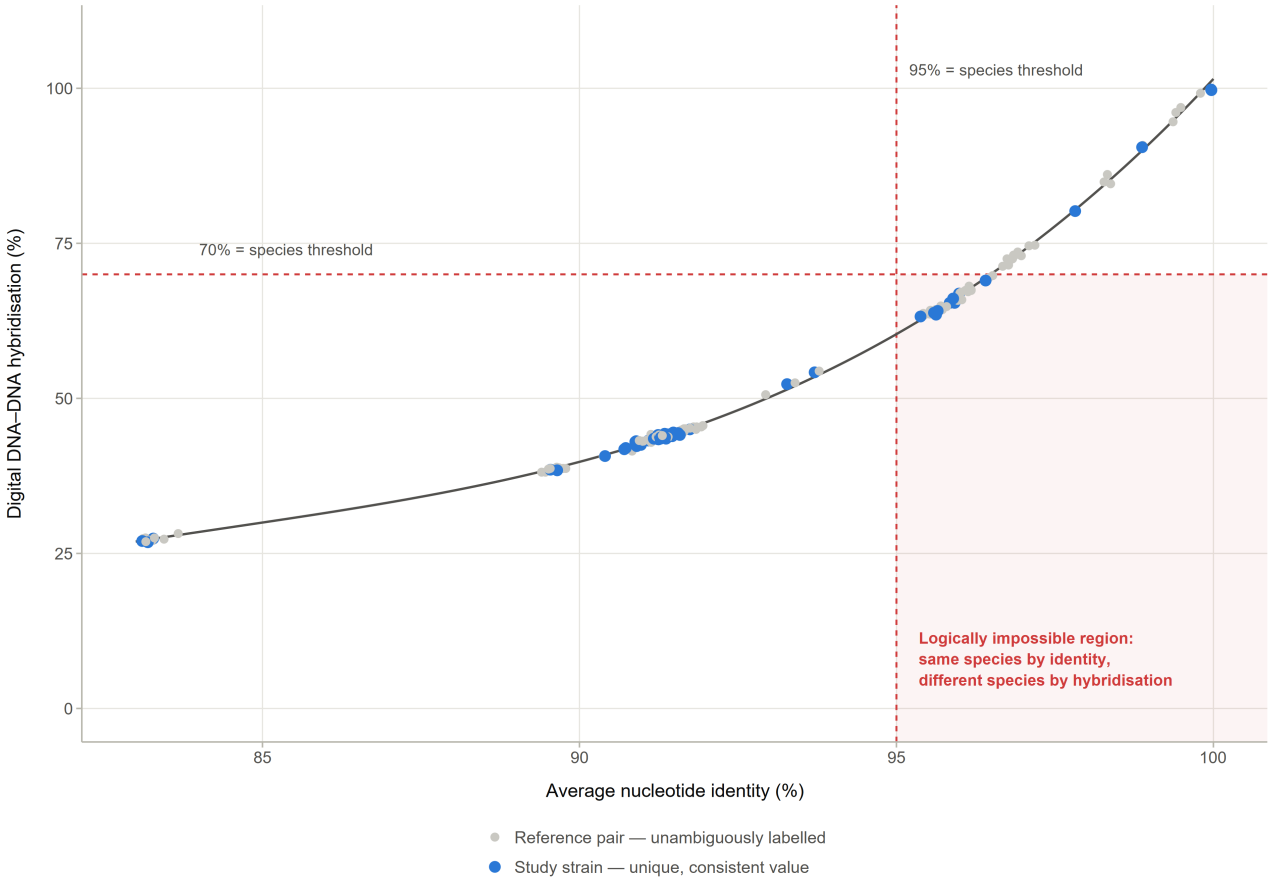


Fig. S1. Average nucleotide identity versus digital DNA–DNA hybridisation, all 190 genome pairs. Agreement between two independent species-delineation criteria across all 190 pairs: ANI against digital DNA–DNA hybridisation (GGDC formula 2). All 190 pairs resolved. The grey line is a degree-two polynomial fit of hybridisation on identity, over which the residual standard deviation is 1.112 percentage points; none of the 54 pairs involving a study strain falls outside three standard deviations of that fit. Dashed lines mark the equivalent species thresholds of the two criteria and the shaded area the region logically inconsistent between them. The two criteria are therefore mutually corroborating rather than merely co-reported.

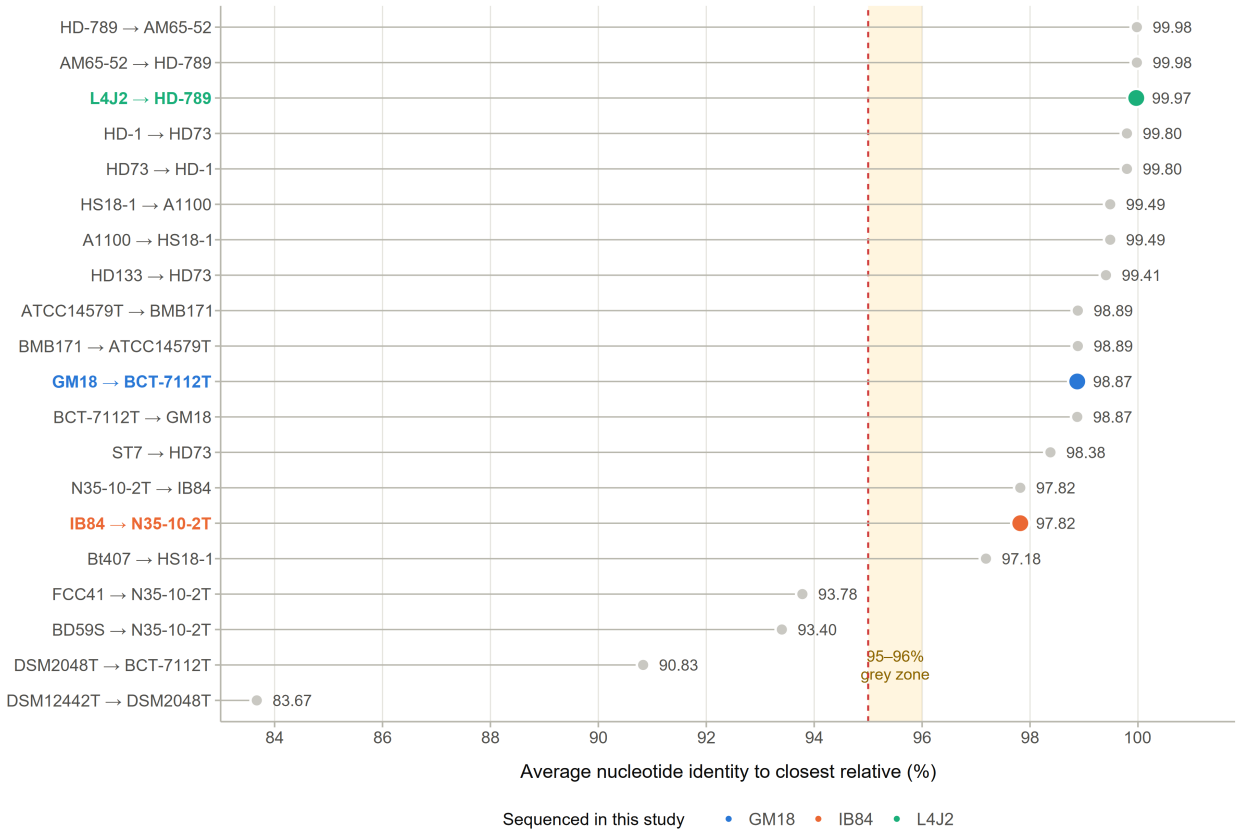


Fig. S2. Closest relative of each genome by average nucleotide identity. Closest relative of each genome by average nucleotide identity, with the amber band marking the 95–96% range in which the criterion is least reliable and the dashed red line the 95% threshold. GM18 and IB84 exceed the threshold and are separated from their second-closest neighbour by 7.29 and 4.11 percentage points respectively. L4J2 is the exception: its two closest relatives, HD-789 and AM65-52, are themselves near-identical to it, so its margin to the second-closest genome is 0.01 points and no single nearest relative can be identified among the three.

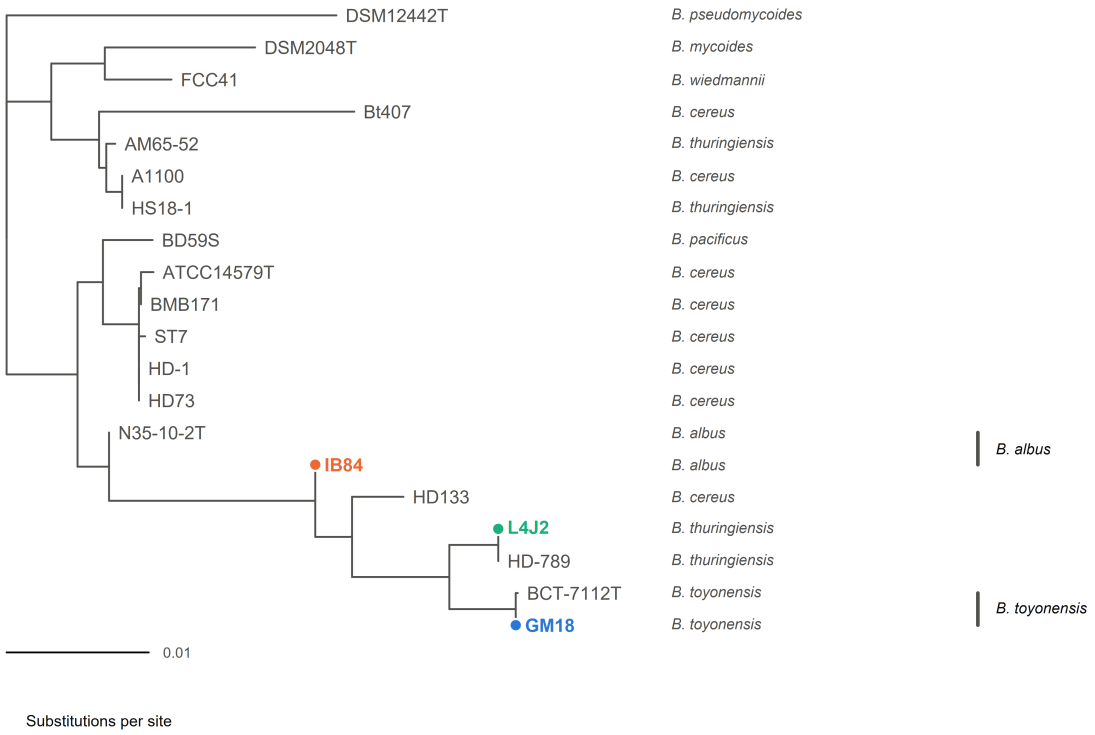


Fig. S3. Maximum-likelihood phylogeny from twenty vertically inherited marker genes. Maximum-likelihood phylogeny from twenty vertically inherited marker genes (VBCG v1.3, RAxML, 100 bootstrap replicates), rooted on the same outgroup to permit comparison. Brackets mark the ANI species blocks, which cease to be clades in this tree: GM18 separates from the B. toyonensis type strain and AM65-52 leaves the L4J2 / HD-789 group. Three genome pairs were declared identical under this marker set, indicating loss of marker resolution rather than an alternative evolutionary history. This figure must not be presented as confirming Fig. 2. Branch lengths are not comparable in magnitude with Fig. 2; only topology is. Scale bar, 0.01 substitutions per site.

