Spatial heterogeneity in genetic relatedness among house sparrows along an urban–rural gradient as revealed

by individual-based analysis

CARL VANGESTEL,\* JOACHIM MERGEAY,† DEBORAH A. DAWSON,‡ VIKI VANDOMME\* and LUC LENS\*

*\*Terrestrial Ecology Unit, Ghent University, K.L. Ledeganckstraat 35, B-9000 Ghent, Belgium,* †*Research Group of Genetic*

*Diversity, Research Institute for Nature and Forest, B-9500 Geraardsbergen, Belgium,* ‡*Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, South Yorkshire S10 2TN, UK*

## Abstract

Understanding factors that shape patterns of kinship in sedentary species is important for evolutionary ecologists as well as conservation biologists. Yet, how patterns of relatedness are hierarchically structured in space remains poorly known, even in common species. Here, we use information from 16 polymorphic microsatellite DNA markers to study how small-scale kinship structure varies among house sparrows (*Passer domesticus*) along an urban–rural gradient. Average levels of relatedness were higher among urban individuals than among individuals from rural areas, suggesting lower rates of dispersal in more built-up habitats. Comparison of observed levels of relatedness with simulated distributions of known kinship values showed that central urban individuals had the highest proportion of closely related conspecifics in their immediate neighbourhood. Spatial auto-correlograms supported this small-scale genetic structure and further indicated stronger effects of genetic drift and ⁄ or limited dispersal in urban populations. Results of this study underscore the importance of individual-level analyses as a complementary approach to traditional population-level analyses when studying genetic population structure over small spatial scales.

*Keywords*: fine-scale genetic structure, kinship, *Passer domesticus*, urbanization

## Introduction

Assessing the proportion of shared alleles that are iden- tical by descent as a measure of genetic relatedness between pairs of individuals (dyads) has proven infor- mative in a wide range of scientific disciplines (Blouin 2003). For instance, captive or in situ breeding pro- grammes can be optimized by implementing informa- tion on the average degree of relatedness into artificial selection schemes (Lynch & Walsh 1998), quantitative geneticists can use the association between genetic and phenotypic similarity to study the extent of trait inheritance in natural populations (Ritland 1996), and knowledge of kin structure can help to unravel the

Correspondence: Carl Vangestel, Fax: +32 0 9 ⁄ 264 87 94; E-mail: Carl.Vangestel@UGent.be

evolutionary ecology of cooperative behaviour in social species (Danchin *et al.* 2008). Temporal or spatial varia- tion in genetic relatedness can further inform ecologists on spatial dispersal scales or non-random and sex- dependent dispersal events (Sweigart *et al.* 1999; Van de Casteele & Matthysen 2006; Zamudio & Wieczorek 2007), on local population dynamics (Ruzzante *et al.* 2001) and on the most appropriate management strate- gies for species of conservation concern.

Often, conservation geneticists infer rates of dispersal and gene flow from patterns of population genetic structure—with ‘populations’ invariably comprising the smallest units of interest. Individual-level genetic analy- ses are much less common in conservation biology (Blouin 2003), despite the fact that such analyses may reveal small-scale patterns of genetic structure that remain cryptic in population-level ones (Zamudio &

Wieczorek 2007). Increased aggregation of close kin in small, isolated populations may foster loss of genetic variation and inbreeding depression and, as such, com- promise population persistence (Keller & Waller 2002). Monitoring spatial heterogeneity in genetic relatedness can therefore be informative when aiming to identify populations susceptible to (future) genetic threats. Now- adays, access to large numbers of polymorphic micro- satellite DNA markers mitigates the high sampling variance generally associated with estimates of related- ness, which has been a major drawback in the past (Lynch & Ritland 1999; Sweigart *et al.* 1999).

Here, we study fine-scale patterns of genetic related- ness in populations of a highly sedentary species along an urban–rural gradient. Despite the increased interest in urban ecology, nourished by the unprecedented rates of urban sprawling and the knowledge that urban spe- cies will comprise a significant component of future glo- bal biodiversity (Chace & Walsh 2006), the population dynamics and evolutionary ecology of urban species remain poorly understood (Shochat *et al.* 2006). Earlier studies related urbanization to changes in species rich- ness (Marzluff *et al.* 2001), nestling condition (Peach *et al.* 2008) or individual-level responses (Rodewald & Shustack 2008), while more recently, it has been shown that community-level interactions such as predator– prey relationships may become decoupled in urban environments because of anthropogenic-induced shifts in predator and vegetation distributions (Chiron & Jul- liard 2007; Weidinger 2009; Rodewald *et al.* 2011). While individual-based analysis of genetic data is believed to provide a powerful tool to study processes underlying these and other patterns, to date, only few studies have effectively performed so in urban settings (Johnson *et al.* 2009).

House sparrows (*Passer domesticus*) have suffered a

dramatic decline in abundance and distribution during the last decades in both urban and rural areas (Hole *et al.* 2002; Chamberlain *et al.* 2007; De Laet & Sum- mers-Smith 2007). While rural population numbers seem to have stabilized (albeit at lower population equi- libria), urban populations continue to plummet (De Laet

& Summers-Smith 2007). Local extinction of small pop- ulations has transformed uniformly distributed, pan- mictic populations (Kekkonen *et al.* 2010) into patchy population networks, most strongly so in highly urban- ized regions (Shaw *et al.* 2008; C. Vangestel unpub- lished data). Such local population extinction can be expected to distort historical levels of population con- nectivity and to result in accumulation of close relatives within small geographical ranges, particularly in less vagile species. House sparrows are among the most sedentary birds with juveniles dispersing in a ‘step- ping-stone’ manner, post-natal dispersal distances being

typically short and adult birds exhibiting high breeding site fidelity (Heij & Moeliker 1990; Anderson 2006; Ke- kkonen *et al.* 2010). Yet, despite its long history as eco- logical model species in behavioural and experimental studies (Lendvai *et al.* 2007; Nakagawa *et al.* 2007), sur- prisingly little is known on patterns of kinship structure in natural populations of the house sparrow.

