Journal for ImmunoTherapy of Cancer

Circular RNA as a source of neoantigens for cancer vaccines

Yi Ren ,[1,2] Thamizhanban Manoharan,[1,2] Beijia Liu,[1] Cyrus Zai Ming Cheng,[1,2] Bei En Siew,[3,4] Wai-Kit Cheong,[4] Kai Yin Lee,[4] Ian Jse-Wei Tan,[4] Bettina Lieske,[4] Ker-Kan Tan,[3,4] Gloryn Chia [1,2]

To cite: Ren Y, Manoharan T, Liu B, *et al*. Circular RNA as a source of neoantigens for cancer vaccines. *Journal for ImmunoTherapy of Cancer* 2024;**12**:e008402. doi:10.1136/jitc-2023-008402

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/jitc-2023-008402).

YR and TM contributed equally.

Accepted 03 March 2024

[1]Department of Pharmacy, National University of Singapore, Singapore
[2]NUS iHealthtech, Singapore
[3]Department of Surgery, National University of Singapore, Singapore
[4]Department of Surgery, National University Hospital, Singapore

Correspondence to
Dr Gloryn Chia;
phagcl@nus.edu.sg

ABSTRACT

Background The effectiveness of somatic neoantigen-based immunotherapy is often hindered by the limited number of mutations in tumors with low to moderate mutation burden. Focusing on microsatellite-stable colorectal cancer (CRC), this study investigates the potential of tumor-associated circular RNAs (circRNAs) as an alternative source of neoepitopes in CRC.

Methods Tumor-associated circRNAs in CRC were identified using the MiOncoCirc database and ribo-depletion RNA sequencing of paired clinical normal and tumor samples. Candidate circRNA expression was validated by quantitative real-time PCR (RT-qPCR) using divergent primers. TransCirc database was used for translation prediction. Human leukocyte antigen binding affinity of open reading frames from potentially translatable circRNA was predicted using pVACtools. Strong binders from messenger RNA-encoded proteins were excluded using BlastP. The immunogenicity of the candidate antigens was functionally validated through stimulation of naïve $CD8^+$ T cells against the predicted neoepitopes and subsequent analysis of the T cells through enzyme-linked immunospot (ELISpot) assay, intracellular cytokine staining (ICS) and granzyme B (GZMB) reporter. The cytotoxicity of T cells trained with antigen peptides was further tested using patient-derived organoids.

Results We identified a neoepitope from circRAPGEF5 that is upregulated in CRC tumor samples from MiOncoCirc database, and two neoepitopes from circMYH9, which is upregulated across various tumor samples from our matched clinical samples. The translation potential of candidate peptides was supported by Clinical Proteomic Tumor Analysis Consortium database using PepQuery. The candidate peptides elicited antigen-specific T cells response and expansion, evidenced by various assays including ELISpot, ICS and GZMB reporter. Furthermore, T cells trained with circMYH9 peptides were able to specifically target and eliminate tumor-derived organoids but not match normal organoids. This observation underscores the potential of circRNAs as a source of immunogenic neoantigens. Lastly, circMYH9 was enriched in the liquid biopsies of patients with CRC, thus enabling a detection-to-vaccination treatment strategy for patients with CRC.

Conclusions Our findings underscore the feasibility of tumor-associated circRNAs as an alternative source of neoantigens for cancer vaccines targeting tumors with moderate mutation levels.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Neoantigen vaccines, designed to target mutant proteins specific to tumors, present a promising avenue for cancer immunotherapy. Existing strategies depend on identifying immunogenic neoantigens from somatic mutations, but this approach often falls short for tumors with a low to moderate mutation burden, emphasizing the necessity for alternative neoantigen sources. Earlier research has revealed compelling associations between circular RNA (circRNA) expression and diverse cancer types. Importantly, certain circRNAs can undergo translation, producing novel peptides distinct from those encoded by their corresponding messenger RNAs. This distinctive characteristic of circRNAs holds significant potential for advancing cancer vaccine development.

WHAT THIS STUDY ADDS

⇒ We identified tumor-associated circRNAs as a novel source of neoantigens in colorectal cancer. The peptides encoded by these circRNAs elicited potent CD8+T cell responses, demonstrating their potential to overcome the limitation in the number of somatic mutations and increase the repertoire of targetable tumor antigens. We validated this using both in vitro assays and patient-derived organoids, demonstrating targeted elimination of tumor cells. Importantly, we detected these circRNAs in liquid biopsies, enabling cancer vaccine development based on readily accessible biomarkers.

INTRODUCTION

The efficacy of adaptive T-cell immune responses against cancer hinges on their capacity to recognize tumor-specific epitopes expressed by malignant cells. Extensive research in the past decade has been dedicated to exploring the importance of neoantigens, which arise from specific non-synonymous mutations within the tumor genome. Several clinical trials have been initiated to assess personalized cancer vaccines that combine neoantigens, with evidence of

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

This study paves the way for a paradigm shift in cancer immunotherapy by:

⇒ Unlocking a new source of tumor antigens: CircRNAs, previously unexplored, offer an alternative for tumors with limited mutations, expanding the targetable antigen landscape.

⇒ Targeting low-mutation tumors: For tumors lacking sufficient mutations for conventional approaches, circRNAs offer a viable alternative, potentially filling a major gap in current immunotherapy strategies.

⇒ Boosting T-cell response: The ability of circRNA-derived antigens to stimulate CD8+ T-cell responses offers a powerful tool for optimizing vaccine design and enhancing vaccine efficacy.

These findings hold immense potentials for broad applicability and tackling previously challenging tumors.

encouraging clinical results, especially in patients with melanoma. Indeed, multiple studies have suggested that T cells targeting neoantigens can induce tumor regression and prevent relapse in patients with resected tumors.[1–3] Somatic mutations such as single nucleotide variants and frameshift mutations are the most common source of personalized neoantigens. However, most cancers are characterized by low or moderate tumor mutational burden (TMB) and the task of identifying immunogenic neoantigens arising from somatic mutations in these cancers is significantly more challenging. Consequently, the efficacy of vaccines may be compromised due to the limited availability of neoantigens for tumors with low or moderate TMB. Therefore, to broaden the spectrum of immunogenic antigens applicable to cancer vaccines targeting these types of cancers, it is imperative to explore alternative sources of antigens. Here, we focused on colorectal cancer (CRC), of which 80–85% are microsatellite stable (MSS) and typically characterized by moderate TMB.[4] Patients with MSS CRC often demonstrate lower response rates to immune checkpoint inhibitors (ICIs). Therefore, cancer vaccines have the potential to enhance the T-cell response to ICIs, offering a potential avenue to improve the effectiveness of treatment in these patients.[5 6]

Circular RNAs (circRNAs) are a class of diverse, covalently closed single-stranded RNAs. CircRNAs are generated through back splicing of exons in linear pre-messenger RNA (mRNA) transcripts.[7] Back splicing of exons can be facilitated by close proximity of splice sites brought together by base pairing of complementary inverted repeats in the introns flanking the exons.[8–11] The process of back splicing generates new exon–exon junctions and translational frames in the circularized RNA, which could potentially give rise to unique cancer-specific peptides or proteins. Studies have suggested that many circRNAs play biological roles in gene regulation by acting as microRNA sponges or protein inhibitors.[12] Notably, some circRNAs exhibit tissue-specific expression[13] and have also been implicated in several diseases including neurological disorders, cardiovascular diseases and cancer.[12] In cancer, studies have demonstrated that some circRNAs can function either as oncogenes or as tumor suppressors and are predictive biomarkers for disease diagnosis and prognosis.[14–18] CircRNAs were also found to display tumor-specific expression and are upregulated in specific cancer types.[19] For instance, a circRNAs, hsa_circ_0091579, derived from the glypican-3 exon 3, was found to be expressed in liver tumors across multiple patients and was not present in healthy liver cells.[20] In CRC, circCCDC66 expression was higher in tumor relative to the precancerous polyps and correlates with poor overall survival.[21] Due to the lack of exposed ends, circRNAs are more resistant to degradation by exonuclease as compared with linear RNA and may exist stably in exosomes and plasma, allowing them to be detected non-invasively in the liquid biopsies. This property and their tissue specificity make circRNAs attractive biomarker and vaccine candidates for cancer diagnosis and treatment.[22 23]

While most of the characterized circRNAs are non-coding and have biological roles in gene regulation, circRNAs were also found to possess internal ribosome entry sites, allowing their translation into proteins or peptides. Importantly, peptidomics analysis demonstrated that circRNAs are translated into proteins or peptides with tissue-specific distribution.[24] Despite the limitations of mass spectrometry (MS), peptidomics analysis validated the protein expression of ~13% (124 of 957) of circRNA with a predicted open reading frame. In contrast, peptidomics data recovered only ~5% of linear mRNA with the same expression level as the mean expression of circRNA. These lines of evidence support the translation of circRNA in cells and the potential of the translated circRNA to be presented as antigens by human leukocyte antigen (HLA).[24] Hence, circRNAs could potentially be a source of antigens for the activation of the adaptive immune response against the cancer. Through the analysis of circRNAs expression in colorectal samples from MiOncoCirc data set[7] and in our clinical CRC samples, we demonstrated that the upregulated circRNAs, circRAPGEF5 and circMYH9, can give rise to HLA-A*11:01-restricted immunogenic neoepitopes. Furthermore, circRNAs neoepitopes-specific T cells were able to recognize and eradicate the CRC cells. Hence, we demonstrate that circRNAs are a potential source of alternative neoantigens for generating cancer vaccines or antigen-specific T cells for the treatment of CRC.

RESULTS

CircRNAs are differentially expressed in CRC in MiOncoCirc data set

In MSS CRC, the number of somatic mutations serving as potential targets for cancer vaccines is limited. To overcome this constraint, we aim to explore alternative sources of cancer-associated antigens. CircRNAs, generated through back splicing of exons, have demonstrated cancer-specific expression patterns.[19–21] A recent report

by Chen *et al* found that circRNA can be presented as antigens bound to HLA. To investigate if circRNA can give rise to immunogenic antigens, we first analyzed public data from the MiOncoCirc to select circRNA that are differentially expressed between CRC and normal tissues.[7] Potentially translatable open reading frames (ORFs) of the upregulated circRNAs (circular ORFs (circORFs)) were obtained from TransCirc, a specialized database that provides comprehensive evidence supporting the translation potential.[25] The binding affinity to HLA-A*11:01, one of most prevalent HLA alleles in East Asia, of circORFs was predicted using netMHCpan4.1.[26] Nine-mer strong binder antigens (IC_{50}<500 nM) were selected and aligned to NCBI non-redundant protein database using BlastP to screen for novel peptides that are not present in the translated products from linear mRNA (figure 1A). 13,761 circRNA were detected in both normal tissues and CRC samples in the MiOncoCirc data set and kept for differential expression analysis using edgeR[27] (figure 1B). To obtain the translation potential, we used genomic locations (GRCh38) of circRNAs as identifiers to map them to TransCirc. 95.96% of circRNAs were uniquely mapped to TransCirc and the rest were excluded from downstream analyses (figure 1C). We identified 61 upregulated and 35 downregulated circRNA in CRC samples (adjusted p value<0.05 and fold change >2 or <−2, respectively) (figure 1D). Based on their expression, samples were clearly clustered (figure 1E).

