Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Resident bacteria and niche chemistry gatekeep host-specific colonization by commensal fungi

**Shipra Vaishnava**

shipra_vaishnava@brown.edu

Brown University https://orcid.org/0000-0002-5377-9586

**Geongoo Han**

Brown University

**Casidhe Hughes**

Brown University

**Alicia Xu**

Brown University

**Article**

**Keywords:**

**Posted Date:** November 13th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11540449/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** There is **NO** Competing Interest.

# Resident bacteria and niche chemistry gatekeep host-specific colonization by commensal fungi

Geongoo Han[1], Casidhe Hughes[1], Alicia Xu[1], Shipra Vaishnava[1,*]

[1] Department of Molecular Microbiology & Immunology, Brown University, Providence, RI 02912, USA

* Corresponding author: shipra_vaishnava@brown.edu

## Abstract

Bacterial colonization resistance against fungi is well established, but its role in fungal commensalism remains unclear. *Kazachstania pintolopesii*, a stomach-tropic fungal commensal, dominates the mouse gastrointestinal (GI) tract despite the presence of resident bacteria. To investigate how mouse-derived strains withstand bacterial colonization resistance, we compared mouse- and guinea pig-derived isolates of this species. Unlike the mouse isolate (Kp-M), the guinea pig isolate (Kp-GP) failed to colonize the mouse GI tract. Resident bacteria imposed this barrier: Kp-GP colonized germ-free mice, and vancomycin-mediated depletion of gastric *Lactobacillus* permitted robust colonization. Cell-free supernatant from aerobically grown *Lactobacillus johnsonii* isolated from the mouse stomach selectively killed Kp-GP through an $H_2O_2$- and low-pH-dependent mechanism, consistent with chemistry supported by the oxygen-containing gastric environment. Genomic and transcriptomic analyses revealed that Kp-GP is triploid, carries deletions in oxidative stress response genes, including the peroxiredoxin gene *AHP1*, and fails to induce *AHP1* expression following supernatant exposure. Pharmacological neutralization of $H_2O_2$ and reduction of gastric acidity allowed Kp-GP colonization in mice. Together, these findings identify gastric *Lactobacillus* and the oxidizing niche they establish as gatekeepers of fungal colonization and reveal that successful establishment by a fungus depends on its ability to withstand bacterial colonization resistance.

## Introduction

Commensal fungi can establish lasting residency in the gastrointestinal (GI) tract alongside complex bacterial communities. How resident bacteria shape fungal establishment, and which fungal traits permit persistence within particular hosts, are central questions in fungal commensalism. Early studies of the mouse stomach provided a foundation for investigating these relationships. In 1967, Savage and Dubos described an indigenous yeast, *Torulopsis pintolopesii*, now classified as *Kazachstania pintolopesii*, that colonized the GI tract and formed dense layers on the secretory gastric epithelium.[1] Savage subsequently showed that gastric *Lactobacillus* and this indigenous yeast occupy distinct gastric niches.[2] This early work provided one of the first descriptions of native fungal commensalism as a mucus-anchored, spatially organized yeast that shares its organ with a bacterial neighbor.

Much of our subsequent understanding of fungal-host interactions has emerged from studies of *Candida albicans*, a human-associated fungus whose GI colonization in mice is commonly studied under antibiotic-treated or germ-free conditions. These models have provided foundational insights into antifungal immunity, fungal morphogenesis, and mucosal defense. Studies in bacteria-replete hosts have also shown that hyphal factors, including candidalysin, promote *C. albicans* establishment under bacterial competition.[3] Complementary studies of indigenous commensal fungi offer an opportunity to extend this understanding to fungal persistence and host interactions within native microbial communities.

Although ultra-clean barrier facilities have depleted indigenous fungal populations from laboratory mice, recent studies have rediscovered an indigenous commensal

murine fungus, *K. pintolopesii*, in the GI tract of wild and pet store mice.[4-6] *K. pintolopesii* can establish colonization in the presence of an intact bacterial microbiota, without requiring antibiotic-induced dysbiosis. In the models examined, its colonization does not elicit systemic neutrophilia or Th17-cell expansion but instead activates type 2 immunity that protects against helminth infection.[4,5] It also promotes intestinal epithelial repair.[7] Together, these findings broaden the ecological and host-response settings available for investigating fungal commensalism and provide an opportunity to examine how an indigenous fungus establishes residency within a native microbial community.

However, comparisons between *K. pintolopesii* and *C. albicans* cannot readily separate the contributions of host association and fungal species identity. The two differ in morphogenesis and mechanisms of host interaction, so differences in establishment could reflect species-specific biology, host association, or both. Comparing isolates of a single species from different hosts offers an approach to examining host-associated variation while holding species identity constant.

*K. pintolopesii* has been reported in several rodents, including mice and guinea pigs.[8] Here we compared a mouse-derived isolate (Kp-M) with a guinea pig-derived isolate (Kp-GP) to investigate their ability to establish within the same mouse gut microbial environment. We combined colonization experiments, analyses of bacterial-fungal interactions, and comparative genomics to examine how resident bacteria shape fungal establishment and which fungal traits might permit persistence despite bacterial defenses. This within-species approach addresses an unresolved question: does an indigenous fungus establish residency because its native microbial community presents little resistance, or because it possesses the traits needed to withstand that resistance?

## Results

### Resident bacteria restrict mouse colonization by guinea pig-derived *K. pintolopesii*

To test whether resident bacteria differentially restrict host-associated isolates of the same fungal species, we compared mouse- and guinea pig-derived *K. pintolopesii* in the mouse GI tract. Our previously characterized mouse-derived isolate, Kp-M, provided a reference for stable colonization in this host.[4] We therefore used Kp-M as a reference to examine establishment by isolates from guinea pigs, a non-murine rodent in which *K. pintolopesii* has been reported in the esophagus and stomach.[9] Internal transcribed spacer (ITS) sequencing of guinea pig fecal samples identified *K. pintolopesii* as the predominant fungal species, consistent with its predominance in mice (Fig. 1a). We recovered colonies from guinea pig feces based on their morphology, and ITS sequencing assigned these isolates, together with Kp-M, to *K. pintolopesii* (Fig. 1b).

We first tested whether the guinea pig-derived isolate Kp-GP could colonize the mouse GI tract (Fig. 1c). In contrast to Kp-M, Kp-GP was not detected in feces at days 7 and 14 after a 3-day exposure in drinking water (Fig. 1d). When provided continuously in drinking water, Kp-M was detected in feces at approximately $10^7$-$10^8$ CFU/g from day 4 onward. Kp-GP was below the limit of detection at day 4 and reached only $10^4$-$10^5$ CFU/g at days 8 and 12 (Fig. 1e,f). To ask whether this difference was specific to these two isolates, we included a second mouse isolate (Kp-MTX) and a second guinea pig isolate (Kp-GP2). Both mouse isolates were readily detected, and neither guinea pig isolate was detected (Fig. 1d).

To distinguish bacterial restriction from an inability to grow in the mouse environment, we compared colonization in germ-free (GF) mice and in GF mice colonized

with the Altered Schaedler Flora (ASF), a defined consortium of eight bacterial species (Extended Data Fig. 1a). In GF mice, Kp-M and Kp-GP reached similar colonization levels at day 14 (~$10^7$ CFU/g). In ASF-colonized mice, Kp-M reached the same level, but Kp-GP was lower by approximately 1.5 $\log_{10}$ (Extended Data Fig. 1b). Thus, the colonization difference emerged in the presence of a defined bacterial community as well as in conventional mice.

We next tested whether antibiotic perturbation of the conventional microbiota permitted Kp-GP establishment. Mice received neomycin or vancomycin in drinking water, which preferentially target aerobic Gram-negative and Gram-positive bacteria, respectively. Antibiotics were given beginning 4 days before Kp-GP administration and continuing throughout the experiment. Kp-GP was given once, in the antibiotic water, on day 0 (Extended Data Fig. 1c). Fecal Kp-GP was measured at days 4, 7, 11, and 14. Vancomycin-treated mice exhibited significantly higher Kp-GP burden than untreated controls, at approximately $10^7$ CFU/g from day 7 onward, similar to the levels of Kp-GP in GF mice (Extended Data Fig. 1b,d). Neomycin-treated mice showed intermediate levels that did not differ significantly from those of untreated controls (Extended Data Fig. 1d). Together, these experiments support a role for resident bacteria in restricting Kp-GP establishment in mice.

**Vancomycin permits luminal recovery of Kp-GP but does not restore gastric mucus-associated localization**

We next asked whether antibiotic-permitted recovery reflected occupation of the same GI niches as Kp-M. We measured fungal loads in the stomach, duodenum, jejunum, ileum, and feces. Kp-GP remained undetectable across all GI tract sections in untreated mice,

even at day 5, whereas vancomycin administration allowed Kp-GP colonization to levels comparable to Kp-M at day 14 (Fig. 2a).

To characterize the bacterial community shifts induced by vancomycin treatments, we performed 16S rRNA gene sequencing on the stomach and fecal samples. Vancomycin administration was accompanied by an increase in an unclassified genus within the family *Enterobacteriaceae* (phylum *Proteobacteria*) in the feces (Extended Data Fig. 2a), and a marked reduction of the genus *Lactobacillus* in the stomach (Fig. 2b).

We then examined microbial localization by fluorescence in situ hybridization (FISH), combining an 18S fungal probe with LGC354A, which targets the *Lactobacillus-Enterococcus* cluster. In untreated mice, Kp-M occupied the mucus layer overlying the gastric mucosa (Fig. 2c), and was found in the colonic lumen alongside resident bacteria (Extended Data Fig. 3b). Kp-GP was not detected in either tissue (Fig. 2c and Extended Data Fig. 3b). Following vancomycin administration, Kp-M remained predominantly associated with the gastric mucus layer, with some cells in the lumen, and retained its luminal distribution in the colon. Kp-GP was recovered in both the stomach and colon but remained luminal, with no detectable cells in the gastric mucus layer (Fig. 2c and Extended Data Fig. 3b). Thus, vancomycin permitted luminal recovery of Kp-GP without restoring the gastric mucus-associated localization observed for Kp-M.

To examine whether carbon-source use could contribute to this spatial difference, we compared growth on defined substrates and mucin. On the Biolog PM1 plate, both Kp-M and Kp-GP grew on glucose, fructose, and mannose, and none of the other sugars tested, including N-acetylglucosamine, fucose, and galactose (Extended Data Fig. 3a). Neither isolate formed colonies on agar containing mucin as the sole carbon source,

whereas both grew on agar containing glucose (Extended Data Fig. 3b). Kp-M used only a few simple sugars such as glucose, which agrees with an earlier report that *K. pintolopesii* assimilated only glucose, mannose and fructose among the compounds tested.[10] These assays detected no difference between the isolates in use of the tested carbon sources and did not support mucin as a sole carbon source for either isolate.

Because Kp-M induces type 2 immunity in mice,[4] we also asked whether Kp-GP reproduced this response under vancomycin treatment, which permitted its recovery from the GI tract. Both isolates were administered in vancomycin-containing water, and immune phenotypes were assessed in peripheral blood on day 13 and the small intestinal lamina propria (SI-LP) on day 14 (Extended Data Fig. 4a). Neither isolate increased blood neutrophils or SI-LP Th17 cells, consistent with our previous findings for Kp-M (Extended Data Fig. 2b,c).[4] Kp-M increased SI-LP Th2 frequencies, whereas Kp-GP did not increase Th2 frequencies above controls (Extended Data Fig. 2d,e). Thus, Kp-GP did not reproduce the Th2 response observed with Kp-M under the tested treatment conditions.

**_Lactobacillus johnsonii_ supernatant preferentially kills Kp-GP through peroxide-dependent activity**

The reduction of gastric *Lactobacillus* during vancomycin treatment prompted us to test whether a mouse stomach-derived *Lactobacillus johnsonii* isolate differentially inhibited Kp-M and Kp-GP. We co-cultured *L. johnsonii* with each fungal isolate (Fig. 3a). Relative to growth in MRS medium alone, *L. johnsonii* co-culture reduced Kp-M growth by about 3-fold and Kp-GP by about 22-fold in CFU/mL (Fig. 3b).

To determine whether this difference depends on cell contact or on secreted factors, we exposed each isolate to cell-free supernatant (CFS) prepared from aerobic *L.*

*johnsonii* cultures (Fig. 3c). In medium containing 60% CFS and 40% fresh MRS, Kp-GP growth was significantly impaired, whereas Kp-M was unaffected (Fig. 3d). Direct bacterial-fungal contact was therefore not required for the differential inhibition. We next distinguished growth inhibition from killing. Undiluted CFS reduced Kp-GP CFU to near the detection limit within 4 h, whereas Kp-M declined more modestly over the same interval (Fig. 3e). CFS was therefore fungicidal toward Kp-GP and, to a lesser extent, Kp-M.

The mouse gastric lumen is acidic and, unlike the largely anaerobic distal gut, retains detectable oxygen (~25 mmHg),[11] and is dominated by *Lactobacillus*. Several antifungal mechanisms of *Lactobacillus* have been described, including $H_2O_2$-mediated oxidative stress, organic acid secretion and consequent acidification, secretion of other antimicrobial peptides, and competition for niche and nutrients.[12] Because CFS alone reproduced the difference, neither direct contact nor competition for nutrients was required in this assay.

Proteinase K treatment did not rescue Kp-GP, whereas catalase treatment prevented killing (Fig. 3f). Raising the CFS pH from 4.6 to 6.2 also prevented killing (Fig. 3f). Thus, activity was sensitive to catalase and pH neutralization but was not relieved by proteinase K under the conditions tested. To assess acidity and peroxide separately, we measured fungal growth at low pH or across a range of $H_2O_2$ concentrations. Both isolates grew normally at pH 3, showing that acidity alone was insufficient to reproduce killing under those conditions (Extended Data Fig. 5a). Both also grew normally at 0.5, 1, and 2 mM $H_2O_2$. Kp-GP growth was significantly impaired at 3 mM, whereas Kp-M remained largely unaffected up to 4 mM before declining sharply at 5 mM (Extended Data Fig. 5b).

CFS from anaerobic *L. johnsonii* cultures killed neither isolate, although a slight growth difference remained between them (Fig. 3g, h). Aerobic CFS contained approximately 0.8 mM $H_2O_2$, which was not detected in anaerobic CFS (Extended Data Fig. 5c). Anaerobic CFS was more acidic than aerobic CFS, with pH values of 3.9 and 4.6, respectively, indicating that acidity alone did not account for killing. Adding 0.5 mM $H_2O_2$ to anaerobic CFS reproduced the preferential killing of Kp-GP observed with aerobic CFS, whereas 1 mM $H_2O_2$ killed both isolates (Fig. 3i). By contrast, 0.5 mM $H_2O_2$ in yeast nitrogen base medium at pH 4 with 2% glucose killed neither isolate (Extended Data Fig. 5d). Together, these experiments identify peroxide-dependent killing in *L. johnsonii* CFS, while showing that the tested peroxide concentration at low pH was insufficient in defined medium. Additional components or properties of the bacterial supernatant were therefore required to reproduce the killing phenotype.

**Guinea pig-associated *K. pintolopesii* populations occur beyond the sampled animals, and guinea pig isolates carry candidate stress-response deletions**

Having established differential susceptibility to bacterial factors, we asked whether Kp-GP-like populations occurred beyond the animals sampled here and whether the guinea pig isolates carried genomic differences that might contribute to susceptibility. We first screened publicly available fecal metagenomes acquired from farm guinea pigs of 13 provinces in China (PRJDB38973).[13] After excluding detections that could not be separated from index crosstalk on shared sequencing lanes, *K. pintolopesii* sequences were detected in 21 of the 81 runs, from 8 of 13 provinces (Fig. 4a).

To assess similarity to Kp-GP, we analyzed 21 positive runs at positions where all copies of the Kp-GP genome carry one allele, and Kp-M carries another. Among them,

four runs from Hunan carried the Kp-GP allele at 62.1 to 83.3% of informative observations, while the other runs fell between 44.6 and 59.9%, and the seven mouse isolates between 0.78 and 10.41% (Fig. 4b). A *K. pintolopesii* resembling Kp-GP across the genome is therefore present in farmed guinea pigs in Hunan, in four independent animals, so it is not confined to the animals our isolates came from.

We next compared the genomes of four *K. pintolopesii* isolates from this study, two mouse-derived (Kp-M and Kp-MTX) and two guinea pig-derived (Kp-GP and Kp-GP2), with five publicly available genomes of *K. pintolopesii* isolated from mice (C89-2, ID1LA, IDNYC, QC1, and W26G).[5] The guinea pig isolates formed a clade separate from all seven mouse isolates, despite sharing more than 98% nuclear average nucleotide identity (ANI) (Extended Data Fig. 6a) and largely conserved synteny (Extended Data Fig. 6b). Because ANI reflects aligned sequences rather than regions absent from one genome, we also searched directly for differences in genomic content.

We mapped reads from each isolate separately to the Kp-GP and Kp-M assemblies and examined relative read depth. To interpret copy-number changes, we assessed ploidy by *k*-mer analysis and by flow cytometry calibrated to the *k*-mer result, as ploidy had not previously been reported for *K. pintolopesii*. Kp-GP and Kp-GP2 were triploid, whereas four of the seven mouse isolates, including Kp-M and Kp-MTX, were diploid. Ploidy could not be determined for the remaining three mouse isolates (Extended Data Fig. 6c-e).

For each isolate, we normalized read depth to its genome-wide median to infer how many copies are present and scanned the assembly with sliding windows of 1 to 20 kb. Regions that were depleted or elevated in one host group but not the other were taken

as candidates. The bidirectional scan returned six deletions and four absences, and one in each direction was junction-supported: the *HOG1-AHP1* interval and the *CHK1* interval (Fig. 4c).

