Decomposing variation in immune response in a wild rodent population

Klara M. Wanelik[1*], Mike Begon[1], Elena Arriero[2,3], Janette E. Bradley[2], Ida M. Friberg[4], Joseph A. Jackson[4], Christopher H. Taylor[2], Steve Paterson[1]

[1] Institute of Integrative Biology, University of Liverpool, Liverpool, United Kingdom

[2] School of Life Sciences, University of Nottingham, Nottingham, United Kingdom

[3] Department of Biodiversity, Ecology and Evolution, University Complutense of Madrid, Madrid, Spain

[4] School of Environment and Life Sciences, University of Salford, Salford, United Kingdom

[*] Corresponding author

Email: klara.wanelik@liverpool.ac.uk (KMW)

Abstract

Individuals vary in their immune response and, as a result, some are more susceptible to infectious disease than others. Little is known about which components of immune pathways are responsible for this variation, but understanding these underlying processes could allow us to predict the outcome of infection for an individual, and to manage their health more effectively. In this study, we describe transcriptome-wide variation in immune response (to a standardised challenge) in a wild population of field voles (*Microtus agrestis*). We find that this variation can be categorised into three main types. We also identify markers, across these three categories, which display particularly strong individual variation in response. This work shows how a simple standardised challenge performed on a natural population can reveal complex patterns of natural variation in immune response.

Introduction

Individuals vary in their immune response. Within a population, some individuals may fail to make protective immune responses following either natural infection or vaccination and so are especially vulnerable to infectious disease[1–4]. Defining the patterns of such variability will enhance our ability to manage the health of individuals – especially those that are most susceptible to infectious disease in human, livestock or wildlife populations.

Studies in laboratory mice are the cornerstone of immunology and have provided a detailed understanding of the molecular and cellular pathways by which immune responses are effected. This impressive mechanistic understanding, however, has only been achieved by minimising genetic and environmental variation within a laboratory setting. Where laboratory

studies have examined the effects of variability – in genetics, microbiota or diet – both qualitative and quantitative differences in immune responses have been observed, with consequent effects on infection[5–7]. Nevertheless, natural variability cannot be fully reproduced in the laboratory, which has led to a recent effort to characterise the immune response in wild populations of mice or other rodents. Recent work in mice from agricultural and other anthropogenic settings is consistent with the expectation that exposure to complex environments greatly alters immune function[8]. New populations of memory T cells, present only in non-laboratory mice, have also been identified[9].

One commonly used measure of an immune response is to assess the amount of one or more markers (e.g. transcripts or proteins) produced by a population of cells following stimulation by an immune agonist. From this *ex vivo* assay, one can gain insight into the types of immune response that could be made to a pathogen *in vivo*. Such responses depend on the cell types, the time points and the immune agonist used. Nevertheless, for any molecular marker with such a response, individuals, in natural populations especially, could exhibit different marker abundances prior to and/or following stimulation, leading to differences in their response to stimulation (here defined as the difference between marker abundances prior to and following stimulation). Furthermore, the most useful (and interesting) markers, in terms of understanding why individuals vary in their ability to mount a successful immune response, will be those for which response is most variable among individuals. In the laboratory, cell populations are usually controlled, or at least well defined, so a difference in the abundance of a particular marker can be attributed to differences in the activity of a particular cell type. However, natural variability in the abundance of a marker, and by extension in the response of individuals in the wild, could result from (i) differences in the composition of cell populations, and/or (ii) differences in the activity levels of particular cell types. Both of these

components have the potential to shape the way an individual responds to immune challenge in the wild. Our intention here is not to distinguish between the two, but rather to propose a categorisation of responses, however generated.

We use a wild population of field voles (*Microtus agrestis*) to examine naturally occurring patterns of individual variation in immune response, across the transcriptome, as a first step towards furthering our understanding of the processes driving these patterns. The field population we study, in Kielder Forest Northumberland, has been the subject of extensive previous study on population ecology and pathogen dynamics[10–13]. Therefore, it allows us to place our existing understanding of more established immunological mechanisms (largely derived from the closely related laboratory mouse, *Mus musculus*) into a well-described, real-world context.

We describe three main categories of immune response: (i) uncorrelated response, (ii) constant response and (iii) baseline-dependent response (depicted in Fig. 1). We also identify markers, across these categories, which show particularly high inter-individual variability in response. We suggest that such categorisation is useful in organising natural immune variation, since little is known about which components of immune pathways are responsible for natural variability in immune response, or about the nature and possible causes of such variability. Indeed, this categorisation is not limited to the components of conventional immune pathways. The ability of an immune response to effect protection against infection, for example, will be supported by a variety of non-immune functions, that will also be activated following stimulation by an agonist, and vary to a greater or lesser extent among individuals within a natural population. By identifying the components (whether conventionally immunological or not) that are likely to be responsible for natural variability

in immune response, and by describing the nature of their variability, we are laying the groundwork for exploring the processes, whether genetic or environmental, which drive inter-individual variation in immune response.

Results

Stimulation with an immune agonist causes a widespread response

Spleen cells from sixty-two field voles were split into two populations per individual vole. One population was stimulated with anti-CD3 and anti-CD28 antibodies, while the other was kept as an unstimulated control (hereafter referred to as the baseline). 1150 transcripts (5% of all genes in the field vole genome and 85% of informative genes, those genes which were more strongly expressed; see Methods) fell into one or more of the response categories set out in Fig. 1. As expected, given that these antibodies are known to stimulate T-cell proliferation[14], they were enriched with transcripts (hereafter markers) associated with the T-cell receptor (TCR) signalling pathway ($n = 27$; $p < 0.001$; Functional Enrichment Analysis performed in DAVID; see Methods) and other T cell-related terms: positive regulation of T-cell proliferation ($n = 12$; $p < 0.03$), TCR complex ($n = 7$; $p < 0.001$), positive thymic T-cell selection ($n = 7$; $p < 0.01$), negative thymic T-cell selection ($n = 6$; $p = 0.03$) and alpha-beta TCR complex ($n = 5$; $p < 0.001$). For the majority of these markers, a significant positive linear relationship was found between baseline and stimulated abundance ($n = 844$). Only a single marker, *Fam193b*, demonstrated a significant negative linear relationship between baseline and stimulated abundance.

