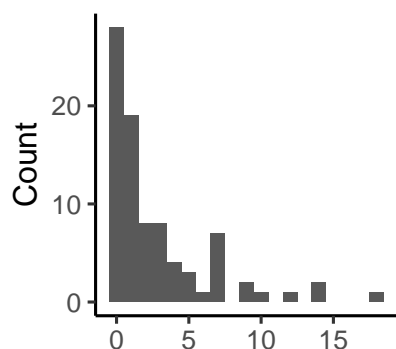## IGHV2−5*02

324 sequences assigned
189 (58.3%) exact matches, in which:
185 unique CDR3
8 unique J

IGHV2−10*01

1 sequences assigned
1 (100%) exact matches, in which:
1 unique CDR3
1 unique J

IGHV2−26*01

102 sequences assigned
57 (55.9%) exact matches, in which:
55 unique CDR3
6 unique J

IGHV2−70*01

97 sequences assigned
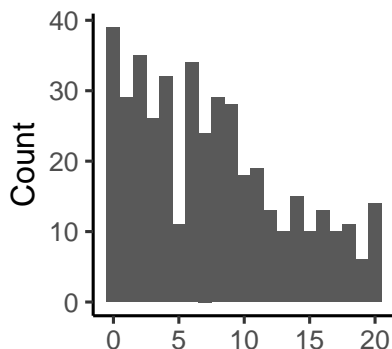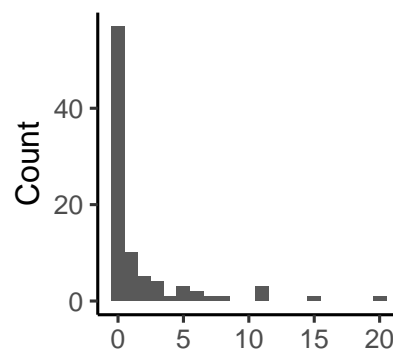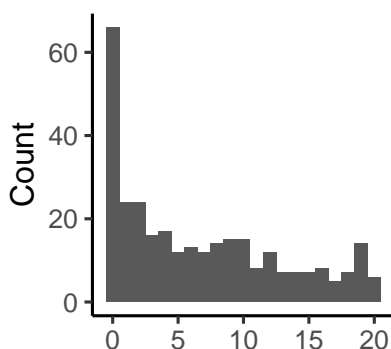28 (28.9%) exact matches, in which:
27 unique CDR3
4 unique J

IGHV3−7*01

556 sequences assigned
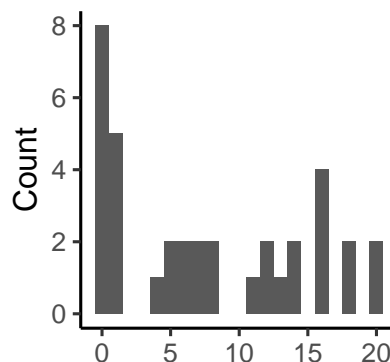39 (7%) exact matches, in which:
39 unique CDR3
7 unique J

IGHV3−9*01

352 sequences assigned
66 (18.8%) exact matches, in which:
66 unique CDR3
5 unique J

IGHV3−11*01

153 sequences assigned
26 (17%) exact matches, in which:
26 unique CDR3
6 unique J

IGHV3−11*05

140 sequences assigned
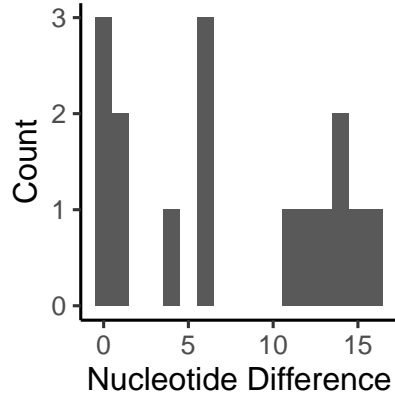23 (16.4%) exact matches, in which:
23 unique CDR3
5 unique J