In a pilot study along an urban–rural gradient near the city of Ghent, Belgium, a preliminary principal coor- dinate analysis indicated that urban house sparrow populations may be genetically separated from rural ones and that genetic erosive effects may be restricted to the most central urban population (C. Vangestel *et al.*, unpublished). Compared to other populations, the latter one was characterized by the lowest number of distinct alleles while retaining a comparable level of heterozygosity, a signature characteristic of populations that went through a genetic bottleneck (Luikart & Corn- uet 1998). These results supported the hypothesis that urban and rural house sparrow populations act as inde- pendent demographic entities (De Laet & Summers- Smith 2007). Building on these results, we here estimate individual levels of pairwise relatedness to quantify genetic structuring at two hierarchically ordered spatial scales, i.e. a local scale (peripheral vs. central popula- tions) and a regional scale (urban vs. rural populations), and address the following research questions: (i) Do observed distributions of relatedness coefficients differ from simulated ones under the assumption of pan- mixia?; (ii) Are average levels of relatedness lower among individuals from different locations than among individuals from the same location?; (iii) Do mixed dis- tributions of various kinship categories differ between locations?; and (iv) Does genetic distance correlate with geographical distance?

## Material and methods

### Study site and species

Data were collected in the greater Ghent area (156 km2; 237 000 inhabitants, of which 82 584 reside within the

7.64 km2 city centre) and in an adjacent rural area near

the village of Zomergem (38.8 km2; 8150 inhabitants) *ca.* 12 km northwest of Ghent (Fig. 1). Ratios of built-up area to total area of grid cells (each GIS grid cell mea- suring 90 000 m2 on the ground; ARCGIS v9.2) in Zomer- gem and Ghent were 0–0.10 (rural) and >0.10 (urban), respectively (Vangestel *et al.* 2010). To study small-scale variation in patterns of kinship, rural and urban areas (regional scale) were subdivided in a ‘central’ and ‘peripheral’ zone (local scale) each (Fig. 1). In the urban area, the central zone comprised the area with the high- est ratio of built-up area to total area of grid cells

Upon capture, standard morphological measurements were taken, and a small sample of body feathers for DNA analysis was collected (see Vangestel *et al.* 2010 for details).

Fig. 1 Central (filled symbols) and peripheral (open symbols) study plots within an urban (circles) and rural (triangles) area near the city of Ghent, Belgium. The inner contour encom- passes the city centre of Ghent, and the outer contour encom- passes the surrounding municipalities.

(>0.30; compared to 0.11–0.30 for the peripheral zone), which corresponded to areas characterized by ‘a high proportion of houses, pavements and commercial build- ings, low percentage of vegetation and high population densities’ (central zone) and ‘of predominantly houses with small ⁄ large gardens containing an abundant vege- tation and lower population densities’ (peripheral zone), respectively (classification following Heij 1985). The central and peripheral zones in the rural area were of comparable size and spatial configuration as the urban zones, yet did not differ in ratio of built-up area (both 0–0.10; Vangestel *et al.* 2010). Between 2003 and 2009, a total of 690 adult house sparrows from four ‘central’ and nine ‘peripheral’ populations in both urban and rural areas were captured with mist-nets (Fig. 1). Populations were sampled over variable time spans (2.7 ± 1.3 years). However, as the length of the sampling period was not correlated with average (*r*p = )0.02, *P* = 0.93) or variation (*r*p = 0.08, *P* = 0.69) in pairwise relatedness, genotypes were pooled over years.

### DNA extraction, PCR and genotyping

We applied a Chelex resin-based method (InstaGene Matrix; Bio-Rad) (Walsh *et al.* 1991) to extract genomic DNA from ten body feathers per individual. Sixteen mi- crosatellite markers (both traditional ‘anonymous’ mi- crosatellites and markers based on expressed sequence tags) were selected based on their polymorphism and stutter profile (see Vangestel *et al.* (2011) for an exhaus- tive description of each marker). Polymerase chain reac- tions were organized in four multiplex sets, and compatibility between primer pairs was checked with AutoDimer (Vallone & Butler 2004). The first multiplex reaction contained *Pdo*l*1* (Neumann & Wetton 1996), *Pdo32*, *Pdo47* (D.A. Dawson *et al.* in preparation) and *TG04-012* (Dawson *et al.* 2010); the second one con- tained *Pdo*l*3* (Neumann & Wetton 1996), *Pdo*l*5* (Griffith *et al.* 1999), *TG13-017* and *TG07-022* (Dawson *et al.* 2010); the third multiplex reaction contained *Pdo10* (Griffith *et al.* 2007), *Pdo16, Pdo19, Pdo22* (D.A. Dawson *et al.* in preparation) and *TG01-040* (Dawson *et al.* 2010); the last set consisted of *Pdo9* (Griffith *et al.* 2007), *TG01-148* and *TG22-001* (Dawson *et al.* 2010). PCR were performed on a 2720 Thermal Cycler (Applied Biosys- tems) in 9 lL volumes and contained approximately 3 lL genomic DNA, 3 lL QIAGEN Multiplex PCR Mas- termix (Qiagen) and 3 lL primermix [concentrations were 0.1 lM (*Pdo*l*1*), 0.12 lM (*TG01-148*), 0.16 lM

(*Pdo10, Pdo19, Pdo22, Pdo32 and TG04-012*) and 0.2 lM (*Pdo*l*3*, *Pdo*l*5, Pdo9, Pdo16, Pdo47, TG01-040, TG07-022,*

*TG13-017 and TG22-001*)]. The PCR profile included an initial denaturation step of 15 min at 95 °C, followed by

35 cycles of 30 s at 94 °C, 90 s at 57 °C and 60 s at 72 °C, followed by an additional elongation step of 30 min at 60 °C and an indefinite hold at 4 °C. Prior to genotyping, samples were quantified using a ND1000 spectrometer (Nanodrop Technologies) and adjusted to a standard concentration of 10 ng ⁄ lL. Negative controls during extraction and PCR were included to rule out contamination of reagents. PCR products were visual- ized on an ABI3730 Genetic Analyzer (Applied Biosys- tems), an internal LIZ-600 size standard was applied to determine allele size, and fragments were scored using the software package GENEMAPPER v4.0. All loci under study were autosomal as inferred from the chromo- somal location of their homologues on the genome of the zebra finch, *Taeniopygia guttata* (C. Vangestel *et al.* unpublished data). Data have been deposited in the Dryad data repository (doi: 10.5061/dryad.3q0d3rb4).

