# Cladogenesis of the European brown hare (Lepus europaeus Pallas, 1778)

Joerns Fickel & Heidi C. Hauffe & Elena Pecchioli & Ramon Soriguer & Ljiljana Vapa & Christian Pitra

Abstract A substantial portion of today's biodiversity is attributed to the climatic oscillations of the Pleistocene Ice Ages. Gradual but dramatic climate changes were accompanied by expansion, contraction, and isolation of populations, promoting the accumulation of genome differences and adaptations in refugial populations and resulting in allopatric differentiation in a variety of taxa. In the present study, partial mitochondrial DNA sequences of the widely distributed European brown hare (Lepus europaeus) were analyzed to test whether the species' present genetic structure is the result of postglacial re-colonization of Europe from Asia Minor (clade A) and the Balkans (clade B) only, as suggested previously, or if additional refugia are

Electronic supplementary material The online version of this article (doi:10.1007/s10344-008-0175-x) contains supplementary material, which is available to authorized users.

C ommunicated by W. Lutz

J. Fickel (\*): C. Pitra Department of Evolutionary Genetics, Leibniz-Institute for Zoo and Wildlife Research, Alfred-Kowalke-Str. 17, 10315 Berlin, Germany e-mail: Fickel@IZW-Berlin.de

H. C. Hauffe E. Pecchioli Centro di Ecologia Alpina, 38040 Viote del Monte, Bondone, Italy

R. Soriguer Especies Cinegeticas y Plaga, Estacion Biologica de Donaña C.S.I.C., Apartado 1056, 41013 Seville, Spain

L. Vapa Department of Genetics, Trg Dositeja Obradovica 2, 21000 Novi Sad, Serbia likely. Analyses indicated the presence of an additional refugium (Italy, clade I). The genealogic network of Italian hares displayed the tree-like structure expected from refugial populations, whereas central European brown hare haplotypes revealed a clear star-phylogeny indicative of past-bottleneck population growth. This population size expansion, which was confirmed by mismatch analysis, was estimated to have occurred ~50–55 thousand years ago (kya). The divergence of clade A\* from the remaining matrilines is estimated at 239 kya, whereas the divergence of the ancestors of clades B\* and I from A\* occurred about 128 kya.

Keywords Biogeography Mitochondrial DNA Glacial refugia

# Abbreviations

- HT haplotype
- CE central Europe
- I Italy
- ky thousand years
- kya thousand years ago
- A\* clade A with additional sequences
- B\* clade B with additional sequences

## Introduction

Pleistocene ice ages had a profound influence on today's biodiversity (Haffer 1969; Hewitt 1996, 1999, 2004; Roy et al. 1996). Not only did glaciers reshape the landscape and reduce the inhabitable land mass, but repeated glacier contractions and expansions (Adams et al. 1999; Dynesius and Jansson 2000) were paralleled by range changes in many taxa, resulting in subdivision and isolation of

populations and in accumulation of genome differences and adaptations between them (e.g., Cooper et al. 1995; Taberlet et al. 1998 and references therein; Fickel et al. 2007a; Frölich et al. 2007). In the last few decades, many biogeographic studies have been carried out to investigate the influences of Quarternary cold periods on intraspecific genetic structure patterns of European terrestrial vertebrate biota (e.g., Cooper et al. 1995; Suchentrunk et al. 2000; Fedorov and Stenseth 2001; Brito 2005; Djan et al. 2006), because such studies gain increasing importance for species conservation issues (Cook et al. 2001; Fickel et al. 2007b). Two principal scenarios were proposed to account for present patterns of genetic variability in these species: (a) postglacial re-colonization from periglacial refugial areas and (b) survival in and dispersal from local refugia within glaciated regions. For the first scenario, three main European refugia were suggested: the Iberian peninsula, the Italian (Appenine) peninsula, and the Balkans/Asia Minor (Taberlet et al. 1998; Aksu et al. 1999; Hewitt 1996, 1999; Branco et al. 2002). For the second scenario, various local refugia were suggested. Although initially controversial (Tzedakis et al. 2002, 2003; Stewart 2003), there is increasing evidence (Lister 2004) that the latter scenario is applicable to a number of temperate northern hemisphere species, including plants (Willis and Whittaker 2000; Tzedakis et al. 2002), small mammals such as pygmy and common shrew, bank and root vole, long-tailed field mouse, Eurasian red squirrel, and artic lemming (e.g., Bilton et al. 1998; Brunhoff et al. 2003; Stewart and Lister 2001), and large mammals such as red deer (Lister 1984). These two scenarios are expected to have profoundly different genetic consequences. For example, populations that re-colonized the European continent from the same refugial source populations following the same glacial period should be genetically similar (among others depending on time since expansion, population bottlenecks, and impact of genetic drift), both to each other and to populations that remained in the periglacial refugia from which they had spread. However, (small) populations of species that survived in buffered local microclimates between the ice sheets are expected to be highly differentiated due to genetic drift (Hewitt 1996, 1999; Bilton et al. 1998; Stewart and Lister 2001; Lister 2004).

Pleistocenic climatic oscillations are believed to have profoundly affected the demographics of the European brown hare (Lepus europaeus). This medium-sized mammal evolved in a temperate, steppe-like climate but is now associated with and adapted to agricultural landscapes. Presently, the species is distributed throughout Palearctic, Nearctic, Oriental, and Neotropical regions. It has also been introduced by humans to Australia, New Zealand, southern Siberia, the Far East, and to North and South America (Dietrich 1984, 1985). The species is found throughout the entire European continent (Lincoln 1974), from the northern shores of the Mediterranean Sea to 60°N latitude in Scandinavia (Thulin et al. 1997).

Recently, it has been suggested that the contemporary genetic population structure of the European brown hare is attributed to a postglacial spread from the Balkans and Asia Minor (Suchentrunk et al. 2000; Kasapidis et al. 2005). However, preliminary analyses of mitochondrial DNA (mtDNA) haplotypes (HTs) of the European brown hare from western and central Europe (Fickel et al. unpublished) are inconsistent with this interpretation. Thus, the present study investigates the HT distribution in the European brown hare metapopulation to evaluate the fit of the demographic scenario suggested above and to identify possible additional retreat areas from which the brown hare may have spread postglacially. Such additional refugia could include Iberia, Italy, and/or more cryptic microrefugia (Stewart and Lister 2001).

Disentangling the demographic history of this species is particularly challenging, because the immense interest in the brown hare as a game animal has resulted in extensive translocations across Europe for restocking hunted or diseased animals and/or for outbreeding purposes (Fraguglione 1963, 1971; Schneider 1978; Spittler 1996; Pierpaoli et al. 1999; Gortázar et al. 2007), including the potential refugial areas of Italy and Spain (Pierpaoli et al. 1999; Soriguer, pers. comm.). Between 1900 and 1939, hares ware also translocated to Germany but on a much smaller scale; for example, in 1936/37, Germany imported 8,259 live hares (with an unknown number of released hares) from Poland, the Czech Republic, and Hungary (Spittler 1996).

## Materials and methods

Tissue and/or blood was sampled from 820 brown hares from 97 sites across Europe: Spain (n=16, six sites), Italy (n=46, 12 sites), Poland (n=23, five sites), Sweden (where L. europaeus was introduced in the 19th century; Lönnberg 1905; n=10, ten sites), Serbia (n=46, 11 sites), Slovakia (n=99, 11 sites), and Germany (n=580, 42 sites; Fig. 1, 1)Table S1). Wherever possible, samples were taken from hares bagged in remote areas, and care was taken to avoid sampling from populations that have been recently and/or obviously restocked; tagged individuals were not sampled. DNA was extracted from either 200 µl whole blood or 30mg tissue using the DNeasy® spin columns (Qiagen, Germany). A 5'-end fragment (413 bp) of the variable mtDNA control region (d-loop) was amplified as described previously (Fickel et al. 1999) using primers LeCtrlH (5'-GGGCGATCTTAGGGTTATGG-3') and LeCtrL (5'-ATC CAAGTAACTTGTCACTATTG-3'). Amplicons were se-

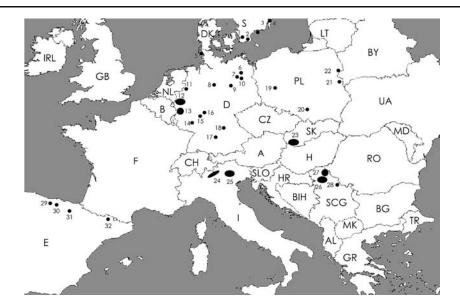


Fig. 1 Map of Lepus europaeus collection sites and haplotypes found (see Table S1). Abbreviations are \$ (Sweden): 1 Genarp c90; Kageröd c7; Svalöv c7; 2 Glemingebro c94; Hammenhög c9; Skivarp c72; Tomelilla c9; Ystad c9; 3 Ementorp; 4 Södra Sandby c89; D (Germany): 5 Föhr c9, c10, c110-c112; 6 Koppelberg c32; 7 Bullenwinkel c91, c92; Rathenow c17; Seeschlag c91; 8 Bierbergen c7, c17, c42, c59, c98; Rötzum c7, c98; 9 Aschersleben c1, c7, c56; Hettstedt c7, c17, c44, c57-c60, c65, c66; Straßfurt c59, c63, c64; 10 Winkel c32; 11 Burgsteinfurt c17, c25; 12 Anholt c7, c15, c17, c18; Boisheim c2, c7, c13, c29-c32; Buldern c17-c19, c23, c34; Dormagen c7, c14, c17, c19; Heinsberg c5, c6, c7, c11, c34; Kleinenbroich c7, c21, c22, c24; Orsoy c7, c11, c17; Ossenberg c2, c7, c11, c16; Roxelbrock c11, c18, c19, c23; Stommeln c15, c20, c26; Wachtendonk c5-c7, c11, c65, c67-c70; Walbeck c7, c13, c38; 13 Bedburg c5, c13, c36; Dünstekoven c4, c5, c7-c10; Heimerzheim c5, c7-c11, c17, c22, c37, c71; Hellenburg c5, c7, c35; Müddersheim c7, c11; Ollheim c5-c7, c12; 14 Biebesheim c5, c7, c54, c61; Meckenheim c7, c8, c33; 15 Hattersheim c7, c50, c53; Leeheim c5, c7, c19, c47, c62; Rendel c7; 16 Goddelau c55; Mensfelden c45; Oberhone c7, c16, c46, c48, c49, c51, c52; 17 Marxheim c7, c49; 18 Rummelsberg c32; PL (Poland): 19 Czempiń c1, c3, c17, c28, c39-

