PNAS
www.pnas.org

Supplementary Information for

Intestinal microbes influence development of thymic lymphocytes in early life

Maria Ennamorati[1,¶], Chithirachelvi Vasudevan[1,¶], Kara Clerkin[1,¶], Stefan Halvorsen[2], Smriti Verma[1], Samira Ibrahim[1], Shaniah Prosper[1], Caryn Porter[2], Vladimir Yeliseyev[3], Margot Kim[4], Joseph Gardecki[4], Slim Sassi[2, 5], Guillermo Tearney[4], Bobby J. Cherayil[1,6], Lynn Bry[3], Brian Seed[2], Nitya Jain[1, 2, 6, *]

[1]Mucosal Immunology and Biology Research Center, Massachusetts General Hospital *for* Children, 114 16th Street, Charlestown, MA 02129, USA
[2]Center for Computational and Integrative Biology, Massachusetts General Hospital, 185 Cambridge Street, Boston MA 02114, USA
[3]Massachusetts Host-Microbiome Center, Brigham & Women's Hospital, Harvard Medical School, 75 Francis St., Boston, MA 02115, USA
[4]Wellman Center for Photomedicine, Massachusetts General Hospital, 50 Blossom Street, Boston, MA 02114
[5]Orthopedics Department, Harvard Medical School, Boston, MA 02115, USA
[6]Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA

Correspondence to Nitya Jain

Email: njain@ccib.mgh.harvard.edu

This PDF file includes:

Supplementary text
Figures S1 to S14

www.pnas.org/cgi/doi/10.1073/pnas.1915047117

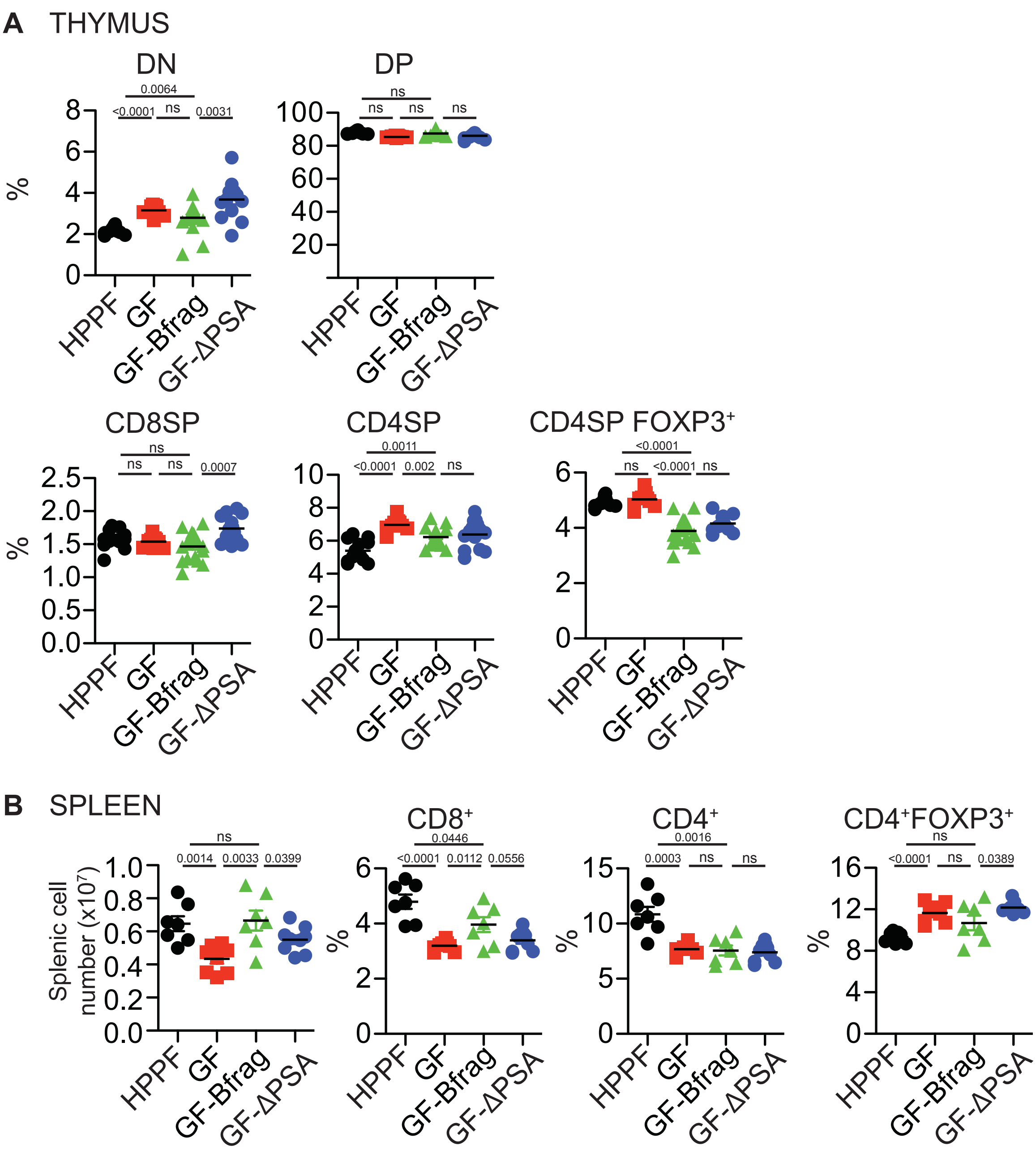

Figure S1: Distribution of major T cell subsets in thymus and spleen of HPPF, GF, GF-Bfrag and GF-ΔPSA pups.

A. Frequency of thymic DN ($CD4^{neg}CD8a^{neg}$), DP ($CD4^{+}CD8a^{+}$), CD8SP ($CD4^{neg}CD8^{+}$), CD4SP ($CD4^{+}CD8a^{neg}$) and CD4SP $FOXP3^{+}$ cells (gated on CD4SP) (HPPF n= 12; GF n=8; GF-Bfrag n=18; GF-ΔPSA n=14).

B. Total splenic cellularity and frequency of $CD8^{+}$, $CD4^{+}$ (gated on Live) and $CD4^{+}FOXP3^{+}$ cells (gated on CD4+ T cells) (HPPF n= 7; GF n=8; GF-Bfrag n=7; GF-ΔPSA n=9).

Data are from 2-3 independent experiments for each group. Bars are Mean ± SEM.

