TITLE:
Frequent Aneuploidy in Primary Human T Cells Following CRISPR-Cas9 cleavage

AUTHORS:
A.D. Nahmad[1,2]*, E. Reuveni[3]*, E. Goldschmidt[4]*, T. Tenne[5], M. Liberman[5], M. Horovitz-Fried[1,2], R. Khosravi[6], H. Kobo[7], E. Reinstein[5,8], A. Madi[4#], U. Ben-David[3#], A. Barzel[1,2#]

AFFILIATIONS:
[1]The School of Neurobiology, Biochemistry and Biophysics, Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel
[2]Dotan Center for Advanced Therapies, Tel Aviv Sourasky Medical Center and Tel Aviv University, Tel Aviv
[3]Department of Human Molecular Genetics & Biochemistry, Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
[4]Department of Pathology, Faculty of Medicine, Tel Aviv University, Israel
[5]Medical Genetics Institute, Meir Medical Center, Kfar-Saba, Israel
[6]Single-Cell Genomics Core, Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
[7]Genomics Research Unit, Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel
[8]Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
* These authors contributed equally
# These authors contributed equally
Corresponding author email: adibarzel@gmail.com, ubendavid@tauex.tau.ac.il, asafmadi@tauex.tau.ac.il

SUMMARY:
Multiple clinical trials use site-specific nucleases to disrupt T cell receptor (TCR) genes for allogeneic T cell therapy[1–6]. Here, we applied single-cell RNA sequencing to human T cells, transfected with CRISPR-Cas9, using guide RNAs targeting genes for TCR chains and programmed cell death protein 1. 4 days post-transfection, we found a loss of chromosome 14, harboring the TCRα locus, in up to 9% of the cells, and a chromosome 14 gain in up to 1.4% of the cells. Chromosome 7, harboring the TCRβ locus, was truncated in 9.9% of the cells. Aberrations were validated using fluorescence *in-situ* hybridization and digital droplet PCR. Aneuploidy was associated with reduced proliferation, induced p53 activation and cell death. However, at 11 days post-transfection, 0.9% of T cells still had a chromosome 14 loss. Aneuploidy and chromosomal truncations are thus frequent outcomes of CRISPR-Cas9 cleavage that need be monitored and minimized in clinical protocols.

MAIN:
Engineered T cells are revolutionizing cancer immunotherapy. Notably, five treatments entailing chimeric antigen receptor (CAR) T cells have received regulatory approval in different B cell malignancies[7]. At the same time, hundreds of ongoing clinical trials, for multiple additional indications, are applying different variations of CARs or engineered TCRs in an attempt to improve efficacy, safety and scalability[1]. Importantly, the approved CAR T therapies all entail the cumbersome, expensive and time-consuming manipulation of autologous T cells. Many current efforts are therefore directed at facilitating the engineering and banking of allogeneic CAR/TCR-T cell products. Importantly, the endogenous TCR of engineered allogeneic T cells has to be disrupted in order to prevent graft vs. host disease (GVHD) and this is most often achieved using site-specific nucleases such as meganucleases[2], TALENs[3,4], megaTALs and CRISPR-Cas9[5]. Cleaving the TCR loci can further serve for the directed integration of CAR and engineered TCR genes, allowing uniform expression, enhancing T cell potency and delaying exhaustion[8–11]. Finally, additional genes, beyond the TCR chains, may be disrupted in order to improve T cell function[12], confer drug resistance[13], or avoid checkpoint inhibition[14].

Site-specific nucleases can be highly potent[15], but they are associated with a variety of undesired outcomes. Cas9 can be immunogenic *in vivo*, eliciting both humoral and cellular responses[16]. We recently demonstrated that Cas9 activates the p53 pathway and selects for p53-inactivating mutations[17]. Off-target cleavage can in turn be reduced, but not eliminated[18], and small on-target insertions and deletions (InDels) are common, even when providing a donor template for gene correction or insertion. Importantly, CRISPR-Cas9 cleavage can also lead to gross chromosomal aberrations. In particular, CRISPR-Cas9 cleavage leads to large deletions in early mouse[19,20] and human embryos[21] as well as in embryonic stem cells and induced pluripotent stem cells[22,23]. Moreover, chromosomal truncations were reported following CRISPR-Cas9 cleavage in mouse and human cell lines[24,25], and entire chromosome loss has resulted from CRISPR-Cas9 cleavage in human embryos[26]. Recently, chromothripsis, defined

as multiplicity of gross genomic rearrangements in a one-off cellular crisis, was inferred in mobilized human CD34+ cells following CRISPR-Cas9 cleavage of the BCL11A gene in a clinically relevant setting[27,28], with possible ramifications for the development of gene therapies for hemoglobinopathies. Finally, in the first U.S. clinical trial involving CRISPR-Cas9, Stadtmauer & Fraietta et al. aimed at disrupting the TCR and programmed cell death protein 1 (PDCD1) loci for allogeneic T cell therapy and reported the detection of chromosomal translocations whose frequency decreased over time after infusion into patients[6].

Here, we use the same guide RNA sequences, as did Stadtmauer & Fraietta et al.[6], to target the TCR and PDCD1 loci in primary human T cells with CRISPR-Cas9. Using a novel unbiased approach, based on scRNAseq, corroborated by Fluorescence In-Situ Hybridization (FISH) and digital droplet PCR (ddPCR) analyses, we detect frequent aneuploidy and truncations of the chromosomes harboring the targeted loci.

We hypothesized that double-strand DNA breaks, induced by site-specific nucleases, may sometime result in chromosomal truncations and aneuploidy due to failure of DNA repair. We further conjectured that such adverse outcomes could be detected and monitored, in a high throughput manner, by following concerted changes in gene expression along chromosomes. In order to analyze chromosomal truncations and aneuploidy in a clinically relevant setting, we first performed CRISPR-Cas9 ribonucleoprotein (RNP) electroporation into primary human T cells using a single guide RNA (gRNA) targeting the TCRα locus (Fig. 1A). Importantly, we used the same gRNA sequence as used clinically in T cells by Stadtmauer & Fraietta et al.[6]. As a control, we used an irrelevant gRNA[15], with no matching target in the human genome. The TCRα locus was successfully targeted in more than 50% of the cells (Fig. 1B-C, Extended Data Fig. 1). 4 days post-transfection, samples were subjected to scRNAseq (Methods). The transcriptional downregulation of TCRα was confirmed (Fig. 1D), and the transcriptional landscapes of the cells were used to infer copy number alterations using a well-established method (inferCNV)[29,30]. Strikingly, 5.3% of the cells in the TCRα targeted sample had expression patterns indicating a chromosome 14 loss (Fig. 1E-F). In addition, among these cells, chromosome 14 was an extreme outlier in the mean number of genes with undetected expression (p value< 0.0001, Extended Data Fig. 2). Notably, a chromosome 14 q-arm truncation at the TCRα locus is functionally equivalent to a whole-chromosome aneuploidy, because almost all protein-coding genes on chromosome 14 are coded on the q-arm, distal to the TCRα locus (Fig. 1A). Interestingly, a smaller fraction of the cell population, entailing 0.5-0.9% of the cells, had an apparent gain of chromosome 14 (Fig 1E-F). This functional chromosome 14 gain is assumed to result from mis-segregation of an acentromeric q-arm[27]. In particular, 100 chromosome 14 genes were underexpressed and 107 chromosome 14 genes (out of 270 chromosome 14 genes expressed in T cells) were overexpressed in the cells categorized as harboring a chromosome 14 loss or gain, respectively (Extended Data Fig. 2, Supplementary Table 1). To set a clear statistical threshold for gains and losses, only cells with inferCNV scores >2 standard deviations from the mean score of the population were defined as having a chromosomal aberration (Methods).

In order to corroborate the scRNAseq results, we next performed FISH assays employing red and green probes proximal and distal to the TCRα locus on chromosome 14, respectively. Expectedly, most examined nuclei had two foci of co-localized red and green signals (often co-detected as a yellow signal). Importantly, in three independent FISH iterations, signal patterns indicative of aneuploidy were observed significantly more frequently in TCRα-targeted cells than in control cells (Fig. 2A-B, Extended Data Fig. 3). These cells had either a single focus of co-localized signals or one focus of co-localized signals and another red focus, interpreted as corresponding to a truncation distal to the targeted gene. In particular, the 4% excess rate of these FISH patterns in TCRα-targeted cells, compared to control cells, is in agreement with the scRNAseq analysis. The rate of chromosome 14 breakage and its incomplete repair can further be monitored using ddPCR[31] (Fig. 2C). DNA was collected at different time points after CRISPR-Cas9 RNP electroporation, in order to assess the dynamics of dislinkage between the two sides of the break. As early as 3 hrs after electroporation, the rate of dislinkage has peaked at more than 45% (Fig. 2D-E). Dislinkage rates were greatly reduced by 24 hrs, but remained significantly higher in the TCRα-targeted sample, compared to the control sample, for at least 11 days, indicating incomplete repair. This ddPCR analysis alone cannot discriminate between transient breaks, translocations and deletions of various sizes, including asymmetrical deletions that span primer/probe binding sites. Still, the ddPCR results are consistent with the results from the scRNAseq and FISH analyses. Moreover, the ddPCR assay provides a scalable and cost-effective means to track the dynamics of breakage, repair and aberration. Importantly, we detected similar rates of aneuploidy and truncations when targeting the TCRα locus with CRISPR-Cas9 and an alternative gRNA or with the alternative nuclease CRISPR-Cas12a (Cpf1[32], Extended Data Fig. 4).

We next used t-distributed stochastic neighbor embedding (t-SNE) analysis in order to cluster T cells according to their transcriptional signatures (Fig. 3A, Extended Data Fig. 5). We found a similar distribution, among the clusters, of T cells transfected with CRISPR-Cas9 RNPs entailing either the TCRα-targeting gRNA or an irrelevant

gRNA with no human target (Fig. 3B). Moreover, among cells receiving the TCRα-targeting gRNA, neither cells with chromosome 14 loss nor cells with chromosome 14 gain were found to be considerably enriched in any T cell subset (Fig. 3C). This implies unbiased incidence of aneuploidy, as well as similar selection forces acting on aneuploid cells of various T cell states. Importantly, however, a gene set enrichment analysis (GSEA) detected significant differences in global gene expression patterns between the groups, beyond the expected reduced expression of the genes encoded on chromosome 14 (Extended Data Fig. 6). Specifically, the transcriptional signatures enriched in T cells with a chromosome 14 loss or gain reflected reduced expression of genes associated with the cell cycle and with various metabolic pathways and increased expression of genes associated with p53 pathway activation and with apoptosis (Fig. 3D-F and Supplementary Table 2). Moreover, a larger fraction of cells with a chromosome 14 loss were found in the G1 cell cycle phase compared to treated cells without an aberration (Fig. 3G). We thus hypothesized that longer culturing will reduce the prevalence of aneuploid cells, and we repeated the above experiments comparing the effects of 4 vs. 11 days of T cell culturing following the CRISPR-Cas9 RNP electroporation (Extended Data Fig. 7). While the aberration rates at day 4 were highly consistent with the results of the previous experiment (Extended Data Fig. 7), scRNAseq at day 11 showed a marked reduction in the fraction of aneuploid cells (Fig. 3H, Extended Data Fig. 7). The reduced proportion of aneuploid cells is not due to a fitness cost of TCR disruption alone, as the fraction of $CD3^-$ cells remained high (Extended Data Fig.7). Importantly, 0.9% of the T cells were found to have a chromosome 14 loss even after 11 days of culture (Fig. 3H, Extended Data Fig. 7). The persistence of chromosomal aberrations at day 11 was further corroborated by ddPCR (Fig. 2D-E, Extended Data Fig. 7), while the sensitivity of FISH analysis for the detection of low aneuploidy rates was insufficient (Extended Data Fig. 7). Finally, both cell cycle phase analysis and GSEA indicated that the aneuploid cells continued to have a reduced fitness at day 11, in comparison to normal cells (Extended Data Fig. 7, Extended Data Fig. 8, Supplementary Table 3). Cumulatively, these results indicate that prolonged T cell cultures reduce the risk of transplanting aneuploid cells, but underscore the need for additional mitigation strategies.

Targeting different loci concomitantly with several gRNAs may aggravate the risk for chromosomal aberrations. Indeed, Stadtmauer & Fraietta et al. identified multiple translocations between cleaved chromosomes when co-delivering 3 gRNAs targeting both TCR loci as well as the PDCD1 gene[6]. We next used the same gRNA combination for CRISPR-Cas9 RNP electroporation of primary human T cells (Fig. 4A). We verified cleavage efficiency (Extended Data Fig. 9) and performed scRNAseq in order to analyze gene expression patterns and infer copy number changes along the targeted chromosomes. The transcriptional downregulation of TCRα, TCRβ and PDCD1 was confirmed (Fig. 4B, Extended Data Fig. 9). Importantly, in this experiment, we found expression patterns indicative of a chromosome 14 loss and gain in 9.0% and 1.4% of the cells, respectively (Fig. 4C-D, Extended Data Fig. 10). Strikingly, we further found as many as 9.9% of treated cells to be characterized with chromosome 7 truncations, entailing the TCRβ locus (Fig. 4C-D). Importantly, the extent of the detected aberrations is in good agreement with the chromosomal position of the cleavage sites: cleaving the TCRα locus, near the chromosome 14 centromere, leads to loss of the entire arm and to functional whole-chromosome aneuploidy; cleaving the TCRβ locus, in the middle of the chromosome 7 q-arm, leads to the expected truncations; and cleaving the PDCD1 gene, which resides near the chromosome 2 q-arm telomere, expectedly has a lesser effect on copy number and gene expression (Extended Data Fig. 11). Finally, a gene set enrichment analysis confirmed the transcriptional differences observed in our first experiment in cells that have lost or gained chromosome 14, and identified very similar pathway enrichments in the cells with a chromosome 7 truncation (Fig. 4E-G, Extended Data Fig. 12 and Supplementary Table 4). We did not culture the cells for more than 4 days following multiplex editing, but our data implies that such cultures would show a negative selection against truncated and aneuploid cells, in concordance with cells edited only at the TRAC locus (Fig. 3H, Extended Data Fig. 7).

Aneuploidy is associated with most types of cancers[33]. Of note, chromosome 14 loss is frequently found in early-onset colon cancer[34], and is among the most commonly affected chromosomal regions in ovarian cancer[35]. Chromosome 14 monosomy, trisomy and tetrasomy are also frequently found in meningioma[36]. Similarly, high grade glioma[37] and adult T cell leukemia/lymphoma[38] are often characterized by chromosome 7 trisomy, while chromosome 14 monosomies and tetrasomies are not frequent events in T cell malignancies. To date, no serious adverse events in clinical trials have been attributed to genome editing. However, our findings raise the possibility that Cas9-induced aneuploidy and chromosomal truncations might be associated with increased risk of tumorigenesis. Importantly, the methods we described were optimized for the detection of aneuploidy and large

truncations, therefore underestimating the overall occurrence of genetic alterations. While detected in some cases (Extended Data Fig. 3), translocations and focal deletions may better be assessed using complementary technologies[6].

We describe high aberration rates induced by different nucleases and different gRNAs. It is possible that optimizing the electroporation conditions, or using a clinical-grade electroporator, may reduce the aberration frequency, by affecting the re-cutting rate of repaired DSBs. Better yet, aberration rates can be reduced by using base editors instead of nucleases[39,40]. In the liver, gene insertion by homologous recombination can be achieved without nucleases[41–43]. However, in T cells, nucleases may still be required for site-specific gene insertion of CAR and engineered TCR genes into TCR loci in order to enhance T cell potency and delay exhaustion[8–10]. Further studies are required in order to determine whether sorting for expression of the targeted receptor cassette will help reduce the rate of cells with nuclease-induced aberrations. Alternatively, endogenous genes that are differentially expressed in aberrant T cells may be used for cell sorting prior to adoptive transfer. As a first step towards this aim, we compiled a list of genes whose expression significantly differed between the cells that have lost chromosome 14 and those that have not (Supplementary Table 1), and manually curated the list to identify cell surface proteins. However, our initial attempts to sort against markers of aneuploidy only mildly reduced the aberration rate (Extended Data Fig. 13). Finally, our gene set enrichment analysis suggests that aneuploidy provides fitness disadvantage to the T cells *in vitro* (Fig. 3D-H, Supplementary Table 2-4, Extended Data Fig.7), in line with multiple studies that showed the negative effect of aneuploidy on cell survival and proliferation[33,44,45]. Indeed, prolonged culturing has reduced the prevalence of aneuploid cells, and could be considered in clinical protocols, while also weighing the effects on cell activation, differentiation and exhaustion.