quenced bidirectionally (BigDye<sup>™</sup> cycle sequencing kit, ABI, Darmstadt, Germany) and visualized on an ABI model 3100 automated sequencer. To devise the analysis as pan-European as possible, we consulted GenBank and added 72 mtDNA control region sequences of 98 European brown hares sampled in Greece, Bulgaria, and several Mediterranean islands (Kasapidis et al. 2005, GenBank accession nos.: AY466782–AY466853). All sequences were aligned using ClustalX (v.1.8, Thompson et al. 1997).

Median-joining (MJ) networks were constructed using Network (v.4.1.1.1, Bandelt et al. 1999). Because its HT joining procedure lacks a statistical significance test, we also generated a parsimonious HT cladogram (Templeton et al. 1992; Templeton 1998) using Tcs (v.1.21, Clement et al. 2000), setting gaps as a fifth state and using a parsimony criterion of 95% (Templeton et al. 1992). Tree- and star-like phylogenetic proportions were estimated using likelihood mapping analysis (Strimmer and von Haeseler 1997)

c43; 20 Skadla c2; 21 Mińsk Mazowiecki c7, c105; 22 Białowieza c1, c17, c19, c100, c103-c105; Bielsk Podlaski c102; SK (Slovakia): 23 Hájske c2, c8, c17, c67; Čiližská Radvaň c8, c65, c97; Dlhá nad Váhom c2, c65, c87, c101; Lehnice c2, c8, c17, c65, c67, c95, c99, c107; Močenok c2, c17, c65, c96, c107; Neded c2, c8, c17, c30, c96, c97; Nitra c17; Svätoplukovo c2, c17, c106; Tešedíkovo c2, c30, c32, c65, c97, c108, c109; Trnava c2, c8, c97, c109; Veca c2, c65; I (Italy): 24 Sondrio i16, i22, i23; 25: Arco i8, i21; Brentonico i9, i14; Castello Molina i19; Civezzano i7, i11, i21; Folgaria i6, i13, i15, i19; Mezzocorona i4, i10, i19; Predazzo i20; Primiero i17; Rovereto i1, i13, i18; Trento South i2; Trento North i3, i5; Trento West i12, i13; SCG (Serbia and Montenegro): 26 Čurug c7, c69, c72, c73, c81-c83; Futog c69; Kač c2, c17, c30, c33, c80, c86; Kovilj c7; Krnjesevci c76, c77, c84, c87; Kulpin c8, c69, c85; Žarcovac c75, c78, c79, c87, c88; 27 Kikinda c7, c33; Novo Miloševo c7, c30, c33, c72; Padej c72; 28 Hajdučica c33; E (Spain): 29 Brañosera c2; 30 Espinosa de los Monteros c2; Puerto de los Tornos c2; 31 Navarra Region c7, c17, c74, c97; 32 Bagerque Salardu c2; Son del Pino c2. Abbreviations follow the international country codes at http://users.pandora.be/ worldstandards/carcodes.htm

implemented in Tree-Puzzle (v.5.2, Schmidt et al. 2002, 2004). Support for proportions was assessed by a reliability percentage after 10,000 quartet puzzling steps. Genetic differentiation among all potential clades was tested using Arlequin, which also considers different population sizes by a resampling approach (v.3.0b, Excoffier et al. 2005). To obtain both mean genetic distances within and among clades as well as confidence intervals (CI 95%) around these distances, we used CIProgram (v.1.01b, C.E.McIntosh, <http://www.mnh.si.edu/GeneticsLab/Technical-Page/TechnicalPage.html>) following the method of Steel et al. (1996). To allow the use of a more complex, but probably more realistic nucleotide substitution model than the software-implemented single model of Jukes & Cantor (JC69) correction, we calculated a correction factor as the ratio of the JC69-based mean genetic distance and the mean genetic distance, which was estimated using Mega (v.3.1, Kumar et al. 2004) by applying the

substitution model parameters estimated by Paup\*/Modeltest (Swofford 2002; Posada 2003; see below). All mean divergences were also corrected for lineage sorting by obtaining the net distance ( $\delta$ ) between clades, calculated by subtracting the mean within-clade distances from the mean inter-clade distances (and their CI95%) using the equation:  $\delta \frac{1}{4} \delta_{XY} - 1=2\delta \delta_X \Rightarrow 0$  (Nei 1975, 1987; Wilson et al. 1985; Steel et al. 1996).

To reduce sampling bias in calculations of genetic differentiation among clades (most HTs in clades A and B of Kasapidis et al. (2005) represent one hare only), we also calculated  $F_{ST}$  values by setting all HT frequencies (f) to f=1. In addition, we reduced the influences of (a) an excessively large number of very similar HTs in a starphylogeny and (b) homoplasies in numerous unresolved inter-haplotypic relationships by calculating  $F_{ST}$  values for such groups of HTs that had previously been identified to be essential for network construction using Network's "torso"-function ( $F_{ST(torso)}$ ); these HTs were then used to construct a parsimonious torso-cladogram (Templeton et al. 1992). Parameters required for Mega, Arlequin, and Tree-Puzzle calculations, such as model of nucleotide substitution and site mutation rate heterogeneity, were estimated using a combination of Paup\* (v.4.b10, Swofford 2002) and Modeltest (v.3.6, Posada and Crandall 1998; Posada 2003). The Akaike information criterion (AIC; Akaike 1974), a function of the maximized log-likelihood and the number of estimable parameters, was likewise calculated using Paup\*/Modeltest. In comparison to the commonly used hierarchical likelihood ratio test (hLRT), AIC has the advantage to allow model averaging (inclusion of nested into general models) and assessment of model selection uncertainty (Posada and Buckley 2004). All the following computations were performed using Arlequin: HT and nucleotide diversity, numbers of transitions (Ts) and transversions (Tv), mean number of pairwise nucleotide differences (K) and their variances (V) within clades, and genetic distances among clades following the model of nucleotide substitution recommendations by Paup\*/Modeltest. Neutrality of mutations within the control region, necessary for the assumption of a constant mutation rate  $\mu$  (for limitations see Ford 2002), was previously demonstrated (Pierpaoli et al. 1999; Fickel et al. 2005; Kasapidis et al. 2005).

In order to explore the demographic history of the central European and Italian hare populations, we followed several approaches. First, we calculated the number of generations t since potential expansion using a simplified estimation method with post-expansion population size  $\theta_1$  set to infinity (Rogers 1995): Given the observed mean of pairwise differences ( $\kappa$ ) and its variance (v), pre-expansion population size ( $\theta_1$ ) and the estimated as  $q_0 \frac{1}{4}$  in  $-k \triangleright$  and t  $\frac{1}{4} k - q_0 \frac{1}{4}$  2mt. To avoid an underestimation of expansion time  $\tau$  (Schneider

and Excoffier 1999), we analyzed the mismatch distribution, which is the distribution of numbers of site differences between each pair of sequences in the population (Li 1977; Rogers and Harpending 1992; Rogers 1995) using a generalized non-linear least-square approach (Excoffier et al. 2005). Displayed graphically, recent sudden population expansions or bottlenecks generate unimodally shaped graphs, whereas stable or slowly declining populations generate a variety of multimodal distributions, reflecting the highly stochastic shape of gene-trees within populations at demographic equilibrium (Rogers and Harpending 1992). The method is based on the infinite-site-model, assuming a stepwise expansion from a relatively small but constant population to a large population over t generations (mutational units) in the past:  $t=T/2\mu$  (T: time to expansion,  $\mu$ : mutation rate). The sum of squared deviations (SSD; Schneider and Excoffier 1999) was computed to test goodness-of-fit of the observed mismatch distribution to that expected under the sudden population expansion model. Parametric bootstrapping (10,000 permutations) was applied to obtain confidence intervals around all estimated parameters (Schneider and Excoffier 1999; Excoffier et al. 2005). Generation time in hares was set to 2 years (Marboutin and Peroux 1995). Although seemingly low for a medium-sized mammal with several leverets per litter and several litters per year, it considers the low survival rate of juveniles (Marboutin and Peroux 1995; Spittler 1996). As a second approach to explore the demographic history of hare populations, Fu's F<sub>s</sub> statistics (Fu 1997; 1,000 simulations), likewise based on the infinite-site model, were computed to detect an excess of low-frequency alleles in a growing population as compared with the expected number of alleles in a stationary population, whereby significantly large negative values indicate population expansion (Fu 1997; Excoffier and Schneider 1999). Because assessment of demographic history based on mismatch distribution may be biased in samples with high  $N_{haplotypes}/N_{individuals}$  ratios (R) as seen in several clades, we also assessed the mean within-clade HT coalescence time by dividing the mean within-clade genetic distance (calculated following the Paup\*/Modeltest recommendations and the AIC-estimated gamma distribution shape parameter  $\alpha$ ) and an evolution rate for the hare mtDNA control region of 12.4% per million years (Pierpaoli et al. 1999).

