HHS Public Access

Author manuscript

Immunity. Author manuscript; available in PMC 2021 July 14.

Published in final edited form as:

Immunity. 2020 July 14; 53(1): 172–186.e6. doi:10.1016/j.immuni.2020.06.006.

Neonatal exposure to commensal bacteria-derived antigens directs polysaccharide-specific B-1 B cell repertoire development

J. Stewart New[1], **Brian L.P. Dizon**[1,2], **Christopher Fucile**[3], **Alexander F. Rosenberg**[1,3], **John F. Kearney**[1,*], **R. Glenn King**[1,*]

[1]Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL 35294 United States of America

[2]National Institutes of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892 United States of America

[3]Informatics Institute, University of Alabama at Birmingham, Birmingham, AL 35294 United States of America

Summary

B-1 B cells derive from a developmental program distinct from that of conventional B cells, through B cell receptor (BCR)-dependent positive selection of fetally-derived precursors. Here, we used direct labeling of B cells reactive with the N-acetyl-D-glucosamine (GlcNAc)-containing Lancefield Group A Carbohydrate of *Streptococcus pyogenes* to study the effects of bacterial antigens on the emergent B-1 B cell clonal repertoire. The number, phenotype and BCR clonotypes of GlcNAc-reactive B-1 B cells were modulated by neonatal exposure to heat-killed *S. pyogenes* bacteria. GlcNAc-reactive B-1 clonotypes and serum antibodies were reduced in germ-free mice compared to conventionally raised mice. Colonization of germ-free mice with a conventional microbiota promoted GlcNAc-reactive B-1 B cell development and concomitantly elicited clonally related IgA^+ plasma cells in the small intestine. Thus, exposure to microbial antigens in early life determines the clonality of the mature B-1 B cell repertoire and ensuing antibody responses, with implications for vaccination approaches and schedules.

Graphical Abstract

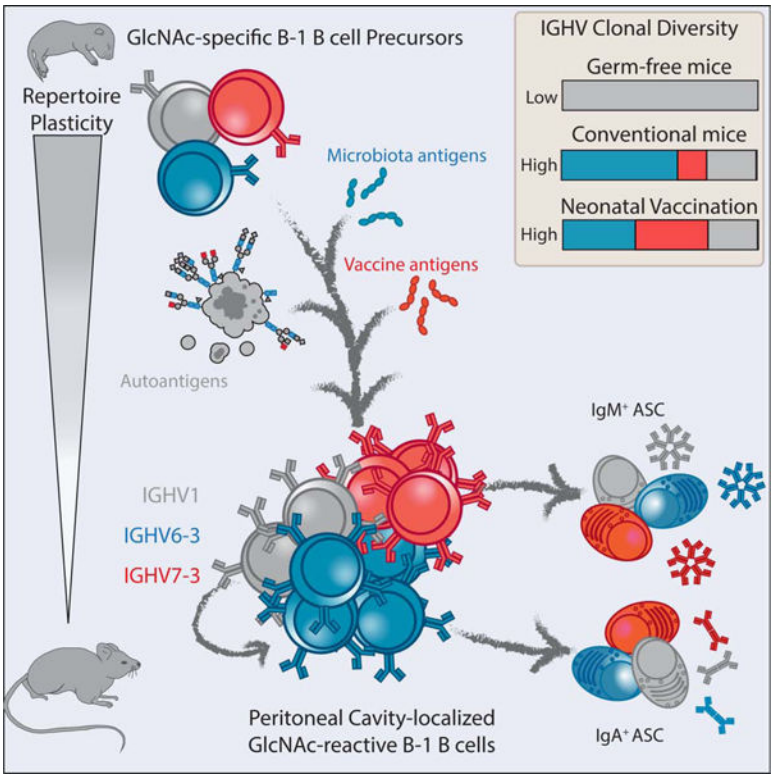
*Corresponding Authors: 1825 University Blvd., Shelby Biomedical Research Building Rm401, Birmingham, AL 35294-2812, jfk@uab.edu, rgking@uab.edu.
Lead Contact: John F. Kearney

Author Contributions

All authors contributed to the design of experiments and participated in the interpretation of experimental results. JSN, BLPD and RGK performed experiments and JFK purified antibodies. AFR and CF developed bioinformatics tools for IG Rep-Seq data analysis and assisted in the analysis and interpretation of sequencing data.

Declaration of Interests

The authors declare no competing interests.



eTOC blurb

Innate-like B-1 B cells express canonical B cell receptors and originate through specialized developmental programs. Through immunoglobulin gene sequencing of single antigen-reactive B cells, New et al. demonstrate a significant role for environmental antigen-dependent selection events in shaping the GAC-reactive B-1 B cell clonal repertoire.

Keywords

Natural Antibodies; B-1 B cells; neonatal immunity; *Streptococcus pyogenes*; N-acetyl-D-glucosamine; microbiota

Introduction

B-1 B lymphocytes are a specialized innate-like subset of B cells in mice with unique functional properties owing to their anatomical localization, B cell receptor expression and cell surface phenotype (Kantor and Herzenberg, 1993). B-1 B cells are derived through the BCR-dependent positive selection of precursors that predominate during fetal and neonatal life and differ from those for conventional B cells (Hayakawa et al., 1999; Montecino-Rodriguez et al., 2006). The B-1 repertoire is enriched in canonical B cell clonotypes which serve essential and non-redundant roles in the context of host immunity through their secreted antibodies (Alugupalli et al., 2004; Baumgarth et al., 2000; Cole et al., 2009; Haas et al., 2005). B-1 B cells additionally contribute to the constitutive production of natural serum antibodies (Choi et al., 2012; Kawahara et al., 2003), which not only provide a front-

line defense against infectious pathogens (Baumgarth et al., 2000; Briles et al., 1981; Jayasekera et al., 2007; Ochsenbein et al., 1999; Rapaka et al., 2010), but also serve to maintain tissue homeostasis through the recognition of conserved antigenic epitopes, such as neoself-epitopes presented on apoptotic cells (Binder et al., 2003; Brandlein et al., 2003; Kearney et al., 2015; Vas et al., 2012). The production of natural IgM by B-1 B cells does not require stimulation by exogenous antigens (Bos et al., 1987; Chou et al., 2009; Haury et al., 1997; Hooijkaas et al., 1984; Ochsenbein et al., 1999; Sigal et al., 1975). However, there is evidence that microbial antigens influence the antigen specificity and clonal content of the pre-immune IgM repertoire (Bos et al., 1989; Kearney et al., 1985; Yilmaz et al., 2014). Despite these indications, direct examination of the phenotypic and clonotypic characteristics of antigen specific B cells at the single cell level are limited, in part due to challenges in identifying such rare cells. In absence of such studies, our understanding of the role environmental antigens play during early B cell ontogeny is incomplete.

Streptococcus pyogenes, commonly known as Group A Streptococcus (GAS), is a human pathogen and a significant global health burden, for which there is a high demand for protective vaccines (Carapetis et al., 2005). The immunodominant N-acetyl-D-glucosamine (GlcNAc) side chains of the Lancefield Group A carbohydrate (GAC), the group-specific component of the cell wall polysaccharide of GAS (McCarty, 1952, 1956), elicit clonally restricted T-independent antibody responses in mice (Lutz et al., 1987). Human GlcNAc-reactive B cell responses are similarly clonally restricted and multiple public idiotypes associated with GAC-reactive BCR have been described (Emmrich et al., 1983; Zenke et al., 1984). While GAC-based immunization strategies confer broadly protective antibody-based immunity against multiple GAS serotypes (Kabanova et al., 2010; Sabharwal et al., 2006), GlcNAc-reactive B cells may additionally modulate immune responses to numerous organisms with GlcNAc-containing cell walls, as we have demonstrated their potential to suppress allergic sensitization of mice to GlcNAc-bearing *Aspergillus fumigatus* (Cywes-Bentley et al., 2013; Kin et al., 2012). However, GAC-reactive antibodies may also recognize autologous GlcNAc epitopes, including apoptosis-associated neoantigens presented on senescent cells and post-translational modifications on the surface of human neuronal cells, that latter of which has been implicated in neuropsychiatric pathological sequelae associated with GAS infection (Kirvan et al., 2003; New et al., 2016). Thus, the factors influencing the establishment of the GlcNAc-reactive B cell clonal repertoire have bearing on safe and efficacious GAS vaccine design as well as the development of therapeutic strategies to treat allergic and autoimmune disease.

Here we examined the impact of exposure to vaccine- and microbiota-derived antigens during the ontogenetic establishment of the B-1 B cell compartment on the clonal repertoire to antigen specific B cells. We used direct visualization of GlcNAc-reactive B cells labeled with fluorescent GAC to investigate the development and distribution of GlcNAc-reactive cells in mice. Single-cell analysis of immunoglobulin heavy variable (IGHV) gene expression in GAC-reactive mouse B cells revealed plasticity in the nascent B-1 B cell repertoire. Neonatal exposure to GAS induced significant and long-lasting phenotypic and clonotypic alterations to the GlcNAc-reactive B cell repertoire. Natural GlcNAc-reactive B-1 B cell development was dependent on microbiota-derived antigens in addition to MyD88 centered pathways, and concomitantly seeded mucosal tissues with B-1-derived IgA-

producing B cells. Collectively, these observations highlight an essential function for microbial antigens in driving the selection and expansion of GlcNAc-reactive B-1 B cell clonotypes as well as production of IgM and indicate that the natural antibody repertoire is fine-tuned by exposure to environmental and microbiota-derived antigens during early life.

Results

Neonatal immunization with Group A Streptococcus expands the GlcNAc-reactive B-1 B cell population and increases resting GlcNAc-reactive serum IgM levels.

To characterize the emergence of N-acetyl-D-glucosamine (GlcNAc)-reactive humoral immunity in mice, we first evaluated serum antibody responses following immunization of C57BL/6 mice with GlcNAc-bearing Group A Streptococcus (GAS, strain J17A4) at various neonatal time-points. Immunization between 14 days-of-age (d14) and d28 resulted in modest elevations in Group A Carbohydrate (GAC)-reactive serum IgM relative to unimmunized eight-week-old adult control mice, whereas mice immunized prior to d14 exhibited no increases in GAC-reactive IgM (Figure S1A). The kinetics of neonatal GlcNAc-reactive IgM responses were delayed and attenuated in magnitude relative to those mounted by adult (d56) mice (Figure 1A). However, ELISpot analysis revealed that adult d14 GAS-immunized mice maintained increased numbers of splenic GAC-reactive IgM^+ antibody-secreting cells (ASCs) into adulthood, compared to naïve controls, (Figure 1B).

We examined GAC-binding (GAC^+) B cells by single-epitope multiple-staining (Townsend et al., 2001) using fluorescently labeled preparations of Lancefield Group A Carbohydrate. GAC^+ B cells were present in the spleen of neonatal mice as early as three-days-of age and, after an initial decline in frequency, increased in abundance until they reached ~0.04% of total $CD19^+$ splenocytes in adult mice (Figure 1C). GAC^+ B cells were enriched in the peritoneal cavity (PerC) of adult mice, constituting 0.25±0.05% of total $CD19^+$ B cells, but were infrequent in the inguinal lymph node (ILN) (Figure 1C–E, Figure S1B). Splenic and PerC GAC^+ B cells were enriched in the B-1 B cell compartment, with most exhibiting a $CD5^-CD43^+CD23^-B220^{lo}$ B-1b B cell phenotype, and a minor population (~20%) exhibiting a $CD5^+CD43^+CD23^-B220^{lo}$ B-1a B cell phenotype (Figure 1D–F, Figures S1C–E). Adult mice immunized with GAS at d14 exhibited an ~4-fold increase in numbers of PerC-localized GAC^+ B-1 B cells over naïve controls, with increased relative representation of CD5-expressing GAC^+ B cells. CD11b (MAC-1) expressing GAC^+ B cells were also more abundant in adult mice immunized with GAS at d14 (Figure 1G). In contrast, numbers of splenic and PerC GAC^+ B-2 B cells were unchanged following d14 immunization (Figure 1 D–F, Figure S1C–G). GAC-reactive serum Ig in 10–12-month-old mice, immunized with GAS at d14, had normalized to levels similar to those of age-matched naïve mice, as had the numbers and phenotypic profiles of GAC^+ B cells present in the spleen and PerC (Figure S1H–K).

We used the monoclonal antibody 1A.1, specific for the IdI-3a idiotope expressed by a subset of A/J mouse-derived GAC-specific hybridomas expressing immunoglobulin heavy chain variable gene 6 family heavy chains (IGHV6, formerly referred to as J606) (Greenspan and Davie, 1985), to probe variations in the GAC-reactive mouse B cell repertoire. Approximately 20% of splenic GAC^+ B cells expressed the IdI-3a idiotope in

naïve eight-week-old C57BL/6 mice and GAS immunization increased the numbers of IdI-3a$^+$ plasmablasts relative to naïve controls, confirming that C57BL/6 mice mount IdI-3a$^+$ B cell responses following GAS immunization (Figure 1H, Figure S1L). However, the frequency of IdI-3a expressing GAC$^+$ B cells was reduced in adult mice following immunization at d14 compared to age-matched naïve mice (Figure 1H). Thus, neonatal immunization with GAS led to increased numbers of GAC$^+$ B-1 B cells in adult mice, that differed from those of naive with respect to their expression of CD5 and the GAC-reactive idiotype IdI3a.

Neonatal immunization with Group A Streptococci results in altered clonal dominance in the GlcNAc-reactive B cell pool.

To investigate the molecular changes underpinning the alterations in IdI-3a idiotype exhibited by GAC-reactive B cells as a result of neonatal exposure to GAS, we sequenced the immunoglobulin heavy chain (IGH) genes of FACS-sorted single GAC$^+$ B cells isolated from naïve, adult- or neonatally-immunized C57BL/6 mice. Splenic GAC$^+$ B cells in naïve adult mice were pauciclonal and dominated by IGHV6–3 expressing B cell clonotypes (Figure 2A, a detailed description of the Circos plot depictions of IGHV sequence data is provided in the STAR Methods section). These findings are consistent with previously characterized GAC-reactive hybridomas (Lutz et al., 1987; Nahm et al., 1982), and multiple public IGH sequences (found in multiple animals) represented in our data set shared sequence homology with GAC-reactive hybridomas generated >30 years ago (Figure S2A) (Phillips and Davie, 1990). Aside from IGHV6–3, we identified a minor population of GAC$^+$ B cells expressing IGHV7–3, as well as other heterogeneous and low-frequency IGHV genes (Figure 2A).

Adult, d56 mice, responded vigorously to GAS immunization (Figure 1A), and GAC$^+$ B cells isolated from mice at the peak of the adult antibody response, 14 days-post-immunization (d.p.i.), were enriched for IGHV6–3 expressing clones compared to naïve animals. GAC$^+$ B cells from mice immunized as adults exhibited expansions of B cells expressing IGHV6–3 encoded BCRs, comprised of IGHD4 or lacking IGHD genes, and predominantly IGHJ3 and IGHJ4 genes (Figure 2B, Figure S2C). Relative to those of naïve mice, pooled GAC$^+$ B cell IGH sequences isolated from d56 immunized mice exhibited reduced clonal diversity (Figure S2E).

In contrast, 8–10-week-old d14 GAS-immunized mice exhibited >five-fold increases in the frequency of IGHV7–3 expressing GAC$^+$ B cells relative to naïve controls (Figure 2C). IGVH7–3 sequences were predominantly germline, utilized almost exclusively IGHD1–1 (DFL16.1), and encoded CDRH3 gene segments that were approximately 50% longer than those of IGHV6–3 receptors (Figure 2C–D). Interestingly, although mice immunized with GAS at d3 did not exhibit increased serum concentrations of GAC-reactive IgM, they shared the enrichment for IGHV7–3 B cell clonotypes (Figure S2B–C). Estimates of clonal diversity in the GAC$^+$ B cell compartment of d14 GAS-immunized mice were intermediate to naïve and d56-immunized mice (Figure S2E). Similar to the expansion of GAC$^+$ B cells, the altered frequencies of IGHV7–3 GAC$^+$ B cells observed in d14 mice waned with time

and had largely normalized between naïve and d14 GAS-immunized mice by 10–12 months-of-age (Figure S2D).