IGHV3−13*01

37 sequences assigned
8 (21.6%) exact matches, in which:
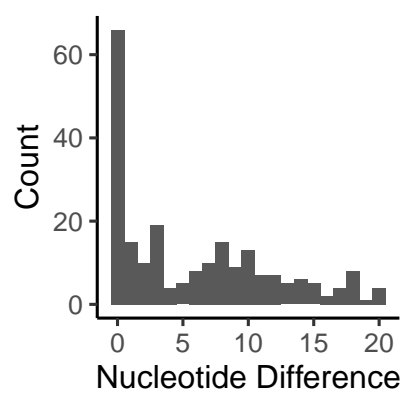8 unique CDR3
3 unique J

IGHV3−13*04

17 sequences assigned
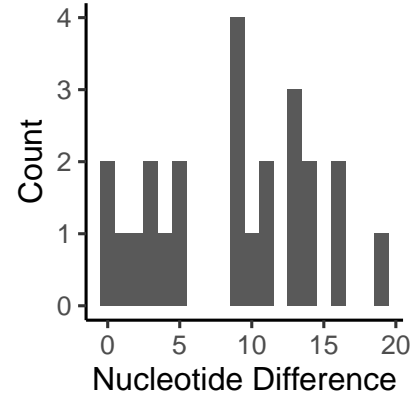3 (17.6%) exact matches, in which:
3 unique CDR3
2 unique J

IGHV3−15*01

254 sequences assigned
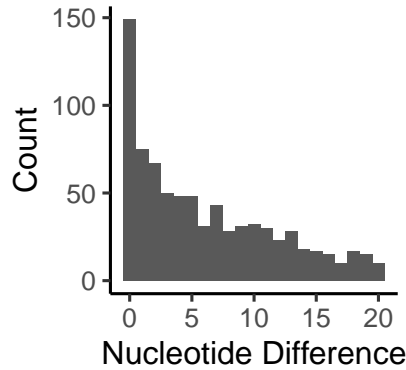66 (26%) exact matches, in which:
65 unique CDR3
6 unique J

IGHV3−20*01

29 sequences assigned
2 (6.9%) exact matches, in which:
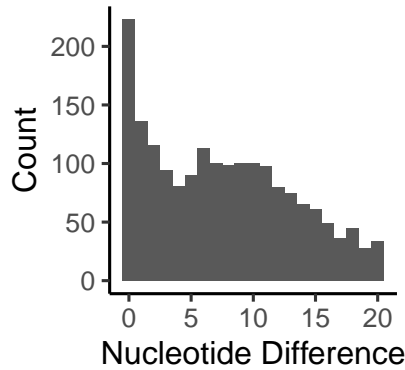2 unique CDR3
2 unique J

IGHV3−21*01

861 sequences assigned
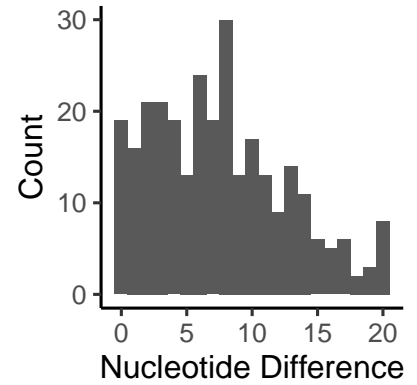149 (17.3%) exact matches, in which:
148 unique CDR3
9 unique J

IGHV3−23*01

2110 sequences assigned
223 (10.6%) exact matches, in which:
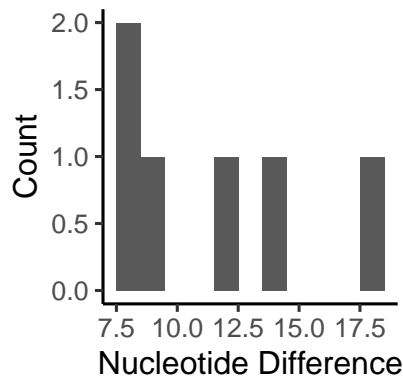221 unique CDR3
8 unique J