# Results

## General analysis

Within the central European (CE) data set (n=807), 112 different mtDNA HTs were identified, 73 of which were new and deposited in the GenBank database (GenBank

accession nos.: AY103494-531, AY154661-6, AY163356-76, EU435412-14). The Italian (I) sample (n=46) consisted of 25 HTs (accession nos.: EU435387-411), one of which (i6) was identical to central European HT c8. The overlapping region between our data set and the sequences retrieved from GenBank (Kasapidis et al. 2005) was 335 bp. These 335-bp fragments, which were used for all subsequent analyses, reduced the number of distinct HTs in I to 17, to 93 in CE and to a total of 70 for the Asia Minor and Balkan clades (A and B, respectively, Kasapidis et al. 2005). Variable positions in the HT sequences of CE+I. compared with the L. europeaus mtDNA genome (Arnason et al. 2002; GenBank accession no.: AJ421471), were identified using Mega and are listed together with absolute HT frequencies in Table S2. MJ-network analysis suggested that six HTs from I and three HTs from CE should be assigned to clades A and B (Kasapidis et al. 2005); thus, the clades A and B with our additional sequences are henceforth called A\* and B\*. The networks/cladograms of all HTs generated by Network and Tcs generated the same cladistic structure ( $A^*$ ,  $B^*$  and CE+I), except that Tcs excluded three small groups of HTs from clade B\* by its 95% parsimony threshold: {c88, B54}, {B52-53}, and {B55-56, c96, i18i22}. It also connected clade B\* and phylogroup {CE+I} by two branches: three substitutions linking HTs c11 and B39 as well as c24 and B49. The latter connection was not recognized by the MJ-network due to a software-introduced median vector. When we applied the 95% parsimony threshold in Tcs, the maximum number of nucleotide differences between pairs of HTs ("step limit") was estimated as 7. Because clades A\* and B\* were separated by more than seven differences, clade A\* was excluded from the cladogram.

Transition/transversion ratios (Ts/Tv) varied considerably among hare lineages. The highest ratios were found in CE (Ts/Tv=21.3) and A\* (Ts/Tv=12.7), while the ratio in all other groups was substantially lower, ranging from 3 to 4.7 (Table 3). Nucleotide diversity ( $\pi$ ) across clades/groups was lowest in CE (0.89%) and highest in B\* (3.58%).

# Italian clade (clade I)

The 46 Italian hares shared 25 HTs with a nucleotide diversity of  $\pi$ =1.5%, a Ts/Tv ratio of 4.67 and a ratio of number of HTs per number of animals of R=0.543 (Table 3). The hLRT and AIC approach favored different models of nucleotide substitution and substitution rate heterogeneity; while hLRT favored the HKY model (Hasegawa et al. 1985) with a gamma distribution shape parameter of  $\alpha$ =0.0143, the AIC advocated the TrN model (Tamura and Nei 1993) with equal substitution rates for all sites and a high proportion of invariable sites (84%). Using the latter test (for rationale, see Posada and Buckley 2004),

we estimated a within-clade mean genetic distance of 1.253%, which, given a mutation rate of 12.4% per million years (Pierpaoli et al. 1999), results in a HT coalescence time of approximately 101 ky. The MJ-network (Bandelt et al. 1999) revealed a clearly tree-like phylogeny, comprising three main HT groups, whose shortest inter-distance was ten mutations (Fig. 2). Likelihood mapping (Schmidt et al. 2002, 2004) estimated the overall tree-like phylogeny proportions as being well above 50% (77.2%), as expected for stable (and potentially old refugial) populations (Strimmer and von Haeseler 1997). The long branches (≥ten mutations) between the three main groups of the MJ-network of the Italian HTs indicated a different phylogenetic history for each of these groups. Alignment and joint MJ-network construction with HTs from clades A and B demonstrated the affiliation of six Italian HTs to these latter clades (i18-i22 to clade B; i23 to clade A, Table S1, Fig. 3). After reassignment of the above HTs to their more appropriate clades A\* and B\* and the collapse of two Italian HTs due to the fragment length reduction to 335 bp, the remaining 17 Italian HTs were re-analyzed using likelihood mapping. The phylogeny was still rather tree-like (56.3%) than star-like (Table 3) with  $\pi = 1.4\%$ ,  $\kappa = 4.67$  and R = 0.515 (n = 33), indicating a relatively stable historical population size. This interpretation was backed by a small negative Fu's  $F_s$  value, ( $F_s = -5.33$ ; Table 3) and by the mismatch distribution analysis. The graph in Fig. 4c shows the same multimodal pattern as that of the mismatch analysis of clades A\* and B\* (Fig. 4a and b). In addition, estimates of pre- and post-expansion population sizes  $\theta_0$  and  $\theta_1$  for clade I were 1.585 and 13.229, respectively, reflecting only an 8.3-fold demographic

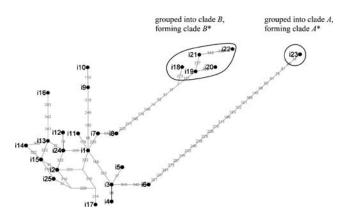


Fig. 2 Median-joining network of 25 HTs from Italy (n=46). Network was constructed using Network v.4.1.1.1. (Bandelt et al. 1999). Frequencies were set to f=1. Gray figures along the branches indicate mutational positions within the 335-bp fragment of the CR I. Long branches compressed for drawing purposes. The tree-like structure consists of three main units, separated from each other by ten or more mutations. These long branches among the three main groups of Italian L. europaeus HTs are indicative for the fact that each of these groups evolved independently. HTs i18–i23 were assigned to clades A and B as indicated (see also Fig. 3)

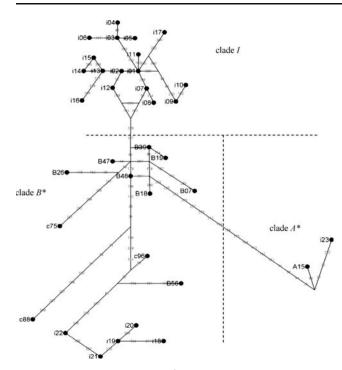
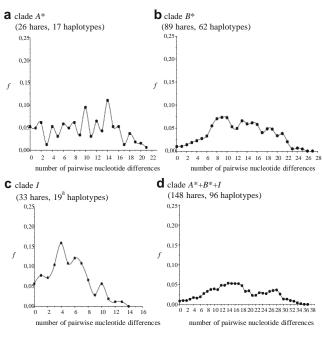


Fig. 3 Simplified MJ-network of Lepus europaeus HTs from Europe. Positions of Italian HTs (335 bp of d-loop) are depicted together with those from Central Europe (c75, c88, c96), assigned to the original clades A and B (Kasapidis et al. 2005). To improve legibility, we reduced the number of HTs depicted from clades A and B. Black dots represent HTs and the initial letters indicate clade assignment; A clade A, B clade B (see text and Kasapidis et al. 2005), i Italy clade, c Central Europe clade. Positions of nucleotide substitutions are marked in gray along branches which were partially compressed. Mutational positions in Figs. 2 and 3 differ due to different data set compositions. Dashed lines symbolize clade limits. Asterisks indicate the addition of sequences from the present work to clades A and B (see text)

increase; much lower than those estimated for clades A\* and B\* (175.2- and 16.4-fold, respectively). The lower demographic increase in clade I, as compared to those in clades A\* and B\*, reflects both the shorter time period available for its evolution, but also takes into account the smaller size of the Italian refugium. For clade I, post-expansion time was estimated to be 100–169 ky, which corresponded well with the one calculated using the mean within-clade genetic distance approach (Table 3).

# Central European haplotypes (clade CE)

The sequences of the 770 CE hares shared 93 HTs (R= 0.121) with a nucleotide diversity of  $\pi$ =0.89%. Both approaches to describe models of nucleotide substitution and substitution rate heterogeneity (hLRT, AIC) favored the TrN model (Tamura and Nei 1993) with a high proportion of invariable sites (59%) and a gamma distribution shape parameter of  $\alpha$ =0.9545. The MJ-Network (Fig. 5) and the



<sup>a</sup>: two haplotypes collapsed after reduction of fragment length to 335 bp

Fig. 4 Observed mismatch distribution in the HT-clades of L. europaeus. HTs were grouped according to the MJ-network. f frequency. X-axis scaling was adapted for drawing purposes. All graphs are multimodal. Superscripted a indicates two haplotypes collapsed after reduction of fragment length to 335 bp

Tcs cladogram both revealed a predominantly star-like phylogeny with several star-like sub-networks. This structure was confirmed by likelihood mapping, which estimated a star-like phylogeny proportion of 58%, a result that indicates a population that either underwent sudden population growth (Rogers and Harpending 1992; Rogers 1995) or experienced a large spatial expansion where neighboring populations exchange more than 50 migrants per generation (Ray et al. 2003). The former interpretation was supported by a highly negative  $F_s$  value ( $F_s$ =-25.55; Table 3) and by a unimodally shaped mismatch distribution graph, both indicative of rapid demographic expansion (Fig. 6). Estimates of pre- and post-expansion population sizes were  $\theta_0 = 0.742$  and  $\theta_1 = 25.454$ , respectively, showing a more than 30-fold total population size increase within an estimated time frame of 50–55 ky (Table 3).