Pockets of polar amino acids flanked by hydrophobic residues were common to the CDRH3 of both IGHV6–3 and IGHV7–3 encoded receptors (Figure 2E); however, IGHV7–3 CDRH3 sequences carried less negative charge than those of IGHV6–3 sequences due to V gene-encoded differences in the 5' junction, and contained multiple IGHD1–1-encoded polar tyrosines (Figure S2H). By contrast, GAC^+ B cell expressing IGHV1 BCR contained CDRH3 enriched for aliphatic residues (Figure S2H). Although N additions were common in both IGHV6–3 and IGHV7–3 sequences (Figure S2I), they were dispensable for the formation of GAC-reactive BCR, as mice deficient in terminal deoxynucleotidyl transferase (TdT) exhibited normal concentrations of GAC-reactive IgM in sera (data not shown).

We next investigated whether IGHV7–3-expressing GAC^+ B cells expanded by d14 GAS-immunization contributed to subsequent immune responses. We examined the memory recall response in mice immunized at d14 and boosted at eight weeks of age, relative to d56-primed mice that received a booster immunization 8 weeks later, after contraction of the primary GAC^+ B cell expansion (Figure S3A, B). Both groups of mice exhibited significant increases in the numbers of GAC^+ ASCs five days following administration of the booster immunization, although mice primed as adults generated more ASCs (Figure S3C). The frequency of GAC^+ B cells was increased 10-fold in both groups of mice, >50% of which expressed the plasmablast marker CD138 (Figure S3D–E). Like the primary response, the secondary response in d56 primed mice resulted in the expansion of IGHV6–3 clonotypes. By contrast, the memory-recall response in mice primed as d14 neonates consisted of a significant proportion of IGHV7–3 clonotypes (Figure 2F. Figure S3F). The frequency of somatic mutations in IGHV6–3 genes increased after secondary immunization, in contrast to neonatal prime-boost where lower level of mutation in IGHV6–3 sequences were observed. Unlike IGHV6–3 sequences, the germline configurations of the IGHV7–3 sequences were maintained following immunization (Figure S3G).

We hypothesized that the expansion of germline IGHV7–3-expressing B cells following d14 immunization with GAS could be explained mechanistically by a different relative abundance of IGHV7–3 clonal precursors during neonatal life. To address this possibility, we examined the emergence of GAC^+ B cells between 14 and 28 days of age. Splenic GAC^+ B cells in 14-day-old neonates exhibited predominantly immature $AA4.1^+IgM^+$ B cell phenotypes, and between 21–28 days of age the frequency of GAC^+ B cells increased in neonates as did the ratio of mature $AA4.1^-$ to immature $AA4.1^+$ GAC^+ B cells (Figure S3H–K). When we examined IGHV gene expression by these emergent neonatal GAC^+ B cells, we found GAC^+ B cells in 14-day-old mice exhibited heterogenous IGHV utilization characterized by a preponderance of IGHV1- and other IGHV- gene-family expressing cells (Figure 2H). IGHV heterogeneity decreased with age as IGHV6–3-expressing clonotypes became progressively enriched in the adult repertoire. IGHV7–3 expressing B cells, although detected at low frequency in neonatal and adult mice, were ultimately diluted by the apparent expansion of IGHV6–3-expressing clones in aged mice (Figure 2H). Annotation of indexed FACS sort profiles with IGHV gene utilization revealed that the mature $AA4.1^-GAC^+$ B cell compartment in d28 mice was enriched for IGHV6–3

clonotypes relative to the AA4.1$^+$GAC$^+$ B cell compartment, with lower frequencies of IGHV1- and other IGHV-family expressing clones (Figure 2H, Figure S3L). Thus, early exposure to GAS antigen expands distinct and normally low-frequency IGHV7–3 GAC-reactive B cell clonotypes promoting their subsequent conservation in the adult B cell repertoire.

Microbial colonization is required for GlcNAc-reactive B-1 B cell development.

The clonal plasticity of emergent GAC-reactive B cells revealed by neonatal immunization led us to investigate whether microbiota-derived antigens influenced the establishment of the GlcNAc-reactive B cell clonal repertoire. Initially, we examined serum from germ-free (GF) C57BL/6 mice which, as previously reported, exhibited reduced class-switched IgA and IgG antibodies but contained IgM at concentrations similar to that of specific-pathogen free (SPF) C57BL/6 mice (Figure S4A) (Bos et al., 1989; Haury et al., 1997; Hooijkaas et al., 1984). Despite quantitative similarities in total IgM, GAC-reactive serum IgM antibody concentrations were reduced 10-fold in GF mice compared to SPF mice (Figure 3A). By contrast, levels of natural Abs reactive with phosphorylcholine (PC) were unchanged (Figure 3A), confirming prior findings (Chou et al., 2009; Etlinger and Heusser, 1986; Sigal et al., 1975). ELISpot analysis confirmed that GAC-reactive ASCs were reduced in both spleen and BM of GF mice (Figure 3B).

GF mice exhibited no apparent abnormalities in spleen, PerC, and ILN B cell subsets (data not shown), yet harbored >10-fold fewer splenic and PerC GAC$^+$ B cells relative to SPF mice (Figure 3C and Figure S4B). Additionally, splenic GAC$^+$ B cells in GF mice exhibited a naïve B-2 B cell phenotype, characterized by high CD45R (B220), IgD, and CD23 expression, and reduced expression of CD43, CD5, and CD36 relative to those of SPF mice (Figure 3C and Figure S4C). IGHV6–3-expressing GAC$^+$ B cells were rare in GF mice, which contained IGHV1 clonotypes that were infrequent and poorly conserved in the repertoires of SPF mice, unlike IGHV6 and IGHV7 encoded GAC-reactive receptors (Figure 3D, Figure S4D–G). DH reading-frame utilization in GAC$^+$ B cells from GF mice was also more heterogeneous than SPF-derived clonotypes, perhaps indicative of a loss of selective pressure on these cells (Figure S4H). Importantly, GlcNAc-reactive Ab responses following intraperitoneal immunization with GAS were blunted in GF mice, in contrast to SPF mice (Figure 3E, Figure S4I–J). Together, these data indicate that development of the GAC-reactive B-1 B cells and GlcNAc reactive antibody production relies on stimulation by the intestinal microbiota, underscoring a critical function for microbiota-derived antigens in selection of B cell specificities that comprise the natural pre-immune repertoire of mice.

Colonization of adult GF mice with conventional microbiota reconstitutes GlcNAc-reactive B-1 B cells

B-1 B cell development in mice occurs predominantly during early life, however B-1 progenitors have been identified in adults (Ghosn et al., 2011; Yoshimoto et al., 2011). We therefore investigated whether timing of colonization by the microbiota influences GlcNAc-reactive B-1 B cell repertoire formation. To this end, we co-housed adult GF mice with SPF mice for eight weeks to generate conventionalized GF (GF→SPF) mice. Deep sequencing of 16S rDNA isolated from the small intestinal microbiota confirmed that the species content

population diversity of GF→SPF was highly similar to that of SPF animals (Figure S5A–C). They also exhibited normal intestinal morphology and normal IgA concentrations in serum and fecal content (Figure S5D–F). We found serum GAC-reactive IgM antibody concentrations as well as numbers of spleen and BM-localized GAC-reactive ASCs were fully restored in GF→SPF mice (Figure 4A, Figure S5G). Titers of GAC-reactive serum IgA were elevated in GF→SPF mice, superseding those of SPF mice, and expanded populations of splenic GAC^+ plasmablasts were readily detected in GF→SPF (Figure 4A, B).

Numbers of PerC-localized GAC^+ B cells increased >58-fold in GF→SPF mice relative to age-matched GF controls and were comprised of $CD5^+$ B-1a and $CD5^-$ B-1b B cells (Figure 4D, Figure S5H–K). GAC^+ B cells of GF→SPF mice expressed IGHV6–3 genes at frequencies similar to those of SPF animals while a small number of IGHV7–3 GAC^+ B cells were identified in the spleen and PerC of GF→SPF animals (Figure 4E, Figure S5L). Clonal diversity of GAC^+ B cells in GF→SPF GAC^+ B cells mice was similar to that of SPF mice and greatly increased over GF mice (Figure S5M), however, GAC^+ B cells bearing mutated IGH6–3 and IGH1-family BCRs were more frequent than in SPF mice which had been colonized from birth. (Figure 4E, Figure S5N). Nonetheless, GF→SPF mice exhibited a high degree of conservation in the GAC^+ BCRs between the spleen and PerC compartments within individual mice, as well between different mice (Figure 4F). Microbial colonization therefore drives the expansion of IGHV6–3 expressing GAC^+ B cells in SPF mice, and the generation of B-1 B cell clonotypes responsive to GlcNAc bearing antigens.

Microbiota-dependent development of GAC-reactive B cells requires innate-immune signaling, but not T cell help

We explored signaling pathways contributing to microbiota-dependent establishment of the GAC-reactive B cell compartment by examining GAC^+ B cells in mouse models deficient in key immune-associated genes. Although anti-GAC responses are reportedly largely T cell independent (Briles et al., 1982), the somatic mutations we observed in GAC^+ IGHV genes led us to investigate the contribution of T lymphocytes to the development of the GAC-reactive B cell repertoire. We observed only minor deficits in GAC^+ B cell numbers and serum IgM in T cell receptor beta chain and delta chain deficient mice ($Tcrb^{-/-}Tcrd^{-/-}$) confirming that T cell help was not required (Figure 5A–C).

We next examined MyD88-deficient ($MYD88^{-/-}$) mice (Hou et al., 2008), which lack the adaptor protein essential for signaling through all innate toll-like receptors (TLR) but TLR3. Numbers of GAC^+ B cells were lower in the spleens and markedly reduced in the PerC of $MYD88^{-/-}$ mice, which also exhibited a deficiency in GAC-reactive serum IgM relative to SPF mice (Figure 5A–C). Similar to GlcNAc-reactive IgM, PC-reactive IgM was modestly decreased in $MYD88^{-/-}$ mice (Figure 5D). We therefore examined mice deficient in toll-like receptor 9 (TLR9), a sensor of unmethylated microbial CpG motifs, previously implicated in the development of B-1b B cells and production of immune-modulating self-reactive IgM in lupus-prone mice (Stoehr et al., 2011). TLR9-deficiency, however, did not recapitulate the GAC-reactive B cell deficiency observed in $MYD88^{-/-}$ mice, suggesting innate sensing of bacterial DNA was dispensable for GAC^+ B-1 B cell development. PC-reactive IgM levels

were, however, unexpectedly high in TLR9-deficient mice, indicating that TLR signaling pathways may differentially regulate innate-like B cell responses.

Nucleotide-binding oligomerization domain (NOD) are cytosolic sensors of peptidoglycan expressed by B cells and other phagocytes that can promote the survival of ligand-experienced immature B cells during B cell development (Hayakawa et al., 2017; Strober et al., 2006). As GlcNAc is a core carbohydrate of many peptidoglycan structures (Vollmer et al., 2008), we hypothesized that mice with defective NOD signaling would exhibit deficiencies in GAC-reactive B cells. This was not the case however, as GAC^+ B-1 B cell numbers were unaffected in *NOD2*$^{-/-}$ mice (Kobayashi et al., 2005). In fact, GAC-reactive serum IgM was increased in *NOD2*$^{-/-}$ mice relative to C57BL/6 animals (Figure 5A–C). We further examined mice deficient in Receptor Interacting Serine/Threonine Kinase 2 (RIP-2), the downstream signal transducer for NOD1 and NOD2 (Kobayashi et al., 2002). *RIPK2*$^{-/-}$ mice exhibited similar increases in GAC-reactive serum IgM, while GAC^+ B cell numbers did not differ significantly from WT mice (Figure 5A–C). The establishment of GAC^+ B cells therefore appears augmented by Myd88-signaling mediated by TLR-dependent signaling, while the secretion of IgM by GAC^+ ASCs is further modulated by innate pathogen sensing pathways, including the NOD-family/RIP-2 signaling axis.

Intestinal microbiota drives GAC-reactive B cell recruitment to the small intestinal Lamina Propria

We next investigated the outcome of microbial colonization on mucosal GAC^+ B cell responses, as B-1 B cells contribute to the production of secretory IgA at the intestinal mucosa (Bunker et al., 2015; Kroese et al., 1989; Macpherson et al., 2000). Initially, we examined supernatants from solubilized preparations of intestinal contents from WT C57BL/6 mice for GlcNAc-reactive IgA. Similar to previous reports (Bunker et al., 2015), lumenal IgA was abundant in the small intestine (SI), but was comparatively scarce in the cecum and colon (Figure 6A). IgA was not detected in *Rag*$^{-/-}$ mice that lack adaptive immunity and was significantly reduced in both *Tcrb*$^{-/-}$*Tcrd*$^{-/-}$ and *MYD88*$^{-/-}$ mice (Figure 6A). We detected robust binding of lumenal IgA in the jejunum and ileum of the SI in WT mice to GAC and $GlcNAc_{35}BSA$ antigens while, like total IgA, GlcNAc-reactive IgA was significantly decreased in both *MYD88*$^{-/-}$ mice and *Tcrb*$^{-/-}$*Tcrd*$^{-/-}$ mice (Figure 6B).

We used $Blimp1^{YFP}$ mice, which harbor a yellow fluorescent protein (YFP) reporter gene within one *Prdm1* (Blimp1) gene locus, to assess the distribution of IgA plasma cells in the mouse gut. IgA^+ B cells were readily detected in both the spleen and PerC of $Blimp1^{+/YFP}$ mice, although surface IgA^+ B cells were a small fraction of splenic plasma cells, and the PerC was devoid of YFP^{Hi} plasma cells (Figure S6A). IgA^+ B cells were clearly enriched in mesenteric lymph nodes, Peyer's patches (PP), small intestinal lamina propria (SI LP), and cecal and colonic tissue digests relative to the spleen and PerC, yet IgA^+ YFP-expressing plasma cells were uniquely abundant in the SI LP (Figure S6B–C). GAC^+ B cell populations detected in gut associated-lymphoid tissues (GALT) and gut tissue-digests were less defined by single epitope, multiple-staining detection than those of spleen and PerC, but were preferentially enriched in SI LP and PP where they represented 1.21±1.04% and 0.19±0.03% of the total B cell compartment, respectively (Figure 6C–D, Figure S6E, F). SI

LP GAC$^+$ B cells were predominantly surface IgA$^+$, CD138$^+$ and YFP$^+$ cells, affirming their plasma cell identity (Figure S6G–I).

To confirm that the SI LP was enriched for GAC-reactive plasma cells, we plated immune cells from systemic lymphoid compartments, GALT, and gut tissue digests for ELISpot analysis. We detected numerous GAC-reactive IgA$^+$ ASCs in SI LP tissue digests, which comprised 0.34±0.46% of total CD19$^+$ B cells plated (Figure 6E, F). GAC$^+$ B cells were more sparsely distributed in PP, cecum and colon tissue, however we detected GAC$^+$ IgA$^+$ ASCs in the spleen and MLN of SPF C57BL/6 mice, though they represented a fairly low-frequency of total CD19$^+$ cells relative to the SI LP (Figure 6E, F). Conventionally raised C57BL/6 mice therefore harbor high frequencies of SI LP-localized GAC$^+$ IgA$^+$ B cells that secrete GlcNAc-specific IgA into the intestinal lumen at homeostasis.

