Figure S1

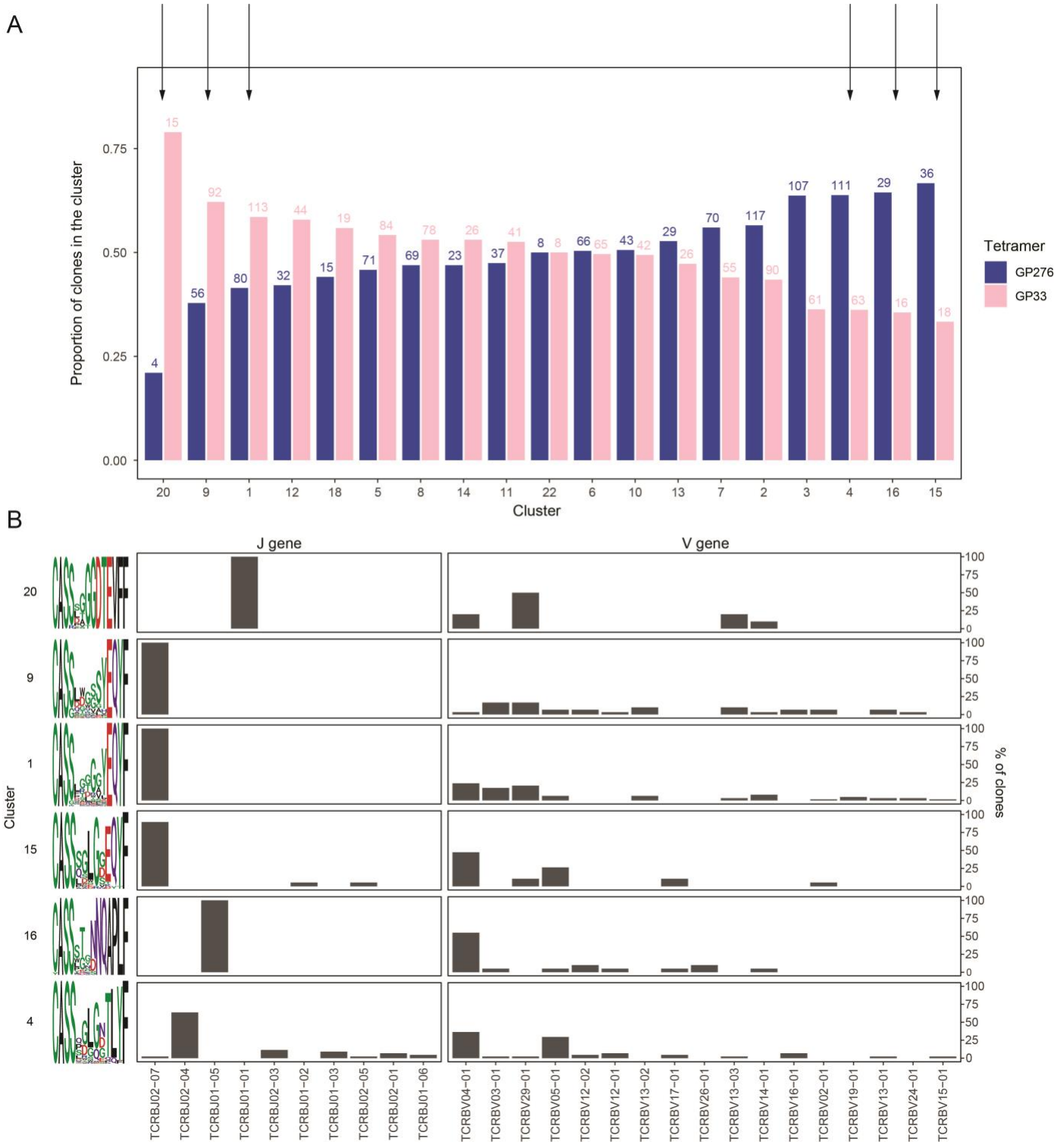

A
Proportion of clones in the cluster
0.00
0.25
0.50
0.75
20 9 1 12 18 5 8 14 11 22 6 10 13 7 2 3 4 16 15
Cluster
4 56 80 32 15 71 69 23 37 8 66 43 29 70 117 107 111 29 36
15 92 113 44 19 84 78 26 41 8 65 42 26 55 90 61 63 16 18
Tetramer
GP276
GP33
B
J gene
V gene
Cluster
20
9
1
15
16
4
% of clones
0 25 50 75 100
TCRBJ02−07
TCRBJ02−04
TCRBJ01−05
TCRBJ01−01
TCRBJ02−03
TCRBJ01−02
TCRBJ01−03
TCRBJ02−05
TCRBJ02−01
TCRBJ01−06
TCRBV04−01
TCRBV03−01
TCRBV29−01
TCRBV05−01
TCRBV12−02
TCRBV12−01
TCRBV13−02
TCRBV17−01
TCRBV26−01
TCRBV13−03
TCRBV14−01
TCRBV16−01
TCRBV02−01
TCRBV19−01
TCRBV13−01
TCRBV24−01
TCRBV15−01

Figure S1. GP33 and GP276 tetramer reactive cells cluster differently. (A) CDR3s from LCMV clone 13 GP33-reactive and GP276-reactive CD8 T cells were clustered. For each cluster, the proportion of sequences from GP33- and GP276-reactive cells is shown, numbers above the bars depict the absolute number of clones. The arrows represent the top three communities significantly enriched for GP33- (left) and GP276-reactive CDR3s. **(B)** Motif analysis and V/J gene usage for the six clusters identified above.

