nature microbiology

Article

https://doi.org/10.1038/s41564-024-01859-8

Differential stress responsiveness determines intraspecies virulence heterogeneity and host adaptation in *Listeria monocytogenes*

Received: 21 March 2024

Accepted: 14 October 2024

Published online: 22 November 2024

Check for updates

Lukas Hafner[1], Enzo Gadin[1], Lei Huang[1], Arthur Frouin[2], Fabien Laporte[2], Charlotte Gaultier[1], Afonso Vieira[1], Claire Maudet[1], Hugo Varet[3], Alexandra Moura[1,4], Hélène Bracq-Dieye[4], Nathalie Tessaud-Rita[4], Mylène Maury[1,4], Melody Dazas[1], Rachel Legendre[3], Pauline Gastineau[1], Yu-Huan Tsai[1], Jean-Yves Coppée[5], Caroline Charlier[1,4,6], Etienne Patin[7], Rayan Chikhi[8], Eduardo P. C. Rocha[9], Alexandre Leclercq[4], Olivier Disson[1], Hugues Aschard[2] & Marc Lecuit[1,4,6]

Microbial pathogenesis is mediated by the expression of virulence genes. However, as microbes with identical virulence gene content can differ in their pathogenic potential, other virulence determinants must be involved. Here, by combining comparative genomics and transcriptomics of a large collection of isolates of the model pathogen *Listeria monocytogenes*, time-lapse microscopy, in vitro evolution and in vivo experiments, we show that the individual stress responsiveness of *L. monocytogenes* isolates determines their respective levels of virulence in vivo and reflects their degree of host adaptation. The transcriptional signature that accounts for the heterogeneity in the virulence of *L. monocytogenes* species is mediated by the stress response regulator SigB and driven by differential stress responsiveness. The tuning of SigB pathway responsiveness is polygenic and influenced by multiple, individually rare gene variations. This study reveals an overarching determinant of microbial virulence, challenging the paradigm of accessory virulence gene content as the major determinant of intraspecies virulence heterogeneity.

Bacterial pathogenesis is mediated by the expression of virulence factors, most of which are encoded by core genes and are therefore shared by all isolates of a given species[1–3]. However, virulence can vary within a species[4–11]. Accessory genes are a source of virulence heterogeneity, by modulating the pathogenic potential of a given clade[4,7–9,12,13]. However, microbes with identical virulence genes can differ in virulence[4,14], suggesting the involvement of other determinants of intraspecies virulence heterogeneity.

Listeria monocytogenes is an environmental bacterial pathogen that alternates between saprophytic and host-adapted lifestyle[15–17]. It is a major foodborne human and zoonotic pathogen[18–20] able to cross the intestinal barrier, disseminate systemically and be shed back in the intestinal lumen and faeces[21,22]. It can cross the blood–brain barrier, resulting in meningoencephalitis and the placental barrier, leading to fetal and neonatal infection[19,20]. The genus *Listeria* includes pathogenic and non-pathogenic species[23], and while

A full list of affiliations appears at the end of the paper. e-mail: marc.lecuit@pasteur.fr

listeriosis is rare, *Listeria* asymptomatic faecal carriage is common and associated with pathogenic species[15]. *L. monocytogenes* is a highly clonal but heterogeneous species[24]: Based on studies conducted in France and other Western countries[3,4,25–27], CC1, CC4 and CC6 are the clonal complexes (CCs) most associated with a clinical origin, while CC9 and CC121 are the most common among food isolates and very rarely cause listeriosis[4]. Clinically associated CCs are overrepresented in severe listeriosis cases, and isolates belonging to these CCs are more virulent in vivo[28] than those from food-associated CCs[4]. While previous studies have evidenced this *L. monocytogenes* intraspecies virulence heterogeneity[4,29,30], its extent, biological relevance and underlying mechanisms are unknown.

L. monocytogenes virulence genes are regulated by two key transcriptional regulators, SigB and PrfA. The stress response factor SigB controls a large regulon mainly containing stress response genes[31–35], the virulence genes *inlA* and *inlB* involved in host invasion and immune escape[36,37], and *prfA*. PrfA controls the transcription of the *Listeria* pathogenicity island-1 (LIPI-1), *inlA*, *inlB* and *prfA* itself[32,38,39]. The major *L. monocytogenes* virulence genes are part of its core genome, under purifying selection[3,21,30], yet loss-of-function mutations in these genes can be observed across the *L. monocytogenes* species[30]. However, variations in core virulence genes do not fully explain the heterogeneity of *L. monocytogenes* virulence[4,30]. In addition, while some lineage- or CC-specific accessory genes have been associated with virulence[4,12], there is no overarching mechanism that fully explains the virulence heterogeneity of *L. monocytogenes*.

L. monocytogenes virulence favours its faecal shedding[21,29,37], illustrating its link to *L. monocytogenes* host adaptation[15,17,29,40–42], while reduced virulence is associated with saprophytism[29]. In the life history of a bacterium, selective pressures on segregating populations can lead to phenotypic differences[14,43,44] and, in the long term, intraspecies heterogeneity and ultimately speciation. If *L. monocytogenes* host association is mediated by virulence genes and saprophytic growth selects for their loss, the recent life history of isolates could have influenced their virulence level, resulting not only in inter-clonal but also intra-clonal virulence heterogeneity. Given that *L. monocytogenes* gene content is highly similar and very homogeneous for isolates of the same CC, such intra-clonal virulence heterogeneity is unlikely to be driven by accessory virulence genes[3,40]. Harnessing a unique and comprehensive collection of *L. monocytogenes* isolates, we investigated the molecular mechanism(s) underlying inter- and intra-clonal virulence heterogeneity.

Results

Closely related isolates differ in virulence

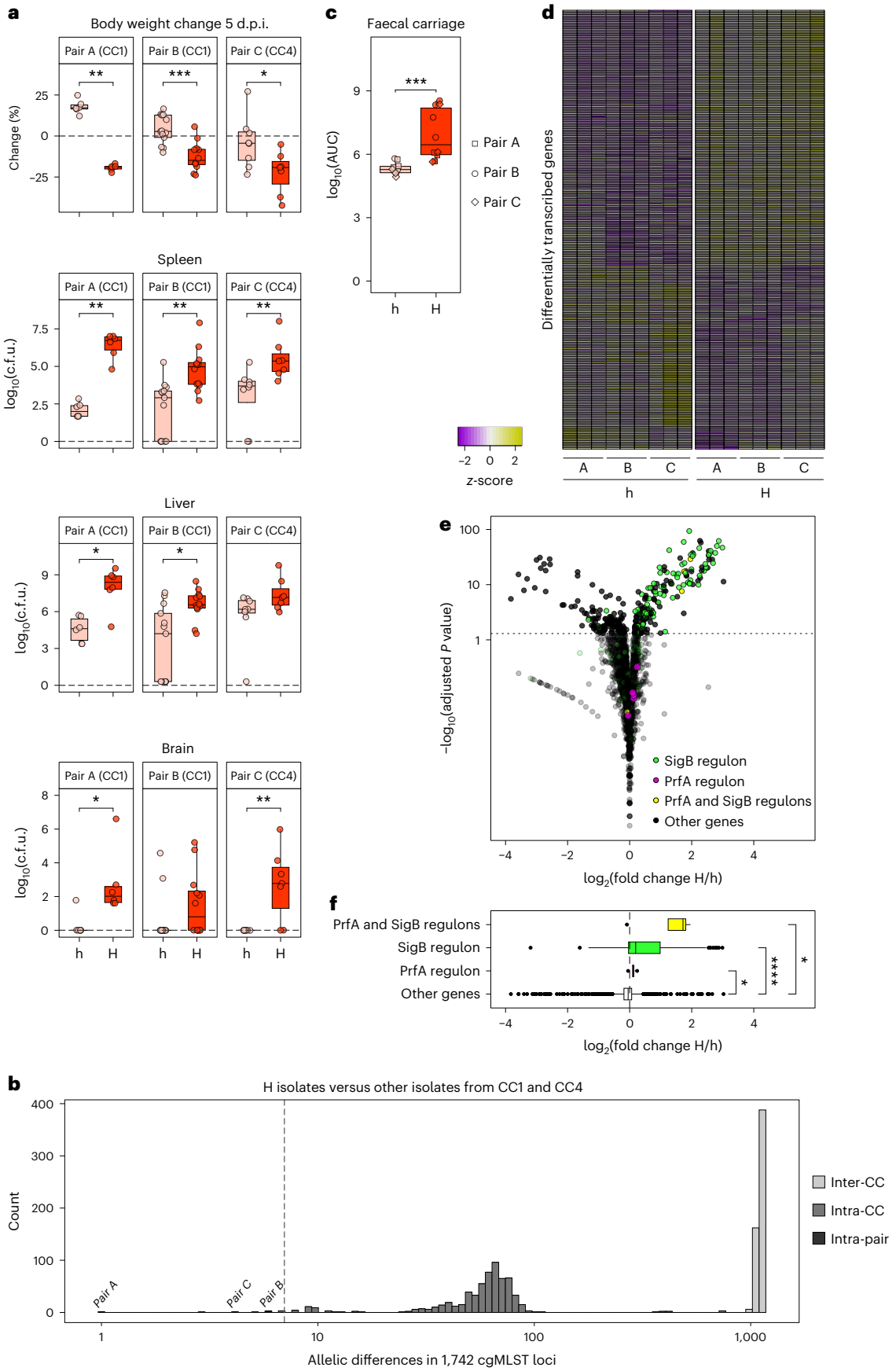
We showed previously that isolates from clinically associated CCs are more virulent than those from food-associated CCs[4] in a humanized mouse model of orally acquired listeriosis[28]. Virulence is also heterogeneous within clinically associated CCs, whereas isolates from food-associated CCs are mostly hypovirulent, as assessed by infection-induced weight loss (Extended Data Fig. 1a). To decipher the molecular mechanisms underlying intra-clonal virulence heterogeneity, we assessed the virulence of pairs of closely related isolates (the same cgMLST type, <7 allelic differences in 1,748 core genes[3]) and identified 3 pairs from clinically associated CCs (hereafter referred to as A, B and C) that showed a significant intra-pair virulence difference (Fig. 1a,b). For each pair, the more virulent isolate is hereafter designated H and the less virulent isolate h. H and h isolates do not differ in in vitro growth in rich medium (Extended Data Fig. 1b). At a sublethal oral inoculum, H and h differed in their capacity to colonize and persist in the mouse gut and to be transmitted back to the environment, as assessed by faecal shedding (Fig. 1c and Extended Data Fig. 1c,d). There are no common genetic differences between the respective H and h isolates across these H/h pairs (that is, the genetic differences between H and h of a given pair do not overlap with those of the other pairs), and all known *L. monocytogenes* virulence genes are conserved in these 3 H/h pairs (see Supplementary Table 1 for an exhaustive list of genetic differences).

Closely related isolates differ in SigB pathway activity

As no common genetic differences accounted for the differences in virulence between H and h isolates, we hypothesized that common transcriptional signatures might be associated with H and h isolates. We therefore performed an RNA sequencing (RNA-seq) experiment with the 3 H/h pairs during exponential growth in brain heart infusion (BHI) at 37 °C and identified 308 differentially regulated transcripts: 179 genes had higher transcript levels in H, and 129 genes had higher transcript levels in h (Fig. 1d,e and Supplementary Table 2). Over all H/h pairs, genes with higher transcript levels in H were highly enriched in SigB regulon genes (74/179, $\chi^2 = 335.11$, $P < 10^{-15}$), and SigB regulon genes had significantly higher transcript levels in H (Fig. 1f; $P < 10^{-15}$). As expected from these transcriptomic data, the InlA-dependent invasion of cultured cells expressing the InlA receptor human E-cadherin was significantly higher in H than in h isolates, as was the expression level of the SigB-dependent protein OpuCA (Extended Data Fig. 1e,f and Supplementary Information). To assess whether any other regulon or transcription factor could account for our finding, we searched the promoter regions of significantly upregulated genes in H for enriched DNA motifs (compared with all other promoter regions). We found that 90/179 genes with higher transcript levels in H contained a motif similar to the SigB binding site at the −10 box[45,46] ($E < 10^{-66}$), of which 52 genes overlapped were also part of our curated SigB regulon list (52/74, $\chi^2 = 16.401$, $P = 5.13 \times 10^{-5}$). Given that no other sequence was enriched in the 100 bp upstream sequence of these genes (Extended Data Fig. 1g), we concluded that the transcriptional difference is mainly driven by SigB (Fig. 1d–f).

Fig. 1 | SigB activity differs between closely related but differently virulent *L. monocytogenes* isolates. a, Body weight change and c.f.u. count in spleen, liver and brain 5 d.p.i. after oral inoculation with 2×10^8 c.f.u. from 3 cgMLST pairs of hyper- (H) and hypovirulent (h) isolates from CC1 and CC4 ($n = 6$ per isolate for pair A, $n = 13$ and $n = 12$ for pair B, $n = 8$ and $n = 7$ for pair C). *P* values for body weight change, spleen, liver and brain for pair A are 2.2×10^{-3}, 4.8×10^{-3}, 1.3×10^{-2} and 1.2×10^{-2}; for pair B are 1.9×10^{-4}, 1.2×10^{-3}, 1.4×10^{-2} and 1.1×10^{-1}; and for pair C are 4×10^{-2}, 6.5×10^{-3}, 7.2×10^{-1} and 7.2×10^{-3}, respectively. **b**, Relatedness of isolates estimated by allelic differences in cgMLST loci. H isolates were compared with the h isolates and other isolates belonging to CC1 and CC4. The comparisons within the three pairs are indicated. The dashed line indicates the cut-off of 7 allelic differences which defines a cgMLST type. **c**, Area under the curve of c.f.u. after oral inoculation with 2×10^7 c.f.u. of either h or H strains from days 1–7 post inoculation (Extended Data Fig. 1c,d), $P = 1.3 \times 10^{-4}$. **d**, Differentially transcribed genes between H and h in RNA-seq experiment ($n = 3$) of isolates from **b** in BHI in mid-exponential phase ($OD_{600} = 0.8$). Heat map shows replica-adjusted and variance-stabilized read counts normalized by *z*-score per significantly different gene. Each column represents a replicate ($n = 3$) in one isolate ($n = 6$), and each line represents a significantly different gene. **e**, Volcano plot showing fold change in transcripts between H and h isolates on the *x* axis and negative $\log_{10}$ of adjusted *P* values on the *y* axis. Dashed line shows significance threshold, and colour indicates genes belonging to SigB (green), PrfA (red) or SigB and PrfA (yellow) regulons. **f**, Box plot shows fold change between H and h isolates in genes belonging to SigB, PrfA or both regulons, compared with all other genes. *P* value for SigB regulon is 3.02×10^{-24}, for PrfA regulon is 3.2×10^{-2} and for genes which are in both regulons is 3.32×10^{-3}. For **a**, **b** and **f**, statistical comparison between groups was performed with two-sided Wilcoxon rank-sum test and per group against all other groups 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 box plots, the hinges represent the first and third quartiles of the distribution. The central line represents the median value. The whiskers extend from the hinge to the largest or smallest value no further than 1.5 × IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles).

https://doi.org/10.1038/s41564-024-01859-8



https://doi.org/10.1038/s41564-024-01859-8

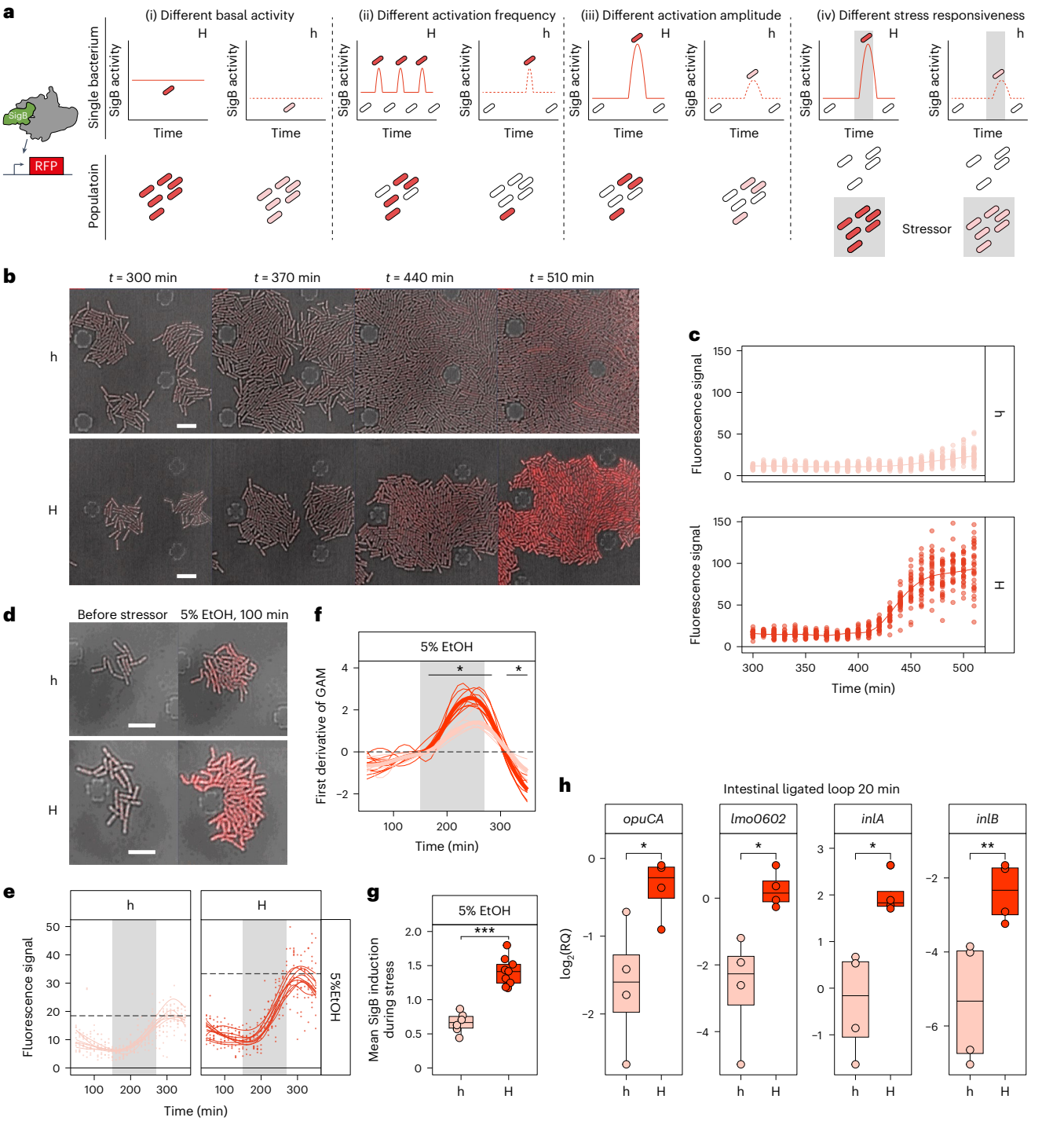


Fig. 2 | Hypervirulent SigB$_{high}$ isolates are more responsive to stressors. **a**, Schemes depicting how SigB activity could differ at the single-cell level accounting for the difference in bulk RNA with help of the SigB reporter system (left side). The top row shows signal at the single-cell level for each scenario, while the bottom row depicts the population level. RFP, red fluorescent protein. **b**,**c**, Time-lapse fluorescence microscopy of pair A H/h isolates with a SigB-dependent RFP reporter system, *pSigB::tdTomato* (white scale bars, 5 µM) (**b**) and its quantification in arbitrary units of fluorescence intensity (a.u.) (**c**). A subset of the points is shown. Each point represents a single cell in the field of view. The line shows a general additive model fitted to all points. **d**. The same as in **b** but before and after 100 min of presence of 5% EtOH. **e**, Quantification of several fields of view ($n = 9$ for H, $n = 6$ for h) as in **c**. The dashed line shows the mean value of fluorescence signal of the population mean per field of view per isolate at the end of stress induction. **f**, The first derivative of each general additive model (GAM) based on a.u. data from **e**; the horizontal line indicates where the two groups are significantly different ($P < 0.05$). **g**, The mean values of these a.u. derivatives over the time that the stressor was present as a readout for SigB responsiveness per field of view ($n = 9$ for H, and $n = 6$ for h). **h**, RT-qPCR of the same H and h isolates retrieved from intestinal ligated loops 20 min post inoculation. Relative quantity values are relative to housekeeping gene *gyrB* and have been $\log_2$-transformed. Each point corresponds to an independently infected ligated loop ($n = 4$ per isolate, for *opuCA*, *lmo0602*, *inlA* and *inlB*; $P = 4.3 \times 10^{-2}$, 1.6×10^{-2}, 1.2×10^{-2} and 6.7×10^{-3}, respectively). Statistical comparison between groups performed with two-sided Wilcoxon rank-sum test with Benjamini–Hochberg correction for multiple test. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. For box plots, the hinges represent the first and third quartiles of the distribution. The central line represents the median value. The whiskers extend from the hinge to the largest or smallest value no further than 1.5 × IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles).

