Stratified Dispersal and Increasing Genetic Variation during the Invasion of Central Europe by the Western Corn Rootworm, *Diabrotica virgifera virgifera*

M. Ciosi [1], N. J. Miller [2], S. Toepfer[3], A. Estoup [4] and T. Guillemaud [1]

Running head: *D. v. virgifera* expansion in Central Europe

Author affiliations

[1] INRA, UMR 1301 IBSV (INRA / Université de Nice Sophia Antipolis / CNRS). 400 Route des Chappes. BP 167 - 06903 Sophia Antipolis cedex. France. E-Mail: M. Ciosi: ciosi@sophia.inra.fr and T. Guillemaud: guillem@sophia.inra.fr

[2] USDA-ARS, Corn Insects and Crop Genetics Research Unit, Genetics Laboratory, Iowa State University, Ames, Iowa 50011, USA. E-mail: nicholas.miller@ars.usda.gov

[3] CABI Europe – Switzerland, c/o Plant Protection Directorate, Rarosi ut 110, H -6800 Hodmezovasarhely, Hungary. E-Mail: stoepfer@gmx.net

[4] INRA, UMR CBGP (INRA / IRD / Cirad / Montpellier SupAgro), Campus international de Baillarguet, CS 30016, F-34988 Montferrier-sur-Lez cedex, France. Email: bourguet@supagro.inra.fr

Type of paper: Original Research

Key words: Dispersal, invasive species, range expansion, microsatellites, spatial genetic pattern, long-distance dispersal

Corresponding author: Thomas Guillemaud: INRA, UMR 1301 IBSV (INRA / Université de Nice Sophia Antipolis / CNRS). 400 Route des Chappes. BP 167 - 06903 Sophia Antipolis cedex. France. E-mail: guillem@sophia.inra.fr. Tel: +33 492 386 481 ; Fax: +33 492 386 401

Acknowledgments: For assistance with field sample acquisition in Central and South-Eastern Europe, we thank Drazen Dzoic, Marianna Scuzc, Ibolya Hatala-Zseller, Szeoeke Kalman, Tamas Hegyi, Laszle Vasas, Szilvia Kleineizel, Eva Komlosi Szabone, Lorinc Andrassy, Maria Veszelka Szantone, Marija Ivezic and Mirjana Brzmez. We also thank Lorenzo Furlan, Sharon Cheek, Joe Ostoja-Starzewski, Peter Baufeld, Dieter Eberhard and Ullrich Benker for their assistance with sample acquisition in other parts of Europe and Aurélie Blin, Pascal Chavigny and Alexandra Auguste-Maros for technical assistance. This work was funded by the *Biodiversité* program of the French *Agence Nationale de la Recherche* (project #ANR-06-BDIV-008-01).

Abstract

Invasive species provide opportunities for investigating evolutionary aspects of colonization processes, including initial foundations of populations and geographic expansion. Using microsatellite markers and historical information, we characterized the genetic patterns of the invasion of the western corn rootworm (WCR), a pest of corn crops, in its largest area of expansion in Europe: Central and South-Eastern (CSE) Europe. We found that the invaded area probably corresponds to a single expanding population resulting from a single introduction of WCR and that gene flow is geographically limited within the population. In contrast to what is expected in classical colonization processes, an increase in genetic variation was observed from the center to the edge of the outbreak. Control measures against WCR at the center of the outbreak may have decreased effective population size in this area which could explain this observed pattern of genetic variation. We also found that small remote outbreaks in southern Germany and north-eastern Italy most likely originated from long-distance dispersal events from CSE Europe. We conclude that the large European outbreak is expanding by stratified dispersal, involving both continuous diffusion and discontinuous long-distance dispersal. This latter mode of dispersal may accelerate the expansion of WCR in Europe in the future.

INTRODUCTION

Biological invasions may cause ecological (reviewed in Olden et al. 2004), health (e.g. Ruiz et al. 2000) and economic problems (Pimentel et al. 2001). In addition to these practical considerations, biological invasions provide opportunities to investigate various aspects of the colonization process, including initial foundations of populations and patterns of geographic expansion into new areas. At the time of introduction, the number of founders affects the chances of a population becoming established, through demographic (e.g. Allee effect, Drake and Lodge 2006) or genetic effects (e.g. inbreeding depression, Elam et al. 2007). After establishment, dispersal is crucial, because it determines the rate of spatial expansion and influences the genetic structure of the expanding populations and, hence, their adaptation to new environments (Sax et al. 2005). Dispersal may be required to bring novel combinations into marginal habitats which can present an adaptive challenge. However, gene flow may limit expansion, because genetic homogenization may limit adaptation at the margins of the spatial distribution (Kirkpatrick and Barton 1997).

Molecular genetic markers can be used to determine population genetic structure, making it possible to reconstruct the introduction and expansion patterns of invasive populations, thereby providing insight into aspects of introduction and spatial expansion (e.g. Estoup et al. 2004 for the cane toad; Williams et al. 2007 for the Brazilian peppertree). Such knowledge may have major implications for the management of invasive pests. For instance, the eradication of newly founded populations at the colonization front of an expanding outbreak can slow the expansion of invasive species, as shown for some invasive noxious weeds (e.g. Moody and Mack 1988) and insects (e.g. Johnson et al. 2006).

The univoltine western corn rootworm (WCR), *Diabrotica virgifera virgifera* LeConte (Coleoptera: Chrysomelidae) is native to North America and is one of the most destructive

76 pests of cultivated corn, *Zea mays* L. This insect rapidly expanded its range in North America
77 during the last century (Gray et al. 2009) and was recently introduced into Europe, where it
78 was first observed near Belgrade, Serbia, in 1992. An international network has since been set
79 up and provides an annually updated, detailed description of the distribution and expansion of
80 WCR in Europe (Kiss et al. 2005). After its introduction, WCR rapidly spread throughout
81 Central and South-Eastern (CSE) Europe, extending its range by up to 100 km per year
82 (Baufeld and Enzian 2001). The continuously expanding CSE European outbreak now
83 extends from Austria to Ukraine and from southern Poland to northern Bulgaria. A number of
84 isolated outbreaks have been detected almost every year since 1998, in various countries
85 including Italy, France, Switzerland, Belgium, the United Kingdom, the Netherlands and
86 Germany (Anonymous 2007; Edwards and Kiss 2007). Recent population genetic studies
87 have provided evidence for repeated transatlantic introductions of this insect, accounting for
88 the initiation of at least some of these European outbreaks, including that in CSE Europe
89 (Ciosi et al. 2008; Miller et al. 2005). However, the source populations have yet to be
90 identified for some outbreaks.

91 Most WCR movements are probably local (Spencer et al. 2005). However, several studies
92 have reported the dispersal of WCR over large distances, partly through active long-distance
93 flights, but largely through passive transport by wind, thunderstorms and human-mediated
94 transportation (Coats et al. 1986; Grant and Seevers 1989; Gray et al. 2009; Spencer et al.
95 2005). The history of WCR range expansion in the US and Europe is also consistent with the
96 long-distance mobility of WCR (Gray et al. 2009). Thus, both local diffusion and long-
97 distance dispersal have probably played a role in the expansion of WCR in CSE Europe. The
98 isolated outbreak of Friuli (north-eastern Italy) first detected in 2003 is currently the only
99 outbreak for which long-distance dispersal of WCR from the large CSE European expanding
100 outbreak has been demonstrated (Ciosi et al. 2008; Miller et al. 2005). However, more recent

outbreaks in southern Germany in 2007 — near Passau in Bavaria in south-eastern Germany and near Frickingen in Baden-Württemberg in south-western Germany — have yet to be analyzed and may also result from the establishment of remote colonies originating from CSE Europe.

We examined here the spatial genetic structure and characterized the expansion process of WCR within its largest area of expansion in Europe: the CSE European outbreak. We combined a specific transect sampling scheme along the WCR expansion with analyses of microsatellite data and historical information concerning WCR expansion. We addressed the following specific questions: (i) Does the genetic evidence suggest single or multiple introductions of WCR in CSE Europe? (ii) Does it support successive founder events during expansion? (iii) Does it indicate stratified dispersal, involving both local movements and long-distance dispersal? (iv) Do the recent outbreaks in southern Germany constitute examples of long-distance dispersal events?

MATERIALS AND METHODS

Sample collection

We collected WCR samples in CSE Europe, at 19 sites in five countries, in 2003 (see details in Figure 1 and Table S1). Samples were collected along two almost linear transects running from the area in which WCR was first detected (near Belgrade Airport in Serbia) to the expansion front in 2003. These transects are referred to hereafter as the "western" and "eastern" transects. The western transect consisted of nine samples taken along a transect running from Belgrade to the eastern end of Austria. The eastern transect consisted of 10 samples taken along a transect running from Belgrade to eastern Hungary. The sample from

124 Stara-Pazova in Serbia is common to both transects. Careful monitoring of the CSE European
125 outbreak made it possible to assign a precise year of first observation to each sample (Kiss et
126 al. 2005; I. Hatala-Zsellér, personal communication, Csongrád County Plant and Soil
127 Protection Service, Hódmezövásárhely, Hungary; I. Sivcev, personal communication,
128 Institute for Plant Protection and Environment, Zemun, Serbia; I. Grozea, personal
129 communication, Banat University, Timisoara, Romania). For each transect, the sampling sites
130 were carefully chosen so as to obtain an even distribution of the years of first observation
131 between 1992 (first observation of WCR in Europe) and 2002 (last year before sampling).
132 One sample from the western part of the invaded area infested in 1999 (the Turanovac
133 sample) was added to the two transects. At each sampling site, we sampled 30 to 50 adult
134 beetles from maize fields (typically, we could find from less than one to ten adult beetles per
135 maize plant in CSE Europe). A sampling site was defined as a collection area of less than 100
136 m x 100 m. In most cases, sampling was carried out with an aspirator device or a funnel with
137 an attached gauze bag, into which beetles where shaken from the plants. Sex pheromone-
138 baited transparent sticky traps (PAL, Csalomon® family of traps, Hungarian Academy of
139 Sciences, Budapest, Hungary) were used in areas infested in 2001 and 2002 because WCR
140 population densities were low in these areas.

141 Samples from two outbreaks detected in 2007 in southern Germany near Passau in Bavaria
142 and near Frickingen in Baden-Württemberg were included (Table S1 and Figure 1), to test the
143 hypothesis of recent long-distance dispersal events from the CSE European area. The Passau
144 and Frickingen samples were collected with the sex pheromone-baited transparent sticky traps
145 (PAL) used for WCR monitoring in Europe (Kiss et al. 2005). For identification of the source
146 populations of the CSE European samples and the two German samples, we also included a
147 substantial number of European and American samples in our analyses. Samples of WCR
148 from most other European outbreaks were collected from six sites in three countries (see

details in Table S1 and Figure 1) between 2003 and 2005. The European samples from Friuli in north-eastern Italy, Piedmont in north-western Italy, Paris-1, Paris-2, Alsace in eastern France and the UK studied here were those previously analyzed by Ciosi *et al.* (2008). We thus sampled every outbreak detected in Western Europe up to 2007, with the exception of four locations at which beetles were no longer detected after 2003 — Venice, in north-eastern Italy (outbreak detected in 1998), Belgium and the Netherlands (outbreaks detected in 2003) — or at which too few beetles were observed: the Lahr airport site in south-western Germany (outbreak detected in 2007). We added samples from five North American locations (in Mexico, Arizona, Texas, Illinois, and Pennsylvania), previously analyzed by Kim and Sappington (2005a) and Ciosi *et al.* (2008). These five samples were representative of the principal structured population entities of WCR in its native continent (for details see Ciosi et al. 2008).

DNA extraction and microsatellite analyses

All WCR samples were stored in absolute ethanol until DNA extraction. Template material for the polymerase chain reaction (PCR) amplification of microsatellites was obtained with three different protocols. (i) For the Deutsch-Jahrndorf sample, DNA was extracted from the thorax and abdomen of each specimen with the AquaPure genomic DNA kit (Bio-Rad, Hercules, CA, USA). (ii) For the Frickingen and Passau samples, we used the thorax or half the body, cut lengthwise, with the DNeasy tissue kit (Qiagen, Hilden, Germany). (iii) For all other insects, the "salting out" rapid extraction protocol of Sunnucks and Hales (1996) was used to extract DNA from the head of each individual. Individuals were washed at least three times in 0.065% NaCl before each protocol, to remove ethanol from the tissues. For the first two extraction protocols, the body part used was first dissected out, placed in a 1.5 ml microcentrifuge tube, frozen in liquid nitrogen and pulverized with a micropestle. DNA was

extracted from the pulverized material. Six dinucleotide (DVV-D2, DVV-D4, DVV-D11, DVV-D5, DVV-D8, DVV-D9) and two trinucleotide (DVV-T2 and DVV-ET1) microsatellite loci (Kim and Sappington 2005b; Miller et al. 2005) were amplified in two separate multiplex PCR, and were analyzed as described by Miller *et al.* (2007).

Genetic variation within WCR samples

Genetic variation within samples was assessed by determining the mean number of alleles per locus (*A*), the mean expected heterozygosity (*H*) (Nei 1987) and the mean variance of absolute allelic size (*V*). The variables *A* and *H* were calculated with GENECLASS 2 ver. 2.0.g (Piry et al. 2004) and *V* was calculated with DIYABC v.0.7.1 (Cornuet et al. 2008). We also calculated the coefficient of inbreeding F_{IS} with GENEPOP ON THE WEB (Raymond and Rousset 1995b). For comparisons of *A* values between population samples, we estimated allelic richness (*AR*) on the basis of minimum sample size, using the rarefaction method (Petit et al. 1998) implemented in FSTAT 2.9.3 (Goudet 2001). The significance of differences in *AR*, *H* and *V* between samples was assessed with the nonparametric Friedman and Wilcoxon sign rank tests (with locus as a repetition unit). Deviation from Hardy-Weinberg equilibrium was assessed with the probability test approach, using GENEPOP ON THE WEB.

