

17 **Abstract**

18 Asthma is a common allergic airway disease that develops in association with the
19 human microbiome early in life. Both the composition and function of the infant gut
20 microbiota have been linked to asthma risk, but functional alterations in the gut
21 microbiota of older patients with established asthma remain an important knowledge
22 gap. Here, we performed whole metagenomic shotgun sequencing of 95 stool samples
23 from 59 healthy and 36 subjects with moderate-to-severe asthma to characterize the
24 metagenomes of gut microbiota in children and adults 6 years and older. Mapping of
25 functional orthologs revealed that asthma contributes to 2.9% of the variation in
26 metagenomic content even when accounting for other important clinical demographics.
27 Differential abundance analysis showed an enrichment of long-chain fatty acid (LCFA)
28 metabolism pathways which have been previously implicated in airway smooth muscle
29 and immune responses in asthma. We also observed increased richness of antibiotic
30 resistance genes (ARGs) in people with asthma. One differentially abundant ARG was a
31 macrolide resistance marker, *ermF*, which significantly co-occurred with the *Bacteroides*
32 *fragilis* toxin, suggesting a possible relationship between enterotoxigenic *B. fragilis*,
33 antibiotic resistance, and asthma. Lastly, we found multiple virulence factor (VF) and
34 ARG pairs that co-occurred in both cohorts suggesting that virulence and antibiotic
35 resistance traits are co-selected and maintained in the fecal microbiota of people with
36 asthma. Overall, our results show functional alterations via LCFA biosynthetic genes
37 and increases in antibiotic resistance genes in the gut microbiota of subjects with
38 moderate-to-severe asthma and could have implications for asthma management and
39 treatment.

40 **Keywords:** asthma, gut microbiome, shotgun metagenomics sequencing, antibiotic
41 resistance, adults, children

42

43 **Background**

44 Asthma is a common respiratory disease characterized by symptoms of airway
45 obstruction including wheeze, cough, and shortness of breath. In most cases, asthma
46 onsets in early childhood with the development of sensitization to environmental
47 allergens. Ongoing environmental exposures lead to airway inflammation and ultimately
48 result in asthma symptoms manifesting within the first few years of life. Recent findings
49 support the notion that asthma develops in association with the human gut microbiome
50 composition early in life[1, 2]. This finding is supported by 16S rRNA sequencing
51 surveys demonstrating that alterations in the gut microbiota precede asthma
52 development within the first few months of life[1, 3].

53 Early childhood gut microbial communities have been proposed to contribute to
54 asthma by several mechanisms. Epoxide hydrolases encoded by enterococci and other
55 gut bacteria produce the lipokine 12,13-diHOME that predisposes towards atopic
56 sensitization and asthma[3, 4]. Similarly, short-chain fatty acids (SCFAs), produced by
57 the metabolism of dietary fibers by diverse members of the gut microbiota, are thought
58 to protect from asthma through their effect on the host G-protein coupled receptor
59 GPR41, shaping immune cell differentiation in the lungs, and ameliorating allergic
60 airway inflammation[1, 5–8].

61 In addition to microbially-encoded metabolic features, carriage of antibiotic
62 resistance genes (ARGs) within the gut microbiota, termed the resistome, has been

63 associated with asthma risk. In infants, microbial signatures associated with the
64 development of asthma are also associated with increased richness of ARGs in the gut
65 microbiome[9]. These differences in ARG carriage were found to be driven primarily by
66 *E. coli*, which is a common colonizer in the first days of life[9]. These findings are
67 important in understanding the origins of asthma since antibiotic exposure correlates
68 both to the number of ARGs within the gut microbiome[10] and the later development of
69 asthma and other allergic diseases[11–13]. This association between antibiotic
70 exposure and asthma is supported by animal models that found antibiotic treatment
71 worsens allergic airway inflammation (AAI)[14–16].

72 While there is an abundance of data supporting the idea that asthma
73 susceptibility is associated with features of the gut microbiota in early childhood, the
74 potential effect of gut microbial functions on asthma later in life remains an important
75 knowledge gap. Since asthma often begins in infancy when the gut microbiota
76 composition is highly unstable, disease-causing microbial functions may not persist into
77 older children and adults. Nevertheless, the gut microbiota in older individuals could
78 underlie the variable manifestations of asthma[17] and may hold valuable prognostic
79 and therapeutic significance.

80 Asthma-associated differences in later childhood and adult gut microbial
81 communities have already been noted in several reports. Studies in preschool-aged
82 children have noted distinct taxonomic composition of gut microbial communities in
83 asthmatic subjects compared to healthy controls[2]. These differences are reported to
84 include reductions in *Akkermansia muciniphila*[18], *Faecalibacterium prausnitzii*[19] as
85 well as *Roseburia* species[20]. Functional characterization of microbial communities by

86 whole metagenomic sequencing from an older population of asthmatic women[19] has
87 shown that pathways related to lipid and amino acid metabolism, as well as
88 carbohydrate utilization were enriched in asthmatics. In contrast, microbial pathways
89 involved in the production of SCFAs, like butyrate, were enriched in the healthy cohort
90 of the same study[19]. These findings are supported by a complementary study
91 designed to test the effect of probiotic supplementation on asthma that found an
92 association of improved asthma symptoms with SCFA biosynthesis as well as
93 tryptophan metabolism pathways in the adult gut microbiota[21].

94 Here, we describe an analysis of whole metagenomic sequencing data from a
95 cohort of 36 subjects with physician-diagnosed, moderate-severe asthma along with a
96 matched cohort of 59 healthy controls. This study tests the hypothesis that the gut
97 metagenome harbors signatures of asthma later in life. Our results identify global
98 differences in metagenomic functions between healthy and asthmatic subjects and
99 reveal an enrichment in long-chain fatty acid biosynthetic pathways. We also find an
100 increased richness of ARGs in asthmatics and co-occurrence of ARGs with known
101 bacterial virulence factors, suggesting a potential relationship between antibiotic
102 exposure and pathogen colonization in asthmatics.

103

104 **Methods**

105 **MARS Study Population**

106 The Microbiome and Asthma Research Study (MARS) consisted of 104 subjects
107 from the St Louis, MO USA area that are either healthy or had physician-diagnosed
108 moderate-to-severe asthma. This study included an adult cohort (ages 18-40 years) and

109 pediatric cohort (ages 6-10 years). As described in previous manuscripts[22, 23], 9
110 patients were disqualified or did not donate stool samples. The remaining 95 patients
111 donated stool samples either at home or at the recruitment visit and were evaluated with
112 a clinical questionnaire to gather relevant metadata. Stool samples were kept at -20°C
113 and delivered within 24 hours to the study site, Kau Lab at Washington University
114 School of Medicine, where they were stored at -80°C for no more than three years until
115 processing for DNA isolation. All recruitment, follow up, and sample acquisition occurred
116 between November 2015 and December 2017.

117

118 **Fecal DNA Isolation**

119 Frozen human stool samples were pulverized in liquid nitrogen using a pestle
120 and mortar. We then homogenized the stool in a mixture of phenol, chloroform, and
121 isoamyl alcohol with a bead beater using sterilized zirconium and steel beads as
122 previously described[24] to extract crude DNA. We purified the fecal DNA with a 96-well
123 QIAGEN PCR Clean up kit and quantitated by measuring the absorbance at 260/280
124 nm. Sample DNA concentrations were normalized to 0.5 ng/mL. Neither depletion of
125 human DNA sequence nor enrichment of microbial or viral DNA was performed. No
126 experimental quantification like a spike-in were used.

127

128 **Whole Metagenomic Sequencing of Fecal Communities**

129 To generate fecal metagenomic sequencing data, we adapter-ligated libraries by
130 tagmentation using an adaptation of the Nextera Library Prep kit (Illumina, cat. No. FC-
131 121-1030/1031)[25]. Individual libraries were then purified with AMPure XP SPRI beads,

132 quantitated using Quant-iT (Invitrogen, cat. Q33130), and then combined in an
133 equimolar ratio. We confirmed that each library was adequately represented in the
134 combined library by preliminary sequencing on a MiSeq instrument at the Washington
135 University in St. Louis Center for Genome Sciences to assess the evenness of the
136 library. Once the quality of the library was assured, we sequenced the combined library
137 on a NovaSeq 6000 S4 with 2x150 bp chemistry to achieve an average of 3.4 Giga-
138 base-pairs (Gb) per sample. NovaSeq services and data demultiplexing were performed
139 by the Genome Technology Access Center at the McDonnell Genome Institute (St
140 Louis, MO). All samples were tagmented simultaneously and sequenced on the same
141 run to avoid batch effects.

142

143 **Processing of sequencing data**

144 Metagenomic raw demultiplexed reads were then processed to (1) remove
145 spurious human sequences (human reference database was hg37dec_v0.1.1), (2)
146 remove low quality sequences, and (3) trim remaining adapter content using Kneaddata
147 v. 0.10.0 (huttenhower.sph.harvard.edu/kneaddata) bypassing the tandem repeat finder
148 step (" -bypass-trf"). FastQC (fastqc v0.11.7) and MultiQC (multiqc v1.2) with default
149 settings were used to create quality reports and visualize processing steps. See Figure
150 S1A and Table S1 for number of reads dropped per processing step. After trimming and
151 filtering, no samples had adaptor content, overrepresented sequences, or an average
152 sequence quality score below Phred 24. Estimated metagenome coverage was
153 calculated with Nonpareil[26, 27] (version 3.4.1) via the online querying tool at
154 <http://enve-omics.ce.gatech.edu/nonpareil/submit>.

155

156 **Read-based metagenome profiling**

157 To obtain functional information about the metagenomic contents of fecal
158 samples, we processed samples using HUMAnN[28] v3.0.0 on filtered reads with
159 default parameters. The marker gene database used by HUMAnN to identify taxonomic
160 identities was ChocoPhlAn v201901b and the protein database used by HUMAnN to
161 identify functions was the UniRef90 full database v201901b. Alpha diversity analysis of
162 Uniref90 genes and two-sample tests of KEGG orthologs were performed on respective
163 genes that were present (>0 copies per million) in at least 16 out of 95 samples, which
164 was the lowest prevalence cutoff that would allow for Bonferroni corrected Wilcoxon p-
165 values below 0.0001. HUMAnN was used to determine the abundance of metagenomic
166 pathways by mapping UniRef90 genes to the MetaCyc database. We performed
167 differential abundance analysis using the Wilcoxon 2-sample tests on pathways that had
168 a minimum of 10% prevalence.

169 To identify antibiotic resistance genes present in the fecal metagenomes of
170 MARS stools, we used ShortBRED-identify[29] (v0.9.4) with the Comprehensive
171 Antibiotic Resistance Database[30] (downloaded 2021-07-05 16:10:04.04555) and
172 Virulence Factor Database[31] (downloaded Fri Jul 16 10:06:01 2021). ShortBRED-
173 Quantify was run on the filtered reads with default parameters. ARGs or VFs that had
174 an abundance greater than zero in less than 7 out of 95 samples were excluded from
175 downstream analyses. This prevalence cutoff was determined using the binomial
176 distribution to maintain a 95% confidence that enrichment was not due to random
177 chance (using stats::binom in R). In the analyses that compared virulence factor profiles

178 to antibiotic resistance gene profiles, any gene with the same name was excluded from
179 the list of antibiotic resistance and considered a virulence factor only, to prevent
180 spurious results due to co-correlations. Only one gene matched this criterion: *ugd*
181 (UDP-glucose 6-dehydrogenase).

