Efficient human-like antibody repertoire and hybridoma production in trans-chromosomic mice carrying megabase-sized human immunoglobulin loci

Hiroyuki Satofuka
Chromosome Engineering Research Center, Tottori University, Tottori 683-8503

Satoshi Abe
Tottori University

Takashi Moriwaki
Division of Genome and Cellular Functions, Department of Molecular and Cellular Biology, School of Life Science, Faculty of Medicine, Tottori University, Tottori 683-8503

Akane Okada
Division of Genome and Cellular Functions, Department of Molecular and Cellular Biology, School of Life Science, Faculty of Medicine, Tottori University, Tottori 683-8503

Kanako Kazuki
Tottori University

Hiroshi Tanaka
Trans chromosomics, Inc., Tottori 683-8503

Kyoutarou Yamazaki
Biomedical Science, Institute of Regenerative Medicine and Biofunction, Graduate School of Medical Science, Tottori University, Tottori 683-8503

Genki Hichiwa
Chromosome Engineering Research Center, Tottori University, Tottori 683-8503

Kayoko Morimoto
Trans chromosomics, Inc., Tottori 683-8503

Haruka Takayama
Trans chromosomics, Inc., Tottori 683-8503

Yuji Nakayama
Division of Radioisotope Science, Research Initiative Center, Organization for Research Initiative and Promotion, Tottori University, Tottori 683-8503

Shinya Hatano
Division of Host Defense, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812–8582

Yoshihiro Baba
Division of Host Defense, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812–8582

Mitsuo Oshimura
Tottori University
Kazuma Tomizuka
Department of Applied Life Sciences, Laboratory of Bioengineering, Tokyo University of Pharmacy and Life Sciences, Tokyo, 192-0392
Yasuhiro Kazuki (kazuki@tottori-u.ac.jp)
Tottori University

Article

Keywords: trans-chromosomic mice, human immunoglobulin, hybridoma production

Posted Date: December 2nd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-103241/v1

Version of Record: A version of this preprint was published at Nature Communications on April 5th, 2022. See the published version at https://doi.org/10.1038/s41467-022-29421-2.

Efficient human-like antibody repertoire and hybridoma production in trans-chromosomic mice carrying megabase-sized human immunoglobulin loci

Hiroyuki Satofuka[1], Satoshi Abe[2], Takashi Moriwaki[1,3], Akane Okada[3], Kanako Kazuki[1], Hiroshi Tanaka[2], Kyoutarou Yamazaki[4], Genki Hichiwa[1,3], Kayoko Morimoto[2], Haruka Takayama[2], Yuji Nakayama[5], Shinya Hatano[6], Yoshihiro Baba[6], Mitsuo Oshimura[2], Kazuma Tomizuka[7], and Yasuhiro Kazuki[1,3,4*].

[1]Chromosome Engineering Research Center, Tottori University, 86 Nishi-cho, Yonago, Tottori 683-8503, Japan.

[2]Transchromosomics Inc., 86 Nishi-cho, Yonago, Tottori 683-8503, Japan.

[3]Division of Genome and Cellular Functions, Department of Molecular and Cellular Biology, School of Life Science, Faculty of Medicine, Tottori University, 86 Nishi-cho, Yonago, Tottori 683-8503, Japan.

[4]Biomedical Science, Institute of Regenerative Medicine and Biofunction, Graduate School of Medical Science, Tottori University, 86 Nishi-cho, Yonago, Tottori 683-8503, Japan.

[5]Division of Radioisotope Science, Research Initiative Center, Organization for Research Initiative and Promotion, Tottori University, Yonago 683-8503, Japan,

[6]Division of Immunology and Genome Biology, Department of Molecular Genetics, Medical Institute of Bioregulation, Kyushu University, Higashi-ku, Fukuoka, 812-8582, Japan.

[7]Laboratory of Bioengineering, Tokyo University of Pharmacy and Life Sciences, 1432-1 Horinouchi, Hachioji, Tokyo, 192-0392, Japan.

These authors contributed equally: Hiroyuki Satofuka, Satoshi Abe and Takashi Moriwaki.

*Correspondence should be addressed to Y.K. (kazuki@tottori-u.ac.jp).

Trans-chromosomic (Tc) mice carrying mini-chromosomes with human immunoglobulin (Ig) loci have contributed to the development of fully human therapeutic monoclonal antibodies (Abs); however, we previously observed that mitotic instability of human mini-chromosomes in mice has limited the efficiency of hybridoma production. Here, we established a new generation of human Ab producing Tc mice (TC-mAb mice), which stably maintain a mouse-derived engineered chromosome containing the entire human Ig heavy and kappa chain loci in a mouse Ig knockout background. Comprehensive, high-throughput DNA sequencing revealed that the human Ig repertoire, including variable gene use, was well recapitulated in TC-mAb mice. Despite slightly altered B cell development and a delayed immune response, immunized TC-mAb mice exhibited more subsets of antigen-specific plasmablast and plasma cells compared with wild-type mice, leading to high efficiency hybridoma production. Thus, TC-mAb mice offer a valuable platform to obtain fully human therapeutic Abs and to elucidate the regulation of human Ig repertoire formation.

Introduction

In the last two decades, therapeutic antibodies (Abs) have emerged as a highly effective and fast-growing pharmaceutical option. Genetic engineering has been used to develop various methods to overcome the immunogenicity of rodent monoclonal antibodies (mAbs) in humans, which was a critical issue in early clinical trials. Among them, transgenic animals designed to express the human Ab repertoire are widely recognized for their ability to generate fully human mAbs[1,2]. The first generation of humanized immunoglobulin (Ig) mice was developed in the 1990s by random integration of DNA segments containing partial Ig heavy and light chain loci into the chromosomes of endogenous Ig knockout mice[2,3]. These transgenic-knockout approaches revealed that diversification and selection in integrated human Ig loci are under the control of the animal's immune system and that they can undergo natural processes of V(D)J rearrangement, somatic hypermutation (SHM) and class-switching. Furthermore, great effort was made to increase the number of V gene segments in mice, which is essential not only for the production of a diverse repertoire of antigen-specific human antibodies but also for the proper development of the B cell lineage[2]. With such

transgenic-knockout mice, antigen-specific fully human mAbs can be readily produced by well-established hybridoma technology[1].

This transgenic method, however, relies on the random insertion of transgenes meaning that it is difficult to include all regulatory elements[4], and there is the possibility that expression of the inserted human Ig loci will be affected by surrounding sequences. Recently, mice that produce a chimeric antibody comprising human variable and mouse constant regions has been engineered by a sophisticated method of replacing the genomic sequence of the mouse Ig variable region with that of a human Ig variable region[5–8]. In these mice, antigen-specific chimeric antibodies with human Ig variable regions are produced as efficiently as in wild-type (WT) mice and, by having a mouse-derived constant region, improved B cell development was achieved[5,7].

To introduce megabase-sized segments of DNA into mice, we have developed an alternative strategy utilizing a human chromosome as a vector for transgenesis[9]. Using this technology, two transmittable human chromosome fragments, one containing the Ig heavy chain locus (*IGH*, ≈1.5 Mb) and the other containing the kappa light chain locus (*IGK*, ≈2 Mb), were introduced into a mouse strain whose endogenous *Igh* and *Igk* loci were inactivated[10,11]. Hybridomas

producing antigen-specific fully human antibodies were obtained from these trans-chromosomic (Tc) mice. Compared with other models, the double-Tc mice contained the largest fraction of human Ig loci at that time; however, there was some instability of human chromosome 2 (hChr.2)-derived human chromosome fragments containing *IGK*, which contributed to lower efficiency of hybridoma production, which was less than one-tenth of that observed in normal mice[12]. In addition, human Ig repertoire formation that relies on the introduction of entire human Ig loci into mice has not yet been evaluated in double-Tc mice. To solve this issue, a Tc mouse carrying hChr.14-derived fragment (hCF14) containing *IGH* was cross-bred with a YAC-transgenic mouse carrying about 50% of *IGK* segments, resulting in a new mouse strain exhibiting considerably improved hybridoma production[13]. However, subsequent studies revealed mosaicism of hCF14 in various tissues of Tc mice, indicating mitotic instability of the human centromere contained in hCF14[14]. We, therefore, constructed a mouse artificial chromosome (MAC) containing a mouse-derived centromere to improve stability[9] in Tc mice. We demonstrated nearly perfect stability in all tissues of Tc mice, germline transmission to offspring, and expression of introduced exogenous

genes. Thus, the generation of MAC-based, human antibody-producing Tc mice has been anticipated.

Here, we present the generation and detailed analysis of a novel fully human Ab-producing Tc mouse (TC-mAb mouse) carrying a MAC containing entire human *IGH* and *IGK* loci [designated as IGHK-NAC (novel artificial chromosome)]. The TC-mAb mice stably maintained the IGHK-NAC as expected, enabling us to analyse the human Ig repertoire and to provide fully human Abs in the context of entire human Ig loci transplanted into an Ig knockout mouse background. Comprehensive profiling of human antibody repertoires including V(D)J gene usage in combinatorial rearrangement, junctional diversity, and SHM by next-generation sequencing (NGS) revealed human-like repertoire formation in TC-mAb mice, which relies on human-derived cis-regulatory elements contained in the complete *IGH* and *IGK* transloci. Upon immunization, TC-mAb mice increase the number of Ab-producing cells, such as plasmablast and plasma cells, and hybridomas can be efficiently generated that express antigen-specific fully human mAbs. Some retardation in B cell development and a delayed immune response was, however, observed. Thus, TC-mAb mice are a useful tool

for recapitulating the human antibody repertoire for efficient therapeutic drug discovery.

Results

Construction of a novel IGHK-NAC containing fully human Ab genes.

For the production of fully human Ab producing mice, sequential translocation cloning of human *IGK* (on hChr. 2) and *IGH* (on hChr. 14) loci into the MAC vector was conducted using Cre/loxP and FLP/FRT systems[9] (Fig. 1a and Supplementary Fig. 1). The MAC is composed of a native mouse centromere, a loxP site, part of the 3′ region of the hypoxanthine-guanine phosphoribosyl transferase (*HPRT*) gene, and telomeres. In addition, the MAC contained enhanced green fluorescent protein (EGFP) and neomycin resistance genes, driven by the ubiquitous promoters, CAG and PGK, repectively[9].

After inserting loxP and FRT sites proximally and distally of the *IGK* locus on hChr.2p, respectively, the modified hChr.2p was transferred into CHO cells carrying the MAC using microcell-mediated chromosomal transfer (MMCT)[10]. An intended reciprocal translocation between the MAC and the modified hChr.2 by Cre/loxP recombination caused reconstitution of the *HPRT* gene and HAT resistance, which enabled us to select CHO cell lines carrying the MAC with the *IGK* locus (IGK-NAC). Using the same procedure, the distal region covering the *IGH* locus on hChr.14 was sequentially translocated into the IGK-NAC to produce

CHO cells with the MAC carrying both *IGK* and *IGH* loci (IGHK-NAC) as described in the Methods section. The construction of IGHK-NAC was confirmed by genomic PCR analysis and *in situ* fluorescence hybridization (FISH) analysis (Supplementary Fig. 2–5).

Generation of a humanized transchromosomic (Tc) mouse.

To generate a humanized Tc mouse, the IGHK-NAC was transferred from CHO cells to mouse embryonic stem (ES) cells (TT2F: 39, XO) and mouse ES cells, in which endogenous *Igk* and *Igh* had been previously knocked out (6TG-9-mES: 39, XO)[11,15]. The mES hybrids with the desired karyotype containing the IGHK-NAC (40, XO, +IGHK-NAC) were isolated (Supplementary Fig. 6) and utilized to produce chimeric Tc mice. Mice with high coat colour chimerism were mated with endogenous Ig gene knockout mice derived from HKD-mES cells. After confirming the IGHK-NAC was transmitted through the germline, F1 mice were further mated with CD-1 (ICR) mice carrying a Ig lambda light chain locus (*Igλ*) low allele[11] to establish Tc mice carrying the IGHK-NAC with mouse *Igk* and *Igh*-KO, and *Igλ* low background (designated as TC-mAb mice).

The retention rate of the IGHK-NAC in various tissues of the TC-mAb mice was evaluated by FISH analysis and by monitoring EGFP expression. This indicated a high percentage of IGHK-NAC retention (Fig. 1b–d). Notably, flow cytometry analyses revealed that the IGHK-NAC was also stable in blood cells, as reported previously[16]. The average percentage of EGFP-positive peripheral blood mononuclear cells (PBMCs) in F2–12 generation TC-mAb mice was 93.6% (Fig. 1e), confirming high stability of the IGHK-NAC over multiple generations, especially in blood cells. The average percentage of both EGFP- and B220-positive cells was 9.5% in TC-mAb mice (Fig. 1f), indicating the IGHK-NAC-dependent rescue of B cell production in the Ig-knockout mice (HKLD). It should be noted that the number of circulating B220-positive cells in TC-mAb mice with a single IGHK-NAC copy was 50–60% of that observed in double heterozygous mice (hetero HKLD) carrying one copy each of functional human and mouse Ig loci. Functional expression of transcripts from the IGHK-NAC was analysed by RT-PCR using total RNA from various tissues of the TC-mAb mice. Transcripts of human *IGK* and *IGH* were abundant in the thymus, spleen, and at low levels in the intestine (Fig. 1g) where B cells are related to T cell development, B cell maturation, and IgA antibody production, respectively. Taken together, we

successfully generated a strain of humanized IGHK Tc mice that displayed tissue-specific expression of the human Ig-genes on the IGHK-NAC.

Analysis of human Ig heavy chain repertoire in TC-mAb mice.

Immunoglobulin genes are assembled by recombination of germline-encoded gene segments: variable (V), diversity (D), and joining (J) for heavy chains and V and J for light chains. Recent advances in NGS technologies enables comprehensive profiling of human antibody repertoires, including V(D)J gene usage in combinatorial rearrangement, junctional diversity, and SHM at an unprecedented scale. However, reported repertoire analyses in human Ig YAC transgenic mice producing fully human Abs has involved a limited number (~500) of cloned Ig sequences[17]. In this study, to investigate the transcript sequences of human Ig from IGHK-NAC, we employed NGS to interrogate human Ig repertoires in RNA samples extracted from spleen cells of five unimmunized TC-mAb mice, one OVA-immunized TC-mAb mouse, and human PBMCs from five healthy adult donors. Over 1.3 million qualified reads were accumulated from each sample and assembled into merged reads. IgBlast-annotated reads were collated into data sets for subsequent analyses (Supplementary Table 1). The saturation of

clonotype variations was confirmed in the rarefaction curve of each sample. In addition, the annotated reads showed that over 92% of the *IGH* and *IGK* transcripts in TC-mAb mice and in hPBMCs were productive.

Comparison of the use of heavy chain V segments denoted as functional by the IMGT database[18,19] (Fig. 2a and Supplementary Table 2) showed that all of the functional V segments (41 segments) were detected in hPBMCs, unimmunized TC-mAb, and immunized TC-mAb mice, and intriguingly, the broad distribution of V segment usage was very similar among these samples. For instance, of twenty V segments represented at a frequency of over 2% in hPBMCs, 17 were also detected at a frequency of over 2% in unimmunized TC-mAb or immunized TC-mAb mice. In addition, in contrast to mice with human Ig YAC transgenes[17] and humanized VH regions[7], which exhibit biased use of the top five frequently used V segments (over 60%), use of the top five V segments in unimmunized TC-mAb mice was lower than 40%, as seen in PBMCs.

The use of D segments was also mostly similar between TC-mAb mice and hPBMCs. *IGHD3-10* gene segments were the most frequently used segment in both hPBMCs and TC-mAb mice, although their use in TC-mAb mice (33.21%) was significantly higher than that in hPBMCs (11.22%) (Supplementary Table 2).

There was no evidence of preferred recombination between D and J segments that are in close proximity, which was observed in human Ig YAC transgenic mice[3]. All six J segments were used, and the J4 and J6 gene segments were dominantly used in TC-mAb mice and hPBMCs. Furthermore, Circos plots of V, D, and J segment combinations in rearranged Ig heavy chains revealed that the frequency of V/D/J use and the combinations of VH regions were comparable between hPBMCs and TC-mAb mice (Fig. 2b and Supplementary Fig 7), indicating no significant difference in V/D/J gene use and combination in TC-mAb mice with or without immunization. The Shannon-Weaver index, which indicates robust estimates of overall immune-repertoire diversity, also indicated that V(D)J recombination in TC-mAb mice occurred as efficiently as it does in humans and that human-like combinatorial diversity is formed in TC-mAb mice by common recombination mechanisms that rely on human-derived *cis*-regulatory elements in the *IGH* locus (Supplementary Table 3).

CDRH3 length distribution and amino acid composition.

Complementary-determining region 3 of the heavy chain (CDRH3) is located at the V/D/J junctional region of the Ig heavy chain; therefore, it is highly

diverse and a key determinant of specificity in antigen recognition[20]. In hPBMCs and unimmunized TC-mAb mice, the distribution of CDRH3 length closely overlapped and the average length was 14.97 and 14.22 amino acids, respectively, with a broad distribution between 5 and 25 amino acids (Fig. 2c). Nucleic acid addition at V-D and D-J junctions was observed in the shorter distributions in TC-mAb mice compared with hPBMCs (Fig. 2d), which was consistent with a previous report indicating a significant decrease in the number of N additions in Ig YAC transgenic mice carrying part of both *IGH* and *IGK* loci[17]. The amino acid composition of CDRH3 among hPBMCs and TC-mAb mice was almost the same indicating that Ig-gene rearrangement was processed in the human manner (Fig. 2e).

Repertoire analysis of human *IGK* chains in TC-mAb mice.

The human kappa locus has large duplicated clusters of kappa chain variable (VK) segments (J-proximal and J-distal), which are separated by 800 kb. They should contribute to expanding the potential *IGK* repertoire. The *IGKV* distal cluster spans 400 kb and comprises 36 (16 potentially functional) V segments and the proximal cluster spans 600 kb and comprises 40 (18 potentially

functional) V segments. Because of this complex structure the introduction of the entire human *IGK* locus into mice has only been accomplished using our chromosome vector system. This enables evaluation of the *IGK* repertoire, which relies on the native configuration of entire human *IGK* locus. The NGS data indicated that all of the 18 potentially functional proximal V segments were used in both hPBMCs and TC-mAb mice (Fig. 2f). Eleven and 9 of 16 functional distal V segments were also detected in TC-mAb mice and hPBMCs, respectively, though their use was less frequent compared with proximal V segments. These results contrast well with previous data showing highly biased use of V segments in mice with a partial *IGK* YAC transgene[17] or a humanized VK region[5,7]. The Circos plots of V-J segments showed that the frequency use and assembled combinations of the VK region were highly similar between hPBMCs and TC-mAb mice (Fig. 2g and Supplementary Table 2). Also, there was no significant difference in J segment use in TC-mAb mice with or without immunization. Furthermore, CDRL3 length, the number of added nucleotides, and amino acid composition were all comparable between hPBMCs and TC-mAb mice (Fig. 2h-j). Taken together, as seen with the human Ig heavy chain repertoire, the

formation of human-like combinatorial and junctional diversity was achieved in TC-mAb mice carrying a MAC with the entire human Ig kappa locus.

Somatic hypermutation in TC-mAb mice.

SHM is a process in which point mutations accumulate in the variable regions of both the Ig heavy and light chain genes during B cell maturation, thereby enabling the selection of B cells producing high-affinity Abs against immunogens. Based on NGS data of VH and VK sequences, clone lineages for both IgM and IgG were identified to estimate the diversification of the B cell receptor (BCR) repertoire. In this study, the definition of clone lineage referred to a set of B cells that are related by descent, arising from the same V(D)J rearrangement event. The definition of clonotype refers to a single Ab sequence[21]. To detect the position of SHMs leading to amino acid changes in the VH and VK sequences, annotated reads were obtained from unimmunized and immunized TC-mAb mice, and were compared with their germline sequences.

The 70 most frequently used clone lineages of VH (CLH001-070) and the 20 most frequently used VK (CLL001-020) sequences were analysed and compared between immunized and unimmunized TC-mAb mice (Fig. 2k–n). The

SHM in these clone lineages was accumulated in the CDR1, 2, 3, and FR3 regions of both VH and VK sequences. The mutations appeared more frequently in immunized TC-mAb mice, indicating selection with antigen administration. The data of the top 20 clone lineages (CLH001-020 of VH and CLL001-020 of VK) were summarized for the fold change of annotated reads, the average length of CDR3, and the percentage of SHM with at least two mutations in a variable region (Supplementary Table 4). The fold change in annotated reads indicated that specific clone lineages of both VH and VK sequences were expanded during immunization, while the fold change of the VH sequence was more prominent than that of the VK sequence. However, a higher mutation rate in the VK sequence was observed compared with the VH sequence after immunization.

Phylogenetic trees (circular dendrograms) were drawn to show the mutation mapping and expansion of clonotypes based on their amino acid sequence in each clone lineage. Based on the number of copies with more than 100 reads in the OVA-immunized TC-mAb mice, 260 clonal lineages of the VH sequence and 148 lineages of the VK sequence were analysed. The circular dendrogram of the 10 most frequent clonal lineages of the VH sequence and VK sequence is presented in Fig. 2o (and Supplementary Fig. 8). The sequence

diversity of a clone, CLH001, is presented in Fig. 2p. The average fold-expansion among the top 20 clone lineages was 16.0 for VH sequence and 5.7 for VK sequence (Supplementary Table 4). These results indicated that expansion of clonotypes with SHM was induced by immunization. Taken together, immunization of TC-mAb mice results in B cell clonal selection and expansion followed by SHM, leading to antigen-driven diversification of VH and VK sequences, which should contribute to the generation of B cells bearing higher-affinity antibodies.

Serum expression of human Igs in TC-mAb mice.

Unimmunized TC-mAb mice (*n=17*) were examined by enzyme-linked immunosorbent assays (ELISAs) to determine serum concentrations of human and mouse Ig proteins. The average levels of human Ig μ, γ and κ were 518.1, 153.2, and 688.6 μg/ml, respectively (Fig. 3a). As expected, the levels of mouse Ig μ, γ, κ, and λ in unimmunized TC-mAb mice were at significantly lower levels, as reported in double Tc mice[11], compared with the parental WT mice (*n=12*) in which the levels of Ig μ, γ and κ were 248.0, 993.4 and 637.0 μg/ml, respectively (Supplementary Table 5). Although we used the λ1 low allele[22] instead of Ig λ

knockout to reduce mouse Ig λ expression in TC-mAb mice, the result indicates that the human IGK locus of the IGHK-NAC can compete well with the mouse λ1 low allele. The observation that levels of Ig γ were lower than those of Ig μ is consistent with data from double-Tc mice and Ig YAC Tg mice, which rely on human-derived constant region sequences[3,11]. All four γ subclasses (γ1, γ2, γ3, and γ4) were detected and the average concentrations were 82.2, 109.3, 4.6, and 13.3 μg/ml, respectively. In addition, Ig α and Ig ε were also detectable (Fig. 3a–b, and Supplementary Table 5).

When TC-mAb mice (*n=21*) were administrated various antigens, including OVA, recombinant proteins expressed in *Escherichia coli*, and whole mammalian cell extract, the Ig μ, γ, and κ levels and the number of spleen cells were greatly elevated (Fig. 3c–d). In particular, the fold change in the average level of Ig γ (8.6) was greater than that of Ig μ (4.0), which indicates a robust induction of class-switching from Ig μ to Ig γ.

Production of antigen-specific fully human Abs in TC-mAb mice.

To optimize the immunization conditions for the production of antigen-specific human Abs in TC-mAb mice, we first immunized age-matched TC-mAb

mice (*n=2*) and WT (ICR) mice (*n=2*) with OVA using Freund's complete adjuvant. The OVA-specific human Ig γ response in TC-mAb mice was relatively delayed compared with that in WT mice (Fig. 3e), indicating the requirement for two or three additional booster steps to reach a plateau and similar OVA-specific mouse Ig titres to those in WT mice (Supplementary Fig. 9). Then, Trx-EpEX and GST-EpEX, each of which is a fusion protein containing the extracellular domain from human EpCAM, was used as an immunogen and for verifying the serum titre of EpEX-specific Abs, respectively. Trx-EpEX is a fusion protein containing thioredoxin 1 (Trx) as a Tag sequence. GST-EpEX contained glutathione S-transferase (GST) as an alternative to Trx. Following administration of Trx-EpEX in TC-mAb mice (*n=2*), the titres of Trx-EpEX-specific human Abs increased linearly until the second or third booster steps and then increased slightly until the final booster step (Fig. 3g). The titres of GST-EpEX-specific human Ig were substantially increased after the third or fourth booster steps. After the fourth (individual A) and seventh boosters (individual B), each immunized TC-mAb mouse was euthanized and lymphocytes were harvested from the spleen and lymph nodes, respectively, for the production of EpEX-specific human mAbs (Supplementary Table 6). The human Ig-isotypes of these EpEX-specific human

mAbs (80 and 279 clones from individual A and B, respectively) were analysed (Fig. 3h) to validate class-switching in TC-mAb mice. While the majority (73.8%) of hybridoma clones from individual A (four boosters) produced IgM, those from individual B (seven boosters) produced IgG as the major isotype (86.9%). Further analysis showed that the percentages of IgG1, IgG2, IgG3, and IgG4 subclasses in IgG producing hybridomas were 64.3%, 4.7%, 16.5%, and 1.4%, respectively. It should be noted that the number of mAbs obtained from TC-mAb mice was much higher than that produced from Balb/c mice (52 clones) using the same protocols (Supplementary Table 6).

To analyse IGHK-NAC stability in hybridoma cells, native antigen-reactive clones were selected and established by limiting dilution cloning. The cloning success rate was over 96.1% (Supplementary Table 6), revealing that the IGHK-NAC was stable in hybridoma cells.

