Shared genetic diversity across the global invasive range of the Monk parakeet suggests a common restricted geographic origin and the possibility of convergent selection

Pim Edelaar[1,3], Severine Roques[3], Elizabeth A. Hobson[2,4], Anders Gonçalves da Silva[8], Michael L. Avery[5], Michael A. Russello[6], Juan C. Senar[7], Timothy F. Wright[2], Martina Carrete[1,3] & José L. Tella[3]

[1] Department of Molecular Biology and Biochemical Engineering, University Pablo de Olavide, ES-41013 Seville, Spain

[2] Department of Biology, MSC 3AF, New Mexico State University, Las Cruces NM, 88003, USA

[3] Department of Conservation Biology, Estación Biológica de Doñana (CSIC), Avda. Américo Vespucio s/n, 41092 Sevilla, Spain

[4] National Institute for Biological and Mathematical Synthesis, University of Tennessee, Knoxville, TN 37996

[5] U.S. Department of Agriculture, Wildlife Services, National Wildlife Research Center, Gainesville, FL, 32641, USA

[6] Department of Biology, University of British Columbia, Okanagan Campus, Kelowna, British Columbia, V1V 1V7, Canada

[7] Natural History Museum of Barcelona, Passeig Picasso s/n, 08003 Barcelona, Spain

[8] School of Biological Sciences, Monash University, Clayton, Victoria 3800, Australia

Key words: invasion genetics, native origin, population genetic structure, founder effect, bottleneck, selection.

Corresponding author: Pim Edelaar, Department of Molecular Biology and Biochemical Engineering, University Pablo de Olavide, ES-41013 Seville, Spain. Email: edelaar@upo.es

Running title: invasion genetics of Monk parakeets

ABSTRACT

While genetic diversity is hypothesized to be an important factor explaining invasion success, there is no consensus yet on how variation in source populations or demographic processes affects invasiveness. We used mitochondrial DNA haplotypic and microsatellite genotypic data to investigate levels of genetic variation and reconstruct the history of replicate invasions on three continents in a globally invasive bird, the Monk parakeet (*Myiopsitta monachus*). We evaluated whether genetic diversity at invasive sites could be explained by (1) the native source populations from which they were derived, and (2) demographic bottlenecks during introduction. Genetic data indicated a localized source area for most sampled invasive populations, with limited evidence for admixing of native source populations. This pattern largely coincides with historical data on pet-trade exports. However, the invasive populations are genetically more similar than predicted from the export data alone. The extent of bottleneck effects varied among invasive populations. The low genetic diversity, evidence of demographic contraction and restricted source area observed do not support the hypothesis that invasion is favored by the mixing and recombining of genetic variation from multiple source populations. Instead, they suggest that reduced genetic variation through random processes may not inhibit successful establishment and invasion in this species. However, convergent selection across invasive sites could also explain the observed patterns of reduction and similarity in genetic variation and/or the restricted source area. In general, the alternative explanation of intraspecific variation in invasive potential among genotypes or geographic areas is neglected but warrants more attention as it could inform comparative studies and management of biological invaders.

INTRODUCTION

Biological invasions are a major component of global change, with potentially large detrimental effects on public health, agriculture and biodiversity (Sakai *et al*. 2001; Mack *et al*. 2000; Simberloff *et al*. 2013). Identifying the biological attributes of successful invaders is among the most pressing questions still to be answered (Kolar & Lodge 2001; Lockwood *et al*. 2007). Some research has focused on the genetic variability of initial founder populations as a key predictor of invasion success. High genetic variability could increase establishment success if it increased the likelihood that some individuals possessed genetic variants more suited to the new environment (Facon *et al*. 2006, 2008; Kolbe *et al*. 2004; Lavergne & Molofsky 2007; Lee 2002; Roman & Darling 2007; Suarez & Tsutsui 2008). Invasive populations may have high genetic variability if a large number of individuals are introduced or if individuals stem from multiple genetically differentiated native source populations.

Yet, previous studies have uncovered a broad range of patterns regarding the relationship between genetic diversity and invasion success: invasive populations can stem from both single and multiple native sources and can have higher or lower genetic diversity relative to native populations (reviewed in Novak & Mack 2005; Wares *et al*. 2005; Roman & Darling 2007). Because of this lack of consistency, there is no consensus on whether invaders stemming from multiple native origins are more successful than those from single populations or whether demographic bottlenecks may limit a species' invasion success.

Understanding the historical context of an invasion could provide important insights into the role of genetic variability in invasion success. By comparing genetic variability in native and invasive populations it is possible to deduce the demographic and evolutionary changes (including genetic drift and selection) that shaped the introduced population (Dlugosch & Parker 2008; Fonseca *et al*. 2010). However, inferring processes underlying successful invasion

remains analytically challenging, largely because of a lack of information about invasion history (Estoup & Guillemaud 2010). This lack of historical context could lead to errors in the identification of the sources of invasive populations, which are expected to be more likely when populations are minimally structured in their native range or if sampling in the native area has been incomplete or inappropriate. Furthermore, genetic divergence between native and invasive populations may occur rapidly during the invasion process (e.g. through drift or selection) such that divergence might confound inference of the source population(s) (Estoup & Guillemaud 2010). To understand the interaction between genetic diversity and invasive potential, it is critical to obtain information on population genetic structure and composition from both native and invasive ranges, and with a sufficient geographic coverage to track most of the genetic diversity potentially sampled during the invasion process.

Birds probably constitute the best studied taxa to identify life history traits associated with invasion success, given the well-recorded and deliberate worldwide introductions of hundreds of species (e.g. Blackburn *et al*. 2009; Sol *et al*. 2012). However, very little is known regarding the genetic processes linked to successful establishment of exotic bird species (Blackburn *et al*. 2009). One of the most notorious and widespread orders of invasive birds are parrots (Psittaciformes; Blackburn *et al*. 2009). We focus here on the Monk parakeet (*Myiopsitta monachus*), a successful invader with a native range restricted to southern South America and with invasive populations occurring worldwide (Lever 2011, Figure 1). In contrast to past deliberate introductions, these invasions were formed as an unintentional byproduct of the pet trade. Millions of wild-caught parakeets have been transported from their native range to pet shops and homes across the globe, and a number of mostly accidental escapes or small-scale releases resulted in the establishment of new populations (Carrete & Tella 2008; Russello *et al*. 2008).

