

LIGAND DISCOVERY AND GENE DELIVERY VIA RETROVIRAL SURFACE DISPLAY

RELATED APPLICATION

This application claims the benefit under 35 U.S.C. § 119(e) of the filing date of U.S. Provisional Application Serial No. 62/851,889, entitled “LIGAND DISCOVERY AND GENE DELIVERY VIA RETROVIRAL SURFACE DISPLAY” and filed on May 23, 2019, the entire contents of which are incorporated herein by reference.

BACKGROUND

It has been well established that retroviruses (*e.g.*, lentiviruses) can have their natural tropism redirected to targets of interest via pseudotyping a virus to express the envelope protein of a different virus. This is most commonly accomplished via pseudotyping with the VSV glycoprotein, which targets the LDL receptor and therefore enables viral entry into a wide range of cells. Recently, groups have shown that if lentiviruses are pseudotyped with envelope proteins from paramyxoviruses such as measles virus or nipah virus, and if mutations are created that abolish native tropism, C-terminal fusions to these viruses allow receptor-mediated entry to the target cell of choice.

SUMMARY

Herein, the inventors have surprisingly demonstrated that a combination of mutations to abolish native function (*e.g.*, tropism) and overexpression of a second membrane protein allows for that second protein to function as the basis for viral entry. These discoveries, as described herein, enable new and innovative methodologies, for example, to screen of cells that are notoriously challenging to screen for specific antigens and function (*e.g.*, T cells), and to deliver nucleic acids to target cells in a target-specific manner.

Some aspects of the disclosure provide compositions of a retrovirus (*e.g.*, a lentivirus) comprising (i) a nucleic acid comprising a non-viral membrane-bound protein comprising a structure: S-ETD-MBD-IRES-R, wherein S encodes a signal sequence, ETD encodes an extracellular targeting domain; MBD encodes a membrane-bound domain, IRES encodes an internal ribosome entry site, and R encodes a reporter; and (ii) a mutated viral envelope protein comprising at least one mutation that diminishes its native function.

Some aspects of the disclosure provide compositions of a retrovirus (*e.g.*, a lentivirus), comprising (i) a CD80 protein domain; and (ii) a mutated viral envelope protein comprising at least one mutation that diminishes its native function. The CD80 protein domain may be an extracellular domain. In some embodiments, the CD80 extracellular domain binds to a receptor on a target cell.

Some aspects of the disclosure provide methods of screening a population of cells, the method comprising (i) providing a retrovirus (*e.g.*, a lentivirus) comprising a viral envelope protein comprising at least one mutation that diminishes its native function, a non-viral membrane-bound protein comprising a membrane-bound domain and an extracellular targeting domain, and a nucleic acid encoding a reporter; (ii) combining the retrovirus with a population of cells; and (iii) sorting the population of cells based on the presence or absence of the reporter. In some embodiments, the retrovirus (*e.g.*, a lentivirus) comprises a nucleic acid comprising a non-viral membrane-bound protein comprising a structure: S-ETD-MBD-IRES-R, wherein S encodes a signal sequence, ETD encodes an extracellular targeting domain; MBD encodes a membrane-bound domain, IRES encodes an internal ribosome entry site, and R encodes a reporter.

In some embodiments, the cells are somatic cells (*e.g.*, antigen-specific cells, *e.g.*, T cells or B cells). In some embodiments, the cells are isolated from a subject (*e.g.*, a human subject). In some embodiments, the cells are isolated from the blood or a tumor of a subject. In some embodiments, the cells are maintained in liquid culture prior to being combined with the retrovirus.

In some embodiments, the viral envelope protein is a VSV-G envelope protein, a measles virus envelope protein, a nipah virus envelope protein, or a cocal virus G protein. A VSV-G envelope protein may be a mutated at one or more of any one of H8, K47, Y209, and/or R354. A measles virus envelope protein may be a mutated at one or more of any one of Y481, R533, S548, and/or F549. A nipah virus envelope protein may be a mutated at one or more of any one of E501, W504, Q530, and/or E533. A cocal virus G protein may be a mutated at K64 and/or R371.

In some embodiments, the non-viral membrane-bound protein comprises a Major Histocompatibility Complex (MHC) protein. In some embodiments, the extracellular targeting domain is a protein (*e.g.*, interleukin-13), a peptide, or an antibody (*e.g.*, anti-CD19

antibody, an anti-TCR antibody, or an anti-CD3 antibody). In some embodiments, the reporter is a fluorescent protein (*e.g.*, green fluorescent protein, yellow fluorescent protein, red fluorescent protein) or an antibiotic resistance marker. In some embodiments, a linker is positioned between the membrane-bound domain and the extracellular targeting domain. A linker may be a rigid linker (*e.g.*, a PDGFR stalk or a CD8 α stalk), a flexible linker (*e.g.*, comprising an amino acid sequence comprising GAPGAS (SEQ ID NO: 5) or GGGGS (SEQ ID NO: 7)), or an oligomerized linker (*e.g.*, an IgG4 hinge or an amino acid sequence that can form a tetrameric coiled coil).

In some embodiments, the retrovirus is combined with the population of cells in (ii) for one minute to seventy-two hours and at a temperature ranging from 4°C to 42°C. In some embodiments, the retrovirus and the population of cells are combined in (ii) in the presence of (a) a cell culture media, optionally RPMI or DMEM cell culture media; (b) a buffered saline solution, optionally phosphate-buffered saline or HEPES-buffered saline; and/or (c) an enhancer of retroviral transduction, optionally heparin sulfate, polybrene, protamine sulfate, and/or dextran. In some embodiments, the extracellular targeting domain is capable of binding to a cognate protein (*e.g.*, a protein receptor) that is present on the cell surface of a subset of the population of cells. In some embodiments, the population of cells is washed between (ii) and (iii) (*e.g.*, using phosphate-buffered saline (PBS), *e.g.*, to remove the retrovirus from the population of cells). In some embodiments, sorting the population of cells is performed using fluorescence-activated cell sorting, single-cell next-generation sequencing, or antibiotic selection.

In some embodiments, the methods further comprise a second retrovirus, wherein the second retrovirus comprises a different extracellular targeting domain and/or a different reporter compared to the first retrovirus.

Other aspects of the disclosure provide methods of delivering a nucleic acid (*e.g.*, a gene of interest, *e.g.*, that encodes a protein) to a cell, the method comprising (i) providing a retrovirus comprising the nucleic acid, a viral envelope protein comprising at least one mutation that diminishes its native function, and a non-viral membrane-bound protein comprising an extracellular targeting domain that is capable of binding to a cognate ligand of the cell; and (ii) contacting the retrovirus with the cell, thereby delivering the nucleic acid to the cell. In some embodiments, the retrovirus enters or infects the cell during (ii).

Some aspects of the disclosure provide methods of delivering a nucleic acid to a cell, the method comprising (i) providing a retrovirus comprising the nucleic acid, a viral envelope protein comprising at least one mutation that diminishes its native function, and a CD80 protein domain; and (ii) contacting the retrovirus with the cell, thereby delivering the nucleic acid to the cell.

Yet other aspects of the disclosure provide methods of detecting an interaction between a retrovirus and an cell, the method comprising: (i) contacting a sample comprising the retrovirus and an cell with an antibody, wherein the retrovirus comprises a viral envelope protein comprising at least one mutation that diminishes its native function, a non-viral membrane-bound protein comprising an extracellular targeting domain, and wherein the antibody binds to the extracellular targeting domain of the retrovirus; (ii) optionally removing unbound antibody from the sample; and (iii) imaging the sample to detect whether the antibody-retrovirus complex is bound to the cell.

In some embodiments, the antibody further comprises a fluorescent label, optionally wherein the antibody is covalently linked to a fluorescent label. In some embodiments, the sample is imaged in (iii) using confocal or fluorescence microscopy.

Some aspects of the disclosure provide libraries of retroviruses, comprising a plurality of unique retroviruses, wherein each unique retrovirus comprises a viral envelope protein comprising at least one mutation that diminishes its native function, a non-viral membrane-bound protein comprising a membrane-bound domain and an extracellular targeting domain (*e.g.*, comprising at least 5, at least 10, at least 15, at least 20, or at least 50 amino acids), and a nucleic acid encoding a reporter, and wherein each unique retrovirus comprises a different and unique extracellular targeting domain.

In some embodiments, a library is capable of being screened against a population of antigen-specific cells, optionally wherein the antigen-specific cells are B cells or T cells. A library may comprise at least 10^2 , at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 , at least 10^8 , at least 10^9 , or at least 10^{10} unique retroviruses. In some embodiments, each different and unique extracellular targeting domain is generated through site-directed mutagenesis.

Some aspects of the disclosure provide populations of cells, wherein a subset of the population of cells contain a retrovirus comprising: a viral envelope protein comprising at

least one mutation that diminishes its native function, a non-viral membrane-bound protein comprising a membrane-bound domain and an extracellular targeting domain, and a nucleic acid encoding a reporter. In some embodiments, a subset of the population of cells (*e.g.*, antigen-specific cells, *e.g.*, B cells or T cells) contains a retrovirus as described herein. In some embodiments, the subset of the population of cells contains a retrovirus inside each cell of the subset. The subset of the population that contain the retrovirus may be isolated and/or sorted from the cells of the population that do not contain the retrovirus.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 depicts an exemplary schematic of a retrovirus of the current invention interacting with a cognate ligand of a target cell.

FIG. 2 depicts an exemplary schematic of a protein comprising a signal sequence (SP), an extracellular targeting domain (peptide), a membrane-bound domain (MHC), an internal ribosome entry site (IRES), and a reporter (GFP); and the ability of a retrovirus comprising said protein to interact with a target cell.

FIGs. 3A-3B depict graphs showing retroviral entry into a target cell. **FIG. 3A** demonstrates that the binding interaction of an extracellular targeting domain present in a retrovirus to a cognate ligand of a target cell leads to entry of the retrovirus and subsequent expression of reporter (GFP). **FIG. 3B** demonstrates that contacting a retrovirus with a non-target cell does not lead to entry of the retrovirus.

FIG. 4 depicts graphs showing the ability of retroviral entry into target cells at variable affinities between an extracellular targeting domain and its cognate ligand.

FIGs. 5A-5B depict graphs showing that T cell signaling is active during retroviral entry or infection. **FIG. 5A** demonstrates that the binding interaction of an extracellular targeting domain present in a retrovirus to a cognate ligand of a target cell leads to entry of the retrovirus and T cell signaling activation. **FIG. 5B** demonstrates that contacting a retrovirus with a non-target cell does not lead to entry of the retrovirus or T cell signaling activation.

FIG. 6 depicts an exemplary schematic of a pMHC retroviral library screening methodology.

FIGs. 7A-7B depict graphs showing retroviral entry into a target cell. **FIG. 7A**

demonstrates that the binding interaction of an interleukin-13 (IL-13) protein domain present in a retrovirus to an IL-13 protein receptor of a target cell leads to entry of the retrovirus and subsequent expression of reporter (GFP). **FIG. 7B** demonstrates that contacting a retrovirus comprising an IL-13 protein domain with a non-target cell that does not express an IL-13 protein receptor does not lead to entry of the retrovirus.

FIG. 8 depicts three exemplary retroviruses, each comprising a different viral envelope protein.

FIG. 9 depicts the ability of a retrovirus comprising a Nipah viral envelope protein and an anti-CD19 scFv antibody to infect and enter CD19⁺ Raji cells.

FIG. 10 depicts the ability of a retrovirus comprising a mutated VSV-G viral envelope protein and an anti-CD19 scFv antibody to infect and enter CD19⁺ Raji cells, while not infecting CD19⁻ Jurkat cells. A retrovirus comprising a wild-type VSV-G viral envelope protein and an anti-CD19 scFv antibody infects both CD19⁺ Raji cells and CD19⁻ Jurkat cells.

FIG. 11 depicts exemplary linkers for use in nucleic acids and proteins comprising membrane-bound and extracellular targeting domains.

FIG. 12 depicts graphs showing the effect of short protein linkers (PDGFR; GAPGAS (SEQ ID NO: 5); CAR ECD; top, left to right) and long protein linkers (CD8 α ; Hinge domain; tetrameric coiled coil; bottom, left to right) on retroviral entry.

FIG. 13 depicts graphs showing that retroviruses comprising mutated VSV-G viral envelope proteins activate T cell signaling (right panel) and that retroviruses comprising wild-type VSV-G viral envelope proteins do not activate T cell signaling (left panel), as determined by expression of CD69.

FIG. 14 depicts graphs showing that retroviruses having variable affinities between their extracellular targeting domains and a T cell cognate ligand are able to activate T cell signaling.

FIG. 15 depicts graphs showing the ability of retroviral entry into target cells at variable affinities, from picomolar to micromolar binding affinities, between an extracellular targeting domain and its cognate ligand.

FIGS. 16A-16B depict graphs demonstrating the ability of retroviral entry into target cells at varying concentrations of virus. **FIG. 16A** shows the ability of NYESO-1 displaying

lentiviruses to transduce IG4-expressing T cells. SL-1 displaying lentiviruses do not transduce IG4-expressing T cells. **FIG. 16B** shows the mean fluorescence provided by the interaction between NYESO-1 displaying lentiviruses and IG4-expressing T cells.

FIG. 17 depicts a graph that demonstrates the impact of protein linker oligomerization on retroviral transduction.

FIGs. 18A-18B depict the ability to detect a binding interaction between a retrovirus comprising an anti-CD19 scFv antibody and a target cell comprising CD19 using a fluorescently labeled antibody that binds to the retrovirus. **FIG. 18A** depicts an exemplary binding schematic. **FIG. 18B** depicts graphs showing ability of retroviral constructs to interact with a target cell.

FIG. 19 depicts the ability of SL9 peptide displaying lentiviruses to transduce 868 TCR-expressing T cells ('868') and the ability of CMV NLV peptide displaying lentiviruses to transduce C7 TCR-expressing T cells ('CMV C7').

FIG. 20 depicts the ability of NYESO-1 displaying lentiviruses to selectively transduce T cells (target cells) expressing an IG4 T cell receptor (TCR) variant (binding affinities of ~26 pM or ~32 μ M for NYESO-1) relative to cells not expressing an IG4 TCR (off-target) at ratios of off-target cells to target cells as high as 100,000:1.

FIG. 21 depicts the ability of lentiviruses displaying varying targeting peptides (CMV NLV; EBV; SL9; GL9; NYESO-1) to selectively transduce T cells that express targeting peptide-specific TCRs (on-target) relative to T cells that do not express the targeting peptide-specific TCRs (off-target).

FIGs. 22A-22B depict the ability of varying lentiviruses (GL9 peptide targeting; NYESO-1 peptide targeting; wild-type VSV) to transduce primary T cells that express GL9-specific TCRs and T cells that express an IG4 TCR variant. **FIG. 22A** depicts the transduction rate for each lentivirus. **FIG. 22B** depicts graphs showing FACS data for GL9 virus and NYESO-1 virus.

FIG. 23 depicts graphs showing that primary NYESO-1-reactive cells are specifically infected.

FIG. 24 depicts graphs showing that NYESO-1 targeted virus specifically infects expanded primary cells.

FIG. 25 depicts graphs showing the ability of a retrovirus comprising a mutated VSV-

G viral envelope protein and a CD80 domain to specifically infect Jurkat T cells, relative to a B cell line.

FIG. 26 depicts the ability of retroviruses comprising a mutated VSV-G viral envelope protein and either an anti-TCR-alpha beta antibody (H57) or an anti-CD3 antibody (C11) to infect of TCR-transduced 58^{-/-} mouse cell lines.

FIG. 27 depicts the ability of a retrovirus comprising either a mutated VSV-G viral envelope protein or a mutated cocl viral envelope protein; and an anti-CD19 scFv antibody to infect and enter CD19⁺ Raji cells, while not infecting CD19⁻ Jurkat cells. A retrovirus comprising a wild-type VSV-G viral envelope protein and an anti-CD19 scFv antibody infects both CD19⁺ Raji cells and CD19⁻ Jurkat cells.

DETAILED DESCRIPTION OF THE INVENTION

Herein are provided new and innovative methods, for example, to screen cells that are notoriously challenging to screen for specific antigens and function (*e.g.*, T cells), and to deliver nucleic acids to target cells in a target-specific manner. In some embodiments, described herein are systems that enable, for example, repertoire-scale analysis of T cell receptor (TCR)-peptide-Major Histocompatibility Complex (pMHC) specificity, a previously intractable bottleneck, as previously described methods required considerable effort to determine what a single T cell clone can recognize (*e.g.*, as in a typical immune response). In some embodiments, described herein are retrovirus-based systems that repurpose viral tropism as a method of selecting for molecular interactions and replace the binding functions of wild-type virus surface proteins with those of protein variants of interest, for example, by encoding these protein variants on the corresponding transfer plasmid used to make the virus, thereby ensuring that the resulting virus displays the protein variant on its surface and packaging the corresponding genetic sequence. As such, when the virus enters a target cell (*e.g.*, bearing a receptor that binds the displayed extracellular targeting domain of the protein variant), cell entry results in integration of the genetic sequence of the displayed protein into the genome of the target cell.

Previous approaches for studying T cell specificity required a combination of generated T cell lines, recombinant expression of T cell receptors, and/or the individual validation of T cell binding or activity via a candidate antigen-based approach. Each of these

elements provided an inherent limitation in the throughput of T cells or antigens screened. For example, yeast display based methods to deorphanize T cell receptors alleviated the bottleneck of number of antigens examined (with the ability to screen $>10^8$ ligands), but were still severely limited by the need to recombinantly express TCRs. The current strategies of the present invention described herein represents a tremendous advance in the study of T cell specificity and screening of T cells by allowing for screening of $>10^8$ ligands and without need for recombinant TCR expression.

Retroviruses

Described herein are retroviruses comprising a viral envelope protein comprising at least one mutation that diminishes its native function, a non-viral membrane-bound protein comprising a membrane-bound domain and an extracellular targeting domain, and a nucleic acid encoding a reporter. In some embodiments, a retrovirus comprises a viral envelope protein comprising at least one mutation that diminishes its native function and a non-viral membrane-bound protein comprising a membrane-bound domain and an extracellular targeting domain.

The retrovirus disclosed herein comprise one or more elements derived from a retroviral genome (naturally-occurring or modified) of a suitable species. Retroviruses include 7 families: alpharetrovirus (Avian leucosis virus), betaretrovirus (Mouse mammary tumor virus), gammaretrovirus (Murine leukemia virus), deltaretrovirus (Bovine leukemia virus), epsilonretrovirus (Walleye dermal sarcoma virus), lentivirus (Human immunodeficiency virus 1), and spumavirus (Human spumavirus). Six additional examples of retroviruses are provided in U.S. Patent No. 7,901,671.

In some embodiments, a retrovirus is a lentivirus. Lentivirus is a genus of retroviruses that typically gives rise to slowly developing diseases due to their ability to incorporate into a host genome. Modified lentiviral genomes are useful as viral vectors for the delivery of a nucleic acids to a host cell. Host cells can be transfected with lentiviral vectors, and optionally additional vectors for expressing lentiviral packaging proteins (e.g., VSV-G, Rev, and Gag/Pol) to produce lentiviral particles in the culture medium.

Retrovirus and lentivirus constructs are well known in the art and any suitable retrovirus can be used to construct the retrovirus (or a plurality or library of retroviruses) as

described herein. Non-limiting examples of retrovirus constructs include lentiviral vectors, human immunodeficiency viral (HIV) vector, avian leucosis viral (ALV) vector, murine leukemia viral (MLV) vector, murine mammary tumor viral (MMTV) vector, murine stem cell virus, and human T cell leukemia viral (HTLV) vector. These retrovirus constructs comprise proviral sequences from the corresponding retrovirus.

