





# *Listeria monocytogenes* faecal carriage is common and depends on the gut microbiota

Lukas Hafner<sup>1,8</sup>, Maxime Pichon<sup>2,3,7,8</sup> , Christophe Burucoa<sup>2,3,7</sup>, Sophie H. A. Nusser<sup>1</sup>, Alexandra Moura<sup>1,4</sup>, Marc Garcia-Garcera<sup>5,9</sup>  & Marc Lecuit<sup>1,4,6,9</sup>  

*Listeria* genus comprises two pathogenic species, *L. monocytogenes* (*Lm*) and *L. ivanovii*, and non-pathogenic species. All can thrive as saprophytes, whereas only pathogenic species cause systemic infections. Identifying *Listeria* species' respective biotopes is critical to understand the ecological contribution of *Listeria* virulence. In order to investigate the prevalence and abundance of *Listeria* species in various sources, we retrieved and analyzed 16S rRNA datasets from MG-RAST metagenomic database. 26% of datasets contain *Listeria sensu stricto* sequences, and *Lm* is the most prevalent species, most abundant in soil and host-associated environments, including 5% of human stools. *Lm* is also detected in 10% of human stool samples from an independent cohort of 900 healthy asymptomatic donors. A specific microbiota signature is associated with *Lm* faecal carriage, both in humans and experimentally inoculated mice, in which it precedes *Lm* faecal carriage. These results indicate that *Lm* faecal carriage is common and depends on the gut microbiota, and suggest that *Lm* faecal carriage is a crucial yet overlooked consequence of its virulence.

<sup>1</sup>Institut Pasteur, Université de Paris, Inserm U1117, Biology of Infection Unit, 75015 Paris, France. <sup>2</sup>University Hospital of Poitiers, Infectious Agents Department, Bacteriology and Infection Control Laboratory, 86021 Poitiers, France. <sup>3</sup>Université de Poitiers, Faculté de Médecine et de Pharmacie, EA 4331, 86022 Poitiers, France. <sup>4</sup>Institut Pasteur, National Reference Center and WHO Collaborating Center *Listeria*, 75015 Paris, France. <sup>5</sup>University of Lausanne, Department of Fundamental Microbiology, 1015 Lausanne, Switzerland. <sup>6</sup>Necker-Enfants Malades University Hospital, Division of Infectious Diseases and Tropical Medicine, APHP, Institut Imagine, 75006 Paris, France. <sup>7</sup>Present address: Université de Poitiers, Faculté de Médecine et de Pharmacie, Inserm U1070, 86022 Poitiers, France. <sup>8</sup>These authors contributed equally: Lukas Hafner, Maxime Pichon. <sup>9</sup>These authors jointly supervised this work: Marc Garcia-Garcera, Marc Lecuit. ✉email: [marc.lecuit@pasteur.fr](mailto:marc.lecuit@pasteur.fr)

Infectious disease symptoms can favour the transmission of pathogenic microorganisms and hence select genes that induce these symptoms (e.g. cough induced by *Mycobacterium tuberculosis*<sup>1</sup>). However, asymptomatic host colonisation can also favour microbial transmission, and thereby select for genes involved in host–microbe association that may also be involved in the development of opportunistic infections. *Listeria monocytogenes* (*Lm*) and *L. ivanovii* can cause infection in humans and other mammals including cattle<sup>2,3</sup>, leading to fetal–placental infection, abortion and meningo–encephalitis, in contrast to other *Listeria* species which are non-pathogenic. *Lm* is known to alternate between a saprophytic and a host-associated lifestyle during which it expresses so-called virulence factors that mediate tissue invasion and within-host dissemination<sup>4</sup>. Most of these virulence factors are part of *Lm* core genome and therefore under purifying selection<sup>5–7</sup>. *Lm* most virulent clonal complexes are also the most adapted to mammalian gut colonisation<sup>8</sup> and *Lm* can be released from infected tissues back to the intestinal lumen<sup>6,9,10</sup>, indicating that virulence may ultimately promote *Lm* faecal carriage and thereby play a major role in its dissemination.

*Lm* is a common contaminant of foodstuffs, and each human individual in Western countries is estimated to be exposed to *Lm* multiple times per year<sup>11</sup>. Yet the incidence of microbiologically proven invasive human listeriosis is extremely low, with 0.28 and 0.6 cases per 100,000 people in the United States and Europe, respectively<sup>12,13</sup>. This implies that in most cases, human exposure to *Lm* leads to either absence of clinically detectable infection and/or clinically silent gut colonisation, suggesting that *Lm* virulence genes are likely not selected for their capacity to induce clinically-overt disease. There have been reports of *Lm* asymptomatic faecal carriage among vertebrates including humans and cattle<sup>14–23</sup>, and almost all studies have suggested that the prevalence of *Lm* carriage is below 1%<sup>18,19,23</sup>. However, these studies were mostly based on culture-based methods<sup>18,19,23</sup>, which are less sensitive than molecular detection methods like PCR and sequencing<sup>20,21,24</sup>. Only one study suggested *Lm* faecal carriage to be above 1% by using PCR to detect *Lm*<sup>20</sup>. Large molecular studies on the distribution of *Listeria* species in mammals and the environment are not available<sup>25–27</sup>.

In this work, we compare the distribution of pathogenic and non-pathogenic *Listeria* species in 16S rRNA gene datasets from diverse origins. We show that *Lm* is more host-associated than non-pathogenic *Listeria* species, and, in contrast to non-pathogenic *Listeria* species, is also present in the faeces of healthy humans, both in publicly available 16S rRNA datasets and in an independent cohort of asymptomatic individuals. A specific microbiota signature is associated with *Lm* faecal carriage, both in humans and experimentally inoculated mice, in which it precedes *Lm* faecal carriage. Asymptomatic carriage might represent an important outcome of *Lm* virulence, in addition to clinically-overt disease<sup>28</sup>.

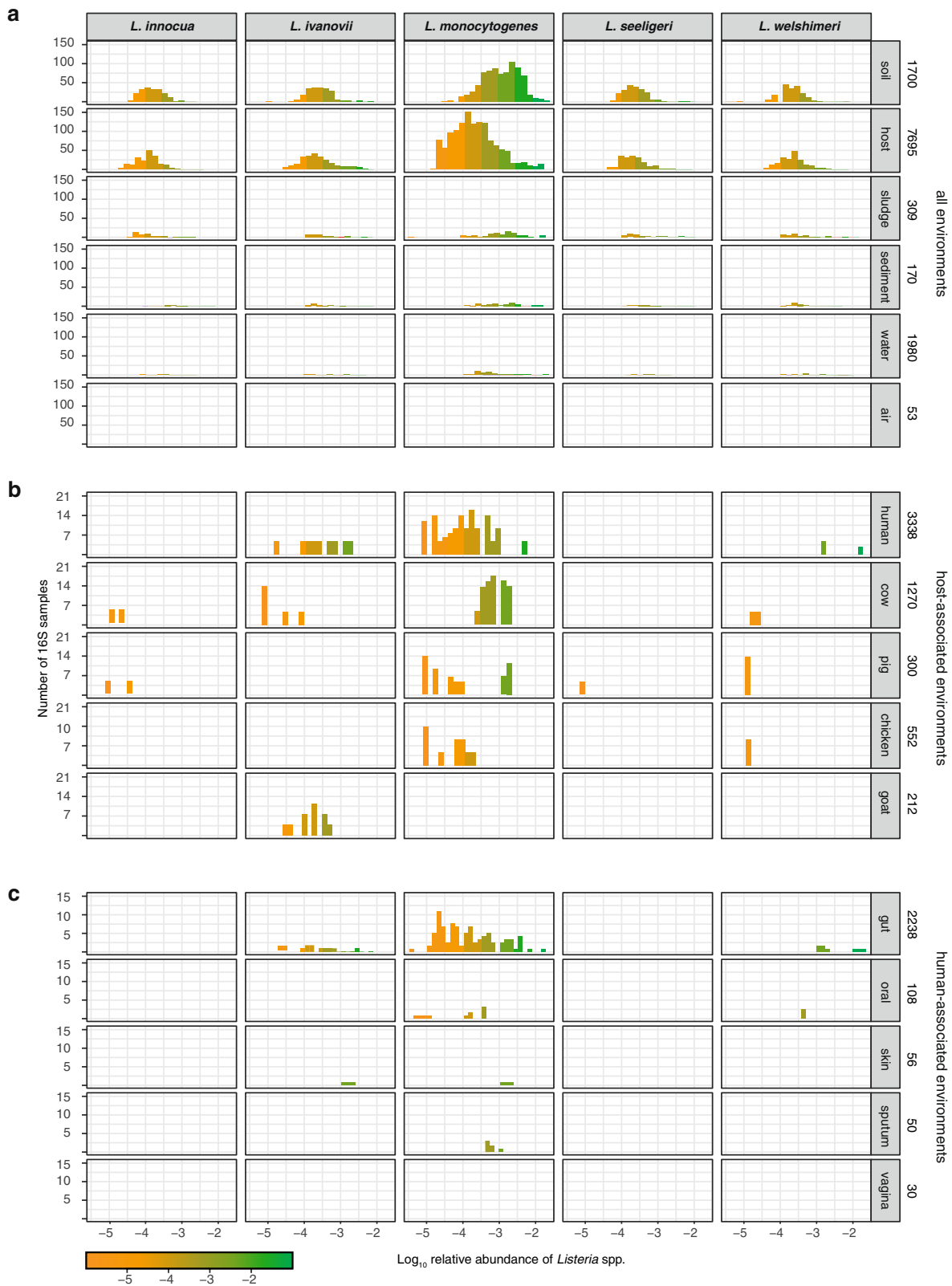
## Results

***Listeria monocytogenes* is more host-associated than non-pathogenic *Listeria* species.** Ecological sampling is influenced by a priori assumptions about potential niches<sup>29,30</sup>. Here we circumvented this limitation by assessing *Listeria* species distribution in publicly available metagenomic datasets from the large MG-RAST database<sup>31</sup>, to which high-quality metadata are associated, and retrieved 2490 full metagenomes and 11,907 16S rRNA high-quality datasets (see Methods). We assessed the impact of *Listeria* pathogenic potential on its ecological distribution by comparing the relative abundance (proportion of a species in a given sample, henceforth expressed as fractions, x-axis, Fig. 1) and prevalence (occurrence of a species in samples

