

An in-library ligation strategy and its application in CRISPR/Cas9 screening of high-order gRNA combinations

Zhike Lu^{1,2,3,†}, Ke Ni^{2,3,4,†}, Yingying Wang^{2,3,5,†}, Yangfan Zhou^{1,2,3}, Yini Li^{2,3,5}, Jianfeng Yan^{2,3,4}, Qingkai Song^{2,3,5}, Min Liu^{2,3,4}, Yujun Xu^{2,3,4}, Zhenxing Yu^{2,3,5}, Tiannan Guo^{2,3,4} and Lijia Ma^{2,3,4,*}

¹College of Life Sciences, Zhejiang University, Hangzhou 310058, China, ²Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, Zhejiang 310024, China, ³Institute of Biology, Westlake Institute for Advanced Study, Hangzhou, Zhejiang 310024, China, ⁴Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang 310024, China and ⁵School of Life Sciences, Fudan University, Shanghai 200433, China

Received April 06, 2022; Revised May 12, 2022; Editorial Decision May 12, 2022; Accepted May 13, 2022

ABSTRACT

Simultaneous targeting multiple genes is a big advantage of CRISPR (clustered regularly interspaced short palindromic repeats) genome editing but challenging to achieve in CRISPR screening. The crosstalk among genes or gene products is a common and fundamental mechanism to ensure cellular stability and functional diversity. However, the screening approach to map high-order gene combinations to the interesting phenotype is still lacking. Here, we developed a universal in-library ligation strategy and applied it to generate multiplexed CRISPR library, which could perturb four pre-designed targets in a cell. We conducted *in vivo* CRISPR screening for potential guide RNA (gRNA) combinations inducing anti-tumor immune responses. Simultaneously disturbing a combination of three checkpoints in CD8+ T cells was demonstrated to be more effective than disturbing *Pdcd1* only for T cell activation in the tumor environment. This study developed a novel in-library ligation strategy to facilitate the multiplexed CRISPR screening, which could extend our ability to explore the combinatorial outcomes from coordinated gene behaviors.

INTRODUCTION

CRISPR-mediated genetic screening is a powerful tool to dissect the functional units that contribute to specific phenotypes. The applications of disturbing single targeting gene have revealed important biological insights in various

studies (1). Meanwhile, simultaneously disturbing multiple targets have unique advantages in identifying targets for outcomes underlying high-order gene interactions. However, the exploration in this direction was still hindered by technical challenges.

The cooperative behavior of genes widely exists in cells. In many cases, disturbing a single gene is insufficient to direct the interested phenotype, even though that single gene dose contribute to the phenotype. For example, sets of transcription factors crosstalk with each other to orchestrate the invasion-metastasis cascade in cancer progression (2); Combinatorial use of inhibitors or blocking antibodies has been proved more effective in immunotherapy against many cancer types (3); more than one receptor is engaged in the cellular invasion of virus, therefore blocking single receptor is insufficient to protect host cells from infection (4).

Recently, some approaches have been established to enable combinatorial genetic screening via disturbing multiple targets. For example, gRNA pairs were randomly assembled to identify gene interactions affecting ovarian cell growth (5,6) and to explore drug pairs corresponding to kill K562 leukemia cells (7). To screen for pre-designed pairs, people could take advantage of high throughput oligo synthesis platforms, which could facilitate the direct synthesis of 2–3 gRNA expression cassettes as a single oligo (8–14). However, when the order of combination goes high, it is neither possible nor necessary to go through all combinations by random assembling. Moreover, for a high-order combination of gRNA, the gRNA expression cassettes must be assembled in a controlled manner in library to enable a screening for high-order pre-designed combinations. Theoretically, longer oligo synthesis is still possible, but the error rate and cost quickly increase along with the length, mak-

*To whom correspondence should be addressed. Tel: +86 571 86048622; Fax: +86 571 86048622; Email: malijia@westlake.edu.cn

†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

ing it not preferred and practical. Alternatively, the crRNA (CRISPR RNAs) for the Cpf1 editing has the advantage of shorter unit length and may fit in up to three units into a single oligo sequence with proper design (15,16). However, the limited options of its PAM sequences, especially around promoter regions with high GC content, restricted it from being a universal solution for high-order combinatorial genetic screening.

Here, we developed an in-library ligation approach to accurately ligate thousands of sequences to their specific counterparts. We demonstrated the in-library ligation strategies by preparing 4gRNA-combo (combination of four gRNAs) libraries, which simultaneously perturb four pre-designed targets in a single cell. Furthermore, one *in vitro* and one *in vivo* CRISPR screenings were conducted based on this type of libraries. The 4gRNA-combo libraries facilitated the discovery of high-order gene coordination that are challenging in high throughput. This study provided a universal strategy to express multiple pre-designed gRNAs from one vector. The strategy opened a new avenue for applying the CRISPR/Cas9 screening to facilitate the scalable and programmable perturbation on gene combinations.

MATERIALS AND METHODS

Multiplexed screening library construction

Overhang designing. The specificity of the in-library ligation highly depends on the complementary overhang sequences between pairs of oligos in library. In order to optimize the performance, we introduced multiple criteria to filter randomly generated 21-nt sequences, until reached the required number of sequences. Both sequence characteristics and duplex characteristics were considered, and the following criteria were applied: (i) the sequence GC content is between 45% to 60%, and the T_m is between 60°C and 65°C; (ii) the predict energy of the sequence secondary structure must be less than -3 kcal/mol predicted by the RNAfold program (17); (iii) the sequence must not present any recognition sites for restriction enzymes used in the following cloning steps; (iv) the sequence must possess more than five mismatches compared to any other sequence in the pool, and at least one mismatch locating within the four nucleotides at either 5' or 3' ends and (v) the predict energy of duplex structure with any sequence in the pool must be less than -15 kcal/mol predicted by the RNAfold RNA duplex program (17).

In-library ligation. The oligo pool was synthesized by GenScript (CustomArray). The immune gene library contained 12,472 oligos (6,236 pairs), 142-nt in length. The oligo pool was amplified as two sub-pools, and each sub-pool was amplified in twenty-four 50 ul PCR reactions. The PCR reactions were set as the following conditions: 25 ul NEBNext Ultra II Q5 Master Mix (NEB, M0544S), 2.5 ul forward primer (10 uM), 2.5 ul reverse primer (10 uM), 1 ul template (2.6 ng/ul oligo pool), and nuclease-free H₂O up to 50 ul. As shown in Figure 1A, primer F-BsrDI-1 and primer R-BsrDI-1-biotin were used in the L reaction; primer F-BsrDI-2-biotin and primer R-BsrDI-2 were used in the R reaction. The primer sequences were listed in the Supplementary Table S1. The PCR program was set as

the following condition to amplify the two sub-pools separately: (i) 98°C 30 s; (ii) 4 cycles of 98°C 10 s, 64°C 30 s, 72°C 30 s; (iii) 16 cycles of 98°C 10 s, 69°C 30 s, 72°C 30 s; (iv) 72°C 2 mins. The PCR products from each sub-pool were then combined and concentrated using Amicon 3K device (Millipore, UFC500324) according to the manufacturer's instruction. The concentrated products were size separated on 2.5% agarose gel. The gel slice with the targeted size was extracted using the QIAGEN Gel Extraction kit (QIAGEN, 28706), purified with phenol-chloroform, and further purified with the QIAquick Nucleotide Removal Kit (QIAGEN, 28306). The purified products from each sub-pool were then digested by nicking endonuclease Nb.BsrDI (NEB, R0648L) at 60°C for 4 h. The biotinylated fragments were cleaned up using the Dynabeads MyOne Streptavidin C1 (Thermo, 65001), and the longer fragments were retained in the sample and each exposed a 21-nt single-stranded overhang. During purification, the NaCl concentration of Binding and Washing buffer was adjusted to 0.25 M. And the biotinylated fragments were bound to the beads at 60°C with 800 RPM shaking for 1 h. The tube was moved to a 60°C water bath and beads were separated with a strong magnet underwater. After collecting the supernatant, which contained the digested product, 100 ul nuclease-free H₂O was added to resuspend the beads and incubated at 62°C for 30 mins, in order to recover the supernatant again. Next, all supernatants were combined and purified with the QIAquick Nucleotide Removal Kit (QIAGEN, 28306) according to the manufacture's instruction.

To perform in-library ligation, the digested products with the 21-nt overhangs were annealed with their pre-designed counterparts with the presence of HiFi Taq DNA ligase (NEB, M0647S). The ligation reaction was set as the following conditions: 350 ng sub-pool 1, 350 ng sub-pool 2, 5 ul 10x HiFi Taq DNA ligase buffer, 2 ul HiFi Taq DNA ligase, and nuclease-free H₂O up to 50 ul. The thermocycling conditions of ligation was set as the following: (i) 10 cycles of 70°C 30 s, 65°C 30 mins, 60°C 10 mins, 55°C 10 mins, 50°C 10 mins; (ii) 4°C hold. Following ligation, 5 ul T7E1 (NEB, M0302L) was added to the ligation product and incubate at 37°C for 30 mins. After digestion, 4 ul 0.5M EDTA (Invitrogen, 15575020) was added to inactivate the T7E1. After purification with the 1.2x AMPure XP beads (Beckman, A63882), the final ligation products were ready for the library cloning.