182 Microbial composition was determined with MetaPhlAn 3.0[28] which is included
183 in the HUMAnN pipeline described[28]. MaasLin[32] (Maaslin2_1.5.1) was used in R to
184 find taxa of any taxonomic level that correlated with asthma by setting asthma as a fixed
185 effect and setting age group and race as random effects.

186 For PERMANOVA analyses, BMI class refers to two stratifications: Non-obese
187 (underweight, healthy, or overweight) and obese determined for adults by BMI cutoffs
188 and for pediatric patients by BMI-for-age percentile as defined by the Centers for
189 Disease Control and Prevention (see
190 [cdc.gov/healthyweight/assessing/bmi/childrens_bmi/about_childrens_bmi.html](https://www.cdc.gov/healthyweight/assessing/bmi/childrens_bmi/about_childrens_bmi.html)). Race
191 was reported by the subject and split into the two categories of Caucasian and non-
192 Caucasian.

193

194 **Metagenome Assemblies**

195 Filtered reads were assembled into contigs using spades[33] (v3.14.0) with the
196 “meta” flag and k-mers lengths as follows: -k 21,33,55,77. The resulting scaffolds
197 achieved an average N50 of 3525 \pm 178 bp, an average L50 of 7192 \pm 372 and an
198 average total length of 136.8 \pm 4.5 Mbp as measured by QUAST (v 4.5) [34, 35] (see
199 Table S1). Determination of *ermF* location was performed by aligning the 801-bp coding
200 sequence of *ermF* from CARD[30] to all scaffolds. Scaffolds containing BLAST hits with

201 98% identity or higher to the full-length CARD *ermF* sequence were further annotated
202 by Prokka (v1.14.5) to find open reading frames and annotate them. Manual BLAST
203 was used to annotate “hypothetical protein” open reading frames for the contexts of
204 *ermF* hits.

205

206 **Statistics**

207 R version 3.6.3 was used for all analyses downstream of HUMAnN and
208 ShortBRED, and for data visualization. Wilcoxon tests with false discovery rate multiple
209 testing correction or Type II ANOVAs were used to determine statistically significant
210 differences with the `car::Anova` package in R. PERMANOVAs were performed in R
211 using the `vegan::adonis` package with default settings and 100,000 iterations. The
212 following symbols were used to designate significance: * $p < 0.05$, ** $p < 0.01$, *** $p <$
213 0.001 and the following for q values (FDR-adjusted p-values): * $q < 0.2$, ** $q < 0.05$.

214

215 **Results**

216 **Whole metagenomic shotgun sequencing of fecal samples from adults and** 217 **children with asthma and healthy controls**

218 We performed whole metagenomic sequencing on fecal samples from asthmatic
219 subjects and healthy controls taking part in the Microbiome & Asthma Research Study
220 (MARS), which we have previously described[22, 23]. MARS participants were recruited
221 from the St. Louis, Missouri area and included pediatric (6-10 years) and adult (18-40
222 years) age groups. All asthma cohort patients had a physician diagnosis of moderate-to-
223 severe asthma, and history of allergic sensitization as evidenced by positive skin testing

224 or serum specific-IgE to one or more common aeroallergens. In total, we analyzed 95
225 patient stool samples including 17 adults and 19 school-aged participants with asthma,
226 and 40 adults and 19 school-aged participants without asthma.

227 NovaSeq S4 sequencing of our libraries yielded 1.69 billion paired-end reads
228 translating to a total of approximately 500 Gigabases (Gb). After filtering for read quality,
229 dropping host contaminants, and trimming adaptor content, we achieved 1.23 billion
230 paired-end reads and an average 3.4 Gb per stool sample with a range of 0.4-9.9
231 Gb/sample (Figure S1A). Neither host contamination nor sequencing depth differed
232 between asthma and healthy cohorts (t-test $p=0.2$ and 0.7 , Table S1). All samples
233 achieved an estimated average metagenomic coverage of at 89% (range of 61-98%)
234 with the annotation-free redundancy-based metagenome coverage estimator,
235 Nonpareil[26] (Figure S1B). Further, estimated metagenome coverage was not different
236 between the asthma and healthy cohorts, although we noted coverage was slightly
237 reduced in the pediatric cohort (Figure SB, Table S1). We employed the read-based
238 annotation pipeline, HUMAnN[28] to determine the abundance of genes and functional
239 pathways in the stool metagenomes. We found that the most abundant functional
240 pathways (Figure S1C) across all MARS participants are involved in essential
241 processes of gut microbes such as starch degradation and glycolysis, demonstrating
242 that our sequencing captured core functions of the gut metagenome, as expected.
243 Taken together, we concluded that our sequencing is of sufficient depth and quality to
244 be used for further analyses.

245

246 **Gut taxonomic composition differs between people with and without asthma**

247 We first leveraged the clade marker annotation tool, MetaPhlAn[28], to analyze
248 the taxonomic composition of the study participants. We found dominate genera typical
249 in gut microbiota communities including *Bacteroides* (phylum Bacteroidota) and
250 *Faecalibacterium* (phylum Bacillota) (Figure S1D). Simpson alpha diversity was slightly
251 higher in the asthma cohort even when taking read depth and age group into account
252 (Figure S1E). Bray-Curtis dissimilarity (Figure 1F) was shifted between the asthma and
253 healthy cohorts ($p < 0.0004$, $R^2 = 0.029$) even when accounting for other covariates
254 including age ($p < 0.001$, $R^2 = 0.032$), race ($p = 0.0006$, $R^2 = 0.026$), recent antibiotic usage
255 ($p = 0.9$, $R^2 = 0.006$), read depth ($p = 0.2$, $R^2 = 0.013$), obesity ($p = 0.7$, $R^2 = 0.008$), sex ($p = 0.4$,
256 $R^2 = 0.011$), and tobacco exposure ($p = 0.2$, $R^2 = 0.012$) by sequential PERMANOVA
257 (Figure 1G). There was also no significant interaction between asthma status and age
258 group ($p = 0.8$, $R^2 = 0.007$), or between asthma status and recent antibiotic usage ($p = 0.6$,
259 $R^2 = 0.009$) (Figure S1G). To determine differentially abundant taxa, we tested the fixed
260 effect of asthma along with the random effects of age group and race in a general linear
261 model[32] and found *Eubacterium rectale* and *Prevotella copri* were enriched in the
262 healthy cohort (Figure S1H, Table S2). All of these findings are consistent with 16S
263 rRNA sequencing performed in a previous study[23] which lent us further confidence
264 that our sequencing data was suitable for functional profiling.

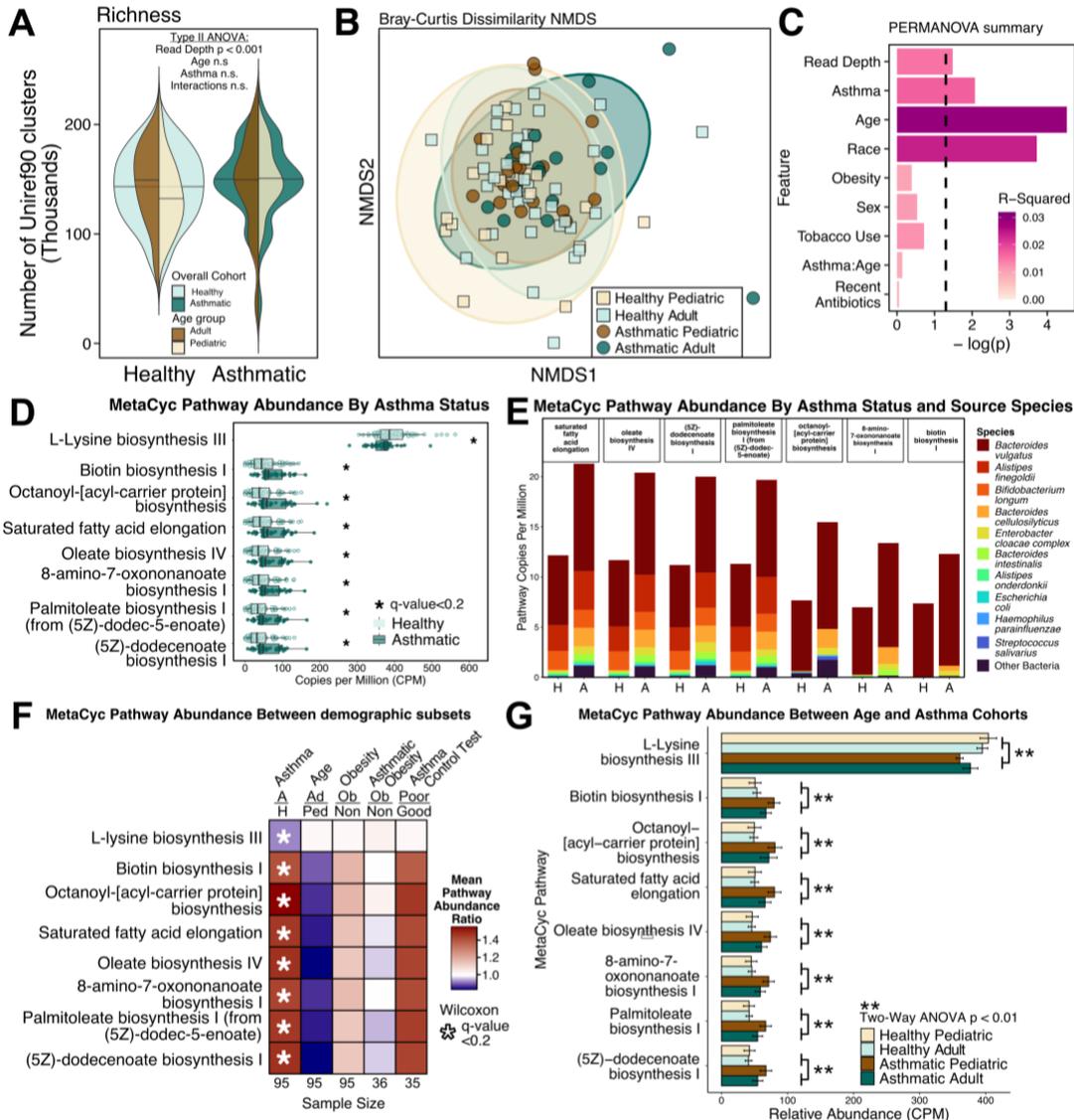
265

266 **Fatty acid metabolism pathways are enriched in the gut metagenomes of people** 267 **with asthma**

268 Given that our samples had adequate coverage to capture expected taxonomic
269 shifts, we started interrogating the differences in metagenomic functions of the gut

270 microbiota attributable to asthma status. The alpha diversity of genes (UniRef90
271 clusters) was neither different between the asthma and healthy cohorts nor between the
272 pediatric and adult cohorts, suggesting that our gene profiling reached a similar total
273 number of genes in both cohorts (Figure 1A). Using PERMANOVA, we noted that, even
274 while accounting for significant covariates of age ($p < 0.001$, $R^2 = 0.029$), race ($p < 0.001$,
275 $R^2 = 0.024$), and read depth ($p = 0.03$, $R^2 = 0.015$), asthma status also significantly
276 impacted gut microbiome functional composition ($p = 0.008$, $R^2 = 0.017$; Figure 1B, C). We
277 note that age group's interaction term with asthma did not significantly contribute to the
278 variance in beta diversity, suggesting that the influence of asthma and age on beta
279 diversity is non-overlapping. These findings support the idea that the gut metagenomic
280 content of people with asthma is different than that of healthy individuals, even when
281 accounting for other clinical sources of interpersonal gut microbiome variation.

