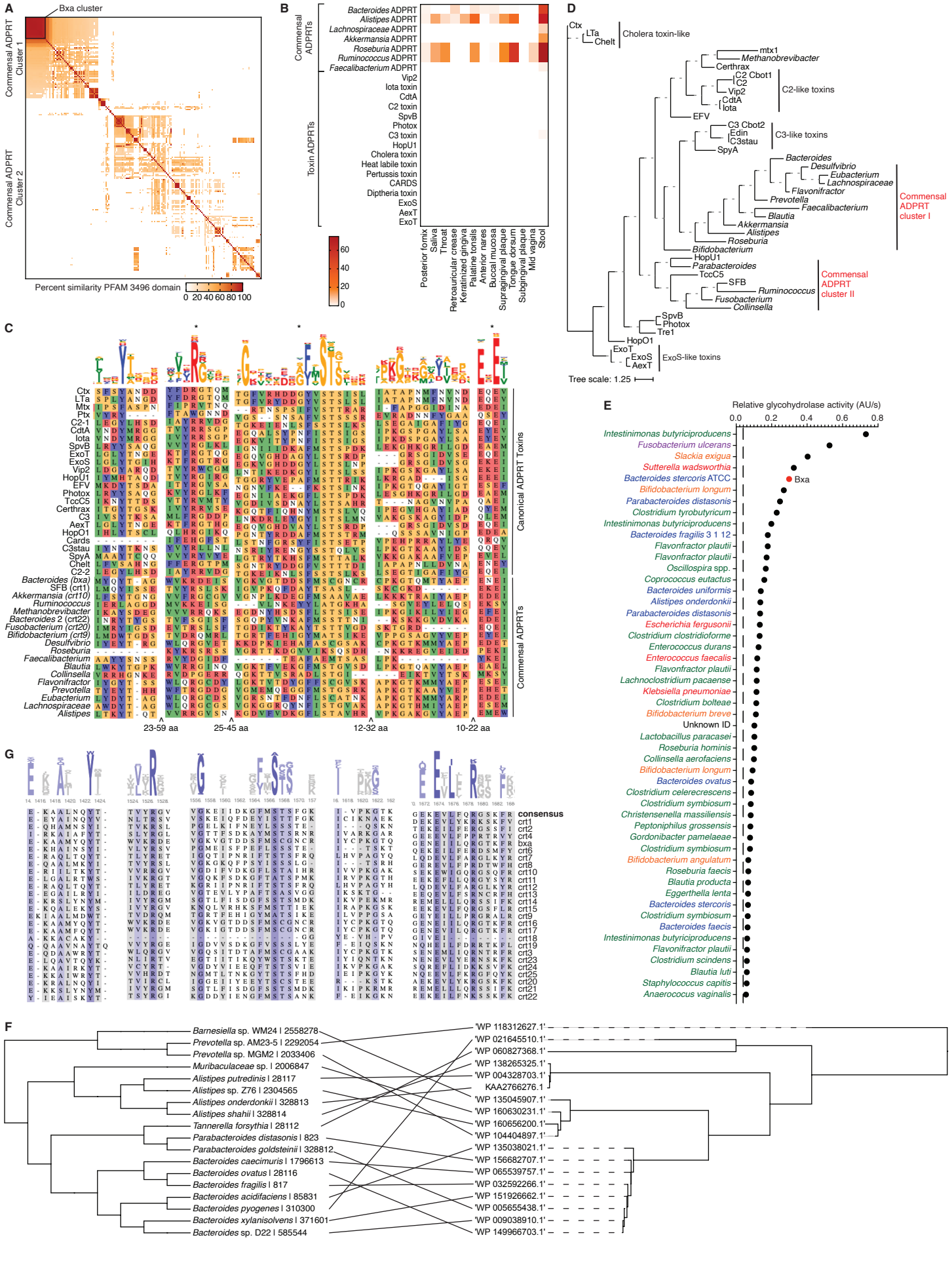
A
Bxa cluster
Commensal ADPRT Cluster 1
Commensal ADPRT Cluster 2
Percent similarity PFAM 3496 domain
0 20 40 60 80 100
B
Commensal ADPRTs
Toxin ADPRTs
Bacteroides ADPRT
Alistipes ADPRT
Lachnospiraceae ADPRT
Akkermansia ADPRT
Roseburia ADPRT
Ruminococcus ADPRT
Faecalibacterium ADPRT
Vip2
Iota toxin
CdtA
C2 toxin
SpvB
Photox
C3 toxin
HopU1
Cholera toxin
Heat labile toxin
Pertussis toxin
CARDS
Diptheria toxin
ExoS
AexT
ExoT
60
40
20
0
Posterior fornix
Saliva
Throat
Retroauricular crease
Keratinized gingiva
Palatine tonsils
Anterior nares
Buccal mucosa
Supragingival plaque
Tongue dorsum
Subgingival plaque
Mid vagina
Stool
D
Ctx
LTa
ChelT
Cholera toxin-like
mtx1
Methanobrevibacter
Certhrax
C2 Cbot1
C2
Vip2
CdtA
Iota
C2-like toxins
EFV
C3 Cbot2
Edin
C3stau
SpyA
C3-like toxins
Bacteroides
Desulfvibrio
Eubacterium
Lachnospiraceae
Flavonifractor
Prevotella
Faecalibacterium
Blautia
Akkermansia
Alistipes
Roseburia
Commensal ADPRT cluster I
Bifidobacterium
HopU1
Parabacteroides
TccC5
SFB
Ruminococcus
Fusobacterium
Collinsella
Commensal ADPRT cluster II
SpvB
Photox
Tre1
HopO1
ExoT
ExoS
AexT
ExoS-like toxins
Tree scale: 1.25
C
Ctx
LTa
Mtx
Ptx
C2-1
CdtA
Iota
SpvB
ExoT
ExoS
Vip2
HopU1
EFV
Photox
TccC5
Certhrax
C3
AexT
HopO1
Cards
C3stau
SpyA
ChelT
C2-2
Bacteroides (bxa)
SFB (crt1)
Akkermansia (crt10)
Ruminococcus
Methanobrevibacter
Bacteroides 2 (crt22)
Fusobacterium (crt20)
Bifidobacterium (crt9)
Desulfvibrio
Roseburia
Faecalibacterium
Blautia
Collinsella
Flavonifractor
Prevotella
Eubacterium
Lachnospiraceae
Alistipes
Canonical ADPRT Toxins
Commensal ADPRTs
23-59 aa
25-45 aa
12-32 aa
10-22 aa
E
Relative glycohydrolase activity (AU/s)
0.0
0.2
0.4
0.6
0.8
Intestinimonas butyriciproducens
Fusobacterium ulcerans
Slackia exigua
Sutterella wadsworthia
Bacteroides stercoris ATCC
Bifidobacterium longum
Parabacteroides distasonis
Clostridium tyrobutyricum
Intestinimonas butyriciproducens
Bacteroides fragilis 3 1 12
Flavonfractor plautii
Flavonfractor plautii
Oscillospira spp.
Coprococcus eutactus
Bacteroides uniformis
Alistipes onderdonkii
Parabacteroides distasonis
Escherichia fergusonii
Clostridium clostridioforme
Enterococcus durans
Enterococcus faecalis
Flavonfractor plautii
Lachnoclostridium pacaense
Klebsiella pneumoniae
Clostridium bolteae
Bifidobacterium breve
Unknown ID
Lactobacillus paracasei
Roseburia hominis
Collinsella aerofaciens
Bifidobacterium longum
Bacteroides ovatus
Clostridium celerescrens
Clostridium symbiosum
Christensenella massiliensis
