Integrative biology defines novel biomarkers of resistance to strongylid infection in horses

Guillaume Salle, Cécile Canlet, Jacques Cortet, Christine Koch, Joshua Malsa, Fabrice Reigner, Mickaël Riou, Noémie Perrot, Alexandra Blanchard, Núria Mach

HAL Id: hal-03284457
https://hal.inrae.fr/hal-03284457v1

Submitted on 12 Jul 2021

scientific reports

Check for updates

OPEN

Integrative biology defines novel biomarkers of resistance to strongylid infection in horses

Guillaume Sallé[1]✉, Cécile Canlet[2], Jacques Cortet[1], Christine Koch[1], Joshua Malsa[1], Fabrice Reigner[3], Mickaël Riou[4], Noémie Perrot[4], Alexandra Blanchard[5] & Núria Mach[6]

The widespread failure of anthelmintic drugs against nematodes of veterinary interest requires novel control strategies. Selective treatment of the most susceptible individuals could reduce drug selection pressure but requires appropriate biomarkers of the intrinsic susceptibility potential. To date, this has been missing in livestock species. Here, we selected Welsh ponies with divergent intrinsic susceptibility (measured by their egg excretion levels) to cyathostomin infection and found that their divergence was sustained across a 10-year time window. Using this unique set of individuals, we monitored variations in their blood cell populations, plasma metabolites and faecal microbiota over a grazing season to isolate core differences between their respective responses under worm-free or natural infection conditions. Our analyses identified the concomitant rise in plasma phenylalanine level and faecal *Prevotella* abundance and the reduction in circulating monocyte counts as biomarkers of the need for drug treatment (egg excretion above 200 eggs/g). This biological signal was replicated in other independent populations. We also unravelled an immunometabolic network encompassing plasma beta-hydroxybutyrate level, short-chain fatty acid producing bacteria and circulating neutrophils that forms the discriminant baseline between susceptible and resistant individuals. Altogether our observations open new perspectives on the susceptibility of equids to strongylid infection and leave scope for both new biomarkers of infection and nutritional intervention.

Infection by gastro-intestinal nematodes is a major burden for human development worldwide as they both affect human health[1] and impede livestock production[2]. Worldwide reports of anthelmintic drug failures against nematodes of veterinary importance have accumulated[3], threatening the sustainability of livestock farming in some areas. The same pattern applies in horses whereby widespread benzimidazole and pyrantel failures against cyathostomin populations have been reported[4–8]. In addition, reduced efficacy of macrocyclic lactones has been observed in the USA[6], Brazil and western Europe[9].

These small strongyles are mostly responsible for growth retardation in young animals, diarrhoea being frequently reported among clinical symptoms, along with colic[10,11]. Hosts become infected through ingestion of infective third stage larvae on pasture. These L3 larvae then migrate to their host caecum where they enter the mucosa and submucosa to complete molting into the fourth larval stage (L4), after two distinct stages (early and late L3)[12]. Duration of this process can vary in time, up to two years, as a result of an optional hypobiosis at the early L3 stage[13]. Following molting, they will resume their development to the adult reproducing stages in the lumen of the caecum and colon. The massive emergence of developing larval stages from the caeco-colic mucosa can cause a larval cyathostominosis syndrome[14]. This seasonal syndrome is characterized by the sudden onset of chronic diarrhoea associated with oedema and weight loss[14], and remains a leading cause of parasite-mediated death[15].

Factors contributing most to the selection of drug-resistant cyathostomin populations in equids remain uncertain[6,7]. However, significant and heritable between-horse variation in strongylid egg excretion has been reported in both domestic[16,17] and wild horse populations[18]. These observations hence suggest that some horses are able to limit parasite egg production, the mechanisms of which remain to be uncovered but likely apply to two broad categories, i.e. reduction of overall worm burden or reduction of worm fertility as found for ruminant

[1]INRAE, Université de Tours, UMR 1282 Infectiologie et Santé Publique, 37380 Nouzilly, France. [2]INRAE, Université de Toulouse, ENVT, INP-Purpan, UPS, UMR 1331 Toxalim, 31027 Toulouse, France. [3]INRAE, UE 1297 Physiologie Animale de l'Orfrasière, 37380 Nouzilly, France. [4]INRAE, UE 1277 Plateforme d'Infectiologie Expérimentale, 37380 Nouzilly, France. [5]Pancosma, Rolle, Switzerland. [6]INRAE, Université Paris-Saclay, AgroParisTech, Génétique Animale et Biologie Intégrative, 78350 Jouy-en-Josas, France. ✉email: Guillaume.Salle@inrae.fr

trichostrongylids[19]. While Faecal Egg Count (FEC) is a poor correlate of actual worm burden in horses[20], it can serve as a marker trait for horse resistance quantification as reported elsewhere[21]. As such, horse resistance to strongylid infection is defined as the ability to control strongyle life-cycle. This expands the strict concept of host resistance—limiting parasite worm burden[22]—by accounting for epidemiological consequences of reduced pasture contamination with free-living stages. It is also the only way resistance could be tracked as measuring strongyle worm burden is hampered by ethical and material limitations associated with the culling of large numbers of horses otherwise. Nonetheless, heritable variation in FEC leaves scope for restricting drug application to the most susceptible horses, thereby alleviating the selection pressure on parasite populations. To date, the genetic architecture of this trait has not been defined in equids, although indications from ruminant species would be in favour of a polygenic architecture[23–25] defining a stronger type 2 cytokinic polarization in resistant individuals[26,27].

Identifying biomarkers of this intrinsic resistance would both contribute to understanding the host-parasite relationship and to defining relevant biomarkers for use in the field. Current targeted-selective treatment schemes are based on FEC that has suboptimal sensitivity and remains time-consuming despite recent advances that should ease egg detection[28]. As a result, its uptake in the field varies widely across countries and remains limited[7,29] despite being cost-effective[30,31]. To date, limited alternative biomarkers have been identified. Alteration in serum albumin level and decrease in circulating fructosamine were the main features found in cyathostomin infected ponies[32]. Independent observations concluded that mixed strongyle infection was associated with mild inflammatory perturbations[33]. We previously highlighted that susceptible ponies had lower monocyte but higher lymphocyte counts than resistant individuals upon natural strongylid infection[34]. More susceptible individuals also exhibited differential modulation of their faecal microbiota, including enrichment for the *Ruminococcus* genus[34], corroborating independent observations of alterations in the gut microbiota composition of infected horses[34–37].

In horses as in other host-parasite systems, limited efforts have been made to isolate compositional shifts in plasma metabolites following parasite nematode infection. Beyond murine models of helminth infection[38,39], implementation of this technology could define a urinary biomarker of infection by *Onchocerca volvulus* in humans[40]. This was however not reproduced in other cohorts of patients[41]. In livestock species, a single study has applied metabolomic profiling on horse faecal matter to identify biomarkers of infection by parasitic nematodes but found little differences between horses with contrasting levels of strongylid infection[35].

In any case, these observations remain limited to individual host compartments and do not provide an integrated perspective of the physiological underpinnings associated with susceptibility to infection. Because we are aiming to distinguish between individuals before the onset of clinical signs, the biological signals may be subtle. To this respect, integration of systems biology data—that consider multiple high-dimensional measures from various host compartments—is expected to better identify the multiple features defining a given physiological state[42].

Under this assumption, we combined and analysed metabolomic, metagenomic and clinical data collected on a selected set of intrinsically resistant and susceptible ponies to identify the physiological components underpinning their resistance potential. Our data defined a strongylid infection signature built around lower circulating monocytes, enriched plasma phenylalanine concentration and higher *Prevotella* load in faecal microbiota that we could replicate in independent cohorts. We also identified an immunometabolic signature centered on neutrophils that best discriminated between resistant and susceptible individuals across strongyle-free or natural infection conditions. These results begin to define the physiological bases supporting the intrinsic resistance potential to cyathostomin infection in equids.

Results

This experiment was based on a set of individual Welsh ponies with divergent resistance to strongyle infection, measured as their ability to limit eggs excretion. During the experiment (2015), we produced metabolomic data that we present herein and integrated this plasma-related dataset with previously described faecal bacteria profiles and clinical parameters[34]. We analysed the data produced for each group under worm-free conditions (day 0) or following natural infection (day 132) to (1) identify biomarkers of infection, (2) establish a holistic view of the physiological underpinnings of the resistance potential to strongyle infection in horses.

Pony divergence toward strongyle infection is significant and sustained.

We selected 20 female Welsh ponies with divergent susceptibility to strongylid infection. Their susceptibility potential was predicted from their past FEC history (at least three FEC records over two years between 2010 and 2015). During the experiment (2015) and following natural infection (day 132), 17 ponies displayed FEC values in good agreement with their predicted potential. In that case, 8 of the susceptible ponies (**TS**) had FEC above the considered 200 eggs/g cut-off for treatment (average FEC = 419 ± 149 eggs/g) at day 132, and 9 resistant ponies (**TR**) were below this threshold (average FEC = 56 ± 77 eggs/g). Our prediction hence achieved an accuracy of 85%. Other ponies that did not match expectations had either higher susceptibility (850 eggs/g for the predicted resistant individual) or too low FEC (0 and 50 eggs/g for the two susceptible individuals).