IGHV3−30*03

316 sequences assigned
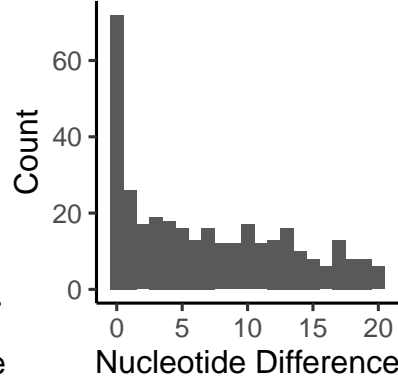19 (6%) exact matches, in which:
19 unique CDR3
4 unique J

IGHV3−30−3*03
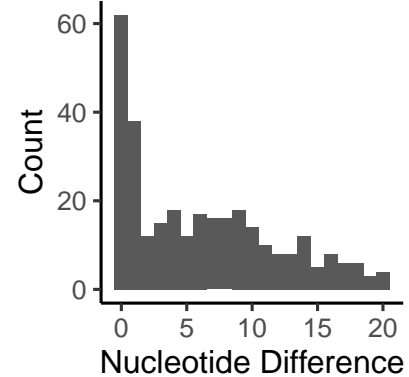
10 sequences assigned
No exact matches.

IGHV3−30*04

376 sequences assigned
72 (19.1%) exact matches, in which:
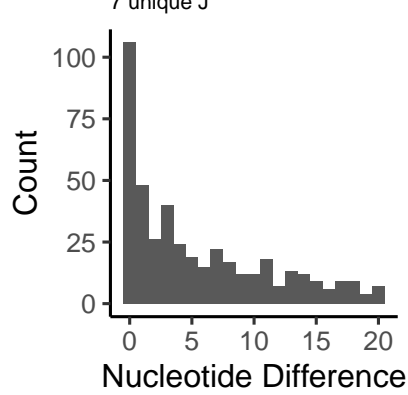70 unique CDR3
8 unique J

IGHV3−30*18

353 sequences assigned
62 (17.6%) exact matches, in which:
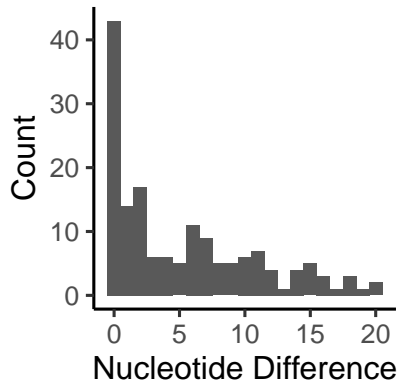61 unique CDR3
8 unique J

IGHV3−33*01

497 sequences assigned
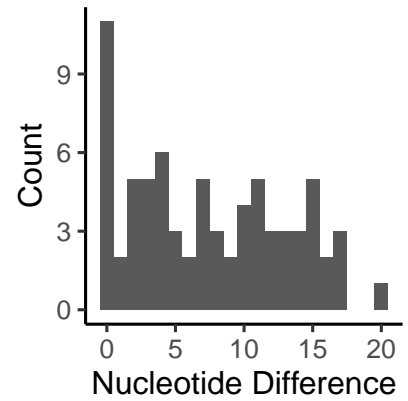106 (21.3%) exact matches, in which:
105 unique CDR3
7 unique J

IGHV3−33*06

169 sequences assigned
43 (25.4%) exact matches, in which:
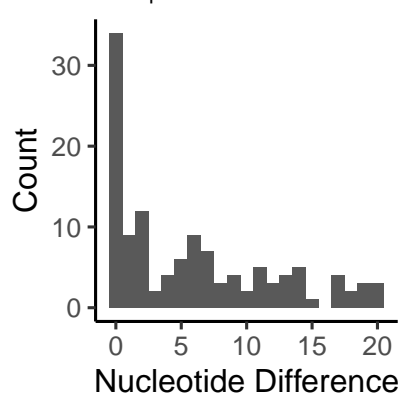42 unique CDR3
7 unique J

