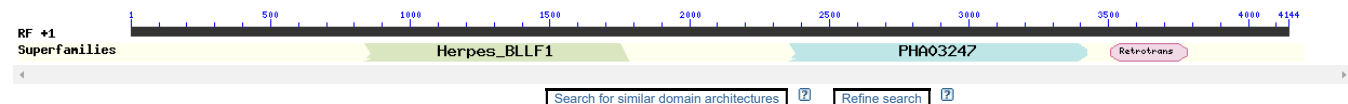


Fig. S1 Comparison of the RTL9 protein in the 10 representative eutherian species

Multiple alignment of the RTL9 protein sequences in 10 eutherian species: the mouse, rat, human, chimpanzee, dog, horse, elephant, manatee, armadillo and sloth. The identical amino acids in the 10 species are depicted with an asterisk (*) below the alignment. Three regions corresponding to the herpes BLLF1 super family, PHA03247 super family and capsid domain of sushi-ichi GAG (Fig. 1A) are shown as gray, orange and blue boxes, respectively. Maximal and minimal GAG-like regions are shown as red boxes and red dashed boxes, respectively. The end of RTL9 Δ C, just before the maximal GAG-like region, is indicated by a red arrow with a short vertical line on its head. Details are provided in the Comparative Genome analysis in the Materials and Methods section.

View Concise Results ▾ ?

Graphical summary Zoom to residue level [show extra options »](#)

?

Name	Accession	Description	Interval	E-value
Retrotransposon gag super family	cd14948	Retrotransposon gag protein; Gag or Capsid-like proteins from LTR retrotransposons. There is a ...	3505-3780	7.31e-08
Retrotransposon gag protein; Gag or Capsid-like proteins from LTR retrotransposons. There is a central motif GGXXEXXXXXXXLXXKH that is common to Retroviridae gag-proteins, but is poorly conserved.				

The actual alignment was detected with superfamily member [pfam03732](#):

Pssm-ID: 455293 Cd Length: 97 Bit Score: 51.56 E-value: 7.31e-08

NM_001040434.2 1169 LVS^{1.1}SL^{1.1}H^{1.1}-GA^{1.1}ER^{1.1}W^{1.1}-S^{1.1}L^{1.1}Q^{1.1}M^{1.1}E^{1.1}V^{1.1}G^{1.1}N^{1.1}P^{1.1}I^{1.1}S^{1.1}S^{1.1}D^{1.1}-K^{1.1}A^{1.1}F^{1.1}L^{1.1}R^{1.1}sqg^{1.1}I^{1.1}Y^{1.1}D^{1.1}S^{1.1}L^{1.1}E^{1.1}I^{1.1}D^{1.1}I^{1.1}L^{1.1}S^{1.1}A^{1.1}V^{1.1}L^{1.1}C^{1.1}H^{1.1}P^{1.1}K^{1.1}Q^{1.1}G^{1.1}K^{1.1}S^{1.1}V^{1.1}R^{1.1}Q^{1.1}Y^{1.1}A^{1.1}T^{1.1}D^{1.1}F^{1.1}L^{1.1}L^{1.1}R^{1.1}-1241
 Cdd:pfam03732 2 LAVHSLRGAALTWKLSVARSIDAFDSWDEIKDAFLKR--FFPSTKIDLLRNELSLRGTGTSVREYVERFKRLARa1 77

NM_001040434.2 1242 **HL**SWSDAIL**RTRF**LEGLSE 1260
Cdd:pfam03732 78 p**H**GRDEEAL**ISAF**LRGLRP 97

PHA03247 super family [cl33720](#) large tegument protein UL36; Provisional
large tegument protein UL36; Provisional

2359-3423 1.61e-06

The actual alignment was detected with superfamily member [PHA03247](#):

Pssm-ID: 223021 [Multi-domain] Cd Length: 3151 Bit Score: 53.02 E-value: 1.61e-06

NM_001040434.2 787 LVRPPASGEIAPHSRTPYVGTISAPHMITTTASGVMTSPMKTSVPVSESATLLRPTDSGVMSITPLTRTPASRAKSRQMA 866
 Cdd:PHA03247 2698 LADPPPPPTPECPAPHALVSATLPGGAAARQASPALPAAPAPPVAPGATPGGPAPRPPTTAGPPAPAPPAAPAA 2777

NM_001040434.2 867 TAcgdmcp I PVRAPATAGISPSVRSPASS TstlIrr PSDGAVTAE LERVL GP AQFAAMTPGEMSKPL MRA —SAPGTT 944
 Cdd:PHA03247 2778 GP PRLTRPAVASLSESRESLPS PWDPADPAVLAPAAALPPAASPAGLPPPTSA_{qp}TAPPPP 2842

NM_001040434.2 945 **T**M**L****S****P****M****T****S****G****E****M****S****P****L****M****K****T****T****P****S****G****T****M****S****T****L****Q****T****K****V****M****S****R****A****T****S****L****P****Q****P****r****n****a****a****s****g****i****a****n****P****P****R****A****P****A****S****G****A****S****T****L****M****R****V****S****S****G****M** 1021
 Cdd:PHA03247 2843 **P****G****P****P****P****S****L****P****G****S****V****A****P****g****g****d****V****R****R****R****P****P****S****R****S****P****A****K****P****A****A****P****P****V****R****L****A****R****P****A****V****S****R****S****T****E****F****A****L****P****D****Q****E****R****P****P****G****M** 2912

NM_001040434.2 1022 MSTPLLGATSGGMSMQMPPPTSGDMFSLMRSPAGGI MSTPQtafGIMTPLLNKATDSEASTSHTRftapgskSTPH 1101
 Cdd:PH03247 2913 APPPPPPQPPPPPPQPPPPPPPPRQPLAPTTDPAGAGEPSGAVPQPNLGLVPRGVPFRFRVPPQPA 2983

NM_001040434.2 1102 MTSTAPEMKTPPPKEVPSPFGMLTPALCYLLEEQEAAAGSS 1141
 Cdd:PHA03247 2984 PSREAPASSTPLTGHSLSRVSSWASSLLAHEETDPPPVSS 3023

Fig. S2-1

Herpes_BLLF1 super family [cl37540](#) Herpes virus major outer envelope glycoprotein (BLLF1); This family consists of the BLLF1 ... 838-1782 5.67e-03
 Herpes virus major outer envelope glycoprotein (BLLF1); This family consists of the BLLF1 viral late glycoprotein, also termed gp350/220. It is the most abundantly expressed glycoprotein in the viral envelope of the Herpesviruses and is the major antigen responsible for stimulating the production of neutralising antibodies in vivo.