Previous studies have focused on determining the geographic origins and source populations for invasive Monk parakeets. An analysis comparing mitochondrial DNA (mtDNA) control region sequences between invasive populations in the United States of America (USA) and native populations in South America concluded that the source for USA invasive populations is likely in the northern region of Argentina, but that unsampled populations may have also contributed to the invasion (Russello *et al*. 2008). Although mtDNA is useful in detecting the historical origin(s) of an invasion in cases where there is sufficient geographic structure in the native range, it provides limited power to infer demographic and genetic processes during and after invasion. A subsequent study based on hypervariable microsatellite loci revealed that high propagule pressure and long-range dispersal in the invasive range likely contributed to Monk parakeet invasion success in the USA (Gonçalves da Silva *et al*. 2010). It remains unknown whether inferences from the USA populations apply to invasive populations elsewhere in the world, or, alternatively, whether these invasive populations have distinct invasion histories.

In this study, we aim to unravel the global invasion history of the Monk parakeet, both in terms of geographic origins and demographic processes. We combined the mtDNA haplotype and nuclear microsatellite data previously collected from populations in the native range in South America and the invasive range in the USA (Russello *et al*. 2008; Gonçalves da Silva *et al*. 2010) with newly-collected data from a broadly expanded sampling of the native range (including the previously unsampled southern portion) and that of invasive populations from two other continents (Europe and Africa). Our goal was to evaluate whether genetic variation observed in established invasive populations could be explained by (1) the number, identity, and characteristics of native source populations from which invasive populations were derived, or by (2) effects of demographic bottlenecks during the introduction. We also

explore whether invasion histories differ between North America and Europe. Additionally, we compare the results obtained by our genetic approach with detailed spatio-temporal historical records on the Monk parakeet pet trade. We place our results in the context of the role that genetic diversity may play in promoting invasion success. Finally, we discuss the extent to which natural selection might have influenced genetic variation and patterns in our putative neutral markers, and the potential importance of selection within the context of invasive species biology.

MATERIALS AND METHODS

The first published records of escaped Monk parakeets in Spain are from 1975, when the species established in Barcelona (Batllori & Nos 1985), followed by establishment on Canary Islands (Tenerife) in 1980, Madrid in 1985, Mallorca in 1986, and Zaragoza in 1991 (Carrete, Anadon & Tella, unpubl. data). In the USA, the first records of established populations are from the 1960's, with separate populations becoming established in Florida in 1969 (Owre 1973), New Jersey in 1970 (Niedermyer & Hickey 1977), and Connecticut in 1973 (Olivieri & Pearson 1992). However, the data from the long-term annual Audubon Christmas Bird Count (CBC, http://netapp.audubon.org/cbcobservation/) indicate that initial populations in New Jersey and Connecticut may have gone extinct or nearly so, and were subsequently augmented or reestablished in the late 1980s/early 1990s. All these dates of establishment should be viewed in the context of the life-history of the species: we estimate life expectancy of full-grown parakeets to be about five years based on survival rates (Conroy & Senar 2009), whereas young birds are nearly two years old when they first reproduce (Martín & Bucher 1993). Historical records suggest that all of these introductions were independent of each other, although all had their original source in animals moved from South America by the pet trade.

144 Likewise, there are no indications of exchange or transfer among different sites within either
145 Spain or the USA, or between continents as reported by the CITES Trade Data Base
146 (www.cites.org).

147 ***Sampling***

148 Samples were collected at 22 sites: 14 in the native range in South America, four in the
149 invasive range in Europe (Spain), one from an African island, and three in the invasive range in
150 North America (USA) (Table 1). In Spain we also sampled recently imported wild-caught birds
151 provided by three pet shops / pet owners (Pet Shops). This sample can be considered a rare
152 sampling of an invader during the transport stage of invasion process, prior to potential
153 introduction into the novel range. Sampling locations are further specified in Table 1 and
154 Figure 2, and additional information on the USA samples and several South American samples
155 can be found in Russello *et al*. (2008) and Gonçalves da Silva *et al*. (2010). Newly collected
156 blood samples from wild individuals were collected by venipuncture and preserved in ethanol
157 before extraction. DNA isolation followed standard phenol–chloroform extraction protocols
158 (Sambrook *et al*. 1989) or Qiaquick DNEasy DNA extraction kits (Qiagen). For museum samples
159 from Boquerón, Paraguay (collection of Estación Biológica de Doñana-CSIC, Spain, collected in
160 the 1960´s), DNA isolation was carried out in a laboratory free from PCR products and
161 especially designated for museum samples. For these last samples, four independent PCR
162 replicates were performed for both mitochondrial and microsatellite markers.

163 ***Mitochondrial DNA***

164 We amplified and sequenced a 439-bp fragment of the control region for all 23 populations
165 following Russello *et al*. (2008) and Eberhard *et al*. (2001). Polymerase chain reaction (PCR)

amplification and cycling conditions were as follows: denaturation for 2 min at 94°C, followed by 30 cycles of denaturation at 95°C for 30 s, annealing at 56°C for 30 s, and an extension at 72°C for 90 s. PCRs consisted of 4 µl of DNA extract (40-60 ng of DNA) in a final volume of 20 µl, containing 1.5 mM MgCl2, 0,25mM dNTPs, 2 pmol each primer, 0.5 unit of *Taq* polymerase (Bioline). Amplified products were sequenced on an automated sequencer (ABI 3100, Applied Biosystems, Foster City, CA). Sequence data were edited and aligned in SEQUENCHER 4.5 (Gene Codes Corporation, Ann Arbor, MI) and Bioedit (Hall 1999) and manually checked. Sequences were aligned with previously published sequences in GenBank (Russello *et al*. 2008) to determine haplotype identity. Haplotype diversity (H_D) was calculated after Nei & Tajima (1981).

Nuclear microsatellites

A total of seven microsatellite markers developed by Russello *et al*. (2007) were used in this study and analyzed in 16 populations (Table 1). PCRs were carried out in 25 µl using 12.5 µL of QIAGEN Multiplex PCR master mix, 6 µL of RNase free water (provided with the QIAGEN master mix), 2.5 µl of Primers mix (4 µL of each primer at a final concentration of 2 µM) and 4 µl of DNA template (40-60 ng of DNA). Cycling parameters were as follows: 5 min at 95 °C and 30 sec at 95 °C, 90 sec at 55 °C, 30 sec at 72 °C repeated 32 times followed by 30 min at 60 °C. PCR products were run on 1.5% agarose gels and a posteriori on an ABI3100 DNA analyzer to determine DNA sizes. GENEMAPPER v1.90 (SoftGenetics LLC®) was used to score alleles and genotypes. Allele assignments were calibrated using samples of one population analyzed in both laboratories.

Departures from linkage equilibrium and Hardy–Weinberg equilibrium (HWE) were tested using exact tests based on Markov chains (10,000 de-memorizations, 1000 batches,

189 5000 iterations per batch), as implemented in GENEPOP on the web (Raymond & Rousset
190 1995; Rousset 2008). The inbreeding coefficient (F_{IS}) and unbiased expected heterozygosity
191 (H_E) were estimated using GENETIX v.4.03 (Belkhir *et al*. 2004). Allelic richness corrected for
192 sample size was determined using HP-RARE (Kalinowski 2005).