Taken together, we achieved highly efficient generation of antigen-specific human Ab producing hybridomas in TC-mAb mice, even though two or more booster immunization steps were required for optimal responses. These results also suggest that high population subsets of antigen-specific B cells are contained in the spleen of immunized TC-mAb mice.

We then measured total Ig levels in the serum of TC-mAb individual B (which received seven booster administrations) (Fig. 3i). In agreement with the increase in the titre of antigen Trx-EpEX-specific Ab (Fig. 3g), the serum concentrations of human Ig μ, γ, and κ were robustly elevated after the second booster step. The human Ig μ level reached a plateau after the second booster step and the human Ig γ level exceeded that of human Ig μ after the fifth booster step, which is consistent with the majority of antigen-specific mAbs isolated from this TC-mAb mouse (individual B) consisting of different IgG subclasses.

Characteristics of mAbs obtained from TC-mAb mice.

Analysis of amino acid sequences of mAbs obtained from TC-mAb mice indicated a high degree of humanness, which was determined using T20 score analyser[23]. This confirmed that hybridoma clones carried fully human Ig heavy and kappa chains (Supplementary Fig. 10). The affinity of human mAbs against EpCAM and AMIGO2 was estimated using surface plasmon resonance. The K_D values of human mAbs have a substantial affinity against target proteins in the nanomolar range (Supplementary Table 7).

B-cell development and antigen-specific B cells in TC-mAb mice.

The B cell lineages in bone marrow and spleen of TC-mAb mice were analysed by flow cytometry using combinations of Abs against cell surface markers (Supplementary Table 8)[24,25] and compared among the five age-matched TC-mAb mice and WT mice (Supplementary Table 9). The subset populations of Pro-B and Pre-B, and recirculating B cells were indistinguishable but that of immature B cells was decreased by half in TC-mAb mice compared with WT mice (Fig. 4a and Supplementary Table 9). In the spleen, the percentages of transitional B and follicular B cells in lymphocytes were decreased, but that of marginal zone B cells was maintained (Fig. 4b and Supplementary Table 9). In human, marginal zone B cells in the spleen is evidence of SHM in Ig-genes and are a reservoir of memory B cells[26]; therefore, these results indicate that antigen-specific B cells were effectively produced in the spleen of TC-mAb mice.

The percentage of germinal centre (GC) B, follicular and, plasma blasts (PB) and plasma cells (PC) were also analysed with or without immunization using OVA as antigen (Supplementary Table 10). In immunized TC-mAb mice, we detected GC formation in spleen (Fig .4c) and the percentage of GC B cells was

slightly decreased compared with that of WT mice (Fig. 4d). With administration of the antigen, the percentages of Ab-producing cells, including PB and PC, were significantly larger in TC-mAb mice than in WT mice (Fig. 4e). There was a 4.7-fold increase (3.9% vs 0.8%) in the unimmunized state and a 2.7-fold increase (6.5% vs 2.4%) in the OVA-immunized state, respectively.

A high percentage of PB and PC subsets supported the efficient production of antigen-specific Abs in anti-sera of TC-mAb mice (Fig. 3), which was consistent with expansion of antigen-inducing clonotypes (Supplementary Fig. 8). A fluorescence-labelled antigen was recently described that detects B cell subsets expressing antigen-specific Ab[27–30]. Thus, we employed a combination of a fluorescent-labelled anti-human IgG and antigen (OVA) to detect the subset populations of the antigen-specific Ab producing B cells in the spleen (Fig. 4f). Surprisingly, the percentage of OVA-specific B cells in the IgG-positive fraction of lymphocytes was very high in TC-mAb mice (72.3%), but low in WT mice (6.1%) (Supplementary Table 10). These results were consistent with the high production rate of antigen-specific hybridoma cells (Supplementary Table 6) and high subset populations of PB and PC in TC-mAb mice (Fig. 4e). These results indicate that

TC-mAb mice have desirable features for the production of antigen-specific, therapeutic mAbs.

Discussion

In this study, a new generation of fully human Ab producing mice (TC-mAb mice) was established that overcome the instability of Ig-transloci in previously generated double Tc mice[11]. Our repertoire analysis of human Ab-producing mice by NGS revealed that the entire human heavy and kappa chain loci were used in the mouse similarly to those in hPBMCs, including the use of Ig-gene segments, combinations of V(D)J rearrangement, production of long-length CDRH3, and amino acid composition of CDRH3. Comparison with Ig-gene segment use in human/mouse chimeric Ab-producing mice (genetically chimeric mice) revealed that the Ig-gene repertoire was slightly or significantly different from that of hPBMCs[5,7,8,17] and TC-mAb mice. In addition, a high frequency of SHM in CDR regions was detected in both VH and VK regions (Fig. 2k-n). Therefore, our analysis of TC-mAb mice strongly suggests that carrying the entire human Ig locus of *IGH* and *IGK* was necessary and sufficient to faithfully reproduce the human Ab rearrangement process in the mouse. These results are consistent with those of a previous report[31], indicating that the human gene expression pattern was recapitulated by the human trans-chromosome in mouse cells and that the genetic sequence was largely responsible for directing

transcriptional programs in homologous tissues. It is clear that “genomically” humanized animals, generated by Tc technology using mouse artificial chromosomes, is a powerful tool to verify the function of gene loci[9].

Some unique profiles of B cell development and Ab production were also observed in Tc-mAb mice, consistent with reports showing species incompatibility of Ig-genes between human and mouse[3,11]. For instance, a lower concentration of IgG (Fig. 3a and Supplementary Table 5) and altered B cell development (Fig. 4a–e and Supplementary Table 9 and 10) in TC-mAb mice compared with WT mice as well as the requirement of two or three additional booster steps to elicit optimal immune response were observed (Fig. 3e). Intriguingly, our detailed analyses showed high titre antigen-specific Abs containing all human subclasses (Fig. 3b), and large populations of antigen-specific B cell subsets (Fig. 4g) with SHM and affinity maturation (Fig. 2o and Supplementary Fig. 8), suggesting that various antigen-specific human Abs were efficiently produced in TC-mAb mice. As expected, a large antigen-specific B cell subset in the spleen of TC-mAb mice led to high efficiency hybridoma production. Because unimmunized TC-mAb mice showed low Ig γ concentration (Supplementary Table 5) and few IgG-positive subsets (Fig. 4g and Supplementary Fig. 11), an environment might have existed

in which naïve B cells and antigen-specific memory B cells could easily respond to antigen and provide IgG-positive B cells by homeostatic proliferation[32,33] or other unknown mechanisms. This speculation is supported by the high production rate of EpCAM-specific mAbs from TC-mAb mice (Supplementary Table 6), although further investigations are needed to understand the underlying mechanisms, which could have important implications for the development of high efficiency therapeutic mAb production. The TC-mAb mice have only one allele of human *IGH* and *IGK* loci in the IGHK-NAC; therefore, inter-mating of TC-mAb mice to generate mice carrying two copies of IGHK-NAC may improve the immune response and lymphocyte production rate. Alternatively, swapping hIgG class switch regulator elements from human to mouse in the IGHK-NAC may also improve these responses[34]. We have summarized the key characteristics of TC-mAb mice for producing antigen-specific mAbs and have compared them with other fully human Ab-producing mice (Supplementary Table 11). Our results show that the fully human Ab-producing mice (TC-mAb mice) described here will be very useful for developing therapeutic Abs[27–30].

Methods

Cell culture.

A MAC vector was used to generate IGHK-NAC[9]. The MAC contained a mouse centromere, EGFP flanked by HS4 insulators, PGK-neo, loxP site-3′-HPRT, PGK-puro, and telomeres. Chicken DT40 cells containing hChr.2 or hChr.14 were maintained at 40°C in RPMI 1640 medium supplemented with 10% foetal bovine serum (FBS), 1% chicken serum, 50 μM 2-mercaptoethanol, and 1.5 mg/mL G418. *Hprt*-deficient Chinese hamster ovary [CHO (*Hprt*$^{-/-}$)] and CHO K1 cells were maintained at 37°C in Ham's F-12 nutrient mixture (Invitrogen, Carlsbad, CA, USA) supplemented with 10% FBS. CHO cells containing the IGHK-NAC were maintained in medium with 800 μg/mL G418. Mouse embryonic fibroblasts (MEFs) were isolated from embryos at 13.5 days postcoitum (d.p.c.). MEFs were grown in Dulbecco's modified Eagle's medium (DMEM) (Sigma-Aldrich, St. Louis, MO, USA) containing 10% FBS. Parental mouse ES cell line (TT2F), endogenous Ig KO ES cell subline (6TG-9) and microcell hybrid TT2F and 6TG-9 clones were maintained on mitomycin C (Sigma-Aldrich)-treated Jcl:ICR (CLEA Japan, Tokyo, Japan) MEFs and neomycin-resistant MEFs (Oriental Yeast Co., Ltd., Tokyo, Japan), respectively, as feeder layers in DMEM

with 18% FBS (Hyclone Laboratories, Logan, UT, USA), 1 mM sodium pyruvate (Invitrogen), 0.1 mM non-essential amino acids (Invitrogen), 0.1 mM 2-mercaptoethanol (Sigma-Aldrich), 2 mM L-glutamine (Invitrogen), and 1000 U/mL leukaemia inhibitory factor (Funakoshi, Tokyo, Japan).

FISH.

Trypsinized cells and homogenized tissue samples were incubated for 15 min in 0.075 M KCl, fixed with methanol and acetic acid (3:1), and then slides were prepared using standard methods. FISH analyses were performed using fixed metaphase or interphase spreads of each cell hybrid using digoxigenin-labelled (Roche, Basel, Switzerland) DNA [human COT-1 DNA/mouse COT-1 DNA (Invitrogen), mouse minor satellite DNA and IGK-BAC (CH17-405H5 and CH17-216K2)] and biotin-labelled DNA [human COT-1 DNA/mouse COT-1 DNA, IGH-BAC (CH17-262H11, CH17-212P11 and RP11-731F5) and each part of the targeting vector], essentially as described previously[10]. Chromosomal DNA was counterstained with DAPI (Sigma-Aldrich). Images were captured using an AxioImagerZ2 fluorescence microscope (Carl Zeiss GmbH, Jena, Germany).

Generation of chimeric mice and Tc mice.

For the Tc line, chimeric mice were produced from ES (IGHK-NAC) cell lines. Briefly, ES cells were injected into eight-cell-stage embryos derived from ICR mice (CLEA, Tokyo, Japan) and then transferred into pseudopregnant ICR females. Chimeric mice showing 100% coat-colour chimerism were used to generate mice for germline transmission. The IGHK-HKD Tc mice were further mated with the KO mice, in which the endogenous *Igh* and *Igk* were knocked out, and CD-1 mice (Charles River, Kanagawa, Japan) for the Igλ low allele[11]. Resultant IGHK-HKLD mice were used in FISH and FCM analyses. The fully humanized IGHK mice were used in RT-PCR and several functional assays. Representative data from these assays are shown in each figure. All animal experiments were approved by the Animal Care and Use Committee of Tottori University.

RT-PCR.

Total RNA from Tc tissue specimens was prepared using ISOGEN (Nippon Gene, Tokyo, Japan), treated with RNase-free DNase I (Wako Pure Chemicals, Osaka, Japan), and purified using RNeasy columns (Qiagen, Hilden,

Germany), in accordance with the manufacturer's instructions. First-strand cDNA synthesis was performed using random hexamers and SuperScript III reverse transcriptase (Invitrogen). Primer pairs for the detection of human Ig-gene expression were as follows: Vk1BACK/Ck and CH4BACK/Cmu-1[11]. GAPDH (RPC1/2) was used as an internal control. The primer sequences for RT-PCR analyses are described in Supplementary Table 12. cDNAs from C57BL/6 and ICR tissues were used as negative controls. PCR was performed with cDNA using AmpliTaq Gold (PerkinElmer, Waltham, MA, USA). Amplified fragments were resolved by electrophoresis on 2% agarose gels, followed by staining with ethidium bromide.

Deep sequencing analysis of Ab-coding transcripts.

An NGS analysis was performed using the unbiased TCR/BCR repertoire analysis technology developed by Repertoire Genesis, (Osaka, Japan). In brief, unbiased adaptor-ligation PCR was performed as previously described[35]. Total RNA was converted to cDNA with Superscript III reverse transcriptase (Invitrogen). The BSL-18E primer (5′-AAA GCG GCC GCA TGC TTT TTT TTT TTT TTT TTT VN-3′) containing polyT18 and a *Not*I site was used for cDNA

synthesis. After cDNA synthesis, double-stranded (ds)-cDNA was synthesized with *Escherichia coli* DNA polymerase I (Invitrogen), *E. coli* DNA ligase (Invitrogen), and RNase H (Invitrogen). The ds-cDNA was blunted with T4 DNA polymerase (Invitrogen). A P10EA/P20EA (5′-GGG AAT TCG G/TA ATA CGA CTC CGA ATT CCC-3′) adaptor was ligated to the 5′ end of the ds-cDNA and then cut with a *Not*I restriction enzyme. After removal of the adaptor and primer with a MinElute Reaction Cleanup Kit (Qiagen), PCR was performed with KAPA HiFi DNA polymerase (Kapa Biosystems, Wilmington, MA, USA) using an IgG constant region-specific primer (CG1: 5′-CAC CTT GGT GTT GCT GGG CTT-3′, for BCR) and P20EA (5′-TAA TAC GAC TCC GAA TTC CC-3′). The PCR conditions were as follows: 98°C (20 s), 65°C (30 s), and 72°C (1 min) for 20 cycles. The second PCR was performed with either CB2 (5′-AGG CAG TAT CTG GAG TCA TTG AG-3′) or CG2 (5′-TCC TGA GGA CTG TAG GAC AGC-3′) and P20EA primers using the same PCR conditions. Amplicons were prepared by the amplification of the products from the second PCR using P22EA-ST1 (5′-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCT AAT ACG ACT CCG AAT TCC C-3′) and either CB-ST1-R (5′-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG GCT CAA ACA CAG CGA CCT C-3′) or CG-ST1-R (5′-TCG

TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TGA GTT CCA CGA CAC CGT CAC-3′). After PCR amplification, index (barcode) sequences were added by amplification with Nextera XT index kit v2 setA (Illumina Inc., San Diego, CA, USA). Equal molar concentrations of the indexed amplicon products were mixed and quantified by a Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). Sequencing was performed using the Illumina MiSeq paired-end platform (2 × 300 bp).

Data analyses.

All paired-end reads were classified by index sequences. Sequence assignment was performed by determining the sequences with the highest identity in a dataset of reference sequences from the international ImMunoGeneTics information system (IMGT) database (http://www.imgt.org). Data processing, assignment, and merging were performed automatically using a repertoire analysis software program originally developed by DNA Chip Research Inc. (Tokyo, Japan).

Annotated sequence reads were defined as distinct sequence reads within the population of merged sequence reads that had been identified as a

BCR gene. The copy numbers of identical annotated reads in each sample were automatically counted using the RG software program and then ranked numerically. According to IMGT nomenclature, the CDR3 nucleotide sequences from a conserved cysteine at position 104 (Cys104) to a conserved phenylalanine at position 118 (Phe118) and the following glycine (Gly119) were translated to deduce amino acid sequences.

Circos analysis.

Circos software was selected for this study for its high data-to-ink ratio and for its ability to clearly display relational data. Circos open-source software was obtained from www.circos.ca. The V(D)J region recombination data were reformatted using the R statistical programming language to comply with Circos data file requirements. Library sizes were normalized with Circos ideogram (circumference segments) scaling and sizing, permitting comparison of individual subgroups within libraries as well as across disparate libraries. Links, drawn from a V region to its observed J region recombinant partner, were used to demonstrate the frequency of recombination, with thicker links indicative of higher frequencies of recombination. The ideogram space allotted to the V region

subgroup corresponds to the frequency of its observation relative to other subgroups. Analysis of V(D)J recombination was conducted with an additional stacked histogram track on each Circos diagram. This track illustrates the relative proportion of each V(D)J recombination as a fraction of the total number of D region sequences observed.

Detection of somatic hypermutations in human VH and VK regions.

Mutations were detected by comparison with germline sequence at every nucleotide position (around 315 nucleotides) and were calculated as a percentage. For example, when 10 reads were recorded in the same clone lineage and three reads had a point mutation at the same nucleotide position, the mutation rate at that position was described as 30%. The mutation rate at every position in the same clone lineage was integrated as a clone linage mutation rate. In addition, there were different lengths of Ab variable regions in annotated reads; therefore, an index of variable region length was set that was the average variable region length in a clone lineage and converted to 100. In this way, clone lineages could be compared at the same magnitude and the index was similar to the amino acid position in the variable region.

Diversity index.

To estimate BCR diversity in deep sequence data, the Shannon-Weaver index (H′) was calculated using the following formula:

$$H' = -\sum_{i=1}^{S} \frac{ni}{N} \ln \frac{ni}{N}$$

where N is the total number of sequence reads, n_i is the number of *i*th annotated read, and S is the species number of annotated reads[36]. The greater the H′ value, the greater the sample diversity.

Phylogenetic analysis of the human Ab repertoire.

Phylogenetic trees (circular dendrograms) were created by alignment of convergent CDRH3 and CDRL3 amino acid sequences using the multiple sequence alignment program and the Neighbour-Joining method[37]. Furthermore, two phylograms of unimmunized and OVA-immunized TC-mAb mice were assembled within one phylogram based on their amino acid sequences. In addition, copies comprising the same CDRH3 sequences were counted and overlaid on the leaves of circular dendrograms and are shown with a maximum

of 50 reads; thereby the larger the number of reads, the larger the circle in the leaves.

Antigens.

Ovalbumin (OVA) was obtained from Sigma (A7641). To obtain the Trx-EpEX recombinant protein, the extracellular domain of human EpCAM (NM_002354) was amplified by PCR using primers (Forward; 5′-AAA GAT ATC GGA TCC TCA GGA AGA ATG TGT CTG TGA-3′ and Reverse; 5′-ATA AAG CTT TTT TAG ACC CTG CAT TGA GAA TTC-3′), and subcloned into pET32b (Merck Millipore, Burlington, MA, USA) using *Eco*RV and *Hind*III restriction enzymes (resulting in pET32b-EpEX). To obtain the GST-EpEX recombinant fusion protein, pGEX6P1 (GE Healthcare, Chicago, IL, USA) was modified by insertion of the synthesized DNA (5′-ACG AGA TCT GCC ATG GAC AAG CTT GTC GAC ACG AGC TCG AAT TCG GAT CCC CCG GGG CTC GAG CAC CAC CAC CAC CAC CAC TGA GCT GAG CGG CCG CTC A-3′) using *Bgl*II and *Not*I restriction enzymes (resulting in pGEX-MCS-His). The amplified EpEX fragment was also cloned into pGEX-MCS-His. To produce the AMIGO2 extracellular domain, a DNA fragment of the whole AMIGO2 region (NM_001143668) was amplified by PCR

using primers (Forward; 5′-GCG AAG CTT GTG TGC CCC ACC GCT TGC AT-3′ and Reverse; 5′-GCG CTC GAG TGT GTT AAA TGC CTC ATG AGC ATG GG-3′), and was subcloned into pET32b (Merck Millipore) using *Hin*dIII and *Xho*I restriction enzymes (resulting in pET32b-AMIGO2-EX). AMIGO2-EX was hard to express in *Escherichia coli* gami B pLysS (DE3); therefore, pET32b-AMIGO2-EX was digested with *Eco*RI (an *Eco*RI site is located near the upstream end of the leucine-rich repeats sequence), blunted using Blunting high, and then digested with *Eco*RV (at a site upstream of AMIGO2-EX) to eliminate the leucine-rich repeats sequence. Therefore, this vector consisted of the Ig-like domain of AMIGO2 (named pET32b-AMIGO2-Ig). After transformation of *E. coli* gami B pLysS (DE3) with each vector, the recombinant proteins were expressed by induction with 1.0 mM Isopropyl-β-D(-)-thiogalactopyranoside (WAKO) in LB medium. Transformation using the empty vector (pET32b) was also carried out to produce the Tag protein for use in hybridoma screening as a negative control. After harvesting cells and sonication, the recombinant proteins were obtained as inclusion-bodies. Following solubilization with 6 M guanidine hydrochloride (WAKO) in PBS with 0.1 mM glutathione (oxide form) and 1 mM glutathione (redox form), recombinant protein was purified using Ni-NTA columns with elution

using 100 mM imidazole containing 6 M guanidine hydrochloride. After dialyzing the eluted fraction against PBS containing 0.4 M arginine, samples were diluted to about 1 mg/ml and stored at -30°C. The construction of the pGEX-MCS-His vector and expression and purification of GST-AMIGO2-Ig were carried out using the same procedure as for GST-EpEX.

Immunization.

Protein antigens were prepared in PBS or in PBS containing 0.4 M arginine at 1 mg/ml, and the volume corresponding to the desired amount of protein was increased to an injectable volume with PBS or PBS containing 0.4 M arginine. This volume was then mixed 1:1 (v/v) with either Freund's or Sigma adjuvant (Sigma Adjuvant S6322; Sigma CFA F5881, Sigma: prepared according to manufacturer's instructions). For viscous adjuvants, the solution was mixed by repeated passage through a syringe until a smooth emulsion was formed (over about 30 min on ice). Injections were performed on 6-week-old male and female mice using a 1 ml glass syringe and a 27-gauge needle. Prime and boost injections were intraperitoneal (*i.p.*) every 2 weeks. Volumes varied depending on the injection route and experimental requirements and were in accordance with

the relevant JP Home Office animal license for the procedure. Final boosts were delivered without adjuvant intravenously (*i.v.*) via the tail vein.

Serum concentration of Abs.

The concentrations of human Igs such as hIgM, hIgG, hIgκ, hIgA, and hIgE, and mouse Igs such as mIgM, mIgG, mIgκ, mIgλ were assayed using sandwich ELISA. The concentration of hIgM was assayed using a mouse monoclonal anti-human IgM Ab (Bethyl Laboratories, Montgomery, TX, USA) immobilized on 96-well plates, Nunc MaxiSorp (Thermo) and detected with peroxidase-conjugated mouse anti-human IgM Ab (Bethyl Laboratories). Similarly, hIgG, hIgκ, hIgA, hIgE, mIgM, mIgG, mIgκ, and mIgλ were assayed using capture and detector Abs listed in Supplementary Table 13. The samples, standard, and Ab conjugates were diluted with sample/conjugate buffer (50 mM Tris, 0.14 M NaCl, 1% BSA, 0.05% Tween 20). 3,3',5,5'-tetramethylbenzidine (TMB) (Nacalai Tesque, Kyoto, Japan) was used as substrate, and absorbance at 450 nm was measured using a spectrophotometer (BioTek instruments, Winooski, VT, USA). The IgG subclasses were determined using an IgG Subclass Human ELISA Kit (Invitrogen) according to the manufacturer's instructions.

Serum titre determination.

Serum bleeds taken about 3 days after antigen boost were analysed by ELISA. 96-well immunoassay plates (Nunc Maxisorp) were coated with 100 μl/well of antigen at 0.5 μg/ml in PBS containing 0.4 M arginine overnight at 4˚C. Plates were washed three times with PBS-T (0.05% v/v) and blocked with PBS containing 5% skimmed milk (Difco) for 30 min at room temperature. After being washed again as above, 100 μl of serially diluted serum samples in TBS-T were added to wells and incubated for 1 h at room temperature. After incubation, plates were again washed as above and incubated with 100 μl of anti-human IgG (H+L)-HRP conjugate added at 1/50,000 dilution in TBS-T for 30 min at room temperature. Plates were washed again as above and developed using 100 μl *o*-phenylenediamine dihydrochloride and stopped using 25 μl 1 M H_2SO_4. Absorbance was read at 492 nm.

Subclass determination.

The subclasses of obtained human mAbs were determined using antigen-specific ELISA using horseradish peroxidase-conjugated secondary Abs

specific for human IgG(H+L), IgG1, IgG2, IgG3, IgG4 and IgM, and human Igκ and mouse Igλ. Alternatively, subclasses were determined using the Iso-Gold™ Rapid Human Antibody Isotyping Kit (BioAssay Works, Ijamsville, MD, USA) according to the manufacturer's instructions.

Humanness score of mAbs.

The amino acid sequence of obtained Abs was analysed using the T20 scoring method, which was developed to calculate the humanness of mAb variable region sequences[23]. A Blast search of the variable region was performed against the T20 Cutoff Human Database available at http://abanalyzer.lakepharma.com. The T20 score for an Ab is obtained from the average of the percent identities of the top 20 matched human sequences. To be considered not immunogenic, T20 scores of the FR and CDR sequences must be above 79, and T20 scores for the FR sequences only must be above 86. Scores near or above these values are predicted to be of low immunogenicity.

Cell staining and flow cytometry.

To evaluate the phenotype of TC-mAb mice, we compared them with age-matched WT mice having a similar genetic background. Bone marrow and spleen tissue, and PBMCs were isolated from adult male and female mice (6–20 weeks of age) using aseptic procedures. Single-cell suspensions were prepared from the bone marrow, spleen, and PBMCs. Samples were stained with Abs (Supplementary Table 8) and analysed using a CytoFLEX S (Beckman Coulter, Brea CA, USA). All staining reactions were incubated at 4°C for 30 min using $1x10^6$ cells in 100 μl staining buffer (PBS with 5% FBS:BD Biosciences Brilliant stain buffer; 1:1) (Franklin Lakes, NJ, USA).

Detection of antigen-specific B cells.

For PBMC analysis, blood was collected from the retro-orbital sinus of anesthetized animals into EDTA-coated paediatric tubes (Kabe Labortechnik GmbH, Nümbrecht, Germany). Total white blood cells were counted using a haemocytometer (HIRSCHMANN GmbH, Fluorn-Winzeln, Germany). OVA was biotinylated using the EZ-Link NHS-LC-LC-Biotinylation kit (Pierce, Rockford, IL, USA) in accordance with the manufacturer's instructions, and dialyzed against PBS.