In conclusion, inferring copy number alterations from scRNAseq data is an unbiased high-throughput method to identify nuclease-induced chromosomal aberrations. Our discovery of frequent aneuploidy and chromosomal truncations in human T cells targeted using CRISPR-Cas9 and clinical gRNA sequences, highlights potential oncogenic risks and underscores the need for mitigation strategies in order to allow the safe application of nucleases in adoptive T cell transfer and beyond.

REFERENCES:

1. Weber, E. W., Maus, M. V. & Mackall, C. L. The Emerging Landscape of Immune Cell Therapies. *Cell* **181**, 46–62 (2020).
2. MacLeod, D. T. *et al.* Integration of a CD19 CAR into the TCR Alpha Chain Locus Streamlines Production of Allogeneic Gene-Edited CAR T Cells. *Mol. Ther.* **25**, 949–961 (2017).
3. Philip, L. P. B. *et al.* Multiplex genome-edited T-cell manufacturing platform for 'off-the-shelf' adoptive T-cell immunotherapies. *Cancer Res.* **75**, 3853–3864 (2015).
4. Qasim, W. *et al.* Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Sci. Transl. Med.* **9**, 1–10 (2017).
5. Osborn, M. J. *et al.* Evaluation of TCR Gene Editing Achieved by TALENs , CRISPR / Cas9 , and megaTAL Nucleases. *Mol. Ther.* **24**, 570–581 (2016).
6. Stadtmauer, E. A. *et al.* CRISPR-engineered T cells in patients with refractory cancer. *Science (80-.).* **367**, (2020).
7. Braendstrup, P., Levine, B. L. & Ruella, M. The long road to the first FDA-approved gene therapy: chimeric antigen receptor T cells targeting CD19. *Cytotherapy* **22**, 57–69 (2020).
8. Eyquem, J. *et al.* Targeting a CAR to the TRAC locus with CRISPR/Cas9 enhances tumour rejection. *Nature* **543**, 113–117 (2017).
9. Roth, T. L. *et al.* Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature* **559**, 405–409 (2018).
10. Schober, K. *et al.* Orthotopic replacement of T-cell receptor α- and β-chains with preservation of near-physiological T-cell function. *Nat. Biomed. Eng.* (2019). doi:10.1038/s41551-019-0409-0
11. Mansilla-Soto, J. *et al.* HLA-independent T cell receptors for targeting tumors with low antigen density. *Nat. Med.* (2022). doi:10.1038/s41591-021-01621-1
12. Shifrut, E. *et al.* Genome-wide CRISPR Screens in Primary Human T Cells Reveal Key Regulators of Immune Function. *Cell* **175**, 1958-1971.e15 (2018).
13. Valton, J. *et al.* A Multidrug-resistant Engineered CAR T Cell for Allogeneic Combination Immunotherapy. *Mol. Ther.* **23**, 1507–1518 (2015).
14. Rupp, L. J. *et al.* CRISPR / Cas9-mediated PD-1 disruption enhances anti-tumor efficacy of human chimeric

antigen receptor T cells. *Sci. Rep.* **737**, 1–10 (2017).
15. Nahmad, A. D. *et al.* Engineered B cells expressing an anti-HIV antibody enable memory retention, isotype switching and clonal expansion. *Nat. Commun.* 1–10 (2020). doi:10.1038/s41467-020-19649-1
16. Charlesworth, C. T. *et al.* Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* **25**, (2019).
17. Enache, O. M. *et al.* Cas9 activates the p53 pathway and selects for p53-inactivating mutations. *Nat. Genet.* **52**, 662–668 (2020).
18. Lazzarotto, C. R. *et al.* CHANGE-seq reveals genetic and epigenetic effects on CRISPR–Cas9 genome-wide activity. *Nat. Biotechnol.* (2020). doi:10.1038/s41587-020-0555-7
19. Adikusuma, F. *et al.* Large deletions induced by Cas9 cleavage. *Nature* **560**, E8–E9 (2018).
20. Papathanasiou, S. *et al.* Whole chromosome loss and genomic instability in mouse embryos after CRISPR-Cas9 genome editing. *Nat. Commun.* **12**, 5855 (2021).
21. Alanis-lobato, G., Zohren, J., Mccarthy, A., Fogarty, N. M. E. & Kubikova, N. Frequent loss-of-heterozygosity in CRISPR-Cas9 – edited early human embryos. *Proc. Natl. Acad. Sci. USA* (2021). doi:10.1073/pnas.2004832117
22. Weisheit, I. *et al.* Detection of Deleterious On-Target Effects after HDR-Mediated CRISPR Editing. *Cell Rep.* **31**, 107689 (2020).
23. Boutin, J. *et al.* CRISPR-Cas9 globin editing can induce megabase-scale copy-neutral losses of heterozygosity in hematopoietic cells. *Nat. Commun.* **12**, 1–12 (2021).
24. Przewrocka, J., Rowan, A., Rosenthal, R., Kanu, N. & Swanton, C. Unintended on-target chromosomal instability following CRISPR/Cas9 single gene targeting. *Ann. Oncol.* **31**, 1270–1273 (2020).
25. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, (2018).
26. Zuccaro, M. V. *et al.* Allele-Specific Chromosome Removal after Cas9 Cleavage in Human Embryos. *Cell* **183**, 1650-1664.e15 (2020).
27. Leibowitz, M. L. *et al.* Chromothripsis as an on-target consequence of CRISPR–Cas9 genome editing. *Nat. Genet.* **53**, 895–905 (2021).
28. Urnov, F. D. CRISPR–Cas9 can cause chromothripsis. *Nat. Genet.* **53**, 765–769 (2021).
29. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science (80-.).* **344**, 1396–1402 (2014).
30. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science (80-.).* **352**, 189–196 (2016).
31. Puig, M. *et al.* Determining the impact of uncharacterized inversions in the human genome by droplet digital PCR. *Genome Res.* **30**, 724–735 (2020).
32. Zetsche, B. *et al.* Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **163**, 759–771 (2015).
33. Ben-David, U. & Amon, A. Context is everything: aneuploidy in cancer. *Nat. Rev. Genet.* **21**, 44–62 (2020).
34. Mourra, N. *et al.* High Frequency of Chromosome 14 Deletion in Early-Onset Colon Cancer. *Dis. Colon Rectum* **50**, 1881–1886 (2007).
35. Bandera, C. A. *et al.* Deletion Mapping of Two Potential Chromosome 14 Tumor Suppressor Gene Loci in Ovarian Carcinoma. *Cancer Res.* **57**, 513–516 (1997).
36. Tabernero, M. D. *et al.* Characterization of Chromosome 14 Abnormalities by Interphase In Situ Hybridization and Comparative Genomic Hybridization in 124 Meningiomas. *Anat. Pathol.* 744–751 (2005). doi:10.1309/D7U997XD2PHBCQCN
37. Lopez-Gines, C. *et al.* Association of chromosome 7, chromosome 10 and EGFR gene amplification in glioblastoma multiforme. *Clin. Neuropathol.* **24**, 209–218 (2005).
38. Kamada, N. *et al.* Chromosome Abnormalities in Adult T-Cell Leukemia/Lymphoma: A Karyotype Review Committee Report1. *Cancer Res.* **52**, 1481–1493 (1992).
39. Webber, B. R. *et al.* Highly efficient multiplex human T cell engineering without double-strand breaks using Cas9 base editors. *Nat. Commun.* (2019). doi:10.1038/s41467-019-13007-6
40. Anzalone, A. V., Koblan, L. W. & Liu, D. R. Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* **38**, 824–844 (2020).
41. Barzel, A. *et al.* Promoterless gene targeting without nucleases ameliorates haemophilia B in mice. *Nature* **517**, 360–364 (2015).
42. Porro, F. *et al.* Promoterless gene targeting without nucleases rescues lethality of a Crigler-Najjar syndrome

mouse model. *EMBO Mol. Med.* **9**, 1346–1355 (2017).
43. Chandler, R. J. *et al.* Promoterless, Nuclease- Free Genome Editing Confers a Growth Advantage for Corrected Hepatocytes in Mice With Methylmalonic Acidemia. *Hepatology* **73**, 2223–2237 (2021).
44. Rutledge, S. D. *et al.* Selective advantage of trisomic human cells cultured in non- standard conditions. *Sci. Rep.* **6**, (2016).
45. Sheltzer, J. M. *et al.* Single-chromosome Gains Commonly Function as Tumor Suppressors. *Cancer Cell* **31**, 240–255 (2017).
46. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data Resource Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).
47. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
48. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).
49. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
50. Brinkman, E. K., Chen, T., Amendola, M. & Van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, 1–8 (2014).

FIGURES:

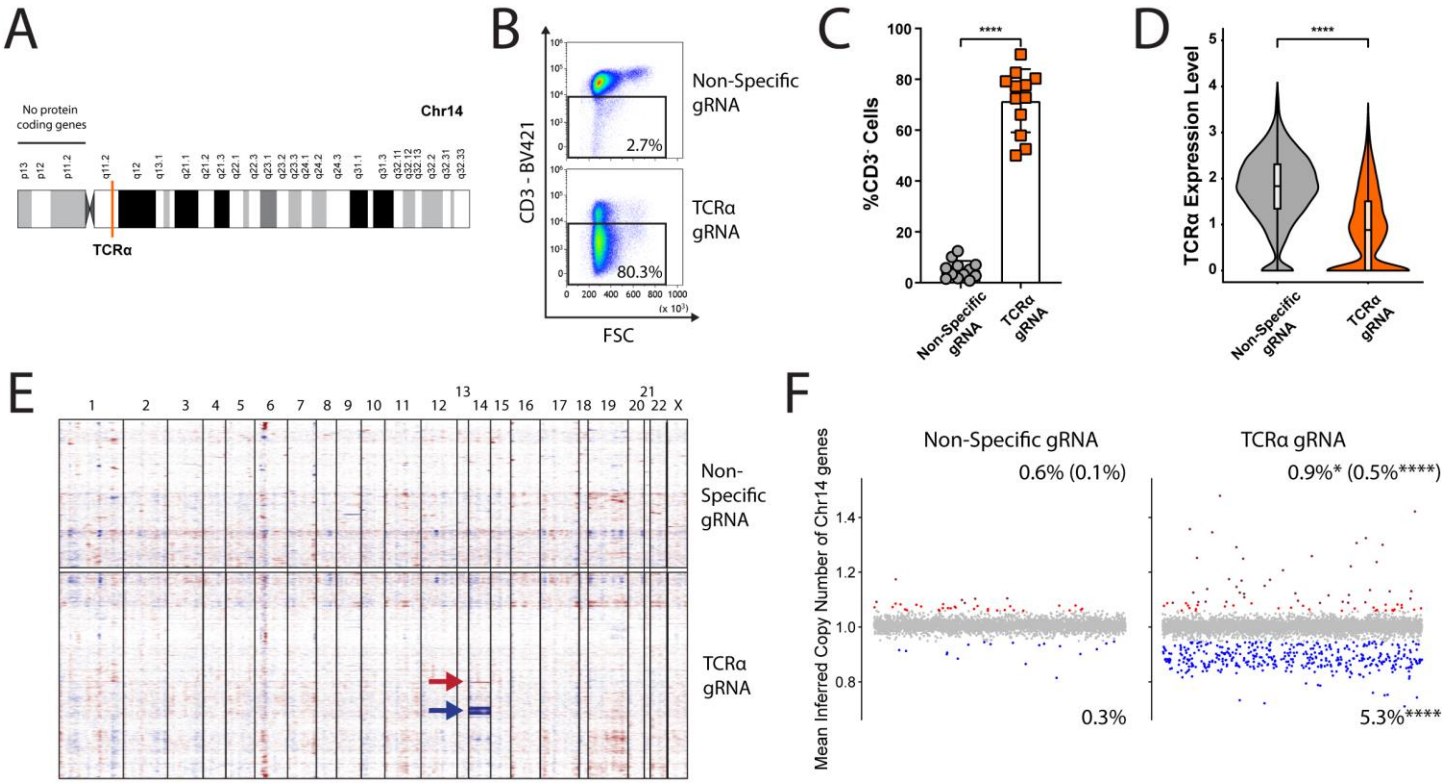


Figure 1: Targeting the TCRα locus using CRISPR-Cas9 leads to chromosome 14 aneuploidy. **A.** Schematic depiction of human chromosome 14 (Chr14). The target locus is indicated in orange. **B.** Flow cytometry example of TCR ablation in primary human T cells 4 days following CRISPR-Cas9 RNP electroporation. Cells were electroporated with Cas9 and either a non-specific gRNA or a TCRα-targeting gRNA. **C.** Quantification of B. Each dot represents an independent experiment. Mean value, standard deviation and individual experiments are indicated. n=12, ****, p<0.0001, two-sided unpaired Wilcoxon test. **D.** A reduction in TCRα expression is evident in the scRNAseq. The violin plots correspond to TCRα expression level in cells treated with either a non-specific gRNA or the TCRα-targeting gRNA. Upper and lower boundaries, median and quartiles are indicated. ****, p<0.0001, two-sided unpaired Wilcoxson test. **E.** A heat map depicting gene copy numbers inferred from scRNAseq analysis following treatment with either a non-specific or TCRα-targeting gRNAs. Each line represents an individual cell. The color coding indicates an increase (red) or decrease (blue) in copy number of genes along the chromosomes, ordered in columns. Arrows indicate gain (red) or loss (blue) of chromosome 14. **F.** Each dot represents the mean inferred copy

number of genes coded on chromosome 14 in each cell treated with a non-specific gRNA (left) or a TCRα-targeting gRNA (right). Cells are marked with dots spread along the x-axis. The dots are colored red and blue when corresponding to cells with a chromosome 14 gain or loss respectively, if their mean inferred gene copy number is >2 standard deviations (blue and red), or 3 standard deviations (dark red), from the population's mean. n=6055 and 7700 for non-specific (left) or TCRα-targeting (right) gRNA treated cells, respectively. *, p=0.0148 and ****, p<0.0001 for Fisher's exact test comparing chromosome 14 gain or loss, respectively, between cells treated with the TCRα gRNA and cells treated with a non-specific gRNA, with a cutoff of 2 standard deviations. ****, p<0.0001 for Fisher's exact test comparing chromosome 14 gain between cells treated with the TCRα gRNA and cells treated with a non-specific gRNA, with a cutoff of 3 standard deviations, for which the frequency of cells is indicated in parentheses.