IGHV3−43*01

86 sequences assigned
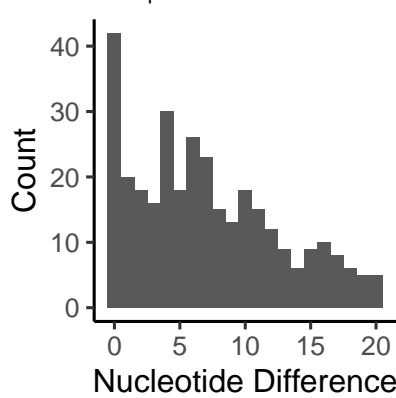11 (12.8%) exact matches, in which:
11 unique CDR3
3 unique J

IGHV3−48*02

130 sequences assigned
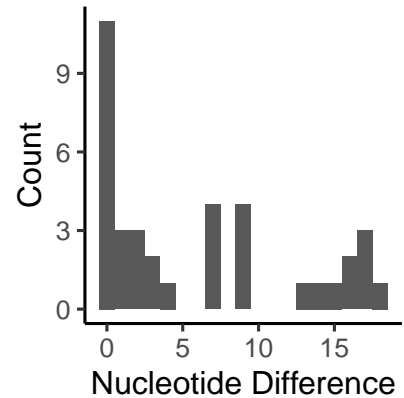34 (26.2%) exact matches, in which:
34 unique CDR3
4 unique J

IGHV3−48*04

358 sequences assigned
42 (11.7%) exact matches, in which:
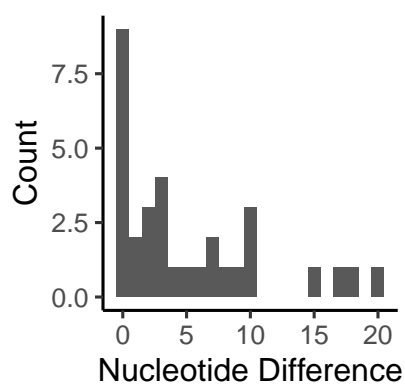42 unique CDR3
9 unique J

IGHV3−49*03

45 sequences assigned
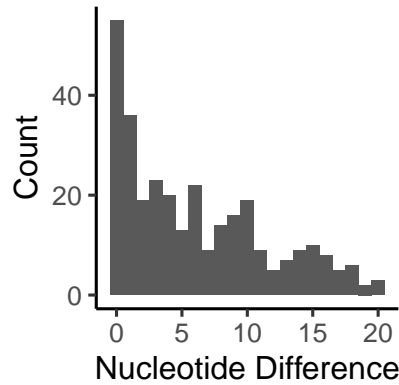11 (24.4%) exact matches, in which:
11 unique CDR3
4 unique J

IGHV3−49*05

36 sequences assigned
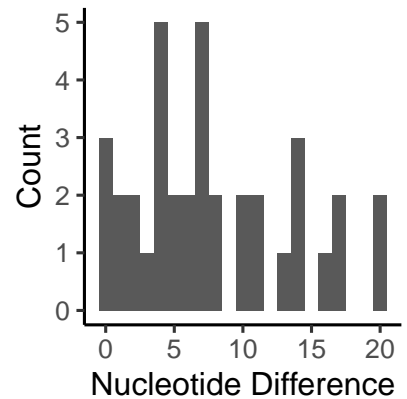9 (25%) exact matches, in which:
9 unique CDR3
4 unique J

IGHV3−53*01

352 sequences assigned
55 (15.6%) exact matches, in which:
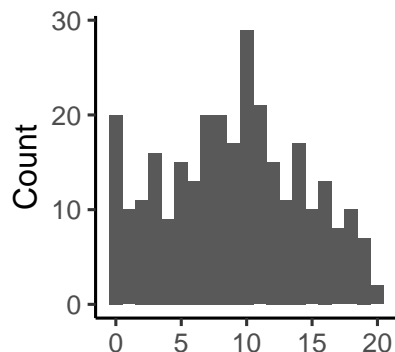55 unique CDR3
7 unique J

