Population fragmentation drives up genetic diversity in signals of individual identity

Calvin Dytham\textsuperscript{1} and Michael D.F. Thom\textsuperscript{2} *

\textsuperscript{1} Department of Biology, University of York, York, YO10 5DD, UK calvin.dytham@york.ac.uk
\textsuperscript{2} School of Biological and Marine Sciences, University of Plymouth, PL14 3JS, UK michael.thom@plymouth.ac.uk Ph: +44 1752 584 473 (* corresponding author)

Submitted to \textit{Oikos}

Running title: Genetic diversity in fragmented populations

Keywords: agent-based model; individual-based model; genetic diversity; individual recognition; mate choice; population fragmentation; sexual selection; dispersal

\textbf{Acknowledgments}: Thanks to Chris Thomas, Michael Bottery, Jacob Davis and Miguel Franco for comments on the draft manuscript.
Abstract

Many species advertise their unique identity to conspecifics using dedicated individuality signals: one familiar example is human faces. But how unique in the global population do these signals need to be? While human faces are highly variable, each person interacts with many fewer individuals than are found in the total population. This raises the question of how evolutionary mechanisms drive up population-wide diversity when selection occurs at such a local level. We use an individual-based model in which individuals broadcast their identity and quality in separate, genetically-coded signals. Mimicking, for example, scent marking mammal species, females in the model assess males using the quality signal, then attempt to relocate the highest quality male using his identity signal. We ask how population fragmentation affects genetic diversity in the individual identity-signalling region under sexual selection, predicting one of two opposing outcomes: (1) divided populations evolve fewer signal variants globally, since repetition of signals is not costly when individuals interact only with local conspecifics, or (2) stochasticity in mutation and selection cause divergence among subpopulations, increasing the global number of signal variants. We show that local selection drives up global genetic diversity substantially in fragmented populations, even with extremely low rates of dispersal. Because new signal variants arise by mutation and then sweep through their subpopulation, a fragmented population has more global signal variation. This result furthers our understanding of how high levels of diversity in individuality signals are maintained.
Introduction

Individual recognition – the ability to identify conspecifics to the level of the individual – appears to be a widespread ability in species from a broad range of taxonomic groups (e.g. humans: (Sheehan and Nachman 2014), wasps: (Sheehan and Tibbetts 2010), mice: (Hurst et al. 2001), lobsters: (Karavanich and Atema 1998), birds: (Medvin et al. 1993)). Across these groups, individuals benefit from being recognized because accurate recognition carries fitness benefits: these include the maintenance of complex social hierarchies (Tibbetts 2002), facilitating mate choice (Aquiloni and Gherardi 2010, Cheetham et al. 2007), ensuring accurate provision of parental care (Jouventin and Aubin 2002) and recognition of neighbouring territory holders (Hurst et al. 2005) or colony mates (Sheehan and Tibbetts 2009). Where benefits to the signaller exist, we expect selection to drive the evolution of individuality signals (Johnstone 1997, Tibbetts and Dale 2007) through negative frequency-dependent selection on rare signal types (Dale et al. 2001). Human faces are a probable example of individuality signalling diversity that has arisen under selection for rarity (Sheehan and Nachman 2014).

While there is evidence that complex social interactions can drive the evolution of diversity in identity signals (Tibbetts and Dale 2007), little is known about how much variability evolves in different systems. There are good reasons to expect that not every individual in a population needs to have a unique identifier. While human faces, for example, appear to offer an almost unlimited number of unique variants, and show little overlap within populations (Sheehan and Nachman 2014), individual recognition may still be beneficial even when there is some overlap among individuals’ identity signals (Dale et al. 2001). Some identification errors might be acceptable, and, even in humans, receivers are often confused.
by similar-looking faces (Tibbetts and Dale 2007). Curiously, Tibbetts and Dale (2007) inadvertently emphasize this point by making just such an error, mislabelling two of the five pictured human faces in their figure legend. So, acceptability of occasional identification errors means that not every individual needs to be globally unique. Second, we might also find shared signals in the population as a consequence of limits to the combinatorial diversity available in the signalling system (e.g. because it is coded by a single locus). Finally, some degree of signal sharing might be expected to evolve because in most cases an individual interacts with only a small proportion of the total population. For instance, there appear to be vastly more human faces than are required for day-to-day human interactions, meaning combinatorial diversity in faces far exceeds what is needed to maintain social processes.

