Development of specific molecular markers to distinguish and quantify broomrape species in a soil sample from infected field

Radi Aly[1]*, Vinay K. Bari[1], Avishai Londner[1], Kocherman Marina[1], Jackline Abu Nassar[1], Leena Taha-Salaime[1], Eizenberg Hanan[1] and Ran Lati[1]

[1]Department of Plant Pathology and Weed Research, Newe Ya'ar Research Center, Agricultural Research Organization (ARO), Israel.

*Corresponding Author: Radi Aly

radi@volcani.agri.gov.il

https://orcid.org/0000-0002-5670-3798

https://orcid.org/0000-0002-6038-9586

Keywords: Plant parasite, *Phelipanche aegyptiaca, Orobanche* spp., ribulose-1,5-bisphosphate carboxylase (*rbcL1*), Internal transcribed spacer (*ITS*).

Conflict of interest: The authors have declared no conflict of interest

Abstract

Broomrapes (*Orobanche* and *Phelipanche*) are obligate holoparasites that cause heavy damage to numerous crops, reducing the yield and its quality. The parasite develops in the soil and exerts the greatest damage prior to its emergence; therefore the majority of field loss may occur before diagnosis of infection. Because of the parasite tiny seed size (200 to 300 μm) and dormancy for several decades in the field, it is very difficult to diagnose the parasite by conventional methods. Therefore, to restrict the parasite seeds spread and contamination to other commercial fields, development of DNA-based molecular markers to identify and quantify broomrape species in a soil sample is much needed. In this study, we developed a specific molecular marker (RbcL-M) based on *rbcL1* (large subunit of the ribulose-bisphosphate carboxylase) gene from *Orobanche crenata* to differentiate between *Orobanche crenata* and *Orobanche cumana.* Likewise, a specific marker (ITS100) based upon unique sequences in the internal transcribed spacer (ITS) regions of the nuclear ribosomal DNA of *Phelipanche aegyptiaca* to quantify three species of the parasite (*P.aegyptiaca*, *O.crenata* and *O.cumana*) in a soil sample was developed. Genomic DNA was extracted from soil samples artificially infested with broomrape seeds or tissue of *P.aegyptiaca*, *O.cumana* and *O.crenata* and subjected to PCR analysis. RbcL-M marker successfully amplified a PCR product (1300bp) when *O.crenata* seeds or tissues (collected from several locations in Israel) were added to the soil samples. The same marker amplified a PCR product (1000bp) when *O. cumana* seeds or tissues were added to the soil samples. RbcL-M marker did not amplify soil samples with seeds or tissues of *P.aegyptiaca* or any soil-borne DNA. Furthermore, using ITS-100 marker and Real-Time PCR analysis, allowed quantitative diagnostic of the parasite in a soil sample from infected sunflower field. As expected the universal internal control primer (UCP-555) amplified a PCR product (555bp) when genomic DNA extracted from soil samples with or without broomrape tissues. The development of an efficient, simple and

48 robust molecular marker to detect and distinguish between broomrape species, has a
49 significant insights on assessment the level of infestation and planning eradication program of
50 the parasite in a field crop.

51 **INTRODUCTION**

52 Broomrapes (*Orobanche* and *Phelipanche* spp.) are member of Orobanchaceae family
53 (Parker C., 1993), is an underground root parasite mostly affecting agricultural crop plants
54 (McNeal et al., 2013). These parasitic plants are unable to synthesize their photosynthetic
55 assimilate due to lack of chlorophyll; and totally depend on host plants for the germination
56 and growth by feeding through haustorium (Westwood et al., 2010a). The Orobanchaceae
57 family contains the largest number of parasitic species that form haustoria in roots, for
58 example *Orobanche, Phelipanche* species (Yoshida et al., 2016). Haustoria are special organ
59 of parasitic plant which invades host tissue and serve as connecting bridge that allows the
60 parasite to obtain water, mineral and nutrient from host (Joel and Losnergoshen, 1994;
61 Westwood et al., 2010b).The natural habitats of broomrapes are the warm and temperate
62 regions of the northern hemisphere, including the Mediterranean region (Parker C., 1993),
63 however branched broomrape *O.ramosa* native to central and south-western Europe, and
64 considered as a major threat to agricultural crops. They cause extensive damage by reducing
65 the yield of all economically important crops belongs to Solanaceae, Fabaceae, Brasicaceae,
66 Asteraceae and Apiaceae family (Joel, 2006; Parker C., 1993; Westwood et al., 2010b). In
67 Israel the significance of infested fields has increased dramatically during the last decade,
68 causing heavy crop damage or even total yield losses.

69 In general, parasitized crops suffer from reductions in total biomass at the greatest expense to
70 their productive tissue (Joel, 2006). Broomrapes cause many problems for agriculture farmers
71 because of their tiny seed size (200-300µm) which makes them difficult to detect in harvested

crop seeds and in agriculture soil (Joel, 1987). Moreover, small broomrape germinates and grows only in the presence of a susceptible host. In addition, small broomrape seeds can survive more than a decade in the soil (Habimana et al., 2014). Therefore, identification and quantification of broomrape seeds in a soil sample urgently required and will be helpful in prevention of crop damages by these parasites (Joel et al., 1996). The conventional methods of detection and differentiation such as seed coat morphological feature (Joel, 1987), Random amplified polymorphic DNA technique (RAPD) (Katzir et al., 1996; Paran et al., 1997), Intersimple sequence repeat (ISSR) (Benharrat et al., 2002) and high resolution melting analysis (Rolland et al., 2016) has some limitation to distinguish and quantify broomrapes exist in the soil. Recent progress in sequencing broomrape plastid genome, which was made available, providing new insight to develop molecular marker for specific broomrape species identification (Cusimano and Wicke, 2016; Wicke et al., 2013). Genetic or DNA based molecular marker such as internal transcribe spacer (ITS), plastidial ribulose bisphosphate carboxylase (Rubisco) large subunit (*rbcL1*) and maturase K (*matK*) are among the most widely used molecular markers in evolutionary, taxonomy, phylogenetic and diversity identification of many organisms including *Orobanche*, have been established in recent years (Agarwal et al., 2008; Park et al., 2008; Schneeweiss et al., 2004).

In order to develop a complementary methodology for mapping parasitic plants in the field, a procedure includes: a known geographical information systems (GIS) for soil sampling that will characterize the spatial variation in the field (Eizenberg et al., 2012), and at the same time, use of molecular markers to identifying broomrapes in a soil sample. Those molecular markers would assist detection and diagnosis of broomrapes species and population level in the soil. In this study we developed a simple and easy method based on *rbcL1* gene to differentiate between three broomrape species *O.crenata, O.cumana* and *P.aegyptiaca*, the

96 most common and destructive weed in Israel. We also developed a specific molecular marker
97 to quantify the parasite seeds in soil sample in an infected sunflower field.