The retrovirus described herein may comprise the viral elements such as those described herein from one or more suitable retroviruses, which are RNA viruses with a single strand positive-sense RNA molecule. Retroviruses comprise a reverse transcriptase enzyme and an integrase enzyme. Upon entry into a target cell, retroviruses utilize their reverse transcriptase to transcribe their RNA molecule into a DNA molecule. Subsequently, the integrase enzyme is used to integrate the DNA molecule into the host cell genome. Upon integration into the host cell genome, the sequence from the retrovirus is referred to as a provirus (e.g., proviral sequence or provirus sequence). The retroviral vectors described herein may further comprise additional functional elements as known in the art to address safety concerns and/or to improve vector functions, such as packaging efficiency and/or viral titer. Additional information may be found in US20150316511 and WO2015/117027, the relevant disclosures of each of which are herein incorporated by reference for the purpose and subject matter referenced herein. Additional information for lentiviruses can be found in, e.g., WO2019/056015, the relevant disclosures of which are incorporated by reference herein for this particular purpose.

In some embodiments, lentiviruses are able to be targeted to target-specific cells via the pMHC-TCR interaction or any other protein-protein cell-to-cell interaction. In some embodiments, T cells with a known and relevant specificity can be enhanced (in the case of cancer or infection) or ablated (in the case of autoimmunity) without affecting other T cells, dramatically limiting the risk of off-target effects. In some embodiments, lentiviruses may encode an extracellular domain to target any other surface-expressed molecule on a target cell.

Viral envelope protein

The retroviruses described herein comprise a viral envelope protein comprising at least one mutation that diminishes its native function (e.g., wild-type function of a non-

mutated viral envelope protein). In some embodiments, a viral envelope protein is any viral envelope protein of any retrovirus (*e.g.*, lentivirus). A viral envelope protein may be a VSV-G envelope protein, a measles virus envelope protein, a nipah virus envelope protein, or a cocal virus G protein. In some embodiments, the native function that is diminished by a mutation of a viral envelope protein is viral tropism (*e.g.*, ability to infect cells, bind to cells, *etc.*)

In some embodiments, a viral envelope protein comprising at least one mutation that diminishes its native function is a mutated VSV-G envelope protein. In some embodiments, a viral envelope protein comprising at least one mutation that diminishes its native function is a mutated measles virus envelope protein. In some embodiments, a viral envelope protein comprising at least one mutation that diminishes its native function is a mutated nipah virus envelope protein. In some embodiments, a viral envelope protein comprising at least one mutation that diminishes its native function is a mutated cocal virus G protein.

In some embodiments, a mutated VSV-G envelope protein comprises a mutation at H8, K47, Y209, and/or R354. In some embodiments, a mutated VSV-G envelope protein comprises a H8A, K47A, K47Q, Y209A, R354A, and/or R354Q mutation. In some embodiments, a mutated VSV-G envelope protein is as described in Nikolic et al., “Structural basis for the recognition of LDL-receptor family members by VSV glycoprotein.” *Nature Comm.*, 2018, 9:1029, the relevant disclosures of which are incorporated by reference herein for this particular purpose.

In some embodiments, a mutated measles virus envelope protein comprises a mutation at Y481, R533, S548, and/or F549. In some embodiments, a mutated measles virus envelope protein comprises a Y481A, R533A, S548L, and/or F549S mutation.

In some embodiments, a mutated Nipah virus envelope protein comprises a mutation at E501, W504, Q530, and/or E533. In some embodiments, a mutated measles virus envelope protein comprises a E501A, W504A, Q530A, and/or E533A mutation.

In some embodiments, a mutated cocal virus G protein comprises a mutation at K64 and/or R371. In some embodiments, a mutated cocal virus G protein comprises a mutation at K64Q and/or R371A.

In some embodiments, the mutated envelope protein is derived from any other enveloped virus including but not limited to baculovirus, herpes simplex virus (HSV),

cytomegalovirus (CMV), lymphocytic choriomeningitis virus (LCMV), Epstein-Barr virus (EBV), vaccinia virus, Hepatitis A, B, or C virus, vaccinia virus, alphavirus, dengue virus, yellow fever virus, Zika virus, influenza virus, hantavirus, Ebola virus, rabies virus, human immunodeficiency virus (HIV), coronavirus, and other members of rhabdoviridae.

In some embodiments, a viral envelope protein comprising at least one mutation comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more mutations. In some embodiments, a viral envelope protein comprising at least one mutation comprises a nucleotide sequence and/or amino acid sequence that is at least 50%, 60%, 70%, 80%, 90%, 95%, or 97% identical to a wild-type viral envelope protein. In some embodiments, a viral envelope protein comprising at least one mutation that diminishes its native function retains less than 95%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, or 10% of the function of a wild-type viral envelope protein. In some embodiments, a viral envelope protein comprising at least one mutation lacks all of its native function. In some embodiments, a retrovirus comprising a viral envelope protein comprising at least one mutation that diminishes its native function comprises less than 95%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, or 10% of the cellular infectivity of a retrovirus comprising a wild-type viral envelope protein.

Non-viral membrane-bound protein

The retroviruses described herein comprise a non-viral membrane-bound protein. A non-viral membrane-bound protein may comprise a membrane-bound domain and an extracellular targeting domain. In some embodiments, a non-viral membrane-bound protein is a chimeric protein comprising sequences from at least two different proteins. In some embodiments, a non-viral membrane-bound protein is a full-length or truncated protein comprising sequence from a single protein.

A membrane-bound domain is a protein or peptide that has an amino acid sequence that enables the protein or peptide to be fully or partially embedded or associated with the membrane (*e.g.*, envelope) of the retrovirus. In some embodiments, a membrane-bound domain enables presentation and delivery of the extracellular targeting domain to the extracellular environment. In some embodiments, a membrane-bound domain comprises an intracellular domain, a transmembrane domain, and/or an extracellular domain. In some embodiments, a membrane-bound domain comprises an intracellular domain and a

transmembrane domain. In some embodiments, the membrane-bound domain comprises a Major Histocompatibility Complex (MHC) protein or fragment thereof. A MHC protein may be a Class I or Class II MHC protein.

In some embodiments, a membrane-bound domain comprises 10-50, 10-100, 25-100, 50-200, 50-150, 100-500, 100-250, 250-500, or any reasonable number of total amino acids.

In some embodiments, a retrovirus present in a library of retroviruses comprises the same membrane-bound domain as some or all of the other retroviruses in the library. In some embodiments, each retrovirus present in a library of retroviruses comprises a different membrane-bound domain relative to some or all of the other retroviruses in the library.

In some embodiments, an extracellular targeting domain is any protein or peptide that has an amino acid sequence and is a binding partner for a target molecule or ligand (*e.g.*, a cognate protein) on a cell surface. When present in the extracellular environment beyond the interior of the retrovirus, an extracellular targeting domain is capable of binding to a target cell. In some embodiments, an extracellular targeting domain binds or targets to a cognate protein or ligand (*e.g.*, a protein receptor present on a target cell) that is present on the cell surface of a cell or a subset of a population of cells. In some embodiments, an extracellular targeting domain binds to a cognate protein or ligand that is present on the cell surface of a single T cell or a subset of a population of T cells. In some embodiments, a binding interaction between an extracellular targeting domain of a retrovirus and a cognate protein or ligand of a cell enables the retrovirus to enter the cell (*e.g.*, an antigen-specific cell, *e.g.*, a T cell).

In some embodiments, an extracellular targeting domain comprises 10-50, 10-100, 25-100, 50-200, 50-150, 100-500, 100-250, 250-500, or any reasonable number of total amino acids. In some embodiments, an extracellular targeting domain comprises at least 5, at least 10, at least 15, at least 20, or at least 50 amino acids.

In some embodiments, an extracellular targeting domain is a protein, an antibody or peptide. In some embodiments, an antibody is a full-length antibody, an antibody fragment, a nanobody, or a single chain antibody (scFv). In some embodiments, an extracellular targeting domain is an antibody that binds to a cognate protein of a target cell. In some embodiments, an extracellular targeting domain is an antibody that binds to a B-cell or T cell antigen. In some embodiments, an extracellular targeting domain is an anti-CD19 antibody

(*e.g.*, an antibody that binds to CD19). In some embodiments, an extracellular targeting domain is an antibody that binds to any cell surface molecule. In some embodiments, an extracellular targeting domain is an antibody that binds to a lineage marker (*e.g.*, CD3, CD20, integrins, or other receptors), phenotypic markers (PD-1, CD25, CD45, or others). In some embodiments, an extracellular targeting domain is a protein or peptide that binds to a receptor (*e.g.*, a receptor that is present on the surface of a target cell). In some embodiments, an extracellular targeting domain is a protein or peptide that binds to a cytokine receptor (*e.g.*, interleukin-13 (IL-13) receptor). In some embodiments, an extracellular targeting domain is a cytokine (*e.g.*, IL-2, IL-6, IL-12, IL-13). In some embodiments, an extracellular targeting domain is a chemokine ligand (*e.g.* CXCL9, CXCL10, CXCL 11, etc.). In some embodiments, an extracellular targeting domain is a cellular receptor, including cytokine receptors (*e.g.* IL-13R α 1, IL-13R α 2, IL-2 receptors, common gamma chain), GPCRs (including chemokine receptors such as CSCR3, CXCR4, etc.), and integrins. In some embodiments, an extracellular targeting domain is a peptide that is displayed by a MHC protein. In some embodiments, non-viral membrane-bound protein comprises a membrane-bound domain comprising a MHC protein or fragment and an extracellular targeting domain comprising a peptide that is displayed by a MHC protein. In some embodiments, an extracellular domain binds to a T cell receptor and/or a B cell receptor. T-cell receptors are expressed in nature on the surface of T-cells usually as alpha/beta and gamma/delta heterodimeric integral membrane proteins, each subunit comprising a short intracellular segment, a single transmembrane alpha-helix and two globular extracellular Ig-superfamily domains. B-cell receptors are transmembrane receptor proteins located on the outer surface of B cells.

In some embodiments, an extracellular targeting domain binds to a target cell or cell surface molecule with a binding affinity of 10^{-9} to 10^{-8} M, 10^{-8} to 10^{-7} M, 10^{-7} to 10^{-6} M, 10^{-6} to 10^{-5} M, 10^{-5} to 10^{-4} M, 10^{-4} to 10^{-3} M, or 10^{-3} to 10^{-2} M. In some embodiments, an extracellular targeting domain binds to a cognate protein or ligand of a target cell with a binding affinity of 10^{-9} to 10^{-8} M, 10^{-8} to 10^{-7} M, 10^{-7} to 10^{-6} M, 10^{-6} to 10^{-5} M, 10^{-5} to 10^{-4} M, 10^{-4} to 10^{-3} M, or 10^{-3} to 10^{-2} M. In some embodiments, the binding affinity between an extracellular targeting domain and a cognate protein or ligand is in the picomolar to nanomolar range (*e.g.*, between about 10^{-12} and about 10^{-9} M). In some embodiments, the

binding affinity between an extracellular targeting domain and a cognate protein or ligand is in the nanomolar to micromolar range (*e.g.*, between about 10^{-9} and about 10^{-6} M). In some embodiments, the binding affinity between an extracellular targeting domain and a cognate protein or ligand is in the micromolar to millimolar range (*e.g.*, between about 10^{-6} and about 10^{-3} M). In some embodiments, the binding affinity between an extracellular targeting domain and a cognate protein or ligand is in the picomolar to micromolar range (*e.g.*, between about 10^{-12} and about 10^{-6} M). In some embodiments, the binding affinity between an extracellular targeting domain and a cognate protein or ligand is in the nanomolar to millimolar range (*e.g.*, between about 10^{-9} and about 10^{-3} M).

As used herein, the term antibody generally refers to a protein that includes at least one immunoglobulin variable domain or immunoglobulin variable domain sequence. For example, an antibody can include a heavy (H) chain variable region (abbreviated herein as V_H), and/or a light (L) chain variable region (abbreviated herein as V_L). In another example, an antibody includes two heavy (H) chain variable regions and/or two light (L) chain variable regions. An antibody can have the structural features of IgA, IgG, IgE, IgD, IgM (as well as subtypes thereof). The V_H and V_L regions can be further subdivided into regions of hypervariability, termed “complementarity determining regions” (“CDR”), interspersed with regions that are more conserved, termed “framework regions” (“FR”). Each V_H and/or V_L is typically composed of three CDRs and four FRs, arranged from amino-terminus to carboxy-terminus in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. The V_H or V_L chain of the antibody can further include a heavy or light chain constant region, to thereby form a heavy or light immunoglobulin chain, respectively. In some embodiments, the antibody is a tetramer of two heavy immunoglobulin chains and two light immunoglobulin chains, wherein the heavy and light immunoglobulin chains are inter-connected by, *e.g.*, disulfide bonds. In IgGs, the heavy chain constant region includes three immunoglobulin domains, CH1, CH2 and CH3.

In some embodiments, a retrovirus present in a library of retroviruses comprises the same extracellular targeting domain as some or all of the other retroviruses in the library. In some embodiments, each retrovirus present in a library of retroviruses comprises a different extracellular targeting domain relative to some or all of the other retroviruses in the library.

In some embodiments, a non-viral membrane-bound protein further comprises a signal sequence (also referred to as a signal peptide or localization sequence). In some embodiments, the signal sequence is at the N- or C-terminal ends of the non-viral membrane-bound protein. A signal sequence functions to translocate the non-viral membrane-bound protein to the membrane (or envelope) of the retrovirus. In some embodiments, a signal sequence is 5-10, 5-15, 10-20, 15-20, 15-30, 20-30, or 25-30 amino acids. In some embodiments, the signal sequence is an Ig Kappa leader sequence (*e.g.*, a murine Ig Kappa leader sequence comprising: METDTLLLWVLLWVPGSTG (SEQ ID NO: 1)) or a B2M signal peptide sequence (*e.g.*, a B2M signal peptide sequence comprising: MSRSVALAVLALLSLSGLEA (SEQ ID NO: 2)). In some embodiments, a retrovirus present in a library of retroviruses comprises the same signal sequence as some or all of the other retroviruses in the library. In some embodiments, each retrovirus present in a library of retroviruses comprises a different signal sequence relative to some or all of the other retroviruses in the library.

In some embodiments, a nucleic acid encoding a non-viral membrane-bound protein further comprises an internal ribosome entry site (IRES). An IRES is an RNA sequence that allows for initiation of translation during protein synthesis. In some embodiments, the IRES is located at or near the C-terminal end. In some embodiments, the IRES is located C-terminal relative to the membrane-bound domain and the extracellular targeting domain. In some embodiments, the IRES is a viral IRES. In some embodiments, the IRES is an IRES that is native to the retrovirus. In some embodiments, the IRES is a sequence derived from encephalomyocarditis virus (EMCV). In some embodiments, a retrovirus present in a library of retroviruses comprises the same IRES as some or all of the other retroviruses in the library. In some embodiments, each retrovirus present in a library of retroviruses comprises a different IRES relative to some or all of the other retroviruses in the library.

In some embodiments, a non-viral membrane-bound protein further comprises a linker positioned between the membrane-bound domain and the extracellular targeting domain. A linker is an amino acid linker and may be a rigid linker, a flexible linker, or an oligomerized linker. A rigid linker is an amino acid sequence that lacks flexibility (*e.g.*, may comprise at least one proline). In some embodiments, a rigid linker comprises a platelet-derived growth factor receptor (PDGFR) stalk or a CD8 α stalk. In some embodiments, a

PDGFR stalk comprises an amino acid sequence comprising AVGQDTQEVIVVPHSLPFK (SEQ ID NO: 3). In some embodiments, a PDGFR stalk comprises an amino acid sequence comprising ASAKPTTTPAPRPPTPAPTIASQPLSLRPEAARPAAGGAVHTRGLDFAK (SEQ ID NO: 4). A flexible linker is an amino acid sequence that has many degrees of freedom (*e.g.*, may comprise a plurality of amino acids with small side chains, *e.g.*, glycine or alanine). In some embodiments, a flexible linker comprises an amino acid sequence comprising GAPGAS (SEQ ID NO: 5). In some embodiments, a flexible linker comprises an amino acid sequence consisting of GAPGSGGGGSGGGGSAS (SEQ ID NO: 6). In some embodiments, a flexible linker comprises an amino acid sequence comprising GGGGS (SEQ ID NO: 7). In some embodiments, a flexible linker comprises an amino acid sequence comprising (GAPGAS)_N (SEQ ID NO: 29) or (G₄S)_N (SEQ ID NO: 30), wherein N is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more. An oligomerized linker is an amino acid that can oligomerize to another related amino acid. In some embodiments, an oligomerized linker is an amino acid sequence that can form a dimer, trimer, or tetramer. In some embodiments, an oligomerized linker comprises an IgG4 hinge domain (*e.g.*, ASESKEYGPPCPPCPAVGQDTQEVIVVPHSLPFK (SEQ ID NO: 8)). In some embodiments, an oligomerized linker comprises an amino acid sequence that can form a tetrameric coiled coil (*e.g.*, ASGGGGSGELAAIKQELAAIKKELAAIKWELAAIKQGAG (SEQ ID NO: 9)). In some embodiments, an oligomerized linker comprises an amino acid sequence that can form a dimeric coiled coil (*e.g.*, ASESKEYGPPCPPCP (SEQ ID NO: 10)).

Reporter

In some embodiments, the retroviruses described herein may comprise a reporter (*e.g.*, a reporter protein). In some embodiments, the retroviruses described herein comprise a nucleic acid encoding a reporter (*e.g.*, a reporter protein). As used herein, a reporter is generally a protein or gene that can be detected when expressed in a retrovirus and/or target cell. In some embodiments, the presence or absence of a reporter in a target cell or a subset of a target cells in a population of cells allows for the ability to sort cells (*e.g.*, using flow cytometry and/or fluorescence-activated cell sorting).

In some embodiments, a reporter is a fluorescent protein. A fluorescent protein may be a green fluorescent protein (GFP), yellow fluorescent protein (YFP), red fluorescent

protein (RFP). A fluorescent protein may be as described in U.S. Patent No. 7,060,869, entitled “Fluorescent protein sensors for detection of analytes”.

In some embodiments, a reporter is an antibiotic resistance marker. In some embodiments, an antibiotic resistance marker is a protein or gene that confers a competitive advantage to a target cell that contains the marker. In some embodiments, the antibiotic resistance marker comprises a hygromycin resistance protein or gene, a kanamycin resistance protein or gene, ampicillin resistant protein or gene, streptomycin resistant protein or gene, or a neomycin resistance protein or gene.

Cells

A cell as described herein may be any bacterial, mammalian, or yeast cell. In some embodiments, a cell is a human, mouse, rat, or a non-human primate cell. In some embodiments, a cell is a somatic cell or a reproductive cell. In some embodiments, a cell is an epithelial cell, a neural cell, a hormone-secreting cell, an immune cell, a secretory cell, a blood cell, an interstitial cell, or a germ cell. In some embodiments, a cell is an antigen-specific cell (*e.g.*, a cell that binds to a specific antigen). In some embodiments, an antigen-specific cell is an immune cell. In some embodiments, an antigen-specific cell is a B cell or a T cell. In some embodiments, a cell is a target cell (*e.g.*, that comprises a cognate protein or ligand capable of being targeted by a retrovirus described herein)

A population of cells as described herein may be any bacterial, mammalian, or yeast cell population. In some embodiments, a population of cells is a population of human, mouse, rat, or non-human primate cells. In some embodiments, a population of cells is a somatic cell population or a reproductive cell population. In some embodiments, a population of cells comprises epithelial cells, neural cells, hormone-secreting cells, immune cells, secretory cells, blood cells, interstitial cells, and/or germ cells. In some embodiments, a population of cells comprises antigen-specific cells (*e.g.*, cells that binds to a specific antigen). In some embodiments, a population of antigen-specific cells comprises immune cells. In some embodiments, a population of antigen-specific cells comprises B cells and/or T cells. In some embodiments, a population of cells comprises a homogenous population of cells. In some embodiments, a population of cells comprises a heterogeneous population of cells.