CircRNAs are a source of neoantigens

56 out of 61 upregulated circRNA were successfully mapped to TransCirc, with 18 of them having an evidence score >2.5. The circORFs from these circRNAs were then submitted to neMHCpan4.1 to predict their binding affinity to HLA-A*11:01. From the list of antigens, we selected four circRNA antigens from three circRNAs (circLDLR, circPOM121, and circRAPGEF5, see tables 1 and 2), for further expression validation and investigation of immunogenicity (figure 2A). We detected expression of the selected circRNAs using divergent primers in the CRC cell lines, HCT116, HT29, LS513, and SW480 (figure 2B), with HT29 showing an overall higher relative expression of circRNAs. The quantitative real-time PCR (RT-qPCR) products were gel electrophoresis separated and excised for Sanger sequencing to detect back-splicing junction (BSJ), to further validate the divergent primers (figure 2C). The presence of BSJ suggested that these circRNAs were indeed expressed in CRC cell lines and the divergent primers were specific to detect circRNAs. Hence, we also detected the expression of these circRNAs in tumor and adjacent normal tissues from our patient with CRC samples, using the same primers (figure 2D). Among them, circRAPGEF5 showed a significant upregulation in tumor samples, whereas circLDLR and circPOM121 exhibited an upregulated trend in tumor samples, which was generally consistent with MiOncoCirc data. Of note, circRAPGEF5 expression was increased in five out of six tumor samples, especially in patients CV025 and CV030, suggesting that it could be a potential candidate.

Neoantigens derived from circRNA are immunogenic

To examine the immunogenicity of the four selected circRNA neoepitopes, we trained naive $CD8^+$ T cells from HLA-$A*11^+$ healthy donor peripheral blood mononuclear cells (PBMCs) using autologous dendritic cells (DCs), pulsed with circRNA neoantigen peptides, as depicted in the schematic (figure 3A). Tetramer staining revealed that approximately 4.74% of the trained T cells exhibited specificity towards pooled neoepitopes, which was absent in the untrained control T cells or when stained with the tetramer loaded with irrelevant peptide derived from the Epstein-Barr virus (EBV) antigen (figure 3B). To quantify antigen recognition, the trained T cells' interferon-γ (IFN-γ) secretion was measured using an enzyme-linked immunospot (ELISpot) assay. Compared with unpulsed K562 HLA-A*11:01 expressing artificial antigen-presenting cells (aAPCs), aAPCs pulsed with the pooled neoepitopes peptides readily stimulated the IFN-γ secretion of trained T cells (figure 3C).

To identify immunogenic peptides within the pool, aAPCs were pulsed with individual peptides and co-cultured with T cells trained with pooled peptides. Two peptides derived from circRAPGEF5 were found to stimulate IFN-γ secretion in T cells (figure 3C,D). Interestingly, these peptides originated from two distinct circORFs in the same circRNA (table 2), that is, RAPGEF5_1, located in ORF 229–91 spanning positions 229–91 after traversing the BSJ, and RAPGEF5_2 from ORF 311–23. This prompted us to conduct a detailed study of their immunogenicity.

To validate the immunogenicity of individual RAPGEF5 neoepitopes, we repeated the DC-T cell training using individual RAPGEF5 peptides instead of pooled peptides. The T cells that underwent stimulation and expansion in the presence of individual RAPGEF5 peptides will be referred to as R1 and R2 T cells henceforth. Notably, we consistently observed the expansion of R2 reactive T cells using naïve T cells from five healthy donors. While we initially detected R1 reactive T cells (figure 3C,D), we were unable to obtain consistent results using naïve T cells from the five healthy donors. Consequently, we focused our efforts on analyzing R2 for downstream investigations.

To validate the specificity and recognition of the R2 trained T cells, we stimulated the T cells with R2 peptide, or a negative control peptide derived from EBV antigen (irrelevant peptide). The stimulation with R2 peptide led to a significant increase in the number of IFN-γ spots in the trained R2 T cells compared with unpulsed and irrelevant EBV peptide controls (figure 3E). Next, we detected the granzyme B (GZMB) release from these T cells using a GZMB reporter cell line, which is consistently green fluorescent protein (GFP) positive and only elicits infrared fluorescent protein (IFP) signal on cleavage by GZMB released by antigen-specific T cells.[28]

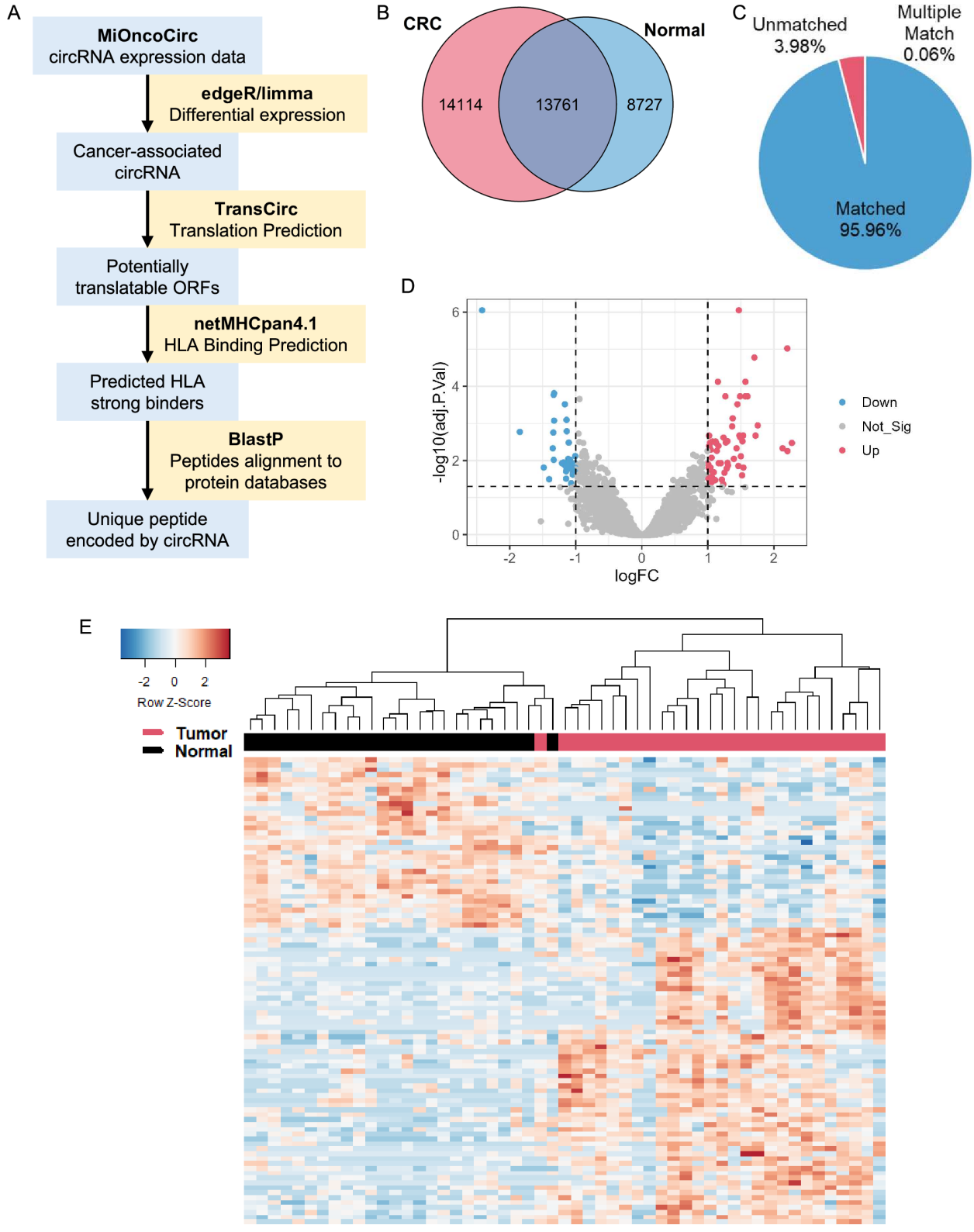

Figure 1 circRNA expression in MiOncoCirc CRC samples. (A) circRNA neoantigen prediction pipeline. circRNA expression data were acquired from MiOncoCirc and differentially expressed circRNA were analyzed using R package edgeR. circRNA were then mapped to TransCirc database for ORF and translation prediction. HLA-A*11:01 binding affinity of potentially translatable ORFs were predicted using netMHCpan4.1 and shortlisted strong binders were aligned to NCBI protein database using BlastP to screen for novel peptides. (B) circRNA detected in CRC and normal samples. Only overlapped circRNA were kept for downstream analyses (C) fraction of circRNA mapped to TransCirc. The genomic locations of circRNA were used as identifiers to map to TransCirc database. Only uniquely matched circRNA were kept. (D) Volcano plot of differentially expressed circRNA. (E) Expression of differentially expressed circRNA in tumor and normal samples. circRNA, circular RNA; CRC, colorectal cancer; HLA, human leukocyte antigen; ORF, open reading frame.

Table 1 Top upregulated circular RNAs in colorectal cancer (fold change >2.5)

TransCirc ID	Fold change	Adjusted p value	Location (hg38)	circBase ID
TC-hsa-DOCK1_0025	4.83	3.36E-03	chr10:126970701–127127764	hsa_circ_0020397
TC-hsa-UBXN7_0044	4.61	5.58E-03	chr3:196391812–196403019	hsa_circ_0001380
TC-hsa-ESRP1_0013	4.60	9.52E-06	chr8:94664696–94665196	hsa_circ_0084927
TC-hsa-ABR_0023	4.38	4.69E-03	chr17:1050049–1100735	hsa_circ_0007919
TC-hsa-CBFA2T2_0037	3.37	1.13E-03	chr20:33619516–33623296	hsa_circ_0003426
TC-hsa-RNF138_0007	3.26	1.67E-05	chr18:32111753–32113860	hsa_circ_0005729
TC-hsa-CUL3_0044	2.97	7.54E-05	chr2:224535527–224557856	hsa_circ_0008309
TC-hsa-PTPN12_0058	2.96	1.86E-04	chr7:77585542–77600806	hsa_circ_0003764
TC-hsa-RNF168_0021	2.90	1.55E-02	chr3:196487398–196488683	–
TC-hsa-UBA2_0004	2.89	2.13E-03	chr19:34430575–34434968	hsa_circ_0005325
TC-hsa-POF1B_0002	2.87	3.03E-03	chrX:85303405–85308216	hsa_circ_0091187
TC-hsa-DYNC1H1_0187	2.86	2.51E-02	chr14:102040235–102040673	hsa_circ_0002398
TC-hsa-RAPGEF5_0092	2.83	7.71E-03	chr7:22291174–22318037	hsa_circ_0001681
TC-hsa-LDLR_0031	2.81	2.41E-03	chr19:11120091–11120522	hsa_circ_0006877
TC-hsa-BAZ1A_0087	2.80	1.86E-04	chr14:34862043–34862322	hsa_circ_0006137
TC-hsa-KDM4C_0023	2.79	2.13E-03	chr9:6880011–6893232	hsa_circ_0001839
TC-hsa-TMEM181_0044	2.77	8.89E-07	chr6:158583953–158589782	hsa_circ_0001663
TC-hsa-PHC3_0107	2.76	1.41E-02	chr3:170145422–170149244	hsa_circ_0001360
TC-hsa-CCSER2_0019	2.73	3.05E-04	chr10:84417770–84425893	hsa_circ_0006956
TC-hsa-ENC1_0005	2.70	4.69E-03	chr5:74634683–74636498	hsa_circ_0006859
TC-hsa-RAPGEF5_0098	2.63	9.04E-03	chr7:22308338–22318037	–
TC-hsa-DNAJB12_0007	2.59	7.30E-04	chr10:72340788–72341170	hsa_circ_0008662
TC-hsa-SAFB2_0024	2.58	1.20E-03	chr19:5604582–5604936	hsa_circ_0000880

The endogenous HLA-A/B/C were knocked out in GZMB reporter cells, and HLA-A*11:01 was transduced to ensure antigen presentation and reduce noise signal. Co-culture of reporter cells pulsed with R2 peptides and trained T cells showed a significant increase in IFP$^+$ population (figure 3F,G), indicating the functional recognition of the neoepitopes by the expanded T cells. As controls, reporter cells either not pulsed with peptides or pulsed with irrelevant EBV peptide showed a significantly lower IFP positive population, suggesting that the R2 T cells specifically recognized the R2 neoepitope. In addition, the trained T cells were restimulated with R2 peptide to examine the production of cytokines, including IFN-γ and tumor necrosis factor α (TNF-α), as well as the cluster of differentiation 137 (CD137) expression, a surrogate marker for T-cell activation[29] (figure 3F,G). As expected, R2 peptide resulted in a significant increase in IFN-γ and TNF-α secretion, as well as CD137 expression, relative to the unpulsed and irrelevant EBV peptide controls. Taken together, these results demonstrate that neoantigens from circRNA, for example, circRAPGEF5, can elicit CD8+T cells response and stimulate expansion, hence, are potential candidates for cancer vaccines.