Spleen cells or PBMCs were stained with anti-human IgG-Fc-PE/Cy7 conjugate or anti-mouse IgG-PE/Cy7 conjugate diluted 1:100, 5 μl of 0.21 mg/ml biotinylated OVA, and 5 μl 0.2 mg/ml streptavidin at 4°C for 30 min using $1x10^6$ cells in 100 μl staining buffer. Stained samples were analysed using a CytoFLEX S.

Immunohistochemistry.

Spleens were fixed with phosphate-buffered 4% paraformaldehyde (Nacalai Tesque, Kyoto, Japan) at 4°C for 2 hr, transferred to 20% sucrose in PBS, frozen in OCT compound (Sakura Finetek) and sectioned. Frozen tissue sections on slides were permeabilized with 50 mM Tris-HCL containing 0.1% Triton-X (pH 8.0) at RT for 10 min and then blocked with Blocking One Histo (Nacalai Tesque) at RT for 10 min. Sections were incubated with a 1:100 dilution of biotin-conjugated anti-CD35 mAb (8C12, BD Biosciences, San Jose, CA, USA) and a 1:100 dilution of Alexa Fluor 647-conjugated GL7 (GL7, BioLegend, San Diego, CA. USA) in TBS-T (1 x Tris-buffered saline and 0.1% Tween 20) containing 5% Blocking One Histo at 4°C overnight. Sections were then incubated with a 1:200 dilution of Alexa Fluor 594-conjugated streptavidin (BioLegend) and 2 μM DAPI (BioLegend) in TBS-T containing 5% Blocking One

Histo at 4°C for 45 min. Coverslips were were mounted with ProLong Gold Antifade reagent (Invitrogen) and sections were analyzed with a Zeiss LSM700 confocal microscope (Carl Zeiss, Oberkochen, Germany).

References

1. Lonberg, N. Human antibodies from transgenic animals. *Nat. Biotechnol.* **23**, 1117–1125 (2005).

2. Brüggemann, M. *et al.* Human Antibody Production in Transgenic Animals. *Arch. Immunol. Ther. Exp. (Warsz).* **63**, 101–108 (2015).

3. Green, L. L. *et al.* Antigen–specific human monoclonal antibodies from mice engineered with human Ig heavy and light chain YACs. *Nat. Genet.* **7**, 13–21 (1994).

4. Volpi, S. A. *et al.* Germline Deletion of Igh 3′ Regulatory Region Elements hs 5, 6, 7 (hs5–7) Affects B Cell-Specific Regulation, Rearrangement, and Insulation of the Igh Locus . *J. Immunol.* **188**, 2556–2566 (2012).

5. Murphy, A. J. *et al.* Mice with megabase humanization of their immunoglobulin genes generate antibodies as efficiently as normal mice. *Proc. Natl. Acad. Sci.* **111**, 5153–5158 (2014).

6. Macdonald, L. E. *et al.* Precise and in situ genetic humanization of 6 Mb of mouse immunoglobulin genes. *Proc. Natl. Acad. Sci.* **111**, 5147–5152 (2014).

7. Lee, E. C. *et al.* Complete humanization of the mouse immunoglobulin loci enables efficient therapeutic antibody discovery. *Nat. Biotechnol.* **32**, 356–363 (2014).

8. Asensio, M. A. *et al.* Antibody repertoire analysis of mouse immunization protocols using microfluidics and molecular genomics. *MAbs* **11**, 870–883 (2019).

9. Moriwaki, T., Abe, S., Oshimura, M. & Kazuki, Y. Transchromosomic technology for genomically humanized animals. *Exp. Cell Res.* **390**, 111914 (2020).

10. Tomizuka, K. *et al.* Functional expression and germline transmission of a human chromosome fragment in chimaeric mice. *Nat. Genet.* **16**, 133–143 (1997).

11. Tomizuka, K. *et al.* Double trans-chromosomic mice: Maintenance of two individual human chromosome fragments containing Ig heavy and kappa loci and expression of fully human antibodies. *Proc. Natl. Acad. Sci.* **97**, 722–727 (2000).

12. Isao Ishida, Kazuma Tomizuka, Hitoshi Yoshida, Tomoyuki Tahara, Nobuaki Takahashi, Atsuko Ohguma, Sonoko Tanaka, Misako Umehashi,

Hiroaki Maeda, Chikateru Nozaki, Ed Halk, N. L. Production of Human Monoclonal and Polyclonal Antibodies in TransChromo Animals. *Cloning Stem Cells* **4**, 91–102 (2002).

13. Kuroiwa, Y. *et al.* Cloned transchromosomic calves producing human immunoglobulin. *Nat. Biotechnol.* **20**, 889–894 (2002).

14. Shinohara, T. *et al.* Stability of transferred human chromosome fragments in cultured cells and in mice. *Chromosom. Res.* **8**, 713–725 (2000).

15. Kuroiwa, Y. *et al.* Manipulation of human minichromosomes to carry greater than megabase-sized chromosome inserts. *Nat. Biotechnol.* **18**, 1086–1090 (2000).

16. Takiguchi, M. *et al.* A novel and stable mouse artificial chromosome vector. *ACS Synth. Biol.* **3**, 903–914 (2014).

17. Longo, N. S., Rogosch, T., Zemlin, M., Zouali, M. & Lipsky, P. E. Mechanisms That Shape Human Antibody Repertoire Development in Mice Transgenic for Human Ig H and L Chain Loci. *J. Immunol.* **198**, 3963–3977 (2017).

18. Lefranc, M. P. Nomenclature of the Human Immunoglobulin Heavy (IGH) Genes. *Exp. Clin. Immunogenet.* **18**, 100–116 (2001).

19. Lefranc, M. P. Nomenclature of the human immunoglobulin kappa (IGK) genes. *Exp. Clin. Immunogenet.* **18**, 161–174 (2001).

20. Xu, J. L. & Davis, M. M. Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. *Immunity* **13**, 37–45 (2000).

21. López-Santibáñez-Jácome, L., Avendaño-Vázquez, S. E. & Flores-Jasso, C. F. The pipeline repertoire for Ig-Seq analysis. *Front. Immunol.* **10**, 1–15 (2019).

22. Kim, J. Y., Kurtz, B., Huszar, D. & Storb, U. Crossing the SJL λ locus into κ-knockout mice reveals a dysfunction of the λ1-containing immunoglobulin receptor in B cell differentiation. *EMBO J.* **13**, 827–834 (1994).

23. Gao, S. H., Huang, K., Tu, H. & Adler, A. S. Monoclonal antibody humanness score and its applications. *BMC Biotechnol.* **13**, 1 (2013).

24. Hardy, R. R., Kincade, P. W. & Dorshkind, K. The Protean Nature of Cells in the B Lymphocyte Lineage. *Immunity* **26**, 703–714 (2007).

25. Chase, F., Chase, F., City, I. & City, I. From the Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111;

and the ’Department ofPathology, University of Iouxt College. *Imaging* **173**, (1991).

26. Dunn-Walters, D. K., Isaacson, P. G. & Spencer, J. MGZ of Human Spleen Is a Reservoir of Memory. *J. Exp. Med.* **182**, 559–566 (1995).

27. Kurosawa, N., Yoshioka, M., Fujimoto, R., Yamagishi, F. & Isobe, M. Rapid production of antigen-specific monoclonal antibodies from a variety of animals. *BMC Biol.* **10**, 80 (2012).

28. Akagi, S., Nakajima, C., Tanaka, Y. & Kurihara, Y. Flow cytometry-based method for rapid and high-throughput screening of hybridoma cells secreting monoclonal antibody. *J. Biosci. Bioeng.* **125**, 464–469 (2018).

29. Dussupt, V. *et al.* Potent Zika and dengue cross-neutralizing antibodies induced by Zika vaccination in a dengue-experienced donor. *Nat. Med.* **26**, 228–235 (2020).

30. Setliff, I. *et al.* High-Throughput Mapping of B Cell Receptor Sequences to Antigen Specificity. *Cell* **179**, 1636-1646.e15 (2019).

31. Wilson, M. D. *et al.* Species-specific transcription in mice carrying human chromosome 21. *Science (80-.).* **322**, 434–438 (2008).

32. Chang, H. D., Tokoyoda, K. & Radbruch, A. Immunological memories of the bone marrow. *Immunol. Rev.* **283**, 86–98 (2018).

33. Haas, A., Zimmermann, K. & Oxenius, A. Antigen-Dependent and -Independent Mechanisms of T and B Cell Hyperactivation during Chronic HIV-1 Infection. *J. Virol.* **85**, 12102–12113 (2011).

34. Matsushita, H. *et al.* Species-specific chromosome engineering greatly improves fully human polyclonal antibody production profile in cattle. *PLoS One* **10**, (2015).

35. Kitaura, K. *et al.* Different somatic hypermutation levels among antibody subclasses disclosed by a new next-generation sequencing-based antibody repertoire analysis. *Front. Immunol.* **8**, 1–11 (2017).

36. Kim, B. R. *et al.* Deciphering diversity indices for a better understanding of microbial communities. *J. Microbiol. Biotechnol.* **27**, 2089–2093 (2017).

37. Yermanos, A. D., Dounas, A. K., Stadler, T., Oxenius, A. & Reddy, S. T. Tracing antibody repertoire evolution by systems phylogeny. *Front. Immunol.* **9**, 1–14 (2018).

38. Lipschultz, C. A., Li, Y. & Smith-Gill, S. Experimental design for analysis of complex kinetics using surface plasmon resonance. *Methods* **20**, 310–318 (2000).

Acknowledgements

We thank Y. Sato (Repertoire Genesis, Inc.) for assistance with generating Cricos plots, frequency analyses of VH and VK, and circular dendrograms. We thank Y. Sumida, E. Kaneda, K. Yoshida, M. Fukino, A. Ashiba, Dr. K. Nakamura, and T. Kurosaki at Tottori University and S. Takehara at Trans Chromosomics Inc. for assistance with generating and maintaining TC-mAb mice, and M. Takami, M. Tanaka, K. Hiramatsu, K. Homma, I. Kanazawa, T. Endo and Y. Wang at Trans Chromosomics Inc., and Y. Nagashima, and M. Morimura at Tottori University for assistance with human Ab production and establishing antigen-specific hybridoma cell lines. We also thank Dr. Hiroyuki Kugoh, Dr. Masaharu Hiratsuka, Dr. Tetsuya Ohbayashi, and Dr. Takahito Ohira at Tottori University for critical discussions. This study was supported in part by the Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS) from the Japan Agency for Medical Research and Development (AMED) under Grant Number JP20am0101124 (Y.K.), the Science and Technology Platform Program for Advanced Biological Medicine from AMED under Grant Number JP20am0401002 (Y.K. and K.T.), the Basic Science and Platform Technology Program for Innovative Biological Medicine from AMED under Grant Number

JP18am0301009 (Y.K.), AMED under Grant Number JP20fk0108141 (Y.K.), and JST CREST Grant Number JPMJCR18S4, Japan (Y.K. and K.T.). We thank Jeremy Allen, PhD, from Edanz Group (https://en-author-services.edanzgroup.com/ac) for editing a draft of this manuscript.

Author contributions

Y.K., H.S., and S.A. planned this study; S.A., K.K., and A.O. performed cell culture and mouse production experiments; K.K., and M.O. performed cytogenetic analyses; H.S., H. Tanaka, K.Y., G.H., K.M., and H. Takayama performed mAb production experiments. S.A., Y.N., T.M., S.H. and Y.B. performed FCM analyses and analysed B-cell development; K.T. contributed to the analysis and discussion of the data; H.S., S.A., T.M., and Y.K. analysed the results, and H.S., S.A., T.M., and Y.K. wrote the manuscript, with contributions from each author; M.O. and T.T. supervised the study.

Competing interests

M.O. is a CEO, employee, and shareholder of Trans Chromosomics, Inc. S.A., H. Tanaka, K.M., and H. Takayama are employees of Trans Chromosomics, Inc., and the other authors declare no conflicts of interest.

Materials & Correspondence

Correspondence and material requests should be addressed to Yasuhiro Kazuki.

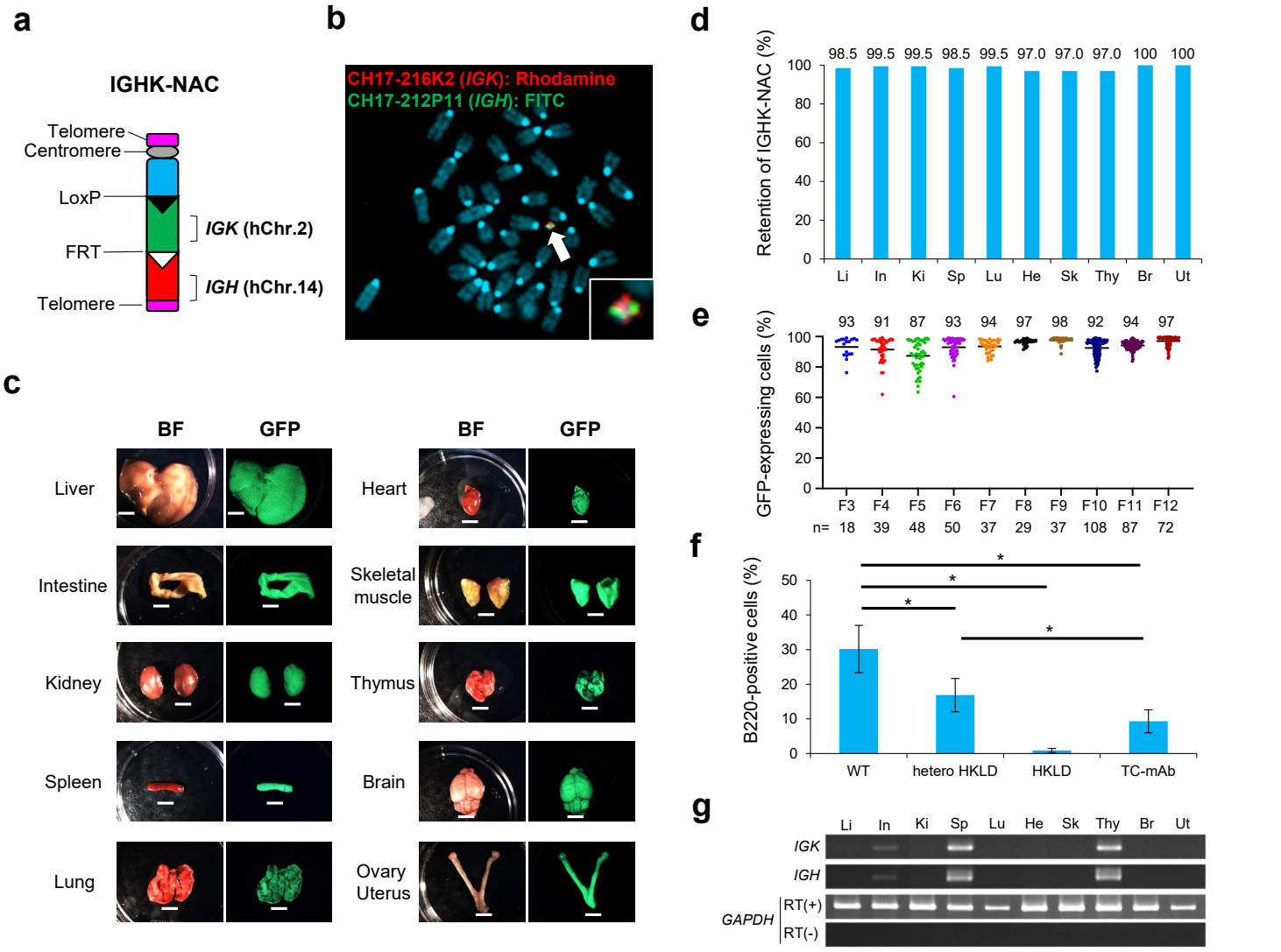


Figure 1| Generation of fully human Ab producing mice.
(a) Representation of the IGHK-NAC structure. Human Ig heavy chain locus (*IGH*) derived from hChr.14 and Ig kappa light chain locus (*IGK*) derived from hChr.2 are represented on a mouse artificial chromosome (MAC). **(b)** Representative image of metaphase FISH analysis with IGK-BAC (CH17-216K2) (red) and IGH-BAC (CH17-212P11) (green) detecting the IGHK-NAC in bone marrow cells from an IGHK-HKLD mouse. Arrow indicates the IGHK-NAC and the inset shows an enlarged image thereof. **(c)** GFP images of different tissues from a Tc mouse carrying the IGHK-NAC. GFP expression indicates the presence of the IGHK-NAC. BF, bright field. Scale bar (5 mm). **(d)** Retention rate of the IGHK-NAC in various Tc mouse tissues analysed by FISH. Li, liver; In, intestine; Ki, kidney; Sp, spleen; Lu, lung; He, heart; Sk, skeletal muscle; Thy, thymus; Br, brain; Ut, uterus. **(e)** Percentage of GFP-expressing cells in peripheral blood mononuclear cells from Tc mice of different generations. **(f)** Percentage of B220-positive cells in the lymphocyte fraction of peripheral blood in age-matched WT (*n=10*), hetero HKLD (*n=10*, *Igh*$^{+/-}$, *Igk*$^{+/-}$ and *Igλ1*$^{+/low}$), HKLD (*n*=10, *Igh*$^{-/-}$, *Igk*$^{-/-}$ and *Igλ1*$^{low/low}$) and Tc-mAb mice (*n*=525). *p<0.001. **(g)** RT-PCR analysis of total RNA from various tissues of the Tc mouse. Li, liver; In, intestine; Ki, kidney; Sp, spleen; Lu, lung; He, heart; Sk, skeletal muscle; Thy, thymus; Br, brain; Ut, uterus.

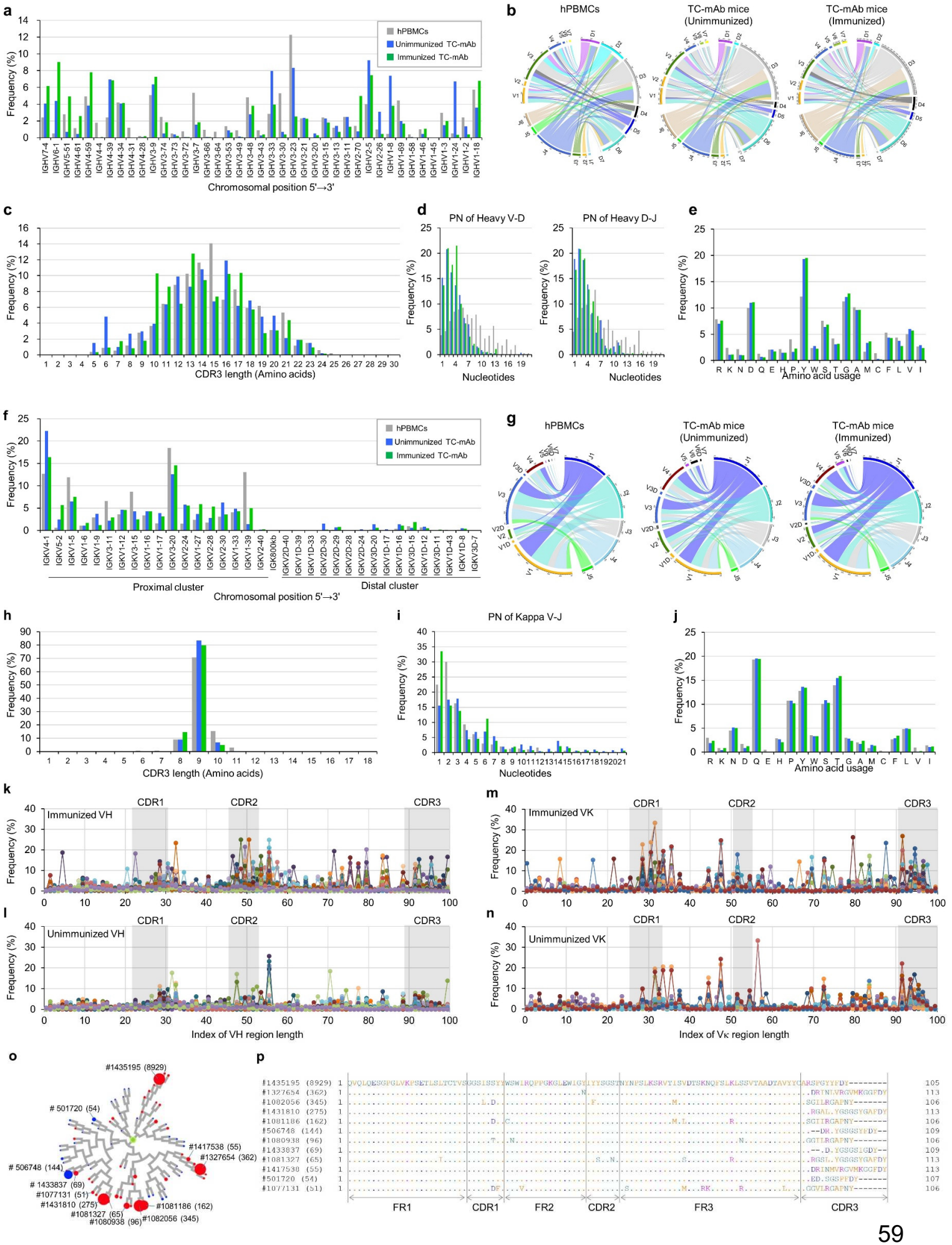

a
hPBMCs
Unimmunized TC-mAb
Immunized TC-mAb
Frequency (%)
Chromosomal position 5'→3'
b
hPBMCs
TC-mAb mice (Unimmunized)
TC-mAb mice (Immunized)
c
CDR3 length (Amino acids)
d
PN of Heavy V-D
PN of Heavy D-J
Nucleotides
e
Amino acid usage
f
Proximal cluster
Distal cluster
g
h
i
PN of Kappa V-J
j
k
Immunized VH
CDR1
CDR2
CDR3
l
Unimmunized VH
Index of VH region length
m
Immunized VK
n
Unimmunized VK
Index of Vκ region length
o
#1435195 (8929)
501720 (54)
#1417538 (55)
#1327654 (362)
506748 (144)
1433837 (69)
#1077131 (51)
#1431810 (275)
#1081186 (162)
#1081327 (65)
#1080938 (96)
#1082056 (345)
p
FR1
CDR1
FR2
CDR2
FR3
CDR3

Figure 2| Repertoire analyses of heavy and light chain variable regions in TC-mAb mice.
(a) Human heavy chain gene utilization in hPBMCs from healthy donors and in TC-mAb mice with and without immunization. Percentage frequency use of V gene segments in healthy human donors (grey), TC-mAb mice with (green), and without immunization (blue) are represented. **(b)** Circos plots comparing VDJ gene association. The gene segments are grouped as subfamilies and are shown together with the first digit of their allele name. Links indicated the relative frequencies of specific VDJ combinations, and wider links indicate higher frequencies of recombination. The relative CDRH3 length **(c)**, P and N nucleotide addition at joining positions **(d)**, and amino acid usage **(e)** within the CDRH3 of productive rearrangements. **(f)** Human VK gene utilization in hPBMCs of healthy donors and in TC-mAb mice with and without immunization. **(g)** Circos plots comparing VJ gene association. Links indicate the relative frequencies of specific VJ combinations. The relative CDRL3 length **(h)**, P and N nucleotide addition at joining positions **(i)**, and amino acid usage **(j)** within the CDRL3 of productive rearrangements. The frequency of SHM at every position of the variable region was calculated as described in the Methods and is represented for VH and VK chains in OVA-immunized TC-mAb mice **(k, m)** and unimmunized TC-mAb mice **(l, n)**. **(o)** The phylogenetic tree of the clone type CLH001. The circular dendrograms of the 20 most frequently used clone lineages were over-laid with the number of copies carrying the same CDRH3 sequences. The leaves of circular dendrograms have a maximum of 50 reads; therefore, the larger the number the larger the circle in the leaves. The colour of the circle indicates with (red) or without (blue) immunization in TC-mAb mice. The amino acid sequences of 12 clonotypes containing more than 50 copies in the CLH001 clone lineage. The clonotype numbers and each copy number are also indicated. (**p**) Amino acid sequence alignment of the 12 clonotype variable regions of CLH001.