Genetic variation between WCR samples

Exact tests of population genic differentiation (Raymond and Rousset 1995a) were carried out with GENEPOP ON THE WEB. As pairwise differentiation tests involve non orthogonal and multiple comparisons, we corrected significance levels with Benjamini & Hochberg's false discovery rate procedure (1995) when necessary. GENEPOP ON THE WEB was also used to calculate Weir & Cockerham's estimator of pairwise F_{ST} (1984), a statistic summarizing genetic differentiation between pairs of populations.

196 ***Clustering analysis of WCR population genetic structure in CSE Europe***

197 Various Bayesian methods of clustering of individuals based on their microsatellite
198 genotypes are currently available to infer population genetic structure in the invaded area of
199 CSE Europe. Because there is little agreement about which method may be most appropriate,
200 we used several of these methods to verify the consistency of the results obtained here.
201 (i) The first approach is implemented in STRUCTURE software (Pritchard et al. 2000). The
202 microsatellite data were converted from GENEPOP to STRUCTURE format with CREATE
203 software v.1.1 (Coombs et al. 2008). We checked for repeatability of the results and
204 convergence of the Markov chain Monte Carlo (MCMC), by carrying out a series of ten
205 replicate runs for each prior value of the number (K) of clusters, set between 1 and 19 (the
206 actual number of samples). Each run consisted of a burn-in of 2×10^4 iterations, followed by
207 10^5 iterations. Two types of inference were performed, using the admixture model of ancestry
208 together with the correlated allele frequencies model (Falush et al. 2003), with and without
209 the use of sampling location as prior information (Hubisz et al. 2009). Default values were
210 maintained for all other parameters. K was estimated at the value with the highest likelihood
211 for the data $P(X|K)$ and with the ΔK statistics of Evanno et al. (2005), which is based on the
212 rate of change in the log-likelihood between successive K values. The results obtained in each
213 analysis were confirmed by performing longer runs with K sets between 1 and 5, with ten
214 replicate runs at each K, 2×10^5 burn-in iterations, and 10^6 iterations. The results of
215 STRUCTURE analyses were then visualized in DISTRUCT v.1.0 (Rosenberg 2004).
216 (ii) In a second approach, we included the spatial coordinates of individuals as prior
217 information using GENELAND v.3.1.4 (Guillot et al. 2005a; Guillot et al. 2005b). We
218 performed 10 replicated analyses to check for convergence, allowing K to vary from 1 to 19
219 clusters and using the following parameters: 10^6 MCMC iterations, maximum rate of Poisson
220 process fixed at 50, maximum number of nuclei in the Poisson-Voronoi tessellation fixed at

450, and an uncertainty associated with the spatial coordinates of 3 km. We used the Dirichlet model of allelic frequencies, as this model has been shown to perform better than the alternative model (Guillot et al. 2005a). We inferred the number of clusters (K) from the modal value of K for the 10 runs.

(iii) Finally, we used an approach implemented in BAPS 5.2 software (Corander et al. 2004; Corander et al. 2003). We estimated K in four different ways, clustering individuals or groups of individuals (samples), and with and without the use of their spatial coordinates as prior information. In each case, we conducted a series of five replicate runs with the *a priori* upper limit for the number of clusters set at 19 (the actual number of samples) for each run.

Geographic and temporal analyses of genetic variation

Each sample of WCR from CSE Europe can be characterized in terms of its *x*-*y* spatial coordinates, its angular coordinate (see below) and the year of first observation of WCR at the sampling site. As the expansion of WCR in CSE Europe is largely radial, the year of first observation provides an indication of the relative location of the sample in the colonization process. The Madaras sample was not included in the temporal analysis, because we considered its year of first observation to be too imprecise. The spatial and temporal features shaping the population genetic structure within the CSE European zone of expansion were investigated with the following procedures:

(a) Genetic isolation by geographic distance was investigated using Rousset's (1997) regression-based framework, by analyzing the correlation between $F_{ST}/(1- F_{ST})$ values and the natural logarithm of geographic distance between samples.

(b) Genetic isolation by temporal distance was investigated by analyzing the correlation between F_{ST} and the difference in year of first observation ($\Delta_1{}^{st}{}_{obs}$) between samples. Note that in each transect, $\Delta_1{}^{st}{}_{obs}$ and the geographic distance between samples were strongly

245 correlated (Spearman's $r = 0.97$ and 0.86, and Mantel tests $p = 1x10^{-5}$ and $4x10^{-5}$ for the
246 western and the eastern transect, respectively).
247 Genetic isolation by geographic and temporal distance were investigated, considering each
248 transect separately.
249 (c) Spatial expansions into new territories may lead to the spread of rare alleles over large
250 geographical areas where they can reach high frequencies (Edmonds et al. 2004; Klopfstein et
251 al. 2006). This phenomenon, called 'gene surfing' (Klopfstein et al. 2006), is caused by
252 genetic drift that prevails over other evolutionary forces at the front of expansion. Empirical
253 (Hallatschek et al. 2007) and simulation (Excoffier and Ray 2008) studies on spatial genetic
254 patterns for a single locus, have recently demonstrated that gene surfing can cause the
255 emergence of genetically homogeneous sectors and thus of a between-sector genetic
256 differentiation. We addressed this issue in the CSE area of WCR expansion, by investigating
257 the genetic isolation by angular distance expected when considering several loci. We
258 attributed to each sample an angle between the line of latitude passing through the origin of
259 the outbreak (Belgrade, Serbia) and the line connecting Belgrade to each sample (Table S1).
260 The pattern of genetic isolation by angular distance was investigated by analyzing the
261 correlation between F_{ST} and differences in angle measured for all pairs of samples. The
262 Belgrade-Airport, Surcin and Stara-Pazova samples were all located geographically close to
263 Belgrade, and their angular coordinates were therefore uncertain. We thus removed these
264 three samples from the genetic isolation by angular distance analysis.
265 Correlations between distance matrices were analyzed with the Mantel test implemented in
266 GENEPOP ON THE WEB.
267 (d) Correlations between genetic variation within samples (*AR*, *H* and *V*) and the year of
268 first observation or geographic distance to the center of the outbreak were assessed by
269 Spearman's rank correlation tests performed with the *spearman.test* function available in the

pspearman library of R (R Development CoreTeam 2008), with the aim of detecting possible decreases in genetic variation during the expansion process.

Source populations of the Frickingen, Passau and Friuli outbreaks

We calculated the mean multilocus individual assignment likelihood of each outbreak sample *i*, — Frickingen and Passau from southern Germany, and Friuli from north-eastern Italy — to each sample of 31 possible source populations *s,* as shown in Table 1 (see Pascual et al. 2007; and Ciosi et al. 2008) with GENECLASS 2 ver. 2.0.g (Piry et al. 2004). For each outbreak considered, the most probable source population was identified as that with both the highest $L_{i\to s}$ value and the lowest F_{ST}-value with the outbreak considered. No ad hoc statistical test has yet been described for the comparison of mean individual assignment likelihoods or F_{ST}. Moreover, non parametric tests, such as the Friedman analysis of variance by rank or pairwise Wilcoxon signed rank test, using locus as a repetition unit, are not sufficiently powerful for such comparisons, due to the limited number of loci and the large inter-locus variance of the statistics used.

We then documented the effect of introduction events on genetic variation within populations, by comparing the Frickingen, Passau and Friuli outbreaks with their identified source populations.

Detection of multiple introductions

The following procedures were used to detect multiple introductions in the CSE European outbreak. We used three procedures based on different assumptions to compensate for a potentially low statistical power and then to avoid false negatives. The first method (a) does not include statistical tests and is sample-centered, the second one (b) includes statistical tests

292 and is individual-centered and the third one (c) is focused on cotemporary gene flow and
293 detection of first generation immigrants.

294 (a) The most probable source population of each CSE European sample was identified by
295 the procedure described above for the Frickingen, Passau and Friuli outbreaks. In this
296 analysis, potential sources were populations with a first observation year ≤ 2003 (i.e. the year
297 of collection of CSE European samples). The Friuli outbreak was excluded as a potential
298 source population from the analysis because it is known to have originated in CSE Europe
299 (Ciosi et al. 2008; Miller et al. 2005), and thus most CSE European samples are expected to
300 wrongly point to Friuli as its most probable source. Most of the non CSE European outbreaks
301 have been observed for the first time after the occurrence of WCR at CSE European sample
302 locations and could thus not be the source of primary introductions in CSE Europe. In this
303 assignment analysis potential sources have thus been considered in two ways: i) North
304 American populations could be the source of multiple introductions in CSE Europe (either
305 primary or secondary) and ii) European outbreaks could be the source of secondary
306 introductions in CSE European locations where WCR had already been observed.

307 (b) Ciosi *et al.* (2008) suggested that the northern USA (represented by the Illinois and
308 Pennsylvania samples) is the sole source population for the CSE European outbreak. We
309 tested this hypothesis for the large number of individuals sampled at various sites within the
310 CSE European area, by calculating the probability of excluding Illinois and Pennsylvania as
311 source populations for each CSE European individual, using GENECLASS 2 ver. 2.0.g.
312 Probabilities were calculated by using the assignment likelihood calculation of Rannala &
313 Mountain (1997), with resampling based on the simulation algorithm of Paetkau *et al.* (2004).
314 Ten thousand individuals were simulated per population and a threshold of 0.05 was used to
315 exclude "northern USA" as the source population of any target individual.

316 (c) Multiple introductions at a single site are expected to leave a genetic signature only if
317 the individuals introduced originate from sources that are genetically differentiated from the
318 outbreak considered. Given the number of loci used, only first-generation immigrants would
319 be detectable (see Rannala and Mountain 1997, for a discussion on the power of statistical
320 tests of assignment). We used GENECLASS 2 ver. 2.0.g (Piry et al. 2004) to detect first-
321 generation immigrants. Given that CSE European individuals were sampled in 2003, this test
322 is specific to introductions in 2002-2003. A thousand individuals were simulated per
323 population, with the algorithm of Paetkau *et al.* (2004) and the assignment likelihood
324 calculation of Rannala & Mountain (1997) was used. The statistic used to identify first-
325 generation immigrants was the L_{home}/L_{max} ratio, where L_{home} is the assignment likelihood of an
326 individual to the population from which it was sampled and L_{max} is the maximum assignment
327 likelihood for this individual over all populations considered. As explained above populations
328 considered as potential sources were those with a first observation year $\leq$ 2003, including the
329 CSE European outbreak itself, with the exception of Friuli which was excluded from the
330 analysis. We considered an individual to be a first-generation immigrant if the probability of
331 the corresponding L_{home}/L_{max} ratio was below alpha=0.05.

332 ***RESULTS***

333 ***Genetic variation within WCR samples***

334 The dataset of CSE European samples showed moderate polymorphism, with a mean of
335 4.625 (SD = 2.973) alleles per locus over all samples. The mean number of alleles was
336 homogeneous over CSE Europe, varying from 3 (SD = 1.195) (*AR* = 2.996, SD = 1.192) in
337 Turanovac to 3.625 (SD = 1.685) (*AR* = 3.307, SD = 1.411) in Kaba (Table S2). Mean
338 expected heterozygosity ranges from 0.442 in Sarbogard to 0.506 in Kunmadaras (Table S2).
339 The mean variance of absolute allelic size varied from 10.366 (SD = 14.685) at Belgrade-

340 Airport to 14.238 (SD = 20.929) in Vardomb (Table S2). No significant differences between
341 samples were detected for any of the statistics summarizing genetic variation within
342 population samples (Friedman's test by rank performed over loci, χ^2=23.17, 19.4 and 24.60,
343 df=18, p= 0.184, 0.368 and 0.136 for *AR*, *H* and *V,* respectively).

344 We found significantly fewer alleles (-55%) and a lower heterozygosity (-26%) in CSE
345 Europe (complete data set) than in Pennsylvania, the putative source population of this
346 outbreak in north-eastern America (Wilcoxon's signed rank tests, p = 0.018 and 0.036 for *AR*
347 and *H,* respectively). The mean variance of absolute allelic size (*V*) did not differ significantly
348 between CSE Europe (complete data set) and Pennsylvania (Table S2; Wilcoxon's signed
349 rank test, p = 0.484).

350 ***Genetic variation between CSE European WCR samples***

351 Most pairwise comparisons (79% before and 87% after correction for multiple comparisons)
352 revealed no significant genetic differentiation (Table S3) and pairwise F_{ST}-values were low
353 (mean = 0.004, SD = 0.01). The sample collected from the site of the first observation within
354 the area studied (i.e. the Belgrade-Airport sample) displayed significant genetic differentiation
355 from all other samples except the Surcin, Stara-Pazova, Vrbas, and Kardoskut samples (Table
356 S3). The three north-eastern samples, Kunmadaras, Kaba, and Bekes, were genetically
357 differentiated from the two central samples (Vrbas and Sajkas) and the three western samples
358 (Bokod, Sarbogard, and Vardomb) (Table S3).