IGHV3−64*01

54 sequences assigned
3 (5.6%) exact matches, in which:
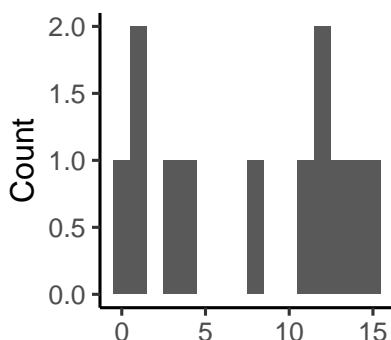3 unique CDR3
1 unique J

IGHV3−66*01

338 sequences assigned
20 (5.9%) exact matches, in which:
20 unique CDR3
6 unique J

IGHV3−72*01
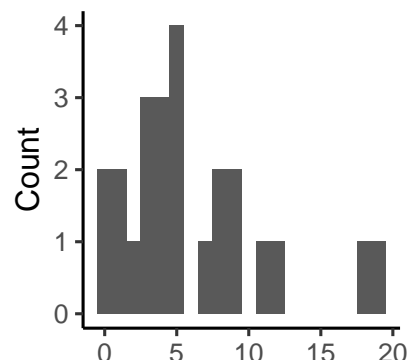
14 sequences assigned
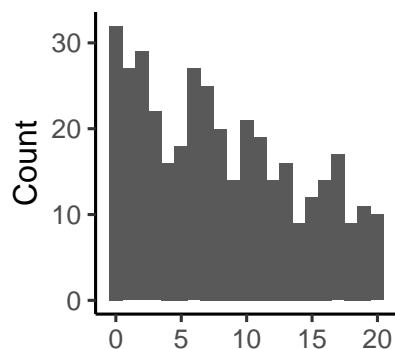1 (7.1%) exact matches, in which:
1 unique CDR3
0 unique J

IGHV3−73*01

31 sequences assigned
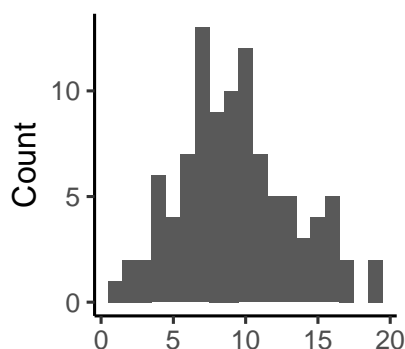2 (6.5%) exact matches, in which:
2 unique CDR3
2 unique J

IGHV3−74*01

482 sequences assigned
32 (6.6%) exact matches, in which:
30 unique CDR3
6 unique J

IGHV3−NL1*01
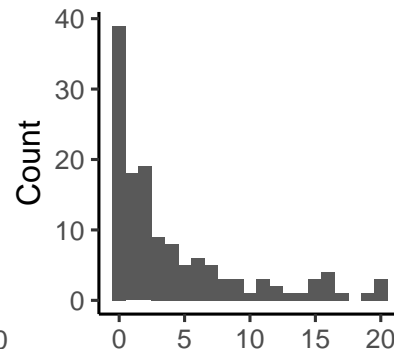
118 sequences assigned
No exact matches.

