An engineering approach to decode immune responses

Bresser, K.

Citation

Bresser, K. (2023, November 15). *An engineering approach to decode immune responses*. Retrieved from https://hdl.handle.net/1887/3663147

Chapter 2.
A mouse model that is immunologically tolerant to reporter and modifier proteins

OR: "How we tricked a mouse into believing it's a jellyfish."
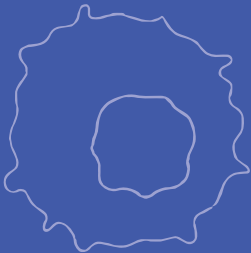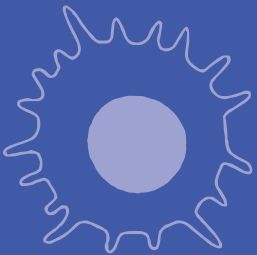
Kaspar Bresser[1,*], Feline E Dijkgraaf[1,*], Colin EJ Pritchard[2], Ivo J Huijbers[2], Ji-Ying Song[3], Jan C Rohr[4], Ferenc A Scheeren[5] and Ton N Schumacher[1,6]

(1) Division of Molecular Oncology & Immunology, Oncode Institute, The Netherlands Cancer Institute, Amsterdam, The Netherlands
(2) Mouse Clinic for Cancer and Aging research (MCCA) Transgenic Facility, The Netherlands Cancer Institute, Amsterdam, The Netherlands
(3) Animal Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
(4) Center for Chronic Immunodeficiency, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany; Center for Pediatrics and Adolescent Medicine, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany.
(5) Department of Medical Oncology, Leiden University Medical Center, Leiden, The Netherlands.
(6) Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden, The Netherlands.

(*) These authors contributed equally to this work

Abstract

Reporter proteins have become an indispensable tool in biomedical research. However, exogenous introduction of these reporters into mice poses a risk of rejection by the immune system. Here we describe the generation, validation and application of a multiple reporter protein tolerant 'Tol' mouse model that constitutively expresses an assembly of shuffled reporter proteins from a single open reading frame. We demonstrate that expression of the *Tol* transgene results in the deletion of CD8+ T cells specific for a model epitope, and substantially improves engraftment of reporter-gene transduced T cells. The Tol strain provides a valuable mouse model for cell transfer and viral-mediated gene transfer studies, and serves as a methodological example for the generation of poly-tolerant mouse strains.

Introduction

Following the advent of standardized genetic editing techniques, researchers have isolated a large collection of reporter and modifier proteins (RPs and MPs, respectively) from a variety of species that have since been instrumental to characterize a wide range of biological processes[1]. For example, RPs are frequently used to tag endogenous proteins or to track the behavior of individual cells *in vivo*[2,3]. In addition, MPs (e.g. Cre recombinase and Cas9) are commonly applied to influence *in vivo* cell behavior through the induction of defined genetic alterations[4]. However, as RPs and MPs are almost invariably derived from non-mammalian species, exogenous introduction of these proteins into immunocompetent mice poses the risk of immunological rejection by host-derived T cells. In line with this, multiple cases of immunological rejection of cells expressing RPs, such as firefly luciferase and eGFP, have been reported[5-9]. Moreover, major histocompatibility complex (MHC) class I-restricted epitopes of luciferase and GFP have been identified, underlining their capacity to induce CD8+ T cell responses[5-8]. Even in the absence of complete immunological rejection of RP- or MP-modified cells, experimental outcomes may potentially be subtly biased through the action of such undesirable immune responses.

To overcome this problem, we set out to engineer a transgene that—once introduced into the genome of a mouse model of choice—prevents the generation of immune responses against RP- or MP-modified cells through the physiological self-tolerance mechanisms. Specifically, CD4+ and CD8+ T cells that carry a self-reactive T cell receptor (TCR) may be inactivated or deleted through multiple mechanisms. First, during their maturation in the thymus, T cells that carry a self-reactive TCR are deleted through negative selection. Second, auto-reactive T cells that escape deletion through central tolerance are kept in check by a mechanism referred to as peripheral tolerance that requires antigen encounter outside of the primary lymphoid organs[9]. In addition, self-reactive CD4+ T cells may develop into regulatory T cells, and play a crucial role in such peripheral tolerance[10]. Thus, in order to generate a poly-tolerant mouse model, constitutive expression of foreign antigens in both the thymus and throughout peripheral tissues would be preferred. However, organism-wide
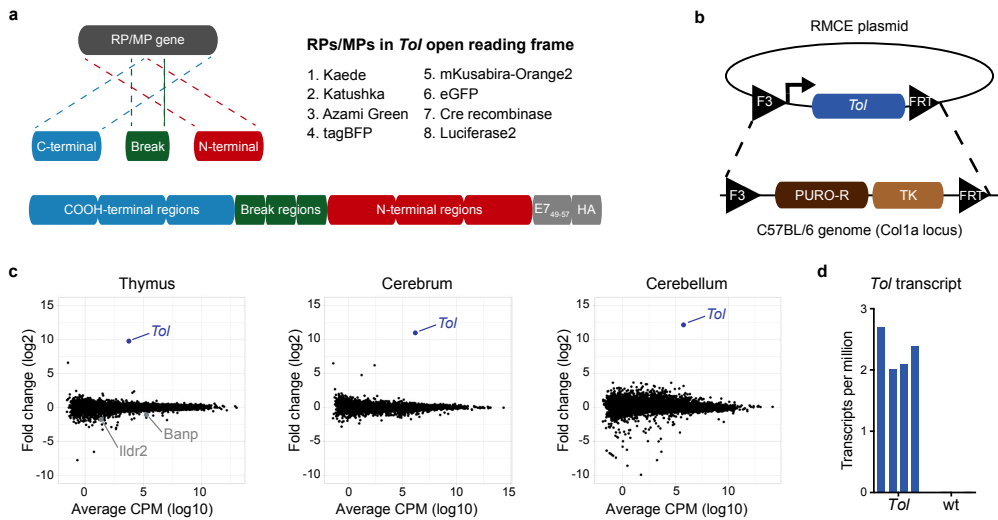
Fig. 1. Generation and validation of the Tol mouse model. a, Top left: cartoon depicting the strategy for reporter gene fragmentation, applied for each RP/MP gene included in the *Tol* ORF. Top right: reporter proteins included in the *Tol* ORF. Bottom: schematic overview of gene fragment placement in the *Tol* ORF. Note that the *Tol* ORF encodes the HPV E7$_{49-57}$ epitope in the COOH-terminal region of the chimeric Tol protein. **b**, Cartoon depicting the targeting strategy of the *Tol* ORF into the *Col1a1* locus through recombinase-mediated cassette exchange. **c**, Differential gene expression analysis of indicated organs of Tol mice relative to WT mice. Scatterplots indicate average log2 fold changes and average counts per million. Genes with a statistically significant change in expression in Tol mice are indicated in grey, and the *Tol* transcript is indicated in blue. **d**, Expression levels of the *Tol* transcript in thymi of Tol and WT mice. Data shown (**c-d**) is aggregated from n=4 (Tol) and n=4 (WT) mice. RP = reporter protein, MP = modifier protein, RMCE = recombinase-mediated cassette exchange, PURO-R = puromycin resistance gene, TK = thymidine kinase gene, CPM = counts per million.