CircRNAs are differentially expressed in clinical samples from patients with CRC

Having presented evidence supporting circRNA as a potential neoantigen source for cancer vaccines in CRC, our next step is to validate these findings using clinical samples obtained from patients with CRC. We performed ribosome-depletion RNA sequencing (RNA-seq) on four sets of paired adjacent normal and tumor samples.

Table 2 Selected neoantigens and corresponding circRNA from MiOncoCirc

Peptide ID	Sequence	Binding affinity (IC50, nM)	circRNA	Number of evidence	Translation evidence score
LDLR_1	SILDRYHQR	129.65	TC-hsa-LDLR_0031	5	3.61
POM121_1	ITKPNVIIK	85.14	TC-hsa-POM121_0010	4	3.04
RAPGEF5_1	RTLNIRLPR	15.42	TC-hsa-RAPGEF5_0098	4	2.75
RAPGEF5_2	ASVNYSIKK	8.2	TC-hsa-RAPGEF5_0098	4	2.75

circRNA, circular RNA.

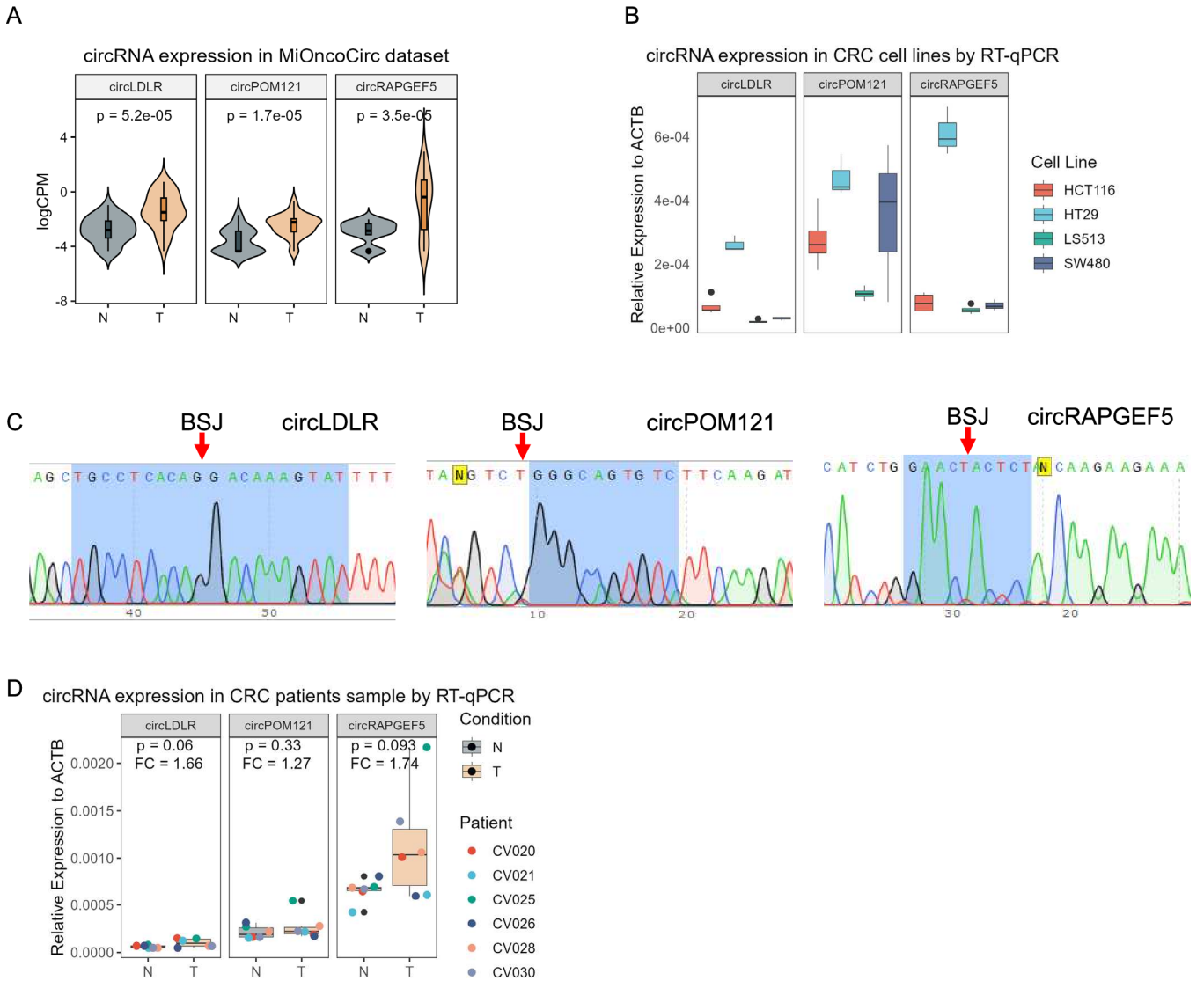

Figure 2 Shortlisted circRNA expression in CRC cell lines and clinical samples. (A) Expression of shortlisted circRNA in CRC (n=28) and normal (n=25) samples in MiOncoCirc data set. (B) Expression of shortlisted circRNA in CRC cell lines by RT-qPCR using divergent primers (n=4 in HCT116 and LS513, n=3 in SW480 and HT29). (C) Sanger sequencing of qPCR products showing the back splicing junction of circRNA. (D) Relative expression of shortlisted circRNA in six pairs of tumor and normal samples from patients with CRC by RT-qPCR. Empirical Bayes moderated t-statistics from edgeR was used in (A). The paired two-sided Student's t-test was used in (D). The boxes in boxplots represented the median ±1 quartile, with the whiskers extending to the largest or smallest non-outlier value within 1.5 times the IQR from the third quartile and first quartile, respectively. BSJ, back-splicing junction; circRNA, circular RNA; CRC, colorectal cancer; RT-qPCR, quantitative real-time PCR; ACTB, actin beta.

CircRNA was identified and quantified using CIRIquant,[30] a bioinformatic tool to accurately characterize and quantify circRNA. The downstream analysis was as shown in figure 1A, except that for HLA binding affinity prediction, pVACbind was employed to increase the sensitivity and specificity[31] (figure 4A). In contrast to netMHCpan alone, pVACbind identifies neoantigens using multiple HLA binding affinity prediction algorithms, including netMHCpan, netMHC, MHCflurry and MHCnuggets.[26 32–34] We were able to detect circRNA from 0.05% to 0.15% of all mapped RNA (figure 4B) using CIRIquant. The differential expression was also performed using CIRIquant (figure 4C). After filtering out circRNAs that were only detected less than 3 samples, 8,772 circRNAs were kept for downstream analyses, among which, 152 were upregulated and 342 were downregulated in tumor samples (logFC>2 or <−2) (figure 4C). Based on differentially expressed circRNAs, normal and tumor samples can be clearly clustered (figure 4D). We successfully mapped 7,882 circRNAs in TransCirc database using their genomic locations as identifiers. To increase the chance of translation, we set the threshold TransCirc evidence score from 2.5 to 3.0 and filtered out 9-mer peptides with a median score >500 (Methods), and finally selected three antigens from one circRNA, that is, circMYH9 (figure 4E and table 3). Of note, peptides MYH9_1 and MYH9_2 shared

