MOLECULAR ECOLOGY

Molecular Ecology (2013) doi: 10.1111/mec.12225

# No evidence for nuclear introgression despite complete mtDNA replacement in the Carpathian newt (*Lissotriton montandoni*)

P. ZIELIŃSKI,* K. NADACHOWSKA-BRZYSKA,*† B. WIELSTRA,‡ R. SZKOTAK,* S. D. COVACIU-MARCOV,§ D. COGĂLNICEANU¶ and W. BABIK*

*Institute of Environmental Sciences, Jagiellonian University, Gronostajowa 7, 30-387 Kraków, Poland, †Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-75236 Uppsala, Sweden, ‡Netherlands Biodiversity Center, P. O. Box 9517, 2300 RA Leiden, The Netherlands, §University of Oradea, Faculty of Sciences, Department of Biology, Universităţii str. 1, 410087 Oradea, Romania, ¶Faculty of Natural Sciences and Agricultural Sciences, University Ovidius Constanţa, Aleea Universităţii nr. 1, corp B, 900470 Constanţa, Romania

## Abstract

**Patterns of interspecific introgression may vary geographically, and the distribution of introgressed variants can yield insight into the historical dynamics of genetic interactions between hybridizing species. Urodele amphibians, often characterized by limited mobility, deep intraspecific genetic structuring and vulnerability to climatic changes, constitute suitable models for such historical inferences. Here, we combine an extensive survey of the mitochondrial (mtDNA) and nuclear (15 microsatellites) genomes in the Carpathian newt, *Lissotriton montandoni* (*Lm*) with species distribution modelling (SDM). Populations of the smooth newt, *L. vulgaris* (*Lv*) from the areas surrounding the *Lm* range were also sampled to test whether gene flow between these hybridizing species extends beyond the area of strict syntopy. The extent of introgression differs dramatically between the mitochondrial genome and the nuclear genome. While multiple, spatially and temporally distinct introgression events from *Lv* resulted in complete mtDNA replacement in *Lm*, there was little evidence of recent interspecific nuclear gene flow in the assayed markers. Microsatellite differentiation within *Lm* defines three units, probably derived from separate glacial refugia, located in the northern, eastern and southern part of the Carpathians. *In situ* survival and range fragmentation of *Lm* are supported by SDM, corroborating the role of the Carpathians as a major refugial area. Our results, in combination with previous reports of extensive introgression of the major histocompatibility complex (MHC) genes, emphasize the complexity of historical gene exchange between *Lm* and *Lv*.**

*Keywords*: genetic structure, introgression, *Lissotriton*, mtDNA replacement, natural hybridization, newt

*Received 29 August 2012; revision received 11 December 2012; accepted 13 December 2012*

## Introduction

The emergence of reproductive isolation is often a prolonged process (Avise *et al.* 1998; Coyne & Orr 2004). If incipient species are not completely isolated geographically, natural hybridization followed by introgression of some genomic fragments may occur (Mallet 2005). Patterns of introgression can yield insight into the history and dynamics of genetic interactions between incipient species (Neafsey *et al.* 2010; Teeter *et al.* 2010; Duvaux *et al.* 2011). Introgression may occur in various parts of the range and at various times. The geographic location of introgression can be inferred from the spatial distribution of introgressed alleles and its timing from interspecific genetic differentiation of introgressed genomic

Correspondence: W. Babik, Fax: +48 12 664 69 12;
E-mail: wieslaw.babik@uj.edu.pl

regions. The extent and outcome of introgression may vary along the contact zone. Due to intraspecific genetic differentiation, genomes of hybridizing species may interact differently in various geographic areas (Good *et al.* 2008; Nolte *et al.* 2009; Teeter *et al.* 2010), ecological conditions may affect the dynamics of hybridization (Vines *et al.* 2003), and the strength of prezygotic reproductive barriers may vary geographically (Zeng *et al.* 2011). Multiple instances of hybridization in various parts of the range may facilitate distinguishing of polymorphisms shared between species due to recent introgression from the retention of ancestral polymorphism (Mims *et al.* 2010; Nevado *et al.* 2011). Polymorphisms due to introgression would be expected to be most frequent in the areas of contact, whereas no such clustering is generally expected in the case of retention of ancestral polymorphism (Grant *et al.* 2005).

In some cases, introgression may be driven by positive selection (Whitney *et al.* 2006; Fitzpatrick *et al.* 2009; Song *et al.* 2011). However, introgression of neutral alleles is the expected outcome if species capable of hybridizing change their ranges (Currat *et al.* 2008). Such range changes have been common in the Northern Hemisphere in response to the climate oscillations during the Quaternary period (Hewitt 2000). Under the model of Currat *et al.* (2008), genes of the resident species would introgress into the invading species and parts of the genome, which experience less intraspecific gene flow are more likely to introgress (Petit & Excoffier 2009). This may partially explain the common observation of asymmetric mtDNA introgression (Petit & Excoffier 2009), although both asymmetric mating preferences and various forms of selection against hybrids may contribute as well (Wirtz 1999; Ballard & Whitlock 2004; Chan & Levin 2005; Toews & Brelsford 2012).

Amphibians, especially urodeles, may provide particularly rich information about patterns of historical gene exchange based on the geographic distribution of introgressed genetic variation. This is because urodeles generally combine a relatively low mobility (Beebee 1996; Smith & Green 2005) with species longevity (Vences & Wake 2007). As a consequence, salamanders typically express deep intraspecific genetic structure (Vences & Wake 2007). These characteristics favour the preservation of signals of geographically and temporally distinct introgression events. Furthermore, the distribution of amphibians is tightly linked to environmental conditions (Buckley & Jetz 2008; Vieites *et al.* 2009); hence, their ranges are strongly affected by climatic changes.

The Carpathian newt (*Lissotriton montandoni*, *Lm*) provides a striking example of a species that may have had its mtDNA totally replaced by mtDNA of its congener, the smooth newt (*L. vulgaris*, *Lv*; Babik *et al.* 2005). This mtDNA replacement did not originate from a single, spatially and temporally restricted introgression event. Events must have been separated in space as mtDNA derived from distinct *Lv* mtDNA lineages is present in various parts of the *Lm* range, which is fully surrounded by the *Lv* range. Introgressions must also have been separated in time as the level of interspecific sequence divergence as well as patterns of interspecific haplotype sharing differ between introgressed mtDNA lineages (Babik *et al.* 2005). In the northern part of its distribution, *Lm* has a specific mtDNA lineage, which is, however, deeply nested in the *L. vulgaris* mtDNA tree, pointing to a relatively old genetic exchange. In the remaining part of the *Lm* distribution, both species share mtDNA lineages and even haplotypes, which signals more recent or ongoing introgression. These distinct events of mtDNA introgression point to a complex history of interactions between the two species. The range of *Lm* is confined to the Carpathians (Fig. 1), and the Carpathian region is increasingly recognized as an important refugium for temperate organisms (Sommer & Nadachowski 2006; Provan & Bennett 2008; Ronikier 2011). Therefore, it is likely that *Lm* survived the Quaternary glaciations within its present range. Range fragmentation into multiple glacial refugia would facilitate distinct, locally acquired mtDNA lineages to reach fixation in different parts of the range producing the

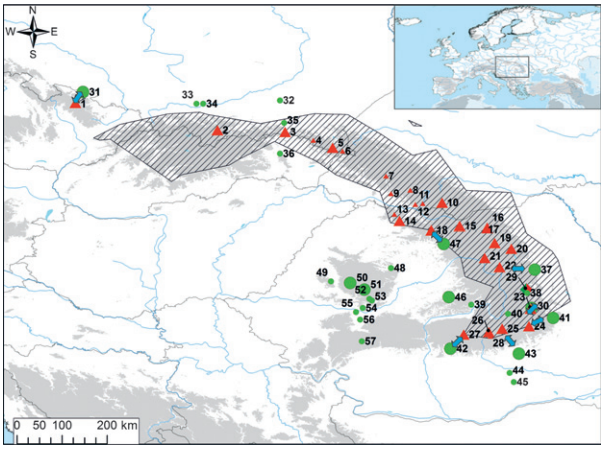


**Fig. 1** The distribution of sampling localities (details in Table 1). Red triangles—*Lissotriton montandoni* (*Lm*); green circles—*L. vulgaris* (*Lv*); green-red diamond (nr 29)—locality where both species co-occur; small symbols—localities reported by Babik *et al.* (2005) for which only mtDNA data are available; large symbols—localities for which both mtDNA and microsatellites or only microsatellites (nr 5 and 29) were studied. The distribution of *L. montandoni* (Zavadil *et al.* 2003 and own unpublished data) is hatched. Areas above 500 m a.s.l. are shaded. Blue arrows indicate the seven population pairs used to test whether pairs of neighbouring *Lm–Lv* populations were less differentiated than random *Lm–Lv* pairs.

observed mtDNA pattern. It should be emphasized, however, that data regarding mtDNA variation have been obtained as a part of larger phylogeographic study, which left the problem of genetic structuring and introgression in *Lm* mtDNA relatively underexplored. For instance, it remains unclear whether mtDNA replacement has indeed been complete, what the detailed distribution of mtDNA lineages in *Lm* is, and how the sequence variation in introgressed mtDNA lineages compares to variation in the respective *Lv* mtDNA lineages. The present study fills these gaps.

A prerequisite for an understanding of the significance of hybridization in this system is assessing to what extent the extensive introgression of mtDNA from *Lv* into *Lm* is mirrored by nuclear genes. Previous studies provided interesting but insufficient information in this respect. First, a fine-scale analysis of a small portion of the hybrid zone, utilizing a few preselected, locally diagnostic or almost diagnostic markers, demonstrated restricted nuclear gene flow (Babik *et al.* 2003). This study was performed in the northern part of *Lm* range, where the two species currently maintain a high degree of distinctiveness in mtDNA (Babik *et al.* 2003, 2005). No comparable data are available from the southern part of *Lm* distribution, where mtDNA introgression appears more recent or ongoing (Babik *et al.* 2005). The differences in the extent and timing of mtDNA introgression between various parts of *Lm* range indicate that differences in the strength of nuclear introgression are also likely to be present and need to be tested at a larger spatial scale, with an unbiased selection of markers. Second, a study examining variation in major histocompatibility complex class II (MHC II) genes in *Lv* and *Lm* found that these genes introgress readily (Nadachowska-Brzyska *et al.* 2012). Thus, various parts of the nuclear genome may differ in their potential for introgression.

The present study assesses, for the first time, the overall extent of nuclear introgression between *Lm* and *Lv* and tests for geographic variation in introgression patterns. We focus on recent and ongoing introgression, as these processes are most likely to leave detectable signatures in the regional patterns of interspecific differentiation. Such signatures may be discovered by examining *Lm* populations across its geographic range and *Lv* populations from surrounding areas. Availability of genetic markers with high per locus information content, such as microsatellites, and broad geographic sampling are instrumental for the success of such an approach.

Next to molecular data, species distribution modelling, the approximation of the ecological requirements of a species based on the range of environmental conditions it experiences at known localities, can provide insight into past range dynamics, especially when models are projected on environmental reconstructions of the past (Svenning *et al.* 2011). By comparing current and past range predictions, distribution shifts and demographic dynamics can be inferred. Insight into the biogeographical history of *Lm*, especially into its response to glacial cycles provided by the synthesis of genetic and species distribution maps, could inform about the dynamics of genetic interaction with *Lv* and thus introgression.

We present an extensive, range-wide survey of variation in mtDNA and nuclear DNA (microsatellites) for *Lm*, including also *Lv* populations from the areas surrounding the *Lm* range. We also provide species distribution models of *Lm* at the Last Glacial Maximum (LGM). This allowed us (i) to test whether nuclear gene flow has been occurring recently along the contact zone between *Lm* and *Lv*, (ii) to test whether patterns of nuclear introgression vary geographically, (iii) to determine intraspecific genetic structuring in *Lm* and test whether patterns of differentiation and variation in mtDNA are paralleled by nuclear DNA, and (iv) to explore whether the geographic distribution of mtDNA and nuclear DNA lends support for a scenario of repeated mtDNA introgression due to multiple refugia located within the present range of *Lm*.