GAC$^+$ Plasma cells of the SI LP share IGH sequence homology with B-1 B cell clonotypes

To further investigate the microbiota-dependent localization of GAC$^+$ B cells to the SI LP, we performed deep sequencing of IGH genes amplified from whole SI LP cDNA libraries prepared from SPF, GF and GF → SPF mice. We obtained tens of thousands of IgA (Range 30,000–130,000, mean=64,866) and IgM sequences (Range 106–111,000, mean=35,646) from most SI LP tissue preparations, and excluded samples with sequence yields outside of these ranges from our analysis (Figure S7A, B). The SI LP B cell repertoire of all mouse groups was comprised of diverse IGHV:IGHJ pairings and multiple IGHV:IGHJ pairings were found to be differentially expressed in SPF or GF→SPF mice relative to GF mice; IGHV:IGHJ gene pairing frequencies were comparatively less-divergent between SPF and GF→SPF mice (Figure S7B,C). Nonetheless, GF mice possessed a diverse SI LP B cell repertoire, and Hill diversity estimates of IgA and IgM were equivalent across mice regardless of their colonization status (Figure S7D). Whereas IgM sequences were largely unmutated, the IgA repertoire exhibited IGHV gene mutations, the degree of which was reduced in GF mice relative to microbiota colonized SPF and GF→SPF animals (Figure S7E).

Grouping IGH sequences into related lineages using previously described strategies revealed that microbial colonization did not significantly alter the distribution of clonal lineage size in the SI LP B cell repertoire (Figure 7A) (Nogales et al., 2018). Determination of the Morisitta-Horn overlap revealed little conservation of SI LP B cell clonotypes between individual mice, however low levels of overlap were observed between the repertoires of SPF and GF→SPF mice that was absent in GF mice (Figure 7B) (Horn, 1966; Morisita, 1959). Visualization of these connections in alluvial plots revealed shared clonotypes in SPF and GF→SPF mice that were often of relatively high-frequency in one mouse but were rarely of high-frequency in both mice (Figure 7C, Figure S7F). Conversely, GF mice exhibited fewer shared clonotypes with the SI LP IGH repertoires of microbiota colonized mice as well as with one another, and similarities to highly expanded clonotypes of GF mice were infrequent; while ~4% of all clonal lineages in our dataset were shared between two or more SPF and GF→SPF, ≤1.0% of clonotypes present in either of these groups were observed in the repertoires of GF mice (Figure 7C, Figure S7F).

We included 570 unique CDRH3 sequences, derived from a reference pool of ~850 FACS-sorted single GAC$^+$ mouse B cell IGH sequences, in our nucleotide-based clustering analysis to identify lineages of GAC$^+$ B cells expressing potentially GlcNAc-reactive IGHV genes in the SI LP. Through this strategy we identified 74 distinct SI LP IGHV clonotypes which clustered with GAC$^+$ B cell derived IGHV sequences, including (5/10) of the most frequent public GAC$^+$ B cell clonotypes which were, in most cases, identical to our CDRH3 query sequences. SI LP lineages bearing homology to GAC$^+$ B cell-derived IGHV sequences were comprised of IGHV6–3, IGHV7–3 and IGHV1 expressing clonotypes, and one or more of these lineages were identified in 6/7 SPF and 9/9 GF→SPF mice. However, no predicted GAC$^+$ B cell clonotypes were identified in the SI LP B cell repertoires of GF mice (0/3) (Figure 7D). A similar analysis, utilizing CDRH3 amino acid sequences derived from previously characterized GAC-reactive hybridomas (Phillips and Davie, 1990), revealed 11 lineages sharing homology with (6/12) hybridoma CDRH3 query sequences in 3/7 SPF and 4/9 GF→SPF mice but were undetected in GF mice (Figure 7E). In some cases, these clonotypes constituted large B cell lineages and, notably, one SPF mouse (SPF Mouse 6) possessed multiple highly expanded GAC$^+$ B cell lineages (Figure 7F, Figure S7G).

We used protein parsimony (Phylip) to construct phylogenetic trees and visualize diversification of expanded GAC$^+$ BCR within individual clonal lineages (Figure 7G) (Felsenstein, 1989, 2005). Although we cannot determine if all sequences represented in these lineages are the progeny of a single parent cell, the resulting phylogenetic analysis provides clear evidence of the accumulation of somatic mutations in GAC$^+$ BCR genes in response to microbiota-derived antigens. In some cases, such as lineages 3970 and 2496 derived from SPF Mouse 6, phylogenetic trees were comprised of both IgM and IgA isotypes, suggesting IgA class-switch recombination occurred *in situ* following recruitment of IgM$^+$ GAC$^+$ precursors (Figure 7G). Collectively, these data indicate that colonization by the microbiota elicits mucosal immune responses that seed the PerC with GAC$^+$ B-1 B cells and the SI LP with clonally related IgA$^+$ ASCs.

Discussion

Numerous reports have demonstrated that the establishment of natural IgM-producing antibody secreting cells is largely independent of exogenous antigens (Bos et al., 1987; Haury et al., 1997; Hooijkaas et al., 1984). Indeed certain Nabs, such as TEPC15$^+$ NAbs reactive with the phospholipid PC, are effectively generated under GF conditions (Chou et al., 2009; Ochsenbein et al., 1999; Sigal et al., 1975). Moreover, microbial colonization is not required for the establishment of B-1-like BCR in the IGH gene repertoires of mice (Yang et al., 2015), further bolstering a model wherein the NAb repertoire is established independently of exogenous antigens and is instead the product of programmed development and clonal selection occurring largely in the context of autologous antigens (Hardy et al., 2004; Hayakawa et al., 1999). However, prior studies have shown that microbial colonization modulates the abundance of carbohydrate-reactive NAbs including those specific for dextran, levan and peptidoglycan-associated epitopes (Bos et al., 1989; Kearney et al., 1985). Our current study similarly demonstrates, through a neonatal immunization model and gnotobiotic mouse models, that bacterial antigens are essential for the establishment of clonal diversity in the GlcNAc-reactive B-1 B cell repertoire.

Interestingly, although positive selection by self-antigens is an accepted model of B-1 B cell differentiation (Hayakawa et al., 1999), GlcNAc-modified autoantigens are insufficient to support the development of GlcNAc-reactive B-1 B cells despite the self-reactivity for autologous carbohydrate modifications exhibited by GlcNAc-reactive B cell clonotypes (New et al., 2016; Turner et al., 1990). This may be due to the insufficient quantity of autologous GlcNAc antigens, the predominantly cryptic nature of GlcNAc epitopes in post translation carbohydrate modifications, or perhaps negative selection of GlcNAc-reactive B cells by autoantigens in the absence of bacteria-derived GlcNAc. A comparable scenario has been described in the case of galactoseα1–3galactoseβ1–4GlcNAc (αGal)-reactive B cells, which are typically negatively selected by autologous αGal antigens in mice, yet readily differentiate into Nab-producing B-1b B cells following genetic ablation of the αGalactosyltransferase enzyme (Ohdan et al., 2000). Collectively, these observations suggest that the development of carbohydrate-reactive innate-like B cells is sensitive to the abundance, accessibility, and autologous or microbial context of their respective cognate antigen (New et al., 2016). Exposure to microbiota-derived antigens nonetheless represents an important priming event for carbohydrate-reactive B cell responses.

Similar to the expansion of low-frequency IdI-3a^{-} GAC^{+} B cells expressing germline encoded IGHV7–3 BCRs described herein, we have previously reported scenarios wherein neonatal exposure to bacterial pathogen-associated antigens elicited clonal responses by PC- and dextran-specific B cells which expressed BCR-restricted idiotypes distinct from those of otherwise naïve or adult-immunized mice (Kearney et al., 2015; Vakil and Kearney, 1991). Our examination of IGHV gene expression by emergent GAC^{+} B cells in neonatal mice reveals that the expansion of IGHV7–3 BCR following neonatal immunization with GAS is not the product of a distinct IGHV gene preference during neonatal development, but is likely born from the lack of clonal competition between GAC-reactive B cells at early neonatal time points. By extension, the clonal dominance of IGHV6–3 expressing B cells exhibited by adult immune response to GAS stems from the prior establishment of IGHV6–3 B cells in the pre-immune repertoire through microbiota-dependent mechanisms. This view is supported by the expansion of IGHV6–3 clonotypes that occurs in aged mice, which outcompete other GAC^{+} B cell clonotypes, including neonatally-expanded IGHV7–3 expressing B cells, over time. We therefore propose a model wherein the B cell repertoire is imprinted during neonatal periods of repertoire plasticity and is subsequently fine-tuned by antigen exposure during life. Interestingly, the observation that GAC^{+} IGHV7–3 clonotypes were expanded in d3 GAS-immunized mice, which did not produce measurable GAC-reactive IgM responses, indicates that the capacity of antigens to modulate clonal selection is decoupled from humoral immune response competency.

Our studies demonstrate that sensing bacterial products through MyD88-dependent pathways contributes to the development of GlcNAc-reactive B-1 B cells. Although TLR signaling is important during clonal selection of immature B cells in mice (Seo et al., 2001), B cell-intrinsic expression of MyD88 is dispensable for positive selection of autoreactive B cells into the primary repertoire (Silver et al., 2006). Thus, MyD88-dependent pathogen sensing by phagocytes and intestinal epithelial cells may be integral to the induction of carbohydrate reactive B cell responses (Friedrich et al., 2017). Alternatively, the influence of MyD88 on GAC^{+} B cell differentiation may stem from the IL-5 dependent role of ST2

(IL33r, *Il1rl1*)-signaling in promoting proliferation and IgM production in B-1 B cells (Komai-Koma et al., 2011). By contrast, NOD2 and RIP-2 were not only dispensable for GlcNAc-reactive B-1 B cell development but appeared to suppress production of GlcNAc-reactive IgM. Previous studies have shown the NOD2 agonist muramyl dipeptide can attenuate TLR2-dependent signaling in splenocytes, and perhaps NOD2/RIPK signaling similarly constrains ASC differentiation by carbohydrate-reactive B cells (Watanabe et al., 2004). The integration of pattern recognition receptor signaling on the dynamics of B-1 B cell activation and production of natural and mucosal antibody represents an intriguing area of further investigation.

It remains unclear if microbial antigens directly influence clonal selection of immature B cells or drive the expansion of rare clonotypes present in the peripheral mature naïve B cell pool. While the degree to which immature transitional B cells migrate to gut tissues in mice is poorly understood, seeding of B cells in the peritoneal cavity does not begin until eight days-of-age and is preceded by the appearance of B cells in the SI mesentery (Hamilton et al., 1994). Mouse SI tissue also harbors B cells undergoing Rag-mediated Ig gene rearrangement, suggesting clonal selection could occur in the context of mucosal antigens (Wesemann et al., 2013). Indeed, in humans, immature B cells localize to the gut beginning *in utero* (Golby et al., 2002) and mucosal responses localized to the small intestine give rise to IgM memory B cells (Magri et al., 2017).

The role for B-1 B cells during T cell-independent secretory IgA responses in the SI is well established (Bunker et al., 2015; Fransen et al., 2015; Kaminski and Stavnezer, 2006; Macpherson et al., 2000), and our results underscore the clonal relationship between secretory IgA-producing plasma cells of the lamina propria and the B-1 compartment. We demonstrated the presence of SI LP-localized B cell lineages that were clonally related to GAC^+ B-1 B cells clonotypes conserved in the PerC and spleen, suggesting GAC^+ B cell lineages expanded in mucosal were disseminated systemically. Our data are consistent with the previously reported relatedness between the B-1b B cell compartment and the SI LP B cell pool (Bunker et al., 2015) and suggest that B-1 B cells may serve as a memory reservoir to reconstitute IgA-expressing clonotypes as occurs following depletion of intestinal plasma cells (Lindner et al., 2012).

Overall, this study exemplifies the dramatic effects of early exposure to vaccine-, bacteria- and microbiota-derived antigens on the establishment of clonal diversity in the GlcNAc-reactive B cell repertoire. Finally, given the significant role of natural IgM in self- tolerance (Binder et al., 2003; Brandlein et al., 2003; Kearney et al., 2015; Kin et al., 2012; Vas et al., 2012), the associations of IgM deficiency with increased rates of autoimmunity and allergic disease (Antar et al., 2008; Louis and Gupta, 2014; Takeuchi et al., 2001), and the predictive power of carbohydrate-reactive NAbs in multiple human pathologies (Dotan et al., 2006; Engelbertsen et al., 2015; Schwarz et al., 2006), these findings may provide an alternative mechanistic explanation for the influence of environmental antigens on allergic and autoimmune disease susceptibility often discussed in context of the hygiene hypothesis (Bach, 2002; Strachan, 1989).

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact—Requests for additional information or resources and reagents should be directed to the Lead Contact, John Kearney (jfk@uab.edu).

Materials Availability—This study did not generate any new unique reagents.

Data and Code Availability—The single cell immunoglobulin heavy chain gene sequencing datasets generated during this study are available at GenBank, under the accession numbers MN661405-MN662226 and MT253110-MT253533. The IG Rep-Seq data set is available at BioProject under the accession number PRJNA587770. Program scripts used during analysis of Immunoglobulin sequencing data are available through previously published analysis software, described and cited where appropriate in the *Immunoglobulin Heavy Chain Sequencing* and *Next-generation Immunoglobulin Repertoire Sequencing (IG Rep-Seq)* subsections.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice: Wild-type C57BL/6J mice were purchased from Jackson Laboratories and maintained in our animal colony. B6.129P2-*Tcrb*^{tm1Mom} *Tcrd*$^{tm1Mom/J}$ (*Tcrb*$^{-/-}$/*Tcrd*$^{-/-}$, Jax Stock #002116) (Mombaerts et al., 1992), B6.129P2(SJL)-*MYD88*$^{tm1.1Defr/J}$ (*MYD88*$^{-/-}$, Jax Stock #009088) (Hou et al., 2008), B6.129S1-*Nod2*$^{tm1Flv/J}$ (*NOD2*$^{-/-}$, Jax Stock #005763) (Kobayashi et al., 2005) and B6.129S1-Ripk2$^{tm1Flv/J}$ (*RIPK2*$^{-/-}$, Jax Stock #007017) (Kobayashi et al., 2002) were all obtained from Jackson Laboratories. C57BL/6-*Tlr9*^{tm1Aki} (*Tlr9*$^{-/-}$) mice were provided by Dr. Chad Steele (UAB). Dr. Eric Meffre (Yale University) provided C57BL/6-*PRDM1*$^{+/YFP}$ (Blimp1^{YFP}) mice. All animal protocols were approved by the Institutional Animal Care and Use Committee at the University of Alabama at Birmingham. Mice were housed in specific pathogen-free facilities and given standard chow and acidified water *ad libitum*. Heat-killed, pepsin-treated vaccine bacteria stocks were prepared as previously described (McCarty and Lancefield, 1955). During immunization of neonatal and perinatal C57BL/6J mice 5×10^7 GAS colony-forming units (c.f.u.) were administered intraperitoneally (i.p.) between three and 28 days of age. During immunization of adult mice 10^8 c.f.u. were administered i.p. at eight weeks-of-age (56 days). Germ-free and gnotobiotic ASF-colonized mice were purchased from the UAB Genetically Engineered and Gnotobiotic Mouse Core or Taconic Farms, and housed in sterile isolators in the UAB Genetically Engineered and Gnotobiotic Mouse Core.

