Supplementary Data
Figure S1

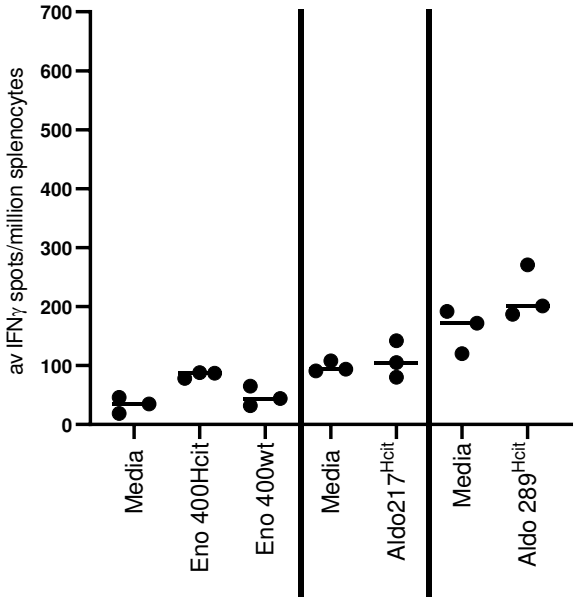


Figure S1. Immunogenicity of Hcit peptides in HLA-DP4 transgenic mice does not correlate with predicted MHCII binding affinity and not all peptides stimulate IFNγ responses.
Ex vivo ELISpot was used to assess IFNγ responses to peptides in HLA-HHDII/DP4 mice. For all studies mice were immunized with three doses of peptide with CpG/MPLA as an adjuvant and responses were assessed at day 21.

Figure S2

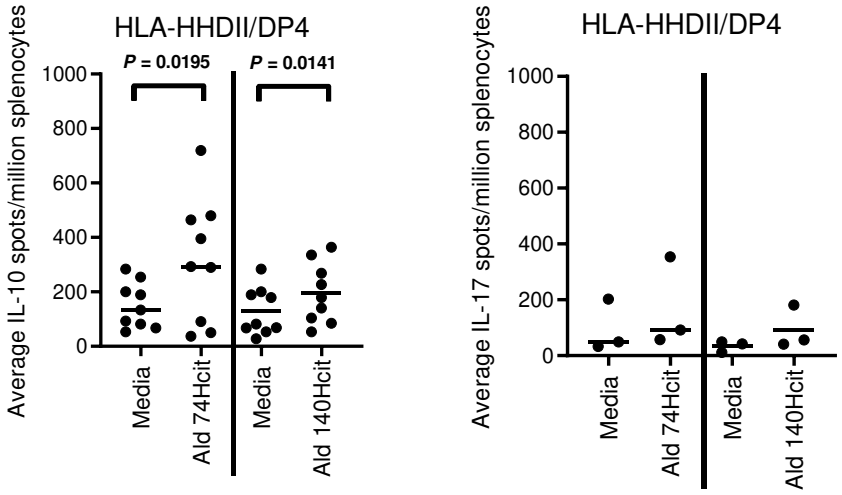
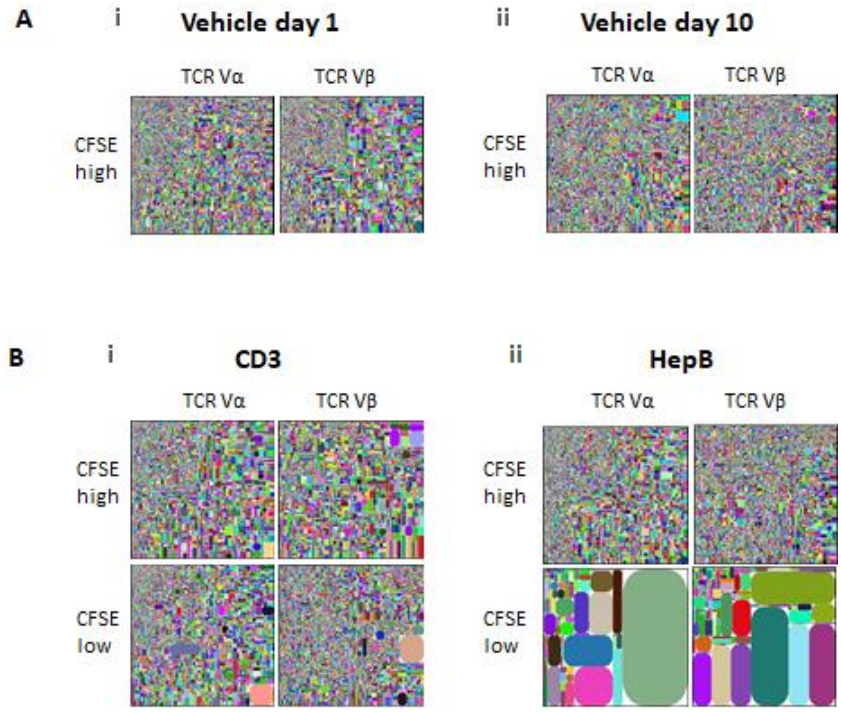


