

Supporting Online Material

Materials and Methods

Insects and bacterial strains.

Diamondback moth larvae (*Plutella xylostella*) were used as hosts throughout this study (Roth strain, a gift of Prof Guy Poppy, University of Southampton). Field and laboratory experiments used third instar larvae reared on Chinese cabbage *Brassica pekinensis* var. “One Kilo SB” at 25 °C

The invertebrate pathogen *Bacillus thuringiensis kurstaki* strain Dt 7.1.o, isolated from wild cabbage on the Dorset coast in the UK, was used throughout this study. This strain has been identified as sequence type 8 according to the *Bacillus cereus* MLST scheme (22) and as a *kurstaki* strain by the sequence of its *hag* flagellin gene using published primers (31). *B. thuringiensis* is formally differentiated taxonomically from its relatives in the *Bacillus cereus* group by the expression of plasmid carried *Cry* genes. While these plasmids can be highly mobile (32), *Cry* genes tend to be carried by a subset of *B. cereus* group genotypes. Nevertheless, bacteria of the same genotype can be found with or without *Cry* toxin expression (22). Within a single locality, 25-75% of isolates in the broader *B. cereus* group may not produce *Cry* toxins (23).

Spontaneous virulence cheats that failed to express *Cry* toxins were created by curing the Dt 7.1.o strain of its *Cry* toxin plasmid using 0.002% SDS in LB broth and culture at high temperature (42°C) (33). Spores from colonies with a translucent morphology were examined for the presence of bipyrimidal *Cry* toxin inclusions using a Coomassie Blue staining procedure (10 minute fix in 100% methanol, 25 minutes stain in 0.05% Coomassie blue in 45% v/v methanol and 10% v/v glacial acetic acid and then a 5 second de-stain in 45% methanol and 10% acetic acid) and oil

immersion microscopy. Strains were confirmed as being plasmid-cured by PCR of Cry1Ac genes. Inocula for experiments were prepared by densely streaking strains on *Bacillus cereus* specific agar (Oxoid, UK) (BcSA) plates and by culturing at 30 °C for 5 days. Spore/Cry toxin suspensions were recovered from plates, suspended in saline (0.85% w/v NaCl) and then centrifuged and washed in three changes of saline. Spore suspensions were pasteurized (65 °C for 20 minutes) before being stored at 10°C prior to use. Spores were used within two weeks of initial isolation. The density of spores in purified suspensions was estimated using their optical density at 625nm (OD/4.1x10⁻⁶ gives spores per µl).

Within-cadaver competition in the laboratory.

This experiment investigated the competition between Cry positive and Cry negative mutants within single larval hosts at a range of frequencies of Cry producers. Insects were infected individually on 3mm diameter leaf discs of Chinese cabbages within 96 well plates. Each leaf disc was inoculated with 500 spores in 1µl of sterile de-ionized water. We imposed five frequency treatments (0, 10, 50, 90 and 100% Cry producers) and a water only control, 60 larvae were infected in each treatment. The frequencies of Cry producers and Cry negative stains were confirmed by plating out experimental inocula and by examination of colony morphology and Coomassie staining of sporulated colonies. Plates containing leaf discs and larvae were maintained at 30°C for 24 hours. Cadavers were recovered from those wells in which larvae had consumed at least 50% of a leaf disc. Cadavers were stored for 7 days at 30 °C in individual microcentrifuge tubes to ensure complete sporulation of *B. thuringiensis*. Subsequently, cadavers were pasteurized (65°C for 20 minutes), then homogenized in 0.5ml saline (0.85% NaCl) and dilution plated onto BcSA following

established protocols (34). Homogenates were replated onto LB agar in order to produce well-separated colonies, and incubated at 30°C for 7 days to ensure complete sporulation. Colonies from cadavers were directly scored for the presence of Cry toxin inclusions (bipyrimidal crystals) in order to quantify the outcome of competition between Cry negative and Cry producing strains. Sporulated colonies were smeared onto microscope slides into 8x4 grids marked with a diamond pencil before Coomassie staining and microscopic examination of Cry toxin production (described above). The fitness of Cry producers and Cry negative mutants was calculated for each larva using the following metric: probability of death in that treatment*(spores produced per cadaver/spores eaten).