Differential stress responsiveness mediates SigB pathway differential activity

In Gram-positive bacteria[32,47,48], SigB pathway activity is induced by environmental stressors through the stressosome signalling cascade[49]. The transcriptional profiling by RNA-seq was performed in mid-exponential phase (Fig. 1d–f), a condition which reflects the basal (that is, non-induced) activity of SigB in the reference strain EGDe[32]. In the model firmicute *Bacillus subtilis*, SigB pathway activity shows a phenotypic stochastic heterogeneity in the absence of stressor and under constant stress level[50] and is synchronous when stress levels change[51]. We hypothesized that the differences in SigB pathway activity we observed in bulk RNA could therefore correspond at the single-cell level to differences in either (i) basal activity, (ii) activation frequency, (iii) peak activity in activated bacteria or (iv) responsiveness to the presence of a stressor (Fig. 2a). We introduced a transcriptional reporter system P_{sigB}::*TdTomato* into paired H/h isolates (pair A) and followed its activity by time-lapse microscopy at the single-cell level. SigB activity was uniform at low bacterial density but increased more rapidly in H than h as a function of bacterial density (Fig. 2b,c). Because bacterial density and stationary phase are stressors and activate SigB in *L. monocytogenes*[32], the observation that SigB activity differs between H and h isolates at higher bacterial density suggests that this difference results from a differential stress responsiveness (hypothesis (iv) in Fig. 2a) and that the stress response of H strains is induced in late exponential phase, as in the RNA-seq experiment at optical density at 600 nm (OD_{600}) = 0.8. Upon exposure to 5% EtOH (Fig. 2d,e, grey background), a stressor known to activate the general stress response in *L. monocytogenes* and *B. subtilis*[51,52], SigB-dependent transcript levels increased more rapidly and to higher levels in H than in h (Fig. 2e–g, $P = 4.0 \times 10^{-4}$), showing that the general stress response system in H is more responsive to stressors than in h. This differential stress responsiveness translates into differential SigB pathway activity between H and h (Fig. 1). We assessed SigB activity in the intestinal lumen of axenic mice by infecting ligated intestinal loops with either H or h isolates and collected the luminal bacteria after 20 min. Again, the transcript level of genes regulated by SigB was higher in H compared with h isolates (Fig. 2h), showing that the transcriptional state of these isolates also differs in vivo, in a condition which is relevant for pathogenesis.

Varying SigB pathway activity affects virulence

Having shown a correlation between differential stress responsiveness and virulence in H/h pairs, we aimed to investigate their causal relationship by genetically manipulating SigB pathway activity. We compared an H isolate with its isogenic mutants overexpressing either RsbV (which activates SigB pathway activity) or RsbW (which inhibits SigB pathway activity) (Fig. 3a,b)[47,49,53], as well as with its isogenic Δ*sigB* loss-of-function mutant. As expected, the transcription levels of *opuCA*, a gene which is only regulated by SigB[54], differed in all mutants compared with the wild-type parental strain, while *hly* transcript levels did not significantly change (Fig. 3c). This resulted in differential InlA-dependent invasion of cultured cells, as *inlA* is regulated by SigB[32] (Fig. 3d). Next, we orally inoculated humanized mice[28] and showed that SigB activity correlates with weight loss at 5 days post inoculation (d.p.i.) ($R^2 = 0.47$, $P < 8.1 \times 10^{-8}$) and c.f.u. counts in inner organs ($0.29 < R^2 < 0.48$, $P < 4.3 \times 10^{-7}$) (Fig. 3e–i), indicating that the relationship between SigB pathway activity and *L. monocytogenes* virulence level is indeed causal.

Differential stress responsiveness is genetically encoded

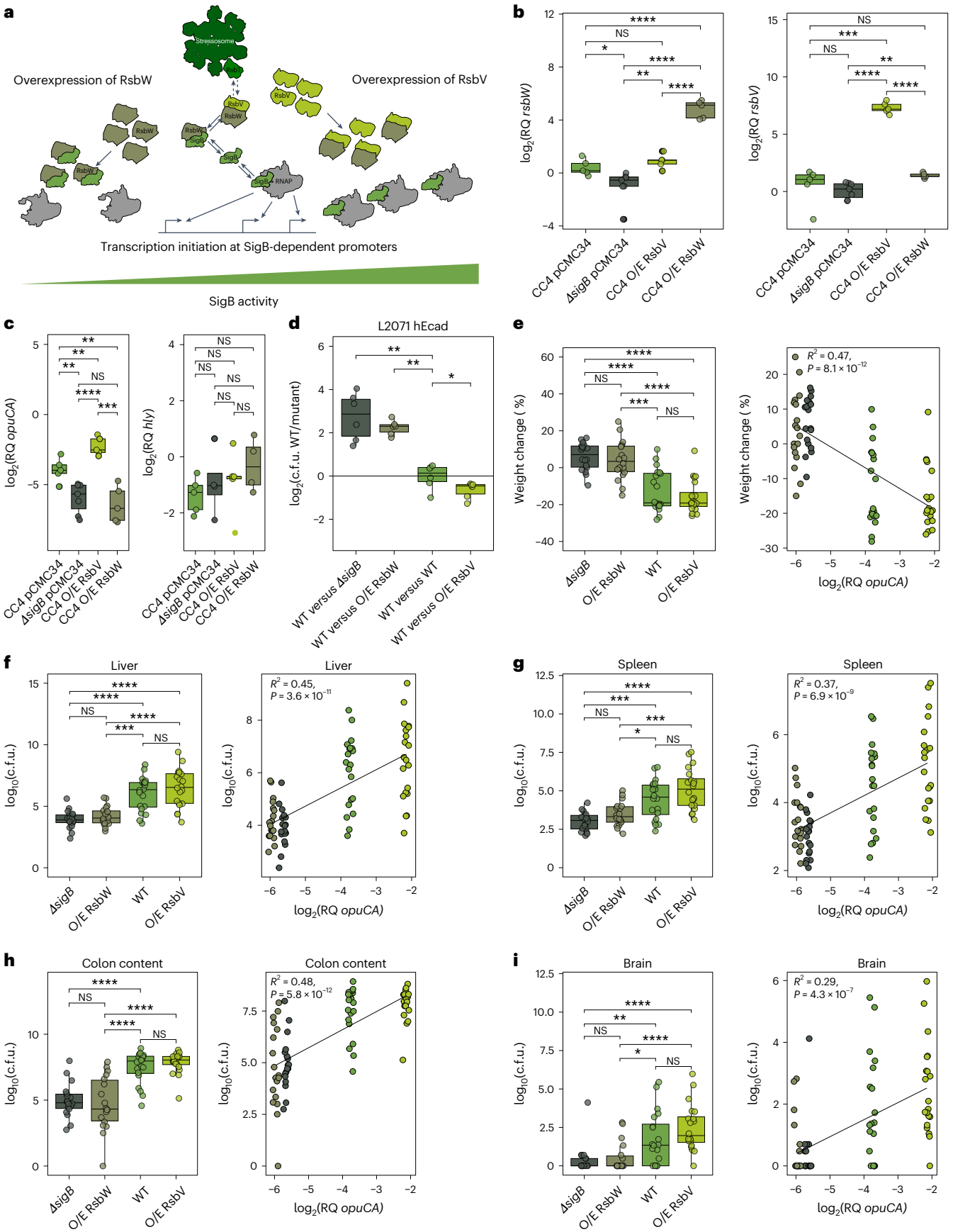
We next investigated how differential SigB pathway activity is determined in H/h pairs. As there were no common genetic difference between the three different H/h pairs, we selected in each pair the genetic difference which we anticipated to have the strongest impact on gene function (frameshift mutations or mutations in essential genes[55], exhaustive list of genetic differences in Supplementary Table 1; Extended Data Fig. 2a). We constructed deletion mutants for (1) *pssZ/lmo1913* (frameshift mutation in H of pair A at amino acid 193/334) which encodes for a putative glycosylhydrolase involved in exopolysaccaride synthesis[56], (2) a trans-complementation for a point mutation in the essential *murC/lmo1605* gene (Ala322Thr mutation in h of pair B) involved in cell wall synthesis[57] and (3) a trans-complementation for a frameshift mutation in *nnrD/lmo1622* (frameshift mutation in H of pair C at amino acid 249/276), an (S)-NAD(P)H-hydrate dehydratase predicted to be part of the stress response[34] (Extended Data Fig. 2b). These genetic modifications indeed influenced SigB pathway activity accordingly, as assessed by transcript levels of four SigB-dependent transcripts (Extended Data Fig. 2b), functionally showing that *pssZ*, *murC* and *nnrD* play a role in modulating SigB pathway activity and account for the difference in SigB pathway activity in the three H/h pairs. As these genetic modifications almost fully explain the transcriptional difference in SigB-dependent genes between H and h strains in these pairs, it is unlikely that epigenetics or other, non-genetic mechanisms play a major role in SigB-dependent virulence heterogeneity[58,59].

Clinically associated CCs are enriched in high SigB activity isolates

Having demonstrated a causal relationship between virulence (Fig. 1) and differential stress responsiveness (Fig. 2) in H/h pairs, we next investigated the relevance of these findings at the species level in *L. monocytogenes* to explain both inter- and intra-clonal virulence heterogeneity. Because virulence levels vary between CCs[4], we hypothesized that the relative proportions of isolates with either low or high SigB responsiveness between CCs may account for their respective clinical association. We therefore evaluated the heterogeneity of SigB pathway activity in a large set of isolates representative of *L. monocytogenes* phylogeny and compared this with that of PrfA, another major transcriptional regulator of *L. monocytogenes*[32] that was not involved in the differential transcriptional signatures identified in the H/h pairs[29]. As a readout of SigB pathway activity, we evaluated the transcript levels of *opuCA* and *lmo0602* from exponentially growing cultures by fluorescent barcode hybridization (Extended Data Fig. 3a–d), both of which are exclusively regulated by SigB[54] and are indeed significantly different between H and h (Extended Data Fig. 3e). We next applied this approach to the *L. monocytogenes* isolate collection of the *Listeria* National Reference Center in France, which includes all food and clinical isolates prospectively

Fig. 3 | Manipulating SigB activity affects virulence. a, Graphical depiction of SigB cascade, adapted from previous work[49], and layout of the used molecular titration system to manipulate SigB activity by overexpressing either RsbV or RsbW. **b**, RT-qPCR of *rsbV* and *rsbW* to confirm overexpression of these factors. For WT and *ΔsigB*, strains carried a genomic insertion of a red fluorescent protein under the same promoter as *rsbV* and *rsbW* overexpression to make up for potential fitness loss. NS, not significant. **c**, RT-qPCR of *opuCA* and *hly*. **d**, Competition assay for InlA-dependent cell entry relative to the WT strain ($n = 6$ for each competition index, $P = 2.2 \times 10^{-3}$, 2.2×10^{-3} and 4.1×10^{-2}). **e**–**i**, Percentage body weight change (**e**) and c.f.u. count in liver (**f**), spleen (**g**), colon content (**h**) and brain (**i**) 5 d.p.i. after oral inoculation with 2×10^8 c.f.u. ($n = 18$ for each derived strain, $n = 20$ for the WT). Relative quantity values for RT-qPCR were $\log_2$-transformed. *P* values and R^2 shown for linear correlations. Each point corresponds to a biological replication, either a single infected mouse or cell culture well, or bacterial culture for RT-qPCR experiments. For box plots, the hinges represent the first and third quartiles of the distribution. The central line represents the median value. The whiskers extend from the hinge to the largest or smallest value no further than 1.5 × IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles). For precise *P* values for in vivo comparisons, see Supplementary Information. For qPCR results, unpaired, two-sided *t*-tests were applied. For all other comparisons, Wilcoxon rank-sum test was applied.

https://doi.org/10.1038/s41564-024-01859-8



https://doi.org/10.1038/s41564-024-01859-8

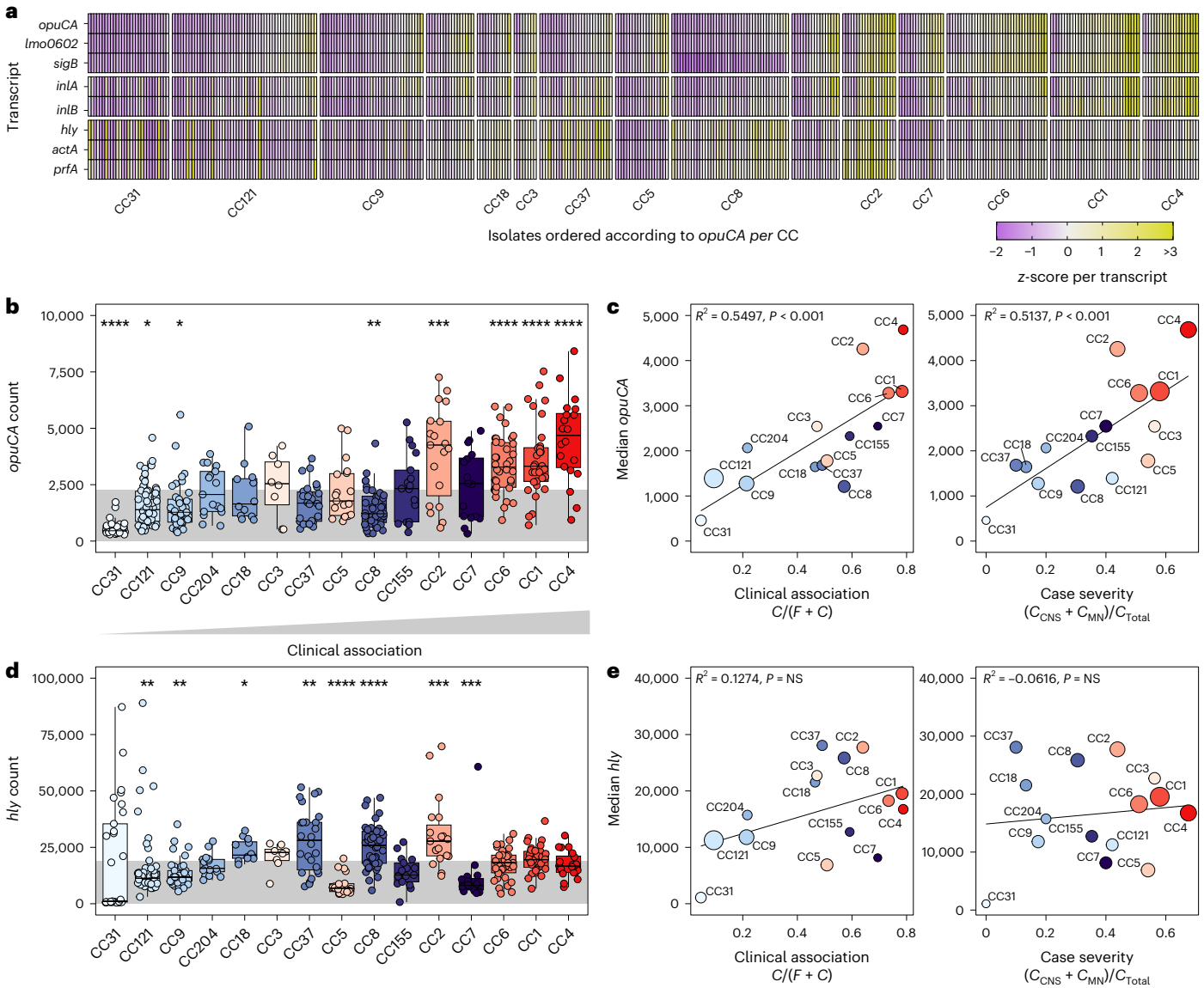


Fig. 4 | SigB activity heterogeneity accounts for inter- and intra-clonal virulence heterogeneity. a, Heat map showing the transcript levels, normalized per transcript, of SigB-dependent genes (top), PrfA-dependent genes (bottom) and co-regulated genes (centre) in 372 food isolates representative of the *L. monocytogenes* species. CCs are ordered by clinical association. **b**, SigB pathway activity measured by *opuCA* transcript level in food isolates from 15 most prevalent CCs. CCs are ordered according to their increasing clinical association from left to right. The grey area represents the mean value across all isolates. *P* values from left to right, 7.2×10^{-13}, 1.6×10^{-2}, 1.13×10^{-2}, 0.64, 0.92, 0.61, 0.19, 0.76407, 2.57×10^{-3}, 0.62, 5.7×10^{-4}, 0.54, 4.5×10^{-6}, 6.6×10^{-6} and 1.7×10^{-6}. **c**, Correlation of the median value of *opuCA* transcript counts of each CC with the clinical association (frequency of clinical samples (C = clinical isolates) in total sampling ($C + F$ = clinical and food isolates), $R^2 = 0.5497$, $P = 9.41 \times 10^{-4}$) and case severity (frequency of CNS and materno-neonatal (MN) cases in all cases (C_{Total}), $R^2 = 0.5137$, $P = 1.59 \times 10^{-3}$). **d**, The same as in **b** but for PrfA activity evaluated by *hly* transcript level. *P* values from left to right, 0.11, 2.60×10^{-3}, 5.1×10^{-3}, 0.89, 1.10×10^{-2}, 0.19, 1.08×10^{-3}, 3.0×10^{-6}, 1.5×10^{-5}, 0.15, 1.1×10^{-4}, 3.1×10^{-4}, 0.65, 7.76×10^{-2} and 0.81. **e**, The same as in **c** but for PrfA activity evaluated by median value of *hly* transcript counts for each CC ($R^2 = 0.1274$, $P = 0.11$ and $R^2 = 0.06158$, $P = 0.67$, respectively). Statistical comparison with two-sided Wilcoxon rank-sum test *per* group against all other groups in **a**, with Benjamini–Hochberg correction for multiple test. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$. For box plots, the hinges represent the first and third quartiles of the distribution. The central line represents the median value. The whiskers extend from the hinge to the largest or smallest value no further than 1.5 × IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles).

collected for surveillance in France. Food isolates are representative of *L. monocytogenes* exposure of the general population, whereas clinical isolates responsible for invasive listeriosis are enriched for virulent isolates[4,29]. To capture the broad diversity of virulence in *L. monocytogenes*, we assessed SigB pathway activity in 382 food isolates from the 15 most prevalent CCs representative of over 80% of the food sampling in France (Extended Data Fig. 4a,b).