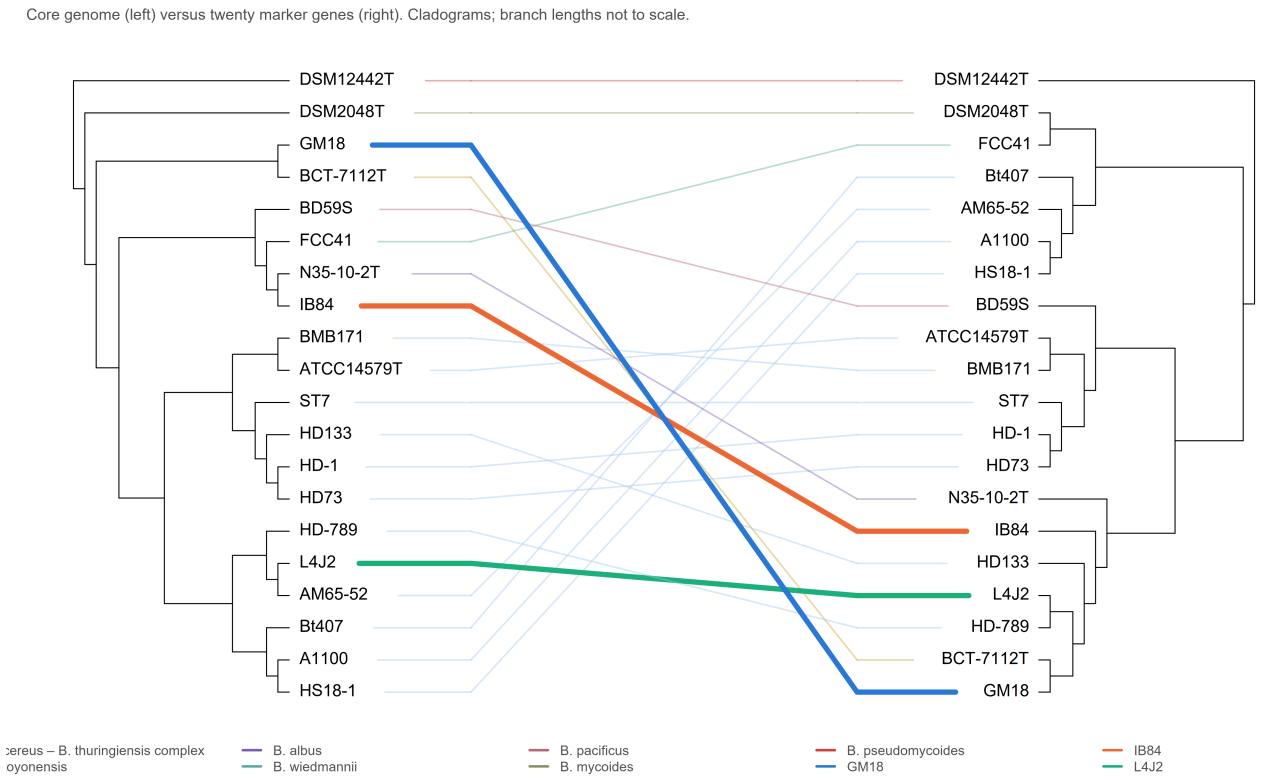


Fig. S4. Topological comparison of the core-genome and marker-gene phylogenies. Face-to-face comparison of the two phylogenies, drawn as cladograms because only topology is comparable between inferences on different character sets. Lines join the same genome in the two trees and their colour indicates the 95% ANI species block. The abundance of crossings reflects extensive topological discordance (Robinson–Foulds distance 26 of a possible 34; 4 of 17 internal nodes shared). The vertical ordering of taxa is chosen by the drawing algorithm to minimise crossings and has no biological meaning. The Robinson–Foulds distance counts unshared bipartitions without weighting by support, treating a 100% bootstrap node and a 23% node alike.

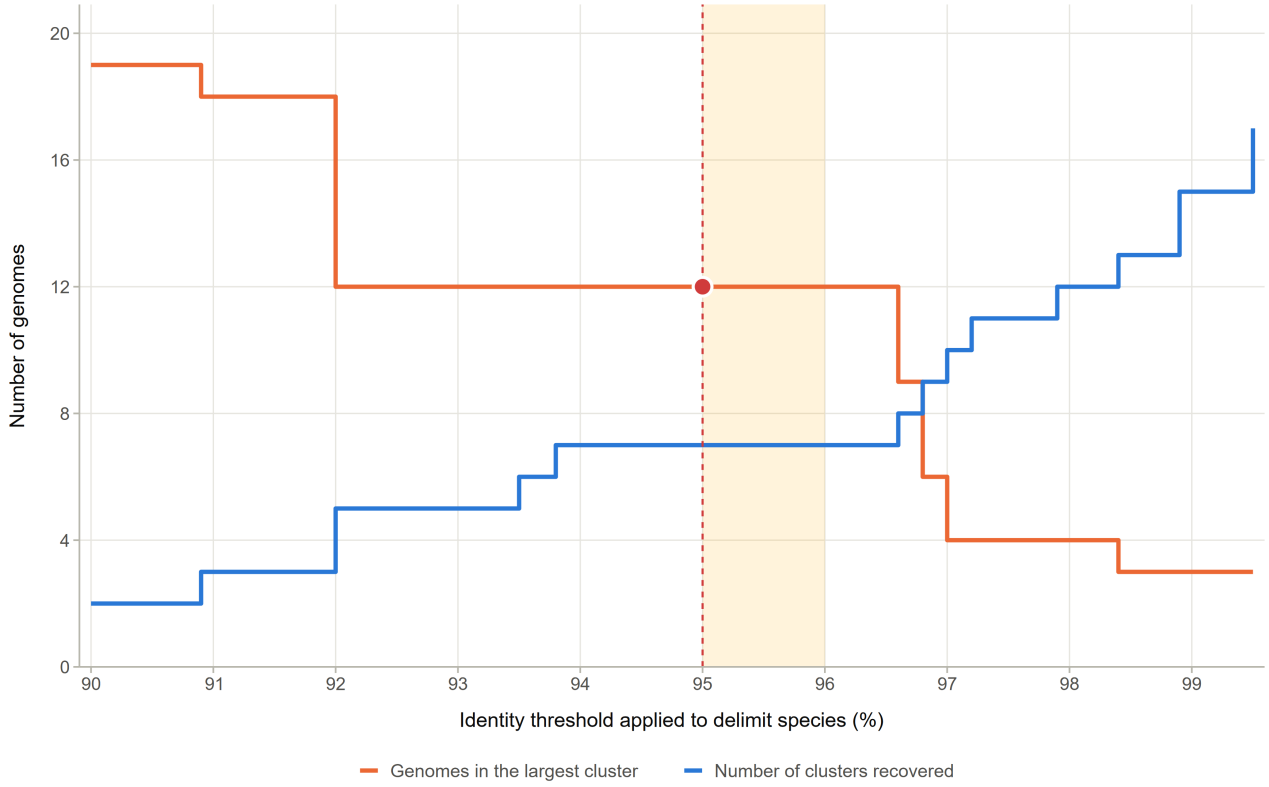


Fig. S5. Species clusters recovered as a function of the identity threshold applied. Single-linkage clustering of the twenty genomes at every ANI threshold between 90.0% and 99.5%. The orange series gives the number of genomes in the largest cluster and the blue series the total number of clusters recovered. The amber band marks the 95–96% range and the dashed line the conventional 95% threshold, at which a single cluster contains twelve of the twenty genomes, including the type strain of B. cereus, four B. thuringiensis reference strains and strain L4J2. GM18 and IB84 persist as two-genome clusters with their respective type strains throughout the range examined. Single linkage chains genomes through intermediates and is deliberately permissive: it is the appropriate choice for testing whether a threshold separates a set, and is not a classification proposal. The two steps in the blue series near 99% reflect the separation of near-clonal pairs and carry no taxonomic meaning.

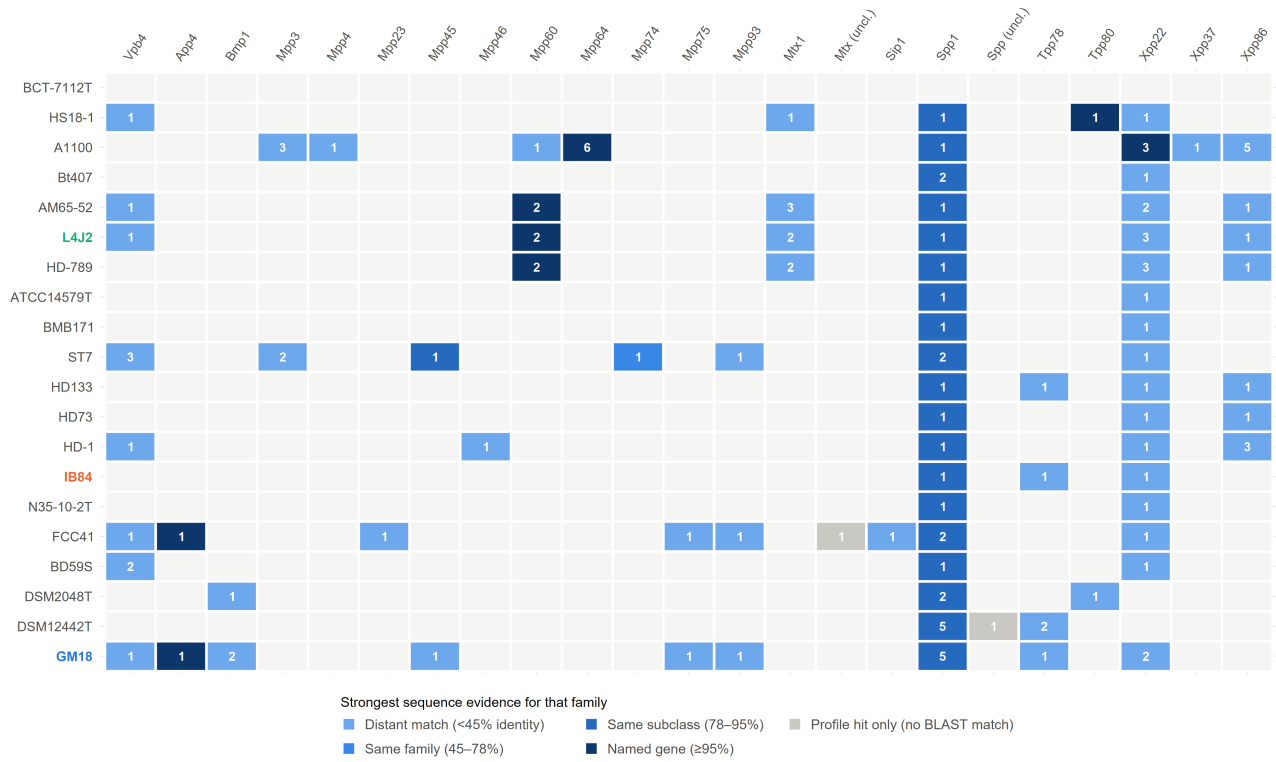


Fig. S6. Parasporin-family gene families carried by each genome. Parasporin-like gene families carried by each genome under the structure-based App/Bmp/Mpp/Mtx/Sip/Spp/Tpp/Xpp nomenclature. Unlike the classical class, this one is nearly universal in the panel, with family Spp1 present in nineteen of the twenty genomes, so the distinguishing feature is the breadth of families and the confidence level reached. Sixty-three per cent of locus entries are distant matches below 45% identity and only eight per cent reach named-gene identity; any figure quoted for parasporin load must state the threshold applied. Selective cytotoxicity cannot be inferred from sequence similarity.