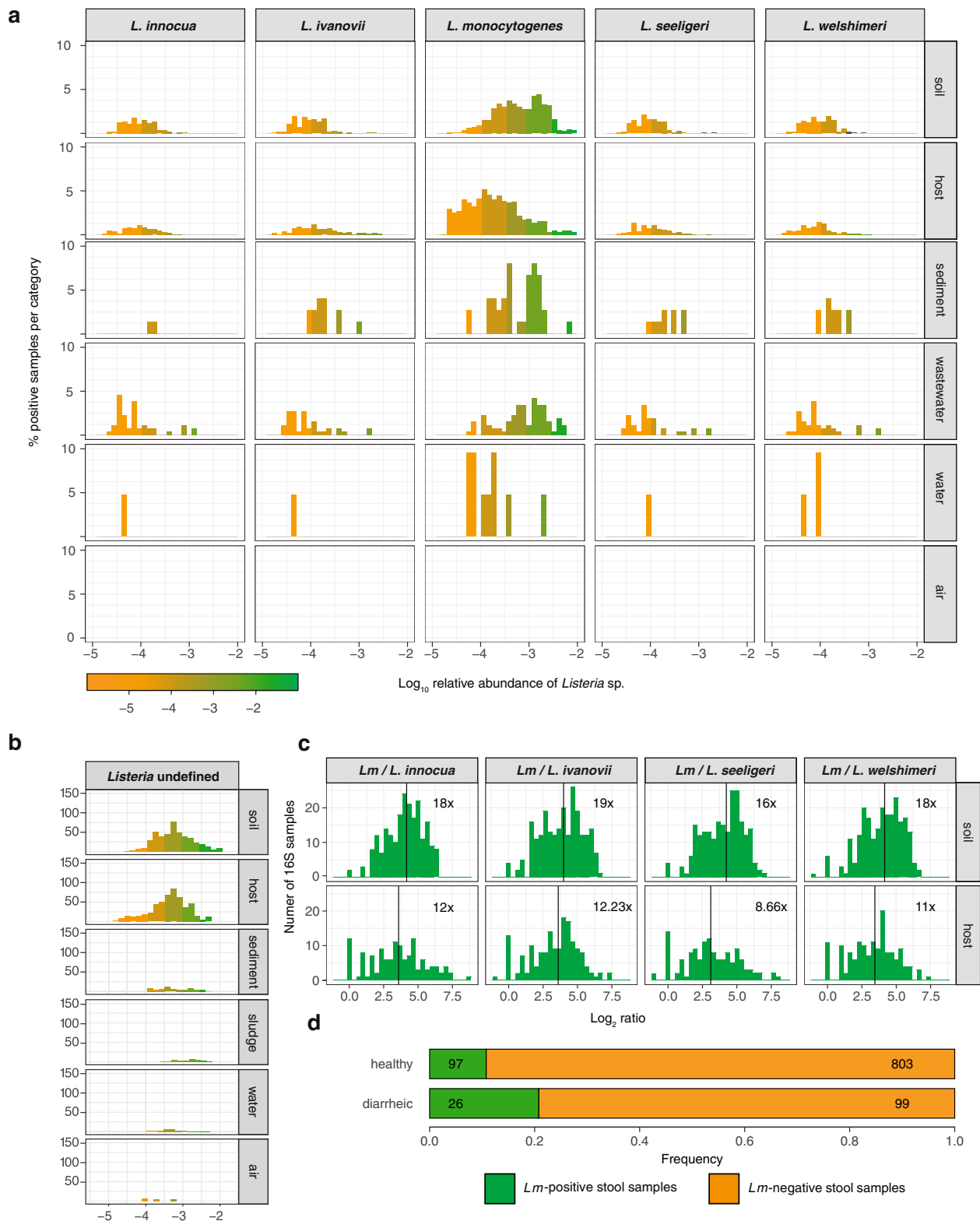
of a given category, y-axis, Fig. 1) of the *Listeria* pathogenic species *Lm* and *L. ivanovii* to that of the non-pathogenic species *L. innocua*, *L. seeligeri* and *L. welshimeri*<sup>32</sup>, termed together as *Listeria sensu stricto* (see Methods).

*Listeria sensu stricto* was detected in 26.06% 16S rRNA datasets (Figs. 1, 2a, b). Note that no positive result could be obtained using our approach (see Materials and Methods) analysing full metagenomes, in line with the relatively low abundance of *Listeria* species<sup>33</sup> and consistent with a higher sensitivity of 16S rRNA gene sequencing compared to full metagenome sequencing for a given sequencing depth<sup>34</sup>. *Lm* was most frequently present in soil (673/1700,  $\geq 39.59\%$ ; mean relative abundance  $1.2 \times 10^{-4}$ ), sludge (70/309,  $\geq 22.65\%$ ), sediment (32/170,  $\geq 18.82\%$ ) and host-associated samples (854/7695,  $\geq 11.10\%$ ; mean relative abundance  $9.0 \times 10^{-5}$ ). Only a few water samples (42/1980,  $\geq 2.12\%$ ) and no air sample (0/53) were positive for any *Listeria* species (Figs. 1a, 2a for normalised data per category). Note that these *Lm* prevalence are likely an underestimate, as the sequenced 16S rRNA region varies among the samples studied and does not always allow to discriminate between *Listeria* species (see Method section), and that the distribution of undefined *Listeria* hits followed a similar distribution to those obtained from discriminant regions of the 16S rRNA gene (Fig. 2b). *Lm* was the most prevalent *Listeria* species in both soil and host-associated environments (Fig. 1a). Further, in samples where more than one *Listeria* species was present, *Lm* was significantly more abundant than other *Listeria* species, both in soil and hosts (Fig. 2c). We next investigated the *Listeria* species host range (Fig. 1b). *Lm* was found to be the most abundant (mean relative abundance  $5.5 \times 10^{-3}$ ) and prevalent in cattle (80/1270;  $\geq 6.30\%$ ), which have indeed been reported as a potential reservoir for *Lm*<sup>26</sup>, especially hypervirulent clonal complexes<sup>3,8,15</sup>. We detected *Lm* in human samples at a similar prevalence to cattle (173/3338;  $\geq 5.18\%$ ), but 40 times less abundantly (mean relative abundance  $1.3 \times 10^{-4}$ ). *Lm* was also frequently found in chicken (mean relative abundance  $3.6 \times 10^{-4}$ , prevalence 28/552,  $\geq 5.05\%$ ) and pig samples (mean relative abundance  $4.7 \times 10^{-4}$ , prevalence 48/300,  $\geq 16\%$ ) but not that of goats (0/212), where only *L. ivanovii* was detected, consistent with the known association of *L. ivanovii* with small ruminants<sup>35</sup>. A high *Lm* prevalence in pigs and wild boars has been reported<sup>36–38</sup>, and pigs might constitute an underappreciated niche for *Lm*. We next investigated the human sampling sites in which *Lm* was present. As expected for a foodborne pathogen, *Lm* was detected in faecal samples (108/2238,  $\geq 4.83\%$ ; 108/1397, 7.73% in samples with discriminatory 16S rRNA sequences), but also in oral samples (7/108,  $\geq 6.48\%$ , Fig. 1c), consistent with reports that *Lm* may colonise both the gut and the oral cavity<sup>39,40</sup>. *Lm* was present in sputum (3/50,  $\geq 6.00\%$ ) and skin samples (2/56,  $\geq 3.57\%$ ) and absent in vaginal samples (0/30), but for these categories, only a few datasets were available for analysis. The non-pathogenic species *L. innocua* and *L. seeligeri* were not detected in any human-associated samples, while *L. ivanovii*, the only other pathogenic *Listeria* species, was detected, albeit far less than *Lm*, second most frequently in human stools (Fig. 1c).

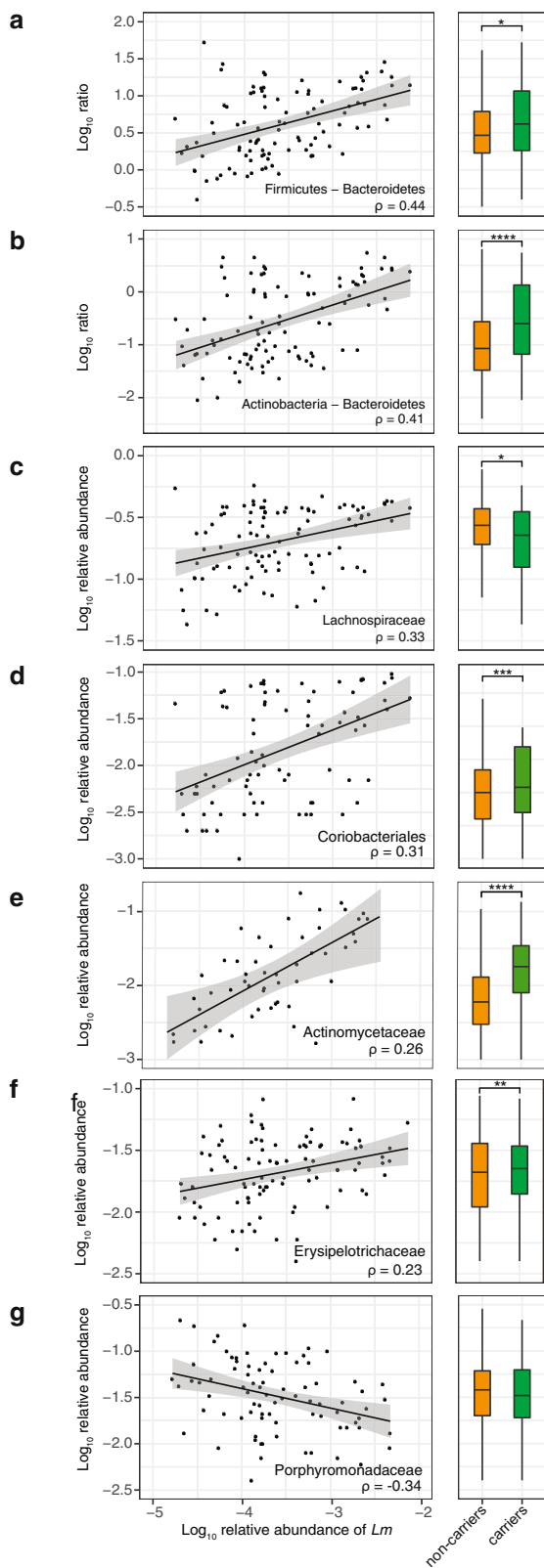
***Lm* faecal carriage is common in humans.** We aimed to replicate the result of frequent *Lm* carriage in humans independently and assessed *Lm* presence by *hly* PCR in the stools of a cohort of 900 healthy and a cohort of 125 diarrhoeic individuals (see Methods). It was detected in 10% (90/900) of healthy human stool samples and 20.8% (26/125) of diarrhoeic stools samples (Fig. 2d and Supplementary Tables 1 and 2). We confirmed that the sequence of the amplicon was that of *hly* in all of the samples we sequenced ( $N = 10$ ) (Supplementary Data 1). The enrichment of *Lm* in



**Fig. 1** *Lm* is more prevalent in host-associated environments than non-pathogenic *Listeria* species. Relative abundance and prevalence of *Listeria sensu stricto* species in 16S rRNA gene datasets in **a** different environments, **b** in selected host datasets (i.e. farm animals and/or known *Lm* reservoirs which were present >100x in our datasets) for which metadata detailing the host species were available and **c** from different sampling sites of healthy human hosts for which detailed metadata on body sampling site were available. Numbers on the right indicate samples per category.



