Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2823467 C 2020/10/27

(11)(21) 2 823 467

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2011/12/21
(87) Date publication PCT/PCT Publication Date: 2012/06/28
(45) Date de délivrance/Issue Date: 2020/10/27
(85) Entrée phase nationale/National Entry: 2013/06/20
(86) N° demande PCT/PCT Application No.: US 2011/066526
(87) N° publication PCT/PCT Publication No.: 2012/088289
(30) Priorités/Priorities: 2010/12/22 (US61/426,170); 2011/03/07 (US61/449,776); 2011/12/02 (US61/566,064)

(51) Cl.Int./Int.Cl. *C12Q 1/68* (2018.01), *A01H 1/00* (2006.01), *A01H 1/04* (2006.01), *A01H 5/00* (2018.01), *A01H 6/20* (2018.01), *C12N 15/11* (2006.01), *C12N 15/29* (2006.01), *C12N 15/82* (2006.01), *C12Q 1/6827* (2018.01), *C12Q 1/6895* (2018.01)

(72) Inventeurs/Inventors:
FALAK, IGOR, CA;
PRIMOMO, VALERIO, CA;
TULSIERAM, LOMAS, CA

(73) Propriétaire/Owner:
PIONEER HI-BRED INTERNATIONAL, INC., US

(74) Agent: TORYS LLP

(54) Titre : LOCUS A CARACTERE QUANTITATIF ASSOCIES A LA RESISTANCE DE CHAMP ENTIER AU SCLEROTINIA ET PROCEDES D'IDENTIFICATION DE CETTE RESISTANCE

(54) Title: QTLS ASSOCIATED WITH AND METHODS FOR IDENTIFYING WHOLE PLANT FIELD RESISTANCE TO SCLEROTINIA

(57) Abrégé/Abstract:
Markers associated with Sclerotinia whole plant field resistance are provided. Methods of identifying Sclerotinia resistant and susceptible plants, using the markers are provided. Methods for identifying and isolating QTLs are a feature of the invention, as are QTLs associated with Sclerotinia whole plant field resistance.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 June 2012 (28.06.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/088289 A3

(51) International Patent Classification:
C12Q 1/68 (2006.01) *A01H 5/10* (2006.01)

(21) International Application Number:
PCT/US2011/066526

(22) International Filing Date:
21 December 2011 (21.12.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/426,170 22 December 2010 (22.12.2010) US
61/449,776 7 March 2011 (07.03.2011) US
61/566,064 2 December 2011 (02.12.2011) US

(71) Applicants *(for all designated States except US)*: **E. I. du Pont de Nemours and Company** [US/US]; 1007 Market Street, Wilmington, Delaware 19898 (US). **PIONEER HI-BRED INTERNATIONAL, INC.** [US/US]; 800 Capital Square 400 Locust Street, Des Moines, Iowa 50309 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **FALAK, Igor** [CA/CA]; 53 Keats Crescent, Guelph, ontario N1G 3B1 (CA). **PRIMOMO, Valerio** [CA/CA]; 6 Marshall Boulevard, Toronto, ontario M6N 2R7 (CA). **TULSIERAM, Lomas** [CA/CA]; 2795 Gulfstream Way, Mississauga, ontario L52 6K2 (CA).

(74) Agent: **CATALANO, Christina, M.**; E. I. du Pont de Nemours and Company, Legal Patent Records Center, 4417 Lancaster Pike, Wilmington, Delaware 19805 (US).

(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 (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, 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).

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))*

(88) Date of publication of the international search report:
27 December 2012

WO 2012/088289 A3

(54) Title: QTLs ASSOCIATED WITH AND METHODS FOR IDENTIFYING WHOLE PLANT FIELD RESISTANCE TO SCLEROTINIA

(57) Abstract: Markers associated with *Sclerotinia* whole plant field resistance are provided. Methods of identifying *Sclerotinia* resistant and susceptible plants, using the markers are provided. Methods for identifying and isolating QTLs are a feature of the invention, as are QTLs associated with *Sclerotinia* whole plant field resistance.

QTLs ASSOCIATED WITH AND METHODS FOR IDENTIFYING WHOLE PLANT FIELD RESISTANCE TO *SCLEROTINIA*

Cross-Reference to Related Applications

Benefit is claimed under 35 U.S.C. § 1.19(e) to the filing dates of U.S. Provisional Application No. 61,426,170, filed December 22, 2010, U.S. Provisional Application No. 61/449,776, filed March 7, 2011 and U.S. Provisional Application No. 61/566,064, filed December 2, 2011 .

Field of the Invention

The present invention relates generally to plant molecular biology. More specifically, it relates to quantitative trait loci (QTLs) associated with whole plant field resistance to *Sclerotinia* in *Brassica*, and use of those QTLs to identify whole plant field resistance to *Sclerotinia* in *Brassica* and other plant species.

Background of the Invention

Sclerotinia infects over 400 species of plants throughout Canada, including numerous economically important crops such as *Brassica* species, sunflowers, dry beans, field peas, lentils, and potatoes (Boland and Hall (1994) Can. J. Plant Pathol. 16:93-108). *Sclerotinia sclerotiorum* is responsible for over 99% of the disease, while *Sclerotinia minor* produces less than 1% of the disease. *Sclerotinia* produces sclerotia, which are irregularly shaped dark overwintering bodies that can endure in soil for four to five years. The sclerotia can germinate carpogenically or myceliogenically depending on the environmental conditions and crop canopies. The two types of germination cause two distinct types of diseases. Sclerotia that germinate carpogenically produce apothecia and ascospores that infect above-ground tissues, resulting in stem blight, stalk rot, head rot, pod rot, white mold and blossom blight of plants. Sclerotia that germinate myceliogenically produce mycelia that can infect root tissues, causing crown rot, root rot and basal stalk rot.

Sclerotinia causes *Sclerotinia* stem rot, also known as white mold, in *Brassica*, including canola. Canola is a type of *Brassica* having a low level of glucosinolates and

erucic acid in the seed. The sclerotia germinate carpogenically in the summer, producing apothecia. The apothecia release wind-borne ascospores that travel up to one kilometer. The disease is *favor*ed by moist soil conditions (at least 10 days at or near field capacity) and temperatures of 15-25°C, prior to and during canola flowering. The spores cannot infect leaves and stems directly. They must first land on flowers, fallen petals, and pollen on the stems and leaves. Petal age affects the efficiency of infection, with older petals better able to effect infection (Heran et al. (1999) "The Effect of Petal Characteristics, Inoculum Density and Environmental Factors on Infection of Oilseed Rape by *Sclerotinia sclerotiorum*" The Regional Institute Ltd. http://www.regional.org.au/au/gcirc/ 3/428.htm). The fungal spores use the flower parts as a food source to germinate and infect the plant.

Brassica can also develop root rot under certain conditions. For example, winter and spring canola occasionally develop root rot during mild winters in Europe (winter canola) and in Georgia, US (spring canola).

The severity of *Sclerotinia* in *Brassica* is variable, and is dependent on the time of infection and climatic conditions (Heran et al., *supra*). The disease is *favor*ed by cool temperatures and prolonged periods of precipitation. Temperatures between 20 and 25°C and relative humidities of greater than 80% are required for optimal plant infection (Heran et al., *supra*). Losses ranging from 5 to 100% have been reported for individual fields (Manitoba Agriculture, Food and Rural Initiatives, 2004). On average, yield losses equal 0.4 to 0.5 times the percentage infection. For example, if a field has 20% infection (20/100 infected plants), then the yield loss would be about 10%. Further, *Sclerotinia* can cause heavy losses in wet swaths.

The symptoms of *Sclerotinia* infection usually develop several weeks after flowering begins. The plants develop pale-grey to white lesions, at or above the soil line and on upper branches and pods. The infections often develop where the leaf and the stem join because the infected petals lodge there. Infected stems appear bleached and tend to shred. Hard black fungal sclerotia develop within the infected stems, branches, or pods. Plants infected at flowering produce little or no seed. Plants with girdled stems wilt and ripen prematurely. Severely infected crops frequently lodge, shatter at swathing, and make swathing more time consuming. Infections can occur in all above ground plant parts especially in dense or lodged stands. Once plants are infected, the mold continues to grow into the stem and invade healthy tissue. New sclerotia are formed to carry the disease over to the next season.

Some varieties of canola with certain morphological traits are better able to withstand *Sclerotinia* infection. For example, Polish varieties *(Brassica rapa)* have lighter canopies and seem to have much lower infection levels. In addition, petal-less varieties (apetalous varieties) do not provide the initial infection source (i.e., the flower petal) and avoid *Sclerotinia* infection to a greater extent (Okuyama et al. (1995) Bulletin of the Tohoku National Agricultural Experiment Station. National Agriculture Research Center, Tsukuba, Ibaraki 305, JAP 3-1-1an. 89: 11-20; Fu (1990) Acta Agriculture Shanghai. Economic Crop Research Institute, Jiangsu Province Academy of Agricultural Sciences, Nanjing 210024, China 6 (3): 76-77. Other examples of morphological traits that confer a degree of reduced susceptibility to *Sclerotinia* in *Brassica* include increased standability, lower petal retention, higher branching (both extent and position), flowering (early start and/or short duration) and early leaf abscission. Jurke and Fernando ("Plant Morphology of Canola and its Effects on *Sclerotinia sclerotiorum* infection in ICPP" 2003 8th International Congress of Plant Pathology, New Zealand) screened eleven canola genotypes for *Sclerotinia* disease incidence. Significant variation in disease incidence was explained by plant morphology and the difference in petal retention was identified as the most important factor. However, these morphological traits alone do not confer resistance to *Sclerotinia* and most canola lines in Canada are considered susceptible to *Sclerotinia*.

The primary means of controlling *Sclerotinia* in infected canola crops is by spraying with fungicide. Typical fungicides used for controlling *Sclerotinia* on *Brassica* include Rovral™/Proline from Bayer and Ronilan™/Lance™ from BASF. If infection is already evident, there is no use in applying fungicide as it is too late to have an effect. Accordingly, growers must assess their fields for disease risk to decide whether to apply a fungicide. This can be done by using a government provided checklist or by using a petal testing kit. Either way, the method is cumbersome and prone to errors.

Numerous efforts have been made to develop *Sclerotinia* resistant spring *Brassica* plants. Built in resistance would be more convenient, economical, and environmentally friendly compared to controlling *Sclerotinia* by application of fungicides. Since the trait is polygenic it would be stable and not prone to changes in efficacy, as fungicides may be. Winter canola is also susceptible to *Sclerotinia*.

Spring canola (*Brassica napus* subsp. oleifera var. annua) differs from winter canola (*Brassica napus* subsp. oleifera var. biennis) primarily in the absence of an obligate

vernalization requirement. Asiatic rapeseed and canola versions have a low to intermediate requirement for vernalization. While winter canola cannot finish its reproduction cycle when planted in the spring, Asiatic material cannot finish its reproduction cycle if planted in late spring, but early spring planting and exposure to cold enables Asiatic material to flower and set seed. In controlled conditions winter material requires 12-14 weeks of vernalization while Asiatic material requires 2-8 weeks. Table 1 summarizes the differences between winter, semi-winter (Asiatic) and spring canola varieties.

Table 1. Main determinations of growth habit in *Brassica napus* materials

Type	Spring*	Spring	Semi Winter	Winter
Growing areas	Canada, Europe	Australia	China, Japan	Europe
Vernalization Requirement	None	None	2-8 weeks Intermediate	12-14 weeks strong or full
Time of seeding	Spring (Increasing Day Length)	Fall (Decreasing Day Length)	Fall (Decreasing Day Length)	Fall (Decreasing Day Length)
Number of days until flowering	30-90	90-150	120-180	150-270

* Canadian, European and Australian spring materials can be planted and grown in any environment or seeding time for spring canola.

Some Chinese cultivars of rapeseed/canola are partially resistant to *Sclerotinia*. For example, ChunYun et al. ((2003) Acta Agronomica Sinica 29 (5): 715-718); HanZhong et al. ((2004) Scientia Agricultura Sinica 37 (1): 23-28); WeiXin et al. ((2002) Chinese Journal of Oil Crop Sciences 24 (3): 47-49); YongJu et al. ((2000) Chinese Journal of Oil Crop Sciences 22 (4): 1-5) describe partially resistant varieties of rapeseed. However, some of these varieties are not canola quality and all of them require vernalization. The partial field resistance in Chinese varieties originated from the rapeseed variety Zhong you 821. Despite improvements in partial resistance in Zhong you 821, its reaction to pathogens is less stable under environmental conditions favorable for development of *Sclerotinia* (Li et al. (1999) "Breeding, inheritance, and biochemical studies on *Brassica napus* cv. Zhongyou 821: Tolerance to *Sclerotinia sclerotiorum* (stem rot)". Proceedings of the 10th International

Rapeseed Congress, Canberra, Australia).

Some Japanese cultivars of rapeseed have partial stem resistance to *Sclerotinia*. Partial stem resistance was detected by indoor tests in comparison with winter canola (Brun et al. (1987) "A field study of rapeseed (*Brassica napus*) resistance to *Sclerotinia sclerotiorum*." 7th International Rapeseed Congress, Poznan, Poland). However, these varieties are not canola quality and are semi-winter types (see Table 1).

Breeding for *Sclerotinia* resistance in canola has been very difficult due to the quantitative nature of this trait. Further, the incorporation of physiological resistance with morphological traits that avoid or reduce infection multiplies the complexity of breeding for resistance. In addition, it has been very difficult to screen for resistance because of the direct environment by genetic (GXE) interaction (i.e., temperature and humidity requirements, as well as microenvironment requirements) with the plant. As stated above, there are few Canadian spring *Brassica* varieties with resistance to *Sclerotinia*, this despite many years of co-evolution and environmental pressure to select for this trait. A level of field resistance in rapeseed (and recently some canola materials) was attained via breeding efforts in China as described with Zhong you 821 (Li et al., *supra*). However, the levels of such partial resistance or tolerance are relatively low and fungicide applications are still recommended on all rapeseed and canola materials in China (verbal communication) (Hu et al. (1999) "Effect of cultural control on rapeseed stem rot (*Sclerotinia sclerotiorum*) in *Brassica napus*." Proceedings of the 10th International Rapeseed Congress, Canberra, Australia). Other breeding efforts included quantitative trait loci analysis (Zhao and Meng (2003) Theoretical and Applied Genetics 106 (4): 759-764), mutagenesis breeding (Mullins et al. (1999) European Journal of Plant Pathology 105 (5): 465-475; Wu et al. (1996) Sichuan Daxue Xuebao (Ziran Kexueban) 33 (2): 201-205; LiangHong et al., 2003, extensive screening efforts (Sedun et al. (1989) Canadian Journal of Plant Science 69 (1): 229-232; Zhao et al. (2004) Plant Disease 88 (9): 1033-1039); and screening for expressed sequence tags (ESTs) (Li et al. (2004) Fungal Genetics and Biology 41 (8): 735-753) to name a few. Several spring canola varieties with moderate tolerance to *Sclerotinia* have been developed (Ahmadi et al. (2000) Seed and Plant 16 (1): Pe127-Pe129, en14; Ahmadi et al. (2000) Introduction of rapeseed (*Brassica napus* L.), cultivar Esteghlal. Seed and Plant 16 (1): Pe127-Pe126, en13; BaoMing et al. (1999) Chinese Journal of Oil Crop Sciences 4: 12-14; and Liu et al. (1991) Scientia Agricultura Sinica 24 (3): 43-49), however the level of

tolerance is low and the lines cannot withstand high disease pressure. Recently, transgenic canola has been developed carrying an oxalic oxidase gene (U.S. Patent No. 6,166,291 and divisional patents thereof) however there are regulatory and social problems associated with transgenic plants. Accordingly, significant technical human intervention is required to breed canola varieties that are resistant to *Sclerotinia*.

More recently, *Brassica* and canola varieties with high levels of resistance to *Sclerotinia* were developed after a long and intensive breeding program (See, for example, WO 2006/135717). This approach is very time and labor intensive, and requires a long time to determine whether the breeding program is successful. The difficulty in breeding for whole plant field resistance to *Sclerotinia* is due, at least in part, to the multigenic nature of this trait.

What is needed in the art and industry is a means to identify genes conferring whole plant field resistance to *Sclerotinia*, using molecular markers. These markers can then be used to tag the favorable alleles of these genes in segregating populations and then employed to make selection for resistance more effective. The present invention provides this and other advantages.

Summary of the Invention

The present invention provides methods and markers for identifying Quantitative Trait Loci ("QTLs") associated with whole plant field resistance or improved whole plant field resistance to *Sclerotinia* in plants.

A first aspect of the invention features a method of identifying a *Brassica* plant or germplasm that exhibits whole plant field resistance or improved whole plant field resistance to *Sclerotinia*. The method comprises detecting in the plant or germplasm at least one allele of at least one quantitative trait locus (QTL) that is associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, wherein the QTL is localized to a linkage group selected from N1, N3, N4, N7, N8, N9, N10, N11, N12, N13, N15, N18 or N19, wherein each linkage group comprises at least one marker that is associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia* with a statistical significance of $p \leq 0.01$, thereby identifying the *Brassica* plant or germplasm that exhibits whole plant field resistance or improved whole plant field resistance to *Sclerotinia*.

In one embodiment, the QTL is localized to a chromosomal interval selected from: (a) an interval flanked by and including (i) markers CA0614 and PE0177 or (ii) markers AG0093 and AG0482 on linkage group N1; (b) an interval flanked by and including markers CA0410 and AG0023 on linkage group N3; (c) an interval flanked by and including markers BG1442 and BG0106 on linkage group N4; (d) an interval flanked by and including markers AG0510 and CA0105 on linkage group N7; (e) an interval flanked by and including markers CA0837 and BG1286 on linkage group N8; (f) an interval flanked by and including (i) markers CA1034 and AG0441or (ii) markers AG0378 and KK66 on linkage group N9; (g) an interval flanked by and including markers BG0228 and PE0131 on linkage group N10; (h) an interval flanked by and including (i) markers CA0120 and CA0163 or (ii) markers CA0120 and CA1097 on linkage group N11; (i) an interval flanked by and including (i) markers BG1321 and CA0991 or (ii) markers CA0753 and PE0250 on linkage group N12; (j) an interval flanked by and including markers CA0603 and CA0736 on linkage group N13; (k) an interval flanked by and including markers PE0286 and AG0369 on linkage group N15; (l) an interval flanked by and including (i) markers BG0278 and CA0636 or (ii) markers UB0315 and CA0739 on linkage group N18; and (m) an interval flanked by and including (i) markers CA1107 and CA0221 or (ii) markers UB0307 and KK98G on linkage group N19.

In another embodiment, the QTL is localized to a chromosomal interval selected from: (a) one or more intervals on linkage group N1, flanked by and including markers (i) AG0093 and PE0203, or (ii) BG0111 and BG1392, or (iii) BG1090 and AG0482, or (iv) BG1090 and PE0203, or (v) CA0614 and BG1392, or (vi) BG0988 and AG0482; or (vii) AG0243 and AG0482; or (viii) AG0243 and BG1453; or BG0988; (b) one or more intervals on linkage group N3, flanked by and including markers (i) BG1197 and AG0023, or (ii) CA0410 and BG1368 or (iii) CA0410 and BG1197; (c) one or more intervals on linkage group N4, flanked by and including markers (i) BG1442 and BG0106, or (ii) UB0181 and BG0106; (d) one or more intervals on linkage group N8, flanked by and including markers (i) BG1449 and BG1062, or (ii) CA0837 and AG0328, or (iii) CA0837 and BG1062, or (iv) CA0837 and BG1101, or (v) CA0837 and BG1286, or (vi) CA0837 and BG1449 or (vii) PE0281 and BG0647; (e) one or more intervals on linkage group N9, flanked by and including markers (i) AG0323 and BG0295, or (ii) CA1034 and AG0378 or (iii) BG1123 and AG0441; (f) one or more intervals on linkage group N10, flanked by and including

markers (i) BG0228 and AG0047, or BG0255 and PE0131; (g) one or more intervals on linkage group N11, flanked by and including markers (i) BG0031 and BG1149, or (ii) BG0031 and BG1230, or (iii) BG0031 and BG1513, or (iv) CA0120 and CA0328, or (v) PE0283 and CA0163, or (vi) PE0324 and PE0283 or (vii) CA0328 and PE0324, or (viii) CA0226 and BG0713, or (ix) CA0233 and CA1080, or (x) CA0233 and AG0370; (h) one or more intervals on linkage group N12, flanked by and including markers (i) BG1321 and CA0991, or (ii) BG1321 and CA1027, or (iii) BG1321 and PE0133, or (iv) PE0063 and CA0991, or (v) PE0133 and CA0991, or (vi) CA1027 and PE0063, or (vii) CA1027 and UB0331, or (viii) CA0423 and PE0250, or (ix) AG0359 and PE0250, or (x) AG0359 and CA0896; (i) one or more intervals on linkage group N13, flanked by and including markers (i) BG0516 and AG0148, or (ii) CA0488 and AG0148, or (iii) CA0488 and CA0736, or (iv) CA0603 and AG0504, or (v) BG1288 and AG0504; (j) one or more intervals on linkage group N15, flanked by and including markers (i) CA0719 and AG0369, or (ii) PE0091 and PE0187, or (iii) PE0286 and AG0369, or (iv) PE0286 and PE0187, or (v) PE0286 and CA0719; (k) one or more intervals on linkage group N18, flanked by and including markers (i) AG0285 and CA0636, or (ii) BG0278 and CA07739, or (iii) CA0739 and CA0636, or (iv) UB0315 and CA0636, or (v) UB0315 and CA0739; and (l) one or more intervals on linkage group N19, flanked by and including markers (i) CA0552 and CA0221, or (ii) CA1107 and CA0552, or (iii) CA1107 and CA0221, or (iv) CA0221 and KK98G, or (v) UB0307 and BG1241, or (vi) BG1241 and KK98G, or (vii) CA0221 and BG1241..

In a particular embodiment, the QTL is localized to a chromosomal interval on linkage group N1, N9, N11, N12, N18 or N19.

In other embodiments, the marker comprises a polymorphism that identifies the at least one allele of the at least one quantitative trait locus (QTL) as being associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, and the detecting comprises identifying the polymorphism. The polymorphism may be, for example, a single nucleotide polymorphism (SNP) or a simple sequence repeat (SSR). In another embodiment of the method of the invention, the detecting comprises detecting at least one marker comprising the polymorphism, selected from AG0023; AG0045; AG0047; AG0070; AG0086; AG0093; AG0125; AG0148; AG0171; AG0203; AG0239; AG0243; AG0272; AG0304; AG0323; AG0324; AG0328; AG0359; AG0369; AG0370; AG0378; AG0391; AG0410; AG0441; AG0477; AG0482; AG0504; AG0510; BG0031; BG0106; BG0111;

BG0119; BG0181; BG0228; BG0255; BG0278; BG0295; BG0452; BG0516; BG0647; BG0651; BG0713; BG0864; BG0869; BG0988; BG1062; BG1090; BG1101; BG1123; BG1127; BG1149; BG1182; BG1197; BG1230; BG1241; BG1244; BG1286; BG1288; BG1321; BG1368; BG1392; BG1442; BG1449; BG1453; BG1513; CA0105; CA0120; CA0163; CA0221; CA0226; CA0233; CA0328; CA0410; CA0423; CA0456; CA0488; CA0546; CA0552; CA0603; CA0614; CA0636; CA0681; CA0719; CA0736; CA0739; CA0753; CA0834; CA0837; CA0896; CA0991; CA1027; CA1032; CA1034; CA1035; CA1066; CA1080; CA1090; CA1097; CA1107; PE0012; PE0017; PE0063; PE0091; PE0131; PE0133; PE0177; PE0187; PE0203; PE0250; PE0281; PE0283; PE0286; PE0324; PE0340; PE0355; UB0015; UB0126; UB0163; UB0181; UB0196; UB0307; UB0315; UB0331; KK66; and KK98G. .

In another embodiment of the method of the invention, the detecting comprises detecting the polymorphism in at least one marker selected from AG0093; AG0304; AG0378; AG0391; AG0482; BG1149; BG1230; BG1241; BG1453; BG1513; CA0120; CA0221; CA0546; CA0739; CA1027; PE0063; PE0203; UB0163; and UB0315.

In other embodiments, the method comprises detecting two or more markers located in two or more different linkage groups, three or more markers located in three or more different linkage groups, four or more markers located in four or more different linkage groups, five or more markers located in five or more different linkage groups, six or more markers located in six or more different linkage groups, seven or more markers located in seven or more different linkage groups, eight or more markers located in eight or more different linkage groups, nine or more markers located in nine or more different linkage groups, ten or more markers located in ten or more different linkage groups, eleven or more markers located in eleven or more different linkage groups, or twelve or more markers located in twelve or more different linkage groups.

In other embodiments, in the method, the detecting comprises amplifying the marker from genomic DNA of the plant or germplasm and determining if the marker comprises the polymorphism associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia*. In other embodiments, the plant is *Brassica napus*; *Brassica juncea*; *Brassica rapa*; *Brassica oleracea*; or *Brassica carinata*. In other embodiments, the plant is spring canola, winter canola, or semi-winter canola. In another embodiment, the whole plant field resistance or improved whole plant field resistance results from decreased

disease incidence compared to a plant lacking the allele of the QTL associated with the whole plant field resistance or improved whole plant field resistance. In another embodiment, the whole plant field resistance or improved whole plant field resistance results from decreased disease severity compared to a plant lacking the allele of the QTL associated with the whole plant field resistance or improved whole plant field resistance. In another embodiment, the plant has whole plant field resistance or improved whole plant field resistance to *Sclerotinia sclerotiorum*.

Another aspect of the invention features a method of introgressing *Sclerotinia* resistance in a second plant by cross pollinating the plant or a progeny identified according to the methods described above with a second plant, wherein the second plant lacks the at least one allele of the at least one QTL detected in the identified plant.

In another aspect, the invention features a method of producing an F1 hybrid seed, wherein the F1 hybrid plant derived from the F1 hybrid seed is resistant to *Sclerotinia*, the method comprising cross pollinating the plant or progeny identified according to the methods described above with a second plant, wherein the second plant lacks the at least one allele of the at least one QTL detected in the identified plant.

In another aspect, the invention features a method of positional cloning of a nucleic acid comprising a quantitative trait locus (QTL) associated with *Sclerotinia* whole plant field resistance or improved whole plant field resistance, the method comprising: providing a nucleic acid from a plant comprising a marker that is associated with *Sclerotinia* whole plant field resistance or improved whole plant field resistance with a statistical significance of $p \leq 0.01$, wherein the QTL is localized to a linkage group selected from N1, N3, N4, N7, N8, N9, N10, N11, N 12, N13, N15, N18 or N19, and wherein the linkage group comprises the marker; and cloning the nucleic acid comprising a quantitative trait locus (QTL) associated with *Sclerotinia* whole plant field resistance or improved whole plant field resistance. (a) an interval flanked by and including (i) markers CA0614 and PE0177 or (ii) markers AG0093 and AG0482 on linkage group N1; (b) an interval flanked by and including markers CA0410 and AG0023 on linkage group N3; (c) an interval flanked by and including markers BG1442 and BG0106 on linkage group N4; (d) an interval flanked by and including markers AG0510 and CA0105 on linkage group N7; (e) an interval flanked by and including markers CA0837 and BG1286 on linkage group N8; (f) an interval flanked by and including (i) markers CA1034 and AG0441or (ii) markers AG0378 and KK66 on linkage group N9; (g) an interval

flanked by and including markers BG0228 and PE0131 on linkage group N10; (h) an interval flanked by and including (i) markers CA0120 and CA0163 or (ii) markers CA0120 and CA1097 on linkage group N11; (i) an interval flanked by and including (i) markers BG1321 and CA0991 or (ii) markers CA0753 and PE0250 on linkage group N12; (j) an interval flanked by and including markers CA0603 and CA0736 on linkage group N13; (k) an interval flanked by and including markers PE0286 and AG0369 on linkage group N15; (l) an interval flanked by and including (i) markers BG0278 and CA0636 or (ii) markers UB0315 and CA0739 on linkage group N18; and (m) an interval flanked by and including (i) markers CA1107 and CA0221 or (ii) markers UB0307 and KK98G on linkage group N19.

In another embodiment, the QTL is localized to a chromosomal interval selected from: (a) one or more intervals on linkage group N1, flanked by and including markers (i) AG0093 and PE0203, or (ii) BG0111 and BG1392, or (iii) BG1090 and AG0482, or (iv) BG1090 and PE0203, or (v) CA0614 and BG1392, or (vi) BG0988 and AG0482; or (vii) AG0243 and AG0482; or (viii) AG0243 and BG1453; or BG0988; (b) one or more intervals on linkage group N3, flanked by and including markers (i) BG1197 and AG0023, or (ii) CA0410 and BG1368 or (iii) CA0410 and BG1197; (c) one or more intervals on linkage group N4, flanked by and including markers (i) BG1442 and BG0106, or (ii) UB0181 and BG0106; (d) one or more intervals on linkage group N8, flanked by and including markers (i) BG1449 and BG1062, or (ii) CA0837 and AG0328, or (iii) CA0837 and BG1062, or (iv) CA0837 and BG1101, or (v) CA0837 and BG1286, or (vi) CA0837 and BG1449 or (vii) PE0281 and BG0647; (e) one or more intervals on linkage group N9, flanked by and including markers (i) AG0323 and BG0295, or (ii) CA1034 and AG0378 or (iii) BG1123 and AG0441; (f) one or more intervals on linkage group N10, flanked by and including markers (i) BG0228 and AG0047, or BG0255 and PE0131; (g) one or more intervals on linkage group N11, flanked by and including markers (i) BG0031 and BG1149, or (ii) BG0031 and BG1230, or (iii) BG0031 and BG1513, or (iv) CA0120 and CA0328, or (v) PE0283 and CA0163, or (vi) PE0324 and PE0283 or (vii) CA0328 and PE0324, or (viii) CA0226 and BG0713, or (ix) CA0233 and CA1080, or (x) CA0233 and AG0370; (h) one or more intervals on linkage group N12, flanked by and including markers (i) BG1321 and CA0991, or (ii) BG1321 and CA1027, or (iii) BG1321 and PE0133, or (iv) PE0063 and CA0991, or (v) PE0133 and CA0991, or (vi) CA1027 and PE0063, or (vii) CA1027 and UB0331, or (viii) CA0423 and PE0250, or (ix) AG0359 and PE0250, or (x) AG0359 and

CA0896; (i) one or more intervals on linkage group N13, flanked by and including markers (i) BG0516 and AG0148, or (ii) CA0488 and AG0148, or (iii) CA0488 and CA0736, or (iv) CA0603 and AG0504, or (v) BG1288 and AG0504; (j) one or more intervals on linkage group N15, flanked by and including markers (i) CA0719 and AG0369, or (ii) PE0091 and PE0187, or (iii) PE0286 and AG0369, or (iv) PE0286 and PE0187, or (v) PE0286 and CA0719; (k) one or more intervals on linkage group N18, flanked by and including markers (i) AG0285 and CA0636, or (ii) BG0278 and CA07739, or (iii) CA0739 and CA0636, or (iv) UB0315 and CA0636, or (v) UB0315 and CA0739; and (l) one or more intervals on linkage group N19, flanked by and including markers (i) CA0552 and CA0221, or (ii) CA1107 and CA0552, or (iii) CA1107 and CA0221, or (iv) CA0221 and KK98G, or (v) UB0307 and BG1241, or (vi) BG1241 and KK98G, or (vii) CA0221 and BG1241.

In a particular embodiment, the QTL is localized to a chromosomal interval on linkage group N1, N9, N11, N12, N18 or N19.

In other embodiments, the marker comprises a polymorphism that identifies the at least one allele of the at least one quantitative trait locus (QTL) as being associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, and the detecting comprises identifying the polymorphism. The polymorphism may be, for example, a single nucleotide polymorphism (SNP) or a simple sequence repeat (SSR). In another embodiment of the method of the invention, the detecting comprises detecting at least one marker selected from AG0023; AG0045; AG0047; AG0070; AG0086; AG0093; AG0125; AG0148; AG0171; AG0203; AG0239; AG0243; AG0272; AG0304; AG0323; AG0324; AG0328; AG0359; AG0369; AG0370; AG0378; AG0391; AG0410; AG0441; AG0477; AG0482; AG0504; AG0510; BG0031; BG0106; BG0111; BG0119; BG0181; BG0228; BG0255; BG0278; BG0295; BG0452; BG0516; BG0647; BG0651; BG0713; BG0864; BG0869; BG0988; BG1062; BG1090; BG1101; BG1123; BG1127; BG1149; BG1182; BG1197; BG1230; BG1241; BG1244; BG1286; BG1288; BG1321; BG1368; BG1392; BG1442; BG1449; BG1453; BG1513; CA0105; CA0120; CA0163; CA0221; CA0226; CA0233; CA0328; CA0410; CA0423; CA0456; CA0488; CA0546; CA0552; CA0603; CA0614; CA0636; CA0681; CA0719; CA0736; CA0739; CA0753; CA0834; CA0837; CA0896; CA0991; CA1027; CA1032; CA1034; CA1035; CA1066; CA1080; CA1090; CA1097; CA1107; PE0012; PE0017; PE0063; PE0091; PE0131; PE0133; PE0177; PE0187;

PE0203; PE0250; PE0281; PE0283; PE0286; PE0324; PE0340; PE0355; UB0015; UB0126; UB0163; UB0181; UB0196; UB0307; UB0315; UB0331; KK66; and KK98G.

In another embodiment of the method of the invention, the detecting comprises detecting at least one marker selected from AG0093; AG0304; AG0378; AG0391; AG0482; BG1149; BG1230; BG1241; BG1453; BG1513; CA0120; CA0221; CA0546; CA0739; CA1027; PE0063; PE0203; UB0163; and UB0315.

In other embodiments, the plant is a whole plant, a plant organ, a plant seed or a plant cell. In other embodiments, the plant is canola. The plant may be, for example, *Brassica napus*, *Brassica juncea*, *Brassica rapa*, *Brassica oleracea*; or *Brassica carinata*. The plant may be, for example, spring canola, winter canola, or semi-winter canola. In another embodiment, the *Sclerotinia* whole plant field resistant plant is resistant to *Sclerotinia sclerotiorum*.

In another aspect, the invention features a method of making a transgenic dicot comprising a quantitative trait locus (QTL) associated with *Sclerotinia* whole plant field resistance or improved whole plant field resistance, the method comprising the steps of: introducing a nucleic acid cloned according to the method described above into a dicot cell; and growing the cell under cell growth conditions. In one embodiment, the QTL is localized to a chromosomal interval selected from: (a) an interval flanked by and including (i) markers CA0614 and PE0177 or (ii) markers AG0093 and AG0482 on linkage group N1; (b) an interval flanked by and including markers CA0410 and AG0023 on linkage group N3; (c) an interval flanked by and including markers BG1442 and BG0106 on linkage group N4; (d) an interval flanked by and including markers AG0510 and CA0105 on linkage group N7; (e) an interval flanked by and including markers CA0837 and BG1286 on linkage group N8; (f) an interval flanked by and including (i) markers CA1034 and AG0441or (ii) markers AG0378 and KK66 on linkage group N9; (g) an interval flanked by and including markers BG0228 and PE0131 on linkage group N10; (h) an interval flanked by and including (i) markers CA0120 and CA0163 or (ii) markers CA0120 and CA1097 on linkage group N11; (i) an interval flanked by and including (i) markers BG1321 and CA0991 or (ii) markers CA0753 and PE0250 on linkage group N12; (j) an interval flanked by and including markers CA0603 and CA0736 on linkage group N13; (k) an interval flanked by and including markers PE0286 and AG0369 on linkage group N15; (l) an interval flanked by and including (i) markers BG0278 and CA0636 or (ii) markers UB0315 and CA0739 on linkage group N18;

and (m) an interval flanked by and including (i) markers CA1107 and CA0221 or (ii) markers UB0307 and KK98G on linkage group N19.

In another embodiment, the QTL is localized to a chromosomal interval selected from: (a) one or more intervals on linkage group N1, flanked by and including markers (i) AG0093 and PE0203, or (ii) BG0111 and BG1392, or (iii) BG1090 and AG0482, or (iv) BG1090 and PE0203, or (v) CA0614 and BG1392, or (vi) BG0988 and AG0482; or (vii) AG0243 and AG0482; or (viii) AG0243 and BG1453; or BG0988; (b) one or more intervals on linkage group N3, flanked by and including markers (i) BG1197 and AG0023, or (ii) CA0410 and BG1368 or (iii) CA0410 and BG1197; (c) one or more intervals on linkage group N4, flanked by and including markers (i) BG1442 and BG0106, or (ii) UB0181 and BG0106; (d) one or more intervals on linkage group N8, flanked by and including markers (i) BG1449 and BG1062, or (ii) CA0837 and AG0328, or (iii) CA0837 and BG1062, or (iv) CA0837 and BG1101, or (v) CA0837 and BG1286, or (vi) CA0837 and BG1449 or (vii) PE0281 and BG0647; (e) one or more intervals on linkage group N9, flanked by and including markers (i) AG0323 and BG0295, or (ii) CA1034 and AG0378 or (iii) BG1123 and AG0441; (f) one or more intervals on linkage group N10, flanked by and including markers (i) BG0228 and AG0047, or BG0255 and PE0131; (g) one or more intervals on linkage group N11, flanked by and including markers (i) BG0031 and BG1149, or (ii) BG0031 and BG1230, or (iii) BG0031 and BG1513, or (iv) CA0120 and CA0328, or (v) PE0283 and CA0163, or (vi) PE0324 and PE0283 or (vii) CA0328 and PE0324, or (viii) CA0226 and BG0713, or (ix) CA0233 and CA1080, or (x) CA0233 and AG0370; (h) one or more intervals on linkage group N12, flanked by and including markers (i) BG1321 and CA0991, or (ii) BG1321 and CA1027, or (iii) BG1321 and PE0133, or (iv) PE0063 and CA0991, or (v) PE0133 and CA0991, or (vi) CA1027 and PE0063, or (vii) CA1027 and UB0331, or (viii) CA0423 and PE0250, or (ix) AG0359 and PE0250, or (x) AG0359 and CA0896; (i) one or more intervals on linkage group N13, flanked by and including markers (i) BG0516 and AG0148, or (ii) CA0488 and AG0148, or (iii) CA0488 and CA0736, or (iv) CA0603 and AG0504, or (v) BG1288 and AG0504; (j) one or more intervals on linkage group N15, flanked by and including markers (i) CA0719 and AG0369, or (ii) PE0091 and PE0187, or (iii) PE0286 and AG0369, or (iv) PE0286 and PE0187, or (v) PE0286 and CA0719; (k) one or more intervals on linkage group N18, flanked by and including markers (i) AG0285 and CA0636, or (ii) BG0278 and CA07739, or (iii) CA0739 and CA0636, or (iv)

UB0315 and CA0636, or (v) UB0315 and CA0739; and (l) one or more intervals on linkage group N19, flanked by and including markers (i) CA0552 and CA0221, or (ii) CA1107 and CA0552, or (iii) CA1107 and CA0221, or (iv) CA0221 and KK98G, or (v) UB0307 and BG1241, or (vi) BG1241 and KK98G, or (vii) CA0221 and BG1241.

In a particular embodiment, the QTL is localized to a chromosomal interval on linkage group N1, N9, N11, N12, N18 or N19.

In other embodiments, the marker comprises a polymorphism that identifies the at least one allele of the at least one quantitative trait locus (QTL) as being associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, and the detecting comprises identifying the polymorphism. The polymorphism may be, for example, a single nucleotide polymorphism (SNP) or a simple sequence repeat (SSR). In another embodiment of the method of the invention, the detecting comprises detecting at least one marker selected from AG0023; AG0045; AG0047; AG0070; AG0086; AG0093; AG0125; AG0148; AG0171; AG0203; AG0239; AG0243; AG0272; AG0304; AG0323; AG0324; AG0328; AG0359; AG0369; AG0370; AG0378; AG0391; AG0410; AG0441; AG0477; AG0482; AG0504; AG0510; BG0031; BG0106; BG0111; BG0119; BG0181; BG0228; BG0255; BG0278; BG0295; BG0452; BG0516; BG0647; BG0651; BG0713; BG0864; BG0869; BG0988; BG1062; BG1090; BG1101; BG1123; BG1127; BG1149; BG1182; BG1197; BG1230; BG1241; BG1244; BG1286; BG1288; BG1321; BG1368; BG1392; BG1442; BG1449; BG1453; BG1513; CA0105; CA0120; CA0163; CA0221; CA0226; CA0233; CA0328; CA0410; CA0423; CA0456; CA0488; CA0546; CA0552; CA0603; CA0614; CA0636; CA0681; CA0719; CA0736; CA0739; CA0753; CA0834; CA0837; CA0896; CA0991; CA1027; CA1032; CA1034; CA1035; CA1066; CA1080; CA1090; CA1097; CA1107; PE0012; PE0017; PE0063; PE0091; PE0131; PE0133; PE0177; PE0187; PE0203; PE0250; PE0281; PE0283; PE0286; PE0324; PE0340; PE0355; UB0015; UB0126; UB0163; UB0181; UB0196; UB0307; UB0315; UB0331; KK66; and KK98G.

In another embodiment, the detecting comprises detecting at least one marker selected from AG0093; AG0304; AG0378; AG0391; AG0482; BG1149; BG1230; BG1241; BG1453; BG1513; CA0120; CA0221; CA0546; CA0739; CA1027; PE0063; PE0203; UB0163; and UB0315.

In another embodiment, the dicot cell is regenerated to form a first plant. In another embodiment, the first plant is crossed with a second plant of the same species. In another

embodiment, the dicot is a soybean, sunflower, canola, or alfalfa. In another embodiment, the dicot is canola, for example, spring canola, winter canola, or semi-winter canola. In another embodiment, the dicot is *Brassica napus*, *Brassica juncea*, *Brassica rapa*, or *Brassica oleracea*. In another embodiment, the *Sclerotinia* whole plant field resistant plant is resistant to *Sclerotinia sclerotiorum*. In other embodiments the whole plant field resistance results from decreased disease incidence or from decreased disease severity compared to a dicot lacking the QTL.

Another aspect of the invention features a method of identifying a candidate nucleic acid comprising a QTL associated with *Sclerotinia* whole plant field resistance from a dicot, the method comprising: providing a nucleic acid cloned according to the methods described above; and, identifying a homolog of the nucleic acid in a dicot.

Another aspect of the invention features a method of marker assisted selection comprising (MAS) of a quantitative trait locus (QTL) associated with whole plant field resistance to *Sclerotinia*, the method comprising the steps of: obtaining a first *Brassica* plant having at least one allele of a marker locus, wherein the marker locus is associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia* with a statistical significance of $p \leq 0.01$; crossing the first *Brassica* plant to a second *Brassica* plant; evaluating the progeny for at least the allele; and selecting progeny plants that possess at least the allele. In one embodiment, the plant is a member of a segregating population. In another embodiment, the marker assisted selection is done via high throughput screening.

Another aspect of the invention features a *Brassica* plant identified by the above method, and progeny thereof, including F1, F2 and F3 progeny.

Another aspect of the invention features an isolated or recombinant nucleic acid comprising a polynucleotide selected from the group consisting of: a sequence selected from any one of marker sequences AG0023 (SEQ ID NO:1); AG0045 (SEQ ID NO:2); AG0047 (SEQ ID NO:3); AG0070 (SEQ ID NO:4); AG0086 (SEQ ID NO:5); AG0093 (SEQ ID NO:6); AG0125 (SEQ ID NO:7); AG0148 (SEQ ID NO:8); AG0171 (SEQ ID NO:9); AG0203 (SEQ ID NO:10); AG0239 (SEQ ID NO:11); AG0243 (SEQ ID NO:12); AG0272 (SEQ ID NO:13); AG0304 (SEQ ID NO:14); AG0323 (SEQ ID NO:15); AG0324 (SEQ ID NO:16); AG0328 (SEQ ID NO:17); AG0359 (SEQ ID NO:18); AG0369 (SEQ ID NO:19); AG0370 (SEQ ID NO:20); AG0378 (SEQ ID NO:21); AG0391 (SEQ ID NO:22); AG0410 (SEQ ID NO:23); AG0441 (SEQ ID NO:24); AG0477 (SEQ ID NO:25); AG0482 (SEQ ID

NO:26); AG0504 (SEQ ID NO:27); AG0510 (SEQ ID NO:28); BG0031 (SEQ ID NO:29); BG0106 (SEQ ID NO:30); BG0111 (SEQ ID NO:31); BG0119 (SEQ ID NO:32); BG0181 (SEQ ID NO:33); BG0228 (SEQ ID NO:34); BG0255 (SEQ ID NO:35); BG0278 (SEQ ID NO:36); BG0295 (SEQ ID NO:37); BG0452 (SEQ ID NO:38); BG0516 (SEQ ID NO:39); BG0647 (SEQ ID NO:40); BG0651 (SEQ ID NO:41); BG0713 (SEQ ID NO:42); BG0864 (SEQ ID NO:43); BG0869 (SEQ ID NO:44); BG0988 (SEQ ID NO:45); BG1062 (SEQ ID NO:46); BG1090 (SEQ ID NO:47); BG1101 (SEQ ID NO:48); BG1123 (SEQ ID NO:49); BG1127 (SEQ ID NO:50); BG1149 (SEQ ID NO:51); BG1182 (SEQ ID NO:52); BG1197 (SEQ ID NO:53); BG1230 (SEQ ID NO:54); BG1241 (SEQ ID NO:55); BG1244 (SEQ ID NO:56); BG1286 (SEQ ID NO:57); BG1288 (SEQ ID NO:58); BG1321 (SEQ ID NO:59); BG1368 (SEQ ID NO:60); BG1392 (SEQ ID NO:61); BG1442 (SEQ ID NO:62); BG1449 (SEQ ID NO:63); BG1453 (SEQ ID NO:64); BG1513 (SEQ ID NO:65); CA0105 (SEQ ID NO:66); CA0120 (SEQ ID NO:67); CA0163 (SEQ ID NO:68); CA0221 (SEQ ID NO:69); CA0226 (SEQ ID NO:70); CA0233 (SEQ ID NO:71); CA0328 (SEQ ID NO:72); CA0410 (SEQ ID NO:73); CA0423 (SEQ ID NO:74); CA0456 (SEQ ID NO:75); CA0488 (SEQ ID NO:76); CA0546 (SEQ ID NO:77); CA0552 (SEQ ID NO:78); CA0603 (SEQ ID NO:79); CA0614 (SEQ ID NO:80); CA0636 (SEQ ID NO:81); CA0681 (SEQ ID NO:82); CA0719 (SEQ ID NO:83); CA0736 (SEQ ID NO:84); CA0739 (SEQ ID NO:85); CA0753 (SEQ ID NO:86); CA0834 (SEQ ID NO:87); CA0837 (SEQ ID NO:88); CA0896 (SEQ ID NO:89); CA0991 (SEQ ID NO:90); CA1027 (SEQ ID NO:91); CA1032 (SEQ ID NO:92); CA1034 (SEQ ID NO:93); CA1035 (SEQ ID NO:94); CA1066 (SEQ ID NO:95); CA1080 (SEQ ID NO:96); CA1090 (SEQ ID NO:97); CA1097 (SEQ ID NO:98); or CA1107 (SEQ ID NO:99); (b) a polynucleotide sequence with at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to a polynucleotide of (a); and (c) a polynucleotide sequence complementary to the polynucleotide sequence of (a) or (b). In one embodiment, the isolated or recombinant nucleic acid is associated with whole plant field resistance to *Sclerotinia*.

In another embodiment, the isolated or recombinant nucleic acid comprising a polynucleotide is selected from the group consisting of: (a) a sequence selected from any one of marker sequences AG0093 (SEQ ID NO:6); AG0304 (SEQ ID NO:14); AG0378 (SEQ ID NO:21); AG0391 (SEQ ID NO:22); AG0482 (SEQ ID NO:26); BG1149 (SEQ ID NO:51);

BG1230 (SEQ ID NO:54); BG1241 (SEQ ID NO:55); BG1453 (SEQ ID NO:64); BG1513 (SEQ ID NO:65); (SEQ ID NO:67); CA0221 (SEQ ID NO:69); CA0546 (SEQ ID NO:77); CA0739 (SEQ ID NO:85); or CA1027 (SEQ ID NO:91); (b) a polynucleotide sequence with at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to a polynucleotide of (a); and (c) a polynucleotide sequence complementary to the polynucleotide sequence of (a) or (b).

Another aspect of the invention features an isolated nucleic acid molecule for detecting a polymorphism in plant DNA associated whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, wherein the nucleic acid molecule comprises at least 15 nucleotides and is identical to a sequence of the same number of consecutive nucleotides in either strand of the plant DNA in a region where the polymorphism is located, wherein the nucleic acid molecule comprises a sequence that is at least 70% , 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to a marker sequence or a fragment of a marker sequence selected from the group consisting of marker AG0023 (SEQ ID NO:1); AG0045 (SEQ ID NO:2); AG0047 (SEQ ID NO:3); AG0070 (SEQ ID NO:4); AG0086 (SEQ ID NO:5); AG0093 (SEQ ID NO:6); AG0125 (SEQ ID NO:7); AG0148 (SEQ ID NO:8); AG0171 (SEQ ID NO:9); AG0203 (SEQ ID NO:10); AG0239 (SEQ ID NO:11); AG0243 (SEQ ID NO:12); AG0272 (SEQ ID NO:13); AG0304 (SEQ ID NO:14); AG0323 (SEQ ID NO:15); AG0324 (SEQ ID NO:16); AG0328 (SEQ ID NO:17); AG0359 (SEQ ID NO:18); AG0369 (SEQ ID NO:19); AG0370 (SEQ ID NO:20); AG0378 (SEQ ID NO:21); AG0391 (SEQ ID NO:22); AG0410 (SEQ ID NO:23); AG0441 (SEQ ID NO:24); AG0477 (SEQ ID NO:25); AG0482 (SEQ ID NO:26); AG0504 (SEQ ID NO:27); AG0510 (SEQ ID NO:28); BG0031 (SEQ ID NO:29); BG0106 (SEQ ID NO:30); BG0111 (SEQ ID NO:31); BG0119 (SEQ ID NO:32); BG0181 (SEQ ID NO:33); BG0228 (SEQ ID NO:34); BG0255 (SEQ ID NO:35); BG0278 (SEQ ID NO:36); BG0295 (SEQ ID NO:37); BG0452 (SEQ ID NO:38); BG0516 (SEQ ID NO:39); BG0647 (SEQ ID NO:40); BG0651 (SEQ ID NO:41); BG0713 (SEQ ID NO:42); BG0864 (SEQ ID NO:43); BG0869 (SEQ ID NO:44); BG0988 (SEQ ID NO:45); BG1062 (SEQ ID NO:46); BG1090 (SEQ ID NO:47); BG1101 (SEQ ID NO:48); BG1123 (SEQ ID NO:49); BG1127 (SEQ ID NO:50); BG1149 (SEQ ID NO:51); BG1182 (SEQ ID NO:52); BG1197 (SEQ ID NO:53); BG1230

(SEQ ID NO:54); BG1241 (SEQ ID NO:55); BG1244 (SEQ ID NO:56); BG1286 (SEQ ID NO:57); BG1288 (SEQ ID NO:58); BG1321 (SEQ ID NO:59); BG1368 (SEQ ID NO:60); BG1392 (SEQ ID NO:61); BG1442 (SEQ ID NO:62); BG1449 (SEQ ID NO:63); BG1453 (SEQ ID NO:64); BG1513 (SEQ ID NO:65); CA0105 (SEQ ID NO:66); CA0120 (SEQ ID NO:67); CA0163 (SEQ ID NO:68); CA0221 (SEQ ID NO:69); CA0226 (SEQ ID NO:70); CA0233 (SEQ ID NO:71); CA0328 (SEQ ID NO:72); CA0410 (SEQ ID NO:73); CA0423 (SEQ ID NO:74); CA0456 (SEQ ID NO:75); CA0488 (SEQ ID NO:76); CA0546 (SEQ ID NO:77); CA0552 (SEQ ID NO:78); CA0603 (SEQ ID NO:79); CA0614 (SEQ ID NO:80); CA0636 (SEQ ID NO:81); CA0681 (SEQ ID NO:82); CA0719 (SEQ ID NO:83); CA0736 (SEQ ID NO:84); CA0739 (SEQ ID NO:85); CA0753 (SEQ ID NO:86); CA0834 (SEQ ID NO:87); CA0837 (SEQ ID NO:88); CA0896 (SEQ ID NO:89); CA0991 (SEQ ID NO:90); CA1027 (SEQ ID NO:91); CA1032 (SEQ ID NO:92); CA1034 (SEQ ID NO:93); CA1035 (SEQ ID NO:94); CA1066 (SEQ ID NO:95); CA1080 (SEQ ID NO:96); CA1090 (SEQ ID NO:97); CA1097 (SEQ ID NO:98); or CA1107 (SEQ ID NO:99). In one embodiment, the nucleic acid is selected from any of SEQ ID NOs: 126-323.

In another embodiment, the marker sequence is AG0093 (SEQ ID NO:6); AG0304 (SEQ ID NO:14); AG0378 (SEQ ID NO:21); AG0391 (SEQ ID NO:22); AG0482 (SEQ ID NO:26); BG1149 (SEQ ID NO:51); BG1230 (SEQ ID NO:54); BG1241 (SEQ ID NO:55); BG1453 (SEQ ID NO:64); BG1513 (SEQ ID NO:65); (SEQ ID NO:67); CA0221 (SEQ ID NO:69); CA0546 (SEQ ID NO:77); CA0739 (SEQ ID NO:85); or CA1027 (SEQ ID NO:91). In one embodiment, the nucleic acid is selected from any of SEQ ID NOs: 136, 137, 152, 153, 166, 167, 168, 169, 176, 177, 226, 227, 232, 233, 234, 235, 252, 253, 254, 255, 258, 259, 262, 263, 278, 279, 294, 295, 306, or 307.

