

# Analysis of integration site distributions and relative clonal abundance for subject pP10

*August 13, 2019*

## Contents

|  |           |
|--|-----------|
| <b>Summary</b>   | <b>2</b>  |
| Is there a rich population of progenitor cells delivering mature cells to the periphery? . . . . . | 2         |
| Do any cell clones account for more than 20% of all clones? . . . . .                              | 2         |
| Are any cell clones increasing in proportion over time? . . . . .                                  | 3         |
| <b>Introduction</b>  | <b>4</b>  |
| <b>Sample Summary</b>  | <b>5</b>  |
| <b>Tracking of clonal abundances</b>   | <b>7</b>  |
| Relative abundance of cell clones . . . . .  | 7         |
| Longitudinal behavior of major clones . . . . .  | 13        |
| Integration sites near particular genes of interest . . . . .                                      | 16        |
| Sample relative abundance heatmap . . . . .  | 19        |
| <b>What are the most frequently occurring gene types in the subject?</b>                           | <b>24</b> |
| <b>Methods</b>   | <b>28</b> |

# Summary

Is there a rich population of progenitor cells delivering mature cells to the periphery?

To provide a simple measure, we ask whether there are  $\geq 1000$  descendants of independent progenitors (i.e. unique integration sites) in minimally fractionated cell specimens (Whole blood, T cells, B cells, NK cells, Neutrophils, Monocytes and PBMC). Cell specimens that pass these criteria are operationally designated Rich.

| Time point | B CELLS | NK CELLS | PBMC | T CELLS | Rich |
|------------|---------|----------|------|---------|------|
| M15        | na      | na       | 105  | na      | No   |
| M125       | na      | na       | 542  | na      | No   |
| M125.7     | 14      | 6        | 224  | 276     | No   |
| M154.6     | na      | na       | 684  | na      | No   |
| M176.4     | na      | na       | 388  | na      | No   |
| M188.4     | na      | na       | 350  | na      | No   |
| M200       | na      | na       | 385  | na      | No   |

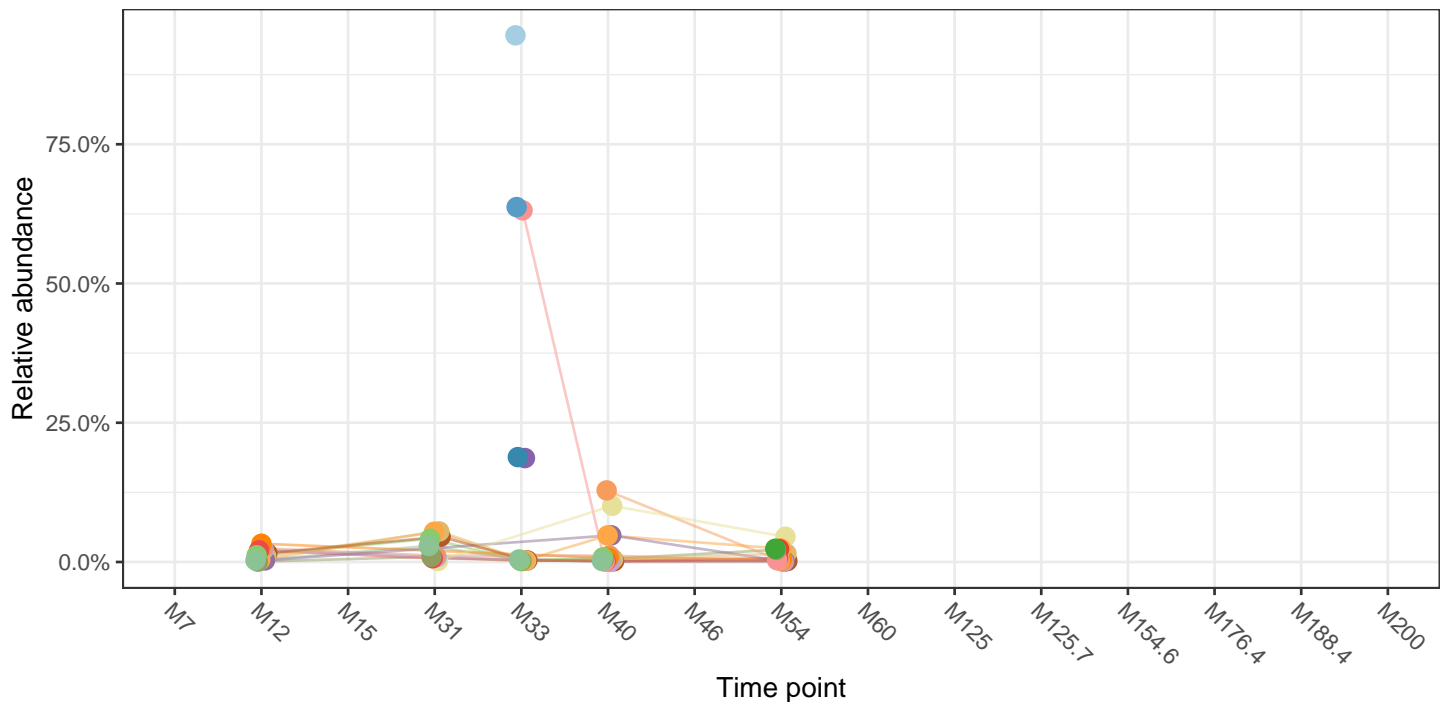
Do any cell clones account for more than 20% of all clones?

For some trials, a reporting criteria is whether any cell clones expand to account for greater than 20% of all clones. The table below highlights samples with relative abundances  $\geq 20\%$  considering only samples with 50 or more inferred cells.

| IntSite        | Abundance | Relative abundance | time point | Cell type       | Nearest gene | Distance (KB) | Nearest oncogene | Distance (KB) |
|----------------|-----------|--------------------|------------|-----------------|--------------|---------------|------------------|---------------|
| chr11-33881660 | 2         | 63.1%              | M33        |                 | LMO2         | 0.00          | LMO2             | 0.00          |
| chr11-33881660 | 2         | 94.6%              | M33        | BLASTBM         | LMO2         | 0.00          | LMO2             | 0.00          |
| chr11-33881660 | 2         | 63.7%              | M33        | BM (CD3+)       | LMO2         | 0.00          | LMO2             | 0.00          |
| chr6-41966616  | 4         | 31.4%              | M125.7     | PBMC            | CCND3        | 0.00          | CCND3            | 0.00          |
| chr6-41966616  | 4         | 25.3%              | M125.7     | T CELLS         | CCND3        | 0.00          | CCND3            | 0.00          |
| chr6-41966616  | 4         | 27.8%              | M125.7     | CD3POSCD45RAPOS | CCND3        | 0.00          | CCND3            | 0.00          |
| chr6-41966616  | 129       | 21.1%              | M154.6     | CD4NAIVE        | CCND3        | 0.00          | CCND3            | 0.00          |
| chr6-41966616  | 161       | 25.6%              | M154.6     | CD8NAIVE        | CCND3        | 0.00          | CCND3            | 0.00          |
| chr6-41966616  | 626       | 21.4%              | M176.4     | PBMC            | CCND3        | 0.00          | CCND3            | 0.00          |
| chr6-41966616  | 534       | 20.9%              | M188.4     | PBMC            | CCND3        | 0.00          | CCND3            | 0.00          |
| chr6-41966616  | 639       | 20.3%              | M200       | PBMC            | CCND3        | 0.00          | CCND3            | 0.00          |

Are any cell clones increasing in proportion over time?

The plot below details the longitudinal sample relative abundances of the most abundant 20 clones where only samples with 50 or more inferred cells are considered.



Data source

- 454
- ▲ Illumina

Clone

- BLASTBM : LMO2 \*~!  
chr11-33881660
- BM (CD3+) : LMO2 \*~!  
chr11-33881660
- BM (CD3+) : SPAG6 \*~!  
chr10+22370708
- NA : ASB2 \*  
chr14-93957427
- NA : BCAT1 \*  
chr12-24931660
- NA : ETV6  
chr12+11954933
- NA : LINC01280 \*~  
chr2+216594615
- NA : LMO2 \*~!  
chr11-33881660
- NA : LMO2 ~!  
chr11+33941890
- NA : PREX2 \*  
chr8+67957311
- NA : RPL14  
chr3+40456837
- NA : RUFY4 \*  
chr2-218056066
- NA : RUNX1 \*~!  
chr21+34988087
- NA : RUNX1 \*~!  
chr21+35040047
- NA : SLC44A2 \*~  
chr19-10603812
- NA : SPAG6 \*~!  
chr10+22370708
- NA : STK24 \*  
chr13-98509246
- NA : TSPAN14 \*  
chr10+80460328
- NA : UBASH3A ~  
chr21-42453277
- NA : USP25  
chr21+15438974

# Introduction

The attached report describes results of analysis of integration site distributions and relative abundance for samples from gene therapy trials. For cases of gene correction in circulating blood cells, it is possible to harvest cells sequentially from blood to monitor cell populations. Frequency of isolation information can provide information on the clonal structure of the population. This report summarizes results for subject pP10 over time points M7, M12, M15, M31, M33, M40, M46, M54, M60, M125, M125.7, M154.6, M176.4, M188.4, M200 in UCSC genome draft .

The samples studied in this report, the numbers of sequence reads, recovered integration vectors, and unique integration sites available for this subject are shown below. We quantify population clone diversity using Gini coefficients, Shannon index, and UC50. The Gini coefficient provides a measure of inequality in clonal abundance in each sample. The coefficient equals zero when all sites are equally abundant (polyclonal) and increases as fewer sites account for more of the total (oligoclonal). Shannon index is another widely used measure of diversity and it accounts for both abundance and evenness of the integration events. Alternatively, the UC50 is the number of unique clones which make up the top 50% of the sample's abundance. For polyclonal samples, one may expect a low Gini coefficient, high Shannon Index, and high UC50 (proportional to the total number of unique sites identified in the sample).

Under most circumstances only a subset of sites will be sampled. We thus include an estimate of sample size based on frequency of isolation information from the SonicLength method (Berry, 2012). The 'S.chao1' column denotes the estimated lower bound for population size derived using Chao estimate (Chao, 1987). If sample replicates were present then estimates were subjected to jackknife bias correction.

We estimate the numbers of cell clones sampled using the SonicLength method (Berry, 2012); this is summarized in the column "Inferred cells". Integration sites were recovered using ligation mediated PCR after random fragmentation of genomic DNA, which reduces recovery biases compared with restriction enzyme cleavage. Relative abundance was not measured from read counts, which are known to be inaccurate, but from marks introduced into DNA specimens prior to PCR amplification using the SonicLength method PMID:22238265.

We quantify population diversity using Gini coefficients, Shannon index, and UC50. The Gini coefficient provides a measure of inequality in clonal abundance in each sample. The coefficient equals zero when all sites are equally abundant (polyclonal) and increases as fewer sites account for more of the total (oligoclonal). Shannon index is another widely used measure of diversity and it accounts for both abundance and evenness of the integration events. UC50 is the number of clones which make up the top 50% of the sample's abundance. For polyclonal samples, one may expect a low Gini coefficient, high Shannon Index, and high UC50 (proportional to the total number of unique sites identified in the sample).