359 ***Clustering analysis of WCR population genetic structure in CSE Europe***

360 Bayesian clustering analyses performed with STRUCTURE provided consistent results over
361 the 10 runs tested for each *K* and over the two models tested, irrespective of the length of the
362 runs. The natural logarithm of the likelihood of the data lnP(X|*K*) slightly increased from *K* =

363 1 to $K = 2$, for which it was maximal (Figure S1A & S1B). Using the ΔK statistics led to the
364 same estimation of K=2 (data not shown). The following analyses of model parameters at $K =$
365 2 indicated an absence of population structure in the area of expansion in CSE Europe
366 (Pritchard et al. 2007; Pritchard et al. 2009). For example, in both analyses, all individuals
367 were admixed and the individual proportions of ancestry from each cluster (Q values) were
368 randomly distributed between samples. In the analysis in which sampling location was not
369 used as prior information, Q values were roughly symmetric (~ 1/2 in each cluster).
370 Moreover, in the analysis in which sampling location was used as prior information, at $K = 2$,
371 99% of individuals were assigned to a single cluster with a mean $Q = 0.89$ (SD = 0.11)
372 (Figure S2) and the value of r, which parameterizes the amount of information carried by the
373 locations, was much greater than 1. This is also consistent with an absence of population
374 structure (Pritchard et al. 2009).

375 GENELAND and BAPS confirmed the absence of genetic structure throughout the expansion
376 area in CSE Europe, with all individuals belonging to a single cluster. Posterior distributions
377 of the estimated number of clusters (K) across the 10 replicates performed in GENELAND
378 displayed a clear mode at $K = 1$ (Figure S1B) and a probability of 1 for $K = 1$ was obtained
379 for all calculation in BAPS.

380 ***Geographic and temporal analyses of genetic variation***

381 Weak but significant genetic isolation by geographic distance was detected only in the
382 eastern transect (Figure 2A; Mantel test on the correlation between $F_{ST}/(1-F_{ST})$ and the natural
383 logarithm of geographic distance between samples, $p = 0.047$, slope = 0.002 for the eastern
384 transect and p= 0.115, slope = 0.005 for the western transect). When combining the
385 probabilities obtained for the western and eastern transects by Fisher's method (Sokal and

386 Rolf 1995, p.794-797), we detected an overall significant genetic isolation by geographic
387 distance ($p = 0.034$).

388 Weak but significant genetic isolation by temporal distance was observed in both transects
389 (Figure 2B; Mantel tests on the correlation between F_{ST} and $\Delta_1^{st}{}_{obs}$, p= 0.022 and 0.020, and
390 slopes ≤ 0.002 for the western and the eastern transect, respectively). After combining the two
391 transect probabilities by Fisher's method, overall genetic isolation by temporal distance was
392 highly significant ($p = 0.004$).

393 Finally, a highly significant but weak correlation between F_{ST} and the differences of angle
394 measured for all pairs of samples was detected (Figure 2C; Mantel test, $p = 7.6\text{x}10^{-3}$; slope =
395 $1.3\text{x}10^{-4}$).

396 No significant correlation between the allelic richness of the samples (*AR*) and the year of
397 first observation or geographic distance from Belgrade was detected in either of the transects
398 (Spearman's $r \leq 0.15$ and $p \geq 0.71$) (Figure 3A & 3B). Allele size variance (*V*) displayed a
399 positive correlation of marginal significance with the year of first observation and with
400 geographic distance from Belgrade, in the western transect (Spearman's $r = 0.65$ and $p = 0.07$
401 for both tests), but not in the eastern transect (Spearman's $r \leq 0.33$ and $p \geq 0.35$) (Figure 3C &
402 3D). Furthermore, heterozygosity (*H*) was positively correlated with the year of first
403 observation and with the geographic distance from Belgrade in the eastern transect
404 (Spearman's $r \geq 0.8$ and $p = 0.01$ for both tests), but not in the western transect (Spearman's
405 $r = 0.09$ and $p = 0.84$ for both tests) (Figure 3E & 3F). When combining the probabilities
406 obtained for the two transects by Fisher's method, *H* displayed a positive correlation of
407 marginal significance with the year of first observation ($p = 0.06$) and a significant positive
408 correlation with geographical distance ($p = 0.03$). When probabilities were combined,
409 correlations based on *AR* and *V* remained non significant ($p > 0.11$).

410 ***Source populations of the Frickingen, Passau and Friuli outbreaks***

411 Comparisons between the Frickingen and Passau samples and all potential source
412 populations in Europe and North America gave small to large F_{ST} estimates, ranging from
413 0.002 between Frickingen and Bokod to 0.345 between Passau and Piedmont (Table 1). All
414 pairwise genetic differentiations tests gave significant results, with the exception of the six
415 comparisons of Frickingen with six sites sampled in the CSE European area: Deutsch-
416 Jahrndorf, Babolna, Bokod, Sarbogard and Vrbas, all from the western transect, and Szeged
417 (Table 1).

418 The CSE European area was identified as the most probable source population for the
419 Frickingen outbreak with the highest $L_{i\rightarrow s}$ and the minimum F_{ST}-values obtained for Vrbas
420 and Bokod, respectively (Table 1). More generally, $L_{i\rightarrow s}$ and F_{ST}-values indicated that the
421 western part of the CSE European area of expansion was the most probable source region for
422 the Frickingen outbreak.

423 For the Passau outbreak, the highest $L_{i\rightarrow s}$ value was obtained for the Frickingen sample
424 (Table 1). By contrast, the minimum F_{ST}-value was obtained for a CSE European sample,
425 Vardomb, identifying this sample site as the potential source of the Passau outbreak (Table 1).
426 We therefore cannot unambiguously identify a single source population for the Passau
427 outbreak. As the Frickingen and many CSE European samples are genetically similar (Table
428 1), similar F_{ST} and $L_{i\rightarrow s}$ values were obtained when assigning Passau to Frickingen or CSE
429 European samples (e.g. Vardomb and Sarbogard). The source of the Passau outbreak may
430 therefore be either the area of expansion in CSE Europe or the Frickingen outbreak.

431 Finally, both the $L_{i\rightarrow s}$ and F_{ST} values identify the area of expansion in CSE Europe as the
432 most probable source of the Friuli outbreak. More precisely $L_{i\rightarrow s}$ values identified Turanovac

433 and F_{ST} values identified Belgrade-Airport or Stara-Pazova, thus indicating that the center and
434 western part of the CSE European area of expansion were the most likely source regions for
435 the Friuli outbreak (Table 1).

436 Building on previous results, it is now possible to document the effect of introduction events
437 on genetic variation within populations, by comparing the Frickingen, Passau and Friuli
438 outbreaks with their identified source populations. For simplicity, data for the Passau, Friuli
439 and Frickingen outbreaks were first compared with the complete CSE European data set. The
440 Passau and Friuli outbreaks were less variable than their putative sources in CSE Europe,
441 whereas the Frickingen outbreak was not. Allelic richness (*AR*) was 16.5% and 44.3% lower
442 in the Passau and Friuli samples, respectively, than in CSE Europe (Wilcoxon's sign rank
443 tests, $p = 0.018$ for both tests). Heterozygosity (*H*) was 26.7% and 38.66% lower in the
444 Passau and Friuli samples, respectively, than in CSE Europe (Wilcoxon's sign rank tests, $p =$
445 0.018 for both tests). By contrast, *AR* and *H* were similar in Frickingen and in CSE Europe
446 (Wilcoxon's sign rank tests, $p = 0.063$ for *AR* and 0.176 for *H*). The mean variance of
447 absolute allelic size (*V*) was similar for the Frickingen, Passau and Friuli samples and for CSE
448 Europe (Wilcoxon's sign rank tests, $p > 0.128$ for each test). Genetic variation within the
449 Passau outbreak did not differ significantly from that in its alternative putative source
450 population, Frickingen (Wilcoxon's signed rank tests on *AR*, *H* and *V*, $p>0.09$ for each test).

451 ***Determination of the number of introductions into CSE Europe***

452 The three procedures used to detect multiple introductions in the CSE European outbreak
453 gave similar results:

454 a) Highest mean multilocus individual assignment likelihood of each sample *i* to each
455 possible source population ($L_{i\rightarrow s}$) clearly identified the northern USA (represented by the
456 Illinois and Pennsylvania samples) as the most probable source of the 19 sampling sites in

457 CSE Europe (Table 2). The highest $L_{i\rightarrow s}$ obtained between each CSE European sample and the
458 northern USA samples was substantially higher than the second highest $L_{i\rightarrow s}$ (with a
459 difference of 0.42 to 1.64 $\log_{10}$ ($L_{i\rightarrow s}$) units). Similar results were obtained with F_{ST}-values for
460 nine sampling sites in CSE Europe. F_{ST}-values identified Alsace as the probable source for
461 the other ten sites (Table 2).

462 b) Northern USA (represented by the Illinois and Pennsylvania samples) was excluded as a
463 possible source population (p<0.05) for only six individuals from CSE Europe (i.e. only 0.8%
464 of the total of 706 individuals). Setting the threshold of this analysis to 0.05, we expected a
465 mean of 0.05 x 706 = 35.3 type I errors in the data set and thus considered the six individuals
466 excluded to be negligible.

467 c) With an alpha of 0.05, none of the 706 individuals sampled in CSE Europe was identified
468 as a first-generation immigrant originating from a genetically differentiated population (p >
469 0.06). With an alpha of 0.2, 13 individuals were identified as first generation immigrants
470 coming from the CSE European outbreak itself. These 13 individuals were not spatially
471 concentrated. For the other 693 individuals, the probability of the corresponding L_{home}/L_{max}
472 ratio was > 0.55.

473 We thus obtained no evidence of multiple introductions from genetically differentiated
474 source populations at different sites within the CSE European area. Moreover, when Belgrade
475 was considered as a possible source, the $L_{i\rightarrow s}$ and F_{ST}-values identified the Belgrade area as
476 the most likely source population for all other CSE European samples (Table 2). Thus, we
477 found no evidence of multiple introductions from the same source population at different sites
478 within the CSE European area.

479

DISCUSSION

By combining a specific sampling scheme along the WCR expansion, microsatellite data and historical information, we characterized the invasion dynamics of WCR in its largest area of expansion in Europe, that of the Central and South-Eastern European outbreak. Within this area, we detected only one genetically homogeneous cluster based on Bayesian analyses, weak genetic differentiation between pairs of samples and weak but statistically significant patterns of genetic isolation by geographic distance and temporal distance. Unexpectedly, it seems that we have highlighted a very small increase in genetic variation from the center to the edge of the outbreak. Finally, we showed that three small geographically distant outbreaks (two in southern Germany and one in north-eastern Italy) most likely originated from CSE Europe. We will present the evidence to suggest that the CSE European outbreak (i) originated from a single introduction, (ii) is expanding through both continuous diffusion and discontinuous long-distance dispersal, and thus through stratified dispersal (Shigesada et al. 1995), and (iii) that human efforts to control WCR may account for the slight increase in genetic variation in the direction of expansion.

A single origin for the CSE European outbreak

Invading populations may result from one or more introduction events. In the case of repeated introductions from genetically differentiated source populations into different locations of a new area, we expect to observe, at least transiently, a mosaic of genetically differentiated patches within the area of expansion (e.g. Genton et al. 2005 for the common ragweed; Voisin et al. 2005 for a brown alga). No such evidence of multiple introductions from genetically differentiated source populations at different sites was observed in the large area of the CSE European outbreak of WCR. Indeed, measurements of interpopulation genetic variation and spatial Bayesian clustering indicated little or no genetic structure in the CSE

504 Europe area of expansion of WCR. This genetic homogeneity is consistent with the
505 homogeneity in susceptibility and resistance to insecticides reported in CSE Europe (Ciosi et
506 al. 2009).

507 The highest $L_{i\to s}$ values identified the northern USA population as the source of all CSE
508 European samples. However, the minimum F_{ST} values identified Alsace (France) as a possible
509 source of some CSE European samples. This result is not particularly surprising given the
510 substantial genetic similarity between the Alsace and northern USA populations (Ciosi et al.
511 2008). However, we consider that the Alsace population is not a likely source population for
512 CSE European sites because (i) the Alsace population was first observed in 2003 (the year of
513 CSE Europe sampling), (ii) population density was very low in Alsace at this time (we
514 analyzed all nine beetles observed), (iii) $L_{i\to s}$ values clearly identify the northern USA
515 population as the source population while F_{ST} do not clearly identify Alsace and (iv) F_{ST}
516 estimate is not corrected to reduce the influence of bottleneck during introduction (Gaggiotti
517 and Excoffier 2000), in contrast to $L_{i\to s}$ (Pascual et al. 2007). Moreover, the hypothesis of a
518 common spatial origin (in northern USA) of individuals sampled in CSE Europe could not be
519 rejected for 700 of the 706 individuals studied. Multiple introductions from the same source
520 population or from genetically similar populations would be expected to be genetically
521 equivalent to a single introduction of a large number of individuals (Roman and Darling
522 2007). However, this would probably not be the case if the multiple introductions occurred at
523 different sites within the invaded area. The hypothesis of multiple introductions from the
524 same or from genetically similar source populations at a single location cannot therefore be
525 rejected for CSE Europe, but we found no evidence of multiple introductions into different
526 locations within the CSE European area. Based on the parsimony principle, this outbreak
527 therefore probably corresponds to a single expanding population originating from a single
528 introduction in the Belgrade area.