In some embodiments, a population of cells is a population of cells isolated from a subject. A subject may be a human subject (*e.g.*, a human subject suffering from a disease), a mouse subject, a rat subject, or a non-human primate subject. In some embodiments, a population of cells is isolated from the blood or a tumor of a subject.

In some embodiments, a population of cells has been previously frozen and thawed (*e.g.*, 1, 2, 3, 4, 5, or more freeze/thaw cycles). In some embodiments, a population of cells are maintained in liquid culture media. In some embodiments, a population of cells have been passaged 1, 2, 3, 4, 5, or more times, using any known method. In some embodiments, a population of cells are maintained in liquid culture media prior to being combined with a retrovirus or plurality of retroviruses. In some embodiments, a population of cells are maintained in liquid culture media after to being combined with a retrovirus or plurality of retroviruses. In some embodiments, a population of cells are maintained in liquid culture media prior to while being combined with a retrovirus or plurality of retroviruses.

In some embodiments, a population of cells comprises any of the retroviruses described herein. In some embodiments, a subset of a population of cells contain any of the retroviruses described herein. In some embodiments, a subset of a population of cells contains the retrovirus inside each cell of the subset (*e.g.*, inside the nucleus of each cell of the subset). In some embodiments, a population of cells or a subset thereof expresses a reporter (*e.g.*, a fluorescent protein or an antibiotic resistance marker). In some embodiments, a population of cells or a subset thereof (*e.g.*, containing a retrovirus) are isolated and/or sorted based on the presence or absence of a reporter. In some embodiments, a subset of a population of cells that contain retrovirus described herein are isolated and/or sorted based on the presence or absence of a reporter away from the cells of the population that do not contain the retrovirus. In some embodiments, at least 50%, 60%, 70%, 80%, 90%, or 95% of a population of cells prior to cell sorting contain a retrovirus. In some embodiments, at least 70%, 80%, 90%, 95%, or 100% of a population of cells contain a retrovirus following isolation and/or sorting based on the presence or absence of a reporter.

Methods of screening

Described herein are methods of screening a population of cells comprising: (i) providing a retrovirus comprising a viral envelope protein comprising at least one mutation

that diminishes its native function, a non-viral membrane-bound protein comprising a membrane-bound domain and an extracellular targeting domain, and a nucleic acid encoding a reporter; (ii) combining the retrovirus with a population of cells; and (iii) sorting the population of cells based on the presence or absence of the reporter. In some embodiments, the retrovirus of (i) comprises a nucleic acid comprising a structure: S-ETD-MBD-IRES-R, wherein S encodes a signal sequence, ETD encodes an extracellular targeting domain; MBD encodes a membrane-bound domain, IRES encodes an internal ribosome entry site, and R encodes a reporter; and a mutated viral envelope protein comprising at least one mutation that diminishes its native function.

As used herein, the term “combining” (which, in some embodiments, is synonymous with the terms “providing” and “contacting”) generally refers to the act of bringing a retrovirus into close, physical contact with a population of cells, such that the extracellular targeting domain of the retrovirus is capable of binding to the cognate ligand present on a subset of cells of the population. In some embodiments, combining of a retrovirus and a population of cells occurs when a solution comprising the retrovirus and a solution comprising the population of cells are mixed. In some embodiments, combining of a retrovirus and a population of cells occurs when a lyophilized retrovirus and a solution comprising the population of cells are mixed. In some embodiments, combining of a retrovirus and a population of cells occurs when a lyophilized retrovirus and a lyophilized population of cells are mixed and reconstituted with a solution. In some embodiments, the cells of the population are maintained in cell culture media, in a monolayer of cells, and/or are attached to a tissue culture plate or petri dish.

Generally, a retrovirus and a population of cells are combined (*e.g.*, physically combined or contacted) for a defined period of time. In some embodiments, a period of time is measured in seconds, minutes, hours or days. In some embodiments, period of time is 0-30 seconds, 15-45 seconds, 30-60 seconds, 45-90 seconds, 60-90 seconds, or 60-120 seconds. In some embodiments, a retrovirus and a population of cells are combined and in contact for 0-30 seconds, 15-45 seconds, 30-60 seconds, 45-90 seconds, 60-90 seconds, or 60-120 seconds. In some embodiments, period of time is 1-2 minutes, 1-5 minutes, 1-10 minutes, 2-10 minutes, 5-10 minutes, 5-20 minutes, 10-20 minutes, 25-30 minutes, 25-60 minutes, 30-45 minutes, 30-40 minutes, 40-60 minutes, 50-70 minutes, or 60-120 minutes. In some

embodiments, a retrovirus and a population of cells are combined and in contact for 1-2 minutes, 1-5 minutes, 1-10 minutes, 2-10 minutes, 5-10 minutes, 5-20 minutes, 10-20 minutes, 25-30 minutes, 25-60 minutes, 30-45 minutes, 30-40 minutes, 40-60 minutes, 50-70 minutes, or 60-120 minutes. In some embodiments, a period of time is 1-2 hours, 1-5 hours, 1-3 hours, 2-5 hours, 3-6 hours, 3-12 hours, 6-12 hours, 12-18 hours, 12-24 hours, 15-30 hours, 18-24 hours, 24-48 hours, 24-36 hours, or 36-50 hours. In some embodiments, a retrovirus and a population of cells are combined and in contact for 1-2 hours, 1-5 hours, 1-3 hours, 2-5 hours, 3-6 hours, 3-12 hours, 6-12 hours, 12-18 hours, 12-24 hours, 15-30 hours, 18-24 hours, 24-48 hours, 24-36 hours, or 36-50 hours. In some embodiments, a period of time is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 5-15 days. In some embodiments, a retrovirus and a population of cells are combined and in contact for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 5-15 days.

In some embodiments, a population of cells are sorted based on the presence or absence of the reporter. In some embodiments, a subset of the population of cells containing the reporter (*e.g.*, express the reporter) are sorted from the remaining subset of the population of cells that do not contain the reporter. In some embodiments, sorting of the population of cells is performed using flow cytometry (*e.g.*, fluorescence-activated cell sorting), next-generation genome sequencing (*e.g.*, single-cell next-generation sequencing), or antibiotic selection.

In some embodiments, the conditions of step (ii) that allow for the retrovirus to have cell-to-cell interactions with a subset of the population of cells comprise combining the retrovirus and the population of cells in the presence of defined solutions, compositions and at specific temperatures. In some embodiments, the retrovirus and the population of cells are combined in the presence of a cell culture media (*e.g.*, RPMI or DMEM cell culture media). In some embodiments, the retrovirus and the population of cells are combined in the presence of a buffered saline solution. In some embodiments, a buffered saline solution is a phosphate-buffered saline or HEPES-buffered saline. In some embodiments, a buffered saline solution comprises bovine serum albumin and/or EDTA. In some embodiments, the retrovirus and the population of cells are combined in the presence of an enhancer of retroviral transduction (*e.g.*, heparin sulfate, polybrene, protamine sulfate, or dextran). In some embodiments, the retrovirus and the population of cells are combined in (ii) at a

temperature ranging from 4°C to 42°C, 4°C to 8°C, 4°C to 10°C, 8°C to 15°C, 10°C to 20°C, 18°C to 23°C, 20°C to 30°C, 25°C to 35°C, 30°C to 40°C, or 37°C to 42°C.

In some embodiments, the methods of screening described herein further comprise washing the population of cells between steps (ii) and (iii) with a wash solution. In some embodiments, a wash solution is any liquid solution that allows for maintenance of healthy cells (*e.g.*, solution comprising neutral pH, low-to-moderate levels of ionic strength). In some embodiments, washing the population of cells removes excess and/or remaining retrovirus from the population of cells. In some embodiments, the population of cells are washed using a cell culture media (*e.g.*, RPMI or DMEM cell culture media). In some embodiments, the population of cells are washed using a buffered saline solution. In some embodiments, a buffered saline solution is a phosphate-buffered saline or HEPES-buffered saline. In some embodiments, a buffered saline solution comprises bovine serum albumin and/or EDTA. In some embodiments, the population of cells are washed at a temperature ranging from 4°C to 42°C, 4°C to 8°C, 4°C to 10°C, 8°C to 15°C, 10°C to 20°C, 18°C to 23°C, 20°C to 30°C, 25°C to 35°C, 30°C to 40°C, or 37°C to 42°C.

In some embodiments, the population of cells are maintained in liquid culture prior to being combined with the retrovirus. In some embodiments, the population of cells are maintained in liquid culture after being combined with the retrovirus. In some embodiments, the population of cells are maintained in liquid culture during the combining step with the retrovirus. In some embodiments, the population of cells are attached to a cell culture plate or petri dish. In some embodiments, the population of cells are maintained in a monolayer, an embryoid body, or any cell aggregate.

In some embodiments, methods of screening comprise the use of a plurality of retroviruses. In certain embodiments, a plurality of retroviruses comprises at least 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , or 10^{12} unique retroviruses. In some embodiments, there may be at least 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , or 10^{12} copies of each unique retrovirus present in a plurality of retroviruses.

In some embodiments, methods of screening comprise screening a population of cells with at least two different, unique retroviruses. In some embodiments, a different, unique retrovirus comprises a different extracellular targeting domain and/or a different reporter. In some embodiments, methods of screening comprise a first retrovirus and a second retrovirus,

wherein the first and second retrovirus comprise different extracellular targeting domains and/or different reporters. In some embodiments, methods of screening comprise screening a population of cells with 2, 3, 4, 5, 6, 7, 8, 9, 10, 50, 100 or more different retroviruses. In some embodiments, methods of screening comprise screening a population of cells with a library of retroviruses. In some embodiments, a library of retroviruses comprises at least 10^2 , at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 , at least 10^8 , at least 10^9 , or at least 10^{10} unique retroviruses.

Library of retroviruses

Described herein are libraries of retroviruses, wherein a library comprises a plurality of unique retroviruses, wherein each unique retrovirus comprises a viral envelope protein comprising at least one mutation that diminishes its native function, a non-viral membrane-bound protein comprising a membrane-bound domain and an extracellular targeting domain, and a nucleic acid encoding a reporter, and wherein each unique retrovirus comprises a different and unique extracellular targeting domain. Also described herein are libraries of cells comprising retroviruses, wherein a library comprises a plurality of unique cells, wherein each unique cell comprises a unique retrovirus.

In some embodiments, libraries include pMHC-encoded (peptide/MHC-encoded) retroviral (*e.g.*, lentiviral) libraries for use in screening populations of T cells. In such libraries, the pMHC displayed on the virus surface will enable T cell infection in a TCR-specific manner. Infected T cells can be collected and sequenced, allowing for the identification of pMHC ligands that can infect a subset of a T cell population of interest and the ability to simultaneously track TCR sequences and reactive pMHC ligands. In some embodiments, pMHC retroviral libraries minimally comprise randomized transfer vectors containing randomized pMHC targeting elements. In some embodiments, randomly derived libraries are generated using degenerate oligonucleotide primers. In some embodiments, targeted libraries that are specific for a unique set of antigens (*e.g.*, all possible viral or bacterial antigens for a particular target of interest – human immunodeficiency virus, tuberculosis TB, *etc.*; or all possible neoantigens for a particular subject) are generated.

In some embodiments, a library is capable of being screened against a population of antigen-specific cells (*e.g.*, B cells or T cells). In some embodiments, a library comprises at

least 10^2 , at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 , at least 10^8 , at least 10^9 , or at least 10^{10} unique retroviruses. In some embodiments, a library comprising unique retroviruses comprises extracellular targeting domains that are at least 5, at least 10, at least 15, at least 20, or at least 50 amino acids in length. In some embodiments, each different and unique extracellular targeting domain is generated through site-directed mutagenesis.

Retroviral or cell libraries can vary in size from hundreds to hundreds of thousands, millions, or more unique retroviruses or unique cells. In some embodiments, the libraries of the disclosure comprise at least 500,000 unique retroviruses or unique cells. The libraries of the invention include retroviral libraries and cellular libraries. A library is a synthetic (i.e., isolated, synthetically produced, free from components that are naturally found together in a cell, purified before being put into the library) collection of members having a common element and at least one distinct element. The library comprises a thousand or more (e.g., at least: 1,000; 2,000; 3,000; 4,000; 5,000; 10,000; 50,000; 100,000; 500,000; 600,000; 700,000; 800,000; 900,000; 1,000,000; 2,000,000; 3,000,000; 4,000,000; or more) members. The upper limit of the library size is defined by the combinatorics of domains or modules providing distinctness or diversity among the members. For instance, an upper limit may be 4,000,000 members. Thus, in some embodiments, the library is highly diverse, and includes at least 500,000 distinct members. The highly diverse library may have a diversity of 10^6 or greater. In some embodiments, a library of retroviruses is generated using site-directed mutagenesis of a nucleic acid described herein. In some embodiments, the site-directed mutagenesis involves the use of primers and a low-fidelity RNA polymerase to allow for randomized mutagenesis of a common nucleic acid as described herein.

Methods of delivering nucleic acid to a cell

Described herein are methods of delivering a nucleic acid to a cell, comprising (i) providing a retrovirus, as described herein, comprising the nucleic acid, a viral envelope protein comprising at least one mutation that diminishes its native function, and a non-viral membrane-bound protein comprising an extracellular targeting domain that is capable of binding to a cognate ligand of the cell; and (ii) contacting the retrovirus with the cell such that the retrovirus enters or infects the cell. In some embodiments, the nucleic acid encodes an mRNA molecule, optionally wherein the mRNA is a gene of interest. In some

embodiments, the nucleic acid encodes a double-stranded RNA, an antisense RNA, a microRNA, or any other RNA molecule. In some embodiments, the gene of interest encodes a protein. In some embodiments, the gene of interest encodes a therapeutic protein (*e.g.*, a protein to compensate for a diseased condition in a subject).

In some embodiments, the nucleic acid is delivered to the cell when the retrovirus enters or infects the cell during step (ii). In some embodiments, the methods of delivering a nucleic acid described herein do not require a transfection agent (*e.g.*, a lipophilic transfection agent such as Lipofectin).

Methods of detection

Described herein are methods of detecting an interaction between a retrovirus and a cell, comprising: (i) contacting a sample comprising the retrovirus and an cell with an antibody, wherein the retrovirus comprises a viral envelope protein comprising at least one mutation that diminishes its native function, a non-viral membrane-bound protein comprising an extracellular targeting domain, and wherein the antibody binds to the extracellular targeting domain of the retrovirus; (ii) optionally removing unbound antibody from the sample; and (iii) imaging the sample to detect whether the antibody-retrovirus complex is bound to the cell.

In some embodiments, the antibody further comprises at least one fluorescent label. In some embodiments, a fluorescent label is a xanthene derivative (*e.g.*, fluorescein, rhodamine, Oregon green, eosin and Texas red), cyanine derivative (*e.g.*, cyanine, indocarbocyanine, oxacarbocyanine, thiocarbocyanine and merocyanine), naphthalene derivative (*e.g.*, dansyl and prodan derivatives), coumarin derivative, oxadiazole derivative (*e.g.*, pyridyloxazole, nitrobenzoxadiazole and benzoxadiazole), pyrene derivative (*e.g.*, cascade blue), oxazine derivative (*e.g.*, Nile red, Nile blue, cresyl violet and oxazine 170), acridine derivative (*e.g.*, proflavin, acridine orange and acridine yellow), arylmethine derivative (*e.g.*, auramine, crystal violet and malachite green), or tetrapyrrole derivative (*e.g.*, porphyrin, phthalocyanine and bilirubin). The fluorescent label may be non-covalently associated with the antibody or covalently linked to the antibody.

In some embodiments, the sample is imaged in step (iii) using confocal or fluorescence microscopy. In some embodiments, methods of detection can be accomplished

using standard microscopy setups (e.g., confocal or fluorescence microscopes). In some embodiments, a sample is detected in an ultra-multiplexed format while imaging using standard confocal or epi-fluorescence microscope.

Nucleic acids

As used herein, the term “nucleic acids” generally refers to multiple linked nucleotides (i.e., molecules comprising a sugar (e.g., ribose or deoxyribose) linked to an exchangeable organic base, which is either a pyrimidine (e.g., cytosine (C), thymidine (T) or uracil (U)) or a purine (e.g., adenine (A) or guanine (G)). Nucleic acids include DNA such as D-form DNA and L-form DNA and RNA, as well as various modifications thereof. Modifications include base modifications, sugar modifications, and backbone modifications.

It is to be understood that the nucleic acids used in retroviruses and methods of the invention may be homogeneous or heterogeneous in nature. As an example, they may be completely DNA in nature or they may be comprised of DNA and non-DNA (e.g., LNA) monomers or sequences. Thus, any combination of nucleic acid elements may be used. The modification may render the nucleic acid more stable and/or less susceptible to degradation under certain conditions. For example, in some instances, the nucleic acids are nuclease-resistant. Methods for synthesizing nucleic acids, including automated nucleic acid synthesis, are also known in the art.

The nucleic acids may comprise modifications in their bases. Modified bases include modified cytosines (such as 5-substituted cytosines (e.g., 5-methyl-cytosine, 5-fluoro-cytosine, 5-chloro-cytosine, 5-bromo-cytosine, 5-iodo-cytosine, 5-hydroxy-cytosine, 5-hydroxymethyl-cytosine, 5-difluoromethyl-cytosine, and unsubstituted or substituted 5-alkynyl-cytosine), 6-substituted cytosines, N4-substituted cytosines (e.g., N4-ethyl-cytosine), 5-aza-cytosine, 2-mercapto-cytosine, isocytosine, pseudo-isocytosine, cytosine analogs with condensed ring systems (e.g., N,N'-propylene cytosine or phenoxazine), and uracil and its derivatives (e.g., 5-fluoro-uracil, 5-bromo-uracil, 5-bromovinyl-uracil, 4-thio-uracil, 5-hydroxy-uracil, 5-propynyl-uracil), modified guanines such as 7 deazaguanine, 7 deaza 7 substituted guanine (such as 7 deaza 7 (C2 C6)alkynylguanine), 7 deaza 8 substituted guanine, hypoxanthine, N2-substituted guanines (e.g. N2-methyl-guanine), 5-amino-3-methyl-3H,6H-thiazolo[4,5-d]pyrimidine-2,7-dione, 2,6 diaminopurine, 2 aminopurine,

purine, indole, adenine, substituted adenines (e.g. N6-methyl-adenine, 8-oxo-adenine) 8 substituted guanine (e.g. 8 hydroxyguanine and 8 bromoguanine), and 6 thioguanine. The nucleic acids may comprise universal bases (e.g. 3-nitropyrrole, P-base, 4-methyl-indole, 5-nitro-indole, and K-base) and/or aromatic ring systems (e.g. fluorobenzene, difluorobenzene, benzimidazole or dichloro-benzimidazole, 1-methyl-1H-[1,2,4]triazole-3-carboxylic acid amide). A particular base pair that may be incorporated into the oligonucleotides of the invention is a dZ and dP non-standard nucleobase pair reported by Yang et al. NAR, 2006, 34(21):6095-6101. dZ, the pyrimidine analog, is 6-amino-5-nitro-3-(1'- β -D-2'-deoxyribofuranosyl)-2(1H)-pyridone, and its Watson-Crick complement dP, the purine analog, is 2-amino-8-(1'- β -D-1'-deoxyribofuranosyl)-imidazo[1,2-a]-1,3,5-triazin-4(8H)-one.

Amino acid substitutions

In some embodiments, the amino acid residue variations are conservative amino acid residue substitutions. As used herein, a "conservative amino acid substitution" refers to an amino acid substitution that does not alter the relative charge or size characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Conservative substitutions of amino acids include substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D.

The "percent identity" of two amino acid sequences is determined using the algorithm of Karlin and Altschul Proc. Natl. Acad. Sci. USA 87:2264-68, 1990, modified as in Karlin and Altschul Proc. Natl. Acad. Sci. USA 90:5873-77, 1993. Such an algorithm is incorporated into the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. J. Mol. Biol. 215:403-10, 1990. BLAST protein searches can be performed with the XBLAST program, score=50, wordlength=3 to obtain amino acid sequences homologous to the protein molecules of interest. Where gaps exist between two sequences, Gapped BLAST can be

utilized as described in Altschul et al., Nucleic Acids Res. 25(17):3389-3402, 1997. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used.

