

SUMMARY

[0006] The methods and compositions described herein enable rapid identification of antigens targeted by an immune response and isolation of the T cells and thus the TCR sequences mediating that response. This rapid identification is a goal of research and drug development in the fields of infectious disease, tumor immunology, autoimmunity, and immunotherapy. The compositions may be applied to identify and to isolate antigen-specific T cells and their encoded MHC I-restricted TCRs for drug development. The identification and characterization of antigen-specific T cells may also be applied in the diagnosis, monitoring and prognosis of immune responses and disease in the context of (pre- and post-) therapeutic treatment. The comPACT approach may also be applied to identify the targets of MHC II-restricted T cell responses. Such responses are central to autoimmune disease, and are emerging as important components of cancer immunotherapy. In addition, the efficiency of comPACT protein secretion from mammalian cells expression data generated in the method described herein is potentially useful as a surrogate metric of proper protein folding and thus the affinity of each neoepitope ligand for its cognate MHC molecule. This technology may then afford the unique ability to refine the MHC binding algorithms for less-well studied MHC alleles, improving other targeted immunotherapies such as neoepitope vaccines. As such, the compositions disclosed herein will have broad application to human health and clear commercial potential.

[0006a] In one aspect, provided herein is a library comprising at least two different polynucleotide molecules, wherein each polynucleotide molecule comprises, in a 5' to 3' orientation, a sequence insert comprising stop codons in each of the different reading frames, a Beta 2 Microglobulin (β 2M) sequence, and a Major Histocompatibility Complex (MHC) allele sequence.

[0006b] In one aspect, provided herein is a library comprising at least two different polynucleotide molecules, wherein each polynucleotide molecule comprises, in a 5' to 3' orientation, a first universal target sequence, a nucleotide sequence encoding an antigenic peptide, a second universal target sequence that is distinct from the first universal target sequence, a Beta 2 Microglobulin (β 2M) sequence, and a Major Histocompatibility Complex (MHC) allele sequence.

[0006c] In one aspect, provided herein is a kit comprising a library as described herein and instructions for use.

[0007] In one ~~aspect~~example, disclosed herein is a polynucleotide molecule comprising, in a 5' to 3' orientation, (i) a first universal target sequence, (ii) a nucleotide sequence encoding an

antigenic peptide, (iii) a second universal target sequence that is distinct from the first universal target sequence, (iv) a Beta 2 Microglobulin (β 2M) sequence, and (v) an Multiple Histocompatibility Complex (MHC) allele sequence. In one ~~embodiment~~[example](#), the polynucleotide molecule comprises, in a 5' to 3' orientation, (i) a promoter sequence, (ii) the first universal target sequence comprising the sequence shown in SEQ ID NO: 3, (iii) the nucleotide sequence encoding an antigenic peptide, wherein the antigenic peptide is a tumor neoantigen, (iv) the second universal target sequence comprising the sequence shown in SEQ ID NO: 4, (v) the β 2M sequence comprising the sequence shown in SEQ ID NO: 106, (vi) the MHC allele sequence comprising a sequence selected from the group consisting of the sequences shown in SEQ ID NOs:109-174, (vii) a first affinity tag sequence comprising the sequence shown in SEQ ID NO: 29, (viii) a protease cleavage site sequence comprising the sequence shown in SEQ ID NO: 31, and (ix) a second affinity tag sequence comprising the sequence shown in SEQ ID NO: 35.

[0008] In one [embodimentexample](#), the antigenic peptide is selected from the group consisting of a tumor antigen, a neoantigen, a tumor neoantigen, a viral antigen, bacterial antigen, phosphoantigen, and a microbial antigen. In another [embodimentexample](#), the antigenic peptide is a neoantigen.

[0009] In some [embodimentexamples](#) a neoantigen is selected by analyzing tumor sequencing data from a subject to identify one or more somatic mutations. In one [embodimentexample](#), the analyzing is performed using an in silico predictive algorithm. In another [embodimentexample](#), the predictive algorithm further comprises an MHC binding algorithm to predict binding between the neoantigen and an MHC allele.

[0010] In some [embodimentexamples](#), the MHC allele is a mammalian MHC allele. In some [embodimentexamples](#), the MHC allele is a human MHC allele. In some [embodimentexamples](#), the MHC allele is a class I HLA allele. In other [embodimentexamples](#), the HLA allele comprises an HLA-A, HLA-B, or HLA-C allele. The HLA allele is an HLA allele of the subject.

[0011] In some [embodimentexamples](#), the HLA allele is selected from the group consisting of: HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*24:02, HLA-A*30:02, HLA-A*31:01, HLA-A*32:01, HLA-A*33:01, HLA-A*68:01, HLA-A*11:01, HLA-A*23:01, HLA-A*30:01, HLA-A*33:03, HLA-A*25:01, HLA-A*26:01, HLA-A*29:02, HLA-A*68:02, HLA-B*07:02, HLA-B*14:02, HLA-B*18:01, HLA-B*27:02, HLA-B*39:01, HLA-B*40:01, HLA-B*44:02, HLA-B*46:01, HLA-B*50:01, HLA-B*57:01, HLA-B*58:01, HLA-B*08:01, HLA-B*15:01, HLA-B*15:03, HLA-B*35:01, HLA-B*40:02, HLA-B*42:01, HLA-B*44:03, HLA-B*51:01, HLA-B*53:01, HLA-B*13:02, HLA-B*15:07, HLA-B*27:05, HLA-B*35:03, HLA-B*37:01, HLA-B*38:01, HLA-B*41:02, HLA-B*44:05, HLA-B*49:01, HLA-B*52:01, HLA-B*55:01, HLA-C*02:02, HLA-C*03:04, HLA-C*05:01, HLA-C*07:01, HLA-C*01:02, HLA-C*04:01, HLA-C*06:02, HLA-C*07:02, HLA-C*16:01, HLA-C*03:03, HLA-C*07:04, HLA-C*08:01, HLA-C*08:02, HLA-C*12:02, HLA-C*12:03, HLA-C*14:02, HLA-C*15:02, and HLA-C*17:01.