Figure S2

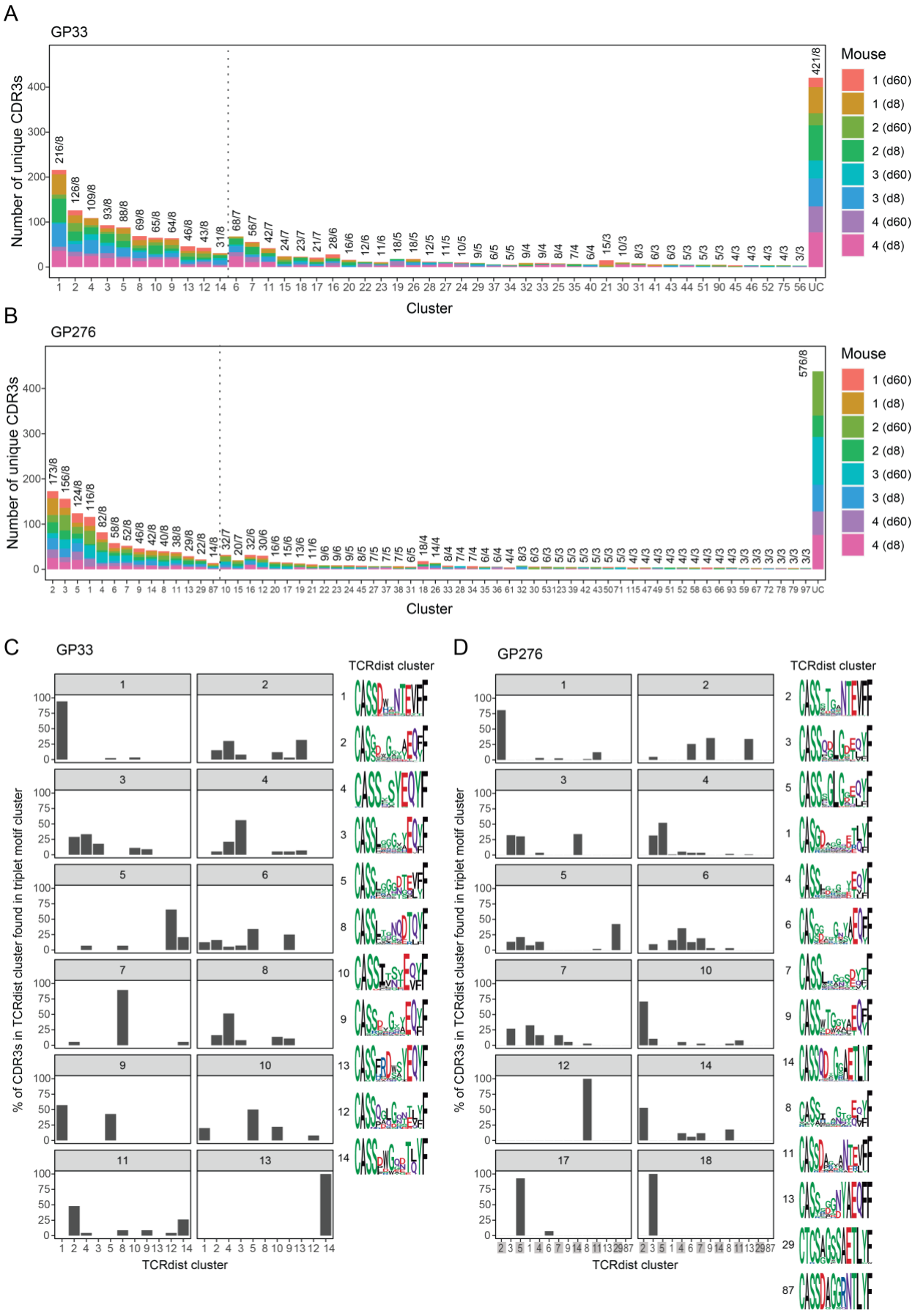

A
GP33
Number of unique CDR3s
Cluster
Mouse
1 (d60)
1 (d8)
2 (d60)
2 (d8)
3 (d60)
3 (d8)
4 (d60)
4 (d8)
B
GP276
Number of unique CDR3s
Cluster
Mouse
C
GP33
TCRdist cluster
% of CDR3s in TCRdist cluster found in triplet motif cluster
TCRdist cluster
D
GP276
TCRdist cluster
% of CDR3s in TCRdist cluster found in triplet motif cluster
TCRdist cluster

Figure S2. Overlap between triplet motif and TCRdist identified clusters. GP33 **(A)** and GP276 **(B)** CDR3 sequences were clustered based on similarity scores calculated using the TCRdist package. Numbers above bars indicate the number of clones in each cluster, followed by the number of mice the cluster was found in. Clones to the left of the vertical dotted line were ubiquitously found amongst all eight mice. Overlap between triplet motif (TM) and TCRdist clusters for the GP33 **(C)** and GP276 repertoires **(D)** is shown, along with amino acid sequence motifs of the TCRdist clusters.

Figure S3


A
GP33
134/8
117/8
69/8
64/8
49/8
