Published in final edited form as:

Nature. 2021 December ; 600(7888): 324–328. doi:10.1038/s41586-021-04144-4.

FAM72A antagonizes UNG2 to promote mutagenic repair during antibody maturation

Yuqing Feng[1,14], **Conglei Li**[1,13,14], **Jessica A. Stewart**[2], **Philip Barbulescu**[1], **Noé Seija Desivo**[3,4], **Alejandro Álvarez-Quilón**[5,6], **Rossanna C. Pezo**[7,8], **Madusha L. W. Perera**[2], **Katherine Chan**[9], **Amy Hin Yan Tong**[9], **Rukshana Mohamad-Ramshan**[2], **Maribel Berru**[1], **Diana Nakib**[1], **Gavin Li**[1], **Gholam Ali Kardar**[10], **James R. Carlyle**[1], **Jason Moffat**[5,9,11], **Daniel Durocher**[5,6], **Javier M. Di Noia**[3,4], **Ashok S. Bhagwat**[2,12], **Alberto Martin**[1]

[1]Department of Immunology, University of Toronto, Toronto, Ontario, Canada.

[2]Department of Chemistry, Wayne State University, Detroit, MI, USA.

[3]Institut de recherches cliniques de Montréal, Montreal, Quebec, Canada.

[4]Molecular Biology Programs, Department of Medicine, University of Montreal, Montreal, Quebec, Canada.

[5]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada.

[6]Lunenfeld–Tanenbaum Research Institute, Sinai Health System, Toronto, Ontario, Canada.

[7]Sunnybrook Health Sciences Center, Toronto, Ontario, Canada.

[8]Department of Medicine, University of Toronto, Toronto, Ontario, Canada.

[9]Donnelly Centre, University of Toronto, Toronto, Ontario, Canada.

[10]Immunology, Asthma and Allergy Research Institute, Tehran University of Medical Sciences, Tehran, Iran.

[11]Institute for Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada.

[12]Department of Biochemistry, Microbiology nd Immunology, Wayne State University School of Medicine, Detroit, MI, USA.

Correspondence and requests for materials should be addressed to Alberto Martin. alberto.martin@utoronto.ca.

Author contributions Y.F. performed the CRISPR screen in collaboration with K.C., A.H.Y.T. and J.M., and characterized *Fam72a*$^{-/-}$ CH12 clones. C.L. characterized *Fam72a*$^{-/-}$ mice and performed drug inhibitor-treated CH12 assays. J.A.S. performed protein purification and in vitro biochemical assays. P.B. generated knockout clones in CH12 cells and performed repair substrate assays. N.S.D. characterized DT40 *Fam72a*$^{-/-}$ clones. A.A.-Q. conducted Bio-ID experiments. R.C.P. performed immunofluorescence analysis. M.L.W.P. performed uracil quantification experiments. M.B. helped with minipreps and maintain colonies for animal work. D.N. and G.A.K. helped generate knockout clones in cell culture. G.L. conducted western blotting in cell culture. Y.F., C.L., J.A.S., N.S.D. and A.A.-Q. performed experiments, analysed data and wrote the manuscript. P.B., R.C.P., M.L.W.P., K.C., A.H.Y.T., R.M.-R., M.B., D.N., G.L., G.A.K. and J.R.C. performed experiments. J.R.C., J.M., D.D., J.M.D.N., A.S.B. and A.M. analysed the data and wrote the manuscript. A.M. conceived, designed and supervised the study.

Competing interests The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-04144-4.

Peer review information *Nature* thanks Uttiya Basu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

[13]Present address: School of Medicine, The Chinese University of Hong Kong, Shenzhen, China.

[14]These authors contributed equally: Yuqing Feng, Conglei Li.

Abstract

Activation-induced cytidine deaminase (AID) catalyses the deamination of deoxycytidines to deoxyuracils within immunoglobulin genes to induce somatic hypermutation and class-switch recombination[1,2]. AID-generated deoxyuracils are recognized and processed by subverted base-excision and mismatch repair pathways that ensure a mutagenic outcome in B cells[3–6]. However, why these DNA repair pathways do not accurately repair AID-induced lesions remains unknown. Here, using a genome-wide CRISPR screen, we show that FAM72A is a major determinant for the error-prone processing of deoxyuracils. *Fam72a*-deficient CH12F3–2 B cells and primary B cells from *Fam72a*$^{-/-}$ mice exhibit reduced class-switch recombination and somatic hypermutation frequencies at immunoglobulin and *Bcl6* genes, and reduced genome-wide deoxyuracils. The somatic hypermutation spectrum in B cells from *Fam72a*$^{-/-}$ mice is opposite to that observed in mice deficient in uracil DNA glycosylase 2 (UNG2)[7], which suggests that UNG2 is hyperactive in FAM72A-deficient cells. Indeed, FAM72A binds to UNG2, resulting in reduced levels of UNG2 protein in the G1 phase of the cell cycle, coinciding with peak AID activity. FAM72A therefore causes U·G mispairs to persist into S phase, leading to error-prone processing by mismatch repair. By disabling the DNA repair pathways that normally efficiently remove deoxyuracils from DNA, FAM72A enables AID to exert its full effects on antibody maturation. This work has implications in cancer, as the overexpression of FAM72A that is observed in many cancers[8] could promote mutagenesis.

UNG2 normally functions in base-excision repair (BER) to replace deoxyuracil (dU) with deoxycytidine (dC) to prevent mutations. Paradoxically, UNG2 promotes transversion mutations at G:C base pairs of AID-induced dUs during somatic hypermutation (SMH)[7]. The mismatch repair (MMR) pathway is also co-opted during antibody diversification by producing mutations at A:T base pairs during the repair of the AID-induced U•G mispairs. Error-free repair of U•G mispairs occurs routinely in mammalian cells, but it remains unknown why B cells do not accurately repair AID-induced dUs. These observations suggest the existence of a factor or multiple factors in B lymphocytes that can subvert DNA repair pathways that normally repair dUs. Accordingly, we conducted a genome-wide pooled CRISPR screen in CH12F3–2 cells (hereafter referred to as CH12 cells), a mouse B cell line that switches from IgM to IgA after exposure to a cocktail of anti-CD40, IL-4 and TGF-β (CIT)[9]. Cas9-expressing CH12 clones (Extended Data Fig. 1a) were combined and transduced with the mouse Toronto KnockOut (mTKO) CRISPR library consisting of 94,528 Cas9 guide RNA (gRNA) sequences targeting 19,463 genes[10]. Transduced cells were selected using puromycin resistance, and then treated with CIT to stimulate switching from IgM to IgA. IgA-positive and IgA-negative cells were sorted by flow cytometry, and gRNA representation was compared to that of unsorted cells by sequencing genomic DNA from each of these populations (Extended Data Fig. 1b). The depletion of gRNAs from IgA-positive cells relative to the unsorted population (Fig. 1a) and the enrichment of gRNAs in the IgA-negative population (Extended Data Fig. 1c) identified several known genes with well-established functions in class-switch recombination (CSR), including *Aicda* (which

encodes AID), *Ung* and members of the MMR pathway (Supplementary Table 1). Among factors that we identified, we focused on the *Fam72a* gene. *Fam72a*, which was originally found to show increased expression in malignant cells, is largely uncharacterized but binds to UNG2[8], although no functional consequence has been established for this interaction.

To validate the role of FAM72A in CSR, we used three different gRNAs to knock out *Fam72a* in CH12 cells in bulk (Extended Data Fig. 1d) and found that all three gRNAs efficiently reduced IgA CSR (Fig. 1b). We also generated *Fam72a*-knockout CH12 clones (Extended Data Fig. 1f) and found that three independent $Fam72^{-/-}$ CH12 clones had markedly reduced *Fam72a* mRNA and IgA CSR compared with two $Fam72a^{+/+}$ wild-type (WT) CH12 clones (Fig. 1c, Extended Data Fig. 1e). CSR completion depends on non-homologous end joining and alternative end joining[2,11–13]. However, these DNA repair pathways were unaffected in FAM72A-deficient CH12 clones (Extended Data Fig. 1g). *Fam72a* mRNA levels were increased approximately 3-fold in lipopolysaccharide (LPS)-stimulated primary mouse B cells ex vivo, and approximately 20-fold in germinal centre B cells in vivo (Fig. 1d), suggesting that *Fam72a* has an important role at this stage of B cell development.

To further test the role of FAM72A in antibody diversification, we generated *Fam72a*-knockout mice (Extended Data Fig. 2). The B cell profiles in bone marrow and spleens were similar between $Fam72a^{-/-}$ mice and littermate controls (Extended Data Figs. 3, 4), indicating that FAM72A is dispensable for B cell development. To test the role of FAM72A in CSR in primary B cells, we isolated splenic B cells from $Fam72a^{-/-}$ and $Fam72^{+/+}$ mice and measured CSR ex vivo: switching to all immunoglobulin isotypes was defective in $Fam72a^{-/-}$ compared with $Fam72^{+/+}$ mice (Fig. 2a). $Fam72^{+/-}$ mice and $Fam72^{+/-}$ CH12 cells had a partial defect in CSR (Extended Data Fig. 5a, b) suggesting that FAM72A was haploinsufficient. To assess antigen-specific CSR, we immunized mice with nitro-phenyl-chicken gamma globulin (NP-CGG) and found that the generation of IgG^{+} nitro-phenyl-specific antibody-secreting cells was reduced in $Fam72a^{-/-}$ compared with $Fam72^{+/+}$ mice (Fig. 2b). $Fam72a^{-/-}$ mice did have a slightly higher number of germinal centre B cells compared with controls, however the germinal centre sizes and T follicular helper (T_{FH}) cell numbers were comparable between both groups (Extended Data Fig. 5c–e). Although there was a slight increase in apoptosis in $Fam72a^{-/-}$ B cells compared with controls, FAM72A deficiency did not affect the expression of *Aicda* and immunoglobulin germline transcripts, cell proliferation or cell cycle progression in B cells (Extended Data Fig. 5f–i). These data indicate that FAM72A is critical for CSR in primary mouse B cells.

To explore the role of FAM72A in SHM, we isolated germinal centre B cells from Peyer's patches of $Fam72a^{-/-}$, $Fam72^{+/+}$ and $Aicda^{-/-}$ mice, and sequenced the J_H4 region (Extended Data Fig. 6a). The proportion of germinal centre B cells in Peyer's patches was similar in $Fam72a^{-/-}$ and $Fam72^{+/+}$ mice (Extended Data Fig. 6b). However, in those germinal centre B cells, the loss of *Fam72a* caused an approximately fourfold reduction in mutation frequency in the J_H4 region, as well as in the *Bcl6* gene (Fig. 2c–e, Extended Data Fig. 6c–f). Moreover, the characteristics of the J_H4 mutations in the $Fam72a^{-/-}$ mice were different: B cells from $Fam72a^{-/-}$ mice had an increased proportion of transversion versus transition mutations at G:C basepairs (Fig. 2f and Extended Data Fig. 6c). This bias

in *Fam72a*$^{-/-}$ mice was the opposite of that in *Ung*$^{-/-}$ mice[7] (Fig. 2f), and more in line with UNG2 overexpression[14]. Collectively, these data suggest that UNG2 is hyperactive in the absence of FAM72A.

We also examined SHM in *Fam72a*$^{-/-}$ CH12 clones by sequencing the region 5′ of the μ-switch region. The mutation frequency was reduced in *Fam72*$^{-/-}$ CH12 clones compared with controls (Fig. 3a, Extended Data Fig. 7a, b). Elimination of *Fam72a* and *Ung* had opposite effects on genomic uracil levels; the CIT stimulation-dependent increase in uracils seen in WT cells was eliminated in *Fam72*$^{-/-}$ cells, whereas loss of *Ung* resulted in even greater increase in uracils upon CIT stimulation (Fig. 3b). These effects were not owing to reduced AID expression (Extended Data Fig. 7c) or reduced germline transcripts in *Fam72*$^{-/-}$ CH12 clones (Extended Data Figs. 5f, 7d). As the reduced mutation frequency could be due to altered dU processing in FAM72A-deficient cells, we tested whether FAM72A impacted AID-induced mutations in the *Ung*$^{-/-}$ background. We generated *Ung*$^{-/-}$ and *Ung*$^{-/-}$*Fam2a*$^{-/-}$ CH12 clones (Extended Data Fig. 1f). As expected, the mutation frequency 5′ of the μ-switch region was increased in the *Ung*$^{-/-}$ CH12 clone compared with WT controls, owing to the lack of repair of AID-induced dU (Fig. 3a). Of note, the loss of *Fam72a* in the *Ung*$^{-/-}$ background did not reduce the mutation frequency (Fig. 3a), suggesting that FAM72A affects SHM through UNG2.

To further test the notion that FAM72A functions with UNG2, we carried out CSR assays in cells doubly deficient in FAM72A and UNG, as well as those doubly deficient in FAM72A and MSH2 (Extended Data Fig. 1f). *Ung*$^{-/-}$ CH12 cells switched at about 3% the rate of WT CH12 cells (Fig. 3c). However, knocking out *Fam72a* in UNG-deficient CH12 cells did not further reduce CSR (Fig. 3c), suggesting that FAM72A is epistatic with UNG2 during CSR. *Msh2*$^{-/-}$ CH12 cells switched at 15% the rate of WT CH12 cells (Fig. 3d). Notably, knocking out either *Ung* or *Msh2* caused more than 85% reduction of CSR, suggesting that similar to SHM[15,16], UNG and MMR pathways work together to promote CSR. Knocking out *Fam72a* in MSH2-deficient CH12 cells did not lead to reduced CSR suggesting that FAM72A is also epistatic with MSH2 during CSR (Fig. 3d).

To confirm a potential epistatic interaction between FAM72A and MSH2 in a physiologically relevant system, we assessed CSR in B cells from *Msh2*$^{-/-}$ and *Fam72a*$^{-/-}$ mice, and mice deficient in both *Msh2* and *Fam72a*. *Msh2*$^{-/-}$, *Fam72a*$^{-/-}$ and *Fam72a*$^{-/-}$*Msh2*$^{-/-}$ B cells exhibited similar defects in CSR (Fig. 3f, Extended Data Fig. 7e) and SHM frequency (Fig. 3g). Although all mutations were reduced in *Msh2*$^{-/-}$ B cells, mutations at A:T base pairs were most affected (Extended Data Fig. 7f), as previously observed[5,6,17,18]. Mutation frequencies in *Fam72a*$^{-/-}$*Msh2*$^{-/-}$ B cells showed a similar decrease overall, but a large increase in dA/dT ratio compared with *Msh2*$^{-/-}$ B cells, probably caused by the increased UNG2 activity (Extended Data Fig. 7g, h), providing insights into how UNG2 promotes A:T mutations independent of MMR during SHM[6,19]. Hence, FAM72A is epistatic with both the MMR and BER during SHM and CSR.

We also tested whether FAM72A affects immunoglobulin gene conversion, which is also initiated by AID[20]. *Fam72a*$^{-/-}$ DT40 clones (Extended Data Fig. 8a) showed an approximately twofold increase in gene conversion (Fig. 3e). Although proliferation was

reduced in *Fam72a*$^{-/-}$ DT40 clones (Extended Data Fig. 8b), analysis of the data using the same number of cellular divisions showed increased gene conversion in *Fam72a*$^{-/-}$ DT40 cells (Extended Data Fig. 8c). Since gene conversion is exclusively mediated by UNG2[21] and not by MMR[22], this result is consistent with FAM72A antagonizing UNG2, since the only sources of DNA breaks would derive from UNG2, and increased UNG2 activity would lead to an increase in the breaks necessary for gene conversion. Collectively, these data show that FAM72A modulates three AID-mediated antibody diversification processes.

A previous report suggested that FAM72A interacts with UNG2[8]. To confirm this interaction, we performed a BioID proximity-labelling assay and found that human FAM72A is proximal to UNG in HEK 293 cells (Extended Data Fig. 9a). To test for a direct interaction, we mixed purified versions of these proteins (Extended Data Fig. 9b) at a 1:1 ratio, and the pull-down of the proteins using either nickel-containing beads or anti-Flag antibodies resulted in co-precipitation of the other protein (Extended Data Fig. 9c).

To test whether FAM72A modulates UNG2 activity, we carried out in vitro uracil excision assays. The addition of mouse FAM72A resulted in a slight reduction in the ability of mouse UNG2 and human UNG2 to excise uracils from three different types of substrates, but had no effect on the activity on *Escherichia coli* UNG (Extended Data Fig. 9d–f). Although FAM72A deficiency had no effect on *Ung2* transcript levels in CH12 cells and mouse B cells (Extended Data Fig. 9g), it did influence UNG2 protein levels, as *Fam72a*$^{-/-}$ CH12 cells had increased UNG2 but not UNG1 protein (Fig. 4a, Extended Data Fig. 9h). This result was confirmed in primary B cells from *Fam72a*$^{-/-}$ mice (Fig. 4b) and in *Fam72a*$^{-/-}$ DT40 cells (Extended Data Fig. 8d).