[0012] In some [embodimentexamples](#), the HLA allele comprises a sequence selected from the group consisting of SEQ ID NOs: 109-174.

[0013] In some [embodimentexamples](#), the β 2M allele is a mammalian β 2M allele. In some [embodimentexamples](#), the β 2M allele is a human β 2M allele. In some [embodimentexamples](#), the β 2M allele comprises the sequence shown in SEQ ID NO: 106.

[0014] In some [embodimentexamples](#), the nucleotide sequence encoding an antigenic peptide is between 20-60, between 20-30, between 25-35, between 20-45, between 30-45, between 40-

60, or between 45-60 nucleotides in length. In one [embodimentexample](#), the nucleotide sequence encoding an antigenic peptide is between 20-30 nucleotides in length.

[0015] In some [embodimentexamples](#), the first universal target sequence is between 4-50, between 4-15, between 15-40, between 15-35, between 15-30, between 20-40, between 25-40, or between 30-40 nucleotides in length. In one [embodimentexample](#), the first universal target sequence is between 25-35 nucleotides in length. In another [embodimentexample](#), the first universal target sequence is at least about 15 nucleotides in length. In another [embodimentexample](#), wherein the first universal target sequence is between 4-6 nucleotides in length. In some [embodimentexamples](#), the second universal target sequence is between 4-50, between 4-15, between 15-40, between 15-35, between 15-30, between 20-40, between 25-40, or between 30-40 nucleotides in length. In one [embodimentexample](#), the second universal target sequence is between 25-35 nucleotides in length. In another [embodimentexample](#), the second universal target sequence is at least about 15 nucleotides in length. In another [embodimentexample](#), the second universal target sequence is between 4-6 nucleotides in length.

[0016] In some [embodimentexamples](#), the first and second universal target sequences comprise polymerase chain reaction (PCR) primer target sequences. In other [embodimentexamples](#), the first and second universal target sequences comprise restriction enzyme cleavage sites.

[0017] In some [embodimentexamples](#), the sequence of the first universal target comprises the sequence shown in SEQ ID NO: 3. In some [embodimentexamples](#), the sequence of the second universal target comprises the sequence shown in SEQ ID NO: 4.

[0018] In some [embodimentexamples](#), the first universal target sequence further comprises a signal sequence. In some [embodimentexamples](#), the signal sequence encodes a signal sequence comprising a Human Growth Hormone signal sequence, a hIG1 Kappa light chain signal sequence, a Beta 2 microglobulin, signal sequence, or an IL2 signal sequence. In some [embodimentexamples](#), the signal sequence comprises a sequence comprising the sequence shown in SEQ ID NOs: 1, 24, 26, or 28. In one [embodimentexample](#), the signal sequence encodes Human Growth Hormone (HGH). In some [embodimentexamples](#), the signal sequence comprises the sequence shown in SEQ ID NO: 1. In one [embodimentexample](#), the signal sequence is between 40-90, 40-60, 45-70, 50-80, 60-90, 55-70, 60-80, or 70-80 nucleotides in length.

[0019] In some [embodimentexamples](#), the 3' end of the polynucleotide sequence further comprises, in a 5' to 3' orientation, a purification cluster sequence comprising (i) a first affinity tag sequence, (ii) a protease cleavage site sequence, and (iii) a second affinity tag sequence. In some [embodimentexamples](#), the first and second affinity tags are selected from the group consisting of: AviTag, streptavidin-tag, polyhistidine (His6)-tag, FLAG-tag, HA-tag, and Myc-tag. In

some [embodimentexamples](#), the first affinity tags comprises a sequence comprising the sequence shown in SEQ ID NO: 29, and the second affinity tag comprises a sequence comprising the sequence shown in SEQ ID NO: 33.

[0020] In some [embodimentexamples](#), the protease cleavage site sequence is selected from the group consisting of: a TEV cleavage site sequence, a thrombin cleavage site sequence, a Factor Xa cleavage site sequence, an enteropeptidase cleavage site sequence, and a rhinovirus 3C protease cleavage site sequence. In one [embodimentexample](#), the protease cleavage site sequence comprises a TEV cleavage site sequence shown in SEQ ID NO: 31. In some [embodimentexamples](#), the first affinity tag sequence encodes for an AviTag peptide, the protease cleavage site sequence encodes for a TEV cleavage site, and the second affinity tag encodes for a His6 peptide. In one [embodimentexample](#), the first affinity tag sequence comprises the sequence shown in SEQ ID NO: 29, the protease cleavage site comprises the sequence shown in SEQ ID NO: 31, and the second affinity tag comprises the sequence shown in SEQ ID NO: 33.

[0021] In some [embodimentexamples](#), the second universal target sequence further comprises a first linker sequence. In one [embodimentexample](#), the first linker sequence comprises the sequence shown in SEQ ID NO: 10, 14, 16, or 18.

[0022] In some [embodimentexamples](#), the polynucleotide further comprises a second linker sequence between the β 2M sequence and the MHC allele sequence. In one [embodimentexample](#), the second linker sequence comprises the sequence shown in SEQ ID NO: 10 or 20.

[0023] In some [embodimentexamples](#), the polynucleotide further comprises a third linker sequence between the MHC allele sequence and the first affinity tag. In one [embodimentexample](#), the third linker sequence comprises the sequence shown in SEQ ID NO: 12 or 22.