Another aspect of the invention features a kit for screening a plant or germplasm for a QTL associated with whole plant field resistance to *Sclerotinia*, comprising a container in which is contained one or more of the isolated nucleic acid molecules of claim as described above; and instructions for screening a plant for the QTL associated with whole plant field resistance or improved whole plant field resistance to *Sclerotinia*.

In one embodiment, the further comprises a buffer. In another embodiment, the kit is for use in high throughput screening and comprises at least one component for such use. In another embodiment the kit is for use in high throughput screening in an integrated system.

In another aspect, the invention features a *Brassica* plant that exhibits whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, comprising alleles favorable to *Sclerotinia* whole plant field resistance in at least 1 QTL localized to a chromosomal interval selected from: (a) an interval flanked by and including (i) markers CA0614 and PE0177 or (ii) markers AG0093 and AG0482 on linkage group N1; (b) an interval flanked by and including markers CA0410 and AG0023 on linkage group N3; (c) an interval flanked by and including markers BG1442 and BG0106 on linkage group N4; (d) an interval flanked by and including markers AG0510 and CA0105 on linkage group N7; (e) an interval flanked by and including markers CA0837 and BG1286 on linkage group N8; (f) an interval flanked by and including (i) markers CA1034 and AG0441or (ii) markers AG0378 and KK66 on linkage group N9; (g) an interval flanked by and including markers BG0228 and PE0131 on linkage group N10; (h) an interval flanked by and including (i) markers CA0120 and CA0163 or (ii) markers CA0120 and CA1097 on linkage group N11; (i) an interval flanked by and including (i) markers BG1321 and CA0991 or (ii) markers CA0753 and PE0250 on linkage group N12; (j) an interval flanked by and including markers CA0603 and CA0736 on linkage group N13; (k) an interval flanked by and including markers PE0286 and AG0369 on linkage group N15; (l) an interval flanked by and including (i) markers BG0278 and CA0636 or (ii) markers UB0315 and CA0739 on linkage group N18; and (m) an interval flanked by and including (i) markers CA1107 and CA0221 or (ii) markers UB0307 and KK98G on linkage group N19.

In another embodiment, the QTL is localized to a chromosomal interval selected from: (a) one or more intervals on linkage group N1, flanked by and including markers (i) AG0093 and PE0203, or (ii) BG0111 and BG1392, or (iii) BG1090 and AG0482, or (iv) BG1090 and PE0203, or (v) CA0614 and BG1392, or (vi) BG0988 and AG0482; or (vii) AG0243 and AG0482; or (viii) AG0243 and BG1453; or BG0988; (b) one or more intervals on linkage group N3, flanked by and including markers (i) BG1197 and AG0023, or (ii) CA0410 and BG1368 or (iii) CA0410 and BG1197; (c) one or more intervals on linkage group N4, flanked by and including markers (i) BG1442 and BG0106, or (ii) UB0181 and BG0106; (d) one or more intervals on linkage group N8, flanked by and including markers (i) BG1449 and BG1062, or (ii) CA0837 and AG0328, or (iii) CA0837 and BG1062, or (iv) CA0837 and BG1101, or (v) CA0837 and BG1286, or (vi) CA0837 and BG1449 or (vii) PE0281 and BG0647; (e) one or more intervals on linkage group N9, flanked by and

including markers (i) AG0323 and BG0295, or (ii) CA1034 and AG0378 or (iii) BG1123 and AG0441; (f) one or more intervals on linkage group N10, flanked by and including markers (i) BG0228 and AG0047, or BG0255 and PE0131; (g) one or more intervals on linkage group N11, flanked by and including markers (i) BG0031 and BG1149, or (ii) BG0031 and BG1230, or (iii) BG0031 and BG1513, or (iv) CA0120 and CA0328, or (v) PE0283 and CA0163, or (vi) PE0324 and PE0283 or (vii) CA0328 and PE0324, or (viii) CA0226 and BG0713, or (ix) CA0233 and CA1080, or (x) CA0233 and AG0370; (h) one or more intervals on linkage group N12, flanked by and including markers (i) BG1321 and CA0991, or (ii) BG1321 and CA1027, or (iii) BG1321 and PE0133, or (iv) PE0063 and CA0991, or (v) PE0133 and CA0991, or (vi) CA1027 and PE0063, or (vii) CA1027 and UB0331, or (viii) CA0423 and PE0250, or (ix) AG0359 and PE0250, or (x) AG0359 and CA0896; (i) one or more intervals on linkage group N13, flanked by and including markers (i) BG0516 and AG0148, or (ii) CA0488 and AG0148, or (iii) CA0488 and CA0736, or (iv) CA0603 and AG0504, or (v) BG1288 and AG0504; (j) one or more intervals on linkage group N15, flanked by and including markers (i) CA0719 and AG0369, or (ii) PE0091 and PE0187, or (iii) PE0286 and AG0369, or (iv) PE0286 and PE0187, or (v) PE0286 and CA0719; (k) one or more intervals on linkage group N18, flanked by and including markers (i) AG0285 and CA0636, or (ii) BG0278 and CA07739, or (iii) CA0739 and CA0636, or (iv) UB0315 and CA0636, or (v) UB0315 and CA0739; and (l) one or more intervals on linkage group N19, flanked by and including markers (i) CA0552 and CA0221, or (ii) CA1107 and CA0552, or (iii) CA1107 and CA0221, or (iv) CA0221 and KK98G, or (v) UB0307 and BG1241, or (vi) BG1241 and KK98G, or (vii) CA0221 and BG1241.

In a particular embodiment, the QTL is localized to a chromosomal interval on linkage group N1, N9, N11, N12, N18 or N19.

Other features and advantages of the invention will be understood from the detailed description and examples that follow.

Detailed Discussion

Overview

The present invention relates to the identification of genetic markers, e.g., marker loci and nucleic acids corresponding to (or derived from) these marker loci, such as probes and amplification products useful for genotyping plants, correlated with *Sclerotinia* whole

plant field resistance. The markers of the invention are used to identify plants, particularly plants of the species *Brassica napus* (*B. napus*) (canola), that are resistant or exhibit improved resistance to *Sclerotinia*. Accordingly, these markers are useful for marker-assisted selection (MAS) and breeding of *Sclerotinia* resistant plants, and for identification of susceptible plants. The markers of the invention are also used to identify and define nucleic acids that are proximal to and/or chromosome intervals corresponding to, or including, quantitative trait loci associated with *Sclerotinia* whole plant field resistance. Quantitative Trait Loci (QTLs) associated with *Sclerotinia* whole plant field resistance are isolated by positional cloning, e.g., nucleic acids proximal to or of genetic intervals defined by a pair of markers described herein, or subsequences of an interval defined by and including such markers. Such isolated QTL nucleic acids can be used for the production of transgenic cells and plants exhibiting improved resistance to *Sclerotinia*. In addition, QTL nucleic acids isolated from one organism, e.g., canola, can, in turn, serve to isolate homologs of QTLs for *Sclerotinia* whole plant field resistance from other susceptible organisms, including a variety of commercially important dicots, such as soybean, alfalfa, sunflower, flax, beans, (for example, white beans), potatoes, peas and peanuts.

Definitions

Units, prefixes, and symbols are denoted in their International System of Units (SI) accepted form. Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation; and amino acid sequences are written left to right in amino to carboxy orientation. Numeric ranges recited within the specification are inclusive of the numbers defining the range and include each integer within the defined range. Nucleotides may be referred to herein by their one-letter symbols recommended by the IUPAC-IUBMB Nomenclature Commission. The terms defined below are more fully defined by reference to the specification as a whole. Section headings provided throughout the specification are provided for convenience and are not limitations to the various objects and embodiments of the present invention.

The term "*Sclerotinia* whole plant field resistance" or "whole plant field resistance to *Sclerotinia*" refers to the resistance of a plant against the plant pathogen *Sclerotinia*, under field conditions or under extreme disease pressure field research conditions (as described, for example, herein and in WO 2006/135717). It reflects the resistance of the entire plant when exposed to *Sclerotinia* under these conditions. In one embodiment, a plant with *Sclerotinia*

whole plant field resistance has a rating of disease development of 5.0 or greater, based on the *Sclerotinia Sclerotiorum* Disease Incidence Severity (SSDIS) rating scale. In other embodiments, a plant with *Sclerotinia* whole plant field resistance has a rating of disease development of 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0 or greater, based on the *Sclerotinia Sclerotiorum* Disease Incidence Severity (SSDIS) rating scale. Ratings of disease development are sometimes expressed in ranges; for instance in a range of 5-6, 6-7, 7-8, 8-9 or in a range of 5-7, 7-9 and so on, or by a number range within integers, such as 5.5-6.5, 5.5-7.5, 6-7.5, 7-8.5, for example. In those instances, a plant with *Sclerotinia* whole plant field resistance has a rating of disease development in the range of at least 5-6, or 6-7, or 7-8, or 8-9, based on the *Sclerotinia Sclerotiorum* Disease Incidence Severity (SSDIS) rating scale.

It will be understood by the skilled artisan that the greater the number (or percentage) of favorable alleles for *Sclerotinia* whole plant field resistance a plant posesses, the greater will be the level of resistance exhibited. This concept can be appreciated by reference to Table 8 herein. In certain embodiments, a plant with *Sclerotinia* whole plant field resistance has a genome containing at least about 50% favorable alleles. In more particular embodiments, a plant with *Sclerotinia* whole plant field resistance has a genome containing at least 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95% or more favorable alleles. The percentage of favorable alleles can also be expressed as a number value. For instance, as shown in Table 8, if a total number of 15 favorable alleles are possible in a certain mapping population, a plant having 12 of those alleles would have 80% favorable alleles. In certain embodiments, the number or percent of favorable alleles in a plant can serve as a rough predictor of the expected level of *Sclerotinia* whole plant field resistance a plant will exhibit.

It will also be understood by the skilled artisan that the QTLs described herein represent regions of the genome comprising genes that contribute to the *Sclerotinia* whole plant field resistance of a plant. Further, each QTL can contribute differently to that resistance level. Thus, breeding efforts are directed to increasing the number of those QTLs, particularly quantitatively significant QTLs, present in the germplasm. Early in a breeding program, fewer QTLs may be present in a particular germplasm, but that number will

increase as the breeding program progresses. Thus, in certain embodiments, a plant exhibiting *Sclerotinia* whole plant field resistance may contain at least 6 of the QTLs described herein. More particularly, the plant may contain at least 7, 8, 9 or 10 of the QTLs described herein. Yet more particularly, the plant may contain 11, 12 or all of the QTLs described herein.

In the present invention, the evaluation for whole plant field resistance, using, for example, extreme disease pressure field research conditions, mimicked growers' field conditions and the worst-case field scenario in canola, requiring two fungicide applications to protect the crop from infection from *Sclerotinia*.

The term "*Sclerotinia* improved whole plant field resistance" or "improved whole plant field resistance to *Sclerotinia*" refers to the increase in resistance of a plant against the plant pathogen *Sclerotinia*, under field conditions or under extreme disease pressure field research conditions (as described, for example, herein and in WO 2006/135717). In one embodiment, a plant with improved whole plant field resistance to *Sclerotinia* is a plant having at least one allele of a QTL associated with whole plant field resistance to *Sclerotinia* and having rating of disease development of 5.0 or higher based on the SSDIS scale compared to a plant that does not have the at least one allele. Other embodiments of *Sclerotinia* improved whole plant field resistance parallel those outlined in the description of *Sclerotinia* whole plant field resistance set forth above.

Sclerotinia affects many different tissues of a plant. Natural infection by *Sclerotinia* begins with infection of the flower petal. The disease spreads to the leaves once the infected petals fall onto them. Lesions then develop simultaneously on a number of leaves per plant. These lesions further expand to colonize and wilt the leaf with infection proceeding further towards the stem via leaf petioles. The infection then reaches the stem to develop further in the stem causing premature ripening.

A plant with whole plant field resistance to *Sclerotinia* is a plant that is resistant to the pathway of *Sclerotinia* disease development in all tissues of the plant. Accordingly, a plant with whole plant field resistance to *Sclerotinia* can also be termed a plant with "pathway resistance" or "field pathway resistance" to *Sclerotinia*. The screening methods as described herein are used to identify plants with whole plant field resistance to *Sclerotinia*. These methods are unique compared to other screening methods known in the art for assessing resistance to *Sclerotinia*. Other screening methods known in the art to assess

resistance to *Sclerotinia* examine only part of the plant, and these methods take place at growth stages not associated with natural disease development. For example, Zhou et al. (2003, TAG 106: 759-764) performed phenotyping by inoculating the leaf at the seedling stage and by inoculating the stem at the mature plant stage. Zhou et al. (2006, TAG 112:509-5160) phenotyped based on petiole inoculation on a single plant per line, while Bela et al. (17th Crucifer Genetics Workshop (*Brassica* 2010), September 2010, Saskatoon, Canada) phenotyped based on petiole inoculation on 12 plants per line. Yin et al. (2010, Euphytica, online version: DOI 10.1007/s10681-009-0095-1) utilized three inoculation methods: mycelial toothpick inoculation, mycelial plug inoculation and infected petal inoculation onto cauline leaves. While it may be possible to identify one or more QTLs associated with *Sclerotinia* resistance using such screening methods, these screening methods are not as comprehensive as the screening methods to identify whole plant field resistance to *Sclerotinia* as described herein. This means that while these other screening methods may be used to uncover one or a few QTLs associated with resistance to *Sclerotinia* in a particular tissue of the plant, they cannot be used to identify all of the QTLs associated with *Sclerotinia* resistance throughout the entire plant. In addition, screening methods involving a single tissue, rather than the whole plant, will not be able to detect epistatic effects resulting from genes in different tissues working together to influence *Sclerotinia* resistance.

There are a number of advantages to using a whole plant approach to detecting resistance to *Sclerotinia*. First, this methodology most closely resembles the natural interaction between *Sclerotinia* and plants in the field, and should, therefore, be a superior system in which to identify QTLs associated with resistance to *Sclerotinia*. Second, the whole plant approach allows for a larger number of QTLs to be identified relative to other screening methodologies that only examine one plant tissue. Third, this approach permits analyses of epistatic effects, unlike other screening methods. Fourth, this approach allows actual field performance to be predicted from the data.

The term "quantitative trait locus" or "QTL" refers to a polymorphic genetic locus with at least two alleles that differentially affect the expression of a continuously distributed phenotypic trait, for example, whole plant field resistance to *Sclerotinia* or improved whole plant field resistance to *Sclerotinia*. For example, the QTL may have a favorable allele that

confers, or contributes to, whole plant field resistance to *Sclerotinia* or improved whole plant field resistance to *Sclerotinia.*

The term "favorable allele" is an allele at a particular locus that confers, or contributes to, a desirable phenotype, e.g., whole plant field resistance to *Sclerotinia* or improved whole plant field resistance to *Sclerotinia*, or alternatively is an allele that allows the identification of plants with decreased whole plant field resistance that can be removed from a breeding program or planting ("counterselection"). A favorable allele of a marker is a marker allele that segregates with the favorable phenotype, or alternatively, segregates with the unfavorable plant phenotype, therefore providing the benefit of identifying plants. Alleles that are favorable for whole plant field resistance to *Sclerotinia* or improved whole plant field resistance to *Sclerotinia* are provided, for example, in Tables 7 and 13.

The term "associated with" or "associated" in the context of this invention refers to, e.g., a nucleic acid and a phenotypic trait or a second nucleic acid, that are in linkage disequilibrium, i.e., the nucleic acid and the trait/second nucleic acid are found together in progeny plants more often than if the nucleic acid and phenotype/second nucleic acid segregated separately.

The term "linkage" is used to describe the degree with which one marker locus is associated with another marker locus or some other locus (for example, a QTL). The linkage relationship between a molecular marker and a phenotype is given as a "probability" or "adjusted probability". Linkage can be expressed as a desired limit or range. For example, in some embodiments, any marker is linked (genetically and physically) to any other marker when the markers are separated by less than 50, 40, 30, 25, 20, or 15 map units (or cM). In some aspects, it is advantageous to define a bracketed range of linkage, for example, between 10 and 20 cM, between 10 and 30 cM, or between 10 and 40 cM. The more closely a marker is linked to a second locus, the better an indicator for the second locus that marker becomes. Thus, "closely linked loci" such as a marker locus and a second locus display an inter-locus recombination frequency of 10% or less, preferably about 9% or less, still more preferably about 8% or less, yet more preferably about 7% or less, still more preferably about 6% or less, yet more preferably about 5% or less, still more preferably about 4% or less, yet more preferably about 3% or less, and still more preferably about 2% or less. In highly preferred embodiments, the relevant loci display a recombination frequency of about 1% or less, e.g., about 0.75% or less, more preferably about 0.5% or less, or yet more

preferably about 0.25% or less. Two loci that are localized to the same chromosome, and at such a distance that recombination between the two loci occurs at a frequency of less than 10% (e.g., about 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.75%, 0.5%, 0.25%, or less) are also said to be "proximal to" or "in proximity of" each other. Since one cM is the distance between two markers that show a 1% recombination frequency, any marker is closely linked (genetically and physically) to any other marker that is in close proximity, e.g., at or less than 10 cM distant. Two closely linked markers on the same chromosome can be positioned 9, 8, 7, 6, 5, 4, 3, 2, 1, 0.75, 0.5 or 0.25 cM or less from each other.

The term "linkage disequilibrium" refers to a non-random segregation of genetic loci. This implies that such loci are in sufficient physical proximity along a length of a chromosome that they tend to segregate together with greater than random frequency.

The term "genetically linked" refers to genetic loci that are in linkage disequilibrium and statistically determined not to assort independently. Genetically linked loci assort dependently from 51% to 99% of the time or any whole number value there between, preferably at least 60%, 70%, 80%, 90%, 95% or 99%. Loci or alleles that are inherited in this way are said to be linked, and are referred to as "linkage groups".

The "probability value" or "p-value" is the statistical likelihood that the particular combination of a phenotype and the presence or absence of a particular marker is random. The lower the probability value, the greater the likelihood that a phenotype and a particular marker will co-segregate. In some aspects, the probability value is considered "significant" or "non-significant". In some embodiments, a probability value of 0.05 (p=0.05, or a 5% probability) of random assortment is considered a significant indication of co-segregation. However, an acceptable probability can be any probability of less than 50% (p=0.5). For example, a significant probability can be less than 0.25, less than 0.2, less than 0.15, less than 0.1, less than 0.05, less than 0.01 or less than 0.001.

The term "marker locus" is a specific chromosome location in the genome of a species where a specific marker can be found. A marker locus can be used to track the presence of a second linked locus, e.g., a linked locus that encodes or contributes to expression of a phenotypic trait. For example, a marker locus can be used to monitor segregation of alleles at a locus, such as a QTL or single gene, that are genetically or physically linked to the marker locus.

The term "marker" is a nucleotide sequence or encoded product thereof (e.g., a protein) used as a point of reference. For markers to be useful at detecting recombinations, they need to detect differences, or polymorphisms, within the population being monitored. For molecular markers, this means differences at the DNA level due to polynucleotide sequence differences (e.g., SSRs, RFLPs, FLPs, SNPs). The genomic variability can be of any origin, for example, insertions, deletions, duplications, repetitive elements, point mutations, recombination events, or the presence and sequence of transposable elements. Molecular markers can be derived from genomic or expressed nucleic acids (e.g., ESTs) and can also refer to nucleic acids used as probes or primer pairs capable of amplifying sequence fragments via the use of PCR-based methods. A large number of *Brassica* molecular markers are known in the art, and are published or available from various sources.

Examples of markers are provided, in SEQ ID NOS: 1-125. It will be understood by one skilled in the art that a marker of the present invention may comprise the entire sequence of any one of the sequences set out in SEQ ID NOS: 1-125, or a fragment of such a sequence. The fragment can be, for example, the SSR (as set out, for example, in Table 14, or a sequence that flanks (e.g., those as set out as SEQ ID NOS: 126-325) and includes the SSR. It will also be understood by one skilled in the art that the sequences of markers such as those set out in any of SEQ ID NOS: 1-125 or a fragment of such a sequence will have some variation from line to line. Therefore, the markers of the present invention include sequences that have 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to a sequence as provided in any of SEQ ID NOS: 1-125 or a fragment thereof.

Markers corresponding to genetic polymorphisms between members of a population can be detected by methods well-established in the art. These include, e.g., DNA sequencing, PCR-based sequence specific amplification methods, detection of restriction fragment length polymorphisms (RFLP), detection of isozyme markers, detection of polynucleotide polymorphisms by allele specific hybridization (ASH), detection of amplified variable sequences of the plant genome, detection of self-sustained sequence replication, detection of simple sequence repeats (SSRs), detection of single nucleotide polymorphisms (SNPs), or detection of amplified fragment length polymorphisms (AFLPs). Well established methods are also known for the detection of expressed sequence tags (ESTs) and

SSR markers derived from EST sequences and randomly amplified polymorphic DNA (RAPD).

The term "molecular marker" may be used to refer to any type of nucleic acid based marker, or an encoded product thereof (e.g., a protein) used as a point of reference when identifying a linked locus. A marker can be derived from genomic nucleotide sequences or from expressed nucleotide sequences (e.g., from a spliced RNA, a cDNA, etc.), or from an encoded polypeptide. The term also refers to nucleic acid sequences complementary to or flanking the marker sequences, such as nucleic acids used as probes or primer pairs capable of amplifying the marker sequence. A "molecular marker probe" is a nucleic acid sequence or molecule that can be used to identify the presence of a marker locus, e.g., a nucleic acid probe that is complementary to a marker locus sequence. Alternatively, in some aspects, a molecular marker probe refers to a probe of any type that is able to distinguish (i.e., genotype) the particular allele that is present at a marker locus. Nucleic acids are "complementary" when they specifically hybridize in solution, e.g., according to Watson-Crick base pairing rules. Any suitable marker detection technology may be used to identify such a hybridization marker, e.g., SSR technology is used in the examples provided herein.

A "marker allele", alternatively an "allele of a marker locus", can refer to one of a plurality of polymorphic nucleotide sequences found at a marker locus in a population that is polymorphic for the marker locus.

The term "interval" refers to a continuous linear span of chromosomal DNA with termini that are typically defined by and including molecular markers.

The terms "nucleic acid," "nucleotide", "polynucleotide," "polynucleotide sequence" and "nucleic acid sequence" refer to single-stranded or double-stranded deoxyribonucleotide or ribonucleotide polymers, or chimeras thereof. As used herein, the term can additionally or alternatively include analogs of naturally occurring nucleotides having the essential nature of natural nucleotides in that they hybridize to single-stranded nucleic acids in a manner similar to naturally occurring nucleotides (e.g., peptide nucleic acids). Unless otherwise indicated, a particular nucleic acid sequence of this invention optionally encompasses complementary sequences, in addition to the sequence explicitly indicated. The term "gene" is used to refer to, e.g., a cDNA and an mRNA encoded by the genomic sequence, as well as to that genomic sequence.

The term "homologous" refers to nucleic acid sequences that are derived from a common ancestral gene through natural or artificial processes (e.g., are members of the same gene family), and thus, typically, share sequence similarity. Typically, homologous nucleic acids have sufficient sequence identity that one of the sequences or its complement is able to selectively hybridize to the other under selective hybridization conditions. The term "selectively hybridizes" includes reference to hybridization, under stringent hybridization conditions, of a nucleic acid sequence to a specified nucleic acid target sequence to a detectably greater degree (e.g., at least 2-fold over background) than its hybridization to non-target nucleic acid sequences and to the substantial exclusion of non-target nucleic acids. Selectively hybridizing sequences have about at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity with each other. A nucleic acid that exhibits at least some degree of homology to a reference nucleic acid can be unique or identical to the reference nucleic acid or its complementary sequence.

The term "isolated" refers to material, such as a nucleic acid or a protein, which is substantially free from components that normally accompany or interact with it in its naturally occurring environment. The isolated material optionally comprises material not found with the material in its natural environment, e.g., a cell. In addition, if the material is in its natural environment, such as a cell, the material has been placed at a location in the cell (e.g., genome or subcellular organelle) not native to a material found in that environment. For example, a naturally occurring nucleic acid (e.g., a promoter) is considered to be isolated if it is introduced by non-naturally occurring means to a locus of the genome not native to that nucleic acid. Nucleic acids that are "isolated" as defined herein, are also referred to as "heterologous" nucleic acids.

The term "recombinant" indicates that the material (e.g., a nucleic acid or protein) has been synthetically (non-naturally) altered by human intervention. The alteration to yield the synthetic material can be performed on the material within or removed from its natural environment or state. For example, a naturally occurring nucleic acid is considered a recombinant nucleic acid if it is altered, or if it is transcribed from DNA that has been altered, by means of human intervention performed within the cell from which it originates. See, e.g., Compounds and Methods for Site Directed Mutagenesis in Eukaryotic Cells,

Kmiec, U.S. Patent No. 5,565,350; *In Vivo* Homologous Sequence Targeting in Eukaryotic Cells; Zarling *et al.*, PCT/US93/03868.

The term "introduced" when referring to a heterologous or isolated nucleic acid refers to the incorporation of a nucleic acid into a eukaryotic or prokaryotic cell where the nucleic acid can be incorporated into the genome of the cell (e.g., chromosome, plasmid, plastid or mitochondrial DNA), converted into an autonomous replicon, or transiently expressed (e.g., transfected mRNA). The term includes such nucleic acid introduction means as "transfection," "transformation" and "transduction."

The terms "SSR" or "simple sequence repeat" refers to a polymorphic locus present in nuclear and organellar DNA that consist of repeating units of 1-6 base pairs in length. Different alleles can have different numbers of the repeating SSR, resulting in different lengths of the alleles, as detectable, for example, by gel electrophoresis after amplification of the allele. For example, a di-nucleotide repeat would be GAGAGAGA and a tri-nucleotide repeat would be ATGATGATGATG. It is believed that when DNA is being replicated, errors occur in the process and extra sets of these repeated sequences are added to the strand. Over time, these repeated sequences vary in length between one cultivar and another. An example of an allelic variation in SSRs would be: Allele A: GAGAGAGA (4 repeats of the GA sequence) and Allele B: GAGAGAGAGAGA (6 repeats of the GA sequence). These variations in length are easy to trace in the lab and allow tracking of genotypic variation in breeding programs.

The term "microsatellite" is an alternative term for SSR.

The term "single nucleotide polymorphism" or "SNP" is a DNA sequence variation occurring when a single nucleotide — A, T, C, or G — in the genome (or other shared sequence) differs between members of a species (or between paired chromosomes in an individual). For example, two sequenced DNA fragments from different individuals, AAGCCTA to AAGCTTA, contain a difference in a single nucleotide. In this case we say that there are two alleles: C and T. Almost all common SNPs have only two alleles.

The term "host cell" means a cell that contains a heterologous nucleic acid, such as a vector, and supports the replication and/or expression of the nucleic acid. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells. The host cells can be monocotyledonous or dicotyledonous plant cells. The dicotyledonous host cell can be, for example, a canola host cell.

The term "transgenic plant" refers to a plant that comprises within its genome a heterologous polynucleotide. Generally, the heterologous polynucleotide is stably integrated within the genome such that the polynucleotide is passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of a recombinant expression cassette. "Transgenic" is used herein to refer to any cell, cell line, callus, tissue, plant part or plant, the genotype of which has been altered by the presence of heterologous nucleic acid including those transgenic organisms or cells initially so altered, as well as those created by crosses or asexual propagation from the initial transgenic organism or cell. The term "transgenic" as used herein does not encompass the alteration of the genome (chromosomal or extra-chromosomal) by conventional plant breeding methods (i.e., crosses) or by naturally occurring events such as random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non-recombinant transposition, or spontaneous mutation.

The term "dicot" refers to the subclass of angiosperm plants also knows as "dicotyledoneae" and includes reference to whole plants, plant organs (e.g., leaves, stems, roots, etc.), seeds, plant cells, and progeny of the same. Plant cell, as used herein includes, without limitation, seeds, suspension cultures, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, and microspores.

The term "crossed" or "cross" in the context of this invention means the fusion of gametes via pollination to produce progeny (i.e., cells, seeds, or plants). The term encompasses both sexual crosses (the pollination of one plant by another) and selfing (self-pollination, i.e., when the pollen and ovule are from the same plant).

The term "introgression" refers to the transmission of a desired allele of a genetic locus from one genetic background to another. For example, introgression of a desired allele at a specified locus can be transmitted to at least one progeny plant via a sexual cross between two parent plants, where at least one of the parent plants has the desired allele within its genome. Alternatively, for example, transmission of an allele can occur by recombination between two donor genomes, e.g., in a fused protoplast, where at least one of the donor protoplasts has the desired allele in its genome. The desired allele can be, e.g., a transgene or a selected allele of a marker or QTL.

The term "extreme disease pressure field research conditions" means controlled disease research conditions as described, for Example, in WO 2006/135717. For example,

extreme disease pressure can be generated with the application of Niger seed carrier mimicking *Sclerotinia*- colonized petals. Natural inoculum may be present in the field as a backup inoculum. The percent disease incidence of the test plants are adjusted to running checks, and given an SSDIS score of 1 to 9. However, under these extreme conditions plants are more susceptible to *Sclerotinia* for at least the following reasons: (1) under extreme disease pressure field research conditions, the plants are subjected to wetness provided by misting irrigation, which is favorable for *Sclerotinia* development, (2) under extreme disease pressure field research conditions the plants are in a semi-enclosed environment due to the artificial canopy, which ensures continuous moist conditions favorable for *Sclerotinia* development, and (3) under extreme disease pressure field research conditions there are six rows of different test plants in each plot, therefore any one row of test plants having a particular morphological phenotype may be surrounded by two different rows of plants with different morphological phenotypes. Accordingly, any benefits from a morphological phenotype that is less conducive to *Sclerotinia* infection (for example high branching) are decreased because any one row may be surrounded by plants having a different morphological phenotype (for example low branching). In contrast, a plant growing under natural field conditions is (1) not enclosed in an artificial canopy, which ensures continuous moisture and (2) is grown in plots surrounded by plants with the same morphological phenotype, which allows all benefits from the morphology to be expressed. Accordingly, selections having a morphology that is less conducive to *Sclerotinia* infection, for example high branching, perform significantly better under natural field conditions compared to extreme disease pressure field research conditions.

It is well established that generating reliable field data on an annual basis is not common. *Sclerotinia* is a potent disease but it only develops during wet summers with moderate temperatures. A number of issues become critical in screening for *Sclerotinia* resistance in the field in years when the conditions of *Sclerotinia* are sub-optimal. Duration of wetness, water quality, availability of inoculum, and presence of moist or humid microenvironments affect disease development in the crop. Although the methods described are directed to *Brassica*, it is to be understood that the methods may be applied to any plant susceptible to *Sclerotinia* infection via ascospores. This includes sunflower (head rot), safflower (head rot), dry bean (pod rot), dry pea (pod rot), soybean (stem and pod rot),

alfalfa (blossom blight), and lettuce (lettuce drop). Bardin and Huang 2001. See also U.S. Patent application publication 2003/0150016 for *Sclerotinia* effects in soybean.

The critical issues in the field have been resolved as follows:

(a) Appropriate artificial inoculum for continuum of data collection: Since natural inoculum is not always triggered in the field, an inoculum that mimics infection via petals has been developed. The carrier for the fungus can be Niger seed (Guizotia abyssinica-Nyer seed) colonized with *Sclerotinia* and distributed at the time of full petal drop.

(b) Water quality and *Sclerotinia*: Initially, ground water was used to irrigate the *Sclerotinia* colonized fields. However, a lack of infection transfer in years with low rainfall and either high or low temperatures was observed. *In vitro* tests have confirmed that ground water inhibits *Sclerotinia* growth. Through lab and field testing, it was determined that deionizing (DI) water treatment alters the ground water quality sufficiently to prevent inhibition of *Sclerotinia* development. Henceforth, DI water was used to irrigate extreme disease pressure field research plots. In theory, the treated deionized water differs from the original ground water in that the minerals, for example magnesium and calcium (lime), are reduced while the pH is not affected. *Sclerotinia* produces oxalic acid, a diffusible toxin, to aid in the infection process (US 6,380,461). Calcium can bind with oxalic acid to create calcium oxalate. Removal of calcium is very likely the qualitative change in the deionized water that enables growth of *Sclerotinia*. Accordingly, a water source low in minerals or having no minerals, for example reduced or eliminated magnesium and calcium, can be used.

(c) Irrigation operated by leaf-wetness sensors: To enable continuous wetness in the field, leaf wetness sensors (Campbell Scientific) that trigger irrigation only if moisture is lower than a set threshold are used. Optimized irrigation enables disease development and enhances screening for disease resistance. However, excess irrigation may interfere with meaningful evaluation. In particular, in a research setting with rows of unique genotypes in close proximity, lodging of one entry can lead to transfer of the pathogen by plant-to-plant contact and increased disease incidence on a second genotype. Thus the *Sclerotinia* resistance score for the second genotype may underestimate its potential performance in a more homogeneous population. In natural field data trials, excessive irrigation can create a more conducive environment for *Sclerotinia* through an increase in lodging over what is

usual for a given genotype. Thus, the performance of the trial entries may be distorted due to excessive irrigation.

(d) Providing an enclosure to help maintain a microenvironment necessary for disease development: To enable development of disease in dry, hot and/or windy seasons, a netting enclosure may be used.

The new methodologies enable controlled disease development, reliable expression of phenotype, and characterization of many different lines under optimal *Sclerotinia* conditions.

Markers

The present invention provides molecular markers genetically linked to quantitative trait loci ("QTLs") associated with whole plant field resistance to *Sclerotinia*. Such molecular markers are useful for identifying and producing dicotyledonous plants, in particular, such commercially important dicot crops as sunflower, canola, alfalfa, and soybean, resistant, or with improved resistance, to *Sclerotinia*.

Genetic mapping of several hundred molecular markers has developed a genetic linkage map covering approximately 1400 cM (centiMorgans) corresponding to the 19 canola chromosomes. Additional details regarding the nature and use of molecular markers are provided below in the section entitled "MARKER ASSISTED SELECTION AND BREEDING OF PLANTS."

Exemplary marker loci associated with whole plant field resistance to *Sclerotinia* are localized to thirteen linkage groups in *Brassica napus*: N1, N3, N4, N7, N8, N9, N10, N11, N12, N13, N15, N18 and N19. These exemplary marker loci delineate chromosomal intervals including Quantitative Trait Loci (QTLs) associated with phenotypic measures of whole plant field resistance or improved whole plant field resistance to *Sclerotinia*. For example, Tables 5 and 11 herein list markers that localize to linkage groups N1, N3, N4, N7, N8, N9, N10, N11, N12, N13, N15, N18 and N19. Additional primers and probes corresponding to these markers or fragments of these markers can be designed based on the sequence information provided herein.

AG0023 (SEQ ID NO:1); AG0045 (SEQ ID NO:2); AG0047 (SEQ ID NO:3); AG0070 (SEQ ID NO:4); AG0086 (SEQ ID NO:5); AG0093 (SEQ ID NO:6); AG0125 (SEQ ID NO:7); AG0148 (SEQ ID NO:8); AG0171 (SEQ ID NO:9); AG0203 (SEQ ID NO:10); AG0239 (SEQ ID NO:11); AG0243 (SEQ ID NO:12); AG0272 (SEQ ID NO:13);

AG0304 (SEQ ID NO:14); AG0323 (SEQ ID NO:15); AG0324 (SEQ ID NO:16); AG0328 (SEQ ID NO:17); AG0359 (SEQ ID NO:18); AG0369 (SEQ ID NO:19); AG0370 (SEQ ID NO:20); AG0378 (SEQ ID NO:21); AG0391 (SEQ ID NO:22); AG0410 (SEQ ID NO:23); AG0441 (SEQ ID NO:24); AG0477 (SEQ ID NO:25); AG0482 (SEQ ID NO:26); AG0504 (SEQ ID NO:27); AG0510 (SEQ ID NO:28); BG0031 (SEQ ID NO:29); BG0106 (SEQ ID NO:30); BG0111 (SEQ ID NO:31); BG0119 (SEQ ID NO:32); BG0181 (SEQ ID NO:33); BG0228 (SEQ ID NO:34); BG0255 (SEQ ID NO:35); BG0278 (SEQ ID NO:36); BG0295 (SEQ ID NO:37); BG0452 (SEQ ID NO:38); BG0516 (SEQ ID NO:39); BG0647 (SEQ ID NO:40); BG0651 (SEQ ID NO:41); BG0713 (SEQ ID NO:42); BG0864 (SEQ ID NO:43); BG0869 (SEQ ID NO:44); BG0988 (SEQ ID NO:45); BG1062 (SEQ ID NO:46); BG1090 (SEQ ID NO:47); BG1101 (SEQ ID NO:48); BG1123 (SEQ ID NO:49); BG1127 (SEQ ID NO:50); BG1149 (SEQ ID NO:51); BG1182 (SEQ ID NO:52); BG1197 (SEQ ID NO:53); BG1230 (SEQ ID NO:54); BG1241 (SEQ ID NO:55); BG1244 (SEQ ID NO:56); BG1286 (SEQ ID NO:57); BG1288 (SEQ ID NO:58); BG1321 (SEQ ID NO:59); BG1368 (SEQ ID NO:60); BG1392 (SEQ ID NO:61); BG1442 (SEQ ID NO:62); BG1449 (SEQ ID NO:63); BG1453 (SEQ ID NO:64); BG1513 (SEQ ID NO:65); CA0105 (SEQ ID NO:66); CA0120 (SEQ ID NO:67); CA0163 (SEQ ID NO:68); CA0221 (SEQ ID NO:69); CA0226 (SEQ ID NO:70); CA0233 (SEQ ID NO:71); CA0328 (SEQ ID NO:72); CA0410 (SEQ ID NO:73); CA0423 (SEQ ID NO:74); CA0456 (SEQ ID NO:75); CA0488 (SEQ ID NO:76); CA0546 (SEQ ID NO:77); CA0552 (SEQ ID NO:78); CA0603 (SEQ ID NO:79); CA0614 (SEQ ID NO:80); CA0636 (SEQ ID NO:81); CA0681 (SEQ ID NO:82); CA0719 (SEQ ID NO:83); CA0736 (SEQ ID NO:84); CA0739 (SEQ ID NO:85); CA0753 (SEQ ID NO:86); CA0834 (SEQ ID NO:87); CA0837 (SEQ ID NO:88); CA0896 (SEQ ID NO:89); CA0991 (SEQ ID NO:90); CA1027 (SEQ ID NO:91); CA1032 (SEQ ID NO:92); CA1034 (SEQ ID NO:93); CA1035 (SEQ ID NO:94); CA1066 (SEQ ID NO:95); CA1080 (SEQ ID NO:96); CA1090 (SEQ ID NO:97); CA1097 (SEQ ID NO:98); CA1107 (SEQ ID NO:99); PE0012 (SEQ ID NO:100); PE0017 (SEQ ID NO:101); PE0063 (SEQ ID NO:102); PE0091 (SEQ ID NO:103); PE0131 (SEQ ID NO:104); PE0133 (SEQ ID NO:105); PE0177 (SEQ ID NO:106); PE0187 (SEQ ID NO:107); PE0203 (SEQ ID NO:108); PE0250 (SEQ ID NO:109); PE0281 (SEQ ID NO:110); PE0283 (SEQ ID NO:111); PE0286 (SEQ ID NO:112); PE0324 (SEQ ID NO:113); PE0340 (SEQ ID NO:114); PE0355 (SEQ ID NO:115); UB0015 (SEQ ID NO:116); UB0126 (SEQ ID NO:117); UB0163 (SEQ ID

NO:118); UB0181 (SEQ ID NO:119); UB0196 (SEQ ID NO:120); UB0307 (SEQ ID NO:121); UB0315 (SEQ ID NO:122); UB0331 (SEQ ID NO:123); KK66 (SEQ ID NO:124); and KK98G (SEQ ID NO:125) (sometimes referred to as "the markers exemplified by SEQ ID NOs: 1-125"). contain simple sequence repeat (SSR) polymorphisms or single nucleotide polymorphisms (SNPs) that identify QTLs contributing to *Sclerotinia* whole plant field resistance or improved whole plant field resistance and can be used as markers thereof. It will be appreciated that the number of repeats in the SSR can vary. Favorable alleles that contribute to whole plant field resistance to *Sclerotinia* or improved whole plant field resistance to *Sclerotinia* are provided, for example, in Tables 7 and 13.

It will be noted that, regardless of their molecular nature, e.g., whether the marker is an SSR, AFLP, RFLP, etc., markers are typically strain specific. That is, a particular polymorphic marker, such as the exemplary markers of the invention described above, is defined relative to the parental lines of interest. For each marker locus, resistance-associated, and conversely, susceptibility-associated alleles are identified for each pair of parental lines. Following correlation of specific alleles with susceptibility and resistance in parents of a cross, the marker can be utilized to identify progeny with genotypes that correspond to the desired resistance phenotype. In some circumstance, i.e., in some crosses of parental lines, the exemplary markers described herein will not be optimally informative. In such cases, additional informative markers, e.g., certain linked markers and/or homologous markers are evaluated and substituted for genotyping, e.g., for marker-assisted selection, etc. In the case where a marker corresponds to a QTL, following identification of resistance- and susceptibility-associated alleles, it is possible to directly screen a population of samples, e.g., samples obtained from a seed bank, without first correlating the parental phenotype with an allele.

Linked Markers

Those of skill in the art will recognize that additional molecular markers can be identified within the intervals defined by the above-described pairs of markers. Such markers are also genetically linked to the QTLs identified herein as associated with *Sclerotinia* whole plant field resistance, and are within the scope of the present invention. Markers can be identified by any of a variety of genetic or physical mapping techniques. Methods of determining whether markers are genetically linked to a QTL (or to a specified marker) associated with resistance to *Sclerotinia* are known to those of skill in the art and

include, e.g., interval mapping (Lander and Botstein (1989) Genetics 121:185), regression mapping (Haley and Knott (1992) Heredity 69:315) or MQM mapping (Jansen (1994) Genetics 138:871). In addition, such physical mapping techniques as chromosome walking, contig mapping and assembly, and the like, can be employed to identify and isolate additional sequences useful as markers in the context of the present invention.

Homologous Nucleotide Sequences

In addition, AG0023; AG0045; AG0047; AG0070; AG0086; AG0093; AG0125; AG0148; AG0171; AG0203; AG0239; AG0243; AG0272; AG0304; AG0323; AG0324; AG0328; AG0359; AG0369; AG0370; AG0378; AG0391; AG0410; AG0441; AG0477; AG0482; AG0504; AG0510; BG0031; BG0106; BG0111; BG0119; BG0181; BG0228; BG0255; BG0278; BG0295; BG0452; BG0516; BG0647; BG0651; BG0713; BG0864; BG0869; BG0988; BG1062; BG1090; BG1101; BG1123; BG1127; BG1149; BG1182; BG1197; BG1230; BG1241; BG1244; BG1286; BG1288; BG1321; BG1368; BG1392; BG1442; BG1449; BG1453; BG1513; CA0105; CA0120; CA0163; CA0221; CA0226; CA0233; CA0328; CA0410; CA0423; CA0456; CA0488; CA0546; CA0552; CA0603; CA0614; CA0636; CA0681; CA0719; CA0736; CA0739; CA0753; CA0834; CA0837; CA0896; CA0991; CA1027; CA1032; CA1034; CA1035; CA1066; CA1080; CA1090; CA1097; and CA1107; as well as PE0012; PE0017; PE0063; PE0091; PE0131; PE0133; PE0177; PE0187; PE0203; PE0250; PE0281; PE0283; PE0286; PE0324; PE0340; PE0355; UB0015; UB0126; UB0163; UB0181; UB0196; UB0307; UB0315; UB0331; KK66; and KK98G are useful for the identification of homologous nucleotide sequences with utility in identifying QTLs associated with *Sclerotinia* whole plant field resistance in different lines, varieties, or species of dicots. Such homologous markers are also a feature of the invention.

Such homologous sequences can be identified by selective hybridization to a reference sequence. The reference sequence is typically a unique sequence, such as a unique oligonucleotide primer sequence, EST, amplified fragment (e.g., corresponding to AFLP markers) and the like, derived from any of the marker loci listed herein or its complement.

Two single-stranded nucleic acids "hybridize" when they form a double-stranded duplex. The double stranded region can include the full-length of one or both of the single-stranded nucleic acids, or all of one single stranded nucleic acid and a subsequence of the other single-stranded nucleic acid, or the double stranded region can include a subsequence of each nucleic acid. Selective hybridization conditions distinguish between nucleic acids

that are related, e.g., share significant sequence identity with the reference sequence (or its complement) and those that associate with the reference sequence in a non-specific manner. Generally, selective hybridization conditions are those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Selective hybridization conditions may also be achieved with the addition of destabilizing agents such as formamide. Selectivity can be achieved by varying the stringency of the hybridization and/or wash conditions. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulphate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 60°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C.

Specificity is typically a function of post-hybridization washes, with the critical factors being ionic strength and temperature of the final wash solution. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4 °C lower than the thermal melting point (T_m); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10 °C lower than the thermal melting point (T_m); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20 °C lower than the thermal melting point (T_m).

The T_m is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe. For DNA-DNA hybrids, the T_m can be approximated from the equation of Meinkoth and Wahl ((1984) Anal. Biochem. 138:267-284): $T_m = 81.5\ °C + 16.6\ (\log M) + 0.41\ (\%GC) - 0.61\ (\%\ form) - 500/L$; where M is the molarity of monovalent cations, %GC is the percentage of guanosine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. T_m is reduced by about 1° C for each 1% of mismatching; thus, T_m, hybridization and/or wash conditions can be

adjusted to hybridize to sequences of the desired identity. For example, if sequences with ≥90% identity are sought, the T_m can be decreased 10° C.

Using the equation, hybridization and wash compositions, and desired T_m, those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a T_m of less than 45 °C (aqueous solution) or 32 °C (formamide solution) it is preferred to increase the SSC concentration so that a higher temperature can be used. Hybridization and/or wash conditions can be applied for at least 10, 30, 60, 90, 120, or 240 minutes. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays" Elsevier, New York. General Texts that discuss considerations relevant to nucleic acid hybridization, the selection of probes, and buffer and incubation conditions, and the like, as well as numerous other topics of interest in the context of the present invention (e.g., cloning of nucleic acids that correspond to markers and QTLs, sequencing of cloned markers/QTLs, the use of promoters, vectors, etc.) can be found in Berger and Kimmel (1987) Guide to Molecular Cloning Techniques, Methods in Enzymology vol.152, Academic Press, Inc., San Diego ("Berger"); Sambrook et al., (2001) Molecular Cloning-A Laboratory Manual, 3rd ed. Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor ("Sambrook"); and Ausubel et al., (eds) (supplemented through 2001) Current Protocols in Molecular Biology, John Wiley and Sons, Inc., ("Ausubel").

In addition to hybridization methods described above, homologs of the markers of the invention can be identified *in silico* using any of a variety of sequence alignment and comparison protocols. For the purposes of the ensuing discussion, the following terms are used to describe the sequence relationships between a marker nucleotide sequence and a reference polynucleotide sequence:

A "reference sequence" is a defined sequence used as a basis for sequence comparison with a test sequence, e.g., a candidate marker homolog, of the present invention. A reference sequence may be a subsequence or the entirety of a specified sequence; for example, a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence.

As used herein, a "comparison window" is a contiguous and specified segment, (e.g., a subsequence) of a polynucleotide/polypeptide sequence to be compared to a reference sequence. The segment of the polynucleotide/polypeptide sequence in the comparison window can include one or more additions or deletions (i.e., gaps) with respect to the reference sequence, which (by definition) does not comprise addition(s) or deletion(s), for optimal alignment of the two sequences. An optimal alignment of two sequences yields the fewest number of unlike nucleotide/amino acid residues in a comparison window. Generally, the comparison window is at least 20 contiguous nucleotide/amino acid residues in length, and optionally can be 30, 40, 50, 100, or longer. Those of skill in the art understand that to avoid a falsely high similarity between two sequences, due to inclusion of gaps in the polynucleotide/polypeptide sequence, a gap penalty is typically assessed and is subtracted from the number of matches.

"Sequence identity" or "identity" in the context of two nucleic acid or polypeptide sequences refers to residues that are the same in both sequences when aligned for maximum correspondence over a specified comparison window.

"Percentage sequence identity" refers to the value determined by comparing two optimally aligned sequences over a comparison window. The percentage is calculated by determining the number of positions at which both sequences have the same nucleotide or amino acid residue, determining the number of matched positions, dividing the number of matched positions by the total number of positions in the comparison window, and multiplying the result by 100 to yield the percentage of sequence identity.

When percentage of sequence identity is used in reference to proteins it is recognized that residue positions that are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ by conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences that differ by such conservative substitutions are said to have "sequence similarity" or "similarity". Means for making this adjustment are well-known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a

non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, e.g., according to the algorithm of Meyers and Miller (1988) Computer Applic. Biol. Sci. 4:11-17, e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California, USA).

Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman ((1981) Adv. Appl. Math. 2:482); by the homology alignment algorithm of Needleman and Wunsch ((1970) J. Mol. Biol. 48:443); by the search for similarity method of Pearson and Lipman ((1988) Proc. Natl. Acad. Sci. USA 85:2444); by computerized implementations of these algorithms, including, but not limited to: CLUSTAL in the PC/Gene program by Intelligenetics, Mountain View, California; GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), Madison, Wisconsin, USA; the CLUSTAL program is well described by Higgins and Sharp ((1988) Gene 73:237-244); Higgins and Sharp ((1989) CABIOS 5:151-153); Corpet et al. ((1988) Nucleic Acids Research 16:10881-90); Huang et al. ((1992) Computer Applications in the Biosciences 8: 155-65), and Pearson et al. ((1994) Methods in Molecular Biology 24:307-331).

The BLAST family of programs that can be used for database similarity searches includes: BLASTN for nucleotide query sequences against nucleotide database sequences; BLASTX for nucleotide query sequences against protein database sequences; BLASTP for protein query sequences against protein database sequences; TBLASTN for protein query sequences against nucleotide database sequences; and TBLASTX for nucleotide query sequences against nucleotide database sequences. See, e.g., Current Protocols in Molecular Biology, Chapter 19, Ausubel et al., Eds., (1995) Greene Publishing and Wiley-Interscience, New York; Altschul et al. (1990) J. Mol. Biol. 215:403-410; and, Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402.

Software for performing BLAST analyses is publicly available, e.g., through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is

referred to as the neighborhood word score threshold. These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see, e.g., Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915).

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see*, e.g., Karlin & Altschul (1993) Proc. Nat'l. Acad. Sci. USA 90:5873-5877). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance.

BLAST searches assume that proteins can be modeled as random sequences. However, many real proteins comprise regions of nonrandom sequences that may be homopolymeric tracts, short-period repeats, or regions enriched in one or more amino acids. Such low-complexity regions may be aligned between unrelated proteins even though other regions of the protein are entirely dissimilar. A number of low-complexity filter programs can be employed to reduce such low-complexity alignments. For example, the SEG (Wooten and Federhen (1993) Comput. Chem. 17:149-163) and XNU (Claverie and States (1993) Comput. Chem. 17:191-201) low-complexity filters can be employed alone or in combination.

Unless otherwise stated, nucleotide and protein identity/similarity values provided herein are calculated using GAP (GCG Version 10) under default values.

GAP (Global Alignment Program) can also be used to compare a polynucleotide or polypeptide of the present invention with a reference sequence. GAP uses the algorithm of Needleman and Wunsch ((1970) J. Mol. Biol. 48: 443-453), to find the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. GAP considers all possible alignments and gap positions and creates the alignment with the largest number of matched bases and the fewest gaps. It allows for the provision of a gap creation penalty and a gap extension penalty in units of matched bases. GAP must make a profit of gap creation penalty number of matches for each gap it inserts. If a gap extension penalty greater than zero is chosen, GAP must, in addition, make a profit for each gap inserted of the length of the gap times the gap extension penalty. Default gap creation penalty values and gap extension penalty values in Version 10 of the Wisconsin Genetics Software Package for protein sequences are 8 and 2, respectively. For nucleotide sequences the default gap creation penalty is 50 while the default gap extension penalty is 3. The gap creation and gap extension penalties can be expressed as an integer selected from the group of integers consisting of from 0 to 100. Thus, for example, the gap creation and gap extension penalties can each independently be: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60 or greater.

GAP presents one member of the family of best alignments. There may be many members of this family, but no other member has a better quality. GAP displays four figures of merit for alignments: Quality, Ratio, Identity, and Similarity. The Quality is the metric maximized in order to align the sequences. Ratio is the quality divided by the number of bases in the shorter segment. Percent Identity is the percent of the symbols that actually match. Percent Similarity is the percent of the symbols that are similar. Symbols that are across from gaps are ignored. A similarity is scored when the scoring matrix value for a pair of symbols is greater than or equal to 0.50, the similarity threshold. The scoring matrix used in Version 10 of the Wisconsin Genetics Software Package is BLOSUM62 (see, e.g., Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915).

Multiple alignment of the sequences can be performed using the CLUSTAL method of alignment (Higgins and Sharp (1989) CABIOS. 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise

alignments using the CLUSTAL method are KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

The percentage sequence identity of a homologous marker to its reference marker (e.g., any one of the markers described herein) is typically at least 70% and, rounded upwards to the nearest integer, can be expressed as an integer selected from the group of integers between 70 and 99. Thus, for example, the percentage sequence identity to a reference sequence can be at least 70%, 75%, 80%, 85%, 90%, 95%, 97%, or 99%. Sequence identity can be calculated using, for example, the BLAST, CLUSTALW, or GAP algorithms under default conditions.

Detection of Marker Loci

Markers corresponding to genetic polymorphisms between members of a population can be detected by numerous methods, well-established in the art (e.g., restriction fragment length polymorphisms, isozyme markers, allele specific hybridization (ASH), amplified variable sequences of the plant genome, self-sustained sequence replication, simple sequence repeat (SSR), single nucleotide polymorphism (SNP), or amplified fragment length polymorphisms (AFLP)).