Field experiment.

The field experiment was designed to explore the effect of bacterial density and frequency of cooperators (Cry producers) on the dynamics of Cry producers and Cry negative cheats over a growing season. The experiment took place on Wytham Farm, near Oxford, UK and ran from July to September in 2008. Six week old cabbages (*Brassica oleracea* var. “Wheeler’s Imperial”) were planted out in a 17x24 grid with 1.5 m separation between plants in early July. Plants were bagged with fine polyester netting (1mm mesh) immediately after planting in order to control herbivore populations and exclude natural enemies. One week after planting 35 third instar diamondback moth larvae were added to each plant. After 48 hours each plant was inoculated with *B. thuringiensis* spores in 0.1% Triton X-100 (or a 0.1% Triton X-100 control) using hand-held misters. Plastic barriers were erected between plants during spraying to minimize spray drift. We imposed three density treatments (250, 50 and 10 spores μl^{-1}) and five frequency treatments (0%, 55, 50%, 95% and 100% Cry

producers) in a factorial design and used two controls: insect only (Triton X-100 control); bacteria only (50 spores μl^{-1} of a 50% Cry producer and 50% Cry negative mixture) but no insects.

Immediately after inoculation plants were sampled to establish a time zero (T0) baseline. We cut two leaf samples of approximately 500 mm² from each plant, one from mature fully expanded leaves and one from one of four youngest leaves on the plant. These were placed directly into sterile 30ml polystyrene containers (Sterilin universal) in the field. The following day a pinch of sterile sharp sand and 2ml of saline were added to each universal, which was then vortexed vigorously for 30s. One ml of leaf wash was pasteurized, of which 100 μl was spread onto LB agar containing polymyxin B (100,000 IU litre⁻¹; Oxoid, UK). Pasteurized leaf washes were stored at 10°C and samples diluted and re-plated the following day if required. Colonies of *B. thuringiensis* and *B. cereus* were identified by their smooth, matt-white morphology. Sequencing of *hag* flagellin genes from circa 120 field-isolates confirmed that this was a reliable means of identifying this group. Density of colony-forming units was recorded after 24 hours, and plates were incubated at 30°C for a total of 7 days to ensure complete sporulation, before being scored for Cry toxin production as described above.

At time zero proportions of toxin producers on older and younger leaves were positively correlated within the same plant (Fig S1a). At time point 14, *B. cereus* group bacteria were patchily distributed across plants: 107 leaf samples contained *B. cereus/thuringiensis* bacteria but there were only 21 plants with bacteria on both young and old leaf samples. The correlation between proportion of toxin producers in young and old leaves was still present at time point 14, albeit somewhat weaker (Fig S1b). Given that samples from within the same plant were non-independent, we

adopted the procedure of pooling samples for each plant, when available, in analyses of the dynamics of proportion of toxin producers.

Statistical analysis Analysis of fitness data in the laboratory competition experiment used generalized linear modelling with binomial errors. The primary analyses of the field experiment used maximum likelihood mixed model ANOVA with plant as a random factor nested with time point; time, density and frequency were fixed covariates. We arc-sine square root transformed proportions, to improve normality in preference to using more complex modelling approaches (35). For maximum likelihood models we have reported the degrees of freedom (df) of the model being tested as well the Likelihood ratio (the ratio of the likelihoods of statistical models in the test). Comparison of mixed models with alternative random effects structure used AIC values, models with lower AICs being preferred. Mixed model assumptions (linearity of response versus fitted values; normality of errors) were checked with standard model plots (36). A complementary analysis of fitness in the field experiment used the change in proportion of cheats over time to calculate cheater fitness (17). Thus cheater fitness $V_{cheat} = x_2(1 - x_1) / x_1(1 - x_2)$ where x_1 is the initial proportion of cheats and x_2 is their final proportion. Negative binomial glms were fitted as intercept only models using the MASS package and normalized (Wald) confidence intervals calculated from the standard errors for theta. Significance testing was carried out via sequential deletion of terms from a full model in all analyses. The above analyses were carried out in R v2.6.2 (<http://www.r-project.org>).