EXAMPLES

Example 1. Generation of retrovirus that targets antigen-specific T cells

Targeted lentiviruses were generated by polyethylenimine (PEI) transfection of HEK293T cells with the following plasmids: an envelope plasmid encoding a mutated VSV-G envelope protein comprising K47Q and R354A mutations, at least one helper plasmid (pRRE, pRev, or psPAX2.1), and a transfer plasmid (FIG. 1). The transfer plasmids used encoded a nucleic acid comprising a structure: S-ETD-MBD-IRES-R, wherein S encoded a B2M signal sequence (as provided by SEQ ID NO: 2), ETD encoded a variable extracellular targeting domain (e.g., the cancer-testis antigen NYESO-1); MBD encoded a membrane bound domain (e.g., MHC HLA-A2 domain), IRES encoded an internal ribosome entry site derived from encephalomyocarditis virus (EMCV), and R encoded a green fluorescent protein GFP reporter (FIG. 2).

The resulting viruses were harvested and purified by standard centrifugation techniques, prior to mixing (using pipette mixing) with T cell lines (e.g. Jurkat T cells) expressing T cell receptors (TCRs) specific to known pMHCs in the presence of hexadimethrine bromide .

Mixing HLA-A2-NYESO-1 pMHC-displaying viruses with Jurkat T cells expressing an IG4 T cell receptor (TCR) variant (SEQ ID NO: 27) that recognizes the displayed NYESO-1 with a binding affinity of ~26 pM, efficient infection was observed, with 41.9% of T cells expressing the GFP reporter after mixing (FIG. 3A). This result indicated that 41.9% of T cells in this cell population were infected by the virus. Conversely, mixing HLA-A2-NYESO-1 pMHC-displaying viruses were mixed with cells that did not express an IG4 TCR, minimal infection was observed, with only 1.1% of T cells expressing the GFP reporter after mixing (FIG. 3B).

Mixing HLA-A2-NYESO-1 pMHC-displaying viruses with T cells expressing different versions of the IG4 TCR (wherein the different versions of the IG4 TCR exhibit reduced binding to the NYESO-1 antigen, compared to IG4 TCR variant comprising SEQ ID

NO: 27) demonstrated that the viruses were able to infect T cells even when the binding affinity between the NYESO-1 antigen and IG4 TCR was down to 32 μ M (FIG. 4). A 32 μ M binding affinity between NYESO-1 antigen and an IG4 TCR variant (SEQ ID NO: 24) provided 3.6% transduction; a 84 nM binding affinity between NYESO-1 antigen and a IG4 TCR variant (SEQ ID NO: 25) provided 17.6% transduction; and a 5 nM binding affinity between NYESO-1 antigen and a IG4 TCR variant (SEQ ID NO: 26) provided 18.1% transduction; when measuring for GFP expression.

IG4-expressing T cells exhibited T cell activation when transduced/infected with HLA-A2-NYESO-1-displaying viruses, as evidenced by upregulation of CD69 (FIG. 5A). Conversely, HLA-A2-SL9-displaying viruses did not transduce/infect nor activate IG4-expressing T cells (FIG. 5B).

Example 2. Generation of retroviruses that target IL-13 receptor

A lentivirus comprising an interleukin-13 (IL-13) extracellular targeting domain was generated, as described in Example 1. The IL-13 extracellular targeting domain consisted of full-length IL-13 protein connected to an IgG4 hinge linker protein linker and a PDGFR transmembrane domain comprising a truncated cytoplasmic tail (PDGFR transmembrane domain: VVVISAILALVVLTIISLIILIMLWQKKPR (SEQ ID NO: 28)). When the resultant viruses were mixed with Jurkat cells expressing the IL-13R α 1 receptor, efficient transduction was observed, with 85.2% of the cells expressing the GFP reporter after mixing (FIG. 7A). When the same viruses were mixed with cells not expressing any IL-13 receptors, minimal infection was observed (FIG. 7B).

Additional IL-13 expressing lentiviruses were generated with alternative protein linkers to connect the IL-13 extracellular targeting domain to the PDGFR transmembrane domain. Equivalent volumes of all lentiviruses tested were mixed with Jurkat cells bearing the IL13R α 1 receptor and transduction rates were assessed by GFP reporter (FIG. 12).

Short and long protein linkers were tested for their abilities to allow for cell-to-cell interactions between the extracellular targeting domain (IL-13) and the cognate ligand (IL13R α 1 receptor). Lentiviruses comprising short linkers were marginally able to infect Jurkat cells, with PDGFR linker (AVGQDTQEVIVPHSLPFK (SEQ ID NO: 3)) enabling infection of 6.1% of cells; GAPGAS linker (GAPGAS (SEQ ID NO: 5)) enabling infection of

1.0% of cells; and a CAR ECD linker (ASESKYGPPCPPCP (SEQ ID NO: 10)) enabling infection of 1.6% of cells. Lentiviruses comprising long linkers were highly successful at infecting Jurkat cells, with CD8a stalk linker (ASAKPTTTPAPRPPTPAPTIASQPLSLRPEAARPAAGGAVHTRGLDFAK (SEQ ID NO: 4)) enabling infection of 88.5% of cells; IgG4 hinge linker (ASESKYGPPCPPCPAVGQDTQEIVIVPHSLPFK (SEQ ID NO: 8)) enabling infection of 85.2% of cells; and an oligomerized linker comprising an amino acid sequence that can form a tetrameric coiled coil (*e.g.*, ASGGGSGELAAIKQELAAIKKELAAIKWELAAIKQGAG (SEQ ID NO: 9)) enabling infection of 38.5% of cells.

Example 3. Generation of retroviruses that target CD19

A lentivirus pseudotyped with a Nipah virus F protein and a Nipah virus G protein fused to an anti-CD19 scFv was able to transduce CD19⁺ Raji B cells (cells that express CD19), with 14.5% of cells expressing the GFP reporter after mixing (FIG. 9).

A lentivirus pseudotyped with VSV-mut (mutated VSV-G) and an anti-CD19 scFv fused to an IgG hinge domain, a truncated PDGFR stalk, and PDGFR transmembrane domain was able to transduce CD19⁺ Raji B cells. Application of 1 μ L of VSV-mut lentivirus (200x concentrated) to the CD19⁺ Raji B cells provided 10.0% infected cells (based on expression of GFP reporter); 10 μ L VSV-mut provided 10.7% infected cells. (FIG. 10). The same VSV-mut lentivirus construct did not transduce/infect CD19⁻ Jurkat cells (cells that do not express CD19), demonstrating that the lentivirus is specific for cells expressing the cognate ligand of its extracellular targeting domain (anti-CD19 scFv).

A lentivirus pseudotyped with VSV-wt (wild-type VSV-G) and an anti-CD19 scFv fused to an IgG hinge domain, a truncated PDGFR stalk, and PDGFR transmembrane domain was able to transduce CD19⁺ Raji B cells and CD19⁻ Jurkat cells. This comparative result demonstrates the lentivirus comprising a mutated VSV-G (wherein the mutations diminish the native function of VSV-G protein) is selective for transducing/infecting cells that express a cognate ligand targeted by the extracellular targeting domain of the lentivirus.

Example 4. Ability of targeted lentiviruses to induce T cell signaling

Viruses were assessed to determine whether infection via TCR targeting resulted in

evidence of TCR signaling. While non-targeted VSV-wt viruses efficiently infected T cells, they did not activate T cells, as assessed by CD69 expression levels, compared to controls (FIG. 13, left panel). However, when HLA-A2-NYESO-1 displaying viruses were used to target Jurkat T cells expressing the high-affinity wild-type 1G4 TCR, upregulation of CD69 was observed in transduced cells, indicating TCR signaling occurred during infection (FIG. 13, right panel). These results were replicated when mixing HLA-A2-NYESO-1 displaying viruses with T cells expressing the lower-affinity mutant 1G4 TCR constructs described in Example 1 (FIG. 14)

Example 5. Lentiviral transduction through the TCR at physiological affinities

A population of Jurkat T cells was purified on the basis of expression of IG4 TCR before being contacted with HLA-A2-NYESO-1 displaying viruses. The viruses efficiently transduced the purified populations of Jurkat T cells expressing each of the affinity variants of the 1G4 TCR, based on expression of the GFP reporter (FIG. 15). A 32 μ M binding affinity between NYESO-1 antigen and an IG4 TCR variant provided 51.3% transduction; a 84 nM binding affinity between NYESO-1 antigen and an IG4 TCR variant provided 84.8% transduction; a 5 nM binding affinity between NYESO-1 antigen and an IG4 TCR variant provided 82.3% transduction; and a 26 nM binding affinity between NYESO-1 antigen and an IG4 TCR variant provided 73.2% transduction.

The differences among different IG4 TCR affinity variants were further investigated by performing a dose response to incubate varying amounts (0.5-10 μ L of virus) of the HLA-A2-NYESO-1 displaying viruses with cells. NYESO-1 displaying lentiviruses transduce IG4-expressing T cells, even at high viral load (FIG. 16A). Conversely, SL-1 displaying lentiviruses (negative controls) do not transduce IG4-expressing T cells (SL-1 and IG4 TCR are not cognate binding partners).

The protein architecture of the membrane-bound NYESO-1 targeting protein was altered to test the ability of oligomerized NYESO-1 displaying constructs to transduce T cells. NYESO-1 displaying expressing lentiviruses were generated with alternative protein linkers to connect the NYESO-1 extracellular targeting domain to transmembrane domain were generated (a protein linker comprising the IgG4 hinge (dimer); and an oligomerized linker comprising an amino acid sequence that can form a tetrameric coiled coil (tetramer)).

As demonstrated in FIG. 17, lentiviruses expressing these alternative architectures were able to transduce T cells expressing the 26 pM and 32 μ M IG4 TCR variants.

Example 6. Virus binding assay

Viruses were pseudotyped with either (1) wild-type VSV-G (VSV-G); (2) measles virus F protein and receptor-blinded H protein fused to an anti-CD19 scFv and a FLAG tag (MV- α CD19); (3) Nipah virus F protein and receptor-blinded G protein fused to an anti-CD19 scFv and a FLAG tag (NiV- α CD19); or (4) Nipah virus F and receptor-blinded G protein with a FLAG tag but no scFv (NiV-dead). Each of these viruses was mixed with CD19⁺ Raji B cells and subsequently exposed to a fluorescently-labeled antibody against the FLAG tag. As demonstrated in FIGs. 18A-18B, the viruses that expressed an anti-CD19 scFv (MV- α CD19 and NiV- α CD19) bound to the Raji B cells. The viruses that did not express an anti-CD19 scFv (VSV-G and NiV-dead) did not bind to the Raji B cells.

Example 7. Screening a library of retroviruses against population of T cells

A library of retroviruses (10^7 - 10^{10} unique retroviruses) comprising a mutated VSV-G envelope protein, a non-viral membrane-bound protein comprising a MHC membrane-bound domain, a variable IL-13 extracellular targeting domain, and a GFP reporter.

First, a library of nucleic acids encoding unique non-viral membrane-bound proteins comprising a MHC membrane-bound domain and a variable IL-13 extracellular targeting domain, is prepared using PCR primers that include degenerate codons at positions known to interact with IL-13 receptors and induce random mutagenesis in the IL-13 extracellular targeting domain. The resulting nucleic acid constructs are assembled into viral transfer plasmids with fluorescent reporters by standard molecular cloning techniques. These plasmids ("transfer plasmids"), along with helper plasmids and an envelope plasmid encoding mutated VSV-G envelope protein (or an equivalent pseudotype) are transfected into retroviruses to generate the library of retroviruses. The viruses are purified and mixed with populations of cells (*e.g.* T cell lines, HEK293 cell lines) that express an IL-13 receptor of interest. The viruses are incubated with the populations of cells at an appropriate temperature (*e.g.*, 37°C) for an appropriate period of time (*e.g.*, 1-48 hours). Mixing is performed in standard cell culture media with hexadimethrine bromide. After 24-48 hours, the cells are

sorted based on the expression of the GFP reporter. High-expressing cells are retained and the sorting process is repeated as needed. After these rounds of selection by sorting, the retained cells are lysed and the cellular RNA and DNA are isolated for analysis using next-generation sequencing methods to determine which IL-13 variants were able to mediate viral entry.

Example 8. Generation of additional retroviruses that targets antigen-specific T cells

Additional targeted lentiviruses were generated using the protocols described in Example 1.

As shown in FIG. 19, mixing (i) lentivirus comprising a mutated VSV-G envelope protein and pMHC displaying a SL9 peptide and (ii) 868 TCR-expressing T cells demonstrated that the viruses were able to transduce and infect T cells even at low amounts of added virus (~70% transduction at 1 μ L of virus). Conversely, mixing of these viruses with off-target Jurkat cells led to low levels of transduction (less than 5% up to 10 μ L of virus added). A similar experiment

As also shown in FIG. 19, mixing (i) lentivirus comprising a mutated VSV-G envelope protein and pMHC displaying a cytomegalovirus (CMV) NLV peptide and (ii) C7 TCR-expressing T cells demonstrated that the viruses were able to transduce and infect T cells (~25% transduction at 10 μ L of virus). Conversely, mixing of these viruses with off-target Jurkat cells led to low levels of transduction (less than 5% up to 10 μ L of virus added).

Example 9. Generation of retroviruses comprising targeted pMHCs stabilized by disulfides

Additional targeted lentiviruses were generated using the protocols described in Example 1. These additional lentiviruses further comprised a disulfide situated within the pMHC in order to stabilize the structure of the pMHC.

As shown in FIG. 21, mixing (i) lentivirus comprising a mutated VSV-G envelope protein and a disulfide-stabilized pMHC displaying a targeting peptide and (ii) target cells expressing a cognate receptor demonstrated that the targeted viruses were able to transduce and infect target cells. Conversely, mixing of these targeted viruses with off-target cells caused no, or very limited (0.1%) transduction.

Specifically, a lentivirus comprising a disulfide-stabilized CMV NLV pMHC transduced on-target C7 TCR-expressing T cells at a 21.8% transduction rate; a lentivirus comprising a disulfide-stabilized EBV pMHC transduced on-target AS01 cells at a 4.8% transduction rate; a lentivirus comprising a disulfide-stabilized SL9 pMHC transduced on-target 868 TCR-expressing T cells at a 22.1% transduction rate; and a lentivirus comprising a disulfide-stabilized NYESO-1 pMHC transduced on-target IG4 TCR-expressing T cells at a 19.0% transduction rate.

Example 10. Targeted viruses transduce primary cells

A primary T cell line specific for GL9 (presented by HLA-A2) was specifically and efficiently transduced by viruses having a mutated VSV-G envelope protein and displaying GL9 (FIGs. 22A-22B). About 50% of primary T cells were transduced by these GL9-targeting viruses. Conversely, these GL9-targeting viruses only transduced 0.3% of IG4 TCR-expressing T cells. The IG4 TCR-expressing T cells were specifically and efficiently transduced by viruses having a mutated VSV-G envelope protein and displaying NYESO-1 (55.9% transduction). Furthermore, the targeted virus infects the primary cells more efficiently than a wild-type VSV-G virus. These results indicate that the targeted viruses can even infect unstimulated primary cells.

Similarly, a primary T cell line specific for NYESO (presented by HLA-A2) was specifically and efficiently transduced by viruses having a mutated VSV-G envelope protein and displaying NYESO-1 (FIGs. 23-24). Conversely, GL9-targeting viruses were unable to efficiently transduce the primary T cell line specific for NYESO. Collectively, these data demonstrate that primary NYESO-1-reactive cells, including expanded primary cells, can be specifically infected by NYESO-targeted viruses.

Example 11. Generation of retroviruses comprising CD80 domain

A virus pseudotyped with a mutated VSV-G envelope protein and comprising a CD80 extracellular domain was generated using the protocol described in Example 1. This virus was able to specifically and efficiently infect Jurkat T cells (25.5% transduction), relative to B cells (0.0% transduction) (FIG. 25).

Similarly, a virus pseudotyped with a mutated VSV-G envelope protein and

comprising a NYESO-1 pMHC and a CD80 extracellular domain was generated using the protocol described in Example 1. This virus was able to specifically and efficiently infect Jurkat T cells (13.5% transduction), relative to B cells (0.2% transduction) (FIG. 25). The presence of the CD80 domain enabled the transduction of this virus into the Jurkat T cells, as demonstrated by the inability of a virus pseudotyped with a mutated VSV-G envelope protein and only comprising the NYESO-1 pMHC to infect these Jurkat T cells (0.9% transduction).

Collectively, these data show that the presence of CD80 on the virus surface mediates specific infection of T cells and demonstrates that CD80 could be used to generally target viruses to T cells.

Example 12. Murine anti-CD3 antibody mediates infection of TCR-transduced 58 α β ⁻ cells

Display of antibodies that are specific to the murine TCR constant region (H57 antibody) or the murine CD3 (2C11 antibody) on the surface of a virus enabled the viruses to infect mouse T cell lines. A virus comprising a mutated VSV-G viral envelope protein and the anti-TCR antibody provided 11.0% transduction of 58 α β ⁻ mouse T hybridoma cells; and a virus comprising a mutated VSV-G viral envelope protein and the anti-CD3 antibody provided 12.8% transduction of 58 α β ⁻ mouse T hybridoma cells.

Example 13. Generation of viruses comprising dead cocal virus G protein

A lentivirus was pseudotyped with a cocal virus G protein (Cocal-dead; amino acid comprising SEQ ID NO: 53) comprising mutations to reduce its infectivity. These mutations, at K64Q and R371A of the cocal virus G protein, were analogous to those mutations used in the VSV-dead variant described in Example 1. The cocal-dead virus further comprised a displayed anti-CD19 scFv antibody.

This cocal-dead virus was able to transduce CD19⁺ Raji B cells (8.6% transduction), as shown in FIG. 27, in a similar manner as a lentivirus comprising VSV-dead protein and displayed anti-CD19 scFv antibody.

These data demonstrates that the viral targeting strategy is highly robust and that any viral envelope protein (*e.g.*, mutated VSV-G, mutated Nipah envelope, mutated Measles

envelope, mutated cocal viral envelope) that can be mutated to reduce its infectivity can be effectively used.

OTHER EMBODIMENTS

All of the features disclosed in this specification may be combined in any combination. Each feature disclosed in this specification may be replaced by an alternative feature serving the same, equivalent, or similar purpose. Thus, unless expressly stated otherwise, each feature disclosed is only an example of a generic series of equivalent or similar features.

From the above description, one skilled in the art can easily ascertain the essential characteristics of the present invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions. Thus, other embodiments are also within the claims.

EQUIVALENTS

While several inventive embodiments have been described and illustrated herein, those of ordinary skill in the art will readily envision a variety of other means and/or structures for performing the function and/or obtaining the results and/or one or more of the advantages described herein, and each of such variations and/or modifications is deemed to be within the scope of the inventive embodiments described herein. More generally, those skilled in the art will readily appreciate that all parameters, dimensions, materials, and configurations described herein are meant to be exemplary and that the actual parameters, dimensions, materials, and/or configurations will depend upon the specific application or applications for which the inventive teachings is/are used. Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific inventive embodiments described herein. It is, therefore, to be understood that the foregoing embodiments are presented by way of example only and that, within the scope of the appended claims and equivalents thereto, inventive embodiments may be practiced otherwise than as specifically described and claimed. Inventive embodiments of the present disclosure are directed to each individual feature, system, article, material, kit, and/or method described herein. In addition, any combination of two or more such features,

systems, articles, materials, kits, and/or methods, if such features, systems, articles, materials, kits, and/or methods are not mutually inconsistent, is included within the inventive scope of the present disclosure.

All definitions, as defined and used herein, should be understood to control over dictionary definitions, definitions in documents incorporated by reference, and/or ordinary meanings of the defined terms.