The first deletion spanned 13.8 kb and contained tRNA-Arg, *HOG1*, *CCW12_2*, *AHP1*, and the uncharacterized gene FUN_002907 (Extended Data Fig. 7a). This deletion was present in both guinea pig isolates and absent from all mouse isolates (Fig. 4d). Read depth inside the interval was 0.51 and 0.72 of the genome median in Kp-GP and Kp-GP2, consistent with the loss of one of three copies in a triploid genome (Fig. 4d and Extended Data Fig. 7b). The deletion was present between two copies of a 316-bp repeat, and reads spanning the junction were found in both Kp-GP and Kp-GP2 (Fig. 4d and Extended Data Fig. 7c,d). *HOG1* encodes a stress-activated MAP kinase,[14] and *AHP1* encodes a peroxiredoxin.[15] Given the greater $H_2O_2$ sensitivity of the guinea pig isolates (Fig. S5B), this interval represented a candidate contributor to differential susceptibility.

Mapping to the Kp-M assembly detected a second deletion in guinea pig isolates. This 1.9 kb deletion removed the start codon and residues 1 to 520 of the 570-amino-acid CHK1 protein, and it included the entire kinase domain (Extended Data Fig. 8a). No read mapped within the deleted interval in either Kp-GP or Kp-GP2, consistent with loss from all three copies (Extended Data Fig. 8b-d). *CHK1* encodes a DNA damage checkpoint kinase,[16] and the gene was present in all seven mouse isolates (Extended Data Fig. 8e).

We also examined regulatory differences that would not alter gene content by comparing stress response element (STRE) and Yap1-binding sites in stress-response gene promoters. Eleven genes carried promoter-site differences shared by both guinea pig isolates and in the opposite state in all mouse isolates. These promoter regions carried

17 differences (10 at STRE sites and 7 at Yap1 binding sites, 9 losses and 8 gains), and all 17 were uniform across all genome copies of every isolate (Extended Data Fig. 9). Because gains and losses were mixed, the analysis identified site-specific changes rather than a directional shift in these regulatory elements.

**Transcript abundance supports the deletion of stress response genes in the Kp-GP genome**

We next asked whether the genomic differences were accompanied by differences in transcript abundance. The inferred *HOG1-AHP1* deletion affected one copy of five neighboring genes, whereas the *CHK1* deletion affected all copies in Kp-GP and Kp-GP2. To assess transcription within these intervals and at genes with altered promoter sites, we performed RNA sequencing of Kp-M and Kp-GP before and after exposure to *L. johnsonii* CFS.

Before CFS exposure, transcript abundance for the four protein-coding genes in the *HOG1-AHP1* interval, *AHP1*, *HOG1*, *CCW12_2*, and FUN_003454 (FUN_002907 in Kp-GP assembly; Extended Data Fig. 8a), was lower in Kp-GP than in Kp-M (Fig. 5a). We then corrected expression for the inferred copy-number difference. After correction, *CCW12_2* and FUN_003454 did not differ between isolates, whereas *AHP1* and *HOG1* remained lower in Kp-GP, with $Padj < 0.001$ for each gene (Fig. 5a).

*CCW12_1*, a paralog of *CCW12_2* located on a different scaffold, was more highly expressed in Kp-GP, indicating that reduced expression was not shared by both paralogs. *AHP1* was particularly relevant to the peroxide-dependent killing phenotype because it encodes a peroxiredoxin. Kp-GP combined an inferred reduction in *AHP1* copy number with lower expression after copy-number correction. *HOG1* showed the same pattern,

although both genes reside within the same deletion interval. These results identify reduced dosage and transcript abundance at candidate stress-response genes without establishing their causal contribution to susceptibility.

Consistent with the *CHK1* truncation, almost no *CHK1* reads were detected in Kp-GP before and after CFS exposure, including reads from the retained 3'-end (Fig. 5b). In Kp-M, *CHK1* expression was higher after exposure, but the increase was not statistically significant (Supplementary Table 1).

We also assessed CFS-responsive expression within the *HOG1-AHP1* interval (Fig. 5c and Supplementary Table 1). *HOG1* did not respond in either isolate. *AHP1* and *CCW12_2* increased significantly after exposure in Kp-M, whereas neither showed a significant increase in Kp-GP. FUN_003454 increased in both isolates. *CCW12_1*, located outside the interval, was not significantly altered in either isolate.

Finally, we asked whether the STRE or Yap1-binding-site differences identified in eleven promoters (Extended Data Fig. 9) were associated with transcriptional differences. Ten genes could be evaluated; *CRH1* was excluded because its annotated gene model appeared to join two genes. None of the ten genes differed significantly between isolates in their response to CFS (Extended Data Fig. 10a). Although several genes differed in baseline expression, the direction did not consistently follow promoter-site gain or loss. Among three genes whose promoters differed only by loss of one STRE site, *HSP26* was lower in Kp-GP, as predicted from site loss, whereas *SOD1* and *CWP2* were higher (Extended Data Fig. 10b). Thus, the promoter-site changes did not provide a consistent explanation for the observed transcriptional differences.

**Catalase and acid suppression partially relieve Kp-GP exclusion in mice**

We returned to the mouse model to test whether interventions targeting peroxide activity and gastric acidity could increase Kp-GP recovery. Because catalase treatment and pH neutralization prevented killing in bacterial supernatant, one group received catalase-supplemented drinking water and another received catalase together with omeprazole, a proton pump inhibitor that lowers gastric acid secretion. Fungal burdens were measured three days after Kp-GP gavage (Fig. 6a).

Kp-GP was below the limit of detection in control mice. With catalase alone, Kp-GP was detected in the feces at $10^4$-$10^5$ CFU/g, significantly above controls, and in the stomach at $10^3$-$10^4$ CFU/g, although not significantly different from controls (Fig. 6b,c). With catalase and omeprazole together, Kp-GP levels were significantly higher than in controls in both stomach and feces, and higher than with catalase alone in the stomach (Fig. 6b,c).

The recovered burdens remained below those observed with vancomycin treatment or Kp-M colonization. Thus, the interventions partially relieved Kp-GP exclusion rather than restoring the phenotype observed with Kp-M. Together, these findings show that targeting peroxide activity and gastric acidity increases short-term Kp-GP recovery in mice, while leaving additional barriers to establishment unresolved.

## Discussion

Successful colonization by an indigenous fungus does not imply that its host microbiota is uniformly permissive. Bacteria have been reported to restrict fungal colonization in mouse models,[17,18] yet *K. pintolopesii*, a murine commensal fungus, colonizes the mouse GI tract in the presence of commensal bacteria.[4,5] Here, a guinea pig-derived isolate of the same species failed to establish sustained colonization in conventional mice after exposure ended. Similar burdens of Kp-M and Kp-GP in GF mice, reduced Kp-GP burden in ASF-colonized mice, and recovery after vancomycin treatment implicated resident bacteria in this difference. Vancomycin treatment, which depleted *Lactobacillus*, allowed Kp-GP to colonize. The CFS of *L. johnsonii* killed Kp-GP in vitro but had little effect on Kp-M. These findings point to *Lactobacillus* and its products as a barrier that Kp-GP cannot tolerate. *Lactobacillus* has long been known to restrict yeast in the mouse stomach, because depleting indigenous lactobacilli allowed the resident yeast *K. pintolopesii* to spread onto the non-secreting epithelium within 24 h, and returning lactobacilli displaced it again.[2] Our comparison suggests that bacterial communities can influence not only where an indigenous fungus resides, but also whether different isolates of that species establish within the same host environment.

This interpretation complements work showing that fungal commensalism depends on both bacterial inhibitory activities and fungal traits that permit persistence. Recent studies have linked microbiota-derived short-chain fatty acids to *C. albicans* colonization through metabolic reprogramming, impaired hexose uptake, and intracellular acidification, and identified valeric acid-dependent intracellular acidification as a mechanism restricting *Candida parapsilosis*.[19,20] Conversely, hyphal factors including candidalysin can promote *C. albicans* establishment in bacteria-replete hosts,

demonstrating that fungal traits associated with pathogenicity can also support commensal competition.[3] Differences in antibiotic-independent persistence and susceptibility to the host antimicrobial peptide CRAMP have also been documented among *C. albicans* isolates.[21] Our study adds a host-associated comparison within an indigenous fungal species, linking differential establishment to bacterial restriction, susceptibility to peroxide-dependent activity, and distinct gastric localization without changing fungal species identity.

*Lactobacillus*-mediated inhibition provides a plausible component of this restriction, although the responsible activities vary across organisms and environments. Vaginal isolates of *Lactobacillus crispatus* and *Lactobacillus fermentum* and their culture supernatants inhibited growth and hyphae formation of *C. albicans.*[22] *L. fermentum* limited the growth of *Candida glabrata* by ergosterol depletion,[23] and *L. johnsonii* isolated from the murine oral cavity damaged the cell wall integrity of *C. albicans* through chitinase activity.[24] In our experiments, killing by the CFS required $H_2O_2$, because catalase abolished it and anaerobic CFS killed neither Kp-M nor Kp-GP. However, $H_2O_2$ alone was not sufficient even at low pH. Adding $H_2O_2$ to anaerobic CFS reduced both isolates in a dose-dependent manner, with Kp-GP more sensitive than Kp-M. The killing of Kp-GP under this condition was comparable to that by aerobic CFS, which is consistent with $H_2O_2$ accounting for most of the difference between aerobic and anaerobic CFS. Because Kp-M was also killed at higher $H_2O_2$, the mouse and guinea pig isolates are different in degree and not in kind. Increasing the pH of aerobic CFS also abolished the Kp-GP killing. Together, these results suggest that CFS contains a second, unidentified factor that acts together with $H_2O_2$ and depends on acidity. Lactic acid and acetic acid are candidates for

this factor, since *Lactobacillus* releases both, but we did not measure their concentrations in the CFS, and other products may also contribute. Measuring them by metabolomics in aerobic and anaerobic CFS would allow lactate and acetate to be tested at the measured concentrations, alone or together with $H_2O_2$, at acidic and neutral pH in Kp-M and Kp-GP. In mice, catalase and omeprazole allowed partial colonization by Kp-GP, which is compatible with a role for $H_2O_2$ and acidity in the stomach, although we did not confirm that these treatments changed $H_2O_2$ or pH in the stomach.

The deletion of *HOG1* and *AHP1* in the guinea pig isolates is a candidate basis for their greater sensitivity to $H_2O_2$ compared with mouse isolates. *HOG1* is required for resistance to $H_2O_2$ in *C. albicans*,[25] whereas *AHP1* is specific for organic peroxides in *S. cerevisiae*,[15] so the two genes may contribute differently. Kp-M carries two copies of this interval, which is the number that remains in the triploid guinea pig isolates after the loss, although at a lower share of their genome. The remaining copies of *HOG1* and *AHP1* were transcribed at lower levels than the copy number alone predicts. After exposure to the CFS, *AHP1* increased in Kp-M but not in Kp-GP, which suggests that Kp-GP does not induce *AHP1* in response to the supernatant. The genomic and transcript data link the deletion to the phenotype but do not show that it is the cause. A second deletion removed *CHK1* from all copies in the guinea pig isolates. Its expression in Kp-M was higher after CFS exposure, although not significantly, so whether *CHK1* contributes to the response to $H_2O_2$ remains unclear and needs further study. Differences at STRE and Yap1 binding sites did not point in one direction, so they do not suggest a general change in stress regulation. Together, these observations identify candidate stress-response differences,

but their causal contribution remains untested because genetic deletion or restoration of the interval has not been established due to lack of genetic tools for this fungus.

The spatial findings distinguish recovery from the GI tract from occupation of a mucosal niche. Kp-M occupied the mucus layer of the stomach, whereas Kp-GP was observed exclusively in the lumen and only after vancomycin treatment. A similar pattern has been described for *Lactobacillus*, in which strains from other hosts reached high numbers in the mouse stomach but, with one exception, did not form layers on the gastric epithelium as mouse strains did.[26,27] In the colon, Kp-M and, after vancomycin treatment, Kp-GP were both observed in the lumen. This difference raises the question of why Kp-M occupies the gastric mucus layer rather than the lumen. One possibility is that vancomycin-sensitive bacteria drive Kp-M into the mucus layer, but this is unlikely because the major population of Kp-M remained in the mucus layer even after vancomycin treatment (Fig. 2C). Another possibility is that Kp-M feeds on components of the mucus layer, but this is also unlikely because the carbon source utilization assays revealed that Kp-M cannot use mucin glycan sugars for its growth (Fig. S4A and B). A third possibility is that the mucus layer protects Kp-M from bacteria, which these results do not exclude. However, small molecules diffuse through the mouse gastric mucus layer,[28] so $H_2O_2$ from bacteria may also reach the fungus there, and the gastric epithelium itself releases $H_2O_2$ through dual oxidases, a process that limits colonization of the mucus layer by *Helicobacter felis* in mice.[29] Kp-M may be exposed to host-derived $H_2O_2$ within the mucus layer, so an isolate that is more sensitive to $H_2O_2$, such as Kp-GP, might be unable to enter the mucus layer even after the bacteria are depleted. However, Kp-M, which tolerates higher concentrations of $H_2O_2$ in vitro, could persist under such conditions.

Why Kp-M is found there rather than in the lumen remains unexplained, and retention in the mucus layer, which could prevent washout with luminal contents, is one possibility that we did not test. How Kp-M enters the mucus layer is unknown. That Kp-M was luminal in the colon suggests that mucus occupancy depends on conditions specific to the stomach. Stable occupancy of the gastric mucus layer may allow indigenous fungi to act as a reservoir that supplies the lower intestine, which could explain how Kp-M persists without antibiotics. These ideas are speculative, because we did not measure $H_2O_2$ in the mucus layer, did not test the contribution of the *HOG1-AHP1* deletion in Kp-GP, and did not examine whether the stomach supplies the lower intestine.

The scope of this study is also limited by the isolates that we could examine. Both guinea pig isolates came from animals housed in the same environment, so we could not conclude whether the failure to colonize mice and the greater sensitivity to $H_2O_2$ are general features of guinea pig-derived *K. pintolopesii*. To partly address this point, we surveyed fecal metagenomes from farmed guinea pigs. *K. pintolopesii* was detected in approximately one-quarter of the surveyed runs, and sequences from four independent guinea pigs in Hunan more closely resembled Kp-GP at the informative positions than did the mouse isolates. These observations support the existence of related populations beyond the source animals, but no run exactly matched Kp-GP, the *HOG1-AHP1* deletion was not established in those populations, and their colonization or susceptibility phenotypes were not tested. Isolates from additional environments and animal sources are needed to test generality. Reciprocal host experiments would also be required to distinguish host-dependent fitness from the isolate differences observed within mice. Accordingly, the present study demonstrates host-associated variation in a shared mouse

environment, not reciprocal host specificity or adaptation. In addition, Kp-GP did not reproduce the Th2 response induced by Kp-M in the small intestine. Whether this difference reflects localization, isolate-specific properties, or other features of the experimental setting remains unresolved; the data do not establish a causal connection between gastric mucus association and intestinal Th2 responses.

Together, our findings support a model in which resident bacteria differentially constrain establishment by host-associated isolates of an indigenous fungal species. Kp-M showed greater tolerance of peroxide-dependent activity in *L. johnsonii* supernatant, while relieving Kp-GP exclusion did not reproduce the gastric mucus association of Kp-M. These results suggest that resident bacteria can act as a gatekeeper that selects among isolates of one fungal species. Such host-specific colonization would be hard to detect by comparing different species, such as *C. albicans* and *K. pintolopesii*, because the many differences between species would confound the comparison. Broader sampling and functional tests will establish whether the same relationships operate across hosts and whether manipulating bacterial activities can alter fungal persistence.

## Methods

### Mice

Conventional wild-type C57BL/6 mice were purchased from The Jackson Laboratory and Taconic Biosciences or were bred in-house at Brown University. Germ-free (GF) mice were of the wild-type C57BL/6 background and were raised and bred in flexible film isolators within a gnotobiotic facility at Brown University. Altered Schaedler flora (ASF) mice were established by colonizing GF mice with ASF. Briefly, ASF fecal pellets (Taconic Biosciences) were resuspended in pre-reduced sterile PBS (50 mg/mL) followed by filtration using the 70 µm cell strainer. 100 µL of filtrate was gavaged into GF mice. Fresh fecal pellets from pet store mice (Texas) and pet store guinea pigs were shipped to Brown University and used for ITS sequencing and fungal isolation. All animal experiments were approved by the Institutional Animal Care and Use Committee (IACUC) at Brown University (24-03-0001).

### Fungal isolation from feces of pet store mice and guinea pigs

To isolate *K. pintolopesii*, fecal pellets from pet store mice and guinea pigs were suspended in sterile PBS and plated on Sabouraud Dextrose (SD) agar containing streptomycin (40 mg/L), followed by incubation for 2 days at 37°C. Colonies morphologically resembling previously isolated mouse *K. pintolopesii* (Kp-M) colonies[4] were selected for further analysis. To identify the isolates, the ITS region was amplified using the Phire Green Hot Start II PCR Master Mix (Thermo Fisher Scientific) with the ITS1-F/ITS4 primer pair (ITS1-F: CTTGGTCATTTAGAGGAAGTAA, ITS4: TCCTCCGCTTATTGATATGC). Amplicons were sequenced via Sanger sequencing (GENEWIZ), and species identification was performed using MOLE-BLAST against the

NCBI ITS RefSeq database. The exported phylogenetic tree from MOLE-BLAST was visualized using the R package ggtree (v3.6.2).

**Fungal culture**

All fungal isolates were cultured in SD medium. A single colony of *K. pintolopesii* was inoculated into SD broth and cultured for 1 day at 37°C with shaking (200 rpm). This culture was subsequently cultured into fresh medium and grown for an additional day under the same conditions.

**Fungal colonization**

For intestinal colonization, mice received *K. pintolopesii* isolates by either a single oral gavage ($2 \times 10^6$ CFU in 200 µL PBS) or administration in drinking water (~$10^6$ CFU/mL) for three consecutive days, after which mice were switched to regular drinking water. For continuous administration, mouse isolate (Kp-M) and guinea pig isolate (Kp-GP) were provided in drinking water (~$10^6$ CFU/mL) renewed every 4 days throughout the experiment.

