1   **Early-life immune expression profiles predict later life health and**

2   **fitness in a wild rodent**

3

4   Klara M Wanelik[1,2*], Mike Begon[1], Janette E Bradley[3], Ida M Friberg[4], Christopher H

5   Taylor[3], Joseph A Jackson[4], Steve Paterson[1]

6

7   [1] Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool,

8   Liverpool, UK

9   [2] Department of Zoology, University of Oxford, Oxford, UK

10   [3] School of Life Sciences, University of Nottingham, Nottingham, UK

11   [4] School of Environment and Life Sciences, University of Salford, Salford, UK

12

13   * Corresponding author

14   **Email:** klara.wanelik@zoo.ox.ac.uk

15

16

17

18

19

20

21

22

23

24

25

1

Abstract

Individuals differ in the nature of the immune responses they produce, affecting disease susceptibility and ultimately health and fitness. These differences have been hypothesised to have an origin in events experienced early in life that then affect trajectories of immune development and responsiveness. Here we investigate early life influences on immune expression profiles using a natural population of field voles, *Microtus agrestis*, in which we are able to monitor variation between and within individuals though time by repeat (longitudinal) sampling of individually marked animals. We analysed the co-expression of 20 immune genes in early life to create a correlational network consisting of three main clusters, one of which (containing *Gata3*, *Il10* and *Il17*) was associated with later life reproductive success and susceptibility to chronic bacterial (*Bartonella*) infection. More detailed analyses supported associations between early life expression of *Il17* and reproductive success later in life, and of increased *Il10* expression early in life and later infection with *Bartonella*. We also found significant association between an *Il17* genotype and the early life expression of *Il10*. Our results demonstrate that immune expression profiles can be manifested during early life with effects that persist through adulthood and that shape the variability among individuals in susceptibility to infection and fitness widely seen in natural populations.

2

51 **Introduction**

52

53 Differences at birth, or experienced during neonatal or juvenile development, may

54 impact an individual's ability as an adult to respond to environmental allergens or

55 pathogens, or to maintain health and reproduction (measures of which we define

56 below). In medical sciences, the developmental origin of health and disease (DOMD)

57 theory – which proposes that adult diseases can be traced back to childhood – has

58 been highly influential since its inception and has provided insight into the aetiology

59 of chronic inflammatory and allergic disorders in humans, such as atopy, asthma and

60 autoimmune diseases[1,2]. Similarly, the development of the immune system during

61 early life is an important influence on subsequent health and fitness[3]. The immune

62 system is shaped by environmental exposure to allergens and pathogens, which, in

63 turn, affect the immune responses made following subsequent exposure. Coupled

64 with genetic polymorphisms affecting immune function[4], early life effects associated

65 with the immune system can give rise to substantial variation among individuals in

66 their immune responses. Such environmental effects can act very early in life, even

67 at birth. For example, children born by caesarean section are exposed to different

68 microbes to vaginal delivery. This has been shown to modify their gut microbiota and

69 the production of pro-inflammatory factors, and ultimately to increase their risk of

70 opportunistic infections[5] and chronic inflammatory disorders in later life e.g. asthma,

71 type 1 diabetes and celiac disease[6,7].

72

73 While effective immune responses are essential to health by protecting individuals

74 against pathogens, immunopathology (such as tissue damage from inflammatory or

75 over-reactive responses) contributes to significant ill-health in chronic inflammatory

3

76 and allergic disorders such as atopy, asthma, autoimmune diseases or indeed

77 infectious diseases such as COVID-19. An important concept is the hygiene

78 hypothesis, which proposes that microbial exposure early in life helps the

79 development of a well-adjusted immune system, the lack of which during childhood

80 in modern societies contributes to the development of immunopathology and chronic

81 immune disorders. Many studies have shown that traditional farming methods where

82 children are widely exposed to microbes from an early age, develop fewer chronic

83 immune disorders than children brought up in modern farms where there are lower

84 levels of microbial exposure. This has been referred to as the 'farm effect'[8,9].

85 Variation in traits associated with immunopathology can also be found in natural

86 populations. In a wild population of moose, malnutrition early in life is linked to both a

87 greater risk of osteoarthritis later in life (a disease associated with immune

88 dysregulation) and reduced life expectancy[10,11]. In a feral population of Soay sheep,

89 heritable variation in levels of anti-nuclear antibodies – which are self-reactive and if

90 maintained at high titres with other markers are indicative of immune dysregulation in

91 humans and dogs – are associated with variations in reproduction and over-winter

92 survivorship[12]. Indeed, natural populations can provide a useful model to understand

93 early life effects on the immune system, because they consist of genetically diverse

94 individuals living in an environment that naturally exposes their immune systems to a

95 wide array of pathogens and allergens. The interaction of different immune

96 components with these environmental stimuli may lead to differences among

97 individuals in their capacity for either effective responses to infection or

98 immunopathologies.

99

100   Previous studies of early-life effects on immunity in natural populations have

101   considered few immune parameters. For example, a number of avian studies have

102   shown associations between broad-brush measures of nestling immune function,

103   such as responsiveness to phytohemagglutinin (PHA) injection, and recruitment[13–16].

104   However, immune responses involve a complex network of interacting signalling and

105   effector molecules, controlling pathways specialised at combating different types of

106   pathogens, or mitigating damage caused by the pathogen or by immunopathology.

107   Activity in one pathway of the immune system may have either positive or negative

108   effects on the activity of another immune pathway. Similarly, the activity of immune

109   pathways will both be influenced by environmental cues, such as infection and

110   nutrition, and, in turn, influence disease resistance, health and fitness. A more

111   nuanced and sophisticated approach to study immunity in natural systems is

112   therefore needed, which considers multi-dimensional immune phenotypes; providing

113   a way to reduce the complexity of these phenotypes, while still acknowledging the

114   links between immune pathways. The availability of genome and transcriptome

115   sequences provides new possibilities to measure immune phenotypes in natural

116   populations, particularly through reverse transcription quantitative PCR (Q-PCR) to

117   measure transcriptional expression of key immune regulators such as cytokines that

118   are indicative of activity within different immune pathways[17].

119

120   Here, we investigate early life influences on immune expression profiles in a natural

121   population of field voles, *Microtus agrestis*. In this study system we are able to

122   sample individuals longitudinally through their lives to measure a range of

123   immunological parameters using Q-PCR and to relate these to subsequent variations

124   in measures of health and fitness, including responses to infection. We do so having

125   initially sought structure in the patterns of expression in our immune network,

126   allowing us to characterise immune expression profiles in terms of appropriate

127   clusters of genes, rather than single markers. Our central aim is to understand the

128   extent to which, (1) the reproductive success of mature individuals (a measure of

129   fitness), and (2) the response to infection of mature individuals (a measure of health)

130   is foreshadowed by immune expression in the early stages of their lives. We note at

131   the outset that the answers to these three questions are likely to vary between

132   different parts of the immune network, and hence a further aim of this study is to

133   begin to elaborate the range of possible patterns, and not to establish a general rule.