METHOD DETAILS

Monoclonal antibodies: The HGAC78 (IgM) GAC-specific hybridoma was generously provided by Dr. Moon Nahm (UAB) (Nahm et al., 1982), and the HGAC39 (IgG3) GAC-specific hybridoma and IdI-3a-specific hybridoma (Clone IA.1) were provided by Dr. Neil Greenspan (Case Western Reserve University) (Greenspan and Davie, 1985). Hybridomas were maintained in serum-free media (Hybridoma SF media, Gibco) and monoclonal antibodies (MAbs) were purified from culture supernatants using prepared GlcNAc-conjugated Sepharose columns; briefly, 10mg of nitrophenyl group of GlcNAc-NP was

dissolved in 0.5 NaHCO3 containing 0.1M sodium hydrosulfite (Na2S2O4) and incubated for two hours at room temperature as in (Bloch and Burger, 1974), prior to addition to cyanogen bromide-activated sepharose (GE Lifesciences). Remaining active cyanogen bromide groups were subsequently blocked with 0.1 M Tris-HCL before the column was extensively washed. Bound antibodies were eluted through addition of soluble GlcNAc, and column eluates were concentrated using AmiconUltra centrifugal filters (Ultracel-10K, Millipore); appropriate molecular weight and antigen-binding properties were confirmed by SDS-PAGE, ELISA. Low endotoxin levels in purified MAb preparations was confirmed routinely by limulus assay (Limulus Amebocyte Lysate Pyrogent, Lonza), and concentrations of MAb stocks were subsequently determined by spectrophotometric analysis (ND-1000, NanoDrop Technologies) prior to storage at 4ºC. In some cases, MAbs were directly conjugated to amine-reactive Alexa Fluor dyes (Life Technologies).

Enzyme-linked immunosorbent assays: Half-area (ELISA) or full area (ELISPOT) 96-well flat-bottom EIA/RIA plates (Costar) were coated overnight at 4ºC with 2 μg/mL of Group A Carbohydrate (10S PG-PS, BD Lee Labs), $GlcNAc_{35}BSA$ (Pyxis Laboratories), or mouse immunoglobulin isotype-specific capture antibodies (goat anti-mouse, Southern Biotechnology). In ELISA assays plates were blocked with PBS+1% BSA. Following adsorption of diluted sera, bound immunoglobulin was detected by alkaline-phosphatase conjugated mouse immunoglobulin isotype-specific secondary antibodies (Southern Biotechnology). Phosphatase reactions were performed at room temperature and stopped via addition of 5N NaOH. Plate absorbances were subsequently read at 405 nm (SPECTROstar Omega, BMG Labtech). Purified HGAC78 (IgM) and HGAC39 (IgG3) MAb preparations were used as standards for quantitation of GlcNAc-reactive serum immunoglobulin. ELISA-based detection of GlcNAc-reactive IgA in titrations of serum and fecal preparations is represented as area under the curve (A.U.C.) calculations. ELISPOT assays for enumeration of ASCs were blocked with PBS+2% gelatin prior to incubation of spleen and BM cells overnight at 37°C in RPMI+10% fetal calf serum. Detection of secreted antibodies was facilitated by alkaline-phosphatase conjugated mouse immunoglobulin isotype-specific secondary antibodies (goat, Southern Biotechnology) and 5-bromo-4-chloro-3-indolyl phosphate substrate reactions.

Flow Cytometry and FACS: All immunostaining was carried out using established conventional methodologies. Briefly, to facilitate analysis of low-frequency antigen-specific B cells, ~10^7 splenocytes were pelleted in 96-well round bottom plates by centrifugation at 1200rpm for 2 minutes and blocked with PBS+2.5% fetal calf serum and 2ug/mL anti-CD16/32 (Ab93) (Oliver et al., 1999). In some cases, non-B cells were excluded from gating analysis through use of a dump cocktail comprised of biotinylated MAbs against CD3, CD11c and F4/80, which were subsequently labeled with Streptavidin PerCP. Dead cells were excluded from analysis by propidium iodide or 7-AAD (Sigma-Aldrich). B cell phenotypes were subsequently identified using a variety of commercially available fluorescence-conjugated monoclonal antibodies. GAC-binding B cells were detected by single epitope multiple staining of B cells (Townsend et al., 2001), using Lancefield Group A Carbohydrate (PGPS-10s, BD Lee Labs), labeled with Alexa Fluor reagents (ThermoFisher Scientific). In some cases, GAC^+ B cells were co-stained with the anti-IdI-3a

idiotype MAb IA.1 (Greenspan and Davie, 1985). Data were acquired on a LSRII or FACSAria (BD) and subsequently analyzed using FlowJo software (Treestar).

Small intestinal lamina propria tissue digests: The isolation of small intestinal lamina propria lymphocytes was completed as describe in (Bunker et al., 2015). Briefly, intestinal pieces were flushed of their contents prior to being cut longitudinally and subsequently cut into ~1cmx1cm pieces. These pieces were washed in RPMI 1% FCS + 1mM EDTA for 20 minutes at 37C with shaking. The intestinal pieces were strained from the solution using a 100μm cell strainer (Fisher Scientific), and the pre-digestion wash step was repeated once. The intestinal pieces were then transferred to digestion solution (RPMI 20% FCS 0.5 mg/ml collagenase A (Roche) 0.1 mg/ml DNAse (Sigma)) and were incubated for 30 minutes at 37C with shaking. The intestinal pieces were again strained from the solution using a 100μm cell strainer, with the supernatant being collected on ice and the remaining tissue being subjected to a second 30min incubation at 37C in digestion solution. The resulting cell supernatants were subsequently pooled, and liberated lymphocytes were isolated by centrifugation over 40% Percoll (GE Lifesciences).

Single-Cell Immunoglobulin Heavy Chain Sequencing: Single GAC-binding B cells were sorted directly into 384-well plates containing 3ul 10mM Tris and 0.75 units/ml of RNAsinPlus (Promega). cDNA was generated from the resulting lysates using the High Capacity cDNA synthesis kit (Applied Biosystems) according to the manufacture's protocol. Rearranged B cell receptor genes were amplified from cDNA by nested PCR using previously described IGHV family and constant region-specific primers (Tiller et al., 2009), and amplification products were subsequently sequenced by Sanger sequencing using an ABI 3730 DNA Analyzer through the Heflin Center for Genomic Science (UAB). FASTA sequence files were submitted to IGMT for IGH gene identification and analyzed using freely available packages through The R Project for Statistical Computing. V-D-J gene usage in GAC^+ mouse B cells was depicted using the Circos software package (Krzywinski et al., 2009), wherein pairing of IGHV (right vertex) and IGHD (left vertex) genes are indicated by connecting arcs, and IGHJ gene usage is indicated in the left inner annulus. Mutation number is indicated in the right inner annulus as shades of gray.

16S Bacteria Phylogenetic analysis: Fecal pellets or pellets from solubilized intestinal contents were collected from mice and stored at −20C until processing. During processing, microbiota samples were thawed on ice prior to extraction of bacterial DNA using *Quick*-DNA Fecal/Soil Microbe Kit minipreps according to the manufacturers protocol (Zymo Research). 16s ribosomal DNA sequences were amplified from bulk fecal DNA using previously described universal primers targeting the V4 regions of the 16s gene locus (Caporaso et al., 2011). Resulting 16s gene amplicons were subsequently sequenced using the MiSeq platform with MiSeq Reagent Kit V2 2×300 (Illumina), according to protocols previously published by the Microbiome Resource at UAB (Kumar et al., 2014), through the Heflin Center for Genomic Science (UAB)

Immunoglobulin repertoire sequencing (IG Rep-Seq): Total RNA was isolated from whole small intestine homogenized in TriZol (Invitrogen), and cDNA was generated using

the High Capacity cDNA kit (Ambion). IGHV gene segments were amplified from cDNA with KOD (Novagen) using primers complimentary to IGHV gene families and the IgA and IgM constant regions, incorporating nucleotide adaptors to allow for barcoding. Following amplification of IGHV genes, amplicons were barcoded with the Illumina TruSeq multiplex PCR kit and sequenced with an Illumina MiSeq at the Heflin Center for Genomic Science (UAB) using Illumina MiSeq V3 2X300 kit. Resulting sequences were filtered for quality (Q>30), assembled, merged using Vsearch, and converted to FASTA files using the Fastq tool kit. FASTA files were submitted to the IMGT/High-Vquest server for IGHV, IGHD, and IGHJ family identification and other bioinformatics analyses (Shi et al., 2014). IGH deep-sequencing analysis was supported using custom in-house developed scripts, and an analysis pipeline described previously (Tipton et al., 2015). Briefly, sequences annotated for VDJ and isotype were clustered into lineages having the same V and J identity, CDR3 length and at least 90% pairwise CDRH3 nucleotide similarity. Predicted GAC-reactive B cell lineages were generated by identifying clonotypes with homology to GAC-reactive single-sorted cell and hybridoma-derived IGH sequences based on the aforementioned lineage clustering criteria (Nogales et al., 2018). Following alignment of these lineages to their respective inferred germline IGHV gene sequences, we generated phylogenetic trees using Phylip's protein parsimony (Protpars) tool (Version 3.695) (Felsenstein, 1989, 2005), with settings 1, 4, and 5 turned-on. The resulting output file was parsed using in-house custom scripts to collapse duplicate sequences into nodes and generate edge files for tree visualization with Cytoscape (Version 3.7.1) (Shannon et al., 2003).

QUANTIFICATION AND STATISTICAL ANALYSIS

For each experiment, the mean and standard deviation (s.d.) or standard error of mean (s.e.m.) were determined for the respective measured parameter, and statistical analyses were performed using Student's t-test (single-factorial comparisons) or one-way or two-way ANOVA when appropriate (multifactorial comparisons), with Tukey's and Bonferroni *post-hoc* tests, respectively. For mouse experiments, sample sizes were not predetermined through statistical methods, but instead chosen based on pilot studies and previously reported results such that appropriate statistical testing could yield significant results. In all cases, samples were tested for similar variance to ensure the assumptions of the statistical comparisons used were met. Blinding was not done in allocation of animals to experimental groups. No specific randomization or exclusion criteria were applied to mouse samples due to use of inbred mouse strains. Analysis was performed with Prism (GraphPad Software).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We acknowledge Enid Keyser, V. Sagar Hanumanthu and the University of Alabama at Birmingham (UAB) Comprehensive Flow Cytometry Center (UAB CFCC, Supported by US National Institutes of Health (NIH) grants P30 AR048311 and P30 AI027767) for assistance in single-cell sorting; the UAB Heflin Center for Genomics Studies Core Laboratory for DNA sequencing; Dr. Casey Morrow and the UAB Microbiome Resource (Supported by The UAB School of Medicine, The Comprehensive Cancer Center (P30 CA013148), The Center for Clinical Translational Science (UL1TR001417)) for assistance in 16s rDNA Microbiome profiling; and Dr. Trenton Schoeb and the Gnotobiotic and Genetically-engineered Mouse Core (Supported by NIH grant G20OD016623) for

gnotobiotic animal resource support. Thanks to Lisa Jia and Jeffrey Sides for assisting with mouse husbandry and hybridoma maintenance, and to Drs. Denise Kaminski and David Woodland for their support in editing this manuscript. This work was supported by NIH grants R01 AI14782, and U01 AI100005 (J.F.K.), F31 AI120500 and T32 A1007051 (J.S.N.), and F30 DK082277 and T32 GM008861 (B.L.P.D).

References

Alugupalli KR, Leong JM, Woodland RT, Muramatsu M, Honjo T, and Gerstein RM (2004). B1b lymphocytes confer T cell-independent long-lasting immunity. Immunity 21, 379–390. [PubMed: 15357949]

Antar M, Lamarche J, Peguero A, Reiss A, and Cole S (2008). A case of selective immunoglobulin M deficiency and autoimmune glomerulonephritis. Clin Exp Nephrol 12, 300–304. [PubMed: 18365138]

Bach JF (2002). The effect of infections on susceptibility to autoimmune and allergic diseases. N Engl J Med 347, 911–920. [PubMed: 12239261]

Baumgarth N, Herman OC, Jager GC, Brown LE, Herzenberg LA, and Chen J (2000). B-1 and B-2 cell-derived immunoglobulin M antibodies are nonredundant components of the protective response to influenza virus infection. The Journal of experimental medicine 192, 271–280. [PubMed: 10899913]

Binder CJ, Horkko S, Dewan A, Chang MK, Kieu EP, Goodyear CS, Shaw PX, Palinski W, Witztum JL, and Silverman GJ (2003). Pneumococcal vaccination decreases atherosclerotic lesion formation: molecular mimicry between Streptococcus pneumoniae and oxidized LDL. Nat Med 9, 736–743. [PubMed: 12740573]

Bloch R, and Burger MM (1974). A rapid procedure for derivatizing agarose with a variety of carbohydrates: its use for affinity chromatography of lectins. FEBS Lett 44, 286–289. [PubMed: 4471977]

Bos NA, Kimura H, Meeuwsen CG, De Visser H, Hazenberg MP, Wostmann BS, Pleasants JR, Benner R, and Marcus DM (1989). Serum immunoglobulin levels and naturally occurring antibodies against carbohydrate antigens in germ-free BALB/c mice fed chemically defined ultrafiltered diet. Eur J Immunol 19, 2335–2339. [PubMed: 2606142]

Bos NA, Meeuwsen CG, Hooijkaas H, Benner R, Wostmann BS, and Pleasants JR (1987). Early development of Ig-secreting cells in young of germ-free BALB/c mice fed a chemically defined ultrafiltered diet. Cell Immunol 105, 235–245. [PubMed: 2434251]

Brandlein S, Pohle T, Ruoff N, Wozniak E, Muller-Hermelink HK, and Vollmers HP (2003). Natural IgM antibodies and immunosurveillance mechanisms against epithelial cancer cells in humans. Cancer Res 63, 7995–8005. [PubMed: 14633732]

Briles DE, Nahm M, Marion TN, Perlmutter RM, and Davie JM (1982). Streptococcal group A carbohydrate has properties of both a thymus-independent (TI-2) and a thymus-dependent antigen. Journal of immunology (Baltimore, Md. : 1950) 128, 2032–2035.

Briles DE, Nahm M, Schroer K, Davie J, Baker P, Kearney J, and Barletta R (1981). Antiphosphocholine antibodies found in normal mouse serum are protective against intravenous infection with type 3 streptococcus pneumoniae. The Journal of experimental medicine 153, 694–705. [PubMed: 7252411]

Bunker JJ, Flynn TM, Koval JC, Shaw DG, Meisel M, McDonald BD, Ishizuka IE, Dent AL, Wilson PC, Jabri B, et al. (2015). Innate and Adaptive Humoral Responses Coat Distinct Commensal Bacteria with Immunoglobulin A. Immunity 43, 541–553. [PubMed: 26320660]

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, and Knight R (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences of the United States of America 108 Suppl 1, 4516–4522. [PubMed: 20534432]

Carapetis JR, Steer AC, Mulholland EK, and Weber M (2005). The global burden of group A streptococcal diseases. Lancet Infect Dis 5, 685–694. [PubMed: 16253886]

Choi YS, Dieter JA, Rothaeusler K, Luo Z, and Baumgarth N (2012). B-1 cells in the bone marrow are a significant source of natural IgM. Eur J Immunol 42, 120–129. [PubMed: 22009734]

Chou MY, Fogelstrand L, Hartvigsen K, Hansen LF, Woelkers D, Shaw PX, Choi J, Perkmann T, Backhed F, Miller YI, et al. (2009). Oxidation-specific epitopes are dominant targets of innate natural antibodies in mice and humans. J Clin Invest 119, 1335–1349. [PubMed: 19363291]

Cole LE, Yang Y, Elkins KL, Fernandez ET, Qureshi N, Shlomchik MJ, Herzenberg LA, Herzenberg LA, and Vogel SN (2009). Antigen-specific B-1a antibodies induced by Francisella tularensis LPS provide long-term protection against F. tularensis LVS challenge. Proceedings of the National Academy of Sciences of the United States of America 106, 4343–4348. [PubMed: 19251656]

Cywes-Bentley C, Skurnik D, Zaidi T, Roux D, Deoliveira RB, Garrett WS, Lu X, O'Malley J, Kinzel K, Zaidi T, et al. (2013). Antibody to a conserved antigenic target is protective against diverse prokaryotic and eukaryotic pathogens. Proceedings of the National Academy of Sciences of the United States of America 110, E2209–2218. [PubMed: 23716675]

Dotan I, Fishman S, Dgani Y, Schwartz M, Karban A, Lerner A, Weishauss O, Spector L, Shtevi A, Altstock RT, et al. (2006). Antibodies against laminaribioside and chitobioside are novel serologic markers in Crohn's disease. Gastroenterology 131, 366–378. [PubMed: 16890590]

Emmrich F, Greger B, and Eichmann K (1983). A cross-reacting human idiotype (B17) associated with antibodies to N-acetyl-D-glucosamine. Specificity, immunoglobulin class association, and distribution in the population. Eur J Immunol 13, 273–278. [PubMed: 6189720]

Engelbertsen D, Vallejo J, Quach TD, Fredrikson GN, Alm R, Hedblad B, Bjorkbacka H, Rothstein TL, Nilsson J, and Bengtsson E (2015). Low Levels of IgM Antibodies against an Advanced Glycation Endproduct-Modified Apolipoprotein B100 Peptide Predict Cardiovascular Events in Nondiabetic Subjects. Journal of immunology (Baltimore, Md. : 1950) 195, 3020–3025.