expression of proteins such as eGFP or Cre is incompatible with their intended use as cell-type specific reporters or modifiers.

Results

Generation and validation of the 'Tol' mouse model

In order to engineer a multiple reporter protein tolerant mouse model, we designed a large chimeric open reading frame (ORF) that encodes 6 fluorescent proteins, firefly luciferase, and Cre-recombinase, all in a scrambled format. Specifically, in order to prevent functional expression of the introduced proteins, each individual gene was split into two fragments, and the resulting set of 16 gene fragments was subsequently assembled in a scrambled order. In addition, to ensure tolerance toward potential T cell epitopes present at the gene breakpoints, the 60bp region surrounding each split site was added (**Fig. 1a**). Finally, to be able to test whether tolerance was induced against epitopes throughout the artificial protein, a control CD8+ T cell epitope derived from the HPV E7 protein (HPV E7$_{49-57}$) was placed at the COOH-terminus. This resulted in a 7,359 bp chimeric '*Tol*' ORF, encoding a protein of ± 275 kDa that covers seven RP and one MP (**Supplementary Fig. 1**).

Recognition of RP- and MP-derived epitopes is dependent on the host MHC haplotype[11,12], and risk of immunological rejection of cells modified with individual RPs and MPs thus varies between mouse strains. To test if the *Tol* cassette could be applied to induce tolerance against proteins of

interest, we set out to assess its functionality in the C57BL/6 strain, a widely used mouse strain in immunological research. In addition, MHC binding predictions indicated that the Tol protein contains 78 potential MHC ligands predicted to bind with high affinity (i.e. netMHC4.0 percentile rank <1%) to the C57BL/6 MHC haplotypes (**Supplementary Data 1**). To establish the transgenic C57BL/6 strain, the *Tol* gene was targeted to the *Col1a1* locus of embryonic stem cells (ESCs) via recombinase mediated cassette exchange (**Fig. 1b**)[13,14]. Successfully modified ESCs were then injected into blastocysts and transferred to pseudopregnant foster mice. The resulting F0 generation was tested for transgene presence (**Supplementary Fig. 2a**), and mice with sufficient chimerism were bred to obtain experimental cohorts of heterozygous ('Tol') and wild-type littermate controls ('WT').

Due to its large size and scrambled design, the Tol protein is expected to be unfolded. Cellular stress induced by accumulation of unfolded proteins has previously been shown to be associated with a variety of pathologies and neurological disorders[15]. To test whether expression of the *Tol* transgene would affect murine development, histopathological evaluation of tissues from Tol and WT mice was performed, revealing no evidence for transgene-induced pathology (**Supplementary Table 1**). As a second test for cellular stress, or potential deregulation of other pathways, that may be induced by *Tol* expression, RNA-sequencing was performed on cerebral and cerebellar tissue of Tol and WT animals. In addition, transcriptional activity in thymic tissue, the site of T cell deletion, was compared between Tol-transgenic and WT mice. Hierarchical clustering analysis of the top expressed genes across tissues showed that WT and Tol mice did not typically cluster separately, suggesting that the *Tol* transgene did not induce major transcriptional changes (**Supplementary Fig. 2b**). Subsequent differential gene expression analysis likewise revealed no significant differentially expressed genes, aside from two marginally downregulated genes in the thymus (**Fig. 1c** and **Supplementary Data 2**). Importantly, the *Tol* transcript could be detected in all organs tested (**Fig. 1c, d**), and aligned reads spanned the entire transgene, demonstrating that the full-length ORF was transcribed (**Supplementary Fig. 2c**).
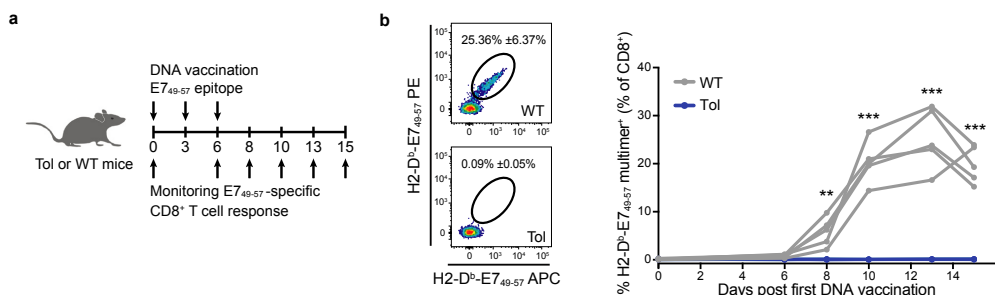