We used MICRO-CHECKER software (Van Oosterhout *et al.* 2004) to identify scoring errors that could be attributed to stuttering, differential amplification of size-variant alleles causing large allele dropout, or the presence of null alleles. All microsatellite loci were checked for Hardy–Weinberg and linkage equilibrium with GENEPOP version 4.0 (Raymond & Rousset 1995; Rousset 2008), and a family-wise error rate of 0.05 was obtained by applying sequential Bonferroni correction (Weir 1990).

### Spatial kinship analysis

Kinship analyses can be assigned to two types of mod-

0.6

0.5

0.4

0.3

**rQG**

0.2

0.1

0.0

–0.1

**PO FS HS UN**

- - - -

- - - -

- - - -

- - - -

els: those that estimate relatedness between individuals as a continuous measure of genome-wide identity by descent and those that assign dyads to discrete relation- ship categories (Weir *et al.* 2006). To date, there is no single best relatedness estimator that outperforms all others. Rather, their performance appears context spe- cific and to rely on the underlying true genetic popula- tion structure (Van de Casteele *et al.* 2001; Wang 2011). For example, while likelihood estimators are less biased than moment estimators when either a large number of individuals or a large set of highly polymorphic mark- ers are used, the former suffer from a larger error of inference (Wang 2011). Ritland (1996) and Lynch & Rit- land (1999) estimators are presumed to be more accu- rate than Queller & Goodnight (1989), Li *et al.* (1993) and Wang (2002) estimators when populations contain mainly unrelated dyads; however, the opposite is often true for samples of highly related dyads (Wang 2011). Yet, as several of the aforementioned factors may inter- act in diverse and complex ways, Van de Casteele *et al.* (2001) recommended the use of simulated data to evalu- ate the appropriateness of each estimator. We per- formed Monte Carlo simulations using COANCESTRY Version 1.0 (Wang 2011) to calculate correlation coeffi- cients between different estimates and true simulated relatedness values (using observed allele frequencies). Based on these simulations, we selected the Queller- Goodnight moment estimator (*r*QG, Queller & Good- night 1989) that yielded a strong correlation between true and estimated values (*r* = 0.82, *P* < 0.001; Fig. 2) within the range reported in previous studies (Van de Casteele *et al.* 2001 and references therein). The asym- metric Queller & Goodnight (1989) index between indi- viduals *x* and *y* (using the former as the reference

individual) was calculated as ^*r* ¼ R*l* R*k* ð*py* -*pkl* Þ ; where -*p*

*QG* R R ð*p* -*p* Þ *kl*

Fig. 2 Association between relatedness estimates of simulated kinship categories (mean value = black dot, standard error = - error bars) and expected theoretical relatedness values (dashed lines); 100 dyads were simulated for each of four relationship categories (U, unrelated; HS, half-siblings; FS, full-siblings; PO, parent–offspring).

allele *k* in the reference individual (1 or 0.5 for homo- zygotes and heterozygotes, respectively), and py equals the frequency of allele *k* in the individual compared (1,

0.5 or 0 for heterozygotes, homozygotes or the absence of allele *k*, respectively). When summed over both indi- viduals, this index transforms into a symmetrical one (Queller & Goodnight 1989).

For each population, we calculated the frequency dis- tribution of relatedness coefficients between all *n*(*n* ) 1) ⁄ 2 dyads using GENALEX version 6.41 (Peakall

& Smouse 2006), while confidence intervals for mean within-population relatedness were estimated via boot- strapping. Likewise, we generated a null distribution of relatedness coefficients among unrelated individuals by randomly permuting genotypes over all populations for 999 times. As corresponding upper and lower 95% intervals represent the range of relatedness values to be expected under random mating, mean values of empiri- cal distributions above the simulated 95% upper limit are considered indicative of non-panmictic conditions. Next, we estimated relatedness coefficients between all possible dyads within central ⁄ peripheral zones or urban ⁄ rural areas, respectively. Mean coefficients of relatedness that exceed the 95% upper limit thereby provide evidence for an increase in relatedness because of the presence of small-scale genetic structure at the particular spatial scale involved.

Given local population structuring, we further

*l k x kl*

equals the population allele frequency of allele

*k* at

expected individuals from common environments (i.e.

locus *l* over all individuals as if they were a single pop- ulation (no substantial effect on relatedness estimates was found when analyses were performed in the urban or rural area separately), *px* equals the frequency of

central ⁄ central) to be more strongly related than dyads containing individuals from mixed environments (i.e. central ⁄ peripheral). We therefore compared differences in mean relatedness between both dyad types and

performed a bootstrap test of 1000 repetitions in COAN- CESTRY Version 1.0 (Wang 2011) to assess the level of sta- tistical significance. To test for ‘urban centre’ effects, we compared mean levels of relatedness in urban and rural centres and assessed whether there was regional hetero- geneity in differences between dyads of common (cen- tral ⁄ central) versus mixed (central ⁄ peripheral) origin. Finally, individuals of urban and rural areas were pooled, and mean levels of relatedness of both zones were compared.

We compared the kin structure of common and mixed dyads by quantifying the relative contribution of unrelated (U), half-sibling (HS), full-sibling (FS) or par- ent–offspring (PO) types of relatedness. Based on observed allele frequencies within, respectively, the urban and rural area, we simulated 1000 pairs of each type with COANCESTRY Version 1.0 (Wang 2011) and esti- mated the proportion of each type to the empirical dis- tribution using a finite Bayesian mixture analysis in which the observed distribution was modelled as a mix of normal distributions (all simulated distributions reached normality). Proportions were drawn from a uniform Dirichlet distribution generating 100 000 pos- terior samples after discarding the initial 50 000 (burn- in) samples. Next, proportions of HS, FS and PO types were pooled into a single ‘close kin’ type, and differ- ences in the proportion of ‘close kin’ between common and mixed dyads were considered significant if the 95% credibility interval did not contain zero. All analy- ses were performed in WINBUGS Version 1.4 (Lunn *et al.* 2000). For the Bayesian analysis, relatedness coeffi- cients were estimated within each area separately.