## Materials and methods

### *Samples*

Adult newts were sampled by dip-netting during the 2009–2011 breeding seasons. Throughout the text, we use terms population and locality interchangeably to refer to a particular water body or an assemblage of water bodies located within a short distance. Animals were released immediately after tailtips were collected. Tissue samples were stored in 95% ethanol until DNA extraction. Altogether, we sampled 29 populations (742 individuals): 18 populations of *Lm* (502 individuals); 10 of *Lv* [192 individuals, including four populations of *L. vulgaris ampelensis* (*Lva*) and six populations of *L. v. vulgaris* (*Lvv*)]; and one population where we found both *Lm* and *Lv* (Fig. 1; Table 1). Samples of *Lm* covered the entire species range, with special emphasis on the Eastern Carpathians in Romania, where we expected to find the highest genetic diversity and where the evidence for recent mtDNA gene flow has been the strongest, based on an earlier phylogeographic study (Babik *et al.* 2005). Populations of *Lv* were sampled in the areas surrounding the *Lm* range to allow testing for the interspecific gene flow at a medium spatial scale, that is, extending beyond the area of strict syntopy.

**Table 1** Sampling sites. For mtDNA, both the total number of available sequences and the number of sequences obtained in the present study (new) are given

| Number | Locality | Country | *N* microsatellites | *N* mtDNA (new) | Species | Latitude (N) | Longitude (E) |
|---|---|---|---|---|---|---|---|
| 1 | Karlova Studánka | CZ | 22 | 8 (–) | *Lm* | 50°04′ | 17°18′ |
| 2 | Łopuszna | PL | 22 | 6 (5) | *Lm* | 49°31′ | 20°08′ |
| 3 | Krempna | PL | 16 | 5 (5) | *Lm* | 49°29′ | 21°29′ |
| 4 | Komańcza | PL | — | 2 (–) | *Lm* | 49°19′ | 22°03′ |
| 5 | Smerek | PL | 21 | – (–) | *Lm* | 49°10′ | 22°26′ |
| 6 | Ustrzyki Górne | PL | — | 5 (–) | *Lm* | 49°06′ | 22°38′ |
| 7 | Maidan | UA | — | 4 (–) | *Lm* | 48°36′ | 23°30′ |
| 8 | Ust′ Chorna | UA | — | 2 (–) | *Lm* | 48°19′ | 23°59′ |
| 9 | Mala Ugolka | UA | — | 3 (–) | *Lm* | 48°15′ | 23°36′ |
| 10 | Dzembronia | UA | 51 | 7 (7) | *Lm* | 48°04′ | 24°37′ |
| 11 | Rakhiv | UA | — | 1 (–) | *Lm* | 48°03′ | 24°14′ |
| 12 | Kobyla Hora | UA | — | 4 (–) | *Lm* | 48°02′ | 24°05′ |
| 13 | Izvoarele | RO | — | 1 (–) | *Lm* | 47°50′ | 23°40′ |
| 14 | Pasul Gutâi | RO | 21 | 4 (4) | *Lm* | 47°42′ | 23°46′ |
| 15 | Şesuri | RO | 12 | 5 (5) | *Lm* | 47°36′ | 24°58′ |
| 16 | Pasul Pascanu | RO | 35 | 5 (5) | *Lm* | 47°34′ | 25°31′ |
| 17 | Sadova | RO | — | 1 (–) | *Lm* | 47°32′ | 25°29′ |
| 18 | Romuli | RO | 20 | 5 (5) | *Lm* | 47°31′ | 24°24′ |
| 19 | Holda | RO | 41 | 5 (5) | *Lm* | 47°16′ | 25°40′ |
| 20 | Petru Vodă | RO | 36 | 5 (5) | *Lm* | 47°09′ | 26°00′ |
| 21 | Secu | RO | 21 | 5 (5) | *Lm* | 46°58′ | 25°28′ |
| 22 | Lacu Roşu | RO | 36 | 5 (5) | *Lm* | 46°47′ | 25°46′ |
| 23 | Pasul Musat 1 | RO | 25 | 6 (6) | *Lm* | 45°57′ | 26°23′ |
| 24 | Penteleu | RO | 24 | 3 (3) | *Lm* | 45°36′ | 26°21′ |
| 25 | Săcele | RO | 22 | 5 (5) | *Lm* | 45°33′ | 25°49′ |
| 26 | Predeal | RO | 50 | 13 (6) | *Lm* | 45°29′ | 25°33′ |
| 27 | Voina | RO | 27 | 7 (7) | *Lm* | 45°26′ | 25°03′ |
| 28 | Baiu | RO | — | 4 (–) | *Lm* | 45°24′ | 25°35′ |
| 29 | Valea Uzului | RO | 48 | – (–) | *Lm&Lv* | 46°20′ | 26°18′ |
| 30 | Pasul Musat 2 | RO | 20 | 6 (6) | *Lvv* | 45°58′ | 26°24′ |
| 31 | Pokrzywna | PL | 16 | 5 (5) | *Lvv* | 50°17′ | 17°27′ |
| 32 | Cieszęciny | PL | — | 1 (–) | *Lvv* | 50°07′ | 21°23′ |
| 33 | Mników | PL | — | 3 (–) | *Lvv* | 50°03′ | 19°43′ |
| 34 | Przegorzały | PL | — | 5 (–) | *Lvv* | 50°03′ | 19°51′ |
| 35 | Załęże | PL | — | 1 (–) | *Lvv* | 49°40′ | 21°28′ |
| 36 | Lipníky | SK | — | 4 (–) | *Lvv* | 49°03′ | 21°23′ |
| 37 | Tazlău | RO | 13 | 5 (5) | *Lvv* | 46°44′ | 26°28′ |
| 38 | Crăcurele | RO | — | 3 (–) | *Lvv* | 46°16′ | 26°19′ |
| 39 | Rupea | RO | — | 5 (–) | *Lvv* | 46°02′ | 25°12′ |
| 40 | Reci | RO | — | 4 (–) | *Lvv* | 45°51′ | 25°56′ |
| 41 | Budeni | RO | 16 | 6 (6) | *Lvv* | 45°46′ | 26°50′ |
| 42 | Vîlsăneşti | RO | 22 | 5 (5) | *Lvv* | 45°09′ | 24°47′ |
| 43 | Plopu | RO | 20 | 5 (5) | *Lvv* | 45°03′ | 26°09′ |
| 44 | Scroviştea | RO | — | 3 (–) | *Lvv* | 44°40′ | 25°58′ |
| 45 | Băneasa | RO | — | 5 (–) | *Lvv* | 44°29′ | 26°03′ |
| 46 | Sighişoara | RO | 25 | 8 (8) | *Lva* | 46°11′ | 24°45′ |
| 47 | Strâmba | RO | 23 | 6 (6) | *Lva* | 47°15′ | 24°39′ |
| 48 | Cluj | RO | — | 5 (–) | *Lva* | 46°46′ | 23°36′ |
| 49 | Briheni | RO | — | 2 (–) | *Lva* | 46°30′ | 22°24′ |
| 50 | Arieşeni | RO | 11 | 5 (5) | *Lva* | 46°28′ | 22°47′ |
| 51 | Cărpiniş | RO | 26 | 9 (5) | *Lva* | 46°20′ | 23°04′ |
| 52 | Izvoru Ampoiului | RO | — | 3 (–) | *Lva* | 46°09′ | 23°10′ |
| 53 | Zlatna | RO | — | 4 (–) | *Lva* | 46°07′ | 23°13′ |
| 54 | Săcărâmb | RO | — | 3 (–) | *Lva* | 45°58′ | 23°02′ |

**Table 1** *Continued*

| Number | Locality | Country | *N* microsatellites | *N* mtDNA (new) | Species | Latitude (N) | Longitude (E) |
|---|---|---|---|---|---|---|---|
| 55 | Deva | RO | — | 4 (–) | *Lva* | 45°53′ | 22°54′ |
| 56 | Călan Băi | RO | — | 2 (–) | *Lva* | 45°44′ | 22°59′ |
| 57 | Câmpu lui Neag | RO | — | 3 (–) | *Lva* | 45°18′ | 23°01′ |

*Lm—Lissotriton montandoni*, *Lva—L. vulgaris ampelensis*, *Lvv—L. v. vulgaris*.

### *Laboratory analyses*

DNA was extracted using the Wizard® Genomic DNA Purification Kit (Promega). The concentration and purity of extracted DNA was estimated using NanoDrop. Fifteen microsatellite loci were scored. Twelve loci were previously described: *Tv3Ca9*, *Tv4Ca9*, *Tv5Ca13* (Johanet *et al*. 2009), *Lm_749*, *Lm_528*, *Lm_632*, *Lm_521*, *Lm_013*, *Lm_870*, *Lm_488* (Nadachowska *et al*. 2010), *Lm_346*, *Lm_AHNC3* (Nadachowska-Brzyska *et al*. 2012), and three were mined from 454 transcriptome sequences (*Lm_ZN5*, *Lm_TDP* and *Lm_8BH*). Loci were amplified in four multiplexes (MPX I-IV), with fluorescently labelled forward primer (Table 2). Multiplex reactions were performed in 8 μL and contained 4 μL of Multiplex PCR Master Mix (QIAGEN), 0.1–0.6 μM of each primer (Table 2) and 30–100 ng of genomic DNA. The following cycling scheme was used: 95 °C/15 min, followed by 29 cycles of 94 °C/30 s, 55 °C/90 s, 72 °C/90 s (MPX I and II) or 35 cycles of 94 °C/30 s, 56 °C/

**Table 2** Microsatellite primer sequences, final concentrations of primers in PCRs and the composition of four multiplex (MPX I–IV) reactions

| Locus | Primer sequence | Final concentration (μM) | PCR multiple × (MPX) | Forward primer label |
|---|---|---|---|---|
| *Lm_749* | F: CCATGGTGGTAGAATAAATGGAA<br>R: AAGACCATTCTTTCTGAGGTATCC | 0.4 | I | PET |
| *Lm_528* | F: CTGGCTTGAAATGCCTTCAT<br>R: AGGGCAGGGCTATACGTCTT | 0.2 | I | NED |
| *Lm_632* | F: CAGAGCAATTTCTAGGCAAGG<br>R: GGCGCTATATCAAACTGCAA | 0.2 | I | HEX |
| *Tv3Ca9* | F: AAATAACTTGTGATTGGGTCATTT<br>R: TGCATATATACTGTATGTTTACTGCA | 0.6 | I | HEX |
| *Lm_521* | F: CATACGGGCACTGAGGTGAT<br>R: GCACAGACATTGATGGCAAA | 0.2 | I | FAM |
| *Tv4Ca9* | F: TTGAGCCAACGCCAATAAAG<br>R: AGGGGGACCTCTACCTTCTG | 0.1 | I | FAM |
| *Lm_013* | F: CTTGGTTCCCAGTGAGGAGA<br>R: GCAAGCCATCCCAAAGTAAG | 0.4 | II | PET |
| *Lm_870* | F: CCACTGCTTTGTGCTGCTAC<br>R: TTTGTCATGGCATTTCCAAC | 0.4 | II | NED |
| *Tv5Ca13* | F: CCACCAGCAGTGCATAAATC<br>R: CACTGGTGATGGCTTTGC | 0.2 | II | HEX |
| *Lm_488* | F: CAGGCAGGGTATTTGCGTAG<br>R: GGTCATTTCCACAACAAGCTC | 0.2 | II | FAM |
| *Lm_346* | F: CCAAAAGACAATTCAAAATGACC<br>R: GGTGATGTCGTTTTAAAGGGTAAC | 0.4 | III | FAM |
| *Lm_ZN5* | F: CGAGTGAAATGTCCAACCAA<br>R: GGAGGAACAATTGACTGGATG | 0.2 | III | VIC |
| *Lm_TDP* | F: ATCCTTGGGCTCTGCAGTTA<br>R: TTGCGTTAGTCCAACTAGATTCTG | 0.6 | III | NED |
| *Lm_AHNC3* | F: TTCCTGTTCTCTGGGGAATG<br>R: GGCACGGACATTTTTAAAGC | 0.1 | IV | FAM |
| *Lm_8BH* | F: TCATACTGTGTGAGTCTATTGTTGGA<br>R: GATTCAGTGCCAGGACACCT | 0.1 | IV | PET |

90 s, 72 °C/90 s (MPX III) or 32 cycles of 94 °C/30 s, 56 °C/90 s, 72 °C/90 s (MPX IV) and a final extension step 72 °C/10 min. The PCR products were electrophoresed on a 3130xl Genetic Analyzer with GeneScan 500 LIZ size standard, and GeneMapper was used for genotyping. The rate of genotyping error was estimated by repeating amplification and the genotyping procedure for 25 randomly selected individuals. Some of the samples analysed here were previously scored for variation in a subset of six loci (Nadachowska-Brzyska *et al.* 2012). However, in the present study, we decided to genotype those loci for all samples to ensure genotyping consistency across the data set, as different dyes had to be used with some loci to efficiently design multiplexes.