As *Fam72a* deletion increases UNG2 protein levels and activity, we assessed whether overexpression of UNG2 in WT CH12 cells decreases CSR and SHM. Indeed, *Ung2* overexpression in WT CH12 cells reduced CSR and SHM in the μ-switch region (Fig. 4c). Overexpression of another uracil DNA glycosylase, SMUG1, has also been shown to reduce SHM in mice[23]. FAM72A expression in WT CH12 cells had only a small impact on SHM, CSR, and UNG2 protein levels (Fig. 4c, Extended Data Fig. 9i). However, expression of FAM72A in *Fam72a*$^{-/-}$ CH12 cells restored CSR and reduced UNG2 levels, whereas expression of FAM72A(W125R), which does not bind to UNG2[8], did not (Fig. 4d). This result indicates that FAM72A exerts its effects on UNG2 post-translationally. We further confirmed that UNG2 is degraded via the proteasome (Extended Data Fig. 9j).

To test whether the expression of *Fam72a* is cell cycle-dependent, we synchronized CH12 cells in G_2/M using the CDK1–cyclin B1 inhibitor RO-3306, and then released the blockade. *Ung* and *Geminin* levels peaked at 7 h after the inhibitor was removed (Fig. 4e), suggesting entry into S phase. However, *Fam72a* transcripts peaked 3 h after the inhibitor was removed (Fig. 4e) suggesting that *Fam72a* level is increased and might affect UNG2 levels in G1. Indeed, UNG2 protein was markedly increased in the G1 phase of the cell cycle in *Fam72a*$^{-/-}$ CH12 cells compared with controls (Fig. 4f, Extended Data Fig. 9k), a stage of the cell cycle that coincides with AID-induced deaminations[24]. These data show that FAM72A handicaps UNG2 in cells, enabling the accumulation of U·G pairs following AID-mediated deamination of dC. While these excess dUs cause SHM through replication

and error-prone BER, the U·G mispairs are also the trigger for aberrant MMR (Extended Data Fig. 10). Thus, FAM72A has a key role in creating the raw material for both SHM and CSR and connects the BER and MMR processes.

The key limitation of using AID-induced dUs for SHM and CSR is that mammalian UNG2 is abundant in the nucleus[25] and has a high turnover rate[26]. A high uracil-excision activity is required because replicative DNA polymerases readily misincorporate dUTP into DNA. Consequently, any dUs created by AID are likely to be quickly excised by UNG2. FAM72A alleviates this problem largely by reducing UNG2 protein levels. This work further suggests that the cell cycle stage at which UNG2 excises uracils is a key determinant of accurate versus mutagenic repair of dUs. Mammalian cells generally express higher level of UNG2 proteins than germinal centre B cells, leading to efficient excision of dU by UNG2 and accurate BER[27]. However, in germinal centre B cells, expression of FAM72A is increased 20-fold (Fig. 1e), largely in the G1 phase (Fig. 4e), which leads to reduced UNG2 protein in the G1 phase (Fig. 4f), precisely when AID produces dUs. As a result, U·G mispairs persist into S phase, where these mispairs are replicated to generate transition mutations at C:G base pairs. As UNG2 is re-expressed in early S phase (Fig. 4f), uracil excision followed by replication produces transition and transversion mutations at C:G base pairs. In addition, the persistence of U·G mispairs into S phase enable detection and processing by the MMR pathway. However, the expression of UNG2 in the S-phase interferes with canonical MMR, leading to the excision of dUs in the MMR-repair track[15,16], generation of abasic sites, and recruitment of polymerase η which introduces mutations at A:T base pairs[28]. When FAM72A is absent in germinal centre B cells, UNG2 protein and activity is unleashed in the G1 phase (Fig. 4f), leading to efficient removal of dU and subsequent accurate BER. Thus, we propose that FAM72A evolved to achieve the optimal level of UNG2 in B cells: low enough to preserve U·G mismatches, but high enough in the S phase of the cell cycle to disrupt canonical MMR. Thus, FAM72A promotes the full hypermutation spectrum, as observed in our study.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-04144-4.

Methods

Mice

Fam72a mutant sperm: C57BL/6N-Fam72a$^{tm1.1(KOMP)Vlcg}$/MbpMmucd mouse sperm, in which a 10,298-bp region between the 131,528,704 and 131,539,001 positions of mouse chromosome 1 was deleted by the insertion of a ZEN-Ub1 cassette, was purchased from Mutant Mouse Resource and Research Centers (UC Davis). The in vitro fertilization of C57BL/6N female mice (Charles River) with *Fam72a* mutant sperm was performed at The Centre for Phenogenomics (Toronto, Canada). *Fam72a* heterozygous mice were bred in our animal facility and maintained under pathogen-free conditions. The *Msh2*-deficient mice

were provided by T. Mak and have been bred in our facility as we previously described[29]. *Fam72*$^{+/-}$*Msh2*$^{+/-}$ heterozygous mice were bred to generate WT, *Fam72a*$^{-/-}$, *Msh2*$^{-/-}$, and *Fam72*$^{-/-}$*Msh2*$^{-/-}$ littermate mice. The experimental procedures were approved by the Animal Care Committee of University of Toronto (Division of Comparative Medicine/CCBR; protocol number 20011472). Mice were housed with humidity 45% on a 14-h light, 10-h dark cycle, with an ambient temperature of approximately 22 °C.

Plasmids

For Cas9-mediated gene editing, gRNAs were cloned into px330 (Addgene, plasmid #42230) or lentiGuide-Puro (Addgene, plasmid #52963). Mouse *Fam72a* was amplified by PCR from CH12 cDNA and cloned into a pRSET-A vector (Thermo Fisher) or the pMX-PIE retroviral vector (gift from B. R. San-Martin). Site-directed mutagenesis was used to construct the FAM72A(W125R) mutant. Mouse *Ung2* was amplified from *Ung2* cDNA (OriGene) and cloned into Flag-HA-pcDNA3.1 vector (Addgene, plasmid #52535) using Gibson assembly kit (New England Biolabs) to create pcDNA3.1-Ung2-polyGly-Flag (Ung2-Flag) expression construct or the pMX-PIE retroviral vector. The Ung2-Flag region was amplified and cloned into a pET28a^{+} expression vector. The coding sequence for human FAM72A (*FAM72A*) was ordered from IDT with gateway-compatible flanking sequences and cloned into pDEST-NLS3×Flag-miniTURBO (a gift from A.-C. Gingras) via gateway cloning. The homologous recombination substrate (DR.GFP) was a gift from M. Jason, non-homologous end joining substrate (EJ5-GFP) was a gift from T. Kohno, and the alternative end joining substrate (EJ2-GFP) was a gift from J. Stark. All the plasmid constructs were verified by Sanger sequencing. Primers were obtained from Invitrogen and are listed in Supplementary Table 2.

Lentiviral and retroviral transduction

To generate lentiviruses, HEK 293T cells were seeded in DMEM with 10% FBS and penicillin-streptomycin, and 18 h later, were transfected with lentiviral packing plasmids pMD2.G (Addgene, plasmid #12259) and psPAX (Addgene, plasmid #12260) along with lentiviral plasmids containing gRNAs of interest in the presence of PEI. The culture supernatants containing retroviruses were obtained by transfecting BOSC23 cells (gift from G. Ehrhardt). WT and *Fam72a*$^{-/-}$ CH12 cells were transduced with an empty retrovirus (pMX-PIE) or retroviruses expressing FAM72A (WT or W125R) or UNG2 cDNA. The lentiviral and retroviral transduction of CH12 cells were performed by centrifugation at 800*g* for 1 h at room temperature in the presence of 8 μg ml^{-1} polybrene.

CRISPR–Cas9 screen in the CH12F3–2 cell line

CH12 cells were cultured as previously described and were mycoplasma free[30]. To generate Cas9-expressing CH12 cells, CH12 cells were transduced with LentiCas9-Blast (Addgene, plasmid #52962) and selected in the presence of 10 μg ml^{-1} Blastcidin for 9 days. Single clones that stably expressed high levels of Cas9 were collected to generate the pooled Cas9-expressing CH12 cells. The timeline that allowed maximal Cas9 editing efficiency in CH12 cells was determined by monitoring CSR defects caused by gRNAs that target AID. Seventy-one million IgA^{-} CH12 cells stably expressing Cas9 were transduced with the lentiviral-based mTKO library[10] consisting of 94,528 gRNA sequences (pLCKO-mTKO) at

a multiplicity of infection (MOI) of 0.6, followed by puromycin selection for 2 days. The pLCKO-mTKO CRISPR library is available at Addgene (pooled libraries #159393). CH12 cells were subsequently stimulated with recombinant TGFβ (R&D systems), recombinant IL-4 (R&D systems), and anti-CD40 (ThermoFisher) for 3 days, as previously described[30]. The cells were stained for surface expression of IgA using PE-conjugated anti-mouse IgA antibody (SouthernBiotech) and sorted with a FACS Aria IIIu and influx sorter (BD Biosciences). The CRISPR sequencing library was prepared as described[31]. Genomic DNA from T0 and T9 (IgA^+, IgA^- and unsorted) samples was extracted using the Wizard Genomic DNA Purification kit (Promega) according to the manufacturer's instructions. To generate sequencing libraries, gRNA barcode sequences were amplified as described previously[10]. The resulting libraries were sequenced on an Illumina HiSeq2500.

CRISPR–Cas9-mediated gene editing in CH12 and DT40 cells

To knock out genes of interest in CH12 cells, the px330 vector containing gRNA against *Fam72a, Ung* or *Msh2* was electroporated into CH12 cells. At 72h post-electroporation, CH12 cells were subcloned by limiting dilution. Three *Fam72a*$^{-/-}$ clones were generated (12A, 2B and 6H) and for the purpose of simplicity, were renamed clones 1, 2 and 3, respectively, while two *Fam72a*$^{+/+}$ clones (20C and 23C, renamed as clones 1 and 2) were used as control. Individual knockout clones were sequenced and validated using quantitative PCR and/or Western blot. DT40 CL18 *Fam72a*$^{-/-}$ cells were obtained by transfecting DT40 CL18 with Nucleofector Kit V four constructs in PX458 (Addgene, plasmid #48138) targeting *Fam72a* (1.5 μg each) and two pBluescript KS(+) constructs containing an antibiotic cassette resistance each (either for blasticidin or hygromycin) flanked by *Fam72a*-homology arms (2 μg each). At 48 h after electroporation, Blasticidin (25 μg ml^{-1}) and hygromycin (2 mg ml^{-1}) were added and cells were plated in 96-well plates (6,000 cells in 20 μl per well). After 7–14 days, cells from wells containing a single colony were expanded. Genomic DNA was isolated and knockout clones were identified by PCR and verified by RNA extraction and PCR with reverse transcription (Extended Data Fig. 8a). All gRNA sequences are listed in Supplementary Table 2.

Measuring expression of *Fam72a* and *Ung* mRNA at different stages of the cell cycle

CH12 cells were synchronized at G2/M phase, as previously described[32]. In brief, *Fam72a*$^{+/+}$ CH12 cells (20×10^4 cells per ml) were cultured for 20 h with 10 μM CDK1/cyclin B1 inhibitor RO-3306 (Sigma-Aldrich), followed by exchange into inhibitor-free culture media. Cells were collected at 0 h (G2/M), 1 h (G2/M), 3 h (G1), 4 h, 7 h (S) and 10 h after RO-3306 release, followed by RNA preparation and quantitative PCR analysis of *Fam72a* and *Ung* mRNA. The corresponding cell cycle stages of CH12 cells at these collection timepoints were previously confirmed by[32] in CH12 cells using DNA staining with propidium iodide (detected by flow cytometry), which was validated by quantitative PCR analysis of *Geminin* mRNA, which accumulates in S phase.

Cycloheximide and MG132 treatment to CH12 cells

To inhibit protein synthesis, CH12 cells were treated for 2–12 h with 10 μg ml^{-1} cycloheximide (Bioshop, catalogue (cat.) no. cyc003). To inhibit proteasome-mediated protein degradation, CH12 cells were treated for 2–12 h with 10 μg ml^{-1} cycloheximide

plus 10 μM MG132 (Sigma-Aldrich, cat. no. 474790). Cells were collected at 2 h, 4 h, 6h, 10 h and 12 h after treatment, followed by UNG2 Western blotting analysis.

Monitoring growth and immunoglobulin gene conversion in DT40 cells

To determine DT40 cells doubling time, cultures of 5×10^4 cells pr ml per condition were set and counted in duplicate by haemocytometer every 6 h for 4 days. Data were analysed with GraphPad Prism and doubling time was determined by nonlinear regression of data to Exponential (Malthusian) growth equation (Extended Data Fig. 8b). Immunoglobulin gene conversion was monitored by the conversion of IgM^- cells to IgM^+ cells by flow cytometry. Fluctuation analysis was performed as described[22] with modifications. In brief, 4 single-cell clones of DT40 WT or Fam72a$^{-/-}$ were stained with anti-IgM-PE (1:40, clone M-1, SouthernBiotech) and 12 populations of 5×10^4 IgM^- cells were sorted for each clone. The proportion of IgM-gain was measured by flow cytometry using the same staining after two weeks of expansion (same time of expansion) and five days later for the Fam72a$^{-/-}$ cells (to achieve the same number of divisions as WT at two weeks). Data were analysed with FlowJo and GraphPad Prism.

Human FAM72A proximity-dependent BioID

HEK 293T cells were seeded in 10-cm dishes, and 24 h later cells were transfected by standard PEI method with pDEST-NLS3×Flag-miniTURBO (1.5 μg) or pDEST-NLS-3×Flag-miniTURBO-FAM72A (3 μg). At 48 h post-transfection, 5 μg ml^{-1} doxycycline was added and 24 h later, cells were exposed to biotin (50 μM) for 1 h. Cells were washed with PBS, scraped off and pelleted. Cell pellet was lysed in 0.5 ml lysis buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM EDTA, 0.5%, NP-40 and 10 mM NaF, 10 mM $MgCl_2$) and 250 U benzonase per sample on ice for 15 min. SDS was added to a final concentration of 0.4% and incubated for 5 min at 4 °C. Lysates were centrifuged at 15,000*g* for 5 min and supernatants were incubated with 50 μl Streptavidin Sepharose High Performance beads (GE Healthcare) for 1 h with rotation. Beads were washed 5 times with lysis buffer and bound proteins were eluted by boiling for 5 min at 96 °C in SDS–PAGE sample buffer and analysed by immunoblotting.

Primary B cell analysis

B cell development in bone marrow and spleen of sex-matched *Fam72a*$^{-/-}$ and *Fam72a*$^{+/+}$ littermates (6–8 weeks old) was evaluated (Supplementary Fig. 2) as described[33]. To induce ex vivo CSR to different immunoglobulin isotypes, splenic B cells were purified with a mouse B cell isolation kit (StemCell Technologies), and then stimulated with LPS (Sigma-Aldrich) in combination with various cytokines, as previously described[33]. Cells were collected at day 4 after stimulation and were stained with antibodies against mouse IgG1 (PE, BD Pharmigen, cat. no. 550083; 1: 150 dilution), IgG2b (PE, SouthernBiotech, cat. no. 1090–09S; 1:150), IgG3 (FITC, BD Pharmigen, cat. no. 553403; 1:100), IgE (FITC, BD Pharmigen, cat. no. 553415; 1:100) or IgA (PE, SouthernBiotech, cat. no. 1040–09; 1:150) to assess CSR. For the cell cycle analysis, splenic B cells were stimulated with LPS for 2.5 days and analysed with the Click-iT Plus EdU Alexa Fluor 647 kit (ThermoFisher Scientific) as described[33]. IgG1 CSR was induced in carboxyfluorescin diacetate succinimidyl ester (CFSE)-pulsed splenic B cells as we previously described[33].

The apoptosis of LPS stimulated splenic B cells was assayed using an APC-conjugated Annexin V Apoptosis Detection Kit (eBioscience). The flow cytometric data were analysed with a FlowJo X10 software.