In agreement with this observation, FEC measured in the TS and TR ponies during the five years preceding the experiment departed significantly from the herd mean (0.6 standard deviations ; $P = 0.01$ and $P = 0.02$ for TS and TR groups respectively; Fig. 1). To ensure that their intrinsic potential was true, we compared their FEC records collected after the experiment took place (between 2015 and 2020) to that of their herd (1436 individual records). We confirmed that their divergence was sustained (− 0.76 and + 0.63 standard deviation from mean in TS and TR ponies respectively; $P = 0.02$ in both cases) throughout the following years (Fig. 1), hence validating their intrinsic potential. This corresponded to an average FEC of 43 eggs/g (min = 0; max = 2200 eggs/g)
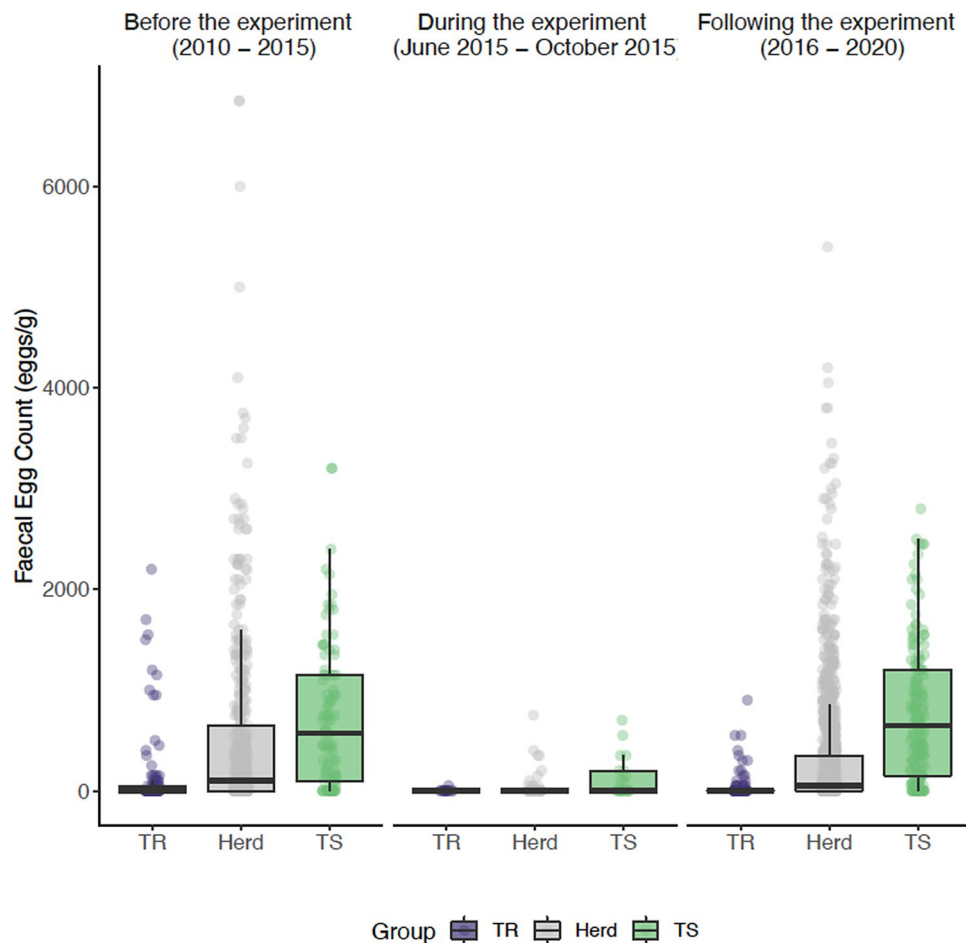
Figure 1. FEC-informed pony potential prediction is robust through time. The distribution of observed faecal egg counts before the experiment (2010–2015), during the experiment (2015) and after the experiment (2015–2020) in the predicted susceptible (8 individuals, 271 records; green) and resistant (9 individuals, 223 records; purple) ponies and their herd unselected counterparts (127 individuals, 1436 measures; grey) is represented. Dots stand for individual measures and boxplots represent the data distribution (mean materialized by a vertical bar within the box that stands for the 25–75% interquartile range).

and 756 eggs/g (min = 0; max = 2800 eggs/g) in TR and TS ponies respectively, while the average herd FEC was 320 eggs/g (min = 0; max = 5400 eggs/g).

Metabolomic profiling highlights the association between plasma phenylalanine level and FEC.   To identify markers of pony intrinsic resistance potential, we measured variation in their plasma metabolites using ^{1}H-NMR between worm-free conditions (day 0) and after natural infection (day 132). This metabolomic profiling of TR and TS ponies throughout the grazing seasons found a total of 791 metabolic buckets, corresponding to 119 unique metabolite signals (Fig. 2, Supplementary Table 1). These included several amino acids, energy metabolism-related metabolites, saccharides, unknown lipids, and organic osmolytes in the plasma. Among these signals, we observed 29 unassigned bins in three main windows ranging between 1.115 and 1.435 ppm, 3.385 and 4.305 ppm or 6.805 and 7.895 ppm (Supplementary Table 1).

The ANOVA—simultaneous component analysis (ASCA) applied to metabolomic time series data identified significant temporal variation ($P < 0.001$) but no differential systematic difference occurred between TR and TS pony metabolomes ($P = 0.54$). The temporal variation was structured around five signals (Supplementary Fig. 1) associated with alkalene moieties from lipids (^{1}H-NMR signal at 5.265–5.355 ppm), sugar moieties of α- and β-glucose (3.455–3.555 ppm) in overlap with proline (3.395–3.445 ppm) and branched-chain amino-acids (BCAAs) such as valine (1.045–1.055 ppm) and leucine (0.965–0.975 ppm). These signals showed mean leverage of 3.4% (ranging between 1.5 and 7.8%) and squared prediction error below 2.2×10^{-5}. Among these, the high-intensity signals ascribable to glucose decreased between the strongyle-free (day 0) and strongyle-infected (day 132) conditions (Supplementary Fig. 1).

Despite the lack of systematic modifications of metabolomes between TR and TS ponies, we sought to identify individual metabolites that would reflect the intrinsic susceptibility status of ponies before any infection took place (day 0) or metabolites that would differentiate individuals in need of treatment at the end of the grazing
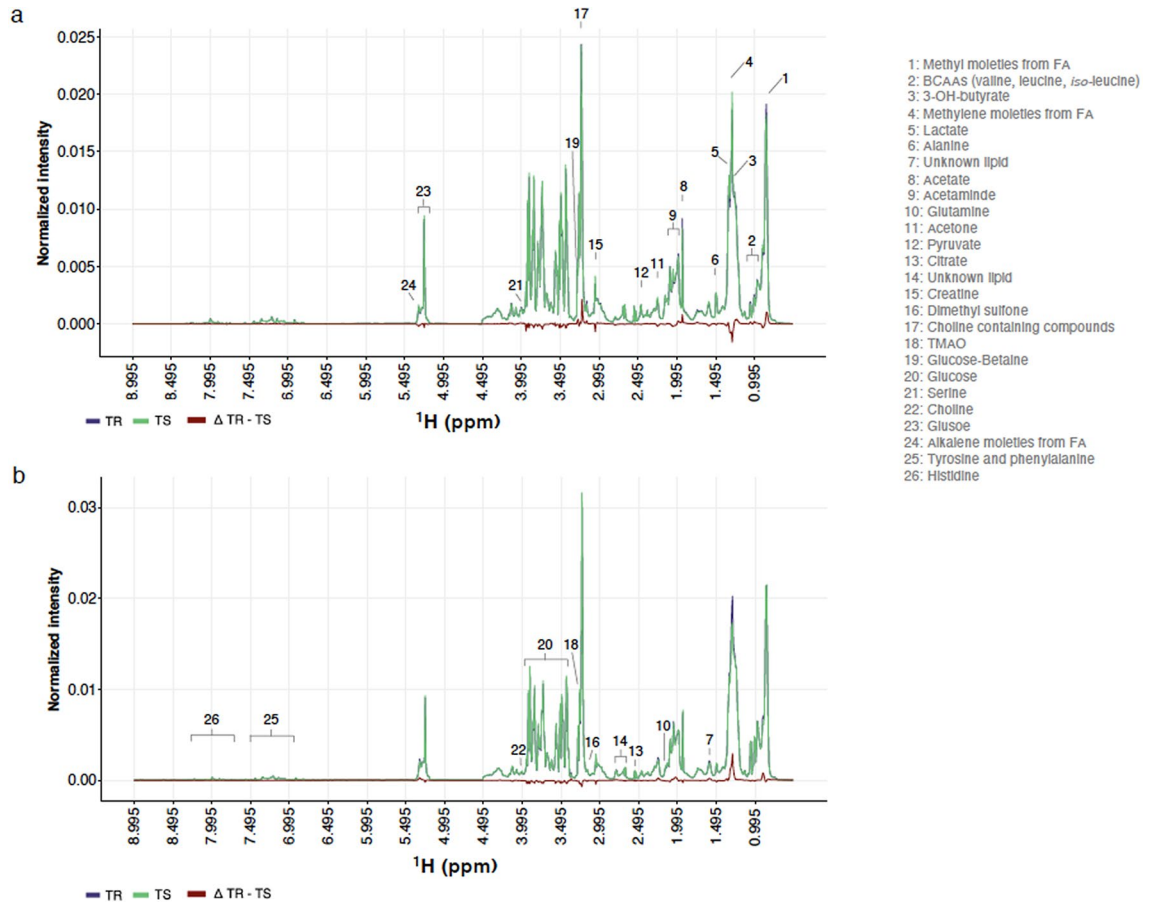
Figure 2. Representative ^{1}H-NMR spectra measured in resistant and susceptible ponies. Group average ^{1}H-NMR signal intensities are plotted against the considered chemical shifts ranging from 0.995 to 8.995 ppm and overlaid for resistant (TR; purple) and susceptible (TS; green) pony groups at day 0 (strongyle-free; panel **a**) or day 132 (strongyle infected; panel **b**). Differential intensity between groups is drawn in red. Associated metabolites are annotated by numbers.

season (day 132). Considering the nominal *P* values of 5%, we identified differences in dimethyl-sulfone and lysine associated signals that were all decreased in the TS pony group (Supplementary Fig. 2a). None of these signals were however significantly correlated with final FEC measured at day 132 (Spearman's ρ ranging between − 0.08 and 0.34, Supplementary Fig. 2b).

The same approach applied to TR and TS ponies after natural infection at day 132 found lower levels of phenylalanine (^{1}H-NMR signals at 3.285–3.305 ppm and 7.345–7.445 ppm) in TR compared to TS ponies (nominal P-value = 6 × 10^{-3}). In addition, an unidentified metabolite between 6.815 and 6.815 ppm (U13) was significantly lower in TR ($P = 0.04$). Phenylalanine signal intensities increased with FEC (Spearman's ρ ranging between 0.55 and 0.60, $P < 0.05$, n = 17). Using the faecal metabolomic data from another independent set of British horses[35], we could validate this signal. In that study and in line with our results, faecal phenylalanine level was significantly increased in horses with higher FEC (*Wilcoxon's test = 27, P value = 0.02*, Fig. 3C).

Overall these results indicate that strongylid infection experienced by TR and TS ponies did not induce metabolome-wide modifications. However, phenylalanine was the most discriminant between TR and TS ponies under infection and stands as a biomarker of the need for treatment.

Multi-compartment data integration identifies circulating monocytes, phenylalanine and faecal *Prevotella* levels as discriminant features under infection. To mine the physiological differences between infected TR and TS ponies more deeply, we applied a data integration framework (sGCC-DA) bringing together clinical, metabolomic and previously analyzed faecal microbiota data from these individuals[34]. The first component of the sGCC-DA better discriminated between TR and TS ponies and retained nine clinical parameters, 15 bacterial genera and ten plasma metabolite signals (Fig. 3, Supplementary Figs. 3 and 4).