Fig. 2. Complete loss of E7$_{49-57}$ immunogenicity in Tol mice. a, Schematic overview of vaccination and monitoring strategy. **b,** Left: Representative flow cytometry plots depicting E7$_{49-57}$-specific CD8$^+$ T cells at day 13 post DNA vaccination in WT and Tol mice. Right: E7$_{49-57}$-specific CD8$^+$ T cell response in blood of WT (gray) and Tol (blue) mice at the indicated timepoints post first DNA vaccination. Lines indicate individual mice. Cells are gated on IR-dye$^-$CD8$^+$ lymphocytes. Depicted data are representative of two independent experiments, including at least 5 mice per group. *P* values were determined by Repeated Measures two-way ANOVA with Sidak's multiple comparisons test to test the significance at each time point. *P < 0.05; **P < 0.01; ***P < 0.001.
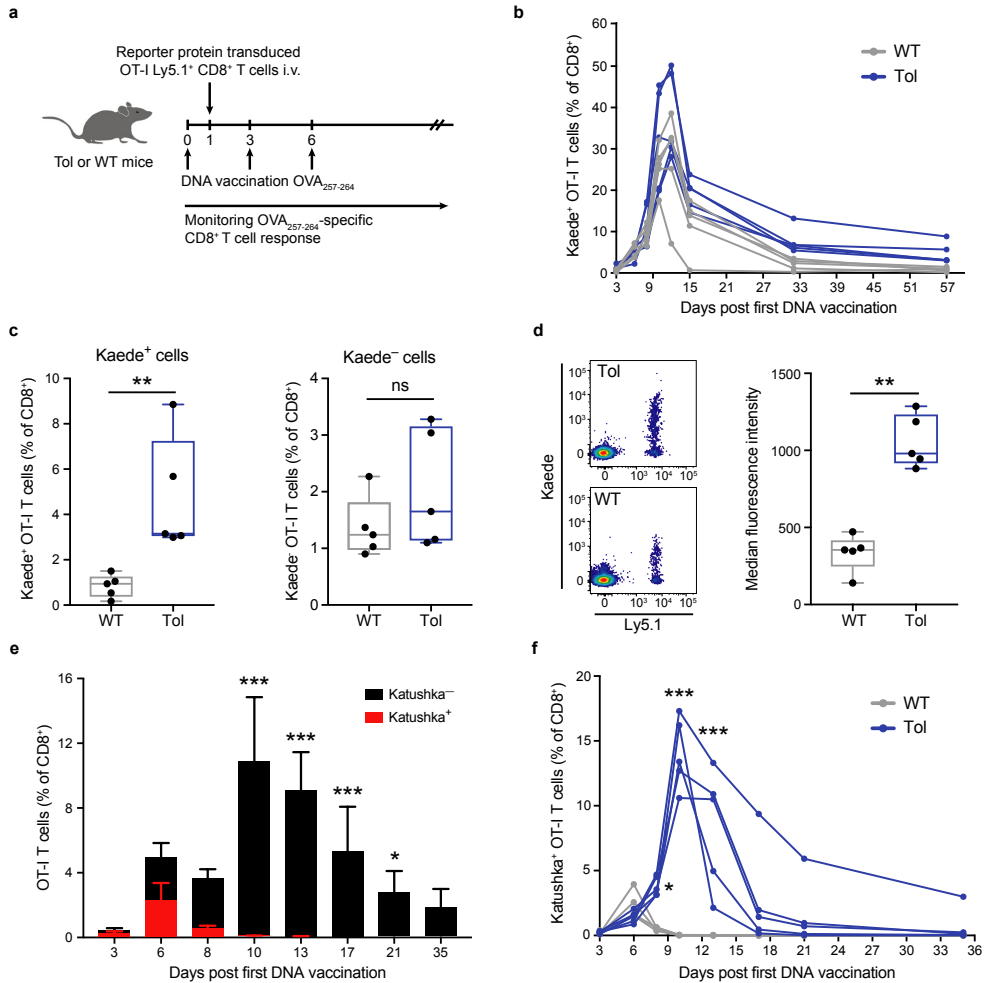
Fig. 3. Enhanced engraftment of Kaede and Katushka expressing cells in Tol mice. a, Schematic overview of experimental design. **b**, Percentage of Kaede+ OT-I T cells in the blood of WT (gray) and Tol (blue) mice at the indicated timepoints post first DNA vaccination. **c**, Percentage of Kaede+ (left) or Kaede- (right) OT-I cells in WT and Tol mice at day 57 after first DNA vaccination. **d,** Left: representative flow cytometry plots depicting Kaede expression in transferred (Ly5.1+) cells at day 57 after first DNA vaccination. Right: median fluorescence intensity (MFI) of Kaede+ OT-I cells in WT and Tol mice at day 57 post first DNA vaccination. **e,** Percentage of Katushka+ (red) and Katushka- (black) OT-I cells in WT animals at the indicated timepoints after first DNA vaccination. **f**, Percentage Katushka+ OT-I cells in WT and Tol mice at the indicated timepoints after first DNA vaccination. Cells (**b-f**) are gated on IR-dye−CD8+ lymphocytes, and plots depict n=5 (Tol) and n=5 (WT) mice. Depicted data are representative of one (**b-d**) or two (**e-f**) independent experiments. Box-plots (**c**, **d**) depict interquartile range, with whiskers representing minimum and maximum values. Dots (**c**, **d**) and lines (**b**, **f**) indicate individual mice. Error bars (**e**) represent SD. *P* values were determined by either two-tailed Mann-Whitney signed rank test (**c**, **d**) or Repeated Measures two-way ANOVA with Sidak's multiple comparisons test (**e**, **f**). *P < 0.05; **P < 0.01; ***P < 0.001, ns = not significant.

Immunological tolerance toward the E7$_{49-57}$ epitope in Tol mice

Having observed the presence of the full-length *Tol* transcript in thymic tissue of Tol mice, we next investigated whether T cell tolerance toward epitopes encoded by the transcript was successfully induced. To this purpose, WT and Tol animals were vaccinated with a DNA vector encoding the

HPV E7$_{49-57}$ epitope, and the development of CD8$^+$ T cell responses against the E7$_{49-57}$ epitope was subsequently monitored in peripheral blood by peptide-MHC multimer staining (**Fig. 2a**). Following DNA vaccination, a large population of E7$_{49-57}$-specific CD8$^+$ T cells was detected in blood of WT mice, with peak E7$_{49-57}$-specific CD8$^+$ T cell frequencies up to 30% of the total CD8$^+$ T cell population (**Fig. 2b, c**). In stark contrast, no E7$_{49-57}$-specific CD8$^+$ T cells were detected above background in *Tol*-transgene positive animals at any timepoint post vaccination (**Fig. 2b, c**). These data indicate strict tolerance of Tol mice toward a T cell epitope that is present at the 3' of the *Tol* ORF, indicating full-length translation of the protein and successful processing by the antigen presentation machinery.