**Fig. 2** *Lm* is more abundant than other *Listeria* species in all environments and *Lm* carriage is common in healthy individuals. **a** Same as Fig. 1a, normalised by category. **b** Relative abundance and prevalence of *Listeria* undefined in 16S rRNA gene datasets in different environments. **c** Log<sub>2</sub> of ratio of *Lm* to each other evaluated *Listeria* species in samples where the species co-occurred. Vertical lines and numbers indicate the mean of the distribution. **d** Prevalence of *Lm* in human faecal samples from healthy ( $n = 900$ ) and diarrhoeic donors ( $n = 125$ ) from France.



**Fig. 3** *Lm* faecal carriage correlates with a specific microbiota signature in humans.

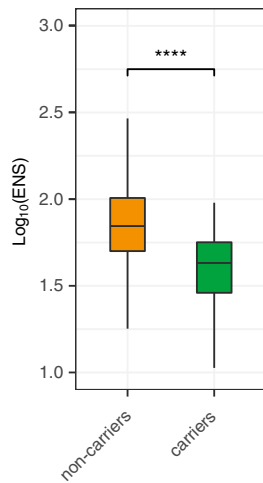
All significant correlations with more than 75 associated samples and  $\rho > 0.2$  or  $\rho < -0.2$  between *Lm* and commensal relative abundance in 108 healthy carrier (left panels) and comparison between carriers and non-carriers ( $n_{\text{non-carriers}} = 2130$ ) for the same groups (right panels). The lines in the left panels corresponds to linear regression models and the grey area to their 95% confidence interval: **a** The ratio of Firmicutes to Bacteroidetes phyla ( $\rho = 0.44$ ,  $p = 2.75 \times 10^{-5}$ ; non-carriers vs. carriers: 0.019; note that *Lm* species was excluded when the relative abundance of Firmicutes was calculated), **b** The ratio of Actinobacteria to Bacteroidetes ( $\rho = 0.414$ ,  $P = 6.1 \times 10^{-5}$ ; non-carriers vs. carriers:  $P = 9 \times 10^{-15}$ ), **c** Lachnospiraceae ( $\rho = 0.326$ ,  $P = 1.25 \times 10^{-3}$ ; non-carriers vs. carriers:  $P = 0.026$ ), **d** Coriobacteriales ( $\rho = 0.314$ ,  $P = 4.01 \times 10^{-2}$ ; non-carriers vs. carriers:  $P = 0.000459$ ), **e** Actinomycetaceae ( $\rho = 0.265$ ,  $P = 7.18 \times 10^{-11}$ ; non-carriers vs. carriers:  $P = 3.093 \times 10^{-48}$ ), **f** Erysipelotrichaceae ( $\rho = 0.226$ ,  $P = 4.51 \times 10^{-2}$ ; non-carriers vs. carriers:  $P = 0.000361$ ), **g** Porphyromonadaceae ( $\rho = -0.337$ ,  $P = 4.28 \times 10^{-3}$ ; non-carriers vs. carriers:  $P = 0.215$ ). The  $\rho$  values are Spearman correlation coefficients. Statistical comparison between carriers and non-carriers were performed with two-sided Wilcoxon rank-sum test with Benjamini-Hochberg correction for multiple test. \*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ , \*\*\*\*  $P < 0.0001$ . For boxplots, the hinges represent the first and third quartile of the distribution. The whiskers extend from the hinge to the largest or smallest value no further than  $1.5 \times \text{IQR}$  from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles).

differential exposure to *Lm*-contaminated food<sup>31</sup>. Neither age nor gender was associated to an asymptomatic carriage (Supplementary Table 2). Of note, our assay cannot distinguish between dead and viable *Lm* and the detection of *Lm* DNA could therefore correspond to dead *Lm*.

***Lm* carriage correlates with a specific gut microbiota signature.**

The gut microbiota is a major line of defence against foodborne pathogens, and several commensals exert a protective effect against enteropathogens<sup>43</sup>, including *Lm*<sup>44</sup>. *Lm* also produces bacteriocins that can alter microbiota composition<sup>45,46</sup>. In order to assess if microbiota composition has an impact on *Lm* faecal carriage in humans and vice versa, we investigated the relative abundance of microbiota taxonomic groups in MG-RAST human faecal samples. To take into account the compositional nature of data of different studies<sup>47</sup>, we calculated the ratios between microbiota phylogenetic groups and *Lm* abundance in the human microbiome datasets where *Lm* is present (Fig. 3 and Supplementary Data 2 and 3). *Lm* abundance correlated with the ratio of abundance of Firmicutes to Bacteroidetes phyla (Fig. 3a left), consistent with the observation that an increase of this ratio correlates with increased susceptibility to *Lm*<sup>48</sup>. This correlation is not due to *Lm* itself, as this species was excluded when the relative abundance of Firmicutes was calculated. The ratio of Actinobacteria to Bacteroidetes also correlated with *Lm* abundance (Fig. 3b left), and Actinobacteria were also significantly enriched compared to Firmicutes and Proteobacteria (Supplementary Data 2). *Lm* abundance also correlated positively at the family and order levels with Lachnospiraceae (Fig. 3c left), Coriobacteriales (Fig. 3d left), Actinomycetaceae (Fig. 3e left), Erysipelotrichaceae (Fig. 3f left) and negatively with Porphyromonadaceae (Fig. 3g left). Erysipelotrichaceae have previously been reported to be elevated in asymptomatic *C. difficile* carriers, which suggests that loss of colonisation resistance is associated with this family<sup>49</sup>. In line with our results, a protective effect of Porphyromonadaceae has also been observed against *Salmonella enterica* serovar Typhimurium<sup>50</sup>, *Enterococcus faecium*<sup>51</sup> and *C.*

diarrhoea samples ( $\chi^2 = 11.702$ ,  $P = 0.0018$ , Benjamini-Hochberg correction) is consistent with the observation that *Lm* can induce diarrhoea<sup>41,42</sup>. The difference in healthy asymptomatic donors in this cohort from France, relative to the 16S rRNA gene datasets from MG-RAST may be due to the different sensitivities of the two methods (targeted *hly* amplification versus total 16S amplification), and sample selection bias reflecting a potential



**Fig. 4 *Lm* faecal carriage correlates with low  $\alpha$ -diversity in humans.**  $\alpha$ -diversity, measured by ENS between carriers and non-carriers ( $n_{\text{carriers}} = 108$ ,  $n_{\text{noncarriers}} = 2130$ ,  $P < 10^{-15}$ ). Statistical comparison was performed with a two-sided Wilcoxon rank-sum test. For boxplots, the hinges represent the first and third quartile of the distribution. The whiskers extend from the hinge to the largest or smallest value no further than  $1.5 \times$  IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles). Points beyond this limit are shown.

*difficile*<sup>52</sup>. The aforementioned significant associations with *Lm* abundance in faecal carriers were also found significant between carriers and non-carriers (Fig. 3, right panels), with the exception of Porphyromonadaceae for which only a trend was observed (Fig. 3g right). For Lachnospiraceae, non-carriers showed a significantly higher prevalence than carriers (Fig. 3c right), reflecting that the comparisons between carriers and non-carriers are prone to study- and sample-dependent biases. Carriers also displayed less diverse microbiomes than non-carriers (Fig. 4), a finding consistent with the observation that  $\alpha$ -diversity is also involved in colonisation resistance<sup>43</sup> against enteropathogens such as *C. difficile*<sup>53</sup>, *Salmonella* or *Shigella*<sup>54</sup>. The overlap between the microbiota features associated with intestinal colonisation by *Lm* and other well-known gut-colonising bacteria is consistent with our finding that *Lm* is frequently present in stools of asymptomatic individuals.

***Lm* faecal carriage depends on gut microbiota.** *Lm* shedding from infected tissues back in the intestinal lumen may favour long-term faecal carriage<sup>6,9,10</sup>, in line with the finding that the most virulent *Lm* clonal complexes are the most adapted to the mammalian gut<sup>8</sup>, and the present observation that non-pathogenic species are not found in stool datasets retrieved from MG-RAST. To study *Lm* faecal carriage and its determinants experimentally, we inoculated mice intravenously with  $5 \times 10^3$  CFUs of *Lm* belonging to the hypervirulent clonal complex-1<sup>8,55</sup>. We observed a cage-dependent faecal carriage in 3/7 cages (11/26 mice). *Lm* could be detected over 30-days post-inoculation, at a time when all mice were asymptomatic. We classified faecal carriage as either heavy ( $>10^6$  CFU/g, six mice in two cages) or light ( $<10^6$  CFU/g, four mice in one cage, together with one noncarrier mouse) (Fig. 5a). In four cages (15 mice), no *Lm* was detected in the faeces 30-days post-inoculation (Fig. 5a). All mice had fully recovered from symptoms developed in the three first days following inoculation as assessed by weight gain, independently of their carrier status (Fig. 5b) and no mice exhibited any detectable clinical sign thereafter, as assessed by daily observation after the inoculation. We also separated mice

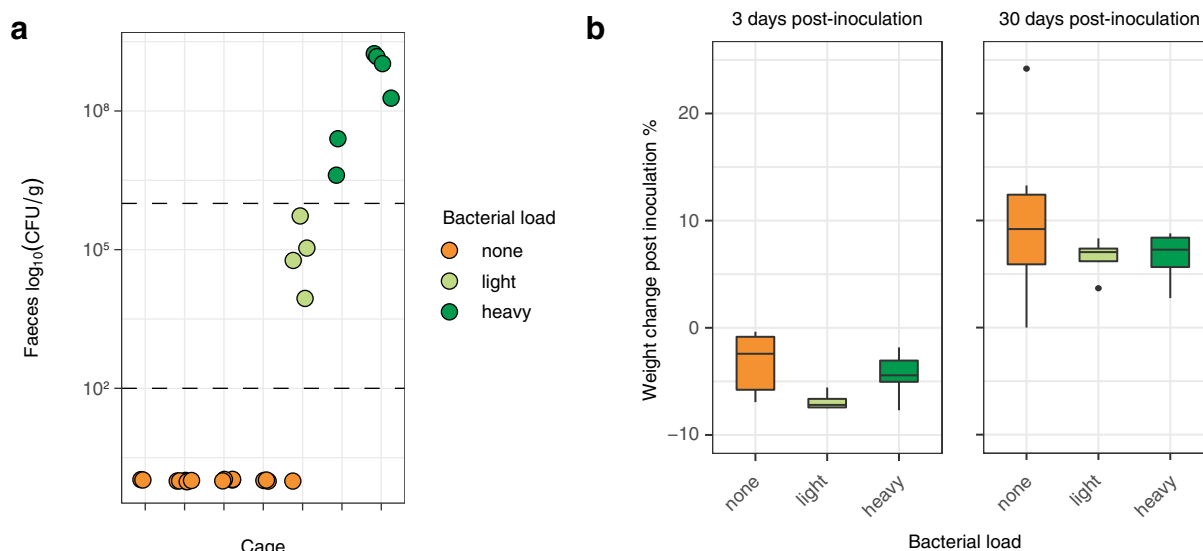
and observed persistent faecal carriage, ruling out that it was resulting from coprophagy.