134

135   **Methods**

136

137   *M. agrestis* were live-trapped from natural populations in Kielder Forest,

138   Northumberland, UK, from 2015-2017 across seven different sites, each a forest

139   clear-cut. At each site, 150-197 Ugglan small mammal traps (Grahnab, Sweden)

140   were laid out in a grid spaced approximately 3-5 m apart. Our study was divided into

141   longitudinal (individuals sampled live, multiple times) and cross-sectional (individuals

142   sampled following culling) components.

143

144   **Ethics statement**

145   All animal procedures were performed with approval from the University of Liverpool

146   Animal Welfare Committee and under a UK Home Office licence (PPL 70/8210 to

147   S.P.).

148

149   For the longitudinal study, on first capture, voles were injected with a Passive

150   Integrated Transponder (PIT) tag (AVID, UK) for unique identification. On first and

151   subsequent captures, a blood sample was taken whenever possible from the tail tip

152   for immune gene expression assays and parasite detection, and on first capture a

153   tail sample was taken for genotyping for pedigree reconstruction (see below). Other

154   basic information was recorded including the snout-vent length, weight and pelage

155   (for aging). Individuals were sampled in this way between 1-13 times. Some of these

156   tagged voles ($n = 510$) were culled later in life for more detailed measurements to be

157   taken (e.g., lens weight for aging purposes; see below), as part of our cross-

158   sectional study.

159

160   Of the voles in the longitudinal study, 624 were first caught prior to breeding (at 6

161   weeks or less). These included juveniles and subadults, aged on weight and pelage.

162   Based on previous work on the same study population, voles weighing 11g or less

163   were assumed to be juveniles aged 2.5 weeks old (mean of 2-3 weeks) and voles

164   weighing more than 11g but no more than 15g were assumed to be juveniles aged

165   4.5 weeks (mean of 3-6 weeks). Those voles weighing more than 15g but aged as

166   juvenile on pelage, and those voles aged as subadult on pelage were assumed to be

167   6 weeks old. Finally, those voles aged as juvenile on pelage but with no weight

168   measurement were assumed to be 4 weeks old (mean of 2-6 weeks)[18,19]. Of these

169   624 voles known to be first caught prior to breeding, 195 had a blood sample taken

170   at first capture.

171

172   Some individuals may have first been caught as subadults but not recorded as such

173   due to their adult-like pelage. In order to identify these individuals (and to increase

7

174   our sample size) we regressed age against eye lens weight for those voles definitely

175   first sampled prior to breeding (for which we had a birth date) that were culled in later

176   life as part of our cross-sectional study ($n = 99$). The relationship between age and

177   eye lens weight has been previously described[20]. A quassipoisson GLM with

178   quadratic and cubic terms for eye lens weight, and log link, provided a good model fit

179   for our data ($r^2 = 0.56$; Supplementary Fig. S1). We used this regression to predict

180   the ages (and hence birth dates) for all other culled voles. Of these, 59 voles were

181   inferred to have first been caught prior to breeding, since their birth date was no

182   more than 6 weeks before their date of first capture, of which 28 had a blood sample

183   taken at first capture. Therefore, the final number of voles first sampled prior to

184   breeding (known and inferred) was 223.

185

186   **Immune gene expression assays**

187   We used SYBR green based Q-PCR to measure the expression levels of a panel of

188   20 immune-associated genes (Table 1) in blood from our longitudinal animals. The

189   choice of our panel of genes was informed by (1) known immune-associated

190   functions in mice, combined with (2) significant sensitivity of gene expression to

191   environmental or intrinsic host variables in our previous studies[21,22] or in a recent

192   differential expression analysis of RNASeq data (not reported here).

193

194

195

196

197

198   **Table 1** Panel of 20 immune-associated genes for which expression levels in blood

199   samples were measured using Q-PCR.

Gene	Protein
Cd8a	T-cell surface glycoprotein CD8 alpha chain
Cd86	T-lymphocyte activation antigen CD86
Cd99l2	CD99 antigen-like protein 2
Foxp3	Forkhead box protein P3
Gata3	GATA binding protein 3
Ifng	Interferon gamma
Irf2	Interferon regulatory factor 2
Irf9	Interferon regulatory factor 9
Tollip	Toll-interacting protein
Retnlg	Resistin-like gamma
Tgfb1	Transforming growth factor beta 1
Il4	Interleukin-4
Il10	Interleukin-10
Il17	Interleukin-17
Il1b	Interleukin-1 beta
Il1rap	Interleukin-1 receptor accessory protein
Il1rn	Interleukin-1 receptor antagonist protein
Apobr	Apolipoprotein B receptor
Orai1	Calcium release-activated calcium channel protein 1
ccn2 (ctgf)	Cellular communication network factor 2

200

201   All primer sets were designed *de novo* in-house and validated (to confirm specific

202   amplification and 100±10% PCR efficiency under assay conditions). *Ywhaz* and *Actb*

203   were employed as endogenous control genes. We extracted RNA from blood

204   conserved in RNAlater using the Mouse RiboPure Blood RNA Isolation Kit

205   (ThermoFisher), according to manufacturer's instructions. RNA extracts were DNAse

206   treated and converted to cDNA using the High-Capacity RNA-to-cDNA™ Kit

207   (ThermoFisher), according to manufacturer's instructions, including reverse

208   transcription negative (RT-) controls for a subsample. SYBR green-based assays

209   were pipetted onto 384 well plates by a robot (Pipetmax, Gilson) using a custom

210   programme and run on a QuantStudio 6-flex Real-Time PCR System

211   (ThermoFisher) at the machine manufacturers default real-time PCR cycling

212   conditions. Reaction size was 10 µl, incorporating 1 µl of template and

213   PrecisionFAST qPCR Master Mix with low ROX and SYBR green (PrimerDesign)

214   and primers at the machine manufacturer's recommended concentrations. We used

215   three standard plate layouts for assaying, each of which contained a fixed set of

216   target gene expression assays and the two endogenous control gene assays (the

217   same sets of animals being assayed on matched triplets of the standard plate

218   layouts). Unknown samples were assayed in duplicate wells and calibrator samples

219   in triplicate wells and no template controls for each gene were included on each

220   plate. Template cDNA (see above) was diluted 1/20 prior to assay. A main calibrator

221   sample (identical on each plate) was created by pooling cDNA from blood samples

222   taken from many different voles from the study site. As *Tollip*, *Il1rap* and *Irf2* were

223   relatively poorly represented in this main calibrator sample, a synthesised 478 bp

224   gene fragment containing the amplification target for each of these genes was used

225   as an additional calibrator sample (at 10×10^5 copies µl^{-1}) in these cases. Samples

226   from different field sampling groups were dispersed across plate triplets, avoiding

227   confounding of plate with the sampling structure. Gene relative expression values

228   used in analyses are RQ values calculated by the QuantStudio 6-flex machine

229   software according to the ΔΔCt method, indexed to the appropriate calibrator

230   samples. Melting curves and amplification plots were individually inspected for each

231   well replicate to confirm specific amplification.

