



- (51) **International Patent Classification:**
C12N 15/82 (2006.01) *C07K 14/325* (2006.01)
- (21) **International Application Number:**
PCT/IB2012/051521
- (22) **International Filing Date:**
29 March 2012 (29.03.2012)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
1026/CHE/2011 30 March 2011 (30.03.2011) IN
- (71) **Applicant (for all designated States except US):** **METAHELIX LIFE SCIENCES LIMITED** [IN/IN]; Plot No. 3, KIADB 4th Phase, Bommasandra, Bangalore, Karnataka 560 099 (IN).
- (72) **Inventors; and**
- (75) **Inventors/Applicants (for US only):** **KAMPLI, Nagaraj**, [IN/IN]; C/o Metahelix Life Sciences Limited, Plot No. 3, KIADB 4th Phase, Bommasandra, Bangalore, Karnataka, 560 099 (IN). **SREERAMULU, Shivakumar**, [IN/IN]; C/o Metahelix Life Sciences Limited, Plot No. 3, KIADB 4th Phase, Bommasandra, Bangalore, Karnataka, 560 099 (IN). **DATTAROY, Tomal, Krishna**, [IN/IN]; C/o Metahelix Life Sciences Limited, Plot No. 3, KIADB 4th Phase,

Bommasandra, Bangalore, Karnataka 560 099 (IN). **RAMANATHAN, Vairamani**, [IN/IN]; C/o Metahelix Life Sciences Limited, Plot No. 3, KIADB 4th Phase, Bommasandra, Bangalore, Karnataka, 560 099 (IN). **NARAYANAN, Kottaram, Krishnadas** [IN/IN]; C/o. Metahelix Life Sciences Limited, Plot No. 3, KIADB 4th Phase, Bommasandra, Bangalore, Karnataka, 560 099 (IN).

- (74) **Agents:** **BHOLA, Ravi** et al.; # 4121/B, 6th Cross, 19th A Main, HAL II Stage [Extension], Bangalore, Karnataka 560 038 (IN).

- (81) **Designated States (unless otherwise indicated, for every kind of national protection available):** AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) **Designated States (unless otherwise indicated, for every kind of regional protection available):** ARIPO (BG, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU,

[Continued on next page]

- (54) **Title:** A NUCLEOTIDE SEQUENCE, EXPRESSION CASSETTE, TRANSGENIC EVENTS, CELLS AND METHODS THEREOF

(57) **Abstract:** The present disclosure is in the area of agricultural biotechnology, wherein two transgenic cotton events, viz., MLS9124 and MLS9878 are disclosed. The present disclosure relates specifically to an expression cassette comprising a nucleotide sequence coding for Cry1C gene in plants of genus *Gossypium*. The disclosure also provides vectors and host cells transformed by the said expression cassette and corresponding methods thereof. The disclosure also comprises method of codon optimizing the said gene, and conferring of resistance to plants of genus *Gossypium* from important pests or insects of cultivated crops causing extensive economic damage.

FIG. 1

FIG. 2

FIG. 3

FIG. 4

FIG. 5

FIG. 6

FIG. 7

FIG. 8

FIG. 9

FIG. 10

FIG. 11

FIG. 12

FIG. 13

FIG. 14

FIG. 15

FIG. 16

FIG. 17

FIG. 18

FIG. 19

FIG. 20

FIG. 21

FIG. 22

FIG. 23

FIG. 24

FIG. 25

FIG. 26

FIG. 27

FIG. 28

FIG. 29

FIG. 30

FIG. 31

FIG. 32

FIG. 33

FIG. 34

FIG. 35

FIG. 36

FIG. 37

FIG. 38

FIG. 39

FIG. 40

FIG. 41

FIG. 42

FIG. 43

FIG. 44

FIG. 45

FIG. 46

FIG. 47

FIG. 48

FIG. 49

FIG. 50

FIG. 51

FIG. 52

FIG. 53

FIG. 54

FIG. 55

FIG. 56

FIG. 57

FIG. 58

FIG. 59

FIG. 60

FIG. 61

FIG. 62

FIG. 63

FIG. 64

FIG. 65

FIG. 66

FIG. 67

FIG. 68

FIG. 69

FIG. 70

FIG. 71

FIG. 72

FIG. 73

FIG. 74

FIG. 75

FIG. 76

FIG. 77

FIG. 78

FIG. 79

FIG. 80

FIG. 81

FIG. 82

FIG. 83

FIG. 84

FIG. 85

FIG. 86

FIG. 87

FIG. 88

FIG. 89

FIG. 90

FIG. 91

FIG. 92

FIG. 93

FIG. 94

FIG. 95

FIG. 96

FIG. 97

FIG. 98

FIG. 99

FIG. 100

FIG. 101

FIG. 102

FIG. 103

FIG. 104

FIG. 105

FIG. 106

FIG. 107

FIG. 108

FIG. 109

FIG. 110

FIG. 111

FIG. 112

FIG. 113

FIG. 114

FIG. 115

FIG. 116

FIG. 117

FIG. 118

FIG. 119

FIG. 120

FIG. 121

FIG. 122

FIG. 123

FIG. 124

FIG. 125

FIG. 126

FIG. 127

FIG. 128

FIG. 129

FIG. 130

FIG. 131

FIG. 132

FIG. 133

FIG. 134

FIG. 135

FIG. 136

FIG. 137

FIG. 138

FIG. 139

FIG. 140

FIG. 141

FIG. 142

FIG. 143

FIG. 144

FIG. 145

FIG. 146

FIG. 147

FIG. 148

FIG. 149

FIG. 150

FIG. 151

FIG. 152

FIG. 153

FIG. 154

FIG. 155

FIG. 156

FIG. 157

FIG. 158

FIG. 159

FIG. 160

FIG. 161

FIG. 162

FIG. 163

FIG. 164

FIG. 165

FIG. 166

FIG. 167

FIG. 168

FIG. 169

FIG. 170

FIG. 171

FIG. 172

FIG. 173

FIG. 174

FIG. 175

FIG. 176

FIG. 177

FIG. 178

FIG. 179

FIG. 180

FIG. 181

FIG. 182

FIG. 183

FIG. 184

FIG. 185

FIG. 186

FIG. 187

FIG. 188

FIG. 189

FIG. 190

FIG. 191

FIG. 192

FIG. 193

FIG. 194

FIG. 195

FIG. 196

FIG. 197

FIG. 198

FIG. 199

FIG. 200

FIG. 201

FIG. 202

FIG. 203

FIG. 204

FIG. 205

FIG. 206

FIG. 207

FIG. 208

FIG. 209

FIG. 210

FIG. 211

FIG. 212

FIG. 213

FIG. 214

FIG. 215

FIG. 216

FIG. 217

FIG. 218

FIG. 219

FIG. 220

FIG. 221

FIG. 222

FIG. 223

FIG. 224

FIG. 225

FIG. 226

FIG. 227

FIG. 228

FIG. 229

FIG. 230

FIG. 231

FIG. 232

FIG. 233

FIG. 234

FIG. 235

FIG. 236

FIG. 237

FIG. 238

FIG. 239

FIG. 240

FIG. 241

FIG. 242

FIG. 243

FIG. 244

FIG. 245

FIG. 246

FIG. 247

FIG. 248

FIG. 249

FIG. 250

FIG. 251

FIG. 252

FIG. 253

FIG. 254

FIG. 255

FIG. 256

FIG. 257

FIG. 258

FIG. 259

FIG. 260

FIG. 261

FIG. 262

FIG. 263

FIG. 264

FIG. 265

FIG. 266

FIG. 267

FIG. 268

FIG. 269

FIG. 270

FIG. 271

FIG. 272

FIG. 273

FIG. 274

FIG. 275

FIG. 276

FIG. 277

FIG. 278

FIG. 279

FIG. 280

FIG. 281

FIG. 282

FIG. 283

FIG. 284

FIG. 285

FIG. 286

FIG. 287

FIG. 288

FIG. 289

FIG. 290

FIG. 291

FIG. 292

FIG. 293

FIG. 294

FIG. 295

FIG. 296

FIG. 297

FIG. 298

FIG. 299

FIG. 300

FIG. 301

FIG. 302

FIG. 303

FIG. 304

FIG. 305

FIG. 306

FIG. 307

FIG. 308

FIG. 309

FIG. 310

FIG. 311

FIG. 312

FIG. 313

FIG. 314

FIG. 315

FIG. 316

FIG. 317

FIG. 318

FIG. 319

FIG. 320

FIG. 321

FIG. 322

FIG. 323

FIG. 324

FIG. 325

FIG. 326

FIG. 327

FIG. 328

FIG. 329

FIG. 330

FIG. 331

FIG. 332

FIG. 333

FIG. 334

FIG. 335

FIG. 336

FIG. 337

FIG. 338

FIG. 339

FIG. 340

FIG. 341

FIG. 342

FIG. 343

FIG. 344

FIG. 345

FIG. 346

FIG. 347

FIG. 348

FIG. 349

FIG. 350

FIG. 351

FIG. 352

FIG. 353

FIG. 354

FIG. 355

FIG. 356

FIG. 357

FIG. 358

FIG. 359

FIG. 360

FIG. 361

FIG. 362

FIG. 363

FIG. 364

FIG. 365

FIG. 366

FIG. 367

FIG. 368

FIG. 369

FIG. 370

FIG. 371

FIG. 372

FIG. 373

FIG. 374

FIG. 375

FIG. 376

FIG. 377

FIG. 378

FIG. 379

FIG. 380

FIG. 381

FIG. 382

FIG. 383

FIG. 384

FIG. 385

FIG. 386

FIG. 387

FIG. 388

FIG. 389

FIG. 390

FIG. 391

FIG. 392

FIG. 393

FIG. 394

FIG. 395

FIG. 396

FIG. 397

FIG. 398

FIG. 399

FIG. 400

FIG. 401

FIG. 402

FIG. 403

FIG. 404

FIG. 405

FIG. 406

FIG. 407

FIG. 408

FIG. 409

FIG. 410

FIG. 411

FIG. 412

FIG. 413

FIG. 414

FIG. 415

FIG. 416

FIG. 417

FIG. 418

FIG. 419

FIG. 420

FIG. 421

FIG. 422

FIG. 423

FIG. 424

FIG. 425

FIG. 426

FIG. 427

FIG. 428

FIG. 429

FIG. 430

FIG. 431

FIG. 432

FIG. 433

FIG. 434

FIG. 435

FIG. 436

FIG. 437

FIG. 438

FIG. 439

FIG. 440

FIG. 441

FIG. 442

FIG. 443

FIG. 444

FIG. 445

FIG. 446

FIG. 447

FIG. 448

FIG. 449

FIG. 450

FIG. 451

FIG. 452

FIG. 453

FIG. 454

FIG. 455

FIG. 456

FIG. 457

FIG. 458

FIG. 459

FIG. 460

FIG. 461

FIG. 462

FIG. 463

FIG. 464

FIG. 465

FIG. 466

FIG. 467

FIG. 468

FIG. 469

FIG. 470

FIG. 471

FIG. 472

FIG. 473

FIG. 474

FIG. 475

FIG. 476

FIG. 477

FIG. 478

FIG. 479

FIG. 480

FIG. 481

FIG. 482

FIG. 483

FIG. 484

FIG. 485

FIG. 486

FIG. 487

FIG. 488

FIG. 489

FIG. 490

FIG. 491

FIG. 492

FIG. 493

FIG. 494

FIG. 495

FIG. 496

FIG. 497

FIG. 498

FIG. 499

FIG. 500

FIG. 501

FIG. 502

FIG. 503

FIG. 504

FIG. 505

FIG. 506

FIG. 507

FIG. 508

FIG. 509

FIG. 510

FIG. 511

FIG. 512

FIG. 513

FIG. 514

FIG. 515

FIG. 516

FIG. 517

FIG. 518

FIG. 519

FIG. 520

FIG. 521

FIG. 522

FIG. 523

FIG. 524

FIG. 525

FIG. 526

FIG. 527

FIG. 528

FIG. 529

FIG. 530

FIG. 531

FIG. 532

FIG. 533

FIG. 534

FIG. 535

FIG. 536

FIG. 537

FIG. 538

FIG. 539

FIG. 540

FIG. 541

FIG. 542

FIG. 543

FIG. 544

FIG. 545

FIG. 546

FIG. 547

FIG. 548

FIG. 549

FIG. 550

FIG. 551

FIG. 552

FIG. 553

FIG. 554

FIG. 555

FIG. 556

FIG. 557

FIG. 558

FIG. 559

FIG. 560

FIG. 561

FIG. 562

FIG. 563

FIG. 564

FIG. 565

FIG. 566

FIG. 567

FIG. 568

FIG. 569

FIG. 570

FIG. 571

FIG. 572

FIG. 573

FIG. 574

FIG. 575

FIG. 576

FIG. 577

FIG. 578

FIG. 579

FIG. 580

FIG. 581

FIG. 582

FIG. 583

FIG. 584

FIG. 585

FIG. 586

FIG. 587

FIG. 588

FIG. 589

FIG. 590

FIG. 591

FIG. 592

FIG. 593

FIG. 594

FIG. 595

FIG. 596

FIG. 597

FIG. 598

FIG. 599

FIG. 600

FIG. 601

FIG. 602

FIG. 603

FIG. 604

FIG. 605

FIG. 606

FIG. 607

FIG. 608

FIG. 609

FIG. 610

FIG. 611

FIG. 612

FIG. 613

FIG. 614

FIG. 615

FIG. 616

FIG. 617

FIG. 618

FIG. 619

FIG. 620

FIG. 621

FIG. 622

FIG. 623

FIG. 624

FIG. 625

FIG. 626

FIG. 627

FIG. 628

FIG. 629

FIG. 630

FIG. 631

FIG. 632

FIG. 633

FIG. 634

FIG. 635

FIG. 636

FIG. 637

FIG. 638

FIG. 639

FIG. 640

FIG. 641

FIG. 642

FIG. 643

FIG. 644

FIG. 645

FIG. 646

FIG. 647

FIG. 648

FIG. 649

FIG. 650

FIG. 651

FIG. 652

FIG. 653

FIG. 654

FIG. 655

FIG. 656

FIG. 657

FIG. 658

FIG. 659

FIG. 660

FIG. 661

FIG. 662

FIG. 663

FIG. 664

FIG. 665

FIG. 666

FIG. 667

FIG. 668

FIG. 669

FIG. 670

FIG. 671

FIG. 672

FIG. 673

FIG. 674

FIG. 675

FIG. 676

FIG. 677

FIG. 678

FIG. 679

FIG. 680

FIG. 681

FIG. 682

FIG. 683

FIG. 684

FIG. 685

FIG. 686

FIG. 687

FIG. 688

FIG. 689

FIG. 690

FIG. 691

FIG. 692

FIG. 693

FIG. 694

FIG. 695

FIG. 696

FIG. 697

FIG. 698

FIG. 699

FIG. 700

FIG. 701

FIG. 702

FIG. 703

FIG. 704

FIG. 705

FIG. 706

FIG. 707

FIG. 708

FIG. 709

FIG. 710

FIG. 711

FIG. 712

FIG. 713

FIG. 714

FIG. 715

FIG. 716

FIG. 717

FIG. 718

FIG. 719

FIG. 720

FIG. 721

FIG. 722

FIG. 723

FIG. 724

FIG. 725

FIG. 726

FIG. 727

FIG. 728

FIG. 729

FIG. 730

FIG. 731

FIG. 732

FIG. 733

FIG. 734

FIG. 735

FIG. 736

FIG. 737

FIG. 738

FIG. 739

FIG. 740

FIG. 741

FIG. 742

FIG. 743

FIG. 744

FIG. 745

FIG. 746

FIG. 747

FIG. 748

FIG. 749

FIG. 750

FIG. 751

FIG. 752

FIG. 753

FIG. 754

FIG. 755

FIG. 756

FIG. 757

FIG. 758

FIG. 759

FIG. 760

FIG. 761

FIG. 762

FIG. 763

FIG. 764

FIG. 765

FIG. 766

FIG. 767

FIG. 768

FIG. 769

FIG. 770

FIG. 771

FIG. 772

FIG. 773

FIG. 774

FIG. 775

FIG. 776

FIG. 777

FIG. 778

FIG. 779

FIG. 780

FIG. 781

FIG. 782

FIG. 783

FIG. 784

FIG. 785

FIG. 786

FIG. 787

FIG. 788

FIG. 789

FIG. 790

FIG. 791

FIG. 792

FIG. 793

FIG. 794

FIG. 795

FIG. 796

FIG. 797

FIG. 798

FIG. 799

FIG. 800

FIG. 801

FIG. 802

FIG. 803

FIG. 804

FIG. 805

FIG. 806

FIG. 807

FIG. 808

FIG. 809

FIG. 810

FIG. 811

FIG. 812

FIG. 813

FIG. 814

FIG. 815

FIG. 816

FIG. 817

FIG. 818

FIG. 819

FIG. 820

FIG. 821

FIG. 822

FIG. 823

FIG. 824

FIG. 825

FIG. 826

FIG. 827

FIG. 828

FIG. 829

FIG. 830

FIG. 831

FIG. 832

FIG. 833

FIG. 834

FIG. 835

FIG. 836

FIG. 837

FIG. 838

FIG. 839

FIG. 840

FIG. 841

FIG. 842

FIG. 843

FIG. 844

FIG. 845

FIG. 846

FIG. 847

FIG. 848

FIG. 849

FIG. 850

FIG. 851

FIG. 852

FIG. 853

FIG. 854

FIG. 855

FIG. 856

FIG. 857

FIG. 858

FIG. 859

FIG. 860

FIG. 861

FIG. 862

FIG. 863

FIG. 864

FIG. 865

FIG. 866

FIG. 867

FIG. 868

FIG. 869

FIG. 870

FIG. 871

FIG. 872

FIG. 873

FIG. 874

FIG. 875

FIG. 876

FIG. 877

FIG. 878

FIG. 879

FIG. 880

FIG. 881

FIG. 882

FIG. 883

FIG. 8



TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

Declarations under Rule 4.17:

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

**“A NUCLEOTIDE SEQUENCE, EXPRESSION CASSETTE, TRANSGENIC
EVENTS, CELLS AND METHODS THEREOF”**

TECHNICAL FIELD

The present disclosure is in the area of agricultural biotechnology, wherein two transgenic cotton events, viz., MLS9124 and MLS9878 are disclosed. The present disclosure relates specifically to an expression cassette comprising a nucleotide sequence coding for Cry1C gene in plants of genus *Gossypium*. The disclosure also provides vectors and host cells transformed by the said expression cassette and corresponding methods thereof. The disclosure also comprises method of codon optimizing the said gene, and conferring of resistance to plants of genus *Gossypium* from important pests or insects of cultivated crops causing extensive economic damage.

BACKGROUND AND PRIOR ART OF THE DISCLOSURE

Pests cause extensive economic damage of crop plants worldwide and necessitate extensive use of chemical pesticides for control. The losses due to damages by insects cause significant losses in crop yields, a situation that the burgeoning population worldwide can ill-afford. The consequent fallout of this extensive pesticide usage is the damage to the ecosystem caused by leaching into the soils as well as by run-off into the water bodies, thus disturbing the flora and fauna. Pesticides, being generally non-preferential for the target organisms, also affect beneficial insects. The residual pesticides in the crops get into the food chain causing severe damages to the living system, including humans where instances of physical deformities have been reported.

Spodoptera litura, a pest of cotton and tobacco, is also known as the Oriental leafworm moth, Cluster caterpillar, Cotton leafworm, Tobacco cutworm, Tropical armyworm, Taro caterpillar, Tobacco budworm, Rice cutworm, and Cotton Cutworm. This moth is found in Asia, with some specific problematic pest population reports occurring in Cambodia, Hong Kong, India, the Pacific islands, Guam, American Samoa, and Hawaii. In Australia, it is found in northern two thirds of the country. It is established in the U.S., and it is a pest of national, regulatory concern. Besides cotton and tobacco, this insect causes extensive damage in many vegetables and is a major pest in corn as well as a minor pest in rice.

One of the most widely used *Bacillus thuringiensis* strains used in the biological control of insect pests of crops is *aizawai* which carries the crystal protein Cry1C. The very high insecticidal activity of this Cry protein on the diamond-back moth and the army worm (*Plutella xylostella*, and *Spodoptera* spp. respectively) combined with the impeccable safety record of the strains carrying this crystal protein has been primarily responsible for this widespread use. Extensive studies on the biosafety of this strain have been done, mostly in the US and details have been published under the joint sponsorship of the UNEP, ILO, WHO and the International Program of Chemical Safety (1).

The US Environmental Protection Agency (EPA) has approved the registration of several formulations, either Bt strains or *Pseudomonas fluorescens* engineered to express Cry1C protein as safe pesticides to be used in all agricultural crops including fruits, vegetables, corn, cotton and nuts; turf; forests; ornamentals; landscape trees; nursery crops. The long history of safe use of these formulations even in crops which are eaten raw like lettuce attest to the high degree of safety of the organism and the Cry1C protein that is expressed in them. Many commercial products by some of the leading companies in the world like Abbot (Xentari, Florbac) and Novartis (Certan) contain the expressed Cry1C protein and are in wide-spread use in many countries (2).

The gene coding for the Cry1C family of proteins has been extensively studied. A search of the GENBANK database shows many sequences of this family of protein deposited. A non-exhaustive list is presented in table 1. In a very significant recent study (3), Cry1C expressing broccoli plants were exposed to the laboratory bred resistant strains of diamond back moth, *Plutella xylostella* in order to understand the effect of the Cry1C protein on the parasitoid, *Diadegma insulare*, which is a natural enemy of *Plutella*. The results of this study provide the first clear evidence of the lack of hazard to a parasitoid by a Bt plant, compared to traditional insecticides.

Gene designation	Accession No.	Reference
Cry1C a1	X 07518	4

Cry1C a2	X 13620	5
Cry1C a3	M 73251	6
Cry1C a4	A 27642	7
Cry1C a5	X 96682	8
Cry1C b1	M 97880	9

Table 1. Selection of Cry1C gene sequences deposited in public databases.

An alternative to the use of biological control has been the implementation of transgenic technology. Crop species have been genetically modified to express different Cry genes to control insects. The earliest commercial disclosure of this technology has been Bollgard that expresses the Cry1Ac gene in cotton and this was developed by Monsanto (10). Bollgard was developed to resist the bollworm *Helicoverpa armigera* and commercial cultivation was practised in the United States in 1996. Subsequently, Bollgard II was developed by Monsanto that had the stacked combination of Cry1Ac and Cry2Ab and was put into commercial cultivation in 2000 to delay the onset of resistance to the Bt toxin by the pest (10, 11).

Studies on effects of two transgenic cotton lines carrying a Cry1A gene on *Spodoptera litura*, were conducted during 2002-2005 in the cotton planting region of the Yangtze River valley of China. Results showed that the *Spodoptera* larvae had low susceptibility to Bt cotton (12). There was no significant difference in larval population densities in conventional and Bt cotton fields. However, the larval populations of the insect on conventional plants treated with chemical insecticides for control of target pest of Bt cotton were significantly lower than that in Bt cotton fields. These results indicated that *Spodoptera litura* had the potential to become a major and alarming pest in Bt cotton fields, and therefore efforts to develop an effective alternative management strategy was needed. The efficacy of Cry1C in controlling insects that have developed resistance to Cry1A toxins in cauliflower is reported (13).