49/8
47/8
41/8
37/8
35/8
32/8
16/8
22/7
15/7
14/6
12/6
11/6
10/6
14/5
10/5
7/5
6/5
6/5
12/4
8/4
7/4
5/4
5/4
5/4
4/4
4/4
5/3
4/3
4/3
4/3
4/3
3/3
3/3
3/3
3/3
3/3
3/3
3/3
4/2
4/2
3/2
3/2
3/2
3/2
3/2
3/2
3/2
3/2
3/2
3/2
3/2
2/2
2/2
2

Figure S3. Ubiquitous clusters comprise most of the repertoire. Barplots depict number of unique CDR3s in each cluster (data from eight mice), amongst GP33-reactive **(A)** and GP276-reactive **(B)** CD8 T cells from liver and spleen of LCMV clone 13 infected mice (d8 and d60 data are pooled). Bar segments show the contribution of individual mice. Numbers above bars depict the number of unique CDR3s in the cluster and the number of mice within which the cluster is found. Clusters above the dotted line are shared by all eight mice. The percentage of the repertoire composed of clones that belong to a ubiquitous cluster ("clustered"), mouse specific clusters ("Sp"; found in seven or fewer mice) and unclustered clones ("UC") at d8 and d60 for GP33-reactive **(C)** and GP276-reactive **(D)** $CD8^+$ T cells.