A 1016-bp fragment of mtDNA comprising 951 bp of the ND2 gene and almost the entire tRNA-Trp gene was amplified and sequenced as previously described (Babik *et al.* 2005). Both newly obtained sequences and those reported by Babik *et al.* (2005) were included in the analyses. We follow the nomenclature of mtDNA lineages introduced by Babik *et al.* (2005) and present a simplified version of the mtDNA phylogeny, which shows relationships of these lineages as a part of Fig. 6.

### *Population genetics analyses*

*Microsatellites.* MicroChecker (van Oosterhout *et al.* 2004) was used to assess the presence of null alleles and scoring errors owing to stuttering or large allele dropout. Allele frequencies for each locus, tests of Hardy–Weinberg equilibrium and tests of linkage disequilibrium (LD) were calculated in GENEPOP 4.1.2 (Rousset 2008); the type I error was controlled using the Bonferroni procedure. Frequencies of null alleles were estimated with the EM algorithm in FreeNA (Chapuis & Estoup 2007). Microsatellite allelic richness, that is, the per-population number of alleles for each locus corrected for sample size was estimated for all *Lm* samples in FSTAT 2.9.3.2 (Goudet 2001). Because the presence of null alleles was inferred for multiple loci, we applied a correction procedure to minimize the effect of null alleles on the estimates of microsatellite allelic richness (Nadachowska-Brzyska *et al.* 2012). Pairwise $F_{ST}$s between populations were computed from allele frequencies adjusted for the presence of null alleles using the excluding null alleles (ENA) correction in FreeNA (Chapuis & Estoup 2007), and multidimensional scaling (MDS) was used for the visualization of the $F_{ST}$ matrix. Using SMOGD (Crawford 2010), we calculated a standardized $F_{ST}$-like measure $G_{ST}''$, which is particularly useful for assessing differentiation between pairs of populations (Meirmans & Hedrick 2011; but see Whitlock 2011), here also MDS was used for the matrix visualization. To determine the most likely number of genetic clusters present in *Lm*, we used the Bayesian clustering method implemented in STRUCTURE 2.3.3 (Pritchard *et al.* 2000; Falush *et al.* 2007; Hubisz *et al.* 2009). For inferring the genetic structure of *Lm*, we used data for all 15 loci (*Lm* data set). For determining genetic clusters present in both species together, we excluded five loci that amplified poorly in *Lv* (*Lm_521*, *Lm_013*, *Lm_488*, *Lm_TDP* and *Tv4CA9*), resulting in a high fraction of missing data representing most likely null homozygotes (*Lm & Lv* data set). The *Lm* data set was analysed in Structure under the admixture model with correlated allele frequencies. For the *Lm* and *Lv* data set, we ran Structure under the admixture model but with uncorrelated allele frequencies, because the populations were expected to have been more divergent on average. In both analyses, the presence of recessive (null) alleles was assumed. Structure uses multilocus genotype data to cluster individuals, without taking into account the population of their origin, into a predefined number of $K$ genetic clusters, each characterized by a set of allele frequencies. The individuals are probabilistically assigned to clusters in such a way that the deviations from Hardy–Weinberg and LD are minimized. For the *Lm* data set, we examined $K$ values from 1 to 12, and for the *Lm & Lv* data set from 1 to 20. For each $K$ value, we performed 10 runs of a million of burn-in steps followed by a million of postburn-in MCMC iterations. To infer the most likely number of clusters in the data, we calculated $\Delta K$, a measure of second-order rate of change in the likelihood of $K$ (Evanno *et al.* 2005), using the online software Structure Harvester (Earl & vonHoldt 2012).

Analysis of molecular variance (AMOVA) in Arlequin (Excoffier & Lischer 2010) was used to partition microsatellite variation into hierarchical levels. We performed an analysis based on grouping suggested by the geographic distribution of populations and by the Structure results. AMOVA components were tested for significance with 1000 permutations. Isolation by distance was assessed by calculating the significance of correlations between linearized pairwise $F_{ST}$ measures and log-geographical distances using Mantel's test as implemented in IBDWS 3.23 (Jensen *et al.* 2005).

To test the presence of interspecific gene flow along the contact zone, we compared, using randomization tests, the average $F_{ST}$ or $G_{ST}''$ between pairs of neighbouring *Lm–Lv* populations (Fig. 1) with the average $F_{ST}$ or $G_{ST}''$ between random pairs of *Lm–Lv* populations.

*mtDNA.* Median joining networks (Bandelt *et al.* 1999) were constructed using NETWORK 4.6 (http://www.

fluxus-engineering.com/sharenet.htm) for each mtDNA lineage present in *Lm* to reveal relationships between mtDNA haplotypes and patterns of haplotype sharing between *Lm* and *Lv*. Haplotype ($h$) and nucleotide ($\pi$) diversities, and their standard deviations were computed in ARLEQUIN 3.5 (Excoffier & Lischer 2010). To test genetic differentiation between species for shared mtDNA lineages, $F_{ST}$ values were computed separately for the G, J and L lineages which represent an overwhelming majority of *Lm* mtDNA (Babik *et al.* 2005), and their statistical significance was tested with 10 000 permutations in Arlequin. For the G and L lineages, only *Lv* populations sampled in the proximity of the *Lm* range were included in the analyses of genetic differentiation to control for the large-scale geographic structuring of mtDNA variation in *Lv* (Fig. 1, Table 1).

To estimate the times of mtDNA introgression, we performed the isolation with migration (IM) analysis for the G and J mtDNA lineages separately using the IMA2 program (Hey & Nielsen 2007; Hey 2010). Unfortunately, we were not able to obtain reliable estimates of the model parameters with migrations switched on. Therefore, we decided to calculate the time of divergence for each mtDNA lineage under the isolation model (migrations switched off). This model would provide the minimum time of divergence for each mtDNA lineage, and if the introgression occurred over a short time period, such estimates would be close to the actual time of introgression, because the process would be then analogous to the split of a single ancestral into two descendant populations. The IMa2 program was run under the infinite sites model of DNA sequence evolution for $2 \times 10^7$ steps following $10^6$ steps of burn-in; genealogies were sampled every 100 steps. Multiple runs converged on virtually identical parameter estimates, and inspection of effective sample sizes and trend plots indicated good mixing. We converted the IM estimates for population sizes and divergence time into appropriate units by applying two mtDNA mutation rates which represent a range of plausible values for this parameter. The divergence of 22 and 29 my between *Lm*/*Lv* and, respectively, *L. italicus* and *L. helveticus* estimated by Nadachowska & Babik (2009) was used to calculate a 'slow mutation rate' of *c.* $3.8 \times 10^{-9}$/year/bp, while a 'fast mutation rate' of $6.4 \times 10^{-9}$ was derived from Weisrock *et al.* (2001). The generation time of 4 years was assumed following Nadachowska & Babik (2009). We do not report the effective population sizes estimated by IMA2, because they are uninterpretable: with migrations switched off, the estimates are to an unknown, and possibly very large degree, dependent on the history of introgression. Thus, these estimates do not, or only minimally, reflect the demography of the recipient species, but rather the strength and geographic heterogeneity of introgression.

### *Species distribution modelling*

We composed a data set of 45 *Lm* occurrences for which global positioning system coordinates were available, including the 19 populations studied genetically. We downloaded bioclimatic variables at 2.5 arcminute resolution (*c.* 5 × 5 km) from the WORLDCLIM database 1.4 (Hijmans *et al.* 2005, http://www.worldclim.org). To obtain realistic and transferable species distribution models, it is recommended to mirror the physiological limitations of the study species and minimize the effects of multicollinearity among data layers (Guisan & Thuiller 2005; Peterson 2011). We therefore selected a subset of bioclimatic variables that (i) showed a Pearson's correlation of $r < 0.7$ and (ii) are deemed biologically meaningful based on the knowledge of life history of the model system: bio10 = mean temperature of warmest quarter, bio11 = mean temperature of coldest quarter, bio15 = precipitation seasonality, bio16 = precipitation of wettest quarter, and bio17 = precipitation of driest quarter. Bioclimatic variables are also available from the WorldClim database for the LGM (~21 Ka). These data are derived from the Paleoclimate Modelling Intercomparison Project phase 2 (Braconnot *et al.* 2007, http://pmip2.lsce.ipsl.fr/) and based on two climate simulations: the Model for Interdisciplinary Research on Climate version 3.2 (MIROC; http://www.ccsr.u-tokyo.ac.jp/kyosei/hasumi/MIROC/techrepo.pdf) and the Community Climate System Model version 3 (CCSM; Collins *et al.* 2006).

We created a species distribution model for *Lm* with Maxent 3.3.3k (Phillips *et al.* 2006). To facilitate extrapolation to a different time frame, we restricted the feature type to hinge features (Elith *et al.* 2010). The environmental range covered by pseudo-absence data, used by Maxent to discriminate presence data from the environmental background, should neither be too narrow nor too broad (VanDerWal *et al.* 2009). A practical solution is to restrict the area from which pseudo-absence is drawn to a buffer around the localities on which the model is calibrated (Warren *et al.* 2010). We used a 200-km radius following VanDerWal *et al.* (2009). We projected the *Lm* species distribution model on the current and LGM climate layers. To determine whether the *Lm* species distribution model performs better than expected by chance, its AUC (i.e. the area under the receiver operating characteristic curve) value was tested for statistical significance against a null model based on AUC values of 99 species distribution models based on random localities (sensu Raes & ter Steege 2007).

## Results

### *Microsatellite variation*

All microsatellite loci were polymorphic in both species. The estimated rate of genotyping error was 4.78%.

The number of alleles per locus ranged from 10 (*Lm_AHNC3*) to 97 (*Lm_521*) with a mean of 41.2 (SD = 24.4). Tests for LD across populations followed by Bonferroni correction produced only one significant result, for the pair *Lm_632-Tv5CA13*. This overall significant LD was due to a highly significant result for a single population (22), which indicates that it may be a consequence of local demography/admixture rather than of physical linkage. Hence, all the markers were regarded as segregating independently. After applying Bonferroni correction, we detected significant deviations from Hardy–Weinberg expectations in 60 of 265 tests (23%) performed for *Lm* populations and in 125 of 425 tests (29%) for all analysed populations (Table S1, Supporting information).

Null alleles were widespread, for each locus they were detected in at least some populations and for five loci in almost all populations at varying frequencies (Table S2, Supporting information). The abundance of null alleles across loci may be related to the large overall effective population size, historical structuring and possibly a high mutation rate and represents biological reality. Therefore, we included all loci in the analyses of *Lm*, applying the appropriate corrections. However, null alleles appeared much more abundant in *Lv*, to the point that we decided to exclude five loci (*Lm_521*, *Lm_013*, *Lm_488*, *Lm_TDP*, *Tv4CA9*) that amplified poorly in *Lv* in comparison with *Lm*, from the analyses including both species.