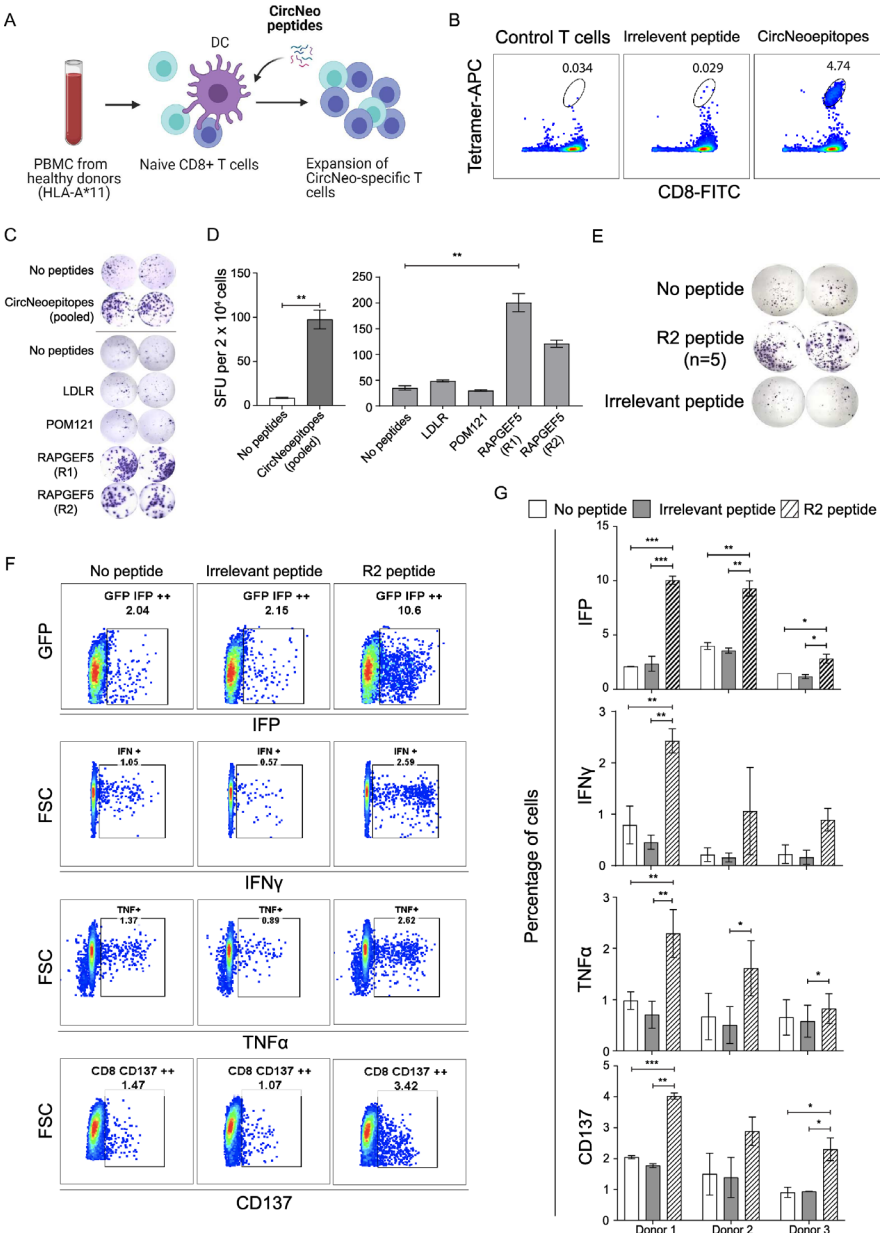

Figure 3 Immunogenicity validation of shortlisted circRNA neoantigens predicted from MiOncoCirc database. (A) Schematics depicting the flow of training HLA-A*11$^+$ naïve CD8$^+$ T cells and the subsequent immunogenicity validation experiments. Created with BioRender.com. (B) Pooled tetramer and CD8 staining of the circRNA neoantigens specific T cells. (C) ELISpot characterization of IFN-γ secretion by circRNA neoantigens specific T cells after stimulation with HLA-A*11:01 expressing aAPCs, in the absence or presence of pooled peptides and as well as individual peptides encoding the circRNA neoantigens. Images were taken and spots were counted using the CTL ImmunoSpot analyzer. (D) Quantification of the number of IFN-γ spots per 20,000 CD8$^+$ T cells in. Two-tailed unpaired Student's t-test was used when comparing two experimental groups (pooled ELISpot) and one-way ANOVA was used when comparing five experimental groups (individual ELISpot) **p<0.01. (E) ELISpot quantification of IFN-γ secretion by RAPGEF5_2 (R2) neoantigen peptides specific T cells after stimulation with HLA-A*11:01 expressing aAPCs, in the absence (top panel) or presence (middle panel) of R2 peptide or the irrelevant EBV peptide (bottom panel). (F) Representative FACS plots of granzyme B reporter validation (first row), ICS of IFN-γ (second row), ICS of TNF-α (third row) and expression of cluster of differentiation 137 (fourth row) in R2 specific T cells in the absence of peptides (first column), presence of irrelevant EBV peptides (second column) and the presence of R2 peptide (third column). (G) Quantification of the representative FACS plots in (F) for trained T cells from three different independent healthy donors. n=3 biological samples per group, each experiment was performed with R2 neoantigen-specific T cells trained independently from three different healthy donors. One-way ANOVA was used for multiple experimental groups with p value adjustment. Data is shown as mean±SEM. *P<0.05; **p<0.01; ***p<0.001. aAPC, artificial antigen-presenting cells; ANOVA, analysis of variance; circRNA, circular RNA; DC, dendritic cell; EBV, Epstein-Barr virus; ELISpot, enzyme-linked immunospot; GFP, green fluorescent protein; HLA, human leukocyte antigen; ICS, intracellular cytokine staining; IFN, interferon; IFP, infrared fluorescent protein; PBMC, peripheral blood mononuclear cell; SFU, spot-forming units; TNF, tumor necrosis factor; FSC, forward scatter; FACS, fluorescence-activated cell sorting; FITC, fluorescein isothiocyanate.

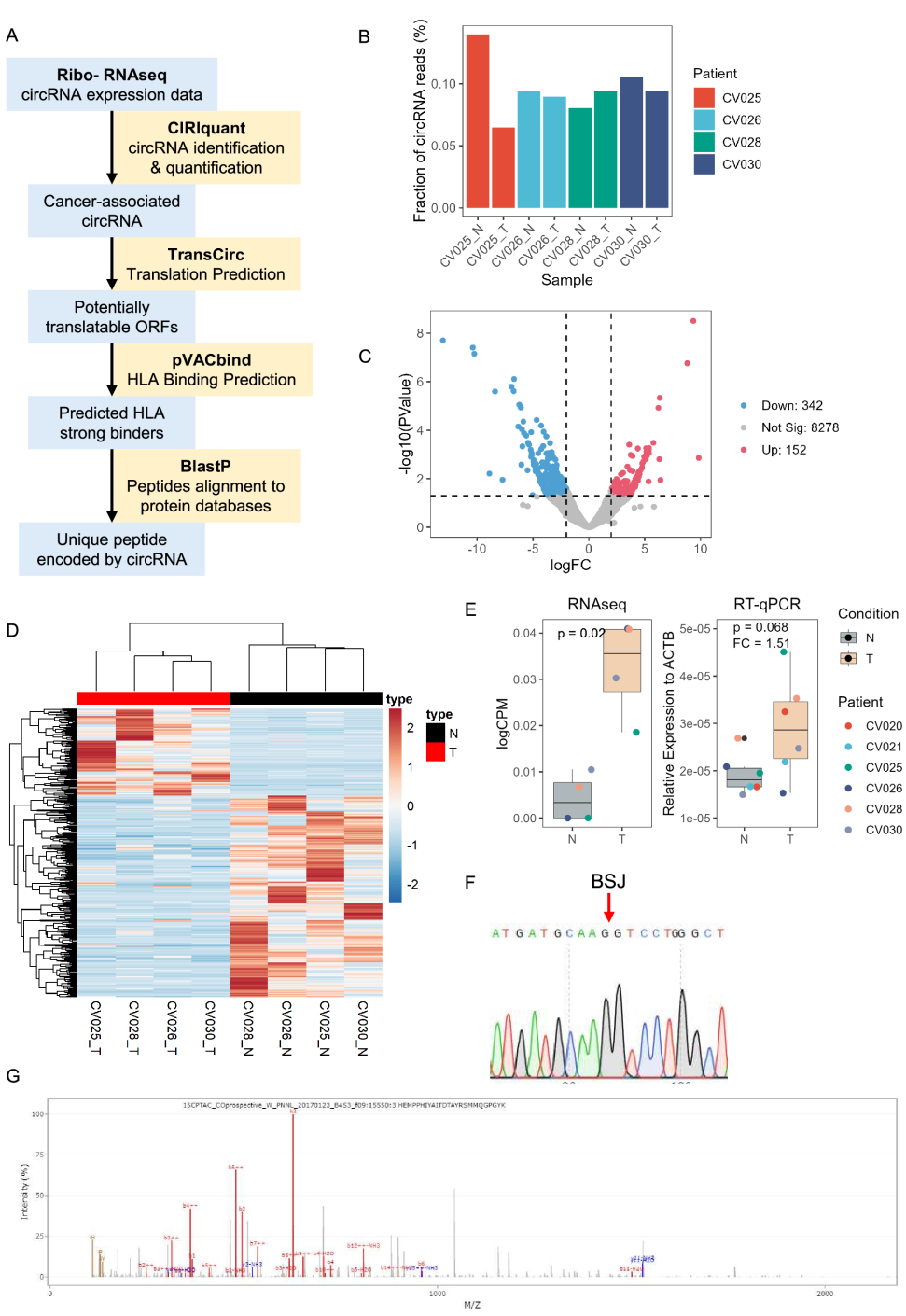

Figure 4 circRNA expression in paired tumor and normal samples from patients with CRC. (A) circRNA neoantigen prediction pipeline. RNA from three pairs of tumor and normal samples were sent for ribosome-depletion RNA-seq. circRNA were identified and quantified using CIRIquant. circRNA were then mapped to TransCirc database for ORF and translation prediction. HLA binding affinity of potentially translatable ORFs were predicted using pVACbind, which employed four algorithms for HLA binding affinity prediction, to increase the sensitivity and specificity. Shortlisted strong binders were aligned to NCBI protein database using BlastP to screen for novel peptides. (B) Fraction of total mapped RNA and circRNA identified from each sample (n=1). (C) Volcano plot of differentially expressed circRNA. (D) Expression of differentially expressed circRNAs in tumor and normal samples from RNA-seq data. (E) Expression of circMYH9. Left, expression data from RNA-seq. Right, Relative expression in six pairs of tumor and normal samples from patients with CRC by RT-qPCR. (F) Sanger sequencing result showing the back splicing junction of circMYH9 from RT-qPCR amplicon. (G) Representative PSM of circMYH9 peptide SMMQGPGYK. Empirical Bayes moderated t-statistics from edgeR and the paired two-sided Student's t-test were used in (E) for RNA-seq and RT-qPCR, respectively. The boxes in boxplots represented the median ±1 quartile, with the whiskers extending to the largest or smallest non-outlier value within 1.5 times the IQR from the third quartile and first quartile, respectively. BSJ, back-splicing junction; circRNA, circular RNA; CRC, colorectal cancer; HLA, human leukocyte antigen; ORF, open reading frame; RNA-seq, RNA sequencing; RT-qPCR, quantitative real-time PCR; CPM, counts per million; ACTB, actin beta.

Table 3 Selected neoantigens and corresponding circRNA from CRC samples

Peptide ID	Sequence	Binding affinity (IC50, nM)	circRNA	Number of evidence	Translation evidence score
MYH9_1	RSMMQGPGY	392.62	TC-hsa-MYH9_0092	5	4.11
MYH9_2	SMMQGPGYK	22.91	TC-hsa-MYH9_0092	5	4.11
MYH9_3	TSSTIRWPR	81.34	TC-hsa-MYH9_0092	5	4.11

circRNA, circular RNA; CRC, colorectal cancer .

eight amino acids, therefore, the peptide MYH9_1, which was predicted to be a weaker binder than MYH9_2, was not tested further. Prior to studying the immunogenicity of these antigens, we first validated the expression of these circRNAs by RT-qPCR in clinical CRC samples. Validation through RT-qPCR confirmed the upregulation of CircMYH9 in five out of six tumor samples (figure 4E). The divergent primer was further validated by Sanger sequencing of the RT-qPCR product and the presence of the BSJ site (figure 4F). When we searched the translation evidence for circMYH9 using PepQuery, 2 out of 312 candidate spectra passed all filters (figure 4G and online supplemental table 1), with one spectrum including the MYH9_2 sequence SMMQGPGYK. This suggests that circMYH9 can indeed be translated into a protein in CRC tumor samples. Additionally, with a lower p value threshold (0.05), 115 spectra were better matched to novel peptides from circMYH9 than unmodified or modified reference protein sequences. These results provided strong translation evidence for circMYH9 and led us to further study the two peptides from circMYH9.

Neoantigens from differentially expressed circRNA in patients with CRC drive tumor-specific killing in patient-derived organoids

Next, we examined the immunogenicity of the two neoepitopes derived from circMYH9, MYH9_2 and MYH9_3. T cells were stimulated and expanded against the pooled MYH9_2 and MYH9_3 peptides (CircNeo). Following this, we assessed the cytotoxicity of these T cells using patient-derived organoids (figure 5A). We determined the antigen specificity by stimulating the expanded T cells with unpulsed aAPCs or aAPCs pulsed with the pooled or individual MYH9 peptides and quantified for IFN-γ secretion. Stimulation with the MYH9_2 peptide, but not the MYH9_3 peptide, resulted in a significant increase in IFN-γ secretion (figure 5B). This was consistent across three different T cells donors (figure 5C). The result correlates with the predicted binding affinity (table 3, IC50 22.91 nM vs 81.34 nM for MYH9_2 and MYH9_3, respectively). Notably, tetramer staining indicated that 2.03% of the trained T cells are specific to the CircNeo peptides and 1.81% are specific to MYH9_2 peptide (figure 5D), whereas no T cells were positively stained with tetramers loaded with the irrelevant EBV peptide. Likewise, tetramer staining of control T cells with CircNeo peptides or MYH9_2 peptide-loaded tetramers did not yield any significant staining. A significant increase in tetramer-positive cells was observed across the three donors (figure 5E), which was consistent with the elevated number of spot-forming units secreting IFN-γ (figure 5C). Thus, we have demonstrated that circMYH9-derived neoantigens, identified from our patient with CRC samples, are highly immunogenic.