Co-housed animals tend to have similar microbiota<sup>56</sup>, therefore the cage dependency of the observed differences in *Lm* carriage suggested that it was mediated by differences in gut microbiota composition. Indeed, heavy, light and noncarrier microbiota differed in microbial richness ( $\alpha$ -diversity) and composition ( $\beta$ -diversity): heavy carriers' microbiota was less diverse than that of light and non-carriers (Fig. 6a), in line with results obtained in humans (Fig. 4).  $\beta$ -diversity analysis showed that faecal carriage groups differed also in composition (PERMANOVA  $P < 0.001$ ): heavy carriers clustered separately from light and non-carriers (Fig. 6b). The difference between the light carrier group and the others reflected the higher homogeneity of the former (Fig. 6b), and the difference between non- and heavy carriers was mainly driven by a different composition in Bacteroidetes and Firmicutes: while OTUs classified as Porphyromonadaceae and Lachnospiraceae, respectively, were enriched in non-carriers, Bacteroidaceae were more present in heavy carriers (Figs. 6c, d and 7a-c).

We next investigated whether these differences in microbiota precede or result from *Lm* carriage (Fig. 8a, b). To this end, we compared microbiota 16S rRNA gene composition before *Lm* inoculation and 30-days post-inoculation. *Lm* inoculation only marginally affected microbiota composition, with only few OTUs (16/710) changes upon *Lm* inoculation. In sharp contrast, very significant OTUs differences were observed between carriers and non-carriers (685/1660) (Fig. 8b). The observation that the  $\beta$ -diversity difference observed between heavy and noncarrier microbiota preceded *Lm* inoculation (Axis 1 in Fig. 8a) suggests that it plays a causative role in the establishment of *Lm* carriage. To demonstrate that *Lm* carriage is indeed driven by the microbiota composition, we aimed to perturb it and assess if it would affect their permissiveness to *Lm* faecal carriage. As we had shown that a lower microbiota  $\alpha$ -diversity is associated with *Lm* faecal carriage both in mice and human 16S rRNA gene datasets, we treated mice with a broad range antibiotic oral cocktail shown to lower  $\alpha$ -diversity<sup>57</sup> and let the microbiota regrow for 4 weeks. We then inoculated these mice and PBS-treated littermates with *Lm*.  $\alpha$ -diversity was indeed significantly lower in antibiotic-treated mice at the time of inoculation, compared to PBS-treated mice (Fig. 8c). Both antibiotic- and PBS-treated mice lost weight upon *Lm* inoculation and recovered similarly (Fig. 8d). However, while none of the PBS-treated mice carried *Lm* in their faeces at 30-days post-inoculation, 6/7 antibiotic-treated mice carried *Lm* (Fig. 8e,  $P = 0.0013$ ). These results demonstrate that permissiveness to *Lm* carriage is driven by the composition of the gut microbiota.

## Discussion

Here we have shown that *Listeria* faecal carriage correlates with virulence: it is common in pathogenic *Listeria* species while it is rare in non-pathogenic species. This finding suggests that establishing faecal carriage, potentially through clinically silent tissue invasion and reseeded of the gut lumen via the gallbladder<sup>9</sup>, is a potential function of virulence genes in *Lm* ecology. Indeed, clinically-overt *Lm* infection is actually rare and is not involved in inter-human horizontal transmission<sup>58</sup>. This also implies that humans are not a focal host<sup>59</sup> for *Lm*. Consistent with this, *Lm* is more prevalent and abundant in cattle than in human stools, which is also in line with our recent report that hypervirulent *Lm* clonal complexes are associated to cattle and dairy products<sup>8</sup>. We also now report that the phylogeography of the hypervirulent *Lm* clonal complex-1 is linked to cattle global trade and farming<sup>60</sup>. Taken together, these observations strongly suggest



**Fig. 5** *Lm* faecal carriage in mice is cage-dependent. **a** CFU/g of the stool of male mice 30 days after an *iv* challenge with *Lm* at  $5 \times 10^3$  CFU from different cages (2–6 mice per cage). Colour indicates carriage group (<100 CFU/g: none, 100– $10^6$  CFU/g: light,  $>10^6$  CFU/g: heavy). Horizontal lines indicate the threshold between the groups. **b** Body weight change of mice after inoculation at 3 days post-inoculation and 30 days post-inoculation according to their carriage group ( $n_{\text{none}} = 16$ ,  $n_{\text{low}} = 4$ ,  $n_{\text{high}} = 6$ ; 3 dpi: none vs. heavy  $P = 0.37$ , low vs. heavy  $P = 0.11$ ; 30 dpi: none vs. heavy  $P = 0.37$ , low vs. heavy  $P = 0.91$ ). The statistical comparison was performed with a two-sided Wilcoxon rank-sum test. For boxplots, the hinges represent the first and third quartile of the distribution. The whiskers extend from the hinge to the largest or smallest value no further than  $1.5 \times \text{IQR}$  from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles). Points beyond this limit are shown.

that cattle constitute a major reservoir where *Lm* virulence is selected for.

We also found that *Lm* is the predominant *Listeria* species in the environment, where it is a saprophyte. *Lm* persistence in food processing plants, away from its vertebrate hosts, is associated with loss of virulence<sup>7,8,55</sup>. That *Lm* is found more abundantly in soil, sludge and sediments than non-pathogenic species (Fig. 1a) suggest that *Lm* regularly transits between its hosts via these environments, while maintaining its host-association capacity, which is mediated by its virulence genes. *Listeria* host-association capacity, therefore, appears as a trait that ensures the ecological success of *Lm* and *L. ivanovii* relative to other *Listeria* species. This does not exclude that *bona fide* virulence factors also contribute to *Lm* saprophytic lifestyle, as shown for ActA, which is involved in biofilm formation<sup>61</sup>. The relative lower prevalence in the environment of non-virulent *Listeria* species *L. innocua*, *L. seeligeri* and *L. welshimeri* which derive from the common virulent ancestor of *Lm* and *L. ivanovii*<sup>62</sup> suggest that they either (i) successfully colonise an environment not sampled in this study, and/or (ii) lost their focal host, and/or (iii) lost their host-association capacity, similar to *Lm* isolates associated with food processing plants which have lost or are in the process of losing virulence<sup>7,8,55</sup>. Future research will have to study the contribution of virulence and host-association to the overall ecological success of *Lm* and other microbial species which, as *Lm*, are widespread in the environment. It will also have to address the relative contribution of host and *Lm* genetics, food habits and intestinal microbiota to the asymptomatic faecal carriage of *Lm*.

## Methods

**Screening of *Listeria* sp. in 16S rRNA datasets.** A summary of the study workflow is represented in Supplementary Fig. 1. We collected 13,749 16S rRNA amplification datasets from MG-RAST from studies with >5 and <50 samples, for studies containing host samples <250 (last accessed: November 2017) as described in ref. 63. When more samples were available, we randomly selected 50 or 250 samples, respectively. We removed those containing non-ribosomal data or less than 2000 sequences using SSU-align v.1.01<sup>64</sup>. This left us with a total of 11,907 rRNA datasets (Supplementary Data 4, 5 for a list of all host species analysed in this study). Sequences shorter than 60 bp were removed. 16S rRNA sequence

