nature biotechnology

Expression of an insecticidal fern protein in cotton protects against whitefly

Anoop Kumar Shukla[1,10], Santosh Kumar Upadhyay[1,9,10], Manisha Mishra[1,2,10], Sharad Saurabh[1,10], Rahul Singh[1,10], Harpal Singh[1,10], Nidhi Thakur[1,2,10], Preeti Rai[1], Paras Pandey[1,2], Aradhana L Hans[1,2], Subhi Srivastava[1], Vikram Rajapure[1], Sunil Kumar Yadav[1,2], Mithlesh Kumar Singh[1,2], Jitendra Kumar[3], K Chandrashekar[1,9], Praveen C Verma[1,2], Ajit Pratap Singh[2,4], K N Nair[2,4], Smrati Bhadauria[5], Muhammad Wahajuddin[6], Sarika Singh[5], Sharad Sharma[5], Omkar[7], Ram Sanmukh Upadhyay[8], Shirish A Ranade[1,2], Rakesh Tuli[1,9] & Pradhyumna Kumar Singh[1,2]

Whitefly (*Bemisia tabaci*) damages field crops by sucking sap and transmitting viral diseases. None of the insecticidal proteins used in genetically modified (GM) crop plants to date are effective against whitefly. We report the identification of a protein (Tma12) from an edible fern, *Tectaria macrodonta* (Fee) C. Chr., that is insecticidal to whitefly (median lethal concentration = 1.49 μg/ml in *in vitro* feeding assays) and interferes with its life cycle at sublethal doses. Transgenic cotton lines that express Tma12 at ~0.01% of total soluble leaf protein were resistant to whitefly infestation in contained field trials, with no detectable yield penalty. The transgenic cotton lines were also protected from whitefly-borne cotton leaf curl viral disease. Rats fed Tma12 showed no detectable histological or biochemical changes, and this, together with the predicted absence of allergenic domains in Tma12, indicates that Tma12 might be well suited for deployment in GM crops to control whitefly and the viruses it carries.

Insects cause substantial crop losses through direct and indirect damage. GM crops such as soybean, corn and cotton that express δ-endotoxin(s) from *Bacillus thuringiensis* (*Bt*) have been planted worldwide[1] to prevent crop losses caused by herbivorous field pests. However, none of the transgenic technologies currently available for cultivation can prevent infestation by sap-sucking pests. This group of insects includes whiteflies, aphids, leafhoppers, mirid bugs and thrips[2]. Whiteflies are polyphagous insects that infest several hundred plant species, including economically important crops[3,4]. They cause damage by sucking phloem sap, resulting in early wilting, premature defoliation, stunted growth and eventually yield loss. Indirect damage to crop plants is caused by the whitefly-excreted honeydew, which promotes fungal growth on leaves and fruit surfaces and thereby reduces the quality and market value of produce. Whitefly also vectors nearly 200 plant viruses that can infect infested plants and cause a complete loss of yield[5,6]. Owing to a reduction in pesticide use in fields planted with *Bt* cotton, the incidence of unaffected pests, including whitefly, and concomitant crop damage have been increasing substantially[7].

Different control strategies based on physical barriers, pesticides, biotic agents and host-plant resistance have been used to combat hemipteran pests in general and whiteflies in particular[6]. Integrated pest management, involving one or more of these strategies, is also practiced in some countries. But none of these strategies can effectively control whitefly. Transgenic crops resistant to whiteflies are desirable because the technology could be supplied as seeds. However, no transgene has been reported to date that provides reliable and stable resistance to whitefly. Lectins from *Allium sativum* (median lethal concentration (LC_{50}) = 8.5 μg/ml) and *Pinnelia ternata* (41% mortality at 400 μg/ml) and double-stranded RNA (dsRNA) targeting V-ATPase mRNA (LC_{50} 3.08 μg/ml) provide some resistance against whiteflies in transgenic crops[8–10]. However, these technologies have not been tested in field trials, possibly because incomplete protection of crops is expected owing to the high LC_{50} of the insecticidal molecule or because of the potential for physiological adaptations in target pests or toxicity to nontarget insects, yield penalties, complex regulatory requirements or predicted allergenicity or toxicity of the protein or transgenic crops to mammals[11].

Ferns and mosses are rarely infested by phytophagous insects in the wild—insect infestation on ferns is 30-fold lower than on flowering plants[12,13]. The mechanisms by which ferns and mosses resist insects are not well understood. Ferns produce phytoecdysones that severely impair insect development and cause molting abnormalities[14]. Extracts of ferns (*Christella parasitica*, *Pteridium aquilinum* and *Hemionitis arifolia*) sprayed on peanut (*Arachis hypogaea*) crops have been found effective against *Helicoverpa armigera* and *Spodoptera litura*, but neither the active molecule(s) in ferns nor the active extract were tested against sap-sucking pests[15].

[1]Genetics and Plant Molecular Biology, Council of Scientific and Industrial Research (CSIR)–National Botanical Research Institute, Lucknow, India. [2]Academy of Scientific and Innovative Research, New Delhi, India. [3]National Agri-Food Biotechnology Institute, Mohali, Punjab, India. [4]Plant Diversity, Systematics and Herbarium Division, CSIR–National Botanical Research Institute, Lucknow, India. [5]Toxicology Division, CSIR–Central Drug Research Institute, Lucknow, India. [6]Pharmacokinetics and Metabolism Division, CSIR–Central Drug Research Institute, Lucknow, India. [7]Ladybird Beetle Research Laboratory, Department of Zoology, University of Lucknow, Lucknow, India. [8]Department of Botany, Banaras Hindu University, Varanasi, India. [9]Present addresses: Department of Botany, Panjab University, Chandigarh, India (S.K.U.), Regional Station, Indian Agricultural Research Institute, Pune, India (K.C.), University Institute of Engineering and Technology, Panjab University, Chandigarh, India (R.T.). [10]These authors contributed equally to this work. Correspondence should be addressed to P.K.S. (pksingh@nbri.res.in).

Received 12 September 2015; accepted 10 August 2016; published online 5 September 2016; doi:10.1038/nbt.3665

We screened diverse ferns to identify proteins that are insecticidal against the whitefly species *B. tabaci*. We identified a candidate protein that was toxic to whitefly from the edible fern *T. macrodonta* and named it Tma12. We then cloned the encoding gene (**Supplementary Note 1**) and constitutively expressed it in transgenic cotton plants (*Gossypium hirsutum* 'Coker 312') using the pMDC32 plant expression vector. One transgenic line (line 14) had a high level of resistance to whitefly (more than 99% control of insect population) and was protected from infection by a cotton leaf curl virus (CLCuKoV-Bu). Subchronic toxicity studies provided preliminary evidence that Tma12 is nontoxic in rats under the tested conditions. This insecticidal protein could provide an opportunity to control whitefly in GM crops.

RESULTS

Screening ferns for insecticidal proteins

Thirty-eight species of fern were screened for proteins with insect toxicity (Online Methods and **Supplementary Fig. 1**). Total soluble proteins (TSP) were prepared from fronds and rhizomes of different fern species and tested for insecticidal activity against whitefly using an *in vitro* feeding assay (**Supplementary Table 1**). Twenty-five TSPs that caused more than 50% mortality in whiteflies in three independent experiments were examined further after thermal inactivation and proteinase K treatment. Four TSPs (from *Drynaria quercifolia* and *T. macrodonta*) that caused >90% mortality and showed >5-fold loss in activity after proteinase K treatment were selected, as these features are a good indicator that the insecticidal properties are from the protein. The activity of the *D. quercifolia* TSP was almost abolished by thermal and proteinase K treatments, whereas activity of the *T. macrodonta* TSP was lost only upon proteinase K treatment. The insecticidal activity in the former was unstable during the purification steps. The *T. macrodonta* TSP was therefore selected for further study. We purified the candidate protein, Tma12, to near homogeneity from the fronds using activity-guided purification (**Fig. 1a–f**). Typically, 7–12 mg of pure Tma12 was recovered from 100 g fronds. Two-dimensional gel electrophoresis and silver staining established the absence of isoforms or other protein contaminants (**Fig. 1g**). The molecular size of Tma12 was ~43 kDa, as determined by analytical size-exclusion chromatography, and its mass was 21.686 kDa, as determined by matrix-assisted laser desorption ionization–tandem time-of-flight (MALDI-TOF/TOF) (**Fig. 1h,i**). These results indicate that Tma12 forms a homodimer in the source