The network of correlations between these features was structured around two major clusters (Fig. 4, Supplementary Figs. 3 and 4). A first core of features built around average daily weight gain (ADG) and four commensal gut bacterial genera (*Anaerovibrio, Fibrobacter, Prevotella,* and *Treponema*) displayed higher levels in TR ponies under infection (Fig. 4, Supplementary Figs. 3 and 4). These features were negatively correlated to a strongylid susceptibility-associated cluster that encompassed neutrophil counts (2.8 ± 0.51 and 3.05 ± 0.8 million cells/mm^3 in TR and TS ponies), the haematocrit (average of 38.74% ± 1.7 vs. 41.45% ± 4.1 in TR and TS ponies)
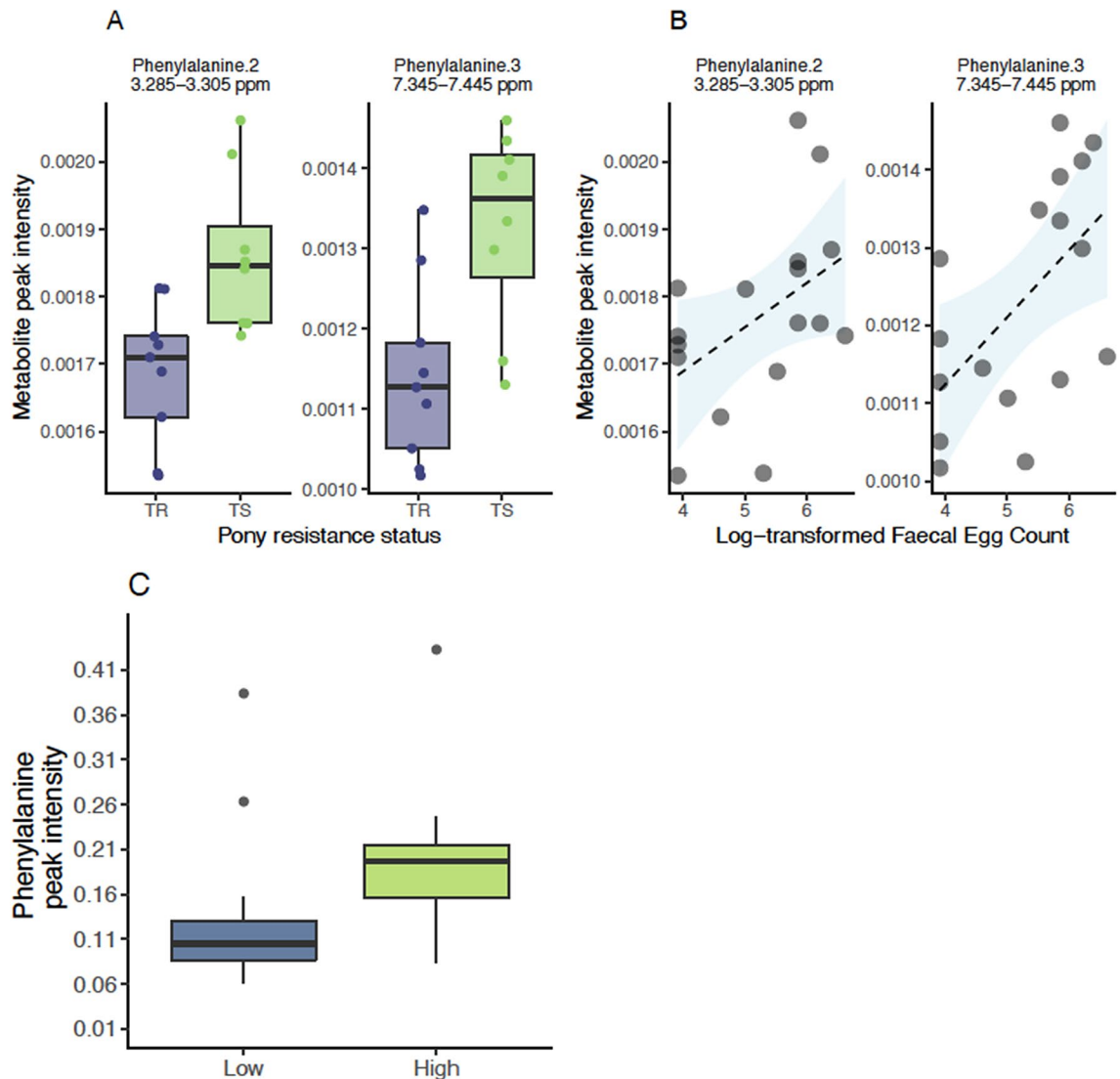
Figure 3. Differential metabolites between infected resistant and susceptible ponies (day 132). Panel (**A**) shows metabolite signal intensity distribution in each pony susceptibility group (purple: resistant, TR; green: susceptible, TS) at day 132. Panel (**B**) shows the relationship between these metabolite signal intensities (Y-axis) at day 0, and matching log-transformed faecal egg count (X-axis) at day 132. Panel (**C**) describes observed phenylalanine levels in the faecal matter of an independent cohort of British horses with low or high FEC.

and the plasma levels of serine and essential amino acids such as phenylalanine, lysine and tryptophan (Fig. 4, Supplementary Figs. 3 and 4). This underscores the association between phenylalanine and FEC (Fig. 3B, C).

Additional features included higher monocyte counts (1.178 ± 0.23 and 1.038 ± 0.24 million cells/mm^3 in TR and TS ponies on average) and higher levels of plasma alkaline phosphatase (ALP; 5.39 ± 0.15 and 5.25 ± 0.12 Units/L) in TR ponies (Fig. 4). However, these parameters displayed least covariation with other parameters (Fig. 4).

Analysis for KEGG pathway enrichment of metabolite signals found significant over-representation of Aminoacyl-tRNA biosynthesis (FDR = 1.1×10^{-4}) underpinned by the presence of alanine, lysine, phenylalanine, serine, and tryptophan (Supplementary Table 2). Of note, variations in alanine and phenylalanine plasma levels were similar to what was previously reported for dengue fever[43], another—yet viral—infectious process (Supplementary Table 2). The most significant enrichment was defined by the *Blautia*, *Coprococcus* and *Ruminococcus* association found in forms of pediatric Crohn's disease[44] (FDR = 2.5×10^{-3}). Two other significant enrichments included infection-mediated perturbations associated with the HIV-1 virus in humans (underpinned by *Anaerovibrio* and *Clostridium* genera, FDR = 1.12×10^{-4}) or with murine model of *Plasmodium* infection (underpinned by *Anaeroplasma* and *Clostridium*, FDR = 0.05).

To validate the association of the most discriminant blood and bacterial features with FEC, we used either a previously published horse data set[35] for bacterial count, or additional blood samples taken from 25 strongylid infected ponies in 2020 (Fig. 5). Out of the five most discriminating bacterial genera identified by our sGCC-DA approach under infection at day 132, *Prevotella* (FDR = 0.19; nominal *P value* = 0.04) also showed significant differences in their mean abundances between Peachey et al.'s horses with low or higher FEC (Fig. 5). Monocyte
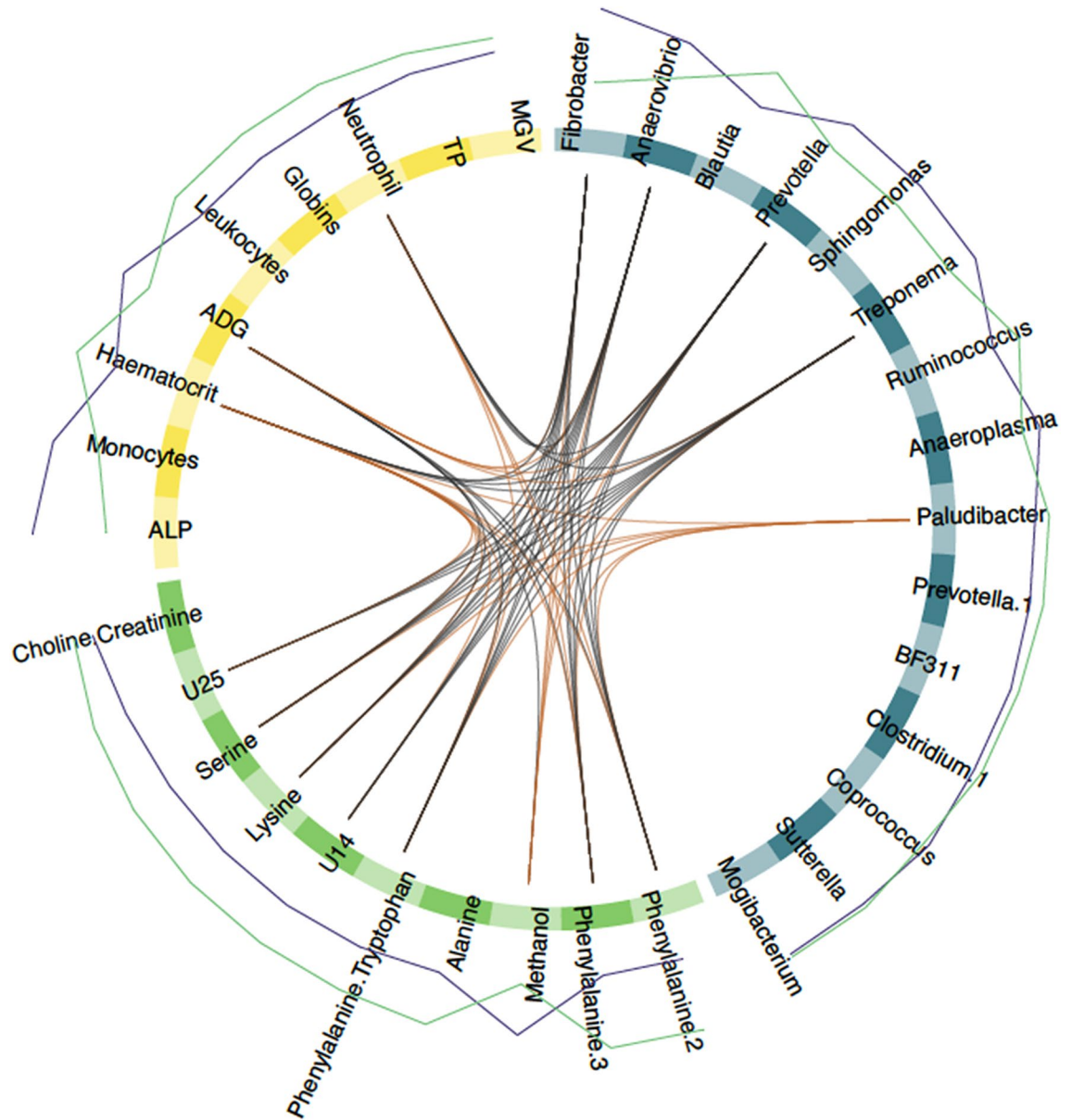
Figure 4. Circos plot showing features best discriminating between infected resistant and susceptible ponies (day 132). For each of the three input data types (clinical data in yellow, bacterial genera in purple and metabolite signals in orange), the features best discriminating between resistant and susceptible ponies are listed. A link is materialized between two features if their shared correlation is above 0.45 (chocolate if positive, grey otherwise). External green and purple lines represent the relative feature level in each pony susceptibility group (resistant, TR: purple; susceptible, TS: green). *MGV* mean globular volume, *TP* total protein, *ALP* alkaline phosphatase, *U14, U25* unknown metabolites 14 and 25.

counts were also negatively correlated with FEC levels in the independent set of ponies (Spearman's $\rho = -0.31$, *P value* = 0.14; Fig. 5).