Cry1C was expressed in corn to harbour resistance to *Spodoptera frugiperda* (15). This was not a codon-optimized gene. Insect resistant transgenic indica rice has been reported where Cry1C was expressed (16). Several patents exist for Cry1C but the

genes protected in these code for the protein found in *Bacillus thuringiensis* subspecies aizawai (17-26).

STATEMENT OF THE DISCLOSURE

Accordingly, the present disclosure relates to a nucleotide sequence set forth as SEQ ID No. 1 or a nucleotide sequence comprising sequence set forth as SEQ ID No.1; an expression cassette set forth as SEQ ID No.3, comprising CaMV 35S promoter, nucleotide sequence set forth as SEQ ID No. 1 or a nucleotide sequence comprising sequence set forth as SEQ ID No.1 and 35S 3' untranslated region; a vector comprising sequence set forth as SEQ ID No. 5, having an expression cassette as claimed above; a transformed host cell comprising the vector as claimed above; a method of obtaining a transformed host cell comprising an expression cassette set forth as SEQ ID No.3, said method comprising acts of: (a) inserting an expression cassette set forth as SEQ ID No.3 into a vector, and (b) transforming a host cell with said vector to obtain the transformed host cell; a transgenic cell comprising an expression cassette as claimed above; a method of obtaining a transgenic plant member of genus *Gossypium*, said method comprising acts of: (a) codon optimizing nucleotide sequence of native Cry1C gene to obtain nucleotide sequence set forth as SEQ ID No.1, (b) inserting an expression cassette as claimed above into a vector and transforming a host cell with said vector to obtain a transformed host cell, and (c) infecting the plant with the transformed host cell followed by tissue culturing to obtain a transgenic plant member of genus *Gossypium*; a transgenic transformation event MLS9124 or MLS9878, said event comprising a nucleotide sequence set forth as SEQ ID No. 1 or a nucleotide sequence comprising sequence set forth as SEQ ID No.1 in a plant member of genus *Gossypium* or any part thereof; a method of detecting presence of a transgene comprising an expression cassette set forth as SEQ ID No.3, in a transgenic plant member of genus *Gossypium*, said method comprising acts of: (a) extracting DNA from the plant source and performing nucleic acid amplification of junction regions of the transgene and the plant to obtain an amplicon, said amplification carried out by primers corresponding to the regions selected from a group comprising left border region of the transgene, right border region of the transgene, left border region of the plant DNA and right border region of the plant DNA or any combination thereof, and (b) detecting and analyzing the amplicon to detect the presence of said transgene in the transgenic plant; primers set forth as SEQ

ID No. 11, SEQ ID No. 12, SEQ ID No. 13, SEQ ID No. 14, SEQ ID No. 15, SEQ ID No. 16, SEQ ID No. 19 and SEQ ID No. 20; and a method of codon optimizing Cry1C gene in plant member of genus *Gossypium*, said method comprising acts of: (a) preparing a codon usage table, for plurality of constitutively expressed genes, on the basis of GC content, TA doublet avoidance in second and third positions of codons, transcription termination signals, splicing signals and polyadenylation signals, from conventionally known constitutively expressed genes in cotton, and (b) using said table for codon optimizing the Cry1C gene.

BRIEF DESCRIPTION OF THE ACCOMPANYING FIGURES

In order that the disclosure may be readily understood and put into practical effect, reference will now be made to exemplary embodiments as illustrated with reference to the accompanying figures. The figure together with a detailed description below, are incorporated in and form part of the specification, and serve to further illustrate the embodiments and explain various principles and advantages, in accordance with the present disclosure where

Figure 1 illustrates pairwise alignment of the codon optimized Cry1CGh gene with the 1st 1890 bp of the Cry1C gene (4).

Figure 2 illustrates a map of the Cry1CGh expression cassette in the transgenic cotton event MLS9124 and MLS9878.

Figure 3 illustrates a map of the binary vector pMH72.

Figure 4 illustrates a map of the plasmid pMH82 with T-DNA containing two Cry 1C expression cassettes repeated in tandem used for the generation of the cotton transgenic events MLS9124 and MLS9878.

Figure 5 illustrates southern hybridization of event MLS9124 digested with EcoRI and HindIII.

Figure 6 illustrates southern hybridization of event MLS9878 digested with EcoRI and HindIII.

Figure 7 illustrates a diagnostic PCR process for the amplification of the left border junction of transgenic cotton event MLS9878.

Figure 8 illustrates a diagnostic PCR process for the amplification of the right border junction of transgenic cotton event MLS9878.

Figure 9 illustrates a diagnostic PCR to locate the unique intermediate region between the two expression cassettes in MLS9124 and MLS9878.

Figure 10 illustrates a comparative efficacy of Cry1C in transgenic cotton events MLS9124 and MLS9878 in *Spodoptera litura* (A) and *Helicoverpa armigera* (B).

Figure 11 illustrates a diagnostic PCR process for the amplification of the right border junction of transgenic cotton event MLS9124.

DETAILED DESCRIPTION OF THE DISCLOSURE

The present disclosure relates to a nucleotide sequence set forth as SEQ ID No. 1 or a nucleotide sequence comprising sequence set forth as SEQ ID No.1.

The present disclosure further relates to an expression cassette set forth as SEQ ID No.3, comprising CaMV 35S promoter, nucleotide sequence set forth as SEQ ID No. 1 or a nucleotide sequence comprising sequence set forth as SEQ ID No.1 and 35S 3' untranslated region.

In an embodiment of the present disclosure, the nucleotide sequence codes for Cry1C gene and corresponding amino acid sequence set forth as SEQ ID No. 2 codes for Cry1C protein.

The present disclosure further relates to a vector comprising sequence set forth as SEQ ID No. 5, having an expression cassette as claimed above.

In an embodiment of the present disclosure, the vector is selected from a group comprising an expression vector, replicable vector, transformation vector, binary vector and shuttle vectors or any combination thereof.

In another embodiment of the present disclosure, the vector is preferably binary vector and wherein the expression cassette is either a single copy sequence or tandem repeat sequence separated by an intervening sequence set forth as SEQ ID No.6.

The present disclosure further relates to a transformed host cell comprising the vector as claimed above.

The present disclosure further relates to a method of obtaining a transformed host cell comprising an expression cassette set forth as SEQ ID No.3, said method comprising

acts of: (a) inserting an expression cassette set forth as SEQ ID No.3 into a vector, and (b) transforming a host cell with said vector to obtain the transformed host cell.

In an embodiment of the present disclosure, the expression cassette comprise CaMV 35S promoter, nucleotide sequence set forth as SEQ ID No. 1 or a nucleotide sequence comprising sequence set forth as SEQ ID No.1 and 35S 3' untranslated region.

In another embodiment of the present disclosure, the nucleotide sequence codes for Cry1C gene and corresponding amino acid sequence set forth as SEQ ID No. 2 codes for Cry1C protein.

In yet another embodiment of the present disclosure, the host cell is *Agrobacterium*.

In still another embodiment of the present disclosure, the transforming of the host cell is carried out using techniques selected from a group comprising electroporation, tri-parental mating, microinjection, genegun method, PEG mediated transfer, Calcium phosphate method, liposome mediated transfer and or any combination thereof.

The present disclosure further relates to a transgenic cell comprising an expression cassette as claimed above.

In an embodiment of the present disclosure, the transgenic cell is a transgenic plant cell exhibiting resistance to lepidopteran insect(s) or pest(s) infestation; and wherein the plant is a member of genus *Gossypium*.

The present disclosure further relates to a method of obtaining a transgenic plant member of genus *Gossypium*, said method comprising acts of: (a) codon optimizing nucleotide sequence of native Cry1C gene to obtain nucleotide sequence set forth as SEQ ID No.1, (b) inserting an expression cassette as claimed above into a vector and transforming a host cell with said vector to obtain a transformed host cell, and (c) infecting the plant with the transformed host cell followed by tissue culturing to obtain a transgenic plant member of genus *Gossypium*.

In an embodiment of the present disclosure, the expression cassette is either a single copy sequence or tandem repeat sequence separated by an intervening sequence set forth as SEQ ID No.6.

In another embodiment of the present disclosure, the host cell is *Agrobacterium*.

In yet another embodiment of the present disclosure, transforming the host cell is carried out using techniques selected from a group comprising electroporation, tri-parental mating, microinjection, genegun method, PEG mediated transfer, Calcium phosphate method and liposome mediated transfer or any combination thereof.

In still another embodiment of the present disclosure, the infection and tissue culturing comprises incubating cotton leaf with the transformed host cell; transferring and incubating the cotton leaf in bacterial selection medium followed by incubation in plant selection medium; transferring to embryogenesis medium till the occurrence of embryogenesis followed by sub-culturing; transferring to suspension medium followed by germination medium; transferring to basal medium after the growth of leaves and root hairs; hardening followed by transfer to soil and greenhouse to obtain the said transgenic plant.

The present disclosure further relates to a transgenic transformation event MLS9124 or MLS9878, said event comprising a nucleotide sequence set forth as SEQ ID No. 1 or a nucleotide sequence comprising sequence set forth as SEQ ID No.1 in a plant member of genus *Gossypium* or any part thereof.

The present disclosure further relates to a method of detecting presence of a transgene comprising an expression cassette set forth as SEQ ID No.3, in a transgenic plant member of genus *Gossypium*, said method comprising acts of: (a) extracting DNA from the plant source and performing nucleic acid amplification of junction regions of the transgene and the plant to obtain an amplicon, said amplification carried out by primers corresponding to the regions selected from a group comprising left border region of the transgene, right border region of the transgene, left border region of the plant DNA and right border region of the plant DNA or any combination thereof, and (b) detecting and analyzing the amplicon to detect the presence of said transgene in the transgenic plant.

In an embodiment of the present disclosure, the primers are selected from a group comprising SEQ ID No. 11 corresponding to the left border region of the transgene , SEQ ID No. 13 corresponding to the right border region of the transgene, SEQ ID No. 12 corresponding to the left border region of the plant DNA and SEQ ID No. 14 corresponding to the right border region of the plant DNA or any combination thereof.

The present disclosure further relates to primers set forth as SEQ ID No. 11, SEQ ID No. 12, SEQ ID No. 13, SEQ ID No. 14, SEQ ID No. 15, SEQ ID No. 16, SEQ ID No. 19 and SEQ ID No. 20.

In an embodiment of the present disclosure, the primers having SEQ ID Nos. 12, 13, 15 and 19 are sense primers and the primers having SEQ ID Nos. 11, 14, 16 and 20 are anti-sense primers.

The present disclosure further relates to a method of codon optimizing Cry1C gene in plant member of genus *Gossypium*, said method comprising acts of: (a) preparing a codon usage table, for plurality of constitutively expressed genes, on the basis of GC content, TA doublet avoidance in second and third positions of codons, transcription termination signals, splicing signals and polyadenylation signals, from conventionally known constitutively expressed genes in cotton, and (b) using said table for codon optimizing the Cry1C gene.

In an embodiment of the present disclosure, the nucleotide sequence is codon optimized for plants of genus *Gossypium*.

A codon optimized gene is necessary to code for the Cry1C protein that would confer resistance to *S. litura* in cotton, tobacco, tomato, cabbage, cauliflower, broccoli, rice and maize. Another object of the disclosure is to incorporate it in a stack with other Bt genes to delay the onset of breakdown of resistance to pests. Cry1C is also responsible to confer resistance to the crop plants to other lepidopteran pests.

In an embodiment of the present disclosure, the cotton plant is selected from a group comprising *Gossypium hirsutum*, *Gossypium barbadense*, *Gossypium arboreum* and *Gossypium herbaceum*. Preferably, the cotton plant is *Gossypium hirsutum*.

Cry1C is a potent insecticidal protein found in several subspecies of *B. thuringiensis* that is effective against a wide range of lepidopteran pests of crops. The range of the crops is very diverse and includes, but is not restricted to, cotton, tobacco, tomato, cauliflower, broccoli, rice and maize. Besides conferring resistance to lepidopteran pests that are not affected by the Cry1A series of insecticidal proteins, this gene, when stacked with any Cry1A protein also helps the delay of onset of resistance to Cry1A group of proteins expressed in transgenic crops in the target pests.

This disclosure relates to two transgenic cotton events that carry a unique Cry1C protein gene that is optimized based on the codon usage for the most constitutively expressed proteins in cotton. This gene, when expressed in tobacco and cotton, demonstrates efficacy against major insect pests hitherto uncontrolled by any Cry1A protein.

The N-terminal 630 amino acid sequence of said Cry1C gene was codon optimized using a codon usage table for highly constitutively expressed genes based on the codon preference for cotton (27).

The Pairwise alignment of the codon optimized Cry1CGh gene with the 1st 1890 bp of the Cry1C gene is provided in figure 1.

Based on the codon preference for cotton (27), a table was created for the most constitutively expressed genes (Table 2):

Codon	Gh Actin		Gh Rubisco		Gh Tubulin		Gh Ribosomal Proteins	
	Frequency per 1000 aa	Probability	Frequency per 1000 aa	Probability	Frequency per 1000 aa	Probability	Frequency per 1000 aa	Probability
CGA	1	0.06	5	0.12	1	0.03	0	0.00
CGC	1	0.06	1	0.02	10	0.31	0	0.00
CGG	0	0.00	1	0.02	0	0.00	1	0.17

CGT	10	0.26	11	0.26	13	0.41	2	0.33
AGA	1	0.06	6	0.14	7	0.22	0	0.00
AGG	5	0.28	18	0.41	1	0.03	3	0.30
	18		42		32		6	
CTA	2	0.07	2	0.04	0	0.00	1	0.20
CTC	7	0.26	13	0.24	6	0.27	0	0.00
CTG	2	0.07	6	0.11	3	0.14	2	0.40
CTT	8	0.40	13	0.24	11	0.30	0	0.00
TTA	4	0.15	1	0.02	0	0.00	0	0.00
TTG	4	0.15	20	0.36	2	0.09	2	0.40
	27		55		22		5	
TCA	1	0.04	4	0.09	6	0.18	1	0.20
TCC	7	0.29	10	0.23	10	0.30	1	0.20
TCG	0	0.00	4	0.09	4	0.12	0	0.00
TCT	5	0.21	4	0.09	7	0.21	1	0.20
AGC	5	0.21	8	0.19	3	0.09	1	0.20
AGT	6	0.25	13	0.40	3	0.09	1	0.20
	24		43		33		5	
ACA	2	0.08	9	0.20	2	0.06	2	0.25
ACC	7	0.29	22	0.39	18	0.53	5	0.60
ACG	1	0.04	4	0.09	1	0.03	0	0.00
ACT	14	0.39	9	0.20	12	0.36	1	0.13
	24		44		33		8	
CCA	7	0.35	13	0.32	10	0.33	1	0.33
CCC	2	0.10	14	0.34	11	0.37	0	0.00
CCG	3	0.15	4	0.10	0	0.00	1	0.33
CCT	8	0.40	10	0.24	9	0.30	1	0.33
	20		41		30		3	
GCA	6	0.19	7	0.10	11	0.15	2	0.67
GCC	7	0.22	26	0.37	21	0.29	0	0.00
GCG	2	0.06	3	0.04	2	0.03	0	0.00
GCT	17	0.33	35	0.49	38	0.54	1	0.33
	32		71		72		3	
GGA	10	0.33	30	0.38	14	0.23	2	0.20
GGC	5	0.17	9	0.11	4	0.07	1	0.10
GGG	3	0.10	12	0.15	2	0.03	2	0.20
GGT	12	0.40	29	0.46	41	0.67	5	0.30
	30		80		61		10	
GTA	0	0.00	2	0.03	0	0.00	1	0.17
GTC	6	0.24	18	0.31	25	0.31	1	0.17
GTG	11	0.40	13	0.22	16	0.20	3	0.30
GTT	8	0.32	26	0.44	39	0.49	1	0.17
	25		59		80		6	
AAA	9	0.45	25	0.45	9	0.19	4	0.17
AAG	11	0.33	31	0.55	38	0.61	20	0.60
	20		56		47		24	
AAC	5	0.36	45	0.79	12	0.63	1	1.00
AAT	4	0.44	12	0.21	7	0.37	0	0.00

	9		57		19		1	
CAA	6	0.60	26	0.76	4	0.22	4	0.57
CAG	4	0.40	8	0.24	14	0.74	3	0.43
	10		34		18		7	
CAC	3	0.33	4	1.00	7	0.54	5	0.33
CAT	6	0.47	0	0.00	5	0.42	1	0.17
	9		4		12		6	
GAA	16	0.27	29	0.59	44	0.74	0	0.00
GAG	12	0.43	19	0.40	38	0.46	2	1.00
	28		48		82		2	
GAC	6	0.27	36	0.59	13	0.23	2	0.27
GAT	16	0.73	28	0.44	44	0.77	1	0.33
	22		64		57		3	
TAC	4	0.27	21	0.40	14	0.45	3	0.75
TAT	11	0.73	10	0.32	17	0.73	1	0.25
	15		31		31		4	
TGC	4	1.00	6	0.77	13	0.72	5	1.00
TGT	0	0.00	3	0.33	5	0.28	0	0.00
	4		9		18		5	
TTC	8	0.73	27	0.73	17	0.49	1	0.33
TTT	3	0.27	9	0.25	18	0.51	2	0.67
	11		36		35		3	
ATA	1	0.03	2	0.04	1	0.03	0	0.00
ATC	10	0.34	31	0.57	10	0.32	1	0.33
ATT	18	0.62	21	0.39	20	0.25	2	0.67
	29		54		31		3	
ATG	15	1.00	37	1.00	18	1.00	1	1.00
TGG	4	1.00	8	1.00	12	1.00	0	1.00

Table: 2 – Codon Optimization usage table.

The designed sequence of the gene is given below as SEQ ID No. 1. The gene was designed using the N-terminal 630 amino acids [SEQ ID No. 2] of the Cry1C protein from *Bacillus thuringiensis* subspecies *entomocidus* (4, 30), unlike the sequences cited in references 17 – 26, which are for *B. thuringiensis* subspecies *aizawai*. Factors that were taken into consideration while designing the gene were GC content, TA doublet avoidance in the second and third positions of codons, transcription termination signal (AGTNNA), splicing signals (5' splice junction: CRG/GTAAGT; 3' junction: YNAG/N) and polyadenylation signals (AATAAA, AATGAA, AATAAT, GATAAA, AATAAG, AATATT).

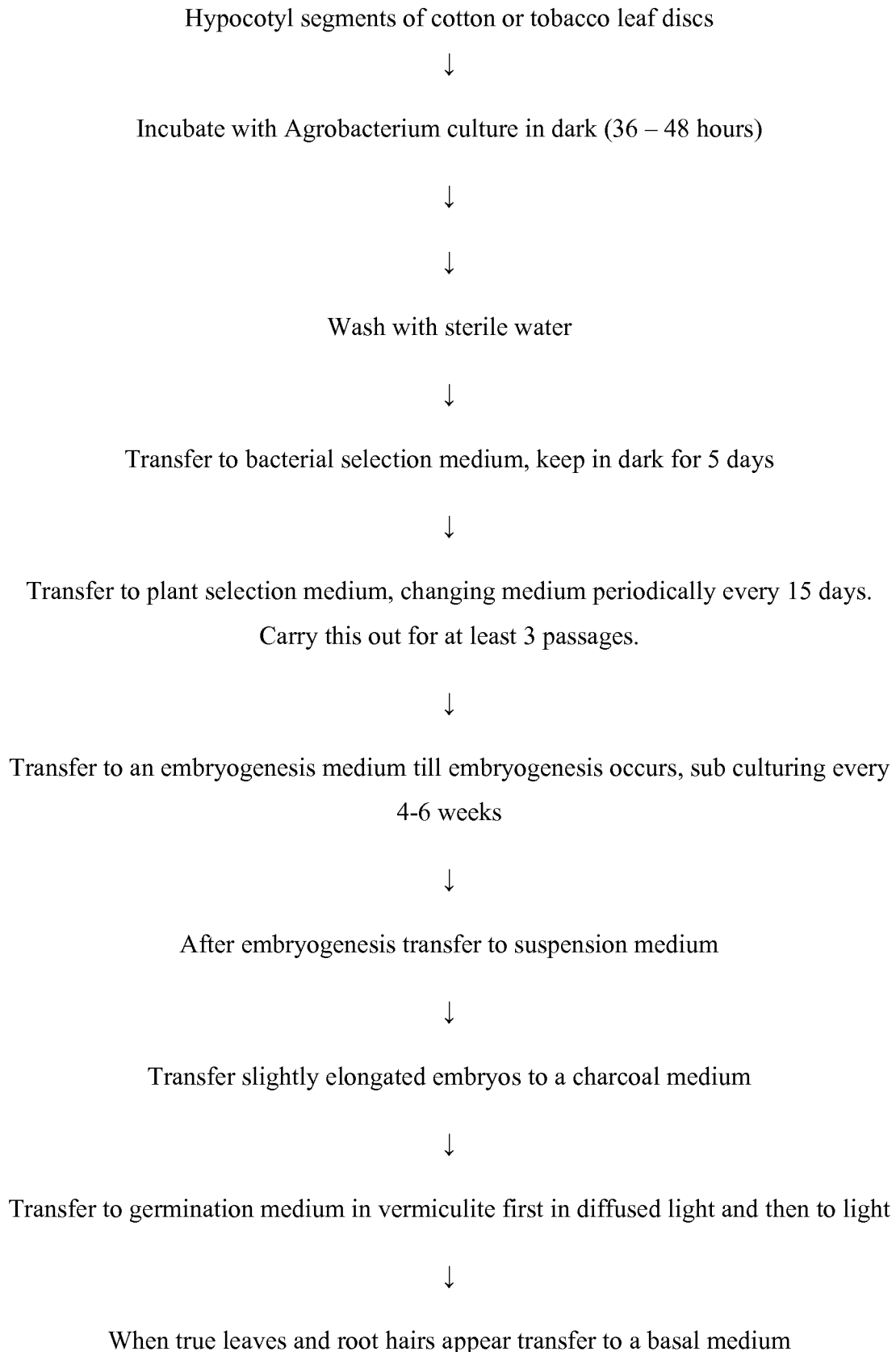
The gene thus obtained was cloned in a binary vector pMH72 (figure 3, SEQ ID No. 4) which was developed to express the nptII gene for plant selection from pGA643 (28) in pPZP201 (29). The gene was cloned to be driven by the cauliflower mosaic virus (CaMV) double 35S promoter (D35SP) and the 35S 3' UTR was used for transcriptional termination. The resultant binary plasmid pMH82 as described in figure 4 and its T-DNA sequence [given in SEQ ID No. 5], was mobilized into *Agrobacterium tumefaciens* EHA105. The mobilization of pMH82 into *Agrobacterium tumefaciens* EHA105 can be carried out using various conventionally known techniques of gene transfer which include microinjection, electroporation, tri-parental mating, genegun method, PEG mediated transfer, Calcium phosphate method, liposome mediated transfer. In an embodiment of the present disclosure, electroporation method is used wherein, electrocompetent *Agrobacterium tumefaciens* cells are prepared and frozen in glycerol at -70°C. The cells are thawed on ice prior to electroporation. Approximately, 50 ng of plasmid DNA is mixed with the electrocompetent cells and transferred to a pre-chilled cuvette. A pulse of 1.8 kV/zero millisecond is applied (feature Ec1 on BioRad MicroPulser Electroporator). The cells are recovered in 1 ml of Luria Bertani Medium for 45 minutes by shaking at 200 rpm at 26°C. The recovered culture is further diluted 25 fold and plated on a minimal medium plated with kanamycin. This plate is incubated at 26°C for 2 days. The transformed cells are selected and used further.

