RESEARCH ARTICLE

A Bayesian approach to conservation genetics of Blanding's turtle (*Emys blandingii*) in Ontario, Canada

**Christina M. Davy · Pedro H. Bernardo ·
Robert W. Murphy**

Abstract Blanding's turtle, *Emys blandingii*, is a globally endangered species with a range centred on the Great Lakes of North America. Several disjunct populations also occur along the East Coast of North America. Previous studies suggest that the Great Lakes portion of the species' range exhibits panmixia. However, *E. blandingii* is restricted to relatively small populations in many areas around the Great Lakes. Therefore, panmixia across large geographic distances in this area is unlikely. Here, we apply Bayesian analyses of population structure to samples collected across southern Ontario ($N = 97$) to test a null hypothesis of panmixia and assess possible management units (MUs), and to estimate rates of gene flow across the study area. Sampled sites in Ontario represent a minimum of four distinct genetic clusters of *E. blandingii*, which we recommend should be considered as independent MUs. Preliminary evidence suggests that further structure may be present in less robustly sampled areas, which deserve further consideration. Genetic diversity at sampled sites is comparable to that reported for other freshwater turtles. Our comparison between this study and previous work confirms that genetic diversity in *E. blandingii* is reduced in disjunct eastern populations compared to populations centred on the Great Lakes. Genetic diversity in *E. blandingii* is not correlated with latitude, and instead may reflect post-glacial dispersal of this species from multiple Pleistocene glacial refugia.

Introduction

Blanding's turtle (*Emys blandingii*, Holbrook 1938) is a moderately sized freshwater species found in the northeastern United States and southern Canada. The main portion of its range is centred on the Great Lakes region. Small, disjunct populations occur in New York, Massachusetts, and Nova Scotia (Fig. 1; Ernst and Lovich 2009). Mean age of maturity in a well-studied Michigan population is 17.5 years, generation time is $\sim$37 years, and longevity exceeds 75 years (Congdon and van Loben Sels 1991; Congdon et al. 1993; Brecke and Moriarty 1989). One consequence of this life history is that populations are sensitive to any increase in the mortality rate of reproductive adults (Congdon et al. 1993), and even small increases in adult mortality (or collection of adults) can cause significant population declines. A number of factors including road mortality, illegal collection, and habitat degradation are currently causing such declines (COSEWIC 2005). Therefore, *E. blandingii* was recently up-listed from Least Concern to Endangered by the International Union for Conservation of Nature (IUCN; van Dijk and Rhodin 2011).
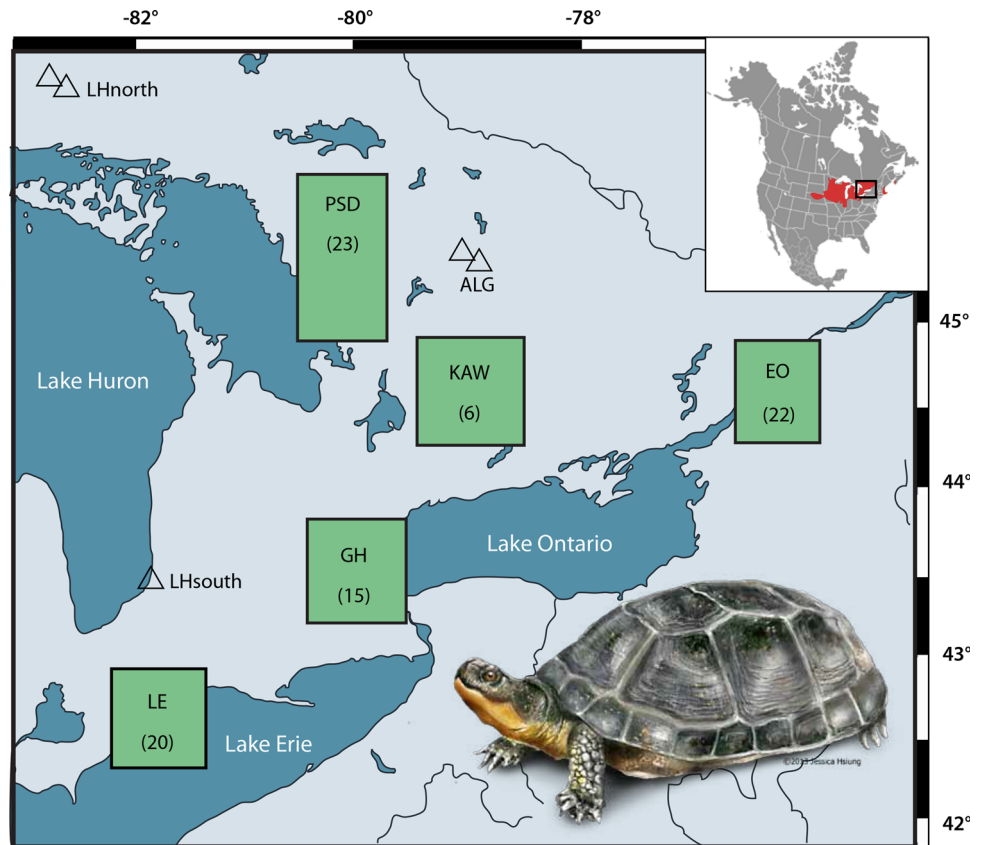
Using random amplified polymorphic DNA (RAPD) markers and microsatellites, Mockford et al. (1999, 2005, 2007) and Rubin et al. (2001) quantified genetic variation in *E. blandingii* across the species' range. Band-sharing analyses of RAPD data showed that the disjunct Nova

C. M. Davy · P. H. Bernardo · R. W. Murphy
Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, ON M5S 3B2, Canada

C. M. Davy (✉)
Natural Resources DNA Profiling & Forensic Centre, Trent University, DNA Building, 2140 East Bank Drive, Peterborough, ON K9J 7B8, Canada
e-mail: christinadavy@trentu.ca

R. W. Murphy
Department of Natural History, Royal Ontario Museum, 100 Queen's Park, Toronto, ON M5S 2C6, Canada

Scotian population differed genetically from central populations (Mockford et al. 1999; Rubin et al. 2001). Within Nova Scotia, analyses of microsatellite data based on F_{ST} values suggested significant differentiation among three subpopulations despite separation by <30 km. However, very little population structure was detected in the main portion of the range based on samples from Minnesota, Wisconsin, Illinois, Michigan, and Ontario (Mockford et al. 2007). Based on these data, Mockford et al. (2007) proposed that *E. blandingii* comprised three evolutionarily significant units (ESUs): (1) the population in Nova Scotia; (2) isolated populations in Massachusetts and New York; and (3) populations centred on the Great Lakes. The ESU concept does not apply to the legal conservation of turtles in either the USA or Canada, where protection is based on the concepts of distinctive population segments (Waples 1991; Rosen 2007) and designatable units (DUs; Green 2005), respectively. In Canada, the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) recognizes *E. blandingii* populations in Nova Scotia and around the Great Lakes as two DUs.