Finally, we assessed patterns of fine-scale genetic structure by quantifying the association between matri- ces of pairwise genetic ⁄ spatial distances (Smouse & Peakall 1999; Peakall *et al.* 2003; Vekemans & Hardy 2004) through spatial auto-correlation analysis in GENA- LEX version 6.41 (Peakall & Smouse 2006). Under a restricted dispersal model, auto-correlograms are pre- dicted to yield positive correlations at short spatial dis- tances, followed by a gradual decrease to zero with increasing geographical distance and a subsequent ran- dom fluctuation of positive and negative values of the correlation coefficient (Smouse & Peakall 1999). The first *x*-intercept is thereby regarded to estimate the extent of non-random genetic structure, that is, to reflect the point at which random stochastic drift replaces gene flow as the key determinant of genetic structure. As this intercept may depend on the true scale of genetic struc- ture, the chosen distance class size and the sample size per distance class (Peakall *et al.* 2003), we performed a second auto-correlation analysis in which we plotted pairwise genetic distances against increasing inclusive distance classes. Here, the distance class at which the

auto-correlation coefficient no longer remains significant (using 999 bootstraps) approximates the true extent of identifiable genetic structure between groups of individ- uals (Peakall *et al.* 2003).

## Results

### Microsatellite data

Loci were highly polymorphic in all populations, and there was no evidence for linkage disequilibrium between any pair of loci. All locus-by-population com- binations were in Hardy–Weinberg equilibrium apart from *Pdo47*, which deviated from Hardy–Weinberg equilibrium in three populations, *Pdo*l*5* in two popula- tions and *Pdo9, Pdo32*, *TG01.148*, *TG07.022* and *TG13- 017* in one population each. There was no evidence for scoring errors because of large allele dropout or stutter, and when rerunning analyses without all or some of these markers, only inclusion or exclusion of *Pdo*l*5* and *Pdo9* had a minor effect on relatedness estimates in population OW, while all other marker-by-population combinations remained unaffected (Fig. 3). As removal of either both loci or all genetic data from population OW did not affect any conclusion of our study, results presented in this paper are based on information from all populations and loci.

### Hierarchical variation in relatedness

Some, but not all, populations showed mean relatedness coefficients outside the limits of the null distribution generated by randomly permuting genotypes over all populations. Pooling populations in either a central or peripheral zone resulted in larger than expected pair- wise relatedness among individuals from the urban cen- tre only. When pooling populations in urban and rural areas, only the former showed evidence of a significant higher degree of relatedness than expected under com- plete panmixia (Fig. 3). While these genetic signatures were not consistent at the smallest spatial scale (popula- tion level), patterns at local and regional scales showed that individuals from the urban area were more strongly related, especially so in the central urban zone. Such pattern can be expected under the assumption of (semi)isolation and lack of ample migration to counter- balance the effects of non-random mating.

### Relatedness in mixed and common dyads

Notwithstanding the low overall level of kinship in our study area, levels of relatedness significantly differed between dyads consisting of individuals from a mixed (central–peripheral) origin compared to those from the

0.25

0.20

0.15

0.10

**rQG**

0.05

0.00

**(a)**

# --------------------------

0.04

0.03

0.02

0.01

**rQG**

0.00

**(b)**

# - -

- -

0.006

0.004

0.002

0.000

–0.002

**rQG**

–0.004

–0.006

–0.008

**(c)**

# - -

–0.05 ------------

–0.10

–0.15

# --------------

–0.01

–0.02

–0.03

- -

-

-

–0.010

–0.012 - -

–0.014

Fig. 3 Mean relatedness values and 95% confidence limits are plotted. Grey bars represent 95% upper and lower expected related- ness values under the assumption of panmixia across all populations. These values were contrasted with relatedness estimates (a) within each population, (b) central and peripheral zones within each area and (c) urban versus rural area, and mean values outside these simulated 95% confidence intervals represent non-panmictic conditions. Population OW with significantly higher *r*QG most likely comprises an outlier because of the performance of loci *Pdo*l*5* and *Pdo9*. Notice the difference in scale between (a), (b) and (c).

central zone only, both in urban and rural populations (all *P* < 0.01). While such local ‘centre’ effect was expected to be strongest in the urban area, this was not supported by our data as the two-factor (origin \* area) interaction was not significant (*P* > 0.05). The higher

**(Regional)**

**\***

level of relatedness among individuals from the urban compared to the rural centre (*P* < 0.01) was consistent with the overall higher relatedness in the entire urban area compared to the rural one (*P* < 0.01) (Fig. 4).

### Bayesian mixture analysis of kinship distributions

The distribution of pairwise relatedness coefficients

0.020 **\***

0.015

0.010

0.005

**rQG**

**(Local )**

**\***

**(Regional)**

**\***

strongly overlapped with a simulated distribution of unrelated individuals (Fig. 5). The estimated percentage (with 95% credibility interval) of close kin equalled 10.8% (9.1–12.6) in the urban centre, while urban dyads of mixed (central–peripheral) origin contained a signifi- cantly lower percentage of 8.9% (8.2–9.6) close kin. In contrast, differences in close kin between both dyad

0.000

–0.005

–0.010

0.25 U

HS FS PO

0.20 Observed

0.15

**Frequency**

Fig. 4 Genetic population structure at a local (central vs. peripheral) and a regional (urban vs. rural) scale. Mean levels of relatedness and error bars (SE) are plotted for several dyad groups: ‘within UC’ and ‘within RC’ refer to common dyads comprising individuals from *u*rban *c*entre or *r*ural *c*entre,

respectively; ‘between UC–UP’ and ‘between RC–RP’ refer to mixed dyads comprising one individual from *u*rban *c*entre or

0.10

0.05

0.00

–0.6 –0.4 –0.2 0.0 0.2 0.4 0.6 0.8

**r**

*r*ural *c*entre, respectively, and one individual from *U*rban *P*eriphery or *R*ural *P*eriphery, respectively; ‘within U’ and ‘within R’ refer to all possible dyads within the *U*rban or *R*ural zone, respectively. The scale at which *r*QG varies resembles that of Fig. 3b.