The majority of genetic markers rely on one or more properties of nucleic acids for their detection. For example, some techniques for detecting genetic markers utilize hybridization of a probe nucleic acid to nucleic acids corresponding to the genetic marker. Hybridization formats include but are not limited to, solution phase, solid phase, mixed phase, or *in situ* hybridization assays. Markers that are restriction fragment length polymorphisms (RFLP), are detected by hybridizing a probe, which is typically a sub-fragment (or a synthetic oligonucleotide corresponding to a sub-fragment) of the nucleic acid to be detected to restriction digested genomic DNA. The restriction enzyme is selected to provide restriction fragments of at least two alternative (or polymorphic) lengths in different individuals, and will often vary from line to line. Determining a (one or more) restriction enzyme that produces informative fragments for each cross is a simple procedure, well known in the art. After separation by length in an appropriate matrix (e.g., agarose) and transfer to a membrane (e.g., nitrocellulose, nylon), the labeled probe is hybridized under conditions that result in equilibrium binding of the probe to the target followed by removal of excess probe by washing.

Nucleic acid probes to the marker loci can be cloned and/or synthesized. Detectable labels suitable for use with nucleic acid probes include any composition detectable by spectroscopic, radioisotopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels include biotin for staining with labeled streptavidin conjugate, magnetic beads, fluorescent dyes, radiolabels, enzymes, and colorimetric labels. Other labels include ligands that bind to antibodies labeled with fluorophores, chemiluminescent agents, and enzymes. Labeling markers is readily achieved such as by the use of labeled PCR primers to marker loci.

The hybridized probe is then detected using, most typically by autoradiography or other similar detection technique (e.g., fluorography, liquid scintillation counter, etc.). Examples of specific hybridization protocols are widely available in the art, see, e.g., Berger, Sambrook, Ausubel, all *supra*.

Amplified variable sequences refer to amplified sequences of the plant genome that exhibit high nucleic acid residue variability between members of the same species. All organisms have variable genomic sequences and each organism (with the exception of a clone) has a different set of variable sequences. Once identified, the presence of specific variable sequence can be used to predict phenotypic traits. Preferably, DNA from the plant serves as a template for amplification with primers that flank a variable sequence of DNA. The variable sequence is amplified and then sequenced.

In vitro amplification techniques are well known in the art. Examples of techniques sufficient to direct persons of skill through such *in vitro* methods, including the polymerase chain reaction (PCR), the ligase chain reaction (LCR), Qβ-replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA), are found in Berger, Sambrook and Ausubel (all *supra*) as well as Mullis et al. ((1987) U.S. Patent No. 4,683,202); PCR Protocols, A Guide to Methods and Applications ((Innis et al., eds.) Academic Press Inc., San Diego Academic Press Inc. San Diego, CA (1990) (Innis)); Arnheim & Levinson ((October 1, 1990) C&EN 36-47); The Journal Of NIH Research (1991) 3, 81-94; Kwoh et al. ((1989) Proc. Natl. Acad. Sci. USA 86, 1173); Guatelli et al. ((1990) Proc. Natl. Acad. Sci. USA 87, 1874); Lomell et al. ((1989) J. Clin. Chem. 35, 1826); Landegren et al. ((1988) Science 241, 1077-1080); Van Brunt ((1990) Biotechnology 8, 291-294); Wu and Wallace ((1989) Gene 4, 560); Barringer et al. ((1990) Gene 89, 117), and Sooknanan and Malek ((1995) Biotechnology 13: 563-564). Improved methods of cloning *in vitro* amplified

nucleic acids are described in Wallace et al., U.S. Pat. No. 5,426,039. Improved methods of amplifying large nucleic acids by PCR are summarized in Cheng et al. (1994) Nature 369: 684, and the references therein, in which PCR amplicons of up to 40kb are generated. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. *See*, Ausubel, Sambrook and Berger, *all supra.*

Oligonucleotides for use as primers, e.g., in amplification reactions and for use as nucleic acid sequence probes are typically synthesized chemically according to the solid phase phosphoramidite triester method described by Beaucage and Caruthers ((1981) Tetrahedron Lett. 22:1859), or can simply be ordered commercially.

Alternatively, self-sustained sequence replication can be used to identify genetic markers. Self-sustained sequence replication refers to a method of nucleic acid amplification using target nucleic acid sequences that are replicated exponentially *in vitro* under substantially isothermal conditions by using three enzymatic activities involved in retroviral replication: (1) reverse transcriptase, (2) Rnase H, and (3) a DNA-dependent RNA polymerase (Guatelli et al. (1990) Proc. Natl. Acad. Sci. USA 87:1874). By mimicking the retroviral strategy of RNA replication by means of cDNA intermediates, this reaction accumulates cDNA and RNA copies of the original target.

Amplified fragment length polymorphisms (AFLP) can also be used as genetic markers (Vos et al. (1995) Nucl. Acids Res. 23:4407. The phrase "amplified fragment length polymorphism" refers to selected restriction fragments that are amplified before or after cleavage by a restriction endonuclease. The amplification step allows easier detection of specific restriction fragments. AFLP allows the detection large numbers of polymorphic markers and has been used for genetic mapping of plants (Becker et al. (1995) Mol. Gen. Genet. 249:65; and Meksem et al. (1995) Mol. Gen. Genet. 249:74.

Allele-specific hybridization (ASH) can be used to identify the genetic markers of the invention. ASH technology is based on the stable annealing of a short, single-stranded, oligonucleotide probe to a completely complementary single-strand target nucleic acid. Detection is via an isotopic or non-isotopic label attached to the probe.

For each polymorphism, two or more different ASH probes are designed to have identical DNA sequences except at the polymorphic nucleotides. Each probe will have exact homology with one allele sequence so that the range of probes can distinguish all the known

alternative allele sequences. Each probe is hybridized to the target DNA. With appropriate probe design and hybridization conditions, a single-base mismatch between the probe and target DNA will prevent hybridization. In this manner, only one of the alternative probes will hybridize to a target sample that is homozygous or homogenous for an allele. Samples that are heterozygous or heterogeneous for two alleles will hybridize to both of two alternative probes.

ASH markers are used as dominant markers where the presence or absence of only one allele is determined from hybridization or lack of hybridization by only one probe. The alternative allele may be inferred from the lack of hybridization. ASH probe and target molecules are optionally RNA or DNA; the target molecules are any length of nucleotides beyond the sequence that is complementary to the probe; the probe is designed to hybridize with either strand of a DNA target; the probe ranges in size to conform to variously stringent hybridization conditions, etc.

PCR allows the target sequence for ASH to be amplified from low concentrations of nucleic acid in relatively small volumes. Otherwise, the target sequence from genomic DNA is digested with a restriction endonuclease and size separated by gel electrophoresis. Hybridizations typically occur with the target sequence bound to the surface of a membrane or, as described in U.S. Patent No. 5,468,613, the ASH probe sequence may be bound to a membrane.

In one embodiment, ASH data are obtained by amplifying nucleic acid fragments (amplicons) from genomic DNA using PCR, transferring the amplicon target DNA to a membrane in a dot-blot format, hybridizing a labeled oligonucleotide probe to the amplicon target, and observing the hybridization dots by autoradiography.

Single nucleotide polymorphisms (SNP) are markers that consist of a shared sequence differentiated on the basis of a single nucleotide. Typically, this distinction is detected by differential migration patterns of an amplicon comprising the SNP on e.g., an acrylamide gel. However, alternative modes of detection, such as hybridization, e.g., ASH, or RFLP analysis are not excluded.

In yet another basis for providing a genetic linkage map, Simple sequence repeats (SSR), take advantage of high levels of di-, tri-,tetra-, penta- or hexa-nucleotide tandem repeats within a genome. Dinucleotide repeats have been reported to occur in the human genome as many as 50,000 times with n varying from 10 to 60 or more (Jacob et al. (1991)

Cell 67:213. Dinucleotide repeats have also been found in higher plants (Condit and Hubbell (1991) Genome 34:66).

Briefly, SSR data are generated by hybridizing primers to conserved regions of the plant genome that flank the SSR sequence. PCR is then used to amplify the nucleotide repeats between the primers. The amplified sequences are then electrophoresed to determine the size and therefore the number of di-, tri-, and tetra-nucleotide repeats. The number of repeats distinguishes the favorable allele from an unfavorable allele. Favorable alleles for whole plant field resistance to *Sclerotinia* or improved whole plant field resistance to *Sclerotinia* are provided, for example, in Tables 7and 13.

Alternatively, isozyme markers are employed as genetic markers. Isozymes are multiple forms of enzymes that differ from one another in their amino acid, and therefore their nucleic acid sequences. Some isozymes are multimeric enzymes containing slightly different subunits. Other isozymes are either multimeric or monomeric but have been cleaved from the proenzyme at different sites in the amino acid sequence. Isozymes can be characterized and analyzed at the protein level, or alternatively, isozymes that differ at the nucleic acid level can be determined. In such cases any of the nucleic acid based methods described herein can be used to analyze isozyme markers.

In alternative embodiments, *in silico* methods can be used to detect the marker loci. For example, the sequence of a nucleic acid comprising the marker can be stored in a computer. The desired marker locus sequence or its homolog can be identified using an appropriate nucleic acid search algorithm as provided by, for example, in such readily available programs as BLAST.

QTL Mapping

Multiple experimental paradigms have been developed to identify and analyze QTLs. In general, these paradigms involve crossing one or more parental pairs, which can be, for example, a single pair derived from two inbred strains, or multiple related or unrelated parents of different inbred strains or lines, which each exhibit different characteristics relative to the phenotypic trait of interest. The parents and a population of progeny are genotyped, typically for multiple marker loci, and evaluated for the trait of interest. In the context of the present invention, the parental and progeny plants are genotyped for any one or more of the molecular markers exemplified herein, or homologs, or alternative markers linked to any one or more of the markers exemplified herein, and evaluated for whole plant

field resistance, or improved whole plant field resistance to *Sclerotinia*. QTLs associated with *Sclerotinia* whole plant field resistance are identified based on the significant statistical correlations between the marker genotype(s) and the resistance phenotype of the evaluated progeny plants. Numerous methods for determining whether markers are genetically linked to a QTL (or to another marker) associated with whole plant field resistance to *Sclerotinia* are known to those of skill in the art and include, e.g., interval mapping (Lander and Botstein (1989) Genetics 121:185), regression mapping (Haley and Knott (1992) Heredity 69:315) or MQM mapping (Jansen (1994) Genetics 138:871). In addition, the following applications provide additional details regarding alternative statistical methods applicable to complex breeding populations that can be used to identify and localize QTLs associated with *Sclerotinia* whole plant field resistance: USSN 09/216,089 by Beavis et al. "QTL MAPPING IN PLANT BREEDING POPULATIONS" and PCT/US00/34971 by Jansen et al. "MQM MAPPING USING HAPLOTYPED PUTATIVE QTLS ALLELES: A SIMPLE APPROACH FOR MAPPING QTLS IN PLANT BREEDING POPULATIONS."

Marker Assisted Selection and Breeding of Plants

A primary motivation for development of molecular markers in crop species is the potential for increased efficiency in plant breeding through marker assisted selection (MAS). Genetic marker alleles, or alternatively, identified QTL alleles, are used to identify plants that contain a desired genotype at one or more loci, and that are expected to transfer the desired genotype, along with a desired phenotype to their progeny. Genetic marker alleles (or QTL alleles) can be used to identify plants that contain a desired genotype at one locus, or at several unlinked or linked loci (e.g., a haplotype), and that would be expected to transfer the desired genotype, along with a desired phenotype to their progeny. The present invention provides the means to identify plants, particularly dicots, e.g., *Brassica*, that are resistant, or exhibit improved resistance, to *Sclerotinia* whole plant field resistance by identifying plants having a specified allele, e.g., at one or more of the markers exemplified herein, or other markers within the intervals set forth herein.. Similarly, by identifying plants lacking a desired allele of the marker, susceptible plants can be identified, and, e.g., eliminated from subsequent crosses. It will be appreciated that for the purposes of MAS, the term marker can encompass both marker and QTL loci as both can be used to identify plants that are resistant or have improved resistance to *Sclerotinia*.

After a desired phenotype, e.g., *Sclerotinia* whole plant field resistance, and a polymorphic chromosomal locus, e.g., a marker locus or QTL, are determined to segregate together, it is possible to use those polymorphic loci to select for alleles corresponding to the desired phenotype - a process called marker-assisted selection (MAS). In brief, a nucleic acid corresponding to the marker nucleic acid is detected in a biological sample from a plant to be selected. This detection can take the form of hybridization of a probe nucleic acid to a marker, e.g., using allele-specific hybridization, southern blot analysis, northern blot analysis, *in situ* hybridization, hybridization of primers followed by PCR amplification of a region of the marker or the like. A variety of procedures for detecting markers are described herein, e.g., in the section entitled "DETECTION OF MARKER LOCI." After the presence (or absence) of a particular marker in the biological sample is verified, the plant is selected, i.e., used to make progeny plants by selective breeding.

Plant breeders need to combine disease tolerant loci with genes for high yield and other desirable traits to develop improved plant varieties. Disease screening for large numbers of samples can be expensive, time consuming, and unreliable. Use of the polymorphic loci described herein, and genetically-linked nucleic acids, as genetic markers for disease resistance loci is an effective method for selecting tolerant varieties in breeding programs. For example, one advantage of marker-assisted selection over field evaluations for disease resistance is that MAS can be done at any time of year regardless of the growing season. Moreover, environmental effects are irrelevant to marker-assisted selection.

When a population is segregating for multiple loci affecting one or multiple traits, e.g., multiple loci involved in resistance to a single disease, or multiple loci each involved in resistance to different diseases, the efficiency of MAS compared to phenotypic screening becomes even greater because all the loci can be processed in the lab together from a single sample of DNA. In the present instance, this means that multiple markers selected from among AG0023; AG0045; AG0047; AG0070; AG0086; AG0093; AG0125; AG0148; AG0171; AG0203; AG0239; AG0243; AG0272; AG0304; AG0323; AG0324; AG0328; AG0359; AG0369; AG0370; AG0378; AG0391; AG0410; AG0441; AG0477; AG0482; AG0504; AG0510; BG0031; BG0106; BG0111; BG0119; BG0181; BG0228; BG0255; BG0278; BG0295; BG0452; BG0516; BG0647; BG0651; BG0713; BG0864; BG0869; BG0988; BG1062; BG1090; BG1101; BG1123; BG1127; BG1149; BG1182; BG1197; BG1230; BG1241; BG1244; BG1286; BG1288; BG1321; BG1368; BG1392; BG1442;

BG1449; BG1453; BG1513; CA0105; CA0120; CA0163; CA0221; CA0226; CA0233; CA0328; CA0410; CA0423; CA0456; CA0488; CA0546; CA0552; CA0603; CA0614; CA0636; CA0681; CA0719; CA0736; CA0739; CA0753; CA0834; CA0837; CA0896; CA0991; CA1027; CA1032; CA1034; CA1035; CA1066; CA1080; CA1090; CA1097; CA1107; PE0012; PE0017; PE0063; PE0091; PE0131; PE0133; PE0177; PE0187; PE0203; PE0250; PE0281; PE0283; PE0286; PE0324; PE0340; PE0355; UB0015; UB0126; UB0163; UB0181; UB0196; UB0307; UB0315; UB0331; KK66; and KK98G or markers homologous or linked thereto can be assayed simultaneously or sequentially in a single sample or population of samples. Thus, any one or more of these markers, e.g., two or more, up to and including all of the established markers, can be assayed simultaneously. In some instances, it is desirable to evaluate a marker corresponding to each of the linkage groups associated with *Sclerotinia* whole plant field resistance.

Another use of MAS in plant breeding is to assist the recovery of the recurrent parent genotype by backcross breeding. Backcross breeding is the process of crossing a progeny back to one of its parents. Backcrossing is usually done for the purpose of introgressing one or a few loci from a donor parent into an otherwise desirable genetic background from the recurrent parent. The more cycles of backcrossing that are done, the greater the genetic contribution of the recurrent parent to the resulting variety. This is often necessary, because tolerant plants may be otherwise undesirable, i.e., due to low yield, low fecundity, or the like. In contrast, strains that are the result of intensive breeding programs may have excellent yield, fecundity or the like, merely being deficient in one desired trait such as resistance to a particular pathogen (e.g., *Sclerotinia* whole plant field resistance).

The presence and/or absence of a particular genetic marker allele, or a homolog thereof, in the genome of a plant exhibiting a preferred phenotypic trait is determined by any method listed above, e.g., RFLP, AFLP, SSR, etc. If the nucleic acids from the plant are positive for a desired genetic marker, the plant can be selfed to create a true breeding line with the same genotype, or it can be crossed with a plant with the same marker or with other desired characteristics to create a sexually crossed hybrid generation.

As mentioned above, the skilled artisan will understand that the QTLs described herein represent regions of the genome comprising genes that contribute to the *Sclerotinia* whole plant field resistance of a plant. Further, each QTL can contribute differently to that resistance level. Thus, breeding efforts are directed to increasing the number of those QTLs,

particularly quantitatively significant QTLs, present in the germplasm. Early in a breeding program, fewer QTLs may be present in a particular germplasm, but that number will increase as the breeding program progresses. Thus, in certain embodiments, a plant exhibiting *Sclerotinia* whole plant field resistance may contain at least 6 of the QTLs described herein. More particularly, the plant may contain at least 7, 8, 9 or 10 of the QTLs described herein. Yet more particularly, the plant may contain 11, 12 or all of the QTLs described herein.

Positional Cloning

The molecular markers of the present invention and nucleic acids homologous thereto, can be used, as indicated previously, to identify additional linked marker loci, which can be cloned by well established procedures, e.g., as described in detail in Ausubel, Berger and Sambrook, *supra*. Similarly, the exemplified markers, as well as any additionally identified linked molecular markers can be used to physically isolate, e.g., by cloning, nucleic acids associated with QTLs contributing to *Sclerotinia* whole plant field resistance. Such nucleic acids, i.e., linked to QTLs, have a variety of uses, including as genetic markers for identification of additional QTLs in subsequent applications of marker assisted selection (MAS).

These nucleic acids are first identified by their genetic linkage to markers of the present invention. Isolation of the nucleic acid of interest is achieved by any number of methods as discussed in detail in such references as Ausubel, Berger and Sambrook, *supra*, and Clark, Ed. (1997) Plant Molecular Biology: A Laboratory Manual Springer-Verlag, Berlin.

For example, positional gene cloning uses the proximity of a genetic marker to physically define an isolated chromosomal fragment that is linked to a QTL. The isolated chromosomal fragment can be produced by such well known methods as digesting chromosomal DNA with one or more restriction enzymes, or by amplifying a chromosomal region in a polymerase chain reaction (PCR), or alternative amplification reaction. The digested or amplified fragment is typically ligated into a vector suitable for replication, e.g., a plasmid, a cosmid, a phage, an artificial chromosome, or the like, and, optionally, expression of the inserted fragment. Markers that are adjacent to an open reading frame (ORF) associated with a phenotypic trait can hybridize to a DNA clone, thereby identifying a clone on which an ORF is located. If the marker is more distant, a fragment containing the

open reading frame is identified by successive rounds of screening and isolation of clones, which together comprise a contiguous sequence of DNA, a "contig." Protocols sufficient to guide one of skill through the isolation of clones associated with linked markers are found in, e.g., Berger, Sambrook and Ausubel, *all supra.*

Nucleic Acids in Proximity to Markers/Isolated Chromosome Intervals

The present invention provides isolated nucleic acids comprising a QTL associated with *Sclerotinia* whole plant field resistance. The QTL is in proximity to a marker described herein and/or is localized within an interval defined by two markers of the present invention wherein each marker flanks the QTL. Such nucleic acids and/or intervals can be utilized to identify homologous nucleic acids and/or can be used in the production of transgenic plants having whole plant field resistance to *Sclerotinia* conferred by the introduced QTL. The nucleic acid and/or chromosome interval comprising a QTL is isolated, e.g., cloned via positional cloning methods outlined above. A chromosome interval can contain one or more ORFs associated with resistance, and can be cloned on one or more individual vectors, e.g., depending on the size of the chromosome interval.

It will be appreciated that numerous vectors are available in the art for the isolation and replication of the nucleic acids of the invention. For example, plasmids, cosmids and phage vectors are well known in the art, and are sufficient for many applications (e.g., in applications involving insertion of nucleic acids ranging from less than 1 to about 20 kilobases (kb). In certain applications, it is advantageous to make or clone large nucleic acids to identify nucleic acids more distantly linked to a given marker, or to isolate nucleic acids in excess of 10-20 kb, e.g., up to several hundred kilobases or more, such as the entire interval between two linked markers, i.e., up to and including one or more centiMorgans (cM), linked to QTLs as identified herein. In such cases, a number of vectors capable of accommodating large nucleic acids are available in the art, these include, yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), plant artificial chromosomes (PACs) and the like. For a general introduction to YACs, BACs, PACs and MACs as artificial chromosomes, *see*, e.g., Monaco and Larin (1994) Trends Biotechnol. 12:280. In addition, methods for the *in vitro* amplification of large nucleic acids linked to genetic markers are widely available (e.g., Cheng et al. (1994) Nature 369:684, and references therein). Cloning systems can be created or obtained from commercially; *see*, for example, Stratagene Cloning Systems, Catalogs 2000 (La Jolla, CA).

Generation of Transgenic Plants and Cells

The present invention also relates to host cells and organisms that are transformed with nucleic acids corresponding to QTLs and other genes identified according to the invention. For example, such nucleic acids include chromosome intervals, ORFs, and/or cDNAs or corresponding to a sequence or subsequence included within the identified chromosome interval or ORF. Additionally, the invention provides for the production of polypeptides corresponding to QTLs by recombinant techniques. Host cells are genetically engineered (i.e., transduced, transfected or transformed) with the vectors of this invention (i.e., vectors that comprise QTLs or other nucleic acids identified according to the methods of the invention and as described above) that include, for example, a cloning vector or an expression vector. Such vectors include, in addition to those described above, e.g., an agrobacterium, a virus (such as a plant virus), a naked polynucleotide, or a conjugated polynucleotide. The vectors are introduced into plant tissues, cultured plant cells or plant protoplasts by a variety of standard methods including electroporation (From et al. (1985) Proc. Natl. Acad. Sci. USA 82;5824), infection by viral vectors such as cauliflower mosaic virus (CaMV) (Hohn et al. (1982) Molecular Biology of Plant Tumors (Academic Press, New York, pp. 549-560); Howell U.S. Patent No. 4,407,956), high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface (Klein et al. (1987) Nature 327;70), use of pollen as vector (WO 85/01856), or use of *Agrobacterium tumefaciens* or *A. rhizogenes* carrying a T-DNA plasmid in which DNA fragments are cloned. The T-DNA plasmid is transmitted to plant cells upon infection by *Agrobacterium tumefaciens*, and a portion is stably integrated into the plant genome (Horsch et al. (1984) Science 233;496; Fraley et al. (1983) Proc. Natl. Acad. Sci. USA 80;4803). The method of introducing a nucleic acid of the present invention into a host cell is not critical to the instant invention. Thus, any method, e.g., including but not limited to the above examples, which provides for effective introduction of a nucleic acid into a cell or protoplast can be employed.

The engineered host cells can be cultured in conventional nutrient media modified as appropriate for such activities as, for example, activating promoters or selecting transformants. These cells can optionally be cultured into transgenic plants. Plant regeneration from cultured protoplasts is described in Evans et al. ((1983) "Protoplast Isolation and Culture," Handbook of Plant Cell Cultures 1, 124-176 (MacMillan Publishing

Co., New York); Davey ((1983) "Recent Developments in the Culture and Regeneration of Plant Protoplasts," Protoplasts, pp. 12-29, (Birkhauser, Basel)); Dale ((1983) "Protoplast Culture and Plant Regeneration of Cereals and Other Recalcitrant Crops," Protoplasts, pp. 31-41, (Birkhauser, Basel)); and Binding ((1985) "Regeneration of Plants," Plant Protoplasts, pp. 21-73, (CRC Press, Boca Raton)).

The present invention also relates to the production of transgenic organisms, which may be bacteria, yeast, fungi, or plants, transduced with the nucleic acids, e.g., cloned QTLs of the invention. A thorough discussion of techniques relevant to bacteria, unicellular eukaryotes and cell culture may be found in references enumerated above and are briefly outlined as follows. Several well-known methods of introducing target nucleic acids into bacterial cells are available, any of which may be used in the present invention. These include: fusion of the recipient cells with bacterial protoplasts containing the DNA, treatment of the cells with liposomes containing the DNA, electroporation, projectile bombardment (biolistics), carbon fiber delivery, and infection with viral vectors (discussed further, below), etc. Bacterial cells can be used to amplify the number of plasmids containing DNA constructs of this invention. The bacteria are grown to log phase and the plasmids within the bacteria can be isolated by a variety of methods known in the art (see, for instance, Sambrook). In addition, a plethora of kits are commercially available for the purification of plasmids from bacteria. For their proper use, follow the manufacturer's instructions (see, for example, EasyPrep™, FlexiPrep™, both from Pharmacia Biotech; StrataClean™, from Stratagene; and, QIAprep™ from Qiagen). The isolated and purified plasmids are then further manipulated to produce other plasmids, used to transfect plant cells or incorporated into *Agrobacterium tumefaciens* related vectors to infect plants. Typical vectors contain transcription and translation terminators, transcription and translation initiation sequences, and promoters useful for regulation of the expression of the particular target nucleic acid. The vectors optionally comprise generic expression cassettes containing at least one independent terminator sequence, sequences permitting replication of the cassette in eukaryotes, or prokaryotes, or both, (e.g., shuttle vectors) and selection markers for both prokaryotic and eukaryotic systems. Vectors are suitable for replication and integration in prokaryotes, eukaryotes, or preferably both. See, Giliman & Smith ((1979) Gene 8:81); Roberts et al. ((1987) Nature 328:731); (Schneider et al. (1995) Protein Expr. Purif. 6435:10); Ausubel, Sambrook, Berger (all *supra*). A catalogue of Bacteria and

Bacteriophages useful for cloning is provided, e.g., by the ATCC, e.g., The ATCC Catalogue of Bacteria and Bacteriophage (1992) Gherna et al. (eds) published by the ATCC. Additional basic procedures for sequencing, cloning and other aspects of molecular biology and underlying theoretical considerations are also found in Watson et al. (1992) Recombinant DNA, Second Edition, Scientific American Books, NY.

Transforming Nucleic Acids into Plants

Embodiments of the present invention pertain to the production of transgenic plants comprising the cloned nucleic acids, e.g., chromosome intervals, isolated ORFs, and cDNAs associated with QTLs, of the invention. Techniques for transforming plant cells with nucleic acids are generally available and can be adapted to the invention by the use of nucleic acids encoding or corresponding to QTLs, QTL homologs, isolated chromosome intervals, and the like. In addition to Berger, Ausubel and Sambrook, useful general references for plant cell cloning, culture and regeneration include Jones (ed.) ((1995) Plant Gene Transfer and Expression Protocols-- Methods in Molecular Biology, Volume 49 Humana Press Towata NJ); Payne et al. ((1992) Plant Cell and Tissue Culture in Liquid Systems John Wiley & Sons, Inc. New York, NY (Payne)); and Gamborg and Phillips (eds) ((1995) Plant Cell, Tissue and Organ Culture; Fundamental Methods Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York) (Gamborg)). A variety of cell culture media are described in Atlas and Parks (eds.) (The Handbook of Microbiological Media (1993) CRC Press, Boca Raton, FL (Atlas)). Additional information for plant cell culture is found in available commercial literature such as the Life Science Research Cell Culture Catalogue (1998) from Sigma- Aldrich, Inc. (St Louis, MO) (Sigma-LSRCCC) and, e.g., the Plant Culture Catalogue and supplement (1997) also from Sigma-Aldrich, Inc. (St Louis, MO) (Sigma-PCCS). Additional details regarding plant cell culture are found in Croy, (ed.) ((1993) Plant Molecular Biology Bios Scientific Publishers, Oxford, U.K.)

The nucleic acid constructs of the invention, e.g., plasmids, cosmids, artificial chromosomes, DNA and RNA polynucleotides, are introduced into plant cells, either in culture or in the organs of a plant by a variety of conventional techniques. Where the sequence is expressed, the sequence is optionally combined with transcriptional and translational initiation regulatory sequences that direct the transcription or translation of the sequence from the exogenous DNA in the intended tissues of the transformed plant.

Isolated nucleic acids of the present invention can be introduced into plants according to any of a variety of techniques known in the art. Techniques for transforming a wide variety of higher plant species are well known and described in the technical, scientific, and patent literature. See, for example, Weising et al. (1988) Ann. Rev. Genet. 22:421-477.

The DNA constructs of the invention, for example, plasmids, cosmids, phage, naked or variously conjugated-DNA polynucleotides, (e.g., polylysine-conjugated DNA, peptide-conjugated DNA, liposome-conjugated DNA, etc.), or artificial chromosomes, can be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation and microinjection of plant cell protoplasts, or the DNA constructs can be introduced directly to plant cells using ballistic methods, such as DNA particle bombardment.

Microinjection techniques for injecting e.g., cells, embryos, callus and protoplasts, are known in the art and well described in the scientific and patent literature. For example, a number of methods are described in Jones (ed.) ((1995) Plant Gene Transfer and Expression Protocols-- Methods in Molecular Biology, Volume 49 Humana Press Towata NJ), as well as in the other references noted herein and available in the literature.

For example, the introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski, et al. (EMBO J. 3:2717 (1984)). Electroporation techniques are described in Fromm, et al. (Proc. Nat'l. Acad. Sci. USA 82:5824 (1985)). Ballistic transformation techniques are described in Klein, et al. (Nature 327:70-73 (1987)). Additional details are found in Jones (1995) and Gamborg and Phillips (1995), *supra*, and in US Patent No. 5,990,387.

Alternatively, *Agrobacterium*-mediated transformation is employed to generate transgenic plants. *Agrobacterium*-mediated transformation techniques, including disarming and use of binary vectors, are also well described in the scientific literature. See, for example, Horsch, et al. (1984) Science 233:496; and Fraley et al. (1984) Proc. Nat'l. Acad. Sci. USA 80:4803 and reviewed in Hansen and Chilton (1998) Current Topics in Microbiology 240:22 and Das (1998) Subcellular Biochemistry 29: Plant Microbe Interactions pp. 343-363.

The DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and

adjacent marker into the plant cell DNA when the cell is infected by the bacteria. See, U.S. Patent No. 5,591,616. Although *Agrobacterium* is useful primarily in dicots, certain monocots can be transformed by *Agrobacterium*. For instance, *Agrobacterium* transformation of maize is described in U.S. Patent No. 5,550,318.

Other methods of transfection or transformation include (1) *Agrobacterium rhizogenes*-mediated transformation (see, e.g., Lichtenstein and Fuller (1987) In: Genetic Engineering, vol. 6, PWJ Rigby, Ed., London, Academic Press; and Lichtenstein; C. P., and Draper (1985) In: DNA Cloning, Vol. II, D. M. Glover, Ed., Oxford, IRI Press); WO 88/02405, published April 7, 1988, describes the use of *A. rhizogenes* strain A4 and its Ri plasmid along with *A. tumefaciens* vectors pARC8 or pARC16 (2) liposome-mediated DNA uptake (see, e.g., Freeman et al. (1984) Plant Cell Physiol. 25:1353), (3) the vortexing method (see, e.g., Kindle (1990) Proc. Natl. Acad. Sci., (USA) 87:1228).

DNA can also be introduced into plants by direct DNA transfer into pollen as described by Zhou et al. ((1983) Methods in Enzymology, 101:433); Hess ((1987) Intern Rev. Cytol. 107:367); and Luo et al. ((1988) Plant Mol. Biol. Reporter 6:165). Expression of polypeptide coding genes can be obtained by injection of the DNA into reproductive organs of a plant as described by Pena et al. ((1987) Nature 325:274). DNA can also be injected directly into the cells of immature embryos and the desiccated embryos rehydrated as described by Neuhaus et al. ((1987) Theor. Appl. Genet. 75:30); and Benbrook et al. ((1986) in Proceedings Bio Expo Butterworth, Stoneham, Mass., pp. 27-54). A variety of plant viruses that can be employed as vectors are known in the art and include cauliflower mosaic virus (CaMV), geminivirus, brome mosaic virus, and tobacco mosaic virus.

Regeneration of Transgenic Plants

Transformed plant cells that are derived by any of the above transformation techniques can be cultured to regenerate a whole plant that possesses the transformed genotype and thus the desired phenotype. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker that has been introduced together with the desired nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans et al. ((1983) Protoplasts Isolation and Culture, Handbook of Plant Cell Culture pp. 124-176, Macmillian Publishing Company, New York); and Binding ((1985) Regeneration of Plants, Plant Protoplasts pp. 21-73, CRC Press, Boca Raton). Regeneration can also be obtained

from plant callus, explants, somatic embryos (Dandekar et al. (1989) J. Tissue Cult. Meth. 12:145; McGranahan, et al. (1990) Plant Cell Rep. 8:512) organs, or parts thereof. Such regeneration techniques are described generally in Klee et al. ((1987)., Ann. Rev. of Plant Phys. 38:467-486). Additional details are found in Payne (1992) and Jones (1995), both *supra*, and Weissbach and Weissbach, eds. ((1988) Methods for Plant Molecular Biology Academic Press, Inc., San Diego, CA). This regeneration and growth process includes the steps of selection of transformant cells and shoots, rooting the transformant shoots and growth of the plantlets in soil. These methods are adapted to the invention to produce transgenic plants bearing QTLs and other genes isolated according to the methods of the invention.

In addition, the regeneration of plants containing the polynucleotide of the present invention and introduced by *Agrobacterium* into cells of leaf explants can be achieved as described by Horsch et al. ((1985) Science 227:1229-1231). In this procedure, transformants are grown in the presence of a selection agent and in a medium that induces the regeneration of shoots in the plant species being transformed as described by Fraley et al. ((1983) Proc. Natl. Acad. Sci. (U.S.A.) 80:4803). This procedure typically produces shoots within two to four weeks and these transformant shoots are then transferred to an appropriate root-inducing medium containing the selective agent and an antibiotic to prevent bacterial growth. Transgenic plants of the present invention may be fertile or sterile.

Plants for the transformation and expression of whole plant filed resistance to *Sclerotinia* associated QTLs and other nucleic acids identified and cloned according to the present invention include, but are not limited to, agronomically and horticulturally important species. Such species include primarily dicots, e.g., of the families: Brassicaceae, Leguminosae (including pea, beans, lentil, peanut, yam bean, cowpeas, velvet beans, soybean, clover, alfalfa, lupine, vetch, lotus, sweet clover, wisteria, and sweetpea); and, Compositae (the largest family of vascular plants, including at least 1,000 genera, including important commercial crops such as sunflower).

Additionally, targets for modification with the nucleic acids of the invention, as well as those specified above, plants from the genera: *Allium*, *Apium*, *Arachis*, *Brassica*, *Capsicum*, *Cicer*, *Cucumis*, *Curcubita*, *Daucus*, *Fagopyrum*, *Glycine*, *Helianthus*, *Lactuca*, *Lens*, *Lycopersicon*, *Medicago*, *Pisum*, *Phaseolus*, *Solanum*, *Trifolium*, *Vigna*, and many others.

Common crop plants that are targets of the present invention include soybean, sunflower, canola, peas, beans, lentils, peanuts, yam beans, cowpeas, velvet beans, clover, alfalfa, lupine, vetch, sweet clover, sweetpea, field pea, fava bean, broccoli, brussel sprouts, cabbage, cauliflower, kale, kohlrabi, celery, lettuce, carrot, onion, pepper, potato, eggplant, and tomato.

In construction of recombinant expression cassettes of the invention, which include, for example, helper plasmids comprising virulence functions, and plasmids or viruses comprising exogenous DNA sequences such as structural genes, a plant promoter fragment is optionally employed to direct expression of a nucleic acid in any or all tissues of a regenerated plant. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1'- or 2'- promoter derived from T-DNA of *Agrobacterium tumefaciens*, and other transcription initiation regions from various plant genes known to those of skill. Alternatively, the plant promoter may direct expression of the polynucleotide of the invention in a specific tissue (tissue-specific promoters) or may be otherwise under more precise environmental control (inducible promoters). Examples of tissue-specific promoters under developmental control include promoters that initiate transcription only in certain tissues, such as fruit, seeds, or flowers.

Any of a number of promoters that direct transcription in plant cells can be suitable. The promoter can be either constitutive or inducible. In addition to the promoters noted above, promoters of bacterial origin that operate in plants include the octopine synthase promoter, the nopaline synthase promoter and other promoters derived from native Ti plasmids. See, Herrara-Estrella et al. ((1983), Nature, 303:209). Viral promoters include the 35S and 19S RNA promoters of cauliflower mosaic virus. See, Odell et al. ((1985) Nature, 313:810). Other plant promoters include the ribulose-1,3-bisphosphate carboxylase small subunit promoter and the phaseolin promoter. The promoter sequence from the E8 gene and other genes may also be used. The isolation and sequence of the E8 promoter is described in detail in Deikman and Fischer ((1988) EMBO J. 7:3315). Many other promoters are in current use and can be coupled to an exogenous DNA sequence to direct expression of the nucleic acid.

If expression of a polypeptide, including those encoded by QTLs or other nucleic acids correlating with phenotypic traits of the present invention, is desired, a polyadenylation region at the 3'-end of the coding region is typically included. The polyadenylation region

can be derived from the natural gene, from a variety of other plant genes, or from, e.g., T-DNA.

The vector comprising the sequences (e.g., promoters or coding regions) from genes encoding expression products and transgenes of the invention will typically include a nucleic acid subsequence, a marker gene that confers a selectable, or alternatively, a screenable, phenotype on plant cells. For example, the marker may encode biocide resistance, particularly antibiotic resistance, such as resistance to kanamycin, G418, bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosluforon, or phosphinothricin (the active ingredient in the herbicides bialaphos or Basta). See, e.g., Padgette et al. (1996) In: Herbicide-Resistant Crops (Duke, ed.), pp 53-84, CRC Lewis Publishers, Boca Raton ("Padgette, 1996"). For example, crop selectivity to specific herbicides can be conferred by engineering genes into crops that encode appropriate herbicide metabolizing enzymes from other organisms, such as microbes. See, Vasil (1996) In: Herbicide-Resistant Crops (Duke, ed.), pp 85-91, CRC Lewis Publishers, Boca Raton) ("Vasil", 1996).

One of skill will recognize that after the recombinant expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed. In vegetatively propagated crops, mature transgenic plants can be propagated by the taking of cuttings or by tissue culture techniques to produce multiple identical plants. Selection of desirable transgenics is made and new varieties are obtained and propagated vegetatively for commercial use. In seed propagated crops, mature transgenic plants can be self crossed to produce a homozygous inbred plant. The inbred plant produces seed containing the newly introduced heterologous nucleic acid. These seeds can be grown to produce plants that would produce the selected phenotype. Parts obtained from the regenerated plant, such as flowers, seeds, leaves, branches, fruit, and the like are included in the invention, provided that these parts comprise cells comprising the isolated nucleic acid of the present invention. Progeny and variants, and mutants of the regenerated plants are also included within the scope of the invention, provided that these parts comprise the introduced nucleic acid sequences.

Transgenic plants expressing a polynucleotide of the present invention can be screened for transmission of the nucleic acid of the present invention by, for example,

standard immunoblot and DNA detection techniques. Expression at the RNA level can be determined initially to identify and quantitate expression-positive plants. Standard techniques for RNA analysis can be employed and include PCR amplification assays using oligonucleotide primers designed to amplify only the heterologous RNA templates and solution hybridization assays using heterologous nucleic acid-specific probes. The RNA-positive plants can then be analyzed for protein expression by Western immunoblot analysis using the specifically reactive antibodies of the present invention. In addition, *in situ* hybridization and immunocytochemistry according to standard protocols can be done using heterologous nucleic acid specific polynucleotide probes and antibodies, respectively, to localize sites of expression within transgenic tissue. Generally, a number of transgenic lines are usually screened for the incorporated nucleic acid to identify and select plants with the most appropriate expression profiles.

One embodiment is a transgenic plant that is homozygous for the added heterologous nucleic acid; i.e., a transgenic plant that contains two added nucleic acid sequences, one gene at the same locus on each chromosome of a chromosome pair. A homozygous transgenic plant can be obtained by sexually mating (selfing) a heterozygous transgenic plant that contains a single added heterologous nucleic acid, germinating some of the seed produced and analyzing the resulting plants produced for altered expression of a polynucleotide of the present invention relative to a control plant (i.e., native, non-transgenic). Back-crossing to a parental plant and out-crossing with a non- transgenic plant are also contemplated.

High Throughput Screening

In one aspect of the invention, the determination of genetic marker alleles is performed by high throughput screening. High throughput screening involves providing a library of genetic markers, e.g., RFLPs, AFLPs, isozymes, specific alleles and variable sequences, including SSR. Such libraries are then screened against plant genomes to generate a "fingerprint" for each plant under consideration. In some cases a partial fingerprint comprising a sub-portion of the markers is generated in an area of interest. Once the genetic marker alleles of a plant have been identified, the correspondence between one or several of the marker alleles and a desired phenotypic trait is determined through statistical associations based on the methods of this invention.

High throughput screening can be performed in many different formats. Hybridization can take place in a 96-, 324-, or a 1524-well format or in a matrix on a silicon chip or other format.

In one commonly used format, a dot blot apparatus is used to deposit samples of fragmented and denatured genomic DNA on a nylon or nitrocellulose membrane. After cross-linking the nucleic acid to the membrane, either through exposure to ultra-violet light or by heat, the membrane is incubated with a labeled hybridization probe. The labels are incorporated into the nucleic acid probes by any of a number of means well-known in the art. The membranes are washed to remove non-hybridized probes and the association of the label with the target nucleic acid sequence is determined.

A number of well-known robotic systems have been developed for high throughput screening, particularly in a 96 well format. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, MA.; ORCA™, Beckman Coulter, Fullerton CA). Any of the above devices are suitable for use with the present invention. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein will be apparent to persons skilled in the relevant art.

In addition, high throughput screening systems themselves are commercially available (*see*, e.g., Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, etc.). These systems typically automate entire procedures including all sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate or membrane in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. The manufacturers of such systems provide detailed protocols for the use of their products in high throughput applications.

In one variation of the invention, solid phase arrays are adapted for the rapid and specific detection of multiple polymorphic nucleotides. Typically, a nucleic acid probe is linked to a solid support and a target nucleic acid is hybridized to the probe. Either the probe, or the target, or both, can be labeled, typically with a fluorophore. If the target is labeled, hybridization is evaluated by detecting bound fluorescence. If the probe is labeled,

hybridization is typically detected by quenching of the label by the bound nucleic acid. If both the probe and the target are labeled, detection of hybridization is typically performed by monitoring a color shift resulting from proximity of the two bound labels.

In one embodiment, an array of probes is synthesized on a solid support. Using chip masking technologies and photoprotective chemistry, it is possible to generate ordered arrays of nucleic acid probes. These arrays, which are known, e.g., as "DNA chips" or as very large scale immobilized polymer arrays (VLSIPS™ arrays) can include millions of defined probe regions on a substrate having an area of about 1 cm^2 to several cm^2.

In another embodiment, capillary electrophoresis is used to analyze a polymorphism. This technique works best when the polymorphism is based on size, for example, AFLP and SSR. This technique is described in detail in U.S. Patent Nos. 5,534,123 and 5,728,282. Briefly, capillary electrophoresis tubes are filled with the separation matrix. The separation matrix contains hydroxyethyl cellulose, urea and optionally formamide. The AFLP or SSR samples are loaded onto the capillary tube and electrophoresed. Because of the small amount of sample and separation matrix required by capillary electrophoresis, the run times are very short. The molecular sizes and therefore, the number of nucleotides present in the nucleic acid sample are determined by techniques described herein. In a high throughput format, many capillary tubes are placed in a capillary electrophoresis apparatus. The samples are loaded onto the tubes and electrophoresis of the samples is run simultaneously. See, Mathies and Huang (1992) Nature 359:167.

Integrated Systems

Because of the great number of possible combinations present in one array, in one aspect of the invention, an integrated system such as a computer, software corresponding to the statistical models of the invention, and data sets corresponding to genetic markers and phenotypic values, facilitates mapping of phenotypic traits, including QTLs. The phrase "integrated system" in the context of this invention refers to a system in which data entering a computer corresponds to physical objects or processes external to the computer, e.g., nucleic acid sequence hybridization, and a process that, within a computer, causes a physical transformation of the input signals to different output signals. In other words, the input data, e.g., hybridization on a specific region of an array is transformed to output data, e.g., the identification of the sequence hybridized. The process within the computer is a set of instructions, or "program," by which positive hybridization signals are recognized by the

integrated system and attributed to individual samples as a genotype. Additional programs correlate the genotype, and more particularly in the methods of the invention, the haplotype, of individual samples with phenotypic values, e.g., using the HAPLO-IM$^+$, HAPLO-MQM, and/or HAPLO-MQM$^+$ models of the invention. For example, the programs JoinMap® and MapQTL® are particularly suited to this type of analysis and can be extended to include the HAPLO-IM$^+$, HAPLO-MQM, and/or HAPLO-MQM$^+$ models of the invention. In addition there are numerous e.g., C/C++ programs for computing, Delphi and/or Java programs for GUI interfaces, and Active X applications (e.g., Olectra Chart and True WevChart) for charting tools. Other useful software tools in the context of the integrated systems of the invention include statistical packages such as SAS, Genstat, and S-Plus. Furthermore additional programming languages such as Fortran and the like are also suitably employed in the integrated systems of the invention.

In one aspect, the invention provides an integrated system comprising a computer or computer readable medium comprising a database with at least one data set that corresponds to genotypes for genetic markers. The system also includes a user interface allowing a user to selectively view one or more databases. In addition, standard text manipulation software such as word processing software (e.g., Microsoft Word™ or Corel Wordperfect™) and database or spreadsheet software (e.g., spreadsheet software such as Microsoft Excel™, Corel Quattro Pro™, or database programs such as Microsoft Access™ or Paradox™) can be used in conjunction with a user interface (e.g., a GUI in a standard operating system such as a Windows, Macintosh or Linux system) to manipulate strings of characters.

The invention also provides integrated systems for sample manipulation incorporating robotic devices as previously described. A robotic liquid control armature for transferring solutions (e.g., plant cell extracts) from a source to a destination, e.g., from a microtiter plate to an array substrate, is optionally operably linked to the digital computer (or to an additional computer in the integrated system). An input device for entering data to the digital computer to control high throughput liquid transfer by the robotic liquid control armature and, optionally, to control transfer by the armature to the solid support is commonly a feature of the integrated system.

Integrated systems for genetic marker analysis of the present invention typically include a digital computer with one or more of high-throughput liquid control software, image analysis software, data interpretation software, a robotic liquid control armature for

transferring solutions from a source to a destination operably linked to the digital computer, an input device (e.g., a computer keyboard) for entering data to the digital computer to control high throughput liquid transfer by the robotic liquid control armature and, optionally, an image scanner for digitizing label signals from labeled probes hybridized, e.g., to expression products on a solid support operably linked to the digital computer. The image scanner interfaces with the image analysis software to provide a measurement of, e.g., differentiating nucleic acid probe label intensity upon hybridization to an arrayed sample nucleic acid population, where the probe label intensity measurement is interpreted by the data interpretation software to show whether, and to what degree, the labeled probe hybridizes to a label. The data so derived is then correlated with phenotypic values using the statistical models of the present invention, to determine the correspondence between phenotype and genotype(s) for genetic markers, thereby, assigning chromosomal locations.

Optical images, e.g., hybridization patterns viewed (and, optionally, recorded) by a camera or other recording device (e.g., a photodiode and data storage device) are optionally further processed in any of the embodiments herein, e.g., by digitizing the image and/or storing and analyzing the image on a computer. A variety of commercially available peripheral equipment and software is available for digitizing, storing and analyzing a digitized video or optical image, e.g., using PC (Intel x86 or pentium chip-compatible DOS™, OS2™ WINDOWS™, WINDOWS NT™ or WINDOWS95™ based machines), MACINTOSH™, LINUX, or UNIX based (e.g., SUN™ work station) computers.

Kits

Kits are also provided to facilitate the screening of germplasm for the markers of the present invention. The kits comprise the polynucleotides of the present invention, fragments or complements thereof, for use as probes or primers to detect the markers for *Sclerotinia* whole plant field resistance or improved whole plant field resistance to *Sclerotinia*. Instructions for using the polynucleotides, as well as buffers and/or other solutions may also be provided to facilitate the use of the polynucleotides. The kit is useful for high throughput screening and in particular, high throughput screening with integrated systems.

Examples

The following experimental methods and results provide additional details regarding specific aspects of protocols and procedures relevant to the practice of the present invention.

The examples, which are provided without limitation to illustrate the claimed invention, involve the application of protocols well known to those of skill in the art, and detailed in the references cited herein. Examples 1-6 describe analyses of one *Brassica napus* mapping population. Examples 7-10 describe analyses of a different *Brassica napus* mapping population.

Example 1: Mapping Population #1

The parents used for the mapping population were 04DHS11418 (a *Sclerotinia* resistant double haploid deposited as ATCC Accession No. PTA-6778) and PHI2004HS1 (a non-resistant spring canola DH line polymorphic to the resistant line, selected for having a similar agronomic phenotype with consistent high susceptibility in the same testing environment where the 04DHS11418 line was selected for resistance) (see Table 2). These lines were used to develop a double haploid mapping population consisting of 186 progeny.

Table 2. Measuring field performance under extreme disease pressure (research trials)

Rating SSDIS**	Category	Disease incidence SSDI%*	Spring Checks	Mapping Parents
1.0	Highly susceptible	≥80	44A89	
1.1 – 2.0	Susceptible	79 – 70	46A65=2	PHI2004HS1
2.1 – 3.0	Moderately susceptible	69 – 60	46A76=3	
3.1 – 4.0		59 – 50		
4.1 – 5.0	Moderately resistant	49 – 40		
5.1 – 6.0		39 – 30		
6.1 – 7.0	Resistant	29 – 20		04DHS11418
7.1 – 8.0		19 – 10		
8.1 – 9.0	Highly resistant	9 - 0		

* SSDI% *Sclerotinia sclerotiorum* Disease Incidence %. SSDI% is UNSSDI (where UNSSDI is the percentage of plants in a population infected with *Sclerotinia)* rating adjusted for a deviation from the expected mean of checks 04DHS11418 (25%) and PHI2004HS1 (75%) as described on Table 4. This rating is used only under controlled extreme disease pressure field research conditions. It is calculated by multiplying the observed SSDI% by Factor X, where Factor X is the factor that brings the average SSDI% of the appropriate checks to 50%. Adjustment for severity is not done.

**SSDIS *Sclerotinia sclerotiorum*. SSDIS is the UNSSDI rating adjusted for a deviation from the expected mean of checks 46A65/46A76 for spring canola under extreme disease pressure This rating is used only under controlled extreme disease pressure field research conditions. It is calculated by multiplying the observed UNSSDI by Factor X, where Factor X is the factor that brings the average SSDI% of the appropriate checks to 50%. Adjustment for severity is done after incidence adjustment.

UNSSDS is a rating of the extent of disease development on an affected plant. Two scales are used in the invention. The Pioneer SSDS scale ranges from 1 (dead) to 9 (no disease) and the Public scale ranges from 0 (no disease) to 5 (dead) plant. For details of the Pioneer SSDS scale, see Table 15 of WO 2006/135717. The Public scale is provided as follows: 0 = no disease; 1 = superficial lesions or small branch affected; 2 = large branch dead; 3 = main stem at least 50% girdled; 4 = main stem girdled but plant produced good seed; 5 = main stem girdled, much reduced yield.

Both parents have good standability. Therefore, standability or lodging resistance is fixed, thus eliminating this variable in the mapping process. The choice of a highly susceptible line resulted in a population without transgressive segregation (i.e., all resistance came from 04DHS11418 and no DH progeny lines were more resistant than the resistant parent). Over a period of four years, the population was phenotyped in the field and genotyped with SSR molecular markers. Phenotyping was carried out as described in WO 2006/135717..

Example 2: RNA Expression Profiling

RNA expression profiling experiments were carried using 04DHS11418, PHI2004HS1, resistant and susceptible bulks (comprised of susceptible and resistant double haploid (DH) progenies). Intact leaf tissues (not inoculated) were used, as well as leaf tissues sampled 6, 24 and 48 hours after leaf inoculation with mycelium. RNA expression profiling was performed by probing the different bulks with a chip consisting of *Brassica* (85,820) and *Arabidopsis* stress-related (17,617) nucleic acid sequences. A number of genes related to *Sclerotinia* disease resistance or to pectin (these genes are associated with cell wall integrity) were upregulated when inoculated with mycelium at 6, 24 and 48 hours. A summary of the results can be found in Table 3. Some of these genes were sequenced in the resistant and susceptible lines to find any SNPs. These SNPs were then added to the genetic map (termed KK).

Table 3. Number of disease related genes upregulated at different time treatments.

Treatment	Pectin-Related Gene Sorting					Disease Resistance-Related Gene Sorting				
	Total	1 Set	2 Set	3 Set	4 Set	Total	1 Set	2 Set	3 Set	4 Set
6 Hour Treatment	25	13	4	7	1	19	11	4	3	1
24 Hour Treatment	65	21	38	6	0	48	38	9	0	1
48 Hour Treatment	60	53	5	2	0	52	43	6	3	0
Total	**150**	87	47	15	1	**119**	92	19	6	2

Example 3: *Sclerotinia* Screening

Disease Scoring

The plants of the double haploid mapping population created as described in Example 1, were rated for disease as described in Table 4. The unadjusted parameters (e.g.,

UNSSDI and UNSSDS) showed year to year variatation due to environmental variation such as positional variation in the field and weather conditions. Such variation would be expected by one skilled in the art.