Peptoniphilus grossensis
Gordonibacter pamelaeae
Clostridium symbiosum
Bifidobacterium angulatum
Roseburia faecis
Blautia producta
Eggerthella lenta
Bacteroides stercoris
Clostridium symbiosum
Bacteroides faecis
Intestinimonas butyriciproducens
Flavonifractor plautii
Clostridium scindens
Blautia luti
Staphylococcus capitis
Anaerococcus vaginalis
Bxa
G
consensus
crt1
crt2
crt4
bxa
crt6
crt7
crt8
crt10
crt11
crt12
crt13
crt14
crt15
crt9
crt16
crt17
crt18
crt19
crt3
crt23
crt24
crt25
crt20
crt21
crt22
F
Barnesiella sp. WM24 | 2558278
Prevotella sp. AM23-5 | 2292054
Prevotella sp. MGM2 | 2033406
Muribaculaceae sp. | 2006847
Alistipes putredinis | 28117
Alistipes sp. Z76 | 2304565
Alistipes onderdonkii | 328813
Alistipes shahii | 328814
Tannerella forsythia | 28112
Parabacteroides distasonis | 823
Parabacteroides goldsteinii | 328812
Bacteroides caecimuris | 1796613
Bacteroides ovatus | 28116
Bacteroides fragilis | 817
Bacteroides acidifaciens | 85831
Bacteroides pyogenes | 310300
Bacteroides xylanisolvens | 371601
Bacteroides sp. D22 | 585544
'WP 118312627.1'
'WP 021645510.1'
'WP 060827368.1'
'WP 138265325.1'
'WP 004328703.1'
KAA2766276.1
'WP 135045907.1'
'WP 160630231.1'
'WP 160656200.1'
'WP 104404897.1'
'WP 135038021.1'
'WP 156682707.1'
'WP 065539757.1'
'WP 032592266.1'
'WP 151926662.1'
'WP 005655438.1'
'WP 009038910.1'
'WP 149966703.1'

Figure S1. Figure S1. ADP-ribosyltransferases are highly prevalent and abundant in commensal strains from the gut microbiome of humans, Related to Figure 1. A) A sequence homology matrix of all ADPRT sequences identified in the IGC Microbiome Database, with percent similarity between the catalytic domain of each shown in the heat map ranging from white (none), to orange (10-60%) to red (>60%). The catalytic domain was determined using the sequences spanning the length of the predicted PFAM 3496 domain. Sequences most similar to Bxa PF 3496 ADPRT domain were used as the reference point and are shown in the top left of the graph. **B)** Prevalence of ADPRTs identified from commensal microbes compared to those identified and characterized as toxin ADPRTs in different body sites, based on metagenomic data from the Human Microbiome Project. Prevalence values range from 0% (white) to 80% (red) and are calculated as an average prevalence per sample. **C)** A catalytic domain-restricted alignment of 25 well-characterized, canonical ADPRT toxins found in pathogens (top) and 18 ADPRT sequences from selected microbiome strains of the most abundant genera