Chapter 5

Proteomic and metabolomic changes driven by spontaneous genomic rearrangements in *Streptomyces coelicolor*

Zheren Zhang, Chao Du, Dennis Claessen, Somayah Elsayed,
Gilles P. van Wezel and Daniel E. Rozen
Institute of Biology, Leiden University, Leiden, the Netherlands

Abstract

Filamentous bacteria from the genus *Streptomyces* are an important source of antibiotics. Previous work has demonstrated that antibiotic production in these bacteria is organized by a division of labor in which a small fraction of cells reduces its own fitness to produce costly antibiotics that benefit the whole colony. However, little is known about the molecular and cellular mechanisms that underlie these phenotypic changes. In this study, we combine proteomics and metabolomics approaches to begin to identify the mechanistic basis of these altruistic traits in mutant strains. We first confirm that mutants have increased antibacterial activity and production and decreased fitness compared to their parental wild-type. Second, we find that proteins from several secondary metabolite biosynthetic gene clusters, including those for actinorhodin, calcium-dependent antibiotic and coelimycin P1, are upregulated in mutants with higher antibacterial activity. Finally, we show that many proteins in pathways that coordinate *Streptomyces* development and sporulation are significantly downregulated in mutants with reduced fitness. These results uncover mechanistic targets driving the trade-off between secondary metabolism and fitness in these multicellular bacteria. They also provide insights into the cellular basis for the division of labor in this species.

Introduction

Streptomycetes are multicellular filamentous soil bacteria with a complex life cycle. Beginning from a germinating spore, bacteria grow by tip extension into the surrounding medium via connected multi-chromosomal compartments known as a vegetative mycelium. When resources are exhausted, the vegetative mycelium undergoes a transition that leads to the formation of aerial structures containing millions of uni-genomic spores (*7–9*). Concurrently, the colony produces specialized metabolites like antibiotics or antifungal agents (*9*, *10*, *12*). Development and antibiotic production have been extensively studied and this work has uncovered many key genes that regulate these phenotypes, including the crucial *bld* and *whi* genes that are involved in aerial growth and sporulation, respectively (*8*, *111*, *185–187*), and biosynthetic gene clusters (BGCs) for four well-known antibiotics produced by *Streptomyces coelicolor*, namely actinorhodin, coelimycin P1, calcium dependent antibiotic (CDA) and prodiginines (*13*, *145*). However, one particularly intriguing aspect of development and antibiotic secretion has eluded understanding. Since the 1960s (*21*), several research groups have recognized that a high frequency of spores from *Streptomyces* colonies have aberrant primary and secondary metabolism, and further that these phenotypes are associated with large deletions at the ends of the *Streptomyces* linear chromosome (*22*, *25–32*). Although the size and frequency of these genome changes have been well characterized, the mechanisms underlying the phenotypic and evolutionary consequences of these deletions remain poorly understood.

We recently sought to address these issues in *S. coelicolor*. We first demonstrated that mutant strains arising as a consequence of genome deletions have dramatically decreased fitness, manifest as a significant reduction in spore numbers, as well as significantly increased and diversified antibiotic production (*166*). We next showed that this trade-off allows colonies to benefit from a type of division of labor for antibiotic production that maximizes the efficiency of both antibiotic and spore production. In analogy with sterile workers in social insect colonies, mutant cells suffer continued genomic decay and eventual extinction after having produced antibiotics. These results highlighted the evolutionary importance of genomic instability for *S. coelicolor* and potentially other streptomycetes yet left unclear the molecular details underlying the altered phenotypes of these mutant strains.

To address these issues, we used mass spectrometry-based approaches to study proteomic and metabolomic changes in mutant cells. By combining omics data with phenotypic assays, we find that large chromosomal deletions: (i) lead to the upregulation of antibiotic biosynthesis proteins which result in antibiotic overproduction; and (ii) lead to the downregulation of well-known developmental proteins involved in both aerial growth and sporulation, possibly contributing to the dramatically reduced fitness of mutant strains. Our results provide key insights into the mechanisms of the negative phenotypic

correlation between *Streptomyces* development and antibiotic secretion and the division of labor coordinating antibiotic production in this species.