Table 4. Field-collected *Sclerotinia* parameters UNSSDI and UNSSDS and their relationship to derived parameters SSDI%, SSDIS (research data) and SSFS (natural/research data).

Trait	**UNSSDI Disease Incidence**	**UNSSDS Disease severity of affected plants**		**SSDI% Based on adjusted UNSSDI** under extreme disease pressure field research conditions	**SSDIS Based on adjusted UNSSDI and UNSSDS** under extreme disease pressure field research conditions	**SSFS Field severity based on both UNSSDI and UNSSDS used in natural field conditions or research**
Scale	0-100%	Pioneer SSDS scale 1=dead 9= no disease Public scale 0=no disease 5=dead plant		0-100% Conversion of UNSSDI and adjustment for checks' UNSSDI	1-9 Conversion of UNSSDI and UNSSDS adjustment for checks' UNSSDI and UNSSDS	0-100% % field impact – quantifies damage in the field irrespective of disease pressure
Usage	General	General		Pioneer only	Pioneer only	General
Adjustments	N/A	N/A		Adjusted to checks	Adjusted to checks	Unadjusted
Hypothetical Examples (HE): Different combinations of disease incidence and disease severity		Pioneer SSDS scale	Public scale	*assuming checks do not deviate for UNSSDI	**assuming checks to do not deviate for UNSSDI and UNSSDS	
HE 1	80	1	5.0	*80	**1.0 (80)	80
HE 2	80	5	2.0	80	2.6 (64)	32
HE 3	50	5	2.0	50	5.0 (40)	20
HE 4	30	7	1.0	30	7.3 (17)	6
HE 5	10	8	1.0	10	8.5 (5)	2

UNSSDI is the percentage of plants in a population infected with *Sclerotinia*.

UNSSDS is a rating of the extent of disease development on an affected plant. Two scales are used in the invention. The Pioneer SSDS scale ranges from 1 (dead) to 9 (no disease) and the Public scale ranges from 0 (no disease) to 5 (dead) plant. For details of the Pioneer SSDS scale, see Table 15 of WO 2006/135717. The Public scale is provided as follows: 0 = no disease; 1 = superficial lesions or small branch affected; 2 = large branch dead; 3 = main stem at least 50% girdled; 4 = main stem girdled but plant produced good seed; 5 = main stem girdled, much reduced yield.

SSDI% is UNSSDI rating adjusted for a deviation from the expected mean of checks 04DHS11418 (25%) and PHI2004HS1 (75%) as described on Table 4. This rating is used only under controlled extreme disease pressure field research conditions. It is calculated by multiplying the observed SSDI% by Factor X, where Factor X is the factor that brings the average SSDI% of the appropriate checks to 50%. Adjustment for severity is not done.

SSDIS is UNSSDI rating adjusted for a deviation from the expected mean of checks 04DHS11418 (25%) and PHI2004HS1 (75%) as described on Table 4. This rating is used only under controlled extreme disease pressure field research conditions. It is calculated by multiplying the observed UNSSDI by Factor X,

where Factor X is the factor that brings the average SSDI% of the appropriate checks to 50%. Adjustment for severity is done after incidence adjustment.

SSFS is a measure of both disease incidence and severity under natural disease pressure in the field. It is calculated as follows: SSFS = [SSDI% x SSDS(0-5 scale)] ÷ 5

Example 4: Genetic Mapping and QTL Analysis

Genetic mapping and QTL analysis were performed using JoinMap v3.0 (Van Ooijen, J.W. and R.E. Voorrips, 2001 JoinMap® 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands). The Kosambi centiMorgan function was used. A QTL was declared if its LOD score exceeded the threshold of 2.0.

Genetic Mapping

Genetic mapping has placed 351 molecular markers to 19 linkage groups (Lg) that correspond to 19 canola chromosomes and public linkage group nomenclature. The linkage map covers ~1400 cM.

QTL Analysis

QTL analysis using simple interval mapping and composite interval mapping (CIM) (Zeng (1994), Genetics 136:1457) identified 7 linkage groups (N1, N7, N9, N11, N12, N18 and N19) contributing to whole plant field resistance to *Sclerotinia*. In addition, regions identified by interval mapping as being associated with *Sclerotinia* resistance were confirmed by single-factor analysis of variance (PROC GLM, SAS Enterprise Guide 4.2) on *Sclerotinia* parameters at the $P \leq 0.01$ significance level. These QTLs are identified in Tables 5 and 6 below. As shown by the "Phenotypic Variation Explained" values in Table 6, some QTLs had a larger effect on *Sclerotinia* resistance than others.

Genetic mapping and QTL analysis were performed using JoinMap v3.0 (Van Ooijen, J.W. and R.E. Voorrips, 2001 JoinMap® 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands). The Kosambi centiMorgan function was used. A QTL was declared if its LOD score exceeded the threshold of 2.0. LOD stands for logarithm of the odds (to the base 10).

Table 5: Markers significantly associated with *Sclerotinia* resistance at P ≤ 0.01.

Linkage Group	Marker	Map Position (cM)	Parameter	Year
N1	AG0093	48.1	UNSSDS	2008
	AG0243	52.2	UNSSDS, SSFS, SSDIS	2008
	BG1453	52.8	UNSSDS, SSDIS	2008
	UB0163	54.3	SSFS, SSDIS	2008
	AG0391	55.4	UNSSDS, SSFS, SSDIS, SSDI%	2008
	AG0045	55.7	SSFS, SSDIS	2008
	AG0304	56.4	SSFS, SSDIS	2008
	PE0203	57.2	SSDIS	2008
	BG0119	58.6	SSFS, SSDIS	2008
	BG0988	60.3	SSDI%	2005
			UNSSDS, UNSSDI, SSFS, SSDI%	2008
	AG0482	66.1	UNSSDS, SSFS, SSDIS, SSDI%	2005
			SSFS, SSDIS	2006
			UNSSDS, UNSSDI, SSFS, SSDIS, ***SSDI%***	2008
			UNSSDS, SSFS, SSDIS	Across years
N7	AG0510	0.0	UNSSDS, UNSSDI, SSFS, SSDIS, SSDI%	2007
			UNSSDS, UNSSDI, SSFS, SSDIS, SSDI%	Across years
	CA0105	0.5	UNSSDS	2006
			UNSSDS, UNSSDI, SSFS, SSDIS, SSDI%	2007
			UNSSDS, UNSSDI, SSFS, SSDIS, SSDI%	Across years
N9	AG0378	0.0	UNSSDI, SSFS	2006
			UNSSDI	Across years
	KK66	4.0	UNSSDI	2006
			UNSSDI	Across years
N11	CA0120	21.9	UNSSDS	Across years
	CA0233	28.8	UNSSDS	2006, Across years
			SSFS	2007
	CA0226	30.7	UNSSDS	2006, Across years
	CA0546	30.7	UNSSDS	2006
			SSFS	2007
			UNSSDS, SSDIS, SSFS	Across years
	BG1149	30.8	UNSSDS	2006, 2007, Across years
			SSFS	2007, Across years
	CA1080	30.8	UNSSDS, SSFS	2006
			SSFS, SSDIS	2007
			UNSSDS, SSFS, SSDIS	Across years
	BG1230	31.3	UNSSDS	2006
			UNSSDS, SSFS	Across years
	AG0370	31.3	UNSSDS	2006
			UNSSDS, SSFS	Across years

	BG0713	31.7	UNSSDS	2006, Across years
			SSFS	2007, Across years
	BG0869	32.9	UNSSDS	2006, Across years
	BG1513	33.3	UNSSDS	2006, Across years
	BG0181	33.4	UNSSDS	2006, Across years
	CA1097	36.2	UNSSDS, SSFS	2006
			UNSSDI, SSFS	2007
			UNSSDI, SSFS	2008
			UNSSDS, SSFS, SSDIS	Across years
N12	CA0753	60.5	UNSSDS	2005, Across years
	CA1027	78.9	UNSSDS, SSFS	2005
			UNSSDS, SSFS, SSDIS, SSDI%	2006
			UNSSDS	Across years
	PE0063	79.3	UNSSDS, SSFS	2005
			UNSSDS, SSFS, SSDIS, SSDI%	2006
			UNSSDS, SSFS	Across years
	UB0331	83.1	UNSSDS, SSFS	2005
			UNSSDS, SSFS, SSDIS, SSDI%	2006
			UNSSDS	Across years
	CA0681	84.7	UNSSDS	2005
			UNSSDS, SSFS, SSDIS, SSDI%	2006
			UNSSDS, SSFS	Across years
	AG0359	93.0	UNSSDS, UNSSDI, SSFS, SSDIS, SSDI%	2006
			UNSSDS, SSFS, SSDIS, SSDI%	Across years
	CA0423	95.6	UNSSDS	2005
			UNSSDS, UNSSDI, SSFS, SSDIS, SSDI%	2006
			UNSSDS, SSFS, SSDI%	Across years
	AG0086	96.1	SSFS	2005
			UNSSDS, UNSSDI, SSFS, SSDIS, SSDI%	2006
			UNSSDS, SSFS, SSDIS, SSDI%	Across years
	CA0896	96.4	UNSSDS, SSFS, SSDIS, SSDI%	2006, Across years
	PE0250	96.6	SSFS	2005
			UNSSDS, UNSSDI, SSFS, SSDIS, SSDI%	2006
			UNSSDS, SSFS, SSDIS, SSDI%	Across years
N18	UB0315	34.8	UNSSDI, SSDI%	2008
	CA0739	42.2	UNSSDI, SSDI%	2008
N19	UB0307	30.1	UNSSDS, SSFS, SSDIS, SSDI%	2005
			UNSSDI, SSFS, SSDIS, SSDI%	2006
			UNSSDI, SSFS, SSDIS, SSDI%	Across years
	CA0221	31.6	UNSSDS, SSFS, SSDIS, SSDI%	2005
			SSDIS	Across years
	BG1241	32.3	UNSSDS, SSFS, SSDIS, SSDI%	2005
			UNSSDI, SSFS, SSDIS	Across years
	KK98G	41.8	UNSSDS, UNSSDI, SSFS, SSDIS, SSDI%	2005
			UNSSDI	2006
			UNSSDI, SSFS, SSDIS, SSDI%	Across years

Table 6: QTLs associated with *Sclerotinia* whole plant field tolerance.

Linkage	Parameter	Year	QTL interval	LOD score	Phenotypic variation explained (%)
N1	SSDI%, SSDIS, SSFS, UNSSDS	2005	BG0988-AG0482	3.2	7.9
	SSFS, SSDIS	2008	AG0243-AG0482	2.4	16.5
	UNSSDS	2008	AG0243-BG1453	2.4	16.6
	SSDI%	2008	BG0988-AG0482	2.6	17.8
	SSDIS	across years	BG0988-AG0482	2.2	5.5
N7	SSDI%, SSDIS, UNSSDS	2007	AG0510-CA0105	3.4	8.7
	SSDIS	across years	AG0510-CA0105	2.3	5.7
N9	UNSSDI	2006, across years	AG0378-KK66	3.4	8.3
N11	UNSSDS	2006	CA0226-BG0713	3.4	7.7
	SSFS	2007	CA0233-CA1080	2.2	5.5
	SSFS, UNSSDS	across years	CA0233-AG0370	3.3	7.9
N12	SSFS	2005	CA1027-PE0063	2.3	5.7
	UNSSDS	2005	CA1027-UB0331	3.0	7.1
	SSDI%, SSDIS, SSFS, UNSSDS, UNSSDI	2006	CA0423-PE0250	4.4	10.3
	SSDI%	across years	CA0423-PE0250	2.5	6.0
	SSDIS, SSFS	across years	AG0359-PE0250	2.7	6.4
	UNSSDS	across years	AG0359-CA0896	3.9	9.2
N18	SSDI%, UNSSDI	2008	UB0315-CA0739	3.2	22.2
	SSDI%, UNSSDI	across years	UB0315-CA0739	2.4	17.2
N19	SSDIS, SSFS	2005	CA0221-KK98G	2.6	6.3
	SSDIS	2006	UB0307-BG1241	2.1	5.0
	SSDI%, SSFS	across years	BG1241-KK98G	2.2	5.4
	SSDIS, UNSSDI	across years	CA0221-BG1241	2.4	5.7

Additional information about the SSR markers flanking the seven QTLs associated with whole field plant resistance to *Sclerotinia* are shown in Table 14 in Example 12. The forward and reverse primer sequences for each marker are also provided. "Repeat" indicates the SSRs or SNPs assocaited with each marker. The positions of the SSRs and SNPS are shown in the sequence information located at the end of the specification.

Additional information about the alleles and allele size of each SSR marker flanking the 7 QTLs associated with whole plant field resistance to *Sclerotinia* is provided in Table 7.

Table 7. The alleles and allele size of each SSR marker flanking the seven *Sclerotinia* (SCL) QTLs.

Linkage Group	Marker	Allele	Allele Size	Favorable allele for SCL
N1	AG0093	a	221	
	AG0093	b	223	yes
	AG0243	a	177	
	AG0243	b	182	
	AG0243	c	193	
	AG0243	d	189	
	AG0243	e	null	yes
	BG1453	a	120	
	BG1453	b	122	
	BG1453	c	130	
	BG1453	d	132	yes
	BG1453	e	134	
	BG1453	f	146	
	BG1453	g	148	
	BG1453	h	152	
	BG1453	i	154	
	BG1453	j	156	
	BG1453	k	172	
	BG1453	l	142	
	BG1453	m	138	
	BG1453	n	158	
	BG1453	o	162	
	BG1453	p	144	
	BG1453	q	136	
	BG1453	r	176	
	BG1453	s	114	
	BG1453	t	160	
	UB0163	b	111	
	UB0163	c	129	yes
	UB0163	e	107	
	UB0163	f	117	
	AG0391	c	139	
	AG0391	a	127	yes
	AG0045	a	168	
	AG0045	b	170	yes
	AG0304	a	163	
	AG0304	b	226	
	AG0304	c	229	yes

	PE0203	a	205	
	PE0203	b	209	yes
	PE0203	c	211	
	PE0203	d	203	
	PE0203	e	207	
	BG0119	a	270	
	BG0119	b	252	yes
	BG0988	a	184	
	BG0988	b	208	
	BG0988	c	200	yes
	BG0988	d	190	
	BG0988	e	208	
	BG0988	f	186	
	AG0482	a	277	yes
	AG0482	b	280	
	AG0482	c	283	
	AG0482	d	286	
	AG0482	e	271	
N7	AG0510	a	272	
	AG0510	b	278	
	AG0510	c	282	yes
	CA0105	a	152	yes
	CA0105	b	170	
N9	AG0378	a	275	
	AG0378	b	281	yes
	AG0378	c	290	
	AG0378	d	293	
	AG0378	e	284	
	AG0378	f	312	
	AG0378	g	299	
	AG0378	h	296	
N11	CA0120	a	138	
	CA0120	b	160	
	CA0120	c	172	yes
	CA0120	d	163	
	CA0120	e	169	
	CA0233	a	298	yes
	CA0226	a	229	
	CA0226	b	250	yes
	CA0226	c	252	
	CA0226	d	221	
	CA0226	e	344	
	CA0546	a	110	

	CA0546	b	120	
	CA0546	c	123	
	CA0546	d	146	yes
	CA0546	e	149	
	CA0546	f	126	
	CA0546	g	144	
	CA0546	h	112	
	BG1149	a	260	yes
	BG1149	b	266	
	BG1149	c	263	
	CA1080	a	300	
	CA1080	b	303	
	CA1080	c	306	
	CA1080	d	325	
	CA1080	e	336	
	CA1080	f	339	yes
	CA1080	g	342	
	CA1080	h	345	
	CA1080	i	348	
	CA1080	j	351	
	CA1080	k	354	
	CA1080	l	357	
	CA1080	m	333	
	CA1080	n	330	
	BG1230	a	252	
	BG1230	b	288	yes
	AG0370	a	283	yes
	AG0370	b	290	
	AG0370	c	298	
	AG0370	d	316	
	AG0370	e	287	
	BG0713	a	222	
	BG0713	b	226	yes
	BG0713	c	228	
	BG0713	d	216	
	BG0869	a	212	
	BG0869	b	216	yes
	BG1513	a	164	yes
	BG1513	b	214	
	BG1513	c	216	
	BG0181	a	210	
	BG0181	b	216	yes
	CA1097	a	245	

	CA1097	b	248	yes
	CA1097	c	251	
	CA1097	d	260	
	CA1097	e	239	
N12	CA0753	a	214	
	CA0753	b	216	
	CA0753	c	281	yes
	CA0753	d	291	
	CA0753	e	218	
	CA1027	a	297	yes
	CA1027	b	300	
	CA1027	c	303	
	CA1027	d	306	
	PE0063	a	114	
	PE0063	b	126	yes
	UB0331	a	123	
	UB0331	b	126	yes
	CA0681	a	264	
	CA0681	b	266	yes
	CA0681	c	268	
	CA0681	d	276	
	AG0359	a	306	
	AG0359	b	330	yes
	AG0359	d	315	
	AG0359	f	312	
	CA0423	a	195	
	CA0423	b	201	
	CA0423	c	204	
	CA0423	d	207	yes
	CA0423	e	210	
	AG0086	a	238	
	AG0086	b	226	yes
	AG0086	c	232	
	CA0896	a	254	yes
	CA0896	b	266	
	CA0896	c	260	
	PE0250	a	239	
	PE0250	b	245	yes
	PE0250	c	269	
N18	UB0315	a	127	
	UB0315	b	131	yes
	UB0315	c	133	
	CA0739	a	220	

	CA0739	b	222	
	CA0739	c	232	
	CA0739	d	234	yes
	CA0739	e	240	
	CA0739	f	224	
	CA0739	g	236	
	CA0739	h	242	
N19	UB0307	a	120	yes
	UB0307	b	126	
	UB0307	c	134	
	UB0307	e	108	
	CA0221	a	271	
	CA0221	b	253	yes
	CA0221	c	265	
	BG1241	a	370	
	BG1241	b	329	yes

Example 5: Validation of the 7 QTLs Associated with Whole Plant Field Resistance to *Sclerotinia*

Simulated validation of the 7 QTLs associated with whole plant field resistance to *Sclerotinia* was performed on 16 *Sclerotinia*-resistant and 10 *Sclerotinia* susceptible breeding lines from Pioneer Hi-Bred's spring canola program under extreme disease pressure field research conditions. This allowed the development of extreme disease conditions every year, regardless of the natural environment.

Referring to Table 8 below, each resistant line was genotyped for SSR markers flanking the seven QTLs identified. In Table 9, the first digit in the QTL group refers to the linkage group and the second digit refers to the QTL number on that linkage group. The number above the marker names represent the positions (in centiMorgans) of the marker on the linkage group. Each allele was denoted as i) present only in resistant lines (light gray), ii) present only in susceptible lines (dark gray), or iii) present in both resistant and susceptible lines (white). The total number of favorable alleles were added for each breeding line by assigning either 1) (allele present only in resistant line), 0.5 (allele is present in both resistant and susceptible) or 0 (allele present only in susceptible line). The percentage of favorable alleles in the resistant lines ranged from 63-90% and only 13-47% in susceptible breeding lines. This correlation indicates that the markers flanking the seven QTLs identified as being associated with *Sclerotinia* in breeding populations can be used to

select for individuals that had the highest number of favorable alleles in breeding populations. Those individuals with the highest percentage of favorable alleles would be selected as good candidates for *Sclerotinia* resistance.

CA 02823467 2013-06-20

Table 8. Comparison of allele scores among three groups: sources of *Sclerotinia* resistance, elite lines with *Sclerotinia* resistance, and elite lines susceptible to *Sclerotinia*. (First number in QTL name refers to linkage group and second number refers to QTL number on that linkage group.

			Sclerotinia QTLs and Flanking SSR Markers															
			Scl_1.1		Scl_7.1		Scl_9.1	Scl_11.1			Scl_12.1			Scl_18.1		Scl_19.1		
			60.3	66.1	0	0.5	0	28.8	30.8	31.3	93	96.1	96.6	34.8	42.2	30.1	31.6	
		SSDIS	BG0988	AG0482	AG0510	CA0105	AG0378	CA0233	BG1149	AG0370	AG0359	AG0086	PE0250	UB0315	CA0739	UB0307	CA0221	Number of favorable alleles	Percent favorable alleles
Sources of *Sclerotinia* Resistance	WC-865	5.5	A	b,d	a,b	a	a	+	a,c	d		a	a,b	a,b	e	a,c	a,b	12.5	83%
	04DHS11418(MappingParent)	6.5	C	a,b	a,c	a	b	+	a	a	b	a,b	a,b	c	a,d	a	a,b	13.5	90%
	SC-1068	6.8	C	a,b	a,c	a	d	+	a	a	b	a	a,b	c	a,d	a,c	b,c	12.0	80%
	SC-1349	7.2	C	b,d	a	a	d	+	a	a	b	a,b	a,b	c	d	a	a,b	12.5	83%
	SC-182	6.0	C	b,d	a,c	a	b,d	+	a	d	b	a,b	a,b	c	a,d	a	a,b	13.5	90%
Elite lines with SCL resistance	SC-391	5.0	C	b,c,d	a,b	a,b	b,d	+		a	b,d	a,b	a,b	b	c,c	a	a,b	11.0	73%
	SC-631	6.0	C	b,d	a,b	b	b	+		a	b,d	a,b	a,b	a,b	a,c	a	a,b,c	11.0	73%
	SC-613	5.0	C	a,b,c	c	a	b	+	a,c	a	a,b	a	a,b	b	a,d	a,c	b,c	12.0	80%
	SC-940	6.0	C	b,d	a,b	b	b,d	+	a	a	b,d	a,b	a,b	b	d	a	a,b	11.5	77%
	SC-942	5.0	B	b,d	c	a	a	+	a	b	b	a,b	a,b	b	c	a	a,b,c	12.0	80%
	SC-1023	5.0	C	b,c	a	a	b	+	a	a	b	a,b	a,b	b	d	a,c	b,c	11.0	73%
	SC-1178	5.0	B	d	a,c	a	a	c	b	b	b	a,b	a,b	b	c	a	b,c	9.5	63%
	SC-1179	6.0	B	b,d	c	a	b	c	c	b	b	a,b	a,b	b	c	a	a,b	10.5	70%
	SC-1180	6.0	B	b,d	a	a	b	+	a	b	b	a,b	a,b	b	a,d	a	a,b	13.0	87%
	SC-1284	6.0	B	b,d	a,c	a	a	+	a	b	b	a,b	a,b	b	a,d	a	a,b	12.5	83%
	SC-1285	7.0	B	b,d	b	b	b	+	a	b	b	a,b	a,b		a,c	a	a,b,c	11.0	73%
Elite susceptible lines	SC-067	1.0	C	b,c	a,b	b	d	c	c	b	a,d	b,c	b,c	b	a,e	a,c	b,c	3.5	23%
	SC-062	2.0	C		b	b	a,d	+	a	b	b	a,b	a,b	a	c	a,c	a,b	7.0	47%
	SC-004	2.0	C	b,c	a,b	b	d	c	b	b	b	a,c	a,b,c	b	c	a,c	b,c	4.0	27%
	SC-101	3.0	C	a,b,c	a,b,c	a,b	a	c	c	c	b	a	b	a	a,e	a	a,b	5.0	33%
	SC-105	2.0	C	b,c	a,b	b	d	c	c	a	b,d	a,b	a,b	b		a,c	a,b	6.5	43%
	SC-112	2.0	C	d	b	b	a	c	b	b	b	a,b	a,b	b		a	b,c	6.0	40%
	SC-129	2.0	C	b,c	c	a	d	c	c	c	b	a,b	a,b	b	d	a,c	b,c	6.5	43%
	SC-139	1.0	C	b,c	a,b	b	a	c	b	b	b	a,b	b	a	c	a	b,c	5.0	33%
	SC-412	1.0	C	b,c	a,b,c	a	d	+	c	b	b,d	a,b	a	a		a,c	b,c	5.5	37%
	PHI2004HS1(MappingParent)	1.5	C	b,c	a,b	b	d	c	c	c	a	b,c	a,b,c	b	a,e	a,c	b,c	2.5	16%

	allele combination present only in resistant lines
	allele combination present in resistant and susceptible lines
	allele combination present only in susceptible lines
	undetermined

Example 6: Introgressing *Sclerotinia* Resistance from Spring *Brassica napus* to Winter *Brassica napus*.

The *Sclerotinia* resistant source 04DHS11418 was crossed to winter canola lines in bi-parental, 3-way or complex crosses (as shown, for example, in Table 9). The F1 cross was then backcrossed to an elite susceptible parent. At the BC1F1 generation, approximately 500 progeny (minimum number required to identify at least one individual with all favorable alleles present) were generated for each cross and submitted for marker analysis using the markers identified as being associated with *Sclerotinia* in spring canola. Each individual sample was examined for presence of the favorable alleles from the *Sclerotinia* resistant line 04DHS11418. The percentage of favorable alleles present in each sample was calculated and the top three from each population were used to cross back to the recurrent parent. This process was repeated again at the BC2 stage. In addition, selections were intermated to develop populations in which individuals could be identified with homozygous desirable *Sclerotinia* alleles.

Table 9. Example of introgression of *Sclerotinia* resistance into winter *B. napus* using marker-assisted selection

Example Set 1 MAS Results

		N9	N7	N7	N18	N18	N1	N1	N11	N19	N12	N12		
	Sample Name	AG 0378	AG 0510	BG 1439	CA 0739	UB 0315	AG 0482	BG 0988	AG 0370	BG 1241	AG 0359	CA 0423	Number of favorable alleles	Percent favorable alleles
Female	04DHS 11418	b	a,c	a,b	a,d	c	a,b	c	a	d	b	b,d		
	WC-058	b	a,b	a	g	a	d	a	d	e	a	a,e		
Male	WC-022	b	a,b	a	g	a	d	e	a	e	a	e		
	WC-663	b	a,b	a	c,g	a	d	a,e	d	e	a	d,e		
	09-CMAS-01-1641	b,d	b	a,b	a,d,g	a,c	b,d	c,e	a	d,e	a,b	b,d,e	9.5	86.4
	09-CMAS-01-1737	b,d	a,b	a,b	a,d,g	a,c	a,d	c,e	a	d,e	a,b	b,d,e	9.5	86.4

Example Set 2 MAS Results

		N9	N7	N7	N18	N18	N1	N1	N11	N19	N12	N12		
	Sample Name	AG 0378	AG 0510	BG 1439	CA 0739	UB 0315	AG 0482	BG 0988	AG 0370	BG 1241	AG03 59	CA 0423	Number of favorable alleles	Percent favorable alleles
Male	04DHS11 418	B	a,c	a,b	a,d	c	a,b	c	a	d	b	b,d		
	WC-058	D	a,b	a	g	a	d	a	d	e	a	a,e		
Female	WC-227	D	a,b	a	d	a	d	b	a	d	a	a,e		

	09-CMAS-01-715	b,d	a,b,c	a,b	a,d	a,c	a,b,d	c,e	a	d	a	a,b,d,e	10.0	90.9
	09-CMAS-01-929	b,d	a,b	a,b	a,d	a,c	a,b,d	c,e	a	d	a,b	a,b,d,e	10.0	90.9

Example Set 3 MAS Results

		N9	N7	N7	N18	N18	N1	N1	N11	N19	N12	N12		
	Sample Name	AG 0378	AG 0510	BG 1439	CA 0739	UB 0315	AG 0482	BG 0988	AG 0370	BG 1241	AG 0359	CA 0423	Number offavorabl ealleles	Percent favorable alleles
Male	04DHS11418	B	a,c	a,b	a,d	c	a,b	c	a	d	b	b,d		
	WC-058	D	a,b	a	g	a	d	a	d	e	a	a,e		
Fe-male	WC-063	D	a,b	a	d	a	d	a	c,d	d	a,c	d,e		
	WC-457	D	a,b	a,b	c	a	b,d	a	c,d	e	a	d,e		
	WC-227	D	a,b	A	d	a	d	e	a	c,d	a	a,e		
	09-CMAS-01-1205	b,d	a,b,c	a,b	a,d	a,c	a,b,d	a,c	c,d	d	a,b	a,b,d,e	9.5	90.5
	09-CMAS-01-1234	b,d	a,b	a,b	a,d	a,c	a,d	c,e	a	d,e	a,b	b,d,e	9.0	85.7
	09-CMAS-01-1241	b,d	a,b,c	a,b	a,d	a,c	a,d	c,e	a	d,e	a,b	d,e	9.0	85.7

Example 7: Mapping Population #2

The parents used for the mapping population were 06DSB13911 (a *Sclerotinia* resistant double haploid) and PHI2008HS1 (a susceptible spring canola DH line polymorphic to the resistant line, selected for having a similar agronomic phenotype with consistent high susceptibility in the same testing environment where the 06DSB13911 line was selected for resistance) (see Table 10). These lines were used to develop a double haploid mapping population consisting of 187 progeny.

Table 10. Measuring field performance under extreme disease pressure (research trials)

Rating SSDIS**	Category	Disease incidence SSDI%*	Spring Checks	Mapping Parents
1.0	Highly susceptible	≥80	44A89=1	
1.1 – 2.0	Susceptible	79 – 70	46A65=2	PHI2008HS1
2.1 – 3.0	Moderately susceptible	69 – 60	46A76=3	
3.1 – 4.0		59 – 50		
4.1 – 5.0	Moderately resistant	49 – 40		
5.1 – 6.0		39 – 30		

6.1 – 7.0	Resistant	29 – 20		
7.1 – 8.0		19 – 10		06DSB13911
8.1 – 9.0	Highly resistant	9 – 0		

* SSDI% *Sclerotinia sclerotiorum* Disease Incidence %. SSDI% is UNSSDI (where UNSSDI is the percentage of plants in a population infected with *Sclerotinia)* rating adjusted for a deviation from the expected mean of checks 06DSB13911 (15%) and PHI2008HS1 (75%). This rating is used only under controlled extreme disease pressure field research conditions. It is calculated by multiplying the observed SSDI% by Factor X, where Factor X is the factor that brings the average SSDI% of the appropriate checks to 45%.

**SSDIS *Sclerotinia sclerotiorum*. SSDIS is the UNSSDI rating adjusted for a deviation from the expected mean of check parents for spring canola under extreme disease pressure. This rating is used only under controlled extreme disease pressure field research conditions. It is calculated by multiplying the observed UNSSDI by Factor X, where Factor X is the factor that brings the average SSDI% of the appropriate checks to 45%. Adjustment for severity is done after incidence adjustment.

UNSSDS is a rating of the extent of disease development on an affected plant. Two scales are used in the invention. The Pioneer SSDS scale ranges from 1 (dead) to 9 (no disease) and the Public scale ranges from 0 (no disease) to 5 (dead) plant. For details of the Pioneer SSDS scale, see Table 15 of WO 2006/135717. The Public scale is provided as follows: 0 = no disease; 1 = superficial lesions or small branch affected; 2 = large branch dead; 3 = main stem at least 50% girdled; 4 = main stem girdled but plant produced good seed; 5 = main stem girdled, much reduced yield.

Both parents have good standability. Therefore, standability or lodging resistance is fixed, thus eliminating this variable in the mapping process. The choice of a highly susceptible line resulted in a population without transgressive segregation (i.e., all resistance came from 06DSB13911 and no DH progeny lines were more resistant than the resistant parent). Over a period of three years, the population was phenotyped in the field and genotyped with SSR molecular markers. Phenotyping was carried out as described in WO 2006/135717 .

Example 8: *Sclerotinia* Screening

Disease Scoring

The plants of the double haploid mapping population created, as described in Example 7, were rated for disease as described in Table 4 of Example 3. The unadjusted parameters (e.g., UNSSDI and UNSSDS) showed year to year variation due to environmental variation such as positional variation in the field and weather conditions. Such variation would be expected by one skilled in the art.

Example 9: Genetic Mapping and QTL Analysis

Genetic mapping and QTL analysis were performed using JoinMap v3.0 (Van Ooijen, J.W. and R.E. Voorrips, 2001 JoinMap® 3.0, Software for the calculation of genetic linkage

maps. Plant Research International, Wageningen, the Netherlands). The Kosambi centiMorgan function was used. A QTL was declared if its LOD score exceeded the threshold of 2.0. LOD stands for logarithm of the odds (to the base 10).

Genetic Mapping

Genetic mapping has placed 278 molecular markers to 19 linkage groups (Lg) that correspond to 19 canola chromosomes and public linkage group nomenclature. The linkage map covers ~1100 cM.

QTL Analysis

QTL analysis using simple interval mapping and composite interval mapping (CIM) (Zeng (1994), Genetics 136:1457) identified 12 linkage groups (N1, N3, N4, N8, N9, N10, N11, N12, N13, N15, N18 and N19) contributing to whole plant field resistance to *Sclerotinia*. In addition, regions identified by interval mapping as being associated with Sclerotinia resistance were confirmed by single-factor analysis of variance (PROC GLM, SAS Enterprise Guide 4.2) on *Sclerotinia* parameters at the $P \leq 0.01$ significance level. These QTLs are identified in Tables 11 and 12 below. As shown by the "Phenotypic Variation Explained" values in Table 12, some QTLs had a larger effect on *Sclerotinia* resistance than others.

Table 11: Markers significantly associated with Sclerotinia resistance at $P \leq 0.01$.

Linkage Group	Marker	Map Position (cM)	Parameter	Year
N1	CA0614	8.2	UNSSDI	2010
			UNSSDS, SSFS, SSDIS	2009
	BG0111	10.9	UNSSDI, SSDI%	2010
	BG1392	22.7	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSDI%, SSDIS	2010
	BG1182	35.6	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSDI%	2010
	BG1090	41.9	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSDI%	2010
	AG0093	46.3	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
	BG1453	53.1	UNSSDS, UNSSDI, SSFS, SSDIS	2009

Linkage Group	Marker	Map Position (cM)	Parameter	Year
N3			UNSSDI, SSFS, SSDI%, SSDIS	2010
	PE0017	53.5	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
	AG0391	56.9	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
	AG0304	56.9	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
	UB0163	57.0	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
	PE0203	57.5	UNSSDS, SSFS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
	AG0482	68.5	UNSSDI	2010
	PE0177	73.6	UNSSDS, SSFS, SSDIS	2009
			UNSSDI	2010
	CA0410	0.0	UNSSDS, SSDIS	2009
			UNSSDI, SSFS	2011
	BG1368	2.2	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS	2011
	BG1197	37.3	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS	2011
			UNSSDI, SSDI%	2010
	AG0272	40.1	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSDI%	2010
	AG0023	40.1	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSDI%	2010
N4	BG1442	0.0	UNSSDS, SSFS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSFS, SSDI%, SSDIS	2011
	UB0181	3.9	UNSSDS, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
	UB0126	10.3	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2011
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
	AG0477	10.8	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSFS, SSDI%, SSDIS	2011
	BG1127	11.0	UNSSDS, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSFS, SSDI%, SSDIS	2011

Linkage Group	Marker	Map Position (cM)	Parameter	Year
	AG0125	14.2	UNSSDS, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSFS, SSDI%, SSDIS	2011
	BG1244	15.0	UNSSDS, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			SSFS	2011
	AG0239	15.2	UNSSDS, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			SSFS	2011
	AG0203	16.1	UNSSDS, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			SSFS	2011
	BG0106	18.5	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			SSFS	2011
N8	CA0837	0.0	UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2009
	BG0647	3.1	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI	2011
	AG0070	3.1	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI	2010
			UNSSDI	2011
	PE0281	3.2	UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2009
			UNSSDI	2011
	AG0324	4.6	UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2009
	BG1101	7.3	UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
	AG0328	15.6	UNSSDS, UNSSDI, SSFS, SSDIS	2009
	BG1449	25.9	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
	BG1062	25.9	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
	AG0410	30.3	UNSSDS, SSFS, SSDIS	2009
	BG1286	33.2	UNSSDS	2009
N9	CA1034	0.0	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
	CA0834	0.0	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
	AG0378	0.1	UNSSDI	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010

Linkage Group	Marker	Map Position (cM)	Parameter	Year
	BG1123	33.4	UNSSDI, UNSSDS, SSFS, SSDIS	2011
	AG0323	45.8	UNSSDI, UNSSDS, SSFS, SSDIS	2011
	BG0295	49.5	UNSSDI, UNSSDS, SSFS, SSDIS	2011
	AG0441	66.6	UNSSDI, UNSSDS, SSFS, SSDIS	2011
N10	BG0228	0.0	UNSSDS, UNSSDI, SSFS, SSDIS	2009
	PE0355	5.1	SSFS	2009
	AG0171	7.1	UNSSDS, SSFS	2009
	BG0651	13.6	UNSSDS, SSFS	2009
	BG0255	17.9	UNSSDS, SSFS	2009
			SSDI%	2011
	AG0047	19.9	UNSSDS, SSFS	2009
			SSDI%	2011
	UB0196	29.4	SSDI%	2011
	UB0015	29.5	SSDI%	2011
	PE0131	36.8	SSDI%	2011
N11	CA0120	20.6	UNSSDS	2010
	BG0452	32.8	UNSSDS	2010
			UNSSDS, UNSSDI, SSFS, SSDIS	2011
	BG0031	34.1	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDS, UNSSDI, SSFS, SSDIS	2011
	CA1035	37.5	UNSSDS, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDS, UNSSDI, SSFS, SSDIS	2011
	CA0546	37.5	UNSSDS, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDS, UNSSDI, SSFS, SSDIS	2011
	CA1032	37.5	UNSSDS, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDS, UNSSDI, SSFS, SSDIS	2011
	BG1149	37.8	UNSSDS, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDS, UNSSDI, SSFS, SSDIS	2011
	BG1230	39.1	UNSSDS, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDS, UNSSDI, SSFS, SSDIS	2011
	BG1513	42.0	UNSSDS	2009

Linkage Group	Marker	Map Position (cM)	Parameter	Year
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDS, UNSSDI, SSFS, SSDIS	2011
	CA0328	42.3	UNSSDS	2010
			UNSSDS, UNSSDI, SSFS, SSDIS	2011
	PE0324	49.1	UNSSDS	2010
			UNSSDS, UNSSDI, SSFS, SSDIS	2011
	PE0283	53.1	UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2009, 2010
	CA0163	56.6	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
N12	BG1321	0.0	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSDI%	2010
			UNSSDI, SSDI%	2011
	PE0133	19.9	UNSSDS, SSFS, SSDI%, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSDI%	2011
	CA0456	21.2	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSDI%	2011
	PE0063	27.2	UNSSDS, SSFS, SSDI%, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSDI%	2011
	CA1027	27.2	UNSSDS, SSFS, SSDI%, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSDI%	2011
	BG0864	28.9	SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSDI%	2011
	CA1090	28.9	SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSDI%	2011
	CA0991	28.9	SSFS, SSDI%, SSDIS	2009
			UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSDI%	2011
N13	CA0603	0.0	UNSSDI, SSFS, SSDI%, SSDIS	2010
	BG1288	2.9	UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDS, SSFS, SSDIS	2011
	CA0488	11.9	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDS, SSFS, SSDIS	2011

Linkage Group	Marker	Map Position (cM)	Parameter	Year
	PE0012	13.1	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDS, SSFS, SSDIS	2011
	PE0340	13.5	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDS, SSFS, SSDIS	2011
	BG0516	22.1	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDS, SSFS, SSDIS	2011
	AG0504	42.8	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSDI%, SSDIS	2010
			UNSSDS, SSFS, SSDIS	2011
	AG0148	55.5	UNSSDS, UNSSDI, SSFS, SSDIS	2009
	CA0736	65.8	UNSSDS, SSFS	2009
N15	PE0286	0.0	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSDI%, SSDIS	2010
			UNSSDI, SSFS, SSDI%, SSDIS	2011
	PE0091	6.7	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSFS, SSDI%, SSDIS	2011
	PE0187	15.2	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSFS, SSDI%, SSDIS	2011
	CA0719	24.8	UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			UNSSDI, SSFS, SSDI%, SSDIS	2011
	AG0369	43.9	UNSSDI, SSFS, SSDI%, SSDIS	2009, 2010
N18	BG0278	12.9	UNSSDI, SSFS, SSDIS	2009
	CA0739	21.6	UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2009
			UNSSDI, SSDI%, SSDIS	2010
	UB0315	27.0	UNSSDS, UNSSDI, SSFS, SSDI%, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
			SSDI%	2011
	CA0636	32.4	UNSSDS, UNSSDI, SSFS, SSDIS	2009
			UNSSDI, SSDI%	2010
			SSDI%	2011
N19	CA1107	0.0	UNSSDI, SSFS, SSDI%, SSDIS	2010
	CA0552	16.1	UNSSDS, SSFS, SSDIS	2009
			UNSSDI	2010

Linkage Group	Marker	Map Position (cM)	Parameter	Year
	CA1066	19.7	UNSSDS, SSFS	2009
	BG1241	27.6	UNSSDS, SSFS, SSDIS	2009
			UNSSDI, SSFS, SSDI%, SSDIS	2010
	CA0221	29.7	UNSSDS, SSFS, SSDIS	2009

Table 12: QTL interval, LOD score and explained phenotypic variation of QTLs associated with *Sclerotinia* whole plant field resistance or improved whole plant field resistance

Linkage Group	Parameter	Year	QTL Interval	LOD Score	Phenotypic Variation Explained (%)
N1	SSFS, SSDIS	2010	AG0093-PE0203	2.9	7.8
	SSFS, SSDIS	2009	BG0111-BG1392	3.5	8.4
	SSDI%	2010	BG0111-BG1392	2.8	7.4
	UNSSDI	2010	BG1090-AG0482	5.0	11.5
	UNSSDI, UNSSDS, SSFS, SSDIS	2009	BG1090-PE0203	3.9	10.5
	SSDI%	2010	BG1090-PE0203	4.7	11.1
	UNSSDS	2009	CA0614-BG1392	5.5	13.0
	UNSSDI	2010	CA0614-BG1392	3.2	12.5
N3	UNSSDS, SSFS, SSDIS	2009	BG1197-AG0023	4.0	9.6
	UNSSDI, SSDI%	2010	BG1197-AG0023	2.2	5.4
	SSFS, SSDIS	2009	CA0410-BG1368	2.8	6.7
	UNSSDI, SSFS	2011	CA0410-BG1197	2.9	3.9
N4	UNSSDS, SSFS	2009, 2010	BG1442-BG0106	6.3	14.7
	UNSSDI, SSDI%, SSDIS	2010	BG1442-BG0106	8.1	18.5
	UNSSDI, SSFS, SSDI%, SSDIS	2011	BG1442-BG0106	7.6	11.1
	SSDIS	2009	UB0181-BG0106	3.4	8.1
N8	UNSSDI	2010	BG1449-BG1062	2.3	6.0
	SSDI%	2009	CA0837-AG0328	2.3	6.5
	SSDIS	2009	CA0837-BG1062	6.6	15.1
	UNSSDI	2010	CA0837-BG1101	2.3	5.6
	UNSSDS, SSFS	2009	CA0837-BG1286	7.5	17.2
	UNSSDI	2009	CA0837-BG1449	3.7	10.5
	UNSSDI	2011	BG0647 -PE0281	2.3	2.6
N9	UNSSDS, SSFS, SSDIS	2009	AG0323-BG0295	4.5	10.6
	UNSSDI, SSDI%	2010	AG0323-BG0295	2.9	6.9
	UNSSDI, SSFS, SSDI%, SSDIS	2010	CA1034-AG0378	5.3	12.2
	UNSSDI, UNSSDS, SSFS, SSDIS	2011	BG1123-AG0441	3.2	5.7

Linkage Group	Parameter	Year	QTL Interval	LOD Score	Phenotypic Variation Explained (%)
N10	UNSSDS, SSFS	2009	BG0228-AG0047	3.3	7.9
	SSDI%	2011	BG0255-PE0131	3.0	4.3
N11	SSFS, SSDIS	2009	BG0031-BG1149	4.4	10.5
	UNSSDS	2009	BG0031-BG1230	4.1	10.5
	UNSSDI, SSDI%	2010	BG0031-BG1230	3.4	8.5
	SSFS, SSDIS	2010	BG0031-BG1513	5.3	12.2
	UNSSDS	2010	CA0120-CA0328	5.4	13.9
	UNSSDI, SSFS, SSDIS	2009, 2010	PE0283-CA0163	5.2	12.2
	UNSSDS	2009	PE0283-CA0163	5.2	12.3
	SSDI%	2010	PE0283-CA0163	2.4	5.6
	SSDI%	2009	PE0324-PE0283	2.2	5.4
	UNSSDS, UNSSDI, SSFS, SSDIS	2011	CA0328-PE0324	5.8	7.2
N12	SSFS, SSDIS	2009	BG1321-CA0991	5.1	13.8
	UNSSDI, SSDI%	2010	BG1321-CA0991	4.6	11.6
	UNSSDS	2009	BG1321-CA1027	3.7	10.2
	UNSSDI	2009	BG1321-PE0133	2.7	6.5
	UNSSDI, SSDI%	2011	BG1231- PE0133	4.1	6.6
	SSDI%	2009	PE0063-CA0991	2.0	5.0
	SSFS, SSDIS	2010	PE0133-CA0991	4.9	12.2
N13	UNSSDI	2009	BG0516-AG0148	2.9	9.7
	SSFS, SSDIS	2009	CA0488-AG0148	7.3	21.7
	UNSSDS	2009	CA0488-CA0736	9.8	27.3
	SSDI%, SSDIS	2010	CA0603-AG0504	7.3	17.3
	UNSSDS, SSFS, SSDIS	2011	BG1288-AG0504	4.6	6.8
N15	SSDI%	2009	CA0719-AG0369	2.7	8.5
	SSFS	2010	PE0091-PE0187	4.2	10.0
	UNSSDI, UNSSDS, SSFS, SSDIS	2009	PE0286-AG0369	7.1	17.8
	UNSSDI, SSDI%, SSDIS	2010	PE0286-PE0187	7.4	17.2
	UNSSDI, SSFS, SSDI%, SSDIS	2011	PE0286-CA0719	8.7	12.1
N18	UNSSDS, SSFS, SSDIS	2009	AG0285-CA0636	8.4	20.0
	UNSSDI	2009	BG0278-CA0779	6.6	16.4
	SSDI%	2009	CA0739-CA0636	2.9	7.0
	UNSSDI	2010	CA0739-CA0636	3.3	8.0
	SSDI%	2011	CA0739- CA0636	2.7	4.1
	SSDI%	2010	UB0315-CA0636	2.6	6.4
N19	SSDIS	2009	CA0552-CA0221	2.9	9.8
	UNSSDI, SSFS, SSDI%, SSDIS	2010	CA1107-CA0552	2.8	8.0
	UNSSDS, SSFS	2009	CA1107-CA0221	4.3	13.5

Additional information about the SSR markers flanking the twelve QTLs associated with whole field plant resistance to *Sclerotinia* is shown in Table 14 of Example 12, where exemplary sets of forward and reverse primer sequences for each SSR are also provided. "Repeat" indicates the SSRs or SNPs associated with each marker. The positions of the SSRs are shown in the sequence information located in Example 12. Additional information about the alleles and allele size of each SSR marker flanking the 12 QTLs associated with whole plant field resistance to *Sclerotinia* is provided in Table 13.

Table 13. The alleles and allele size of each SSR marker flanking the twelve *Sclerotinia* QTLs. as well as favorable allele for *Sclerotinia* (SCL) resistance.