The in-library ligation pilot experiments were performed in the same procedure, except the length of oligo (141-nt), length of overhang sequence (20-nt), PCR primers (Primer PilotPrimer_L_Fwd-MO1 and PilotPrimer_L_Rev-MO1-biov1 were used in the L reaction; primer PilotPrimer_R_Fwd-MO2-biov1 and PilotPrimer_R_Rev-MO2 were used in the R reaction, Supplementary Table S1) and the nicking endonuclease (Nt.BspQI (NEB, R0644L)). And the library for the *in vivo* screening contained 1380 oligos (690 pairs), 146-nt in length.

Plasmid library construction. The in-library ligation products from last step were cloned into a modified lentiGuide Puro backbone (addgene, 52963) with mKate2. The molar ratio of backbone and insert is 1:3.5, and 145 fmol of inserts were used. The 50 uL Golden Gate Assembly (GGA)

reaction also included 0.5 ul of T4 DNA ligase (Thermo, EL0014), 5 ul 10× T4 DNA ligase buffer, 2.5 ul Esp3I (Thermo, ER0451), and nuclease-free H₂O up to 50 ul. The GGA condition was set as (i) 90 cycles of 37°C 5 mins and 22°C 5 mins; (ii) 65°C 30 mins; (iii) 37°C 3 h. Additional 2 ul of Esp3I was added into the reaction right before the 3 h 37°C incubation. One negative control reaction was performed following the same condition except without adding the inserts (GGA1 plasmid library). The GGA reaction products were purified with 0.7× AMPure XP beads (Beckman, A63882) and then dialysis on the MF-Millipore™ Membrane Filter (Sigma, VSWP02500) for 2 h. For each transformation reaction, 2 ul GGA products were electroporated (Eppendorf 2510, 1700 V) with 25 ul electrocompetent cells (Lucigen, 60242-2). One reaction was performed for the sample and one reaction was performed for the negative control. The tube with transformation mixture was recovered for 1 h at 37°C, then spread on two 25 cm × 25 cm LB-ampicillin plate and incubated for 20 h at 30°C. After propagation, colonies were scraped from the plates. Plasmids were extracted using QIAGEN Plasmid Plus Midi Kit (QIAGEN, 12945) according to the manufacturer's instructions. The plasmids (GGA1 library) were purified with phenol–chloroform and ethanol precipitation for the next GGA reaction. NGS library was generated using this step of products to evaluate the library quality.

The second GGA was performed under the same condition except another TypeIIS restriction enzyme BsaI-HF v2 (NEB, R3733) was used, and five 50 ul reactions and one negative control reaction were performed. The molar ratio of the GGA1 library and the 'human Gln-tRNA vector' (The vector and map will be available in addgene) is 1:4.9, and 31 fmol of GGA1 library was used. The transformation, propagation and plasmid library extraction were performed the same way as the preparation for the GGA1 library. The product was called the GGA2 plasmid library.

The third GGA was performed using 1 ul AarI (Thermo, ER1582) per reaction, and five 50 ul reactions and one negative control reaction were performed. The molar ratio of the GGA2 library and the 'human Gly-tRNA vector' (The vector and map will be available in addgene) is 1:3, and 35 fmol of the GGA2 library was used. The transformation, propagation and plasmid library extraction were performed the same way as the preparation for the GGA1 library. The product was called the GGA3 plasmid library.

The fourth GGA was performed with 30 fmol of the GGA3 product, and the molar ratio of the GGA3 product and the 'human Pro-tRNA vector' (The vector and map will be available in addgene) is 1:3.5. In total of five 50 ul reactions and one negative control reaction were performed. Each reaction also included 0.5 ul T4 DNA Ligase (Thermo, EL0014), 5 ul 10× Cutsmart Buffer (NEB, B7204), 2 ul BbsI-HF (NEB, R3539L) for the in vitro screening library or 2 ul SapI (NEB, R0569S) for the in vivo screening library, 0.5 ul ATP (Thermo, R0441), 0.5 ul DTT (Invitrogen, Y00147), and nuclease-free H₂O up to 50 ul. The transformation, propagation and plasmid library extraction were performed the same way as the preparation for the GGA1 library. In order to remove distorted constructs due to recombination, the plasmids were size-selected by electrophoresis (2% agarose, 80 V and 50 mins). The plas-

mids recovered from gel extraction was transformed into the Stable cells (NEB, C3040H). The propagated *Escherichia coli* was harvested, and plasmids were extracted. The product was called the GGA4 plasmid library. NGS libraries were generated using the products from this step to evaluate the library quality.

Across all GGA and transformation steps, in order to maintain the representative of the diverse plasmids in the library, the propagation was performed on two 25 cm × 25 cm plates. Additionally, regular petri dish was used to spread *E. coli* culture in order to estimate the number of colonies, which represented the complexity of the plasmid library. The typical complexity of our libraries is 8 800 000, which represented 1400× coverage to the 6236 4gRNA-combos.

Lentivirus production. The vector used in screening or validation experiment was transfected into 293T cells. Viral supernatant was collected at 48 and 72 h post transfection, passed through 0.45 μm filter, and concentrated via ultracentrifuging at 70 000g for 2 h. The concentrated supernatant was subsequently aliquoted and stored at –80°C.

***In vitro* Jurkat activation screening and validation**

Screening library design. A candidate gRNA pool was firstly generated. All unique 20 bp gRNA spacer with NGG PAM were picked from the human genome (hg38), and the off-target score were calculated according to the mismatch matrix (18). Among them, the spacers targeting 5' of the CDS regions with off-target score ≤0.05 were combined with the Brunello library (18), which in total made a spacer pool for further filtering. To be compatible with the multiple cloning steps in this study, spacer sequences with BsmBI, AarI, BbsI and BsaI recognition sites were removed. Also, spacer sequences with constitutive thymidines (≥4) were also excluded to avoid unexpected transcription termination. These post-filtering spacer sequences were grouped according to the targeting genes and ranked according to the library source and off-target score. The spacers designed by the Brunello come first. The top10 spacers were selected to represent the gRNA pool for each human coding gene. If <10 gRNAs available, all were included.

To find candidate immune response related genes, we used gene annotations from the Gene Ontology (GO) database. In total of 1,599 genes (Supplementary Table S2) that are also presented in the known pathways (Kyoto Encyclopedia of Genes and Genomes, KEGG) or gene families (European Genome-phenome Archive, EGA) were included as targeting genes of the multiplexed CRISPR screening library. In order to determine the combination of gRNAs, we considered the functional relevance of their targeting genes. In brief, for KEGG pathways possessing four or more targeting genes, combinations were generated within pathway by randomly picking four genes (3,492 combinations). For pathways with less than four targeting genes, all genes were picked to make combination with genes from other pathways (180 combinations). For EGA families with more than four or more targeting genes, combinations were generated within families by randomly picking four genes (945 combinations). During the process of

generating combinations, gRNAs of each targeting gene were randomly picked from the gRNA pool as mentioned above. The combination search stopped if the gene coverage, which was denoted as the frequency of one gene exhibited across all combinations, reached 15. In order to balance the coverage across genes, we generated 1569 combinations by picking genes according to their coverage across the established combinations in descending order. Additional 50 negative controls were also included. Finally, in total of 6236 4gRNA-combs were generated (Supplementary Table S2), and each of the 1599 candidate genes were covered by 15 combinations in average (min = 13) in the designed library. The free energy (ΔG) of each oligo (including the spacer sequences of four gRNAs and other accessory sequences) must be less than -48 kcal/mol (17).

Cell culture. The Cas9 coding gene was inserted into the Jurkat cell using lentivirus of lentiCas9-Blast (Addgene, 52962). According to the blasticidin killing curve assay, 2 $\mu\text{g}/\text{ml}$ blasticidin was used for Jurkat-Cas9 selection. After blasticidin selection, the survival cells were collected for single cell sorting performed via BD FACS Fusion flow cytometer (BD Bioscience). Then, monoclones of the Cas9-expressing cells were established at the presence of 2 $\mu\text{g}/\text{ml}$ blasticidin. The Cas9 expression of each monoclonal was confirmed by western blot (Cell Signaling, Mouse anti-Cas9, 7A9-3A3).