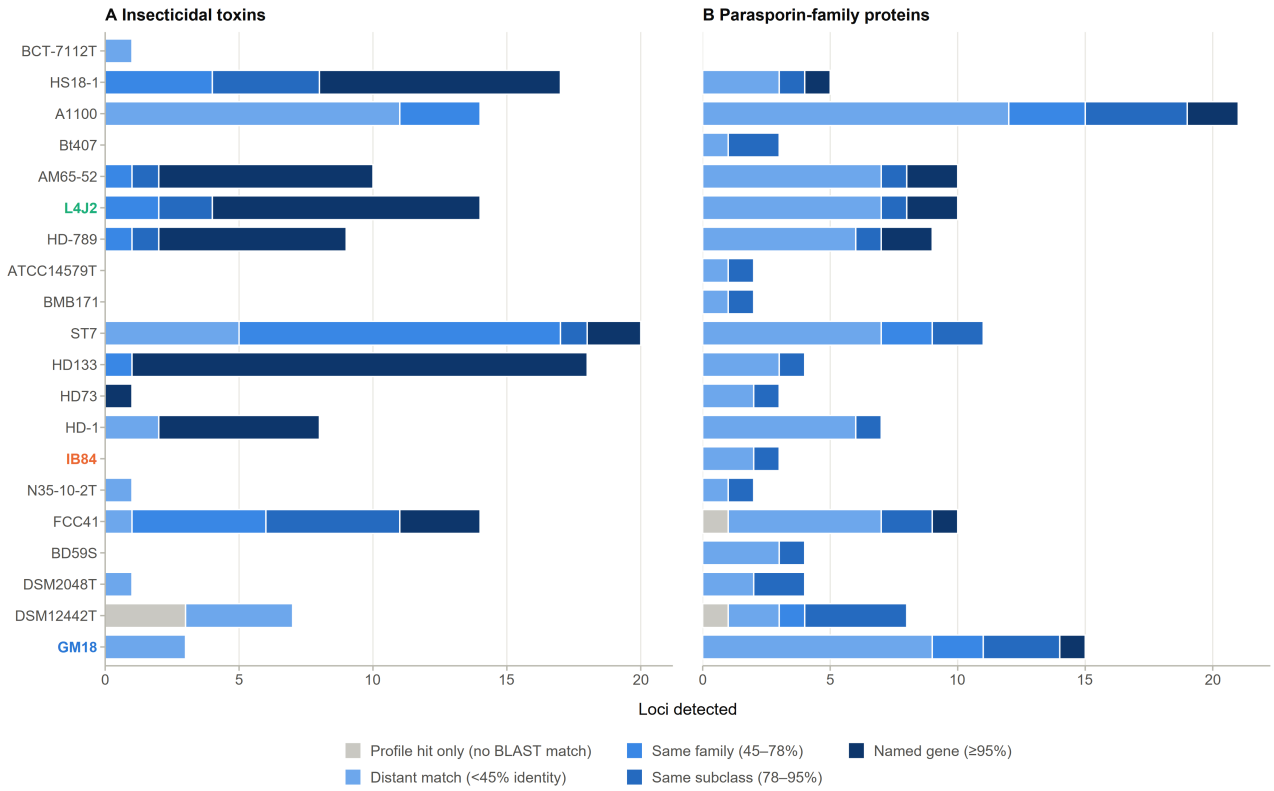


Fig. S7. Sequence-evidence strength behind each toxin locus assignment. Strength of the sequence evidence behind each detected locus, for the classical insecticidal class (a) and the structure-based class (b), broken down by the identity of the best BPPRC match into four tiers plus a category for loci detected only by hidden Markov model profile. The two classes differ markedly: 46% of classical loci reach named-gene identity against only 8% of structure-based loci, of which 63% are distant matches. The tiers are identity thresholds against a reference database, not measures of biological activity; a named locus may be pseudogenised. Horizontal axis scales differ between panels.

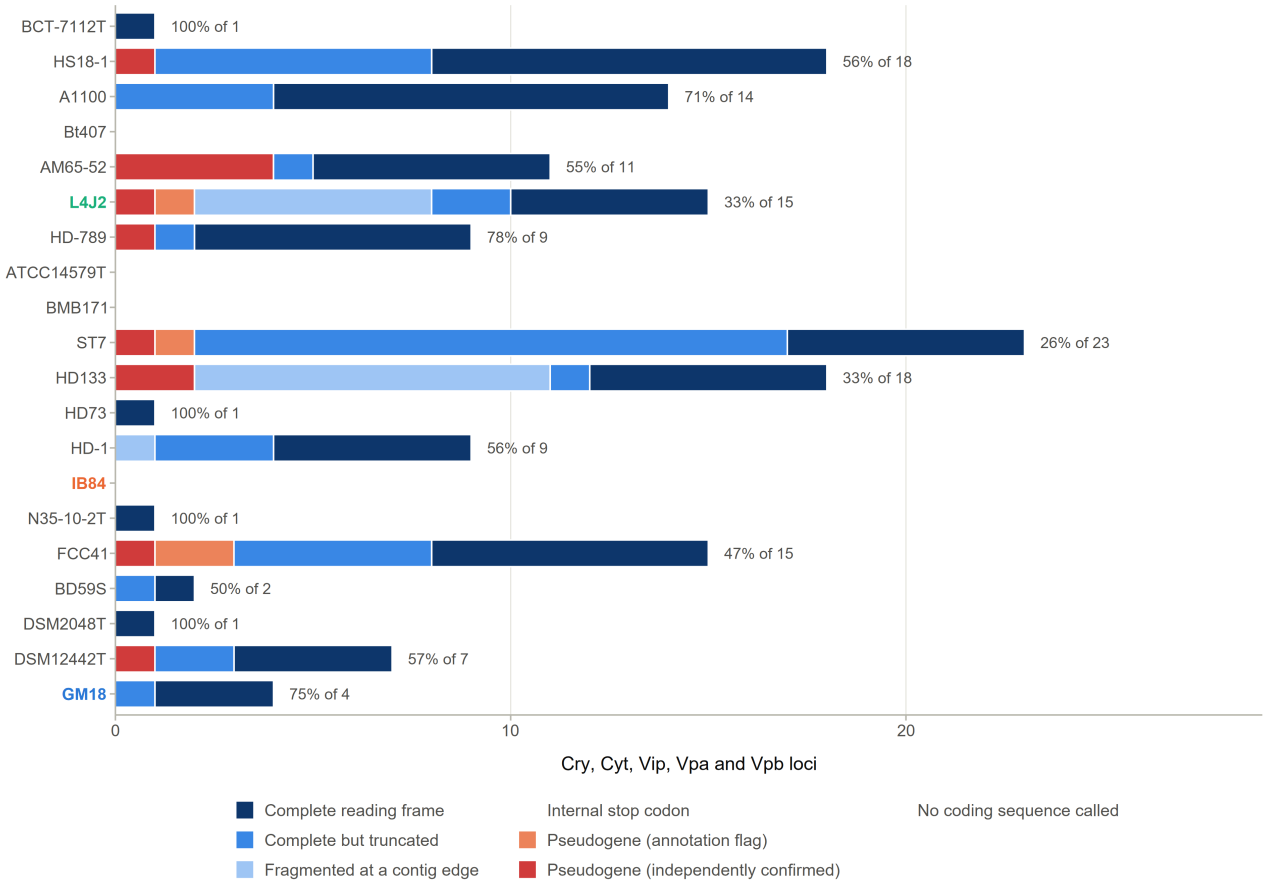


Fig. S8a. Coding integrity of insecticidal toxin and parasporin-family loci in each genome. (a) Coding integrity of insecticidal toxin loci in each genome. (b) Coding integrity of parasporin-family loci in each genome. The 271 detected locus entries were classified by direct translation, checking length, internal stop codons and length ratio against the reference protein, and cross-referencing the result with two independent pseudogene call sets. Each bar corresponds to a genome and its total length to the number of loci detected; the text on the right gives the percentage of loci with an intact frame and the total on which it is computed. Rows follow the order of the core-genome phylogeny. All three loci of IB84 are intact; GM18 retains 12 of 18, with losses attributable to pseudogenisation that improved contiguity did not reverse; L4J2 retains 11 of 24 and includes independently confirmed pseudogenes, in the most fragmented assembly of the panel and the genome with by far the largest integrated prophage burden of the three study strains. Percentages for genomes with very few loci are fragile, which is why the denominator is always printed; the contig-edge category is an assembly artefact rather than a biological observation. The horizontal axis scales differ between panels.

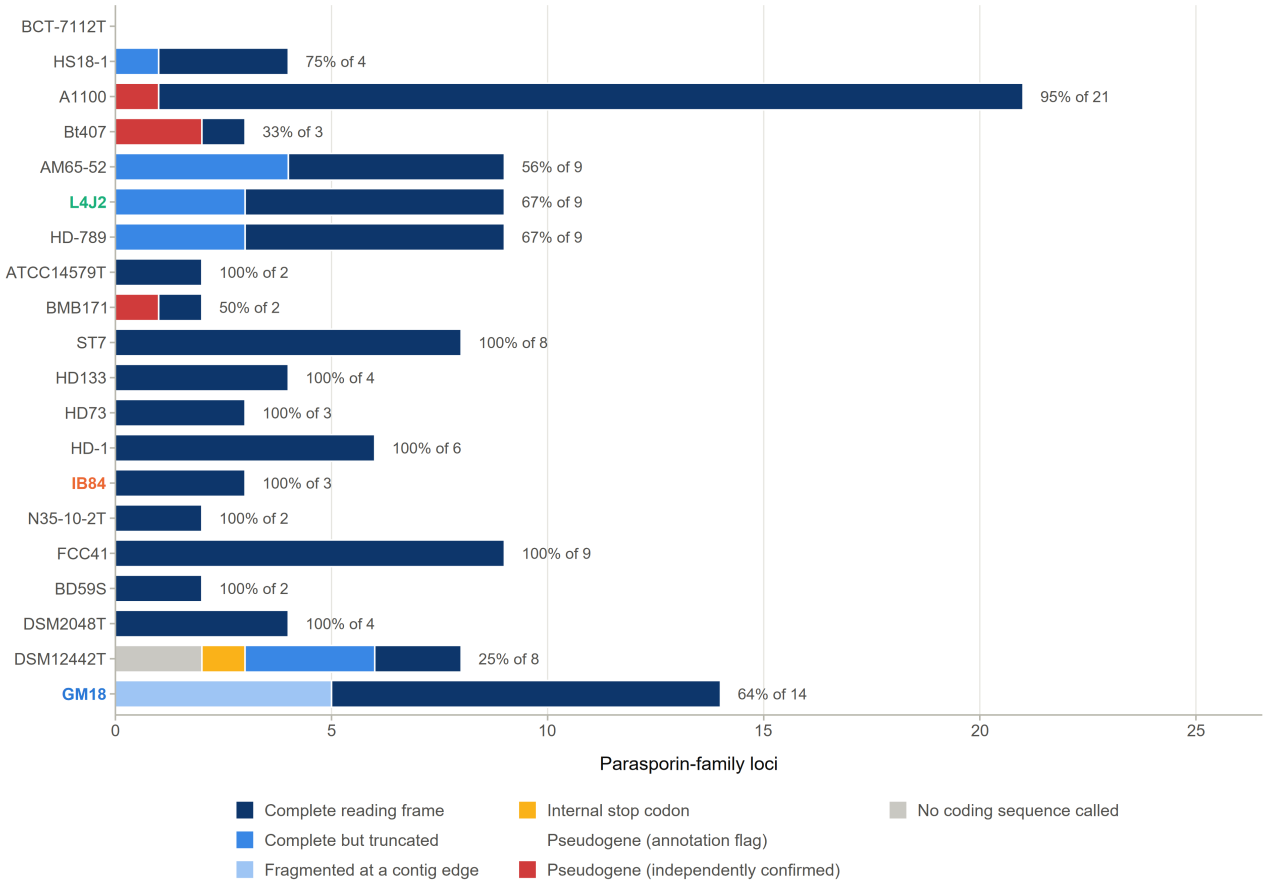


Fig. S8b. (continued) Coding integrity of insecticidal toxin and parasporin-family loci in each genome. (a) Coding integrity of insecticidal toxin loci in each genome. (b) Coding integrity of parasporin-family loci in each genome. The 271 detected locus entries were classified by direct translation, checking length, internal stop codons and length ratio against the reference protein, and cross-referencing the result with two independent pseudogene call sets. Each bar corresponds to a genome and its total length to the number of loci detected; the text on the right gives the percentage of loci with an intact frame and the total on which it is computed. Rows follow the order of the core-genome phylogeny. All three loci of IB84 are intact; GM18 retains 12 of 18, with losses attributable to pseudogenisation that improved contiguity did not reverse; L4J2 retains 11 of 24 and includes independently confirmed pseudogenes, in the most fragmented assembly of the panel and the genome with by far the largest integrated prophage burden of the three study strains. Percentages for genomes with very few loci are fragile, which is why the denominator is always printed; the contig-edge category is an assembly artefact rather than a biological observation. The horizontal axis scales differ between panels.