98 **RESULTS**

99 **Design of *rbcL1* specific marker to differentiate between *O. cumana* and *O. crenata***

100 In our previous study, we reported amplification of ITS region using ITS350 primer; which
101 specifically detect broomrape in soil sample (Aly et al., 2012). However, ITS350 marker does
102 not differentiate between broomrape species because of the high sequence similarity among
103 ITS regions. Therefore, we compared and analysed the *rbcL1* gene sequences from these
104 species using Clustal Omega and designed new molecular marker to differentiate broomrape
105 species such as *O.cumana* and *O.crenata* (Figure S1). To distinguish between the above
106 broomrape species, probes differing at least by five-base mismatches or more were designed.
107 Based on sequence alignment, a molecular marker (RbcL-M) was designed using *rbcL1* gene
108 from *O.cumana* (GenBank accession number: AF090349.1) Soil samples were artificially
109 contaminated with seeds or shoots of *P. aegyptiaca*, *O.crenata* and *O.cumana* and subjected
110 to PCR analysis. By PCR amplification using RbcL-M, we were able to differentiate between
111 *O.crenata* and *O.cumana* according to the length of the PCR products. For *O.crenata* the
112 PCR product size was 1300bp compared to 1000bp for *O. cumana*. However, RbcL-M
113 marker failed to amplify any PCR product while soil samples were artificially contaminated
114 with *P. aegyptiaca* seeds or shoots (Figure 1A). As expected, the universal internal control
115 primers amplified a PCR product (555bp) from soil samples with or without broomrape
116 shoots or seeds (Figure 1B).

117 **Development of specific ITS100 marker to detect seeds of three broomrape species in a**
118 **soil sample**

119 To analyse the number of the broomrape seeds present in a soil sample from infected field,
120 we designed ITS100 primers from the highly conserved region of ITS350 region (Aly et al.,
121 2012). To determine the specificity of the marker to each broomrape specie, we first extracted
122 genomic DNA from soil samples (250mg) containing (0, 1, 10, 25, 50 100 and 250 seeds) of
123 *P.aegyptiaca*, *O.crenata* or *O.cumana* and subjected to PCR analysis. ITS100 primers
124 successfully amplified a PCR product (100bp) (Figure 2A) when *P.aegyptiaca*, *O.crenata* or
125 *O.cumana* were mixed with soil sample, except soil sample without *Orobanche* seed.
126 However, we have got PCR product in all the sample including soil without *Orobanche* seeds
127 using universal primer UCP-555 (Figure 2B).

128 In addition, the specificity of the markers generated in this study were evaluated by testing
129 DNA extracted from sunflower, carrot and tomato leaves and roots corresponding to possible
130 respond to the host plants. Fortunately, the qPCR assay always yielded negative results for
131 these plant species (data not shown). Next to quantify *Orobanche*, we generated standard
132 curve using *O.cumana* genomic DNA and ITS100 marker utilizing qPCR procedure. *O.*
133 *cumana* genomic DNA was purified and serial dilutions (1:10 fold) ranging from 0.001 to 10
134 ng/µl were used for qPCR assay. The standard curve for *O.cumana* exhibited a slope of -1.28
135 and (R^2) of 0.9994 (Figure 2C). The same type of assay was performed with *O.cumana*
136 genomic DNA extracted from soil sample (250 mg) mixed with the parasite seeds (0.1, 1, 10
137 and 100 mg seeds) and subjected to qPCR. The standard curve for *O.cumana* seeds mixed
138 with soil sample, exhibited a slope of -1.332 and (R^2) of 0.987 (Figure 2D).

139 **Prediction and quantification of *O. cumana* seeds in sunflower field**

140 Previously, we have developed a protocol that allows extraction of genomic DNA from few
141 tiny seeds of *Orobanche* spp. in a soil sample and we were able to subject the sample DNA to
142 a rapid PCR-assay (Aly et al., 2012). Here, we extended the protocol, allowing for the

143 identification and distinguishing between soil-borne and parasite seed species that are
144 collected from a field soil sample. Following confirmation and validation of the ITS100
145 marker specificity to *O.cumana* seeds, geo-statistics model for soil sampling that
146 characterized the spatial variation in the field, was performed according to (Eizenberg et al.,
147 2012) in a sunflower field located at Havat Eden, Emic Betshan– Israel. This field was
148 infected with *O. cumana* (Figure 3A). Five soil samples were collected from a plot of 1 acre
149 according to the illustration (Figure 3B). Genomic DNA was extracted from each soil sample
150 and subjected to qPCR using ITS100 primers to predict and quantify *O.cumana* DNA in
151 naturally infected field. Then, genomic DNA from soil samples was extrapolated with Ct
152 values to quantify the number of seeds in each sample according to the generated Ct values
153 (Figure 2C and D). The results showed variable quantity of *O.cumana* seeds in the soil
154 sample ranged from 2 to 22 parasite seeds (Figure 3C) in 250mg soil sample processed
155 originally from 1kg soil sample. Sample no. 6 showed the highest density of *O.cumana* seeds
156 (22 seeds) in the sunflower infected field as compared to the other samples. Sample no. 6 was
157 collected from a highly infected area as was shown by the geo-statistics model proposed by
158 (Eizenberg et al., 2012). Concomitantly, we monitored and counted the parasite
159 inflorescences in adjacent to the position were the soil samples were taken. Our results
160 revealed that average density of 0.6 inflorescences /m^2, in a total area of 1acre was recorded.
161 Inflorescences counts in the selected area in the field were also variable ranged from 2 to 15
162 and were less than the seed numbers found in the soil sample when qPCR was applied
163 (Figure 3C).

164 **Evaluation viability of the parasite seeds in sunflower infected field**

165 To determine viability of broomrape seeds in a sunflower infected field, a tetrazolium test
166 (TZ) (Lopez-Granados and Garcia-Torres, 1999; Thorogood et al., 2009) was conducted.
167 According to TZ test, viable parasite seeds tends to have orange to brown colour as a result of

168 its metabolite activity while non-viable tends to have yellow to faint colour (Figure 4A). To
169 evaluate the parasite seed viability, two treatments were assayed: soil samples (organic matter
170 only) collected from sunflower field located at Havat Eden, Emic Betshan, naturally infested
171 with *O. cumana* and pure *O. cumana* seeds (1 mg) prepared from a seed stock. Our results
172 indicate that the highest percentage (% of total) of viable *O. cumana* seeds were found in soil
173 samples collected from the sunflower field (83%) as compared to 25% found in a seed stock
174 (Figure 4B).

175 **DISCUSSION**

176 Parasitic weeds are often exert their greatest damage prior to their emergence; therefore the
177 majority of field loss may occur before diagnosis of infection. Early detection of infection by
178 visual inspection is not possible. This hampers early control strategy efforts and the detection
179 methods used by crop scientists are laborious and time consuming. Most of the methods
180 developed to detect broomrape seeds in soil samples, need several steps such as seed
181 purification (flotation, filtration and binocular microscope observation (Kroschel, 2001).
182 Despite of the great advances in genome sequencing approaches that has already been
183 achieved, our understanding towards genomic evolution of Orobancheaceae is still
184 incomplete due to aberrant evolution of plastid genome of holoparasite (Park et al., 2007).
185 Available DNA sequence data are insufficient to differentiate closely related species, such as
186 the weedy parasites *O.cumana, O.crenata* and *P.aegyptiaca*.