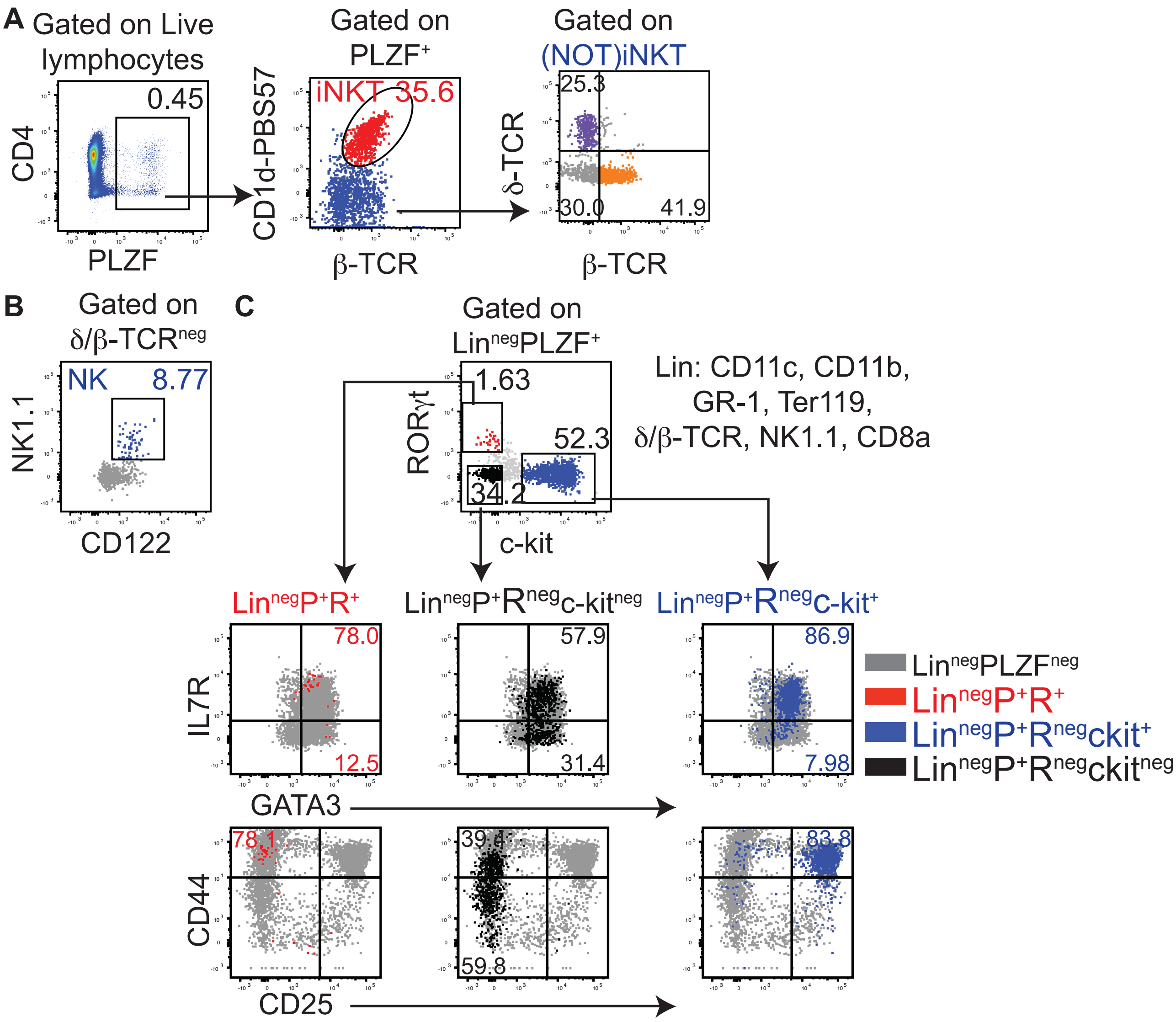

Figure S2: Identification of PLZF expressing cell subsets by flow cytometry
Gating strategy to identify PLZF expressing cell subsets in the thymus of 14 days old C57BL/6 pups. Live and single lymphocytes were analyzed as follows:

A. (Left) Representative flow cytometry plot showing expression of PLZF (y-axis) and CD4 (x-axis) to identify total PLZF$^+$ cells. (Middle) PLZF$^+$ cells were next analyzed for β-TCR (x-axis) and mCD1d-PBS57 tetramer (y-axis) expression to identify mCD1d-PBS57$^+$TCRβ$^+$ iNKT cells. (Right) A (NOT)iNKT gate was used to determine expression of β-TCR (x-axis) and δ-TCR (y-axis) on mCD1d-PBS57^{neg} cells to identify PLZF$^+$γδ T cells and PLZF$^+$ innate-like αβ-T cells.

B. PLZF$^+$ cells were gated on β/δ-TCRneg cells and analyzed for expression of CD122 (x-axis) and NK1.1 (y-axis) to identify NK cells.

C. TCR^{neg}NK1.1^{neg}PLZF$^+$ cells that were also Lin^{neg} (Lin: CD11c, CD11b, Ter119, CD19, GR-1 and CD8a) were analyzed for expression of c-kit (x-axis) and RORγt (y-axis). Cells identified as Lin^{neg}PLZF$^+$RORγt$^+$ (Lin^{neg}P$^+$R$^+$ in red), Lin^{neg}PLZF$^+$c-kit$^+$ (Lin^{neg}P$^+$c-kit$^+$ in blue) and Lin^{neg}PLZF$^+$c-kit^{neg} (Lin^{neg}P$^+$c-kit^{neg} in black) cells were further analyzed for expression of (top row) GATA3 (x-axis) and IL7R (y-axis) and (bottom row) CD25 (x-axis) and CD44 (y-axis).

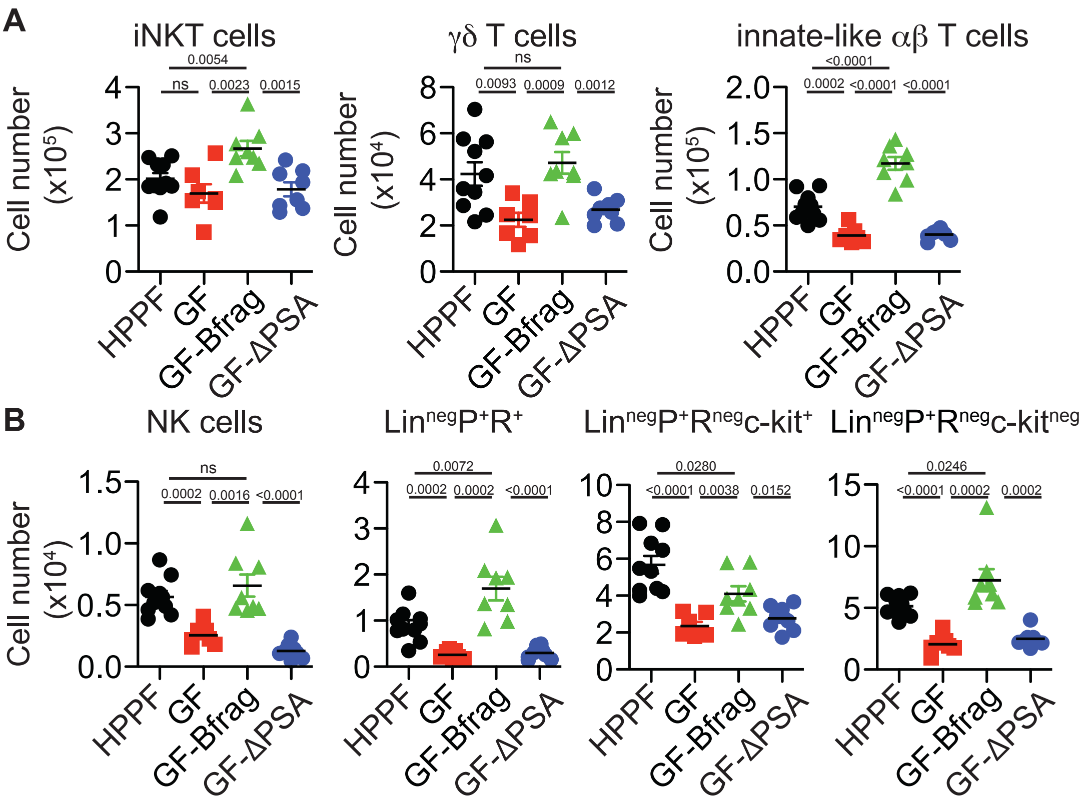

Figure S3: Cell numbers of PLZF-expressing subsets in thymus of HPPF, GF, GF-Bfrag and GF-ΔPSA pups.