193 ***Population structure analyses***

194 The partitioning of the total genotypic variation into different genetic clusters was assessed by
195 two methods. First, we performed a Factorial Component Analysis (FCA) with default settings
196 in GENETIX, which determines the axes of genetic variation that best differentiate among pre-
197 defined populations based on population allelic frequencies. We then plotted the individuals in
198 this genetic space in order to evaluate population overlap. Second, we employed the model-
199 based clustering method implemented in STRUCTURE version 2.3.4 (Pritchard *et al.* 2000), which
200 assigns individuals to clusters that are derived without information on population
201 membership. We ran STRUCTURE for 10 replicate runs each for K = 1-16 using the default
202 parameters for an admixture model, no sampling site information, correlated allele
203 frequencies between populations, a burn-in chain length = 100,000 and a Markov chain Monte
204 Carlo length = 100,000. We used STRUCTURE HARVESTER (Earl & von Holdt 2012) to determine the
205 most likely K following the Evanno method (Evanno *et al.* 2005). The individual population
206 assignment graphs for the 10 replicate runs for the most likely K compiled using CLUMP 1.1.2
207 (Jakobsson &Rosenberg 2007) and default parameters for the Greedy algorithm. The
208 composite assignments were graphically displayed using DISTRUCT 1.1 (Rosenberg 2004).

RESULTS

MtDNA haplotypes

We found 19 haplotypes across our 23 population samples (Figure 2). Six of these haplotypes (32%, haplotypes NH01-6: GenBank accessions KP873200-KP873205) had not been previously reported. Of these, haplotypes NH04 and NH05 showed well-defined polymorphisms (overlapping fluorescence peaks of equal heights) which were maintained even after repeated sequencing of the same individuals. Since duplication of the control region does not occur in this species (Schirtzinger *et al*. 2011), these polymorphisms likely indicate the presence of heteroplasmy in the mitochondrial genome.

Within the native range, populations were diverse and differentiated, and frequencies of haplotypes varied considerably over relatively short distances (Figure 2). An exception to this pattern was a cluster of populations at the southern end of the native range, which were composed of only two haplotypes (Shared01 and NH01, the last one unique to this cluster). These two haplotypes were found in similar proportions, even at relatively distant sites (Figure 2).

Only seven out of the 19 haplotypes (37%) were found in samples from the invasive ranges. All established populations from both the European and North American invasive range were dominated by the same haplotype (Monach1), which occurred in low frequencies in just two native populations (Entre Ríos and Rio Grande do Sul; Figure 2). The population from Canary Islands differed somewhat in that Monach1 was less dominant and haplotype diversity was higher. Haplotype NH05 was unique to the invasive range and was not documented in any of our samples from the native range (Figure 2).

Wild-caught birds sampled in Spanish Pet Shops (i.e. before their potential introduction into the invasive range) were more diverse than invasive populations (Table 1).

233 Interestingly, Monach1 was not the dominant haplotype in the Pet Shops, thus this sampling
234 more closely resembled some of the native populations rather than the invasive populations in
235 Spain (Figure 2). Overall, transient birds (Pet Shops) and invasive populations showed the
236 greatest similarity in haplotype composition with populations from Entre Ríos on the border of
237 Argentina and Uruguay, and Rio Grande do Sul (Brazil) (Figure 2).

238 ***Nuclear microsatellites***

239 Across the 16 populations analyzed (Table 1), expected heterozygosities of the seven loci
240 ranged between 0.51 and 0.70, while rarefied allelic richness (N = 8 individuals) varied
241 between 2.55 and 4.24 (Table 1). Global multi-locus Hardy–Weinberg exact tests detected
242 deviations from equilibrium expectations for only two out of the 16 populations (one invasive,
243 one native). Absolute F_{IS} values averaged across loci were low in all populations (< 0.10;
244 significant, and negative, in only one population), with an average across populations of -
245 0.0096. Loci appeared unlinked as only one comparison in one population remained significant
246 following sequential Bonferroni correction (data not shown).

247 Genetic diversity was highest in the native range, but decreased towards the southern
248 end (Table 1). Invasive populations were overall less diverse, but levels of diversity did vary
249 among populations, with the Canary Islands population being the most diverse (Table 1). The
250 Factorial Correspondence Analysis uncovered structuring of genotypic variation among
251 populations (Figure 3). The first three axes described 47%, 23% and 17% (88% in total) of the
252 total among-population variation. Invasive populations from the USA clustered together with
253 invasive populations from mainland Spain and birds from the Pet Shops. Populations from the
254 southern end of the native range formed another distinct cluster. The remaining populations in

255 the native range also showed similarity, while the population from Canary Islands was distinct
256 but most resembled the northern populations of the native range (Figure 3).

257 The most likely number of clusters inferred from the STRUCTURE analysis was $K = 3$ ($\Delta K =$
258 20, more than twice as large as any other ΔK). The graphical output of individual population
259 memberships for $K = 3$ (Figure 4) showed that a first cluster was formed by individuals that
260 were almost exclusively encountered in the populations from the southern end of the native
261 range. A second cluster was formed by individuals mostly found in populations from the
262 northern end of the native range, from Canary Islands, from the Pet Shops, and, to a lesser
263 extent, from the invasive USA populations (especially Connecticut) and the Madrid population
264 from Europe. A third cluster was formed by individuals mostly found in populations from both
265 the continental European and the North-American invasive range, and, to a lesser extent, from
266 the Pet Shops.

267 ***Relationship between nuclear and mitochondrial variation***

268 Overall, nuclear and mitochondrial genetic diversity appear correlated across populations in
269 both the native and invasive range (Figure 5). Populations from the southern part of the native
270 range have a lower diversity than those from the north for both marker types (Table 1, Figure
271 5). Similarly, populations from the invasive range generally have lower diversity than those
272 from the native range for both marker types: some populations are even fixed for a single
273 mtDNA haplotype. In contrast, the birds from the Spanish Pet Shops have relatively high
274 mitochondrial diversity (Table 1, Figure 5).

275 **DISCUSSION**

276 We used patterns of variation at mtDNA control region sequences and nuclear microsatellites
277 to reconstruct the history of replicate invasions by the South American Monk parakeet on

three continents. Our goal was to evaluate whether genetic variation observed in established invasive populations could be explained by (1) the native source populations from which invasive populations were derived, and (2) genetic effects of demographic bottlenecks during the introduction. Nuclear microsatellite and mtDNA haplotypes both exhibited strong and consistent patterns of geographic structuring. Genetic diversity was highest in the northern parts of the native range. This northern area was identified as the most likely native source for invasive populations, and genetic analyses provide evidence for a single native source for virtually all sampled invasive populations. Nonetheless, genetic diversity varied among invasive populations and was overall lower than in native populations. Although these patterns indicate that genetic bottlenecks likely reduced the diversity of invasive populations compared to the native source, many of these invasive populations are thriving. The low genetic diversity, evidence for bottleneck effects, and the restricted area of native source populations that we observed in this highly successful invader do not support the hypothesis that high genetic variation inherently favors biological invasion, or that invasion is favored by the combining or mixing of genetic variation from multiple source populations. Below we discuss these results in more detail and relate them to known historical patterns of transport of birds via the global pet trade.