187 Genetic or DNA based marker techniques such as ITS or *rbcL1* are routinely being used in
188 evolutionary, taxonomical, phylogenetic and diversity studies (Agarwal et al., 2008; Park et
189 al., 2008; Schneeweiss et al., 2004). Diagnosis and early identification of the parasite species
190 in the field by soil sampling is of great importance to farmers and it is particularly needed due
191 to the host-parasite specific interaction, the difficulty to detect the tiny parasite seeds by

192 conventional methods in contaminated fields and because seeds will only germinate and grow
193 in the presence of a susceptible host. The gene encoding the large subunit of ribulose-1,5-
194 bisphosphate carboxylase/oxygenase (Rubisco; *rbcL1*) is retained as a pseudogene in
195 *Orobanche* and *Phelipanche* (Wicke et al., 2013). In the present work, we designed sets of
196 primers targeting the plastid *rbcL1* gene sequence from *O.cumana* to differentiate between
197 three broomrape species (*O.cumana, O.crenata* and *P.aegyptiaca*) the most common and
198 harmful species in field crops in Israel. Another sets of primers were designed to target ITS
199 sequences (Schneeweiss et al., 2004) to quantify the parasite seeds potentially present in a
200 soil sample to help for seed bank assessment in the field. We found that RbcL-M primers
201 (marker) were useful to distinguish between DNA from *O.cumana* and *O. crenata* producing
202 different amplicon sizes (a specific PCR product about 1300bp for *O. crenata* and 1000bp for
203 *O.cumana*). Subsequent BLASTn with previously published broomrape genomic data
204 (Cusimano and Wicke, 2016) reconfirmed the differences between the two species. In
205 contrast, DNA from *P.aegyptiaca* failed to produce PCR amplicon suggesting possibly that
206 *Phelipanche* species completely lost *rbcL1* gene (Delavault et al., 1995; Leebens-Mack and
207 de Pamphilis, 2002).To allow differentiation between *P.aegyptiaca* from other species, a
208 specific molecular marker was recently designed (Aly et al., 2019), (GenBank accession
209 numbers MK637618-637624).
210 To diagnose and quantify broomrape species in a seed stock (Dongo et al., 2012), different
211 types of nuclear and plastid DNA markers have been proposed. (Schneeweiss et al., 2004)
212 were the first to present molecular phylogenetic analysis using nuclear ITS sequences. ITS-
213 based markers were also used to detect *P.aegyptiaca* seeds in a soil sample (Aly et al., 2012)
214 and quantify contamination of *O.ramosa* and *O. cumana* in crop seed lots (Dongo et al.,
215 2012). Here we developed ITS100 marker that was based on primers consisting of unique
216 sequences in the internal transcribed spacer (ITS) regions of the nuclear ribosomal DNA

217 (nrDNA) of *O. crenata*. ITS-100 marker was used with qPCR assay to quantify *O.cumana*
218 seeds in a soil sample from sunflower field located at Havat Eden, Emic Betshan – Israel.
219 We were able to detect 0.1 mg *O.cumana* seeds in 250 mg soil sample (wt / wt). A detection
220 threshold of 0.1 mg broomrape seeds in 20 g seed samples was previously reported (Dongo et
221 al., 2012). For detection, mapping and quantifying *O. cumana* seeds in the field, we used
222 geographical information systems (GIS) for soil sampling and other advanced technologies
223 for parasitic weed mapping and field history data storage (Eizenberg et al., 2012) followed by
224 qPCR assay. Soil samples (500gm) were collected from the sunflower field, organic material
225 was extracted from each sample ending with 250mg then, genomic DNA was extracted and
226 subjected to qPCR using ITS100 primers. This method allowed specifically detecting and
227 quantifying the DNA of *O.cumana* in a total DNA extract from sunflower soil sample.
228 Accordingly, the results of this assay can be also expressed as the number of parasite seeds
229 per kilogram of soil following extrapolation with the standard curve was prepared. Our
230 results indicate that samples collected from highly infected area (Fig. 4A) according to the
231 geo-statistics model proposed by Eizenberg et al. (2012), were with agreement with our
232 qPCR assay (sample no. 6 showed the highest density of *O.cumana* seeds (22 seeds) in the
233 sunflower infected field as compared to the other samples. However, no correlation was
234 found between qPCR assay (parasite seed number) compared to number of the parasite
235 inflorescences collected from the same location in the sunflower infected field. An
236 explanation for that may be related to the viability of *O.cumana* seeds, genomic DNA from a
237 soil sample will contain viable and non-viable parasite seeds, add to that, we cannot exclude
238 the presence of some related parasite seeds in the same sample therefore, the we counted
239 more parasite seeds using the qPCR assay. Specificity of the qPCR assay was tested against
240 several possible contaminants of soil-borne pathogens by using universal internal control

241 primer UCP-555 (White, 1990) or harvested crop seeds like sunflower and tomato using
242 ITS100. No amplification was observed, confirming the specificity of the marker.
243 The powerful of broomrape-infected field to distribute and contaminate the neighbouring
244 non-infested fields depends on the soil seed-bank and viability of the parasite seeds. Viability
245 of broomrape seeds in the sunflower infected field was determined by tetrazolium test. Our
246 results showed the highest percentage of viable broomrape seeds were found in soil samples
247 collected from the sunflower field (83%) as compared to 25% found in a seed stock. We
248 assume that the highest count recorded in the sunflower field was due to the release and
249 distribution of the fresh parasite seeds by the newly parasite inflorescences showed up
250 through the crop growth. We have to take in consideration that we used an old seed stock (10
251 years old) that doesn't represent newly harvested seed stocks. Our experience with seed
252 germination of *O. cumana* with germination stimulant (GR) could reach more than 90%.
253 Additionally, we cannot exclude the presence of some related parasite seeds in the same
254 sample from previous growth seasons.
255 In this study, we provide a simple, fast and non-expensive approach to distinguish and
256 quantify broomrape seeds exist in a soil sample from a crop field. Molecular markers would
257 assist accurate detection and population level of *Phelipanche* and *Orobanche* spp. in a soil
258 sample and offer numerous advantages over conventional phenotype based alternatives, as
259 they are stable and detectable in all tissues regardless of growth, differentiation or
260 development. These methods could be helpful in precision agriculture, in which they provide
261 answers routinely questioned by the farmers: are there parasite seeds in my crop field? What
262 species? And how much seeds are exist in a soil sample?

263

264

265 **EXPERIMENTAL PROCEDURE**

266 **Plant material**

267 Seeds, shoots and inflorescences of three broomrape species (*P.aegyptiaca, O.cumana and O.*
268 *crenata*) were collected from different Israeli locations (Golan heights, Havat Eden- Emic
269 Betshan and Jordan valley respectively). Seeds were harvested from freshly inflorescences
270 and stored at 4°C while shoots and inflorescences were dried and stored at −80°C until used.
271 Two hundred milligrams of parasite plant seeds/tissue were ground manually in presence of
272 liquid nitrogen with the help of pestle and mortar. Total genomic DNA was extracted using
273 the GeneGET Plant genomic DNA extraction kit according to the manufacturer instructions
274 (Thermo Scientific).

275

276 **ITS and *rbcL1* sequences and primers design**

277 The internal transcribed spacer (ITS) regions of *Phelipanche aegyptiaca* (GenBank-
278 Accession No. AY209327) was used to design ITS100 marker. The PCR product yielded a
279 100bp fragment in length. UCP-555 used as universal internal control primer, which
280 amplifies a region of the small subunit of nrDNA (555 bp) from a wide variety of
281 microorganism such as protists, fungi, and plants. The length of the *rbcL1* sequences was
282 1211bp and 1290bp for *O.cumana* (GenBank: AF090349.1) and *O. crenata* (GenBank:
283 AY582191.1) respectively. Sequence similarities between broomrape species were done
284 using Clustal Omega software. Design of primers for the qPCR detection of broomrape
285 species was based on the alignment of the ITS of nuclear ribosomal DNA regions. All
286 primers were designed with Gene Runner software. PCR amplifications of *rbcL1* sequences
287 were performed with the RbcL-M-F and RbcL-M-R primers using PCR Ready Mix
288 (BioLabs, Israel), and 50ng DNA-template. The following PCR conditions were used:

289 denaturation for 5 min at 95°C; 32 cycles with 30 second at 95°C, 30 second at 55°C, 1.30
290 min at 72°C; and final elongation for 10 min at 72°C.