282 We next considered which metagenomic functions and metabolic pathways may
283 be involved in the differences between asthma and healthy cohorts. We first examined a
284 list of specific metagenomic functions previously implicated in asthma, including genes
285 related to histamine production, 12-13 diHOME biosynthesis, and tryptophan
286 metabolism, but we were unable to identify a difference between cohorts (Figure S2A).
287 To identify pathways that differed between asthma and healthy subjects, we performed
288 a Wilcoxon Rank Sum test with a false discovery rate $q < 0.2$ on the relative abundance
289 of all pathways annotated by the MetaCyc database that were above 10% prevalence
290 within the population. Using these criteria, we found seven pathways that were enriched
291 in asthma and one that was enriched in the healthy cohort out of 312 total pathways
292 (Figure 1D). To determine if these findings were robust to other analysis methods, we



293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309

Figure 1: Gut metagenomes from individuals with asthma show increased genes encoding fatty acid metabolism. **A)** Stacked violin plots of Uniref90 cluster richness (unique Uniref90 cluster with CPM>0) grouped by either healthy and asthma cohort (blue green colors in background) or age (brown colors in foreground). **B)** Non-metric multidimensional scaling plot of Bray-Curtis Dissimilarity distance between Uniref90 (copies per million) profiles. Axis 1 and 2 of five total are shown of an NMDS with stress value 0.09. **C)** Sequential PERMANOVA of Bray-Curtis dissimilarities between Uniref90 profiles. Input order of terms to the test is identical to the order of the barplot from top to bottom. **D)** Relative abundance of MetaCyc pathways that were differentially abundant given a Wilcoxon q value below 0.2 (p-value after FDR correction). **E)** Stacked bar plot of differentially abundant fatty acid metabolism pathways mapped to respective taxa by MetaPhlan3.0/HUMAnN3.0, averaged within asthma or healthy cohorts. **F)** Heatmap of MetaCyc pathway abundance ratios between groups in important clinical demographics: Asthma vs. Healthy, Adult vs. Pediatric, Obese vs Non-Obese, and Well-Controlled Asthmatics vs. Poorly-Controlled Asthmatics. Asterisk denotes a significant differential abundance ($*q < 0.2$) according to Wilcoxon tests controlled for multiple comparison testing within each demographic category. **G)** Differentially abundant MetaCyc pathways plotted as four cohorts: asthma by age with respective Two-Way ANOVAs. Only statistically significant p values shown.

310 performed additional differential abundance approaches on the 312 MetaCyc pathways,
311 including a Wilcoxon test on centered log-transformed counts and ALDEX2, both of
312 which demonstrated that these pathways differed between healthy and asthmatic
313 cohorts (See Table S3). All differentially abundant pathways enriched in patients with
314 asthma were involved in fatty acid synthesis, and included the production of oleate,
315 palmitoleate, (5Z)-dodecenoate, 8-amino-7-oxononanoate, biotin, and octanoyl acyl-
316 carrier protein, as well as saturated fatty acid elongation. In the healthy cohort, only a
317 single L-lysine biosynthesis pathway was enriched.

318 Using taxonomically tiered functional mapping, we determined which taxa were
319 driving the observed differences in asthma-associated pathways. For the L-lysine
320 biosynthesis III pathway which was more abundant in healthy subject, we found that it
321 primarily originated from *Blautia obeum*, Figure S2B). In the case of the asthma-
322 enriched pathways, we found that *Bacteroides vulgatus* and *Alistipes finegoldii* account
323 for the largest fraction of complete fatty acid biosynthesis pathways (Figure 1E, Figures
324 S3C). However, the differential abundance of these asthma-associated pathways was
325 probably not due solely to an enrichment of *B. vulgatus* or *A. finegoldii* in asthma stool
326 since neither species was differentially abundant (maaslin2 q-value=0.58 and 0.25,
327 respectively; See Table S2). Further, the majority of mapped pathways were not
328 attributable to any single species and these unmapped pathway counts made up more
329 of the overall pathway richness than *B. vulgatus* (Wilcoxon q values < 0.05 for all seven
330 pathways; see “Community” stratification in Figure S2C). Taken together, these findings
331 indicate that the differences may be either driven by community-level effort (i.e. distinct
332 steps of the pathway are encoded across more than one species), or that current

333 databases are insufficiently granular to identify the key taxa responsible for these
334 differences.

335 We reviewed the enzymatic steps of each of the eight pathways represented in
336 Figure 1D and found that, of the 78 total reactions in these pathways, only 11 reactions
337 were shared between 2 pathways (Figure S3). The 8-amino-7-oxononanoate
338 biosynthesis I pathway consists of the first 11 reactions of the larger biotin biosynthesis
339 pathway and the latter only has four additional reaction steps past synthesizing 8-
340 amino-7-oxonanoate to produce biotin. Additionally, the (5Z)-dodecenoate pathway can
341 feed directly into the palmitoleate biosynthesis pathway, and that the octanoyl acyl
342 carrier protein pathway shares an upstream substrate (acetoacetyl-acyl carrier protein)
343 with the saturated fatty acid elongation pathway (Figure S3). Together, our findings
344 indicate that long chain fatty acid biosynthesis is differentially abundant in the asthma
345 gut metagenome via related but largely non-redundant pathways.

346 Given the association between obesity with fatty acid metabolism[36] as well as
347 asthma[37–39], we next wanted to determine whether obesity (which we define here as
348 a BMI greater than 30 in adults or a BMI-for-age percentile of greater than 95% in
349 children) confounds the association of microbial fatty acid metabolism with asthma. We
350 compared the abundance of the differentially abundant fatty acid pathways between all
351 non-obese and obese patients and found no significant difference (Figure 1F). Within
352 the asthma cohort, there was similarly no statistically significant difference between the
353 asthmatic obese and asthmatic non-obese patients, suggesting that obesity is not a
354 confounder for the difference we observed in fatty acid metabolism. To determine
355 whether fatty acid metabolism is related to the intensity of asthma symptoms and their

356 effect on everyday life activities, we utilized a validated survey of asthma control (The
357 Asthma Control Test; ACT)[40]. None of the fatty acid pathways were differentially
358 abundant between well-controlled and poorly-controlled asthmatics (Figure 1F). We
359 tested if age group affects the differentially abundant metabolic pathways and found that
360 these pathways were not differentially abundant between age groups alone (Figure 1F).
361 We also tested the impact of asthma and age as independent variables to differentially
362 abundant metabolic pathways using a Two-way ANOVA. We found that, even while
363 taking age into account, these pathways are differentially abundant between asthma
364 and healthy cohorts, but are not different by age or an interaction between asthma and
365 age (Figure 1G, 2-Way ANOVA). Given that the effect of asthma status on differentially
366 abundant metagenomic functions was distinct from that of age, we primarily focused our
367 subsequent analyses on the asthma and healthy cohorts overall, combining age groups.

368

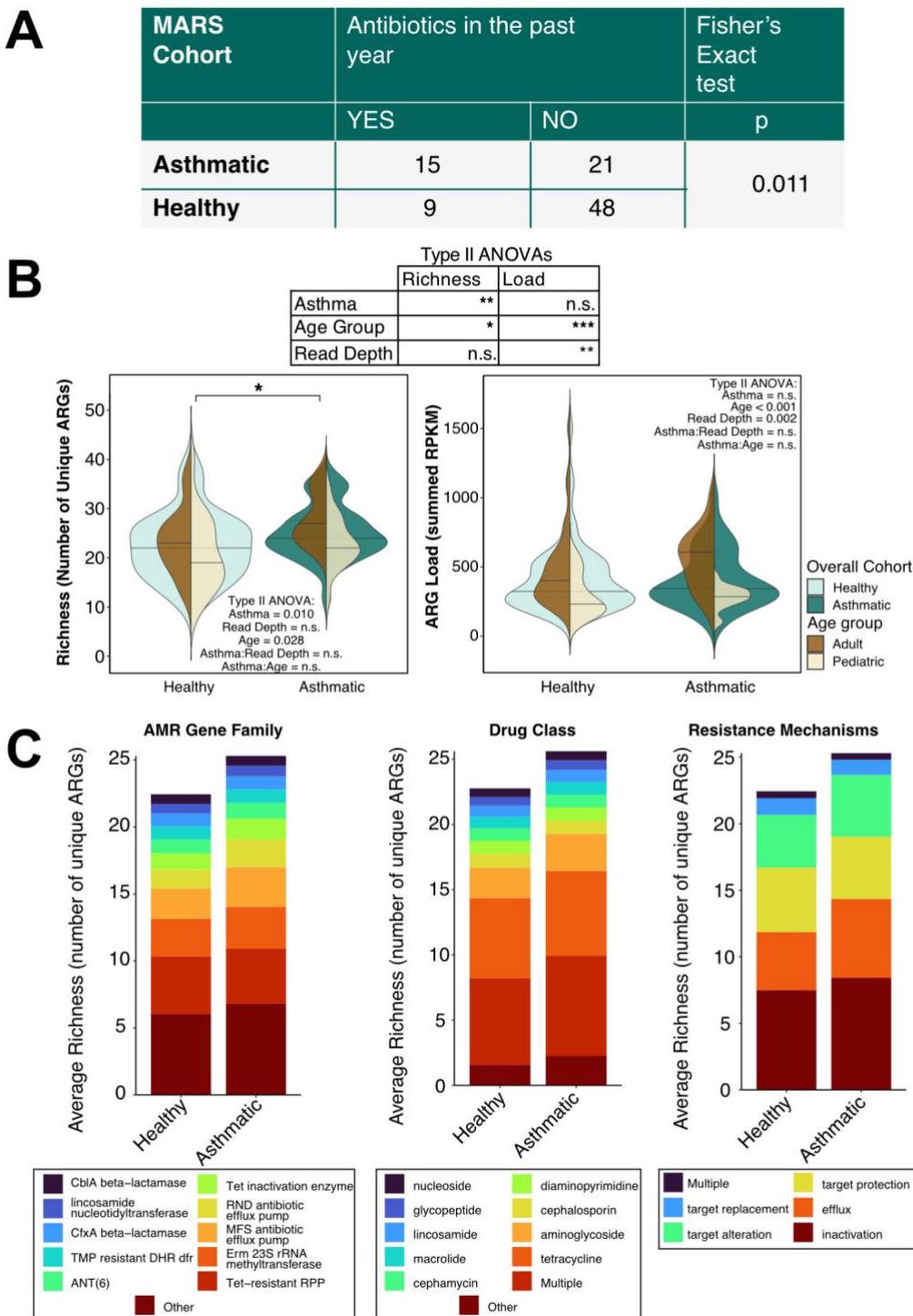
369 **Richness of antibiotic resistance genes is increased in the gut metagenomes of** 370 **people with asthma**

371 Since people with asthma tend to be prescribed antibiotics frequently[41] and
372 oral antibiotic exposure is a risk factor for the acquisition of ARGs in the gut[10], we
373 wanted to determine if the members of our asthma cohort were more likely to have
374 received antibiotics. To test this, we counted how many subjects had taken a course of
375 antibiotics within one year of their participation in the study. As part of the study design,
376 participants could not take antibiotics in the month prior to fecal donation. We found that
377 a greater proportion of the asthma cohort received antibiotics in the past year compared
378 to that of healthy participants (42% of asthma cohort versus 15% of the healthy cohort,

379 Fisher's test, $p=0.011$, Figure 2A). This finding represents evidence of increased
380 antibiotic exposure amongst subjects with asthma in our study.