The actual alignment was detected with superfamily member [pfam05109](#):

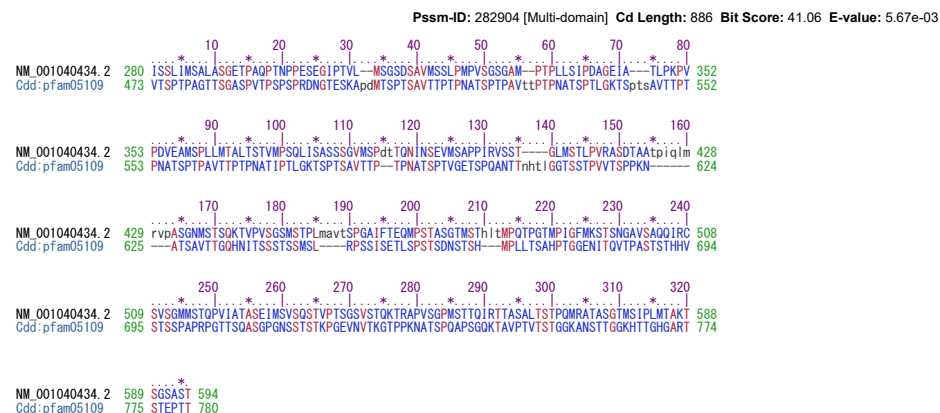


Fig. S2 The result of a conserved domain search using NCBI CDD

Three domains, the retrotransposon gag protein, large tegument protein UL36 and herpes virus major outer envelope glycoprotein (BLLF1), are listed in this order. The identical amino acids in the superfamily motifs and RTL9 are depicted in red.

Fig. S2-2

Mouse RTL9 pfam05109	1	ISSLIMSALASGETPAQPTNPPESGIP TVL--MSGSDSAVMSSLPMPVSGSGAM--PTPLLSIPDAGEIA---TLPKPV	80
	1	VTSTPTAGTTSGASPVTPSPSPRDNGTESKApdMTSPTS AVTTTPNATSPTPAVttTPNATSPTLGKTSptsAVTTPT	80
Mouse RTL9 pfam05109	81	PDVEAMSPLLMTALTSTVMPSQLISASSSGVMSPdtTQNINSEVMSAPP IRVSST---GLMSTLPVRASDTAAtpiqlm	160
	81	PNATSPTPAVTTTPNATIPTLGKTSPTS AVTTP--TPNATSPTVGETSPQANTTnhtlGGTSSTPVVTSPPKN-----	160
Mouse RTL9 pfam05109	161	rvpASGNMSTSQKTVPVSGSMSTPLmavtSPGAIFTEQMPSTASGTMSThltMPQTPGTMPIGFMKSTSN GAVSAQQIRC	240
	161	---ATSAVTTGQHNITSSSTSSMSL----RPSSISETLSPSTSDNSTSH---MPLL TSAHPTGGENITQVTPASTSTHHV	240
Mouse RTL9 pfam05109	241	SVSGMMSTQPVIATASEIMSVSQSTVPTSGSVSTQKTRAPVSGPMSTTQIRTTASALTSTPQMRATASGTMSIPLMTAKT	320
	241	STSSPAPRPGTTSQASGPGNSSTSTKPGEVNVTKGTPPKNATSPQAPSGQKTAVPTVTSTGGKANSTGGKHTTG HGART	320
Mouse RTL9 pfam05109	321	SGSAST	326
	321	STEPTT	326

Herpes_BLLF1 super family
identity 73/326 = 22.4%
Similarity 125/326 = 38.3%

Fig. S3 Pairwise alignment and identity between pfam05109 (BLLF1) and mouse RTL9

The N-terminal part of mouse RTL9 (280 aa – 594 aa) possessing homology to herpes BLLF1 super family motif was re-analyzed using EMBOS Water program and EMBOSS needle program as described in Materials and Methods. One and two dots represent similarities of amino acids, such as weekly and strong similar, respectively.

Mouse RTL9 PHA03247	1	LVRPPASGEIAPHSRTPVYGTISAPHMTTASGVMTMSPMKTSVPVSESATLLRPTDSGVMSIPLTRTPASRAKSRPQMA	80
	:.....:.....:	
	1	LADPPPPPTPEPAPHALVSATPLPPGPAAARQASPALPAAPAPPAVPAGPATPGGPAPPARPPTTAGPPAPAPPAAPAA	80
Mouse RTL9 PHA03247	81	TAcgdmcp1PVRAPATAGISPSPVRSPASSt1stllrrPSDGAVTAELERVLGPAQFAAMTPGEMSKPLMRA--SAPGTT	160
	: : ...	
	81	GP-----PRRLTRPAVASLSESRESLPS-----PWDPADPPAAVLAPAAALPPAASPAGPLPPPTSaqTAPPPP	160
Mouse RTL9 PHA03247	161	TMPLMSPMTSGEMSMP---LMKTTPSGTMSTLQTKVMSSRATSLPQPrnaasgvianPPQRAPASGAGSTPLMRVSGSGM	240
	:.. :.... :.....:..... .: :..	
	161	PGPPPPSLPLGGSVAPggdVRRRPPSRSPAAPKPAAPARPPVRLARP-----AVSRSTESFALPPDQPERPPQSQ	240
Mouse RTL9 PHA03247	241	MSTPLLGATASGGMSMPQMAPPTSGDMFSPLMRSPAPGIMSTPQtafGMTPTLNKATDSGEASTSHTRFtapgskSTPH	320
	 :.... :..... . ..	
	241	APPPPPQPPQPPPPPPQPPPPPPPPRQPPLAPTTPAGAGEPS---GAVPQPWL GALVPGRVAVPRFRV-----PQPA	320
Mouse RTL9 PHA03247	321	MTSTAPEMKTPPPKEVPSFGMLTPALCYLLEEQAARGSS	360
	:..:.....	
	321	PSREAPASSTPPLTGHSLSRVSSWASSLALHEETDPPPV	360

PHA03247 super family
identity 63/360 = 17.5%
similarity 85/360 = 23.6%

Fig. S4 Pairwise alignment and identity between PHA03247 family and mouse RTL9 protein

The middle of mouse RTL9 (787 aa –1141 aa) possessing homology to the PHA03247 super family motif was re-analyzed using the EMBOSS Water program and EMBOSS needle program. One and two dots represent similarities of amino acids, such as weekly and strong similar, respectively.