A. Total numbers of PLZF$^+$ iNKT cells, PLZF$^+$ γδ T cells and PLZF$^+$ innate-like αβ-T cells (HPPF n= 10; GF n=7; GF-Bfrag n=8; GF-ΔPSA n=8).

B. Total numbers of PLZF$^+$ NK cells, Lin^{neg}PLZF$^+$RORγt$^+$ cells, Lin^{neg}PLZF$^+$RORγt^{neg}c-kit$^+$ cells and Lin^{neg}PLZF$^+$RORγt^{neg}c-kit^{neg} cells (HPPF n= 10; GF n=7; GF-Bfrag n=8; GF-ΔPSA n=8).

Data are representative of 2-3 experiments for each group. Bars are Mean ± SEM.

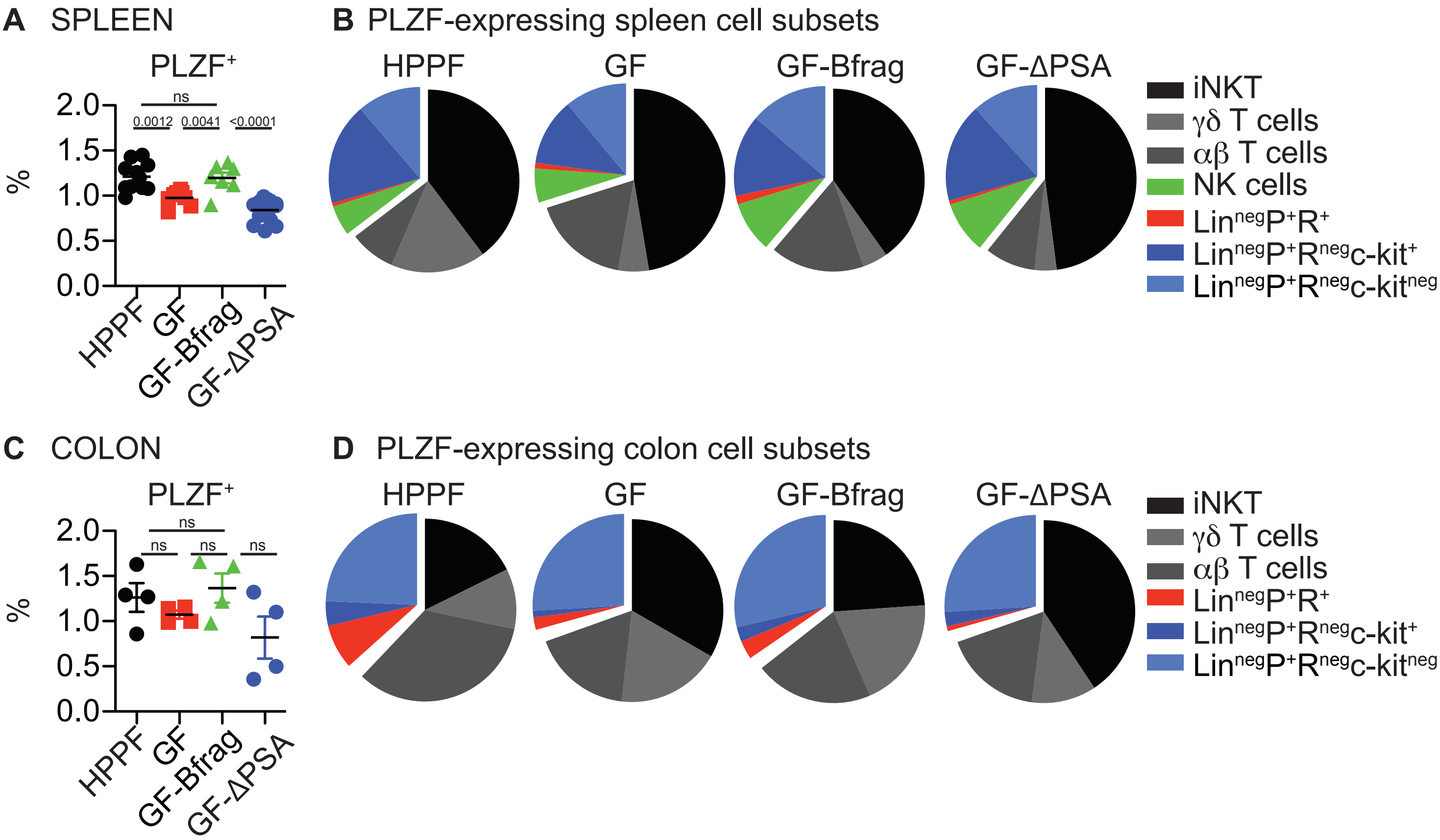

Figure S4: Frequency of PLZF⁺ cells in spleen and colon of HPPF, GF, GF-Bfrag and GF-ΔPSA pups.

A. Frequency of PLZF⁺ cells in spleen (HPPF n= 11; GF n=8; GF-Bfrag n=7; GF-ΔPSA n=17). Data are from of 2-3 experiments in each group.

B. Pie graphs showing distribution of indicated PLZF expressing cell subsets in spleen.

C. Frequency of PLZF⁺ cells in colon (HPPF n= 4; GF n=4; GF-Bfrag n=4; GF-ΔPSA n=4). Data are representative of 2 independent experiments.

