Cytotherapy 000 (2023) 1–7

Contents lists available at ScienceDirect

CYTOTHERAPY

journal homepage: www.isct-cytotherapy.org

Full-length article

Reference gene selection for clinical chimeric antigen receptor T-cell product vector copy number assays

Jinxia Ma[1], Lipei Shao[1], Tatyana Fuksenko[1], Hui Liu[1], Rongye Shi[1], Anh Dinh[1], Steven L. Highfill[1], Nan Zhang[1], Sandhya R. Panch[2], Robert P. Somerville[1], David F. Stroncek[1], Ping Jin[1,*]

[1] *Center for Cellular Engineering, Department of Transfusion Medicine and Center for Cellular Engineering, National Instates of Health Clinical Center, Bethesda, Maryland, USA*
[2] *Department of Medicine (Hematology Division), University of Washington/Fred Hutchinson Cancer Center, Seattle, Washington, USA*

ARTICLE INFO

Article History:
Received 13 July 2022
Accepted 16 February 2023
Available online xxx

Key Words:
chimeric antigen receptor T-cells
droplet digital PCR
vector copy number

ABSTRACT

Background aims: Reference genes are an essential part of clinical assays such as droplet digital polymerase chain reaction (ddPCR), which measure the number of copies of vector integrated into genetically engineered cells and the loss of plasmids in reprogrammed cells used in clinical cell therapies. Care should be taken to select reference genes, because it has been discovered that there may be thousands of variations in copy number from genomic segments among different individuals. In addition, within the same person in the context of cancer and other proliferative disorders, substantial parts of the genome also can differ in copy number between cells from diseased and healthy people. The purpose of this study was to identify reference genes that could be used for copy number variation analysis of transduced chimeric antigen receptor T cells and for plasmid loss analysis in induced pluripotent stem cells using ddPCR.
Methods: We used The Cancer Genome Atlas (TCGA) to evaluate candidate reference genes. If TCGA found a candidate gene to have low copy number variance in cancer, ddPCR was used to measure the copy numbers of the potential reference gene in cells from healthy subjects, cancer cell lines and patients with acute lymphocytic leukemia, lymphoma, multiple myeloma and human papillomavirus–associated cancers.
Results: In addition to the rPP30 gene, which we have has been using in our copy number assays, three other candidate reference genes were evaluated using TCGA, and this analysis found that none of the four gene regions (AGO1, AP3B1, MKL2 and rPP30) were amplified or deleted in all of the cancer cell types that are currently being treated with cellular therapies by our facility. The number of copies of the genes AP3B1, AGO1, rPP30 and MKL2 measured by ddPCR was similar among cells from healthy subjects. We found that AGO1 had copy number alteration in some of the clinical samples, and the number of copies of the genes AP3B1, MKL2 and rPP30 measured by ddPCR was similar among cells from patients with the cancer cell types that are currently being treated with genetically engineered T-cell therapies by our facility.
Conclusions: Based on our current results, the three genes, AP3B1, MKL2 and rPP30, are suitable for use as reference genes for assays measuring vector copy number in chimeric antigen receptor T cells produced from patients with acute leukemia, lymphoma, multiple myeloma and human papillomavirus–associated cancers. We will continue to evaluate AGO1 on our future samples.

Introduction

Chimeric antigen receptor (CAR) T cells are genetically engineered T cells composed of an antigen-recognizing receptor coupled to signaling molecules that activate T cells. Immunotherapy with CAR-T cells has achieved tremendous success in the treatment of hematological B-cell malignancies [1–4]. Although CAR T cells are effective, they also must be safe. Most CAR T cells used in clinical trials make use of retroviral or lentiviral vectors for the introduction for the CAR encoding transgenes into the T-cell genome. The number of transgenes integrated into each cell must be high enough to make the resulting genetically engineered cell potent but low enough to maintain a minimal probability of integrating near an oncogene, which could result in a greater risk of genotoxicity. Therefore, maintaining

* Correspondence: Pin Jin, Center for Cellular Engineering, Department of Transfusion Medicine and Center for Cellular Engineering, National Instates of Health Clinical Center, 10 Center Drive, Building 10, Room 3N317B, Bethesda, MD 20892.
E-mail address: PJin@mail.cc.nih.gov (P. Jin).

https://doi.org/10.1016/j.jcyt.2023.02.010

transgene copy numbers within a safe and effective range is very important to CAR T-cell therapy. Consequently, the US Food and Drug Administration recommended that the CAR vector copy numbers (VCNs) per cell are less than 5 [5,6] and the copy number of transgenes with the CAR T-cell genome must be accurately measured.

Droplet digital polymerase chain reaction (ddPCR) is widely used as a gene copy number determination assay. ddPCR is an end-point PCR assay that uses primers for the amplification of specific DNA fragments and a probe or a fluorescent dye for the amplicon detection. The mixture of ddPCR is partitioned into water-in-oil droplets. Each droplet is a PCR-competent nano-reactor and will be amplified simultaneously. The number of analyte molecules is calibrated to be lower than the number of droplets, and the random distribution of sample fragments into droplets allows calculating the confidence interval of the concentration using Poisson statistics [7,8]. After amplification, the fluorescence of the droplets can be measured to determine the presence or absence of template molecules. Using different fluorescent dyes allows for the simultaneous detection of two or more targets in one droplet. To use ddPCR to measure the number of copies of a specific gene, vector or plasmid, a reference gene also must be analyzed. A single copy gene is usually used as a reference to determine the copies per cell of the vector or gene of interest that are present in a single reaction [9]. ddPCR is used for a variety of in-process and lot release assays that are used for testing genetically modified cell therapies. It frequently is used to determine the number of vector copies present in each cell. ddPCR has been shown to provide more precision and sensitivity than real-time PCR and is widely used for CAR transgene copy number analysis [9–12]. In addition, when using plasmids to produce induce pluripotent stem cells (iPSCs), ddPCR can be used to determine whether the plasmid is no longer present in iPSC colonies.