Viral library production and transduction. The transfer plasmid of 4gRNA-combo library, the pMD2.G (Addgene, 12259) envelope plasmid and the psPAX2 (Addgene, 12260) packaging plasmid were mixed at the mass ratio 5:2:3 and incubated with 250 μM calcium chloride. Equal volume of 2 \times HeBS (280 mM NaCl, 1.5 mM Na_2HPO_4 , 50 mM HEPES (pH 7.05)) was added into the DNA- CaCl_2 and incubated 15 mins at room temperature. The mixture was dropped to HEK293T cells at 80% confluency. Lentiviral supernatant was collected at 48 and 72 h post transfection, filtered through a 0.45 μm filter (Millipore, SLHV033RB), and then concentrated by ultracentrifuging at 70 000g at 4°C for 2 h. A total of 20×10^6 Jurkat-Cas9 cells were infected by the concentrated viral library at $\text{MOI} \leq 0.3$ in RPMI-1640 containing 8 $\mu\text{g}/\text{ml}$ polybrene. A ‘spinection’ was conducted by centrifuging the culture plate at 700 g, 32°C for 2 h. At 48 h post transduction, cellular mKate2 expression, indicating the successful transduction, was verified by flow cytometry (Cytoflex, Beckman). Typical rate of mKate2 positive cell is $\sim 30\%$. In the next 6–10 days, cells grew under 2 $\mu\text{g}/\text{ml}$ puromycin and 2 $\mu\text{g}/\text{ml}$ blasticidin antibiotic selection, and the cell concentration was maintained at $5 \times 10^5/\text{ml}$. During the course of antibiotic selection, the cellular mKate2 expression was monitored by flow cytometry until $>95\%$ of cells were mKate2 positive.

Activation experiment. A total of 6×10^6 successfully infected Jurkat-Cas9 cells were collected as the starting reference (sample ‘Control’ in Figure 2A). Another 30×10^6 Jurkat cells were stimulated by 25 $\mu\text{l}/\text{ml}$ ImmunoCult™ Human CD3/CD28 T Cell Activator (STEMCELL, 10971) in RPMI with 10% FBS and 1 \times Pen-Strep at $5 \times 10^6/\text{ml}$ cell density. After 24 h of stimulation, cells were stained by anti-

CD69 (Biolegend, FN50), CD69+ (top 25%) and CD69– (bottom 25%) cell population were sorted and collected using FACS (Fusion, BD). In total of 5×10^6 cells from each population were collected (sample ‘CD69+’ and ‘CD69–’ in Figure 2B).

Candidate validation. In order to verify the inhibitive effects of gRNA combination, validation experiments were conducted following the same procedure as the large-scale library transduction and activation. Except the starting number of the Jurkat-Cas9 cell was 5×10^5 per each virus transduction experiment. 24 h after stimulation, the percentage of CD69+ cells were examined using flow cytometry (Cytoflex, Beckman). All the flow data were analyzed by Flowjo v10.

NGS library preparation. For the first GGA plasmid library, three 50 μl PCR reactions were performed. In each PCR reaction, we used 18.23 ng plasmid as template, 0.25 μM forward primer, 0.25 μM reverse primer, 25 μl NEBNext Ultra II Q5 Master Mix (NEB, M0544S), and nuclease-free H_2O up to 50 μl . The PCR program was set as: (i) 98°C 30 s; (ii) 6 cycles of 98°C 10 s, 64°C 30 s, 72°C 20 s; (iii) 9 cycles of 98°C 10 s, 72°C 45 s and (iv) 72°C 2 mins. The resulting PCR product was purified with AMPure XP beads (Beckman, A63882). For the purification, 0.4 \times AMPure XP beads were used to remove the large fragment firstly and 0.7 \times AMPure XP beads were used to bind the PCR product. The primers are Fwd-libseq-lib1 and Rev-libseq-TCATCTCC and primer sequences can be found in Supplementary Table S1.

For the fourth GGA plasmid pool (which was used to make large-scale viral library). One NGS library was generated to cover amplicons including the first and the second gRNAs (G12 library), and another NGS library was generated to cover amplicons of the second and the third gRNAs (G23 library) (Supplementary Table S3). For each of the NGS libraries, three 50 μl PCR reactions were performed. In each PCR reaction, we used 4.6 ng plasmid as template, 0.25 μM forward primer, 0.25 μM reverse primer, 25 μl NEBNext Ultra II Q5 Master Mix (NEB, M0544S), and nuclease-free H_2O up to 50 μl . The PCR program was set as: (i) 98°C 30 s; (ii) 16 cycles of 98°C 10 s, 63°C 30 s, 72°C 20 s; (iii) 72°C 2 mins. The resulting PCR product was size-selected from a 1.5% agarose gel via MinElute Gel Extraction Kit (Qiagen, 28606) and purified using 1.2 \times AMPure XP beads (Beckman, A63882). The primers used to amplify the first and second gRNAs are Fwd-libseq-U6 and Rev-libseq-Gly, and the primers for the second and the third gRNAs are Fwd-libseq-Gln and Rev-libseq-Pro. The primer sequences can be found in Supplementary Table S1.

For the ‘Control’, ‘CD69+’ and ‘CD69–’, two NGS libraries were made for each of these samples. For each sample, the genomic DNA (gDNA) harvested from 3M cells were used as template to provide $\sim 500\times$ coverage to the total number of four 4gRNA-combos. The gDNA was extracted from cells using the QIAamp DNA Blood Mini kit (QIAGEN, 51106). One NGS library generated amplicons covering the first and the second gRNAs (G12 library), and another NGS library generated amplicons covering the second and the third gRNAs (G23 library) (Supplementary Ta-

ble S3). For each of the NGS library, multiple 50 µl PCR reactions were performed to use up the gDNAs extracted from 1.5M cells. In each PCR reaction, we used ~550 ng gDNA as template, 0.25 µM forward primer, 0.25 µM reverse primer, 25 µl NEBNext Ultra II Q5 Master Mix (NEB, M0544S), and nuclease-free H₂O up to 50 µl. The PCR program was set as: (i) 98°C 30 s; (ii) 6 cycles of 98°C 10 s, 63°C 30 s, 72°C 20 s; (iii) 20–22 cycles of 98°C 10 s, 72°C 50 s (according to the qPCR) and (iv) 72°C 2 mins. The resulting PCR product was size-selected from a 1.5% agarose gel via MinElute Gel Extraction Kit (Qiagen, 28606) and purified using AMPure XP beads (Beckman, A63881). The primers used to amplify the first and second gRNAs are Fwd-libseq-U6 and Rev-libseq-Gly, and the primers for the second and the third gRNAs are Fwd-libseq-Gln and Rev-libseq-Pro. The primer sequences can be found in Supplementary Table S1.

Data processing. All libraries were sequenced as 150-bp paired-end. The sequencing reads were firstly undergone adapter removal by ‘cutadapt’ (19). The parameter to process the first GGA product NGS library is ‘-n 1 -e 0.11 -O 15 -m 16’, and the parameter to process the other libraries are ‘-n 1 -e 0.11 -O 24 -m 16’. Alignment was conducted using bowtie2 with ‘-np 0 -n-ceil L,0,0.2 -very-sensitive’ (20). Customized references were used according to the sources of the NGS libraries. The successfully aligned reads were assigned to the designed 4gRNA-combos, and the read counts were used to calculate the combination representatives in libraries and the cumulative distributions.

In order to find the 4gRNA-combos targeting genes that are essential to the T cell activation, the normalized read counts of each combination were used to compare their representatives between the CD69+ and CD69- cell populations. Normalizations were conducted according to the depth of sequencing libraries. The G12 library and G23 library of each sample were treated as technical replicates. For each 4-gene combination, the log₂ fold-change of the normalized read counts were calculated between the CD69+ and CD69- samples, and the mean of replicates was used. The *P*-values were calculated from DESeq2 under a negative binomial distribution (21). Combinations that meet *P*-value < 0.01 and log₂FC < -5, or *P*-value < 10⁻¹⁰ and -5 < log₂FC < -2 were labelled in Figure 2C. To pick combinations for further validation, all 4-gene combinations were sorted in descending order according to their log₂FC. Top candidates were validated using individually generated constructs.

***In vivo* immune checkpoint screening**

Screening library design. The design of gRNA pool followed the same principle as the library of *in vitro* Jurkat activation screening. We included a group of six immune checkpoint genes (CP group), a group of four genes involved in the first signaling of T cell activation (TCR group) and a group of five genes of co-stimulatory molecules involved in the secondary signaling (CS group) (Supplementary Table S4). Within each group, all possible 4gRNA-combo, 3gRNA-combo, 2gRNA-combo and single gRNA construct were designed. The unoccupied position was

placed with non-targeting control gRNA. Further, all possible combinations were represented by six groups of gRNAs that are distinct from each other (Supplementary Table S5). Additionally, 84 combos composed only the non-targeting control gRNAs (NT group) were also included and served as negative control (Supplementary Table S5). This screening library composed a total of 101 gene combinations represented by 606 gRNA groups and 84 negative control combinations including on non-targeting gRNAs.

The CRISPR screening and validation. The screening library was constructed following the same steps as the Jurkat activation screening library. The backbone plasmid composed a mKate2 reporter, which were used to isolate the engineered T cells that infiltrated into the tumors.