Enhanced engraftment of fluorescent protein expressing OT-I T cells in Tol animals

To assess whether expression of the shuffled reporter gene fragments in Tol mice resulted in tolerance toward the parental non-shuffled reporter proteins, CD8$^+$ Ly5.1$^+$ OT-I T cells—specific for the OVA$_{257-364}$ peptide—were transduced with a retrovirus encoding the green-to-red photo-switchable reporter protein Kaede. The resulting Kaede$^+$ and Kaede$^-$ OT-I T cells were then transferred into Tol and WT mice, and OT-I T cell frequencies were boosted by administration of an OVA$_{257-364}$-epitope encoding DNA vaccine (**Fig. 3a**). In the first days following adoptive cell transfer, the numbers of Kaede$^+$ T cells increased rapidly and similarly in both WT and Tol mice (day 10, ratio Tol/WT = 1.25, p = 0.421, **Fig. 3b**). Notably, at later time points Kaede$^+$ OT-I T cells formed a readily detectable memory T cell population in Tol mice, but declined to near undetectable numbers in WT mice (day 57, ratio Tol/WT = 5.6, p = 0.008, **Fig. 3b, c**), suggestive of immunological rejection of the introduced Kaede$^+$ cells. As a control, Kaede$^-$ T cells were detected in comparable numbers in WT and Tol mice (day 57, ratio Tol/WT = 1.5, p = 0.309). In addition, remaining Kaede$^+$ T cells in WT mice showed a significantly lower expression of the Kaede protein than Kaede$^+$ T cells present in Tol mice (p = 0.008, **Fig. 3d**), indicating that immune-mediated rejection can both lead to a reduction in cell quantities and selection of cells with lower transgene expression. To assess whether expression of the *Tol* transgene resulted in long-term tolerance toward Kaede, recipients of Kaede$^+$ and Kaede$^-$ OT-I T cells were vaccinated and then challenged by a secondary vaccination 155 days post adoptive cell transfer (**Supplementary Fig. 3a**). In Tol mice, Kaede$^+$ cells again expanded to large numbers after receiving a secondary stimulus, emphasizing ongoing tolerance toward the transferred cells (**Supplementary Fig. 3b**). In contrast, Kaede$^+$ T cells exhibited a stunted proliferative burst, and expressed only low levels of the fluorescent protein in WT animals (**Supplementary Fig. 3c**). The inability to mount a robust secondary response was directly related to Kaede expression, as Kaede$^-$ T cells were able to expand in WT mice (**Supplementary Fig. 3d**).

We next examined the engraftment potential of OT-I T cells that expressed a third foreign entity, the rapid-folding red-fluorescent protein Katushka, in Tol and WT recipient mice. Adoptive transfer of Katushka$^+$ cells induced a vigorous anti-fluorochrome immune response in WT mice, as reflected by the rapid disappearance of Katushka$^+$ OT-I T cells, but not Katushka$^-$ T cells (ratio Kat$^-$/Kat$^+$ cells: day 6 = 1.2; day 10 = 385.4, **Fig. 3e**). Importantly, the observed clearance of Katushka-expressing

cells was abrogated by *Tol* expression, as shown by the presence of an approximately 500-fold larger pool of Katushka⁺ T cells in Tol mice as compared to WT littermates at day 10 post first DNA vaccination ($p = 0.008$, **Fig. 3f**).

To comprehensively evaluate whether cells expressing any of the fluorescent proteins (FPs) encoded by the *Tol* transgene could engraft in Tol mice, we concurrently transferred OT-I T cells that had been transduced with four different FPs (Katushka, BFP, AzamiGreen or mKO2) into either Tol or WT mice. In line with the prior analyses, Katushka⁺ OT-I T cells disappeared from the circulation of WT mice as early as day 10 post-transfer, whereas robust engraftment of Katushka expressing cells was observed in Tol mice (**Supplementary Fig 4**). Analysis of modified T cell frequencies for the other three FPs showed the presence of FP-modified cells in blood of both WT and Tol mice, with a trend toward enhanced engraftment in Tol mice (**Supplementary Fig 4**). Thus, cells modified with 5 different FPs (Kaede, Katushka, BFP, AzamiGreen or mKO2) stably engraft in Tol mice, whereas engraftment of cells expressing at least 2 of these transgenes is severely impaired in WT mice.

Discussion

Manipulation and detection of cellular pathways using modifier and reporter proteins forms a cornerstone of animal research; however, the xenogenic nature of these proteins introduces a substantial risk of recognition by the adaptive immune system. In this work, we describe an approach to fuse multiple RPs/MPs into a large shuffled ORF, thereby perturbing their functionality while maintaining all potentially immunogenic epitopes. Importantly, we show that this approach allows expression in the thymus plus peripheral tissues, that the encoded fusion protein is fully translated, and that this protein translation confers immunological tolerance toward all epitopes tested. Interestingly, comparison of the fate of Katushka⁺ and Kaede⁺ cells in WT mice underlines that immune recognition can either lead to the rapid clearance of transgene expressing cells (i.e. Katushka), or to the gradual selection of cells with lower transgene expression (i.e. Kaede). This second mode of immune rejection may go unnoticed in many experimental settings, thereby representing a hidden confounder. Engraftment of BFP, AzamiGreen and mKO2 transduced OT-I T cells was also observed but was only marginally improved relative to WT mice, suggesting that these proteins have a low immunogenicity in the C57BL/6 strain used in this study. As antigen presentation is MHC restricted, the value of the *Tol* transgene for these FPs may be more profound in other mouse strains, as e.g. also exemplified by the preferential immunogenicity of eGFP in Balb/c mice relative to C57BL/6 strains[6,16,17].