The reference gene that is used for ddPCR copy number variance (CNV) analysis of genetically engineered cells administered in clinical trials must be a sensitive and reliable endogenous gene that can serve as an assay control. Such a reference gene should ideally have only a single-copy within the genome and be non-variant in copy number in both healthy subjects and patients with cancer. Theoretically, any region of the genome that is known or strongly expected to be invariant in copy number can be used as a reference control for the ddPCR assay. However, when working with cancer samples, additional attention is required, as it is known that somatic CNV, gene amplification (copy number gain) and deletion (copy number loss) are common in cancer cells and have been shown to correlate to cancer development and progression [13–15]. CNV also plays an important role in human genetic diversity and has been associated with multiple complex disorders [16]; therefore, close attention is required when selecting reference genes for copy number assays because patients with cancer also may have genetic abnormalities.

The Cancer Genome Atlas (TCGA) Program has molecularly characterized more than 20 000 primary cancer and matched normal samples spanning 33 cancer types. This joint effort between the National Cancer Institute and the National Human Genome Research Institute began in 2006, bringing together researchers from diverse disciplines and multiple institutions.

The development of cancer is driven by the acquisition of somatic genetic alterations, and somatic copy-number alterations (SCNAs) are extremely common in cancer. Genomic Identification of Significant Targets in Cancer (GISTIC) is a tool to identify genes targeted by SCNAs. It identifies likely driver SCNAs by evaluating the frequency and amplitude of observed events [17].

In this study, we selected four reference genes and evaluated all the genomic events among cancers using TCGA (http://www.broadinstitute.org/tcga/home). Furthermore, we compared the number of copies of the four candidate reference genes using cell lines, normal donor cells and patient samples. This approach also could be used to identify other reference control genes that could be used for measuring VCN and plasmid loss during the iPSC manufacturing process.

Materials and Methods

Study design

After we identified four candidate reference genes, the genes were screened for reported deletions, duplications and other abnormalities among patients with cancer using the TCGA database. Then, the number of copies of the candidate genes was measured using ddPCR and compared in cells from healthy subjects and patients with cancer. Genes showing little or no abnormalities in the database and whose copy numbers were found to similar to other potential reference genes among the different cells tested were considered a suitable reference gene.

TCGA copy number data analysis

The GISTIC module identifies regions of the genome that are significantly amplified or deleted across a set of samples. Each aberration is assigned a G-score that considers the amplitude of the aberration as well as the frequency of its occurrence across samples. False discovery rate q-values are then calculated for the aberrant regions, and regions with q-values below a user-defined threshold are considered significant. For each significant region, a "peak region" is identified, which is the part of the aberrant region with greatest amplitude and frequency of alteration. In addition, a "wide peak" is determined using a leave-one-out algorithm to allow for errors in the boundaries in a single sample. The "wide peak" boundaries are more robust for identifying the most likely gene targets in the region. Each significantly aberrant region is also tested to determine whether it results primarily from broad events (longer than half a chromosome arm), focal events, or significant levels of both. The GISTIC module reports the genomic locations and calculated q-values for the aberrant regions. It identifies the samples that exhibit each significant amplification or deletion, and it lists genes found in each "wide peak" region.

Two methods have been used for our reference variation analyses: Gene-Centric GISTIC analyses and Cancer-Centric GISTIC analyses. Gene-Centric GISTIC analyses summarizes the significance of copy number alterations affecting a specific gene or transcript. This program analyzes the gene across the entire dataset. This dataset includes many cancer types and the analysis indicates whether the gene falls within a peak region of alteration, the significance of copy number alterations affecting the gene and the frequency of alteration. Cancer-Centric GISTIC analyses summarizes the number of specimens analyzed on the certain cancer types and the significantly aberrant peak regions of the genome are listed, including the boundaries of the peak region (expected to contain the target gene in 95% of cases), the number of genes within the peak region, the significance of copy number alterations affecting this peak region and the frequency of alteration.

Healthy donor cells

Healthy donor cells were collected at the Department of Transfusion Medicine, Clinical Center, National Institutes of Health. Three healthy donor samples were tested using four reference genes selected in this project.

Normal and cancer cell genomic DNA

Genomic DNA used in this study was purchased from OriGene (Rockville, MD, USA; cat. no. CD564202), American Type Culture Collection (Manassas, VA, USA; cat. no. CRM-CCL-185D), Promega (Madison, WI, USA, cat. no. E493A) and Sigma (St. Louis, MO, USA; cat. no. 11691112001).

Genetically engineered T-cell production

The methods used to produce the CAR T cells are described in detail elsewhere [18,19]. In brief, autologous peripheral blood mononuclear cells collected by apheresis were cultured with an anti-CD3 monoclonal antibody or anti-CD3/CD28 beads and IL-2 to induce T-cell proliferation. The cells were then transduced with a γ-retroviral vector or lentiviral vector that encoded a CAR and were expanded in culture of 7–9 days. At the conclusion of the cultures the CAR T cells were sampled for VCN analysis. A similar method was used to manufacture T-cell receptor (TCR)-engineered T cells [20,21] using a γ-retroviral vector that encoded a TCR recognizing human papillomavirus (HPV) 16 E7 oncoprotein. The assay also makes use of untransduced T cells as a control. These cells were processed in the same manner as the genetically engineered T cells except that they were not transduced.