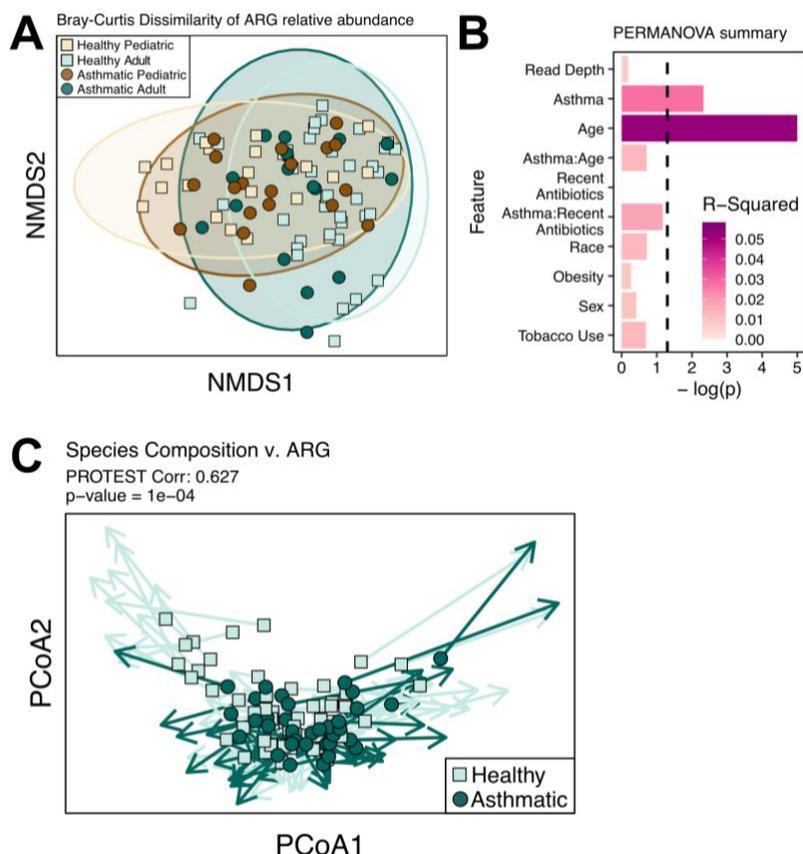
381 We next sought to characterize the gut antibiotic resistome in the asthma and
382 healthy cohorts. To test if the increased antibiotic exposure in the asthma cohort was
383 reflected in the gut resistome, we utilized the ShortBRED pipeline[29] to detect reads
384 mapped to the Comprehensive Antibiotic Resistance Database (CARD)[30]. We first
385 asked whether there were more ARGs in our asthma cohort by summarizing our dataset
386 into richness (Total number of unique ARGs detected per sample) and load (Total sum
387 of ARG RPKM per sample). We found that ARG richness was higher in people with
388 asthma even when accounting for differences due to age ($p=0.03$) and sequencing
389 depth ($p=0.09$ while ARG load was not different between asthma and healthy cohorts
390 ($p=0.4$) when accounting for age ($p<0.001$) and read depth (0.002) (Figure 2B). We note
391 that *E. coli* was not differentially abundant between asthma and healthy cohorts ($p=0.52$,
392 Table S2), so the richness increase we observe in the asthma cohort is not due solely to
393 an increase in *E. coli* relative abundance. These results suggest that there are a higher
394 number of unique ARGs, or a higher diversity, in asthma compared to healthy controls.

395 From our 95 stool samples, we detected 71 unique ARGs, comprising 32
396 antimicrobial resistance families, 29 drug classes, and 7 mechanisms of resistance, with
397 26 ARGs (37% of the total) conferring multi-drug resistance (Figure 2C). Similar to
398 previous studies of gut resistomes, we found that tetracycline resistance markers were
399 the most commonly detected ARGs and inactivation is the most common mechanism of
400 resistance followed by efflux pumps[9] (Figure 2C). Using the abundance data of each
401 detected ARG, we determined that asthma ($p=0.005$, $R^2=0.028$) and age ($p<0.001$,



402
 403 **Figure 2: Gut metagenomes from individuals with asthma harbor an increased richness of**
 404 **antibiotic resistance genes. A)** Table describing short-term antibiotic usage in the MARS cohorts. **B)**
 405 Overlapping violin plots of ARG richness and load by grouped by either healthy and asthma cohort (blue
 406 green colors in background) or age (brown colors in foreground). **C)** Stacked bar plots of average ARG
 407 richness painted by antimicrobial family (AMR), drug class to which the ARG confers resistance, and ARG
 408 resistance mechanism.

409 $R^2=0.053$) were the strongest factors contributing to the variance in ARG beta diversity
 410 even when accounting for important technical and demographic covariates (Figure 3A
 411 and 3B). We next wanted to ascertain to what degree the resistome profile was
 412 determined by microbial composition. We used a Procrustes analysis[42] to compare
 413 compositional data generated from MetaPhlAn[28] to the antibiotic resistome profile
 414 derived from ShortBRED and found that the microbiome composition correlated to the
 415 resistome profile (Figure 3C, PROTEST corr = 0.627, p -value < 0.0001), indicating that
 416 ARG profiles are directly related to bacterial species composition.

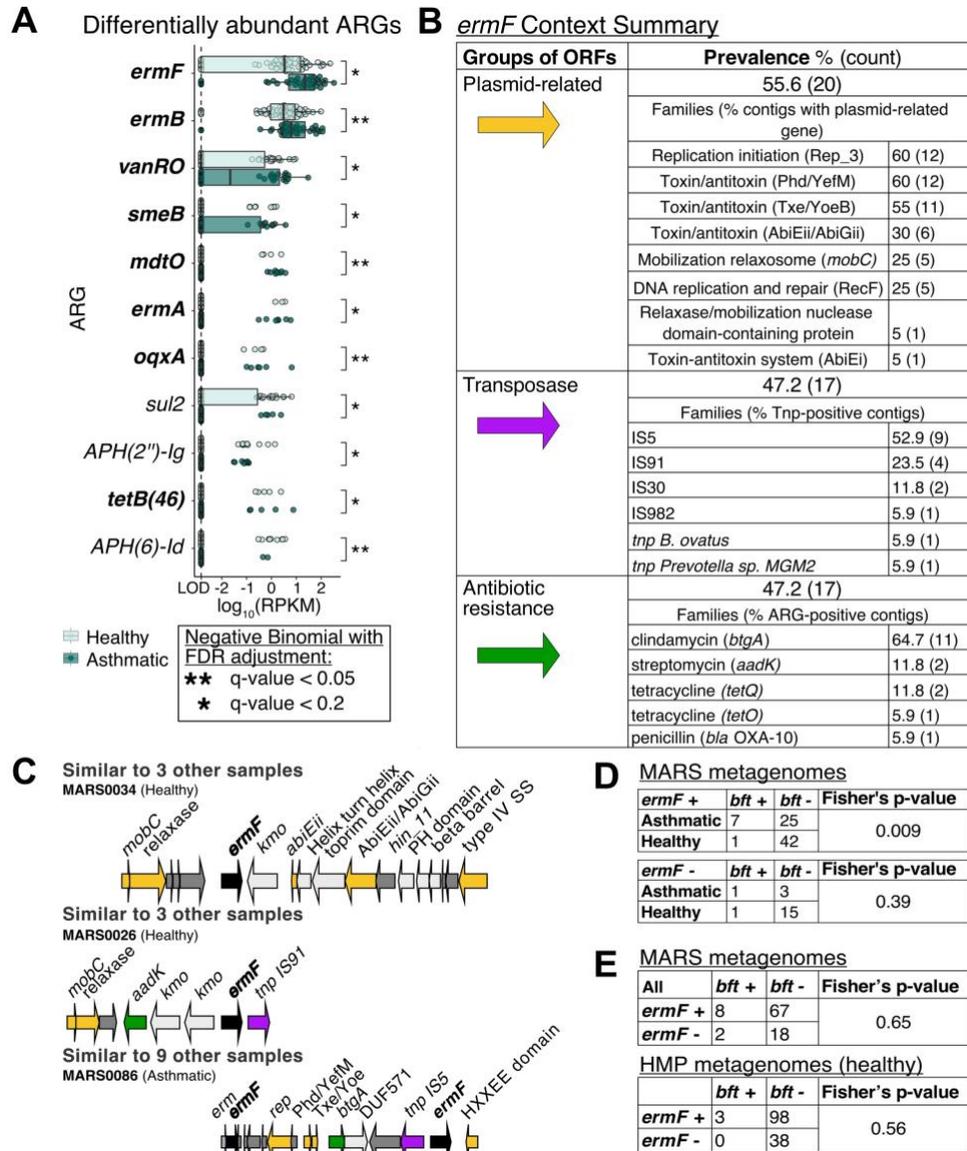


417
 418 **Figure 3: The gut antibiotic resistome is altered in asthma patients. A)** Non-metric Multidimensional
 419 Scaling (NMDS) plot of antibiotic resistome with units in Bray-Curtis dissimilarity of total-sum scaled
 420 RPKM, labeled by asthma and age cohorts. Showing two axes out of five with stress value=0.1. **B)** Effect
 421 of demographic categories on antibiotic resistome data in A (sequential PERMANOVA). **C)** Procrustes
 422 and PROTEST analysis between MetaPhlAn species-level Bray-Curtis dissimilarity distances and CARD
 423 ShortBRED Bray-Curtis dissimilarity distances. Arrows connect the two data points belonging to identical
 424 samples.

425 **Macrolide resistance markers are differentially abundant in asthma**

426 To determine ARGs that are differentially abundant between asthmatic and
427 healthy gut metagenomes, we applied negative binomial tests to the abundance of all
428 ARGs detected in at least 7 samples. This prevalence cutoff was chosen because it is
429 the minimum number of samples needed to detect a difference using a negative
430 binomial distribution. We found that genes encoding resistance to macrolides (*ermF*,
431 *ermB* and *ermA*), vancomycin (*vanRO*), tetracycline (*tet(45)*), as well as multi-drug
432 efflux pumps (*smeB*, *mdtO*, and *oqxA*) were enriched in the asthma cohort (Figure 4A,
433 Table S4). Prominent amongst these was the 23S rRNA methyltransferase *ermF*, which
434 is typically encoded by *Bacteroides* species and confers resistance to macrolides.

435 Next, we explored the genomic context of *ermF* by assembling metagenomic
436 sequencing reads into contigs with metaSPAdes[33] and annotating open reading
437 frames with Prokka[43] and BLAST. We detected full-length *ermF* with 98% or higher
438 identity in 53 out of 95 samples. Out of 53 contigs, the vast majority originated from
439 members of the Bacteroidota, 75.4% originated from the *Bacteroides* genus and 60.3%
440 of them were likely from *B. fragilis* based on the top BLAST homology. Of the contigs
441 that encoded *ermF*, 68% occurred on scaffolds with at least one other open reading
442 frame within ten kilobases (Figure 4B). We found that many *ermF* genes are co-located
443 with genes associated with mobile genetic elements such as transposases, mobilization
444 genes, and toxin/antitoxin systems, as well as with other ARGs like *btgA* which encodes
445 clindamycin resistance (Figure 4B,C). This indicates that *ermF* occurs in multiple
446 different genomic contexts within our cohort and suggests that its presence is not strictly
447 due to propagation of a single *B. fragilis* strain.