291 **Soil sampling and extraction of Genomic DNA**

292 Design and mapping the soil samples in the sunflower tested field were performed according
293 to (Eizenberg et al., 2012). Soil samples were collected from the infected field on May and
294 counting the inflorescences was performed on July. Each sample includes 500gm of a soil
295 sample that was introduced into one-litter container, mixed thoroughly with 1.5L of (13.3 M
296 $CaCl_2$) solution and kept at room temperature for overnight. The next day, the upper phase
297 including the organic materials and broomrape seeds were collected, dried on Whatmann
298 paper and the parasite seeds were obtained by filtration through 50 and 100 mesh. Genomic
299 DNA was extracted by PowerSoil kit (MO BIO Lab. Inc., Loker Ave West, Carlsbad, CA).
300 according to the manufacturer's instructions (Sagova-Mareckova et al., 2008).
301

302 **Quantitative real time PCR of soil sample from the sunflower fields**

303 The quantitative real time PCR was performed using PerfeCTa® SYBR® Green Fast Mix®,
304 ROX™ (Quanta biosciences) and ABI-Prism 7000 Real-Time PCR Detection System
305 (Applied Biosystems) according to the manufacturer's protocol. The qPCR was performed in
306 quantitative reaction with final volume of 10 µl including: 100ng DNA, 5 µl of SYBR Green
307 Fast Mix- ROX and 500nM of each ITS-100 primers. *O.cumana* actin was used as
308 endogenous control and the relative gene expression level was calculated using the $2^{-\Delta\Delta Ct}$
309 method (Livak and Schmittgen, 2001). Two standard curves were prepared: The first curve
310 included standard points from *O.cumana* seeds ranging from 0.001 to 10 ng DNA per tube,
311 were made using 1:10-fold serial dilutions of broomrape DNA and were used for qPCR
312 assay. The second curve was generated by mixing different amounts of genomic DNA from

313 *O.cumana* (0.1 to 100 ng) with total genomic DNA (10 µg) extracted from 250 mg soil
314 sample. Each point on the standard curve was assayed in triplicate.

315 **Viability evaluation of the parasite seeds by Tetrazolium test**

316 Tetrazolium test was performed as described earlier (Lopez-Granados and Garcia-Torres,
317 1999; Thorogood et al., 2009). Briefly, 50-100 seeds were hydrated between paper towel
318 sheets for 2-3 days. After pre-conditioning, they were placed in flasks, covered with a 1%
319 solution of 2,3, 5 triphenyl tetrazolium chloride (TZ) and incubated at 40°C for 2h. After that,
320 the TZ solution was discarded; and the seeds were washed thoroughly with sterile water.
321 Seeds were directly evaluated for viability under the microscope using magnification (40X
322 and 100X) and classified in two categories: viable seeds – orange to brown light coloured or
323 non-viable seeds –yellow to white colour.

324 **ACKNOWLEDGMENTS**

325 Results of this research were supported by the Chief Scientist of the Ministry of Agriculture
326 and Rural Development - Israel, grant No.132-1499-10 (MEZAMALEKET). VKB is grateful
327 to the ARO-Volcani Center, Agricultural Ministry of Israel for providing the Postdoctoral
328 fellowship.

329

330

331

332

333

334

335 **REFERENCES**

336 Agarwal, M., Shrivastava, N., and Padh, H. (2008). Advances in molecular marker techniques and
337 their applications in plant sciences. Plant Cell Rep *27*, 617-631.

338 Aly, R., Eizenberg, H., Kocherman, M., Abu-Nassar, J., Taha, L., and Saadi, I. (2012). Use of ITS
339 nuclear sequences from Phelipanche aegyptiaca as a direct tool to detect single seeds of broomrape
340 species in the soil. Eur J Plant Pathol *133*, 523-526.

341 Aly, R., Lati, R., Abu-Nassar, J., Ziadna, H., Achdari, G., Münchow, C.S.v., Wicke, S., Bari, V.K.,
342 and Eizenberg, H. (2019). The weedy parasite Phelipanche aegyptiaca attacks Brassica rapa var. rapa
343 L. for the first time in Israel. Plant Disease Note.

344 Benharrat, H., Veronesi, C., Theodet, C., and Thalouarn, P. (2002). Orobanche species and population
345 discrimination using intersimple sequence repeat (ISSR). Weed Res *42*, 470-475.

346 Cusimano, N., and Wicke, S. (2016). Massive intracellular gene transfer during plastid genome
347 reduction in nongreen Orobanchaceae. New Phytol *210*, 680-693.

348 Delavault, P., Sakanyan, V., and Thalouarn, P. (1995). Divergent evolution of two plastid genes, rbcL
349 and atpB, in a non-photosynthetic parasitic plant. Plant Mol Biol *29*, 1071-1079.

350 Dongo, A., Leflon, M., Simier, P., and Delavault, P. (2012). Development of a high-throughput real-
351 time quantitative PCR method to detect and quantify contaminating seeds of Phelipanche ramosa and
352 Orobanche cumana in crop seed lots. Weed Res *52*, 34-41.

353 Eizenberg, H., Aly, R., and Cohen, Y. (2012). Technologies for Smart Chemical Control of
354 Broomrape (Orobanche spp. and Phelipanche spp.). Weed Sci *60*, 316-323.

355 Habimana, S., Nduwumuremyi, A., and Chinama, R.J.D. (2014). Management of orobanche in field
356 crops- A review. J Soil Sci Plant Nut *14*, 43-62.

357 Joel, D.M. (1987). Detection and Identification of Orobanche Seeds Using Fluorescence Microscopy.
358 Seed Sci Technol *15*, 119-124.

359 Joel, D.M., Hershenhorn, Y., Eizenberg, H., Aly, R., Ejeta, G., Rich, P.J., Ransom, J.K., Sauerborn, J.
360 and Rubiales, D. (2006). Biology and Management of Weedy Root Parasites. In Horticultural
361 Reviews, New Jersey: Wiley, 267–349.