DNA isolation and qualification

Genomic DNA was extracted from cells from healthy subjects and genetically engineered T-cells and patient untransduced control T-cells using a Qiagen DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). The purity and concentration of the DNA samples were measured using a Nanodrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and Qubit spectrophotometer (Thermo Fisher Scientific). For each sample, 260/280 nm absorbance ratio was approximately 1.8–2.0, and 260/230 ratio was approximately 2.0–2.2.

ddPCR VCN assay

An Auto DG QX200 ddPCR system (Bio-Rad Laboratories, Hercules, CA, USA) was used for all ddPCR experiments analyzing extracted DNA for vector copy numbers. Full protocols are detailed in the following manufacturers' materials:

Bio-Rad Laboratories, Inc. (2019) *ddPCR Supermix for Residual DNA Quantification*, retrieved from http://www.bio-rad.com/webroot/web/pdf/lsr/literature/10048259.pdf;

Bio-Rad Laboratories, Inc. (2019) *Automated Droplet Generator Instruction Manual*, retrieved from http://www.bio-rad.com/webroot/web/pdf/lsr/literature/10043138.pdf; and

Bio-Rad Laboratories, Inc. (2019) *QX200 Droplet Reader and QuantaSoft Software Instruction Manual*, retrieved from www.biorad.com/webroot/web/pdf/lsr/literature/10031906.pdf.

To summarize, ddPCR was performed with the QX200 Droplet Digital PCR system (Bio-Rad). The final concentrations of forward and reverse primers were 900 nmol/L and 250 nmol/L for probes. Droplets were generated using a QX200 droplet generator followed by PCR using a T100 thermal cycler (Bio-Rad) with PCR parameters of 95°C for 10 min; 40 cycles of 30 s at 94°C and 1 min at 60°C, followed by 98°C for 10 min. The plate was loaded into the QX200 droplet reader (Bio-Rad) after PCR amplification. The data were analyzed using QuantaSoft software (Bio-Rad). Premade ddPCR assays of rPP30 (assay ID: dHsaCP2500350, probe fluorophore: HEX), AP3B1 (assay ID: dHsaCP2500348, probe fluorophore: HEX) and AGO1(assay ID: dHsaCP2500349, probe fluorophore: HEX) were purchased from Bio-Rad. rPP30 assay (probe fluorophore: FAM) sequence information was adapted from Mazaika and Homsy [22]; MKL2 assay (probe fluorophore: HEX) sequence information was adapted from Li *et al.* [23]; and both were synthesized by IDT (Integrated DNA Technologies, Inc., San Diego, CA, USA).

Results

Selection of reference genes

Any region of the genome that is known or strongly suspected of being invariant in copy number can be used as a reference control. As candidate genomic reference loci, we chose locus with a low CNV in cancer. We evaluated four genes that we had either previously used in our laboratory as a reference gene, or other sources had suggested were suitable for use as a reference gene. The genes were rPP30, AGO1, AP31 and MKL2. The gene rPP30 (accession number NC_000010.11) is ribonuclease P/MRP subunit p30; it has two copies in diploid cells and the genomic sequence is present on chromosome 10q23.31, and it is widely used as a reference for copy number analysis with others [22,24-26] and is currently used in our own analysis of CAR-T cell products. In addition to rPP30, we chose three other genes: AGO1, AP3B1 and MKL2 located on chromosomes 1, 5 and 16, respectively. AGO1 (EIF2C1) (accession number NC_000001.11) is argonaute RISC component 1, and it is located on chromosome 1p34.3. AP3B1 (accession number NC_000005.10) is adaptor-related protein complex 3 subunit beta 1, and it is located on chromosome 5p14.1. MKL2 (MRTFB) (accession number NC_000016.10) is myocardin-related transcription factor B, and it is located on chromosome 16p13.12. All three genes have two copies in diploid cells. MKL2 gene has been used previously as a reference control by our laboratory as well.

Cancer genome database analysis

After the selection, we analyzed TCGA data to see whether any amplifications and deletions were identified in the regions of interest. The dataset we used was 2015-06-01 stddata__2015_04_02 regular peel-off.

Analysis of AGO1 (chr1:36348809-36389899)

The GISTIC analyses showed that AGO1 is not significantly focally amplified across the entire dataset of 10 844 tumors and is not located within a focal peak region of amplification. It is located 3.99 Mb away from the nearest peak region of amplification and is not within the focal peak region of amplification in any of the individual tumor types. AGO1 is not located within a focal peak region of deletion and is located 9.22 Mb away from the nearest peak region of deletion. We didn't find that the region of AGO1 is amplified or deleted in any of the cancer cell types that are currently being treated with genetically engineered cancer immunotherapies manufacture by our laboratory: acute leukemia, lymphoma, multiple myeloma and HPV-associated cancers.

Analysis of AP3B1 (chr 5:77298149-77590579)

GISTIC analyses showed that AP3B1 is not significantly focally amplified across the entire dataset of 10 844 tumors and is not located within a focal peak region of amplification. AP3B1 is located 40.65 Mb away from the nearest peak region of amplification, and it is not significantly focally amplified in any of the 33 individual cancer types analyzed in the dataset. AP3B1 is not located within a focal peak region of deletion and is located 15.86 Mb away from the nearest peak region of deletion. Like AGO1, we didn't find that the AP3B1 region is amplified or deleted in acute leukemia, lymphoma, multiple myeloma and HPV-associated cancers.

