Oikos

OIK-06182

Sneck, M. E., Rudgers, J. A., Young, C. A. and Miller, T. E. X. 2019. Does host outcrossing disrupt compatibility with heritable symbionts? – Oikos doi: 10.1111/oik.06182

Appendix 1

Appendix methods – Molecular techniques to quantify genetic distance between outcrossed parents (P_1 generation)

Tissue sampling, PCR protocol, and processing microsatellite markers

We used a multiplex approach with nine fluorescently labeled simple sequence repeat (SSR) microsatellite markers to confirm the effectiveness of the greenhouse crossing experiment (Table A2, Fig. A3) and to estimate genetic distance between outcrossed P_1 parents (Fig. 1) (Hayden et al. 2008, Saha et al. 2009). Genomic DNA was extracted from ~10 mg of freshly frozen and lyophilized plant tissue using MagAttract 96 DNA plant core Kit (QIAGEN Inc., Valencia, CA) and analyzed following Takach et al. (2012) to determine the endophyte genotype. PCR was performed in a total volume of 10 µl containing 20 ng of template DNA, 1.0 unit of Promega GoTaq™ DNA Polymerase (Promega Corp., Madison, WI), 2 mM of dNTPs, 2 µl of 5X Colorless GoTaq™ Buffer, 10 µM of the reverse and M13 dye primer, and 5 µM of the forward M13 fluorescently tagged primer. The PCR cycling parameters were: an initial denaturation of 3 min at 95°C, then 6 cycles of 94°C for 45 s, 68°C for 5 min, and 72°C for 60 s followed by 8 cycles of 94°C for 45 s, 58°C for 2 min, and 72°C for 30 s. Then, 25 cycles at 94°C for 45 s, 50°C for 2 min, and 72°C for 30 s followed by a final extension of 72°C for 7 min. To confirm amplification, we ran each PCR product on a 1.5% agarose gel visualized with ethidium bromide and a UV light. Then, 1 µl of amplified and fluorescently labeled PCR product per SSR marker (4 unique markers pooled per well) was suspended in Hi-Di Formamide with 0.5 µl LIZ 500 ladder (Applied Biosystems, Farmingham, MA) and then separated on an Applied Biosystems 3730 capillary sequencer. Raw data were analyzed using GENEMAPPER 4.0 and Peak Scanner 2 (Applied Biosystems). Kentucky 31 (*Festuca arundinacea*) of known endophyte status were used as positive and negative controls.

Estimating genetic distance between P_1 parents

Both P_1 parent species (*Elymus virginicus* and *E. canadensis*) are allopolyploids, with genomes from *Pseudoroegneria* and *Critesion* species (Dewey 1983, Sun et al. 1997). However, estimating

genotypes from allopolyploid organisms is challenging due to potential amplification of multiple genomes from the two progenitor species, which makes assigning heterozygotes impossible without knowing the origin of each amplified sequence (Dufresne et al. 2014). Therefore, to avoid taxonomic issues common to polyploids, analyses were performed using data formats that assume uniform ploidy across individuals.

First, to measure linkage disequilibrium among loci, we calculated a modified index of association robust to small sample sizes ($\bar{r}_d$) (Brown et al. 1980, Agapow and Burt 2001) between the amplified microsatellite loci in the R package *poppr* (Kamvar et al. 2014). Higher measures of $\bar{r}_d$ indicate greater linkage between loci. We uncovered a pair of significantly linked loci ($\bar{r}_d$ = 0.0418, p = 0.001). Removing one of two loci from the dataset significantly reduced linkage disequilibrium ($\bar{r}_d$ = 0.0299, p = 0.061). In total, we identified 29 alleles across the eight remaining loci (mean = 3.75 alleles per locus). We then generated a saturating genotype accumulation curve, which quantifies the amount of power within the data to discriminate between unique individuals given a random sub-sample of n loci (command locus_table and genotype_curve in *poppr* (Kamvar et al. 2014) (Fig. A1). Then, we calculated a pairwise, individual-by-individual (N × N) genetic distance matrix using the binary data set in GenAlEx ver. 6.502 (Peakall and Smouse 2006), which tallies the total number of differences between two genetic profiles (Huff et al. 1993). When multiple individuals per maternal line were genotyped, genetic distance was averaged over all estimates. Using this distance matrix, we conducted a distance-based cluster analysis (PCoA, GenAlEx ver. 6.502), which revealed that the genetic markers used in this study provided adequate resolution to genetically distinguish between *Elymus virginicus* individuals as well as between *E. virginicus* and *E. canadensis* host species (Fig. A2).