**QG**

Fig. 5 Observed (solid line) versus expected (bars) patterns of relatedness in central urban house sparrows. U, unrelated (true *r* = 0); HS, half-siblings (true *r* = 0.125); FS, full-siblings (true r = 0.5); PO, parent–offspring (true *r* = 0.5).

14

\*

12

10

**Proportion close kin (%)**

8

6

4

2

0

Fig. 6 Posterior mean proportions of close kin obtained from a finite Bayesian mixture analysis with error bars representing 95% credibility intervals. Significant differences between dyad types in the urban or rural area are indicated by ‘\*’; ‘within UC’ and ‘within RC’ refer to common dyads comprising indi- viduals from *u*rban *c*entre or *r*ural *c*entre, respectively; ‘between UC–UP’ and ‘between RC–RP’ refer to mixed dyads comprising one individual from *u*rban *c*entre or *r*ural *c*entre, respectively, and one individual from *U*rban *P*eriphery or *R*ural *P*eriphery, respectively.

types did not reach statistical significance in rural popu- lations; however, estimated proportions were signifi- cantly lower than in urban populations [central–central dyads: 5.7% (4.4-7.0); central–peripheral dyads: 5.5%

(4.9–6.1)] (Fig 6).

### Correlations between spatial and genetic distances

Spatial auto-correlograms provided analogous evidence for small-scale restricted gene flow, i.e. significant,

positive genetic correlations among geographically adja- cent populations and a subsequent decline with increas- ing distance (Fig. 7a). Auto-correlograms intersected the *x*-axis between 1500 and 2000 m for urban individu- als and between 500 and 1000 m for rural ones. At lar- ger distances, patterns became chaotic as r tended to fluctuate around zero, hence confirming that the spatial scale of our study was appropriate. Based on their pro- portion of shared genes, individuals became genetically clustered at distances below 1.5–2 km (urban) or 0.5– 1 km (rural), while samples beyond this threshold were considered as genetically independent. Likewise, multi- ple distance class plots indicated non-randomly distrib- uted genotypes at small geographical distances and a gradual decrease towards zero with increasing distance size class (Fig. 7b). Here, auto-correlation coefficients reached zero values at 3.5 km (urban) and 5 km (rural), respectively.

Discussion

Individual-level analysis of relatedness coefficients inferred from microsatellite genotypes revealed small- scale genetic population structure in urban and rural house sparrows that was indicative of highly restricted gene flow beyond distances of 3500 m. Average related- ness was higher among urban individuals, suggesting lower dispersal in more strongly built-up habitats. Comparison of observed levels of relatedness with sim- ulated distributions of known kinship values further showed that central urban sparrows had the highest proportion of closely related conspecifics in their imme- diate neighbourhood. Spatial auto-correlograms sup- ported this small-scale genetic structure and indicated stronger effects of genetic drift and (or) more restricted dispersal in urban populations.

0.12

0.10

0.08

**Autocorrelation (r)**

0.06

0.04

0.02

0.00

–0.02

–0.04

–0.06

**(a)**

**(b)**

0.12

0.10

0.08

0.06

0.04

0.02

0.00

–0.02

**Urban zone Rural zone**

**Distance class (m)**

Fig. 7 Spatial auto-correlation analysis in urban and rural area. (a) Single correlogram depicting the auto-correlation as a function of distance. (b) Multiple distance class plot depicting the effect of different distance size classes on the extent of genetic autocorrelation. Error bars were estimated using 999 bootstraps.

Given the close proximity of neighbouring popula- tions, it is unlikely that low dispersal was the main dri- ver underlying the collapse of the central urban populations. Yet, it may have enhanced negative demo- graphic trends induced by other local drivers, such as increased juvenile mortality owing to lack of insects (Peach *et al.* 2008). The small proportion of house spar- rows that disperse over distances of several kilometres (Anderson 2006) may suffice to homogenize gene pools and decrease levels of relatedness, but be insufficient to reverse negative demographic trends (Lowe & Allen- dorf 2010). Through stochastic simulations, Peach *et al.* (2008) showed that breeding success was insufficient to maintain sustainable house sparrow populations during two of three breeding seasons in central UK, and repro- ductive failure because of low invertebrate abundance was believed to constitute the most likely demographic mechanism for the decline. While successful immigrants from neighbouring populations could have countered (or even reversed) these negative effects, sparrow num- bers continued to decreased by more than 16% (Vincent 2005). It hence appears that small shifts in dispersal rates may substantially affect demographic patterns in highly sedentary species such as the house sparrow.

Population-level genetic analysis based on Wright’s *F*-statistics (or derivatives thereof) earlier provided only weak evidence for regional structuring of house spar- row populations in the same study area (C. Vangestel *et al.* unpublished). While results of the current study appear to contradict such panmictic pattern, *F*-related statistics largely reflect evolutionary outcomes and assume equilibrium conditions that are often not met in natural populations (Whitlock & McCauley 1999). In contrast, patterns of kinship reflect ongoing genetic pro- cesses (Peakall *et al.* 2003) although it is currently unclear whether, and to what extent, distributions of relatedness are also affected by genetic disequilibria. Earlier, Liker *et al.* (2009) did not reveal significant vari- ation in kinship within and between winter feeding flocks of house sparrows in NW Hungary, nor were patterns related to geographical distance. Results from our study provide different, non-exclusive explanations for this apparent lack of small-scale genetic structure. First, Liker and colleagues studied house sparrows in a moderately urbanized area. This may have resulted in a less powerful spatial genetic analysis as we showed that non-random patterns of genetic affinity were weaker in such areas compared to highly urbanized city centres. Second, as mentioned by the authors, despite the fact that their sampling design was most appropriate to relate feeding aggregations to genetic relatedness, the spatial scale of their study (most distant populations separated by ca. 1200 m, majority of the populations within 500 m range) may have fallen below the threshold

for detecting genetic structuring. Third, genetic differen- tiation among populations and feeding flocks may have increased differences in estimates of relatedness between ‘common’ and ‘mixed’ dyads. However, as allele frequencies did not strongly vary among most of our populations, we consider this explanation less likely.