[0024] The 5' end of the polynucleotide sequence may further comprise a promoter sequence linked to the 5' end of the first universal target. In some [embodimentexamples](#), the promoter sequence is a CMV, an EF1 α , or an SV40 promoter. In one [embodimentexample](#), the promoter sequence is a CMV promoter. In one [embodimentexample](#), the promoter sequence is a CMV promoter sequence.

[0025] The 3' end of the polynucleotide sequence may further comprise a polyA sequence. In some [embodimentexamples](#), the polyA sequence is an SV40, hGH, BHG, or rbGlob polyA sequence. In one [embodimentexample](#), the polyA sequence comprises a bGH polyA sequence as shown in SEQ ID NO: 179.

[0026] In one [embodimentexample](#), the polynucleotide molecule comprises, in a 5' to 3' orientation, (i) a promoter sequence, (ii) a first universal target sequence, (iii) a nucleotide sequence encoding an antigenic peptide, (iv) a second universal target sequence that is distinct from the first universal target sequence, (v) a β 2M sequence, (vi) an MHC allele sequence, (vi) a first

affinity tag sequence, (vii) a protease cleavage site sequence, and (viii) a second affinity tag sequence.

[0027] In another [aspectexample](#), described herein is an expression construct comprising the polynucleotide molecule as disclosed herein. In one [embodimentexample](#), the expression construct comprises a plasmid or a viral vector.

[0028] In another [aspectexample](#), described herein is a host cell comprising the polynucleotide molecule or the expression construct as previously described. In one [embodimentexample](#), the polynucleotide is integrated into the cell genome. In another [embodimentexample](#), the polynucleotide is extrachromosomal. In some [embodimentexamples](#), the host cell is a mammalian cell. In some [embodimentexamples](#), the host cell is a human cell. In one [embodimentexample](#), the cell is a stem cell, a tumor cell, an immortalized cell, or a fetal cell. In some [embodimentexamples](#), the host cell is a prokaryotic cell. In some [embodimentexamples](#), the cell is an *Escherichia coli* cell. In some [embodimentexamples](#), the cell expresses a BirA protein or fragment thereof.

[0029] In yet another [aspectexample](#), described herein is a library comprising the polynucleotide molecule or expression construct as described herein, wherein the library comprises greater than or equal to two distinct polynucleotide molecules, wherein each distinct polynucleotide molecule comprises (i) the first universal sequence, (ii) the nucleotide sequence encoding a antigenic peptide, wherein the nucleotide sequence is not the same for each of the greater than or equal to two polynucleotide molecules (iii) the second universal target sequence, (iv) the β 2M sequence, and (v) the MHC allele sequence. In one [embodimentexample](#), the MHC allele sequence is not the same for each of the greater than or equal to two polynucleotide molecules. In another [embodimentexample](#), the library comprises 20 to 500 distinct polynucleotide molecules. In some [embodimentexamples](#), the library comprises at least 66 distinct polynucleotide molecules.

[0030] In some [embodimentexamples](#), the library comprises at least the HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*24:02, HLA-A*30:02, HLA-A*31:01, HLA-A*32:01, HLA-A*33:01, HLA-A*68:01, HLA-A*11:01, HLA-A*23:01, HLA-A*30:01, HLA-A*33:03, HLA-A*25:01, HLA-A*26:01, HLA-A*29:02, HLA-A*68:02, HLA-B*07:02, HLA-B*14:02, HLA-B*18:01, HLA-B*27:02, HLA-B*39:01, HLA-B*40:01, HLA-B*44:02, HLA-B*46:01, HLA-B*50:01, HLA-B*57:01, HLA-B*58:01, HLA-B*08:01, HLA-B*15:01, HLA-B*15:03, HLA-B*35:01, HLA-B*40:02, HLA-B*42:01, HLA-B*44:03, HLA-B*51:01, HLA-B*53:01, HLA-B*13:02, HLA-B*15:07, HLA-B*27:05, HLA-B*35:03, HLA-B*37:01, HLA-B*38:01, HLA-B*41:02, HLA-B*44:05, HLA-B*49:01, HLA-B*52:01, HLA-B*55:01, HLA-C*02:02, HLA-C*03:04, HLA-C*05:01,

HLA-C*07:01, HLA-C*01:02, HLA-C*04:01, HLA-C*06:02, HLA-C*07:02, HLA-C*16:01, HLA-C*03:03, HLA-C*07:04, HLA-C*08:01, HLA-C*08:02, HLA-C*12:02, HLA-C*12:03, HLA-C*14:02, HLA-C*15:02, and HLA-C*17:01 alleles. In some [embodimentexamples](#), the library comprises at least the sequences shown in SEQ ID Nos: 109-174.

[0031] In another [aspectexample](#), described herein is a polypeptide comprising, in an amino to carboxyl terminus orientation, (i) a first universal target peptide, (ii) an antigenic peptide, (iii) a second universal target peptide that is distinct from the first universal target peptide, (iv) a β 2M peptide, and (v) an MHC peptide. In some [embodimentexamples](#), the antigenic peptide is selected from the group consisting of a tumor antigen, a neoantigen, a tumor neoantigen, a viral antigen, bacterial antigen, phosphoantigen, and a microbial antigen. In one [embodimentexample](#), the antigenic peptide is a neoantigen.