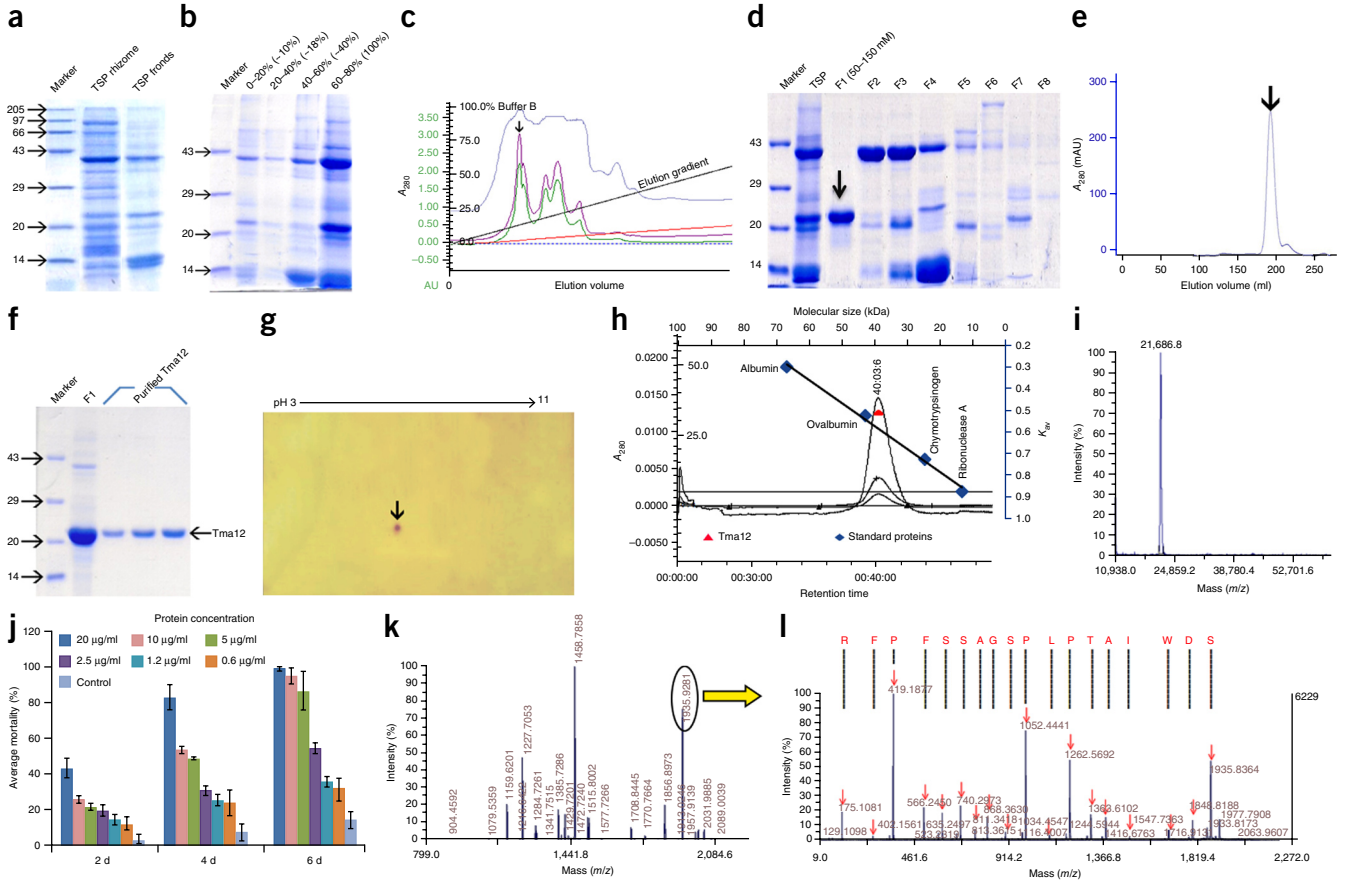

Figure 1 Purification and characterization of Tma12 from *T. macrodonta*. (**a**) TSP from rhizome and fronds of *T. macrodonta*. (**b**) Differential precipitation of TSP from fronds in steps of 20% up to 80% saturation of ammonium sulphate. The 60–80% fraction possessed maximum protein and highest insecticidal activity. (**c**) Separation of 60–80% fraction from **b** on Q Sepharose High Performance column; only the first peak (eluted at 50–150 mM NaCl, indicated by arrow) showed insecticidal activity. (**d**) SDS–PAGE analysis of the eluted fractions; the peak showing insecticidal activity had a single protein (~21 kDa). (**e**) Purification of insecticidal protein from **c** on a Superdex 200 'pre-grade' size-exclusion column. Tma12 eluted as single peak. (**f**) SDS–PAGE analysis showing homogeneously purified Tma12. (**g**) Two-dimensional gel electrophoresis and silver staining showing homogeneity of purified Tma12 and the absence of its isoforms. Experimental isoelectric point (pI) of the protein was 6.2. (**h**) Molecular size analysis of native Tma12 on an analytical Superdex 200 column. (**i**) Resolution of the protein on MALDI-TOF/TOF at molecular mass 21.6 kDa. (**j**) *In vitro* bioassay of purified Tma12 on newly emerged whiteflies; data represent mean ± s.d. Experiments were conducted on protein purified in 4 batches, each time in triplicate with 3 technical replicates (n = 30–50). (**k**) Peptide mass fingerprint of the insecticidal protein. (**l**) *De novo* sequencing of 1,935.9281-Da peptide after 4-sulphophenyl isothiocyanate modification of the N terminus showing deduced amino acid sequence.

plant, and that this homodimer is converted into a monomer by chemical treatment in mass spectrometry (MS) analyses. In the whitefly toxicity assay, Tma12 had an LC_{50} of 1.49 µg/ml (fiducial limits = 0.866–2.239; slope = 1.759; 95% confidence interval (CI) for slope 1.487–2.032; χ^2 for heterogeneity 12.952 calculated; χ^2 significance value 0.012). Tma12 caused >90% mortality at 5.0 µg/ml (**Fig. 1j**) and retained its insecticidal property at 95 °C for 20 min but not longer (30 min or more at 80–95 °C; data not shown). Tma12 peptides identified through MS (**Fig. 1k**) and tandem MS (MS/MS) data did not have significant ($P > 0.05$) similarity with other known proteins in MSDB, Swiss-Prot and NCBInr databases. The amino acid sequence of the Tma12 N terminus, determined by Edman degradation sequencing, was HGSMEDPISRVY. Further MS-based *de novo* sequencing of two peptides (generated after trypsin digestion) revealed amino acid sequences of SDWIATPLPSGASSFPFR and ELIPDGQLCSAGR with a high level of confidence (**Fig. 1l** and **Supplementary Figs. 2–4**).

Characterization of the *tma12* gene

The gene encoding Tma12 (809 bp, GC content = 60%) cloned by rapid amplification of cDNA ends (RACE) (**Supplementary Note 1**) has a 651-bp open reading frame (ORF; GenBank JQ438776) with a 33-bp 5′ untranslated region (UTR) and a 125-bp 3′ UTR. Translation of the ORF yielded a putative 216-aa pre-protein with a 24-aa N-terminal signal peptide (**Supplementary Fig. 4a**). The peptide fragments identified by *de novo* sequencing were present in the final amino acid sequence obtained from the cDNA (**Supplementary Fig. 4b**). This shows that the cDNA encodes the target insecticidal protein.

Chitin-binding and chitinase activity of Tma12

Bioinformatic analysis of Tma12 predicted the presence of a conserved chitin-binding domain of the Chitin_bind_3 superfamily (**Supplementary Fig. 4c**). We found that Tma12 binds chitin beads (**Supplementary Fig. 5a**), consistent with the presence of a chitin-binding domain. Chitinolytic activity was retained up to 50 °C, declined at 55 °C and was abolished at 60 °C (**Supplementary Fig. 5b**). Enzymatic activity with three substrates (colloidal crab chitin; 4-methylumbelliferyl β-D-*N*,*N*′-diacetylchitobioside hydrate; and 4-methylumbelliferyl β-D-*N*, *N*′,*N*″-triacetylchitotriose) established that the chitinolytic and exo- and endochitinase properties of Tma12 followed first-order Michaelis–Menten kinetics (**Supplementary Fig. 5c–e**). The K_m value of Tma12 for colloidal chitin was substantially higher than other reported chitinases (**Supplementary Tables 2** and **3**). We included three commercially available chitinases in our study as controls and examined them for insecticidal potential toward whitefly, but none were toxic to the insect (**Supplementary Table 4**). It is not clear without further investigation whether the chitin-binding and/or chitinase activity of Tma12 is associated with its insecticidal activity. Tma12 expressed in *Escherichia coli* formed inclusion bodies; it was refolded and tested for insecticidal as well as chitinase activity. Conditions for refolding were not optimized, and although chitinase activity was retained, insecticidal activity was reduced ten-fold compared to native protein purified from the fern (data not shown).