MKL2 (chr16:14165195-14360630)

MKL2 is not significantly focally amplified across the entire dataset of 10 844 tumors and is not located within a focal peak region of amplification. MKL2 is not significantly focally amplified in any of the 33 individual cancer types analyzed in TCGA dataset. MKL2 is not significantly focally deleted across the entire dataset of 10 844 tumors and is not located within a focal peak region of deletion. MKL2 is located 6.4 Mb away from the nearest peak region of deletion and is not significantly focally deleted in any of the 33 individual cancer types analyzed in the dataset. Similar to the other reference genes, the MKL2 region doesn't have any deletions or amplifications acute leukemia, lymphoma, multiple myeloma and HPV-associated cancers.

Table 1
Genomic DNA evaluated.

Vendor	Description
OriGene	Liver, carcinoma of liver, hepatocellular
Promega	K562 genomic DNA
American Type Culture Collection	High molecular weight genomic DNA isolated from A-549 lung carcinoma cell line.
Sigma	Human genomic DNA from human blood (buffy coat)

rPP30 (chr 10:92631708-92668312)

GISTIC analyses showed that rPP30 is not significantly focally amplified across the entire dataset of 10 844 tumors and is not located within a focal peak region of amplification. rPP30 is located 15.3 Mb away from the nearest peak region of amplification. rPP30 is not significantly focally amplified in any of the 33 individual cancer types analyzed in our dataset and is not located within a focal peak region of deletion. For the cellular therapies manufactured in our center, we didn't find that the rPP30 region was amplified or deleted in acute leukemia, lymphoma, multiple myeloma and HPV-associated cancers.

Validation of the candidate reference genes using CNV analysis on cells derived from various cancer types

To validate the measurement of CNV using the candidate reference genes and genomic DNA from liver carcinoma, lung carcinoma and K562 cell lines as well as DNAs from healthy donors was analyzed (Table 1). We first analyzed normal DNA samples obtained from Sigma. As expected, ddPCR analysis showed identical copy numbers for the four reference genes (Figure 1).

Next, we tested a cell line known to have a large number of chromosomal abnormalities, K562 cells. The leukemia cell line K562 has hypotriploid karyotype with a modal chromosome number of 67, and there are three normal copies of chromosome 1, three normal copies of chromosome 5, two normal copies of chromosome 10 and three chromosomes 16. In addition, there are also deletions on chromosome 10 [27]. As expected, the analysis ddPCR analysis showed reduced copy numbers of rPP30, which is located on chromosome 10 compared with AGO1, which is located on chromosome 1, AP3B1, which is located on chromosome 5, and MLK2, which is located on chromosome 16 (Figure 1).

Analysis of the liver hepatocellular cells showed similar copies of all four reference genes (Figure 1). The TCGA GISTIC analyses showed that all the chromosome alterations are on chromosome 1 in liver hepatocellular carcinomas, but AGO1 was not on the list of genes which are altered in the liver hepatocellular carcinomas.

Analysis of the lung cancer cell line found that the copy numbers of AGO1 were less than that of AP3B1, rPP30 and MKL2. This indicated that this particular lung cancer cell line has structural chromosomal changes in the AGO1 region (Figure 1). Although TCGA GISTIC analyses didn't show that any of the four candidate genes have alterations in lung cancers, a high frequency of chromosome aberrations has been observed in lung cancer and there are reports of chromosomal changes in lung cancer [28,29] and we found that the copy number of AGO1 was decreased relative to rPP30.

Analysis of VCN of reference genes using healthy subjects

After copy number validation from cancer cell lines, we analyzed cells from healthy subjects. DNA was isolated from peripheral blood mononuclear cells from three healthy subjects and ddPCR was used to assess gene copy numbers of all four reference genes. The measured copy numbers of all four reference genes were similar indicating that all the selected reference genes have two copies in diploid cells (Figure 2).

Analysis of VCN of reference genes in CAR T-cells

We selected three samples from each of the following CAR/TCR products as following: BCMA-CAR (multiple myeloma), CD22-CAR (B-cell malignancies), CD30-CAR (lymphomas), E7-TCR (HPV-associated cancers) and SLAMF-7-CAR (multiple myeloma) (Table 2) for copy number analysis. In addition, control, untransduced T cells produced with each CAR T-cell product also were analyzed. The genes AP3B1, MKL2 and rPP30 showed no variations of the copy numbers (Figure 3), and we also found differences in the copy number of AGO1 compared with the other reference genes on cells derived from various cancer types. We will continue to test AGO1 on future patient samples to see whether this trend holds. The VCN of reference genes in different CAR-T cells summarized in Table 3.

We saw some minor differences in copy number for all three reference genes of interest between CAR T-cell and transduced T-cell samples from the same patients, and we believed that these differences were due to the variation of the amount to DNA in the material used to begin the copy number analysis. To confirm this, we selected the samples for the CAR T-cell and untransduced T-cells that had these copy number variations and performed copy number analysis for the four selected reference genes and calculated/normalized the copy number per cell by using rPP30-FAM primers/probe. We found that all the CAR T-cell and untransduced T-cell samples had similar copy numbers (two copies) per cell (Figure 4).