found in the gut microbiome (bottom). Alignment was performed using the MUSCLE algorithm, where the size of the amino acid residues in each position on the top panel relates to its relative conservation amongst the aligned ADPRTs. The five most conserved regions in the alignment are shown. Numbers in between each amino acid alignment indicate the number of amino acids present between each conserved region of the catalytic domain. An * indicates the canonical R-S-ExE sequences. **D)** A phylogenetic analysis of the ADPRT sequences from (C) was performed using a BLOSUM62 algorithm with the name of the toxin or its genera listed at the leaves of the dendrogram. Dendrogram clusters are labeled based on known ADPRT toxin clusters and new commensal ADPRT clusters. Dotted lines represent different nodes and clustering within the algorithm. **E)** The top 50 relative NAD-glycohydrolase activity values from the fluorimetric assay performed on 300 clinical microbiome strains. The taxonomic information for each strain is listed on the y-axis and the NAD-glycohydrolase value on the x-axis. We used a cutoff of 0.04 (2x standard deviation above control) to determine a positive reaction (dotted line). The phyla of each taxa is denoted by color: Bacteroidetes (blue), Firmicutes (green), Actinobacteria (orange), Proteobacteria (red) and Fusobacteria (purple). Bxa activity is labeled on the graph. Data is representative of 2 independent experiments. **F)** Dual phylogenetic trees (BLOSUM62) generated based taxa phylogeny (left) and the phylogeny of ADPRT sequences (WP numbers added) from the identical strains (right). Black lines connect the ADPRT sequence with the strain it is found in, with strain IDs noted after the taxa information. **G)** A catalytic domain-restricted alignment of the top 25 ADPRTs (*crt* genes) noted in Fig. 1D. Alignment was performed using the MUSCLE algorithm, where the size of the amino acid residues in each position on the top panel relates to its relative conservation amongst the aligned ADPRTs. Conservation of each amino acid is shown as using a colorimetric scale of grey (low conservation), light purple (some conservation) to dark purple (highly conserved residue). The five most conserved regions in the alignment are shown.

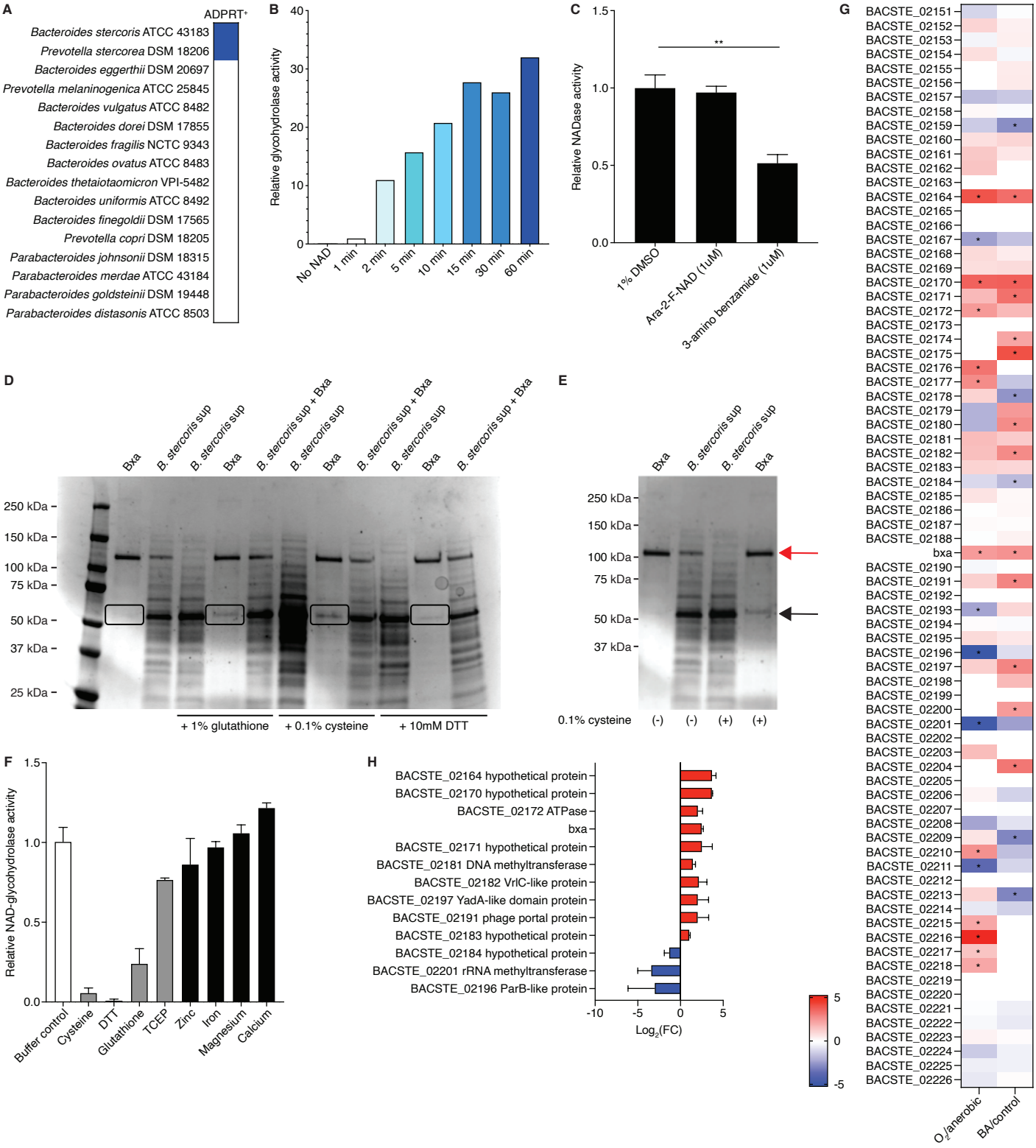