While there are several reasons to expect some degree of signal sharing, we predict a relationship between the number of interacting individuals and the number of signals that evolve. Where dispersal is high, or populations are large, more signal variants will be required to ensure misidentification is rare. Indeed, positive correlations between group size and signal diversity have been reported in bats (Luo et al. 2017) and chickadees (Freeberg 2006), and there is evidence that species-level signal variability is linked with coloniality in swallows (Medvin et al. 1993). There may be a threshold population size beyond which individual identity signals do not evolve, either because of the difficulties of recognizing large numbers of individuals, or because of group instability (Rohwer 1982).

However, beyond these few studies, there is little theoretical analysis of how group size might affect the evolution of signal diversity. In particular, the relationship between the number of signals found within each interacting group and the total signal diversity in the
population, has not been explored. Understanding the dynamics of this relationship should help explain why, for example, most humans interact regularly with only ~150 individuals (Dunbar 1992) and yet there are billions of apparently unique human faces on the planet.

Here we explore the effect of interacting group size on the evolution of individuality signals using an agent-based model of a population subject to different levels of fragmentation. We previously used this technique to show that variation in individual identity signals can arise as a consequence of even very weak sexual selection on male attractiveness (Thom and Dytham 2012). The model simulates a species in which females encounter male quality and identity information in two separately encoded signals that are temporally separated from the signaller. Females subsequently encounter the males and can correctly identify and mate with the highest quality individual only if his identity signal is unique – if it is not, they choose randomly from the males that share the signal. This temporal separation between assessment and mating mimics the mate choice mechanisms found in species that leave scent marks in the environment (Cheetham et al. 2007), broadcast auditory signals (Seddon and Tobias 2010), or in which females observe male contests and subsequently mate with the winner (Aquiloni and Gherardi 2010). Similarly, physical displays of attractiveness in humans – such as ritual jumping by Maasai men (Fink et al. 2019) – are often temporally distinct from subsequent mate choice events in which the chooser recognizes the ‘best’ male from the earlier display.

We predicted that either of two opposing outcomes could emerge from subdivision of the population into patches. First, because the benefits of signal uniqueness are related primarily to local diversity in a fragmented population, signal overlap between patches might not be strongly selected against and each individual signal might be repeated
numerous times at the global scale, thereby reducing global signal variation. Alternatively, because the evolutionary trajectory of each patch is largely determined by the local effects of drift and mutation, global signal diversity might exceed that found in a well-mixed population. We find that sexual selection can maintain local (within-patch) diversity in signalling loci even when the population is highly fragmented. Significantly, summing the signal variants across all patches reveals that population fragmentation increases the global signal diversity by 10 – 15% above that found in non-fragmented populations, revealing a substantial genetic diversity dividend to population subdivision.
Methods

We use an individual-based model of a sexual population with discrete events following Allen & Dytham’s (2009) adaptation of the Gillespie (1977) algorithm for simulating continuous time models. Extending the single, well-mixed population approach of Thom & Dytham (2012), we model a one dimensional ring of discrete patches connected by dispersal. An event can be either a birth (with possible dispersal to an adjacent population) or a death, and time advances after each event. The probability of a death event is density dependent and the population size will show stochastic logistic growth. We use an equilibrium density of 10,000 individuals divided equally across identical patches. The number of patches varies from 1 to 50, so the population ranges from 1 patch with 10,000 equilibrium density, to 2 patches with 5,000 through to 50 patches with an equilibrium density of 200.

Each individual carries a diploid attractiveness locus with alleles that can take any value, and six unlinked, diploid loci with four possible alleles at each locus. We consider the loci in two groups of three. One group controls signalling and the other evolves neutrally under mutation and drift only. There are 1000 possible unique combinations in each group (10 unique combinations at each locus, because genotype AB is phenotypically equivalent to BA), and thus $10^3$ (1000) possible individuality signals.