232

233   Parasite detection

234   We quantified infections by microparasites (*Babesia microti* and *Bartonella* spp.) in

235   blood using a strategy analogous to the host gene expression assays above,

236   targeting pathogen ribosomal RNA gene expression and normalizing to the host

237   endogenous control genes. We included two extra sets of primers in the blood Q-

238   PCR assays described above (in one of the standard plate layouts for each plate

239   triplet above). For *B. microti* we used the forward primer

240   CTACGTCCCTGCCCTTTGTA and reverse primer CCACGTTTCTTGGTCCGAAT

241   targeting the 18S ribosomal RNA gene and for *Bartonella* spp. we used the forward

242   primer GATGAATGTTAGCCGTCGGG and reverse primer

243   TCCCCAGGCGGAATGTTTAA targeting the 16S ribosomal RNA gene. As a

244   calibrator sample we employed the main calibrator sample used for host gene

245   expression in blood (above) in addition to a pool of DNA extracted from 154 blood

246   samples from different *M. agrestis* at our study sites in 2015 and 2016; these DNA

247   extractions were carried out using the QIAamp UCP DNA Micro Kit (Qiagen)

248   following manufacturer's instructions.

249   Relative expression values are presumed to relate to the expression of pathogen

250   ribosomal RNA genes and in turn to the intensity of infection. Pilot testing indicated

251   that the detected relative quantities of parasite RNA correlated closely to visual

252   counts of parasitized cells on blood smears examined under a microscope.

253   Moreover, we validated our diagnostic results by comparing our PCR RQ values to

254   independent data for a subset of cross-sectional voles with mapped genus-level

255   pathogen reads from RNASeq analysis of blood samples ($n = 44$), finding the two

256   data sets strongly corroborated each other (*Bartonella* spp., $r^2 = 0.81$, $p < 0.001$; *B.*

257   *microti*, $r^2 = 0.81$, $p < 0.001$).

258

259   **Genotyping & pedigree reconstruction**

260   We genotyped voles from their tail samples for 346 single nucleotide polymorphisms

261   (SNPs) in 127 genes. See Wanelik *et al.*[23] for details of the approach used to select

11

262   these SNPs. Briefly, we combined information from an external database of human

263   immune genes (Immunome database; http://structure.bmc.lu.se/idbase/Immunome/

264   index.php) with our own existing knowledge of the study system. DNA was extracted

265   from a tail sample taken from the animal using DNeasy Blood and Tissue Kit

266   (Qiagen). Genotyping was then performed by LGC Biosearch Technologies

267   (Hoddesdon, UK; http://www.biosearchtech.com) using the KASP SNP genotyping

268   system. This included negative controls (water) and duplicate samples for validation

269   purposes.

270

271   We used a subset of our SNP dataset (n = 114 SNPs) to reconstruct a pedigree

272   using the R package *Sequoia*[24]. All analyses were performed in R statistical software

273   version 3.5.2[25]. Full details can be found in Wanelik *et al.*[26]. Briefly, we inputted life

274   history information into *Sequoia* where possible (99% of samples were assigned a

275   sex; 54% were assigned a birth month), and generated site-specific pedigrees

276   (assuming no dispersal between sites, each several kilometres apart). We inspected

277   log10 likelihood ratios (LLRs) for parent pairs as recommended in the user manual

278   for *Sequoia*. Almost all LLRs were positive (97% of LLRs) indicating confidence in

279   our assignments. For each individual present in a pedigree (n = 652; see below), the

280   number of offspring was counted to provide a measure of their reproductive success.

281   Half of individuals present in our pedigrees (n = 325) were found to have no

282   offspring. We expect the majority of these to be true zeros (representing actual

283   reproductive failure) as we sampled a large proportion of the total population within

284   clear-cuts. We minimised the chance of false zeros by excluding from the pedigree

285   those individuals (e.g., at the periphery of a study grid) for which we recorded no

286   relatives (including offspring) likely because we had not sampled in the right place.

287   **Statistical methods**

288

289   **Correlation networks**

290   We constructed an expression correlation matrix for all immune genes measured

291   prior to breeding, in early life. We used Spearman Rank correlation coefficients (1) in

292   case of any non-linear relationship between genes, and (2) because our expression

293   data were not normally distributed. There was some missing data in our expression

294   dataset (640/4460 values = 14%). In order to maximise the sample size used to

295   generate each correlation coefficient, we used pairwise complete observations. This

296   resulted in 72% of correlations based on at least 75% of the total data. For each

297   correlation coefficient, we randomly permuted the data 1000 times to calculate a p-

298   value. We also adjusted all p-values for multiple testing using the Benjamini-

299   Hochberg method. We thresholded our correlation network using these corrected p-

300   values – only keeping those edges with a significant corrected p-value ($p \leq 0.05$).

301   Genes were clustered on edge betweenness, using the edge betweenness algorithm

302   in igraph[27]. The edge betweenness score of an edge is a measure of the number of

303   shortest paths that go through it. The algorithm identifies densely connected

304   modules by gradually removing edges with highest edge betweenness scores[28].

305

306   We then ran an exploratory analysis, repeating the steps above to construct a

307   correlation network for all immune genes measured in early life and six additional

308   measures of later life or lifetime success. These were: (1) whether or not an

309   individual was recaptured (a proxy for survival), (2) reproductive success in later life

310   (number of offspring in the pedigree), the proportion of later life infected with two

311   major microparasites in our population, namely (3) the bacterium *Bartonella* spp. and

13

312   (4) the protozoan *Babesia microti*, (5) mean Scaled Mass Index (SMI; a measure of

313   an individual's average condition across their lifetime[29]), and (6) coefficient of

314   variation in SMI (a measure of an individual's resilience in maintaining their condition

315   across their lifetime; with the caveat that this measure of lifetime of success will be

316   less reliable when based on fewer captures; see Supplementary Fig. 2). There was a

317   lack of variation in the first of these measures, whether or not an individual was

318   recaptured (total of 191 recaptured and 32 not recaptured) which led to zero

319   variance warnings and missing correlation coefficients for some pairwise

320   comparisons. This measure was therefore omitted from the network and instead

321   tested for associations with immune parameters of interest using a Fisher's exact

322   test (see below). There were slightly more missing data in our final combined

323   expression and success dataset (1078/5575 values = 19%). As before, we used

324   pairwise complete observations. This time, with 45% of correlations based on at

325   least 75% of the total data.

326

327   **Confirming associations between immune genes and measures of later life**

328   **success**

329   Follow-up analyses were run to confirm the two correlations of interest: (1) a

330   significant positive correlation between *Il17* expression in early life and reproductive

331   success, and (2) a significant positive correlation between *Il10* expression in early

332   life and proportion of later life infected with *Bartonella* spp. (see Results).