125 **There are three main categories of immune response**

126 Three main categories of immune response were identified based on the dependence of an
127 individual's response on its baseline abundance. Each of these categories demonstrates a
128 unique pattern (Fig. 1):

129

130 **Uncorrelated response:** markers for which individuals taken from the wild differ in their
131 baseline abundance, but the responses of different individuals are variable and independent of
132 their baseline, such that the slope of the relationship between baseline and stimulated
133 abundance is not significantly different from zero.

134

135 **Constant response:** markers for which individuals taken from the wild also differ in their
136 baseline abundance, but the responses of different individuals are (approximately) constant
137 and independent of their baseline, such that the slope of the relationship between baseline and
138 stimulated abundance is not significantly different from one and the intercept (indicating the
139 level of response) is significantly greater than zero.

140

141 **Baseline-dependent response:** markers for which individuals taken from the wild again
142 differ in their baseline abundance, but the responses of different individuals vary as a
143 function of their baseline level, either as a linear function of their baseline level (slope
144 significantly different from one), or as a quadratic function of their baseline level, where
145 stimulated levels either increase exponentially as a function of baseline levels or become
146 saturated at some upper limit.

147

We also identified markers, across these three categories, for which variability in baseline and stimulated samples was significantly different, leading to high inter-individual variability in response (see Methods). These can be divided into two categories (Fig. 1):

Convergent response: markers for which variability in baseline abundance is significantly greater than variability in stimulated abundance.

Divergent response: markers for which variability in stimulated abundance is significantly greater than variability in baseline abundance.

The baseline-dependent response category is most common and is significantly enriched in components of conventional immune pathways

The baseline-dependent response category was the most common (Table 1), and included a majority of markers for which stimulated levels were a linear function of baseline levels (n = 539), and a remainder for which they were a quadratic function (n = 160). The majority of quadratic response markers showed evidence for saturation (n = 138), indicating some upper limit on stimulated abundance. The general ontology term for immunity was enriched in the linear response category of markers (n = 20; $p < 0.01$). The TCR signaling pathway was enriched in the quadratic response category (n = 7; p = 0.01; Fig. 2).

The uncorrelated response category is least common and lacks enrichment in components of conventional immune pathways

A number of markers showed no evidence for a relationship between baseline and stimulated abundance (n = 47; Table 1). For the majority of these, mean abundance was significantly greater for stimulated than for baseline samples (n = 39), suggesting that these markers were

(on average) responding to stimulation, but to an individually variable degree, independent of baseline levels. These markers lacked any enrichment for immune-related terms (Fig. 2).

A number of markers, including *Zap70*, show particularly high inter-individual variability in response

For a number of markers, variability in baseline and stimulated abundance was significantly different, leading to high inter-individual variability in response (n = 244). The vast majority of these markers showed a divergent (n = 237), rather than a convergent (n = 7) response (Table 1). Within the (stimulated) TCR signalling pathway, the highest level of variability in individual response, and the highest level of divergence, was demonstrated by *Zap70* (Fig. 3). All convergent markers fell into one of the three main immune response categories. However, over a third of divergent markers (n = 98), did not fall into any of these categories, appearing instead as markers which (on average) did not respond to stimulation (Table 1). Mean abundances for these markers were also not significantly different between stimulated and baseline samples.

Juveniles show more inter-individual variability in response than adults

An age-specific analysis, run separately on samples from mature (n = 43) and juvenile (n = 19) field voles, showed that higher inter-individual variability in immune response (whether divergent or convergent) was more common among juvenile voles (no. divergent markers = 108; no. convergent markers = 6) than mature voles (randomly sampled 1000 times as more samples available; mean no. divergent markers = 50, empirical 95% interval = 0–338.2; mean no. of convergent markers = 0.11, empirical 95% interval = 0–1).

198 **Response to stimulation is not limited to components of conventional immune pathways**

199 Non-immune related terms were enriched in the baseline-dependent response category,
200 including: insulin signalling pathway (n = 9; p = 0.05) and thyroid hormone signalling
201 pathway (n = 8; p = 0.05). The top convergent response marker, *Pdk1*, is also a component of
202 the insulin signalling pathway (Fig. 2).
203

204 ## Discussion

205

206 The need to better understand variation in immune response in natural populations is now
207 widely accepted[15–18]. Our understanding of immune responses in laboratory settings comes
208 from animals that vary little either genetically or in prior experience. By contrast, animals in
209 natural populations vary (perhaps extensively) in both of these. In this study, we describe
210 natural variation in immune response in a wild population of rodents, and find that it can be
211 categorised into a limited number of types. We identify three main categories of immune
212 response: uncorrelated response, constant response and baseline-dependent response. We also
213 identify markers, across these categories, which show particularly high inter-individual
214 variability in response. Our work shows how a simple stimulatory assay performed on a
215 natural population can reveal underlying patterns of natural variation among individuals in
216 immune response.
217

218 The baseline-dependent response category is the largest. Markers in this category show a
219 relationship between baseline and stimulated abundance across individuals, and their
220 response to stimulation is (to a lesser or a greater extent) dependent on their baseline level. In
221 some cases, individuals already expressing the greatest abundance of a marker in their natural
222 setting went on to exhibit the greatest response to stimulation by an agonist. In others, the