529 ***Expansion process in CSE Europe***

530 The expansion process should leave specific genetic signatures in invading populations. The
531 simplest expansion model for invading species, the "wave of advance" model (Fisher 1937),
532 considers dispersal to be a random diffusion process. Under this model, we expect a pattern of
533 genetic isolation by geographic distance, the intensity of which depends on the balance
534 between gene flow and drift (Slatkin 1993). We detected such a pattern of genetic isolation by
535 geographic distance in the CSE European outbreak of WCR, indicating greater gene flow
536 between geographically close than between geographically distant locations, and thus that the
537 dispersal of WCR is spatially limited.

538 Theoretical studies have shown that founder events in populations located at the edge of the
539 expansion may have two consequences: (i) substantial genetic differentiation between the
540 center and the periphery of a colonized area (Excoffier and Ray 2008; Le Corre and Kremer
541 1998), and (ii) a decrease in genetic variation in the direction of colonization (Hallatschek and
542 Nelson 2008; Le Corre and Kremer 1998). A decrease in genetic variation in the direction of
543 colonization has been documented for many organisms (e.g. Prugnolle et al. 2005 for humans;
544 Williams et al. 2007 for Brazilian peppertrees). In CSE Europe, we found population genetic
545 differentiation to be weak and we observed no decrease in genetic variation in the direction of
546 colonization. This suggests an absence of successive major founder events at the front of the
547 invaded area during the expansion process or that the effect of dispersal outweighed the effect
548 of genetic drift related to the founder events. Theoretical and simulation studies have shown
549 that a combination of short- and long-distance dispersal (i.e. stratified dispersal), during
550 geographic expansion may maintain genetic diversity in expanding populations (Bialozyt et
551 al. 2006; Davies et al. 2004; Ibrahim et al. 1996; Le Corre and Kremer 1998). However this
552 dispersal model does not account for the small increase in genetic variation along the axis of

553 colonization suggested by the weak but significant positive correlation of heterozygosity (*H*)
554 with the year of first observation and with the geographic distance from Belgrade in the
555 eastern transect. We argue that human activities, including control measures against WCR in
556 particular, could be responsible for the observed pattern (no decrease in genetic variation and
557 even a small increase in the direction of colonization). Significant damage to crops in Serbia
558 and southern Hungary, has led to the establishment of integrated pest management (IPM),
559 including crop rotation, which may have resulted in a decrease in WCR population density.
560 Such a decrease has been documented in Serbia and southern Hungary following the
561 establishment of IPM (Boriani et al. 2006; Ripka and Princzinger 2001; Sivcev and Stankovic
562 2004; Sivcev et al. 2009). In 2003, the year of sampling, IPM against WCR had been used in
563 Serbia for at least five years (Sivcev and Stankovic 2004; Sivcev et al. 2009) and in southern
564 Hungary for two years (Ripka and Princzinger 2001). Few or no WCR control methods were
565 implemented in central and northern Hungary, and in the sampled areas of Austria and
566 Romania (Kiss et al. 2005). WCR populations may thus have underwent a decrease in size in
567 Serbia and in southern Hungary, with this decrease more persistent in Serbia, but no such
568 decrease in population size was observed further north. These demographic bottlenecks were
569 probably associated with genetic bottlenecks, which lasted longer in Serbia than in southern
570 Hungary. The erosion of the genetic variation induced by these genetic bottlenecks should
571 thus have been stronger in Serbia than in southern Hungary and absent in northern Hungary.
572 This may explain the pattern of absence of decrease and even of small increase in genetic
573 variation from Serbia to northern Hungary suggested by our results. This hypothesis is
574 consistent with the significant genetic differentiation observed between the sample collected
575 at the location at which WCR was first sighted in the study area (i.e. the Belgrade-Airport
576 sample) and almost all other samples. Differences in population densities due to geographic
577 and temporal heterogeneities in control measures may frequently occur in a number of pest

578 species. It is therefore possible that such a pattern of absence of decrease and small increase in
579 genetic variation during colonization may be found in other invading pests for which control
580 strategies are used, above a certain density threshold.

581

582 We observed a pattern of genetic isolation by angular distance together with a significant
583 east-west genetic differentiation in CSE Europe, suggesting that the expanding population is
584 divided into genetically differentiated sectors. The fragmentation of a colonized region into
585 genetically differentiated sectors and the occurrence of allele frequency clines parallel to the
586 colonization front have recently been attributed to a phenomenon called "gene surfing" in
587 theoretical studies (Edmonds et al. 2004; Klopfstein et al. 2006). During "gene surfing", rare
588 alleles or mutations may invade parts of the space not yet colonized due to genetic drift at the
589 edge of an expanding population (Excoffier and Ray 2008; Hallatschek et al. 2007). However,
590 in our case, we observed no evidence of genetic drift at the edge of the expansion (weak
591 population genetic differentiation and no loss of diversity in the direction of colonization.).
592 An anisotropic dispersal of WCR with effective dispersal more frequent in the direction of
593 expansion due to weak competition beyond the front might also account for the observed
594 pattern. Computer simulation-based studies are needed to test this hypothesis.

595 ***WCR is expanding in CSE Europe through stratified dispersal***

596 The Frickingen outbreak in south-eastern Germany resulted from an introduction from the
597 western part of the CSE European WCR area. Similarly, the Friuli outbreak resulted from an
598 introduction of WCR from the center and/or the western part of the CSE European area.
599 Finally, Frickingen and the western part of the CSE European outbreak can both be sources of
600 the Passau population in south-western Germany. These geographically isolated outbreaks
601 thus probably result from long-distance dispersal events from the continuously growing CSE

602 European area, with different degrees of genetic diversity loss. Therefore, WCR appears to
603 expand in CSE Europe through both local continuous dispersal and discontinuous long-
604 distance dispersal. This corresponds to a stratified dispersal (Shigesada et al. 1995) process, in
605 which satellite colonies are founded outside the main expanding population. These colonies
606 eventually merge with the main expanding area and contribute to advance the front
607 (Shigesada et al. 1995).

608 Large distances (>300 km) separate the Friuli and Frickingen outbreaks from their source
609 populations, and the Alps stand between the CSE European area of expansion and both these
610 isolated outbreaks. Moreover, the CSE European outbreak was not the geographically closest
611 population to Friuli (when sampling was performed) and to Frickingen. When they were first
612 observed, these outbreaks were closer to NW Italy and Alsace respectively. Human activities
613 may therefore be the chief means of long-distance WCR dispersal in Europe, although north
614 American studies have also suggested a major role for wind in the long-distance dispersal of
615 WCR (Isard et al. 2004; Onstad et al. 1999). Whatever the means of transport, stratified
616 dispersal seems to have a major impact on the geographic expansion of the species. The
617 growth of satellite colonies and their coalescence with the main expanding population may
618 lead to greater rates of geographic expansion than observed in cases of expansion without
619 long-distance dispersal (Shigesada et al. 1995). Moreover, the longer the length of the
620 expansion front, the more frequent long-distance dispersal events are likely to be. Therefore
621 stratified dispersal may also result in an acceleration of the rate of expansion as the length of
622 the front increases whereas this rate remains constant in a continuous diffusion (Shigesada et
623 al. 1995). Metcalf (1983) reported such an acceleration of the rate of geographic expansion of
624 WCR in the US Corn Belt and Gray *et al.* (2009) concluded that it probably resulted from
625 stratified dispersal. A similar acceleration of the rate of expansion of the CSE European
626 population may therefore be expected in the near future.

Conclusion

Genetic evidence suggest that a single introduction is responsible for the foundation of the Central and South-Eastern European outbreak of the western corn rootworm and that this invasive population is expanding through stratified dispersal. Combined with historical documentation of population size and of the establishment of management strategies, our results also suggest that control measures against the western corn rootworm are probably responsible for genetic bottlenecks at the center of the outbreak. Thus, human activities probably affect the population dynamics of the pest, fortuitously increasing its capacity to disperse or deliberately decreasing pest population densities, thereby globally affecting the population structure of the western corn rootworm.

638

639 ***LITERATURE CITED***

640 Anonymous. 2007. News Related to IWGO Matters, Pages 2 *in* U. Kuhlmann, ed. IWGO Newsletter.
641 Delémont, Switzerland, Harald K. Berger.
642 Baufeld, P., and S. Enzian. 2001. Simulations model for spreading scenarios of western corn rootworm
643 (*Diabrotica virgifera virgifera*) in case of Germany. IWGO Newsletter 22:14-15.
644 Benjamini, Y., and Y. Hochberg. 1995. Controlling the False Discovery Rate - a Practical and
645 Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B-
646 Methodological 57:289-300.
647 Bialozyt, R., B. Ziegenhagen, and R. J. Petit. 2006. Contrasting effects of long distance seed dispersal
648 on genetic diversity during range expansion. Journal of Evolutionary Biology 19:12-20.
649 Boriani, M., M. Agost, J. Kiss, and C. R. Edwards. 2006. Sustainable management of the western corn
650 rootworm, *Diabrotica virgifera virgifera* LeConte (Coleoptera: Chrysomelidae), in infested areas:
651 experiences in Italy, Hungary and the USA. EPPO Bulletin 36:531-537.
652 Ciosi, M., N. J. Miller, K. S. Kim, R. Giordano, A. Estoup, and T. Guillemaud. 2008. Invasion of
653 Europe by the western corn rootworm, *Diabrotica virgifera virgifera*: multiple transatlantic
654 introductions with various reductions of genetic diversity. Molecular Ecology 17:3614-3627.
655 Ciosi, M., S. Toepfer, H. Li, T. Haye, U. Kuhlmann, H. Wang, B. D. Siegfried et al. 2009. European
656 populations of *Diabrotica virgifera virgifera* are resistant to aldrin, but not to methyl parathion.
657 Journal of Applied Entomology 133:307–314.
658 Coats, S. A., J. J. Tollefson, and J. A. Mutchmor. 1986. Study of migratory flight in the western corn
659 rootworm (Coleoptera: Chrysomelidae). Environmental Entomology 15:620-625.
660 Coombs, J. A., B. H. Letcher, and K. H. Nislow. 2008. CREATE: a software to create input files from
661 diploid genotypic data for 52 genetic software programs. Molecular Ecology Resources 8:578-
662 580.
663 Corander, J., P. Waldmann, M. P., and M. J. Sillanpää. 2004. BAPS 2: enhanced possibilities for the
664 analysis of genetic population structure. Bioinformatics 20:2363-2369.
665 Corander, J., P. Waldmann, and M. J. Sillanpää. 2003. Bayesian analysis of genetic differentiation
666 between populations. Genetics 163:367-374.
667 Cornuet, J. M., F. Santos, M. A. Beaumont, C. P. Robert, J. M. Marin, D. J. Balding, T. Guillemaud et
668 al. 2008. Inferring population history with DIY ABC: a user-friendly approach to approximate
669 Bayesian computation. Bioinformatics 24:2713-2719.
670 Davies, S., A. White, and A. Lowe. 2004. An investigation into effects of long-distance seed dispersal
671 on organelle population genetic structure and colonization rate: a model analysis. Heredity 93:566-
672 576.
673 Drake, J. M., and D. M. Lodge. 2006. Allee effects, propagule pressure and the probability of
674 establishment: Risk analysis for biological invasions. Biological Invasions 8:365-375.
675 Edmonds, C. A., A. S. Lillie, and L. L. Cavalli-Sforza. 2004. Mutations arising in the wave front of an
676 expanding population. Proceedings of the National Academy of Sciences of the United States of
677 America 101:975-979.
678 Edwards, C. R., and J. Kiss. 2007. New WCR 2006 General Spread Map for Europe. IWGO
679 Newsletter 28:3.
680 Elam, D. R., C. E. Ridley, K. Goodell, and N. C. Ellstrandt. 2007. Population size and relatedness
681 affect fitness of a self-incompatible invasive plant. Proceedings of the National Academy of
682 Sciences of the United States of America 104:549-552.
683 Estoup, A., M. Beaumont, F. Sennedot, C. Moritz, and J. M. Cornuet. 2004. Genetic analysis of
684 complex demographic scenarios: Spatially expanding populations of the cane toad, *Bufo marinus*.
685 Evolution 58:2021-2036.
686 Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the
687 software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620.