Spatial genetic structuring in native range

We found evidence for strong spatial structuring of genetic diversity. In the native range, genetic diversity decreased along a north-south axis in the native range (Figures 2 and 5, Table 1). The high genetic diversity and structuring at the northern end of the native range suggests that populations are relatively stable here and that dispersal is relatively restricted in this species. Short dispersal distances for this species have been reported in the native range based

on mark-recapture methods (Martín & Bucher 1993); although genetic evidence has suggested longer dispersal events may occur in invasive populations (Gonçalves da Silva *et al.* 2010). In contrast, there is less structuring in the southern end of the native range. There is no evidence that this is due to a difference in dispersal rates. Instead, lack of geographic structure can occur as the result of a recent expansion of the range (Avise 2004). Indeed, such an expansion (filling up a gap in the distribution) has been well-documented for the Pampas region of Argentina (Bucher & Aramburú 2013). Interestingly, the southern populations we sampled lie on opposite sides of this recently invaded area yet are genetically very similar, suggesting that they may be part of a larger expansion that predated the 20th century expansion into the Pampas documented by Bucher & Aramburú (2013). Further sampling is necessary to confirm and clarify this pattern.

When native populations are strongly structured in neutral genetic markers, this typically indicates reduced dispersal among populations. Reduced dispersal generally increases the potential for local adaptation to emerge (Lenormand 2002). In that case, it therefore becomes more important to establish which areas or populations have acted as sources. At the same time, stronger spatial structuring allows for more accurate identification of the origin of invasive populations. However, our results may act as a warning that the degree of population structuring can itself be heterogeneous: structuring is much stronger among northern than among southern native populations (Figures 2,3). Local results on population structuring may therefore not generalize range-wide. We therefore recommend that (in the absence of any other information) studies directed towards inferring source populations start with a very broad but coarse sampling, and then iteratively sample areas at a finer-scale that might contain putative source localities.

Inferring source populations

The strong structuring of native populations allows insight into the invasion pathways of the Monk parakeets. Most sampled native populations can be discounted as potential source localities as the general haplotype composition of invasive populations differed substantially from those in the native range. There are, however, relatively close fits to the haplotype compositions for the native populations of Entre Ríos and Rio Grande do Sul (Figure 2D). This is especially clear for the Monach1 haplotype, which is dominant in all sampled invasive populations but virtually absent in all sampled native populations except for Entre Ríos and Rio Grande do Sul. However, even in these two native populations the Monach1 haplotype is not dominant. This pattern suggests that the source populations could be even more spatially restricted than what our current sampling can resolve, and might lie between the two putative native source populations in Uruguay. Such a restricted source area is also indicated by the microsatellite data, because the sampled invasive populations are genetically quite similar, suggesting they share a similar origin, but are distinct from anything we have sampled in the native range.

Comparison with historical geographical data on transports

Another approach to deduce source areas of biological invasions is the use of historical records on the movements of organisms, if available (Blackburn *et al*. 2009; Estoup & Guillemaud 2010). For the Monk parakeet, natural overseas dispersal events are highly unlikely as this species, like most parrots, is non-migratory (Forshaw 1989). We also find it highly unlikely that this bird would be accidentally transported (e.g., stow-away in a plane). In contrast, close to 1,000,000 wild-caught individuals have been exported across the world to be sold as pets (CITES Trade Data Base, www.cites.org). While the numbers obtained from CITES are only

approximate, our summary of the database indicates that Uruguay has been the main exporter of Monk parakeets in the world from 1980 onwards (Figure 6). This observation corroborates our conclusion based on the genetic data. This conclusion is further supported by the mtDNA haplotype obtained from a single Uruguayan sample (Russello *et al*. 2008). This individual had the Monach2 haplotype, which is the second-most common haplotype across the invasive populations but spatially restricted in the native range (Figure 2D,E). Hence, the historical transport data appear to corroborate our genetic assessment that there is a single main source for most invasive populations, and that it is likely located in Uruguay.

However, the relative proportions of Monk parakeets imported from Uruguay versus Argentina differ considerably between Spain and the USA, and among years (Figure 6). Moreover, data collected by the US Fish & Wildlife Service (Form 3-177 reports) indicate that before 1980 (when at least the invasive Florida population established in the USA), Paraguay was the principal source. Together, these data would predict variation in genetic composition among invasive populations of Monk parakeets, since these became established during a wide temporal window (1969-1991) and in different countries. This prediction contrasts with our observation of high genetic similarity among invasive populations, suggesting a similar origin. We therefore conclude that well-sampled genetic data provides a more comprehensive picture of which native populations actually contributed to invasive populations as it integrates over individuals that may have been transported in different years or from different sources and held in captivity for some time before founding or joining invasive populations. Furthermore, the genetic approach is the only option available for many invasive species for which no historical trade or transport data are available.

Reduced genetic diversity in invasive populations

One striking pattern we recovered was the lower mitochondrial haplotype diversity and microsatellite allelic richness in the invasive populations. This lower genetic variation in invasive populations likely stems from two effects. First, reduced genetic diversity may be a characteristic of the native source population. The strong genetic similarity among invasive populations suggests that their resemblance is due to a common origin; if this source area had low genetic variation to begin with, subsequent invasive populations would also exhibit low genetic variability. Our samples from the native range show that genetic diversity does vary considerably among native populations (Table 1). However, because we do not have population genetic samples that exactly correspond to the inferred native source, this hypothesis cannot yet be tested directly. Second, genetic diversity in both markers is especially low for some populations like Connecticut and Zaragoza (Figure 5), which may be indicative of a demographic bottleneck. In contrast, the Canary Islands population has the highest genetic diversity of all invasive populations (Table 1) and, to the best of our knowledge, is the only deliberately introduced invasive Monk parakeet population involving dozens of released and supplementary-fed individuals (R. Riera, pers. comm.). However, it is worth pointing out that the Canary Islands has a different microsatellite composition and, alternatively, may have been founded from a source population with more genetic diversity.