[0032] In some [embodimentexamples](#), the polypeptide comprises in a 5' to 3' orientation, (i) a promoter peptide comprising the sequence shown in SEQ ID NO: 2, (ii) the first universal target peptide, (iii) the nucleotide sequence encoding an antigenic peptide, wherein the antigenic peptide is a tumor neoantigen, (iv) the second universal target peptide comprising the sequence shown in SEQ ID NO: 15 or 17, (v) the β 2M peptide comprising the sequence shown in SEQ ID NO: 105, (vi) the MHC peptide comprising a sequence shown in SEQ ID Nos: 38-103, (vii) a first affinity peptide comprising the sequence shown in SEQ ID NO: 30, (viii) a protease cleavage site peptide comprising the sequence shown in SEQ ID NO: 32, and (ix) a second affinity peptide comprising the sequence shown in SEQ ID NO: 36.

[0033] In some [embodimentexamples](#) the neoantigen is selected by analyzing tumor sequencing data from a subject to identify one or more somatic mutations. In one [embodimentexample](#), the analyzing is performed using an in silico predictive algorithm. In another [embodimentexample](#), the predictive algorithm further comprises an MHC binding algorithm to predict binding between the neoantigen and an MHC peptide.

[0034] In one [embodimentexample](#), the antigenic peptide is 7-15 amino acids, 7-10, 8-9, 7, 8, 9, 10, 11, 12, 13, 14, or 15 amino acids in length.

[0035] In some [embodimentexamples](#), the MHC peptide is a mammalian MHC peptide. In some [embodimentexamples](#), the MHC peptide is a human MHC peptide. In some [embodimentexamples](#), the MHC peptide is a class I HLA peptide. In other [embodimentexamples](#), the HLA peptide comprises an HLA-A, HLA-B, or HLA-C peptide. In one [embodimentexample](#), the HLA peptide is an HLA peptide of the subject. In one [embodimentexample](#), the HLA peptide comprises a Y84A or a Y84C mutation.

[0036] In some [embodimentexamples](#), the β 2M peptide is a mammalian β 2M peptide. In some [embodimentexamples](#), the β 2M peptide is a human β 2M peptide. In some [embodimentexamples](#), the β 2M

peptide comprises the sequence shown in SEQ ID NO: 105. In another [embodimentexample](#), the β 2M peptide comprises an S88C mutation. In some [embodimentexamples](#), the β 2M peptide comprises the sequence shown in SEQ ID NO: 107, comprising an S88C mutation at amino acid 88.

[0037] In some [embodimentexamples](#), the HLA peptide comprises HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*24:02, HLA-A*30:02, HLA-A*31:01, HLA-A*32:01, HLA-A*33:01, HLA-A*68:01, HLA-A*11:01, HLA-A*23:01, HLA-A*30:01, HLA-A*33:03, HLA-A*25:01, HLA-A*26:01, HLA-A*29:02, HLA-A*68:02, HLA-B*07:02, HLA-B*14:02, HLA-B*18:01, HLA-B*27:02, HLA-B*39:01, HLA-B*40:01, HLA-B*44:02, HLA-B*46:01, HLA-B*50:01, HLA-B*57:01, HLA-B*58:01, HLA-B*08:01, HLA-B*15:01, HLA-B*15:03, HLA-B*35:01, HLA-B*40:02, HLA-B*42:01, HLA-B*44:03, HLA-B*51:01, HLA-B*53:01, HLA-B*13:02, HLA-B*15:07, HLA-B*27:05, HLA-B*35:03, HLA-B*37:01, HLA-B*38:01, HLA-B*41:02, HLA-B*44:05, HLA-B*49:01, HLA-B*52:01, HLA-B*55:01, HLA-C*02:02, HLA-C*03:04, HLA-C*05:01, HLA-C*07:01, HLA-C*01:02, HLA-C*04:01, HLA-C*06:02, HLA-C*07:02, HLA-C*16:01, HLA-C*03:03, HLA-C*07:04, HLA-C*08:01, HLA-C*08:02, HLA-C*12:02, HLA-C*12:03, HLA-C*14:02, HLA-C*15:02, or HLA-C*17:01. In some [embodimentexamples](#), the HLA peptide comprises a sequence selected from the group consisting of SEQ ID NOs: 38-103.

[0038] In some [embodimentexamples](#), the first universal target peptide further comprises a signal peptide. In one [embodimentexample](#), the signal sequence is between 15-45, between 15-30, between 20-45, between 20-30, or between 30-45 amino acids in length. In one [embodimentexample](#), the signal peptide comprises Human Growth Hormone. In some [embodimentexamples](#), the signal peptide comprising a Human Growth Hormone signal peptide, a hIG1 Kappa light chain signal peptide, a Beta 2 microglobulin signal peptide, or an IL2 signal peptide. . In some [embodimentexamples](#), the signal peptide comprises a sequence comprising the sequence shown in SEQ ID NOs: 2, 23, 25, or 27. In some [embodimentexamples](#), the signal peptide comprises the Human Growth Hormone (HGH) signal peptide sequence shown in SEQ ID NO: 2. In one [embodimentexample](#), the second universal target peptide comprises the sequence GGGGSGGGGSGGGGS. In one [embodimentexample](#), the second universal target peptide comprising the sequence shown in SEQ ID NO: 15 or 17.

[0039] In some [embodimentexamples](#), the carboxyl terminus of the polypeptide may further comprise, in an amino to carboxylic terminus orientation, a purification cluster comprising (i) a first affinity peptide, (ii) a protease cleavage site, and (iii) a second affinity peptide. In some [embodimentexamples](#), the first and second affinity peptides are selected from the group consisting of AviTag, strep-tag, polyhistidine (His6)-tag, FLAG-tag, HA-tag, and/or Myc-tag. In some

[embodimentexamples](#), the first affinity peptide comprises a sequence comprising SEQ ID NO: 30 and the second affinity peptide comprises a sequence comprising SEQ ID NO: 34.