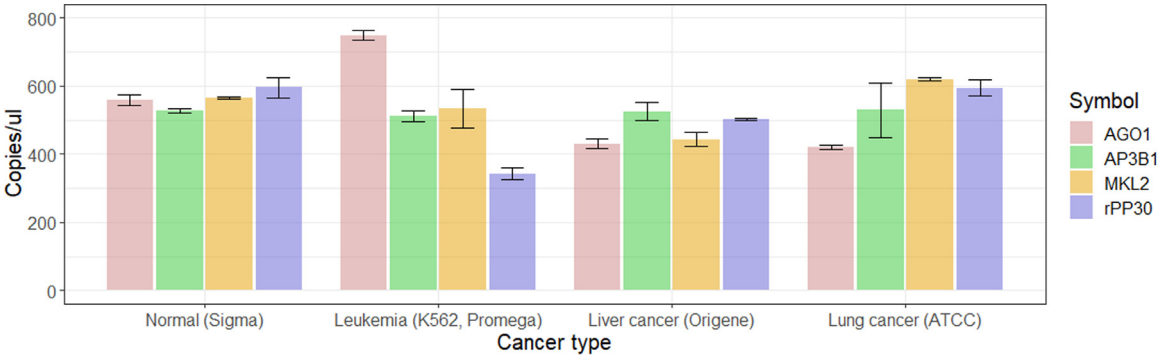

Fig. 1. Gene copy number results from the analysis of DNA from various cancer types. ddPCR analysis of copy numbers of AGO1, AP3B1, MKL2 and rPP30 was performed on liver carcinoma DNA from OriGene, K562 genomic DNA from Promega, lung carcinoma DNA from ATCC and normal human genomic DNA (from human blood buffy coat) from Sigma. The x-axis shows the different DNA samples and y-axis shows the copy numbers per microliter. Normal: normal human genomic DNA; Leukemia: the leukemia cell line K562.

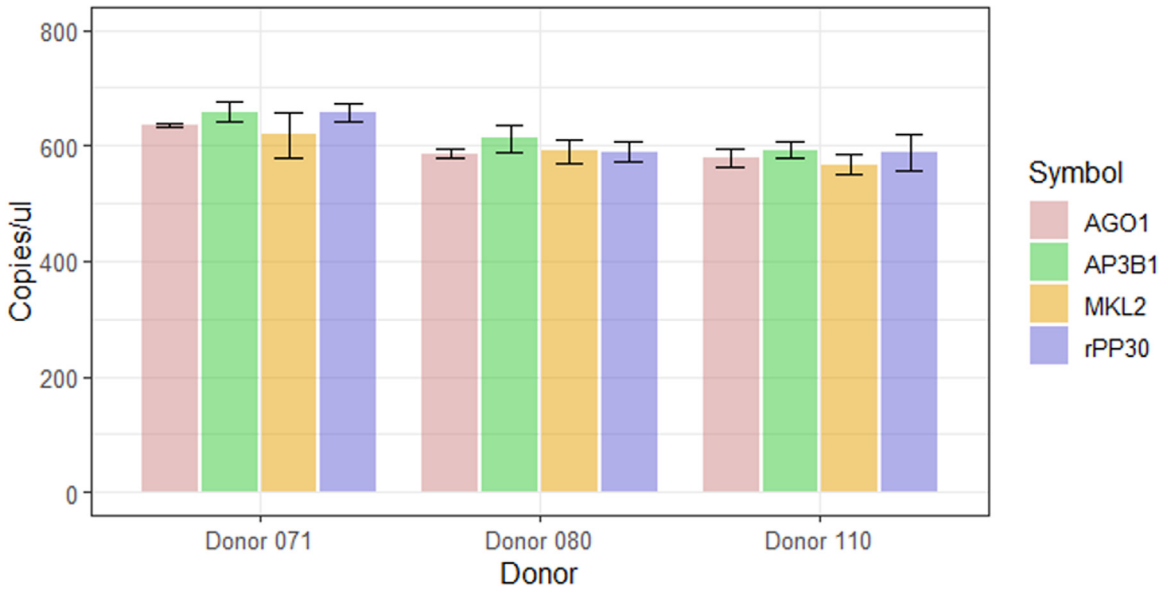

Fig. 2. Gene copy numbers analysis of DNA from normal donors. The number of copies of the four reference genes (AGO1, AP3B1, MKL2 and rPP30) were tested on three normal donor DNA samples: donor 071, donor 080 and donor 110. The x-axis shows the donor name, y-axis shows the copy numbers as copies per microliter.

Table 2
Genetically engineered T-cells evaluated.

	Cell type	Cancer type
1	CD22-CAR T cell	B-cell malignancies
2	E7 TCR-engineered T cell	HPV-associated cancers
3	CD30-CAR T cell	Lymphomas
4	Anti-BCMA CAR T cell	Multiple myeloma
5	SLAMF-7 CAR T cell	Multiple myeloma

CAR, Chimeric antigen receptor; HPV, human papillomavirus.

Discussion

Retroviral and lentiviral vectors can integrate into the host genome and are broadly used to manufacture CAR T cells for clinical application. However, both retro- and lenti-viruses insertional mutagenesis can cause serious side effects, and there is an increased risk of onco-genesis if the VCN per cell is high. Consequently, the accurate and rapid measurement of VCN is an important quality control step required for releasing CAR T-cell products for patient infusion. ddPCR is widely used as a copy number determination assay. This assay requires a reference control gene to quantity transgene copies are present in each cell. The reference control should ideally have a single-copy within the genome and be non-variant in copy number across healthy subjects and patients with cancer. Here, we tested four different reference control genes to see if they were suitable for evaluation of vector copy number of engineered T cells.

We first checked the four reference genes in TCGA, a program generated by National Cancer Institute and the National Human Genome Research Institute, which has molecularly characterized more than 20 000 primary cancers spanning 33 cancer types and matched normal samples. TCGA has Gene-Centric GISTIC Analyses and Cancer-Centric GISTIC Analyses. We didn't find any significant copy number alterations in the loci of the four genes among these 20 000 cancer samples.