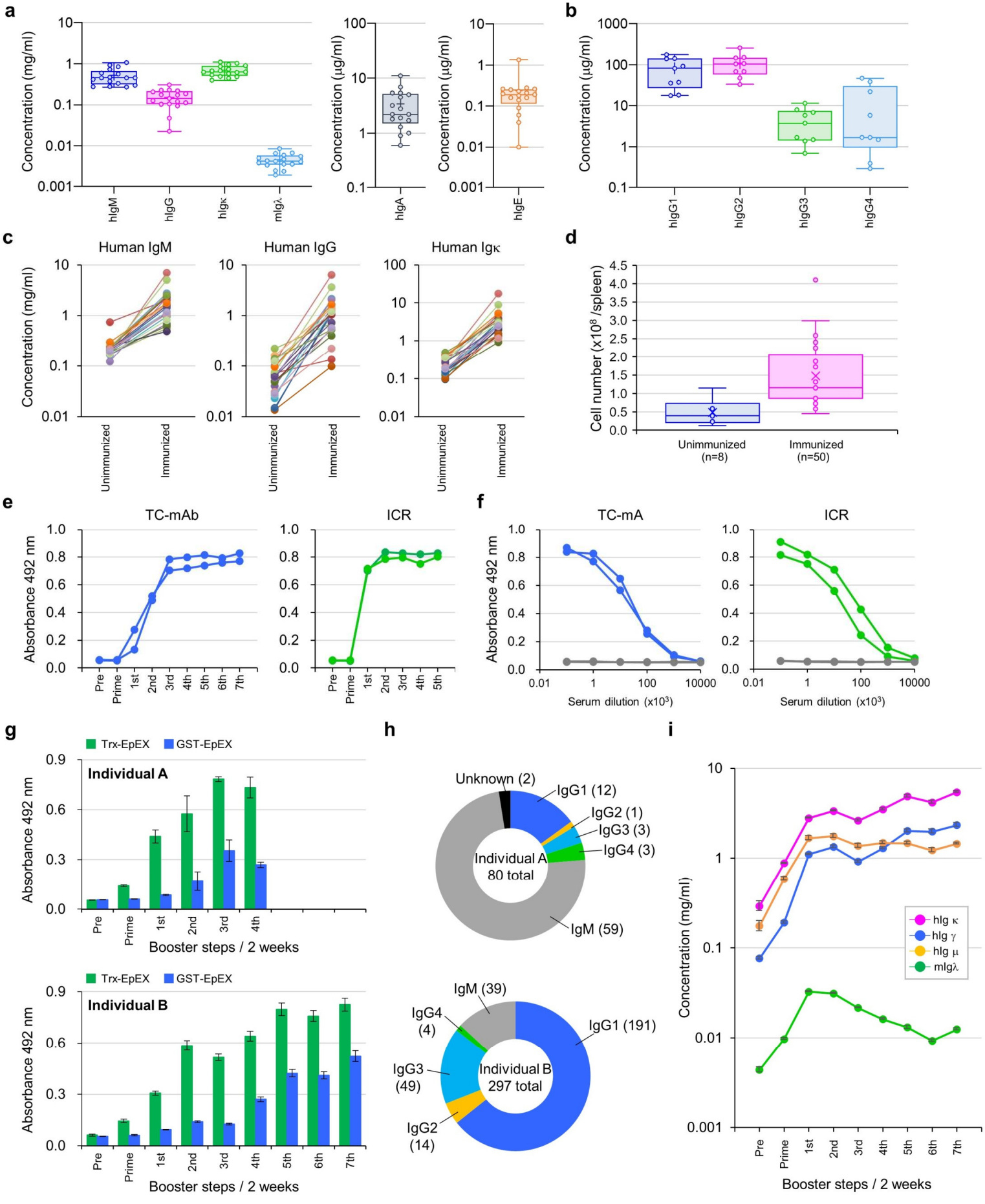

a
Concentration (mg/ml)
hIgM
hIgG
hIgκ
mIgλ
Concentration (μg/ml)
hIgA
Concentration (μg/ml)
hIgE
b
Concentration (μg/ml)
hIgG1
hIgG2
hIgG3
hIgG4
c
Human IgM
Human IgG
Human Igκ
Concentration (mg/ml)
Unimmunized
Immunized
d
Cell number (x10⁸ /spleen)
Unimmunized (n=8)
Immunized (n=50)
e
TC-mAb
ICR
Absorbance 492 nm
Pre
Prime
1st
2nd
3rd
4th
5th
6th
7th
f
TC-mA
ICR
Absorbance 492 nm
Serum dilution (x10³)
g
Trx-EpEX
GST-EpEX
Individual A
Individual B
Absorbance 492 nm
Booster steps / 2 weeks
h
Unknown (2)
IgG1 (12)
IgG2 (1)
IgG3 (3)
IgG4 (3)
Individual A
80 total
IgM (59)
IgM (39)
IgG4 (4)
IgG1 (191)
IgG3 (49)
Individual B
297 total
IgG2 (14)
i
Concentration (mg/ml)
hIg κ
hIg γ
hIg μ
mIgλ
Booster steps / 2 weeks

Figure 3| Production of human Ig in TC-mAb mice .
(a) The serum concentration of different classes of human Igs in TC-mA mice. The concentration of hIgM, hIgG, hIgκ, mIgλ, hIgA and hIgE in TC-mAb mice (*n=17*) was measured. Each symbol represents one serum concentration. **(b)** The serum concentration of human IgG subclasses. The concentrations of IgG1, 2, 3, and 4 were measured in TC-mAb mice (*n=9*). **(c)** The serum concentration of hIgM, IgG, and Igκ before and after immunization of TC-mAb mice. The symbols collected from the same mouse before and after immunization are connected by a line. **(d)** The number of lymphocytes. The numbers of harvested cells were counted under Turk's solution. Each symbol indicates one mouse. **(e)** OVA-specific Ab titre. The titres of Abs against OVA were measured by ELISA using a species-matched secondary Ab. An anti-human IgG-Fc HRP conjugate was used for TC-mAb mice, and an anti-mouse IgG-Fc HRP conjugate was used for ICR. **(f)** Comparison of the anti-sera titres. The anti-sera samples were collected from TC-mAb mice after the seventh booster, and from ICR mice after the fourth booster. Samples were serially diluted by 1/10 from 10^2 to 10^7 and used to determine the OVA-specific titre by ELISA. **(g)** The antigen-specific titre in TC-mAb mice. Individual A (upper graph) was immunized with Trx-EpEX with the prime and four booster administrations. Individual B (lower graph) was immunized with the prime and seven booster administrations. The titres were analysed using the fusion proteins, Trx-EpEX (green) and GST-EpEX (blue). The Trx-EpEX titre indicated induction of antigen-specific Abs, but that of GST-EpEX indicated administration of EpEX-specific Abs. Error bars indicate standard deviation from triplicate measurements. **(h)** Determination of human Ig-classes. The mouse individual and the total number of class-determined clones are indicated in the centre of the circle. **(i)** The serum concentration of different classes of Ab. hIg μ (orange), hIg μ (blue), hIg κ (pink), and mIg λ (green) in the Trx-EpEX immunized TC-mAb mice are represented. Error bars indicate standard deviation from triplicate measurements.

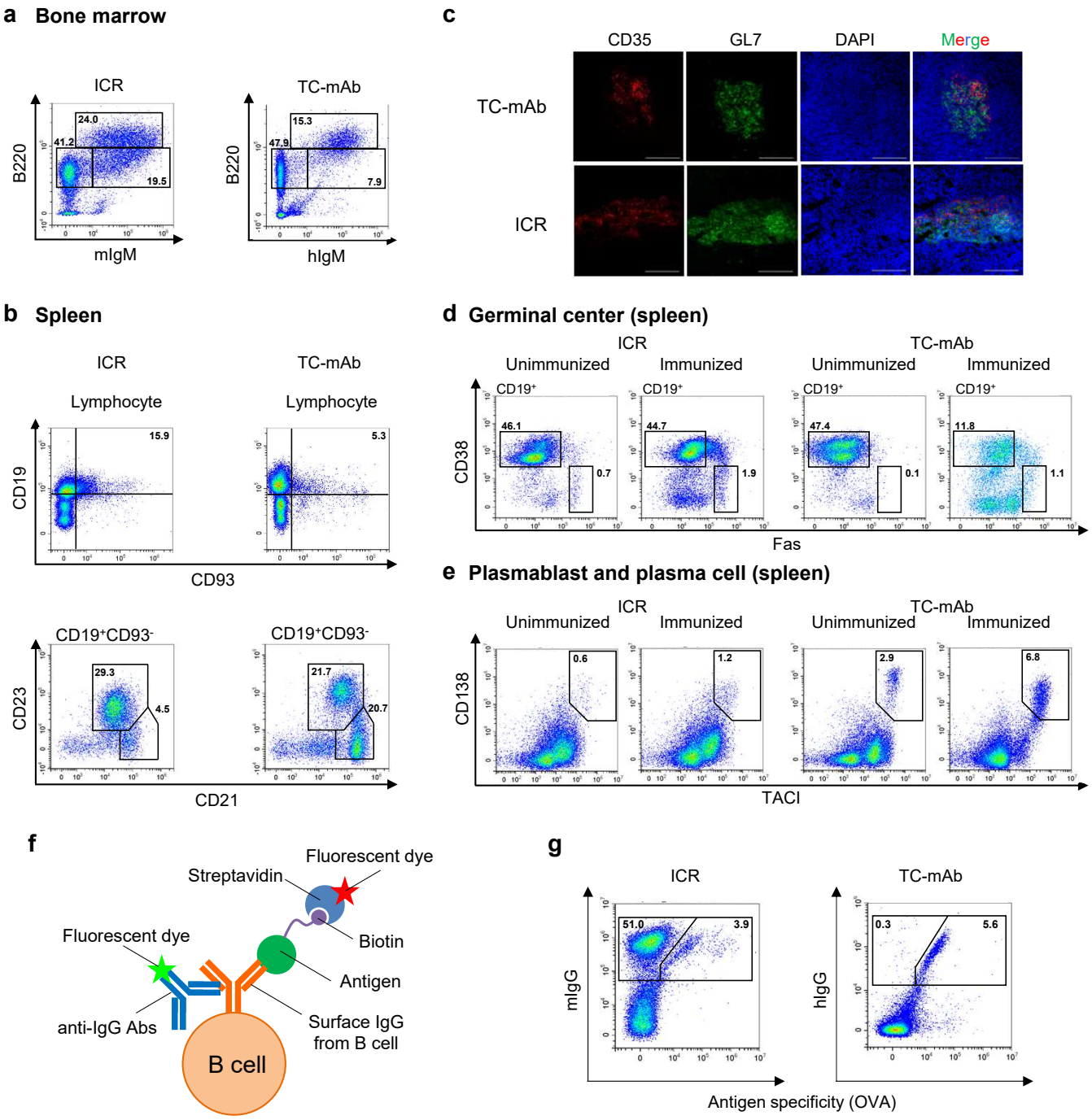


Figure 4| Analysis of B-cell development.
(a) Flow cytometry of B cell subsets of both Pro-B and Pre-B (IgM^-B220^+), immature B (IgM^+B220^+), and mature recirculating B ($IgM^{lo/+}B220^{hi}$) cells in bone marrow. Numbers indicate the percentage of B cells in each subset. The numbers in the panels represent the percentage of lymphocytes. **(b)** The number of cells in the B lineage in the spleen. Flow cytometry of Transitional B ($CD19^+CD93^+$), Follicular B ($CD19^+CD93^-CD21^+CD23^+$), and Marginal zone B ($CD19^+CD93^-CD21^{hi}CD23^-$) cells are represented. **(c)** Representative images of fluorescence immunohistochemistry. The spleens of OVA-immunized TC-mAb mice were harvested and analysed by immunohistochemistry. Sections were stained for follicular dendritic cells (CD35, red), GC B cells (GL7; green), and nuclei (DAPI, blue). The images were taken at ×20 magnification. All scale bars indicate 100 μm. **(d)** The number of cells in the germinal centre-related B lineage in the spleen. Flow cytometry of germinal centre B ($CD19^+CD38^{lo/-}Fas^+$) cells is indicated. **(e)** The number of Ab-producing cells in the spleen. Flow cytometry of both PB and PC ($CD138^+TACI^+$) are represented. The numbers in the panels represent the percentages in the lymphocytes. **(f)** Representation of image to detect antigen-specific Ab expressing B cell subsets. **(g)** Flow cytometry of antigen-specific B cell subsets ($IgG^+antigen^+$). The numbers in the panels represent the percentages of IgG-positive lymphocytes.

Supplementary information

Supplementary Methods

Materials.

CHO (*Hprt*$^{-/-}$) cells (JCRB0218) were obtained from the Japanese Collection of Research Bioresources Cell Bank (JCRB Cell Bank, NIBIOHN, Osaka, Japan). CHO K1 cells (CCL61), DT40 cells (CRL-2111), P3X63Ag8.653 myeloma cells (CRL-1580) and HCT116 cells (CCL-247) purchased from the American Type Culture Collection (ATCC, Manassas, VA, USA). Restriction enzymes and DNA modifying enzymes were purchased from New England Biolabs (Ipswich, MA, USA) and TOYOBO (Tokyo, Japan), respectively. Primers were ordered from Eurofins (Huntsville, AL, USA). *E. coli strains* (DH5α) and [Rosseta-gami B(DE3) pLysS] were purchased from Takara Bio (Shiga, Japan) and Merck Millipore.

Construction of targeting vectors.

Targeting vector for introducing loxP-5′-HPRT into hChr.2. For preparation of the homology arm, DNA obtained from DT40 cells carrying a hChr.2 was used as a PCR template. A 9.5 kb homology arm was amplified using primers: cos138-F6B, 5′-TCG AGG ATC CCA CAT AGA CAT TCA ACC GCA AAG CAG-3′ and cos138-R6B, 5′-TCG AGG ATC CAG GCC CTA CAC ATC AAA AAG TGA AGC AG-3′. The

fragment was cloned into the *Bam*HI site of the pKO Scrambler V901 backbone vector (Lexicon Genetics, Woodlands, TX) (V901-cos138). A PGKHyg-loxP-5′-HPRT fragment obtained by *Asc*I and *Kpn*I treatment was cloned into the *Spe*I site of V901-cos138 by blunt-end ligation (pCos138HL5′-H). Targeting vector for introducing FRT-5′-HPRT into hChr.2. A FRT-5′-HPRT unit with cloning sites was synthesized (pkD9FRT). A CMV-Bsd fragment was inserted into pkD9FRT using *Eco*RI and *Xho*I sites (to give pBsdkD9FRT). A 4.1 kb left arm was amplified with kD-R9La L, 5′-TCG AGC GGC CGC AGG ATC TTT GGG GGA CTG AAT GGG GTG TGC T-3′ and kD-R9La R, 5′-TCG AAC GCG TTG GAA CCC TCA TAC GTT GCT GGT GGA ATG T-3′, and cloned into pBsdkD9FRT using *Not*I and *Mlu*I sites. A 3.2 kb right arm was obtained by PCR using primers: kD-F9 Ra L, 5′-CGA GGA TCC ATT TCT CCA CAT CCT AGC CAA CAC TTG ACA TTT CCT-3′ and kD-F9Ra R, 5′-TCG AGG ATC CGC CAG GGA GAC AGA TGC CAA GTA CGG TTT AG-3′, and inserted to the *Bam*HI site (pBkD9FLR). Targeting vector for introducing FRT-3′-HPRT into hChr.14. DNA from DT40 cells containing a hChr.14 was used as a PCR template for the preparation of the homology arm. An FRT site with cloning sites was synthesized (pSC355FRT). PGKhyg and 3′-HPRT fragments were inserted into pSC355FRT using *Kpn*I/*Cla*I and *Nhe*I/*Mlu*I,

respectively (to give pSC355HF3′H). A 3.8 kb left arm was amplified with primers: NotISC355-F, 5′-TCG AGC GGC CGC GTA CAA TCT TGG ATC ACT ACA ACC TCT GCC TA-3′ and AscISC355-R, 5′-TCG AGG CGC GCC AGG ATT ATA GAT GTG AGC CAT CAC TAA GAC TCC T-3′, and cloned into *Not*I/*Asc*I sites of pSC355HF3′H (to give pSC355HF3′HL). A 4.2 kb right arm was then prepared using primers, SalISC355-F4, 5′-TCG AGT CGA CAG CAC GTT GGG AGG CCA AGG CAG GAG AAT A-3′ and 5′-BamHISC355-R4, 5’-TCG AGG ATC CTG GCT GAC ACA GCC AGT CCC GGA TT-3′ and sub-cloned into pSC355HF3′HL (to give pSC355HF3′HLR).

Modification of hChr.2 and hChr.14 in DT40 cells.

Homologous recombination-proficient chicken DT40 cells (1×10^7) in 0.5 mL RPMI with 25 μg of linearized targeting vector were electroporated at 550 V and 25 μF using a Gene Pulser (Bio-Rad, Hercules, CA, USA). Drug-resistant DT40 clones were selected in 1.5 mg/mL G418, 10 μg/mL blasticidin S or 1.5 mg/mL hygromycin. Homologous recombination in DT40 hybrid clones was identified by PCR analysis using primers described in Supplementary Table 12.

Microcell-mediated chromosomal transfer.

MMCT was performed as described previously[9]. hChr.2-loxPFRT and hChr.14-FRT in DT40 cells were transferred to CHO (MAC) and CHO (*Hprt*$^{-/-}$) cells, respectively, via MMCT. For each transfer, microcell hybrids were selected in medium containing 800 μg/mL G418, 6 μg/mL blasticidin S and 10 μM Ouabain, and 300 μg/mL G418 and 10 μM Ouabain, respectively. CHO IGK-NAC and IGHK-NAC were transferred to CHO (hChr.14-FRT) and CHO K1, and selected with 600 μg/mL G418 and 4 μg/mL blasticidin S, and 800 μg/mL G418, respectively. To transfer the IGHK-NAC to mouse ES cells, CHO K1 cells containing the IGHK-NAC were used as donor microcell hybrids. Briefly, mouse ES cells were fused with microcells prepared from the donor hybrid cells and selected with G418 (250 μg/mL). The transferred IGHK-NAC in mouse ES cells was characterized by PCR and FISH analyses.

DNA transfection.

The Cre-expression vector, pBS185 (Invitrogen), or pCAG-FLPo was transfected into the CHO hybrids containing the MAC vector and the modified hChr.2, or IGK-NAC and modified hChr.14 using Lipofectamine 2000 reagent (Invitrogen)

according to the manufacturer's protocol. After culture for 24 hours in basic growth medium, the cells were cultured in medium containing 1 x HAT (Sigma) and 4–6 μg/mL blasticidin S for selection. Fourteen days later, drug-resistant colonies were picked and expanded for further analyses.

Genomic PCR.

Genomic DNA was extracted from cell lines and Tc mouse tissue specimens using a genomic extraction kit (Gentra System, Minneapolis, MN, USA) and PCR was performed using primers listed in Table S1. Primers for hChr.2 detection were D2S177 F/R, FABP1-F/R, EIF2AK3-F/R, RPIA-F/R, IGKC-F/R, IGKV-F/R, Vk3-2 F/R, and D2S159_1 F/R. Primer pairs for the detection of targeted hChr.2 were cos138 sp L PAGE/cos138 sp R, x6.1 cos RA L/R, kD9 tcLa L/R, and kD9 tcRa L/R. Primer pairs for hChr.14 were MTA1-F3/R3, ELK2P2-F/R, g1(g2)-F/R, CH3F3/CH4R2, and VH3-F/R. Primer pairs for the detection of targeted hChr.14 were 14TarC_La F/R, and 14TarC_Ra F/R. Primer pairs for the detection of recombination junctions were KJneo/PGKr-2, TRANS L1/R1, and PGK-r2/CMVr-1. Mouse *Igκ* and *Igh* KO were confirmed by HKD mCk L1/R1 and mCk L1/R1, and HKD mCmu L1/R1 and mCmu L2/R2, respectively. Igλ low

mutation was confirmed by PCR amplification with mIglc1 VnC L/ J3C1 followed by *Kpn*I digestion. PCR was performed using AmpliTaq Gold (PerkinElmer, Waltham, MA, USA), KOD FX (TOYOBO), or AccuPrime Taq DNA polymerase (Invitrogen, Carlsbad, CA, USA). Amplified fragments were then resolved by electrophoresis on 2% agarose gels followed by staining with ethidium bromide.

IGHK-NAC construction.

Construction of the MAC containing human *IGH* and *IGK* loci involved PCR analyses using primers listed in Supplementary Table 12 and FISH analyses at each step. Human chromosomes 2 and 14 were modified in homologous recombination-proficient chicken DT40 cells for the recombination-mediated translocation. First, a loxP site was inserted proximally to the *IGK* locus on hChr.2p by homologous recombination. The targeting vector was introduced into DT40 cells carrying an intact hChr.2 with a Neo resistance marker by electroporation and drug-resistant clones were obtained in medium containing 1500 μg/mL hygromycin. FISH analysis confirmed independent maintenance of a single copy of hChr.2 containing the loxP unit (DT40 hChr.2loxP) (Supplementary Fig. 2a and 2b). Next, the FRT site was introduced distally to the

IGK locus on hChr.2p. As for the loxP insertion, the targeting vector was introduced into DT40 cells carrying hChr.2loxP and drug-resistant clones were obtained in medium containing 10 μg/mL blasticidin S. FISH analysis confirmed independent maintenance of a single copy of hChr.2 containing an FRT unit (DT40 hChr.2loxPFRT) (Supplementary Fig. 2c and 2d). The modified hChr.2 was transferred into CHO *Hprt*$^{-/-}$ cells carrying the MAC via MMCT[9]. The microcell hybrids were selected in medium containing 800 μg/mL Geneticin and 6 μg/mL blasticidin S. FISH analyses revealed that the MAC and modified hChr.2 were independently and stably maintained in host CHO cells (Supplementary Fig. 3b).

Then, a distal region of hChr.2p from the loxP site including the *IGK* locus was translocated to the MAC by Cre/loxP recombination (Supplementary Fig. 3a). A vector expressing Cre under control of the CMV promoter (pBS185) was transfected by lipofection into CHO cells carrying the MAC and modified hChr.2. An intended reciprocal translocation between the MAC and modified hChr.2 by Cre/loxP recombination caused reconstitution of the *HPRT* gene in the by-product giving HAT resistance, which enabled selection of CHO cell lines carrying the MAC with the *IGK* locus (IGK-NAC) and the by-product. Therefore, drug-resistant clones were obtained by selection in medium containing 1×HAT

and 4 μg/mL blasticidin S. Each recombination junction was detected by PCR analysis and the structure of the IGK-NAC and the by-product was confirmed by FISH analysis (Supplementary Fig. 3b).

Next, we introduced the FRT site proximally to the *IGH* locus on hChr.14q in DT40 cells by homologous recombination (Supplementary Fig. 4a). We did not delete the distal side of the *IGH* locus because the *IGH* locus is located at the very end of hChr.14q. The targeting vector for FRT insertion was introduced by electroporation to DT40 cells carrying an intact hChr.14 with a Neo resistance marker. Drug-resistant clones were obtained by selection in medium containing 1500 μg/mL Hygromycin. FISH analyses confirmed the accurate targeting of hChr.14 in DT40 cells (DT40 hChr.14FRT) (Supplementary Fig. 4b). The modified hChr.14 was transferred from DT40 cells into CHO *Hprt*$^{-/-}$ cells via microcell-mediated chromosomal transfer (MMCT) and microcell hybrids were obtained in selection medium containing 300 μg/mL Geneticin. FISH analysis confirmed that we obtained CHO cells carrying the modified hChr.14 (CHO hChr.14FRT) (Supplementary Fig. 4b). The IGK-NAC was then transferred to CHO hChr.14FRT by MMCT and microcell hybrids were obtained by selection in medium containing 600 μg/mL Geneticin and 6 μg/mL blasticidin S. FISH

analyses confirmed that the IGK-NAC and modified hChr.14 coexisted independently and stably in host CHO cells (Supplementary Fig. 5b). To clone the *IGH* locus into the IGK-NAC, FRT/FLP recombination-mediated reciprocal translocation between the IGK-NAC and modified hChr.14 was performed in CHO *Hprt*$^{-/-}$ cells and HPRT gene reconstruction with the desired product again enabled us to select CHO cells carrying the IGK-NAC with the *IGH* locus (IGHK-NAC) and the by-product. Drug-resistant clones were selected in medium containing 1×HAT and 6 μg/mL blasticidin S. FISH analysis revealed that the IGHK-NAC and the by-product were independently and stably maintained in host CHO cells (Supplementary Fig. 5a and 5b). The resultant IGHK-NAC was transferred to CHO K1 cells to generate donor CHO K1 cells carrying a single desired chromosome, IGHK-NAC, for further MMCT. Microcell hybrids were selected in medium containing 800 μg/mL Geneticin and were monitored by GFP. FISH analysis confirmed that a single copy of IGHK-NAC was independently maintained in host CHO K1 cells (Supplementary Fig. 6a and 6b).

Hybridoma generation.

Spleens and lymph nodes from immunized mice were harvested from euthanized mice, homogenized to single-cell suspensions, and fused with myeloma P3X63Ag8.653 cells using an electro-cell-fusion generator (ECFG21) (Nepagene, Chiba, Japan). Fused hybridoma cells were plated in 96-well plates. After about 14 days of culture, a primary screen of supernatants was performed by ELISA. Hybridoma clones producing EpCAM-specific Abs were identified by ELISA using GST-EpEX following HAT selection. The positive wells were picked and passaged in 96-well plates. Each supernatant was again analysed by ELISA using Tag, Trx-EpEX, and GST-EpEX. Hybridoma clones that reacted with Trx-EpEX and GST-EpEX but not Tag were established by two or more limited dilutions. To produce anti-AMIGO2 mAbs, Trx-AMIGO2-Ig was used as an immunogen, and GST-AMIGO2-Ig was also used for screening AMIGO2-specific mAbs.

Hybridoma screening.

Hybridoma cells producing EpCAM or AMIGO2-specific Abs were identified by ELISA and immunocytochemistry screening. The human colorectal cell line, HCT116, was used for immunocytochemistry screening of anti-EpCAM

mAbs. CHO cells stably transfected with human AMIGO2 were used for screening anti-AMIGO2-specific mAbs. Cultured cells were harvested from a 10 cm dish and resuspended at 2.0 x 10^5 cells/ml. Each well of a 96-well flat-bottom plate (TPP) was seeded with 100 μl of cell suspension. Cells were incubated for 2 days at 37°C in a CO_2 incubator. The culture medium was then removed by aspiration and 100 μl of supernatant containing Abs was added. After incubating for 1 h on ice, plates were washed twice with 150 μl of the medium, and 100 μl of goat anti-Human IgG (H+L) Cross-Adsorbed Secondary Ab (Abcam, Cambridge, UK) diluted 1:400 in the medium was added. Plates were washed twice with 150 μl of the medium, and PBS with 1% v//v FBS was added. Plates were scanned using Keyence BZ-X700 microscopy.

Surface plasmon resonance.

Kinetic analysis was performed using a Biacore T200 (GE Healthcare). Each kinetic run was set up using the kinetic wizard template with six non-zero concentrations (in series), with at least one of the concentrations in duplicate (to check the surface performance), and a zero concentration. A blank immobilized surface was used as a reference surface, which was prepared as described in

the ligand immobilization step but without any ligand. All dilutions were prepared in HBS-EP running buffer (GE Healthcare) at room temperature. Regeneration between each cycle was performed using 10 mM glycine (GE Healthcare) at pH 2.5 for 30 seconds. The sensor chip protein G or CM5 (GE Healthcare) was used to directly capture human Abs of interest. For kinetic analysis of EpCAM Abs, five concentrations of analyte were used (10, 20, 30, 40, and 50 nM). For kinetic analysis of AMIGO2 Abs, five concentrations of analyte were used (6.25, 12,5, 25, 50, and 100 nM). Post-run, the data were evaluated using the 1:1 kinetic binding model in Biacore T200 evaluation software to generate *ka*, *kd*, and K_D[38].