Our study provides a methodological framework through which shuffled transgenes are applied to avoid undesirable immune responses against exogenously introduced proteins. Both the *Tol* ORF and the C57BL/6 Tol strain used in this work are available upon reasonable request (from the corresponding author and the Netherlands Cancer Institute Transgenics Core Facility, respectively). Furthermore, creation of a resource of mouse strains that are tolerant for additional reporter and modifier proteins used in biomedical research would be of considerable value.

Methods

Generation of *Tol* open reading frame and vaccination plasmids

The *Tol* ORF was designed as described in Supplementary Fig. 1 and synthesized by Genscript. The full-length ORF was subcloned into a recombinase-mediated cassette exchange (RMCE) compatible vector, producing the pF3-CAG-Tol-FRT plasmid (Addgene ID: 141349). In the resultant vector, the Tol ORF was positioned downstream of the CAG promoter and flanked by F3 and FRT recombination sites. The entire F3-CAG-Tol-FRT cassette was sequence verified by Sanger sequencing. Katushka, Kaede, AzamiGreen, mKO2 and tagBFP ORFs were shuttled directly into the multiple cloning site of pMP71 via Gibson cloning to generate pMP71-Katushka, pMP71-Kaede, pMP71-AzamiGreen, pMP71-mKO2 and pMP71-tagBFP, respectively. Generation of pVAX- $E7_{49-67}$ has been described previously[22]. pVAX-SIINFEKL (Addgene ID: 141350) was generated by Gibson cloning to produce scarless fusions between SIINFEKL, several CD4+ T cell epitopes (HELP), and the KDEL-signal peptide.

Mice

C57BL/6-Ly5.1 and OT-I mice were obtained from Jackson Laboratories and crossed to obtain C57BL/6-Ly5.1-OT-I donor mice for adoptive transfer experiments. All animals were maintained and bred in the animal department of The Netherlands Cancer Institute and used for experimentation at 7-14 weeks. All animal experiments were approved by the Animal Welfare Committee of the NKI, in accordance with national guidelines.

Generation of Tol transgenic mice

The *Tol* expression cassette was introduced into a locus 3' to the *Col1a1* locus[14] using recombinase-mediated cassette exchange (RMCE)[18] In brief, a C57BL/6J ES cell line was derived from blastocysts[19] and an F3-Puro-deltaTK-FRT cassette was targeted into the ESCs by homologous recombination, as described[14]. The pF3-CAG-Tol-FRT plasmid was co-transfected with pCAGGS-FLPe into the B6J-RMCE ES cells, followed by selection in medium containing fialuridine. Clones were screened by PCR for correct and complete integration of the F3-Tol-FRT cassette into the *Col1a1* locus. Correctly modified ESCs were then injected into B6/NTAC blastocysts, and injected blastocysts were transferred into pseudopregnant B6CBAF1/JRj foster mice, as described previously[13,19]. Resulting chimeras were tested for chimerism by quantitative-PCR (Q-PCR). Positive animals (>0.4 RQ) were crossed with C57BL/6JRj mice (Janvier) to generate experimental cohorts. Presence of the *Tol* cassette was confirmed in experimental cohorts by PCR using the forward 5'-GGAAAGAATCACAACTTACG-3' and reverse 5'-AGAGCATTTCGGTTGAGGCC-3' primers. The Tol strain described in this communication is available upon request from the Netherlands Cancer Institute Transgenics Core Facility.

Histopathology

For histopathological analyses, 2 mm-thick hematoxylin-eosin stained sections were prepared from formalin-fixed, paraffin-embedded murine tissues such as skin, spleen, thymus, lymph nodes, liver, pancreas, gastrointestinal tract, heart, lung, kidneys, testes, ovaries, accessory sex glands, bone marrows (sternum and extremity), and muscles. Sections were evaluated and scored by an animal pathologist blinded to animal genotype.

RNA sequencing

RNA was extracted from the indicated frozen tissues using the RNeasy Mini Kit (Qiagen). Whole transcriptome sequencing samples were prepared with the TruSeq Stranded mRNA Kit (Illumina). Single-end 65 bp sequencing was performed on a HiSeq 2500 System (Illumina). Transcript abundance was calculated using Salmon 0.14.1[20], based on the GRCm38 transcriptome build. The *Tol* transcript sequence was added before read alignment. Differential gene expression analysis was performed using EdgeR 3.9[21].

Production of retroviral supernatants

Phoenix-E packaging cells were seeded at $1.4*10^6$ cells per 10 cm dish in IMDM supplemented with 8% FCS, glutamax, penicillin/streptomycin (Gibco). After 24 hours, medium was refreshed and cells were transfected using FuGene6 (Roche), following the manufacturer's protocol. 48 hours post transfection, viral supernatant was harvested and passed through a 0.22 μm filter (Sigma). Filtered retroviral supernatant was snap-frozen in liquid nitrogen and stored at -80 C.

Generation and adoptive transfer of Kaede, Katushka, BFP, AzamiGreen or mKO2 positive OT-I splenocytes

Spleens from C57BL/6;Ly5.1;OT-I mice were passed through 70 μm strainers to obtain a single cell suspension. Splenocytes were then seeded at $6*10^6$ cells/ml in RPMI 1640 supplemented with 8% FCS, glutamax, penicillin/streptomycin, pyruvate, non-essential amino acids, HEPES, β-mercapto-ethanol (Gibco), 2 μg/ml concanavalin A (Calbiochem) and 1 ng/ml IL-7 (Peprotech). After 48 hours, T cells were harvested and re-seeded on 24-well plates coated with Retronectin (Takara) at $1.5*10^6$ cells per well in RPMI 1640 supplemented with 8% FCS, glutamax, penicillin/streptomycin, pyruvate, non-essential amino acids, HEPES, β-mercapto-ethanol and 50 IU IL-2 (Novartis). Retrovirus containing supernatant was added and cells were transduced by spin-infection at 2,000 rpm for 90 minutes. After 24 hours, cells were harvested and washed twice in PBS. Subsequently, T cells were resuspended in HBSS (Gibco) and $1*10^6$ CD8+ cells were intravenously injected via the tail vein.