In these food isolates, we found an inter- and intra-clonal heterogeneity in both SigB-regulated and PrfA-regulated gene transcripts (Fig. 4a and Extended Data Fig. 5a). Isolates from clinically associated CCs (CC1, CC2, CC4, CC6)[4] showed the highest level of SigB-regulated transcripts, whereas isolates from food-associated CCs (CC9, CC121, CC31)[4] showed the lowest (Fig. 4b,c). Isolates with high PrfA-dependent transcripts were also enriched in certain CCs, such as CC37, CC8 and CC2, but unrelated to their clinical association (Fig. 4d,e). Moreover, the median transcript level per CC of SigB-dependent genes, in contrast to PrfA-dependent genes, correlated with the degree of their clinical association (Fig. 4c,e; $R^2 = 0.5497$, $P < 0.001$ for SigB-dependent gene *opuCA* versus $R^2 = 0.1274$, $P = 0.11$ for PrfA-dependent gene *hly*) and with their association with neurolisteriosis (central nervous system

(CNS)) and maternal–neonatal (MN) cases[4] (Fig. 4d,e; $R^2 = 0.5137$, $P < 0.001$ for SigB-dependent gene *opuCA* versus $R^2 = 0.0616$, $P = 0.67$ for PrfA-dependent gene *hly*). This is consistent with the fact that both clinical association and association with CNS and MN correlate with the relative enrichment of virulent isolates in a given CC[4,29].

As expected, the transcript levels of SigB-regulated genes *sigB*, *opuCA* and *lmo0602* are positively correlated with each other (Extended Data Fig. 5b), as are those of the PrfA-regulated genes *prfA*, *actA* and *hly* (Extended Data Fig. 5b). Consistently, *inlA* and *inlB* transcript levels correlate with both SigB- and PrfA-regulated genes, in line with their transcriptional regulation by both SigB and PrfA[32] (Extended Data Fig. 5b). While SigB- and PrfA-regulated genes do not show a significant correlation in isolates at the species level, they do correlate in isolates from the same CCs, as expected from SigB's role as a transcriptional regulator of *prfA* (Extended Data Fig. 5c–e)[32].

Patients with MN and CNS *L. monocytogenes* infection have, on average, fewer immunosuppressive comorbidities than those with septicaemia[19] and would therefore be expected to be infected by isolates of higher virulence, and thus higher SigB pathway activity, than patients with septicaemia[4]. Indeed, MN and CNS *L. monocytogenes* isolates from a large prospective cohort of patients with listeriosis (MONALISA, an ongoing prospective cohort of listeriosis patients in France)[19] showed higher SigB pathway activity than those from patients with septicaemia (Extended Data Fig. 6a,b; $P = 0.0232$ and $P = 0.0038$, respectively), an association consistent with the causal relationship between SigB pathway activity and *L. monocytogenes* virulence described above.

Differences in SigB pathway activity account for intra-clonal virulence heterogeneity

We next investigated experimentally the impact of SigB pathway activity on virulence. As the broadest heterogeneity in SigB pathway activity is observed in clinically associated CCs (Fig. 4a–e), we randomly selected two pairs of isolates per CC from CC1, CC4 and CC6 that belonged to either the top third or bottom third of SigB pathway activity of their respective CC (Fig. 4b and Extended Data Fig. 6c,d), hereafter referred to as $SigB_{high}$ and $SigB_{low}$ (Extended Data Fig. 7a,b). It is worth noting that these $SigB_{high}$ and $SigB_{low}$ isolates were not chosen based on their genetic similarity within CCs and differed both in gene content and sequence (for gene content, see Supplementary Table 3). To assess whether SigB pathway activity sets *L. monocytogenes* virulence level, we co-inoculated mice orally with both isolates of each pair and assessed their respective bacterial loads at 5 d.p.i. in liver and spleen (reflecting systemic infection), brain (reflecting severe infection) and colon content (reflecting transmission back to the environment)[21,22,29]. $SigB_{high}$ isolates consistently outcompeted $SigB_{low}$ isolates (Extended Data Fig. 7c,d; $P < 0.0001$ in all compartments), showing that $SigB_{high}$ isolates are more virulent in vivo and better transmitted back to the environment than $SigB_{low}$ isolates.

In vitro evolution leads to attenuation of SigB activity

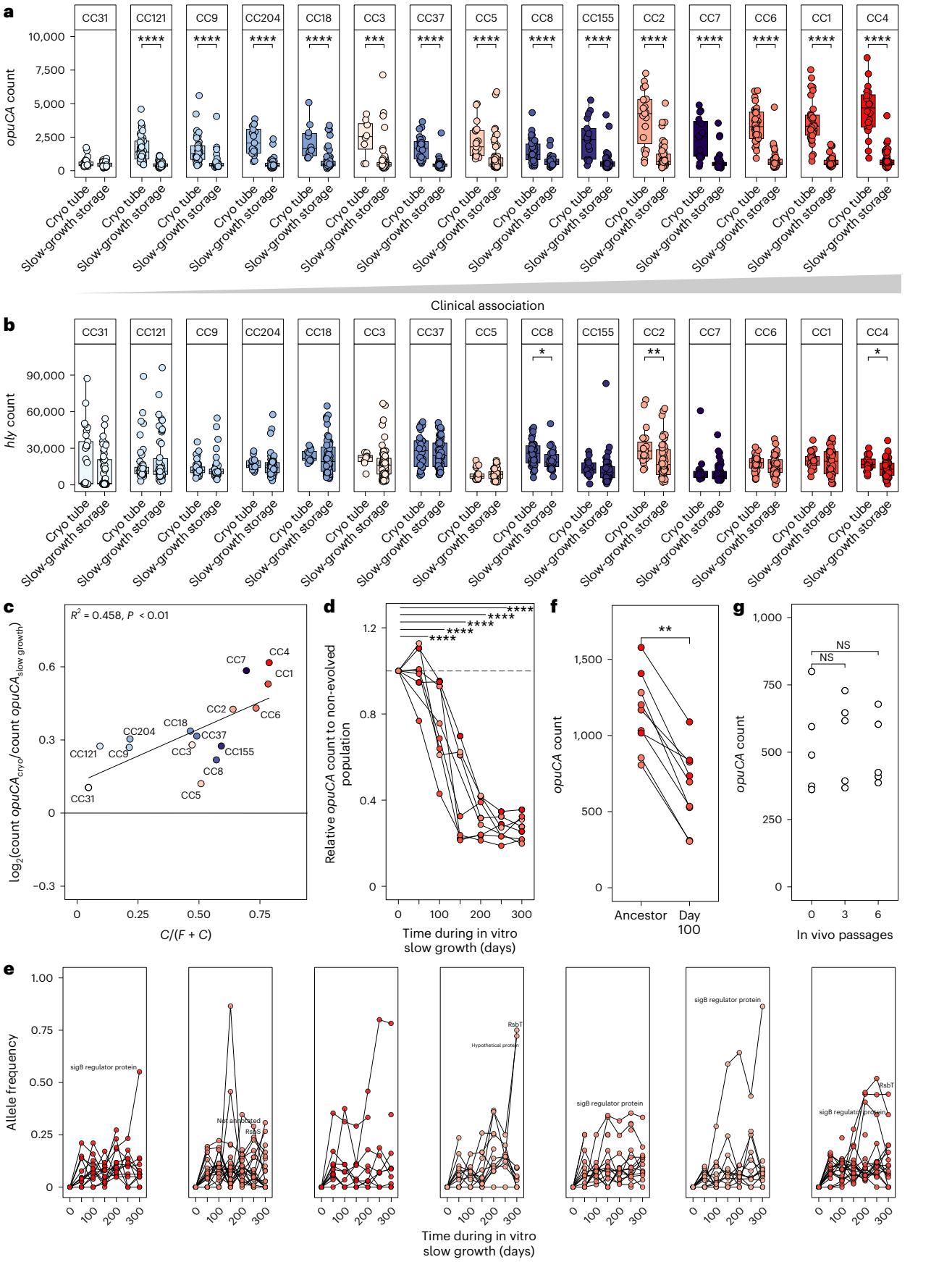
Virulence is linked to faecal carriage in *L. monocytogenes* and *Listeria*[12,15,29,41,42] (Fig. 1c and Extended Data Fig. 1c,d), and low virulence is associated with saprophytic growth in *L. monocytogenes*[4,29,30]. To study how saprophytic growth in a non-stressed condition affects SigB-dependent virulence, we used a collection of isolates stored for more than 2 years at 14 °C in slow-growth tubes, in contrast to the isolate collection stored at −80 °C used so far in this study (Fig. 4 for food isolate collection and Extended Data Fig. 5 for clinical isolate collection). Isolates conserved at 14 °C have a consistently significant lower SigB pathway activity than those stored at −80 °C, in contrast to PrfA regulon (Fig. 5a,b). The difference in mean SigB activity of a given CC between long-term in vitro growth at 14 °C and cryopreservation correlated with its clinical association ($R^2 = 0.4572$, $P = 0.00338$; Fig. 5c). This suggests that growth outside of the host leads to a loss of SigB pathway activity. Isolates from CCs which do not show a loss of SigB during in vitro growth, namely CC31, already have low SigB activity and likely underwent such an adaptation before sampling and storage (Figs. 4 and 5). To confirm that growth away from the host results in loss of SigB activity, we performed an experimental evolution assay with 7 $SigB_{high}$ isolates which were stored at −80 °C and belong to clinically associated CCs. We grew these isolates in vitro at 14 °C for 300 days, and this systematically led to a decrease in SigB pathway activity (Fig. 5d). In the evolved populations, mutations accumulated over time, both in *sigB* and in *rsb* genes that encode SigB regulators (Fig. 5e), which are expected to account at least in part for the resulting loss of SigB pathway activity, as they include missense and frameshift mutations substitutions (Supplementary File 1). To demonstrate that this effect is not limited to this specific growth condition, we performed a daily passage experiment in a rich medium that also resulted in a loss of SigB activity (Fig. 5f). This suggests that, under growth conditions where virulence is not expected to contribute to fitness, that is, away from the host, SigB pathway activity has a cost, leading to positive selection for the loss of SigB pathway activity and possibly to a relaxation of purifying selection on *L. monocytogenes* genes involved in SigB pathway activity. It is worth noting that we cannot exclude that under less stable conditions, for example, fluctuating stress levels, SigB activity would have been maintained. By contrast, an increase in SigB pathway activity and virulence after in vivo passages starting from a $SigB_{low}$ isolate is expected to be excessively rare, given the low mutation rate in *L. monocytogenes*[3] and that the limited number of bacteria which cross the intestinal barrier and disseminate internally[22] imposes a strong bottleneck[22,60]. Together with the moderate 5- to 10-fold fitness advantage per in vivo passage of $SigB_{high}$ over $SigB_{low}$ (Fig. 4d), this makes the emergence and fixation of such a mutant in vivo a very rare event unlikely to be obtained in an experimental setting. Accordingly, we observed no increase in SigB activity after up to six in vivo faecal to oral passages (Fig. 5g).

SigB-dependent stress responsiveness is a polygenic trait

In *L. monocytogenes*, the stressosome senses stress signals[47,49,53] and activates SigB through a partner switching cascade[48]. Genetically manipulating this cascade leads to changes in virulence (Fig. 3 and Extended Data Fig. 2). However, genes encoding this cascade are fully conserved and functional in H/h pairs (Supplementary Table 1) and are highly conserved within CCs and across the *L. monocytogenes*

Fig. 5 | In vitro growth leads to SigB activity attenuation. a, Comparison of transcript counts of *opuCA* per CC between isolates that have been stored in cryo tubes or slow-growth condition. *P* values from left to right, 0.13, 3.8×10^{-15}, 2.9×10^{-10}, 6.6×10^{-9}, 5.5×10^{-6}, 7.5×10^{-4}, 3.5×10^{-14}, 7.7×10^{-6}, 2.7×10^{-5}, 9.9×10^{-8}, 2.0×10^{-7}, 4.5×10^{-7}, 3.1×10^{-16}, 9.4×10^{-13} and 3.1×10^{-9}. **b**, Same as in **a** but for *hly*. *P* values from left to right, 0.4, 0.93, 0.23, 6.7×10^{-2}, 0.2, 6.1×10^{-2}, 0.5, 0.39, 2×10^{-2}, 0.12, 4.4×10^{-3}, 0.38, 0.18, 0.7 and 2.8×10^{-2}. **c**, Correlation of fold change between cryo and slow-growth condition isolates and clinical association per CC, suggesting that SigB activity was reduced the most in the isolates that come from the most clinically associated CCs. $P = 3.38 \times 10^{-3}$. **d**, SigB pathway activity of mixed populations after growth in vitro (*P* values from left to right, 0.19, 1.1×10^{-7}, 1.4×10^{-7}, 1.1×10^{-7}, 1.4×10^{-7}, 1.1×10^{-7}). **e**, allelic differences that were arising in these populations. Alleles with a frequency above 25% at 300 days of experimental evolution were annotated. Mutations in *sigB* or *rsb* genes which reached at least 20% at one point are described in Supplementary Data 1. **f**, Same as in **d** but for 9 populations after daily passages in BHI for 100 days. **g**, Same as in **d** but for in vivo passages from the faeces at 5 d.p.i. Statistical comparison between groups performed with two-sided Wilcoxon rank-sum test with Benjamini–Hochberg correction for multiple test. For experimental evolution, values were compared with the non-evolved isolates (at 0 days of in vitro evolution). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. For box plots, the hinges represent the first and third quartiles of the distribution. The central line represents the median value. The whiskers extend from the hinge to the largest or smallest value no further than 1.5 × IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles).

https://doi.org/10.1038/s41564-024-01859-8



https://doi.org/10.1038/s41564-024-01859-8

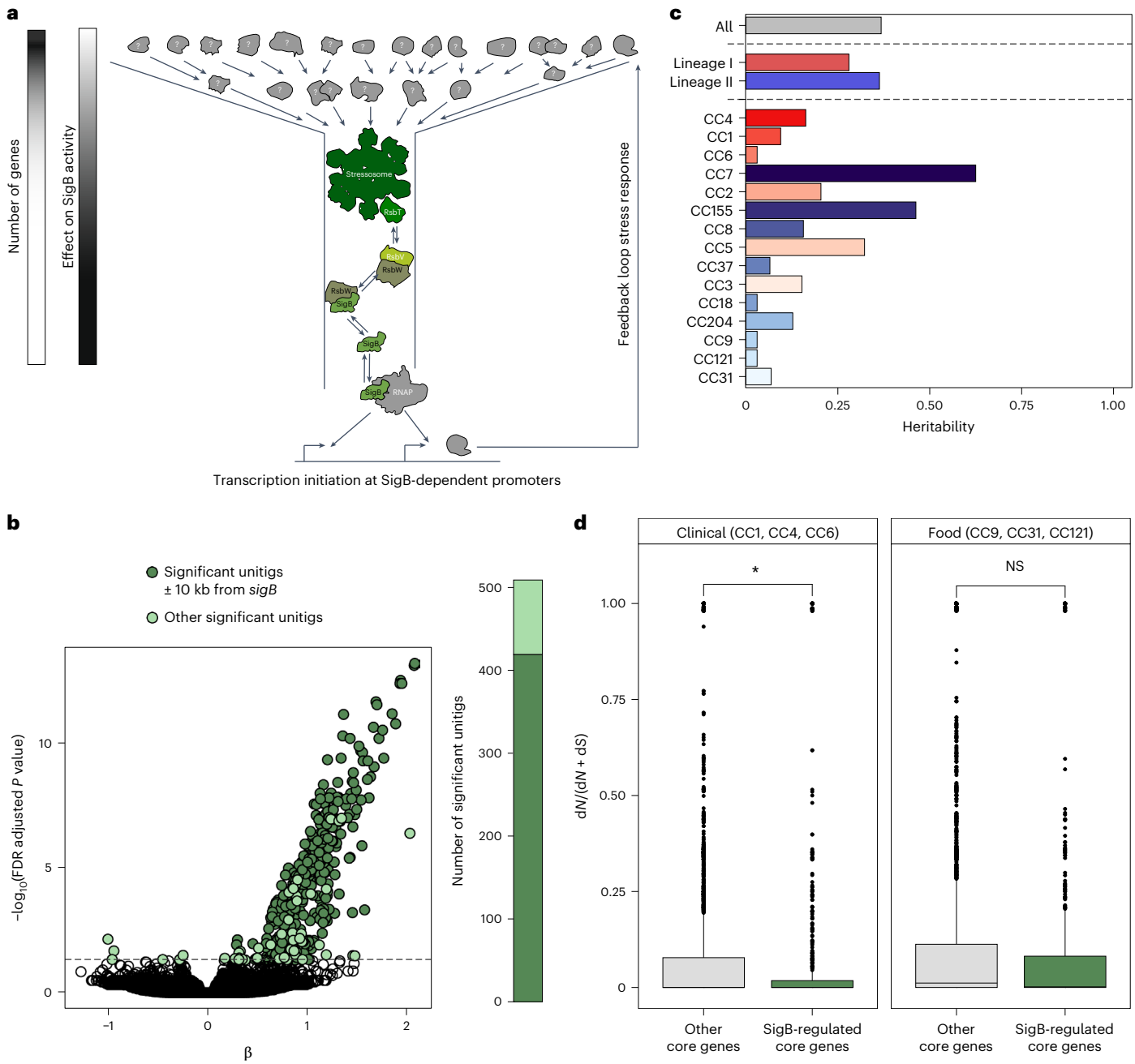


Fig. 6 | SigB activity heterogeneity is polygenic across *L. monocytogenes*. **a**, Schematic representation of stressosome and its upstream regulators. The stressosome cascade forms a molecular funnel for stress signals, and therefore mutations in the genes encoding parts of this cascade will have a strong effect on SigB activity, while most mutations upstream are expected to have weak effects on SigB pathway activity. **b**, Effect size on SigB activity and FDR adjusted *P* value per unitig estimated with ChoruMM[62]. The colour indicates whether a unitig can be mapped to the *sigB* operon or the regions in its proximity in the reference genome EGDe. Bar plot on the right shows the distribution of significant unitigs which could be mapped to EGDe genome (Methods) among categories. **c**, Heritability of SigB pathway activity per CC, lineage and for all of the *L. monocytogenes* species, estimated by different ChoruMM models[62]. **d**, $dN/(dN+dS)$ value for all comparisons between the most prevalent allele of each CC of a given group (Clinical corresponding to CC1, CC4 and CC6, Food corresponding to CC31, CC121, CC9) comparing SigB-regulated ($n = 191$) and other core genes ($n = 1{,}604$). $P = 0.01$ for clinical CCs. Two-sided Wilcoxon test to compare gene dN/dS between all core genes and SigB-regulated core genes. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. For box plots, the hinges represent the first and third quartiles of the distribution. The central line represents the median value. The whiskers extend from the hinge to the largest or smallest value no further than 1.5 × IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles).