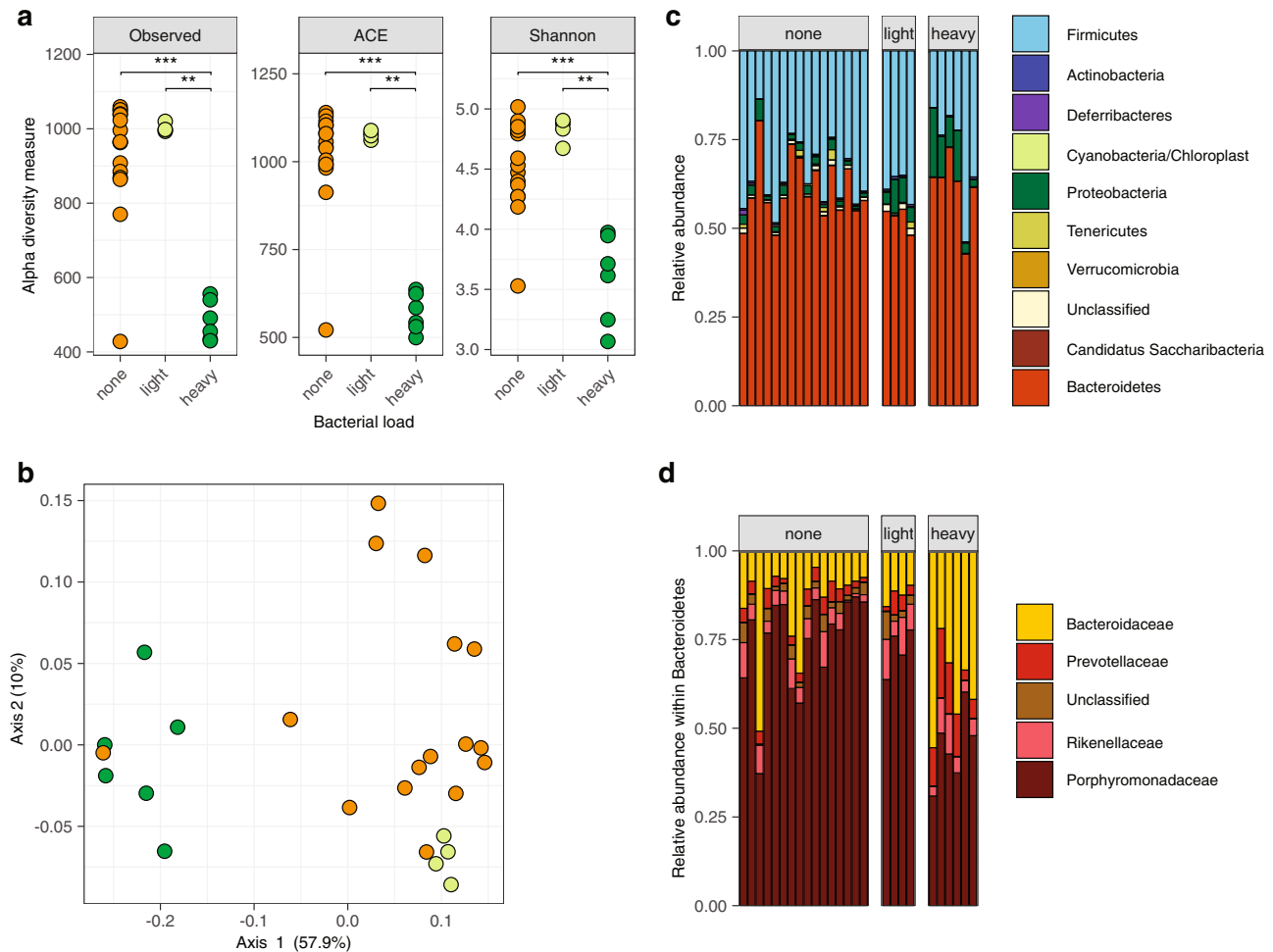
datasets were re-aligned using mafft v. 7.407<sup>65</sup>, and trimmed using trimal v.1.4<sup>66</sup> using the ‘automated1’ algorithm. The resulting trimmed sequences were then clustered within each sample at 99% identity and 90% coverage using the uclust algorithm from usearch v. 10.0.240<sup>67</sup>. Non-redundant reads which were present >3 were considered in the analysis. A representative sequence of each cluster was defined according to the distance to the cluster centroid. Henceforth, we will call these our environmental dataset.

To identify *Listeria* spp. in the environmental dataset we used a maximum likelihood approach. First, *Listeria sensu stricto* 16S rRNA reference sequences (accession numbers X56153, X98527, DQ065846, DQ065845 and X98528) were obtained from Genbank<sup>68</sup> and aligned using mafft with the ‘linsi’ algorithm. The resulting multiple sequence alignment was trimmed using trimal v.1.4. Phylogenetic reconstruction was then performed using IQ-tree v.1.6.5<sup>69</sup> using the GTR model (according to the model test) and 1000 rapid bootstrap iterations. The resulting tree was manually pruned to leave only one representative member of each clade. Environmental sequences were then classified as potential *Listeria* candidates by mapping them against the multiple sequence alignment using the ‘-adfragments’ algorithm of mafft. Sequences with at least 90% identity and 90% coverage to one reference member were kept for further analyses, or otherwise were discarded. The remaining sequences were then assigned to one of the branches of the phylogenetic tree using the evolutionary placement algorithm implemented in RAXML v. 8.2<sup>70</sup>. Environmental sequences assigned to any terminal branches with a maximum likelihood of 0.6 or higher, were classified as the specific *Listeria* species. Otherwise, they were classified as ‘*Listeria* undefined’. Note that this was the case for sequences with a non-discriminative amplicon region at the species level, e.g. V3–V4. In this work, we focused on all *Listeria sensu stricto* species, which are frequently found in the environment (*Lm*, *L. ivanovii*, *L. innocua*, *L. seeligeri* and *L. welshimeri*). We did not include the closest non-pathogenic relative of *Lm*, *L. marthii* since it is only rarely sampled in any environment<sup>71</sup>.

The remaining non-*Listeria* representative sequences were used to construct a global catalogue of operational taxonomic units (OTUs). To do so, the representative sequences of all datasets were grouped and clustered together at 97% identity using usearch, and the frequency of each OTU was calculated on each dataset. Finally, OTU representatives were taxonomically classified at the genus level using the RDP classifier<sup>72</sup>. At the same time, we defined the  $\alpha$ -diversity of each dataset as the Expected Number of Species (ENS). To do so, we did calculate the Shannon diversity index ( $H'$ ) (1):

$$H' = - \sum_{i=1}^R p_i \ln(p_i)$$

where  $p_i$  is the relative frequency of a specific species in the dataset (the number of sequences associated with the species divided by the total number of sequences assigned to species), and  $R$  is the number of datasets. We calculated the ENS as the



**Fig. 6** *Lm* faecal carriage correlates with a specific microbiota signature in mice. **a** Carriage groups differ in  $\alpha$ -diversity, measured by observed species (left), abundance-based coverage estimate (middle) and Shannon index (right), (Observed: none vs. heavy:  $P = 0.00017$ , light vs. heavy:  $P = 0.014$ , ACE: none vs. heavy:  $P = 0.00026$ , light vs. heavy:  $P = 0.0095$ , Shannon: none vs. heavy:  $P = 0.0001$ , light vs. heavy:  $P = 0.0095$ ). **b**  $\beta$ -diversity of mice microbiomes using MDS and Bray-Curtis distance. The colour indicates the carriage group (<100 CFU/g: none, 100–10<sup>6</sup> CFU/g: light, >10<sup>6</sup> CFU/g: heavy). All groups differed in composition (PERMANOVA overall  $P = 0.001$ , heavy/none  $P = 0.006$ , heavy/light  $P = 0.0075$ , light/none  $P = 0.031$ , with Benjamini–Hochberg correction). Light carriers were more homogeneous than other groups (permutation test for homogeneity of multivariate dispersion, heavy/none,  $P = 0.246$ , heavy/light  $P = 0.0160$ , light/none  $P = 0.0193$ , with Benjamini–Hochberg correction) **c** Microbiota composition of mice from Fig. 2c at phyla level and **d** family level within the Bacteroidetes phylum. Statistical comparison performed with two-sided Wilcoxon rank-sum test. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .

exponential of the Shannon diversity (2):

$$\text{ENS} = e^{H'}$$

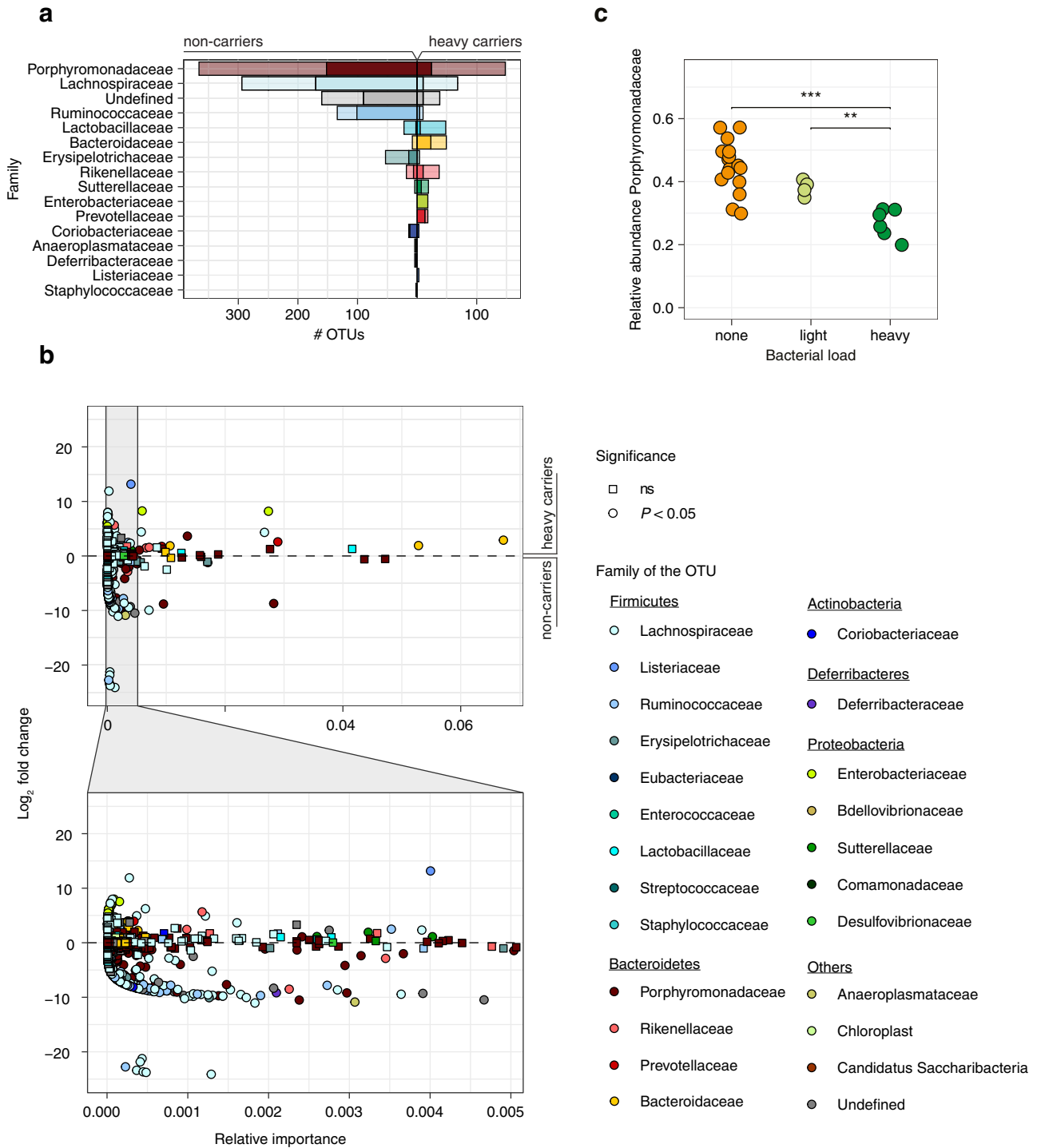
**Statistical analyses on the 16S rRNA datasets.** Given the lack of quantitative data associated with the public datasets, data were treated as compositional. In order to avoid methodological biases when comparing the datasets, only comparisons between ratios were performed. To assess associations between *Lm* abundance and other taxonomic groups, and given the non-parametric nature of the microbiome data, both Spearman correlation and Wilcoxon rank test were used to compare between the relative abundance of *Lm* and the different taxonomic ratios. Benjamini–Hochberg correction was used to avoid type I statistical error.