	131	ELLFRHQPSRFVSDEAKVGFITSLLADKALSWAIAAVDLDPRLLSSDYSAF	180
GAG		:..... ... :.. .. :.. ::	
Mouse RTL9	1148	EIDEEKQMKGFLDDSEKMAFLVSLHLGAAERWSILQMEVGNPISSDNKAF	1197
	181	RREFKAVFEHPTYGEDAASRLALQQGSRVAEYTLFRILAAESRWGET	230
GAG	: .. :.. .. :.. .. :..	
Mouse RTL9	1198	LRRSQGLYDSLSEIDILSAVLCHPKQGKKSVRQYATDFLLARHLSWSDA	1247
	231	ALRSAYRRGLSEAIKDLIVR---DRPSSLNELITLSLQMDERLRERRQER	277
GAG		. :..... .. :.. .. :.. .. :..	
Mouse RTL9	1248	ILRTRFLEGLSEAVTTKMGRIFLKVAGSLKELIDRSLYTECQLAEKD--	1295
	278	AQRAGGSTRQLSHRTSSAPDFSLTSTAAPPPHILLQSPAHPSPRVGEEPM	327
GAG		:. :.. :..	
Mouse RTL9	1296	---SSGNSNQV-----VP-----TS-----CKRNNEEAM	1316
	328	--QIGRSRLSRQEREQRLR-DQLCLYCGNNG-----HFIQACPVRPK	366
GAG		:: .. :.. :..	
Mouse RTL9	1317	ENELG---SQQQTEEHQHVPKRCYYLKEHGDPEESLHDHLRQSAGL-PK	1362
	367	GPAHQ	371
GAG		. :::	
Mouse RTL9	1363	APT ^{Red} NK	1367

Red: Amino Acids encoded in the exon 2
Light blue: CCHC RNA-binding motif

Sushi-ichi GAG
identity 65/255 = 25.5%
similarity 110/255 = 43.1%

Fig. S5 Pairwise alignment and identity between sushi-ichi GAG and mouse RTL9 (1)

A conserved domain search using NCBI CDD indicated that the C-terminal portion of mouse RTL9 (1169 aa –1260 aa) possesses homology to the Capsid motif of sushi-ichi retrotransposon GAG. However, the regions before and after the Capsid motif (1148 aa-1168 aa and 1261 aa-1367 aa) also exhibit homology to GAG (**Minimal GAG-like region**). The AAs in *Rtl9* exon 2 are shown in red. The AAs of the CCHC motif of GAG are shown in light blue. One and two dots represent similarities of amino acids, such as weakly and strong similar, respectively.

GAG	6	TPGQ-----PGGPRTPTLP--SPLERRVEAHSACLSSSLQSELTKAFPT	46
Mouse RTL9	927	TPGEMSKPLMRASAPGTTTLMPLMSPMT-----SGEMS---MPLMKTTPS	967
GAG	47	IQGEISELQSSSQTTSSSTLSALSNOQMSAMATVLASIIQK-----LGSDP-	90
Mouse RTL9	968	--GTMSTLQ--TKVMSSRATSLPQPRNAASGVIANPPQRAPASGAGSTPL	1013
GAG	91	-----GGAAPSEPSLP--LSPRAEPNLASP-----	113
Mouse RTL9	1014	MRVSGSGMMSTPLLGAATSGGMSMPQMAPPTSGDMFSPLMRSPAPGIMST	1063
GAG	114	--RVFG-----GDFDL-----GKGFLHQ-----	129
Mouse RTL9	1064	PQTAFGMTPTLNVKATDSGEASTSHTRFTAPGSKSTPHMTSTAPEMKTTP	1113
GAG	130	-----CELLFRHQPSR-----FVSDE	145
Mouse RTL9	1114	PKEVPSFGMLTPALCYLLEEQAARGSSSVVEEDAEEDIDEEKQMKGFLDDS	1163
GAG	146	AKVGFITSLLDKALSWAIAAVDLDPRLSSDYSAFRREFKAVFEHPTYGE	195
Mouse RTL9	1164	EKMAFLVSLHLGAAERWSILQMEVGNPISSDNKAFLRRSQGLYDSLSEID	1213
GAG	196	DAASRLALQQGSRSVAEYTLFRILAAESRWGETALRSAYRRGLSEAIK	245
Mouse RTL9	1214	ILSAVLCHPKQGKKSVRQYATDFLLARHLWSDAILRTRFLEGLSEAVT	1263
GAG	246	DLIVR---DRPSSLNELITLSLQMDERLRERRQERAQRAGGSTRQLSHRT	292
Mouse RTL9	1264	TKMGRIFLKVAGSLKELIDRSLYTECQLAEKD-----SSGNSNQV----	1304
GAG	293	SSAPDFSLTSTAAPPHILLQSPAHPSPRVGEEPM--QIGRSRLSRQERE	340
Mouse RTL9	1305	--VP---TS-----CKRNNEEAMENELG-----SQQQTE	1328
GAG	341	QRLR-DQLCLYCGNNG-----HFIQACPVRPKGPAHQ	371
Mouse RTL9	1329	EHQHVPKRCYYLKEHGDPQESLHDHLRQSAGL-PKAPTNNK	1367

Sushi-ichi GAG
identity 108/490 = 21.8%
similarity 175/490 = 35.7%

Fig. S6 Pairwise alignment and identity between sushi-ichi GAG and mouse RTL9 (2)

Entire sushi-ichi GAG exhibits homology to mouse RTL9 (927 aa-1357 aa, **Maximal GAG-like region**).