Analysis of Sμ, J_H4 region and *Bcl-6* mutations

For Sμ-region sequencing in CH12 cells, WT, *Fam72a*$^{-/-}$, *Ung*$^{-/-}$ and *Fam72a*$^{-/-}$ *Ung*$^{-/-}$ CH12 cells, and WT CH12 cells retrovirally transduced with UNG2 or FAM72A, were cultured with CIT for 5 days. Corresponding genomic sequences from resting WT CH12 cells was used as the reference sequence. For J_H4 region and *Bcl-6* gene mutation analysis, Peyer's patches were collected from approximately 5-month-old mice of indicated genotypes. Single-cell suspensions of Peyer's patches were prepared by grinding, and then filtered through a 70-μm cell strainer. As previously described[34], cells were stained with antibodies against mouse B220 (PE, clone: RA3–6B2, eBiosciences, 1:300), Fas (biotin, clone: Jo2, BD Pharmigen; 1:100) and GL-7 (eF660, clone: GL-7, eBiosciences; 1:200) in the presence of Fc block (2.4G2), followed by streptavidin-APC-eFluor 780 staining (eBiosciences, cat. no.: 47–4317-82; 1:200), to identify germinal centre B cells. The cells were then stained with 7-AAD (live/dead marker) prior to flow cytometric sorting on an FACS Aria II sorter (BD Biosciences). The sorted germinal centre B cells (7-AAD$^-$B220$^+$GL-7$^+$Fas$^+$) were digested with Proteinase K (Invitrogen) at 55 °C overnight, followed by phenol–chloroform genomic DNA purification. The J_H4 intron region was amplified by PCR using a high-fidelity Platinum SuperFi II DNA polymerase (Invitrogen) with primers listed in Supplementary Table 2. Using the primer sequences and PCR protocol previously described[35], *Bcl6* was amplified from germinal centre B cells of Peyer's patches with SuperFi II DNA polymerase, and *Bcl6* PCR product (~1.2 kb) was purified using a PureLink PCR purification kit (ThermoFisher Scientific). Gel-extracted J_H4 or purified *Bcl6* PCR products were cloned and sequenced at cloned into Zero Blunt vectors (Invitrogen) or Clone JET (Invitrogen) and then transformed into competent DH5α bacteria. The plasmids yielded from individual bacterial clones were sequenced at The Centre for Applied Genomics (Toronto, ON). Sequences were analysed with DNA-STAR software and were aligned with the corresponding genomic sequences as described[17].

Quantitative PCR

RNA extraction was performed using TRIzol (Invitrogen), followed by DNase I treatment (Invitrogen), and reverse transcription with Maxima First strand cDNA synthesis Kit for quantitative PCR (Thermo Fisher) to prepare cDNA. The cDNA samples were analysed by quantitative PCR using PowerUp SYBR Green Master Mix (Applied Biosystems) according to the manufacturers protocol. Primer sequences are listed in Supplementary Table 2.

Purification of mouse UNG2 and FAM72A proteins

The plasmid pRSET_6×His-Fam72a was transformed into BL21(DE3) and mouse FAM72A expression was induced using 0.1 mM IPTG and the cells were grown at 18 °C for additional 18 h. The cells were collected by centrifugation and resuspended in 30 ml lysis buffer (20 mM Tris-Cl pH 8.0, 50 mM NaCl, 10 mM imidazole) supplemented with cOmplete protease inhibitor cocktail. Cells were lysed using a French press and the lysate was clarified by centrifugation. The supernatant was mixed with the Ni-NTA agarose beads

and rotated at 4 °C for 1 h. The beads were successively washed with 20 ml of buffer A (20 mM Tris-Cl pH 8, 50 mM imidazole) containing 50, 250 or 500 mM NaCl. The mouse FAM72A protein was eluted from the beads using 500 µl aliquots of elution buffer (20 mM Tris-Cl pH 8.0, 50 mM NaCl, 250 mM imidazole). The elution fractions found to have the protein when analysed by SDS– PAGE were combined and dialysed using buffer B (20 mM Tris-Cl pH 8.0, 50 mM NaCl, 1 mM DTT, 1 mM EDTA, 10% glycerol). The protein was concentrated to 1 mg ml^{-1} using a 10 kDa MWCO centrifugal filter and stored at −80 °C.

BL21(DE3) cells containing plasmid pET28a-Ung2–6×Gly-Flag were grown, and protein expression was induced using IPTG (0.5 mM) and the cells were broken in the same way as cells expressing FAM72A. Following cell lysis, the cleared lysate was mixed with the anti-Flag M2 magnetic beads (Sigma-Aldrich) and rotated at 4 °C for 18 h. The beads were collected using a DynaMag-2 magnet (Thermofisher) and successively washed with 1×TBS-T containing 50, 250 or 500 mM NaCl. The mouse UNG2 protein was eluted with 500 µl Flag Elution Buffer (3× Flag peptide at 150 ng $µl^{-1}$ in 1× TBS). The elution fractions found to have the protein when analysed by SDS–PAGE were combined, the buffer was exchanged 3 times with UNG2 storage buffer (25 mM HEPES– NaOH (pH 7.4), 200 mM NaCl, 0.01% Triton X-100 and 1 mM TCEP, 20% glycerol) and the protein was concentrated to 5 mg ml^{-1} using a 10 kDa MWCO centrifugal filter and stored at −80 °C. The human UNG2 protein was a gift from B. Weiser (Rowan University School of Osteopathic Medicine).

Co-precipitation of mouse FAM72A and UNG2 proteins

The UNG2 proteins were mixed with mouse FAM72A at a molar ratio of 1:1 in UNG2 buffer (10 mM Tris–HCl (pH 8.0), 100 mM NaCl, 1 mM DTT and 5 mM EDTA) and incubated at 25 °C for 20 min. The proteins were co-precipitated by pulling down UNG2–Flag or 6×His–FAM72A using anti-Flag magnetic beads or Ni-NTA agarose beads, respectively. After an overnight incubation at 4 °C while rotating, the beads were washed 3 times with 1× TBS-T and eluted by the addition of Laemmli buffer and boiling the mixture for 10 min at 95 °C. The input and elution samples were analysed by western blot by probing with mouse anti-Flag M2 antibody (Sigma-Aldrich; 1:2,000), rabbit anti-His tag antibody (Cell Signaling; 1:2,000), and anti-UDG antibody (Santa Cruz; 1:500) followed by the addition of goat anti-rabbit IgG and goat anti-mouse IgG HRP-conjugated antibodies (Cell Signaling; 1:1,000). The protein bands were visualized by the addition of Super signal West Pico Plus chemiluminescence substrate (ThermoFisher) and detected using a FluorChemQ scanner (Cell Biosciences).

In vitro uracil excision assay

DNA oligomers containing uracil used in these assays are listed in the Supplementary Table 2. The mouse UNG2 protein alone or mixed with mouse FAM72A was preincubated at 25 °C for 20 min in UNG2 buffer. This was followed by incubation at 25 °C for 5 min and the reactions were stopped by the addition of NaOH to 0.1 M and boiling at 95 °C for 10 min. Formamide dye was added to 50% v/v and the reaction products were separated on a 15% denaturing polyacrylamide gel. The DNA was visualized by scanning for Cy2 fluorescence using a Typhoon FLA 9500 phosphor imager and band intensities were quantified using ImageJ software. The percent of product was normalized to the negative control sample

(UNG2 without FAM72A) and bar graphs were generated using GraphPad Prism 8 and the statistical significance was determined using Mann–Whitney *U*-test with a 95% confidence interval.

Quantification of genomic uracils

The genomic uracils were analysed in CH12 cells as described[36]. In brief, about 5 μg of the genomic DNA was digested with HaeIII (New England Biolabs) and purified using phenol:chloroform extraction followed by ethanol precipitation. The digested DNA was incubated with AA7 (Sigma-Aldrich, 10 mM final concentration) at 37 °C for 1 h to block pre-existing abasic sites. DNA was then treated with *E. coli* uracil DNA-glycosylase at 37 °C for 30 min followed by 1-h incubation with 2 mM AA6. AA6 tagged DNA was labelled with DBCO-Cy5 (1.7 μM) by shaking the reaction mixture for 2 h at 37 °C in dark. The tagged DNA was purified using DNA Clean and Concentrator Kit (Zymo Research). Fluorescently labelled DNA was transferred on to Zeta probe membrane (Bio Rad) using a dot blot apparatus (Bio-Rad) and the membrane was scanned using a Typhoon 9210 phosphor imager (GE Healthcare). The images were analysed using the ImageJ software. Fluorescence intensity was converted to number of uracils using the CJ236 standard curve as described previously[36].

Western blotting

Western blotting was performed as described[37]. The following antibodies were used: anti-UNG (abcam; cat. no. 245630), anti-AID (Invitrogen; cat. no. 39–2500), anti-MSH2 (Invitrogen; cat. no. 33–7900) or β-actin (Sigma-Aldrich; cat. no. A2066). Densitometry analysis was performed using ImageJ (NIH).

Germinal centre analysis

Fam72a$^{-/-}$ or *Fam72a*$^{+/+}$ littermates (around 8–12 weeks old) were injected intraperitoneally with 100 μg NP_{29}-CGG plus Alum adjuvant, as previously described[33]. At day 8 after immunization, single-cell suspensions of spleen were prepared by grinding and then filtered through a 70-μm cell strainer. As described[34], cells were stained with antibodies against mouse CD45.2 (APC-eF780, clone: 104, eBiosciences; 1:100), CD19 (BV605, clone: 6D5, BioLegend; 1:200), PE-Cy7 dump antibodies (CD4, clone: RM4–5; CD8, clone: 53–6.7; CD11c, clone: N418; F4/80, clone: BM8, all from eBiosciences; 1:200), IgD (eF450, clone: 11–26C, eBiosciences; 1:100), Fas (biotin, clone: Jo2, BD Pharmigen) and GL7 (eF660, clone: GL7, eBiosciences) in the presence of Fc block, and then stained with streptavidin–FITC (eBiosciences, cat. no. 11–4317-87; 1:200), to identify germinal centre B cells (live $CD45^+CD4^-CD11c^-F4/80^-CD19^+IgD^-GL\text{-}7^+Fas^+$) (Supplementary Fig. 2); to identify T_{FH} cells (live $CD4^+B220^-CXCR5^+PD\text{-}1^+$intracellular $Bcl\text{-}6^+$) in spleen, cells were labelled with Live/Dead Aqua (Life Technologies) and then stained with antibodies against mouse B220 (BV605, clone: RA3–6B2, BioLegend; 1:200), CD4 (PE-Cy7, clone: RM4–5, eBiosciences; 1:200), PD-1 (APC, clone: J43, eBiosciences; 1:100) and CXCR5 (biotin, clone: SPRCL5, eBiosciences; 1:50) in the presence of Fc block, and then stained with streptavidin-APC-eFluor 780 (1:100) (Supplementary Fig. 2). Cells were subsequently permeabilized and fixed using perm/fix buffer from eBiosciences, and intracellular staining with anti- Bcl-6 (PE, clone: K112–91, BD Pharmigen; 1:20) was performed as previously

described[34]. The stained cells were analysed by flow cytometric assays and data were collected using FACS DIVA software.

For the histological analysis of germinal centres, PE-anti-IgD and FITC-PNA were used to stain 7 μM OCT-embedded splenic sections, as described[38]. In brief, frozen splenic sections from *Fam72a*$^{-/-}$ and *Fam72a*$^{+/+}$ mice were fixed in pre-cooled acetone and stained with FITC-PNA (Vector Laboratories cat. no. 405705; 1:100), PE-anti-IgD (BioLegend cat. no. 405705; 1:200) and DAPI. Fluorescence microscopy was performed using a Zeiss AX10 microscope, EC Plan-NeoFluar 10×/0.3 Ph1 objective and Axiocam 506 mono camera using filters for DAPI, GFP and Ds-red. Images were analysed using Zeiss image processing software ZEN2 blue edition. Random fields containing germinal centres were imaged and the diameters of germinal centres contained within each field (in microns) were measured and recorded.

Statistical analysis

All analyses were performed in GraphPad Prism software. Otherwise noted, two-tailed unpaired Student's *t*-test was used.

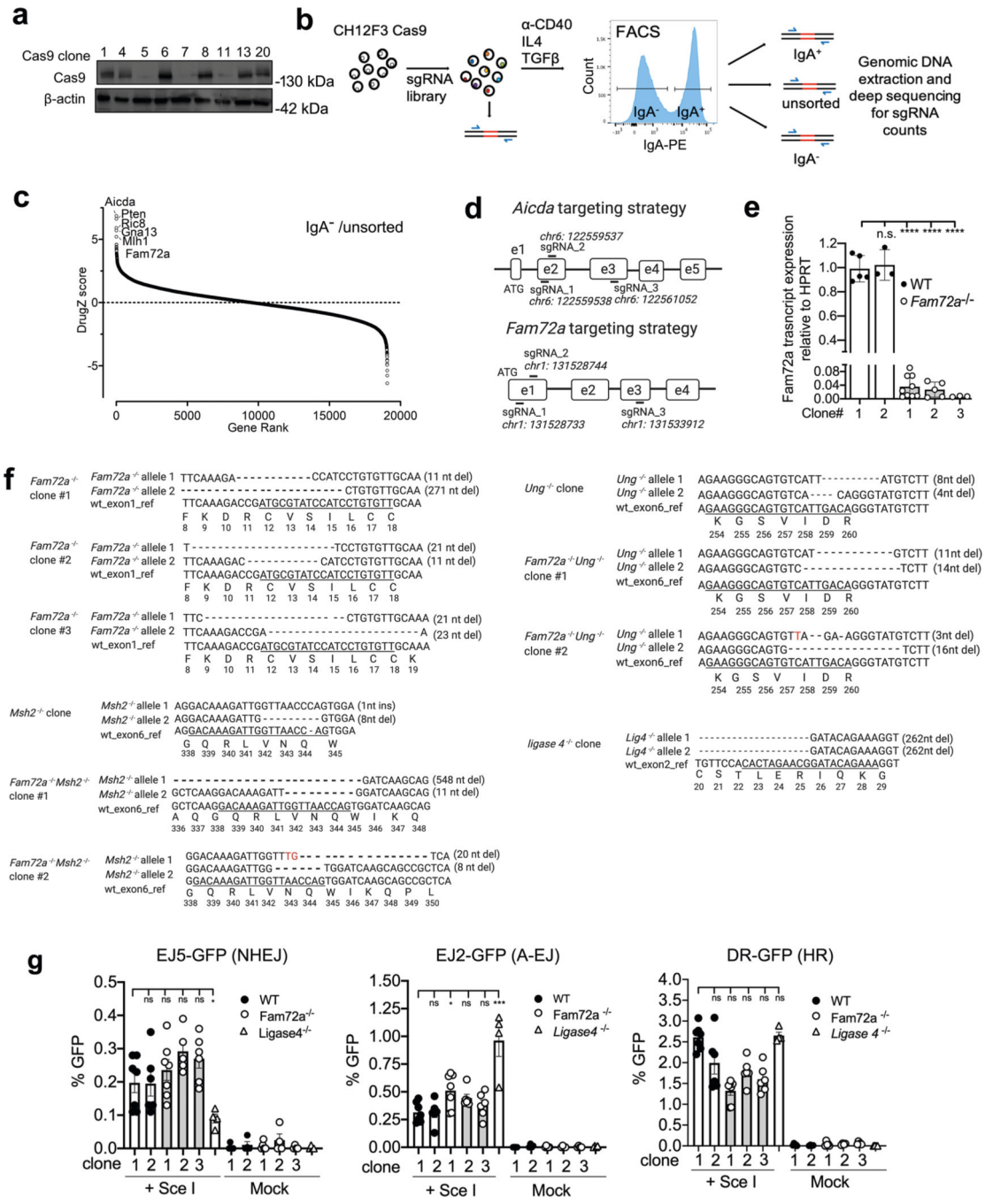
Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The raw sequence reads from CRISPR screening have been deposited in the NCBI Gene Expression Omnibus under accession number GSE183706. Source data are provided with this paper.

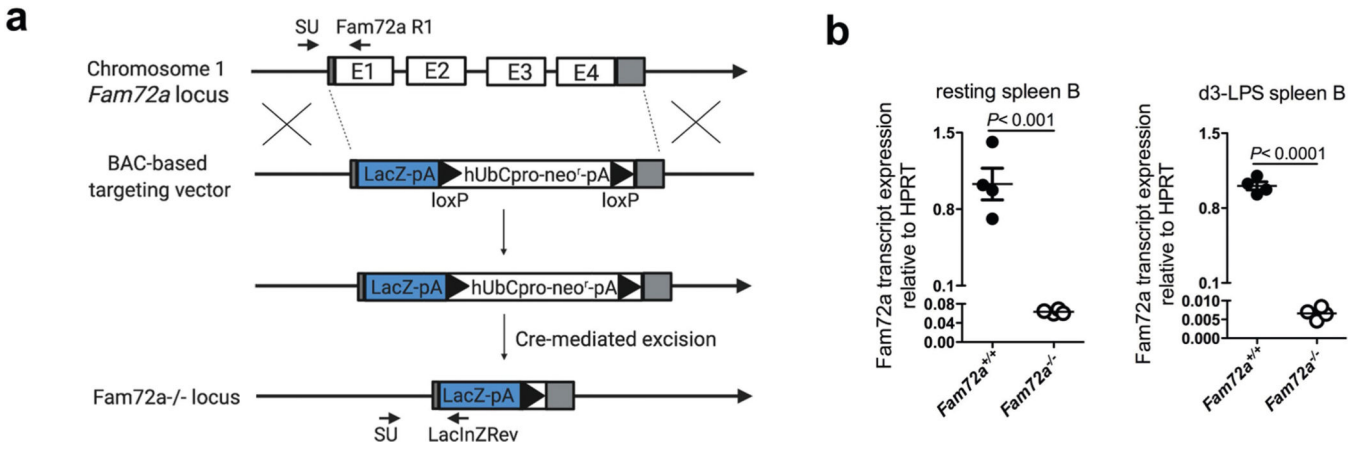
Extended Data



Extended Data Fig. 1 |. Validating the role of FAM72A during CSR in CH12 cells.