D. Pie graphs showing distribution of indicated PLZF expressing cell subsets in colon.
Bars are Mean ± SEM.

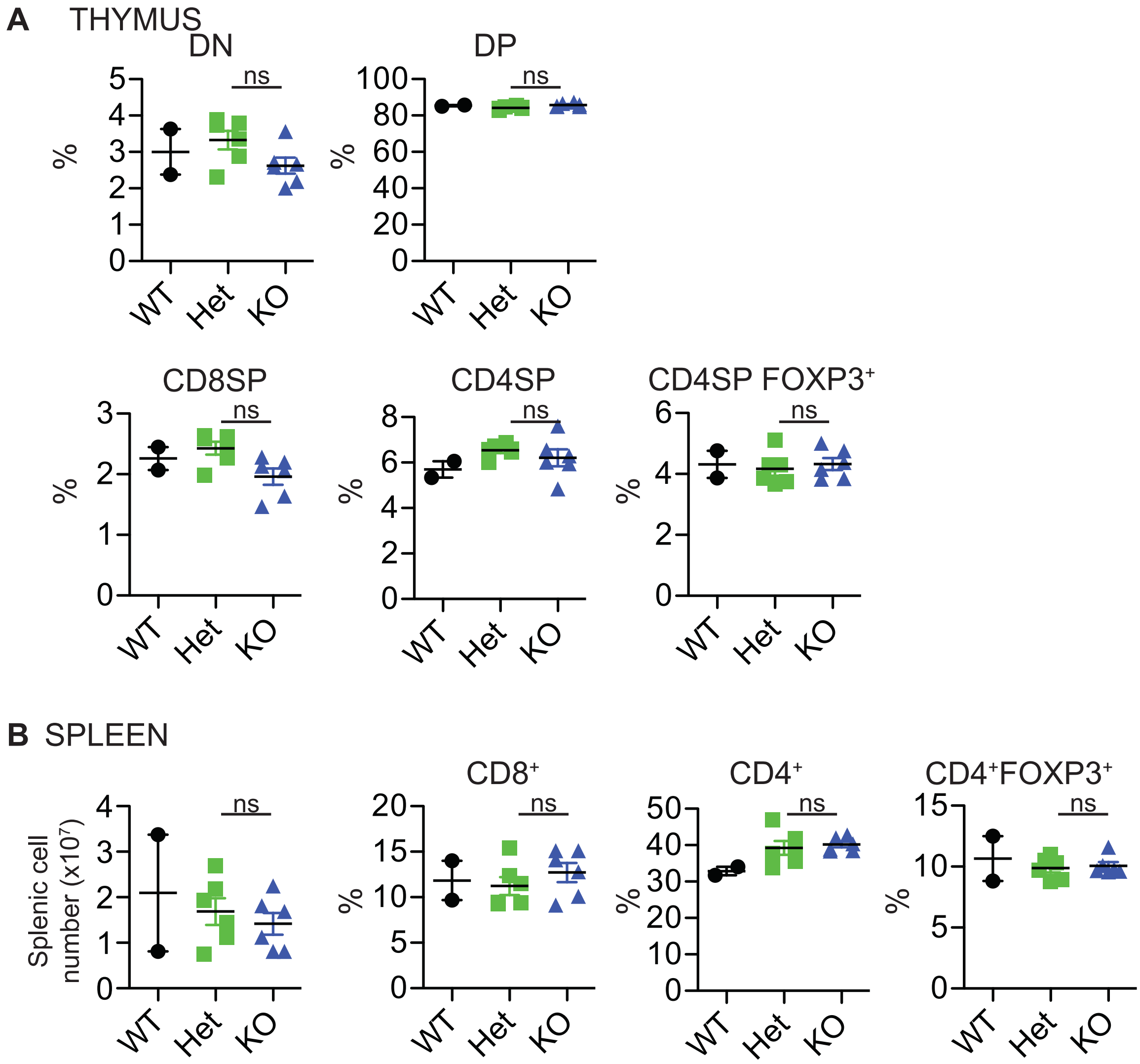

Figure S5: Distribution of major T cell subsets in thymus and spleen of infant *Tlr2*$^{-/-}$ mice

A. Frequency of thymic DN, DP, CD8SP, CD4SP and CD4SP FOXP3^{+} cells (WT n= 2; HET n=6; KO n=6).

B. Total splenic cellularity and frequency of CD8^{+}, CD4^{+} and CD4^{+}FOXP3^{+} cells (WT n= 2; HET n=6; KO n=6).

Data are from 2 experiments. Bars are Mean $\pm$ SEM.

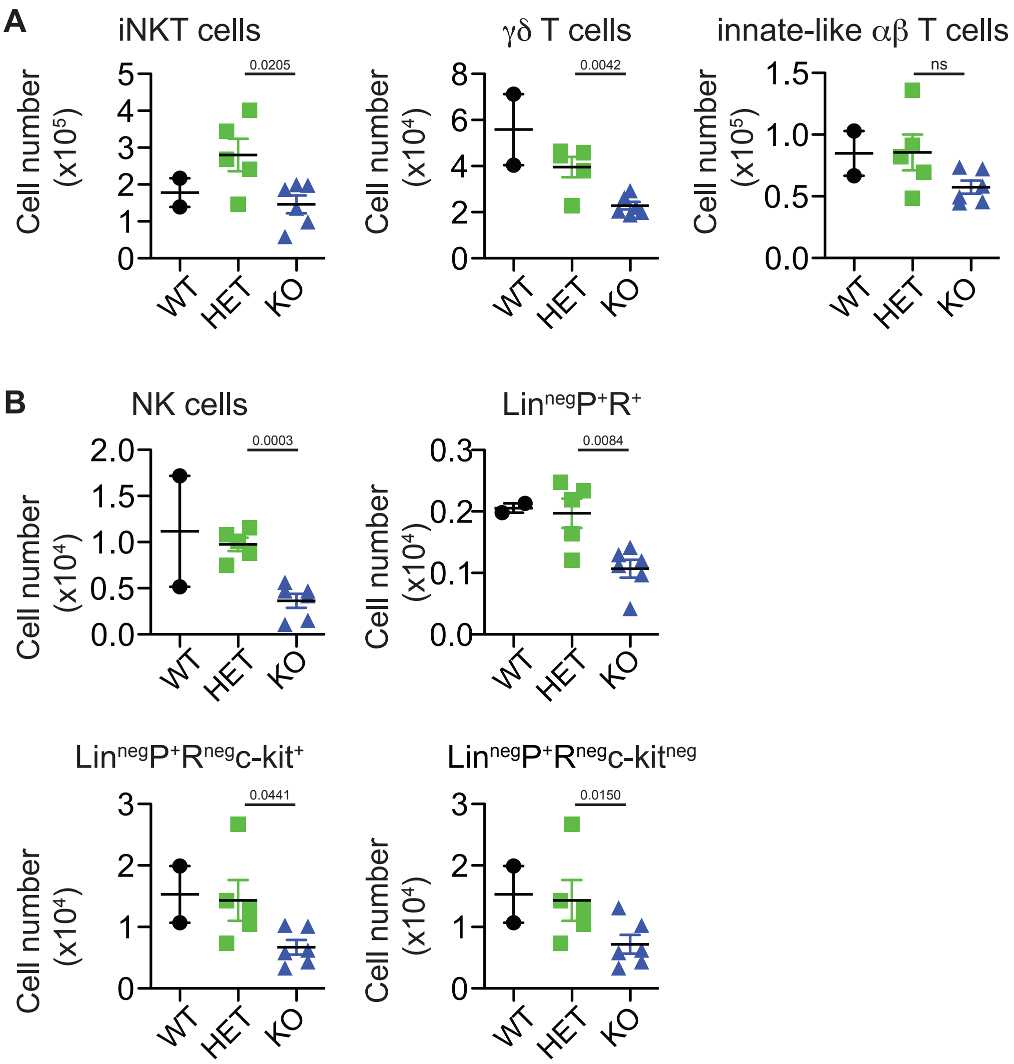

Figure S6: Cell numbers of PLZF-expressing cell subsets in thymus of infant *Tlr2*$^{-/-}$ mice.

A. Total numbers of PLZF^{+} iNKT cells, PLZF^{+} γδ T cells and PLZF^{+} innate-like αβ-T cells (WT n= 2; HET n=5; KO n=6).

B. Total numbers of PLZF^{+} NK cells, Lin^{neg}PLZF^{+}RORγt^{+} cells, Lin^{neg}PLZF^{+}RORγt^{neg}c-kit^{+} cells and Lin^{neg}PLZF^{+}RORγt^{neg}c-kit^{neg} cells (WT n= 2; HET n=5; KO n=6).

Data are from 2 experiments. Bars are Mean ± SEM.