Results

Strains with spontaneous genomic rearrangements show increased antibacterial activity and decreased fitness

We selected five fully sequenced mutant strains from an earlier study (*166*), each containing different sized genomic rearrangements (Fig. 1A); two of these strains contain large genomic deletions that lead to arginine auxotrophy. We measured colony forming units (CFU) and antibiotic activity on SFM agar by spotting colonies on a cellophane disc, to mirror conditions used for the omics analyses below. Consistent with earlier results from assays carried out on agar without cellophane, these experiments revealed significant increases in antibacterial activity and significant reductions in CFU production (Fig. 1B), with a pronounced negative correlation between these traits (fig. S1. $F_{1,4} = 14.74$, $r^2 = 0.787$, $P = 0.018$).

Proteome and metabolome levels vary in different strains

To study the effects of genomic rearrangements on the global expression profiles of proteins and secondary metabolites, we applied liquid chromatography (LC)-MS/MS-based proteomics and metabolomics. Proteomic analyses identified on average 1435 proteins (ranging from 1165 to 1648) in eighteen samples (6 strains × 3 replicates); replicate samples are highly correlated, confirming the reproducibility of these measurements (fig. S2). Principal component analysis (PCA) of both proteomics and metabolomics results indicate a partitioning into three mutant classes, each distinct from the WT (Fig. 2). These results indicate that the proteomes and secondary metabolomes of these strains have been altered due to genomic rearrangements, consistent with the phenotypic results in Fig. 1.

Proteins affecting antibiotic production in mutant strains

PCA score plots (Fig. 2) show excellent correlations between samples without taking phenotypic results (antibacterial activity and CFU production) into consideration. To identify specific proteins that are associated with antibiotic production, we used a PLS approach (partial least squares/projection to latent structures) which examines the correlation between proteomics profiles and the halo size, indicative of antibiotic production, of each strain. As with the PCA score plot of the proteomics experiments (Fig. 2A), different strains from three classes are distinctly partitioned from each other and from the WT, as shown in the PLS score plot ($R^2 = 0.97$) (Fig. 3A). To assess which proteins contribute to increased antibacterial activity, we used a variance importance in projection