**Screening of *Listeria* sp. in full metagenomes.** We followed the approach described in Garcia-Garcera et al.<sup>73</sup>. In brief, we retrieved around 3000 metagenomes (Supplementary Data 6) from MG-RAST. To characterise the presence of *Listeria*, reads were mapped against a discriminative concatenate of genes, here the seven housekeeping genes used for multilocus sequence typing in *Lm*<sup>74</sup>. Since no reads could be mapped to these genes concatenate, no further downstream analysis has been performed. The absence of *Listeria* in full metagenomes can either be due to the low relative abundance of *Listeria* per se which does not permit it to be identified, or previous filtering of reads from species with low abundance.

**Detection of *Lm* in human faecal samples.** Determination of faecal carriage of *Lm* was performed by PCR amplifying *hly*<sup>24</sup>. We evaluated the performance of this method using artificial samples that mimic natural stool (Supplementary Table 1). Briefly, tenfold dilutions of the ATCC *Lm* strain (ATCC BAA751) in saline buffer (10<sup>8</sup> to 10<sup>1</sup> CFU/mL) were diluted in a 1:1 ratio in a PCR-negative stool sample conserved on eNat (Copan, Italy) before extraction. Extraction was performed with EasyMag (bioMérieux, Marcy-l’Etoile, France) according to manufacturer’s recommendations. PCR assays were performed in triplicate on a CFX96 system (BioRad, CA, USA) as described<sup>24</sup>. *Lm* was considered present when at least two of three PCR assays were positive. *Lm* detection threshold was 10<sup>6</sup> CFU/ml of stool.

Tested stool samples originated from two cohorts collected and stored on eNat (Copan): (i) the Hepystool cohort that included samples ( $n = 900$  samples, 2015–2016) from non-diarrhoeic patients (inclusion criteria described in ref. <sup>75</sup>) and (ii) stool samples from diarrhoeic patients, received at the Infectious Agents Department of the University Hospital of Poitiers, France. DNA was extracted on EasyMag (bioMérieux, Marcy-l’Etoile, France) according to manufacturer’s recommendations then amplified in triplicate. All samples which were at least once positive on the first triplicate were subjected to a second triplicate and were considered as positive when again detected at least once. The significance of the associations of the items listed in Supplementary Table 2 with *Lm* presence was assessed as follows: Student’s *t*-test for quantitative data and  $\chi^2$ -test for qualitative data except for birthplaces where a Fisher exact test was used.



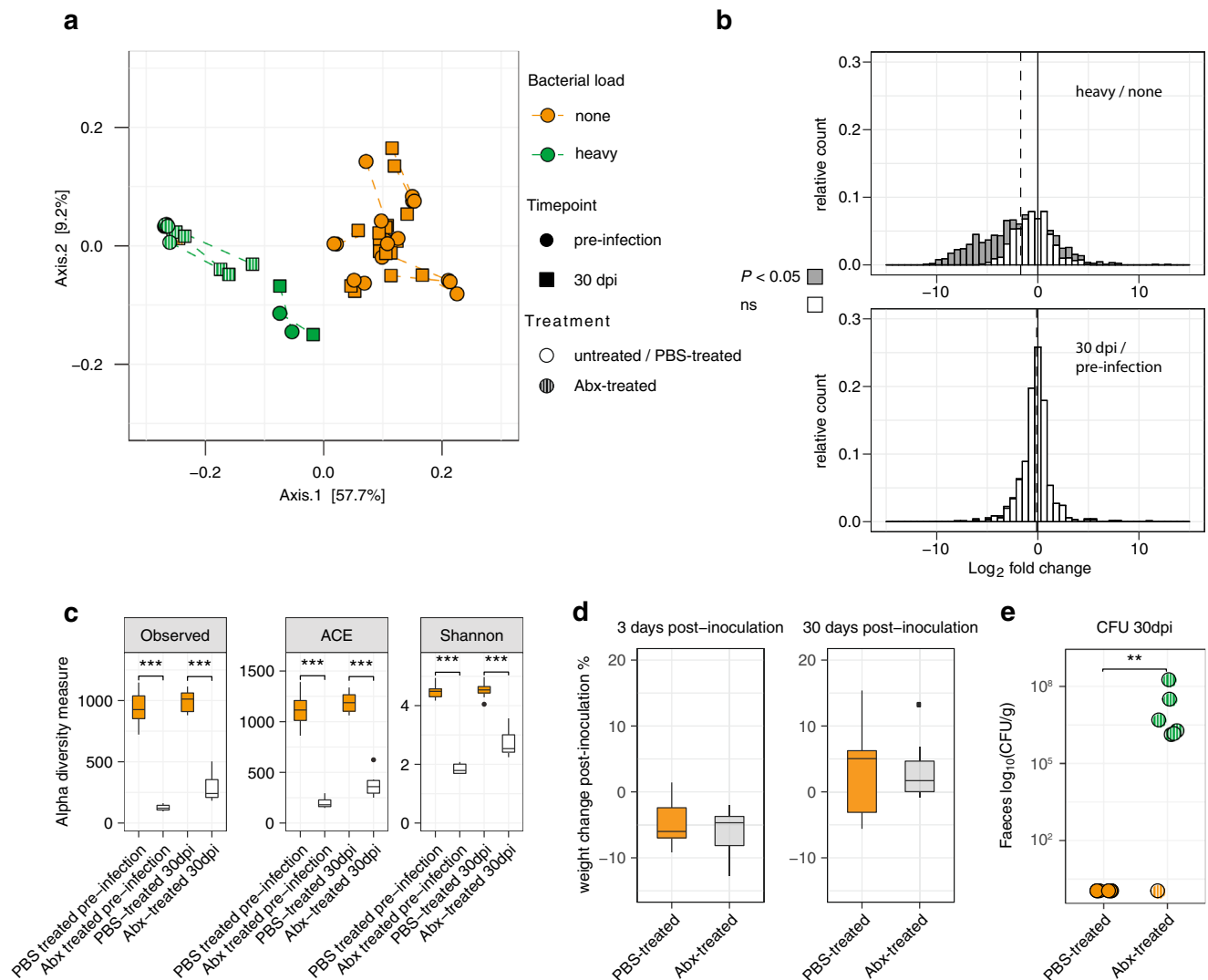


**Fig. 7 Microbiota composition in mice differs according to *Lm* faecal carriage.** **a** Count of all OTUs enriched in non-carriers (left) or heavy carriers (right) identified by DESeq2. Opaque bars represent significantly different OTUs (after Benjamini–Hochberg correction). **b** Mean abundance and fold change of all OTUs between carrier and noncarrier mice determined by DESeq2. Opaque points significantly differ between the two groups ( $P$  after Benjamini–Hochberg correction). **c** Relative abundance of Porphyromonadaceae in 16S rRNA data from mice from different carriage groups ( $n_{\text{none}} = 16$ ,  $n_{\text{low}} = 4$ ,  $n_{\text{heavy}} = 6$ , none vs. heavy:  $P = 0.0001$ , light vs. heavy:  $P = 0.0095$ ). The statistical comparison was performed with a two-sided Wilcoxon rank-sum test. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

Note that no culture-based identification was applicable given that the eNat protocol conserves nucleic acid and is bactericidal within 30 min (see manufacturer’s instructions for details). We sequenced the *hly* gene from position 327 to 829 with the primers CAAAATAATGCAGACATCCAAG and CTTTAGTAACAGCTTTGCCG for a subset ( $N = 10$ ) of *Lm* positive samples to ensure the specificity of the *hly* PCR assay. Sequences from Sanger

sequencing with CAAAATAATGCAGACATCCAAG and EGD-*e hly* sequence retrieved from <https://bigsd.b.pasteur.fr/listeria/listeria.html> were aligned with clustal omega<sup>76</sup>.

**Mouse colonisation experiments.** Seven- to 11-week-old male mice (C57BL/6 mEad E16P KI<sup>77</sup>) were infected intravenously in the tail vein as previously



**Fig. 8** *Lm* faecal carriage depends on the microbiota. **a** Comparison of microbiota pre-inoculation and 30-days post-inoculation.  $\beta$ -diversity of mice microbiomes using MDS and Bray-Curtis distance. The colour indicates the carriage group ( $<100$  CFU/g: none,  $>10^6$  CFU/g: heavy), the shape of the timepoint (round: Pre-inoculation, square: 30-days post-inoculation) and opacity of the treatment (plain: untreated or PBS-treated, striped: antibiotic treated) **b** Distribution of fold change determined by DESeq2 of OTUs between heavy and non-carriers (top) and pre- and post-infection (bottom). Dark bars indicate significantly differently present OTUs identified by DESeq2 ( $P < 0.05$  after Benjamini-Hochberg correction). The dashed bar indicates the median of the respective distribution. **c** Comparison of  $\alpha$ -diversity, measured by observed species (left), abundance-based coverage estimates (middle) and Shannon index (right) between antibiotic- and PBS-treated mice 4 weeks after antibiotic treatment/before inoculation and 4 weeks after *Lm* inoculation. ( $n_{\text{Abx-treated}} = 7$ ,  $n_{\text{PBS-treated}} = 9$ , PBS-treated vs. Abx-treated pre-infection  $P = 0.00017$ , 30dpi  $P = 0.00017$ ). **d** Body weight change of mice at 3 days post-inoculation and 30-days post-inoculation according to their treatment group ( $n_{\text{PBS-treated}} = 9$ ,  $n_{\text{Abx-treated}} = 7$ , PBS-treated vs. Abx-treated 3 dpi:  $P = 0.76$ , 30dpi:  $P = 0.84$ ). **e** CFU/g of the stool of female mice 30 days after an *iv* challenge with *Lm* at  $5 \times 10^3$  CFU. Mice were either treated with antibiotics or PBS 4 weeks prior to inoculation. Statistical comparison performed with two-sided Wilcoxon rank-sum test (PBS-treated vs. Abx-treated:  $P = 0.0013$ ). The statistical comparison was performed with a two-sided Wilcoxon rank-sum test. For boxplots, the hinges represent the first and third quartile of the distribution. The whiskers extend from the hinge to the largest or smallest value no further than  $1.5 \times$  IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles). Points beyond this limit are shown.

described<sup>77</sup>. Fraternities were kept together in a cage during the whole experiment, except when separated to exclude that carriage was due to coprophagy. To quantify carriage at 30-days post-inoculation, faeces were collected from each individual mouse and weighted before being homogenised in 2 ml of PBS. CFU count was performed by serial dilution of homogenised faeces on ALOA plates (Biomérieux) as described in ref. <sup>8</sup>. Separate faecal pellets were collected pre-inoculation and/or 30-days post-inoculation and stored at  $-20^\circ\text{C}$  for DNA extraction for 16S rRNA gene sequencing. DNA was extracted with DNeasy PowerSoil Pro Kit (Qiagen).