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
N1	CA0614	a	160	
	CA0614	b	178	yes
	CA0614	c	188	
	CA0614	d	196	
	BG0111	a	141	yes
	BG0111	b	146	
	BG0111	c	147	
	BG0111	d	144	
	BG1392	a	251	yes
	BG1392	b	257	
	BG1182	a	299	
	BG1182	b	301	
	BG1182	c	303	yes
	BG1182	d	347	
	BG1182	e	345	
	BG1090	a	268	yes
	BG1090	b	272	
	AG0093	a	221	
	AG0093	b	223	yes
	BG1453	a	120	
	BG1453	b	122	
	BG1453	c	130	
	BG1453	d	132	yes
	BG1453	e	134	
	BG1453	f	146	

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
	BG1453	g	148	
	BG1453	h	152	
	BG1453	i	154	
	BG1453	j	156	
	BG1453	k	172	
	BG1453	l	142	
	BG1453	m	138	
	BG1453	n	158	
	BG1453	o	162	
	BG1453	p	144	
	BG1453	q	136	
	BG1453	r	176	
	BG1453	s	114	
	BG1453	t	160	
	PE0017	a	81	yes
	PE0017	b	84	
	PE0017	c	78	
	PE0017	d	87	
	AG0391	a	127	yes
	AG0391	b	130	
	AG0391	c	139	
	AG0304	a	163	
	AG0304	b	226	
	AG0304	c	229	yes
	UB0163	b	111	
	UB0163	c	129	yes
	UB0163	e	107	
	UB0163	f	117	
	PE0203	a	205	
	PE0203	b	209	
	PE0203	c	211	yes
	PE0203	d	203	
	PE0203	e	207	
	AG0482	a	278	
	AG0482	b	281	
	AG0482	c	284	
	AG0482	d	287	yes
	AG0482	e	272	
	PE0177	a	197	
	PE0177	b	199	yes
	PE0177	c	201	
	PE0177	d	205	
N2	CA0410	a	140	

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
	CA0410	b	142	yes
	CA0410	c	148	
	CA0410	d	154	
	BG1368	a	128	yes
	BG1368	b	131	
	BG1197	a	262	
	BG1197	b	267	yes
	BG1197	c	272	
	BG1197	d	286	
	AG0272	a	151	yes
	AG0272	b	157	
	AG0272	c	163	
	AG0023	a	128	
	AG0023	b	131	yes
	AG0023	c	139	
	AG0023	d	145	
	AG0023	e	148	
N4	BG1442	a	232	yes
	BG1442	b	238	
	BG1442	c	250	
	BG1442	d	258	
	BG1442	e	266	
	BG1442	f	242	
	BG1442	g	254	
	BG1442	h	230	
	UB0181	a	147	
	UB0181	b	283	
	UB0181	c	369	
	UB0181	d	379	yes
	UB0181	e	275	
	UB0181	f	285	
	UB0181	g	289	
	UB0126	a	216	yes
	UB0126	b	220	
	UB0126	c	230	
	AG0477	a	268	
	AG0477	b	276	
	AG0477	c	278	
	AG0477	d	289	yes
	AG0477	e	292	
	AG0477	f	285	
	BG1127	a	284	
	BG1127	b	290	

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
	BG1127	c	304	yes
	BG1127	d	306	
	BG1127	e	292	
	BG1127	f	294	
	BG1127	g	302	
	AG0125	a	212	
	AG0125	b	238	
	AG0125	c	255	
	AG0125	d	264	yes
	AG0125	e	267	
	AG0125	f	225	
	BG1244	a	247	
	BG1244	b	284	yes
	BG1244	c	275	
	AG0239	a	295	
	AG0239	b	311	
	AG0239	c	314	yes
	AG0203	a	206	
	AG0203	b	209	
	AG0203	c	215	
	AG0203	d	221	yes
	AG0203	j	213	
	AG0203	k	203	
	AG0203	l	227	
	BG0106	a	216	
	BG0106	b	286	
	BG0106	c	310	yes
	BG0106	d	288	
	BG0106	e	276	
	BG0106	f	284	
	BG0106	g	280	
N8	CA0837	a	257	
	CA0837	b	269	
	CA0837	c	253	yes
	CA0837	d	279	
	CA0837	e	283	
	BG0647	a	247	yes
	BG0647	b	250	
	BG0647	c	259	
	BG0647	d	272	
	BG0647	e	275	
	BG0647	f	244	
	AG0070	a	270	yes

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
	AG0070	b	274	
	AG0070	c	272	
	PE0281	a	179	
	PE0281	b	209	yes
	PE0281	c	211	
	PE0281	d	213	
	PE0281	e	233	
	PE0281	f	221	
	PE0281	g	215	
	PE0281	h	236	
	PE0281	i	203	
	PE0281	j	194	
	PE0281	k	206	
	PE0281	l	249	
	AG0324	a	226	yes
	AG0324	b	229	
	AG0324	c	244	
	BG1101	a	210	
	BG1101	b	219	yes
	AG0328	a	222	
	AG0328	b	228	yes
	AG0328	c	255	
	AG0328	d	258	
	AG0328	e	267	
	AG0328	f	270	
	AG0328	g	279	
	AG0328	h	276	
	AG0328	i	281	
	AG0328	j	273	
	AG0328	k	287	
	BG1449	a	132	
	BG1449	b	134	
	BG1449	c	158	
	BG1449	d	160	yes
	BG1449	e	164	
	BG1449	f	156	
	BG1449	g	170	
	BG1062	a	188	
	BG1062	b	210	
	BG1062	c	212	yes
	BG1062	d	186	
	BG1062	e	214	
	BG1062	f	216	

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
	BG1062	g	224	
	BG1062	h	208	
	BG1062	i	196	
	BG1062	j	176	
	BG1062	k	206	
	BG1062	l	218	
	BG1062	m	202	
	BG1062	o	182	
	BG1062	p	222	
	AG0410	a	322	
	AG0410	b	325	
	AG0410	c	328	yes
	AG0410	d	334	
	AG0410	e	337	
	AG0410	f	331	
	BG1286	a	156	
	BG1286	b	159	
	BG1286	c	162	yes
	BG1286	d	165	
N9	CA1034	a	275	yes
	CA1034	b	290	
	CA1034	c	293	
	CA1034	d	306	
	CA1034	e	309	
	CA1034	f	321	
	CA1034	g	299	
	CA1034	h	284	
	CA1034	i	278	
	CA1034	j	315	
	CA1034	k	296	
	CA1034	l	324	
	CA1034	m	327	
	CA1034	n	287	
	CA0834	a	274	yes
	CA0834	b	289	
	CA0834	c	292	
	CA0834	d	301	
	CA0834	e	304	
	CA0834	f	307	
	CA0834	g	298	
	CA0834	h	319	
	CA0834	i	283	
	CA0834	j	313	

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
	CA0834	k	295	
	CA0834	l	277	
	CA0834	m	286	
	CA0834	n	325	
	CA0834	o	322	
	AG0378	a	275	
	AG0378	b	281	
	AG0378	c	290	
	AG0378	d	293	yes
	AG0378	e	284	
	AG0378	f	312	
	AG0378	g	300	
	AG0378	h	295	
	BG1123	a	202	
	BG1123	b	216	
	BG1123	c	220	yes
	BG1123	d	252	
	BG1123	e	254	
	BG1123	f	256	
	BG1123	g	222	
	BG1123	h	226	
	BG1123	i	234	
	BG1123	j	258	
	BG1123	k	218	
	BG1123	l	224	
	BG1123	m	264	
	BG1123	n	248	
	AG0323	a	210	
	AG0323	b	216	yes
	AG0323	c	219	
	AG0323	d	231	
	AG0323	e	213	
	AG0323	f	222	
	AG0323	g	201	
	AG0323	h	228	
	AG0323	i	195	
	BG0295	a	286	yes
	BG0295	b	288	
	BG0295	c	283	
	AG0441	a	292	yes
	AG0441	b	301	
	AG0441	c	283	
N10	BG0228	a	137	yes

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
	BG0228	b	143	
	BG0228	c	146	
	PE0355	a	217	
	PE0355	b	223	yes
	PE0355	c	225	
	PE0355	d	237	
	PE0355	e	229	
	AG0171	a	241	
	AG0171	b	245	yes
	AG0171	c	247	
	BG0651	a	230	
	BG0651	b	245	
	BG0651	c	253	
	BG0651	d	256	yes
	BG0651	e	269	
	BG0651	f	248	
	BG0651	g	251	
	BG0651	h	275	
	BG0651	i	272	
	BG0255	a	180	
	BG0255	b	183	
	BG0255	c	186	yes
	AG0047	a	293	yes
	AG0047	b	311	
	AG0047	c	319	
	UB0196	a	281	
	UB0196	b	287	
	UB0196	c	293	
	UB0196	d	295	yes
	UB0196	e	283	
	UB0196	f	275	
	UB0196	g	271	
	UB0196	h	277	
	UB0196	i	265	
	UB0196	j	289	
	UB0015	a	241	
	UB0015	b	243	
	UB0015	c	253	yes
	UB0015	e	251	
	UB0015	f	255	
	PE0131	b	144	

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
	PE0131	c	148	yes
	PE0131	d	152	
	PE0131	e	150	
	PE0131	f	154	
N11	CA0120	a	138	
	CA0120	b	160	
	CA0120	c	172	yes
	CA0120	d	163	
	CA0120	e	169	
	BG0452	a	197	
	BG0452	b	209	yes
	BG0452	c	212	
	BG0452	d	215	
	BG0452	e	221	
	BG0452	f	194	
	BG0452	g	191	
	BG0031	a	225	
	BG0031	b	228	
	BG0031	c	237	yes
	CA1035	a	255	
	CA1035	b	258	
	CA1035	c	282	
	CA1035	d	294	
	CA1035	e	297	yes
	CA1035	f	285	
	CA1035	g	306	
	CA1035	h	300	
	CA0546	a	110	
	CA0546	b	120	
	CA0546	c	123	
	CA0546	d	146	yes
	CA0546	e	149	
	CA0546	f	126	
	CA0546	g	144	
	CA0546	h	112	
	CA1032	a	203	yes
	CA1032	b	211	
	BG1149	a	260	yes
	BG1149	b	266	
	BG1149	c	263	
	BG1230	a	252	
	BG1230	b	288	yes
	BG1513	a	164	yes

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
	BG1513	b	214	
	BG1513	c	216	
	CA0328	a	237	
	CA0328	b	240	
	CA0328	c	252	yes
	CA0328	d	255	
	CA0328	e	258	
	CA0328	f	234	
	CA0328	g	264	
	PE0324	a	258	
	PE0324	b	270	yes
	PE0283	a	149	yes
	PE0283	b	167	
	PE0283	c	173	
	PE0283	d	176	
	CA0163	a	311	
	CA0163	b	317	yes
N12	BG1321	a	197	yes
	BG1321	b	200	
	BG1321	c	330	
	PE0133	a	131	
	PE0133	b	141	
	PE0133	c	147	yes
	PE0133	d	153	
	PE0133	e	133	
	CA0456	a	179	
	CA0456	b	185	yes
	PE0063	a	114	yes
	PE0063	b	126	
	CA1027	a	297	
	CA1027	b	300	yes
	CA1027	c	303	
	CA1027	d	306	
	BG0864	a	169	yes
	BG0864	b	175	
	BG0864	c	185	
	BG0864	d	191	
	BG0864	e	193	
	BG0864	f	197	
	BG0864	g	195	
	BG0864	h	187	
	CA1090	a	282	
	CA1090	b	288	yes

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
	CA1090	c	292	
	CA1090	d	295	
	CA1090	e	279	
	CA0991	a	162	
	CA0991	b	165	yes
N13	CA0603	a	179	
	CA0603	b	182	yes
	BG1288	a	199	
	BG1288	b	205	yes
	CA0488	a	202	
	CA0488	b	217	
	CA0488	c	237	
	CA0488	d	231	yes
	CA0488	e	234	
	CA0488	f	240	
	PE0012	a	115	
	PE0012	b	134	yes
	PE0340	c	279	yes
	PE0340	d	281	
	PE0340	e	277	
	PE0340	f	255	
	PE0340	g	257	
	BG0516	a	165	yes
	BG0516	b	170	
	BG0516	c	179	
	BG0516	d	154	
	BG0516	e	182	
	BG0516	f	173	
	BG0516	g	158	
	AG0504	a	320	
	AG0504	b	332	yes
	AG0504	c	338	
	AG0148	a	268	
	AG0148	b	272	
	AG0148	c	280	yes
	AG0148	d	286	
	CA0736	a	324	yes
	CA0736	b	474	
N15	PE0286	a	186	
	PE0286	b	194	yes
	PE0091	a	152	
	PE0091	b	164	
	PE0091	c	176	

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
	PE0091	d	180	yes
	PE0187	a	176	
	PE0187	b	178	
	PE0187	c	180	
	PE0187	f	182	yes
	PE0187	i	184	
	CA0719	a	300	
	CA0719	b	304	yes
	AG0369	a	180	
	AG0369	b	184	yes
	AG0369	c	186	
N18	BG0278	a	239	
	BG0278	b	241	yes
	CA0739	a	220	
	CA0739	b	222	
	CA0739	c	232	
	CA0739	d	234	yes
	CA0739	e	240	
	CA0739	f	224	
	CA0739	g	236	
	CA0739	h	242	
	UB0315	a	127	
	UB0315	b	133	
	UB0315	c	131	yes
	BG0278	a	239	
	BG0278	b	341	yes
	CA0636	a	257	
	CA0636	b	263	yes
N19	CA1107	a	225	
	CA1107	b	228	yes
	CA1107	c	328	
	CA1107	d	216	
	CA0552	a	192	
	CA0552	b	195	yes
	CA0552	c	204	
	CA0552	d	207	
	CA1066	a	196	
	CA1066	b	217	
	CA1066	c	232	
	CA1066	d	235	
	CA1066	e	238	
	CA1066	f	241	yes
	CA1066	g	244	

Linkage Group	SSR Marker	Allele Name	Allele Size (bp)	Favorable Allele for SCL Resistance
	BG1241	d	329	yes
	BG1241	e	370	
	CA0221	a	253	yes
	CA0221	b	265	
	CA0221	c	271	

Example 10: Introgressing *Sclerotinia* Resistance from Spring *Brassica napus* to Winter *Brassica napus.*

The *Sclerotinia* resistant source 06DSB13911 is crossed to winter canola lines in bi-parental, 3-way or complex crosses. The F1 cross is then backcrossed to an elite susceptible parent. At the BC1F1 generation, approximately 800-1000 progeny (minimum number required to identify at least one individual with all favorable alleles present) are generated for each cross and submitted for marker analysis using the markers identified as being associated with *Sclerotinia* in spring canola. Each individual sample is examined for the presence of the favorable alleles from the *Sclerotinia* resistant line 06DSB13911. The percentage of favorable alleles present in each sample is calculated and the top three from each population are used to cross back to the recurrent parent. This process is repeated again at the BC2 stage. In addition, selections are intermated to develop populations in which individuals can be identified with homozygous desirable *Sclerotinia* alleles.

Example 11: Use of *Sclerotinia* Resistant Lines for Hybrid Seed Production

Sclerotinia resistant lines with scores of 5 and higher for SSDIS are selected for use in hybrid seed production and hybrid testing. Production of these seeds can be done according to methods known to the skilled person and are described, for example, in WO 2006/135717.

Example 12: Marker Sequences Containing Polymorphisms, and Exemplary Primers

Set forth below is sequence information for markers of QTLs significantly associated with *Sclerotinia* whole plant field resistance at $P \leq 0.01$, as set forth in the foregoing examples. In the sequences, n = an unknown nucleotide; underlined sequences indicate the primer sequences from Table 14 below and sequences in brackets indicate polymorphic regions (SSRs, SNPs).

AG0023 (SEQ ID NO:1)

CGAATTCGCCCTTCTCTTGCTTAGATCTGGACTAACTACTTCnnAAAGAAAACATTnnnTTAATGTTTAT
GTCGAATGTCATTTATGCTGAACAAAATAACCTTGAAAATATGTTCTGTAGGCTAAAGTTGGGAGAGAGA
AGGAGGTTGAAGAGATTTTGTCAAGATTGCGAGGAGAAAATTCTGATGTATCAGATGAGGCAGGAGAGAT
ATTAGTAAGCATATATATGCATGAATAATCATATGATCAATGTATATATTTTTTACTTCACAATATTTTG
ATGATCATCAGGCATATACAGAACATGTTAAACAACAAGGAGATGATCGCGGTTTCCTCAAGTTGTTTCA
GCGAAAATACGCGTTCTCACTTACTGTAATT[CTTCTTCTTCTTCTTCTTCTTCTTTTTCTTCTTCTTCT
TCTTCTT]TAATAACCCGTTTGGTTTACACAGATTGGAGTTGTTCTTATAGCTTTGCCTCAACTTGGAGG
TCTTAGTGGTTATTCTTTTTACACTGAGTCCATTTTCATATCTACAGGTAnnnTAACTCTTACTTCTTCA
ACAAAATCTTGATTTTTATATATTTATTTACCGTAACGATAATTGTTGATAATTACGnnnATCAGGTGTA
TCGAGTGATGTTGGATTCATATCGACATCTATAGTTC

AG0045 (SEQ ID NO:2)

ACGAATTCGCCCTTCTCTTGCTTAGATCTGGACTAnnnnnTGATTTGCCCGCTATGTTCGACGGGTGGAG
ATTTTAGTTTTACTTCCTCGATCTGATTGTATGGGTTGGGAGTAGGGTCTAATATATCAACTGCGAGTGT
ATGTTCGTTTCCTCCTCAGTTTCGAAGTTGGGTTCTTATGTGTTTAGCCTAAGnnnCTGTGAnnnGnTAG
TTTTTTTTTAATCAGTTCCAACAGGATTCATTTCAGGnnnTTGGAACTTGTGTATATGTGTTAGCCTGAG
ATCTCTGTAGTGTCCGGAAATGATATTTnnnnATTATCATTAATTTAGTTCGAAGnATGAAGCTCAGTGT
TGTTGGACTTGTGTATATGGAGCTCGAAGAGTGAAGCTCAGTGCGTTTTCATCTGAGGATGATGATGATG
GAGCTAATGTGCTGAGCAATGAGAACTCGAGATGATAAGGCTTGAGGGACATGCCAGTGAGT[GAAGAAA
]CCGTCGGGCTATAGCTTAGT[GAAGAAGAAGAAGAA]GAGCTCGTGGAGTGATCAAATTTGCAGGTATG
CCCAAACTTGCCAATCCCACATTGTGGAGAATGGCTGCATTTTTACCACAAAGCTGTTTCTGTGGAGCCA
AAAATGAATGGAGGATAGTAAAACAGAACGTCATAATCAAATCAAGAAATTTTAACTTTTTTTGTCAGCA
CAAATTTnnnCTTTATCTTTAATTATTTAC

AG0047 (SEQ ID NO:3)

CGAATTCGCCCTTCTCTTGCTTAnnnnCTGGACTAACAACAATTCCAAAATACTAATTCACAAACTTTGT
TTACAATCCAAAGAAAATCCGCTCTTTTGAAGCGCGGATCAAGATCTAGTGTTATAATATATCTAGAACA
TGGGAGTTTGGTCCAATGAACTACTGTATAGTTCTATCGAAATTTTTGAGTGATAAGATTGAAGCTCCAG
CACTCACTTATCTATTTGAGAAGCAAATAATAGAAAAAGAAGTAGATTTGAGGAAGAGATGATGGAGTTG
AACAAGGAGCTTTAAGATTTGAGTTCTGACAGTGTAGAAGCTGCAATACTGAGGCACCAAGGAAGAAAT[
CATCATCATCATCATCATCATCATCAT]CAAGAATTAGTTTCAGTTCATATCCACAACCATTTTTTCTTT
CAAAGAAATTTGCTGGTAGTAATTTTGAAGTTGTAAATTTTACATTTTCAGTGTTTCATTTTTCTCACGT
TTTCTTAATAATTGTTTACTTGCCAAATGATTCCATCACTTGGAAACTCACTATTGTTTGACATTTTGGT
GTGCTTAAGTGACTCTTTTCGAGTATTCATACATTATAGAAATTGTTTGGGACAACAGGTAAGAATTGCT
TGGCACAAGTAATGGCATCCCTCCCTGCAAATATATATAAATATTACAGTTGTCCTGGAACTTTTnnnnT
CTATCCTCTGCTGACAGGATGAGATATATGCATATAGAATATTAACTTCnnTCnGCCCGTATGTTCATGG
ATGnnnAGCTCCAT

AG0070 (SEQ ID NO:4)

CGAGGAGTTGAATGACCCTGACTGTACTTTGGCCTCGAGACAGTCCCATCAAGAATAATTTACTGGGTCG
ATTTTTATTTTTAATTCTGGTCGAGCCAACTCCGAACTGGTCGAGCGGGATTTTTTAATTCTGGTCGACC
AAAATCATATCCGCTCGTGAGGGTCTTTACAACCACCATCACCACACTCGGACGATCACCCCACCACCAC
TTGGACGA[CCACCACCACCACCA]CCGGCGGCTCGGCTAGCTCTCGGGGGGCTCGCGGCGAG[GAGAGA
]GGAAGATATCCNACGGAAAGAGAAAAGAGAGG[GAGAGAGAGAGAGA]GGCGTGAGAGAAGAAGAGAGA
AAAGGAAAAGAGAAGCTTGACGGCTAGGGTTTCCTAGTCTCTATAAATTCCTGCAGAGCTTCACTCAAGT
TTCAGAATGAGAGAAGAGTAAGAGGAGGCAGCTTCATTTATAGAAACAGGAGGAAACCCTAGGTCATTTA
CCCTAATGGGCTGCAGTCTTAATGGGCTCTCCTTAAGAAAATTTTGGGCTAGGAACCGGGACGTTACAAT
AATGCTTCTTATGAATATGTCTGAGTAGTTCTTTTGTTAGATTTAGGGTTCTTCAAGGGGTGAATTATGG
TTTGCTANATTTATATTGTTGTTTGTGTGATTT

AG0086 (SEQ ID NO:5)

GCTCGCCGACTTCGGAGTCGCCTCGCCGCTCGATATCCCTTTCAGCCTCGCCTCCATCTCTTTTCCCCAA
ACTCTAGGCTGTTGCTGTTGCGTCGCCGCCGCCGCCGCCGTGGCTGTTATCGAGCTATTTGATCTACCGT
ACAGCATTTTAAACCGTTGATCAGATTCGGGATCAGACTTTGTCGTCACCGGAGGGCTCTTGATCGGCGG
TTGCACTTCCCCCTCCGTACACGGCGTACAATGTCGGTAAGCACCGGAAGCTTTCAGAGCCATATCTTTG
AGCTGAGAATGAATTTACGAAAATACCCTTGATCAGTATAGAGAATGACAAGAGGTGGAGGATGAGCAAA
[GAGAGAGAGAGAGAGAGAGA]AGTCTACCTGAGATGTTAGAGATTTGGCTTGCTTGGAATCCGGATCGT
CGGGTTGACCCGAGGTTTCATCGCCGGCTCGCTTCGAACGAGCTATACAAGTCAGCATTTTCCGGCAGCT
GCTGTTTCTTGGTAATGTGATTTTGTTTCTTCTCTTTTTGGATACG[GAGAGA]CAGTAGATGCTGTCAG
TTTCTAACTTTGGTTTGTGTTGTGTGTTTGGTCATGGTGCTCTTTTTATGTTTTATACTCACTTTACCAN
NGAAAACGGTTCCATTTTTTTAA

AG0093 (SEQ ID NO:6

TAGGAGATGAGATGTACTGTTGCTTAGGGCTCTTATTTCTCTTGAAACTAGAATAAGCTGCCATCGGGTC
GGTGTAATAATCAAACCTTGGCTTATCATACGATTCCTGCTGATGTATGGATGCTTCAGCCAAGGGATTT
GAGAGGTGACTTGTGTTCATAGAGGTTCCAAGCTCTGTAGAACCATCATTCTCTG[CAGCAGCAGCAGCA
GCAG]CTTCCATCCGCATTGCTTTTAGCATTT[CTTTTCTTTT]CTCTGAATCTTCCATTACTGCTCAGC
TTCAAAGCTAATCAACTACAAAAATATAAACTTTTTTTCGAAATTATCAATCGAATCGCACCAAAAGAGC
TAAGATCTCCACGCGAGAACAATCTAACTAACCCTAAACCCCCAAATTATCCCAAACTCTGTACGGATAC
TCAAATTGGAAAAGCGAAATTGAGAGGATGCTAACCTTGGTTTACTCAACTTCTTCACTTCCTGGTCGCC
AGAGGTAGAGGATGAATGACAAGTGAAAACCCAGAACACGATGATGACGACAACNAAGCCTCCACAAATA
AATATAANACCCGGTTCGTGTTCGACCGTGTTTTTCCNATTAAAACCGGTTTACGGCGATNAGAATCATA
AACCAAATACGATNATCACGAAGGGTGACGATTAANACGAGACTTCCCAAAACCGGTTCGT

AG0125 (SEQ ID NO:7)

CTGTTGAGGGGAGGAAACAAGAGCCTGGGAGGAGAACTCCTTGCTGGGGAAGACGAGATCATTCTCCTCA
GAGGCAATGGATTCACCTAAGACCACAGCGTTTAACTGAGAGATCTTGCCGGCACCTGAGGGTAGCAGAG
ACATGGACTCCACATGCCTTAGGTGATTGGATGACATTGTCTTACACCGGAGAAGTTTGTCCAACGGAGA
TGATCTGCCACACCCTTACAAGTCAGATGTCATTTGAATAAAATTTAAAACAAAACCACAAATGTCTTTT
TGACTTATTTATCAAAACTGCCTAAACCCCAAACCCAAT[CATCATCATCATCATCATCAT]AACCATAT
TCATCAATCATTCTATCATTATTGTCATCATTGGATCAGATTATTCATTCACCTTTGAGAAGCCGGAAGA
ATCCGAGATCCAAGTGATCCGCTTGTTTTCAGATCCTGAAGAAACAAAAACAGATCANAGGCGAATATTC
TTTTTTGATTACNATCAGATCATAAGAAGAAGAANAGATTGAAACTTTCGTANACCCAAAACATATCATT
ATGACNAAAGATCACATCTTTAACTCCNATGATCCCTAAGATTCGACTTACAGGTCGAGAACGAAGAGAG
GAAATTTTTTGAAAAATTGTAAGAAGGGGCG

AG0148 (SEQ ID NO:8)

CTTGAGAGAGAGATTGAAAGTGAGCTGCGCCAAGAACCAAAACGGACATATCACAAGGCTTACTTTAGCA
ACACGCATCCTTGTCCAACGTCCCTTCTCATCTAAACCTTTAGCACAAATCTCACCACTAACATCAAGCT
CTACACTTTCCACACCAGATTCAATCCCACTACCATCTCCCTCCTCTTCTCTAACAACAACCCCTGATCC
TTCAGACACATCTGACA[CCACCACCACACCACCTCCTCCT]AGCAAAACATCCTTCTGTGACTCAGCCG
ACTCCTGCAAATGAGACCTTTTGCTTCTCCGAAACAGTATACTCCCATACGCATCTGCTAAATACCGACC
GATCTCGAAGAGAGGAAGGTTAGTCCAGATCTAAGCAAGAGAAGGGCGAATTCGCGGCCGCTAAATTCAA
TTCGCCCTATAGTGAGTCGTATTACAATTCACTGGACCGTCGTTTTACAACGACATGACTGGGAAAACCC
TGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCATCTGGCGTAATAGCTAACAGGCC
CGGACCGATCGCCCTTCCCAACAGTTGCGCAGCCTATACGTACGGCGGATTAAGGTTTACACCTATCAGA
GAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATATTATTGACACGCCGGGG

AG0171 (SEQ ID NO:9)

CACTTTGCATAAATACTTTTACGAGCAATTTTAAAAAAAAATTCTAAAAATGTCTATTATTTGTGGACTT
GGAATAGCCGTTTATCTGCTTTGATTTTGTCGTTTCTTAAATCAAAGTCCTAATCAGGGTCCTTATATAA
GATCAGTCCTCTAGAATCTGAATAGCTTTTTAAGACAAAAAAAAAACAAAACAGATTAGAGTCCGAATCG
GACTCGAACATCTCCAATTTAACTTCTATCTTTTTTTTTTCTAAAATAAAATGTAAAATAAAAATATTTT
AATTGTATGAAAAATTGCATTCAATAGCTAAAAAAAAATAAAAATTCAATACATAACTAAAATCCCACTT
T[CTCTCCTCTTTTCTCTTCCTATCTCTCTCTTCTCTCTCTCT]AAAAATCTAATTTTTCTTTTTTTTTC
TGGTTATTCCCTAAATAAGCCCTAATTGTATTCTATTTTCACTCTAAAAAATAGCTTGATTTTATAAATA
GAATAATTCATTTGTTTTTTTAAAAATAAATTATCATTAGAATATAATTTAACTTTATTATAAAATTATT
CTCTTTTAGAGCAAAAAAATAAAATAAACCATTAGAAATTGTTTTAGAGAAATCATCAGTCAAAATCTCA
CNNNTTCCACTATTTAGTTTCATCTCTNNNGCAATCAGACGTAAGAACACAAAAACATATGTTAT

AG0203 (SEQ ID NO:10)

CAGCCATTCTCTATGGCCTTGGTGACCATGGAGATCAAGTGCCTTAACAGCAACCGGTTGGGCTTCTAAA
CCCGGTTTAACCTTGTCATCAATGAACCCTTTGTAAACTGGCCCAAACCCTCCTTCTCCCAGCATGTTAC
TTCTTGAGAAATTATGCGTAATAACTCTCAGCTCAGACAAGGTGAACATACGAAGCTTTTGAGATGTGGA
GGAGTTTGAGAGGTCATCCATGACCGACATGGGCGAGCTTGGATCACTTATGTCCGATAACGACAGCCTC
TTGATCACCGGACAAGTTCTTATTTTCATTGCGTTGCCCCTCTCCTCTACTTCGTATCTACTCGCGTTCT
TTGTCCTGTAACATCCTAAAAACAGAGATGTCAATGATGT[CTTCTTGTTCTT]GGTTACTGCCATTTT[
CTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT]CAACTTTTTGGAGAA
AATGGAAGATAGATAGTGTTTTACTCTTTTTGATGTATCTTTTGTAGATTGCGTNNNNNTNNNAATTAAG
GGTGTTTTAGTTAACCTACTTGTCTTCAAAGTCCTGTATTTATAGNGGTTGTTGTGTCTTATTCGTAGTG
ACTACCTTGGAACTTTCTNNNGACATATACTGTN

AG0239 (SEQ ID NO:11)

AAGCAACCTACGGTGTTCTCTTATTACCGAGTGCTATTAGGCTTGTCTTAATGTCCTAAACAATTACGTT
TTTACTTTGAATGTGAACCTGCTCTTGAACGAAATATCTTCATGACTTCATATCATTTGTTTAATTCTAC
TACTGTCTGGTTAAAAATGGTTGTATGTTGCATATATTGCTATTTTAACCACATGTAAGAAAAGAGAAAC
GATCCCACGTTTATACTCAAATTCAAGAATCCAACACAAATCCTCTCACAATTCTATTAGATTGGAAGAA
AGAAACTCATAAAATAATTATAAATAACAAAAGGACAAGTAAATAGAGTCTATCTGGAACATAAAAACAC
TAACGGGTCTTGTGGGGTTTACAGA[TTTTCTTCTTTCTCTTCTTGACAGATTTGGGTGTCTTGTTCTTC
CTAGCCTCTCTCTTCTCTTCTTTTACTTTCTTCTTGAGCTCTTTCCTTGCTGCTTTCTTGTCTT]CANGA
CCCAACTCCTCTTCACTCTCACTTCCACTTTCCTCTTCTTCNNNNTCACTTTCTTCTTCTCCTTCTCCTT
TTTCTTCCTCTTTTCCCTCTTCAACATCAACCTTTACTNNCTGCTCATCTACTTTAGGANNNCATCCTTA
NNN

AG0243 (SEQ ID NO:12)

TTTAGAATCAACACAAGACATAGNCAANTTGATTTTGGCTACTGGGTATTAATAATCCCCAAAATCACAT
CCTTCAAACCGAACGAACAATCAACAAGAAAATAAGAGCTACGATTCAGAAAAAGCCTATGATCAAAACG
CCTAAAATAACNNNAAAAAAANNNGGAGCAATTTTAAACAAATGGAATTGAATAGTATGAGATGAGATGA
GAAACAAAAGAGAAAGCAGTGTGCATAGATCATCAGGGAAGCTAACCTGAAATGATTTGTTGATTGGGGG
ATGAGATTGTTGGAAGGGGACAGAGA[GAAGAAGAA]GGTTTGCTTGAAGACTCGAAAATTAAAGCTTGT
TAAG[GAAGAAGAA]GAG[GAAGAAGAA]GAGGATATAAATTGACATGGACCTATTAAATGCCCATTTTG
TTCTGNTTATTTACTTAAGATTGCCACTATGACCTTTGACTTTTGGACGGCGNNTGTAGCTAAGCTACTG
TTTCTTCATTAATCACGCTTGCCATGATTAGTTTTTTTTTTCCTCCTATAGNNTTCATANNTAGCCCGAA
ATTACTGACTTTTATGAGATAAAGATCGTATTTTTTTATTTCTTANNGTTTAATACCCT

AG0272 (SEQ ID NO:13)
CTACTTCCAAAAGAAAACATTAAATTAATGTTTATGTCGAATGTCATTTATACTGAACAAAATAACCTTG
AAAATATGTTCTGTACGCTAAAGTTGGGAGAGAGAAGGAGGTTGAAGAGATTTTGTCAAGATTGCGAGGA
GAAAATTCTGATGTATCAGATGAAGCAGGAGAGATATTAGTAAGCATATATATGCATGAATAATCATATG
ATCAATGTATATATTTTTTACTTCATAATATTTTGATGATCATCACGCATATACAGAACATGTTAAACAA
CAACGAGATGATCGCGGTTTCCTCAAGTTGTTTCAGCGAAAATACACGTTCTCACTTACTGTAATT[CTT
CTTCTTCTTCTTCTTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT]TAATAACCCNNNTGGTTTACACAG
ATTGGAGTTGTTCTTATAGCTTTGCCTCAACTTGGAGGTCTTAGTGNGTATTCTTTTTACACTGAGTCCA
TTTTCATATCTACAGGTAAAATAATTCTTCTTCTTNNNCAAAATATTGATTTTTATATATTTATTTACCT
TAACGATAATTGTTGATAATTACNNNNATCACGTGTATCGAGTGATGTTGGATTCATATCGACATCTATA
GTTCNNNNNTTACCGATTTCGAGTGACCTTGTTTAGAGTTC

AG0304 (SEQ ID NO:14)
ACAAACATAATTGCAATTAAACGGATAGTAAGGGTCACAGATCACACAATACTGCATCGAAGTTTTGTCA
TCAACACAAGTGCGCATCGTTTCATTCTTTTCTTTCTTCCGGCTTACCTGAGCCCGGCCGTGGCACAATC
TTCTTCAACAGACAGCGTTTAAATAAAACTTAACTTGGTAGGGCTGAGGATTCAAGAATCATTTCTTGTA
ATTCACTGGCACATCGTCGTCATCTTCTTCAAGATCACTAAACGTTACATCTTCATCCTCATCCACAACA
TGTTTGGTTGCTACGGTGGAGCTAACAGTTCCTGCATTATCTTCATCCTTCAACCAATCATCAA[CATCA
TCATCATCATCATCAT]CGACTTGAACGTCAATAACTCTAGGTGAAGATCCAGTGACTGGCTTGTCATGG
GGCGCTGGTGCTGGTTTTTCTTCAATCACAGGCTTGTCAACAACTTGTATCTCCTTGCTCTCGATTGGGT
GCTTCTCCGTCTCAACCTCNNNTGTATCACTCAAATGTATTGTTTCCANNNAGATGNANTTGACTGNNNC
TGGTGAAGACACAGTAAGTGGTTCCTCATTGGCA

AG0323 (SEQ ID NO:15)
AATCATCCTAAATTTACCTAACCCATGGATTCAAAAGGTAACTAACTACTCGCTAATAGACATGATAAAC
CCAAAACCAATAGTGATTGAAGGTTAATCATGATTAGTGATCAAGATAATCCAATAAACACAAGACAAAG
ATGAAAAGAGGCTAAAGATTGAATCTTTCACCAAAATATGTTCATGTCTAGAGAAAACAAGATAGATCCT
AAGAATCTAACAATACTAAAAGCATGATAGTAAGCCCTCTAAGCGTGTCCACGTAAGTTAATATATTCAG
CTAATCAGAGATTACTAGCTATTTTGCCATGTCATAACAATTTTAAGTCGACCAATACAAAAACCGGGCA
AGGCGTCTGGGCCATTAGTAATATCCAGTGGCCAATACGAAACCCATCTCATTAATATCAAATCTCCAAT
GAAAGCCATTATCGTGGCGACTCTTCTTTTT[CATCATCATCATCATCCTCATCATCATCATCATCATCA
TCAT]CGCTTGGGATCACAACAATTTCCTGTTAGCACAACCCACTCTCCATCAATCAATCAGGGTCTTTA
CTACTCTTTTCATGCTTTCGTNTCAACTCCCTTTGTTTATCCTCCTATATAAATCATTGAATATCNNNAT
TTTGATCCAAG

AG0324 (SEQ ID NO:16)
TAGTCTCACCAACTCCAAACTTGTTAACATCTGAAGGAGTCTGAATATTCTCCTCCT[CATCATCATCAT
CATCATCATCATCATCATCAT]CACAATCAACAACTTCAACCTTCTTCTTATTGTTATCATCATCCTTTA

CTAATTTCTCCTCATCACAATCAATAACTTCAATGACTGAATCTTGAGAAACTGCTTCTTCTTCTTCTTC
CTCCAAATCGATAAACTCTTTATCTTTGGTAAGGAACCTGAAGGCATTCAAAGCCGATCTCTTGGCGTTA
TCATACTCGCGAGTGAAAAGCTCCCCCTTCTCGCAGCAATCGTTAGGCTTAGCCGTGGGTGATCCTCCGT
ATATAAGGTTCGCTCTGCTANNAGCGTGGATTGCTCGTCTAAGTGGAGTTTTGGCGTCGGGATACCTCGA
AATCCTCGAAGATGGCTTGTTGGGATCCTGAGACATGGTTCCGAAGGAGAATCTGCGTTTCTTCGGAGCT
TGGAAGTAGGGCGAATGTCAGCGGAATTGGGCGGTGGAAGGGTGGTTGGAGTATAAGGAATCTTCGCTGC
GCTTGCGATTGATGGCGACNGCGCTCATTTNNNGAGTCGATCACTGAACCCTANNGATTGGGAGATCGAC
GNNNGGAGAGGAACCATAAGAGTNGAGA

AG0328 (SEQ ID NO:17)

TTGCCAAAACATTTAACCAGGTGAGCACTTAAACTCTTGTCTGGCCCAAAAAAAAAAGAGTGAGACTGTA
TAGAGGATCAAGCCAACAGTAGATGAGGAAGAGGAAGGCCATGTAATCTCTAATCCACAACGATCTTGAT
TACCCATATAGGTCCATTGACTTTGAATCTTAATTTCAGAACATCAACAAATCTTCATCTTTACTAAAAT
TACAAAAAATCTTTTAACTTTTTAATTTTTGAAAAAAATACATATACACACATACAGCTAGTCTCTTACG
AAACACTACACAACTAGATAACTCCAAACATTTACAACTGAAAGTTTATCAGCTTGGAAAATCATCACTC
AGATTTCTTGTGGAACTTCACGGAGTCTATCAAGTGTATTAACAATCTCACTCAGACAAGCGATGAGCTC
GTCTCCTTGCTGTCTCACAAACTTTAACTTGTCCTCAATGCCAACATCATCCTCTGTTTCACTTCCATTT
GCTAGATTTCTCCCGACTATTGCATGTATCATGTCAGCTTTCTCACCTAGCTTTGCTTCCAACGCTTTAA
CCTCTTCTTCTATGCTCATCGTTT[CTTCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
CTT]CTACATCANNNGAAATTTCTACCACCGGCTTCTTCTTCACTTCCATGGTTTTATTCTTCTGATGGN
NNGCTTTTGCACG

AG0359 (SEQ ID NO:18)

CTCTTGCTTAGNTCTGGACTAACCCATATCCGAAAAGNTTCGAGGACCGACTCGAAAACCCAACCCGAAG
ATCTGCACGCCTAGGCCTAGTTTACAAGCAAGCCTACCAAAATATGTCAACTCGTTAAAAGCCTTTTAAC
CTGTCTGGTTCGGTGCACGGTTCAATTCCCGGTTTAGTTGTAACCGGTTTGTGATTGCTCAAAACCCTAG
TCGTCACCCTTTTTTATCATTATTGTGAACAAGTAGTCACCTCTACAAGTAAAACCTTAAACCCTATTGA
GCGAGTAGCAGAGCGCAGCAAGAAGAAACAAAACCAAAATATGAGACCACCACGTGGCGGCGGAAGCTTC
AGAGGAAGAGGAGGAAGAGATGGCAGCGGACGCGGAGGTGGCGGACGTTTTAATCGTGGAGGTGGCCGCT
TT[GGTGGTGGTGGTGGTGGT]GGCTGGCGTGACGAAGGACCTCCCGACCAAGTCGTNNNTTCGTTTTCT
CTCCTCTCTTGGTTTTCGCTCTCACTTTACAGCTCAAGCAGAAGTCTTTATTAACAAAAGTTGTCACCTT
TGACAAATTAGTCTTATCCCTTTGTTAGAGTCATCTTTAAGTTAAAGGTATAAACTTTGTGAAGTTATTC
GTTATGACAAAGTTTCTTTCTTTCGTTGGGTTATAACAGAAGTTGCAACGTTTGTTCATGCTNN

AG0369 (SEQ ID NO:19)

TTTTGAAGGAGGTTGCTTGGGTGATTTTTGATGAGATTCATTACATGAAGGATAGGGAGAGAGGGGTTGT
TTGGGAGGAGAGTATTATTTTCTTGCCGCCTGCTATTAAGATGGTTTTTCTTTCGGCCACGATGTCTAAT
GCTACTGAGTTTGCGGAGTGGATTTGCTATCTGCATAAGCAGCCGTGTCACGTGGTGTATACGGACTTTA
GGCCCACGCCTCTGCAGCATTATGCTTTTCCTATGGGTGGGAGTGGGCTGTACCTTGTAGTTGATGAGAA
TGAGCAGTTTAGAGAGGCTAATTTCATTAAGATGCATGATACTTTCCCAAAACCAAAATCTGAGGGGAAA
AAGAGTGCAAATGGCAAATCA[GGTGGTAGGGGCGGCGCTAAAGGTGGTGGCGGCGGC]GGTGGTGATTC
TGATGTTTACAAAATTGTAAA

AG0370 (SEQ ID NO:20)

TAAAAAATACCTTAAATATATAAAAAATCTATCTTTGTCGAACAAGTAAAAAATTTAAAACATCTTACTT
TTGGAAACGAGGGAATATGATATTTTGGAATGAATCAAAATTGACAATCACCTTGTATAGAACCCATAGG
TTCGTGGATTCGCGGTCCTCACCACTAATAGGTTGGACCTGGTTAAGAATCCTTGCACACAGAAATAAAC
TTAACCATTGCCGCCTTACATTGTATTCCAAATTTGTTAATTACCGGCCAACAACACAATTATGTTATCT
CCATTATTACAACCACCCGCCGCAATAATTATCTCAATCA[GTTTTGTTTTGTTTT]TATTTATATTCAA

ACGGAATATCGTTATTTAATTATTAGAGCTTTAAATAATCTATATAAAGTCCATACAATTTTTGTTTATG
GAAATAGACACTACAAACGCGTATTTTCAGTTTTTTTTTCATATGGAGAGAACTACCAATCAGTTGAAAG
AAAAAAGAGAACTACCAATCTACCATATAATATAAAAACAATAATAGTATTAAAAAAAGGAGAGGGCAAC
GGAACGGGACGGAAGAGAATGGAAATGGTTACGTTTATAATAGCAATGATCTGTTGAACAGCTTATGACA
CCTCACTCTGCGCTTGCTTCCATTCCTCATCTCTCTCTCTCTCTNNNAACTCTCTACGAAACCCTACCTT
CTTCN

AG0378 (SEQ ID NO:21)
TACAAGACAGACGGACATAATAAGAAGAAATGCAAATAGACAAGTCCTGAGAGAGAGTATCCAAAATAGA
GATTTTAAAATCTGTCACTAACTTTTAGGCATAGCTTGTTTAGTTCGCTTAGTCCCTCCTTTCTTCACAA
AACCAAAACCAAAAAAAAAAAGATGAGAGAGAGCAGTTTGTTGATAAAGAAGCAAATGAAATGTAACTTA
CTTTTCTACCGGCGGTGGTGGTTGCTGGTGGAGTTGCTGCTGCAATTGAGTAACGTCACAACAAAAAGGA
AGATGGAAATGAAAAAAAGGAGGATTCAATGGATATCAACAAAAACGTGTTAGAAAGACTCACCACTCTG
TTGAGCTTCACCACCTTGGCTGCCTCCTCTTCTCGTCGGGGTCTACACATTTCAATAGATTCGTCAACAA
CAGTAACGAGATTGCGTCGAAACCTAACTCAGAAAAAAAAAAGAGATGACGTACCGCACTCGGACCTG[CC
GCCGCCGCCGCCG]AAGATGAAGCAGTCGCATCTACAAATTTCAGATGCCAAATTAGGGTTTAACCTAGA
AAATAAAAATATCAATAAGGCAAAGAGAGAGAGAGAGAGTACTAGTTGGATTGCGATCT

AG0391 (SEQ ID NO:22)
GTCTTCTCACGGCCGTACGGATAACGCCGTGAAAAACCACTGGAACTCGACGCTCAAGAGGAAATGCTTA
GGCGGCGGTGGAGATGGTAATCTCATTGTGATGAGGAC[GGAGGAGGAGGAGGA]TCAGGATCGGCGGAA
GAAGAGGAGATCGGTGAGCTCTGAGTCTGCTACTCCGGTGGACACTGGGTTGTACATGAGCCCGGAGAGT
CCCACCGGAATCCCATCTCCGCCGTCTCCGGTTGATGCTCAGCTTTTAAAACCAATGGCGATGCCGTCAC
CGGTGGAAATGTCTTCGGTGGAGGAGGATCCGACAGCGTCATTGAGCCTGTCACTGTCACTTCCTGGTCC
TGATGTCAGACAGGAGTTGAAGAACGCGGGTTCGAAACACAACTCGTTGCTGTTTCCCCGGTTTGGGAGT
CAAATGAAAATTAATGTTGAGGAGAGAGGAGAAGCACGTGTTGGACATAAAGCTGAGTTTTTGACGGTGG
TGCAAGAGATGATTAAGGTGGAAGTGAGGAGTTATATGGCGGAGATGCAAAAAAATAGCGGTGGTGGCGG
TGGTGAATTCATCGTCAGTGGTTTTTATGATGCCGGCAACGGCGGTTTCAGGGATAGTGGG

AG0410 (SEQ ID NO:23)
AAAAGGAAAGTTTAAGACTTTAAGCTTTACCTGATGATCCATATCGGGGAAAATCGCGGCGAGGTGATCG
AGGAGAAGC[GAGGAGGAGGCAGGAGGAGGAGGA]GGGGGAAAACGCGGCGGAGAAGAGGATGAGAAGTA
ACGGAGTTTCTTGGAGACGGGAGGGGAAGAAGCGGCGGACAAATCCTCGAACAGAGATCTCTTGCTACCG
CAAACAATCGCAGACATGTTATCTGCTTCCACCTTCTTTTTCTTCTTCTTTCTTCCTTCCTTCAGATCTC
AACCTTTCCTTTTTGTTTGGTTTTTTTTTTTCCTTTTT]CCTCTAATCCATCTCTGATCTGTTTCTGTCG
GAAACCAAGCAAAAAAAAGTCAAAACACATCGGATCTTCTTCCGCATCTAAATAGATCCAACAACCCGGA
CTCGGATTCAAAT

AG0441 (SEQ ID NO:24)
AAAAGACGTTGACTTGATTTGATATGCCATAGAGCTAAACCCTAATTGAATTTCAATCAATTAGGGGTAA
AGATCTCTCAATCTACGAACAAAAGATCTAATATTTACAGTCAAAATCTACGGAAAACACAAAGAAAAGG
CTTACGGCGACTGGGCTGCGAGGAGCGCGATTCTGATTACGGATCGGCATCTTCCGATTCTGGATCCGAG
CTAGGCGATACGAAAGATGATTTCTTCCGGAAACCTCCGATTGAGTAACAAGAATCTCGACAGAAGTTGT
TTCTTCTCAGTTAGAGAGAAGAGATTAGGCTTCGGCCCTTTTTGTGATTTTGAGAAGGAT[GAGAGAGAG
AGAGA]GT[GGAGGAGGAAGGAGGAGGAGGAGGAGGA]GCCTTTGTTATTTTGAAAGTTTGAAAATA
GATCTTGAGAATAATTGTAACGTTACTCTTGGTCCTCTATATGCTTATTTATTTATTCCACTAATACTTT
ATAAGGTATATGGGCTTTATATGGACTATAATCTCGGCCCATCTATGTTAAACTAATCCGTAATTTTCTT
GGTTTTTTTTAAACTTGCGCGCTCCTTAATTTGAAT

AG0477 (SEQ ID NO:25)

ACTTTTTATAACCGACACTTAAATCAAAACTTGAAAAATAGCATCAATTAGATTTGTAACGGAGTATCAT
CAATCATCAAGAAACAACAATCTTGTAGGTGAGTAAATAAAAGATACCGTGAATAATGTCAACAATCGTA
ATCTCATACCACTAATACGTAATTAAAGAAAATAATCATATAATTAGGGAGATAATGTTGGGAATCTTAA
TCGTATAATCAGAAGCGTATTCATTTCATTACAAATTGATTCTCTTGTCATTTGTTATAT[AATAATAAT
]AAAAAAAACGTTAAATCAATTCAAACTAAACCTT[CTCTCTCTCTCTCTCTCT]TTCTATTTCGCTCAT
CATCATTTTATCTGATGAATACGCCCAATTGAAATCCTTTCCTTATCAACTCAAATTGAGTTTTCAAAAT
TATTCAATTTTCGGATCTCCGTAGATTTGCTCGGCGGAGGAGGAGGAAGGATGGCTCAGTTGGCGGCGGC
GGCGGGGAGGAGAATAGGGGATTACGCGGTGGGAAGACAAATCGGGTCGGGTTCGTTTTCGGTGGTGTGG
GAAGGGAGGCATCTGGGAGATGGAAACGTGGTTGTAATCAAGGAGATAGCCATGGCGAGGCTTAGTAAGA
AGTTGCAAGATAGTCTCATGTCCGAGATTATCATCTTGAGGAA

AG0482 (SEQ ID NO:26)

ACCCAAACGAATTGCTCTGTCCGTAGAAAGAACAGGCTCGGGAGCTGAGT[GGTGGTGGTGGTGGT]GGA
GAAGCGACGGTGGACCATCCGGGAACGAGTGCAGCGAGAGACGGAGATCTTGACTCGGAGGAGCTTCCGT
CGAGGAGCCAACCACCGGGAAAAACGACTCCGAATCCATCGACGGCGGAAGAAAACTCGGAAGCTCCGCC
ACTCCGTCGAATCCACCGGACCGTGCACCACCGAACTGAAGCTGTTCCCGCTG[CGGCGGCGGCGG]CGA
CGGAGATATTTTAGTTTTGGCGGCGGTTCTTCTCGGTTTAGCGTTTGCGGCGGCGTTGCGAACAGCGTCG
GCGGGACTCCACGGAGGAAGCTGAGCGAGCTCGTCGATGGAAGTCTGAGCCTTTCTGATCAGCCAGTCAA
CGGCTTTGCTCGGTCGGTCGAAGCCAAGGCGGTCTTGAACGTCGTAGAACTGAATCGCCGTGTGAGCCGA
TAGCCTCACGCGCCGGTCACGTGGCCCTTTGGCCGTGCAGACTTTGCTGTGCCGGTCTTTTCTCCCCGTC
GACCGCACAATGTGACCTCCTTGCACCTCCACTATCTCGTCTGACGCAGCGCGGTGCCTCATTGAAGAAG
GCNGNNNGGGTTGAGGGGTGGAGGAAGTGGTGAGCTTCGTCGTGGTCGTCTGCCATTGGTTGAGCATAC

AG0504 (SEQ ID NO:27)

ACAACTTTGAAGTGTGAATAGAGTAAAAGATTCAATCTTTCATATCAAAAGACTAACCTAGACTCGAACT
CACGGATCTCAGCAAGTTCTTTCCCCATCAAATCCACCCACTGCACCTTCTTCTTCTTCTCCCTTCTCTC
ACCATCCTCTGAATCTAAAGTTTCCTTTTTTAGACTACTCTTCAGGATCTCTCCATTACTTTGACCTTCC
TCTAAGGCACAATCTTCTTCCTTTCCCTCTTCTTCTGTTCCCTCATTCACCAAACTATCAACGTGATCAA
CAACTTCCTCTACTTGTGCATCAACATACTGAT[CATCATCATCACCGTCAT]CAACAACCGAAGAGACT
TGA[GGAGGAGGAGGAGGAGGAGGAGGAGGA]GTGTCCTCCAATTTCAGAGAACCAGATGCGTAGATGTG
AGGAGAAGGCTTACAGAAGCAGATGAAAGAAGGGCACTGGATCTTACAAAGCAAAACCCTCATCAAAACA
TATGTTCCAATCATCAACCAATTCAACAAGATCTCTTTTGTCTTTGGCAAAGTTAGAAACTTTGTGTGCC
CATTGATATGCCCAGATTGAGAAAAGGAAACACTTTTGATTTCTGAATAAAAAGTAGAAACAGAGCAGCA
AAGAAGTTAGTATATATCTCTCGTCTTGGATAGAATCCAAAGACCATAAATAACGAGTTGATCAGATGAN
NNAGCAGACAAA

AG0510 (SEQ ID NO:28)

ACTGGTCTCAATGGAGGTGGTGGAGGAGGAGGAGGAAGAATACGACTGGAATCTCTTTTAGACCTTGTAG
AGAGTTAAAAGATAGTTTTAAGAAACAAATGTCAATCTCTCAAAATAGGAATACTATACTCTTTACCTGT
GAGATAGCTGAGCAAAATCATCATCTGATTCATCGTCATCATCATGATTGATGCTGACAAGTTGAGCT[G
GAGGAGGAGGAGGAGGAGGA]GCTGTTGCACCGCCTCCATTTGAAGGAACAGAGGCGATATCGTCATGAC
GCTGAAGAACACGCTGCAAGTTATCGTTCAATGCTAATCCCTGGCACAGAAGCTCCTCGTCTCTGAAGAA
AAGAAACATCATCAAAAGGGGATGAAACATCAGGTGAT[GAAGCAAGAAGAAGAAGAA]AACTCACGTGG
TGGTGTTGACAAGAGTCATCACACGTTTCTGATAGGT

BG0031 (SEQ ID NO:29)

CGAGGAAGCAT[AGGAGGAGGAGGAGG]AAGCAGTTTGAGTGTTTGGAGGAGATGCCTGAGG[AGAGAGA
G]GAAGGGAGGGAGTCGGACGAAGGCGTGTGTTTTTGCACGTGCAGCGGAGGAGAATCCGGATCCGAAGA
CGGAGAGGGAGAGTGGGAGTCGGTGGAATGGACGGCTGAGATGGAGGCGGAGGCTGAGGGAATGGGATGG
GCCGTTGATTTGGGGATTTGGGTTATGTGTTTAGGTGTGGGCTACTTGGTGTCCAAAGCCTCAACTAAAA
CCTTGAGAGGTGGAGGAAGGAGAAGAAGATCAAAAAGTTTCTTTTAGAGTTCTCTGTAATCAGTCAGTCT
AGTTGTTCAATAACGTTCTAATGTAATAGTACAGATCAATAAACCATAAATGTAAAACAATCCATGATTT
TGAATACCAAGAGTCGCACGAGTTCCATTTTATTTGAGAGCATAGAACAATAAACTTTCTCCTCTGACCT
GATGAACTAAGGCAAGTTCATGCAAGAATCTAATGAATGCAAGCAATCAAGTACGTCAAATCATATTGCA
TTTACAAATTATACAAATACACAAAGGATCCAAAAAGTGCCTTCTCCCTTTTCTTACTAACAATAATAAT
AATGCAGCAAAAAGGAATAAAAGTTTATCAAAAACGTGTGATGATAATTCAATGTAAATAAGCAAATATG
TGGAGAGCT

BG0106 (SEQ ID NO:30)

GTAGCCTTGTGTGAGTTGGAACCAGACTTTCCTGTTTCCCTGCCTTGTCTCAAGGTTATGCATTTAGAGA
GAGTTATAGCTAACCTTGAGAGGCTTATAACTAGCTGCCCTGTTCTTGAAAAGTTAACCATAATCAGGGA
TTCTTTTGAAGTTCTCGAAATTATGTGTGTGCGCTCCAAGTCTTTAAAAAGTTTGGCTCTACTGATTGAA
GCTTCTGATACTGATCTCTTAGAAGATCACGATTTGGAGATCGATGCCCCAAAGCTTGAGCGTATGAGTC
TCTGTGATCACTTATCCAGAAGCATCGTTATACACAGTATTGCTCCCTCTGCAGTGGTACAGATCGATGT
TAACTTTAAT[AGGGAGGG]TGGTGATACATTATTGGACCAA[GATGATGATGATGATGATGATGATGAT
]TCCAAGAGAACTATGATCCGTAATTTCCTAACCGGGATATCCACAGTCAGCCTCATGAAGATCTCCTCT
GATACTCTACAGGTAC

BG0111 (SEQ ID NO:31)

GTAATCATTTCTTTGTTATCTCTCTTTCCATGATCGTCCGTCCAAGAGATATGTAATTGGCGTTGTTTGA
TTCTGCAATCCGTACAATCCATTTCTAGCTGTTAATCTGAATATAGCCATCTTATTAGACTGAAATCTAA
GCGCCTGGATGGGGTGGTTTTATTTTCATTTTGACTTTTGGCGTTTGGTTTTCAGATCTTTAAGATAT[G
ATGATGATGATGATGATGATGAT]GAAAATGATGAGATTTAGATTTTACTG[ACCACC]CTTTTTTTTTT
TTTGTCTTTACGTTTCTTTCAGCTCAATTCAGAGAAGAGCCCTTTTCAACGTACTTATGCAGCTCAGGTA
AATTTCATGTTTATCTGACACTTGTCTAGTAATGTGTGATACAATCTAAGAATGTAAATCTTACAATTGT
GATAAAAATATTCTCTCTCGTGTTTAGATAAAAAGATGTGGAGAGATGGCAC

BG0119 (SEQ ID NO:32)

GTGCGGGTTCGAGCAGCTCTCAGCGCTCGCGGAGGGAGGCATGAACGTGGCCAGGCTCAACATGTGCCAC
GGCACTCGCGACTGGCACCGTGACGTCATCCGCAGCGTCAGGAGGCTCAATGAGGAGAAAGGATTCGCGG
TCGCGATCATGATGGATACCGAAGGTAGCGAGATTCACATGGGAGATCT[CGGCGG]CGAGGCCTCGGCT
AAAGCAGAGGTTCCTTCCTCTTCTTGAAATCTT[GATGATGATGATGATGATGATGAT]GCAT[GTTGTT
]AATCAGATTATTGGATATAATCCGGTTTAGTTAGAGACCGGTTTAGTTAG[ATTAATTA]TGGTTAAGT
TTCTTTTTGCTTAATCATGTATATAAAGAAATGTTAACACAGATGAGGTTTTTGTAGGATGGTGAGGTTT
GGACGTTTACCGTTAGAGCTTTTGATTCGTCTCGTCCTCAACGTACCATTAGTGTGAGTTATGATGGTTT
CGCTGAAGGTAATGTGTCTTTTTTTTTTGTGTTATGAAAGCATCAAGTGGATGTGAGTATGAGATGGGGA
TCGATTTTTTTTTTTTTTTTGTGATTTCAGATGTAAGAGTTGGTGATGAGCTTCTTGTTGATGGTGGAAT
GGTTAGATTTGATGTGATTGAGAAGATTGGTTCCGATGTGAAGTGTCTGTGTACTGACCCTGGGCTGTTG
CTTCCTCGAGCTAACTTGACTTTCTGGAGAGATGGGAGTCTTGTAC