Transgenic cotton expressing *tma12*

An *Agrobacterium tumefaciens*–based plant expression cassette (pNBRI719) was used in which *tma12* was expressed constitutively from a CaMVE35S promoter (**Fig. 2a**) in the vector pMDC32. Transformation of cotton (*G. hirsutum* 'Coker 312') with pNBRI719 resulted in 87 somatic cotton embryos. Of these, 28 embryos developed into plantlets, and 16 of the plantlets reached maturity (**Supplementary Fig. 6**). PCR and qRT-PCR analyses confirmed the presence of *tma12*

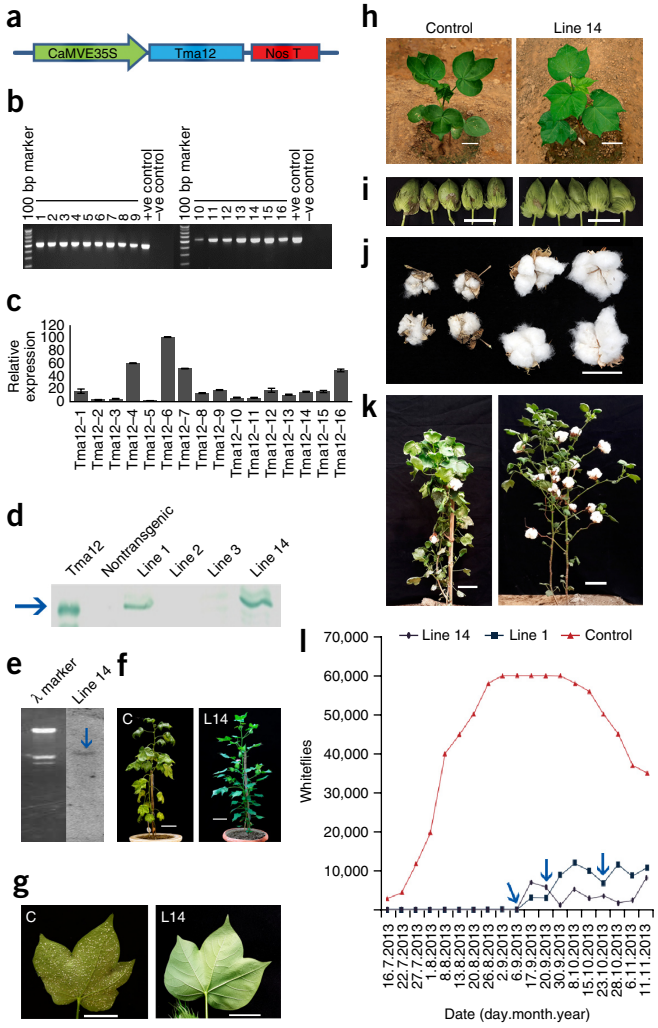

Figure 2 Transgenic cotton plants expressing Tma12 resist whitefly. (**a**) Gene expression cassette (pNBRI719) used for *Agrobacterium*-mediated transformation. (**b**) PCR analysis of putative transgenic cotton lines with *tma12*-specific primers showing presence of the transgene in all lines. +ve, positive control (plant expression cassette); –ve, negative control (genomic DNA of vector-transformed transgenic plant). (**c**) qRT-PCR analysis of putative transgenic lines showing variable expression of the transgene. Experiments were repeated 3 times, each time with 3 technical replicates per line; error bars, mean ± s.d. ($n = 3$). (**d**) Immunoblot analysis of Tma12 expression in T1 transgenic lines 1, 2, 3 and 14. (**e**) Southern blot analysis of line 14. Presence of a band of ~7 kb (arrow) established single copy insertion. (**f**) Nontransgenic control (C) and line 14 (L14) after whitefly challenge in greenhouse. (**g**) Whitefly infestation on nontransgenic control leaf and its absence in L14 leaf. (**h**) Leaf morphology in nontransgenic control and transgenic L14 plant. (**i**,**j**) Size comparisons of premature (**i**) and mature (**j**) bolls of control (left) and transgenic (right) plants. (**k**) Boll setting in transgenic and nontransgenic plants. (**l**) Whitefly population on transgenic lines 1 and 14 and nontransgenic control ($n = 24$ plants). Total numbers of whiteflies counted on indicated dates. Adult whiteflies were introduced to transgenic plants at three time points (arrows). Whitefly numbers >2,500/plant were uncountable and considered to be 2,500 for plotting the data. Scale bars, 5 cm (**g**–**j**) and 10 cm (**f** and **k**).

and its transcripts in all 16 putative transgenic lines (**Fig. 2b**,**c**). Of the 16 mature lines, 4 (lines 1, 2, 3 and 14) advanced to T1 generation. Immunoblot analyses of the T1 lines confirmed expression of Tma12 in lines 1 and 14 (**Fig. 2d**). The level of expression was ~0.01% of TSP. Lines 1 and 14 were advanced to the T2 generation, analyzed by

Table 1 Yield analysis of transgenic cotton line 14 in a contained field trial

	Bolls/plant (± s.d.)	Boll weight/plant (g ± s.d.)	Seed weight/plant (g ± s.d.)	Lint weight/plant (g ± s.d.)
Transgenic	16.62 ± 4.62	31.13 ± 7.94	21.85 ± 5.93	9.28 ± 2.10
Pesticide-protected nontransgenic	15.48 ± 3.87	29.87 ± 5.88	21.13 ± 5.19	8.74 ± 1.38
Nontransgenic	9.29 ± 2.53	14.15 ± 3.31	9.48 ± 2.41	4.62 ± 1.25
F/*P*[a]	22.956/0.0001	51.974/0.0001	44.740/0.0001	51.917/0.0001

[a]*F* and *P* values from one-way ANOVA.

Event 14 (T3 generation) and nontransgenic cotton plants ($n = 21$, 7 plants × 3 rows) were grown (January–June 2014) in independent plots of size 4 × 7 m. In another plot of the same size, nontransgenic cotton plants were grown and protected with three sprays of Confidor (imidacloprid, 0.04% v/v in water). Bolls were harvested after 180 d. Data were analyzed by one-way ANOVA followed by pairwise comparisons using Tukey's *post hoc* test.

PCR for homozygosity and by immunoblotting to confirm transgene expression (data not shown). Line 14 was selected for further study on the basis of higher insect resistance, superior yield performance (**Supplementary Table 5a**) and better fiber quality (**Supplementary Table 5b**). Southern blot analysis of the T3 generation showed that line 14 carried a single copy of *tma12* (**Fig. 2e**). Expression of Tma12 in various parts of plants was examined in the T5 generation. The *tma12* transcript was present in large amounts in root and less so in the stem, fruit wall and leaf. Low *tma12* transcription was also detected in petal, anther and square (**Supplementary Fig. 7a**). However, Tma12 protein was detectable only in the leaf (0.01–0.02% of TSP), stem (0.001%) and seeds (barely detectable, 0.0002%) and not in other plant parts, including the root (**Supplementary Fig. 7b**), suggesting tissue-specific post-transcriptional differences in expression.

Whitefly resistance of cotton lines expressing Tma12

Whitefly resistance and yield performance of transgenic cotton lines were evaluated simultaneously. Nine lines with high *tma12* expression in the T0 generation (including lines 1 and 14) showed substantial (>90% population reduction) protection against whitefly challenges in a greenhouse study (**Fig. 2c** and **Supplementary Fig. 6**). Insect feeding on nontransgenic plants caused chlorosis and drooping of leaves in the greenhouse, whereas the selected event (line 14) showed normal growth (**Fig. 2f**). Line 14 was almost completely insect resistant up to the T4 generation in contained field trials. Whitefly infestation on control plants caused thickening, darkening and curling of leaves, sooty mold formation and stunted growth leading to a reduction in the size and number of bolls (**Fig. 2g–k**). An average yield loss of 50–80% was recorded across all field trials and all crop seasons (**Supplementary Table 5a**). Conversely, transgenic cotton plants appeared healthy and did not show symptoms of whitefly infestation. Whitefly count on high-expressing lines (1 and 14) never exceeded 20 per plant (**Fig. 2l**), whereas the infestation on nontransgenic plants increased rapidly and became uncountable (>2,500 insects/plant) in 40 d. Even after three rounds of inoculation, the whitefly population did not increase on transgenic plants. Leaves of transgenic plants grown in contained field tests had no visible symptoms of cotton leaf curl disease, unlike control plants. PCR analysis of control plants revealed the presence of rounds of treatment CLCuKoV-Bu and its associated alpha- and beta-satellites, whereas transgenic plants did not contain viral DNA (**Supplementary Fig. 8**). The yield of the transgenic plants was comparable to that of pesticide-protected nontransgenic plants (**Table 1**). Therefore Tma12 offers similar benefit to cotton as that provided by three rounds of treatment with a chemical pesticide.