Figure S2. *Bacteroides* ADPRTs are expressed, secreted and enzymatically active proteins, Related to Figure 3. A) A list of the 16 sequenced type strains from the *Bacteroides* or *Parabacteroides* deposited in the ATCC or DSMZ biobank and whether a putative ADPRT gene is present in the deposited genome sequence (blue is present). **B)** Relative NAD-glycohydrolase activity increases over baseline control from *B. stercoris* supernatants exposed to a minimal Tris buffer for 60 min. Samples were taken and analyzed at the 1 min, 2 min, 5 min, 10 min, 15 min, 30 min and 60 min intervals. **C)** Relative NAD-glycohydrolase activity of purified Bxa with either 1% DMSO, 1μM Ara-2-F-NAD or 1μM 3-amino benzamide added to the assay. Statistical analysis performed using a Student's t-test; **p<0.01. Data is representative of 2 independent experiments. Error bars represent the mean +/- SEM. **D)** An SDS-PAGE gel of (left to right) purified Bxa, *B. stercoris* supernatant (*B. stercoris* sup) or *B. stercoris* supernatant plus purified Bxa (*B. stercoris* sup + Bxa). Each of the 3 samples are either unexposed to reducing agents (farthest left) or exposed to 1% glutathione, 0.1% cysteine or 10mM DTT (left to right). Black boxes highlight the band appearing at ~55 kDa. Image is representative of 3 independent experiments. **E)** An SDS-PAGE gel of (left to right) purified Bxa or *B. stercoris* supernatant (*B. stercoris* sup) either treated (+) or untreated (-) with 0.1%

cysteine. Black arrow points to the 55 kDa observed band; red arrow points to the 105 kDa Bxa protein. Image is representative of 3 independent experiments. **F)** Relative NAD-glycohydrolase activity of purified Bxa treated with 0.1% cysteine, 10mM DTT, 1% glutathione, 0.01% TCEP, 100 μM zinc, 5mM iron, 1mM magnesium or 1mM calcium. Values are computed as relative to the buffer control (white), which is normalized to 1. **G)** A heat-map of the fold-change differences over control for phage-associated genes from *B. stercoris* grown in BHI and either exposed to oxygen for 1 hr or bile acids (BA) for 2 hrs before RNA-sequencing (n=3). Genes that are significantly up- or down-regulated are denoted with an asterisk (FDR<0.05). A full list of differentially expressed genes in this experiment is provided in File S1. Error bars represent the mean +/- SEM. **H)** All the Bxa phage-associated genes (y-axis) that were consistently upregulated (red) or downregulated (blue) at a $\log_2$ fold-change (x-axis) >2 in the oxygen and bile acid conditions were plotted. Based on RNA-sequencing data (n=3) from (G). Other consistently co-expressed genes between the conditions can be found in File S1. Error bars represent the mean +/- SEM.