BG0181 (SEQ ID NO:33)

GTAACATATACAAATACTTCTAGGAATCAATCGAAATATATATTTCATATCGCAATTTCACAATACTGTT
GAACTTACAAACGTGTATAATTACACCATTTTTTTACACAAAATCTTTAACATGTCGATTTCTTATACCA
TTTGTAATTAACTCAACATATTTTTTTAACTAAATCAGCCTCGCCAATTTGTGTTGGTTTACGGAACCGG
TACAAATATTGTTGGCCTGGCCGTTATTAATTTCAAATGATTGATTCATAGGTAACATGAGAAGTTTGGA

GAGCTTACTAACGAAAGCAGGAGCGGAGACATTGCCATTGGCAGAGCAACGAAGTACGCCTTGAATATTG
ATGAGACCTAAAATGCCTCCAAGGAC[ACCACCACCACC]AAGAATCCCACCAAGGCCACCATTACCAAG
AAGCCCCCCTACTAGGCCACCATTACCAAGAAGGCCACCTAGGCC[ACCACCACC]AAGAAGGCCACCTA
GGCC[ACCACCACCACCACCACCACCACCACC]AAGAAGGCCACCTAGACCCCCAAGCTGAGCCTTA
ACCATTGGAGACACCATCACGAGACACACGAAGATCAGTGAGAAGGTTATGCGTTTGTTCTCAAGCATTG
TCATGTTCTTGG

BG0228 (SEQ ID NO:34)
GTATCTATCTCCTCTTGCCTAAATCACACCATGACTGACTTTCCCAAAATAACCTAGAGATCCAGAAAGA
ACGGAGGAAAGAAAGAAAAAATGGAGGAGACGAAGCCATTGGTAGGGAACCATCCCCAGCAAC[AGCAGC
]AG[CAACAACAACAACAACAACAGCAG]CTCCTGTATCAACACCAATTACAACAGAGACAGCAACAGAT
GCTTCTATTACAGCAGTTGCAGAAACAGCAACAACAACAAGCCGCCATGTCTAGGTTCCCCTCCAACATC
GACGTTCATCTCCGACCTCCAGGGTCAATCCAGACCCGACCAATTGTTCCCCCTCAGCAGCAGAACCCTA
ATCCCAACCCTAGCTTGGGACAGCCTACACCGAATCTTCAGCAGCAGCAGCAGCAGCAGCAACAGCAGGT
TGTAGCGAGTCAGCAGATGCTGCAGCAGCAGCAACAACAGCAGCAGCAGAAGTTGATGCGTCCTTTGAAT
CACATCGAGCTTCAATTCGCTTATCAGGACGCTTGGCGTGTCTGCCACCCTGATTTCAAGCGACCTTTCT
CTTCTCTCGAAGACGCTTGCGAAAGGTTCAGTTCTAATTTTTATCTAATTACATTTGTCTTTTTGAGATA
TTTCCTTAAATAAAATCGGTTATAGACAATCTCATCCGTTCAATCTTATTTCAGGCTATCGTGTGATATA
TGCATACGGGTCTTGTGATCTTTGAAATGAAACATTGATCTGTTAATGACTTACTTACTGGTCATATCTG
CAACTTGTATGTTCTTCTTTAGTTCGTGTTTGGTATTATGGTGATGATATCTGTTAGCCTTTTCGTTAAT
TTCTATACTTCTTTTCATTGATATTGTTTGTGTTAGATCCAATAGATCCTGCTTCTTTTGGTGTTCGTGC
GAAACTTAAATCTCTTTCTGAGTTTAGTGTGGTTGATTTTATATTATTTTTGTCATCTAATGTGGTTGAT
TTAGAATTACAAAACTTTGTGATTGTTTCCTATTTTAGTATAACCACCTGATTCACTGATACTGATAATT
ATTCCCTGACTTTTATATTTATGCTAAAAGTTTACAACTTTACATTAGCATATTATTGGTTTTATTAGAT
ACATTTGTTGCCTTGATTGAACATTTCTGTATATTGTTTGTTTTATCTTACCTCATAC

BG0255 (SEQ ID NO:35)
CTAGTGGCTACAAATCCAACTGTCGGTTCTCACTTGGGAGACCCAGGTTTGGATCTATCAAGTTTAAAAA
TCCAAACTCAGCACAGCCTTTCATTTCTGAAACAAGAAAAGAGATG[GAAGAAGAA]GATCAAGAAGCCA
AAAGCATCGGTTTCAGGGAG[GAAGAAGAAGAAGAAGAA]GATTAT[GATGAT]GGAGCTAAGGGTATTG
ATCTAGAAGGAGAAGAGAAGAAGCATATATGCTGTGAATGTGGCAAACGTTTCAAGTCAGGCAAGGCGTT
AGGTGGCCATAAAAGGATCCATGTGCTCGAAACTCGCAAATTCTCAATGGTGAGACCGAAGATGGTGGTG
ACGTCTGGTGCGGTTGCGGTTGCGGTTGGTAGATCTGATGAGCAGAGAGATGATTTCGAAGTTGATTGCT
GTGTTTGTCATAAGAAGTTTACATCGATGAAGGCTTTGTCTGGACACATGAGGTTTCATCCAGACAGAGG
ATGGAAAGGTGTTTTGCCTCCTCATCATCCACTTGATGATCATCATGGTGGGGAGTTTATAAGCTCCGAT
TACGATGATGATGCTGATTATGATTATCATGAGGATGATGATTATGAGAACTCGGAGTTATGGGATATTA
ATCGTTGGGAATTGGACAACGTGGTTGACCTTAAGGACTCGATCAAAGAAGGATGGACGGTGACAGGAAA
GAGAGGAAGGAGAAGTGCTTTGAAGATTGATGAACCTGATGATATTGATGCTAAGGATCTATTGTTCTTA
GCTACTACAGCAGAATCTGTCGATGCTGCAGAGACTTGTTGTGATTCGCTTTTGGGGGAAGAGATGATGA
TGAAGAAGAGGAAAAAGAAGAAGAAAAGATTGTCTGAGATGGAGAAAGAGTCATCATCTAGTCATGGTCA
TCATCAGCTTGAGGTTGGTGATGCTGCTGAGGGAGGTGGCGGTGCAC

BG0278 (SEQ ID NO:36)
GTATATGTCTTTGGTTATTTTTTTTGGTATCCAAATAACCGTAAATAAAAATTAAAAATGGCCCGTTTTC
CCTCGGATAAAAAAATTGTAGAGTTTAAATCATGTCTTTTAAAACCATGGGAGCAAAATCAAAGGAAGAG
AGAAGATAAATTAAATGGTGGCTGTTCAGTTGTTTAGCTGGAAGACATTGATTCTTCTACCTTCACAAGC
TTCAAGACATAAGGGTTTCACTTCTTTTAACAGGTTTTTAATCTGT[CTTCTTCTTCTTCTTCTTCTTCT
TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTATTATTATTATTATT]AT
AATAGTTTCAAGTTTCTGAAAAACAATTGATTCCATGGTGGTGCATGTGTTTTACAAGATATCTCACTGA

AAATTAACTTTGTTGCAGAACATTGAGTTTGCACTCTCTGCCTTCAAATGGGATTGATTCTTTTAGATCC
CGAGGTGAGGAGGCTCTGAAACACATTCCACGTCTTAATGTCCTTCCTCTCAACAAAGACTCATACTTTC
ATACTATCATATTTTCATAATTTCATTATTACAGGAACCTTCAGAGTCAAATCTCAAAAGACAGGAGACA
CAGAGTCATCTACTTCCAACTTGAATCAACCTAATGATTTAAAATCCAAATTCCATAAGGTGCGTGTGTG
TCATGCATGTCTTTACTTTTTTTATCTAATGATTTACTTAATGCTTTATGTTATAATCTTTCTTAATATA
CATATCTGCAGAGTCTCCAATATAAACTTGTACTAGGATGCATCCCACTGTATGCGGTATCGAGAATTGT
ACAAAAGATCATTCATGGGCTTCCACTCCACATTCAGAACTCAGTAGGGGCTGGCTTGCCTTTTGCTTGT
GCATCAGACTCTCTGAATAAACCATCTTTAAGTGGTATCAAATGGAGTCTTGCAAGGTTCTTTTTCCTGT
TCAATATTCGGCTCGAGAAGAACGTTGCTAC

BG0295 (SEQ ID NO:37)

CTGTTCCGTTTAACTATGCTCGCACCTCCATTATCTCTCCTCTTTCATAACTCTCTCTCCTCCTTCTTTC
TTCTCCACATCTCTCCGATTTCATCGCTAGAATTCTCCACCGATTCTTAAGGTATGTTTTATCTTCACTT
CAACTCTTGTCGGAATTCACTCTCCTTGCCTGTCTGAAACTTTCCATTTGCAGATCTGTAAAACTTTCTA
TTTGTGTTTCCTCCTTTCCGTAGATCGAGAAGAAACGATGACTTCAACGG[AGGGAGGG]ATACGATCCC
TCTTGTCTC[TCCTCCTCCTCCTCCTCC]TTCTCTTATCCATAACCACTCTAATCTCAGCCGCTGACTAC
ACACCCACCGACAAAATCCTCTTAAACTGCGGCGGCTCCTCCGACCTAACCGACACAGATAACAGAACAT
GGATCCCCGATGTCAAATCCAAGTTCCTGTCTTCCTCCGGAGACTCCAAAACATCCCCCGCCGCAACACA
AGACCCCTCCGTCCCCACCGTCCCTTACATGTCCGCCAGAATCTTCAGATCTCCCTTCACTTACTCCTTC
CCGGTCGCCTCAGGTATTGGTTCAATCCTGGTTTAGTAATTGTACTTTGGTTTACTCATTTCCGGTTTAC
TAAACACTTTTCCCTATCACAGGTCGCAAGTTCGTGCGTCTCTACTTCTACCCCAACTCCTACGACAGCC
TCAACGCAACCAACTCCCTCTTCTCCCTCTCCTCAGGACCCTACACTCTTCTCAAAAACTTCAGCGCCGC
TCAAACCTCCCAGGCGTTGAACTACGCTCACATCATCAAAGAGTTCGTAGTCAACGTCGAAGGTGGGACC
TTAAACATAACCTTCACACCAGAGTCAACGCCTTCTAACGCCTACGCCTTCGTCAACGGTATCGAAGTAA
CTTCGATGCCTGATATCTACAGTAGCGCCGACGGGACGTTGACCGTTGTAGGGACTTCTAGTGGCGTCAC
GATCGATAACACCACCGCTCTCGAGAATGTCTACAGGCTCAACGTCGGCGGGAACGACATCTCTCCTTCT
GCTGACACCGGTTTGTTTAGGTCTTGGTACGATGATCAGGATTACATCTTCGCCGCGAGTCTCGGTATCC
CCGAGACA

BG0452 (SEQ ID NO:38)

GGTCTGAGATATATCCTCGAGGGTTGTCCTAAACTAGAGAAGCTTGGGATCAGGGACAGTCCCTTTGGTG
ATGTTGGACTGCGCTCTGGGATGCATAGGTATAACGACATGAGGTTTGTTTGGATGTCGTCATGTCGGTT
ATCCCGGGGAGCCTGCAGGGACATTGCTCATACTCTGCCTAGTG[TGGTGGTGG]AGGCGTTTGGGTCA[
GATGATGATGATGATGATGAT]GACGAAGACGACAATGCAGATTATGTGGAGACGTTGTACATGTATCGG
TCCCTTGATGGCCCAAGGAAGGATGCTCCAAAGTTTGTAACAATTTTATGAAGACAAGCTTAGAGAAAGC
AGGAGCTGAAGTAGAAGAGAATGTGTGTTTGTATGATTGTTTGTACCATTTGATTTGATTGGCTCCCCTC
TGTTTTTGGATTTGTCTTGTACCAAGAAAGAGTGAAGAGTCAGTGAAGAAAGAGGTTGTTTGTGGAAGTC
AAAGAATGAAACTTTTATTATTTGTGTGTAATCAAGAATATGATTTTACAGCCATTTCACGATTATTTTT
GTCTACAAGAAGTATTGGTTATACATTACATTATAAGATCTTCACCAATCTTGACTTCGTCCTCCATCAG
CAGATGCTCTAAGGTGTCGATGAAAGCAGTAACTTTCTCCAAGCTCTTCTCATCAAGCCTTGGGACCGTG
TGGCCCTTGGGATGATGGACCACCACCGGATTCTTGAAGGAATCTATCAGCTCAGTTCCGTAAGGTTTCA
AAAAATCAGTCTCTCCTGCAAAGAAAAAACTCATTTTTCACATTGAAATTTGCAAACCAGATATACAATT
TAGTAGGTCATCAAATTACCTAGAAAGTGGAGGGAGGGAATGTCCATGGTAGACGAATACGCATCCTTCG
CCACCTTGGTGGATTTGAACATAGCTCCTCCAATAATTATGATAAACTTGATCTTTGGTACTTTCTGGAG
TGCAATTCCCTGCAATATAAAATATAATTCTAAGATAATGTAATGCGATTTCCCAACGCAAAAGCAACAC
TACTGACGTACCTTAGCTTGCAGTCCTGGTAATCCTCCAGACAATATTGCACCCTGCAAAATTAACATAG
AGATATATTATTAGATCTTATATAAGAAACTGTTAAATGAGAAATGAAGCAATTTTGTAATTAGAGTACC
TGAGAAAAGCCAATGAGACCATCAAAGGGACCAAGCTCGATCATACGATCCTCTAAATACTCCAAACATT
TCTCGAAATTCG

BG0516 (SEQ ID NO:39)
GATGTGTTCTTCATTGTATCTAGCAGAAGCTTGGTCAACAGAAAATGGCCTGAAACAT[GATGATGATGA
TGATGATGATGATGATGAT]GAGACTATAAAACTTAGGACAAAGGTA[TAATAA]TC[TTGGTTTGGT]T
TCTCTTAGCTCACCTAGATGGTTAGTTGCGAATTGCAGCTCAATATTGTCCTTAGAGAGCATGAAAGGAC
ATGCCATTACCCCAGCGTTGTTGCTAACAAGATTTGAGAGATTACAAAACATTAAAACCGTCACAAAACA
CTAGACATGAACTACTGTGTTTCGAGAGCTTACATCAAGATGTTTAGTGGAAGACCAGTAGATTTGTAGT
CAGATGCAAATCTCCTGACAGATTCAATTGAGCTGAGATCTAACTCCATGACGTCGAGTTTAGCACCAGG
GACTTGATTGAGGATATCTTGCTTAACTTTAGCACCGGAGACAGTGTTCCTCACCGCCATAACC

BG0647 (SEQ ID NO:40)
GCACATATGTCCGCACCTGTACAAAACCGCCTCGACCTGCGTTTCGTCGCAGACGCAGCATTTGCGTTTC
ATTGGGTTTTCTCTGTAAACCGATTGTTGCAAGCTCGCGTTAGCATCCAAACACGTTTTGACAGAATCTC
GTAGTAAGGACATTTCTTGTTGAAGCTGTTGGATCTGTGTTCTCATATCGGTTATCAGCTCCATTTCCTG
AAAACGTTTTAAAGCGGTTCAAAGATTTTACTATTCTACTAGTTGGGGTTTGCGAGTTTTCTATGCAATA
ACAAGAAATCGAAAATTACTTACATGTGAAGGAGGATTGTGAACAGACAAGACAGGAGTGGAAGTTACTT
CGGTGTCTTGACAACTCCATGATCCTGCAGGAGACGATGCAAAGATGGGCGAAGAAGACGATCTGCTTGA
GTCATCTCTATCGTTTTGTTCTTCACCTTCCGTTGATGGCTCTTCTTCAGTTTCCTCCGCAGTGTCATCT
CTATGTTCTTCCT[CTTCTTCTTCTTCTTCTTCTTCTT]GTTGCAATTCCCATGATTCAGAATGCTTTTT
CGAATGTGTCTGCAGACGAGACATCATGAGCCTATCGATCTGATCTCGTAACCCGCTCTCGAGAAAGTCT
GTCACTGTTCTTCTGTAATAGCAAGAAAATATTTATCTTCTTAGTTAATGGTTTAACAAATAAGAAAAGG
GATTTGTTGAATCGATGTTGCGTACCGCTCAAGGAGTCT

BG0651 (SEQ ID NO:41)
AATGCATAACAAAAGATTTGAACCCGGGTCTTTGGTCAAACAATAATCATCCTAAATTTATGCTAATAGT
GATTCTTTTGTTAGCCACTGAACACAAACTCT[CTTCTTCTTCTTCTTCTTCTTCTTCTTCTT]CCT
CTTTCCTCTGCACTCTCTCCGACACAAGACGGCGGTCAACGGAGTCCTTGTCGGTCAAATGATCCCTAAG
GACGAAGGAGGAGTTGTGGAGATTTCCGATTCTGTTCCGCTCTTTTGCTCCAACCTCGCTCTCCTTCCTC
CTCTCTAGATCTCGCTCATCATGGTCGCTCTCACTACATAAGTTTTTGAAATTGAATATTGAAAAACTTA
GGATCTGAGTGCACTGTTGCGAATTCTCAATATTGTTGTTCTGTAGCTGTGTTTGGGAGAGAGGCAGTGT
CTGTAATAC

BG0713 (SEQ ID NO:42)
ACGGTTCAGCACAGTAAAAAAAAAGTTTTTTTGACTTTTTTTTCTTTGACCGCCAAAGAAGACGAAAATG
AGTCTTTGAGAAAATCACAAAAAAGAAGAAGAAGAAAAATGAATCCTTTTTGTTTCTTCTGCACAGAATC
TTCTT[CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
CTCTCTCTCTCTCTCTCT]TTCTTGAGGTTTCTTTTCCTCCACGATTCCTCGTCCCTCTTGCTTCTGTGT
GATCGATTTTGGTGAAATTGAGCTGAGTGTATCTGTCCGCCGAGGCCTTTTGTTCACTGTTCAATTCAAC
ATCAGATCAATTTTAGGGGCTTTCAGTCAAAGATCGCTGCTTTGGTGTAAGTTTGAATTTGGGTAACTGA
ATGAATGTGATCTTTGGTTCCAGTTCATGTAATTATGTTTGATTGACTGGGAAAGTATCATCCTTTATTA
CGGATTGTAAACATTTAAGGTTGAATCTTAACATTAGCACCATTTGGATTCGAATTTGTTTGGTGGGTTT
GGCTTTAGATCCATAAGCAAGCTTATGAGCTCTTAAAGTTATGTTGTTTTTTTTTGCTTAAGCCATTCAA
ACTGATGAGATATACTCTCTTTGTCTTGCTTCCTAGGTTTGTGATTTTAGTATAGAATCCTGTTATCATG
GATGAACACAATAGGAATCCATTTGCAAGTGCAAGCGGAAGAGCAAGTGGAAGTACAAGTGTGAGTTCCA
ACTCCAGTTTTAGTAGCAGCGTGGCGGATACAGAGGATGATCAAACCATTGC

BG0864 (SEQ ID NO:43)

TGCAAAGGAAGCAGGTGTAGCAGCTCAAGCTTATGAAGCTCTAAAGACACTGAGAGAAAAAAAAACATCT
GCAAAGTGGTAAACAAACTCTTCTTATTTCACACAACACATGGTAAAGAAAATACTTTTTCATGGAGAAT
AAGAAGAAGAAGAAGCTAAATGCGTTGCGTTGCAGGTG[GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
AGAGA]AGCGCATAAGTAGTGTTTGTGTGTTGGTGATTTTCTATTTGGAAACTCTTTTGTAAGCAATAAT
CTCAGATGCTAAAGCCATTGTATTTATTGCTCACTTCATTTTACAGCCAAACTAAGTTTTAAAAACTGAA
AATATAAAACGCTAAAATTTTCTTTGGTTGACATCAGCATAATATAAATTTAGCTTACTCCCTCGATTCA
ACAATACAAAAAAAACGACATAAGTTTGAGTTTACATGCTTTCAACCAATAAAATGGAACTCTTTATCAT
AAAATAACAGTCAACGTATTATTAAGTCCAAACCACCACAAACCAATATTTGCACAAATAAAAGTTTCCA
ACCTTAGCTGCCACTATAAAGTTATAAACCACCATCCAAAGTCCATTATTTTAAGATAGATTTCGTACGG
TAC

BG0869 (SEQ ID NO:44)

TCGTATAAAATAAAATTCTGAACAAAAATAATTATATAATTTCAAATTGCGGCTCAAAATCTATTATTTT
TAAAAACTCAACAAAATTGTATATGGGCCGATGAAGCCCAAGTATTTTAATTACCGTAAAGGAGGGTTTG
AGTCGGCCACAAATCAAGGAATTATTTC[CTCTCTCTCTCTCT]CCTGTGACGAGTTG[CTCTCTCTCTC
TCTCTCT]CGTCTCGTCCGCGCTCCGAAGAAATTTCACAGATTCCTGTCATGTCTTCCGGCGGAAACTCT
ACCCTCTCCAACGTCGAAAAGATGTTCTTCTGTTACCAGTGCAATCGCACAGTCACCATCTCAATCTCCT
CCTCCTCCGACGATCCTTTCTGCCCTCGCTGCTCCGGTGGGTTTCTAGAAGAATACGACGAGCCAAACCC
TAATCCGCCCCCAAATCTCAACCCTCTCGGTTTCCTCCCCATGGCCGATCCTTTCTCCACCCTGCTCCCG
CTCCTATTCGGCTCCTCCTCCTCTCCTCCTTCCTCCACGAACCAGAGCTTCTTCGGCCAGAATCAGCACC
CTCCTCGCGGCGGAGCTTTCGATCCGGTGTCGTTTCTCCAGAACCATCTCCAGCACCTGCAATCCAGCGG
CACTCACGTCCAGTTCGTGGTGGAGGATCATCCCTCGGATCCGTTTGGCCGGATGCCGGGGAACATGGGG
GACTACTTCTTCGGCCCTGGCCTCGAGCA

BG0988 (SEQ ID NO:45)

AACGGTTTTGTATAAATAGTATATTCTATATATGTATGCATATAATCTTTTTTCGTAACTTAAAAGGATT
AAACCGGATTTATTAAAGACACAAATCTAACTTCCAGATGAGAGGTGCAATACACATATGGATTATTTTC
CAGATATTTAAATGGACCATAAATATAGACCCATAACCGCGTGGCCACATATGGAACTAATGATTTCGCA
CTAGAAGGGAATCGATTCCTGACCTGAACCAACAGGACAATTCCTCCTCTAGCGGAAACCATTAAGCCAC
CACAACATGGTTTTAAACAAAAAATTGTACGCATCTGCGTGGCTTACTATTAAAACATCTCTATCTCTCT
CTTAAAATACATCAAGAGTATAATGAGAGATATCTCAGTTTCATGTAGTAAGACAAAACCCAAGACTCCA
ACCGGAAAATTCCAACCCTAAGAGGCAAACTAAATTTCATTGTACAATAAAATAATTAATGCTATTCAGT
TTTCTAAAAGCAGATTTAAGTCTCTAACTCCAATTTTCCAT[CTCTCTCTCTCTCTCTCTCTCTCTCTCT
CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
CT]AAATCCCCACTAGGATTATGGGAACTCACGTCCTCGTTTAATGCGATTCATGACTCCTCAAAGCCCA
GCATTCCCACTCTGCAAATTACTAGTACCTCTTAGTCTTAATTACCATTTGACCAATCT

BG1062 (SEQ ID NO:46)

GTAATGCGGCAGGACAGCCCTCCTCGGAGTCCACTTCATGGAGGAGCATACTATTCATCCAGTGATGATG
ATAACCACTCCACCTACCTCTTCCCAGAAATTGGCACCCCAACTCGTTCCATCCCAGTCTCCGCCAACAC
CACTGTATGAAT[CTCTCTCTCTCTCTCTCTATCTCTCT]TTCACCATTGTTTTTATGATCTTATGGACC
TTAATAAATAAACATATGCAGCCTGTTCACCACAACTACCAAATCATTGCGGTGGAAACCTACGAGCAAG
AGAAGCAGTACGAGCCACCGGAGCTAGCGGACGAGTCACAGAGCTTCTCGATCCAGGAGATCGCCAAAAT
GCGAGGACTCAAGGAAGAGAGCCAATCGATGATCTCCGAGTCCTAC

BG1090 (SEQ ID NO:47)

GTGAAACCGGTTCTGGAGAACTCTAAGGTTGTTTTGAAAGATCGCAAAAGAGAGTGGAAGCGGAATTAGG
GTTCTGGTTATGGCAAGGAGATGAACCGGAAGGAGGTACGAATCCTTTGGGAAGCATAACAGAAAACGTA

TCCGGTCGGACCGGTCGGTTAAAACCGTTACTGTTTTTCTTGGTCATCATCTTGAAGTAGTTGGCACGGT
GACGAAGACCACGGCGACGAATGCTATGGCGGTAGTAGTAGTAGTGATGGTTGTAACGAAGACGACGGTG
CTGATCACGTTGCCGCGGGTTCAAGATAAACGGCGTCATTTTCTTGATGTAGACGAGTTGGTGCGGTGAT
TCGTCGCCGTTGAGGTAGGTTTCTGCCTTTTGTGTAGATACTCTTGTTTCTTGATTCGATAATGATGAGG
ATGATGATGATGGCGATGATGGTAAT[GATGATGATGAT]GGTGAGATAGGGAAGACGAGAATGAGAATG
AGAGAGATGATGAAGCATTTGACAAGGTTGTGTTTCATCAAAACATCCATTGCGATT[GAGAGAGAGA]G
GGAGTAGGACTTTTGGTTTAATAGAGAGAGGGAGAGTAAAGATGAAACAAAAAGATGTGAGCGAGGCAAC
TATAACAAATCTTGGTATGGCGTCTAAATAATTCGTTTAGTTATTCGAATTTTAATTAATTTTAGTATGA
TTTTTGATTGCGTATAATTTGGAAATTAGTTGGGCTTTTTGTTGGTCTGAGGC

BG1101 (SEQ ID NO:48)
CACATGCTTGTGGATAAAT[CATCATCATCATCATCATCATCAACATCATCATCATCATCATCAT]CAAT
ATCAAATATATGGTAAGTCCATTTTCATTTAGCTTTCAGTAAAACTGTTAATCTATGCATTCGATAATTA
AGAGAATCAAACGAATTGTGTTTGCAACATTATAATTAATGGTTGAAATTCATTAAGAATATTTAGTTTG
GGTTTTCTCATTTTCATACAAACATTATCCATGCATACGGTTGGTCATTAGGTTTTGAAAATATATGAAA
TCAGAAACATTTTAATTTTTTTTAATGTAATTTGAAAGCATACAAGTTATGTATATTAACTTTGTGTAAT
TTGAAAGCATACAACTTATGTATATTAACTTTTCAAAATTTGGACTATAAATAAATATTTCTTTGATCTG
CCCAAAATCACAAAAGATTCTTTTACAAGATAAACTGTATCTTTTACTCTCTTTTTTGTCAATACTGTAT
GTTTCACTTGTCACGAATTTGCATTCAAATAACTATGTAGCAGCACATTATGATAAAGTTGGAAGTGTAT
GAATAAATTGATAATGTAGATTGTAGGGTGAGAAGTTAAAAAAAATGAGTAATTTTTAGGGGCCAAATGT
ATTTTCGTATAAATTAAGGGTGGAAACATGAAAATTAGATTTTTTATGTCCGAACTACCCACTGACTTGT
CCGAAGTCCGT

BG1123 (SEQ ID NO:49)
GTTCAAAGGCATATTTATATTATATTTAAATTGGGACCAAGATTTGCTTTGGGACAAGCTGTGCCCCACG
ACTTTCTCGCTAGTGCTCTCTGGTCGCTTCTCCTTCTAGAGACCAACCATTTCCACCAACTCCGTTTTCA
GTTCACACCATGCCCACCACTGCATCAGTTAGTTGATATGAGCCCAACTTCTTTCTTCACTGTTTAACAA
AATGGACTGGTCAACACAGTCTCTGTCACACCCGAGAATTCTAATGTGGTGGACACAATCTTCACTAGGC
CACCTTTTGTCACCAGTCTCTCTCTCTCTTTTCCTGTTTTGATCCTTCCATAAGATTAAACCTTTATGGT
TACTACCATATTATAACGATCTCGGTGGTGGTAGCGTAGCCCAAAGATGATGATCCGAAACTGAATGTAA
ACTATGTACCAAA[GAGAGAGAGAGAGAGAGAGAGAGAGAGAGA]GG[GAGAGAGAGAGAGAGAGAGA
GAGAGAGAGAGAGA]GTAATAATTAAAACAAATGGGACAAATTAACCCCCC

BG1127 (SEQ ID NO:50)
CCGCCATAACAAAAATCTTCCCACAAGCGTGTAGAGATCTG[GAGAGAGAGAGAGAGAGAGAGAGA]GCT
TTCAAGTGATTAAATCCAAAAGTAATAAAGAGAAGACGGAGAAACTAAAGTGACGCGCCCCCTTCCTAAC
GGATTATATTTATTTTATTCTTATATATTTATGGGCTTATTGCAGCAATAGCCATATTTGAAATGAAAAT
TAAGAGAGTAGCCATGATGTTGACATAATGTACTCACTGTCTTTTTACAATTTTACTAGCCGGTTATACC
TTTGTAGGAAACAGGTTCCCAGTTCCTTTAACTAAAGTAAACGATGTGGTGATTTACTGACCCATAGTAA
CAATGAGAGTATTTTAGCAACGCCTAAAATTAAAATGAAAGGAAGGAAAACATTCTATAGAGATGAAAAT
ATAAAAAAAACAGAAGTGTAAAAGAAAGAACGTTACAAACGGAGAATGCATGGATCGTAATGCTGATGCC
AAAATATGGAAATAGTTCCACTTCAAATAGAATACACATAGTATAACAATAGTTTAAAGTTTGTCACCGC
TATGTCATATGAGAATATTTTCCATTCTATCGGATATAGATCAGTTTATATTTACTAATATAATCAC

BG1149 (SEQ ID NO:51)
GTAGAGTGATAAGAGAACATCTTGTCTGACTCATGCTTCTTTTCTGCTATGTTAGGCTATTGCTCTTAAT
GATCTGCCTTCCACTAACGTATCGAGACTGGGGCTTGCTCTTAACTTATCTCTTTTCTACTATGAGACTC
TCATATCAACTAAAGCTGCGCGTAAGATCGCAAAGGCGGTATGTTGTTGCTTTCTCTCATTTAGTATTTT
GGTTTATGTTATGCGATTATCATCTATTCTCCCAGATGCTCTGTTTGATAAAACTTAATGCTTTTCTTTC

TTTTTTTTTGTTCTGACAATCGTTACTGTTCAATTGTATTCATATTGTGGCATAAATATGTATATGTTGC
TACACTTCCCTGTGGGTGTGCAATCTTCATATGATATAGTAATGGTTTGCAGATTGCTTATCATTTGGAA
GATAGATATTGTAATTGATTATGATGATGATGTGCATAATTTGGAAAGTAGAGCCTATCGTTATTCCCTT
ACACTAATGGAATTATATATTGATGATGTTCCAAATTTTTTAAATATGATGAATT[GATGATGATGATGA
TGATGAT]GGTAGGCTTTCGAAGCGTCAATAACAGAAATGCACGCAGTGAGAGAGGAATCATACGAGCAA
ACTGCATTGATCACGAATCTTATCCTTGACCGTATCACCCCTCTGG

BG1182 (SEQ ID NO:52)

GTGTTTCTAAGCACTTTTTTTTTATAATCAAATCACATACAGCAAACATAAACATGAGTCTCAGCTTCAA
GAGCCAATGAAATTAGCTTCCTTTATAATATTCAAGAACTAACCAGTTTCACTTCTACTAATCCTCGGCG
CACTGATGATGTTTCTAACTAAATTGGATACTAAAGAACGTAGCTTTTCACCTCGAATCAAACAACTTGA
AAACCAAAACAATCTAAACGAAATTTCATAACCTAAGGAGGATCGGAAACTAAAATTTCTACATCGGAAT
CGAATCGACGCGAAGTGAAACGAAGATCGAT[AGAGAGAGAGAGAGAGAGAG]GACTCACTCGCCAGGAG
AAGACATGTTCGTCGATTTCGAAGATCCGATTGATTCAGAAGCGGAGAACAATCCAAGTTTTTTATTGAG
AGAGCACCGAACAAACTCTCTCCCTAGAACGTTCCTTCCCAGCTTCTCTACAAATCACTTGTTCCGCGAC
TGCGTATCTTATCCAATCATGTCTTGCCACG

BG1197 (SEQ ID NO:53)

BG1197 (SEQ ID NO:224)

GAACATCCTGCTGTTTCAGTTCTATTTTTCTTGTGTAGTCAAAAAGACATTATTACCCCATTGGAAATTA
CAACAACACATTAGTTGTGAGCGCCAAGATAAGGCTAAAGACGTAAAAACGCTCTGAGTATTCATTCTTT
CAGGTCAGGTTCAGAAACTAGTTTCGTTT[CATTTCATTTCATTTCATTTCATTT]CATCCACCTCCTCT
TCACTTGAGAAGTTCTGTCTTTTGCGATCCTTGTCATTTTTGTAAAGGTGAGTCGATCTATATATGGTCA
CTAGTATTCTGGAAATGATGCTATTTTAATACTCAGTTCGAACATTCTGTTATCAAATCCGGTTCTAGTT
AGTTGTTCGCGGGATAGGGTTTGCTTGAGATCATTTCGCTTCTTTATTTTTTTTAATGTCACTGATGGAT
CTGGTAATCTTCCTTATCGAATTAGGAAAATGAATCTGTATTAAGTGGACTAATCTCAAATCTAGGTAAA
AAAAATGGGAGGAGGAGGAGGAGGAGAAGGAGTTGGCGATTTCAGAGCCAAAGTATGGAGCATGTCTGGT
GGGCCTTACTGTAGGCCCAAGCACTGGCGTCGCAACACCGCCTTTGCAATGCTCGGCGTTTTCCTTGTCT
GCATCCCCATTGCCATGAAGTCTGCCGAGCTCGAGG

BG1230 (SEQ ID NO:54)

AGAGAGACCTCCAGTCACCTCGTCTCTTCAGGCCTCTTGTGTTTCCTCCAACTTCCTTTACCAAAAAAAA
ACAAATCAAAATCAGATTCAAAGGAGAGAAAGAGAGAGGGAGAGAGCACTACAAGAGTGGAAAAGAAGAG
AATCAGGTCGTG[GAGAGAGAGAGAGAGATGG]CG[GATGGTGGTGG]TGATGAATCTGAGATGCGATGG
TGGTCGTGATGAATCTGAGATGC[GATGGTGGTGG]TGAGGAATGATGGCGGATGGGAAAGATGGC[GAT
GGTGGTGGTGG]TGACGAGTGAATGAGCGG[TGGTGGTGG]TGACGAGTGGGAGGAGAGATGGCGGTAGT
GGTGGTGGTGGTGATGAGGAGGTCAAACCTGATGGATTGGAGGAGAAAAGGAGGCGTCACAAAGAGAGAG
AGAGAGATTTGTGTGTTAGGTTAAAGATTGCACATTCAGAAATGTGCTTAGACAATGATCTGAAGTGGTC
TTGGTCGAGGTAGTCCGTACATGTCCGTACACAGTGC

BG1241 (SEQ ID NO:55)

GGATGGGATGAGTCAGCTGCTGGTGATAGGCCCAGTCGAGTTTCAGTTTGGGACATTGAACCAGTTTTAA
CTCCTTTCTACATATGTCCTCCTCCATTTTTTCGACCTCGGTTTGCTGGACAACCAGGAATGCCAGGTAA
AGTCTTTGTACAGTTTCATTTTGCACATCATCTTTGAATCTCCTTAGAGATGGCAATTCTGGTGGTCTTG
CAGATGATGGGACTGACATGGAGTCTGCGTTGAAGAGAGCAATGCCGTGGCTTGACAATGGCCTAGAGAT
GAAGGACCCTTCCAGTACGATATTTCCTGGTCTGAGTTTAGTTCAGTGGATGAGTATGCAACAGCAGAAC
GGCCAGGTCCCTTCTGCCGCTGCACAGCCTGGTTTCTTCCCGTCAATGCTCCCTCCAACCGCGGCTCTGC
ACAACAATCTTGGCGGGGCTGATGATTCCTCAAAGTTACTGAGCTTTCAGGCGCCTCCAGGGGGGGTTTC
CTCATCAAACCTCCAATTTAACAAACCGAATCCGCAAGCGGCAATGTCCCAGTTACCTCAGCCACCAACT

ACGTTGTCCCAACAACAGCAGCTGCAGCAGTTGTTGCACTCCTCTTTGAACCATCAGCAGCAGCAGCAAT
CACAGCCTCAGCAACCACAGTCGTTGCAGCAACAACAACAACCGCAATCCCTGC[AACAACAAC]AATCA
CTG[CAGCAGCAACAAC]AATCACTACTG[CAGCAGCAGCAGCAACAAC]AATCTCTGCAGCAGCAGCAG
CAACAACAATCTCTGCAGCAA

BG1244 (SEQ ID NO:56)

ATGATGATGAAATAGCTCTGAAGAAGAAGTTAATTAAGGAATTGTTGCTGTCTAATTAGGTGTTCTTGTT
GTTTGGTTAATTATGTTTGGTTCTCGGATTTGAAAGCTCTGTTAAAGAGCTTCAGTTTTAACTTTAATTA
TCGGATTTGAAAGCTCTGTGAAGAGCTTTATTTTTCACTTTATCTGTAATTGTTCTCCTGTTCTTGATGA
TATAAAATATTTAAGTTGTTCTTGTGTTGTTCAGTTATATTTACAGTTGTTGTTTATGATATATCATGTT
TCTTTGTCTTGTAGAGAAGTCACGGAGTCCACAGAGATGCTTGGACTGAAGGGAGTCACGGAGTCCACAG
AGAGATGTCATGTACCATGTCTTGTAGTGTGCAGGGTCTGTAACGAGTCACGGACCATGTGTTTGTATGT
GTCAGTATGTGTTTGTACGTGTCTTGTATGTGTCACAGAGTCCATGTTTTTGTTTGTGTCTGTATGTGTT
TGTATGTGTCGATGTCTTGTAGTCACGGACAGTATTTTTGTAGTCACGGACTTTTACCAAACTCATCTTC
TATTTATAACATCAATCTCATCTTCTATTTATATCAACCTTCTCTCTCGAAACATATACAACGAACTCTT
CTTCTCTGCTTTACAACAACAAACT[CTTCTT]CT[CTTCTT]AACAACAACAAACT[CTTCTT]ACCAT
ATTTATATTTTTTCCCCTTATTATAAACACCAAAACCATCTTTATAAAAACTTTATATGG[CTTCTTCTT
]CTC[ATGATGATGATG]CGTTTG[ATGATG]CATTTG[ATGATG]TTTTTG[ATGATG]TCTATGATCA
ATATTTTGATCAAGCATTTGAGAATTTGACCATTTGTCGTGATCAAGAAGAACGAAGAAAGAAAAGAAAA
AAACGAGCGTATATCGAAAGACATCGTGAGGAA

BG1286 (SEQ ID NO:57)

AGTCCGAGACAGAGTATGCTA[AGCAGCAGCAGCAGCAGCAGC]TGAATACTGCATATGATGCGTCACAG
ACAAATGCTCAGAATCAGATGCAGAATCTTGCTTCTTTATCAAATGTGATGGTAAGCTACATGTGCATTA
TTCATATTTGAAGTGATCCACCAATGACATTCTCCAATGGCATTGCTAACATTGGTACTTTGTTTGTGTG
TTTTTGACTCAGCAGGGATATCCACACTCAGATCCCAACAGTTTATTGGCACAAAACGCTAGGGAGCTTG
AGTTCCAGTATTCCAATTTTGCACAGTCTATGCAGTCAAGAAATAGCAATAATGCTTCTTCACTTGGTGG
TCAAAGCATTTCCATGCCAGAGGTAAATAACCACTTTTGTCTTCTTTTTTTTTAAGAAACACAAGATGTC
TTGTTAATTAGGTTTTGCTCGACTATGGAGTGATCTATATGTATCCAAATCTATACAACAAGAGGAATTT
ATATGATTTTGATTATATATTTTCTTACATTGTAGGCGCCCCGAGGCAGTGGAATCCAAGCGACGCAGCA
AAACTTACAAGGTGCTAATATCGCCACTGGACCAGCACTTCCTCAACAGCTT

BG1288 (SEQ ID NO:58)

ATAAGCATACCAATGAAGATATCAAAGAATGCAATATGTATGTTTTGTGTTGTGAAAGCTAAAGGATTCT
ACTTTATTTGTGTTGAGTGATGGTTCTTTAGTTTGGTGTTAATGTCTTGTGAATTGTGTTTGGCAGGTAC
AAGCTAGGTTTTTGTCCCAACGGTCCTGATTGTCGGTACAGGCACGCGAAGCTGCCTGGACCGCCGCCTC
CAGTTGAGGAAGTTCTTCAGAAGATACAGCAGCTGACTTCGTATAATTACGGGCCTAATAGATTCTATCA
GCCACGGAACGCTGCTCCGCAGTTGGGAGATAGTAATAAGCCTCAGGTGCAAGTTCAGACGCAAGAGGCG
GGTAACTTGC[AGCAGCAGCAGCAGCAGCAGCAGC]AACAACCTCAGCAGTCACAACATCAGGTCAGCCA
GACTCAGACACAAAACACTGCTGACCAAACGTCTCATCCTTTGCCTCGTGGGGTAAATAGGTGTGTTCAG
AGTTTCTAAAGTTTTTAATTGGGTTGTGTAAACTATGCTTCTGTATATCTGTCAAGACATTGTTTATTG

BG1321 (SEQ ID NO:59)

ATGAAGATTGATGTATATTCGAATTTATAAAGTCTACGTTTAGTAAAGGTATACAAATCAGAGGTCTGAA
TTTGTTCAACTTCCTCCTCATTCCCCCATCCCCAAAAGAATCCGAGTTTTTTTGGATCAAGCCTATATAG
ATCCAAAAACCAACATAATGGCCCATTAAAGATGCATAGACTCGAACCAAACCGGATTAATACACTGCGG
GTGAAACCGGTTTGGGAATTTTCACAATTGACTGAAGAATCAGGGTTTAAGGAGAAGTCACAGACCCAGG
[AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGCAGAAG]AATGGAGTCAGAGCACCAAACGA
TGGAACAGTTCCTACGATGGGCAGCAGAGCTTGGCGTATCAGATTCCATCGATCCTTCTCGATCTCAAGA

TTCATGTCTCGGCCATTCCCTTTCCGTCGCCGACTTCCCTCTCGCCGGCGGGTGCGTAGAAACAAAAATC
ACATCTTTTTATCATTCAAATTCCTAAACTTTTTCGACCATTGATGGGAAACTAGGAGAGGGTTGGGGGC
TGTT

BG1368 (SEQ ID NO:60)

CTTCAAGTTTCTATTATTATCAGTTCAGGAAGTGACACTCACTAGACCAACAGAAGAAGAAAAAAATCAA
CATACAGAAAACAGATAAGCACTGCTCATATTAATCATGAATCGTTCAACAAATTTGATCCGAACATTAC
AGAAACTATACGTGTTTGATCCAACAACGAAAGGAGCACAAACAAAATGAGATCAATACGATCGTTCTTC
ATTGTCGTTCTATTACAAAACTGTGCTTGCTTTGTTGGTTCGAACTCGAACATACAACAACATAGATAGT
TATGTCGGGATATACTTATTTATATTTAGATTTAATTATGGATAACGACGGCGAGAGATTCTCGGCGACG
GAATATCAACTGTTTCGCGATGAATGCTTCGATCGTTTTCTGAAACTCTTCGTTTGTCAGATTATCCTCC
GGAAACGGCACTTCTTCGCACGACGCGACGGTTTCACGGCACTTCTCCGTCTCCGATCGCCGGAGCGTAG
GTTTCGTCACCGTCTCCGGACTCTGTTTCGCAGAGATTTCCGTTTTGCTTCTTTTATAGACCTT[CGTCG
TCGTCGTCGT]CGGAGGATGATCATTCGTCGATTC

BG1392 (SEQ ID NO:61)

CTCAATCTTTGTGTGGGTGGAGCAATACATAAATTTTGTGTGGTTGGAAGGTTATGAAATGAAACTTTTA
GTGGAAACGTGAATAGATCATCCAACTTATTTTGTGTGTTTTAAGAAAGATTAGATGAATTTGACTCAGC
TCATTGGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTTCACAGAGCCCATCGAATCCTATGCGCGTGTGAA
AAGCACGATCCAATCACGAAGCTAAAATCTTCAGCTTCGTTGTATAAAAAAAACTTATTGAAACAAACCT
CAAATTCCAATTACACCCTTGACAGCGATACACACTCTCTCTCTCTCCAACTAAAACATATCTGGAAATT
ATAAATAAAATTTATACTTTATCTGGTAACCCATCAAATAAAGCTATTAGTCACATAATAGATGACAAAA
AAAAAACAAATAAAGAAAATTTAGGAAACAAATCTACTGAGATTAGGCTGTAAATCATACGTATATCTTT
CCCGTATACAGAGTGCCGTTTTAAGTATAATGTCGACACGTGTCGGTCAGAGGCTCGGCTTCCAAGGGTA
AGATTGTAAAATCACGATCGTCATCTCTCTTTAAGAATTTCCAGAGTGCT[GAGAGAGAGAGAGAGAGAG
AGAGAGAGAGA]GGTGCTTTCCCATAGCCATTCACGTC[GAGAGAGAGAGAGAGAGA]GGAAGGAGATGG
AGGATATACAGGAGGAAGAGAACGGTACGGACGAGGAGGTGCTGGGATCGAGCTTGACCATGGAGAAAGT
GGCGGCAGCTAAGCAGTACATCGAGAATCACTACAAAGCTCAGAATAAGAACATTCAGGAGAGGAAAGAG
AGGTATTATAAATCGTCTCTTTCGTTGAGTGAGAGATTTGAGATTTGGAATATCGTTTTTTTTTAGAGAC
TAGTTAGGGCGAAATTAGTTGCGTGAGCTTTGATTAGTCTCTCGTATTTGATGATAATCATGGT

BG1442 (SEQ ID NO:62)

GAACAATAACCTGATCAAGGCTCTGCTCAGCTGCTTTACGCTCGTCGGGATTGGGACTGCAAGCAGCCGC
AGCGATGATTACTGCCAGGTTAGACAGATCCATCGGGAAAATTCAGCAACAGATTCTCCGAGGAATAGTC
GGCGTGTCAAGAATTTTCCGGTGAATCGAAGAGGCGAGGAGGAAGAAGATGACCGACGAAGACGAGGACG
AG[GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA]GGTATAAGAAGGAAGGTCTAAGCTAGGGTTTT
AATGTATGGTTCGTTGGTCTGTTATGGTAAGTCTATCGCACGCACGCGCGTGGTAAAAAAGTGAAAAAAA
AGAAAAATCGACGTGAGACACGATACACAACCCAGACTTGCCTGGACCTCTAGTCACCTATTTATTTTCA
CCGCCTGCTCGTTTAGTTAGTTACGGTCAATCGATTGACTTTTGGTTATTTTCTATGTGTTTTCAATATA
ATCAAATTCAAATGATTTTTTAATCAAATCAAATGTAAATAATAAATAATAAAAATCCAAATGGAATTAA
TTAGTTTAAACTTATTAAACTATTTTGTCATCTTTTTAGTTATTTAAGAATTATATTAAAACTTGTAAAA
TTCTAGTTACAAATATAATTTAATGCATAAATCTAATGAATTTGGGAGAAAATAG

BG1449 (SEQ ID NO:63)

GGTTTCCGCCGCAATGATTTGGTAGTTGTGGTGAACAGGCTGCATATGTTTATTTATTAAGGTCCATAAG
ATCATAGAAACAATGGTGAAA[GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA]TTCATAC
AGTGGTGTTGGCGGAGACTGGGATGGAACGAGTTGGGGTGCCAATTTCTGGGAAGAGGTAGGTGGAGTGG
TTATCATCATCACTGGATGAGTAGTATGCTCCTCCATGAAGTGGACTCCGAGGAGGGCTGTCCTGCCGCA
TAACGTAGCCATCTGGCTCATACATTGCTTCCATTTGCTCATCATTATGATTGTTGCTGCGACGGTCAGG

CTGCGATTTGTTTAGGAGAGCTGATTATTTATTGTCAAGAATATGTTTTATCACTAGAGAGAAGCTCAGG
CTTGAGAATGTTGTTTGAGTACCGTTCTGTATTGTTGTGGTGGTGGCGAACGGGTCGCAGCTGGATGAAA
TGGGCTTGGTGGAGAGCTGTTTCGTGGTGTTACTTCGCCGTCTCGCTCATACACTTCCTCGGATTGCTGA
TTATCAGGGTGGTTTCTACGACGATCAGGCTGTTAGAATTAGAATAACGAACTTTATTGTCAGGTTAAGT
GAATCACTTTATCTGTTTCTGGAAAATAAGGCTGAGTTTTATGGTACCGTTCTGTATGGTTGTGGTGGTG
GTGGCGAACGTCTTGCGGCTGGATGAAAGGGGCTCGGTGGAGAGCTGGTTCGTGGTGTTACCTCGCCG

BG1453 (SEQ ID NO:64)
AGTGCTTTAGGAGAGGAGGATCTTGATACAAAATTGTTCTCAAGAATTGAGAAGAGGAAGAAGAAAAAAG
GGTATTAAAAGGAAGAAGAGGAGGGCTTTATTAATCATGTATTGCTAAAGAGGAAGATGAGGAACTCATC
ACCTTCCTTCCTTCAATGGCAATGAAATGAAAA[GAGAGAGATATGAATGAGAGA]GTGAGTGAAAAA[G
AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA]GTTGGGTCTACCATGA
AAAGCAGAGGAGGTGGGTCAAAGCAAAAGTTTTCGACTCTTTTTTTAGGGCTTTTCAATTTCCTTTTTTT
CTTTACTTTCCATGTTTGTATATTTCCCAAAACTCGAATTCATACACAAGTATCTCGGGAAAACGGCTTA
TTCATGCCCGAACTAGGGGTTGCTGAGAGAATACATACCTCAACTTTTACTGCAAGTCGAAACATAC

BG1513 (SEQ ID NO:65)
ATAGGTTCTGTGTCGTAACCATTAGAGTCTTAAAACCGTCAATTTCGATTCTCACATCCAAAAGGTTTTA
ACATTTATATATATAGTAAGATGGGTAAAGTTTGTAACCGCTCGTGTATAAAAATCAGGTTTACGAAAAT
CGCATTACTTTTGATTTTCTGATCGAAACAGAGCTTCGAAAAGGAACTACTACGCAGTAAATAAACTTAC
TTGAAGAAACGAAACTTACTTCGAAAAGGAATTACTTTCTGAAGGTTGCGATAGCGAAGAGAAACTT[GA
GAGAGAGAGAGAGAGA]GCGAGCGAGAGAGATATGAGAGTGATACCGCGGCGGAAATGAAGAAAAAAAAA
AAAGGTTTAGGGTTTAACGACGACTGTTGCAAGTTGTAACCTTTGACTTGTTTTTTTTTAATAAATCTTT
TTTTCTTTAAATTAAAAGAATAAACTCTAGAGTGGAAACCCCCTGATAAGTAATAATGTTTCAGTTCCGA
CCCCAAAGTAAAGTTTCAAATAATTTAACCCAACTTTATTGTAAATAAAAAAAATATTTCGATGAAACGA
ATATGTGCAAAATTTCATATATCCATAATTCATTACGATGTGTTTTATCAAAAAAAAAATCTTTT

CA0105 (SEQ ID NO:66)
GGTTAGTCATCCAAGAACAGAAAAATCTCTAAAGAAAACAAACAAACCTTGATGTAATAAGCCTTGCGAA
TCTCTTCCTCAGAAGCGGAGGGAGTGACACCAAGAACATCATAATATACTGTTTCCTTCACCATGATCGG
ATCAAAGCAAGAGATGTAACTTTTTTGTGTAGATAAGAAAATAAAGGTTGGTTTTGTGGATTGTGTGTGA
AGCTTTTAGGAGATATGGG[GAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA]GTTGGGACACAAG
AGAGGTAGGGTGTGTGATCTGCTTGAAGAAGCAAATAAAGAGATGTCTTTACAGTTATGCACTTTTGATT
TAATAAATAAAAAACTTTATAGCGGGGAGGCTACACTACACTTTCACCATCTCTTTTTAACTGTCCACTC
AATTGCTTTATTGATCTCATGCCTCCTTTTTTATTATTCCATTGCTTTTCTGTATTGTTGAACATACTGA
AGAAAAAGAAGATGATGATTTTTGCAATACGATTGTGATCTGGTGTATCTTATCATTTAGCCAAATGGAT
ATTAAGTTGGAAAAAAATTCAAAAATATAGGTCCTACCGGGAGTCGAACCCAGGTCGCTGGATTCAAAGT
CCAGAGTGCTAACCACTACACCATAGAACCTTGTTGTCTTAACTTTACTTTATTTATTTATAATGAAACA
TTATAT