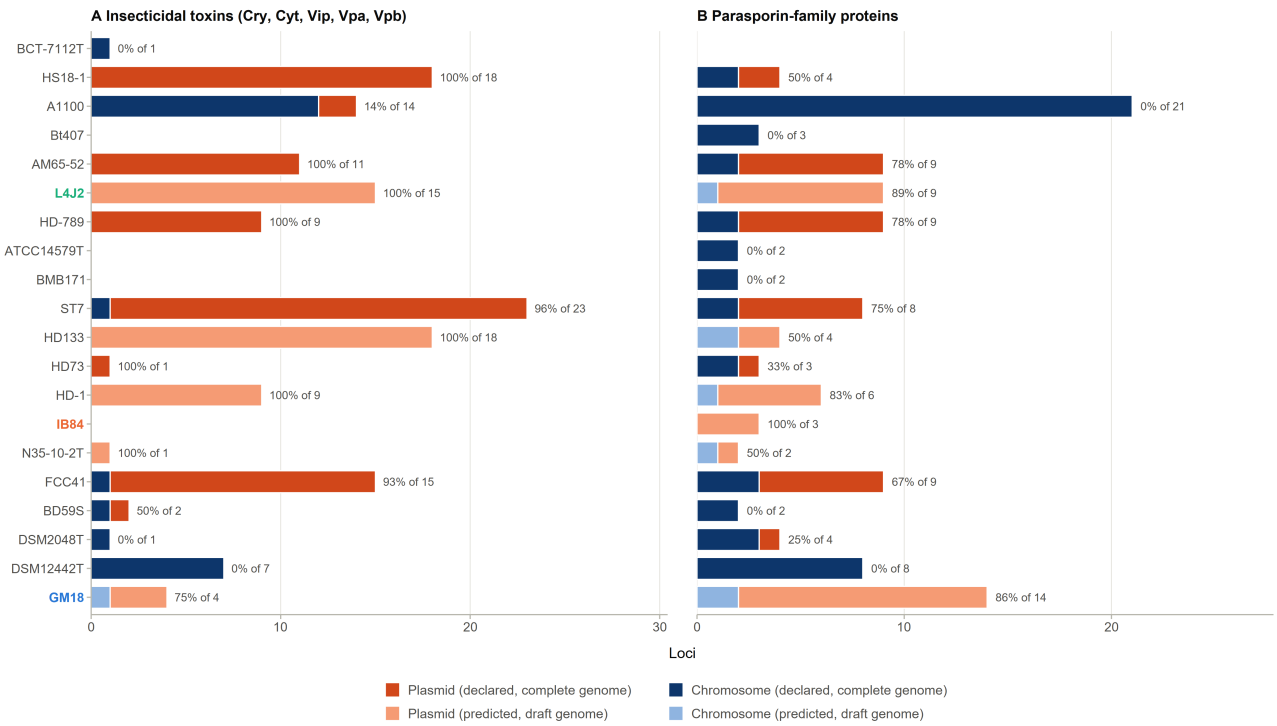


Fig. S9. Replicon location of insecticidal and parasporin loci in each genome. Replicon location of toxin loci in each genome, separated by protein class. Loci were assigned to chromosome or plasmid by combining contig classification with RFPlasmid in the draft genomes, the declared sequence headers in the complete genomes, and an independent assignment with MOB-suite v3.1.9. The text on the right gives the percentage of loci on plasmids and the class total for that genome. In aggregate, 185 of 271 locus entries (68.3%) reside on plasmids. Replicon assignment is declared in the sequence record only for the 14 closed genomes; for the six drafts it is a prediction. That a gene lies on a plasmid does not imply it is mobilisable, which depends on the plasmid carrying conjugation or mobilisation systems (Fig. S14).

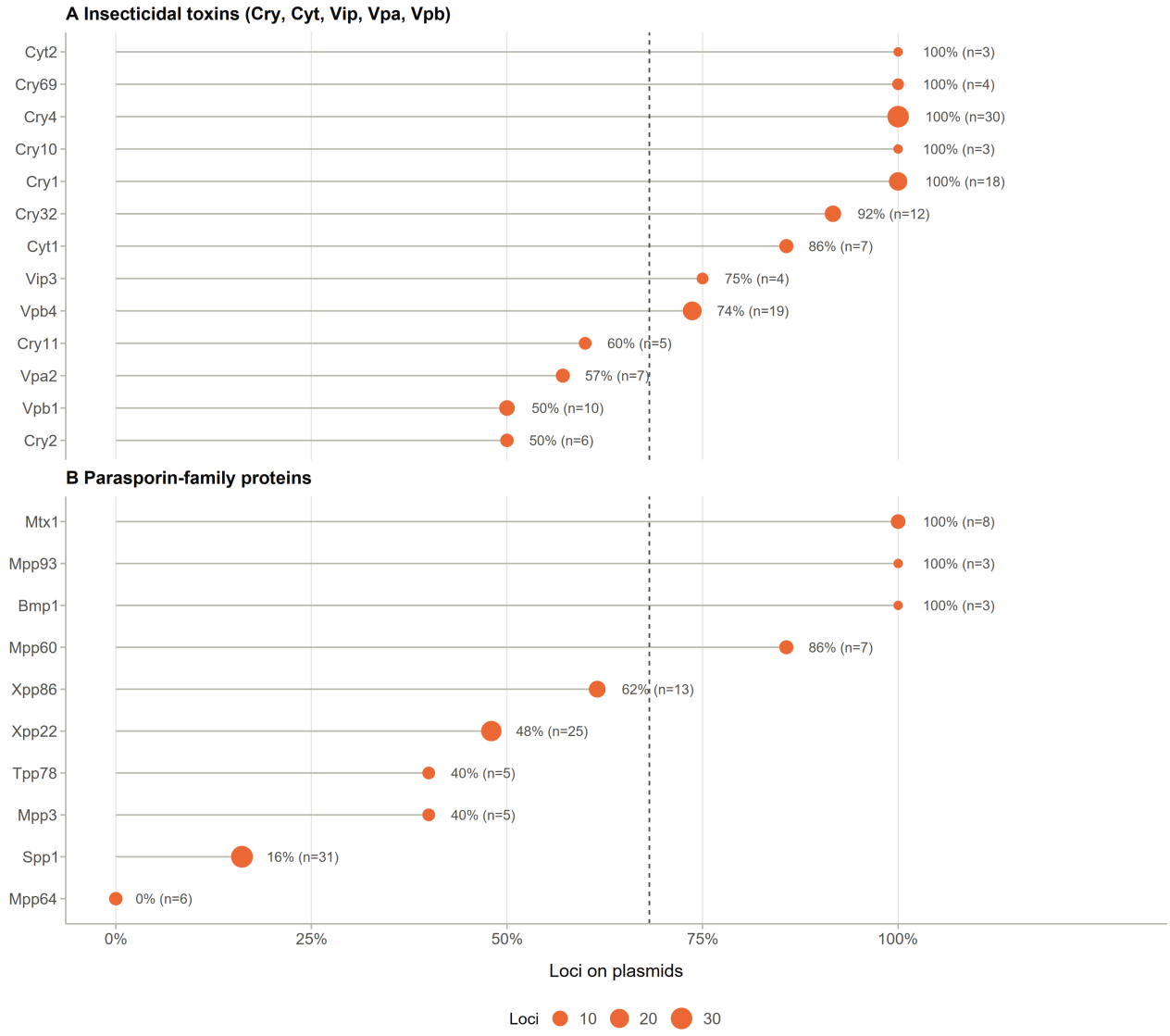


Fig. S10. Proportion of each gene family located on plasmids. Proportion of the loci of each gene family located on plasmids, aggregated across the twenty genomes, for the classical (a) and structure-based (b) classes. Point size is proportional to the number of loci and the dashed line marks the overall proportion of 68.3%. Families with fewer than three assigned loci are excluded, since small families give unstable proportions. The aggregation mixes complete genomes, where the replicon is declared, with drafts, where it is predicted. A family that is plasmid-borne in this panel need not be so throughout the group.

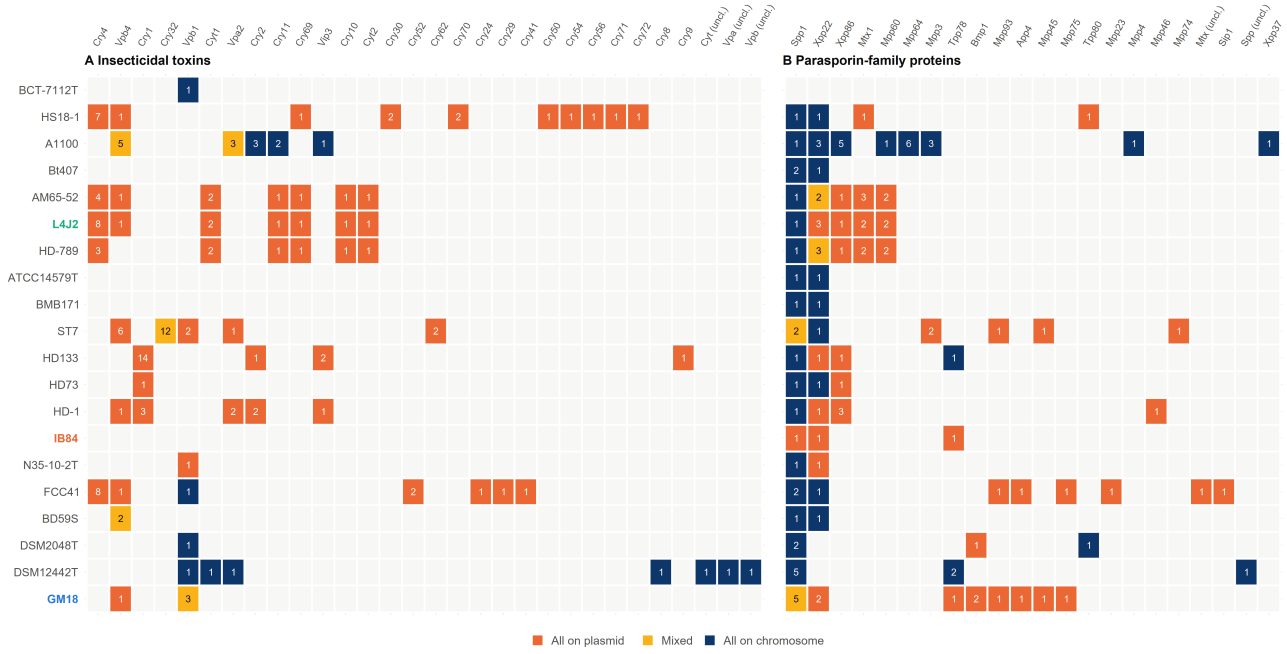


Fig. S11. Replicon location of each toxin family in each genome. Replicon location of each toxin family in each genome, showing locus counts and whether they reside on plasmid, chromosome or both. White cells indicate that the genome does not carry the family. A family may sit on different replicons in different genomes, which is what is expected of genes that move. Cells containing a single locus admit no intermediate category.

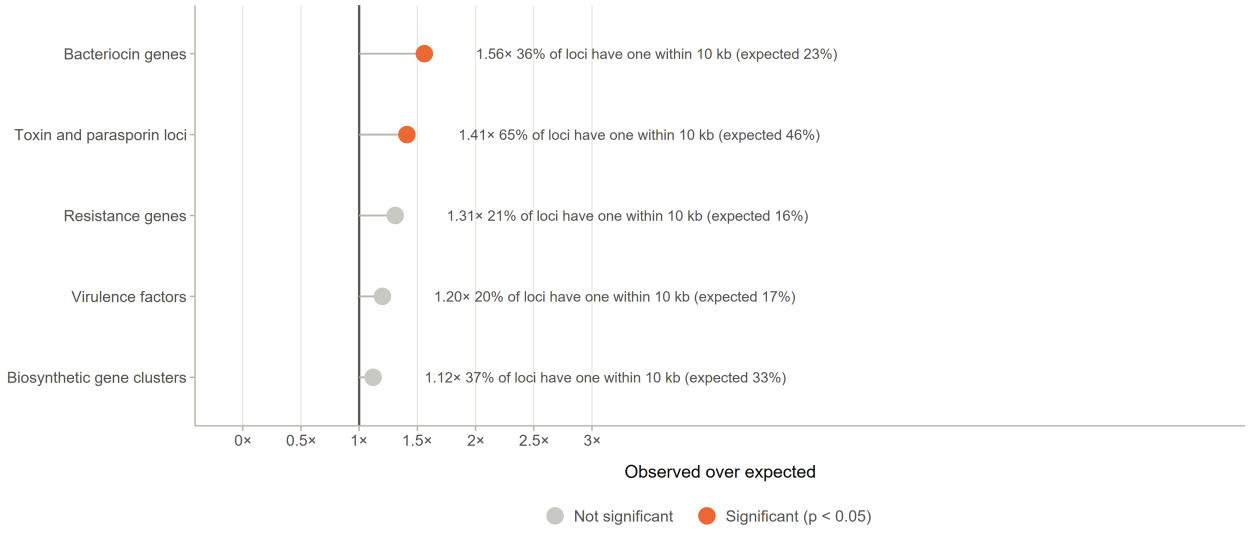


Fig. S12. Insertion sequences within 10 kb of each class of locus. For each class, the ratio of the observed number of loci with at least one insertion sequence within 10 kb to the number expected under random placement within the same replicon, shown with the observed and expected percentages. Elements were detected with ISEScan v1.7.3 and named by comparison with the ISfinder reference database; the analysis is restricted to the complete genomes, where contiguity permits reliable distance measurement, and loci close to one another are merged so that a single neighbourhood is not counted repeatedly. Toxin and parasporin-like loci lie closer to insertion sequences than expected, with 168 of them within 10 kb of an insertion sequence against 46% expected, whereas virulence factors, resistance genes and biosynthetic clusters do not depart from expectation; bacteriocin genes also reach significance, though on a far smaller locus base. Association does not establish direction. The 10 kb window is one choice among three examined, which are nested and therefore not independent, and p-values are not corrected for multiple comparisons across the category, replicon and window combinations tested.