Table A1**.** Sample size of experimental crosses between maternal and paternal plant hosts collected from across populations. All sampled populations were from state parks. For exact collection locations, see our previous publication (Sneck et al. 2017). Number of crosses represent experimental replicates from each endophyte transmission experiment in the greenhouse (F_1) and common garden (F_2), respectively.

Cross type	Maternal population	Paternal population	F_1 sample size (greenhouse crosses)	F_2 sample size (common garden plants)
Within-population	Davis Mountains, TX	South Llano, TX	1	0
	McKinney Falls, TX	Bona Dea, AK	1	0
	McKinney Falls, TX	Keystone, OK	1	0
	McKinney Falls, TX	South Llano, TX	1	3
	McKinney Falls, TX	Mother Neff, TX	1	0

	McKinney Falls, TX	Palmetto, TX	14	8
	McKinney Falls, TX	Fall River, KS	1	0
	McKinney Falls, TX	San Angelo, TX	1	0
	McKinney Falls, TX	Lake Texoma, OK	1	0
	Keystone, OK	McKinney Falls, TX	1	0
	South Llano, TX	McKinney Falls, TX	4	2
	South Llano, TX	Keystone, OK	1	1
	South Llano, TX	Palmetto, TX	1	0
	South Llano, TX	San Angelo, TX	4	1
	Lake Murray, OK	McKinney Falls, TX	1	0
	Lake Murray, OK	Mother Neff, TX	1	0
	Lake Murray, OK	San Angelo, TX	1	0
	Lake Murray, OK	Toronto, KS	1	0
	Mother Neff, TX	McKinney Falls, TX	3	12
	Mother Neff, TX	Palmetto, TX	2	10
	Mother Neff, TX	Lake Texoma, OK	1	1
	Palmetto, TX	McKinney Falls, TX	17	4
	Palmetto, TX	Keystone, OK	2	0
	Palmetto, TX	South Llano, TX	1	0
	Palmetto, TX	San Angelo, TX	3	2
	San Angelo, TX	Bona Dea, AK	1	0
	San Angelo, TX	McKinney Falls, TX	5	1
	San Angelo, TX	Keystone, OK	2	1
	San Angelo, TX	Mother Neff, TX	1	1
	San Angelo, TX	Palmetto, TX	7	1
	San Angelo, TX	Toronto, KS	1	0
	Lake Texoma, OK	Lake Thunderbird, OK	1	0
	Toronto, KS	South Llano, TX	1	1
Between-population	McKinney Falls, TX	McKinney Falls, TX	9	1
	Mother Neff, TX	Mother Neff, TX	1	1
	Palmetto, TX	Palmetto, TX	6	3
	San Angelo, TX	San Angelo, TX	4	7
Between-species	McKinney Falls, TX	Davis Mountains, TX	1	0
	McKinney Falls, TX	Keystone, OK	1	0
	McKinney Falls, TX	Slide Rock, AZ	1	0
	McKinney Falls, TX	South Llano, TX	7	3
	McKinney Falls, TX	Lake Thunderbird, OK	1	0
	Keystone, OK	South Llano, TX	1	0
	South Llano, TX	Davis Mountains, TX	3	3
	South Llano, TX	South Llano, TX	3	3
	Lake Murray, OK	Davis Mountains, TX	2	0
	Lake Murray, OK	South Llano, TX	1	0
	Mother Neff, TX	Davis Mountains, TX	1	1
	Mother Neff, TX	South Llano, TX	3	1
	Mother Neff, TX	Lake Thunderbird, OK	1	1
	Palmetto, TX	Davis Mountains, TX	4	1

Palmetto, TX	South Llano, TX	8	0
Palmetto, TX	Slide Rock, AZ	1	0
San Angelo, TX	Davis Mountains, TX	3	9
San Angelo, TX	McKinney Falls, TX	1	0
San Angelo, TX	Keystone, OK	1	0
San Angelo, TX	South Llano, TX	4	1
Lake Texoma, OK	Davis Mountains, TX	1	1
Toronto, KS	Davis Mountains, TX	1	0

Table A2. Microsatellite markers used to estimate genetic distance between outcrossed parents modified from Saha et al. 2009. Primer 61 was dropped from final analyses due to linkage with primer 59. An annealing temperature of 58°C was used for all primers. Nei's gene diversity was calculated to determine the expected heterozygosity per locus.