Analysis of relatedness

The presence of large Cry toxin parasporal inclusions in Cry toxin producing strains of *B. thuringiensis* facilitates the differentiation of *B. cereus* group strains into toxin producers and non-producers. Proportions of toxin producers at the level of leaf and the plant can be used to readily calculate relatedness at this locus directly, without having to resort general population genetic markers that are typically used to infer population structure. Scoring relatedness based on phenotype is preferable to general comparisons of genetic similarity in this study because Cry toxin production is effectively encoded at a single locus, namely on a large plasmid (37). More importantly, the ability of Cry toxin encoding plasmids to move horizontally between distinct genotypes indicates that overall genetic similarity is not a reliable indicator of whether or not neighbouring strains will express toxins (32, 38). Nevertheless, these phenotypic data can be treated with the same standard quantitative methods used to calculate relatedness from genetic markers (39, 40). We used the following modifications: we included actors as being in the pool of recipients of cooperation as Cry toxins will benefit all individuals in a local patch; allele frequencies of the cooperator gene (p_{ij}) were treated as 1 within individuals expressing Cry toxins (since bacteria are haploid). We used the recommended correction for sample size (40) and summed calculation over individuals within leaves in each population using:

$$\frac{\sum_i^{n_i} \sum_j^{n_j} (p_i - \bar{p}_{(-i)})}{\sum_i^{n_i} \sum_j^{n_j} (1 - \bar{p}_{(-i)})}$$

where i represents each group (leaf samples in this case), j is each individual in that group, and p_i is the frequency of toxin production in each group, $\bar{p}_{(-i)}$ is the mean frequency of toxin producers in the whole population (excluding the contribution of group i). All samples in which the frequency of toxin producers was greater than zero

were classified as containing potential actors. This direct phenotypic method ignores the substantial potential variation in *Cry* genes seen in *B. thuringiensis* (which can confer substantial differences in pathogen host range) and assumes that all *Cry* toxin producers are potentially infectious for our Lepidopteran host. In practice this is a realistic assumption: a substantial majority of *B. thuringiensis* isolates from terrestrial plants express *Cry1Aa*, *Ab* or *Ac* toxins which are highly effective against our host, *P. xylostella*, and the presence of bipyrimidal crystals, in particular, is good indicator of activity against Lepidoptera (41-43). Standard errors and confidence intervals for relatedness were obtained by jackknifing over cabbage plants (44). In addition, plant level relatedness was used in mixed effect ANOVA models; these data were log transformed ($\ln(r+1)$) to improve normality, as per previous studies (45).

Text

Supplementary analyses

Efficacy of manipulative treatments in the field experiment

Several analyses were carried out to examine whether our experimental treatments had a measurable effect on the microbial community on our experimental plants. Firstly, we confirmed that our density treatment resulted in significant variation in observable colony forming units (cfu) per leaf sample at time point 0 (log-transformed spray dose - $F_{1,198} = 23.4$, $P < 0.0001$)(Figure S2a). Mixed model analysis of log bacterial density over the four time points indicated that densities declined with a decelerating quadratic function over the course of the experiment and reached a minimum at time point 28, which corresponded to a crash in host populations (Fig S2b). Density treatments (fitted as a covariate) had a significant

effect on the intercepts ($t = 5.22$, $P < 0.0001$), and on the linear ($t = -5.52$, $P < 0.0001$) and quadratic terms ($t = -4.65$, $P < 0.0001$) of the fitted model.