Invasive success versus genetic diversity

The low genetic diversity, evidence for bottleneck effects, and the restricted area of native source populations that we observed in this highly successful invader do not support the hypothesis that high genetic variation inherently favors biological invasion, or that invasion is favored by the combining of genetic variation from multiple source populations (Blackburn *et*

392 *al*. 2009; Facon *et al*. 2006, 2008; Kolbe *et al*. 2004; Lavergne & Molofsky 2007; Lee 2002;
393 Roman & Darling 2007; Suarez & Tsutsui 2008). Instead, we find that a single, spatially
394 restricted source area likely has given rise to virtually all successful invasive populations across
395 different continents, with little evidence for admixture of multiple native source populations.
396 Furthermore, our results suggest that this restricted native source population most likely had
397 reduced genetic variability to begin with, and that bottlenecks during invasion reduced this
398 variation even more. Nonetheless, the invasive populations are viable and have high initial
399 population growth rates. As an extreme example, the Zaragoza population from Spain is
400 thought have been established by perhaps as little as 2 or 3 individuals in 1991, is fixed for a
401 single haplotype and has the lowest nuclear heterozygosity and allelic richness that we
402 detected across our sampling. Yet, this population grew to a size of over 1,000 in 15 years,
403 which means an average population growth rate of nearly 50% per year (Carrete, Anadon &
404 Tella unpubl. data). Even if the number of founders was higher, a growth rate of >20% was
405 likely experienced. Hence, we can conclude that high genetic diversity per se is not critical for
406 successful establishment in this species. Instead, there might be particular traits that are
407 characteristic for this species that make it such a successful invader. These may include the
408 capacity to build its own nest instead of relying on cavities for breeding, tolerance of human
409 disturbance, and dietary flexibility (Strubbe & Matthysen 2009; Carrete & Tella 2011; Bucher &
410 Aramburú 2013). Nonetheless, high propagule pressure (close to 1 million individuals
411 exported) will have also facilitated invasion.

412 ***Might selection explain observed genetic patterns?***

413 The dominance of a single haplotype (Monach1) in all independently-established continental
414 invasive populations compared to the low frequency of this haplotype in native populations

(Figure 2) is striking. In addition it has a higher frequency in invasive populations than in the transient (pre-establishment) Pet Shops sample (Figure 2). Similarly, it is predominant in the populations from Connecticut and Canary Islands (Figure 2) even though these populations are distinct from other invasive populations with regards to microsatellite variation (especially Canary Islands; Figures 3 and 4). These observations could be interpreted as a signature that natural selection favors this haplotype within invasive populations, putatively linked to specific variants within non-recombining mitogenomic coding regions. If convergent selection is acting on invasive populations, what are the underlying drivers? Climates and associated vegetations vary greatly across the invasive range, with an average winter temperature of 18°C on the subtropical Canary Islands versus -3°C in cold-temperate Connecticut, suggesting that such factors are not driving convergent selection. (As an aside, it does appear as if populations exposed to lower average winter temperatures (Connecticut, New Jersey, Zaragoza) have lost more genetic diversity than populations with higher temperatures (Florida, Canary Islands, Mallorca; Table 1, Figure 5). One interpretation could be that colder climates have caused greater demographic bottlenecks, e.g. due to mortality related to cold spells. An independent set of populations would be needed to properly test this suggestive pattern).

One aspect that all invasive populations do share is that they occur in urban environments, which have been shown to exert selection on genes related to behavior in other avian populations (Mueller *et al*. 2013). Future comparisons of invasive and native populations that sample more widely across the genome may help detect whether specific genes have responded to selection (e.g. Puzey & Vallejo-Marín 2014) imposed by the novel urban settings and whether any of these are functionally linked to the Monach1 haplotype.

Alternatively, the haplotype Monach1 might be dominant in the invasive range because it is already dominant in a restricted but unsampled source area that we inferred using both

marker types. It is notable that the historic trade-data document that exports originated from a broad area involving several countries (Paraguay, Argentina, Uruguay), yet we do not see a genetic signal of such diverse origins in the invasive populations. This disparity indicates that Monk parakeets from some source areas (e.g. Paraguay) failed to establish. It further suggests that there might be some characteristics particular to Monk parakeets from a restricted subset of the native range from which exports originated that is favored by selection in the novel range, for example a certain (potentially behavioral) urban phenotype. In general, this scenario suggests that having propagules originate from more areas would increase the likelihood that some suitable individuals have been introduced, favoring establishment and subsequent invasion.

Even though our data do not currently permit strong inferences regarding selection and its potential contributions towards shaping observed patterns, we do feel that it provides an alternative explanation that warrant future testing with new genomic approaches. At present, the role of selection in invasion success is often neglected. A limited number of intraspecific studies have shown that invasive potential may differ considerably between introduced populations from the same species (e.g. Ciosi *et al*. 2008; Kang *et al*. 2007; Kelly *et al*. 2006). We argue (see also Carrete *et al*. 2012) that taking into account intra-specific variation in invasive potential may yield further insights, additional options for effective management of biological invasions, and improved prediction of the potential range limits of invaders (e.g. when based on climatic niche modeling).

ACKNOWLEDGEMENTS

We thank Kevin Burgio, Miguel Santillán, the Laboratorio de Ecología Molecular, and the Estación Biológica de Doñana for providing help and samples; Jessica Eberhard, landowners in

462 Argentina, Paraguay and Spain, James Lindsay (Florida Power and Light), Eric Tillman and
463 Monte Chandler (USDA), Ana Anechina (Ayuntamento de Zaragoza), R. Riera and G. Blanco for
464 helping us to obtain samples; Enrique Bucher and Rosana Aramburú for information and
465 discussion. Sampling protocols in Argentina were approved by the New Mexico State
466 University Institutional Animal Care and Use Committee (protocol no. 2006-027) and permits
467 for sampling in Entre Ríos, La Pampa and Buenos Aires provinces were provided by the
468 Argentinean Department of Natural Resources. Funding for the project was provided by P08-
469 RNM-4014 and Fundación Repsol (to JLT and MC), the Spanish Ministry of Economy and
470 Competitiveness (RYC-2011-07889 and CGL2012-35232 with support from the ERDF to PE, and
471 CGL2012-38262 to JCS), Canadian Foundation for Innovation (MR), National Institutes of
472 Health (S06GM008136 to TFW), and the National Institute for Mathematical and Biological
473 Synthesis (Postdoctoral Fellowship to EAH, through NSF #DBI-1300426 with additional support
474 from the University of Tennessee, Knoxville). Barcelona data collection was partially supported
475 by Servei de Vigilància i Control de Plagues Urbanes from Agència de Salut Pública de
476 Barcelona. Further support was provided by the European COST Action “ParrotNet” (ES1304).

477

REFERENCES

Avise JC (2004) *Molecular Markers, Natural History, and Evolution*. Sinauer Kluwer Academic Publishers.