Name	Forward primer	Reverse primer	Expected size	Repeats	Allele per locus	Nei's gene diversity
19	TGGATTTGCAATTAGCCTCA	GCTCGTGTATGGCCTTCAAT	176-530	ta	2	0.39
22	ATGATGTCCGAGGAGGAGAA	CATCATGATCCAGTGCCTTG	184-266	agg	3	0.63
32	ACGGTCTGTACCGTGGATGT	GCTGTAGACTCAGCCGAACC	288-330	ctg	4	0.57
50	GATGGACGAAGGCTTCTTTG	AGCCGAACCTGAACTCAGAC	177-287	cag	4	0.72
59	TTTGCACTCTCGGACCTAGC	CGGTACACCTTCTGCACCTT	288-290	ga	4	0.73
61	GTCGCCGGAGAAGAGAAGAG	AACGCTAGCCGTGATGACTT	127-142	ag	4	0.59
78	TCCTAAGCAGAGCTCGATCC	GAGGTTGGCGAACTTCCTC	164-216	ga	4	0.72
113	CAATGGTGGTGCAAGAAATG	AGAGAGCAAGGAGGAAGAAACC	153-248	ct	5	0.71
142	ACTTGCCGGAGAAGAAGCTC	ATACAGGAGGAGGAGGAGCAG	185-304	aga	4	0.73