All references, patents and patent applications disclosed herein are incorporated by reference with respect to the subject matter for which each is cited, which in some cases may encompass the entirety of the document.

The indefinite articles “a” and “an,” as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to mean “at least one.”

The phrase “and/or,” as used herein in the specification and in the claims, should be understood to mean “either or both” of the elements so conjoined, i.e., elements that are conjunctively present in some cases and disjunctively present in other cases. Multiple elements listed with “and/or” should be construed in the same fashion, i.e., “one or more” of the elements so conjoined. Other elements may optionally be present other than the elements specifically identified by the “and/or” clause, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, a reference to “A and/or B”, when used in conjunction with open-ended language such as “comprising” can refer, in one embodiment, to A only (optionally including elements other than B); in another embodiment, to B only (optionally including elements other than A); in yet another embodiment, to both A and B (optionally including other elements); etc.

As used herein in the specification and in the claims, “or” should be understood to have the same meaning as “and/or” as defined above. For example, when separating items in a list, “or” or “and/or” shall be interpreted as being inclusive, i.e., the inclusion of at least one, but also including more than one, of a number or list of elements, and, optionally, additional unlisted items. Only terms clearly indicated to the contrary, such as “only one of” or “exactly one of,” or, when used in the claims, “consisting of,” will refer to the inclusion of exactly one element of a number or list of elements. In general, the term “or” as used herein shall only be interpreted as indicating exclusive alternatives (i.e. “one or the other but not both”) when preceded by terms of exclusivity, such as “either,” “one of,” “only one of,” or

“exactly one of.” “Consisting essentially of,” when used in the claims, shall have its ordinary meaning as used in the field of patent law.

As used herein in the specification and in the claims, the phrase “at least one,” in reference to a list of one or more elements, should be understood to mean at least one element selected from any one or more of the elements in the list of elements, but not necessarily including at least one of each and every element specifically listed within the list of elements and not excluding any combinations of elements in the list of elements. This definition also allows that elements may optionally be present other than the elements specifically identified within the list of elements to which the phrase “at least one” refers, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, “at least one of A and B” (or, equivalently, “at least one of A or B,” or, equivalently “at least one of A and/or B”) can refer, in one embodiment, to at least one, optionally including more than one, A, with no B present (and optionally including elements other than B); in another embodiment, to at least one, optionally including more than one, B, with no A present (and optionally including elements other than A); in yet another embodiment, to at least one, optionally including more than one, A, and at least one, optionally including more than one, B (and optionally including other elements); etc.

It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one step or act, the order of the steps or acts of the method is not necessarily limited to the order in which the steps or acts of the method are recited.

SEQUENCES

>Kappa leader sequence, amino acid (SEQ ID NO: 1): METDTLLLWVLLLWVPGSTG

>B2M signal peptide sequence, amino acid (SEQ ID NO: 2): MSRSVALAVLALLSLSGLEA

>PDGFR short stalk, amino acid (SEQ ID NO: 3): AVGQDTQEIVVPHSLPFK

>PDGFR long stalk, amino acid (SEQ ID NO: 4):

ASAKPTTTPAPRPPTPAPTIASQPLSLRPEAARPAAGGAVHTRGLDFAK

>Short flexible linker, amino acid (SEQ ID NO: 5): GAPGAS

>Long flexible linker, amino acid (SEQ ID NO: 6): GAPGSGGGGSGGGGSAS

>Short flexible linker, amino acid (SEQ ID NO: 7): GGGGS

>IgG4 hinge domain, amino acid (SEQ ID NO: 8):

ASESKYGPPCPPCPAVGQDTQEVIVVPHSLPFK

>Tetrameric coiled coil, amino acid (SEQ ID NO: 9):

ASGGGSGELAAIKQELAAIKKELAAIKWELAAIKQGAG

>Dimeric coiled coil, amino acid (SEQ ID NO: 10): ASESKEYGPCCPPCP

>Wild-type VSV-G envelope protein (with leader sequence), DNA sequence (SEQ ID NO: 11):

atgaagtgcctttgtacttagcctttttattcattgggggtgaattgcaagttcaccatagttttccacacaacaaaaaggaaactggaaaaatgttcctt
ctaattaccattattgcccgtcaagctcagatttaaattggcataatgacttaataggcacagccatacaagtcaaaatgcccagagtcacaaggcta
ttcaagcagacgggttgatgtgcatgcttccaaatgggtcactactgtgatttccgctggtatggaccgaagtataaacacagtcacccgatcctt
cactccatctgtagaacaatgcaaggaaagcattgaacaacgaaacaaggaaactggctgaatccagggttccctcctcaaaagtgtggatatgca
actgtgacggatgccgaagcagtgattgtccaggtgactcctccatgtgctggtgatgaatacacaggagaatgggttgattcacagttcatcaa
cggaaaatgcagcaattacatgccccactgtccataactctacaacctggcattctgactataaggtcaaaaggctatgtattctaactcatttcc
atggacatcaccttctcagaggacggagagctatcatcctgggaaggagggcacaggggttcagaagtaactacttgcctatgaaactggag
gcaaggcctgcaaaatgcaactgcaagcattggggagtcagactccatcaggtgtctggttcgagatggctgataaggtatcttctgctgcagc
cagattccctgaatgccagaagggtcaggtatctctgctccatctcagacctcagtggtgatgaagctaaatcaggacgttgagaggtatcttgatta
ttccctctgccaaagcctggagcaaaatcagagcgggtcttccaatctctcagtggtatctcagctatcttgcctctaaaaaccagggaaccggtc
ctgctttaccataatcaatgtaccctaaaatactttgagaccagatacatcagagtcgatattgctgctcctcaatcctcaagaatggcggatgat
cagtggaactaccacagaagggaactgtgggatgactgggcaccatataagagcgtggaattggaccaatggagttctgaggaccagttca
ggatataagtttctttatcatgattggacatggtatgttgactccgatcttcatcttagctcaaaaggctcaggtgttcgaacatcctcattcaagac
gctgcttcgcaactcctgatgatgagagtttattttgtgatactgggctatccaaaatccaatcagctgtgtagaagggttggttcagtagttggaa
aagctctattgctctttttcttatcatagggtaatcattgactattctgtgtctccgagttggtatccatcttgcattaaatgaagcacaccaagaa
aagacagatttatacagacatagatgaaccgacttggaagtaa

>Wild-type VSV-G envelope protein (with leader sequence), amino acid sequence (SEQ ID NO: 12):
MKCLLYLAFLFIGVNCKFTIVFPHNQKGNWKNVPSNYHYCPSSDLNWHNDLIGTAIQVKMP
KSHKAIQADGWMCHASKWVTTCDFRWYGPKYITQSIRSFTPSVEQCKESIEQTKQGTWLN
PFPPQSCGYATVTDAAEVIVQVTPHHVLVDEYTGWVDSQFINGKCSNYICPTVHNSTTWHS
DYKVKGLCDSNLISMDITFFSEDGELSSLGKEGTGFRSNYFAYETGGKACKMQYCKHWGVR
LPSGVWFEMADKDLFAAARFPECPEGSSISAPSQTSVDVSLIQDVERILDYSLCQETWSKIRAG
LPISPVDSLAPKNPGTGAFTIINGTLKYFETRYIRVDIAAPILSRMVGMISSGTTTERELWDD
WAPYEDVEIGPNGVLRRTSSGYKFLYMIGHGMLDSDLHLSKAQVFEHPHIQDAASQLPDDE
SLFFGDTGLSKNPIELVEGWFWSSWKSSIASFFFIIGLIIGLFLVLRVGIHLCKIKLKHKKRQIYTDI
EMNRLGK

>Wild-type VSV-G envelope protein, amino acid sequence (SEQ ID NO: 13):

KFTIVFPHNQKGNWKNVPSNYHYCPSSDLNWHNDLIGTAIQVKMPKSHKAIQADGWMCH
ASKWVTTCDFRWYGPKYITQSIRSFTPSVEQCKESIEQTKQGTWLNPGFPPQSCGYATVTDAA
EVIVQVTPHHVLVDEYTGWVDSQFINGKCSNYICPTVHNSTTWHSKYKVKGLCDSNLISM
DITFFSEDGELSSLGKEGTGFRSNYFAYETGGKACKMQYCKHWGVRVLPSPGVWFEMADKDLF
AAARFPECPEGSSISAPSQTSVDVSLIQDVERILDYSLCQETWSKIRAGLPISPVDSLAPKNP
GTGAFTIINGTLKYFETRYIRVDIAAPILSRMVGMISSGTTTERELWDDWAPYEDVEIGPNGVL
RTSSGYKFLYMIGHGMLDSDLHLSKAQVFEHPHIQDAASQLPDDESLFFGDTGLSKNPIEL
VEGWFWSSWKSSIASFFFIIGLIIGLFLVLRVGIHLCKIKLKHKKRQIYTDIEMNRLGK

>VSV-G envelope protein (with leader sequence), DNA sequence (SEQ ID NO: 14):

atgaagtgcctttgtacttagcctttttattcattgggggtgaattgcaagttcaccatagttttccacacaacaaaaaggaaactggaaaaatgttcctt
ctaattaccattattgcccgtcaagctcagatttaaattggcataatgacttaataggcacagccttacaagtcaaaatgcccagagtcacaaggctat
tcaagcagacgggttgatgtgcatgcttccaaatgggtcactactgtgatttccgctggtatggaccgaagtataaacacagtcacccgatcctt

actccatctgtagaacaatgaaggaaagcattgaacaaacgaacaaggaacttggtgaatccaggcttccctcctcaaaagtgtgatatgcaa
 ctgtgacggatgccgaagcagtgattgtccagggtgactcctcaccatgtgctggttgatgaatacacaggagaatgggttgattcacagttcatcaac
 ggaaaaatgcagcaattacatatgccccactgtccataactctacaacctggcattctgactataaggtaaaagggtatgtgattctaacctcatttcca
 tggacatcaccttctctcagaggacggagagctatccctgggaaaggagggcacagggttcagaagtaactacttctgtatgaaactggagg
 caaggcctgcaaaaatgcaatactgcaagcattggggagtcagactcccatcagggtgtctggttcagatggctgataaggatctcttctgcagcc
 agattccctgaatgccagaagggtcaagtatctctgctccatctcagacctcagtggtgatgaagtctaattcaggacgttgagaggatcttgattatt
 ccctctgcaagaacctggagcaaaatcagagcggtcttccaatctctcagtggtatctcagctatcttctcctaaaaaccagggaaccgggtcc
 tgccttcaccataatcaatggtaccctaaaatacttggagaccagatacatcagagtcgatattgctgctccaatctctcaagaatggtcggaatgac
 agtggaaactaccacagaagccgaactgtgggatgactgggcaccatataagacgtggaaattggaccaatggagtctgaggaccagttcag
 gatataagtttctttatcatgattggacatggtatgttgactccgacttcatcttagctcaaaaggctcaggtgttcgaacatctcacattcaagacg
 ctgcttcgaactctctgatgatgagagttatttttgggtgatactgggctatccaaaaatccaatcagctttagaagggttggtcagtagttgga
 agctctattgctctttttctttatcatagggttaatcattggactattcttggttctccagttggtatccatcttgcaatfaaataagcacaccaagaaa
 agacagatttatacagacatagatgaaccgacttggaagtaa

>VSV-G envelope protein (with leader sequence), amino acid sequence (SEQ ID NO: 15):

MKCLLYLAFLFIGVNCKFTIVFPHNQKGNWKNVPSNYHYCPSSSDLNWHNDLIGTALQVKM
 PQSHKAIQADGWMCHASKWVTTCDFRWYGPKYITQSIRSFTPSVEQCKESIEQTKQGTWLN
 PGPFPQSCGYATVTDAAEVIVQVTPHHVLVDEYTG EWVDSQFINGKCSNYICPTVHNSTTWHS
 DYKVKGLCDSNLISMDITFFSEDGELSSLGKEGTGFRSNYFAYETGGKACKMQYCKHWGVR
 LPSGVWFEMADKDLFAAARFPECPEGSSISAPSQTSVDVSLIQDVERILDYSLCQETWSKIRAG
 LPISPVDLSYLAPKNPGTGPAFTIINGTLKYFETRYIRVDIAAPILSRMVGMMISGTTTEAELWDD
 WAPYEDVEIGPNGVLRTSSGYKFPLYMIGHGMLDSDLHLSSKAQVFEHPHIQDAASQLPDDE
 SLFFGDTGLSKNPIELVEGWFWSSWKSSIASFFFIIGLIIGLFLVLRVGIHLCKIKLKHTKKRQIYTDI
 EMNRLGK

>I41L/K47Q/R354A VSV-G envelope protein, amino acid sequence (SEQ ID NO: 16):

KFTIVFPHNQKGNWKNVPSNYHYCPSSSDLNWHNDLIGTALQVKMPQSHKAIQADGWMCH
 ASKWVTTCDFRWYGPKYITQSIRSFTPSVEQCKESIEQTKQGTWLNPGFPFPQSCGYATVTDAAE
 AVIVQVTPHHVLVDEYTG EWVDSQFINGKCSNYICPTVHNSTTWHSKYKVKGLCDSNLISM
 DITFFSEDGELSSLGKEGTGFRSNYFAYETGGKACKMQYCKHWGVRVRLPSGVWFEMADKDLF
 AAARFPECPEGSSISAPSQTSVDVSLIQDVERILDYSLCQETWSKIRAGLPISPVDLSYLAPKNP
 GTGPAFTIINGTLKYFETRYIRVDIAAPILSRMVGMMISGTTTEAELWDDWAPYEDVEIGPNGVL
 RTSSGYKFPLYMIGHGMLDSDLHLSSKAQVFEHPHIQDAASQLPDDESLLFFGDTGLSKNPIEL
 VEGWFSSWKSSIASFFFIIGLIIGLFLVLRVGIHLCKIKLKHTKKRQIYTDIEMNRLGK

>K47Q/R354A VSV-G envelope protein, amino acid sequence (SEQ ID NO: 17):

KFTIVFPHNQKGNWKNVPSNYHYCPSSSDLNWHNDLIGTALQVKMPQSHKAIQADGWMCH
 ASKWVTTCDFRWYGPKYITQSIRSFTPSVEQCKESIEQTKQGTWLNPGFPFPQSCGYATVTDAAE
 AVIVQVTPHHVLVDEYTG EWVDSQFINGKCSNYICPTVHNSTTWHSKYKVKGLCDSNLISM
 DITFFSEDGELSSLGKEGTGFRSNYFAYETGGKACKMQYCKHWGVRVRLPSGVWFEMADKDLF
 AAARFPECPEGSSISAPSQTSVDVSLIQDVERILDYSLCQETWSKIRAGLPISPVDLSYLAPKNP
 GTGPAFTIINGTLKYFETRYIRVDIAAPILSRMVGMMISGTTTEAELWDDWAPYEDVEIGPNGVL
 RTSSGYKFPLYMIGHGMLDSDLHLSSKAQVFEHPHIQDAASQLPDDESLLFFGDTGLSKNPIEL
 VEGWFSSWKSSIASFFFIIGLIIGLFLVLRVGIHLCKIKLKHTKKRQIYTDIEMNRLGK

>Exemplary wild-type measles envelope protein (with leader sequence), DNA sequence (SEQ ID NO: 18):

ATGGGCAGCCGGATCGTGATCAACCGGGAGCACCTGATGATCGACCGGCCCTACGTGCT
 GCTGGCCGTGCTGTTCTGATGTTCTGAGCCTGATCGGCTTGCTAGCCATTGCTGGAAT
 CCGGCTGCACAGAGCCGCATCTACACCGCCGAGATCCACAAGAGCCTGAGCACCAACC
 TGGACGTGACCAACAGCATCGAGCATCAGGTCAAGGACGTGCTGACCCCCCTGTTTAAG

ATCATCGGCGACGAAGTGGGCCTGCGGACCCCCCAGAGATTCACCGACCTGGTCAAGTT
CATCAGCGACAAGATCAAGTTCCTGAACCCCGACCGGGAGTACGACTTCCGGGACCTGA
CCTGGTGCATCAACCCCCCGAGCGGATCAAGCTGGACTACGACCAGTACTGCGCCGAT
GTGGCCCGCGAGGAAGTGAATGCATTGGTGAAGTCAACTCTACTGGAGACCAGAAC
AACCAATCAGTTCCTAGCTGTCTCAAAGGGAAACTGCTCAGGGCCCACTACAATCAGAG
GTCAATTCTCAAACATGTCGCTGTCCCTGTTAGACTTGTATTTAGGTCGAGGTTACAATGT
GTCATCTATAGTCACTATGACATCCCAGGGAATGTATGGGGGAACTTACCTAGTGGA
AGCCTAATCTGAGCAGCAAAAGGTCAGAGTTGTCACAACTGAGCATGTACCGAGTGTTT
GAAGTAGGTGTTATCAGAAATCCGGGTTTGGGGGCTCCGGTGTTCATATGACAACTAT
CTTGAGCAACCAGTCAGTAATGATCTCAGCAACTGTATGGTGGCTTTGGGGGAGCTCAA
CTCGCAGCCCTTTGTACGGGGAAGATTCTATCACAATCCCTATCAGGGATCAGGGAAA
GGTGTCAAGCTCGTCAAGCTAGGTGTCTGGAATCCCCAACCGACATGCAATCC
TGGGTCCCCTTATCAACGGATGATCCAGTGATAGACAGGCTTTACCTCTCATCTCAGAG
GGTGTATCGCTGACAACCAAGCAAAATGGGCTGTCCCGACAACACGAACAGATGACAA
GTTGCGAATGGAGACATGCTTCCAACAGGCGTGTAAGGGTAAAATCCAAGCACTCTGCG
AGAATCCCGAGTGGGCACCATGAAGGATAACAGGATTCCTTCATACGGGGTCTTGTCTG
TTGATCTGAGTCTGACAGTTGAGCTTAAATCAAAATTGCTTCGGGATTCGGGCCATGTG
TCACACACGGTTCAGGGATGGACCTATACAAATCCAACCACAACAATGTGTATTGGCTG
ACTATCCCGCCAATGAAGAACCTAGCCTTAGGTGTAATCAACACATTGGAGTGGATACC
GAGATTCAAGGTTAGTCCCtatCTCTTCAcAgtCCCAATTAAGGAAGCAGGCGGAGACTGCC
ATGCCCCAACATACCTACCTGCGGAGGTGGATGGTGTATGTCAAACCTCAGTTCCAATCTGG
TGATTCTACCTGGTCAAGATCTCCAATATGTTTTGGCAACCTACGATACTTCCcgGGTTGA
ACATGCTGTGGTTTATTACGTTTACAGCCCAAGCCGCTCATTTTCTTACTTTTATCCTTTTA
GGTTGCCTATAAAGGGGGTCCCCATCGAATTACAAAGTGAATGCTTCACATGGGACCAA
AAACTCTGGTGCCGTCACCTTCTGTGTGCTTGGGACTCAGAATCTGGTGGACATATCACT
CACTCTGGGATGGTGGGCATGGGAGTCAGCTGCACAGTCACCCGGGAAGATGGAACCAA
TGACTACAAAGACGATGACGACAAGTGA

>Exemplary wild-type measles envelope protein, amino acid sequence (SEQ ID NO: 19):