Etlinger HM, and Heusser CH (1986). T15 dominance in BALB/c mice is not controlled by environmental factors. Journal of immunology (Baltimore, Md. : 1950) 136, 1988–1991.

Felsenstein J (1989). PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics 5, 164–166.

Felsenstein J (2005). PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Fransen F, Zagato E, Mazzini E, Fosso B, Manzari C, El Aidy S, Chiavelli A, D'Erchia AM, Sethi MK, Pabst O, et al. (2015). BALB/c and C57BL/6 Mice Differ in Polyreactive IgA Abundance, which Impacts the Generation of Antigen-Specific IgA and Microbiota Diversity. Immunity 43, 527–540. [PubMed: 26362264]

Friedrich C, Mamareli P, Thiemann S, Kruse F, Wang Z, Holzmann B, Strowig T, Sparwasser T, and Lochner M (2017). MyD88 signaling in dendritic cells and the intestinal epithelium controls immunity against intestinal infection with C. rodentium. PLoS Pathog 13, e1006357. [PubMed: 28520792]

Ghosn EE, Sadate-Ngatchou P, Yang Y, Herzenberg LA, and Herzenberg LA (2011). Distinct progenitors for B-1 and B-2 cells are present in adult mouse spleen. Proceedings of the National Academy of Sciences of the United States of America 108, 2879–2884. [PubMed: 21282663]

Golby S, Hackett M, Boursier L, Dunn-Walters D, Thiagamoorthy S, and Spencer J (2002). B cell development and proliferation of mature B cells in human fetal intestine. Journal of leukocyte biology 72, 279–284. [PubMed: 12149418]

Greenspan NS, and Davie JM (1985). Serologic and topographic characterization of idiotopes on murine monoclonal anti-streptococcal group A carbohydrate antibodies. Journal of immunology (Baltimore, Md. : 1950) 134, 1065–1072.

Haas KM, Poe JC, Steeber DA, and Tedder TF (2005). B-1a and B-1b cells exhibit distinct developmental requirements and have unique functional roles in innate and adaptive immunity to S. pneumoniae. Immunity 23, 7–18. [PubMed: 16039575]

Hamilton AM, Lehuen A, and Kearney JF (1994). Immunofluorescence analysis of B-1 cell ontogeny in the mouse. Int Immunol 6, 355–361. [PubMed: 7514440]

Hardy RR, Wei CJ, and Hayakawa K (2004). Selection during development of VH11+ B cells: a model for natural autoantibody-producing CD5+ B cells. Immunol Rev 197, 60–74. [PubMed: 14962187]

Haury M, Sundblad A, Grandien A, Barreau C, Coutinho A, and Nobrega A (1997). The repertoire of serum IgM in normal mice is largely independent of external antigenic contact. Eur J Immunol 27, 1557–1563. [PubMed: 9209510]

Hayakawa K, Asano M, Shinton SA, Gui M, Allman D, Stewart CL, Silver J, and Hardy RR (1999). Positive selection of natural autoreactive B cells. Science 285, 113–116. [PubMed: 10390361]

Hayakawa K, Formica AM, Zhou Y, Ichikawa D, Asano M, Li YS, Shinton SA, Brill-Dashoff J, Nunez G, and Hardy RR (2017). NLR Nod1 signaling promotes survival of BCR-engaged mature B cells through up-regulated Nod1 as a positive outcome. The Journal of experimental medicine 214, 3067–3083. [PubMed: 28878001]

Hill MO (1973). Diversity and Evenness: A Unifying Notation and Its Consequences. Ecology 54, 427–432.

Hooijkaas H, Benner R, Pleasants JR, and Wostmann BS (1984). Isotypes and specificities of immunoglobulins produced by germ-free mice fed chemically defined ultrafiltered “antigen-free” diet. Eur J Immunol 14, 1127–1130. [PubMed: 6083871]

Horn HS (1966). Measurement of “Overlap” in comparative ecological studies. The American Naturalist 100, 419–424.

Hou B, Reizis B, and DeFranco AL (2008). Toll-like receptors activate innate and adaptive immunity by using dendritic cell-intrinsic and -extrinsic mechanisms. Immunity 29, 272–282. [PubMed: 18656388]

Jayasekera JP, Moseman EA, and Carroll MC (2007). Natural antibody and complement mediate neutralization of influenza virus in the absence of prior immunity. J Virol 81, 3487–3494. [PubMed: 17202212]

Kabanova A, Margarit I, Berti F, Romano MR, Grandi G, Bensi G, Chiarot E, Proietti D, Swennen E, Cappelletti E, et al. (2010). Evaluation of a Group A Streptococcus synthetic oligosaccharide as vaccine candidate. Vaccine 29, 104–114. [PubMed: 20870056]

Kaminski DA, and Stavnezer J (2006). Enhanced IgA class switching in marginal zone and B1 B cells relative to follicular/B2 B cells. Journal of immunology (Baltimore, Md. : 1950) 177, 6025–6029.

Kantor AB, and Herzenberg LA (1993). Origin of murine B cell lineages. Annu Rev Immunol 11, 501–538. [PubMed: 8476571]

Kawahara T, Ohdan H, Zhao G, Yang YG, and Sykes M (2003). Peritoneal cavity B cells are precursors of splenic IgM natural antibody-producing cells. Journal of immunology (Baltimore, Md. : 1950) 171, 5406–5414.

Kearney JF, McCarthy MT, Stohrer R, Benjamin WH Jr., and Briles DE (1985). Induction of germline anti-alpha 1–3 dextran antibody responses in mice by members of the Enterobacteriaceae family. Journal of immunology (Baltimore, Md. : 1950) 135, 3468–3472.

Kearney JF, Patel P, Stefanov EK, and King RG (2015). Natural antibody repertoires: development and functional role in inhibiting allergic airway disease. Annu Rev Immunol 33, 475–504. [PubMed: 25622195]

Kin NW, Stefanov EK, Dizon BL, and Kearney JF (2012). Antibodies generated against conserved antigens expressed by bacteria and allergen-bearing fungi suppress airway disease. Journal of immunology (Baltimore, Md. : 1950) 189, 2246–2256.

Kirvan CA, Swedo SE, Heuser JS, and Cunningham MW (2003). Mimicry and autoantibody-mediated neuronal cell signaling in Sydenham chorea. Nat Med 9, 914–920. [PubMed: 12819778]

Kobayashi K, Inohara N, Hernandez LD, Galan JE, Nunez G, Janeway CA, Medzhitov R, and Flavell RA (2002). RICK/Rip2/CARDIAK mediates signalling for receptors of the innate and adaptive immune systems. Nature 416, 194–199. [PubMed: 11894098]

Kobayashi KS, Chamaillard M, Ogura Y, Henegariu O, Inohara N, Nunez G, and Flavell RA (2005). Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract. Science 307, 731–734. [PubMed: 15692051]

Komai-Koma M, Gilchrist DS, McKenzie AN, Goodyear CS, Xu D, and Liew FY (2011). IL-33 activates B1 cells and exacerbates contact sensitivity. Journal of immunology (Baltimore, Md. : 1950) 186, 2584–2591.

Kroese FG, Butcher EC, Stall AM, Lalor PA, Adams S, and Herzenberg LA (1989). Many of the IgA producing plasma cells in murine gut are derived from self-replenishing precursors in the peritoneal cavity. Int Immunol 1, 75–84. [PubMed: 2487677]

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, and Marra MA (2009). Circos: an information aesthetic for comparative genomics. Genome Res 19, 1639–1645. [PubMed: 19541911]

Kumar R, Eipers P, Little RB, Crowley M, Crossman DK, Lefkowitz EJ, and Morrow CD (2014). Getting started with microbiome analysis: sample acquisition to bioinformatics. Curr Protoc Hum Genet 82, 18 18 11–29. [PubMed: 25042718]

Lindner C, Wahl B, Fohse L, Suerbaum S, Macpherson AJ, Prinz I, and Pabst O (2012). Age, microbiota, and T cells shape diverse individual IgA repertoires in the intestine. The Journal of experimental medicine 209, 365–377. [PubMed: 22249449]

Louis AG, and Gupta S (2014). Primary selective IgM deficiency: an ignored immunodeficiency. Clin Rev Allergy Immunol 46, 104–111. [PubMed: 23760686]

Lutz CT, Bartholow TL, Greenspan NS, Fulton RJ, Monafo WJ, Perlmutter RM, Huang HV, and Davie JM (1987). Molecular dissection of the murine antibody response to streptococcal group A carbohydrate. The Journal of experimental medicine 165, 531–545. [PubMed: 3102673]

Macpherson AJ, Gatto D, Sainsbury E, Harriman GR, Hengartner H, and Zinkernagel RM (2000). A primitive T cell-independent mechanism of intestinal mucosal IgA responses to commensal bacteria. Science 288, 2222–2226. [PubMed: 10864873]

Magri G, Comerma L, Pybus M, Sintes J, Llige D, Segura-Garzon D, Bascones S, Yeste A, Grasset EK, Gutzeit C, et al. (2017). Human Secretory IgM Emerges from Plasma Cells Clonally Related to Gut Memory B Cells and Targets Highly Diverse Commensals. Immunity 47, 118–134 e118. [PubMed: 28709802]

McCarty M (1952). The lysis of group A hemolytic streptococci by extracellular enzymes of Streptomyces albus. II. Nature of the cellular substrate attacked by the lytic enzymes. The Journal of experimental medicine 96, 569–580. [PubMed: 13022851]

McCarty M (1956). Variation in the group-specific carbohydrate of group A streptococci. II. Studies on the chemical basis for serological specificity of the carbohydrates. The Journal of experimental medicine 104, 629–643. [PubMed: 13367334]

Montecino-Rodriguez E, Leathers H, and Dorshkind K (2006). Identification of a B-1 B cell-specified progenitor. Nat Immunol 7, 293–301. [PubMed: 16429139]

Morisita M (1959). Measuring of the dispersion and analysis of distribution patterns. Memoires of the Faculty of Science, Kyushu University, Series E. Biology. 2, 215–235.

Nahm MH, Clevinger BL, and Davie JM (1982). Monoclonal antibodies to streptococcal group A carbohydrate. I. A dominant idiotypic determinant is located on Vk. Journal of immunology (Baltimore, Md. : 1950) 129, 1513–1518.

New JS, King RG, and Kearney JF (2016). Manipulation of the glycan-specific natural antibody repertoire for immunotherapy. Immunol Rev 270, 32–50. [PubMed: 26864103]

Nogales A, Piepenbrink MS, Wang J, Ortega S, Basu M, Fucile CF, Treanor JJ, Rosenberg AF, Zand MS, Keefer MC, et al. (2018). A Highly Potent and Broadly Neutralizing H1 Influenza-Specific Human Monoclonal Antibody. Sci Rep 8, 4374. [PubMed: 29531320]

Ochsenbein AF, Fehr T, Lutz C, Suter M, Brombacher F, Hengartner H, and Zinkernagel RM (1999). Control of early viral and bacterial distribution and disease by natural antibodies. Science 286, 2156–2159. [PubMed: 10591647]

Ohdan H, Swenson KG, Kruger Gray HS, Yang YG, Xu Y, Thall AD, and Sykes M (2000). Mac-1-negative B-1b phenotype of natural antibody-producing cells, including those responding to Gal alpha 1,3Gal epitopes in alpha 1,3-galactosyltransferase-deficient mice. Journal of immunology (Baltimore, Md. : 1950) 165, 5518–5529.

Oliver AM, Grimaldi JC, Howard MC, and Kearney JF (1999). Independently ligating CD38 and Fc gammaRIIB relays a dominant negative signal to B cells. Hybridoma 18, 113–119. [PubMed: 10380010]

Phillips NJ, and Davie JM (1990). Idiotope structure and genetic diversity in anti-streptococcal group A carbohydrate antibodies. Journal of immunology (Baltimore, Md. : 1950) 145, 915–924.

Rapaka RR, Ricks DM, Alcorn JF, Chen K, Khader SA, Zheng M, Plevy S, Bengten E, and Kolls JK (2010). Conserved natural IgM antibodies mediate innate and adaptive immunity against the

opportunistic fungus Pneumocystis murina. The Journal of experimental medicine 207, 2907–2919. [PubMed: 21149550]

Rognes T, Flouri T, Nichols B, Quince C, and Mahe F (2016). VSEARCH: a versatile open source tool for metagenomics. PeerJ 4, e2584. [PubMed: 27781170]

Sabharwal H, Michon F, Nelson D, Dong W, Fuchs K, Manjarrez RC, Sarkar A, Uitz C, Viteri-Jackson A, Suarez RS, et al. (2006). Group A streptococcus (GAS) carbohydrate as an immunogen for protection against GAS infection. J Infect Dis 193, 129–135. [PubMed: 16323141]

Schaedler RW, Dubs R, and Costello R (1965). Association of Germfree Mice with Bacteria Isolated from Normal Mice. The Journal of experimental medicine 122, 77–82. [PubMed: 14325475]

Schwarz M, Spector L, Gortler M, Weisshaus O, Glass-Marmor L, Karni A, Dotan N, and Miller A (2006). Serum anti-Glc(alpha1,4)Glc(alpha) antibodies as a biomarker for relapsing-remitting multiple sclerosis. J Neurol Sci 244, 59–68. [PubMed: 16480743]

Seidl KJ, MacKenzie JD, Wang D, Kantor AB, Kabat EA, Herzenberg LA, and Herzenberg LA (1997). Frequent occurrence of identical heavy and light chain Ig rearrangements. Int Immunol 9, 689–702. [PubMed: 9184914]

Seo S, Buckler J, and Erikson J (2001). Novel roles for Lyn in B cell migration and lipopolysaccharide responsiveness revealed using anti-double-stranded DNA Ig transgenic mice. Journal of immunology (Baltimore, Md. : 1950) 166, 3710–3716.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, and Ideker T (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13, 2498–2504. [PubMed: 14597658]

Shi B, Ma L, He X, Wang X, Wang P, Zhou L, and Yao X (2014). Comparative analysis of human and mouse immunoglobulin variable heavy regions from IMGT/LIGM-DB with IMGT/HighV-QUEST. Theor Biol Med Model 11, 30. [PubMed: 24992938]

Sigal NH, Gearhart PJ, and Klinman NR (1975). The frequency of phosphorylcholine-specific B cells in conventional and germfree BALB/C mice. Journal of immunology (Baltimore, Md. : 1950) 114, 1354–1358.

Silver K, Ferry H, Crockford T, and Cornall RJ (2006). TLR4, TLR9 and MyD88 are not required for the positive selection of autoreactive B cells into the primary repertoire. Eur J Immunol 36, 1404–1412. [PubMed: 16703567]

Stoehr AD, Schoen CT, Mertes MM, Eiglmeier S, Holecska V, Lorenz AK, Schommartz T, Schoen AL, Hess C, Winkler A, et al. (2011). TLR9 in peritoneal B-1b cells is essential for production of protective self-reactive IgM to control Th17 cells and severe autoimmunity. Journal of immunology (Baltimore, Md. : 1950) 187, 2953–2965.