Recent studies highlight concerns with determining population structure based on F_{ST}. For example, Jost (2008) demonstrated that F_{ST} and related measures of diversity do not necessarily measure actual population differentiation. He proposed an alternative, more accurate measure (D_{est}). Jost (2008) also pointed out that statistical significance of F_{ST} was primarily a factor of sample size and may be biologically meaningless. Howes et al. (2009) summarized further concerns with analyses based on F_{ST}, including that assumptions of these analyses often were not met in natural populations (Whitlock and McCaughley 1999) and that F_{ST} does not reflect contemporary gene flow (Paetkau et al. 2004). F_{ST} and related measurements can provide information about historical migration rates, but they are not appropriate measures of population differentiation or structure (Jost 2009).

Howes et al. (2009) applied Bayesian methods that do not rely on F_{ST} to the data of Mockford et al. (2005), to study population connectivity in *E. blandingii* in Nova Scotia. They demonstrated moderate historic and current gene flow among the three known subpopulations, and clustered the two nearest subpopulations together as a genetically continuous unit. Bayesian analysis of three *E. blandingii* populations separated by <60 km in Illinois also detected no evidence of population structure (Banning-Anthonysamy 2012). Despite these results, panmixia across the main portion of the species' range is unlikely due to potential dispersal barriers such as long, high escarpments, or large, upland areas containing only unsuitable habitat. *E. blandingii* is dependent on wetland habitats but has high dispersal ability, and individuals may travel >10 km overland

between wetlands or to find a suitable nest site (Power 1989). Thus, population structure in this species may occur on a relatively large geographic scale (>100 km).

Although COSEWIC currently considers the "Great Lakes/St. Lawrence population," comprised of all *E. blandingii* in Ontario and Quebec, as a single unit for management and recovery purposes (COSEWIC 2005), the hypothesis of panmixia in *E. blandingii* in Ontario has not been tested. A large portion of the core range of *E. blandingii* occurs in Ontario, but previous studies included only 11 samples from St. Lawrence Islands National Park (Mockford et al. 2007), in south-eastern Ontario on the U.S./Canadian border. Presence-absence data show that the distribution of *E. blandingii* in Ontario is not continuous (Ontario Nature Herpetofaunal Atlas, http://www.ontarionature.org/protect/species/reptiles_and_amphibians/map_blandings_turtle.html). Gaps in occurrence records may reflect historic or current barriers to gene flow, and further suggest that populations may not be panmictic across the province. Isolated, single records of occurrence may also represent individuals translocated from other populations, or released illegal pets.

Emys blandingii reaches the northern limit of its range in Ontario. Founder effects during recolonization of northern areas following the retreat of ice sheets can cause populations reaching their northern limits in recently glaciated areas to exhibit lower genetic diversity than more southern populations (e.g. Galbraith 2008). As a result, Ontario populations of *E. blandingii* may show lower genetic variation than populations at lower latitudes. Mockford et al. (2007) documented reduced genetic diversity in isolated eastern populations of *E. blandingii* in Nova Scotia, presumably also caused by founder effects and subsequent isolation during the last Ice Age.

Here, we apply three Bayesian analyses and a principal coordinates analysis to investigate population structure in *E. blandingii* across >500 km in southern Ontario. We investigate the level of population structure and genetic diversity at sampled sites to test the hypothesis of panmixia in Ontario populations of *E. blandingii*. We also compare genetic variation (heterozygosity) among populations in Ontario and the populations studied by Mockford et al. (2007) to test two hypotheses: (a) that variation will be lower in disjunct eastern populations than in populations around the Great Lakes, as suggested by Mockford et al. (2005); and (b) that variation will decrease with proximity to the northern limit of the species' range.

Methods

We collected DNA from *E. blandingii* across southern Ontario between 2008 and 2011, with additional samples contributed by other researchers and government biologists (Fig. 1). Exact locations are withheld at the request of the Ontario Ministry of Natural Resources, to prevent an increase in collection pressure on populations. Turtles were captured by hand or in hoop traps baited with sardines at sites Lake Erie (LE), Golden Horseshoe (GH), Eastern Ontario (EO), Algonquin Park (ALG), Lake Huron South (LHsouth) and Lake Huron North (LHnorth). Blood was sampled by caudal venipuncture with a sterile syringe and blotted onto FTA cards (Whatman Inc., Clifton, NJ) for storage. Sampled individuals were released at their initial capture site. Extraction of DNA from FTA cards followed Smith and Burgoyne (2004).

In Parry Sound District (site PSD), muscle tissue was collected from road-killed individuals. At sites near the Kawartha Lakes (KAW), blood samples were taken from turtles injured on local highways and rehabilitated at the Kawartha Turtle Trauma Centre (Peterborough, Ontario). These blood samples were stored in heparin before analysis. Extraction of DNA from muscle and heparinised blood followed the phenol–chloroform procedure of Sambrook et al. (1989) and extracted DNA was cleaned with EtOH precipitation. Four additional blood samples were collected from captive *E. blandingii* at Scales Nature Park (Orillia, ON) that were from Ontario, but whose exact locations of origin were unknown.

We amplified samples at four microsatellite loci developed for *E. blandingii* (Eb09, Eb11, Eb17 and Eb19; Osentoski et al. 2002). These loci were used by Mockford et al. (2007) and we included them to facilitate direct comparison of within-population genetic diversity between studies. We amplified an additional 13 loci from *Glyptemys muhlenbergii* that cross-amplified in *E. blandingii* (GmuB08, GmuD16, GmuD21, GmuD28, GmuD55, GmuD70, GmuD87, GmuD88, GmuD89, GmuD90, GmuD93, GmuD107 and GmuD121; King and Julian 2004). Amplification and allele scoring followed Davy et al. (2012), using the locus-specific annealing temperatures listed in Table 1. Genotyping error was assessed by including positive controls with each PCR reaction and re-amplifying ∼6 % of the samples.

Evidence for null alleles and long allele drop-out was assessed with MICRO-CHECKER (van Oosterhout et al. 2004) using 1,000 iterations. Frequency of null alleles was calculated with the method of Brookfield (1996). We calculated the number of alleles per locus, observed heterozygosity (H_O) and expected heterozygosity (H_E) in GENALEX (Peakall and Smouse 2006). Allelic richness was rarefacted to correct for unequal sample sizes in HP-RARE (Kalinowski 2004, 2005). Linkage disequilibrium and deviations from Hardy–Weinberg equilibrium (HWE) were tested in GENEPOP v.4.0.1 (Raymond and Rousset 1995; Rousset 2008). Significance levels were corrected for multiple comparisons following Rice (1989).