To determine if the MYH9_2-specific T cells were able to recognize and eradicate tumors, we derived tumor and normal organoids from the patient's resected tumor and normal adjacent tissue, respectively.[35] Both the tumor and normal organoids were co-cultured with the MYH9_2 peptide-trained T cells for 24 hours, cell death was measured by caspase-3/7 activity and propidium iodide (PI) staining via flow cytometry. Notably, tumor organoids co-cultured with the T cells trained with the MYH9_2 peptide showed an increase in cell death, compared with the normal organoids, as observed by confocal microscopy (figure 5F). The increase of cell death was partially blocked by an anti-HLA blocking antibody, suggesting major histocompatibility complex-I-dependent recognition of tumor organoids by trained T cells. To measure the cytotoxic activity of the T cells against the organoids, we dissociated tumor organoids into single cells, co-cultured them with the trained T cells, and assessed the cell death using flow cytometry. The cell death was more pronounced when the effector:target ratio was increased from 5:1 to 10:1 (figure 5G,H). Thus, our data have shown that T cells trained against MYH9_2 peptides can specifically target and kill the tumor organoids, indicating that MYH9_2 neoantigens could serve as candidates for CRC cancer vaccine. In addition, we also observed similar dose-dependent killing effects when tumor organoids were co-cultured with circRAPGEF5 peptide (R2) trained T cells (figure 5I). Collectively, our findings based on clinical CRC samples indicate that neoantigens originating from circRNAs, such as circMYH9, which exhibit upregulation in CRC tumors, possess immunogenicity and may serve as a source of personalized or common neoantigens to enhance the repertoire of targetable tumor antigens.

circMYH9 is enriched in patient with CRC blood samples

Recent studies have reported that circRNAs can be detected in human blood samples, which are implicated with CRC and act as potential biomarkers for diagnosis.[36 37] Therefore, we wondered if circMYH9 could be found in our patient with CRC blood samples. To this end, we isolated plasma from two patient blood samples (CV026 and CV028), extracted cell-free RNA from plasma and sent them for RNA-seq as we did for tissue samples (figure 6A). Consistent with our hypothesis,

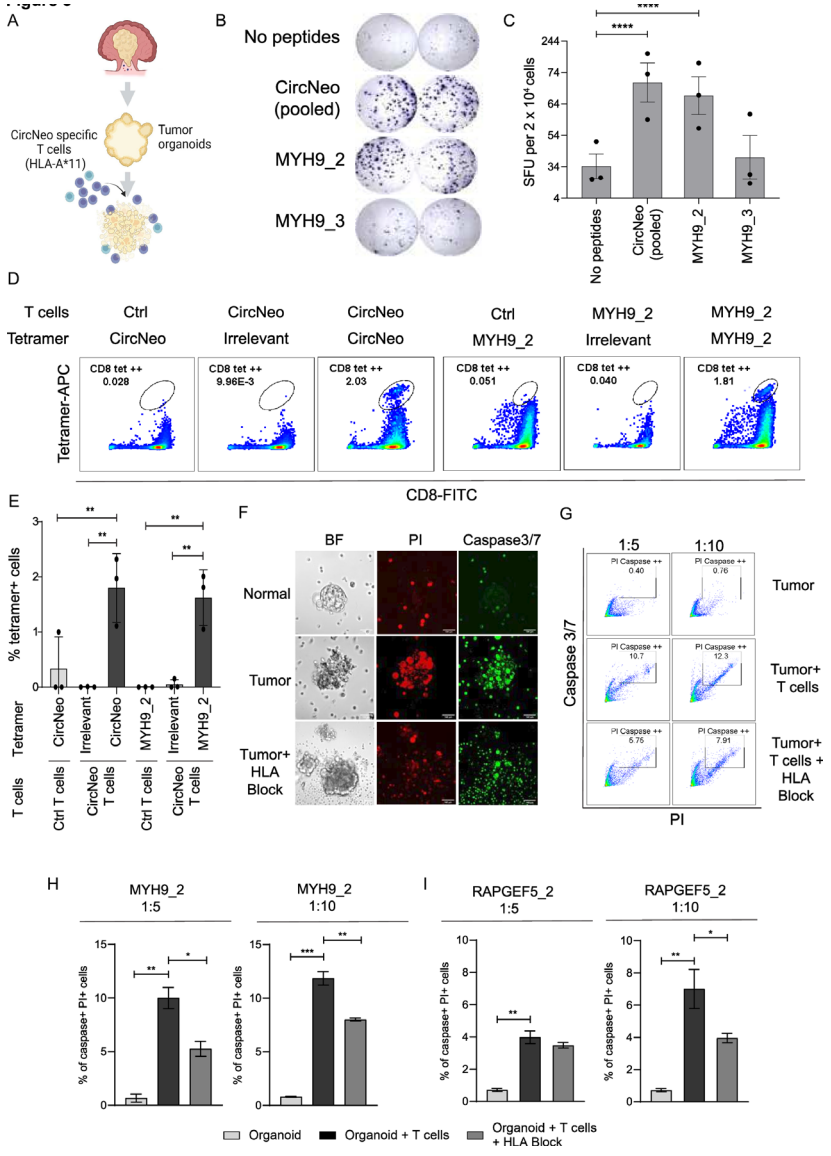

Figure 5 Immunogenicity of shortlisted potential circular RNA neoantigens predicted from patients with CRC. (A) Schematics depicting antitumor effects of trained antigen-specific HLA-A*11 CD8$^+$ T cells on patient-derived organoids. Created with BioRender.com. (B) Enzyme-linked immunospot characterization of IFN-γ secretion by circRNA neoantigens specific T cells after stimulation with HLA-A*11:01 expressing artificial antigen-presenting cells, in the absence (first row) or presence (second row) of CircNeo peptides (MYH9_2 & MYH9_3 pool), as well as individual peptides encoding MYH9_2 (third row) or MYH9_3 (fourth) neoepitopes. Images were taken and spots were counted using the CTL ImmunoSpot analyzer. (C) Quantification of the number of IFN-γ spots per 20,000 CD8$^+$ T cells in. n=3 biological samples per group, each experiment was performed using independently trained T cells from three different healthy donors. One-way ANOVA was used for multiple experimental groups with p value adjustment. Data is shown as mean±SEM. *P<0.0001. (D) Tetramer staining showing the percentage of the control T cells specific to CircNeo peptides (first panel) or the MYH9_2 peptide (fourth panel), CircNeo trained T cells specific to the irrelevant EBV peptide (second panel) or CircNeo peptides (third panel), as well as MYH9_2 trained T cells specific to the irrelevant EBV peptide (fifth panel) and MYH9_2 peptide (last panel). (E) Quantification of tetramer percentages from CircNeo and MYH9_2 trained T cells in (D). n=3 biological samples per group, each experiment was performed using independently trained T cells from three different healthy donors. One-way ANOVA was used for multiple experimental groups with p value adjustment. Data is shown as mean±SEM. **P<0.01. (F) Representative images of organoid-neoantigen specific T cells killing assay showing normal organoid co-cultured with MYH9_2 specific T cells (first row), paired tumor organoid co-cultured with MYH9_2 specific T cells (second row) and tumor organoid co-cultured with MYH9_2 specific T cells in the presence of HLA blockade (third row). Scale bar, 100 µm. (G–I) Dose-dependent killing of tumor organoids by MYH9_2 or RAPGEF5_2 trained CD8$^+$ T cells. Tumor organoids were dissociated into single cells or small clusters prior to co-culture with CD8$^+$ T cells at the effector: target ratio 5:1 and 10:1, and analyzed by flow cytometry. The representative flow cytometry plots from circMYH9_2 T cells were shown in (G). n=2. The two-sided Student's t-test was used to compare between each two groups. Data is shown as mean±SEM. *P<0.05; **p<0.01; ***p<0.001. ANOVA, analysis of variance; APC, antigen presenting cell; circRNA, circular RNA; EBV, Epstein-Barr virus; HLA, human leukocyte antigen; IFN, interferon; PI, propidium iodide; SFU, spot-forming units; FITC, fluorescein isothiocyanate; BF, bright field.

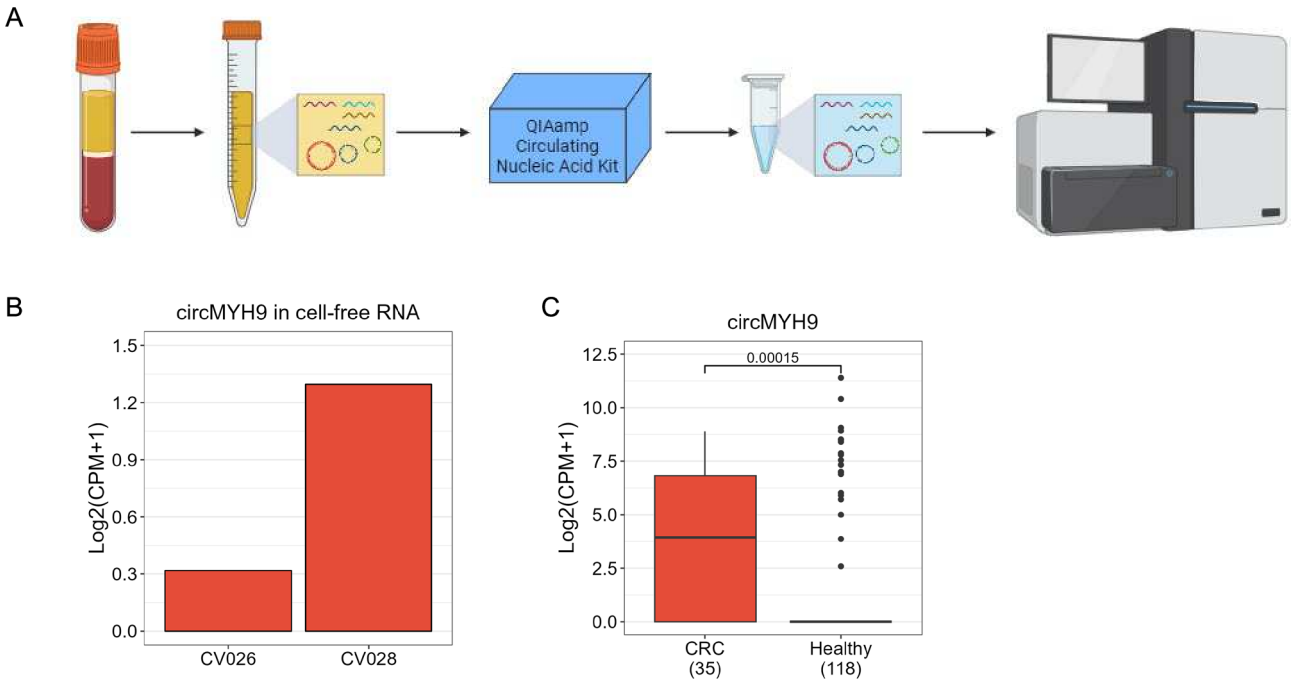

Figure 6 circMYH9 expression in blood samples. (A) Schematics depicting the workflow of cell-free RNA extraction and RNA-seq. Created with BioRender.com. (B) Expression of circMYH9 in two patient with CRC blood samples by RNA-seq (n=1). (C) Expression of circMYH9 in patient with CRC and healthy donor blood samples in exoRbase V.2.0. Wilcoxon's test was used to compare the expression levels between CRC and the healthy group. CRC, colorectal cancer; RNA-seq, RNA sequencing; CPM, counts per million.

circMYH9 was found in both samples (figure 6B). To further examine if circMYH9 is enriched in patient with CRC blood samples than healthy donor blood samples, we explored the exoRbase V.2.0, a repository of RNA-seq data of extracellular vesicles (EVs) in human blood.[38] Interestingly, circMYH9 was not detected in most healthy blood EVs, whereas it was found in most patient with CRC blood samples (figure 6C). This finding suggests that circMYH9 is detectable in the plasma of patients with CRC, potentially enabling a vaccination strategy based on detection. In such a strategy, cancer-associated circRNAs identified from liquid biopsy could be directly employed to predict immunogenic neoantigens for cancer vaccines.