Strachan DP (1989). Hay fever, hygiene, and household size. BMJ 299, 1259–1260. [PubMed: 2513902]

Strober W, Murray PJ, Kitani A, and Watanabe T (2006). Signalling pathways and molecular interactions of NOD1 and NOD2. Nature reviews. Immunology 6, 9–20.

Takeuchi T, Nakagawa T, Maeda Y, Hirano S, Sasaki-Hayashi M, Makino S, and Shimizu A (2001). Functional defect of B lymphocytes in a patient with selective IgM deficiency associated with systemic lupus erythematosus. Autoimmunity 34, 115–122. [PubMed: 11905841]

Tiller T, Busse CE, and Wardemann H (2009). Cloning and expression of murine Ig genes from single B cells. J Immunol Methods 350, 183–193. [PubMed: 19716372]

Tipton CM, Fucile CF, Darce J, Chida A, Ichikawa T, Gregoretti I, Schieferl S, Hom J, Jenks S, Feldman RJ, et al. (2015). Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. Nat Immunol 16, 755–765. [PubMed: 26006014]

Townsend SE, Goodnow CC, and Cornall RJ (2001). Single epitope multiple staining to detect ultralow frequency B cells. J Immunol Methods 249, 137–146. [PubMed: 11226471]

Turner JR, Tartakoff AM, and Greenspan NS (1990). Cytologic assessment of nuclear and cytoplasmic O-linked N-acetylglucosamine distribution by using anti-streptococcal monoclonal antibodies. Proceedings of the National Academy of Sciences of the United States of America 87, 5608–5612. [PubMed: 2116002]

Vakil M, and Kearney JF (1991). Functional relationship between T15 and J558 idiotypes in BALB/c mice. Dev Immunol 1, 213–224. [PubMed: 1726556]

Vas J, Gronwall C, Marshak-Rothstein A, and Silverman GJ (2012). Natural antibody to apoptotic cell membranes inhibits the proinflammatory properties of lupus autoantibody immune complexes. Arthritis Rheum 64, 3388–3398. [PubMed: 22577035]

Vollmer W, Blanot D, and de Pedro MA (2008). Peptidoglycan structure and architecture. FEMS Microbiol Rev 32, 149–167. [PubMed: 18194336]

Watanabe T, Kitani A, Murray PJ, and Strober W (2004). NOD2 is a negative regulator of Toll-like receptor 2-mediated T helper type 1 responses. Nat Immunol 5, 800–808. [PubMed: 15220916]

Wesemann DR, Portuguese AJ, Meyers RM, Gallagher MP, Cluff-Jones K, Magee JM, Panchakshari RA, Rodig SJ, Kepler TB, and Alt FW (2013). Microbial colonization influences early B-lineage development in the gut lamina propria. Nature 501, 112–115. [PubMed: 23965619]

Won WJ, Bachmann MF, and Kearney JF (2008). CD36 is differentially expressed on B cell subsets during development and in responses to antigen. Journal of immunology (Baltimore, Md. : 1950) 180, 230–237.

Won WJ, Foote JB, Odom MR, Pan J, Kearney JF, and Davis RS (2006). Fc receptor homolog 3 is a novel immunoregulatory marker of marginal zone and B1 B cells. Journal of immunology (Baltimore, Md. : 1950) 177, 6815–6823.

Won WJ, and Kearney JF (2002). CD9 is a unique marker for marginal zone B cells, B1 cells, and plasma cells in mice. Journal of immunology (Baltimore, Md. : 1950) 168, 5605–5611.

Yang Y, Wang C, Yang Q, Kantor AB, Chu H, Ghosn EE, Qin G, Mazmanian SK, Han J, and Herzenberg LA (2015). Distinct mechanisms define murine B cell lineage immunoglobulin heavy chain (IgH) repertoires. Elife 4, e09083. [PubMed: 26422511]

Yoshimoto M, Montecino-Rodriguez E, Ferkowicz MJ, Porayette P, Shelley WC, Conway SJ, Dorshkind K, and Yoder MC (2011). Embryonic day 9 yolk sac and intra-embryonic hemogenic endothelium independently generate a B-1 and marginal zone progenitor lacking B-2 potential. Proceedings of the National Academy of Sciences of the United States of America 108, 1468–1473. [PubMed: 21209332]

Zenke G, Eichmann K, and Emmrich F (1984). Characterization of a major human antibody clonotype (1A) by monoclonal antibodies to combining site-associated idiotopes. Eur J Immunol 14, 164–170. [PubMed: 6199214]

Highlights

- GAC-reactive B cells exhibit the phenotype and localization patterns of B-1 B cells
- Neonatal S. pyogenes immunization expands minor IGHV7-3 GAC-reactive BCR clonotypes
- Establishment of IGHV6-3 GAC-reactive B cell clonal dominance is microbiota-dependent
- Microbiota-dependent expansion of GAC-reactive B cells seeds the SI LP with ASCs

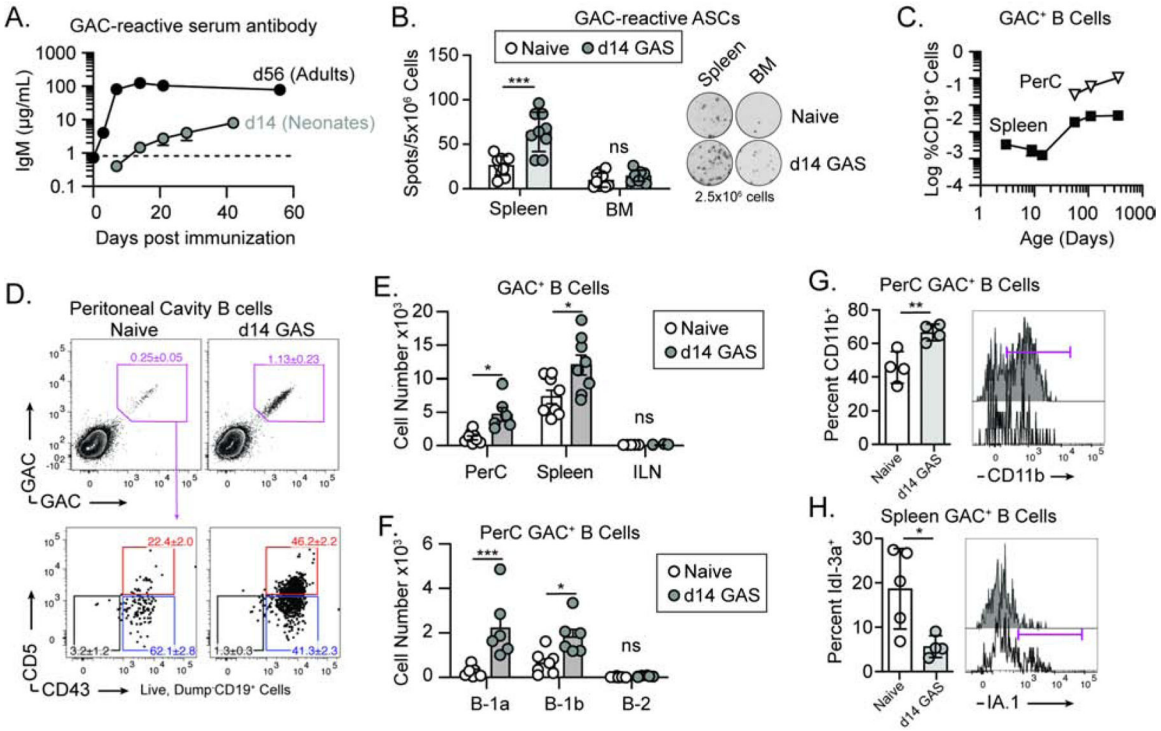


Figure 1. Neonatal B cell responses to Group A Streptococcus are associated with B-1 B cell expansion.

(A) Mean GAC-reactive serum IgM antibody concentrations in C57BL/6 mice following immunization as 14-day-old pups (*grey*) or as 56-day-old adults (*black*), relative to levels in naïve adult mice (*dotted line*). (B) GAC-reactive antibody-secreting cell (ASC) numbers in spleen and BM of naïve (*white*, n=10) or d14 GAS-immunized mice (*grey,* n=9). Data are mean ± s.d. with individual data points overlaid, and representative ELISpot images embedded; ns (not significant), ***p<.001 by two-way ANOVA. (C) GAC^+ B cell frequencies in spleen and peritoneal cavity (PerC) of C57BL/6 mice at neonatal and adult timepoints. (D) Phenotypic analysis of PerC GAC-binding B cells in naïve (*top*) and d14 GAS-immunized (*bottom*) mice; embedded frequencies are mean ± s.e.m. for n=7 naïve and n=8 d14 GAS-immunized mice, pooled from two replicate experiments. (E) Numbers of GAC^+ B cells found in PerC, spleen and ILN of naïve (*white,* n=7) and d14 GAS-immunized (*grey,* n=8) adult mice. (F) Phenotypic distribution of PerC-localized GAC-binding B cells defined as $CD5^+CD43^+CD23^-$ B-1a B cells, $CD5^-CD43^+CD23^-$ B-1b B cells or $CD5^-CD43^-CD23^+$ B-2 B cells. Data are pooled from two replicate experiments; ns (not significant), * p<.05, ***p<.001 by two-way ANOVA with Bonferroni's multiple comparisons test. (G) CD11b expression by GAC^+ B cells in naïve (n=4, *white*) and d14 GAS-immunized (n=4, *grey*) adult mice. (H) IdI-3a idiotype expression by GAC^+ B cells in naïve (n=5) and d14 GAS-immunized (n=4) mice. Data are mean ± s.d. with individual data points from representative experiments completed three times; *p<.05, ** p<.01 by two-tailed t-test. See also Figure S1.

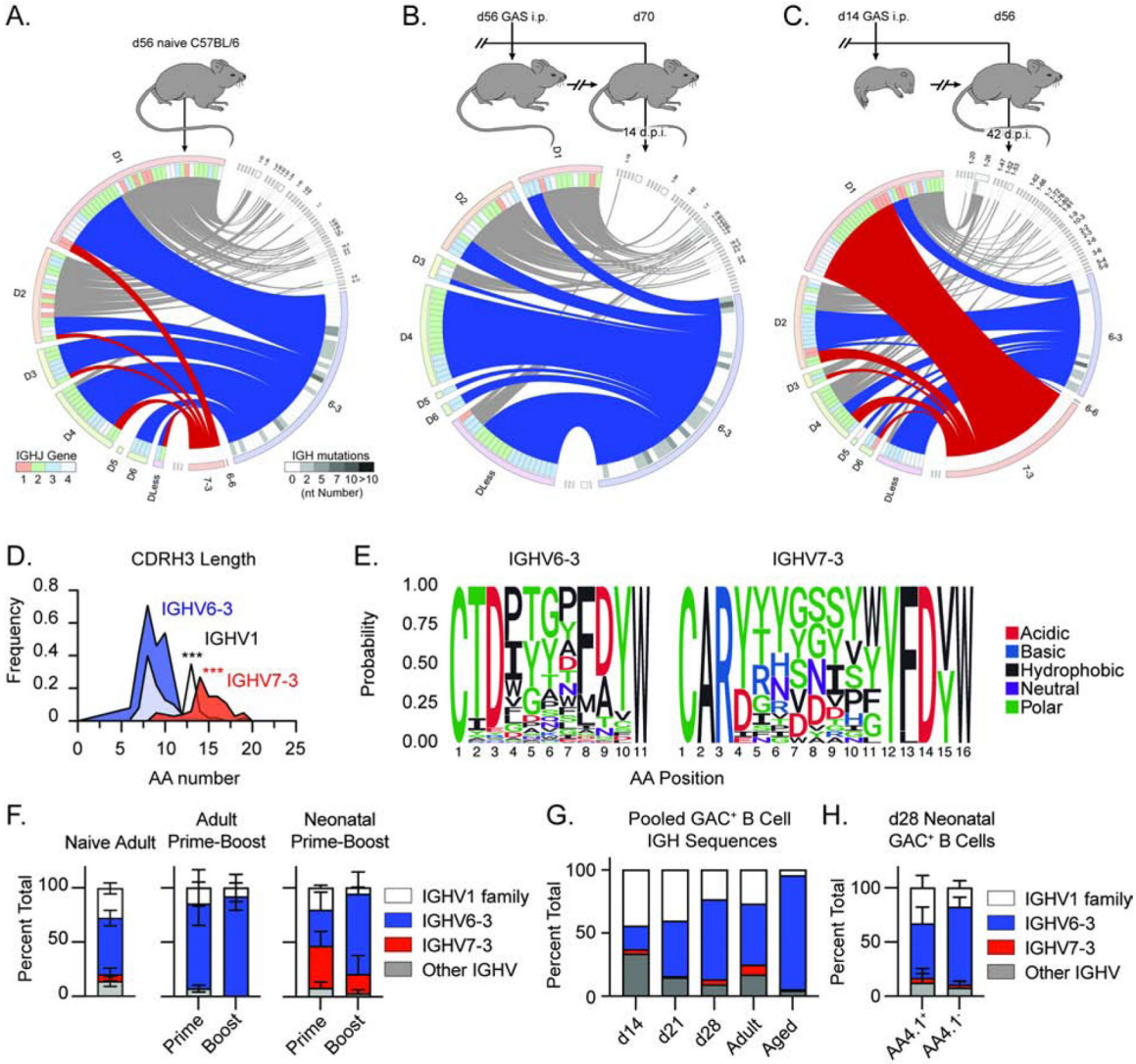


Figure 2. Early immune responses to GAS result in altered clonal dominance in the GAC-reactive B cell pool.

Immunoglobulin heavy chain gene-sequencing of single GAC$^+$ B cells depicted as Circos plots (See Experimental Procedures). (A) GAC$^+$ B cells (n=94) sorted from naïve adult C57BL/6 mice, (B) Adult C57BL/6 mice were immunized with GAS at d56 and GAC$^+$ B cells were sorted 14 days post-immunization (d.p.i.) (n=74), (C) GAC$^+$ B cells sorted from d56 resting adult C57BL/6 mice previously immunized as pups at d14 (n=87). IGH sequences are pooled from naïve (n=4), d56-immunized (n=4) or d14-immunized (n=4) mice from one representative experiment completed twice. Ribbon color coresponds IGHV identity (IGHV6 *blue*, IGHV7–3 *red*, all other IGHVs *grey*), IGHJ usage and IGHV mutations are depicted by inlayed heatmaps. (D) CDRH3 length distribution of pooled IGH gene sequences of identifiied IGHV families of GAC$^+$ B cells. (E) CDRH3 consensus logos of of the mode CDRH3 length for pooled IGHV6–3 and IGHV7–3 IGH gene sequences. See also Figure S2. (F) Relative IGHV gene abundance in recall responses elicited from adult and d14 GAS-immunized mice following a prime-boost regimen. See also Figure S3. (G)

Relative abundance of IGHV gene utilization of GAC^+ B cells in d14, d21 and d28 neonates, adult (8–10wk) mice and aged adult (10–12mo) mice. (H) Relative IGHV gene abundance in $AA4.1^+$ immature and $AA4.1^-$ mature GAC^+ B cells isolated from d28 neonates. See also Figure S3.