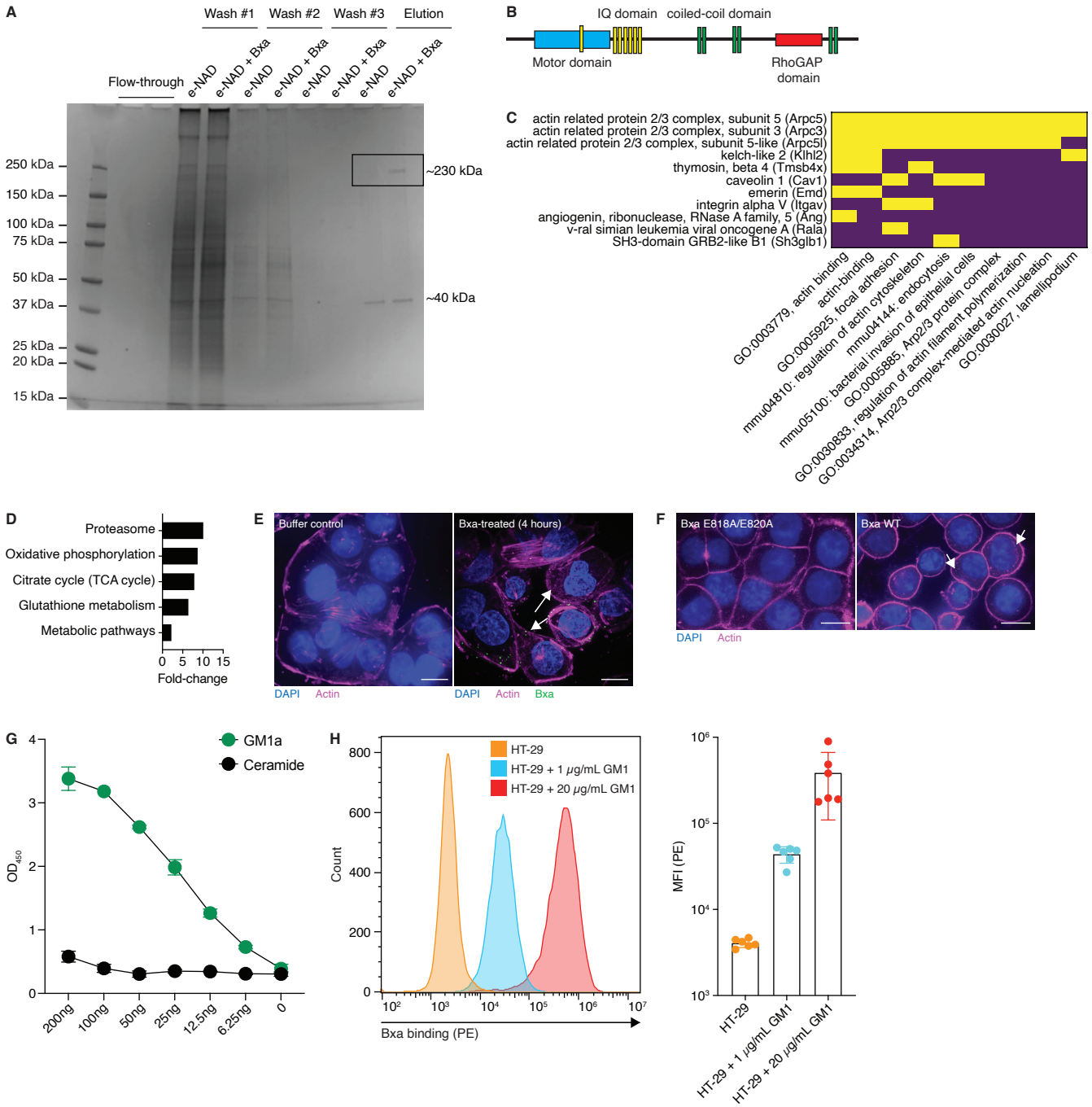

Figure S3. *Bacteroides* ADPRTs disrupt the actin cytoskeleton, focal adhesions and cell migration in epithelial cells, Related to Figure 4. A) SDS-PAGE gel of the proteins immunoprecipitated with an anti-ethenoadenosine antibody from HT-29 cell lysates treated with e-NAD alone or e-NAD plus Bxa. The elution from the IP beads is shown in the far right lanes. The lanes to the left show protein bands from the wash of the beads during the IP. Protein mass of the ladder is shown on the left. The box shows the dominant band seen in the gel, at ~230kDa, which is the differential between e-NAD alone and e-NAD plus Bxa. Image is representative of 3 independent experiments. **B)** A schematic of non-muscle myosin II protein domains, representative of the structure of the myosin-9, myosin-10 and myosin-14 proteins identified in the IP-MS analysis. Noted in red is the Rho-GAP domain, which is unique to the non-muscle myosin II proteins. **C)** Genes found within an actin cytoskeleton gene expression module as identified by GSEA (y-axis) and the related GO terms associated with the genes (x-axis). Presence of each gene within the GO term is denoted by yellow (present) or purple (not present). **D)** The most significantly altered pathways in Bxa-treated murine monolayers, based upon GSEA of each gene that was significantly downregulated (FC>2; FDR <0.01) after Bxa treatment. The x-axis describes the fold-change of the pathway over expected values from a control-treated sample. **E)** Representative images of HT-29 cells after 4 hrs of being treated with Bxa (right) or a buffer control (left). The cell nucleus (blue) is labeled with DAPI and actin (pink) is labeled with an SiR stain. Bxa is in green. Arrows indicate cells with actin-cytoskeleton disruption. 