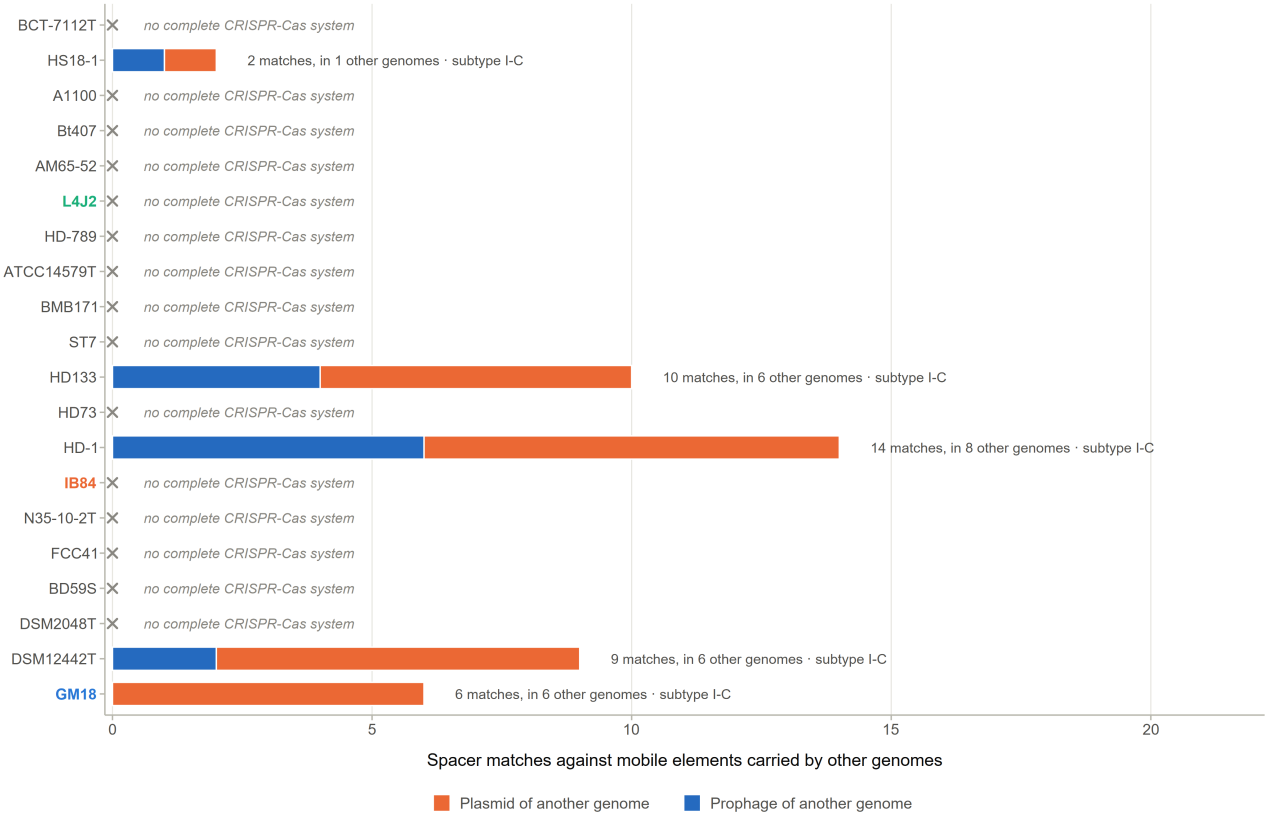


Fig. S13. Foreign mobile elements recorded in the CRISPR memory of each genome. Foreign mobile elements recorded in the CRISPR memory of each genome. Spacers were extracted with CRISPRCasTyper v1.9.0 and CRISPRDetect v2.4 and compared against the plasmids and prophages carried by the other genomes of the panel using CRISPRTarget and BLASTN, retaining only matches supported by both methods; matches against a genome's own elements are excluded. All twenty genomes are shown: the fifteen lacking a complete CRISPR-Cas system appear without a bar because they lack the machinery to record any memory at all, not because they encountered nothing. Of the seventy-one matches recovered, most correspond to plasmids of other genomes, documenting effective circulation of plasmid elements between lineages of the group. The analysis establishes neither the direction nor the date of any transfer, and a false-positive rate was bounded with a shuffled-spacer control.

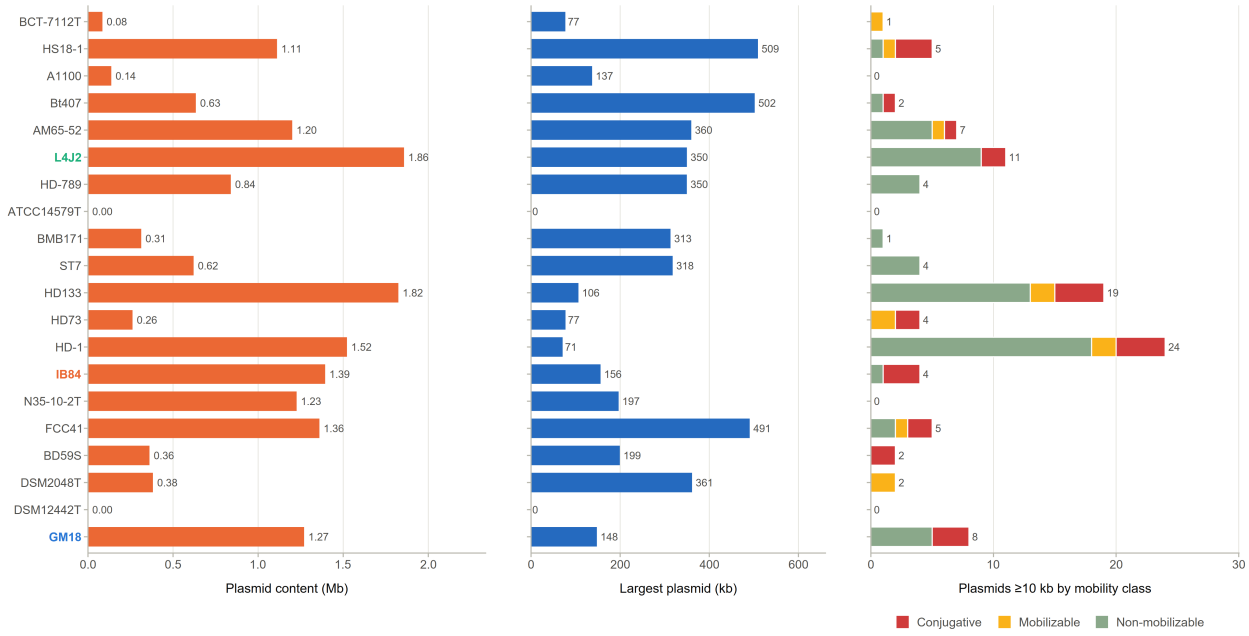


Fig. S14. Plasmid content, largest plasmid and plasmid mobility class per genome. Plasmid content, largest plasmid and mobility class per genome. Plasmids were reconstructed and typed with MOB-suite v3.1.9, with independent typing by PlasmidFinder. The left panel gives total plasmid content in megabases, the centre panel the size of the largest individual plasmid in kilobases, and the right panel the number of plasmids of at least 10 kb classified as conjugative, mobilisable or non-mobilisable. Reconstruction from draft assemblies is incomplete, so counts are lower bounds, plasmid reconstruction from drafts being incomplete for at least 10 kb in the complete genomes. Mobility classification is predictive, based on the presence of the genetic determinants of conjugation rather than on a demonstration of transfer. The three panels use different units and should not be compared with one another.

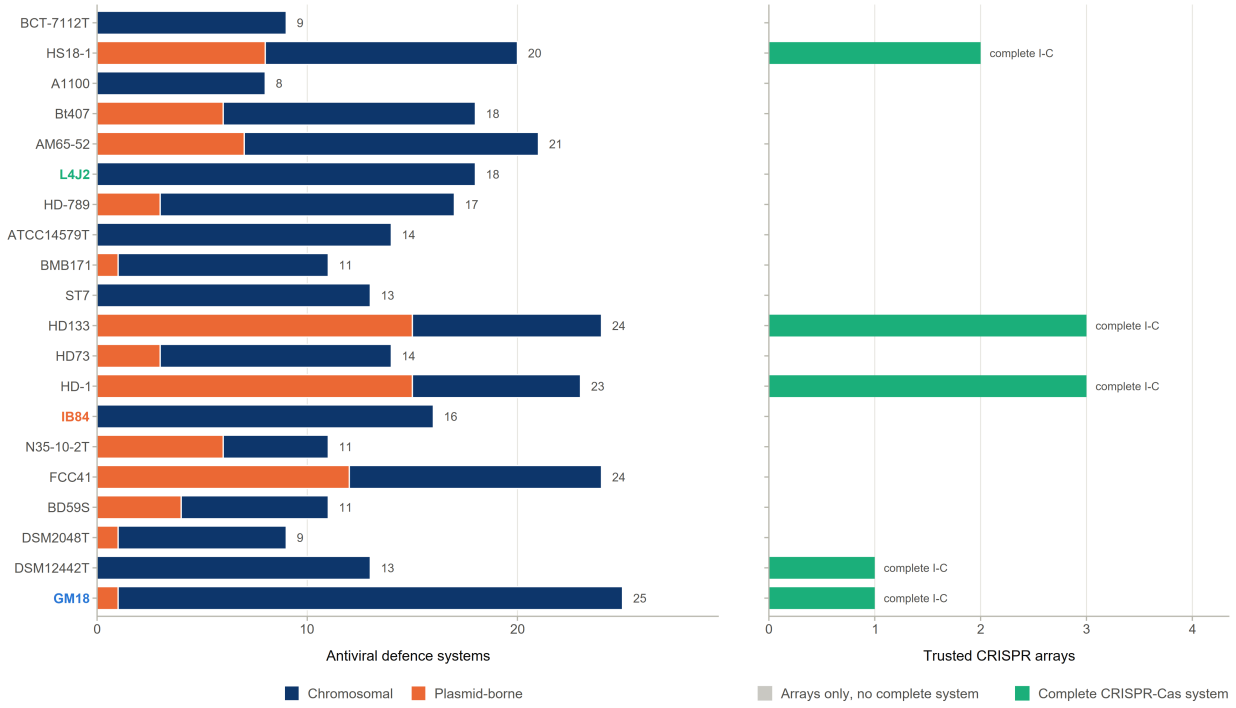


Fig. S15. Antiviral defence systems and CRISPR-Cas content of each genome. Antiviral defence systems and CRISPR-Cas content per genome. The left panel gives the number of defence systems identified by reconciling DefenseFinder v3.0.0 and PADLOC v2.0.0, split by chromosomal or plasmid residence. The right panel gives the number of confident CRISPR repeats detected with CRISPRDetect v2.4 and marks in colour the genomes that additionally possess a complete CRISPR-Cas system typed with CRISPRCasTyper v1.9.0, with its subtype. Only five of the twenty genomes possess a complete system and all are subtype I-C; among the study strains only GM18, which also holds the largest number of plasmid-resident defence systems of the panel. Absence of a complete system is the majority state in this panel and must not be read as a deficiency of a particular strain. Detection counts are lower bounds, reflecting rapidly expanding databases.