448
 449 **Figure 4: Resistance gene *ermF* is differentially abundant in diverse genomic contexts of gut**
 450 **resistomes belonging to individuals with asthma. A)** Boxplots of antibiotic resistance gene (ARG)
 451 abundance by cohort on log-scale. Showing only ARGs present in at least 7 out of 95 samples and have
 452 q-values less than 0.2. A pseudocount of 0.0015 RPKM (designated as the limit of detection "LOD") was
 453 used for the negative binomial tests. Bolded genes are enriched in the asthma cohort while non-bolded
 454 are enriched in the healthy cohort. **B)** Summary of *ermF* contexts on contigs from metagenomic
 455 assemblies that had at least one detectable open reading frame flanking the *ermF* within 10 kilobases. **C)**
 456 Three representative *ermF* context maps generated in GeneSpy. **D)** Count tables of *ermF*+ (top) and
 457 *ermF*- (bottom) MARS fecal samples split by *bft* presence and asthma status. Both tables showing two-
 458 sided p-value. **E)** Count tables of metagenomes (top from MARS and bottom from Human Microbiome
 459 Project) split by the presence of *B. fragilis* toxin (*bft*) and *ermF*. Top: one-sided p-value shown; bottom:
 460 two-sided p-value shown.

461 **People with asthma have a distinct set of co-existing pairs of antibiotic resistance**
 462 **genes and virulence factors in the gut metagenome**

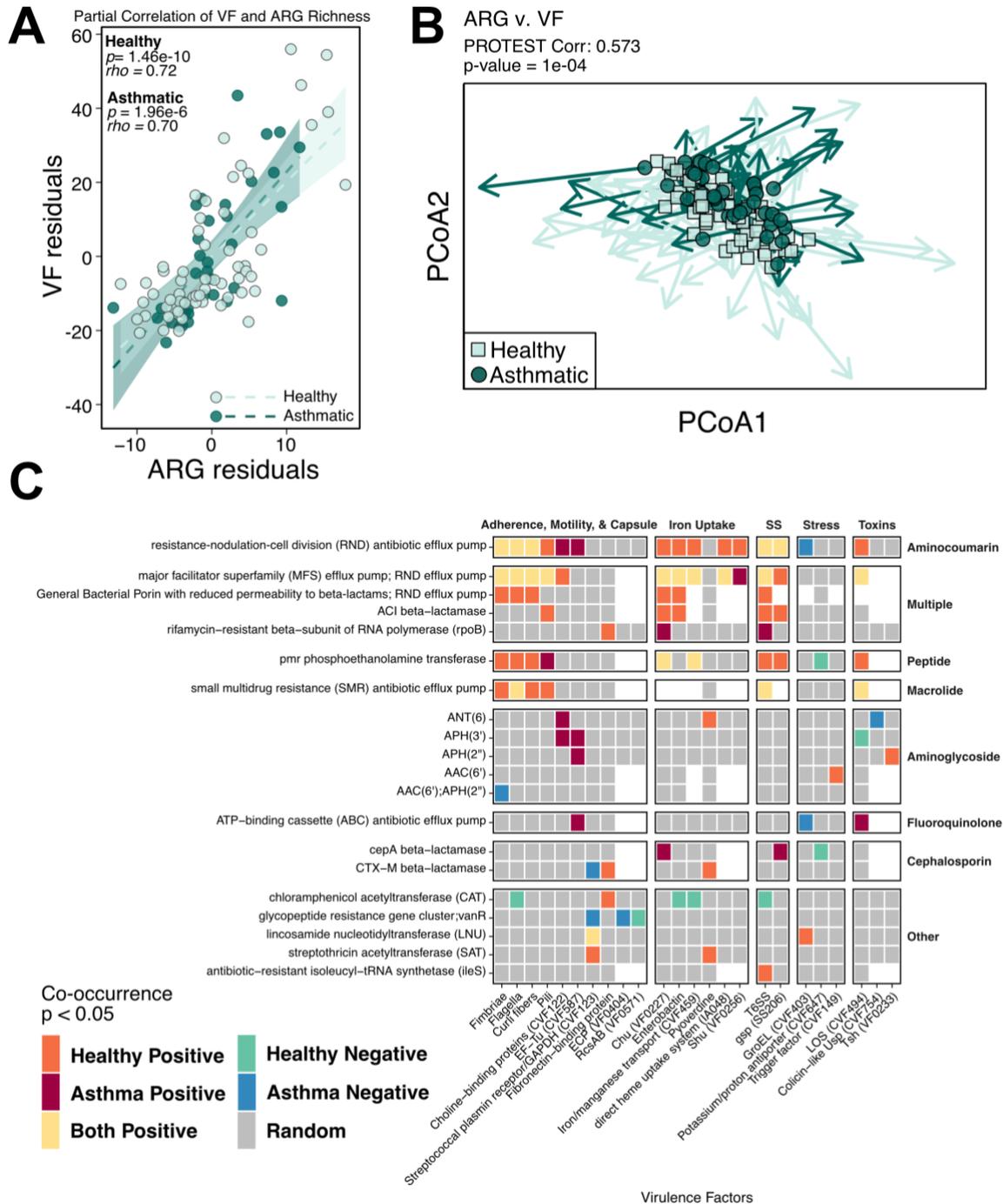
463 In our prior work on this same cohort of patients, we found that, compared to
464 healthy subjects, a greater portion of asthma subjects were colonized with *B. fragilis*
465 strains harboring the virulence factor *B. fragilis* toxin (*bft*), which we showed has the
466 potential to shape inflammation in the lung[23]. Given that our resistome analysis
467 pointed to an enrichment of a *B. fragilis* ARG, we wanted to test whether the *ermF* gene
468 is co-selected with *bft*. When only taking metagenomes encoding *ermF* into account, we
469 observed an enrichment of *bft* prevalence in the asthma cohort (Figure 4D, $p=0.009$). In
470 contrast, among metagenomes with no detectable *ermF*, there is no enrichment of *bft* in
471 the asthma cohort (Figure 4D, $p=0.39$). When reviewing the entire MARS population, we
472 found no statistically significant co-occurrence of *ermF* with *bft* (Figure 4E; one-tailed
473 Fishers test $p>0.05$) and this was consistent with healthy gut metagenomes from the
474 Human Microbiome Project (Figure 4E, Fisher's test $p=0.56$). However, in our MARS
475 samples, we did not find any instances where *bft* and *ermF* occurred on the same
476 scaffold, so it remains unclear whether these two genes are encoded within the same *B.*
477 *fragilis* strain or within two separate strains. Nevertheless, these results suggest that the
478 environment supporting the gut microbiota of asthmatic individuals presents
479 opportunities or niches for *ermF* and *bft* to co-occur.

480 To explore the possibility that virulence traits and ARGs are linked in the gut
481 microbiota, we characterized virulence factor (VF) content of all samples using the
482 Virulence Factor Database[31] and compared these data to the antibiotic resistome
483 profiles. We did not find the same overall shift in the virulence factor beta diversity
484 between asthma and healthy that we observed with the resistomes (Figure S4A-C), but
485 we did find differentially abundant VFs belonging to capsule and peritrichous flagella VF

486 families (Table S5, q values <0.2). Further, we found that microbiota composition is
487 highly correlated with virulence factor profile (Figure S4D, Spearman correlation
488 coefficient=0.61, $p<0.0001$). Given that microbiota composition strongly affects both VF
489 and ARG content, we used a partial correlation between VF and ARG richness to test
490 our hypothesis while removing the effect of total metagenomic content. We found a
491 positive partial correlation between VF and ARG richness in both the asthma and
492 healthy cohorts (Figure 5A). Similarly, virulence factor and resistome beta diversity
493 profiles were also positively correlated (Figure 5B, Spearman correlation coefficient=0.574,
494 $p=1e-4$). Together, our results suggest that these two microbial features, virulence and
495 antibiotic resistance, are closely linked within the gut metagenome.

496 We next performed a co-occurrence analysis to uncover other linked virulence
497 and antibiotic resistance traits that could be important in gut ecology. We found
498 numerous co-occurring VF-ARG pairs in MARS gut metagenomes (Figure 5C, $p<0.05$).
499 Several of these positively co-occurring pairs were shared between the two cohorts
500 (yellow), suggesting that these relationships are not dependent on asthma status. In
501 contrast, many pairs specifically co-occur in one cohort and may indicate microbial
502 interactions important in asthma but not healthy gut metagenomes (Figure 5C). In
503 summary, we found that VF and ARG presence is linked in the gut metagenome and
504 that people with asthma have a distinct set of co-occurring functions compared to
505 healthy people.

506 While our co-occurrence analysis between VFs and ARGs demonstrated multiple
507 examples of virulence and antibiotic resistance traits found in the same gut
508 metagenome, this analysis does not indicate if these genes are present in a single



509
510 **Figure 5: Asthma patients have unique sets of virulence factor and antibiotic resistance gene**
511 **associations. A)** Partial correlations split by asthma status between virulence factor richness and ARG
512 richness after accounting for species richness. **B)** Procrustes and PROTEST analysis between Bray-
513 Curtis dissimilarity distances of virulence factors and CARD resistomes. Arrows connect the two data
514 points belonging to identical samples. **C)** Heatmap of statistically significant (cooccur R package $p < 0.05$)
515 co-occurrence relationships between all VFs and ARGs. Colors indicate direction of co-occurrence and in
516 which cohort(s) the respective effect was detected. Grey squares mark pairs with no statistically
517 significant co-occurrence. White squares were pairs filtered out due to a lack of observed co-occurrence.

518 organism. To obtain a more granular view of VF-ARG co-occurrence, we limited our
519 analysis to look for VF-ARG pairs that could be encoded by the same species. This
520 analysis showed that the asthma cohort had a greater number of ARGs ($p=0.007$ and
521 0.01) and VFs ($p=0.005$ and 0.09) annotated as coming from *Klebsiella pneumoniae*
522 and *Escherichia coli*, respectively (Figure S5A). Individual co-occurrences attributable to
523 each of these species are summarized in Figure S5B and show that *cepA*, encoding a
524 beta-lactamase, and *chuU*, a VF involved in iron acquisition, are both putatively
525 encoded by *E. coli* and co-occur in asthmatics, suggesting that the metagenome-wide
526 co-occurrence of CepA and Chu families observed in Figure 5C may be due to
527 enrichment within one or more *E. coli* strains harboring these VF/ARG pairs. Together,
528 our co-occurrence analyses show that there appear to be multiple co-occurring VFs and
529 ARGs, similar to *B. fragilis*-encoded *bft* and *ermF*, in the gut metagenome and within
530 putative individual species that could be important for asthma. The cohort-specific co-
531 occurring VF-ARG pairs found here could serve as candidates for future studies of
532 asthma gut microbiome ecology.