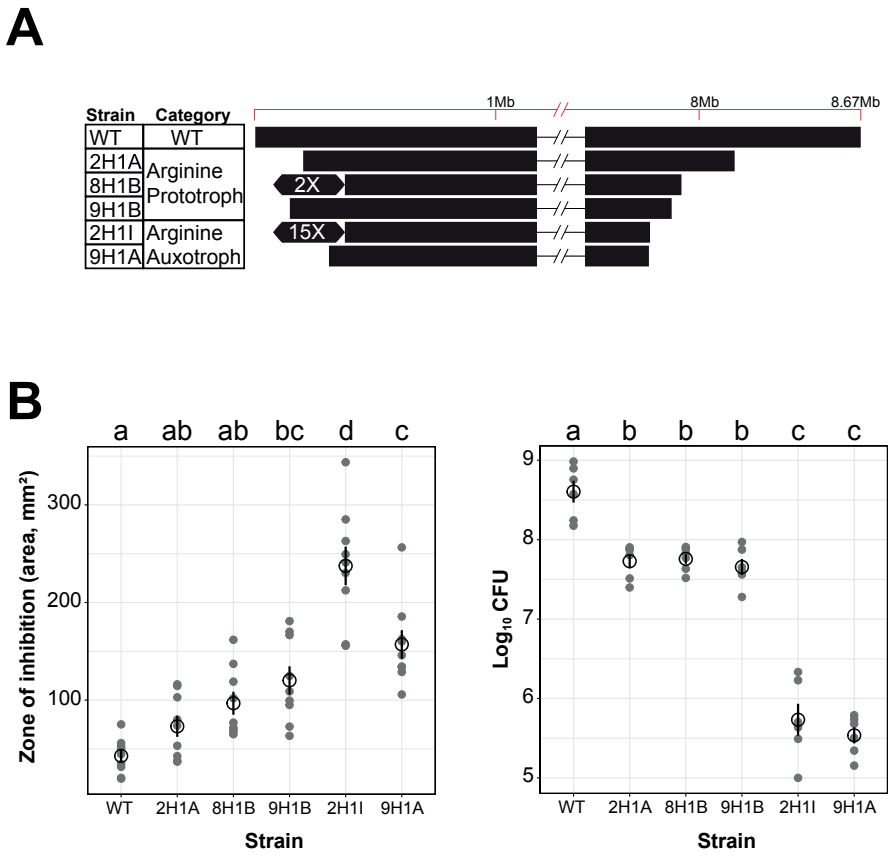

Fig. 1. Overview of the six strains used in this study. Traits shown include (**A**) genomic organization and the category different strains belong to, (**B**) antibacterial activity represented by zone of inhibition against *B. subtilis* (left) and fitness represented by CFU production (right). Values are shown as mean ± SEM. Letters indicate the statistical difference clarified by one-way ANOVA followed by Tukey's tests ($P < 0.05$).

(VIP) plot. High VIP values indicate a larger contribution of that protein to the overall PLS model. Proteins with a VIP>1.4 were ranked by their regression coefficient (table S1 and S2), which reveals that proteins encoded by the BGCs for actinorhodin, coelimycin P1 and CDA, are upregulated in mutant strains and positively correlated with their antibacterial activity (Fig. 3B). Fig. 3C details expression levels of the proteins involved in all four antibiotic biosynthesis pathways compared to the WT strain. These results indicate that, with the exception of prodiginines, most proteins in these BGCs are upregulated in mutants, while none of the proteins in these pathways are downregulated.

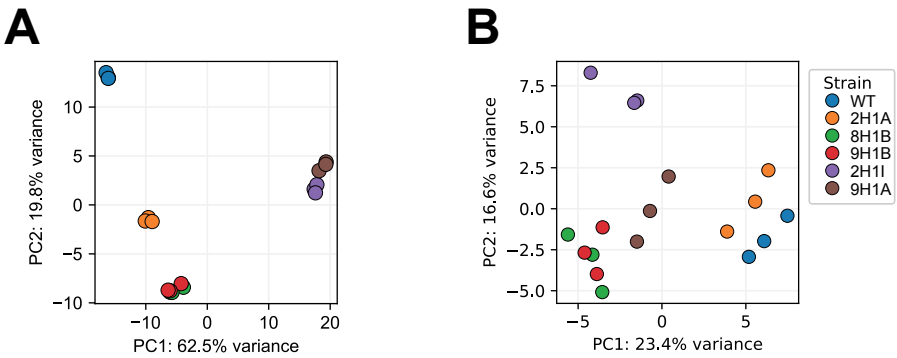

Fig. 2. PCA score plots based on different omics data from six strains. (**A**) Proteomics (**B**) Metabolomics.

Because upregulated BGCs in proteomes perfectly predict increased production of their corresponding compounds (*141*), our proteomics data suggest that mutant strains will produce more antibiotics in terms of quantity and diversity versus the WT. To confirm this prediction, we analyzed metabolomics data to investigate the quantity of the antibiotics indicated above based on their MS information. This revealed that three mutants produced significantly more actinorhodin compared to WT (two-sample *t* tests, all Bonferroni adjusted $P < 0.05$) while no significant difference in undecylprodigiosin production was observed (Fig. 3D and fig. S3), consistent with the proteomics results (Fig. 3B and C). Although we were unable to identify coelimycin P1 and CDA in our extraction and testing conditions, possibly due to the low yield in our extraction method, further experiments revealed that strains have an increased halo size against *B. subtilis* in calcium-supplemented media. This supports the idea that elevated levels of CDA are present in the mutant strains (fig. S4), consistent with the increased expression of proteins in this biosynthetic pathway. These data demonstrate that the increased killing by mutant colonies is caused by the upregulation of proteins underlying production of several different antibiotics.