Hypocotyl bits of cotton variety Coker are infected with the *Agrobacterium* cultures harbouring the plasmid pMH82. Embryogenic calli are obtained and regenerated following which regenerants are assayed for resistance to *S. litura*, *Earias vitella* and *Helicoverpa armigera*.

The uniqueness of the construct in pMH82 arises from 2 aspects, viz., the presence of dual expression cassettes [each expression cassette set forth as SEQ ID No. 3] and a unique 57 bp region in between the two expression cassettes, as given in SEQ ID No. 6.

Map of the Cry1CGh expression cassette in the transgenic cotton event MLS9124 and MLS9878 is provided in figure 2.

A Flow chart depicting the course of events for obtaining transgenic cotton plants is provided as below:





Harden the plants in coco peat cum perlite



Transfer to soil and into the greenhouse

Two independent transgenic cotton events, viz., MLS9124 and MLS9878, were further studied by Southern hybridization (31), analysis of the site of integration using TAIL-PCR (32), ELISA for studying levels of protein expression and bioassays by challenging the events with *Spodoptera litura*.

For Southern hybridization, 20 µg genomic DNA was digested with either HindIII or EcoRI overnight and separated on a 0.8% agarose gel in TBE. The gel was transferred to Nylon membrane by downward capillary transfer and hybridized to either a 2.317 kb nptII probe [SEQ ID No. 7] or a 1.954 kb Cry1CGh probe[SEQ ID No. 8]. The probes were labeled with α -³²P dCTP using a randomly primed extension with Klenow and hybridized overnight with formamide. Washed gels were exposed to Kodak X-OMAT films.

Southern Hybridization

Southern hybridization of event MLS9124 digested with EcoRI and HindIII (figure 5); comprises:

- (a) Hybridization with probe I (NptII probe: EcoRI-HindIII fragment from pMH82-2317 bp)
- (b) Hybridization with probe II (Cry1CGh probe: XhoI fragment from pMH82-1954 bp)

Lane No.	Sample	Panel A (kb)	Panel B (kb)
1	9124 Plant #1 HindIII digest	>2.4 kb	2.979kb, 5.729kb
2	9124 Plant #1 EcoRI digest	8.039kb	2.979kb, 8.039kb
3	9124 Plant #5 HindIII digest	>2.4 kb	2.979kb, 5.729kb

4	9124 Plant #5 EcoRI digest	8.039kb	2.979kb, 8.039kb
---	----------------------------	---------	------------------

Table: 3 - Locations of the reference sizes of the λ DNA digested with HindIII are given between the two panels [event MLS9124]

Southern hybridization of event MLS9878 digested with EcoRI and HindIII (Figure 6); comprises:

- (a) Hybridization with probe I (NptII probe: EcoRI-HindIII fragment from pMH82-2317 bp)
- (b) Hybridization with probe II (Cry1CGh probe: XhoI fragment from pMH82-1954 bp)

Expectations:

Restriction Enzyme	Panel A	Panel B
9878/ <i>EcoRI</i>	8040 bp	8040 bp, 2980 bp
9878/ <i>HindIII</i>	> 2644 bp	5730 bp, 2980 bp

Table: 4 – Expected locations of the reference sizes of the λ DNA digested with HindIII are given between the two panels [event MLS9878]

The results are as illustrated in Figure 6, wherein the legends for both panels are provided as below:

1	λ / <i>HindIII</i> Marker
2	Positive control (Probe)
3	Cotton Event 9878 – undigested
4	Cotton Event 9878 / <i>EcoRI</i>
5	Cotton Event 9878 / <i>HindIII</i>
6	Negative control (Non-transgenic)/ <i>EcoRI</i>

Table: 5 - Data of samples loaded in the gel (figure 6)

The two transgenic events were further analyzed by PCR for the presence of the unique region as described in SEQ ID No. 6 between the 2 expression cassettes. The primer sequences are described in SEQ ID No. 15 (forward primer) and SEQ ID No. 16 (reverse primer). The expected amplicon size was 975 bp.

Diagrammatic representation of diagnostic PCR process for the amplification of the left and right border junction of the transferred DNA [T-DNA] within the transgenic cotton events MLS9878 are illustrated in figures 7 and 8 respectively.

Transgenic cotton events not harbouring the T-DNA of the plasmid pMH82 were also subjected to this PCR and these samples did not amplify.

The results are as illustrated in Figure 9, wherein the legend for the gel loaded with different samples is provided as below:

1	Negative control – water
2	Negative control - Non transgenic cotton genomic DNA
3	Positive control - pMH82
4	1Kb Gene Ruler (MBI Fermentas)
5	MLS9878
6	MLS9124
7	Transgenic Cotton not transformed with pMH82 – 1
8	Transgenic Cotton not transformed with pMH82 – 2
9	Transgenic Cotton not transformed with pMH82 – 3

Table: 6 - Data of samples loaded in the gel (figure 9)

The 975 bp amplicon from both the events MLS9124 and MLS9878 is sequenced and is set forth as SEQ ID NO. 17 which shows the presence of the unique 57 bp region described in SEQ ID No. 6.

Junction Analysis

In an embodiment of the present disclosure, junction analysis studies for both the transgenic cotton events are carried out at a lab scale.

The left and right border junction sequence of the transgenic cotton event MLS9878 are set forth as SEQ ID Nos. 9 and 10, respectively.

T-DNA specific primer amplifying the left border junction sequence of the transgenic cotton events MLS9878 is set forth as SEQ ID No. 11, whereas the T-DNA specific primer amplifying the right border junction sequence of the transgenic cotton events MLS9878 is set forth as SEQ ID No. 13.

The corresponding cotton genome specific primer amplifying the left border junction sequence of the transgenic cotton event MLS9878 is set forth as SEQ ID No. 12, whereas the cotton genome specific primer amplifying the right border junction sequence of the transgenic cotton event MLS9878 is set forth as SEQ ID No. 14.

The right border junction sequence of the transgenic cotton event MLS9124 is set forth as SEQ ID No. 18. The corresponding T-DNA specific primer amplifying the right border junction sequence of the transgenic cotton events MLS9124 is set forth as SEQ ID No. 19 and the cotton genome specific primer amplifying the right border junction sequence of the transgenic cotton event MLS9124 is set forth as SEQ ID No. 20.

Diagrammatic representation of diagnostic PCR process for the amplification of the right border junction of transgenic cotton events MLS9124 is illustrated in figure 11, wherein the legend for the gel loaded with different samples is provided as below:

M	1 Kb ladder
A	pMH82
B	Non-transgenic cotton
C	Cotton Cry1CGh Event MLS9124
D	Water (PCR control)

Table: 7 - Data of samples loaded in the gel (figure 11)

Spodoptera litura and *Helicoverpa armigera* Bioassays:

Transgenic cotton events MLS9124 and MLS9878 were assayed for resistance to *S. litura* and *Helicoverpa armigera* up to the T4 generation and mortality up to 100% was observed to be carried through generations.

First fully expanded terminal leaves from unsprayed plots expressing various transgenes are bioassayed for bioactivity against *Helicoverpa armigera* and *Spodoptera litura* larvae. These bioassays were carried out at time intervals of 64, 78, 92, 106, 120 and 134 days after planting, in Non transgenic event (control), Event MLS9124 and Event MLS9878.

Length of the sampled leaves is measured approximately 5 cm across and is removed with as much petiole attached as possible. Leaves from each replicate are kept separate in labeled paper sacks and placed in plastic cooler containing cooling packs for transport to the laboratory.

Individual leaves are placed in a Petri dish having wet whatman filter paper and infested with a single day old fed larva (10 dishes/plot) for a total of 40 larvae/plot. Bioassays are kept at a constant temperature (25-30°C and relative humidity greater than 50 %) prior to assessment. After five days after exposure (DAE), larvae are prodded with a camel-hair brush and considered alive if coordinated movement is observed. Larvae from each replicate are scored according to the criteria in Table 8:

Table 8: Scoring criteria for bioassays

Condition	Score
Dead L0	0 25
Alive L1	1
Alive L2	2
Alive L3	3

Explanation for scoring:

L0 – Dead larvae

L1 – larvae with morbid growth

L2 – approximately 50% growth of L3

L3 – identical to control larvae (normal and healthy)

Percent survival and development of larvae is analyzed using analysis of variance (ANOVA). Insect bioassay results of the two transgenic cotton events MLS9124 and MLS9878 for efficacy against *Spodoptera litura* are summarized in Figure 10 (A) and 10 (B). The data is captured at different stages of the growth cycle of the transgenic cotton events. In an embodiment of the present disclosure, the bar diagram for each event per insect has one bar for the average worm size. This is calculated as follows:

Each category of survival score is assigned a factor which is designated as-

L1 – 0.5

L2 – 1

L3 – 2

Further, $(\text{number of L1} \times 0.5) + (\text{number of L2} \times 1) + (\text{number of L3} \times 2) = \text{total worm size}$.

$\text{Total worm size} / \text{Total number of survived larvae} = \text{Average worm size}$

In an embodiment, the total number of survived larvae is 40, wherein L1 is 20, L2 is 11 and L3 is 9.

Therefore the total worm size is $(20 \times 0.5) + (11 \times 1) + (9 \times 2) = 39$ and the average worm size is-

$\text{Total worm size} / \text{Total number of survived larvae} = 39/40 = 0.975$.

In an exemplary embodiment, the average worm size for constructing a bar diagram is calculated as a percentage of the L3 factor, i.e., 2. Therefore, if the value of average worm size is found to be 0.975, it is represented in the bar diagram as $(0.975/2) \times 100 = 48.75\%$.

Earias vitella Bioassays:

Squares (flower buds) of approximately 4-6 mm thickness from unsprayed plots were bioassayed for activity against *Earias vittella* larvae.

Sampled squares are removed with as much petiole attached as possible. Squares from each replicate are kept separate in labeled paper sacks and placed in plastic cooler containing cooling packs for transport to the laboratory.

Individual squares are placed into a Petri dish having wet whatman filter paper and infested with a two day old fed larva for a total of 40 larvae/trait. Bioassays are kept at a constant temperature (25-30°C and relative humidity greater than 50%) prior to assessment. After five days after exposure (DAE), larvae are prodded with a camel-hair brush and considered alive if coordinated movement is observed.

Insect bioassay results of the two transgenic cotton events MLS9124 and MLS9878 for efficacy against *Earias vitella* are summarized in Tables 9 and 10 respectively.

Table-9: *Earias vitella* bioassay with respect to cotton event MLS9124

Insect	1st assay	2nd assay	3rd assay	Average Mortality (%)
<i>Earias vitella</i> assay	4/4	4/4	4/4	100

Table-10: *Earias vitella* bioassay with respect to cotton event MLS9878

Insect	1st assay	2nd assay	Average Mortality (%)
<i>Earias vitella</i> assay	16/16	16/16	100

Protein expression data:

The expression of the Cry1C protein as estimated by quantitative ELISA was in the range 0.41 to 4.67 µg/gm fresh weight of leaf. This is explained in Table 11 below.

	MLS9124		MLS9878	
	Average	Range	Average	Range
	µg protein per gram fresh weight			
Leaf	1.41	1.00 –	1.88	0.41 –

		1.93		4.67
Seed		-	9.4	-

Table: 11 – ELISA protein expression data.

REFERENCES:

1. <http://www.inchem.org/documents/ehc/ehc/ehc217.htm>
2. Bt-commercial products – www.sipweb.org/bacteria.htm
3. Chen et al (2008) PLoS ONE, 3(5): e 2284.
4. Honee et al, (1988) Nucl. Acids Res., 16: 6240.
5. Sanchis et al, (1989) Mol. Microbiol., 3: 229-238.
6. US Patent 5246852.
7. European Patent 0400246.
8. Strizhov, (1996) Unpublished (Genbank Accession Number X96682).
9. Kalwan et al, (1993) Appl. Environ. Microbiol., 59:1131-1137.
10. Perlak et al (2001) Plant J., 27:489-501.
11. Gore et al (2001) J. Econ. Entomol., 94, 1445-1451.
12. Wan et al (2008) Environ. Entomol., 37: 1043-1048.
13. Cao et al (2003) J. New Seeds, 5:193-207.
14. Strizhov, et al (1996) Proc. Natl. Acad. Sci. USA, 93:15012-15017.
15. Valicenta et al (2005) Proceedings of the 6th Pacific Rim Conference on the Biotechnology of *Bacillus thuringiensis* and its Environmental Impact (Victoria BC, 2005)
16. Tang et al (2006) Mol. Breeding, 18:1-10.
17. US Patent 5914318
18. US Patent 5942664
19. US Patent 6033874
20. US Patent 6153814
21. US Patent 6177615
22. US Patent 6313378
23. US Patent 6423828
24. US Patent 6809078
25. US Patent 6825006
26. US Patent 7256017
27. Codon Usage Database (<http://www.kazusa.or.jp/codon/>)

28. An, G. (1986) Plant Physiol., 81:86-91.
29. Hajdukiewicz, P., Svab, Z., Maliga, P. (1994) Plant Mol. Biol., 25:989-994.
30. Genbank accession number CAA30396.
31. Sambrook, J., Russell, D.W. (2001) Molecular Cloning – a laboratory manual.
3rd Edition. CSHL Press, Cold Spring Harbor, New York, USA.
32. Sessions et al (2002) Plant Cell, 14:2985-2994.

SEQUENCE LISTING

```

<110>  METAHELIX LIFESCIENCES PRIVATE LIMITED
<120>  A NUCLEOTIDE SEQUENCE, EXPRESSION CASSETTE TRANSGENIC EVENTS,
      CELLS AND METHODS THEREOF
<130>  IP16850
<160>  20
<170>  PatentIn version 3.5
<210>  1
<211>  1893
<212>  DNA
<213>  Bacillus thuringiensis

<220>
<221>  gene
<222>  (1)..(1893)

<220>
<221>  gene
<222>  (1)..(1893)
<223>  Synthetic Sequence

<400>  1
atggaagaga acaatcaaaa ccagtgcaatt ccatacaact gcctcagtaa tcctgaagag      60
gttcttctcg atggtgaaag gatctccacc ggaaacagtt ccattgacat tagtctttcc      120
ctcgtgcaat tccttgtag taacttcgtc cccggtggtg gattcctcgt tggctttatc      180
gattttgtgt ggggtattgt tggaccatcc cagtgggatg ctttcctcgt ccaaattgag      240
cagcttatca atgagcgtat tgccgagttc gctaggaacg ctgccattgc taacctcgaa      300
ggctttggta ataacttcaa catctatggt gaggccttta aggaatggga ggaagaccct      360
aataaccccg agactcgtac caggggtgatt gatcgtttca ggattctcga tggacttctc      420
gaacgtgaca tcccaagttt caggatttcc ggtttcgagg ttctcttctt cagtgtctac      480
gccccagctg ctaaccttca cctcgccatt cttcgtgatt ccgttatctt tggtgaaagg      540
tggggactca ccactattaa tgtgaacgag aactataatc gtcttattag gcacatcgat      600
gaatacgtg accactgcgc taacacctat aaccgtgggc tcaataacct tcccaaaagt      660
acctaccagg attggattac ttataacagg ctccgtaggg atcttaccct caccgttctt      720
gacattgccg ctttcttccc aaattacgat aaccgtaggt atcctatcca accgctcggt      780
cagctcactc gtgagggttta caccgatcca ctattaact tcaatcctca actccagtcc      840
gtggctcaac ttccacctt taacgttatg gaaagttcca ggattcgtaa cccacacctc      900
ttcgacatcc ttaataacct cactattttc accgattggt tcagtgtcgg aagggaacttt      960
tattggggtg gtcaccgtgt tatttccagt cttatcggag gtggtaacat tacctcccct      1020

```

```

atctacggaa gggaggccaa ccaggaaccc ccacgtagtt tcactttcaa cggtcctgtg      1080
ttcaggaccc tctccaatcc cacccttcgt ctccttcaac agccttggcc tgctcccca      1140
tttaacctca ggggtgttga gggagtcgag ttcagtactc ctaccaactc cttcacctat      1200
cgtggtaggg gtactgttga tagtcttacc gagctccccc cagaagacaa ttccgtgcct      1260
ccccgtgagg gatacagtca caggctttgc cacgctacct tcgttcaacg ttccggtact      1320
ccattttctca ccaccggtgt cgttttcagt tggactgata ggtccgccac cttaccaac      1380
actatcgatc ctgaacgtat taaccagatt cccctcgtga agggattcag ggtttgggg      1440
ggtaccagtg tcatcaccgg accaggtttc actggtggag acattcttcg taggaatacc      1500
tttggtgatt tcgtttccct ccaagtgaac attaacagtc ctatcaccca gcgttatagg      1560
cttcgtttca ggtacgcttc cagtcgtgat gctagggtta ttgtcctcac tggtgccgct      1620
tccaccggag ttggtggtca agtgagtgtt aatatgcccc ttcagaagac tatggagatt      1680
ggagaaaacc tcacttcccg taccttcagg tataccgact ttagtaaccc attctccttc      1740
cgtgctaatc ctgatatcat tggatttagt gagcaacccc ttttcggtgc cggatccatc      1800
agttccggtg aactctacat tgataagatt gagatcattc ttgctgacgc tacttttgaa      1860
gccgagagtg atctcgaaag ggctcagaag tga      1893

```

```

<210> 2
<211> 630
<212> PRT
<213> Bacillus thuringiensis

```

```

<220>
<221> PEPTIDE
<222> (1)..(630)

```

```

<400> 2

```

```

Met Glu Glu Asn Asn Gln Asn Gln Cys Ile Pro Tyr Asn Cys Leu Ser
 1              5              10              15

Asn Pro Glu Glu Val Leu Leu Asp Gly Glu Arg Ile Ser Thr Gly Asn
          20              25              30

Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Phe Leu Val Ser Asn
      35              40              45

Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val Trp
      50              55              60

Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile Glu
65              70              75              80

Gln Leu Ile Asn Glu Arg Ile Ala Glu Phe Ala Arg Asn Ala Ala Ile
          85              90              95

Ala Asn Leu Glu Gly Leu Gly Asn Asn Phe Asn Ile Tyr Val Glu Ala
          100              105              110

Phe Lys Glu Trp Glu Glu Asp Pro Asn Asn Pro Glu Thr Arg Thr Arg
      115              120              125

Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp Ile
      130              135              140

```

Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val Tyr
 145 150 155 160
 Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val Ile
 165 170 175
 Phe Gly Glu Arg Trp Gly Leu Thr Thr Ile Asn Val Asn Glu Asn Tyr
 180 185 190
 Asn Arg Leu Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala Asn
 195 200 205
 Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln Asp
 210 215 220
 Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val Leu
 225 230 235 240
 Asp Ile Ala Ala Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro Ile
 245 250 255
 Gln Pro Val Gly Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Leu Ile
 260 265 270
 Asn Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe Asn
 275 280 285
 Val Met Glu Ser Ser Arg Ile Arg Asn Pro His Leu Phe Asp Ile Leu
 290 295 300
 Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn Phe
 305 310 315 320
 Tyr Trp Gly Gly His Arg Val Ile Ser Ser Leu Ile Gly Gly Gly Asn
 325 330 335
 Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro Arg
 340 345 350
 Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Asn Pro Thr
 355 360 365
 Leu Arg Leu Leu Gln Gln Pro Trp Pro Ala Pro Pro Phe Asn Leu Arg
 370 375 380
 Gly Val Glu Gly Val Glu Phe Ser Thr Pro Thr Asn Ser Phe Thr Tyr
 385 390 395 400
 Arg Gly Arg Gly Thr Val Asp Ser Leu Thr Glu Leu Pro Pro Glu Asp
 405 410 415
 Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His Ala
 420 425 430
 Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Val Val
 435 440 445

Phe Ser Trp Thr Asp Arg Ser Ala Thr Leu Thr Asn Thr Ile Asp Pro
 450 455 460
 Glu Arg Ile Asn Gln Ile Pro Leu Val Lys Gly Phe Arg Val Trp Gly
 465 470 475 480
 Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu
 485 490 495
 Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln Val Asn Ile Asn
 500 505 510
 Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg Tyr Ala Ser Ser
 515 520 525
 Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala Ser Thr Gly Val
 530 535 540
 Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys Thr Met Glu Ile
 545 550 555 560
 Gly Glu Asn Leu Thr Ser Arg Thr Phe Arg Tyr Thr Asp Phe Ser Asn
 565 570 575
 Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly Ile Ser Glu Gln
 580 585 590
 Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu Leu Tyr Ile Asp
 595 600 605
 Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu Ala Glu Ser Asp
 610 615 620
 Leu Glu Arg Ala Gln Lys
 625 630

<210> 3
 <211> 2926
 <212> DNA
 <213> *Bacillus thuringiensis*

<220>
 <221> gene
 <222> (1)..(2926)
 <223> Synthetic Sequence

<400> 3
 atggtggagc acgacactct cgtctactcc aagaatatca aagatacagt ctcagaagac 60
 caaagggcta ttgagacttt tcaacaaagg gtaatatcgg gaaacctcct cggattccat 120
 tgcccagcta tctgtcactt catcaaaagg acagtagaaa aggaagggtg cacctacaaa 180
 tgccatcatt gcgataaagg aaaggctatc gttcaagatg cctctgccga cagtgggtccc 240
 aaagatggac cccacccac gaggagcatc gtggaaaaag aagacgttcc aaccacgtct 300
 tcaaagcaag tggattgatg tgataacatg gtggagcacg acactctcgt ctactccaag 360
 aatatcaaag atacagtctc agaagaccaa agggctattg agacttttca acaaagggta 420
 atatcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat caaaaggaca 480

gtagaaaagg aaggtggcac ctacaaatgc catcattgcg ataaaggaaa ggctatcggt	540
caagatgcct ctgccgacag tgggtccaaa gatggacccc caccacgag gagcatcggt	600
gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga tatctccact	660
gacgtaaggg atgacgcaca atcccactat ccttcgcaag accttctct atataaggaa	720
gttcatttca tttggagagg acacgctgaa atcaccagtc tctctctaca aatctatctc	780
tctcgagatc tatcgatgca tgccatggaa gagaacaatc aaaaccagtg cattccatac	840
aactgcctca gtaatcctga agaggttctt ctcgatgggt aaaggatctc caccggaac	900
agttccattg acattagtct ttccctcggt caattccttg ttagtaactt cgtccccggt	960
gggtgattcc tcgttgggtc tatcgatttt gtgtggggta ttgttgacc atcccagtg	1020
gatgctttcc tcgtccaaat tgagcagctt atcaatgagc gtattgccga gttcgctagg	1080
aacgctgcca ttgctaacct cgaaggtctt ggtaataact tcaacatcta tgttgaggct	1140
tttaaggaat gggaggaaga ccctaataac cccgagactc gtaccagggt gattgatcgt	1200
ttcaggattc tcgatggact tctcgaacgt gacatcccaa gtttcaggat ttccggtttc	1260
gaggttctc ttctcagtg ctacgccccaa gctgctaacc ttcacctgc cattcttctg	1320
gattccgtta tctttggtga aaggtgggga ctcaccacta ttaatgtgaa cgagaactat	1380
aatcgtctta ttaggcacat cgatgaatac gctgaccact gcgctaacac ctataaccgt	1440
ggcttcaata accttcccaa aagtacctac caggattgga ttacttataa caggctcgt	1500
agggatctta cctcaccgt tcttgacatt gccgctttct tcccaaatta cgataaccgt	1560
aggtatccta tccaaccgt cggtcagctc actcgtgagg ttacaccga tccacttatt	1620
aacttcaatc ctcaactcca gtccgtggct caacttccca cctttaacgt tatggaaagt	1680
tccaggattc gtaaccaca cctcttcgac atccttaata acctactat tttcaccgat	1740
tggttcagt tcggaaggaa cttttattgg ggtgggtacc gtgttatttc cagtcttctc	1800
ggaggtggta acattacct ccctatctac ggaagggagg ccaaccagga acccccacgt	1860
agtttactt tcaacgggtc tgtgttcagg acctctcca atcccacct tcgtctctt	1920
caacagcctt ggcttctcc cccatttaac ctcaggggtg ttgagggagt cgagttcagt	1980
actcctacca actccttcac ctatcggtg aggggtactg ttgatagtct taccgagctc	2040
ccccagaag acaattccgt gcctccccgt gagggataca gtcacaggct ttgccacgt	2100
accttcttc aacgttccgg tactccattt ctcaccaccg gtgtcgtttt cagttggact	2160
gataggtccg ccaccttac caacactatc gatcctgaac gtattaacca gattccccctc	2220
gtgaagggat tcagggtttg ggggtgtacc agtgtcatca ccggaccagg tttcactggt	2280
ggagacattc ttcgtaggaa tacctttggt gatttcgttt ccctccaagt gaacattaac	2340
agtcctatca cccagcgtta taggcttctg ttcaggtagc cttccagtcg tgatgctagg	2400
gttattgtcc tactggtgc cgcttcacc ggagttggtg gtcaagttag tgtaatatg	2460
ccccctcaga agactatgga gattggagaa aacctactt cccgtacctt caggatatac	2520
gacttttagta acccattctc cttccgtgct aatcctgata tcattggtat tagtgagcaa	2580
cccccttctg gtgccggatc catcagttcc ggtgaactct acattgataa gattgagatc	2640
attcttgctg acgctacttt tgaagccgag agtgatctcg aaagggctca gaagtgacct	2700
gggagctcga attcgaagct tctgcagacg cgtccctcga gtttctccat aataatgtgt	2760
gagtagttcc cagataaggg aattaggggt cctatagggt ttcgctcatg tgttgagcat	2820

ataagaaacc cttagtatgt atttgtatatt gtaaaatact tctatcaata aaattttctaa 2880
 ttctctaaaac caaaatccag tactaaaatc cagatccccc gaatta 2926