The tumor cells. The Hepa1-6 cells was transduced with H-2K^b-OVA₂₅₇₋₂₆₄-expressing lentivirus. And monoclones were validated with the H-2K^b-OVA₂₅₇₋₂₆₄ expression via flow cytometry. The resulted cell line was named as Hepa1-6-H-2K^b-OVA₂₅₇₋₂₆₄. The established Hepa1-6-H-2K^b-OVA₂₅₇₋₂₆₄ cells were further transduced with a lentiviral vector (lenti-EF-1α-luciferase-T2A-BSD) for luciferase stable expression.

The animal experiments. The primary T cells were isolated from OT-I or Cas9+OT-I mice, which were bred from OT-I and Cas9 mouse ordered from the Jackson Laboratory. The tumor was inoculated to the NOD-*Prkdc*^{scid} *Il2rg*^{null}/Shjh mice ordered from Shanghai Jihui Laboratory Animal Care. The T cell donor mice were 10–12 weeks old. The tumor recipient mice were 6–8 weeks old. All mice were housed in standard individually ventilated and pathogen-free conditions in the laboratory facility of the Westlake University, under that animal protocol (AP#21-016-MLJ). All mice were used in accordance with Institutional Animal Care and Use Committee (IACUC) guidelines for Westlake University.

T cell isolation and culture. Splens were isolated from Cas9+OT-I mice, followed by mashing through 40 µm filter and RBCs lysis (BD Pharm Lyse). CD8+ T cells were purified by negative selection via CD8a+ T cell isolation Kit (Miltenyl). Cells were stimulated with 100 U/ml recombinant human IL-2 (Peprotech), 1 µg/ml anti-mouse CD3ε (Ultraleaf, Clone 145-2C11, Biolegend) and 0.5 µg/ml anti-mouse CD28 (Ultraleaf, Clone 37.51, Biolegend) and cultured in RPMI-1640 with 10% FBS, 10 mM HEPES (Gibco), 100 µM non-essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), 50 µM β-mercaptoethanol (Sigma), 50 U/ml penicillin and 50 µg/ml streptomycin (Gibco).

T cell transduction, transduction efficiency test and gene editing efficiency test. After *ex vivo* stimulation for 24 h, CD8+ T cells were transduced with lentivirus in the presence of polybrene at 8 µg/ml during spinfection at 2000g for 2 h at 32°C. At 48h after transduction, T cells were collected for transduction efficiency test via flow cytometry and adoptive transfer. In validation experiment, CD8+ T cells were transduced with lentivirus for 2 times at 24 and

48 h after isolation. At 24 h after second transduction, T cells were collected for transduction efficiency test via flow cytometry and adoptive transfer after sorting via FACS. The gene editing efficiency was tested in T cells with *Pdcd1-Adora2a-Ctla4* combined disruption. At 48 h after second transduction, mKate2+ T cells were sorted via FACS and pelleted for gDNA extraction. Then, the gRNA target sequences of each gene were amplified by two-step PCR for NGS sequencing. The list of oligos used in gene editing efficiency test were included in Supplementary Table S1.

Antigen specificity test for OT-I T cells. OT-I CD8+ T cells were co-cultured with either tumor cells or tumor expressing H-2K^b-OVA₂₅₇₋₂₆₄ cells for 2 and 48 h. In the 2 h test, cells were co-cultured at the presence of anti-CD107a (Biolegend, 1D4B). After 2h, all cells were collected and stained with anti-CD8a (Biolegend) for degranulation analysis via flow cytometry (Cytoflex, Beckman). After 48 h, all cells were collected and stained with anti-CD8a, PI and Annexin V (Biolegend) for target cell apoptosis analysis via flow cytometry (Cytoflex, Beckman). All FCM data were analyzed by Flowjo 10.

Screening experimental workflow. Hepa1-6 cells expressing H-2K^b-OVA₂₅₇₋₂₆₄ were mixed with matrigel (1:1 volume) and injected subcutaneously into the right flank of NPSG mice at 1×10^6 /recipient. At d12 after tumor cell inoculation, 1×10^7 CD8+ T cells with screening library transduction (5–10% mKate2+ cells in total cells) were adoptively transferred into each recipient via *i.v.* injection. Meanwhile, $2-3 \times 10^6$ CD8+ T cells with screening library transduction were frozen as starting reference (SR). Weight loss and tumor size was measured at d0 and d7 after T cell injection. At d7 after injection, tumor was collected and cut into small fragments. After consecutively mashing through 100 μ m and 40 μ m filters, RBCs in the cell suspension were lysed. Then, the tumor infiltrating CD8+ T cells were enriched by density gradient centrifugation via Lymphprep (StemCell). Cells at the interface were carefully collected and washed by PBS. Then, the cells were resuspended into PBS and stained with anti-mouse CD8a for 30 min on ice. Finally, CD8+ mKate2+ tumor infiltrating lymphocytes (TILs) were sorted via FACS (Fusion, BD). A total of 20 000–40 000 CD8+ mKate2+ TIL could be collected per tumor. TIL from three to four recipient mice were mixed together and pelleted with carrier cells (WT Raji cell) at 1:50 (CD8+ T cells: carrier cells) for genomic DNA extraction.

Genomic DNA extraction and NGS library preparation. Genomic DNA extraction was performed using TIANamp Genomic DNA kit (TIANGEN) and finally resuspended in 50 μ l nuclease free water. To prepare the gRNA NGS library for the SR sample, all gDNA were amplified on thermocycling with parameters of 98°C for 30 s, 20–22 cycles of (98°C for 10 s, 64°C for 30 s, 72°C for 20 s), 72°C for 2 mins. To prepare the gRNA NGS library for the TIL sample, two-step amplification was applied. In the first step, PCR reaction (400–800 ng DNA input per reaction, 2–4 reactions per sample) was performed using Ultra II Q5 Master Mix (NEB) with thermocycling parameters as 98°C for 30 s, 28–

30 cycles of (98°C for 10 s, 60°C for 30 s, 72°C for 20 s), 72°C for 2 mins. And the PCR condition and primers of the second step follows the condition of the SR library preparation, but with 8–10 cycles. All primers used in NGS library preparation were listed in Supplementary Table S1. Both G12 and G23 NGS libraries were prepared for the TIL sample as stated in the GGA4 NGS library preparation methods part.

Validation of candidates. The multiplexed CRISPR knockout vector that contained a *Pdcd1-Adora2a-Ctla4* gRNA tandem cassette and mKate2 reporter was generated (mKate2+). Meanwhile, vector contained a *Pdcd1-NTC-NTC* gRNA tandem cassette or *Pdcd1-Ctla4-NTC* gRNA tandem cassette as well as tagBFP reporter were created as control (tagBFP+). Hepa1-6 cells expressing H-2K^b-OVA₂₅₇₋₂₆₄ with luciferase were mixed with matrigel (1:1 volume) and injected subcutaneously into the right flank of NPSG mice at 1×10^6 /recipient. At d11–d12 after tumor cell inoculation, 1×10^6 mKate2+ or tagBFP+ Cas9+OT-I CD8+ T cells were sorted via FACS and adoptively transferred into each recipient via intravenous injection. Weight loss and tumor size was measured every 3 days after T cell injection. Meanwhile, tumor size was monitored weekly by *in vivo* imaging via PHOTON IMAGERTM OPTIMA, in which luciferin was administered intraperitoneally 5 mins prior to signal collection.

Data analysis. In order to find the effective 4gRNA-combo that enhance the capacity of the CD8+ T cell-mediated tumor elimination *in vivo*, the normalized read counts of each combination were used to compare their representatives between the TIL and SR samples. The normalizations were conducted according to the depth of sequencing libraries. We calculated both the fold-change and the *P*-value for each 4gRNA-comb. TIL and SR were treated as two samples, and G12 library and G23 library of each sample were treated as technical replicates. We used the \log_2 fold-change of G12 and G23 between the TIL and SR library to pick combinations for validations, which could be explained as $\log_2((\text{mean of TIL three batches G12} + 1)/(\text{mean of SR three batches G12} + 1))$ and $\log_2((\text{mean of TIL three batches G23} + 1)/(\text{mean of SR three batches G23} + 1))$.