## Relationships among clades

From clade-pairwise  $F_{ST}$  values (Table 1) it appears that differentiation between clades CE and I is greater than between I and B\*. However, when HT frequencies were set to 1 to correct for potential sampling errors, these relationships changed. As expected,  $F_{ST}$  values were reduced (but remained still highly significant) and differentiation between clades CE and I was now less than that between B\*

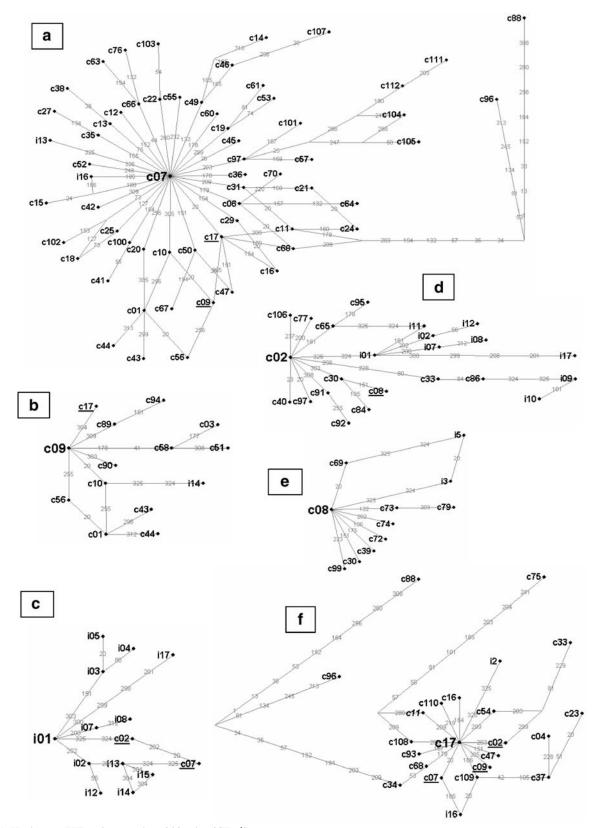


Fig. 5 Haplotype (HT) sub-networks within the  $\{CE+I\}$  group. a Haplotype network, which contains HT c07 estimated to be the ancestral HT of all central European HTs using Tcs (v.1.21, Clement et al. 2000). b–f Sub-networks of a, constructed and drawn separately for increased legibility. All sub-networks are connected, and the central

HT of each sub-network is highlighted. HTs are underlined if they themselves are central stars in their "own" sub-network. Gray figures along branches between HTs indicate substitution positions within the 335-bp d-loop fragment. Networks are partially compressed. HTs c75, c88 and c96 belong to clade B\* (see text)

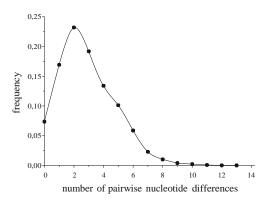


Fig. 6 Mismatch distribution among 93 Lepus europaeus haplotypes (CE; n=770)

and 1 ( $F_{ST}$ =0.293 and 0.291 versus 0.366 and 0.358, respectively). In both comparisons (observed and equal HT frequencies),  $F_{ST(torso)}$  values for 1 versus B\* were larger than they were between I and CE (Table 2). Table 1 also shows that the largest genetic distances were found between clade A\* and the other clades: the largest between clades A\* and CE, followed by I and B\*. We, therefore, conclude that cladogenesis in European brown hares began with the divergence of the ancestors of clades A\*, B\* and I in two major matrilineages leading to clade A\* and the ancestors of {B\*+1}.

This conclusion was supported by the estimation of clade coalescence times. We calculated the coalescence time for clade A\* and phylogroup {B\*+1} using the estimated gamma distribution of  $\alpha$ =0.0724. Mean genetic within-clade distances were then 6.56% for clade A\* (3.4% based on  $\alpha$ ~1 in Kasapidis et al. 2005) and 3.32% for {B\*+1}, respectively. Given a mutation rate of  $\mu$ =1.24×

Table 1 Pairwise  $\mathsf{F}_{ST}$  values among HT groups (all haplotypes c onsidered)

Clade/ group	CE	I	B*	CE+I+B*
I	0.388 [0.291]			
B*	0.553 [0.388]	0.377 [0.358]		
A*	0.861 [0.770]	0.756 [0.713]	0.603 [0.582]	0.795 [0.639]

Calculations are based on the TrN nucleotide substitution model (Tamura and Nei 1993) using  $\alpha$ =0.9719. Asterisks indicate that original clades A and B (Kasapidis et al. 2005) were extended by central European and Italian haplotypes. Brackets enclose F<sub>ST</sub> values at HT-frequency f=1. All F<sub>ST</sub> values were significantly different from zero.

CE Central Europe 93 haplotypes/770 hares, I clade Italy 17 HTs/33 hares, B\* clade 62 HTs/82 hares, A\* clade 17 HTs/26 hares, CE+ clades B\*+I 171 HTs/892 hares

Table 2 Pairwise  $F_{ST(torso)}$  values among clades/groups of haplotypes

Clade/group	CE <sub>torso</sub>	Itorso	
I <sub>torso</sub> B* <sub>torso</sub>	0.437 [0.295] 0.670 [0.504]	0.566 [0.522]	

Calculations are based on the TrN nucleotide substitution model (Tamura and Nei 1993) using  $\alpha$ =0.9719. Asterisks indicate that original clades A and B (Kasapidis et al. 2005) were extended by central European and Italian haplotypes. Brackets enclose F<sub>ST</sub> values at HT-frequency f=1. All F<sub>ST</sub> values were significantly different from zero. Clade A\*<sub>torso</sub> contained only 2 haplotypes, which were removed by the parsimonious cladogram construction and thus were not considered for the F<sub>ST(torso)</sub> calculation. MJ-network-torso haplotypes were (clade B\*) B1, B4, B5, B6, B7, B16, B18, B19, B22, B26, B29, B30, B39, B43, B47, B48; (clade l) i1, i2, i3, i5, i11, i13, i14, i15, i16; (group CE) c1, c2, c6, c7, c8, c9, c10, c11, c16, c17, c18, c19, c20, c21, c22, c24, c25, c29, c30, c31, c46, c47, c48, c50, c54, c56, c61, c65, c68, c69, c97, c104, c108, and c109.

CE Central Europe 93 haplotypes/770 hares, I clade Italy 17 HTs/33 hares, B\* clade 62 HTs/82 hares, A\* clade 17 HTs/26 hares, CE+ clades B\*+1 171 HTs/892 hares, Torso  $F_{ST}$  values calculated from the remaining haplotypes after application of the "torso"-function of the network construction algorithm (Bandelt et al. 1999) applied to reduce the bias caused by the numerous star phylogenies in the complete network, done by removal of terminal haplotypes, CE<sub>torso</sub> 34 haplotypes/558 individuals, I<sub>torso</sub> 10 HTs/24 individuals, B\*<sub>torso</sub> 17 HTs/23 individuals

 $10^{-7}$  mutations per nucleotide and generation (Pierpaoli et al. 1999), this corresponds to a mean HT coalescence time of approximately 529.3 ky for clade A\* and of 267.7 ky for {B\*+1}. Mean genetic distance between clades was 7.905%, with a TrN-corrected (Tamura and Nei 1993) confidence interval of CI<sub>95</sub> = 5.69%-10.983% (calculated correction factor = 1.14). Consideration of intra-clade evolution (=lineage sorting; Nei 1975, 1987; Wilson et al. 1985; Steel et al. 1996) led to a net between clade distance of 4.942%, corresponding to a divergence time between A\* and {B\*+1} of 239 ky (CI<sub>95</sub>=60.2-486.8 ky).

Mean genetic distances within clades B\* and I were 3.176% and 1.25%, respectively, corresponding to mean haplotypic coalescence times of 256.1 and 100.6 ky, respectively. The mean genetic distance between clades B\* and I was 3.824% (TrN-corrected  $CI_{95}=2.386-5.499\%$ ; calculated correction factor=1.067) and after consideration of lineage sorting it was 1.591% (TrN-corrected  $CI_{95}=$  0.153–3.267%). This corresponds to a mean divergence time between these clades of ~128.3 ky ( $CI_{95}=12.4-263.5$  ky).

#### Discussion

Late Pleistocene climate changes in Europe

During the Pleistocene (1.800–11 kya), central and northern Europe experienced four major cold periods (glacials): Chromerian complex (860–540 kya, marine isotopic stage

MIS22-MIS14), Elsterian 1-4 (480-400 kya, MIS12), Saalian (380-130 kya, MIS10-MIS6), and Weichselian (113-11 kya, MIS4-MIS2; Gibbard and van Kolfschoten 2004). The climatic changes in Europe during the Late Pleistocene (120-11 kya) consisted of two major cold and arid glacial periods, during which most of northern Europe (north of 52°N) as well as the Alps, Apennines, Pyrenees, Carpathians, Caucasus and other mountainous areas were covered by thick ice sheets (Hewitt 1996, 2004; Giraudi and Frezzotti 1997; Taberlet et al. 1998; Adams et al. 1999). The central continent was generally exposed to permafrost and tundra vegetation (Frenzel et al. 1992; Hewitt 2004 and references therein). The first major Late Pleistocene cold period lasted from 120 to 55 kya and reached its minimum temperatures (Lower Pleniglacial) ~70 kya. It was followed by a predominantly warm and moist period from ~55-30 kya. The second major cold phase, lasting from  $\sim 37$  to 10 kva, reached its most extreme stage (Upper Pleniglacial or Last Glacial Maximum) ~22–15 kya (Adams et al. 1999). Besides these major phases, many sudden and often short-lived warm events (interstadials) were interspersed between the generally colder conditions that prevailed between 130 and 10 kya (Adams et al. 1999). Since 18 kya, the general trend has been towards warming, although another colder period occurred between 11.6 and 10.3 kya ("Younger Dryas", Taylor et al. 1997).