DNA vaccination

One day prior to vaccination with DNA encoding MHC-II class restricted helper epitopes[22], plus either the $OVA_{257-264}$ epitope ('Help-OVA') or the $E7_{49-67}$ epitope ('Help-E7'), hair of mice was removed from hind legs using Veet depilation cream (Reckitt Benckiser). Primary DNA vaccination was subsequently performed on day 0, 3 and 6, as previously described[23]. In brief, a droplet of 15 µl of a 2 µg/µl DNA solution in 10 mmol/L Tris pH 8.0 and 1 mmol/L EDTA pH 8.0 was applied on both the inside and outside of the leg, and vaccination was performed using a rotary tattoo device with a sterile disposable 9-needle bar (MT.DERM) oscillating at a frequency of 100 Hz for 1 min with a needle depth of 1 mm. For secondary vaccinations, mice received a single DNA tattoo with 20 µl of the 2 µg/µl plasmid solution on both the inside and outside of the leg, at >155 days after start of primary vaccination.

Analysis of antigen-specific T cell responses by flow cytometry

Blood samples were collected from the tail vein at the indicated time points. Erythrocytes were removed by treatment with NH_4Cl buffer, and cells were washed and then stained with anti-CD45.1-APC (A20, Thermo Fisher Scientific), anti-CD8-PerCp-Cy5.5 (eBioH35-17.2, Thermo Fisher Scientific), near-IR dye (Thermo Fisher Scientific) and indicated MHC multimers. MHC multimers were produced in-house by UV-induced ligand exchange and subsequent labeling with BV421, APC or PE (Thermo Fisher Scientific), as described previously[24]. Flow cytometry data were acquired on a Fortessa (BD Biosciences) and analyzed in FlowJo (version 10.4.2) according to the gating strategy shown in Supplementary Fig. 5.

Statistics and reproducibility

Statistical analyses were performed in Prism (GraphPad), comparing groups of mice with the two-tailed Mann-Whitney test. Results were regarded as statistically significant at a P-value of <0.05.

Data availability

Mouse cDNA sequences (GRCm38) used for Salmon alignment were downloaded from Ensembl (http://www.ensembl.org/biomart/martview). RNA-seq data have been deposited, and are available from the Gene Expression Omnibus under accession number GSE147757. All source data underlying the graphs and charts presented in the main figures are presented in Supplementary Data 3.

## Acknowledgements

Author contributions

K.B. generated the *Tol* targeting vector and other vectors, and performed monitoring of T cell responses. F.E.D. performed adoptive cell transfer and vaccination experiments, and performed monitoring of T cell responses. J.-Y.S. evaluated histopathological data. J.C.R., K.B., F.E.D. and F.A.S. contributed to design of the *Tol* ORF. I.J.H. and C.E.J.P. generated the genetically engineered Tol mouse model. K.B., F.E.D., F.A.S. and T.N.S. contributed to experimental design and prepared the manuscript with input of all co-authors.