Downregulation of developmental proteins are linked to the fitness decline

Following the same PLS approach as above, we next sought to identify proteins whose expression changes correlated with CFU. We first confirmed that strains were partitioned from each other and from the WT (R^2=0.99), as shown in Fig. 4A. Next, we ranked proteins with high VIP (VIP>1.4) based on their regression coefficients (Fig. 4B; table S3 and S4), which reveals that several proteins essential for development, cell division and sporulation strongly correlate with reduced fitness and are significantly downregulated

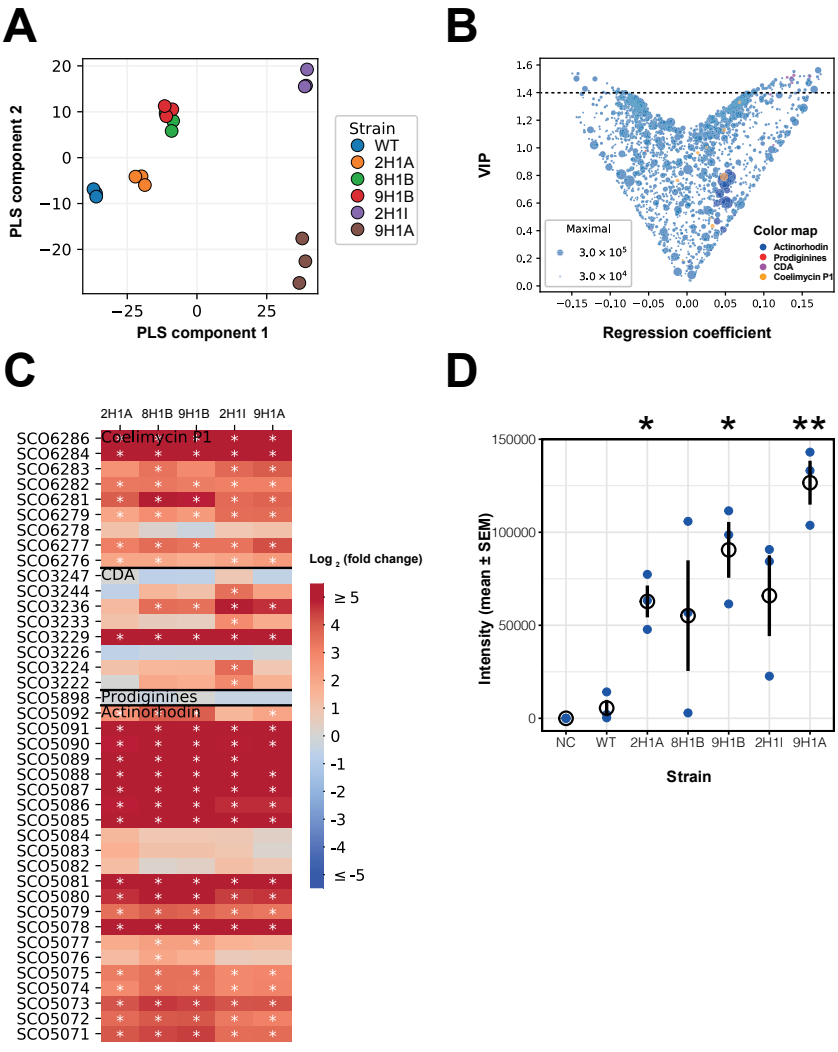

Fig. 3. Antibiotic production increases due to the upregulation of proteins in antibiotic biosynthesis pathways. (**A**) PLS score plot based on proteomics data. (**B**) V-plot of VIP and regression coefficient. (**C**) Log_2 (fold change) of proteins in mutants compared to WT. Asterisks indicate significance according to one-sample *t* tests ($P < 0.05$ and log_2 (fold change) > 1). (**D**) Actinorhodin production quantified by LC-MS/MS. Asterisks indicate significant differences compared to WT based on two-sample *t* tests with Bonferroni adjusted *P* value (* $P < 0.05$, ** $P < 0.01$).