Use of auto-correlation analysis revealed that proxi- mate sparrow pairs were genetically more strongly cor- related than randomly chosen pairs, while this was not true for more distant ones. Average estimates of relat- edness were higher in urban populations, suggesting stronger local population structure in highly built-up areas. Strong genetic drift and ⁄ or weak gene flow may cause substantial genetic correlation between house sparrows within these areas. Positive genetic structuring was detectable up to 1.5–2 km in urban populations, while rural sparrows showed genetic independence at 0.5–1 km. However, as *x*-intercepts of auto-correlograms (representing shifts from gene flow to genetic drift as main driver of population differentiation) are believed to depend on the sampling scheme applied (Vekemans

& Hardy 2004), it is recommendable to use multiple dis- tance size plots to detect the true extent of genetic struc- ture (Peakall *et al.* 2003), even if sampling schemes are considered comparable as in our study. Based on this second analysis, distances over which gene flow was still detectable tended to be larger for rural (5 km) than for urban (3.5 km) populations (no statistical test applied). Because genetic drift was expected to be stron- ger in urban populations (characterized by lower census population sizes; C. Vangestel, unpublished), this may have caused the observed difference between both cor- relograms, possibly in synergy with low dispersal rates. In the urban area, but not in the rural one, the propor- tion of close kin (modelled as a dichotomous variable) differed between common and mixed dyads, suggesting that ‘centre’ effects were stronger in the former. How- ever, no such pattern was apparent for mean estimates of relatedness (if modelled as a continuous variable) as ‘centre’ effects were present in both the urban and rural area and effect sizes did not differ. As overall level of relatedness was lowest in the rural area, shifting from a continuous to a dichotomous variable may have reduced the level of variation in kinship in this area, thereby causing the apparent discordance between both analyses.

With the exception of central urban dyads, propor- tions of close kin were smaller than those reported in Liker *et al.* (2009) where all flock assemblages except one (0.056) had mean proportion of close kin ranging from 0.11 to 0.19. Such difference may have been real or caused by differences in the study design (i.e. scale over which populations were sampled) or analytical

methods applied. As opposed to Liker *et al.* (2009) who assigned dyads to particular classes of kinship based on likelihood algorithms (Milligan 2003), we applied a Bayesian mixture approach that exploited the full range of data and is considered to be less prone to large errors of inference associated with indi- vidual kinship estimates (Lynch & Ritland 1999; Swei- gart *et al.* 1999). When applying a likelihood method comparable to Liker *et al.* (2009), estimated proportions of close kin increased to 0.15–0.21 (data not shown), however, without affecting any of the conclusions of our study.

Integrating individual-based genetic estimates with ecological and demographic data can help to address key ecological questions. For example, as described ear- lier, it may allow to assess to what extent heterogeneity in reproductive success along urban–rural gradients coincides with differences in spatial (relatedness) struc- ture, which, in turn, may affect breeding output. Also, while often assumed in population genetic models, dis- persal may not be entirely random but depend on indi- vidual behavioural types (Dingemanse *et al.* 2003). If so, assays of behaviour of known genotypes may help to unravel to what extent interactions between individual personality types, tendencies to disperse and non-ran- dom genetic samples from source populations can drive small-scale population structure. Finally, predators such as sparrowhawks (*Accipiter nisus*) have only recently invaded city centres after a long period of absence (Bell *et al.* 2010). Such extended periods of predator release may result in natural selection acting against anti-pred- ator strategies (Blumstein & Daniel 2005), thereby ham- pering rapid behavioural adjustments when predators reappear (Steadman 2006). Estimation of predation rates on known genotypes during post-fledging dispersal may allow to assess whether, and to what extent, pre- dation risks can drive spatial population structure. Yet, individual-based analysis of relatedness is still not widely applied to quantify small-scale population struc- ture in natural populations. Positive examples include the studies by Sweigart *et al.* (1999), Ruzzante *et al.* (2001), and Van de Casteele & Matthysen (2006) on relatedness-based estimates of genetic population struc- ture in the wildflower *Mimulus guttatus*, brown trout (*Salmo trutta*) and great tits (*Parus major*), respectively. In the latter case, clustering algorithms based on Hardy–Weinberg and linkage disequilibria earlier failed to reveal small-scale population structure (Pritchard *et al.* 2000). Likewise, individual-based genetic analyses showed substantial among-patch movement in pikas (*Ochotona princeps*), while mark–recapture analysis did not pick up such signal (Peacock 1997; see Spong & Creel 2001 for an example where both methods pro- vided consistent results).

Despite these positive results, inferring small-scale population structure from relatedness values needs to be done with caution as distributions of coefficients of relatedness are not exclusively shaped by the level of connectivity among adjacent populations. For instance, similarity in the direction of dispersal among siblings as a result of parental control (Matthysen *et al.* 2005, 2010; see Massot & Clobert 2000 for other examples) can affect kinship assemblages, and temporal or spatial variation in family behaviour may induce fine-scale kin- ship structure. While still hypothetical, this may in par- ticular affect highly sedentary species such as the house sparrow that show a strong preference to reselect previ- ous occupied nesting locations (Vincent 2005). Apart from kin-related dispersal, house sparrows also tend to show fewer aggressive interactions during foraging against kin (To´ th *et al.* 2009). Such behaviour, too, may result in differential dispersal between family groups, for example, through differences in access to high-qual- ity areas for offspring from high and low dominant sta- tus parents. Applying individual-level genetic analysis as a complementary approach to population-level study, such as shown in this paper, will ultimately allow to test whether, and to what extent, human- induced shifts in habitat quality can affect fine-scale patterns of relatedness through complex, behavioural interactions.