[0040] In some [embodimentexamples](#), the protease cleavage site is a TEV cleavage site, a thrombin cleavage site, a Factor Xa cleavage site, an enteropeptidase cleavage site, or a rhinovirus 3C protease cleavage site. In some [embodimentexamples](#), the purification cluster comprises an AviTag epitope, a TEV cleavage site, and a His6 epitope. In some [embodimentexamples](#), the protease cleavage site comprises the TEV cleavage site sequence shown in SEQ ID NO: 32. In some [embodimentexamples](#), the purification cluster comprises the AviTag peptide sequence shown in SEQ ID NO: 30, the TEV cleavage site sequence shown in SEQ ID NO: 32, and the His6 peptide sequence shown in SEQ ID NOs: 34 or 36. In some [embodimentexamples](#), the purification cluster comprises two or more copies of the His6 peptide sequence shown in SEQ ID NO: 34. In some [embodimentexamples](#), the purification cluster comprises the sequence shown in SEQ ID NO: 37.

[0041] In some [embodimentexamples](#), the second universal target peptide further comprises a linker comprising the sequence shown in SEQ ID NO: 9, 11, 15, or 17.

[0042] In some [embodimentexamples](#), the polypeptide comprises a second linker sequence between the β 2M sequence and the MHC allele sequence. In some [embodimentexamples](#), the second linker sequence comprises the sequence shown in SEQ ID NO: 9 or 19. In some [embodimentexamples](#), the polypeptide further comprising a third linker sequence between the MHC allele sequence and the first affinity tag. In some [embodimentexamples](#), the third linker sequence comprises the sequence shown in SEQ ID NO: 13 or 21. In some [embodimentexamples](#), the polypeptide is biotinylated.

[0043] In one [embodimentexample](#), the polypeptide comprises in an amino to carboxyl terminus orientation, (i) the first universal target peptide, (ii) the antigenic peptide, (iii) the second universal target peptide, (iv) the β 2M peptide, (v) the MHC peptide, (vi) the first affinity tag peptide, (vii) the protease cleavage site, and (viii) the second affinity tag peptide. In some [embodimentexamples](#), the polypeptide is biotinylated.

[0044] In another [aspectexample](#), disclosed herein is a library comprises greater than or equal to two distinct polypeptide molecules, wherein the distinct polypeptide molecule comprises (i) the first universal target peptide, (ii) the antigenic peptide, wherein the antigenic peptide is not the same for each of the greater than or equal to two polypeptide molecules, (iii) the second universal target peptide that is distinct from the first universal target peptide, (iv) the β 2M peptide, and (v) the MHC peptide. In one [embodimentexample](#), the MHC peptide is not the same for each of the greater than or equal to two polypeptide molecules. In another [embodimentexample](#), the

library comprises 20 to 500 distinct polypeptide molecules. In some [embodimentexamples](#), the library comprises at least 66 distinct polypeptides.

[0045] In some [embodimentexamples](#), the library comprises HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*24:02, HLA-A*30:02, HLA-A*31:01, HLA-A*32:01, HLA-A*33:01, HLA-A*68:01, HLA-A*11:01, HLA-A*23:01, HLA-A*30:01, HLA-A*33:03, HLA-A*25:01, HLA-A*26:01, HLA-A*29:02, HLA-A*68:02, HLA-B*07:02, HLA-B*14:02, HLA-B*18:01, HLA-B*27:02, HLA-B*39:01, HLA-B*40:01, HLA-B*44:02, HLA-B*46:01, HLA-B*50:01, HLA-B*57:01, HLA-B*58:01, HLA-B*08:01, HLA-B*15:01, HLA-B*15:03, HLA-B*35:01, HLA-B*40:02, HLA-B*42:01, HLA-B*44:03, HLA-B*51:01, HLA-B*53:01, HLA-B*13:02, HLA-B*15:07, HLA-B*27:05, HLA-B*35:03, HLA-B*37:01, HLA-B*38:01, HLA-B*41:02, HLA-B*44:05, HLA-B*49:01, HLA-B*52:01, HLA-B*55:01, HLA-C*02:02, HLA-C*03:04, HLA-C*05:01, HLA-C*07:01, HLA-C*01:02, HLA-C*04:01, HLA-C*06:02, HLA-C*07:02, HLA-C*16:01, HLA-C*03:03, HLA-C*07:04, HLA-C*08:01, HLA-C*08:02, HLA-C*12:02, HLA-C*12:03, HLA-C*14:02, HLA-C*15:02, and HLA-C*17:01 polypeptides.

[0046] In some [embodimentexamples](#), the polypeptide is attached to a particle, wherein the particle is a surface, a nanoparticle, a bead, or a polymer. In one [embodimentexample](#), the polypeptide is attached via a linker to the particle. In one [embodimentexample](#), the particle is a nanoparticle and the nanoparticle is a magnetic nanoparticle or a polystyrene nanoparticle. In one [embodimentexample](#), the magnetic nanoparticle comprises magnetic iron oxide. In another [embodimentexample](#), the bead is an agarose bead or a sepharose bead. In some [embodimentexamples](#), the polypeptide attached to the particle further comprises a fluorophore. In one [embodimentexample](#), the fluorophore is attached, with or without a linker, to the particle.

[0047] In some [embodimentexamples](#), a library of polypeptides is attached to particles wherein the library comprises greater than or equal to two distinct polypeptide molecules, wherein the antigenic peptide is not the same for each of the greater than or equal to two polypeptide molecules, and wherein each distinct polypeptide is attached to a particle. In one [embodimentexample](#), the MHC peptide is not the same for each of the greater than or equal to two polypeptide molecules. In one [embodimentexample](#), the library further comprises a unique defined barcode sequence operably associated with the identity of each distinct polypeptide. In another [embodimentexample](#), the library comprises 20 to 500 distinct polypeptides. In another [embodimentexample](#), the library comprises at least 66 distinct polypeptides.