333

334   We had information about both *Il17* expression in early life *and* reproductive success

335   for 131 voles. Our measure of reproductive success was zero-inflated (107/131 or

336   82% zero values) and was simply coded as either reproduced or not. We then ran a

337   logistic regression to model the probability of reproducing. *Il17* expression in early

338   life was also found to be zero-inflated (117/131 or 89% zero values) and was simply

339   coded as either expressed or not. We included birth month as a (continuous)

340   covariate in the model, given that autumn-born voles have a lower chance of

341   reproducing than spring-born voles[30]. Other (categorical) covariates included in the

342   model as fixed effects were: sex, whether or not an individual was culled for the

343   cross-sectional component of this study (again, reducing the opportunity to

344   reproduce) and the year in which an individual was born. We also included a random

345   effect for the site in which an individual was born.

346

347   We had information about *Il10* expression in early life *and* future *Bartonella* infection

348   for 104 voles. *Bartonella* infection status was assessed multiple times for the majority

349   of individuals in the longitudinal component of the study (mean = 2.3; range = 1-6).

350   Therefore, we ran a logistic regression, with a random effect for individual, to model

351   the probability of an individual being infected throughout their life. The response

352   variable was whether (1) or not (0) an individual tested positive at a particular time.

353   As was the case for *Il17*, *Il10* expression in early life was found to be zero-inflated

354   (89/104 or 86% zero values) and was simply coded as either being expressed or not.

355   Other individual-specific covariates, considered potential drivers of future *Bartonella*

356   infection and included as fixed effects were: sex, whether or not an individual was

357   targeted for anti-parasite treatment (which may have affected the flea vectors of

358   *Bartonella*), whether or not they were already infected with *Bartonella* in early life and

359   birth year. We also included a random effect for birth site.

360

361 Logistic regressions were run using the R package glmmADMB[31,32]. All covariates

362 were tested for independence using variance inflation factors (all VIFs < 3). Full

363 submodel sets were generated from each global model, including all fixed and

364 random terms of interest, using the MuMIn package (Bartoń 2016). All candidate

365 models were then evaluated and ranked on relative fit using the Akaike Information

366 Criterion, AIC. The best model, with the lowest AIC, is reported in the text and

367 figures.

368

369 **Testing the age-specificity of these associations**

370 To test whether these associations are strictly early-life or also present in adult

371 (breeding) samples, we re-ran the same best models for reproductive success and

372 proportion of later life infected with *Bartonella* on adult blood samples. We

373 maximised the sample size for these analyses by including all individuals sampled as

374 adults, whether or not they were first sampled prior to breeding and therefore had an

375 approximate birth date. For this reason, we were unable to account for birth month

376 (reproductive success) and birth year (reproductive success and *Bartonella*

377 infection). The best models were otherwise unchanged. After omitting those samples

378 with missing data, we had information about both *Il17* expression in adult life *and*

379 reproductive success for 357 voles, and information about both *Il10* expression in

380 adult life and *Bartonella* infection for 798 voles. For individuals with more than one

381 adult sample, we selected one adult sample at random from which to measure

382 expression. To ensure that our results were robust to different random draws, we

383 repeated this process 1000 times, re-running the best models on 1000 random

384 draws. We report the number of random draws for which each association was

385 significant ($p \leq 0.05$) compared to that expected by chance. We also tested for an

16

386   association between whether or not an individual expressed each immune parameter

387   of interest early in life, and whether or not they expressed it in later life using a

388   Fisher's exact test.

389

390   **Testing for associations between immune genes of interest and genetic**

391   **polymorphisms**

392   We looked for associations between the immune parameters of interest (*Il10* and *Il17*

393   expression in early life) and genetic polymorphism. In order to avoid difficulties in

394   interpretation due to multiple testing, we looked only at polymorphisms in the

395   immune genes found within immune cluster 3 (see Results and Fig. 1; *Il17* = 1 SNP,

396   *Il10* = 2 SNPs, *Gata3* = 2 SNPs). We used the R package hapassoc (for *Il10* and

397   *Gata3*, for which we had information about more than one SNP) and the package

398   SNPassoc (for *Il17*, for which we had information about a single SNP). Both

399   hapassoc and SNPassoc models assumed an additive genetic model.

400

401   <u>**Results**</u>

402

403   **Three main clusters of immune genes are visible in early life**

404   We constructed a correlation network for 20 immune genes whose expression was

405   measured in our voles prior to breeding. Following thresholding, one immune gene,

406   *Il4*, dropping out of the network. Using clustering analysis we identified three main

407   immune clusters in early life: (1) immune cluster 1 (*Il1rap, Il1rn, Il1b, Orai1, Retnlg,*

408   *Apobr, Cd8a, Ifng, Tgfb1, Cd86*), (2) immune cluster 2 (*Irf2, Irf9, Foxp3, Tollip*) and

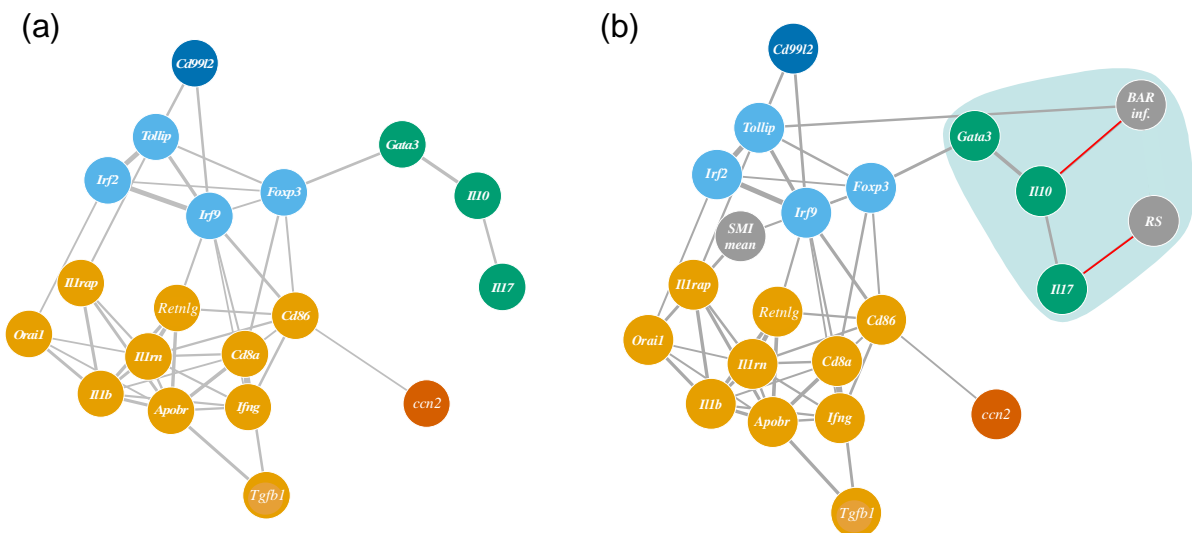409   (3) immune cluster 3 (*Gata3, Il10, Il17*; Fig. 1a).

410

17

411

412 **Fig. 1** Network showing significant correlations (and associated nodes) of magnitude

413 ≥ 0.2 between (a) immune parameters in early life (coloured nodes) and (b) immune

414 parameters in early life and measures of future success (grey nodes). Nodes without

415 any significant correlations of magnitude ≥ 0.2 are not shown in either (a) or (b).

