**Supplementary Materials:**

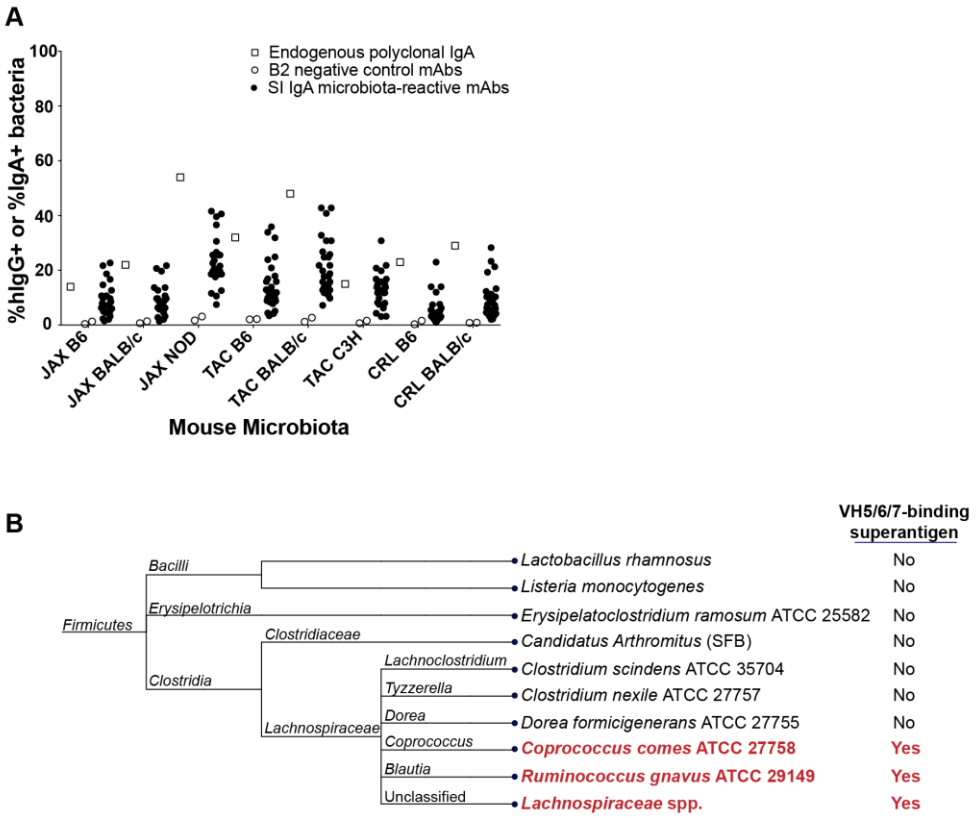
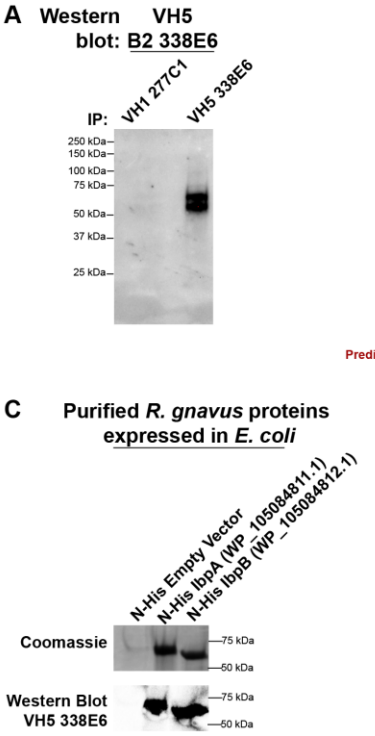



**Fig. S1. mAb reactivity to mouse microbiota and cladogram of individual bacterial strains analyzed. (A)** Bacterial flow cytometry analysis assessing endogenous polyclonal IgA (open squares) or reactivity of negative control B2 mAb (open circles) or microbiota-reactive SI IgA mAbs (closed circles) against fecal microbiota from mice of indicated strains and vendors. Each circle represents a distinct mAb. JAX = Jackson Laboratories. TAC = Taconic Biosciences. CRL = Charles River Laboratories. B6 = C57BL/6. **(B)** Cladogram analysis of individual bacterial strains analyzed for superantigen-type reactivity. Strains with superantigen-type binding are

shown in red. Taxonomy for each strain was determined using the NCBI Taxonomy browser. Tree generated using PhyloT.