533

534 Discussion

535 In this study, we present an exploratory analysis of fecal whole metagenomic
536 sequencing contrasting subjects with moderate-to-severe asthma to a group of healthy
537 controls to identify disease-associated microbial genes with the strongest likelihood of
538 affecting disease. Our sequencing and subsequent analyses revealed that the
539 functional content of individuals with asthma differed significantly from that of healthy
540 controls. We found an enrichment of functions associated with saturated and mono-

541 unsaturated fatty acids, including oleate, palmitoleate, 5(Z)-dodecenoate, biotin, 8-
542 amino-oxononanoate, saturated fatty acid elongation, and octanoyl acyl carrier protein
543 pathways. Currently, the functional significance of gut bacterial synthesis of these long-
544 chain fatty acids (LCFA) to asthma has not been well defined. Excess LCFAs, usually
545 studied in the context of dietary fat intake, have been associated with metabolic
546 diseases including diabetes, obesity, and atherosclerosis risk[38] but is also linked to
547 asthma risk in adults[37–39, 44]. Increasing recognition that obesity predisposes to
548 asthma has motivated investigation of the impact of fatty acids on airway biology and
549 has shown that LCFA signaling through free fatty acid receptor 1 (FFAR1, also called
550 GPR40) induces airway smooth muscle cell contraction and proliferation, both of which
551 are important components of asthma pathophysiology[38, 45]. Notably, a study that
552 sequenced airway microbes in children with cystic fibrosis implicated a similar list of
553 LCFA production pathways during exacerbations, suggesting that microbially produced
554 LCFAs may influence airway physiology[46]. To our knowledge, the potential for gut
555 microbes to contribute to the amount of free fatty acids available to the lung has not yet
556 been defined, however, LCFAs are readily absorbed into the circulation[47] and could
557 plausibly reach the airways. Further, previous studies have shown the effect of SCFA
558 (e.g. acetate, butyrate, propionate) produced by gut microbes to directly alter lung
559 inflammation via GPR41 (FFAR3)[7, 8]. While our study did not find a direct enrichment
560 of SCFA production pathways in the healthy cohort as has been previously reported[19],
561 we did observe that lysine biosynthesis was enriched. Since lysine may serve as a
562 precursor to the SCFA butyrate[48], SCFAs may still be more abundant in our healthy
563 cohort but may be subject to transcriptional regulation that would not be detected by

564 metagenomic DNA sequencing. Together, our metabolic pathway analyses of the gut
565 metagenome demonstrate a positive association between LCFAs produced by gut
566 microbes and asthma, in contrast to the negatively associated SCFAs.

567 In addition to metabolic alterations, analysis of the gut resistome demonstrated
568 that subjects with asthma had a distinct ARG composition. In a recently published
569 prospective gut metagenomic study of infants, asthma-associated taxonomic signatures
570 were associated with a higher number of ARGs[9]. These differences in the resistome
571 were largely driven by a single species of bacteria, *E. coli*, and reveals that acquisition
572 of ARGs in subjects with asthma may begin in early childhood and could affect asthma
573 development. In our study of older subjects with established asthma, we similarly found
574 a higher richness of ARGs that is associated with asthma in both school-aged children
575 and adults, supporting the idea that increased ARG carriage may persist in people with
576 asthma throughout life. Based on our resistome annotation, however, ARGs in our
577 cohort were likely from a diverse assemblage of bacteria in contrast to what was
578 observed in infants. This is likely due to differences in gut dynamics between age
579 groups. The infant microbiome is heavily shaped by limited available niches in the
580 developing gut, which favor transient, facultative anaerobes like *E. coli* [9], whereas the
581 gut resistome in older subjects reflects selective pressures experienced over a lifetime.
582 One important consequence of increased richness of ARGs in people with asthma is
583 that it may promote persistence of some bacterial strains[49, 50] and contribute to the
584 taxonomic differences in the gut microbiota between asthma and healthy people[2, 23].

585 While asthma was among the important factors accounting for a significant
586 amount of the variance in ARG beta diversity, we found that recent antibiotic exposure

587 (within the past year) was not. Notably, no participant in our cohort received a course of
588 antibiotics in the month prior to fecal sampling since this could have confounded our
589 analyses on asthma-associated microbial community changes. Previous studies have
590 shown that the gut microbiota recovers in approximately a month after perturbation from
591 antibiotics in healthy adults[51]. We interpret these findings to mean that recent
592 exposure (within 1 - 12 months) to antibiotics does not drastically change the resistome,
593 whereas repeated exposures over time may be more important for driving the
594 population-wide shifts we observed in our cohort[50].

595 Of the ARGs found to be enriched within asthmatic resistomes, the ARG *ermF*,
596 encoding resistance to macrolide antibiotics, was especially prominent amongst the
597 asthmatic cohort. While we did not collect data on the antibiotic drug classes, number of
598 courses and their duration, or the reason for prescription of antibiotics, our subjects
599 received, it is likely that our asthma population has been exposed to macrolides.
600 Macrolide antibiotics, including clarithromycin and azithromycin, are commonly
601 prescribed for upper and lower airway infections which disproportionately affect people
602 with asthma[52]. This class of antibiotics, particularly azithromycin, have been a focus of
603 special concern for driving antibiotic resistance due to their frequent usage and
604 pharmacological properties[53–55]. Nevertheless, azithromycin has been noted to have
605 beneficial effects in asthma, and some[56], but not all[57], studies suggest that
606 azithromycin may prevent exacerbations in asthmatics. Given the interest in
607 azithromycin as a treatment modality in asthma, there will be an urgent need for
608 additional studies to determine the robustness of the association between asthma and
609 macrolide ARG accumulation in the gut to inform parameters for antibiotic selection and

610 prescription in people with asthma.

611 Additional exploration of the gut metagenomes revealed potential co-selection in
612 people with asthma for *B. fragilis* genes *ermF* and *bft* (*B. fragilis* toxin), the latter of
613 which is more prevalent in fecal samples from the asthma compared to healthy
614 cohort[23]. Untargeted analysis of gut resistomes revealed multiple examples of
615 virulence factor and ARG co-occurrence as well as positive correlations between ARG
616 and VF richness in people with and without asthma. Our findings are consistent with
617 previous reports that found correlations between VFs and ARG richness and VF-ARG
618 cooccurrence relationships in both gut metagenomes[58] and human-associated
619 bacterial genomes[59]. Our findings also add to these studies by demonstrating that,
620 while the correlation between VF and ARG richness does not appear to be any stronger
621 in the asthma cohort after taking gene richness into account, the two MARS cohorts do
622 not have identical sets of statistically significant co-occurring VF-ARG pairs. These data
623 suggest that people with asthma may be experiencing different selection pressures from
624 that of healthy people, leading to accumulation of a distinct set of virulence and
625 antibiotic determinants. Given that antibiotics induce gut inflammation through the
626 disruption of the gut microbiota[60], and strains encoding virulence factors such as *bft*
627 are known to thrive in an inflammatory environment[61], one plausible model for the
628 apparent accumulation of distinct VF-ARG pairs is that antibiotic treatment not only
629 selects for ARGs[10, 50], but simultaneously selects for VFs. Together with evidence
630 that virulence determinants, such as *bft*, are associated with airway inflammation[23],
631 our model implies that heightened antibiotic treatment may contribute to the
632 manifestations of asthma via co-selection for VFs and ARGs. Considering that prenatal

633 and early life antibiotic exposure is linked to asthma risk[12, 60], this model could be
634 used to test whether the initial events driving VF and ARG co-occurrence start with the
635 first vertical transmission events in very early life.

636 Our study has several limitations that constrain the scope of our claims. First,
637 MARS is an exploratory, cross-sectional study with only a moderate number of subjects
638 recruited from a single site, which is less ideal for identifying disease-associated
639 microbiome differences[62]. As a result, our study had limited statistical power to detect
640 less prevalent or abundant functions. Second, our study focused on school-aged and
641 older subjects with moderate-to-severe asthma, and thus our findings may not be
642 applicable to other younger populations or those with less severe disease. These
643 population differences may explain why we were unable to identify statistically
644 significant differences in microbial metabolic pathways identified from other studies
645 including bile acid metabolism[1], epoxide hydrolases[4], histamine metabolism[63, 64],
646 or tryptophan metabolism[65, 66] (Figure S2A). Third, the factors driving the shift in gut
647 bacterial metabolism to LCFA biosynthesis and whether gut microbiome enrichment of
648 this pathway is sufficient to change the hosts' LCFA profile is not known. Collecting
649 blood to interrogate host metabolism as well as dietary information at the time of fecal
650 sample collection would have helped to disentangle the effects of diet on host and gut
651 microbiota metabolism. Fourth, a record of the frequency and class of antibiotics
652 administered to our participants would have allowed us to confirm whether macrolide
653 administration associates with the enrichment of *ermF* in our asthma cohort and
654 whether a higher diversity of antibiotic usage correlates with ARG richness. It is likely
655 that antibiotic exposures accumulated throughout life contribute to the resistome, and a

656 complete catalog of exposures is critical to determine patterns of antibiotic prescription
657 most likely to account for the ARG associations to asthma found in this study. Fifth, as
658 with all metagenomic sequencing studies, we are limited by annotation bias in existing
659 databases. This is a concern for our virulence factor and antibiotic resistance profiling
660 especially, where we rely on the database to predict source species for ARGs and VFs.
661 We also recognize that the databases we used for these two analyses are biased
662 towards well-studied human pathogens rather than commensals or opportunistic
663 pathogens. However, we note that other investigators have reported similar co-
664 occurrence of ARGs and VFs[58, 59], and co-selection of these features is biologically
665 plausible.

666 Despite these constraints on the scope of our study, we provide evidence that
667 there is an increased production of LCFA and an increased richness of ARGs encoded
668 by the gut microbiota in people with asthma. These findings could have applications in
669 the care of patients with asthma. If LCFA pathways are shown to play a causal role in
670 airway inflammation in future studies, microbiota-directed therapeutics in the form of
671 dietary interventions or probiotics, could be developed to modify gut microbial
672 metabolism to protect against asthma. Additionally, our resistome findings add to the
673 growing concern over antibiotic resistance in patients with asthma by suggesting that
674 antibiotic administration may also contribute to gut carriage of virulence factors that can
675 alter airway inflammation. Ultimately, our study shows that the gut microbiota of school-
676 aged and older subjects with moderate-to-severe asthma harbor important functional
677 alterations that could serve as a foundation for future studies investigating how gut
678 microbial functions affect pulmonary diseases.

679 **Conclusions**

680 Asthma is an airway disease that affects the everyday lives of millions of people
681 and accounts for approximately 1.5 million emergency room visits yearly in the US[67]
682 Both antibiotic usage and gut microbiota dysbiosis have been linked to the development
683 of asthma, however, little is known about the specific gut microbial functions associated
684 with asthma, particularly in older populations. In this study, we characterized the gut
685 microbiota of school-aged children and adults with moderate-to-severe asthma and
686 uncovered asthma-associated microbial functions that may contribute to disease
687 features. We found that people with asthma have an increase in gut microbial genes
688 associated with long-chain fatty acid metabolism as well as an accumulation of antibiotic
689 resistance genes, both of which may have practical consequences for monitoring and
690 treatment of asthma.

691

692 **List of Abbreviations**

693 Allergic airway inflammation (AAI), antibiotic resistance gene (ARG), long-chain fatty
694 acid (LFCA), short-chain fatty acid (SCFA), virulence factor (VF).

695

696 **Declarations**

697 *Ethics approval and consent to participate*

698 This study was approved by the Washington University Institutional Review
699 Board (IRB# 201412035). Written informed consent documents were obtained from all
700 MARS subjects or their legal guardians.

701 *Consent for publication*

702 Not applicable.

703 *Availability of data and material*

704 The metagenomic sequencing dataset generated during the current study are
705 available at European Nucleotide Archive (<https://www.ebi.ac.uk/ena/browser/home>)
706 under project accession number PRJEB56741. Demographic data needed to reproduce
707 results can be found in this manuscript (Table S1). A full record of all statistical analyses
708 is included as a PDF document generated by knitr in R[68] in Additional File 1. A
709 STORMS (Strengthening The Organizing and Reporting of Microbiome Studies)
710 checklist[69] is available at doi: 10.5281/zenodo.7492635.

711 *Competing interests*

712 The authors declare that they have no competing interests.

713 *Funding*

714 A.L.K. is funded by the AAAAI Foundation Faculty Development Award and the
715 NIH K08 AI113184. A.H-L received funding through the NIH T32 GM007200 and F30
716 DK127584. The funding bodies listed here were not involved in the design of the study
717 and collection, analysis, and interpretation of data and writing of the manuscript.