Figure S4

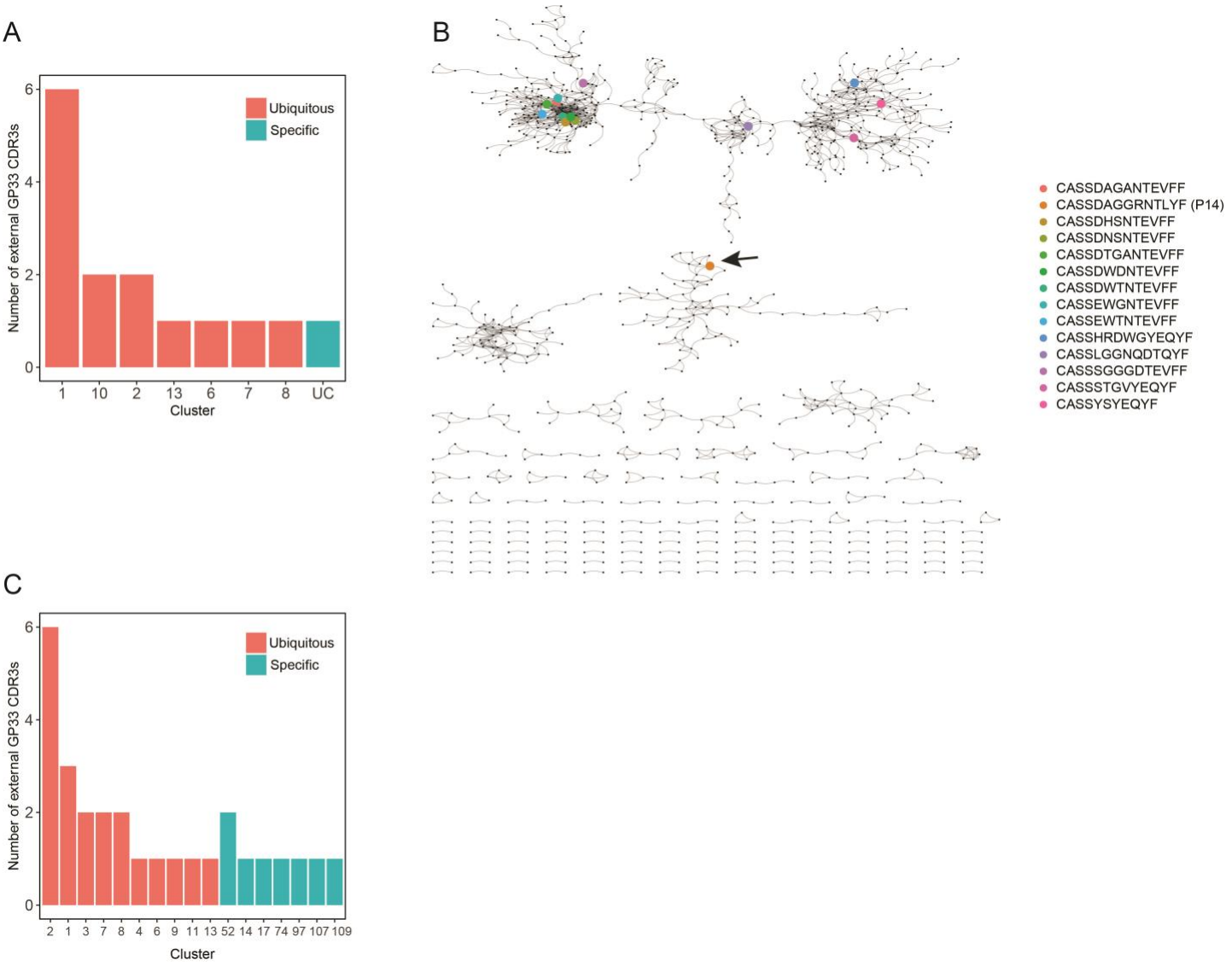


Figure S4. GP33 CDR3s from other datasets commonly map to ubiquitous clusters. Publicly available GP33-reactive CDR3 sequences were identified (n=123), of which 15 were found in the Chang et al. dataset. **(A-B)** Cluster distribution **(A)** and network plot localization **(B)** are shown. **(C)** Cluster distribution of the remaining 108 CDR3 sequences that are not found in the Chang et al. dataset. External sequences were mapped to clusters based on sequence similarity.