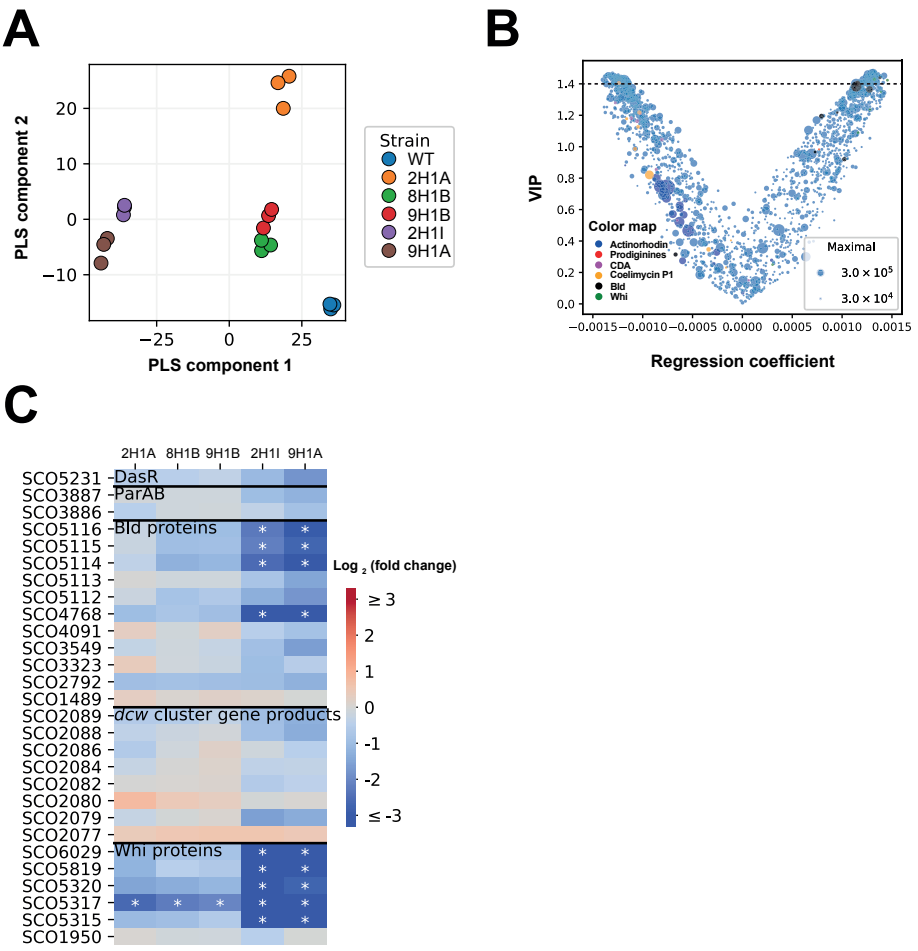

Fig. 4. Decreased fitness correlates to downregulation of developmental proteins. (**A**) PLS score plot based on proteomics data. (**B**) V-plot of VIP and regression coefficient. (**C**) Log_2 (fold change) of proteins in mutants compared to WT. Asterisks indicate the significance according to one-sample *t* tests ($P < 0.05$).