718 *Author Contributions*

719 N.G.W. and A.L.K. conceptualized the work. L.B.B. and A.L.K. planned the
720 clinical study. N.G.W., A.H-L., D.J.S., and A.L.K. analyzed the data and drafted the
721 manuscript. All authors interpreted the data, and read and approved the manuscript.

722 *Acknowledgements*

723 We would like to thank our clinical study coordinators Tarisa Mantia, Caitlin
724 O'Shaughnessy, and Shannon Rook; the physicians of Washington University Pediatric

725 and Adolescent Ambulatory Research Consortium, especially Dr. Jane Garbutt; the
726 Volunteer for Health registry; and the MARS participants and their families. We would
727 also like to acknowledge the Center for Genome Sciences Sequencing Core and
728 Genome Technology Access Center at the McDonnell Genome Institute for sequencing
729 services. The authors also thank Dr. Leyao Wang and members of the Kau lab for
730 critical reading of the manuscript.

731 **References**

- 732 1. Arrieta M-C, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch S, et al.
733 Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci*
734 *Transl Med.* 2015;7:307–152.
- 735 2. Hufnagl K, Pali-Schöll I, Roth-Walter F, Jensen-Jarolim E. Dysbiosis of the gut and
736 lung microbiome has a role in asthma. *Semin Immunopathol.* 2020;42:75–93.
- 737 3. Fujimura KE, Sitarik AR, Havstad S, Lin DL, Levan S, Fadrosch D, et al. Neonatal gut
738 microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat*
739 *Med.* 2016;22:1187–91.
- 740 4. Levan SR, Stamnes KA, Lin DL, Panzer AR, Fukui E, McCauley KK, et al. Elevated
741 faecal 12,13-diHOME concentration in neonates at high risk for asthma is produced by
742 gut bacteria and impedes immune tolerance. *Nat Microbiol.* 2019;4:1851–61.
- 743 5. Roudot C, Frei R, Ferstl R, Loeliger S, Westermann P, Rhyner C, et al. High levels of
744 Butyrate and Propionate in early life are associated with protection against atopy.
745 *Allergy.* 2018; June:1–11.
- 746 6. Cait A, Hughes MR, Antignano F, Cait J, Dimitriu PA, Maas KR, et al. Microbiome-
747 driven allergic lung inflammation is ameliorated by short-chain fatty acids. *Mucosal*
748 *Immunol.* 2018;11:785–95.
- 749 7. Trompette A, Gollwitzer ES, Yadava K, Sichelstiel AK, Sprenger N, Ngom-Bru C, et
750 al. Gut microbiota metabolism of dietary fiber influences allergic airway disease and
751 hematopoiesis. *Nat Med.* 2014;20:159–66.
- 752 8. Zaiss MM, Rapin A, Lebon L, Dubey LK, Mosconi I, Sarter K, et al. The Intestinal
753 Microbiota Contributes to the Ability of Helminths to Modulate Allergic Inflammation.

- 754 Immunity. 2015;43:998–1010.
- 755 9. Li X, Stokholm J, Brejnrod A, Vestergaard GA, Russel J, Trivedi U, et al. The infant
756 gut resistome associates with *E. coli*, environmental exposures, gut microbiome
757 maturity, and asthma-associated bacterial composition. *Cell Host Microbe*. 2021;;1–13.
- 758 10. Ramirez J, Guarner F, Bustos Fernandez L, Maruy A, Sdepanian VL, Cohen H.
759 Antibiotics as Major Disruptors of Gut Microbiota. *Front Cell Infect Microbiol*. 2020;10
760 November:1–10.
- 761 11. Kozyrskyj AL, Ernst P, Becker AB. Increased risk of childhood asthma from
762 antibiotic use in early life. *Chest*. 2007;131:1753–9.
- 763 12. McKeever TM, Lewis SA, Smith C, Collins J, Heatlie H, Frischer M, et al. Early
764 exposure to infections and antibiotics and the incidence of allergic disease: A birth
765 cohort study with the West Midlands General Practice Research Database. *J Allergy*
766 *Clin Immunol*. 2002;109:43–50.
- 767 13. Hoskin-Parr L, Teyhan A, Blocker A, Henderson AJW. Antibiotic exposure in the first
768 two years of life and development of asthma and other allergic diseases by 7.5 yr: A
769 dose-dependent relationship. *Pediatr Allergy Immunol*. 2013;24:762–71.
- 770 14. Russell SL, Gold MJ, Willing BP, Thorson L, McNagny KM, Finlay BB. Perinatal
771 antibiotic treatment affects murine microbiota, immune responses and allergic asthma.
772 *Gut Microbes*. 2013;4:158–64.
- 773 15. Yang X, Feng H, Zhan X, Zhang C, Cui R, Zhong L, et al. Early-life vancomycin
774 treatment promotes airway inflammation and impairs microbiome homeostasis. *Aging*
775 (Albany NY). 2019;11:2071–81.
- 776 16. Borbet TC, Pawline MB, Zhang X, Wipperman MF, Reuter S, Maher T, et al.

- 777 Influence of the early-life gut microbiota on the immune responses to an inhaled
778 allergen. *Mucosal Immunol.* 2022;15:1000–11.
- 779 17. Wenzel SE. Asthma phenotypes: The evolution from clinical to molecular
780 approaches. *Nat Med.* 2012;18:716–25.
- 781 18. Michalovich D, Rodriguez-Perez N, Smolinska S, Pirozynski M, Mayhew D, Uddin S,
782 et al. Obesity and disease severity magnify disturbed microbiome-immune interactions
783 in asthma patients. *Nat Commun.* 2019;10.
- 784 19. Wang Q, Li F, Liang B, Liang Y, Chen S, Mo X, et al. A metagenome-wide
785 association study of gut microbiota in asthma in UK adults. *BMC Microbiol.* 2018;18:1–
786 7.
- 787 20. Chiu CY, Chan YL, Tsai MH, Wang CJ, Chiang MH, Chiu CC. Gut microbial
788 dysbiosis is associated with allergen-specific IgE responses in young children with
789 airway allergies. *World Allergy Organ J.* 2019;12:100021.
- 790 21. Liu A, Ma T, Xu N, Jin H, Zhao F, Kwok L-Y, et al. Adjunctive Probiotics Alleviates
791 Asthmatic Symptoms via Modulating the Gut Microbiome and Serum Metabolome.
792 *Microbiol Spectr.* 2021;9:1–17.
- 793 22. Jaeger N, McDonough RT, Rosen AL, Hernandez-Leyva A, Wilson NG, Lint MA, et
794 al. Airway Microbiota-Host Interactions Regulate Secretory Leukocyte Protease Inhibitor
795 Levels and Influence Allergic Airway Inflammation. *Cell Rep.* 2020;33.
- 796 23. Wilson NG, Hernandez-Leyva A, Rosen AL, Jaeger N, McDonough RT, Santiago-
797 Borges J, et al. The asthma gut microbiota influences lung inflammation in gnotobiotic
798 mice. *bioRxiv.* 2022;:2022.08.09.502549.
- 799 24. Kau AL, Planer JD, Liu J, Rao SS, Yatsunencko T, Trehan I, et al. Functional

800 characterization of IgA-targeted bacterial taxa from undernourished Malawian children
801 that produce diet-dependent enteropathy. *Sci Transl Med*. 2015;7:1–15.

802 25. Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, Kishony RK.
803 Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One*.
804 2015;10:1–15.

805 26. Rodriguez-R LM, Konstantinidis KT. Nonpareil: a redundancy-based approach to
806 assess the level of coverage in metagenomic datasets. *Bioinformatics*. 2014;30:629–35.

807 27. Rodriguez-R LM, Gunturu S, Tiedje JM, Cole JR, Konstantinidis KT. Nonpareil 3:
808 Fast Estimation of Metagenomic Coverage and Sequence Diversity. *mSystems*. 2018;3.

809 28. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, et al.
810 Integrating taxonomic, functional, and strain-level profiling of diverse microbial
811 communities with biobakery 3. *Elife*. 2021;10:1–42.

812 29. Kaminski J, Gibson MK, Franzosa EA, Segata N, Dantas G, Huttenhower C. High-
813 Specificity Targeted Functional Profiling in Microbial Communities with ShortBRED.
814 *Nature*. 2012;486:207–14.

815 30. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al.
816 CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic
817 resistance database. *Nucleic Acids Res*. 2019;48 October 2019:517–25.

818 31. L C, J Y, J Y, Z Y, L S, Y S, et al. VFDB: a reference database for bacterial
819 virulence factors. *Nucleic Acids Res*. 2005;33 Database issue.

820 32. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, et al. Multivariable
821 association discovery in population-scale meta-omics studies. *PLoS Comput Biol*.
822 2021;17.

- 823 33. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al.
824 SPAdes: A new genome assembly algorithm and its applications to single-cell
825 sequencing. *J Comput Biol.* 2012;19:455–77.
- 826 34. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUASt: Quality assessment tool for
827 genome assemblies. *Bioinformatics.* 2013;29:1072–5.
- 828 35. Mikheenko A, Valin G, Prjibelski A, Saveliev V, Gurevich A. Icarus: visualizer for de
829 novo assembly evaluation. *Bioinformatics.* 2016;32:3321–3.
- 830 36. Brayner B, Kaur G, Keske MA, Perez-Cornago A, Piernas C, Livingstone KM.
831 Dietary Patterns Characterized by Fat Type in Association with Obesity and Type 2
832 Diabetes: A Longitudinal Study of UK Biobank Participants. *J Nutr.* 2021;151:3570–8.
- 833 37. Wendell SG, Baffi C, Holguin F. Fatty acids, inflammation, and asthma. *J Allergy*
834 *Clin Immunol.* 2014;133:1255–64.
- 835 38. Mizuta K, Matoba A, Shibata S, Masaki E, Emala CW. Obesity-induced asthma:
836 Role of free fatty acid receptors. *Jpn Dent Sci Rev.* 2019;55:103–7.
- 837 39. Scott HA, Gibson PG, Garg ML, Wood LG. Airway inflammation is augmented by
838 obesity and fatty acids in asthma. *Eur Respir J.* 2011;38:594–602.
- 839 40. Schatz M, Sorkness CA, Li JT, Marcus P, Murray JJ, Nathan RA, et al. Asthma
840 Control Test: Reliability, validity, and responsiveness in patients not previously followed
841 by asthma specialists. *J Allergy Clin Immunol.* 2006;117:549–56.
- 842 41. Snyder BM, Patterson MF, Gebretsadik T, Cacho F, Ding T, Turi KN, et al.
843 Association between asthma status and prenatal antibiotic prescription fills among
844 women in a Medicaid population. *J Asthma.* 2021;59:2100–7.
- 845 42. Mardia K, Kent J, Bibby J. Mardia, Kent, Bibby - 1979 - Multivariate Analysis.pdf.