Figure S5

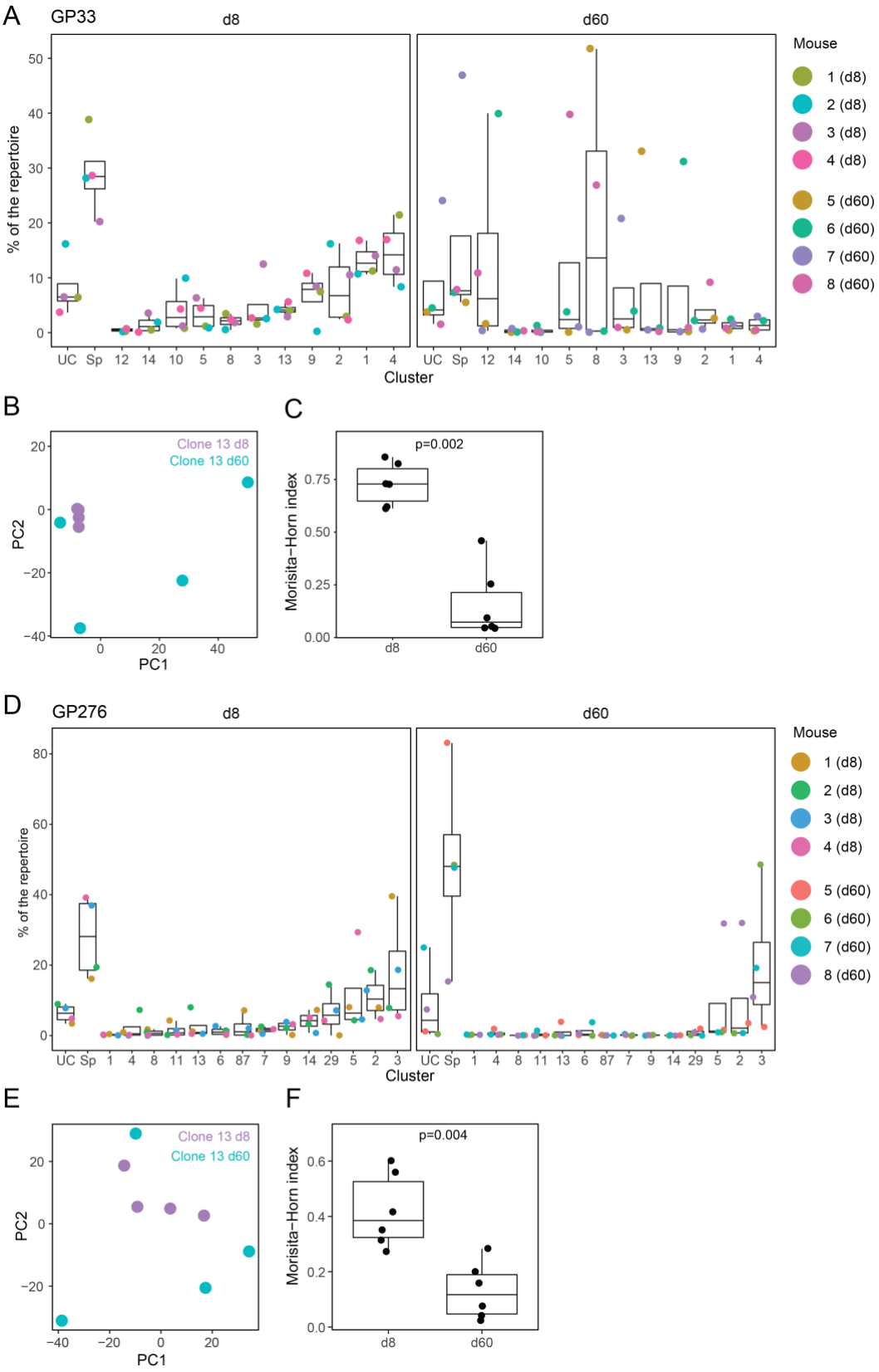

A
GP33
d8
d60
% of the repertoire
Cluster
UC Sp 12 14 10 5 8 3 13 9 2 1 4
Mouse
1 (d8)
2 (d8)
3 (d8)
4 (d8)
5 (d60)
6 (d60)
7 (d60)
8 (d60)
B
Clone 13 d8
Clone 13 d60
PC1
PC2
C
p=0.002
Morisita−Horn index
d8
d60
D
GP276
d8
d60
% of the repertoire
Cluster
UC Sp 1 4 8 11 13 6 87 7 9 14 29 5 2 3
Mouse
1 (d8)
2 (d8)
3 (d8)
4 (d8)
5 (d60)
6 (d60)
7 (d60)
8 (d60)
E
Clone 13 d8
Clone 13 d60
PC1
PC2
F
p=0.004
Morisita−Horn index
d8
d60

Figure S5. Repertoire evolution amongst GP276 and GP33 reactive CD8 T cells. (A) Percentage of the GP33 reactive CD8 T cell repertoire occupied by each ubiquitous TCRdist cluster at d8 and d60. Each point is one mouse. The proportion of the repertoire occupied by CDR3βs belonging to non-ubiquitous clusters (mouse specific, Sp; found in 7 or fewer mice) or not allocated to a cluster (unclustered; UC) are also shown. **(B)** Principal components analyses (PCA) of GP33 reactive CDR3β cluster abundance at d8 and d60. **(C)** Repertoire similarity in terms of cluster abundance between mice at d8 and d60, measured by calculating the Morisita-Horn index between all pairs of mice at each timepoint. Equivalent analyses were done for CDR3βs from GP276 reactive CD8 T cells in **(D-F)**. P-values were calculated by two-tailed Wilcoxon rank-sum tests.