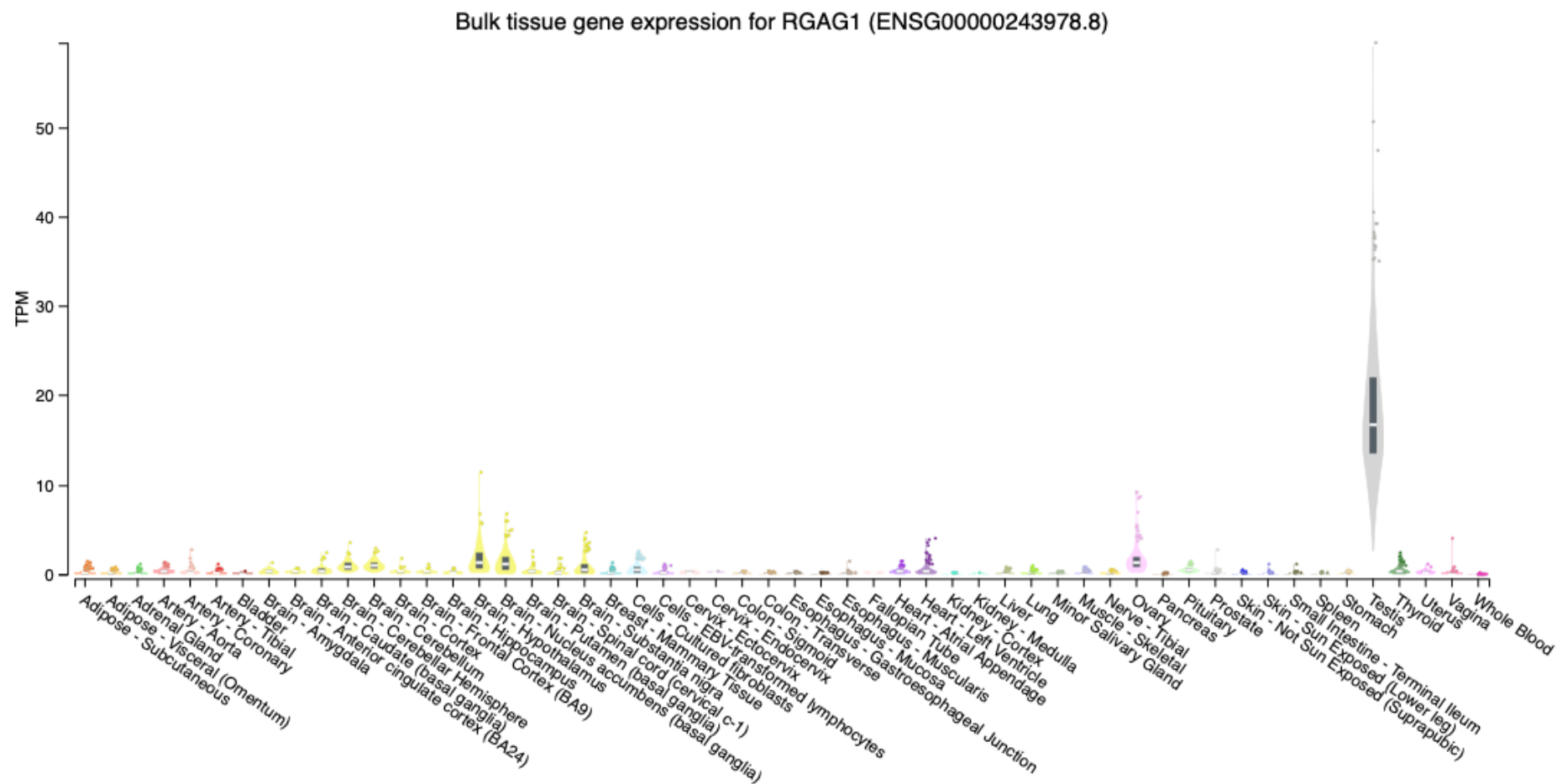


Fig. S7 *RTL9* (*RGAG1*) expression in human tissues and organs

See the original data at <https://gtexportal.org/home/gene/RGAG1> (accessed on 2 June 2023)

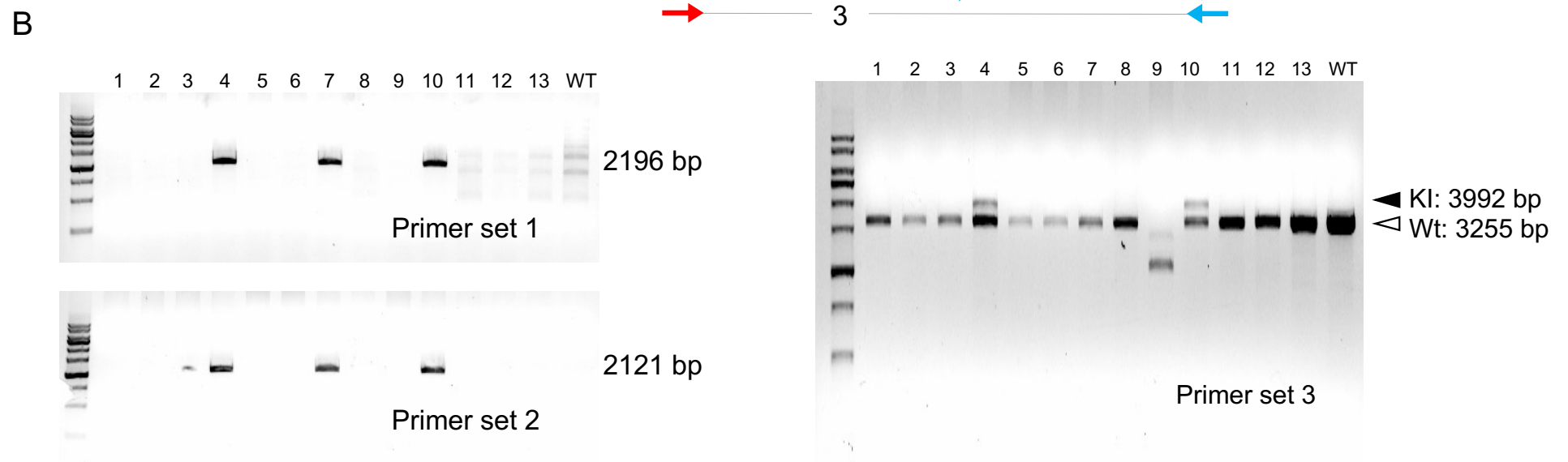
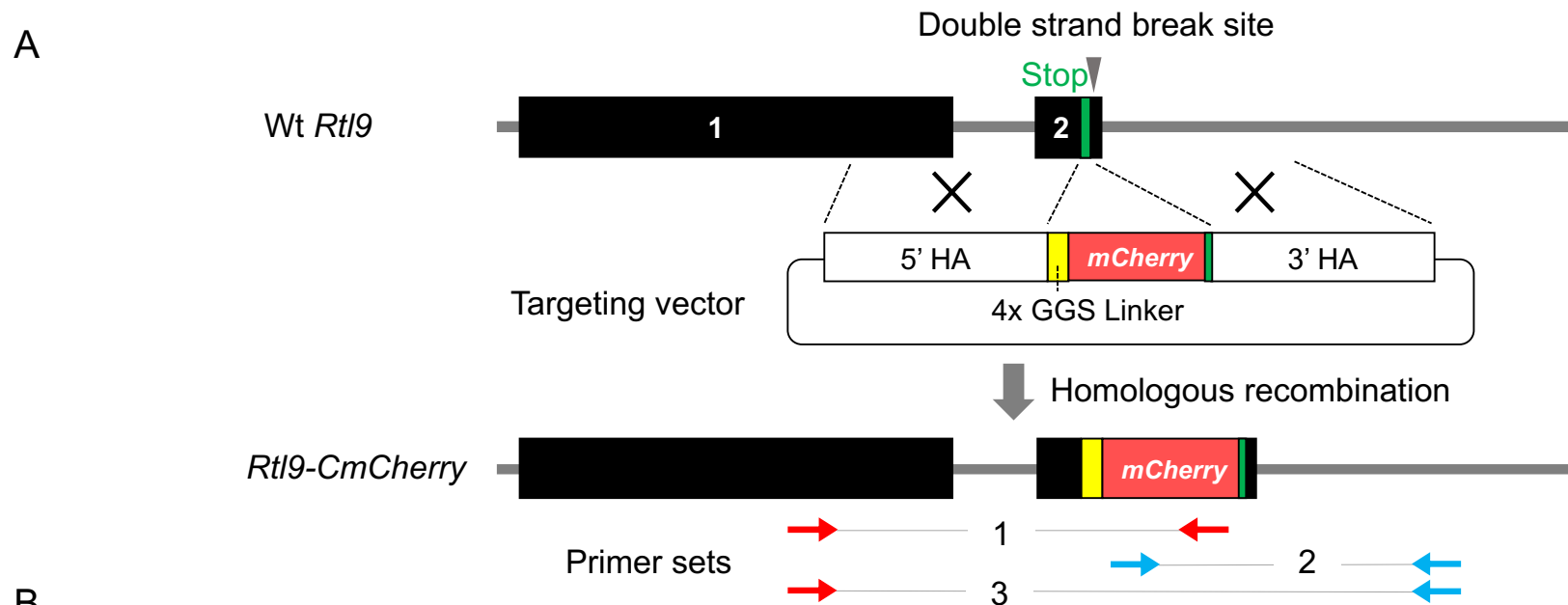