416 Width of an edge represents exact magnitude of the Spearman rank correlation

417 coefficient. Three main clusters of immune parameters emerge based on edge

418 betweenness (indicated by the colour of nodes in (a): orange, light blue and green).

419 One of these (cluster with green nodes and also highlighted in green) was closely

420 associated with two measures of future success (*Bartonella* infection and

421 reproductive success) and was investigated further, in particular the edges

422 highlighted in red in (b).

423

424 **Exploratory analysis points to immune cluster 3 being closely associated with**

425 **measures of later life success**

426 We then ran an exploratory analysis, repeating the same process as above, but

427 adding six measures of later life success. Two measures of later life success

18

428   (coefficient of variation in SMI and proportion of later life infected with *Babesia*) were

429   not correlated with any other measure and dropped out of the network. Both immune

430   clusters 1 and 2 were associated with mean SMI. This included a significant negative

431   correlation between *Orai1* expression in early life (immune cluster 1) and mean SMI

432   (*rho* = -0.33; corrected *p* < 0.01), and a significant positive correlation between

433   *Foxp3* expression in early life (immune cluster 2) and mean SMI (*rho* = 0.23;

434   corrected *p* < 0.05). However, the immune cluster with the most associations with

435   measures of later life success was immune cluster 3. Two measures of later life

436   success (reproductive success and the proportion of later life infected with *Bartonella*

437   spp.) were associated with immune cluster 3 (cluster highlighted in green in Fig. 1b).

438   This included a significant positive correlation between *Il17* expression in early life

439   and reproductive success (*rho* = 0.22; corrected *p* = 0.04), and a significant positive

440   correlation between *Il10* expression in early life and proportion of later life infected

441   with *Bartonella* spp. (*rho* = 0.22; corrected *p* = 0.05; edges highlighted in red in Fig.

442   1b).

443

444   **Follow-up analysis confirms that *Il17* expression in early life is associated with**

445   **increased odds of reproducing later in life**

446   A logistic regression confirmed that whether or not an individual expressed *Il17* in

447   early-life was significantly associated with their probability of reproducing in later life

448   (odds ratio = 5.10; 95% CI = 1.28 – 20.32; *p* = 0.02). Other fixed effects which

449   appeared in the best model for, and were significantly associated with, the probability

450   of reproducing in later life were the year in which an individual was born, and the

451   month in which it was born (Fig. 2a). There was no significant association between

452   *Il17* expression in later (breeding) life and reproductive success (*n* = 7/1000 random

453   samples for which p ≤ 0.05; many fewer than the 50/1000 expected by chance)

454   indicating that this association was specific to early-life expression of *Il17*. Consistent

455   with this result, we found no association between whether or not an individual

456   expressed *Il17* early in life, and whether or not they expressed *Il17* later in life

457   (Fisher's exact test; odds ratio = 0.64; 95% CI = 0.18 – 2.03; *p* = 0.45).

458

459   ***Il10* expression in early life is associated with an increased proportion of later**

460   **life infected with *Bartonella* spp. and decreased odds of recapture (a proxy for**

461   **survival)**

462   A logistic regression confirmed that whether or not an individual expressed *Il10* in

463   early-life was significantly associated with their probability of being infected with

464   *Bartonella* in later life (odds ratio = 9.98; 95% CI = 1.29 – 77.08; *p* = 0.003). Year in

465   which an individual was born was the only other fixed effect present in the best

466   model for, and was also significantly associated with, *Bartonella* infection (Fig. 2b).

467   The association between *Il10* expression in later life and probability of infection was

468   also significant (*n* = 993/1000 random samples for which p ≤ 0.05; many more than

469   the 50/1000 expected by chance) indicating that this association was not specific to

470   early-life expression. We found no association between whether or not an individual

471   expressed *Il10* early in life, and whether or not they expressed *Il10* later in life

472   (Fisher's exact test; odds ratio = 1.05; 95% CI = 0.25 – 3.90; *p* = 1.00). *Bartonella*

473   infection status in early life did not appear in the best model for explaining the

474   probability of infection in later life, but there was a significant association between

475   this variable and *Il10* expression in early life (Fisher's exact test; odds ratio = 5.90;

476   95% CI = 1.38 – 53. 41; *p* = 0.01). In addition, we found a significant negative

20

477   association between *Il10* expression in early life and probability of recapture (a proxy

478   for survival; Fisher's exact test, odds ratio = 0.39; 95% CI = 0.15 – 1.03; *p* = 0.05).

479

(a) **Reproductive success**   (b) *Bartonella* **infection**
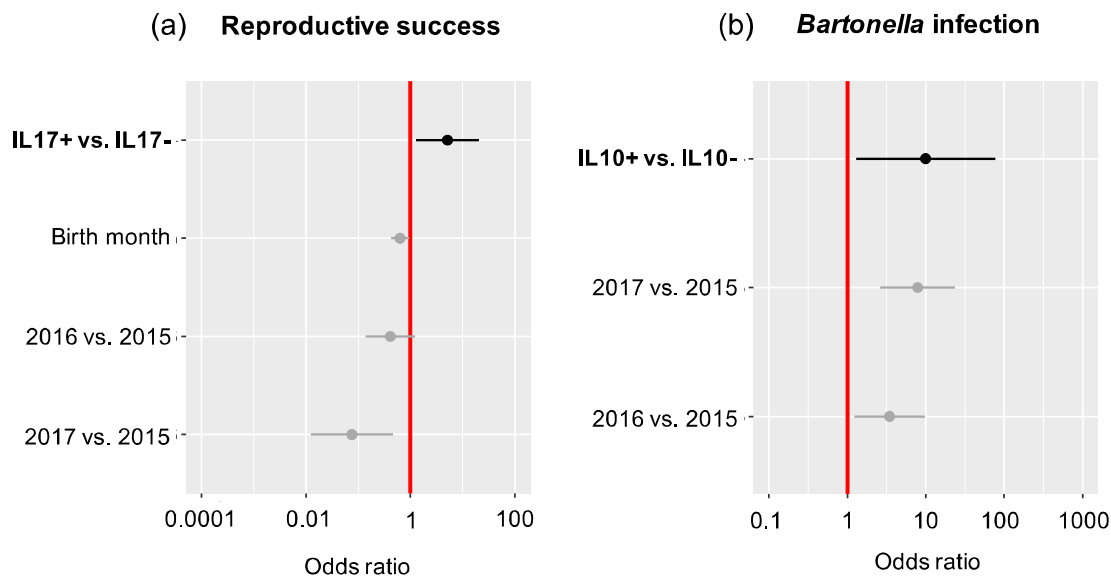
480

481   **Fig. 2** Odds of later-life success given immune parameters in early-life. In bold: (a)

482   odds of reproduction for individuals expressing *Il17* in early life (IL17+) compared to

483   individuals not expressing *Il17* (IL17-), and (b) odds of *Bartonella* infection for

484   individuals expressing *Il10* in early life (IL10+) compared to individuals not

485   expressing *Il10* (IL10-). Not in bold: odds of reproduction or infection associated with

486   other fixed effects which also appeared in best models: year in which an individual

487   was born (2015, 2016 or 2017) and month in which an individual was born (treated

488   as continuous variable). All estimates taken from best models (see text). Error bars

489   represent 95% confidence intervals.