223 opposite was true, and these individuals exhibited the smallest response to stimulation.
224 Similarly, previous work on humans has identified baseline (transcriptional) predictors of
225 influenza vaccination response[19,20]. These differences in baseline level could be driven by
226 either genetic variation or individual differences in past experience. In humans, genetic
227 determinants of baseline immune cell population frequencies have been identified[21]. Even
228 though the stimulation we describe here was not antigen specific, previous challenge by a
229 parasite might also lead to changes in the baseline T-cell population within an individual's
230 spleen, affecting its response to any subsequent challenge. In fact, we find that voles infected
231 with *Babesia microti* (a blood parasite, common in our population[22]) have larger spleens than
232 uninfected voles[13]. This prior experience may prime an individual, enabling a greater
233 response to subsequent challenge (e.g. slope greater than one; Fig. 1). However, individuals
234 may also have an upper limit on the number of immue cells they have available[23,24]. An
235 individual that is already mounting an immune response to a parasite, and has a large number
236 of activated T cells, could therefore respond less to a similar challenge than an
237 'immunologically naïve' individual (slope less than one; Fig. 1). Membership of the baseline-
238 dependent response category recapitulates the known biology of the immune response (being
239 highly enriched for immune ontogeny terms). In doing so, it validates the approach we use
240 here, as a way of identifying markers of immune significance.
241
242 In some cases, individuals varied in their natural abundance of a marker but their response
243 was unrelated to this. They did nevertheless respond to stimulation, with the majority of these
244 markers occurring at a significantly higher mean abundance in stimulated samples than in
245 baseline samples. This uncorrelated response category, which contains a moderate number of
246 markers, also lacks any enrichment for immune-related ontology terms. This suggests that
247 markers in this category are not conventional immune markers but could be of immune

248 significance. We warn against omitting such markers from studies of immune response in the
249 laboratory. They could play an important part in our understanding of the immune response,
250 indicating for example, genetic variation in response among individuals, which is
251 independent of baseline level.
252
253 Cutting across this categorisation, a large number of markers displayed a pattern in which
254 variation between individuals was particularly strong. We describe two types of such
255 markers, both of which could be used in future studies as indicators of natural variability in
256 immune response. Markers in the less common, convergent, response category showed much
257 greater variation naturally than following stimulation. This pattern may be characteristic of
258 markers showing variable levels of prior activation, coupled with some maximum or
259 optimum abundance, and resulting in a stabilisation of the immune response across the
260 population following stimulation. We found that convergent patterns were more common
261 among juvenile voles. This could suggest that they are more constrained in the energy they
262 have available (as a result of the competing energetic demands of growth and development)
263 or the number of immune cells they have available (as a result of a developing immune
264 system). Either resource constraint could result in a maximum abundance, making them more
265 inclined to converge. Due to the costly nature of the immune response, individuals often
266 trade-off their investment in different arms of the immune system[25,26]. Different types of
267 immune response are therefore likely to be associated with different optimum abundances (or
268 regions) and an individual already mounting an immune response, but to a different type of
269 challenge (associated with different cell types), may respond by down-regulating expression.
270
271 Divergent markers, which were more common, showed much greater variation following
272 stimulation than there was naturally. This pattern may be characteristic of (but not limited to)

markers showing genetic variation in response to the agonist, independent of baseline levels e.g. subsets of animals that appear similar but respond more strongly to stimulation than others. Our own recent work, where we found an association between polymorphism in a single gene and a marker of a more tolerant immune response[27], is an example of such genetic variation in immune response. Further supporting this hypothesis, here, we found more divergent markers among juvenile voles than mature voles. Younger voles are expected to have less variable exposure histories, as a result of their shorter life spans, making it easier to detect genetic effects. Equally, though, divergent patterns could be the result of differences in early life experiences. One would also expect these to be more easily detectable in juveniles.

The divergent category (predominantly) included markers for which individuals made (on average) the same response to stimulation and markers that did not respond (on average) to stimulation. Standard differential expression analysis would miss the individual variation present in the former group, and would fail to pick up the latter group of markers altogether. Both warn against looking at average (population-level) response, and point instead, to the value of looking at individual-level differences in immune response. This is particularly important because divergent markers may act as critical regulators of pathways. For example, *Zap70*, which demonstrates particularly high levels of variability in individual response and is centrally located in the TCR signalling pathway, interacts with many other markers (Fig. 3). We suggest that *Zap70* expression could be used as a marker of response in larger studies. Indeed, it is already linked to major seasonal immune variation in wild fish[28] and is being used as a prognostic marker for B-cell chronic lymphocytic leukemia in humans, with potential implications for determining a patient's treatment path (recently reviewed in Liu *et al.*[29]). Other potential prognostic (or diagnostic) factors which may have been missed using

298 standard differential expression analyses may be present in this category and warrant further
299 investigation.
300
301 The immune response categories we describe here are based on spleen cells stimulated with
302 anti-CD28 and anti-CD3 antibodies and sampled at 24 hours. However, the relative
303 frequency of the response categories reported here may vary depending on the choice of
304 agonist and/or time point. For example, markers are known to follow different response
305 trajectories, with some immediately responding and reaching peak activation, and others
306 taking longer to reach this point[30]. Sampling at a later time point, then, when the 'slower'
307 markers have reached peak activation, may lead to more convergence than reported here. In
308 order to fully account for this temporal variation, multiple time points need to be averaged
309 across. We argue that both time-specific and averaged responses are of functional
310 significance, but hope others will extend our work. We use RNASeq here in order to give a
311 broad view of the immune response. Single-cell RNASeq could be used to quantify
312 differences in individual response resulting solely from differences in cell-specific activity.
313 Previous work has shown that transcript levels generally correlate with protein levels across
314 genes[31]. However, more work is needed to confirm these response categories at the functional
315 level[32]. In future, Q-PCR or protein-level data could be used in order to include weakly
316 expressed markers, which were excluded here as a result of the heteroscedasticity inherent in
317 RNASeq data.
318
319 Markers that responded to stimulation were not limited to immune pathways as
320 conventionally defined. They included, for example, markers involved in the insulin
321 signalling pathway. This is in line with previous studies, which suggest that insulin plays a
322 key role in coordinating an organism's response to infection, influencing, in particular, the

allocation of resources[33,34]. One of these markers, *Pdk1*, was also among the top convergent markers. This could be representative of the high levels of variability in the (baseline) nutritional status of individuals in a natural population, coupled with an upper limit on the processes involved in glucose metabolism.