B

**Mass spectrometry analysis, Sudarsam et al. 2007 genome**

| Identified proteins (Top 10) | Accession number | Spectral counts p55 | p65 | p75 |
|---|---|---|---|---|
| Ribonuclease Y | tr\|A7B2G9\|A7B2G9_RUMGN | 55 | 68 | 38 |
| Phosphoenolpyruvate carboxykinase (ATP) | tr\|A7B1Y0\|A7B1Y0_RUMGN | 46 | 60 | 17 |
| Translation initiation factor IF-2 | tr\|A7B8A5\|A7B8A5_RUMGN | 40 | 52 | 76 |
| Inosine-5'-monophosphate dehydrogenase | tr\|A7B3H0\|A7B3H0_RUMGN | 33 | 43 | 2 |
| Putative ribose/galactose/methyl galactoside import ATP-binding protein | tr\|A7B0I8\|A7B0I8_RUMGN | 7 | 43 | 2 |
| Aldehyde-alcohol dehydrogenase | tr\|A7B514\|A7B514_RUMGN | 34 | 42 | 49 |
| Pyruvate-flavodoxin oxidoreductase | tr\|A7B798\|A7B798_RUMGN | 32 | 40 | 44 |
| Ribonuclease J | tr\|A7B397\|A7B397_RUMGN | 17 | 39 | 8 |
| SPFH/Band 7/PHB domain protein | tr\|A7B8N4\|A7B8N4_RUMGN | 46 | 38 | 3 |
| Protein translocase subunit SecA | tr\|A7B7L6\|A7B7L6_RUMGN | 23 | 35 | 48 |

**Mass spectrometry reanalysis, Beaulaurier et al. 2018 genome**

| Identified proteins (Top 10) | Accession number | Spectral counts p55 | p65 | p75 |
|---|---|---|---|---|
| Ribonuclease Y | WP_004842699.1 | 55 | 68 | 38 |
| Phosphoenolpyruvate carboxykinase (ATP) | WP_004842366.1 | 46 | 60 | 17 |
| Translation initiation factor IF-2 | WP_004844757.1 | 45 | 59 | 83 |
| Predicted cell wall localization — Hypothetical protein | WP_105084811.1 | 42 | 43 | 26 |
| Predicted cell wall localization — Hypothetical protein | WP_105084812.1 | 41 | 43 | 25 |
| Bifunctional acetaldehyde-CoA/alcohol dehydrogenase | WP_004843573.1 | 34 | 43 | 49 |
| IMP dehydrogenase | WP_004843034.1 | 33 | 43 | 2 |
| Sugar ABC transporter ATP-binding protein | WP_004841622.1 | 8 | 43 | 2 |
| Pyruvate-flavodoxin oxidoreductase | WP_009244813.1 | 31 | 40 | 44 |
| RNase J family beta-CASP ribonuclease | WP_004843012.1 | 17 | 39 | 8 |

**Fig. S2. Immunoprecipitation and mass spectrometry analysis and Coomassie and western blot analysis of purified proteins. (A)** Western blot analysis of immunoprecipitate obtained using beads linked to VH1 mAb 277C1 or VH5 mAb 338E6. Western blot was probed with VH5 338E6. **(B)** Mass spectrometry analysis of 55, 65, or 75 kDa bands cut from a silver stained gel run using the immunoprecipitate shown in panel A. In the top panel, peptides were analyzed against the Sudarsam et al. 2007 *R. gnavus* ATCC 29149 genome. In the bottom panel, peptides were reanalyzed against the Beaulaurier et al 2018 genome. **(C)** Coomassie gel analysis (top) or Western blot analysis (bottom) of *R. gnavus* superantigen proteins or empty vector control expressed in *E. coli* and purified.