We assessed genetic differentiation among sample sites by calculating absolute differentiation (D_{est}, Jost 2008) of

Table 1 Genetic diversity at 12 microsatellite loci for 97 *E. blandingii* from southern Ontario

	Temp.	N	k	H_O	H_E	PI	PISibs
GmuB08	58	96	7	0.406	0.403	0.378	0.643
GmuD16	56	95	13	0.811	0.818	0.055	0.357
GmuD21	58	95	3	0.253	0.230	0.617	0.789
GmuD28	61	97	16	0.845	0.862	0.034	0.328
GmuD55	56	96	13	0.792	0.816	0.056	0.356
GmuD87	54	88	11	0.659	0.723	0.123	0.419
GmuD88	58	96	11	0.792	0.848	0.040	0.336
GmuD93	58	95	4	0.421	0.552	0.294	0.548
GmuD107	58	96	11	0.771	0.854	0.038	0.332
GmuD121	58	94	8	0.766	0.725	0.103	0.413
Eb17	58	95	6	0.705	0.742	0.109	0.406
Eb19	58	92	4	0.478	0.704	0.140	0.433
Total		94.583 ± 0.701	8.917 ± 1.209	0.642 ± 0.057	0.690 ± 0.057	0.000	0.000

Temp. (optimal annealing temperature (°C) determined from temperature gradients of initial PCR reactions) sample size (N), allelic richness (k), observed and expected heterozygosity (H_O, H_E) and two measures of probability of identify (PI, PISibs) are shown for each locus. Total values show mean ± standard error for N, k, Ne, H_O and H_E, and PI/PIsibs values with all loci included

all sites with $N \geq 15$ in SMOGD (Crawford 2010). For purposes of comparison with previous studies we also used FSTAT (Goudet 1995) to calculate pairwise F_{ST} and assessed significance with 10,000 randomizations. Isolation by distance (IBD, a significant correlation between geographic and genetic distance, Wright 1943) was assessed using IBDWS (Jensen et al. 2005) with an input matrix of D_{est} and pairwise distances (km) between sites with $N \geq 15$. We estimated average historical number of migrants per generation between populations (Nm) based on F_{ST} following Barton and Slatkin (1986). We also estimated recent directional historic migration rates (m) between pairs of sampled sites using a Bayesian approach in BAYESASS v.1.3 (Wilson and Rannala 2003), considering this an estimate of migration rates across approximately the past 1–5 generations, or 185 years, following Chiucchi and Gibbs (2010). BAYESASS analysis involved 10^6 generations with a burn-in of 10 %, sampling the chain every 1,000 generations. We considered migration to be significantly directional if the 95 % confidence intervals of the estimates for each direction did not overlap, following Howes et al. (2009).

Within-population heterozygosity (H_O) from Ontario sites was compared to H_O values reported in Mockford et al. (2007) using an independent-samples *t* test in SPSS v.20.0 (IBM-SPSS, Chicago, IL, USA) after testing normality of the data. Comparisons were made for each of the two loci successfully amplified in both studies, and across all sampled loci. We compared heterozygosity between sampled sites in Ontario and the "western" sites from Mockford et al. (2007); (all sampled sites west of the Appalachian Mountains). We also compared sites east of the Appalachian Mountains to western populations, combining study sites from Ontario with western sites from

Mockford et al. (2007). Pearson's correlation coefficient was used to test for significant relationships between latitude and H_O among sites surrounding the Great Lakes.

Population structure was assessed by Bayesian inference in STRUCTURE v.2.3.4 (Pritchard et al. 2000) and TESS v.2.3.1 (Chen et al. 2007). STRUCTURE considered possible K values (number of genetically distinct populations) from one to six with eight independent runs at each value of K. We used the LOCPRIOR function to consider the sample site of each individual in the analysis. Each run involved 750,000 generations with a burn-in of 75,000 generations. The model assumed correlated allele frequencies (Falush et al. 2003) and historical admixture between populations (Pritchard et al. 2000).

We compiled the output of the STRUCTURE runs with STRUCTURE HARVESTER v.0.6.92 (Earl and vonHoldt 2012) and used two methods to estimate K, the most probable number of genetically distinct populations represented in the data. The increase in pr(X|K), the probability of the data given a particular value of K, typically plateaus at the most likely value of K (Pritchard et al. 2000). The ad hoc ΔK method (Evanno et al. 2005) implemented in STRUCTURE HARVESTER was also used to calculate the second-order rate of change in log likelihood between successive values of K, which typically peaks at the appropriate value of K. We used the Greedy and LargeKGreedy algorithms in CLUMPP v.1.1.2 (Jakobsson and Rosenberg 2007) to permute and combine results from independent runs. Genetic clusters identified by STRUCTURE were visualized with DISTRUCT v.1.1 (Rosenberg 2004).

TESS analyses assumed an admixture model (Durand et al. 2009) and included the specific geographic location (GPS coordinates) of each sampled individual. TESS considered possible K_{max} values (maximum possible number

of populations represented by the data) from two to eight, with ten independent runs at each K_{max}. Runs included 50,000 sweeps with a burn-in of 10,000 sweeps. Data were assessed to ensure convergence and individual assignments were visualized with CLUMPP and DISTRUCT. We considered the most likely K to occur at the point of inflection of the decreasing deviance information criterion (DIC) values and at which the number of distinct clusters stabilized (Chen et al. 2007).

Assignment tests were conducted in GENECLASS v.2.0 (Piry et al. 2004) using the Bayesian method of Rannala and Mountain (1997), with 100,000 iterations and a Type I error level of 0.05. This duplicates the analyses conducted by Howes et al. (2009), allowing a reasonable level of comparison between studies. Assignment tests considered only sampling areas with six or more samples. Individual samples from other sites and samples of unknown origin were then assessed by the program as "unknown", and assigned to the most similar sampling area. Population structure was also visualized with a principal coordinate analysis (PCoA) in GENALEX, based on D_{est} for sampled sites and on Nei's unbiased genetic distance for individuals.

Results

Loci Eb09, Eb11, GmuD70, GmuD89, and GmuD90 either did not amplify, or could not be scored consistently despite multiple adjustments of PCR conditions. Thus, 12 loci were used for analyses. In total, 116 samples were collected but several yielded degraded DNA and were successfully amplified at only 5 or 6 loci. These samples were excluded and a total of 97 individuals (91 individuals from known locations) were genotyped at >10 loci and included in the final analysis.

Duplicate amplification of positive controls yielded identical genotypes, and the influence of genotyping error on our dataset was therefore considered negligible. MICRO-CHECKER found evidence for potential null alleles at three loci (Eb19, GmuD93 and GmuD107). However, when the four largest samples were tested independently, the results were inconsistent among sites; only EO and GH showed evidence for potential null alleles, and only at locus Eb19.

Deviations from HWE were detected at locus Eb19 in PSD, GH, and EO, but not in KAW or LE. Evidence for LD was detected across the entire dataset between two pairs of loci: GmuD55–GmuD107 and GmuD28–GmuD107. However, LD was not detected when testing sampled areas independently and we therefore accepted the null hypothesis of linkage equilibrium. PI and PIsibs decreased to <0.01 with the inclusion of three and six loci, respectively. The 12 loci exhibited 3–16 alleles (mean 8.917, SD = 4.187). Observed heterozygosity (H_O) ranged from 0.253 at locus GmuD21 to 0.845 at locus GmuD28 (Table 1).