For colonization with Kp-GP, antibiotic water was prepared using neomycin (1 mg/mL) or vancomycin (0.5 mg/mL), each supplemented with 2.5% glucose. Control mice received 2.5% glucose water alone. Antibiotic treatment began four days prior to fungal inoculation. On the day of inoculation, *K. pintolopesii* was added directly to the antibiotic water at ~$10^6$ CFU/mL. The following day, this water was replaced with fresh antibiotic water without fungi, and fresh antibiotic water was subsequently provided twice weekly for the rest of the experiment.

Before inoculation, fungal cultures were washed twice with sterile PBS and resuspended in PBS or drinking water, depending on the administration route, at the appropriate concentration.

**Fungal burden quantification**

To determine fungal burden, fecal pellets and intestinal contents were resuspended in sterile PBS supplemented with ampicillin (500 mg/L), streptomycin (250 mg/L), and kanamycin (225 mg/L). Tissue samples were homogenized in the same antibiotic-supplemented PBS. All samples were serially diluted in the same antibiotic-supplemented PBS and plated on SD agar containing streptomycin (40 mg/L), followed by incubation for 2 days at 37°C.

**Immune cell isolation and flow cytometry**

Peripheral blood lymphocytes and small intestinal lamina propria lymphocytes were isolated as previously described.[4,30,31] The isolated cells were stained with the corresponding cocktail of fluorescently labeled antibodies (CD4, CD11b, CD45, CD90.2, B220, Ly6C, Ly6G, and ST2) and biotin-conjugated anti-mouse lineage (CD3ε and CD19) and streptavidin. Cell viability was determined using Fixable Viability Dye eFluor 450 (Thermo Fisher Scientific). After surface staining, cells were fixed and permeabilized using eBioscience Foxp3/Transcription Factor Staining Buffer Set (Thermo Fisher Scientific), followed by intracellular staining with the corresponding cocktail of fluorescently labeled antibodies (Foxp3, GATA3, and RORγT). The data were acquired on the Aurora flow cytometer (Cytek) and were analyzed with FlowJo 10 software (BD).

**Carbon source utilization assay**

Utilization of 95 carbon sources was tested using Biolog PM1 microplates according to the manufacturer's protocol. Plates were incubated in a BioTek Epoch 2 microplate reader at 37°C with shaking, and the $OD_{750}$ was read every 15 min for 24 h. The growth of each isolate on each substrate was calculated as the difference in $OD_{750}$ between 24 h and the first reading, and the corresponding change in the negative control well was subtracted.

Yeast nitrogen base without amino acids (Millipore Sigma) agar was prepared with no carbon source, 1% (w/v) glucose, or 1% (w/v) type III mucin from porcine stomach. Kp-M and Kp-GP cultures ($OD_{600}$ 1.0) were serially diluted, and 5 µL of each dilution was spotted onto the plates. The plates were then incubated at 37°C for 5 days and colony growth was recorded.

### ***Lactobacillus* co-culture and cell-free supernatant assays**

*Lactobacillus johnsonii* was isolated from the mouse stomach and used for all co-culture and killing assays described below. For all experiments, overnight cultures of Kp-M and Kp-GP were inoculated into fresh SD medium and grown until $OD_{600}$ reached 1.0. An overnight culture of *L. johnsonii* was inoculated into fresh De Man, Rogosa, and Sharpe (MRS) medium and grown until $OD_{600}$ reached 1.0.

***Preparation of cell-free supernatant (CFS).*** CFS of *L. johnsonii* was prepared under aerobic or anaerobic conditions. For aerobic conditions, glycerol stocks were inoculated into fresh MRS medium at a 1:1000 ratio in tubes (20 mL medium in a 50 mL tube with a loose cap), followed by incubation for 21 h at 37°C with shaking (200 rpm). For anaerobic conditions, glycerol stocks were inoculated into fresh MRS medium at the same ratio in culture tubes with a GasPak EZ Anaerobe Container System Sachet (BD) and incubated statically for 21 h at 37°C. Supernatants were collected by centrifugation at 5,000 rpm for

10 min at 4°C and filtered through a 0.22 μm syringe filter. The $H_2O_2$ concentrations in aerobic and anaerobic CFS were measured using the Amplex Red Hydrogen Peroxide/Peroxidase Assay Kit (Thermo Fisher Scientific), following the manufacturer's protocol.

***Co-culture.*** Kp-M or Kp-GP (final $OD_{600}$ 0.1) was either grown with *L. johnsonii* (final $OD_{600}$ 0.05) in fresh MRS medium or grown in diluted aerobic CFS (60% CFS, 40% fresh MRS) for 18 h at 37°C with shaking (200 rpm). For $H_2O_2$ dose-response experiments, hydrogen peroxide was added to fresh MRS medium at final concentrations of 0.5, 1, 2, 3, 4, and 5 mM, and Kp-M or Kp-GP (final $OD_{600}$ 0.1) was cultured in each condition for 18 h at 37°C with shaking (200 rpm). Cultures were serially diluted in sterile PBS and plated on SD agar containing streptomycin (40 mg/L), followed by incubation for 2 days at 37°C for CFU enumeration.

***Killing assay*.** To assess fungal survival, Kp-M or Kp-GP (final $OD_{600}$ 0.1) was inoculated into aerobic or anaerobic *L. johnsonii* CFS and was cultured for 4 h at 37°C with shaking (200 rpm). Cultures were sampled at 0, 2, and 4 h, serially diluted in sterile PBS, and plated on SD agar containing streptomycin (40 mg/L) for CFU enumeration. For proteinase K treatment, aerobic CFS was treated with proteinase K (final concentration 0.1 mg/mL) for 1 h at 37°C. The sample was then heated for 10 min at 95°C to inactivate the enzyme prior to the assay. For catalase treatment, aerobic CFS was treated with catalase (final concentration 200 U/mL) for 30 min at 37°C before the killing assay. HCl or NaOH was used to adjust the pH of *L. johnsonii* CFS and yeast nitrogen base (with 2% glucose). For $H_2O_2$ add-back experiments, $H_2O_2$ was added to anaerobic CFS at final concentrations of 0, 0.25, 0.5, or 1 mM, and fungal survival was assessed at 2 h.

## Genome analysis of *K. pintolopesii* isolates

***DNA extraction and Sequencing.*** Fungal genomic DNA was extracted from fungal culture of Kp-M, Kp-MTX, Kp-GP, and Kp-GP2 as described in Jiang et al.[32] Briefly, fungal cells (~$10^8$ CFU) were collected by centrifugation at 4,000 rpm for 5 min. Cell pellets were suspended in 50 mM Tris buffer (pH 7.5) supplemented with 1 mM EDTA, 0.2% β-mercaptoethanol, and 1,000 units/mL of lyticase (Millipore Sigma), followed by incubation for 30 min at 37°C. Genomic DNA was then extracted using a Quick-DNA Fungal/Bacterial Miniprep Kit (Zymo Research).

The extracted DNA samples were submitted to SeqCenter or Plasmidsaurus for library construction and sequencing. For Illumina sequencing, libraries were prepared using the Illumina DNA prep kit with IDT for Illumina UDI indices and sequenced on an Illumina NextSeq 2000 or NovaSeq X Plus (2 × 151 bp paired-end). For Oxford Nanopore Technology (ONT) sequencing, samples were prepared using a Genomic DNA by Ligation kit (Oxford Nanopore) and sequenced on a Nanopore GridION or PromethION.

***Genome assembly and annotation.*** Adapter and low-quality sequences were trimmed from Illumina reads using Trimmomatic (v0.39).[33] ONT reads were trimmed and filtered using Porechop (v0.2.4)[34] and NanoFilt (v2.8.0).[35] *k*-mer frequencies were counted from Illumina reads using Jellyfish (v.2.3.1)[36] and visualized using GenomeScope[37] to estimate genome size.

*De novo* assembly was performed using Flye (v.2.9)[38] based on ONT data. The assembly was subsequently polished with Illumina reads using medaka (v1.6.0; https://github.com/nanoporetech/medaka), Racon (v1.5.0),[39] and Pilon (v1.24).[40] Assembly quality was assessed using QUAST (v5.0.0)[41] and BUSCO (v5.2.2).[42] Genome

annotation was performed using the Funannotate pipeline (v1.8.9; https://github.com/nextgenusfs/funannotate). Mitochondrial genomes were removed from the assembly before genome analysis. Genome assemblies and raw sequencing reads of five mouse-derived *K. pintolopesii* isolates (C89-2, ID1LA, IDNYC, QC1, and W26G) were downloaded from the NCBI database (PRJNA962598),[5] and included them in comparative genome analysis. NCBI currently lists these isolates as *Arxiozyma pintolopesii*,[43] but we use the taxon name *Kazachstania pintolopesii* throughout this paper.

***Average nucleotide identity (ANI) and synteny.*** ANI was computed with fastANI (v1.34).[44] Whole-genome synteny between Kp-M and Kp-GP was determined with minimap2 (v2.28).[45] Alignments shorter than 3 kb were excluded. Kp-M scaffolds were ordered by number and concatenated, and Kp-GP contigs were placed by the position of their alignment anchor on the Kp-M scaffolds.

***Public metagenome analysis.*** 81 available metagenomic runs from farmed guinea pigs (PRJDB38973)[13] were screened at 1 M read pairs for *K. pintolopesii*. Reads were mapped with bwa mem (v0.7.18; https://github.com/lh3/BWA) to the Kp-GP assembly with its mitochondrial genome. A read was counted as Kp-confident when it was a primary alignment with mapping quality of at least 20, an aligned length of at least 100 bp, and an identity of at least 97%, on a nuclear contig of at least 20 kb. The mitochondrial sequence was kept in the reference so that mitochondrial reads would not be forced onto nuclear contigs, and reads on it were tabulated separately because mitochondrial copy number does not track the nuclear genome. A run was called positive when it carried at least 5 Kp-confident reads in at least 5 distinct 1 kb bins, leaving out one recurrent conserved bin. Because index hopping can move a small number of reads between libraries

sequenced on one lane, positive runs carrying 20 or fewer confident reads were removed when they shared a sequencing lane with a run carrying 100 or more.

To compare the Kp-confident reads of each positive run, the 21 positive runs were aligned to the Kp-M assembly. A position in the Kp-M assembly was considered informative when all three genome copies of Kp-GP carried the same allele and that allele differed from the Kp-M base. This filter generated 100,371 single-nucleotide positions throughout the genome, and indels were not called. Reads were aligned singly to the Kp-M assembly with bwa mem and scored at the informative positions. Each observation was quantified as the percentage of observations carrying the Kp-GP allele. The nine isolates of this study were subsampled to 500,000 reads and used as controls.

***Structural variant scan.*** Reads of isolates were mapped to a reference assembly, and the depth was compared across the reference in two directions. Using the Kp-GP assembly as the reference, we looked for regions lost or gained in the guinea pig isolates, and using the Kp-M assembly as the reference, we looked for regions lost in them. Depth was calculated in non-overlapping 1 kb bins on contigs of at least 20 kb with samtools (v1.23.1)[46] at a mapping quality of at least 20, and each isolate's depth was divided by its own genome-wide median. Bins were combined into sliding windows that advance by 1kb. The Kp-GP-referenced scan used windows of 1, 5, 9, and 20 kb, and the Kp-M-referenced scan used 1 kb windows only. The reference distribution for each scale came from contigs of at least 100 kb, excluding the 20 kb at each end.

In the Kp-GP-referenced scan, a window was a deletion candidate when both guinea pig isolates fell below the 1$^{st}$ percentile of their own distribution and all mouse isolates were at or above its 5$^{th}$ percentile, and an increase candidate when both guinea

pig isolates exceeded the 99$^{th}$ percentile and all mouse isolates were at or below its 95$^{th}$ percentile. In the Kp-M-referenced scan, a window was an absence candidate when normalized depth was below 0.15 in both guinea pig isolates and each of the other mouse isolates was at or above its 5$^{th}$ percentile. Windows that overlapped or lay within 1 kb of each other were merged into one interval. To avoid artifacts from repetitive or poorly assembled regions, candidates were excluded from the junction test if they lay within 5,000 bp of a contig end, if a self-BLAST against the same assembly covered at least 30% of the interval at 98% identity or higher, or if ONT read depth exceeded three times its genome-wide median in Kp-M or Kp-MTX. Candidates that passed these exclusions were called interpretable.

***Promoter analysis.*** The promoters of 70 stress-response genes were examined to find differences in binding sites for the stress response element (STRE) and for Yap1 between mouse and guinea pig isolates (Supplementary Table 2). The promoter region of each gene was defined as the 700 bp upstream of the start codon. Promoter regions were aligned across the nine isolates with MAFFT (v7.526).[47] Binding sites for the STRE and for Yap1 were called at $p < 0.001$ against a zero-order nucleotide background estimated from promoter sequences, using a position weight matrix scorer that reproduced the output of FIMO (v5.4.1).[48] The STRE matrix was the union of the Msn2 and Msn4 matrices (JASPAR MA0341.1 and MA0342.1), and Yap1 sites used JASPAR MA0415.2. The scanner reproduced FIMO output on the same sequences without differences. Hits of the same motif family on the same strand that overlapped by at least one base pair were merged into one site, and an interval called on both strands was counted once. A site was treated as the same site across isolates when the alignment columns it occupied

overlapped. A difference was counted when the allele carried by both guinea pig isolates was fixed in reads, and all mouse isolates carried the other state.

**Ploidy determination**

Ploidy of *K. pintolopesii* isolates was determined by *k*-mer analysis of Illumina reads and by flow cytometry.

***k-mer analysis.*** 21-mer frequency histograms were generated from Illumina reads with KMC (v3.2.4).[49] Ploidy was assigned with Smudgeplot (v0.5.3)[50] applied to FastK (v1.2; https://github.com/thegenemyers/FASTK) 21-mer tables, with the lower coverage cutoff set at the local minimum of the KMC histogram between the error peak and the main coverage peak. Smudgeplot classified each pair of covarying *k*-mers by the coverage ratio of its two alleles, which is 1:1 in a diploid heterozygote and 1:2 in a triploid heterozygote. Ploidy was not determined if Smudgeplot detected no heterozygous *k*-mer pairs.

***Flow cytometry.*** Fungal cells were collected and fixed in 70% ethanol overnight at 4°C. Fixed cells were washed twice in 50 mM Tris buffer A (50 mM Tris, 5 mM EDTA, pH 8.0), resuspended in 2 mg/mL RNase A in buffer A, and incubated for 2 h at 37°C. Cells were then pelleted, resuspended in 5 mg/mL pepsin in 55 mM HCl, and incubated for 1 h at 37°C. After two washes in buffer B (50 mM Tris, 5 mM EDTA, pH 7.5), cells were resuspended in 1 µM SYTOX Green Nucleic Acid Stain in buffer B, stained overnight at 4°C protected from light. Samples were analyzed without a further wash. All steps were performed in a 96-well U-bottom plate, with centrifugation at 500 × *g* for 5 min at 4°C. Each isolates was prepared as one unstained control and three stained replicates. Data were acquired on the Aurora flow cytometer (Cytek) and were analyzed with FlowJo 10

682 software (BD). After singlet gating, the G1 peak was identified as the lower of the two
683 major peaks in the SYTOX Green histogram, the second peak lying at approximately
684 twice its channel value, consistent with G2/M. The median G1 signal of each isolate was
685 normalized to the assembly-based monoploid genome size and expressed relative to Kp-
686 GP and Kp-GP2, which were set to 1.00, because the ploidy of those isolates was
687 determined as triploid (3n) from *k*-mer analysis. On this scale, a diploid isolate is expected
688 at 0.667 and a haploid isolate at 0.333. Because the relationship between SYTOX Green
689 signal and DNA content depends on the species, calibration was performed within *K.*
690 *pintolopesii* and not against a heterologous reference. *Saccharomyces cerevisiae*
691 BY4741 (haploid, 1n) was included to confirm that the assay resolves a known haploid
692 and was not used for calibration. Flow cytometry therefore measured ploidy relative to the
693 *k*-mer-assigned triploid reference and did not test the ploidy of Kp-GP and Kp-GP2
694 independently. Ploidy was assigned to the nearest expected value.

695 **DNA extraction for microbiome and mycobiome analysis**

696 The stomach was flushed with sterile PBS to collect luminal contents, and the flushed
697 stomach was then scraped to collect mucosal contents. Luminal and mucosal fractions
698 were pooled and centrifuged at 5,000 rpm for 10 min before DNA extraction. Fecal
699 samples were directly collected from the rectum. Samples (~50 mg) were suspended in
700 50 mM Tris buffer (pH 7.5) supplemented with 1 mM EDTA, 0.2% β-mercaptoethanol,
701 and 1,000 units/mL of lyticase (Millipore Sigma), followed by incubation for 30 min at 37°C.
702 After incubation, genomic DNA was extracted from samples using a Quick-DNA
703 Fecal/Soil Microbe Microprep Kit (Zymo Research) according to the manufacturer's
704 protocol.

**16S rRNA gene sequencing and bacterial community analysis**

The V4/V5 region of the bacterial 16S rRNA gene was amplified using the Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific) with the 518F/926R primer pair (518F: CCAGCAGCYGCGGTAAN, 926R: CCGTCAATTCNTTTRAGT), each conjugated with an overhang adapter sequence for Illumina MiSeq. DNA libraries were constructed and sequenced on the Illumina MiSeq platform (2 × 300 bp paired-end) at the Rhode Island Genomics and Sequencing Center.

The raw sequencing reads were processed using the QIIME 2 pipeline (v2023.9).[51] DADA2 (v1.28.0)[52] was used to perform denoising and to construct the amplicon sequence variants (ASVs) table. The pre-trained Naïve Bayes classifier trained on the SILVA database (v132)[53] was used for taxonomic assignment. Singletons and all features annotated as mitochondria or chloroplasts were removed from the table, and the abundance of bacterial taxa was expressed as a percentage of the total 16S rRNA gene sequences.