Figure S2. Homocitrullinated aldolase responses are Th1 phenotype.
Ex vivo ELISpot was used to assess IL-10 (A) and IL-17 (B) responses to aldolase peptides in HLA-HHDII/DP4 mice. For all studies mice were immunized with three doses of peptide with CpG/MPLA as an adjuvant and responses were assessed at day 21. Significant *P* values are shown for peptide compared to media only control stimulation.

Figure S3



C

PEPTIDE/CONDITION	SAMPLE	Diversity Index (D50) TRA	TRB
Vehicle	Day 1	16.1	15.3
	Day 10	9.2	18.2
CD3	$CD4^+$ $CFSE^{High}$	8.3	8.1
	$CD4^+$ $CFSE^{Low}$	7.9	13.8
HepB	$CD4^+$ $CFSE^{High}$	12.2	16.3
	$CD4^+$ $CFSE^{Low}$	0.4	0.4

Figure S3. TCRα and β CDR3 tree maps for non-proliferating vehicle (media) only control (A), proliferating and non-proliferating CD4+ cells in response to anti-CD3 polyclonal stimulus and stimulation with Hepatitis B peptide epitope (B). Diversity index was calculated for each condition (C).

Figure S4

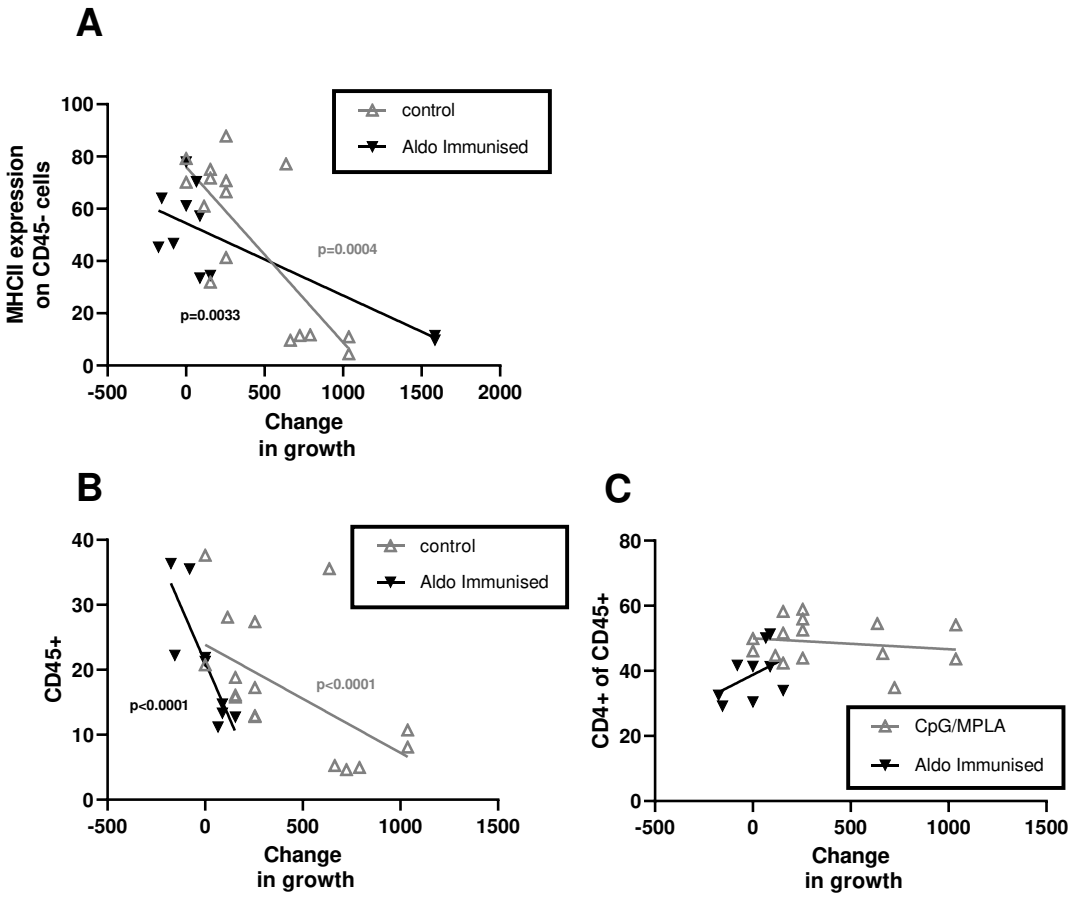


Figure S4. Correlation of change in tumour volume between time of immunization and tumour analysis and MHCII expression on CD45- cells (A) or CD45 (B) and CD4 (C) cell infiltrates. MHCII expression on CD45- cells includes all tumors, whereas the two immunized mice which showed no regression were excluded from further analysis due to lack of immune response as shown by minimal MHCII expression. Correlation was assessed by linear regression analysis and significant *P* values are shown.