References

1. Ghim, C.-M., Lee, S. K., Takayama, S. & Mitchell, R. J. The art of reporter proteins in science: past, present and future applications. BMB Reports 43, 451–460 (2010).
2. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. Science 263, 802–805 (1994).
3. Shaner, N. C. et al. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nature biotechnology 22, 1567–1572 (2004).
4. Orban, P. C., Chui, D. & Marth, J. D. Tissue- and site-specific DNA recombination in transgenic mice. PNAS 89, 6861–6865 (1992).
5. Davey, G. M. et al. Identification of a MHC I-restricted epitope of DsRed in C57BL/6 mice. Mol. Immunol. 53, 450–452 (2013).
6. Han, W. G. H., Unger, W. W. J. & Wauben, M. H. M. Identification of the immunodominant CTL epitope of EGFP in C57BL/6 mice. Gene Ther. 15, 700–701 (2008).
7. Gambotto, A. et al. Immunogenicity of enhanced green fluorescent protein (EGFP) in BALB/c mice: identification of an H2-Kd-restricted CTL epitope. Gene Ther. 7, 2036–2040 (2000).
8. Limberis, M. P., Bell, C. L. & Wilson, J. M. Identification of the murine firefly luciferase-specific CD8 T-cell epitopes. Gene Ther. 16, 441–447 (2009).
9. Xing, Y. & Hogquist, K. A. T-cell tolerance: central and peripheral. Cold Spring Harb Perspect Biol 4, a006957–a006957 (2012).
10. Wyss, L. et al. Affinity for self antigen selects Treg cells with distinct functional properties. Nature immunology 17, 1093–1101 (2016).
11. Stripecke, R. et al. Immune response to green fluorescent protein: implications for gene therapy. Gene Ther. 6, 1305–1312 (1999).
12. Sellers, R. S., Clifford, C. B., Treuting, P. M. & Brayton, C. Immunological variation between inbred laboratory mouse strains: points to consider in phenotyping genetically immunomodified mice. Vet. Pathol. 49, 32–43 (2012).
13. Huijbers, I. J. et al. Using the GEMM-ESC strategy to study gene function in mouse models. Nat Protoc 10, 1755–1785 (2015).
14. Beard, C., Hochedlinger, K., Plath, K., Wutz, A. & Jaenisch, R. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis 44, 23–28 (2006).
15. Hetz, C. & Saxena, S. ER stress and the unfolded protein response in neurodegeneration. Nat Rev Neurol 13, 477–491 (2017).
16. Aoyama, N. et al. Transgenic mice that accept Luciferase- or GFP-expressing syngeneic tumor cells at high efficiencies. Genes Cells 23, 580–589 (2018).
17. Skelton, D., Satake, N. & Kohn, D. B. The enhanced green fluorescent protein (eGFP) is minimally immunogenic in C57BL/6 mice. Gene Ther. 8, 1813–1814 (2001).
18. Prosser, H. M., Koike-Yusa, H., Cooper, J. D., Law, F. C. & Bradley, A. A resource of vectors and ES cells for targeted deletion of microRNAs in mice. Nature biotechnology 29, 840–845 (2011).
19. Huijbers, I. J. et al. Rapid target gene validation in complex cancer mouse models using re-derived embryonic stem cells. EMBO Mol Med 6, 212–225 (2014).
20. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. Nature methods 14, 417–419 (2017).
21. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140 (2010).
22. Ahrends, T. et al. CD27 Agonism Plus PD-1 Blockade Recapitulates CD4+ T-cell Help in Therapeutic Anticancer Vaccination. Cancer research 76, 2921–2931 (2016).
23. Bins, A. D. et al. A rapid and potent DNA vaccination strategy defined by in vivo monitoring of antigen expression. Nature medicine 11, 899–904 (2005).
24. Toebes, M., Rodenko, B., Ovaa, H. & Schumacher, T. N. M. Generation of peptide MHC class I monomers and multimers through ligand exchange. Curr Protoc Immunol Chapter 18, Unit 18.16–18.16.20 (2009).
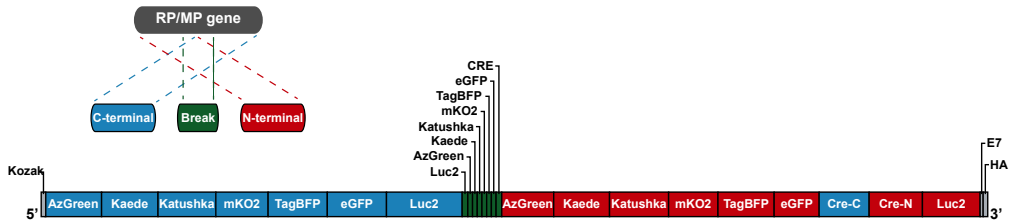
Supplementary Fig. 1. Design of the Tol open reading frame. Cartoon depiction of the Tol open reading frame (ORF), with the position of each gene fragment depicted. AzGreen = Azami Green, mKO2 = mKusabira Orange2, TagBFP = TagBlue fluorescent protein, eGFP = enhanced green fluorescent protein, Luc2 = luciferase2, E7 = HPV E749-57 epitope.
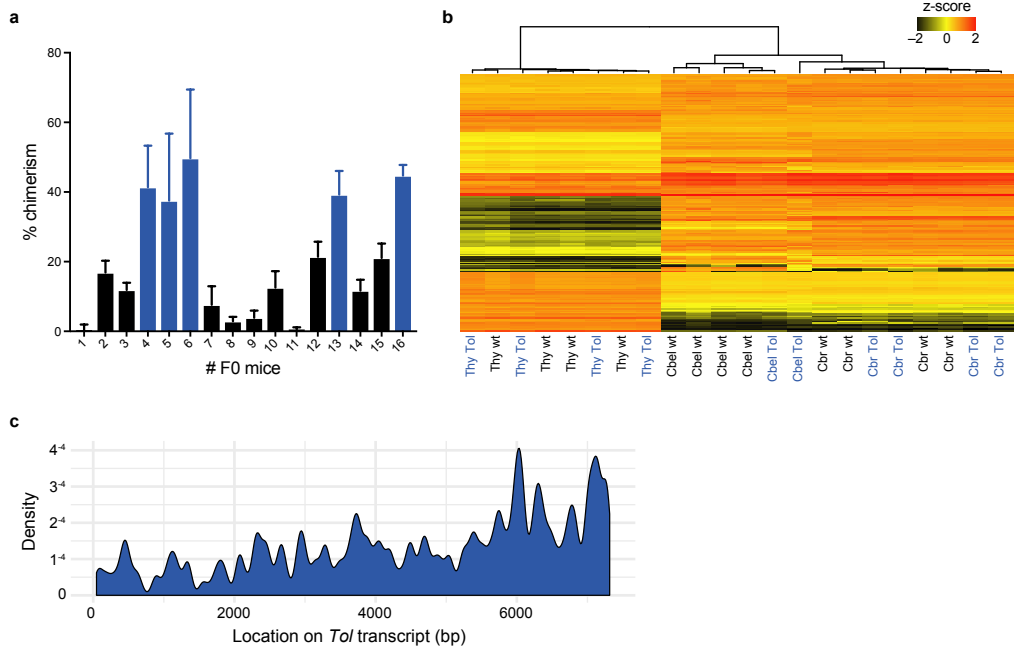