**ITS sequencing and fungal community analysis**

The ITS1 region of the fungal DNA was amplified using Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific) with the ITS1-F/ITS2 primer pair (ITS1-F: CTTGGTCATTTAGAGGAAGTAA, ITS2: GCTGCGTTCTTCATCGATGC),[54,55] each conjugated with an overhang adapter sequence for Illumina MiSeq. DNA library construction and sequencing were performed as described above for 16S rRNA gene sequencing.

The raw sequencing reads were processed using the QIIME 2 pipeline (v2023.9).[51] The reads were trimmed using q2-itsxpress.[56] Trimmed sequence reads were denoised,

and sequence variants were identified by DADA2 (v1.28.0).[52] Taxonomy was assigned using a pre-trained Naïve Bayes classifier trained on the UNITE database (v8.2).[57] Singletons were removed from the table, and the abundance of fungal taxa was expressed as a percentage of total sequence reads.

**RNA-seq analysis**

Overnight cultures of Kp-M and Kp-GP were inoculated into fresh SD medium and grown until $OD_{600}$ reached 1.0. Kp-M or Kp-GP (final $OD_{600}$ 0.1) was inoculated into the aerobic *L. johnsonii* CFS and was incubated for 20 min at 37°C with shaking (200 rpm). Kp-M and Kp-GP were sampled before (T0) and after the CFS exposure (T20), and the pellets were collected by centrifugation at 4,000 rpm for 5 min. Cell pellets were suspended in 50 mM Tris buffer (pH 7.5) supplemented with 1 mM EDTA, 1 M sorbitol, 0.2% β-mercaptoethanol, and 1,000 units/mL of lyticase (Millipore Sigma), followed by incubation for 30 min at 37°C. The pellets were collected at 300 × *g* for 5 min at 4°C, and RNA was then extracted using an RNeasy mini plus kit (Qiagen).

The extracted RNA samples were submitted to Plasmidsaurus for library construction and sequencing. The mRNA was enriched by poly(A) selection and sequenced as 3′ end RNA-seq (single-end with 14 nt unique molecular identifiers) on an Illumina NovaSeq X Plus, and read lengths were 50 to 94 bp. Reads were trimmed using Trimmomatic (v0.41),[33] and aligned to the Kp-M assembly using the STAR aligner (v2.7.10b). Reads were then deduplicated by unique molecular identifiers with UMI-tools (v1.1.6)[58] and counted with featureCounts (v2.1.1).[59] Each gene model was extended at its 3‘-end by up to 500 bp before counting. DESeq2 (v1.38.1)[60] was used to identify differentially expressed genes when *Padj* < 0.05 using the design isolate + condition +

isolate:condition, with Kp-M and T0 as reference. Log2 fold changes were shrunken using the R package ashr (v2.2.63). Gene names are those of the closest *S. cerevisiae* ortholog, and gene functions are inferred from *S. cerevisiae*. Differences in gene copy number between the two isolates were corrected using genomic read depth. For each gene, we calculated the ratio of DNA read depth in Kp-GP to that in Kp-M and smoothed it with a running median over five adjacent genes. Genes were treated as copy number altered when at least two adjacent genes showed an absolute log2 ratio of 0.4 or more, which is close to the smallest one-copy difference between a diploid and a triploid genome. Correction factors were restricted to a range of log2 ratio -2 to 2.

**Fluorescence *in situ* hybridization (FISH)**

Stomach and colon tissues were fixed in Methacarn (Bioworld) and embedded in paraffin. Sections (7 µm) were deparaffinized in xylene and rehydrated through a graded ethanol series (95% and 90%) to $H_2O$. Sections were then hybridized with a universal fungal probe (Alexa 555-CTCTGGCTTCACCCTATTC) targeting the 18S rRNA gene. A probe targeting low-G+C Gram-positive bacteria (LGC354A; Alexa 647-TGGAAGATTCCCTACTGC), including the *Lactobacillus-Enterococcus* cluster, was used to visualize these bacteria. FITC-conjugated wheat germ agglutinin (WGA) and Ulex europaeus agglutinin (UEA) (Lectin Kit I, Vector Laboratories) were used to stain the mucus layer. DAPI was used to counterstain nuclei.

**Catalase and omeprazole administration**

Mice received one of three treatment combinations: regular drinking water with vehicle (control), catalase-supplemented drinking water with vehicle, or catalase-supplemented drinking water with omeprazole, starting one day before Kp-GP inoculation until the day

of sacrifice (day 3). Kp-GP was inoculated into mice by a single oral gavage ($2 \times 10^6$ CFU in 200 µL PBS).

Catalase was dissolved in autoclaved tap water at 2,000 U/mL and filtered through a 0.22 µm filter. Catalase-supplemented water was covered with foil to protect it from light and provided daily.

Omeprazole was dissolved in 10% DMSO in PBS and filtered through a 0.22 µm filter. Omeprazole (40 mg/kg) or vehicle (10% DMSO in PBS) was administered intraperitoneally in a volume of 300 µL once daily.

**Statistical analysis**

All statistical analyses and plotting were performed in R (v4.2.2; https://www.R-project.org) and Prism (v11.1.0; GraphPad). Normality of the data was assessed using the Shapiro-Wilk test. If data were normally distributed, parametric tests were used (unpaired two-tailed t-test for two groups; one-way ANOVA with Tukey's multiple comparisons test for more than two groups). If data were not normally distributed, non-parametric tests were used (Mann-Whitney test for two groups; Kruskal-Wallis test with Dunn's multiple comparisons test for more than two groups). *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

**Data availability**

Raw sequence reads from 16S and ITS sequencing have been deposited in the NCBI Sequence Read Archive (SRA) under accession number PRJNA1544112. The draft genome sequences of *K. pintolopesii* isolates and raw sequencing reads have been deposited in the NCBI GenBank and SRA under accession number PRJNA954910.

Bulk RNA-seq data have been deposited in the NCBI GEO database under accession number GSE350879.

## Code availability

The code used for the genomic analyses in this study is available on GitHub at https://github.com/goo9han/kp-m-gp-genome-analysis.[61]


## Acknowledgements

Schematics were created with BioRender.com.

## Funding

This work was supported by NIDDK R01DK113265 and P20GM109035 (pilot project).


## Author information

### *Authors and Affiliations*


**Department of Molecular Microbiology & Immunology, Brown University, Providence, RI 02912, USA**

Geongoo Han, Casidhe Hughes, Alicia Xu, Shipra Vaishnava


### *Contributions*

Conceptualization: S.V. and G.H.; methodology: S.V., G.H., and C.H.; investigation: S.V., G.H., C.H., and A.X.; resources: S.V., and G.H.; funding acquisition: S.V.; supervision: S.V.; writing – original draft: S.V., and G.H.; writing – review and editing: S.V., and G.H.

### *Corresponding authors*

Correspondence to Shipra Vaishnava

## Ethics declarations

### *Competing interests*

The authors declare no competing interests.

## References


1 Savage, D. C. & Dubos, R. J. Localization of indigenous yeast in the murine stomach. *J Bacteriol* **94**, 1811-1816 (1967). https://doi.org:10.1128/jb.94.6.1811-1816.1967

2 Savage, D. C. Microbial interference between indigenous yeast and lactobacilli in the rodent stomach. *J Bacteriol* **98**, 1278-1283 (1969). https://doi.org:10.1128/jb.98.3.1278-1283.1969

3 Liang, S. H. *et al.* The hyphal-specific toxin candidalysin promotes fungal gut commensalism. *Nature* **627**, 620-627 (2024). https://doi.org:10.1038/s41586-024-07142-4

4 Han, G. *et al.* A host-adapted commensal fungus from pet store mice drives type 2 immunity and cross-kingdom protection. *iScience* **29** (2026). https://doi.org:10.1016/j.isci.2026.117169

5 Liao, Y. *et al.* Fungal symbiont transmitted by free-living mice promotes type 2 immunity. *Nature* (2024). https://doi.org:10.1038/s41586-024-08213-2

6 Oh, J. H. *et al.* Naturalized immune responses are stable over years in a colony of laboratory mice with wild-derived microbiota. *Immunity* **58**, 2305-2319 e2305 (2025). https://doi.org:10.1016/j.immuni.2025.07.004

7 Gao, Y. *et al.* Fungal commensal promotes intestinal repair via its secreted peptide in mice. *Nat Microbiol* **11**, 476-491 (2026). https://doi.org:10.1038/s41564-025-02233-y

8 Kurtzman, C. P. *et al.* Multigene phylogenetic analysis of pathogenic candida species in the Kazachstania (Arxiozyma) telluris complex and description of their

ascosporic states as Kazachstania bovina sp. nov., K. heterogenica sp. nov., K. pintolopesii sp. nov., and K. slooffiae sp. nov. *J Clin Microbiol* **43**, 101-111 (2005). https://doi.org:10.1128/JCM.43.1.101-111.2005

9 Yoon, C. S. & Tannock, G. W. Microbes associated with the epithelium of the oesophagus and stomach of guinea pigs (Cavia porcellus). *Can J Microbiol* **24**, 1099-1101 (1978). https://doi.org:10.1139/m78-182

10 Artwohl, J. E. & Savage, D. C. Determinants in microbial colonization of the murine gastrointestinal tract: pH, temperature, and energy-yielding metabolism of Torulopsis pintolopesii. *Appl Environ Microbiol* **37**, 697-703 (1979). https://doi.org:10.1128/aem.37.4.697-703.1979

11 Friedman, E. S. *et al.* Microbes vs. chemistry in the origin of the anaerobic gut lumen. *Proc Natl Acad Sci U S A* **115**, 4170-4175 (2018). https://doi.org:10.1073/pnas.1718635115

12 Vazquez-Munoz, R. & Dongari-Bagtzoglou, A. Anticandidal Activities by Lactobacillus Species: An Update on Mechanisms of Action. *Front Oral Health* **2**, 689382 (2021). https://doi.org:10.3389/froh.2021.689382

13 Shi, Y. *et al.* Extensive cross-species transmission of pathogens and antibiotic resistance genes in mammals neglected by public health surveillance. *Cell* **188**, 6591-6605 e6514 (2025). https://doi.org:10.1016/j.cell.2025.08.016

14 Hohmann, S. Osmotic stress signaling and osmoadaptation in yeasts. *Microbiol Mol Biol Rev* **66**, 300-372 (2002). https://doi.org:10.1128/MMBR.66.2.300-372.2002

15 Lee, J., Spector, D., Godon, C., Labarre, J. & Toledano, M. B. A new antioxidant with alkyl hydroperoxide defense properties in yeast. *J Biol Chem* **274**, 4537-4544 (1999). https://doi.org:10.1074/jbc.274.8.4537

16 Sanchez, Y. *et al.* Control of the DNA damage checkpoint by chk1 and rad53 protein kinases through distinct mechanisms. *Science* **286**, 1166-1171 (1999). https://doi.org:10.1126/science.286.5442.1166

17 Fan, D. *et al.* Activation of HIF-1alpha and LL-37 by commensal bacteria inhibits Candida albicans colonization. *Nat Med* **21**, 808-814 (2015). https://doi.org:10.1038/nm.3871

18 Guinan, J., Wang, S., Hazbun, T. R., Yadav, H. & Thangamani, S. Antibiotic-induced decreases in the levels of microbial-derived short-chain fatty acids correlate with increased gastrointestinal colonization of Candida albicans. *Sci Rep* **9**, 8872 (2019). https://doi.org:10.1038/s41598-019-45467-7

19 Mishra, A. A. *et al.* Microbiota-derived short-chain fatty acids mediate Candida albicans gastrointestinal colonization resistance. *Cell Host Microbe* **34**, 1302-1319 e1311 (2026). https://doi.org:10.1016/j.chom.2026.05.007

20 Yasuma-Mitobe, K. *et al.* Intracellular acidification by microbiota-derived valeric acid facilitates trans-kingdom ecology limiting Candida parapsilosis colonization. *Cell Host Microbe* **34**, 1320-1332 e1326 (2026). https://doi.org:10.1016/j.chom.2026.05.008

21 McDonough, L. D. *et al.* Candida albicans Isolates 529L and CHN1 Exhibit Stable Colonization of the Murine Gastrointestinal Tract. *mBio* **12**, e0287821 (2021). https://doi.org:10.1128/mBio.02878-21

22 Jang, S. J., Lee, K., Kwon, B., You, H. J. & Ko, G. Vaginal lactobacilli inhibit growth and hyphae formation of Candida albicans. *Sci Rep* **9**, 8121 (2019). https://doi.org:10.1038/s41598-019-44579-4

23 Zangl, I. *et al.* Limosilactobacillus fermentum Limits Candida glabrata Growth by Ergosterol Depletion. *Microbiol Spectr* **11**, e0332622 (2023). https://doi.org:10.1128/spectrum.03326-22

24 Vazquez-Munoz, R., Thompson, A., Sobue, T. & Dongari-Bagtzoglou, A. Lactobacillus johnsonii is a dominant Lactobacillus in the murine oral mucosa and has chitinase activity that compromises fungal cell wall integrity. *mBio* **15**, e0241624 (2024). https://doi.org:10.1128/mbio.02416-24

25 Alonso-Monge, R. *et al.* The Hog1 mitogen-activated protein kinase is essential in the oxidative stress response and chlamydospore formation in Candida albicans. *Eukaryot Cell* **2**, 351-361 (2003). https://doi.org:10.1128/EC.2.2.351-361.2003

26 Kotarski, S. F. & Savage, D. C. Models for study of the specificity by which indigenous lactobacilli adhere to murine gastric epithelia. *Infect Immun* **26**, 966-975 (1979). https://doi.org:10.1128/iai.26.3.966-975.1979

27 Lin, J. H.-C. & Savage, D. C. Host specificity of the colonization of murine gastric epithelium by lactobacilli. *FEMS Microbiology Letters* **24**, 67-71 (1984).

28 Baumgartner, H. K. & Montrose, M. H. Regulated alkali secretion acts in tandem with unstirred layers to regulate mouse gastric surface pH. *Gastroenterology* **126**, 774-783 (2004). https://doi.org:10.1053/j.gastro.2003.11.059

29 Grasberger, H., El-Zaatari, M., Dang, D. T. & Merchant, J. L. Dual oxidases control release of hydrogen peroxide by the gastric epithelium to prevent Helicobacter felis infection and inflammation in mice. *Gastroenterology* **145**, 1045-1054 (2013). https://doi.org:10.1053/j.gastro.2013.07.011

30 Beura, L. K. *et al.* Lymphocytic choriomeningitis virus persistence promotes effector-like memory differentiation and enhances mucosal T cell distribution. *J Leukoc Biol* **97**, 217-225 (2015). https://doi.org:10.1189/jlb.1HI0314-154R

31 Grizotte-Lake, M. *et al.* Commensals Suppress Intestinal Epithelial Cell Retinoic Acid Synthesis to Regulate Interleukin-22 Activity and Prevent Microbial Dysbiosis. *Immunity* **49**, 1103-1115 e1106 (2018). https://doi.org:10.1016/j.immuni.2018.11.018

32 Jiang, T. T. *et al.* Commensal Fungi Recapitulate the Protective Benefits of Intestinal Bacteria. *Cell Host Microbe* **22**, 809-816 e804 (2017). https://doi.org:10.1016/j.chom.2017.10.013

33 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014). https://doi.org:10.1093/bioinformatics/btu170

34 Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genom* **3**, e000132 (2017). https://doi.org:10.1099/mgen.0.000132

35 De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M. & Van Broeckhoven, C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* **34**, 2666-2669 (2018). https://doi.org:10.1093/bioinformatics/bty149

36 Marcais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**, 764-770 (2011). https://doi.org:10.1093/bioinformatics/btr011

37 Vurture, G. W. *et al.* GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics* **33**, 2202-2204 (2017). https://doi.org:10.1093/bioinformatics/btx153

38 Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* **37**, 540-546 (2019). https://doi.org:10.1038/s41587-019-0072-8

39 Vaser, R., Sovic, I., Nagarajan, N. & Sikic, M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res* **27**, 737-746 (2017). https://doi.org:10.1101/gr.214270.116

40 Walker, B. J. *et al.* Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* **9**, e112963 (2014). https://doi.org:10.1371/journal.pone.0112963

41 Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072-1075 (2013). https://doi.org:10.1093/bioinformatics/btt086

42 Manni, M., Berkeley, M. R., Seppey, M., Simao, F. A. & Zdobnov, E. M. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Mol Biol Evol* **38**, 4647-4654 (2021). https://doi.org:10.1093/molbev/msab199

43 Liu, F. *et al.* Saccharomycetaceae: delineation of fungal genera based on phylogenomic analyses, genomic relatedness indices and genomics-based synapomorphies. *Persoonia* **52**, 1-21 (2024). https://doi.org:10.3767/persoonia.2024.52.01

44 Jain, C., Rodriguez, R. L., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* **9**, 5114 (2018). https://doi.org:10.1038/s41467-018-07641-9

45 Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100 (2018). https://doi.org:10.1093/bioinformatics/bty191

46 Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10** (2021). https://doi.org:10.1093/gigascience/giab008

47 Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772-780 (2013). https://doi.org:10.1093/molbev/mst010

48 Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017-1018 (2011). https://doi.org:10.1093/bioinformatics/btr064

49 Kokot, M., Dlugosz, M. & Deorowicz, S. KMC 3: counting and manipulating k-mer statistics. *Bioinformatics* **33**, 2759-2761 (2017). https://doi.org:10.1093/bioinformatics/btx304

50 Ranallo-Benavidez, T. R., Jaron, K. S. & Schatz, M. C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat Commun* **11**, 1432 (2020). https://doi.org:10.1038/s41467-020-14998-3

51 Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**, 852-857 (2019). https://doi.org:10.1038/s41587-019-0209-9

52 Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581-583 (2016). https://doi.org:10.1038/nmeth.3869

53 Yilmaz, P. *et al.* The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res* **42**, D643-648 (2014). https://doi.org:10.1093/nar/gkt1209

54 White, T. J., Bruns, T., Lee, S. & Taylor, J. in *PCR Protocols: A Guide to Methods and Applications* (eds M. A. Innis, D. H. Gelfand, J. J. Sninsky, & T. J. White) 315-322 (Academic Press, 1990).