## Acknowledgements

We wish to thank H. Matheve, T. Cammaer and C. Nuyens for field assistance, and A. Krupa and G. Horsburgh for labo- ratory assistance. We thank T. Burke for supporting this work and helpful discussions, E. Matthysen, F. Hendrickx and two anonymous reviewers for insightful comments on an earlier draft. Genotyping was performed at the NERC Biomolecular Analysis Facility funded by the Natural Environment Research Council, UK. Fieldwork and laboratory analyses were financially supported by research projects 01J01808 (Ghent University) and G.0149.09 (Research Foundation- Flan- ders).

## References

Anderson TR (2006) *Biology of the Ubiquitous House Sparrow*. Oxford Univ. Press, Oxford.

Bell CP, Baker SW, Parkes NG, Brooke MD, Chamberlain DE (2010) The role of the Eurasian sparrowhawk (Accipiter nisus) in the decline of the house sparrow (*Passer domesticus*) in Britain. *Auk*, 127, 411–420.

Blouin MS (2003) DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology and Evolution*, 18, 503–511.

Blumstein DT, Daniel JC (2005) The loss of anti-predator behaviour following isolation on islands. *Proceedings of the Royal Society of London*, 272, 1663–1668.

Chace JF, Walsh JJ (2006) Urban effects on native avifauna: a review. *Landscape and Urban Planning*, 74, 46–69.

Chamberlain DE, Toms MP, Cleary-McHarg R, Banks AN (2007) House sparrow (*Passer domesticus*) habitat use in urbanized landscapes. *Journal of Ornithology*, 148, 453–462.

Chiron F, Julliard R (2007) Responses of songbirds to magpie reduction in an urban habitat. *Journal of Wildlife Management*, 71, 2624–2631.

Danchin E, Giraldeau L-A, Ce´zilly F (2008) *Behavioural Ecology*. Oxford University Press, Oxford.

Dawson DA, Horsburgh GJ, Kupper C *et al.* (2010) New

methods to identify conserved microsatellite loci and develop primer sets of high cross-species utility—as demonstrated for birds. *Molecular Ecology Resources*, 10, 475–494.

De Laet J, Summers-Smith JD (2007) The status of the urban House Sparrow *Passer domesticus* in North-Western Europe: a review. *Journal of Ornithology*, 148, S275–S278.

Dingemanse NJ, Both C, van Noordwijk AJ, Rutten AL, Drent PJ (2003) Natal dispersal and personalities in great tits (*Parus major*). *Proceedings of the Royal Society of London*, 270, 741–747. Griffith SC, Stewart IRK, Dawson DA, Owens IPF, Burke T (1999) Contrasting levels of extra-pair paternity in mainland and island populations of the house sparrow (*Passer domesticus*): is there an ‘island effect’? *Biological Journal of the*

*Linnean Society*, 68, 303–316.

Griffith SC, Dawson DA, Jensen H *et al.* (2007) Fourteen polymorphic microsatellite loci characterized in the house sparrow *Passer domesticus* (Passeridae, Aves). *Molecular Ecology Notes*, 7, 333–336.

Heij CJ (1985) *Comparative ecology of the house sparrow Passer domesticus in rural, suburban and urban situations*. PhD Thesis, Vrije Universiteit Amsterdam, Netherlands.

Heij CJ, Moeliker CW (1990) Population dynamics of dutch House Sparrows in urban, suburban and rural habitats. In: *Granivorous Birds in the Agricultural Landscape* (eds Pinowski J, Summers-Smith JD), pp. 59–85. Polish Scientifi Publishers, Warszawa.

Hole DG, Whittingham MJ, Bradbury RB *et al.* (2002) Widespread local house sparrow extinctions—agricultural intensification is blamed for the plummeting populations of these birds. *Nature*, 418, 931–932.

Johnson JB, Peat SM, Adams BJ (2009) Where’s the ecology in molecular ecology? *Oikos*, 118, 1601–1609.

Kekkonen J, Seppa P, Hanski I *et al.* (2010) Low genetic differentiation in a sedentary bird: house sparrow population genetics in a contiguous landscape. *Heredity*, 106, 183–190.

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. *Trends in Ecology and Evolution*, 17, 230–241.

Lendvai AZ, Giraudeau M, Chastel O (2007) Reproduction and modulation of the stress response: an experimental test in the house sparrow. *Proceedings of the Royal Society B-Biological Sciences*, 274, 391–397.

Li CC, Weeks DE, Chakravarti A (1993) Similarity of DNA fingerprints due to chance and relatedness. *Human Heredity*, 43, 45–52.

Liker A, Bokony V, Kulcsar A *et al.* (2009) Genetic relatedness in wintering groups of house sparrows (*Passer domesticus*). *Molecular Ecology*, 18, 4696–4706.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology*, 19, 3038–3051.

Luikart G, Cornuet J-M (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology*, 12, 228–237.

Lunn DJ, Thomas A, Best N, Spiegelhalter D (2000) WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, 10, 325–337.

Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. *Genetics*, 152, 1753–1766.

Lynch M, Walsh B (1998) *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, Massachusetts.

Marzluff JM, Bowman R, Donnelly R (2001) *Avian Ecology and Conservation in an Urbanizing World*. Kluwer Academic Publishers, Norwell, Massachusetts.

Massot M, Clobert J (2000) Processes at the origin of similarities in dispersal behaviour among siblings. *Journal of Evolutionary Biology*, 13, 707–719.

Matthysen E, Van de Casteele T, Adriaensen F (2005) Do sibling tits (*Parus major, P. caeruleus*) disperse over similar distances and in similar directions? *Oecologia*, 143, 301–307.

Matthysen E, Van Overveld T, Van de Casteele T, Adriaensen F (2010) Family movements before independence influence natal dispersal in a territorial songbird. *Oecologia*, 162, 591– 597.