Batllori X, Nos R (1985) Presencia de la Cotorrita gris (*Myiopsitta monachus*) y de la Cotorrita de collar (*Psittacula krameri*) en el área metropolitana de Barcelona. *Miscellània Zoològica*, **9**, 407-411.

Belkhir K, Borsa P, Chikhi L *et al.* (2004) GENETIX 4.05, logiciel sous Windows pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Blackburn TM, Lockwood JL, Cassey PB (2009) *Avian Invasions: the Ecology and Evolution of Exotic Birds*. Oxford University Press, Oxford, UK.

Bucher EH, Aramburú RM (2014) Land-use changes and Monk parakeet expansion in the Pampas grasslands of Argentina. *Journal of Biogeography,* **6**, 1160-1170.

Carrete M, Tella J (2008) Wild-bird trade and exotic invasions: a new link of conservation concern? *Frontiers in Ecology and Environment*, **6**, 207–211.

Carrete M, Tella JL (2011) Inter-individual variability in fear of humans and relative brain size of the species are related to contemporary urban invasion in birds*. PloS One*, **6**, e18859.

Carrete M, Edelaar P, Blas J *et al.* (2012) Don´t neglect pre-establishment individual selection in deliberate introductions. *Trends in Ecology and Evolution*, **27**, 67-68.

Ciosi M, Miller NJ, Kim KS *et al.* (2008) Invasion of Europe by the western corn rootworm, *Diabrotica virgifera virgifera*: multiple transatlantic introductions with various reductions of genetic diversity. *Molecular Ecology*, **17**, 3614–3627.

Conroy MJ, Senar JC (2009) Integration of demographic analyses and decision modeling in support of management of invasive Monk parakeets, an urban and agricultural pest. In: *Modeling Demographic Processes in Marked Populations* (eds. Thompson DL *et al*.), pp. 491–510. Springer Science+Business Media LCC, New York, USA.

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**, 431–449.

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

Eberhard JR, Wright TF, Bermingham E (2001) Duplication and concerted evolution of the mitochondrial control region in the parrot genus *Amazona*. *Molecular Biology and Evolution*, **18**, 1330-1342.

Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology*, **19**, 4113-4130.

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.

Facon B, Genton BJ, Shykoff J *et al.* (2006) A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology and Evolution*, **21**, 130–135.

Facon B, Pointier JP, Jarne P *et al.* (2008) High genetic variance in life-history strategies within invasive populations by way of multiple introductions. *Current Biology*, **18**, 363–367.

Fonseca DM, LaPointe DA, Fleischer RC (2000) Bottlenecks and multiple introductions: population genetics of the vector of avian malaria in Hawaii. *Molecular Ecology*, **9**, 1803–1814.

Forshaw J M (1989) *Parrots of the World*. Landsdowne Editions, Chipping-Norton, Australia.

Gonçalves da Silva A, Eberhard JR, Wright TF *et al.* (2010) Genetic evidence for high propagule pressure and long-distance dispersal in Monk parakeet (*Myiopsitta monachus*) invasive

526 populations. *Molecular Ecology*, **19**, 3336-3350.
527 Hall TA (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis
528 program for Windows 95 / 98 / NT. *Nuclic Acids Symposium Series*, **41**, 95-98.
529 Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for
530 dealing with label switching and multimodality in analysis of population structure.
531 *Bioinformatics*, **23**, 1801-1806.
532 Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on
533 measures of allelic richness. *Molecular Ecology Notes*, **5**, 187-189.
534 Kang M, Buckley YM, Lowe AJ (2007) Testing the role of genetic factors across multiple
535 independent invasions of the shrub Scotch broom (*Cytisus scoparius*). *Molecular Ecology*,
536 **16**, 4662–4673.
537 Kelly DW, Muirhead JR, Heath DD *et al.* (2006) Contrasting patterns in genetic diversity
538 following multiple invasions of fresh and brackish waters. *Molecular Ecology*, **15**, 3641–
539 3653.
540 Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. *Trends in Ecology*
541 *and Evolution*, **16**, 199–204.
542 Kolbe JJ, Glor RE, Schettino LRG *et al.* (2004) Genetic variation increases during biological
543 invasion by a Cuban lizard. *Nature*, **431**, 177–181.
544 Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the
545 success of an invasive grass. *Proceedings of the National Academy of Sciences of the United*
546 *States of America*, **104**, 3883–3888.
547 Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology and Evolution*, **17**,
548 386–391.
549 Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology and*
550 *Evolution*, **17**, 183–189.
551 Lever C (2005) *Naturalised Birds of the World*. T & AD Poyser, London, UK.
552 Lockwood J, Hoopes M, Marchetti M (2007) Invasion Ecology. Blackwell Publishing, Oxford, UK.
553 Mack RN, Simberloff D, Mark Lonsdale W *et al.* (2000) Biotic invasions: causes, epidemiology,
554 global consequences, and control. *Ecological Applications*, **10**, 689-710.
555 Martín LF, Bucher EH (1993) Natal dispersal and first breeding age in Monk parakeets. *Auk*,
556 **110**, 930-933.
557 Mueller JC, Partecke J, Hatchwell BJ *et al.* (2013) Candidate gene polymorphisms for
558 behavioural adaptations during urbanization in blackbirds. *Molecular Ecology*, **22**, 3629-
559 3637.
560 Nei M, Tajima F (1981) DNA polymorphism detectable by restriction endonucleases. *Genetics*,
561 **97**, 145-63.
562 Neidermyer WJ, Hickey JJ (1977) The monk parakeet in the United States, 1970-75. *American*
563 *Birds*, **31**, 273-278.
564 Novak SJ, Mack RN (2005) Genetic bottlenecks in alien plant species. In: *Species Invasions:*
565 *Insights into Ecology, Evolution and Biogeography* (eds. Sax DF, Stachowicz JJ, Gaines SD),
566 pp. 201–228. Sinauer Associates Inc, Sunderland, USA.
567 Olivieri A, Pearson L (1992) Monk parakeets in Bridgeport, Connecticut. *The Connecticut*
568 *Warbler*, **12**, 104-111.
569 Owre OT (1973). A consideration of the exotic avifauna of southeastern Florida. *Wilson*
570 *Bulletin*, **85**, 491-500.
571 Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus
572 genotype data. *Genetics*, **155**, 945-959.