Summary statistics for all sampling sites are shown in Table 2. Pairwise values of D_{est} ranged from 0.010 to 0.156 (mean = 0.083, SD = 0.044, Table 3). Values of F_{ST} ranged from 0.039 to 0.099 (mean 0.072, SD = 0.021). Based on F_{ST}, Nm among sites averaged 2.432, and Nm between each pair of sites ranged from 0.095 (GH–PSD) to 3.380 (PSD–EO; Table 3). Recent gene flow (m) among sites estimated with BAYESASS averaged 0.0363. Pairwise m ranged from 0.011 (EO to KAW) to 0.084 (KAW to PSD; Table 3). Only one pair of sites (KAW and PSD) had significantly asymmetrical estimates of gene flow, inferred from non-overlapping confidence intervals: estimated migration from KAW to PSD over 14 times greater than migration from PSD to KAW. No significant IBD was detected among the four sites with $N > 12$ ($Z = 194.900$, $r = 0.233$, $p = 0.301$).

The first axis of the PCoA of sampling sites accounted for 60.19 % of total variation. This axis separated sites LE and GH from PSD, KAW, and EO (Fig. 2). When the PCoA was conducted at the individual level, individuals clustered by site. Overlap between the edges of these clusters indicated that differentiation in this dataset may be occurring along clines rather than along sharply defined boundaries.

GENECLASS assigned individuals from LE, GH, PSD, KAW, and EO to their area of origin with 69 % accuracy (Table 4). Samples from KAW were assigned to PSD ($N = 5$) or GH ($N = 1$). When KAW was removed from assignment tests, overall accuracy increased to 79 %. The two samples from the north shore of Lake Huron were assigned to PSD. The two samples from ALG and the sample from the south shore of Lake Huron were not assigned to any sampled clusters ($p < 0.01$).

The DIC in the TESS analysis decreased gradually from $K_{max} = 2$ with no clear point of inflection (mean $\Delta DIC = 54.2$, Fig. 3a). Individual q-matrices stabilized at $K_{max} = 2$. No clearly defined new clusters appeared at higher values of K_{max}, although potential admixture from a third population became apparent in site EO at $K_{max} = 3$. The first resolved population included LE, GH, and LHsouth (mean q = 0.969, SD = 0.094). The second population included all other samples (mean q = 0.709, SD = 0.377). One sample from PSD was assigned with approximately equal probability to both populations (0.493 vs. 0.507).

STRUCTURE resolved the same two populations as TESS at K = 2 (Fig. 3b). At K = 3, LE and LHsouth (mean q = 0.939, SD = 0.049) separated from population GH (mean q = 0.821, SD = 0.160) with evidence of admixture remaining between the two clusters. At K = 4, EO (mean q = 0.832, SD = 0.1) separated from a final population consisting of PSD, KAW, LHnorth and ALG (mean q = 0.776, SD = 0.071).

Heterozygosity data were normally distributed and Levene's test indicated equal variances ($F = 0.09$, $p = 0.927$).

Table 2 Number of alleles (number of private alleles in parentheses) and observed and expected heterozygosities (H_O and H_E) for 97 *E. blandingii* sampled across southern Ontario and genotyped at 12 microsatellite loci

		ALG	EO	LHnorth	PSD	GH	KAW	LE	LHsouth
GmuB08	Number of alleles	2	4 (1)	3 (1)	5 (1)	4	3	2	1
Estimated null allele	H_O	0.500	0.591	0.500	0.522	0.214	0.333	0.150	–
Frequency = 0.00	H_E	0.375	0.583	0.625	0.436	0.199	0.292	0.139	–
	N	2	22	2	23	14	6	20	1
GmuD16	Number of alleles	3	8	2	8	9 (1)	7	7	1
Estimated null allele	H_O	1.000	0.905	0.500	0.826	0.800	0.667	0.800	–
Frequency = 0.00	H_E	0.625	0.796	0.375	0.751	0.767	0.819	0.743	–
	N	2	21	2	23	15	6	20	1
GmuD21	Number of alleles	1	2	2	2	2	3 (1)	2	2
Estimated null allele	H_O	0.000	0.091	1.000	0.174	0.286	0.800	0.350	–
Frequency = 0.00	H_E	0.000	0.087	0.500	0.159	0.245	0.580	0.289	–
	N	2	22	2	23	14	5	20	1
GmuD28	Number of alleles	2	9 (1)	3	11	8	5	10	2
Estimated null allele	H_O	0.000	0.818	1.000	0.870	0.733	1.000	0.900	–
Frequency = 0.00	H_E	0.500	0.789	0.625	0.843	0.791	0.722	0.851	–
	N	2	22	2	23	15	6	20	1
GmuD55	Number of alleles	4	9	2	8	5	8 (2)	6	1
Estimated null allele	H_O	1.000	0.773	0.500	0.773	0.800	1.000	0.850	–
Frequency = 0.000	H_E	0.750	0.784	0.375	0.826	0.664	0.819	0.711	–
	N	2	22	2	22	15	6	20	1
GmuD87	Number of alleles	3	5 (1)	2	8 (3)	6	3	5 (1)	2
Estimated null allele	H_O	0.500	0.636	1.000	0.696	0.818	0.250	0.579	–
Frequency = 0.000	H_E	0.625	0.727	0.500	0.641	0.736	0.656	0.677	–
	N	2	22	2	23	11	4	19	1
GmuD88	Number of alleles	3	8 (1)	4	9	7	6	8	2
Estimated null allele	H_O	1.000	0.818	1.000	0.913	0.500	0.500	0.800	–
Frequency = 0.00	H_E	0.625	0.790	0.750	0.855	0.640	0.778	0.800	–
	N	2	22	2	23	14	6	20	1
GmuD93	Number of alleles	2	4 (1)	2	2	2	2	3	1
Estimated null allele	H_O	0.500	0.455	0.500	0.522	0.286	0.667	0.400	–
Frequency = 0.148	H_E	0.375	0.567	0.375	0.491	0.408	0.500	0.531	–
	N	2	22	2	23	14	6	20	1
GmuD107	Number of alleles	2	8	4	9	6	6	7	2
Estimated null allele	H_O	1.000	0.773	1.000	0.783	0.643	0.833	0.750	–
Frequency = 0.073	H_E	0.500	0.721	0.750	0.823	0.694	0.694	0.659	–
	N	2	22	2	23	14	6	20	1
GmuD121	Number of alleles	3	7	3	6	5	6	5	1
Estimated null allele	H_O	0.500	0.818	1.000	0.762	0.867	1.000	0.600	–
Frequency = 0.000	H_E	0.625	0.751	0.625	0.718	0.598	0.800	0.484	–
	N	2	22	2	21	15	5	20	1
Eb17	Number of alleles	2	4	2	5	5 (1)	3	5	1
Estimated null allele	H_O	0.500	0.591	0.500	0.636	0.733	0.600	0.950	–
Frequency = 0.000	H_E	0.375	0.699	0.375	0.636	0.709	0.460	0.696	–
	N	2	22	2	22	15	5	20	1
Eb19	Number of alleles	2	4	2	4	4	3	3	2
Estimated null allele	H_O	0.500	0.364	0.500	0.591	0.250	0.600	0.550	–
Frequency = 0.234	H_E	0.375	0.673	0.375	0.577	0.642	0.660	0.594	–
	N	2	22	2	22	12	5	20	1