DISCUSSION

The limited presence of somatic mutations in tumors with low to moderate mutational burden poses a significant challenge in the discovery of immunogenic neoantigens, impeding the progress of neoantigen vaccine development. Neoantigens, generated from somatic mutations, play a crucial role in initiating a potent antitumor immune response. However, tumors with a restricted number of somatic mutations offer a limited pool of targetable neoantigens. As a result, the efficacy of neoantigen vaccine development is diminished in such cases. To overcome this obstacle, it becomes imperative to explore alternative sources of tumor-specific or tumor-associated antigens. These alternative antigens can include RNA dysregulation in cancer cells, such as splicing abnormalities, intron retention, and intronic polyadenylation, which give rise to antigens capable of eliciting immune responses.[39–41]

In this study, we demonstrated that there are circRNAs that were upregulated in CRC tumors using the public database MiOncoCirc and our clinical samples, in line with the previous studies.[42–44] By using the comprehensive circRNA translation database, TransCirc, we filtered upregulated circRNAs that have a higher potential to be translated into peptides or proteins. The potential ORFs were submitted to predict their binding affinity to HLA-A*11:01, a common HLA allele in Asian. The predicted strong binders were further screened to remove peptides from existing proteins, using BlastP. Next, we validated the immunogenicity of these potential antigens from upregulated circRNAs. Notably, we were able to stimulate and expand $CD8^+$ T cells against antigens translated from unique novel ORFs resulting from circRNAs, specifically circRAPGEF5 from MiOncoCirc and circMYH9 in our colorectal clinical samples. The enriched T cells targeting neoantigens derived from circRAPGEF5 and circMYH9 successfully recognized and eliminated cancer cells and tumor organoids, respectively. These findings underscore the potential utility of these antigens in cancer vaccines or antitumor T-cell receptor (TCR) therapies.

By expanding our search beyond somatic mutations, we have established circRNA as an alternative source of cancer antigens capable of eliciting a $CD8^+$ T-cell response. This discovery broadens the repertoire of tumor antigens available for targeted immunotherapy. Importantly, we observed the upregulation

of circMYH9 in all four tumor samples compared with adjacent normal tissues, suggesting its prevalence across multiple patients with CRC. Moreover, circMYH9 can also be detected in patient blood samples as cell-free circRNA, and is enriched in blood samples from patients with CRC, compared with healthy donors. This implies that antigens derived from cancer-upregulated circRNAs may be shared among patients, enabling the detection and development of off-the-shelf vaccines. Of note, circMYH9 expression levels were elevated in a few healthy donors. Thus, it will be of great interest to investigate if the circMYH9 is implicated in colon-associated diseases or CRC progression, to comprehensively identify biomarkers that can help early diagnosis and prevention of such diseases.

Nevertheless, we also noted that the expression levels of circRNAs are generally low, potentially compromising the detection and sensitivity required to identify cancer-associated circRNAs. To address this issue, we propose that RNase R enrichment could be advantageous over ribosome-depletion RNA-seq during library preparation, enhancing the sensitivity of circRNA detection. Furthermore, the advancement of next-generation sequencing technologies, particularly long-read sequencing techniques like nanopore sequencing, holds promise for detecting full-length RNA and a greater number of cancer-associated circRNAs,[45] thus improving accuracy and detection. Thus, with the advancement in sequencing technologies, we could potentially detect a greater number of cancer-associated circRNAs that could give rise to immunogenic antigens as alternative candidates for cancer vaccines.

In conclusion, the study demonstrates that cancer-associated circRNAs can generate immunogenic antigens, triggering a $CD8^+$ T-cell response and expansion. This discovery underscores an alternative source of antigens that can be harnessed to expand the range of targetable tumor antigens.

METHODS

Cell lines

HCT116 (CCL-247), SW480 (CCL-228), LS513 (CRL-2134), HT29 (HTB-38), HEK293T (CRL-3216) and K562 (CCL-243) cell lines were purchased from the American Type Culture Collection. All cell lines were tested to be negative for *Mycoplasma* contamination by PCR. All cells were cultured in a 37°C incubator with 5% CO_2. All cells except for LS513 and K562, were cultured in high glucose Dulbecco's Modified Eagle Medium (DMEM, Cytiva, #SH30243.01) supplemented with 10% fetal bovine serum (FBS, Thermo Fisher Scientific, #10437028) and 1% penicillin/streptomycin (P/S, Thermo Fisher Scientific, #15140122). LS513 and K562 cells were cultured in Roswell Park Memorial Institute medium (RPMI 1640, Cytiva, #SH30027.01) supplemented with 10% FBS and 1% P/S.

CLINICAL SAMPLES

Healthy donor blood samples were obtained from the Singapore Health Sciences Authority blood bank (Application #202006–06). Human PBMCs were isolated blood samples using Ficoll-Paque (Cytiva, #17144003). The study involved six patients diagnosed with CRC who underwent surgery at National University Hospital. After resection, samples from both tumor and adjacent normal colon tissues were collected and immediately transferred for organoids derivation and DNA/RNA extraction. Pre-surgery blood samples were collected and plasma was isolated by centrifuging the blood for cell-free RNA extraction.

Patient-derived organoids establishment

For tumor organoids, tumor dissociates were isolated from fresh CRC tumor tissues. Briefly, tissues were cut into 5 mm × 5 mm fragments and washed with 10 mL cold chelation buffers (5.6 mM Na_2HPO4, 8 mM KH_2PO4, 96.2 mM NaCl, 1.6 mM KCl, 43.4 mM sucrose, 54.9 mM sorbitol, 0.5 mM DTT, 1% P/S, 50 µg/mL gentamicin, 2.5 µg/mL amphotericin B, pH 7.0–7.3) four times, with 5 min rotation each time. The washed tissue fragments were minced into 1 mm^3 pieces using a scalpel and incubated in 5 mL digestion buffer (2 mg/mL collagenase P (Sigma # 11249002001), 0.5 mg/mL hyaluronidase (STEMCELL #07461), 10 µM Y-27632 (Sigma #Y0503) in Hanks' Balanced Salt Solution (HBSS, Thermo Fisher Scientific #14025092) for 15 min in a 37°C shaker. After incubation, tumor dissociates were diluted with 20 mL chelation buffer, filtered through a 70 µm cell strainer and centrifuged at 500 G for 5 min. After spinning, tumor dissociates was washed with 20 mL chelation buffer. For seeding, 1 million live cells were resuspended in 200 µL Cultrex UltiMatrix Reduced Growth Factor Basement Membrane Extract (BME) (R&D Systems, #BME001-05) and seeded as 8–10 domes in a well of a warmed 6-well plate. The plate was inverted and incubated in a 37°C incubator for 20 min. Once BME was polymerized, 2 mL CRC organoids culture media, supplemented with 10 µM Y-27632, were added.

For normal organoids, crypts were isolated using EDTA chelation from paired fresh normal colon tissues. Similarly, tissues were cut into 5 mm × 5 mm fragments and washed with 10 mL cold chelation buffers four times. The washed tissue fragments were incubated in 20 mL chelation buffer, supplemented with 2 mM EDTA, for 30 min with rotation. After incubation, crypts were released with vigorous shaking. The supernatant containing the crypts was transferred to a new tube and centrifuged to pellet down the crypts. The crypts were then washed with 20 mL chelation buffer supplemented with 2% FBS twice and resuspended in 5 mL chelation buffer. The quality and quantity of crypts were examined under the microscope. 2000 intact crypts were resuspended in 200 µL BME and seeded. CRC organoids culture media, supplemented with 10 µM Y-27632 and 10 µM CHIR99021 (Sigma #SML1046), were added.

L-WRN conditioned media production

L-WRN cells were seeded at 1:10 ratio in 25 mL culture media (without G418 and hygromycin B) into 15 cm dishes. The media were discarded after 3 days and cells were washed with 10 mL media. 25 mL fresh media were added, and cells were cultured for 24 hours. The conditioned media were collected and centrifuged at 2,000 g for 5 min, decanted into a 1 L sterile bottle and stored at 4°C. Then fresh media were added, and the collection was repeated 12 times. After the last collection, conditioned media were pooled and filtered using 0.22 µm filter units, aliquoted into 50 mL tubes and frozen at −30°C.

CircRNA antigen prediction

For public circRNA, all circRNA expression data, sequencing depths and metadata were downloaded from MiOncoCirc (https://mioncocirc.github.io/download/). Samples from the analysis cohort COLO were extracted as CRC tumor samples, whereas samples with the sample name Human_Normal_Tissues were extracted as normal samples. Two tumor samples (TP_2175-capt-SI_12702-H3Y7VADXX and TP_2263-capt-SI_14737-H2VNJBCXY) were excluded due to low or no circRNA reads, resulting in 28 tumor and 25 normal samples left for downstream analyses. Genomic locations (chr:start-stop) were used as circRNA identifiers. The counts matrix was normalized to sequencing depth using edgeR, and differential expression between tumor and normal samples was analyzed using limma. Upregulated circRNAs in tumor (fold change >2 and adjusted p value<0.05) were then mapped to TransCirc database to obtain translation information. Evidences_score >2.5 was set as the threshold to filter out circRNAs with low translation potential. The predicted ORFs from the remaining circRNAs were submitted to NetMHCpan4.1 to predict the 9-mer peptide binding affinity to HLA-A*11:01. Strong binders (binding affinity <500 nM) were aligned to human non-redundant protein sequences (nr) using NCBI blastp, to exclude those from parental mRNA-encoded proteins. Last, 9-mer peptides from ORFs that start with leucin (CTG or TTG), or are infinite were also excluded.

For our CRC samples, four pairs of normal and tumor RNA samples and two cell-free RNA samples were sent to Azenta for ribosomal depletion RNA-seq, with 100 M paired-end 150 bp reads for each sample. RNA-seq libraries were prepared using Ribo-off rRNA Depletion Kit (Vazyme, #N406) following the manufacturer's instruction. The raw reads were trimmed with TrimGalore (V.0.6.7) with flag -q 5. CircRNAs were quantified using CIRIquant from trimmed reads and differential expression was analyzed using CIRI_DE_replicate module. The output was further filtered for circRNAs that were detected in at least three samples. The downstream workflow was the same as above except that the Evidences_score threshold was increased to three and HLA binding affinity was predicted using the pVACbind module in pVACtools (V.3.0.3). Median binding affinity from four algorithms (netMHC, netMHCpan, MHCflurry, and MHCnuggetsI) was used to select strong binders (<500 nM).

DNA and RNA extraction

Patient with CRC sample DNA and RNA was extracted using AllPrep DNA/RNA Mini Kit (Qiagen, #80204). Human PBMC DNA was extracted using Wizard Genomic DNA Purification Kit (Promega, #A1125). Cell line RNA were extracted using E.Z.N.A. Total RNA Kit I (Omega Bio-Tek, #R6834). Cell-free RNA from blood samples was extracted using QIAamp Circulating Nucleic Acid Kit (Qiagen, # 55114).

RT-qPCR and Sanger sequencing

Complementary DNA (cDNA) was synthesized using Maxima First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, #K164) from 5 µg of total RNA. CircRNA and housekeeping gene (actin beta, ACTB) expression were quantified by RT-qPCR using SYBR Select Master Mix (Thermo Fisher Scientific, #4472919) with 50 ng cDNA per 10 µL reaction as input. Primers for RT-qPCR were listed in online supplemental table 2. RT-qPCR products were separated by agarose gel electrophoresis, excised, and extracted using E.Z.N.A. Gel Extraction Kit (Omega Bio-Tek, #D2500). Purified DNA was sequenced using RT-qPCR primers to look for back-splicing junctions. For each sample, the RT-qPCR was performed in triplicate and the threshold cycle values (Ct) were obtained using Design and Analysis Software (Thermo Fisher Scientific, V.2.6). Mean Ct values were processed according to the comparative Ct method and normalized to the expression of the housekeeping gene ACTB.