A random individual is chosen from the global population of N individuals and an event type (either birth or death) is chosen at random. Time moves on an average of 1/2N of a time step after each event. If birth is selected and the chosen individual is female, she chooses a mate from a random selection of 10 individuals from within the same patch. The focal female either selects a male on attractiveness criteria (see below) or is assigned one at
random, with a probability of 0.5 for each. If no males are encountered there is no birth, but if a male is encountered then the female produces a single offspring. At birth, the new individual is randomly assigned a sex (even sex ratio), inherits one allele for each of the six marker loci from each parent, and one attractiveness allele from each parent. There is no linkage. There is an independent chance of mutation for marker and attractiveness alleles. For signalling or neutral region mutations, there is a 1:1000 chance that one allele of 12 will mutate to one of the three different states. This represents a 1:6000 mutation rate per locus, which is of the order used in other simulation models (Roff 1998). Our mutation rate of 1:500 per locus for attractiveness is substantially higher because we assume that attractiveness is the product of numerous alleles across the genome, and thus that the mutational target is relatively large (Hunt et al. 2004). When an attractiveness allele mutates, a random number from a normal distribution with a mean of -0.02 and standard deviation of 0.02 is added to the existing allele with the result that the majority of mutations have negative effects on attractiveness. Following Thom & Dytham (2012), there is no upper limit on attractiveness. After birth the individual has a probability, set by the dispersal rate, of moving to an adjacent patch. Patches are arranged in a ring, and dispersal in a clockwise or anti-clockwise direction is equally likely. We use dispersal rates of 0, 0.0001, 0.001, 0.01, 0.1 and 0.5. Individuals have no more than one dispersal event during their lifetime.

We used a discrimination rate of 50% for simulations here, based on previous simulations of this system (Thom and Dytham 2012) — females choose the best male in half their mating opportunities; the rest of the time, they select randomly from males they have encountered. When discriminating, a female chooses a mate based on male attractiveness using the sum of the two attractiveness alleles. The female rejects all males that have
signalling loci different from those of the most attractive male, and then chooses a mate at
random from those remaining. There is no other effect of attractiveness or signalling on
fecundity, dispersal or mortality, and females assess male attractiveness without error.

Populations are initiated with the number of individuals equal to the equilibrium density
(10,000) spread randomly across patches. At initiation, each individual has an equal chance
of being male or female. All markers are set to the same value (i.e., there is initially no
variation in signalling or neutral loci) and each attractiveness allele is assigned a random
value drawn from a uniform distribution between 0 and 1. We describe a “time step” as the
period when the number of possible events is twice the population size. We used 100
realisations for each parameter set tested (the ‘neutral’ model, with female discrimination
rate set to 0, was replicated 40 times). Simulations ran for 50,000 time steps, by which time
population dynamics had settled into an equilibrium state.

Statistical models were performed using data from the end of the model run. We collected
data on signal diversity at two scales – global and local. Global signal diversity is the total
number of signal variants found in the entire population, and local signal diversity is the
mean number of signal variants in each patch. Effect sizes for local signal diversity are thus
the mean of means, as we used each model replicate as a statistical replicate in analyses. To
test the effect of increasing levels of fragmentation on signal variability, we conducted linear
models with signal number (either global or local) as the response variable and the number
of patches in the system as a factor – these analyses were performed pairwise, with each
level of fragmentation compared to both (a) the panmictic one-patch system and (b) the
next level of fragmentation to identify any threshold where increasing population
subdivision ceased to have any effect. To assess whether the sexual selection mechanism
was specifically driving up variation in the identity signalling system we compared the number of signal and neutral variants at the end of the model run using paired T tests. All analyses were performed in R version 3.3.2 (R Core Team 2017).

**Results**

Because the number of signals within a patch is limited by the number of individuals available to carry them, within-patch signal number is lower than in the panmictic system (across dispersal rates; all $F_{1,198} > 19.2$, all $p << 0.001$), as it is in the non-signalling regions invisible to selection (all $F_{1,198} > 5.0$, all $p < 0.027$ except 2 patches vs 1 patch at dispersal of 0.1 [$F_{1,198} = 2.0$, $p = 0.154$] and 0.5 [$F_{1,198} = 2.9$, $p = 0.093$]; Figure 1). However, signalling loci, which are under selection through female choice, retained higher levels of variation than non-signalling loci at all levels of fragmentation and dispersal (paired t-tests, all $t_{99} > 15.2$, all $p << 0.001$), even in the most conservative case of the 50-patch system and no dispersal (mean ± SE signalling variants per patch = $3.1 ± 0.04$; non-signalling variants per patch = $2.0 ± 0.02$; paired t-test $t_{99} = 26.2$, $p << 0.001$). Thus, sexual selection maintains positive selection on male signal rarity even when local population size is small and dispersal is extremely rare (see also Figure S1). We confirmed the expected isolation-by-distance in $F_{ST}$ values between pairs of patches (Figure S2). Tracking the spread of signals in a single replicate confirmed that genetic diversity was maintained by negative frequency dependence, ensuring that the number of signals present in the population remains diverse over time (Figure S3). By contrast, in the non-signalling region of the genome invisible to selection, drift leads to rapid changes in the frequency of the most abundant genotype, and in relatively small numbers of genotypes dominate in the population at any time.
The effects of fragmentation on evolutionary diversity across a species can be understood by investigating the global (population-wide) number of signal variants under different regimes. Global diversity in the signalling region remained significantly higher than genetic diversity of the neutral region across all dispersal and fragmentation levels (all t_{99}>14.3, all p <<0.001), demonstrating that population fragmentation does not break down the mechanism of sexual selection maintaining signal diversity at a global scale. Even more strikingly, at low to intermediate dispersal rates, the number of global signal variants significantly increased at intermediate levels of fragmentation compared to the levels of diversity seen in the single-patch system (Figure 1, hollow arrowheads). At the lowest non-zero dispersal rate of 0.0001, the global number of signal variants peaked at a value 10% higher than that found in the single patch system. At 0.001, 0.01 and 0.1 dispersal rates the peak was 13-15% higher than in a single patch system (all F_{1,198}>13.8, all p < 0.001). Population fragmentation was associated with lower global signal variation only in the absence of dispersal (Figure 1, top axis rug).
Even when populations become highly fragmented and subpopulation size is small, female choice of males they individually recognize can drive the evolution of genetic diversity in the signalling region. Indeed, fragmentation drives genetic variation in the population substantially above that of unfragmented populations, suggesting a potentially important role for population subdivision in maintaining evolutionary diversity.