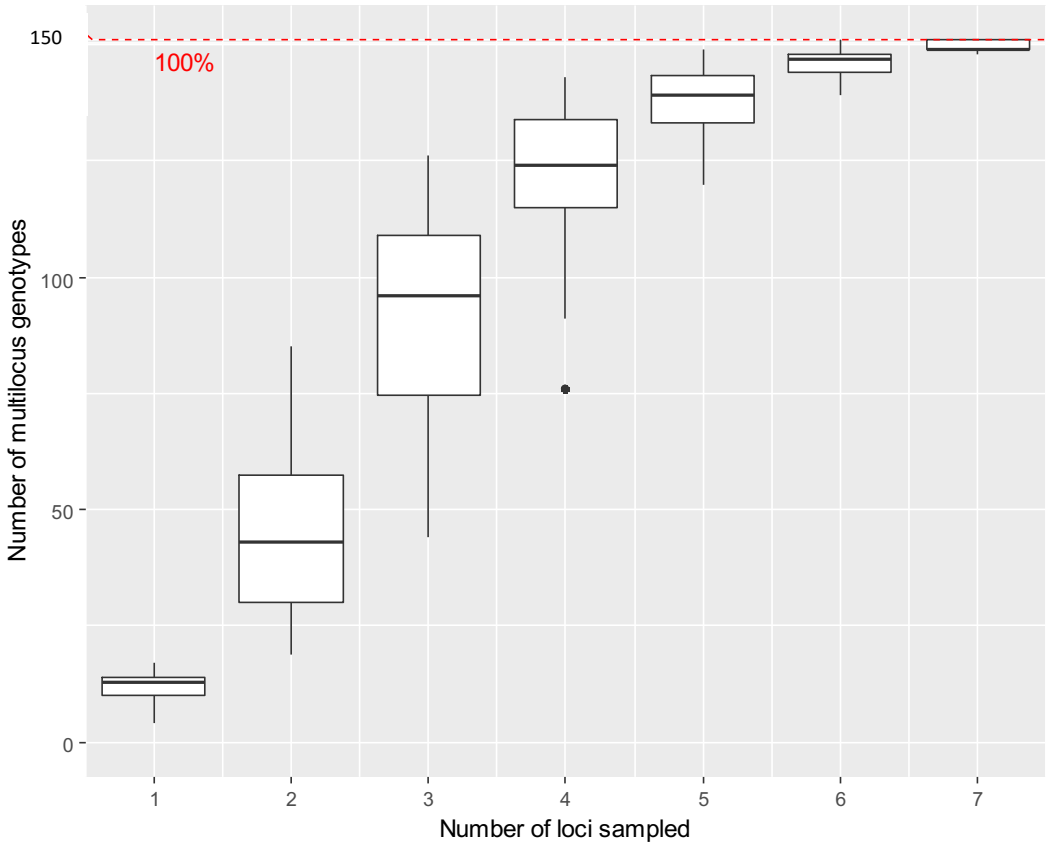

Figure A1. Genotype accumulation curve of the eight microsatellite loci used to quantify genetic distance between outcrossed parents. Boxplots show the number of unique genotypes estimated at each number of sampled loci. The saturating curve indicates that the number of loci used were sufficient to discriminate between individuals. The red dotted-line indicates the number of multilocus genotypes across all 8 loci. Samples were bootstrapped 1000 times to create this distribution.

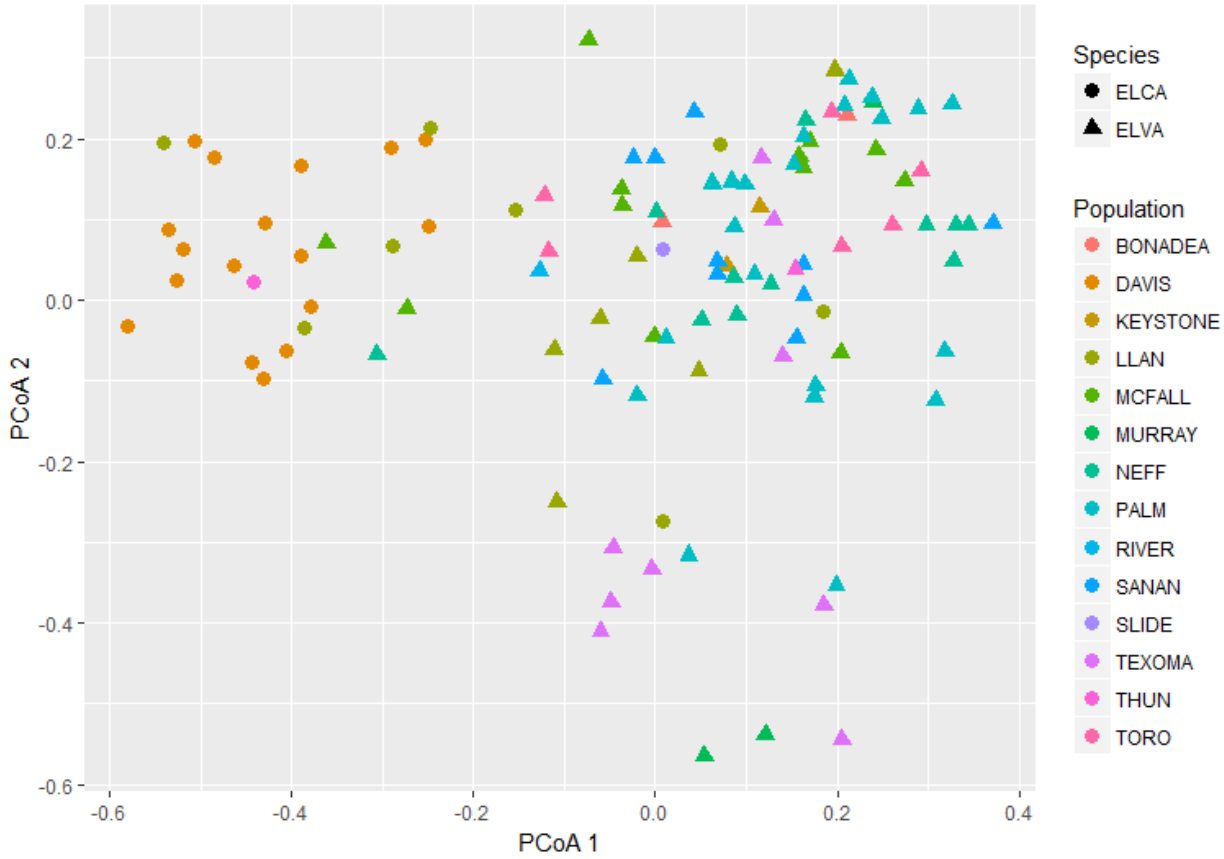

Figure A2. Principle coordinate analysis (PCoA) representing the genetic distance between P_1 individuals within a host species and between host species (ELCA (filled circles) = *Elymus canadensis*; ELVA (filled triangles) = *E. virginicus*). This PCoA displays the relationship between collected individuals based upon multivariate genetic distances estimated from eight microsatellite loci, wherein spatial distance between points (spread between principle coordinates 1 and 2) is analogous to estimated genetic distance. Point colors correspond to their collection location listed in Table A1.