MGSRIVINREHLMIDRPYVLLAVLFVFMFLSLIGLLAIAGIRLHRAAIYTAIEIHKSLSTNLDVTNS
IEHQVKDVLTPLFKIIGDEVGLRTPQRFTDLVKFISDKIKFLNPDREYDFRDLTWCINPPERIKL
DYDQYCADVAEEELMNALVNSTLLETRTTNQLAVSKGNCSGPTTIRGQFSNMSLSLLDLYL
GRGYNVSSIVTMTSQQMYGGTYLVEKPNLSSKRSELSQLSMYRVFEVGVIRNPGLGAPVFH
MTNYLEQPVSNLNSCMVALGELKLAALCHGEDSITIPYQSGSGKVSFQLVKLGWKSPTD
MQSWVPLSTDDPVIDRLYLSSHRGVIADNQAKWAVPTTRTDDKLRMETCFQQACKGKIQUAL
CENPEWAPLKDNRIPIYGVLSVDLSLTVELKIKIASGFGPLITHGSGMDLYKSNHNNVYWLTI
PPMKNLALGVINTLEWIPRFKVSPLYFTVPIKEAGGDCHAPTYPAEVDGDVKLSSNLVILPG
QDLQYVLATYDTSRVEHAVVYVYSPSRFSFYFPFRLPIKGVPIELQVECFTWDQKLWCRH
FCVLADSESGGHITHSGMVGVMGVSCTVTREDGTNDYKDDDDK

>Exemplary mutant measles envelope protein, DNA sequence (SEQ ID NO: 20):

ATGGGCAGCCGGATCGTGATCAACCGGGAGCACCTGATGATCGACCGGCCCTACGTGCT
GCTGGCCGTGCTGTTCTGTGATGTTCTGAGCCTGATCGGCTTGCTAGCCATTGCTGGAAT
CCGGCTGCACAGAGCCGCCATCTACACCGCCGAGATCCACAAGAGCCTGAGCACCAACC
TGGACGTGACCAACAGCATCGAGCATCAGGTCAAGGACGTGCTGACCCCCCTGTTTAAG
ATCATCGGCGACGAAGTGGGCCTGCGGACCCCCCAGAGATTCACCGACCTGGTCAAGTT
CATCAGCGACAAGATCAAGTTCCTGAACCCCGACCGGGAGTACGACTTCCGGGACCTGA
CCTGGTGCATCAACCCCCCGAGCGGATCAAGCTGGACTACGACCAGTACTGCGCCGAT
GTGGCCCGCGAGGAAGTGAATGCATTGGTGAAGTCAACTCTACTGGAGACCAGAAC
AACCAATCAGTTCCTAGCTGTCTCAAAGGGAAACTGCTCAGGGCCCACTACAATCAGAG
GTCAATTCTCAAACATGTCGCTGTCCCTGTTAGACTTGTATTTAGGTCGAGGTTACAATGT

GTCATCTATAGTCACTATGACATCCCAGGGAATGTATGGGGGAACCTTACCTAGTGAAAA
AGCCTAATCTGAGCAGCAAAAGGTCAGAGTTGTCACAACCTGAGCATGTACCGAGTGT
GAAGTAGGTGTTATCAGAAATCCGGGTTTGGGGGCTCCGGTGTTCATATGACAACTAT
CTTGAGCAACCAGTCAGTAATGATCTCAGCAACTGTATGGTGGCTTTGGGGGAGCTCAAA
CTCGCAGCCCTTTGTACGGGGAAGATTCTATCACAATTCCCTATCAGGGATCAGGGAAA
GGTGTGAGCTTCCAGCTCGTCAAGCTAGGTGTCTGGAAATCCCCAACCGACATGCAATCC
TGGGTCCCCTTATCAACGGATGATCCAGTGATAGACAGGCTTTACCTCTCATCTCACAGA
GGTGTATCGCTGACAACCAAGCAAAATGGGCTGTCCCGACAACACGAACAGATGACAA
GTTGCGAATGGAGACATGCTTCCAACAGGCGTGTAAGGGTAAAATCCAAGCACTCTGCG
AGAATCCCGAGTGGGCACCATGAAGGATAACAGGATTCCTTCATACGGGGTCTTGTCTG
TTGATCTGAGTCTGACAGTTGAGCTTAAAATCAAAATTGCTTCGGGATTCGGGGCCATTGA
TCACACACGGTTCAGGGATGGACCTATACAAATCCAACCACAACAATGTGTATTGGCTG
ACTATCCCGCCAATGAAGAACCTAGCCTTAGGTGTAATCAACACATTGGAGTGGATAACC
GAGATTCAAGGTTAGTCCCGCGCTCTTCAATGTCCCAATTAAGGAAGCAGGCGGAGACT
GCCATGCCCCAACATACCTACCTGCGGAGGTGGATGGTGTATGTTCAAACTCAGTTCCAATC
TGGTGATTCTACCTGGTCAAGATCTCCAATATGTTTTGGCAACCTACGATACTTCCGCGGT
TGAACATGCTGTGGTTTATTACGTTTACAGCCCAAGCCGCTCATTTTCTTACTTTTATCCT
TTTAGGTTGCCTATAAAGGGGGTCCCCATCGAATTACAAGTGGAATGCTTCACATGGGAC
CAAAAACCTCTGGTGCCGTCACTTCTGTGTGCTTGCGGACTCAGAATCTGGTGGACATATC
ACTACTCTGGGATGGTGGGCATGGGAGTCAGCTGCACAGTCACCCGGGAAGATGGAAC
CAATGACTACAAAGACGATGACGACAAGTGA

>Exemplary mutant measles envelope protein, amino acid sequence (SEQ ID NO: 21):

MGSRIVINREHLMIDRPYVLLAVLFVFMFLSLIGLLAIAGIRLHRAAIYTAIEHKSLSTNLDVTNS
IEHQVKDVLTPLFKIIGDEVGLRTPQRFTDLVKFISDKIKFLNPDREYDFRDLTWCINPPERIKL
DYDQYCADVAEEELMNALVNSTLLETTRTNQFLAVSKGNCSGPTTIRGQFSNMSLSLLDLYL
GRGYNVSSIVTMTSQGMYGGTYLVEKPNLSSKRSELSQLSMYRVFEVGVIRNPGLGAPVFH
MTNYLEQPVSNLNSNCMV ALGELKLAALCHGEDSITIPYQSGSGKGVSFQLVKLG VWKSP TD
MQSWVPLSTDDPVIDRLYLSSH RGVIADNQA KW AVPTTRTDDKLRMETCFQ QACKGKI QAL
CENPEWAPLKDNRIPSYGVLSVDLSLTVELKIKIASGFGPLITHGSGMDLYKSNHNNVYWLTI
PPMKNLALGVINTLEWIPRFKVSPALFNVPIKEAGGDCHAPT YLPAEVDGDVKLSSNLVILPG
QDLQYVLATYDTS AVEHAVVYYVYSPSRFSFYFPFRLPIKGVPIELQVECF TWDQKLWCRH
FCVLADSESGGHITHSGMVG MVGSCTVTREDGTNDYKDDDDK

>Exemplary mutant Nipah envelope protein, DNA sequence (SEQ ID NO: 22):

ATGAAGAAGATCAACGAGGGCCTGCTGGACAGCAAGATCCTGAGCGCCTTCAACACCGT
GATTGCCCTGCTGGGCTCTATCGTGATCATCGTGATGAACATCATGATCATCCAGAACTA
CACCCGGTCCACCGACAACCAGGCCGTGATTAAGGATGCTCTGCAGGGAATCCAGCAGC
AGATCAAAGGCCTGGCCGACAAGATCGGCACAGAGATCGGCCCTAAGGTGTCCCTGATC
GACACCAGCAGCACCATCACAATCCCCGCCAATATCGGACTGCTGGGAAGCAAGATCAG
CCAGAGCACC GCCAGCATCAACGAGAACGTGAACGAGAAGTGCAAGTTCACCCCTGCCTC
CACTGAAGATCCACGAGTGCAACATCAGCTGCCCAATCCTCTGCCATT CAGAGAGTACA
AAAAACAGACAGAGGGCGTGTCCAATCTCGTGGGCCTGCCTAACAAACATCTGCCTGCAG
AAAACCAGCAACCAGATCCTGAAGCCTAAGCTGATCTCCTACACACTGCCCGTCGTGGG
CCAGAGCGGCACCTGTATTACAGATCCTCTGCTGGCCATGGACGAGGGCTACTTTGCCTA
CAGCCACCTGGAAAGAATCGGCAGCTGTAGCCGGGGAGTGTCCAAGCAGAGAATCATCG
GCGTGGGCGAAGTGCTGGATAGAGGCGACGAAGTGCCAGCCTGTT CATGACCAATGTG
TGGACCCCTCCTAATCCTAACACCGTGTACCCTGCAGCGCCGTGTACAACAACGAGTTC
TACTACGTGCTGTGCGCCGTGTCCACAGTGGGCGACCCTATCCTGAACAGCACCTATTGG
AGCGGCAGCCTGATGATGACCAGACTGGCCGTGAAGCCCAAGAGCAATGGCGGCGGATA
CAACCAGCATCAGCTGGCCCTGCGGTCCATCGAGAAGGGCAGATACGACAAAGTGATGC

CTTACGGCCCCAGCGGCATCAAGCAAGGCGATACCCTGTACTTTCCCGCCGTGGGATTTC
 TCGTGCGGACCGAGTTCAAGTACAACGACAGCAACTGCCCCATCACCAAGTGCCAGTAC
 AGCAAGCCCCGAGAACTGCAGACTGAGCATGGGCATCAGACCCAACAGCCACTACATCCT
 GAGAAGCGGCCTGCTGAAGTACAACCTGAGCGACGGCGAGAACCCCAAGGTGGTGTTC
 TCGAGATCAGCGACCAGCGGCTGTCTATCGGCAGCCCCTCCAAGATCTACGACTCTCTGG
 GCCAGCCAGTGTTCTACCAGGCCAGCTTTAGCTGGGACACCATGATCAAGTTTCGGCGACG
 TGCTGACCGTGAATCCCCTGGTGGTCAACTGGCGGAACAATACCGTGATCAGCCGGCCTG
 GCCAGTCTCAGTGCCCCAGATTCAATACCTGTCCTGCCATTTGCGCCGAAGGCGTGTACA
 ATGACGCTTCTCTGATCGATCGGATCAACTGGATCTCTGCCGGCGTGTTCCTGGACTCTA
 ATGCCACAGCCGCCAATCCTGTGTTACCGTGTTCAAGGACAATGAGATCCTGTATCGGG
 CCCAGCTGGCCTCCGAGGACACAAATGCCAGAAAACAATCACCAACTGCTTTCTGCTCA
 AGAACAAGATCTGGTGCATCAGCCTGGTGGAAATCTACGACACCGGCGACAACGTGATC
 AGGCCCAAGCTGTTTCGCCGTGAAGATCCCTGAGCAGTGTACAGGCGGCGGAGGATCTGG
 CGGAGGTGGAAGCGGAGGCGGTGGATCTGCTAGCGATTACAAGGATGACGACGATAAGT
 GA

>Exemplary mutant Nipah envelope protein, amino acid sequence (SEQ ID NO: 23):

MKKINEGLLDKILSAFNTVIALLSIVIMNIMIIQNYTRSTDNQAVIDDALQIQQIKGLA
 DKIGTEIGPKVSLIDTSSTITIPANIGLLGSKISQSTASINENVNEKCKFTLPLKIHECNISCPNPL
 PFREYRPQTEGVSNLVGLPNNICLQKTSNQLKPKLISYTLPVVGQSGTCITDPLLAMDEGYFA
 YSHLERIGSCSRGVSKQRIIGVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYF
 VLCAVSTVGPILNSTYWSGLMMTRLAVKPKSNGGGYNQHLALRSIEKGRYDKVMPYGP
 SGIKQGDITLYFPAVGFLVRTEFKYNDSPITKQYKSKPENCRLSMGIRPNSHYILRSGLLY
 NLSDGENPKVVFIEISDQRLSIGSPSKIYDSLQGPVIFYQASFSWDTMIKFGDVLTVNPLVVNW
 RNNTVISRPGQSQCPRFNTCPAICAEGVYNDAFLIDRINWISAGVFLDSNATAANPVFTVFKD
 NEILYRAQLASEDTNAQKTTITNCFLKNKIWCISLVEIYDTGDNVIRPKLFAVKIPEQCTGGGG
 SGGGSGGGGSASDYKDDDDK

>IG4 TCR (Variant that binds to NYESO-1 antigen with 32 μ M binding affinity), amino acid sequence (SEQ ID NO: 24):

MSIGLLCCAALSLLWAGPVNAGVTQTPKFQVLKTGQSMTLQCAQDMNHEYMSWYRQDPG
 MGLRLIHYSVGAGITDQGEVPNGYNVSRSTTDFPLRLLSAAPSQTSVYFCASSYVGNTGELF
 FGEGRSLTVLEDLKNVFPPEVAVFEPSEAEISHTQKATLVCLATGFYPDHVELSWVNGKEV
 HSGVSTDPQPLKEQPALNDSRYCLSSRLRVSAFTWQNPRNHFRCQVQFYGLSENDEWTQDR
 AKPVTQIVSAEAWGRADCGFTSESYQQGVLSATILYEILLGKATLYAVLVSAVLMMAMVKR
 KDSRGGSATNFSLLKQAGDVEENPGPMETLLGLLILWLQLQWVSSKQEVTPAALSVPEGE
 NLVLNCSFTDSAIYNLQWFRQDPGKGLTSLLLIQSSQREQTSGRNLNASLDKSSGSSTLYIAASQ
 PGDSATYLCAVRPTSGGSYIPTFGRGTSLVHPYIQNPDPVYQLRDSKSSDKSVCLFTDFDSQ
 TNVSQSKDSDVYITDKTVLDMRSMDFKSNSAVAWSNKSDFACANAFNNSIIPEDTFFPSPSS
 CDVKLVEKSFETDTNLFQNLVIGFRILLKLVAGFNLLMTLRLWSSAAA

>IG4 TCR (Variant that binds to NYESO-1 antigen with 84 nM binding affinity), amino acid sequence (SEQ ID NO: 25):

MSIGLLCCAALSLLWAGPVNAGVTQTPKFQVLKTGQSMTLQCAQDMNHEYMSWYRQDPG
 MGLRLIHYSVGAGITDQGEVPNGYNVSRSTTDFPLRLLSAAPSQTSVYFCASSYLGNTGELF
 FGEGRSLTVLEDLKNVFPPEVAVFEPSEAEISHTQKATLVCLATGFYPDHVELSWVNGKEV
 HSGVSTDPQPLKEQPALNDSRYCLSSRLRVSAFTWQNPRNHFRCQVQFYGLSENDEWTQDR
 AKPVTQIVSAEAWGRADCGFTSESYQQGVLSATILYEILLGKATLYAVLVSAVLMMAMVKR
 KDSRGGSATNFSLLKQAGDVEENPGPMETLLGLLILWLQLQWVSSKQEVTPAALSVPEGE
 NLVLNCSFTDSAIYNLQWFRQDPGKGLTSLLLIQSSQREQTSGRNLNASLDKSSGSSTLYIAASQ
 PGDSATYLCAVRPMIGGTIYPTFGRGTSLVHPYIQNPDPVYQLRDSKSSDKSVCLFTDFDSQ

TNVSQSKSDVYITDKTVLDMRSMDFKSNSAVAWSNKSDFACANAFNNSIIPEDTFFPSPSS
CDVKLVEKSFETDTNLFQNLVIGFRILLKLVAGFNLLMTLRLWSSAAA

>IG4 TCR (Variant that binds to NYESO-1 antigen with 5 nM binding affinity), amino acid sequence (SEQ ID NO: 26):

MSIGLLCCAALSLLWAGPVNAGVTQTPKFQVLKTGQSMTLQCAQDMNHEYMSWYRQDPG
MGLRLIHYSVGAGTTDRGEVPNGYNVSRSTIEDFPLRLLSAAPSQTSVYFCASSYVGDGTGELF
FGEGRSLTVLEDLKNVFPPEVAVFEPSEAEISHTQKATLVCLATGFYDPDHVELSWVNGKEV
HSGVSTDPQPLKEQPALNDSRYCLSSRLRVSAFTWQNPNNHFRQCQVQFYGLSENDEWTQDR
AKPVTQIVSAEAWGRADCGFTSESYQQGVLSATILYEILLGKATLYAVLVSAVLMLMAMVKR
KDSRGGSATNFSLLKQAGDVEENPGPMETLLGLLILWLQLQWVSSKQEVTPAALSVPEGE
NLVLNCSFTDSAIYNLQWFRQDPGKGLTSLLLIQSSQREQTSGRNLNASLDKSSGSSTLYIAASQ
PGDSATYLC AVRPLLDGTIYPTFGRGTSILVHPYIQNPDPVYQLRDSKSSDKSVCLFTDFDSQ
TNVSQSKSDVYITDKTVLDMRSMDFKSNSAVAWSNKSDFACANAFNNSIIPEDTFFPSPSS
CDVKLVEKSFETDTNLFQNLVIGFRILLKLVAGFNLLMTLRLWSSAAA

>IG4 TCR (Variant that binds to NYESO-1 antigen with 26 pM binding affinity), amino acid sequence (SEQ ID NO: 27):

MSIGLLCCAALSLLWAGPVNAGVTQTPKFQVLKTGQSMTLQCAQDMNHEYMSWYRQDPG
MGLRLIHYSVAIQTTDQGEVPNGYNVSRSTIEDFPLRLLSAAPSQTSVYFCASSYLNGTGELFF
GEGSRLTVLEDLKNVFPPEVAVFEPSEAEISHTQKATLVCLATGFYDPDHVELSWVNGKEVH
SGVSTDPQPLKEQPALNDSRYCLSSRLRVSAFTWQNPNNHFRQCQVQFYGLSENDEWTQDRA
KPVTQIVSAEAWGRADCGFTSESYQQGVLSATILYEILLGKATLYAVLVSAVLMLMAMVKRK
DSRGGSATNFSLLKQAGDVEENPGPMETLLGLLILWLQLQWVSSKQEVTPAALSVPEGEN
LVLNCSFTDSAIYNLQWFRQDPGKGLTSLLLITPWQREQTSGRNLNASLDKSSGSSTLYIAASQP
GDSATYLC AVRPLLDGTIYPTFGRGTSILVHPYIQNPDPVYQLRDSKSSDKSVCLFTDFDSQT
NVSQSKSDVYITDKTVLDMRSMDFKSNSAVAWSNKSDFACANAFNNSIIPEDTFFPSPSS
CDVKLVEKSFETDTNLFQNLVIGFRILLKLVAGFNLLMTLRLWSSAAA

>868 TCR (Beta Chain-P2A-Alpha Chain), amino acid sequence (SEQ ID NO: 31):

MSIGLLCCAALSLLWAGPVNADAGVTQSPHILIKTRGQVTLRCSKPKQGHDTVSWYQQALG
QGPQFIFQYEEEEERQRGNFPDRFSGHQFPNYSSSELNVNALLGDSALYLCASSDTVSYEQYF
GPGTRTLTVTEDLKNVFPPEVAVFEPSEAEISHTQKATLVCLATGFYDPDHVELSWVNGKEVH
SGVSTDPQPLKEQPALNDSRYCLSSRLRVSAFTWQDPRNNHFRQCQVQFYGLSENDEWTQDRA
KPVTQIVSAEAWGRADCGFTSESYQQGVLSATILYEILLGKATLYAVLVSAVLMLMAMVKRK
DSRGGSATNFSLLKQAGDVEENPGPMETLLGLLILWLQLQWVSSKEVEQNSGPLSVPEGAIA
SLNCTYSRGSQSFFWYRQYSGKSPELIMFIYSNGDKEDGRFTAQLNKASQYISLLIRDSKLS
SATYLC AVRTPNSGYALNFGKGTSLVTPHIQKPDPAVYQLRDSKSSDKSVCLFTDFDSQTNV
SQSKSDVYITDKTVLDMRSMDFKSNSAVAWSNKSDFACANAFNNSIIPEDTFFPSPSSCDV
KLVEKSFETDTNLFQNLVIGFRILLKLVAGFNLLMTLRLWSS

>CMV C7 TCR, amino acid sequence (SEQ ID NO: 33):