RESULTS

Pair-specific in-library ligation mediated by long overhang

To conduct multiplexed CRISPR screening, we firstly established a framework to enable the controllable ligations only between pre-designed pairs in a pooled sequence library, which we named pair-specific in-library ligation. Inspired by the typical ligation reaction between digestion products of restriction enzymes, we decided to ligate sequences based on the specific and efficient anneals between complementary overhangs. To facilitate the pair-specific fragments assembling in library scale, long and sufficient amount of overhang sequences are required. We considered that fragments could be ligated to their counterparts in a library if long overhangs were explicitly designed to assemble each pair of fragments. The customized long over-

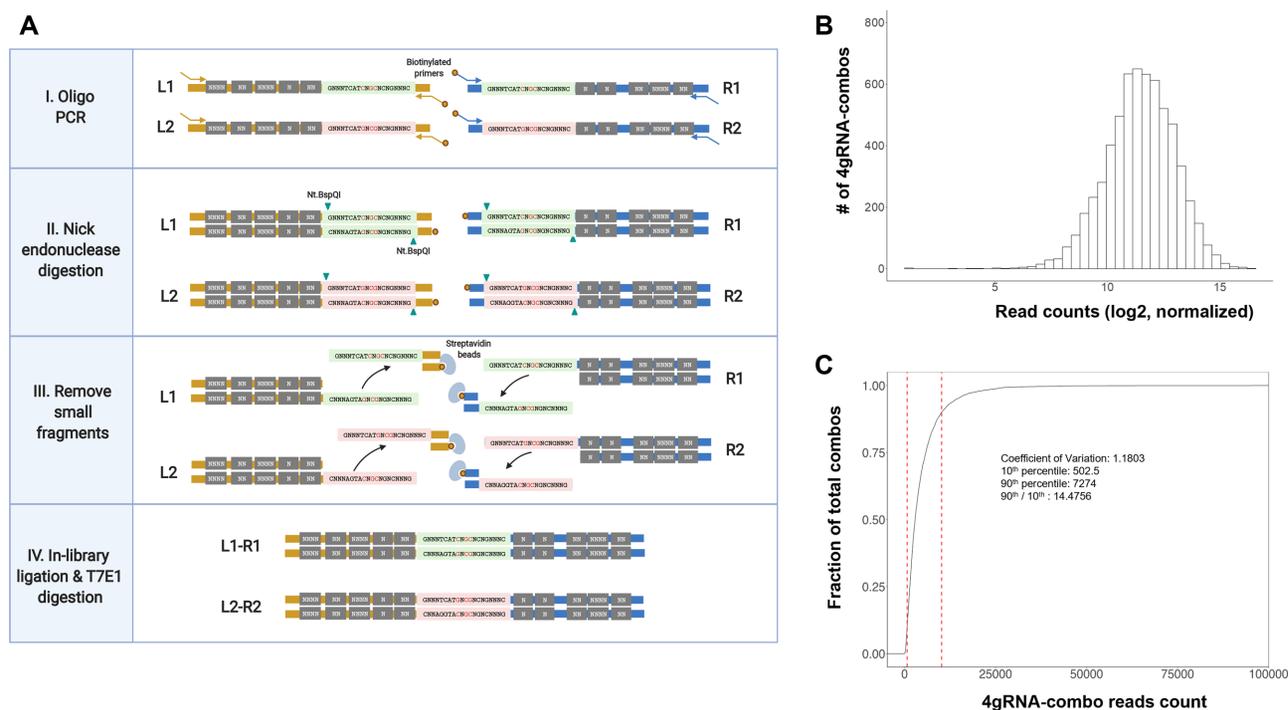


Figure 1. Pair-specific in-library ligation and multiplexed CRISPR library construction. (A) Schema of the pair-specific in-library ligation. The L1, L2, R1 and R2 are four synthesized oligos with both fixed and degenerated nucleotides. The oligos were pooled as a library and amplified in two separate PCR reactions (Step I). The PCR products from each reaction were enzyme digested (Step II) and cleaned up (Step III). The products from the Step III were pooled again, and the in-library ligation and T7E1 digestion were conducted (Step IV). At the end, only products from L1 and R1 and products from L2 and R2 were successfully ligated. Yellow and blue lines: oligos with different fixed sequences. Dark grey rectangle: degenerated nucleotides in oligos. Light green and pink rectangles: sequences served as complementary regions in ligation (both the fixed and degenerated nucleotides were designed in complementary regions). (B) Library quality of the plasmid library. The histogram shows the read counts of the gRNA combinations. (C) Library quality of the plasmid library. The cumulative distribution of read counts of the gRNA combinations.

hang could provide great sequence diversity to differentiate the counterpart from the non-counterpart in large-scale library design and increase the ligation specificity between sequences.

With this rationale, we designed a pilot experiment to create long overhangs for paired sequences and then annealed those sequences specifically. Each paired sequence encoded two gRNAs from a 4gRNA combination (Figure 1A). We designed and synthesized four 141-nt oligos (two pairs) named L1, R1, L2 and R2. These oligos were mosaicked with degenerated nucleotides (Ns), which mimicked the characteristics of diverse sequences in a pooled library, while the fixed nucleotides helped to determine whether the ligation products assembled correct counterparts (see Methods and Supplementary Table S1). As illustrated (Figure 1A), the oligos were firstly amplified by two pairs of biotinylated primers in two separate reactions (Figure 1A, step I). The double-stranded amplicons all carried biotins and served as the substrate of the nick endonuclease in the next step. The nick endonuclease could generate two staggered nicks, one on the top strand and the other on the bottom strand (Figure 1A, step II). In the next step, the biotinylated fragments were removed by the streptavidin beads, and the 20-nt long overhangs were exposed (Figure 1A, step III). The four cleavage products, which carried different overhangs, were then pooled to conduct the in-library ligation (Figure 1A, step IV). The fragments were ligated in the li-

brary mediated by their pair-specific long overhang with the presence of HiFi Taq DNA Ligase. The HiFi Taq DNA Ligase ensures the ligation could be conducted at a relatively high temperature (60°C), thus eliminating annealing between unpaired overhangs. The resulted ligation products were further treated with T7E1 to remove error products. In our design, the overhang of the L1 products was only fully complementary with the R1 products, while L2's was only complementary with R2's.

Finally, the products were examined by Sanger sequencing. Among the sixteen colonies we checked, eight were correct ligation products of L1 and R1, and the other eight were originated from L2 and R2 ligation (two of them were incomplete sequences) (Supplementary Figure S1). Importantly, no swapped products were found in the sixteen colonies. Noticeably, in the 20-nt overhang region, only three nucleotides were designed differently between the L1–R1 pair and L2–R2 pair. This demonstrated the high specificity of the in-library ligation when the pooled sequences are distinguishable only at a few nucleotides.

Together, by sophisticatedly designed pair-specific long overhang, we precisely assembled four gRNAs encoded in two different oligos, which suggested the possibility of conducting high-order combinatorial CRISPR screening using a commonly available oligo pool.

Overhang design and multiplexed CRISPR library construction

The pilot study demonstrated that long overhangs could enable pair-specific ligations. For a multiplexed gRNAs library with thousands of combinations, we need thousands of pair-specific overhang sequences to facilitate the precise assembling of gRNA combinations. To do so, we generated random sequences and filtered them by matching multiple criteria that might influence the ligation efficiency and downstream cloning (Supplementary Figure S2). The overhang sequences for the in-library ligation must meet all the following criteria: (i) proper GC content and T_m ; (ii) no strong secondary structure; (iii) without recognition sites for restriction enzymes used in the following cloning steps; (iv) possessing more than five mismatches compared to any other sequence in the pool, and at least one mismatch locating within the four nucleotides at either 5' or 3' ends and (v) no strong duplex structure with any other sequence in the pool (see Materials and Methods). Adhere to these criteria, we designed 6236 combination-specific overhang sequences to enable the in-library ligation of 12 472 oligos in a pool.

Similar to the pilot experiment, the gRNA sequences of each 4gRNA-combo were synthesized from a pair of oligos, which were then amplified and ligated in a library. The ligation products were cloned into a lentiviral backbone, followed by three sequential Golden Gate Assembly to insert the scaffolds and tRNAs downstream of the spacers 1, 2 and 3 (Materials and Methods, Supplementary Figure S3). We applied the polycistronic tRNA-gRNA (PTG) method to express four gRNAs simultaneously (22). Compared to other strategies for expressing multiple gRNAs, the PTG system utilizes the conserved endogenous RNA-processing system and requires no additional endonuclease expression (Supplementary Figure S4). We chose three different human tRNAs from previous publications to eliminate the potential recombination (23).

We examined the library quality in multiple steps throughout library preparation, including the plasmid library after the first golden gate (GGA1) and the final plasmid library after the fourth golden gate (GGA4) (Supplementary Table S3). High throughput sequencing data confirmed that the plasmid library faithfully represented all designed 4gRNA-combos under a uniform distribution (Supplementary Figure S5A). With 20 million sequencing reads, 98.6% of the first golden gate assembly products were mapped to the references (Supplementary Table S3), and all pre-designed combinations were well covered ($CV = 1.09$) (Supplementary Figure S5B). These data indicated that the in-library ligation efficiently and accurately assembled all pre-designed gRNAs from two sub-pools. The final library also exhibited excellent performance, with 91.4–94.7% mappable reads and coefficient variance 1.18 (Figure 1B and C). Collectively, these data demonstrated a high-quality multiplexed gRNA library that was achieved by massively parallel in-library ligation and sequential golden gate assemblies.