Whitefly life cycle is disrupted by Tma12

In clip-cage experiments, feeding of whiteflies on transgenic leaves (line 14) resulted in reduction in number of insects in the next generation. In contrast, the insect population increased rapidly on nontransgenic leaves (data not shown). We examined the developmental stages of whitefly after inoculation on cotton plants in mesh chambers. Microscopic examination of the nontransgenic leaf revealed whiteflies of various developmental stages (**Fig. 3a–d**); however, insects fed on transgenic leaves showed interference in the life cycle, as evidenced by substantially reduced numbers of eggs, abnormal egg-laying patterns, failure of most of the eggs to hatch, emergence of fewer nymphs and

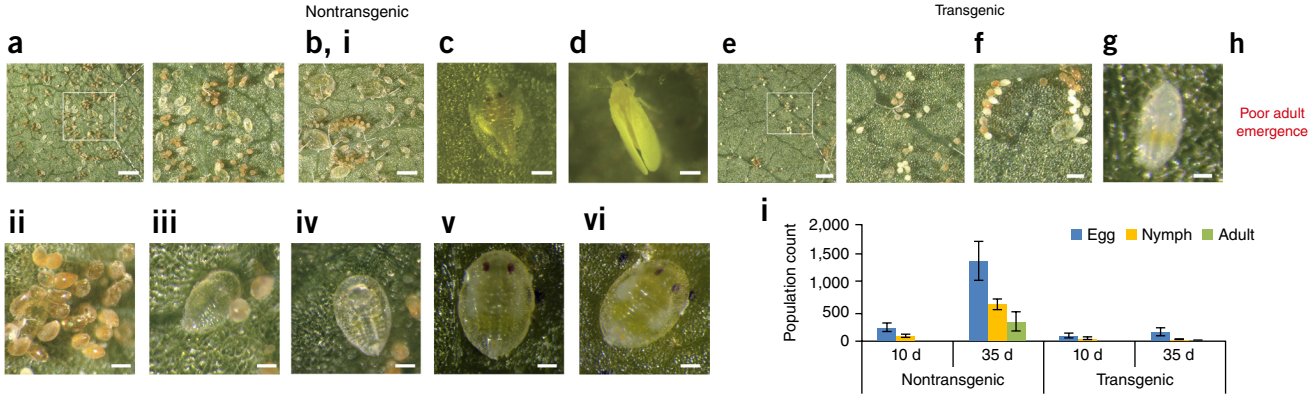

Figure 3 Tma12 expression in cotton interferes with the whitefly life cycle. (**a**–**h**) Developmental stages of *B. tabaci* on nontransgenic (**a**–**d**) and transgenic (**e**–**h**) cotton plants grown in pots and challenged with newly emerged whiteflies, examined under stereomicroscope 7 d after challenge. (**a**) Eggs on leaves (left), enlarged (right) to show mature and fertile eggs (golden brown) advancing to nymphal stages. (**b**, **i**–**vi**) Eggs hatching into nymphs. Shown are mature eggs in cluster (**i**) and magnified views of mature eggs (**ii**); newly hatched nymphs (**iii**) and advanced nymphal stages (**iv**–**vi**). (**c**) Fourth instar nymphs with visible wings. (**d**) Adult whitefly. (**e**–**g**) Abnormalities in the life cycle of whitefly fed on transgenic cotton plants. (**e**) Reduced egg-laying on transgenic leaf (left), enlarged (right) to show higher proportion of sterile (white) eggs. (**f**) Poor egg-maturation efficiency, indicated by low proportion of golden-brown eggs. (**g**) Emergence of nymphs. (**h**) Poor adult emergence. (**i**) Average number of eggs and nymphs per leaf and adults per plant at 10 and 35 d after release of 40 adult whiteflies in mesh chamber. Error bars, mean ± s.d. ($n = 6$ plants). Scale bars, 1 mm (**a**,**e**), 500 µm (**b**, **i**), 100 µm (**b**, **iii**–**v** and **c**), 200 µm (**d**) or 20 µm (**b**, **ii**).

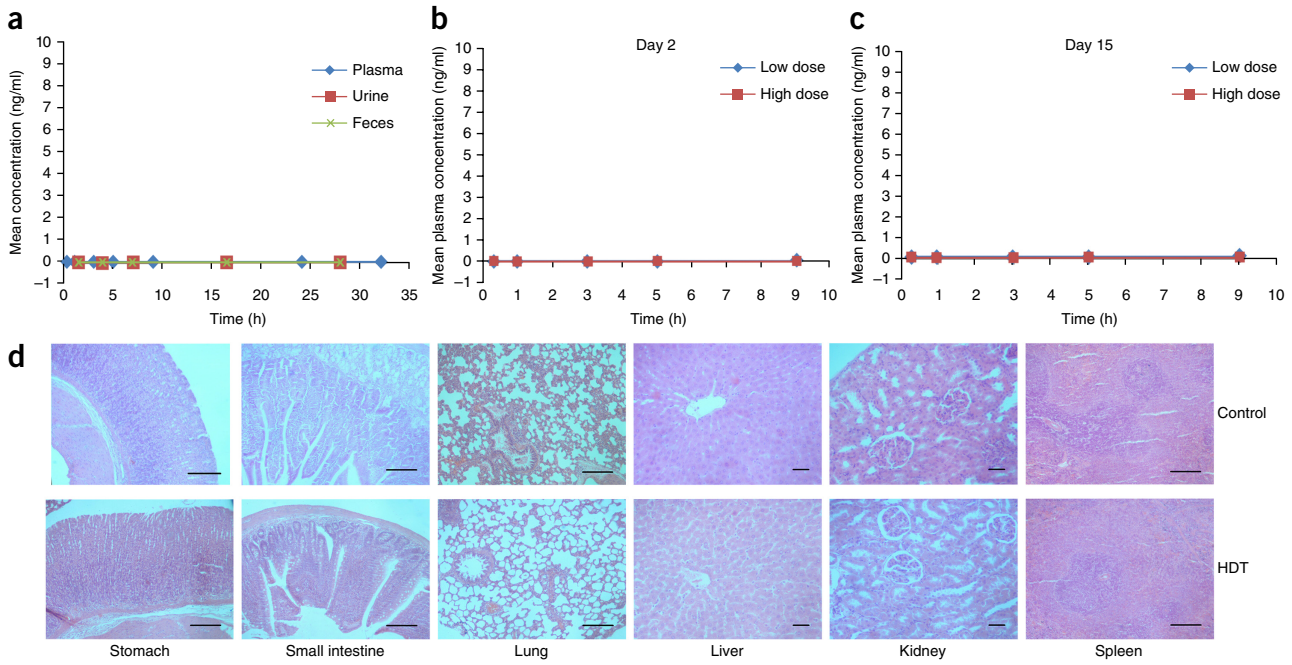

Figure 4 Toxicological study of Tma12 in rats. (**a**) Concentration of Tma12 in plasma, urine and feces at predetermined time points over 32 h in rats fed a diet supplemented with 10.35 mg/kg Tma12 per day (high-dose treated). (**b**,**c**) Mean concentrations of Tma12 in plasma on day 2 (**b**) and day 15 (**c**) in low-dose-treated (3.45 mg/kg/day) rats and high-dose-treated rats. (**d**) Hematoxylin and eosin staining of tissues of control (top) and high-dose-treated (HDT) rats (bottom) on day 15. Images were captured at 40× (stomach, small intestine, lung and spleen) and 100× (liver and kidney) magnification. Scale bars; 10 µm (liver and kidney) or 20 µm (stomach, small intestine, lung and spleen). Experiments were conducted with $n = 5$ males and 5 females per group (low dose, high dose and control).

very few emerged adults (**Fig. 3e**–**h**). Numbers of eggs, nymphs and adults increased rapidly on nontransgenic plants but remained low (<1%) on the transgenic plants (**Fig. 3i**).

Tma12 cotton is resistant to whitefly only

Unlike whitefly, the growth and population size of nontarget chewing pests *H. armigera* and *S. litura* (**Supplementary Table 6a** and **Supplementary Fig. 9**) and sucking pests *Aphis gossypii* (cotton aphid) and *Phenacoccus solenopsis* (cotton mealybug) (**Supplementary Fig. 10**) were unaffected on transgenic plants (line 14). *Menochilus sexmaculatus* (ladybird beetle), a beneficial predatory insect, was fed on cotton aphids reared on line 14. No adverse effect was recorded, and the insect reproduced normally (**Supplementary Table 6b**).

Safety tests in rodents

Systemic absorption and potential toxicity of Tma12 were assessed by subchronic exposure in Sprague–Dawley rats at two different oral doses (3.45 and 10.35 mg per kg of body weight (mg/kg) per day) for 14 d. Tma12 was not detected in plasma, urine or feces after oral administration of a single high dose (10.35 mg/kg) (**Fig. 4a**). The protein was undetectable in plasma at all time points throughout the 15 d of observations (**Fig. 4b**,**c**), indicating that it had negligent systemic distribution. There were no significant differences from reported values in the hematological, serum biochemical or urine profiles of untreated or treated rats (**Supplementary Table 7a**–**g**). On the time scale of these experiments, Tma12 had no adverse effects on kidney, liver or general metabolic function in male or female rats (**Supplementary Table 7a**–**g**). Histological examination of liver, lung, small intestine, stomach layer and kidney sections showed normal architecture with no signs of adverse effects. Most notably, administration of Tma12 did not cause splenic abnormality, indicating that Tma12 is unlikely to be immunotoxic (**Fig. 4d**).

DISCUSSION

Proteins that can control whiteflies effectively in transgenic crops would be invaluable, provided that they do not affect host plant biology and are safe to humans and other nontarget organisms. In search of such proteins, we screened 38 ferns for protein-based insecticidal activity and identified *T. macrodonta* as the most promising source. TSP from rhizomes as well as fronds had insecticidal activity. The target protein was purified from fern fronds. Activity-guided purification yielded the insecticidal protein Tma12. We note that Tma12 was isolated from an edible plant, which could make it a more promising candidate for the development of whitefly-resistant GM crops than lectins[8,9]. Although the use of RNA interference to control whitefly has been suggested[10], the multigenic nature of target genes and nonspecific recognition make it a poor choice for GM crop development.