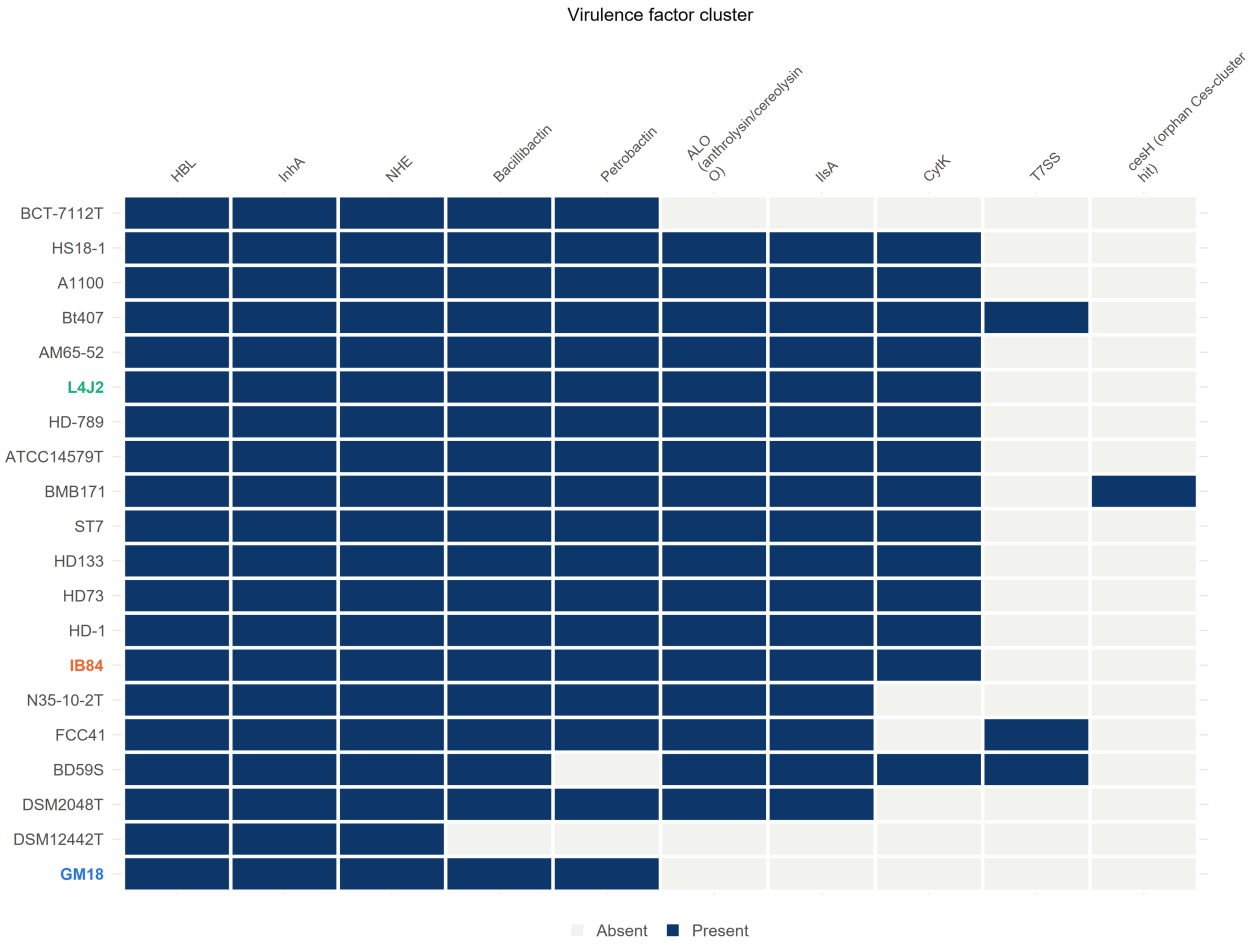


Fig. S16. Virulence factor clusters detected in each genome. Virulence factor clusters detected in each genome (ABRicate v1.4.0 against VFDB, April 2026 release, after identity and coverage filtering). Columns are ordered by decreasing frequency in the panel. The repertoire is largely uniform and includes, in most genomes, the enterotoxigenic determinants characteristic of the Bacillus cereus group, also present in the reference strains used commercially as bioinsecticides. Presence of a virulence gene is not virulence: this figure must not be cited as a risk assessment, and detection depends on database coverage biased towards well-studied pathogens.

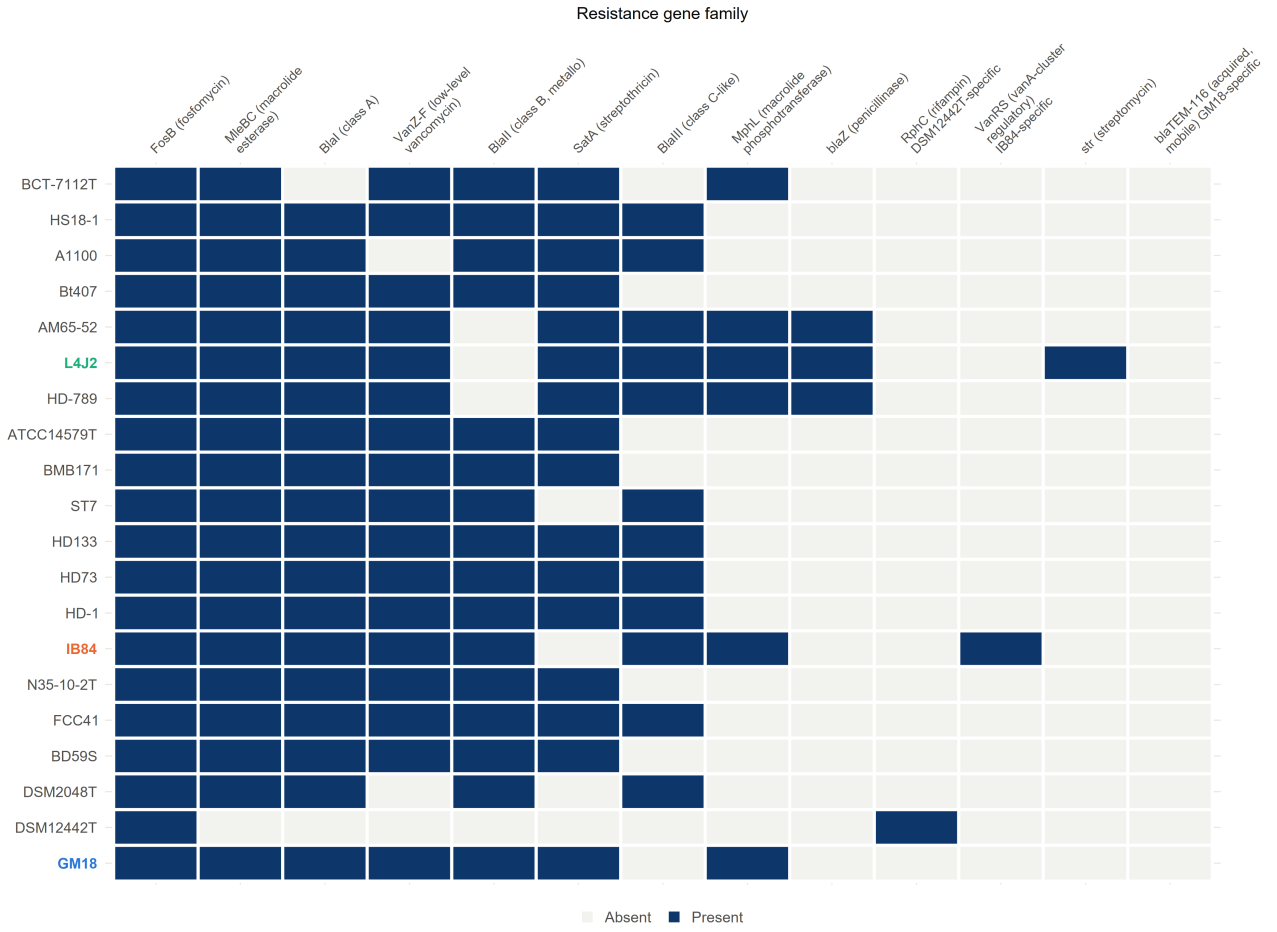


Fig. S17. Antimicrobial resistance gene families detected in each genome. Antimicrobial resistance gene families detected in each genome (ABRicate v1.4.0 against CARD, ResFinder, ARG-ANNOT, MEGARes and the NCBI catalogue, with calls reconciled by genomic coordinate to avoid counting one locus several times). The repertoire is dominated by intrinsic chromosomal β-lactamases of the Bacillus cereus group. Presence of a resistance gene is not phenotypic resistance, which can be established only by susceptibility testing, and this figure does not distinguish intrinsic from acquired determinants; the distinction relevant to biosafety is plasmid location.

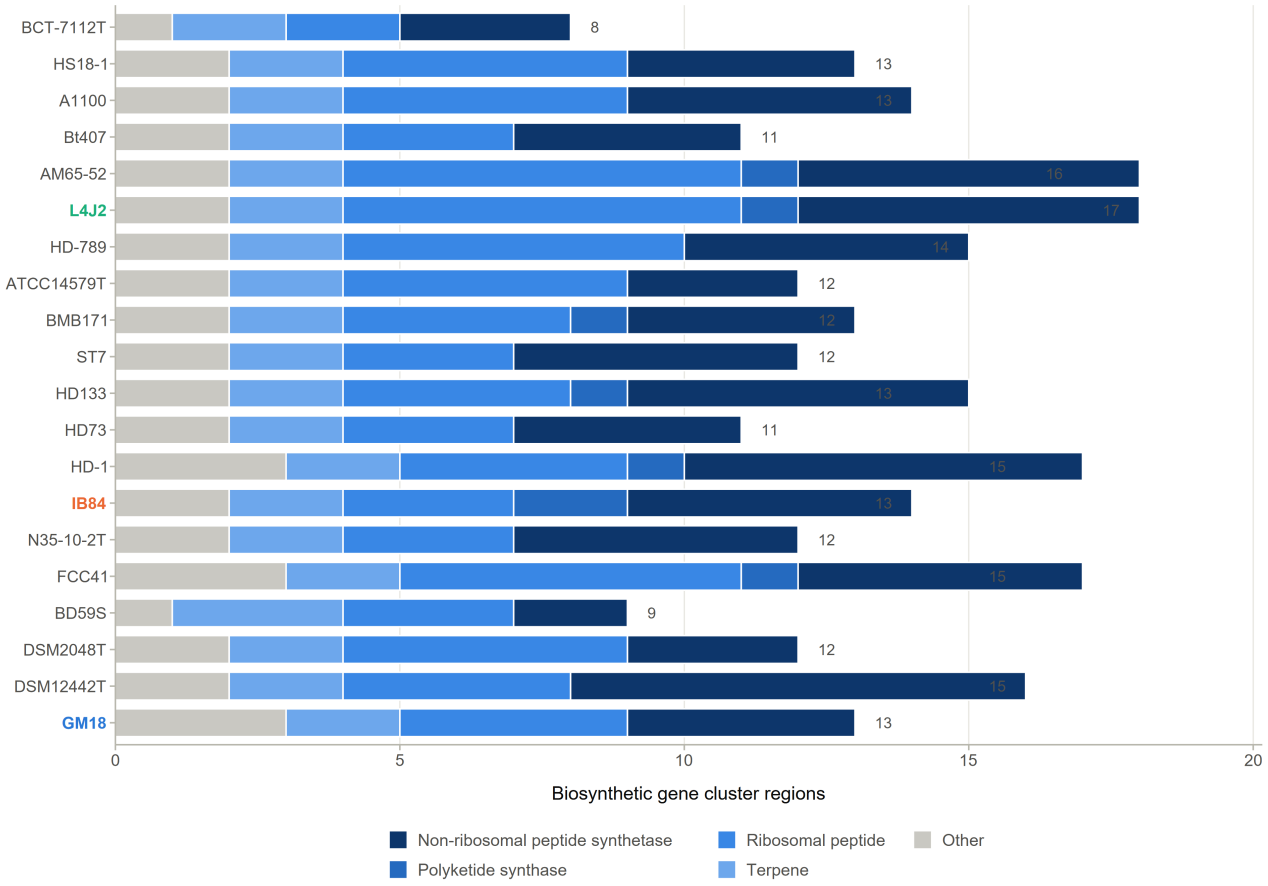


Fig. S18. Biosynthetic gene clusters by class in each genome. Biosynthetic gene clusters per genome by chemical class (antiSMASH v8.0.4, full database). L4J2 has the highest number of regions in the panel, a trait shared with its closest relative AM65-52 and therefore characteristic of the clade rather than of the strain. A predicted region does not imply production of the metabolite, and region counts covary with assembly contiguity, so GM18's value should be read with that caveat.