490

491   ***Il10* expression in early life is associated with a single SNP in the *Il17* gene**

492   We looked for associations between the immune parameters of interest (*Il10* and *Il17*

493   expression in early life) and genetic polymorphism. We found a significant

21

494   association between a single SNP in the *Il17* gene and *Il10* expression in early life

495   under an additive model, with the probability of expressing *Il10* significantly

496   increasing with the number of T alleles (odds ratio = 1.74; 95% CI = 1.03 – 2.94; *p* =

497   0.04; Fig. 3). *Il17* genotype only explained a small proportion of variation in *Il10*

498   expression (r^2 = 0.026).

499

Early life   **Later life**

Bartonella infection

***Il17* genotype** ⟶ *Il10* expression

Recapture (proxy for survival)

Il17 expression ⟶ Reproductive success

500

501

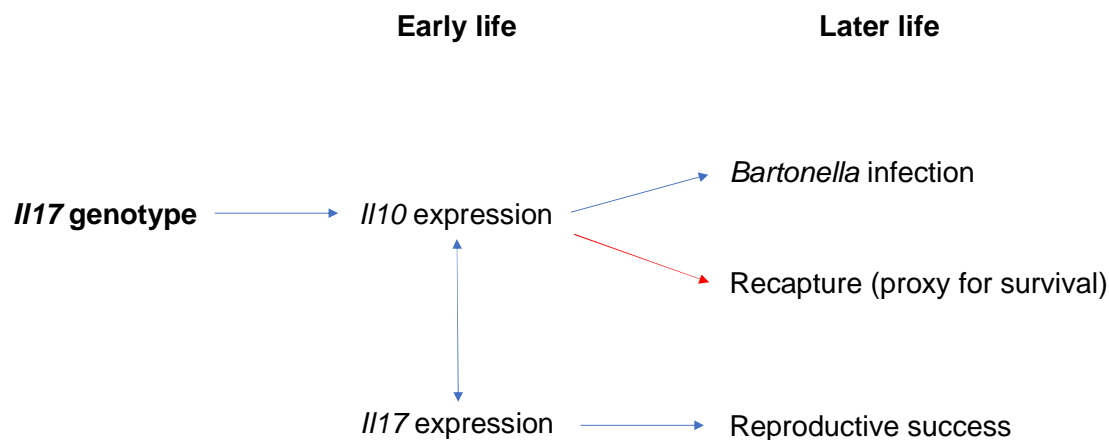502   **Fig. 3** Summary diagram of confirmed associations. Blue arrows indicate positive

503   associations, red arrows indicate negative associations. One-way arrows indicate a

504   likely causal association, two-way arrows indicate a correlation only.

505

506   **Discussion**

507

508   We have used a wild rodent system to provide evidence that, (1) the reproductive

509   success (a measure of fitness), and (2) the response to infection of mature

510   individuals (a measure of health) is foreshadowed by immune expression in the early

511   stages of their lives. We have also found evidence to suggest that these patterns

512   vary between different parts of the immune network.

22

513

514 We were able to derive three clusters of genes related to immune function based on

515 co-expression data and it was unsurprising to see genes involved in different

516 pathways, and with mixed pro- and anti-inflammatory functions, included in the same

517 clusters. This is expected given the complexity of immune regulatory networks. For

518 example, some responses might increase at the same time as their negative

519 feedbacks, or responses within one pathway might lead to host states that favour the

520 expression of another pathway. Nonetheless, we saw very clear functional biases,

521 with the largest cluster containing predominantly pro-inflammatory markers such as

522 *Il1* signalling cascade members[33], *Ifng*[34], *Tgfb1*[35] and *Retnlg*[36], but also genes with

523 more nuanced functions in the promotion and suppression of inflammation, including

524 *Tgfb1*[37] itself and *Orai1*[38]. This "pro-inflammatory" cluster was distinct from a second

525 cluster containing predominantly immune down-regulatory markers including

526 *Foxp3*[39], *Tollip*[40] and *Irf2*[41], but also I*rf9*[42] a driver of interferon effector responses. A

527 third cluster containing *Gata3*, *Il10* and *Il17* may give new insights into immune

528 function, given that it contained a counterintuitive set of genes and was, at the same

529 time, the cluster best linked to host health and fitness. We have previously found

530 GATA3 to be associated with immune tolerance phenotypes, and the association

531 with the anti-inflammatory cytokine IL10 provides further support of this finding[43,44].

532 *Il10* expression is, in turn, positively associated with *Il17* expression. IL17, which we

533 found to be an early-life indicator of reproductive success, is typically associated with

534 a pro-inflammatory immunopathology but may also be protective against the

535 numerous primary microbial insults that young voles, lacking acquired immunity, may

536 encounter[45]. The correlation between IL10 and IL17 may thus be one whereby IL10

537 is produced in response to IL17 in order to safely regulate its inflammatory effects

538   and prevent immunopathology[46–49]. We argue that wild rodent studies can help us to

539   define which of the possible immunological interactions known from laboratory

540   studies predominate and shape the structure of immune phenotypes in individuals

541   exposed to pathogens, nutritional challenge or other stressors within the natural

542   environment. As in the case of IL10 and IL17, wild rodents may also give insights

543   into immune regulation within a natural context.

544

545   Having identified the correlational structure of immunological traits in a wild

546   population, it is then a natural extension to include other phenotypic traits. We used

547   longitudinal data to address our aim to determine how immune phenotypes may be

548   laid down in early life to affect later life health and fitness. We found two pathways

549   through which immune phenotypes in early life act on later life traits (Fig. 3). In the

550   first, increased expression of *Il10* in blood samples taken during early life (i.e., prior

551   to breeding) was subsequently associated with both a ten-fold increased

552   susceptibility to *Bartonella* infection later in life and a three-fold reduction in

553   survivorship. The association of a genotypic effect on *Il10* expression, a

554   polymorphism in *Il17*, suggests that early-life effects may be, at least in part, laid

555   down at birth in this case. A challenge for future work will be to tease out causation

556   from correlation in this case. An adverse *Il17* genotype may make an individual more

557   likely to become infected with *Bartonella*, with infection stimulating *Il10* expression.

558   Alternatively, that genotype may predispose an individual to express *Il10* early in life

559   (and perhaps lead to other immune-expression variations) with that pattern of

560   immune expression increasing the likelihood of *Bartonella* infection. The positive

561   association between *Il10* expression in the earliest samples and *Bartonella* infection

562   at that time supports but cannot positively establish the former, since *Bartonella*

563   infection before even the earliest samples is possible. Further, it is unclear whether it

564   is *Bartonella* infection or the pattern of immune expression that reduces survival.