Secondly, we confirmed that experimental manipulation of the frequency of strains producing Cry toxins persisted until time point 14. Negative frequency dependence in toxin production could be caused by the replacement of our experimental bacterial sprays (with a wide range in frequencies of Cry toxin production) with the natural bacterial community (with a mean frequency between our experimentally imposed extremes). A return to a mean intermediate frequency of Cry toxin production in all treatments at time point 14 would therefore be sufficient to produce negative frequency dependence by chance. If experimental populations were entirely replaced by the natural community with a single mean any effect of experimental treatment at time 0 should be lost by time point 14. The data on the proportion of colonies producing Cry toxins at time point 14 were over-dispersed so we applied a conservative analysis using a beta-binomial error structure (package “aod”). While the experimental treatments initiated with high frequencies of toxin producers (50, 95 and 100 %) were relatively homogeneous (pooling these factor levels resulted in no significant loss of deviance, $\chi^2 = 1.071$, $df = 2$, $P = 0.58$) we could reject the null hypothesis of random colonization because there was significant variation in toxin production among the remaining experimental treatments, including the spray-free controls ($\chi^2 = 11.43$, $df = 4$, $P = 0.022$; Fig. S3). This indicates that experimental treatments, rather than purely random colonization by naturally occurring strains, affected the frequencies of toxin production at time point 14. In addition to the significant effects on frequency of toxin producers, experimental manipulation also led to significant variation in population structure (ie relatedness) that persisted until time point 14 (see below).

We also tested how the presence of host larvae affected the dynamics of competition between Cry producers and cheats over the first three time points of the experiment. This analysis only used data from plants inoculated with a frequency of 50:50 Cry toxin producers: cheats. There was some evidence that the mean frequency of Cry producers was slightly higher in the host addition treatment (post hoc analysis $t = -2.034$, $P = 0.0498$, Fig. S4). However, there was a more rapid decline in the frequency of Cry toxin producers when hosts were present (likelihood ratio = 4.70, $P = 0.03$; Fig S4) suggesting that the faster replication rates afforded by pathogenic niche allowed the cheats to out-compete Cry producers more rapidly.

Relationship between toxin frequency and bacterial population size

A positive relationship between bacterial density and proportion of toxin producers was only evident for bacteria on young leaf material at time point 56 (Fig S5a, leaf age*proportion interaction $F_{1,50}=7.56$, $P = 0.0083$). While this regression was sensitive to a single data point of high influence, a simpler analysis comparing the log densities with or without toxin producers found a more robust interaction with leaf age (Fig S5b, $F_{1,50}=7.45$, $P = 0.0087$, or $F_{1,49}=5.73$, $P = 0.021$ with the reduced data set). Patches of bacteria containing toxin producers also had higher densities than patches without toxin producers at day 28 (mixed model ANOVA, $df=6$, Likelihood ratio = 7.13, $P = 0.0076$) although there was no significant interaction with leaf age at this time point (mixed model ANOVA, $df=8$, Likelihood ratio = 0.0014, $P = 0.97$).

One explanation for the absence of a positive relationship between density and proportion of toxin producers at time point 56 on older leaves was that the majority of

bacterial transmission was occurring on the young leaves where larvae were concentrated. Older leaves are also more subject to haphazard colonization by spores from the soil via rain splash. We explored the natural variation in spatial structure of bacterial populations on young and older foliage (fully expanded leaves) on the final time point of the experiment, assuming that there would be little effect of experimental treatments 56 days after the application of inoculum. We calculated dispersion parameters and their standard errors by fitting negative binomial models to bacterial count data: bacterial were more dispersed on older cabbage leaves (index of dispersion $\phi = 0.387$, $L1 = 0.273$, $L2 = 0.501$ [normalized 95% CLs]) than on younger cabbage leaves ($\phi = 0.104$, $L1 = 0.074$, $L2 = 0.133$ [normalized 95% CLs]).

Relatedness

Relatedness remained high over the course of the field experiment despite fluctuations in the overall frequency of toxin producers: population level relatedness peaked when the frequency of toxin producers in the population crashed at time point 28, the time point with the lowest bacterial population density (Fig S2b). Mixed model analyses of log-transformed plant level relatedness (Fig 4a) were in close agreement with the jack-knifed 99% confidence intervals of relatedness calculated for each frequency treatment (Fig S6). Frequency treatments with high proportions of toxin producers in applied sprays (100% and 95%) produced high plant-level relatedness at time point 0, while lower frequencies of toxin producers (50% or less) led to decreasingly lower relatedness at time point 0 (Fig S6). When relatedness was high initially it decreased rapidly, while when it was low initially relatedness increased, as might be expected given that the definition of relatedness includes the frequency of cooperators in the numerator (Fig S6).