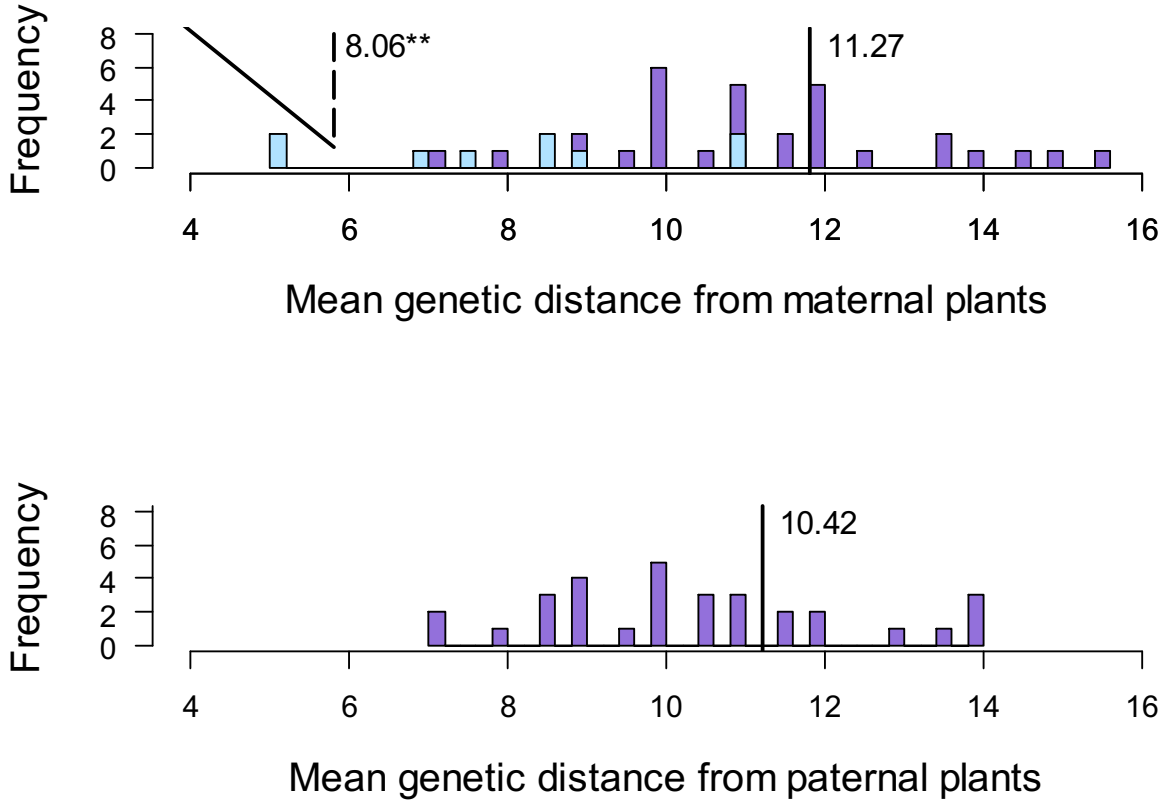

Figure A3. Totest whether our experimental crosses successfully altered the genetic background of outcrossed offspring, outcrossed (n = 33) and self-pollinated (n = 7) individuals were genotyped using eight microsatellite markers. We found that outcrossed offspring were similarly genetically distant from both parents (purple) . Also, compared to self-fertilized offspring (blue), outcrossed offspring were significantly more distant from the maternal plant. Solid (outcrossed) and dashed (selfed) vertical bars and adjacent numbers indicate the mean genetic distance of each group.