in mutant strains (Fig. 4C). This also reveals a clear negative correlation between CFU and the expression of antibiotic biosynthetic gene clusters, consistent with a trade-off between these functions (Fig. 4B). Key downregulated developmental proteins include gene products of the *bld* genes that regulate the transition to aerial growth, the *dcw* gene cluster that coordinates cell wall formation and cell division, and the *whi* genes that are required for sporulation. Mutations in these pathways have significant negative impacts on *Streptomyces* fitness (*8*, *111*, *188–193*), suggesting that the downregulation we observe here is likely causally related to decreased CFU. In addition to these proteins,

we also examined the expression of proteins involved in different developmental stages in *Streptomyces* (*9*), including ParAB, that regulate DNA condensation and segregation (*194–196*), Ssg proteins that control cell division (*197*, *198*), DasR that globally regulates development and antibiotic production (*109*), and proteins that are related to formation of the surface layers that envelope aerial structures including SapB (*199–202*), chaplins (*203*, *204*) and rodlins (*205*, *206*) (Fig. 4C). We futher observed a pronounced reduction in SCO7036 (ArgG) (table S4), as expected given that two strains (2H1I and 9H1A) are arginine auxotrophs, owing to deletions in this region of the genome (*172*). Interestingly, we find that proteins from the arginine and pyrimidine biosynthesis clusters negatively correlate with fitness (table S4), indicating that changes in primary metabolism can also be a vital factor in influencing the fitness of mutant strains (*142*). Taken together, these data demonstrate the connection between these developmental proteins and fitness differences of different strains.

Discussion

We previously showed that genomic rearrangements occur frequently in *S. coelicolor* and that these changes reduce fitness while increasing antibacterial activity. This inverse correlation leads to a division of labor within the colony, whereby some cells produce spores while others produce antibiotics, but bear the costs of doing so. In this study, we sought to understand the causes of this trade-off by integrating phenotypic and omics data, based on expressed proteins and metabolites. We have found that the upregulation of proteins in antibiotic biosynthesis clusters directly contributes to enhanced antibacterial activity, which is reflected in the overproduction of antibiotics. Although we only detected actinorhodin and undecylprodigiosin in our extracts, we found significant upregulation of CDA and coelimycin P1 clusters at proteomic level, suggesting that this division of labor involves a combination of antibiotics that can suppress different competing bacteria. This combination of antibiotics is likely to be active against different species by targeting different essential functions in diverse species. Because genomic rearrangements occur spontaneously, leading to a variety of genome sizes that each have different effects, it is possible that these behaviors can also be considered a bet-hedging strategy that *Streptomyces* uses to overcome rapidly fluctuating biotic threats.

This study, and our previous work, indicated that increased production of antibiotics is coupled to a decrease in fitness. More specifically, the number of spores was reduced up to 1000-fold. Morphological development has been studied for decades, and the formation of aerial hyphae and spores depend, amongst others, on the classical *bld* and *whi* genes, respectively (*8*, *111*, *129*, *207*). Here we find that the levels of BldM and WhiH are reduced in mutants with low fitness. BldM is an atypical response regulator that is crucial for the formation of aerial hyphae (*208*, *209*). Notably, BldM is one of the last proteins that is formed in the developmental *bld* cascade, while BldKE, which is also reduced, is

known as an early stage regulator (*201*). Taken together, these suggest that genomic rearrangements have global effects on both early and late development stages.

Trade-offs in microbes can be mediated by mutations in genes, often global regulators, that have pleiotropic effects on different traits. For example, mutations in transporters or efflux pumps can simultaneously affect resource uptake and antibiotic import and export (*210*). It was therefore surprising that in *Streptomyces*, the trade-off between fitness and antibiotic production can be explained by gross, and apparently random, genomic deletions. The linear chromosome of *Streptomyces* consists of a core region containing essential genes that are conserved across species, and two so-called "dispensable" arms that contain a diverse set of genes, both within and across species (*34*, *35*). Our results challenge the assumption that these genomic regions are non-essential. Although their deletions are not lethal, they significantly impact several key phenotypes, including sporulation and the regulation of secondary metabolites. The massive size of these deletions suggests that it will be challenging to identify the causal factors leading to increased antibiotic production and reduced fitness. Do deletions lead to developmental stress that, in turn, causes the overproduction of antibiotics, if both functions are coordinated by general stress responses? Or is it the reverse, that deletions lead to the derepression of antibiotics, and that metabolic constraints imposed by secondary metabolites compromise development? At present this remains unclear. However, in either case, our study proves that proteomics is a powerful tool to identify causal factors that underlie these trade-offs. Similar methods, and their integration with other omics approaches, offer potential to extend studies to a wide range of *Streptomyces* species, in the lab and in nature, to test if proteins tied to trade-offs that arise from deletions are conserved.