55 Gardes, M. & Bruns, T. D. ITS primers with enhanced specificity for basidiomycetes--application to the identification of mycorrhizae and rusts. *Mol Ecol* **2**, 113-118 (1993). https://doi.org:10.1111/j.1365-294x.1993.tb00005.x

56 Rivers, A. R., Weber, K. C., Gardner, T. G., Liu, S. & Armstrong, S. D. ITSxpress: Software to rapidly trim internally transcribed spacer sequences with quality scores for marker gene analysis. *F1000Res* **7**, 1418 (2018). https://doi.org:10.12688/f1000research.15704.1

57 Nilsson, R. H. *et al.* The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res* **47**, D259-D264 (2019). https://doi.org:10.1093/nar/gky1022

58 Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res* **27**, 491-499 (2017). https://doi.org:10.1101/gr.209601.116

59 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014). https://doi.org:10.1093/bioinformatics/btt656

60 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014). https://doi.org:10.1186/s13059-014-0550-8

61 Han, G. Analysis code for the comparative genomics of *Kazachstania pintolopesii* from mice and guinea pigs. *GitHub*, https://github.com/goo9han/kp-m-gp-genome-analysis (2026).

## Figures

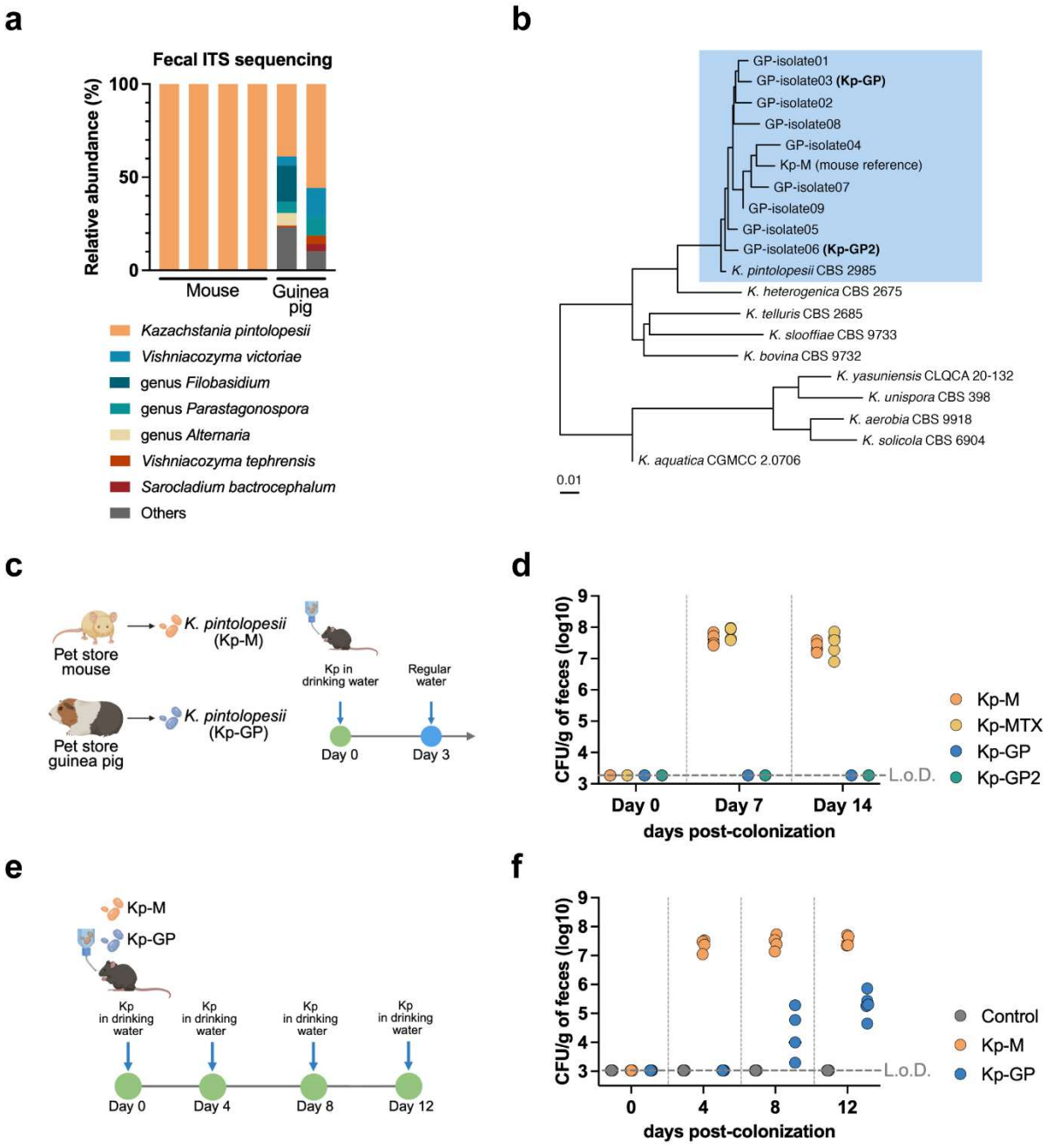



**Fig. 1: A guinea pig-derived *K. pintolopesii* isolate fails to colonize the mouse GI tract. a,** Fecal ITS sequencing of pet store mice ($n$ = 4) and guinea pigs ($n$ = 2). Bars indicate the relative abundance of fungal species. **b,** MOLE-BLAST result of guinea pig isolates based on the ITS sequences. MOLE-BLAST result was visualized as a

phylogenetic tree. Kp-M sequence was used as a reference for the mouse-derived *K. pintolopesii* isolate. **c,** Scheme of the Kp-M and Kp-GP colonization in conventional laboratory mice. Mice received *K. pintolopesii* isolates by administration in drinking water for three consecutive days, after which mice were switched to regular drinking water. **d,** Fungal burden in the fecal samples of Kp-M (*n* = 5), Kp-MTX (*n* = 5), Kp-GP (*n* = 4), and Kp-GP2 (*n* = 5) at days 7 and 14 post-colonization. **e,** Scheme of the continuous Kp-M and Kp-GP administration. **f,** Fungal burden in the fecal samples of Kp-M (*n* = 4) and Kp-GP (*n* = 5) at days 4, 8, and 12 post-colonization during the continuous administration. L.o.D., limit of detection.

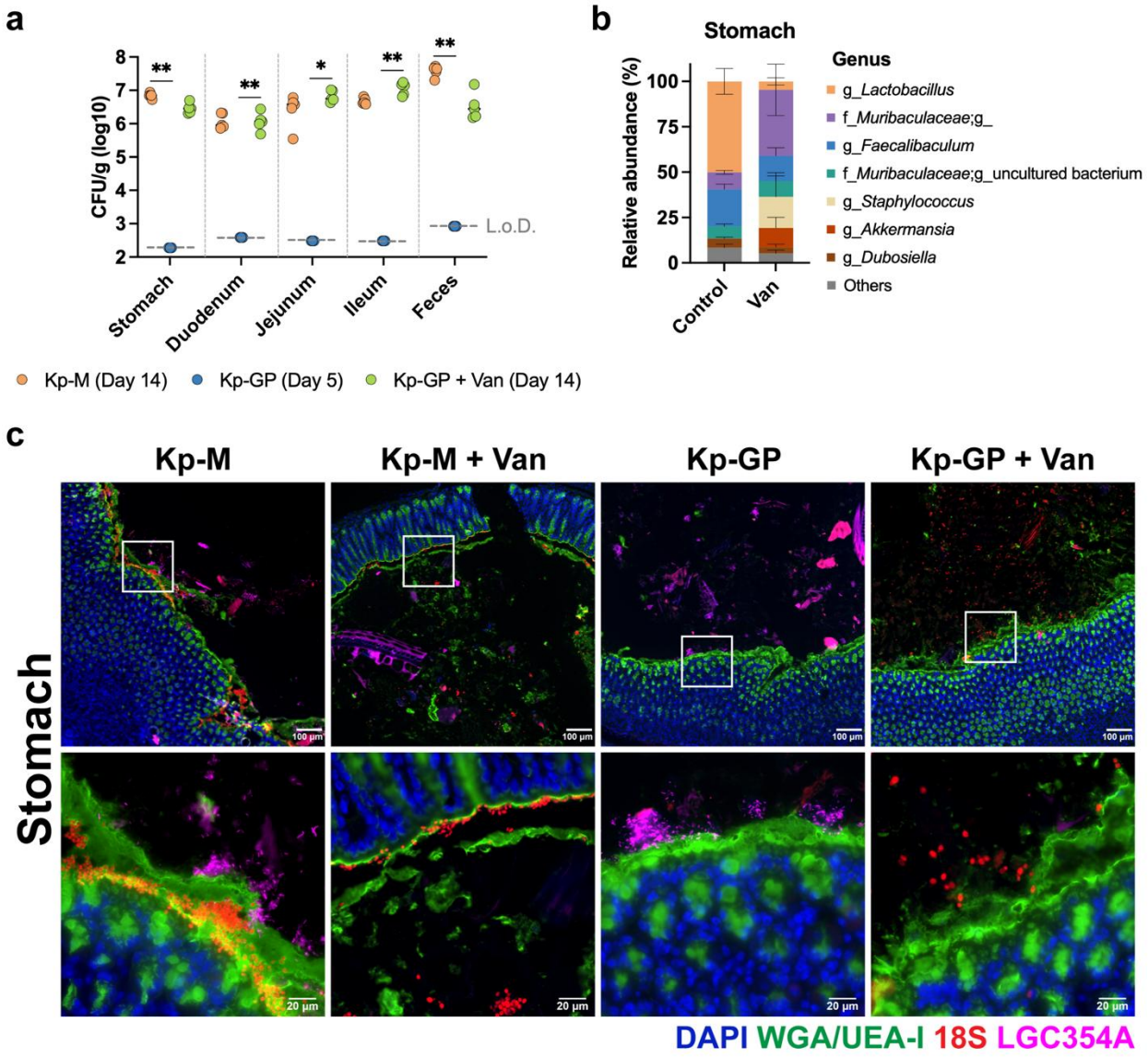



**Fig. 2: Vancomycin depletes gastric *Lactobacillus* and allows luminal but not mucosal colonization by Kp-GP. a,** Fungal burden in the GI tract of mice colonized with Kp-M (day 14; *n* = 5), Kp-GP (day 5; *n* = 3), and Kp-GP with vancomycin (Kp-GP + Van, day 14; *n* = 5). **b,** Relative abundance of bacterial genera in the stomach of control and vancomycin-treated mice from 16S rRNA gene sequencing (day 4; *n* = 3 per group). Bars indicate mean ± SEM. **c,** Representative images of the stomach of mice colonized with Kp-M or Kp-GP, with or without vancomycin by fluorescence in situ

hybridization (FISH). The lower row shows the boxed areas in the upper row. DAPI (blue) marks nuclei, WGA/UEA-I (green) marks the mucus layer, the 18S probe (red) marks fungi, and the LGC354A probe (magenta) marks the *Lactobacillus-Enterococcus* cluster. Scale bars, 100 μm for the upper row and 20 μm for the lower row. L.o.D., limit of detection. Significance was determined by the Kruskal-Wallis test with Dunn's multiple comparisons test. **P* < 0.05 and ***P* < 0.01.

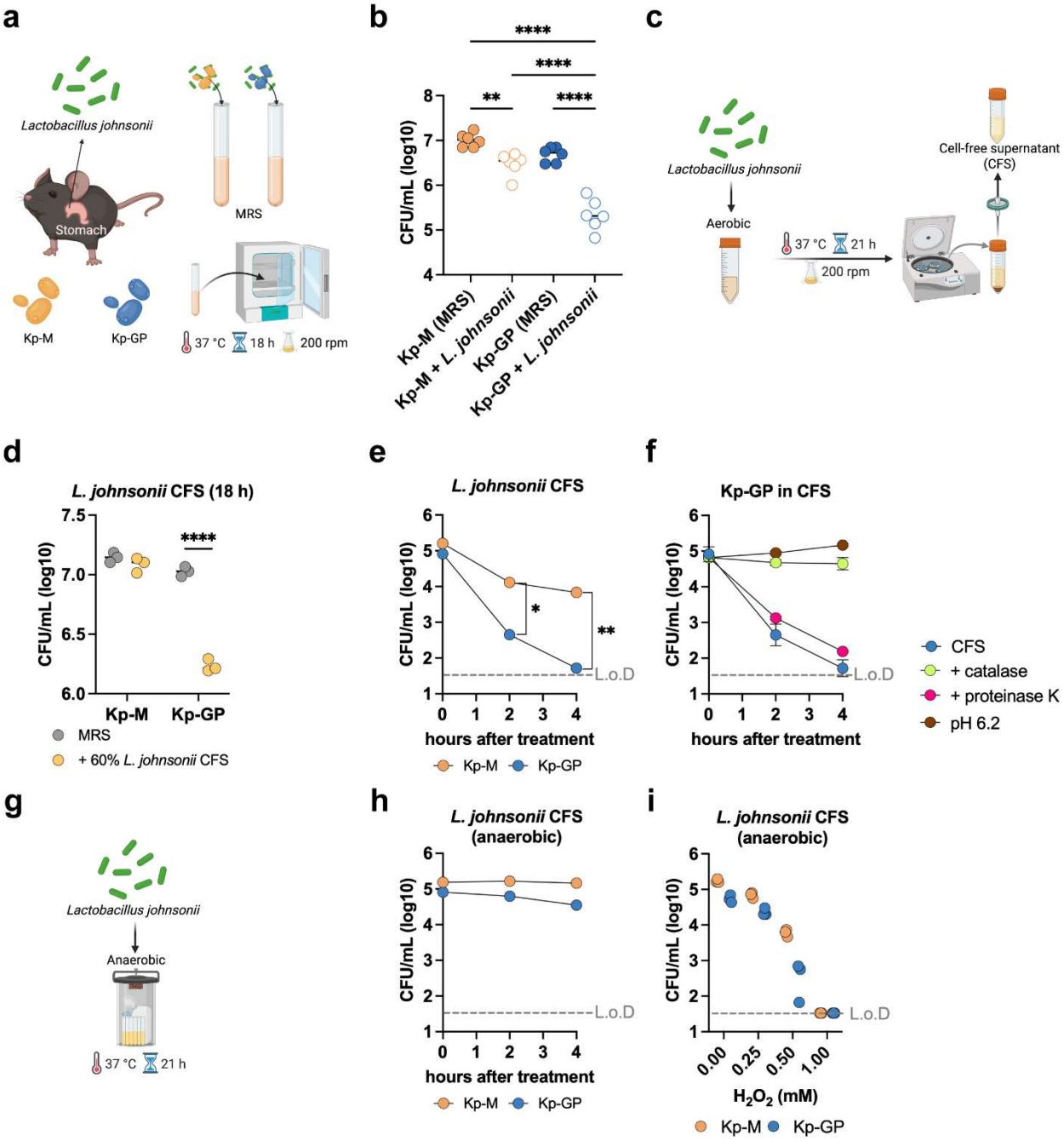



**Fig. 3: *L. johnsonii*-derived peroxide kills Kp-GP, but not Kp-M. a,** Scheme of the co-culture of Kp-M or Kp-GP with *L. johnsonii*, which was isolated from the mouse stomach. **b,** Growth of Kp-M and Kp-GP after co-culture with *L. johnsonii* for 18 h at 37°C. **c,** Preparation of aerobic cell-free supernatant (CFS) of *L. johnsonii*. **d,** Growth of Kp-M and Kp-GP in MRS medium or in 60% aerobic CFS for 18 h at 37°C. **e,**

Comparison of the survival of Kp-M and Kp-GP in aerobic CFS. **f,** Survival of Kp-GP in aerobic CFS, aerobic CFS treated with catalase or proteinase K, or aerobic CFS adjusted to pH 6.2. **g,** Preparation of anaerobic *L. johnsonii* CFS. **h,** Comparison of the survival of Kp-M and Kp-GP in anaerobic CFS. **i,** Survival of Kp-M and Kp-GP after 2 h in anaerobic CFS to which $H_2O_2$ was added at the concentrations shown. L.o.D., limit of detection. Significance between groups was determined by one-way ANOVA with Tukey's multiple comparisons test for **(b)**, and by an unpaired two-tailed t-test for **(d), (e),** and **(h)**. *$P < 0.05$, **$P < 0.01$, and ****$P < 0.0001$.