CA0120 (SEQ ID NO:67)
TTTAAGCCAAAATGTCATAACACAACAAAATGAGACAATAATAACATTACTGTAACAAATACATAGTTTC
TAATTAGAACAAAGACTAAACCAGACCAAGAGAAAAGTCGACAACAACTTTTAACTCTGTCCTT[CCACC
ATCATCATCATCATCATCATCATCATCATCATCATCATCA]TCCTCATAACTTATTGTTGTACCAGAACA
CACCTTTCTTCTCACCTTGCCTATCCGGTTCAACATAGATACACTCCTTCGCCTCCCTCCACATCGCCTT
AACCACCGGCGTTCCATCAAACTGGTAATACTCTCCAAGTATCGGCTTTATCGCCTTGGTCGCTTCCATC
GCGTTATAATGCGGCATCGTCGAGAACAGATGATGCGCCACGTGCGTGTCCGTGATGTTATGAAACACCT
TGTTCAAGATTCCATAGTCTCTATCCACAGTAGCCAAAGCTCCTCTCAACCAATCCCACTCCGAAGAATA
TAGTGAGGCAGCGAAGGGTGCGTGTGCTGCAAGTAAGTGATCAAGACGAGGAAACAGTTGACAATCATAA

GCGGAACTCCGTAGACACAGACCATCGAGGCCACTCCTCGCGAACCAGCGTAGCGGTAGAGACCGTAACA
TACGGAGAGGACGCCAGCGTCAGAGATGTATATCTGGAGACGCTCGCGGTCGTTGTAGATGGGAGCGTTC
GGGTGGAAATGGCAAGCGAAACCGTCGCTGTAAGGTCTTCCAGAGACGTTGAAGGCTAAGTACAACGGCC
AGCCGAGCGTGAACTGGACGGTTAGCATCACCGTGCGTCCTAGCGGG

CA0163 (SEQ ID NO:68)

TGAGCTATTAAAACTACATTATCTTAACGTCATATCAATTTGACATTTGCTTAATATTCATTTTCTTAGA
ACCGCTTGTCAAAAGCTTTCCCAGTTTTCCATTAACTAAAGAATGCTATCAGGAGAAGTTTCTGAAATTA
GATCATCATCATCATCATTTCTTAGCAACTTTTCTGAGATTGAGATCATATATTATCATCACCATCATCA
CCACCACCATCATCATAATCATATTTGCTTAGCAAATTTTTCTAAGAATCGTATATTATAACCACAAAAT
CTATATTTACTAACTTACAAGATAGATCCCATAAATTTATAACATTCTGCGATTACTCATTCCCTATATA
AATAACGTTCCATCTATTATATCCTACATTAT[CATCATCATCATCAT]CACCATCACAAT[CATCATCA
T]CACCATCACAATCATCACCATCACAGTCATCATTTTTTCATAGCAAACTTACAATTCGAAGAAACGAG
CGCCAAAACATCCGACCTTCTCCAGCAAGACTGAATCCAAAAATCCGAAATCGACAACATCTCCAGCTCA
TCACGAACCCTAGGCAGCCACACCCGCACGAATTCGACATCGGTAAGGTACAATTCGTCGGAATCGTCTT
CGACCAAGTCGTCGAACGACATCCCGCTGAGCGCCTTATCAACCCCAACCGATATGACCCTAACGTTGCG
AGAATCTTCGATCAGATTCCTGAAGACCGTCTTGAACGGCGTGACGGAGGTCGAGCGCGATTTGGAGTAC
CGTGACAACGTGCATAGG

CA0221 (SEQ ID NO:69)

TGTTCCTCTGTTTCTTCAACCTCAAAAGCTCTTTCCACAGAACAGTGCAGTTCGGTGGACGAGGAGAAGG
GTCGTCGTCAAGAAACCAGCTTCCTCTGTTATTATTCTCCTGCTCCTTGGTGCTGTTGTGAACAGTCTCC
TCTTTCTTAACTCTACGAT[CTTCTTCTTCTTCTTCTTCTTCTT]CTACTACTTCTACTTCGTCTACGTC
GACGACTTTTGTTTTCAAGGTAACGTTTTGAAGCTTCTCCGAGTGTTTGGCTTGCCAGAAAGGAAGCAGC
TTTCCTTCGGAGAACAGCTCGTCTGCGGCGGTGAGCATCGTTTGTGTGTTCGACAGAAACTCAAAGTCTC
CAGCTTTCACTTGTTCTTCTTTTCCCCTTAGGAGATTCTCAGGGTTGATGCAGATGTAGTCTCCCTCGCT
GTCTGATGATGACAGATCGGCGGAGAAGGAAATGCGAGGTCCTTCCGTCGTGAAAACCATCGTAGCCTCC
GCCGTTTCCGCTACTACCATGATCGTACAAATGTGTAGTTATGAAGTGAAAGACAAATCAAGTGGA

CA0226 (SEQ ID NO:70)

TCTTAAATTTAAATTCTATTCCTCAAACACTAAATCTTAAATCTACACTCTATATCTAATACTCTATACC
ACAAATTTAAACTCTATATACAAAATCATAAACTCAATCTTTACAAACTTATAATAGAAACTTTAAATTT
AAACCTAAGTATATTATAAAACTCAAATTATATACTTAATCCTAAATCTTAACCCTAACTCATAAACCAC
ATATTTCATAAAATATTAAATCTTAATTTTTAAATAAATTATAGTTCCATAAATAAATTAAAATTTCAAA
TAAAAAATTTAGATTTTAAATTTGAAATTATGATATCAAAGTATTTAAAACTTAAATAATTTTTATAATA
GCTATAAATAAATAAGAAGATAATTTGTATTGTTTTTTTATACCATTGATTGATTTGAATCAAAGACTCA
AGATAGCTCTTGTATCTATTTTCGCCTTTTTTTCTTATCGGTAGTTGTTGTTTATGGCATGGATCACCTG
CACCCTTAGATAATATTGAACCAGAGATTAATTGTTCTTTTATTCTTTTTTTTTTAATTTACTTTTCTCA
GATCTACGAAA[GAGAGAGAGAGAAGAGA]TGGAGTTCAAGGTAGAGAAGGAGAACGCGACGGCTGTTCG
TC[ACCACCACCACCACCACCACCACCACC]ATCGTTCGTCACTACCACCTTCGCTTCTCAGATACGTCT
TCACCGGAGTCGCCAGAACCACCGTCACGCTCGTCATAACCAACATCGCTCGTTCCCACAACCACCGCCA
TCTCTCCCATCAGCCATCGCTCGTCACCACCACATCACTTCTCAGATCCGTCTTCACCGGAGTCGCCACC
ACAACCGTCACGCTCCTCATAA

CA0233 (SEQ ID NO:71)

TTTTTTTGGTTCTTTTAACTTTTAATAATTAATTCAATAATCTGTAACCTCTCTATGTATCTCTTCCTCA
TTGTATCTGATTGAGTTTGAATCAGTTTTTGAGCAGGACGTGGTGATGCCAGAAGATCTAGCGAATGTCC
TTAGGACAGCGAAAGAGATTGTCGTTGCCACAGTCCTTCCCGTCACACTTTGCTTTTTGTACATCTCTAT
GCCTTGACGCCGCCGTAGGAACACGCTATCAAAGCTTGCGGGGTTGGAGGGAGCTCGACGACGGAGGAGG

CAAGCATGGTCTCGTCAGATACAACAGAGTCAGCATTATCTCAACATTCCGGTCATGTAAGCGGTGTAGT
CGTTTATACAAATTTGATTTTCAGATCTGAGATTCGCTTTTTGACTTTACAGTTTTTGTGTATATTTTTG
TAGGATAGGATGAAGGG[GAAGAAGAAGAA]GGAGGAGACGAAGACGAGAGGCTGTGTGGTTGGCGTCAG
TTATAT[CACCACCACCACCACCACCACCAC]TACCGTCACGCTTGTCACCACCATGCTCCTTTCCATTC
GAGGCGGCTGGAATCTTTTTTTTTCTAGGTTTAGA

CA0328 (SEQ ID NO:72)
ATGGTACAAACAATCATACAGACACGCATTCTCTTCTAGTTCAGCTGCTGCTTTCTCTAAGCTTGTCTTC
ATAAAATTGTTACAAACTTTGGAGCATCCTTCCTTGGGCCATCAAGGGACCGATACATGTACAACGTCTC
CACATAATCTGCATTGT[CGTCTTCGTCGTCGTCGTCGTCTTCGTCGTCGTCGTCATCATCAT]CTGACC
CAAACACCT[CCACCACCA]CACTAGGCAGAGTATGAGCAACATCCCTGCAGGCTCCCCGGGACAACCTA
CATGACGACATCCAAACAAACCTCATCTCGTTATACCTATGCATACCAGAGCGCAGTCCAACATCACCAA
AGGGACTGTCCCTGATCTCAAGCTTCTCTAGTTTAGGACACCCCTCGAGGATATATCTCAGACCCATGTC
ACTGTCCCCTGCAAAAGCTACAGATAGAGTACGTATCAGTTTCCCATACTCTCCTATAAGGCTAAAGGCT
TGGTCCGTTAGTAATCCAGATACTGCAAGCCTGGTTAGCTTCTTGCAGTTTTTAACAATGGCGCCAAATC
CATCGTCCATTGGCTTCCTTGTCACGTGGTCAGGCCTATGGCGACCCATTATGCAAAGCCTAAACACGGT
AAGCTGGGGACAGTTCTCAGACATGGCTGTCACAGCTACATTTGTCATCCGCTGGCAGAAGTAGAGAATA
GACTCAAGTTTCTTACAACCTTTCTGAAATTGCTTGGAGGCCTAATCCCGAGACAGGACCTTCACTGTCT
TCACTAGGATCCAAAGGGAAAATCCCTAGCTCACGGAGCTCCTTGCATG

CA0410 (SEQ ID NO:73)
CGCCGCGTCTCCTCACTCTCCGGATTACTCTCCGTCTGAATCCTCTCCTTCTCGCTCGCGATCTCCCTCT
CCTCCTTCCCGCGACGCTCCCTACCGTCTCCGATCGAAAGCCGCCGCCGCCTCCGCGAATCAAGGAGCTG
GTGGTAATCCATCGGGAAGCCGTACTACTAGGAGCCGTCAACAAGCTGGGAACATCCGTACGTTCGCCGA
TCTGAACCGTTCCCCCGCTGACGGCGCGGATAGTGATTCCGACGAAGGCCAAGAGTACTATACTGGTGGA
CAGAGGAGGTAAAATTGTGTTTATATTGAATGATCATAAACTGAGTAATGTGGAATCATGGAGAATTGTG
CTATTGATTGTTTGTGTTGGCTTCTCTTTAGCTAATGGATTGGGCCTTGTGTGTTTAGTGGGATGATGGT
TCAAGATCCTACTAAGAAAGCAAAAGATGTTGATGCACTCTTTGAGCAAGCTAGGCTTTCAGCTGTGGAC
AGGCCTGTTGAGCCATCGAGATCAGCTTCTACAAGCTTCACTGGAGCTTCTAAGATGTTATCTGGTGAGC
CTGTTCCCTCTGCTACTCCT[CAGCAGCAGCAGCAGCAG]CAAGACCAGCCTCAGTTGGTTATGCACACC
ATCACTTTCTGG

CA0423 (SEQ ID NO:74)
CCCTATTTCGCTGAATCTGCTTTCTAACCCTAATTTTCTCGATTTTTCTGCTCAAGCGTGTTGGCAATGT
CGGAGGACATGGTGATGCATTTCTCCTCCAATTCCTCCAATCAGTCCGATCACTCCCTGCCCGACAAAAT
CGCGAAGCTCGAGGCTCGCTTGACCGGCAAAACCGCCTCCTCCGCCAAGCCGCAGCCTC[AGCAGCAGCA
GCAGC]TCTCCGTCTGGTCATCTGCTTCCGCCCCTGCCAAAGTCGCGGCGGGTTCGTCGGATGTCTCTAT
CAGTGATTCCGACGACGAGGTAACTTCCGATGATTTTTTTTTATTATTTTTTTTTTTAAGATTTGATGTC
TAATAGTATTCTCGTTGTTACTACTGTCTCAGAACACAGGAGATTTCCTGATCCGAGCAAATACCAAGAA
GCGCCAGAAAGTTCAAGACTTTAACAACAACAACTCCACTCTTGTTGATCATGCTGAGGTAGTGAATTTT
CAGTTTAAATATCGATCTTTTCGTCCCTTGCCTGGTTCGTAGTTATATTGATATGGTAACTAAGGTTGTG
CGATACTGAAACAATCTGATATGATGCAAGTTTTGTATTCCCTTTTGATGAATTATTATAATGTCGAAAT
TGAAGCCGCAAGAGGCAGCATATGATGGAAGGAAAAACGACGCTGAGAACCAGACAGGCGTCGATGTGAG
TAAGAAGAAGCAAGGTCGAGGTCGAGGTTCATC

CA0456 (SEQ ID NO:75)
CTTCTGTTAGAATTCTACCG[TTGTTGTTGTTGTTGTTG]TCTTGGTTGTCTTAGAAGC[TCAATCA
A]CGCCTCCGCCTTTAGCTTAGCTCGACTCTTACTACTCTGCGACAAAG[CCACCA]CAGGAGCAATCGC
ACCTTCTCGCGCCACCATGGTTCGATACACCACACT[CTCCTC]ACAAAGCTGCAGCAATATCGACACGC

CCATCTCCTTCTGCCTCTGCGTTCCCACCTCCACTATCTCCACAAGCACCGGAACTCCTCCTTCCTCCAC
CACCGCCGGCTTCGACTCCGGCGCCGACATCAGCAGATTCATCACGTACGCCGATTTATCCACCATGTTC
GAATCGAAATCCGCCATCAGCTCCACGAGCGGCTTCATAACTCCCGATTCCACGGCCCTGGTCTTGTTCT
CCTTGGCCGAGCAGAGCGAGTAAAGAGCCGTCGCCGCGTCCTTCTTCCCCCTGAACCCGCCGGTTTCCAG
AAGGTTCACCAAGTGAGGAATCGCTCCGGATCTCCCGATCGCGATCTTGTTGTCTTCGATCTGCGATAGG
CGGAGGAGAGCGCAGGCGGCGTTCTCTTTCGCCGTCGGCGTTCCCGATTTCAAAACCCTAACGAGCGGTT
TAATCGCGCCGGAGGAAGCGATCAGCTCCTTGGTCTCGTCGCAGAGGGAGAGGTTCAGCACAGCGGTG

CA0488 (SEQ ID NO:76)

GCCAAGCTCTCCGACTTGTCACCGGTGCCAAGAAGCAGCCACTGAGAAAATAAAACTCATTCAAGATTCA
AAATCCTTGTGTGCTTCTTCAATGCCATTTTAGTTTGACTTCTTCATTTGCTACAGTTCATTAGTTATTT
CCTTATTTGCAAAAGAGCCCTCGAGTTTGTTAGAAACGTGAAATAAAGCCATTAAATACCAATTCCCTCC
ACTTTGAAGGGGTTTTGAATATCTTTCCCTCGACTCCAAAATCCTCGCCGGCGATAAGCAAACCCTAGAT
TCGATTCGCCGTCTGTTCATCCAGCAATGTCGTCGTTCAATCCATTCTCTACCCCACAGCGACATCAGCA
GACGCCTCAGCCGCAGAGCATCTCCTTCTTCTCGCCACCGCAGAGCACTCC[CTTCTT]CTCTCAACTG[
CAACAA]CAGCAAACGCCGTCGTTTCAGCCGCACCAGTTCCAGCAG[CAACAACAACAACAACAACAA]A
GTCAGCAGCAGCTGTATTTGTTCACGAACGATCAAGCTCCGGCGAGTTACAGCACCGAATGGGAGTGATC
TTTCATCCGATTCTCAGAAACTTCTCCTTGAGATTGAGTATTGCTCTTCTTTC

CA0546 (SEQ ID NO:77)

ATCAAACCATAGAAATACATGACTTACTACATTCACTCATCTGTGTTGGAAAACTTTTCAAATTTATTAT
ATCTTAATTTATATTATTTACAAATGTTTATAATTGCATGATTTCAATTATCCCCCATCACAACATATTT
TAAAAAATTTAAAAATTATTTTTAAGATATACAATATGAGAAGATTTTCAGAAGGCTTCTATGAGTATGT
TCTTAAAAATACATTCTATTTTTTTTTTTTGGTCTAATGGACTATTTATAATTTCAGTAGCATTTTAGAT
TAATTTTGCATTTGATCCATGAGGTATATCTTTGTGTTTAAAACCAAGTTTTAGGTTATATTTGGAAATT
TCCTCTTGATAGTTTGAAGGTTTGAAGTTTTGATGCGGATAGCAATGGATAATAAAACGGATTTTGGATC
TAGGACAATAATTCGTCCATCTCCTACGTGGGGTCTTTAGTGATAATGAAAAAACTCTTCTGGTAAAAAC
AAAATGTTTTAATAAATATGGGGCTCATCCATAAGTGAAAAATACCTCTCTTCTTCACTGCAAATGAATT
ATAAACCCCTTCCTTATCCACACACACACAGACTTGTTCGCTCTCTTAAACCCCT[GAAGAGGAAGAAGA
AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA]GGCGAATCATGCAGATTTGCCAAGCAGCGGTAACC
TTCACCTTCACGAACCCAACAAACCCTAATTTCTGCAAACCCAAACCTCTCTTCCCAAGCTTCCAACCCC
CTCGCCGCGTCGCCTTGCCGCCATGCCGTGGCTTCAGCTCCGACGAGTTCCCCGTCGACGAAACCTTCCT
CGAGAAATTCGGACCCCAGGACAAAGACACAGAGGACGAAGC

CA0552 (SEQ ID NO:78)

GATTTAAAATGACAATTTTTTATTGGTTGGTTCCGCTAGGGGGTGAACCAAGAATAACT[CTTCTT]TTA
TTTCTACGTTCTCATTCTTT[CTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT]CTACGATTTTTAATT
TCATTCAATGAAACAACAAAGTAGATCTGATTTTTATTTGAGTTTGGGTCCAAGAAGTGAAGAAAAATAT
TGGAGAGGAGGATCGACGCCTCTGTTCAATAGCCATGGAAACTCTGTACGCTTCCTCTCAAGCTCCGTGG
AGAAGAATACCAGGAAGACGGCGAGCAAGCTCCGTGGAGAAGAAGACCACGAAGACGGCGAGCGACTATG
AAGGTGGCTGGATGCAACGACAAGTACGCCTCATCTCTGAAGCTTGACTCTCAAATCCCACAGCGAAGAA
GAAGAATCAAGGCGGATGTGAAGAGCATAGAAGAGACAACGACGAAGGTGGTGGTGGGGCTGGTGAGGAC
TGGAAGATGAAGAAGACCACGTAGGTGGTGCCATGGTGGTGTGCGGCGGCGTACAAGATGAAGAAGACGA
CATATGGTAGTTAATTAGAAATTAGGTTTAAGCTTGGTTTAGGGTTTTGGTTTATTTGGTTTGGTTTTAG
TACTTTTTTATCTAATTAGATTTTTTTTAATATTTTTGTAAAACAATTAA

CA0603 (SEQ ID NO:79)

CATCATCATCCATTCATCATCATCATCATCATCATCATCATCAGCCACCTTTCTACTTTGTCTGTTT
CACAACTGCCCAATCTACCTCCCAAAGTCGTCTCTTCCATGTGATACACTTCGCTTGGCTTCTTCTGCCT

TTAGCTGTTCCCGTCTCAACGTTCGTACGTTAAAGGTAAAGT[CTTCTTCTTCTTCTTCTTCCCTT]
ACGTATTTTCGTTTTCCATCTAAAGATTCATTCTCTCCTCCGAGTTTCGTCCCCTGTCTACTCTGTTTCT
GTGATGTTGACCTCTCTCTTAAGCTGATCTGATATGTGTTCTTCTTCCTCTTTGATGCTTCTGTCTCTGT
AATTCTTTGACTACTTTAGATATTTTATCTTATGGGTTTCATTAAACTCGCACAAAGCTCGTGACTTTGA
GTTATATAACCAGTTCAGCTCTATTAAAGTTTTCTTGTAGACCAAACACTCATGAGTTACAGTGTCTTGT
TCTTAATCTTCCTTTTGACTATTTTATGAAAAGTTCTTGATCTTCGTTACTTTTCAATAGTCTGATTC

CA0614 (SEQ ID NO:80)
CGAGGGACACGAGGATAGGACCTGGGATGCCTTGAGTGCAGGCTGTGCAGCACCATAGGCGGGCACCTTT
TTAATATTGTATGTTGTAATATTTCATCTAAAATTATAAGATAATAGGGTATACATAACTTATTTGCGTG
TAAATAGATCTCATTTCTACATTTGAGAATCATGAAAAATATATATGTTCCAAGTGGTTCTGCAATGTGT
TAAA[TATATATATATATATATATATATA]GATATTATTTTCAATTAATATACTCACAAAGTTGGTTA
TGACTTATAGTACAAAACAAAATGTGGAGTTCATTAACTACACGAAACCCATTTGTCCACAATATTGAAG
TAGTCTTTTGTGATGATTGACATAAATTCTCATTTTAATTGCCCTTTATTGGGATAGCTGACAACAACAA
AATAAGTAATTCTTTTCAGATTTGAGAAAATTTCAACTACATATGTAAACAATTCAAAGAACATAAATAT
AATAAGAAATGTGCACAAAAAAAAATATAGAGATATATAAGAAATAGGTAAATGAGGCCAAAGATTGTTG
TTATATAGAAAGCAAGTCACTGCATCATAATATCATGTGGTGGTTCAACTTTATGACGATAGTGAATAGG
TCCTCTCTTACGCCTTGAGATTTTGTTTCCGTTGAAGCTGCAGATAACACACCTACACCTATATCTTGGA
TGATAGTGATGATGATAATTATGATGATGAGGATGATGACGCAAGTGATGATGGTTTTTCCTAAGAGCGC
AACAAACCGAGCTGTCTGATACC

CA0636 (SEQ ID NO:81)
CTCGGAAGAACGAATCGGGTCAGCTCAATGCTCCATCGGGTCAACCTTATCTCTACTCGGATCTCTCTCT
CGGGTCGAACTTCACTTGCGCGACGCGAAGCTCGAACAGCTCTGTGACTTGCTGGGGAGGAGGAGCGGAG
AGGTTCAACAATGTAACCGAAAAGATCTCATTCGAGTCAGTTACATCCGGGTCGGGTCTAATCTGCGGGT
TGATATCCGGTAACCTCTCGGTCATGTGTTGGAGCCCTAATAACTTCTCAAGAATCTTCCTTCCTTTCCC
AGATATCTTACCAGGTCCTTGCGTTGAATCATCTATTTGCAAATGTGGTGTGTATCCACGATCTGATCAG
CTATGCTCCGGCTCGGGTTCGATCTGCAGCAAATGCAAAATCTCA[CCTCCTCCACAACCACCATCACCA
CCACCACCACCA]CCGTCAGATTCATCTCCATCTCCATCTCCGCCGCCGTCGAAGGCGTTAACGAGAGGA
TTACTAGCGTTTGCGATCGTTGGATCAGTAGGAGCGTTTGCAGGGATATGCAGTGTGGTGTACTGTTTGT
GGACCGGAGCTTTCTTGGGGAAAGAGAAAGTTCATAACTCGGTTCAACCGACGATAACCCGCGGCGGTTC
GAGTACCCGGTCAAGCAGCTCGCCGCCTTCTCGGTCCTTGACGAATAGACGTCAGGGATCGAGAATATTT
TCGATGAGAA

CA0681 (SEQ ID NO:82)
TATACTGCTAACTATTTACATGTAAAATCTCTGTCAATTATCTTTTCCTATTCTATACAATTTTCCACAG
TTATTTTTGTAGTTCTGTTGCTGAACTTGAAGCTGAGTTTGTGGTGAAGAAACAATAACACCAAACACAG
CATATCACCTCCTTCCTTCTTCTAATGCATCTCTTCTCCTCTTCAACCCCACACTGGATGCGAACCTGTT
GTTCTTGCTTCAATCGTTTGAGCTTGGATTTTTTCAATCTCCATCTCCCCTCCTCTTCCACCAACATTAC
TCTCTTCATCCCCATCCAATACCAACTTCTCAAACCTACACTGTTTTGTTGCAAAACACACAGTCTCAGC
CACATCATTAAAAAAAAGGAAGAGG[AAGGAAGGAAGGAAGG]AAAGTTTTACATTTTCACATTCCCATA
AGGGTAGGTGAGACGGCTGATACATTCGCAACTCTGTTTGAGACACGTACAGCTGATGGTTGTTATCCAC
CACACGAGGGGCCTTTCCTTTCTTAGTTGCTTTCCACGGCTCCTCAGCACTCTGCTTTATAACAACCTCC
GGAAACACGTTACTTCCAGAACCCGAAGTTGATCCACCTCCTCCATATACGTCACTCTGCTGATCTTGCA
TCTCTCTTCTAGATCGGTTCTGGACCATGTACCACCTTCGTATCCTCACAACCGGTCCTTCAGTCTCTGC
ACCAGACTCGGAATCAATCCCCGTCATCACAAAGCCTCACCACCATCCCCAGCCTTCAACGCGTACTGTG
TCATACATCCAGAAGGAGCAAAAACAGTAGATTCCTCGTTACGCTATTATCAGAAGCAGTA

CA0719 (SEQ ID NO:83)

GCTACATATGTCCCTAAAGAGTGAAACTAAAGCGAATTGCGAGGATTATTGAACAAGTTTCCTTCCAACT
TTCTAACGAATCAGCCATCATAGTAGCTCGCAATCAACATTTAGTTTCTGGGAAGATGAACAAACACAAA
TTACCCAAGAACACGAGACACCCAGAACATAAACAAAATCAAATACATCATGAATCCGATTAAAAAGAAC
GAAGATGGAGCAAAGTACCTTTTCTCGATTCGACTTGGAGAGAAACTCGAACGAAAGGGAATAAAACCCG
AGGAGTGACTTAATTGGGTCACATAATTTTGTTAACCGGAAAGTTACCGAACCGGAATCATACAGCTCGT
TGTGTAGT[GGTTGGTTGGTTGGTT]TTACAACTTCCACAGACTAAAAATGACATGAAAAATTAATCAAT
TATTTACCTGAAATGTACGATTAGCCAACAATTAGTTCTGTTATTCATAACAAAAGAAAACATTTTAATT
ACAGAGGTGAACGTATCCTAAAGAGAAATCTTTTTTTTCAAAACAATTAAACTTCCATTCATTAACATTA
ACCATCGCAAATACAAATCAAGGTCCAATCACACATATACGACTCAGACTCAGGATCTGACTGGTTCAAA
CGCAGC

CA0736 (SEQ ID NO:84)
GCTTTTATCCCGGTTTAATAACTTTGAAATTCAATCCATCGACGTGGTCGGTTCATGGTCGAACCAATTA
CTGGACCCAACCCGACTATACAACCGGTTCATGGTCGAACCCGGTCCAACCATCGGGTCGGTCCGGTTTT
AAAAACACTGCTCTAAATGGAAAAATATAGTCCTTTTTAACTATTGTTTAATTCAAAATCTGTTGACTAT
AGTGATGGATAATACAATTTATATGGGATTTGAATTTATGTATTAGATGTAAAAATTGAAAAAGAAAACA
TATTATATACGCTCTACGAGCTTTTTAAATAGATTTATTGGACCTAAGTATTTCATAAGTTTTGAAAACA
TGGGCTTACAAAACCTTTTTAACGATGTTCAACCCGGGCTTAGACAAACTTTATGAGATCACGGCTAGGC
CTTTGAGTGCTATTATTTTATTTTATTTATTTCTTGAATTTTAGGGATTAATAAATGTGAGAAGGAGTAG
ATAGTACATAATTAGAGATTGATGGAACAAATTGCAATAATTTAAAAGTAAAAGGATTTAAAATGCAAAA
AAAAATATGAGGACACATGTCAACAAACCCTCCTTCTATATGTCATAAGAAGGGAAAAAATCAACTTTAT
ATAT[ATAGATAGATAGATA]GATATGTATAATTTGGGATAAAAGTCGGTATAGCCGTACAGGCGTTTGT
GCGATCAATCGGTCATCAACTAAACAAAATTTTAAAATGATTTTTTAAACAAAAAAAAATATTATTTAAT
ATTTATTAAATAATTTGCAATTTTTAATAAAAAATAGTTTTCATATGGGATAAAATTTATCAATCTCATC
TACTATATAAA

CA0739 (SEQ ID NO:85)
GATCTTATCACCAAATTTTATATTGTCACGTTTAAGTATATCTTTCGTAGAAACATATTTTCCTCGAAAT
GAAACTATATGATATAATATTTTTTTTGTCTAAACATTTTTATACTAAAAACTGATAAGATTATTGTTGG
TAACTACAATTATATTTACCTTGATAAATATATAAAG[ATATATATATATATATATATATATATGTATGT
ATGTATGTATGTATGTAT]ATATATCATCTTTTATTGAATTTGGATAATAGCAGATTAATTAATATTTTT
TAGATAATGATAATATATAATTAAATTTTGATTTACTCAATTATTATTTATGCAAGTTTAACTTTTATTT
TTGGGTGATTTATTATTGTATATGAATATATAAATATATTATGAATAAATATAATTATCTAATTATTAAG
CATTATAAATATAATTATTCATTAAATGTAAAATGACTCTAATTACTCTAGTTTTTAATGCGATAGTTCA
GATCAAAAATATCAATCAGAAACTAATAATATCACAATTTTATATTAGAGTATATTTGTTTAATTAACTA
ACTAATACAATCTGTGAATTTGTATCATTACCAGAAACAACCAATTGAGCAAGTCGGTTAAAAAGTTCAT
GCCATGTTTTAATTTTTGAGCTCATACATTTTTCATTTACTCAGGATTCAC

CA0753 (SEQ ID NO:86)
GAGTATCAGCTCTACAACATCGTCCGGAAGCAATTGCAACTCTTCGCGATGCCTTTTCAGTGTTCTTGTT
TTGAACGGAAACATCCTTCTCGTCCTCGTCGTCTCTAAATTTTATTGAATTTGATCAACAGATCTGATAT
ATATATATAGATACAGGCAGCTAAGGAATCTGGAAACACAAAAAAAAAAAAAAAGAGAAAATTTCCTCTCC
GTTTAAGTAAGATTTCCTTTTTTGAATTTAAACAGAATCGAAACATCAAATC[TAATAATAATAATAA]T
ACAAAT[ATACATACATACATAC]ATATTACCTCAGACTCAGGCAATGAACAAGCCTTTTCTAATCCTCA
GGAATCATCCA[TCTCTCTCTCTCTCTCTCTC]GTTTTGTTTGTGAGCATCGATG[CGGCGGCGG]CGCT
TAGACAAAGACATCTCATCGGGACGCCTCTTTTACAACTCCTCCTCCTGTCTTCTTTTGGGCTTTCTGTA
AGGCCCGACCCGGTTTCCCTTAACGCCGGTACGTCCTTAGTTCGCTTACCTCGACCAAACTGGCCCTATC
CGAATTTATTCTCTTAACTTAGGATTATTATGCAATTTTCCTCTAAGAGGTTCAGTTCAGTACACAAGGT

TCGCTAAGTCTAATCCAGCAACTTAGCAGTCTACTAGTAATCGCAGCATAATGAACATGTACCTACTGCC
TCTGTACTTTGGTATCTGATAATCCATCCATACACTCCTTCA

CA0834 (SEQ ID NO:87)
GCCATTTTCCTGATTCGAATCCAGATACTGCATAATAAATTTGAGAATAATAATCACTTTTTATTCACTG
CGACATTGTAAGAGTGACTGTTATTCATAAGGCTGTTACTGTTAGGGTCGAAGGCAACTATTATTCTTTT
TTTTGTTAAAGCCTTTATTCTTTCTTTTTCTTTTTGAATCTTTAGTTCGTAAATATTCTCTTTCATATTC
ATAAAAAATACACAACACAACATATGTATTACTATTAGAGGCATAACCATTAACATTGGATTTATTGAGG
TTAGTAATTATTATGGTTGTTTGACAACAAAAAAAAAAGTAATTTTTTTTTGAGCAAACAAAAAGTAATT
ATCTGACAATAGTAGAAACTAAAAAATGCAACCATGCAATACGTGGTTTTATAATCATTCTATTGTTAAA
TATGATGAT[AATAATAATAAT]GAT[AATAATAATAATAATAATAATAATAATAATAATAATAAT]
ATTACAGAATGTTGATGTAATAAACAAAAATAGTTTGTTAGCTAACGCCTCAGATCGATCAATGAGTAAT
TCATTCAGTTACCACATAAAGAAACAAATAAAAACTATGATAAAAAAGTTTTGACATCATTTTTTATTGA
CATGTCAATATGTGATAATACACTCTCTGCAGCAGTGACAACAAATACTACAAAC

CA0837 (SEQ ID NO:88)
TTGGAGTCCTCCGACTTGGTCTTGATACAAGATTGTGAGGAAATCACTGTCCGTGTGTGGCATCAAACCA
TACACCTCCGATGGTTTCGGACATGGTGGATAACGGTTCATCCTTAGATAACATGTGTTTCGCACACAGG
TTTTTTTGAAGAAACTTGATTTCCGTCCTGATTTCTCTGCAAGGACCTCTGCCAATGAATATGCCAGAGC
CTCGGATTCTGAAGCAAAATTTTCCATTGTTGAGCTGCGATATAAAATGTACATATTATAACTAAGGTTA
ATTTATTATAGAGACAAATAAATCATGTTAAATAAATTAGGTGAAATAATTGCGAAAGCCATGAACCTTT
CTTGTTCCTTTTGTTTAATCCAAGCCATGAATTGTTCATTTGTTTTAACGTAGAATTGCTAAGATTTTTT
TTTTTGTAAACCATGAATTGAAGTTATGATAAGAAAAGAAATGGAAAAT[ATTATTATTATTATTATTAT
TATTATT]TAATGGTAGAATGATATAGTATAAATAATTATTTCACGGTAATTATTTGATTTGGTAGAAAA
TTGCGGAAATTATTTGATTTGCTGATTTTTTTTGTGAAGAACAAAATTCACCTAACAAAAGAAACACGTA
AGATTATTTGTTATTTGTGGTACATAT

CA0896 (SEQ ID NO:89)
TAAATTTACATGTAAATTACCACGAAACATTTTCTTTGTAACTTTACTACGACCTTACTACGAAATTCAG
TTTTGTCGTAAAATCGTAGTAATTTTCTCGTAAATTTACGAGGAATATATTTCCTCGTAATTTTTCCTTG
TTATAGGCATGTTTTCTTGTAGTGTTTATGTTGCCGTTGTTCTGACCACGATAGTTATGATACCTTTGTG
ATTTTCTGGTCACATATTCACTTAATTATTTTGTATGCTGACATACCTCATGGGAGGTTCGCTTGATATA
AATCATCACTTACAAACAAAAAATATTCATAAAAAAAAATATTCACACGTTTACAAAATCAAAAAGAGTT
ATATATAAATAGCTAT[AATAATAATAAT]GATACTAATATTAATAACAGT[AATAATAATAAT]GTTTA
GAAAGCTAAACAACAAGGATTAGAACATGTATTTTTACAATTGCAAAAACAACAACAAAGTCGTAGCTTA
GGACATTTAAAACAAGATGACCATTTGATCTTGCAGTTGCAGCTGCAACATGAGCTCTTCTTATTAAGAC
ATGATGGTCGACTGCAACTGCGGAAACATGTGGGTATGCAACAACATAAGTCCGGACAAGAACAACAGCA
ACCTAGGAAGCATGAACAGCTCGGGCAGCTCAGGCATTTGGGGCAGCTCGGTTTCGGGCAACAACAGCTG
TTTGAGCAGCAGGACCCGTTGCAGCATTTGGAACAACGCAGCTACAACATGTGGAAGTGCAGCATTTTGG
CTTCCTCAGATGGCACGAACACTCGGGTCGG

CA0991 (SEQ ID NO:90)
GGCCGACGGCGATCTGTATCATCGCCCTGTGGTGGAATTCCTCCTCCGAGGGTATAGAAGGAAGCTAACT
TGGTCTGCTTGTGAAGCTAAAATGAAAGGCTCGAATTTGTTGTACCTTCGTCCACCGTTGACATCAATAA
CACCGAATTTCTTAGACCGAACACCTCTGTTGACGACGGGGCAATGGAATGTATGAAAAACTTAAGGCAG
TTTCAGGTTACGTTTGCAATGGATTATTCTCAATTGGTGAAGATGGTTTTCTGAACCAGATGAATGACCA
GCATTTGAAAGCTACCTGGAAGATATTAAGCTTTTGCGACGAAGTTTCGTCAACTCAGATATTATTCATG
TTCATAGGGCGGAGAACATAAGGGCGGATAGCTTGGCACACGTAGTACTCAGAAACAACCGTCTTTCGTC
GTGCATATGGACGCAGAGTTGCCACATTGGTTTACAGAGTCTACATGAGTCTGTAAATATTTGCTGTTAA

[AATAATAATAATAATAATAAT]ATATATCTGTCTATCAATTTTTAAAACACAATAAGTTTACGGTATAT
TTTTCATTGAATAGATTGTTTTCAACTTTCACATGTATTTGTATCTTCTTCTATATATATATTTTCAGAT
TATTATTTCATTATTANAATCGTAACAATATGTATAAAAATTAGTAAAATATTGTTTTGTTGTCATATTC
AAAGATA

CA1027 (SEQ ID NO:91)
AGCGGCGATCTGATTCTCGCCCTGTGGTGGATTCTTCTTTCTTTAGTCTTTCCATTATTCTATGACGGTG
TAATTCCCTATATATAAAAGGCTCCTTATATTTATGAATAATATAGAAACATAGATTTCATTACGACTAT
ATTATTAGTATATCAGTCTAGGCGTTTACCAATACCAATATACTTAATATATTTAGTATAATATCTTATG
ATTTACAATTATTTTCATATGATTTTGTACTATAATATGTCAATTATTATAATTTATAAAAAACTT[ATT
CATTATTTATTATTATTATTATT]ATAAACCTACAACCTTTCAACTTAATTAGAATTCACAACCTTTAGA
ATTAATTGAGATTCTTATTATTAATAGATATTATAATCTTTTAAATGGTATATAAGATAATCACCACGGT
ACTAGAAAGCCTAGAGCCAAAGCACCGCCTAAGCCGCCGCCTAGAACAATTACCTAATTTAAAGAAAAAC
TAATACTTATATTTGATTTTGAAATTTTATTAAACTTTGCAAAAAAAGAAGAAGATGGAAACATGTTAGA
AACATATATCCAAATATAAAAATATAAGAATAATTTTATAAAAATTAATGATTAAAAACATATGCAAGAT
TTCGTATGAAAAAACTATTCTGCACAAAAATAATTTATAATATTAGTTTAATATTTACATATTTC

CA1032 (SEQ ID NO:92)
NGGCGACGGCGATTGATTCTCGCCCTGTGGTGGAATTCCAGTTTGACCAGGACTGTCGTAAGCTTAGTTT
AATTTCACCAGCAGACACCGACTCAATAGCACCCTGGAAAAAACACACTCAAACCAGAAGCAAATGAACT
ATAATAGTCCAAGTAGAAGAAACACAATCAATCATCCAAGAAAAGATACTACTACATCACCAACAATACT
GCTAGATAATGTAAAAAATGGACAGAAGAAATAAAACTACACTGGTCTTCCACCGAAAGAGTCCAAATAG
AAACACAAGGAATAAAGCAAAAGAAAACTAAAATTACCATAGCACTAGCAAGATAATGTAAAAACCTACA
TTGATCTTCTACAGAAGCAGTTTGTTTTATTTTTTCTCCGTTTAGAGAATTTTGGGGTGCTTCTCACCTT
ATTGAACTTGACGACGACATCCCTGAGGCATTTCCAACCGCCAAAACGGAACACAACAGATGCTCCCAGC
ACTCGGCTAAGAATCCATGCAAAGAATCTTGAACAGAGTCTGGAGAGTCAAAATG[AAATAAATAAATAA
ATATAATAATAATAATAATAAT]AAGACCACTATAGCAGCATAGTCCAGCAGCTAAATCATGCAATCTCA
GCTACTGAAGGAAATTAGAGAATGTGCAAACCGAACTANAATCATCACTAGAACTAACTCACACGAAGAT
CATCCACAAGACCATGGAAAGAATCAGGAAC

CA1034 (SEQ ID NO:93)
AACGGAGACTGATATCTCGCCCTGTGGTGGAATTCTCTGACAATAGTAGAAACTAAAAAATGCAACCATG
CAATACGTGGTTTTATAATCATTCTATTGTTAAATATGATGAT[AATAATAATAATGATAATAATAATAA
TAATAATAATAATAATAATAATAATAATAAT]ATTACAGAATGTTGATGTAATAAACAAAAATAGTTTGT
TAGCTAACGCCTCAGATCGATCAATGAGTAATTCATTCAGTTACAACATAAAGAAACAAATAAAAACTAT
GATAAAAAAAGTTTTGACATCATTTTTTATTGACATGTCAATATGTGATAATACACTCTCTGCAGCAGTG
ACAAACAAATACTACAAACTCTTATTTTTAATCGTTCAAAGATAAGAGTCTATACTAGTAGACTAGAAAG
TGGGGGGAAACAAATAAATTTAGGAGGATTCATTGACAATTTAAGAAGACATTTTTGATACGCCTCGTCT
TATTAGAATTGGGAATGGCCTATGGAGAGGATATGAATGTGATGGGCATAGTGATAAGGTAGAGGAGATA
ATGCAGAAAAGCGAGAAGAAGAATCTTAAACTATCATTTATGAATTATGAGTTAACCTCAGAAAGCCAGT
TTACAAAAAAAAAAAATTATGATATCTCCACTCGTTTCTATTAACTTATTCCTCCATGATTGGTCGTTTT
TGTAAACTTCTGATGATTC

CA1035 (SEQ ID NO:94)
AGCGGAGATTGATTCTCGCCACTGTGGTGGAATTCAGTAGCTATGTGCTATTAGTGCATTGATTGTTCTT
TTGTGTGGTGTAATAGACACCTGTTTGTTCCATGCTAGAGCTAGGCCTAAATTTTTGTAGTGCTATTAAC
TAAGTCAGTGGTTTGTGGTTTAGCATCCCATACCTCACTGAGTGACTCCCTTATTGCTCACCCCTCCTTC
GTTCTCCCAGGTGAGACCGACAATCATGAGTGATTTTATCGGATTGGTACTTTTGAGCTTTTATCGTTAC
TGAGCTTTTAGACCTTTGGACTTTTATCTTTTATGCTATTTCATATTTCAGACTTTCGGTTTTATATTGC

TATCTATATTTCAGATGTTATCGGACCTTCTGATATTGACTTTTGTATTATGAAGTGGAG[ATTATTATT
ATTATTATTATTATTATTATTATTATT]ACTAGATTCCTTTTCCGCGCTACGCGCGGATAGTATCTTATA
AATTTTAAATTTATTTTTAAAAAAAAAAATATTTAAAGTTTAATTTACATTATTTTATACCAAAAATCAC
AAATATGGCTAAGAATTGGTTGATTTTATTTGTGATTTTTTGACAATATTAAATTGATTTATTTATTGCT
CATAGTTAACAGATTTGTTTGATTGGTTTTCAGTTCATAGTAATGTATAGTATTATATTTGGTGAGTGTA
TATA

CA1066 (SEQ ID NO:95)
TCTCAGCGCTGAGCTCCATGTGGTGGAATTCATCCTAAATCTCTTTGATGCGATTGATCTCTTTTATAAC
TCATTTTGACTCTTTTGGAATTGGAAGAGATGGACTGGACGGCTATTGCAAAGCCTTTGGAAGAGATGCA
AAGGAGCTGGCTATTGCAATCTTCTACGCTCAAGAATAAGAAGACATATGTCAAGAAGAAGAAGATATAT
GCCTGGACTCTTGCCTTCATCGGCGTACTTGTGGTTATTGCATTTAGTTTGAACATAAAGCTCTTAGGGG
CTCATGCATAACGCTTTCTTAATTAGCTCTGTTTTTTCCAACTGATGTTTACTCTTTCTGAT[ATTATTA
TTATTATTATTATT]AATTGTTAGTTGCTGTTGACCGGTTGAGGCTGTTTTTGAGCCCTGGATATCTTTC
TGAGTTGAGAAAAATTTCATAACCAAATCGAGATTGTTATGTGCTCTTTCTTGCCTCCTTTCAACAAGTT
TAATAGAACCAAAGGCAAATAGTTTGTCTTTATTCTATACTAGGATTGCGAATCCCCGCGGTCCGCGGGG
AAAAAAAGATGTTTTACCGCAAAAAAAATGATGTTAAAACTTAAATGTAATAGTAAAATTTTAGTTTGTA
ACAATCAACCGGTTGAATGGTTAATAATCAAATATTGCAATGTTAACTATTAAAAATTACAGTGGAACAT
TTAAAAGTTGTGAATATTTATATAAAA

CA1080 (SEQ ID NO:96)
AGCGGCAATTGTATCTCGCCCTGTGGTGGAATTCTGTGATGATCAAAAACAAGTTTCAATCCAAATATCC
TGATTTTGCAAATTATTTGACAAAATCCACATCTCTAATGGTGTTAGAACACTAATAAG[ATTATTATTA
TTATTATTATTATTATTATTATTATTATTATTATT]AACAGTCTTTCCAAAATAAATGATTGATAAAATA
TTGAAAAAGAAGGCAAAGAAGAGGTTGAGAAACAATTCTTATTTCAAAATTTTACAAAAATACAAATTGT
TCGCGTAACATTTTCATTTTCTATTTAGTTTAATTTTTGTCATTTAAAATTATCTTGTTGGTGTTGTTGG
CGTAAGCCCAAAACCGATGTAGCCTACACTGGGCCAATCTCCTGCGCAAGCCCAAGACATAAAGCATTAG
GGTTTTGTTGCTAGCTCATATGTAAACAAAACTTAAGCTATCTTGTTGCCTAAGGTTTTAAGTTTTCTAA
GATACAAAGCTTGTACATACACAAGCTAGATCATAGTTGTGATCACCTCTGTACTCTCTTATTCATAGTG
AAGTTTGGGAGGACAGTCTCCCACGAGACGTACCGGTTAGAGGCCGGGAACTCGTTAAATTGTGTGTGTT
CTTATTGCTTTAGTTTAATCTCTTCTTAAACAACCATAAGCATGATAAGAACTAGTTA

CA1090 (SEQ ID NO:97)
GGAAACGGCAGTCTGATATCATCGCCACTGTGGTGGAATTCTAACATAACATATGATTGATAGTACATTG
TATATTCTAATTCAAATTAAAAAATATAAATATACAAATATAAATCACCAATATAGTATTTTTACATATT
AATTTTATAATGCTTTATTCTAATATTTTCTTATACCTACTTATTAATTTTTAACTTATTAATTCTAATC
AATGGTCTATTCTGTATTTATTCTCATAAAGAGAAAATAAAGATCTACCAGAAATTGATATTTATGTACA
TTCATTACATGTACAGTAATAAGACCCACATAATCATTTTGTTTTAGCTATGGCTCGTGTTAGGAAAAAT
CTATAAGATGTATCTAATAGTGTGTTGCAATACTAGACTACCAGTTGATCCAGTTGTTTCAAATAATTAG
TTATGTTTCGGAAACTTTGTAGATTGGCTTATTTTCCATGCAGCTTTTTGTTACGACAGAACAAATTACG
CATAAACCTTTAGGCTGAGCAAAGTTGATGACTTTAACCAAGATTGAGTCTTGAATATGTGTCATTACAT
CGAAATATCGAATGTTAAAAATA[TAATAATAATAATAATAATAATAATAATAA]CAATAACGGGTTCAT
ATATCATTACAATTATAACGTAATAGCTGAAAATTCAAAATTGACTAAAATAATATTATGACCCGTCCCT
TTTTATGGTTCCCCCGTTCGTGTATTTGCATTGTTGGCCGTTGGTGATCCCATGNCGCACTTACCCTCCA
AGTCTTCA

CA1097 (SEQ ID NO:98)
TCTGGAGCTGAAGTCCATGTGGTGGAATTCTGATTTTTGAAAAAATTAAGCGTTATTTTTGTGATTTTTG
ACTTTGAGTGCTAATTTGGAAACAAAAACTTGATTTAGAGATATTTTTGTCTTTTTTTCTTCTCCAAGCG

TGTATCTTTATTTTTATTTTTTATAATTAAACAGGCGGCTTTTTATCCTAATTCAATTCAGGTGGGGTTT
TGTTCTTTTACACATGTCAATTTTGTTCTTTTAATGGTAAAAATTTAAGATAAATTATAAACTGAACCGG
AATCGCAATTGGTAATAAACTGAACAAAATTCCTAATAAGATATATTCCTGAAAAAATCCCCGGAAGATT
TTGATACGTATTAAAATCATATTAAGTTTGAAATATCAAGTTTTATATAATAAGATATACATTATATGCA
ACATCTTTGAATAACTCTCAACCTTTTGGTCATATCACAATAAAGTGGTGGAGCTTTTTCCAGTTACTGA
TGAATGAGTTAAAAAGTACTTAAGTTGCAATAATCTATTCATATTCCATGATCAAAAGCTCTTACAAGAA
ACAAAAGATTACATGAAAATGTCCAAAAGGGTACTTT[ATTATTATTATTATTATTATTATTATT]ATCA
GGATTGAGACCCACGTATACGTAATAAGAAAATATATAACTAATAGCGAATGCTACCTTATGTCATATAC
ACGTAAACACATCCCACTGGTCTTGGCAACACAAGGTGTCATCCTTCTCTTAAACATTCAAC

CA1107 (SEQ ID NO:99)

AGACGGAAATCTGATATCATCGCCACTGTGGTGGAATACTAAACAAGGCATATAGCATAATATTATTTCA
TGATATAAAAGCAAAAAAAAAATAGAATAATTAAATATACAATAAAAAAAATAAACAATAACTAAATATA
CAATAGCAAAAATGAAAAAAACTAAATGAAACATCTATCTGAAAAATGTATAATAAATAAATAATAAGTA
AATATATAATATGAGAAATAAAAATATTACACTAAATATCTATCGTAATATTAAAATAAAAATGGAGGTG
GAGGCGTTAATATGGGCAATGGAGTGTATGAGGAATTTGCGTCAGTTTCATGTCACGTTTGCAACAGATT
TTCCTCAATTGGTGAAGATGGTTTGAGAACCAGAAAAATGACCAGCATTTGAAAGTTATCTAGAAGACAT
CAAGATTTTGAAAGAAAGTTTCATCAACTCAGAGATCATTCATGTACCTCGGACGGAGAATTTAAGAGCG
GAGAGTCTAGCACGTAGTGTCAAAAAACATTTGTCTTTCATCGTTCACATGGATTTAGAGTTACCAGTTA
GGTTTACAAAGTCGGTATGAGTCTGTAAAAGTCGATTACAA[AATAATAATAATAATAATAATAATAATA
ATAAT]AAGTAAATATATGATACAAAAAATAGAAAAACTACATTGAAAATGTGTAGTAAAATAAATAATA
AATAAATATAAAATATGATGACAAAAAAATGACAAAAAAGTAAATATAAAATATAACAT

PE0012 (SEQ ID NO:100)

ATACGCCAAGCTTCAATCAAAGGAGGGAAGTGGTGAGAATACAAACCTAGCCTTATTCTCTCTATGTATC
TTCCTCAACTCTGCATCTTCCAACTGTGCCTGAAAAAAGAAAAGCAAAACCCATTAGACGCTAAGCTAAT
GCAATTTCGAGTTTAATGTTTCAGCTTAATCCACATAAAGACGGAAACATACCTGCTCCTGGGTGAATTC
CTCCGGTGCACCGTCGGAGTCTGAATCGGAGTTGTGATCTTTGTTATCCGACATCTGATTATTAT[CTTC
TTCTTCTTCTTCTTCTTCTTCTT]CGTTGCTTTCTTCTCCCTCAGACGACACACACTANGGTTTAAC
GGCTCTTCAGTTTCTCAAAAACAGAAGATTTCTATTCTGAGAGTTAATTGCTTCTCTCCTTTATGGTGGA
TTCTATTGGGAAGCTTGCATGCCTGCAGGTCGACTCTAGANGATCCCCGGGTACCGAGCTCGAATTCACT
GGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGGCGTTACCCAACTTAATCGCCTTGCAGCACAT
CCCCTTTCGCCAGCTGGGGTTANTANCGAAAAGGCCGCACCGATCGCCTTCCAACANTTGCGCANCTGAA
TGGCGAATGGCGCCTGATGCGGTATTTTCTCTTACCTCTGTGCGGTATTTCCNCCGCNTATGGTGCCTCT
CANTACAATCTGCCTGATGNCGCNTANTTAACCANCCCGAAANCCGCCANNNCCGCTGAANCCCCTGAAG
GGGNTGTTCGGCNCCGGNATCCGCTTTAAAAAAANCTGTTAACGTCTCCGGAACCGCTTTTTTCAAGGTT
TTCCCGTCTCNCNAAN

PE0017 (SEQ ID NO:101)