Our results will help to identify mechanisms driving the trade-off between antibiotics and fitness. They also lead to questions regarding the mechanisms of division of labor in these bacteria. Deletions appear to arise stochastically, possibly as a by-product of programmed cell death. But they may also arise via a regulated process that is responsive to environmental cues. For example, colonies may increase the proportion of cells containing deletions if there are more competing bacteria nearby, requiring an antibiotic "cocktail", while this fraction could decline in the absence of competition. Similar adjustments of division of labor are seen in social insects (*123*). Characterizing the plasticity of division of labor in *Streptomyces* remains an important aim for future work. However, what is already clear is that the irreversibility of these deletions, together with their large and progressively negative impacts, implies that this division of labor is established independently in each colony.

Genomic deletions occur universally in many streptomycetes with linear chromosomes (*20*, *33*, *134*, *136*, *170*, *211*). Some authors have suggested that linear chromosomes may be more mutable, although this has been disputed (*212–214*). It may also be that

chromosome linearity creates recombination hotspots at chromosome ends that causes increased variation in these regions. If high variability is coupled to intrinsic instability, this would serve to maximize within colony phenotypic heterogeneity. In the present case, and perhaps in all *Streptomyces*, the result is a division of labor that is favored by a strong functional trade-off between secondary metabolism and development that is evident phenotypically and in terms of protein and metabolite expression.

Materials and methods

Bacterial strains and culturing conditions

Strains used in this chapter are from previous study and are stored in 20% glycerol at -80°C. During culturing, all strains were diluted to 10^7 CFU ml^{-1}. We then plated 1µl onto the SFM agar covered with a piece of sterile cellophane. SFM agar contains 20 g mannitol, 20 g agar and 20 g soya flour per liter of water. Bacteria were cultivated at 30 °C.

CFU estimation

Colonies grown on cellophane were harvested after 5 days by aliquoting 5 ml Milli-Q water onto the cellophane surface followed by gently scaping all the spores. The 5 ml water was then filtered through a cotton-plugged syringe to remove mycelial fragments. CFU were estimated by plating serial dilutions.

Antibacterial activity assay

Overlay soft agar was prepared by adding 300 µl freshly grown *B. subtilis* (OD_{600}=0.4-0.6) to 5 ml soft LB media containing 0.7% agar, 1% tryptone, 0.5% yeast extract and 1% NaCl. For testing the effects of calcium ions in antibacterial activity, 20 mM calcium nitrate was added to the soft agar (*215*). Each plate was overlaid with 5 ml soft agar on top and incubated at 30 °C overnight. Pictures were taken and the sizes of the zone of inhibition were measured using ImageJ.

Extraction of the proteome

At the time of protein extraction, all mycelia were scraped from cellophane disc and put into 2 ml Eppendorf tubes containing metal beads. Tubes were snap frozen in liquid nitrogen and the proteins were extracted and processed as described before (*166*). Briefly, frozen mycelium was disrupted using Tissuelyser II (QIAGEN, Germany) and then dissolved in disrupting buffer (4% SDS, 0.06 M DTT, 100 mM Tris-HCl pH 7.6, 50 mM EDTA). Cell debris was then removed by centrifugation, and proteins were precipitated using chloroform-methanol method (*151*). Protein pellets were then dissolved in RapiGest SF surfactant (Waters), reduced using DTT, and treated with iodoacetamide. Trypsin (recombinant, proteomics grade, Roche) was then added to digest the proteins. RapiGest SF surfactant

was then degraded by acidification. Resulted peptide solution was then desalted using STAGE-Tips as described by Rappsilber *et al.* (*216*). LC-MS/MS measurement was performed using nanoACQUITY UPLC system (Waters) connected in line with Synapt G2-Si HDMS mass spectrometer (Waters) using an UDMS^{E} method set up as described previously (*154*, *166*).