Supplementary Table 1| Data collection of NGS.

Species	**TC-mAb mice**				**hPBMCs**	
Sample collection	Unimmunized *n=5*		Immunized *n=1*		Healthy donor (Asians) 426 male/female, ages: 18-54	
Analysed chain	IgG and IgM	Igκ	IgG and IgM	Igκ	IgG and IgM	Igκ
Original reads	2,811,364	1,398,612	3,567,368	2,540,916	1,662,560	1,483,520
Merged reads	1,256,816	782,060	1,863,712	1,845,272	627,292	1,082,276
Annotated reads	312,705	195,281	465,431	460,851	156,798	270,485
CDR3 reads	280,739	188,313	450,749	438,134	139,745	262,936
Productive reads (%)	263,500 (93.9)	180,629 (95.9)	418,032 (92.7)	419,601(95.8)	127,103 (91.0)	248,872 (94.7)

Merged reads represented the number of NGS reads recovered from preprocessing with merge paired-end reads of original reads. Annotated reads and CDR3 reads were determined as human Ig-sequence identified by IgBlast. Productive reads (%) indicated the percentage of productive amino acid sequence with in-frame and contained no stop codon.

Heavy chain

Supplementary Table 2| Frequency usage of V-gene segments.

	hPBMC	Unimmunized	Immunized
IGHV6-1	0.53	4.39	9.02
IGHV1-2	2.42	1.35	0.41
IGHV1-3	2.96	1.49	2.00
IGHV4-4	1.77	0.04	0.05
IGHV7-4	2.44	4.07	6.16
IGHV2-5	4.02	9.25	7.46
IGHV3-7	5.34	1.55	1.84
IGHV1-8	0.45	7.39	3.80
IGHV3-9	5.09	6.36	7.25
IGHV3-11	2.45	2.45	1.28
IGHV3-13	1.18	1.41	0.68
IGHV3-15	2.40	2.29	1.79
IGHV1-18	5.74	3.61	6.79
IGHV3-20	0.02	0.50	0.28
IGHV3-21	2.32	2.36	2.26
IGHV3-23	12.27	8.34	2.52
IGHV1-24	0.53	6.70	0.34
IGHV2-26	0.98	3.07	0.44
IGHV4-28	0.21	0.07	0.18
IGHV3-30	5.32	0.69	0.38
IGHV4-31	1.19	0.05	0.01
IGHV3-33	2.84	7.98	3.95
IGHV4-34	4.21	4.08	4.14
IGHV4-39	2.41	6.94	6.85
IGHV3-43	0.90	0.27	0.37
IGHV1-45	0.01	0.01	0.03
IGHV1-46	1.00	0.45	1.06
IGHV3-48	4.82	2.80	3.79
IGHV3-49	0.89	0.16	0.06
IGHV5-51	2.80	0.69	4.90
IGHV3-53	1.38	0.87	0.61
IGHV1-58	0.39	0.05	0.01
IGHV4-59	4.92	3.83	7.81
IGHV4-61	1.15	0.44	2.58
IGHV3-64	0.72	0.04	0.01
IGHV3-66	0.95	0.08	0.08
IGHV1-69	4.42	1.98	1.67
IGHV2-70	1.41	0.75	4.97
IGHV3-72	0.74	0.08	0.07
IGHV3-73	0.49	0.41	0.20
IGHV3-74	2.46	0.50	1.83

Supplementary Table 2| Frequency usage of D-gene segments.

	hPBMC	Unimmunized	Immunized
IGHD1-1	2.49	1.47	0.55
IGHD2-2	9.96	1.21	0.78
IGHD3-3	2.73	0.45	0.79
IGHD4-4	0.00	0.00	0.00
IGHD5-5	0.00	0.00	0.00
IGHD6-6	3.03	0.06	0.09
IGHD1-7	0.88	0.22	0.14
IGHD2-8	0.97	0.20	0.10
IGHD3-9	5.41	8.93	6.54
IGHD3-10	11.79	33.46	30.38
IGHD4-11	0.37	0.11	0.24
IGHD5-12	3.81	2.47	2.56
IGHD6-13	9.84	11.96	7.69
IGHD1-14	1.98	1.96	1.26
IGHD2-15	4.50	1.78	0.70
IGHD3-16	3.82	1.87	1.35
IGHD4-17	4.27	2.28	7.05
IGHD5-18	2.92	0.94	2.02
IGHD6-19	6.30	16.77	21.08
IGHD1-20	0.37	0.29	0.24
IGHD2-21	2.23	0.86	2.29
IGHD3-22	7.30	1.29	2.60
IGHD4-23	2.44	1.42	1.67
IGHD5-24	4.64	2.21	0.62
IGHD6-25	0.20	0.08	0.13
IGHD1-26	7.48	5.57	8.01
IGHD7-27	0.24	2.14	1.14

Supplementary Table 2| Frequency usage of J-gene segments.

	hPBMC	Unimmunized	Immunized
IGHJ1	0.68	3.02	3.17
IGHJ2	3.33	5.09	5.10
IGHJ3	4.98	9.96	10.76
IGHJ4	42.61	40.72	52.35
IGHJ5	6.87	5.74	10.47
IGHJ6	41.53	35.48	18.15

(continued)

Kappa chain

Supplementary Table 2| Frequency usage of V-gene segments.			
	hPBMC	Unimmunized	Immunized
IGKV4-1	12.68	22.26	16.41
IGKV5-2	0.60	2.42	5.66
IGKV1-5	11.92	6.52	7.50
IGKV1-6	1.12	1.09	1.71
IGKV1-9	2.93	3.70	1.22
IGKV3-11	6.54	2.18	2.94
IGKV1-12	3.59	4.62	4.59
IGKV3-15	8.71	4.23	2.50
IGKV1-16	3.34	4.32	4.25
IGKV1-17	1.59	3.86	3.13
IGKV3-20	18.47	12.57	14.56
IGKV2-24	1.53	5.75	5.54
IGKV1-27	2.37	3.65	5.86
IGKV2-28	1.73	2.76	5.33
IGKV2-30	3.10	6.21	3.56
IGKV1-33	4.06	4.87	4.31
IGKV1-39	13.05	1.41	4.99
IGKV2-40	0.06	0.18	0.25
800 kbp			
IGKV2D-40	0.00	0.00	0.00
IGKV1D-39	0.00	0.00	0.00
IGKV1D-33	0.00	0.00	0.00
IGKV2D-30	0.00	1.54	0.21
IGKV2D-29	0.12	0.67	0.80
IGKV2D-28	0.00	0.00	0.00
IGKV2D-24	0.05	0.23	0.10
IGKV3D-20	0.19	1.37	0.57
IGKV1D-17	0.00	0.24	0.01
IGKV1D-16	0.36	1.39	1.12
IGKV3D-15	0.92	0.59	1.83
IGKV1D-12	0.67	0.82	0.43
IGKV3D-11	0.04	0.05	0.19
IGKV1D-43	0.14	0.02	0.01
IGKV1D-8	0.12	0.48	0.43
IGKV3D-7	0.00	0.00	0.00

Supplementary Table 2| Frequency usage of J-gene segments.			
Sample	hPBMC	Unimmunized	Immunized
IGKJ1	29.07	37.40	40.56
IGKJ2	29.79	23.88	21.83
IGKJ3	11.12	15.26	12.57
IGKJ4	22.47	18.91	19.13
IGKJ5	7.57	4.56	5.90

The relative usage of human variable regions in hPBMCs and TC-mAb mice was indicated. The genes segments annotated as an open reading frame or pseudogene[18,19] were excluded.

Supplementary Table 3| Data Shannon-Weaver Index (H').

	hPBMCs	Immunized	Unimmunized
HV	3.371	3.129	3.143
HD	3.018	2.305	2.300
HJ	1.386	1.209	1.384
LV	2.594	2.910	3.051
LJ	1.495	1.442	1.452

The Shannon-Weaver diversity indexes were represented based on communication theory. The indexes were calculated using productive reads indicated in Supplementary Table 1.

Supplementary Table 4| The 20 most frequently used clone lineages in heavy chain.

				Annotated Reads			SHM%		Average length of CDR3	
Clone Lineage	V	D	J	Immunized	Unimmunized	Fold Change	Immunized	Unimmunized	Immunized	Unimmunized
CLH001	IGHV4-59*01	IGHD3-10*01	IGHJ4*02	11108	591	18.8	39.3	29.1	10.8	13.6
CLH002	IGHV6-1*01	IGHD3-10*01	IGHJ6*02	8352	741	11.3	43.6	55.9	16.1	16.6
CLH003	IGHV7-4-1*02	IGHD6-19*01	IGHJ4*02	7235	340	21.3	31.7	31.8	12.9	11.1
CLH004	IGHV6-1*01	IGHD6-19*01	IGHJ4*02	6442	1144	5.6	47.9	40.2	9.8	11.7
CLH005	IGHV3-9*01	IGHD4-17*01	IGHJ6*02	6797	68	100.0	42.9	45.6	16.1	16.8
CLH006	IGHV2-5*02	IGHD6-19*01	IGHJ4*02	5603	894	6.3	18.5	20.4	14.4	12.8
CLH007	IGHV4-39*01	IGHD1-26*01	IGHJ4*02	5439	278	19.6	84.3	23.0	10.0	9.7
CLH008	IGHV6-1*01	IGHD3-10*01	IGHJ4*02	4829	596	8.1	44.0	45.1	13.8	12.6
CLH009	IGHV2-70*01	IGHD3-10*01	IGHJ6*02	5050	159	31.8	23.9	27.0	18.6	17.8
CLH010	IGHV2-70*01	IGHD6-19*01	IGHJ6*02	4948	245	20.2	72.5	41.2	18.0	17.0
CLH011	IGHV1-18*01	IGHD1-26*01	IGHJ6*02	4742	179	26.5	38.4	38.5	16.8	17.1
CLH012	IGHV2-5*02	IGHD3-10*01	IGHJ4*02	4858	8594	0.6	59.1	16.3	15.8	12.6
CLH013	IGHV1-18*01	IGHD3-9*01	IGHJ6*02	3921	1237	3.2	41.7	38.7	21.4	19.2
CLH014	IGHV4-59*01	IGHD3-10*01	IGHJ6*02	2763	986	2.8	55.7	31.0	17.8	18.2
CLH015	IGHV1-18*01	IGHD5-12*01	IGHJ6*02	3461	200	17.3	35.1	82.0	18.0	18.1
CLH016	IGHV3-9*01	IGHD3-10*01	IGHJ6*02	2927	1532	1.9	64.4	61.3	17.6	18.3
CLH017	IGHV3-9*01	IGHD6-19*01	IGHJ6*02	1845	939	2.0	67.4	77.2	16.6	19.4
CLH018	IGHV3-33*01	IGHD6-19*01	IGHJ4*02	2610	1009	2.6	52.1	44.7	11.4	12.8
CLH019	IGHV2-5*02	IGHD5-18*01	IGHJ4*02	2799	166	16.9	30.8	18.1	12.9	11.8
CLH020	IGHV3-21*01	IGHD3-10*01	IGHJ6*02	2399	665	3.6	47.1	52.6	19.9	20.5
Average						16.0	47.0	41.0	15.4	15.4

Supplementary Table 4| The 20 most frequently used clone lineages in kappa chain.

				Annotated Reads			SHM%		Average length of CDR3	
Clone Lineage	V	D	J	Immunized	Unimmunized	Fold Change	Immunized	Unimmunized	Immunized	Unimmunized
CLL001	IGKV4-1*01	-	IGKJ1*01	37330	12723	2.9	29.0	35.2	9.0	9.0
CLL002	IGKV3-20*01	-	IGKJ1*01	28397	8806	3.2	70.4	52.9	8.3	9.0
CLL003	IGKV1-5*03	-	IGKJ1*01	17544	4416	4.0	41.6	28.1	8.8	8.6
CLL004	IGKV5-2*01	-	IGKJ2*01	13024	1354	9.6	40.0	33.8	9.0	9.1
CLL005	IGKV4-1*01	-	IGKJ2*01	12437	8426	1.5	41.8	25.6	8.8	9.0
CLL006	IGKV2-24*01	-	IGKJ1*01	10555	4171	2.5	53.6	34.5	8.6	9.0
CLL007	IGKV1-27*01	-	IGKJ1*01	11035	2504	4.4	52.3	45.3	9.0	9.0
CLL008	IGKV3-20*01	-	IGKJ2*01	10693	3618	3.0	71.6	46.5	8.9	8.9
CLL009	IGKV4-1*01	-	IGKJ4*01	8613	4254	2.0	52.1	33.6	9.0	9.0
CLL010	IGKV5-2*01	-	IGKJ1*01	8688	1566	5.5	62.7	33.1	8.4	8.4
CLL011	IGKV1-39*01	-	IGKJ4*01	8209	450	18.2	88.7	37.8	9.0	8.9
CLL012	IGKV1-27*01	-	IGKJ4*01	8017	970	8.3	92.6	55.4	9.0	9.0
CLL013	IGKV3-20*01	-	IGKJ4*01	6929	2694	2.6	82.9	31.6	8.7	8.6
CLL014	IGKV2-28*01	-	IGKJ1*01	5997	644	9.3	78.6	56.8	9.0	9.0
CLL015	IGKV3-20*01	-	IGKJ3*01	6552	1847	3.5	56.7	38.2	8.5	9.0
CLL016	IGKV1-12*01	-	IGKJ1*01	6561	882	7.4	86.5	49.2	9.1	9.0
CLL017	IGKV2-28*01	-	IGKJ3*01	5085	335	15.2	44.6	66.0	9.0	9.0
CLL018	IGKV1-5*03	-	IGKJ2*01	5565	2138	2.6	48.8	27.2	8.7	8.4
CLL019	IGKV1-12*01	-	IGKJ3*01	5049	1440	3.5	55.8	49.3	8.9	9.0
CLL020	IGKV2-30*02	-	IGKJ4*01	5013	1328	3.8	65.0	33.4	9.0	9.0
Average						5.7	60.8	40.7	8.8	8.9

Relative frequency of top 20 clone lineages (CLH001-020 of VH and CLL001-020 of VK) were summarized with the fold change of annotated reads, the average length of CDR3, and the percentage of SHM with at least two mutations in a variable region.

Supplementary Table 5| Serum Ig-levels in TC-mAb mice.

	Average (μg/ml)	SD	Range
Unimmunized TC-mAb mice			
Number of mice=17			
hIgμ	518.1	242.2	270.6-1067.8
hIgγ	153.2	72.6	22.8-306.4
hIgκ	688.6	209.6	396.3-1095.2
hIgα	3.4	2.7	0.6-11.1
hIgε	0.2	0.3	0.0-1.3
mIgμ	21.3	7.1	9.2-36.1
mIgγ	11.5	11.2	0.9-45.7
mIgκ	29.7	9.3	20.6-57.8
mIgλ	4.6	1.7	2.0-8.6
Number of mice=9			
hIgγ1	82.2	59.3	17.8-176.0
hIgγ2	109.3	68.9	33.3-257.7
hIgγ3	4.6	3.8	0.7-11.6
hIgγ4	13.3	18.3	0.3-47.2
Immunized TC-mAb mice			
Number of mice=21			
hIgμ	2,097	1,396.0	485.5-5534.2
hIgγ	1,322	1,044.8	318.5-5070.4
hIgκ	4,122	2,980.9	923.0-14103.6
Unimmunized ICR mice			
Number of mice=12			
mIgμ	248.0	130.8	107.8-511.3
mIgγ	993.4	992.6	141.3-1779.8
mIgκ	637.0	546.5	124.6-1842.7
OVA-immunized TC-mAb mice			
Number of mice=5			
hIgμ	2,457	1,496.6	1281.8-5041.6
hIgγ	1,087	605.0	44.6-1573.8
OVA-immunized ICR mice			
Number of mice=8			
mIgμ	636.1	320.7	191.1-1208.9
mIgγ	12,835	3,679.0	3679.0-19597.5

Supplementary Table 6| Monoclonal antibody production in mice.

Mouse	TC-mAb		Balb/c
	Individual A	Individual B	
Immunization	Primary and 4 boosters	Primary and 7 boosters	Primary and 4 boosters
Lymphocytes (cells/mouse)	1.87×10^8	2.40×10^8	3.0×10^8
Lymphocytes/cell fusion (cells)	0.9×10^8	0.9×10^8	1.5×10^8
ELISA positive well after HAT selection (well)	303	428	122
ELISA positive well after second screening (clones)	80	349	52
Determination of subclass (clones)	80	297	N.D.
Bound to native antigen (clones)	20	52	14
Cloned by limiting dilution (clones) (Success rate of cloning)	20 (100%)	50 (96.1%)	14 (100%)

N.D. : not determined

Production and screening steps of hybridoma cell lines producing anti-EpCAM mAbs are summarized. The numbers of hybridoma cells (clones) are indicated when Balb/c and TC-mAb mice were immunized with the Trx-EpEX protein as antigen. A portion of the harvested lymphocytes was used for cell fusion, as indicated in the 'lymphocytes/cell fusion' row. After selection with HAT medium, the antigen-positive wells, which contained a few of hybridoma cell colonies, were screened and hybridoma cells were picked from those wells. After a second screening to check the antigen-specific binding of Abs, positive hybridoma cells were analysed for subclass determination. After analysing the binding activity of Abs to native EpCAM on the cell surface of human HCT116 cells, the positive mAb-producing hybridoma cells were established by the limiting dilution procedure. The number of established clones is indicated in the last row.

Supplementary Table 7| K_D values of obtained mAbs.

Target protein	Clone name	ka (1/Ms)	kd (1/s)	KD (M)
EpCAM	hZAK1C008	7.60×10^{4}	4.88×10^{-3}	6.42×10^{-8}
	hZAK1C015	1.92 x 104	5.18 x 10-4	2.70×10^{-8}
	hZAK2C018	12.70×10^{4}	11.29×10^{-3}	8.85×10^{-8}
	hZAK3C014	2.51×10^{4}	5.13×10^{-4}	2.05×10^{-8}
	hZAK3C101	8.10×10^{4}	13.93×10^{-3}	1.71×10^{-8}
AMIGO2	hTNK1C006	5.77×10^{4}	5.11×10^{-4}	8.86×10^{-9}
	hTNK1C017	2.93×10^{4}	2.67×10^{-4}	9.11×10^{-9}
	hTNK1C032	3.57×10^{4}	2.45×10^{-4}	6.86×10^{-9}
	hTNK1C041	2.20×10^{4}	2.92×10^{-4}	1.33×10^{-8}
	hTNK1C099	2.21×10^{4}	3.43×10^{-4}	1.55×10^{-8}

K_D values of mAbs isolated from TC-mAb mice immunized with recombinant protein EpCAM or AMIGO2, determined by SPR. The recombinant proteins containing the whole region of the extracellular domain were applied to the Ab captured sensor chip, respectively. And the K_D values were determined by kinetic titration procedure.

Supplementary Table 8| Antibody reagents used in experiments.

Antibody	Target species	fluorochrome	Dilution	Supplier	Clone
B220	Mouse	BV650	1:100	Biolegened	RA3-6B2
CD19	Mouse	BUV395	1:100	BD Biosciences	1D3
IgM	Human	PE/CF594	1:100	BD Biosciences	G20-127
	Mouse	PE/CF594	1:100	BD Biosciences	R6-60.2
IgD	Human	PE/Cy7	1:100	Biolegened	IA6-2
	Mouse	PE/Cy7	1:100	Biolegened	11-26c.2a
CD93	Mouse	PE	1:100	Biolegened	AA4.1
CD21	Mouse	PE/Cy7	1:100	Biolegened	7E9
CD23	Mouse	BV650	1:100	BD Biosciences	B3B4
CD38	Mouse	BV650	1:100	BD Biosciences	90/CD38
Fas	Mouse	PE/CF594	1:100	BD Biosciences	Jo2
CD138	Mouse	PE/Cy7	1:100	Biolegened	281-2
TACI	Mouse	PE	1:100	Biolegened	8F10
CD4	Mouse	PE	1:100	Biolegened	GK1.5
PD-1	Mouse	PE/Cy7	1:100	Biolegened	29F.1A12
CXCR5	Mouse	BV650	1:100	Biolegened	L138D7
IgG	Human	PE/Cy7	1:100	Biolegend	M1310G05
	Mouse	PE/Cy7	1:100	Biolegend	Poly4053
CD35	Mouse	biotin	1:100	BD Biosciences	8C12
GL7	Mouse	Alexa Fluor 647	1:100	Biolegend	GL7

Supplementary Table 9| B cell development (Unimmunized state).

Mice	Antibody combination	ICR (n=5) (±SD)	TC-mAb mice (n=5) (±SD)
Bone marrow			
Total cells (x10^6)	FSC, SSC	38.1 (±12.7)	20.0 (±0.8)
Lymphocytes (x10^6)	FSC-W, FSC-H	5.8 (±1.3)	1.5 (±0.7)
% Lymphocyte			
Pro-B and Pre-B	$B220^+IgM^-$	42.5 (±7.3)	39.8 (±7.1)
Immature B	$B220^+IgM^+$	27.8 (±2.8)	14.1 (±1.7)
Recirculating B	$B220^{hi}IgM^+$	11.7 (±3.7)	11.1 (±6.0)
Spleen			
Total cells (x10^6)	FSC, SSC	132.0 (±3.2)	27.0 (±0.8)
Lymphocytes (x10^6)	FSC-W, FSC-H	85.0 (±23.5)	16.4 (±5.6)
% Lymphocyte			
Transitional B	$CD19^+CD93^+$	15.5 (±8.3)	5.3 (±2.9)
Follicular B	$CD19^+CD93^-CD21^+CD23^+$	31.3 (±6.3)	15.6 (±4.1)
Marginal zone B	$CD19^+CD93^-CD21^{hi}CD23^-$	8.7 (±5.1)	8.9 (±6.5)

Data represent the average ± SD of five independent experiments from different animals.

Supplementary Table 10| Data collection of B cell development (Comparison between unimmunized and immunized state).

Mice	Antibody combination	ICR		TC-mAb mice	
		Unimmunized (n=5)	Immunized (n=5)	Unimmunized (n=5)	Immunized (n=5)
Spleen (Germinal center, Plasma cell)					
Total cells (x10^6)	FSC, SSC	132.0 (±3.2)	429.0 (±6.5)	27.0 (±0.8)	122.0 (±5.4)
Lymphocytes (x10^6)	FSC-W, FSC-H	84.6 (±25.1)	231.6 (±28.9)	17.0 (±6.6)	54.4 (±18.6)
%Lymphocyte					
Naive B	$CD19^+CD38^{hi}Fas^-$	50.8 (±6.0)	43.5 (±5.9)	32.9 (±8.0)	20.5 (±5.4)
Germinal center B	$CD19^+CD38^{lo/-}Fas^+$	0.9 (±1.1)	3.6 (±2.1)	0.5 (±0.5)	0.9 (±0.5)
Plasma blast, plasma cell	$CD138^+$, $TACI^+$	0.8 (±0.2)	2.4 (±1.0)	3.9 (±1.1)	6.5 (±2.7)
Helper T	$CD4^+$	25.7 (±4.3)	26.7 (±2.8)	38.8 (±6.7)	48.9 (±6.9)
Follicular helper T	$CD4^+CXCR5^+PD1^+$	0.9 (±0.9)	2.8 (±1.2)	0.8 (±0.7)	1.1 (±0.3)
Antigen specific B cell					
Spleen					
Total cells (x10^6)	FSC, SSC	132.0 (±3.2)	429.0 (±6.5)	27.0 (±0.8)	122.0 (±5.4)
IgG producing cell (x10^6)	IgG^+	49.1 (±16.5)	128.2 (±20.7)	0.2 (±0.1)	2.0 (±1.3)
%IgG-positive cells					
Non-specific cell	IgG^+OVA^-	98.5 (±0. 4)	93.9 (±2)	56.6 (±9.9)	27.7 (±21.3)
OVA-specific cell	IgG^+OVA^+	1.5 (±0.4)	6.1 (±2)	43.4 (±9.9)	72.3 (±21.3)

Data represent the average ± SD of five independent experiments from different animals.

Supplementary Table 11| Characteristics of TC-mAb mice for antibodies production.