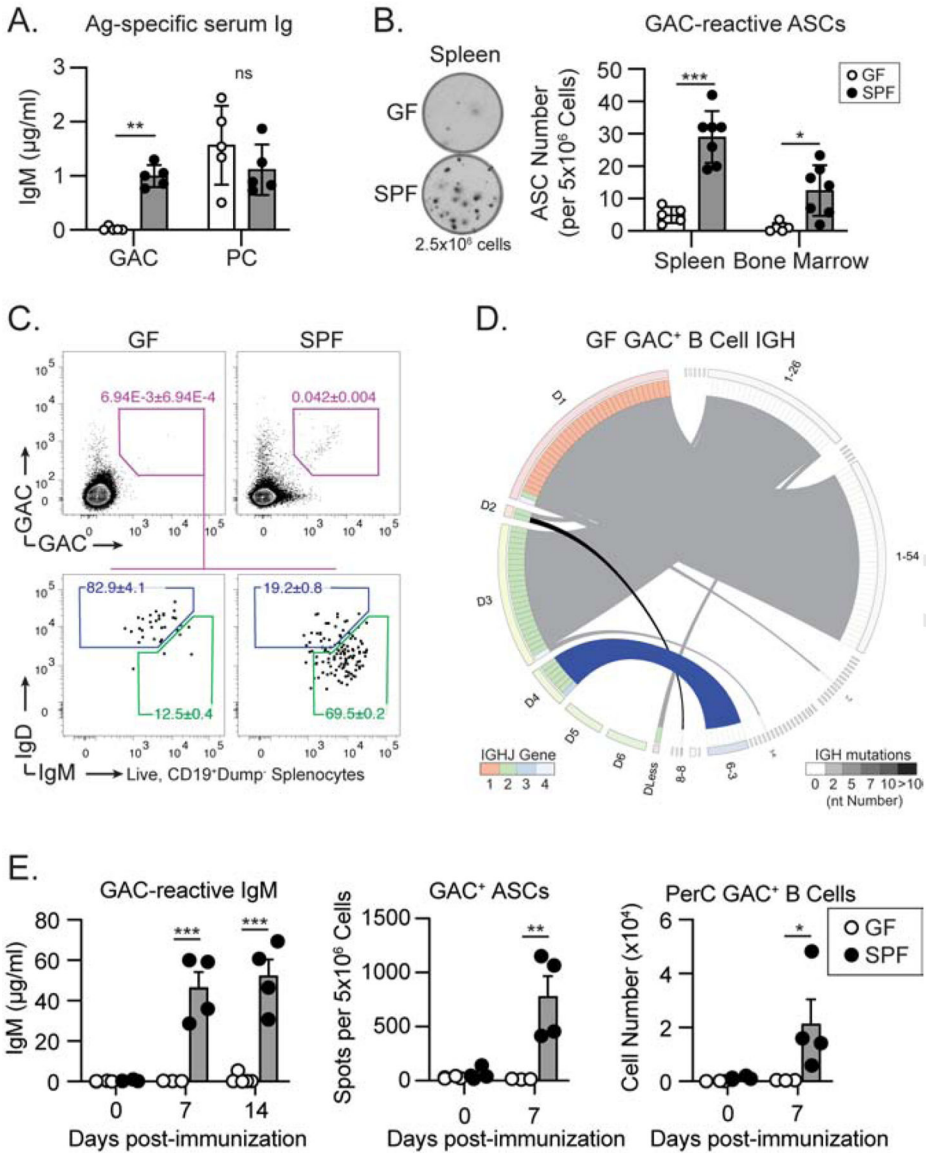


Figure 3. Germ-free mice exhibit defective generation of GlcNAc-reactive B-1 B cells and NAb-producing ASCs.

(A) Serum concentration of GAC- and phosphorylcholine (PC)-reactive IgM antibody in adult germ-free (GF) mice (*white*, n=5) and specific pathogen-free (SPF) mice (*black*, n=5). Data are mean ± s.d. with individual data points overlaid, from one representative experiment completed three times; (B) Numbers of spleen and BM GlcNAc-reactive ASC in 10–12-week-old GF (*white*, n=5) and SPF mice (*black*, n=7). Data are mean ± s.e.m. with individual data points overlaid, pooled from two-independent experiments; ns (not significant), *p<.05, **p<.01, ***p<.001 by two-way ANOVA with Bonferroni's multiple comparisons test. (C) Flow-cytometric enumeration of GAC-reactive B cells in GF and SPF mice (*top*), and representative surface IgD/IgM expression profiles of each group (*bottom*); embedded frequencies are mean ± s.e.m. for n=5 GF and n=7 SPF mice, pooled from two replicate experiments. (D) IgH gene configurations in single GAC^{+} B cells (n=71) pooled from GF mice (n=4), represented as a Circos plot. (E) Quantitation of GlcNAc-reactive

serum IgM antibody responses, spleen ASC numbers and PerC-localized GAC^+ B cells in GF mice (*white*) and SPF mice (*black*) 7–14 days post-immunization. Data are pooled from two replicate experiments; *p<.05, **p<.01, ***p<.001 by two-way ANOVA with Bonferroni's multiple comparisons test. See also Figure S3.

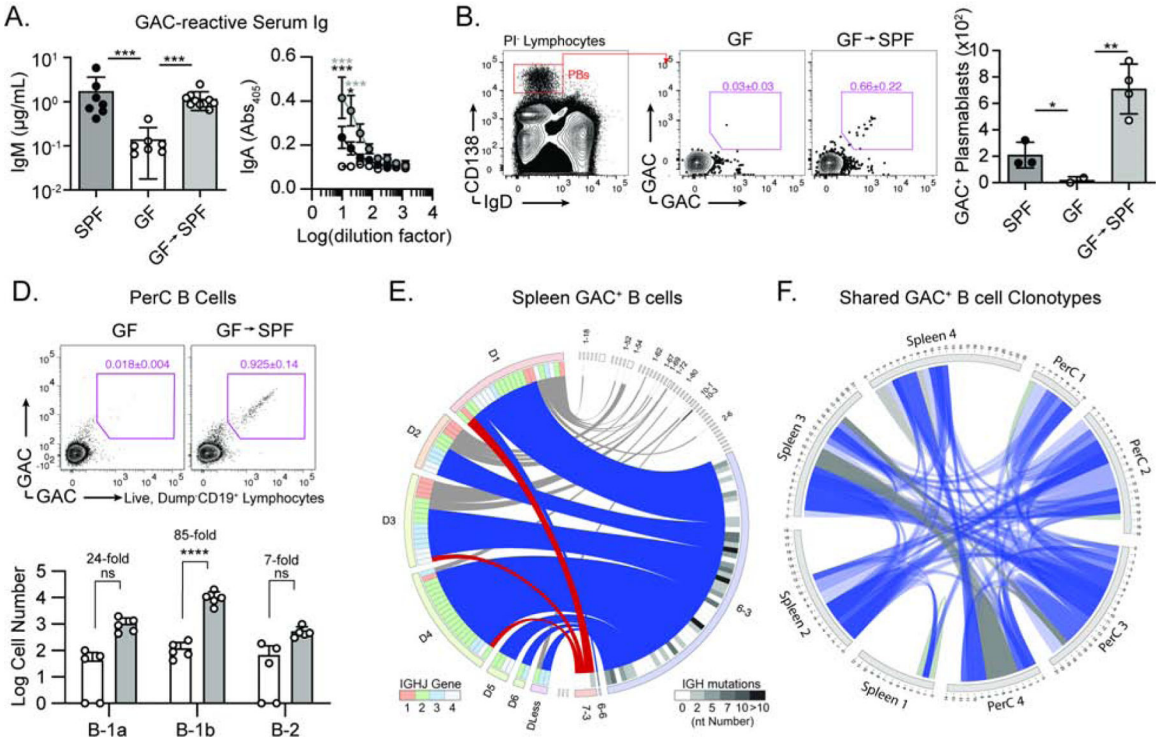


Figure 4. Reconstitution of the microbiota restores GlcNAc-reactive B-1 B cells in GF mice.

GAC-reactive serum IgM (A) and IgA (B) concentrations in SPF (*black*), GF (*white*) and conventionalized (GF→SPF) (*grey*) mice, following eight weeks of co-housing with SPF mice. (C) GAC-reactive B cells within the $IgD^{-}CD138^{+}$ splenic plasmablast compartment of GF→SPF mice; data are mean ± s.e.m. from one representative experiment completed twice; *p<.05, ***p<.001 by one-way ANOVA with Tukey's post-test. (D) Representative flow cytometric plots and enumeration of PerC GAC^{+} B-1 B cells recovered from GF and GF→SPF mice. Data are mean ± s.e.m. pooled from one representative experiment, with individual data points overlaid. (E) Circos plot of IgH gene usage by sorted splenic GAC^{+} B cells (n=81), pooled from n=4 GF→SPF mice. (F) Circos plot depicting GAC^{+} B cell clonotypes shared between the spleen and PerC of individual GF→SPF mice; outer labels depict mouse and tissue and inner labels depict sequence ID. IGHV6–3 are depicted as *blue* ribbons, all other IGHV genes are shown in *grey*. See also Figure S4.

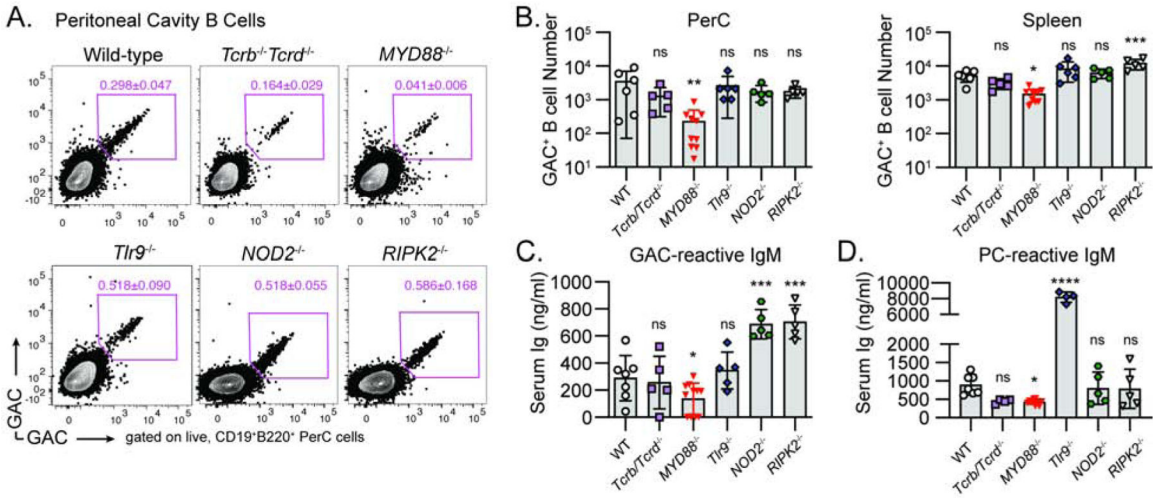


Figure 5. MyD88-signaling contributes to the development of GAC-binding B-1 B cells and GAC-reactive serum antibody.

(A) Representative flow cytometric analyses of PerC-localized GAC-binding B cells and (B) enumeration of PerC and splenic GAC-binding B cells in 8-week-old WT C57BL/6 mice (n=6), TCR β/δ-deficient mice (n=5), MyD88-deficient mice (n=10), TLR9-deficent mice (n=6), NOD2-deficient mice (n=5) and RIP2-deficient mice (n=5). Data are mean ± s.d. with individual data points overlaid, pooled from two-independent experiments; ns (not significant), *p<.05, **p<.01, ***p<.001 by one-way ANOVA with Dunnet's post-test (compared to WT) (C-D) Quantitation of GAC-reactive (C) and phosphorylcholine-reactive (D) serum IgM in each respective strain. Data are mean ± s.d. with individual data points overlaid, pooled from two-independent experiments; ns (not significant), *p<.05, ***p<.001, ***p<.001 by two-tailed student's t-test

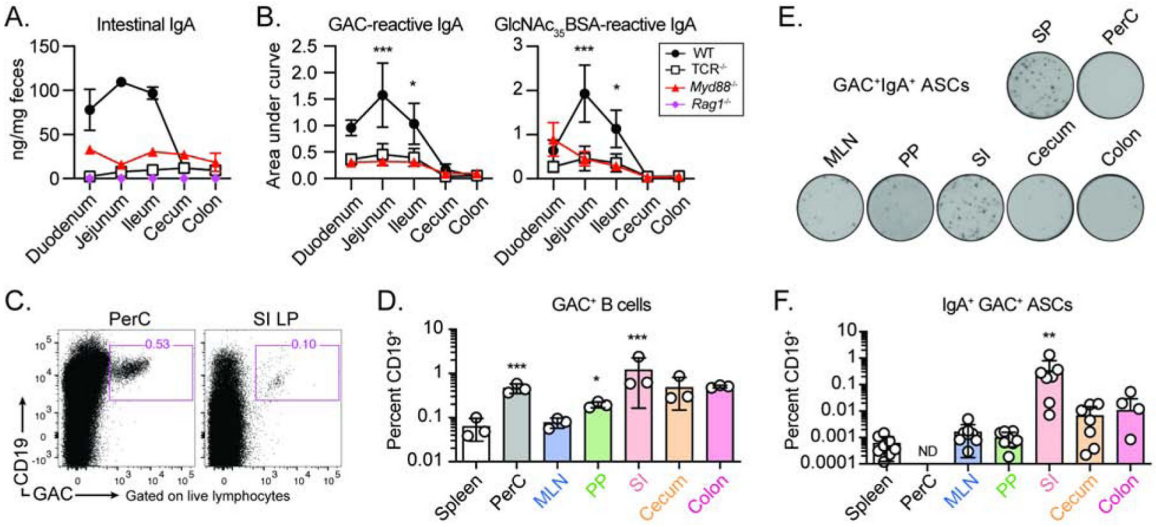


Figure 6. Mouse SI LP harbors GAC-reactive antibody-secreting cells at steady state.

Quantitation of (A) free lumenal total IgA, and (B) GAC-reactive IgA (*left*) and $GlcNAc_{35}BSA$-reactive IgA (*right*) in duodenum, jejunum, ileum, cecum and colon of 8-week-old WT, Rag1-deficient, TCRβ/δ-deficient and MyD88-deficient C57BL/6 mice. Data are mean ± s.e.m. pooled from two replicate experiments; ***p<.001 by two-way ANOVA with Bonferroni's multiple comparisons test, (C) Representative flow cytometric analysis of GAC^+ B cells localized in PerC or small intestinal lamina propria (SI LP) lymphocytes. (D) Frequency of GAC^+ B cells in systemic and gut-associated lymphoid tissues of n=3 adult C57BL/6 mice. Data are mean ± s.e.m., from one representative experiment, completed twice; *p<.05, ***p<.001 by One-way ANOVA with Tukey post-test, relative to spleen frequency. (E) Representative ELISPOT analysis of IgA^+GAC^+ ASCs represented in systemic immune tissues and GALT. (F) Estimated numbers of GAC-specific IgA^+ ASCs in each tissue represented as a frequency of total $CD19^+$ cells (ND, not detected). Data are each replicate, with floating bars depicting mean ± s.e.m., from n=7 adult C57BL/6 mice, pooled from two replicate experiments; **p<.01 by One-way ANOVA with Tukey post-test, relative to spleen frequency. See also Figure S5.