IGHV4−4*07
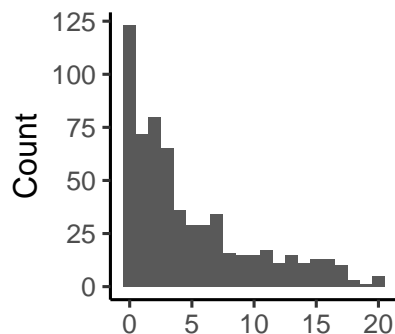
163 sequences assigned
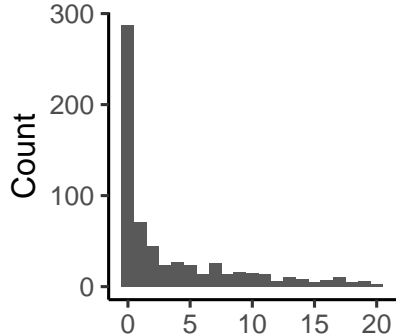39 (23.9%) exact matches, in which:
35 unique CDR3
7 unique J

IGHV4−31*03

694 sequences assigned
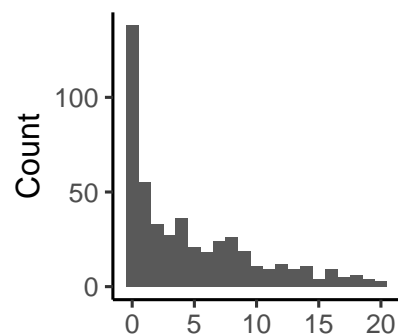123 (17.7%) exact matches, in which:
121 unique CDR3
9 unique J

IGHV4−34*01

692 sequences assigned
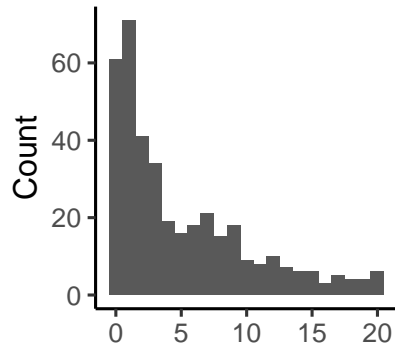287 (41.5%) exact matches, in which:
282 unique CDR3
10 unique J

IGHV4−34*01_T300C

526 sequences assigned
138 (26.2%) exact matches, in which:
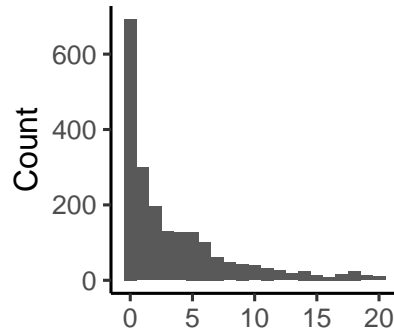138 unique CDR3
8 unique J

10

IGHV4−38−2*02

427 sequences assigned
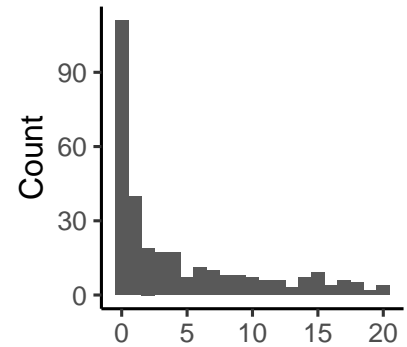61 (14.3%) exact matches, in which:
55 unique CDR3
7 unique J

IGHV4−39*07

2229 sequences assigned
693 (31.1%) exact matches, in which:
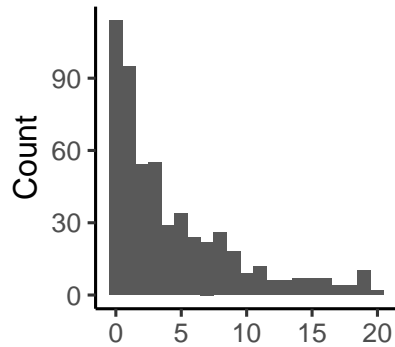691 unique CDR3
11 unique J

IGHV4−59*01

340 sequences assigned
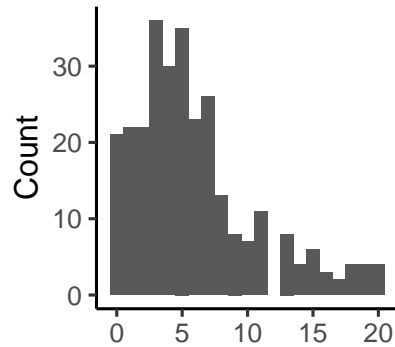111 (32.6%) exact matches, in which:
108 unique CDR3
8 unique J