362 Joel, D.M., and Losnergoshen, D. (1994). The Attachment Organ of the Parasitic Angiosperms
363 Orobanche Cumana and O-Aegyptiaca and Its Development. Can J Bot *72*, 564-574.
364 Joel, D.M., Portnoy, V., and Katzir, N. (1996). Identification of single tiny seeds of Orobanche using
365 RAPD analysis. Plant Mol Biol Rep *14*, 243-248.
366 Katzir, N., Portnoy, V., Tzuri, G., Joel, D.M., and Castejon-Munoz, M. (1996). Use of random
367 amplified polymorphic DNA (RAPD) markers in the study of the parasitic weed Orobanche. Theor
368 Appl Genet *93*, 367-372.
369 Kroschel (2001). A Technical Manual for Parasitic Weed Research and Extension. Springer,
370 Dordrecht.
371 Leebens-Mack, J., and de Pamphilis, C. (2002). Power analysis of tests for loss of selective constraint
372 in cave crayfish and nonphotosynthetic plant lineages. Mol Biol Evol *19*, 1292-1302.
373 Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time
374 quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408.
375 Lopez-Granados, F., and Garcia-Torres, L. (1999). Longevity of crenate broomrape (Orobanche
376 crenata) seed under soil and laboratory conditions. Weed Sci *47*, 161-166.
377 McNeal, J.R., Bennett, J.R., Wolfe, A.D., and Mathews, S. (2013). Phylogeny and origins of
378 holoparasitism in Orobanchaceae. Am J Bot *100*, 971-983.
379 Paran, I., Gidoni, D., and Jacobsohn, R. (1997). Variation between and within broomrape
380 (Orobanche) species revealed by RAPD markers. Heredity *78*, 68-74.
381 Park, J.M., Manen, J.F., Colwell, A.E., and Schneeweiss, G.M. (2008). A plastid gene phylogeny of
382 the non-photosynthetic parasitic Orobanche (Orobanchaceae) and related genera. J Plant Res *121*,
383 365-376.
384 Park, J.M., Manen, J.F., and Schneeweiss, G.M. (2007). Horizontal gene transfer of a plastid gene in
385 the non-photosynthetic flowering plants Orobanche and Phelipanche (Orobanchaceae). Mol
386 Phylogenet Evol *43*, 974-985.
387 Parker C., R.C.R. (1993). Parasitic weeds of the world : biology and control (Wallingford, Oxon :
388 CAB International).

389 Rolland, M., Dupuy, A., Pelleray, A., and Delavault, P. (2016). Molecular Identification of
390 Broomrape Species from a Single Seed by High Resolution Melting Analysis. Front Plant Sci *7*.
391 Sagova-Mareckova, M., Cermak, L., Novotna, J., Plhackova, K., Forstova, J., and Kopecky, J. (2008).
392 Innovative methods for soil DNA purification tested in soils with widely differing characteristics.
393 Appl Environ Microbiol *74*, 2902-2907.
394 Schneeweiss, G.M., Colwell, A., Park, J.M., Jang, C.G., and Stuessy, T.F. (2004). Phylogeny of
395 holoparasitic Orobanche (Orobanchaceae) inferred from nuclear ITS sequences. Mol Phylogenet Evol
396 *30*, 465-478.
397 Thorogood, C.J., Rumsey, F.J., and Hiscock, S.J. (2009). Seed viability determination in parasitic
398 broomrapes (Orobanche and Phelipanche) using fluorescein diacetate staining. Weed Res *49*, 461-
399 468.
400 Westwood, J.H., Yoder, J.I., Timko, M.P., and dePamphilis, C.W. (2010a). The evolution of
401 parasitism in plants. Trends in Plant Science *15*, 227-235.
402 Westwood, J.H., Yoder, J.I., Timko, M.P., and dePamphilis, C.W. (2010b). The evolution of
403 parasitism in plants. Trends Plant Sci *15*, 227-235.
404 White, T.J., T. D. Bruns, S. B. Lee, and J. W. Taylor (1990). Amplification and direct sequencing of
405 fungal ribosomal RNA Genes for phylogenetics.
406 Wicke, S., Muller, K.F., de Pamphilis, C.W., Quandt, D., Wickett, N.J., Zhang, Y., Renner, S.S., and
407 Schneeweiss, G.M. (2013). Mechanisms of Functional and Physical Genome Reduction in
408 Photosynthetic and Nonphotosynthetic Parasitic Plants of the Broomrape Family. Plant Cell *25*, 3711-
409 3725.
410 Yoshida, S., Cui, S.K., Ichihashi, Y., and Shirasu, K. (2016). The Haustorium, a Specialized Invasive
411 Organ in Parasitic Plants. Annu Rev Plant Biol *67*, 643-667.

412

413

414

415 **Figure legends**

416 **Figure 1. (A)** Specificity of RbcL-M marker and PCR amplification patterns in three
417 broomrape species (*P.aegyptiaca, O.cumana and O.crenata*). Soil samples (250 mg) were
418 artificially contaminated with broomrape seeds (100) or 10 mg shoots of *P.aegyptiaca,*
419 *O.crenata and O.cumana* and subjected to PCR analysis using RbcL-M primers. Genomic
420 DNA was extracted and 100 ng were used for PCR amplification. A PCR product of 1300bp
421 was obtained from *O.crenata* seeds and shoots, A PCR product of 1000bp was obtained from
422 *O.cumana* seeds and shoots. However, RbcL-M marker failed to amplify PCR product with
423 soil sample alone or with soil contaminated with *P.aegyptiaca* seeds or shoots. **(B)** The same
424 samples were subjected to PCR detection using universal internal control primers (UCP-555).
425 Arrows indicates the PCR product sizes.

426 **Figure 2.** PCR amplification patterns of ITS100 primers in *O.cumana* seeds **(A)** gDNA was
427 extracted from soil samples (250 mg) containing the parasite seeds (1, 10, 25, 50, 100 and
428 250 seeds). For each sample, 100 ng gDNA were used for PCR amplification. The same
429 samples were subjected to PCR detection using universal internal control primers (UCP-555).
430 Arrows indicates the PCR product sizes **(B)** Standard curves generated for qPCR to quantify
431 *O.cumana* seeds in a soil sample. standard curve generated by plotting the value of threshold
432 cycle value (Ct) against log of the amount of template DNA (ng/µl) from *O.cumana.***(C)**
433 Detection and quantification of *O. cumana* template DNA (ng) from known amount of the
434 parasite seeds mixed with extracted soil sample DNA (10 µg). Each point on the standard
435 curve was assayed in triplicate.

436 **Figure 3.** Quantification of *O.cumana* seeds in an infected sunflower field by qPCR and
437 manually counting the inflorescences adjacent to the soil sample. **(A)** represnts a section of
438 the sunflower infected field. Red arrows show *O.cumana* inflorescences. **(B)** Representing a

439 (GIS) model (Eizenberg et al., 2012) for soil sampling and characterize the spatial variation
440 in sunflower field infected with *O.cumana* in Havat Eden. Five soil samples (2, 4, 6, 8, 10)
441 of 500gm were collected from the infected field on May. Each sample was collected from a
442 depth of 0 to 20 cm in a total area of 1x1 meter. The blue spots indicate location of the
443 sample in the infected field. Red color in the plot represents high density of the parasite
444 inflorescences. Monitoring infectivity by counting the inflorescences was performed on July
445 and restricted to the total area 1x1 meter of the five soil samples. **(C)** quantification of *O.*
446 *cumana* seeds in a soil sample by qPCR (black bars) and monitoring the parasite
447 inflorescences counts in fields (Grey bars). The data are the means of three soil sample or
448 biological replicates. Vertical lines indicate SD of three independent measurements.

449 **Figure 4.** Evaluation the viability of *O.cumana* seeds by Tetrazolium test. **(A).** Tetrazolium
450 test was used to visually differentiate between metabolically active tissue from (cleaned
451 viable seed (a), viable seed with organic matter from an infected sunflower field (c), and
452 inactive tissues from (cleaned non-viable seed (b), non-viable seed with organic mater from
453 an infected sunflower field (d). Images were acquired using fluorescence microscope using
454 40X and 100X magnification. **(B)** The same test was also performed to evaluate percentage
455 of viability of *O.cumana* seeds in soil samples from the infected field (83%) as compared to
456 viability of the seeds from cleaned seed bank (25%). The data are the means of three separate
457 experiments.