Figure S6

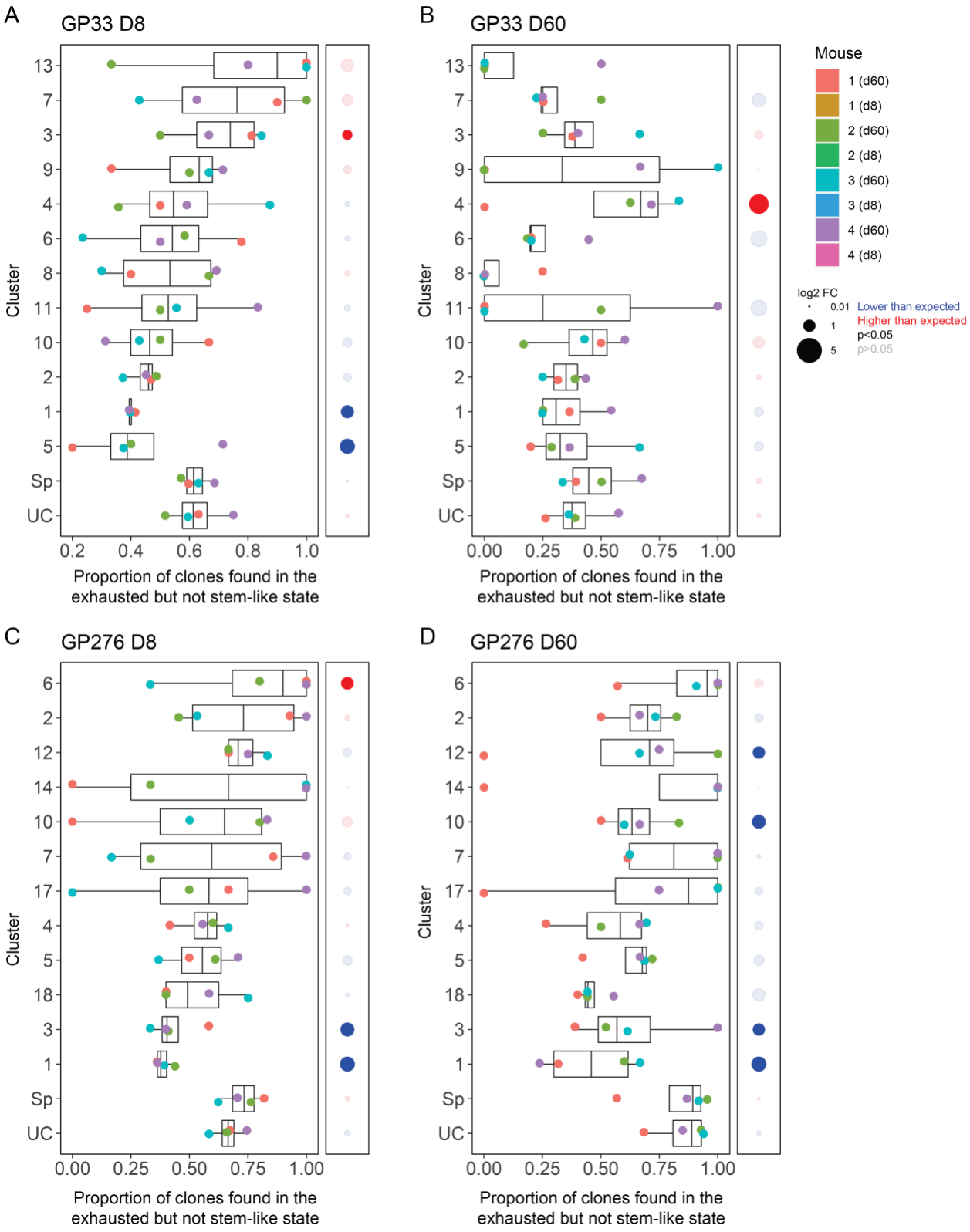

A
GP33 D8
B
GP33 D60
C
GP276 D8
D
GP276 D60
Cluster
Sp
UC
Proportion of clones found in the exhausted but not stem-like state
Mouse
1 (d60)
1 (d8)
2 (d60)
2 (d8)
3 (d60)
3 (d8)
4 (d60)
4 (d8)
log2 FC
0.01
1
5
Lower than expected
Higher than expected
p<0.05
p>0.05