[0048] In another aspect, described herein is a kit comprising the composition of any of the above [embodimentexamples](#) and instructions for use.

[0049] In another [aspectexample](#), disclosed herein is a method of manufacturing a polynucleotide molecule comprising the steps of: (a) obtaining a first polynucleotide sequence comprising, in a 5' to 3' orientation, (i) a first universal target sequence comprising a restriction site, (ii) a second universal target sequence comprising a restriction site that is not the same as the first universal target sequence, (iii) a β 2M sequence, and (iv) an MHC allele sequence; (b) obtaining a second polynucleotide comprising, in a 5' to 3' orientation (i) a portion of the first universal target sequence, (i) a sequence encoding an antigenic peptide, and (iii) a portion of the second universal target sequence; (c) obtaining a third polynucleotide comprising, in a 5' to 3' orientation, (i) the reverse complement of a portion of the second universal target sequence, (ii) the reverse complement of the sequence encoding an antigenic peptide, and (iii) the reverse complement of a portion of the first universal target sequence; (d) mixing the second and third polynucleotides together such that the complementary sequences anneal; (e) performing a restriction digest on the first polynucleotide sequence with at least one restriction enzyme; and (f) ligating the first and second polynucleotides together by mixing the digested first polynucleotide and annealed second and third polynucleotides together with DNA ligation reaction reagents.

[0050] In some [embodimentexamples](#), the method further comprises phosphorylating the 5' nucleotide of the second and third polynucleotides after the complementary sequences anneal, wherein the phosphorylating comprises incubating the first or second and third polynucleotides with an enzyme. In some [embodimentexamples](#), the enzyme is T4 kinase. In some [embodimentexamples](#), the first polynucleotide is an expression construct. In some [embodimentexamples](#), comprising inserting the manufactured polynucleotide into an expression construct.

[0051] In some [embodimentexamples](#), the first polynucleotide further comprises a promoter sequence. In one [embodimentexample](#), the promoter sequence is a CMV, an EF1 α , or an SV40 promoter.

[0052] In another [embodimentexample](#), the 3' end of the second polynucleotide sequence further comprises, in a 5' to 3' orientation, a purification cluster sequence comprising (i) a first affinity tag sequence, (ii) a protease cleavage site sequence, and (iii) a second affinity tag sequence.

[0053] In one [embodimentexample](#), the 3' end of the second polynucleotide sequence further comprises a polyA sequence.

[0054] In some [embodimentexamples](#), obtaining the third and fourth polynucleotide sequence steps further comprises obtaining a predictive data set of a tumor sequencing data from a subject. In one [embodimentexample](#), obtaining the third and fourth polynucleotide sequence steps further

comprises predicting the sequence encoding the antigenic peptide from the tumor sequencing data. In another [embodimentexample](#), obtaining the third and fourth polynucleotide sequence steps further comprises synthesizing the polynucleotide encoding the antigenic peptide based on the predicted sequence.

[0055] In some [embodimentexamples](#), the method further comprises manufacturing a library comprising greater than or equal to two distinct polynucleotide molecules, wherein the distinct polynucleotide molecule comprises (i) the first universal target sequence, (ii) the nucleotide sequence encoding an antigenic peptide, wherein the antigenic peptide is not the same for each of the greater than or equal to two polypeptide molecules, (iii) the second universal target sequence that is distinct from the first universal target sequence, (iv) the β 2M peptide, and (v) the MHC allele sequence. In one [embodimentexample](#), the MHC allele is not the same for each of the greater than or equal to two polynucleotide molecules. In another [embodimentexample](#), the library comprises 40 to 500 distinct polynucleotide molecules.

[0056] In some [embodimentexamples](#), the method further comprises inserting the manufactured polynucleotide into an expression construct. In one [embodimentexample](#), the method further comprises manufacturing a library comprising greater than or equal to two expression constructs comprising distinct polynucleotide molecules, wherein the distinct polynucleotide molecule comprises (i) the first universal target sequence, (ii) the nucleotide sequence encoding an antigenic peptide, wherein the nucleotide sequence is not the same for each of the greater than or equal to two polynucleotide molecules, (iii) the second universal target sequence that is distinct from the first universal target sequence, (iv) the β 2M sequence, and (v) the MHC allele sequence. In one [embodimentexample](#), the MHC allele is not the same for each of the greater than or equal to two polynucleotide molecules. In another [embodimentexample](#), the library comprises 20 to 500 distinct expression constructs. In some [embodimentexamples](#), the library comprises at least 66 distinct polynucleotide molecules. In some [embodimentexamples](#), the MHC allele is selected from the sequences shown in the group consisting of SEQ ID NOs: 109-174.

[0057] In some [embodimentexamples](#), the method further comprises expressing a polypeptide molecule from the polynucleotide molecule. In some [embodimentexamples](#), the polynucleotide is transfected or transduced into a cell. In one [embodimentexample](#), the polynucleotide is integrated into the cell genome. In another [embodimentexample](#), the polynucleotide remains extrachromosomal in the cell. In some [embodimentexamples](#), the host cell is a mammalian cell or a human cell. In one [embodimentexample](#), the cell is a stem cell, a tumor cell, an immortalized cell, or a fetal cell. In some [embodimentexamples](#), the cell is a prokaryotic cell. In some [embodimentexamples](#), the cell is an *Escherichia coli* cell.

[0058] In one [embodimentexample](#), the method further comprises quantifying the expression of the polyprotein. In some [embodimentexamples](#), the expressed and non-expressed polyprotein sequences are used to refine the antigenic peptide prediction analysis step of the method. In one

[embodimentexample](#), the method further comprises biotinylating the polyprotein.