In vitro screening for T cell activation

The 6236 combinations targeted 1599 genes, and each combination contains four gRNAs that target four different genes (see Materials and Methods, Supplementary Table

S2). We reasoned that genes involved in the same pathway or the same gene family might exhibit more functional relevance and lead to genetic compensation when only one of them is functionally disrupted. Therefore, we hypothesized that disturbing multiple genes in the same pathway (24) or gene family (25) might help identify new candidates that share coordinated behavior than disturbing one single target (Methods, Supplementary Figure S6 and S7, Supplementary Table S2).

Next, we applied a screening of canonical T cell activation to demonstrate the performance of the 4gRNA-combo CRISPR library (Methods, Figures 2A, Supplementary Figure S8–S10, Supplementary Table S2). We calculated the ratio of normalized read counts between the CD69+ and CD69– T cells for each combination and used that to evaluate the perturbation of CRISPR KO to the T cell receptor (TCR) signal transduction. We firstly looked up the ratios of combinations from the TCR signaling pathway, the salivary secretion pathway and the pre-designed non-targeting controls (Supplementary Table S6). As expected, combinations in the TCR signaling pathway largely shifted towards the low CD69+/CD69– ratio side, which means cells edited by these combinations enriched in the CD69– sample (Figure 2B). Meanwhile, combinations involved in the ‘Salivary secretion pathway’, which has no known interaction with the TCR signaling transmitting, exhibited a tight and uniform distribution largely overlapped with the negative control group (Figure 2B). We ranked all combinations by the CD69+/CD69– ratio to identify top candidates (Methods, Figure 2C). Among the top twelve combinations enriched in the CD69– cell population, three of them were expected to be essential to the T cell activation signal transduction, as they were either relevant to the T cell activation pathway or contained subunits of the TCR complex. To validate the candidates from the screening, we repeated the T cell activation experiment with T cells engineered by each of the individual combinations (Supplementary Table S7). Most of the combinations showed a reduction in terms of the percentage of activated T cells compared to the control, and six of them show statistical significance (see Methods, Supplementary Figure S11). Overall, these data demonstrated that the multiplexed CRISPR perturbation is an effective strategy for identifying functional and combinatorial gene sets responsible for phenotypic outcomes.

In vivo screening for combinatorial checkpoint blockades to boost T cells

The combined immunotherapy promotes anti-tumor immunity by targeting multiple immune repressors that work in complementary and nonredundant mechanisms (26). To identify potential candidates for combined immunotherapy, we applied the 4gRNA-combo multiplexed library in an *in vivo* screening for the boosted tumor-infiltrating T cells (TILs). Following the multiplexed CRISPR library construction strategy, we genetically engineered CD8+ T cells collected from OT-I mice, which were further injected into recipient mice inoculated with Hepa1-6 cells with stable H-2K^b-OVA₂₅₇₋₂₆₄ expression (Figure 3A and Supplementary Figure S12). To investigate the cooperative anti-tumor efficacy, we engineered T cells to target one to four gRNAs si-

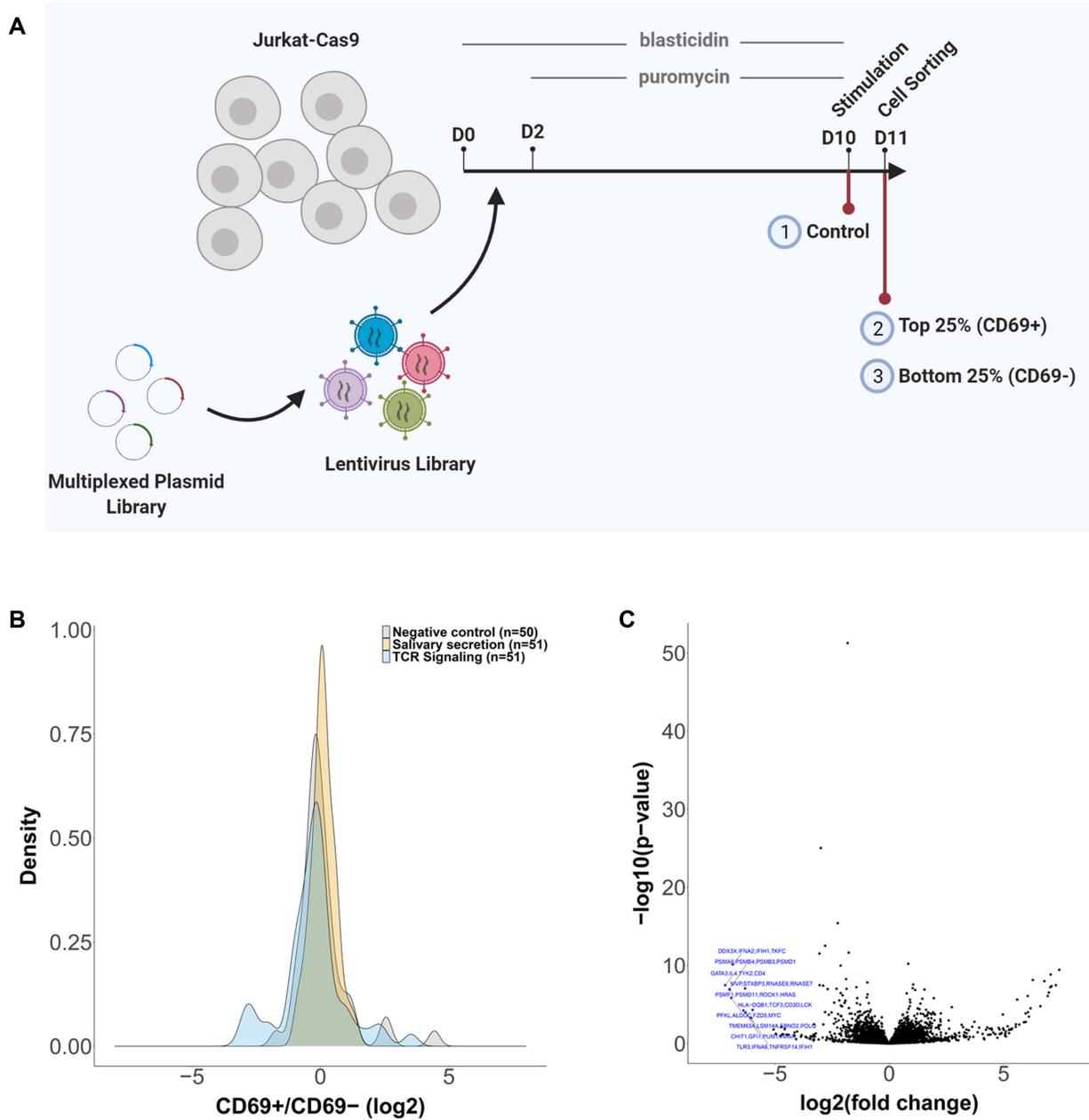


Figure 2. *In vitro* screening for T cell activation by the 4gRNA-comb library. (A) Schema of the Jurkat activation screening. The Jurkat cells stably expressing Cas9 were infected with the multiplexed CRISPR library containing 6236 4gRNA-combos targeting 1599 genes. The cells were undergone antibiotic selections for ten days. Then a canonical T cell activation experiment was conducted to screen for the engineered Jurkat cells expressing more CD69. (B) Distribution of combinations between the CD69+ and CD69- cell populations. The cell distribution between the CD69+ and the CD69- populations were plotted. For each combination, the ratio of normalized read counts within these two populations were calculated and log2 transformed. Three subgroups of combination were plotted and compared. As expected, the distribution of 51 combinations raised from the TCR signaling pathways (kegg.hsa04660) was shifted towards the CD69- cell population. And the 51 combinations raised from the salivary secretion (kegg.hsa04970) were distributed around the centered, and largely overlapped with the 50 combinations from the negative control subgroup. (C) Volcano plot of the Jurkat screening results. The screening outcome of each combination was quantified and illustrated in a volcano plot. The combinations labeled in blue meet cutoff of P -value < 0.01 and $\log_2FC < -5$. The combinations labeled in black meet cutoff of P -value $< 10^{-10}$ and $-5 < \log_2FC < -2$.