688 Excoffier, L., and N. Ray. 2008. Surfing during population expansions promotes genetic revolutions
689 and structuration. Trends in Ecology & Evolution 23:347-351.
690 Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus
691 genotype data: Linked loci and correlated allele frequencies. Genetics 164:1567-1587.
692 Fisher, R. A. 1937. The wave of advance of advantageous genes. Annals of Eugenics 7:355-369.
693 Gaggiotti, O. E., and L. Excoffier. 2000. A simple method of removing the effect of a bottleneck and
694 unequal population sizes on pairwise genetic distances. Proceedings of the Royal Society of
695 London Series B-Biological Sciences 267:81-87.
696 Genton, B. J., J. A. Shykoff, and T. Giraud. 2005. High genetic diversity in French invasive
697 populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of
698 introduction. Molecular Ecology 14:4275-4285.
699 Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version
700 2.9.3). Updated from Goudet (1995). Available at:
701 http://www2.unil.ch/popgen/softwares/fstat.htm.
702 Grant, R. H., and K. P. Seevers. 1989. Local and long-range movement of adult western corn
703 rootworm (Coleoptera: Chrysomelidae) as evidenced by washup along southern Lake Michigan
704 shores. Environmental Entomology 18:266-272.
705 Gray, M. E., T. W. Sappington, N. J. Miller, J. Moeser, and M. O. Bohn. 2009. Adaptation and
706 Invasiveness of Western Corn Rootworm: Intensifying Research on a Worsening Pest. Annual
707 Review of Entomology 54:303-321.
708 Guillot, G., A. Estoup, F. Mortier, and J. F. Cosso. 2005a. A spatial statistical model for landscape
709 genetics. Genetics 170:1261-1280.
710 Guillot, G., F. Mortier, and A. Estoup. 2005b. GENELAND: a computer package for landscape
711 genetics. Molecular Ecology Notes 5:712-715.
712 Hallatschek, O., P. Hersen, S. Ramanathan, and D. R. Nelson. 2007. Genetic drift at expanding
713 frontiers promotes gene segregation. Proceedings of the National Academy of Sciences of the
714 United States of America 104:19926-19930.
715 Hallatschek, O., and D. R. Nelson. 2008. Gene surfing in expanding populations. Theoretical
716 Population Biology 73:158-170.
717 Hubisz, M. J., D. Falush, M. Stephens, and J. K. Pritchard. 2009. Inferring weak population structure
718 with the assistance of sample group information. Molecular Ecology Resources 9:1322 -1332.
719 Ibrahim, K. M., R. A. Nichols, and G. M. Hewitt. 1996. Spatial patterns of genetic variation generated
720 by different forms of dispersal during range expansion. Heredity 77:282-291.
721 Isard, S. A., J. L. Spencer, T. R. Mabry, and E. Levine. 2004. Influence of atmospheric conditions on
722 high-elevation flight of western corn rootworm (Coleoptera: Chrysomelidae). Environmental
723 Entomology 33:650-656.
724 Johnson, D. M., A. M. Liebhold, P. C. Tobin, and O. N. Bjornstad. 2006. Allee effects and pulsed
725 invasion by the gypsy moth. Nature 444:361-363.
726 Kim, K. S., and T. W. Sappington. 2005a. Genetic structuring of western corn rootworm (Coleoptera:
727 Chrysomelidae) populations in the United States based on microsatellite loci analysis.
728 Environmental Entomology 34:494-503.
729 —. 2005b. Polymorphic microsatellite loci from the western corn rootworm (Insecta: Coleoptera:
730 Chrysomelidae) and cross-amplification with other *Diabrotica* spp. Molecular Ecology Notes
731 5:115-117.
732 Kirkpatrick, M., and N. H. Barton. 1997. Evolution of a species' range. American Naturalist 150:1-23.
733 Kiss, J., C. R. Edwards, H. K. Berger, P. Cate, M. Cean, S. Cheek, J. Derron et al. 2005. Monitoring of
734 western corn rootworm (*Diabrotica virgifera virgifera* LeConte) in Europe 1992-2003, Pages 29-
735 39 *in* S. Vidal, U. Kuhlmann, and C. R. Edwards, eds. Western Corn Rootworm: Ecology and
736 Management. Cambridge, MA USA, CABI Publishing.
737 Klopfstein, S., M. Currat, and L. Excoffier. 2006. The fate of mutations surfing on the wave of a range
738 expansion. Molecular Biology and Evolution 23:482-490.
739 Le Corre, V., and A. Kremer. 1998. Cumulative effects of founding events during colonisation on
740 genetic diversity and differentiation in an island and stepping-stone model. Journal of
741 Evolutionary Biology 11:495-512.

742 Metcalf, R. L. 1983. Implications and prognosis of resistance to insecticides, Pages 703-733 *in* G. P.
743 Georghio, and T. Saito, eds. Pest Resistance to Pesticides. New York, NY, Plenum.
744 Miller, N., A. Estoup, S. Toepfer, D. Bourguet, L. Lapchin, S. Derridj, K. S. Kim et al. 2005. Multiple
745 transatlantic introductions of the western corn rootworm. Science 310:992.
746 Miller, N. J., M. Ciosi, T. W. Sappington, S. T. Ratcliffe, J. L. Spencer, and T. Guillemaud. 2007.
747 Genome scan of *Diabrotica virgifera virgifera* for genetic variation associated with crop rotation
748 tolerance. Journal of Applied Entomology 131:378-385.
749 Moody, M. E., and R. N. Mack. 1988. Controlling the spread of plant invasions: The importance of
750 nascent foci. Journal of Applied Ecology 25:1009-1021.
751 Nei, M. 1987, Molecular Evolutionary Genetics. New York, Columbia University Press.
752 Olden, J. D., N. LeRoy Poff, M. R. Douglas, M. E. Douglas, and K. D. Fausch. 2004. Ecological and
753 evolutionary consequences of biotic homogenization. Trends in Ecology & Evolution 19:18-24.
754 Onstad, D. W., M. G. Joselyn, S. A. Isard, E. Levine, J. L. Spencer, L. W. Bledsoe, C. R. Edwards et
755 al. 1999. Modeling the spread of western corn rootworm (Coleoptera: Chrysomelidae) populations
756 adapting to soybean-corn rotation. Environmental Entomology 28:188-194.
757 Paetkau, D., R. Slade, M. Burden, and A. Estoup. 2004. Genetic assignment methods for the direct,
758 real-time estimation of migration rate: a simulation-based exploration of accuracy and power.
759 Molecular Ecology 13:55-65.
760 Pascual, M., M. P. Chapuis, F. Mestres, J. Balanya, R. B. Huey, G. W. Gilchrist, L. Serra et al. 2007.
761 Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey
762 using ABC methods. Molecular Ecology 16:3069-3083.
763 Petit, R. J., A. El Mousadik, and O. Pons. 1998. Identifying populations for conservation on the basis
764 of genetic markers. Conservation Biology 12:844-855.
765 Pimentel, D., S. McNair, J. Janecka, J. Wightman, C. Simmonds, C. O'Connell, E. Wong et al. 2001.
766 Economic and environmental threats of alien plant, animal, and microbe invasions. Agriculture
767 Ecosystems & Environment 84:1-20.
768 Piry, S., A. Alapetite, J. M. Cornuet, D. Paetkau, L. Baudouin, and A. Estoup. 2004. GENECLASS2:
769 A software for genetic assignment and first-generation migrant detection. Journal of Heredity
770 95:536-539.
771 Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using
772 multilocus genotype data. Genetics 155:945-959.
773 Pritchard, J. K., X. Wen, and D. Falush. 2007. Documentation for *structure* software: Version 2.2.
774 Chicago, USA: University of Chicago. Available at: http://pritch.bsd.uchicago.edu/structure.html.
775 —. 2009. Documentation for *structure* software: Version 2.3. Chicago, USA: University of Chicago.
776 Available at: http://pritch.bsd.uchicago.edu/structure.html.
777 Prugnolle, F., A. Manica, and F. Balloux. 2005. Geography predicts neutral genetic diversity of human
778 populations. Current Biology 15:R159-R160.
779 R Development CoreTeam. 2008.R: A language and environment for statistical computing.R
780 Foundation for Statistical Computing, Vienna, Austria.
781 Rannala, B., and J. L. Mountain. 1997. Detecting immigration by using multilocus genotypes.
782 Proceedings of the National Academy of Sciences of the United States of America 94:9197-9201.
783 Raymond, M., and F. Rousset. 1995a. An exact test for population differentiation. Evolution 49:1280-
784 1283.
785 —. 1995b. Genepop (version. 1.2), a population genetics software for exact tests and ecumenicism.
786 Journal of Heredity 86:248-249.
787 Ripka, G., and G. Princzinger. 2001. Monitoring of western corn rootworm (*Diabrotica virgifera*
788 *virgifera* LeConte) in Hungary in 2001. IWGO Newsletter 22:27-28.
789 Roman, J., and J. A. Darling. 2007. Paradox lost: genetic diversity and the success of aquatic
790 invasions. Trends in Ecology & Evolution 22:454-464.
791 Rosenberg, N. A. 2004. DISTRUCT: a program for the graphical display of population structure.
792 Molecular Ecology Notes 4:137-138.
793 Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation
794 by distance. Genetics 145:1219-1228.

795 Ruiz, G. M., T. K. Rawlings, F. C. Dobbs, L. A. Drake, T. Mullady, A. Huq, and R. R. Colwell. 2000.
796 Global spread of microorganisms by ships - Ballast water discharged from vessels harbours a
797 cocktail of potential pathogens. Nature 408:49-50.
798 Sax, D. F., J. J. Stachowicz, and S. D. Gaines. 2005, Species Invasions: Insights into Ecology,
799 Evolution, and Biogeography. Sunderland, MA, U.S.A., Sinauer Associates.
800 Shigesada, N., K. Kawasaki, and Y. Takeda. 1995. Modeling stratified diffusion in biological
801 invasions. American Naturalist 146:229-251.
802 Sivcev, I., and S. Stankovic. 2004. Population level changes of western corn rootworm in SERBIA.
803 IWGO Newsletter 25:10.
804 Sivcev, I., S. Stankovic, M. Kostic, N. Lakic, and Z. Popovic. 2009. Population density of Diabrotica
805 virgifera virgifera LeConte beetles in Serbian first year and continuous maize fields. Journal of
806 Applied Entomology 133:430-437.
807 Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution
808 47:264-279.
809 Sokal, R. R., and F. J. Rolf. 1995, Biometry. The Principles and Practice of Statistics in Biological
810 Research. New York., W.H. Freeman and Company.
811 Spencer, J. L., E. Levine, S. A. Isard, and T. R. Mabry. 2005. Movement, dispersal and behaviour of
812 western corn rootworm adults in rotated maize and soybean fields, Pages 121-144 *in* S. Vidal, U.
813 Kuhlmann, and C. R. Edwards, eds. Western Corn Rootworm: Ecology and Management.
814 Cambridge, MA USA, CABI Publishing.
815 Sunnucks, P., and D. F. Hales. 1996. Numerous transposed sequences of mitochondrial cytochrome
816 oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). Molecular Biology and
817 Evolution 13:510-524.
818 Voisin, M., C. R. Engel, and F. Viard. 2005. Differential shuffling of native genetic diversity across
819 introduced regions in a brown alga: aquaculture vs. maritime traffic effects. Proceedings of the
820 National Academy of Sciences of the United States of America 102:5432-5437.
821 Weir, B. S., and C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure.
822 Evolution 38:1358-1370.
823 Williams, D. A., E. Muchugu, W. A. Overholt, and J. P. Cuda. 2007. Colonization patterns of the
824 invasive Brazilian peppertree, *Schinus terebinthifolius*, in Florida. Heredity 98:284-293.
825
826

Table 1: Pairwise estimates of F_{ST} (Weir and Cockerham 1984) and mean individual assignment likelihood ($L_{i\to s}$) of the Passau, Frickingen and Friuli samples for each potential source population

Geographic area	Potential source population	Location	1st obs.	Passau, Germany		Frickingen, Germany		Friuli, Italy	
				$L_{i\to s}$	F_{ST}	$L_{i\to s}$	F_{ST}	$L_{i\to s}$	F_{ST}
North American range	Mexico	Durango, Mexico	<1940	15.989	0.284	16.604	0.229	16.664	0.319
	Arizona	Willcox, Arizona	<1974	15.229	0.213	15.559	0.164	18.62	0.276
	Texas	New Deal, Texas	<1980	9.135	0.177	9.427	0.132	10.048	0.226
	Illinois	Champaign, Illinois	<1974	7.504	0.158	7.848	0.115	9.175	0.229
	Pennsylvania	Bellefonte, Pennsylvania	<1985	7.669	0.170	7.734	0.116	8.153	0.218
Western European isolated outbreaks	Piedmont	Oleggio, Italy	2000	15.092	0.345	13.652	0.242	15.354	0.357
	Paris-1	Roissy Airport, France	2002	11.379	0.299	11.775	0.241	14.494	0.439
	UK	Slough, United Kingdom	2003	8.174	0.191	8.877	0.134	11.301	0.285
	Friuli	Buttrio, Italy	2003	11.052	0.239	11.466	0.168	-	-
	Alsace	Mulhouse Airport, France	2003	7.881	0.176	8.678	0.106	9.44	0.267
	Paris-2	Pierrelaye, France	2004	11.350	0.220	12.054	0.177	12.777	0.278
	Passau	Passau, Germany	2007	3.378	-	5.438	0.063	5.891	0.239
	Frickingen	Frickingen, Germany	2007	**4.100**	0.063	4.270	-	5.144	0.168
Central and South Eastern European area of Expansion	Belgrade-Airport^{W}	Belgrade Airport, Serbia	1992	5.087	0.104	5.626	0.061	4.863	**0.116**
	Surcin^{E}	Surcin, Serbia	1992	4.328	0.045	4.98	0.027	5.209	0.153
	Stara-Pazova$^{W\&E}$	Stara Pazova, Serbia	1993	4.519	0.066	4.883	0.024	4.696	**0.116**
	Vrbas^{W}	Vrbas, Serbia	1995	4.184	0.051	**4.537**	0.004 ^{Ø}	4.934	0.143
	Aleksa-Santic^{W}	Aleksa Santic, Serbia	1996	4.263	0.053	4.613	0.006	5.025	0.147
	Vardomb^{W}	Vardomb, Hungary	1997	4.120	**0.039**	4.619	0.006	5.153	0.146
	Sarbogard^{W}	Sarbogard, Hungary	1999	4.146	0.051	4.576	0.012 ^{Ø}	4.806	0.136
	Bokod^{W}	Bokod, Hungary	2000	4.241	0.051	4.548	**0.002** ^{Ø}	5.32	0.155
	Babolna^{W}	Babolna, Hungary	2001	4.299	0.055	4.626	0.008 ^{Ø}	5.267	0.167
	Deutsch-Jahrndorf^{W}	Deutsch Jahrndorf, Austria	2002	4.287	0.051	4.609	0.004 ^{Ø}	5.184	0.159
	Sajkas^{E}	Sajkas, Serbia	1994-1995	4.286	0.057	4.705	0.015	4.957	0.145
	Szeged^{E}	Szeged, Hungary	1995	4.327	0.056	4.655	0.009 ^{Ø}	5.118	0.157
	Kardoskut^{E}	Kardoskut, Hungary	1996	4.276	0.045	4.906	0.020	5.066	0.149
	Kondoros^{E}	Kondoros, Hungary	1997	4.432	0.060	4.869	0.016	5.353	0.156
	Bekes^{E}	Bekes, Hungary	1998	4.829	0.087	5.214	0.043	5.245	0.158
	Kunmadaras^{E}	Kunmadaras, Hungary	1999	4.667	0.077	4.989	0.027	5.38	0.153
	Madaras^{E}	Madaras, Romania	1997-1999	4.474	0.063	4.770	0.013	5.329	0.153
	Kaba^{E}	Kaba, Hungary	2000	4.729	0.068	5.299	0.047	5.267	0.164
	Turanovac	Turanovac, Croatia	1999	4.182	0.042	4.759	0.017	**4.64**	0.125
Most likely source population				**Vardomb** ^{W}		**Vrbas** ^{W}	**Bokod** ^{W}	**Tura-novac**	**Belgrade-Airport**^{W} **or Stara-Pazova**$^{W\&}$

Note: $-\log_{10}$ values of the $L_{i\to s}$ are indicated. ^{Ø} indicates nonsignificant pairwise differentiation exact tests. ^{W}: sample from the western sampling transect. ^{E}: sample from the eastern sampling transect. 1st obs.: First observation year. Highest $L_{i\to s}$ and minimum F_{ST} are indicated in bold typeface. The most likely source population for each sample is indicated in the last row.