Puzey J, Vallejo-Marín M (2014) Genomics of invasion : diversity and selection in introduced populations of monkeyflowers (*Mimulus guttatus*). *Molecular Ecology*, **23**, 4472–4485.
Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.
Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.
Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology and Evolution*, **22**, 454–64.
Rosenberg NA (2004). DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137-138.
Russello MA, Saranathan V, Buhrman-Deever S *et al.* (2007) Characterization of polymorphic microsatellite loci for the invasive monk parakeet (*Myiopsitta monachus*). *Molecular Ecology Notes*, **7**, 990-992.
Russello MA, Avery ML, Wright TF (2008) Genetic evidence links invasive Monk parakeet populations in the United States to the international pet trade. *BMC Evolutionary Biology*, **8**, 217.
Sakai AK, Allendorf FW, Holt JS *et al.* (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics*, **32**, 305–332.
Sambrook J, Fritschi EF, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York, USA.
Schirtzinger EE, Tavares ES, Gonzales LA *et al.* (2012) Multiple independent origins of mitochondrial control region duplications in the order Psittaciformes. *Molecular Phylogenetics and Evolution*, **64**, 342-356.
Simberloff D, Martin JL, Genovesi P *et al.* (2013) Impacts of biological invasions: what's what and the way forward. *Trends in Ecology and Evolution*, **28**, 58-66.
Sol D, Maspons J, Vall-Llosera M *et al.* (2012) Unraveling the life history of successful invaders. *Science*, **337**, 580-583.
Suarez AV, Tsutsui ND (2008) The evolutionary consequences of biological invasions. *Molecular Ecology*, **17**, 351–60.
Strubbe D, Matthysen E (2009) Establishment success of invasive ring-necked and Monk parakeets in Europe. *Journal of Biogeography*, **36**, 2264-2278.
Wares JP, Hughes AR, Grosberg RK (2005) Mechanisms that drive evolutionary change: insights from species introductions and invasions. In: *Species Invasions: Insights into Ecology, Evolution and Biogeography* (eds Sax DF, Stachowicz JJ, Gaines SD), pp. 229–257. Sinauer Associates Inc, Sunderland, USA.

Data Accessability

Control region DNA sequences: GenBank accessions KP873200-KP873205 (this study) and EU545521-EU545537 (Russello et al. 2008). Haplotypes of all individuals, haplotype frequenties per population, haplotype alignment, and sampling locations and microsatellite genotypes: Dryad doi: http://dx.doi.org/10.5061/dryad.5pr61.

Author Contributions

JLT, SR, MAR, TFW, MC and PE designed the research. All authors contributed to performing the research. SR, AGS, PE, TFW, EAH and MAR analyzed and graphically illustrated the data. PE

wrote the first and final version of the paper with all authors (especially EAH and TFW) contributing to revisions.

Table 1. Overview of populations (full name, country and abbreviation) sampled from the invasive and native ranges. N: number of individuals sampled, H_E: unbiased expected heterozygosity, AR: rarefied allelic richness, H_D: haplotype diversity, Year: approximate year of introduction based on published observations of first continued presence of Monk parakeets at the locality.

		Nuclear microsatellites			MtDNA haplotypes		
Population	**Abbreviation**	**N**	**H_E**	**AR**	**N**	**H_D**	**Year**
Connecticut (USA)	CNCT	19	0.58	2.93	9	0.00	1973/1985?
New Jersey (USA)	NWJY	*NA*	*NA*	*NA*	11	0.55	1970/1990s?
Florida (USA)	FLRD	91	0.63	3.26	43	0.54	1969
Zaragoza (Spain)	ZRGZ	21	0.51	2.55	20	0.00	1991
Madrid (Spain)	MADR	23	0.64	3.47	28	0.27	1985
Barcelona (Spain)	BARC	102	0.61	3.16	91	0.31	1975
Mallorca (Spain)	MALL	40	0.63	3.25	9	0.42	1986
Canary Islands (Spain)	CANR	28	0.65	3.53	21	0.66	1980
Pet Shops (Spain)	PETS	8	0.58	3.20	8	0.71	-
Mato Grosso (Brazil)	MTGS	*NA*	*NA*	*NA*	5	0.90	-
Tucumán (Argentina)	TUCU	*NA*	*NA*	*NA*	5	0.00	-
Concepción (Paraguay)	CCEP	*NA*	*NA*	*NA*	11	0.55	-
Santiago del Estero (Argentina)	SEST	*NA*	*NA*	*NA*	5	1.00	-
Boquerón (Paraguay)	BOQR	7	0.69	4.24	9	0.69	-
Corrientes (Argentina)	CRRT	*NA*	*NA*	*NA*	13	0.73	-
Entre Ríos (Argentina)	ENRS	49	0.70	3.80	37	0.83	-
Rio Grande do Sul (Brazil)	RGSL	*NA*	*NA*	*NA*	6	0.53	-
Algarrobo (Argentina)	ALGA	20	0.57	3.08	10	0.47	-
General San Martín (Argentina)	SMRT	11	0.56	3.23	12	0.41	-
Buenos Aires (Argentina)	BAIR	19	0.57	3.22	12	0.30	-
Parque Luro (Argentina)	LURO	43	0.58	3.09	9	0.50	-
General Rondeau (Argentina)	RDEA	19	0.57	3.16	10	0.53	-
Mayor Buratovich (Argentina)	BURT	9	0.62	3.53	10	0.47	-

622 **Figure legends**

623

624 **Figure 1.** Native range (blue, approximate) and established invasive populations (red, non-
625 exhaustive, including some oceanic islands) of the Monk parakeet.

626

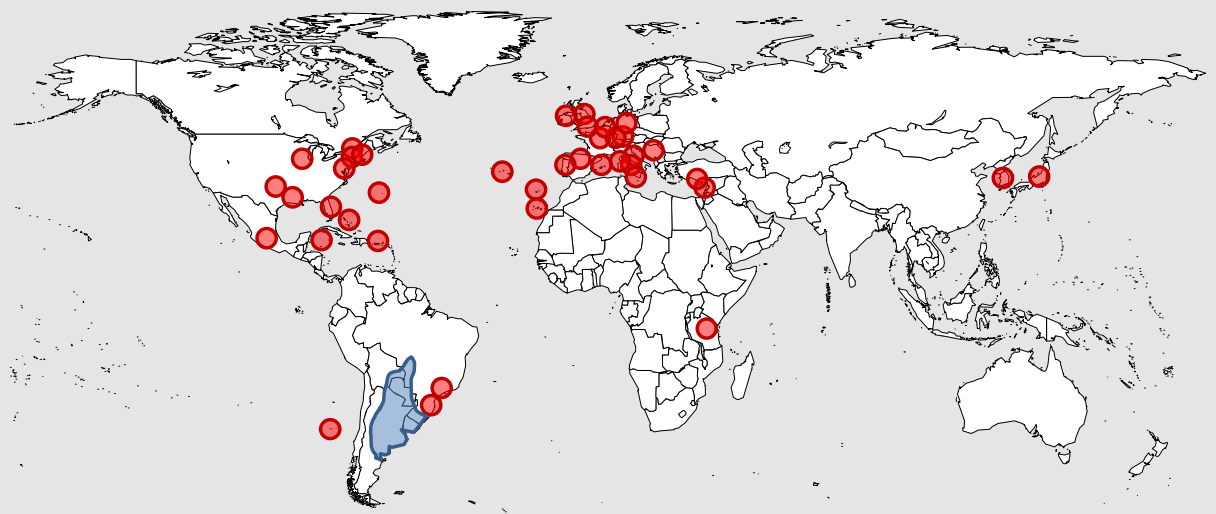
627 **Figure 2.** Overview of mtDNA variation across the native and invasive range of *M. monachus*.
628 A. Distribution of sampled populations across the entire native range (indicated by the dotted
629 line). B. Location of sampled populations in the USA. C. Location of sampled populations in
630 Spain. Wild-caught birds sampled in Pet Shops (in between uptake and potential introduction)
631 do not have a location. D. Haplotype frequencies in each population. The names at the bottom
632 indicate each haplotype, whereas the size of the bubble is proportional to the number of
633 individuals with this haplotype. (Ordering or similarity in color does not refer to haplotype
634 relatedness). E. Proportions of each haplotype across the native range (blue bars, ordered
635 from highest to lowest) and invasive range (red bars). The full names of abbreviated sampling
636 sites are given in Table 1.