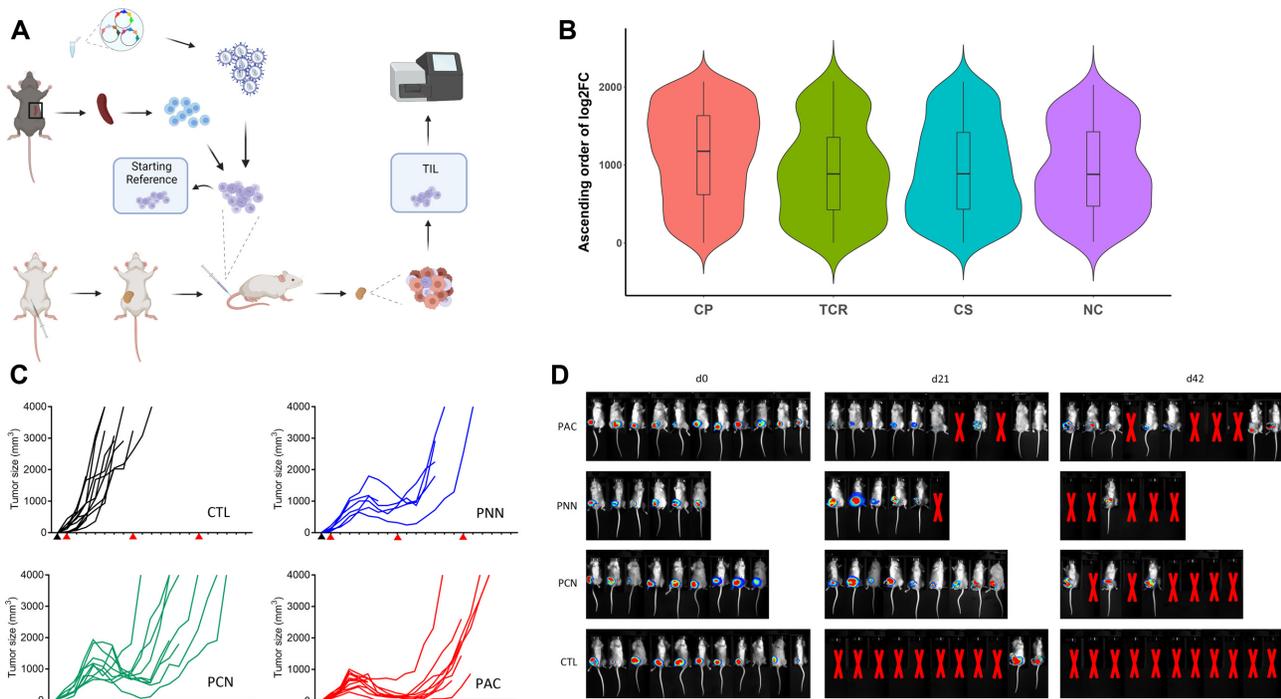


Figure 3. *In vivo* screening for combinatorial checkpoint blockades to boost T cells. (A) Schema of *in vivo* screening. CD8⁺ T cells collected from OT-I mice, which were infected by screening library and further injected into recipient mice inoculated with Hepa1-6 cells with stable H-2K^b-OVA_{257–264} expression. (B) Ranks of the log₂FC of the engineered T cells across different groups. More T cells from the CP group were enriched into the tumor samples compared to T cells from the groups of TCR, CS and NC. CP: checkpoint group; TCR: T cell receptor group; CS: co-stimulatory molecule group; NC: negative control group. (C) Tumor size curve for mice receiving OT-I CD8⁺ T cells with combined *Adora2a*, *Ctla4* and *Pdcd1* disruption (PAC), combined *Ctla4* and *Pdcd1* disruption (PCN) or only *Pdcd1* disruption (PNN), and the mice that did not receive CD8⁺ T injection (CTL). The black triangle indicated the day of tumor cell line inoculation; the red triangles indicated d0, d21 and d42 after T cell injection. Tumor sizes were recorded every 3 days. The number of mice in each group was: 11 in PAC, 6 in PNN, 9 in PCN and 12 in CTL. (D) *In vivo* imaging for mice receiving OT-I CD8⁺ T cells with PAC, PCN or PNN disruption, and the mice that did not receive CD8⁺ T injection (CTL). The red crosses indicated the dead mice or mice sacrificed because of tumor size limitation (≤ 4000 mm³).

multaneously. The T cells were then screened for activation capability in the tumor environment. At the endpoint of the screening, the engineered T cells successfully presented in the tumors were isolated by FACS and subjected to NGS characterization.

This *in vivo* screening library included six well-characterized checkpoint genes and saturated all fifty-six possible combinations, composed of fifteen 4gRNA-combos, twenty 3gRNA-combos, fifteen 2gRNA-combos and six single-gRNA combos (CP group) (Supplementary Table S4). For each combo, we used non-targeting control gRNAs to fill the unoccupied positions if the targeting gRNA is less than four. For comparison, we also included combos targeting two other groups of genes; one included four genes involved in the first signaling of T cell activation (fifteen combos, TCR group), the other included five co-stimulatory molecules involved in the secondary signaling of T cell (thirty combos, CS group) (Supplementary Table S4). T cells engineered by combos from the TCR group and CS group should be incapable of T cell activation. Together, we included 101 distinct combinations targeting one to four genes. For each distinct combination, we designed a group of six gRNA-combos in the library to eliminate the biases of individual guide RNA (Supplementary Tables S4 and S5). Another eighty-four combos composed only non-targeting control gRNAs, which served as negative

control (NT group) (Supplementary Table S4 and S5). The screening was conducted in three independent batches.

We calculated a log₂ transformed fold-change (log₂FC) to indicate the relative abundance of each combo in the TIL relative to the engineered T cells before being injected into the recipient mice (SR, represented ‘starting reference’) (see Materials and Methods). We hypothesized that the T cells enriched in tumors gain functions relevant to anti-tumor immunity, which were reflected by the gRNA combos with high log₂FC values. As expected, most T cells did not successfully enrich in tumors and show negative log₂FC values (Supplementary Figure S13). Among the four groups (CP, TCR, CS and NC), more T cells from the CP group show positive log₂FC values (Figure 3B). This result aligned with the expected function of the engineered T cells and suggested the effectiveness of the screening model.

With these metrics, we ranked all combos according to their enrichment from three screening batches and identified a top candidate of 3gRNA-combo that simultaneously targeted *Pdcd1*, *Adora2a* and *Ctla4* (noted as PAC hereafter) (Supplementary Figure S14). Among all combinations, the PAC combo exhibited the best reproducibility across three independent batches of screening and different groups of gRNAs. We also examined the other 4gRNA-combos included gRNAs targeting these three genes and found that only this specific combination could maximumly

activate the infiltrated T cells in tumors (Supplementary Figure S15). This result implied the importance of identifying the precise combination of targets, as the anti-tumor ability of T cells might not be positively strengthened by knocking out more checkpoint genes.

Next, we performed validation experiments to confirm the screening results. We included T cells knocked out only at the *Pdcd1* loci (PNN), at *Pdcd1* and *Ctla4* loci (PCN), and at *Pdcd1*, *Ctla4* as well as *Adora2a* loci (PAC). The knockout efficiencies of each gRNA were confirmed before *in vivo* validation (Supplementary Figure S16). We also shuffled gRNAs when making combos to eliminate biases from individual gRNA. The engineered T cells were injected intravenously into the recipient mice inoculated with Hepa1–6 cancer cells expressing H-2K^b-OVA_{257–264}. After the T cell therapy, the weight loss of the mice and the tumor size were monitored for eight weeks. We found that the growth of tumor size of the PNN, PCN and PAC groups was controlled at different levels. Among them, T cells engineered by the PAC combination showed the best anti-tumor immune responses compared to T cells engineered by PCN or PNN, which were reflected by the tumor size and the survival rate of the mice (Figure 3C and D & Supplementary Figure S17). The combinatorial therapy targeting these three checkpoints at protein level has been previously recognized, in which an antagonist of A2AR was used with antibodies for PD-(L)1 and CTLA-4 for a better tumor elimination in animal model (27). These results suggested that the multiplexed CRISPR screening was an effective way to look for candidates for potential combinatorial immune checkpoint blockades.

DISCUSSION

A high throughput method to systematically characterize the function of high-order gene combinations is of great interest to study complicated biological systems. The CRISPR screening has provided convenient and unprecedented opportunities to dissect gene function in mammals. However, limited by the oligo length of pool synthesis in regular practice, performing combinatorial genetic screening remains challenging and costly.

This study reported a massively parallel in-library ligation approach, which precisely assembled thousands of sequences with their counterparts in a library. Through combination-specific overhangs, this method provided a universal solution for controlled sequences assembling in a pooled library. Although four gRNAs were multiplexed in this report, more could fit in if more sequences were assembled the same way. In cases where the PAM sequence is not a concern, more crRNA units could also be multiplexed using Cpf1-mediated screening.

By multiplexing gRNAs, the complexity of the screening library is also decreased, which brought another advantage of the in-library ligation strategy. In many cases, especially when performing genome editing in primary cells, the screening scale is limited by the number of cells that could decently cover the library complexity. By assembling more gRNA expression cassettes into a single vector, a difficult screening could be performed by first identifying the candidate combination from a limited number of cells. Fur-

thermore, the candidate combinations could be further dissected to pinpoint the causal subset or identify the cooperative effects among subsets.

We demonstrated the multiplexed CRISPR screening strategy in two demonstrations. One started with a large number of candidate genes, and the library was constructed with 6236 pre-designed 4gRNA-combos to target 1599 genes. Although pre-knowledge is needed to design this kind of library, specific hypotheses could be tested with this strategy when complete randomization is unnecessary or impractical. In another *in vivo* screening, we started with a short candidate gene list but saturated all possible gRNA combinations targeting the candidate genes. This screening helped us to go through the entire candidate space unbiasedly.