# European refugia

During glacials, most temperate species, including the European brown hare, survived in refugia from which they colonized formerly uninhabitable areas during interglacials and interstadials. As pointed out earlier, three major refugial areas are recognized: the Iberian Peninsula, the Balkanic refugium (Albania, Greece, extending into Asia Minor) and the Italian Peninsula (cf. Bennett et al. 1991; Taberlet et al. 1998). Depending on the extend of the ice sheet and the length of glacial periods, minor or micro-refugia may have also existed across the continent.

During glacials, when the Adriatic sea level was 120–135 m lower than at present (Clark and Mix 2002), the Balkans and the Italian peninsula were connected across the northern, shallow part of the Adriatic Sea, potentially allowing dispersal and gene flow between populations from these two refugia. During Pleniglacials, a landbridge emerged between the Aegean and the Black Sea connecting Asia Minor and Europe (Aksu et al. 1999) rendering eastern refugia even more accessible for L. europaeus. However, several palynological and macrofossil plant studies as well as bone fossil analyses have also indicated the existence of buffered local microclimates north of the classical refugia in southern Europe within which temperate mammals could have survived during glacial periods (Bilton et al. 1998; Stewart and Lister 2001; Brunhoff et al. 2003; Lister 2004; Fig. 7). Although no common phylogeographic distribution patterns have been found across taxa, there seems to be a common feature in the current genetic structure of numerous species across the European continent; that is, geographic subdivision with parapatric genomes (spatial assortment) and suture-zones between them (Hewitt 1996, 2004; Taberlet et al. 1998; Gavrilets et al. 2000). The two most commonly recognized suture-zones are the Alps and the Pyrenees (Remington 1968; Taberlet et al. 1998). Other suture-zones lack apparent topographic barriers; e.g., one in Scandinavia (across central Sweden and Norway; Hewitt 2004), another one along the border between France and Germany (Taberlet et al. 1998), and a southern suture-zone connecting the Pleniglacial refugia of the Balkans and Asia Minor (Hewitt 1999).

In general, the current geographical distribution of the three ancient European brown hare clades (A\*, B\*, and I) confirms the existence of the main southern European refugial areas. In addition, the overlapping distribution of L. europaeus clades A and B in northern Greece (Thrace) and Bulgaria (Kasapidis et al. 2005) corresponds to the Balkans-Asia Minor suture-zone. As for the evolution of CE, the question emerges which one of the refugia (Italian, Balkanic) harbored its source population? Following the  $F_{ST}$  statistics (Tables 1 and 2), we consider the Italian clade as the most likely source population for the colonization of central European territories (see also below). Two scenarios are conceivable for this process: CE hares either colonized central Europe from I during early intervals of climatic improvement and survived at their northern new locations during subsequent climatic oscillations or their pattern of genetic variability is due to expansion since the last glacial stage. The genealogical structure of CE (Fig. 5) is predominantly star-like, emphasizing the relative short coalescence time of this clade and rendering the existence of cryptic micro-refugia unlikely. This is supported by the fact that existing CE hare populations do not display a substantial accumulation of mutations, which would distinguish them from their possible source populations as would be expected to occur by genetic drift and/or selection if they had survived in northern refugia for several interglacials. Instead, on the basis of derived haplotypes, this clade very likely evolved directly and fairly recently from I (see below).

Although it is unlikely that northern remnants of an earlier advance from the Italian (and/or Balkanic) refugium formed "secondary" refugia (sensu Brewer et al. 2002) north of the Alps for subsequent expansion waves, the worsening climatic conditions during glacials would have caused either the extinction of these pioneer populations north of the Alps or their retreat to their southern source refugia. In the latter case, repeated genetic admixture of neo-CE, I and/or B may have hindered lineage divergence among these three clades. At this time, we are unable to distinguish between these two scenarios.

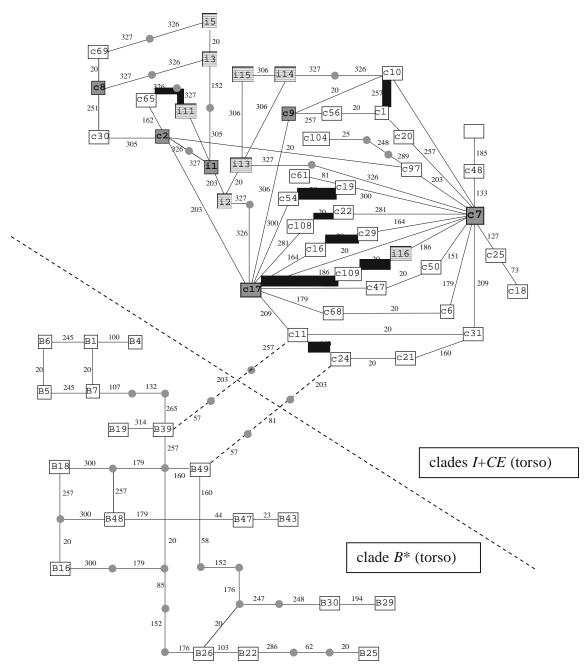


Fig. 7 Parsimonious cladogram of HTs from the MJ-Network-torso of L. europaeus from Europe. Connecting lines represent substitutions; figures beside lines indicate their position within the 335-bp d-loop fragment. Gray dots indicate mutations introduced by the algorithm (Tcs; v.1.21, Clement et al. 2000). HT labels: B clade B, c Central Europe, i Italy. Upper network: dark boxes mark star-HTs from sub-

A clear tree-like pattern can be seen in the evolution of the three ancient major clades (Figs. 2 and 3). We hypothesize that phylogroup  $\{B^*+I\}$  evolved after an advance to the Northwest followed by allopatric divergence from the ancestors of clade A\*. Because the territorial overlap between the distributions of clades A\* and B\* is small (Kasapidis et al. 2005), we can also rule out that

networks (Fig. 5). Gray stripes underlay HTs from Italy. Lower network: asterisk indicates extension of original clade B (Kasapidis et al. 2005). Both clade networks are separated by three mutations. Due to its excessively large genetic distance (step limit >7) clade A\* was excluded from the cladogram

northbound advancing "colonists" from clade A\* returned to their southern source populations when climatic conditions deteriorated; instead, they either went extinct or survived farther north or west, respectively.

Identification of a suture-zone between clades B\* and I will only be possible after more thorough sampling in Italy and the Balkans.

Table 3 Analyses performed on the European brown hare (Lepus europaeus) data sets

Analysis	Clade A*	Clade B*	Clade I	Clade $\{B^*+I\}$	CE	All
Number of animals n	26	89	33	122	770	918
Number of usable sites [bp] <sup>a</sup> Diversity indices	335	335	334	336	335	335
Haplotypes identified <sup>b</sup>	$17 (16+1^{\circ})$	$62 (54+5^{c}+3^{d})$	17	79	93 $(92+1^{b})$	188 <sup>b</sup>
Transitions/transversions (Ti/Tv ratio)	38/3 (12.67)	58/19 (3.05)	14/3 (4.67)	66/21 (3.14)	64/3 (21.3)	101/25 (4.04)
Nucleotide diversity :	0.0284	0.0358	0.014	0.0317	0.0089	0.0184
Haplotype diversity h	0.948	0.99	0.932	0.9898	0.926	0.948
R (number of haplotypes/n)	0.654	0.697	0.515	0.647	0.121	0.205
Likelihood mapping						
Proportion of tree-like phylogeny [%]	56.5	64.5	56.3	63.3	42.0	57.7
Akaike information criterion (A	AIC)					
Model of nucleotide substitution	K81uf+G	TrN+I+G	TrN+I	TrN+I+G	TrN+I+G	TrN+I+G
Number of substitution	6	6	6	6	6	6
types −lnL	734.0917	1 210 0675	591 0296	1 402 0229	1 107 6510	2 452 2280
AIC	1,480.1835	1,319.9675	581.9386	1,493.0228	1,107.6519	2,453.3389
	,	2,653.9351 0.649	1,173.8772 0.8446	3,000.0457	2,229.3037	4,920.6777
Proportion of invariable sites	0			0.6053	0.5917	0.5431
Gamma distribution shape parameter ! <sup>e</sup>	0.0724	1.0267	Equal site rates	0.9306	0.9545	0.9719
Within-clade mean genetic distance <sup>f</sup>	0.06564	0.03176	0.01248	0.0332	0.04702	0.03596
Mean haplotype coalescence time [ky]	529.3	256.1	100.6	267.7	_	
Sudden population expansion						
Fu's F <sub>s</sub>	-2.71 <sup>g</sup>	-24.32	-5.33	-24.33	-25.55	
$p(F_{Ssim} \leq F_{Sobs}), 1,000$	0.024	$< 10^{-5}$	0.001	$< 10^{-5}$	$< 10^{-5}$	
permutations						
Mismatch distribution						
- (CI <sub>95</sub> )	9.53 (5.33/13.55)	12 (8.86/19.9)	4.67 (2.5/7.56)	10.67 (8.46/14)	2.91 (1.71/4.46)	
Variance (V) of .	30.32	26.19	9.119	19.15	3.557	
C (CI <sub>95</sub> )	13.67 (6.26/20.7)	8.94 (4.64/22)	4.15 (1.62/11)	10.1 (8/15.9)	2.3 (1.27/5.1)	
$\theta_0$ (CI <sub>95</sub> )	0.106 (0.00/5.26)	5.06 (0.00/18.3)	1.58 (0.00/3.65)	1.93 (0.00/4.42)	0.742	
$\theta_1$ (CI <sub>95</sub> )	18.57 (8.18/108.8)	82.81 (36/3,777)	13.23 (6.79/4,628)	54.87 (39.85/439)	25.45 (7.38/6,735)	
$\theta_1/\theta_0$ ratio	175.2	16.4	8.3	28.4	34.3	
Goodness-of-fit tests						
Sum of squared deviation (SSD)	0.0084	0.00244	0.01168	0.00096	0.00088	
$p(SSD_{sim} \ge SSD_{obs})$	0.818	0.471	0.0396	0.770	0.616	
Post-expansion time [kya]						
to $\mathfrak{b}; \mathfrak{h}_0$ $\mathfrak{h}_{\mathbf{p}}$	119.9 (109-131)	198.1 (189-229)	61.6 (n.c152)	186.1 (124-281)	50.7 (37.3-n.c.)	
10% variation of µ [kya]	99.1–145.2	171.4–254	n.c169.3	113.2-311.9	33.9–n.c.	
t(C) from mismatch distribution <sup>i</sup> (CI <sub>95</sub> )	329 (151-499)	215.2 (112-529)	100.2 (39–266)	242.4 (192–382)	55.3 (31-122)	
10% variation of µ [kya]	137.1–554.2	101.6-588.2	35.6-295.3	174.5-424.3	27.8-136.1	