The 16 transgenic cotton lines developed in this study had variable *tma12* expression. One line (event 14) showed excellent control of whitefly and superior agronomic parameters throughout four generations. Its lack of effect on other cotton pests, including *H. armigera*, *S. litura*, *A. gossypii* and *P. solenopsis* indicates that Tma12 is specific for whitefly. The transgenic cotton did not affect ladybird beetle, an important predator of whitefly and a beneficial insect. Transgenic plants grew normally (**Supplementary Fig. 11**), and yield was on par with that of pesticide-protected control plants (**Supplementary Fig. 12**). Seed germination, photosynthesis rate, plant biomass and flowering time were comparable between the transgenic and control plants. This indicates that Tma12 expression does not impose a detectable penalty on growth or yield. Transgenic cotton plants grown in the field also had no visible symptoms of whitefly-borne cotton leaf curl disease. The absence of the CLCuV genome and its associated satellites is most likely due to the control of the whitefly vector.

Although transgenic cotton did not cause mortality in adult whiteflies, it arrested their advancement to the next generation, likely because Tma12 accumulation in phloem was ~100 ng/ml, ~19-fold lower than the LC_{50} value. Feeding on transgenic cotton interfered with insect egg laying and early developmental stages of nymphs (**Fig. 3**). The number of eggs, egg-laying pattern and nymphal development were severely affected. Feeding whitefly with sublethal doses of Tma12 *in vitro* showed similar effects (data not shown). The mechanism of action of Tma12 remains to be elucidated. However, the chitinolytic activity of Tma12 may not be responsible for its insecticidal activity, as the former is thermosensitive but the latter is not.

Safety of a candidate protein is essential for its deployment in GM crops for commercial cultivation. Although cotton is a fiber crop, its seed meal cake is used as feed for cattle. Our toxicity studies in rats at two doses (3.45 and 10.35 mg/kg/day) of the purified protein (which corresponds to 4.5 and 13.5 kg, respectively, of cottonseed meal cake fed daily to 5.5-month-old calves of 60 kg, on an equivalence basis) revealed no adverse effects over 14 d. Hematological and biochemical parameters of control and treated rats remained within the reference range[16]. The sensitivity of Tma12 to pepsin (**Supplementary Fig. 13**) and the absence of known allergenic motifs suggest that Tma12 might be a promising candidate for the development of GM crops. Nevertheless, detailed safety studies are required to meet GM regulations. Recently, a spider neurotoxin (Hv1a) was used to kill aphid species (LC_{50} 80 µg/ml), using the coat protein of luteovirus as the delivery vehicle[17]. Despite its high LC_{50} value, the fusion protein provided resistance against aphids in transgenic *Arabidopsis*. However, the safety of such a technology for unregulated field deployment is uncertain, given the potential toxicity of a neurotoxin. In contrast, the source fern used in our study is a known edible plant, consumed as a vegetable by humans and used in traditional remedies for various human diseases[18–20]. Therefore *tma12* is a promising candidate gene that could be pyramided with *Bt* toxins to develop GM crops with resistance to whitefly and other herbivorous pests. Other crop plants susceptible to whitefly infestation could now be tested to evaluate whether Tma12 has potential as a transgenic technology on a broader scale.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. GenPept: amino acid sequence has been deposited under accession code AFR32946. GenBank: mRNA sequence has been deposited under accession code JQ438776.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

The authors acknowledge H.K. Yadav, R. Kamal and S. Sharma for help in statistical analysis; M. Lal, Rajesh K. Srivastava, S.H.M. Abidi, A. Bano (deceased) and Rakesh K. Srivastava for technical assistance; and A.C. Little for photography. The authors thank the Council of Scientific and Industrial Research, India, for funding under the Supra Institutional Project (SIP05), the Network Projects (NWP03 and PlaGen BSC0107) and the Council of Scientific and Industrial Research–National Botanical Research Institute for providing research facilities. A.K.S., S.K.U., R.S., M.M., S. Saurabh, H.S., N.T., P.P., A.L.H. and V.R. are supported by CSIR. P.R. and S.K.Y. are supported by the Indian Council of Medical Research. S. Srivastava is supported by the University Grant Commission. R.T. is supported by a Department of Science and Technology (DST) JC Bose National Fellowship.

AUTHOR CONTRIBUTIONS

A.K.S.: development of cotton transformation protocols, transgenic cotton lines and acquisition of data. S.K.U.: purification of Tma12 and characterization, cloning of *tma12*. M.M.: MS analysis of Tma12, study on whitefly life cycle on transgenic cotton in clip cage and manuscript preparation. S. Saurabh, R.S., H.S., P.R.: bioprospection of ferns for insecticidal activity, establishment of screening protocol, identification of the potential ferns, execution of experiments and manuscript preparation. N.T.: clip cage assay, analysis of transgenic cotton lines and manuscript preparation. S. Srivastava: analysis of transgenic cotton lines and manuscript preparation. A.L.H., P.P.: study of whitefly life cycle on transgenic cotton in the field. V.R., M.K.S.: evaluation of transgenic cotton in field trials. S.K.Y.: chitin-binding assay, analysis of impact of Tma12 on whitefly at sublethal doses. J.K.: virus studies. K.C.: designing and execution of whitefly bioassays and statistical analysis. P.C.V.: designing and supervision of whitefly life cycle study on transgenic cotton and transgenic cotton development. A.P.S.: collection, identification, mass multiplication and in-house maintenance of ferns. K.N.N.: review of experimental design and writing of manuscript. S.B., M.W., S. Singh, S. Sharma: biosafety experiments. O.: designing and execution of experiments on beneficial insect ladybird beetle. R.S.U.: supervised A.K.S. in cotton tissue culture and field evaluation of transgenic cotton. S.A.R.: critical review of manuscript. R.T.: critical review of results, intellectual input and manuscript editing. P.K.S.: study concept and experimental design, data analysis, study supervision and manuscript writing and editing.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Sanahuja, G., Banakar, R., Twyman, R.M., Capell, T. & Christou, P. *Bacillus thuringiensis*: a century of research, development and commercial applications. *Plant Biotechnol. J.* **9**, 283–300 (2011).
2. Chougule, N.P. & Bonning, B.C. Toxins for transgenic resistance to hemipteran pests. *Toxins (Basel)* **4**, 405–429 (2012).
3. Dubey, A.K. & Ko, C.C. Whitefly (*Aleyrodidae*) host plants list from India. *Orient. Insects* **42**, 49–102 (2008).
4. Legg, J. in *Whitefly and Whitefly-borne Viruses in the Tropics: Building a Knowledge Base for Global Action* (eds. Anderson, P.K. & Morales, F.J.) 15–23 (CIAT Publications, 2005).
5. Navas-Castillo, J., Fiallo-Olivé, E. & Sánchez-Campos, S. Emerging virus diseases transmitted by whiteflies. *Annu. Rev. Phytopathol.* **49**, 219–248 (2011).
6. Horowitz, A.R., Antignus, Y. & Gerling, D. in *The Whitefly, Bemisia tabaci (Homoptera: Aleyrodidae) Interaction with Geminivirus-Infected Host Plants.* (ed. Thompson, W.M.O.) 293–322 (Springer, 2011).
7. Bergé, J.B. & Ricroch, A.E. Emergence of minor pests becoming major pests in GE cotton in China: what are the reasons? What are the alternatives practices to this change of status? *GM Crops* **1**, 214–219 (2010).
8. Das, S. *et al.* A mannose binding lectin from leaves of *Allium sativum* effective against whitefly, and process for its preparation. Indian patent 228783 (2009).
9. Jin, S., Zhang, X. & Daniell, H. *Pinellia ternata* agglutinin expression in chloroplasts confers broad spectrum resistance against aphid, whitefly, Lepidopteran insects, bacterial and viral pathogens. *Plant Biotechnol. J.* **10**, 313–327 (2012).
10. Thakur, N. *et al.* Enhanced whitefly resistance in transgenic tobacco plants expressing double stranded RNA of *V-ATPase A* gene. *PLoS One* **9**, e87235 (2014).
11. Doorn, A.V. & Vos, M.D. Resistance to sap-sucking insects in modern-day agriculture. *Front. Plant Sci.* **4**, 222 (2013).
12. Hendrix, S.D. An evolutionary and ecological perspective of the insect fauna of ferns. *Am. Nat.* **115**, 171–196 (1980).
13. Markham, K., Chalk, T. & Stewart, C.N. Jr. Evaluation of fern and moss protein based defenses against phytophagous insects. *Int. J. Plant Sci.* **167**, 111–117 (2006).
14. Balasubramanian, R., Selvaraj, P. & Sahayaraj, K. Partial purification and characterization of phytoecdysone from *Christella parasitica* (L.) and screening its pesticidal properties on lepidopteran pests. *J. Biopesticides* **1**, 201–205 (2008).
15. Sahayaraj, K. & Selvaraj, P. Field evaluation of three fern extracts on *Helicoverpa armigera* (Hubner) and *Spodoptera litura* (Fab.) incidence and infestation and groundnut production. *Legume Res.* **36**, 84–86 (2013).
16. Petterino, C. & Argentino-Storino, A. Clinical chemistry and haematology historical data in control Sprague-Dawley rats from pre-clinical toxicity studies. *Exp. Toxicol. Pathol.* **57**, 213–219 (2006).
17. Bonning, B.C. *et al.* Toxin delivery by the coat protein of an aphid-vectored plant virus provides plant resistance to aphids. *Nat. Biotechnol.* **32**, 102–205 (2013).
18. Uprety, Y., Poudel, R.C., Asselin, H. & Boon, E. Plant biodiversity and ethnobotany inside the projected impact area of the Upper Seti Hydropower Project, Western Nepal. *Environ. Dev. Sustain.* **13**, 463–492 (2011).
19. Debendra, S.D. Indigenous vegetables of Nepal for biodiversity and food security. *Int. J. Biodivers. Conserv.* **5**, 98–108 (2013).
20. Nwosu, M.O. Ethnobotanical studies on some pteridophytes of southern Nigeria. *Econ. Bot.* **56**, 255–259 (2002).