(**a**) Western blot analysis to identify Cas9-expressing CH12 cell subclones after transduction with a lentiviral vector expressing Cas9. Experiment was repeated 2 times independently with similar results. (**b**) Schematic of the CRISPR/Cas9 sorting screen in CH12 cells. CH12 cells stably expressing Cas9 were transduced with a mouse gRNA library that contains ~5 gRNA per gene. Transduced populations were treated with CIT cocktail (CD40

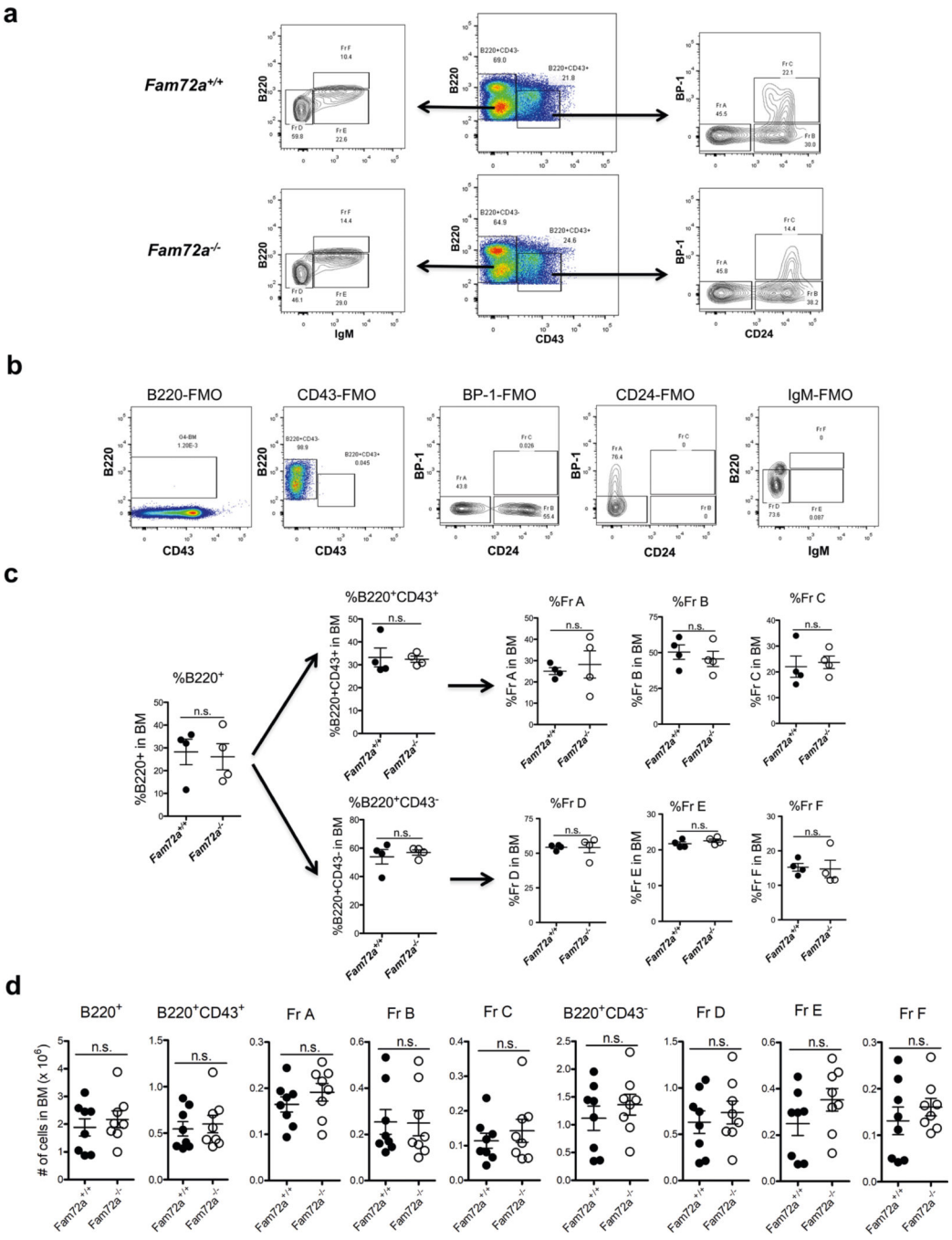
ligand, IL4, TGFβ) to induce CSR from IgM to IgA, followed FACS sorting to separate the switched (IgA^+) *versus* the unswitched (IgA^-) population. Genomic DNA from the initially transduced cells (T0), and from unsorted, IgA^-, IgA^+ population was isolated, and sequenced. (**c**) gRNA was ranked using a NormZ plot (standard deviations from the mean) comparing IgA^- cells to the unsorted population. (**d**) Guide RNA (gRNA) targeting strategy against mouse *Aicda* and *Fam72a* genes to validate the role of FAM72A in CSR in bulk CH12 cells expressing Cas9. (**e**) Quantification of *Fam72a* mRNA relative to HPRT in wild-type (WT) and $Fam72a^{-/-}$ CH12 clones by qPCR. (**f**) Sequenced *Fam72a, Msh2, Ung, and ligase4* alleles in $Fam72a^{-/-}$, $Ung^{-/-}$, $Fam72a^{-/-}Ung^{-/-}$, $Msh2^{-/-}$, $Fam72a^{-/-}Msh2^{-/-}$, and $Ligase\ 4^{-/-}$ CH12 clones generated using CRISPR/Cas9. Underlined sequence denotes gRNA target sequence, with the wildtype amino acid sequence indicated at the bottom. (**g**) Assessing the role of FAM72A in DNA double-strand break repair pathways. WT, $Fam72a^{-/-}$ and $Ligase\ 4^{-/-}$ CH12 clones stably expressing EJ5-GPF, EJ2-GFP, and DR-GFP substrates that measure non-homologous end joining (NHEJ), alternative end joining (A-EJ), and homologous recombination (HR), respectively. Cells were mock transfected or transfected with yeast endonuclease I-SceI expressing vector, pCBA-SceI. GFP expression was monitored by flow cytometry three days post-transfection. Data were presented as mean ± SEM and were analyzed using two-tailed unpaired Student t test. Left panel: p=0.028 for the comparison between WT clone 1 and $Lig4^{-/-}$ clone; middle panel: p=0.014 for the comparison between WT clone I and $Fam72a^{-/-}$ clone 1; p=0.0003 for the comparison between WT clone I and $Lig4^{-/-}$ clone; ns: not significant. Data are representative of 3 independent experiments



Extended Data Fig. 2 |. Generation of $Fam72a^{-/-}$ mice.

(**a**) Schematic representation of *Fam72a* gene disruption strategy. The whole coding sequence of *Fam72a* was replaced with a LacZ/neo cassette by homologous recombination using a bacterial artificial chromosome (BAC)-based targeting vector. The floxed neo cassette was removed by further breeding to a ubiquitous Cre mouse strain. Strain development was done at MMRRC (UC, Davis). PCR genotyping primer sequence can be found in Supplementary Table 2. (**b**) qPCR analysis of *Fam72a* mRNA from resting spleen B cells or spleen B cells that were *ex vivo* stimulated with LPS for 3 days (n= 4 mice per group). Data were presented as mean ± SEM and were analyzed using two-tailed unpaired

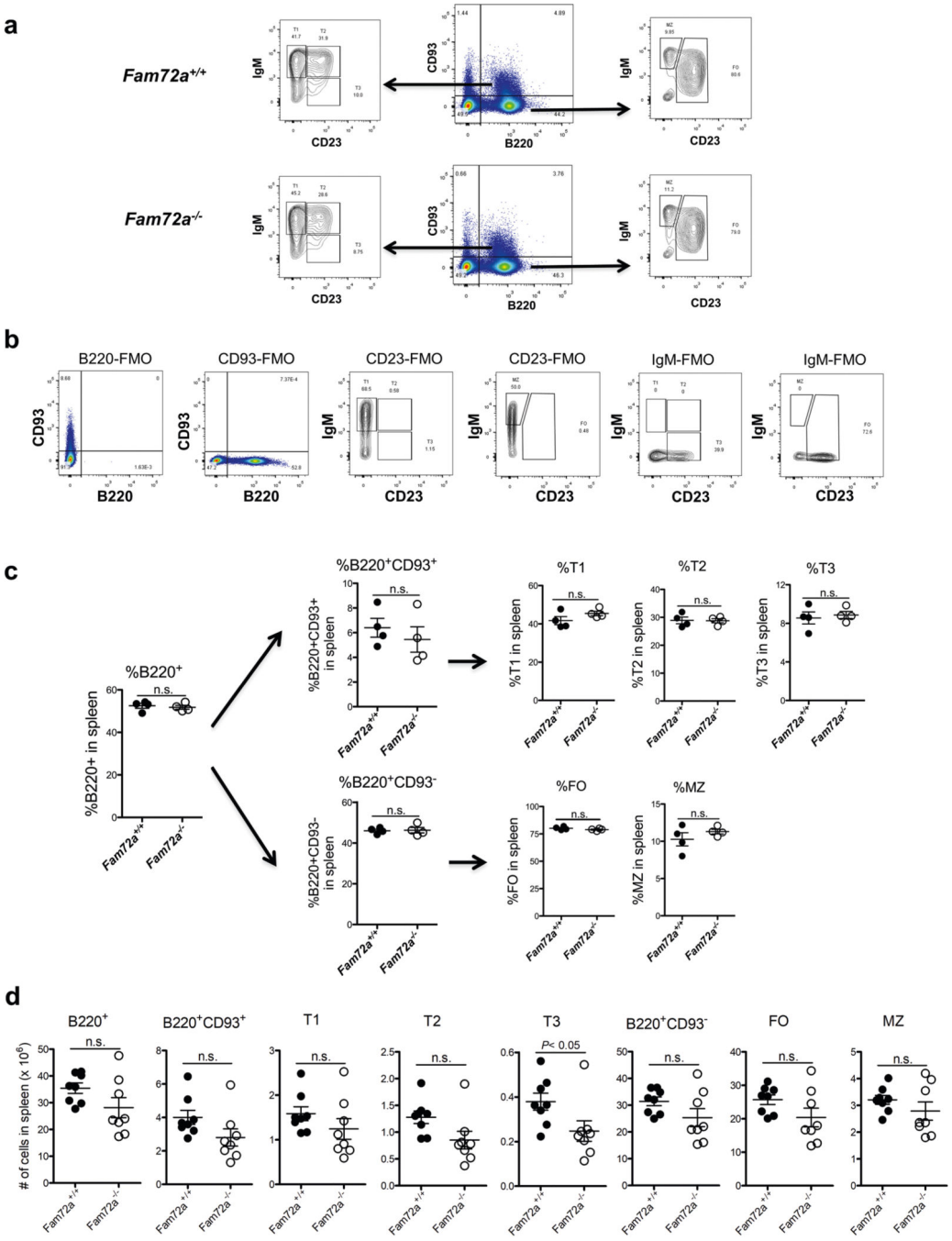
Student t test (p= 0.0006 for the comparison in the left panel). Data are representative of 2 independent experiments.



Extended Data Fig. 3 |. Comparable B cell profiles in the bone marrow of *Fam72a*$^{-/-}$ and *Fam72a*$^{+/+}$ littermate mice.

(**a**) Representative FACS plots of B cells derived from the bone marrow of *Fam72a*$^{-/-}$ or *Fam72a*$^{+/+}$ littermate mice. (**b**) Full minus one (FMO)-derived background staining for B220, CD43, BP-1, CD24 and IgM, respectively. (**c**) The frequencies of indicated B cell fractions in the bone marrow of *Fam72a*$^{-/-}$ or *Fam72a*$^{+/+}$ littermates (n= 4 mice per group).

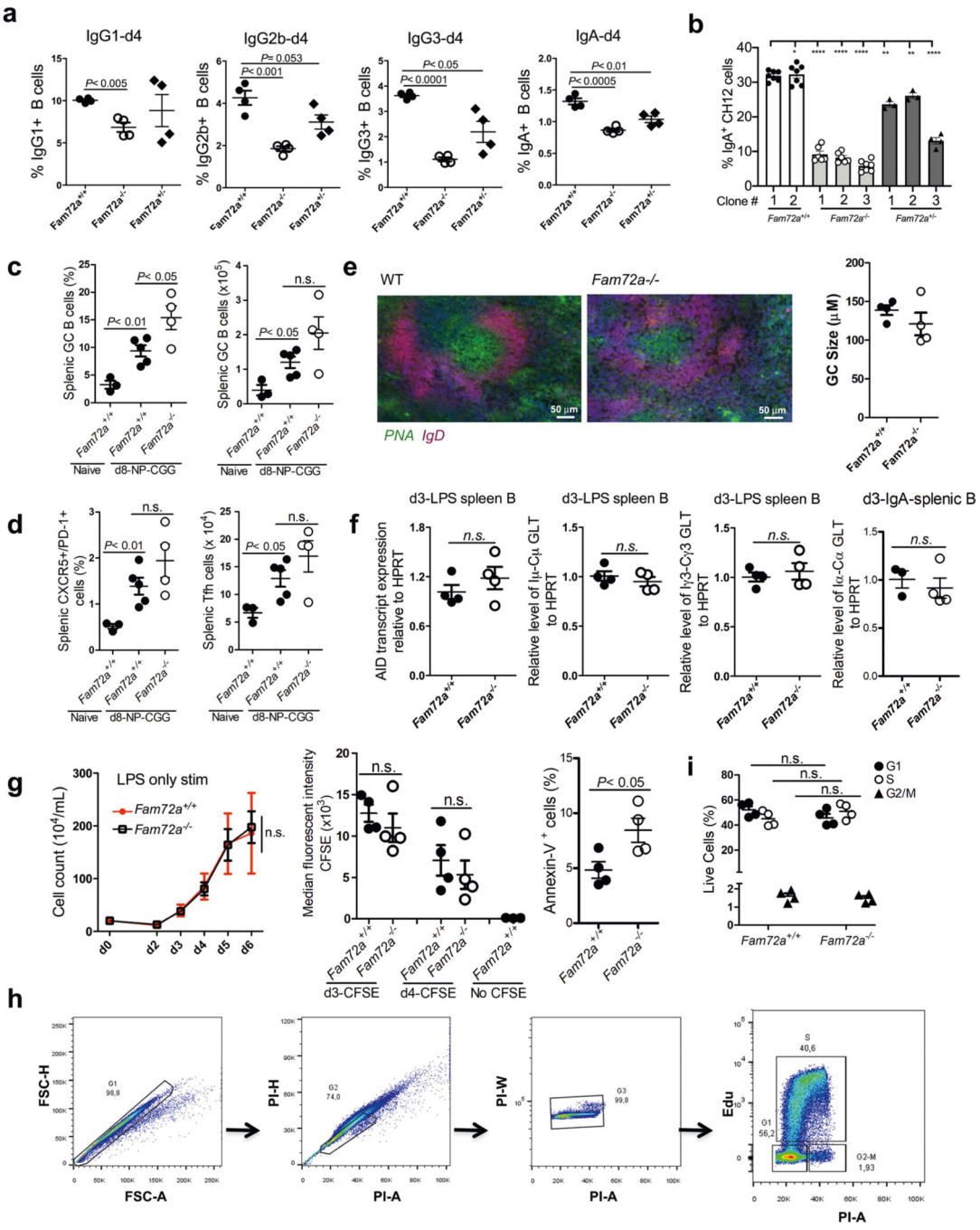
Data were presented as mean ± SEM and were analyzed using two-tailed unpaired Student t test (ns: not significant). Data are representative of 2 independent experiments. (**d**) Same as **c** except that total cell numbers were reported.



Extended Data Fig. 4 |. Comparable B cell profiles in the spleen of *Fam72a*$^{-/-}$ and *Fam72a*$^{+/+}$ littermates.