Figure S5

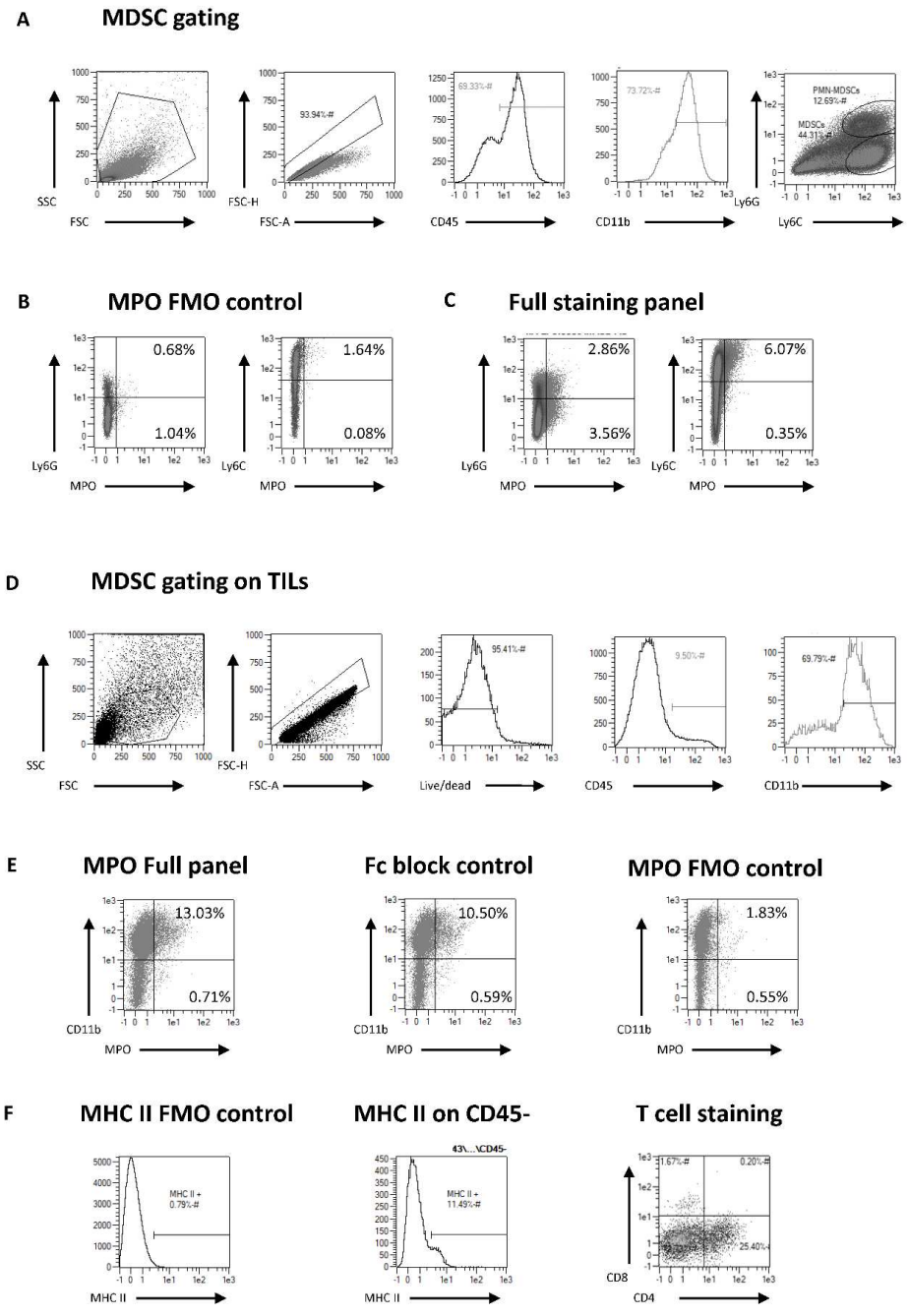

A
MDSC gating
SSC
FSC
FSC-H
FSC-A
93.94%-#
CD45
69.33%-#
CD11b
73.72%-#
Ly6G
Ly6C
PMN-MDSCs
12.69%-#
MDSCs
44.31%-#
B
MPO FMO control
Ly6G
MPO
0.68%
1.04%
Ly6C
1.64%
0.08%
C
Full staining panel
2.86%
3.56%
6.07%
0.35%
D
MDSC gating on TILs
SSC
FSC
FSC-H
FSC-A
Live/dead
95.41%-#
CD45
9.50%-#
CD11b
69.79%-#
E
MPO Full panel
CD11b
MPO
13.03%
0.71%
Fc block control
10.50%
0.59%
MPO FMO control
1.83%
0.55%
F
MHC II FMO control
MHC II
MHC II +
0.79%-#
MHC II on CD45-
43\...\CD45-
MHC II +
11.49%-#
T cell staining
CD8
CD4
11.67%-#
0.20%-#
25.40%-#

Figure S5. Example gating strategy for MDSCs
For MDSC analysis sequential gating was performed as shown (A), examples taken from *in vitro* grown MDSCs. FSC/SSC gated cells were doublet excluded using FSC-A/FSC-H gating. Events were then gated to include CD45+ then CD11b+ cells. MDSC populations were identified based on the combination of Ly6G and Ly6C. Fluorescent minus one (FMO) control shows Ly6C or Ly6G staining versus MPO on CD45/CD11b gated cells from MPO FMO (B) or fully stained samples (C). When analyzing TILs gating strategy was adapted to include live/dead gating (D) and Fc blocking was performed to confirm the specificity of the staining (E). Live/CD45- cells were assessed for MHC II staining compared to FMO control, and live/CD45+ cells for the presences of CD4 and CD8 (F).

Table S1. Peptide sequences.