ONLINE METHODS

Screening of ferns for insecticidal proteins. Thirty-eight species of ferns were collected from the CSIR-NBRI fernery and Pachmarhi Biosphere Reserve (Madhya Pradesh, India). Total soluble protein (TSP) was prepared from rhizomes (35 species) and fronds of each species and tested for insecticidal activity against whitefly (*B. tabaci)* adults. 1 g frond/rhizome was homogenized in 4 ml protein extraction buffer (20 mM HEPES, pH 8.0; 0.5 mM DTT; 1 mM EDTA; 10% glycerol; 1 mM phenylmethylsulfonyl fluoride and 1 mM benzamidine), spun (12,000 × *g*, 15 min, 4 °C) and dialyzed against 20 mM Tris-HCl (pH 8.0) at 4 °C. The protein was measured by Bradford reagent, and insect bioassay was performed (200 µg fern protein/ml diet), following the protocol of Upadhyay *et al.*[21]. The samples that showed >50% mortality to whiteflies in three independent experiments were selected. Proteinaceous nature of the insecticidal principle was examined by feeding the samples after (i) heat inactivation at 95 °C for 20 min and (ii) proteinase K digestion at 55 °C (1:50 enzyme/ protein) for 30 min.

Purification of Tma12 protein from *T. macrodonta* fronds. The activity-guided screening showed the presence of insecticidal protein in the fronds of *T. macrodonta.* TSP from its fronds was then fractionated to purify the insecticidal protein. The TSP was fractionated by raising saturation with ammonium sulfate to 80% at 4 °C in steps of 20%. The precipitated protein was recovered by centrifugation (12,000 × *g*, 10 min, 4 °C), solubilized in 20 mM Tris-HCl (pH 8.5), dialyzed and spun (38,000 × *g,* 30 min, 4 °C). The supernatant was examined for insecticidal activity. The fraction with the highest activity was filtered through a 0.22-µm filter and applied onto a Q Sepharose High Performance (HP) column (5 ml, GE Healthcare) pre-equilibrated with 20 mM Tris-HCl (pH 8.5). The column was washed until the OD_{280} reached below 0.002 and stabilized. The bound proteins were eluted with increasing linear gradient of NaCl (0 to 1 M) in 20 column volumes. The peak fractions were resolved on 12% denaturing SDS–PAGE. Each fraction was dialyzed and tested for insecticidal activity. Fractions possessing insecticidal activity were pooled, concentrated and resolved on size-exclusion column (Superdex 200, prep grade, 330 ml, GE Healthcare) in 20 mM Tris-HCl (pH 8.0) containing 150 mM NaCl. The desired protein eluted as single peak that was collected and analyzed on SDS–PAGE. The purified protein was also analyzed on a silver-stained two-dimensional gel electrophoresis gel. Molecular size of the protein was estimated on analytical Superdex 200 column (24 ml). Purified protein (0.1 mg in 0.5 ml) was resolved at the flow rate of 0.5 ml/min. Albumin (~66 kDa), ovalbumin (~43 kDa), chymotrypsinogen (~25 kDa) and RNase A (13.7 kDa) were used as protein standards.

Insect bioassay. Newly emerged whitefly adults (30–50) were challenged with Tma12 at six concentrations (0.6, 1.2, 2.5, 5, 10 and 20 µg/ml). The protein was mixed in a sucrose- and yeast extract–based diet and fed to the flies across parafilm membrane for 6 d, as described in Upadhyay *et al.*[21]. The experiments were conducted on protein purified from four batches, sampled across the seasons, each time in triplicate with three technical replicates. The percentage mortality was calculated for each experiment and pooled for average percentage mortality calculation. The average percentage mortality of all the four experiments was used in the calculation of LC_{50} by Probit analysis on SPSS software version 22.0. The analysis builds a model for mortality and gives LC_{50} value with defined confidence limits (fiducial limits). It also analyzes χ^2 for the heterogeneity observed in actual mortality at increasing concentration of the toxin across the four experiments.

Mass spectrometric analysis. The mass of intact protein was determined on MALDI–TOF/TOF mass spectrometer (4800, ABI, USA) in linear mode, using CHCA matrix. For MS/MS analysis, the protein band was excised from a Coomassie Brilliant Blue (R-250)–stained gel and digested (in-gel) with trypsin. The peptides were recovered from the gel pieces in 50% acetonitrile (ACN) containing 1% trifluoroacetic acid (TFA), lyophilized and processed for analysis, as described previously[22].

N-terminal sequencing. The purified protein was outsourced to Proteomics International Pvt. Ltd. for sequencing of amino acids at the N terminus.

***De novo* sequencing.** Peptides of the insecticidal protein were sequenced on MALDI-TOF/TOF. The peptides extracted after tryptic digestion were captured on C_{18} (C_{18} ZipTips, Millipore), modified at their N terminus and analyzed on a mass spectrometer. In brief, the ZipTip was rinsed with ACN, and equilibrated in 0.1% TFA. For absorption onto C_{18}, the peptides were dissolved in 0.1% TFA and bound to the ZipTip by repeated pipetting. The ZipTip was washed with water, followed by freshly prepared 0.5 M *O*-methyl isourea hydrogen sulfate (MIS) in 0.25 M Na_2CO_3 (pH 11.7). The washed ZipTip was then incubated in the same reagent at 37 °C for 3 h to facilitate the conversion of the ε-amino group of terminal lysine residues to guanidyl groups. The tip was washed with 0.1% TFA to remove excess reagent. Freshly prepared solution of 4-sulfophenyl iso-thio-cyanate (SPITC) in 50 mM Tris-HCl (pH 8.2) was loaded onto the tip. The bound peptides were allowed to react at 50 °C for 2 h and add sulfate group on N terminus. Excess reagent was washed with 0.1% TFA, and the derivatized peptides were eluted with 80% ACN containing 0.5% TFA. The eluted peptides were lyophilized and dissolved in 50% ACN containing 0.1% TFA. Derivatized and underivatized peptides were spotted in different wells of the MALDI plate, and MS analysis was performed. Peptides whose derivatization was successful (identified by comparing with the mass of underivatized peptides) were selected for MS/MS analysis. The spectra were analyzed manually and sequences were derived by calculating the difference in the mass of adjacent peaks.

Cloning of the insecticidal gene (*tma12*) and sequence analyses. Degenerate primers were designed on the basis of peptide sequences. The complete gene sequence was obtained by 5′ and 3′ RACE using BD SMART RACE cDNA Amplification Kit. Details of cloning are given in **Supplementary Note 1**. Nucleotide and amino acid sequences were analyzed by blastn and blastp (NCBI). The presence of the conserved domain and of the signal peptide was determined through Conserved Domain Database search on NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and the SignalP 4.1 Server (http://www.cbs.dtu.dk/services/SignalP/), respectively.

Assay for chitin-binding and chitinolytic activity in Tma12. Chitin-binding assays were performed with chitin magnetic bead (NEB) following the manufacturer's protocol. Chitinolytic activity was measured with fluorometric chitinase assay kit (Sigma-Aldrich) following the manufacturer's instructions. K_m and V_{max} of the purified protein were calculated using Sigma plot (10.0).