The mean allelic richness per locality ranged from 4.6 to 9.0 and was lowest in populations 2 and 27 and highest in 10 and 14 (Table S3, Supporting information). A clear geographic pattern was observed for allelic richness: populations from the northern (1–3, 5) and southern (23–27) part of the species range had lower allelic richness compared with populations from the central part of the range in the Eastern Carpathians (10, 14–16, 18–22; Mann–Whitney *U*-test, $P = 0.007$ for northern–eastern comparison, $P = 0.003$ for eastern–southern comparison, $P = 0.0004$ northern and southern combined compared with eastern; Table S3, Fig. S1, Supporting information). The abundance of private alleles did not show a clear geographic pattern (Fig. S1, Supporting information).

### *Structuring of microsatellite variation in Lissotriton montandoni*

Genetic differentiation between *Lm* populations was significant in 150 of 153 pairwise comparisons and ranged from negligible 0.002 ($F_{ST}$[ENA]) between populations 15 and 16 or 0.055 ($G_{ST}''$) between 3 and 5 to highly significant with the highest differentiation ($F_{ST}$[ENA] = 0.276, $G_{ST}'' = 0.814$) between 1 and 27, two populations at the, respectively, north-western and southern extremes of the species range (Tables S4 and S5, Supporting information). The MDS plot of pairwise $F_{ST}$ and $G_{ST}''$ suggested the presence of three genetically differentiated groups of populations with distinct geographic distributions in the northern, eastern and southern parts of the Carpathians (Figs 2 and 3 and Fig. S2, Supporting information). The presence of three genetic clusters within *Lm* was corroborated by the results of Structure analysis, which produced the highest $\Delta K$ value for $K = 3$ (Fig. 3 and Fig. S3, Supporting Information) and showed relatively little admixture

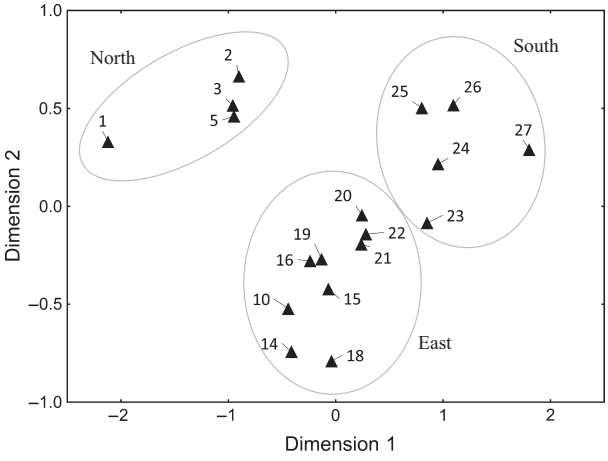


**Fig. 2** Two-dimensional scaling of the matrix of pairwise $G_{ST}''$ between *Lissotriton montandoni* populations.

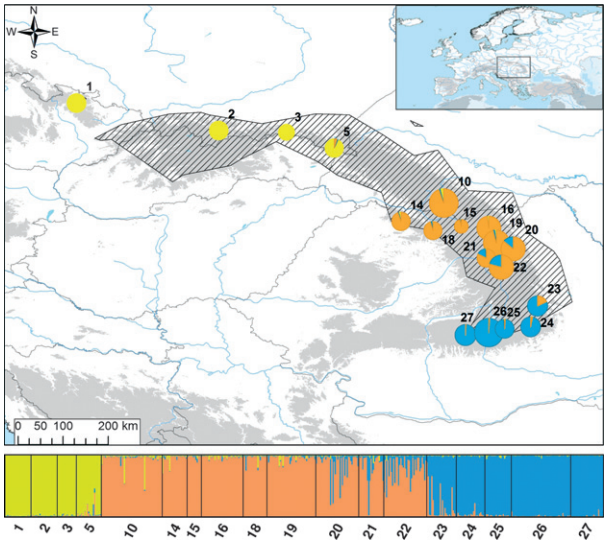


**Fig. 3** Genetic structure of *Lissotriton montandoni* populations inferred by Structure from 15 microsatellite loci for $K = 3$. Symbol areas are proportional to sample sizes. The distribution of *L. montandoni* (Zavadil *et al.* 2003 and own unpublished data) is hatched. Areas above 500 m a.s.l. are shaded.

between clusters, that is, most populations were classified entirely or almost entirely to a single cluster. Genetic clusters were not only distinct but also substantially differentiated from each other as revealed by AMOVA, with 6.9% of microsatellite variation distributed among clusters as compared to 4.9% between populations within clusters (Table 3).

As could have been expected from the significant among-population differentiation, there was a clear pattern of isolation by distance observed both at the level of the entire species ($r = 0.86$, $P < 10^{-3}$, Mantel test, Fig. S4, Supporting Information) and within genetic clusters (Fig. S5, Supporting information).

### *Microsatellite differentiation between Lissotriton montandoni and L. vulgaris*

All analyses involving both species were performed with 10 microsatellite loci (see above). Pairwise genetic differentiation between *Lm* and *Lv* populations was in all cases strong and highly significant (Tables S6 and S7, Supporting information), and, consequently, separation of both species was evident in MDS plots (Fig. 4

**Table 3** Results of the Analysis of Molecular Variance (AMOVA) performed for two data sets

| Source of variation | d.f. | Sum of squares | % of variation explained | *P* |
|---|---|---|---|---|
| *Lissotriton montandoni*—three groups of populations as the highest hierarchical level (15 loci) | | | | |
| Among groups | 2 | 265.908 | 6.93 | $<10^{-4}$ |
| Among populations | 15 | 295.043 | 4.91 | $<10^{-4}$ |
| Within populations | 986 | 4767.389 | 88.16 | $<10^{-4}$ |
| *Lissotriton montandoni* and *L. vulgaris*—species as the highest hierarchical level (10 loci) | | | | |
| Among species | 1 | 211.218 | 7.79 | $<10^{-4}$ |
| Among populations within species | 26 | 612.031 | 9.27 | $<10^{-4}$ |
| Within populations | 1360 | 4911.034 | 82.93 | $<10^{-4}$ |

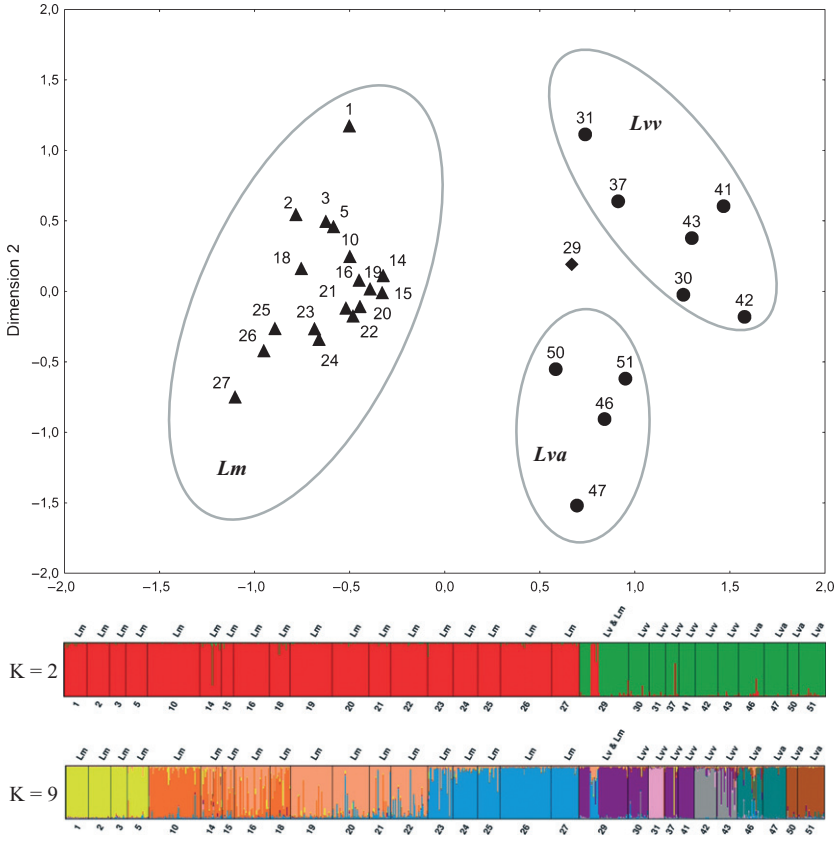


**Fig. 4** Two-dimensional scaling of the matrix of pairwise $G_{ST}''$ between *Lissotriton montandoni* (*Lm*) and *L. vulgaris* (*Lva*, *L. v. ampelensis*; *Lvv*, *L. v. vulgaris*), populations and genetic structure inferred by Structure for $K = 2$ and $K = 9$.

and Fig. S6, Supporting information). Within *Lv*, subspecies *Lva* and *Lvv* also form two somewhat differentiated groups. Additional insights were provided by the Structure analysis. Not only was the major division into two species confirmed (Fig. 4) but also only very few individuals of mixed ancestry were identified. Interestingly, even in locality 29, where both species co-occur in a single pond, only two substantially admixed individuals were identified. Such results indicate that the overall nuclear gene flow between species is currently very limited along the contact zone.

Although $\Delta K$ unequivocally identified two major genetic clusters corresponding to species, the highest probability of data was obtained for $K = 9$, which allowed to identify further genetic structure within both species. Four clusters were detected in *Lm*, with an additional split within the eastern group in comparison with the Structure results based on 15 loci. Each of the five clusters within *Lv* comprised mainly geographically close populations. An AMOVA revealed that 7.8% of variation was distributed between species, while 9.3% was distributed between populations within species, both components are highly significant (Table 3). However, the extent of interspecific differentiation varied among loci. Locus by locus AMOVA revealed that for three loci (*Lm_749*, *Lm_632*, *Lm_870*), the interspecific level, although significant, explained only a very small percentage of variation (Table 4) and MDS plots of $F_{ST}$ and $G_{ST}''$ confirmed little separation between species based on these three loci (Fig. S8, Supporting information).

To further check for the possibility of interspecific gene flow along the contact zone, we tested whether pairs of neighbouring *Lm–Lv* populations were less differentiated than random *Lm–Lv* pairs. We selected seven pairs of neighbouring populations (separated by the distances of 1–61 km; Fig. 1, blue arrows) and compared the average $F_{ST}$ and $G_{ST}''$ values computed for such pairs with the distribution of the averages computed for 1000 random samples of seven interspecific population pairs. Pairs of neighbouring populations did not show less differentiation than random interspecific pairs either for all 10 microsatellites ($G_{ST}''$: $P = 0.619$; $F_{ST}$: $P = 0.736$) or for three loci least differentiated between species ($G_{ST}''$: $P = 0.872$; $F_{ST}$: $P = 0.870$).

### *mtDNA variation, structuring and interspecific differentiation*

A total of 141 new mtDNA sequences from 26 populations were obtained in the current study and together with those reported by Babik *et al*. (2005), and 246 mtDNA sequences from 55 populations were analysed. The genetic composition of populations and basic statistics describing sequence variation are presented in Table S8 (Supporting information). The correlation between genetic distance measures based on mtDNA ($F_{ST}$) and microsatellites ($G_{ST}''$, 10 loci) was significant but rather low ($r = 0.273$, $P = 0.005$, Mantel test). The matrix of pairwise $F_{ST}$ is given in Table S9 (Supporting information). These tables are, however, difficult to interpret, because high heterogeneity in polymorphism and differentiation reflects mainly the fact that in many populations, two divergent mtDNA lineages were present, whereas in others only a single lineage occurred. Therefore, a more informative analysis of mtDNA variation and differentiation within and between species was performed at the level of major lineages. Three mtDNA lineages are common in *Lm*: G, I and J, and two additional lineages, F and L occur locally (Figs 5 and 6). Only lineage I, found in the northern part of the species range (Poland and Ukraine), appears specific to *Lm*, whereas the remaining lineages are found in *Lv* as well.