species[3], in contrast to in vitro evolved isolates that are kept in a fixed non-host environment and eventually lose SigB-mediated regulation (Fig. 5). As in isolates which did not undergo in vitro experimental evolution, the stressosome cascade and *sigB* are generally conserved; the genetic variations underlying SigB-dependent differential virulence in these isolates are expected to be found in genes acting upstream of the stressosome, which serves as a molecular funnel for stress signals (Fig. 6a). While only few of these genes have been identified[61], they are expected to be numerous, given (1) the multiplicity of stress signals, (2) the integrative nature of the stressosome[47–49,53] and (3) the expected high number of SigB-regulated genes whose products also feedback on SigB pathway activity (Fig. 6a)[32]. Yet, the individual effect

of mutations in the various genes acting upstream of the stressosome is expected to be subtle compared to the drastic effect of mutations in the genes encoding SigB signalling cascade (Fig. 5d,e). To identify the genetic determinants of SigB pathway responsiveness, we performed a genome-wide association study (GWAS) on 1,597 *L. monocytogenes* genomes spanning its genetic diversity, using the SigB pathway activity value from our large transcriptional profiling as a continuous phenotypic variable (Fig. 6b; Methods)[62]. To associate phenotypic and genetic variations, we compacted all genomes into unitigs (that is, maximally long sequences in a pangenome de Bruijn graph[63]). Given the very clonal nature of *L. monocytogenes* population structure[3,24] (that is, long branches and little intra-clonal variation[64]), most unitigs in our dataset have a low minor allele frequency within a given clone, limiting the statistical power of the analysis (Extended Data Fig. 8). Consequently, only unitigs with a strong effect size, such as mutations in *sigB*, *rsb* genes or other putative unknown upstream regulators, are identifiable with this approach (Extended Data Fig. 8). Indeed, we mainly found significantly associated unitigs (Supplementary Table 4; Methods) in the *sigB* or *rsb* genes (Fig. 6b, right panel), which occurred mainly in 14 °C stored isolates and in our in vitro evolution experiment. We estimated that a dataset containing up to two orders of magnitude more isolates would be required to assess the significance of the association of unitigs with a moderate to low effect on SigB pathway activity (Extended Data Fig. 8a). While the GWAS detected little to no signal beyond *sigB* or *rsb* genes, it estimated a high heritability of the phenotype beyond those genes (Fig. 6c and Extended Data Fig. 8c), especially within CCs that show a broad range in SigB activity, such as CC7. This underscores that SigB-dependent stress responsiveness is a polygenic trait shaped by multiple individually rare and/or weak genetic factors, consistent with our observation that none of the H/h pairs (Fig. 1 and Supplementary Table 1) shared common genetic differences.

SigB regulon is under stronger purifying selection in clinically associated CCs

Having shown that the GWAS' statistical power is insufficient to identify rare single allelic variants associated with SigB pathway activity, we hypothesized that the common genetic signature of isolates with similar life histories and thus subjected to similar selective pressures might be more readily detectable by examining larger evolutionary distances, at the CC rather than isolate level, and by considering the SigB regulon as a whole rather than individual mutations. Thus, we probed the conservation of the predominant allelic variant per CC of SigB-regulated genes compared with that of other core genes between clinically and food-associated CCs, with the hypothesis that the SigB regulon is under greater purifying selection in clinically associated CCs. We found that SigB-regulated genes have fewer non-synonymous mutations compared to all mutations than other core genes in clinically associated CCs ($P = 0.01$) but not in food-associated CCs (Fig. 6d), suggesting a higher purifying selection on the general stress response in clinically associated CCs. Taken together, these results suggest that SigB pathway responsiveness is a polygenic trait shaped by the life history of individual isolates, which reflects their ecology, as highlighted by the selective pressure on SigB-regulated core genes in clinically associated CCs (Fig. 6d).

Discussion

The extent of intraspecies microbial virulence heterogeneity and its underlying mechanisms are only partially explained by differences in virulence gene content[4,65]. Here we discovered that differential stress responsiveness is an overarching determinant of *L. monocytogenes* virulence level and host adaptation at the isolate level, which drives the heterogeneity of *L. monocytogenes* virulence. The genetic determinants of stress responsiveness that control virulence are multiple and individually rare, consistent with the integrative nature of the stressosome and reflecting the evolutionary paths of individual *L. monocytogenes* isolates and their respective adaptation to the host[10,11,17,29,40,42,49,66]. It is worth noting that the existence of an overarching mechanism to fine-tune virulence does not exclude the contribution of accessory genes to virulence heterogeneity, as shown for *lls*, LIPI-4 and other accessory genes[4,12,13,67–69].

Virulence can impose important costs on environmental pathogens[11,32,65,70], but it can also benefit them by participating in their transmission[10,11,66,71]. Indeed, while virulence genes are under purifying selection in natural *L. monocytogenes* isolates[3,21], we demonstrate experimentally that prolonged slow growth of *L. monocytogenes* in non-stressful conditions leads to attenuation of SigB pathway activity, consistent with previous observations[72,73], and virulence. This suggests that CCs enriched in $SigB_{low}$ isolates are adapted to a lifestyle away from the host, as for CC9, CC121[4,29] and most strikingly CC31, which has the lowest SigB activity and is almost never responsible for clinical infection[4]. Conversely, clinically associated hypervirulent CCs such as CC1, CC4 and CC6 (ref. 4) are enriched in $SigB_{high}$ isolates, are associated with vertebrate hosts, particularly cattle[17,29,42], and have a higher capacity to colonize the gut[18,29,40,41]. At the intra-clonal level, the individual life histories of *L. monocytogenes* isolates and the conditions under which they have thrived result in different levels of stress responsiveness, which in turn shape the *L. monocytogenes* virulence landscape and host adaptation[19]. Consequently, the extent of loss of stress responsiveness of $SigB_{low}$ isolates belonging to clinically associated CCs may reflect the time during which they have evolved away from a host. It is worth noting that our study was limited to food and clinical isolates, and investigation of environmental and non-human host-associated *L. monocytogenes* isolates would allow to complement our current findings and potentially identify ecological contexts that favour either high or low SigB activity.

Environmental pathogens sense the host environment and express virulence genes only when they serve an ecological function, namely, microbial replication within the host and the resulting selection of the fittest for release into the environment[10,11,14,21,66,71,74,75], and otherwise restrict their expression to limit metabolic costs[11,32,65,70]. In *L. monocytogenes*, the general stress response system serves as a stress-sensing system that also controls *L. monocytogenes* virulence[32,47,48]. Indeed, SigB regulates the major virulence genes *inlA*, *inlB* and *prfA*[32,33]. The importance of SigB responsiveness in regulating *L. monocytogenes* virulence is highlighted by our recent finding that higher levels of InlB expression result in increased systemic spread during infection[37], and a similar quantitative effect occurs for higher levels of InlA expression, likely resulting in the higher bacterial loads upon crossing the intestinal barrier of $SigB_{high}$ isolates, and thereby contributing to hypervirulence. The host's chemical, physical and immune defences against infection constitute stressful conditions to which microbes must adapt to associate with the host. It is therefore consistent that the general stress pathway regulates genes involved in host association, as it does in *L. monocytogenes* (*inlA* and *inlB*, *prfA* and indirectly LIPI-1 through PrfA)[32] and the host-associated bacterial species *Staphylococcus aureus* (for example, *sarA* and adhesins)[76], and that its differential responsiveness affects virulence and host adaptation as we have demonstrated here.

Fine-tuning the responsiveness of the general stress pathway emerges as a prime target for optimizing virulence gene expression to adapt to the host environment, as we demonstrate here for *L. monocytogenes*. Our findings challenge the conventional view of virulence heterogeneity being mainly driven by accessory genes. Therefore, gene content and comparative genomics may not be sufficient for in silico virulence prediction for microbial surveillance purposes and might need to be complemented with comparative transcriptomics.

Adaptive evolution is mainly studied at the level of gene function and expression level[77], but we show here that fine-tuning the responsiveness of transcriptional regulators is a very powerful and gradual mechanism for niche adaptation, consistent with similar

findings of adaptive evolution through transcriptional regulation in a Gram-negative bacterial pathogen[78]. Bacteria with multiple lifestyles are expected to fine-tune their transcriptional programs according to their ecological niches, potentially leading to transcriptional heterogeneity. The degree of such intraspecies heterogeneity depends on (i) the environmental changes which isolates are exposed to, (ii) the degree of differential niche preferences between isolates and (iii) the genetic plasticity of the transcriptional program regulation. We expect that such multi-layered transcriptional heterogeneity is not limited to bacteria and that a deeper understanding of the transcriptional networks, their regulation and the selective pressures on their upstream regulators will reveal comparable multilayered adaptive evolutionary processes across the tree of life.

Methods

Isolate selection and information

Isolates were collected by the French National Reference Center for *Listeria* in the context of clinical listeriosis cases in France, which are mandatory to be declared, and food controls. In this study, we used non-redundant food isolates (n = 382) and clinical isolates from the MONALISA study (n = 584) stored at −80 °C and additional isolates (n = 738) from a 14 °C storage collection. For the assessment of transcript levels in food isolates, isolates were selected from the 15 most prevalent food CCs (Extended Data Fig. 4a) in a de-doubled sampling from France for the period of August 2016 until July 2018. Isolates were selected to reflect prevalence in the food sampling and genetic diversity inside CCs. The main isolates and mutants used in this study can be found in Supplementary Table 5. Gene content in Supplementary Table 6 for $SigB_{high/low}$ isolates has been generated with Prokka v.1.14.5 and Roary v.3.11.2 [79,80]. The full isolate list is available upon request. Growth curves were established after 100× dilution of overnight cultures in 200 µl BHI in a shaking 96-well microplate in a Tecan Infinite M Nano at 37 °C and were measured during 400 min.

Mouse experiments

All in vivo experiments were carried out according to the Institut Pasteur guidelines for laboratory animals' husbandry and in compliance with European regulation 2010/63 EU. In vivo data for Extended Data Fig. 1 were retrieved from source data from ref. 4. All procedures were approved by the Animal Ethics Committee of Institut Pasteur, authorized by the French Ministry of Research and registered under APAFIS 14644-2018041116183944 and 39306-2022111412359059 v2. *L. monocytogenes* isolates were grown in BHI overnight, diluted 20× and grown until OD_{600} = 0.8 before being washed with PBS and diluted in PBS to 10^9 c.f.u. ml^{-1}. Seven- to 11-week-old female mice (C57BL/6 mEcad E16P KI (ref. 28)) were orally inoculated by gavage, as previously described[4,28]. Colony-forming unit loads at 5 days post-infection were assessed after homogenization of organs and serial dilutions in PBS and c.f.u. counts on BHI plates. For intestinal content, ALOA (Agar *Listeria* according to Ottaviani and Agosti) plates (BIOMERIEUX AEB150072 and AEB684420) were used, which are selective for *Listeria* and which allow to identify *L. monocytogenes*.

For co-inoculations, *L. monocytogenes* isolates were transformed with either pCMC34 (TdTomato) or pCMC44 (GFP), both described in ref. 37. To avoid a bias due to a difference in virulence due to the plasmid, for each co-inoculation both combinations were performed. We have previously shown that these plasmids do not have different fitness costs in vivo or in vitro[37].

For ligated loop experiments, 8- to 11-week-old axenix-free mice (C57BL/6 mEcad E16P KI (ref.28)) were used. Deep anaesthesia was induced with a mix of ketamine (100 mg per kg body weight; Imalgene 1000; Virbac) and xylazine chlorhydrate (10 mg per kg body weight; Rompun; Virbac), combined with analgesia with buprenorphine (0.1 mg per kg body weight, Vetergesic multidose 0.3 mg ml^{-1}). The skin was cleaned with Surfa Safe Anios premium, a laparotomy was performed, the small intestine was exposed and an ileal loop of 1.5 cm was prepared by clamping. *L. monocytogenes* isolates were grown in BHI overnight, diluted 20× and grown until OD_{600} = 0.8 before being washed with PBS (Gibco, Thermo Fisher Scientific 10010023) and diluted in PBS to 5.10^9 c.f.u. ml^{-1}. About 200 µl of inoculum (10^9 c.f.u. per loop) was injected into the loop. After 20 min, mice were killed by cervical dislocation. Intestinal loops were collected, with the content collected in a 1.5 ml tube that was immediately snap frozen in liquid nitrogen.

Concerning housing conditions, 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 that maintains a constant temperature of 22 ± 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 were covered with stainless steel grids and non-woven filter caps.

Protein expression levels

To assess the expression level of SigB regulon members, western blots against OpuCA were performed for all H/h strain pairs. Overnight cultures in BHI were collected by centrifugation, and bacterial pellets were then incubated with B-PER Complete Bacterial Extraction Reagent (Thermo Fisher Scientific, 89821) supplemented with lysozyme and DNase I for lysis. Protein concentrations were measured by using the Qubit Protein Assay Kit (Thermo Fisher Scientific, Q33212). About 17 µg of total protein was mixed with reducing sample buffer (NuPage, Invitrogen, Thermo Fisher Scientific NP032B), loaded on a Mini-PROTEAN TGX precast gel (4–15%, Bio-Rad 4561041) for electrophoresis and then transferred to a nitrocellulose membrane. The membranes were blocked for 2 h with 5% non-fat milk diluted in PBS-Tween 1% and incubated overnight at 4 °C with the primary polyclonal antibody directed against OpuCA (dilution 1:2,000)[81]. After 1 h of incubation with the secondary antibody (goat anti-chicken IgY H and L chain (HRP), Abcam ab6877, dilution 1:2,000), immunodetection was performed by using a chemiluminescence kit (Amersham ECL Prime, GE Healthcare, Sigma Aldrich GERPN2232), and signals were detected using the PXi imaging system (Syngene).

Quantification was performed by using Fiji[82], and OpuCA (signal at about 45kDA) quantification was normalized by the non-specific band at ~70 kDa, which corresponded to the equal protein input.

Live microscopy imaging of SigB-dependent reporter system

For strain construction, the promoter of *lmo2230* was inserted into pPL2 with TdTomato (pCMC34 (ref. 37)). This SigB reporter system was generated by digestion of pCMC34 with SacI and EagI (NEB R3156, R3505). The promoter of *lmo2230* of EGDe was amplified by PCR (Supplementary Table 5 for primers) and inserted into the digested plasmid by ligation with T4 ligase (NEB, M0202M). For microscopy, overnight cultures of bacteria were diluted 20× and grown in BHI until mid-exponential phase (OD_{600} = 0.8) before being washed and loaded on CellASIC ONYX microfluidic plates for bacterial cell culture (Sigma, B04A-03) with the CellASIC ONYX system, according to the manufacturer's instructions. After verifying that bacteria were loaded, they were grown 1–2 h before the start of imaging. Images were taken every 10 min. Observation was performed at 37 °C under a Zeiss LM710 microscope and acquisition with the ZEN v2.3 (black) software. At the first field of view per strain at every time point, the image was autofocused with fluorescence as input by the ZEN software (Zen Black v.2.3). For each field of view, a z-stack was taken. For the analysis, each stack was focused using the Brenner criteria, and bacteria were segmented and identified with ilastik v1.4[83]. For stress induction, EtOH was used, as it is a known inducer of SigB activity[51,52] and compatible with the experimental set-up.

Cell invasion assays

Gentamicin assays were performed as described previously[84]. In brief, L2071 or L2071 hEcad cells were seeded in DMEM (Gibco, 31966) + 10% FBS in 96-well plates, each containing 20,000 cells at 1 day before inoculation. Cells were then washed with DMEM. *L. monocytogenes* isolates were grown overnight at 37 °C before being co-inoculated with either L2071 or L2071 hEcad cells for 1 h in DMEM at 37 °C. After 1 h, cells were washed with DMEM + 10% FBS + gentamicin (Sigma Aldrich, G1272) 400 µl ml^{-1} and subsequently incubated with DMEM + 10% FBS + gentamicin 400 µl ml^{-1} for 40 min before washes with DMEM gentamicin 200 µl ml^{-1} and subsequent incubation for 1 h 20 min with DMEM gentamicin 200 µg ml^{-1}. After PBS washes, cells were lysed by incubation in cold water, and bacterial load was assessed by serial dilutions in PBS and c.f.u. count on BHI plates. InlA-dependent invasiveness was assessed by the c.f.u. count in L2071 hEcad − c.f.u. count in L2071. Statistical analysis between isolates from the same cgMLST pair were performed with paired Wilcoxon test and paired by experimental replicate.