<210> 4
 <211> 9449
 <212> DNA
 <213> *Bacillus thuringiensis*

<220>
 <221> gene
 <222> (1)..(9449)
 <223> Synthetic Sequence

<400> 4
 agtacttttga tccaacccct ccgctgctat agtgcagtcg gcttctgacg ttcagtgcag 60
 ccgtctttctg aaaacgacat gtcgcacaag tcctaagtta cgcgacaggc tgccgccctg 120
 ccttttttctt ggcgtttttct tgtcgcgtgt tttagtcgca taaagtagaa tacttgcgac 180
 tagaaccgga gacattacgc catgaacaag agcgccgccg ctggcctgct gggctatgcc 240
 cgcgtcagca ccgacgacca ggacttgacc aaccaacggg ccgaactgca cgcggccggc 300
 tgcaccaagc tgttttccga gaagatcacc ggcaccaggc gcgaccgcc ggagctggcc 360
 aggatgcttg accacctacg ccctggcgac gttgtgacag tgaccaggct agaccgcctg 420
 gcccgcagca cccgcgacct actggacatt gccgagcgca tccaggaggc cggcgcgggc 480
 ctgcgtagcc tggcagagcc gtgggcccgc accaccacgc cggccggccg catggtgttg 540
 accgtgttcg ccggcattgc cgagttcgag cgttccctaa tcatcgaccg cacccgagc 600
 gggcgcgagg ccgccaaggc ccgaggcggtg aagtttgccc cccgccctac cctcaccgcg 660
 gcacagatcg cgcacgcccg cgagctgacg gaccaggaag gccgcaccgt gaaagaggcg 720
 gctgcactgc ttggcgtgca tcgctcgacc ctgtaccgcg cacttgagcg cagcgaggaa 780
 gtgacgcca ccgaggccag gcggcgcggt gccttcctg aggacgcatt gaccgaggcc 840
 gacgccctgg cggccgcca gaataaacgc caagaggaac aagcatgaaa ccgcaccagg 900
 acggccagga cgaaccgttt ttcattaccg aagagatcga ggcgagatg atcgcgccg 960
 ggtacgtgtt cgagccgccc gcgcacgtct caaccgtgcg gctgcatgaa atcctggccg 1020
 gtttgtctga tgccaagctg gcggcctggc cggccagctt ggccgctgaa gaaaccgagc 1080
 gccgccgtct aaaaaggtga tgtgtatttg agtaaaacag cttgcgtcat gcggtcgtg 1140
 cgtatatgat gcgatgagta aataaacaac tacgcaaggg gaacgcatga aggttatcgc 1200
 tgtacttaac cagaaaggcg ggtcaggcaa gacgaccatc gcaaccatc tagcccgcg 1260
 cctgcaactc gccggggccg atgttctgtt agtcgattcc gatccccagg gcagtgcgcg 1320
 cgattgggcg gccgtgcggg aagatcaacc gctaaccgtt gtcggcatcg accgcccagc 1380
 gattgaccgc gacgtgaagg ccatcggccg gcgcgacttc gtagtgatcg acggagcgcc 1440
 ccaggcgcg gacttggtg tgtcccgat caaggcagcc gacttcgtgc tgattccggt 1500
 gcagccaagc ccttacgaca tatgggccac gccgcacctg gtggagctgg ttaagcagcg 1560
 cattgaggtc acggatggaa ggctacaagc ggctttgtc gtgtcgcggg cgatcaaagg 1620
 cacgcgcac gccgggtgagg ttgccgaggc gctggccggg tacgagctgc ccattcttga 1680
 gtcccgatc acgcagcgcg tgagctaccc aggcactgcc gccgccggca caaccgttct 1740
 tgaatcagaa cccgaggggc acgctgcccg cgagggtccg gcgctggccg ctgaaattaa 1800

atcaaaactc	atttgagtta	atgaggtaaa	gagaaaatga	gcaaaagcac	aaacacgcta	1860
agtgccggcc	gtccgagcgc	acgcagcagc	aaggctgcaa	cgttggccag	cctggcagac	1920
acgccagcca	tgaagcgggt	caactttcag	ttgccggcgg	aggatcacac	caagctgaag	1980
atgtacgcgg	tacgccaagg	caagaccatt	accgagctgc	tatctgaata	catcgcgcag	2040
ctaccagagt	aatgagcaa	atgaataaat	gagtagatga	atttttagcgg	ctaaaggagg	2100
cggcatggaa	aatcaagaac	aaccaggcac	cgacgccgtg	gaatgcccc	tgtgtggagg	2160
aacgggcggg	tggccaggcg	taagcggctg	ggttgtctgc	cggccctgca	atggcactgg	2220
aacccccaa	cccaggaat	cggcgtgacg	gtcgcaaac	atccggcccc	gtacaaatcg	2280
gcgcggcgct	gggtgatgac	ctggtggaga	agttgaaggc	cgcgaggcc	gccagcggc	2340
aacgcatcga	ggcagaagca	cgccccgggt	aatcgtggca	agcggccgct	gatcgaatcc	2400
gcaaagaatc	ccggcaaccg	ccggcagccg	gtgcgccgtc	gattaggaag	ccgccccagg	2460
gcgacgagca	accagatfff	ttcgttccga	tgctctatga	cgtgggcacc	cgcatagtc	2520
gcagcatcat	ggacgtggcc	gttttccgtc	tgctgaagcg	tgaccgacga	gctggcgagg	2580
tgatccgcta	cgagcttcca	gacgggcacg	tagaggtttc	cgaggggccg	gccggcatgg	2640
ccagtgtgtg	ggattacgac	ctggtactga	tggcggtttc	ccatctaacc	gaatccatga	2700
accgataccg	ggaaggggaag	ggagacaagc	ccggccgcgt	gttccgtcca	cacgttgccg	2760
acgtactcaa	gttctgccgg	cgagccgatg	gcggaaagca	gaaagacgac	ctggtagaaa	2820
cctgcattcg	gttaaacacc	acgcacgttg	ccatgcagcg	tacgaagaag	gccaagaacg	2880
gccgcctggg	gacggtatcc	gaggggtgaag	ccttgattag	ccgctacaag	atcgtaaaga	2940
gcgaaaccgg	gcggccggag	tacatcgaga	tcgagctagc	tgattggatg	taccgcgaga	3000
tcacagaagg	caagaaccgg	gacgtgctga	cggttcaccc	cgattacttt	ttgatcgatc	3060
ccggcatcgg	ccgttttctc	taccgcctgg	cacgccgcgc	cgaggcaag	gcagaagcca	3120
gatggttgtt	caagacgatc	tacgaacgca	gtggcagcgc	cggagagttc	aagaagttct	3180
gtttcacctg	gcgcaagctg	atcgggtcaa	atgacctgcc	ggagtacgat	ttgaaggagg	3240
aggcggggca	ggctggcccc	atcctagtca	tgcgctaccg	caacctgatc	gagggcgaag	3300
catccgccgg	ttcctaattg	acggagcaga	tgctagggca	aattgcccta	gcaggggaaa	3360
aaggtcgaaa	aggtctcttt	cctgtggata	gcacgtacat	tgggaacca	aagccgtaca	3420
ttgggaaccg	gaacccgtac	attgggaacc	caaagccgta	cattgggaac	cggtcacaca	3480
tgtaagtga	tgatataaaa	gagaaaaaag	gcgatttttc	cgctaaaaac	tctttaaacc	3540
ttattaaaac	tcttaaaacc	cgcctggcct	gtgcataact	gtctggccag	cgcacagccg	3600
aagagctgca	aaaagcgcct	acccttcggg	cgctgcgtc	cctacgcccc	gccgcttcgc	3660
gtcggcctat	cgcgcccgct	ggccgctcaa	aaatggctgg	cctacggcca	ggcaatctac	3720
cagggcgcgg	acaagcccg	ccgtcgccac	tcgaccggcg	gcgcccacat	caaggcaccc	3780
tgctcgcgc	gtttcgggtg	tgacggtgaa	aacctctgac	acatgcagct	cccggagacg	3840
gtcacagctt	gtctgtaagc	ggatgccggg	agcagacaag	cccgtcaggg	cgcgtcagcg	3900
ggtgttggcg	ggtgtcgggg	cgagccatg	acccagtcac	gtagcgatag	cggagtgtat	3960
actggcttaa	ctatgcggca	tcagagcaga	ttgtactgag	agtgacccat	atgcggtgtg	4020
aaataaccgca	cagatgcgta	aggagaaaa	accgcatcag	gcgctcttcc	gcttcctcgc	4080

tcactgactc	gctgcgctcg	gtcgttcggc	tgcggcgagc	ggtatcagct	caactcaaagg	4140
cggtataacg	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaaag	4200
gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	4260
gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aacccgacag	4320
gactataaag	ataccaggcg	tttccccctg	gaagctccct	cgtagcgctct	cctgtttcga	4380
ccctgccgct	taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	4440
atagctcacg	ctgtaggtat	ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	4500
tgcacgaacc	ccccgttcag	cccgaccgct	gcgccttctc	cggtaaactat	cgctttgagt	4560
ccaacccggt	aagacacgac	ttatcgccac	tggcagcagc	caactggtaac	aggattagca	4620
gagcgaggta	tgtaggcggg	gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	4680
ctagaaggac	agtatttggt	atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	4740
ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	cggtgggtttt	tttgtttgca	4800
agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	ttttctacgg	4860
ggtctgacgc	tcagtggaaac	gaaaactcac	gttaagggat	tttgggtcatg	catgatatat	4920
ctcccaattt	gtgtagggct	tattatgcac	gcttaaaaaat	aataaaaagca	gacttgacct	4980
gatagtttgg	ctgtgagcaa	ttatgtgctt	agtgcatcta	atcgcttgag	ttaacgccgg	5040
cgaagcggcg	tcggcttgaa	cgaatttcta	gctagacatt	atttgccgac	taccttggtg	5100
atctcgcctt	tcacgtagtg	gacaaattct	tccaactgat	ctgcgcgcga	ggccaagcga	5160
tcttcttctt	gtccaagata	agcctgtcta	gcttcaagta	tgacgggctg	atactgggcc	5220
ggcaggcgct	ccattgccca	gtcggcagcg	acatccttcg	gcgcgatttt	gccggttact	5280
gcgctgtacc	aaatgcggga	caacgtaagc	actacatttc	gctcatcgcc	agcccagtcg	5340
ggcggcgagt	tccatagcgt	taagggttca	tttagcgcct	caaataagatc	ctgttcagga	5400
accggatcaa	agagttcctc	cgccgctgga	cctaccaagg	caacgctatg	ttctcttgct	5460
tttgtcagca	agatagccag	atcaatgtcg	atcgtggctg	gctcgaagat	acctgcaaga	5520
atgtcattgc	gctgccattc	tccaaattgc	agttcgcgct	tagctggata	acgccacgga	5580
atgatgtcgt	cgtagcacaac	aatggtgact	tctacagcgc	ggagaatctc	gctctctcca	5640
ggggaagccg	aagtttccaa	aaggctggtg	atcaaagctc	gccgcgttgt	ttcatcaagc	5700
cttacgggtca	ccgtaaccag	caaatacaata	tactgtgtg	gcttcaggcc	gccatccact	5760
gcggagccgt	acaaatgtac	ggccagcaac	gtcggttcga	gatggcgctc	gatgacgcca	5820
actacctctg	atagttgagt	cgatacttcg	gcgatcaccg	cttcccccat	gatgtttaac	5880
tttgttttag	ggcgactgcc	ctgctgcgta	acatcgttgc	tgctccataa	catcaaakat	5940
cgaccacagg	cgtaacgcgc	ttgctgcttg	gatgcccag	gcatagactg	tacccccaaa	6000
aaacatgtca	taacaagaag	ccatgaaaac	cgccactgcg	ccgttaccac	cgctgcgttc	6060
ggtcaagggt	ctggaccagt	tgcgtgacgg	cagttacgct	acttgcatga	cagcttacga	6120
accgaacgag	gcttatgtcc	actgggttcg	tgcccgaatt	gatcacaggc	agcaacgctc	6180
tgtagctggt	acaatcaaca	tgctaccctc	cgcgagatca	tccgtgtttc	aaacccggca	6240
gcttagttgc	cgttcttccg	aatagcatcg	gtaacatgag	caaagtctgc	cgctttacaa	6300
cggctctccc	gctgacgccg	tcccggactg	atgggctgcc	tgtatcgagt	ggtgattttg	6360
tgccgagctg	ccggtcgggg	agctgttggc	tggctggtgg	caggatatat	tgtggtgtaa	6420

acaaattgac gcttagacaa cttataaca cattgcggac gtttttaatg tactgaatta	6480
acgccgaatt gctctagcca atacgcaaac cgctctccc cgcgcgttgg ccgattcatt	6540
aatgcagctg gcacgacagg tttcccgact ggaaagcggg cagtgcgcgc aacgcaatta	6600
atgtgagtta gctcactcat taggcacccc aggcctttaca ctttatgctt ccggctcgta	6660
tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aacagctat gaccatgatt	6720
acgaattcga gctcccaccg cgggtttctg gagtttaatg agctaagcac atacgtcaga	6780
aaccattatt gcgcgttcaa aagtcgccta aggtcactat cagctagcaa atatttcttg	6840
tcaaaaatgc tccactgacg taccataaat tccccctcgg atccaattag agtctcatat	6900
tcactctcaa tccaaataat ctgcaatggc aattacctta tccgcaactt ctttacctat	6960
ttccgcccgg atccgggcag gttctccggc cgcttgggtg gagaggctat tcggctatga	7020
ctgggcacaa cagacaatcg gctgctctga tgccgcgtg ttccggctgt cagcgcaggg	7080
gcgcccgggt ctttttgtca agaccgacct gtccggtgcc ctgaatgaac tgcaggacga	7140
ggcagcgcgg ctatcgtagc tggccacgac gggcggttct tgccgagctg tgctcgacgt	7200
tgtcactgaa gcgggaaggg actggctgct attgggcgaa gtgcccgggc aggatctcct	7260
gtcatctcac cttgctcctg ccgagaaagt atccatcatg gctgatgcaa tgcggcggct	7320
gcatacgctt gatccggcta cctgcccatt cgaccaccaa gcgaaacatc gcatcgagcg	7380
agcacgtact cggatggaag ccggtcttgt cgatcaggat gatctggacg aagagcatca	7440
ggggctcgcg ccagccgaac tgctcgccag gctcaaggcg cgcagcccg acggcgagga	7500
tctcgtcgtg acccatggcg atgcctgctt gccgaatatc atggtggaaa atggccgctt	7560
ttctggattc atcgactgtg gccggctggg tgtggcggac cgctatcagg acatagcggt	7620
ggctaccggt gatattgctg aagagcttgg cggcgaatgg gctgaccgct tcctcgtgct	7680
ttacggtatc gccgctcccg attcgcagcg catcgccttc tatcgccttc ttgacgagtt	7740
cttctgagcg ggactctggg gttcgaaatg accgaccaag cgacgcccaa cctgccatca	7800
cgagatttcg attccaccgc cgcttctat gaaagggttg gcttcggaat cgttttccgg	7860
gacgccggct ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcaccacc	7920
gatccaacac ttacgtttgc aacgtccaag agcaaataga ccacgaacgc cggaagggtg	7980
ccgcagcgtg tggattgctg ctcaattctc tcttgaggga atgcaatgat gaatatgata	8040
ctgactatga aactttgagg gaatactgcc tagcaccgtc acctcataac gtgcatcatg	8100
catgccctga caacatggaa catcgctatt tttctgaaga attatgctcg ttggaggatg	8160
tcgaggcaat tgcagctatt gccaaaatcg aaataaccct cacgcatgca ttcataata	8220
ttattcatgc ggggaaaggc aagattaatc caactggcaa atcatccagc gtgattggta	8280
acttcagttc cagcgacttg attcgttttg gtgctacca cgttttcaat aaggacgaga	8340
tgggtggagta aagaaggagt gcgtcgaagc agatcgttca aacatttggc aataaagttt	8400
cttaagattg aatcctgttg ccggctctgc gatgattatc atataatttc tgttgaaata	8460
cgtaagcat gtaataatta acatgtaatg catgacgtta tttatgagat gggtttttat	8520
gattagagtc ccgcaattat acatttaata cgcgatagaa aacaaaatat agcgcgcaaa	8580
ctaggataaa ttatcgcgcg cgggtgtcatc tatgttacta gatcgatcaa acttcggtac	8640
tgtgtaatga cgatgagcaa tcgagaggct gactaacaaa aggtatgccc aaaaacaacc	8700

tctccaaact	gtttcgaatt	ggaagtttct	gctcatgccg	acaggcataa	cttagatatt	8760
cgcgggctat	tcccactaat	tcgtcctgct	ggtttgccg	aagataaatc	agtgcattct	8820
cttacaagtt	cctctgtctt	gtgaaatgaa	ctgctgactg	ccccccaaga	aagcctcctc	8880
atctcccagt	tggcggcggc	tgatacacca	tcgaaaaccc	acgtccgaac	acttgataca	8940
tgtgcctgag	aaataggcct	acgtccaaga	gcaagtcctt	tctgtgctcg	tcggaaattc	9000
ctctcctgtc	agacggtcgt	gcgcatgtct	tcggttgatg	aagcttatcg	ataccgtcgg	9060
ggatcctcta	gagtcgacct	gcaggcatgc	aagcttgcca	ctggccgctg	ttttacaacg	9120
tcgtgactgg	gaaaaccctg	gcgttaccca	acttaatcgc	cttgagcac	atcccccttt	9180
cgccagctgg	cgtaatagcg	aagaggcccg	caccgatcgc	ccttcccaac	agttgcgcag	9240
cctgaatggc	gaatgagctt	gagcttggtg	cagattgtcg	tttccgcct	tcagtttaaa	9300
ctatcagtgt	ttgacaggat	atattggcgg	gtaaacctaa	gagaaaagag	cgtttattag	9360
aataacggat	attttaaaagg	gcgtgaaaag	gtttatccgt	tcgtccattt	gtatgtgcat	9420
gcccaaccaca	gggttcccct	cgggatcaa				9449