The immune categories we presented here, therefore, highlight markers not traditionally associated with immune functions, and offer a promising avenue for identifying potential prognostic (or diagnostic) factors for disease, like *Zap70*. They also point to both genetics and prior experience as drivers of natural variation in immune response. Our future work will further decompose this natural variation into that driven by these two components.

Methods

Field methods

Sixty-two field voles were collected between July and October 2015 to assay expression by RNASeq. These voles came from four sites in Kielder Forest, Northumberland (55°130N, 2°330W). Each site contained a trapping grid of regularly spaced traps (at approx. 5 m intervals) and was also used for other components of a larger field study (for more details see Wanelik *et al.*[27]).

Ethics statement

All animal procedures, carried out as part of this field study, were performed with approval from the University of Liverpool Welfare Committee and under the authority of the UK Home Office (Animals (Scientific Procedures) Act 1986) project licence number PPL

347 70/8210 to SP. Voles were killed by a rising concentration of CO_2 followed by
348 exsanguination.
349
350 **Cell culture methods**
351 Splenocyte cultures from each vole were split into two populations, one of which was
352 stimulated with anti-CD3 antibodies (Hamster Anti-Mouse CD3e, Clone 500A2 from BD
353 Pharmingen) and anti-CD28 antibodies (Hamster Anti-Mouse CD28, Clone 37.51 from BD
354 Tombo Biosciences) at concentrations of 2 μg/ml and of 1 μg/ml respectively for 24hr, and
355 the other was left as an unstimulated control to act as a reference level. We refer to this
356 reference level as the baseline, and control samples as baseline samples. However, it is
357 important to note that this level will vary for an individual, not only on a day to day basis, but
358 throughout its life. Culture conditions were otherwise equivalent to those used in Jackson *et*
359 *al*. (2011)[35]. Costimulation with anti-CD3 and anti-CD28 antibodies was used to selectively
360 promote the proliferation of T cells[14], our assumption being that this would reflect the
361 potential response of T-cell populations *in vivo*. Cell populations within splenocyte cultures
362 were variable but left undefined.
363
364 **RNASeq preparation and mapping**
365 RNA was extracted using Invitrogen PureLink kits. Following extraction, cDNA libraries
366 were prepared using Illumina RiboZero kits and libraries were constructed with NEBNext
367 Ultra directional RNA library prep kit according to the manufacturers protocols. Samples
368 were sequenced to produce 2 x 75 bp paired-end reads on an Illumina HiSeq4000 platform.
369 Adaptor sequences were removed with CUTADAPT version 1.2. and further trimmed with
370 SICKLE version 1.200 (minimum window quality score of 20). This resulted in a mean
371 library size of 18 million (range = 5–50 million) paired-end reads.

High-quality reads were mapped against a draft genome for *M. agrestis* (GenBank Accession no. LIQJ00000000) using TOPHAT version 2.1.0, and a set of predicted gene models was generated using BRAKER. Mapped reads were counted using FEATURECOUNTS. Further analysis was performed on counts of mapped reads for each gene in R version 3.4.2[36]. These count data were initially filtered to remove unexpressed genes (those genes with fewer than three counts per million across all samples; $n = 13$). Following filtering, library sizes were recalculated and data were normalised to represent counts per million (cpm). These data were found to be correlated with quantitative PCR (Q-PCR) data (see Supplementary Fig. 1). No correction for gene length was necessary as all analyses were based on comparisons across (rather than within) samples. Transcript abundance for a particular gene here represents a single, functional measure of its activity across some, undefined, cell population. In order to maximise the power of our analysis to identify biologically relevant patterns, we focussed on those genes which were expressed at an informative level in the spleen prior to and/or following stimulation ($n = 1350$ or 6%). Genes expressed at a mean level greater than 200 cpm were considered informative. As weakly expressed genes were removed (minimising heteroscedasticity), log-transformation of data was unnecessary (Supplementary Fig. 2).

Statistical analysis

Genes for which a response to stimulation was observed across individuals were identified, and, as elaborated in the Results, categorised on the basis of (i) the dependence of an individual's response on its baseline level, and (ii) the degree of inter-individual variability in response across individuals.

Baseline-dependence of response. The dependence of an individual's response on its baseline level was quantified by testing the relationship between that individual's baseline abundance (cpm_{base}) and its stimulated abundance (cpm_{stim}) using a linear regression, taking the form

$$cpm_{stim} \sim cpm_{base}$$

as well as a quadratic regression, taking the form

$$cpm_{stim} \sim cpm_{base} + cpm_{base}^{2}$$

For approximately one third of genes (n = 466), the residuals from both of these regressions deviated significantly from the assumptions of normality and/or homoscedasticity, and a non-parametric Kendall–Theil linear regression was fitted instead. Regression fits varied from gene to gene (R^2 ranging from <0.001 to 0.85).