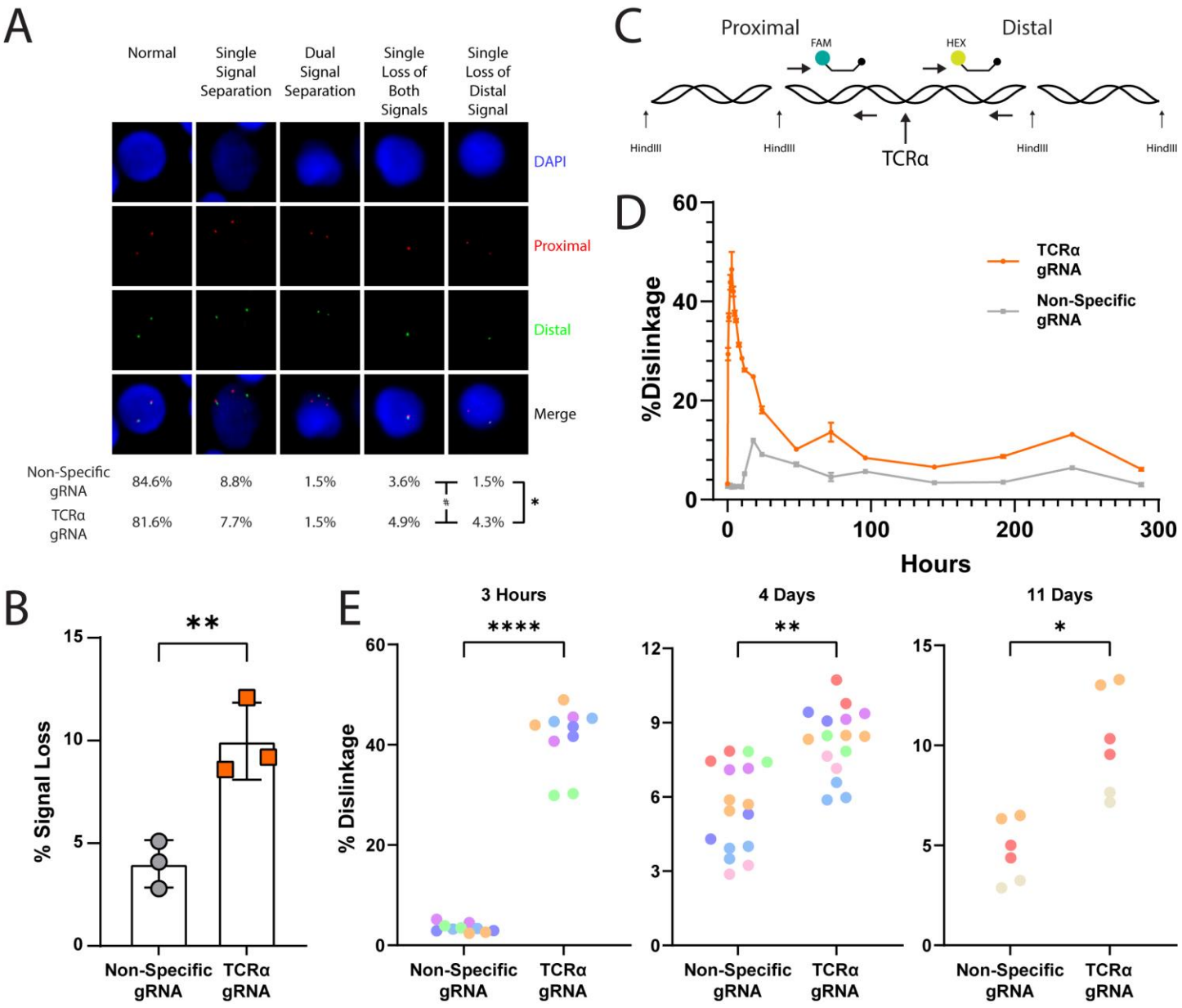


Figure 2: FISH and ddPCR analyses of Chromosome 14 aberrations. **A.** FISH analysis of cells treated with either a non-specific or a TCRα-targeting gRNA. In the pictures are examples of all signal patterns found among analyzed cells. The frequency of each signal pattern, among cells electroporated with the TCRα-targeting gRNA or a non-specific gRNA, is presented below the pictures. n = 273 and n = 506 for cells electroporated with a non-specific and TCRα-targeting gRNA, respectively. *, p=0.0129 Fisher's exact test for loss of distal signal comparison. #, p=0.0486 Fisher's exact test for loss of either both signals or only the distal signal. **B.** Quantification of A. from 3 independent experiments, 4 days following electroporation. Signal loss designates loss of either both signals or only the distal signal. Mean and Standard Deviations are indicated. **, p=0.0093 two-sided unpaired t-test. **C.** Schematic depiction

of the ddPCR experiment. Genomic DNA is sheared using the HindIII restriction enzyme. The two sets of primers and probe are on the same restriction fragment and are separated only upon CRISPR-Cas9 cleavage. **D.** Dislinkage, calculated based on the rate of single and dual color droplets (see methods, is analyzed at different time points following CRISPR-Cas9 electroporation with either a TCRα-targeting gRNA (orange) or a non-specific (gray) gRNA from a single experiment. Mean and standard deviation for technical replicates are indicated. **E.** Dislinkage at either 3 hours (left), 4 days (middle) or 11 days (right) following CRISPR-Cas9 electroporation. Each color represents a different human donor. Each dot represents a technical replication. n=3-7, ****, p<0.0001 and *, p=0.0085, *, p=0.0474 two-sided unpaired t-test.

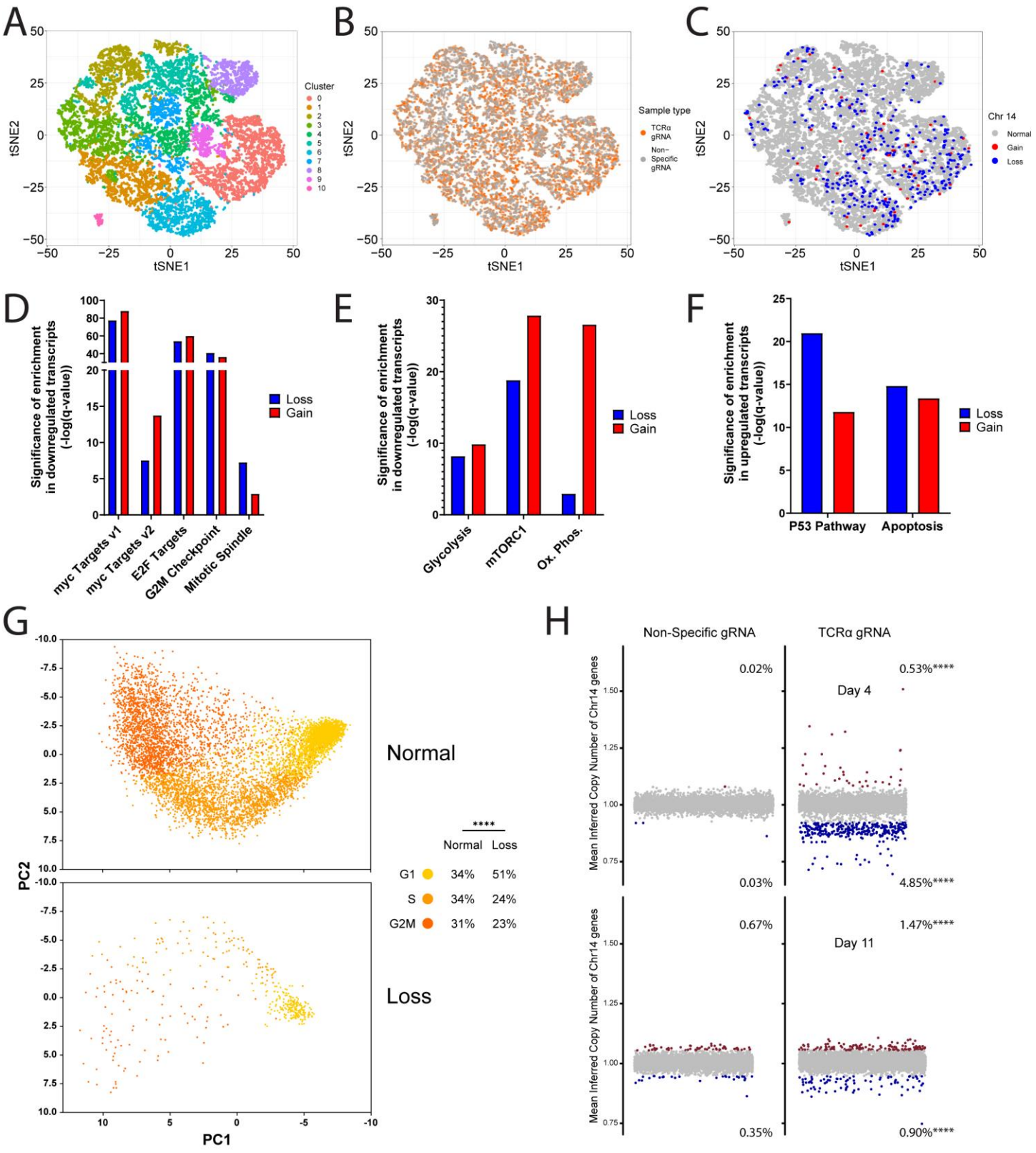


Figure 3: Global gene expression analysis following genome editing. **A.** scRNAseq data from primary human T cells showing 13,755 cells as individual dots. Data is displayed by tSNE with each cluster colored differently. **B.** Distribution of cells electroporated with a non-specific (gray) or TCRα-targeting (orange) gRNA among the clusters. **C.** Distribution of cells with a chromosome 14 loss (blue) or gain (red) among the clusters. **D-F.** Gene set enrichment analysis for 'Hallmark' gene sets associated with cell cycle (D), metabolism (E), p53 expression and apoptosis (F) in cells with chromosome 14 loss (blue) and or gain (red). **G.** Principal component (PC) analysis of cell cycle phase among cells treated with a TCRα-targeting gRNAs and characterized, as in Fig. 1F, as having a chromosome 14 loss (bottom) or not (normal, up). Legend and fraction of cells in each cell cycle phase is indicated on the right. ****, $p<0.0001$ Fisher's exact test for G1 phase compared to S/G2M phases. n = 7289 and n = 411 for cells without or with a chromosome 14 loss, respectively. **H.** Each dot represents the mean inferred copy number of genes coded on chromosome 14 in each cell treated with a non-specific gRNA (left) or a TCRα-targeting gRNA (right). Cells are marked with dots spread along the x-axis. The dots are colored red and blue when corresponding to cells with a chromosome 14 gain or loss respectively, if their mean inferred gene copy number is >3 standard deviations (dark blue and dark red), from the population's mean. For Day 4, n=8642 and 6592 for cells transfected with a non-specific (left) or TCRα-targeting (right) gRNA, respectively. For Day 11, n=8642 and 9197 for cells transfected with a non-specific (left) or TCRα-targeting (right) gRNA, respectively. The expression levels in cells treated with a non-specific gRNA 4 days after CRISPR-Cas9 electroporation served as a reference in both analyses. ****, $p<0.0001$ for Fisher's exact test comparing chromosome 14 gain or loss between cells treated with the TCRα gRNA and cells treated with a non-specific gRNA.

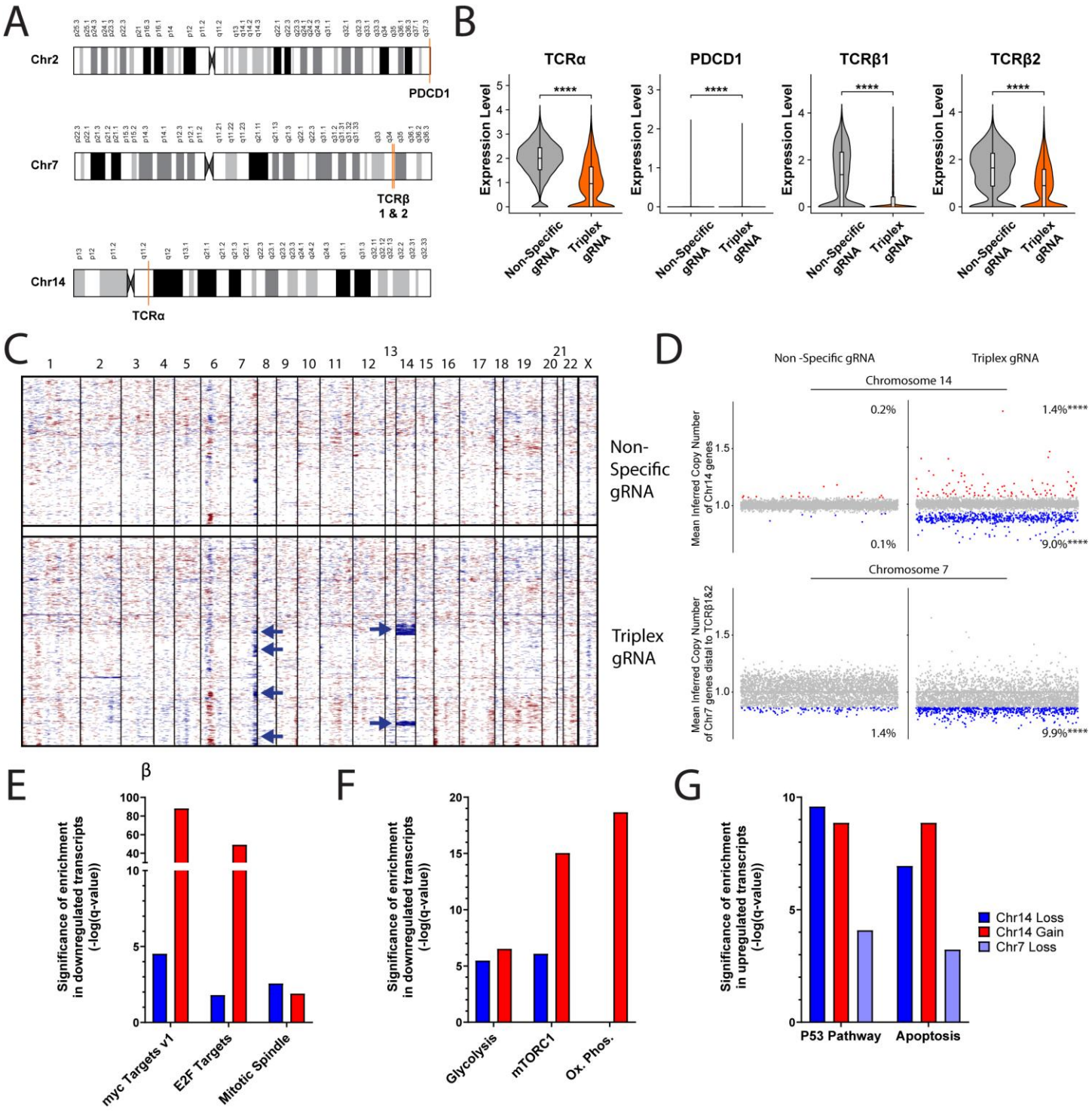


Figure 4: Targeting different loci concomitantly with several gRNAs aggravates chromosomal aberrations. **A.** Schematic depiction of human chromosomes 2, 7 and 14. Target loci are shown in orange. A single gRNA targets both TCRβ1 and TCRβ2. **B.** RNA expression of TCRα, PDCD1, TCRβ1 and TCRβ2 in cells treated with either a non-specific gRNA or a combination of TCRα, TCRβs and PDCD1- targeting gRNAs. ****, p<0.0001, two-sided unpaired Wilcoxon test. **C.** A heat map depicting gene copy numbers inferred from scRNAseq analysis following treatment with either a non-specific gRNA or a combination of TCRα, TCRβs and PDCD1- targeting gRNAs. Each line represents an individual cell. The chromosomes are ordered in columns and the color coding indicates an increase (red) or decrease (blue) in copy number. Blue arrows indicate loss of chromosome 14 or chromosome 7 segments. **D.** Each dot represents the mean inferred copy number of genes coded on the relevant chromosome in each cell treated with a non-specific gRNA (left) the combination of TCRα, TCRβs and PDCD1- targeting gRNAs (right). For chromosome 7, the mean inferred copy number of genes coded distal to the TCRβ1&2 is analyzed. Cells are marked with dots spread along the x-axis. The dots are colored red and blue when corresponding to cells with a chromosome 14 gain or loss

respectively, if their mean inferred gene copy number is >2 standard deviations away from the population's mean. n=8619 and 6326 for cells treated with a non-specific gRNA (left) or a combination of TCRα, TCRβs and PDCD1-targeting gRNAs (right), respectively. ****, p<0.0001 for Fisher's exact test comparing chromosome gain or loss between cells treated with the TCRα/β gRNA and cells treated with a non-specific gRNA. **E-G.** Gene set enrichment analysis of 'Hallmark' gene sets associated with cell cycle (E), metabolism (F), p53 expression and apoptosis (G) in cells with chromosome 14 loss (blue) and in cells with chromosome 14 gain (red).

METHODS:

T cell editing

Whole blood was obtained with donor consent from the Israeli Blood Bank (Magen David Adom, Sheiba Medical Center) in accordance with Tel Aviv University Review Board. PBMCs were extracted using Lymphocyte Separation Medium (mpbio) and cryopreserved until subsequent use. Following thawing, cells were activated for 24-48hrs with 1ug/ml anti-human CD28 (biogems) and anti-human CD3 (biogems). Cells were cultured in MEM-Alpha (Biological Industries) supplemented with 10% Heat Inactivated FCS (Sigma), 50IU rhIL-2 (Peprotech) and Pen/Strep.
For RNP electroporation, 18pmol of Alt-R spCas9 Nuclease V3 (IDT) and 66pmol of Alt-R CRISPR-Cas9 gRNA (IDT) per 1E6 cells per target were complexed in Buffer T (IDT). Experiments involving a triple targets (TCRα, TCRβ, PD-1), quantities of Cas9:gRNA complexes were tripled for a non-specific control. For experiments aiming at reducing the aneuploid cell fraction we used Alt-R sp HiFi Cas9 Nuclease V3 (IDT). For experiments involving CRISPR-Cas12a, Alt-R L.b. Cas12a (Cpf1) Ultra (IDT) and crRNA were used at mass ratios equivalent to those used in Cas9 electroporations, when 1E6 cells per target were complexed in Buffer T (IDT). gRNA and crRNA sequences are provided in Supplementary Table 5. Cells were harvested, washed and electroporated at 1600v, 10ms, 3pulses in 10ul Buffer T and subsequently grown in culture media devoid of P/S. For prolonged culture, live cells were purified using Lymphocyte Separation Medium (mpbio) every other day.