Asterisks indicate after addition of haplotypes from the Italian and central European sample.

Models of nucleotide substitution: TrN Tamura and Nei (1993), K81uf Kimura-3-parameter (Kimura 1981) with unequal base frequencies, +I with invariable sites, +G among-site rate variation with gamma distribution shape parameter  $\alpha$ , n.c. not calculable because of (v<.) or (.<\_0), ky thousand years, kya thousand years ago,  $\mu$  mutation rate, t time in number of generations,  $_0$  pre-expansion population size,  $_1$  post-expansion population size,  $_1$  mean number of pairwise differences, Cl<sub>95</sub> 95% confidence interval, gray column "all" estimates are demographically uninformative

<sup>a</sup> Length differences due to indel

<sup>b</sup> Haplotype c8 and i6 are identical in the sequence investigated.

<sup>c</sup> From Italy, assigned to clade by MJ-network

<sup>d</sup> From central Europe, assigned to clade by MJ-network

<sup>e</sup> The smaller the  $\alpha$ , the more uneven the mutation rate (uniform mutation rates correspond to  $\alpha = \infty$ ).

<sup>f</sup> Calculated using the  $\alpha$  estimation of AIC

<sup>g</sup> Value not significantly different from zero, because p>0.2 (Fu 1997)

<sup>h</sup> Post-expansion population size set to ∞

 $^{\circ}$ C=2µt (Rogers 1995), t=2 years generation time (Marboutin and Peroux 1995), µ=1.24×10<sup>-7</sup> substitutions/site×generation [kya] (Pierpaoli et al. 1999).

Cladogenesis in L. europaeus

Ts/Tv ratios varied considerably among hare lineages (Table 3). Such variation is indicative of significant substitution rate heterogeneity, as confirmed by models of observed nucleotide substitution (Table 3). However, nucleotide diversity ( $\pi$ ) also varied considerably among lineages, suggesting either substitution saturation (when sequence divergence is high, e.g., in clade A\*) or overestimation of the Ts/Tv ratio when sequence divergence is low (Yang 1994; Yang and Yoder 1999). The latter seems to apply to CE (Table 3); however, its low R value (R=0.121) contradicts such a scenario (Yang and Yoder 1999), leading us to conclude that substitution saturation is more likely.

Net divergence estimates between clades (3.8-4.9%) as well as relatively high nucleotide diversity estimates within clades (Table 3) indicate that the distinct phylogeographic structure of this species began to evolve prior to the latest glaciation. From its current matrilinear distribution pattern it seems plausible that L. europaeus expanded into Europe from Asia Minor. This expansion probably occurred in at least four phases (Fig. 8). Beginning as a panmictic population (phase I), brown hares began to diverge (phase II) into two major matrilines between 60 and 487 kya (mean: 239 kya; see also Kasapidis et al. 2005: 275 kya). One of these lineages evolved to today's clade A\*, whereas the other comprised the ancestors of  $\{B^*+I\}$ . The evolution of these lineages may have occurred in the Mindel-Riss interglacial (isotopic stage MIS7), a period estimated to have occurred ~410-380 kya or during the Riss-Wuerm interglacial (MIS5e) at ~130-110 kya (Kukla 2000; Gibbard and van Kolfschoten 2004). Our coalescence time estimate for the oldest brown hare clade (clade A\*: 529 ky) is relatively short compared to the estimated age of the species (~3 million years; Pierpaoli et al. 1999). The reason for this discrepancy may lie in the fact that either

undiscovered lineages older than clade A\* still exist on Asia Minor and/or further east, or that intermediate A\*-HTs have not yet been sampled.

At 12–263 kya (mean: 128 kya), the ancestors of phylogroup  $\{B^*+I\}$  began to diverge and to evolve into today's clades B\* and I (phase III). This phase most likely took place in the Riss-Wuerm interglacial (MIS5e, ~130-110 kya) or in the beginning of the current interstadial. In light of the large confidence limits around the time estimates, the timescale of these first three phases has to be viewed cautiously. The timescale of the final phase (phase IV) concerns the differentiation of clade CE. Firstly, the star-like phylogenetic structure of CE supports population expansion without subsequent admixture with other clades, but does not support survival and differentiation during interglacials in cryptic refugia (Forster 2004). Therefore, we conclude that expansion and colonization has probably occurred since the last ice age, as indicated by divergence times (Table 3). In addition, paleontological records suggest that Central Europe was not colonized by L. europaeus before 10–8 kya, when a rise in temperature, receding glaciers, and expanding vegetation generated favorable living conditions. The first safely classifiable post-Pleistocene L. europaeus fossils north of the Alps appear ~8 kya (Musil 1985; Döhle 1999), with successive northward distribution expansion by brown hares only occurring in the wake of native forest clearing and agricultural landscape changes (Fischer and Heinrich 1983; Benecke 1994; Döhle 1999), which provided open habitats necessary for L. europaeus to successfully compete with Lepus timidus, leading to the ecological replacement of the latter species (Gaiduk 1982; Benecke 1994; Döhle 1999).

Our results (Figs. 5 and 7, Tables 1 and 2) also support the hypothesis that clade I was the source clade for the northward expansion leading to the establishment of CE by a strong sudden population expansion. As seen in clades A\*,

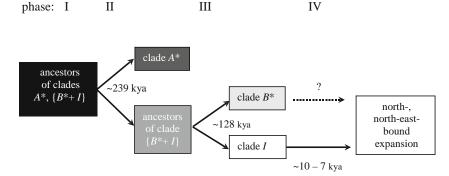


Fig. 8 Proposed chronology of Lepus europaeus cladogenesis in Europe. Phase I Most likely panmictic ancestral population of Lepus europaeus. Phase II ~239 kya (Hoxnian interglacial, MIS11, Gibbard and van Kolfschoten 2004): Ancestors of clades A\*, B\*, and I diverged

in two main matrilines, later forming clade A\* and the ancestors of clades {B\*+1}. Phase III ~128 kya (Eemian interglacial, MIS5e, Kukla 2000): ancestors of clade {B\*+1} diverged in clades B\* and I. Phase IV Re-colonization of Central Europe about 10–7 kya

B\*, and I, refugial populations are expected to display a high intraspecific diversity (much higher than in recently colonized northern areas) and to contain ancestral HTs (Cooper et al. 1995; Hewitt 1996, 2004). Although the direct comparison between CE (n=770) and I (n=33) is biased due to very different sample sizes, we regard the high number of different HTs per sampled individuals in I (R=0.515 and  $\pi$ =0.014, as compared to R=0.121 and  $\pi$ =0.0089 in CE) as an indirect indicator of higher intraspecific diversity in present or past Italian hare populations. Conversely, given the estimated mutation rate in brown hares of 12.4% per million years (Pierpaoli et al. 1999), the high intraspecific diversity in CE (h=0.926, Table 3) could not have accumulated within the short post-Pleistocene period of ~9 ky, indicating that some HTs currently sampled in CE derived from ancient HTs in I (previously existing or unsampled), which either became extinct in the course of the population expansions/contractions or did not contribute equally to the expanding population. Support for the assumption that CE evolved from I comes from the combined analysis of haplotypes from CE+l versus haplotypes from clade B\*, in which the former are shown to be well-separated from all B\*-HTs (upper part of Fig. 7). However, the mean CE intra-clade genetic distance of 3.597% leads to a HT coalescence time estimation of ~290 ky, which contradicts the assumption of a post-Pleistocene colonization of central Europe. To resolve this inconsistency, two explanations seem possible. Firstly, higher lineage-specific mutation rates (Meyer et al. 1999; Gissi et al. 2000) may have accompanied genetic divergence. However, an unreasonably high mutation rate would be necessary for CE to generate the observed HT diversity within only  $\sim$ 9 ky. Even under the likely assumption that the 109 HTs of all CE sub-networks did not evolve sequentially but rather simultaneously, the resulting coalescent time of  $\sim$ 99 ky as estimated from the network torsi (Fig. 7) would still be too high. Thus, a second explanation appears much more plausible, whereupon many of the HTs detected in CE did not evolve within this clade but rather constitute ancestral HTs (ancestral polymorphisms that evolved in the Italian refugium) but have either gone extinct or have not yet been detected in the Italian endogenous source population.