Inter-individual variability in response. Inter-individual variability in response was quantified by comparing the coefficient of variation (CV) for baseline abundances across individuals (CV_{base}) and the CV for stimulated abundances across individuals (CV_{stim}). As response is defined as the difference between baseline and stimulated abundance, a large difference in their CVs, either

$$CV_{base} > CV_{stim}$$

or

$$CV_{stim} > CV_{base}$$

419

420 indicates a high level of variability in response. A relationship between gene-wise mean
421 expression levels and CV is typically found in RNASeq data, with low mean transcript
422 abundance being strongly associated with high variability[37]. As we restricted our analysis to
423 informative genes only, excluding those genes with low mean abundance, it was not
424 necessary to account for this relationship (Supplementary Fig. 2). Asymptotic tests for the
425 equality of CVs were run using the cvequality package. All *p*-values were corrected for
426 multiple testing using the Benjamini-Hochberg method[38].

427

428 **Functional annotation.** Functional enrichment analyses were run using The Database for
429 Annotation, Visualization and Integrated Discovery (DAVID) version 6.8[39,40]. Benjamini-
430 Hochberg corrected *p*-values and gene counts are reported alongside ontology terms,
431 including Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways to indicate their
432 level of enrichment[41–43].

433

434 **Age-specific analysis.** In order to begin to investigate the relative importance of genetic
435 variation versus prior stimulation for shaping patterns of variation in immune response, the
436 same analysis was performed separately on juvenile and mature voles. As we had more
437 samples from mature voles (n = 43) than juvenile voles (n = 19), we randomly sampled the
438 mature population (with replacement) 1000 times and averaged across these samples. The
439 number (juveniles) or mean number (matures) of genes in each of these age classes is
440 presented in the text.

441

442

443 **References**

444

445 1. Sarre, C. *et al*. Comparative immune responses against *Psoroptes ovis* in two cattle
446 breeds with different susceptibility to mange. *Vet. Res*. **46** (2015).
447 2. Mahler, M., Janke, C., Wagner, S. & Hedrich, H. J. Differential susceptibility of
448 inbred mouse strains to *Helicobacter pylori* infection. *Scand. J. Gastroenterol.* **37,**
449 267–278 (2002).
450 3. Eugui, E. M. & Allison, A. C. Malaria infections in different strains of mice and their
451 correlation with natural killer activity. *Bull. World Health Organ.* **57,** 231–238 (1979).
452 4. Brodin, P. *et al.* Variation in the human immune system is largely driven by non-
453 heritable influences. *Cell* **160,** 37–47 (2015).
454 5. Sellers, R. S., Clifford, C. B., Treuting, P. M. & Brayton, C. Immunological variation
455 between inbred laboratory mouse strains: points to consider in phenotyping genetically
456 immunomodified mice. *Vet. Pathol.* **49,** 32–43 (2012).
457 6. Lochmiller, R. L., Vestey, M. R. & Nash, D. Gut associated lymphoid tissue in the
458 cotton rat (*Sigmodon hispidus*) and its response to protein restriction. *J. Wildl. Dis.* **28,**
459 1–9 (1992).
460 7. Hooper, L. V, Littman, D. R. & Macpherson, A. J. Interactions between the microbiota
461 and the immune system. *Science* **336,** 1268–1273 (2012).
462 8. Abolins, S. R., Pocock, M. J., Hafalla, J. C., Riley, E. M. & Viney, M. E. Measures of
463 immune function of wild mice, *Mus musculus*. *Mol. Ecol.* **20,** 881–892 (2011).
464 9. Beura, L. K. *et al.* Normalizing the environment recapitulates adult human immune
465 traits in laboratory mice. *Nature* **532,** 512–516 (2016).
466 10. Burthe, S. *et al.* Cowpox virus infection in natural field vole *Microtus agrestis*
467 populations: delayed density dependence and individual risk. *J. Anim. Ecol.* **75,** 1416–

1425 (2006).

11. Smith, A., Telfer, S., Burthe, S., Bennett, M. & Begon, M. Trypanosomes, fleas and field voles: ecological dynamics of a host-vector--parasite interaction. *Parasitology* **131,** 355–365 (2005).

12. Telfer, S. *et al.* Species interactions in a parasite community drive infection risk in a wildlife population. *Science* **330,** 243–246 (2010).

13. Taylor, C. H. *et al.* Physiological, but not fitness, effects of two interacting haemoparasitic infections in a wild rodent. *Int. J. Parasitol.* **48,** 463–471 (2018).

14. Frauwirth, K. A. & Thompson, C. B. Lymphocyte signal transduction activation and inhibition of lymphocytes by costimulation. *Cancer Res.* **109,** 295–299 (2002).

15. Pedersen, A. B. & Babayan, S. A. Wild immunology. *Mol. Ecol.* **20,** 872–880 (2011).

16. Viney, M. & Riley, E. M. The immunology of wild rodents: current status and future prospects. *Front. Immunol.* **8,** 1–9 (2017).

17. Turner, A. K. & Paterson, S. Wild rodents as a model to discover genes and pathways underlying natural variation in infectious disease susceptibility. *Parasite Immunol.* **35,** 386–395 (2013).

18. Jackson, J. A. *et al.* Immunomodulatory parasites and toll-like receptor-mediated tumour necrosis factor alpha responsiveness in wild mammals. *BMC Biol.* **7,** 16 (2009).

19. Tsang, J., Schwartzberg, P., Kotliarov, Y. & Al., E. Global analyses of human immune variation reveal baseline predictors of post-vaccination respones. *Cell* **157,** 499–513 (2014).

20. IPC-CHI Signatures Project Team & HIPC-I Consortium. Multicohort analysis reveals baseline transcriptional predictors of influenza vaccination responses. *Sci. Immunol.* **2,** eaal4656 (2017).