Integration positions are reported with the format (nearest gene, chromosome, +/-, genomic position) where the nearest gene is the nearest transcriptional boundary to the integration position, '+' refers to integration in the positive orientation and '-' refers to integration in the reverse orientation. Reported distances are signed where the sign indicates if integrations are upstream (-) or downstream (+, no sign) of the nearest gene. Nearest genes possess additional annotations described in the table below.

| Symbol | Meaning   |
|--------|---|
| *      | site is within a transcription unit                 |
| ~      | site is within 50kb of a cancer related gene        |
| !      | nearest gene was associated with lymphoma in humans |

# Sample Summary

The table below provides population statistics for each analyzed sample. Occasionally multiple samples from the same cell fraction and time point are analyzed where only the sample with greatest number of inferred cells is considered in this report.

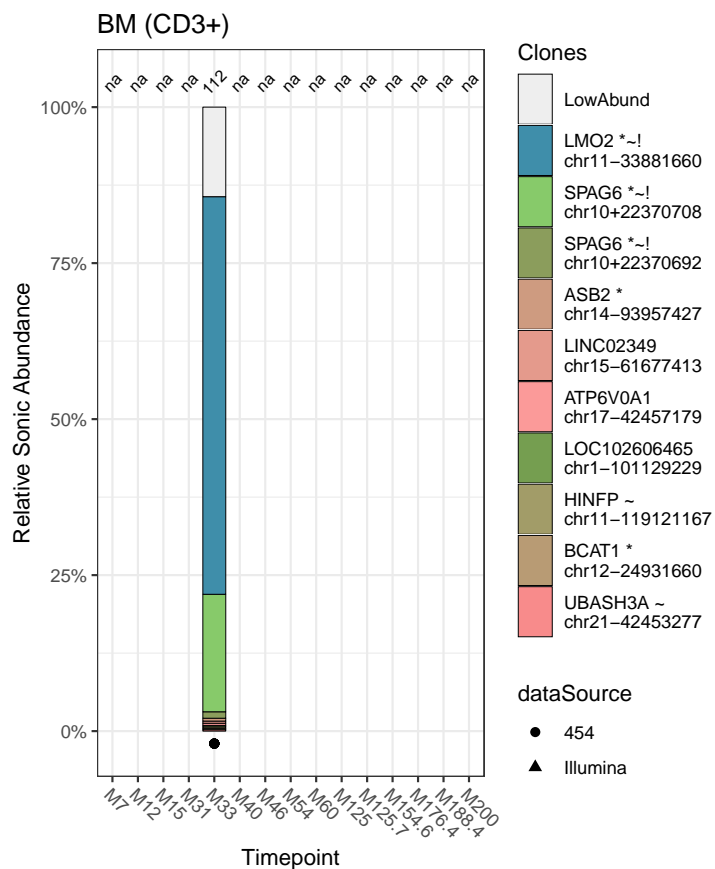
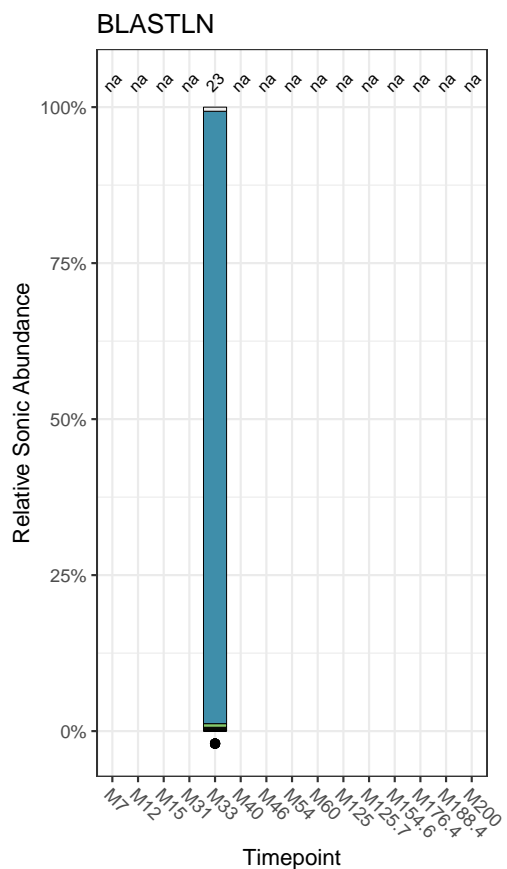
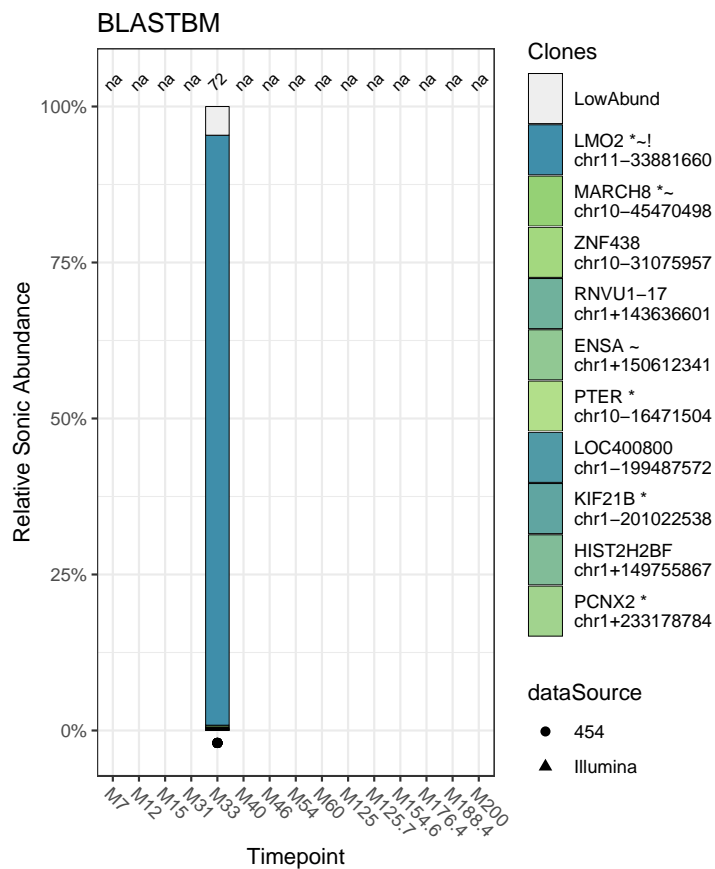
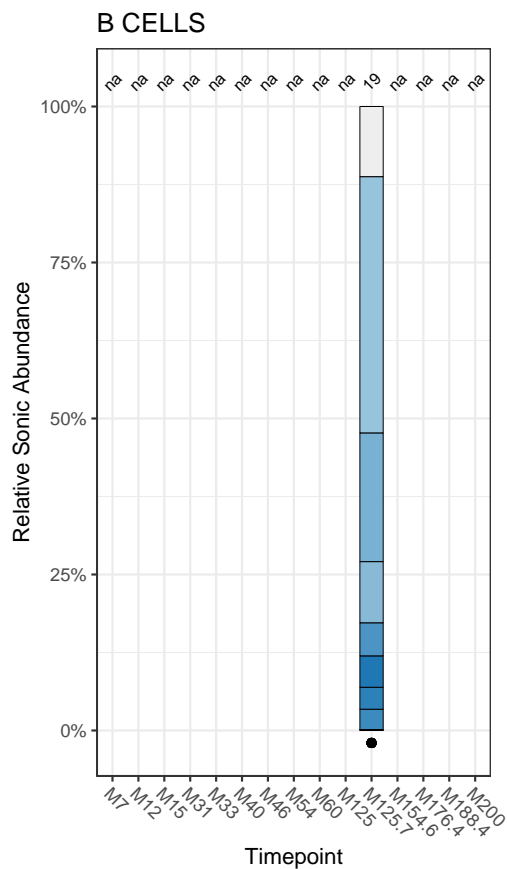
| GTSP     | dataSource | Patient | Timepoint | CellType        | TotalReads | InferredCells | UniqueSites | Gini  | Chao1 | Shannon | Pielou | UC50 | Included | VCN |
|----------|------------|---------|-----------|-----------------|------------|---------------|-------------|-------|-------|---------|--------|------|----------|-----|
| LS-22    | 454        | pP10    | M7        | CD19            | 19         | 1             | 1           | 0.000 | 1     | 0.00    | NaN    | 1    | yes      | NA  |
| LS-24    | 454        | pP10    | M7        | CD3             | 2,767      | 10            | 10          | 0.000 | 55    | 2.30    | 1.000  | 6    | yes      | NA  |
| LS-28    | 454        | pP10    | M7        | CD56            | 2,834      | 12            | 12          | 0.000 | 78    | 2.48    | 1.000  | 7    | yes      | NA  |
| LS-29    | 454        | pP10    | M7        | GRAN            | 118        | 2             | 2           | 0.000 | 3     | 0.69    | 1.000  | 2    | yes      | NA  |
| LS-13    | 454        | pP10    | M12       | NA              | 2,101.5    | 441           | 395         | 0.094 | 1,902 | 5.94    | 0.994  | 175  | yes      | NA  |
| LS-31    | 454        | pP10    | M12       | PBL             | 2,022      | 410           | 393         | 0.040 | 4,310 | 5.96    | 0.997  | 189  | no       | NA  |
| LS-34    | 454        | pP10    | M12       | PBL             | 2,085.5    | 438           | 394         | 0.091 | 1,987 | 5.94    | 0.994  | 176  | yes      | NA  |
| LS-38    | 454        | pP10    | M15       | PBMC            | 2,043      | 105           | 105         | 0.000 | 5,565 | 4.65    | 1.000  | 53   | yes      | NA  |
| LS-14    | 454        | pP10    | M31       | NA              | 1,807      | 435           | 386         | 0.100 | 1,696 | 5.92    | 0.993  | 169  | yes      | NA  |
| LS-23    | 454        | pP10    | M31       | CD3             | 1,473.5    | 395           | 371         | 0.057 | 2,772 | 5.89    | 0.996  | 174  | yes      | NA  |
| LS-25    | 454        | pP10    | M31       | CD3+            | 1,785      | 429           | 381         | 0.099 | 1,690 | 5.90    | 0.993  | 167  | yes      | NA  |
| LS-15    | 454        | pP10    | M33       | NA              | 879        | 114           | 104         | 0.079 | 501   | 4.61    | 0.994  | 48   | yes      | NA  |
| LS-19    | 454        | pP10    | M33       | BLASTBM         | 2,478      | 72            | 71          | 0.014 | 1,278 | 4.26    | 0.999  | 36   | yes      | NA  |
| LS-20    | 454        | pP10    | M33       | BLASTLN         | 2,591      | 23            | 21          | 0.079 | 78    | 3.01    | 0.990  | 10   | yes      | NA  |
| LS-21    | 454        | pP10    | M33       | BM (CD3+)       | 871        | 112           | 102         | 0.081 | 483   | 4.59    | 0.993  | 47   | yes      | NA  |
| LS-16    | 454        | pP10    | M40       | NA              | 3,750.5    | 366           | 303         | 0.146 | 995   | 5.65    | 0.989  | 121  | yes      | NA  |
| LS-35    | 454        | pP10    | M40       | PBL             | 3,722      | 360           | 298         | 0.146 | 983   | 5.63    | 0.989  | 119  | yes      | NA  |
| LS-32    | 454        | pP10    | M46       | PBL             | 4,619      | 676           | 614         | 0.082 | 3,028 | 6.39    | 0.995  | 277  | yes      | NA  |
| LS-17    | 454        | pP10    | M54       | NA              | 2,860      | 609           | 490         | 0.162 | 1,460 | 6.12    | 0.988  | 186  | yes      | NA  |
| LS-36    | 454        | pP10    | M54       | PBL             | 2,830      | 611           | 492         | 0.162 | 1,503 | 6.13    | 0.988  | 187  | yes      | NA  |
| LS-33    | 454        | pP10    | M60       | PBL             | 4,423.5    | 149           | 142         | 0.045 | 1,273 | 4.94    | 0.997  | 68   | yes      | NA  |
| GTSP0519 | Illumina   | pP10    | M125      | PBMC            | 573,270    | 3,761         | 542         | 0.766 | 988   | 4.64    | 0.737  | 15   | yes      | 1.6 |
| GTSP0141 | 454        | pP10    | M125.7    | PBMC            | 3,635      | 342           | 224         | 0.268 | 612   | 5.26    | 0.972  | 60   | yes      | NA  |
| GTSP0143 | 454        | pP10    | M125.7    | T CELLS         | 2,431      | 438           | 276         | 0.276 | 638   | 5.47    | 0.973  | 76   | yes      | NA  |
| GTSP0144 | 454        | pP10    | M125.7    | B CELLS         | 2,642      | 19            | 14          | 0.214 | 42    | 2.52    | 0.957  | 5    | yes      | NA  |
| GTSP0145 | 454        | pP10    | M125.7    | NK CELLS        | 2,396      | 12            | 6           | 0.000 | 78    | 2.48    | 1.000  | 7    | yes      | NA  |
| GTSP0146 | 454        | pP10    | M125.7    | CD3POSCD45RAPOS | 6,224.5    | 77            | 63          | 0.162 | 302   | 4.05    | 0.978  | 25   | yes      | NA  |
| GTSP0147 | 454        | pP10    | M125.7    | CD3POSCD45ROPOS | 6,020.5    | 246           | 188         | 0.203 | 742   | 5.12    | 0.977  | 66   | yes      | NA  |
| GTSP0856 | Illumina   | pP10    | M154.6    | PBMC            | 795,247    | 6,373         | 684         | 0.796 | 1,242 | 4.74    | 0.726  | 15   | yes      | NA  |
| GTSP0857 | Illumina   | pP10    | M154.6    | CD4MEMORY       | 417,303    | 867           | 330         | 0.546 | 812   | 4.92    | 0.849  | 28   | yes      | NA  |
| GTSP0858 | Illumina   | pP10    | M154.6    | CD4NAIVE        | 361,232    | 610           | 118         | 0.664 | 204   | 3.69    | 0.774  | 8    | yes      | NA  |
| GTSP0859 | Illumina   | pP10    | M154.6    | CD8MEMORY       | 595,117    | 944           | 180         | 0.737 | 465   | 3.66    | 0.705  | 5    | yes      | NA  |