Interestingly, the three HTs that were assigned to clade  $B^*$  (c75, c88, c96) came from animals sampled in Serbia and southern Slovakia, respectively, potentially indicating a most northern clade  $B^*$  distribution rim along Slovenia–Austria–Slovakia.

# Conclusions

The current clade distribution of L. europaeus indicates that this species survived the Pleistocene glaciation periods in at

least three southern European refugia: the Italian peninsula, the Balkans, and Asia Minor. The genealogy of central European HTs does not support the long-term survival of brown hares in northern micro-refugia during glacial periods. Cladogenesis of the European brown hare predates the Last Glacial Maximum and was probably initiated by one or more northward expansion waves from Asia Minor or further east, when the first genetic lineage split occurred between ancestors of today's clade A\* and the ancestors of phylogroup {B\*+1}, followed by the divergence of B\* and I. Genetic and paleontological evidence suggests that colonization of Central Europe by L. europaeus took place in the post-Pleistocene in the wake of anthropogenic landscape changes.

Whether this species also had an Iberian refugium or not remains unanswered. A high percentage of the Spanish hare samples originally available for this study (n=32) were unsuitable due to the presence of introgressed L. timidus mtDNA (Thulin et al. 1997; Alves et al. 2003). The HTs of the remaining specimens (n=16) fitted very well into the starlike phylogeny of CE and thus did not match the criteria for refugial HTs. Therefore, we are at present unable to distinguish whether these specimens are offspring of allochthonous hares or represent the edge of the central European haplogroup. Whether the analysis of Y-chromosomal haplotypes (Wells et al. 2001; Putze et al. 2007) will contribute to solve this question remains to be seen in the future.

Acknowledgments We are very grateful to all people (especially hunters and forest wardens) who contributed to this study. T. Noventa and A. Schmidt offered expert technical assistance. U. Peschel, S. Blottner, M. Faßbender (all IZW), and M. Putze (Friedrich-Schiller University Jena) provided numerous hare samples, as did D. Faber (University Gießen, Germany), Z. Pielowski (Research Institute of the Polish Hunter's Association Czempiń, Poland), J. Slamečka (Czech Academy of Landscape Ecology, Brno), and K. Schmidt (Mammal Research Institute of the Polish Academy of Sciences, Białowieza). Spanish hares were kindly provided by the Scientific Collection of the Doñana Biological Station (CSIC) and the CSIC-Navarra Foral Goverment project, Spain. Special thanks goes to Dr. N. Benecke (German Archeological Institute) for his help regarding the paleontology of L. europaeus. HCH and EP were funded by the Centro di Ecologia Alpina and the Research Fund of the Autonomous Province of Trento (Project Faunagen). We also thank the two anonymous reviewers for their suggestions and comments on a previous version of the manuscript. All experiments carried out in the course of this study complied with the existing law in Germany.

### References

- Adams J, Maslin M, Thomas E (1999) Sudden climate transitions during the Quarternary. Prog Phys Geogr 23:1–36
- Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Contr 19:716–723
- Alves PC, Ferrand N, Suchentrunk F, Harris DJ (2003) Ancient introgression of Lepus timidus mtDNA into L. granatensis and L. europaeus in the Iberian Peninsula. Mol Phyl Evol 27:70–80
- Aksu AE, Hiscott RN, Yaşar D (1999) Oscillating Quarternary water levels of the Marmara Sea and vigorous outflow into the Aegean

Sea from the Marmara Sea–Black Sea drainage corridor. Marine Geology 153:275–302

- Arnason U, Adegoke JA, Bodin K, Born EW, Esa YB, Gullberg A, Nilsson M, Short RV, Xu X, Janke A (2002) Mammalian mitogenomic relationships and the root of the eutherian tree. Proc Natl Acad Sci USA 99:8151–8156
- Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48
- Benecke N (1994) Archäozoologische Studien zur Entwicklung der Haustierhaltung in Mitteleuropa und Südskandinavien von den Anfängen bis zum ausgehenden Mittelalter. Schr Ur-FruÉhgesch 46:1–451
- Bennett KD, Tzedakis PC, Willis KJ (1991) Quarternary refugia of north European trees. J Biogeography 18:103–115
- Bilton DT, Mirol PM, Mascheretti S, Fredga K, Zima J, Searle JB (1998) Mediterranean Europe as an area of endemism for small mammals rather than a source for northwards postglacial colonization. Proc Royal Soc B 265:1219–1226
- Branco M, Monnerot M, Ferrand N, Templeton AR (2002) Postglacial dispersal of the European rabbit (Oryctolagus cuniculus) on the Iberian peninsula reconstructed from nested clade and mismatch analyses of mitochondrial DNA genetic variation. Evolution 56:792–803
- Brewer S, Cheddadi R, de Beaulieu J-L, Reille M (2002) The spread of deciduous Quercus throughout Europe since the last glacial period. Forest Ecol Manage 156:27–48
- Brito PH (2005) The influence of Pleistocene glacial refugia on tawny owl genetic diversity and phylogeography in western Europe. Mol Ecol 14:3077–3094
- Brunhoff C, Galbreath KE, Fedorov VB, Cook JA, Jaarola M (2003) Holarctic phylogeography of the root vole (Microtus oeconomus): implications for the late Quarternary biogeography of high latitudes. Mol Ecol 12:957–968
- Clark PU, Mix AC (2002) Ice sheets and sea level of the Last Glacial Maximum. Quarternary Science Rev 21:1–7
- Clement MD, Posada MD, Crandall KA (2000) Tcs: a computer program to estimate gene genealogies. Mol Ecol 9:1657–1660
- Cook JA, Bidlack AL, Conroy CJ, Demboski JR, Fleming MA, Runck AM, Stone KD, MacDonald SO (2001) A phylogeographic perspective on endemism in the Alexander Archipelago of southeast Alaska. Biol Conserv 97:215–227
- Cooper SJ, Ibrahim KM, Hewitt GM (1995) Postglacial expansion and genome subdivision in the European grasshopper Chorthippus parallelus. Mol Ecol 4:49–60
- Dietrich U (1984) Beitrag zum Status des europäischen Feldhasen (Lepus europaeus Pallas 1778) im südlichen Chile. Z Jagdwiss 30:256–259 (in German)
- Dietrich U (1985) Populationsökologie des in Argentinien eingebürgerten europäischen Feldhasen (Lepus europaeus). Z Jagdwiss 31:92–102 (in German)
- Djan M, Obreht D, Vapa L (2006) Polymorphism of mtDNA regions in brown hare (Lepus europaeus) populations from Vojvodina (Serbia and Montenegro). Eur J Wildl Res 52:288–291
- Döhle H-J (1999) Die nacheiszeitliche Verbreitung einiger Wildsäugetiere (Carnivora, Leporidae) in Mittel-und Norddeutschland. Säugetierkundl Mitt 44:154–161 (in German)
- Dynesius M, Jansson R (2000) Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. Proc Natl Acad Sci USA 97:9115–9120
- Excoffier L, Schneider S (1999) Why hunter-gatherer populations do not show signs of Pleistocene demographic expansions. Proc Natl Acad Sci USA 96:10597–10602
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinformatics Online 1:47–50