493 21. Orrù, V. *et al.* Genetic variants regulating immune cell levels in health and disease.
494 *Cell* **155,** 242–256 (2013).
495 22. Telfer, S. *et al.* Parasite interactions in natural populations: insights from longitudinal
496 data. *Parasitology* **135,** 767–781 (2010).
497 23. Sheldon, B. C. & Verhulst, S. Ecological immunology: costly parasite defences and
498 trade-offs in evolutionary ecology. *Trends Ecol. Evol.* **11,** 317–321 (1996).
499 24. Lochmiller, R. L. & Deerenberg, C. Trade-offs in evolutionary immunology: just what
500 is the cost of immunity? *Oikos* **88,** 87–98 (2000).
501 25. Reiner, S. L. Development in motion: helper T cells at work. *Cell* **129,** 33–36 (2007).
502 26. Abbas, A. K., Murphy, K. M. & Sher, A. Functional diversity of helper T
503 lymphocytes. *Nature* **383,** 787–793 (1996).
504 27. Wanelik, K. M. *et al.* A candidate tolerance gene identified in a natural population of
505 field voles (*Microtus agrestis*). *Mol. Ecol.* **27,** 1044–1052 (2018).
506 28. Brown, M. *et al.* Seasonal immunoregulation in a naturally-occurring vertebrate. *BMC*
507 *Genomics* **17** (2016).
508 29. Liu, Y., Wang, Y., Yang, J., Bi, Y. & Wang, H. ZAP-70 in chronic lymphocytic
509 leukemia: a meta-analysis. *Clin. Chim. Acta* **483,** 82–88 (2018).
510 30. Tate, A. T. & Graham, A. L. Dissecting the contributions of time and microbe density
511 to variation in immune gene expression. *Proc. R. Soc. B Biol. Sci.* **284** (2017).
512 31. Tian, Q. *et al.* Integrated genomic and proteomic analyses of gene expression in
513 mammalian cells. *Mol. Cell. Proteomics* **3,** 960–969 (2004).
514 32. Viney, M. E., Riley, E. M. & Buchanan, K. L. Optimal immune responses:
515 immunocompetence revisited. *Trends Ecol. Evol.* **20,** 665–669 (2005).
516 33. Garcia, N. W. *et al.* Exogenous insulin enhances humoural immune responses in short-
517 day, but not longday, Siberian hamsters (*Phodopus sungorus*). *Proc. R. Soc. B Biol.*

Sci. **277,** 2211–2218 (2010).

34. Lee, S. *et al.* Modulatory upregulation of an insulin peptide gene by different pathogens in *C. elegans*. *Virulence* **9,** 648–658 (2018).

35. Jackson, J. A. *et al.* The analysis of immunological profiles in wild animals: a case study on immunodynamics in the field vole, *Microtus agrestis*. *Mol. Ecol.* **20,** 893–909 (2011).

36. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. (2016).

37. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15,** 1–17 (2014).

38. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57,** 289–300 (1995).

39. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37,** 1–13 (2009).

40. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4,** 44–57 (2009).

41. Kanehisa, M. & Goto, S. KEGG; Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28,** 27–30 (2000).

42. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44,** D457–D462 (2016).

43. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45,** D353–

D361 (2017).

Acknowledgements

The authors wish to thank those involved in obtaining and processing samples from the field: Rebecca Turner, Lukasz Lukomski, Stephen Price, William Foster, Ann Lowe and Anna Thomason. They also wish to thank the Forestry Commission for access to the study sites and the Centre for Genomic Research at the University of Liverpool for sequencing samples.

Author contributions

M.B., J.E.B., J.A.J. and S.P. designed the study. E.A. undertook the stimulatory assays. K.M.W. analysed the data. All authors wrote the manuscript.

Competing interest statement

The authors declare no competing financial interests.

Figure captions

Fig. 1 Different categories of immune response. These are based on two overlapping sets of criteria, baseline-dependence of response (blue) and inter-individual variability in response (yellow background). Arrows represent individual immune responses. No response (for reference): markers for which individuals (on average) show no response to stimulation (intercept not significantly different from zero; slope not significantly different from one). Uncorrelated response: markers for which responses of different individuals are variable and independent of their baseline level (slope not significantly different from zero). Constant

response: markers for which the responses of different individuals response are (approximately) constant and independent of their baseline (intercept significantly greater than zero; slope not significantly different from one). Baseline-dependent response: markers for which responses of different individuals vary as a function of their baseline level, either as a linear function of their baseline (slope significantly different from one; slope greater than one is depicted but could equally be less than one), or as a quadratic function of their baseline (a saturating function is depicted but could equally be exponential). Convergent response: markers for which the coefficient of variation (CV) for baseline abundances is significantly greater than the CV for stimulated abundances across individuals ($CV_{base} > CV_{stim}$). Divergent response: markers for which CV for stimulated abundances is significantly greater than CV for baseline abundances across individuals ($CV_{stim} > CV_{base}$). Both convergent and divergent markers depicted as, but not limited to, markers for which response is uncorrelated.

Fig. 2 Top 10 markers and enriched ontology terms in each immune response category. Each box represents a category of immune response (as in Fig. 1). For each category, top 10 annotated markers for which we had the most confidence in their categorisation (markers were ranked on R^2 and *p*-values) are listed, one or two of these are represented in plots showing stimulated versus baseline abundances across individuals (solid line indicates significant relationship between baseline and stimulated abundance; dashed line indicates slope equal to one for reference). In the case of the convergent category, which only included a total of six annotated markers, all markers are listed. Ontology terms of interest, from an enrichment analysis preformed on all markers within a category (where possible), are also included (immune-related terms in black).