40X magnification, scale bars=10μm. **F)** Representative images of HT-29 cells treated with 200nM WT Bxa (right) or Bxa E818A/E820A (left) for 8 hrs. Actin (pink) and nuclei (blue) are shown. Arrows point to cells of interest with cytoskeleton disruption. 40X magnification, scale bars=10μm. **G)** Bxa binding measured by absorbance at 450 nm (y-axis) to an ELISA plate coated with either 10 µg/mL of GM1 ganglioside (green) or 10 µg/mL of C18:0 ceramide (black). The concentration of Bxa plated is listed from 200 to 0ng (x-axis). Data is representative of 2 independent experiments. **H)** A histogram of the mean fluorescent

intensity (MFI, x-axis) of 200nM of PE-labeled Bxa bound to HT-29 cells alone or with increasing concentrations of GM1 ganglioside. The y-axis is the number of cells. Individual points (n=6) from the histogram plot are depicted in the graph to the right, showing MFI (y-axis) according to treatment group (x-axis). Error bars represent the mean +/- SEM. Data is representative of 3 independent experiments.

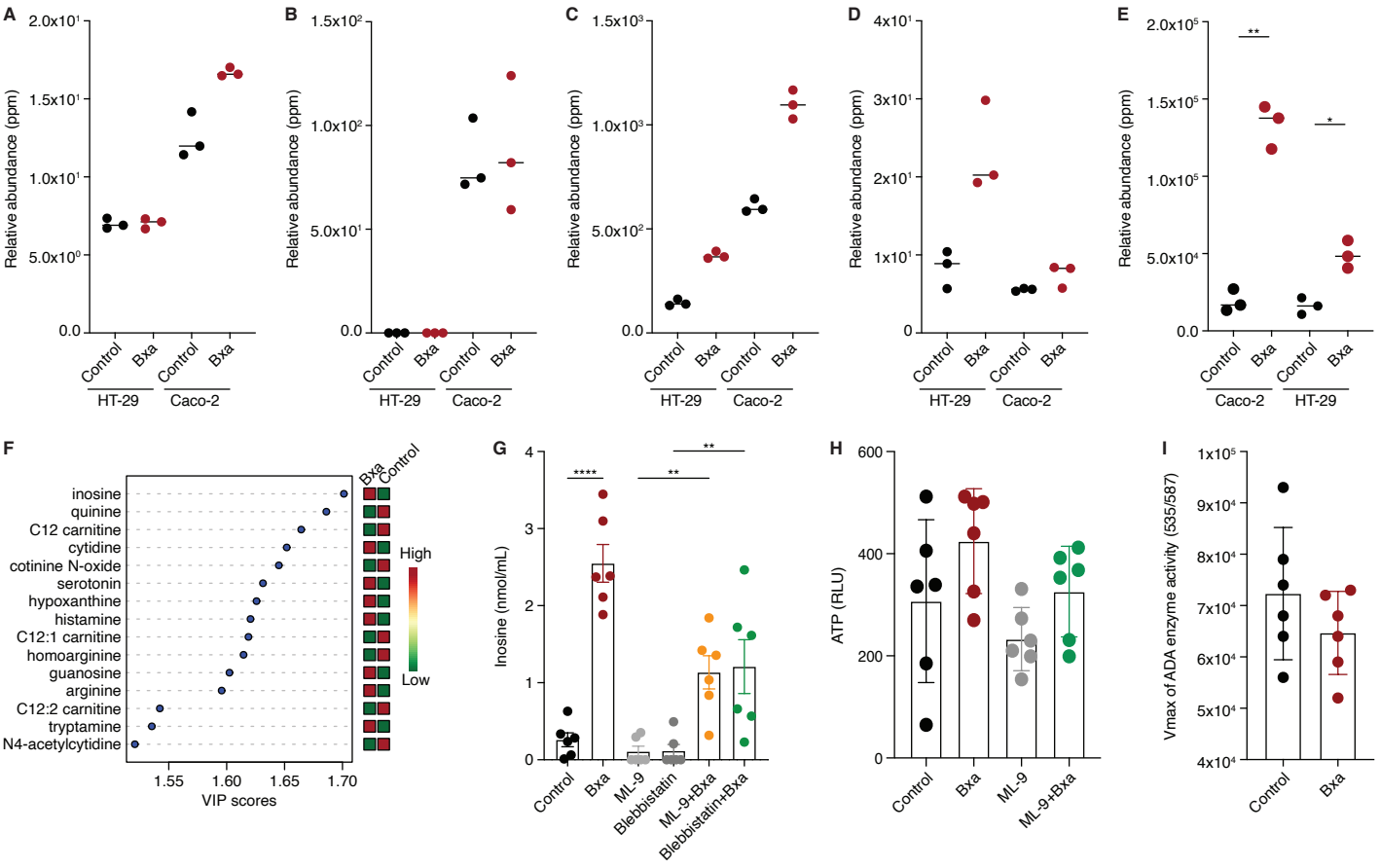

Figure S4. Bxa induces the secretion of inosine from epithelial cells, Related to Figure 5. Plots of A) cytosine, **B)** thymidine, **C)** cytidine, **D)** adenosine and **E)** guanosine relative abundance in each epithelial cell type (HT-29 and Caco-2) supernatant, comparing treatment with Bxa and buffer control. **F)** A partial least squares discriminant analysis (PLS-DA) of the top 15 most statistically significant metabolites increased (red) or decreased (green) after Bxa treatment of HT-29 cells. Metabolites are scored based on VIP (x-axis) and listed in order of significance (y-axis, from top to bottom). **G)** The concentration of inosine secreted by HT-29 epithelial cells after 200nM Bxa treatment compared to control-treated cells with or without 2 μM ML-9, a myosin light chain kinase inhibitor, or 10 μM blebbistatin, a non-muscle myosin II inhibitor. Data is representative of 2 independent