References

31. D. Xu, J.-C. Côté, *Applied and Environmental Microbiology* **74**, 5524 (2008).
32. G. F. L. T. Vilas-Boas, L. A. Vilas-Boas, D. Lereclus, O. M. N. Arantes, *FEMS Microbiology Ecology* **25**, 369 (1998).
33. K. Bernhard, H. Schrempf, W. Geobel, *J. Bacteriol.* **133**, 897 (1978).
34. B. Raymond, D. Davis, M. B. Bonsall, *Journal of Invertebrate Pathology* **96**, 151 (2007).
35. B. M. Bolker *et al.*, *Trends Ecol Evol* **24**, 127 (2009).
36. M. J. Crawley, *The R Book*. (John Wiley, Chichester, 2007).
37. J. M. Gonzalez, B. C. Carlton, *Plasmid* **3**, 92 (1980).
38. J. M. Gonzalez, B. J. Brown, B. C. Carlton, *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **79**, 6951 (1982).
39. A. Grafen, *Oxford Surv Evol Biol* **2**, 28 (1985).
40. D. C. Queller, K. F. Goodnight, *Evolution* **43**, 258 (1989).
41. M. Bizzarri, A. Prabhakar, A. Bishop, *Microb. Ecol.* **55**, 619 (2008).
42. P. H. Damgaard, B. M. Hansen, J. C. Pedersen, J. Eilenberg, *J. Appl. Microbiol.* **82**, 253 (1997).
43. B. M. Hansen, P. H. Damgaard, J. Eilenberg, J. C. Pedersen, *Journal of Invertebrate Pathology* **71**, 106 (1998).
44. D. C. Queller, *Animal behaviour* **47**, 569 (1994).
45. W. D. Brown, L. Keller, *Proceedings of the Royal Society of London. Series B: Biological Sciences* **267**, 1751 (2000).

Figure S1. Relationship between proportion of toxin producers between young and old leaves within plants.