Fig. 3 Map of the T-cell receptor signalling KEGG pathway for *Mus musculus*, with the colour of nodes representing level of inter-individual variability in response to stimulation with anti-CD3 and anti-CD28 antibodies in *Microtus agrestis*. Namely the *p*-value from an asymptotic test for the equality of variance in gene expression levels for baseline and stimulated samples (range = < 0.001–0.97). Dark blue indicates high inter-individual variability in response, whereas light blue or white indicates low inter-individual variability in response. Grey nodes represent genes for which no information is available, either because they are unannotated in the *M. agrestis* genome, or because they are weakly expressed in the spleen.

Tables

Table 1 Table summarising the number of markers found in each of the three main categories of immune response. For each of these categories, the number of convergent and divergent markers is shown.

Immune response category	Total no. markers	No. convergent	No. divergent
Uncorrelated	47	2	1
Constant	306	0	91
Baseline-dependent	699	5	47

response
increase in level
decrease in level

No response

$y = mx + c$

m = 1

c = 0

stimulated level

baseline level

Baseline-dependence

Uncorrelated response

m = 0 (not shown)

stimulated level

baseline level

Constant response

m = 1

c > 0

stimulated level

baseline level

Baseline-dependent response

Linear m ≠ 1

stimulated level

Quadratic

$y = ax^2 + bx + c$

stimulated level

baseline level

Variability

Convergent response

$CV_{base} > CV_{stim}$

CV

CV

stimulated level

Divergent response

$CV_{stim} > CV_{base}$

CV

CV

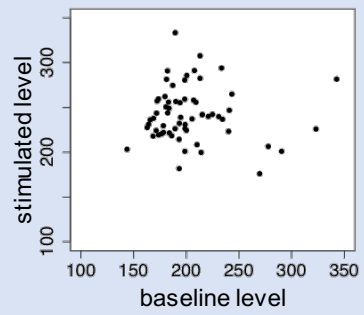
stimulated level

baseline level

Baseline-dependence

Variability

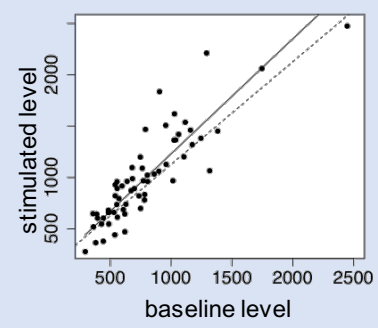
Uncorrelated response



Mon2 *Glyr1* *Setx* *Baz2b* *Clasp1*
Smarcc2 *Rbm5* *Ap3d1* *Zfp280d* *Zfp445*

mRNA splicing (4)

Constant response

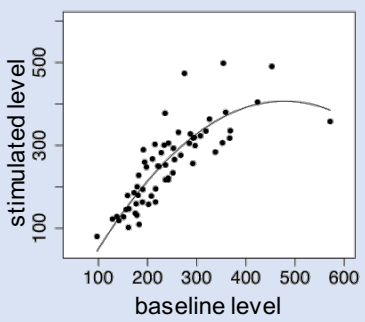


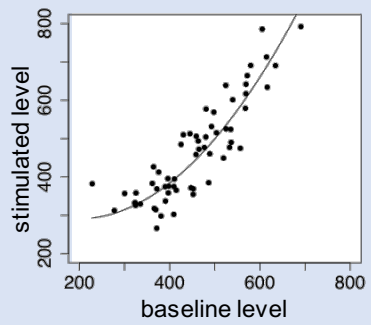
Lyst *Mgat4a* *Rgs9* *Fyn* *Ddr1*
Rapgef1 *Swap70* *Ckap5* *Sik3* *Hfe*

TCR pathway (9)
Natural killer cell mediated cytotoxicity (8)
Fc epsilon RI signaling pathway (7)
Fc gamma R-mediated phagocytosis (7)
Insulin signaling pathway (9)
thyroid hormone signaling pathway (8)

Baseline-dependent response

Quadratic

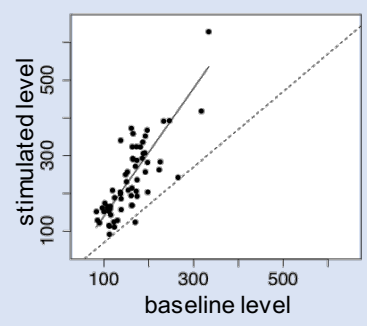


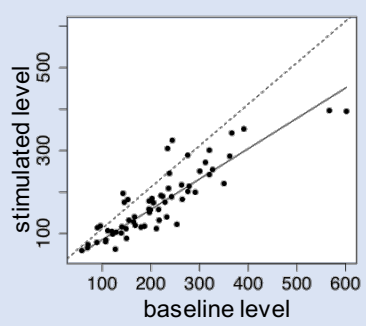


Cd247 *Snrk* *Cpd* *Itga6* *Cd3d*
Fam135a *Card11* *Traf1* *Sik2* *Ralgapa2*

TCR pathway (7)
Biological rhythms (6)
RNA-mediated gene silencing (4)

Linear



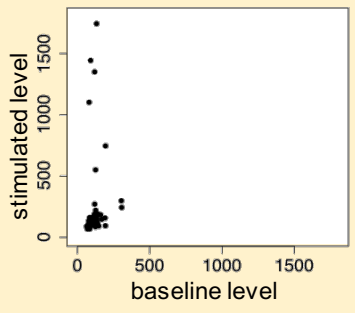


Jchain *Klri1* *Ikzf2* *Gbp2/2b* *Man2a2*
Cybb *Ncf2* *Cd8a* *Tcf4* *Txndc5*

Immunity (20)
Leukocyte transendothelial migration (14)
T Cytotoxic Cell Surface Molecules (7)
HTLV-I infection (20)
Leukocyte cell-cell adhesion (7)
Immunological synapse (6)
Antigen processing and presentation of endogenous peptide antigen via MHC class I (4)

