**SUPPLEMENTAL METHODS**

**Insertion Site Search Validation**

Vector insertion events were simulated by randomly sampling positions from the human genome (hg38), excluding alt contigs and sex chromosomes, between 20 - 30 times per simulation library. Bisulfite sequencing reads (125bp, paired-end) were then simulated from the vector insertion contigs and mapped to a combined hg38 and MSCV transgenic TCR reference (BSBolt v0.1.2). Six simulation libraries, three MART-1 TCR and three NY-ESO-1 TCR, were created. The vector insertion detection pipeline was then validated against the known integration positions from the simulated reads. Correct identification of alignments that spanned the genome or vector sequence as a split (1 or more spanning bases) or discordant (complete read end mapped) was assessed for each library for various minimum alignment scores (40, 80, 120, 160, 200, 240, 280). With a minimum alignment score of 160, 29.2% (STD 1.41%) of MART-1 TCR and 36.6% (STD. 2.17%) NY-ESO-1 alignments were called correctly as vector spanning reads with correct mapping coordinates, and 0.183% (STD 0.0947%) and 0.301% (STD 0.0141%) alignments were called incorrectly. At minimum, a minimum alignment score of 160 at least 80 read bases must map for a valid alignment. Simulated split reads with fewer than this threshold are undetectable, resulting in a large proportion of simulated reads being unobserved. However, when a vector spanning is called the vast majority are called correctly. The complete simulation and validation pipeline can be found at <https://github.com/NuttyLogic/Epigenetic_Suppression_of_Transgenic_T-cell_Nowicki.2020/tree/master/VectorInsertionValidation>.