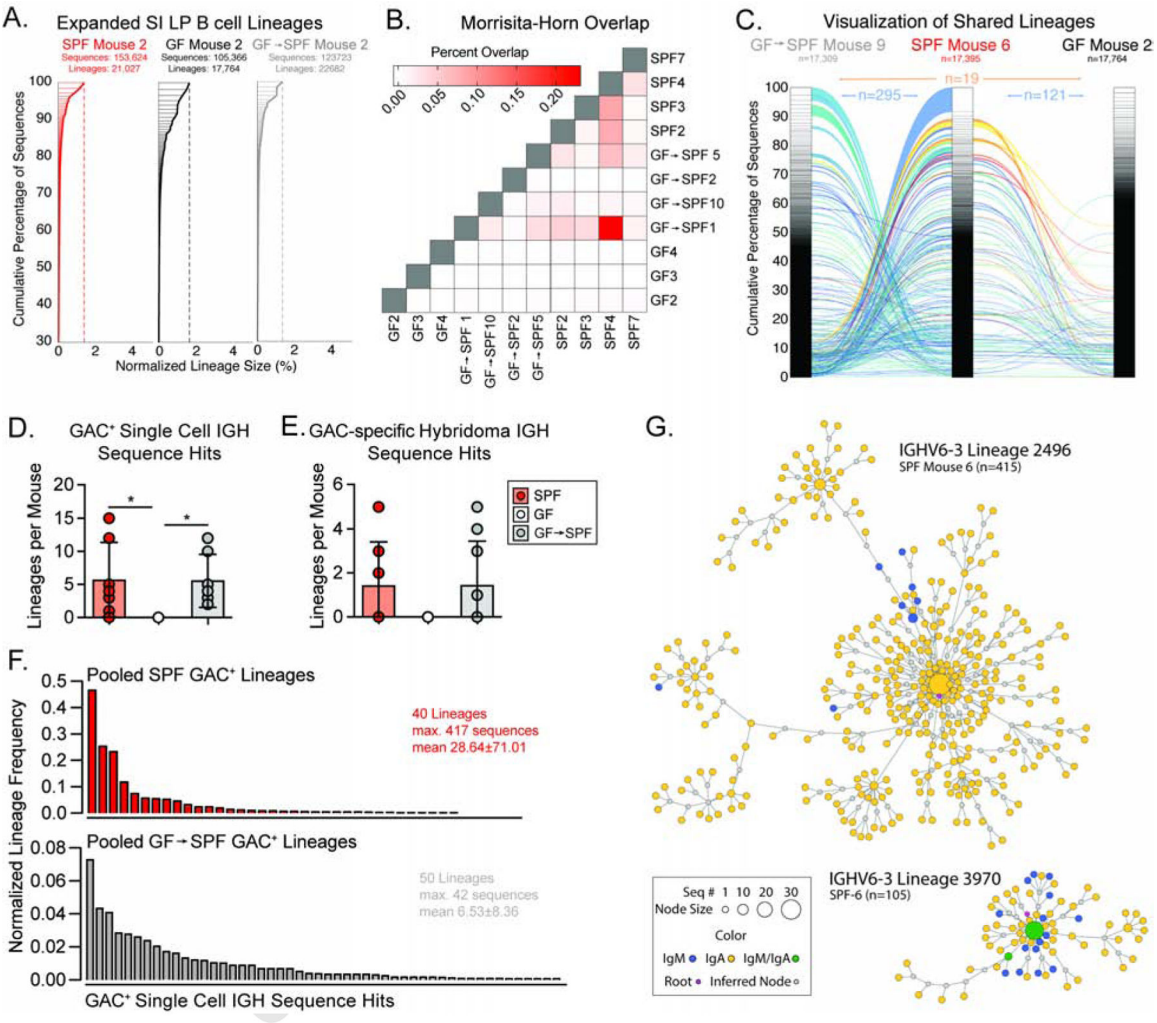


Figure 7. Microbiota colonization induces expansion of clonally conserved lineages of GAC-reactive B cells.

(A) Frequency of B cell lineages from one representative SPF, GF and GF→SPF mice. (B) Heat map of the Morisita-Horn overlap of the SI LP B cell repertoires of individual mice of the indicated groups. (C) Alluvial plot of clonotypes shared between the indicated SPF, GF→SPF and GF samples. Number of distinct lineages present in each sample are indicated below sample name. Clonotypes between only neighboring samples are indicated by *blue-green* shaded ribbons, and clonotypes present in all three samples are indicated by *orange-red* shaded ribbons; number of shared clonotypes are indicated. (D) Enumeration of B cell lineages identified in SI LP tissues from SPF, GF, and SPF GF mice sharing homology with spleen and PerC GAC$^+$ B cell clonotypes identified by single-cell sorting (*Top*) or previously characterized GAC$^+$ hybridomas (*Bottom*). (E-F) Size distribution of B cell lineages (identified in (D)) pooled from SPF (*Top*) or GF→SPF (*Bottom*) mice which share homology to single GAC$^+$ B cell IGH sequences (E) or previously characterized GAC$^+$ hybridomas (F). Lineage trees depicting probable relationship and diversification of SI LP-localized IGH sequences identified as probable GAC-binding B cell clonotypes. The number of unique sequences is indicated by node size, and the isotype of nodes is color-coded. The

root sequence, shown in fuchsia, is defined as the germline IGHV and IGHJ genes combined with the consensus CDHR3 sequence of the node. See also Figure S6.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-Lancefield Group A Carbohydrate (clone HGAC78, μ)	(Nahm et al., 1982)	N/A
Mouse monoclonal anti-Lancefield Group A Carbohydrate (clone HGAC39, γ3)	(Greenspan and Davie, 1985)	N/A
Rat monoclonal anti-IdI-3a idiotype (clone IA.1, γ1)	(Greenspan and Davie, 1985)	N/A
Goat polyclonal anti-Mouse IgM, unlabeled	Southern Biotech	CAT NO 1021-01
Goat polyclonal anti-Mouse IgA, unlabeled	Southern Biotech	CAT NO 1040-01
Goat polyclonal anti-Mouse IgG, unlabeled	Southern Biotech	CAT NO 1030-01
Goat polyclonal anti-Mouse IgM, alkaline phosphatase	Southern Biotech	CAT NO 1021-04
Goat polyclonal anti-Mouse IgA, alkaline phosphatase	Southern Biotech	CAT NO 1040-04
Goat polyclonal anti-Mouse IgG, alkaline phosphatase	Southern Biotech	CAT NO 1030-04
Goat polyclonal anti-Mouse Kappa, PE	Southern Biotech	CAT NO 1050-09
Goat polyclonal anti-Mouse Lambda, PE	Southern Biotech	CAT NO 1060-09
Goat polyclonal anti-Mouse IgA, AlexaFluor 647	Southern Biotech	CAT NO 1040-31
Rat monoclonal anti-Mouse CD16/CD32 (Clone Ab93)	(Oliver et al., 1999)	N/A
Hamster monoclonal-anti Mouse CD3 (Clone 145-2C11), biotin	BD Biosciences	Cat: 553059
Hamster monoclonal-anti Mouse CD11c (Clone HL3), biotin	BD Biosciences	Cat: 553800
Rat monoclonal-anti Mouse F4/80 (Clone RM8), biotin	eBioscience	REF 13-4801-81
Rat monoclonal anti-Mouse CD19 (Clone 1D3), PE-CF594	BD Biosciences	Cat: 562291
Rat monoclonal anti-Mouse IgD (Clone 11-26c.2a), BV510	BD Biosciences	Cat: 563110
Rat monoclonal anti-Mouse CD43 (Clone S7), FITC	BD Biosciences	Cat: 553270
Rat monoclonal anti-Mouse CD43 (Clone S7), PE	BD Biosciences	Cat: 553271
Rat monoclonal anti-Mouse CD23 (Clone B3B4), PE	BD Biosciences	Cat: 553139
Rat monoclonal anti-Mouse CD5 (Clone 53-7.3), PECy7	eBioscience	REF 25-0051-81
Rat monoclonal anti-Mouse B220 (Clone 1D3), PE	BD Biosciences	Cat: 553089
Rat monoclonal anti-Mouse B220 (Clone 1D3), APCCy7	BD Biosciences	Cat: 553092
Rat monoclonal anti-Mouse CD11b (Clone M1/70), BV711	BD Biosciences	Cat: 563168
Rat monoclonal anti-Mouse CD11b Clone M1/70), APC-Cy7	BD Biosciences	Cat: 557657
Rat monoclonal anti-Mouse CD93 (clone AA4.1), FITC	BD Biosciences	Cat: 559156
Rat monoclonal anti-Mouse CD138 (clone 281-2), PE	BD Biosciences	Cat: 561070
Rat monoclonal anti-Mouse CD138 (clone 281-2), BV605	BD Biosciences	Cat: 563147
Rat monoclonal anti-Mouse CD138 (clone 281-2), PECy7	BioLegend	Cat: 142514
Rat monoclonal anti-Mouse CD23 (Clone B3B4), PECy7	Invitrogen	REF 25-0232-82
Rat monoclonal anti-Mouse CD93 (clone AA4.1), PeCy7	Invitrogen	REF 25-5892-81
Rat monoclonal anti-Mouse IgM (clone 11/41), PeCy7	Invitrogen	REF 25-5790-82
Monoclonal anti-Mouse CD36 (clone MZ1), AlexaFluor 488	(Won et al., 2008)	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Monoclonal anti-Mouse FcRL5 (clone MZ2) AlexaFluor 488	(Won et al., 2006)	N/A
Monoclonal anti-Mouse CD9 (clone MZ3) AlexaFluor 488	(Won and Kearney, 2002)	N/A
Streptavidin, PerCP	BD Biosciences	Cat: 554064
Bacterial and Virus Strains		
Streptococcus pyogenes: Grouping strain J17A4; GAS	ATCC	ATCC 12385
Biological Samples		
None involved in study		
Chemicals, Peptides, and Recombinant Proteins		
S.pyogenes peptidoglycan, 10S	BD-Lee Labs Inc.	Cat: 210866
GlcNAcylated BSA, $GlcNAc_{35}BSA$	Pyxis Laboratories	NAGB281541
Phosphorylcholine BSA, PC-BSA	Biosearch Technologies	PC-1011-10
4-Nitrophenyl-N-acetyl-beta-D-glucosamine	BIOSYNTH	Cat: N-4050
Sodium hydrosulfite	Sigma Aldrich	Cat: 157953
Phosphatase Substrate	Sigma Aldrich	Cat: P4744
5-bromo-4-chloro-3-indolyl-phosphate	Roche	REF 11383221001
7-AAD	BD	Cat: 51-68981E
Collagenase A	Roche	REF 10103586001
Deoxyriobonuclease I	Sigma Aldrich	Cat: DN25
RNAsinPlus	Promega	N2111
KOD Hot Start DNA Polymerase	Sigma Millipore	71086-4
TRIzol Reagent	Invitrogen	REF 15596018
Gelatin	Sigma	G1890
Bovine Serum Albumin	Fischer Scientific	BP1605
Fetal Calf Serum	Atlanta Biologicals	S11150
Cyanogen bromide activated Sepharose 4 Fast Flow	GE LifeSciences	17-0981-01
Percoll	GE LifeSciences	17-0891-01
Critical Commercial Assays		
High Capacity cDNA synthesis kit	Applied Biosystems	4368813
TruSeq CD Index adaptors	Integrated DNA Technologies	D501-D508 D701-D712
MiSeq Reagent Kit v2 2X300	Illumina	MS-102-3003
123count eBeads Counting Beads	Invitrogen	REF 01-1234-42
Quick-DNA Fecal/Soil Microbe Kit	Zymo Research	D6010
Deposited Data		
Single Mouse GAC^+ B cell immunoglobulin heavy chain gene sequencing data	GenBank	MN661405-MN662226, MT253110-MT253533
Mouse Small intestine B cell repertoire sequencing data	BioProject	PRJNA587770

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Cell Lines		
None involved in study		
Experimental Models: Organisms/Strains		
Mouse: wild-type C57BL/6	The Jackson Laboratory	JAX: 000664
Mouse: TCRβ/δ$^{-/-}$: B6.129P2-Tcrb^{tm1Mom} Tcrd^{tm1Mom}/J	The Jackson Laboratory	JAX: 002116
Mouse: *MYD88*$^{-/-}$: B6.129P2(SJL)-*MYD88*$^{tm1.1Defr}$/J	The Jackson Laboratory	JAX: 009088
Mouse: *NOD2*$^{-/-}$: B6.129S1-*NOD2*^{tm1Flv}/J	The Jackson Laboratory	JAX: 005763
Mouse*: RIPK2*$^{-/-}$, B6.129S1-RIPK2^{tm1Flv}/J	The Jackson Laboratory	Jax: 007017
Mouse: *TLR9*$^{-/-}$, C57BL/6-*Tlr9*^{tm1Aki}	Dr. Chad Steele (UAB)	
Mouse: Blimp1^{YFP}, C57BL/6-*PRDM1*$^{+/YFP}$	Dr. Eric Meffre (Yale University)	
Mouse: Germ-free wild-type C57BL/6; C57BL/6NTa	Taconic Farms	
Mouse: Germ-free wild-type C57BL/6	UAB Genetically Engineered and Gnotobiotic Mouse Core	
Mouse: Altered Schaedler's Flora-colonized gnotobiotic wild-type C57BL/6	UAB Genetically Engineered and Gnotobiotic Mouse Core	
Oligonucleotides		
Primer: IGHV Forward: 5' Ms VHE GGGAATTCGAGGTGCAGCTGCAGGAGTCTGG	(Seidl et al., 1997) Integrated DNA Technologies	
Primer: mIGHM Reverse GCTAGGTACTTGCCCCCTGTCC	Integrated DNA Technologies	
NGS mVH1a CCCTACACGACGCTCTTCCGATCTCAGGTGTCCACTCCCAGGTCC	Integrated DNA Technologies	
NGS mVH1b CCCTACACGACGCTCTTCCGATCTCAGGTGTCCTCTCTGAGGTCCAG	Integrated DNA Technologies	
NGS mVH2 CCCTACACGACGCTCTTCCGATCTGCTGTGTCCTRTCCCAGGTGCA	Integrated DNA Technologies	
NGS mVH3 CCTACACGACGCTCTTCCGATCTCTGGTATCCTGTCTGATGTSCAGCTTC	Integrated DNA Technologies	
NGS mVH6 CCCTACACGACGCTCTTCCGATCTAAGGKGTCCAGAGTGAGGTGAAGC	Integrated DNA Technologies	
NGS mVH7 CCCTACACGACGCTCTTCCGATCTATGGTATCCAGTGTGAGGTGAAGCTGG	Integrated DNA Technologies	
NGS mVH8 CCCTACACGACGCTCTTCCGATCTCATATGTCCTGTCCCAGRTTACTCTGAAAG	Integrated DNA Technologies	
NGS mVH9 CCCTACACGACGCTCTTCCGATCTCAAAGTGCCCAAGCACAGATCCAG	Integrated DNA Technologies	
NGS mVH14 CCCTACACGACGCTCTTCCGATCTCAGGGGTCAATTCAGAGGTTCAGC	Integrated DNA Technologies	

REAGENT or RESOURCE	SOURCE	IDENTIFIER
NGS IGHM Reverse ATCTCGTATGCCGTCTTCTGCTTGAGGGGGCTCTCGCAGGAGACGAGG	Integrated DNA Technologies	
NGS IGHA Reverse GAGTTCAGACGTGTGCTCTTCCGATCTGCCGAAAGGGAAGTAATCGTGAAT	Integrated DNA Technologies	
NGS IGHG GAGTTCAGACGTGTGCTCTTCCGATCGCTCAGGGAAATAGCCCTTGAC	Integrated DNA Technologies	
Universal 16S V4 5' primers AATGATACGGCGACCACCAGATCTACACTATGGTAATTGTGTGCCAGCMGCCGCGGTAA	(Caporaso et al., 2011)	
Universal 16S V4 3' primers TAATCTWTGGGVHCATCAGGCCGACTGACTGANNNNNNTAGAGCATACGGCAGAAGA CGAAC	(Caporaso et al., 2011)	
Recombinant DNA		
None involved in study		
Software and Algorithms		
FlowJo V10	BD	
Prism V8	Graphpad	
The R Project for Statistical Computing		
Circos Software Package	(Krzywinski et al., 2009)	
IMGT/High-Vquest	(Shi et al., 2014)	
Vsearch	(Rognes et al., 2016)	
IgSeq, B cell receptor repertoire sequencing analysis pipeline	(Tipton et al., 2015)	
Phylip (Version 3.695)	(Felsenstein, 1989, 2005)	
Cytoscape (Version 3.7.1)	(Shannon et al., 2003)	
Other		
N/A		