Epitope	Sequence Human (Mouse)	IEDB predicted binding scores for wild type proteins			WT DP4 predicted core regions
		DP4	DR1	DR4	
Aldolase					
Aldo74-93^{Hcit}	IGGVILFHETLYQ-**hcit**-ADDGRP	2.51	33.91	30.22	FHETLYQ**K**A, LFHETLYQ**K**
Aldo140-157^{Hcit}	**hcit**-DGADFA-**hcit**-WRCVL-**hcit**-IGEH	22.1	59.86	49.8	FA**K**WRCVL**K**, A**K**WRCVL**K**I, DFA**K**WRCVL
Aldo217-235^{Hcit}	LSDHHI(V)YLEGTLL-**Hcit**-PNMVT	4.92	13.4	4.87	YLEGTLL**K**P, IYLEGTLL**K**
Aldo238-256^{Hcit}	HACTQ-**hcit**-FSH(N)EEIAMATVTA	13.18	46.73	4.11	TQ**K**FSHEEI, **K**FSHEEIAM
Aldo289-307^{Hcit}	**Hcit**-CPLL-**Hcit**-PWALTFSYGRALQ	14.38	24.99	11.15	**K**PWALTFSY, LL**K**PWALTF
BIP					
BIP328-346^{Hcit}	EELNMDLFRSTM-**hcit**-PVQ-**hcit**-VL	19.99	28.99	6.65	LFRSTM**K**PV, FRSTM**K**PVQ
Bip562-579^{Hcit}	RNELESYAYSL-**hcit**-NQIGD-**hcit**	21.01	32.43	12.81	LESYAYSL**K**, YAYSL**K**NQI
Enolase					
Eno156-176^{Hcit}	GSHAGN-**hcit**-LAMQEFMILPVGAA(S)	1.2	34.28	12.67	**K**LAMQEFMI, HAGN**K**LAMQ
Eno400-420^{Hcit}	RSERLA-**Hcit**-YNQLLRIEEELGS	11.8	42.27	23.15	RLA**K**YNQLL, A**K**YNQLLRI
Cyk8					
Cyk371-388^{Hcit}	LREYQELMNV-**hcit**-LALDIEI	24.36	6.74	4.42	LMNV**K**LALD, ELMNV**K**LAL
HepB					
HepB181-192	GFFLLTRILTIPQ	7.01	5.6	1.2	FLLTRILTI

Analysis is based on human sequences.

Where human and mouse sequences are not homologous, amino acids found in mouse sequences are shown in parenthesis.

Table S2. Healthy donor and patient data.
Healthy Donor

Healthy Donors

ID	Smoking status	HLA type							
BD0001	Non-smoker	HLA A: *02, *32	HLA B *08,*44	HLA C: *04, *07	HLA DRB1: *03, *07	HLA DRB3: *01	HLA DRB4: *01	HLA DQA1: *02, *05 HLA DQB1: *01	HLA DPB1: *04