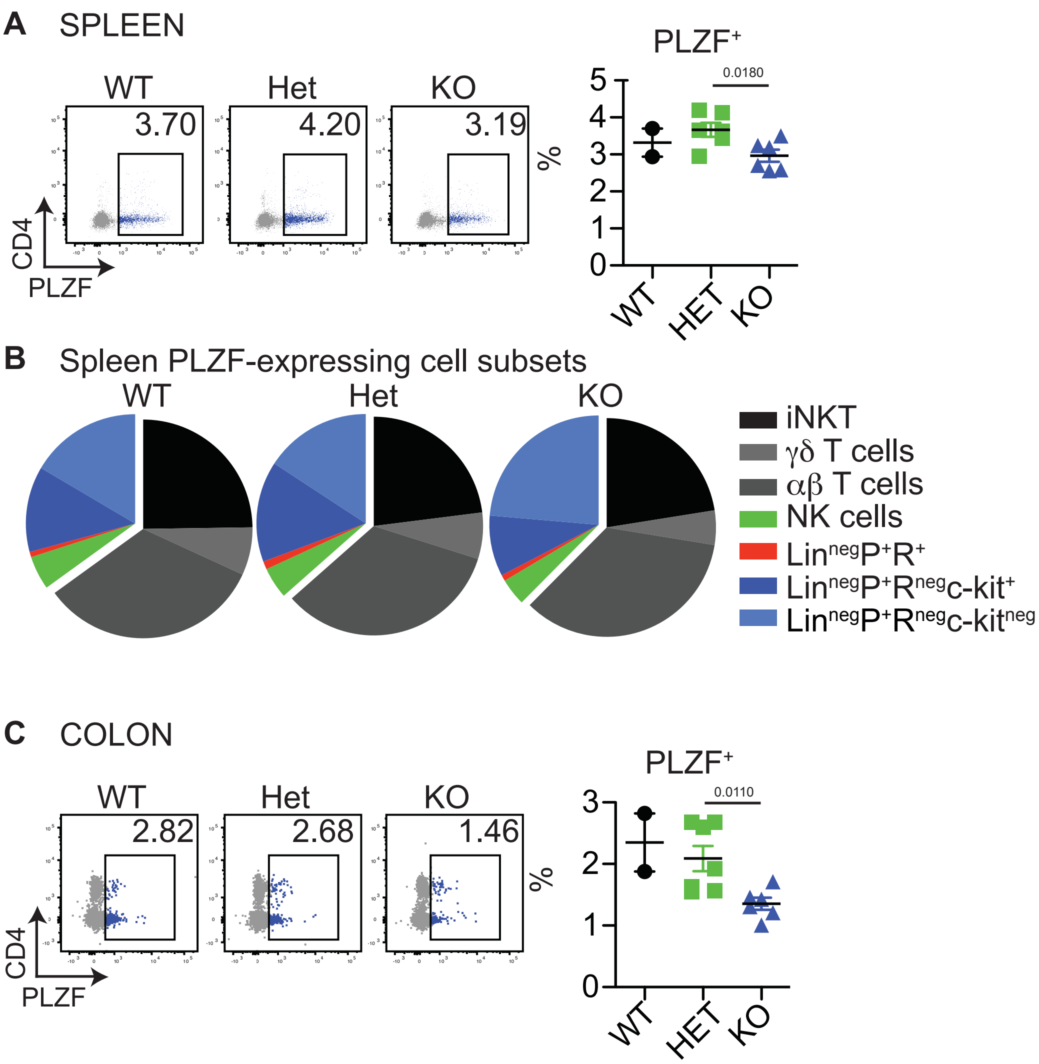

Figure S7: Frequency of PLZF expressing subsets in spleen and colon of infant *Tlr2*$^{-/-}$ mice.

A. Frequency of PLZF^{+} cells in spleen (WT n= 2; HET n=6; KO n=6).

B. Pie graphs showing distribution of indicated PLZF expressing cell subsets in spleen.

C. Frequency of PLZF^{+} cells in colon (WT n= 2; HET n=5; KO n=6).

Data in A-C are from 2 independent experiments. Bars are Mean ± SEM.

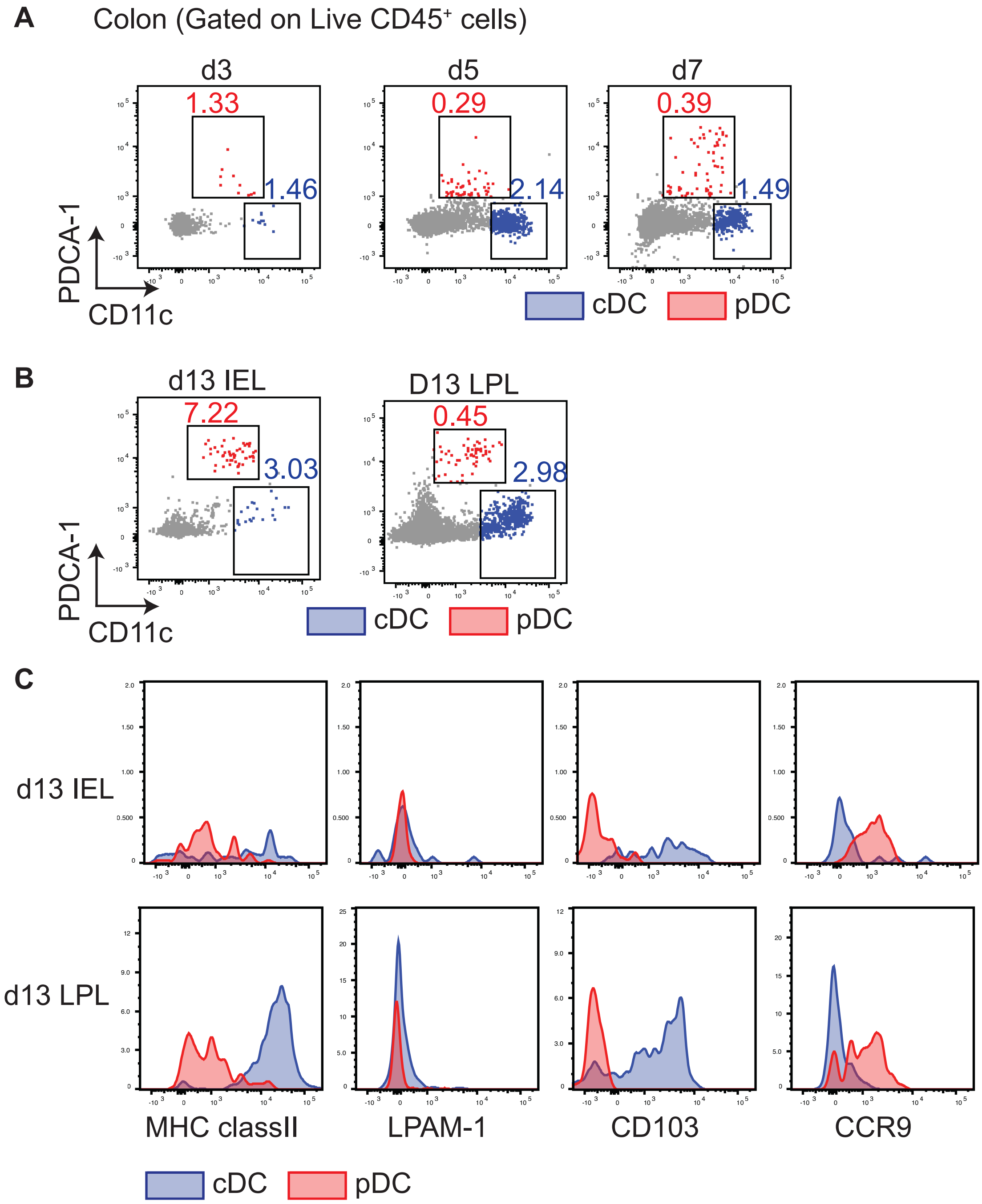

Figure S8: Colonic pDCs in conventionally-housed infant C57BL/6 mice.