The trend was however opposite for the neutrophil population (Spearman's $\rho = -0.63$, *P value* = 9×10^{-4}; Fig. 5) and no relationship between FEC and haematocrit was found in this independ set of individuals (Spearman's $\rho = -0.07$, *P-value* = 0.7; Fig. 5).

These findings hence retain increased *Prevotella* abundance in faecal material as a conserved signal in equids with reduced FEC, and monocyte counts appears to be a good predictor of FEC level.

Short-chain fatty acid producing bacteria, plasma lysine and circulating neutrophils recapitulate pony intrinsic potential across conditions. Using the same analytical framework, we aimed to identify features that would best define the intrinsic pony resistance potential across worm-free or natural infection conditions.
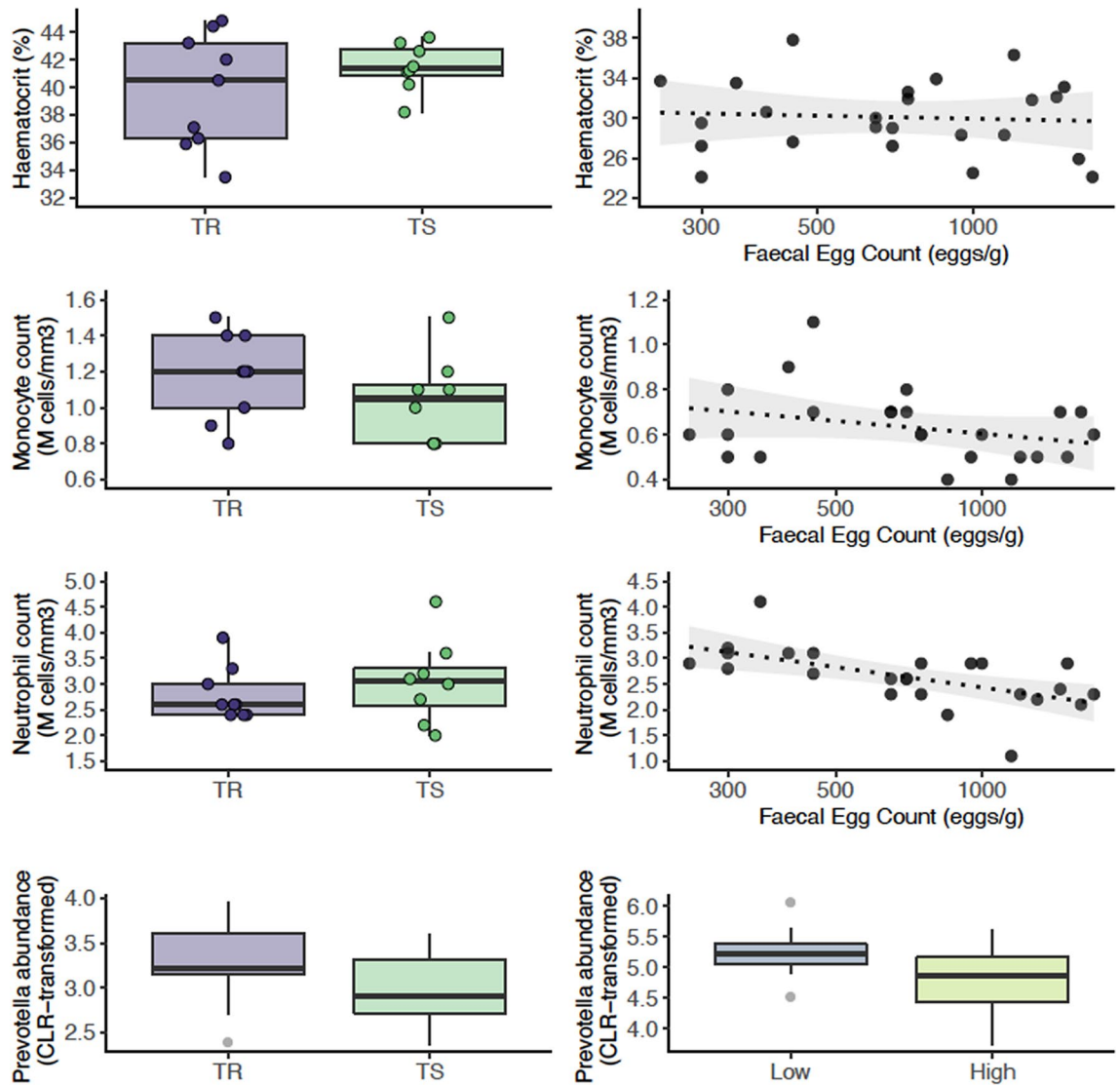
Figure 5. Validation of the most discriminant features between resistant and susceptible ponies under strongylid infection. Each row presents the feature level in resistant and susceptible ponies (left panels) and the association with faecal egg count recorded in an independent set of ponies or horses (for the *Prevotella* genus). *CLR* centered log-transformed. The figure highlights the significant positive relationship between circulating monocyte count and faecal abundance of the *Prevotella* genus in independent individuals.

The sGCC-DA applied on records measured under worm-free conditions (day 0) identified reduced circulating neutrophil and leukocyte counts in the TR ponies as the most discriminant clinical parameters (Supplementary Figs. 5, 6, and 7). Cell counts of these two populations were tightly linked with plasma levels of 1-methylhistidine (7.785 ppm; reduced in TR ponies), lysine (1.445–1.465 and 1.845–1.915 ppm; increased in TR ponies) and β-hydroxybutyrate (4.125–4.135 ppm; increased in TR ponies). In addition to these parameters, maximal covariance was obtained for a few genera from the Actinomycetia class (order Corynebacteriales), namely *Dietzia, Gordonia, Mycobacterium* and *Sacharopolyspora* (order Pseudonocardiales), that all showed higher relative abundances in resistant ponies, as well as the candidate genus BF311 within Bacteroidetes (i.e. *BF311*; Supplementary Fig. 5). On the opposite, higher relative abundance of butyrate-producing Clostridia, specifically *Anaerofustis, Coprococcus* and *Ruminococcus* were found in TS ponies (Supplementary Fig. 5). These genera showed positive correlations with circulating lymphocytes and neutrophil counts (ranging between 0.51 and 0.68 for lymphocyte counts and between 0.38 and 0.51 for neutrophil counts respectively; Supplementary Figs. 5 and 7). Of note, the presence of *Prevotella* and *Desulfovibrio* among the set of covarying features defined significant enrichments compatible with *Plasmodium* infection in mice or *Schistosoma haematobium* infection in humans (FDR = 0.04 in both cases; Supplementary Table 2).

Combining these differential features between pony groups under worm-free conditions with that found after strongylid infection retained a core set of seven markers consistently discriminating TR and TS ponies across infection conditions (Supplementary Fig. 8). For these markers, we aimed to identify differential trends between both groups across infection conditions.
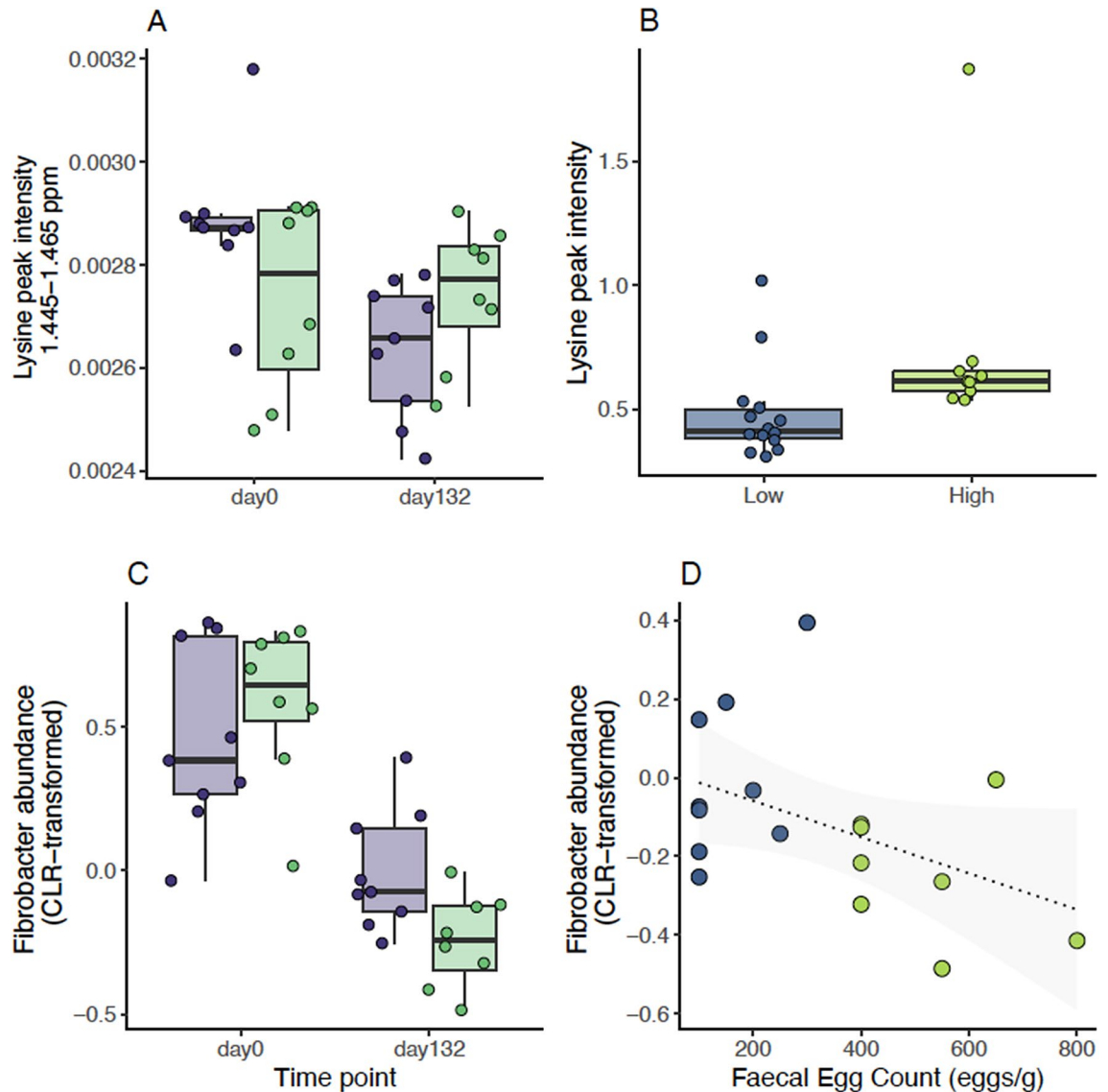
Figure 6. Discriminant features between resistant and susceptible ponies across infection conditions. Measured values of lysine signal intensity (**A**) and fibrobacter abundance (**C**) are plotted across susceptibility groups (TR: resistant, in blue; TS: susceptible, in green) and time points (day0: before infection; day132: under strongylid infection, 132 days after the onset of grazing season). Lysine intensity measured on faecal material in an independent cohort of British horses with low or high faecal egg count is represented in panel (**B**). Panel (**D**) shows the relationship between fibrobacter count in TR and TS ponies at day 132 and measured FEC.