<210> 5
 <211> 11885
 <212> DNA
 <213> *Agrobacterium tumefaciens*

<220>
 <221> gene
 <222> (1)..(11885)
 <223> Synthetic Sequence

<400> 5	
tggcaggata	tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60
gacgttttta	atgtactgaa ttaacgccga attgctctag ccaatacgca aaccgcctct 120
ccccgcgcgt	tggccgattc attaatgcag ctggcagcac aggtttcccg actggaaagc 180
gggcagtgag	cgcaacgcaa ttaatgtgag ttagctcact cattaggcac ccaggccttt 240
acactttatg	cttccggctc gtatgtttgt tggaattgtg agcggataac aatttcacac 300
aggaaacagc	tatgaccatg attacgaatt cgagctccac cgcggtttc tggagttaa 360
tgagctaagc	acatacgtca gaaaccatta ttgcgcgttc aaaagtcgcc taaggctact 420
atcagctagc	aaatatttct tgtcaaaaat gctccactga cgtaccataa attccccctg 480
gtatccaatt	agagtctcat attcactctc aatccaaata atctgcaatg gcaattacct 540
tatccgcaac	ttctttacct atttccgccc ggatccgggc aggttctccg gccgcttggg 600
tggagaggct	attcggctat gactgggcac aacagacaat cggtgtctct gatgccgccc 660
tgttcgggct	gtcagcgagc gggcgcccgg ttctttttgt caagaccgac ctgtccggtg 720
ccctgaatga	actgcaggac gaggcagcgc ggctatcgtg gctggccacg acgggcgttc 780
cttgcgagc	tgtgctcgac gttgtcactg aagcgggaag ggactggctg ctattgggcg 840
aagtgcgggg	gcaggatctc ctgtcatctc acctgtctcc tgccgagaaa gtatccatca 900
tggctgatgc	aatgcggcgg ctgcatacgc ttgatccggc tacctgccca ttcgaccacc 960
aagcgaaaca	tcgcatcgag cgagcacgta ctcggatgga agccggtctt gtcgatcagg 1020
atgatctgga	cgaagagcat caggggctcg cgccagccga actgttcgcc aggtcaagg 1080
cgcgcatgcc	cgacggcgag gatctcgtcg tgacctatgg cgatgcctgc ttgccgaata 1140

tcattggtgga	aaatggccgc	ttttctggat	tcattcgactg	tggccggctg	ggtgtggcgg	1200
accgctatca	ggacatagcg	ttggctaccc	gtgatattgc	tgaagagctt	ggcggcgaat	1260
gggctgaccg	cttcctcgtg	ctttacggta	tcgccgctcc	cgattcgcag	cgcattgcct	1320
tctatgcct	tcttgacgag	ttcttttgag	cgggactctg	gggttcgaaa	tgaccgacca	1380
agcgacgccc	aacctgccat	cacgagattt	cgattccacc	gccgccttct	atgaaagggt	1440
gggcttcgga	atcgttttcc	gggacgccgg	ctggatgatc	ctccagcgcg	gggatctcat	1500
gctggagttc	ttcgcccacc	ccgatccaac	acttacgttt	gcaacgtcca	agagcaaata	1560
gaccacgaac	gccggaaggt	tgccgcagcg	tgtggattgc	gtctcaattc	tctcttgacg	1620
gaatgcaatg	atgaatatga	tactgactat	gaaactttga	gggaataactg	cctagcaccg	1680
tcacctcata	acgtgcatca	tgcatgccct	gacaacatgg	aacatcgcta	tttttctgaa	1740
gaattatgct	cgttggagga	tgtcgcggca	attgcagcta	ttgccaaaat	cgaaataccc	1800
ctcacgcatg	cattcatcaa	tattattcat	gcggggaaaag	gcaagattaa	tccaactggc	1860
aatcatcca	gcgtgattgg	taacttcagt	tccagcgact	tgattcgttt	tggtgctacc	1920
cacgttttca	ataaggacga	gatggtggag	taaagaagga	gtgcgtcgaa	gcagatcggt	1980
caaacatttg	gcaataaagt	ttcttaagat	tgaatcctgt	tgccggtctt	gcgatgatta	2040
tcatataatt	tctgttgaat	tacgttaagc	atgtaataat	taacatgtaa	tgcatgacgt	2100
tatttatgag	atgggttttt	atgattagag	tcccgcatt	atacatttaa	tacgcgatag	2160
aaaacaaaat	atagcgcgca	aactaggata	aattatcgcg	cgcggtgtca	tctatgttac	2220
tagatcgatc	aaacttcggt	actgtgtaat	gacgatgagc	aatcgagagg	ctgactaaca	2280
aaaggatgac	ccaaaaacaa	cctctccaaa	ctgtttcgaa	ttggaagttt	ctgctcatgc	2340
cgacaggcat	aacttagata	ttcgcgggct	attcccacta	attcgtcctg	ctggtttgcg	2400
ccaagataaa	tcagtgcac	tccttacaag	ttcctctgtc	ttgtgaaatg	aactgctgac	2460
tgccccccaa	gaaagcctcc	tcattctcca	gttggcgggc	gctgatacac	catcgaaaac	2520
ccacgtccga	acatttgata	catgtgcctg	agaaataggc	ctacgtccaa	gagcaagtcc	2580
tttctgtgct	cgtcggaaat	tcctctcctg	tcagacggtc	gtgcgcatgt	cttgcgttga	2640
tgaagcttat	cgataccgtc	ggggatcctc	tagtggcgta	atcatgggtca	tagctgtttc	2700
ctgtgtgaaa	ttgttatccg	ctcacaattc	cacacaacat	acgagccgcg	gaagcataaa	2760
gtgtaaagcc	tggggtgcct	aatgagtgag	ctaactcaca	ttaattgcgt	tgcgctcact	2820
gcccgccttc	cagtcgggaa	acctgtcgtg	ccagctgcat	taatgaatcg	gccaacgcgc	2880
ggggagaggc	ggtttgcgta	ttgggcgctc	ttccgcttcc	tcgctcactg	actcgtcgcg	2940
ctcggtcggt	cggctgcggc	gagcggatc	agctcactca	aaggcggtaa	tacggttatc	3000
cacagaatca	ggggataacg	caggaaagaa	catgtgagca	aaaggccagc	aaaaggccag	3060
gaaccgtaaa	aaggccgcgt	tgctggcggt	tttccatagg	ctccgcccc	ctgacgagca	3120
tcacaaaaat	cgacgtcaa	gtcagagggtg	gcgaaacccg	acaggactat	aaagatacca	3180
ggcgtttccc	cctggaagct	ccctcgtgcg	ctctcctgtt	ccgaccctgc	cgcttaccgg	3240
atacctgtcc	gcctttctcc	cttcgggaag	cgtggcgctt	tctcaatgct	cacgctgtag	3300
gtatctcagt	tcggtgtagg	tcgttcgctc	caagctgggc	tgtgtgcacg	aacccccgt	3360
tcagcccgac	cgctgcgcct	tatccggtaa	ctatcgtctt	gagtccaacc	cggtaaagaca	3420
cgacttatcg	ccactggcag	cagccactgg	taacaggatt	agcagagcga	ggtatgtagg	3480

cgggtgctaca	gagttcttga	agtgggtggcc	taactacggc	tacactagaa	ggacagtatt	3540
tggtatctgc	gctctgctga	agccagttac	cttcggaaaa	agagttggta	gctcttgatc	3600
cggcaaaaa	accaccgctg	gtagcgggtg	tttttttggt	tgcaagcagc	agattacgcg	3660
cagaaaaaaa	ggatctcaag	aagatccttt	gatcttttct	acgggggtctg	acgctcagtg	3720
gaacgaaaac	tcacgttaag	ggattttggg	catgagcttg	cgccgtccc	tcaagtcagc	3780
gtaatgctct	gccagtgtta	caaccaatta	accaattctg	attagaaaa	ctcatcgagc	3840
atcaaatgaa	actgcaattt	attcatatca	ggattatcaa	taccatattt	tgaaaaagc	3900
cgtttctgta	atgaaggaga	aaactcaccg	aggcagttcc	ataggatggc	aagatcctgg	3960
tatcgggtctg	cgattccgac	tcgtccaaca	tcaatacaac	ctattaattt	cccctcgtca	4020
aaaataaggt	tatcaagtga	gaaatcacca	tgagtgcga	ctgaatccgg	tgagaatggc	4080
aaaagtttat	gcattttctt	ccagacttgt	tcaacaggcc	agccattacg	ctcgtcatca	4140
aaatcactcg	catcaaccaa	accgttattc	attcgtgatt	gcgctgagc	gagacgaaat	4200
acgcgatcgc	tgtaaaaagg	acaattacaa	acaggaatcg	aatgcaaccg	gcgaggaac	4260
actgccagcg	catcaacaat	attttcacct	gaatcaggat	attcttctaa	tacctggaat	4320
gctgtttttc	cggggatcgc	agtgggtgag	aaccatgcat	catcaggagt	acggataaaa	4380
tgcttgatgg	tcggaagagg	cataaattcc	gtcagccagt	ttagtctgac	catctcatct	4440
gtaacatcat	tggcaacgct	acctttgcca	tgtttcagaa	acaactctgg	cgcatcgggc	4500
ttcccatata	atcggtagat	tgtcgcacct	gattgcccga	cattatcgcg	agccatttta	4560
taccatata	aatcagcatc	catgttgga	tttaatcgcg	gcctcgtcga	gcaagacggt	4620
tcccgttgaa	tatggctcat	aacacccctt	gtattactgt	ttatgtaagc	agacagtttt	4680
attgttcatg	atgatataat	tttatcttgt	gcaatgtaac	atcagagatt	ttgagacaca	4740
acgtgcccc	ccccatgaca	tttaacctata	aaaataggcg	tatcacgagg	ccctttcgtc	4800
tcgcgcgttt	cgggtgatgac	ggtgaaaacc	tctgacacat	gcagctccc	gagacgggtca	4860
cagcttgtct	gtaagcggat	gccggggagca	gacaagccc	tcagggcgcg	tcagcgggtg	4920
ttggcgggtg	tcggggctgg	cttaactatg	cgcatcaga	gcagattgta	ctgagagtgc	4980
accataaaat	tgtaaacggt	aatattttgt	taaaattcgc	gttaaatttt	tgtaaataca	5040
gctcattttt	taaccaatag	gccgaaatcg	gcaaaatccc	ttataaatca	aaagaatagc	5100
ccgagatagg	gttgagtgtt	gttccagttt	ggaacaagag	tccactatta	aagaacgtgg	5160
actccaacgt	caaagggcga	aaaaccgtct	atcagggcga	tgggccacc	cgatttagag	5220
cttgacgggg	aaagccggcg	aacgtggcga	gaaaggaagg	gaagaaagcg	aaaggagcgg	5280
gcgctagggc	gctggcaagt	gtagcgggtca	cgctgcgcgt	aaccaccaca	cccgcgcgc	5340
ttaatgcgcc	gctacagggc	gcgtactatg	gttgctttga	cgtatgcggg	gtgaaatacc	5400
gcacagatgc	gtaaggagaa	aataccgcat	caggcgccat	tcgccattca	ggctgcgcaa	5460
ctgttgggaa	gggcgatcgg	tgccggccctc	ttcgctatta	cgccagctgg	cgaaaggggg	5520
atgtgctgca	aggcgattaa	gttgggtaac	gccagggttt	tcccagtcac	gacgttgtaa	5580
aacgacggcc	agtgaattgt	aatacgactc	actatagggc	gaattgggga	tcgatccact	5640
agagcagctt	gccaacatgg	tgagacacga	cactctcgtc	tactccaaga	atatcaaaga	5700
tacagtctca	gaagacaaa	gggctattga	gacttttcaa	caaagggtaa	tatcgggaaa	5760

cctcctcgga	ttccattgcc	cagctatctg	tcacttcac	aaaaggacag	tagaaaagga	5820
aggtggcacc	tacaaatgcc	atcattgcga	taaaggaaag	gctatcgttc	aagatgcctc	5880
tgccgacagt	ggccccaaag	atggaccccc	accacgagg	agcatcgtgg	aaaaagaaga	5940
cgttccaacc	acgtcttcaa	agcaagtgga	ttgatgtgat	aacatggtgg	agcacgacac	6000
tctcgtctac	tccaagaata	tcaaagatac	agtctcagaa	gaccaaaggg	ctattgagac	6060
ttttcaacaa	agggtaatat	cgggaaacct	cctcggattc	cattgcccag	ctatctgtca	6120
cttcatcaaa	aggacagtag	aaaaggaagg	tggcacctac	aaatgccatc	attgcgataa	6180
aggaaaggct	atcgttcaag	atgcctctgc	cgacagtgg	cccaaagatg	gacccccacc	6240
cacgaggagc	atcgtggaaa	aagaagacgt	tccaaccacg	tcttcaaagc	aagtggattg	6300
atgtgatata	tccactgacg	taagggatga	cgacaatcc	cactatcctt	cgcaagacct	6360
tcctctatat	aaggaagttc	atttcatttg	gagaggacac	gctgaaatca	ccagtctctc	6420
tctacaaatc	tatctctctc	gagatctatc	gatgcatgcc	atggaagaga	acaatcaaaa	6480
ccagtgcatt	ccatacaact	gcctcagtaa	tcctgaagag	gttcttctcg	atggtgaaag	6540
gatctccacc	ggaaacagtt	ccattgacat	tagtctttcc	ctcgtgcaat	tccttgtag	6600
taacttcgtc	cccgggtggtg	gattcctcgt	tggctttatc	gattttgtgt	ggggtattgt	6660
tggaccatcc	cagtgggatg	ctttcctcgt	ccaaattgag	cagcttatca	atgagcgtat	6720
tgccgagttc	gctaggaacg	ctgccattgc	taacctcgaa	ggtcttggtg	ataacttcaa	6780
catctatggt	gaggctttta	aggaatggga	ggaagaccct	aataaccccg	agactcgtac	6840
cagggtgatt	gatcgtttca	ggattctcga	tggacttctc	gaacgtgaca	tcccaagttt	6900
caggatttcc	ggtttcgagg	ttcctcttct	cagtgtctac	gccaagctg	ctaaccttca	6960
cctcgccatt	cttcgtgatt	ccgttatctt	tggtgaaagg	tggggactca	ccactattaa	7020
tgtgaacgag	aactataatc	gtcttattag	gcacatcgat	gaatacgtg	accactgcgc	7080
taacacctat	aaccgtggtc	tcaataacct	tcccaaaagt	acctaccagg	attggattac	7140
ttataacagg	ctccgtaggg	atcttaccct	caccgttctt	gacattgccg	ctttcttccc	7200
aaattacgat	aaccgtaggt	atcctatcca	accgctcggt	cagctcactc	gtgaggttta	7260
caccgatcca	cttattaact	tcaatcctca	actccagtcc	gtggctcaac	ttcccacctt	7320
taacgttatg	gaaagttcca	ggattcgtaa	cccacacctc	ttcgacatcc	ttaataacct	7380
cactattttc	accgattggt	tcagtgtcgg	aaggaaactt	tattgggggtg	gtcaccgtgt	7440
tatttccagt	cttatcgag	gtggtaacat	tacctcccct	atctacggaa	gggaggccaa	7500
ccaggaaccc	ccacgtagtt	tcactttcaa	cggtcctgtg	ttcaggaccc	tctccaatcc	7560
cacccttcgt	ctccttcaac	agccttggtc	tgtccccca	tttaacctca	ggggtgttga	7620
gggagtcgag	ttcagtactc	ctaccaactc	cttcacctat	cgtaggtagg	gtactgttga	7680
tagtcttacc	gagctcccc	cagaagacaa	ttccgtgcct	ccccgtgagg	gatacagtca	7740
caggctttgc	cacgtacct	tcgttcaacg	ttccggtact	ccatttctca	ccaccggtgt	7800
cgttttcagt	tggactgata	ggtccgccac	ccttaccac	actatcgatc	ctgaacgtat	7860
taaccagatt	cccctcgtga	agggattcag	ggtttgggg	ggtaccagtg	tcaccaccgg	7920
accaggtttc	actggtggag	acattcttcg	taggaatacc	tttgggtgatt	tcgtttccct	7980
ccaagtgaac	attaacagtc	ctatcaccca	gcgttatagg	cttcgtttca	ggtacgcttc	8040
cagtcgtgat	gctaggggtta	ttgtcctcac	tgggtccgct	tccaccggag	ttggtggtca	8100

agtgagtgtt aatatgcccc ttcagaagac tatggagatt ggagaaaacc tcacttcccg	8160
taccttcagg tataccgact ttagtaaccc atttctccttc cgtgctaata ctgatatacat	8220
tggatattagt gagcaacccc ttttcggtgc cggatccatc agttccggtg aactctacat	8280
tgataagatt gagatcattc ttgctgacgc tactttttgaa gccgagagtg atctcgaaag	8340
ggctcagaag tgacccggga gctcgaattc gaagcttctg cagacgcgtc cctcgagttt	8400
ctccataata atgtgtgagt agttcccaga taagggaatt aggggttccta taggggtttcg	8460
ctcatgtgtt gagcatataa gaaaccctta gtatgtatatt gtatttgtaa aatacttcta	8520
tcaataaaat ttctaattcc taaaaccaa atccagtact aaaatccaga tccccgaat	8580
taattcggcg ttaattcaga atcggatccc cctcggacta gagcagcttg ccaacatggt	8640
ggagcacgac actctcgtct actccaagaa tatcaaagat acagtctcag aagaccaaaag	8700
ggctattgag acttttcaac aaagggtaat atcgggaaac ctctcggat tccattgccc	8760
agctatctgt cacttcatca aaaggacagt agaaaaggaa ggtggcacct acaaagcca	8820
tcattgcat aaaggaaagg ctatcgttca agatgcctct gccgacagtg gtcccaaaga	8880
tggaccccca cccacgagga gcatcgtgga aaaagaagac gttccaacca cgtcttcaaa	8940
gcaagtggat tgatgtgata acatggtgga gcacgacact ctctgttact ccaagaatat	9000
caaagataca gtctcagaag accaaagggc tattgagact tttcaacaaa gggtaatatc	9060
gggaaacctc ctcgatttcc attgcccagc tatctgtcac ttcatacaaaa ggacagtaga	9120
aaaggaaggt ggcacctaca aatgccatca ttgcgataaa ggaaaggcta tcgttcaaga	9180
tgctctgcc gacagtggc ccaaagatgg acccccaccc acgaggagca tcgtggaaaa	9240
agaagacgtt ccaaccacgt cttcaaagca agtggattga tgtgatattc cactgacgt	9300
aagggatgac gcacaatccc actatccttc gcaagacctt cctctatata aggaagtcca	9360
tttcatttgg agaggacacg ctgaaatcac cagtctctct ctacaaatct atctctctcg	9420
agatctatcg atgcatgcca tggaagagaa caatcaaaac cagtgcattc catacaactg	9480
cctcagtaat cctgaagagg ttcttctcga tggtgaaagg atctccaccg gaaacagttc	9540
cattgacatt agtctttccc tcgtgcaatt ccttgttagt aacttcgtcc ccggtggtgg	9600
attcctcgtt ggtcttatcg attttgtgtg gggatattgt ggaccatccc agtgggatgc	9660
tttcctcgtc caaattgagc agcttatcaa tgagcgtatt gccgagttcg ctaggaacgc	9720
tgccattgct aacctcgaag gtcttggtta taacttcaac atctatgttg aggcctttta	9780
ggaatgggag gaagacctc ataaccgga gactcgtacc agggtgattg atcgtttcag	9840
gattctcgat ggacttctcg aacgtgacat cccaagtttc aggatttccg gtttcgaggt	9900
tcctcttctc agtgtctacg cccaagctgc taaccttcac ctgcgcatc ttcgtgattc	9960
cgttatcttt ggtgaaaggt ggggactcac cactattaat gtgaacgaga actataatcg	10020
tcttattagg cacatcgatg aatacgctga cactgcgct aacacctata accgtggtct	10080
caataacctt cccaaaagta cctaccagga ttggattact tataacaggc tccgtaggga	10140
tcttacctc accgttcttg acattgccgc tttcttccca aattacgata accgtaggta	10200
tcctatccaa cccgtcggtc agctcactcg tgagggttac accgatccac ttattaactt	10260
caatctcaa ctccagtccg tggctcaact tcccaccttt aacgttatgg aaagttccag	10320
gattcgtaac ccacacctc tcgacatcct taataacctc actattttca ccgattgggt	10380

```

cagtgtcggga aggaactttt attggggtgg tcaccgtgtt atttccagtc ttatcggagg 10440
tggtaacatt acctccccta tctacggaag ggaggccaac caggaacccc cacgtagttt 10500
cactttcaac ggtcctgtgt tcaggaccct ctccaatccc acccttcgtc tccttcaaca 10560
gccttggcct gctcccccatt ttaacctcag ggggtgttag ggagtcgagt tcagtactcc 10620
taccaactcc ttcacctatc gtggtagggg tactgttgat agtcttaccg agtccccccc 10680
agaagacaat tccgtgcctc cccgtgaggg atacagtcac aggccttgcc acgtacctt 10740
cgttcaacgt tccggtactc catttctcac caccggtgtc gttttcagtt ggactgatag 10800
gtccgccacc cttaccaaca ctatcgatcc tgaacgtatt aaccagattc ccctcgtgaa 10860
gggattcagg gtttggggtg gtaccagtgt catcaccgga ccaggtttca ctggtggaga 10920
cattcttcgt aggaatacct ttggtgattt cgtttccctc caagtgaaca ttaacagtcc 10980
tatcaccag cgttataggc ttcgtttcag gtacgcttcc agtcgtgatg ctagggttat 11040
tgtcctcact ggtgccgctt ccaccggagt tgggtgtcaa gtgagtgtta atatgccctt 11100
tcagaagact atggagattg gagaaaacct cacttcccgt accttcaggt ataccgactt 11160
tagtaacca ttctccttcc gtgctaattc tgatatcatt ggtattagt agcaaccctt 11220
tttcggtgcc ggatccatca gttccggtga actctacatt gataagattg agatcattct 11280
tgctgacgct acttttgaag ccgagagtga tctcgaaagg gtcagaagt gaccgggag 11340
ctcgaattcg aagcttctgc agacgcgtcc ctcgagttt tccataataa tgtgtgagta 11400
gttcccagat aagggaatta gggttcctat agggtttcgc tcatgtgttg agcatataag 11460
aaacccttag tatgtatttg tatttgtaaa atacttctat caataaaatt tctaattcct 11520
aaaacaaaaa tccagtacta aaatccagat cccccgaatt aattcggcgt taattcagaa 11580
tcggatcccc ctcggactag agtcgacctg caggcatgca agcttggcac tggccgtcgt 11640
tttacaacgt cgtgactggg aaaaccctgg cgttacccaa cttaatcgcc ttgcagcaca 11700
tccccctttc gccagctggc gtaatagcga agaggccgc accgatcgcc cttcccaaca 11760
gttgcgcagc ctgaatggcg aatgagcttg agcttggatc agattgtcgt ttcccgctt 11820
cagtttaaac tatcagtgtt tgacaggata tattggcggg taaacctag agaaaagagc 11880
gttta 11885

```

```

<210> 6
<211> 57
<212> DNA
<213> Bacillus thuringiensis

```

```

<220>
<221> misc_feature
<222> (1)..(52)

```

```

<220>
<221> misc_feature
<222> (1)..(57)

```

```

<400> 6
ttcggcggtta attcagaatc ggatccccct cggactagat tagagcagct tgccaac 57

```

```

<210> 7
<211> 2317
<212> DNA
<213> Bacillus thuringiensis

```

<220>
 <221> prim_transcript
 <222> (1)..(2317)

<400> 7
 aattcgagct ccaccgcggg tttctggagt ttaatgagct aagcacatac gtcagaaacc 60
 attattgctg gttcaaaagt cgcctaaggt cactatcagc tagcaaatat ttcttgtaaa 120
 aaatgctcca ctgacgtacc ataaattccc ctcggtatcc aattagagtc tcatattcac 180
 tctcaatcca aataatctgc aatggcaatt accttatccg caacttcttt acctatttcc 240
 gcccggatcc gggcagggtt tccggccgct tgggtggaga ggctattcgg ctatgactgg 300
 gcacaacaga caatcggtc ctctgatgcc gccgtgttcc ggctgtcagc gcaggggagc 360
 ccggttcttt ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca 420
 gcgcggctat cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgct 480
 actgaagcgg gaagggactg gctgctattg ggcaagtgc cggggcagga tctcctgtca 540
 tctcaccttg ctctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat 600
 acgcttgatc cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca 660
 cgtactcgga tgggaagccg tcttgctgat caggatgatc tggacgaaga gcatcagggg 720
 ctgcgccag ccgaactgtt cgccaggctc aaggcgcgca tgcccagcg cgaggatctc 780
 gtcgtgacct atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct 840
 ggattcatcg actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggt 900
 acccgatgata ttgctgaaga gcttgccggc gaatgggctg accgcttcct cgtgctttac 960
 ggtatcgccg ctcccgattc gcagcgcac gccttctatc gccttcttga cgagttcttc 1020
 tgagcgggac tctgggggtt gaaatgaccg accaagcgac gcccaacctg ccatcacgag 1080
 atttcgattc caccgccgcc ttctatgaaa ggttgggctt cggaatcggt ttccgggacg 1140
 ccggctggat gatcctccag cgcggggatc tcatgctgga gttcttcgcc caccgccatc 1200
 caacacttac gtttgcaacg tccaagagca aatagaccac gaacgccgga aggttgccgc 1260
 agcgtgtgga ttgctgtcct attctctctt gcaggaatgc aatgatgaat atgatactga 1320
 ctatgaaact ttgagggaat actgcctagc accgtcacct cataacgtgc atcatgcatg 1380
 ccctgacaac atggaacatc gctatttttc tgaagaatta tgctcgttgg aggatgtcgc 1440
 ggcaattgca gctattgcca aaatcgaaat acccctcacg catgcattca tcaatattat 1500
 tcatgccccg aaaggcaaga ttaatccaac tggcaaatca tccagcgtga ttggttaactt 1560
 cagttccagc gacttgattc gttttggtgc taccacggtt ttcaataagg acgagatggt 1620
 ggagtaaaga aggagtgcgt cgaagcagat cgttcaaaca tttggcaata aagtttctta 1680
 agattgaatc ctgttgccgg tcttgcatg attatcatat aatttctggt gaattacggt 1740
 aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt ttttatgatt 1800
 agagtccgcg aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaaactag 1860
 gataaattat cgcgcgcggg gtcatttatg ttactagatc gatcaaaactt cggctactgtg 1920
 taatgacgat gagcaatcga gaggtgact aacaaaagg atgccccaaa acaacctctc 1980
 caaactgttt cgaattggaa gtttctgctc atgccgacag gcataactta gatattcgcg 2040
 ggctattccc actaattcgt cctgctgggt tgccgcaaga taaatcagtg catctcctta 2100
 caagttcttc tgtcttgtga aatgaactgc tgactgcccc ccaagaaagc ctctcatct 2160

cccagttggc	ggcggctgat	acaccatcga	aaacccacgt	ccgaacactt	gatacatgtg	2220
cctgagaaat	aggcctacgt	ccaagagcaa	gtcctttctg	tgctcgtcgg	aaattcctct	2280
cctgtcagac	ggtcgtgcgc	atgtcttgcg	ttgatga			2317