In subpopulations as small as 200 individuals, sexual selection on male quality drives genetically-coded signal diversity higher than that found in an equivalent neutral genome region. Although the mechanism of selection modelled here is relatively weak – in only 50% of matings do females even attempt to discriminate the best male, and they investigate only 10 individuals before choosing – it was sufficiently effective to counteract the loss of allelic diversity through drift and to increase signal diversity across a range of demographic conditions (dispersal and local population size). We conclude that the evolution of individual variation, at least under this mechanism, does not appear to be prevented by small local population sizes. In small populations we find that the absolute number of signal variants is low: in the case with no dispersal and 50 patches there were only 3.1 signal variants per patch, meaning 65 individuals in each patch shared the same signal on average. Even at this high level of signal sharing, the mechanism of selection we describe here drives the evolution of greater signal diversity in signalling than non-signalling regions. In more biologically-plausible, intermediate parameter sets we see much lower rates of signal sharing (e.g. at 50 patches and dispersal of 0.1, there are 90 signal variants and just 2.2 individuals on average with each signal variant). While the number of signals in any population is constrained by either the number of carriers or the total combinatorial
diversity available from the signalling system, we have shown that selection can maintain
variation in both local and global signal numbers across a large range of population
fragmentation levels. Sexual selection is thus a robust mechanism for the evolution of
individuality signals.

More importantly, we find that global signal diversity is dramatically enhanced when the
population is subdivided. This contradicts our expectation that the rescue of rare alleles by
negative frequency dependence would be most effective in a panmictic population. Instead,
global signal diversity is elevated by population fragmentation by three mechanisms. First,
in a subdivided population there are many ‘best’ males (as many as there are patches), and
that the absolute quality required to be the local best is lower when the population is more
subdivided. Second, with many ‘best’ males the likelihood of a high-quality male also
carrying a rare signal variant is improved (since 1000 signal variants are possible in our
system, but in the most subdivided population there are only 200 individuals), giving more
opportunities for the selection mechanism to gain traction. Finally, in a subdivided
population any relatively high-quality individual that disperses will be more likely to possess
a rare signal variant in the destination population, increasing its chances of spreading
through selection on rarity and quality.

There are a number of examples of signal characteristics varying with geography, including
in chimpanzee calls (Mitani et al. 1999), in major urinary protein expression among
subspecies of house mice (Hurst et al. 2017, Sheehan et al. 2019), in human faces (Guo et al.
2014), and in intraspecific bird song dialects (Baker and Cunningham 1985). While this
gerographic diversity can develop under a number of processes, our model predicts such
variation in fragmented populations of species in which there is temporal separation of
mate assessment and mating. One such system in which this hypothesis might be tested in the future is in birds, where our data suggest that lekking species might avoid the negative genetic diversity effects of fragmentation: there is indeed some evidence that lekking grouse do not always suffer the expected decline in genetic diversity with population fragmentation (Bush et al. 2011, Segelbacher et al. 2008). Thus our model describes a mechanism for understanding of the paradox in which lek mating species maintain higher than expected genetic diversity in the face of sexual selection (Kotiaho et al. 2007).

Counterintuitively, our result suggests that fragmentation may in fact elevate genetic diversity in such a system, at least in signalling regions and among linked loci.