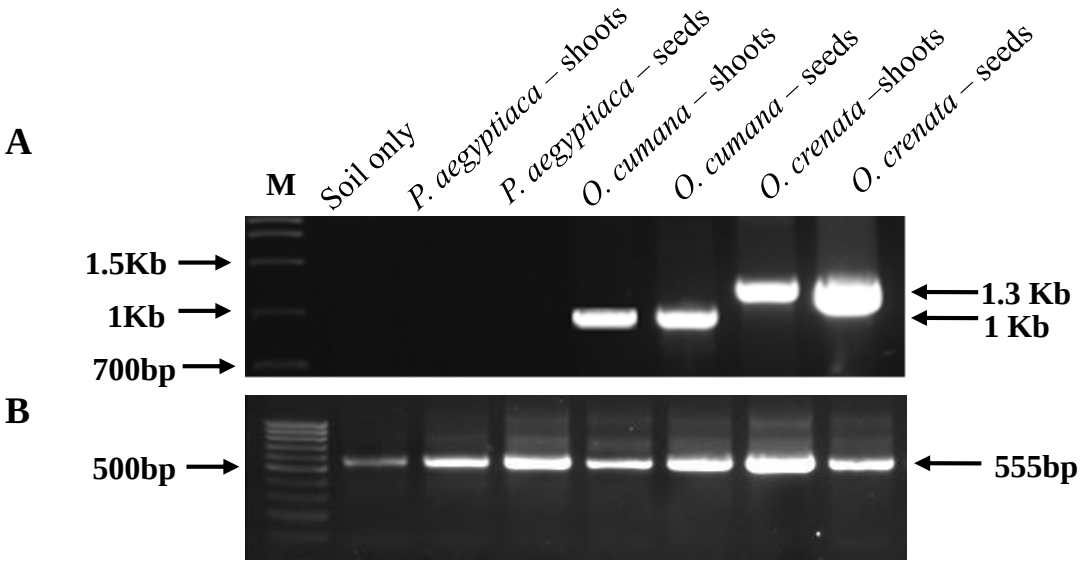


Figure 1_Radi Aly et al., 2019

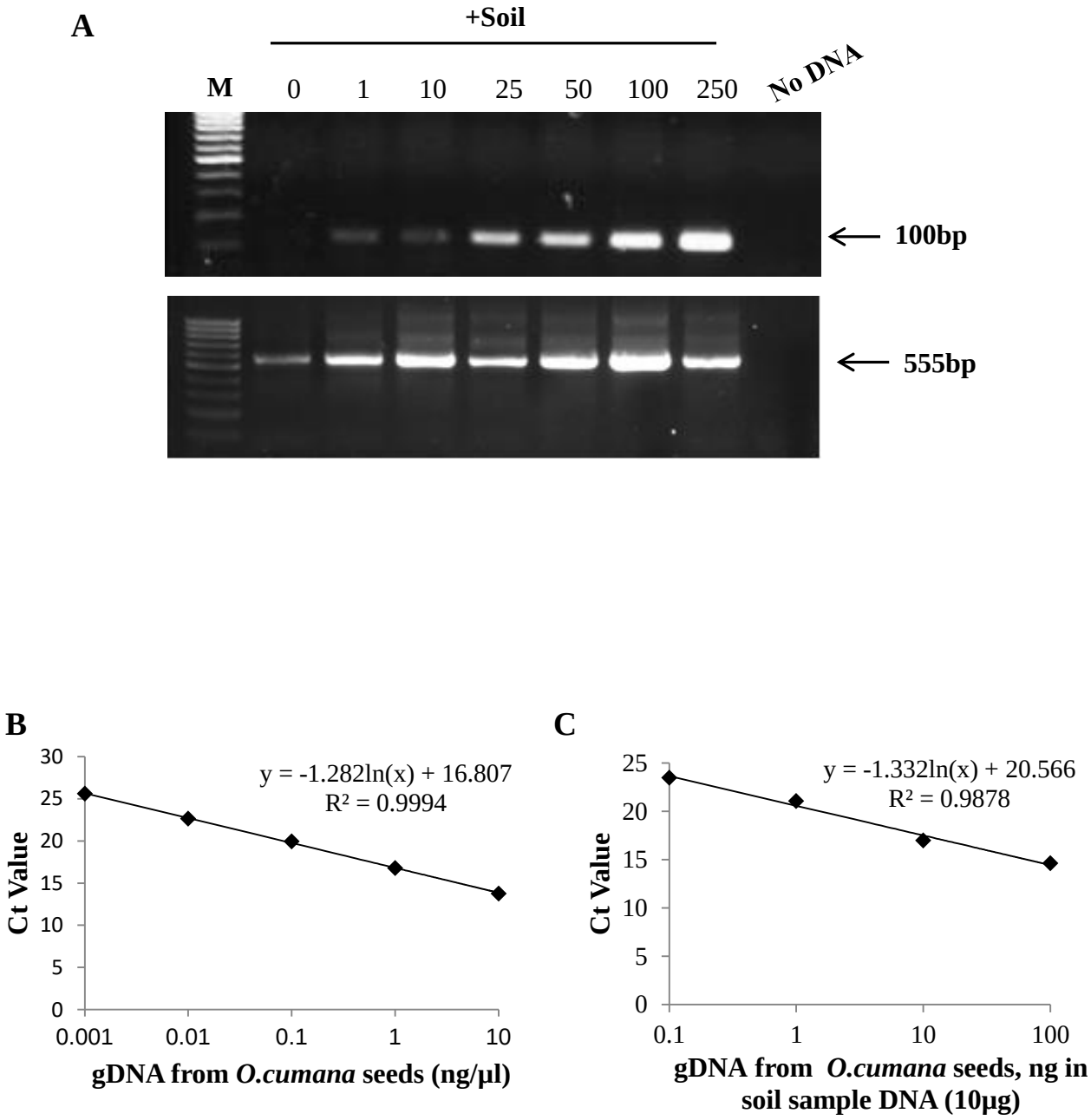
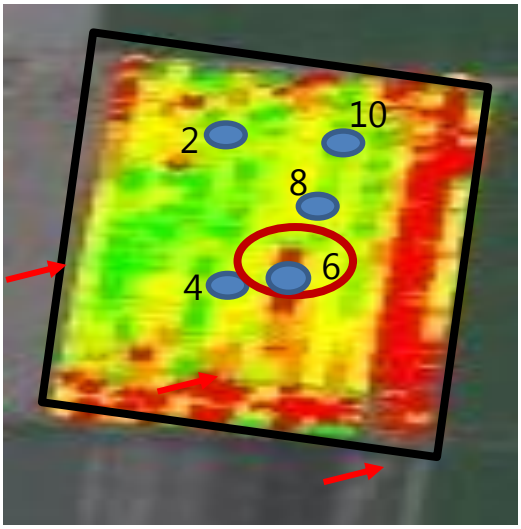