Fig. S8 Making of *Rtl9-mCherry* KI mouse

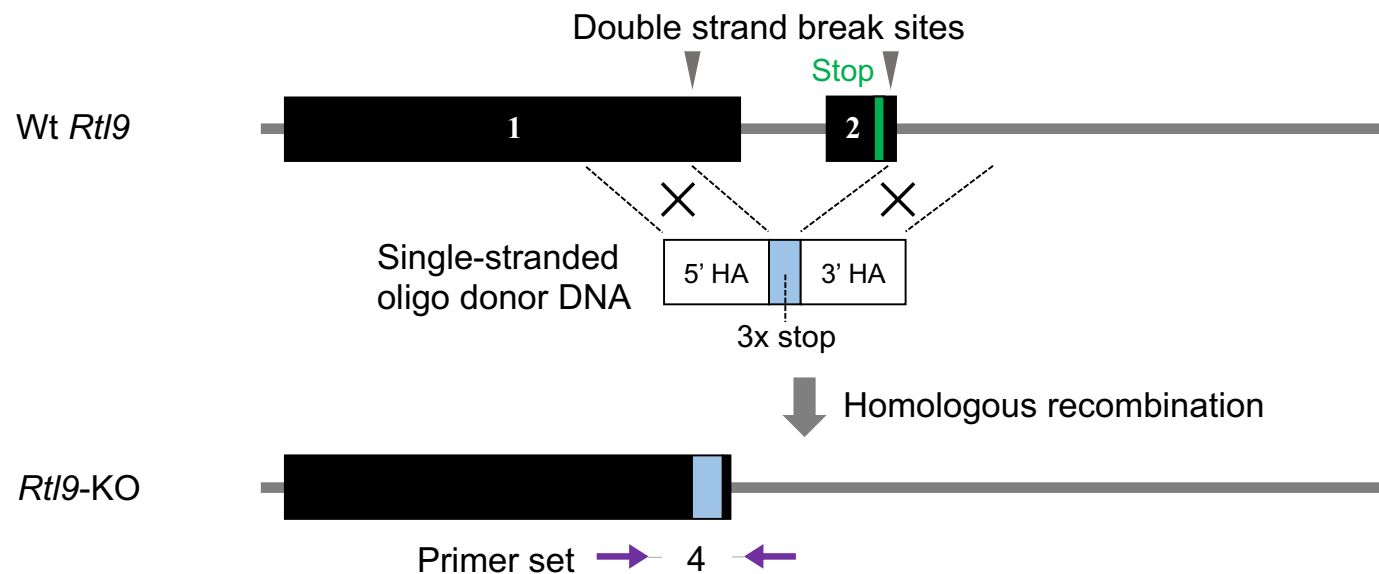
Fig. S8 Generation of the *Rtl9CmC* KI mouse

A) Schematic representation of the generation of the *Rtl9CmC* KI mouse. *Rtl9* exon 1 and 2 are numbered and shown in black. The stop codon is shown in green. The genomic locus near the stop codon in exon2 of *Rtl9* was cut out using the CRISPR/Cas system and homologous recombination was induced between the Wt *Rtl9* locus and a targeting vector containing 5' and 3' homology arms (5' HA and 3' HA), a 4x GGS linker and the mCherry coding sequence. The reconstructed *Rtl9CmC* has mCherry at the C-terminus. The position of the PCR primer sets for genotyping is indicated (red and light blue arrows). Cross symbols means homologous recombination. **B)** The results of genotyping 13 founder (F0) mice. Genomic PCR was performed using the primer sets 1 to 3 shown in A. The number (#) of F0 mice is indicated above the gel images. Wild type (WT) mouse genome was used as a control. Based on the genotyping PCR data, the offspring of F0 mouse #4 and #10 were used for further experiments.

PCR primer sets for genotyping of KI and KO mice

Mutant	Primer set name	Direction	Sequence
<i>Rtl9-CmCherry</i>	Primer set 1	F	GGAATGATGTCCACGCCACTA
		R	CTTCAGCTTCAGCCTCTGCT
	primer set 2	F	CCTGTCCCCTCAGTTCATGT
		R	CCTAGACTATTGGACCAGAGG
	Primer set 3	F	GGAATGATGTCCACGCCACTA
		R	CCTAGACTATTGGACCAGAGG
<i>Rtl9-KO</i>	Primer set 4	F	GAGAACACCAGCTTCTAGAGC
		R	GGGAGTTCAGAACCTCATACAC