BD002	Non-smoker	HLA A: *02, *29	HLA B *44,*51		HLA DRB1: *04, *13	HLA DRB3: *01	HLA DRB4: *01	HLA DQB1: *03, *06	HLA DPB1: *04
BD0007	Non-smoker	HLA A: *01, *32	HLA B *08, *15	HLA C: *07	HLA DRB1: *03, *13	HLA DRB3: *01, *03	-	HLA DQB1: *02, *06	HLA DPB1: *04
BD0010	Non-smoker	HLA A: *02, *11	HLA B *40, *44	HLA C: *03, *16	HLA DRB1: *13, *16	HLA DRB3: *03	HLA DRB5: *02	HLA DQB1: *05, *06	HLA DPB1: *04, *10
BD0014	N/A	N/A							
BD0015	Non-smoker	HLA A: *03, *24	HLA B: *07, *15	HLA C: *03, *07	HLA DRB1: *04, *15	HLA DRB4: *01	HLA DRB5: *01	HLA DQB1: *03, *06	HLA DPB1: *04
BD0016	Ex-smoker	HLA A: *01, *02	HLA B: *8, *44	HLA C: *05, *07	HLA DRB1: *03, *15	HLA DRB5: *01	HLA DRB3: *02	HLA DQB1: *02, *06	HLA DPB1: *01, *04
BD0017	Non-smoker	N/A							
BD0022	Non-smoker	HLA A: *01, *02	HLA B: *35, *50	HLA C: *06, *12	HLA DRB1: *04, *07	-	-	HLA DQB1: *02, *03	HLA DPB1: *02, *04
BD0025	Ex-smoker	HLA A *02, *29	HLA B *07, *57	HLA C: *06, *07	HLA DRB1: *01, *07	HLA DRB4: *01	-	HLA DQB1: *03, *05	HLA DPB1: *03, *13
BD0038	Non-smoker	HLA A: *26, *33	HLA B: *40, *58	HLA C: *03	HLA DRB1: *09, *11	HLA DRB3: *02	HLA DRB4: *01	HLA DQB1: *03	HLA DPB1: *04, *05
BD0041	Ex-smoker	HLA A: *01, *24	HLA B: *07, *40	HLA C: *03, *07	HLA DRB1: *04, *11	-	-	HLA DQB1: *03	HLA DPB1: *02, *04
BD0044	Non-smoker	N/A							
BD0051	Non-smoker	HLA A: *11, *68	HLA B: *07, *15	HLA C: *05, *07	HLA DRB1: *11, *15	-	-	HLA DQB1: *06	HLA DPB1: *04, *19
BD0050	Non-smoker	HLA A: *24, *26	HLA B *35, *34	HLA C: *04, *06	HLA DRB1: *07, *11	HLA DRB3: *02	HLA DRB4: *01	HLA DQA1: *02, *05 HLA DQB1: *02, *03	HLA DPB1: *02, *04
BD0095	Non-smoker	N/A							
BD0150	Non-smoker	N/A							