Supplementary Fig. 2. Validation of the Tol mouse model. a, Bar chart depicting the percentage of chimerism in F0 animals, as determined by quantitative PCR. Blue bars represent animals used for subsequent breeding. **b**, Hierarchical clustering analysis performed on RNA sequencing data from thymus, cerebrum and cerebellum of WT (black) and Tol (blue) mice. Complete-linkage clustering was performed on the top 1,000 most highly expressed genes in the dataset. Thy = thymus, Cbel = cerebellum, Cbr = cerebrum. **c**, Density of aligned reads along the Tol ORF. Data shown in (b-c) are representative of n=4 (Tol) and n=4 (WT) mice. PURO-R = puromycin resistance gene, TK = thymidine kinase gene, CPM = counts per million.
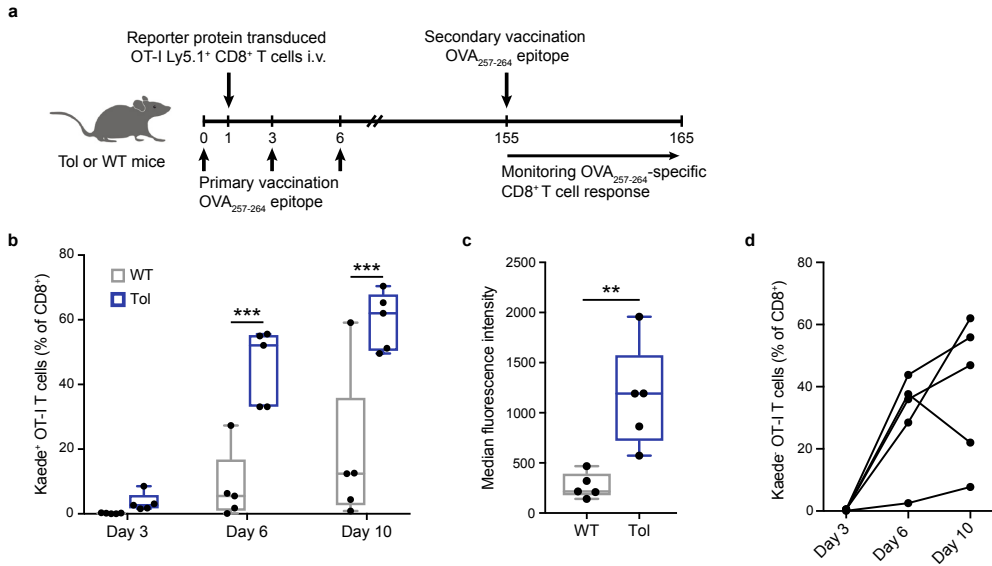
Supplementary Fig. 3. Enhanced recall potential of Kaede+ OT-I T cells in Tol mice. a, Schematic overview of adoptive cell transfer, vaccination and monitoring strategy. **b, c,** Percentage and MFI of Kaede+ OT-I T cells after secondary vaccination. Note that both the number of Kaede+ OT-I cells and the Kaede expression level are increased in Tol mice relative to WT mice. **d**, Percentage of Kaede- OT-I T cells after secondary vaccination. Cells (b-d) are gated on IR-dye–CD8+ lymphocytes, and data depict n=5 (Tol) and n=5 (WT) mice. Box-plots (b, c) depict interquartile range, with whiskers representing minimum and maximum values. Dots (b, c) and lines (d) indicate individual mice. P values were determined by either Repeated Measures two-way ANOVA with Sidak's multiple comparisons test (b) or two-tailed Mann-Whitney signed rank test (c). **P < 0.01; ***P < 0.001.
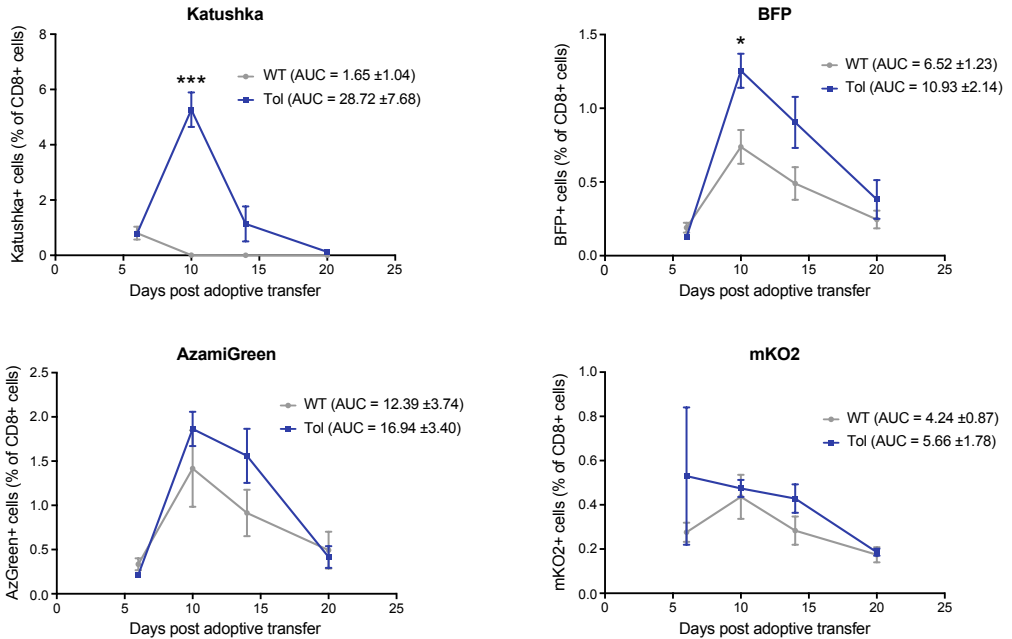
Supplementary Fig. 4. Analysis of engraftment of Katushka, BFP, AzamiGreen and mKO2 modified T cells in WT and Tol mice. Percentage FP+ OT-I T cells detected in the blood of WT and Tol mice at indicated time-points. Lines represent group means of Tol (n=5) and WT (n=7) mice, error bars depict standard errors of the mean. Mean area under the curve (AUC) is given for each group. P values were determined by Repeated Measures two-way ANOVA with Sidak's multiple comparisons test to test the significance at each time point. *P < 0.05; **P < 0.01; ***P < 0.001.
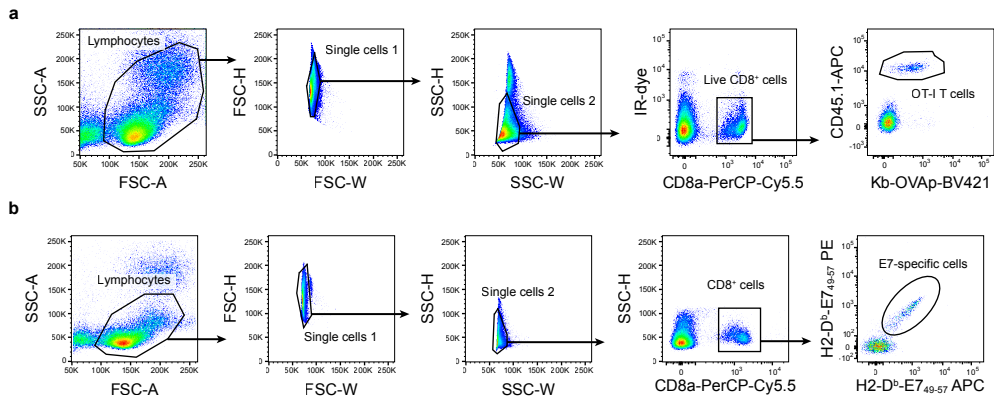
Supplementary Fig. 5. Gating strategy for the detection of antigen-specific T cell responses in blood samples by flow cytometry. a, Gating strategy applied to detect transferred OT-I T cell populations. **b,** Gating strategy applied to detect endogenous E7-specific T cell responses.