For experiments with antibiotic treatment, 5-week-old female mice (C57BL/6 mEcad E16P KI<sup>77</sup>) received orally every 12 h during 4 days PBS (as control) or the following antibiotics (similarly to<sup>57</sup>): ampicillin (100 mg/kg, Sigma-Aldrich A9618), vancomycin (50 mg/kg, Sigma-Aldrich 75423), metronidazole (100 mg/kg,

Abcam ab141218), neomycin (100 mg/kg, Gibco 11811-031), amphotericin B (1 mg/kg, ApexBio Technology B1885). The mice were inoculated, as described above, 4 weeks after the antibiotic treatment.

**16S rRNA gene analysis in mice.** DNA from faeces has been isolated with DNeasy PowerSoil Kit (Qiagen, Cat. No. 47016) accordingly to the manufacturer's instructions. The V4 region has been amplified and sequenced with the primers CCTACGGGNGGCWGCAG and GACTACNVGGGTWCTAATCC using the Illumina MiSeq workflow at the biomics platform at the Institut Pasteur, Paris. Analysis has been performed with micca v.1.7.2<sup>78</sup>, using the RDP classifier<sup>79</sup> and unioise3 for clustering<sup>80</sup>. Forward and reverse reads were merged with a minimum overlap of 100 bp and 30 maximum allowed mismatches. Forward and reverse primers were removed and reads were trimmed to 400 nucleotides

using the micca workflow. Reads with an expected error rate above 0.75% were excluded. Reads were grouped in sequence variants by unoise<sup>30</sup> and chimeric sequences were removed. Sequence variants were classified with RDP<sup>79</sup>, which uses VSEARCH to match sequences with the reference database<sup>81</sup>. Statistical were performed with R and the phyloseq v.1.34.0, vegan and microbiome libraries<sup>82,83</sup>.  $\alpha$ -diversity has been calculated by the number of observed species, abundance-based coverage estimator (ACE)<sup>84</sup> and Shannon index<sup>85</sup>.  $\beta$ -diversity between samples has been calculated with MDS of Bray–Curtis dissimilarities<sup>86</sup>. PERMANOVA and homogeneity between microbiome groups were calculated with Adonis and Betadispers from the vegan library<sup>87</sup>. Differentially present OTUs between groups were identified with DESeq2 v.1.30.1<sup>88</sup>. To identify OTUs that differ between pre-inoculation and 30 dpi, the ‘~mouse\_id + timepoint’ design was used.

**Housing conditions of mice.** Animals were maintained in a facility which is licensed by the French Ministry of Agriculture (agreement B 75-15-01, 04, 05, 06, 07, 08, 09, 11, A 75-15-13 dated 22 May 2008 and A 75-15-27 dated 12 November 2004). The facility has central air conditioning equipment which maintains a constant temperature of  $22 \pm 2$  °C. Air is renewed at least 20 times per hour in animal rooms. Light is provided with a 14:10 h light:dark cycle (6:30 a.m. to 8:30 p.m.). Animals were kept in polypropylene or polycarbonate cages which comply with European regulations in terms of floor surface per animal. All cages are covered with stainless steel grids and non-woven filter caps.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

Primary sequencing data are available on the Sequence Read Archive under the entry PRJNA642013. 16S rRNA datasets for screening of *Listeria* sp. were retrieved from MG-RAST.

### Code availability

Code for screening of a microbe of interest in 16S rRNA gene data is published in ref. <sup>73</sup>.

Received: 13 January 2021; Accepted: 3 November 2021;

Published online: 24 November 2021

### References

- Ruhl, C. R. et al. *Mycobacterium tuberculosis* Sulfolipid-1 activates nociceptive neurons and induces cough. *Cell* **181**, 293–305 (2020). e11.
- Charlier, C. et al. Clinical features and prognostic factors of listeriosis: the MONALISA national prospective cohort study. *Lancet Infect. Dis.* **17**, 510–519 (2017).
- Dreyer, M. et al. *Listeria monocytogenes* sequence type 1 is predominant in ruminant rhombencephalitis. *Sci. Rep.* **6**, 36419 (2016).
- Toledo-Arana, A. et al. The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* **459**, 950–956 (2009).
- Moura, A. et al. Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. *Nat. Microbiol.* **2**, 1–10 (2016).
- Louie, A., Zhang, T., Becattini, S., Waldor, M. K. & Portnoy, D. A. A multiorgan trafficking circuit provides purifying selection of *Listeria monocytogenes* virulence genes. *MBio* **10**, e02948–19 (2019).
- Mauray, M. M. et al. Spontaneous loss of virulence in natural populations of *Listeria monocytogenes*. *Infect. Immun.* **85**, e00541-17 (2017).
- Mauray, M. M. et al. Hypervirulent *Listeria monocytogenes* clones’ adaption to mammalian gut accounts for their association with dairy products. *Nat. Commun.* **10**, 3619 (2019).
- Zhang, T. et al. Deciphering the landscape of host barriers to *Listeria monocytogenes* infection. *Proc. Natl Acad. Sci. USA* **114**, 6334–6339 (2017).
- Melton-Witt, J. A., Rafelski, S. M., Portnoy, D. A. & Bakardjiev, A. I. Oral infection with signature-tagged *Listeria monocytogenes* reveals organ-specific growth and dissemination routes in guinea pigs. *Infect. Immun.* **80**, 720–732 (2012).
- Ricci, A. et al. *Listeria monocytogenes* contamination of ready-to-eat foods and the risk for human health in the EU. *EFSA J.* **16**, e05134 (2018).
- Scallan, E. et al. Foodborne illness acquired in the United States—Major pathogens. *Emerg. Infect. Dis.* **17**, 7–15 (2011).
- Pohl, A. M. et al. Differences among incidence rates of invasive listeriosis in the U.S. FoodNet population by age, sex, race/ethnicity, and pregnancy status, 2008–2016. *Foodborne Pathog. Dis.* **16**, 290–297 (2019).
- Lamont, R. J. & Postlethwaite, R. Carriage of *Listeria monocytogenes* and related species in pregnant and non-pregnant women in Aberdeen, Scotland. *J. Infect.* **13**, 187–193 (1986).
- Esteban, J. I., Oporto, B., Aduriz, G., Juste, R. A. & Hurtado, A. Faecal shedding and strain diversity of *Listeria monocytogenes* in healthy ruminants and swine in Northern Spain. *BMC Vet. Res.* **5**, 2 (2009).
- Schuchat, A., Swaminathan, B. & Broome, C. V. Epidemiology of human listeriosis. *Clin. Microbiol. Rev.* **4**, 169–183 (1991). vol.
- Kampelmacher, E. H. & Van Noorle Jansen, L. M. Listeriosis in humans and animals in the Netherlands (1958–1977). *Zentralbl. Bakteriol. A* **246**, 211–227 (1980).
- Müller, H. E. *Listeria* isolations from feces of patients with diarrhea and from healthy food handlers. *Infection* **18**, 97–100 (1990).
- Macgowan, A. P., Marshall, R. J. & Mackay, I. M. *Listeria* faecal carriage by renal transplant recipients, haemodialysis patients and patients in general practice: Its relation to season, drug therapy, foreign travel, animal exposure and diet. *Epidemiol. Infect.* **106**, 157–166 (1991).
- Grif, K. et al. Prevalence and characterization of *Listeria monocytogenes* in the feces of healthy Austrians. *Wien. Klin. Wochenschr.* **113**, 737–742 (2001).
- Grif, K., Patscheider, G., Dierich, M. P. & Allerberger, F. Incidence of fecal carriage of *Listeria monocytogenes* in three healthy volunteers: a one-year prospective stool survey. *Eur. J. Clin. Microbiol. Infect. Dis.* **22**, 16–20 (2003).
- Schuchat, A. et al. Gastrointestinal carriage of *Listeria monocytogenes* in household contacts of patients with listeriosis. *J. Infect. Dis.* **167**, 1261–1262 (1993).
- Sauers, B. D. et al. Low prevalence of *Listeria monocytogenes* in human stool. *J. Food Prot.* **68**, 178–181 (2005).
- Le Monnier, A., Abachin, E., Beretti, J. L., Berche, P. & Kayal, S. Diagnosis of *Listeria monocytogenes* meningoencephalitis by real-time PCR for the hly gene. *J. Clin. Microbiol.* **49**, 3917–3923 (2011).
- Ivanek, R., Gröhn, Y. T. & Wiedmann, M. *Listeria monocytogenes* in multiple habitats and host populations: review of available data for mathematical modeling. *Foodborne Pathog. Dis.* **3**, 319–336 (2006).
- Nightingale, K. K. et al. Ecology and transmission of *Listeria monocytogenes* infecting ruminants and in the farm environment. *Appl. Environ. Microbiol.* **70**, 4458–4467 (2004).
- Linke, K. et al. Reservoirs of *Listeria* species in three environmental ecosystems. *Appl. Environ. Microbiol.* **80**, 5583–5592 (2014).
- Rohmer, L., Hocquet, D. & Miller, S. I. Are pathogenic bacteria just looking for food? Metabolism and microbial pathogenesis. *Trends Microbiol.* **19**, 341–348 (2011). vol.
- Culina, A., Crowther, T. W., Ramakers, J. J. C., Gienapp, P. & Visser, M. E. How to do meta-analysis of open datasets. *Nat. Ecol. Evol.* **2**, 1053–1056 (2018). vol.
- Gurevitch, J., Koricheva, J., Nakagawa, S. & Stewart, G. Meta-analysis and the science of research synthesis. *Nature* **555**, 175–182 (2018). vol.
- Meyer, F. et al. The metagenomics RAST server - A public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinforma.* **9**, 386 (2008).
- Guillet, C. et al. Human listeriosis caused by *Listeria ivanovii*. *Emerg. Infect. Dis.* **16**, 136–138 (2010).
- Segata, N. On the road to strain-resolved comparative metagenomics. *mSystems* **3**, e00190-17 (2018).
- Tessler, M. et al. Large-scale differences in microbial biodiversity discovery between 16S amplicon and shotgun sequencing. *Sci. Rep.* **7**, 6589 (2017).
- Ramage, C. P., Low, J. C., McLauchlin, J. & Donachie, W. Characterisation of *Listeria ivanovii* isolates from the UK using pulsed-field gel electrophoresis. *FEMS Microbiol. Lett.* **170**, 349–353 (2006).
- Kanuganti, S. R., Wesley, I. V., Reddy, P. G., McKean, J. & Hurd, H. S. Detection of *Listeria monocytogenes* in pigs and pork. *J. Food Prot.* **65**, 1470–1474 (2002).
- Yokoyama, E., Saitoh, T., Maruyama, S. & Katsube, Y. The marked increase of *Listeria monocytogenes* isolation from contents of swine ceum. *Comp. Immunol. Microbiol. Infect. Dis.* **28**, 259–268 (2005).
- Wacheck, S., Fredriksson-Ahomaa, M., König, M., Stolle, A. & Stephan, R. Wild boars as an important reservoir for foodborne pathogens. *Foodborne Pathog. Dis.* **7**, 307–312 (2010).
- Zundel, E. & Bernard, S. *Listeria monocytogenes* translocates throughout the digestive tract in asymptomatic sheep. *J. Med. Microbiol.* **55**, 1717–1723 (2006).
- Fredriksson-Ahomaa, M., Gerhardt, M. & Stolle, A. High bacterial contamination of pig tonsils at slaughter. *Meat Sci.* **83**, 334–336 (2009).
- Aureli, P. et al. An outbreak of febrile gastroenteritis associated with corn contaminated by *Listeria monocytogenes*. *N. Engl. J. Med.* **342**, 1236–1241 (2000).
- Dalton, C. B. et al. An outbreak of gastroenteritis and fever due to *Listeria monocytogenes* in milk. *N. Engl. J. Med.* **336**, 100–106 (1997).