Quantification of proteomics

Identification and quantification of the proteins was done as described previously (*166*). Briefly, raw data from all samples were first analysed using the vender software ProteinLynx Global SERVER (PLGS) version 3.0.3. Default protein identification workflow with an additional variable modification, acetyl in N-terminal, was used. Reference protein database was downloaded from GenBank with the accession number NC_003888.3. The resulted dataset was imported to ISOQuant version 1.8 for label-free quantification (*154*). TOP3 quantification was filtered to remove identifications meet these two criteria: 1. identified in lower than 70% of samples of each strain and 2. sum of TOP3 value less than 1×10^5. Cleaned quantification data was further subjected to DESeq2 package (version 1.22.2) (*217*) for variance stabilizing transformation (vst) and PCA was conducted using the normalized data.

Extraction of metabolites

The agar with mycelia was cut into small pieces and extracted with ethyl acetate (or methanol) by soaking in the solvent overnight at room temperature. The extracts were washed twice with 30 ml of water and then dried in a fume hood at room temperature for 24 h until further analyses.

Liquid chromatography tandem-mass spectrometery (LC-MS/MS)

LC-MS/MS acquisition was performed using Shimadzu Nexera X2 UHPLC system, coupled to Shimadzu 9030 QTOF mass spectrometer, equipped with a standard ESI source unit, in which a calibrant delivery system (CDS) is installed. The dry extracts were dissolved in 90% methanol to a final concentration of 1 mg ml^{-1}, and 2 µL were injected into a Waters ACQUITY HSS C18 column (1.8 µm, 100 Å, 2.1 × 100 mm). The column was maintained at 30 °C, and run at a flow rate of 0.5 ml min^{-1}, using 0.1% formic acid in H_2O as solvent A, and 0.1% formic acid in acetonitrile as solvent B. A gradient was employed for chromatographic separation starting at 5% B for 1 min, then 5 – 85% B for 9 min, 85 – 100% B for 1 min, and finally held at 100% B for 4 min. The column was re-equilibrated to 5% B for 3 min before the next run was started. The LC flow was switched to the waste the first 0.5 min, then to the MS for 13.5 min, then back to the waste to the end of the run.

All the samples were analyzed in positive (negative) polarity, using data dependent

acquisition mode. In this regard, full scan MS spectra (m/z 100 – 2000, scan rate 20 Hz) were followed by three data dependent MS/MS spectra (m/z 100 – 2000, scan rate 20 Hz) for the three most intense ions per scan. The ions were selected when they reach an intensity threshold of 1000, isolated at the tuning file Q1 resolution, fragmented using collision induced dissociation (CID) with collision energy ramp (CE 20 – 50 eV), and excluded for 0.05 s (one MS scan) before being re-selected for fragmentation.

Multivariate and statistical analyses

Multivariate analyses (PCA and PLS) were done by using scikit-learn (version 0.22) package in Python (version 3.76) following the instruction on the website. Partial least squares/ projection to latent structures (PLS) regression is applied to investigate proteins relevant to a corresponding phenotypic traits (antibacterial activity or fitness).

All other statistical analyses were performed in R (v 3.6.2). Shapiro-Wilk test was performed to confirm the normal distribution of data in the same group and Bartlett's test was performed to assess the equality of variances between groups. ANOVA was performed followed by a Tukey's honestly significant difference (HSD) post hoc test. Two-sample Student's *t* test or Welch's *t* test was performed to compare the difference between groups.

Proteome data availability

The proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD014413 and 10.6019/PXD014413 (*160*, *161*).