Figure S6. Differential abundance of phenotypes between clusters. Proportion of clones in each cluster found only in the exhausted subset at d8 and d60 in the GP33 **(A)** or GP276 **(B)** repertoires. Data for non-ubiquitous clusters (mouse specific, Sp; found in 7 or fewer mice) and CDR3s not allocated to a cluster (unclustered; UC) are also shown. Circles to the right of the boxplots indicate the fold change of cluster abundance compared to what is expected by chance (red, greater than expected; blue, less than expected), and whether this difference is statistically significant.

Figure S7

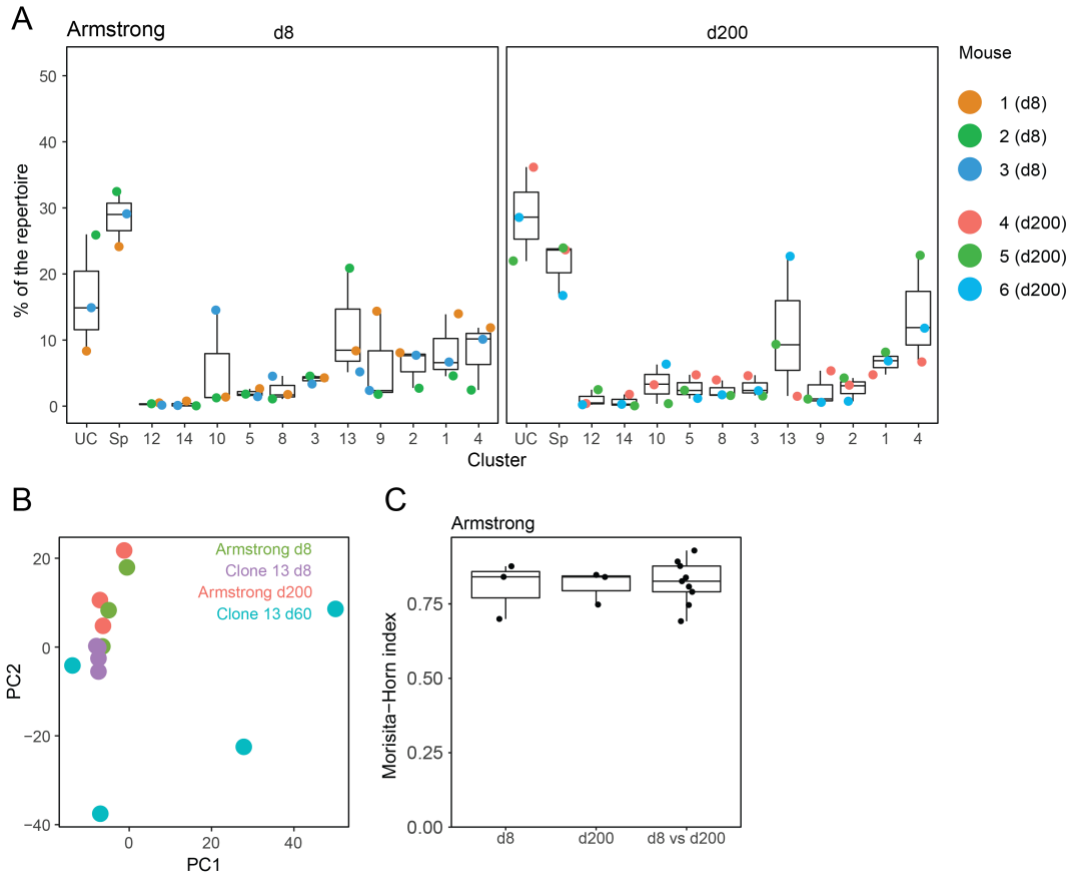


Figure S7: TCR cluster usage in acute infection. (A) Percentage of the GP33 reactive CD8 T cell repertoire occupied by each ubiquitous TCRdist cluster early in the course of Armstrong infection (day 8) and following clearance of infection (day 200). Each point is one mouse. The proportion of the repertoire occupied by TCRs belonging to non-ubiquitous clusters (mouse specific, Sp; found in 5 or fewer mice) or not allocated to a cluster (unclustered; UC) are also