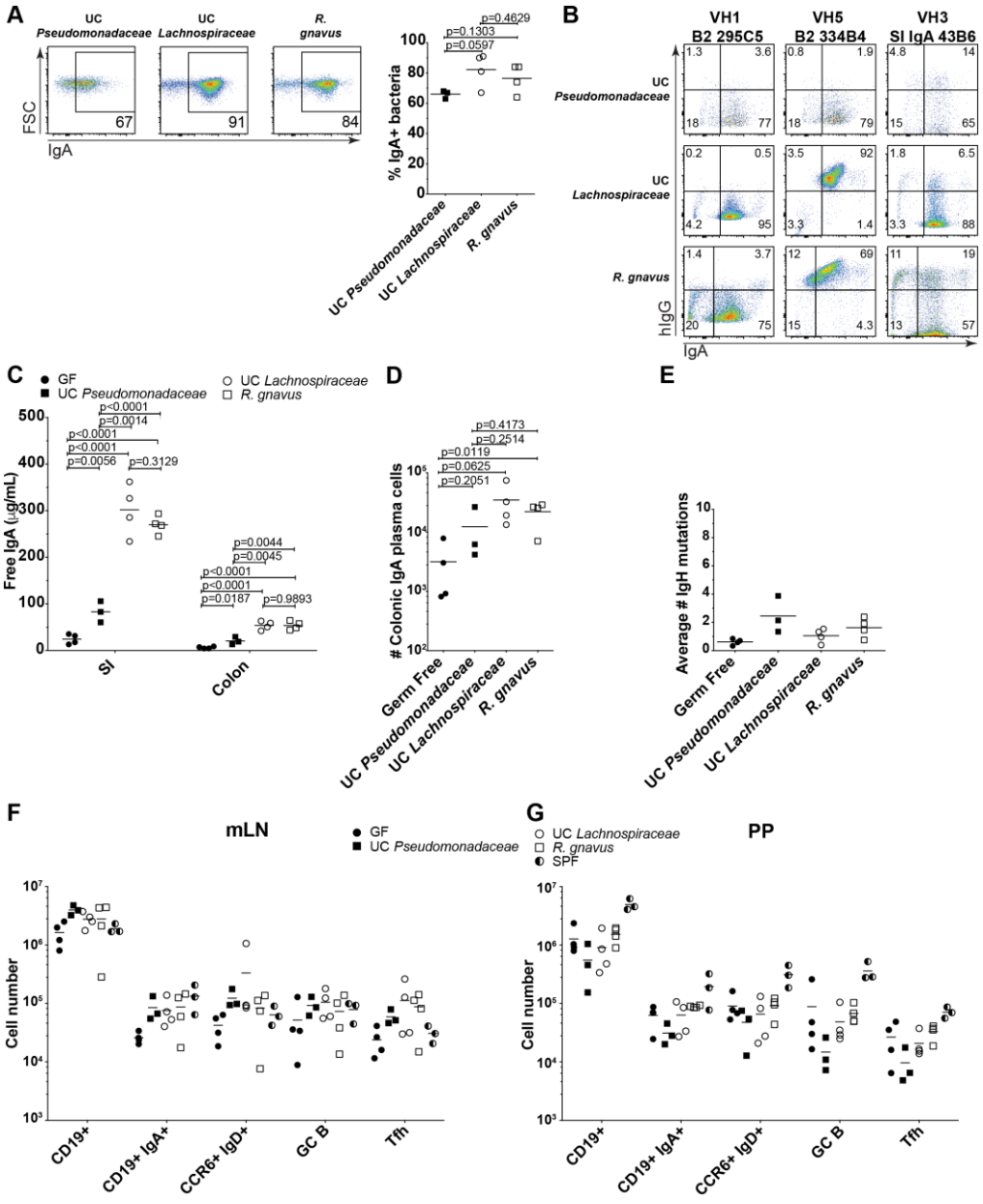



**Fig. S3. Additional analyses of GF or gnotobiotic monocolonized mice. (A)** Representative bacterial flow cytometry analysis (left) and summary plot (right) of polyclonal IgA coating in

indicated groups of mice. Data compiled from three independent experiments. **(B)** mAb and polyclonal IgA bacterial staining analysis of indicated groups with indicated mAbs. Feces were pooled from multiple mice from each group prior to staining. **(C)** Free luminal IgA analysis as measured by ELISA. Data compiled from three independent experiments. One-way ANOVA $p<0.0001$ for both SI and colon. **(D)** Colonic LP IgA PC numbers in indicated groups. Data compiled from three independent experiments. **(E)** Average number of IgH somatic mutations as determined by repertoire analysis of sorted IgA PCs from each mouse in the indicated groups. **(F)** Absolute number of indicated populations in mLN or **(G)** PPs of indicated groups of mice. SPF, specific pathogen free. Data compiled from three independent experiments. In all panels, p values were calculated by unpaired t test.

**Table S1. Distribution of superantigens across human metagenomes.** Number and percent of human metagenomes from indicated countries of origin with detectable *R. gnavus* (whole genome) or each of the two superantigen genes determined using a detection threshold of ≥0.25.

| Country of origin | Number of metagenomes (MGs) analyzed | Total number of short reads | Avg. number of short reads per MG | *R. gnavus* detected | Superantigen *ibpA* WP_105084811.1 detected | Superantigen *ibpB* WP_105084812.1 detected |
|---|---|---|---|---|---|---|
| USA | 150 | 6,545,670,628 | 43,637,804.2 | 64 (42%) | 64 (42%) | 61 (41%) |
| China | 74 | 938,343,512 | 12,680,317.7 | 32 (43%) | 32 (43%) | 32 (43%) |
| Fiji | 172 | 5,610,872,642 | 32,621,352.6 | 16 (9%) | 12 (7%) | 10 (7%) |
| Tanzania | 38 | 336,157,473 | 8,846,249.29 | 2 (5%) | 2 (5%) | 2 (5%) |

**Data file S1. List of mAbs used in this study.** Excel file listing the name and origin of mAbs used in this study.

**Data file S2. List of metagenomes analyzed.** Excel file listing the identification and accession numbers and quality-passed read information for the metagenomes analyzed in Fig. 5 and Table S1.

**Data file S3. Metagenomic recruitment analysis of *R. gnavus* genes including its superantigens.** Raw data from read recruitment analyses of the 3520 genes in the *R. gnavus* ATCC 29149 genome against the 434 metagenomes studied. The superantigen genes are labeled at positions 3060 and 3062.