Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection

- Bruker Compass DataAnalysis 4.3 was used for chromatography and mass spectrometry. Bruker TopSpin 4.0 was used for NMR data.
- Geneious Prime 2021 was used for sequence data. Tecan iControl 2.0 was used for absorbance (cell viability) measurements. Collection of X-ray diffraction data was processed with the program package XDS version February 5, 2021.

Data analysis

- Raw genome sequencing data were trimmed using Trimmomatic 0.39. Genomes were assembled with SPAdes 3.10.1 and annotated with Prokka 1.12. Completed genome sequences were analyzed and viewed in Geneious Prime 2021. Biosynthetic gene clusters were annotated by antiSMASH 5.0, exploration by BIG-FAM 1.0.0, and classification by BIG-SCAPE 1.0.0 with the MIBiG repository 2.0 reference and PFAM database 32.0. Cytoscape 3.7.2 was used to visualize the BiG-SCAPE network. Pangenome analysis and visualization were performed by anvi'o 6.1. Statistical data of biosynthetic gene clusters were analyzed and evaluated using Origin 2020b and Excel of Microsoft office 365.
- Compound production in a wild-type strain or mutant, as well as MS/MS were analyzed by Bruker Compass DataAnalysis 4.3 and MetabolicDetect 2.1. IC50 values (cytotoxicity and IOC proteasome inhibition) were calculated based on sigmoidal fitting in GraphPad Prism 8.4.3 or 9.0.2. Data from insect immunity bioassays were subjected to one-way ANOVA using PROC GLM of SAS program (SAS Institute, 1989) for continuous variables. All results were plotted by using Sigma plot 12.0. Conventional crystallographic rigid body, positional, and temperature factor refinements were carried out with REFMACS 5.0.32. Protein crystal was determined and analyzed by CCP4 Program Suite 7.1.0.16. Model building was performed with Coot 0.8.7.

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

Policy information about availability of data
All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The genome sequence data that support the findings of this study are available in NCBI GenBank database under accession nos. AYSJ00000000, CP011104.1, CP016176.1, FO704550, FOVO0100000, JADEUF00000000, JAGIUD0000000, JAGJIP00000000, JAGJIQQ00000000, JAGJJRO00000000, JAGJSS00000000, JAGJJWT00000000, JAGJIVV00000000, JAGJIW00000000, LOICDO000000, LOMYO000000, MKGQ00000000, MKGR00000000, MUBJ00000000, MUBK00000000, NC_005126.1, NC_014228.1, NIB500000000, NIBT000000000, NIBU000000000, NIBV000000000, NITY000000000, NIUZ000000000, NJAH000000000, NJAD000000000, NJAK000000000, NJC000000000, NJCH000000000, NJGH000000000, NKT000000000, NKTU000000000, NKNQ000000000, NKNR000000000, XE30TX1_c0031_NODE_001, Xdo_HBL131_1.region001, Xdo_FRM16.1.region005, Xbo_NC_013892.1.region004, Pbod_HBL135_17.region001, Pb6_6.region001, Plum_TTO1_1.region008, Phtr_P11.23.region001, Plau_IT4.1_12.region001, Plau_HB1.3_105.region001, Plau_ENO1_24_scfc009.region001, Pbod_DE6.1_24.region001, Plau_DE2_2_108.region001, Phsb_1.region001, Pbod_LI_007.region001, Pbod_CN4_25_scfc0020.region001, Paeg_BT4T_5.19.region001, P_tem_1.region017 et al.) that exists throughout 45 Xenorhabdus and Photorhabdus genomes was excluded in the analysis, since it turned out that its annotation by antiSMASH 5.0 is a false positive and early reports suggest that this cluster is responsible for ribosomal methylthiolation (doi: 10.1074/mcp.M110.001199 and doi: 10.1021/bi00939a). Two BGCs, Xdo_HBL131_4.region001 and Xdo_HBL135_2.region002 encoding the biosynthesis of gliobactins in X. douceitae HBL131 and Ptem_HBL135_2.region002 encoding the biosynthesis of riiwipptides in P. temperata HBL135 were artificially integrated into their respective genome by CRAGE previously (doi: 10.1038/s41564-019-0573-8), and thus the two BGCs were also excluded in our analysis. The exclusion criteria were pre-established at the outset of the study.

Field-specific reporting

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

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

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

Life sciences study design

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

Sample size
For bacterial genomes, no sample-size calculation was performed. The involved strains (covering all strains in our collection and representing almost all Xenorhabdus and Photorhabdus taxonomy) for pangenome and sequence similarity network analysis were chosen in order to maximize our ability to obtain a comprehensive biosynthetic gene cluster atlas. Sample sizes of compound production in strains and bioassays were based on previous work (https://www.nature.com/articles/s41589-019-01246-1, https://www.nature.com/articles/s41598-019-56892-z, https://sfamjournals.onlinelibrary.wiley.com/doi/10.1111/1462-2920.13845).

