

Expanded View Figures

V

IGHV4-59/D3-3/J6	1	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGY
GL-CDR3rev1 Steichen	1	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGY
GL-CDR3rev4 Escolano	1	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGY
PGT121 bnAB	1	QMQLQESGPGLVKPSSETLSLTCSVSGASISDSYWSWIRRSPGKGLEWIGY

IGHV4-59/D3-3/J6	51	IYYSGSTNYPNPSLKSRVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
GL-CDR3rev1 Steichen	51	IYYSGSTNYPNPSLKSRVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
GL-CDR3rev4 Escolano	51	IYYSGSTNYPNPSLKSRVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
PGT121 bnAB	51	VHKSGDTNYSPLKSRVNLSDTSKNQVSLSLVAATAADSGKYYCAR

D

J

IGHV4-59/D3-3/J6	95	----ITIFGVVII---YYYYYMDV
GL-CDR3rev1 Steichen	98	TLHGITIFGVVAFKEYYYYYYMDV
GL-CDR3rev4 Escolano	98	TQQGKRIYGVVSFGDYYYYYYMDV
PGT121 bnAB	98	TLHGRRYIGIVAFNEWFTYFYNDV

bold – sequence was used for generating the KI mouse presented here

V- variable region; D - diversity region; J - joining region

Figure EV1. Alignment of inferred germline PGT121 heavy chain sequences with full PGT121 bnAb heavy chain sequences.

Sequences were aligned in Jalview using ClustalO and post-processed in BOXSHADE (http://www.ch.embnet.org/software/BOX_form.html). Relevant publications, from which these sequences originate, are listed: (i) GL-CDR3rev1 Steichen—PMID: 27617678, (ii) GL-CDR3rev4 Escolano—PMID: 27610569.

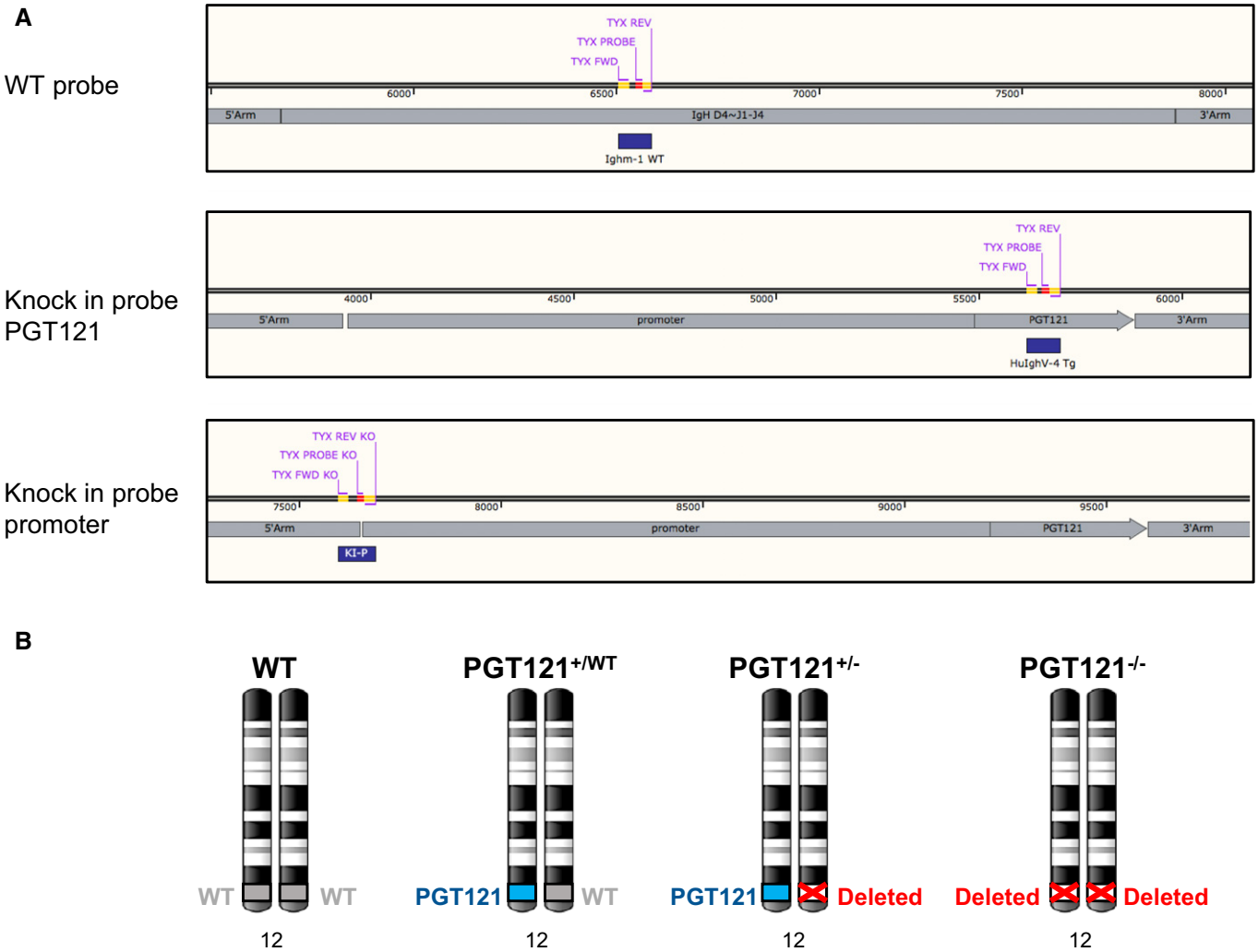


Figure EV2. TransnetYX probes design and KI mice named.