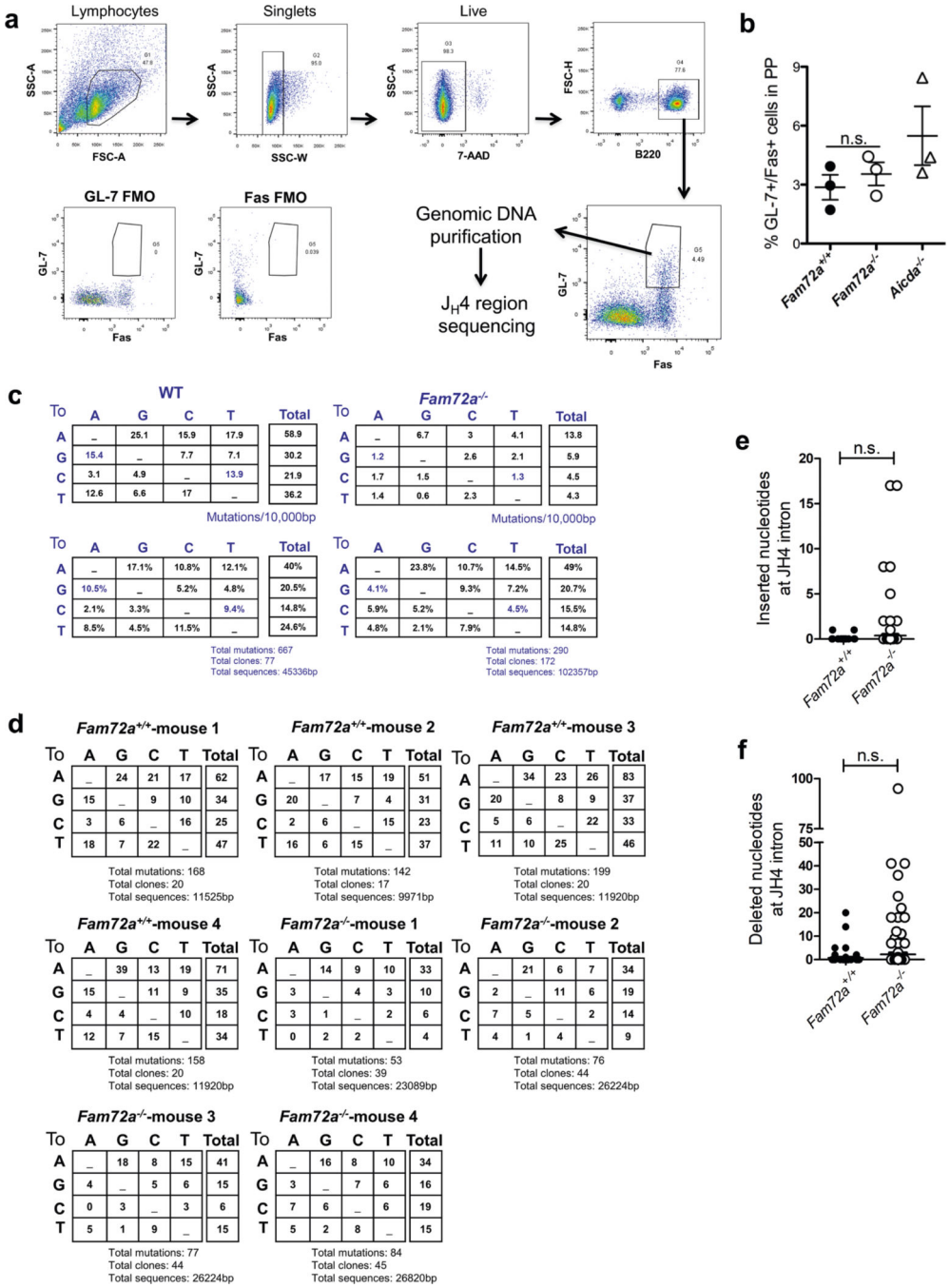
(**a**) Representative FACS plots of B cells in the spleen of *Fam72a*$^{-/-}$ or *Fam72a*$^{+/+}$ littermate mice. (**b**) FMO-derived background staining for B220, CD93, CD23 and IgM, respectively. (**c**) The frequencies of indicated B cell subsets in the spleen of *Fam72a*$^{-/-}$ or *Fam72a*$^{+/+}$

littermates (n= 4 mice per group). Data were presented as mean ± SEM and were analysed using two-tailed unpaired Student t test (ns: not significant). Data are representative of 2 independent experiments. MZ: marginal zone B cells; FO: follicular B cells. (**d**) Same as **c** except that total cell numbers were reported.



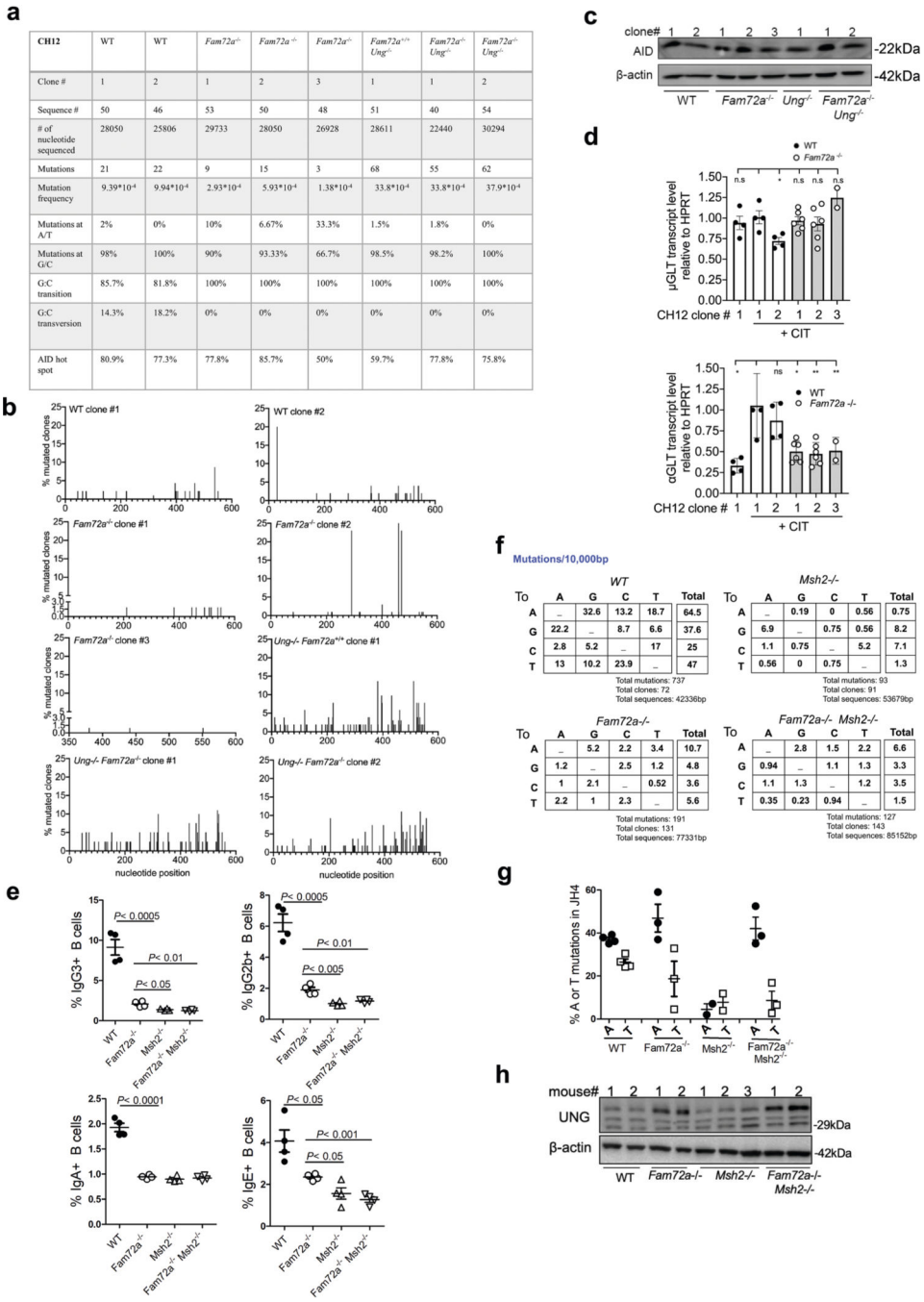
Extended Data Fig. 5 |. CSR and germinal center formation analysis from *Fam72a*$^{-/-}$ *Fam72a*$^{+/-}$ and WT mice and CH12 clones, as well as AID and germline transcripts, cell proliferation and cell cycle profile of splenic B cells.

(**a**) Analysis of *ex vivo* CSR in splenic B cells from *Fam72a*$^{+/+}$, *Fam72a*$^{+/-}$, and *Fam72a*$^{-/-}$ mice (n=2 mice per group; duplicate assays per mouse). For IgG1 panel, p= 0.0014 for the comparison between *Fam72a*$^{+/+}$ and *Fam72a*$^{-/-}$ group; for IgG2b panel, p= 0.0005 for the comparison between *Fam72a*$^{+/+}$ and *Fam72a*$^{-/-}$ group; for IgG3 panel, p= 0.0143 for the comparison between *Fam72a*$^{+/+}$ and *Fam72a*$^{+/-}$ group; for IgA panel, p= 0.0002 for the comparison between *Fam72a*$^{+/+}$ and *Fam72a*$^{-/-}$ group, and p= 0.0074 for the comparison between *Fam72a*$^{+/+}$ and *Fam72a*$^{+/-}$ group. (**b**) CH12 clones of indicated genotype were treated with CIT for 2 days and analyzed for CSR to IgA. Data are representative of 3 independent experiments. (**c**) Characterization of the humoral response in NP-CGG immunized WT and *Fam72a*$^{-/-}$ mice. Splenic germinal center (GC) B cell analysis in WT and *Fam72a*$^{-/-}$ mice after NP-CGG immunization for 8 days. Y-axis in the left panel shows the frequency of GC (GL-7^{+}Fas^{+}) among live CD45^{+}CD4^{-}CD8^{-}CD11c^{-}F4/80^{-}CD19^{+}IgD^{-} splenic cells (n= 3–5 mice per group). For the left panel, p= 0.0059 for the comparison between naïve and d8-immunized *Fam72a*$^{+/+}$ group, and p= 0.0308 for the comparison between d8-immunized *Fam72a*$^{+/+}$ and *Fam72a*$^{-/-}$ group; for the right panel, p= 0.0176 for the comparison between naïve and d8-immunized *Fam72a*$^{+/+}$ group. (**d**) Same as (c), except that follicular helper T (Tfh) cells were examined. Y-axis in the left panel shows the frequency of CXCR5^{+}PD-1^{+} among live CD4^{+}B220^{-} splenic cells (n= 3–5 mice per group). For the left panel, p= 0.0123 for the comparison between naïve and d8-immunized *Fam72a*$^{+/+}$ group; for the right panel, p= 0.0254 for the comparison between naïve and d8-immunized *Fam72a*$^{+/+}$ group. (**e**) Representative immunofluorescence analysis of splenic GCs post NP immunization in WT and *Fam72a*$^{-/-}$ mice. Magnification 20x. Spleens were collected day 8 post immunization. Cryosection were stained with PNA-FITC and IgD-PE. Right panel: The average GC size for each mouse is reported, in which ~2–6 individual GCs were analyzed per mouse (n=4 mice per group). (**f**) qPCR analysis of AID mRNA, Iμ-Cμ and Iγ3-Cγ3 germline transcripts of d3-LPS stimulated splenic B cells, as well as Iα-Cα germline transcripts of splenic B cells stimulated to switch to IgA from *Fam72a*$^{-/-}$ or *Fam72a*$^{+/+}$ littermate mice (n= 3–4 mice per group). Data are representative of 2 independent experiments. (**g**) Evaluation of cell proliferation of LPS-stimulated splenic B cells from *Fam72a*$^{-/-}$ or *Fam72a*$^{+/+}$ littermate mice (n=4 mice per group), and data was tested by two-way ANOVA and representative of 2 independent experiments (left panel). Proliferation was also assessed by CFSE dilution (Middle panel), and apoptosis by Annexin V-staining (right panel; n=3–4 mice per group; p= 0.0341). (**h**) Gating strategy for cell cycle analysis. (**i**) The compiled cell cycle analysis of *Fam72a*$^{-/-}$ or *Fam72a*$^{+/+}$ splenic B cells that were stimulated by LPS for 3 days (n= 4 mice per group). Data are representative of 2 independent experiments. Data were presented as mean ± SEM and were analyzed using two-tailed unpaired Student t test (ns: not significant).



Extended Data Fig. 6 |. Analysis of germinal center B cells and mutations from Peyer's patches. (**a**) Gating strategy used to sort germinal center B cells from Peyer's patches for *J_H4* region and *Bcl6 gene* sequencing. (**b**) The frequencies of germinal center B cells in Peyer's patches from naïve *Fam72a*$^{+/+}$, *Fam72a*$^{-/-}$ or *Aicda*$^{-/-}$ mice (n= 3 mice per group). Data are representative of 2 independent experiments. Data were presented as mean ± SEM and were analyzed using two-tailed unpaired Student t test (ns: not significant). (**c**) The compiled spectrum of unique top-strand (coding-strand) mutations shown as mutation frequencies (top panel) and percentage of total mutations (bottom panel). These two panels represented the

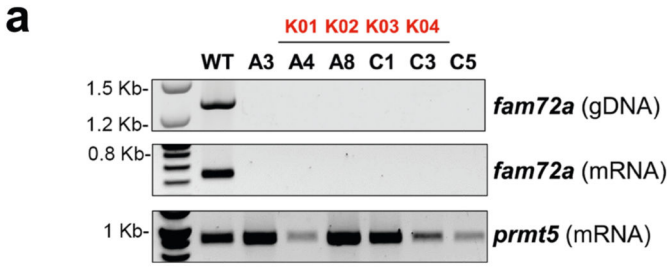
pooled data from 4 *Fam72a*$^{+/+}$ and 4 *Fam72a*$^{-/-}$ mice, respectively. Data shown in the upper panels showing mutation frequencies for each individual mutation type were analyzed using a MannWhitney test, and all types of mutations were significantly reduced in the *Fam72a*$^{-/-}$ mice, except for C to A mutations. (**d**) Analysis of SHM in the J_H4 region of germinal center B cells for each *Fam72a*$^{+/+}$ or *Fam72a*$^{-/-}$ mouse. (**e**) Size of insertion in the J_H4 region is shown (n= 4 mice per group). In addition, there was also no statistical significance in the frequency of insertions between *Fam72a*$^{+/+}$ and *Fam72a*$^{-/-}$ mice (data not shown). (**f**) Size of deletions in the J_H4 region is shown (n= 4 mice per group). In addition, there was also no statistical significance in the frequency of deletions between *Fam72a*$^{+/+}$ and *Fam72a*$^{-/-}$ mice (data not shown).

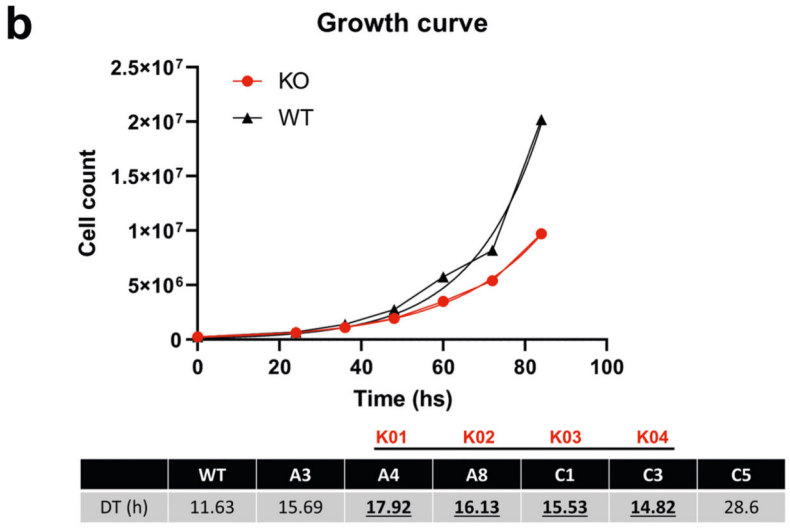


Extended Data Fig. 7 |. Characterizing the effect of FAM72A-deficiency on mutations.