565

566   We also found a significant association between I*l10* expression in breeding adults

567   and the proportion of later life infected with *Bartonella*. It is perhaps not surprising

568   that we found this association, as both *Il10* expression and *Bartonella* infection in this

569   case were measured in contemporaneous samples (taken in later life) and our

570   previous work shows that *Il10* expression increases soon after acquisition of

571   *Bartonella*[50]. Furthermore, we found no significant association between *Il10*

572   expression early and later in life, suggesting that the later-life association was not

573   driven by its early-life counterpart.

574

575   The second pattern that we found is for increased *Il17* expression early in life to be

576   associated with a subsequent five-fold increased likelihood of reproduction. By

577   contrast, *Il17* expression in breeding adults showed no such association with

578   reproduction, and indeed, no association either between *Il17* expression early and

579   later in life. In contrast to *Il10* expression, no genetic effects on *Il17* expression were

580   detected; at least from the deliberately limited set of cytokine polymorphisms

581   considered, which included *Il17* polymorphism itself. This suggests that early life

582   expression of *Il17* may be the consequence of an environmental effect leading to an

583   immune phenotype that acts on juveniles and influences their development as they

584   transition into reproductively mature adults. This could be, for example, a bacterial

585   infection or composition of the gut microbiome early in life.

586

587   Hence, despite the contrasts between the patterns for *Il10* and *Il17* expression, both

588   appear to be examples where the die has been cast early: immune expression early

589   in life, however determined, is a powerful predictor of the future life course of the

590   individual concerned. Previous evidence for this has been scarce. A number of

591   studies have shown an association between broad-brush measures of immune

592   expression in nestling birds and recruitment from the nest[13–16], an almost

593   contemporaneous effect. Bowers *et al.*[13] went a step further, and showed an

594   association between a broad-brush measure of immune expression (PHA

595   responsiveness), longevity and reproductive success. Here we provide a more

596   refined analysis, and describe two examples in which differences detectable early in

597   life in the expression of distinct immune pathways play out much later in life, or at

598   least over a much more extended period. *Il10* differences confirm a genetic basis for

599   immune expression that, as previously shown in this system and others, act through

600   the network of interacting cytokines such that polymorphism at one cytokine gene

601   can affect expression of other cytokines and immune phenotypes[23,51–56], with these

602   effects further percolating to affect health and fitness more broadly in later life. By

603   contrast, the *Il17* differences appear to be more reminiscent of the Jesuit motto "Give

604   me the boy until he is seven and I will give you the man": that independently of any

605   genetic differences, an individual's experiences early in life may, in this case through

606   the medium of immune expression, profoundly influence their later health and

607   fitness.

608

609

610

611

26

References

612   **References**

613

614   1.   Barker, D. J. P. The origins of the developmental origins theory. *J. Intern. Med.*

615   **261**, 412–417 (2007).

616   2.   Wadhwa, P., Buss, C., Entringer, S. & Swanson, J. Developmental Origins of

617   Health and Disease: Brief History of the Approach and Current Focus on

618   Epigenetic Mechanisms. *Semin. Reprod. Med.* **27**, 358–368 (2009).

619   3.   Dowling, D. J. & Levy, O. Ontogeny of early life immunity. *Trends Immunol.* **35**,

620   299–310 (2014).

621   4.   Brodin, P. *et al.* Variation in the human immune system is largely driven by

622   non-heritable influences. *Cell* **160**, 37–47 (2015).

623   5.   Stewart, C. J. *et al.* Temporal development of the gut microbiome in early

624   childhood from the TEDDY study. *Nature* **562**, 583–588 (2018).

625   6.   Khafipour, E. & Ghia, J. Mode of Delivery and Inflammatory Disorders. *J.*

626   *Immunol. Clin. Res.* **1**, 1004 (2013).

627   7.   Shao, Y. *et al.* Stunted microbiota and opportunistic pathogen colonization in

628   caesarean-section birth. *Nature* **574**, 117–121 (2019).

629   8.   Ober, C., Sperline, A., E, M. & Vercelli, D. Immune Development and

630   Environment: Lessons from Amish and Hutterite Children. *Curr. Opin.*

631   *Immunol.* **48**, 51–60 (2017).

632   9.   Von Mutius, E. & Vercelli, D. Farm living: Effects on childhood asthma and

633   allergy. *Nat. Rev. Immunol.* **10**, 861–868 (2010).

634   10.   Peterson, R. O., Vucetich, J. A., Fenton, G., Drummer, T. D. & Larsen, C. S.

635   Ecology of arthritis. *Ecol. Lett.* **13**, 1124–1128 (2010).

636   11.   Woodell-May, J. E. & Sommerfeld, S. D. Role of Inflammation and the Immune

637 System in the Progression of Osteoarthritis. *J. Orthop. Res.* **38**, 253–257

638 (2020).

639 12. Graham, A. L. *et al.* Fitness correlates of heritable variation in antibody

640 responsiveness in a wild mammal. *Science* **330**, 662–665 (2010).

641 13. Bowers, E. K. *et al.* Neonatal body condition, immune responsiveness, and

642 hematocrit predict longevity in a wild bird population. *Ecology* **95**, 3027–3034

643 (2014).

644 14. Cichoń, M. & Dubiec, A. Cell-mediated immunity predicts the probability of

645 local recruitment in nestling blue tits. *J. Evol. Biol.* **18**, 962–966 (2005).

646 15. López-Rull, I., Celis, P., Salaberria, C., Puerta, M. & Gil, D. Post-fledging

647 recruitment in relation to nestling plasma testosterone and

648 immunocompetence in the spotless starling. *Funct. Ecol.* **25**, 500–508 (2011).

649 16. Moreno, J. *et al.* Nestling cell-mediated immune response, body mass and

650 hatching date as predictors of local recruitment in the pied flycatcher *Ficedula*

651 *hypoleucai. J. Avian Biol.* **36**, 251–260 (2005).

652 17. Turner, A. K. & Paterson, S. Wild rodents as a model to discover genes and

653 pathways underlying natural variation in infectious disease susceptibility.

654 *Parasite Immunol.* **35**, 386–395 (2013).

655 18. Begon, M. *et al.* Effects of abundance on infection in natural populations: field

656 voles and cowpox virus. *Epidemics* **1**, 35–46 (2009).

657 19. Gebert, S. Patterns in helminth infections in natural populations of rodents.

658 (University of Liverpool, 2008).

659 20. Rowe, F., Bradfield, A., Quy, R. & Swinney, T. Relationship Between Eye Lens

660 Weight and Age in the Wild House Mouse (*Mus musculusi*). *J. Anim. Ecol.* **22**,

661 55–61 (1985).

662 21. Jackson, J. A. *et al.* The analysis of immunological profiles in wild animals: A

663   case study on immunodynamics in the field vole, *Microtus agrestis. Mol. Ecol.*

664   **20**, 893–909 (2011).

665 22. Jackson, J. A. *et al.* An immunological marker of tolerance to infection in wild

666   rodents. *PLoS Biol.* **12**, e1001901 (2014).