A 3 TaqMan probes, Ighm-1 WT, HulghV-4 Tg, and KI-P designed for genotyping.
B Schematic showing nomenclatures of WT and PGT121 KI mice according to genotyping results.

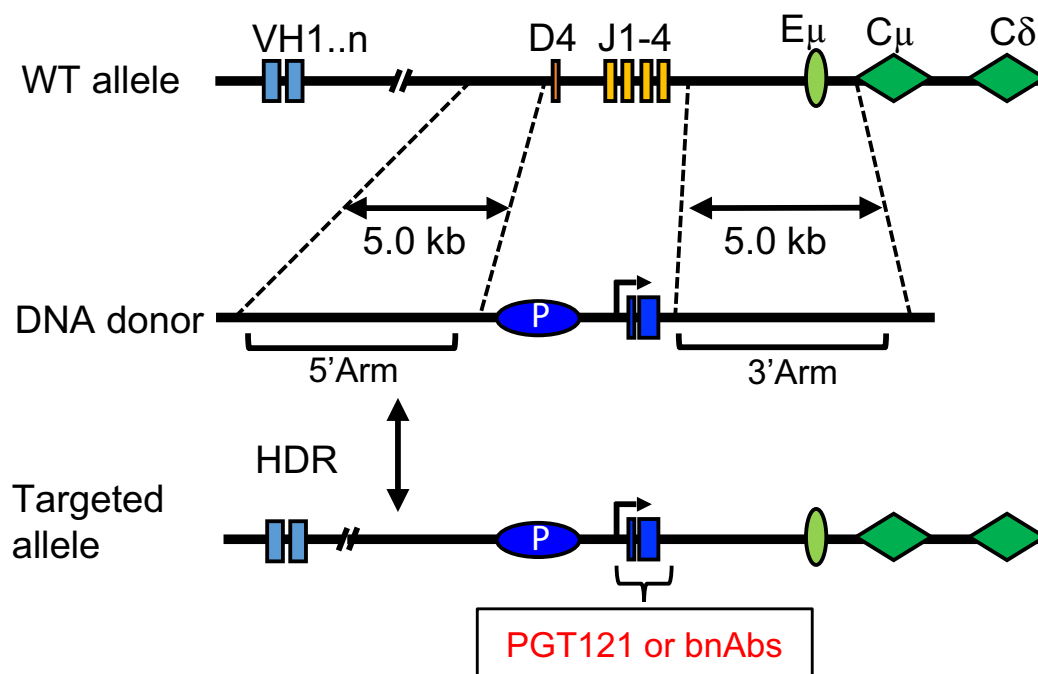


Figure EV3. Schematics of elongated DNA donor arms.

The original arms of DNA donor are 3.9 kb and 2.6 kb in length. We extended both homology arms to 5 kb for later microinjections.

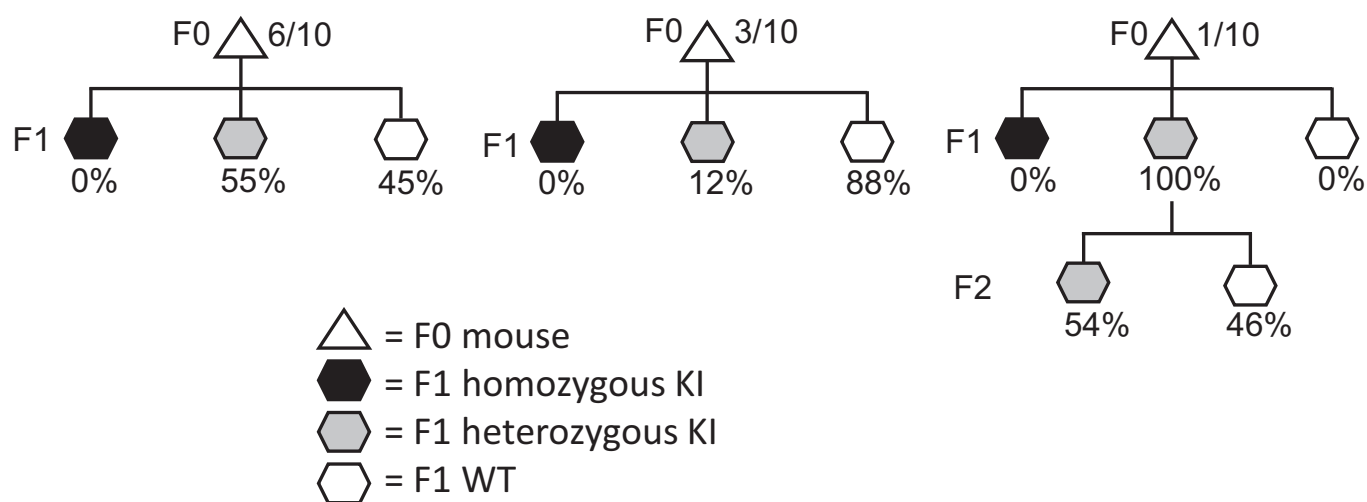


Figure EV4. Genetic tree of KI mice depicting the F0, F1, and F2 generations.