Table 2 continued

	ALG	EO	LHnorth	PSD	GH	KAW	LE	LHsouth
Mean H_O	0.583	0.636	0.75	0.672	0.578	0.688	0.64	–
Mean HE	0.479	0.664	0.521	0.646	0.591	0.648	0.598	–
Ar	–	5.09	–	5.25	4.8	–	4.64	–
PAr	–	0.62	–	0.39	0.53	–	0.33	–

Loci Gmu—from King and Julian (2004). Loci Eb—from Osentoski et al. (2002). Acronyms for sampling areas are defined in Fig. 1. Estimated frequency of a null allele is based on analysis of the entire data set following Brookfield (1996). No loci showed consistent evidence for null alleles when sampling areas were analyzed independently

H_O observed heterozygosity, H_E expected heterozygosity, *Ar* allelic richness, *PAr* private allelic richness

Table 3 Genetic differentiation of *E. blandingii* among sites in Ontario with $N \geq 15$

Sampling site	Approximate distance (km)	D_{est}	F_{ST}	Nm, based on F_{ST}	Short-term gene flow (m, estimated in BAYESASS)	
					m1	m2
Lake Erie–Golden Horseshoe	151	0.0569	0.0664	1.35	0.0149 ± 0.0143	0.0481 ± 0.0282
Lake Erie–Parry Sound District	310	0.0623	0.0621	1.442	0.0416 ± 0.0379	0.0505 ± 0.037
Lake Erie–Eastern Ontario	516	0.0997	0.0888	1.029	0.0131 ± 0.0127	0.0323 ± 0.0247
Golden Horseshoe–Parry Sound District	266	0.1556	0.0955	0.952	0.0294 ± 0.0216	0.014 ± 0.0129
Golden Horseshoe–Eastern Ontario	367	0.1434	0.0992	1.013	0.0258 ± 0.0193	0.018 ± 0.0162
Parry Sound District–Eastern Ontario	337	0.0638	0.0397	3.380	0.0835 ± 0.0418	0.0589 ± 0.033
Kawartha Lakes–Lake Erie	364	–	–		0.0293 ± 0.031	0.0118 ± 0.0112
Kawartha Lakes–Golden Horseshoe	200	–	–		0.0271 ± 0.0253	0.0145 ± 0.0136
Kawartha Lakes–Parry Sound District	164	–	–		0.1631 ± 0.0509	0.0109 ± 0.0105
Kawartha Lakes–Eastern Ontario	195	–	–		0.0285 ± 0.026	0.0111 ± 0.0106

All F_{ST} values were significant ($p < 0.05$). Estimated migration rates include (1) Nm, the average number of historical migrants based on F_{ST} following Barton and Slatkin (1986), and (2) short-term gene flow (m; estimated mean ± standard deviation) between each set of sites (including KAW, with $N = 6$) was estimated in BAYESASS: m1 is estimated gene flow from the first site listed to the second, and m2 is the estimated migration rate in the opposite direction

Data from the two loci used both in this study and by Mockford et al. (2007); (Eb17 and Eb19) were combined for comparison. Observed heterozygosity in the Great Lakes portion of the species' range was significantly higher at locus Eb17 ($t = -3.621$, d.f. = 15, $p = 0.003$) but not at locus Eb19 ($t = -1.823$, d.f. = 15, $p = 0.088$). When mean heterozygosity across all loci used in both studies was compared, H_O was significantly higher in western populations ($t = -3.749$, d.f. = 15, $p = 0.002$) than in the disjunct eastern populations. No difference in heterozygosity occurred between the western populations sampled by Mockford et al. (2007) and the populations sampled in this study ($t = -0.413$, df = 10, $p = 0.688$). Latitude and H_O/H_E were not correlated (Pearson's correlation = 0.056, $N = 11$, $p = 0.869$).

Discussion

Our analyses identify a minimum of two genetic populations and four subpopulations of *E. blandingii* in southern Ontario, and refute the hypothesis of panmixia. Intensive urban development and expanding road networks make current migration between these four subpopulations nearly impossible, but our results indicate that historic migration rates were also low enough for genetic drift to cause differentiation among populations. As a result, assignment tests can identify individuals to their area of origin with relatively high accuracy, especially considering the small sample sizes available for this study. Diversity does not decrease with latitude in this species. Our data further support the finding of Mockford et al. (2007) that the isolated populations of *E. blandingii* east of the Appalachian Mountains show reduced diversity relative to western populations.

Management units (MUs) are typically delimited based on both genetic and demographic data (Moritz 1994; Palsbøll et al. 2006). We propose four tentative MUs for *E. blandingii* in Ontario: LE, GH, Georgian Bay-Parry Sound District, and EO. STRUCTURE analyses and assignment tests show that these units are genetically differentiated, and distribution data

Fig. 2 Principal coordinates analysis (PCoA) for 91 *E. blandingii* sampled from across southern Ontario based on 12 microsatellite loci. PCoA of sampling areas (**a**, **b**) supports the population structure inferred by Bayesian clustering analyses (Fig. 3). PCoA of individuals (**c**) shows that individuals largely cluster by sampling site (inferred population) in coordinate space
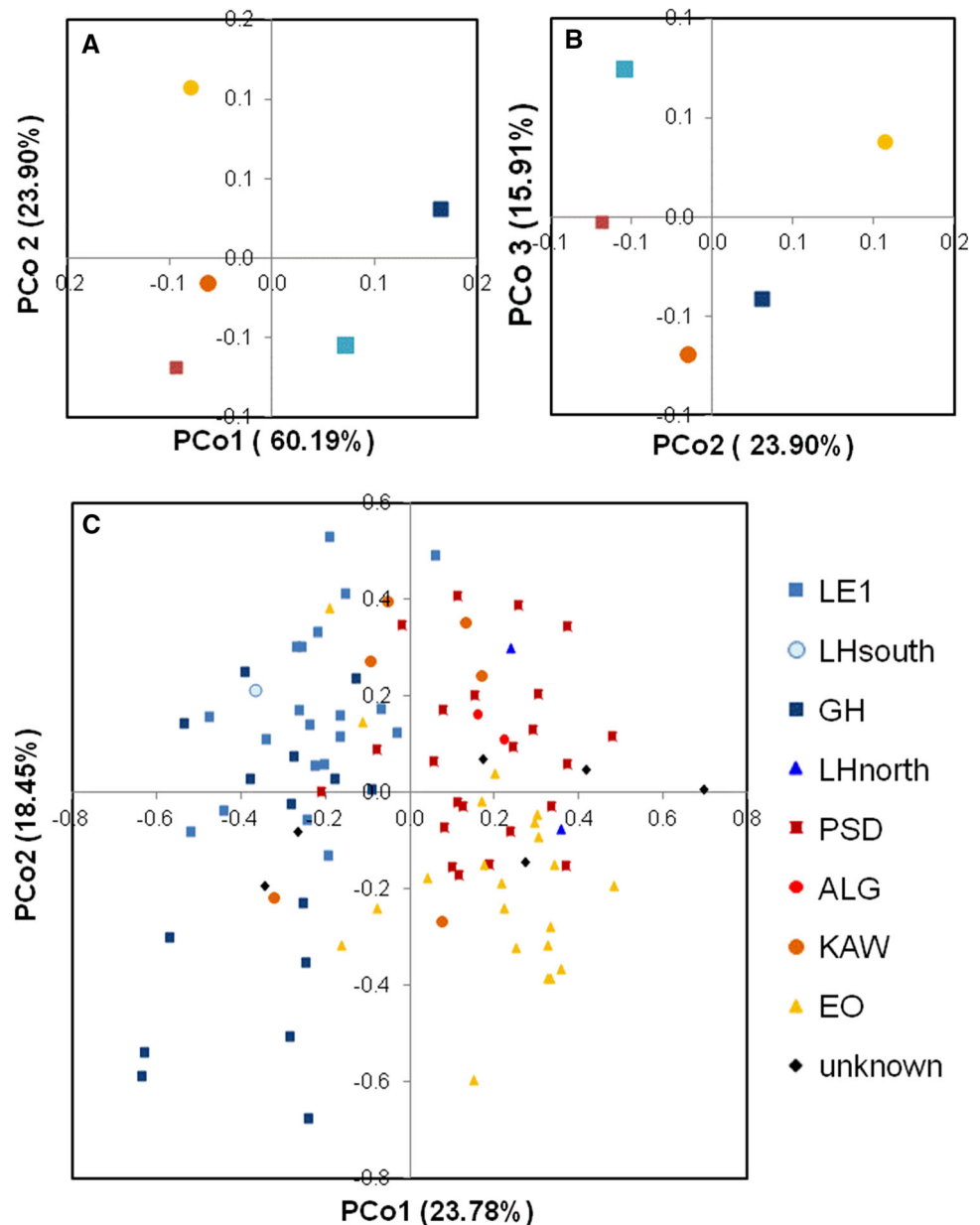