<210> 8
 <211> 1954
 <212> DNA
 <213> *Bacillus thuringiensis*

<220>
 <221> prim_transcript
 <222> (1)..(1954)

<400> 8	
tcgagatcta	tcgatgcatg ccatggaaga gaacaatcaa aaccagtgc
60	ttccatacaa
ctgcctcagt	aatcctgaag aggttcttct cgatggtgaa aggatctcca ccggaaacag
120	
ttccattgac	attagtcttt ccctcgtgca attccttggt agtaacttcg tccccggtgg
180	
tggattcctc	gttggcttta tcgattttgt gtgggggtatt gttggaccat cccagtggga
240	
tgctttcctc	gtccaaattg agcagcttat caatgagcgt attgccgagt tcgctaggaa
300	
cgctgccatt	gctaacctcg aaggcttgg taataacttc aacatctatg ttgaggcttt
360	
taaggaatgg	gaggaagacc ctaataacct cgagactcgt accaggggtga ttgatcgttt
420	
caggattctc	gatggacttc tcgaacgtga catcccaagt ttcaggattt ccggtttcga
480	
ggttcctctt	ctcagtgtct acgccccagc tgctaacctt cacctcgcca ttcttcgtga
540	
ttccgttatc	tttgggtgaaa ggtggggact caccactatt aatgtgaacg agaactataa
600	
tcgtcttatt	aggcacatcg atgaatacgc tgaccactgc gctaacacct ataaccgtgg
660	
tctcaataac	cttccccaaa gtacctacca ggattggatt acttataaca ggctccgtag
720	
ggatcttacc	ctcaccttct ttgacattgc cgctttcttc ccaaattacg ataaccgtag
780	
gtatcctatc	caaccgctcg gtcagctcac tcgtgagggt tacaccgatc cacttattaa
840	
cttcaatcct	caactccagt ccgtgggtca acttcccacc tttaacgtta tggaaagttc
900	
caggattcgt	aaccacacc tcttcgacat ccttaataac ctactattt tcaccgattg
960	
gttcagtgtc	ggaaggaact tttattgggg tggtcacctg gttatttcca gtcttatcgg
1020	
aggtggtaac	attacctccc ctatctacgg aaggaggagg aaccaggaac cccacgtag
1080	
tttactttc	aacggctctg tggtcaggac cctctccaat cccacccttc gtctcttca
1140	
acagccttgg	cctgctcccc catttaacct caggggtggt gagggagtcg agttcagtac
1200	
tcctaccaac	tccttcacct atcggtgtag ggggtactgt gatagtctta ccgagctccc
1260	
cccagaagac	aattccgtgc ctccccgtga gggatacagt cacaggcttt gccacgctac
1320	
cttcgttcaa	cgttccggta ctccatttct caccaccggt gtcgttttca gttggactga
1380	
taggtccgcc	acccttacca aactatcga tcctgaacgt attaaccaga ttccccctgt
1440	
gaagggattc	aggggttggg gtggtaccag tgatcatcacc ggaccagggt tctactggtg
1500	
agacattctt	cgtaggaata cctttggtga tttcgtttcc ctccaagtga acattaacag
1560	
tcctatcacc	cagcgttata ggcttcgttt caggtagcgt tccagtcgtg atgctagggt
1620	
tattgtctc	actggtgccg cttccaccgg agttggtggt caagtgagtg ttaatatgcc
1680	
ccttcagaag	actatggaga ttggagaaaa cctcacttcc cgtaccttca ggtataccga
1740	
ctttagtaac	ccattctcct tccgtgctaa tcctgatatc attggtatta gtgagcaacc
1800	

```

ccttttcggt gccg gatcca tcagttccgg tgaactctac attgataaga ttgagatcat    1860
tcttgctgac gctacttttg aagccgagag tgatctcgaa agggctcaga agtgacccgg    1920
gagctcgaat tcgaagcttc tgcagacgcg tccc                                1954

```

```

<210> 9
<211> 848
<212> DNA
<213> Gossypium hirsutum

```

```

<220>
<221> gene
<222> (1)..(848)

```

```

<400> 9
ccagcaggat cagctggtac tcgcaagca tctagattca ttccacacaa catacgagcc    60
ggaagcataa agtgtaaagc ctgggggtgcc taatgagtga gtaacacac attaatgctg    120
ttgcgctcac tgcccgtttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc    180
ggccaacgcg cggggagagg cggtttgctg attggctaga gcaattcggc gttaattcag    240
tacattaatt caattctgca cggagatgag gaagggtggag tggagttgat gtggacgggg    300
tctaaccttc ttcgctcatt tattttttat aatggctctt atgacatggg atctcagggg    360
cctatgttca cttggaatcg ggttaacctt tttagcaacc cgatcgggca atttgtaatg    420
atctatggaa tgtgatcgct cctgaaagta acgtaggcac cttcctagac ccaagtctga    480
tcgtcaaccg attttattgt ctgtcttcaa ggccgcacat tactactaat cgaccaattt    540
gtttcttagc tagttgggta ttgcaccctc gattccgagg tcttgtacaa gaaactcgga    600
actggactca ggatattgtt actaacttgg atgattttac tttgaaagtc taagagtggg    660
ataaatgtgt ttatggtaac atctttgcaa ttgggttggg ttcagaaagc tctcaaatat    720
tttgacacca atcgactcca aggacaggaa caagtaccta aagttgagct tgagagcatc    780
ttgagtcagg aagagatcct ccggtaccag aaaccgagaa gcgattggct tcttaattggc    840
gataggaa                                848

```

```

<210> 10
<211> 853
<212> DNA
<213> Gossypium hirsutum

```

```

<220>
<221> gene
<222> (1)..(853)

```

```

<400> 10
cgatgaagat cgccagctct atacgactca ctatagggaa agcttgcatg caggcctctg    60
cagtcgacgg gcccgggatc cgattgcact ggccgtcggt ttacaacgcg tgactgggaa    120
aaccctggcg ttacacaact taatcgctt gcagcacatc cccctttcgc cagctggcgt    180
aatagcgaag aggcccgcac cgatcgccct tcccaacagt tgccgagcct gaatggcgaa    240
tgagcttgag cttggatcag attgtcggtt cccgccttca gtttaaacta tcagtgtttg    300
attgtctccc ttcccttccc ggtatcggtt catggattcg gttagatggg aaaccgccat    360
cagtaccagg tcgtaatccc acacactggc catgccggcg ggccctgcgg aaacctctac    420
gtgcccgctt ggaagctcgt agcggatcac ctcgccagct cgtcgctcac gcttcgacag    480

```

acggaaaacg gccacgtcca tgatgctgcg actaccgcgg gtgcccacgt catagagcat 540
 cggaacgaaa aaatctgggt gctcgtcgcc cttgggcggc ttcctaatacg acggcgacc 600
 ggctgccggc gggttgccggg attctttgcg gattcgatca gcggccgctt gccacgattc 660
 accggggcgt gcttctgcct cgatgcgttg tcgctgggcg gcctgcgcgg ccttcaactt 720
 ctccaccagg tcatcaccca gcgccgcgc gatttgtacc gggccggatg gtttgcgacc 780
 gctcacgccg attcctcggg ctcggggggt ccagtgccat tgcagggccg gcagacaacc 840
 cagccgctta cgc 853

<210> 11
 <211> 26
 <212> DNA
 <213> *Gossypium hirsutum*

<220>
 <221> primer_bind
 <222> (1)..(26)

<400> 11
 gaccttaggc gacttttgaa cgcgca 26

<210> 12
 <211> 23
 <212> DNA
 <213> *Gossypium hirsutum*

<220>
 <221> primer_bind
 <222> (1)..(23)

<400> 12
 agtagacgag gagttgattg cga 23

<210> 13
 <211> 25
 <212> DNA
 <213> *Gossypium hirsutum*

<220>
 <221> primer_bind
 <222> (1)..(25)

<400> 13
 aacagttgcg cagcctgaat ggcga 25

<210> 14
 <211> 24
 <212> DNA
 <213> *Gossypium hirsutum*

<220>
 <221> primer_bind
 <222> (1)..(24)

<400> 14
 atcaagaaca accaggcacc gacg 24

<210> 15
 <211> 24
 <212> DNA
 <213> *Gossypium hirsutum*

<220>
 <221> primer_bind
 <222> (1)..(24)

<400> 15
 gggtttcgct catgtgttga gcat 24

<210> 16
 <211> 28
 <212> DNA
 <213> *Gossypium hirsutum*

<220>
 <221> primer_bind
 <222> (1)..(28)

<400> 16
 aactgcagtg cgaaagctcg agagagat 28

<210> 17
 <211> 975
 <212> DNA
 <213> *Gossypium hirsutum*

<220>
 <221> misc_feature
 <222> (1)..(975)

<400> 17
 gggtttcgct catgtgttga gcatataaga aacccttagt atgtatttgt atttgtaaaa 60
 tacttctatc aataaaattt ctaattccta aaaccaaagt ccagtactaa aatccagatc 120
 ccccgaatta attcggcggt aattcagaat cggatcccc tcggactaga ttagagcagc 180
 ttgccaacat ggtggagcac gacactctcg tctactccaa gaatatcaaa gatacagtct 240
 cagaagacca aagggtctatt gagacttttc aacaaagggt aatatcggga aacctcctcg 300
 gattccattg cccagctatc tgtcacttca tcaaaaggac agtagaaaag gaagggtggca 360
 cctacaaatg ccatcattgc gataaaggaa aggctatcgt tcaagatgcc tctgccgaca 420
 gtggtcccaa agatggaccc ccaccacga ggagcatcgt ggaaaaagaa gacgttccaa 480
 ccacgtcttc aaagcaagtg gattgatgtg ataacatggt ggagcacgac actctcgtct 540
 actccaagaa tatcaaagat acagtctcag aagaccaaag ggctattgag acttttcaac 600
 aaagggtaat atcgggaaac ctctcggat tccattgccc agctatctgt cacttcatca 660
 aaaggacagt agaaaaggaa ggtggcacct acaaatgcc tcatcgcat aaaggaaagg 720
 ctatcggttc agatgcctct gccgacagt gtcccaaaga tggaccccca cccacgagga 780
 gcatcggtga aaaagaagac gttccaacca cgtcttcaaa gcaagtggat tgatgtgata 840
 tctccactga cgtaagggt gacgcacaat cccactatcc ttcgcaagac cttcctctat 900
 ataaggaagt tcatttcatt tggagaggac acgctgaaat caccagtctc tctctacaaa 960
 tctatctctc tcgag 975

<210> 18
 <211> 1373
 <212> DNA
 <213> *Gossypium hirsutum*

<220>
 <221> gene
 <222> (1)..(1373)

```

<400> 18
gcactggccg tcgttttaca acgtcgtgac tgggaaaacc ctggcgttac ccaacttaat      60
cgcccttgag cacatcccc tttcgccagc tggcgtaata gcgaagaggc ccgcaccgat      120
cgcccttccc aacagttgag cagcctgaat ggcgaaatgag cttgagcttg gatcagattg      180
tcgtttctcc gccgtcattc attaatgcac tggcacatga ttgtgaaccg aaaaacctag      240
cccagggttta ttaggccctt gaagaaagct gcagacactg agaaaatatt aatatttgat      300
attgtaattt taatggtttt catccttttc ttgtatttaa atttgaaatt tattttttat      360
agtttgtttg gtataacttg gtctctttat tttttaacta ttattaaact attccgatta      420
aatgcatatg caaatttttc ttaaaaatat tccttttaat ttatatgtac tctttttcag      480
gttgaaataa gtgtttttta caattaatgg tattataaga gtagaaatat caaattatgc      540
aaaattaaaa taaataaaaa acaactaaat accaaattta aacacagtag atacttaata      600
tatattcatg tgatcataat atataaaatt ttgtaattta agaataaaaag aagttgtaaa      660
tgttggagat gcaaacatag ctctttcatt agccacattc ctttctccaa ttgtttacct      720
acctgcaaag acaattgcaa tatcggatca tctcaatcat ataccattgt tcttcaacaa      780
atcttttact gccatcatgg tgttatatct ttcacgaaaa ataattattgt tctctccaat      840
tattaacaag aatccccact ccaaagttgg atatcgttcc taacttacac acttgtctcc      900
tttttagaat gtctttccaa atccacgagt ctgatgcctc aagggtttct ttatttttgt      960
actattttgc tttgaaagcc ttggtgatta gtagatttgg gttggttgag agacgccaca     1020
cctgttttgc taaaagtgtt ttattcatca aagcaatagg tctaatacct agacctcctt     1080
caactttagg actacagatt gtagtctaata tacacaaatg catgcatgta ccatgactcc     1140
actagaaagc tctcaaactt ggtcaatctc atcacaaata tagcgacagg ctttgaaaac     1200
agaacatgta aagaaagtgg atttaatgag ggtgagtcta cctcattatg aaagacactt     1260
cgcccttcaa tgtgccactt tagcttgaat cttgtcaatg atgttttaga atagctcttt     1320
tatcactgat ctgcgaaaca ccaaactcga tatcttgctg aaaaactcga gcc              1373

```

```

<210> 19
<211> 26
<212> DNA
<213> Gossypium hirsutum

```

```

<220>
<221> primer_bind
<222> (1)..(26)

```

```

<400> 19
cgttacccaa cttaatcgcc ttgcag      26

```

```

<210> 20
<211> 30
<212> DNA
<213> Gossypium hirsutum

```

```

<220>
<221> primer_bind
<222> (1)..(30)

```

```

<400> 20
agaaaccctt gaggcacag actcgtggat      30

```

We Claim:

1. A nucleotide sequence set forth as SEQ ID No. 1 or a nucleotide sequence comprising sequence set forth as SEQ ID No.1.
2. An expression cassette set forth as SEQ ID No.3, comprising CaMV 35S promoter, nucleotide sequence set forth as SEQ ID No. 1 or a nucleotide sequence comprising sequence set forth as SEQ ID No.1 and 35S 3' untranslated region.
3. The nucleotide sequence as claimed in claim 1 and the expression cassette as claimed in claim 2, wherein the nucleotide sequence codes for Cry1C gene and corresponding amino acid sequence set forth as SEQ ID No. 2 codes for Cry1C protein.
4. A vector comprising sequence set forth as SEQ ID No. 5, having an expression cassette as claimed in claim 2.
5. The vector as claimed in claim 4, wherein the vector is selected from a group comprising an expression vector, replicable vector, transformation vector, binary vector and shuttle vectors or any combination thereof.
6. The vector as claimed in claims 4 and 5, wherein the vector is preferably binary vector and wherein the expression cassette is either a single copy sequence or tandem repeat sequence separated by an intervening sequence set forth as SEQ ID No.6.
7. A transformed host cell comprising the vector as claimed in claim 4.
8. A method of obtaining a transformed host cell comprising an expression cassette set forth as SEQ ID No.3, said method comprising acts of:
 - a) inserting an expression cassette set forth as SEQ ID No.3 into a vector; and
 - b) transforming a host cell with said vector to obtain the transformed host cell.
9. The method as claimed in claim 8, wherein the expression cassette comprise CaMV 35S promoter, nucleotide sequence set forth as SEQ ID No. 1 or a nucleotide sequence comprising sequence set forth as SEQ ID No.1 and 35S 3' untranslated region.

10. The method as claimed in claim 9, wherein said nucleotide sequence codes for Cry1C gene and corresponding amino acid sequence set forth as SEQ ID No. 2 codes for Cry1C protein.
11. The method as claimed in claim 8, wherein the host cell is *Agrobacterium*.
12. The method as claimed in claim 8, wherein the transforming of the host cell is carried out using techniques selected from a group comprising electroporation, microinjection, genegun method, PEG mediated transfer, Calcium phosphate method, liposome mediated transfer and tri-parental mating or any combination thereof.
13. A transgenic cell comprising an expression cassette as claimed in claim 2.
14. The transgenic cell as claimed in claim 13, wherein the transgenic cell is a transgenic plant cell exhibiting resistance to lepidopteran insect(s) or pest(s) infestation; and wherein the plant is a member of genus *Gossypium*.
15. A method of obtaining a transgenic plant member of genus *Gossypium*, said method comprising acts of:
 - a) codon optimizing nucleotide sequence of native Cry1C gene to obtain nucleotide sequence set forth as SEQ ID No.1;
 - b) inserting an expression cassette as claimed in claim 2 into a vector and transforming a host cell with said vector to obtain a transformed host cell; and
 - c) infecting the plant with the transformed host cell followed by tissue culturing to obtain a transgenic plant member of genus *Gossypium*.
16. The method as claimed in claim 15(b), wherein the expression cassette is either a single copy sequence or tandem repeat sequence separated by an intervening sequence set forth as SEQ ID No.6.
17. The method as claimed in claim 15, wherein the host cell is *Agrobacterium*.
18. The method as claimed in claim 15, wherein the transforming the host cell is carried out using techniques selected from a group comprising electroporation, microinjection, genegun method, PEG mediated transfer, Calcium phosphate method and liposome mediated transfer or any combination thereof.
19. The method as claimed in claim 15 (c), wherein the infection and tissue culturing comprises incubating cotton leaf with the transformed host cell; transferring and incubating the cotton leaf in bacterial selection medium followed by incubation in plant selection medium; transferring to

embryogenesis medium till the occurrence of embryogenesis followed by sub-culturing; transferring to suspension medium followed by germination medium; transferring to basal medium after the growth of leaves and root hairs; hardening followed by transfer to soil and greenhouse to obtain the said transgenic plant.

20. A transgenic transformation event MLS9124 or MLS9878, said event comprising a nucleotide sequence set forth as SEQ ID No. 1 or a nucleotide sequence comprising sequence set forth as SEQ ID No.1 in a plant member of genus *Gossypium* or any part thereof.
21. A method of detecting presence of a transgene comprising an expression cassette set forth as SEQ ID No.3, in a transgenic plant member of genus *Gossypium*, said method comprising acts of:
 - a. extracting DNA from the plant source and performing nucleic acid amplification of junction regions of the transgene and the plant to obtain an amplicon, said amplification carried out by primers corresponding to the regions selected from a group comprising left border region of the transgene, right border region of the transgene, left border region of the plant DNA and right border region of the plant DNA or any combination thereof;
 - b. detecting and analyzing the amplicon to detect the presence of said transgene in the transgenic plant.
22. The method as claimed in claim 21, wherein the primers are selected from a group comprising SEQ ID No. 11 corresponding to the left border region of the transgene, SEQ ID No. 13 corresponding to the right border region of the transgene, SEQ ID No. 12 corresponding to the left border region of the plant DNA and SEQ ID No. 14 corresponding to the right border region of the plant DNA or any combination thereof.
23. Primers set forth as SEQ ID No. 11, SEQ ID No. 12, SEQ ID No. 13, SEQ ID No. 14, SEQ ID No. 15, SEQ ID No. 16, SEQ ID No. 19 and SEQ ID No. 20.
24. The primers as claimed in claim 23, wherein the primers having SEQ ID Nos.12, 13, 15 and 19 are sense primers and the primers having SEQ ID Nos.11, 14, 16 and 20 are anti-sense primers.
25. A method of codon optimizing Cry1C gene in plant member of genus *Gossypium*, said method comprising acts of:

- a) preparing a codon usage table as described in Table 2, for plurality of constitutively expressed genes, on the basis of GC content, TA doublet avoidance in second and third positions of codons, transcription termination signals, splicing signals and polyadenylation signals, from conventionally known constitutively expressed genes in cotton; and
 - b) using said table for codon optimizing the Cry1C gene.
26. The nucleotide sequence as claimed in claim 1, the expression cassette as claimed in claim 2 and the method as claimed in claim 9, wherein the nucleotide sequence is codon optimized for plants of genus *Gossypium*.