- 846 1979;:521.
- 847 43. Seemann T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*.
848 2014;30:2068–9.
- 849 44. Nagel G, Linseisen J. Dietary intake of fatty acids, antioxidants and selected food
850 groups and asthma in adults'. *Eur J Clin Nutr*. 2005;59:8–15.
- 851 45. Mizuta K, Zhang Y, Mizuta F, Hoshijima H, Shiga T, Masaki E, et al. Novel
852 identification of the free fatty acid receptor FFAR1 that promotes contraction in airway
853 smooth muscle. *Am J Physiol - Lung Cell Mol Physiol*. 2015;309:L970–82.
- 854 46. Felton E, Burrell A, Chaney H, Sami I, Koumbourlis AC, Freishtat RJ, et al.
855 Inflammation in children with cystic fibrosis: contribution of bacterial production of long-
856 chain fatty acids. *Pediatr Res*. 2021;90:99–108.
- 857 47. Niot I, Poirier H, Tran TTT, Besnard P. Intestinal absorption of long-chain fatty acids:
858 Evidence and uncertainties. *Prog Lipid Res*. 2009;48:101–15.
- 859 48. Vital M, Howe AC, Tiedje JM. Revealing the bacterial butyrate synthesis pathways
860 by analyzing (meta)genomic data. *MBio*. 2014;5:1–11.
- 861 49. Yassour M, Vatanen T, Siljander H, Hämäläinen AM, Härkönen T, Ryhänen SJ, et
862 al. Natural history of the infant gut microbiome and impact of antibiotic treatment on
863 bacterial strain diversity and stability. *Sci Transl Med*. 2016;8.
- 864 50. Schwartz DJ, Langdon AE, Dantas G. Understanding the impact of antibiotic
865 perturbation on the human microbiome. *Genome Med*. 2020;12:1–12.
- 866 51. Palleja A, Mikkelsen KH, Forslund SK, Kashani A, Allin KH, Nielsen T, et al.
867 Recovery of gut microbiota of healthy adults following antibiotic exposure. *Nat Microbiol*.
868 2018;3:1255–65.

- 869 52. Juhn YJ. Risks for infection in patients with asthma (or other atopic conditions): Is
870 asthma more than a chronic airway disease? *J Allergy Clin Immunol.* 2014;134:247-
871 257.e3.
- 872 53. Malhotra-Kumar S, Lammens C, Coenen S, Van Herck K, Goossens H. Effect of
873 azithromycin and clarithromycin therapy on pharyngeal carriage of macrolide-resistant
874 streptococci in healthy volunteers: a randomised, double-blind, placebo-controlled
875 study. *Lancet.* 2007;369:482–90.
- 876 54. Doan T, Hinterwirth A, Worden L, Arzika AM, Maliki R, Abdou A, et al. Gut
877 microbiome alteration in MORDOR I: a community-randomized trial of mass
878 azithromycin distribution. *Nat Med.* 2019;25:1370–6.
- 879 55. Doan T, Arzika AM, Hinterwirth A, Maliki R, Zhong L, Cummings S, et al. Macrolide
880 Resistance in MORDOR I — A Cluster-Randomized Trial in Niger. *N Engl J Med.*
881 2019;380:2271–3.
- 882 56. Gibson PG, Yang IA, Upham JW, Reynolds PN, Hodge S, James AL, et al. Effect of
883 azithromycin on asthma exacerbations and quality of life in adults with persistent
884 uncontrolled asthma (AMAZES): a randomised, double-blind, placebo-controlled trial.
885 *Lancet.* 2017;390:659–68.
- 886 57. Brusselle GG, VanderStichele C, Jordens P, Deman R, Slabbynck H, Ringoet V, et
887 al. Azithromycin for prevention of exacerbations in severe asthma (AZISAST): a
888 multicentre randomised double-blind placebo-controlled trial. *Thorax.* 2013;68:322–9.
- 889 58. Escudeiro P, Pothier J, Dionisio F, Nogueira T. Antibiotic Resistance Gene Diversity
890 and Virulence Gene. *mSphere.* 2019;4:1–13.
- 891 59. Pan Y, Zeng J, Li L, Yang J, Tang Z, Xiong W, et al. Coexistence of Antibiotic

- 892 Resistance Genes and Virulence Factors Deciphered by Large-Scale Complete
893 Genome Analysis. *mSystems*. 2020;5.
- 894 60. Strati F, Pujolassos M, Burrello C, Giuffrè MR, Lattanzi G, Caprioli F, et al.
895 Antibiotic-associated dysbiosis affects the ability of the gut microbiota to control
896 intestinal inflammation upon fecal microbiota transplantation in experimental colitis
897 models. *Microbiome*. 2021;9:1–15.
- 898 61. Casterline BW, Hecht AL, Choi VM, Bubeck Wardenburg J. The *Bacteroides fragilis*
899 pathogenicity island links virulence and strain competition. *Gut Microbes*. 2017;0976:1–
900 10.
- 901 62. Walter J, Armet AM, Finlay BB, Shanahan F. Establishing or Exaggerating Causality
902 for the Gut Microbiome: Lessons from Human Microbiota-Associated Rodents. *Cell*.
903 2020;180:221–32.
- 904 63. Barcik W, Pugin B, Westermann P, Perez NR, Ferstl R, Wawrzyniak M, et al.
905 Histamine-secreting microbes are increased in the gut of adult asthma patients. *J*
906 *Allergy Clin Immunol*. 2016;138:1491-1494.e7.
- 907 64. Barcik W, Pugin B, Brescó MS, Westermann P, Rinaldi A, Groeger D, et al.
908 Bacterial secretion of histamine within the gut influences immune responses within the
909 lung. *Allergy Eur J Allergy Clin Immunol*. 2019;74:899–909.
- 910 65. Licari A, Fuchs D, Marseglia G, Ciprandi G. Tryptophan metabolic pathway and
911 neopterin in asthmatic children in clinical practice. *Ital J Pediatr*. 2019;45:1–4.
- 912 66. Van der Leek AP, Yanishevsky Y, Kozyrskyj AL. The kynurenine pathway as a novel
913 link between allergy and the gut microbiome. *Front Immunol*. 2017;8 NOV:1374.
- 914 67. Asthma-related emergency department visits 2010–2018 | CDC.

915 https://www.cdc.gov/asthma/asthma_stats/asthma-ed-visits_2010-2018.html. Accessed
916 28 Dec 2022.

917 68. Xie Y. knitr: A Comprehensive Tool for Reproducible Research in {R}. In: Stodden
918 V, Leisch F, Peng RD, editors. Implementing Reproducible Computational Research.
919 Chapman and Hall/CRC; 2014.

920 69. Mirzayi C, Renson A, Furlanello C, Sansone SA, Zohra F, Elsafoury S, et al.
921 Reporting guidelines for human microbiome research: the STORMS checklist. Nat Med.
922 2021;27:1885–92.

923

924 Supplemental Figure and Data Table Captions

925

926 Figures:

927 **Figure S1: MARS whole metagenomic shotgun sequencing captures essential**
928 **functions and taxonomic shifts of the asthma gut microbiota. A)** Summary of select
929 sequencing statistics from NovaSeq shotgun metagenomic sequencing and subsequent
930 filtering steps. **B)** Boxplot of redundancy-based estimated metagenome coverage (%)
931 as calculated by running the forward reads through the Nonpareil tool. Split into asthma
932 and age group and Two-way Type II ANOVA results shown. **C)** Bar plot of MetaCyc
933 pathway copies per million (CPM) in all MARS samples annotated by HUMAnN
934 pipeline, with horizontal length representing mean and bars the standard error. For all
935 panels: N= 20 healthy children, 39 healthy adults, 19 asthmatic children, 17 asthmatic
936 adults. **D)** Relative abundance stacked barplots of top abundant bacterial genera split
937 by age group and asthma cohort. **E)** Simpson alpha diversity boxplots split by asthma
938 and age group cohorts (2-Way Type II ANOVA). **F)** NMDS of Bray-Curtis Dissimilarity of
939 species-level relative abundance grouped by age and asthma. **G)** Sequential
940 PERMANOVA to test effect of demographics on beta diversity. Terms were input into
941 the test as ordered from top to bottom of barplot. Dotted vertical line represents a p
942 value of 0.05. Color scale is mapped to the R² value. **H)** Arcsine transformed relative
943 abundance boxplots of differentially abundant species as determined by Maaslin2 with
944 age group and race modeled as random effects. For all panels: N= 20 healthy children,
945 39 healthy adults, 19 asthmatic children, 17 asthmatic adults.

946

947 **Figure S2: KEGG orthologs, KEGG pathways, and differentially abundant**
948 **MetaCyc fatty acid pathways. A)** Relative abundance of KEGG orthologs previously
949 implicated in asthma. Copies per million (CPM) are counts normalized by gene size and
950 read depth, then total-sum-scaled to one million. **B-C)** Stacked bar plots of differentially
951 abundant pathways mapped to respective taxa including “Community” bin which
952 accounts for the remaining reads that mapped to the pathway but not to any single
953 species by MetaPhlan3.0/HUMAnN3.0, averaged within asthma or healthy cohorts. **B)**
954 L-lysine biosynthesis III pathway. Only top 13 taxa shown in addition to Community
955 category. **C)** Seven fatty acid metabolism pathways differentially abundant in the
956 asthma cohort. Only top 9 taxa shown in addition to Community category. Stars
957 represent a q value < 0.05 of Wilcoxon tests between the Community pathway richness
958 and *B. vulgatus*-encoded pathway richness. For all panels: N= 59 healthy, 36 asthmatic
959 individuals.

960

961 **Figure S3: Pathway collage for differentially abundant MetaCyc pathways.** “PWY-
962 6519: 8-amino-7-oxononanoate biosynthesis I” is completely overlapping with “BIOTIN-
963 BIOSYNTHESIS-PWY: biotin biosynthesis I” and its steps are highlighted in blue text.
964 Pathway collage made on MetaCyc browser tool.

965

966 **Figure S4: Gut virulence factor ecology shifts with age group but not asthma**
967 **cohort. A)** Total-sum scaled RPKM Bray-Curtis Dissimilarity Non-metric
968 Multidimensional Scaling (NMDS) plot labeled by asthma and age cohorts. Showing two
969 axes out of 5 with stress value=0.09. **B)** Effect of demographic categories on virulence

970 factor profile in A (by sequential PERMANOVA, input terms ordered from top to bottom
971 of barplot). **C**) Stacked violin plots of virulence factor alpha diversity grouped by either
972 healthy and asthma cohort (blue green colors in background) or age (brown colors in
973 foreground). Two-Way ANOVA results shown in table above plot. **D**) Procrustes plot
974 and PROTEST analysis between virulence factor profile Bray-Curtis dissimilarity
975 distances and Metaphlan species relative abundance Bray-Curtis dissimilarity
976 distances. Arrows connect the two data points belonging to identical samples. For all
977 panels: N= 20 healthy children, 39 healthy adults, 19 asthmatic children, 17 asthmatic
978 adults.

979
980 **Figure S5: Asthma-associated ARG richness and ARG-VF co-occurrence**
981 **relationships are observed within *K. pneumoniae* and *E. coli*.** **A**) Richness bar plots
982 between antibiotic resistance genes (ARGs) and virulence factors (VFs) grouped by
983 asthma status. **B**) Heatmap of the co-occurrence of each VF/ARG pair colored by the
984 direction in which (positively or negatively co-occurring) and the cohort for which
985 (asthma vs. healthy) the pair had a p-value less than 0.05 via R cooccur function. Blank
986 squares were pairs filtered out due to a lack of observed co-occurrence. SS: secretion
987 systems. (N=36 Healthy, 59 Asthmatic)

988
989 **Tables:**

990 Table S1: Fecal shotgun metagenomics filtering and assembly summary statistics
991 Table S2: Maaslin 2 Analysis of Metaphlan Community Composition
992 Table S3: Comparison of MetaCyc Pathway Differential Abundance Analyses
993 Table S4: Negative Binomial tests of antibiotic resistance gene abundance (RPKM)
994 between healthy and asthma cohorts
995 Table S5: Negative Binomial tests of virulence factor abundance (RPKM) between
996 healthy and asthma cohorts

997
998 **Additional Files**

999 **Additional File 1:** Statistical Analyses for 'The gut metagenome harbors metabolic and
1000 antibiotic resistance signatures of moderate-to-severe asthma' knitr document