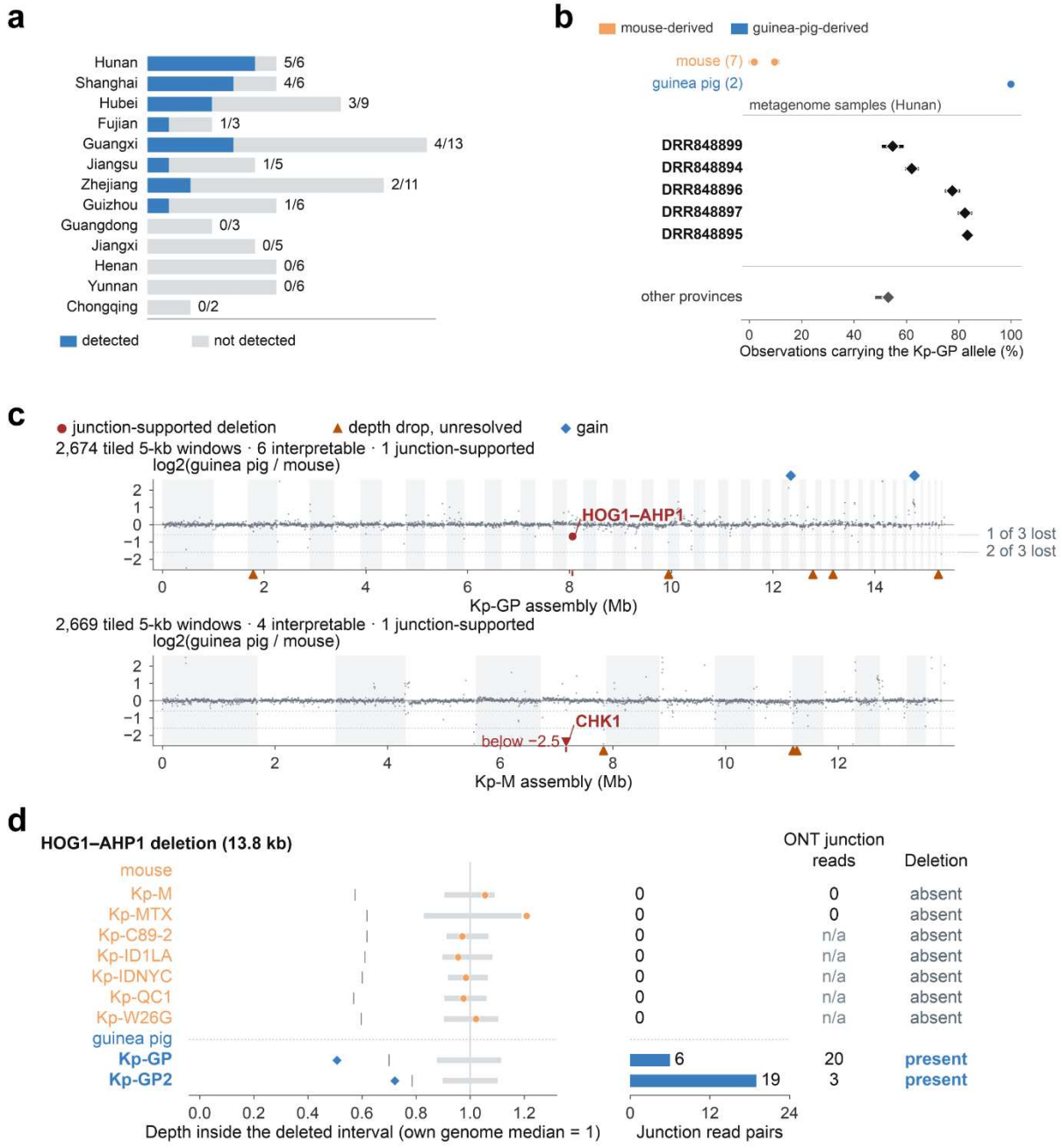



**Fig. 4: Guinea pig metagenomes resemble Kp-GP, and the guinea pig isolates carry a deletion of *HOG1* and *AHP1*. a,** Detection of *K. pintolopesii* in public metagenome runs from farmed guinea pigs in 13 provinces of China (PRJDB38973). Blue bars represent runs in which *K. pintolopesii* was detected, and gray bars represent the remaining runs. **b,** Share of informative positions that carry the Kp-GP allele in the

metagenomic runs. Informative positions are those at which Kp-GP and Kp-M carry different alleles. Bars show 95% bootstrap intervals. **c,** Read depth of the guinea pig isolates relative to the mouse isolates in 5-kb windows along the Kp-GP (top) and Kp-M (bottom) assemblies, plotted as log2(guinea pig/mouse). The red circle marks the deletion supported by junction reads, red triangles mark unresolved depth drops, and blue diamonds mark gains. **d,** Read depth inside the 13.8-kb HOG1-AHP1 interval relative to the genome-wide median of each isolate (left) and the numbers of Nanopore reads and Illumina read pairs that span the junction (right). Orange dots show mouse isolates and blue diamonds show guinea pig isolates. Gray bars show the 5th to 95th percentile of 9-kb windows across the same genome. n/a, no Nanopore reads.

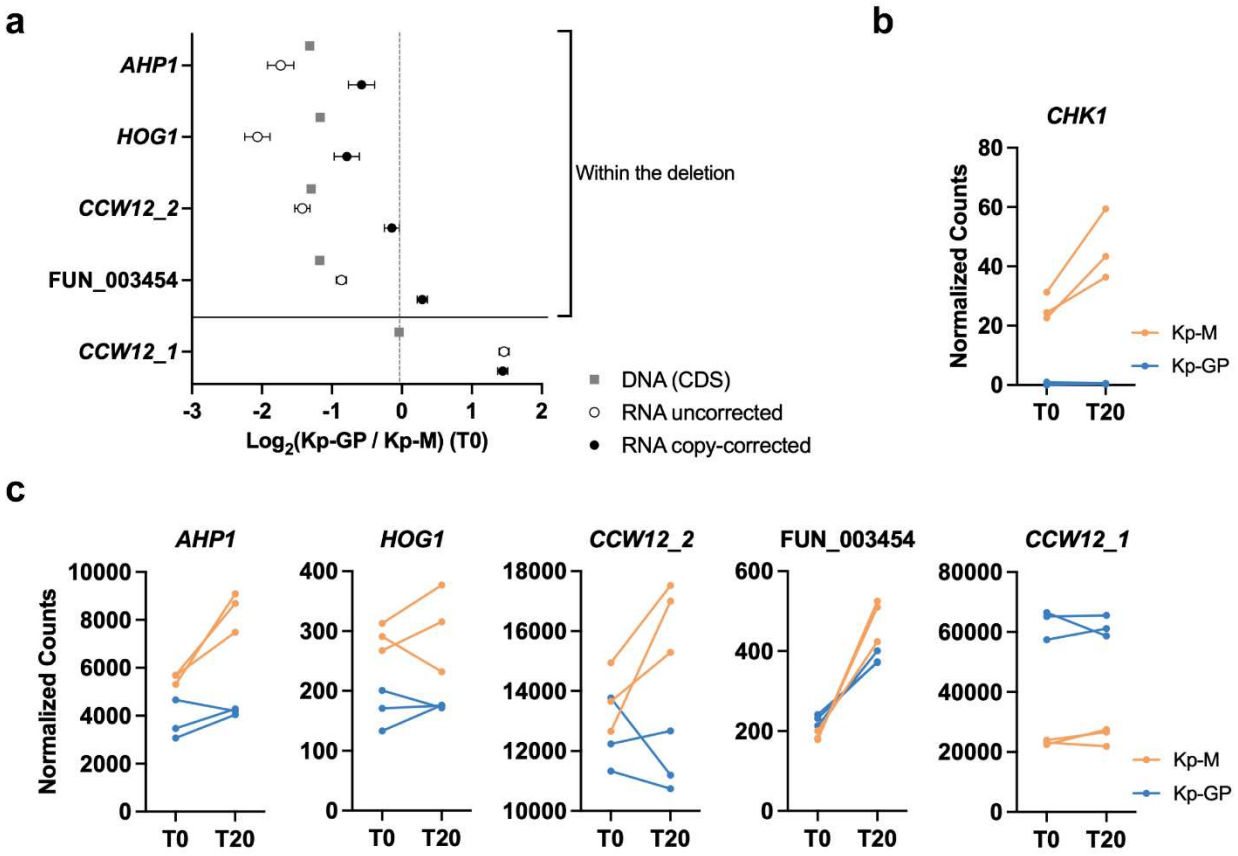



**Fig. 5: Transcript levels reflect the *HOG1-AHP1* and *CHK1* deletions.** RNA-seq results of Kp-M and Kp-GP before (T0) and 20 min after (T20) exposure to aerobic *L. johnsonii* CFS (*n* = 3 for each isolate per time point). **a,** DNA read depth and expression of the four protein-coding genes in the *HOG1-AHP1* interval at T0, with *CCW12_1* outside the interval for comparison. Data are shown as the log2 ratio of Kp-GP to Kp-M. RNA is shown without correction (open circle) and corrected for gene copy number (black circle), and DNA read depth is shown as a gray square. Bars indicate mean ± SEM. **b,** Expression of *CHK1* at T0 and T20. **c,** Expression of the genes in **(a)** at T0 and T20. Differential expression was tested with DESeq2, and *Padj* < 0.05 was considered significant.

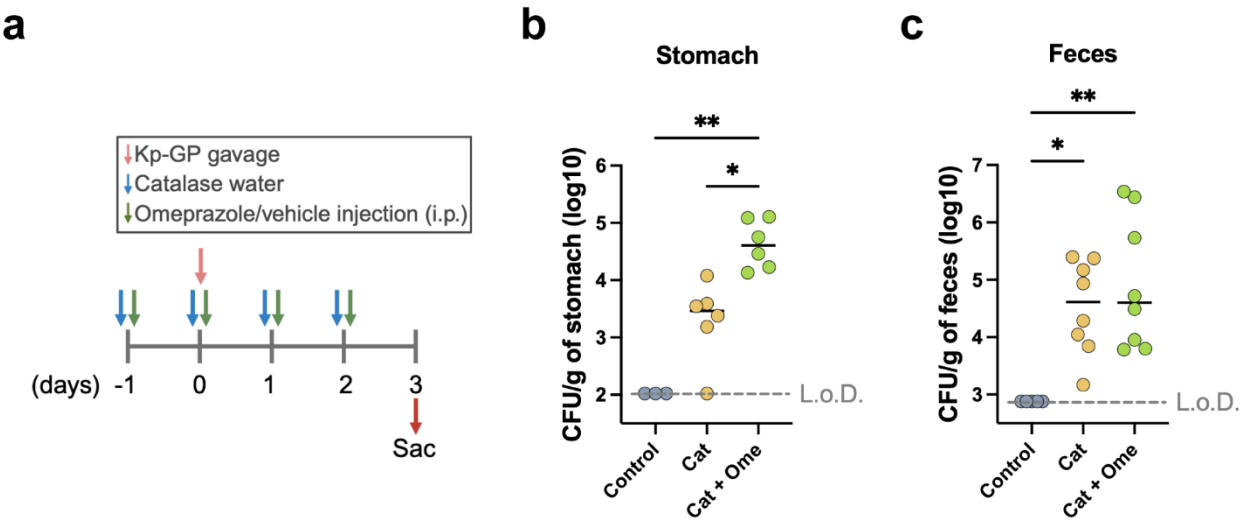
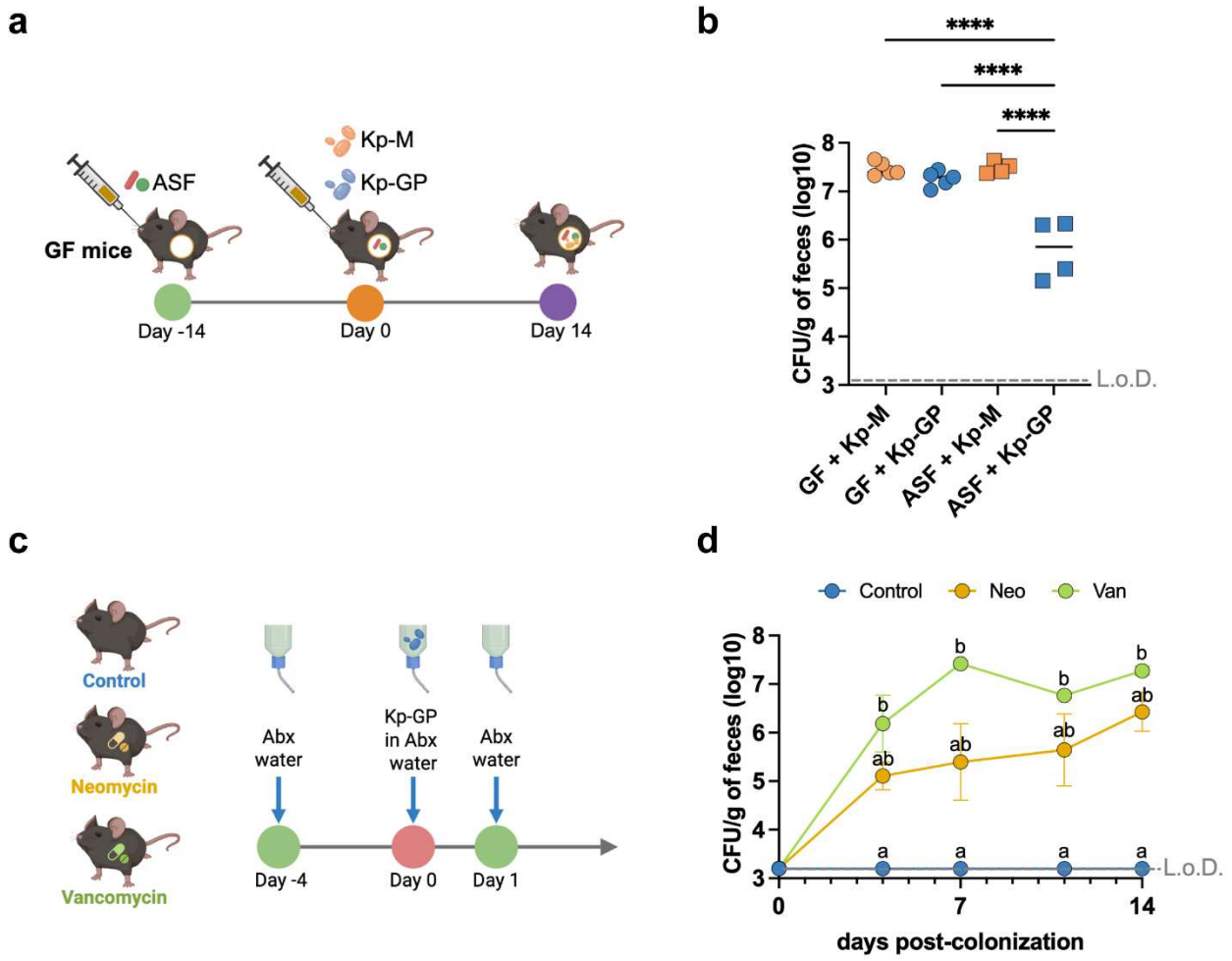


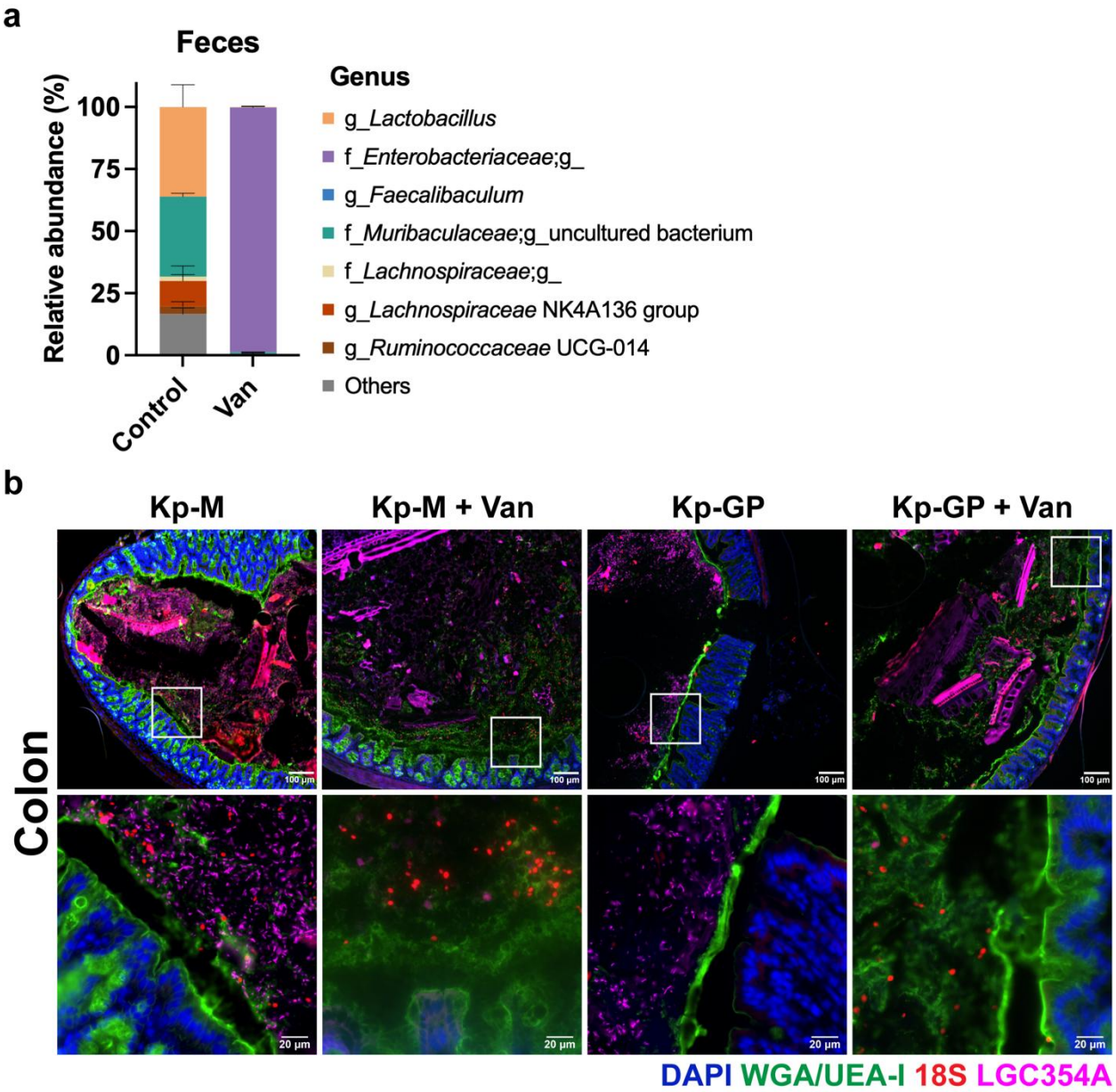

**Fig. 6: Catalase and omeprazole allow partial colonization of the mouse GI tract by Kp-GP. a,** Scheme of the catalase and omeprazole treatment. Kp-GP was administered by oral gavage on day 0, and mice were sacrificed on day 3. **b, c,** Fungal burden in the stomach **(b)** and feces **(c)** of control mice (*n* = 3 for stomach and *n* = 5 for feces) and mice given catalase (Cat; *n* = 6 for stomach and *n* = 8 for feces) or catalase and omeprazole (Cat + Ome; *n* = 6 for stomach and *n* = 8 for feces). L.o.D., limit of detection. Significance was determined by the Kruskal-Wallis test with Dunn's multiple comparisons test. **P* < 0.05 and ***P* < 0.01.