Table 2: Weir & Cockerham's estimator of pairwise F_{ST} (1984) and mean individual assignment likelihood ($L_{i\rightarrow s}$) of each Central and South-Eastern European sample of the western corn rootworm to each potential source population and to both samples from Belgrade area.

Sample names	Potential source populations									Belgrade area	
	North America					Europe					
	Mexico	Arizona	Texas	Illinois	Pennsylvania	Piedmont, Italy	Paris-1, France	Alsace, France	UK	Belgrade-Airport	Surcin
Belgrade-Airport^{W}	16.391 (0.224)	16.657 (0.167)	8.959 (0.103)	8.270 (0.109)	**7.589 (0.095)**	12.925 (0.197)	11.021 (0.257)	9.002 (0.118)	8.572 (0.126)	-	5.201 (0.014)
Surcin^{E}	16.555 (0.213)	16.894 (0.159)	9.188 (0.109)	8.386 (0.106)	**7.991 (0.104)**	15.022 (0.236)	11.499 (0.234)	8.893 (0.112)	8.951 (0.124)	5.346 (0.014)	-
Stara-Pazova$^{W\&E}$	17.323 (0.204)	17.096 (0.148)	9.236 (0.098)	8.495 (0.098)	**8.221** (0.095)	14.941 (0.217)	11.942 (0.228)	8.943 **(0.093)**	9.302 (0.120)	5.308 (0.005)	5.110 (-0.002)
Vrbas^{W}	16.190 (0.214)	15.511 (0.142)	8.897 (0.101)	7.827 (0.093)	**7.610** (0.091)	13.296 (0.204)	10.900 (0.217)	8.364 **(0.085)**	8.628 (0.109)	5.234 (0.017)	4.964 (0.006)
Aleksa-Santic^{W}	17.251 (0.210)	16.314 (0.144)	9.065 (0.100)	8.026 (0.092)	**7.879** (0.092)	14.543 (0.219)	11.352 (0.209)	8.842 **(0.087)**	8.888 (0.109)	5.442 (0.024)	5.124 (0.002)
Vardomb^{W}	16.893 (0.205)	16.245 (0.145)	9.315 (0.108)	**8.186** (0.095)	8.200 (0.100)	15.358 (0.236)	11.602 (0.213)	8.810 **(0.080)**	9.034 (0.116)	5.772 (0.039)	5.337 (0.007)
Sarbogard^{W}	16.739 (0.234)	16.112 (0.167)	9.158 (0.119)	8.285 (0.115)	**8.000** (0.113)	14.279 (0.227)	11.510 (0.240)	8.424 **(0.098)**	9.035 (0.132)	5.246 (0.020)	4.887 (0.005)
Bokod^{W}	16.198 (0.209)	15.077 (0.142)	9.003 (0.105)	7.887 (0.096)	**7.553** (0.097)	13.531 (0.225)	10.895 (0.216)	8.404 **(0.087)**	8.850 (0.118)	5.346 (0.026)	5.005 (0.005)
Babolna^{W}	16.727 (0.209)	16.232 (0.148)	9.284 (0.107)	8.278 (0.098)	**8.194** (0.104)	14.799 (0.231)	11.384 (0.208)	8.671 **(0.089)**	9.182 (0.114)	5.772 (0.035)	5.126 (0.000)
Deutsch-Jahrndorf^{W}	15.829 (0.196)	15.371 (0.135)	8.984 (0.098)	7.967 (0.087)	**7.882** (0.092)	14.265 (0.221)	10.967 (0.203)	8.607 **(0.082)**	9.059 (0.107)	5.689 (0.030)	5.164 (-0.001)
Sajkas^{E}	17.029 (0.221)	16.423 (0.158)	9.224 (0.115)	8.261 **(0.110)**	**8.220** (0.112)	15.338 (0.246)	11.875 (0.235)	8.844 (0.111)	9.471 (0.136)	5.530 (0.029)	5.009 (0.000)
Szeged^{E}	16.563 (0.198)	15.979 (0.138)	9.147 (0.093)	8.196 (0.082)	**7.887** (0.083)	14.454 (0.209)	11.416 (0.198)	8.935 **(0.082)**	9.044 (0.100)	5.669 (0.028)	5.208 (-0.002)
Kardoskut^{E}	16.252 (0.200)	16.406 (0.143)	9.124 (0.101)	8.338 (0.099)	**7.868 (0.097)**	14.697 (0.228)	11.757 (0.235)	8.997 (0.105)	9.036 (0.117)	5.375 (0.015)	5.043 (-0.007)
Kondoros^{E}	17.366 (0.204)	16.748 (0.139)	9.273 (0.091)	8.513 (0.088)	**8.210 (0.081)**	15.026 (0.198)	11.865 (0.206)	9.459 (0.096)	9.124 (0.097)	5.783 (0.019)	5.415 (-0.001)
Bekes^{E}	16.498 (0.192)	16.606 (0.137)	9.009 (0.087)	8.434 **(0.082)**	**8.025** (0.086)	15.058 (0.220)	11.453 (0.208)	9.487 (0.100)	9.254 (0.106)	5.640 (0.024)	5.271 (0.004)
Kunmadaras^{E}	16.467 (0.184)	16.254 (0.131)	9.163 (0.083)	8.404 (0.077)	**8.025 (0.077)**	14.663 (0.203)	11.535 (0.201)	9.683 (0.093)	9.158 (0.095)	5.714 (0.017)	5.354 (0.000)
Madaras^{E}	16.671 (0.192)	16.626 (0.137)	9.319 (0.091)	8.462 (0.084)	**8.181 (0.083)**	14.914 (0.207)	11.572 (0.202)	9.355 (0.086)	9.117 (0.098)	5.714 (0.018)	5.271 (-0.006)
Kaba^{E}	16.516 (0.196)	17.341 (0.148)	9.307 (0.104)	8.777 **(0.098)**	**8.142** (0.100)	16.312 (0.249)	11.768 (0.220)	9.559 (0.109)	9.566 (0.123)	5.836 (0.036)	5.421 (0.003)
Turanovac	16.598 (0.189)	16.729 (0.136)	9.233 (0.097)	8.294 (0.090)	**7.912** (0.094)	14.955 (0.225)	12.011 (0.215)	8.700 **(0.084)**	9.285 (0.112)	5.520 (0.023)	5.055 (0.000)

Note: Potential sources were populations with a first observation year ≤ 2003 (i.e. the year of collection of CSE European samples). The Friuli outbreak was excluded from the analysis because it is known to have originated in CSE Europe, and thus most CSE European samples are expected to wrongly point to Friuli as its most probable source. $-\log_{10}$ of the $L_{i\rightarrow s}$ are indicated and F_{ST} are in parentheses. ^{W}: sample from the western sampling transect. ^{E}: sample from the eastern sampling transect. For each CSE European sample the maximum $L_{i\rightarrow s}$ and minimum F_{ST} with North American and European samples are indicated in bold typeface. -: not suitable.

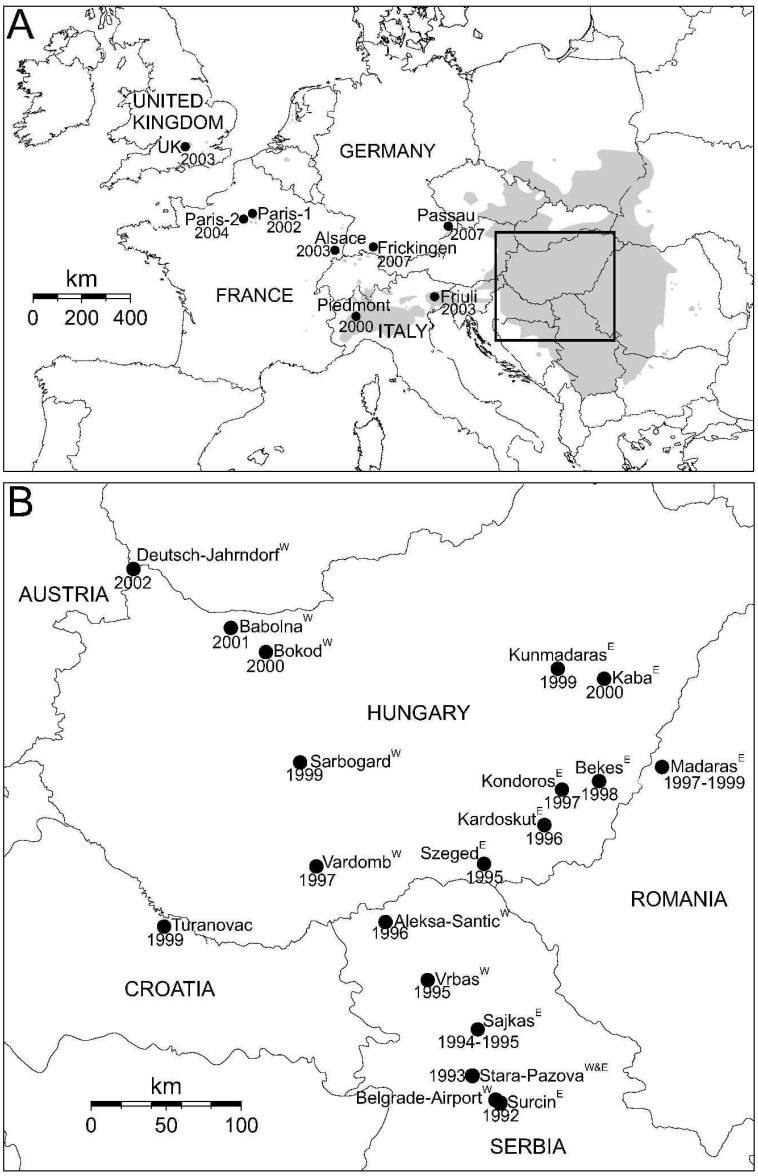


Figure 1: Location of sampling sites and geographic distribution of the western corn rootworm (WCR) in 2007, together with year of first observation: A) In Europe; distribution area, shown in gray, is defined as areas in which WCR has been observed for at least 1 year (Edwards and Kiss 2007). B) In the Central and South-Eastern European area; W: sample from the western sampling transect. E: sample from the eastern sampling transect. The names of the countries in which insects were collected are shown in capital letters
112x173mm (600 x 600 DPI)

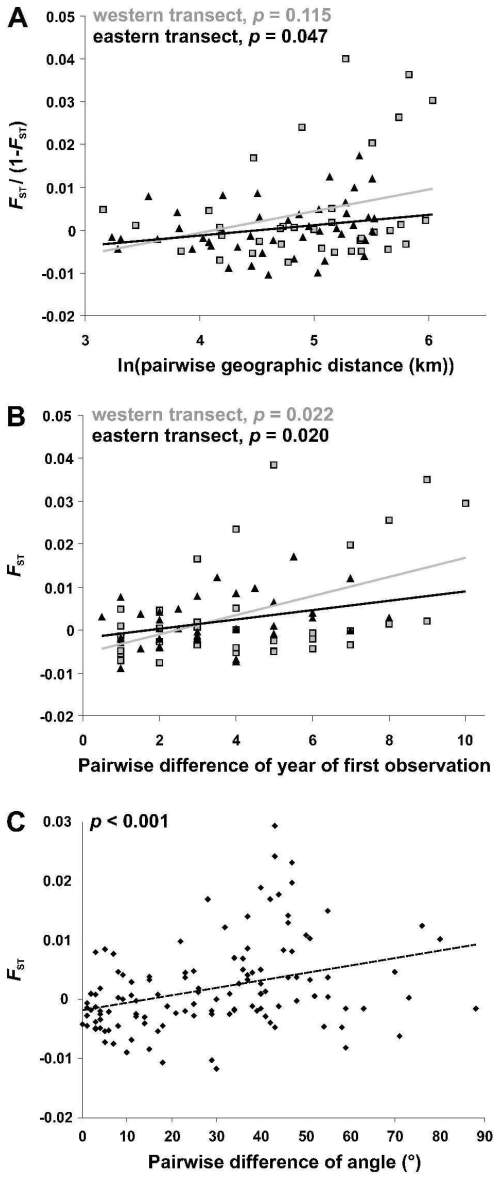


Figure 2: Patterns of genetic isolation by geographic (panel A), temporal (panel B) and angular distance (panel C) among Central and South-Eastern (CSE) European sample sites of the western corn rootworm. Linear regression lines and p-values for Mantel tests analyzing the correlation between genetic and geographic, temporal or angular distance are shown. In panels A and B, gray and black items correspond to the western and the eastern transects, respectively.
80x191mm (600 x 600 DPI)