Table 4 Accuracy of Bayesian assignment test results in GENECLASS

Sampled population	Assigned population				
	LE	GH	PSD	KAW	EO
LE	*0.8*	*0*	0.2	0	0
GH	*0.07*	*0.67*	0.26	0	0
PSD	0	0	*0.74*	0.04	0.22
KAW	0	0.17	0.83	*0*	*0*
EO	0	0	0.27	0	*0.73*

Values represent the proportion of individuals from each sampled population assigned to each population. Values in bold indicate the proportion of individuals from each sampled population assigned correctly to their source population. Italics values indicate the two larger genetic clusters identified by TESS and STRUCTURE

(Ontario Nature Reptile and Amphibian Atlas) imply that these four proposed MUs may be isolated demographically. These four areas are unlikely to qualify as DUs under Canadian law (Green 2005) because we are not aware of empirical evidence that they experience significantly different risks of extinction. However, both the LE and GH MUs are increasingly isolated in declining patches of habitat, and DU designations may become appropriate in the future. The Georgian Bay-Parry Sound District MU is widely distributed and probably represents a large population. The same is likely true in EO, although it is currently not possible to estimate population size across these larger units. Our sampling was limited, and further sampling is required to clarify the exact boundaries of these MUs.

Fig. 3 Population structure inferred by Bayesian inference for 91 *E. blandingii* collected across southern Ontario. **a** TESS results showing decreasing deviance information criterion (DIC) with increasing values of K_{max}. **b** STRUCTURE results, mean estimated ln probability of the data [L(K)] for increasing values of K, and ΔK, the second order rate of change of L(K) following Evanno et al. (2005). Site abbreviations are explained in Fig. 1
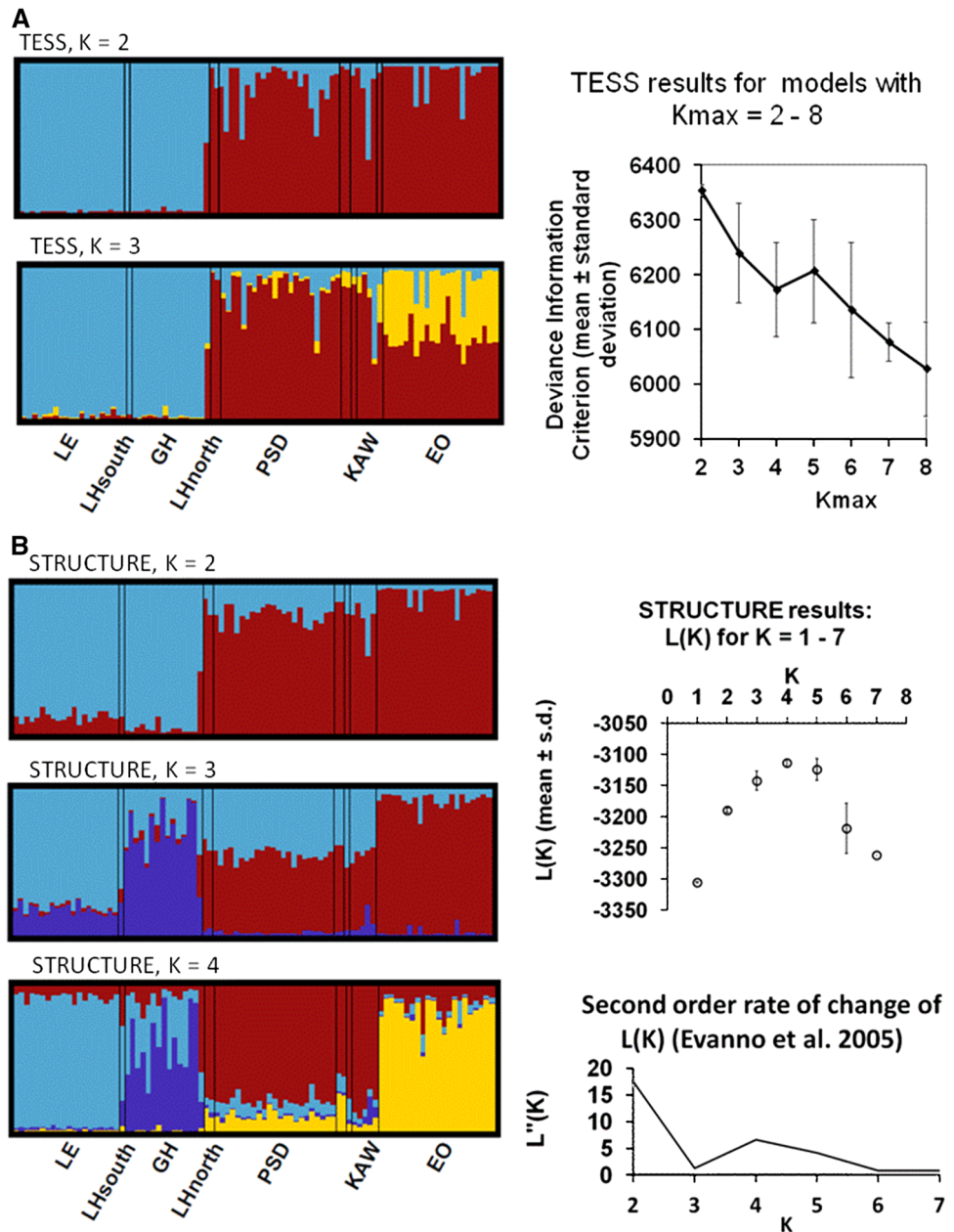
Assignment tests and clustering analyses provide preliminary evidence about connectivity between the proposed MUs and the four under-sampled sites (LHsouth, LHnorth, KAW and ALG). The northern extreme of the species' range (LHnorth) may be genetically continuous with the population in PSD. Clustering analyses and strong assignment of individuals from KAW to PSD implies similar genetic continuity between these two areas (Table 4; Fig. 3), which are connected by large remaining areas of suitable habitat. The samples from LHsouth and ALG were not strongly associated with any of the four proposed MUs and also represent isolated occurrences (Ontario Nature Reptile and Amphibian Atlas). None of the undersampled sites can be assumed to be part of the four tentative MUs until further evidence is available.