shown. **(B)** Principal components analyses (PCA) of GP33 reactive TCR cluster abundance at d8 and d200. PCA results for the GP33 reactive repertoires of clone 13 infected mice at d8 and d60 are shown for comparison. **(C)** Evaluation of repertoire similarity of mice at d8, d200 and d8 compared to d200 post infection, measured by calculating the Morisita-Horn index between pairs of mice.

Figure S8

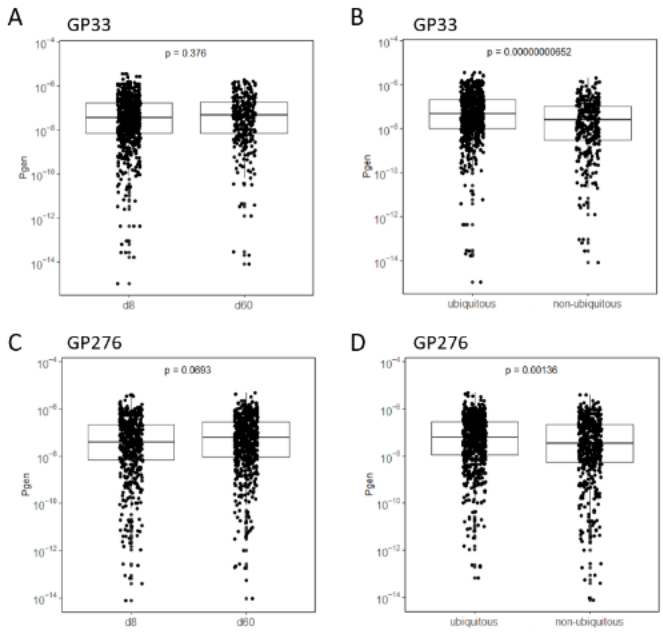


Figure S8: Probability of generation of TCR sequences. Probability of generation (pgen) values of TCR sequences from GP33 (A) and GP276 (C) reactive clones were compared for clones obtained at d8 and d60 of clone 13 infection. Pgen values of TCR sequences belonging to ubiquitous and non-ubiquitous clones reactive to GP33 (B) and GP276 (D) are shown.

Figure S9

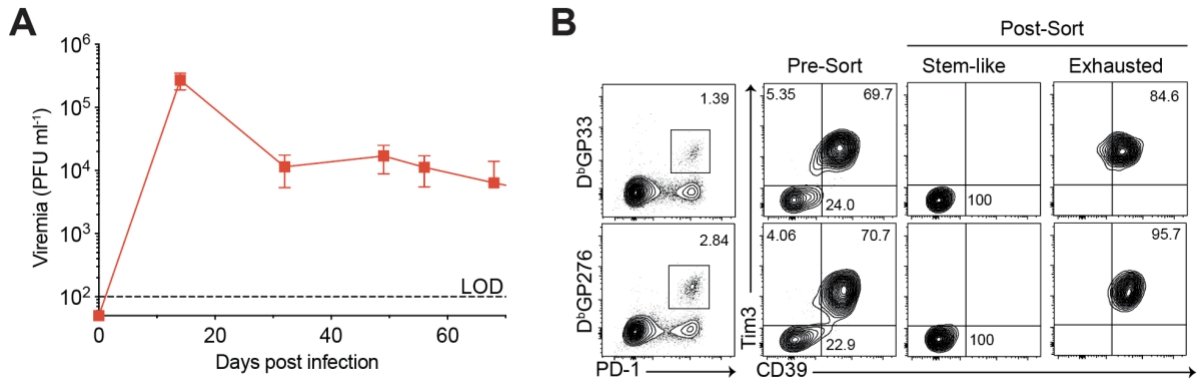


Figure S9. Chronic LCMV viremia and CD8+ T cell subset sorting strategy. (A) Kinetics of viral load of LCMV clone 13 infected mice. **(B)** Flow cytometry plots show the strategy used to sort LCMV-specific CD8+ T cell subsets based on Tim3 and CD39 for TCR sequencing.