Flow Cytometry

Cells were harvested, washed, and resuspended in Cell Staining Buffer (Biolegend) containing 1/100 diluted anti-human CD3 or anti-human TCRα/β (Biolegend), both targeting the TCR complex. Staining was performed for 15mins at room temperature in the dark. Finally, cells were washed and data acquisition was performed on an Attune NxT Flow Cytometer (life Technologies).

FACS

Cell sorting was performed 3 days following electroporation. Cells were harvested, washed and resuspended in PBS supplemented with 5% BSA containing 1/100 diluted anti-human CD70 and anti-human CD52 (Biolegend). Staining was performed for 15 mins at room temperature in the dark. Subsequently, cells were washed and resuspended in PBS supplemented with 0.5% BSA. During Fluorescence Activated Cell Sorting, cells were collected in MEM-Alpha (Biological Industries) supplemented with 0.5% Heat Inactivated FCS (Sigma) and P/S. Acquisition was performed on a BD FACSAria III (BD Biosciences). After collection, live cells were purified using Lymphocyte Separation Medium (mpbio), washed and cultured in MEM-Alpha (Biological Industries) supplemented with 10% Heat Inactivated FCS (Sigma), 50IU rhIL-2 (Peprotech) and P/S. Purity of the sorted cells can be found in Extended Data Fig. 13.

Single-cell RNA sequencing

Edited T cells were harvested and live cells were purified using Lymphocyte Separation Medium (mpbio), washed and resuspended in PBS supplemented with 0.5% BSA to achieve optimal concentration of approximately 1000 cells per microliter. Cells were counted and viability assessed manually in Trypan Blue 0.4% (Biological Industries). 17,000 cells were loaded on Next GEM Chip (10x Genomics). Libraries were prepared at the Single-Cell Genomics Core, Faculty of Medicine, Tel Aviv University, using the 10x Genomics Chromium Controller in conjunction with the single-cell 3′ v3.1 kit, protocol revision D. The cDNA synthesis, barcoding, and library preparation were then carried out according to the manufacturer's instructions. Briefly, cDNA amplification was performed for 11 cycles. Sample index PCR was performed for 13 cycles using Chromium i7 Sample Indices. Resulting libraries were quantified and analyzed by Qubit and Tapestation. Libraries were sequenced on the NextSeq 500 platform (Illumina), following the manufacturer's protocol, using a NextSeq 500/550 High Output Kit v2.5 (75 Cycles) kit (Illumina). Sequencing was performed at the Genomics Research Unit, at the Life Sciences Inter-Departmental Research Facility Unit, Tel-Aviv University.

scRNA-seq gene expression pre-processing

Raw BCL files for the DNA sequencing data were processed using Cellranger DNA (version 5.0.1). Data were aligned to the 10X Genomics GRCh38 genome. Results were visualized in the Loupe scDNA Browser (version 5.0.0). Raw gene expression data were extracted from the Seurat object as recommended in the "Using 10x data" section (inferCNV of the Trinity CTAT Project, https://github.com/broadinstitute/inferCNV).

InferCNV

InferCNV was used to infer copy number changes from the gene expression profiles[29,30] The non-targeted T cell population was used as the reference, and the CRISPR/Cas9-targeted population was tested, with the following parameters: "denoise", default hidden markov model (HMM) settings, and a value of 0.1 for "cutoff".

Identification of aberrant cells

For each cell, the mean of inferCNV scores was calculated across genes and plotted. The PDCD1 gene resides near the chromosome 2 q-arm telomere and only 15 genes, expressed in T cells, reside between the TCRα gene and the chromosome 14 centromere. Therefore, for chromosome 14 and chromosome 2, all of the expressed genes on the respective chromosomes were used for the analysis. For chromosome 7, the 47 expressed genes that reside distal to TCRβ were used. Cells with a mean lower or higher >2 standard deviations from the mean of the population were determined as cells with a loss or a gain, respectively.

Simulation analysis for downregulated genes

In order to determine whether cells categorized as harboring a chromosome 14 loss of a chr7 distal loss, also had a significant increase in the fraction of zero expression calls (that is, whether these regions are enriched with genes not detected at all by scRNAseq), the ratio between the number of genes from each chromosome with expression = 0 and expression > 0 for each cell population (loss vs. non-loss) was calculated. The fold change of the proportion of zeros calls between the normal and aberrant cells was determined, and 10,000 simulations were then performed, selecting an equivalent number of random genes from other chromosomes. An empirical p-value was determined by comparing the empirical values to the simulated values.

Differential gene expression analysis

The "FindMarkers" package from the Seurat library[46] was used to detect the differentially expressed (DE) genes between two groups of cells (logFC of 0.25). The function receives two identities of clusters in the data set and a value for the minimum percentage that is required for a feature to be detected in either of the two groups of cells. The minimum percentage value that we used is 0.25 and p-value were calculated using Wilcoxon test.

The first comparison was made to detect DE genes between the cells that have undergone loss in the TCRα gRNA and the cells that did not show loss or gain in the same treatment group. The second was to detect DE genes between the cell that have undergone gain in the TCRα gRNA and the cells that did not show gain or loss in the same treatment group.

Gene Set Enrichment Analysis

The lists of differentially expressed genes between each two conditions were determined using Seurat as described above. These lists were subjected to gene set enrichment analysis (GSEA) using the GSEA-MSigDB portal (https://www.gsea-msigdb.org/gsea/msigdb/). The analysis was run using the following curated gene sets: "Hallmark", "KEGG", "GO biological process" and "positional" gene sets from MSigDB[47–49].

ddPCR

For digital droplet PCR, whole genomic DNA was extracted from cells using Gentra PureGene Tissue Kit (Qiagen). In order to remove sheared genomic fragments, resulting eluates were further purified using AmpureXP beads (Beckman Coulter) at a 0.5:1 ratio. DNA fragmentation by digestion was performed in reaction, using 66ng of purified genomic DNA and 10U HindIII-HF (NEB) in ddPCR Supermix for Probes (BioRad). Thermo-cycling reaction was performed as per manufacturer recommendation. Sequences for the primers and probes can be found in Supplementary Table 5. Reactions were performed using a QX200 Droplet Digital PCR System (Bio-Rad). To analyze for dislinkage, we used the following equation as per the resulting Quantasoft (BioRad) Linkage (*Linkage*) and Concentration (C_{HEX}, C_{FAM}) values:

$$\%\ Dislinkage = \ 100 - \left(\frac{Linkage}{\frac{C_{HEX} + C_{FAM}}{2}} x100 \right)$$

For dislinkage followup, multiple electroporations of treated cells were pooled and then divided in separate wells for collection at each time point. Cells were seeded at 1E6 cells/ml in MEM-Alpha (Biological Industries) supplemented with 10% Heat Inactivated FCS (Sigma) and 50IU rhIL-2 (Peprotech).

Nucleic acid manipulations

For T7E1 assays, PCR amplification was performed on Gentra PureGene Tissue Kit (Qiagen) extracted genomic DNA. 200-500ng of genomic DNA was amplified using PrimeStar MAX (Takara) for 35 cycles. Primers for these reactions can be found in Supplementary Table 5. Resulting amplicons were denatured and reannealed in a thermocycler before nuclease reaction using T7 Endonuclease 1 (New England Biolabs) at 37C for 20min. Resulting fragments were analyzed by agarose gel electrophoresis and quantified using Biovision (Vilber Lourmat) using a rolling ball for background subtraction. Efficiency was calculated using the following equation:

$Cleavage\ efficiency = 100$ x $(1 - (1 - fraction\ cleaved)^{\frac{1}{2}}))$

For TIDE analysis, PCR amplicons were subjected to purification by AmpureXP beads (Beckman Coulter) at a 1:1 ratio. Sanger sequencing was performed at the DNA Sequencing Unit, Tel Aviv University. Samples were compared using TIDE (https://tide.nki.nl/)[50].

Fluorescence in-situ hybridization

Fluorescence in situ hybridization (FISH) analysis was performed following the manufacturer's instructions (Cytocell) on interphase human T cells from peripheral blood spreads' using the TRACD breakapart probe. Images were captured using GenASIs imaging system.

Statistics

For FISH, genomic DNA cleavage efficiency and flow cytometry knock out efficiency, statistical analyses were performed on Prism (GraphPad). For t-tests on dislinkage by ddPCR, for each donor, technical replicates were averaged and t-test was performed on the averaged values. Each figure legend denotes the statistic used, central tendency and error bars.

ACKOWLEDGEMENTS:

We thank Natalie Zelikson (TAU), for helpful discussions and reviewing the manuscript. We thank Rani Elkon for computational support. We thank the IDRFU, GRU and SICF units, Tel Aviv University for logistic support and council. This research was funded by the European Research Council (A.B. and U.B.-D.), the Gertner Institute Scholarship, the Yoran Institute Scholarship, the SAIA Foundation (A.D.N.), SCGC Tel Aviv University (A.B. and U.B.-D.), the Israel Cancer Research Development (A.M) and Gesher Award (U.B.-D.), the Azrieli Faculty Fellowship (U.B.-D.), the Alon Fellowship for Outstanding Young Scientists, Israel Council for Higher Education (A.M.), the Edmond J. Safra Center fellowship for Bioinformatics at Tel-Aviv University (E.G.).

AUTHOR CONTRIBUTIONS:

A.D.N. designed, performed and analyzed the experiments; E.R. and E.G. analyzed the single-cell RNA sequencing data; T.T. and M.L. performed fluorescence in-situ hybridization; M.H-F. helped with sample culturing and processing; R.K. and H.K. performed single-cell RNA sequencing; E.R. supervised fluorescence in-situ hybridization; A.M. and U.B.-D. supervised single-cell RNA sequencing analysis; A.D.N. and A.B. drafted the manuscript, and revised it together with A.M. and U.B.-D; U.B.-D. and A.B. supervised the study; A.B. conceptualized the study.

COMPETING INTERESTS:

None

CODE AVAILABILITY

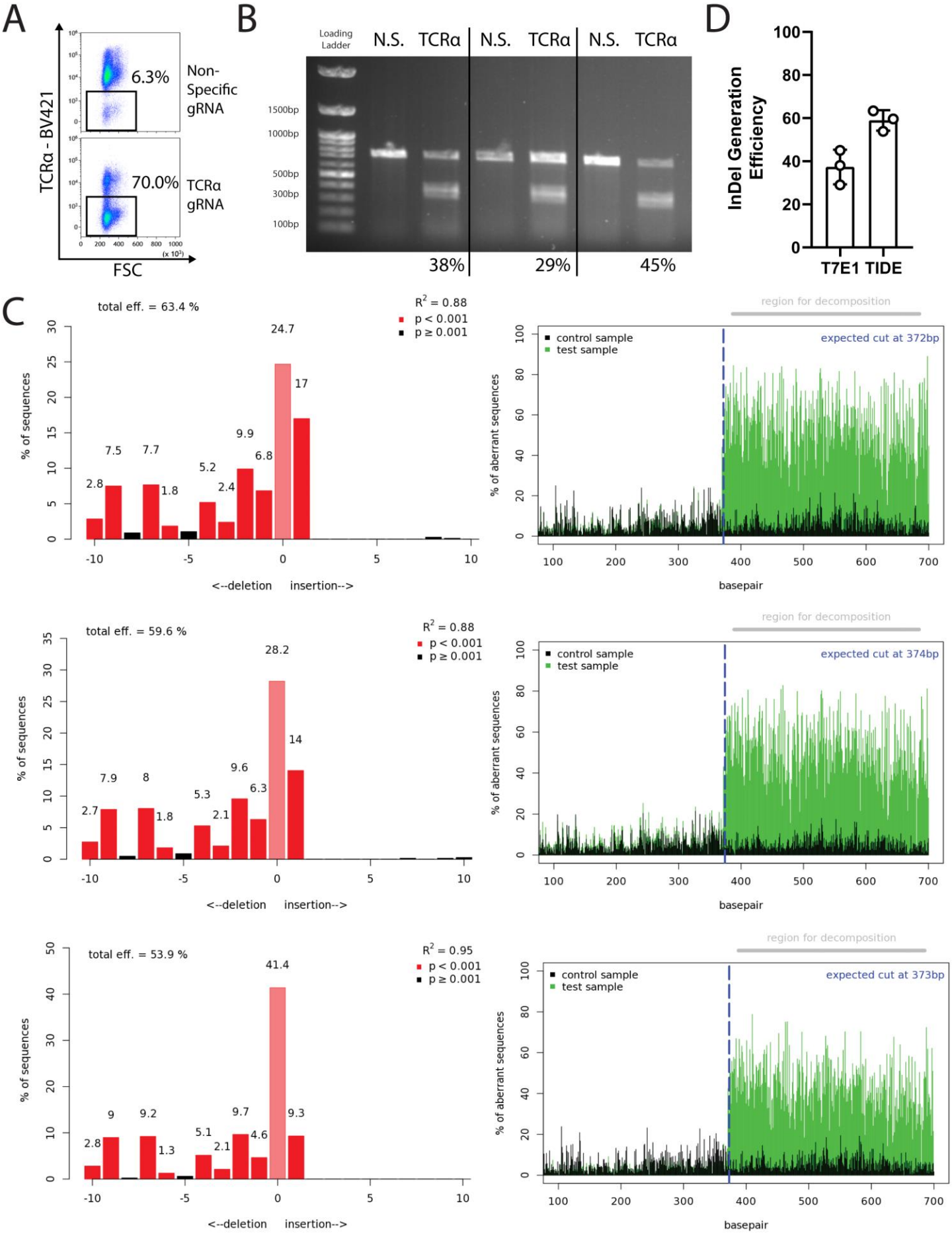
The software used for individual and integrated analyses are described and referenced in the individual sections in Methods.

DATA AVAILABILITY

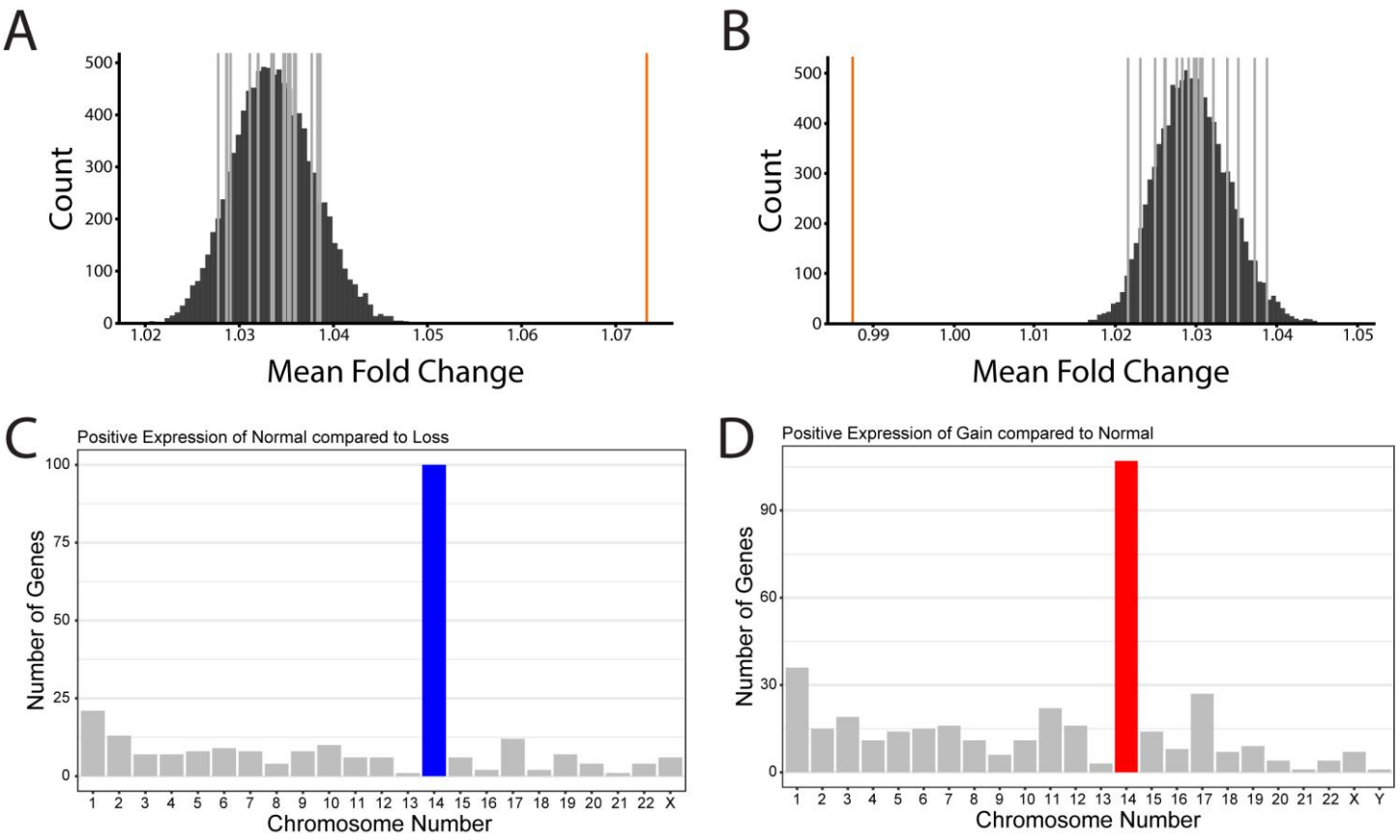
All datasets are available within the article and its Supplementary Information. 10X scRNAseq data have been deposited to SRA with BioProject accession number PRJNA759387.