Transcript level screen and real-time quantitative PCR

We assessed transcript levels either by RNA hybridization with fluorescent barcodes or real-time quantitative PCR (RT-qPCR). Bacteria were grown in 2 ml BHI for 1 h 50 min until mid-exponential phase (corresponding to OD_{600} = 0.8). Normal growth was assessed by optical assessment. Bacteria were pelleted and frozen in liquid nitrogen before being lysed in lysis tubes with 0.1 mm zirconium oxide beads (Ozyme, P000919-LYSK0-A.0) in a precellys 24 (Bertin) in 450 µl lysate buffer (10% glucose, 12.5 mM Tris, 75 mM EDTA (Sigma Aldrich E5391)) and 500 µl acid phenol pH 4.3 (Sigma-Aldrich P4682). The aqueous phase was used as cell lysate and input for the downstream application. The RNA concentration of the cell lysate was measured by Qubit broad range (Thermo Fisher Scientific Q10210) and diluted to 9 ng µl^{-1}. The transcript levels, based on hybridization between gene-specific probes (ordered from IDT) and RNA molecules, were then measured with the nanostring plexset system (Nanostring, 121200001, SH0004, 100054) according to the manufacturer's instructions. Probes were hybridizing for 17 h at 67 °C. Probe sequences can be found in Supplementary Table 5 and have been designed to maximize the match between probe sequence and all known alleles of evaluated genes. Three housekeeping genes, which are frequently used as references in qPCR and which did not vary between H and h strains in the RNA-seq experiment (Supplementary Table 2), *gyrA*, *gyrB* and *rpoB*, were used to normalize transcript counts with the manufacturer's software Nanostring nSolver v.4.0.66. Six samples with high counts for all target genes were used to normalize between batches of Nanostring plexset probes, according to the manufacturer's instructions. The two isolates from pair B were used as controls for the assay for every batch of isolates that were tested. Experimental batches constituted 40 different strains and the control isolates. RT-qPCR assays were performed as described previously[37]. In short, after reverse transcription, complementary DNA was quantified on the BI 7500 Fast real-time PCR system (Applied Biosystems) using the Power SYBR Green PCR master mix (Applied Biosystems, Thermo Fisher Scientific A25742). qPCR primer sequences can be found in Supplementary Table 5. Transcript levels are expressed as relative quantities (determined by differential Ct) compared with the housekeeping gene *gyrB*. For mutants, data are expressed relative to the wild-type H strain.

GWAS and heritability

We conducted a GWAS to screen for sequences associated with SigB pathway activity measure. Given the limited level of heterogeneity per CC in our dataset of natural isolates (Fig. 5) and to increase the power of our analysis, we also included data from in vitro evolved isolates, resulting in a total of 1,631 isolates (Fig. 6a), of which 34 were excluded during clustering. The GWAS was performed using ChoruMM v1.0, a multi-component linear mixed model (LMM) that we recently developed[62], and applied to short contiguous sequence of nucleotides called unitigs, extracted from the pangenome of isolates of interest using u-counter[63]. Briefly, the raw sequencing data of isolates is split into *k*-mers and assembled together using De Bruijn graphs based on overlapping *k*-mers. The graph is then compacted to construct unitigs, which are sequences corresponding to maximal unambiguous paths. From our isolates sequencing, we obtained 530,726 unitigs. We used chorumm's built-in algorithm to both remove potential outlier isolates and filter unitigs with extreme frequencies ($n < 5$ or $n > n_{total} - 5$, corresponding to the unitigs being present in less than 5 genomes, or in all but less than 5 genomes) to avoid false positives in downstream analysis (see ref. 62 for details). After filtering, we retained 1,597 strains and 441,091 unitigs, which correspond to 194,403 unique unitig distribution patterns, that is, when two unitigs were present in exactly the same isolates, they would correspond to the same pattern of presence/absence.

The multi-component LMM approach allows for fine modelling of the population structure of *L. monocytogenes* and can output both estimates of heritability per genetic component modelled and single unitig association test. In all analyses, we included an intercept and binarized variables corresponding to the lineage, CC, experimental batches and storage conditions as fixed effect. We first considered models including an extended number of genetic random terms modelling at either the cohort, lineage or CC level to derive heritability. For the GWAS screening, we used a simplified model including only lineages as random effect and unitigs as fixed effect. The unitig patterns were tested independently using a Wald test, assuming the test statistic follows a χ^2 distribution with 1 degree of freedom under the null hypothesis. For more detail, see ref. 62. Multiple tests were adjusted with Benjamini–Hochberg correction. Significant unitigs (Supplementary Table 4) were mapped on closed reference genomes of EGDe and hypervirulent CC1, CC4 and CC6 isolates with BLAST v.2.13.0[85].

SigB pathway activity variance explained by top associated unitigs was computed as the difference in adjusted R^2 between linear genetic models including unitigs of interest and a baseline model. The baseline model included variables corresponding to binarized versions of the experimental conditions, samples storage condition and *Listeria* lineage. As for the GWAS and heritability estimation, the variance explained was derived using rank-inverse transformed sigB read counts. Conversely, we did not included CC in the baseline model, as those variables are expected to be highly correlated with some of the associated unitigs. Unitigs were included using a forward procedure based on their contribution to SigB pathway variance, starting with the largest contributors to the less contributing ones. We identified 573 significant unitigs (Supplementary Table 4), of which 510 could be mapped onto EGDe reference genome.

For comparison purposes, we also derived the unadjusted R^2. This metric does not account for potential overfitting due to a large number of predictors, so that the R^2 estimation is expected to be biased upward. Nevertheless, this analysis confirmed that significant unitigs that contribute less to SigB pathway variance carry negligible additional information, meaning that all the significant unitigs together cannot fully explain the estimated heritability. Note that we did not use a LMM in this estimation because of the lack of gold standard for the derivation of variance explained by fixed effect from multi-component mixed models.

Experimental evolution

For in vitro experimental evolution under constant slow-growth conditions, eight isolates from the top third SigB activity of clinically associated CCs were selected. Isolates were stored at 14 °C in a glucose-free solid media (Bio-Rad 356-3683), sampled every 50 days and plated on BHI plates before being stored in cryo tubes. For serial passage experiments, 3 replicas of a $SigB_{high}$ isolate per hypervirulent CC was grown in

200 µl of BHI in a 96-well plate. About 10 µl was passed into fresh BHI every day for 100 days. For in vivo evolution, mice were inoculated with 2×10^8 c.f.u., as described above. Bacterial population at 5 d.p.i. were isolated from faeces by plating on ALOA and used for a next cycle of infection. For this, they were grown from ALOA in 5 ml BHI overnight and used as described above; 6 cycles were performed in total. Transcript levels of evolved populations were then assessed as described in 'Transcript level screen and real-time quantitative PCR' and compared with the non-evolved population.

Trans-expression mutant construction

For *rsbV* and *rsbW* overexpression, the *rsbV* and *rsbW* genes were amplified by PCR from CLIP2009/00558 genome (CC4; see Supplementary Table 5 for primer sequences and strain list) and cloned into the pCMC12 vector[37] between the EcoRV and SalI restriction sites (NEB R0195M, R3138M). In the case of *rsbV*, blunt-end ligation without EcoRV restriction was used due to a pre-existing restriction site for this enzyme in the gene. The pCMC12 plasmid is a suicide vector which integrates at the tRNAArg-attBB site as a single copy and allows the high expression of the cloned genes in a constitutive manner under the control of a strong promoter.

For *murC* and *nnrD* overexpression, the *murC* operon and *nnrD* gene were respectively amplified (see Supplementary Table 5 for primer sequences) from each member of the pair B and pair C (strain list in Supplementary Table 5). The amplicons were cloned under a strong promoter into the pCMC12 vector[22] between the *EcoRV* and *SalI* restriction site, using the Gibson assembly (NEB E2611) technic according to the provided protocol by the company.

The plasmids with the cloned gene of interest were then electroporated into previously prepared electrocompetent cells of the respective strain. Clone selection was done followed by sanger sequencing to verify the success of the cloning.

Construction of deletion mutants for lmo1913/pssZ

Deletion mutants were constructed as previously described[86] (see Supplementary Table 5 for primer sequences). The flanking regions of *lmo1913* were amplified by PCR from pair A (CC1). Once amplified, the flanking regions were purified and joined to the pLR16-pheS plasmid[87] by Gibson assembly (NEB E2611) between *Bam*HI and *Kpn*I restriction sites. The plasmids were then purified and electroporated into previously prepared electrocompetent bacteria. After plasmid integration and excision by sequence homology, deletions were verified by PCR inside and outside of the target genes. Deletion mutants were then confirmed by whole-genome sequencing to conserve those without off-target mutations.

Core genome d*N*/(d*N* + d*S*) analysis

For all core genes, d*N*/(d*N* + d*S*) between most common alleles per CC were calculated. For core genome analysis, all genetic elements were predicted for every genome with Prokka[80]. Genes were then grouped by Roary[79]. Genes with amino acid identity >95% were grouped together. Genes that were present in at least 99% of all isolates in this study (except evolved strains) were considered as core genes. For every CC and every core gene, the most common allele was selected. For each gene, CC major alleles were compared as follows: (1) alleles were aligned with each other with MAFFT v.7.511 (ref. 88) with '–auto'; (2) d*N* and d*S* (the number of non-synonymous and synonymous differences) were calculated by PAML v.4.10.0 (ref. 89), used through codeml in python from biopython v.1.84 (refs. 90,91), in pairwise mode. Comparisons which failed due to differences in sequence length were excluded resulting in comparisons in 1,795 core genes; (3) results were parsed and divided in comparisons between clinically associated CCs (CC1, CC4 and CC6) and food-associated CCs (CC9, CC31 and CC121). Comparisons between SigB-regulated and other core genes were performed with a Wilcox rank-sum test.

Reporting summary

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

Data availability

All datasets generated and/or analysed during the current study are available from the corresponding author. RNA-seq data are available on Sequence Read Archive under the accession number PRJNA1097069. Closed genomes of H/h isolates from pairs are available on bigsdb.pasteur.fr, and raw reads are released on Sequence Read Archive (for accession code, see Supplementary Table 5). Genome data analysed in this study were generated in the context of the epidemiological surveillance of listeriosis in France. As mentioned in the 'Décret no. 2016–806 du 16 juin 2016 relatif aux centres nationaux de référence pour la lutte contre les maladies transmissibles' and in the 'Arrêté du 16 juin 2016 fixant le cahier des charges des centres nationaux de références pour la lutte contre les maladies transmissibles', all samples collected at the French National Reference Center for *Listeria* as well as the data generated from these samples belong to the French Government and constitute a 'national collection of biological resources of interest to public health', which has to remain under medical and industrial confidentiality, and therefore they cannot be made publicly available.

References

1. Bosi, E. et al. Comparative genome-scale modelling of *Staphylococcus aureus* strains identifies strain-specific metabolic capabilities linked to pathogenicity. *Proc. Natl Acad. Sci. USA* **113**, E3801–E3809 (2016).
2. Roumagnac, P. et al. Evolutionary history of *Salmonella* Typhi. *Science* **314**, 1301–1304 (2006).
3. Moura, A. et al. Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. *Nat. Microbiol.* **2**, 16185 (2016).
4. Maury, M. M. et al. Uncovering *Listeria monocytogenes* hypervirulence by harnessing its biodiversity. *Nat. Genet.* **48**, 308–313 (2016).
5. Stewart, R. M. K. et al. Genetic characterization indicates that a specific subpopulation of *Pseudomonas aeruginosa* is associated with keratitis infections. *J. Clin. Microbiol.* **49**, 993–1003 (2011).
6. Mageiros, L. et al. Genome evolution and the emergence of pathogenicity in avian *Escherichia coli*. *Nat. Commun.* **12**, 765 (2021).
7. Vasquez-Rifo, A., Veksler-Lublinsky, I., Cheng, Z., Ausubel, F. M. & Ambros, V. The *Pseudomonas aeruginosa* accessory genome elements influence virulence towards *Caenorhabditis elegans*. *Genome Biol.* **20**, 1–22 (2019).
8. Tazi, A. et al. The surface protein HvgA mediates group B streptococcus hypervirulence and meningeal tropism in neonates. *J. Exp. Med.* **207**, 2313–2322 (2010).
9. Nienaber, J. J. C. et al. Methicillin-susceptible *Staphylococcus aureus* endocarditis isolates are associated with clonal complex 30 genotype and a distinct repertoire of enterotoxins and adhesins. *J. Infect. Dis.* **204**, 704–713 (2011).
10. Diard, M. & Hardt, W.-D. Evolution of bacterial virulence. *FEMS Microbiol. Rev.* **41**, 679–697 (2017).
11. Brown, S. P., Cornforth, D. M. & Mideo, N. Evolution of virulence in opportunistic pathogens: generalism, plasticity, and control. *Trends Microbiol.* **20**, 336–342 (2012).
12. Quereda, J. J. et al. Listeriolysin S is a streptolysin S-like virulence factor that targets exclusively prokaryotic cells in vivo. *mBio* **8**, e00259–17 (2017).
13. Rolhion, N. et al. A *Listeria monocytogenes* bacteriocin can target the commensal *Prevotella copri* and modulate intestinal infection. *Cell Host Microbe* **26**, 691–701.e5 (2019).

14. Laabei, M. et al. Evolutionary trade-offs underlie the multi-faceted virulence of *Staphylococcus aureus*. *PLoS Biol.* **13**, e1002229 (2015).
15. Hafner, L. et al. *Listeria monocytogenes* faecal carriage is common and depends on the gut microbiota. *Nat. Commun.* **12**, 6826 (2021).
16. Sauders, B. D. et al. Diversity of *Listeria* species in urban and natural environments. *Appl. Environ. Microbiol.* **78**, 4420–4433 (2012).
17. 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).
18. Dreyer, M. et al. *Listeria monocytogenes* sequence type 1 is predominant in ruminant rhombencephalitis. *Sci. Rep.* **6**, 36419 (2016).
19. Charlier, C. et al. Clinical features and prognostic factors of listeriosis: the MONALISA national prospective cohort study. *Lancet Infect. Dis.* **17**, 510–519 (2017).
20. Charlier, C. et al. Neonatal listeriosis presentation and outcome: a prospective study of 189 cases. *Clin. Infect. Dis.* **74**, 8–16 (2022).
21. 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).
22. Zhang, T. et al. Deciphering the landscape of host barriers to *Listeria monocytogenes* infection. *Proc. Natl Acad. Sci. USA* **114**, 6334–6339 (2017).
23. Orsi, R. H. & Wiedmann, M. Characteristics and distribution of *Listeria* spp., including *Listeria* species newly described since 2009. *Appl. Microbiol. Biotechnol.* **100**, 5273–5287 (2016).
24. Ragon, M. et al. A new perspective on *Listeria monocytogenes* evolution. *PLoS Pathog.* **4**, e1000146 (2008).
25. Lee, S. et al. *Listeria monocytogenes* source distribution analysis indicates regional heterogeneity and ecological niche preference among serotype 4b clones. *mBio* **9**, e00396–18 (2018).
26. Bertrand, S. et al. Diversity of *Listeria monocytogenes* strains of clinical and food chain origins in Belgium between 1985 and 2014. *PLoS ONE* **11**, e0164283 (2016).
27. Fagerlund, A. et al. Whole-genome sequencing analysis of *Listeria monocytogenes* from rural, urban, and farm environments in Norway: genetic diversity, persistence, and relation to clinical and food isolates. *Appl. Environ. Microbiol.* **88**, e0213621 (2022).
28. Disson, O. et al. Modeling human listeriosis in natural and genetically engineered animals. *Nat. Protoc.* **4**, 799–810 (2009).
29. Maury, M. M. et al. Hypervirulent *Listeria monocytogenes* clones' adaption to mammalian gut accounts for their association with dairy products. *Nat. Commun.* **10**, 2488 (2019).
30. Maury, M. M. et al. Spontaneous loss of virulence in natural populations of *Listeria monocytogenes*. *Infect. Immun.* **85**, e00541–17 (2017).
31. Ribeiro, V. B. et al. Contributions of σB and PrfA to *Listeria monocytogenes* salt stress under food relevant conditions. *Int. J. Food Microbiol.* **177**, 98–108 (2014).
32. Toledo-Arana, A. et al. The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* **459**, 950–956 (2009).
33. Nadon, C. A., Bowen, B. M., Wiedmann, M. & Boor, K. J. Sigma B contributes to PrfA-mediated virulence in *Listeria monocytogenes*. *Infect. Immun.* **70**, 3948–3952 (2002).
34. Oliver, H. F., Orsi, R. H., Wiedmann, M. & Boor, K. J. *Listeria monocytogenes* σB has a small core regulon and a conserved role in virulence but makes differential contributions to stress tolerance across a diverse collection of strains. *Appl. Environ. Microbiol.* **76**, 4216 (2010).
35. Liu, Y., Orsi, R. H., Boor, K. J., Wiedmann, M. & Guariglia-Oropeza, V. Home alone: elimination of all but one alternative sigma factor in *Listeria monocytogenes* allows prediction of new roles for σB. *Front. Microbiol.* **8**, 289905 (2017).
36. Lecuit, M. et al. A transgenic model for listeriosis: role of internalin in crossing the intestinal barrier. *Science* **292**, 1722–1725 (2001).
37. Maudet, C. et al. Bacterial inhibition of Fas-mediated killing promotes neuroinvasion and persistence. *Nature* **603**, 900–906 (2022).
38. Johansson, J. et al. An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. *Cell* **110**, 551–561 (2002).
39. Reniere, M. L. et al. Glutathione activates virulence gene expression of an intracellular pathogen. *Nature* **517**, 170–173 (2015).
40. Moura, A. et al. Emergence and global spread of *Listeria monocytogenes* main clinical clonal complex. *Sci. Adv.* **7**, eabj9805 (2021).
41. 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).
42. Palacios-Gorba, C. et al. Ruminant-associated *Listeria monocytogenes* isolates belong preferentially to dairy-associated hypervirulent clones: a longitudinal study in 19 farms. *Environ. Microbiol.* **23**, 7617–7631 (2021).
43. Mikonranta, L., Friman, V.-P. & Laakso, J. Life history trade-offs and relaxed selection can decrease bacterial virulence in environmental reservoirs. *PLoS ONE* **7**, e43801 (2012).
44. Alizon, S., Hurford, A., Mideo, N. & Van Baalen, M. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *J. Evol. Biol.* **22**, 245–259 (2009).
45. Sultan, I., Fromion, V., Schbath, S. & Nicolas, P. Statistical modelling of bacterial promoter sequences for regulatory motif discovery with the help of transcriptome data: application to *Listeria monocytogenes*. *J. R. Soc. Interface* **17**, 5273–5287 (2020).
46. Kazmierczak, M. J., Mithoe, S. C., Boor, K. J. & Wiedmann, M. *Listeria monocytogenes* σB regulates stress response and virulence functions. *J. Bacteriol.* **185**, 5722–5734 (2003).
47. Ferreira, A., Gray, M., Wiedmann, M. & Boor, K. J. Comparative genomic analysis of the sigB operon in *Listeria monocytogenes* and in other Gram-positive bacteria. *Curr. Microbiol.* **48**, 39–46 (2004).
48. Hecker, M., Pané-Farré, J. & Uwe, V. SigB-dependent general stress response in *Bacillus subtilis* and related Gram-positive bacteria. *Annu. Rev. Microbiol.* **61**, 215–236 (2007).
49. Pané-Farré, J., Quin, M. B., Lewis, R. J. & Marles-Wright, J. Structure and function of the stressosome signalling hub. *Subcell. Biochem.* **83**, 1–41 (2017).
50. Locke, J. C. W., Young, J. W., Fontes, M., Jiménez, M. J. H. & Elowitz, M. B. Stochastic pulse regulation in bacterial stress response. *Science* **334**, 366–369 (2011).
51. Young, J. W., Locke, J. C. W. & Elowitz, M. B. Rate of environmental change determines stress response specificity. *Proc. Natl Acad. Sci. USA* **110**, 4140–4145 (2013).
52. Chaturongakul, S. & Boor, K. J. σB activation under environmental and energy stress conditions in *Listeria monocytogenes*. *Appl. Environ. Microbiol.* **72**, 5197 (2006).
53. Williams, A. H. et al. The cryo-electron microscopy supramolecular structure of the bacterial stressosome unveils its mechanism of activation. *Nat. Commun.* **10**, 3005 (2019).
54. Fraser, K. R., Sue, D., Wiedmann, M., Boor, K. & O'Byrne, C. P. Role of σB in regulating the compatible solute uptake systems of *Listeria monocytogenes*: osmotic induction of opuC is σB dependent. *Appl. Environ. Microbiol.* **69**, 2015–2022 (2003).
55. Fischer, M. A. et al. *Listeria monocytogenes* genes supporting growth under standard laboratory cultivation conditions and during macrophage infection. *Genome Res.* **32**, 1711–1726 (2022).
56. Köseoğlu, V. K. et al. *Listeria monocytogenes* exopolysaccharide: origin, structure, biosynthetic machinery and c-di-GMP-dependent regulation. *Mol. Microbiol.* **96**, 728–743 (2015).