Plasmid

Lentiviral plasmids pMDLg/pRRE (Addgene #12251), pMD2.G (Addgene #12259) and pRSV-Rev (Addgene #12253) were a gift from Didier Trono. pLV-EF1a-IRES-Neo was a gift from Tobias Meyer (Addgene #85139). pcDNA3.1 iCasper T2A HO1 was a gift from Xiaokun Shu (Addgene #64278). AAVS1-CAG-hrGFP was a gift from Su-Chun Zhang (Addgene # 52344). AAVS1-TALEN-L and AAVS1-TALEN-R were a gift from Danwei Huangfu (Addgene #59025 and #59026).

The gBlock encoding HLA-A*11:01 coding sequence was cloned using EcoRI (New England Biolabs #R3101L) and BamHI (New England Biolabs #R3136L) into pLV-EF1a-IRES-Neo to generate lentivirus.

The gBlocks encoding GZMB reporter sequence and caspase-resistant Inhibitor of Caspase Associated Dnase (cr-ICAD) sequence[28] were synthesized by Integrated DNA Technologies (IDT). First, the gBlock encoding GZMB reporter was cloned into pcDNA3.1 iCasper T2A HO1 vector to replace the caspase cleavage site with the GZMB cleavage site, using BamHI and EcoRI. Then the GZMB reporter sequence was amplified from pcDNA3.1 iCasper T2A HO1 vector, digested using SalI (New England Biolabs #R3138S) and EcoRV (New England Biolabs #R3198S), and cloned into AAVS1-CAG-hrGFP

vector to replace the hrGFP sequence (AAVS1-CAG-GZMB Reporter). The gBlock encoding cr-ICAD was directly cloned into AAVS1-CAG-hrGFP vector using SalI and EcoRV (AAVS1-CAG-cr-ICAD).

The gBlock sequences and primer sequences were listed in online supplemental table 2.

HLA haplotyping

Healthy donor HLA haplotypes were detected by PCR using HLA-A*02, HLA-A*11 and HLA-A*24 specific primers (online supplemental table 2). APC primers were added as a positive control. PCR reactions were carried out using GoTaq G2 Green Master Mix (Promega, #M7822) with 800 nM primers, under the following conditions: 2 min at 96°C; 6 cycles of 30 s at 96°C, 45 s at 69°C, and 45 s at 72°C; 21 cycles of 30 s at 94°C, 45 s at 65°C, and 45 s at 72°C; and 5 cycles of 30 s at 96°C, 60 s at 55°C, and 2 min at 72°C. HLA haplotypes were determined by the presence of specific PCR products after gel electrophoresis (HLA-A*02, 975 bp; HLA-A*11, 522 bp; HLA-A*24, 559 bp; and APC, 251 bp).

PepQuery

MS evidence for circRNA translation was searched using PepQuery web portal (http://pepquery2.pepquery.org/). The two ORFs from circRAPGEF5 and one ORF from circMYH9 were submitted as novel protein sequences. Two Clinical Proteomic Tumor Analysis Consortium (PDC000116 and PDC000109) data sets and one The Cancer Genome Atlas (PDC000111) data set were searched. Swissprot_human_20220527 database was used as a reference for filtering. Hyperscore was calculated for each spectrum.

Generation of lentivirus

HEK293T cells were seeded in a 10 cm dish 1-day prior to transfection at 70% confluency. 40 µg of pLV-EF1a-HLA-A*11:01-IRES-Neo, pMDLg/pRRE, pMD2.G, and pRSV-Rev (mass ratio 4:2:1:1) were transfected using Lipofectamine 3000 Transfection Reagent (Thermo Fisher Scientific # L3000015) into HEK293T cells. Media were changed 6 hours after transfection. The media containing virus were collected at 24, 48, and 72 hours after transfection, pooled and centrifuged using ultracentrifuge (100,000 g, 2 hours at 4°C). Virus pellets were resuspended in phosphate-buffered saline (PBS) and kept in −80°C.

HLA knockout

The CRISPR-Cas9 guide RNAs (crRNAs) targeting HLA-A/B/C genes (online supplemental table 2) were synthesized by IDT. Transfection of Cas9:gRNA ribonucleoprotein (RNP) complex was performed using Lipofectamine CRISPRMAX Cas9 Transfection Reagent (Thermo Fisher Scientific #CMAX00003) following the manufacturer's instruction in HEK293T. Briefly, each crRNA was annealed with CRISPR-Cas9 tracrRNA (IDT #1072532) and incubated with Cas9 Nuclease (IDT #1081060) in OptiMEM media (Thermo Fisher Scientific, #31985070), in the presence of Cas9 PLUS reagent. The assembled RNP complex was then mixed with CRISPRMAX reagent in OptiMEM media and reversely transfected into the HEK293T cells. Three days after transfection, cells were harvested and stained with anti-human HLA-A/B/C antibody (BioLegend, #311413) for flow cytometry (BD LSRFortessa Cell Analyzer). HLA-A/B/C knockout cells were sorted using BD FACSAria Fusion Cell Sorter.

Generation of aAPC

The HLA-null K562 cells were used to generate aAPC. Cells were transduced with CD80 (Origene, #RC206540L1V), CD86 (Origene, #RC217341L1V) and HLA-A*11:01 using spinoculation. Briefly, cells were mixed with appropriate amounts of virus and 8 µg/mL polybrene (Merck #TR-1003-G), centrifuged at 800 G for 1 hour at 32°C and cultured for 1 day. Media containing virus were replaced the next day to expand cells. CD80, CD86 and HLA triple positive cells were sorted using BD FACSAria Fusion Cell Sorter.

Generation of GZMB reporter cells

HEK293T HLA KO cells were transfected with AAVS1-TALEN-L, AAVS1-TALEN-R plasmid and AAVS1-CAG-cr-ICAD (mass ratio 1:1:1) using the Lipofectamine 3000. The transfected cells were then selected with puromycin (Thermo Fisher Scientific #A1113803) at 1 µg/mL to get stable cells expressing cr-ICAD. The HEK293T-HLA KO-cr-ICAD cells were then transfected with AAVS1-TALEN-L, AAVS1-TALEN-R plasmid and AAVS1-CAG-GZMB Reporter (mass ratio 1:1:1). Stable cells expressing the GZMB Reporter were selected by sorting for GFP positive cells. Last, HEK293T-HLA KO-cr-ICAD-GZMB Reporter cells were transduced with the HLA-A*11:01 using lentivirus. Stable cells expressing HLA-A*11:01 were selected by staining for HLA and sorted.

DCs-naïve CD8$^+$ T cells co-culture

Classical monocytes were isolated from healthy donor PBMCs using EasySep Human Monocyte Isolation Kit (STEMCELL, #19359). Monocytes were then plated, differentiated and matured using ImmunoCult Dendritic Cell Culture Kit (STEMCELL, #10985). Nine-mer neoantigen peptides were ordered from GenScript and added to the DCs at a concentration of 100 µg/mL during the maturation step. DCs were then lifted from the wells by pipetting up and down gently using PBS with 2% Human Serum (HS, Sigma #H3667) and 1% P/S, at 22 hours post maturation. Naïve CD8$^+$ T cells were isolated from the PBMCs using EasySep Human Naïve CD8$^+$ T Cell Isolation Kit (STEMCELL #19258). Matured DCs were then co-cultured with the naïve CD8$^+$ T cells at a 1:4 ratio (DC: naïve CD8$^+$ T cell) in RPMI 1640 with 10% HS, 1% P/S, 30 ng/mL interleukin (IL)-21 (Miltenyi #130-095-768), 40 ng/mL IL-7 (Miltenyi #130-095-361) and 40 ng/mL IL-15 (Miltenyi #130-095-762) for 3 days in a 48-well plate. On day 4, the expansion media (RPMI 1640 with 10% HS, 2% P/S, 200 IU/mL IL-2 (Miltenyi #130-097-744), 40 ng/

mL IL-7 and 40 ng/mL IL-15), were added to the co-cultured cells. Subsequently, the expansion media were added once every 3 days for a total of 14 days to allow for the proliferation and expansion of neoantigen-specific $CD8^+$ T cells. On day 14 days, the expanded T cells were used for immunogenicity validation assays.

GZMB reporter assay

One day before the co-culture, GZMB reporter cells were seeded in a 24-well. The neoantigen peptides of interest were added to the reporter cells at a final concentration of 100 µg/mL. On the day of the co-culture, neoantigen-specific T cells were added to the respective wells of the GZMB reporters at a ratio of 1:4 (reporter:T cell). The plate was then placed in a 37°C incubator for 6 hours. Then the co-cultured cells were lifted from the wells by pipetting up and down and resuspended in the flow buffer (1% BSA in PBS). The cells were pelleted down by centrifuging at 300 g for 5 min and the supernatant was discarded. The cell pellets were then resuspended in the flow buffer with 4′,6-diamidino-2-phenylindole (DAPI, Abcam, #ab228549, 1:5000) and analyzed using a flow cytometer. Live cells were gated using the DAPI stain and multiplets were excluded. From the single cells, GFP-positive GZMB reporter cells were gated and IFP positive population was calculated.

ELISpot assay

ELISpot assays were performed following the manufacturer's instructions. Briefly, plates from Human IFN-γ ELISpot PLUS kit (Mabtech, #3420-4APT-10) were washed with PBS and blocked with PBS containing 10% FBS for 30 min at 37°C. After incubation, the plates were washed five times with PBS. 20,000 neoantigen expanded T cells were stimulated with 2000 K562 aAPC in the absence (control) or presence of neoantigen or EBV peptides for 24 hours. The cells were then discarded, and the wells were washed five times with PBS. The wells were then incubated with biotin-conjugated anti-IFN-γ primary antibody for 2 hours followed by a PBS wash and then incubated with streptavidin-alkaline phosphatase (ALP) secondary antibody for 1 hour. After incubation, 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium (BCIP/NBT)-plus substrate was added and incubated at room temperature for 15 min. Plates were dried overnight, and the spots were scanned and analyzed using CTL ImmunoSpot S6 Entry M2 Analyzer.

Tetramer staining

Flex-T HLA-A*11:01 monomers UVX (BioLegend, #280007) was irradiated with ultraviolet light (UV) in the presence of the peptides. The monomers were then assembled into tetramers with APC-conjugated streptavidin following the manufacturer's instructions. The neoantigen-expanded T cells were stained with peptides-loaded tetramers and anti-CD8 antibody (BioLegend, #344718). After staining, cells were pelleted down by centrifuging at 300 g for 5 min and the supernatant was discarded. The cell pellets were then resuspended in the flow buffer with PI (Sigma-Aldrich, #P4170, 1:1000) and analyzed using a flow cytometer.

Intracellular cytokine staining

One day before co-culture, aAPC were pulsed with peptides at a final concentration of 100 µg/mL. Peptide pulsed aAPC were then mixed with the respective neoantigen expanded T cells at a ratio of 1:4 (aAPC:T cell), in the presence of CD107a antibody (BD Biosciences, #641581, 1:200). The plates were incubated for 2 hours at 37°C, then GolgiStop (BD Biosciences, #554715, 1:1,500) and GolgiPlug (BD Biosciences, #555029, 1:1,000) were added and the plates were incubated for another 6 hours. After incubation, the co-cultured cells were stained with LIVE/DEAD Fixable Near-IR stain (Thermo Fisher Scientific, #10119, 1:2000) and anti-CD8 antibody (BioLegend, #344718, 1:200) for 30 min on ice. The cells were washed with the flow buffer and centrifuged at 500 g for 5 min three times, then fixed and permeabilized with Cytofix/Cytoperm Kit (BD Biosciences, #554714) for 20 min on ice and washed three times. The cells were then stained with anti-TNF-α (BioLegend, #502930, 1:200), anti-IFN-γ (BioLegend, #506504, 1:200) and anti-CD8 antibodies for 30 min on ice, washed three times and then analyzed via flow cytometry.