## Extended Data

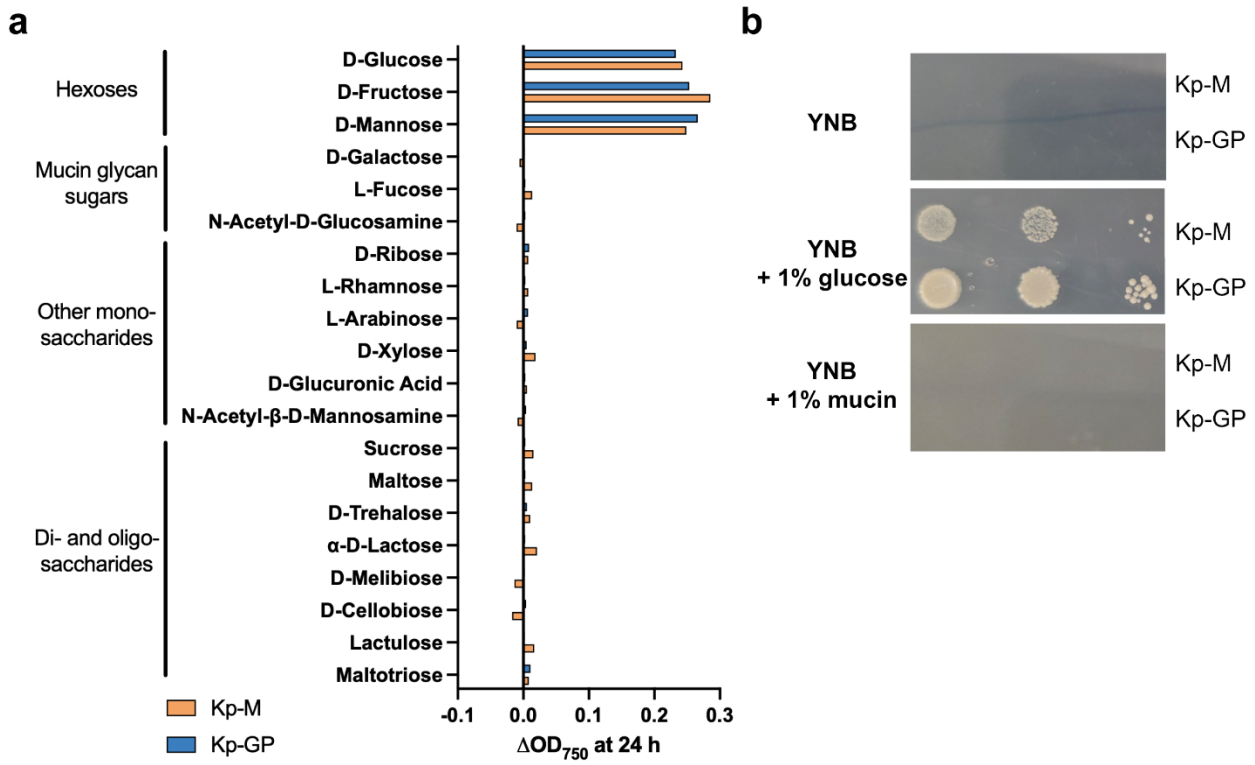



**Extended Data Fig.1: Colonization of Kp-GP in GF mice, ASF mice, and antibiotic-treated mice. a,** Scheme of Kp-M and Kp-GP colonization in germ-free (GF) mice (*n* = 5 per group) and GF mice colonized with Altered Schaedler Flora (ASF; *n* = 4 per group). **b,** Fungal burden in the fecal samples on day 14 in the GF and ASF mice. Significance between groups was determined by one-way ANOVA with Tukey's multiple comparisons test. ****$P < 0.0001$. **c,** Scheme of the antibiotic treatment. Conventional mice received neomycin or vancomycin in drinking water starting four days before Kp-GP colonization (*n* = 4 per group). **d,** Fungal burden in the fecal samples in control, neomycin- (Neo), and vancomycin-treated mice (Van). Data are presented as mean ± SEM. Groups that do not share a letter differ significantly (Significance was

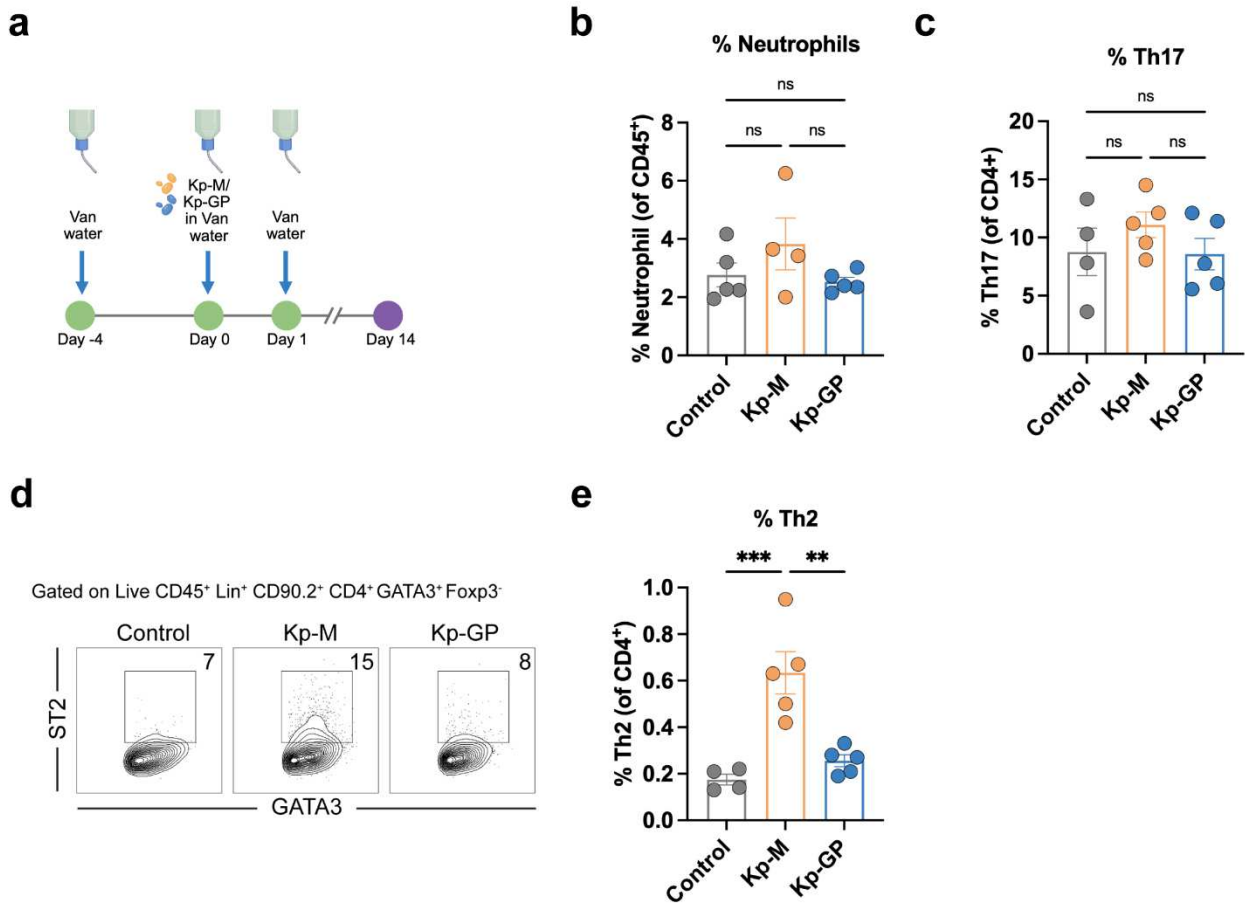
determined at $P < 0.05$ by the Kruskal-Wallis test with Dunn's multiple comparisons test). L.o.D., limit of detection.



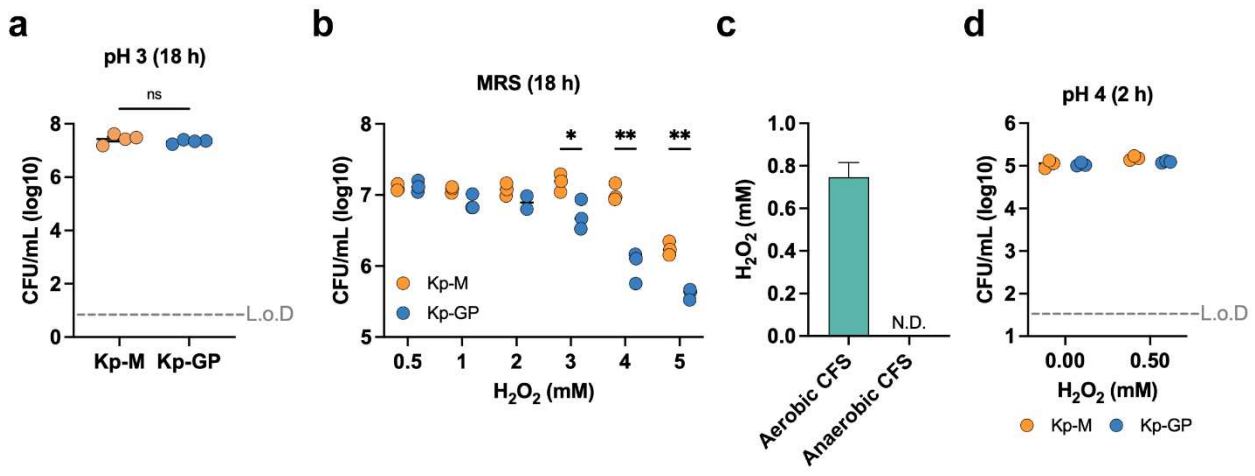

**Extended Data Fig.2: Fecal bacterial community and colonic localization of Kp-M and Kp-GP with and without vancomycin. a,** Relative abundance of bacterial genera in the feces of control and vancomycin-treated mice from 16S rRNA gene sequencing (day 4; *n* = 3 per group). Bars indicate mean ± SEM. **b,** Representative images of the colon of mice colonized with Kp-M or Kp-GP, with or without vancomycin by fluorescence in situ hybridization (FISH). The lower row shows the boxed areas in the upper row. Colors are as in Fig. 2c. Scale bars, 100 µm for the upper row and 20 µm for the lower row.



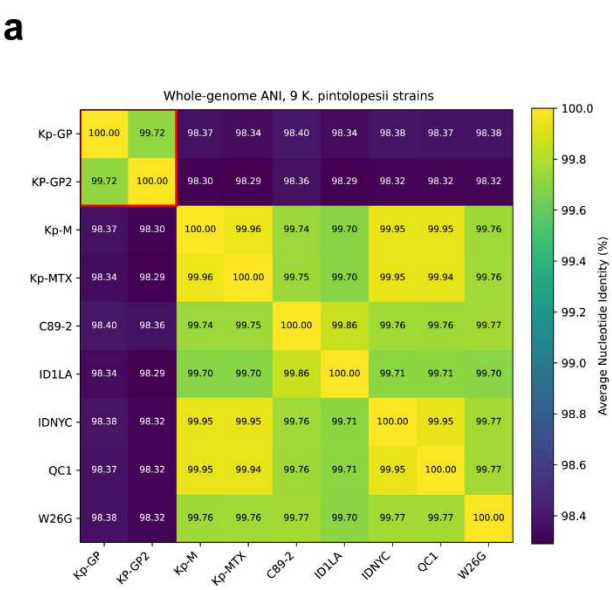

**Extended Data Fig.3: Carbon source utilization assay of Kp-M and Kp-GP. a,** Growth on the carbon sources of Biolog PM1 plates as the change in $OD_{750}$ after 24 h (*n* = 1). **b,** Spot growth of Kp-M and Kp-GP on yeast nitrogen base (YNB) agar alone or with 1% glucose or 1% mucin.

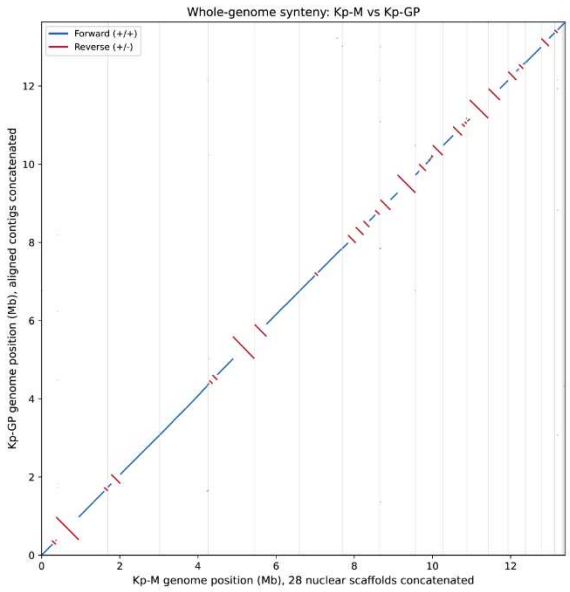


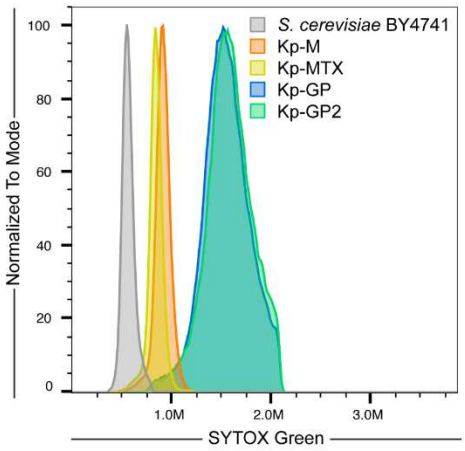

**Extended Data Fig.4: Immune cell analysis of mice colonized with Kp-M or Kp-GP under vancomycin. a,** Scheme of the vancomycin treatment. Conventional mice received vancomycin in drinking water starting four days before fungal colonization. Peripheral blood was analyzed on day 13 and small intestinal lamina propria (SI-LP) on day 14. **b,** Frequency of circulating neutrophils in the peripheral blood at 13 days post-colonization of control (*n* = 5), Kp-M (*n* = 4), and Kp-GP-colonized mice (*n* = 5). Neutrophil populations were gated as $CD45^+$ $B220^-$ $CD11b^+$ $Ly6G^+$. **c,d,e,** Frequencies of Th17 **(c)** and Representative flow cytometry plots **(d)** and frequency of Th2 **(e)** in the SI-LP of control (*n* = 4), Kp-M (*n* = 5), and Kp-GP-colonized mice (*n* = 5). Th17 populations were gated as live $CD45^+$ $Lin^+$ $CD90.2^+$ $CD4^+$ $Foxp3^-$ $ROR\gamma T^+$. Th2 populations were gated as live $CD45^+$ $Lin^+$ $CD90.2^+$ $CD4^+$ $GATA3^+$ $Foxp3^-$ $ST2^+$. Lin: CD3ε, CD19. Bars indicate mean ± SEM.

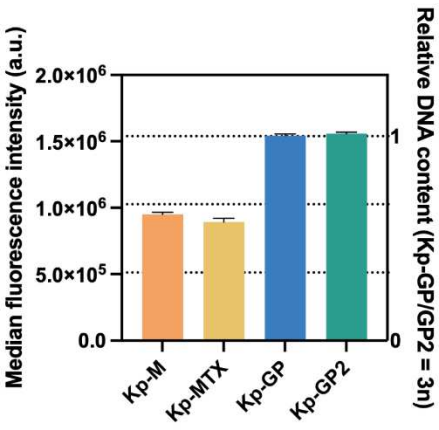



**Extended Data Fig.5: Effects of low pH and $H_2O_2$ on Kp-M and Kp-GP. a,** Growth of Kp-M and Kp-GP after 18 h of culture at pH 3. **b,** Growth of Kp-M and Kp-GP after 18 h of culture in MRS medium with 0.5 to 5 mM $H_2O_2$. **c,** $H_2O_2$ in aerobic and anaerobic *L. johnsonii* CFS. N.D., not detected. Bars indicate mean ± SEM. **d,** Survival of Kp-M and Kp-GP after 2 h in YNB medium at pH 4 with or without 0.5 mM $H_2O_2$. L.o.D., limit of detection. Significance between groups was determined by an unpaired two-tailed t-test for **(a)** and by Mann-Whitney test for **(b)** and **(c)**. ns, not significant. *$P$ < 0.05 and **$P$ < 0.01.