IGHV4−59*01_T288C

606 sequences assigned
114 (18.8%) exact matches, in which:
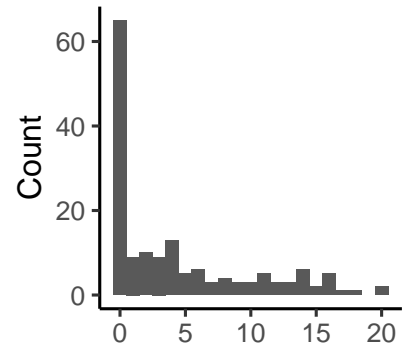114 unique CDR3
10 unique J

IGHV4−61*01

322 sequences assigned
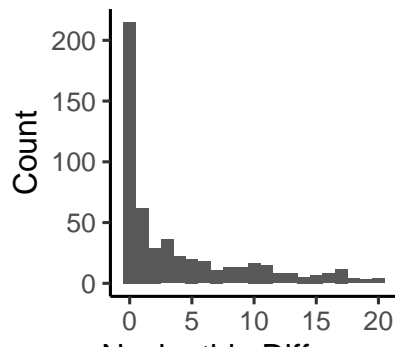21 (6.5%) exact matches, in which:
21 unique CDR3
4 unique J

IGHV5−10−1*01

171 sequences assigned
65 (38%) exact matches, in which:
61 unique CDR3
9 unique J

IGHV5−51*01
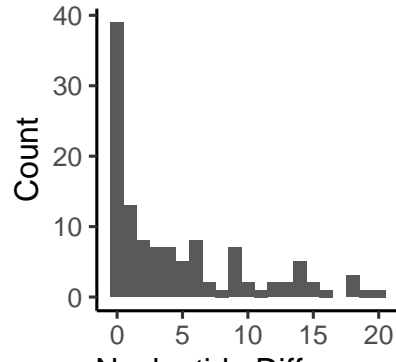
574 sequences assigned
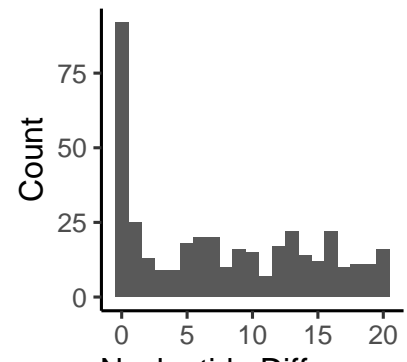215 (37.5%) exact matches, in which:
209 unique CDR3
10 unique J

IGHV6−1*01

124 sequences assigned
39 (31.5%) exact matches, in which:
39 unique CDR3
7 unique J

IGHV7−4−1*02

540 sequences assigned
92 (17%) exact matches, in which:
91 unique CDR3
6 unique J

3 Allele usage in potential haplotype anchor genes

4 Configuration settings

Repertoire file: TWO01A_naive_genotyped.tsv

Germline reference file: IMGT_REF_GAPPED.fasta

Novel allele file: TWO01A_naive_novel_ungapped.fasta

Species: Homosapiens

Chain: IGHV

Segment: V

Inferred gene IGHV4 34 01_T300C gapped using IGHV4 34 01: CAGGTGCAGCTACAGCAGTGGGGCGCA. . . GGACTGTTC

Inferred gene IGHV4 59 01_T288C gapped using IGHV4 59 01: CAGGTGCAGCTGCAGGAGTCGGGC-CCA. . . GGACTGGTGAAGCCTTCGGAGACCCTGTCCCTCACCTGCACTGTCTCTGGTGGCTCCATC. . . . . AGT

Warning: IGHV3 30 3 03, IGHV3 30 04 have identical germline sequences.

Haplotyping analysis is based on gene IGHJ6 alleles 02:03