Highly variable response

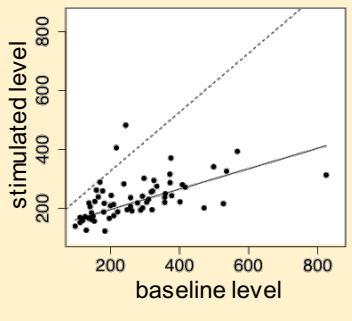
Divergent



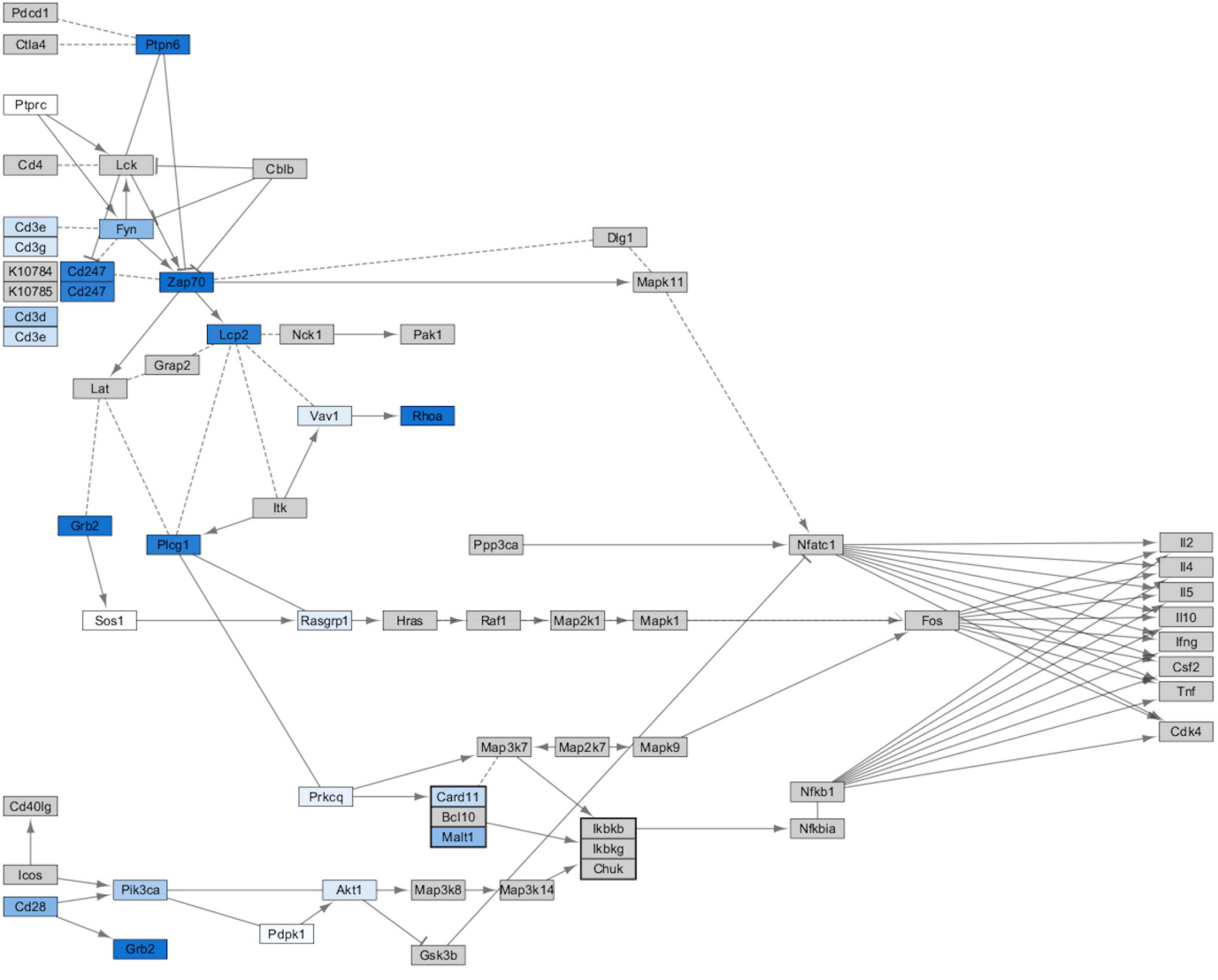
Mdn1 *Xpo1* *Akap13* *Serbp1* *Ptbp1*
Jun *Ddx21* *Usp28* *Hsph1* *Sik2*

Immunological synapse (5)
Viral carcinogenesis (12)
Epstein-Barr virus infection (11)
Heat shock protein 70 family (4)
Transcription factor Jun (3)
Toxin transport (6)
Positive regulation of telomere maintenance via telomerase (5)

Convergent



Pdk1 *Ero1l* *Pja2* *Lonp1* *Mon2* *Tcirg1*


Pdcd1
Ctla4
Ptpn6
Ptprc
Cd4
Lck
Cblb
Cd3e
Cd3g
Fyn
K10784
K10785
Cd247
Cd247
Zap70
Dlg1
Mapk11
Cd3d
Cd3e
Lcp2
Nck1
Pak1
Grap2
Lat
Vav1
Rhoa
Itk
Grb2
Plcg1
Ppp3ca
Nfatc1
Il2
Il4
Il5
Il10
Ifng
Csf2
Tnf
Cdk4
Sos1
Rasgrp1
Hras
Raf1
Map2k1
Mapk1
Fos
Map3k7
Map2k7
Mapk9
Prkcq
Card11
Bcl10
Malt1
Nfkb1
Ikbkb
Ikbkg
Chuk
Nfkbia
Cd40lg
Icos
Cd28
Pik3ca
Akt1
Map3k8
Map3k14
Pdpk1
Grb2
Gsk3b