Supplementary Information and Extended Data Figures are available for this paper.

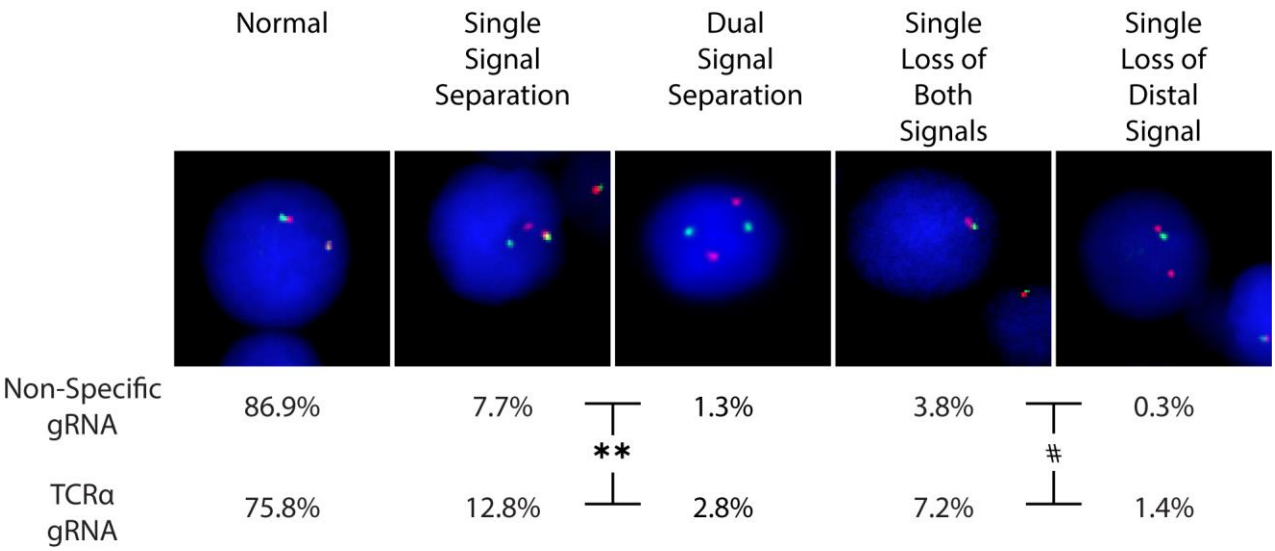
EXTENDED DATA FIGURES:


A
TCRα - BV421
FSC
6.3%
Non-Specific gRNA
70.0%
TCRα gRNA
B
Loading Ladder
N.S.
TCRα
N.S.
TCRα
N.S.
TCRα
1500bp
1000bp
500bp
300bp
100bp
38%
29%
45%
D
InDel Generation Efficiency
T7E1
TIDE
C
total eff. = 63.4 %
R² = 0.88
p < 0.001
p ≥ 0.001
% of sequences
<--deletion insertion-->
region for decomposition
control sample
test sample
expected cut at 372bp
% of aberrant sequences
basepair
total eff. = 59.6 %
R² = 0.88
expected cut at 374bp
total eff. = 53.9 %
R² = 0.95
expected cut at 373bp

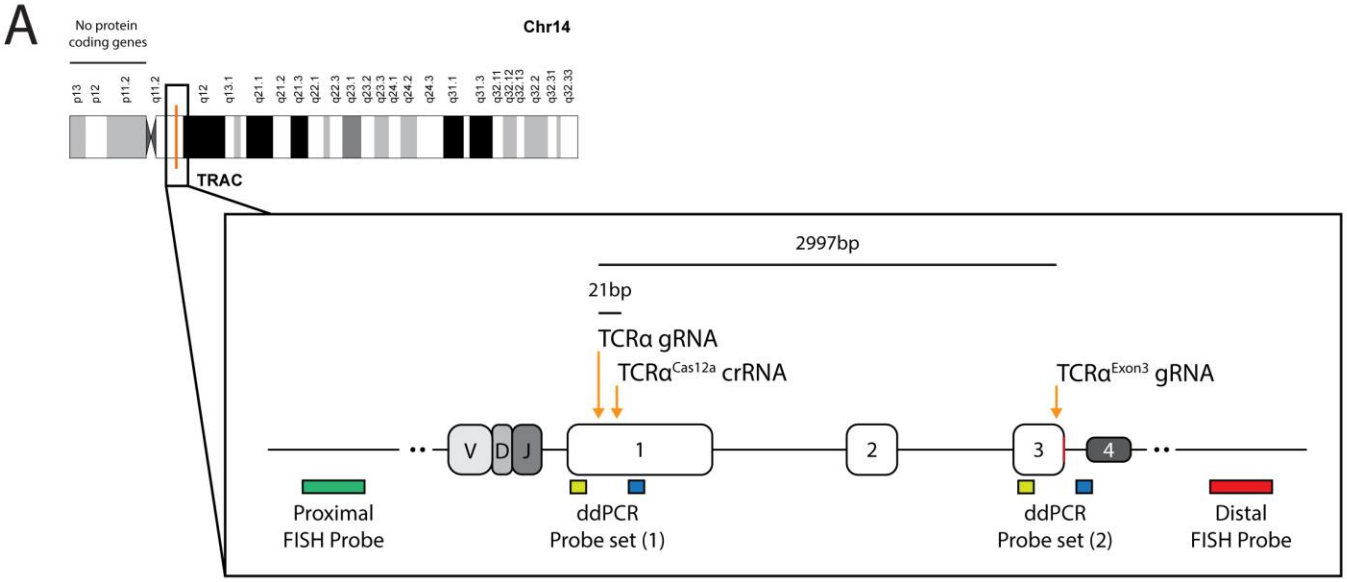
Extended Data Figure 1: Quantification of InDels produced by CRISPR-Cas9 activity at the TCRα locus indicates efficient cleavage. **A.** Flow cytometry example of TCR ablation in primary human T cells following CRISPR-Cas9 RNP electroporation. Cells were electroporated with Cas9 and either a non-specific gRNA or a TCRα-targeting gRNA. **B.** T7 Endonuclease 1 (T7E1) assay for three independent experiments. For each experiment, one lane for non-Specific gRNA (N.S.) and one lane for TCRα-targeting gRNA treated cells are presented. Experiments are separated by a black bar. Percentages, presented below the lanes, refer to cleavage efficiency as inferred by densitometric analysis. Loading Ladder and relative sizes are indicated on the left. Unprocessed scan can be found in Supplementary Data 1. **C.** TIDE Analysis for the same experiments as in B. In the left panel, the height of each bar corresponds to the rate of sequences having the given number of nucleotides added or deleted. The right panel depicts the rate of sequence misalignments at each position of the PCR fragment amplified from the TCRα locus of cells treated with either the TCRα-targeting gRNA (green) or a non-Specific gRNA (black) **D.** Quantification of B and C.

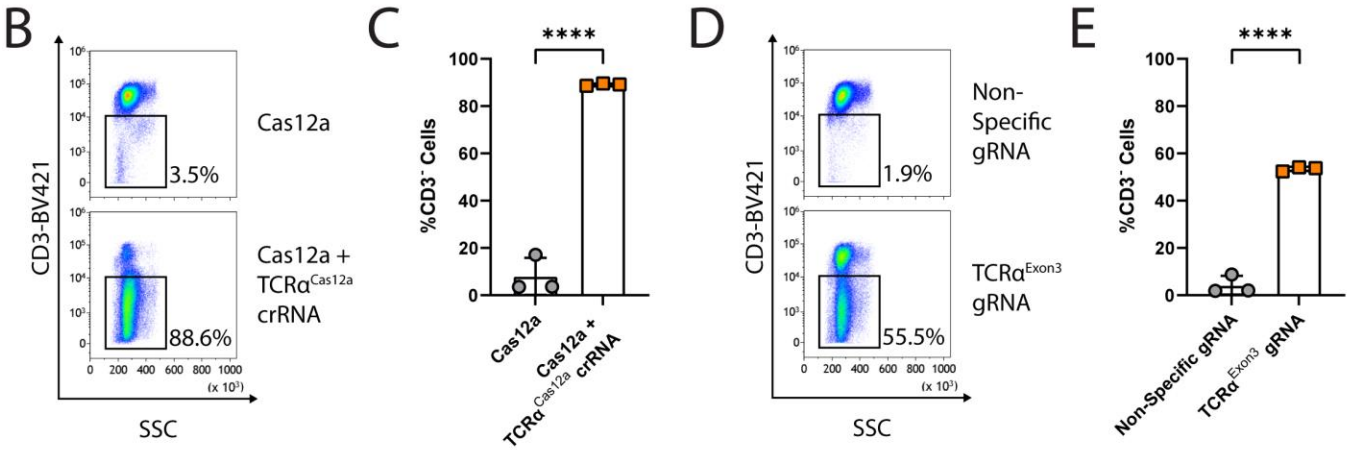


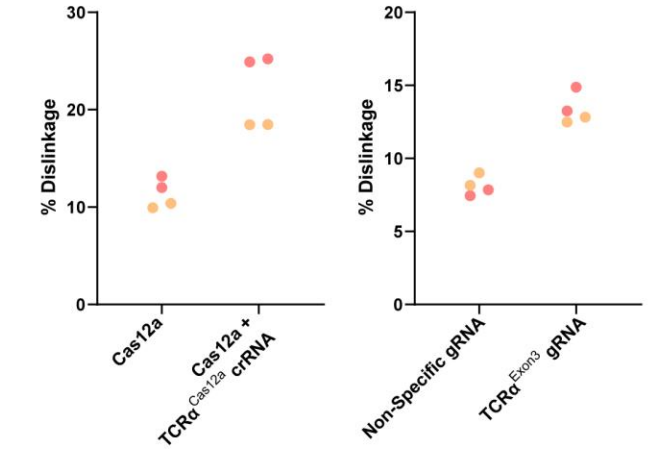
Extended Data Figure 2: A-B. Enrichment in the number of genes with no detected expression among cells identified as having a chromosome 14 loss (A) or gain (B) (Fig. 1F). The x-axis represents the fold-change in the number of genes with no detected expression between cells with or without a chromosome 14 loss, based on the InferCNV analysis (Fig. 1F). The dark gray lines represent the empirical values obtained for each chromosome, except for chromosome 14. The orange line is the empirical value for chromosome 14. The black bars are the results of 10,000 permutations. **C.** Number of differentially expressed genes in each chromosome, as compared between cells with or without a chromosome 14 loss (see Fig. 1F). **D.** Number of differentially expressed genes in each chromosome, as compared between cells with or without a chromosome 14 gain (see Fig. 1F).

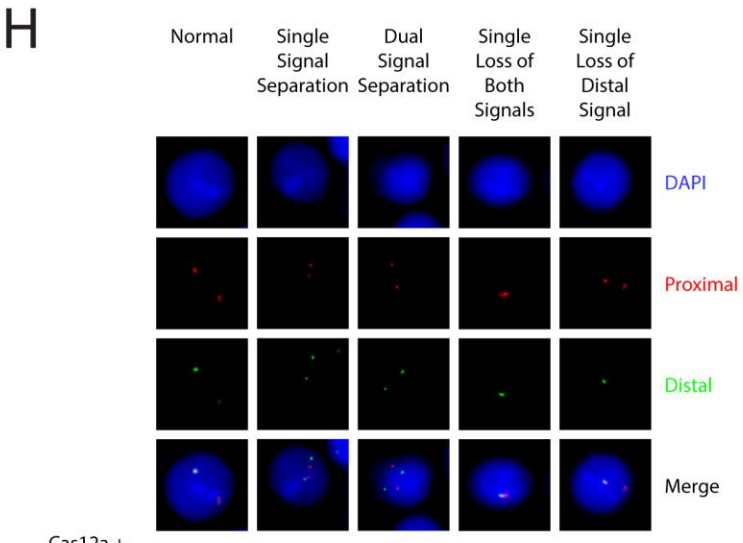


Extended Data Figure 3: FISH analysis of cells treated with either a non-specific or a TCRα-targeting gRNA. In the pictures are examples of all signal patterns found among analyzed cells. The frequency of each signal pattern, among cells electroporated with the TCRα-targeting gRNA or a non-specific gRNA, is presented below the pictures. n = 313 and n = 360 for non-specific and TCRα targeted cells, respectively. #, p= 0.0115 Fisher's exact test for loss of either both signals or the distal signal. ** p =0.0048 for either single or dual Signal Separation.