[0059] In some [embodimentexamples](#), the method comprises purifying the polyprotein. The purifying step may comprise affinity chromatography. In one [embodimentexample](#), the affinity chromatography comprises immobilized metal affinity chromatography, comprising a support, a chelator and a divalent metal. The chelator for the affinity chromatography can be nitrotriacetic acid (NTA) or iminodiacetic acid (IDA). In some [embodimentexamples](#) the divalent metal is selected from the group consisting of: Nickel (Ni), Cobalt (Co), Copper (Cu), and Iron (Fe). In one [embodimentexample](#), the chelator is NTA and the divalent metal is Ni. In other [embodimentexamples](#), the support is selected from the group consisting of: agarose beads, sepharose beads, and magnetic beads. In some [embodimentexamples](#), the method further comprises quantifying the biotinylation level of the purified polyprotein.

[0060] In some [embodimentexamples](#) the method further comprises attaching the polypeptide to a particle, wherein the particle is a surface, a nanoparticle, a bead, or a polymer. In one [embodimentexample](#), the polypeptide is attached via a linker to the particle. In one [embodimentexample](#), the particle is a nanoparticle and the nanoparticle is a magnetic nanoparticle or a polystyrene nanoparticle. In one [embodimentexample](#), the magnetic nanoparticle comprises magnetic iron oxide. In another [embodimentexample](#), the bead is an agarose bead or a sepharose bead. In some [embodimentexamples](#), the polypeptide attached to the particle further comprises a fluorophore. In one [embodimentexample](#), the fluorophore is attached, with or without a linker, to the particle.

[0061] In another [embodimentexample](#), the method further comprises producing a library comprising greater than or equal to two distinct polyproteins attached to at least one particle, wherein the distinct polypeptide comprises (i) the first universal target peptide, (ii) the antigenic peptide, wherein the antigenic peptide is not the same for each of the greater than or equal to two polypeptide molecules, (iii) the second universal target peptide that is distinct from the first universal target peptide, (iv) the β 2M peptide, and (v) the MHC peptide. In one [embodimentexample](#), the MHC peptide is not the same for each of the greater than or equal to two polypeptide molecules. In another [embodimentexample](#), the library comprises 20 to 500 distinct single polyproteins. In another [embodimentexample](#), the library comprises at least 66 distinct single polyproteins. In another [embodimentexample](#), the MHC peptide is selected from the sequences shown in the group consisting of SEQ ID NO: 39-104.

[0062] In another [aspectexample](#), disclosed herein is a method of manufacturing a polynucleotide molecule comprising the steps of: (a) obtaining a first polynucleotide sequence comprising, in a 5' to 3' orientation, (i) a first universal target sequence comprising a restriction site, (ii) a second universal target sequence comprising a restriction site that is not the same as the first universal target sequence, (iii) a β 2M sequence, and (iv) an MHC allele sequence; (b) obtaining a second polynucleotide comprising, in a 5' to 3' orientation, (i) the first universal target sequence, (i) a sequence encoding an antigenic peptide, and (iii) the second universal target sequence; (c) performing a restriction digest on the first polynucleotide sequence with at least one restriction enzyme; (d) performing a restriction digest on the second polynucleotide sequence with at least one restriction enzyme; and (e) ligating the first and second polynucleotides together by mixing the digested first and second polynucleotides together with DNA ligation reaction reagents.

[0063] In another [aspectexample](#), described herein is a method of manufacturing a polynucleotide molecule comprising the steps of: (a) obtaining a first polynucleotide sequence comprising a first universal target sequence; (b) obtaining a second polynucleotide sequence comprising, in a 5' to 3' orientation, (i) a second universal target sequence, (ii) a β 2M sequence, and (iii) an MHC allele sequence; (c) obtaining a third polynucleotide sequence comprising, in a 5' to 3' orientation, a sequence encoding an antigenic peptide and the second universal target sequence; (d) obtaining a fourth polynucleotide sequence comprising, in a 5' to 3' orientation, the reverse complement of the antigenic peptide and the reverse complement of the first universal target sequence; (e) combining the first, second, third, and fourth polynucleotides in a solution; (f) adding polymerase chain reaction (PCR) reagents; and (g) performing a PCR reaction wherein a complementary region of the second universal target sequence in the second and third polynucleotides anneal to each other and a complementary region of the first universal target sequence in the first and fourth polynucleotides anneal to each other; and wherein the annealed sequences provide priming sequences for a PCR extension and amplification reaction.

[0064] Also disclosed herein is a method for isolating an antigen specific T cell, the method comprising the steps of: (a) providing a polypeptide comprising, in an amino terminus to carboxyl terminus orientation, (i) a first universal target peptide, (ii) an antigenic peptide, (iii) a second universal target peptide that is distinct from the first universal target peptide, (iv) a β 2M peptide, and (v) an MHC peptide, wherein the polypeptide is linked to one particle; (b) providing a sample known or suspected to comprise one or more T cells; (c) contacting the polypeptide with the sample, wherein the contacting comprises providing conditions