First, circulating neutrophil and leukocyte counts showed significant increase following infection ($P = 0.02$ and 4×10^{-4} respectively), but this increase was not different between TR and TS ponies ($P = 0.5$ in both cases; Supplementary Fig. 8 and Supplementary Table 3). This trend was also not corroborated in another independent set of individuals (Fig. 5).

Second, plasma lysine level was significantly different between infected TR and TS ponies ($P = 0.002$) and matched independent observations made in faecal samples from another cohort of British horses (Fig. 6). Our temporal records also supported a differential decrease of plasma lysine levels between TR and TS ponies from worm-free to natural infection conditions ($P = 0.02$).

Third, sGCC-DA recognized three dominant gut bacteria, namely *Fibrobacter, Ruminococcus,* and *Treponema,* and one rare microbial taxon (*Coprococcus*) as keystone species whose covariation formed a distinct signature between the two pony groups (Fig. 6, Supplementary Fig. 8). Using linear regression modeling, we found a significant decrease in faecal abundance of *Treponema* ($P = 1.7 \times 10^{-2}$) upon infection that was shared across pony groups (Supplementary Fig. 8). On the contrary, the observed rise in *Fibrobacter* abundance at day 132 was slightly milder in the TS ponies (-0.73 ± 0.26, FDR $= 4.4 \times 10^{-2}$; Fig. 4, Supplementary Table 3).

Lysine and *Fibrobacter* would hence mark the intrinsically higher sensitivity of TS ponies.

Discussion

Inter-individual variation in resistance to strongylid infection offers the opportunity to reduce drug selection pressure by selectively treating the most susceptible individuals. To date, the physiological markers associated with the host resistance are to be defined. Here, we selected resistant and susceptible ponies based on their FEC and confirmed that their divergence in egg excretion was sustained over a 10-year time window. Using this unique set of individuals, we performed the first detailed investigations of the variations occurring in three different compartments (blood, plasma and faeces) to identify the underpinnings of their differential resistance to strongylid infection. Our findings are twofold. First, we found support for defining phenylalanine, circulating monocytes and faecal abundance of *Prevotella* as biomarkers of patent strongylid infection in equids. Second, we identified an immunometabolic network encompassing β-hydroxybutyrate level, short-chain fatty acid producing bacteria and circulating neutrophils that recapitulates the intrinsic resistance of ponies to patent infection under both worm-free and natural infection conditions.

For practical purposes, resistance was defined as the ability to limit egg excretion in the host as precise worm burden quantification after pony culling was not possible. To remove as many encysted larvae as possible, we applied a moxidectin treatment four months before the start of the trial. This was sufficient to ensure negative FEC in all but one susceptible individual at day 0. This indicates that some encysted larval stages had resumed their life-cycle in the meantime, as a result of the imperfect efficacy of that drug against this particular life-stage stage (between 60 and 95% efficacy according to studies)[45]. While the exact contribution of this life-stage to the reported parameter variations remains unknown, their putative presence would not affect our conclusions. Indeed, if they contribute to observed FEC or if ability to limit larval burden and FEC are well correlated traits, the experimental design would account for their contribution. However, if variation in larval burden across individuals is independent from their ability to control FEC, reported parameter variations would be partially confounded by the presence of developing larval stages. Current knowledge cannot however disentangle this point.

Elevation of plasma phenylalanine concentration has long been recognized as a metabolic consequence of bacterial and viral infections[46]. Similar observation has also been made in the faecal matter of infected horses[35] and we could replicate it from plasma samples. Of note, our data highlighted the covariation of this essential amino acid with average daily weight gain and the concomitant increase in plasma alkaline phosphatase concentration. This is compatible with strongyle infection reducing muscle protein synthesis (in line with observed reduced average daily weight gain), thereby increasing the extracellular release of phenylalanine and its subsequent uptake by the liver (whose activity was tracked by the increased alkaline phosphatase level). This model would match the theoretical framework derived from other infectious processes[46]. Altogether, plasma phenylalanine represents another diagnostic option for strongylid infection. The sensitivity and specificity of this marker remains to be determined in a larger cohort for field use in association with monocyte count and faecal *Prevotella* abundance.

The concomitant increase in the plasma level of this aromatic amino acid with the monocyte count decrease upon infection in susceptible individuals also opens new perspectives on the pathophysiology associated with strongylid infection in equids. Indeed, observations in humans suggested that an antibacterial compound derived from phenylalanine and produced by the gut microbiota could alter monocyte cell migration[47]. Specifically, D-phenyllactic acid can favour monocyte recruitment through the binding of their ydroxy-carboxylic acid (HCA) 3 G-protein coupled receptor[47]. While the only lactate-associated HCA1 receptor exists in equids, the possibility remains that phenylalanine derivatives could recruit immune cells[48]. Under this speculative model, serum phenylalanine (increased in infected individuals), would enter the gut lumen before subsequent transformation into phenyllactic acid by gut bacteria. Bacteria-derived phenyllactic would then be absorbed into the portal vein and serve as a recruitment signal for circulating monocytes. Further studies on this receptor would confirm its role in any differential immunometabolic regulations between the resistant and susceptible pony groups.

Under patent infection, susceptible individuals exhibited decreased abundance of faecal *Prevotella*. Effects of the *Prevotella* genus are still debated[49]. Experimental colonization of germ-free mice by *Prevotella* promoted the decrease of IL-18 interleukin expression, neutrophil recruitment at the site of infection and gut inflammation[49]. *Prevotella* was also associated with susceptibility to *Plasmodium* infection in mice[50], increased in patients infected by the human trematode *S. haematobium*[51] and in the colon of pigs infected by *Trichuris suis*[52]. A *Prevotella*-led inflammatory state would hence define the reduced strongyle infection observed in resistant ponies. This would contradict past findings in mice showing the detrimental association between IL-18 and *Trichuris muris* infection[53]. Because this cytokine seems to be sensitive to its environment, there is scope for microbiota-based regulation[54] that would differ between resistant and susceptible individuals.

In our attempt to discriminate between resistant and susceptible ponies, we identified a feature network with contrasting proinflammatory abilities. First, resistant ponies exhibited higher plasma lysine levels across conditions. Lysine is an essential amino acid for metabolism and immune response. Although the role of its derivative on the immune system is yet to be clarified[55], it is a natural ligand of the G protein-coupled receptor family C group 6 (GPRC6) receptor[56] that can modulate Th-2 response and antibody production by B cells[57]. Second, butyrate-producing bacteria and other bacteria able to metabolize β-hydroxybutyrate into butyrate (i.e. *Coprococcus*) defined differential baselines between pony susceptibility groups under worm-free conditions. The former is known to favour an anti-inflammatory state by inhibiting the neutrophil inflammasome and the release of pro-inflammatory cytokines like IL-1β and IL-18[58,59]. Butyrate binds free fatty acid receptors, like FFAR2 that is enriched on neutrophil cell surface[48], thereby promoting their intestinal recruitment[60]. Third, faecal *Fibrobacter* was a core discriminating feature of the intrinsic resistance potential. The drastic reduction of faecal *Fibrobacter* abundance in susceptible ponies upon strongylid infection mirrored past observations made in pigs infected with *Trichuris suis,* for which the reduction occurred irrespective of the worm load[61]. In goats, abundance of this genus was negatively correlated with the pro-inflammatory cytokines TNFα that is produced by monocytes[62]. *Fibrobacter* bacteria are key cellulose degraders that produce succinate subsequently converted into propionate,

another short-chain fatty acid able to modulate the recruitment of monocytes and neutrophils[48]. The reduction of this bacterial genus could hence result in a pro-inflammatory state favourable to the development of strongyle infection in susceptible ponies.

This immuno-metabolic network tying the host metabolites with its gut microbiota also points towards a key role of neutrophils in the definition of strongylid susceptibility. In line with this strand of evidence, we found a consistent covariation between circulating neutrophils and intrinsic susceptibility to strongylid infection. This relationship was however not validated in another independent population with higher excretion levels than that found during our 2015 experiment. The role played by neutrophils hence remains to be fully characterized as for other host-helminth models[63,64]. Transient neutrophilia (in the range of 9×10^9 cells per L) was previously described as the only haematological variation in ponies subjected to experimental infection by cyathostomins[32]. But neutrophils are not typically associated with type 2 immunity—that is effective against helminths[65]—and they do not play a role against the infection by *Trichinella spiralis*, a clade I nematode[66]. However, neutrophils can cooperate with macrophages to bring nematode infection under control as found for *Litomosoides sigmodontis*[67], *Strongyloides stercoralis*[68] and *Heligmosomoides polygyrus*[69]. They also appear to be regulated by the type-2 cytokinic environment to prevent damages to the host[63]. In horses, neutrophils respond to stimulation by IL-4—a type 2 cytokine—by an increase in the pro-inflammatory cytokines TNF-α and IL-8 but a decrease in IL1-β[70] that could be detrimental to the anti-helminth response. The covariation between neutrophil counts and the enhanced susceptibility in horses may thus result from qualitative differences. Application of single-cell RNAseq on neutrophil populations from the two pony backgrounds could provide significant advances in the understanding of their respective properties. Quantification of cytokines and reactive oxidative species production after in vitro exposure to infective strongylid larvae could also unravel distinct properties and putative bias toward a more effective response in resistant individuals.

Overall, this work suggests that FEC records from at least two years are sufficient to define the resistance potential of an equid to strongyle infection. It also proposes phenylalanine, monocyte counts and faecal *Prevotella* abundance as biomarkers of infection. Last, we identify a neutrophil-centered network tying together gut microbiota members and a few metabolites as the discriminant baseline between susceptible and resistant individuals. These results open novel perspectives for the understanding of strongylid susceptibility in equids and for nutritional modulation of these infections.

Materials and methods

This study relied on measurements gathered during a previous study[34]. Measured features are summarized herein and additional ¹H-NMR and FEC parameters collected during this experiment are described. Data were analysed with R[71] version 4.0.2 unless stated otherwise.

Selection and monitoring of 20 ponies with divergent susceptibility profile to strongyle infection.
Ponies were selected from the experimental herd (INRAE UE PAO, Nouzilly, France) according to their FEC history recorded since 2010. They usually graze from April to November according to meteorological conditions and maintained indoors under worm-free conditions (240 m² pen with slatted floors) otherwise. Because worm burden could not be determined under our setting, we considered resistance as the ability to limit egg excretion and select ponies based on that criterion. Briefly, individual pony random effect was estimated from individual records after correction for environmental fixed effects including year, season, age at sampling, time since last treatment and last anthelmintic drug received. This was implemented only for ponies on which data was collected three times a year, over at least two years. Using these estimates, the ten most susceptible (S) and ten most resistant (R) ponies were retained for this study. Median FEC over the past 5 years were 800 and 0 eggs/g for S and R ponies respectively[34]. These two groups of female Welsh ponies were matched for age (mean group ages of 3.63 years ± 1.33 and 4.44 years ± 1.30 in the R and S groups respectively).