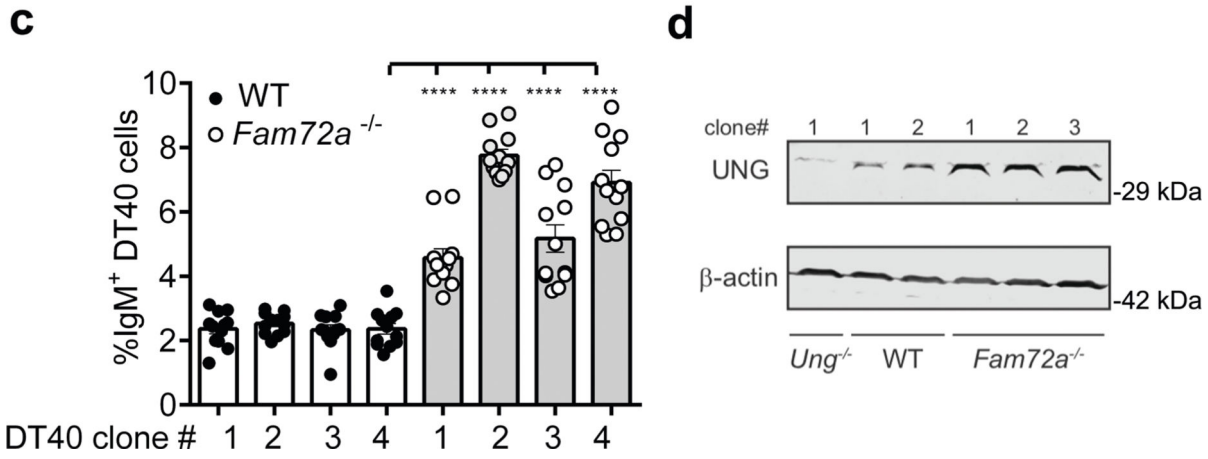
(**a**) Mutation characteristics at the 5′ Sμ region was analyzed in WT, *Fam72a*$^{-/-}$, *Ung*$^{-/-}$, and *Ung*$^{-/-}$ *Fam72a*$^{-/-}$ CH12 clones after 5 days of CIT treatment. Mutations at WRC and GYW motifs are considered AID hotspot mutation, where W=A/T, R=A/G, and Y=C/T. Mutation frequency was calculated from total mutations pooled from three independent experiment divided by total number of nucleotide sequenced. Data is summarized in Fig. 3a. (**b**) Distribution of mutations at the 5'μ switch region in CH12 cells of the indicated genotype after 5 days of CIT stimulation. (**c**) AID expression levels of WT,

Fam72a$^{-/-}$, *Ung*$^{-/-}$, and *Fam72a*$^{-/-}$*Ung*$^{-/-}$ CH12 cells as assessed by Western blot analysis. β-ACTIN was used as loading control. Experiment was repeated 3 times independently with similar results. (**d**) Sμ germline transcripts (μGLT) and Sα germline transcript (αGLT) of the indicated CH12 clones before and after stimulation with CIT for 2 days. Data are representative of 3 independent experiments. (**e**) Analysis of *ex vivo* CSR to IgG2b, IgG3, IgA and IgE using splenic B cells from WT, *Fam72a*$^{-/-}$, *Msh2*$^{-/-}$, or *Fam72a*$^{-/-}$*Msh2*$^{-/-}$ littermate mice (n=2 mice per group; duplicate assays per mouse). Data are representative of 2 independent experiments. For IgG3 panel, p= 0.0004 for the comparison between WT and *Fam72a*$^{-/-}$ group, p= 0.011 for the comparison between *Fam72a*$^{-/-}$ and *Msh2*$^{-/-}$ group, and p= 0.005 for the comparison between *Fam72a*$^{-/-}$ and *Fam72a*$^{-/-}$*Msh2*$^{-/-}$ group; for IgG2b panel, p= 0.0003 for the comparison between WT and *Fam72a*$^{-/-}$ group, p= 0.0024 for the comparison between *Fam72a*$^{-/-}$ and *Msh2*$^{-/-}$ group, and p= 0.0059 for the comparison between *Fam72a*$^{-/-}$ and *Fam72a*$^{-/-}$*Msh2*$^{-/-}$ group; for IgE panel, p= 0.0184 for the comparison between WT and *Fam72a*$^{-/-}$ group, p= 0.0326 for the comparison between *Fam72a*$^{-/-}$ and *Msh2*$^{-/-}$ group, and p= 0.0005 for the comparison between *Fam72a*$^{-/-}$ and *Fam72a*$^{-/-}$*Msh2*$^{-/-}$ group. (**f**) The compiled spectrum of unique top-strand (coding-strand) mutations shown as mutation frequencies per 10, 000bp. The data were pooled from 4 WT mice, 3 *Fam72a*$^{-/-}$, 2 *Msh2*$^{-/-}$, or 3 *Fam72a*$^{-/-}$*Msh2*$^{-/-}$ littermate mice, respectively. (**g**) The frequency of A or T mutations in the J_H4 region of WT, *Fam72a*$^{-/-}$, *Msh2*$^{-/-}$, or *Fam72a*$^{-/-}$*Msh2*$^{-/-}$ littermate mice (n=2–4 mice per group). (**h**) Western blots measuring for UNG expression level in *ex vivo* LPS-activated spleen B cells from WT, *Fam72a*$^{-/-}$, *Msh2*$^{-/-}$, and *Fam72a*$^{-/-}$*Msh2*$^{-/-}$ mice. Experiment was repeated 2 times independently with similar results.





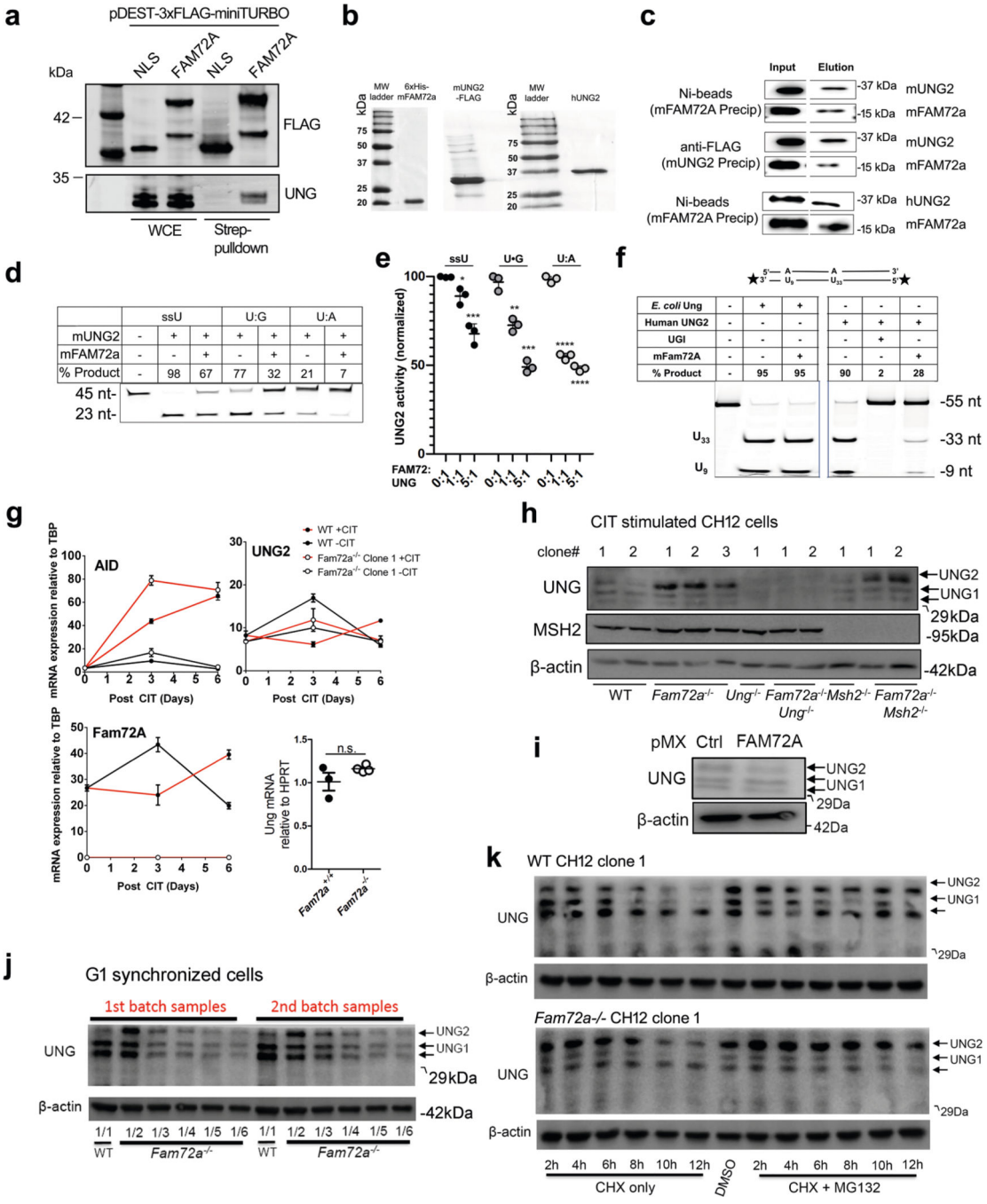
			K01	K02	K03	K04	
	WT	A3	A4	A8	C1	C3	C5
DT (h)	11.63	15.69	17.92	16.13	15.53	14.82	28.6



Extended Data Fig. 8 |. Characterization of chicken DT40 *Fam72a*$^{-/-}$ clones.

(**a**) *Fam72a*$^{-/-}$ DT40 clones were generated using CRISPR/Cas9. Depletion of *Fam72a* was confirmed at the genomic DNA and mRNA level in 6 different DT40 clones. All the clones were mostly IgM^{-} as assessed by flow cytometry and only 4 clones (renames K01–04) were picked for fluctuation analysis. *Prmt5* was used as a control for mRNA extraction. (**b**) Growth curve analysis in WT and *Fam72a*$^{-/-}$ DT40 cells. The average doubling time (DT) of WT and *Fam72a*$^{-/-}$ clones in cultures for each individual clone is shown in the bottom panel and revealed reduced growth kinetics for all *Fam72a*$^{-/-}$ clones (**c**) Fluctuation

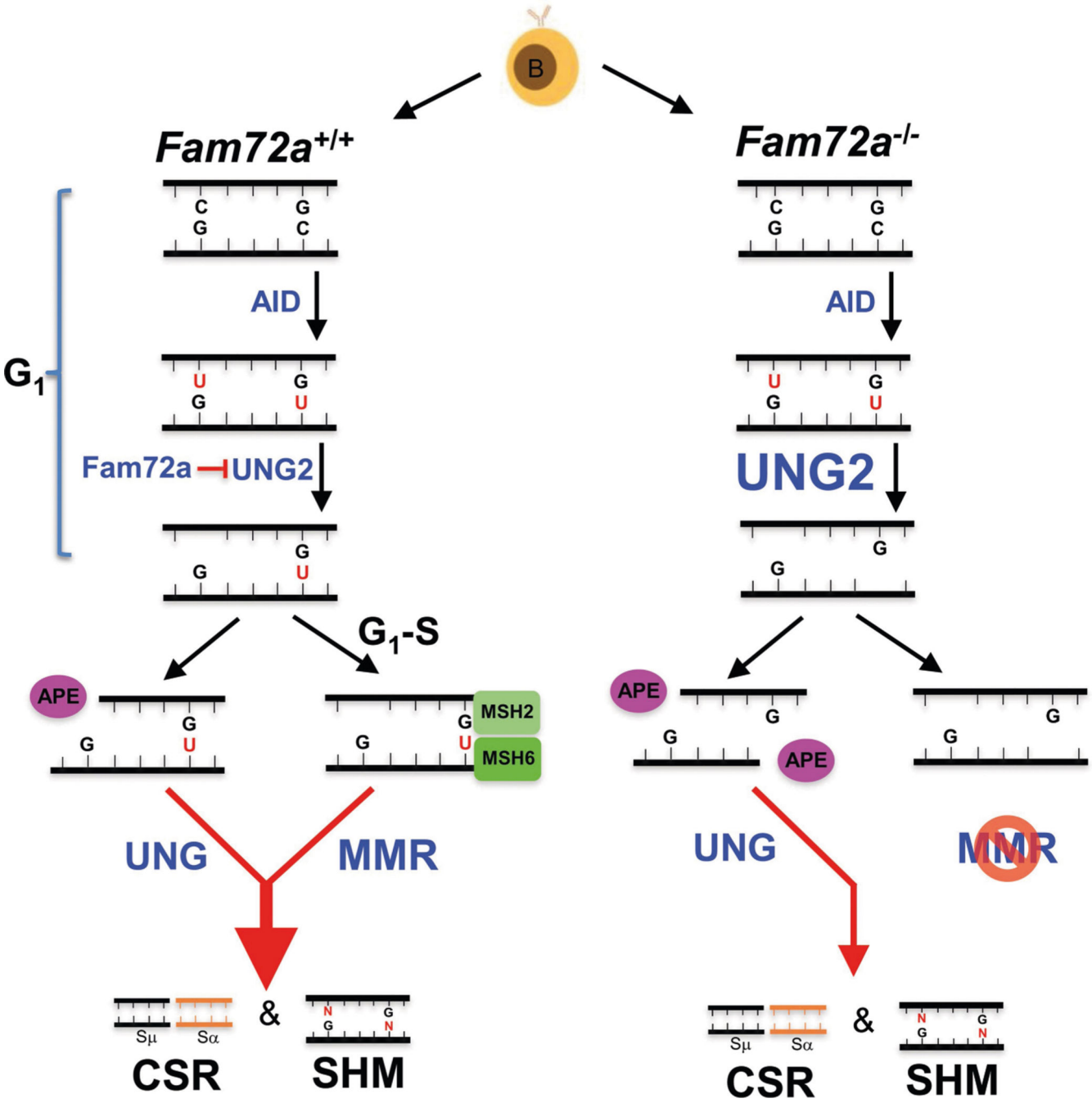
analysis for gene conversion in WT and *Fam72a*$^{-/-}$ DT40 cells based on same number of cell divisions. Data are representative of 3 independent experiments. (**d**) Western blots for UNG were carried out on the indicated clones.



Extended Data Fig. 9 |. FAM72A binds to and inhibits UNG2, leading to reduced UNG2 protein level.

(**a**) Streptavidin affinity purification from 293T cell lysates expressing N-terminal 3xFLAG-miniTURBO tagged FAM72A upon proximity biotinylation. FLAG-miniTURBO-SV40-NLS (NLS) was used as control. Strepavidin bound proteins and whole-cell extracts

(WCEs) were immunoblotted with the indicated antibodies. (**b**) Purified mFAM72A and UNG2 proteins were electrophoresed on an SDS-PAGE gel and stained with Coomassie Brilliant Blue dye. The 6XHis-mFAM72A (20 kDa), mUNG2-polyGlyFLAG (35kDa), and hUNG2 (35kDa) are seen next to the precision plus protein standard (Bio-Rad). (**c**) Murine UNG2 and FAM72A proteins were mixed in equal molar amounts and pulled-down using polyHis-tag or FLAG-tag on mFAM72A and mUNG2 proteins, respectively. The input and elutions were analyzed by Western blot using anti-His tag or anti-FLAG tag antibodies. Cross-species pull-down of mFAM72A with hUNG2 is also shown: the hUNG2 protein was detected using anti-UDG antibody. (**d**) Purified murine UNG2 and Fam72A proteins were pre-incubated at a 1:5 molar ratio, respectively, then reacted with either a single-stranded uracil substrate (ssU) or double-stranded uracil substrates with either U:G or U:A pairs. Reaction products were separated on a denaturing polyacrylamide gel. (**e**) Quantification of three independent experiments of the type shown in Extended Data Fig. 9d using 0:1, 1:1 or 5:1 molar ratios of mFAM72A to mUNG2. (**f**) Inhibition of hUNG2, but not *E. coli* UNG by mFAM72A. Purified *E. coli* UNG or hUNG2 proteins were pre-incubated with a five-fold molar excess of mFAM72A and then reacted with a double-stranded DNA substrate, with uracils at position 9 and 33 in a 55 bp DNA. Both the ends of the uracil-containing oligomer are labeled with 6-FAM (shown as stars). Consequently, two major excision products (U9 and U33) are observed. Uracil glycosylase inhibitor (Ugi) was used in one of the reactions. (**g**) The mRNAs for *Aicda, Ung2*, and *Fam72a* genes were quantified from WT (clone 1) and *Fam72a*$^{-/-}$ (clone 1) cell lines that had been stimulated with the CIT cocktail (+CIT) or untreated (-CIT). Mean values and standard deviations of three independent qRT-PCR reactions are shown for each treatment. The expression was normalized to the reference gene *TBP*. *Ung* transcript levels were also quantified in resting primary mouse B cells by qRT-PCR (bottom right panel; n=4 per group). Data were presented as mean ± SEM and were analyzed using two-tailed unpaired Student t test (ns: not significant). (**h**) Western blots for UNG, MSH2, and β-ACTIN (as control) in CIT-stimulated CH12 clones of the indicated genotype. UNG1 and UNG2 are indicated on the gel. Experiment was repeated 3 times independently with similar results. (**i**) Western blot analysis for UNG in WT CH12 cells expressing the empty vector or overexpressing FAM72A. Experiment was repeated 3 times independently with similar results. (**j**) WT CH12 clone#1 and *Fam72*$^{-/-}$ CH12 clone#1 were treated with CHX or CHX plus MG132 for the indicated time, and lysates were probed with anti-UNG or anti-β-actin antibodies. Data are representative of 3 independent experiments. (**k**) Lysates from G_1-synchronized *Fam72a*$^{-/-}$ CH12 cells were diluted to determine the level of increase of UNG2 protein compared to undiluted WT CH12 cells, then probed for UNG protein by western blot. These blots suggest a 3–4 fold increase in UNG2 protein in *Fam72a*$^{-/-}$ CH12 cells compared to controls. Data are representative of 3 independent experiments.