Figure 2_Radi Aly et al., 2019

A

B



C

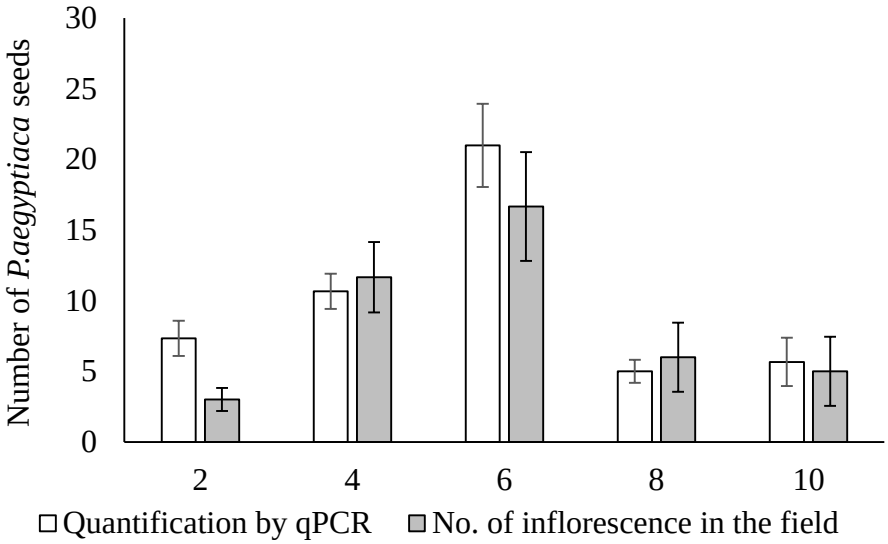
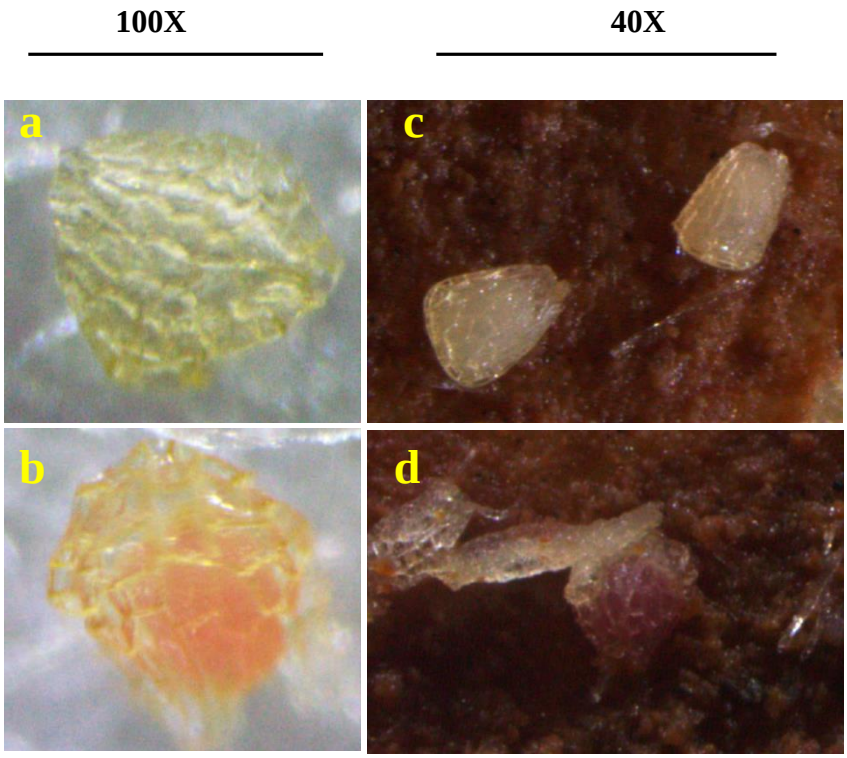


Figure 3_Radi Aly et al., 2019

A



B

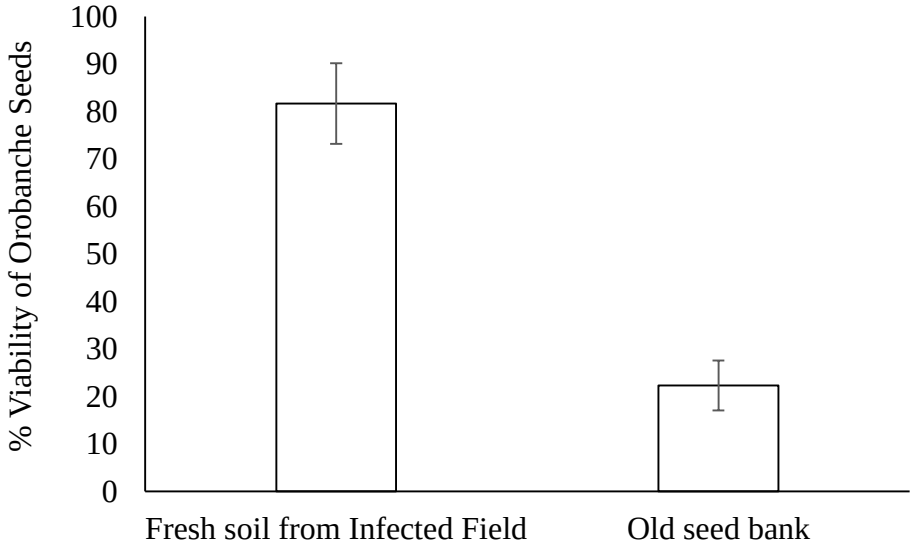


Figure 4_Radi Aly et al., 2019

Supplementary Information

Development of specific molecular markers to distinguish and quantify broomrape species in agricultural field soil sample

Radi Aly[1]*, Vinay K. Bari[1], Avishai Londner[1], Kocherman Marina1, Jackline Abu Nassar[1], Leena Taha-Salaime[1], Eizenberg Hanan[1] and Ran Lati[1]

[1]Department of Plant Pathology and Weed Research, Newe Ya'ar Research Center, Agricultural Research Organization (ARO), Israel.