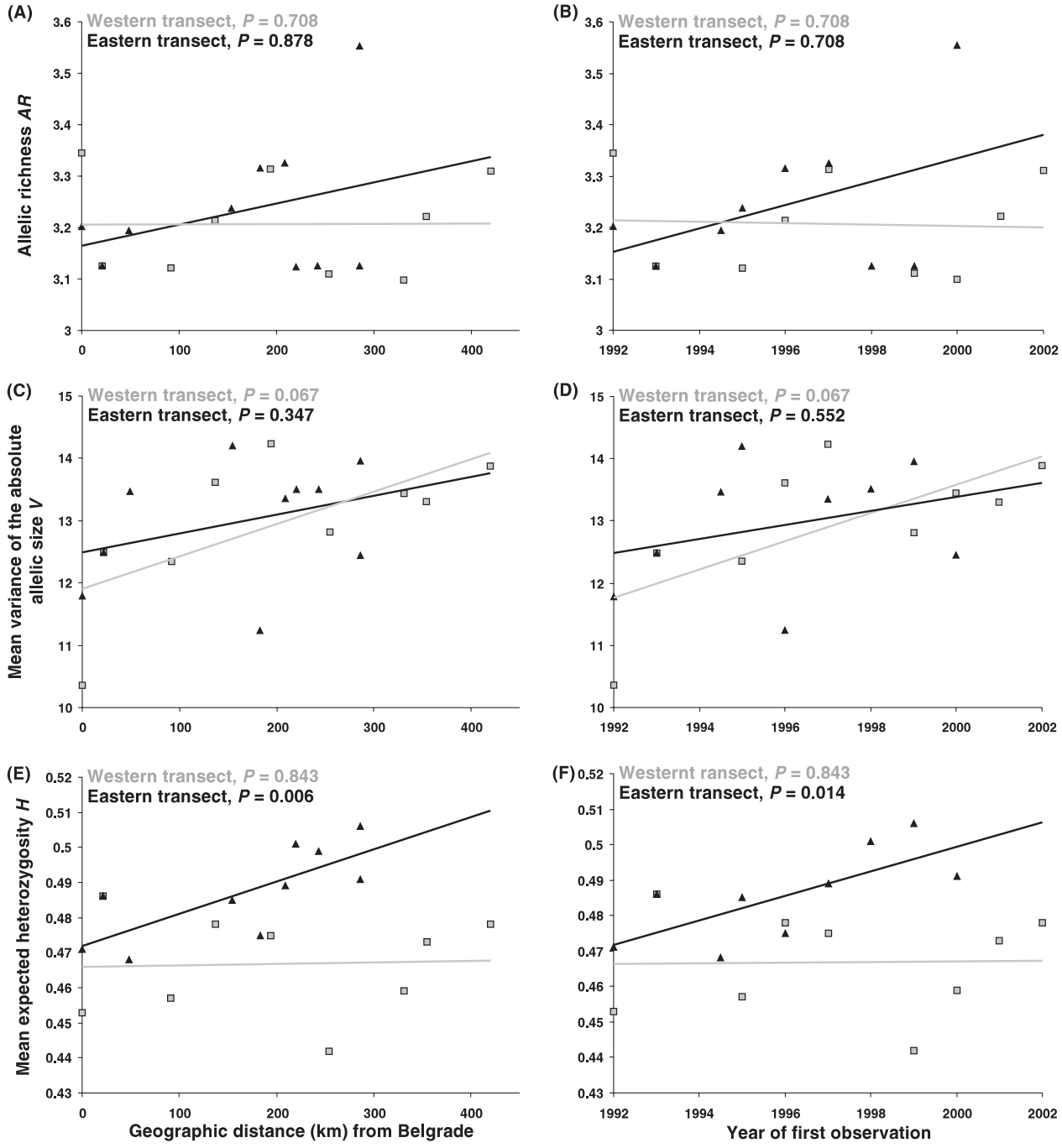


Figure 3 Correlations between genetic variation parameters within sampled sites [allelic richness *AR* based on minimum sample size (*N* = 28 in Belgrade-Airport and Kaba for loci DVV-D11 and DVV-ET1 respectively), using the rarefaction method (panels A and B), mean variance of the absolute allelic size *V* (panels C and D) and mean expected heterozygosity *H* (panels E and F)] and their geographic distance to the center of the Central and South-Eastern European area of expansion (panels A, C and E) or year of first observation (panels B, D and F). Linear regression lines and *P*-values for Spearman's rank correlation tests are shown. Gray and black items correspond to the western and eastern sampling transects, respectively.

of the Friuli outbreak. More precisely $L_{i\rightarrow s}$ values identified Turanovac and F_{ST} values identified Belgrade-Airport or Stara-Pazova, thus indicating that the center and western part of the CSE European area of expansion were the most likely source regions for the Friuli outbreak (Table 1).

Supplementary Figure legends

Figure S1.
Estimated number of populations from STRUCTURE (A) and GENELAND (B) analyses.
A) Mean (± SD) natural logarithm of the likelihood of the data [LnP(X|*K*)] over 10 STRUCTURE replicated runs for each value of the putative number of clusters (*K*). Black triangles and the solid line show the results of the analysis, performed wit STRUCTUREv.2.2, that does not use the sampling locations as prior information. Open squares and the dashed line show the results of the analysis, performed wit STRUCTUREv.2.3.1, that uses the sampling locations as prior information.
B) Posterior density distribution of the number of clusters (*K*) estimated from the highest-probability GENELAND run (among ten).

Figure S2
Estimated population structure from both STRUCTURE analyses for *K* = 2 to *K* = 5.
Each individual is represented by a thin horizontal line divided into *K* coloured segments that represent the individual's estimated membership fractions in *K* clusters. Black lines separate individuals from different sample localities. Each plot, produced with DISTRUCT (Rosenberg 2004), is based on the highest-probability run (among ten) at that value of *K*. ^{W}: sample from the western sampling transect. ^{E}: sample from the eastern sampling transect.

Table S1: Description of the western corn rootworm samples used in this study.

Geographic area	Sample name	Location	1st obs.	*N*	Collection year	GPS Coordinates	angle
North American range	Mexico	Durango, Mexico	<1940	14	2001	24°00'N 104°25'W	-
	Arizona	Willcox, Arizona	<1974	40	1998	32°03.83'N 109°43.86'W	-
	Texas	New Deal, Texas	<1980	51	2004	33°44.23'N 101°50.23'W	-
	Illinois	Champaign, Illinois	<1974	60	2003	40°06.98'N 88°14.62'W	-
	Pennsylvania	Bellefonte, Pennsylvania	<1985	62	2003	40°54.80'N 77°45.88'W	-
Western European isolated outbreaks	Piedmont	Oleggio, Italy	2000	40	2003	45°35.8'N 8°38.25'E	-
	Paris-1	Roissy Airport, France	2002	19	2003	49°0.28'N 2°33.97'E	-
	UK	Slough, United Kingdom	2003	36	2005	51°28.4'N 0° 32.1'W	-
	Friuli	Buttrio, Italy	2003	27	2003	46°00.683'N 13°20.083'E	-
	Alsace	Mulhouse Airport, France	2003	9	2003	47°36.0'N 7°29.0'E	-
	Paris-2	Pierrelaye, France	2004	74	2004	49°01'N 2°09'E	-
	Passau	Passau, Germany	2007	28	2007	48°34.8'N 13°26.4'E	-
	Frickingen	Frickingen, Germany	2007	32	2007	47°49'N 9°15'E	-
Central and South Eastern Europe area of expansion	Belgrade-Airport[W]	Belgrade Airport, Serbia	1992	38	2003	44°49.391'N 20°15.459'E	-
	Surcin[E]	Surcin, Serbia	1992	41	2003	44°48.334'N 20°17.669'E	-
	Stara-Pazova[W&E]	Stara Pazova, Serbia	1993	40	2003	44°.57.949'N 20°02.911'E	-
	Vrbas[W]	Vrbas, Serbia	1995	34	2003	45°33.406'N 19.39.848'E	60°
	Aleksa-Santic[W]	Aleksa Santic, Serbia	1996	40	2003	45°55.014'N 19°18.888'E	57°
	Vardomb[W]	Vardomb, Hungary	1997	39	2003	46°14.467'N 18°41.905'E	52°
	Sarbogard[W]	Sarbogard, Hungary	1999	39	2003	46°52'N 18°34'E	60°
	Bokod[W]	Bokod, Hungary	2000	35	2003	47°30.587'N 18°16.114'E	63°
	Babolna[W]	Babolna, Hungary	2001	35	2003	47°38.900'N 17°57.604'E	61°
	Deutsch-Jahrndorf[W]	Deutsch Jahrndorf, Austria	2002	38	2003	48°00'N 17° 06'E	56°
	Sajkas[E]	Sajkas, Serbia	1994-1995	48	2003	45°15.359'N 20°06.159'E	75°
	Szeged[E]	Szeged, Hungary	1995	31	2003	46°13.987'N 20°08.955	86°
	Kardoskut[E]	Kardoskut, Hungary	1996	34	2003	46°30.138'N 20°44.332'E	98°
	Kondoros[E]	Kondoros, Hungary	1997	37	2003	46°45.874'N 20°46.527'E	100°
	Bekes[E]	Bekes, Hungary	1998	38	2003	46°44.787'N 21°08.858'E	107°
	Kunmadaras[E]	Kunmadaras, Hungary	1999	35	2003	47°25'N 20°48'E	97°
	Madaras[E]	Madaras, Romania	1997-1999	39	2003	46°49.859'N 21°42.050'E	115°
	Kaba[E]	Kaba, Hungary	2000	32	2003	47°20.972'N 21°12.030'E	103°
	Turanovac	Turanovac, Croatia	1999	33	2003	45°53.039'N 17°23.311'E	27°

Note: [W]: sample from the western sampling transect. [E]: sample from the eastern sampling transect. 1st obs.: First observation year. *N*: number of individuals analysed per sample. angle: angle between the line of latitude passing through the origin of the outbreak (Belgrade, Serbia) and the line connecting Belgrade to the sample. -: not applicable.

Table S2: Statistics summarizing genetic variation within each western corn rootworm sampled population in America and Europe.

	A			
Sample name	**Direct Count**	***AR***	***H***	***V***
Pennsylvania	7.500 (5.043)	5.710 (3.265)	0.644 (0.175)	13.825 (17.518)
Belgrade-Airport[W]	3.375 (1.685)	3.203 (1.474)	0.453 (0.290)	10.366 (14.685)
Surcin[E]	3.250 (1.488)	3.123 (1.337)	0.471 (0.285)	11.784 (16.065)
Stara-Pazova[W&E]	3.125 (1.356)	3.116 (1.348)	0.486 (0.285)	12.485 (18.010)
Vrbas[W]	3.125 (1.356)	3.073 (1.312)	0.457 (0.315)	12.352 (18.14)
Aleksa-Santic[W]	3.250 (1.488)	3.146 (1.369)	0.478 (0.295)	13.608 (19.487)
Vardomb[W]	3.375 (1.685)	3.195 (1.446)	0.475 (0.303)	14.238 (20.929)
Sarbogard[W]	3.125 (1.356)	3.013 (1.337)	0.442 (0.315)	12.814 (18.238)
Bokod[W]	3.125 (1.356)	3.012 (1.226)	0.459 (0.295)	13.442 (18.957)
Babolna[W]	3.250 (1.488)	3.135 (1.349)	0.473 (0.290)	13.309 (18.521)
Deutsch-Jahrndorf [W]	3.375 (1.685)	3.199 (1.451)	0.478 (0.309)	13.884 (19.569)
Sajkas[E]	3.250 (1.488)	3.117 (1.342)	0.468 (0.285)	13.463 (19.214)
Szeged[E]	3.250 (1.581)	3.140 (1.423)	0.485 (0.310)	14.196 (20.003)
Kardoskut[E]	3.375 (1.685)	3.174 (1.404)	0.475 (0.304)	11.241 (15.900)
Kondoros[E]	3.375 (1.685)	3.219 (1.477)	0.489 (0.288)	13.355 (19.330)
Bekes[E]	3.125 (1.356)	3.105 (1.336)	0.501 (0.290)	13.509 (18.304)
Kunmadaras[E]	3.125 (1.356)	3.108 (1.332)	0.506 (0.282)	13.954 (19.517)
Madaras[E]	3.125 (1.356)	3.109 (1.335)	0.499 (0.297)	13.509 (18.853)
Kaba[E]	3.625 (1.685)	3.307 (1.411)	0.491 (0.269)	12.447 (16.339)
Turanovac	3.000 (1.195)	2.996 (1.192)	0.490 (0.284)	11.538 (16.143)
Passau	2.75 (1.282)	2.615 (1.187)	0.35 (0.261)	11.331 (18.410)
Frickingen	3.000 (1.195)	2.881 (1.258)	0.435 (0.337)	14.108 (21.037)
Friuli	1.75 (0.707)	1.744 (0.694)	0.293 (0.246)	3.833 (8.328)
Piedmont	4.25 (3.151)	/	0.42 (0.295)	10.869 (15.546)
Paris-1	3.75 (1.753)	/	0.51 (0.151)	9.431 (10.266)
Paris-2	3.75 (1.581)	/	0.534 (0.168)	10.479 (11.918)
Alsace	4.625 (1.996)	/	0.581 (0.25)	12.432 (16.433)
UK	5.75 (3.77)	/	0.612 (0.221)	12.55 (15.915)

Note: [W]: sample from the western sampling transect. [E]: sample from the eastern sampling transect. *S*tandard deviations across loci are shown in parentheses. ***A*:** average number of alleles per locus. *A* is given by direct counts (*DC*) and as allelic richness (*AR*). *AR* is based on minimum sample size (*N*=16 in Friuli for locus DVV-ET1). *H*: mean expected heterozygosity (Nei 1987). *V*: mean variance of the absolute allelic size. /: the statistics is not reported because the corresponding population was not concerned by statistical comparisons of genetic diversity. Significant deviation from Hardy Weinberg Equilibrium was observed for Stara-Pazova only ($p<0.001$).