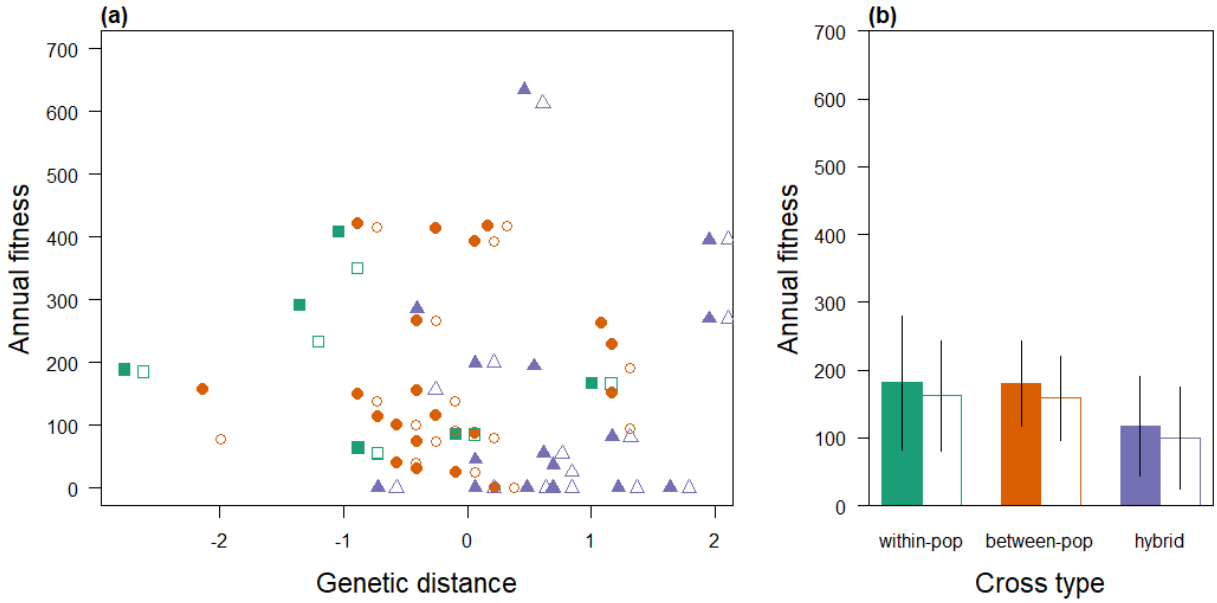

Figure A4. Fitness estimation excluding three high-fitness inter-population outliers for hosts (expected number of seeds per seed per year; filled shapes / bars) and symbionts (expected number of E+ seeds per seed from an E+ plant; open shapes / bars), shown with respect to neutral genetic distance (a) and qualitative cross type (b). Layout and content as in Fig. 6.

References

Agapow, P. M. and Burt, A. 2001. Indices of multilocus linkage disequilibrium. – Mol. Ecol. Notes 1: 101–102.

Brown, A. H. D. et al. 1980. Multilocus structure of natural populations of *Hordeum spontaneum.* – Genetics 96: 523–536.

Dewey, D. R. 1983. Historical and current taxonomic perspectives of *Agropyron*, *Elymus*, and related genera. – Crop Sci. 23: 637–642.

Dufresne, F. et al. 2014. Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. – Mol. Ecol. 23: 40–69.

Hayden, M. J. et al. 2008. Multiplex-ready PCR: A new method for multiplexed SSR and SNP genotyping. – BMC Genomics 9: 80.

Huff, D. R. et al. 1993. RAPD variation within and among natural populations of outcrossing buffalograss (*Buchloe dactyloides* L.). – Theor. Appl. Genet. 86: 927–934.

Kamvar, Z. N. et al. 2014. *Poppr*: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. – PeerJ 2: e281.

Peakall, R. and Smouse, P. E. 2006. Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. – Mol. Ecol. Notes 6: 288–295.

Saha, M. C. et al. 2009. Genetic variation within and among wildrye and populations from the southern Great Plains. – Crop Sci. 49: 913.

Sneck, M. E. et al. 2017. Variation in the prevalence and transmission of heritable symbionts across host populations in heterogeneous environments. – Microb. Ecol. : 1–14.

Sun, G.-L. et al. 1997. Analysis of tetraploid *Elymus* species using wheat microsatellite markers and RAPD markers. – Genome 40: 806–814.

Takach, J. E. et al. 2012. Genotypic and chemotypic diversity of *Neotyphodium* endophytes in tall fescue from Greece. – Appl. Environ. Microbiol. 78: 5501–5510.