Cancer patient

ID	Smoking status	Indication/Treatment
LG6	Ex-smoker	Lung adenocarcinoma/currently none, previous chemotherapy and checkpoint inhibitors
LG8	Ex-smoker	Lung adenocarcinoma/ Tyrosine Kinase inhibitor and Steroid
LG9	Ex-smoker	Lung adenocarcinoma/ Chemotherapy
LG10	Smoker	Lung adenocarcinoma/ Checkpoint inhibitor
LG12	Ex-smoker	SCLC/Chemotherapy
LG18	Ex-smoker	Lung adenocarcinoma/treatment naïve
LG19	Ex-smoker	Lung adenocacinoma/ checkpoint inhibitor
OV19	Unknown	Ovarian serous adenocarcinoma (low grade)/ post surgery and pre systemic treatment
OV21	Unknown	Ovarian serous adenocarcinoma/ Finished chemotherapy
OV22	Unknown	Peritoneal adenocarcinoma/ treatment naïve
BR7	Unknown	Triple negative breast cacner

Table S3. Details of antibodies used.

Antibody target	Target species	Format	Clone	Supplier	Dilution
CD45	Mouse	eFlour 450	30-F11	Thermofisher	1/50
CD11b	Mouse	PE-Cy7	M1/70	Thermofisher	1/50
Ly6C	Mouse	APC	HK1.4	Thermofisher	1/50
Ly6G	Mouse	FITC	RB6-8C5	Thermofisher	1/50
MPO	Mouse	PE	8F4	HyCult Biotech	1/10
Homocitrulline	N/A	N/A	Polyclonal	Cayman Chemical	1/50
Secondary	Rabbit	Alexa-647	Polyclonal	Abcam	1/1000
CD4	Human	Pe-Cy5	RMC 4-5	Thermofisher	1/50
CD8	Human	eFlour 450	53-6.7	Thermofisher	1/50
HLA-DR/DP/DQ	Human	APC-Vio770	REA332	Miltenyi	1/50
Ly6G	Mouse	purified	1A8	Biolegend	N/A
Ly6C	Mouse	purified	Monts	BioXcell	N/A
CD4	Human	purified	RPA-T4	Abcam	N/A
CD4	Human	purified	OKT-4	BioXcell	N/A
CD8	Mouse	purified	2.4	BioXcell	N/A
Live/dead	N/A	Zombie Yellow	N/A	Biolegend	1/50
Fc block	Mouse	N/A	N/A	Miltenyi	1/50