Bayesian assignment of individuals to Ontario populations had higher general accuracy (0.67–0.80) than that reported for the three Nova Scotia subpopulations (0.37–0.73; Howes et al. 2009), probably due to increased population structure and greater geographic distances in Ontario. This difference may also be an artefact of the different loci, and different number of loci, used in the two studies.

Interestingly, results from both STRUCTURE and TESS suggest possible past translocations of individuals between some subpopulations (Fig. 3; individuals with an ∼50 %

probability of membership to two populations may be first-generation offspring of migrants who mated with residents). This result may also be due to other causes (for example, natural migration from an unsampled population), but incidental and unrecorded translocations of individual turtles by members of the public occurs regularly in Ontario (F. Ross, pers. comm.; S. Gillingwater, pers. comm.; C. Davy, unpublished data). This is a separate issue from collection for the commercial pet trade, which also threatens *E. blandingii* (van Dijk and Rhodin 2011). There are no data on the frequency of these casual translocations and we cannot be certain that the observed admixture resulted from translocations, but the possibility should not be discounted.

The range of Nm values calculated for Ontario (0.952–3.380) is slightly lower than that reported from three sites in Nova Scotia (1.76–5.80; Mockford et al. 2005), but our estimates of recent gene flow (0.011–0.163) are comparable (0.012–0.16 in Nova Scotia; Howes et al. 2009). In both regions, no strong, overall directionality of gene flow among populations is apparent. Further comparative analyses of gene flow based on data from *E. blandingii* and other, co-occurring species could explicitly test hypotheses about potential barriers to gene flow on the current and historic landscape.

A negative correlation between genetic diversity and latitude is expected in turtles in post-glacial landscapes, due to founder effects during northward range expansions following the last ice age (Galbraith 2008). Diversity in *E. blandingii* does not support this prediction, possibly because this species has a compressed latitudinal range and may have used East–West re-colonization routes as well as moving northward. Fossil evidence places *E. blandingii* in southern Indiana 15–14 ka BP, and fossils are also known from Indiana and Michigan 6–4 ka BP (Holman 1992). Some populations might have used Pleistocene refugia in the southern Atlantic plain (Bleakney 1958). Others may have persisted near the Great Lakes throughout the Wisconsonian ice age, rapidly recolonizing the area from multiple directions as the ice sheets retreated (Holman 1992). Reduced genetic diversity (H_O) in the disjunct populations east of the Appalachian Mountains remains evident when our samples from Ontario are considered, further supporting the hypothesis that these represent relict populations stranded in glacial refugia during the last ice age (Mockford et al. 2005, 2007).

Our data suggest that loss of genetic diversity is not an immediate threat to *E. blandingii*. We detected private alleles at several sites, but no alleles are fixed at any loci. Heterozygosity is comparable to that reported for other populations of turtles (summarized by Vargas-Ramirez et al. 2012). Long generation times may have slowed the loss of genetic diversity in declining populations of *E. blandingii*. However, this species continues to experience

significant population fragmentation and decline across its range (van Dijk and Rhodin 2011). If this trend continues, the impacts of reduced genetic diversity will eventually become unavoidable.

Current conservation efforts for *E. blandingii* should mitigate demographic threats by reducing adult mortality (Congdon et al. 2008) and increasing recruitment of juveniles (and eventually mature adults) to populations. Increasing population size can prevent further loss of genetic diversity in threatened populations (Frankham et al. 2002), and simultaneously targets demographic threats to a population. Measures that can increase habitat and population connectivity should explicitly consider genetic structure. For example, anthropogenic features that fragment habitat and cause mortality (e.g. highways, urban development) also reduce gene flow among population fragments. This effect can be mitigated using tools such as wildlife underpasses, or corridors of suitable habitat. Conversely, actions such as translocations that involve moving individuals across the landscape should also explicitly consider genetic population structure. Mixing of genetic populations can have serious consequences for fitness if locally adapted genes or co-adapted gene complexes are disrupted (outbreeding depression, Templeton 1986). There is no empirical evidence to suggest local genetic adaptations in *E. blandingii*, but minimal genetic data exist for this species and the possibility should not be discounted.

We have clearly demonstrated that *E. blandingii* is not panmictic across Ontario, even on a relatively small scale (approximately 500 km between the most distant sites). Future analyses may reveal similar levels of population structure across the rest of the central range of *E. blandingii*, or even further ESUs beyond those already identified by Mockford et al. (2007). Such information would be crucial to effective conservation and recovery of this endangered species. Further analyses of current and historic gene flow in *E. blandingii* can also clarify the post-glacial colonization routes of turtles into Canada, and test the role of landscape features in shaping the current distribution of *E. blandingii* across its range.

References

Banning-Anthonysamy WJ (2012) Spatial ecology, habitat use, genetic diversity, and reproductive success: measures of connectivity of a sympatric freshwater turtle assemblage in a fragmented landscape. Dissertation, University of Illinois at Urbana-Champaign

Barton NH, Slatkin M (1986) A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. Heredity 56:409–416

Bleakney JS (1958) A zoogeographic study of the amphibians and reptiles of eastern Canada. Natl Mus Can Bull 155:1–119

Brecke BJ, Moriarty JJ (1989) Emydoidea blandingii (Blanding's turtle). Longevity. Herpetol Rev 20:53

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. Mol Ecol 5:453–455

Chen C, Durand E, Forbes F, François O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. Mol Ecol Notes 7:747–756

Chiucchi JE, Gibbs HL (2010) Similarity of contemporary and historical gene flow among highly fragmented populations of an endangered rattlesnake. Mol Ecol 19:5345–5358

Congdon JD, van Loben Sels RC (1991) Growth and body size in the Blanding's turtles (Emydoidea blandingii): relationships to reproduction. Can J Zool 69:239–245