Figure S1. Multiple sequence alignment of *O.cumana* and *O.crenata rbcL1* gene using online tool CLUSTAL Omega

```
                 RbcL-M-F
O.cumana        ------------------------------------------------------------ 0
O.crenata       ATGTCACCACAAACAGAAACTAAAGCAAGTGTTGGATTCAAAGCGGGTGTTAAAGATAAA 60

O.cumana        ------------------------------------------------------------ 0
O.crenata       GAGTACAAATTGACTTATTATATTCCTGAATACAAAACTAAAGATACTGATATTTTAGCA 120

O.cumana        ------------------------------------------------------------ 0
O.crenata       GCATTTCGAGTAACTCCTCAACCTGAAGTTCCGCCTGAACAGGGGCTGCAGTAGCTGCCG 180

O.cumana        ------------------------------------------------------------ 0
O.crenata       AATCTTCTATTGGTACATGGACAACTGTGTGGACTGATGGACCCTTGATCGTTACAAAGG 240

O.cumana        -------------------------------TATCGAGTA----GACCTTGTGTTGTTATC 26
O.crenata       TCGATACTACCACATCACATCGAGCCCAGTTCTGGAGAAACAGATCAATATATTTGTTAT 300
                                               * * *** *      *  * * **  *
                                                                RbcL-M-F
O.cumana        AGAATTTTTAATTCATAAGTTTTAGGGAGAAATTTATGTCACC--------ACAAACAGA 78
O.crenata       ATAGCTTACCCTTTAGAC-TTTTTTGAGGAAGGTTCTGTTACTAACATGTTTACTTCCAT 359
                *  *  **    ** * *  ****  *  ***  ** *** **              *
O.cumana        AACTAAAGCAAGTGTTGGATTCAAAGTCCTGCGTGATCTATGTCTGGAATATCTGCGAAT 138
O.crenata       TATAGGAAATATATTTGGATTTAAAGCCCTGCGTGCTCTACGTCTGGAAGATCTACGAAT 419
                 *    *   *   ******* **** ******** **** ******** **** *****
O.cumana        CCCTCCTGCTTATATTAAAA---TTTTCCAAGGCCCGCTCCATGGTATTCAAGTTTAAAG 195
O.crenata       TCCTCCTGCTTATATTAAAATTATTTTCCAAAGCCCGCCCCATGAGATTAAAGTTGAAAT 479
                 *******************   ******** ****** *****  *** ***** ***
O.cumana        AGATAAATTAAACAAGTATGGTCGTCCCCTGTTGGGATGTATTATTAAACCTAAA----T 251
O.crenata       AGATAAATTGAACAAGTATGGTCGTCCCCTGTTGGGATGTACTATTAAATAAACCTAAAT 539
                ********* ******************************* *******   *      *
O.cumana        TGGGGTTATCCACTAAAAACTATGGGGGAACAGTTTATGAATGTCTTCGCGGTGGACTTG 311
O.crenata       TAGGATTATCCGC--------------------------------------TAAAAACC 560
                * ** ****** *                                       *  *
O.cumana        ATTGTACCAAAGATGATGAGAACGTAAATTCCCAGCCATTTATGCGCTGGAGAGATCGTT 371
O.crenata       ATGGGAGAACAGTTGATGAGAACGTAAACTCCCAGCCATTTATGCGCTGGAGAAATCGTT 620
                ** * *  * ** *************** ************************ ******
O.cumana        TTTTATTTTGTGCTGAAGCAATTTATAAATCACAGGCTGAAACAGGCGAAATTAAAGGCC 431
O.crenata       TCTTATTTTGTGCCGAAGCAATTTATAAAACACAGGCTGAAACAGGCGAAATCAAAGGTC 680
                * *********** *************** ********************** ***** *
O.cumana        ATTACTTGAATACTACTGCAGGTACATGCGAGGAAATGATCAAAAGAGCTGTCTTTGCTA 491
O.crenata       ATTACTTGAATACTACTGCAGGTACATGTGAGGAAATGATCAAAAGAGCTGTTTTTGCCA 740
                ***************************** ************************ ***** *
O.cumana        GAGAATTGGGAGTTCCTATTATAA--ATGTACGACTACTTAACAGTAGGATTCACTGCAA 549
O.crenata       GAGAATTGGGAGTTCCTATTATAATAATGCACGACTACTTAACAGGAGGATTCACTTCAA 800
                ************************  *** *************** ********** ***
O.cumana        ATACTAGCTTGGCTCATTATTGCCGTAATAATGTCC-----TACTTCTCACCGTGCAATG 604
O.crenata       ATACTAGCTTGGCTCATTATTGCCATAATAATGGCCTGGCCTACTTCTTCACCTGCAATG 860
                ************************ ******** **     *******   * *******
O.cumana        CATGTAGTTATTGATAGACAGAAGAACCATAGTATACATTTCCGTTGTACTAGCTAAAGC 664
O.crenata       CATGCAGTTATTTATAGACAGAAGAACCATGGTATACATTTCC-GTGTACTAGCTAAAGC 919
                **** ******* ***************** ************  ***************
O.cumana        GTTACGTATGTCTGGTGGAGATCATATTCACTCTAGGACCGTAGTAGGTAAACTTGAAGG 724
O.crenata       ATTACGTATGTCTGGTGGAGATCATATTCACTTTGGGACTGCAGTAGGTAAACTTGAAGG 979
                ******************************* * **** * ******************
```

```
O.cumana       AGAAAGAGACATTACTTTGGGCTTTGTTGATTTATTGCGTGATGATTTTATTAA-AAAGA 783
O.crenata      AGAAAGAGACATTACTTTTGACTTTGTTGATTTATTGCGTGATGATTTTATTGAAAAAGA 1039
               ****************** * ****************************** * *****
O.cumana       TCGAAGTCACGGTATTTATTTTACTCAAGATTGGGTTTCTTTACCAGGTGTTTTTACTGT 843
O.crenata      TCGAAGTCGCGATATTTATTTCACCCAAGATTGGGTTTCTCTACCAGGTGTTATTCCTGT 1099
               ******** ** ********* ** *************** *********** ** ****
O.cumana       GGCTTCAGGGGGTATTCACGTTTGGCATATGCCTTCATCT-GACGGAGATCTTTGGGGAT 902
O.crenata      GGCTTCAGGGGGTATTCACGTTTGGCATATGCCTTCCCTGACCGCGAGATCTTTGGGGAT 1159
               ************************************         ***************
O.cumana       GATTCCGTACTACAGTTTGACGGAGGAACTTTAGGACATCCTTAGCGTAATGCACCAGGC 962
O.crenata      GATTCCATACTACAGTTTGGCGGAGGAACTTTAGGACATCCTTGGGGTAATGCACCAGGT 1219
               ****** ************ ********************** * *************
O.cumana       GCTCTGTAGCTAATCGAGTAGCTATAGAAGCATGTGTACAAGCTCGTAATGAAGGATGTA 1022
O.crenata      GC--TGTAGCTAATCGAGTAGCTATATAAGCATATGTACAAGCTCGTAATGAAGGACGTG 1277
               **  ********************** ****** ********************** **

                                                                RbcL-M-R
O.cumana       ATCTTGCTACTGAGGGGAATGCAATTATACGCGAGGCTAGGAAATGGAGCCCTGAACTAG 1082
O.crenata      ATCTTGCTGCTGAGGGTAATTATA-----CGTGAGGCTAGCAAAGGGAGTCCTGAACTAG 1332
               ******** ******* ***   *     ** ******** *** **** **********
                                                                RbcL-M-R

O.cumana       TTGCTGTTTGTGAGGTATGTAAAGATATCAAATTTGAGTTTAAAGCAGTCGATACTTTGG 1142
O.crenata      TTGCTG------------------------------------------------------ 1338
               ******
O.cumana       GTAAGTGTAAGTAAGATAACATTACTCTTCATTCTCTTAATTGAATTTCAATTAAATTCG 1202
O.crenata      ------------------------------------------------------------ 1338

O.cumana       GCTTAATC   1210
O.crenata      --------   1338
```

Table S1. List of oligo's used in this study

Primer's name	**Sequence 5 ⟶ 3'**	**Use in the study**
UCP555-F	GTAGTCATATGCTTGTCTC	Universal internal control
UCP555-R	GGCTGCTGGCACCAGACTTGC	
RbcL-M-F	ATGTCACCACAAACAGAAAC	Molecular marker to identify *O.cumana* and *O.crenata*
RbcL-M-R	CAGCAACTAGTTCAGGCTCC	
ITS100-F	CAACGGATATCTCGGCTCTC	qRT-PCR to quantify broomrape
ITS100-R	TTGCGTTCAAAGACTCGATG	
ACT1-F	ATGGGCCAGAAAGATGCATATGTT	Housekeeping gene used for qRT-PCR expression normalization
ACT1-R	GTGTGATGCCAAATTTTCTCCATGT	