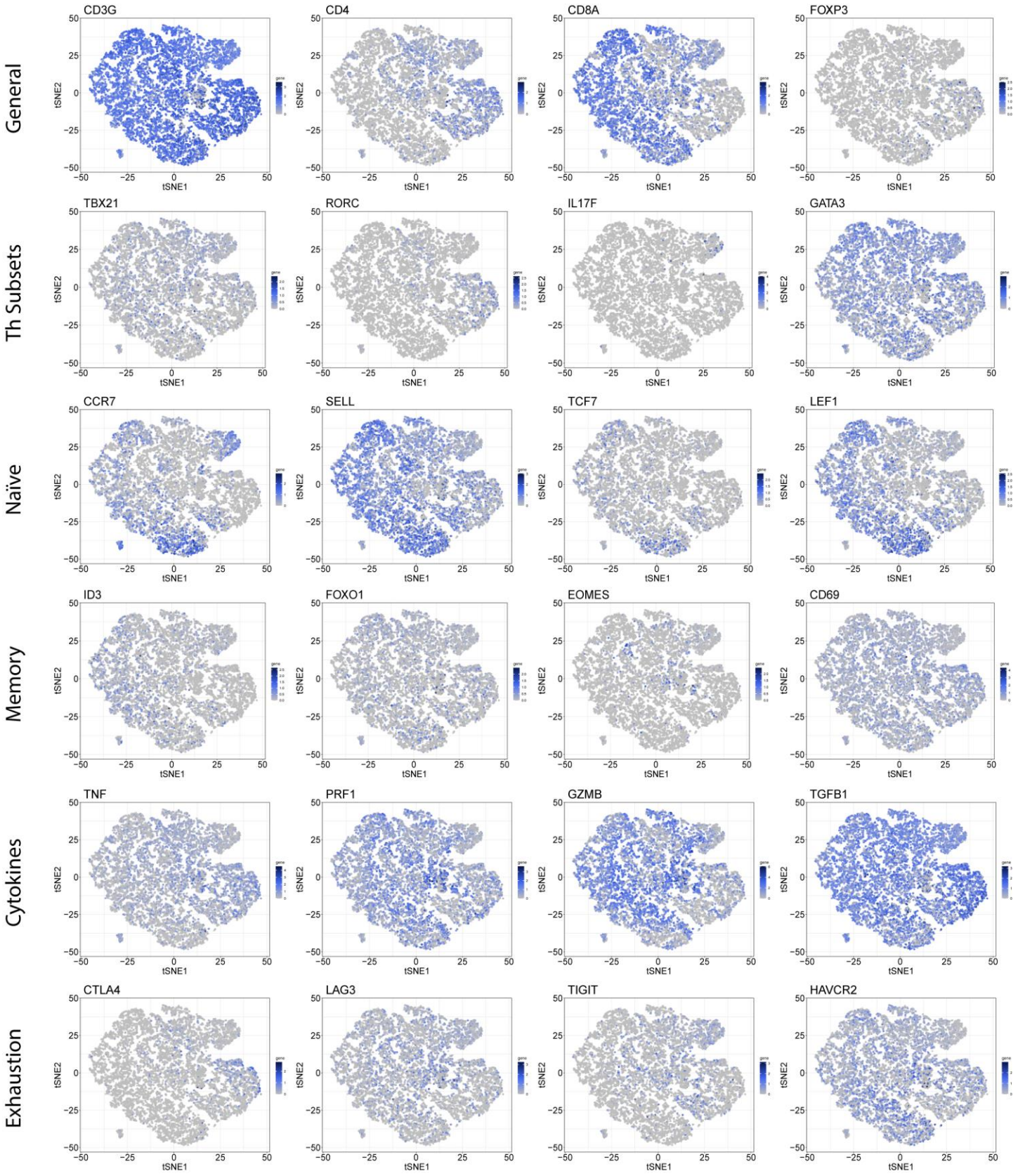




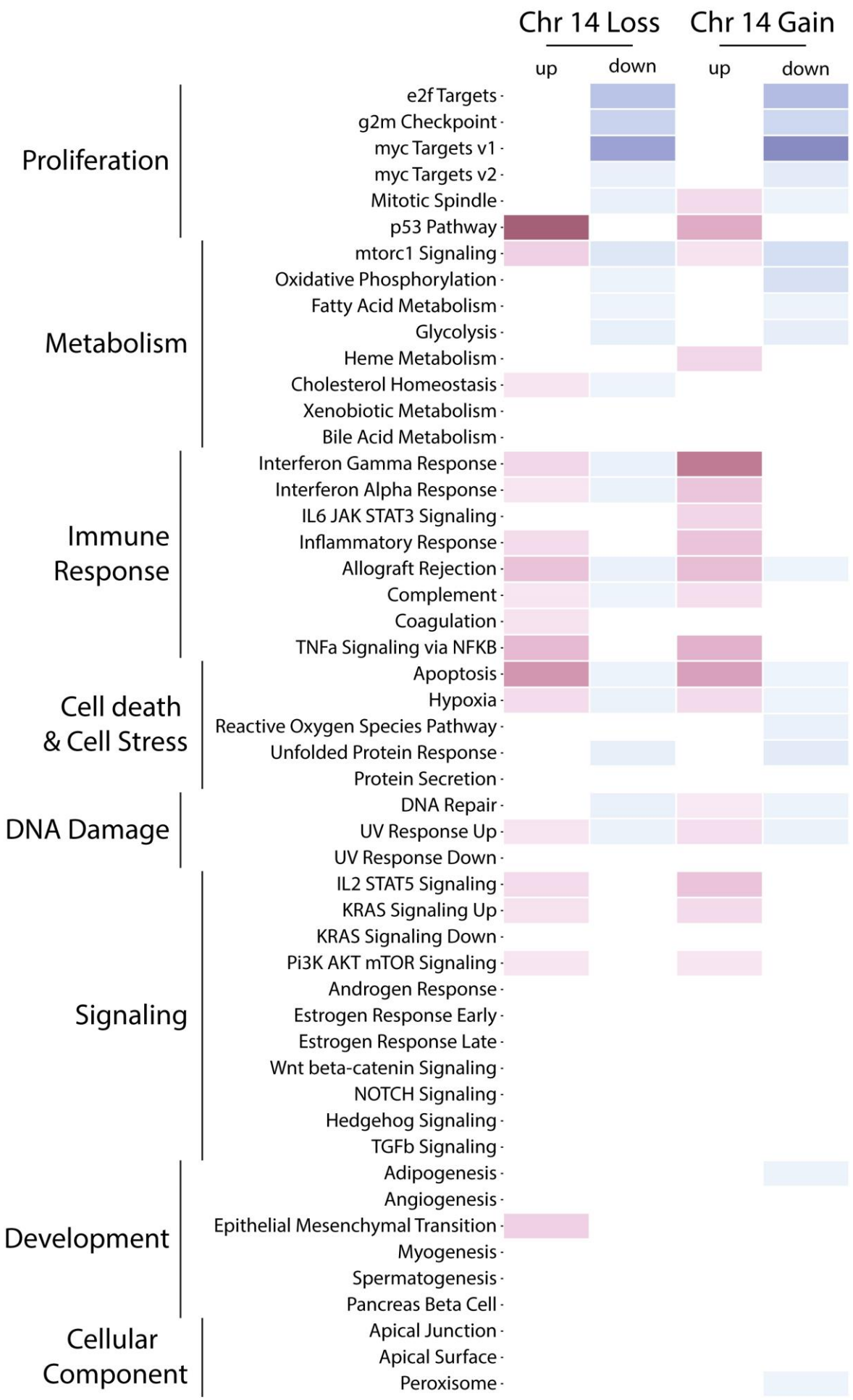
	Normal	Single Signal Separation	Dual Signal Separation	Single Loss of Both Signals	Single Loss of Distal Signal
Cas12a + TCRα^{Cas12a} crRNA	79.3%	8.7%	2.2%	4.8%	5.0%
Cas12a	87.6%	6.2%	2.4%	3.0%	0.9%
TCRα^{Exon3} gRNA	78.9%	6.8%	3.7%	6.1%	4.4%
Non-Specific gRNA	89.7%	5.6%	1.9%	1.4%	1.4%
TCRα gRNA	77.1%	7.7%	3.1%	7.2%	4.9%

##, ####, **, n.s.

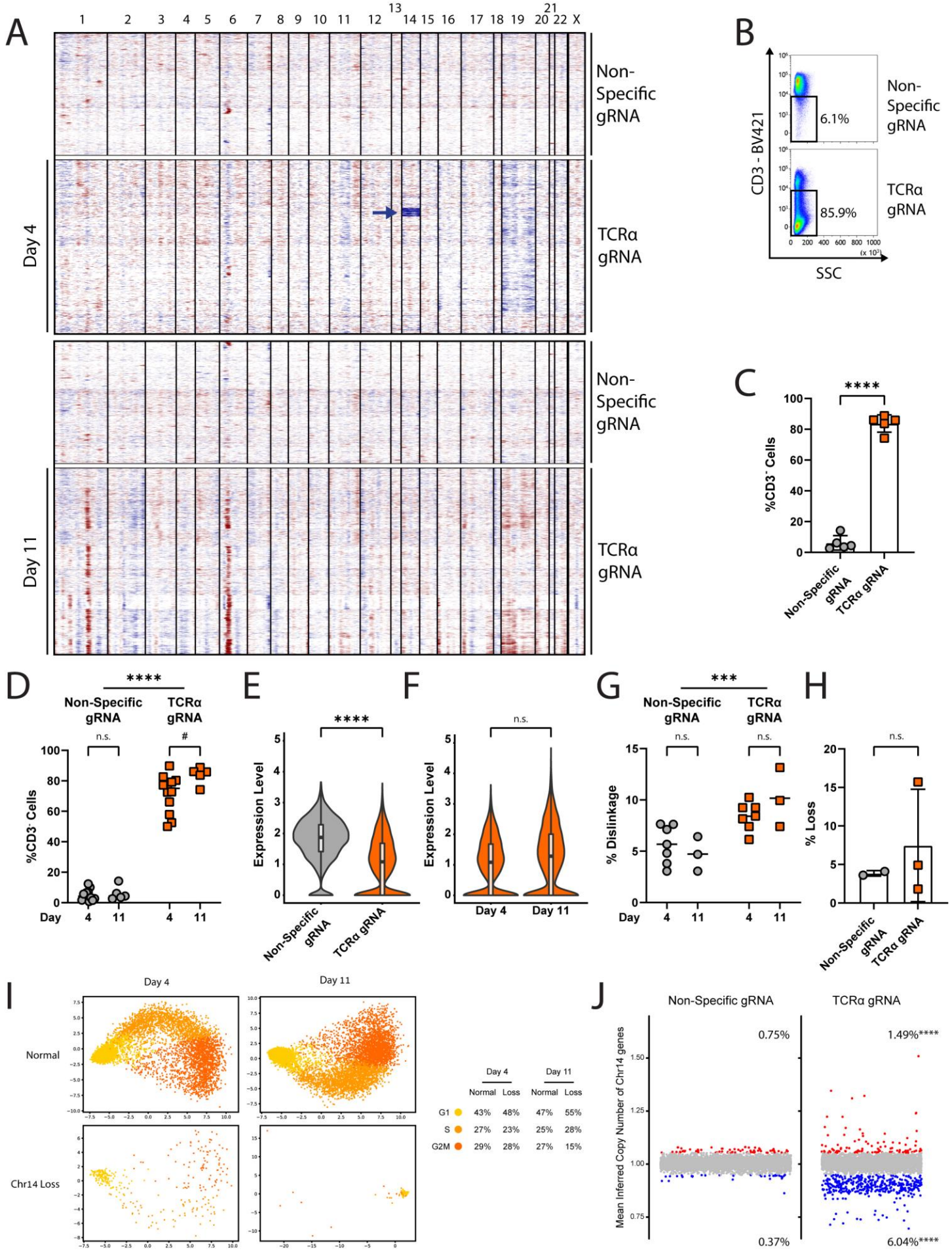
Extended Data Figure 4: A. Schematic depiction of human chromosome 14 (Chr14). The target locus is indicated in orange. The zoomed in box indicates more specifically, the different target sites of each gRNA/crRNA used in this study. ddPCR and FISH probe sets are indicated below as colored yellow/blue or green/red boxes, respectively. Numbers above indicate distances, in base pairs (bp), between the targets. **B.** Flow cytometry example of TCR ablation in primary human T cells, 4 days following CRISPR-Cas12a RNP electroporation. Cells were electroporated with Cas12a alone or with Cas12a and the relevant crRNA. **C.** Quantification of B. Each dot represents an independent experiment. Mean value, standard deviation and individual experiments are indicated. n=3, ****, p<0.0001, two-sided unpaired Wilcoxon test. **D.** Flow cytometry example of TCR ablation in primary human T cells following CRISPR-Cas9 RNP electroporation. Cells were electroporated with Cas9 and either a non-specific gRNA or the TCRα-targeting gRNA in exon3 ($TCR\alpha^{exon3}$) as in A. **E.** Quantification of B. Each dot represents an independent experiment. Mean value, standard deviation and individual experiments are indicated. n=3, ****, p<0.0001, two-sided unpaired Wilcoxon test. **F-G** ddPCR dislinkage analysis 4 days following CRISPR-Cas12a electroporation (F) or CRISPR-Cas9 electroporation with the $TCR\alpha^{exon3}$ gRNA (G). Each color represents a different human donor. Each dot represents a technical replication. **H.** FISH analysis of cells treated with either Cas12a only, Cas12a + $TCR\alpha^{Cas12a}$ crRNA or Cas9 + a $TCR\alpha^{exon3}$-targeting gRNA. In the pictures are examples of all signal patterns found among analyzed cells. The frequency of each signal pattern among the electroporated cells is presented below the pictures. n = 338, n = 357 and n = 427 for cells electroporated with Cas12a only, Cas12a + $TCR\alpha^{Cas12a}$ crRNA or Cas9 + a $TCR\alpha^{exon3}$-targeting gRNA, respectively and n=358 and n=388 for cells electroporated with Cas9 + non-specific gRNA or Cas9 + TCRα targeting gRNA (as in Fig. 1A), respectively. **, p=0.0012 for Fisher's exact test comparing loss of distal signal between cells transfected with Cas12a only and cells transfected with Cas12a + $TCR\alpha^{Cas12a}$ crRNA. ##, p=0.0015 for Fisher's exact test comparing loss of either both signals or only the distal signal between cells transfected with Cas12a only and cells transfected with Cas12a + $TCR\alpha^{Cas12a}$ crRNA. ####, p<0.0001 for Fisher's exact test comparing loss of distal signal between cells transfected with $TCR\alpha^{exon3}$ gRNA and cells transfected with a non-specific gRNA and for Fisher's exact test comparing loss of distal signal between cells transfected with TCRα gRNA and cells transfected with a non-specific gRNA. n.s., p>0.05 for Fisher's exact test. comparing loss of distal signal between cells transfected with $TCR\alpha^{exon3}$ gRNA and cells transfected with TCRα gRNA. **, p=0.0069 for Fisher's exact test comparing loss of only the distal signal between cells transfected with $TCR\alpha^{exon3}$ gRNA and cells transfected with a non-specific gRNA. **, p=0.0031 for Fisher's exact test comparing loss of only the distal signal between cells transfected with the TCRα gRNA to cells transfected with a non-specific gRNA.



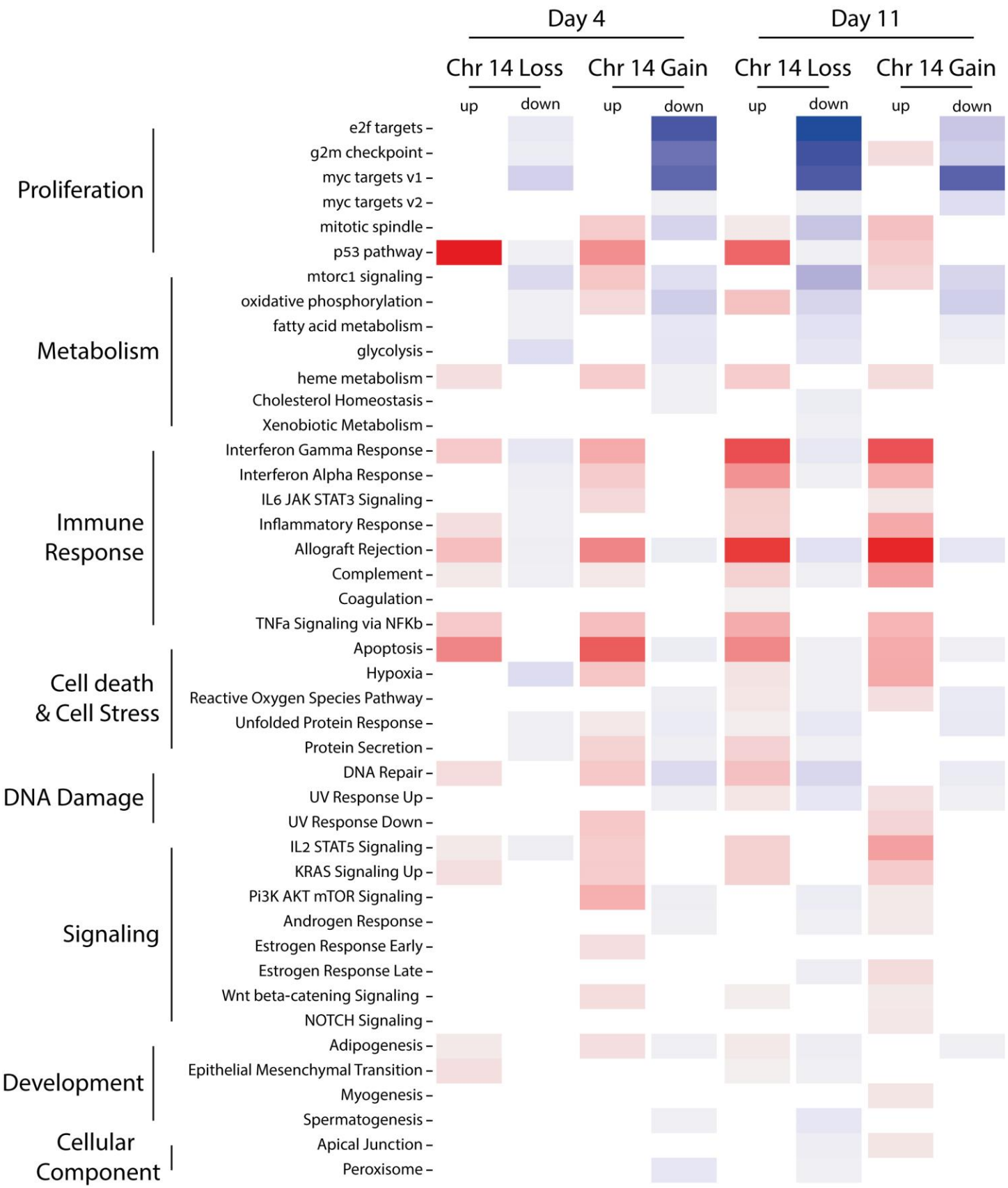
Extended Data Figure 5: Selected gene expression patterns across the cells. Each t-SNE plots represent all the cells in the experiment (Fig. 3A-C). Each plot represents the expression pattern of a different gene, indicated on the top left, among the clusters. A darker shade of blue corresponds to higher RNA expression. For each row, a title for the type of markers is indicated on the left.


Chr 14 Loss
Chr 14 Gain
up
down
up
down
Proliferation
e2f Targets
g2m Checkpoint
myc Targets v1
myc Targets v2
Mitotic Spindle
p53 Pathway
Metabolism
mtorc1 Signaling
Oxidative Phosphorylation
Fatty Acid Metabolism
Glycolysis
Heme Metabolism
Cholesterol Homeostasis
Xenobiotic Metabolism
Bile Acid Metabolism
Immune Response
Interferon Gamma Response
Interferon Alpha Response
IL6 JAK STAT3 Signaling
Inflammatory Response
Allograft Rejection
Complement
Coagulation
TNFa Signaling via NFKB
Cell death & Cell Stress
Apoptosis
Hypoxia
Reactive Oxygen Species Pathway
Unfolded Protein Response
Protein Secretion
DNA Damage
DNA Repair
UV Response Up
UV Response Down
Signaling
IL2 STAT5 Signaling
KRAS Signaling Up
KRAS Signaling Down
Pi3K AKT mTOR Signaling
Androgen Response
Estrogen Response Early
Estrogen Response Late
Wnt beta-catenin Signaling
NOTCH Signaling
Hedgehog Signaling
TGFb Signaling
Development
Adipogenesis
Angiogenesis
Epithelial Mesenchymal Transition
Myogenesis
Spermatogenesis
Pancreas Beta Cell
Cellular Component
Apical Junction
Apical Surface
Peroxisome

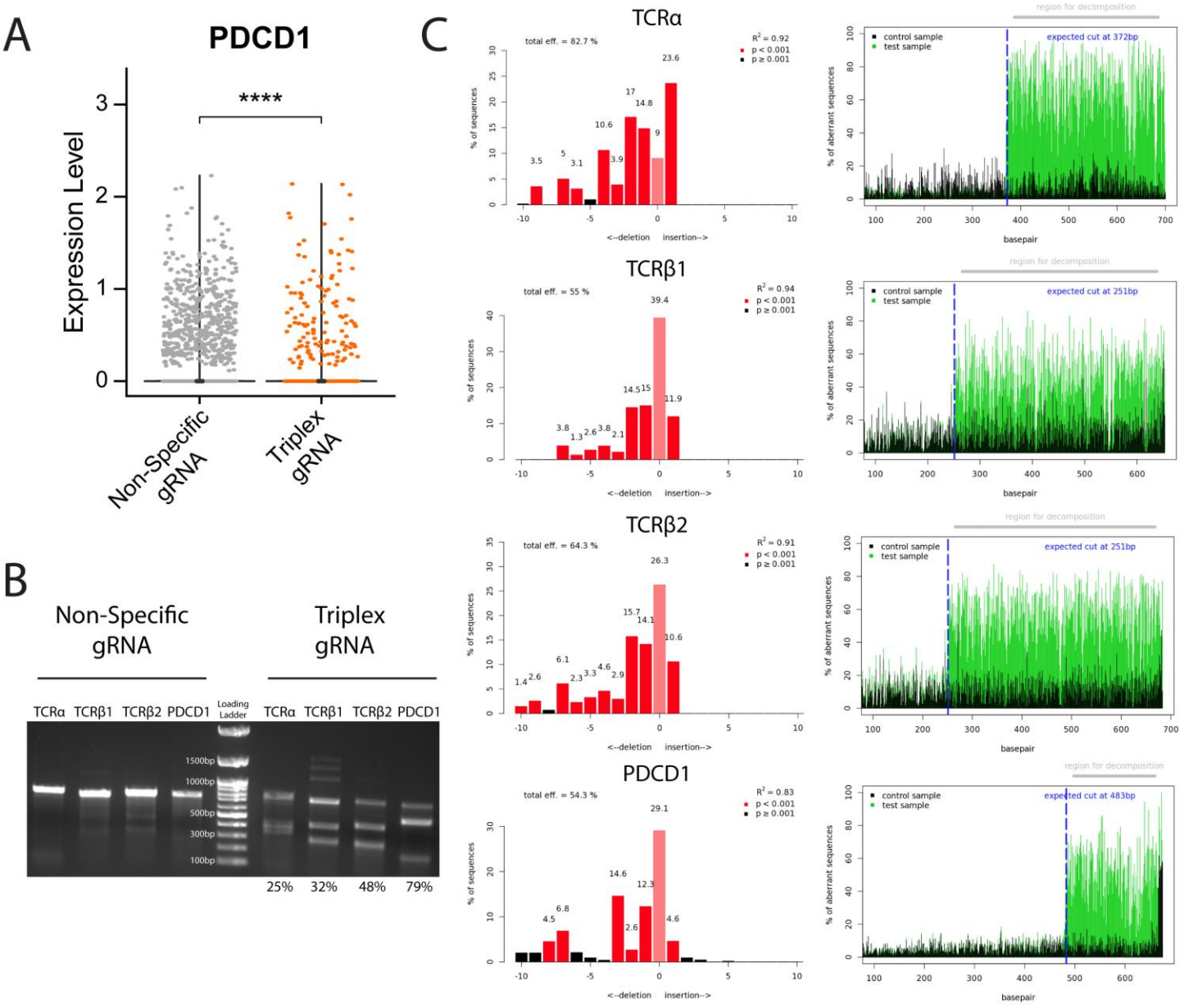
Extended Data Figure 6: A gene set enrichment analysis between cells that have lost a copy of chromosome 14 (Fig. 1F), to cells without these chromosomal aberrations. The 50 'Hallmark' MSigDB gene sets are shown. Gene sets enriched in up-regulated genes are depicted in red and those enriched in down-regulated genes are depicted in blue. Values are scaled to -log(FDR) of the enrichments. The full enrichment scores are shown in Supplementary Table 2.


A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X
Day 4
Non-Specific gRNA
TCRα gRNA
Day 11
Non-Specific gRNA
TCRα gRNA
B
CD3 - BV421
SSC
6.1%
85.9%
Non-Specific gRNA
TCRα gRNA
C
%CD3- Cells
****
Non-Specific gRNA
TCRα gRNA
D
%CD3- Cells
****
n.s.
#
Non-Specific gRNA
TCRα gRNA
Day
4 11 4 11
E
Expression Level
****
Non-Specific gRNA
TCRα gRNA
F
Expression Level
n.s.
Day 4
Day 11
G
% Dislinkage
***
n.s.
n.s.
Non-Specific gRNA
TCRα gRNA
Day
4 11 4 11
H
% Loss
n.s.
Non-Specific gRNA
TCRα gRNA
I
Day 4
Day 11
Normal
Chr14 Loss
Day 4 Day 11
Normal Loss Normal Loss
G1 43% 48% 47% 55%
S 27% 23% 25% 28%
G2M 29% 28% 27% 15%
J
Non-Specific gRNA
TCRα gRNA
Mean Inferred Copy Number of Chr14 genes
0.75%
0.37%
1.49%****
6.04%****

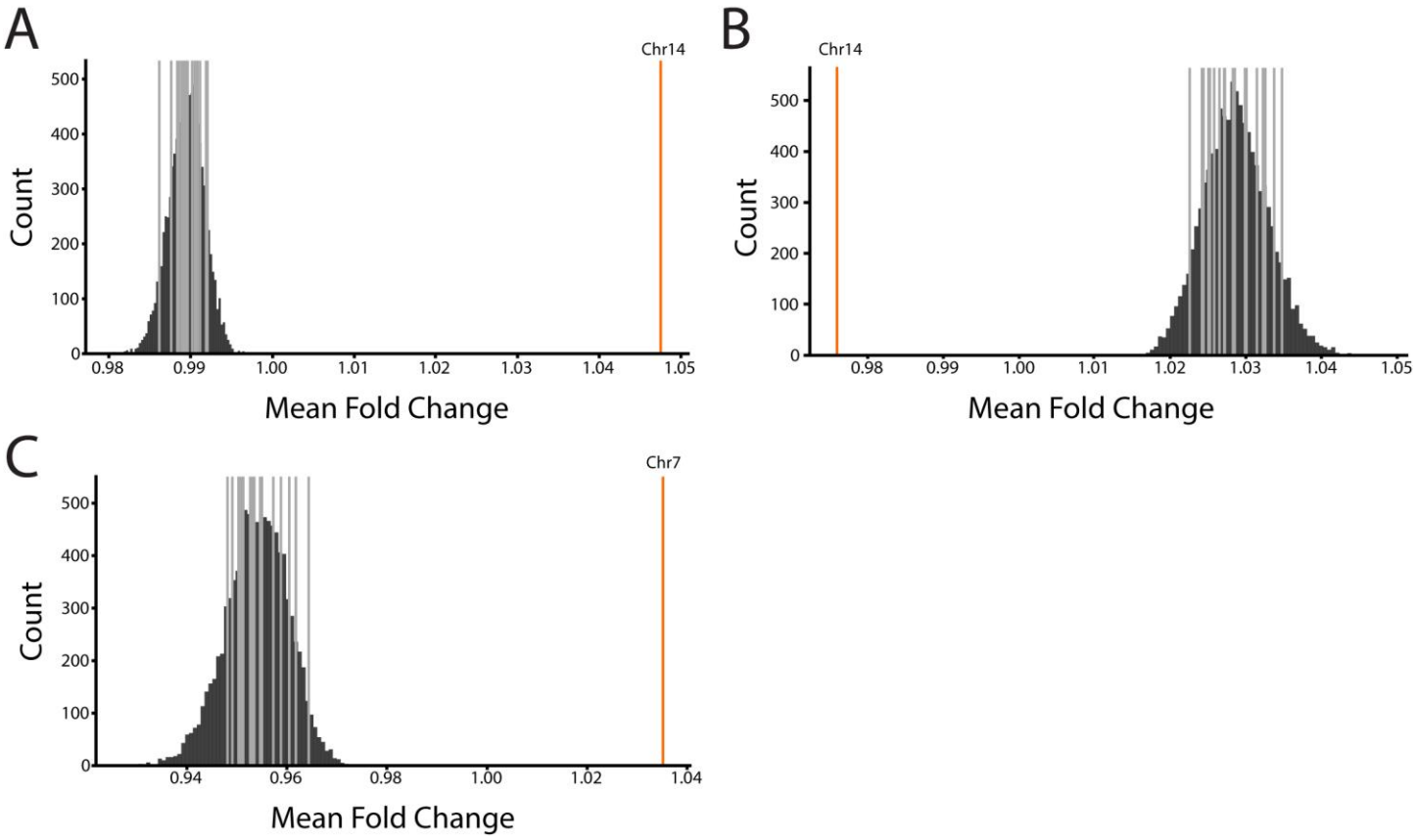
Extended Data Figure 7: A. Heat map depicting gene copy numbers inferred from scRNAseq analysis following treatment with the TCRα-targeting gRNA 4 days (above) or 11 days (below) after CRISPR-Cas9 electroporation, as presented in Fig. 3H. The expression levels in cells treated with a non-specific gRNA 4 days after CRISPR-Cas9 electroporation served as a reference in both analyses. Each line represents an individual cell. The Chromosomes are ordered in columns and the color coding indicates an increase (red) or decrease (blue) in copy number of genes along the chromosomes (x-axis). **B.** Flow cytometry example of TCR ablation in primary human T cells, measured by CD3 staining, 11 days following CRISPR-Cas9 RNP electroporation. Cells were electroporated with Cas9 and either a non-specific gRNA or a TCRα-targeting gRNA. **C.** Quantification of B. Each dot represents an independent experiment. Mean value, standard deviation and individual experiments are indicated. n=3, ****, $p<0.0001$, two-sided unpaired t-test. **D.** CD3 expression from flow cytometry data at either 4 days or 11 days following CRISPR-Cas9 electroporation. n=5-12, ****, $p<0.0001$ for Two-Way ANOVA. #, $p=0.0410$ and ns, $p>0.05$ Tukey's multiple comparison. This plot includes data presented in Fig. 1C. **E.** A reduction in TCRα expression is evident in the scRNAseq, 4 days following CRISPR-Cas9 electroporation. The violin plots correspond to the TCRα expression level in cells treated with either a non-specific gRNA or the TCRα-targeting gRNA. Upper and lower boundaries as well as median and quartiles are indicated. ****, $p<0.0001$, two-sided unpaired Wilcoxson test. **F.** A reduction of TCRα expression is stable in the scRNAseq, 11 days following CRISPR-Cas9 electroporation. The violin plots correspond to TCRα expression level in cells treated with the TCRα-targeting gRNA at either 4 days or 11 days following treatment. Upper and lower boundaries as well as median and quartiles are indicated. n.s., $p>0.05$, two-sided unpaired Wilcoxson test. **G.** ddPCR dislinkage at either 4 days or 11 days following CRISPR-Cas9 electroporation. n=3-7. Each dot represents the mean of replicates from an independent experiment, ***, $p=0.0003$ and ns, $p>0.05$ for Two-Way ANOVA and Tukey's multiple comparison, respectively. This plot includes the means of data presented in Fig. 2E. **H.** Quantification of FISH signals, from three independent experiments of T cells 11 days following CRISPR-Cas9 electroporation. Signal loss designates loss of either both signals or only the distal signal. Mean and Standard Deviations are indicated. n=2-3, ns, non-significant two-sided unpaired t-test. **I.** Principal component (PC) analysis of cell cycle phase among cells treated with a TCRα-targeting gRNAs and characterized, as in Fig. 3H, as having a chromosome 14 loss (bottom) or not (normal, up) for cells 4 days following treatment (left) or 11 days following treatment (right). Legend and fraction of cells in each cell cycle phase is indicated on the right. **J.** Each dot represents the mean inferred copy number of genes coded on chromosome 14 in each cell treated with a non-specific gRNA (left) or a TCRα-targeting gRNA (right). Cells are marked with dots spread along the x-axis. The dots are colored red and blue when corresponding to cells with a chromosome 14 gain or loss respectively, at day 4 following treatment, if their mean inferred gene copy number is >2 standard deviations (blue and red), from the population's mean. ****, $p<0.0001$ for Fisher's exact test comparing chromosome 14 gain or loss between cells treated with the TCRα gRNA and cells treated with a non-specific gRNA. Data with inferred gene copy number >3 standard deviations (blue and red), from the population's mean is presented in Fig. 3H.



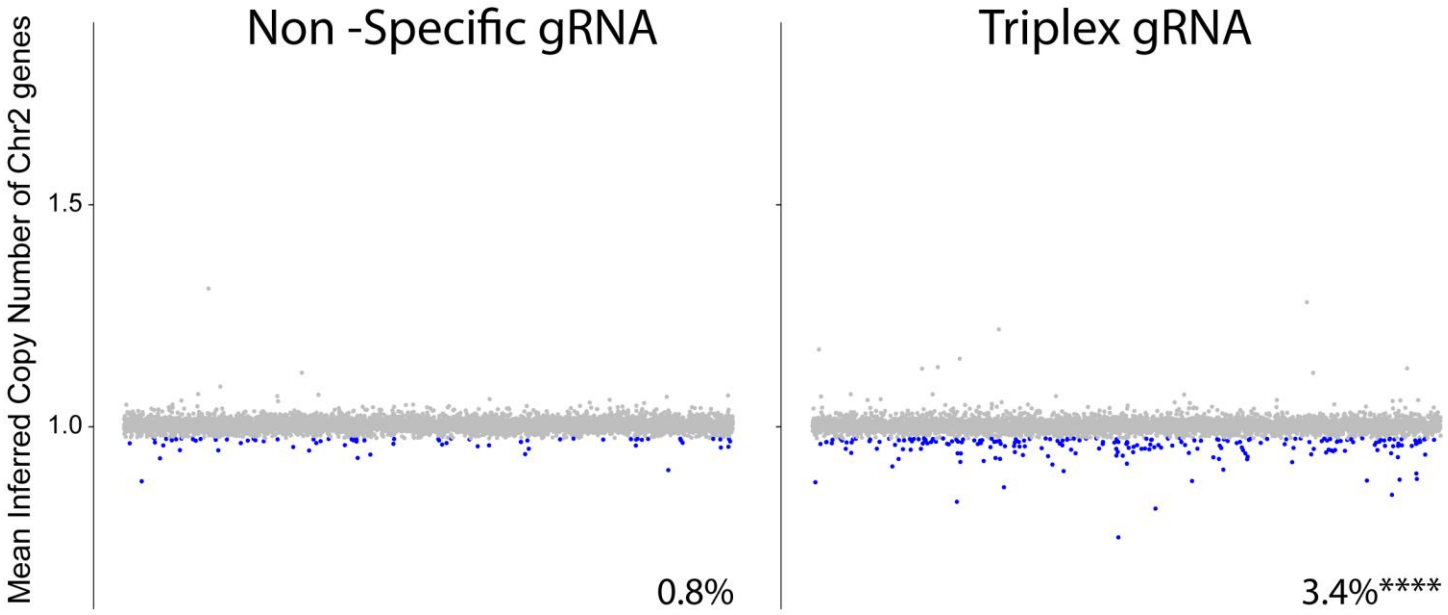
Extended Data Figure 8: A gene set enrichment analysis between cells that have lost or gained a copy of chromosome 14, 4 or 11 days following treatment, compared to cells without these chromosomal aberrations. The 50 'Hallmark' MSigDB gene sets are shown. Gene sets enriched in up-regulated genes are depicted in red and those enriched in down-regulated genes are depicted in blue. Values are scaled to -log(FDR) of the enrichments. The full enrichment scores are shown in Supplementary Table 4.



Extended Data Figure 9: Quantification of InDels produced by CRISPR-Cas9 activity at the TCRα, TCRβ and PDCD1 loci indicates efficient cleavage. **A.** RNA expression of PDCD1 in cells treated with either a non-specific gRNA or a combination of TCRα, TCRβs and PDCD1- targeting gRNAs. ****, p<0.0001, two-sided unpaired Wilcoxon test. RNA expression presented as violin can be found in Fig. 4B. **B.** T7E1 assay for the 4 genomic target loci. Cells treated with a non-Specific gRNA (4 left lanes) are compared to cells treated with the combination of the TCRα, TCRβ and PDCD1-targeting gRNAs (4 right lanes). In each lane, a different locus is analyzed. Calculated efficiency is indicated below. Loading Ladder and relative sizes are indicated in the middle. Unprocessed scan can be found in Supplementary Data 1. **C.** TIDE Analysis for the same experiments as in A. The rows represent (in this order, from top to bottom) the TCRα, TCRβ1, TCRβ2 and PDCD1 loci. In the left panels, the height of each bar corresponds to the rate of sequences having the given number of nucleotides added or deleted. The right panels depict the rate of sequence misalignments at each position of the PCR fragment amplified from the target locus of cells treated with either a non-specific gRNA (black) or a combination of the TCRα, TCRβ and PDCD1- targeting gRNAs (green).



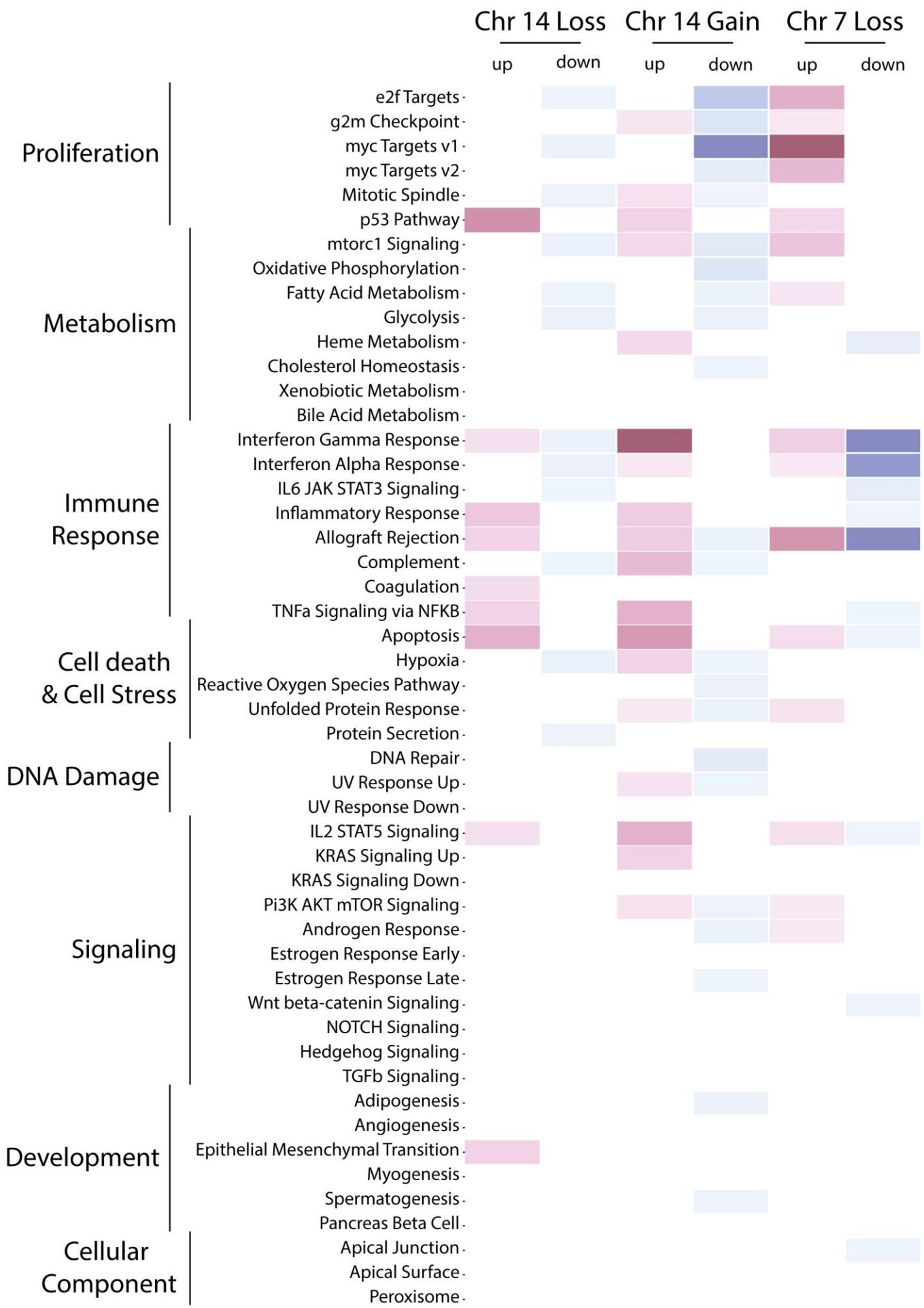
Extended Data Figure 10: A-C. Enrichment in the number of genes with no detected expression among cells identified as having a chromosome 14 loss (A), a chromosome 14 gain (B) or a chromosome 7 truncation (C) (Fig.4D). The x-axis represents the fold-change in the number of genes with no detected expression between cells with or without a chromosome 14 loss or gain or a chromosome 7 truncation, based on the InferCNV analysis (Fig. 4D). The dark gray lines represent the empirical values obtained for each chromosome, except for chromosome 14 or 7 in the respective plots. The orange line is the empirical value for chromosome 14 or chromosome 7. The black bars are the results of 10,000 permutations.



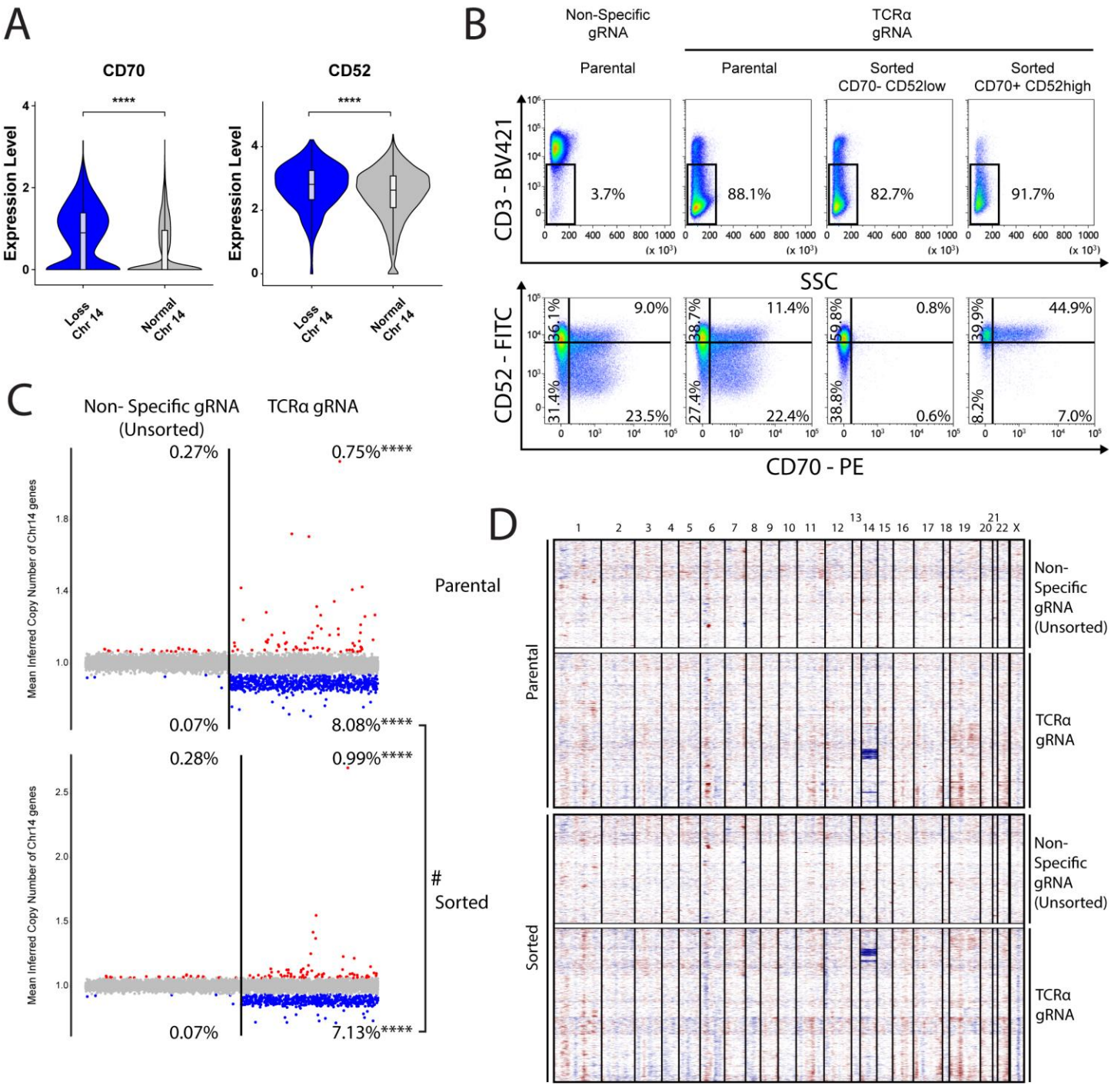
Extended Data Figure 11: Differential gene expression analysis for chromosome 2. Each dot represents the mean inferred copy number of genes coded on chromosome 2 in each cell treated with a non-specific gRNA (left) or a combination of the TCRα, TCRβ and PDCD1- targeting gRNAs (right). Cells are marked with dots spread along the x-

axis. The dots are colored blue when corresponding to cells with a chromosome 2 loss, if their mean inferred gene copy number is >2 standard deviations from the population's mean. n=8619 and 6326 for cells treated with a non-specific gRNA (left) or a combination of the TCRα, TCRβ and PDCD1- targeting gRNAs (right), respectively. ****, $p<0.0001$ for Fisher's exact test comparing chromosome 2 loss between cells treated with the PDCD1 gRNA and cells treated with a non-specific gRNA.

.


Chr 14 Loss
Chr 14 Gain
Chr 7 Loss
up
down
up
down
up
down
Proliferation
e2f Targets
g2m Checkpoint
myc Targets v1
myc Targets v2
Mitotic Spindle
p53 Pathway
Metabolism
mtorc1 Signaling
Oxidative Phosphorylation
Fatty Acid Metabolism
Glycolysis
Heme Metabolism
Cholesterol Homeostasis
Xenobiotic Metabolism
Bile Acid Metabolism
Immune Response
Interferon Gamma Response
Interferon Alpha Response
IL6 JAK STAT3 Signaling
Inflammatory Response
Allograft Rejection
Complement
Coagulation
TNFa Signaling via NFKB
Cell death & Cell Stress
Apoptosis
Hypoxia
Reactive Oxygen Species Pathway
Unfolded Protein Response
Protein Secretion
DNA Damage
DNA Repair
UV Response Up
UV Response Down
Signaling
IL2 STAT5 Signaling
KRAS Signaling Up
KRAS Signaling Down
Pi3K AKT mTOR Signaling
Androgen Response
Estrogen Response Early
Estrogen Response Late
Wnt beta-catenin Signaling
NOTCH Signaling
Hedgehog Signaling
TGFb Signaling
Development
Adipogenesis
Angiogenesis
Epithelial Mesenchymal Transition
Myogenesis
Spermatogenesis
Pancreas Beta Cell
Cellular Component
Apical Junction
Apical Surface
Peroxisome

Extended Data Figure 12: A gene set enrichment analysis between cells that have lost a copy of chromosome 14, or a segment of chromosome 7, to cells without these chromosomal aberrations. The 50 'Hallmark' MSigDB gene sets are shown. Gene sets enriched in up-regulated genes are depicted in red and those enriched in down-regulated genes are depicted in blue. Values are scaled to -log(FDR) of the enrichments. The full enrichment scores are shown in Supplementary Table 4.



Extended Data Figure 13: Using the T cell surface markers CD70 and CD52 to sort out aberrant cells. **A.** The violin plots show gene expression of CD70 (left) and CD52(right), for cells treated with the combination of the TCRα, TCRβ and PDCD1- targeting gRNAs and having a normal gene expression (grey) or a gene expression pattern indicating a chromosome 14 loss(blue). **** $p<0.0001$, unpaired two-tailed Wilcoxon test. **B.** Flow Cytometry of cells at day 4

following treatment with Cas9 and the TCRα targeting gRNA of cells sorted a day earlier for $CD70^-$ and $CD52^{low}$ or $CD70^+$ $CD52^{high}$. Plots marked as "parental" include cells that went through the sorter, but were not sorted for a specific expression of the CD70 or CD52. **C.** Each dot represents the mean inferred copy number of genes coded on chromosome 14 in each cell treated with a non-specific gRNA (left) or the TCRα-targeting gRNA (right). InferCNV analysis performed 4 days following electroporation. Cells are marked with dots spread along the x-axis. The dots are colored red and blue when corresponding to cells with a chromosome 14 gain or loss respectively, if their mean inferred gene copy number is >2 standard deviations (blue and red), from the population's mean. n=8642 and 8970 for "parental" cells after treatment with a non-specific gRNA (left) or the TCRα-targeting gRNA(right), respectively. n=8642 and 7598 for sorted cells after treatment with a non-specific gRNA (left) or TCRα-targeting gRNA (Right), respectively. The same dataset was used as a reference in both the InferCNV analyses (non-specific gRNA, Unsorted). ****, $p<0.0001$ for Fisher's exact test comparing chromosome 14 gain or loss between cells treated with the TCRα gRNA and cells treated with a non-specific gRNA (unsorted). #, $p=0.0257$ for Fisher's exact test comparing chromosome 14 loss between the parental cells treated with the TCRα gRNA and sorted cells treated with the TCRα gRNA. **D.** Heat map depicting gene copy numbers inferred from scRNAseq analysis following sorting of TCRα-targeted cells for either alive ("Parental", above) or $CD70^-$ $CD52^{low}$ expression (Sorted, below). Each line represents an individual cell. The Chromosomes are ordered in columns and the color coding indicates an increase (red) or decrease (blue) in copy number of genes along the chromosomes (x-axis). The same dataset was used as a reference in both the InferCNV analyses (non-specific gRNA, Unsorted).