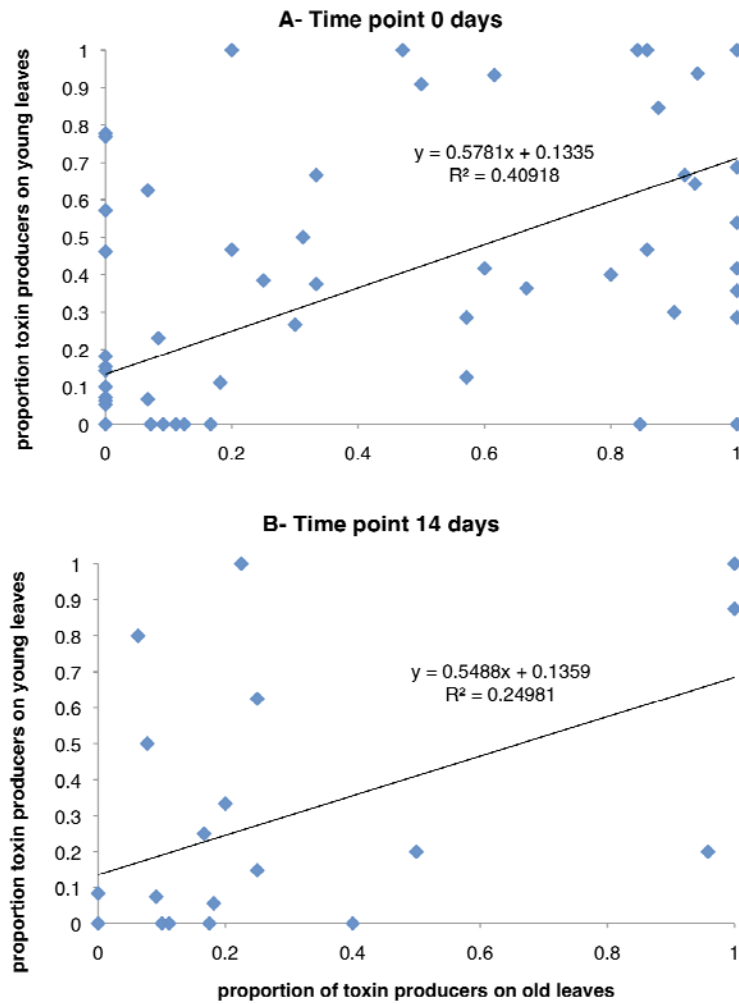


Figure S2. Variation in density of *B. cereus* group bacteria from leaf washes at time-point 0 (A) and over the whole experiment (B). Data in (A) are means $\pm$ SEM, data in (B) are log10 transformed data with a fitted quadratic model.