CLAIMS

1. A library comprising at least two different polynucleotide molecules, wherein each polynucleotide molecule ~~comprising~~ comprises, in a 5' to 3' orientation, a sequence insert comprising stop codons in each of the different reading frames, a Beta 2 Microglobulin (β 2M) sequence, and a ~~Multiple~~ Major Histocompatibility Complex (MHC) allele sequence.
2. The ~~polynucleotide~~ library of claim 1, wherein the sequence insert is flanked by a first universal target sequence and/or a second universal target sequence.
3. The ~~polynucleotide~~ library of claim 1 or 2, wherein each polynucleotide molecule further comprises ~~ing~~ a nucleotide sequence encoding an antigenic peptide positioned within the sequence insert.
4. A library comprising at least two different polynucleotide molecules, wherein each polynucleotide molecule ~~comprising~~ comprises, in a 5' to 3' orientation, a first universal target sequence, a nucleotide sequence encoding an antigenic peptide, a second universal target sequence that is distinct from the first universal target sequence, a Beta 2 Microglobulin (β 2M) sequence, and a ~~Multiple~~ Major Histocompatibility Complex (MHC) allele sequence.
5. The ~~polynucleotide~~ library of claim 3 or 4, wherein the antigenic peptide is a neoantigen.
6. The ~~polynucleotide~~ library of any one of claims 1-5, wherein the MHC allele is a human HLA allele.
7. The ~~polynucleotide~~ library of claim 6, wherein the HLA allele is selected from the group consisting of: HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*24:02, HLA-A*30:02, HLA-A*31:01, HLA-A*32:01, HLA-A*33:01, HLA-A*68:01, HLA-A*11:01, HLA-A*23:01, HLA-A*30:01, HLA-A*33:03, HLA-A*25:01, HLA-A*26:01, HLA-A*29:02, HLA-A*68:02, HLA-B*07:02, HLA-B*14:02, HLA-B*18:01, HLA-B*27:02, HLA-B*39:01, HLA-B*40:01, HLA-B*44:02, HLA-B*46:01, HLA-B*50:01, HLA-B*57:01, HLA-B*58:01, HLA-B*08:01, HLA-B*15:01, HLA-B*15:03, HLA-B*35:01, HLA-B*40:02, HLA-B*42:01, HLA-B*44:03, HLA-B*51:01, HLA-B*53:01, HLA-B*13:02, HLA-B*15:07, HLA-B*27:05, HLA-B*35:03, HLA-B*37:01, HLA-B*38:01, HLA-B*41:02, HLA-B*44:05, HLA-B*49:01, HLA-B*52:01, HLA-B*55:01, HLA-C*02:02, HLA-C*03:04, HLA-C*05:01, HLA-C*07:01, HLA-C*01:02, HLA-C*04:01, HLA-C*06:02, HLA-C*07:02, HLA-C*16:01, HLA-C*03:03, HLA-C*07:04,

HLA-C*08:01, HLA-C*08:02, HLA-C*12:02, HLA-C*12:03, HLA-C*14:02, HLA-C*15:02, and HLA-C*17:01.

8. The ~~polynucleotide~~-library of claim 6, wherein the HLA allele comprises a sequence selected from the group consisting of SEQ ID NOs: 109-174.

9. The ~~polynucleotide~~-library of any one of claims 1-8, wherein the β 2M allele is a human β 2M allele.

10. The ~~polynucleotide~~-library of any one of claims 2-9, wherein the first and second universal target sequences comprise a polymerase chain reaction (PCR) primer target site.

11. The ~~polynucleotide~~-library of any one of claims 2-10, wherein the first and second universal target sequences comprise a restriction enzyme cleavage site.

12. The ~~polynucleotide~~-library of any one of claims 1-10, wherein each polynucleotide molecules further ~~comprising~~-comprises a nucleotide sequence encoding a signal sequence.

13. The ~~polynucleotide~~-library of any one of claims 1-11, wherein the 3' end of ~~each~~the polynucleotide sequence further comprises a purification cluster sequence comprising a first affinity tag sequence, a protease cleavage site sequence, and a second affinity tag sequence.

14. The ~~polynucleotide~~-library of any one of claims 2-13, wherein the first universal target sequence comprises the sequence set forth in SEQ ID NO: 3 and the second universal target sequence comprises the sequence set forth in SEQ ID NO: 4.

~~15. An expression construct comprising the polynucleotide molecule of any one of claims 1-14.~~

~~16. A library comprising at least two polynucleotide molecules of any one of claims 1-14 or at least two expression constructs of claim 15.~~

~~17.~~15. The library of ~~claim 16~~any one of claims 1-14 comprising at least 66 ~~distinct~~ polynucleotide molecules ~~or 66 distinct expression constructs.~~

~~18. A polypeptide encoded by any one of claims 1-14.~~

~~19. The polypeptide of claim 18, wherein the MHC allele encodes for an HLA peptide comprising a Y84A or a Y84C mutation.~~

~~20. — A library comprising at least two polypeptides of claim 17 or 18.~~

~~21.~~ 16. A kit comprising ~~a polynucleotide of any one of claims 1-14, an expression construct of claim 15, a library of polynucleotides of claims 16 or 17, a polypeptide of claim 18 or 19, or a library of polypeptides of claim 20~~ any one of claims 1-15 and instructions for use.

~~22. — A method for isolating an antigen-specific T cell, the method comprising:~~

- ~~a. — preparing a plurality of particles, wherein each particle comprises a polypeptide of any one of claim 18 or 19;~~
- ~~b. — obtaining a sample known or suspected to comprise one or more T cells;~~
- ~~c. — contacting the plurality of particles with the sample, wherein the contacting comprises providing conditions sufficient for a single T cell to bind the polypeptide attached to the particle, and~~
- ~~d. — isolating the single T cell associated with the particle.~~

~~23. — A method for assaying a T cell interaction with an antigen, the method comprising:~~

- ~~a. — preparing a plurality of particles, wherein each particle comprises a polypeptide of any one of claims 18 or 19;~~
- ~~b. — obtaining a sample known or suspected to comprise one or more T cells;~~
- ~~c. — contacting the plurality of particles with the sample, wherein the contacting comprises providing conditions sufficient for a single T cell to bind the polypeptide attached to the particle, and~~
- ~~d. — treating the particles such that the labeled polypeptides are released and~~
 - ~~e. — detecting release of the labeled polypeptides from the T cells to assay the T cell interaction with the antigen sequence.~~