57. Mahapatra, S., Crick, D. C. & Brennan, P. J. Comparison of the UPD-N-acetylmuramate:L-alanine ligase enzymes from *Mycobacterium tuberculosis* and *Mycobacterium leprae*. *J. Bacteriol.* **182**, 6827–6830 (2000).
58. Beaulaurier, J., Schadt, E. E. & Fang, G. Deciphering bacterial epigenomes using modern sequencing technologies. *Nat. Rev. Genet.* **20**, 157–172 (2019).
59. Sánchez-Romero, M. A. & Casadesús, J. The bacterial epigenome. *Nat. Rev. Microbiol.* **18**, 7–20 (2019).
60. Abel, S., Wiesch, P. A., zur, Davis, B. M. & Waldor, M. K. Analysis of bottlenecks in experimental models of infection. *PLoS Pathog.* **11**, e1004823 (2015).
61. Ondrusch, N. & Kreft, J. Blue and red light modulates SigB-dependent gene transcription, swimming motility and invasiveness in *Listeria monocytogenes*. *PLoS ONE* **6**, e16151 (2011).
62. Frouin, A. et al. ChoruMM: a versatile multi-components mixed model for bacterial-GWAS. Preprint at *bioRxiv* https://doi.org/10.1101/2023.03.28.534531 (2023) .
63. Jaillard, M. et al. A fast and agnostic method for bacterial genome-wide association studies: bridging the gap between k-mers and genetic events. *PLoS Genet.* **14**, e1007758 (2018).
64. Lees, J. A. et al. Fast and flexible bacterial genomic epidemiology with PopPUNK. *Genome Res.* **29**, 304–316 (2019).
65. Mazzuoli, M.-V. et al. The CovR regulatory network drives the evolution of Group B *Streptococcus* virulence. *PLoS Genet.* **17**, e1009761 (2021).
66. Sokurenko, E. V., Gomulkiewicz, R. & Dykhuizen, D. E. Source–sink dynamics of virulence evolution. *Nat. Rev. Microbiol.* **4**, 548–555 (2006).
67. Yin, Y. et al. A hybrid sub-lineage of *Listeria monocytogenes* comprising hypervirulent isolates. *Nat. Commun.* **10**, 4283 (2019).
68. Drolia, R., Tenguria, S., Durkes, A. C., Turner, J. R. & Bhunia, A. K. *Listeria* adhesion protein induces intestinal epithelial barrier dysfunction for bacterial translocation. *Cell Host Microbe* **23**, 470–484.e7 (2018).
69. Liu, D. et al. Cell-surface anchoring of *Listeria* adhesion protein on *L. monocytogenes* is fastened by internalin B for pathogenesis. *Cell Rep.* **42**, 112515 (2023).
70. Vasanthakrishnan, R. B. et al. PrfA regulation offsets the cost of *Listeria* virulence outside the host. *Environ. Microbiol.* **17**, 4566–4579 (2015).
71. Alizon, S. & Michalakis, Y. Adaptive virulence evolution: the good old fitness-based approach. *Trends Ecol. Evol.* **30**, 248–254 (2015).
72. Guerreiro, D. N. et al. In vitro evolution of *Listeria monocytogenes* reveals selective pressure for loss of SigB and AgrA function at different incubation temperatures. *Appl. Environ. Microbiol.* **88**, e0033022 (2022).
73. Lee, B. H. et al. Exploring *Listeria monocytogenes* transcriptomes in correlation with divergence of lineages and virulence as measured in *Galleria mellonella*. *Appl. Environ. Microbiol.* **85**, e01370–19 (2019).
74. Casadevall, A. & Pirofski, L. A. Benefits and costs of animal virulence for microbes. *mBio* **10**, e00863–19 (2019).
75. Levin, B. R. & Bull, J. J. Short-sighted evolution and the virulence of pathogenic microorganisms. *Trends Microbiol.* **2**, 76–81 (1994).
76. Bischoff, M. et al. Microarray-based analysis of the *Staphylococcus aureus* σB regulon. *J. Bacteriol.* **186**, 4085–4099 (2004).
77. Payne, J. L. & Wagner, A. The causes of evolvability and their evolution. *Nat. Rev. Genet.* **20**, 24–38 (2018).
78. Weimann, A. et al. Evolution and host-specific adaptation of *Pseudomonas aeruginosa*. *Science* **385**, eadi0908 (2024).
79. Page, A. J. et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693 (2015).
80. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
81. Utratna, M., Cosgrave, E., Baustian, C., Ceredig, R. H. & O'Byrne, C. P. Effects of growth phase and temperature on σB activity within a *Listeria monocytogenes* population: evidence for RsbV-independent activation of σB at refrigeration temperatures. *Biomed. Res. Int.* **2014**, 641647 (2014).
82. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
83. Berg, S. et al. ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226–1232 (2019).
84. Lecuit, M., Ohayon, H., Braun, L., Mengaud, J. & Cossart, P. Internalin of *Listeria monocytogenes* with an intact leucine-rich repeat region is sufficient to promote internalization. *Infect. Immun.* **65**, 5309–5319 (1997).
85. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
86. Monk, I. R., Gahan, C. G. M. & Hill, C. Tools for functional postgenomic analysis of *Listeria monocytogenes*. *Appl. Environ. Microbiol.* **74**, 3921–3934 (2008).
87. Argov, T., Rabinovich, L., Sigal, N. & Herskovits, A. A. An effective counterselection system for *Listeria monocytogenes* and its use to characterize the monocin genomic region of strain 10403S. *Appl. Environ. Microbiol.* **83**, e02927–16 (2017).
88. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772 (2013).
89. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
90. Cock, P. J. A. et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423 (2009).
91. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME suite. *Nucleic Acids Res.* **43**, W39–W49 (2015).

Acknowledgements

We thank the members of the Biology of Infection Unit for helpful discussions including Kieran Toms for proofreading. We thank Martin Lukačišin, Idan Yelin and Roy Kishony for helpful comments on the manuscript. We thank Conor O'Byrne (University of Galway, Ireland) for kindly providing the OpuCA antibody. L. Hafner was supported by the Pasteur-Paris University International Ph.D. Program, funded by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska–Curie grant agreement number 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, Laboratoire d'Excellence Integrative Biology of Emerging Infectious Diseases and the Fondation Le Roch-Les Mousquetaires. M.L. is a senior member of Institut Universitaire de France. H.A. laboratory is supported by the INCEPTION program (Investissement d'Avenir grant ANR-16-CONV-0005).

Author contributions

M.L. initiated and coordinated the project, together with L. Hafner. L. Hafner and M.L. designed the study. L. Hafner, L. Huang, E.G., C.M., C.G. and O.D. conducted the in vivo experiments in mice, with help from H.B.-D., A.V., Y.-H.T. and P.G. L. Hafner, L. Huang, E.G., C.M., C.G., O.D. and A.V. performed in vitro invasion assays. C.M., C.G., L. Huang, E.G., H.B.-D. and L. Hafner performed RT-qPCR experiments. L. Huang conducted the RNA-seq experiments, with corresponding data analysis performed by H.V., R.L. and J.-Y.C., with help from L. Huang, L. Hafner and A.M. N.T.-R., H.B.-D., C.C., M.M., L. Hafner and A.L. handled the bacteria collection and associated metadata. L. Hafner conducted the experimental evolution experiments, transcript level

screen, live imaging and the corresponding data analyses. E.P. and E.P.C.R. helped with genetic and evolution experiments analyses. F.L., A.F. and H.A. performed the GWAS, with the help of R.C., M.M. and L. Hafner. L. Hafner, E.G., C.G., A.V. and M.D. performed strain and plasmid constructions. L. Hafner and M.L. wrote the manuscript, with the input of the other co-authors.

Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41564-024-01859-8.

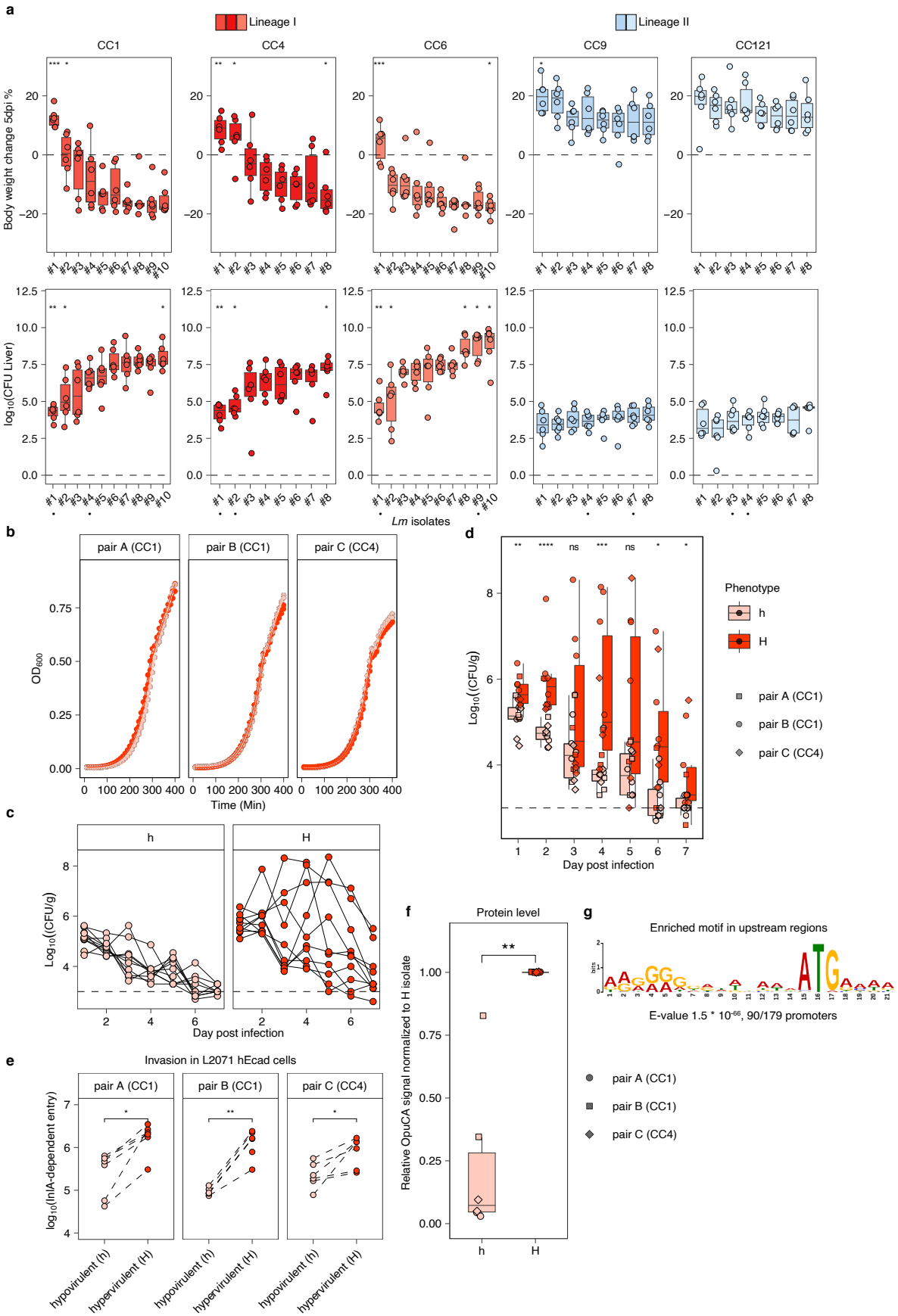
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41564-024-01859-8.

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

Peer review information *Nature Microbiology* thanks Arun Bhunia, Renato Orsi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

[1]Biology of Infection Unit, Institut Pasteur, Université Paris Cité, Inserm U1117, Paris, France. [2]Statistical Genetics Unit, Institut Pasteur, Université Paris Cité, CNRS USR375, Paris, France. [3]Bioinformatics and Biostatistics Hub, Institut Pasteur, Université Paris Cité, Paris, France. [4]National Reference Center and WHO Collaborating Center Listeria, Institut Pasteur, Paris, France. [5]Transcriptome et Epigenome Platform, Biomics, Center for Technological Resources and Research, Institut Pasteur, Université Paris Cité, Paris, France. [6]Necker-Enfants Malades University Hospital, Division of Infectious Diseases and Tropical Medicine, Institut Imagine, APHP, Paris, France. [7]Human Evolutionary Genetics Unit, Institut Pasteur, Université Paris Cité, CNRS UMR2000, Paris, France. [8]Sequence Bioinformatics Group, Institut Pasteur, Université Paris Cité, Paris, France. [9]Microbial Evolutionary Genomics Unit, Institut Pasteur, Université Paris Cité, CNRS UMR3525, Paris, France. ✉e-mail: marc.lecuit@pasteur.fr