- Fedorov VB, Stenseth NC (2001) Glacial survival of the Norwegian lemming (Lemmus lemmus) in Scandinavia: inference from mitochondrial DNA variation. Proc Royal Soc B 268:809–814
- Fickel J, Lieckfeldt D, Pitra C (1999) Analyse der genetischen Diversität und Struktur in benachbarten Populationen des Feldhasen (Lepus europaeus Pallas, 1778). Z Jagdwiss 45:230– 237 (in German)
- Fickel J, Schmidt A, Putze M, Spittler H, Ludwig A, Streich WJ, Pitra C (2005) Genetic structure of populations of European brown hare: implications for management. J Wildlife Manage 69:760–771
- Fickel J, Lieckfeldt D, Ratanakorn P, Pitra C (2007a) Distribution of haplotypes and microsatellite alleles among Asian elephants (Elephas maximus) in Thailand. Eur J Wildl Res 53:298–303
- Fickel J, Wagener A, Ludwig A (2007b) Semen cryopreservation and the conservation of endangered species. Eur J Wildl Res 53:81–89
- Fischer K, Heinrich WD (1983) Skelettreste von Lepus europaeus PALLAS, 1778 (Lagomorpha, Mammalia) aus dem fossilen Tierbautensystem von Pisede bei Malchin. Wiss Z Humboldt-Univ Berl, Math-natwiss Reihe 32:713–717 (in German)
- Forster P (2004) Ice Ages and the mitochondrial DNA chronology of human dispersals: a review. Phil Trans R Soc Lond B 359:255– 264
- Fraguglione D (1963) Introduction artificielle et tentatives d'acclimatisation du lièvre commun (Lepus europaeus Pallas 1778) (de par le monde). Royal Saint Hubert Club belgique, Belgium, pp 29– 30 42–46 52–53 (in French)
- Fraguglione D (1971) Les problèmes posés par les entreprises de repeuplement avec le lièvre commun (Lepus europaeus Pallas 1778). Union Internationale des Biologistes du Gibier, Actes du X<sup>e</sup> Congrès, France, pp 491–505 (in French)
- Frenzel B, Pécsi M, Velichko AA (1992) Atlas of Paleoclimates and Paleoenvironments of the Northern Hemisphere. Late Pleistocene– Holocene. Geographical Research Institute, Hungarian Academy of Sciences. Gustav-Fischer Verlag, Budapest, Stuttgart, p 153
- Frölich K, Fickel J, Ludwig A, Lieckfeldt D, Streich WJ, Jurčík R, Slamecka J, Wibbelt G (2007) New variants of European brown hare syndrome virus (EBHSV) strains in free-ranging European brown hares (Lepus europaeus) from Slovakia. J Wildl Dis 43:89–96
- Ford MJ (2002) Applications of selective neutrality tests to molecular ecology. Mol Ecol 11:1245–1262
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915–925
- Gaiduk VE (1982) Some aspects of the ecology of closely related sympatric lagomorph species (Lagomorpha). Plenum Publishing Corporation (1983): pp. 401–405. Translated from Ékologiya 6:55–60 (1982)
- Gavrilets S, Li H, Vose MD (2000) Patterns of sparapatric speciation. Evolution 54:1126–1134
- Gibbard P, van Kolfschoten T (2004) The Pleistocene and Holocene Epochs. In: Gradstein FM, Ogg JG, Smith AG (eds) A geologic time scale, Chapter 22. Cambridge University Press, Cambridge, UK, pp 441–452
- Giraudi C, Frezzotti M (1997) Late Pleistocene glacial events in the Central Apennines, Italy. Quarternary Research 48:280–290
- Gissi C, Reyes A, Pesole G, Saccone C (2000) Lineage-specific evolutionary rate in mammalian mtDNA. Mol Biol Evol 17: 1022–1031
- Gortázar C, Ferroglio E, Höfle U, Frölich K, Vicente J (2007) Diseases shared between wildlife and livestock: a European perspective. Eur J Wildl Res 53:241–256
- Haffer J (1969) Speciation in Amazonian forest birds. Science 165:131-137
- Hasegawa M, Kishino K, Yano T (1985) Dating the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174

- Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. Biol J Linn Soc 58:247–276
- Hewitt GM (1999) Postglacial re-colonization of European biota. Biol J Linn Soc 68:87–112
- Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quarternary. Phil Trans R Soc Lond B 359:183–195
- Kasapidis P, Suchentrunk F, Magoulas A, Kotoulas G (2005) The shaping of mitochondrial DNA phylogeographic patterns of the brown hare (Lepus europaeus) under the combined influence of late Pleistocene climatic fluctuations and anthropogenic translocations. Mol Phyl Evol 34:55–66
- Kimura M (1981) Estimation of the evolutionary distances between homologous nucleotide sequences. Proc Natl Acad Sci USA 78:454–458
- Kukla G (2000) The last interglacial. Science 287:987–989
- Kumar S, Tamura K, Nei M (2004) Mega3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163
- Li WH (1977) Distribution of nucleotide differences between two randomly chosen cistrons in a finite population. Genetics 85:331-337
- Lincoln GA (1974) Reproduction and March madness in the Brown hare, Lepus europaeus. J Zool Lond 174:1–14
- Lister AM (1984) Evolutionary and ecological origins of British deer. Proc Royal Soc Edinburgh 82:205–229
- Lister AM (2004) The impact of Quarternary Ice Ages on mammalian evolution. Phil Trans R Soc Lond B 359:221–241
- Lönnberg E (1905) On hybrids between Lepus timidus L. and Lepus europaeus Pallas from southern Sweden. Proc Zool Soc London 1:278–287
- Marboutin E, Peroux R (1995) Survival pattern of European hare in a decreasing population. J Appl Ecol 32:809–816
- Meyer S, Weiss G, von Haeseler A (1999) Pattern of nucleotide and rate heterogeneity in the hypervariable regions I and II of human mtDNA. Genetics 152:1103–1110
- Musil R (1985) Die Fauna der Magdalénien-Siedlung Oelknitz. Weimarer Monographien zur Ur-und Frühgeschichte 17:1–43
- Nei M (1975) Molecular population genetics and evolution. North-Holland and American Elsevier, Amsterdam 288
- Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York 512
- Pierpaoli M, Riga F, Trocchi V, Randi E (1999) Species distinction and evolutionary relationships of the Italian hare (Lepus corsicanus) as described by mitochondrial DNA sequencing. Mol Ecol 8:1805–1817
- Posada D (2003) Using Modeltest and Paup\* to select a model of nucleotide substitution. In: Baxevanis AD, Davison DB, Page RDM, Petsko GA, Stein LD, Stormo G (eds) Current protocols in bioinformatics. Wiley, New York, pp 6.5.1–6.5.14
- Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of the AIC and Bayesian approaches over likelihood ratio tests. Syst Biol 53:793–808
- Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics Appl Note 14:817–818
- Putze M, Nürnberg S, Fickel J (2007) Y-chromosomal markers for the European brown hare (Lepus europaeus, Pallas 1778). Eur J Wildl Res, 53:257–264
- Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. Mol Biol Evol 20:76–86
- Remington CL (1968) Suture-zones of hybrid interaction between recently joined biotas. Evol Biol 2:321–428
- Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. Evolution 49:608–615
- Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9:552–569

- Roy K, Valentine JW, Jablonski D, Kidwell SM (1996) Scales of climatic variability and time averaging in Pleistocene biotas: implications for ecology and evolution. TREE 11:458–453
- Schmidt HA, Strimmer K, Vingron M, von Haeseler A (2002) Tree-Puzzle—maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics Appl Note 18:502–504
- Schmidt HA, Strimmer K, von Haeseler A (2004) Tree-Puzzlemaximum likelihood analysis for nucleotide, amino acid and two state data. Available online http://www.tree-puzzle.de
- Schneider E (1978) Der Feldhase. BLV Verlagsgesellschaft München, 198 pp
- Schneider S, Excoffier L (1999) Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. Genetics 152:1079–1089
- Spittler H (1996) Der Hase in der Vergangenheit, Gegenwart und Zukunft. Diana-Verlag, Vettelschoß, Germany 265
- Steel M, Cooper A, Penny D (1996) Confidence intervals for the divergence time of two clades. Syst Biol 45:127–134
- Stewart JR (2003) Comment on "Buffered tree population changes in a Quaternary refugium: evolutionary implications". Science 299:825a
- Stewart JR, Lister AM (2001) Cryptic northern refugia and the origins of the modern biota. TREE 11:608–613
- Strimmer K, von Haeseler A (1997) Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proc Natl Acad Sci USA 94:6815–6819
- Suchentrunk F, Michailov C, Markov G, Haiden A (2000) Population genetics of Bulgarian brown hares Lepus europaeus: allozymic diversity at zoogeographical crossroads. Acta Theriologica (Warszawa) 45:1–12
- Swofford DL (2002) Paup\*. Phylogenetic Analysis Using Parsimony (\*and other Methods). Version 4.0 beta version. Sinauer Associates, Sunderland, Massachusetts, 144 pp (updates available at http://paup.csit.fsu.edu)
- Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. Mol Ecol 7:453–464
- Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526
- Taylor KC, Mayewski PA, Alley RB, Brook EJ, Gow AJ, Grootes PM, Meese DA, Saltzman ES, Severinghaus JP, Twickler MS, White JWC, Whithlow S, Zielinski GA (1997) The Holocene– Younger Dryas transition recorded at Summit, Greenland. Science 278:825–827
- Templeton AR (1998) Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. Mol Ecol 7:381–397
- Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132:619–633
- Thulin GC, Jaarola M, Tegelström H (1997) The occurrence of mountain hare mitochondrial DNA in wild brown hares. Mol Ecol 6:463–467
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl Acids Res 24:4876–4882
- Tzedakis PC, Lawson IT, Frogley MR, Hewitt GM, Preece RC (2002) Buffered tree population changes in a Quarternary refugium: evolutionary implications. Science 297:2044–2047 (with technical comments in Vol. 299:825a–825b, 2003)
- Tzedakis PC, Lawson IT, Frogley MR, Hewitt GM, Preece RC (2003) Reply to comment on "Buffered vegetation changes in a Quaternary refugium: evolutionary implications". Science 299:825b

Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva I, Blue-Smith J, Jin L, Su B, Pitchappan R, Shanmugalakshmi S, Balakrishnan K, Read M, Pearson NM, Zerjal T, Webster MT, Zholoshvili I, Jamarjashvili E, Gambarov S, Nikbin B, Dostiev A, Aknazarov O, Zalloua P, Tsoy I, Kitaev M, Mirrakhimov M, Chariev A, Bodmer WF (2001) The Eurasian Heartland: a continental perspective on Y-chromosome diversity. Proc Natl Acad Sci USA 98:10244–10249

Willis KJ, Whittaker RJ (2000) The refugial debate. Science 287:1406-1407

- Wilson AC, Cann RL, Carr SM, George M Jr, Gyllensten UB, Helm-Bychowski K, Higuchi RC, Palumbi SR, Prager EM, Sage RD, Stoneking M (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. Biol J Linn Soc 26:375–400
- Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol 39:306–314
- Yang Z, Yoder AD (1999) Estimation of the transition/transversion rate bias and species sampling. J Mol Evol 48:274–283