Ponies were housed in November 2014. While indoors they received hay ad libitum and 600 g concentrate per animal per day (Tellus Thivat Nutrition Animale Propriétaire, Saint Germain de Salles, France) as outlined previously[34]. This concentrate contained oats and oat bran, wheat straw, barley, alfalfa, sugar beet pulp, molasses, salt and calcium carbonate and was enriched with vitamins and minerals[34]. To remove any intercurrent nematode infection, they were administered moxidectin and praziquantel (Equest Pramox°, Zoetis, Paris, France, 400 µg/kg of body weight of moxidectin and 2.5 mg/kg of praziquantel) on February 20th 2015 and kept indoors until they were turned out on pasture on June 20th 2015. Then, they remained on a 7.44 ha pasture from that date (day 0) to the end of October 2015 to observe the effects of natural infection in both groups (day 132). Pasture botanical analysis identified *Festuca arundinacea*, *Phleum prateonse*, *Poa abbreviate*, *Holcus lanatus*, and *Dactylis glomerata*.

At day 0 (4 months after moxidectin treatment), a single pony was FEC-positive (150 eggs/g). Although we did not explicitly determine egg reappearance period, this supports high efficacy level for moxidectin in this population in line with early observations after moxidectin release to the market[72]. It is however likely that some encysted larvae remained before the start of the pasture season. Nonetheless and because we aimed for a comparison between R and S groups at the end of pasture season, we decided to treat all ponies to reset their egg excretion. This was thought of as a conservative measure towards our hypothesis that would prevent unfair comparison between the S groups that harboured two FEC-positive individuals early in the season relative to the R group. Pyrantel pamoate (Strongid° paste, Zoetis, Paris, France; single oral dose of 1.36 mg pyrantel base per kg of body weight) was chosen as it is effective (horses were all FEC negative 20 days after treatment on July 30th 2015) and short-lived, thereby leaving scope for reinfection and FEC measure in late summer[72].

To identify biomarkers of infection, blood samples were taken at 0 and 132 days after the start of the trial to compare pony metabolomes before (indoors and under worm-free conditions) and after the pasture season

(under strongylid infection).. These were matched with FEC, ADG, haematological and blood biochemistry parameters and faecal microbiota as previously outlined[34].

All experiments were conducted in accordance with EU guidelines and French regulations (Directive 2010/63/EU, 2010; Rural Code, 2018; Decree No. 2013-118, 2013) in compliance with the ARRIVE guidelines. All experimental procedures were evaluated and approved by the Ministry of Higher Education and Research (APAFIS# 2015021210238289_v4, Notification-1). Procedures involving horses were evaluated by the ethics committee of the Val de Loire (CEEA VdL, committee number 19) and took place at the INRAE, Experimental unit of Animal Physiology of the Orfrasière (UE-1297 PAO, INRAE, Centre de Recherche Val de Loire, Nouzilly, France).

Clinical parameters measurement and modelling.
FEC were performed on 5 g of fresh faecal material after intra-rectal sampling and diluted in 70 mL of NaCl solution (d = 1.2) following a modified McMaster technique with a sensitivity of 50 eggs/g[73].

Haematological and serum biochemical parameters were recorded at PFIE (UE-1277 PFIE, INRAE, https://doi.org/10.15454/1.5535888072272498e12). Haematological parameters were determined after 15-min stirring at room temperature with a MS9-5 Haematology Counter* (digital automatic haematology analyzer, Melet Schloesing Laboratories, France). Recorded parameters encompassed erythrocyte, micro- and macro-erythrocyte counts, associated mean corpuscular volume, haematocrit, and mean corpuscular haemoglobin concentration. In addition, circulating thrombocytes were counted as well as leukocytes including lymphocytes, monocytes, neutrophils, basophils and eosinophils. For further analysis, an albumin to globulin ratio (AGR) was also performed.

In addition, serum biochemical parameters were quantified using Select-6 V rings with the M-Scan II Biochemical analyser (Melet Schloesing Laboratories, France). These parameters included albumin, cholesterol, globulin, glucose, alkaline phosphatase (ALP), total proteins (TP) and urea concentrations.

Normality of haematological and biochemistry parameters was tested with the Shapiro–Wilk's test (*shapiro.test()* function in R) and variables showing test value below 0.90, i.e. FEC, glucose, ALP and AGR concentrations, and eosinophil count were log-transformed (Supplementary Table 4). Validation of monocyte and neutrophil counts were obtained from a group of 25 ponies in October 2020 using the same setting.

Proton nuclear magnetic resonance (¹H NMR) data acquisition and processing.
Blood was collected in heparin coated tubes. Whole blood drawn for plasma generation was refrigerated immediately at 4 °C to minimize the metabolic activity of cells and enzymes and kept the metabolite pattern almost stable. After clotting at 4 °C, plasma was separated from blood cells and subsequently cryopreserved at − 80 °C and shipped as a single batch for ¹H-NMR profiling. Plasma samples were thawed on ice, 200 μL were mixed with 500 μL of deuterium oxide (D_2O) containing sodium trimethylsilylpropionate (TSP, 1 mM). Samples were vortexed, centrifuged (5500 g; 10 min; 4 °C) and 600 μL of supernatant were transferred into 5 mm NMR tubes.

The ¹H-NMR analysis was performed on a Bruker Avance III HD spectrometer (Bruker, Karlsruhe, Germany) operating at 600.13 MHz, and equipped with a 5 mm reversed ^1H-^{13}C-^{15}N-^{31}P cryoprobe connected to a cryoplatform. ¹H NMR spectra were acquired using a Carr-Purcell-Meiboom-Gill (CPMG) spin echo pulse sequence with a 2-s relaxation delay. The spectral width was set to 20 ppm and 128 scans were collected with 32 k points. Free induction decays were multiplied by an exponential window function (LB = 0.3 Hz) before Fourier Transform.

Spectra were manually phase and baseline corrected using Topspin 3.2 software (Bruker, Karlsruhe, Germany). All spectra were referenced to TSP (d 0 ppm). The spectral data were imported in the Amix software (version 3.9, Bruker, Rheinstetten, Germany) to perform data reduction in the region between 9.0 and 0.5 ppm with a bucket width of 0.01 ppm. The region between 5.1 and 4.5 ppm corresponding to water signal was excluded and data were normalized to the total intensity of the spectra.

¹H-NMR data were filtered from noise-related signals and finally consisted in 791 buckets ranging from 8.585 to 0.505 ppm (Supplementary Table 1). Buckets were annotated based on similarity of chemical shifts and coupling constants between plasma samples and reference compounds. The comparison was performed between one dimensional analytical data and reference compounds acquired under the same analytical conditions in our internal database, as well as from public databases like the Human Metabolome Database (HMDB, http://hmdb.ca/) and the Biological Magnetic Resonance Data Bank (http://www.bmrb.wisc.edu/). Manually curation was performed to isolate buckets corresponding to the same metabolite. Intensities of these metabolite-specific buckets were summed together to define metabolite-associated signals (hereafter referred to as "signals"), yielding 119 signals for further analysis.

To prevent spurious signals linked to age differences between individuals, metabolite levels at D0 were regressed upon pony age. None of the 119 metabolite signal intensities showed variation associated with pony ages at the considered cut-off (FDR < 5%).

16S faecal microbiota data for data integration.
Total microorganism's DNA was extracted from aliquots of frozen fecal samples (200 mg), using E.Z.N.A.* Stool DNA Kit (Omega Bio-Tek, Norcross, Georgia, USA). The V3-V4 16S rRNA gene amplification and sequencing have been described elsewhere[34].

For data integration purposes (described in next paragraph), 16S rRNA gene sequencing data generated from these ponies were re-analysed using qiime2 v.2020.2[74]. Adapters and primers were removed from sequencing data using cutadapt v2.1[75]. Trimmed data was subsequently imported into qiime2, and denoised using the Divisive Amplicon Denoising Algorithm from dada2[76]. In this workflow, Operational Taxonomic Units (OTU, sequence cluster defined by their dissimilarity level) are replaced by so-called amplicon sequence variant (ASV) that matches observed genetic variation in microorganism 16S rRNA gene amplicons instead of relying on a clustering operation[77]. Reads were quality filtered (maximal expected error of 1 for both reads), chimera trimmed (following the default "consensus" option) and the reads trimmed (6 and 20 bp for forward and reverse reads

respectively) to yield 3,054,206 reads. ASVs were subsequently filtered to retain that found in at least two individuals and supported by five reads, before alignment and phylogeny building with mafft and fasttree respectively. ASVs were subsequently assigned taxonomy using a naive Bayes classifier trained on the green genes reference database (gg_13_8) clustered at 99% similarity. This workflow left 6208 ASVs that were aggregated to 91 genera with phyloseq (v.1.32) for subsequent analysis. Total sum scaling normalization was applied to each taxon abundances for subsequent data integration analysis.

The raw sequences of the gut metagenome 16S rRNA targeted locus are available in NCBI under the Sequence Read Archive (SRA), with the BioProject number PRJNA413884 and SRA accession numbers from SAMN07773451 to SAMN07773550.

Statistical analysis and data integration. Our experiment focused on resistant and susceptible ponies monitored during a pasture season. Resistance in this context was defined as the ability to limit parasite egg excretion. Using this experimental design, we interrogated their respective metabolomes (1) to characterize metabolite trajectories throughout a single pasture season (June 2015 to Octobre 2015), (2) to identify biomarkers that would predict the pony intrinsic resistance to strongyle infection, (3) to identify biomarkers that could be used to differentiate between ponies reaching FEC cut-off for anthelmintic treatment (200 eggs/g) at the end of pasture season. We restricted this analysis to ponies whose predicted resistance status matched the observed FEC value at the end of pasture season, leaving nine true resistant (**TR**) and eight true susceptible (**TS**) ponies respectively. Statistical analyses applied for each of these three objectives are presented below.

Multivariate analysis of longitudinal metabolomic data in TR and TS ponies. First, we aimed to characterize metabolomic modifications occurring in TR and TS ponies throughout a grazing season. To analyse our longitudinal data, we implemented an ANOVA-Simultaneous Component Analysis (ASCA)[78,79] with the MetaboAnalyst R package v.3.0.3[80]. This analysis first partitions the variance contained in the metabolite data across the factors of interest (susceptibility group and time point including 0, 24 or 132 days after onset of the pasture season) and their interactions, thereby correcting the data for these effects. Simultaneously, a PCA is applied to each partition for dimensionality reduction ultimately isolating the metabolites contributing most to each effect. Significance of each effect was tested by 1000 permutations. Following ASCA, 37 outlier metabolite signals were identified and subsequently removed from further data integration analysis leaving 319 signals for subsequent analysis.