		1	50
Cry1CGh	(1)	ATGGAAGAGAACAAATCAAAAACCAGTGCATTCATACAACTGCCTCAGTAA	
x07518	(1)	ATGGAGGAAAATAATCAAAAATCAATGCATACTTACAAATGTTTTAAGTAA	
Consensus	(1)	ATGGA GA AA AATCAAAA CA TGCAT CC TACAA TG T AGTAA	
		51	100
Cry1CGh	(51)	TCCTGAAGAGGTTCTTTCCTCAGTGGTGAAAGGATCTCCACCGGAAACAGTT	
x07518	(51)	TCCTGAAGAAGTACTTTTGGATGGAGAACGGATATCAACTGGTAATTCAAT	
Consensus	(51)	TCCTGAAGA GT CTT T GATGG GAA GGAT TC AC GG AA T	
		101	150
Cry1CGh	(101)	CCATTGACATTAGTCTTTCCTCCTGCAATTCCTGTGTAGTAACTTTCGTC	
x07518	(101)	CAATTGATATTTCTCTGTCACCTTGTTCAGTTCTGSTATCTAACTTTGTA	
Consensus	(101)	C ATTGA ATT TCT TC CT GT CA TT CT GT TAACTT GT	
		151	200
Cry1CGh	(151)	CCCGGTGGTGGATTCCCTCGTTGGTCTTATCGATTTTGTGIGGGGTATTGT	
x07518	(151)	CCAGCGGGGAGGATTTTATGTTGGATTAATAGATTTTGTATGGGGAATAGT	
Consensus	(151)	CC GG GG GGATT T GTTGG T AT GATTTTGT TGGGG AT GT	
		201	250
Cry1CGh	(201)	TGGACCATCCAGTGGGATGCTTTCTCGTCCAAATTGAGCAGCTTATCA	
x07518	(201)	TGGCCCTTCTCAATGGGATGCAATTTCTAGTACAAATTGACAATTAATTA	
Consensus	(201)	TGG CC TC CA TGGGATGC TT CT GT CAAATTGA CA T AT A	
		251	300
Cry1CGh	(251)	ATGAGCGTATTTGCCGAGTTCTCTAGGAACGCTGCCATTGCTAACCTCGAA	
x07518	(251)	ATGAAAGAATAAGCTGAATTTGCTAGGAATGCTGCTATTGCTAATTAGAA	
Consensus	(251)	ATGA G AT GC GA TT GCTAGGAA GCTGC ATTGCTAA T GAA	
		301	350
Cry1CGh	(301)	GGTCTTGGTAATAACTTCAACATCTATGTTGAGGCTTTTAAAGGAATGGGA	
x07518	(301)	GGATTAGGAACAATTTCAATATATATGEGGAAGCAATTTAAGAATGGGA	
Consensus	(301)	GG T GG AA AA TTCAA AT TATGT GA GC TTTAA GAATGGGA	
		351	400
Cry1CGh	(351)	GGAAGACCTTAATAACCCGAGACTCTTACCAGGGTGATTGATCGTTTTCA	
x07518	(351)	AGAAGATCCTAATAATCCAGAAACCAAGACCAGAGTAATTGATCGCTTTC	
Consensus	(351)	GAAGA CCTAATAA CC GA AC G ACCAG GT ATTGATCG TT	
		401	450
Cry1CGh	(401)	GGATTCTCGATGGACTTCTTCGAACCTGACATCCCAAGTTTCAGGATTTCC	
x07518	(401)	GTATACTTGATGGGCTACTTGAAGGGACATTCCTTCGTTTCAATTTCT	
Consensus	(401)	G AT CT GATGG CT CT GAA G GACAT CC TT G ATTTT	
		451	500
Cry1CGh	(451)	GGTTTCGAGGTTCTCTTCTCAGTCTCTACGCCAAGCTGCTAACCTTCA	
x07518	(451)	GGATTTCGAGTACCCCTTTATCCCTTTATGCTCAAGCGGCCAATCTGCA	
Consensus	(451)	GG TT GA GT CC CTT T GT TA GC CAAGC GC AA CT CA	
		501	550
Cry1CGh	(501)	CCTCCCATTTCTTCGTGATTCCTTATCTTTGGTGAAAGGTGGGGACTCA	
x07518	(501)	TCTAGCTATATTAAGAGATTTCTTAATTTTGGAGAAAGATGGGGATTTGA	
Consensus	(501)	CT GC AT T G GATTC GT AT TTTGG GAAAG TGGGGA T A	
		551	600
Cry1CGh	(551)	CCACTATTAATGTGAACGAGAACTATAATCCTCTATTAGGCACATCGAT	
x07518	(551)	CAACGATAAATGTCAATGAAAACATAATAGACTAATTAGGCATATTGAT	
Consensus	(551)	C AC AT AATGT AA GA AACTATAAT G CT ATTAGGCA AT GAT	

FIGURE: 1

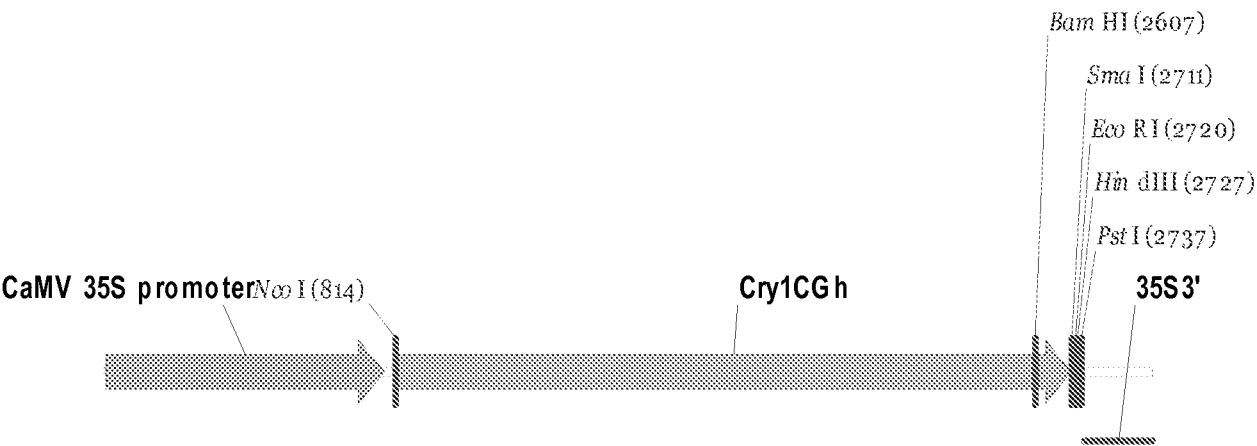
		601	650
Cry1CGh	(601)	GAATACGCTGACCACTGCCCTAACACCTATAACCGTGGTCTCAATAACCT	

x07518	(601)	GAATATGCTGATCACTGTGCAAAATACGTATAATCGGGGATFAAATAATTT	
Consensus	(601)	GAATA GCTGA CACTG GC AA AC TATAA CG GG T AATAA T	
		651	700
Cry1CGh	(651)	TCCCAAAAGTACCTACCGAGGATTGGATTACTTATAACAGCCTCCGTAGGG	
x07518	(651)	ACCGAAATCTACGTATCAAGATTGGATAACATATAATCGATTACGGAGAG	
Consensus	(651)	CC AAA TAC TA CA GATTGGAT AC TATAA G T CG AG G	
		701	750
Cry1CGh	(701)	ATCTTACCTCACCCTTCTGTGACATTGCCCGCTTTCTTCCCAAATTACGAT	
x07518	(701)	ACTTAACATTGACTGTATAGATATCGCCCGCTTTCTTCCAAACTATGAC	
Consensus	(701)	A T AC T AC GT T GA AT GCCGCTTTCTT CCAA TA GA	
		751	800
Cry1CGh	(751)	AACCGTAGGTATCTATCCAAACCGTCGGTCAGCTCACTCGTGAGGTTTA	
x07518	(751)	AATAGGAGATATCCAAATTCAGCCAGTTGGTCAACTAACAGGGAGTTTA	
Consensus	(751)	AA G AG TATCC AT CA CC GT GGTCA CT AC G GA GTTTA	
		801	850
Cry1CGh	(801)	CACCGATCCACTTATTAACCTTCAATCTCAACTCCAGTCCGTGGCTCAAC	
x07518	(801)	TACGGACCCATTAAATTAATTTAATCCACAGTTACAGTCTGTAGCTCAAT	
Consensus	(801)	AC GA CCA T ATTAA TT AATCC CA T CAGTC GT GCTCAA	
		851	900
Cry1CGh	(851)	TTCCCACTTTTAAAGTTATGSAAGTTCCAGGATTCCTAACCCACACCTC	
x07518	(851)	TACCTACTTTTAAAGTTATGGAGAGCAGCCCAATTAGAAATCCTCATTTA	
Consensus	(851)	T CC AC TTTAACGTTATGGA AG C G ATT G AA CC CA T	
		901	950
Cry1CGh	(901)	TTCGACATCCCTAATAACCTCACTATTTTCACCGATTGGTTTCACTGTCGG	
x07518	(901)	TTTGATATATGAATAATCTTACAATCTTTACGGAATTGGTTTAGTGTTGG	
Consensus	(901)	TT GA AT T AATAA CT AC AT TT AC GATTGGTT AGTGT GG	
		951	1000
Cry1CGh	(951)	AAGGAACTTTTATTGGGGTGTCTACCGTCTTATTTCCAGTCTTATCGGAG	
x07518	(951)	ACGCAATTTTTATTGGGGAGGACATCGAGTTAATATCTAGCCTTATAGGAG	
Consensus	(951)	A G AA TTTTATTGGGG GG CA CG GT AT TC AG CTTAT GGAG	
		1001	1050
Cry1CGh	(1001)	GTGGTAACATTACCTCCCTATCTACGGAAGGGAGGCCAACCAGGAACCC	
x07518	(1001)	GTGGTAACATAACATCTCTATATATGGAAGAGAGGCCAACCAGGAGCCT	
Consensus	(1001)	GTGGTAACAT AC TC CCTAT TA GGAAG GAGGC AACCAGGA CC	
		1051	1100
Cry1CGh	(1051)	CCACGTAGTTTCACTTTCAACCGTCTGTGTTTCAGGAACCTCTCCAATCC	
x07518	(1051)	CCAAGATCCCTTACTTTTAATCGACCGGTATTTAGGACTTTATCAATCC	
Consensus	(1051)	CCA G TT ACTTT AA GG CC GT TT AGGAC T TC AATCC	
		1101	1150
Cry1CGh	(1101)	CACCCCTCGTCTCCCTCAACAGCCTTGGCCTGCTCCCCATTAAACCTCA	
x07518	(1101)	TACTTTACGATTATACAGCAACCTTGGCCAGCGCCACCATTTAATTTAC	
Consensus	(1101)	AC T CG T T CA CA CCTTGGCC GC CC CATTTAA T	
		1151	1200
Cry1CGh	(1151)	GGGGTGTTCAGGGGAGTCAAGTTTCAGTACTCTACCAACTCTTACACCTAT	
x07518	(1151)	GTGGTGTTCAGGGGAGTCAAAATTTCTACACCTACAAATAGCTTTACGTAT	
Consensus	(1151)	G GGTGTTGA GGAGT GA TT TAC CCTAC AA CTT AC TAT	
		1201	1250
Cry1CGh	(1201)	CGTGGTAGGGCTACTGTTGATAGTCTTAACGAGCTCCCTCCAGAGACAA	
x07518	(1201)	CGAGGAAGAGGTACGTTGATCTCTTAACGAGCTCCCTCCAGAGACAA	
Consensus	(1201)	CG GG AG GGTAC GTTGAT T T AC GA T CC CC GA GA AA	
		1251	1300
Cry1CGh	(1251)	TTCCGTGGCTCCCGTGAGGGGATACAGTCACAGGCCTTCCACGGTACCT	
x07518	(1251)	TAGTGTGGCAACCTCCCGAGGATATAGTCATCGTTTATGTCTGTGCAACTT	

FIGURE: 1 (cont'd)

Consensus	(1251)	T GTGCC CC CG GA GGATA AGTCA G T TG CA GC AC T	1301	1350
Cry1CGh	(1301)	TCGTTCAACGTTCCGGTACTCCATTTCTCACCACCGGTGTGCTTTTCAGT		
x07518	(1301)	TTGTTCAAAGATCTGGAACACCTTTTTFACCAACTGGTGTAGTATTTTCT		
Consensus	(1301)	T GTTCAA G TC GG AC CC TTT T AC AC GGTGT GT TT T	1351	1400
Cry1CGh	(1351)	TGGACTSATAAGGTCCGTCACCCCTTACCAACACTATCGATCCTGAACGTAT		
x07518	(1351)	TGGACCGATCGTAGTCCAACTCTTACAAATACAATTGATCCAGAGAGAAAT		
Consensus	(1351)	TGGAC GAT G GC AC CTTAC AA AC AT GATCC GA G AT	1401	1450
Cry1CGh	(1401)	TAACCAAGATTCCCTTCGTGAAGGGATTTCAGGGTTTGGGGTGGTACCAGTG		
x07518	(1401)	TAATCAAATACTTTAGTGAAAGGATTTAGAGTTTGGGGGGGCACTCTG		
Consensus	(1401)	TAA CA AT CC T GTGAA GGATT AG GTTTGGGG GG ACC TG	1451	1500
Cry1CGh	(1451)	TCATCACCGGACCAGGTTTCACTGGTGGAGACATTCTTCTGTAGGAATAAC		
x07518	(1451)	TCATTACAGGACCAGGATTTACAGGAGGGATATCCTTUGAAGAAATAAC		
Consensus	(1451)	TCAT AC GGACCAGG TT AC GG GG GA AT CTTCTG AG AATACC		
			1501	1550
Cry1CGh	(1501)	TTTGGTGATTTCTTTTCCTTCCAAGTGAACTTAACAGTCTTATCACCCA		
x07518	(1501)	TTTGGTGATTTCTTAATCTCAACAGTCAATATTAATTACCAATTACCCA		
Consensus	(1501)	TTTGGTGATTT GT TC CT CAAGT AA ATTAA CC AT ACCCA	1551	1600
Cry1CGh	(1551)	GCGTTATAGGCTTCGTTTCAGGTACGCTTCCAGTCTGATGCTAGGTTA		
x07518	(1551)	AAGATACCTTTAAGATTTCTGTTACGCTTCCAGTAGGGATGCACAGTTA		
Consensus	(1551)	G TA G T G TT G TACGCTTCCAGT G GATGC G GTTA	1601	1650
Cry1CGh	(1601)	TTGTCCTCACCTGGTGGCGCTTCCACCAGGAGTTGGTGGTCAAGTGAGTGT		
x07518	(1601)	TAGTATTAAACAGGAGCGGCATCCACAGGAGTGGGAGGCCAAGTTAGTGTA		
Consensus	(1601)	T GT T AC GG GC GC TCCAC GGAGT GG GG CAAGT AGTGT	1651	1700
Cry1CGh	(1651)	AATATGCCCTTTCAGAAAGACTATGGAGATTGGAGAAACCTCACTTCCCG		
x07518	(1651)	AATATGCCCTTTCAGAAACTATGGAAATAGGGGAGAACTTAACATCTAG		
Consensus	(1651)	AATATGCC CTTACAGAA ACTATGGA AT GG GA AAC T AC TC G	1701	1750
Cry1CGh	(1701)	TACCTTCAGGTATAACCGACTTTAGTAACCCATTCTCCCTCCGTGCTAATC		
x07518	(1701)	AACATTTAGATATAACCGATTTTAGTAATCCTTTTTCATTTAGAGCTAATC		
Consensus	(1701)	AC TT AG TATACCGA TTTAGTAA CC TT TC TT G GCTAATC	1751	1800
Cry1CGh	(1751)	CTGATATCATTGGTATTAAGTGAGCAACCCCTTTTCGGTGGCGGATCCATC		
x07518	(1751)	CAGATATAATTGGGATAAGTGAAACACCTCTATTGGTGGCAGGTTCTATT		
Consensus	(1751)	C GATAT ATTGG AT AGTGA CAACC CT TT GGTGC GG TC AT	1801	1850
Cry1CGh	(1801)	AGTTCCGGTGAACCTCTACATTGATAAGATTGAGATCATTCTTGCTGACGC		
x07518	(1801)	AGTAGCGGTGAACCTTATATAGATAAAATTGAAATTATTCTAGCAGATGC		
Consensus	(1801)	AGT CGGTGAAC T A AT GATAA ATTGA AT ATTCT GC GA GC	1851	1890
Cry1CGh	(1851)	TACTTTTGAAGCCGAGAGTGATCTCGAAAGGGCTCAGAAG		
x07518	(1851)	AACATTTGAAGCAGAAATCTGATTTAGAAAGAGCACAAAG		
Consensus	(1851)	AC TTTGAAGC GA TGAT T GAAAG GC CA AAG		

FIGURE: 1 (cont'd)