experiments. **H)** ATP release from HT-29 epithelial cells measured by relative luminescence units (RLU) after 200nM Bxa treatment compared to control-treated cells with or without 2 μM ML-9. **I)** The enzymatic activity of ADA measured by Vmax during a 30 minute kinetic assay on HT-29 cells treated with 200nM Bxa or a buffer control. For all figures in this panel, *p<0.05, **p<0.01, ***p<0.001; ****p<0.0001; one-way ANOVA analysis. Error bars represent the mean +/- SEM. Representative of 2 independent experiments; n=3-6.

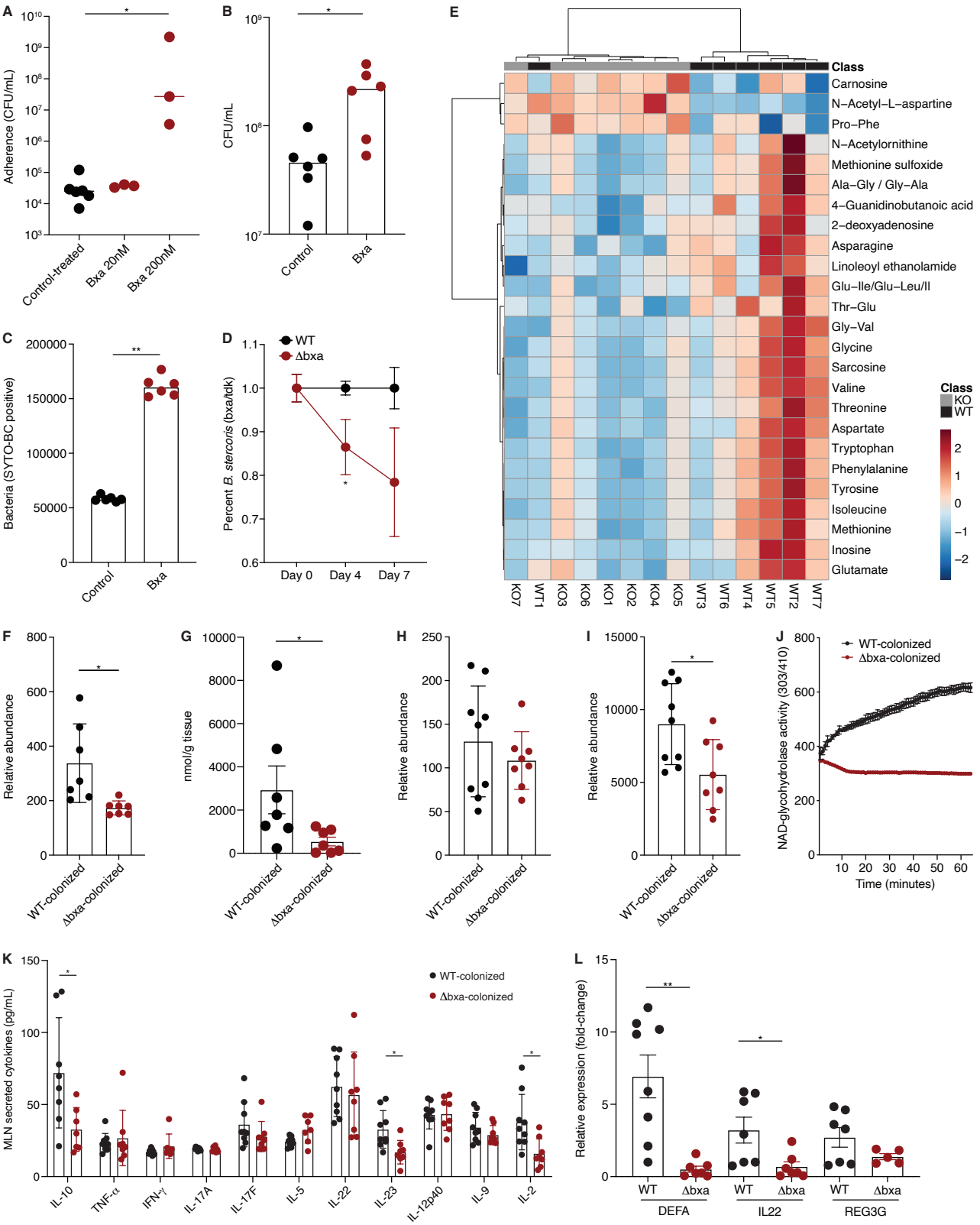

Figure S5. Presence of an ADPRT in *Bacteroides* enhances bacterial adherence to epithelial cells, biofilm formation and colonization of the intestinal epithelium, Related to Figure 6. A) A plot which shows the number of *B. stercoris* CFUs able to adhere to Caco-2 epithelial cells in either control-treated (black), 20nM Bxa-treated or 200nM Bxa-treated (red) Caco-2 cells. Data is representative of 3 independent experiments. **B)** The adherence of Δ*bxa B. stercoris* as measured by CFUs to HT-29 cells treated with Bxa or a buffer control. **C)** The adherence to *B. stercoris* with and without Bxa treatment is plotted using mean fluorescent intensity of SYTO-BC positive bacteria able to adhere to HT-29 cells. **D)** Relative abundance of WT or Δ*bxa* measured using qPCR of RNA collected from stool on days 0, 4 and 7 of a competition experiment, where each strain was administered to mice at a 1:1 ratio. The data show relative abundance of live bacteria by comparing *bxa* expression to a 16S internal control (WT strain; black) and *tdk* expression to a 16S internal control (Δ*bxa* strain; red) and normalized based on percentage of each. A higher number indicates more of the WT strain compared to the Δ*bxa* strain in the stool of mice; n=7 mice. **E)** A heat map with clustering of the top 25 differentially abundant