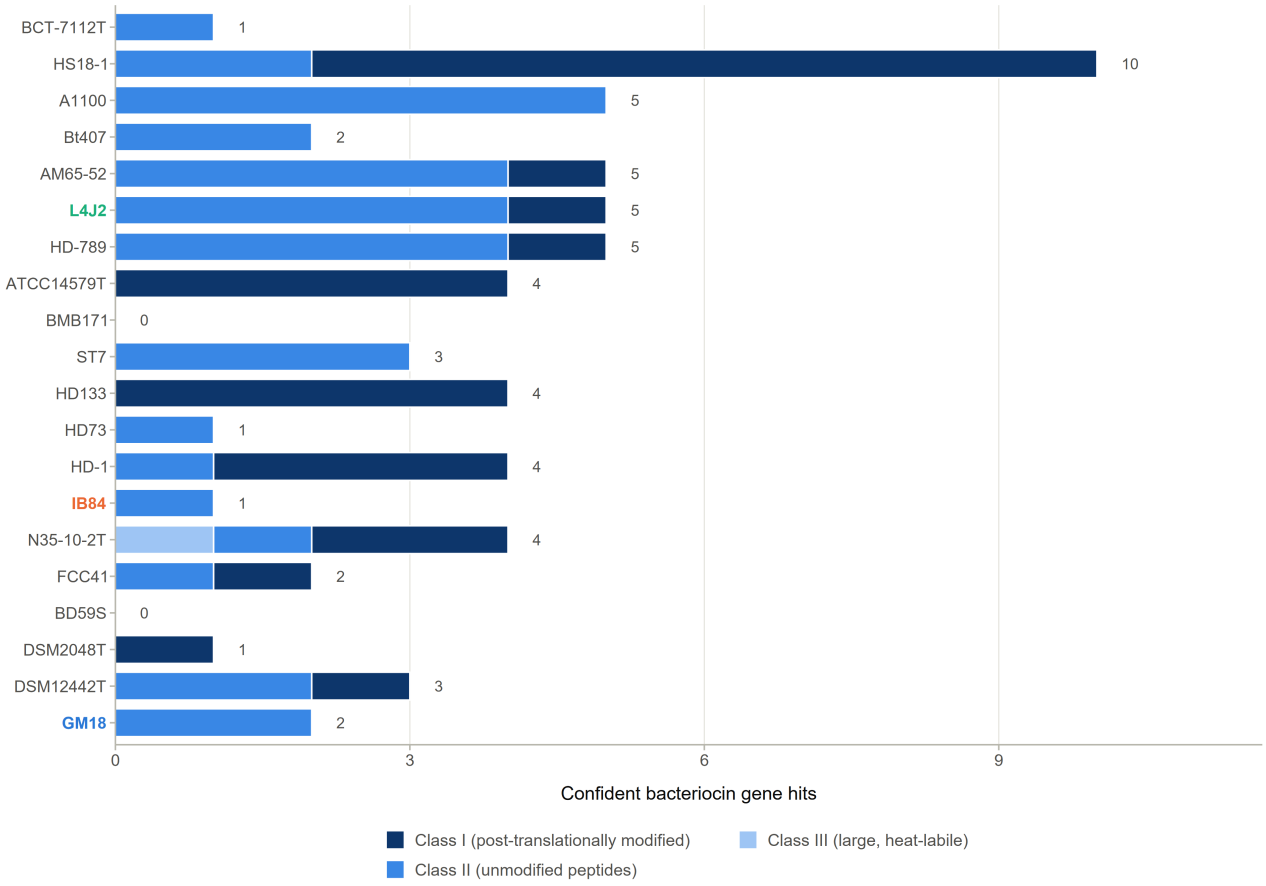


Fig. S19. Bacteriocin genes by structural class in each genome. Bacteriocin genes per genome by structural class, from homology search against the class I, II and III reference protein panel of BAGEL4 and cross-checked against independently detected biosynthetic regions. The repertoire is modest and uniform across the panel. A sequence match does not imply bacteriocin production or antimicrobial activity, and the method detects only bacteriocins with sufficient similarity to the reference panel.

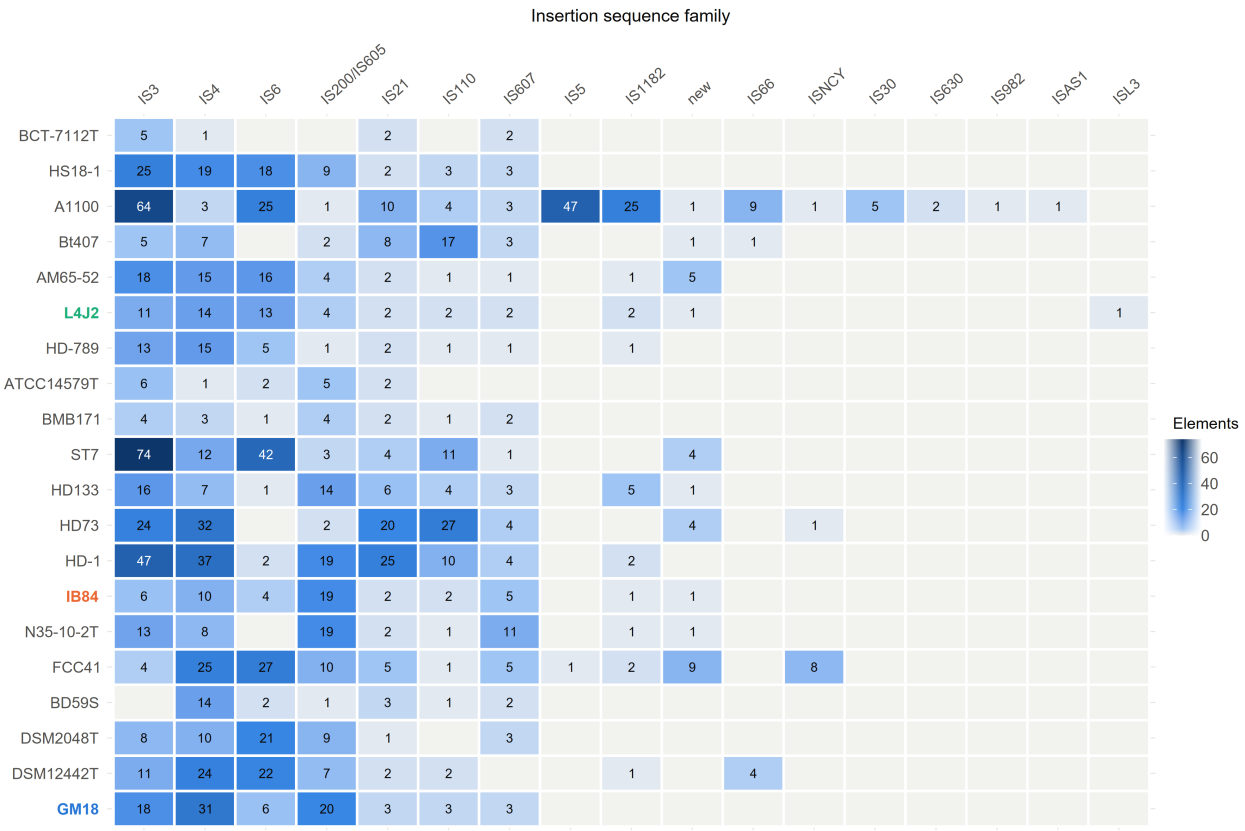


Fig. S20. Insertion sequence families in each genome. Insertion-sequence families present in each genome (ISEScan v1.7.3, named against ISfinder, complemented by a TnCentral search for transposons outside this class). Element density is substantially higher on plasmids than on chromosomes and is associated with pseudogenisation of toxin loci. In fragmented assemblies the count is inflated, because each contig edge generates truncated calls; GM18's value must not be compared directly with those of complete genomes. Detection does not distinguish active elements from degraded, immobile copies.

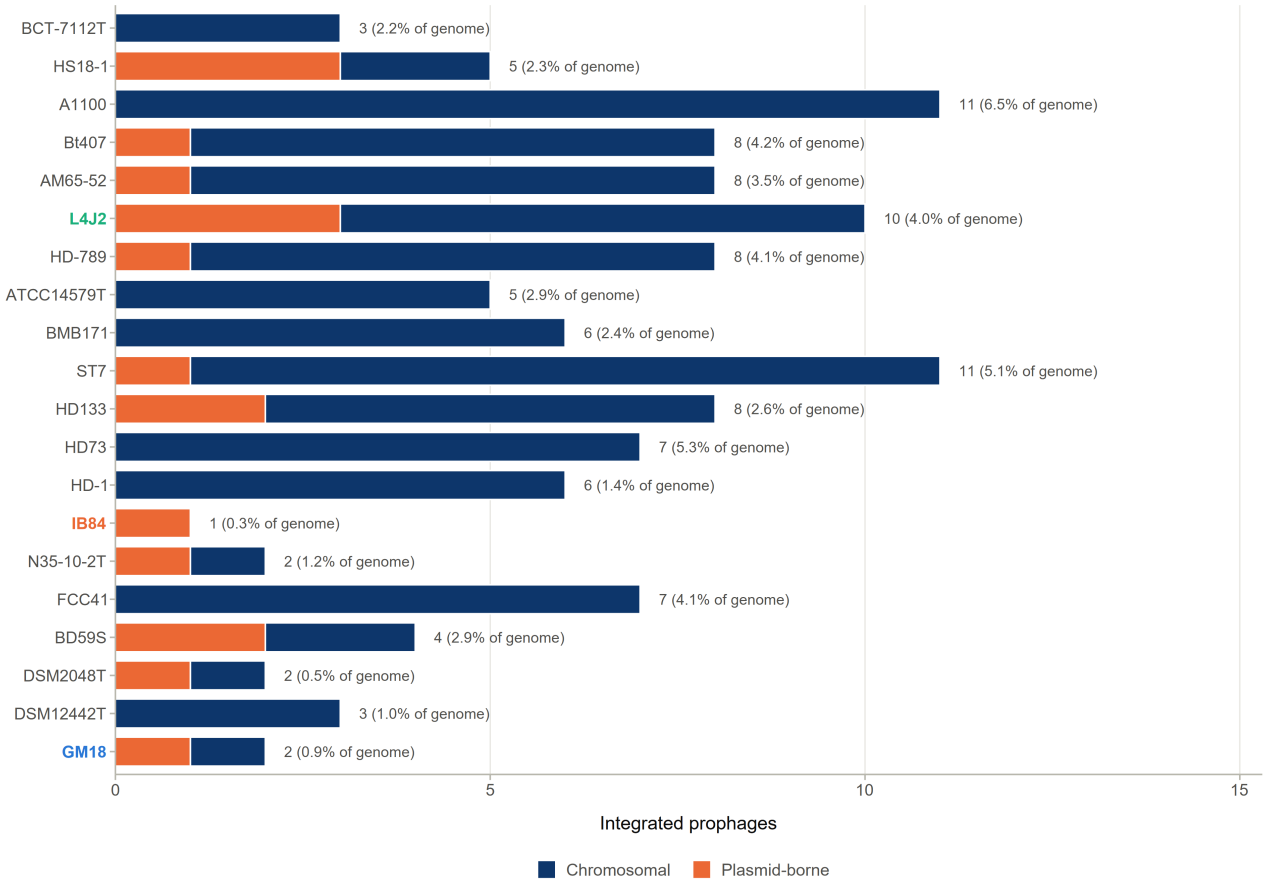


Fig. S21. Integrated prophages per genome and their replicon location. Integrated prophages per genome and their replicon location, from independent detection with geNomad v1.12.0 and PhiSpy v5.0.10 using the Bacillus model, with region completeness assessed by CheckV v1.1.1. L4J2 carries the highest prophage load of the panel, including several prophages integrated into plasmids. Prophage detection is predictive: it identifies regions with viral signature, not demonstrably inducible phages, and region boundaries are approximate.