Table S3: Weir & Cockerham's estimator of pairwise F_{ST} (1984) between western corn rootworm sampled populations in the Central and South-Eastern European expanding area.

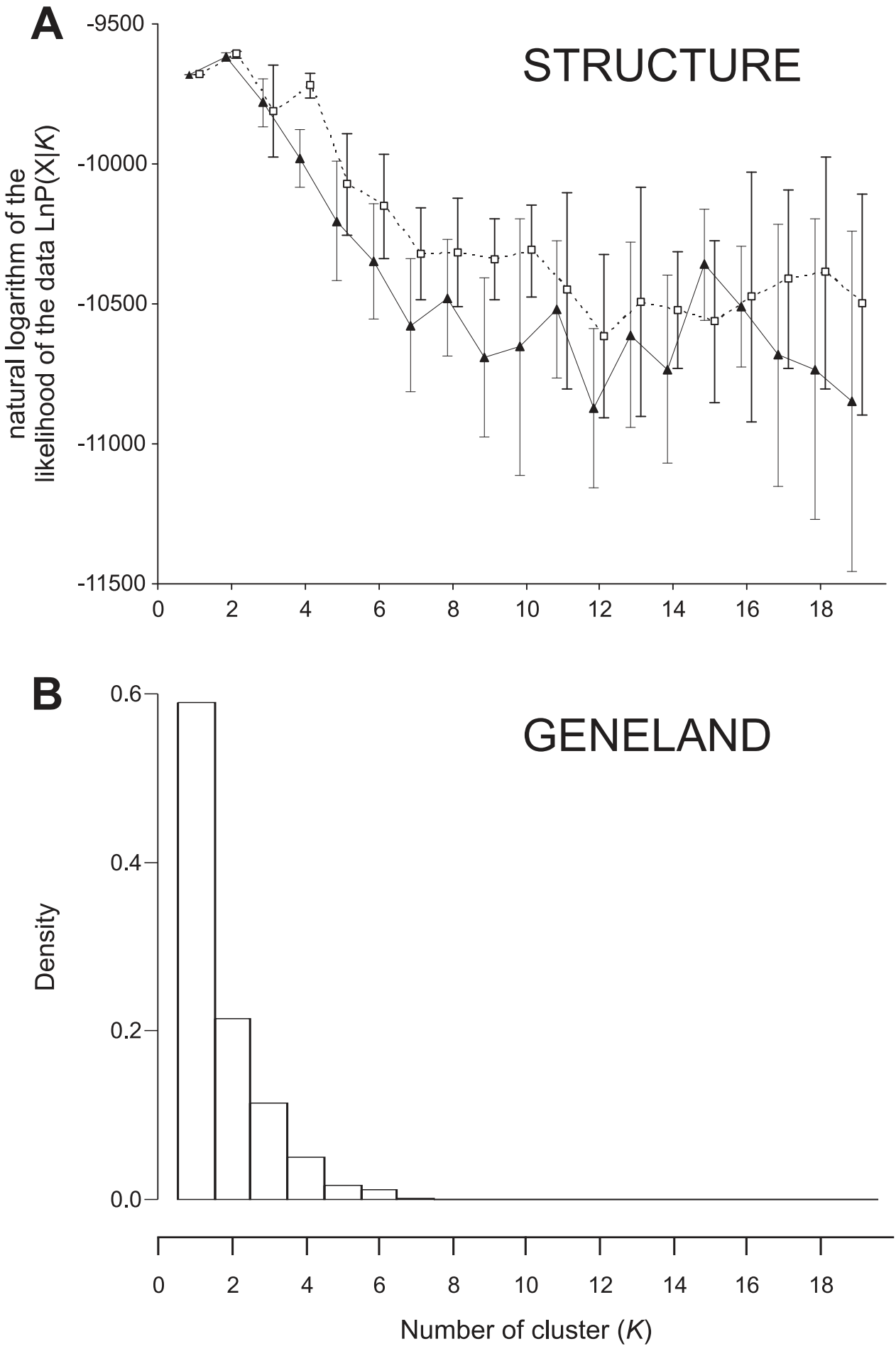
	Belgrade-Airport W	Surcin E	Stara-Pazova W&E	Vrbas W	Aleksa-Santic W	Vardomb W	Sarbogard W	Bokod W	Babolna W	Deutsch-Jahrndorf W	Sajkas E	Szeged E	Kardoskut E	Kondoros E	Bekes E	Kunmadaras E	Madaras E	Kaba E
Surcin^{E}	0.014*																	
Stara-Pazova$^{W\&E}$	0.005	-0.002																
Vrbas^{W}	0.017*	0.006	0.001															
Aleksa-Santic^{W}	**0.024***	0.002	-0.003	-0.005														
Vardomb^{W}	**0.039***	0.007*	0.005	0	-0.007													
Sarbogard^{W}	**0.020***	0.005*	-0.002	-0.004	0.001	0.005												
Bokod^{W}	**0.026***	0.005	0	-0.005	-0.005	0.001	-0.001											
Babolna^{W}	**0.035***	0	0.001	-0.001	-0.005	0	-0.003	0.001										
Deutsch-Jahrndorf^{W}	**0.030***	-0.001	0.002	-0.003	-0.004	-0.003	0.002	-0.008	-0.005									
Sajkas^{E}	**0.029***	0	-0.004	0.004	-0.004	0.005	0.003	-0.002	-0.003	-0.001								
Szeged^{E}	**0.028***	-0.002	0.002	0.001	-0.003	-0.002	0.002	-0.002	-0.003	-0.012	0.003							
Kardoskut^{E}	0.015	-0.007	0	-0.001	0.001	0.004	0.005	0.003	0.004	-0.004	0.004	-0.003						
Kondoros^{E}	**0.019***	-0.001	0	-0.002	-0.005	0	0.003	0.003	-0.002	-0.001	0.005	-0.004	-0.002					
Bekes^{E}	**0.024***	0.004	0.006	**0.020***	0.011*	**0.015***	**0.023***	**0.018***	0.013*	0.003	0.012*	-0.002	0.004	0.008*				
Kunmadaras^{E}	**0.017***	0	0.003	0.009	0.005	0.008	**0.014***	0.007	0.007	-0.003	0.010*	-0.007	-0.001	-0.004	-0.009			
Madaras^{E}	**0.018***	-0.006	-0.002	0	-0.005	-0.002	0.004	0.001	-0.005	-0.008	0.001	-0.010	-0.005	-0.008	-0.005	-0.011		
Kaba^{E}	**0.036***	0.003	0.012	**0.029***	0.014*	0.010*	**0.024***	0.019*	0.017*	0.008	**0.017***	0.001	0.009	0.008	-0.002	-0.002	0	
Turanovac	**0.023***	0	-0.002	-0.002	0	-0.001	0.001	0.005	-0.002	-0.002	0.004	-0.002	-0.006	0	0.010*	0.005	-0.002	0.012

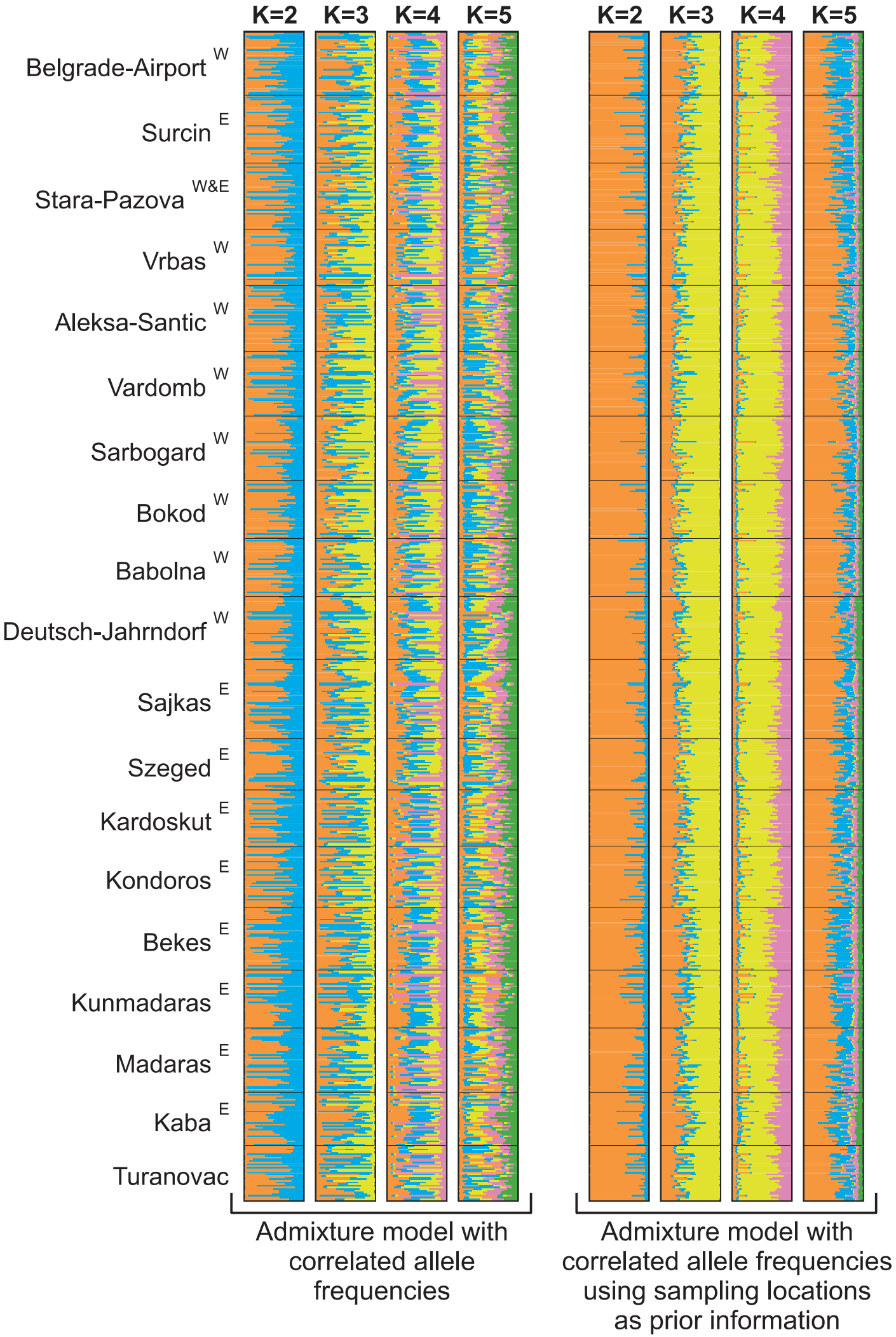
Note: Significant differentiation tests are indicated by an asterisk. Tests that remained significant after correction for multiple comparisons are indicated in bold. . ^{W}: sample from the western sampling transect. ^{E}: sample from the eastern sampling transect.

LITERATURE CITED IN THE SUPPLEMENTARY INFORMATION

Nei, M. 1987, Molecular Evolutionary Genetics. New York, Columbia University Press.

Rosenberg, N. A. 2004. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes 4:137-138.


A
STRUCTURE
natural logarithm of the likelihood of the data LnP(X|K)
-9500
-10000
-10500
-11000
-11500
0
2
4
6
8
10
12
14
16
18
B
GENELAND
Density
0.6
0.4
0.2
0.0
0
2
4
6
8
10
12
14
16
18
Number of cluster (K)


K=2
K=3
K=4
K=5
K=2
K=3
K=4
K=5
Belgrade-Airport W
Surcin E
Stara-Pazova W&E
Vrbas W
Aleksa-Santic W
Vardomb W
Sarbogard W
Bokod W
Babolna W
Deutsch-Jahrndorf W
Sajkas E
Szeged E
Kardoskut E
Kondoros E
Bekes E
Kunmadaras E
Madaras E
Kaba E
Turanovac
Admixture model with correlated allele frequencies
Admixture model with correlated allele frequencies using sampling locations as prior information