**Table 4** Results of the locus by locus AMOVA

| | Among species | | Among populations within species | | Within populations | |
|---|---|---|---|---|---|---|
| Locus | % variation explained | *P* | % variation explained | *P* | % variation explained | *P* |
| *Lm_749* | 1.29 | 0.0059 | 7.29 | $<10^{-4}$ | 91.42 | $<10^{-4}$ |
| *Lm_528* | 4.90 | $<10^{-4}$ | 5.38 | $<10^{-4}$ | 89.72 | $<10^{-4}$ |
| *Lm_632* | 2.18 | $<10^{-4}$ | 8.55 | $<10^{-4}$ | 89.28 | $<10^{-4}$ |
| *Tv3CA9* | 21.77 | $<10^{-4}$ | 9.50 | $<10^{-4}$ | 68.73 | $<10^{-4}$ |
| *Lm_870* | 0.76 | 0.0225 | 6.23 | $<10^{-4}$ | 93.01 | $<10^{-4}$ |
| *Tv5CA13* | 13.95 | $<10^{-4}$ | 10.73 | $<10^{-4}$ | 75.32 | $<10^{-4}$ |
| *Lm_346* | 12.79 | $<10^{-4}$ | 9.46 | $<10^{-4}$ | 77.75 | $<10^{-4}$ |
| *Lm_ZN5* | 4.16 | 0.0068 | 17.50 | $<10^{-4}$ | 78.34 | $<10^{-4}$ |
| *Lm_AHNC3* | 6.93 | $<10^{-4}$ | 10.44 | $<10^{-4}$ | 82.63 | $<10^{-4}$ |
| *Lm_8BH* | 9.43 | $<10^{-4}$ | 10.01 | $<10^{-4}$ | 80.56 | $<10^{-4}$ |

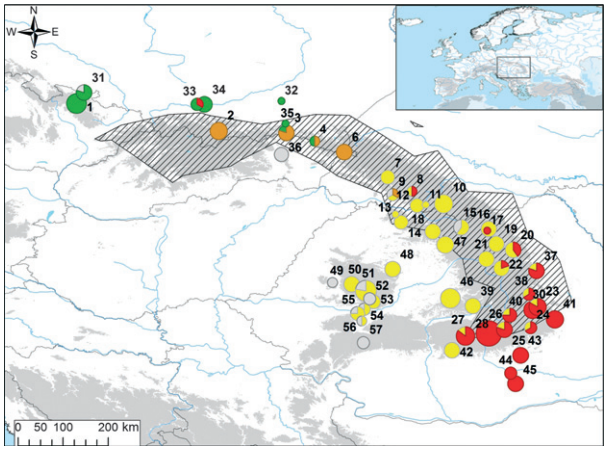


**Fig. 5** Geographic distribution of mtDNA lineages. Colours correspond to particular mtDNA lineages as used by Babik *et al.* (2005): grey—F, red—G, brown—I, yellow—J, green—L. Symbol areas are proportional to sample sizes. The distribution of *Lissotriton montandoni* (Zavadil *et al.* 2003 and own unpublished data) is hatched. Areas above 500 m a.s.l. are shaded.

Our extensive sampling did not detect any additional mtDNA lineages, which could have been identified as 'original' *Lm* mtDNA. The geographic distributions of lineages G and J, found in *Lm* in Romania and Ukraine, overlap partially, with haplotypes from both lineages co-occurring in seven localities. However, the G lineage is frequent in the southern part of the Romanian *Lm* range, whereas lineage J is mainly or exclusively present in the remaining part of the Romanian and in the Ukrainian Carpathians (Fig. 5). Thus, the distribution of the G and J lineages is roughly concordant with the genetic clusters identified in microsatellites, but the degree of their geographic overlap is larger. In *Lv*, the G and J lineages have only slightly overlapping distributions: lineage J (corresponding mainly to the *Lva* subspecies) occurs in Romania west of the Carpathians and lineage G predominates in the lowlands at the outer side of the Carpathian belt (Fig. 5).

Lineages G and J show dissimilar patterns of interspecific differentiation: it is not significant in lineage G ($F_{ST}$ = 0.021, $P$ = 0.14), but substantial in lineage J ($F_{ST}$ = 0.117, $P < 10^{-4}$). Haplotype sharing between species occurred in both G and J lineages, and haplotypes private to each species were intermingled in the networks (Fig. 6). The time of the interspecific split estimated by IMa2 under the isolation model, which should thus be interpreted as a minimum time, was approximately two times younger for the G lineage (*c.* 20 or 33 kya depending on the mutation rate) than for the J lineage (*c.* 39 or 65 kya; Table 5). Interestingly, for neither lineage, the highest posterior density of divergence time overlapped zero, indicating no signal of contemporary mitochondrial gene flow. Haplotypes from the *Lv* lineage L are found in the isolated *Lm* population from the Eastern Sudetes (locality 1) and, rarely, together with haplotypes from the I lineage in the Polish *Lm* populations. Interspecific $F_{ST}$ was not significant for this lineage ($F_{ST}$ = 0.010, $P$ = 0.081), and some haplotypes were shared between species (Fig. 6).

Nucleotide diversity differed substantially between the mtDNA lineages present in *Lm* (Table 6). Of the major lineages, the lowest diversity was found in I, followed by G and J, where it was an order of magnitude higher than in I. Nucleotide diversity in lineage G was lower in *Lm* than in *Lv*, whereas in lineage J, nucleotide diversities in both species were similar (Table 6). The relatively high nucleotide diversity in *Lm* was caused by the presence of haplotypes from the L2 and L3 sublineages in the Eastern Sudetes (1), whereas the Polish *Lv* populations had haplotypes only from the L3 sublineage. Haplotype diversity (Table 6 and Table S8, Supporting information) is probably not particularly informative about levels of variation because the sampling scheme may affect haplotype diversity more strongly than nucleotide diversity due to a high incidence of slightly divergent haplotypes private to individual populations.

### *Species distribution modelling*

The species distribution model for *Lm* performs statistically significantly better than random expectation (Fig 7). Figure 7 depicts the predicted suitability of the Carpathian region for *Lm* at present and at the LGM. The currently predicted range of *Lm* closely follows the Carpathian range (cf. Fig 1). At the LGM, the Carpathian region in general is predicted to have been suitable for *Lm*. At a smaller scale, differences are apparent. The two different climate simulations used for the LGM (MIROC and CCSM) paint a different picture. Both agree on suitable area being more widely distributed at lower elevation. Only the MIROC climate simulation suggests that part of the current distribution area was not suitable at the LGM and that the *Lm* distribution was fragmented at the time. It should be noted, however, that also according to the CCSM climate simulation, the current range was less suitable at the LGM.

## Discussion

### *Introgressions from Lissotriton vulgaris to L. montandoni resulted in complete mtDNA replacement*

This study shows, using extensive sampling, that all mtDNA present in *Lm* is derived from *Lv*—no new, previously unknown mtDNA lineages, which could

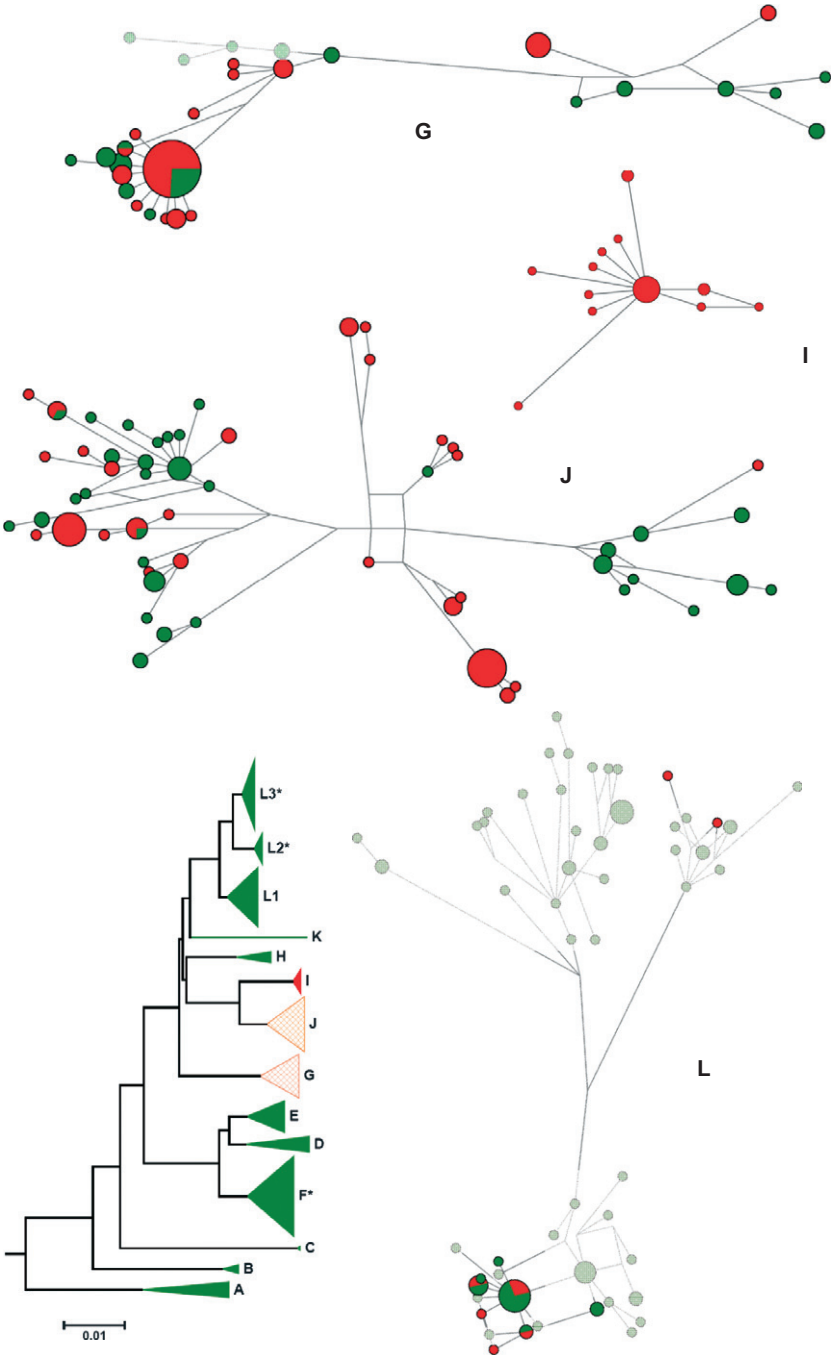


**Fig. 6** Median joining networks showing *Lissotriton montandoni* (*Lm*)and *L. vulgaris* (*Lv*) haplotypes from mitochondrial lineages G, I, J and L. Within each lineage, node sizes are proportional to haplotype frequencies. Red—haplotypes found in *Lm*; green—haplotypes found in *Lv*; for shared haplotypes, respective colours correspond to frequencies of the haplotype in each species. For the G and L lineages, haplotypes found in *Lv* populations far from the *Lm* range are shaded. The simplified mtDNA tree depicting relationship of all mtDNA lineages found in *Lm*/*Lv* (adapted from Babik *et al.* 2005) is shown in the lower left corner; red—the lineage unique to *Lm*; green—lineages found exclusively or almost exclusively in *Lv* (asterisks indicate lineages in which a limited introgression into *Lm* has been detected); orange pattern—the lineages extensively shared by both species.

have represented the species' 'original' mtDNA, were detected in *Lm*. Our results suggest that this range-wide introgression occurred via multiple, spatially and temporally separated episodes of hybridization. The geographic distributions of introgressed mtDNA lineages overlap to a larger extent than previously thought (Babik *et al.* 2005), with extensive admixture of the G and J lineages in the *Lm* range. The mtDNA lineages present in *Lm* show dissimilar patterns of interspecific differentiation. The I lineage is present only in *Lm*, and its relatively old introgression was inferred from the mtDNA phylogeny (Babik *et al.* 2005). The J lineage occurs in both species, and some haplotypes are shared, but significant interspecific differentiation was observed. Finally, the G lineage, which appears to have introgressed more recently, is not significantly differentiated between the two species: they share haplotypes extensively. Another feature of mtDNA introgression readily noticed in lineages G and J is a substantial diversity in both species, indicating that introgression of each lineage was not a single, spatially