	Wild type mice	TC-mAb mice	Double Tc mice	KM mouse	Tc cattle	XenoMouse
Carrier of human Ig locus	—	IGHK-NAC	HCF14 and HCF2	HCF14 and *IGK* gene mini-locus	istHAC	non-targeting insertion
human Ig-genes	—	Full length of *IGH* and *IGK*	Full length of *IGH* and *IGK*	Full length of *IGH* and partial *IGK*	Full length of *IGH*, *IGK*, and *IGL*	Partial length of *IGH* and *IGK*
Stability of Ig-genes	Stable (endogenous)	Stable	Instable	Instable (*IGH*)	Stable	Stable
IgG concentration in anti-sera	Normal	Decrease	Decrease	—	Normal	Decrease
Immune response	Normal	Delayed	—	—	—	—
Hybridoma production frequency	Normal	High	One-tenth of normal	Improved against Double Tc mice	—	—
Antibody titre	Normal	Normal	Normal or weak	—	Normal	Normal or weak
IgG subclass (unimmunized)	Normal	IgG1≈2>3,4	IgG1>2>3,4	—	—	—
IgG subclass (Immunized with protein)	Normal	IgG1>3>2>4	—	—	—	—
Human Ig repertoires	—	Closely similar to hPBMCs	—	—	—	Some differences
Spleen size and number of lymphocytes	Normal	Decrease	—	—	—	—
B cell development	Normal	Adverse	Adverse	—	—	Adverse
Antigen-specific B cells frequency	Normal	High	—	—	—	—

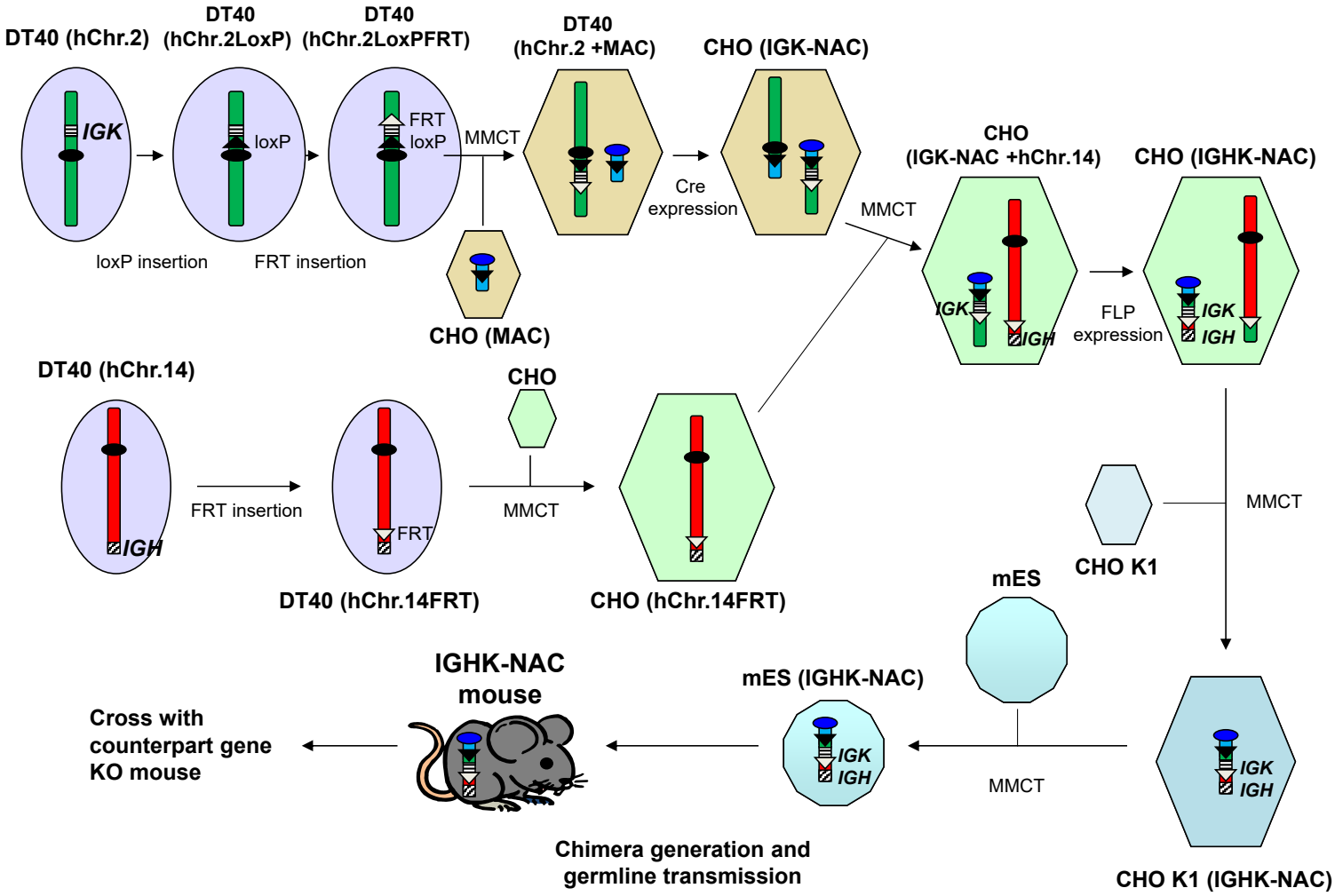
-: There is no known indication.

Supplementary Table 12| Primer list.

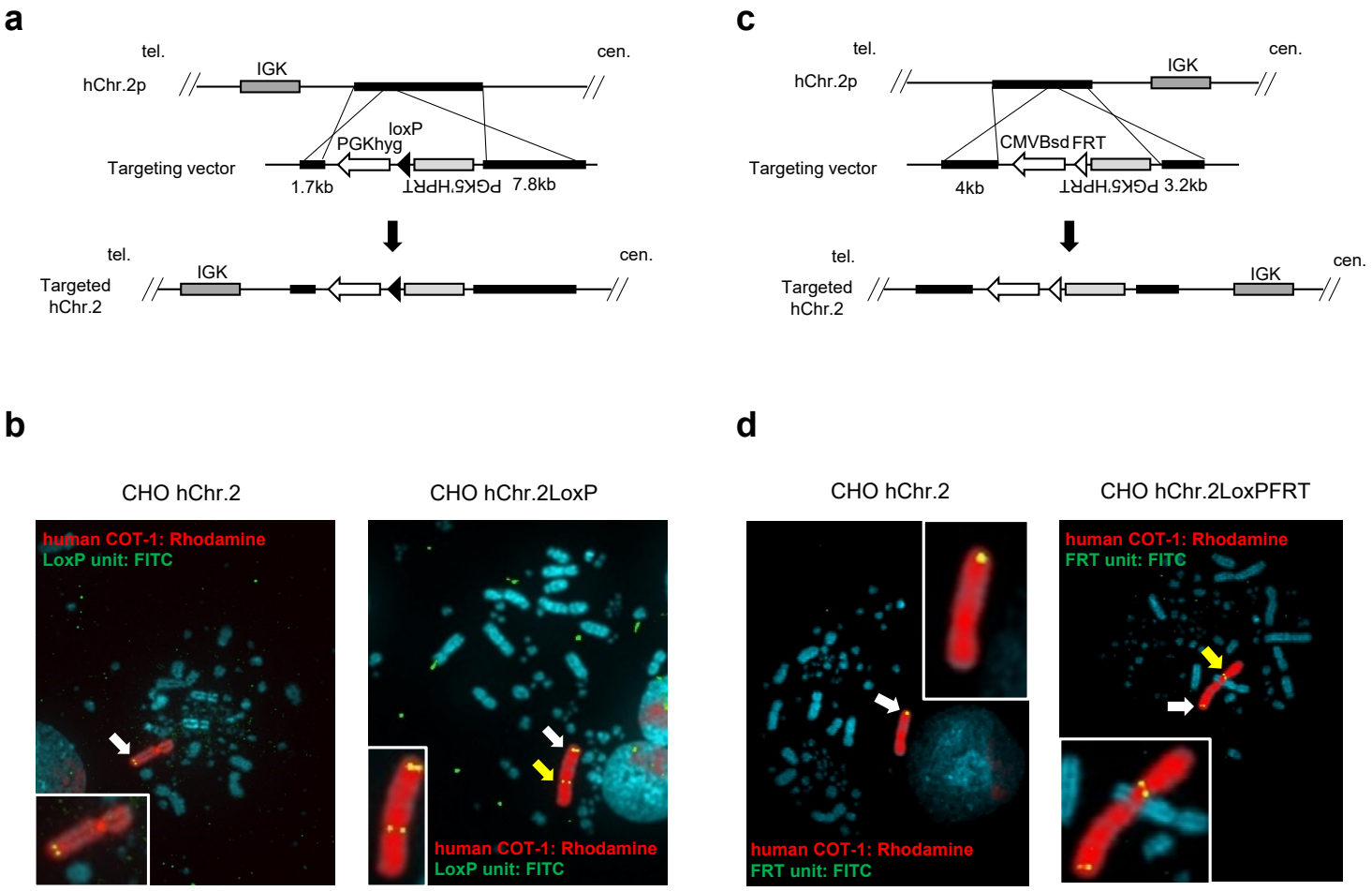
	Gene or aim	Primer name (forward)	Forward primer (5'-3')	Primer name (reverse)	Reverse primer (5'-3')	Product size
Genomic PCR	hChr.2: arm for loxP insertion	cos138-F6B	TCGAGGATCCCACATAGACATTCAACCGCAAAGCAG	cos138-R6B	TCGAGGATCCAGGCCCTACACATCAAAAAGTGAAGCAG	9.5 kb
	hChr.2: loxP targeting check	cos138 sp L	CTGAGAAGAGTCATTGTTTATGGTAGACT	cos138 sp R	ATCCCCATGTGTATCACTGGCAAACTGT	4,143 bp
	hChr.2: loxP targeting check	x6.1cosRa L	GGGGAATAAACACCCTTTCCAAATCCTC	x6.1cosRa R	ACCAAGTAACCGATCAAACCAACCCTTG	9,117 bp
	D2S177	D2S177 F	AGCTCAGAGACACCTCTCCA	D2S177 R	CTGTATTAGGATACTTGGCTATTGA	292 bp
	FABP1	FABP1-F	TATCAAGGGGGTGTCGGAAATCGTG	FABP1-R	ACTGGGCCTGGGAGAACCTGAGACT	444 bp
	EIF2AK3	EIF2AK3-F	AGGTGCTGCTGGGTGGTCAAGT	EIF2AK3-R	GCTCCTGCAAATGTCTCCTGTCA	695 bp
	RPIA	RPIA-F	CTTACCCAGGCTCCAGGCTCTATT	RPIA-R	CTCTACCTCCCTACCCCATCATCAC	631 bp
	IGKC	IGKC-F	TGGAAGGTGGATAACGCCCT	IGKC-R	TCATTCTCCTCCAACATTAGCA	377 bp
	IGKV	IGKV-F	AGTCAGGGCATTAGCAGTGC	IGKV-R	GCTGCTGATGGTGAGAGTGA	156 bp
	IGKV	Vk3-2 F	CTCTCCTGCAGGGCCAGTCA	Vk3-2 R	TGCTGATGGTGAGAGTGAACTC	169 bp
	D2S159_1	D2S159_1 F	CTCTAACTGAATCAAGGGAATGAAC	D2S159_1 R	AGCAGTTTGAGTTTAGGATGAAGG	201 bp
	hChr.2: arm for FRT insertion	kD-R9La L	TCGAGCGGCCGCAGGATCTTTGGGGGACTGAATGGGGTGTGCT	kD-R9La R	TCGAACGCGTTGGAACCCTCATACGTTGCTGGTGGAATGT	4.0 kb
	hChr.2: arm for FRT insertion	KD-F9Ra L	CGAGGATCCATTTCTCCACATCCTAGCCAACACTTGACATTTCCT	KD-F9Ra R	TCGAGGATCCGCCAGGGAGACAGATGCCAAGTACGGTTTAG	3.2 kb
	hChr.2: FRT targeting check	kD9 tcLa L	TGAGAACACAGGGGTCTCCATTCTGACT	kD9 tcLa R	ACAATCAACAGCATCCCCATCTCTGAAG	4,951 bp
	hChr.2: FRT targeting check	kD9 tcRa L	GACGTGCTACTTCCATTTGTCACGTCCT	kD9 tcRa R	TGGTCACTGAAGCTTTCCATCTGCTCTT	3,538 bp
	Cre-loxP recombination	TRANSL1	TGGAGGCCATAAACAAGAAGAC	TRANSR1	CCCCTTGACCCAGAAATTCCA	409 bp
	Cre-loxP recombination	kj neo	CATCGCCTTCTATCGCCTTCTTGACG	PGKr-2	ATCTGCACGAGACTAGTGAGACGTGCTA	~600 bp
	hChr.14: arm for FRT insertion	NotISC355-F	TCGAGCGGCCGCGTACAATCTTGGATCACTACAACCTCTGCCTA	AscISC355-R	TCGAGGCGCGCCAGGATTATAGATGTGAGCCATCACTAAGACTCCT	3.8 kb
	hChr.14: arm for FRT insertion	SalISC355-F4	TCGAGTCGACAGCACGTTGGGAGGCCAAGGCAGGAGAATA	BamHISC355-R4	TCGAGGATCCTGGCTGACACAGCCAGTCCCGGATT	4.2 kb
	hChr.14: FRT targeting check	14TarC_La F	AGCAATTAGGGCCTGTGCATCTCACTTT	14TarC_La R	CCAGCTCATTCCTCCCACTCATGATCTA	4,151 bp
	hChr.14: FRT targeting check	14TarC_Ra F	CATCTGGAGTCCTATTGACATCGCCAGT	14TarC_Ra R	CTTATTCCTCCTTCTGCCCACCCTTCAT	5,026 bp
	MTA1	MTA1-F3	AGCACTTTACGCATCCCAGCATGT	MTA1-R3	CCAAGAGAGTAGTCGTGCCCCTCA	486 bp
	ELK2P2	ELK2P2-F	CCCACTTTACCGTGCTCATT	ELK2P2-R	ATGAAGGTCCGTGACTTTGG	540 bp
	IGHG1	g1(g2)-F	ACCCCAAAGGCCAAACTCTCCACTC	g1(g2)-R	CACTTGTACTCCTTGCCATTCAGC	520 bp
	IGHV3-74	VH3-F	AGTGAGATAAGCAGTGGATG	VH3-R	CTTGTGCTACTCCCATCACT	247 bp
	IGHM	CH3F3	AGGCCAGCATCTGCGAGGAT	CH4R2	GTGGCAGCAAGTAGACATCG	326 bp
	Flp-FRT recombination	TRANSL1	TGGAGGCCATAAACAAGAAGAC	TRANSR1	CCCCTTGACCCAGAAATTCCA	409 bp
	Flp-FRT recombination	PGKr-2	ATCTGCACGAGACTAGTGAGACGTGCTA	CMVr-1	CCTATTGGCGTTACTATGGGAACATACG	444 bp
RT-PCR	IgK	Vk1BACK	GACATCCAGCTGACCCAGTCTCC	Ck	CAGAGGCAGTTCCAGATTTC	
	IgM	VH4BACK	CAGGTGCAGCTGCAGGAGTCGGG	Cmu-1	CAGGAGAAAGTGATGGAGTC	
	GAPDH	RPC1	CCATCTTCCAGGAGCGAGA	RPC2	TGTCATACCAGGAAATGAGC	722 bp

Supplementary Table 13| Antibodies for determining serum immunoglobulin concentrations.

		Antibody	Labeling	Dilution	Supplier	Clone
hIgμ	Capture	Goat anti-Human IgM	-	1:100	Bethyl Laboratories	Goat polyclonal A80-100A-11
	Detector	Goat anti-Human IgM	HRP	1:75000	Bethyl Laboratories	Goat polyclonal A80-100P-37
hIgγ	Capture	Goat anti-human IgG-Fc	-	1:100	Bethyl Laboratories	Goat polyclonal A80-104A-9
	Detector	Goat anti-human IgG-Fc	HRP	1:150000	Bethyl Laboratories	Goat polyclonal A80-104P-90
hIgκ	Capture	Goat anti-Human Ig kappa	-	1:300	Bethyl Laboratories	Goat polyclonal A80-115A-6
	Detector	Goat anti-Human Ig kappa	HRP	1:150000	Bethyl Laboratories	Goat polyclonal A80-115P-43
hIgα	Capture	Goat anti-Human IgA	-	1:100	Bethyl Laboratories	Goat polyclonal A80-102A-6
	Detector	Goat anti-Human IgA	HRP	1:75000	Bethyl Laboratories	Goat polyclonal A80-102P-26
hIgε	Capture	Goat anti-Human IgE	-	1:100	Bethyl Laboratories	Goat polyclonal A80-108A-15
	Detector	Goat anti-Human IgE	HRP	1:75000	Bethyl Laboratories	Goat polyclonal A80-108P-35
mIgμ	Capture	Goat anti-Mouse IgM	-	1:100	Bethyl Laboratories	Goat polyclonal A90-101A-22
	Detector	Goat anti-Mouse IgM	HRP	1:75000	Bethyl Laboratories	Goat polyclonal A90-101P-34
mIgγ	Capture	Goat anti-Mouse IgG-Fc	-	1:100	Bethyl Laboratories	Goat polyclonal A90-131A-16
	Detector	Goat anti-Mouse IgG-Fc	HRP	1:100000	Bethyl Laboratories	Goat polyclonal A90-131P-38
mIgκ	Capture	Goat anti-Mouse Ig kappa	-	1:100	Bethyl Laboratories	Goat polyclonal A90-119A-14
	Detector	Goat anti-Mouse Ig kappa	HRP	1:100000	Bethyl Laboratories	Goat polyclonal A90-119P
mIgλ	Capture	Goat anti-Mouse Ig lambda	-	1:100	Bethyl Laboratories	Goat polyclonal A90-121A-12
	Detector	Goat anti-Mouse Ig lambda	HRP	1:100000	Bethyl Laboratories	Goat polyclonal A90-121P-18

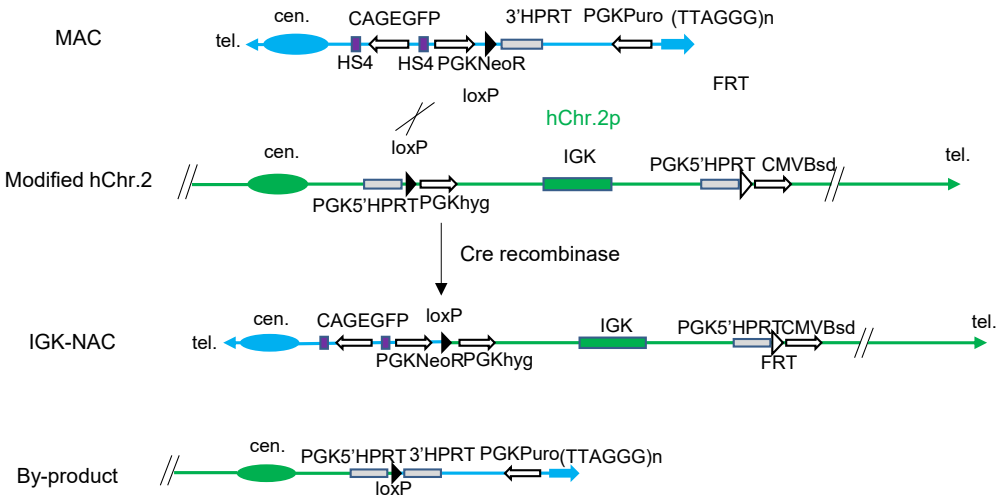


Supplementary Figure 1| Schematic depiction of the generation of fully human Ab-producing mice via chromosome engineering.
Chromosome modifications were performed in chicken DT40 cells carrying the entire hChr.2 *IGK* locus or hChr.14 *IGH* locus. Following IGK-NAC construction and its transfer into CHO cells carrying a hChr.14, the IGHK-NAC was constructed. The IGHK-NAC was transferred to mES cells through CHO K1 cells to produce the chimeric mice. Mice carrying the IGHK-NAC were mated with endogenous Ig-gene knockout mice to produce fully human Ab-producing mice.

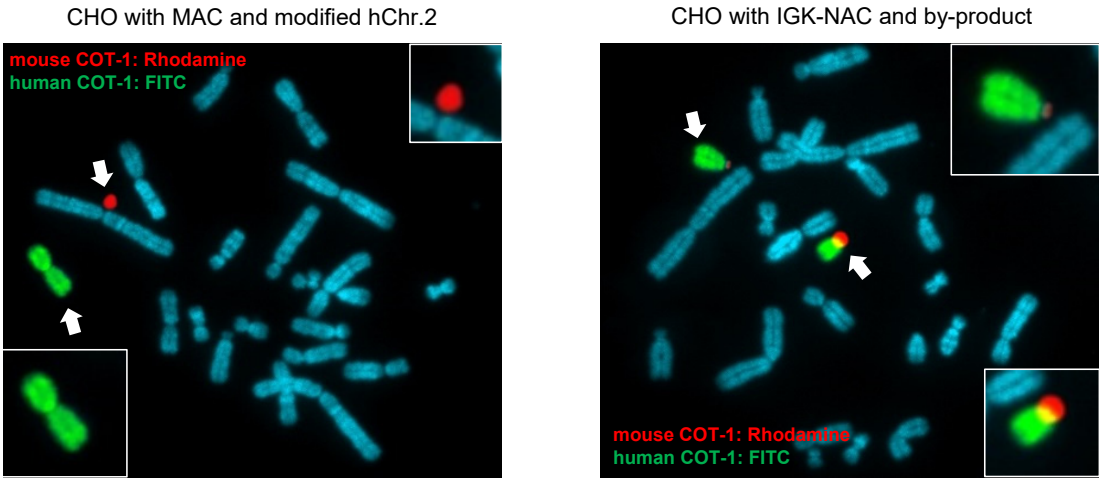


Supplementary Figure 2| Modification of hChr.2 in chicken DT40 cells.
(a) Schematic representation of the targeted integration of the loxP site into the proximal side of the IGK locus on hChr.2p. **(b)** FISH analysis detecting the loxP unit integrated into hChr.2 in DT40 cells carrying an intact hChr.2 (Left) and loxP (Right). Red indicates hChr.2. White and yellow arrows indicate the neomycin resistance marker and the loxP unit, respectively. Insets show enlarged images. **(c)** Strategy for the targeted integration of the FRT site into hChr.2. **(d)** FISH analysis detecting the FRT unit integrated into hChr.2 in DT40 cells carrying hChr.2loxP (Left) and DT40 cells carrying hChr.2loxPFRT (Right). Red indicates hChr.2. White and yellow arrows indicate the neomycin resistance marker and the FRT unit, respectively. Insets show enlarged images.

a

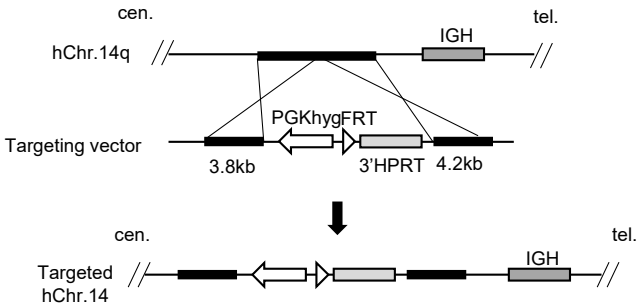


b

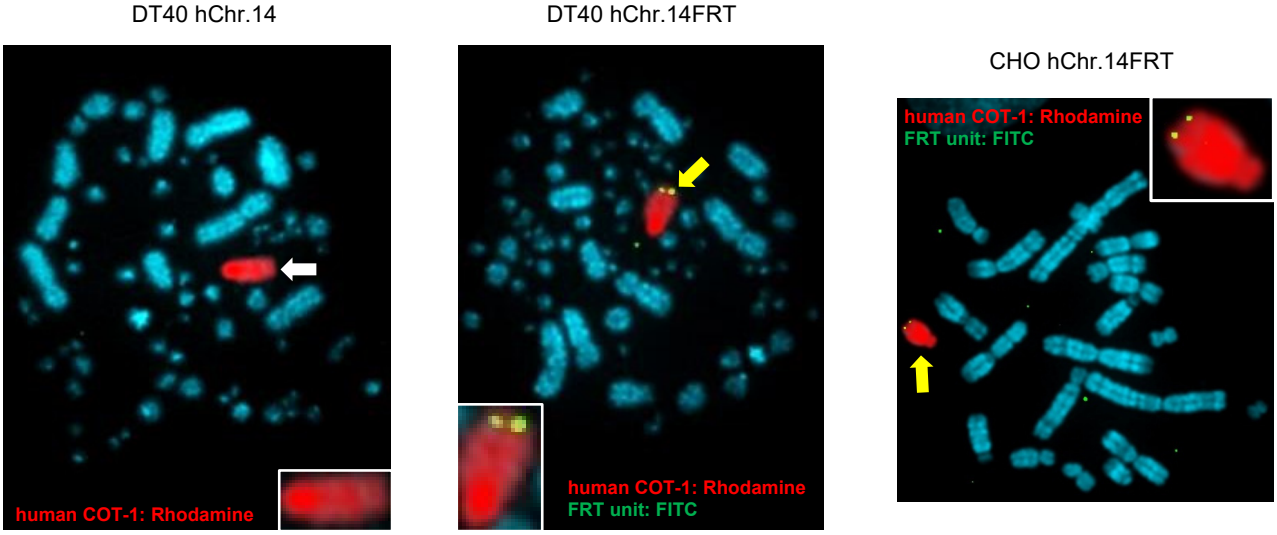


Supplementary Figure 3| Construction of IGK-NAC in CHO cells via Cre/loxP recombination.
(a) Strategy for the generation of IGK-NAC via Cre/loxP recombination-mediated translocation. Recombination between the MAC and the modified hChr.2 generates the IGK-NAC and a by-product. **(b)** FISH analysis confirmed the generation of the IGK-NAC in CHO cells carrying the MAC and modified hChr.2 (Left) and CHO cells carrying the IGK-NAC and by-product (Right). Red and green indicate the MAC and hChr.2, respectively. Insets show enlarged images.