A. Representative flow cytometry dot plots showing expression of CD11c (x-axis) and PDCA-1 (y-axis) in colonic lymphocyte preparations from 3, 5 and 7 days old C57BL/6 pups. Frequency of pDCs ($CD11^{+}PDCA\text{-}1^{+}$) and cDCs ($CD11c^{hi}PDCA\text{-}1^{neg}$) is shown.

B. Representative flow cytometry dot plots showing expression of CD11c (x-axis) and PDCA-1 (y-axis) in colonic intra-epithelial lymphocyte (IEL) and lamina propria lymphocyte (LPL) preparations from 13 days old C57BL/6 pups. Frequency of pDCS and cDCs is shown.

C. Histogram overlays of cDCs (blue) and pDCs (red) in IEL and LPL fractions showing expression of MHC class II, LPAM-1, CD103 and CCR9.

Data are representative of 2 independent experiments.

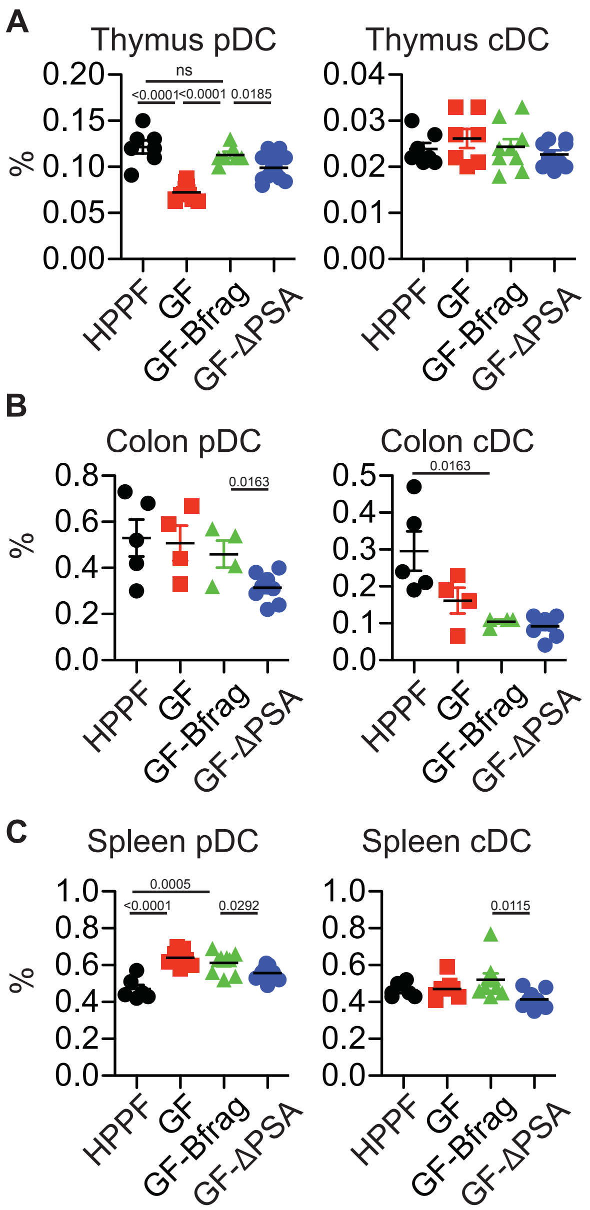

Figure S9: Distribution of pDCs and cDCs in colon, spleen and thymus of d14 GF and monocolonized mice

A-C. Frequency of pDCs and cDCs in the thymus (HPPF n= 7; GF n=12; GF-Bfrag n=7; GF-ΔPSA n=15), colon (HPPF n= 5; GF n=4; GF-Bfrag n=4; GF-ΔPSA n=8) and spleen (HPPF n=6; GF n=7; GF-Bfrag n=9; GF-ΔPSA n=9) is shown.

Data are representative of at least 2 independent experiments. Bars are Mean ± SEM.