CGCCNNGCTTCCTTTAAGAGTTAATTTCATACAAATAAGCCCACTAAACAGCGTTGTTTAAACTTTAAAC
CCTGTCTTCTTACCAAAGCTCCTCTTCTTACATAGTAGCTCTGCTAAACCTTCGCCGCAATAAACCCAAA
GTCGTAAGAAACCGTCGCCATGCCTGCTCTTACGCGTAACAAGCAGAAGGGAGCTAAGTCGCAGACTCCT
CCACTGATTAAGCGGACTAAATCGAATCCCACGCCTCCACCGAAGAAGGCGATGAAGTCCCGTAAGCCTC
CGTTGAAGAAACAGAGGAAAGGTGTTTCGGATGAGAAGCCTGAAGTTTCTAATGATGAGGAGGAAGAGGA
AGAAGAAGAAGAAGAAGAAGTGAGTGAAGAGTCTGATGACGGGAGATGAATTGGGTTCTGACCTTTTCTC
AGATGGTGACG[AAGAAGAAGAAGAAGAAGAAGAAGAAG]ATGATATAGAGCCTTCGGATGACGACTTTC
TTGGTGGTAGCGATGAGGAAAAGGGAACTTTGGGTTCTGATTCTGACTCTGATGAGTCAGATAAGCTTGC
ATGCCTGCAGGTCGACTCTAGANGATCCCGGGTACCGAGCTCGAATTCCTGGNCGTCGTTTTACNACGTC
GTGACTGGGAAAACCTGGCGTTACCACTTAATCGCCTTGCAGCACATCCCCTTTCGCCNGCTGGCGTTNT

ACCNAAAAGGCCGCNCCGATCGCCTTCCACAGTTGCGCNCCTGAATGGCGAATGGGGCCTGGATGCGGTT
TTTTCCCCTTCCCCTCTGTTGCGGTTTTCCNCCGCCTNTGGTGCCTCTCNTTCATCTGCNCTGATGCCCC
TTNTTTANCCNNCCCGAACCCGCCANNCCCGTGAACCC

PE0063 (SEQ ID NO:102)
ATCTCTATCGCTATACTTGTCACTTTGTTCACTTTCTCAGCAGTTCCTTCCACATGGTGATGCAACCATC
TCGCCATTACACTCAGCAAGCCTCCCTTAGCCTTCTTGCATCGTTCTGACCTGTCGTTATATCCATAATT
CTTTGCAATTTTTGTCATTCT[CTTCTTCTTCTTCTTCTTCCTCTTCCTCTTCCTCTTCCTCTTCCTCTT
CTTCTT]TATGAACATGAGCAGCCATTTCCTGTCAACTCTTTGACGGATGTGCACCAATTGACAAGAGCT
TACTCTGTTACACCACTCCAACAAACTCTTCCTTCGTCTCTTAGATATTTTCTCCATTTGTATTCCCAGT
GTCTCAAAGTTATCATTATCCTTTTCGCTCAAAGTATCATCCAAATGCTTTTCAATATCTAGAAAAGTTT
TTTTCCATCTCTTTCTCTTCCTTGTTACTCTTACACCTCCCTCAACCATTTGATCGAACACATGGTGGGC
ATACTTCTCCTTCTTGGTCTCAAGCAACTCCTCTCTTGCATCNTTACCCCTCTCAAAGT

PE0091 (SEQ ID NO:103)
CGGCAATAATGGACCACTGGTTTGGTTGAACCACTAAAAAGAGTGC[GTGTGTGTGTGTGTGTGTGTGTG
TGT]GAGGGACTCTATTTAAAAGCACTGCTTAACTCAATAATTATTCCATCGCTCCAAAATAAAANAGAA
TAGCTAAAAGATGGCTCTCGAAGTCTGCGTGAAAGCCGCCGTTGGTGCCCCTGATGCTCTCGGCGACTGT
AACTTCCCTTCTCTCTCTCTAGCTCTTTTTTTTTATATCAGATTATGATCTCTGATGATCTTCAAATGTA
AAATTTATAATACATGATTTGTCTGTCGTTTCAGG

PE0131 (SEQ ID NO:104)
CCGATAAGAGCCATCATCTCGAGGAGGGGTATTAAGGGATATGTATCCCACATTGGAAAATCAATGGGAC
ATTAAGTAATATATAAAGGGTTAGGGCCAATCCACTAATAGCCAATTGGTTTTGAGTTGGAAGCCCATAA
TAAACCCGAATCTAACAAGATTTTAGATTGATTAAGGAAATTAATATATTATATGCAATATTTCATGGTT
AACGTCAAAATAAGTCCAATTTATAAACAAGCGGATAAACATTTCCCTATATATATGGGAAGGTTTGTGG
TTGCCAAACTCAAAGCACATTGGGTCTTATCTCTCTCTAA[CACACACACACACACACACA]AACTCACG
TATATATTTAGAGCTA[GAGAGAGAGA]TGGGTGAAGAGATGAAAGAAGTGAGAGTAATCGAGGAGTGGT
CTCCGGTTATAGTAATGGTGAT

PE0133 (SEQ ID NO:105)
CCTCCTGTGCCAAGTTTTACAAGGCACCCAACCTGTGAACCTAAATTGCTTTGAAAGAGAAGTTTCTCTA
TCTATAT[CACACACACACACACACACACA]AATCTAATCTTTCTCTTTCAACCGTAAATTTTGCTCACC
ACCAAGGCAAGTTTCCTTCTTTTGCTCCCCTAGCAATATTAATTGCTACTAAAATATCTTGCTAAGGGTA
ACCAAATCTTGCTTCATTCCTCTGTAATATCACAGAAAGAAACTAAAATTTAGGGTTTTTTTTTGGGTTC
TTTCCATGTGATGTGAGCATTTTTGGGTGAGAAAGATGAAGACTATAATTAAGTTAGGGATTGGGTTGAG
TTTGGTGTTTGGGTTTCTTCTCTTAGCACTTATTGCAGAAGTCTATTACCTTCTGAGATGGAAGAAGCAC
AAGAAGAGAGTCATAAGCCAAGAGAGTGAGGAAGAGAAAGAAGAAGAGCAACAACAACAAACTGGGT

PE0177 (SEQ ID NO:106)
TACCCAAGCTTCCACGATCGGAGGGATCTATCGGGACTTCTCGGAGATTTGTTTCACGGTGATGTCTATT
GATCTTTGGCTCTCTGTTTCGATTGTTGTTGTC[TTCTTCTTCTTCTTTTTCTTCTTCTTC]TCGAGAGG
TTGCGGGGTTTTGAAATCGTCTCCTTCGTCCATGTAGTCGTTTTGGTGTTTGTCCGCCAT[GAGAGAGAG
AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA]GCGAAGACGTTACGAAAAACTTCGATAGAGAGTATA
AGAGAGAGATGCTGAAACTGCTTAAACCCTAATTTTGATCGAGTGTGTTTTGGGAAATTTGCAGAGTAAG
TCCTTATATTTGAGCCGAATTAATTAAAGTAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGG
GTACCGAGCTCGAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAAC
TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTTATAGCGAANAGGCCCGCACCGATCGCCC

TTCCCAACAGTTGCGCNCCTGAATGGCGAATGGCCCTGATGCNGTATTTCTCCTTACNCNTCTGTGCGGT
ATTCCCNCCGCATATGGTGCCTCTCNTTACATCTGCTCTGAAGCCGCNTNTTTANCCAGCCCGACACCCG
CCAACACCCGCTGACCCCCCTGAAGGGCTTGTCTGCCCCCGGGNTCCCTTNCAAACAACTGTTNACCGNC
CCCGGGAACCGCNTNTTTCAAANGTTTCCCCCGCCTCCCCGAAACCCCCAAAAAAAGGGGCNCTNANACC
CCNNTTTTNNGGGTTANGTCNGAAAANAANGGTTCCTAAANTTCGGGGGGCCTTN

PE0187 (SEQ ID NO:107)
ATCTAGGACCCCATGATCCGAATAAGGATAATAAAAAAATGGATCTGCCGAATAAGTATCTGGATAACTT
AAAATTCCCTAGATACCCCCCCCCCGC[CACACACACACACACACACACA]TCAATTTTCCTTGTAAT
TTTTCTCTTGTTCTCTATTTTTCTAAACTCCAATAAAGCAAGTCTTTAACATATACTCCACCTTTATTTG
AGTAAATAATCATGGATTTAATCTCTAAAGTGAAGGACACTTTGTTTATGTTTTCTGTTTTTCTAAATTG
TAAAATCTATTTTCTACCTTTTTAATGTGCTAATTTTAGGAAAAATTATATCAATATTTTGTGTCGTAAT
TAAATCTGTCAACATGAAGTAAATCTGTGTCAAAAAGAAAAAAAAATCTATAGAAACATAATTAAAGTAA
ATGTATGAACATATAAAATAAATCTATGAATGATGTATAAATCTATCAAAATTAAATAAATATGTGG

PE0203 (SEQ ID NO:108)
CTTTAATGCTTAAAGCTCGTTCCAGCTCAGGGGTTGGGAGTTTAAGCCTAAAACTCGAGATCTGCTCATC
CGCAGTTGCTGATTTGGCAGCTTATTCACCTTCGCATACCTACAAATCAACGAATACAACGCAAACGTTC
CTCCTG[CACACACACACACACACACACA]TATACTTAGAACCACATAACACCACTTTTTACAAAAAAAA
ACAAGGATTAGGGATGTAATATTATTACCTTCGCCATTGTCGTTGGCTTTAAGGACAACAAAGACATACT
TTGCTAAAGGAATGACAGCAATGGTGTAGATAACGAGAGACAAGGCACCAAGAACATCAACTTCTGATCT
GATAGGAACTTTACTGAAGACATCACTAAACACATACAAAGGGCTTGTTCCCATGTCTCCATACACAACA
CCTAACGTCTGAAACGCTATCCCAATCGT

PE0250 (SEQ ID NO:109)
ACTCAATAACATCAGGCTTTTCTGGATGCAATCTTGCATTGGTTTGGTCCAGTTTGCATCGGGACAAATA
TTGTTGCTGGTGTATATGGGATCGGAATCTTTGCTCCGACATTCTCTTCGGTGTCTT[CATCATCATCAT
CATCATCATCATCATCATCATCAT]CAGTATCAGTATCATCTCTTATTCCAAACTTCTTATATTGTA
GATGTTCTTGACCGGAGTGGCAAATGAAGACATCTCATATATGGCATTGCTTTTAGGGTTTTTAAGAAGA
TGTGGTGTTGCTATGTTGGTCTTTGATGCAACCACGGTGGTCGGAATGGGCGAGAATGGTCTGCAAACTA
ATGGAGATTGTAGGTGTTCACATGGGTTCATCGACTATTACTTTGGGAGAAGTCCAACAACATCAAGCCC
TCTTCTTCCTGCAGGAAAATAAAATATTTCGTTCATCAATTAAAGTAAGAGAATTAACTCATCACTATGG
TGATATGTTAGTTTCTGTTTATTGTAAGACTTAAAATTACCAG

PE0281 (SEQ ID NO:110)
ACAAAACCGCTTCCACCTGCGTTTCGTCGCAGNCAAGCATTTGCGTTTCATTGGGTTTTCTCGGTGAACC
AAGCTCGCGTTAGTATCCAAACATGTTTTCACAGAATCTCGTAGTAACGACATTTCTTGTTGAAGTTGTT
GGATCTGTGTTCTCATTTCGCTTATCAGCTCCATTTCCTGAAACGTTTTTAATTAAGCGGTTCAATGATT
CTCTTCTATGGGGATAACAATAAAATCTAAAAACCTTACAGGTGACGGAGGGTTGTGAACAGACAAGACA
GGAGTTGAAGTTACTTCAGTGTCTTGACAACTCCATGATCCTGCAGGAGACGATGCAAAGATCGGCGAAG
AAGACGAGTCATCTCCATCGTCTTGCTCTTCACCTTCTTCAGTAAATGGCT[CTTCTTCTTCTTCTTCTT
CTTCTTCTTCTTCTTCTTCTTCTT]CAGTTCCTCCACATTTTCATTCTATGGTCTTCCTCTTCTTCTTCA
TGTTGCAATTCCCATTTTTCAGAATGTTTGTTCGAATGTGTCTGCACACGAGACATCATGAGCCTATCGA
TCTGATCTCTTAAACCGGTCTGGGAGAANGTCTGTGACCGGTCCTCTGTAATATCCAAAAACCACACATT
TTTCTTAGTNAATGGGTACCCAATTAGGAAAATGGGATTTAAAAGATTGGATCAG

PE0283 (SEQ ID NO:111)

AATGAACGAGACAGGAGGTGCCACCCTATTGTCTTTGATGGACCAGAGTCTAGCTATCCGGTAAACAGTA
TTTTATAACAAAGACATGATCATGAAATGATTTTTTTCTTCTGAAGTTTAACTGATGACTCATATATCTA
TATCTGACTAGTTCATCGTGGACATGGAGCATTCACAGGACCAGAGCACACATCGAAGCTGACATACAAA
TGAATCTCTTATTGCCTCTAGTTTCACACGTAAATATTCAGTTTCTAG[GATGATGATGATGAATTGATG
ATGATGATGACGATGATGAT]CAGTCAAAAGTACTGTAAATTGATGGTTATGGTTGTCTTGGCTTTGCTT
AATCATGT

PE0286 (SEQ ID NO:112)
ACAAGTTACCATTTGAGGATATATCAGAGAAGAAGGAGTTGCTTGAAGATGACGAGAAAACCAAAAAGAA
GATGAGTTCTAATGGTCGTTGGTACGAGGAGCTTGATGTCTTCATAGAGAAACCTGAAACTGGTGTTCTT
ACTGGTGATGGTGCTGTGGTGGACGCATGACTGGGAACGAACCTGTTGATGGTGACGAGTTGGATGTTGA
GCAACAAGATGATAATTCTGATGGTGATCATGGTGATCATGAAGCAG[GAGAGAGTGAAGA]TGAGTATC
AAGCGAGTGATGAATCTGATAAAGAAGAGGATATTGACAGAAATTTTGAAGAGGATGTTGAGATGTTCCA
GGGATGAGAACTACGATGGAGGAGATTCCAGACGAGGAGGAGGTATATTCTGACACGGAGGAGTCATCTG
ATGATGAAGAGGAACAAGCTGAGAAGGATGCTAATAGGGGTGAATTAGATGGCATTTTTAAGTCTTAGGC
AGGAANTTGCAATGCCTGCAAGTCGACCTCTAGAGGATNCCCGGGTACCGAGCTCGAATTTCCACTGGGC
CGTCCGTTTTACAACGTCCGNGACTGGGAAAAACCCTGGGGTTAACCCAACTTAATCGCCTTTCAGCACA
TNCCCCNTTCGCCANGNTNGGGGTAATAGCCNANAAGGCCCGCAACCGATNGGNCCTTTCCCAANAGTNG
CCGCACCTNAAATGGNGNATTGGCGCCTTANGNGGGAANTTTNNCCTTANGNATT

PE0324 (SEQ ID NO:113)
ACATAAGCCCTTTTTTATTATCTCTGCATATCATTACATTCATTTTATGTCACATATGTTTATTGCTCTTC
TCTTCAGATTACTATTACATCGCAAGTAAAACAAAAGAGTTAGAAAATAAAGTAAACACTCCATACATAG
TCAAAGTATCTCCATTACTCCTCTTCTTCGTGTTAACAAGTCTTTAGGCGTTTCTAAACCGCAGAAACCA
TCATAGCCGGTGATGCACCAACCATCAAGT[CTTCTTCTTCTTCATCATCATCATCATCATCAT]CCTCT
GCTTCCCACATGAAATGAGCGTATGATCCCAAAACCATACTACAAAAGTCACAAACCTTTAACATTCTGA
AAAAAAAACTCATCAAAGAATCCAAACTCTACATATAACATAACATACCAATCATCAGGAGAAGCGTTAA
CAGCTTGATCAAAGTAACACTGAGCTCTCTTCTCATCTCTCTTCGTCTCCCAAATCAGCTTCCCATACAT
CGACAACGCTTCACCATCACCTGGATCCGCAAGTATAGCTCTCCCGTAATACTCCTCCG

PE0340 (SEQ ID NO:114)
CTTAGTGACCCAAAAGCCATTGGTGTATGATAGAAAAGTTAGTTAAATACCGTTACTCGCAAGGAAGACC
ACACATTTTTTAATTCTATCTCACTTAGTCAGACCAGCTCGGATCCTTCTCTAGAAC[CACACACACACA
CACACA]CTCAGAGTGAGAGATTCATCAATGGCGGTTTCTTGCAGCCACTCATCGATTCTCTTGCCCCCA
ACCACCTCCTCCGTTGGCTTCAACCGCTTCCCTTGTCTCCAAACGCTGCGTTTCAAATCCAGAAACGTTT
ATCAGAAAGCGAGGATCTCTACAGTGTCGGCGTCATCTTCACGGTCTCTCGAAGCTCTGATCTTCGACTG
CNACGGTGTGATACTCGAATCGGAGAATCTACACCGTC

PE0355 (SEQ ID NO:115)
TCAAAGAGCACTTACAAGGATCCAGACGATGGAAGGCAACGATTCTTACTCGAACTTGAGTTCATTCAGT
GTCTCGCGAATCCTACTTACATACACTGTAAGCTCTTATGATTCCTTATCACATAGTATCTACTTATAGC
ATTTAGGAAGTGATAAGAGATCTT[GTGTGTGTGTGTGTGTGTGTGT]TTTATGCTCTATGATGAACTTA
CCACTTAGCTTTTNGATTCTGTTTTGGCAGACCTAGCACAGAATCGTTATTTTGAAGATGAAGCATTTAT
TGAATACTTGAAGTATCTTCAGTATTGGCAGCGACCAGAGT

UB0015 (SEQ ID NO:116)
ACTTCAGTGGTCGAAAATCAAAATATTCTTCCATCATTTTAGTTTTTTTTTTTCTCTATGTTCGCATCAA
GAAAACGAAATGAAAGGGATTATAAAAGGAAGAAGAACTTGTGAATCACGGTAAGTTTCGGGGTTTGTTG

TGAGGAGATTTCGAGAGAATCAAGAATAAAATTATATCACGAGATTTTTTTGTTTGAAGTGAGAAAGAAA
TCAAAGATTTTATTTTTTCTCTTTTGGTGAGTGATA[GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA]CGTGTTTTGGAACTACGG
TGATTTTTACTATTTTGATGATGTTTTCAACTTTGAAGAAGACCTTCTCTCATGCTCACTCTTAGCATCC
TCCTCATTTATAGGATTAGATGGGAGAGAGAGAGCGTTTTAGCCATTAATACTTTAATAACAAAATGAAA
AATCTGATATTAACATTTCTTTTTTCACTTCTCCATCAGTGGCATTTTCGATATTTT

UB0126 (SEQ ID NO:117)
ACCTCCTGATGACTGCTTAAACAGCGCCTGCGAAGATCTGGATTCTGTAGTTAACCAGGCTAGGGAGTTC
TTAGAGGACTGGTCCCCAAAGTTGAGCAAGCTCTTTGGTGTAAGTTGATGAACAAGCTCTCATTTTCAGT
TTTCTTT[CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
CTCTCT]TTGCATCATTCATTGAGTTGTGTGTGTGCAGGTGTTTCACTCCGAGCTTTTGTTGGAGAAGGT
CCAGACTTGTTCACTGGAGATTAATCGCATACTTCTTCAGTTATCACAGTCAAGTCCTGTAACTTCAAGT
GT

UB0163 (SEQ ID NO:118)
ACAGAAACAGTAACATCAACACACACAACAAACAGCTCGCGAAATGAATTACAGATTCCTCTCCGAAATC
AAAACAGGAAACGGACACA[GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGAGA]AGGTGA
TACCGTCGAGAGGTTTGATGTTGCCGTCGCGGCGATCGACGGAGAAGCCTTCATGAGGCGAATCGACGGG
TTTGACGACGT

UB0181 (SEQ ID NO:119)
ACAGACGCGATGATGACACTTCTGGTCCTAAGCGTTGTGTTACGTTTGATACTCCTGCGCTTGTTTACTT
GGAGTATTCTGATGTGGTTGCTGATAAGTATGAGAATCTGAGTTTGGACAGCTTGGTTGAAGTTAGGCTT
GATCTTCAGTTGACTGCAGATCAAATCATGCGCAAGAATGCTACAGACAGTGTTGGTTTTGTTCCCGGTG
ATGTTTCAACTTTGTTCATGGGGGTCAAGAACGTCAAGATCCTCTGCTTATCTCCTGATTCTTTAGATGT
GAGTCCAGTCCTTTTTAAGTTAGCTTCATCACTGTGTAGCATTTGTTTTTTTTTTAAATTTGATTGGTTA
GTGATGATACAAAATATTTGATTCTG[GTGTGTGTGTGTGTGT]TTGTGAAAGTTCAGTCCCTTTAACTT
AGCTTCATGAGTGTGTAGCCTTTGTTTTTTAATTGGTTACTGATGATATGGTGTGTGTGTGTGTGTGTGT
GTTTCAGACGCTCTACTACCGTGGTGGTGACATGCCGGTGTTCAACAATCTGATTT

UB0196 (SEQ ID NO:120)
ACATGAGAACAAGATGGGTTCGAATTACCTCTAGCCTAGATCTGGATCTGGAAACAAGAGACAAGGGAGA
GACGAGATCTTGTCACCACCACCAATCGGCTGCCACCACCACCACCACAACACGGCGCCAGCGAAAGGGA
GATA[GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
AGAGAGA]AGAAGAAGAAGAGAGAAAAGGAAAAGAGAAGCTTGATGGCTAGGGTTTCTTAGTCTCTCTAA
ATCTCTGCAGGGCTTTGCTCAAGTTTCAGAATGAGAGAAAAAAGAGAGGAGGCAACTTTATTTATAGGAA
ATGGAGGGAACCCTAGGTCATTTACCTTAATGGGCTGCAGTCCTAACGAGCTCTCGTTAAAAAAATTTGG
GCCGGGTATCGGGATGTTACACTAACGGTGTGTGGCGATGAAGGCTCTTCGACTCTCAAAATTAATGATG
TCCATAACTAAATAAAAACTACTCGACTTTATTAAGATATAGCTTCAATGATTTAAAATTAAATATAGAA
CTCT

UB0307 (SEQ ID NO:121)
ACCTNTGGGTAAGTAACTGTGGTGGC[CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT]CAATACAC
TCTTCACTTAATAAATGTGAANACGTTAACTNGTTTCTTTTNTCACTTCTCAGTTATGTAGCTCCAGAGT
ATGCGAACTCTGATCTTCTGAATGAGAAAAGTGATGTCTATAGCTTTGGTGTTGT

UB0315 (SEQ ID NO:122)

CACGAAAGCAGGCCCCACCCAATAAGCGATGAGCTGTATATTTATTTTGTCTTGTTTTCACAAAAAATAA
CCCTTCATGTTTACAGTTAATTACACAACAGCCCCTTTCTTTCCTCCATGACCAACGACAAGGTCGAATT
T[CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT]CCGTCGTCTTCATTCAATCTATCTCAG
TGATTTACTCGCAATAGAAGTCGCCTCTTAATCTCTCGAGAGAGAAGCTCAAGT

UB0331 (SEQ ID NO:123)
CGTGGACTAACGCTCGTGTGCAGGAAACGATGTTCGTGAAAAGGTATCCGATCAGAGGAGCCTCCGCCGG
TAAAAACCCTTCG[CCGCCGCCGCCTCCG]TTGAATGGTAATAACTCTTGTTCTCGCTTTCTCTTCAAAC
TCCCTTTTTTTTCTCTGATTATTTTTGTTGGTTA

KK66 (SEQ ID NO:124)
GAATGGGAAGCATCACAATGATAATGCTAATGGCGGTTTTGGTCTGGTCCATTACTCTAGAGACCTGCAT
TGCTAGAAGAGGAAGACATTGGAGACATAACCACCGAAGCTCCTC[A/T]GACTTGTCTGATTCCTTGTC
AAGCAAGAAACCAAAAAGCCACAGTCACCACCACAGCTCTCA[C/T]AACAACAACCATAATCATCACCA
CAAGTCTAAACCTAAACCAAA[A/G]CCAAAGCTGAAAACGCCGCCAAAAAGTGACCACA[A/C]TAAAT
CTCCGGTGGTTTCACCGCCACCAAAAGTCCAACCACCGTCTCTTCCGCCGCCAAAGGGATCCAAAGTTTT
CAATGTGATGGATTTTGGCGCAAAGGGTGATGGCAAATGTGATGACACTAAGTCGTTTGAAGCGGCTTGG
GCAGCAGCTTGCAAAGTGGAGGCATCCATGATGATCATACCGCCTGAATACACTTTCCTTGTGGGTCCAA
TCTCATTCTCTGGTCCTTATTGTCAAGCTAACATTGTGTTTCAGCTTGATGGTACTATTATAGCTCCAAC
GGATTCAAAATCATGGGGAAAAGGGTTAATGTGGTGGCTTGAATTCACAAAGCTGAAAGGAATTAAAGTA
CAAGGTAAAGGTGTGATTGATGGAAGAGGCTCTGGT

KK98G (SEQ ID NO:125)
GACAGAGATAGCCCTAACTTAGTCACTCTCTCTCACACACACTCCAGTTCAAAGTTCAAA[A/C]AATGG
CTCCTCCACAGAAGCTCTTTCTCGCCGCCATTGTCGCTGCCGTCATTGTAGCCGCCACCACCGGATATGC
ACCTAATAGTGCTGCGGAAGATATTGTGCATTCCTCATGCGTGCACGCGAGCTATCCATCGCTATGCGTC
CGTACACTCTCTACCTACTC[C/T]GGTCCAACCATCACAAACCGTCGCGAGCTAGCTCAAGCCGCCGTC
AAGATAAGCCTCTCCCACGCTCGAGCAGC[C/T]GCTAAGAAACTCGCGGCTGTGAGAGAAACCGTGGG[
A/G]AAGAAACGGGTGAAAGCGGCGGTTGTGGACTGCGTGGAGATGATTGGAGACTCGGTGGACGAGCTG
[A/C]GCCGCACGCTAGGCGTTTTAAAGCATCT[A/C]CACGTTTCGGGCGTTTCGGCGAACGAGTTCA[
A/G]GTGGCAGATGAGCAACGCGCAGACGTGGGCTAGTGCGGCGTTGACGGATGACGACACGTGTCTCGA
TGGGTTTAAAGGGGTCGAGGGTAAGGTTAAAACGGAGGTGAAGCA[G/T]TGGATGACGAAAGTGGCGAG
GGTTAC[A/G]AGCAACGCGCTTTACATGATCAACCAGCTAGATGAATCACGTGGCTAGCCCCACGTAGT
ACGTTCTTGATGTTATGATGTGCTTGTCCTAATGGACAGTTATGATTTGGTGTTAGTTTTTTTCGTGTTT
GCTTAATTGCGAGTTATCTACTATTTAAAAATGAGAGGCATTGTCCTTTTAAGTAGTTCTGATAATGGTA
TACTAAATAAATGGTTTATCTCTTTTTTCGGACGGTATGTCATTGTATCGTATTGTGTTGTTCCCTTCGG
ATTCGATAGCATGTGATTTTGTCTTGACGTGTAGTAGCGCCTTGGCTGAGCTAATGCTCTAAATAAAAGT
TTTAAGTGGC

Table 14 below sets forth additional information about the markers of QTLs associated with whole field plant resistance to *Sclerotinia*, as well as exemplary sets of forward and reverse primer sequences for each polymorphic region.

Table 14. List of SSR and SNP markers and primer sequences used for amplification of loci associated with *Sclerotinia* whole plant field resistance

Marker	Repeat	Forward Primer Sequence	Seq ID NO	Reverse Primer Sequence	Seq ID NO
AG0023	(CTT)	ACAAGGAGATGATCGCGGTTTC	126	CCAATCTGTGTAAACCAAACGGG	127
AG0045	(GAA)-(GAA)	ATAAGGCTTGAGGGACATGCCA	128	TGGCTCCACAGAAACAGCTTTG	129
AG0047	(CAT)	GAAGCTGCAATACTGAGGCACC	130	GCAATTCTTACCTGTTGTCCCAAA	131
AG0070	(CCA)-(GA)-(GA)	AACTGGTCGAGCGGGATTTTTT	132	TAGGAAACCCTAGCCGTCAAGC	133
AG0086	(GA)	CAATGTCGGTAAGCACCGGAAG	134	TGCCGGAAAATGCTGACTTGTA	135
AG0093	(CAG)-(CTTTT)	GCTTCAGCCAAGGGATTTGAGA	136	AGCTCTTTTGGTGCGATTCGAT	137
AG0125	(CAT)	CCACATGCCTTAGGTGATTGGA	138	TTCTTCCGGCTTCTCAAAGGTG	139
AG0148	(CCA/T)	CATCCTTGTCCAACGTCCCTTC	140	TTCCTCTCTTCGAGATCGGTCG	141
AG0171	(CT)	GGACTCGAACATCTCCAATTTAACT	142	TGAAAATAGAATACAATTAGGGCTT	143
AG0203	(CTT)-(CTT)	GCGTTGCCCCTCTCCTCTACTT	144	CGCAATCTACAAAAGATACATCAAAAAG	145
AG0239	(CTT)-(CTT)	AAGAAAAGAGAAACGATCCCACG	146	TGAGAGTGAAGAGGAGTTGGGTC	147
AG0243	(GAA)-(GAA)-(GAA)	GATTGGGGGATGAGATTGTTGG	148	GCCGTCCAAAAGTCAAAGGTCA	149
AG0272	(CTT)	AGATGATCGCGGTTTCCTCAAG	150	GAGGCAAAGCTATAAGAACAACTCCA	151
AG0304	(CAT)	CATGTTTGGTTGCTACGGTGGA	152	GAGGTTGAGACGGAGAAGCACC	153
AG0323	(CAT)	GACCAATACAAAAACCGGGCAA	154	TTGATGGAGAGTGGGTTGTGCT	155
AG0324	(CAT)	GTCTCACCAACTCCAAACTTGTTAA	156	CAGGTTCCTTACCAAAGATAAAGAG	157
AG0328	(CTT)	AGCTCGTCTCCTTGCTGTCTCA	158	GGAAGTGAAGAAGAAGCCGGTG	159
AG0359	(GGT)	TGCTCAAAACCCTAGTCGTCACC	160	TGAGAGCGAAAACCAAGAGAGGA	161
AG0369	(GGT/C)	GAGTGGGCTGTACCTTGTAGTTGA	162	TTTGTAAACATCAGAATCACCACC	163
AG0370	(GTTTT)	TCCATTATTACAACCACCCGCC	164	GTTCCGTTGCCCTCTCCTTTTT	165
AG0378	(CCG)	GTTGCTGGTGGAGTTGCTGCT	166	TTTGTAGATGCGACTGCTTCATCTT	167
AG0391	(GGA)	CGACGCTCAAGAGGAAATGCTT	168	CAGTGTCCACCGGAGTAGCAGA	169
AG0410	(GGA)-(GGA)	TGATCCATATCGGGGAAAATCG	170	TTTTTGCTTGGTTTCCGACAGA	171
AG0441	(GA)-(GGA)	CCGGAAACCTCCGATTGAGTAA	172	GGATTAGTTTAACATAGATGGGCCG	173
AG0477	(AAT)-(CT)	GGGAGATAATGTTGGGAATCTTAATCG	174	ACTGAGCCATCCTTCCTCCTCC	175
AG0482	(GGT)-(CGG)	TCCGTAGAAAGAACAGGCTCGG	176	GAACCGCCGCCAAAACTAAAAT	177
AG0504	(CAT)-(GGA)	ATCAAATCCACCCACTGCACCT	178	GCCTTCTCCTCACATCTACGCA	179

Marker	Repeat	Forward Primer Sequence	Seq ID NO	Reverse Primer Sequence	Seq ID NO
AG0510	(GGA)-(GAA)	TCATCTGATTCATCGTCATCATCA	180	TGACTCTTGTCAACACCACCACG	181
BG0031	(AGG)-(AG)	GAGGAAGCATAGGAGGAGGAG	182	ACATAACCCAAATCCCCAAAT	183
BG0106	(AGGG)-(GAT)	CGAAATTATGTGTGTGCGCTCC	184	CCCGGTTAGGAAATTACGGATCA	185
BG0111	(GAT)-(ACC)	TCATTTTGACTTTTGGCGTTTGG	186	TGCATAAGTACGTTGAAAAGGGCTC	187
BG0119	(CGG)-(GAT)-(GTT)-(ATTA)	ATGAGGAGAAAGGATTCGCGGT	188	CGGTAAACGTCCAAACCTCACC	189
BG0181	(ACC)-(ACC)-(ACC)	GCAGAGCAACGAAGTACGCCTT	190	TCGTGATGGTGTCTCCAATGGT	191
BG0228	(AGC)-(CAA/CAG)	GAGACGAAGCCATTGGTAGGGA	192	CGGCTTGTTGTTGTTGCTGTTT	193
BG0255	(GAA)-(GAA)-(GAT)	CCAAACTCAGCACAGCCTTTCA	194	ACGTTTGCCACATTCACAGCAT	195
BG0278	(CTT)-(ATT)	TGGCTGTTCAGTTGTTTAGCTGGA	196	CCCATTTGAAGGCAGAGAGTGC	197
BG0295	(AGGG)-(TCC)	CACATCTCTCCGATTTCATCGC	198	AAGAGGATTTTGTCGGTGGGTG	199
BG0452	(TGG)-(GAT)	TCCCTTTGGTGATGTTGGACTG	200	CACATAATCTGCATTGTCGTCTTCG	201
BG0516	(GAT)-(TAA)-(TTGGT)	TGGTCAACAGAAAATGGCCTGA	202	TGGCATGTCCTTTCATGCTCTC	203
BG0647	(CTT)	ATGCAAAGATGGGCGAAGAAGA	204	CGAGAGCGGGTTACGAGATCAG	205
BG0651	(CTT)	TGCATAACAAAAGATTTGAACCCG	206	GGAGCAAAAGAGCGGAACAGAA	207
BG0713	(CT)	ACCGCCAAAGAAGACGAAAATG	208	CTCGGCGGACAGATACACTCAG	209
BG0864	(GA)	GAAGCTAAATGCGTTGCGTTGC	210	TTTGGCTGTAAAATGAAGTGAGCA	211
BG0869	(CT)-(CT)	GAGTCGGCCACAAATCAAGGAA	212	TCGGAGGAGGAGGAGATTGAGA	213
BG0988	(CT)	CCCAAGACTCCAACCGGAAAAT	214	TCGCATTAAACGAGGACGTGAG	215
BG1062	(CT)	CCAACTCGTTCCATCCCAGTCT	216	CGGTGGCTCGTACTGCTTCTCT	217
BG1090	(GTA)-(GA)	CCGTTGAGGTAGGTTTCTGCCTT	218	TTGCCTCGCTCACATCTTTTTG	219
BG1101	(CAT)	ACATGCTTGTGGATAAATCATCAT	220	GAAAATGAGAAAACCCAAACTAAA	221
BG1123	(GA)	AGGCCACCTTTTGTCACCAGTC	222	GGGGGGTTAATTTGTCCCATTT	223
BG1127	(GA)	CCGCCATAACAAAAATCTTCCC	224	TGTTTCCTACAAAGGTATAACCGGC	225
BG1149	(GAT)	TTGCTACACTTCCCTGTGGGTG	226	CCTCTCTCACTGCGTGCATTTC	227
BG1182	(AG)	CGGCGCACTGATGATGTTTCTA	228	GGGAGAGAGTTTGTTCGGTGCT	229
BG1197	(CATTT)	TTGTGAGCGCCAAGATAAGGCT	230	AAACCCTATCCCGCGAACAACT	231
BG1230	(GA/TGG)	GGCCTCTTGTGTTTCCTCCAAC	232	CACCACTACCGCCATCTCTCCT	233
BG1241	(CAG/AAC/CAG)	TTAACAAACCGAATCCGCAAGC	234	GATTGTTGTTGCTGCTGCTGCT	235
BG1244	(CTT)-(CTT)-(CTT)-(AGT)	CAACGAACTCTTCTTCTCTGCTTTACA	236	TCACGACAAATGGTCAAATTCTCA	237
BG1286	(AGC)	GTCCGAGACAGAGTATGCTAAGC	238	TCATTGGTGGATCACTTCAAATA	239
BG1288	(AGC)	GATTCTATCAGCCACGGAACGC	240	CACCTATTTACCCCACGAGGCA	241
BG1321	(AAG)	TGCATAGACTCGAACCAAACCG	242	TCTGATACGCCAAGCTCTGCTG	243
BG1368	(CGT)	CGACGGTTTCACGGCACT	244	TCATCCTCCGACGACGAC	245
BG1392	(GA)	GCTCGGCTTCCAAGGGTAAGAT	246	GATGTACTGCTTAGCTGCCGCC	247

Marker	Repeat	Forward Primer Sequence	Seq ID NO	Reverse Primer Sequence	Seq ID NO
BG1442	(GA)	CGAGGAGGAAGAAGATGACCGA	248	ACGAGCAGGCGGTGAAAATAAA	249
BG1449	(GA)	AGTTGTGGTGAACAGGCTGCAT	250	CCAACTCGTTCCATCCCAGTCT	251
BG1453	(GA)-(GA)-(GA)	TCACCTTCCTTCCTTCAATGGC	252	TGACCCACCTCCTCTGCTTTTC	253
BG1513	(GA)	TGAAGGTTGCGATAGCGAAGAG	254	TTGGGGTCGGAACTGAAACATT	255
CA0105	(GAA)	TGTTTCCTTCACCATGATCGGA	256	CACACCCTACCTCTCTTGTGTCCC	257
CA0120	(CCA)-(TCA)	GACTAAACCAGACCAAGAGAAAAGTCG	258	TGTTGAACCGGATAGGCAAGGT	259
CA0163	(CAT)-(CAT)	CCATCATCACCACCACCATCAT	260	CTTGCTGGAGAAGGTCGGATGT	261
CA0221	(CTT)	GGAGAAGGGTCGTCGTCAAGAA	262	TCGAACACACAAACGATGCTCA	263
CA0226	(GA)-(ACC)	CATGGATCACCTGCACCCTTAG	264	GCGACTCCGGTGAAGACGTATC	265
CA0233	(GAA)-(CAC)	GGCAAGCATGGTCTCGTCAGAT	266	AAAAAAAGATTCCAGCCGCCTC	267
CA0328	(CGT)-(CAT)-(CCA)	ACTTTGGAGCATCCTTCCTTGG	268	GATGTTGGACTGCGCTCTGGTA	269
CA0410	(CAG)	CTGTTGAGCCATCGAGATCAGC	270	CCAGAAAGTGATGGTGTGCATAA	271
CA0423	(AGC)	AATCAGTCCGATCACTCCCTGC	272	GGAAGTTACCTCGTCGTCGGAA	273
CA0456	(TTG)-(TCAA)-(CCA)-(CTC)	TTCTGTTAGAATTCTACCGTTGTTG	274	AGCTTTGTGAGGAGAGTGTGGT	275
CA0488	(CTT)-(CTT)-(CAA)-(CAA)	CAGCAATGTCGTCGTTCAATCC	276	CTGTAACTCGCCGGAGCTTGAT	277
CA0546	(GAA)	CCCCTTCCTTATCCACACACACA	278	TGGGTTCGTGAAGGTGAAGGTT	279
CA0552	(CTT)-(CTT)	CGCTAGGGGGTGAACCAAGAAT	280	GCGTCGATCCTCCTCTCCAATA	281
CA0603	(CTT)	CCTCCCAAAGTCGTCTCTTCCA	282	GTAGACAGGGGACGAAACTCGG	283
CA0624	(TA)	TTCCAAGTGGTTCTGCAATGTG	284	TCAGCTATCCCAATAAAGGGCAA	285
CA0636	(CCT)-(CCA)	CCAGGTCCTTGCGTTGAATCAT	286	ACCACACTGCATATCCCTGCAA	287
CA0682	(AAGG)	CCTCTTCAACCCCACACTGGAT	288	TTGCGAATGTATCAGCCGTCTC	289
CA0729	(GGTT)	CCGAACCGGAATCATACAGCTC	290	TGCGTTTGAACCAGTCAGATCC	291
CA0736	(ATAG)	CCGGGCTTAGACAAACTTTATGAG	292	ATTGATCGCACAAACGCCTGTA	293
CA0739	(AT)-(GTAT)	AAAAACTGATAAGATTATTGTTGGTAAC	294	CCAAAAATAAAAGTTAAACTTGCATA	295
CA0753	(TAA)-(ATAC)-(TC) -(CGG)	CAGGCAGCTAAGGAATCTGGAAA	296	TAAAAGAGGCGTCCCGATGAGA	297
CA0834	(AAT)-(AAT)	AAAATGCAACCATGCAATACGTG	298	GTTTGTAGTATTTGTTGTCACTGCTGC	299
CA0837	(ATT)	TGCGAAAGCCATGAACCTTTCT	300	TGTTCTTCACAAAAAAAATCAGCAA	301
CA0896	(AAT) -(AAT)	CTCATGGGAGGTTCGCTTGATA	302	GCAACTGCAAGATCAAATGGTCA	303
CA0991	(AAT)	CACATTGGTTTACAGAGTCTACATGA	304	ACAAATACATGTGAAAGTTGAAAACA	305
CA1027	(ATT)	TCAGTCTAGGCGTTTACCAATACCA	306	CTTAGGCGGTGCTTTGGCTCTA	307
CA1032	(AAT)-(AAT)	TTGAACTTGACGACGACATCCC	308	GATTGCATGATTTAGCTGCTGGA	309
CA1034	(AAT)-(AAT)	AAAATGCAACCATGCAATACGTG	310	TTTGTAGTATTTGTTTGTCACTGCTGC	311
CA1035	(ATT)	CTCCCAGGTGAGACCGACAATC	312	CGTAGCGCGGAAAAGGAATCTA	313
CA1066	(ATT)	CTCTTGCCTTCATCGGCGTACT	314	ATCCAGGGCTCAAAAACAGCCT	315
CA1080	(ATT)	TGACAAAATCCACATCTCTAATGGTG	316	AGGCTACATCGGTTTTGGGCTT	317

Marker	Repeat	Forward Primer Sequence	Seq ID NO	Reverse Primer Sequence	Seq ID NO
CA1090	(TAA)	CCATGCAGCTTTTTGTTACGACA	318	GGCCAACAATGCAAATACACGA	319
CA1097	(ATT)	AAAGTGGTGGAGCTTTTTCCAGTT	320	GCCAAGACCAGTGGGATGTGTT	321
CA1107	(AAT)	TAGTGTCAAAAAACATTTGTCTTTCA	322	TTGTCATTTTTTTGTCATCATATTTT	323
PE0012	(CTT)	CACCGTCGGAGTCTGAAT	324	GAGCCGTTAAACCNTAGTGTG	325
PE0017	(AAG)	ATTGGGTTCTGACCTTTTCTC	326	CTTTTCCTCATCGCTACCAC	327
PE0063	(CTT)	TCGTTCTGACCTGTCGTTAT	328	GGAAATGGCTGCTCATGTT	329
PE0091	(GT)	CGGCAATAATGGACCACTGG	330	CGGCTTTCACGCAGACTTCG	331
PE0131	(CA)-(GA)	TATGGGAAGGTTTGTGGTTGC	332	CACTCCTCGATTACTCTCACT	333
PE0133	(CA)	TCCTGTGCCAAGTTTTACAAG	334	GGTTACCCTTAGCAAGATATT	335
PE0177	(TTC)-(GA)	TCTATTGATCTTTGGCTCTCT	336	CGTAACGTCTTCGCTCTC	337
PE0187	(CA)	GACCCCATGATCCGAATA	338	AAGACTTGCTTTATTGGAGTT	339
PE0203	(CA)	TACAACGCAAACGTTCCT	340	TTGATGTTCTTGGTGCCT	341
PE0250	(CAT)	TGGTGTATATGGGATCGG	342	GTTTGCAGACCATTCTCG	343
PE0281	(CTT)	GAGACGATGCAAAGATCG	344	TGCAGACACATTCGAACA	345
PE0283	(GAT)	CATTCACAGGACCAGAGC	346	CAAAGCCAAGACAACCAT	347
PE0286	(GA)-(GA)	ATGGTGACGAGTTGGATG	348	CCTCGTCTGGAATCTCCT	349
PE0324	(CTT)-(CAT)	CAGAAACCATCATAGCCG	350	TGATTTGGGAGACGAAGA	351
PE0340	(CA)	CGTTACTCGCAAGGAAGA	352	TTCGAGAGACCGTGAAGA	353
PE0355	(GT)	ATGGAAGGCAACGATTCT	354	TTCTGTGCTAGGTCTGCC	355
UB0015	(GA)	TCGGGGTTTGTTGTGAGG	356	GAGGAGGATGCTAAGAGTGAGC	357
UB0126	(CT)	ATGACTGCTTAAACAGCGCC	358	CTTCTCCAACAAAAGCTCGG	359
UB0163	(GA)	ACACACAACAAACAGCTCGC	360	AACATCAAACCTCTCGACGG	361
UB0181	(GT)	AAGAACGTCAAGATCCTCTGC	362	ACCACCACGGTAGTAGAGCG	363
UB0196	(GA)	CATGAGAACAAGATGGGTTCG	364	CTGAAACTTGAGCAAAGCCC	365
UB0307	(CT)	TGGGTAAGTAACTGTGGTGGC	366	AGAGTTCGCATACTCTGGAGC	367
UB0315	(CT)	TCCATGACCAACGACAAGGTC	368	AAGAGGCGACTTCTATTGCG	369
UB0331	(CCG)	GTGTGCAGGAAACGATGTTC	370	GGGAGTTTGAAGAGAAAGCG	371
KK66	-	CCACAGTCACCACCACAGCTCTCAT	372	GGCGGTGAAACCACCGGAGATTTAG	373
KK98G	-	GCCTCTCCCACGCTCGAGCAGCC	374	CCCTCGCCACTTTCGTCATCCAA	375

While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be clear to one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention. For example, all the techniques, methods, compositions, apparatus and systems described above may be used in various combinations.

Claims

What is claimed is:

1. A method of identifying a *Brassica* plant or germplasm that exhibits whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, the method comprising detecting in the plant or germplasm at least one allele of at least one quantitative trait locus (QTL) that is associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, wherein the QTL is localized to linkage group N15, wherein said linkage group comprises at least one marker that is associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia* with a statistical significance of $p \leq 0.01$, thereby identifying the *Brassica* plant or germplasm that will exhibit whole plant field resistance or improved whole plant field resistance to *Sclerotinia*;

   wherein the QTL is localized to a chromosomal interval flanked by and including markers PE0286 and AG0369 on linkage group N15.

2. The method of claim 1, wherein the QTL is localized to a chromosomal interval selected from one or more intervals on linkage group N15, flanked by and including markers (i) CA0719 and AG0369, or (ii) PE0091 and PE0187, or (iii) PE0286 and PE0187, or (iv) PE0286 and CA0719.

3. The method of claim 1, wherein the marker comprises a polymorphism that identifies the at least one allele of the at least one quantitative trait locus (QTL) as being associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, and the detecting comprises identifying the polymorphism.

4. The method of claim 3, wherein the polymorphism is a single nucleotide polymorphism (SNP) or a simple sequence repeat (SSR).

5. The method of claim 4, wherein the detecting comprises detecting at least one marker selected from AG0369; CA0719; PE0091; PE0187; and PE0286.

6. The method of claim 5, comprising detecting two or more markers located in two or more different linkage groups.

7. The method of claim 5, wherein the detecting comprises amplifying the marker from genomic DNA of the plant or germplasm and determining if the marker comprises the polymorphism associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia*.

8. The method of claim 1, wherein the plant is *Brassica napus*; *Brassica juncea*; *Brassica rapa*; *Brassica oleracea*; or *Brassica carinata*.

9. The method of claim 8, wherein the plant is *Brassica napus* (canola).

10. The method of claim 9, wherein the plant is spring canola.

11. The method of claim 9, wherein the plant is winter canola.

12. The method of claim 9, wherein the plant is semi-winter canola.

13. The method of claim 1, wherein the whole plant field resistance or improved whole plant field resistance results from decreased disease incidence compared to a plant lacking the allele of the QTL associated with the whole plant field resistance or improved whole plant field resistance.

14. The method of claim 1, wherein the whole plant field resistance or improved whole plant field resistance results from decreased disease severity compared to a plant lacking the allele of the QTL associated with the whole plant field resistance or improved whole plant field resistance.

15. The method of claim 1, wherein the plant has whole plant field resistance or improved whole plant field resistance to *Sclerotinia sclerotiorum*.

16. A method of introgressing *Sclerotinia* resistance in a second plant by (i) cross pollinating the identified plant or a progeny thereof of claim 1 with a second plant, and (ii) screening

for the at least one allele of the at least one QTL, wherein the second plant lacks the at least one allele of the at least one QTL detected in the identified plant.

17. A method of producing an F1 hybrid seed, wherein the F1 hybrid plant derived from the F1 hybrid seed is resistant to *Sclerotinia*, the method comprising (i) cross pollinating the identified plant or progeny thereof of claim 1 with a second plant, and (ii) screening for the at least one allele of the at least one QTL, wherein the second plant lacks the at least one allele of the at least one QTL detected in the identified plant.

18. Use of the identified plant or a progeny thereof of claim 1, for introgressing *Sclerotinia* resistance from the identified plant or progeny thereof into a second plant.

19. Use of the identified plant or a progeny thereof of claim 1, for producing an F1 hybrid seed, wherein the F1 hybrid plant derived from the F1 hybrid seed is resistant to *Sclerotinia*.

20. A method of positional cloning of a nucleic acid comprising a quantitative trait locus (QTL) associated with *Sclerotinia* whole plant field resistance or improved whole plant field resistance, the method comprising:

    (a) providing a nucleic acid from a plant comprising a marker that is associated with *Sclerotinia* whole plant field resistance or improved whole plant field resistance with a statistical significance of $p \leq 0.01$, wherein the QTL is localized to linkage group N15, and wherein the linkage group comprises the marker, and wherein the QTL is localized to a chromosomal interval flanked by and including markers PE0286 and AG0369 on linkage group N15; and

    (b) cloning the nucleic acid comprising a quantitative trait locus (QTL) associated with *Sclerotinia* whole plant field resistance or improved whole plant field resistance.

21. A method of making a transgenic dicot comprising a quantitative trait locus (QTL) associated with *Sclerotinia* whole plant field resistance or improved whole plant field resistance, the method comprising the steps of:

(a) introducing a nucleic acid cloned according to the method of claim 20 into a dicot cell; and

(b) growing the cell under cell growth conditions.

22. A method of identifying a candidate nucleic acid comprising a QTL associated with *Sclerotinia* whole plant field resistance or improved whole plant field resistance from a dicot, the method comprising:

(a) providing a nucleic acid cloned according to the method of claim 20; and,

(b) identifying a homolog of the nucleic acid in a dicot.

23. A method of marker assisted selection (MAS) of a quantitative trait locus (QTL) associated with whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, the method comprising the steps of:

(a) obtaining a first *Brassica* plant having at least one allele of at least one quantitative trait locus (QTL) associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, wherein the QTL is localized to linkage group N15, wherein said linkage group comprises at least one marker that is associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia* with a statistical significance of $p \leq 0.01$, wherein the QTL is localized to a chromosomal interval flanked by and including markers PE0286 and AG0369 on linkage group N15;

(b) crossing the first *Brassica* plant to a second *Brassica* plant;

(c) evaluating the progeny for the allele associated with the whole plant field resistance or improved whole plant field resistance to *Sclerotinia*; and

(d) selecting progeny plants that possess the allele.

24. The method of claim 23 wherein the plant is a member of a segregating population.

25. The method of claim 23 wherein the marker assisted selection is performed using high throughput screening.

26. A *Brassica* plant cell from a *Brassica* plant identified by the method of claim 23.

27. A *Brassica* plant cell from a progeny plant of the *Brassica* plant identified by the method of claim 23.

28. The *Brassica* plant cell from a progeny plant according to claim 27, wherein the progeny is selected from F1, F2, and F3 progeny.

29. An isolated or recombinant nucleic acid comprising a polynucleotide selected from:

(a) a sequence of marker sequence AG0369 (SEQ ID NO:19);

(b) a polynucleotide sequence with at least 90% sequence identity to the polynucleotide of (a); and

(c) a polynucleotide sequence complementary to the sequence of (a) or (b).

30. The isolated or recombinant nucleic acid of claim 29, wherein the nucleic acid is associated with whole plant field resistance or improved whole plant field resistance to *Sclerotinia*.

31. Use of an isolated or recombinant nucleic acid comprising a polynucleotide selected from:

(a) a sequence of marker sequence AG0369 (SEQ ID NO:19) and CA0719 (SEQ ID NO: 83);

(b) a polynucleotide sequence with at least 70% sequence identity to the polynucleotide of (a); and

(c) a polynucleotide sequence complementary to the sequence of (a) or (b)

for marker assisted selection of a *Brassica* plant or germplasm that exhibits whole plant field resistance or improved whole plant field resistance to *Sclerotinia*.

32. An isolated nucleic acid molecule for detecting a polymorphism in plant DNA associated with whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, wherein the nucleic acid molecule comprises at least 15 nucleotides and is identical to a sequence of the same number of consecutive nucleotides in either strand of the plant DNA where the polymorphism is located, wherein the nucleic acid molecule comprises a sequence that is at least 70% identical to a marker sequence or a fragment of a marker sequence selected from AG0369 (SEQ ID NO:19) and CA0719 (SEQ ID NO:83).

33. A kit for screening a plant or germplasm for a QTL associated with whole plant field resistance to *Sclerotinia*, comprising a container in which is contained:

(a) one or more of the isolated nucleic acid molecules of claim 32; and

(b) instructions for screening a plant for the QTL associated with whole plant field resistance or improved whole plant field resistance to *Sclerotinia*.

34. The kit of claim 33, comprising at least one component for high throughput screening the plant or germplasm for the QTL.

35. A *Brassica* plant cell from a *Brassica* plant that exhibits whole plant field resistance or improved whole plant field resistance to *Sclerotinia*, comprising alleles favorable to *Sclerotinia* whole plant field resistance in a QTL localized to a chromosomal interval flanked by and including markers PE0286 and AG0369 on linkage group N15.