Differential analysis of signal metabolite intensities between TR and TS ponies and between ponies in need of treatment. Second, we aimed to identify biomarkers that would either be predictive of the intrinsic resistance potential of an individual (TR vs. TS comparison at day 0) or would distinguish between individuals in need of treatment (TR vs. TS comparison at day 132). To fulfill these two aims, we respectively performed Student's *t* tests on the 319 retained metabolite signals between TR and TS pony baselines or between individuals showing FEC below or above 200 eggs/g at day 132. To account for multiple testing, nominal *P*-values were adjusted using the Benjamini–Hochberg correction (FDR) as implemented in the *p.adjust* function (stats package v.4.0.2). Spearman's correlations were estimated using the *rcorr* function from the Hmisc package v.4.4–1.

Data integration to identify biomarkers associated with parasite resistance or need of treatment. As a complementary approach for biomarker identification, clinical (including blood cell population profile, blood biochemistry and average daily weight gain), metabolomic (319 signals) and faecal microbiota data were integrated to identify correlated signals associated with intrinsic resistance potential or need of treatment. This approach also has the potential to uncover biological signals that would be missed when considering each dataset independently[42]. We ran two analyses to extract the features from each dataset with best discriminant ability between TR and TS ponies either before the pasture season on one hand, or between TR and TS ponies in need of treatment at day 132 on the other hand.

In this analysis, ASV counts estimated from faecal microbiota data were aggregated at the genus level within each time point of interest (day 0 or day 132) using the *tax_glom* function of the phyloseq package v.1.32.0. They were then filtered to retain those reaching 5% prevalence (n = 45 and 41 at day 0 and day 132 respectively), and normalized with the centered log-ratio transformation of the mixOmics package[81]. ^{1}H NMR data consisted in the 319 metabolite signal intensities retained following filtering and outlier identification with ASCA. Clinical data consisted of 18 parameters. At day 0, ADG was not considered as no variation occurred leaving 17 parameters. FEC was not considered as it was used to define the groups to be compared.

Data integration was performed following the DIABLO (data integration analysis for biomarker discovery using latent variable approaches for Omics studies) framework[82] as implemented in the mixOmics package v.6.12.2. This algorithm implements a sparse generalized canonical correlation discriminant analysis (sGCC-DA)[42,82]. Briefly, DIABLO performs feature selection, thereby retaining the only bacterial genera, metabolite signal or clinical parameters with best discriminative ability between groups. Using this sparse dataset, DIABLO then seeks for latent components (linear combinations of features from each dataset, i.e. ^{1}H-NMR, faecal microbiota, and clinical parameters) that simultaneously explains as much as possible of the covariance between input datasets and the status of interest, *i.e.* pony resistance potential before or under infection[42,82]. We applied this analysis to discriminate between TR and TS ponies either before the onset of pasture season (day 0) or at pasture turnout (day 132). In each case, the number of features to be retained was determined by cross-validation analysis (10 × fivefold) with the *tune.block.splsda()* function exploring grids of 10–20 genera with increments of 5, 10 to 80 metabolite signals with increments of 10 and 5 to 10 clinical parameters with increment of 1. The

correlation matrix between each input datasets was determined after running a partial least square analysis using the *pls()* function[81].

Following data integration, we applied linear regression models to quantify group and time variation in the levels of seven features that distinguished between TR and TS ponies both at day 0 and day 132. These most discriminant features had either an absolute contribution of 0.25 on the first sGCC-DA axis. For metabolite signal intensities, no metabolite matched this condition and all features contributing to the 1st axis at day 0 and 132 were retained.

Enrichment analysis. To isolate biological pathways associated with discriminant ^{1}H NMR signals and bacterial taxa, enrichment analyses were run using the MetaboAnalyst v5.0[80] and MicrobiomeAnalyst[80] web interfaces respectively. Significant metabolite signals were tested for enrichment against KEGG[83] annotated metabolites and disease related blood biomarkers, while enrichment analysis on discriminant bacterial genera were run using taxa collections associated with aging or disease. Any enrichment with a FDR below 5% was deemed significant.

Validation in an independent set of data. We validated our biological signal on microbial and metabolomic data using the data from Peachey et al.[35] as an independent data set because they combined both 16S rRNA gene amplicon sequencing with metabolomic analyses. While they performed ^{1}H-NMR on faecal material, we wanted to evaluate how our results obtained from blood ^{1}H-NMR could match theirs. Raw data were retrieved and processed as ours, but following Peachey et al.' read truncation parameters[35]. This process left 3,305,166 sequences assigned to 5233 ASVs. ASV counts per sample and ASV taxonomic assignments from Clark et al. 2018 and Peachey et al. 2019 are available under the github repository.

Blood parameters of interest were validated in an independent set of 24 ponies with high FEC (821 eggs/g on average, ranging between 250 and 1700 eggs/g). Increased alkaline phosphatase had already been described in infected horses[32] and was not considered for validation.

Data availability

R script and raw data matrices are available. All the phenotypic and H-NMR data have been uploaded under the github repository available under the https://github.com/guiSalle/STROMAEQ repository. The raw sequences of the gut metagenome 16S rRNA targeted locus are available in NCBI under the Sequence Read Archive (SRA), with the BioProject number PRJNA413884 and SRA accession numbers from SAMN07773451 to SAMN07773550.

References

1. G. B. D. DALYs & Hale Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet* **390**, 1260–1344 (2017).
2. Nieuwhof, G. J. & Bishop, S. C. Costs of the major endemic diseases of sheep in Great Britain and the potential benefits of reduction in disease impact. *Anim. Sci.* **81**, 23–29 (2005).
3. Kaplan, R. M. & Vidyashankar, A. N. An inconvenient truth: Global worming and anthelmintic resistance. *Vet. Parasitol.* **186**, 70–78 (2012).
4. Lyons, E. Population-S benzimidazole- and tetrahydropyrimidine-resistant small strongyles in a pony herd in Kentucky (1977–1999): Effects of anthelmintic treatment on the parasites as determined in critical tests. *Parasitol. Res.* **91**, 407–411 (2003).
5. Näreaho, A., Vainio, K. & Oksanen, A. Impaired efficacy of ivermectin against *Parascaris equorum*, and both ivermectin and pyrantel against strongyle infections in trotter foals in Finland. *Vet. Parasitol.* **182**, 372–377 (2011).
6. Nielsen, M. K. *et al.* Anthelmintic efficacy against equine strongyles in the United States. *Vet. Parasitol.* **259**, 53–60 (2018).
7. Sallé, G. *et al.* Risk factor analysis of equine strongyle resistance to anthelmintics. *Int. J. Parasitol. Drugs Drug Resist.* **7**, 407–415 (2017).
8. Tzelos, T. *et al.* Strongyle egg reappearance period after moxidectin treatment and its relationship with management factors in UK equine populations. *Vet. Parasitol.* **237**, 70–76 (2017).
9. Peregrine, A. S., Molento, M. B., Kaplan, R. M. & Nielsen, M. K. Anthelmintic resistance in important parasites of horses: Does it really matter?. *Vet. Parasitol.* **201**, 1–8 (2014).
10. Debeffe, L. *et al.* Negative covariance between parasite load and body condition in a population of feral horses. *Parasitology* **143**, 983–997 (2016).
11. Rubenstein, D. I. & Hohmann, M. E. Parasites and social behavior of island feral horses. *Oikos* **55**, 312 (1989).
12. Nielsen, M. K., Martin, A. N., Scare, J. A. & Steuer, A. E. Precision and spatial variation of cyathostomin mucosal larval counts. *Vet. Parasitol.* **290**, 109349 (2021).
13. Gibson, T. E. The effect of repeated anthelmintic treatment with phenothiazine on the faecal egg counts of housed horses, with some observations on the life cycle of Trichonema spp. in the horse. *J. Helminthol.* **27**, 29–40 (1953).
14. Giles, C. J., Urquhart, K. A. & Longstaffe, J. A. Larval cyathostomiasis (immature trichonema-induced enteropathy): A report of 15 clinical cases. *Equine Vet. J.* **17**, 196–201 (1985).
15. Sallé, G. *et al.* Compilation of 29 years of postmortem examinations identifies major shifts in equine parasite prevalence from 2000 onwards. *Int. J. Parasitol.* https://doi.org/10.1016/j.ijpara.2019.11.004 (2020).
16. Kornaś, S. *et al.* Estimation of genetic parameters for resistance to gastro-intestinal nematodes in pure blood Arabian horses. *Int. J. Parasitol.* **45**, 237–242 (2015).
17. Scheuerle, M. C., Stear, M. J., Honeder, A., Becher, A. M. & Pfister, K. Repeatability of strongyle egg counts in naturally infected horses. *Vet. Parasitol.* **228**, 103–107 (2016).
18. Gold, S. *et al.* Quantitative genetics of gastrointestinal strongyle burden and associated body condition in feral horses. *Int. J. Parasitol. Parasites Wildl.* **9**, 104–111 (2019).
19. Stear, M. J., Park, M. & Bishop, S. C. The key components of resistance to *Ostertagia circumcincta* in lambs. *Parasitol. Today Pers. Ed.* **12**, 438–441 (1996).