667 23. Wanelik, K. M. *et al.* A candidate tolerance gene identified in a natural

668   population of field voles (*Microtus agrestis*). *Mol. Ecol.* **27**, 1044–1052 (2018).

669 24. Huisman, J. Pedigree reconstruction from SNP data: parentage assignment,

670   sibship clustering and beyond. *Mol. Ecol. Resour.* **17**, 1009–1024 (2017).

671 25. R Core Team. R: A language and environment for statistical computing. R

672   Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-

673   project.org/. (2018).

674 26. Wanelik, K. *et al.* IgE receptor polymorphism predicts divergent, sex-specific

675   inflammatory modes and fitness costs in a wild rodent. *biorXiv* (2019).

676 27. Csardi, G. & Nepusz, T. The igraph software package for complex network

677   research. *InterJournal, Complex Syst. 1695* (2006).

678 28. Newman, M. E. J. & Girvan, M. Finding and evaluating community structure in

679   networks. *Phys. Rev. E* **69**, 026113 (2004).

680 29. Peig, J. & Green, A. J. New perspectives for estimating body condition from

681   mass/length data: The scaled mass index as an alternative method. *Oikos*

682   **118**, 1883–1891 (2009).

683 30. Wang, M., Ebeling, M. & Hahne, J. Relevance of body weight effects for the

684   population development of common voles and its significance in regulatory risk

685   assessment of pesticides in the European Union. *Environ. Sci. Eur.* **31**, (2019).

686 31. Fournier, D. A. *et al.* AD Model Builder: Using automatic differentiation for

687   statistical inference of highly parameterized complex nonlinear models. *Optim.*

688   *Methods Softw.* **27**, 233–249 (2012).

689   32.   Skaug, H., Fournier, D., Bolker, B., Magnusson, A. & Nielsen, A. Generalized

690   Linear Mixed Models using 'AD Model Builder'. R package version 0.8.3.3.

691   (2016).

692   33.   Dinarello, C. A. Overview of the IL-1 family in innate inflammation and acquired

693   immunity. *Immunological Reviews* **281**, 8–27 (2018).

694   34.   Schroder, K., Hertzog, P. J., Ravasi, T. & Hume, D. A. Interferon-γ: an

695   overview of signals, mechanisms and functions. *J. Leukoc. Biol.* **75**, 163–189

696   (2004).

697   35.   Han, G., Li, F., Singh, T. P., Wolf, P. & Wang, X. J. The pro-inflammatory role

698   of TGFβ1: A paradox? *Int. J. Biol. Sci.* **8**, 228–235 (2012).

699   36.   Nagaev, I., Bokarewa, M., Tarkowski, A. & Smith, U. Human resistin is a

700   systemic immune-derived proinflammatory cytokine targeting both leukocytes

701   and adipocytes. *PLoS One* **1**, e31 (2006).

702   37.   Li, M. O. & Flavell, R. A. TGF-β: A master of all T cell trades. *Cell* **134**, 392–

703   404 (2008).

704   38.   Shaw, P. J. & Feske, S. Regulation of lymphocyte function by ORAI and STIM

705   proteins in infection and autoimmunity. *J. Physiol.* **590**, 4157–4167 (2012).

706   39.   Shevach, E. M. Mechanisms of Foxp3+ T Regulatory Cell-Mediated

707   Suppression. *Immunity* **30**, 636–645 (2009).

708   40.   Zhang, G. & Ghosh, S. Negative regulation of toll-like receptor-mediated

709   signaling by Tollip. *J. Biol. Chem.* **277**, 7059–7065 (2002).

710   41.   Blanco, J. C. G. *et al.* Interferon regulatory factor (IRF)-1 and IRF-2 regulate

711   interferon γ- dependent cyclooxygenase 2 expression. *J. Exp. Med.* **191**,

712   2131–2144 (2000).

713   42. Jefferies, C. A. Regulating IRFs in IFN driven disease. *Front. Immunol.* **10**,

714   (2019).

715   43. Coomes, S. M. *et al.* CD4 + Th2 cells are directly regulated by IL-10 during

716   allergic airway inflammation. *Mucosal Immunol.* **10**, 150–161 (2017).

717   44. Couper, K. N., Blount, D. G. & Riley, E. M. IL-10: The Master Regulator of

718   Immunity to Infection. *J. Immunol.* **180**, 5771–5777 (2008).

719   45. Xu, S. & Cao, X. Interleukin-17 and its expanding biological functions. *Cell.*

720   *Mol. Immunol.* **7**, 164–174 (2010).

721   46. Li, C., Corraliza, I. & Langhorne, J. A defect in interleukin-10 leads to

722   enhanced malarial disease in *Plasmodium chabaudi chabaudii* infection in

723   mice. *Infect. Immun.* **67**, 4435–4442 (1999).

724   47. Gazzinelli, R. *et al.* In the Absence of Endogenous IL-10, Mice Acutely Infected

725   with *Toxoplasma gondiii* Succumb to a Lethal immune Response Dependent

726   on CD4+ T Cells and Accompanied by Overproduction of 11-12, IFN-y, and

727   TNF-cw. *J. Immunol.* **157**, 798–805 (1996).

728   48. Gu, Y. *et al.* Interleukin 10 suppresses Th17 cytokines secreted by

729   macrophages and T cells. *Eur. J. Immunol.* **38**, 1807–1813 (2008).

730   49. Kulcsar, K. A., Baxter, V. K., Greene, I. P. & Griffin, D. E. Interleukin 10

731   modulation of pathogenic Th17 cells during fatal alphavirus encephalomyelitis.

732   *Proc. Natl. Acad. Sci. U. S. A.* **111**, 16053–16058 (2014).

733   50. Taylor, C. H. *et al.* Physiological, but not fitness, effects of two interacting

734   haemoparasitic infections in a wild rodent. *Int. J. Parasitol.* **48**, 463–471

735   (2018).

736   51. Dinarello, C. A. Historical insights into cytokines. *Eur. J. Immunol.* **37**, 34–45

737   (2007).

738   52.   ter Horst, R. *et al.* Host and Environmental Factors Influencing Individual

739   Human Cytokine Responses. *Cell* **167**, 1111-1124.e13 (2016).

740   53.   Li, Y. *et al.* A Functional Genomics Approach to Understand Variation in

741   Cytokine Production in Humans. *Cell* **167**, 1099-1110.e14 (2016).

742   54.   Salnikova, L. E. *et al.* Cytokines mapping for tissue-specific expression, eQTLs

743   and GWAS traits. *Sci. Rep.* **10**, 1–14 (2020).

744   55.   Turner, A. K., Begon, M., Jackson, J. a., Bradley, J. E. & Paterson, S. Genetic

745   Diversity in Cytokines Associated with Immune Variation and Resistance to

746   Multiple Pathogens in a Natural Rodent Population. *PLoS Genet.* **7**, e1002343

747   (2011).

748   56.   Viney, M. & Riley, E. M. The immunology of wild rodents: Current status and

749   future prospects. *Front. Immunol.* **8**, 1–9 (2017).

750

Acknowledgements

760

761