MGTRLLFWVAFCLLGADHTGAGVSQSPSNKVTEKGDVELRCDPISGHTALYWYRQRLGQ
GLEFLIYFQNSAPDKSGLPSDRFSAERTGESVSTLTQRTQQEDSAVYLCASSQTQLWETQY
FGPGTRLLVLEDLKNVFPPEVAVFEPSEAEISHTQKATLVCLATGFYDPDHVELSWVNGKEV
HSGVSTDPQPLKEQPALNDSRYCLSSRLRVSAFTWQNPNNHFRQCQVQFYGLSENDEWTQDR
AKPVTQIVSAEAWGRADCGFTSESYQQGVLSATILYEILLGKATLYAVLVSAVLMLMAMVKR
KDSRGGSATNFSLLKQAGDVEENPGPMETLLGLLILWFHLDVSSILNVEQSPQSLHVQE
GDSTNFTCSFPSSNFYALHWYRWETAKSPEALFVMTLNGDEKKKGRISATLNTKEGYSYLYI
KGSQPEDSATYLC AFITGNQFYFGTGTSLTVIPNIQNPDPVYQLRDSKSSDKSVCLFTDFDSQ

TNVSQSKSDSVYITDKTVLDMRSMDFKSNSAVAWSNKSDFACANAFNNSIIPEDTFFPSPESP
CDVKLVEKSFETDTNLFQNLVIGFRILLKLVAGFNLLMTLRLWSSAAA

>SL9 pMHC (SL9 peptide portion bolded), amino acid sequence (SEQ ID NO: 35):

MSRSVALAVLALLSLSGLEAS**LYNTVATL**GGGASGGGGSGGGGSIQRTPKIQVYSRHPAENG
KSNFLNCYVSGFHPSDIEVDLLKNGERIEKVEHSDLSFSKDWSFYLLYYTEFTPTKDEYACR
VNHVTLSPQKIVKWDRDMGGGGSGGGGSGGGGSGGGGSHSMRYFFTSVSRPGRGEPRFIAV
GYVDDTQFVRFDSDAASQRMEPRAPWIEQEGPEYWDGETRKKVKAHSQTHRVDLGLTRGAY
NQSEAGSHTVQRMYGCDVGSWDRFLRGYHQYAYDGKDYLKEDLRSWTAADMAAQTTK
HKWEAAHVAEQLRAYLEGTCVEWLRRYLENGKETLQRTDAPKTHMTHHAVSDHEATLRC
WALSFPYAEITLTWQRDGEDQTQDTELVETRPAGDGTQKWA AVVVPSGQEQRYTCHVQH
EGLPKPLTLRWEPSQPTIPVGHAGLVLFGAVITGAVVA AVMWRRKSS

>CMV pMHC (CMV NLV peptide portion bolded), amino acid sequence (SEQ ID NO: 37):

MSRSVALAVLALLSLSGLEAN**LVP**MTV**ATV**GGGASGGGGSGGGGSIQRTPKIQVYSRHPAEN
GKSNFLNCYVSGFHPSDIEVDLLKNGERIEKVEHSDLSFSKDWSFYLLYYTEFTPTKDEYACR
RVNHVTLSPQKIVKWDRDMGGGGSGGGGSGGGGSGGGGSHSMRYFFTSVSRPGRGEPRFIAV
VGYVDDTQFVRFDSDAASQRMEPRAPWIEQEGPEYWDGETRKKVKAHSQTHRVDLGLTRGA
YNQSEAGSHTVQRMYGCDVGSWDRFLRGYHQYAYDGKDYLKEDLRSWTAADMAAQTTK
KHKWEAAHVAEQLRAYLEGTCVEWLRRYLENGKETLQRTDAPKTHMTHHAVSDHEATLRC
CWALSFPYAEITLTWQRDGEDQTQDTELVETRPAGDGTQKWA AVVVPSGQEQRYTCHVQ
HEGLPKPLTLRWEPSQPTIPVGHAGLVLFGAVITGAVVA AVMWRRKSS

>disulfide trap MHC comprising Y84C mutation(bolded) and position 2 of the linker being a C (bolded); shown with the HGH Signal Peptide (underlined) and the GL9 peptide (italicized), amino acid sequence (SEQ ID NO: 39):

MATGSRTSLLLA**FGLLCLPW**LOEGSAGILGFVFTLGCSSGGSGGGSGGGGSIQRTPKIQVYSRH
PAENGKSNFLNCYVSGFHPSDIEVDLLKNGERIEKVEHSDLSFSKDWSFYLLYYTEFTPTKDEYACR
EYACRVNHVTLSPQKIVKWDRDMGGGGSGGGGSGGGGSGGGGSHSMRYFFTSVSRPGRG
EPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQEGPEYWDGETRKKVKAHSQTHRVDLGL
TRG**CYNQ**SEAGSHTVQRMYGCDVGSWDRFLRGYHQYAYDGKDYLKEDLRSWTAADM
AAQTTKHKWEAAHVAEQLRAYLEGTCVEWLRRYLENGKETLQRTDAPKTHMTHHAVSDH
EATLRCWALSFPYAEITLTWQRDGEDQTQDTELVETRPAGDGTQKWA AVVVPSGQEQRYT
CHVQHEGLPKPLTLRWEPSQPTIPVGHAGLVLFGAVITGAVVA AVMWRRKSS

>GL9 pMHC (GL9 peptide portion bolded), amino acid sequence (SEQ ID NO: 41):

MSRSVALAVLALLSLSGLEA**GILGFVFTL**GGGASGGGGSGGGGSIQRTPKIQVYSRHPAENG
KSNFLNCYVSGFHPSDIEVDLLKNGERIEKVEHSDLSFSKDWSFYLLYYTEFTPTKDEYACR
VNHVTLSPQKIVKWDRDMGGGGSGGGGSGGGGSGGGGSHSMRYFFTSVSRPGRGEPRFIAV
GYVDDTQFVRFDSDAASQRMEPRAPWIEQEGPEYWDGETRKKVKAHSQTHRVDLGLTRGAY
NQSEAGSHTVQRMYGCDVGSWDRFLRGYHQYAYDGKDYLKEDLRSWTAADMAAQTTK
HKWEAAHVAEQLRAYLEGTCVEWLRRYLENGKETLQRTDAPKTHMTHHAVSDHEATLRC
WALSFPYAEITLTWQRDGEDQTQDTELVETRPAGDGTQKWA AVVVPSGQEQRYTCHVQH
EGLPKPLTLRWEPSQPTIPVGHAGLVLFGAVITGAVVA AVMWRRKSS

>CD80, amino acid sequence (SEQ ID NO: 43):

MGHTRRQGTSPSKCPYLNFFQLLVLAGLSHFCSGVIHVTKEVKEVATLSCGHNVSVVEELAQT
RIYWQKEKKMVLTMMSGDMNIWPEYKNRTIFDITNNLSIVILALRPSDEGTYESCVLKYEKD
AFKREHLAEVTLVSKADFPTPSISDFEIPTSNIRRIICSTSGGFPEPHLSWLENGEELNAINTTVS
QDPETELYAVSSKLDNFMTTNHSFMCLIKYGHILRVNQTFNWNTTKQEHFPDNLPSWAITLI
SVNGIFVICCLTYCFAPRCRE

>CD86, amino acid sequence (SEQ ID NO: 45):

MDPQCTMGLSNILFVMAFLLSGAAPLKIQAYFNETADLPCQFANSQNQSLSELVVFWDQEN
LVLNEVYLGKEKFDVHVKYMGRTSFDSWTLRLHNLQIKDKGLYQCIHHKKPTGMIRIH
QMNSELSVLNFSQPEIVPISNITENVYINLTCSHIGYPEPKKMSVLLRRTKNSTIEYDGVMMQK
SQDNVTELYDVSISLSVSFPDVTNMTIFCILETDKTRLLSSPFSIELEDPQPPDHIPWITAVLPT
VIICVMVFCLILWKWKKKKR

>mouse anti-TCR Beta Clone H57-597 Fab antibody (light chain-P2A-heavy chain, PDGFR transmembrane domain), amino acid sequence: (SEQ ID NO: 47):

METDTLLLWVLLLWVPGSTGADYKDDDDKDIQMTQSPSSLPASLGDRVTINCQASQDISNYL
NWYQQKPGKAPKLLIYYTNKLADGVPSRFSGSGSGRDSSTFTISSESEDIGSYQCQYYNYP
WTFPGTKLEIKRADAKPTVSIFPPSSEQLGTGSATLVCFVNNFYPKDINVWKVVDGSEKRD
GVLQSVTDQDSKDYSLSTSLTKADYERHNLTYCEVTHKTSTAAIVKTLNRNECGSGAT
NFSLLKQAGDVEENPGPMVPCTLLLLLAAALAPTQTRAEVQLVESGGGLVQPGKSLKLSCEA
SGFTFSGYGMHWVRQAPGRGLESVAYITSSSINIKYADAVKGRFTVSRDANKNLLFLQMNIL
KSEDTAMYYCARFDWDKNYWGQGTMTVTVSSAKTTAPSVYPLAPACDSTTSTNTVTGLCL
VKGYFPEPVTVIWNSGALTSGVHTFPSVLHSLGLYSLSSSVTVPSSTWPSQTVTCNVAHPASST
TVDLKIEAVGQDTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQKKPR

>mouse anti-CD3 Epsilon Clone 145-2C11 Fab antibody (light chain-P2A-heavy chain, PDGFR transmembrane domain), amino acid sequence (SEQ ID NO: 49):

METDTLLLWVLLLWVPGSTGADYKDDDDKYEIQPSSASVTVGETVKITCSGDQLPKNFAY
WFQKSDKNILLIYMDNKRPSGIPERFSGSTSGTTATLTISGAQPEDEAAYYCLSSYGDND
LVFGSGTQLTVLRGPKSSPKVTVFPPSPEELRTNKATLVCLVNDYFPGSATVTWKANGATIN
DGVKTTKPSKQGNMTSSYLSLTADQWKSHNRVSCQVTHEGETVEKSLSPAELGSGATN
FSLKQAGDVEENPGPMVPCTLLLLLAAALAPTQTRAEVYLVESGGDLVQPGSSLKVSAAAS
GFTFSDFWMYWVRQAPGKGLEWVGRIKNIPNNYATEYADSVRGRFTISRDDSRNSIYLQMN
RLRVDDTAIYYCTRAGRFDHFDYWGQGTMTVTVSSATTTAPSVYPLAPACDSTTSTDTVTGL
CLVKGYFPEPVTVSWNSGALTSGVHTFPSVLHSLGLYSLSSSVTVPSSTWPKQPITCNVAHPAS
STKVDKKIEPRAVGQDTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQKKPR*

>Cocal Virus Glycoprotein, amino acid sequence: (SEQ ID NO: 51):

MNLLLLTFIVLPLCSHAKFSIVFPQSQKGNWKNVPSSYHYCPSSSDQNWHNDLLGITMKVKM
PKTHKAIQADGWMCHAAKWITTCDFRWYGPKYITHSIHQPTSEQCKESIKQTKQGTWMSP
GFPPQNCGYATVTDVAVVVQATPHHVLVDEYTGWIDSQFPNGKCETEECTVHNSTVWY
SDYKVTGLCDATLVDTEITFFSEDGKKESIGKPNTGYRSNYFAYEKGDVKCMNYCKHAGV
RLPSGVWFEFVDQDVYAAAKLPECPVGATISAPTQTSVDVSLILDVERILDYSLCQETWSKIR
SKQPVSPVDLSYLA PKNPGTGPAFTIINGTLKYFETRYIRIDIDNPIISKMVGKISGSQTERELWT
EWFPEYEGVEIGPNGILKTPTGYKFPLFMIGHGMLDSDLHKTSAEVEFHPHLAEAPKQLPEEE
TLFFGDTGISKNPVELIEGWFWSSWKSTVVTFFFAIGVFILLYVVARIVIAVRYRYQGSNNKRIY
NDIEMSRFRK*

>Cocal Virus Glycoprotein, DNA sequence: (SEQ ID NO: 52):

ATGAACCTTCTGCTGCTCACGTTTATCGTACTCCCGTTGTGCTCTCATGCGAAATTTTCAA
TAGTCTTTCTCAGTCCCAGAAAGGGAATTGGAATAATGTTCCCTCCAGTTACCACTATT
GTCCCTCTCCTCTGACCAAACTGGCACAATGACTTGCTCGGGATTACAATGAAAGTAA
AGATGCCGAAAACCCATAAAGCCATACAGGCGGATGGGTGGATGTGTCACGCTGCGAAG
TGGATCACTACATGCGATTTCCGGTGGTATGGCCCTAAGTACATTACACACTCTATCCAT
AGCATACAGCCGACATCAGAGCAATGCAAAGAGAGTATTAAACAGACCAAAACAAGGGA
CATGGATGAGCCCTGGCTTTCACCTCAGAATTGTGGGTACGCGACCGTCACGGATAGTG

TCGCTGTTGTGGTGCAGGCCACGCCACATCACGTACTCGTAGATGAATATACTGGTGAAT
 GGATCGACTCCCAATTCCCGAATGGGAAATGTGAGACGGAAGAGTGCGAAACAGTGCAT
 AACTCAACCGTTTGGTATTCCGATTACAAGGTTACTGGTCTTTGCGACGCCACCCTCGTG
 GATACCGAGATCACGTTTTTTAGTGAGGATGGCAAGAAAGAGTCAATAGGCAAACCTAA
 TACTGGCTACCGGAGTAACTATTTTCGCTTACGAGAAGGGTGACAAGGTATGTAAAATGA
 ACTATTGCAAGCATGCGGGAGTGCGACTCCCCAGTGGGGTATGGTTCGAATTTGTTGACC
 AAGACGTATACGCCGCTGCGAAGTTGCCAGAATGCCCCGTAGGCGCGACCATTTCAGCA
 CCTACCCAAACGTCCGTTGACGTCTCCTTGATACTGGATGTAGAGCGAATCCTGGACTAC
 AGTCTCTGCCAGGAAACGTGGTCAAAAATAAGAAGTAAGCAGCCAGTTTCACCCGTGGA
 TCTGTCTTATCTGGCGCCAAAAAACCCGGGCACGGGCCCTGCTTTTACCATAATTAACGG
 AACGCTTAAATACTTCGAAACCCGCTACATTAGAATCGATATAGACAATCCTATTATCAG
 CAAGATGGTAGGGAAGATATCTGGGTCTCAAACGGAGCGAGAATTGTGGACGGAGTGGT
 TCCCTTATGAGGGAGTGGAATTTGGGCCAACGGGATCCTCAAGACCCCAACGGGTTAC
 AAGTTCCCTCTGTTTATGATCGGCCATGGCATGTTGGACAGTGACTTGACAAAAACATCT
 CAGGCAGAGGTTTTCGAACATCCACATTTGGCGGAGGCGCCCAAGCAACTTCAGAGA
 AGAAACTCTCTCTTTGGAGATACAGGCATTTCAAAAAATCCTGTAGAATGATAGAAGG
 GTGGTCTCTCCTGGAAATCAACGGTTGTCACGTTTTTCTTTGCAATAGGCGTATTATA
 CTCCTGTACGTCGTAGCCCGCATTGTGATCGCAGTACGATACAGATACCAGGGCAGTAAC
 AATAAACGCATATATAATGACATCGAAATGTCAAGGTTCCGAAAGtga

>Cocal-dead (mutations to ablate native tropism **bolded** in protein sequence; these are K64Q and R371A, counting from the start codon), amino acid sequence: (SEQ ID NO: 53):

MNLLLLTFIVLPLCSHAKFSIVFPQSQKGNWKNVPSSYHYCPSSDQNWHDLLGITMKVKM
PQTHKAIQADGWMCHAAKWITTCDFRWYGPKYITHSIHSIQPTSEQCKESIKQTKQGTWMSP
 GFPPQNCGYATVTDVAVVVQATPHHVLVDEYTG EWIDSQFPNGKCETEECETVHNSTVWY
 SDYKVTGLCDATLVDTEITFFSEDGKKESIGKPNTGYRSNYFAYEKGDVKMNYCKHAGV
 RLPSGVWFEFVDQDVYAAKLPECPVGATISAPTQTSVDVSLILDVERILDYSLCQETWSKIR
 SKQVPSPVDLSYLA PKNPGTGPAFTIINGTLKYFETRYIRIDIDNPIISKMVGKISGSQTEAELW
 TEWFPYEGVEIGPN GILKTPTGYKFPLFMIGHGMLDSDLHKTSQAEVFEHPLAEAPKQLPEE
 ETLFFGDTGISKNPVELIEGW FSSWKSTVVTFFFAIGVFILLYVVARIVIAVRYRYQGSNNKRI
 YNDIEMSRFRK

>Cocal-dead (mutations to ablate native tropism), DNA sequence: (SEQ ID NO: 54):

ATGAACTTTCTGCTGCTCACGTTTATCGTACTCCCGTTGTGCTCTCATGCGAAATTTTCAA
 TAGTCTTTCTCAGTCCCAGAAAGGGAATTGGAAAAATGTTCCCTCCAGTTACCACTATT
 GTCCCTCCTCCTCTGACCAAACTGGCACAATGACTTGCTCGGGATTACAATGAAAGTAA
 AGATGCCGcagACCCATAAAGCCATACAGGCGGATGGGTGGATGTGTCACGCTGCGAAGT
 GGATCACTACATGCGATTTCCGGTGGTATGGCCCTAAGTACATTACACACTCTATCCATA
 GCATACAGCCGACATCAGAGCAATGCAAAGAGAGTATTAACAGACCAAAACAAGGGAC
 ATGGATGAGCCCTGGCTTTCCACCTCAGAATTGTGGGTACGCGACCGTCACGGATAGTGT
 CGCTGTTGTGGTGCAGGCCACGCCACATCACGTACTCGTAGATGAATATACTGGTGAATG
 GATCGACTCCCAATTCCCGAATGGGAAATGTGAGACGGAAGAGTGCGAAACAGTGCATA
 ACTCAACCGTTTGGTATTCCGATTACAAGGTTACTGGTCTTTGCGACGCCACCCTCGTGG
 ATACCGAGATCACGTTTTTTAGTGAGGATGGCAAGAAAGAGTCAATAGGCAAACCTAAT
 ACTGGCTACCGGAGTAACTATTTTCGCTTACGAGAAGGGTGACAAGGTATGTAAAATGAA
 CTATTGCAAGCATGCGGGAGTGCGACTCCCCAGTGGGGTATGGTTCGAATTTGTTGACCA
 AGACGTATACGCCGCTGCGAAGTTGCCAGAATGCCCCGTAGGCGCGACCATTTCAGCAC
 CTACCCAAACGTCCGTTGACGTCTCCTTGATACTGGATGTAGAGCGAATCCTGGACTACA
 GTCTCTGCCAGGAAACGTGGTCAAAAATAAGAAGTAAGCAGCCAGTTTCACCCGTGGAT
 CTGTCTTATCTGGCGCCAAAAAACCCGGGCACGGGCCCTGCTTTTACCATAATTAACGGA
 ACGCTTAAATACTTCGAAACCCGCTACATTAGAATCGATATAGACAATCCTATTATCAGC

AAGATGGTAGGGAAGATATCTGGGTCTCAAACGGAGgccGAATTGTGGACGGAGTGGTTC
CCTTATGAGGGAGTGGAAATTGGGCCCCAACGGGATCCTCAAGACCCCAACGGGTACAA
GTTCCCTCTGTTTATGATCGGCCATGGCATGTTGGACAGTGACTTGCACAAAACATCTCA
GGCAGAGGTTTTTCGAACATCCACATTTGGCGGAGGCGCCCAAGCAACTTCCAGAAGAAG
AAACTCTCTTCTTTGGAGATACAGGCATTTCAAAAAATCCTGTAGAACTGATAGAAGGGT
GGTTCTCTTCCTGGAAATCAACGGTTGTCACGTTTTTCTTTGCAATAGGCGTATTTATACT
CCTGTACGTCGTAGCCCGCATTGTGATCGCAGTACGATACAGATACCAGGGCAGTAACA
ATAAACGCATATATAATGACATCGAAATGTCAAGGTTCCGAAAGtga

CLAIMS

What is claimed is:

1. A method of screening a population of cells, the method comprising:
 - (i) providing a retrovirus comprising a viral envelope protein comprising at least one mutation that diminishes its native function, a non-viral membrane-bound protein comprising a membrane-bound domain and an extracellular targeting domain, and a nucleic acid encoding a reporter;
 - (ii) combining the retrovirus with a population of cells; and
 - (iii) sorting the population of cells based on the presence or absence of the reporter.
2. A method of screening a population of cells, the method comprising:
 - (i) providing a retrovirus comprising a nucleic acid comprising a non-viral membrane-bound protein comprising a structure: S-ETD-MBD-IRES-R, wherein S encodes a signal sequence, ETD encodes an extracellular targeting domain; MBD encodes a membrane-bound domain, IRES encodes an internal ribosome entry site, and R encodes a reporter; and a mutated viral envelope protein comprising at least one mutation that diminishes its native function;
 - (ii) combining the retrovirus with a population of cells; and
 - (iii) sorting the population of cells based on the presence or absence of the reporter.
3. The method of either one of claims 1 or 2, wherein the cells are somatic cells, optionally antigen-specific cells.
4. The method of claim 3, wherein the antigen-specific cells are T cells or B cells.
5. The method of any one of claims 1 to 3, wherein the retrovirus is a lentivirus.