A



B

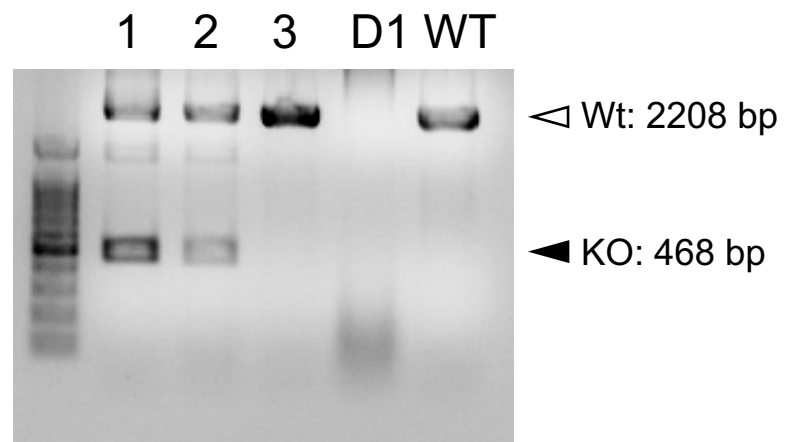


Fig. S9-1 Making of *Rtl9* KO mouse

C



Fig. S9-2 Making of *Rtl9* KO mouse

Fig. S9 Making of *Rtl9* KO mouse

A) Schematic representation of the generation of the *Rtl9* KO mouse. *Rtl9* exons 1 and 2 are numbered and shown in black. The stop codon is shown in green. The genomic locus encoding alanine 921 (A921) and downstream of the *Rtl9* stop codon was cut using the CRISPR/Cas system, and homologous recombination was induced between the Wt *Rtl9* locus and single-stranded oligo donor DNA containing a sequence of 3 stop codons (3x stop) and 5' and 3' homology arms (5' HA and 3' HA). In the reconstructed *Rtl9*-KO allele, the genomic region from glutamine 922 to the stop codon of *Rtl9* was deleted and replaced by a 3x stop. The position of the PCR primer set used for genotyping is indicated as purple arrows. **B)** Results of genotyping of 3 surviving and 1 dead (D1) F0 mice. Genomic PCR was performed with the primer set 4 shown in A. The number (#) of F0 mice, including D1, is indicated above the gel image. Wild type (WT) mouse genome was used as a control. **C)** Illustrated results of Sanger sequencing using the founder (F0) mice #1 and 2 shown in B. In F0 mouse #1 (top), a 3x stop was inserted at the expected genomic position just after A921 of *Rtl9*. In #2 (bottom), 12 extra bases (highlighted in red) were unexpectedly inserted just before a stop codon (highlighted in grey). Based on the Sanger sequencing data, the offspring of F0 mouse #1 were used for further experiments. The PCR primer sets for genotyping the KO mice are presented in the Fig. S6 legend.

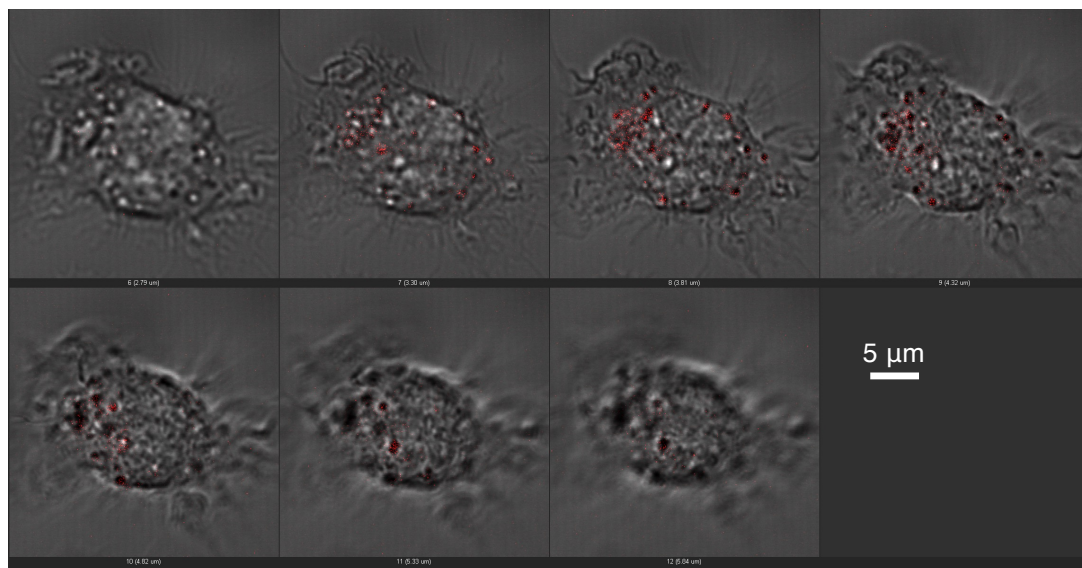
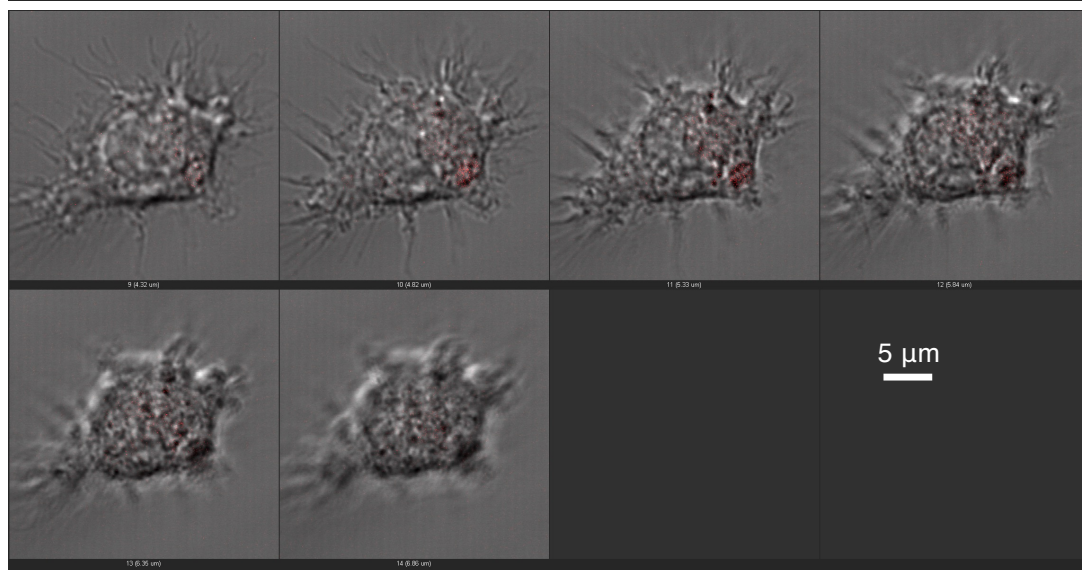


Fig. S10 mCherry signals in isolated microglia isolated from *Rtl9CmC* KI brain

Microglia were isolated from P0 neonatal brain and cultured for 10 days. The microglial cells were collected from the cultured Petri dishes (mixed glia cultures) by tapping (Lian et al., 2016; Irie et al., 2022).