We then selected a few cancer types for testing. It is known that common chromosomal aberrations have been identified in lung carcinoma [28,29] and hepatocellular carcinoma [30,31]. Our analysis

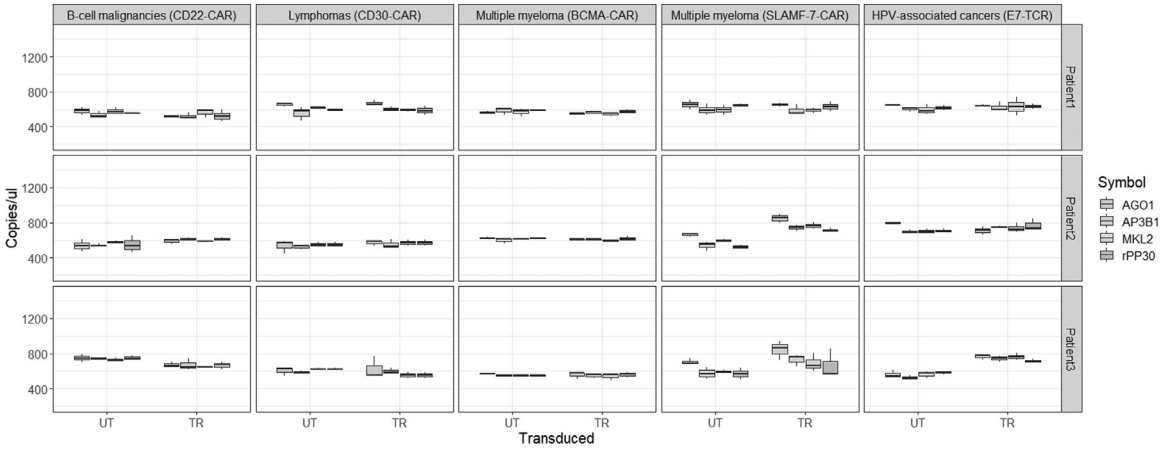

Fig. 3. Gene copy number analysis in genetically engineered T-cell products. Copy number of AGO1, AP3B1, MKL2 and rPP30 was tested in five different types of genetically engineered T cells. For each type of CAR T-cell three sets of CAR T-cells and associated untransduced T-cell samples were tested. HPV E7-TCR engineered T-cells and CD22-, CD30-, BCMA- and SLAMF7-CAR T-cells were evaluated. The x-axis indicates T cells transduced or not. The y-axis represents copy number per microliter. TR, transduced sample; UT, untransduced sample.

Table 3
Gene copy number analysis in genetically engineered T-cell products and untransduced controls.

Assay	Patient	B-cell malignancies (CD22-CAR)				Lymphomas (CD30-CAR)				Multiple myeloma (BCMA-CAR)				Multiple myeloma (SLAMF-7-CAR)				HPV-associated cancers (E7-TCR)			
		UT		TR		UT		TR		UT		TR		UT		TR		UT		TR	
		Mean copies	SD	Mean copies	SD	Mean copies	SD	Mean copies	SD	Mean copies	SD	Mean copies	SD	Mean copies	SD	Mean copies	SD	Mean copies	SD	Mean copies	SD
AGO1	Patient1	584	42.3	515	7.0	653	23.5	671	33.4	566	13.3	548	5.2	655	52.6	653	19.0	647	4.0	642	14.8
	Patient2	541	67.5	589	30.1	535	76.7	571	27.7	625	9.1	610	7.0	659	20.8	850	52.4	796	14.0	710	64.3
	Patient3	750	62.2	672	30.0	604	47.4	625	128.2	572	5.6	557	37.1	705	33.9	847	107.3	565	45.4	765	32.4
AP3B1	Patient1	539	39.3	522	33.0	554	77.8	606	25.9	585	39.4	565	17.3	595	64.5	588	58.1	603	27.2	626	51.3
	Patient2	543	15.5	611	18.5	523	21.7	553	50.1	598	27.6	609	9.8	539	50.1	742	32.5	698	22.6	752	12.7
	Patient3	744	12.2	671	64.3	589	12.7	601	31.9	554	8.7	548	24.4	575	70.7	731	65.5	526	28.9	742	32.5
MLK2	Patient1	579	41.5	564	49.7	618	7.8	593	16.4	572	41.9	547	20.9	595	54.2	590	32.0	596	56.6	632	103.5
	Patient2	577	16.9	590	7.5	552	26.5	573	36.6	615	6.8	595	14.2	598	19.1	766	39.2	701	27.8	739	49.1
	Patient3	730	21.9	646	2.1	627	11.8	552	33.5	549	10.0	540	39.0	596	18.5	690	101.1	564	36.9	766	39.2
rPP30	Patient1	554	5.5	525	65.7	595	14.8	588	49.6	588	6.0	574	22.1	646	18.6	633	52.0	618	31.2	634	30.7
	Patient2	551	97.0	616	14.6	552	26.5	573	36.6	624	3.5	618	25.9	524	22.5	717	17.1	709	17.4	769	68.8
	Patient3	754	22.2	667	42.8	627	11.8	552	33.5	552	9.1	562	29.8	572	66.5	663	167.0	584	19.3	717	17.1

CAR, chimeric antigen receptor; SD, standard deviation; TR, transduced sample; UT, untransduced sample.

found that the lung carcinoma cell line showed altered AGO1 copy numbers, which indicate chromosome structural changes in this type of cancer. The chromosome 1, which the AGO1 gene located, may have deletions in this particular lung cancer cell line. Cell line K562 has 67 chromosomes, and it has many types of genome structural alterations [27]. Our analysis of the number of copies of these four genes in the K562 cell line was consistent with the reported chromosomal abnormalities in the cell line.

We tested the copy numbers of the reference genes using normal donor samples and found, as expected, that there was no significant difference in copy number among the four genes.