Data exclusions
A putative thiopeptide BGC (clusters: Xsus_1.region006, Xse_2.region003, Xsto_4.region001, Xpbo_30.3_21.region001, Xmir_10.region001, Xmau_6.region001, Xjap_NZ_FOVO01000011.region001, Xish_1.region003, Xhom_ANU1.region005, Xhom_ANU2.region003, Xets_11.region001, Xsne_2.region002, XenKL20_c00108_NODE_12.region001, Xek_19.region001, Xehl_28.region001, Xe30TX1_c0031_NODE_38.region001, Xdo_HBL131_1.region001, Xdo_FRM16.1.region005, Xbo_NC_013892.1.region004, Ptem_HBL135_17.region001, Pb6_6.region001, Plum_TTO1_1.region008, Ptem_P11.23.region001, Plau_IT4.1_12.region001, Plum_ILP_35_scfc001.region001, Pbod_HU2.3_20.region001, Plau_HB1.3_105.region001, Plau_ENO1_24_scfc009.region001, Pbod_DE6.1_24.region001, Plau_DE2_2_108.region001, Phsb_1.region001, Pbod_LI_007.region001, Pbod_CN4_25_scfc0020.region001, Paeg_BT4T_5.19.region001, P_tem_1.region017 et al.) that exists throughout 45 Xenorhabdus and Photorhabdus genomes was excluded in the analysis, since it turned out that its annotation by antiSMASH 5.0 is a false positive and early reports suggest that this cluster is responsible for ribosomal methylthiolation (doi: 10.1074/mcp.M110.001199 and doi: 10.1021/bi00939a). Two BGCs, Xdo_HBL131_4.region001 and Xdo_HBL135_2.region002 encoding the biosynthesis of gliobactins in X. douceitae HBL131 and Ptem_HBL135_2.region002 encoding the biosynthesis of riiwipptides in P. temperata HBL135 were artificially integrated into their respective genome by CRAGE previously (doi: 10.1038/s41564-019-0573-8), and thus the two BGCs were also excluded in our analysis. The exclusion criteria were pre-established at the outset of the study.

Replication
In general, all experiments were performed at least three independent times with representative data shown. All attempts to repeat compound determination and bioassays were successful. By comparing Xenorhabdus and Photorhabdus genome sequencing data available in NCBI, we were able to assess the reproducibility of the results (that is, the biosynthetic gene clusters) from each study (https://www.nature.com/articles/s41589-019-0246-1, https://www.nature.com/articles/s41598-019-56892-z, https://sfamjournals.onlinelibrary.wiley.com/doi/10.1111/1462-2920.13845).

Randomization
For all insect immunity assays, Spodoptera exigua fifth instar larvae (L5) aged day 1 (L5D1) were randomly allocated into experimental groups. For Galleria infection assays, the larvae were purchased and bred in-house to similar weight/size. Those deemed equivalents were randomly assigned to the experimental groups.

Blinding
Blinding was not relevant to this study, as we analyzed all sequenced genomes in our strain collection.

Reporting for specific materials, systems, and methods

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

- A description of any restrictions on data availability
- A list of figures that have associated raw data
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
**Eukaryotic cell lines**

Policy information about cell lines

- **Cell line source(s)**: DSMZ (ACC 180)
- **Authentication**: HepG2 cells were obtained from DSMZ and did not undergo additional authentication procedures.
- **Mycoplasma contamination**: negative
- **Commonly misidentified lines (See ICLAC register)**: No commonly misidentified cell lines were used in the study.

**Animals and other organisms**

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

- **Laboratory animals**: The beetle armyworm (Spodoptera exigua) was used in insect immunity assays. The wax moth (Galleria mellonella) was used for Xenorhabdus szentirmaii and the mutants thereof infection assays. Since we used only the larva of insects, the sex could not be determined until they become adults. The lepidopteran insect larvae of S. exigua were reared on an artificial diet (https://agris.fao.org/agris-search/search.do?recordID=KR19910051407) at 25 ± 2°C and relative humidity of 60 ± 5% with 16 h:8 h (L:D) photoperiod. Under these rearing conditions, S. exigua underwent five larval instars (L1-L5) before pupation. They continue being in the L5 phase for three days and then they become prepupa. In all insect immunity experiments, we used L5 day 1 (L5D1). Adults were provided with 10% sucrose for oviposition. G. mellonella larvae were purchased from Zoohaus Haindl, Frankfurt am Main.

- **Wild animals**: The study did not involve wild animals.

- **Field-collected samples**: S. exigua larvae were collected from Welsh onion (Allium fistulosum L.) field in Andong, Korea in 1994. The colony was reared for more than 26 years in the laboratory under the conditions described above. Insects were reared in the laboratory under conditions of 25 ± 2°C constant temperature, 16:8 h (L: D) photoperiod, and 60 ± 5% relative humidity.

- **Ethics oversight**: No ethical approval or guidance was required because no insect pests were used in this study.

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