CD137 expression

Peptide pulsed aAPC were then co-cultured with the respective neoantigen expanded T cells at a ratio of 1:4 (aAPC:T cell) for 24 hours. After incubation, the cells were stained with LIVE/DEAD Fixable Near-IR stain, anti-CD8 and anti-CD137 antibody (BioLegend, #309810, 1:200) for 30 min on ice. The cells were then washed three times and analyzed via flow cytometry.

Organoid killing assay

The CRC normal and tumor organoids were cultured for about 2 weeks till they reached sufficient confluency and size. On day 1, both normal and tumor organoids were harvested using the cultrex organoid harvesting solution (R&D Systems, #2700-100-01), following the manufacturer's instruction. The harvested organoids were either dissociated into single cells or small clusters (for fluorescence-activated cell sorting (FACS)), or directly plated (for confocal microscopy) and incubated in a non-tissue culture treated 6-well plate in organoid culture media for 24 hours. On the day 2, IFN-γ (STEMCELL, #78141.1) was added to the wells. 24 hours after IFN-γ treatment, the organoids were collected and stained with CellTracker Deep Red Dye (Thermo Fisher Scientific, #C34565). The neoantigen-specific T cells were stained with CellTracker Blue CMAC Dye. After staining, both the organoids and T cells were counted, and mixed together at 1:5 ratio (organoid:T cells) in the co-culture media (RPMI1640 with 5% HS, 1% P/S, 10 µM Y-27632 and CellEvent Caspase-3/7 Green reagent). Anti-human HLA-A/B/C monoclonal blocking antibody (BioLegend,

311428) was added to the respective wells to block HLA-mediated recognition. The plates were placed in the incubator for 72 hours and the cells were processed according to the protocol as described previously.[46]

Statistical analyses

All quantitative data represent the mean±SD unless otherwise stated. Statistical significance was accessed with paired or unpaired two-tailed Student's t-tests using Prism V.8 (GraphPad) or R. For all figures: *p<0.05; **p<0.01; ***p<0.001.

Acknowledgements We thank and acknowledge Singapore Health Sciences Authority for providing healthy donor blood samples, and Mr Khairul Rifdi Bin Khairul Sani and Mr Brandon Kee Bing Rui for their assistance to TM.

Contributors GC conceived and supervised the study. BES, W-KC, KYL, IJ-WT, BL, and K-KT provided clinical samples. YR processed clinical samples, conducted the bioinformatics analyses and RT-qPCR-related experiments. TM, BL, and CZMC performed experiments related to immunology. GC, YR and TM wrote the manuscript. GC is responsible for the overall content as guarantor.

Funding GC was supported by National Research Foundation (NRF) Singapore, under the NRF fellowship (NRF-NRFF12-2020-0007) and Institute for Health Innovation & Technology (iHealthtech). TM and BL were supported by NUS Research Scholarship from the Singapore Ministry of Education.

Competing interests No, there are no competing interests.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by NUS DSRB 2020/01343 and NUS-IRB LH-20-026E. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iDs
Yi Ren http://orcid.org/0000-0002-3809-0967
Gloryn Chia http://orcid.org/0000-0002-2099-719X

REFERENCES

1 Linnemann C, van Buuren MM, Bies L. High-throughput epitope discovery reveals frequent recognition of neo-antigens by CD4+ T cells in human melanoma. *Nat Med* 2015;21:81–5.
2 Verdegaal EME, de Miranda NFCC, Visser M, *et al*. Neoantigen landscape dynamics during human melanoma-T cell interactions. *Nature* 2016;536:91–5.
3 Zacharakis N, Chinnasamy H, Black M, *et al*. Immune recognition of somatic mutations leading to complete durable regression in metastatic breast cancer. *Nat Med* 2018;24:724–30.
4 Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. *J Clin Oncol* 2005;23:609–18.
5 Fabrizio DA, George TJ, Dunne RF, *et al*. Beyond microsatellite testing: assessment of tumor mutational burden identifies subsets of colorectal cancer who may respond to immune checkpoint inhibition. *J Gastrointest Oncol* 2018;9:610–7.
6 Schrock AB, Ouyang C, Sandhu J, *et al*. Tumor mutational burden is predictive of response to immune checkpoint inhibitors in MSI-high metastatic colorectal cancer. *Ann Oncol* 2019;30:1096–103.
7 Vo JN, Cieslik M, Zhang Y, *et al*. The landscape of circular RNA in cancer. *Cell* 2019;176:869–81.
8 Conn SJ, Pillman KA, Toubia J, *et al*. The RNA binding protein quaking regulates formation of circRNAs. *Cell* 2015;160:1125–34.
9 Ashwal-Fluss R, Meyer M, Pamudurti NR, *et al*. circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell* 2014;56:55–66.
10 Liang D, Wilusz JE. Short intronic repeat sequences facilitate circular RNA production. *Genes Dev* 2014;28:2233–47.
11 Zhang X-O, Wang H-B, Zhang Y, *et al*. Complementary sequence-mediated exon circularization. *Cell* 2014;159:134–47.
12 Kristensen LS, Andersen MS, Stagsted LVW, *et al*. The biogenesis, biology and characterization of circular RNAs. *Nat Rev Genet* 2019;20:675–91.
13 Memczak S, Jens M, Elefsinioti A, *et al*. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 2013;495:333–8.
14 Weng W, Wei Q, Toden S, *et al*. Circular RNA ciRS-7-A promising prognostic biomarker and a potential therapeutic target in colorectal cancer. *Clin Cancer Res* 2017;23:3918–28.
15 Li R-C, Ke S, Meng F-K, *et al*. Cirs-7 promotes growth and metastasis of esophageal squamous cell carcinoma via regulation of miR-7/HOXB13. *Cell Death Dis* 2018;9:838.
16 Yu L, Gong X, Sun L, *et al*. The circular RNA Cdr1As act as an oncogene in hepatocellular carcinoma through targeting miR-7 expression. *PLoS ONE* 2016;11:e0158347.
17 Okholm TLH, Nielsen MM, Hamilton MP, *et al*. Circular RNA expression is abundant and correlated to aggressiveness in early-stage bladder cancer. *NPJ Genom Med* 2017;2:1–14.
18 Smid M, Wilting SM, Uhr K, *et al*. The circular RNome of primary breast cancer. *Genome Res* 2019;29:356–66.
19 Wang S, Zhang K, Tan S, *et al*. Circular RNAs in body fluids as cancer biomarkers: the new frontier of liquid biopsies. *Mol Cancer* 2021;20:13.
20 Zhang C, Zhang C, Lin J, *et al*. Circular RNA Hsa_Circ_0091579 serves as a diagnostic and prognostic marker for hepatocellular carcinoma. *Cell Physiol Biochem* 2018;51:290–300.
21 Hsiao K-Y, Lin Y-C, Gupta SK, *et al*. Noncoding effects of circular RNA CCDC66 promote colon cancer growth and metastasisCircCCDC66 promotes colon cancer progression. *Cancer Res* 2017;77:2339–50.
22 de Fraipont F, Gazzeri S, Cho WC, *et al*. Circular RNAs and RNA splice variants as biomarkers for prognosis and therapeutic response in the liquid biopsies of lung cancer patients. *Front Genet* 2019;10:390.
23 Zhou X, Liu H-Y, Wang W-Y, *et al*. Hsa_Circ_0102533 serves as a blood-based biomarker for non-small-cell lung cancer diagnosis and regulates apoptosis in vitro. *Int J Clin Exp Pathol* 2018;11:4395–404.
24 Chen C-K, Cheng R, Demeter J, *et al*. Structured elements drive extensive circular RNA translation. *Mol Cell* 2021;81:4300–18.
25 Huang W, Ling Y, Zhang S, *et al*. Transcirc: an interactive database for translatable circular RNAs based on multi-omics evidence. *Nucleic Acids Res* 2021;49:D236–42.
26 Reynisson B, Alvarez B, Paul S, *et al*. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif Deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* 2020;48:W449–54.
27 Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;26:139–40.
28 Kula T, Dezfulian MH, Wang CI, *et al*. T-scan: a genome-wide method for the systematic discovery of T cell epitopes. *Cell* 2019;178:1016–28.
29 Wolfl M, Kuball J, Ho WY, *et al*. Activation-induced expression of CD137 permits detection, isolation, and expansion of the full repertoire of CD8+ T cells responding to antigen without requiring knowledge of epitope specificities. *Blood* 2007;110:201–10.
30 Zhang J, Chen S, Yang J, *et al*. Accurate quantification of circular RNAs identifies extensive circular Isoform switching events. *Nat Commun* 2020;11:90.
31 Hundal J, Kiwala S, McMichael J, *et al*. pVACtools: a computational toolkit to identify and visualize cancer neoantigens. *Cancer Immunol Res* 2020;8:409–20.

32 Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* 2016;32:511–7.
33 O'Donnell TJ, Rubinsteyn A, Bonsack M, *et al*. MHCflurry: open-source class I MHC binding affinity prediction. *Cell Syst* 2018;7:129–32.
34 Shao XM, Bhattacharya R, Huang J. High-throughput prediction of MHC class I and II neoantigens with MHCnuggets. *Cancer Immunol Res* 2020;8:396–408.
35 Magré L, Verstegen MMA, Buschow S, *et al*. Emerging organoid-immune co-culture models for cancer research: from oncoimmunology to personalized Immunotherapies. *J Immunother Cancer* 2023;11:e006290.
36 Dong L, Lin W, Qi P, *et al*. Circulating long RNAs in serum extracellular vesicles: their characterization and potential application as biomarkers for diagnosis of colorectal cancer. *Cancer Epidemiol Biomarkers Prev* 2016;25:1158–66.
37 Li J, Song Y, Wang J, *et al*. Plasma circular RNA panel acts as a novel diagnostic biomarker for colorectal cancer detection. *Am J Transl Res* 2020;12:7395–403.
38 Lai H, Li Y, Zhang H, *et al*. exoRBase 2.0: an atlas of mRNA, lncRNA and circRNA in extracellular vesicles from human biofluids. *Nucleic Acids Res* 2022;50:D118–28.
39 Pan Y, Kadash-Edmondson KE, Wang R, *et al*. RNA dysregulation: an expanding source of cancer immunotherapy targets. *Trends Pharmacol Sci* 2021;42:268–82.
40 Li M, Wang Y, Wu P, *et al*. Application prospect of circular RNA-based neoantigen vaccine in tumor immunotherapy. *Cancer Lett* 2023;563:216190.
41 Xia J, Li S, Ren B, *et al*. Circular RNAs as a potential source of neoepitopes in cancer. *Front Oncol* 2023;13:1098523.
42 Hsiao K-Y, Lin Y-C, Gupta SK, *et al*. Noncoding effects of circular RNA CCDC66 promote colon cancer growth and metastasis. *Cancer Res* 2017;77:2339–50.
43 Xu H, Wang C, Song H, *et al*. RNA-seq profiling of circular RNAs in human colorectal cancer liver metastasis and the potential biomarkers. *Mol Cancer* 2019;18:8.
44 Zeng K, Chen X, Xu M, *et al*. CircHIPK3 promotes colorectal cancer growth and metastasis by sponging miR-7. *Cell Death Dis* 2018;9:417.
45 Zhang J, Hou L, Zuo Z, *et al*. Comprehensive profiling of circular RNAs with nanopore sequencing and CIRI-long. *Nat Biotechnol* 2021;39:836–45.
46 Cattaneo CM, Dijkstra KK, Fanchi LF, *et al*. Tumor organoid-T-cell coculture systems. *Nat Protoc* 2020;15:15–39.