Development of transgenic cotton plants. Transgenic cotton lines were developed through genetic transformation of *G. hirsutum* 'Coker 312'. Regeneration and transformation were performed by the protocols modified from Kumar *et al.*[23]. Hypocotyls were cut into ~1-cm pieces and used for *A. tumefaciens*–mediated transformation. *Agrobacterium* cells were grown to OD_{600} of 1.4, harvested (3,000 × *g*, 4 °C, 10 min), resuspended in 100 ml induction medium (NH_4Cl, 1 g/l, $MgSO_4$, 0.3 g/l, KCl, 0.15 g/l, $CaCl_2$, 10 mg/l, $FeSO_4.7H_2O$, 2.5 mg/l, KH_2PO_4, 0.272 g/l, MES, 0.39 g/l, glucose, 5 g/l, pH 6.0) containing 100 µM acetosyringone and incubated in a shaker (175 r.p.m., 27 °C, 4 h). Cells were again harvested and suspended in MSO + 100 µM acetosyringone and incubated (150 r.p.m., 26 °C, 2 h). The explants were dipped in induced *Agrobacterium* suspension and incubated for 20 min with gentle shaking before transfer to MSP for co-cultivation (21 °C, 48 h) in the dark. These were then washed with cefotaxime (250 mg/l) for 5 min, rinsed with sterile water, dried on blotting paper, transferred to MSP containing antibiotics (augmentin 250 mg/l, hygromycin 50 mg/l) and incubated (28 °C, 16 h light–8 h dark, 60 µmol/m^2/s light) for *in vitro* regeneration. De-differentiation of cells was initiated after two cycles of subculture (21 d each). The calli were transferred to MSA containing antibiotics for proliferation and selection (two cycles). Proliferated calli were transferred to MSA without antibiotics for embryogenesis (three cycles). The embryos were transferred to 0.5× MSP for root and shoot development. Embryos with root and shoot were transferred to 0.5× MSP in Planton boxes (Tarsons) for plant development. The plants (called T0 generation) were transferred to pots containing soil mix (vermiculite/soilrite/soil, 1:1:2) and grown in a greenhouse for hardening and development.

Molecular analysis of transgenic plants. Genomic DNA was isolated from young leaves of putative transgenic cotton plants (T0 generation)

using Qiagen's DNeasy Plant mini kit according to the manufacturer's instructions. 10 ng genomic DNA was subjected to PCR analysis, using end primers (**Supplementary Note 1**). Relative expression of the transgene (*tma12* mRNA) in T0 generation was evaluated by RT-qPCR with gene-specific forward (5′-TGGAGGAGTTGCCGTTCATC-3′) and reverse (5′-GTGGTGCCGGTGATTTGG-3′) primers by calculating the comparative threshold cycle ($\Delta\Delta Ct$). Ubiquitin amplified with primers forward 5′-AGCTC GGATACGATTGATAACG-3′ and reverse 5′- GAAGACGAAGAACAAG GGGAAG-3′ was used for the normalization of ΔCt values of the target gene. Relative abundance (fold-change value) was calculated from the expression ratio. Expression of Tma12 protein in transgenic lines in T1 and subsequent generations was estimated by immunoblot analysis. Rabbit anti-Tma12 polyclonal IgG was used as the primary antibody, and anti-rabbit antibody-HRP conjugate as the secondary antibody in 1:1,000 and 1:20,000 dilutions, respectively. The number of inserts of the transgene per genome was established by Southern blot analysis of the selected transgenic line (event 14). Genomic DNA was isolated from leaves of a 1-month-old plant after holding it in dark for 48 h. The DNA was digested with HindIII, transferred onto an N^+ membrane (GE Healthcare) and hybridized with ^{32}P-labeled probe.

Expression of *tma12* in different plant parts of transgenic line 14 was estimated in T5 generation by qRT-PCR. The protein expression was assessed by immunoblot analysis. For immunoblot analysis, 20 g root, leaf, stem and developing boll wall were collected. TSP from each tissue was differentially precipitated with ammonium sulfate in steps of 20%. Tma12 was detected in 40–60% ammonium sulfate fraction. Tma12 was enriched on an ion-exchange column and analyzed by immunoblot analysis. The purified Tma12 was electrophoresed on SDS–PAGE and transferred onto a PVDF membrane (Merck Millipore) using CAPS buffer (Sigma-Aldrich). The membrane was blocked with 1% BSA in TBS containing 0.05% Tween-20 (TBST) for 1 h. The membrane was incubated with anti-Tma12 developed in rabbit at 1:1,000 dilutions in TBST containing 0.5% BSA for 2 h. The membrane was washed three times with TBST and incubated with HRP-conjugated anti-Rabbit IgG (Sigma-Aldrich) at a 1:20,000 dilution for 2 h. The membrane was washed with TBST, and the blot was developed with HRP chromogenic substrate (TMB) (Invitrogen). The quantity was estimated by comparison with the intensity of 5, 10, 20, 30, 40 and 50 ng purified Tma12 run as standard.

Field evaluation of transgenic lines and selection of whitefly-resistant events. Twenty-eight putative transgenic lines were grown in a greenhouse (28° ± 2 °C, relative humidity 60 ± 5%). Sixteen of those reached maturity. These were challenged with whitefly (reared on cotton plants in controlled conditions), and their numbers were counted weekly. Seeds of the four lines that were advanced to T1 generation were analyzed for Mendelian segregation by germination in the presence of antibiotic. The selected plants were also tested for presence of the transgene by PCR. The field evaluation of transgenic cotton for agronomic performance was done in two seasons (January–June and July–December). The T1 and T3 generations were evaluated in January–June (off season), and the T2 and T4 generations were evaluated in July–December. The nursery of the plants was maintained in a greenhouse during January–February and May–June.

Experimental layout for field trials. Contained field trials to evaluate insect pest resistance in GM crops were set up as per the guidelines by India's Review Committee on Genetic Manipulation (http://igmoris.nic.in/files/cd_ibsc/files/transgenic.pdf) and described previously[24], with minor modifications approved by Institutional Biosafety Committee, CSIR-National Botanical Research Institute, India. The field (40 × 32 m) was divided into 12 experimental plots; eight of size 12 × 7 m and four of 7 × 4 m. The plots were covered with ~1-mm nylon mesh to allow the exchange of air. The plots were covered with a polythene sheet during heavy rain. Plants received ~80% of solar radiation. The layout of the contained field is shown in **Supplementary Note 1**.

T1 generation. The segregating population in T1 generation was grown in a greenhouse. The transgene-positive plants were transferred to contained field in March 2013. The plants (n = 21; 3 rows of 7 plants) of each transgenic line (lines 1, 2, 3 and 14) were grown in two plots of 12 × 7 m. The control plants (n = 48; 8 rows of 6 plants) were grown in the third plot. The transgenic and control plants were challenged with ~20 adult whiteflies/plant (collected from field) and examined for whitefly infestation–related signs. The number of whiteflies per plant and the yield (boll number, boll weight, seed and lint weight) were recorded periodically.

T2 generation. Two lines (1 and 14) were advanced to the T2 generation. Homozygous plants of each line (n = 24; 8 rows of 3 plants) were grown in a single plot of 12 × 7 m in July 2013. Control plants (n = 48; 8 rows of 6 plants) were grown in an adjacent plot. Whiteflies were released on transgenic and control plants as described above, and their numbers were tracked 15 d after inoculation. To examine the impact of an outburst of whiteflies on transgenic cotton, approximately 250 whiteflies were released on each transgenic plant three times, at the interval of three weeks (not released on control plants). The number of insects was counted weekly. The total number of insects per plant was considered to be uncountable when it exceeded 2,500.

T3 generation. Line 14 was selected as the primary 'event' for detailed studies. An experiment was designed to compare the protection offered by Tma12 in transgenic cotton with that offered by chemical pesticide. Transgenic plants (n = 21; 3 rows of 7 plants) were grown in a 4 × 7 m plot during January–June 2014. Control plants (n = 21; 3 rows of 7 plants) were grown in two plots of the same size. All three plots were challenged with adult whiteflies (20/plant). One of the control plots was protected with chemical pesticide (Confidor (imidacloprid), 0.04% v/v in water, three sprays at intervals of 3 weeks).

T4 generation. Event 14 was evaluated again in July–December 2014 in two plots (n = 21 and 48 plants; 3 rows of 7 plants and 8 rows of 6 plants). Twenty-one plants grown in the smaller plot were used for evaluating whitefly resistance and yield performance; 48 plants from the bigger plot were used for other experiments.

Throughout the trials, observations were made to record the number of whiteflies, symptoms related to the cotton leaf curl disease, plant health and lint yield. A few representative photographs of the field of transgenic cotton in T1 to T4 generations are shown in **Supplementary Figures 11** and **12**.

Yield data (boll number and boll, seed and lint weight) of all four generations were statistically analyzed. For the T1 and T2 generations, the comparison was done using one-way ANOVA followed by Dunnet's *post hoc* test for pair wise comparisons. Data from the T3 and T4 generations were analyzed by Student's *t*-test. The yield data of transgenic, pesticide-protected nontransgenic and nontransgenic cotton plants in the T3 generation were analyzed by one-way ANOVA followed by Tukey's *post hoc* test for multiple comparisons.

Generation-wise experimental setup in contained fields. See **Supplementary Note 1**.