https://doi.org/10.1038/s41564-024-01859-8

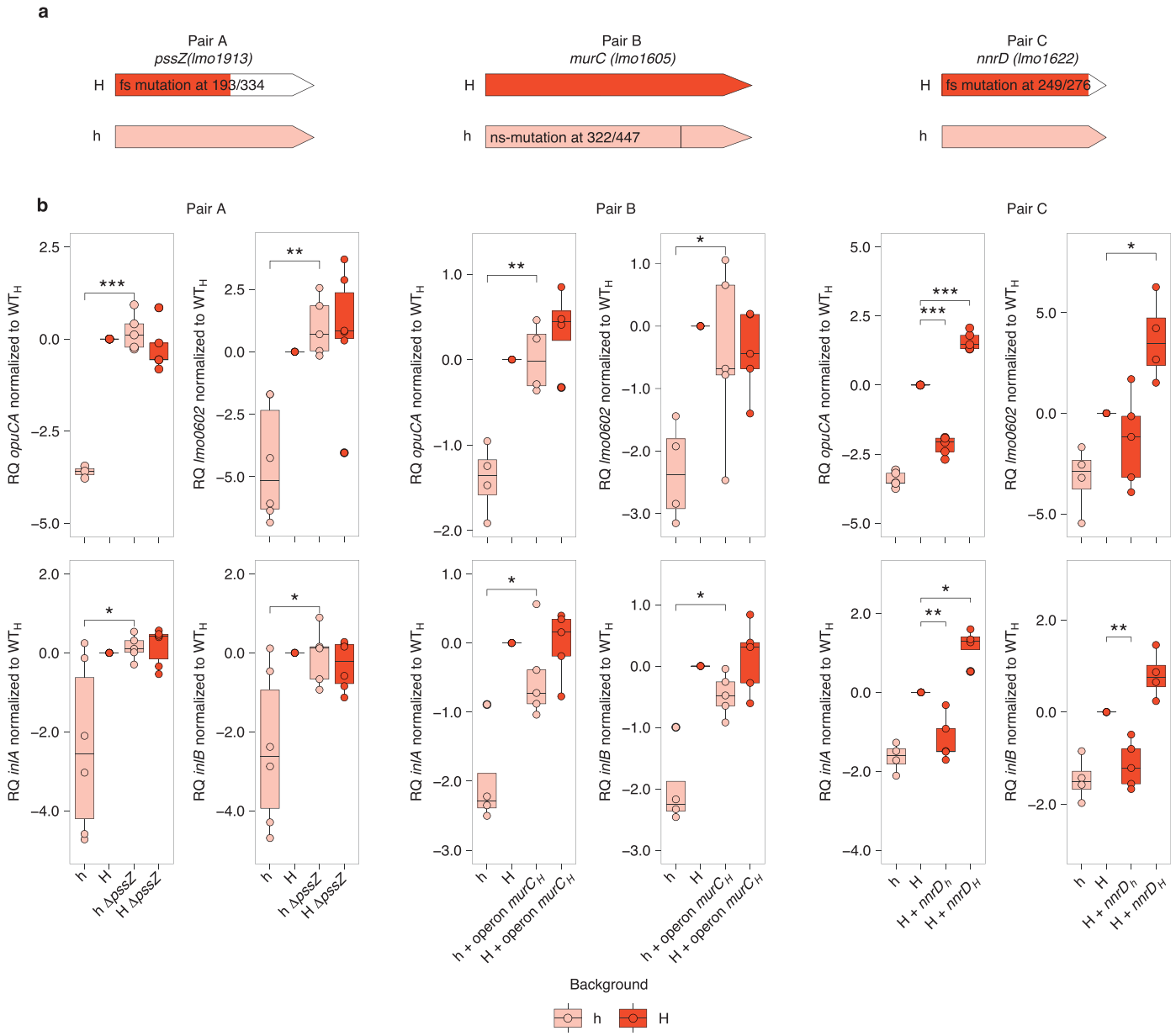


Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41564-024-01859-8

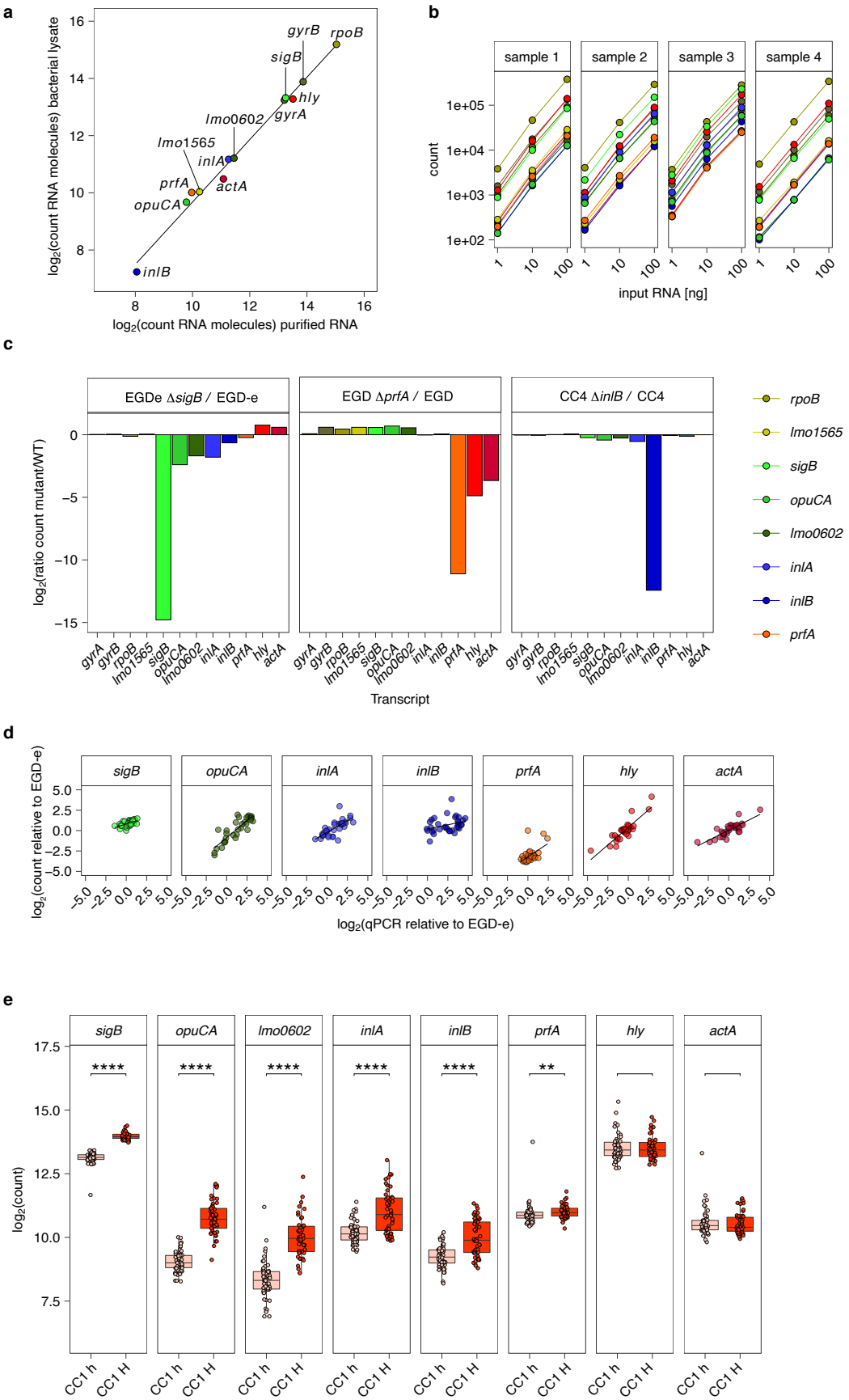
Extended Data Fig. 1 | Characterization of H and h isolates. a. Body weight change of mice 5 dpi (top) and CFU in the liver (bottom) after oral inoculation with $2x10^8$ CFU of isolates from CC1, CC4, CC6, CC9 and CC121. Each boxplot represents 6 individual mice inoculated with the same *Lm* isolate shown grouped per CC as in[4]. Food isolates are indicated by a dot below legend (2 isolates *per* CC). Statistical analysis has been performed for each isolate against all other isolates of the same CC with Wilcoxon test (*P* values from left to right, for CC1 = (3.37×10^{-4}, 4.83×10^{-2}, 0.31, 0.69, 0.45, 0.74, 0.20, 8.58×10^{-2}, 0.12), for CC4 = (5.2×10^{-3}, 2.15×10^{-2}, 0.48, 0.75, 0.25, 0.17, 0.39, 3.77×10^{-2}), for CC6 = (8.5×10^{-3}, 0.27, 0.11, 0.90, 0.58, 0.33, 0.12, 0.23, 0.31, 4.8×10^{-2}), for CC121 = (4.9×10^{-2}, 0.11, 0.63, 0.88, 0.36, 0.30, 0.61, 0.50) and for CC9 = (0.17, 0.67, 0.79, 0.45, 0.59, 0.40, 0.50, 0.45)). Of note, not all isolates from Fig. 1 belong to this dataset[4]. **b**. Growth curve of isolates from Fig. 1. in a microplate (Methods). Each curve corresponds to a single isolate, either H (dark red) or h (light red) from Fig. 1. **c**. CFU in feces *per* dpi after inoculation with h or H strains from Fig. 1. Line indicates detection minimum for 0.05g of feces. **d**. As (c) but shown as boxplots (n = 3 per isolate). Each point corresponds to the fecal carriage of a single infected mouse at the respective day with *P*-values from day 1 to 7 = (5.20×10^{-3}, 2.2×10^{-5}, 0.19, 9.9×10^{-4}, 1.08×10^{-2} and 2.42×10^{-2}). **e**. Cell invasion assay with isolates from Fig. 1 (n=6, lines indicate experiments which were performed in parallel). Cells were infected with overnight cultures of *Lm* and treated with gentamicin after 1h, before being inoculated with gentamicin for 2h. InlA-dependent entry was evaluated by difference in CFU in entry in L2071 hEcad cells and L2071 cells (For pairs A, B and C *P*-values = 2.6×10^{-2}, 4.30×10^{-3} and 2.6×10^{-2}). **f**. Quantification of 2 Western blots for each H and h isolate against SigB-dependent protein OpuCA (H *vs* h isolates, $P = 2.8\times10^{-3}$). Each sample came from an independent culture of the respective isolate. EF-Tu was used to check for amount of loaded protein **g**. DNA motifs significantly enriched (identified by MEME[82]) in 100bp upstream and first 10bp of genes significantly more transcribed in H isolates against all other gene upstream regions and corresponding E-value. The motif is similar to the -10 box described for SigB dependent promoters, GGGWAW, at variable distance from the start codon ATG[45,46]. Two-sided Wilcoxon rank-sum test and *per* group against all other groups in (a), between groups of the same day in (c) and paired two-sided Wilcoxon rank-sum test between results from same experiment in (d) and unpaired for H against h isolates in (f). * $P<0.05$, ** $P<0.01$, *** $P<0.001$, **** $P<0.0001$. 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 x IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles).

https://doi.org/10.1038/s41564-024-01859-8



Extended Data Fig. 2 | SigB pathway differential responsiveness is genetically encoded. a. Schematic depiction of genetic difference between H and h strain for all pairs. Mutation positions are indicated as amino acid residues. Filled bar shows sequence similarity between alleles and vertical thin lines indicate non-synonymous variations. **b**. SigB pathway activity of deletion and complementation mutants measured by RT-qPCR of *opuCA*, *lmo0602*, *inlA* and *inlB* (n = 6 for Pair A, n = 4 for Pair B and C). Only relevant statistical comparisons are shown (*i.d.* between ancestor and mutant strains). Each point corresponds to an independent culture of the respective strain or mutant. Values were normalized to the H strain of each pair. RQ values were $\log_2$-transformed. Two-sided t-test was applied for statistical comparisons. * 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. For boxplots, the hinges represent the first and third quartile of the distribution. The central line represents the median value. The whiskers extend from the hinge to the largest or smallest value no further than 1.5 x IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles)). For precise P-values, see Supplementary Information.

https://doi.org/10.1038/s41564-024-01859-8

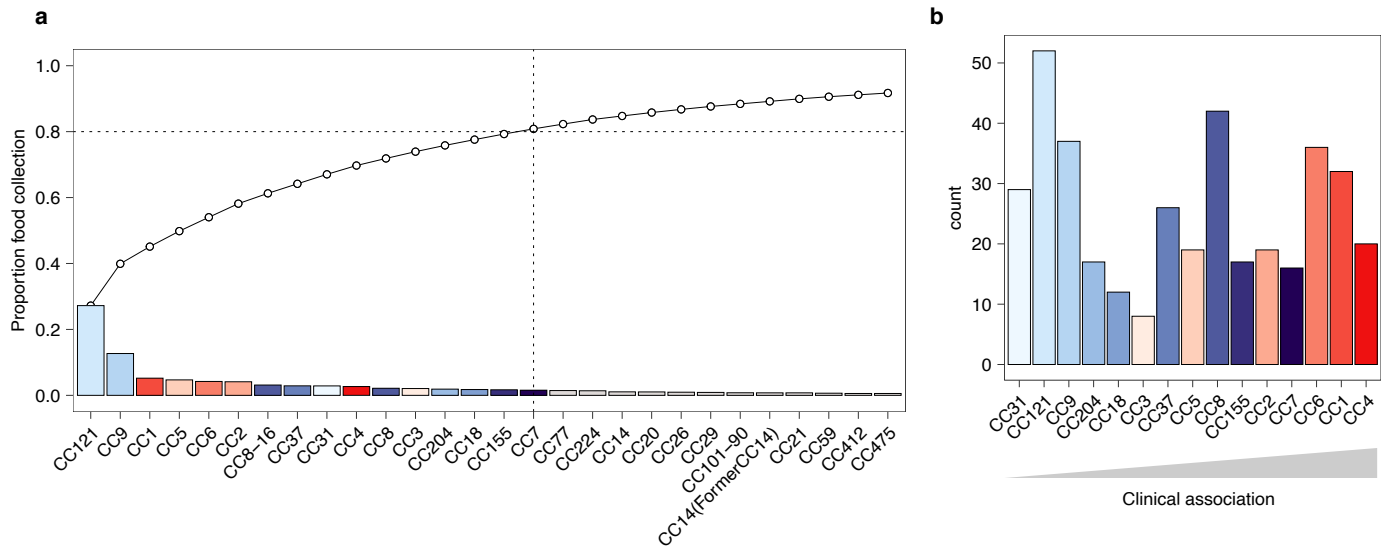


Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41564-024-01859-8

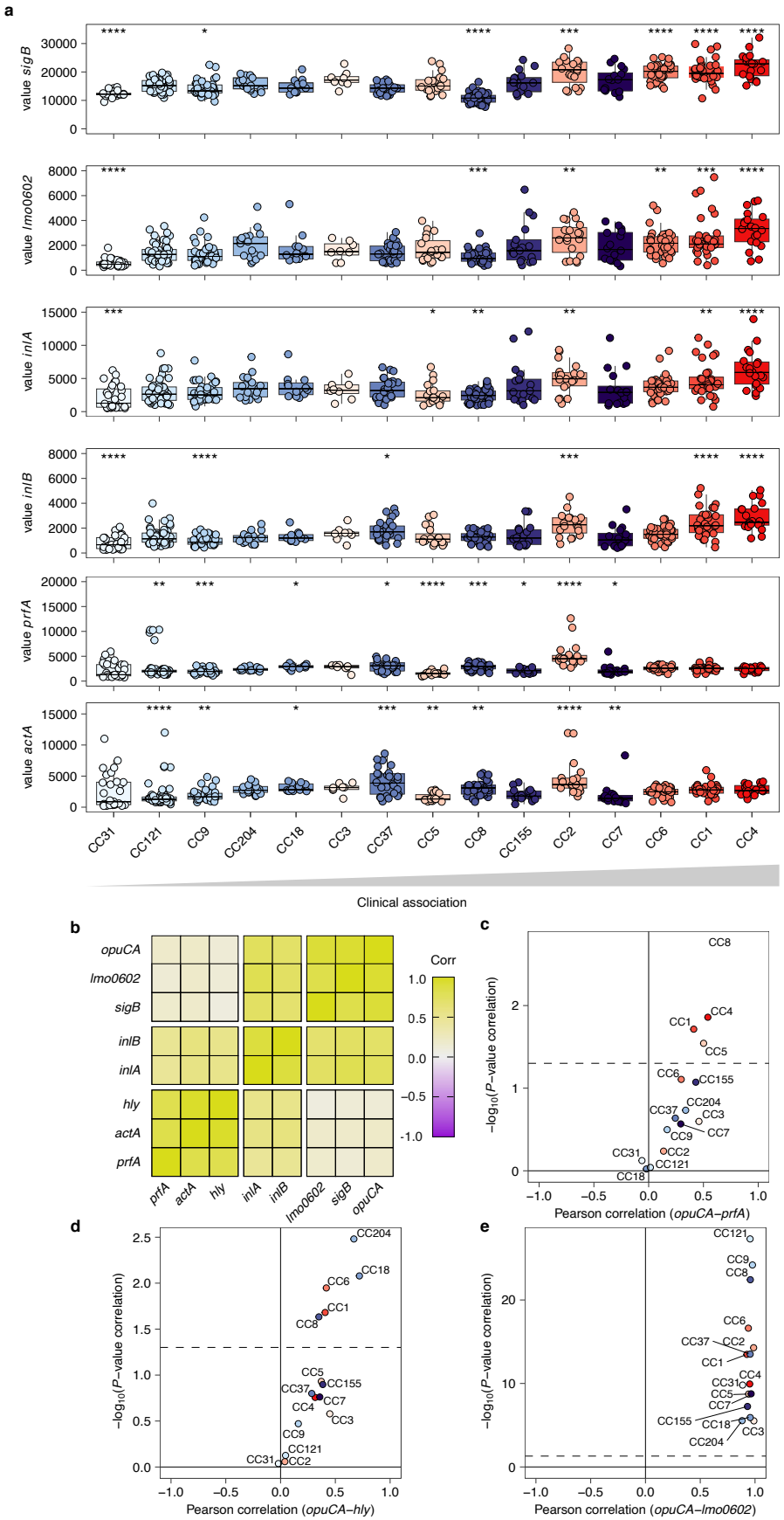
Extended Data Fig. 3 | Setup and validation of mid-throughput pathway activity assay. a. Correlation of $\log_{10}$(Nanostring read counts) for all evaluated genes between cell lysate and purified RNA (R^2=0.989, P=1.6x10^{-11}). This shows that we could use cell lysates instead of purified RNA for our pathway activity experiments **b**. Titration of RNA in 4 samples, showing that result is not dependent on RNA concentration between 1 and 100 ng. Optimal concentration for coverage of the nanostring cartridge of 9 ng of RNA *per* sample was established based on these data. **c**. Evaluation of individual probes in EGDe, EGD, which has a constitutively active PrfA, EGD *ΔprfA* (left), EGDe and EGDe *ΔsigB* (center) and a CC4 and CC4 *ΔinlB*, demonstrating the specificity of our assay. The housekeeping genes *gyrA*, *gyrB*, and *rpoB* were used for normalization. **d**. Relation between nanostring counts in cell lysate and RT-qPCR on cDNA results with primers from Table 3 in 34 *Lm* isolates. Values are expressed as $\log_2$ of the relative quantity to EGDe from the same experiment. **e**. Log2 of transcript levels of pair B from Fig. 1 across different experiments of isolates batches from Fig. 2, which has been used as a control in the SigB pathway activity screen. Each point corresponds to an independent culture of the respective strain. For boxplots, the hinges represent the first and third quartile of the distribution. The central line represents the median value. The whiskers extend from the hinge to the largest or smallest value no further than 1.5 x IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles).

https://doi.org/10.1038/s41564-024-01859-8



Extended Data Fig. 4 | Food sampling for SigB pathway activity screen. **a**. Proportion of each CC in food sampling of the *Listeria* National Reference Center in France (colored bars) and their cumulative value (open circles). The dashed line delimits on the left the 15 most prevalent CCs, which cover >80% of the food sampling. These CCs have been considered in this study. **b**. Number of isolates *per* CC considered in Fig. 5a-d.

https://doi.org/10.1038/s41564-024-01859-8

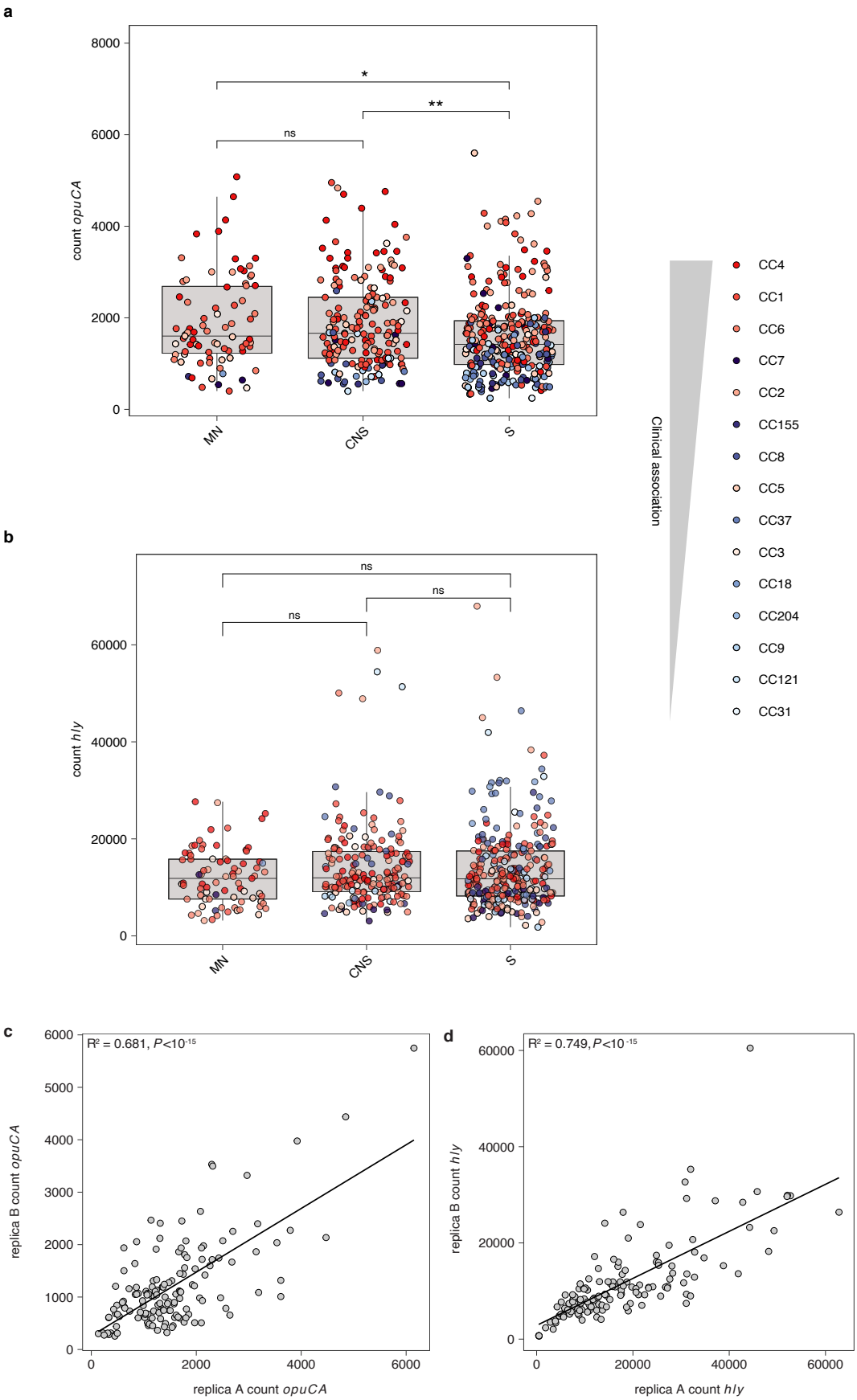


Extended Data Fig. 5 | See next page for caption.