The reference genes of AP3B1, MKL2 and rPP30 copies were not alternated in the clinical CAR T-cell and untransduced T-cell samples we tested, which means that these three reference genes can be used for our ongoing clinical assays. AGO1 copy number had variation in some of the patient samples, so this gene should be tested on more patient samples in the future. Since patients with the same type of cancer may have different chromosomal abnormalities, it may be worthwhile to either use multiple reference controls to make sure the reference controls work properly and there are no structural chromosome alterations. Clinical-grade whole-genome sequencing has been studied on colorectal cancer patients [32] and it has the potential to become the standard of care within the clinic. Chromosome karyotyping and clinical-grade whole-genome sequencing may be needed before we start copy number analyses.

Our study has implications for other similar applications. When using plasmids to produce iPSCs, at the completion of the manufacturing process the iPSC clones must be tested to document that the plasmid is no longer present. Currently, GAPDH was used for iPSC plasmid loss assay testing as an internal control. AP3B1, MKL2 and rPP30 could also be used as internal controls for iPSC plasmid loss testing.

In conclusion, based on our current experiment results, no copy number alterations were found in, AP3B1, MKL2 and rPP30 in B-cell malignancies, HPV-associated cancers, lymphomas and multiple myeloma cells manufactured in our center. AGO1 along with the other housekeeping genes will be tested in our laboratory on our future CAR and TCR engineered T-cell products to confirm these results. For new CAR and TCR-engineered T-cell products, multiple reference control genes may be tested on each patient's product to make sure that the reference control genes work properly.

Author Contributions

Conception and design of the study: JM, PJ and DS. Acquisition of data: JM, TF, HL, JM, TF, RS, AD, SP and NZ. Analysis and interpretation of data: JM, LS, RPS and SH. Drafting or revising the manuscript: AD, SP, NZ, RPS and SH. All authors have approved the final article.

Data availability

Patient-related data not included in the paper were generated as part of completed or ongoing clinical trials and may be subject to patient confidentiality. Therefore, these data may be restricted. With regard to sequencing data, raw fastq files will upload into public GEO datasets upon manuscript was accepted. Any other data that support the findings of the study are available from the corresponding author upon reasonable request.

Declaration of Competing Interest

The authors have no commercial, proprietary or financial interest in the products or companies described in this article.

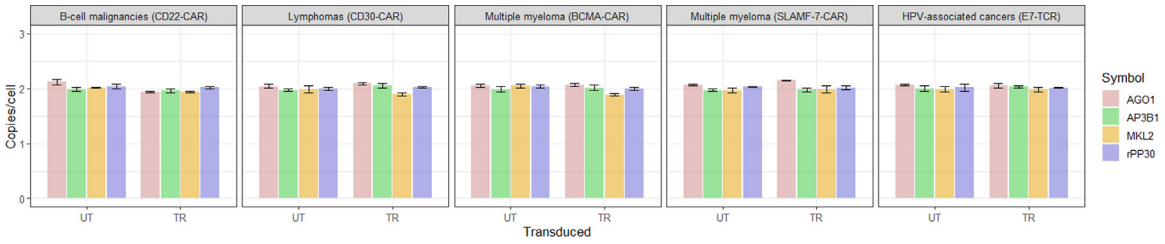

Fig. 4. Gene copy numbers in genetically engineered T cells normalized to rPP30. Genetically engineered T-cell samples and the associated untransduced control T cells were analyzed for copy number of three reference genes (AP3B1, rPP30 and MKL2) and the values were normalized to rPP30. HPV E7 TCR engineered T cells and CD22-, CD30-, BCMA- and SLAMF7-CAR T-cells were evaluated. For each type of genetically engineered T-cell, samples from 3 patients were tested. TR, transduced sample; UT, untransduced sample.

Funding

This research is supported in part by the Intramural Research Program of the National Institutes of Health, Clinical Center.

References

[1] Scheuermann RH, Racila E. CD19 antigen in leukemia and lymphoma diagnosis and immunotherapy. Leuk Lymphoma 1995;18:385–97.
[2] Kochenderfer JN, Wilson WH, Janik JE, Dudley ME, Stetler-Stevenson M, Feldman SA, Maric I, Raffeld M, Nathan DA, Lanier BJ, et al. Eradication of B-lineage cells and regression of lymphoma in a patient treated with autologous T cells genetically engineered to recognize CD19. Blood 2010;116:4099–102.
[3] Pillai V, Muralidharan K, Meng WZ, Bagashev A, Oldridge DA, Rosenthal J, Van Arnam J, Melenhorst JJ, Mohan D, DiNofia AM, et al. CAR T-cell therapy is effective for CD19-dim B-lymphoblastic leukemia but is impacted by prior blinatumomab therapy. Blood Advances 2019;3:3539–49.
[4] Imura Y, Ando M, Kondo T, Ito M, Yoshimura A. CD19-targeted CAR regulatory T cells suppress B cell pathology without GvHD. JCI Insight 2020: 5.
[5] Hollyman D, Stefanski J, Przybylowski M, Bartido S, Borquez-Ojeda O, Taylor C, Yeh R, Capacio V, Olszewska M, Hosey J, et al. Manufacturing validation of biologically functional T cells targeted to CD19 antigen for autologous adoptive cell therapy. J Immunother 2009;32:169–80.
[6] Zhao Y, Stepto H, Schneider CK. Development of the first World Health Organization lentiviral vector standard: toward the production control and standardization of lentivirus-based gene therapy products. Hum Gene Ther Methods 2017;28:205–14.
[7] Basu AS. Digital assays part I: partitioning statistics and digital PCR. SLAS Technol 2017;22:369–86.
[8] Pinheiro LB, Coleman VA, Hindson CM, Herrmann J, Hindson BJ, Bhat S, Emslie KR. Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. Anal Chem 2012;84:1003–11.
[9] Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Anal Chem 2011;83:8604–10.
[10] Lu A, Liu H, Shi R, Cai Y, Ma J, Shao L, Rong V, Gkitsas N, Lei H, Highfill SL, et al. Application of droplet digital PCR for the detection of vector copy number in clinical CAR/TCR T cell products. J Transl Med 2020;18:191.
[11] Lou Y, Chen C, Long X, Gu J, Xiao M, Wang D, Zhou X, Li T, Hong Z, Li C, et al. Detection and quantification of chimeric antigen receptor transgene copy number by droplet digital PCR versus real-time PCR. J Mol Diagn 2020;22:699–707.
[12] Haderbache R, Warda W, Hervouet E, da Rocha MN, Trad R, Allain V, Nicod C, Thieblemeont C, Boissel N, Varlet P, et al. Droplet digital PCR allows vector copy number assessment and monitoring of experimental CAR T cells in murine xenograft models or approved CD19 CAR T cell-treated patients. J Transl Med 2021;19:265.
[13] Shlien A, Malkin D. Copy number variations and cancer susceptibility. Curr Opin Oncol 2010;22:55–63.
[14] Shlien A, Malkin D. Copy number variations and cancer. Genome Med 2009;1:62.
[15] Sansregret L, Vanhaesebroeck B, Swanton C. Determinants and clinical implications of chromosomal instability in cancer. Nat Rev Clin Oncol 2018;15:139–50.
[16] Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. Annu Rev Med 2010;61:437–55.
[17] Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol 2011;12:R41.
[18] Ali SA, Shi V, Maric I, Wang M, Stroncek DF, Rose JJ, Brudno JN, Stetler-Stevenson M, Feldman SA, Hansen BG, et al. T cells expressing an anti-B-cell maturation antigen chimeric antigen receptor cause remissions of multiple myeloma. Blood 2016;128:1688–700.
[19] Stroncek DF, Ren J, Lee DW, Tran M, Frodigh SE, Sabatino M, Khuu H, Merchant MS, Mackall CL. Myeloid cells in peripheral blood mononuclear cell concentrates inhibit the expansion of chimeric antigen receptor T cells. Cytotherapy 2016;18:893–901.
[20] Jin JJ, Gkitsas N, Fellowes VS, Ren JQ, Feldman SA, Hinrichs CS, Stroncek DF, Highfill SL. Enhanced clinical-scale manufacturing of TCR transduced T-cells using closed culture system modules. Journal of Translational Medicine 2018: 16.
[21] Doran SL, Stevanovic S, Adhikary S, Gartner JJ, Jia L, Kwong MLM, Faquin WC, Hewitt SM, Sherry RM, Yang JC, et al. T-cell receptor gene therapy for human papillomavirus-associated epithelial cancers: a first-in-human, phase I/II study. Journal of Clinical Oncology 2019;37:2759-+.
[22] Mazaika E, Homsy J. Digital droplet PCR: CNV analysis and other applications. Curr Protoc Hum Genet 2014;82. 7 24 21-13.
[23] Li D, Li N, Zhang YF. Dropulic: persistent polyfunctional chimeric antigen receptor T cells that target glypican 3 eliminate orthotopic hepatocellular carcinomas in mice (vol 158, 2250, 2020). Gastroenterology 2021;160. 1433-1433.
[24] Kluwe L. Digital PCR for discriminating mosaic deletions and for determining proportion of tumor cells in specimen. Eur J Hum Genet 2016;24:1644–8.
[25] Imaizumi T, Yamamoto-Shimojima K, Yamamoto T. Advantages of ddPCR in detection of PLP1 duplications. Intractable Rare Dis Res 2019;8:198–202.
[26] Wen T, Zhang X, Lippuner C, Schiff M, Stuber F. Development and evaluation of a droplet digital PCR assay for 8p23 beta-defensin cluster copy number determination. Mol Diagn Ther 2021;25:607–15.
[27] Naumann S, Reutzel D, Speicher M, Decker HJ. Complete karyotype characterization of the K562 cell line by combined application of G-banding, multiplex-fluorescence in situ hybridization, fluorescence in situ hybridization, and comparative genomic hybridization. Leuk Res 2001;25:313–22.
[28] Berrieman HK, Ashman JN, Cowen ME, Greenman J, Lind MJ, Cawkwell L. Chromosomal analysis of non-small-cell lung cancer by multicolour fluorescent in situ hybridisation. Br J Cancer 2004;90:900–5.
[29] Grigorova M, Lyman RC, Caldas C, Edwards PA. Chromosome abnormalities in 10 lung cancer cell lines of the NCI-H series analyzed with spectral karyotyping. Cancer Genet Cytogenet 2005;162:1–9.
[30] Wong N, Lai P, Pang E, Leung TW, Lau JW, Johnson PJ. A comprehensive karyotypic study on human hepatocellular carcinoma by spectral karyotyping. Hepatology 2000;32:1060–8.
[31] Midorikawa Y, Yamamoto S, Ishikawa S, Kamimura N, Igarashi H, Sugimura H, Makuuchi M, Aburatani H. Molecular karyotyping of human hepatocellular carcinoma using single-nucleotide polymorphism arrays. Oncogene 2006;25:5581–90.
[32] Stodolna A, He M, Vasipalli M, Kingsbury Z, Becq J, Stockton JD, Dilworth MP, James J, Sillo T, Blakeway D, et al. Clinical-grade whole-genome sequencing and 3′ transcriptome analysis of colorectal cancer patients. Genome Med 2021;13:33.