Milligan BG (2003) Maximum-likelihood estimation of relatedness. *Genetics*, 163, 1153–1167.

Nakagawa S, Ockendon N, Gillespie DOS, Hatchwell BJ, Burke T (2007) Assessing the function of house sparrows’ bib size using a flexible meta-analysis method. *Behavioral Ecology*, 18, 831–840.

Neumann K, Wetton JH (1996) Highly polymorphic microsatellites in the house sparrow *Passer domesticus*. *Molecular Ecology*, 5, 307–309.

Peach WJ, Vincent KE, Fowler JA, Grice PV (2008) Reproductive success of house sparrows along an urban gradient. *Animal Conservation*, 11, 493–503.

Peacock MM (1997) Determining natal dispersal patterns in a population of North American pikas (*Ochotona princeps*) using direct mark-resight and indirect genetic methods. *Behavioral Ecology*, 8, 340–350.

Peakall R, Smouse PE (2006) Genalex 6: genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288–295.

Peakall R, Ruibal M, Lindenmayer DB (2003) Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, *Rattus fuscipes*. *Evolution*, 57, 1182– 1195.

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

Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers. *Evolution*, 43, 258–275.

Raymond M, Rousset F (1995) Genepop (version- 1.2)—population-genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86, 248–249.

Ritland K (1996) A marker-based method for inferences about quantitative inheritance in natural populations. *Evolution*, 50, 1062–1073.

Rodewald AD, Shustack DP (2008) Urban flight: understanding individual and population-level responses of Nearctic- Neotropical migratory birds to urbanization. *Journal of Animal Ecology*, 77, 83–91.

Rodewald AD, Kearns LJ, Shustack DP (2011) Anthropogenic resource subsidies decouple predator-prey relationships. *Ecological Application*, 21, 936–943.

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

Ruzzante DE, Hansen MM, Meldrup D (2001) Distribution of individual inbreeding coefficients, relatedness and influence of stocking on native anadromous brown trout (*Salmo trutta*) population structure. *Molecular Ecology*, 10, 2107–2128.

Shaw LM, Chamberlain D, Evans M (2008) The House Sparrow *Passer domesticus* in urban areas: reviewing a possible link between post-decline distribution and human socioeconomic status. *Journal of Ornithology*, 149, 293–299.

Shochat E, Warren PS, Faeth SH, Mcintyre NE, Hope D (2006) From patterns to emerging processes in mechanistic urban ecology. *Trends in Ecology & Evolution*, 21, 186–191.

Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity*, 82, 561–573.

Spong G, Creel S (2001) Deriving dispersal distances from genetic data. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 268, 2571–2574.

Steadman DW (2006) *Extinction and Biogeography of Tropical Pacific Birds*. University of Chicago Press, Chicago, Illinois.

Sweigart A, Karoly K, Jones A, Willis JH (1999) The distribution of individual inbreeding coefficients and pairwise relatedness in a population of *Mimulus guttatus*. *Heredity*, 83, 625–632.

To´ th Z, Bokony V, Lendvai AZ *et al.* (2009) Effects of relatedness on social-foraging tactic use in house sparrows. *Animal Behaviour*, 77, 337–342.

Vallone PM, Butler JM (2004) Autodimer: a screening tool for primer-dimer and hairpin structures. *BioTechniques*, 37, 226– 231.

Van de Casteele T, Matthysen E (2006) Natal dispersal and parental escorting predict relatedness between mates in a passerine bird. *Molecular Ecology*, 15, 2557–2565.

Van de Casteele T, Calbusera P, Matthysen E (2001) A comparison of microsatellite-based pairwise relatedness estimators. *Molecular Ecology*, 10, 1539–1549.

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

Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, 13, 921–935.

Vincent K (2005) *Investigating the causes of the decline of the*

*urban house sparrow Passer domesticus population in Britain*. PhD Thesis, De Montfort University, UK.

Walsh PS, Metzger DA, Higuchi R (1991) Chelex-100 as a medium for simple extraction of dna for pcr-based typing from forensic material. *BioTechniques*, 10, 506–513.

Wang J (2002) An estimator for pairwise relatedness using molecular markers. *Genetics*, 160, 1203–1215.

Wang J (2011) COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources*, 11, 141–145.

Weidinger K (2009) Nest predators of woodland open-nesting songbirds in central Europe. *Ibis*, 151, 352–360.

Weir B (1990) *Genetic Data Analysis: Methods for Discrete Population Genetic Data*. Sinauer Associates, Sunderland, Massachusetts.

Weir BS, Anderson AD, Hepler AB (2006) Genetic relatedness analysis: modern data and new challenges. *Nature Review Genetics*, 7, 771–780.

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration*: F*st „ 1 ⁄ (4 Nm + 1). *Heredity*, 82, 117–

125.

Zamudio KR, Wieczorek AM (2007) Fine-scale spatial genetic structure and dispersal among spotted salamander (*Ambystoma maculatum*) breeding populations. *Molecular Ecology*, 16, 257–274.

C.V.’s research focuses on patterns of neutral and adaptive genetic variation underlying key ecological and evolutionary processes. J.M. is researcher in ecology and evolutionary biol- ogy, with a particular focus on extinction debt and the mainte- nance of genetic diversity in small populations. In addition, he studies priority effects in population genetics and community ecology. D.A.D.’s research involves the development of micro- satellite markers and their application in studies of reproduc- tive behaviour, population structure and avian genome organisation. V.V. is interested in the application of neutral and adaptive genetic markers to study demographic and evo- lutionary processes in vertebrate and invertebrate species. L.L. is interested in the evolutionary ecology and conservation genetics of fragmented populations.

Vangestel C, Braeckman BP, Matheve H, Lens L (2010)

Constraints on home range behaviour affect nutritional condition in urban house sparrows (*Passer domesticus*).

*Biological Journal of the Linnean Society*, 101, 41–50.

Vangestel C, Mergeay J, Dawson DA, Vandomme V, Lens L (2011) Developmental stability covaries with genome-wide and single-locus heterozygosity in house sparrows. *PLoS ONE*, 6, e21569.