13

20. Nielsen, M. K., Baptiste, K. E., Tolliver, S. C., Collins, S. S. & Lyons, E. T. Analysis of multiyear studies in horses in Kentucky to ascertain whether counts of eggs and larvae per gram of feces are reliable indicators of numbers of strongyles and ascarids present. *Vet. Parasitol.* **174**, 77–84 (2010).
21. Bishop, S. C. & Stear, M. J. Modeling of host genetics and resistance to infectious diseases: Understanding and controlling nematode infections. *Vet. Parasitol.* **115**, 147–166 (2003).
22. Råberg, L., Graham, A. L. & Read, A. F. Decomposing health: Tolerance and resistance to parasites in animals. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 37–49 (2009).
23. Sallé, G. *et al.* A genome scan for QTL affecting resistance to *Haemonchus contortus* in sheep. *J. Anim. Sci.* **90**, 4690–4705 (2012).
24. Kemper, K. E. *et al.* The distribution of SNP marker effects for faecal worm egg count in sheep, and the feasibility of using these markers to predict genetic merit for resistance to worm infections. *Genet. Res. Camb.* **93**, 203–219 (2011).
25. McRae, K. M., McEwan, J. C., Dodds, K. G. & Gemmell, N. J. Signatures of selection in sheep bred for resistance or susceptibility to gastrointestinal nematodes. *BMC Genom.* **15**, 637 (2014).
26. Sallé, G. *et al.* Functional investigation of a QTL affecting resistance to *Haemonchus contortus* in sheep. *Vet. Res.* **45**, 68 (2014).
27. Terefe, G. *et al.* Immune response to *Haemonchus contortus* infection in susceptible (INRA 401) and resistant (Barbados Black Belly) breeds of lambs. *Parasite Immunol.* **29**, 415–424 (2007).
28. Slusarewicz, P. *et al.* Automated parasite faecal egg counting using fluorescence labelling, smartphone image capture and computational image analysis. *Int. J. Parasitol.* **46**, 485–493 (2016).
29. Relf, V. E., Morgan, E. R., Hodgkinson, J. E. & Matthews, J. B. A questionnaire study on parasite control practices on UK breeding Thoroughbred studs. *Equine Vet. J.* **44**, 466–471 (2012).
30. Lester, H. E. *et al.* A cost comparison of faecal egg count-directed anthelmintic delivery versus interval programme treatments in horses. *Vet. Rec.* **173**, 371 (2013).
31. Sallé, G., Cortet, J., Koch, C., Reigner, F. & Cabaret, J. Economic assessment of FEC-based targeted selective drenching in horses. *Vet. Parasitol.* **214**, 159–166 (2015).
32. Murphy, D. & Love, S. The pathogenic effects of experimental cyathostome infections in ponies. *Vet. Parasitol.* **70**, 99–110 (1997).
33. Andersen, U. V. *et al.* Physiologic and systemic acute phase inflammatory responses in young horses repeatedly infected with cyathostomins and *Strongylus vulgaris*. *Vet. Parasitol.* **201**, 67–74 (2014).
34. Clark, A. *et al.* Strongyle infection and gut microbiota: Profiling of resistant and susceptible horses over a grazing season. *Front. Physiol.* **9**, 272 (2018).
35. Peachey, L. E. *et al.* Dysbiosis associated with acute helminth infections in herbivorous youngstock—Observations and implications. *Sci. Rep.* **9**, 1–16 (2019).
36. Walshe, N. *et al.* Removal of adult cyathostomins alters faecal microbiota and promotes an inflammatory phenotype in horses. *Int. J. Parasitol.* **49**, 489–500 (2019).
37. Daniels, S. P., Leng, J., Swann, J. R. & Proudman, C. J. Bugs and drugs: A systems biology approach to characterising the effect of moxidectin on the horse's faecal microbiome. *Anim. Microbiome* **2**, 38 (2020).
38. Wang, Y. *et al.* Systems metabolic effects of a necator americanus infection in *Syrian hamster*. *J. Proteome Res.* **8**, 5442–5450 (2009).
39. Wu, J.-F. *et al.* Metabolic alterations in the hamster co-infected with *Schistosoma japonicum* and *Necator americanus*. *Int. J. Parasitol.* **40**, 695–703 (2010).
40. Globisch, D. *et al.* *Onchocerca volvulus*—Neurotransmitter tyramine is a biomarker for river blindness. *Proc. Natl. Acad. Sci.* **110**, 4218–4223 (2013).
41. Lagatie, O. *et al.* Evaluation of the diagnostic potential of urinary N-acetyltyramine-O, β-glucuronide (NATOG) as diagnostic biomarker for *Onchocerca volvulus* infection. *Parasit. Vectors* **9**, 302 (2016).
42. The EPIC Consortium *et al.* Dynamic molecular changes during the first week of human life follow a robust developmental trajectory. *Nat. Commun.* **10**, 1–14 (2019).
43. Klassen, P., Fürst, P., Schulz, C., Mazariegos, M. & Solomons, N. W. Plasma free amino acid concentrations in healthy Guatemalan adults and in patients with classic dengue. *Am. J. Clin. Nutr.* **73**, 647–652 (2001).
44. Gevers, D. *et al.* The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* **15**, 382–392 (2014).
45. Xiao, L., Herd, R. P. & Majewski, G. A. Comparative efficacy of moxidectin and ivermectin against hypobiotic and encysted cyathostomes and other equine parasites. *Vet. Parasitol.* **53**, 83–90 (1994).
46. Wannemacher, R. W. Key role of various individual amino acids in host response to infection. *Am. J. Clin. Nutr.* **30**, 1269–1280 (1977).
47. Peters, A. *et al.* Metabolites of lactic acid bacteria present in fermented foods are highly potent agonists of human hydroxycarboxylic acid receptor 3. *PLoS Genet.* **15**, e1008145 (2019).
48. Husted, A. S., Trauelsen, M., Rudenko, O., Hjorth, S. A. & Schwartz, T. W. GPCR-mediated signaling of metabolites. *Cell Metab.* **25**, 777–796 (2017).
49. Iljazovic, A. *et al.* Perturbation of the gut microbiome by Prevotella spp. enhances host susceptibility to mucosal inflammation. *Mucosal Immunol.* **14**, 113–124 (2021).
50. Villarino, N. F. *et al.* Composition of the gut microbiota modulates the severity of malaria. *Proc. Natl. Acad. Sci.* **113**, 2235–2240 (2016).
51. Ajibola, O. *et al.* Urogenital schistosomiasis is associated with signatures of microbiome dysbiosis in Nigerian adolescents. *Sci. Rep.* **9**, 829 (2019).
52. Stolzenbach, S. *et al.* Dietary inulin and *Trichuris suis* infection promote beneficial bacteria throughout the porcine gut. *Front. Microbiol.* **11**, 312 (2020).
53. Helmby, H., Takeda, K., Akira, S. & Grencis, R. K. Interleukin (IL)-18 promotes the development of chronic gastrointestinal helminth infection by downregulating IL-13. *J. Exp. Med.* **194**, 355–364 (2001).
54. Reynolds, L. A. *et al.* MyD88 signaling inhibits protective immunity to the gastrointestinal helminth parasite *Heligmosomoides polygyrus*. *J. Immunol.* **193**, 2984–2993 (2014).
55. Kelly, B. & Pearce, E. L. Amino assets: How amino acids support immunity. *Cell Metab.* **32**, 154–175 (2020).
56. Wellendorph, P. *et al.* Deorphanization of GPRC6A: A promiscuous l-α-amino acid receptor with preference for basic amino acids. *Mol. Pharmacol.* **67**, 589–597 (2005).
57. Quandt, D., Rothe, K., Baerwald, C. & Rossol, M. GPRC6A mediates alum-induced Nlrp3 inflammasome activation but limits Th2 type antibody responses. *Sci. Rep.* **5**, 16719 (2015).
58. Youm, Y.-H. *et al.* The ketone metabolite β-hydroxybutyrate blocks NLRP3 inflammasome-mediated inflammatory disease. *Nat. Med.* **21**, 263–269 (2015).
59. Goldberg, E. L. *et al.* β-Hydroxybutyrate deactivates neutrophil NLRP3 inflammasome to relieve gout flares. *Cell Rep.* **18**, 2077–2087 (2017).
60. Sina, C. *et al.* G protein-coupled receptor 43 is essential for neutrophil recruitment during intestinal inflammation. *J. Immunol.* **183**, 7514–7522 (2009).
61. Wu, S. *et al.* Worm burden-dependent disruption of the porcine colon microbiota by *Trichuris suis* infection. *PLoS ONE* **7**, e35470 (2012).
62. Jiao, J. *et al.* Shifts in host mucosal innate immune function are associated with ruminal microbial succession in supplemental feeding and grazing goats at different ages. *Front. Microbiol.* **8**, 1655 (2017).

63. Heeb, L. E. M., Egholm, C., Impellizzieri, D., Ridder, F. & Boyman, O. Regulation of neutrophils in type 2 immune responses. *Curr. Opin. Immunol.* **54**, 115–122 (2018).
64. Allen, J. E., Sutherland, T. E. & Rückerl, D. IL-17 and neutrophils: Unexpected players in the type 2 immune response. *Curr. Opin. Immunol.* **34**, 99–106 (2015).
65. Sorobetea, D., Svensson-Frej, M. & Grencis, R. Immunity to gastrointestinal nematode infections. *Mucosal Immunol.* **11**, 304–315 (2018).
66. Michels, C. E. *et al.* Neither interleukin-4 receptor α expression on CD4$^+$ T cells, or macrophages and neutrophils is required for protective immunity to *Trichinella spiralis*. *Immunology* **128**, e385–e394 (2009).
67. Al-Qaoud, K. M. *et al.* A new mechanism for IL-5-dependent helminth control: Neutrophil accumulation and neutrophil-mediated worm encapsulation in murine filariasis are abolished in the absence of IL-5. *Int. Immunol.* **12**, 899–908 (2000).
68. Bonne-Année, S. *et al.* Human and mouse macrophages collaborate with neutrophils to kill larval *Strongyloides stercoralis*. *Infect. Immun.* **81**, 3346–3355 (2013).
69. Anthony, R. M. *et al.* Memory TH2 cells induce alternatively activated macrophages to mediate protection against nematode parasites. *Nat. Med.* **12**, 955–960 (2006).
70. Lavoie-Lamoureux, A. *et al.* IL-4 activates equine neutrophils and induces a mixed inflammatory cytokine expression profile with enhanced neutrophil chemotactic mediator release ex vivo. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **299**, L472–L482 (2010).
71. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2016).
72. Boersema, J. H., Eysker, M. & van der Aar, W. M. The reappearance of strongyle eggs in the faeces of horses after treatment with moxidectin. *Vet. Q.* **20**, 15–17 (1998).
73. Raynaud, J. P. Study of the efficiency of a quantitative coproscopic technic for the routine diagnosis and control of parasitic infestations of cattle, sheep, horses and swine. *Ann. Parasitol. Hum. Comp.* **45**, 321–342 (1970).
74. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
75. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.J.* **17**, 10 (2011).
76. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
77. Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **11**, 2639–2643 (2017).
78. Jansen, J. J., Hoefsloot, H. C. J., van der Greef, J., Timmerman, M. E. & Smilde, A. K. Multilevel component analysis of time-resolved metabolic fingerprinting data. *Anal. Chim. Acta* **530**, 173–183 (2005).
79. Smilde, A. K. *et al.* ANOVA-simultaneous component analysis (ASCA): A new tool for analyzing designed metabolomics data. *Bioinform. Oxf. Engl.* **21**, 3043–3048 (2005).
80. Chong, J. *et al.* MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. *Nucl. Acids Res.* **46**, W486–W494 (2018).
81. Rohart, F., Gautier, B., Singh, A. & Lê Cao, K.-A. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* **13**, e1005752 (2017).
82. Singh, A. *et al.* DIABLO: An integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinform. Oxf. Engl.* **35**, 3055–3062 (2019).
83. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucl. Acids Res.* **28**, 27–30 (2000).

Acknowledgements

Author contributions

G.S.: analyzed the data, drafted the manuscript. C.C.: 1H-NMR data acquisition and annotation. J.C., C.K., J.M.: sampling and parasitological data acquisition. F.R.: pony management and sampling. M.R., N.P.: haematological and biochemistry data acquisition. A.B.: raised funding, project management. N.M.: analyzed the data, drafted the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-93468-2.

Correspondence and requests for materials should be addressed to G.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.