https://doi.org/10.1038/s41564-024-01859-8

Extended Data Fig. 5 | Counts of other SigB and PrfA-dependent transcripts and their correlations. a. Transcription screen results shown as $\log_{10}$(read counts) for other genes across CCs for isolates from Fig. 4. Statistical comparison with two-sided Wilcoxon rank-sum test *per* group against all other groups. As in Fig. 4, each point corresponds to a single isolate from the respective CC (n = 372). **b**. Correlation over all samples of transcript levels of all evaluated genes. SigB regulated genes correlate strongly together, PrfA regulated genes as well and *inlA* and *inlB* correlate with both. **c.-e**. Pearson correlation of transcript counts *per* CC on x-axis and corresponding negative $\log_{10}$(*P*-value) without correction for multiple testing for *opuCA-lmo0602*, *opuCA-hly* and *opuCA-prfA*, respectively. The horizontal dashed line indicates *P*=0.05. For boxplots, the hinges represent the first and third quartile of the distribution. For boxplots, the hinges represent the first and third quartile of the distribution. The central line represents the median value. The whiskers extend from the hinge to the largest or smallest value no further than 1.5 x IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles).

https://doi.org/10.1038/s41564-024-01859-8

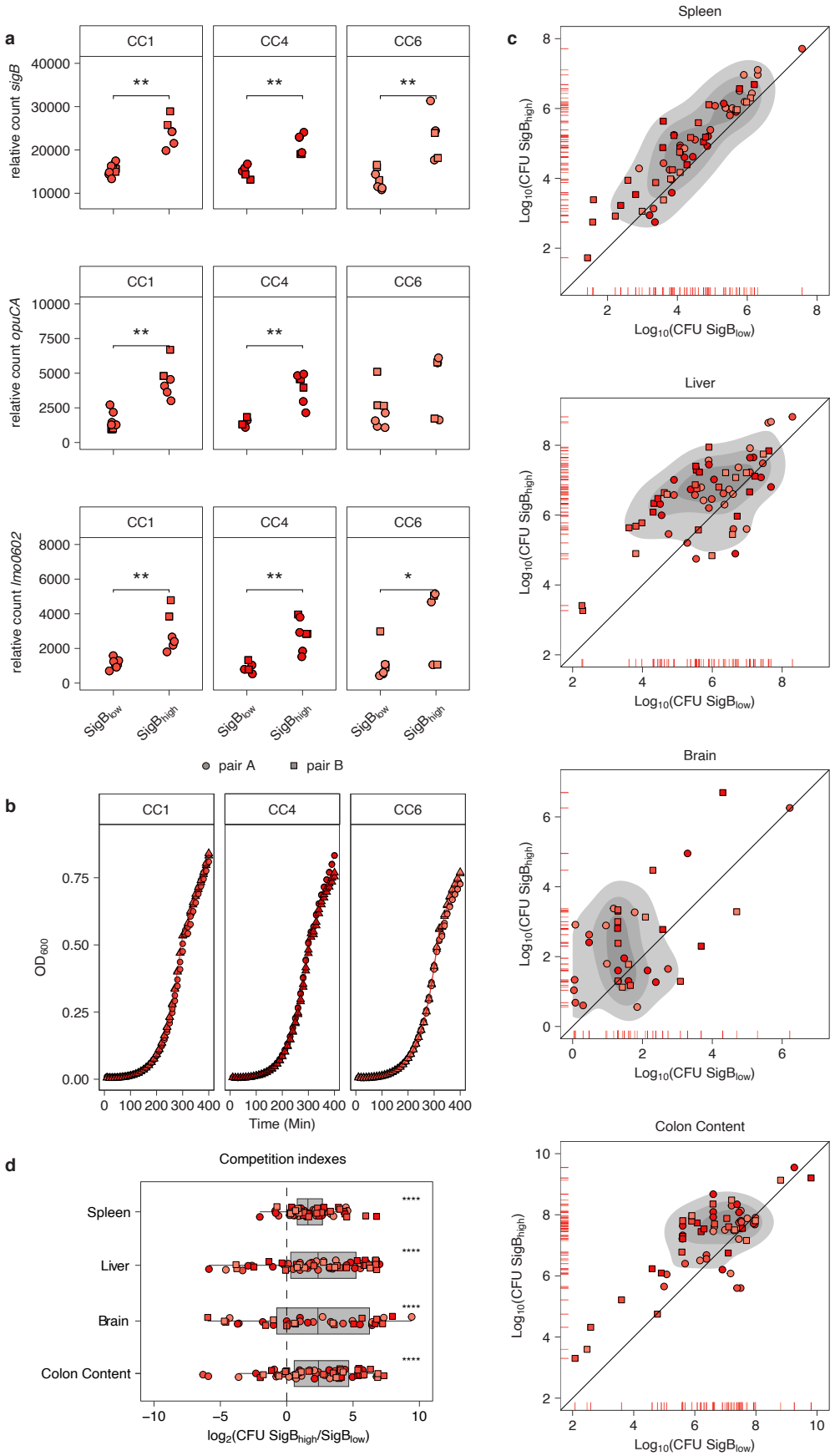


Extended Data Fig. 6 | See next page for caption.

https://doi.org/10.1038/s41564-024-01859-8

Extended Data Fig. 6 | SigB and PrfA activity in clinical isolates. Transcript levels of **a**. *opuCA* (*P*-values for MN *vs*. CNS = 0.9, for MN *vs*. S = 1.5×10^{-2} and for CNS *vs*. S = 1.3×10^{-3}) and **b**. *hly* (*P*-values for MN *vs*. CNS = 0.13, for MN *vs*. S = 0.3 and for CNS *vs*. S = 0.4) of clinical *Lm* isolates from maternal-neonatal (MN), central nervous system (CNS) or septic (S) infections. Each point corresponds to a single isolate (n = 79 for MN, n = 183 for CNS and n = 288 for S). Confirming our results, we selected for each CC 2 food isolates from the bottom, the center and the top third of SigB activity and reproduced twice to test the robustness of our assay. Pearson correlation of additional replicas A and B for **c**. *opuCA* and **d**. *hly* to assess biological reproducibility of SigB pathway activity without correction for multiple testing. For boxplots, the hinges represent the first and third quartile of the distribution. The central line represents the median value. The whiskers extend from the hinge to the largest or smallest value no further than 1.5 x IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles).

https://doi.org/10.1038/s41564-024-01859-8

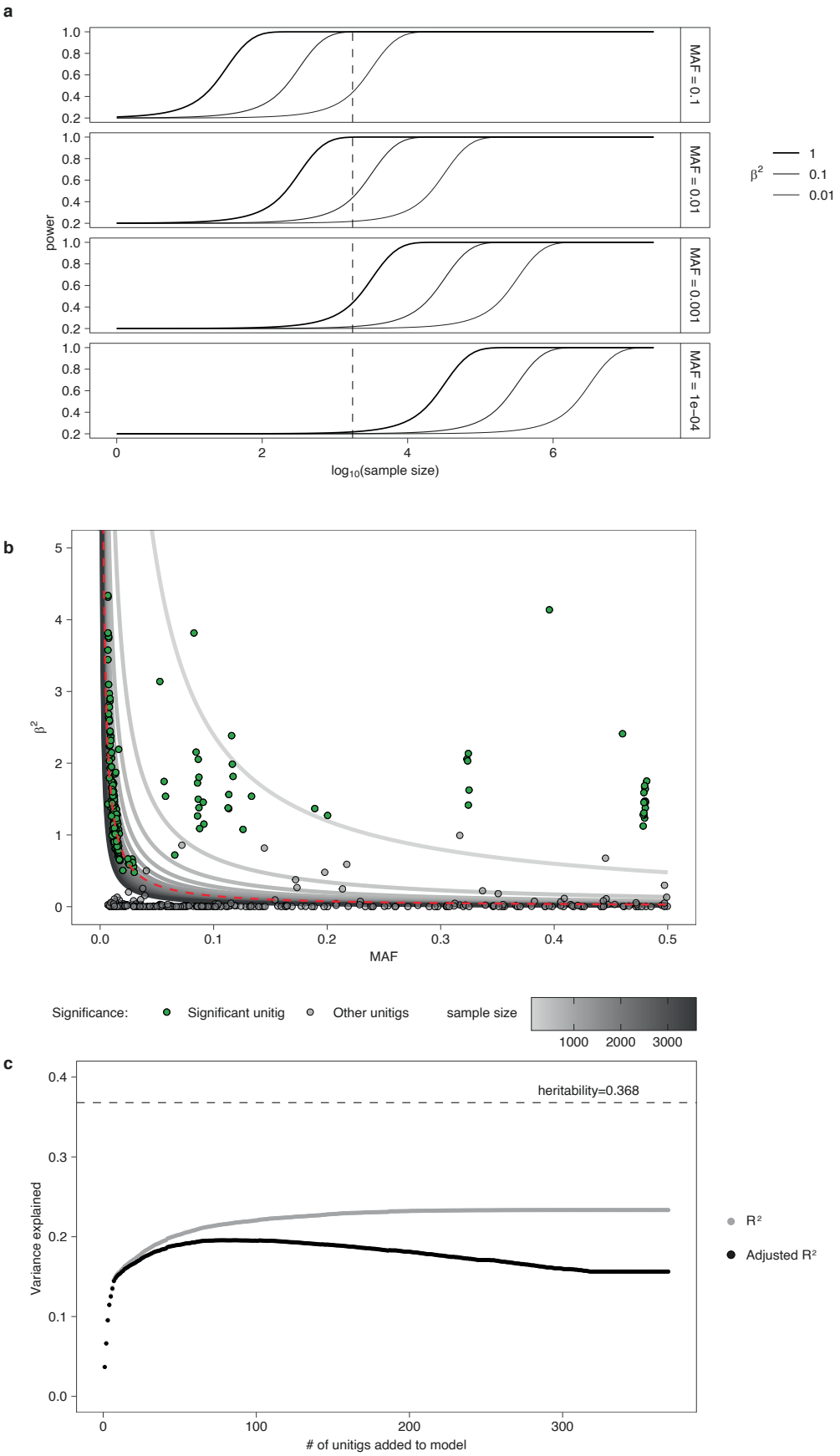


Extended Data Fig. 7 | See next page for caption.

https://doi.org/10.1038/s41564-024-01859-8

Extended Data Fig. 7 | SigB pathway activity predicts intraclonal virulence differences. a. SigB pathway activity for 2 isolates of CC1, CC4 and CC6 from either the top ($SigB_{high}$) or bottom ($SigB_{low}$), which have been selected for co-infections. *P*-values for *sigB*, *opuCA* and *lmo0602* in CC1 = 2.4×10^{-3}, 2.4 × 10-3, 2.4×10^{-3} in CC4 = 8×10^{-3}, 5.7×10^{-3}, 5.7×10^{-3}, in CC6 = 2.5×10^{-3}, 0.15, 3.0×10^{-2}. **b**. Growth curve of isolates from C. Triangle symbolize $SigB_{low}$ and circles $SigB_{high}$. **c**. CFUs of $SigB_{low}$ (x-axis) and $SigB_{high}$ (y-axis) isolates from co-infections with each 10^8 CFUs. 2 pairs *per* CC were tested (circle/square) at 5dpi in liver, spleen, brain and colon content. Marks along axis show distribution. Shades of distribution show quartiles of distribution. **d**. Competition indexes, calculated as the ratio of $SigB_{high}$ to $SigB_{low}$ *per* mouse and organ from co-inoculations with each 10^8 CFUs. Each point corresponds to an organ coming from an independently infected mouse. 2 pairs (circle/square) *per* CC (color as in a-d) of $SigB_{high}$ *vs* $SigB_{low}$ were tested, each in 6 mice. Competition indexes at 5dpi in liver, spleen, brain and colon content. *P*-values in Spleen = 2.75×10^{-10}, in Liver = 3.75×10^{-6}, in Brain = 3.93×10^{-3} and in Colon Content = 9.71×10^{-7}. For boxplots, the hinges represent the first and third quartile of the distribution. The central line represents the median value. The whiskers extend from the hinge to the largest or smallest value no further than 1.5 x IQR from the respective hinge (where IQR is the inter-quartile range or distance between the first and third quartiles).

https://doi.org/10.1038/s41564-024-01859-8



Extended Data Fig. 8 | See next page for caption.

https://doi.org/10.1038/s41564-024-01859-8

Extended Data Fig. 8 | GWAS power is not sufficient to capture genetic determinants of SigB pathway activity in our dataset. a. Estimation of effect on statistical power of higher effective population sizes for different minor allele frequencies (MAF) of unitigs (*i.d.* the lower frequency of presence or absence of a unitig in the tested dataset). Dashed vertical line indicates the size of the current dataset. **b**. Estimation of statistical power of each unitig in our dataset according to its frequency and effect size. All significant unitigs are shown (green), together with a random subset of non-significant unitigs (grey). The red dashed line indicates the sample size of the actual dataset. **c**. Cumulative heritability estimation in a linear model with all significant unitigs in +/- 10kb to *sigB* shown as R^2 or adjusted R^2 *vs* total heritability estimation from GWAS (dashed line, Methods).

nature portfolio

Corresponding author(s): Marc Lecuit

Last updated by author(s): Oct 10, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Confocal microcopy Image acquisition was performed with Zen Black v2.3
Data analysis	Data analysis was performed with R, image analysis was performed with Ilastik 1.4.0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The datasets generated during and/or analysed during the current study are available from the corresponding author. Genome information can be obtained on bigsdb.pasteur.fr or from the correspodning author. Certain genomes from isolates cannot be shared due to legal limitations as explained in the data availability statement.

A data availability statement has been added to the manuscript:
All datasets generated during and/or analyzed during the current study are available from the corresponding author. RNA-seq data is available on SRA under the accession number PRJNA1097069. Closed genomes of H/h isolates from pairs are available on bigsdb.pasteur.fr and raw reads are released on SRA (for accession code, see Supplementary Table 5). Genome data analyzed in this study were generated in the context of the epidemiological surveillance of listeriosis in France. As mentioned in the "Décret no. 2016–806 du 16 juin 2016 relatif aux centres nationaux de référence pour la lutte contre les maladies transmissibles" and in the "Arrêté du 16 juin 2016 fixant le cahier des charges des centres nationaux de références pour la lutte contre les maladies transmissibles", all samples collected at the NRCL as well as the data generated from these samples belong to the French Government and constitute a "national collection of biological resources of interest to public health", which has to remain under medical and industrial confidentiality, and therefore they cannot be made publicly available.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	NA
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	NA
Recruitment	Clinical isolates of Listeria monocytogenes were obtained through the activity of the French Center for Listeriosis
Ethics oversight	NA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For in vivo experiments, no sample size calculation was performed, and sample sizes were chosen based on preliminary experiments to enable statistically sound analysis. In practice, this means that between 6 - 14 animals were used per group, for large to small effect sizes, respectively. For the transcriptomics screen, a minimum number of 10 isolates per CC was estimated based on the differences between H and h isolates. For microscopy, the sample number was driven by image acquisition and quality. For qPCR experiments, no sample size estimation was performed.
Data exclusions	Data were excluded when positive and/or negative controls were not as expected for qPCRs or when mice reached ethical endpoints before the pre-determined timepoints of killing: they were terminated early without being included in data analysis.
Replication	All data was replicated at least twice; for transcriptomics screen, a representative subgroup was used for replication (2x)
Randomization	Isolates were randomized across groups for transcriptomics screen. For in vivo experiments, mice were randomized across cages. For image analysis, randomization was not possible, but same segmentation and quantification has been applied to all images. For Western Blots and growth curves, this was not relevant, as samples were processed in the same batches.
Blinding	All isolates in the transcriptomics screen were given a identifier and randomized. For image analysis of reporter system, blinding was not possible, but all data was analyzed in a batch and afterwards associated with isolate. For other experiments, blinding was not possible.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Primary antibody for WB : OpuCA (diluted 1:2000) Secondary Antibody: Goat anti-chicken IgY H and L chain (HRP), ref : ab6877 from abcam (diluted 1:2000)
Validation	delta SigB strain, which does not express OpuCA was used, as in the original publication of this antibody (DOI: 10.1155/2014/641647)

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	Lab intern stock of L2071 and L2071 hEcad cells. The establishement of this cell lines are described in Lecuit et al 1999, https://doi.org/10.1093/emboj/18.14.3956
Authentication	None of the cell lines were identified, but positive and negative controls were used to assess their specificity
Mycoplasma contamination	Cell lines were negative
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell line was used in this study

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	7-12 weeks-old female mice were used of the strain C57BL/6 mEcad E16P KI.
Wild animals	No wild animals were used
Reporting on sex	All animals used in in vivo study were female to reduce inter-individual differences and reduce therefore the number of animals needed to make statistically sound conclusions.
Field-collected samples	No field collected samples were involved.
Ethics oversight	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, authorized by the French Ministry of Research and registered under APAFIS #14644-2018041116183944 and #39306-2022111412359059 v2.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA