

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 distinct Major Histocompatibility Complex (MHC) allele sequence.

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

[0007] In one 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 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.

CLAIMS

1. 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 distinct Major Histocompatibility Complex (MHC) allele sequence.
2. The 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 library of claim 1 or 2, wherein each polynucleotide molecule further comprises a nucleotide sequence encoding an antigenic peptide positioned within the sequence insert.
4. The library of claim 3, wherein the antigenic peptide is a neoantigen.
5. The library of any one of claims 1-4, wherein the MHC allele is a human HLA allele.
6. The library of claim 5, 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.
7. The library of claim 5, wherein the HLA allele comprises a sequence selected from the group consisting of SEQ ID NOs: 109-174.
8. The library of any one of claims 1-7, wherein the β 2M allele is a human β 2M allele.

9. The library of any one of claims 2-8, wherein the first and second universal target sequences comprise a polymerase chain reaction (PCR) primer target site.
10. The library of any one of claims 2-9, wherein the first and second universal target sequences comprise a restriction enzyme cleavage site.
11. The library of any one of claims 1-9, wherein each polynucleotide molecule further comprises a nucleotide sequence encoding a signal sequence.
12. The library of any one of claims 1-10, wherein the 3' end of each 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.
13. The library of any one of claims 2-12, 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.
14. The library of any one of claims 1-13 comprising at least 66 polynucleotide molecules.
15. A kit comprising a library of any one of claims 1-14 and instructions for use.