Congdon JD, Dunham AE, van Loben Sels RC (1993) Delayed sexual maturity and demographics of Blanding's turtles (Emydoidea blandingii): implications for conservation and management of long-lived organisms. Conserv Biol 7:826–833

Congdon JD, Graham TE, Herman TB, Lang JW, Pappas MJ, Brecke BJ (2008) Emydoidea blandingii (Holbrook 1838)—Blanding's turtle. In: Rhodin AGJ, Pritchard PCH, van Dijk PP, Saumure RA, Buhlmann KA, Iverson JB (eds) Conservation biology of freshwater turtles and tortoises: a compilation project of the IUCN/SSC Tortoise and Freshwater Turtle Specialist Group. Chelon Res Monogr 5:015.1–015.12. doi:10.3854/crm.5.015.blandingii.v1.2008. http://www.iucn-tftsg.org/cbftt/. Accessed 4 Feb 2013

COSEWIC (2005) COSEWIC assessment and update status report on the Blanding's Turtle Emydoidea blandingii in Canada. Committee on the Status of Endangered Wildlife in Canadal, Ottawa, viii +40 pp. www.sararegistry.gc.ca/status/status_e.cfm. Accessed 10 Jan 2012

Crawford NG (2010) SMOGD: software for the measurement of genetic diversity. Mol Ecol Res 10:556–557

Davy CM, Leifso AE, Conflitti IM, Murphy RW (2012) Characterization of 10 novel microsatellite loci and cross-amplification of two loci in the snapping turtle (Chelydra serpentina). Conserv Genet Resour 4:695–698

Durand E, Jay F, Gaggiotti OE, François O (2009) Spatial inference of admixture proportions and secondary contact zones. Mol Biol Evol 26:1963–1973

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361

Ernst CH, Lovich JL (2009) Turtles of the United States and Canada, 2nd edn. Johns Hopkins University, Baltimore

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Falush D, Stevens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1615

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to conservation genetics. Cambridge University, Cambridge

Galbraith DA (2008) Population biology and population genetics. In: Steyermark AC, Finkler MS, Brooks RJ (eds) The biology of the snapping turtle. Johns Hopkins University, Baltimore, pp 168–180

Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Green DM (2005) Designatable units for status assessment of endangered species. Conserv Biol 19:1813–1820

Holman JA (1992) Late quaternary herpetofauna of the central Great Lakes region, U.S.A.: zoogeographical and paleoecological implications. Quat Sci Rev 11:345–351

Holbrook JE (1838) North American herpetology; or, a description of the reptiles inhabiting the United States. In: Dobson J (ed), vol 3, 1st edn. Philadelphia, pp 122

Howes BJ, Brown JW, Gibbs HL, Herman TB, Mockford SW, Prior KA, Weatherhead PJ (2009) Directional gene flow patterns in disjunct populations of the black ratsnake (Pantherophis obsoletus) and the Blanding's turtle (Emydoidea blandingii). Conserv Genet 10:407–417

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806

Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. BMC Genet 6:13. v.3.23. http://ibdws.sdsu.edu/. Accessed 10 Jan 2011

Jost L (2008) G_{ST} and its relatives do not measure differentiation. Mol Ecol 17:4015–4026

Jost L (2009) D vs. G_{ST}: response to Heller and Siegismund (2009) and Ryman and Leimar. Mol Ecol 18:2088–2091

Kalinowski ST (2004) Counting alleles with rarefaction: private alleles and hierarchical sampling designs. Conserv Genet 5:539–543

Kalinowski ST (2005) HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. Mol Ecol Notes 5:187–189

King TL, Julian SE (2004) Conservation of microsatellite DNA flanking sequences across 13 Emydid genera assayed with novel bog turtle (Glyptemys muhlenbergii) loci. Conserv Genet 5:719–725

Mockford SW, Snyder M, Herman TB (1999) A preliminary examination of genetic variation in a peripheral population of Blanding's turtle, Emydoidea blandingii. Mol Ecol 8:323–327

Mockford SW, McEachern L, Herman TB, Snyder M, Wright JM (2005) Population genetic structure of a disjunct population of Blanding's turtle (Emydoidea blandingii) in Nova Scotia, Canada. Biol Conserv 123:373–380

Mockford SW, Herman TB, Snyder M, Wright JM (2007) Conservation genetics of Blanding's turtle and its application in the identification of evolutionarily significant units. Conserv Genet 8:209–219

Moritz C (1994) Defining 'evolutionarily significant units' for conservation. Trends Ecol Evol 9:373–375

Osentoski MF, Mockford S, Wright JM, Snyder M, Herman TB, Hughes CR (2002) Isolation and characterization of microsatellite loci from the Blanding's turtle, Emydoidea blandingii. Mol Ecol Notes 2:147–149

Paetkau D, Slade R, Burden M, Estoup A (2004) Direct, real-time estimation of migration rate using assignment methods: a simulation-based exploration of accuracy and power. Mol Ecol 13:55–65

Palsbøll PJ, Bérubé M, Allendorf FW (2006) Identification of management units using population genetic data. Trends Ecol Evol 22:11–16

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol 6:288–295

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GeneClass2: a software for genetic assignment and first-generation migrant detection. J Hered 95:536–539

Power TD (1989) Seasonal movements and nesting ecology of a relict population of Blanding's turtles (Emydoidea blandingii) in Nova Scotia. M.Sc. Thesis, Acadia University, Wolfville

Pritchard JK, Stephens M, Donnelly PJ (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proc Natl Acad Sci USA 94:9197–9221

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225

Rosen T (2007) The endangered species act and the distinct population segment policy. Ursus 18:109–116

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes 4:137–138

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Mol Ecol Res 8:103–106

Rubin CS, Warner RE, Bouzat JL, Paige KN (2001) Population genetic structure of Blanding's turtles (Emydoidea blandingii) in an urban landscape. Biol Conserv 99:323–330

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning—a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, New York

Smith LM, Burgoyne LA (2004) Collecting, archiving and processing DNA from wildlife samples using FTA® databasing paper. BMC Ecol 4:4. http://www.biomedcentral.com/1472-6785/4/4. Accessed 12 Nov 2010

Templeton AR (1986) Coadaptation and outbreeding depression. In: Soulé M (ed) Conservation biology: the science of scarcity and diversity. Sinauer, Sunderland, pp 105–116

van Dijk PP, Rhodin AGJ (2011) Emydoidea blandingii. In: IUCN 2011. IUCN red list of threatened species. Version 2011.2. http://www.iucnredlist.org. Accessed 15 Dec 2011

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Vargas-Ramirez M, Stuckas H, Castaño-Mora OV, Fritz U (2012) Extremely low genetic diversity and weak population differentiation in the endangered Colombian river turtle Podocnemis lewyana (Testudines: podocnemididae). Conserv Genet 13:65–77

Waples RS (1991) Pacific salmon, Oncorhynchus spp. and the definition of "species" under the endangered species act. Mar Fish Rev 53:11–22

Whitlock MC, McCaughley DE (1999) Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4\ Nm + 1)$. Heredity 82:117–125

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163:1177–1191

Wright S (1943) Isolation by distance. Genetics 28:114–138