**Table 5** Estimates of the split time (which corresponds to the minimum time of introgression) and the demographic parameters for the G and J mtDNA lineages under the isolation model (migrations set to zero) implemented in IMa2. *t*—time of divergence in years; HPD95Lo and HPD95Hi—the lower and upper bounds of the estimated 95% highest posterior density (HPD) interval

| Lineage | *t* |
|---|---|
| | Slow mutation rate: $3.8 \times 10^{-9}$/bp/year ($3.7 \times 10^{-6}$/locus/year) |
| G (HPD95Lo-HPD95Hi) | $3.27 \times 10^4$ ($9.87 \times 10^3$–$9.54 \times 10^4$) |
| J (HPD95Lo-HPD95Hi) | $6.54 \times 10^4$ ($2.96 \times 10^4$–$1.35 \times 10^5$) |
| | Fast mutation rate: $6.4 \times 10^{-9}$/bp/year ($6.3 \times 10^{-6}$/locus/year) |
| G (HPD95Lo-HPD95Hi) | $1.97 \times 10^4$ ($5.95 \times 10^3$–$5.75 \times 10^4$) |
| J (HPD95Lo-HPD95Hi) | $3.94 \times 10^4$ ($1.79 \times 10^4$–$8.13 \times 10^4$) |

restricted event, as such a 'founder event' would cause limited mtDNA variation in the recipient species. Alternatively, variation might have recovered after introgression through mutation, which we deem unlikely because of insufficient time. The IMa2 analysis of the isolation model provided the minimum time estimates of introgression events; if introgressions of individual mtDNA lineages were temporary restricted, these could represent the actual timing. Inferred introgression was twice as old for the J lineage compared with the G lineage, but introgression events of both lineages were firmly placed within the last glacial period (115–15 kya), regardless of the mutation rate used. Time of introgression for the G lineage, estimated under the fast mutation rate, corresponded very well to the LGM. Thus, the climatic oscillations and long periods of unfavourable climate during the last glacial period probably affected the interaction between species in a way that promoted introgression. Interestingly, despite extensive interspecific haplotype sharing in the G lineage, the IMa2 analysis did not provide evidence for contemporary mtDNA gene flow—in such a case, under the isolation model, we would expect the highest posterior density of the splitting time to overlap zero.

Partial replacement of the original mtDNA due to introgression from a related species is a common phenomenon (reviewed in Petit & Excoffier 2009; Toews & Brelsford 2012), but complete replacement across the whole range appears rarer (Melo-Ferreira *et al.* 2012; Toews & Brelsford 2012 and references therein). Many possible scenarios leading to the complete replacement of the *Lm* mtDNA could be proposed, differing in the complexity and the relative importance attributed to various mechanisms and factors. Assessing the likelihood of such scenarios and thus ranking them according to plausibility are difficult (Toews & Brelsford 2012). We would like to emphasize, however, that the observed patterns of introgression and available information about the *Lm*/*Lv* system are largely consistent with recent results on the expected patterns of neutral introgression during biological invasions (Currat *et al.* 2008). The Currat *et al.* study modelled the genetic composition of populations at neutral markers in a spatially explicit setting and showed that introgression from the resident to the invading species is the rule. The neutral model of Currat *et al.* (2008) predicts that if some markers introgress more readily than others, these are the markers that show lower intraspecific gene flow. Petit & Excoffier

**Table 6** Intraspecific mtDNA variation within individual mitochondrial lineages

| Species | mtDNA | npop | *N* | nhap | npriv | *S* | π | SD(π) | *h* | SD(*h*) |
|---|---|---|---|---|---|---|---|---|---|---|
| Lm | G | 10 | 39 | 11 | 9 | 21 | 0.0045 | 0.0025 | 0.72 | 0.07 |
| Lv | G | 9 | 29 | 14 | 12 | 24 | 0.0061 | 0.0033 | 0.92 | 0.03 |
| Lm | J | 19 | 57 | 25 | 23 | 57 | 0.0100 | 0.0051 | 0.91 | 0.02 |
| Lv | J | 14 | 54 | 34 | 32 | 63 | 0.0107 | 0.0055 | 0.98 | 0.01 |
| Lm | I | 5 (12) | 23 | 12 | 12 | 14 | 0.0016 | 0.0011 | 0.81 | 0.08 |
| Lm | L | 3 | 10 | 7 | 5 | 19 | 0.0065 | 0.0038 | 0.91 | 0.08 |
| Lv | L | 5 | 13 | 6 | 4 | 6 | 0.0010 | 0.0008 | 0.64 | 0.15 |

For *Lv* lineages, G and L only populations surrounding *Lm* range were included (Fig. 1, Table 1). For the *Lm*, I lineage also sequences available from the hybrid zone in southern Poland reported by Babik *et al.* (2005) were included to increase the sample size. *Lm*—*Lissotriton montandoni*; *Lv*—*L. vulgaris*; *npop*—number of populations; *n*—sample size; *nhap*—number of mtDNA haplotypes; *npriv*—number of private mtDNA haplotypes; *S*—number of segregating sites; π—nucleotide diversity; SD—standard deviation; *h*—haplotype diversity.

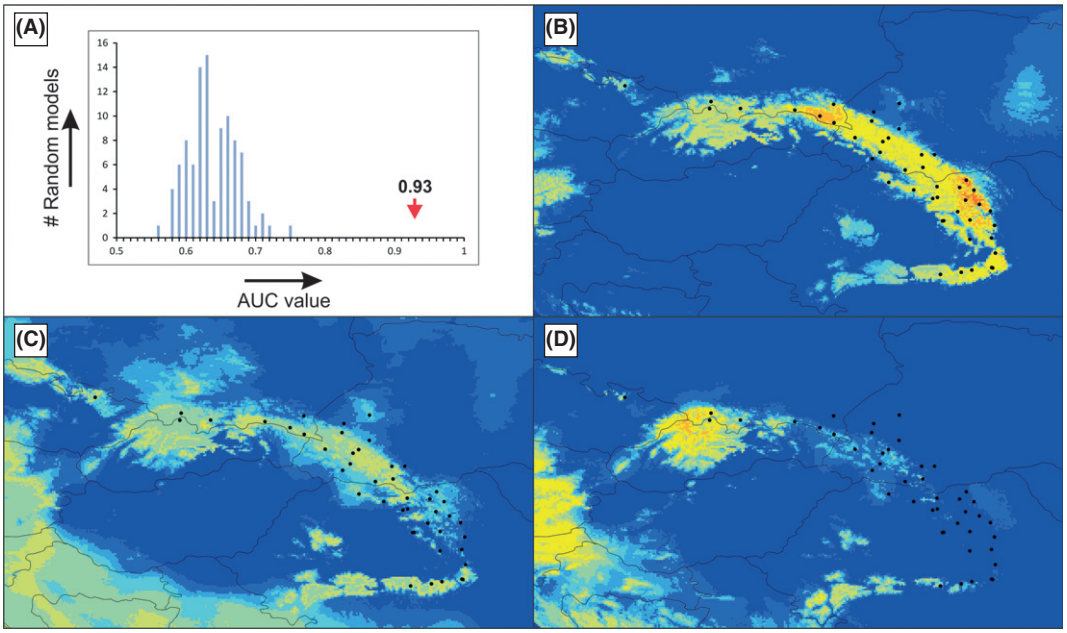


**Fig. 7** Results of the species distribution modelling for *Lissotriton montandoni*. (A) performance of the model in comparison with 99 random models; (B) current predicted distribution; (C) predicted distribution during the Last Glacial Maximum (LGM) based on the Community Climate System Model version 3 (CCSM) climate simulation; (D) predicted distribution during the LGM based on the Model for Interdisciplinary Research on Climate version 3.2 (MIROC) climate simulation In B–D, the warmer the colour of a grid cell, the higher its suitability is; black dots indicate the 46 localities with available GPS coordinates which were used to approximate the current species range.

(2009) argue that the empirical patterns are largely consistent with these expectations (but see Toews & Brelsford 2012). Little is known about sex-biased dispersal in newts, but data for the Alpine newt indicate male-biased juvenile dispersal (Joly & Grolet 1996), suggesting male-biased dispersal is also plausible in *Lm*/*Lv*, which would imply limited intraspecific mtDNA gene flow.

Distribution modelling shows range shifts of *Lm*, with the LGM distribution less extensive than the current distribution. The range changes probably also involved an altitudinal component: *Lm* would descend into lower elevations invading the range of *Lv*, and mtDNA introgression from *Lv* could occur during such invasions. An additional factor facilitating unidirectional introgression could be the fact that in laboratory and semi-natural conditions, *Lv* females mate more readily with heterospecific males than *Lm* females (Michalak *et al.* 1997; Michalak & Rafinski 1999). The relative importance of these factors could be tested in hybrid zones located in the northern part of *Lm* range (where mtDNA is diagnostic for each species), in the areas where land use changes have occurred. If the neutral mechanism is more important, introgression should always occur into the expanding species. If the differences in mate choice prevail, introgression into the expanding species should be weaker or absent in situations when *Lv* is expanding to the area previously inhabited by *Lm* (e.g. following deforestation). Because the process of invasion could occur simultaneously across large areas, introgression could have been massive, resulting in the substantial diversity currently observed in the *Lm* G and J lineages. *Lm* populations at higher elevations, containing 'original' mtDNA, would go extinct as climatic conditions deteriorated, and the subsequent recolonization by introgressed *Lm* populations would produce the currently observed pattern of the range-wide mtDNA replacement. The absence of *Lv* in more northern areas during the last glacial period, inferred from patterns of mtDNA variation (Babik *et al.* 2005), may explain the lack of recent introgression in the northern part of *Lm* distribution occupied by the mtDNA lineage I. The introgression of the I lineage most probably occurred during one of the previous glacial periods. The isolated *Lm* population (locality 1) from the Eastern Sudetes, completely introgressed with the *Lv* L mtDNA lineage, may constitute another example of invasion-related introgression. This population clusters unambiguously in microsatellites with the remaining northern *Lm* populations, indicating a recent common ancestry. Hence, the introgression of the L lineage could have occurred postglacially, during the expansion into the eastern Sudetes through the areas already occupied by the *Lv* L lineage. Additional support to this scenario comes from the hybrid zone was analysed by Babik *et al.* (2003) in

Poland, where each species shows diagnostic mtDNA. In this hybrid zone, a substantial frequency of *Lv* mtDNA was observed in *Lm* populations inhabiting re-forested areas, where *Lm* recently replaced *Lv*.

The largely neutral scenario presented above is speculative, and its rigorous testing remains challenging, but the available data are in line with this 'null' (Currat *et al.* 2008) hypothesis. On the other hand, the presence of multiple introgressed mtDNA lineages and their geographic distribution speak against adaptive introgression—it is difficult to imagine how would such divergent lineages all present selective advantage, unless the original mtDNA in *Lm* accumulated deleterious mutations due to drift during periods of small population sizes so that each introgressed *Lv* lineage would be more fit than the original *Lm* mtDNA (Ballard & Whitlock 2004). The pattern of mtDNA introgression does not support strong negative epistatic interactions between the mitochondrial and nuclear genomes, unless one supposes that relevant nuclear genes have introgressed as well, which is certainly a testable hypothesis.

### *No evidence for recent introgression of microsatellite loci between the two newt species*

Contrary to the mtDNA situation, we found little evidence for recent genetic exchange in 15 microsatellite loci, which may be considered a first approximation of the degree of genome-wide interspecific gene flow. A similar result was found in a fine-scale analysis of a hybrid zone in Poland, where species are strongly isolated reproductively and introgression is restricted (Babik *et al.* 2003). However, in Poland current mtDNA gene flow is also restricted, and no signature of a recent large-scale mtDNA introgression was detected. Additionally, the markers used by Babik *et al.* (2003) did not constitute a random sample, as diagnostic and almost diagnostic loci were overrepresented. The strength of overall genomic isolation may thus have been overestimated, if substantial genomic heterogeneity in the level of introgression occurs. We expected that a signal of stronger exchange in the nuclear genome may be detected in the Romanian part of the range, where hybridization, as evidenced by the mtDNA, has been more recent and likely more extensive. Also the use of randomly selected polymorphic markers should provide a more balanced view. We have, however, not found any evidence for geographical variation in the amount of nuclear gene flow at the spatial scale investigated, and a population in the Romanian Carpathians (29), where both species co-occur, exhibits a bimodal distribution of genotypes similar to that observed in the Polish hybrid zone. Massive introgression of the L mtDNA lineage into the marginal *Lm* population in Sudetes has not been accompanied by introgression of microsatellites, further emphasizing the lack of genome-wide interspecific gene flow.