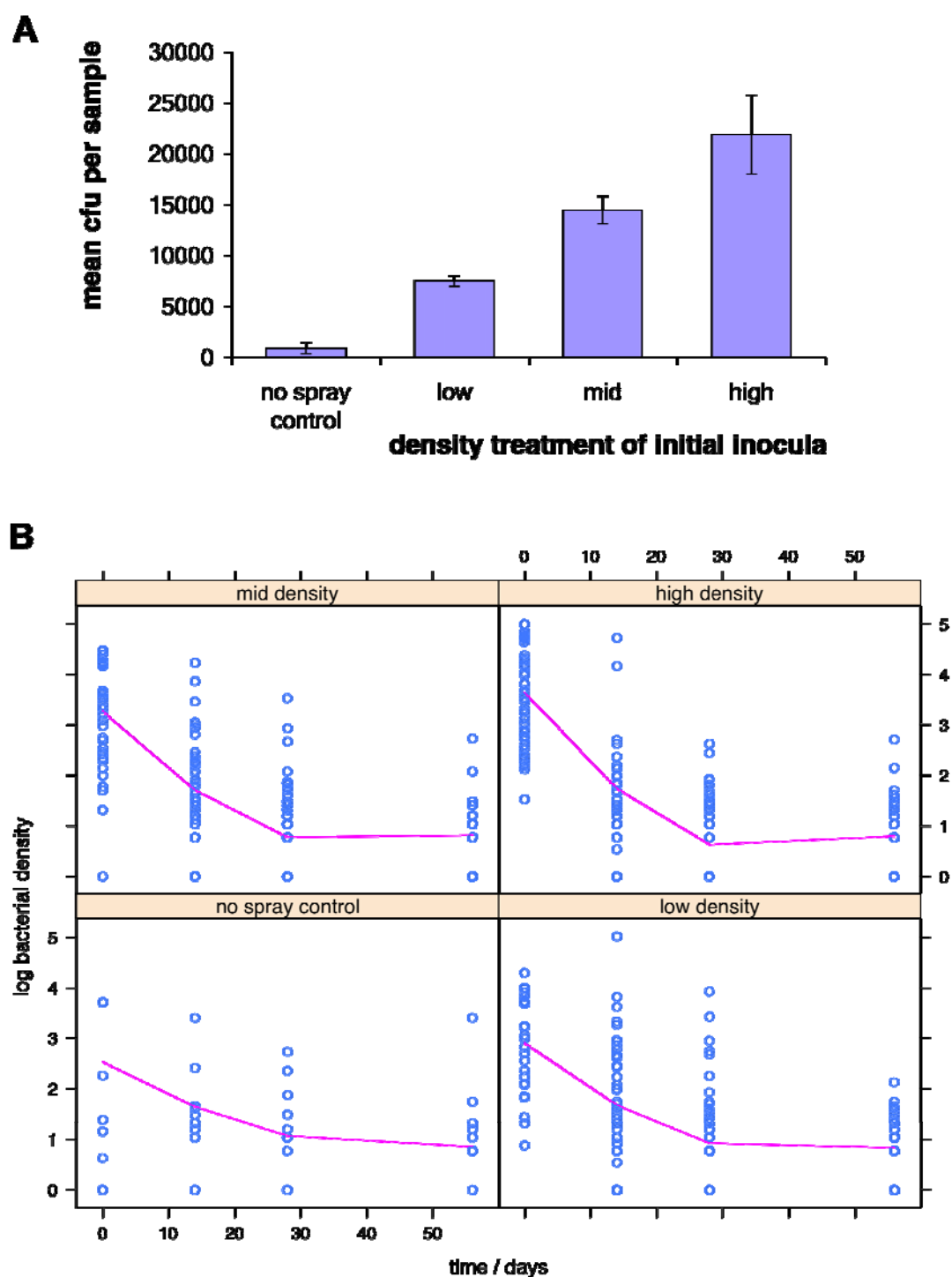


Figure S3. Effect of experimental treatment on mean proportion of toxin producers at time point 14 in the field experiment. Data are mean proportions $\pm$ SEM (calculated according to the normal approximation to the binomial).

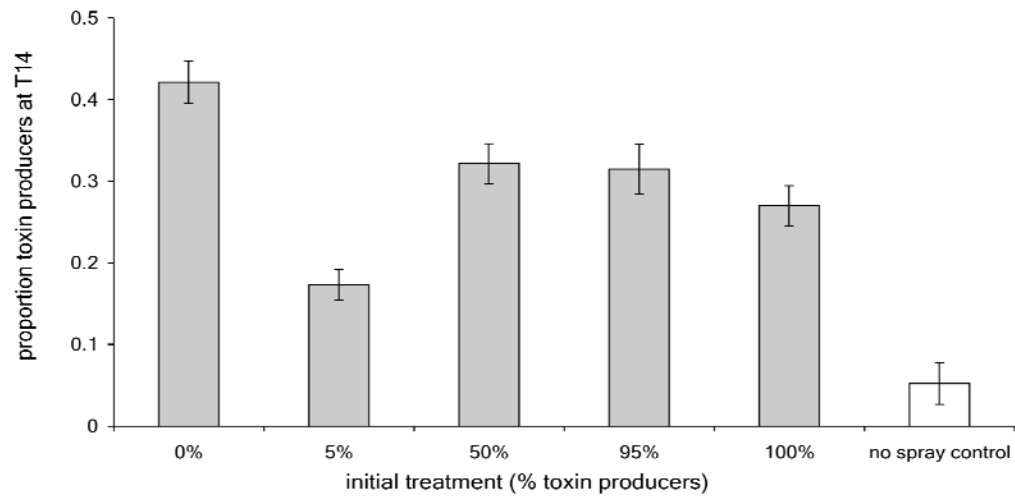


Figure S4. Effects of hosts on dynamics of toxin production in the field: data are from plants inoculated with a 1:1 ratio of toxin producers and non-producers at time zero; triangles and solid lines represent plants with hosts; crosses and dashed lines represent no host control plants. Data are proportions of toxin producers per plant, but were analysed after arc-sine transformation.