Extended Data Fig. 10 |. Model for the role of FAM72A in SHM and CSR.

FAM72A antagonizes UNG2 in the G_1 phase of the cell cycle, leading to reduced processing of the AID-induced dU in *Ig* genes, which leads to increased G:U mismatches. The increased G:U lesions can either be replicated to produce transition mutations at G:C basepairs, or can engage the mismatch repair system to enhance mutagenesis, or provide DNA lesions, in collaboration with UNG2 that are required for CSR. In the absence of FAM72A, the protein expression and enzymatic activity of UNG2 are enhanced, resulting in increased excision of dU, which favors faithful repair by base excision repair pathway,

as evidenced by diminished mutation frequencies in both μ switch region and J_H4 region. Furthermore, the accelerated removal of dU in the context of FAM72A deficiency fails to recruit the mismatch repair system in the G_1 or S phase, which is required for both CSR and SHM. One question that arises, is why doesn't CSR increase with increased UNG2 protein/activity in *Fam72a*$^{-/-}$ B cells? The key is the engagement of the MMR system by FAM72A to induce CSR. First, deleting FAM72A in *Msh2*$^{-/-}$ CH12 cells and *Msh2*$^{-/-}$ spleen B cells moderately increases CSR (Fig. 3d, f) consistent with an increased UNG2 level and activity inducing larger number of breaks to facilitate CSR. Importantly, CSR is not increased to WT levels. For SHM, the increased UNG2 activity in *Fam72a*$^{-/-}$ B cells would cause an increase in ssDNA breaks in the V-region, but ultimately lead to faithful repair of the dUs, thereby erasing many AID-induced uracil lesions. Second, removal of either *Ung* or *Msh2* leads to a defect in CSR that is not additive suggesting an interaction between these two proteins during CSR (Fig. 3c, d). For SHM, UNG and MMR pathways converge to induce error-prone repair. The exact mechanism likely involves the interruption of faithful MMR repair by uracil excision by UNG2. The interaction between these both pathways has not been formally shown in CSR, although there was evidence for this notion in the literature (e.g. see figure 3d in[6]). As to how MMR and UNG2 work together to induce CSR is not clear, and requires more work, but the mechanism might be similar to that hypothesized to what occurs during SHM. For CSR, we therefore hypothesize that FAM72A antagonizes UNG2, leading to the accumulation of U:G mispairs that can be engaged by the MMR system to induce CSR.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank P. Poussier and M. Shulman for critical review of the manuscript. Y.F., C.L. and A.A.-Q. are recipients of the Canadian Institutes of Health Research Postdoctoral Fellowship. N.S.D. is the recipient of a doctoral award from the FRQ-S (Fonds de recherche Santé Québec). J.A.S. was supported by a competitive graduate research assistantship from the Wayne State University and R.M.-R. was supported by a Thomas C. Rumble fellowship. J.M.D.N. is a Merit scholar from the Fonds de recherche de Quebec– Santé and is supported by CIHR (PJT-155944). J.M. is supported from a CIHR project grant (CBT-438323) and holds a Canada Research Chair in Functional Genomics. D.D. is a Canada Research Chair (Tier I) and work in the D.D. laboratory was supported by a grant from the CIHR (FDN143343). A.S.B. was supported by a National Institutes of Health grant (1R21AI144708) and Bridge Funding grant from Wayne State University. A.M. is supported by grants from the CIHR (PJT-153307 and PJT-156330).

References

1. Muramatsu M et al. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell 102, 553–563 (2000). [PubMed: 11007474]
2. Feng Y, Seija N, JM DIN & Martin A AID in antibody diversification: there and back again. Trends Immunol. 41, 586–600 (2020). [PubMed: 32434680]
3. Cascalho M, Wong J, Steinberg C & Wabl M Mismatch repair co-opted by hypermutation. Science 279, 1207–1210 (1998). [PubMed: 9469811]
4. Di Noia J & Neuberger MS Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. Nature 419, 43–48 (2002). [PubMed: 12214226]

5. Wiesendanger M, Kneitz B, Edelmann W & Scharff MD Somatic hypermutation in MutS homologue (MSH)3-, MSH6-, and MSH3/MSH6-deficient mice reveals a role for the MSH2-MSH6 heterodimer in modulating the base substitution pattern. J. Exp. Med 191, 579–584 (2000). [PubMed: 10662804]
6. Rada C, Di Noia JM & Neuberger MS Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. Mol. Cell 16, 163–171 (2004). [PubMed: 15494304]
7. Rada C et al. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. Curr. Biol 12, 1748–1755 (2002). [PubMed: 12401169]
8. Guo C et al. Ugene, a newly identified protein that is commonly overexpressed in cancer and binds uracil DNA glycosylase. Cancer Res. 68, 6118–6126 (2008). [PubMed: 18676834]
9. Muramatsu M et al. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. J. Biol. Chem 274, 18470–18476 (1999). [PubMed: 10373455]
10. Lawson KA et al. Functional genomic landscape of cancer-intrinsic evasion of killing by T cells. Nature 586, 120–126 (2020). [PubMed: 32968282]
11. Pan-Hammarstrom Q et al. Impact of DNA ligase IV on nonhomologous end joining pathways during class switch recombination in human cells. J. Exp. Med 201, 189–194 (2005). [PubMed: 15657289]
12. Ward IM et al. 53BP1 is required for class switch recombination. J. Cell Biol 165, 459–464 (2004). [PubMed: 15159415]
13. Yan CT et al. IgH class switching and translocations use a robust non-classical end-joining pathway. Nature 449, 478–482 (2007). [PubMed: 17713479]
14. Perez-Duran P et al. UNG shapes the specificity of AID-induced somatic hypermutation. J. Exp. Med 209, 1379–1389 (2012). [PubMed: 22665573]
15. Frieder D, Larijani M, Collins C, Shulman M & Martin A The concerted action of Msh2 and UNG stimulates somatic hypermutation at A.T base pairs. Mol. Cell. Biol 29, 5148–5157 (2009). [PubMed: 19596785]
16. Thientosapol ES. et al. . Proximity to AGCT sequences dictates MMR-independent versus MMR-dependent mechanisms for AID-induced mutation via UNG2. Nucleic Acids Res. 45, 3146–3157 (2017). [PubMed: 28039326]
17. Martin A et al. Msh2 ATPase activity is essential for somatic hypermutation at a-T basepairs and for efficient class switch recombination. J. Exp. Med 198, 1171–1178 (2003). [PubMed: 14568978]
18. Phung QH et al. Increased hypermutation at G and C nucleotides in immunoglobulin variable genes from mice deficient in the MSH2 mismatch repair protein. J. Exp. Med 187, 1745–1751 (1998). [PubMed: 9607916]
19. Delbos F, Aoufouchi S, Faili A, Weill JC & Reynaud CA DNA polymerase eta is the sole contributor of A/T modifications during immunoglobulin gene hypermutation in the mouse. J. Exp. Med 204, 17–23 (2007). [PubMed: 17190840]
20. Arakawa H, Saribasak H & Buerstedde JM Activation-induced cytidine deaminase initiates immunoglobulin gene conversion and hypermutation by a common intermediate. PLoS Biol. 2, E179 (2004). [PubMed: 15252444]
21. Di Noia JM & Neuberger MS Immunoglobulin gene conversion in chicken DT40 cells largely proceeds through an abasic site intermediate generated by excision of the uracil produced by AID-mediated deoxycytidine deamination. Eur. J. Immunol 34, 504–508 (2004). [PubMed: 14768055]
22. Campo VA et al. MSH6- or PMS2-deficiency causes re-replication in DT40 B cells, but it has little effect on immunoglobulin gene conversion or on repair of AID-generated uracils. Nucleic Acids Res. 41, 3032–3046 (2013). [PubMed: 23314153]
23. Di Noia JM, Rada C & Neuberger MS SMUG1 is able to excise uracil from immunoglobulin genes: insight into mutation versus repair. EMBO J. 25, 585–595 (2006). [PubMed: 16407970]
24. Wang Q et al. The cell cycle restricts activation-induced cytidine deaminase activity to early G1. J. Exp. Med 214, 49–58 (2017). [PubMed: 27998928]

25. Cappelli E et al. Rates of base excision repair are not solely dependent on levels of initiating enzymes. Carcinogenesis 22, 387–393 (2001). [PubMed: 11238177]
26. Krusong K, Carpenter EP, Bellamy SR, Savva R & Baldwin GS A comparative study of uracil-DNA glycosylases from human and herpes simplex virus type 1. J. Biol. Chem 281, 4983–4992 (2006). [PubMed: 16306042]
27. Krokan HE et al. Error-free versus mutagenic processing of genomic uracil–relevance to cancer. DNA Repair 19, 38–47 (2014). [PubMed: 24746924]
28. Zeng X et al. DNA polymerase eta is an A–T mutator in somatic hypermutation of immunoglobulin variable genes. Nat. Immunol 2, 537–541 (2001). [PubMed: 11376341]
29. Belcheva A et al. Gut microbial metabolism drives transformation of MSH2-deficient colon epithelial cells. Cell 158, 288–299 (2014). [PubMed: 25036629]
30. Ramachandran S et al. The SAGA deubiquitination module promotes DNA repair and class switch recombination through ATM and DNAPK-mediated γH2AX formation. Cell Rep. 15, 1554–1565 (2016). [PubMed: 27160905]
31. Aregger M, Chandrashekhar M, Tong AHY, Chan K & Moffat J Pooled lentiviral CRISPR–Cas9 screens for functional genomics in mammalian cells. Methods Mol. Biol 1869, 169–188 (2019). [PubMed: 30324523]
32. Sarno A et al. Uracil–DNA glycosylase UNG1 isoform variant supports class switch recombination and repairs nuclear genomic uracil. Nucleic Acids Res. 47, 4569–4585 (2019). [PubMed: 30838409]
33. Li C et al. The H2B deubiquitinase Usp22 promotes antibody class switch recombination by facilitating non-homologous end joining. Nat. Commun 9, 1006 (2018). [PubMed: 29520062]
34. Li C et al. Early-life programming of mesenteric lymph node stromal cell identity by the lymphotoxin pathway regulates adult mucosal immunity. Sci. Immunol 4, aax1027 (2019).
35. Liu M et al. Two levels of protection for the B cell genome during somatic hypermutation. Nature 451, 841–845 (2008). [PubMed: 18273020]
36. Siriwardena SU, Perera MLW, Senevirathne V, Stewart J & Bhagwat AS A tumor-promoting phorbol ester causes a large increase in APOBEC3A expression and a moderate increase in APOBEC3B expression in a normal human keratinocyte cell line without increasing genomic uracils. Mol. Cell. Biol 39, e00238–18 (2019).
37. So CC, Ramachandran S & Martin A E3 ubiquitin ligases RNF20 and RNF40 are required for double-stranded break (DSB) repair: evidence for monoubiquitination of histone H2B lysine 120 as a novel axis of DSB signaling and repair. Mol. Cell. Biol 39, e00488–18 (2019).
38. Boulianne B et al. AID-expressing germinal center B cells cluster normally within lymph node follicles in the absence of $FDC^{-}M1^{+}$ $CD35^{+}$ follicular dendritic cells but dissipate prematurely. J. Immunol 191, 4521–4530 (2013). [PubMed: 24068672]

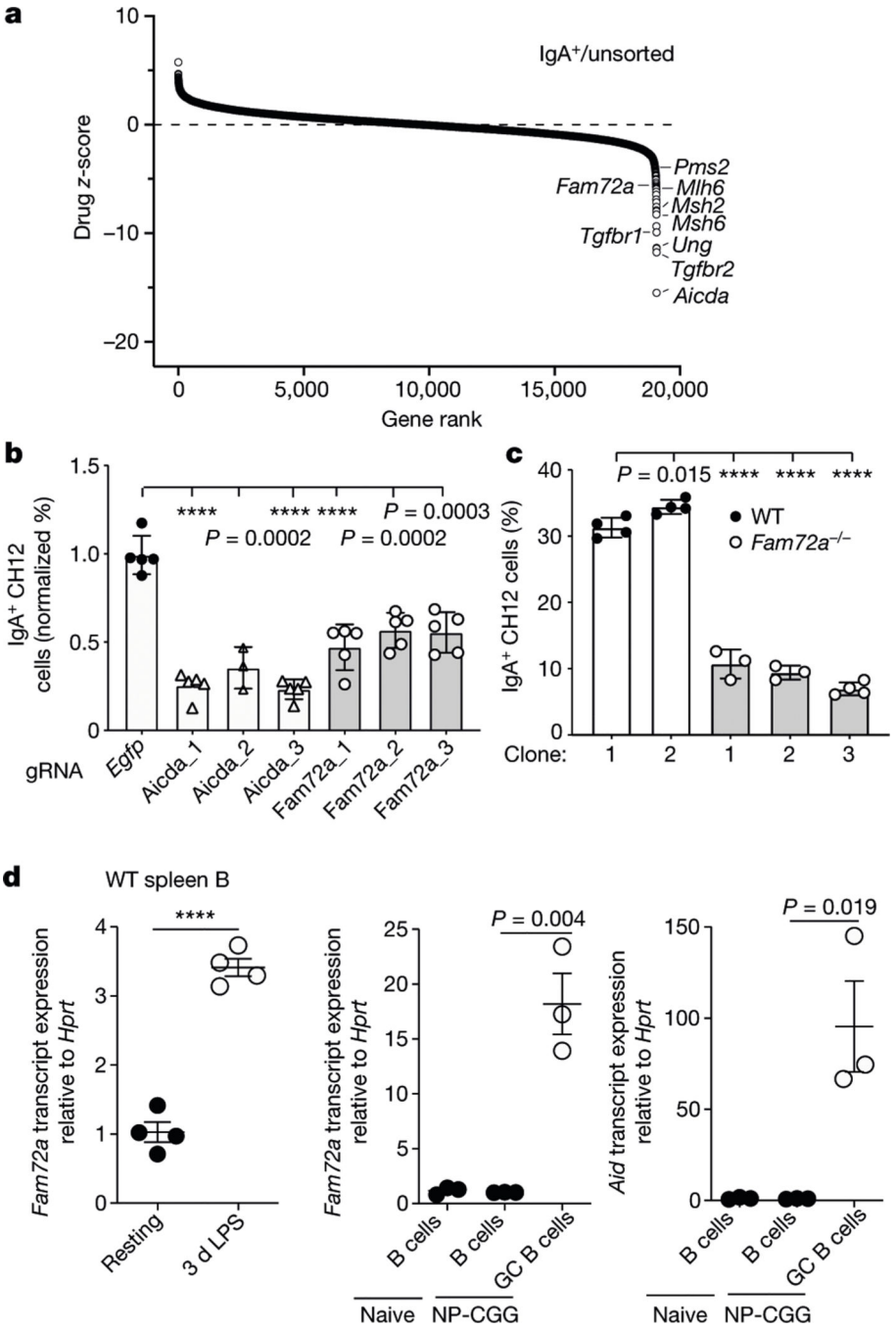


Fig. 1 |. Whole-genome CRISPR screen identified FAM72A as a factor required for CSR.
a, gRNAs were ranked using a NormZ plot comparing IgA$^+$ cells to the unsorted population. **b**, gRNAs targeting different areas of *Fam72a* and *Aicda* genes in Cas9-expressing CH12 cells followed by IgA-CSR assessment on day 2 of CIT treatment, measured relative to the CSR in CH12 cells expressing gRNAs targeting eGFP. **c**, CSR was assessed in the indicated CH12 clones on day 2 of CIT treatment. **d**, Left, mouse splenic B cells were stimulated with LPS ex vivo for 3 days and analysed by quantitative PCR (qPCR) for *Fam72a* (n = 4 mice per group). *Fam72a* (middle) and *Aicda* (right) expression assessed by qPCR from naive

splenic B cells ($B220^{+}GL\text{-}7^{-}Fas^{-}$) and germinal centre B cells (GC B; $B220^{+}GL\text{-}7^{+}Fas^{+}$) from NP_{29}-CGG immunized WT mice at day 8 after immunization ($n = 3$ mice per group). Data in **b–d** are mean ± s.e.m. representative of at least two independent experiments and were analysed by two-tailed unpaired Student's *t*-test. NS, not significant ($P > 0.05$); ****$P < 0.0001$.