In this context, the three microsatellite loci showing little interspecific differentiation deserve attention. The overall lack of differentiation is not accompanied by higher similarity of neighbouring interspecific populations, speaking against recent introgression. Thus, all 15 loci support a scenario of no recent or ongoing interspecific nuclear gene flow between *Lm* and *Lv*. It is possible that the three least differentiated loci did actually introgress in the past, but gene flow has homogenized allele frequencies in the recipient species so that the geographic signal of introgression has been lost. Other theoretically possible reasons for the lack of differentiation include similarity due to ancestral polymorphism or extensive allele size homoplasy resulting from constrains on the mutational process, but these two possibilities are less plausible for quickly mutating, highly polymorphic loci. We cannot, however, dismiss the possibility that one or more of these microsatellite loci are linked to genes under balancing selection, which would slow down interspecific differentiation of allele frequencies.

Even if representative of the average behaviour of nuclear markers, microsatellites do not tell the complete story of interspecific nuclear genetic exchange, as demonstrated by extensive introgression of MHC class II genes between the two species (Nadachowska-Brzyska *et al.* 2012), MHC genes may, however, be atypical with respect to introgression—they are a prime example of genes evolving under balancing selection and both theory and limited experimental evidence indicate that such genes should be among the last to stop flowing between incipient species as the strength of barriers against gene flow increases (Schierup *et al.* 2000; Castric *et al.* 2008; Abi-Rached *et al.* 2011). This is because, if balancing selection operates through the rare allele advantage, introgressed alleles immediately enjoy a selective advantage due to their rarity, which facilitates their establishment in the recipient species, even if interspecific gene flow is too limited for effective introgression of neutral markers. Recent genome-wide studies of human–Neanderthal (Green *et al.* 2010) and brown bear–polar bear (Miller *et al.* 2012) patterns of gene flow indicate that the admixture resulting from hybridization may affect a relatively small portion (on the order of a few percent) of the genome. In such cases, dense sampling of variation across the genome is needed to detect signals of introgression.

Several mechanisms have been proposed to explain the commonly observed weaker introgression of nuclear genes compared with mtDNA (Wirtz 1999; Ballard & Whitlock 2004; Chan & Levin 2005; Petit & Excoffier

2009). If numerous regions involved in incompatibilities causing reduced fitness of hybrids are scattered across the genome, they form a barrier to gene flow, which impedes introgression of linked neutral variants (Barton & Bengtsson 1986). Because mtDNA is not physically linked to the nuclear genome, it may introgress more freely (Martinsen *et al.* 2001; Funk & Omland 2003). Unequal abundance of hybridizing species and frequency-dependent prezygotic isolation may cause females of the rarer species to mate more often with males of the more abundant species leading to both stronger and asymmetric mtDNA introgression compared with nuclear genes (Wirtz 1999; Chan & Levin 2005). Finally, if introgression is neutral and the consequence of invasion, introgression from the resident to the invading species is the rule and expected to be stronger in the parts of the genome experiencing reduced interspecific gene flow, as is often the case with mtDNA (Petit & Excoffier 2009).

A previous analysis of the Polish *Lm*/*Lv* hybrid zone did not detect a clinal pattern of allele frequencies that would indicate strong genomic incompatibilities between species (Babik *et al.* 2003). However, this may reflect a low power rather than a lack of genomic incompatibilities. A low power may have resulted from the dominant role of assortative mating and extinction–recolonization events in shaping the composition of the hybrid zone as well as from a limited number of markers. Taking into account the old divergence and substantial genetic differentiation between the two species, as evidenced by the current study, incompatibilities between their nuclear genomes are likely. Closer insights into the nature of interactions between genomes of hybridizing species may be obtained by examining differential introgression of various genomic segments in hybrid zones using a large number of markers (Payseur 2010). The availability of a genetic map would greatly facilitate such studies.

### *Genetic clusters within Lissotriton montandoni are probably derived from separate refugia*

Genetic structure of the *Lm* populations is consistent with long-term survival of this species in the Carpathian region as suggested by Babik *et al.* (2005). The importance of the Carpathians as a refugium for temperate biota, located further to the north than canonical refugia, has now been relatively well documented (reviewed in Sommer & Nadachowski 2006; Provan & Bennett 2008; Ronikier 2011), and refugia have been identified there also for terrestrial poikilotherm vertebrates (Ursenbacher *et al.* 2006; Fijarczyk *et al.* 2011). The results of species distribution modelling confirm that climatic conditions during the LGM would have allowed survival of *Lm* over a substantial portion of the Carpathians. Especially the results based on the MIROC climate model are consistent with isolation of the populations in the southern part of the Carpathians and fragmentation also in the eastern and northern part of the range. Similar patterns of genetic differentiation between populations from the northern, eastern and southern parts of the Carpathians have been reported for high altitude plants *Campanula alpina* (Ronikier 2011), *Hypochaeris uniflora* (Mraz *et al.* 2007), and differentiation between northern and eastern-south Carpathians was described for the caddisfly *Drusus discolor* (Pauls *et al.* 2006).

Unfortunately, with the data at hand, it is also difficult to estimate the age of the nuclear genetic differentiation within *Lm*. First, mtDNA is not very useful in this respect as the timing of mtDNA introgression may have little to do with nuclear differentiation within *Lm*. Second, microsatellite data have proven notoriously difficult to estimate divergence times and overcoming these difficulties requires the availability of extensive sequence data to guide and validate calibration of the microsatellite clock (Sun *et al.* 2009). Therefore, we postpone proposing an absolute temporal perspective on differentiation until additional sequence data from the nuclear genome become available.

### *Genome-wide polymorphism and divergence data are needed to unravel the history of genetic exchange between species*

The lack of similarities in microsatellite allele frequencies between neighbouring interspecific populations suggests little recent interspecific gene flow in the nuclear genome. These data, however, do not provide conclusive evidence for the lack of nuclear genetic exchange between the two newt species. First, as discussed above, only a relatively minor fraction of the genome may be subject to gene flow, but this could still encompass a large number of genes of potential evolutionary significance (Miller *et al.* 2012). Second, little evidence for contemporary gene flow does not preclude the possibility that gene flow was more extensive in the past. For exploring the first possibility, truly genome-wide polymorphism and divergence data are essential. As sequencing of the vast newt genomes is currently not feasible, transcriptome sequencing is the method of choice to provide thousands of SNPs to be examined for signals of interspecific gene flow, using models based on the allele frequency spectrum (Gutenkunst *et al.* 2009; Ellison *et al.* 2011). Signals of older gene flow may be effectively extracted using gene genealogies even if the signal based on geographic distribution of shared variation close to the areas of hybridization has dissipated (Pinho & Hey 2010; Marko & Hart 2011).

## Conclusions

Our study shows that introgression, a consequence of natural hybridization between the two closely related species of newts, *Lm* and *Lv*, differs dramatically between the mitochondrial and the nuclear genome. While mtDNA replacement in *Lm* is complete, resulting from multiple, spatially and temporally distinct introgression events from *Lv*, there is little evidence of recent interspecific nuclear gene flow as assessed by 15 microsatellite loci. Microsatellite differentiation within *Lm* defines three genetic units, which most likely correspond to three glacial refugia, located in the northern, eastern and southern part of the Carpathians. *In situ* survival and range fragmentation of *Lm* are supported by the results of species distribution modelling, adding important evidence to the emerging view of the Carpathians as a major refugial area. The results of the present study, in combination with previous reports of extensive introgression of MHC genes, emphasize the complexity of historical gene exchange between *Lm* and *Lv*.

## Acknowledgements

We are grateful to Maciek Bonk, Magda Herdegen, Marcin Liana and Jacek Radwan who helped in sample collection and to Michał Stuglik for help in data analysis. Three anonymous reviewers provided valuable comments. The work was supported by the Polish National Science Center grant nr 8171/B/P01/2011/40 to WB and by the Jagiellonian University grants: DS/MND/WBiNoZ/INoŚ/27/2012 to PZ and DS/WBiNoZ/INoŚ/762/11; KN is the recipient of a 2011 Annual Stipend for Young Scientists START from the Foundation for Polish Science.

## References

Abi-Rached L, Jobin MJ, Kulkarni S *et al.* (2011) The shaping of modern human immune systems by multiregional admixture with archaic humans. *Science*, **334**, 89–94.

Avise JC, Walker D, Johns GC (1998) Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society. B, Biological Sciences*, **265**, 1707–1712.

Babik W, Szymura JM, Rafinski J (2003) Nuclear markers, mitochondrial DNA and male secondary sexual traits variation in a newt hybrid zone (*Triturus vulgaris* × *T. montandoni*). *Molecular Ecology*, **12**, 1913–1930.

Babik W, Branicki W, Crnobrnja-Isailovic J *et al.* (2005) Phylogeography of two European newt species—discordance between mtDNA and morphology. *Molecular Ecology*, **14**, 2475–2491.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Barton N, Bengtsson BO (1986) The barrier to genetic exchange between hybridising populations. *Heredity*, **57**, 357–376.

Beebee TJC (1996) *Ecology and Conservation of Amphibians*. Chapman & Hall, London.

Braconnot P, Otto-Bliesner B, Harrison S *et al.* (2007) Results of PMIP2 coupled simulations of the mid-holocene and last glacial maximum—part 1: experiments and large-scale features. *Climate of the Past*, **3**, 261–277.

Buckley LB, Jetz W (2008) Linking global turnover of species and environments. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 17836–17841.

Castric V, Bechsgaard J, Schierup MH, Vekemans X (2008) Repeated adaptive introgression at a gene under multiallelic balancing selection. *PLoS Genetics*, **4**, e1000168.

Chan KMA, Levin SA (2005) Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution*, **59**, 720–729.

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular biology and evolution*, **24**, 621–631.

Collins WD, Bitz CM, Blackmon ML *et al.* (2006) The Community Climate System Model Version 3 (CCSM3). *Journal of Climate*, **19**, 2122–2143.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer, Sunderland.

Crawford NG (2010) Smogd: software for the measurement of genetic diversity. *Molecular Ecology Resources*, **10**, 556–557.

Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. *Evolution*, **62**, 1908–1920.

Duvaux L, Belkhir K, Boulesteix M, Boursot P (2011) Isolation and gene flow: inferring the speciation history of European house mice. *Molecular Ecology*, **20**, 5248–5264.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330–342.

Ellison CE, Hall C, Kowbel D *et al.* (2011) Population genomics and local adaptation in wild isolates of a model microbial eukaryote. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 2831–2836.

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

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Fijarczyk A, Nadachowska K, Hofman S *et al.* (2011) Nuclear and mitochondrial phylogeography of the European firebellied toads *Bombina bombina* and *Bombina variegata* supports their independent histories. *Molecular Ecology*, **20**, 3381–3398.

Fitzpatrick BM, Johnson JR, Kump DK *et al.* (2009) Rapid fixation of non-native alleles revealed by genome-wide SNP

analysis of hybrid tiger salamanders. *BMC Evolutionary Biology*, **9**, 176.

Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 397–423.

Good JM, Hird S, Reid N *et al.* (2008) Ancient hybridization and mitochondrial capture between two species of chipmunks. *Molecular Ecology*, **17**, 1313–1327.

Goudet J (2001) FSTAT, version 2.9.3, a program to estimate and test gene diversities and fixation indices. Available from http://www.unil.ch/izea/softwares/fstat.html.