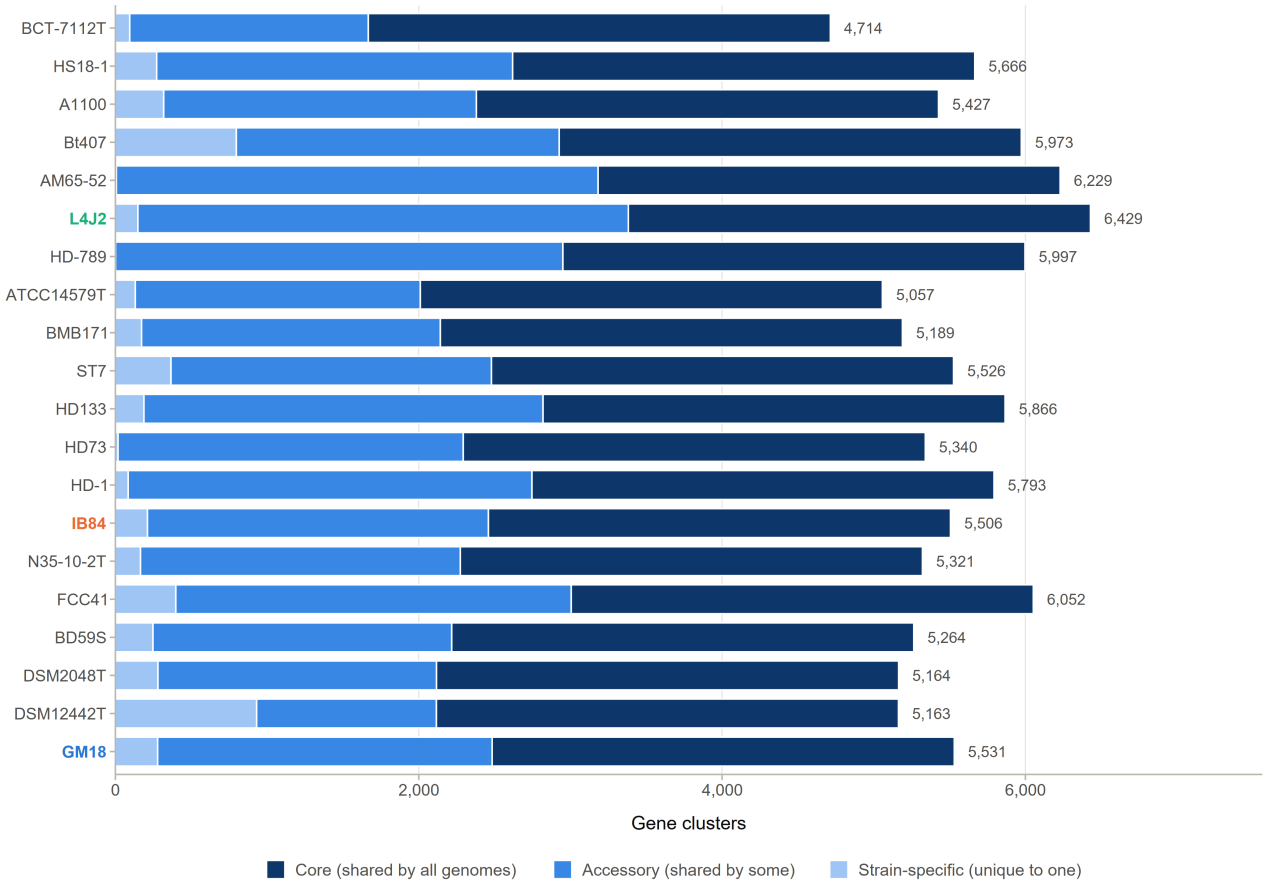


Fig. S22. Pangenome compartments in each genome. Partition of the pangenome in each genome into core, accessory and strain-specific gene clusters (Panaroo v1.8.0 strict mode, complemented by the independent BPGA v1.3 partition). The size of the strain-specific compartment depends on the number of genomes in the panel and is not transferable to other studies; it also includes fragments and annotation artefacts in fragmented assemblies, which affects GM18 most.

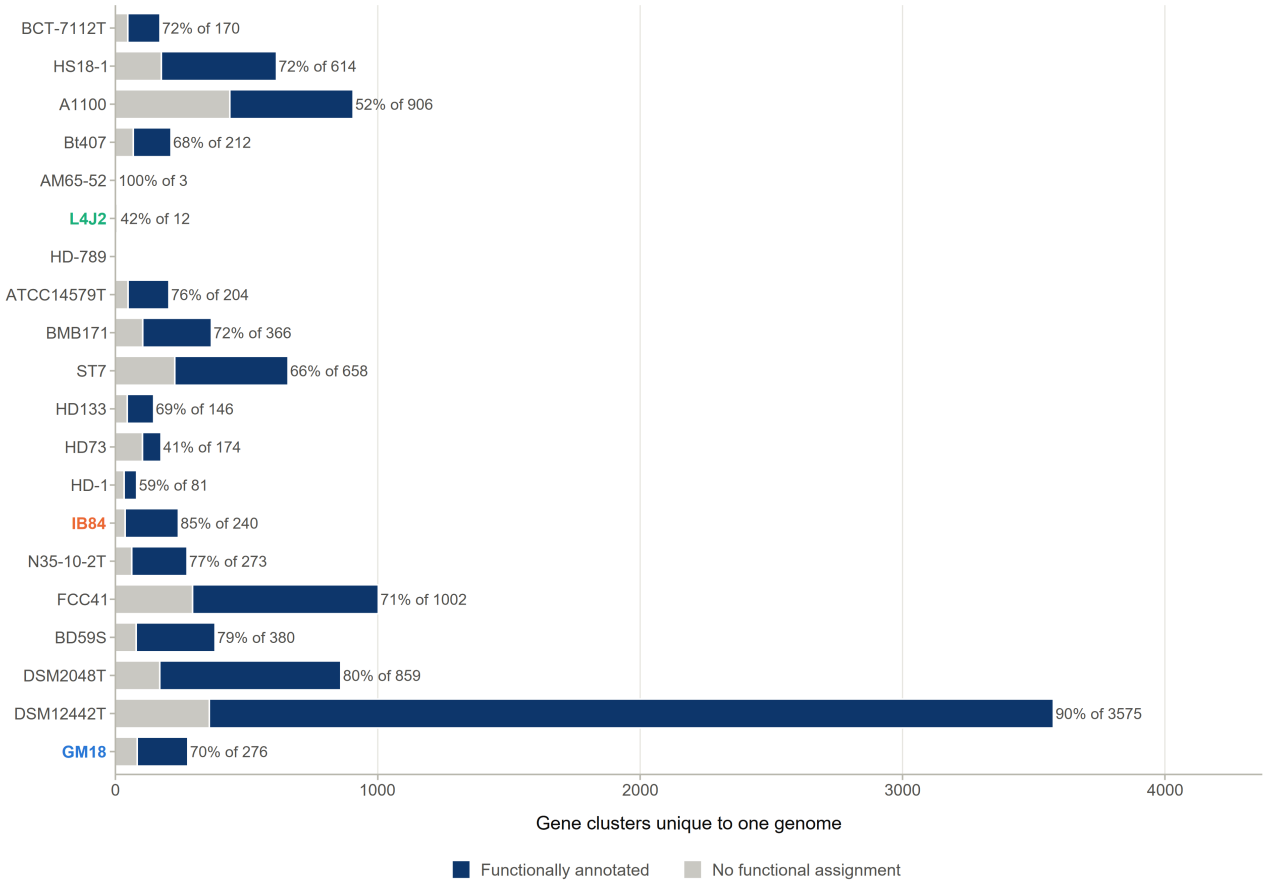


Fig. S23. Strain-specific genes and their functional annotation. Strain-specific gene clusters per genome and the fraction with a functional assignment from eggNOG-mapper v2.1.15. In fragmented assemblies, genes split across contig boundaries inflate this count; L4J2, with 799 contigs, is the most affected genome of the panel. Lack of functional annotation does not imply novelty.

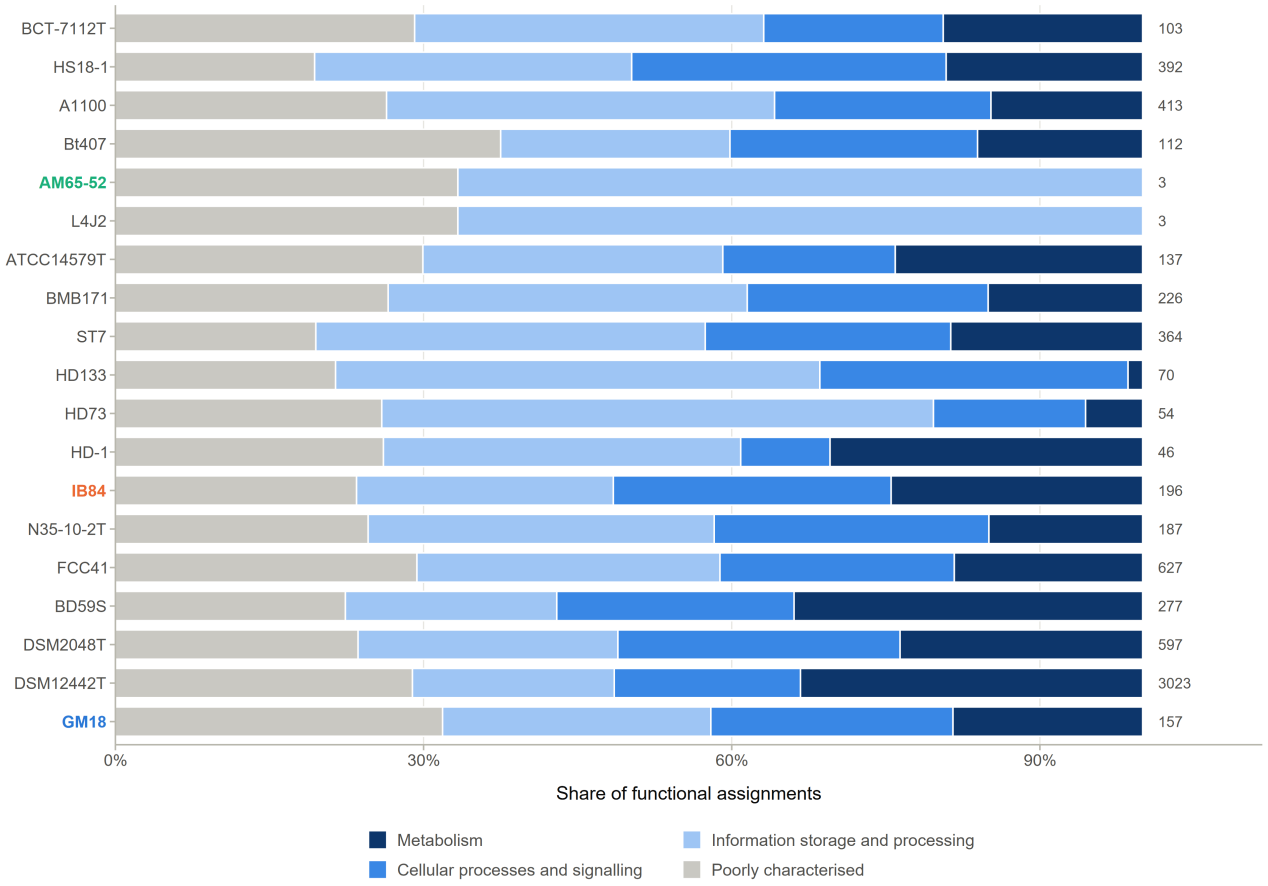


Fig. S24. Functional categories of the strain-specific genes. Functional categories of strain-specific genes, grouped into the four conventional divisions of the orthologous-group system. The denominator is the number of category assignments rather than the number of genes, since a gene with several category letters contributes to each. Poorly characterised categories dominate in every genome, as is usual for this compartment.