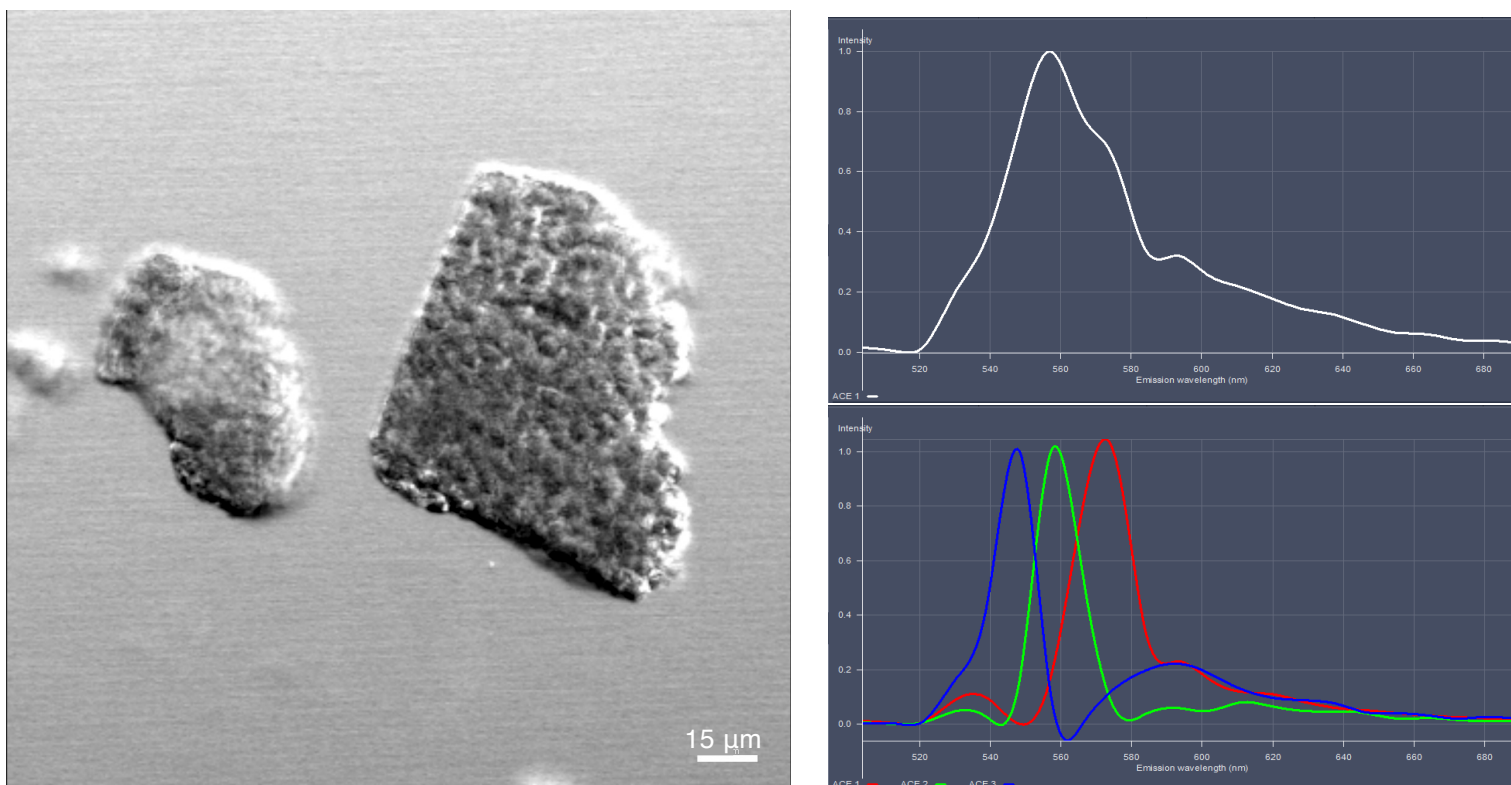


Fig. S11 The autofluorescence spectrum of Zymosan

Left: Transmission image. Right: Autofluorescence spectrum. Single (top) and/or the strongest signal (Maximum peak emission fluorescence wavelength: 570 nm) among 3 representative signals were used to detect zymosan.