metabolites observed in the stool of mice mono-colonized with either WT or Δ*bxa B. stercoris*. Metabolites were increased (red) or decreased (blue) after Bxa treatment on a $\log_2$ scale. **F)** Relative abundance of inosine measured using targeted metabolomics (LC-MS) in the stool of germ-free mice colonized with either WT or Δ*bxa B. stercoris*. Representative of 2 independent experiments; n=7 mice. **G)** The concentration of inosine measured from colonic epithelial cells from the stool of germ-free mice colonized with either WT or Δ*bxa B. stercoris*. Representative of 2 independent experiments; n=7 mice. Relative abundance of **H)** inosine or **I)** guanosine measured using targeted metabolomics (LC-MS) in the stool of ASF mice colonized with either WT or Δ*bxa B. stercoris*. Representative of 2 independent experiments; n=9 mice. **J)** NAD-glycohydrolase activity detected in the stool of either WT- or Δ*bxa*-colonized mice. Supernatants from the stool were normalized for protein content from each mouse. Representative of 2 independent experiments, n=7 mice per group. **K)** The concentration of IL-10, TNF-α, IFN-γ, IL-17A, IL-17F, IL-5, IL-22, IL-23, IL-12p40, IL-9 and IL-2 measured in the supernatant of immune cells isolated from the mesenteric lymph nodes (MLNs) of germ-free mice colonized with either WT (black) or Δ*bxa B. stercoris* (red). MLN cells were plated 100,000 per well in a 96-well plate and left for 18 hrs to analyze natural secretion of cytokines. Representative of 2 independent experiments; n=8-9 mice. **L)** Relative gene expression in the proximal colonic tissue of *DEFA*, *IL22* or *REG3G* measured by qPCR in WT- or Δ*bxa*-colonized germ-free mice. All statistical analyses in this figure were performed using a one-way ANOVA for multiple comparisons or Mann-Whitney U-test for two comparisons, *$p<0.05$, **$p<0.01$. Error bars represent the mean +/- SEM.