Grant PR, Grant BR, Petren K (2005) Hybridization in the recent past. *The American Naturalist*, **166**, 56–67.

Green RE, Krause J, Briggs AW *et al.* (2010) A draft sequence of the neandertal genome. *Science*, **328**, 710–722.

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, **5**, e1000695.

Hewitt G (2000) The genetic legacy of the quaternary ice ages. *Nature*, **405**, 907–913.

Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, **27**, 905–920.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 2785–2790.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. *BMC Genetics*, **6**, 13.

Johanet A, Picard D, Garner TWJ *et al.* (2009) Characterization of microsatellite loci in two closely related Lissotriton newt species. *Conservation Genetics*, **10**, 1903–1906.

Joly P, Grolet O (1996) Colonization dynamics of new ponds, and the age structure of colonizing alpine newts, *Triturus alpestris*. *Acta Oecologica*, **17**, 599–608.

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, **20**, 229–237.

Marko PB, Hart MW (2011) The complex analytical landscape of gene flow inference. *Trends in Ecology and Evolution*, **26**, 448–456.

Martinsen GD, Whitham TG, Turek RJ, Keim P (2001) Hybrid populations selectively filter gene introgression between species. *Evolution*, **55**, 1325–1335.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F ST and related measures. *Molecular Ecology Resources*, **11**, 5–18.

Melo-Ferreira J, Boursot P, Carneiro M *et al.* (2012) Recurrent introgression of mitochondrial DNA among hares (*Lepus* spp.) revealed by species-tree inference and coalescent simulations. *Systematic Biology*, **61**, 367–381.

Michalak P, Rafinski J (1999) Sexual isolation between two newt species, *Triturus vulgaris* and *T. montandoni* (Amphibia, Urodela, Salamandridae). *Biological Journal of the Linnean Society*, **67**, 343–352.

Michalak P, Grzesik J, Rafinski J (1997) Tests for sexual incompatibility between two newt species, *Triturus vulgaris* and *Triturus montandoni*: no-choice mating design. *Evolution*, **51**, 2045–2050.

Miller W, Schuster SC, Welch AJ *et al.* (2012) Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. *Proceedings of the National Academy of Sciences*, **109**, E2382–E2390.

Mims MC, Hulsey CD, Fitzpatrick BM, Streelman JT (2010) Geography disentangles introgression from ancestral polymorphism in Lake Malawi cichlids. *Molecular Ecology*, **19**, 940–951.

Mraz P, Gaudeul M, Rioux D *et al.* (2007) Genetic structure of *Hypochaeris uniflora* (Asteraceae) suggests vicariance in the Carpathians and rapid post-glacial colonization of the Alps from an eastern Alpine refugium. *Journal of Biogeography*, **34**, 2100–2114.

Nadachowska K, Babik W (2009) Divergence in the face of gene flow: the case of two newts (Amphibia: Salamandridae). *Molecular Biology and Evolution*, **26**, 829–841.

Nadachowska K, Flis I, Babik W (2010) Characterization of microsatellite loci in the Carpathian newt (*Lissotriton montandoni*). *The Herpetological Journal*, **20**, 107–110.

Nadachowska-Brzyska K, Zielinski P, Radwan J, Babik W (2012) Interspecific hybridization increases MHC class II diversity in two sister species of newts. *Molecular Ecology*, **21**, 887–906.

Neafsey DE, Lawniczak MKN, Park DJ *et al.* (2010) SNP genotyping defines complex gene-flow boundaries among African malaria vector mosquitoes. *Science*, **330**, 514–517.

Nevado B, Fazalova V, Backeljau T, Hanssens M, Verheyen E (2011) Repeated unidirectional introgression of nuclear and mitochondrial DNA between four congeneric tanganyikan cichlids. *Molecular Biology and Evolution*, **28**, 2253–2267.

Nolte AW, Gompert Z, Buerkle CA (2009) Variable patterns of introgression in two sculpin hybrid zones suggest that genomic isolation differs among populations. *Molecular Ecology*, **18**, 2615–2627.

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

Pauls SU, Lumbsch HT, Haase P (2006) Phylogeography of the montane caddisfly *Drusus discolor*: evidence for multiple refugia and periglacial survival. *Molecular Ecology*, **15**, 2153–2169.

Payseur BA (2010) Using differential introgression in hybrid zones to identify genomic regions involved in speciation. *Molecular Ecology Resources*, **10**, 806–820.

Peterson AT (2011) Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography*, **38**, 817–827.

Petit RJ, Excoffier L (2009) Gene flow and species delimitation. *Trends in Ecology and Evolution*, **24**, 386–393.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Pinho C, Hey J (2010) Divergence with gene flow: models and data. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 215–230.

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

Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, **23**, 564–571.

Raes N, ter Steege H (2007) A null-model for significance testing of presence-only species distribution models. *Ecography*, **30**, 727–736.

Ronikier M (2011) Biogeography of high-mountain plants in the Carpathians: an emerging phylogeographical perspective. *Taxon*, **60**, 373–389.

Rousset F (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Schierup MH, Vekemans X, Charlesworth D (2000) The effect of subdivision on variation at multi-allelic loci under balancing selection. *Genetical Research*, **76**, 51–62.

Smith MA, Green DM (2005) Dispersal and the metapopulation paradigm in amphibian ecology and conservation: are all amphibian populations metapopulations? *Ecography*, **28**, 110–128.

Sommer RS, Nadachowski A (2006) Glacial refugia of mammals in Europe: evidence from fossil records. *Mammal Review*, **36**, 251–265.

Song Y, Endepols S, Klemann N *et al.* (2011) Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. *Current Biology*, **21**, 1296–1301.

Sun JX, Mullikin JC, Patterson N, Reich DE (2009) Microsatellites are molecular clocks that support accurate inferences about history. *Molecular Biology and Evolution*, **26**, 1017–1027.

Svenning J-C, Fløjgaard C, Marske KA, Nógues-Bravo D, Normand S (2011) Applications of species distribution modeling to paleobiology. *Quaternary Science Reviews*, **30**, 2930–2947.

Teeter KC, Thibodeau LM, Gompert Z *et al.* (2010) The variable genomic architecture of isolation between hybridizing species of house mice. *Evolution*, **64**, 472–485.

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907–3930.

Ursenbacher S, Carlsson M, Helfer V, Tegelstrom H, Fumagalli L (2006) Phylogeography and Pleistocene refugia of the adder (*Vipera berus*) as inferred from mitochondrial DNA sequence data. *Molecular Ecology*, **15**, 3425–3437.

VanDerWal J, Shoo LP, Graham C, Williams SE (2009) Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecological Modelling*, **220**, 589–594.

Vences M, Wake DB (2007) Speciation, species boundaries and phylogeography of amphibians. In: *Amphibian Biology, Vol. 6, Systematics* (eds Heatwole H & Tyler M), pp. 2613–2669. Surrey Beatty & Sons, Chipping Norton, Australia.

Vieites DR, Nieto-Roman S, Wake DB (2009) Reconstruction of the climate envelopes of salamanders and their evolution through time. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19715–19722.

Vines TH, Kohler SC, Thiel M *et al.* (2003) The maintenance of reproductive isolation in a mosaic hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*. *Evolution*, **57**, 1876–1888.

Warren DL, Glor RE, Turelli M (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, **33**, 607–611.

Weisrock DW, Macey JR, Ugurtas IH, Larson A, Papenfuss TJ (2001) Molecular phylogenetics and historical biogeography among salamandrids of the "true" salamander clade: rapid branching of numerous highly divergent lineages in *Mertensiella luschani* associated with the rise of Anatolia. *Molecular Phylogenetics and Evolution*, **18**, 434–448.

Whitlock MC (2011) $G_{ST}$" and D do not replace $F_{ST}$. *Molecular Ecology*, **20**, 1083–1091.

Whitney KD, Randell RA, Rieseberg LH (2006) Adaptive introgression of herbivore resistance traits in the weedv sunflower *Helianthus annuus*. *The American Naturalist*, **167**, 794–807.

Wirtz P (1999) Mother species-father species: unidirectional hybridization in animals with female choice. *Animal Behaviour*, **58**, 1–12.

Zavadil V, Pialek J, Dandova R (2003) *Triturus montandoni* (Boulenger, 1880)—Karpatenmolch. In: *Handbuch der Reptilen und Amphibien Europas* (ed. Bohme W), pp. 657–706. Aula, Wiebelsheim.

Zeng YF, Liao WJ, Petit RJ, Zhang DY (2011) Geographic variation in the structure of oak hybrid zones provides insights into the dynamics of speciation. *Molecular Ecology*, **20**, 4995–5011.

---

P.Z., a PhD student at the Jagiellonian University, is interested in the patterns of historical and contemporary gene flow and intraspecific differentiation between diverging populations; a part of this study constituted his MSc project; K.N.-B. is interested in the genomics of speciation; currently she is a postdoc at the Uppsala University studying speciation genomics in birds; B.W. is interested in speciation and gene flow and uses newts as a model system; R.S. is a biotechnology student at the Jagiellonian University; a part of this study constituted his BSc project; S.D.C.-M. is interested in zoogeography and ecology of Romanian herpetofauna; D.C. is a Professor at the University Ovidius Constanta interested in the life history, evolution and conservation of amphibians; W.B. uses molecular tools to study historical processes at and below the species level.

---

## Data accessibility

DNA sequences: GenBank accessions KC297508–KC297565.

IMa2 input files, microsatellite data and coordinates of localities and bioclimatic data used for species distri-

bution modelling: DRYAD entry doi: 10.5061/dryad.g216n.

## Supporting information

Additional supporting information may be found in the online version of this article.

**Fig. S1** Allelic richness (light blue) and the private allele richness (deep blue) in *L. montandoni* populations.

**Fig. S2** Two-dimensional scaling of the matrix of pairwise $F_{ST}$ between *L. montandoni* populations.

**Fig. S3** $\Delta K$, a measure of second-order rate of change in the likelihood of $K$ for various $K$ values; *L. montandoni* populations, 15 microsatellite loci.

**Fig. S4** The relationship between the straight-line geographic distance and genetic differentiation ($G_{ST}''$) for all *L. montandoni* populations.

**Fig. S5** The relationship between the straight-line geographic distance and genetic differentiation ($G_{ST}''$) for the northern, eastern and southern *L. montandoni* genetic clusters.

**Fig. S6** Two-dimensional scaling of the matrix of pairwise $F_{ST}$ between *L. montandoni* and *L. vulgaris* populations; 10 microsatellite loci.

**Fig. S7** $\Delta K$, a measure of second-order rate of change in the likelihood of $K$ for various $K$ values; *L. montandoni* and *L. vulgaris* populations, 10 microsatellite loci.

**Fig. S8** Two-dimensional scaling of the matrix of pairwise $G_{ST}''$ (A) and $F_{ST}$ (B) between *L. montandoni* and *L. vulgaris* populations; three microsatellite loci showing little interspecific differentiation.

**Table S1** (a). Results of the tests of the conformance of genotype frequencies to Hardy–Weinberg expectations; uncorrected $P$ values are given. (b) Results of the tests of the conformance of genotype frquencies to Hardy–Weinberg expectations.

**Table S2** Frequencies of null alleles estimated by FreeNA for each locus and each population.

**Table S3** Allelic richness for each locus and each population.

**Table S4** Pairwise $G_{ST}''$ values for *L. montandoni* populations; 15 microsatellite loci.

**Table S5** Pairwise $F_{ST}$ values for *L. montandoni* populations; 15 microsatellite loci; values not significantly different from zero in red.

**Table S6** Pairwise $G_{ST}''$ values for *L. montandoni* and *L. vulgaris* populations; 10 microsatellite loci.

**Table S7** Pairwise $F_{ST}$ values for *L. montandoni* and *L. vulgaris* populations; 10 microsatellite loci; values not significantly different from zero are in red.

**Table S8** mtDNA variation in populations.

**Table S9** Pairwise $F_{ST}$ betwen *L. montandoni* and *L. vulgaris* populations calculated for mtDNA sequences; populations with a minimum of four available sequences were included.