Figure S10: RNA-seq analysis of thymic pDCs from *B. fragilis* (GF-Bfrag) or PSA mutant *B. fragilis* (GF-ΔPSA) monocolonized GF mice.															
	a1	**a2**	**a3**	**a4**	**a5**	**a6**	**a7**	**a8**	**a9**	**a10**	**a11**	**a12**	**a13**	**a14**	**a15**
1	Gm37347	Fam102a	Cyb5r1	Gpr174		Vcam1	Ccr7	C3	Cmah	Xpnpep3	Smco4	Zfp324		Xist	Spata2l
2	Gm13919	Ormdl1	Stxbp5	Mir378b		Zfp46	Gm14636	Csf1r	Gm38014	Slc28a2	Cd151			Pla2g2d	Rpp25l
3	Emid1	A530099J19Rik	H2-Ob	Tmem2		Rpain	Casz1	Arhgef2	Pou2af1		Dnajc11			Oprd1	Stard4
4	Gm8369	Fastkd5	Gm16982	Bend4		Nod1	Zc2hc1a	Sepp1	RP24-351O18.4		Tnfrsf21			Tnfrsf11a	Gm16337
5	Gm15903	Mcm8	Ndrg1	Gm13528		Meaf6	Timm22	Col27a1	Gm16310		Wdfy3			Wdr41	Ydjc
6	Gm38387	RP23-214I6.2	Ppap2a	Slc30a1		Nav1		Psen2	Slc43a2		Rab31			Taf5l	Il18
7	Slc36a3	Acot6	Aftph	Igkc		Atxn1		Xdh	Gm11131		Vps37c			Adam22	Oip5
8	Gdpd1		E130311K13Rik			Ccdc71		Phf11c	Pkd1l3		Fmn1			RP23-231L15.4	RP23-132D7.2
9	Gm24452		Fasn			Tiam1		Igsf3	B230317F23Rik		Ms4a6b			Clec4n	Adora2a
10	A630072L19Rik		Aire			Papss2		Rps13-ps1	Zbtb11os1		Gpnmb			Rnf145	Pla2g12a
11	Gtf3c5		Zak			Cd63		Plekhg2	Exd2		Clstn1			Slc12a2	Cd99l2
12			Ptpn22			Arrb1		Plekhn1	Tmem140		Il18bp			Ttn	Gm14730
13			Pygl			Anxa4		Bcor	Ppp1r9a		Rtp4			Nid1	
14			Bsdc1			Lrp5		Siglec1	RP23-158A20.7		Clec7a			Fam26f	
15			Itpripl2			Chst2		Itga9	Ccdc93		Med4			Rbbp8	
16			Matr3			Zdhhc24		Egr1	2210016L21Rik		Adck5			Oas2	
17			Tmem236			Gfra2		Ppm1g	Slc16a5		Bcl9			Tmprss6	
18			Rras2			Lpl		Hist1h2bl	1700001D01Rik		Galnt10			Anxa3	
19			F2rl2			Neurl1b		Taf11	Gm17146		Ostm1			Igf2r	
20			Gm4876			Polr3e		Iah1	Ldoc1l		Spg21			Adamdec1	
21			Lcorl			Scamp1		Furin	Zfp948		Rnf144b			Fgl2	
22			Klf8			Gca		Arl4c	RP23-454A14.10		Agrn			Xcr1	
23			Gm34727			Nlrp1b		Myo6	Wbscr16		Top3a			Rab3il1	
24			Tmem164			Gla		Kcnn4	Ccdc88c		Olfr433			Mex3b	
25			Disc1			Mcm10		Abcc5	Elmod3		Psme3			Slc37a3	
26			Gm37699			Hnrnpll		Gm14446	Cbr2		Asah2			Epb4.1l3	
27			Adap1			Slc7a8		Cyp4v3	Atxn7		Pex5			Brap	
28			Taf4b			Ddx60		Strip2	9230114K14Rik		Lyz2			Stxbp1	
29			Nccrp1			Hif1an		Cdc20	Gm14212		Ptms			Anpep	
30			Rarg			Itga6		Tbxas1	Malsu1		Tmem26			Ccl12	
31			Pet100			2900026A02Rik		Ptgs1	Msh5		Selo			Dusp16	
32			Phf19			Tut1		Gins2	Iffo1					Il6st	
33			Apol7c			Focad		Ifi204	Bsn					Ifnlr1	
34			Zfp277			Ttc39b		Osbpl3	Vsig10					Arhgef40	
35			Ptpn11			Pid1		Gm28933	Pias2					Ccdc91	
36			1700056E22Rik			Mllt6		RP23-2C16.2	Fsd1l					Dusp19	
37			Slc25a30					Ttyh3	Gm38147					9330175E14Rik	
38			Pabpc4					Zfp949	Rabl2					Gm37570	
39			F830112A20Rik					Mogs	L2hgdh					RP24-176P19.6	
40			H2-DMb2					Rpf2	I7Rn6					Scaf1	
41			Mboat1					Ddx49	Strbp					Asb2	
42			Tgif2					C1qc	Prdm11					Nab2	
43								Mtf1	Gm22639					9130221H12Rik	
44								Cisd1	Gm37368					Zfp366	
45								Kctd6	Adam33					Acp5	
46								Phtf2	Bcl7a					Slc36a1	
47								Zfp54	Rfx2					Trmt61a	
48								A730011C13Rik	Gm37873					Utp14b	
49								AI846148	Ubxn11					Sigmar1	
50								Clec4a1	Zfp524					Zfp180	
51								A430033K04Rik	Slc25a23					Gm15965	
52								Wdr78	Nudt7					Serpinf1	
53								Map4k5	Csrp2					Slc41a3	
54								Pck2	Ccl2					C77080	
55								Zfp874a	RP23-422D12.3					Nfrkb	
56								Gnal	Dok3					Doc2g	
57								Zfp395	Cyp26c1					Ggh	
58								Dnajc17	Impa1					Gid8	
59								Gbp2	Gm4258					Dnajb2	
60								Cdkn1a	Pla2g4f					Repin1	
61								Rpl36a	Nanos1					Timd4	
62								Fn1	Gm16675					Gas2l1	
63								Anxa1						Spred3	
64								Hmox1						Fosl2	
65								Pgap1						Loxl3	
66								Klri2						Lfng	
67								Abcc3						Haus4	
68								Timm23						Gm37590	
69								Trim23						Arntl	
70								Zcchc9						Rhobtb1	
71								Myof						2610012C04Rik	
72								Gm28466						Vps18	
73								Gcfc2						Gm29151	
74								Kcnj10						Acot11	
75								Rab9						Msh2	
76								Inip						Pld1	
77								Lig3						2810429I04Rik	
78								Fcgr1						Gm22733	
79														Jazf1	
80														Vipr1	
81														Fzd5	
82														Lrrc20	
83														Phldb1	
84														Aif1	
85														Polrmt	
86														Ms4a7	
87														Orm3	
88														Serf2	
89														Wdr34	
90														Prdm1	
91														Ect2	
92														Rgs3	
93														Stap2	
94														Pitpnm1	
95														Oas1b	
96														Lpcat4	
97														Chmp7	
98														Mettl13	
99														Mical2	
100														Gp49a	
101														Gm13710	
102														Olfr99	

Figure S10: RNA-seq analysis of thymic pDCs from GF, GF-Bfrag and GF-∆PSA monocolonized mice

Genes in each numbered section (superscript) of the Venn diagram in **Figure 3F**.

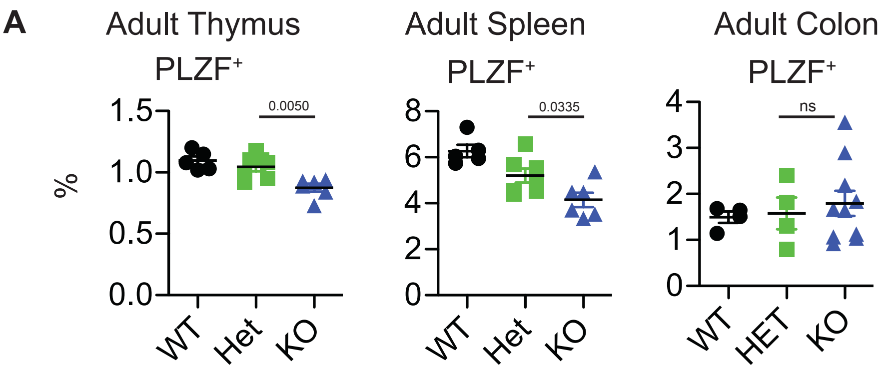

Figure S11: Frequency of PLZF expressing subsets in thymus, spleen and colon of adult *Tlr2*$^{-/-}$ mice.

A. Frequency of PLZF^{+} cells in thymus, spleen and colon of adult (d48-55) mice (Thymus, Spleen: WT n= 5; HET n=7; KO n=6; Colon: WT n= 4; HET n=4; KO n=10).

Bars are Mean ± SEM.

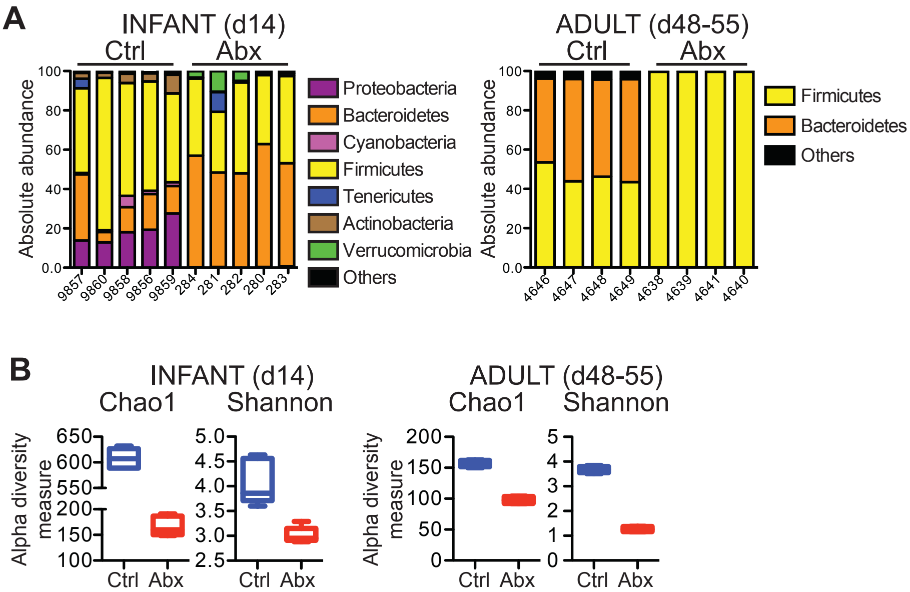

Figure S12: Colonic microbial diversity of antibiotic treated infant and adult mice

A. 16S sequencing of colonic content for bacterial diversity from infant (Ctrl n=5; Abx n=5) and adult mice that were treated with antibiotics in early life (Ctrl n=4; Abx n=4) was performed. Absolute abundance at phylum level is shown.