Bioassay of transgenic cotton with nontarget insects. Event 14 transgenic cotton plants (T5 generation) were challenged with nontarget insects *Helicoverpa armigera* (cotton bollworm), *Spodoptera litura* (cotton leafworm), *Aphis gossypii* (cotton aphid), *Phenacoccus solenopsis* (cotton mealybug) and *Menochilus sexmaculatus* (ladybird beetle).

Fully expanded leaves (third to fifth from the top) of transgenic or control cotton plants were excised and challenged with neonatal larvae of *H. armigera* or *S. litura* in an insect culture room. Leaves were replaced every 2 d, and insect mortality was recorded. Because the transgenic plant did not result in any mortality until day 7, the weight of the surviving larvae was recorded. The bioassay was conducted in three biological replicates; each replicate consisted of three leaves of the control and transgenic plants. Data were statistically analyzed by Student's *t*-test.

Aphids were released on three transgenic and control plants each. Twenty aphids (fourth instar) were released on each plant and their population was recorded on day 7 and day 22.

For mealybug, three leaves of transgenic or control cotton plants each were challenged with six mature, healthy female insects (two per leaf). Their numbers were counted on day 6 and day 24. The bioassays of aphids and mealybugs were conducted in three biological replicates; each replicate consisted of three plants or leaves, as stated above. Three leaves per plant were tagged for counting the aphid population. The data were statistically analyzed by Student's *t*-test. To assess the impact of Tma12 cotton on beneficial insects, ladybird beetle was selected as the representative example. Ten newly hatched larvae of ladybird were fed *ad libitum* on cotton aphids, reared on control or transgenic cotton plants. The larvae were fed in 100-mm Petri dishes, and

doi:10.1038/nbt.3665

observations on larval development and larval weight were taken every 24 h until adult emergence. Oviposition in the next generation were also studied by continuing the feeding study. Ten days after emergence, adults of opposite sex were paired. After mating, the females were isolated and kept singly in the Petri dishes. Their fecundity and egg viability were recorded for 10 d. The data were statistically analyzed by Student's *t*-test.

Study of life cycle of whitefly on transgenic cotton plants. Transgenic cotton (line 14) and control cotton plants (12 each) were grown in pots. At the five- to six-leaf stage, the plants were covered individually with 100-µm mesh. Forty newly emerged adult whiteflies were released on each plant. After 10 d, the whiteflies were removed mechanically. Parts of the plants were sacrificed to score the number of eggs and nymphs under stereomicroscope. The experiment was repeated after 35 d with other leaves on the plant.

Assessment of cotton leaf curl virus (CLCuV). Total DNA (100 ng) isolated from the leaves of transgenic and control plants (11 each) was used as template for rolling-circle amplification (RCA) and PCR to amplify the genome of begomoviruses and the satellites. RCA was performed using the TempliPhi amplification kit (GE Healthcare). Phusion High-Fidelity DNA Polymerase (Thermo Scientific, USA) was used in the PCR. The RCA products were digested with BamHI, HindIII, SacI and SalI restriction endonucleases (New England BioLabs) to characterize begomoviruses and the satellites on the basis of sizes of the fragments. The RCA restriction products were cloned in a pDRIVE cloning vector and sequenced using M13 forward and reverse primers. The sequence was used for designing of primers for the amplification of begomoviruses. The primer pair CLB forward and CLB reverse (GAAGAATGTTCATCGTGATAGGT and ACAGTCGCCGTACTGGGCTCAT) was designed to amplify the viral genome of a cotton leaf curl virus (CLCuKoV-Bu). Alpha- and beta-satellites were amplified by PCR using the primer pairs nanofor and nanorev and β01 β04, respectively, as described previously[25]. The PCR products were cloned into pDRIVE vectors (Qiagen). Two RCA and two PCR clones for the virus and the alpha- and beta-satellites in the virus-positive samples were sequenced.

Safety study of Tma12 on rats. The safety study was conducted on male and female Sprague–Dawley rats. CSIR–Central Drug Research Institute is registered under the Committee for the Purpose of Control and Supervision of Experiments on Animals (CPCSEA) with registration number 34/GO/ReBi/S/99 CPCSEA. The animals for the study were approved by the Institutional Animal Ethics Committee with number IAEC/2012/91 (9 October 2012) and procured from the animal facility of the institute. The safety study was conducted as per Schedule Y guidelines.

The subchronic (repeat dose) toxicity study was performed in Sprague–Dawley rats. The dose was calculated on the basis of average daily intake of cottonseed meal by cattle. It is typically 450 g for a calf of 60 kg body weight (~7.5 g/kg)[26]. It is equivalent to ~3.45 g cottonseed protein/kg. As transgenic cotton accumulates Tma12 at 0.01% of TSP, its calculated amount in feed was estimated at 0.345 mg/kg/day. Eight-week-old animals were used in the study. Initial weight of male rats was between 132 ± 11.2 and 144.2 ± 12.5, and that of females was between 132.8 ± 7.92 and 137.8 ± 8.41 (g). They were fed with 10- and 30-fold higher doses of Tma12 (3.45 and 10.35 mg/kg/day, respectively) for 14 d. On an equivalence basis, these doses correspond to 4.5 and 13.5 kg cottonseed meal cake fed to 5.5-month-old calves of 60 kg. Rats administered vehicle (PBS) served as controls. Each of the three groups comprised 5 male and 5 female rats. The presence of Tma12 was analyzed in plasma, urine and feces in high-dose-treated animals by ELISA using anti-Tma12 antibody at predetermined time points up to 32 h. The animals were fed for 14 d, and Tma12 was estimated in plasma on days 2 and 15 at predetermined time points. Nearly 60 different histological, hematological and blood- and urine-related biochemical and organ-development-related parameters were recorded. The toxicity data were subjected to one-way ANOVA (**Supplementary Table 7a–g**), and where the ANOVA was significant, data were analyzed by Bonferroni *post hoc* test. In rare cases where the ANOVA analysis suggested significant differences between Tma12-fed and control rats, the values were compared with published reference ranges[27]. In all cases, despite statistical differences, the parameters were well within the population range, thus indicating that the fern protein may be safe to mammals.

Pepsin-digestion assay. Simulated gastric fluid (SGF) was prepared by dissolving NaCl (122.8 mg) in distilled water (59.2 ml) and adjusting the pH to 1.2 with concentrated HCl. Pepsin (Sigma-Aldrich) was used to prepare SGF. The assay was performed as per standard protocol[28].

Allergenicity assessment. To predict the allergenicity risk of Tma12, the amino acid sequence was searched against an allergen database using Allermatch as per FAO/WHO guidelines. The parameters used were: sliding window of 80 aa, default 6-aa short subsequences match and full fasta alignment.

Experimental methods. *Blinding.* The insect bioassay during screening of ferns and during purification process was conducted in a 'single-blind' manner. Calculation of LC_{50} of purified Tma12 was double blinded. Control samples (containing BSA or buffer alone) were mixed randomly with the experimental samples. For contained field experiments, T2–T4 generation trials were blinded to various extents. In T2 generation, the entomology team had no information about transgenic and control plants while counting insects. Experiments with the T3 and T4 generations were double blind. For the experiment shown in **Table 1**, the pesticide-protected control was known to investigators, and the transgenic and other control were not.

Randomization. For animal studies, randomization was done according to body weight.

Research materials. The purified Tma12, the gene in the expression cassette and seeds of the transgenic cotton lines are available from the CSIR-NBRI institutional biosafety committee for independent evaluation of the technology.

21. Upadhyay, S.K. *et al.* RNA interference for the control of whiteflies (*Bemisia tabaci*) by oral route. *J. Biosci.* **36**, 153–161 (2011).
22. Upadhyay, S.K. *et al.* Interaction of *Allium sativum* leaf agglutinin with midgut brush border membrane vesicles proteins and its stability in *Helicoverpa armigera*. *Proteomics* **10**, 4431–4440 (2010).
23. Kumar, M., Shukla, A.K., Singh, H. & Tuli, R. Development of insect resistant transgenic cotton lines expressing *cry1EC* gene from an insect bite and wound inducible promoter. *J. Biotechnol.* **140**, 143–148 (2009).
24. Sharma, H.C. *Biotechnological Approaches for Pest Management and Ecological Sustainability* (CRC Press, 2009).
25. Kumar, J., Kumar, A., Roy, J.K., Tuli, R. & Khan, J.A. Identification and molecular characterization of begomovirus and associated satellite DNA molecules infecting *Cyamopsis tetragonoloba*. *Virus Genes* **41**, 118–125 (2010).
26. Stewart, L. & Rossi, J. *Using Cotton Byproducts in Beef Cattle Diets (Bulletin 1311)* (University of Georgia Cooperative Extension, 2010).
27. Petterino, C. & Argentino-Storino, A. Clinical chemistry and haematology historical data in control Sprague-Dawley rats from pre-clinical toxicity studies. *Exp. Toxicol. Pathol.* **57**, 213–219 (2006).
28. Fu, T.J., Abbott, U.R. & Hatzos, C. Digestibility of food allergens and nonallergenic proteins in simulated gastric fluid and simulated intestinal fluid—a comparative study. *J. Agric. Food Chem.* **50**, 7154–7160 (2002).