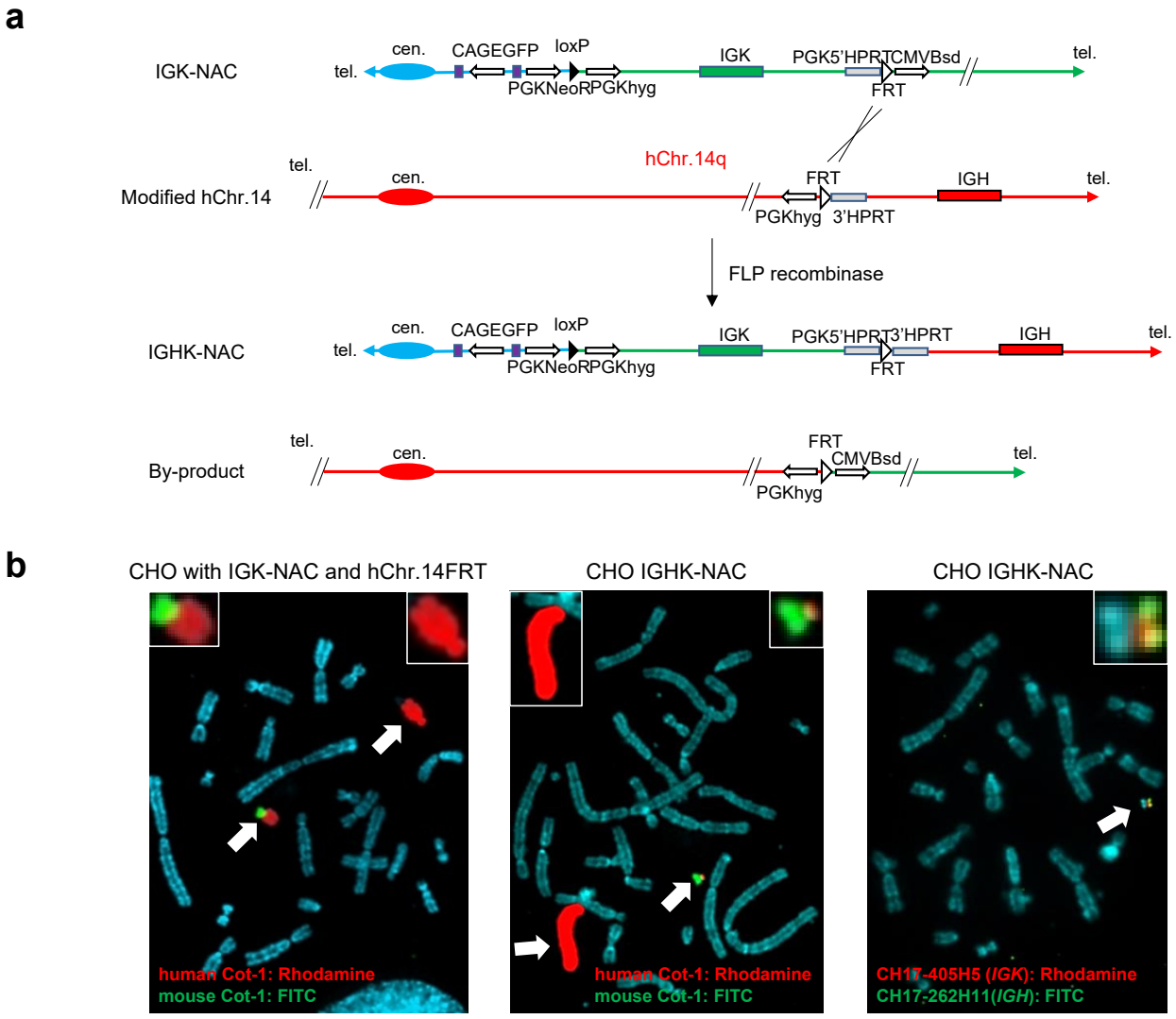
a



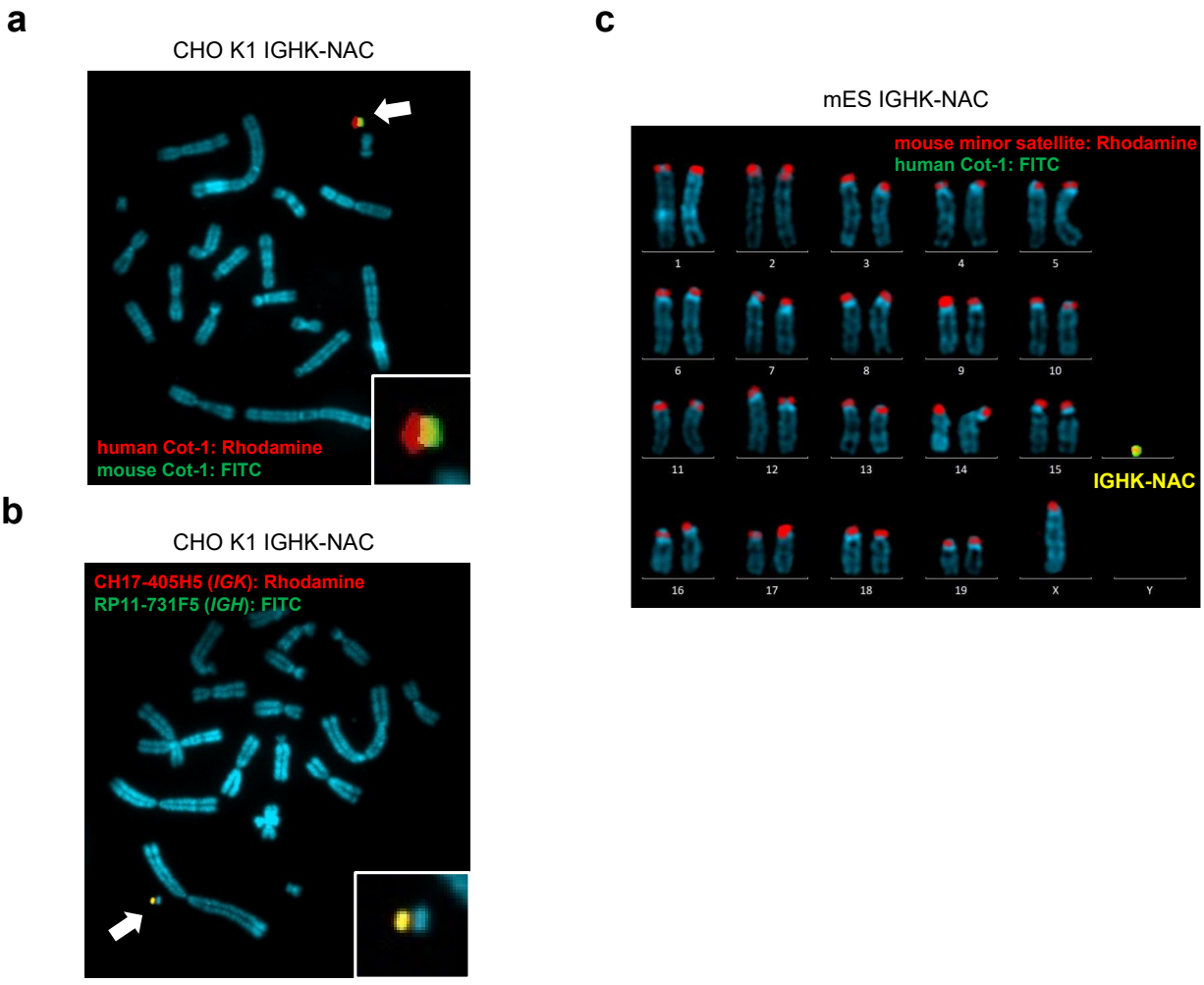
b



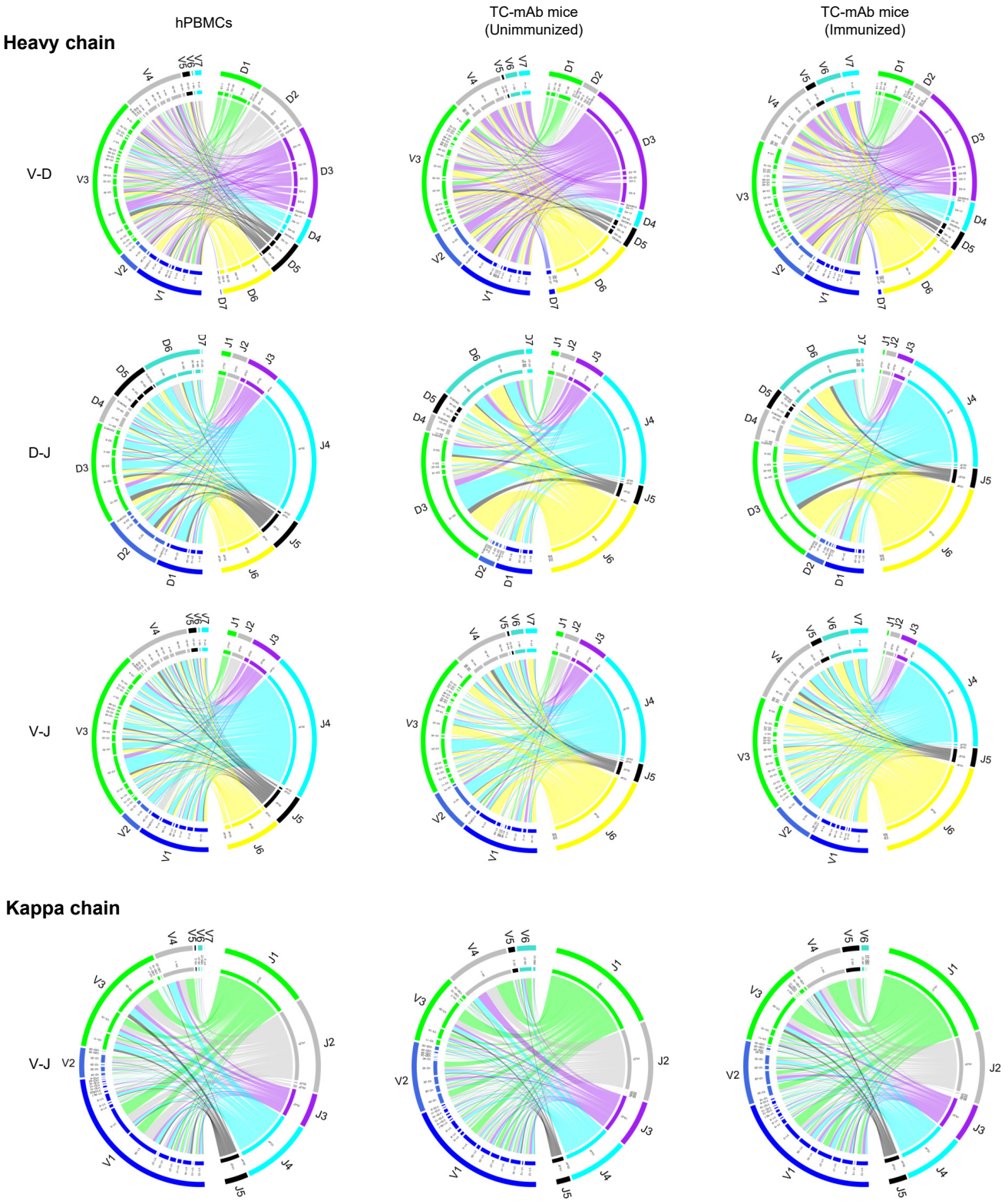
Supplementary Figure 4| Modification of hChr.14 in DT40 cells and transfer of modified hChr.14 to CHO cells.
(a) Strategy for the targeted integration of the FRT site into hChr.14. **(b)** FISH analysis detected targeted integration of the FRT unit into hChr.14 and transfer of hChr.14FRT to CHO cells. DT40 with intact hChr.14 (left), DT40 with hChr.14FRT (middle) and CHO with hChr.14FRT (right). Red and green indicate hChr.14 and FRT units, respectively. Insets show enlarged images.



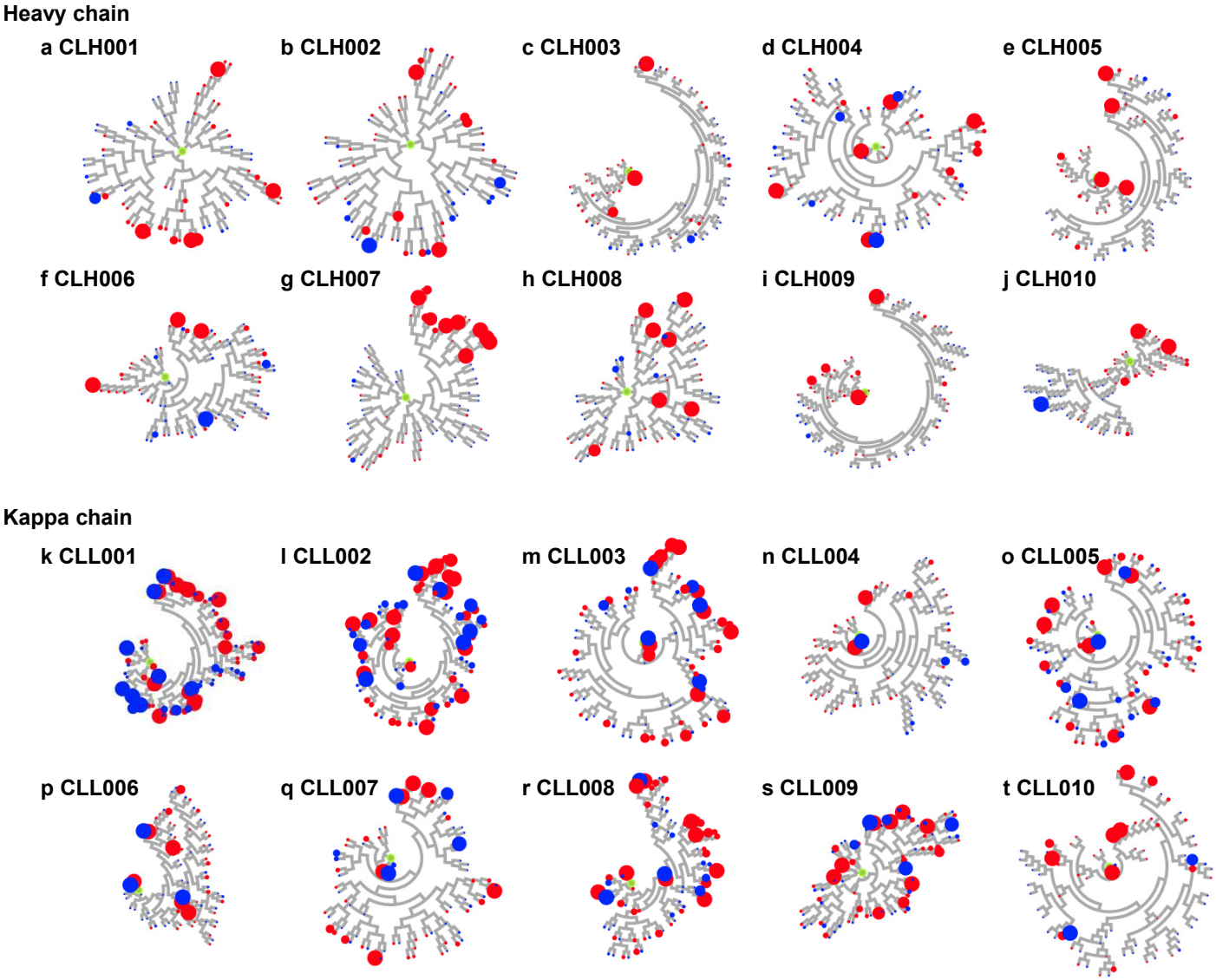
Supplementary Figure 5| Construction of IGHK-NAC in CHO cells via Flp/FRT recombination.
(a) Strategy for the generation of IGHK-NAC via Flp/FRT recombination-mediated translocation. Recombination between IGK-NAC and modified hChr.14 generates IGHK-NAC and a by-product. **(b)** FISH analysis confirmed the generation of IGHK-NAC in CHO cells carrying the IGK-NAC and modified hChr.14 (Left) and CHO cells carrying the IGHK-NAC and by-product (Middle and right). Red and green indicate human genome and MAC, respectively (Left and middle). Red and green indicate parts of the *IGK* and *IGH* regions, respectively (Right). Insets show enlarged images.



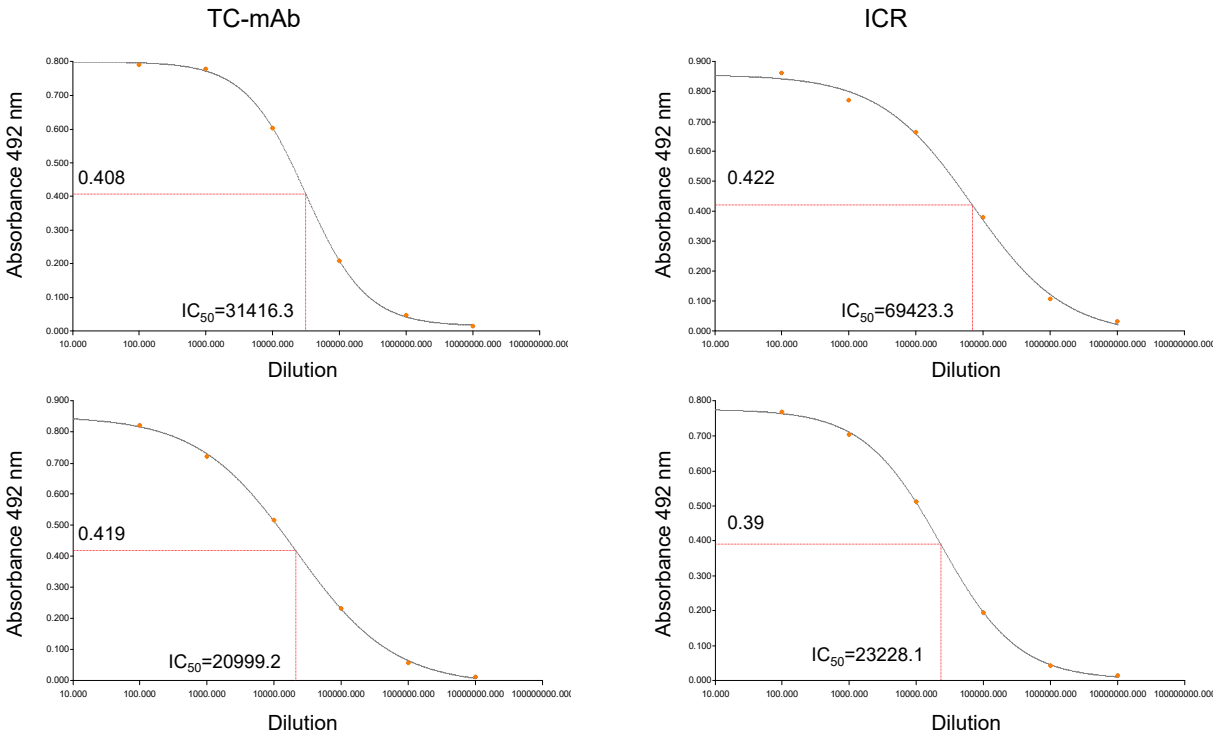
Supplementary Figure 6| Transfer of the IGHK-NAC to mES cells through CHO K1 cells.
(a) FISH analysis of CHO K1 carrying the IGHK-NAC. Green and red indicate the MAC and human genome, respectively. **(b)** FISH analysis of CHO K1 carrying the IGHK-NAC. Red and green indicate parts of the *IGK* and *IGH* regions, respectively. **(c)** Representative karyotype image obtained by FISH analysis of mES cells carrying the IGHK-NAC. Red and green indicate the mouse minor satellite and the human genome, respectively.



Supplementary Figure 7| Comparison of V(D)J gene association in TC-mAb mice.
The circos plots were represented to compare of V(D)J genes detailed association. Outermost tracks mark the boundaries of each V, D, or, J region subfamily in the circos plot. Internal tracks indicate the relative frequencies of subgroups within subfamilies. Highlighted links indicate combinations that constitute 1% or more of all sequences observed. Links indicate the relative frequencies of specific V-D, D-J, and V-J combinations of heavy chain and V-D combinations of light chain. The wider links indicate higher frequencies of recombination. The circos plots of hPBMCs, unimmunized TC-mAb mice, and OVA-immunized TC-mAb mice were represented.

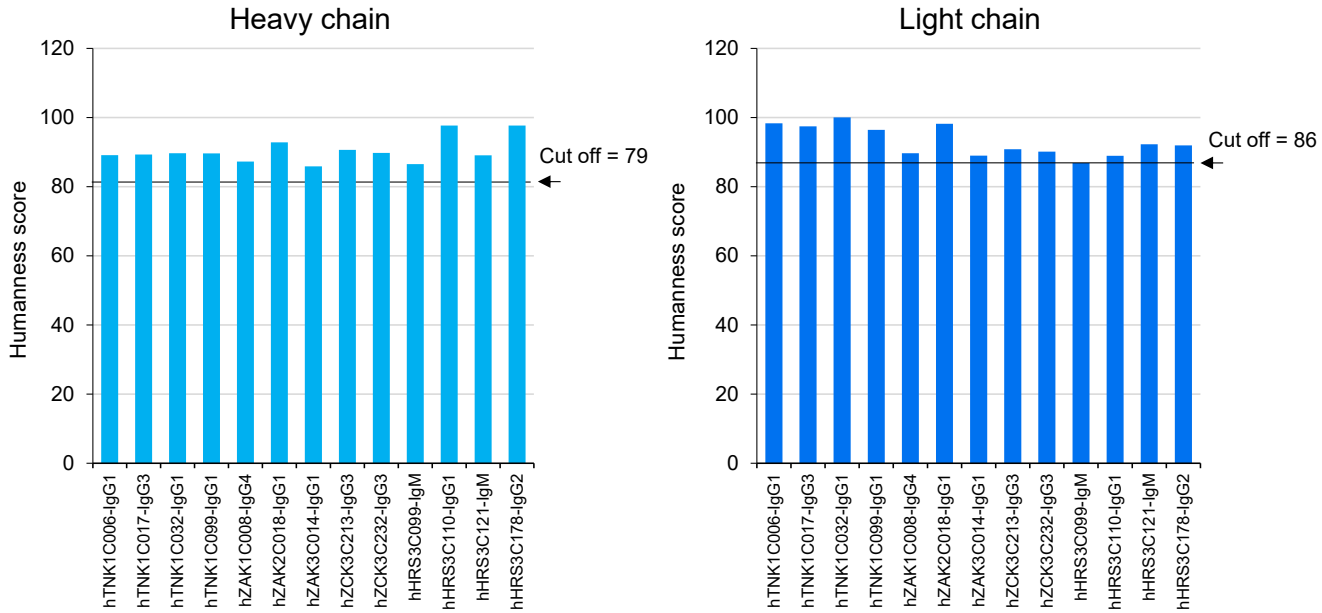


Supplementary Figure 8| Circular dendrogram and convergent sequences.
The phylogenetic trees of the CDRH3 and CDRL3 and amino acid convergent sequences were represented. The circular dendrograms of Top 20 frequently used clone lineages were over-rayed with the number of copies having the same CDR3 sequences in VH and VK, respectively. The results of CLH001-020 of heavy chain **(a-j)** and CLL001-020 of kappa chain **(k-t)** were represented. The leaves of circular dendrograms with a maximum of 50 reads, thereby the larger number drew the larger circle in the leaves. The color of the circle was indicated with (red) or without (blue) immunization in TC-mAb mice.



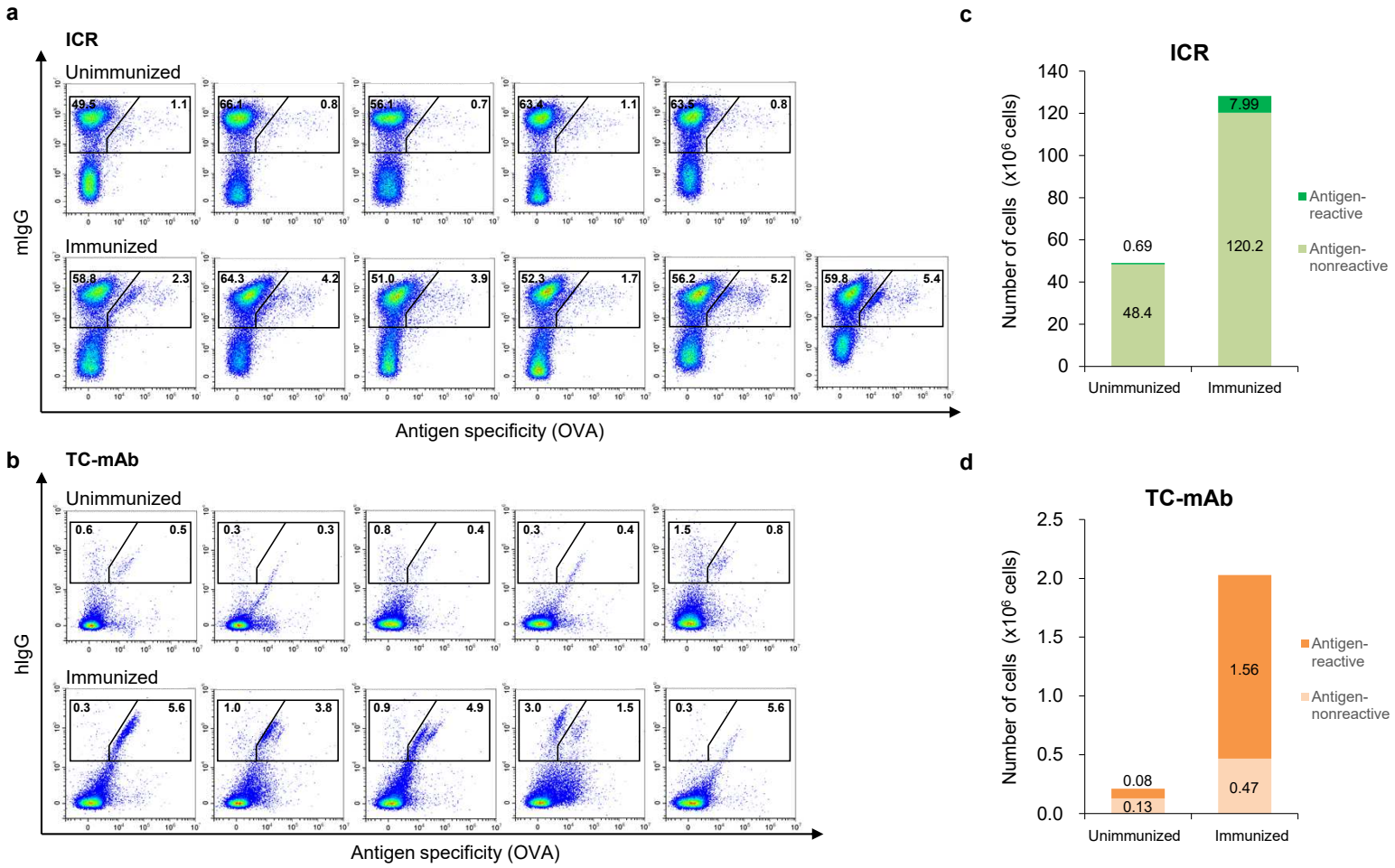
Curve Name	Y (Absorbance 492 nm)	X (Concentration μg/ml)
TC	0.408	31416.3
	0.419	20999.2
ICR	0.422	69423.3
	0.390	23228.1

Supplementary Figure 9| Comparison of anti-sera affinities.
The effective concentration 50 (EC_{50}) of anti-sera obtained from TC-mAb mice (the two graphs on the left side) and ICR (the two graphs on the right side) are indicated. The titres of serially diluted OVA-specific anti-sera showed that the EC_{50} of respective anti-sera was calculated at a dilution of 2–7 in 10^4 by a 4-parameter logistic curve fitting model. The table is a summary of each absorbance and Ab concentration at EC_{50}.