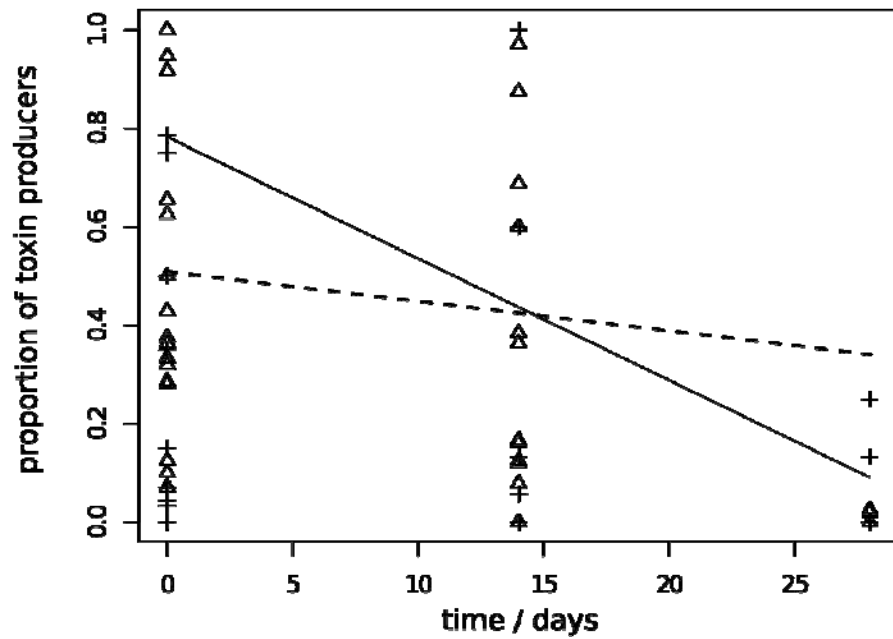


Figure S5. The relationship between proportion of toxin producers and bacterial density (of the *B. cereus* group) in the final time point (day 56). (A) A sunflower plot showing data and fitted models for young leaves (in green) and for mature leaves (in gray). Multiple overlapping data points are indicated by the line or "petals" emanating from each data point: the number of petals indicates the number of overlapping data points. (B) bacterial densities on mature and young leaves for populations with and without Cry toxin producers. Data are means $\pm$ SEs.

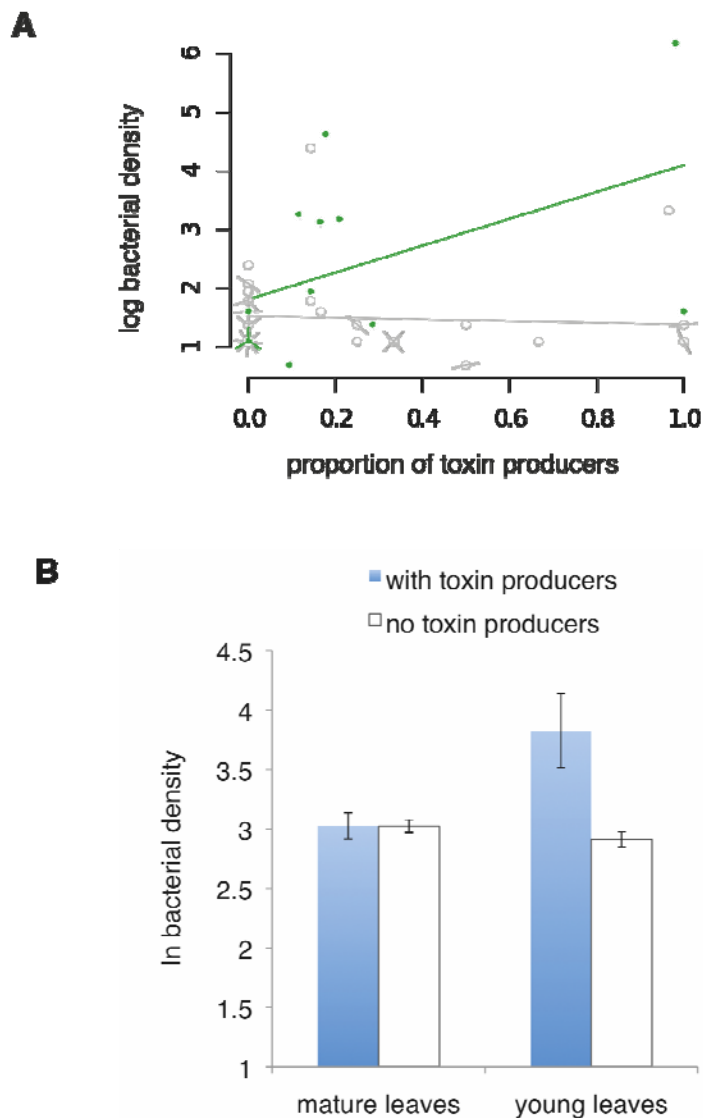


Figure S6. Treatment level relatedness with jack-knifed 99% confidence limits calculated for the first two time points of the field experiment.