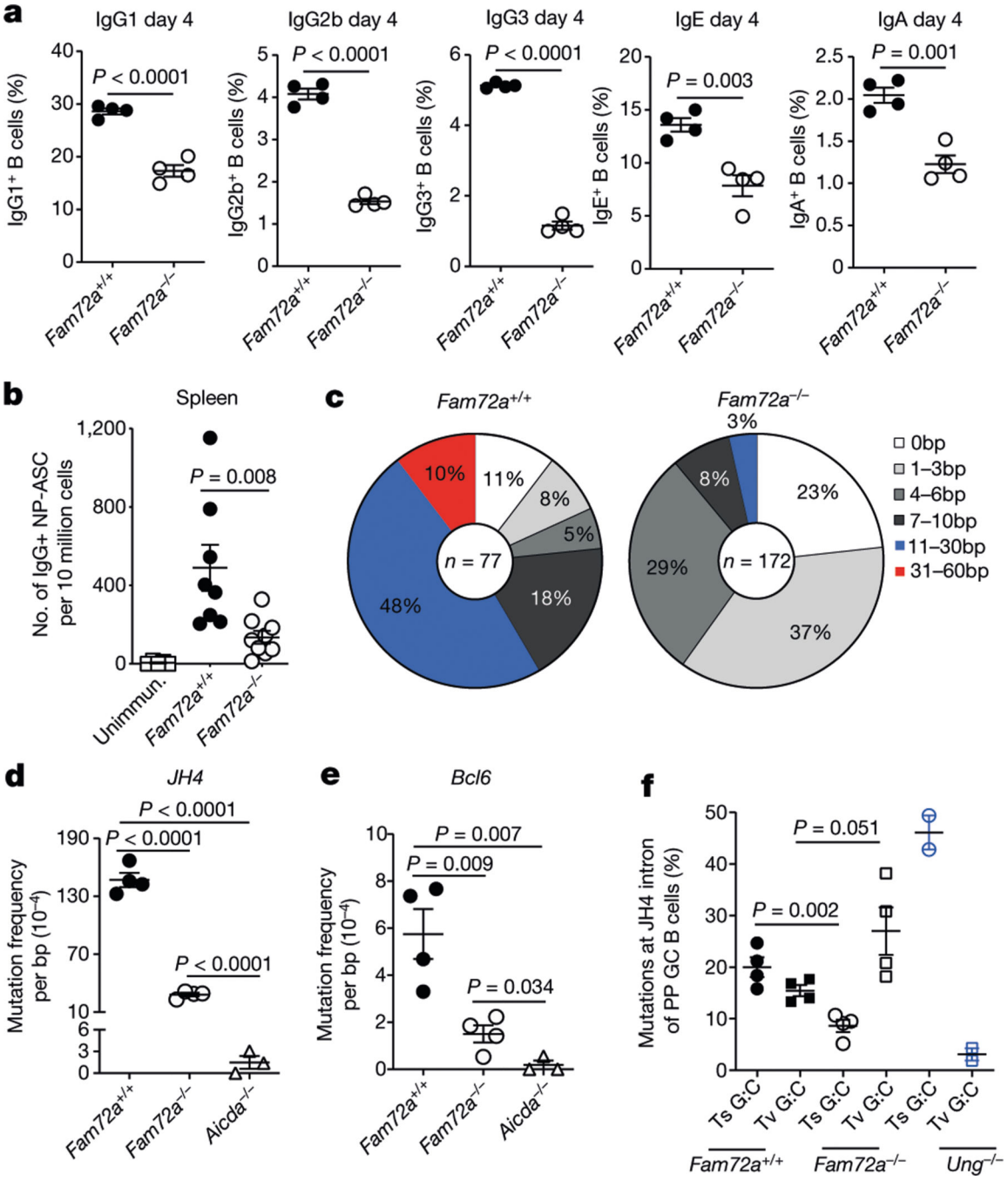


Fig. 2 |. *Fam72a*$^{-/-}$ mice exhibit defects in CSR and SHM.

a, Analysis of ex vivo CSR to indicated isotypes using splenic B cells from *Fam72a*$^{-/-}$ or *Fam72a*$^{+/+}$ littermate mice (n = 4 mice per group). Data are representative of 2 independent experiments. **b**, Number of IgG^{+} NP-specific antibody-secreting cells (NP-ASC) from mice immunized with NP-CGG (day 22 after immunization) or unimmunized control (unimmun.) (n = 5–9 mice per group; data were pooled from 2 independent experiments). **c**, Analysis of SHM profiles in the J_H4 region of germinal centre B cells from Peyer's patches of *Fam72a*$^{-/-}$ or *Fam72a*$^{+/+}$ littermate mice (n = 4 mice per group). The total number of

sequences analysed (centre of pie charts) and proportions of sequences with indicated numbers of mutations are shown. **d**, **e**, The frequency of unique mutations in the J_H4 region (**d**) and the Bcl6 gene (**e**) of germinal centre B cells per mouse ($n = 3$ mice for *Aicda*$^{-/-}$ group). **f**, Frequencies of transitional (Ts) and transversional (Tv) mutations at G:C pairs in the J_H4 region for *Fam72a*$^{-/-}$ or *Fam72a*$^{+/+}$ littermate mice ($n = 4$ mice per group). Transitional G:C mutations include G>A and C>T mutations, and transversional G:C mutations include G>C or T and C>A or G mutations. The data for *Ung*$^{-/-}$ mice ($n = 2$ mice) were obtained from ref. [7]. Data in **a**, **b**, **d–f** are mean ± s.e.m. and were analysed using two-tailed unpaired Student's *t*-test.

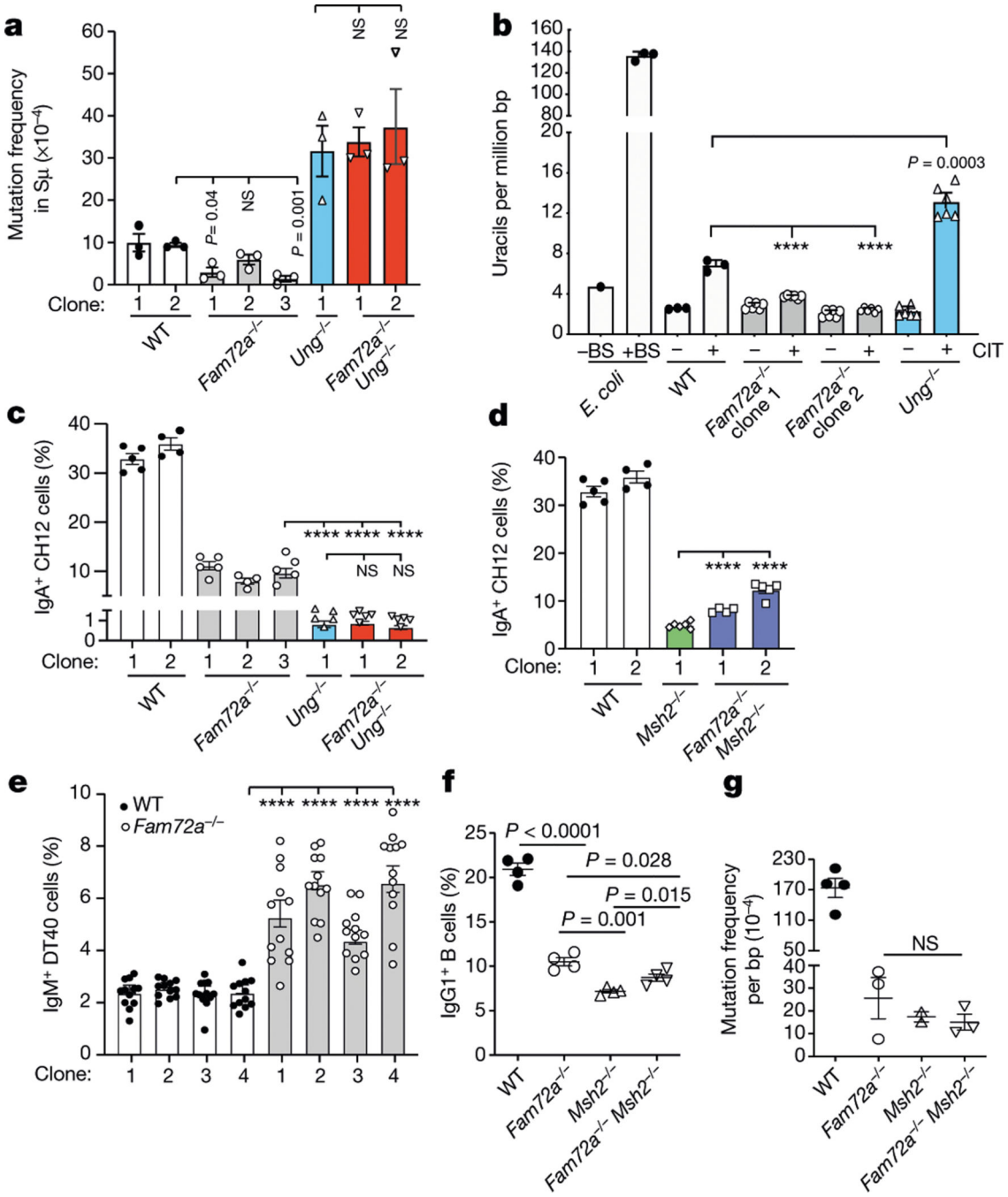


Fig. 3 |. FAM72A is epistatic with UNG and MMR during CSR.

a, Sμ mutation frequency in WT, *Fam72a*$^{-/-}$, *Ung*$^{-/-}$ and *Fam72a*$^{-/-}$*Ung*$^{-/-}$ CH12 cells at day 5 after CIT treatment. **b**, Quantification of genomic uracils in DNA isolated from WT CH12 cells, *Fam72a*$^{-/-}$ CH12 clones 1 and 2, and *Ung2*$^{-/-}$ CH12 cells. **c**, CSR in WT, *Fam72a*$^{-/-}$, *Ung*$^{-/-}$ and *Fam72a*$^{-/-}$*Ung*$^{-/-}$ clones at day 2 after CIT treatment. **d**, CSR in WT, *Msh2*$^{-/-}$ and *Fam72a*$^{-/-}$*Msh2*$^{-/-}$ CH12 clones at day 2 after CIT treatment. Experiments shown in **c**, **d**, were performed at the same time. **e**, Fluctuation analysis of gene conversion in WT and *Fam72a*$^{-/-}$ DT40 clones at two weeks of expansion. **f**,

Analysis of ex vivo CSR to IgG1 using splenic B cells from WT, *Fam72a*$^{-/-}$, *Msh2*$^{-/-}$ and *Fam72a*$^{-/-}$*Msh2*$^{-/-}$ littermate mice (n = 2 mice per group; duplicate assays per mouse). **g**, The frequency of unique mutations in the J_H4 region of germinal centre B cells from WT, *Fam72a*$^{-/-}$, *Msh2*$^{-/-}$ and *Fam72a*$^{-/-}$*Msh2*$^{-/-}$ littermate mice (n = 2–4 mice per group). Data in **a–g** are mean ± s.e.m. and were analysed using two-tailed unpaired Student's t-test. Data in **a**, **c–f** are representative of at least two independent experiments.

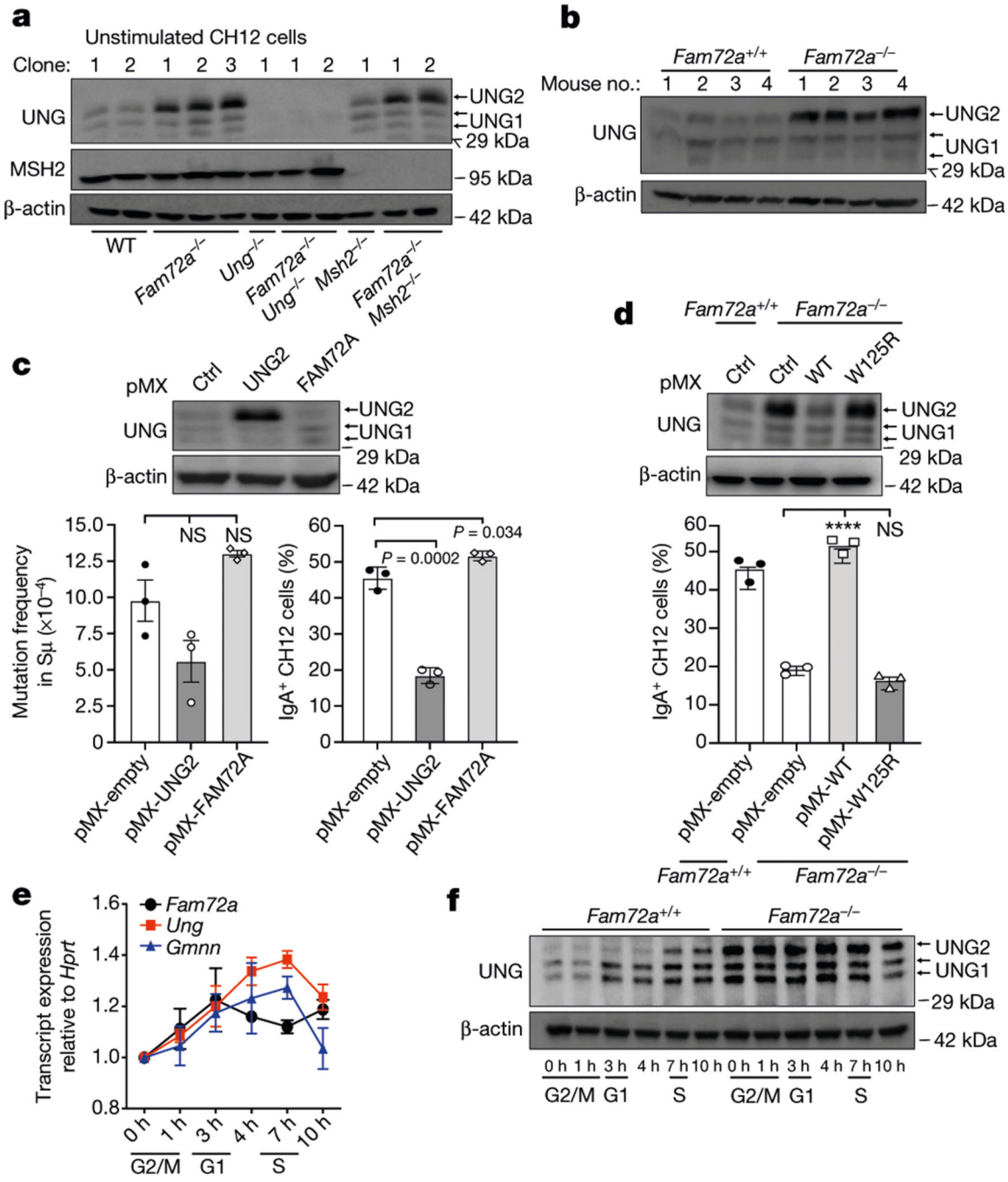


Fig. 4 |. FAM72A binds to UNG2 and leads to reduced UNG2 protein levels.

a, Western blots showing UNG and MSH2 protein in unstimulated CH12 cells of the indicated genotypes. **b**, Western blots showing UNG in primary B cells isolated from WT and *Fam72a*$^{-/-}$ mice. **c**, WT CH12 cells (clone 1) were transduced with empty pMX vector (ctrl) or vectors containing UNG2 or FAM72A cDNA. SHM in the μ switch region (bottom left), IgA CSR (bottom right) and UNG protein expression (top) were examined at day 5 after CIT treatment. **d**, IgA CSR in *Fam72*$^{-/-}$ CH12 cells (clone 1) transduced with empty pMX vector or vectors for expression of WT FAM72A or FAM72A(W125R) at day 3

after CIT treatment (bottom). Top, western blot analysis of UNG in cells of the indicated genotypes transduced with the indicated retroviruses. **e**, **f**, *Fam72a*$^{+/+}$ CH12 cells were synchronized at G2/M phase and then collected at 0 h (G2/M), 1 h (G2/M), 3 h (G1), 4 h, 7 h (S) and 10 h after RO-3306 treatment. **e**, qPCR analysis of *Fam72a*, *Ung* or *Gmnn* (which encodes geminin and accumulates at S phase) mRNA. **f**, Western blot of UNG. Data in **a–f** are representative of at least two independent experiments.