		Heavy chain		Light chain	
Clone name	**Subclass**	VH gene	CDR3H length	VL gene	CDR3L length
hTNK1C006	IgG1	IGHV3-33*01|D3-10*01|J4*02	16	IGKV2-29*02|J3*01	8
hTNK1C017	IgG3	IGHV4-59*01|D3-10*01|J6*02	19	IGKV4-1*01|J2*01	9
hTNK1C032	IgG1	IGHV3-33*01|D3-10*01|J4*02	16	IGKV2-29*02|J3*01	8
hTNK1C099	IgG1	IGHV4-59*01|D3-10*01|J6*02	19	IGKV4-1*01|J2*01	9
hZAK1C008	IgG4	IGHV3-13*01|D7-27*01|J6*02	14	IGHKV3-20*01|J3*01	9
hZAK2C018	IgG1	IGHV3-13*01|D7-27*01|J6*02	14	IGKV3-11*01|KJ2*01	9
hZAK3C014	IgG1	IGHV3-73*02|D3-10*01|J6*02	19	IGKV2D-29*02|J1*01	9
hZCK3C213	IgG3	IGHV3-13*01|D5-24*01|J6*02	14	IGKV1-27*01|J3*01	9
hZCK3C232	IgG3	IGHV3-13*01|D5-24*01|J6*02	14	IGKV1-27*01|J3*01	9
hHRS3C099	IgM	IGHV3-9*01|D1-26*01|J6*02	14	IGKV4-1*01|J2*01	10
hHRS3C110	IgG1	IGHV3-48*02|D6-13*01|J4*02	9	IGKV1-16*01|J4*01	9
hHRS3C121	IgM	IGHV3-13*01|D1-1*01|J6*02	14	IGKV3-20*01|J3*01	9
hHRS3C178	IgG2	IGHV3-48*02|D6-13*01|J4*02	9	IGKV4-1*01|J2*01	9

Supplementary Figure 10| Humanness of mAbs.
Humanness scores of mAb isolated from TC-mAb mice were indicated. The nucleotide sequences of VH and VK region of mAbs isolated from TC-mAb mice were determined for anti-EpCAM mAbs (hZAK1A008 and 014) and anti-AMIGO2 mAbs (hTNK1A006, B017a and C032a). The T20 scores were determined using the T20 score analyser. To make a judgment as a human sequence, cut off values (Heavy chain was 79 and κ chain was 86) was used according to the threshold set by authors. The thresholds were indicated as arrows. The table was the summary of antibodies sequence including their subclass, matched germline genes of VH and VK, and CDR length.



Supplementary Figure 11| Detection of antigen-specific B cells in spleen and PBMCs.
(a, b) Flow cytometry of B cell subsets of antigen (OVA)-specific B cells (IgG^+OVA^+). Representative images of spleen and PBMCs were collected from five different ICR **(a)** and TC-mAb mice **(b)** with or without immunization. Upper panel indicates the images of unimmunized mice and the lower panel indicates that of unimmunized mice. The numbers in the panels represent the percentage of IgG-positive cells among lymphocytes. **(c, d)** Absolute number of IgG-positive cells and antigen-specific cells. The average number of antigen-reactive and -nonspecific cells in the lymphocytes of ICR **(c)** and TC-mAb mice **(d)** are represented. The numbers in the figures represent the number of IgG-positive cells.

Figures

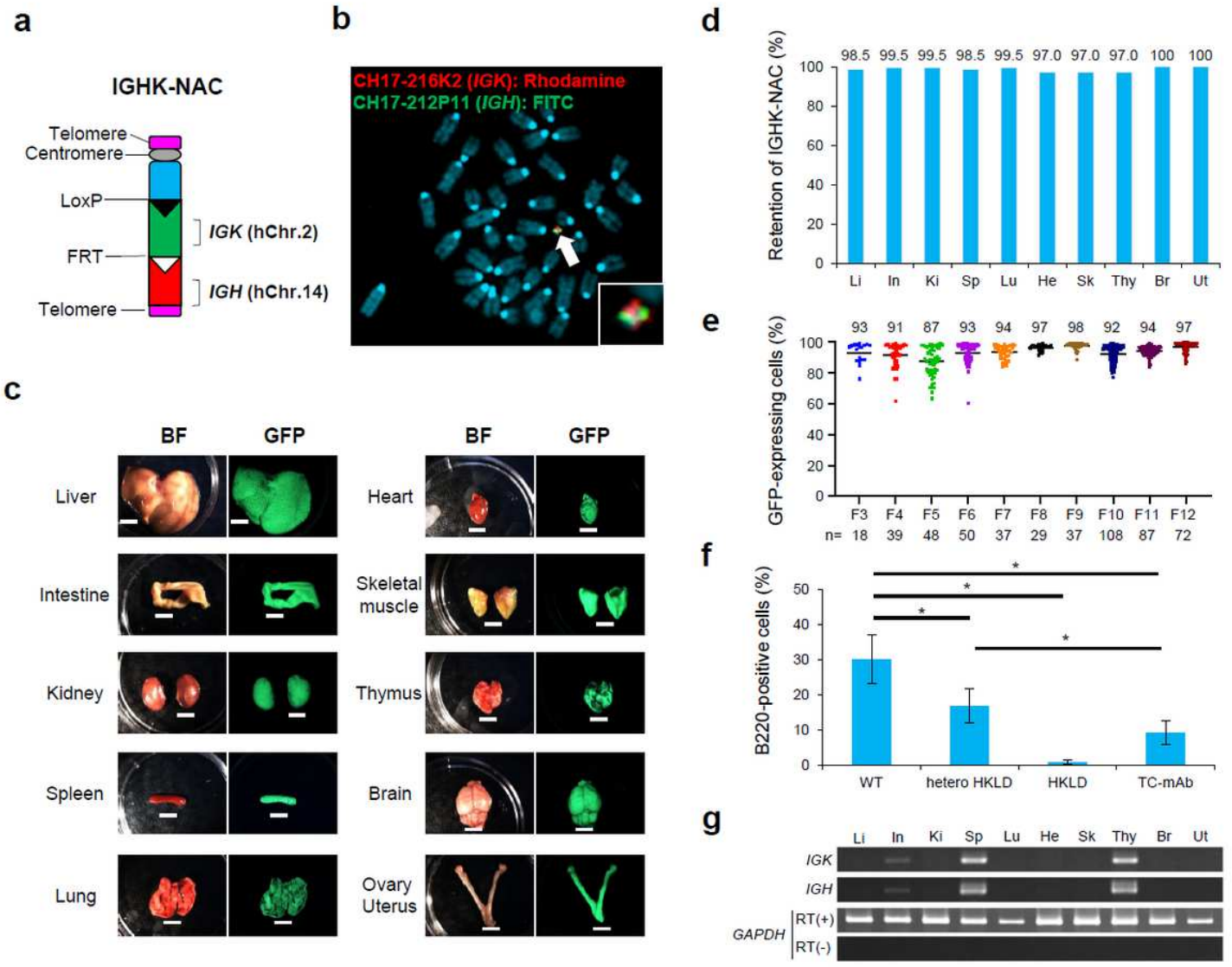


Figure 1

Generation of fully human Ab producing mice. (a) Representation of the IGHK-NAC structure. Human Ig heavy chain locus (IGH) derived from hChr.14 and Ig kappa light chain locus (IGK) derived from hChr.2 are represented on a mouse artificial chromosome (MAC). (b) Representative image of metaphase FISH analysis with IGK-BAC (CH17-216K2) (red) and IGH-BAC (CH17-212P11) (green) detecting the IGHK-NAC in bone marrow cells from an IGHK-HKLD mouse. Arrow indicates the IGHK-NAC and the inset shows an enlarged image thereof. (c) GFP images of different tissues from a Tc mouse carrying the IGHK-NAC. GFP expression indicates the presence of the IGHK-NAC. BF, bright field. Scale bar (5 mm). (d) Retention rate of the IGHK-NAC in various Tc mouse tissues analysed by FISH. Li, liver; In, intestine; Ki, kidney; Sp, spleen; Lu, lung; He, heart; Sk, skeletal muscle; Thy, thymus; Br, brain; Ut, uterus. (e) Percentage of GFP-expressing cells in peripheral blood mononuclear cells from Tc mice of different generations. (f) Percentage of B220-positive cells in the lymphocyte fraction of peripheral blood in age-matched WT

(n=10), hetero HKLD (n=10, Igh+/-, Igk+/- and Igλ1+/low), HKLD (n=10, Igh-/-, Igk-/- and Igλ1low/low) and Tc-mAb mice (n=525). *p<0.001. (g) RT-PCR analysis of total RNA from various tissues of the Tc mouse. Li, liver; In, intestine; Ki, kidney; Sp, spleen; Lu, lung; He, heart; Sk, skeletal muscle; Thy, thymus; Br, brain; Ut, uterus.

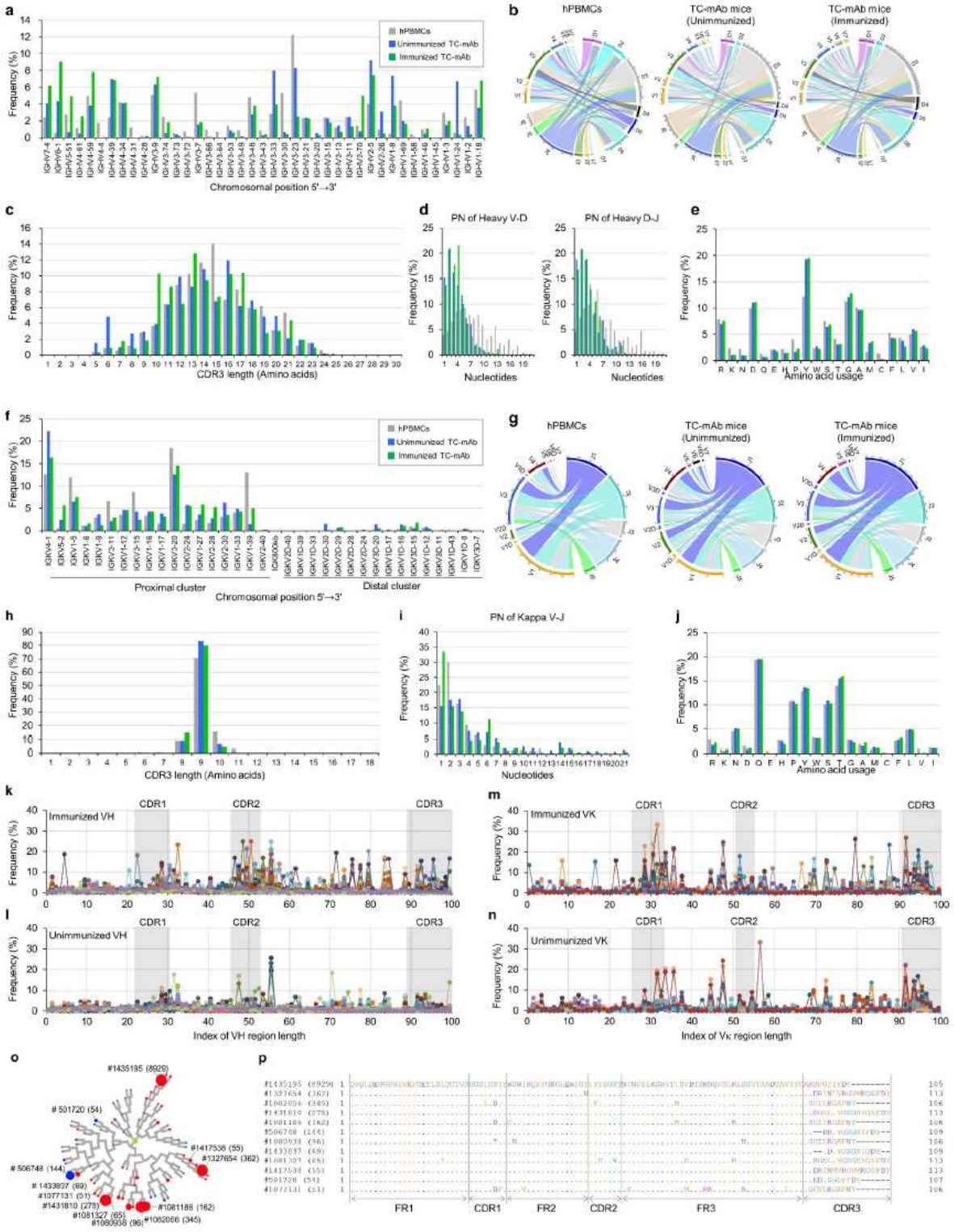


Figure 2

Repertoire analyses of heavy and light chain variable regions in TC-mAb mice. (a) Human heavy chain gene utilization in hPBMCs from healthy donors and in TC-mAb mice with and without immunization. Percentage frequency use of V gene segments in healthy human donors (grey), TC-mAb mice with (green), and without immunization (blue) are represented. (b) Circos plots comparing VDJ gene association. The gene segments are grouped as subfamilies and are shown together with the first digit of their allele name. Links indicated the relative frequencies of specific VDJ combinations, and wider links indicate higher frequencies of recombination. The relative CDRH3 length (c), P and N nucleotide addition at joining positions (d), and amino acid usage (e) within the CDRH3 of productive rearrangements. (f) Human VK gene utilization in hPBMCs of healthy donors and in TC-mAb mice with and without immunization. (g) Circos plots comparing VJ gene association. Links indicate the relative frequencies of specific VJ combinations. The relative CDRL3 length (h), P and N nucleotide addition at joining positions (i), and amino acid usage (j) within the CDRL3 of productive rearrangements. The frequency of SHM at every position of the variable region was calculated as described in the Methods and is represented for VH and VK chains in OVA-immunized TC-mAb mice (k, m) and unimmunized TC-mAb mice (l, n). (o) The phylogenetic tree of the clone type CLH001. The circular dendrograms of the 20 most frequently used clone lineages were over-laid with the number of copies carrying the same CDRH3 sequences. The leaves of circular dendrograms have a maximum of 50 reads; therefore, the larger the number the larger the circle in the leaves. The colour of the circle indicates with (red) or without (blue) immunization in TC-mAb mice. The amino acid sequences of 12 clonotypes containing more than 50 copies in the CLH001 clone lineage. The clonotype numbers and each copy number are also indicated. (p) Amino acid sequence alignment of the 12 clonotype variable regions of CLH001.

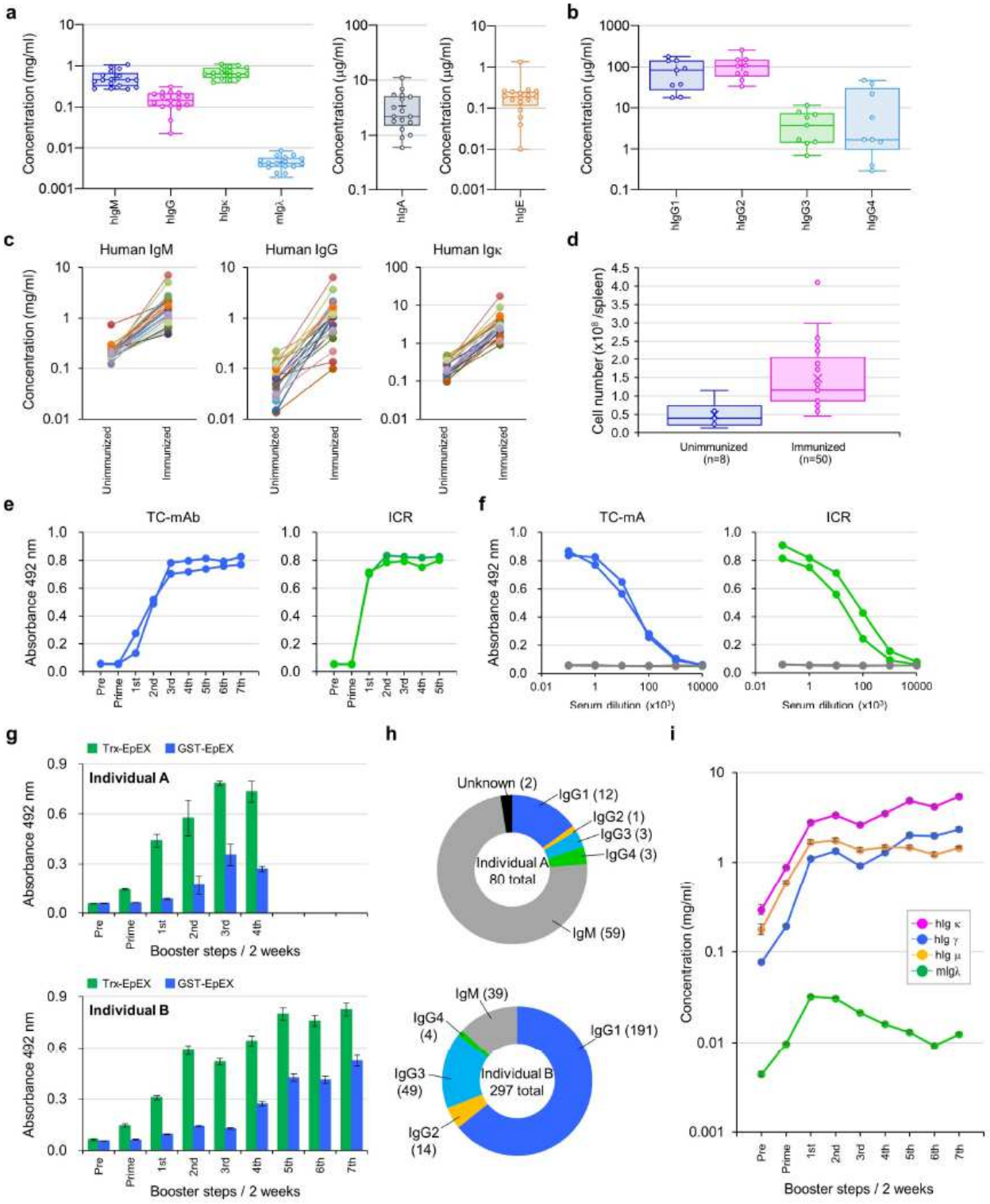


Figure 3

Production of human Ig in TC-mAb mice . (a) The serum concentration of different classes of human Igs in TC-mA mice. The concentration of hIgM, hIgG, hIgκ, mIgλ, hIgA and hIgE in TC-mAb mice (n=17) was measured. Each symbol represents one serum concentration. (b) The serum concentration of human IgG subclasses. The concentrations of IgG1, 2, 3, and 4 were measured in TC-mAb mice (n=9). (c) The serum concentration of hIgM, IgG, and Igκ before and after immunization of TC-mAb mice. The symbols

collected from the same mouse before and after immunization are connected by a line. (d) The number of lymphocytes. The numbers of harvested cells were counted under Turk's solution. Each symbol indicates one mouse. (e) OVA-specific Ab titre. The titres of Abs against OVA were measured by ELISA using a species-matched secondary Ab. An anti-human IgG-Fc HRP conjugate was used for TC-mAb mice, and an anti-mouse IgG-Fc HRP conjugate was used for ICR. (f) Comparison of the anti-sera titres. The anti-sera samples were collected from TC-mAb mice after the seventh booster, and from ICR mice after the fourth booster. Samples were serially diluted by 1/10 from 102 to 107 and used to determine the OVA-specific titre by ELISA. (g) The antigen-specific titre in TC-mAb mice. Individual A (upper graph) was immunized with Trx-EpEX with the prime and four booster administrations. Individual B (lower graph) was immunized with the prime and seven booster administrations. The titres were analysed using the fusion proteins, Trx- EpEX (green) and GST-EpEX (blue). The Trx-EpEX titre indicated induction of antigen-specific Abs, but that of GST-EpEX indicated administration of EpEX-specific Abs. Error bars indicate standard deviation from triplicate measurements. (h) Determination of human Ig-classes. The mouse individual and the total number of class-determined clones are indicated in the centre of the circle. (i) The serum concentration of different classes of Ab. hIg μ (orange), hIg μ (blue), hIg κ (pink), and mIg λ (green) in the Trx-EpEX immunized TC-mAb mice are represented. Error bars indicate standard deviation from triplicate measurements.

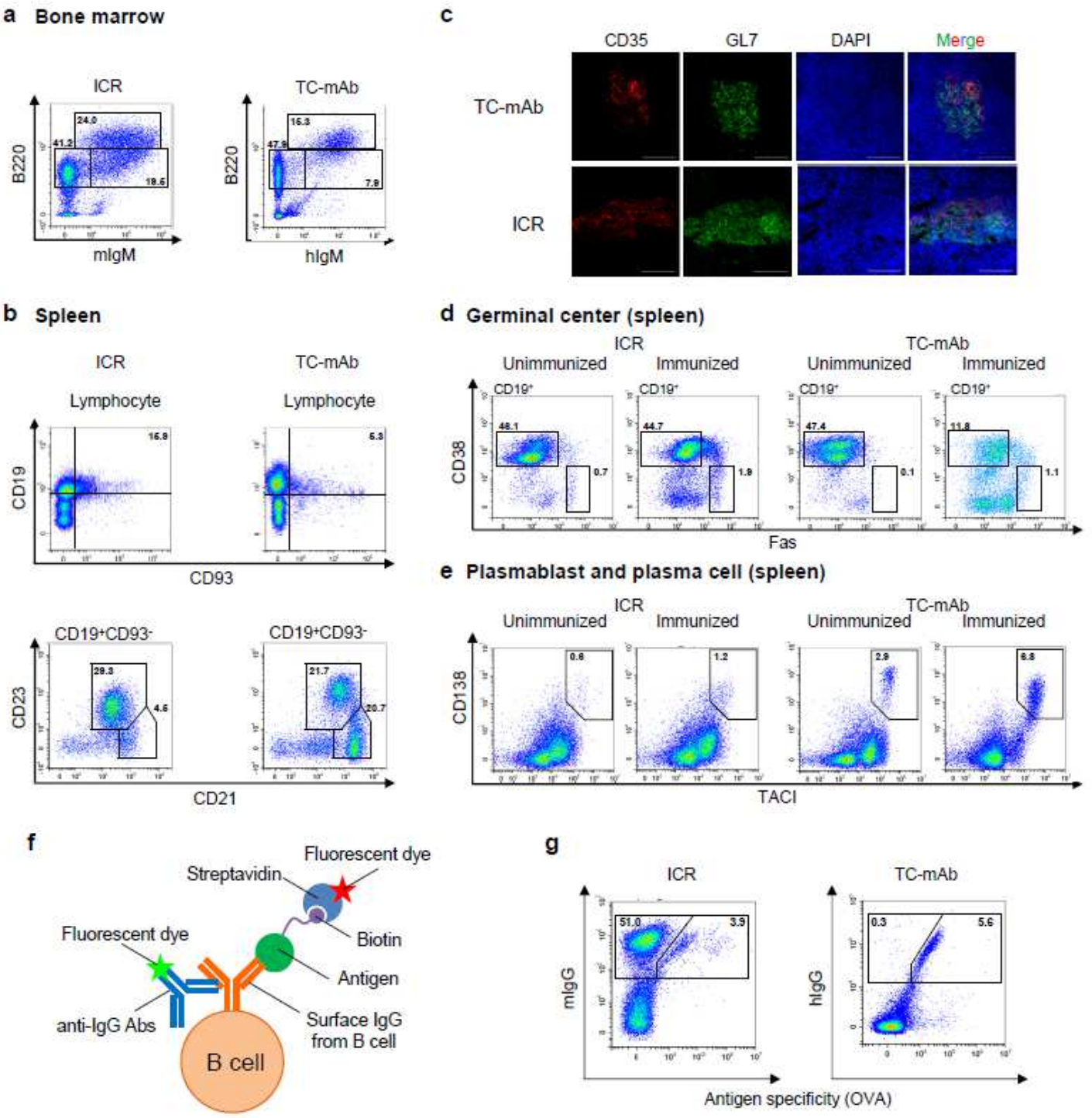


Figure 4

Analysis of B-cell development. (a) Flow cytometry of B cell subsets of both Pro-B and Pre-B (IgM-B220+), immature B (IgM+B220+), and mature recirculating B (IgMlo/+B220hi) cells in bone marrow. Numbers indicate the percentage of B cells in each subset. The numbers in the panels represent the percentage of lymphocytes. (b) The number of cells in the B lineage in the spleen. Flow cytometry of Transitional B (CD19+CD93+), Follicular B (CD19+CD93-CD21+CD23+), and Marginal zone B (CD19+CD93-CD21hiCD23-) cells are represented. (c) Representative images of fluorescence immunohistochemistry.

The spleens of OVA-immunized TC-mAb mice were harvested and analysed by immunohistochemistry. Sections were stained for follicular dendritic cells (CD35, red), GC B cells (GL7; green), and nuclei (DAPI, blue). The images were taken at ×20 magnification. All scale bars indicate 100 μm. (d) The number of cells in the germinal centre-related B lineage in the spleen. Flow cytometry of germinal centre B (CD19+CD38lo/-Fas+) cells is indicated. (e) The number of Ab-producing cells in the spleen. Flow cytometry of both PB and PC (CD138+TACI+) are represented. The numbers in the panels represent the percentages in the lymphocytes. (f) Representation of image to detect antigen-specific Ab expressing B cell subsets. (g) Flow cytometry of antigen-specific B cell subsets (IgG+antigen+). The numbers in the panels represent the percentages of IgG-positive lymphocytes.