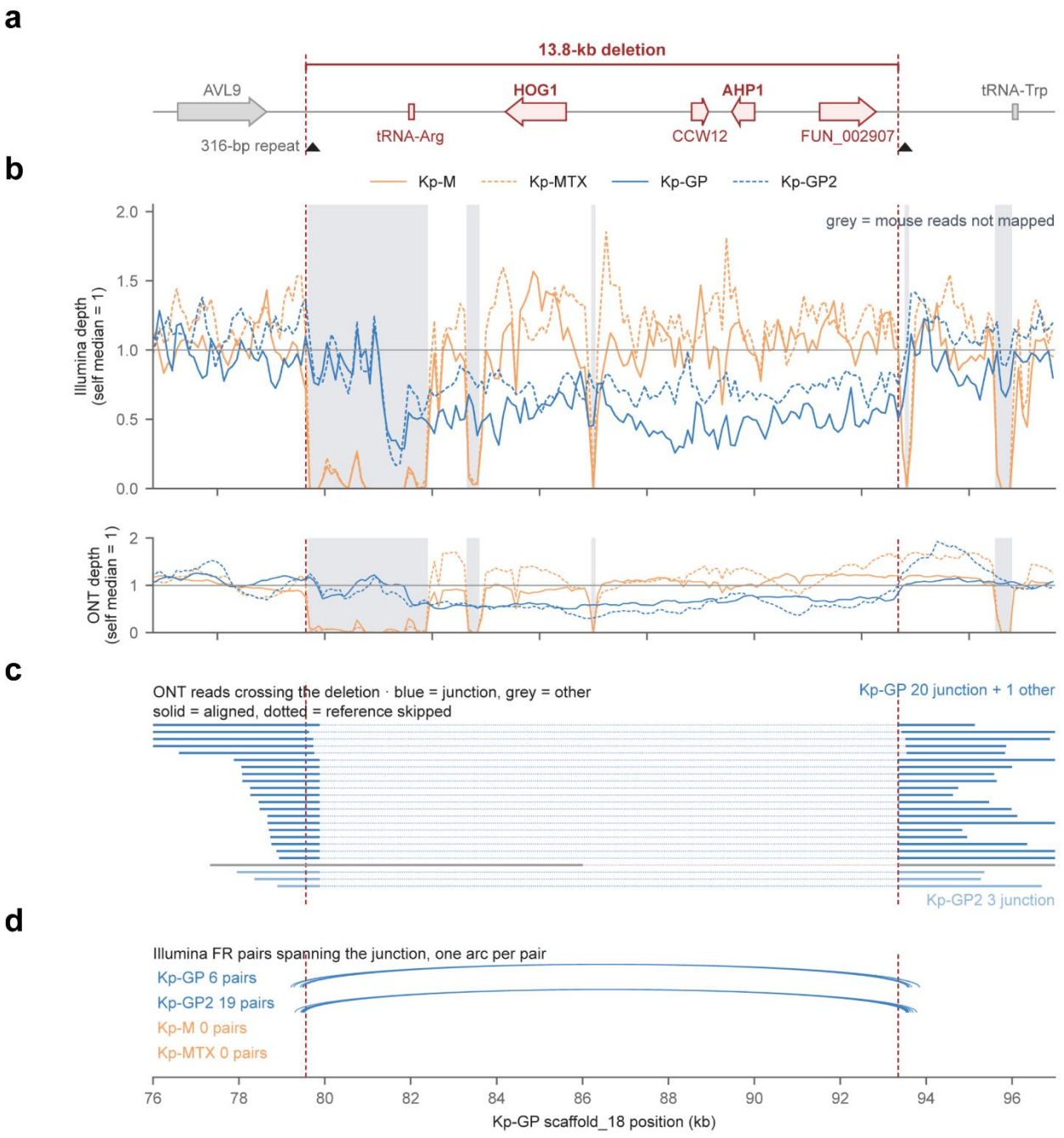




e

| Strain | Host of origin | k-mer spectrum / Smudgeplot | Relative DNA content (flow cytometry) | Assigned ploidy |
|---|---|---|---|---|
| Kp-M | Mouse | Not determinable | 0.61 ± 0.02 | 2n |
| Kp-MTX | Mouse | Not determinable | 0.58 ± 0.03 | 2n |
| ID1LA | Mouse | 2n (AB) | Not analyzed | 2n |
| C89-2 | Mouse | 2n (AB) | Not analyzed | 2n |
| W26G | Mouse | Not determinable | Not analyzed | Undetermined |
| IDNYC | Mouse | Not determinable | Not analyzed | Undetermined |
| QC1 | Mouse | Not determinable | Not analyzed | Undetermined |
| Kp-GP | Guinea pig | 3n (AAB) | 1.00 ± 0.01 | 3n |
| Kp-GP2 | Guinea pig | 3n (AAB) | 1.00 ± 0.01 | 3n |

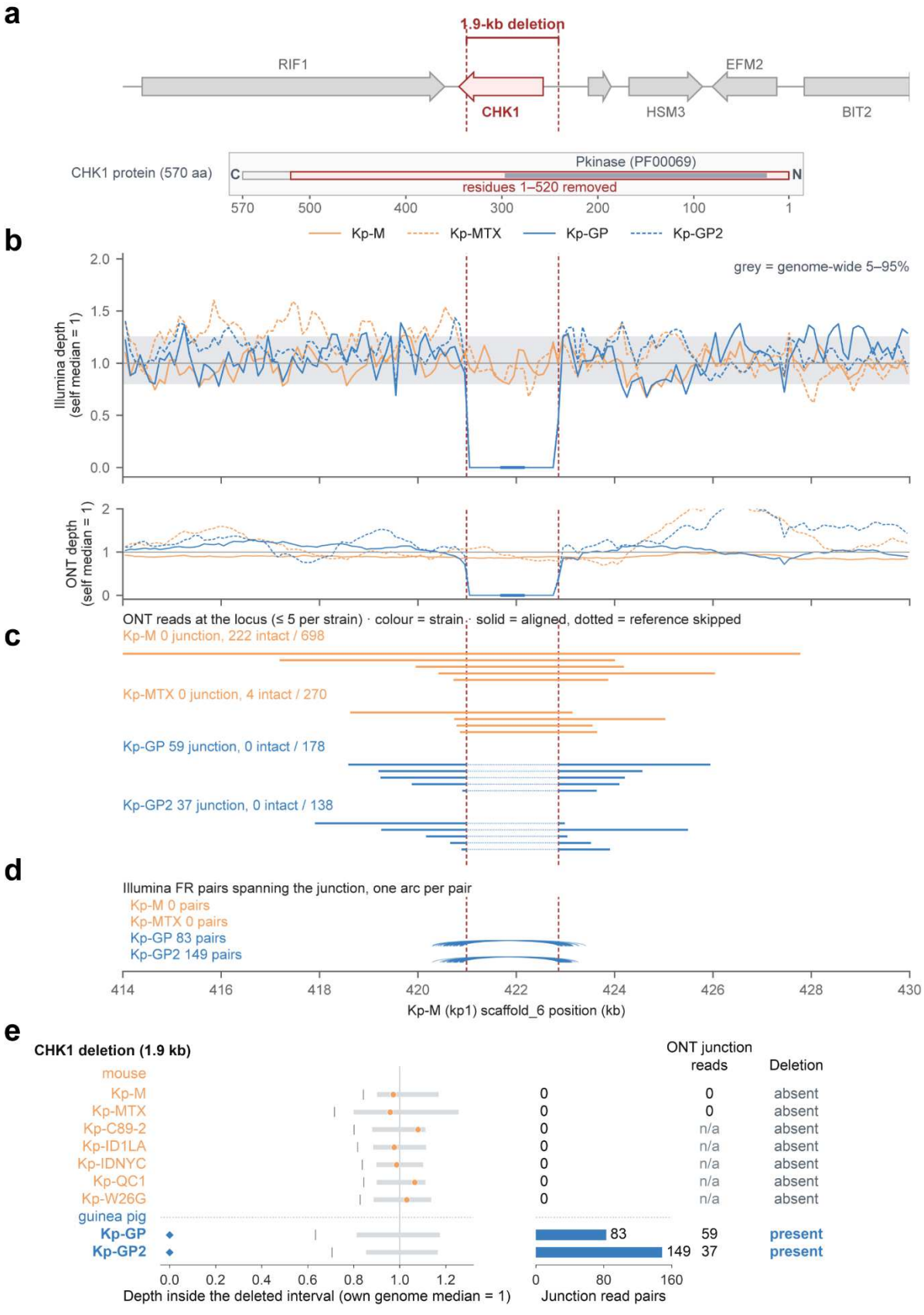
**Extended Data Fig. 6: Genome comparison and ploidy of the *K. pintolopesii* isolates. a,**

Average nucleotide identity (ANI) among the nine isolates. **b,** Synteny between the Kp-M and Kp-GP assemblies. **c,** SYTOX Green fluorescence of Kp-M, Kp-MTX, Kp-GP, Kp-GP2, and *S. cerevisiae* BY4741 as a reference. **d,** Median fluorescence of the G1 peak (*n* = 3 per isolate). The right y-axis shows DNA content relative to Kp-GP and Kp-GP2, set to 1.00 as triploid. Bars indicate mean ± SEM. **e,** Ploidy assigned from the *k*-mer spectrum and flow cytometry. Values are the mean ± SEM.

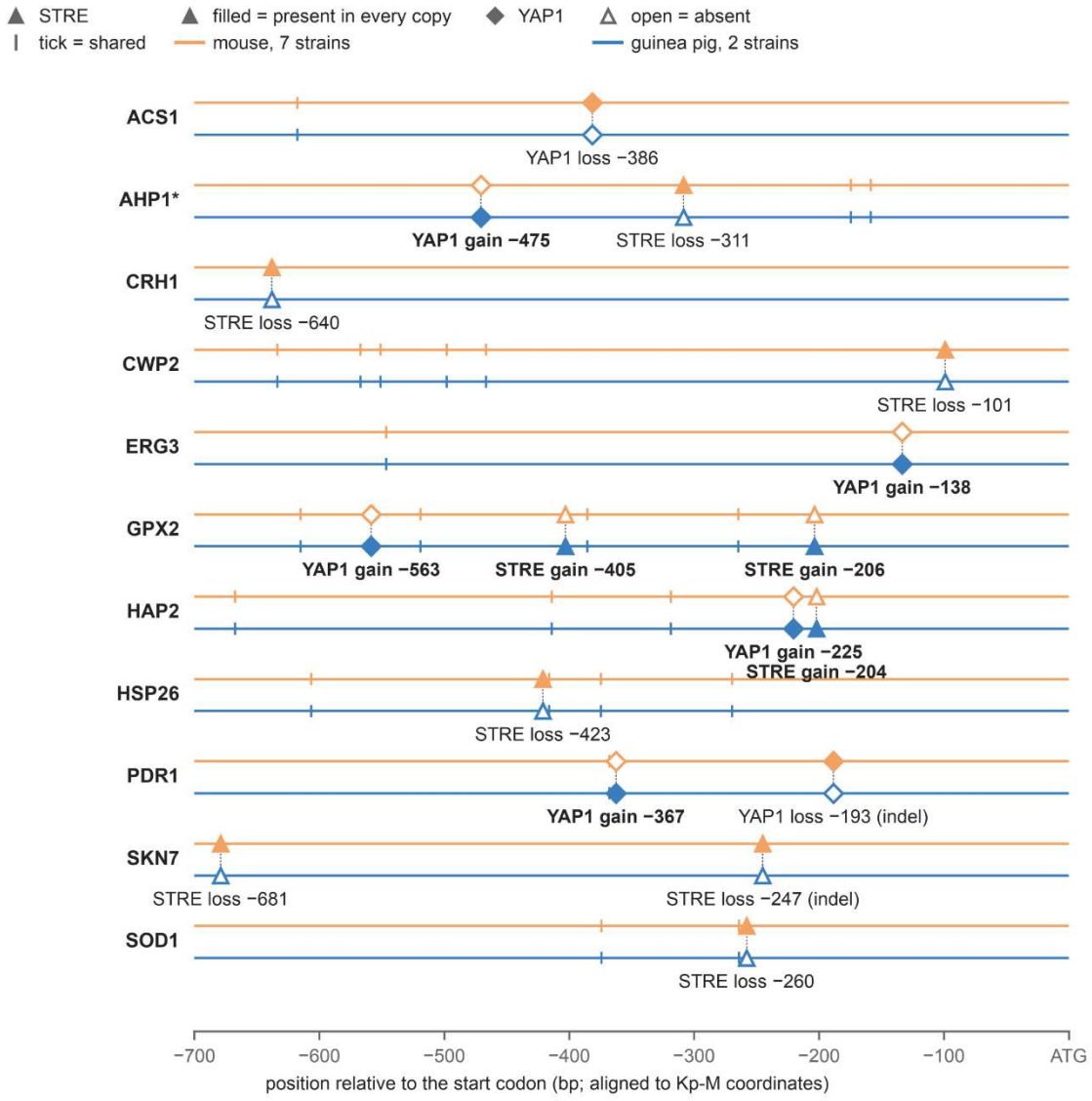



**Extended Data Fig. 7: Sequencing read analysis of the *HOG1-AHP1* interval. a,** Gene map on Kp-GP scaffold 18. The deletion lies between two copies of a 316-bp repeat (black triangles) and removes tRNA-Arg, *HOG1*, *CCW12_2*, *AHP1*, and FUN_002907. **b,** Illumina (top) and Nanopore (bottom) read depth relative to the median of each isolate. Gray areas mark positions

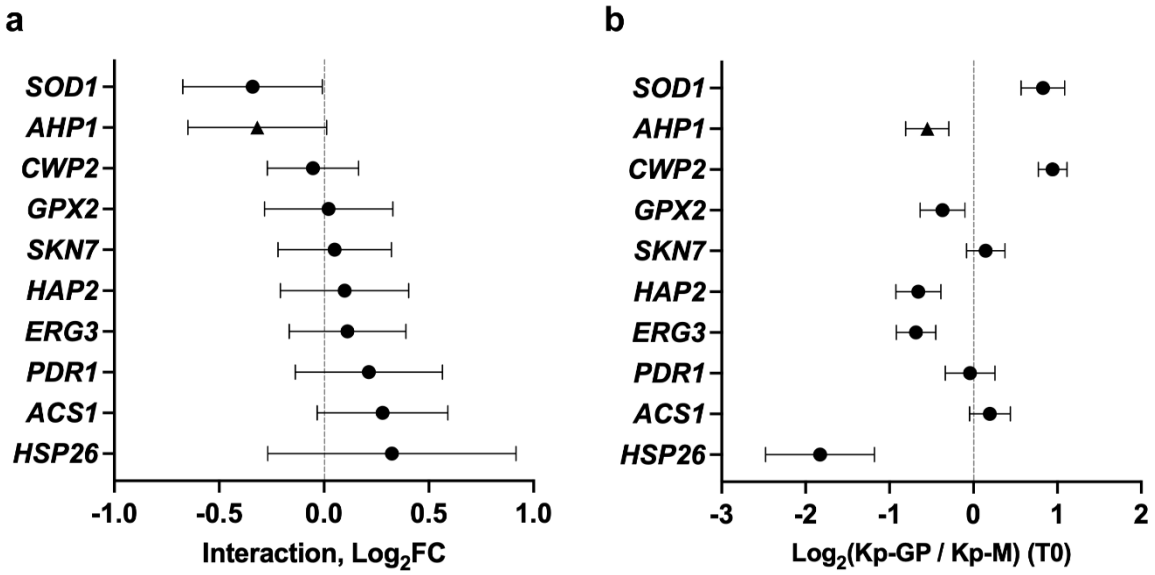
at which mouse reads could not be mapped. **c,** Nanopore reads that cross the deletion. Blue lines span the junction and gray lines do not. **d,** Illumina read pairs that span the junction.


a
1.9-kb deletion
RIF1
CHK1
HSM3
EFM2
BIT2
CHK1 protein (570 aa)
C
N
Pkinase (PF00069)
residues 1–520 removed
570
500
400
300
200
100
1
b
Kp-M
Kp-MTX
Kp-GP
Kp-GP2
grey = genome-wide 5–95%
Illumina depth (self median = 1)
ONT depth (self median = 1)
c
ONT reads at the locus (≤ 5 per strain) · colour = strain · solid = aligned, dotted = reference skipped
Kp-M 0 junction, 222 intact / 698
Kp-MTX 0 junction, 4 intact / 270
Kp-GP 59 junction, 0 intact / 178
Kp-GP2 37 junction, 0 intact / 138
d
Illumina FR pairs spanning the junction, one arc per pair
Kp-M 0 pairs
Kp-MTX 0 pairs
Kp-GP 83 pairs
Kp-GP2 149 pairs
414
416
418
420
422
424
426
428
430
Kp-M (kp1) scaffold_6 position (kb)
e
CHK1 deletion (1.9 kb)
ONT junction reads
Deletion
mouse
Kp-M
Kp-MTX
Kp-C89-2
Kp-ID1LA
Kp-IDNYC
Kp-QC1
Kp-W26G
guinea pig
Kp-GP
Kp-GP2
0
0
absent
0
0
absent
0
n/a
absent
0
n/a
absent
0
n/a
absent
0
n/a
absent
0
n/a
absent
83
59
present
149
37
present
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Depth inside the deleted interval (own genome median = 1)
0
80
160
Junction read pairs

**Extended Data Fig. 8: Sequencing read analysis of the *CHK1* locus. a,** Gene map on Kp-M scaffold 6 and the CHK1 protein with its kinase domain (Pfam PF00069). The deletion removes the start codon and residues 1 to 520 of 570. **b,** Illumina (top) and Nanopore (bottom) read depth relative to the median of each isolate. The gray band shows the 5th to 95th percentile of genome-wide depth. **c,** Nanopore reads at the locus, at most five per isolate. Numbers show junction reads and intact reads. **d,** Illumina read pairs that span the junction. **e,** Read depth inside the deleted interval and junction reads for the nine isolates, as in Fig. 4d.




**Extended Data Fig. 9: STRE and Yap1-binding sites in the promoters of stress-response genes.** Stress response element (STRE, triangles) and Yap1-binding sites (diamonds) in the 700 bp upstream of the start codon, aligned to Kp-M coordinates. Orange lines indicate the mouse isolates and blue lines indicate the guinea pig isolates. Filled symbols mark sites present in every copy, open symbols mark absent sites, and ticks mark shared sites. Bold labels mark gains and regular labels mark losses in the guinea pig isolates. *One of the three copies of *AHP1* is deleted in guinea pig isolates.




1185

1186 **Extended Data Fig. 10: RNA-seq analysis of genes with STRE or Yap1-binding site**

1187 **differences. a,** Isolate-by-treatment interaction on expression after 20 min exposure to *L.*

1188 *johnsonii* CFS for ten of the eleven genes in Extended Data Fig. 9, shown as log2 fold change.

1189 *CRH1* was left out because its gene model appears to join two genes. No interaction was

1190 significant at *Padj* < 0.05. **b,** Log2 ratio of Kp-GP to Kp-M at T0 for the same genes. Bars

1191 indicate mean ± SEM. The triangle: One of the three copies of *AHP1* is deleted in guinea pig

1192 isolates.

1193

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarytables.xlsx