| GTSP     | dataSource | Patient | Timepoint | CellType | TotalReads | InferredCells | UniqueSites | Gini  | Chao1 | Shannon | Pielou | UC50 | Included | VCN  |
|----------|------------|---------|-----------|----------|------------|---------------|-------------|-------|-------|---------|--------|------|----------|------|
| GTSP0860 | Illumina   | pP10    | M154.6    | CD8NAIVE | 379,054    | 628           | 121         | 0.680 | 202   | 3.62    | 0.754  | 8    | yes      | NA   |
| GTSP1534 | Illumina   | pP10    | M176.4    | PBMC     | 697,563    | 2,923         | 388         | 0.787 | 759   | 4.10    | 0.688  | 9    | yes      | 0.64 |
| GTSP2228 | Illumina   | pP10    | M188.4    | PBMC     | 845,657    | 2,558         | 350         | 0.778 | 720   | 4.10    | 0.700  | 10   | yes      | 0.44 |
| GTSP2887 | Illumina   | pP10    | M200      | PBMC     | 1,284,394  | 3,155         | 385         | 0.801 | 685   | 4.04    | 0.679  | 8    | yes      | 0.40 |

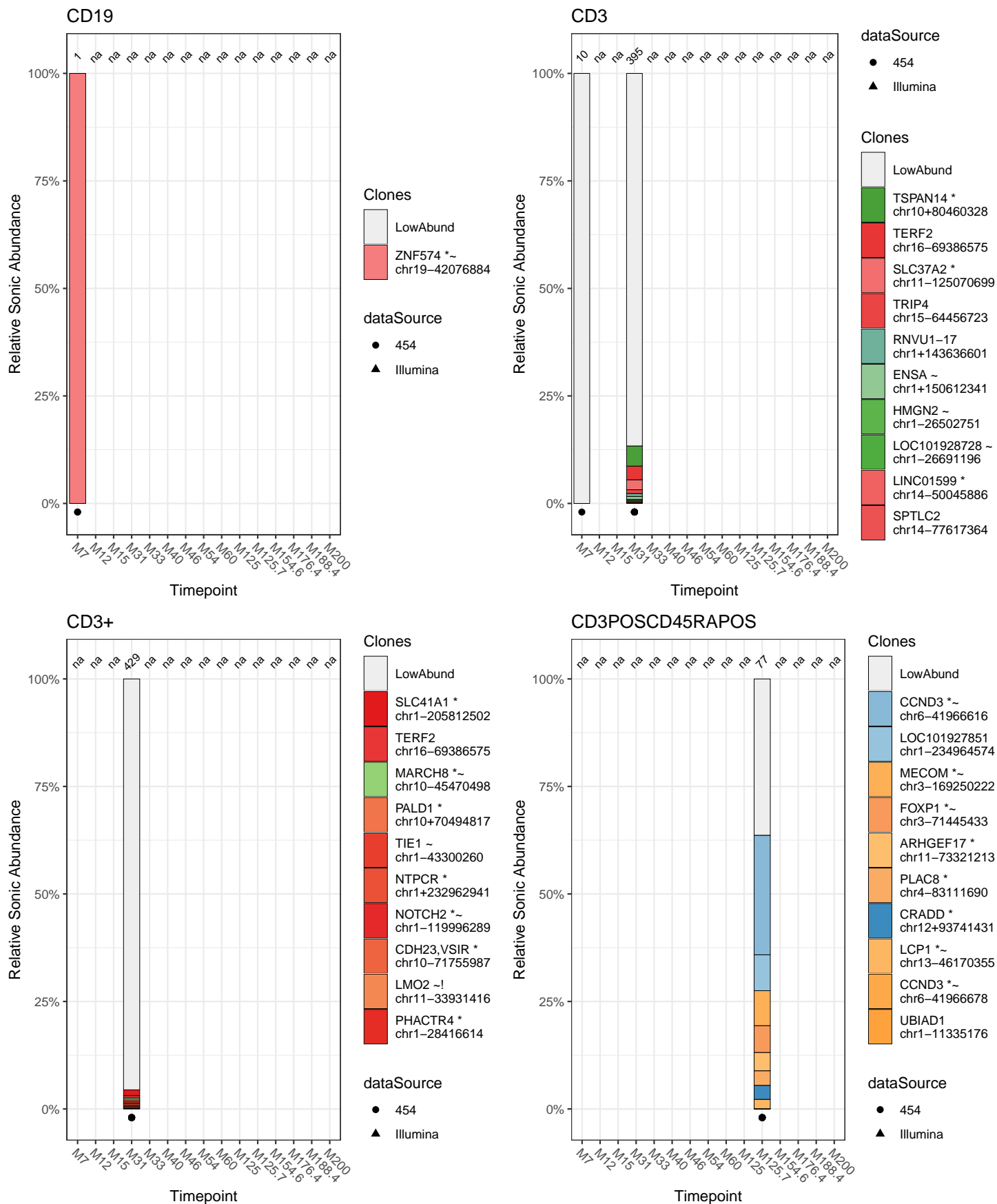
# Tracking of clonal abundances

## Relative abundance of cell clones

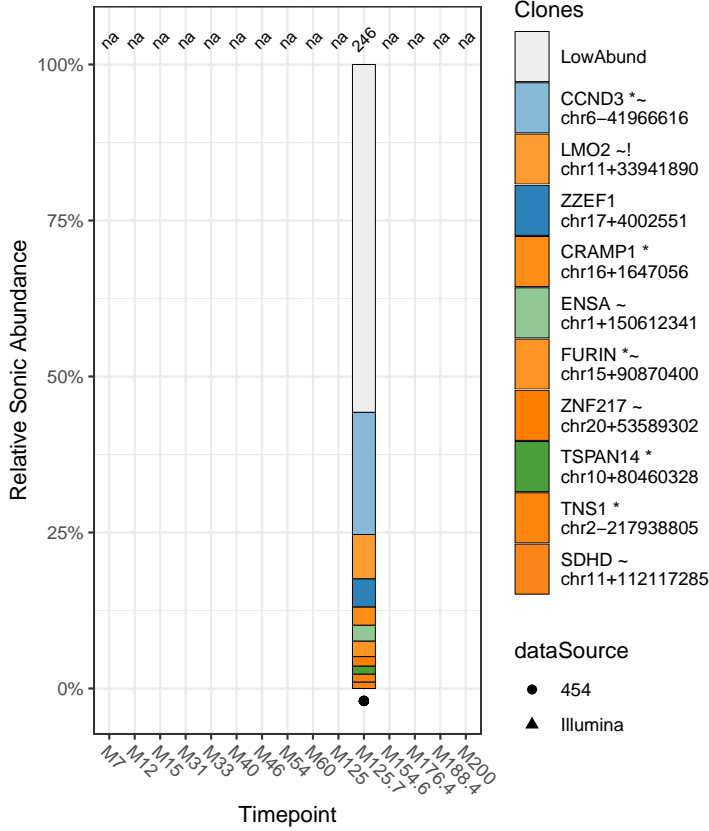
The relative abundances of cell clones is summarized in the stacked bar plots below. The cell fraction studied is named at the top of each plot and the time points are marked at the bottom. The different bars in each panel show the major cell clones, as marked by integration sites where the x-axis indicates time points and the y-axis is scaled by proportion of the total cells sampled. The top 10 most abundant clones from each cell type have been named by the nearest gene while the remaining sites are binned as low abundance (LowAbund; grey). The total number of genomic fragments used to identify integration sites are listed atop of each plot. These fragments are generated by restriction endonucleases in 454 sequencing experiments and by sonic shearing in Illumina sequencing experiments. Relative abundances are calculated using the total number of reads associated with clones in 454 sequencing experiments while the number of unique sonic breaks is used in Illumina sequencing experiments.



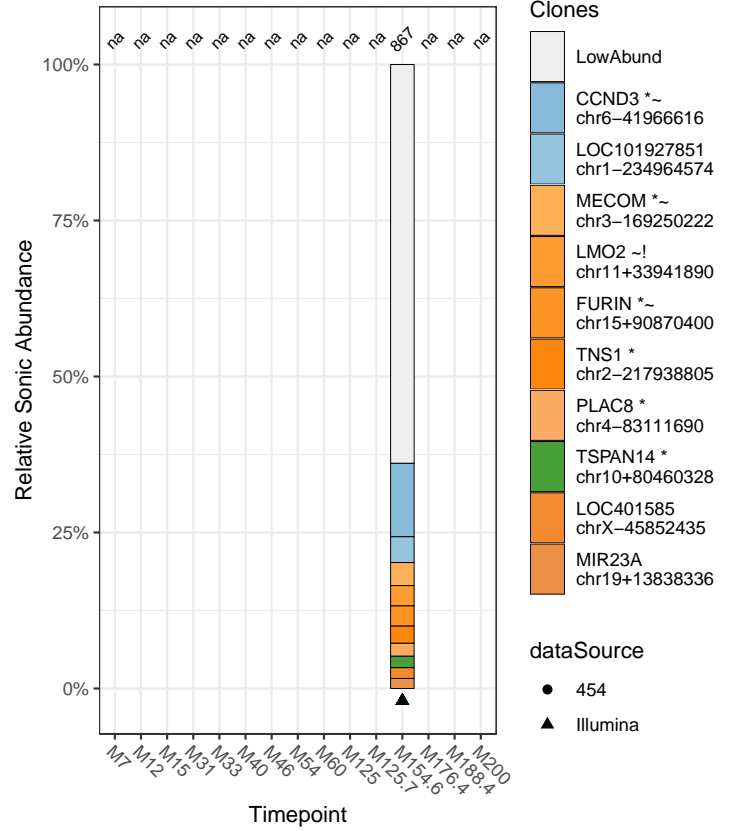




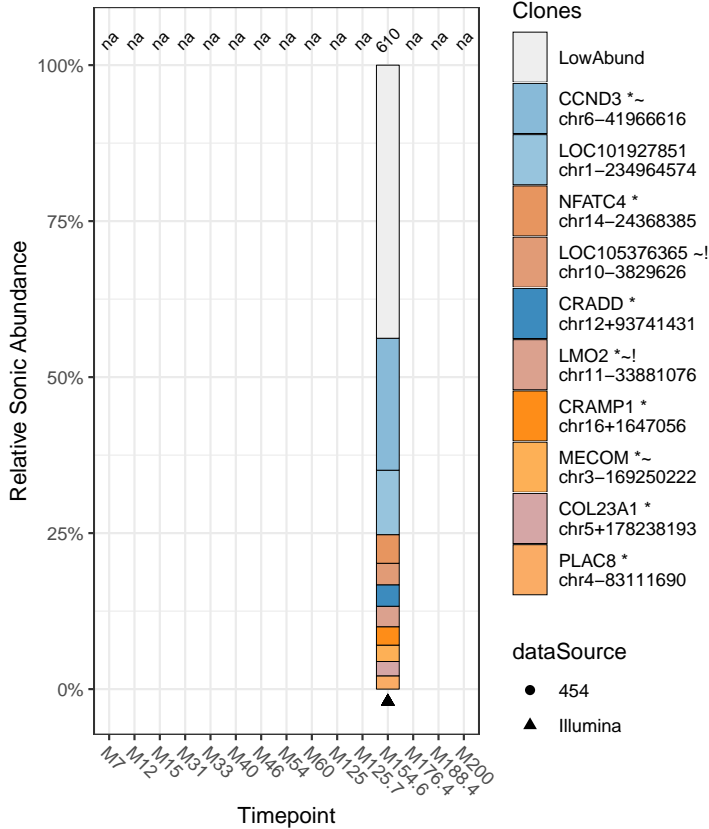
### CD3POS CD45RO POS



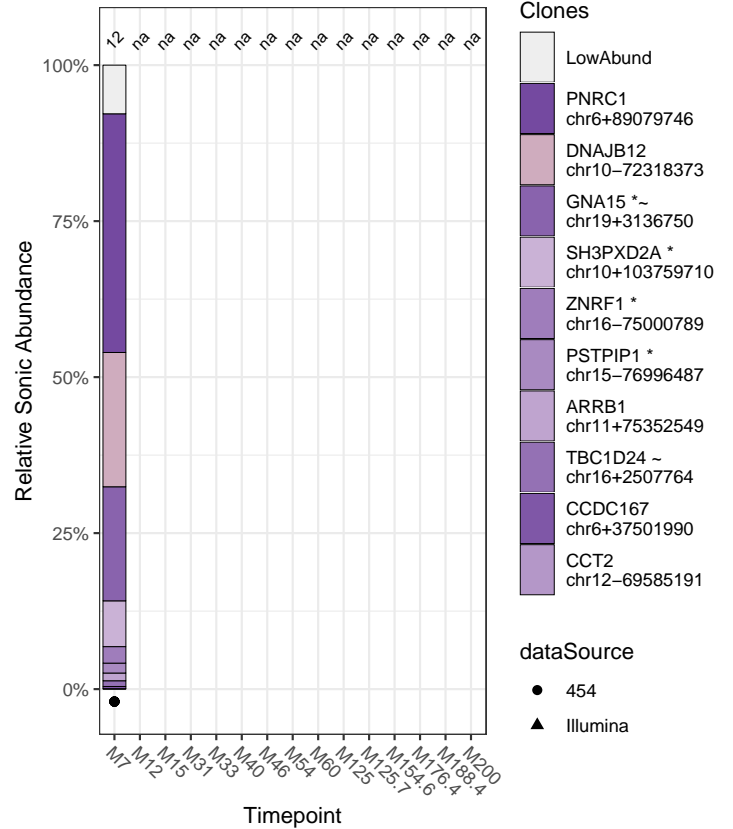
### CD4MEMORY

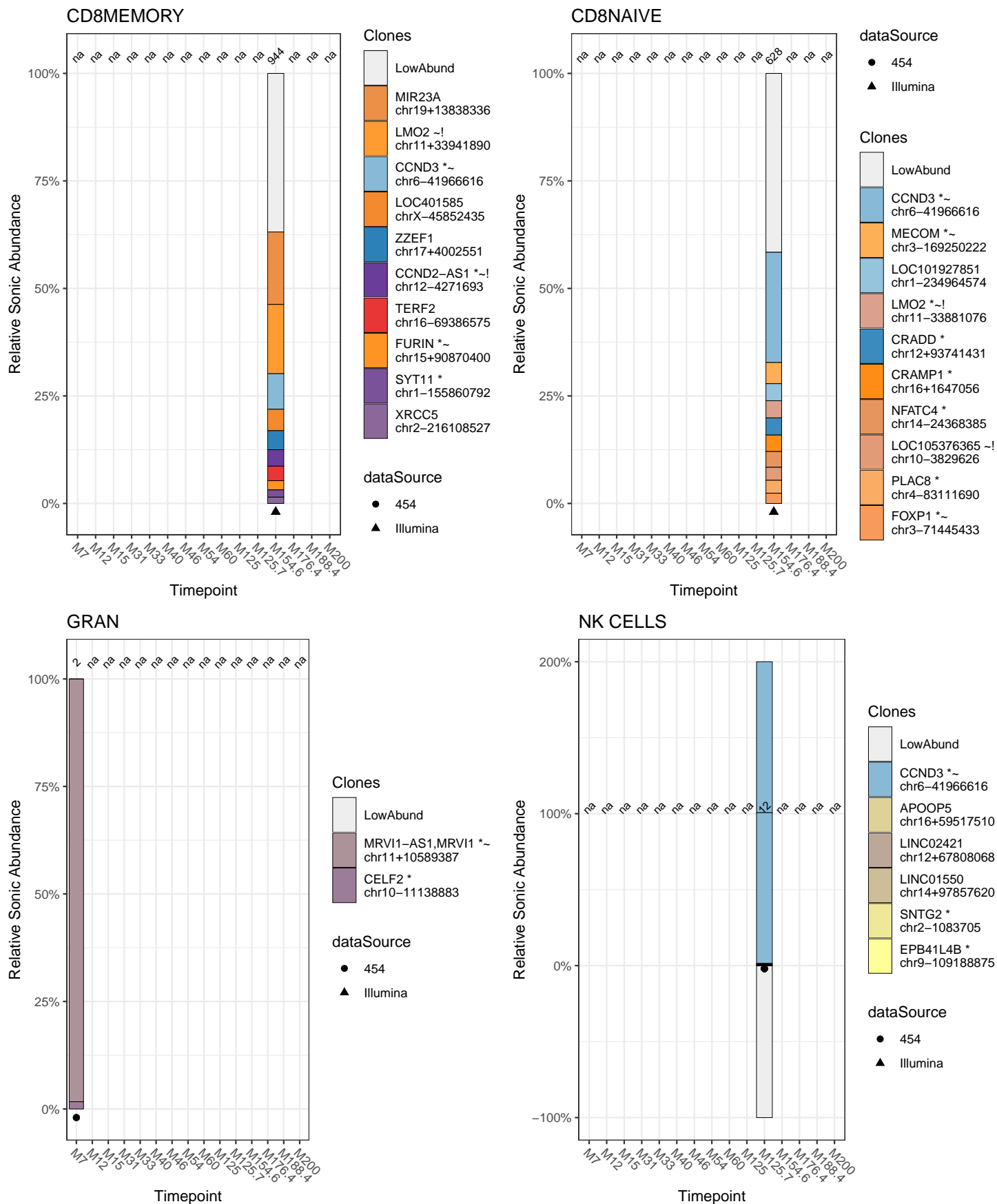


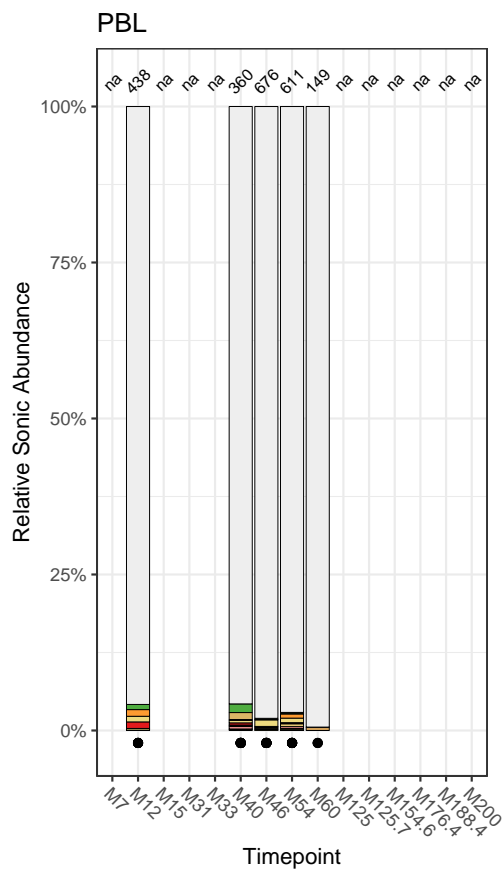
### CD4NAIVE



### CD56





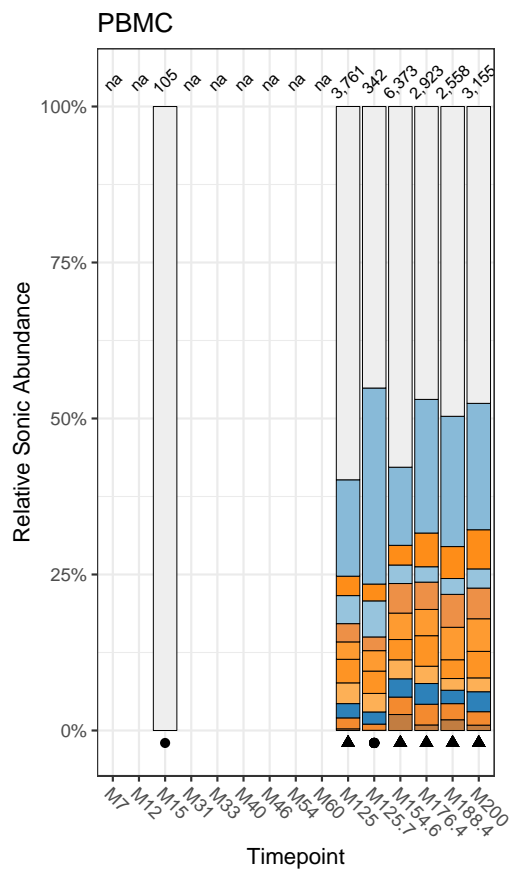


dataSource

- 454
- ▲ Illumina

Clones

- LowAbund
- LOC101928728 ~ chr1-26691196
- CPEB4 \* chr5-173916732
- LMO2 ~! chr11+33941890
- LMO2 ~! chr11+33931259
- SLC41A1 \* chr1-205812502
- CYP1B1 chr2+38093891
- CLEC9A \* chr12-10054630
- ATP6V0A1 chr17-42457179
- BTG2 ~ chr1+203322876
- STMN3 ~ chr20+63654538

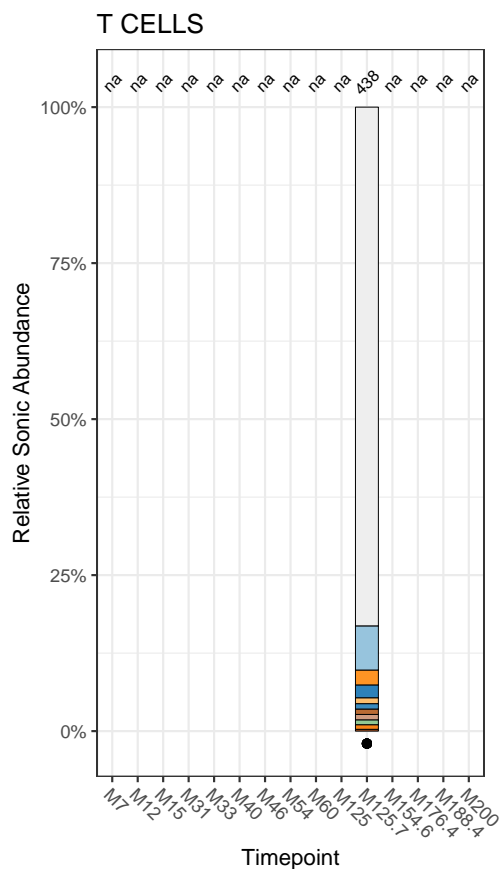


Clones

- LowAbund
- CCND3 \*~ chr6-41966616
- CRAMP1 \* chr16+1647056
- LOC101927851 chr1-234964574
- MIR23A chr19+13838336
- LMO2 ~! chr11+33941890
- FURIN \*~ chr15+90870400
- MECOM \*~ chr3-169250222
- ZZEF1 chr17+4002551
- LOC401585 chrX-45852435
- MIR6076 chr14+49969555

dataSource

- 454
- ▲ Illumina



Clones

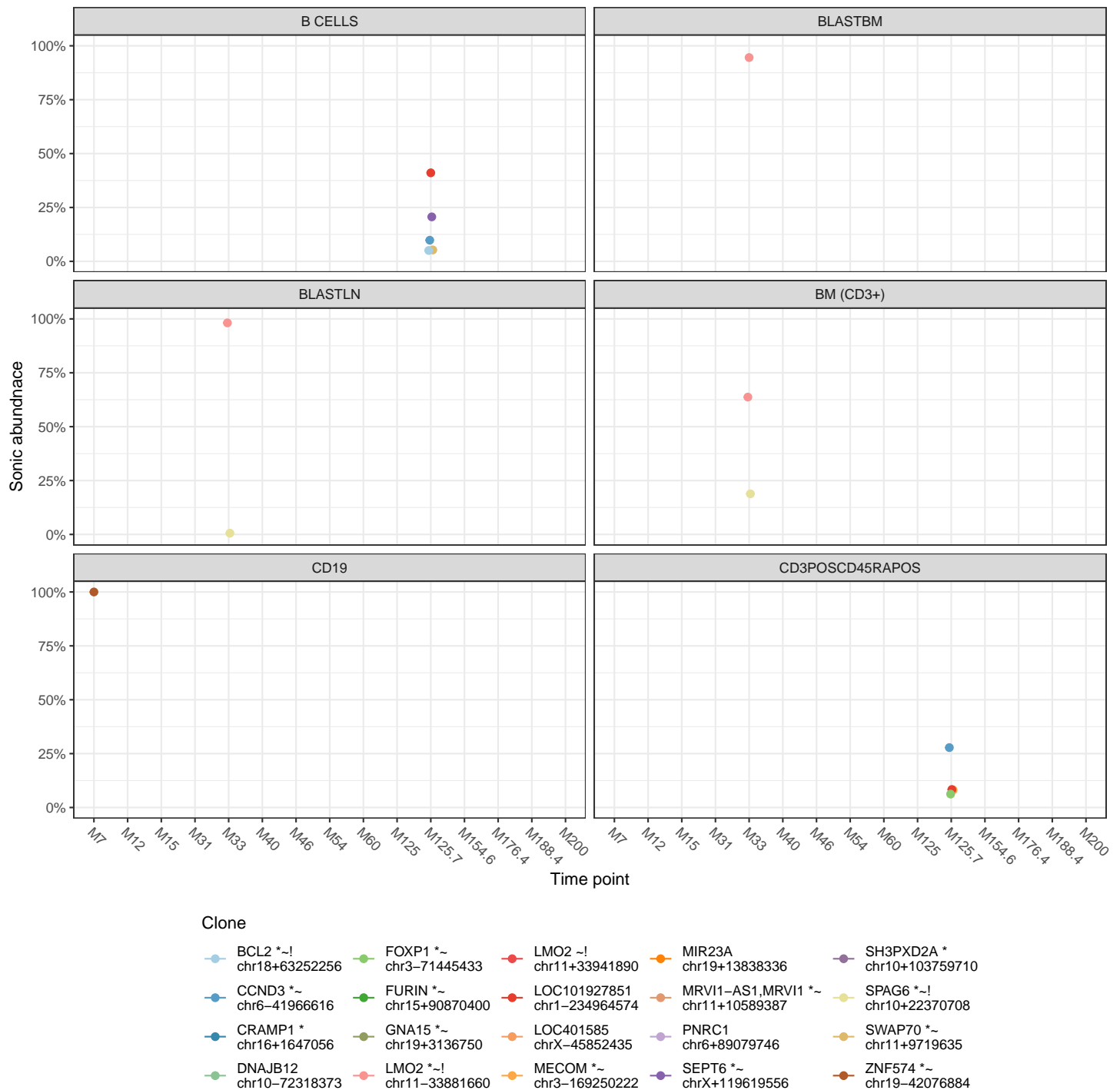
- LowAbund
- LOC101927851 chr1-234964574
- FURIN \*~ chr15+90870400
- ZZEF1 chr17+4002551
- ARHGEF17 \* chr11-73321213
- CRADD \* chr12+93741431
- ZNF219 \*~ chr14-21104531
- ASB2 \* chr14-93957427
- ENSA ~ chr1+150612341
- SDHD ~ chr11+112117285
- LINC01588 \* chr14-49999840

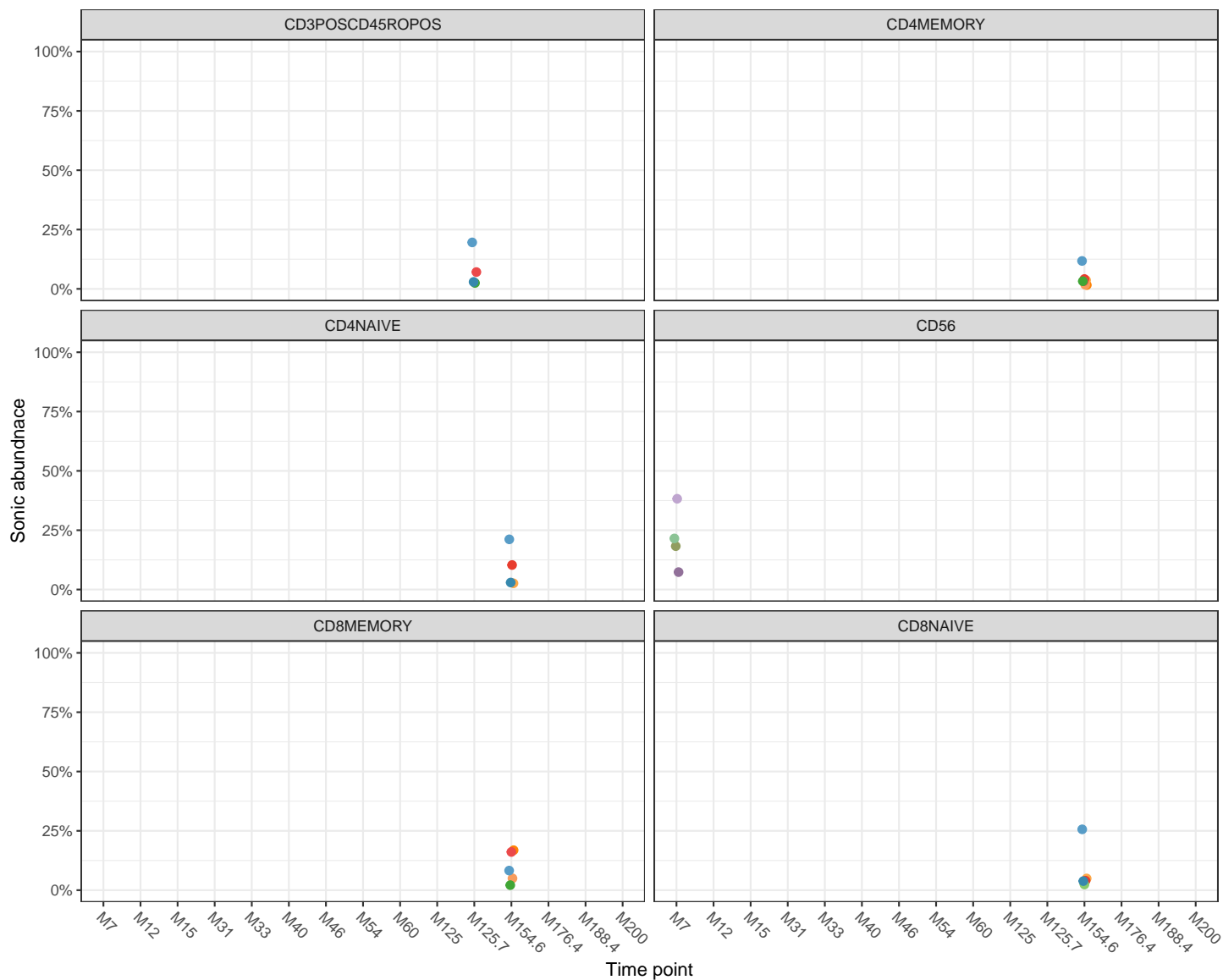
dataSource

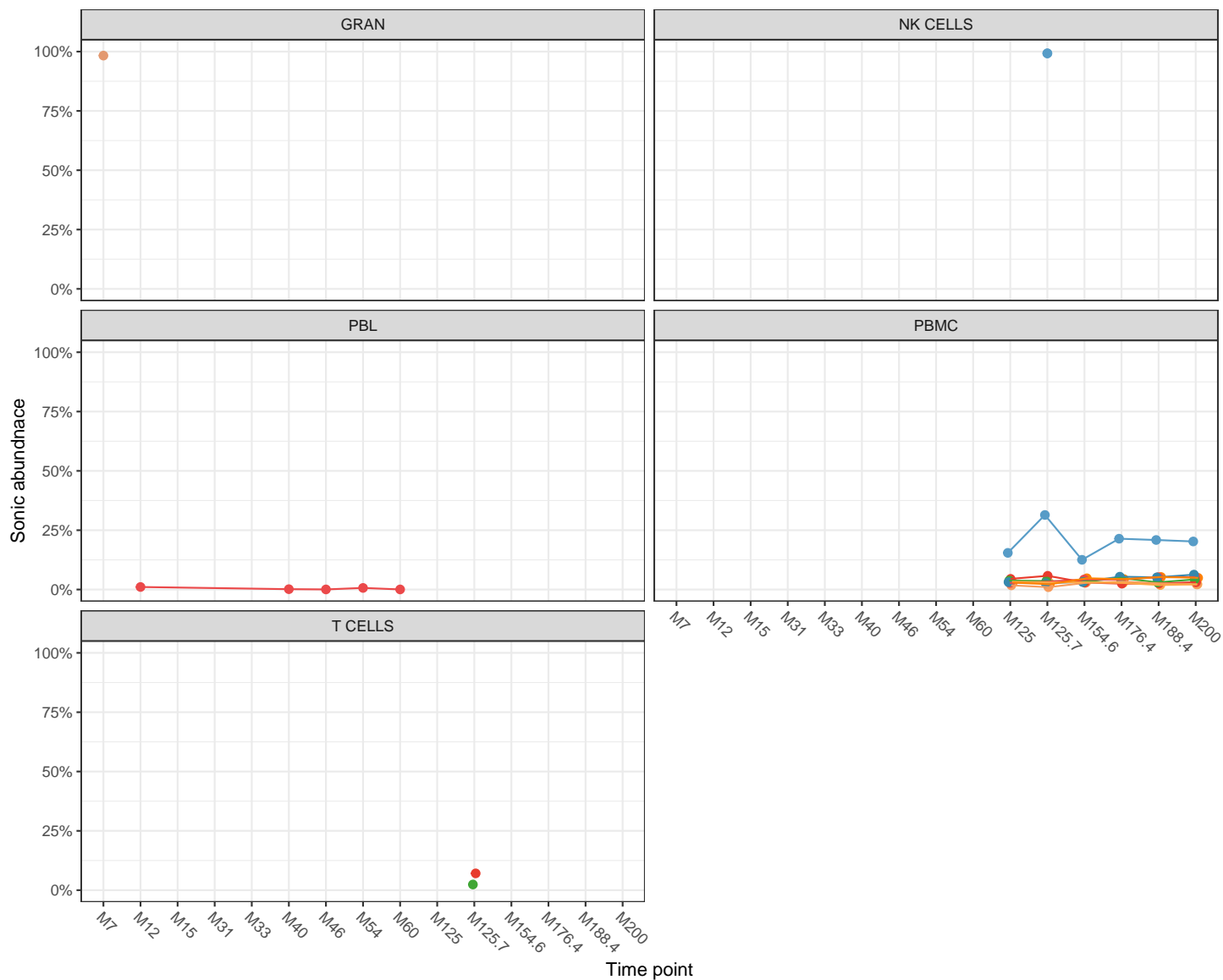
- 454
- ▲ Illumina

## Longitudinal behavior of major clones

When multiple time points are available, it is of interest to track the behavior of the most abundant clones across different cell types. A plot of the relative abundances of the most abundant 20 clones is shown below. For cases where only a single time point is available, the data is plotted as unlinked points.





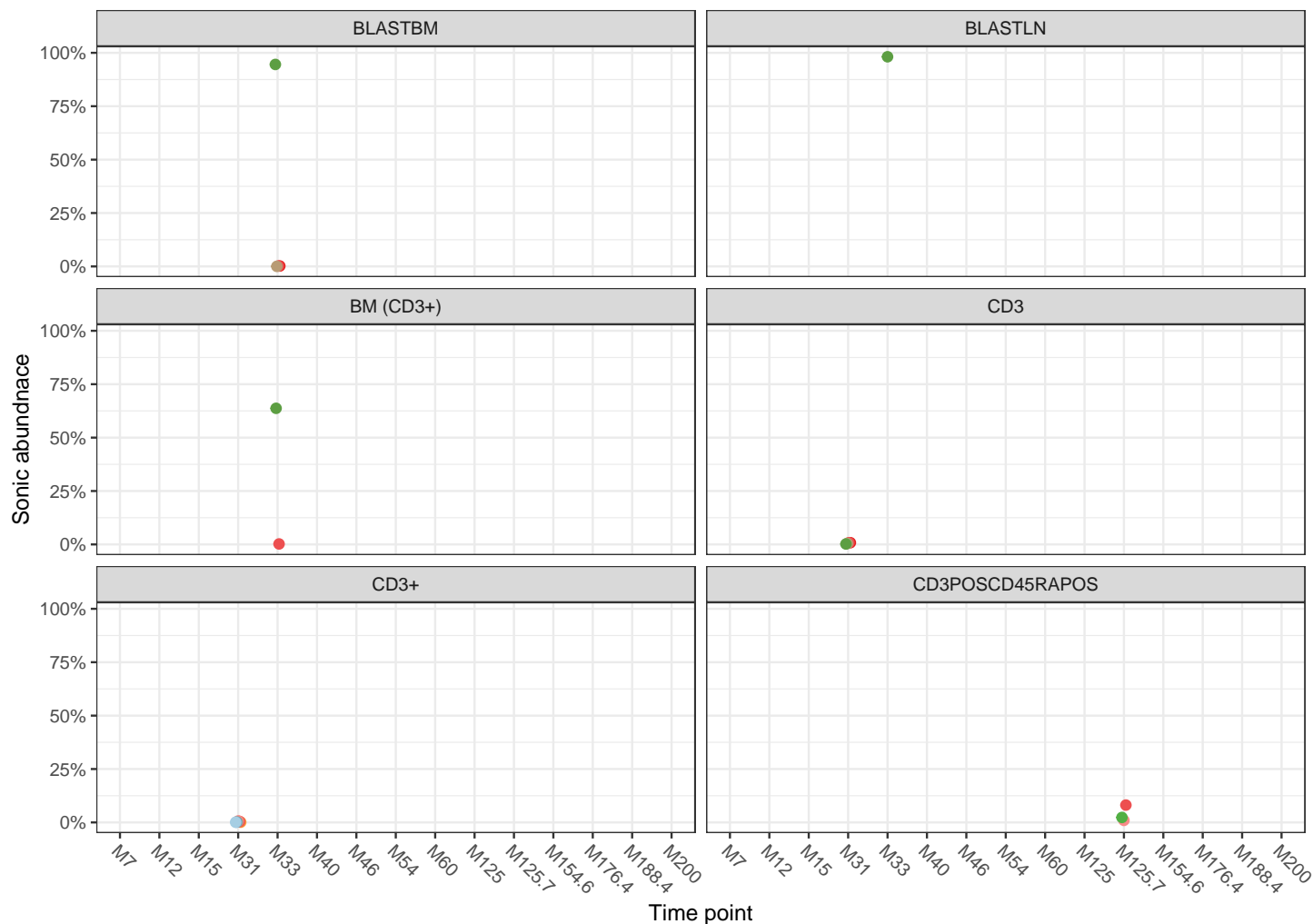


#### Clone

|                            |                            |                                |                                      |                               |
|----------------------------|----------------------------|--------------------------------|--------------------------------------|-------------------------------|
| BCL2 *~!<br>chr18+63252256 | FOXP1 *~<br>chr3-71445433  | LMO2 ~!<br>chr11+33941890      | MIR23A<br>chr19+13838336             | SH3PXD2A *<br>chr10+103759710 |
| CCND3 *~<br>chr6-41966616  | FURIN *~<br>chr15+90870400 | LOC101927851<br>chr1-234964574 | MRVI1-AS1,MRVI1 *~<br>chr11+10589387 | SPAG6 *~!<br>chr10+22370708   |
| CRAMP1 *<br>chr16+1647056  | GNA15 *~<br>chr19+3136750  | LOC401585<br>chrX-45852435     | PNRC1<br>chr6+89079746               | SWAP70 *~<br>chr11+9719635    |
| DNAJB12<br>chr10-72318373  | LMO2 *~!<br>chr11-33881660 | MECOM *~<br>chr3-169250222     | SEPT6 *~<br>chrX+119619556           | ZNF574 *~<br>chr19-42076884   |

## Integration sites near particular genes of interest

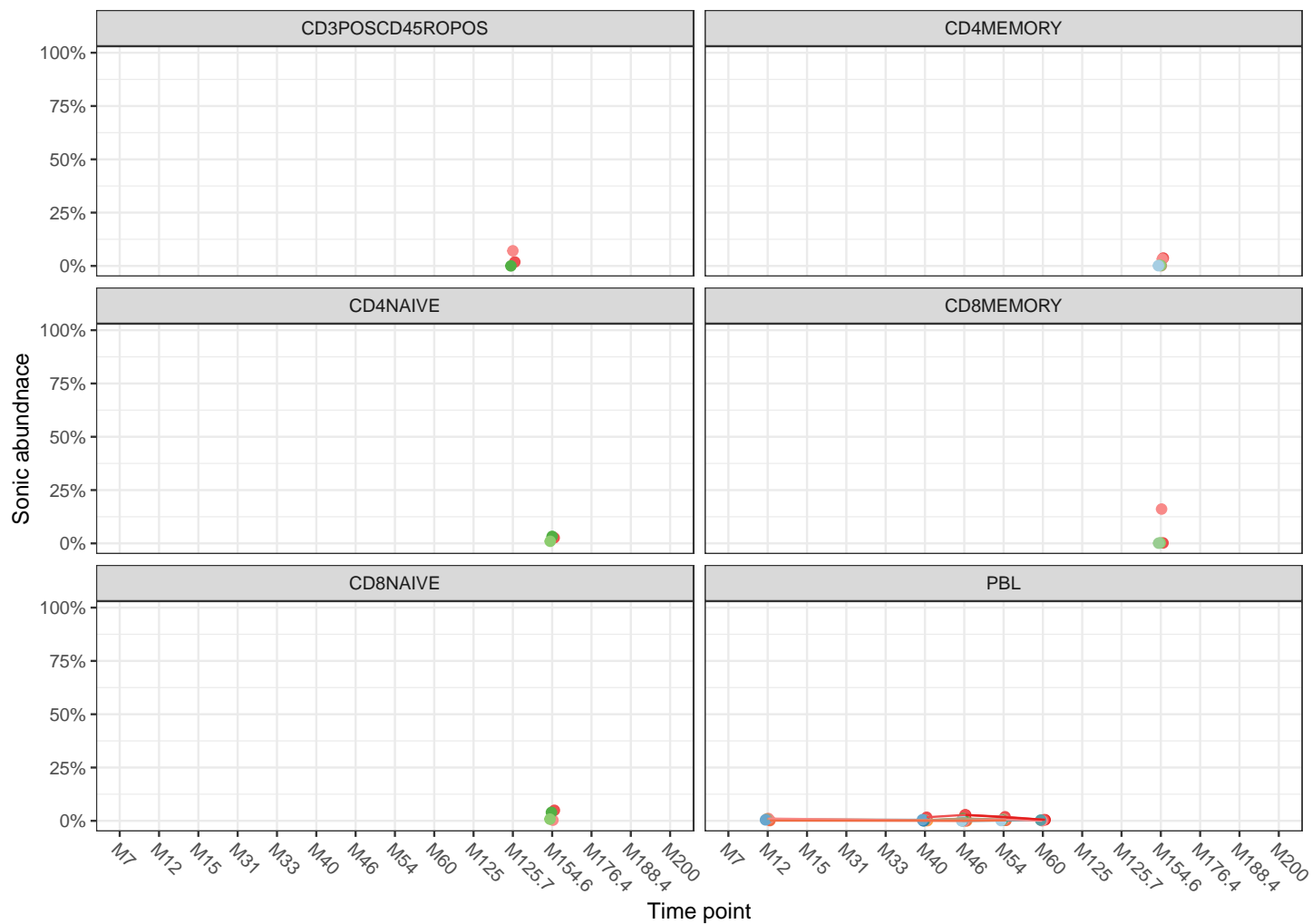
Integration sites near genes that have been associated with adverse events are of particular interest. Below are longitudinal relative abundance plots that focus on the most abundant 5 clones whose nearest genes are LMO2, IKZF1, CCND2, HMGA2, and MECOM.



### Clone

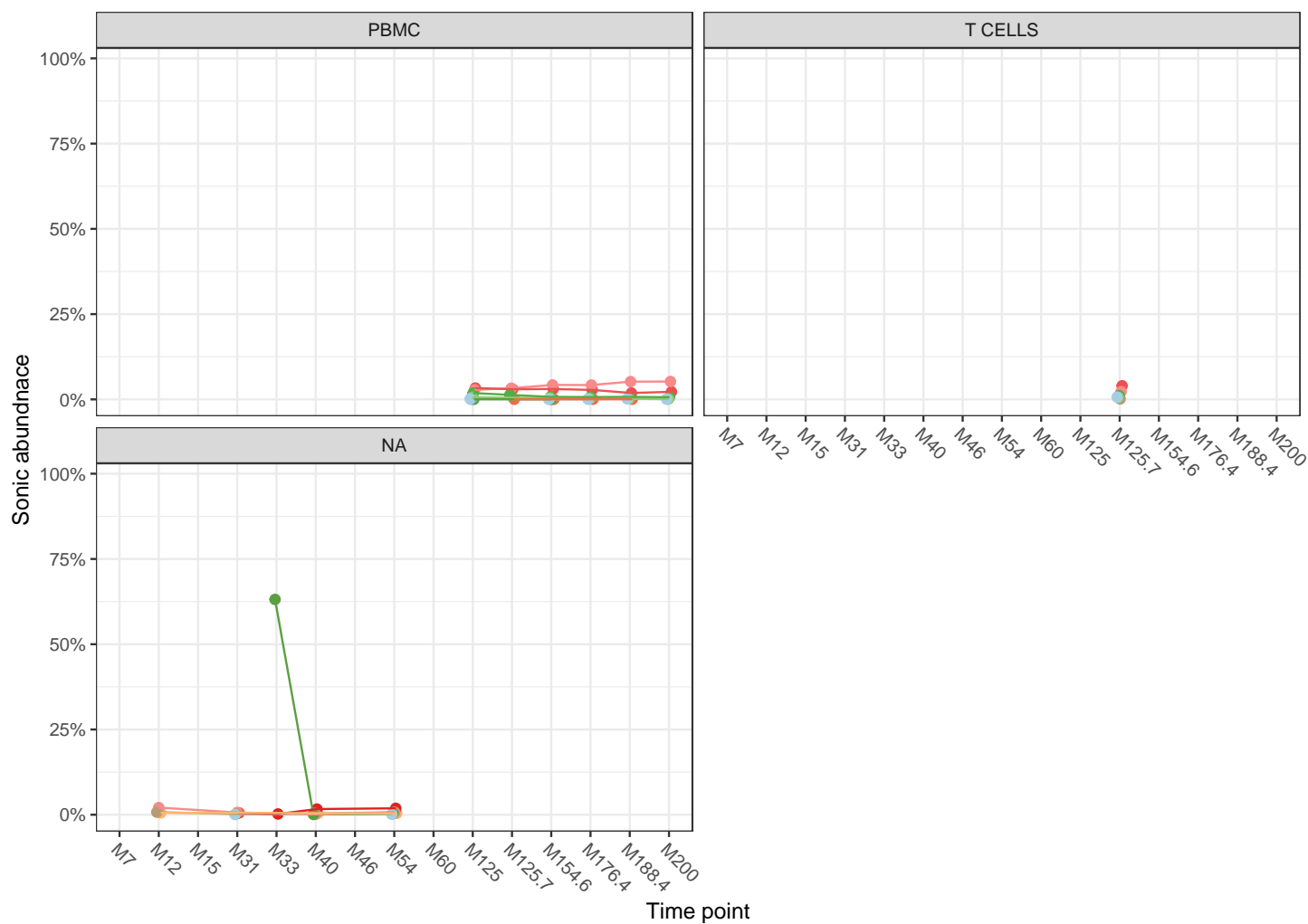
- |                            |                            |                            |                            |                            |
|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| HMGA2 *~<br>chr12+65896462 | IKZF1 *~<br>chr7+50307832  | LMO2 *~!<br>chr11-33881076 | LMO2 ~!<br>chr11+33941890  | MECOM *~<br>chr3-169657381 |
| IKZF1 *~<br>chr7-50305238  | IKZF1 *~<br>chr7+50337919  | LMO2 *~!<br>chr11-33881660 | MECOM *~<br>chr3-169250222 | MECOM *~<br>chr3-169657389 |
| IKZF1 *~<br>chr7-50305266  | LMO2 *~!<br>chr11-33868930 | LMO2 ~!<br>chr11+33931259  | MECOM *~<br>chr3-169250230 | MECOM *~<br>chr3+169249981 |





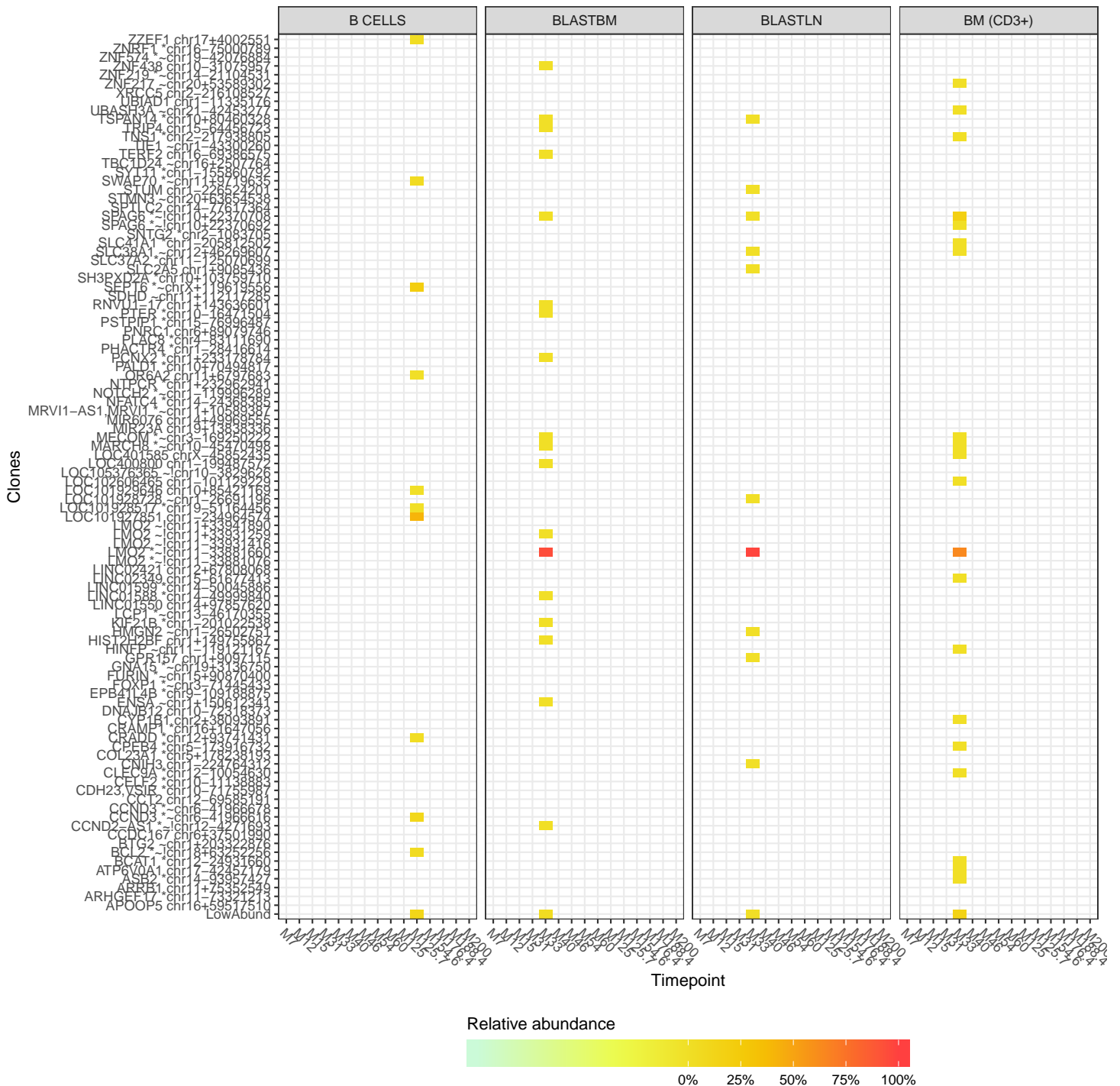
#### Clone





Sample relative abundance heatmap

Alternatively, the relative abundances of the most abundant 10 clones from each cell sampled type can be visualized as a heat map.











# What are the most frequently occurring gene types in the subject?

The word clouds below illustrate the nearest genes of the most abundant clones from each sample where the numeric ranges represent the upper and lower clonal abundances.

CD3  
M7 1:1

MLLT3 \*~  
LOC100505851  
RMDN3  
NCOR2 \*~  
LOC101929538 \*~  
CCND2-AS1 ~!  
VMP1 \*~  
ADD1 \*~  
PNRC1

CD19  
M7 1:1

ZNF574 \*~

CD56  
M7 1:1

PNRC1  
CCDC167  
TBC1D24 ~  
PSTPIP1 \*~  
SH3PXD2A \*~  
DNAJB12  
ARRB1  
CCT2  
ZNR1 \*~  
GNA15 \*~  
FBXL18

GRAN  
M7 1:1

CELF2 \*

NA  
M12 1:3

LMO2 ~!  
PLEKHA6 \*~  
NAV1, IPO9-AS1 \*~

PBL  
M12 1:3

LMO2 ~!  
SLC41A1 \*~  
BTG2 ~



PBMC  
M15 1:1

DUSP6  
NRIP1 MSRB3 \*  
AP2A2 \*~  
IFITM3 ~  
LOC101927851  
GATAD2B ~  
RCSD1 SF1 ~  
RAB18  
MUS81 ~MN1 ~  
VSIG10 \*  
SOCS2-AS1 ~

CD3  
M31 1:2

CYBB \* MIR23A  
PTP4A2 TRIP4 SPOCD1  
JAK1 \*~ LINC01599 \*  
GATAD2B \* ENSA ~  
PLEKHA6 \*  
FCER1A RNVU1-17  
DNM3 \* HMGN2 ~ PALD1 \*  
LOC101928728 ~  
TSPAN14 \* CD48  
SLC37A2 \* PP2672  
STIM1 \* SPTLC2 GBAP1 \*~  
TERF2 PDE4DIP \*~

CD3+  
M31 1:3

SLC41A1 \*  
NOTCH2 \*~  
TERF2

NA  
M31 1:3

PLEKHA6 \*  
NOTCH2 ~  
TERF2

BLASTBM  
M33 1:2

HLA-DMA \*~!  
PPARGC1A \* LINC02427  
MECOM \*~  
CDHR2 C15orf39 RBM47 \*  
IRF2 \* CCND2-AS1 \*~!  
RNF166 \* LOC101927964 \*  
ZNF438  
LMO2 \*~! PCNX2 \* GPCPD1  
UBASH3B \* RNVU1-17 NRXN1  
PPM1H \* LOC400800  
CBX2 \*~ LMO2 \*~! NRXN1  
TERF2 KIF21B \* TSPAN14 \*  
LINC01570 ENSA ~ SPAG6 \*~!  
PTER \*  
MARCH8 \*~ LOC283710  
SNX10 \* LMO2 ~! LINC01629  
LINC01588 \* UMODL1  
ADAM28 RBM11 \* SEMA6A \*  
BRINP1 \* UMODL1 \*  
PPARGC1A \* NSUN7

BLASTLN  
M33 1:2

SNHG4, MATR3, SNORA74D \*  
SMARCA1 \*  
SLC38A1 ~  
SLC2A5  
UXS1 \* LOC101928728 ~  
STK24 \* STUM  
GPR157 CNH3 UXS1  
SPAG6 \*~!  
LMO2 \*~!  
HMGN2 ~ DIRC1 ~  
TSPAN14 \*  
IRF2BPL DIRC1 ~  
MMP25, MMP25-AS1 \*  
LINC01280 ~  
HAS2-AS1 ~

BM (CD3+)  
M33 1:2

SLC38A1 ~ LINC02397 TNS1 \*  
MSI2 \*~ LOC102724050 \*  
FFAR2 LINC02349  
MORN3 \* ASB2 \* GAS7 \*~  
SLC41A1 \* PIFO LMO2 \*~!  
C1orf94 SPAG6 \*~!  
USP6NL \*  
LOC102606465  
SPAG6 \*~! FSIP1 \*  
CD82 \*~  
B2M ~ HINFP ~ KLF6 ~!  
BCAT1 \* B4GALNT4  
ATP6V0A1 AHS1 \*  
NR1D1 ~  
UBASH3A ~ CNN2 \*

NA  
M33 1:2

TCIRG1 ~ UBASH3A ~  
CRYBG2 ~ LINC02349  
FSIP1 \* BCAT1 \* KCNMA1 \*  
JAK1 \*~ HINFP ~ LTA4H  
KLF6 ~! SPAG6 \*~! KLF2  
LOC102606465 \*  
LPA \* SPAG6 \*~! LMO2 ~!  
C1orf94 \*  
LMO2 \*~! SHE \*  
ASB2 \* MTOR \*~  
ATP6V0A1 GAS7 \*~  
RUFY4 \*  
FFAR2 PLEKHA6 \* MRPL16  
FKBP5 \* LINC01226 ZNF343

NA  
M40 1:4

MECOM \*~  
SH3BP5 \*  
ATP6V0A1  
HMGN2 ~  
CPEB4 \*  
CLEC9A \*  
CYP11B1-AS1 \*  
STMN3 ~  
LOC392452 ~

PBL  
M40 1:4

MECOM \*~  
SH3BP5 \*  
ATP6V0A1  
LOC101928728 ~  
**CPEB4 \***  
CLEC9A \*  
CYP1B1  
STMN3 ~  
LOC401585

PBL  
M46 1:2

EHD1 \*~  
LMO2 \*~!  
TRIP4 IL2RA ~  
LOC101928728 ~  
APLN HMGN2 ~  
VAV3 \*~  
VAV3 \*~DOK2  
SPOCD1  
PTP4A2  
TSPAN14 \*  
LMO2 ~! CHSY1

NA  
M54 1:4

UBASH3A ~  
RNF125 \*  
CDRT15L2  
LINC01599 \*  
PBX3 \*~ NINJ2 \*  
CLEC9A \*  
PCNX2 \*  
**LMO2 ~!**  
HMGN2 ~HK2 \*  
STAT6 \*~  
LCP1 \*~  
TRIP4  
ARRDC4 \*  
LDLRAD4 \*  
CLHC1 ~

PBL  
M54 1:4

RNF125 \*  
LDLRAD4 \*  
ARRDC4 \*  
LINC01599 \*  
NINJ2 \*  
CLEC9A \*  
LOC101927851  
**LMO2 ~!**AK1  
LOC101928728 ~  
STAT6 \*~ RTN4 \*~  
CRADD \* HK2  
LCP1 \*~  
TRIP4  
CDRT15L2

PBL  
M60 1:2

LINC01588 \* TSPAN14 \*  
OR10AD1 PIK3CD ~ LINC02288  
TDRD3 LOC102606465 XYLT1 \*  
LCP1 \*~ **IKZF1 \***~ RUNX3 ~  
JAK1 \*~ MAML3 NINJ2 \*  
JAK1 \*~ MAML3 \*  
RUFY4 \*  
HIF1A-AS1  
HOXB-AS1 \*  
HLA-DMA \*~! WSB2  
SPATA48 \*~! DPEP2  
ZNRB2-AS2 \* CFHR5  
ABCA9 \* TMEM201 ~ STAT6 \*~  
LOC101927964 \* HOXB-AS3 ~

PBMC  
M125 5:581

MIR23A  
TSPAN14 \*  
**ZNF217 ~**  
SDHD ~ ZZE1  
LMO2 ~!  
ENSA ~  
FURIN \*~  
CRAMP1 \*  
TNS1 \* TERF2  
**CCND3 \***~  
ACYP2 \*  
PLAC8 \*

B CELLS  
M125.7 1:3

**CCND3 \*~**  
LOC101927851

CD3POSCD45RAPOS  
M125.7 1:4

FOXP1 \*~  
**CCND3 \***~  
LOC101927851  
PLAC8 \*

CD3POSCD45ROPOS  
M125.7 1:4

MIR23A  
TSPAN14 \*  
**ZNF217 ~**  
SDHD ~ ZZE1  
LMO2 ~!  
ENSA ~  
FURIN \*~  
CRAMP1 \*  
TNS1 \* TERF2  
**CCND3 \***~  
ACYP2 \*  
PLAC8 \*

NK CELLS  
M125.7 1:1



PBMC  
M125.7 1:4



T CELLS  
M125.7 1:4



CD4MEMORY  
M154.6 2:102



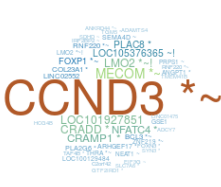
CD4NAIVE  
M154.6 1:129



CD8MEMORY  
M154.6 1:159



CD8NAIVE  
M154.6 1:161



PBMC  
M154.6 9:798



PBMC  
M176.4 3:626



PBMC  
M188.4 3:534

PBMC  
M200 3:639



## Methods

All coordinates are on human genome draft hg38.

Detailed methods can be found these publications:

- Bioinformatics. 2012 Mar 15; 28(6): 755–762.
- Mol Ther Methods Clin Dev. 2017 Mar 17; 4: 17–26.
- Mol Ther Methods Clin Dev. 2017 Mar 17; 4: 39–49.

Analysis software:

- INSPIRED v1.1 (<http://github.com/BushmanLab/INSPIRED>)

Report generation software:

- subjectReport v0.1 (<http://github.com/everettJK/geneTherapySubjectReport>)