Cry1CGh expression cassette in MLS9124 and MLS9878

2936 bp

FIGURE: 2

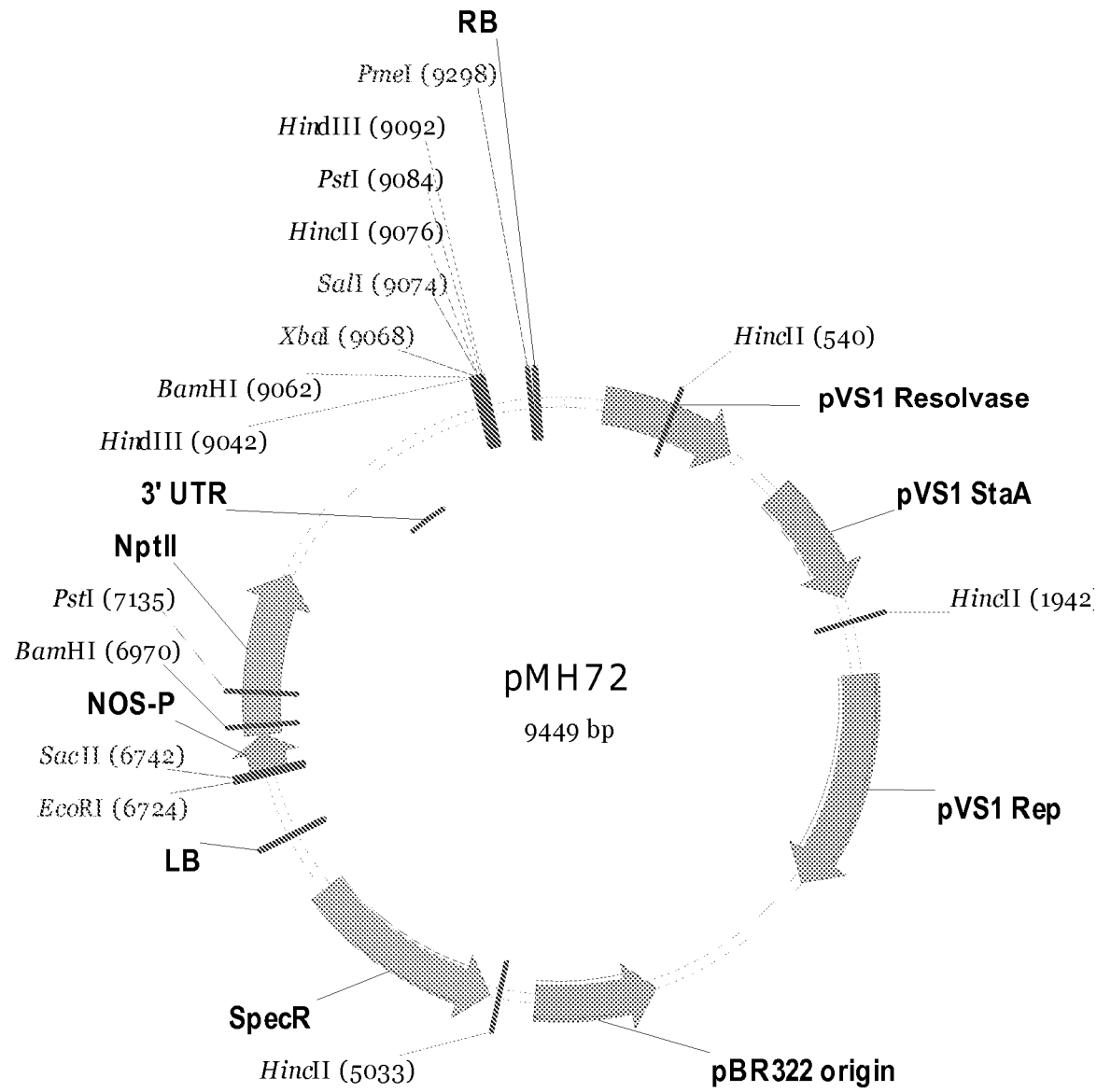


FIGURE: 3

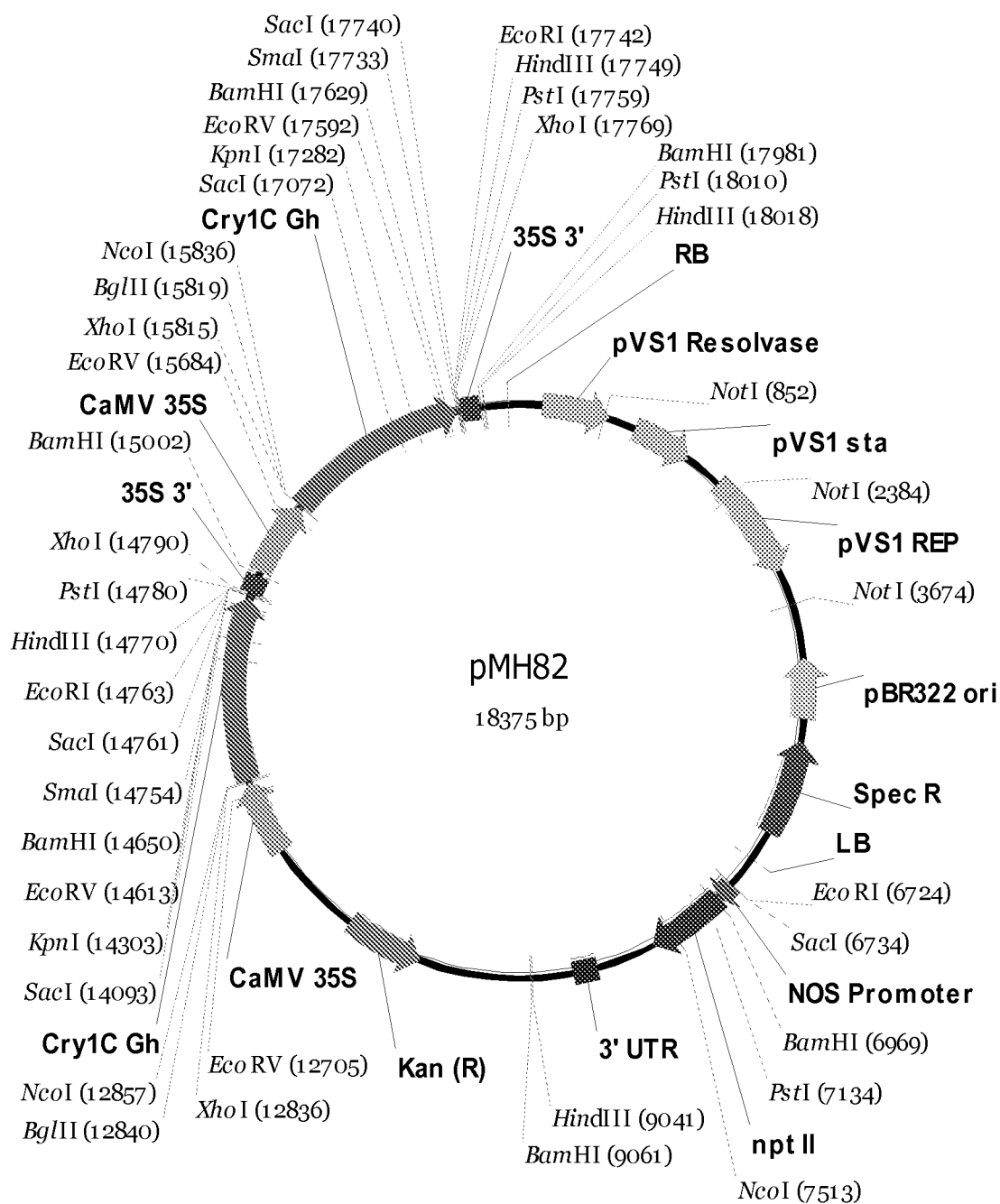


FIGURE: 4

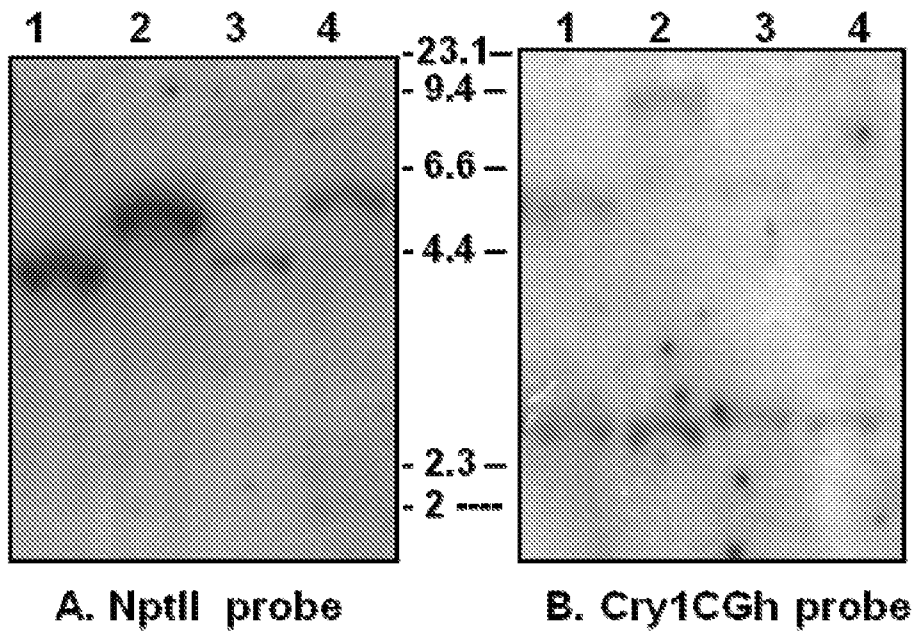


FIGURE: 5

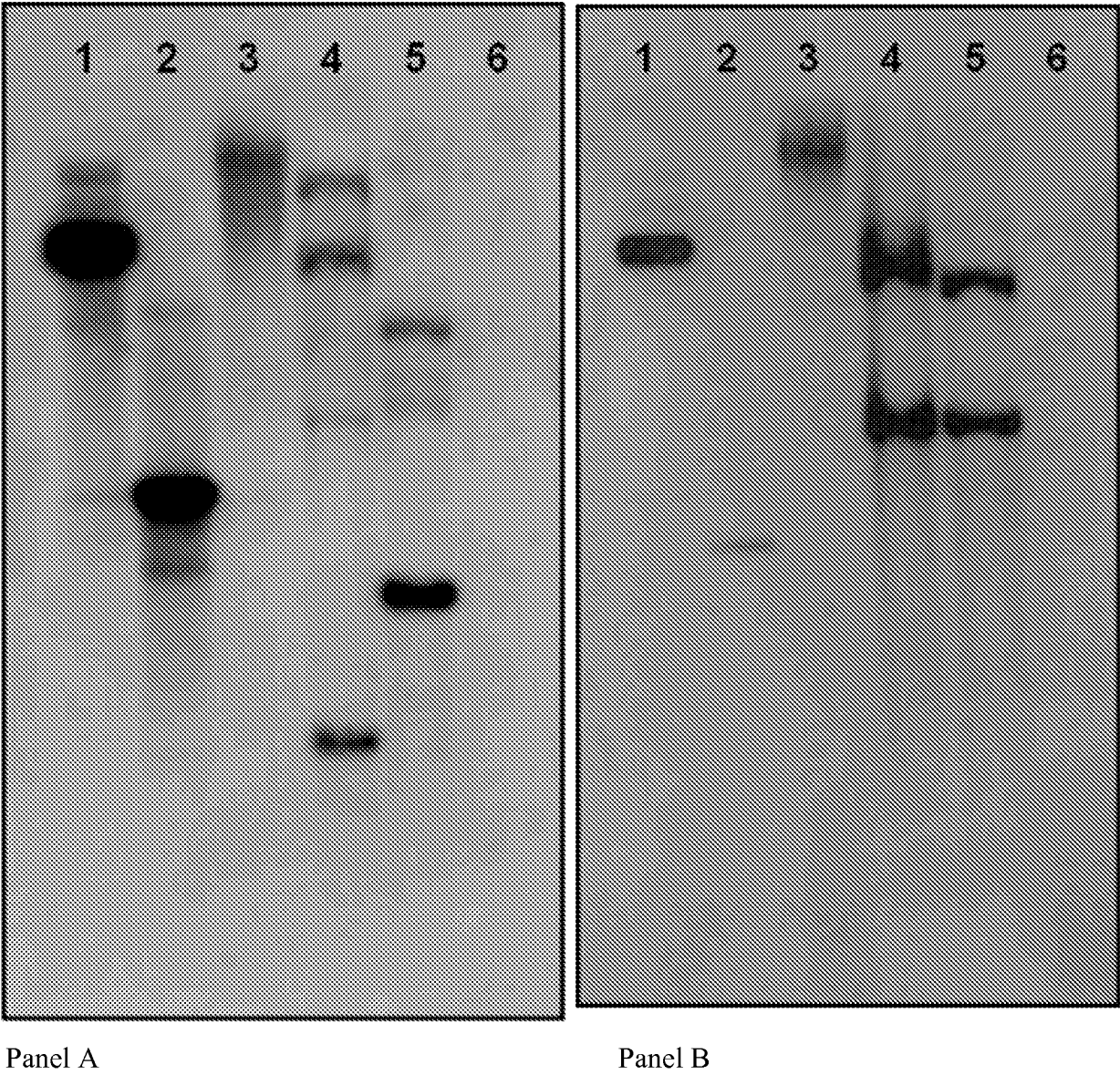


FIGURE: 6

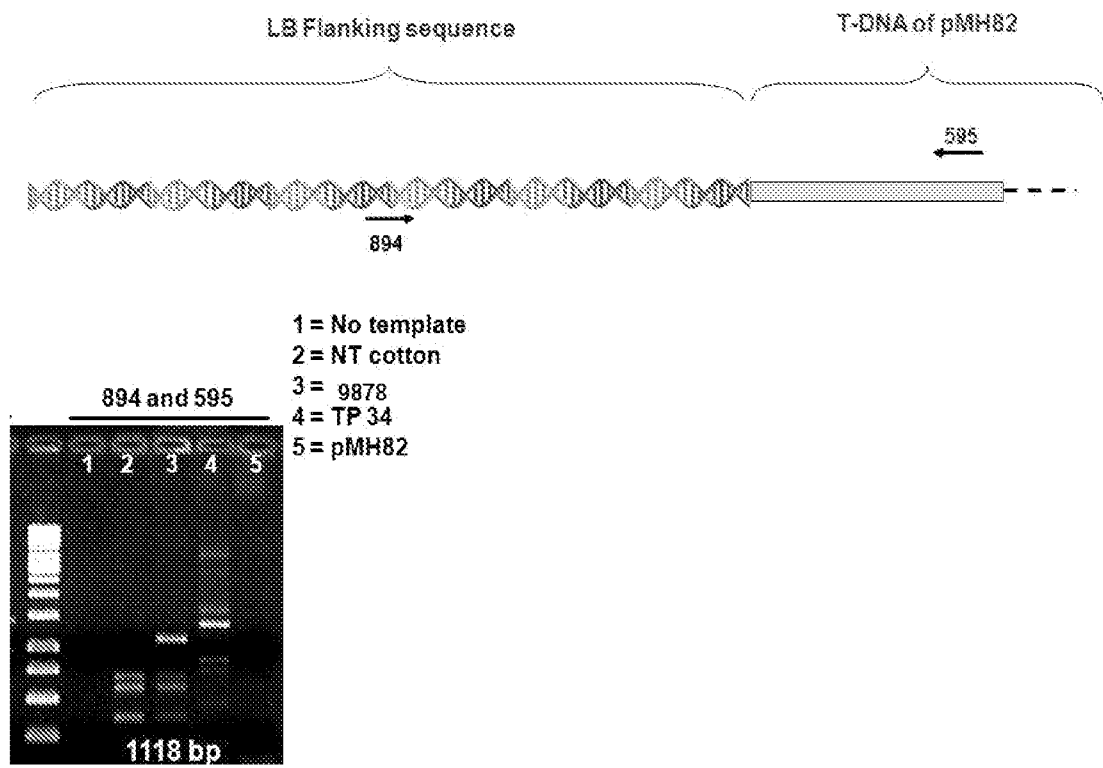


FIGURE: 7

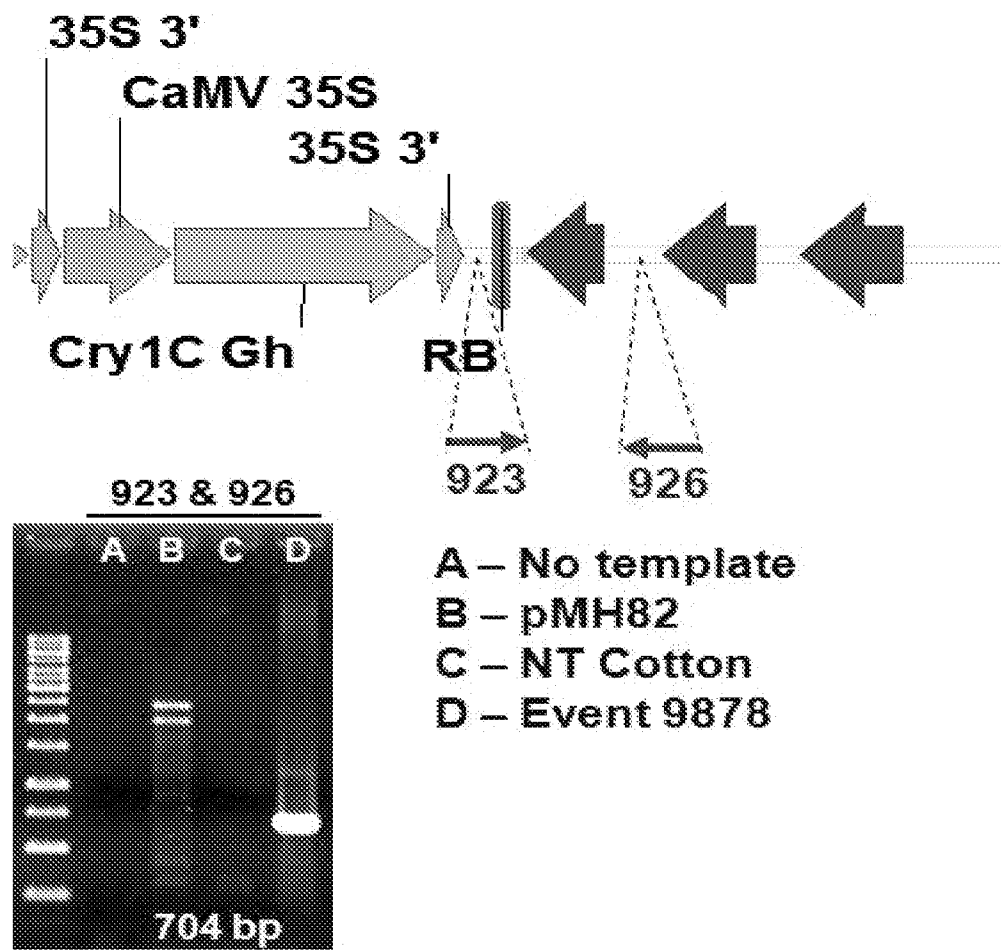


FIGURE: 8

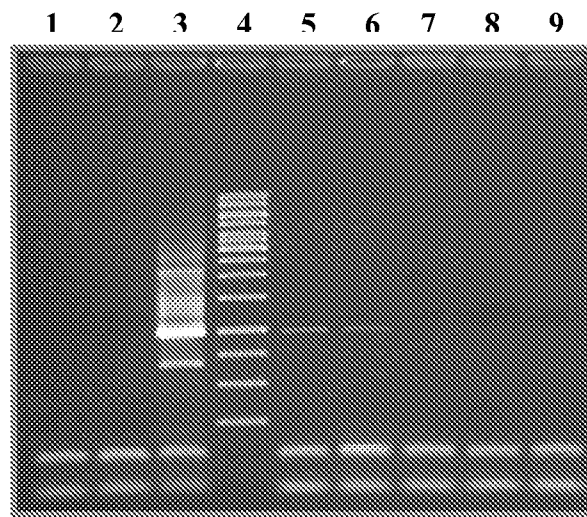
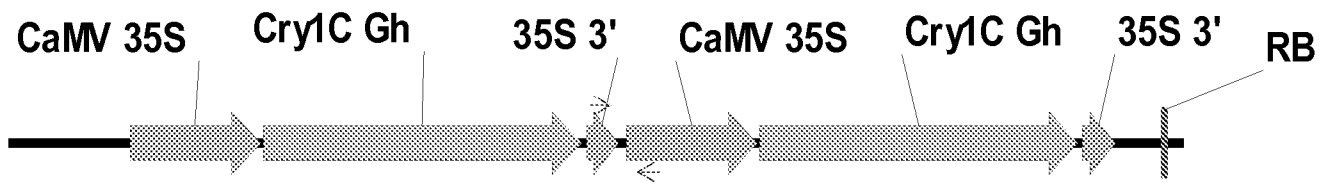
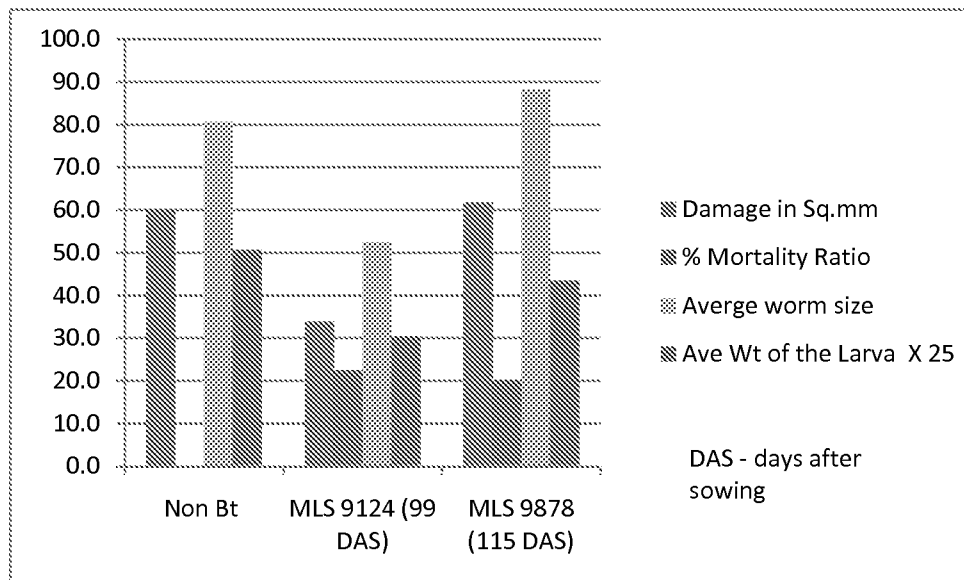


FIGURE: 9

(A)



(B)

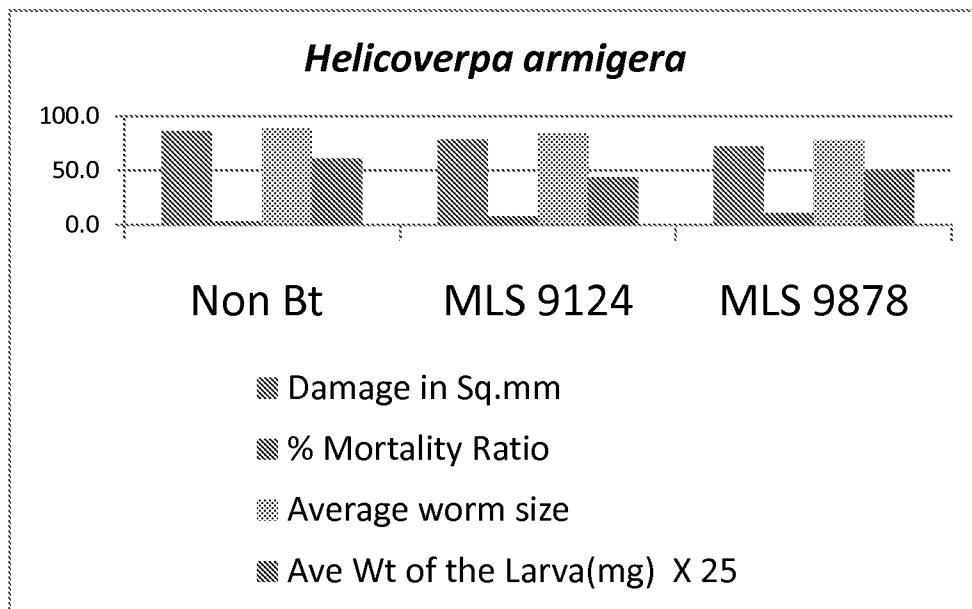


FIGURE: 10

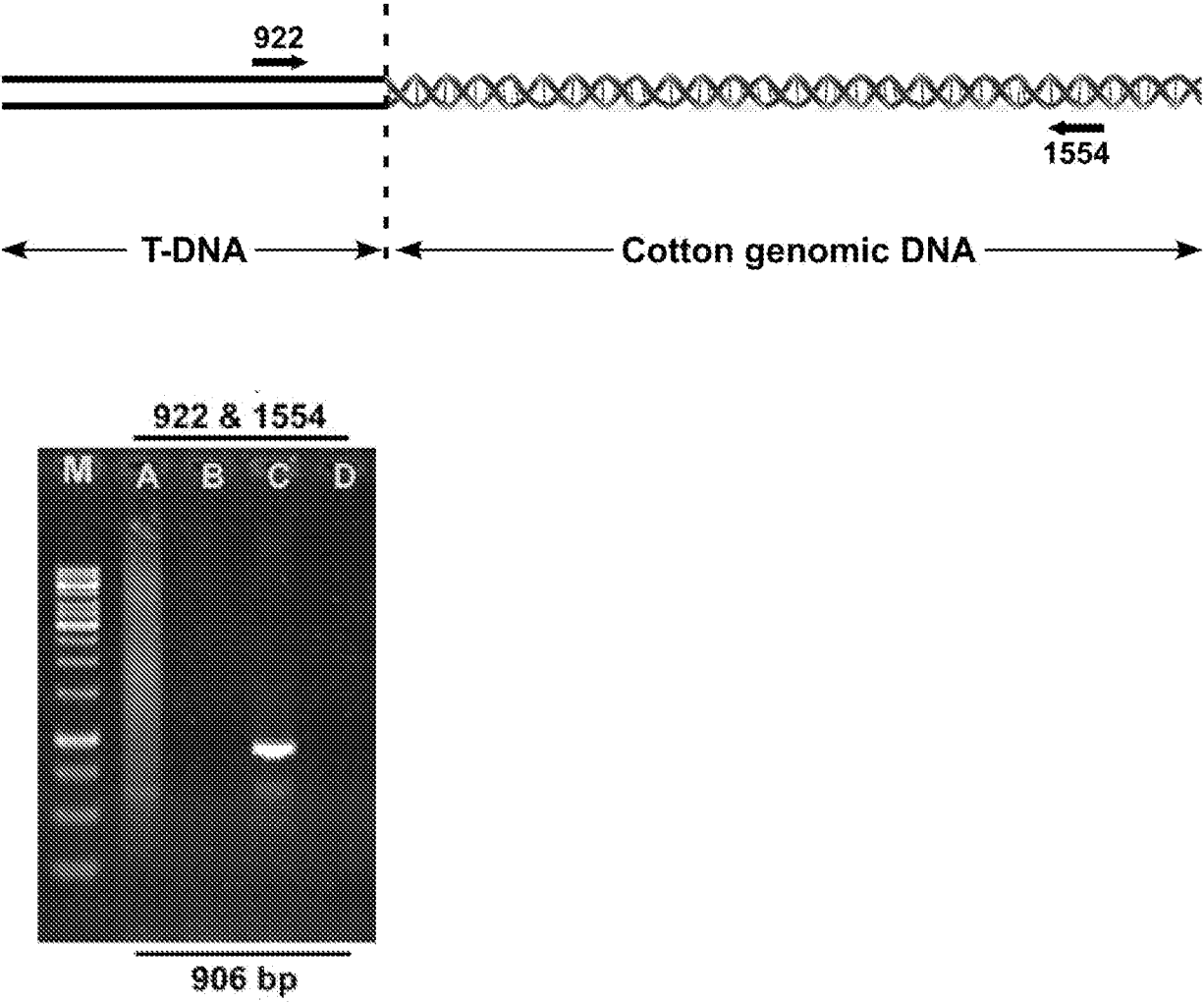


FIGURE: 11

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2012/051521

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/82 C07K14/325
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EP0-Internal, WPI Data, BIOSIS, CHEM ABS Data, Sequence Search, EMBASE, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98/15567 A2 (VITALITY BIOTECHNOLOGIES LTD [IL]; STRIZHOV NICOLAI [DE]; KONCZ CSABA) 16 April 1998 (1998-04-16) claims 1-9; figure 2	1-26
X	EP 0 408 403 A1 (PLANT GENETIC SYSTEMS NV [BE]) 16 January 1991 (1991-01-16) page 18 - page 23; claims 1-19 ----- -/--	1-26

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

26 July 2012

Date of mailing of the international search report

24/08/2012

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Marchesini, Patrizia

INTERNATIONAL SEARCH REPORT

International application No

PCT/IB2012/051521

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	STRIZHOV N ET AL: "A SYNTHETIC CRYIC GENE, ENCODING A BACILLUS THURINGIENSIS DELTA- ENDOTOXIN, CONFERS SPODOPTERA RESISTANCE IN ALFALFA AND TOBACCO", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, NATIONAL ACADEMY OF SCIENCES, US, vol. 93, no. 26, 24 December 1996 (1996-12-24), pages 15012-15017, XP002064600, ISSN: 0027-8424, DOI: 10.1073/PNAS.93.26.15012 the whole document -----	1-26
X	WO 95/06730 A1 (SANDOZ LTD [CH]; SANDOZ AG [DE]; SANDOZ AG [AT]; BOSCH HENDRIK JAN [NL] 9 March 1995 (1995-03-09) claims 1-25 -----	1-26
Y	WO 2009/056155 A1 (BAYER BIOSCIENCE NV [BE]; STEINKELLNER HERTA [AT]; STRASSER RICHARD [A] 7 May 2009 (2009-05-07) sequence 17 -----	1-26
Y	ROH JONG YUL ET AL: "Bacillus thuringiensis as a specific, safe, and effective tool for insect pest control", JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY, vol. 17, no. 4, April 2007 (2007-04), pages 547-559 URL, XP002680840, the whole document -----	1-26

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/IB2012/051521

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9815567	A2	16-04-1998	AR 008885 A1	23-02-2000
			AR 013845 A1	31-01-2001
			AU 4635597 A	05-05-1998
			DE 69711667 D1	08-05-2002
			EP 0935603 A2	18-08-1999
			EP 0948626 A2	13-10-1999
			NZ 335358 A	23-02-2001
			US 6043415 A	28-03-2000
			US 6110668 A	29-08-2000
			WO 9815567 A2	16-04-1998
			WO 9815630 A2	16-04-1998
EP 0408403	A1	16-01-1991	AT 109511 T	15-08-1994
			AU 623530 B2	14-05-1992
			AU 5724590 A	07-01-1991
			CA 2032481 A1	01-12-1990
			CL 14792004 A1	03-06-2005
			DE 69011232 D1	08-09-1994
			DE 69011232 T2	08-12-1994
			DK 0408403 T3	28-11-1994
			EP 0400246 A1	05-12-1990
			EP 0408403 A1	16-01-1991
			ES 2060975 T3	01-12-1994
			HK 1008050 A1	30-04-1999
			JP 3325564 B2	17-09-2002
			JP H04500161 A	16-01-1992
			US 5866784 A	02-02-1999
			US 6172281 B1	09-01-2001
			US 2007150983 A1	28-06-2007
			WO 9015139 A1	13-12-1990
WO 9506730	A1	09-03-1995	AT 329033 T	15-06-2006
			AU 691522 B2	21-05-1998
			AU 7693294 A	22-03-1995
			BR 9407377 A	16-07-1996
			CA 2168011 A1	09-03-1995
			DE 69434758 T2	09-11-2006
			EP 0719335 A1	03-07-1996
			ES 2262136 T3	16-11-2006
			IL 110829 A	20-08-2006
			JP H09502343 A	11-03-1997
			PL 313317 A1	24-06-1996
			RU 2210593 C2	20-08-2003
			TW 430669 B	21-04-2001
			US 5736131 A	07-04-1998
			US 6204246 B1	20-03-2001
			WO 9506730 A1	09-03-1995
			ZA 9406770 A	04-03-1996
WO 2009056155	A1	07-05-2009	AU 2007360682 A1	07-05-2009
			CA 2704108 A1	07-05-2009
			EP 2205728 A1	14-07-2010
			US 2010242128 A1	23-09-2010
			WO 2009056155 A1	07-05-2009