6. The method of any one of claims 1 to 5, wherein the viral envelope protein is a VSV-G envelope protein, a measles virus envelope protein, a nipah virus envelope protein, or a cocal virus G protein.
7. The method of claim 6, where the at least one mutation of the VSV-G envelope protein is a mutation selected from the group consisting of H8, K47, Y209, and R354.
8. The method of claim 6, where the at least one mutation of the measles virus envelope protein is a mutation selected from the group consisting of Y481, R533, S548, and F549.
9. The method of claim 6, where the at least one mutation of the nipah virus envelope protein is a mutation selected from the group consisting of E501, W504, Q530, and E533.
10. The method of claim 6, where the at least one mutation of the cocal virus G protein is a mutation selected from the group consisting of K64 and R371.
11. The method of any one of claims 1 to 10, wherein the non-viral membrane-bound protein comprises a Major Histocompatibility Complex (MHC) protein.
12. The method of any one of claims 1 to 11, wherein the extracellular targeting domain is a protein, a peptide, or an antibody.
13. The method of claim 12, wherein the protein is interleukin-13 or a CD80 protein domain.
14. The method of claim 12, wherein the antibody is an anti-CD19 antibody, an anti-TCR antibody, or an anti-CD3 antibody.
15. The method of any one of claims 1 to 14, wherein the reporter is a fluorescent protein (*e.g.*, green fluorescent protein, yellow fluorescent protein, red fluorescent protein) or an antibiotic resistance marker.

16. The method of any one of claims 1 to 15, wherein a linker is positioned between the membrane-bound domain and the extracellular targeting domain.
17. The method of claim 16, wherein the linker is a rigid linker, optionally comprising a PDGFR stalk or a CD8 α stalk.
18. The method of claim 16 wherein the linker is a flexible linker, optionally comprising an amino acid sequence comprising GAPGAS (SEQ ID NO: 5) or GGGGS (SEQ ID NO: 7).
19. The method of claim 16, wherein the linker is an oligomerized linker, optionally comprising an IgG4 hinge or an amino acid sequence that can form a tetrameric coiled coil.
20. The method of any one of claims 1 to 19, wherein the retrovirus is combined with the population of cells in (ii) for one minute to seventy-two hours.
21. The method of any one of claims 1 to 20, wherein the retrovirus and the population of cells are combined in (ii) in the presence of:
- a) a cell culture media, optionally RPMI or DMEM cell culture media;
 - b) a buffered saline solution, optionally phosphate-buffered saline or HEPES-buffered saline; and/or
 - c) an enhancer of retroviral transduction, optionally heparin sulfate, polybrene, protamine sulfate, and/or dextran.
22. The method of any one of claims 1 to 21, wherein the retrovirus and the population of cells are combined in (ii) at a temperature ranging from 4°C to 42°C.
23. The method of any one of claims 1 to 22, wherein the extracellular targeting domain is capable of binding to a cognate protein that is present on the cell surface of a subset of the population of cells.

24. The method of claim 23, wherein the cognate protein is protein receptor.
25. The method of either one of claims 23 or 24, wherein the extracellular targeting domain of a single retrovirus binds to the cognate protein that is present on the cell surface of a single T cell during (ii), leading to entry of the retrovirus into the T cell.
26. The method of any one of claims 1 to 25, further comprising washing the population of cells between (ii) and (iii), optionally using phosphate-buffered saline (PBS), optionally removing the retrovirus from the population of cells.
27. The method of any one of claims 1 to 26 wherein sorting the population of cells is performed using fluorescence-activated cell sorting, single-cell next-generation sequencing, or antibiotic selection.
28. The method of any one of claims 1 to 27, wherein the population of cells is a population of cells isolated from a subject, optionally wherein the subject is a human.
29. The method of claims 28, wherein the population of cells are isolated from the blood or a tumor of a subject.
30. The method of any one of claims 1 to 29, wherein the population of cells are maintained in liquid culture prior to being combined with the retrovirus.
31. The method of any one of claims 1 to 30, further comprising a second retrovirus, wherein the second retrovirus comprises a different extracellular targeting domain and/or a different reporter compared to the first retrovirus.
32. The method of any one of claims 2 to 31, wherein:
- a) the signal sequence comprises SEQ ID NO: 1 (METDTLLLWVLLLWVPGSTG) or SEQ ID NO: 2 (MSRSVALAVLALLSLSGLEA); and/or

b) the IRES encodes a sequence derived from encephalomyocarditis virus (EMCV).

33. A retrovirus, comprising

i) a nucleic acid comprising a non-viral membrane-bound protein comprising a structure: S-ETD-MBD-IRES-R, wherein S encodes a signal sequence, ETD encodes an extracellular targeting domain; MBD encodes a membrane-bound domain, IRES encodes an internal ribosome entry site, and R encodes a reporter; and

ii) a mutated viral envelope protein comprising at least one mutation that diminishes its native function.

34. The retrovirus of claim 33, wherein the retrovirus is a lentivirus.

35. The retrovirus of either one of claims 33 or 34, wherein the viral envelope protein is a VSV-G envelope protein, a measles virus envelope protein, a nipah virus envelope protein, or a cocal virus G protein.

36. The retrovirus of claim 35, where the at least one mutation of a VSV-G envelope protein is a mutation selected from the group consisting of H8, K47, Y209, and R354.

37. The retrovirus of claim 36, where the at least one mutation of the measles virus envelope protein is a mutation selected from the group consisting of Y481, R533, S548, and F549.

38. The retrovirus of claim 36, where the at least one mutation of the nipah virus envelope protein is a mutation selected from the group consisting of E501, W504, Q530, and E533.

39. The retrovirus of claim 36, where the at least one mutation of the cocal virus G protein is a mutation selected from the group consisting of K64 and R371.

40. The retrovirus of any one of claims 33 to 39, wherein the non-viral membrane-bound protein comprises a Major Histocompatibility Complex (MHC) protein.
41. The retrovirus of any one of claims 33 to 40, wherein the extracellular targeting domain is a protein, a peptide, or an antibody.
42. The retrovirus of claim 41, wherein the protein is interleukin-13 or a CD80 protein domain.
43. The retrovirus of claim 41, wherein the antibody is an anti-CD19 antibody, an anti-TCR antibody, or an anti-CD3 antibody.
44. The retrovirus of any one of claims 33 to 43, wherein the reporter is a fluorescent protein (*e.g.*, green fluorescent protein, yellow fluorescent protein, red fluorescent protein) or an antibiotic resistance marker.
45. The retrovirus of any one of claims 33 to 44, wherein a linker is positioned between the membrane-bound domain and the extracellular targeting domain.
46. The retrovirus of claim 45, wherein the linker is a rigid linker, optionally comprising a PDGFR stalk or a CD8 α stalk.
47. The retrovirus of claim 45, wherein the linker is a flexible linker, optionally comprising an amino acid sequence comprising GAPGAS (SEQ ID NO: 5) or GGGGS (SEQ ID NO: 7).
48. The retrovirus of claim 45, wherein the linker is an oligomerized linker, optionally comprising an IgG4 hinge or an amino acid sequence that can form a tetrameric coiled coil.

49. The retrovirus of any one of claims 33 to 48, wherein the extracellular targeting domain is capable of binding to a cognate protein that is present on the cell surface of a cell, optionally wherein the cell is an antigen-specific cell (*e.g.*, T cell or B cell).

50. The retrovirus of any one of claims 33 to 49, wherein the signal sequence comprises SEQ ID NO: 1 (METDTLLLWVLLLWVPGSTG) or SEQ ID NO: 2 (MSRSVALAVLALLSLSGLEA).

51. The retrovirus of any one of claims 33 to 50, wherein the IRES encodes a sequence derived from encephalomyocarditis virus (EMCV).

52. A method of delivering a nucleic acid to a cell, the method comprising:

- (i) providing a retrovirus comprising the nucleic acid, a viral envelope protein comprising at least one mutation that diminishes its native function, and a non-viral membrane-bound protein comprising an extracellular targeting domain that is capable of binding to a cognate ligand of the cell; and
- (ii) contacting the retrovirus with the cell, thereby delivering the nucleic acid to the cell.

53. The method of claim 51, wherein the nucleic acid encodes a gene of interest, optionally wherein the gene of interest encodes a protein.

54. The method of either one of claims 52 or 53, wherein the cognate ligand is a protein, optionally a protein receptor.

55. The method of any one of claims 52 to 54, wherein the retrovirus enters or infects the cell during (ii).

56. The method of any one of claims 52 to 55, wherein the cell is a somatic cell, optionally an antigen-specific cell.

57. The method of claim 56, wherein the antigen-specific cell is a T cell or a B cell.
58. The method of any one of claims 52 to 57, wherein the retrovirus is a lentivirus.
59. The method of any one of claims 52 to 58, wherein the viral envelope protein is a VSV-G envelope protein, a measles virus envelope protein, a nipah virus envelope protein, or a coxsa virus G protein.
60. The method of claim 59, where the at least one mutation of a VSV-G envelope protein is a mutation selected from the group consisting of H8, K47, Y209, and R354.
61. The method of claim 59, where the at least one mutation of the measles virus envelope protein is a mutation selected from the group consisting of Y481, R533, S548, and F549.
62. The method of claim 59, where the at least one mutation of the nipah virus envelope protein is a mutation selected from the group consisting of E501, W504, Q530, and E533.
63. The method of claim 59, where the at least one mutation of the coxsa virus G protein is a mutation selected from the group consisting of K64 and R371.
64. The method of any one of claims 52 to 63, wherein the non-viral membrane-bound protein comprises a Major Histocompatibility Complex (MHC) protein.
65. The method of any one of claims 52 to 64, wherein the extracellular targeting domain is a protein, a peptide, or an antibody.
66. The method of claim 65, wherein the protein is interleukin-13 or a CD80 protein domain.

67. The method of claim 65, wherein the antibody is an anti-CD19 antibody, an anti-TCR antibody, or an anti-CD3 antibody.
68. The method of any one of claims 52 to 67, wherein a linker is positioned between the membrane-bound domain and the extracellular targeting domain.
69. The method of claim 68, wherein the linker is a rigid linker, optionally comprising a PDGFR stalk or a CD8 α stalk.
70. The method of claim 68, wherein the linker is a flexible linker, optionally comprising an amino acid sequence comprising GAPGAS (SEQ ID NO: 5) or GGGGS (SEQ ID NO: 7).
71. The method of claim 68, wherein the linker is an oligomerized linker, optionally comprising an IgG4 hinge or an amino acid sequence that can form a tetrameric coiled coil.
72. The method of any one of claims 52 to 71, wherein the extracellular targeting domain is capable of binding to a cognate protein that is present on the cell surface of a T cell.
73. A method of detecting an interaction between a retrovirus and an cell, the method comprising:
- (i) contacting a sample comprising the retrovirus and an cell with an antibody, wherein the retrovirus comprises a viral envelope protein comprising at least one mutation that diminishes its native function, a non-viral membrane-bound protein comprising an extracellular targeting domain, and wherein the antibody binds to the extracellular targeting domain of the retrovirus;
 - (ii) optionally removing unbound antibody from the sample; and
 - (iii) imaging the sample to detect whether the antibody-retrovirus complex is bound to the cell.
74. The method of claim 73, wherein the antibody further comprises a fluorescent label, optionally wherein the antibody is covalently linked to a fluorescent label.

75. The method of either one of claims 73 or 74, wherein the sample is imaged in (iii) using confocal or fluorescence microscopy.
76. The method of any one of claims 73 to 75, wherein the cell is a somatic cell, optionally an antigen-specific cell.
77. The method of claim 76, wherein the antigen-specific cell is a T cell or a B cell.
78. The method of any one of claims 73 to 77, wherein the retrovirus is a lentivirus.
79. The method of any one of claims 73 to 78, wherein the viral envelope protein is a VSV-G envelope protein, a measles virus envelope protein, a nipah virus envelope protein, or a coxal virus G protein.
80. The method of claim 79, where the at least one mutation of a VSV-G envelope protein is a mutation selected from the group consisting of H8, K47, Y209, and R354.
81. The method of claim 79, where the at least one mutation of the measles virus envelope protein is a mutation selected from the group consisting of Y481, R533, S548, and F549.
82. The method of claim 79, where the at least one mutation of the nipah virus envelope protein is a mutation selected from the group consisting of E501, W504, Q530, and E533.
83. The method of claim 79, where the at least one mutation of the coxal virus G protein is a mutation selected from the group consisting of K64 and R371.
84. The method of any one of claims 73 to 83, wherein the non-viral membrane-bound protein comprises a Major Histocompatibility Complex (MHC) protein.

85. The method of any one of claims 73 to 84, wherein the extracellular targeting domain is a protein or a peptide.
86. The method of claim 85, wherein the protein is interleukin-13 or a CD80 protein domain.
87. The method of any one of claims 73 to 86, wherein a linker is positioned between the membrane-bound domain and the extracellular targeting domain.
88. The method of claim 87, wherein the linker is a rigid linker, optionally comprising a PDGFR stalk or a CD8 α stalk.
89. The method of claim 87, wherein the linker is a flexible linker, optionally comprising an amino acid sequence comprising GAPGAS (SEQ ID NO: 5) or GGGGS (SEQ ID NO: 7).
90. The method of claim 87, wherein the linker is an oligomerized linker, optionally comprising an IgG4 hinge or an amino acid sequence that can form a tetrameric coiled coil.
91. A library of retroviruses, comprising a plurality of unique retroviruses,
wherein each unique retrovirus comprises a viral envelope protein comprising at least one mutation that diminishes its native function, a non-viral membrane-bound protein comprising a membrane-bound domain and an extracellular targeting domain, and a nucleic acid encoding a reporter, and
wherein each unique retrovirus comprises a different and unique extracellular targeting domain.
92. The library of claim 91, wherein the library is capable of being screened against a population of antigen-specific cells, optionally wherein the antigen-specific cells are B cells or T cells.

93. The library of claim 91 or 92, wherein the library comprises at least 10^2 , at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 , at least 10^8 , at least 10^9 , or at least 10^{10} unique retroviruses.
94. The library of any one of claims 91 to 93, wherein each extracellular targeting domain comprises at least 5, at least 10, at least 15, at least 20, or at least 50 amino acids.
95. The library of any one of claims 91 to 94, wherein the non-viral membrane-bound protein comprises a Major Histocompatibility Complex (MHC) protein.
96. The library of any one of claims 91 to 95, wherein each different and unique extracellular targeting domain is generated through site-directed mutagenesis.
97. A population of cells, wherein a subset of the population of cells contain a retrovirus comprising: a viral envelope protein comprising at least one mutation that diminishes its native function, a non-viral membrane-bound protein comprising a membrane-bound domain and an extracellular targeting domain, and a nucleic acid encoding a reporter.
98. A population of cells, wherein a subset of the population of cells contains a retrovirus of any one of claims 33 to 51.
99. The population of claim 97 or 98, wherein the cells are antigen-specific cells, optionally B cells or T cells.
100. The population of any one of claims 97 or 98, wherein the subset of the population of cells contains the retrovirus inside each cell of the subset, optionally inside the nucleus of each cell of the subset.
101. The population of any one of claims 97 to 100, wherein the subset of the population expresses the reporter, optionally wherein the reporter is a fluorescent protein.

102. The population of any one of claims 97 to 101, wherein the subset of the population that contain the retrovirus are isolated and/or sorted from the cells of the population that do not contain the retrovirus.

103. A retrovirus, comprising

- i) a CD80 protein domain; and
- ii) a mutated viral envelope protein comprising at least one mutation that diminishes its native function.

104. The retrovirus of claim 103, wherein the retrovirus is a lentivirus.

105. The retrovirus of either one of claims 103 or 104, wherein the viral envelope protein is a VSV-G envelope protein, a measles virus envelope protein, a nipah virus envelope protein, or a cocal virus G protein.

106. The retrovirus of claim 105, where the at least one mutation of a VSV-G envelope protein is a mutation selected from the group consisting of H8, K47, Y209, and R354.

107. The retrovirus of claim 105, where the at least one mutation of the measles virus envelope protein is a mutation selected from the group consisting of Y481, R533, S548, and F549.

108. The retrovirus of claim 105, where the at least one mutation of the nipah virus envelope protein is a mutation selected from the group consisting of E501, W504, Q530, and E533.

109. The retrovirus of claim 105, where the at least one mutation of the cocal virus G protein is a mutation selected from the group consisting of K64 and R371.

110. The retrovirus of any one of claims 103 to 109, wherein the CD80 protein domain is an extracellular domain, optionally wherein the extracellular domain binds to a receptor on a target cell.

111. The retrovirus of any one of claims 103 to 110, further comprising a non-viral membrane-bound protein comprising an extracellular targeting domain that is capable of binding to a cognate ligand of a cell.

112. The retrovirus of claim 111, wherein the non-viral membrane-bound protein comprises a Major Histocompatibility Complex (MHC) protein.

113. The retrovirus of claim 111 or 112, wherein the extracellular targeting domain is a protein, a peptide, or an antibody.

114. A method of delivering a nucleic acid to a cell, the method comprising:

- (i) providing a retrovirus comprising the nucleic acid, a viral envelope protein comprising at least one mutation that diminishes its native function, and a CD80 protein domain; and
- (ii) contacting the retrovirus with the cell, thereby delivering the nucleic acid to the cell.

115. The method of claim 114, wherein the nucleic acid encodes a gene of interest, optionally wherein the gene of interest encodes a protein.

116. The method of either one of claims 114 or 115, wherein the retrovirus enters or infects the cell during (ii).

117. The method of any one of claims 114 to 116, wherein the cell is a somatic cell, optionally an antigen-specific cell.

118. The method of claim 117, wherein the antigen-specific cell is a T cell or a B cell.

119. The method of any one of claims 114 to 118, wherein the retrovirus is a lentivirus.
120. The method of any one of claims 114 to 119, wherein the viral envelope protein is a VSV-G envelope protein, a measles virus envelope protein, a nipah virus envelope protein, or a coxsa virus G protein.
121. The method of claim 120, where the at least one mutation of a VSV-G envelope protein is a mutation selected from the group consisting of H8, K47, Y209, and R354.
122. The method of claim 120, where the at least one mutation of the measles virus envelope protein is a mutation selected from the group consisting of Y481, R533, S548, and F549.
123. The method of claim 120, where the at least one mutation of the nipah virus envelope protein is a mutation selected from the group consisting of E501, W504, Q530, and E533.
124. The method of claim 120, where the at least one mutation of the coxsa virus G protein is a mutation selected from the group consisting of K64 and R371.
125. The method of any one of claims 114 to 124, wherein the CD80 protein domain is an extracellular domain, optionally wherein the extracellular domain binds to a receptor on the cell.