Finally, from the immune checkpoint combination screening, the PAC combo was identified and verified in independent validation. The candidate combo has also been reported in another study, in which researchers successfully induced anti-tumor responses in a mouse model that used the blocking antibodies for PD-(L)1 and CTLA-4 and antagonist for A2AR (27). In the clinic, combined immune therapy has been considered as a promising direction to overcome resistance in cancer immunotherapy (28,29). And the multiplexed CRISPR screening strategy provided a systematic approach to investigate potential combinations in high-order for immunotherapy.

Together, this novel in-library ligation approach addressed the challenge of multiplexing pre-designed gRNAs for CRISPR/Cas9 screening and opened a new avenue for discovering more sophisticated and complicated cellular processes in high throughput.

DATA AVAILABILITY

GEO accession GSE149945 and GSE202653 (<https://www.ncbi.nlm.nih.gov/geo/>). FlowRepository (<https://flowrepository.org/>): FR-FCM-Z56P, FR-FCM-Z56T, FR-FCM-Z56Q, FR-FCM-Z56S, FR-FCM-Z56U, FR-FCM-Z56E.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the Flow Cytometry Facility and Genomics Facility of the Westlake Biomedical Research Core Facilities and Laboratory of Animal Resource Center of Westlake University for their assistances in conducting this work. We thank Dr Dangsheng Li and Dr Hongtao Yu for valuable discussions and suggestions. The flowchart illustrations were made from Biorender.

FUNDING

National Science Foundation of China [31950575 to L.M.]; Basic Research Foundation of Zhejiang Province for Distinguished Young Scholar [LR21C060001 to L.M.]; Basic Research Foundation of Zhejiang Province [LQ19C050003 to

M.L.]; Westlake Education Foundation and Tencent Foundation (to L.M.). Funding for open access charge: National Science Foundation of China [31950575 to L.M.]; Basic Research Foundation of Zhejiang Province for Distinguished Young Scholar [LR21C060001 to L.M.]; Basic Research Foundation of Zhejiang Province [LQ19C050003 to M.L.]; Startup Funding to L.M. from the Westlake Education Foundation.

Conflict of interest statement. Zhike Lu, Dr Ke Ni and Dr Lijia Ma are co-founders of AIdit, a biotech company focusing on AI-assisted CRISPR Therapy. Dr Ke Ni conducted this work when she worked as a postdoc in Dr Lijia Ma's lab at Westlake University and is now a full-time employee of AIdit during the submission of this work.

REFERENCES

- Pickar-Oliver, A. and Gersbach, C.A. (2019) The next generation of CRISPR-Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.*, **20**, 490–507.
- Hanahan, D. and Weinberg, R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646–674.
- Seung, E., Xing, Z., Wu, L., Rao, E., Cortez-Retamozo, V., Ospina, B., Chen, L., Beil, C., Song, Z., Zhang, B. *et al.* (2022) A trisppecific antibody targeting HER2 and T cells inhibits breast cancer growth via CD4 cells. *Nature*, **603**, 328–334.
- Jackson, C.B., Farzan, M., Chen, B. and Choe, H. (2022) Mechanisms of SARS-CoV-2 entry into cells. *Nat. Rev. Mol. Cell Biol.*, **23**, 3–20.
- Wong, A.S., Choi, G.C., Cheng, A.A., Purcell, O. and Lu, T.K. (2015) Massively parallel high-order combinatorial genetics in human cells. *Nat. Biotechnol.*, **33**, 952–961.
- Wong, A.S., Choi, G.C., Cui, C.H., Pregernig, G., Milani, P., Adam, M., Perli, S.D., Kazer, S.W., Gaillard, A., Hermann, M. *et al.* (2016) Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 2544–2549.
- Han, K., Jeng, E.E., Hess, G.T., Morgens, D.W., Li, A. and Bassik, M.C. (2017) Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. *Nat. Biotechnol.*, **35**, 463–474.
- Zhu, S., Li, W., Liu, J., Chen, C.H., Liao, Q., Xu, P., Xu, H., Xiao, T., Cao, Z., Peng, J. *et al.* (2016) Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. *Nat. Biotechnol.*, **34**, 1279–1286.
- Diao, Y., Fang, R., Li, B., Meng, Z., Yu, J., Qiu, Y., Lin, K.C., Huang, H., Liu, T., Marina, R.J. *et al.* (2017) A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. *Nat. Methods*, **14**, 629–635.
- Shen, J.P., Zhao, D., Sasik, R., Luebeck, J., Birmingham, A., Bojorquez-Gomez, A., Licon, K., Klepper, K., Pekin, D., Beckett, A.N. *et al.* (2017) Combinatorial CRISPR-Cas9 screens for de novo mapping of genetic interactions. *Nat. Methods*, **14**, 573–576.
- Boettcher, M., Tian, R., Blau, J.A., Markegard, E., Wagner, R.T., Wu, D., Mo, X., Biton, A., Zaitlen, N., Fu, H. *et al.* (2018) Dual gene activation and knockout screen reveals directional dependencies in genetic networks. *Nat. Biotechnol.*, **36**, 170–178.
- Liu, Y., Cao, Z., Wang, Y., Guo, Y., Xu, P., Yuan, P., Liu, Z., He, Y. and Wei, W. (2018) Genome-wide screening for functional long noncoding RNAs in human cells by Cas9 targeting of splice sites. *Nat. Biotechnol.*, **36**, 1203–1210.
- Najm, F.J., Strand, C., Donovan, K.F., Hegde, M., Sanson, K.R., Vaimberg, E.W., Sullender, M.E., Hartenian, E., Kalani, Z., Fusi, N. *et al.* (2018) Orthologous CRISPR-Cas9 enzymes for combinatorial genetic screens. *Nat. Biotechnol.*, **36**, 179–189.
- Chow, R.D., Wang, G., Ye, L., Codina, A., Kim, H.R., Shen, L., Dong, M.B., Errami, Y. and Chen, S.M. (2019) In vivo profiling of metastatic double knockouts through CRISPR-Cpf1 screens. *Nat. Methods*, **16**, 405–408.
- Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E.M., Winblad, N., Choudhury, S.R., Abudayyeh, O.O., Gootenberg, J.S. *et al.* (2017) Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.*, **35**, 31–34.
- Gier, R.A., Budinich, K.A., Evitt, N.H., Cao, Z., Freilich, E.S., Chen, Q., Qi, J., Lan, Y., Kohli, R.M. and Shi, J. (2020) High-performance CRISPR-Cas12a genome editing for combinatorial genetic screening. *Nat. Commun.*, **11**, 3455.
- Gruber, A.R., Lorenz, R., Bernhart, S.H., Neubock, R. and Hofacker, I.L. (2008) The Vienna RNA websuite. *Nucleic Acids Res.*, **36**, W70–W74.
- Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R. *et al.* (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.*, **34**, 184–191.
- Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.*, **17**, 10–12.
- Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
- Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
- Xie, K., Minkenberg, B. and Yang, Y. (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 3570–3575.
- Knapp, D., Michaels, Y.S., Jamilly, M., Ferry, Q.R.V., Barbosa, H., Milne, T.A. and Fulga, T.A. (2019) Decoupling tRNA promoter and processing activities enables specific Pol-II Cas9 guide RNA expression. *Nat. Commun.*, **10**, 1490.
- Kanehisa, M. and Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*, **28**, 27–30.
- Lappalainen, I., Almeida-King, J., Kumanduri, V., Senf, A., Spalding, J.D., Ur-Rehman, S., Saunders, G., Kandasamy, J., Caccamo, M., Leinonen, R. *et al.* (2015) The European Genome-Phenome Archive of human data consented for biomedical research. *Nat. Genet.*, **47**, 692–695.
- Postow, M.A., Chesney, J., Pavlick, A.C., Robert, C., Grossmann, K., McDermott, D., Linette, G.P., Meyer, N., Giguere, J.K., Agarwala, S.S. *et al.* (2015) Nivolumab and ipilimumab versus ipilimumab in untreated melanoma. *N. Engl. J. Med.*, **372**, 2006–2017.
- Willingham, S.B., Ho, P.Y., Hotson, A., Hill, C., Piccione, E.C., Hsieh, J., Liu, L., Buggy, J.J., McCaffery, I. and Miller, R.A. (2018) A2AR antagonism with CPI-444 induces antitumor responses and augments efficacy to Anti-PD-(L)1 and anti-ctla-4 in preclinical models. *Cancer Immunol. Res.*, **6**, 1136–1149.
- Sharma, P., Hu-Lieskovan, S., Wargo, J.A. and Ribas, A. (2017) Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell*, **168**, 707–723.
- Khair, D.O., Bax, H.J., Mele, S., Crescioli, S., Pellizzari, G., Khiabany, A., Nakamura, M., Harris, R.J., French, E., Hoffmann, R.M. *et al.* (2019) Combining immune checkpoint inhibitors: established and emerging targets and strategies to improve outcomes in melanoma. *Front. Immunol.*, **10**, 453.