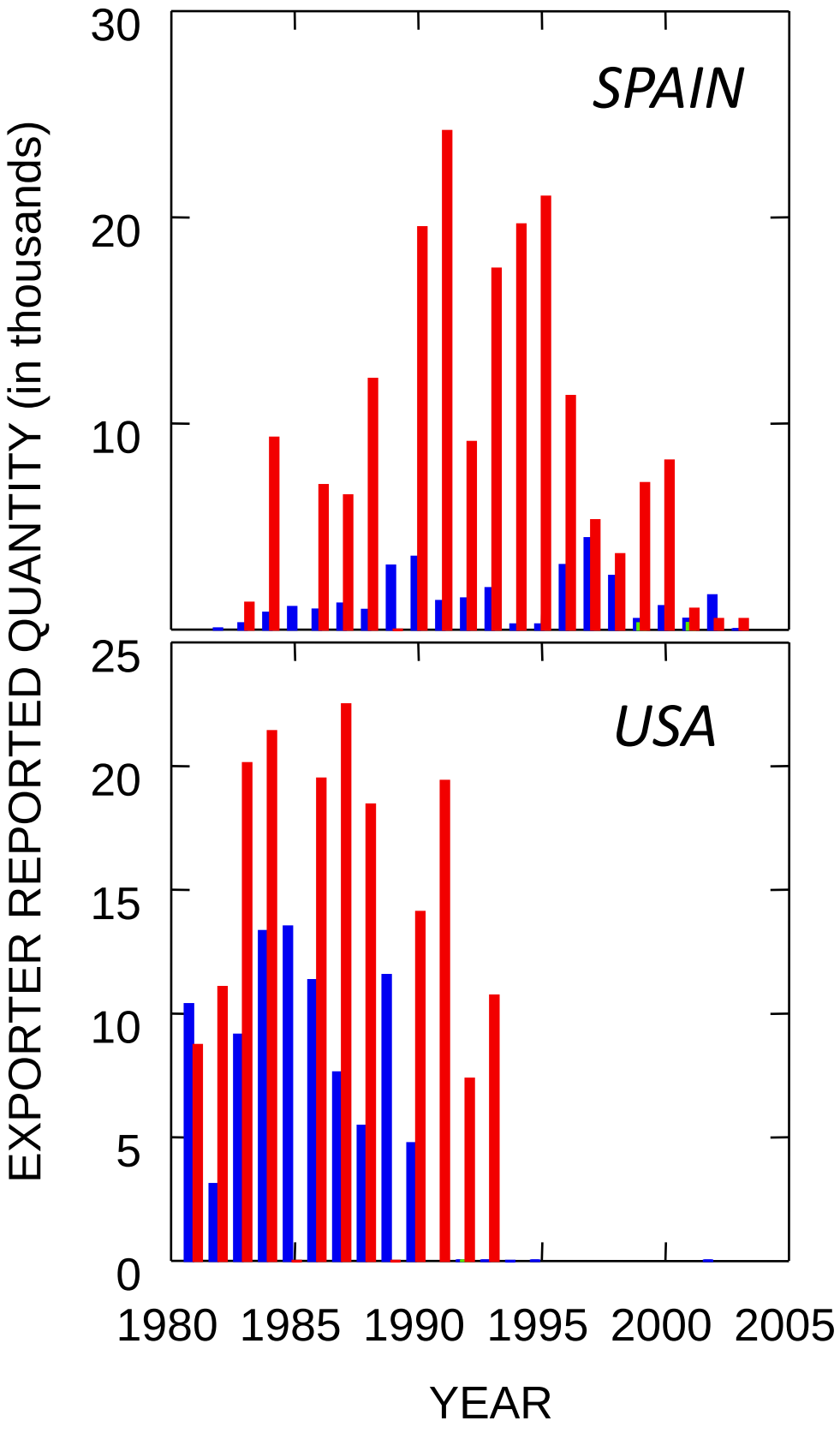
637

638 **Figure 3.** Microsatellite divergence across the native and invasive range, as determined by
639 Factorial Correspondence Analysis. Plotted are individual genotypes along the three axes that
640 best differentiate the genetic divergence among populations. Colored ellipses indicate the
641 approximate ranges of *a priori* and *a posteriori* determined groups (blue for native groups, red
642 for invasive groups).

643

644 **Figure 4.** Individual population membership coefficients estimated by the program STRUCTURE
645 for K = 3 as the most likely number of clusters. Bottom labels refer to each sampled location.
646 Top labels indicate *a priori* population groupings (pale and dark blues for native populations;
647 red for invasive populations; purple for intermediate captive wild birds in Pet Shops). Note that
648 the three clusters uncovered by STRUCTURE correspond well to our *a priori* population
649 groupings, with Canary Islands as the largest exception.
650
651 **Figure 5.** Nuclear (microsatellite) and mitochondrial (control region) genetic diversity in native
652 (blue dots) and invasive populations (red dots), showing how correlated reductions in diversity
653 occur going from native to invasive populations. The sample of birds from the Spanish Pet
654 Shops (representing the transport phase of invasion) is indicated separately in green.
655
656 **Figure 6.** Difference in number of Monk parakeets exported from potential invasion source
657 areas to Spain and the USA (red: Uruguay, blue: Argentina, green: Paraguay – just a few
658 around the year 2000; CITES Trade Data Base, www.cites.org). Note that export data are
659 missing from Uruguay in 1985 and 1989, and that exports to the USA largely stopped in 1994.




SPAIN
USA
EXPORTER REPORTED QUANTITY (in thousands)
30
20
10
25
20
15
10
5
0
1980
1985
1990
1995
2000
2005
YEAR