B. Chao1 and Shannon Index showing alpha diversity for complexity of microbiota within infant and adult groups.

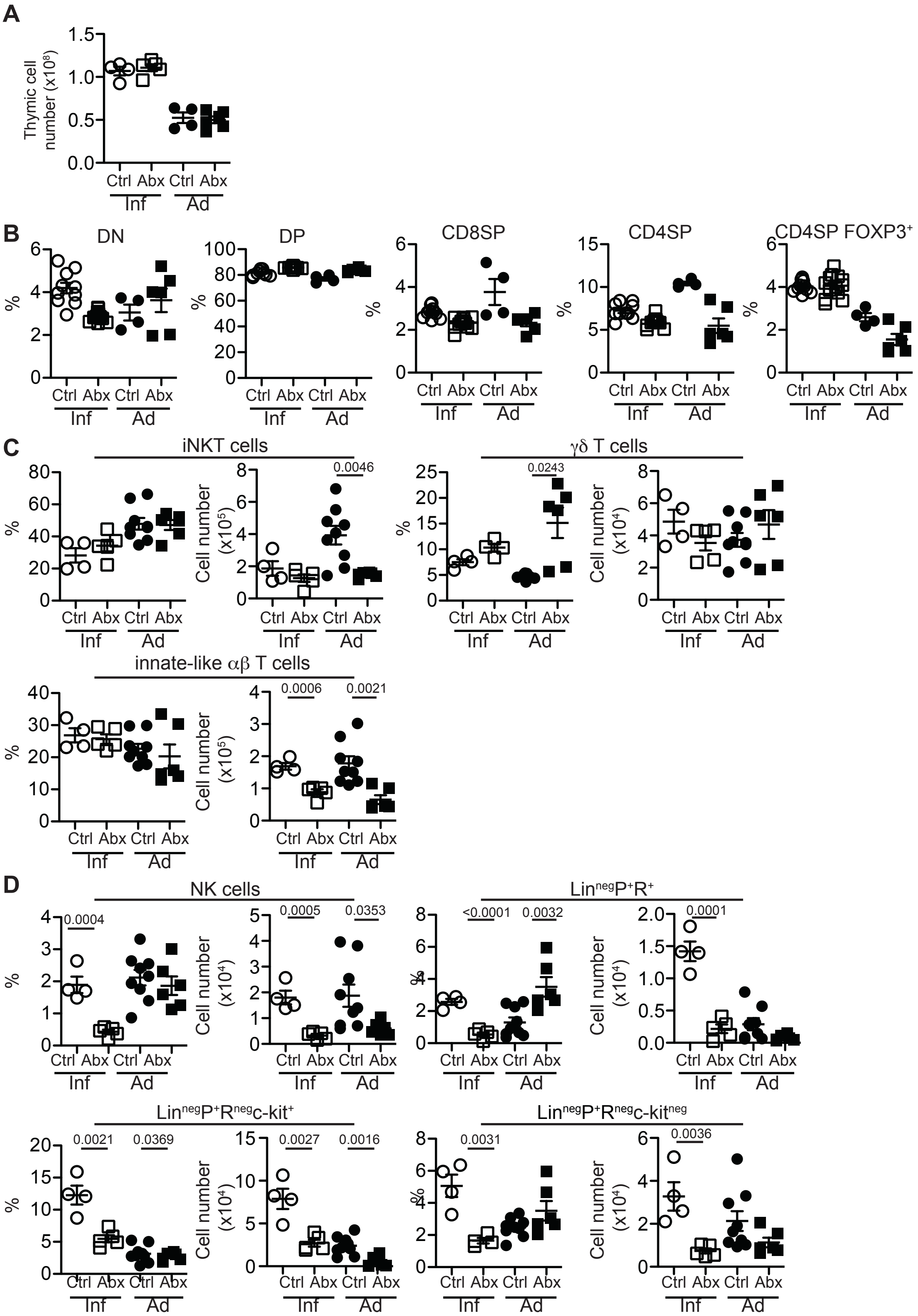

Figure S13: Distribution of major T cell subsets and PLZF expressing cell subsets in thymus of antibiotics treated infant and adult mice.

A. Total thymic cellularity (Inf-Ctrl n=4; Inf-Abx n=5; Ad-Ctrl n=4; Ad-Abx n=6).

B. Frequency of thymic DN, DP, CD8SP, CD4SP and CD4SP FOXP3⁺ cells (Inf-Ctrl n=10; Inf-Abx n=13; Ad-Ctrl n=4; Ad-Abx n=6).

C. Frequency and numbers of PLZF⁺ iNKT cells, PLZF⁺ γδ T cells and PLZF⁺ innate-like αβ-T cells (Inf-Ctrl n=4; Inf-Abx n=5; Ad-Ctrl n=9; Ad-Abx n=6).

D. Frequency and numbers of PLZF⁺ NK cells, $Lin^{neg}PLZF^{+}ROR\gamma t^{+}$ cells, $Lin^{neg}PLZF^{+}ROR\gamma t^{neg}c\text{-}kit^{+}$ cells and $Lin^{neg}PLZF^{+}ROR\gamma t^{neg}c\text{-}kit^{neg}$ cells in the thymus (Inf-Ctrl n=4; Inf-Abx n=5; Ad-Ctrl n=9; Ad-Abx n=6).

Data are from 2 independent experiments for each group. Bars are Mean ± SEM.

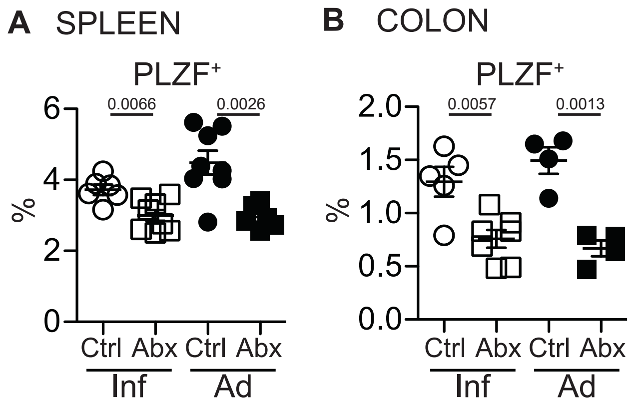

Fig. S14: Frequency of PLZF⁺ cells in the spleen and colon of antibiotic treated infant and adult mice.

A. Frequency of PLZF⁺ cells in spleen (Inf-Ctrl n=6; Inf-Abx n=8; Ad-Ctrl n=8; Ad-Abx n=6).

B. Frequency of PLZF⁺ cells in colon (Inf-Ctrl n=5; Inf-Abx n=7; Ad-Ctrl n=4; Ad-Abx n=4).

Data are representative of 2 independent experiments. Bars are Mean ± SEM.