Six out of ten F0 mice revealed the Mendelian transmission of the heavy chain. Three out of ten F0 founders exhibited non-Mendelian transmission. One of our founders has at least more than one insertion of our KI heavy chain. White triangles—F0 mouse; black hexagon—F1 homozygous KI; gray hexagon—F1 heterozygous KI; white hexagon—F1 WT; the number of mice is shown next to the triangles. The percentages of animals of each genotype detected are shown below each of the hexagons.

A

BG18 heavy chain germline

ID	IGHV	IGHD	IGHJ	HCDR3(aa)	Length
BG18	Hu 4-4*02	Hu 3-3*01	6*02	CARNAITIFGVVALGEYYYYGMDVW	23

BG18 knock-in mice

ID	IGHV	IGHD	IGHJ	HCDR3(aa)	Length
m1.A09	Mu 1-52*01	Mu 2-4*01	Hu 6*02	CARGIYYDYDARNAITIFGVVALGEYYYYGMDVW	32
m1.F05	Mu 1-42*01	Mu 2-4*01	Hu 6*02	CARQAIYYDYDEGNAITIFGVVALGEYYYYGMDVW	33
m2.A04	Mu 1-80*01	Mu 2-5*01	Hu 6*02	CARSGYSNPPRNAITIFGVVALGEYYYYGMDVW	31
m3.B08	Mu 2-2*01	Mu 1-1*01	Hu 6*02	CARGGLPRNAITIFGVVALGEYYYYGMDVW	28
m3.B10	Mu 3-6*01	Mu 1-1*01	Hu 6*02	CARGRGVVALGEYYYYGMDVW	19
m3.C06	Mu 1-75*01	Mu 1-1*01	Hu 6*02	CARSLFTIFGVVALGEYYYYGMDVW	23
m3.C11	Mu 5-17*01	Mu 2-10*02	Hu 6*02	CARGMGYGNITIFGVVALGEYYYYGMDVW	27
m3.F02	Mu 1-55*01	Mu 1-1*01	Hu 6*02	CARAPIYYYGVVALGEYYYYGMDVW	23

B

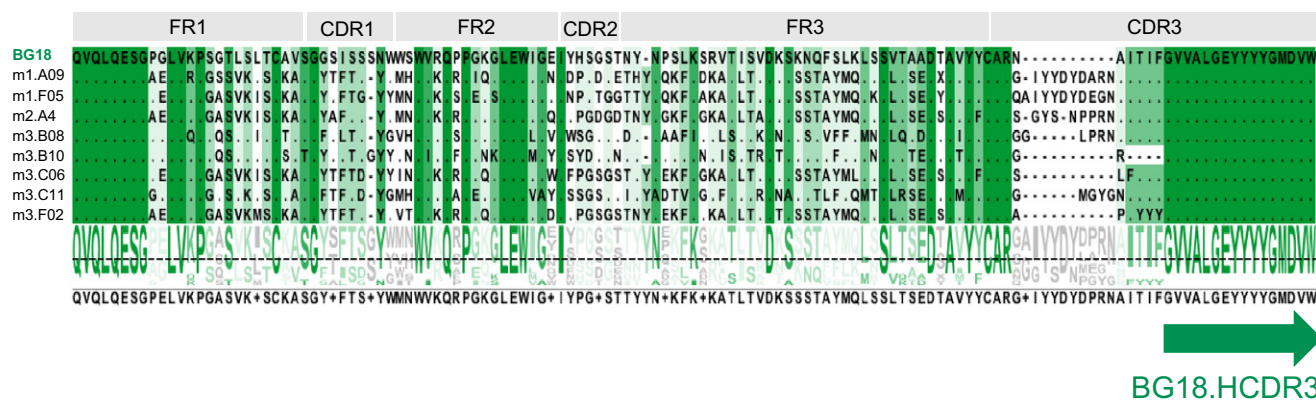


Figure EV5. Sequence analysis of BG18 sequences recovered from spleen cells.

- A Splenic B cells from three BG18 mice were stained for B220 and IgM. B cells were single-cell sorted for antibody heavy chain and subjected to PCR amplification and sequencing. The murine IGHV sequences isolated from splenic B cells exhibited CDR3 motif of BG18, depicted in green lettering.
- B The eight murine IGHV sequences (from Fig 8A; BG18^{+/+}) were aligned to the germline-reverted BG18 heavy chain (green). The CDR3 motif GVVALGEYYYYGMDVW of BG18 was found in all of these IGHV sequences (indicated by green arrow). CDR: complementarity-determining region, FR: framework region.