43. Buffie, C. G. & Pamer, E. G. Microbiota-mediated colonization resistance against intestinal pathogens. *Nat. Rev. Immunol.* **13**, 790–801 (2013). vol.
44. Becattini, S. et al. Commensal microbes provide first line defense against *Listeria monocytogenes* infection. *J. Exp. Med.* **214**, 1973–1989 (2017).
45. Rohlion, N. et al. A *Listeria monocytogenes* bacteriocin can target the commensal *Prevotella copri* and modulate intestinal infection. *Cell Host Microbe* **26**, 691–701 (2019). e5.
46. Quereda, J. J. et al. Bacteriocin from epidemic *Listeria* strains alters the host intestinal microbiota to favor infection. *Proc. Natl Acad. Sci. USA* **113**, 5706–5711 (2016).
47. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. Microbiome datasets are compositional: and this is not optional. *Front. Microbiol.* **8**, 2224 (2017). vol.
48. Heras, V. L. et al. Short-term consumption of a high-fat diet increases host susceptibility to *Listeria monocytogenes* infection. *Microbiome* **7**, 7 (2019).
49. Rea, M. C. et al. Clostridium difficile carriage in elderly subjects and associated changes in the intestinal microbiota. *J. Clin. Microbiol.* **50**, 867–875 (2012).
50. Ferreira, R. B. R. et al. The intestinal microbiota plays a role in salmonella-induced colitis independent of pathogen colonization. *PLoS ONE* **6**, e20338 (2011).
51. Ubeda, C. et al. Intestinal microbiota containing *Barnesiella* species cures vancomycin-resistant *Enterococcus faecium* colonization. *Infect. Immun.* **81**, 965–973 (2013).
52. Schubert, A. M., Sinani, H. & Schloss, P. D. Antibiotic-induced alterations of the murine gut microbiota and subsequent effects on colonization resistance against *Clostridium difficile*. *MBio* **6**, e00974 (2015).
53. Vincent, C. et al. Reductions in intestinal Clostridiales precede the development of nosocomial *Clostridium difficile* infection. *Microbiome* **1**, 1–11 (2013).
54. Singh, P. et al. Intestinal microbial communities associated with acute enteric infections and disease recovery. *Microbiome* **3**, 45 (2015).
55. Maury, M. M. et al. Uncovering *Listeria monocytogenes* hypervirulence by harnessing its biodiversity. *Nat. Genet.* **48**, 308–313 (2016).
56. Neff, E. P. Littermate wanted: standardizing mouse gut microbiota requires more than cohousing. *Lab Anim.* **48**, 197 (2019). vol.
57. Zarrinpar, A. et al. Antibiotic-induced microbiome depletion alters metabolic homeostasis by affecting gut signaling and colonic metabolism. *Nat. Commun.* **9**, 1–13 (2018).
58. de Graaf, M. et al. Sustained fecal-oral human-to-human transmission following a zoonotic event. *Curr. Opin. Virol.* **22**, 1–6 (2017). vol.
59. Papkou, A. et al. The genomic basis of red queen dynamics during rapid reciprocal host–pathogen coevolution. *Proc. Natl Acad. Sci. USA* **116**, 923–928 (2019).
60. Moura, A. et al. Emergence and global spread of *Listeria monocytogenes* main clinical clonal complex. *Sci. Adv.* (in press, 2021).
61. Travier, L. et al. ActA Promotes *Listeria monocytogenes* aggregation, intestinal colonization and carriage. *PLoS Pathog.* **9**, e1003131 (2013).
62. Schmid, M. W. et al. Evolutionary history of the genus *Listeria* and its virulence genes. *Syst. Appl. Microbiol.* **28**, 1–18 (2005).
63. Garcia-Garcera, M. & Rocha, E. P. C. Community diversity and habitat structure shape the repertoire of extracellular proteins in bacteria. *Nat. Commun.* **11**, 758 (2020).
64. Nawrocki, E. *Structural RNA Homology Search and Alignment using Covariance Models*. Ph.D. thesis Washington Univ. (2009).
65. Katoh, K., Misawa, K., Kuma, K. I. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
66. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
67. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
68. Benson, D. A. et al. GenBank. *Nucleic Acids Res.* **41**, D36–D42 (2013).
69. Nguyen, L. T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
70. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
71. Sauders, B. D. et al. Diversity of *Listeria* species in urban and natural environments. *Appl. Environ. Microbiol.* **78**, 4420–4433 (2012).
72. Cole, J. R. et al. The ribosomal database project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res.* **31**, 442–443 (2003).
73. Garcia-Garcera, M., Touchon, M., Brisse, S. & Rocha, E. P. C. Metagenomic assessment of the interplay between the environment and the genetic diversification of *Acinetobacter*. *Environ. Microbiol.* **19**, 5010–5024 (2017).
74. Ragon, M. et al. A new perspective on *Listeria monocytogenes* evolution. *PLoS Pathog.* **4**, e1000146 (2008).
75. Pichon, M. et al. Diagnostic accuracy of a noninvasive test for detection of *Helicobacter pylori* and resistance to clarithromycin in stool by the AmpliDiag H. Pylori+clarir real-time PCR assay. *J. Clin. Microbiol.* **58**, e01787–19 (2020).
76. Madeira, F. et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).
77. Disson, O. et al. Modeling human listeriosis in natural and genetically engineered animals. *Nat. Protoc.* **4**, 799–810 (2009).
78. Albanese, D., Fontana, P., De Filippo, C., Cavalieri, D. & Donati, C. MICCA: a complete and accurate software for taxonomic profiling of metagenomic data. *Sci. Rep.* **5**, 9743 (2015).
79. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
80. Edgar, R. & Flyvbjerg H. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* **31**, 3476–3482 (2015).
81. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **2016**, e2584 (2016).
82. McMurdie, P. J. & Holmes, S. Shiny-phyloseq: Web application for interactive microbiome analysis with provenance tracking. *Bioinformatics* **31**, 282–283 (2015).
83. Lahti, L. & Shetty, S. Tools for microbiome analysis in R. Version 1.9.19. (2017).
84. Chao, A. & Lee, S. M. Estimating the number of classes via sample coverage. *J. Am. Stat. Assoc.* **87**, 210–217 (1992).
85. Hill, M. O. Diversity and evenness: a unifying notation and its consequences. *Ecology* **54**, 427–432 (1973).
86. Bray, J. R. & Curtis, J. T. An ordination of the upland forest communities of Southern Wisconsin. *Ecol. Monogr.* **27**, 325–349 (1957).
87. Oksanen, J. et al. vegan: community ecology package. (2019).
88. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).

## Acknowledgements

We thank Hélène Bracq-Dieye for technical help, Georges Michel Haustant, Cédric Fund, Elodie Turc, Laure Lemée, Biomics Platform, C2RT, Institut Pasteur, Paris, France, supported by France Génomique (ANR-10-INBS-09-09) and IBISA for 16S sequencing, Auguste Fourné and Amandine Brunet for PCR assays and Sandrine Isaac, Henrik Salje and Olivier Disson for critical reading. L.H. is supported by the Pasteur-Paris University (PPU) International Ph.D. Programme, funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665807, the 'Ecole Doctorale FIRE-Programme Bettencourt' of the CRI Paris and the Fondation Pasteur Suisse. M.L. laboratory is funded by Institut Pasteur, Inserm, the European Research Council and Laboratoire d'Excellence Integrative Biology of Emerging Infectious Diseases. M.L. is a member of Institut Universitaire de France.

## Author contributions

M.L. initiated and coordinated the project. L.H., A.M., M.G.-G. and M.L. designed the study. M.G.-G. collected and analyzed the public metagenomic data. C.B. and M.P. assessed the prevalence of *Lm* in stool donors' cohorts. L.H. conducted the in vivo experiments in mice, together with S.H.A.N. L.H. conducted the mouse 16S rRNA analysis. L.H. and M.L. wrote the manuscript, M.G.-G. and A.M. commented and edited on it.

## Competing interests

The authors declare no competing interests.

## Ethical statement

Animal experiments were performed according to the Institut Pasteur guidelines for laboratory animals' husbandry and in compliance with European regulation 2010/63 EU. All procedures were approved by the Animal Ethics Committee of Institut Pasteur, authorised by the French Ministry of Research and registered under #14644-2018041116183944. The stool donor cohorts received ethical approval from the regional Committee for the Protection of People (CPP Ouest III) and from the National Commission for Protection of Personal data on 23 November 2015. All patients were informed before inclusion and their consent was obtained before analysis.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-27069-y>.

**Correspondence** and requests for materials should be addressed to Marc Lecuit.

**Peer review information** *Nature Communications* thanks Henk Den Bakker, Sophia Kathariou and the other, anonymous, reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021