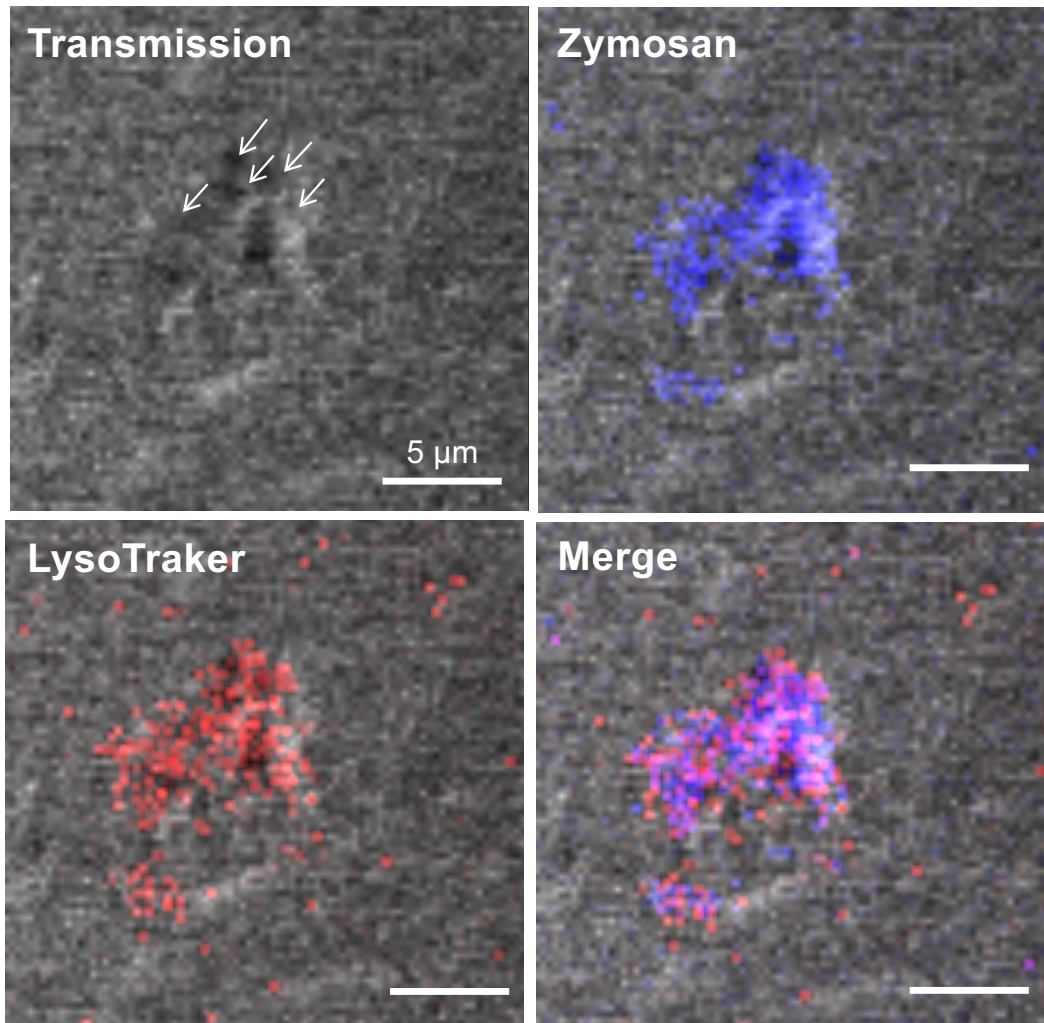


Fig. S12 Zymosan is visible in the transmission image

Zymosan incorporated in lysosomes in the *Rtl9* KO brain. Top, left: A transmission image. Arrows indicate zymosan aggregates. Top, right: zymosan autofluorescence (light blue). Bottom, left: The LysoTracker signal (red). Bottom, right: A merged image. Zymosan is distinguishable by its shape in the transmission image(s) in addition to its specific fluorescence signal. The same samples as in Fig. 3E were analyzed.

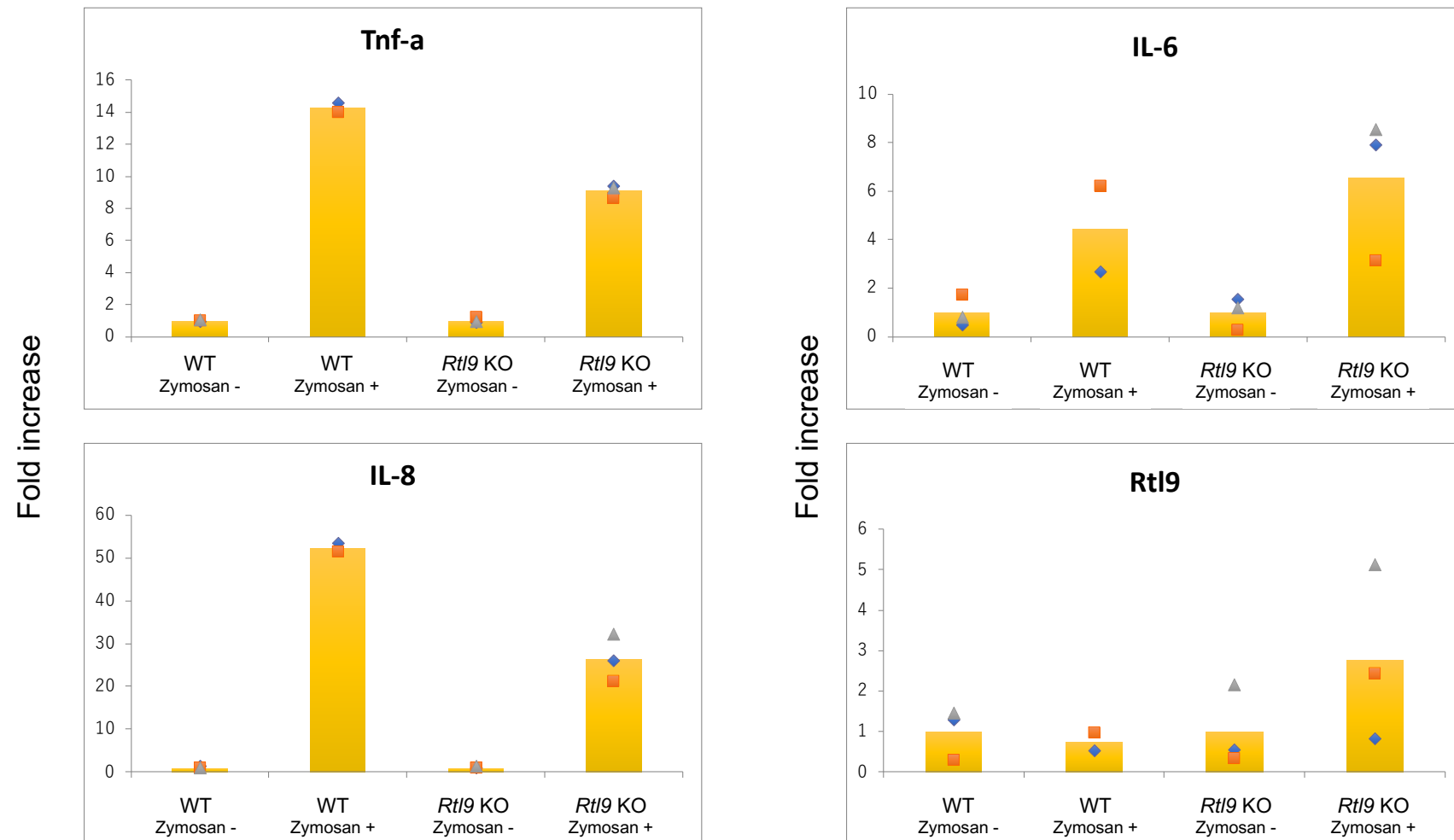


Fig. S13 Zymosan induces *Tnfa*, *Il6* and *Il8* expression in both WT and *Rtl9* KO microglia

Fig. S13 Zymosan induces *Tnfa*, *Il6* and *Il8* expression in both WT and *Rtl9* KO microglia

Tnfa, *Il6*, *Il8* and *Rtl9* mRNA were analyzed at 0 min (no administration) and 30 min after zymosan administration to WT and *Rtl9* KO microglia (n=3 each, shown in different colors and shapes). The expression levels are averaged with non-administration samples as 1. *Rtl6* primers were designed for the 3'-UTR, and were therefore also present in the *Rtl9* KO microglia. It should be noted that the *Rtl9* mRNA itself increases upon zymosan administration in *Rtl9* KO microglia, presumably due to feedback regulation of the RTL9 protein.

The PCR primers used were as follows:

Tnfa, 5'-GACAAGGCTGCCCCGACTACG-3' (forward) and 5'-CTTGGGGCAGGGGCTCTTGAC-3' (reverse); *Il6*, 5'-AGTTGCCTTCTTGGGACTGA-3'(forward) and 5'-CCTCCGACTTGTGAAGTGGT-3' (reverse); *Il8*: 5'-CTCTCAAGGGCGGTCAAAAAGTT-3' (forward) and 5'-TCAGACAGCGAGGCACATCAGGTA-3'(reverse); *Rtl9*: 5'-TCACCTACATGCCTGTGACC-3' (forward) and 5'-CAACAACACCACATTGTTACGG-3' (reverse).