The rate of dispersal has substantial effects on patterns of genetic diversity in our model, as it does in fragmented wild populations (Riginos et al. 2014). With no dispersal we see the effects of drift vs. mutation and frequency dependence, and global signal diversity is not enhanced by population fragmentation. With a high dispersal rate the system behaves as a single, panmictic population. At intermediate dispersal, signals that are attached to high quality males increase in frequency, and thus increase their probability of spilling over into adjacent populations where the strength of positive selection will increase. Interestingly, there were quite striking effects of both dispersal and fragmentation on mean population quality (which was uncapped): the lowest rates of quality evolution were in the most fragmented populations with low dispersal, the highest rates in relatively unfragmented populations with high dispersal (approaching a panmictic system). This matches the prediction that selection should operate more effectively in larger populations where the impact of drift is reduced.
One counteracting pressure that we expect to see is “impersonation”, where an individual with low attractiveness, but whose individuality signal matches that of a high-quality male, gains ‘unearned’ reproductive output. Because females choose at random from within the pool of males that carries the best male’s signal, unattractive males are only likely to obtain matings from discriminating females if they are in this pool. This kind of identification error did happen in our system, although ‘unearned’ reproductive success was rare (< 10% of matings) except in very fragmented populations with very low dispersal rates (Figure S4).

This type of mimicry might be particularly likely to occur in systems that allow some signal plasticity (e.g. vocalizations: Hile et al. 2000). In our model, the most likely cost of inadvertent signal copying is that when an impersonated signal spills out into neighbouring patches the mean quality of the dispersers will be lower because of the imposter, and the spread will thus be weaker than it would be in the absence of impersonation. Of course, the risk of impersonation would be reduced with a larger signal set – we allowed 1000 signal variants, but this may be rather conservative compared to the number available with more loci or alleles contributing to the signal, or if there is variation not only in genotype but also in relative expression (e.g. Sheehan et al. 2016).

While our model simulates the type of social environment described by Sheehan & Bergman (2016), where an animal moves from one social group to another, the system described here does not require the accumulation of specific information about individuals following repeated encounters, since the female assesses quality and ‘memorizes’ matching individuality information simultaneously. The model operates purely through a series of instantaneous mate choice decisions by females. Much more complex mechanisms than this undoubtedly occur in species with complex social systems where repeated encounters and
memorization of individual-specific traits are an alternative mechanism explaining the evolution of individual recognition (Tibbetts and Dale 2007).

Previous studies of individual recognition have identified this process as an underappreciated mechanism for maintaining polymorphism (Sheehan and Tibbetts 2010). However, there has been little exploration — or even reporting — of the effects of individual recognition on species-wide genetic variation. Here we show that population fragmentation drives up global variation in signalling regions by between 10 and 15% above that expected in a non-fragmented system even when only half of the females are discriminating. This finding contrasts with the many examples in which habitat fragmentation is bad for diversity (Hanski 2015). Although the idea that physically isolated populations undergo separate evolutionary trajectories is not in itself surprising, the strength of the effect we demonstrate here, and the degree to which selection has an effect even in very small subpopulations, are potentially significant for conservation. This would particularly be the case if genetic diversity in non-signalling regions piggybacked on this increased diversity through, for example, linkage. Our result thus adds to the evidence for positive effects of habitat fragmentation on biodiversity (Fahrig 2003, Fahrig 2017, Fahrig et al. 2019).

Aquiloni, L. and Gherardi, F. 2010. Crayfish females eavesdrop on fighting males and use smell and sight to recognize the identity of the winner. – Anim. Behav. 79: 265-269.


Fahrig, L., Arroyo-Rodríguez, V., Bennett, J. R., Boucher-Lalonde, V., Cazetta, E., Currie, D. J.,
Eigenbrod, F., Ford, A. T., Harrison, S. P., Jaeger, J. A. G., Koper, N., Martin, A. E., Martin, J.-
L., Metzger, J. P., Morrison, P., Rhodes, J. R., Saunders, D. A., Simberloff, D., Smith, A. C.,
Fink, B., Butovskaya, M. L. and Shackelford, T. K. 2019. Assessment of physical strength from
Freeberg, T. M. 2006. Social complexity can drive vocal complexity: group size influences
Gillespie, D. T. 1977. Exact stochastic simulation of coupled chemical reactions. – The
Guo, J., Tan, J., Yang, Y., Zhou, H., Hu, S., Hashan, A., Bahaxar, N., Xu, S., Weaver, T. D., Jin,
face. – J. Hum. Evol. 75: 143-152.
59: 1209-1218.
Hurst, J. L., Beynon, R. J., Armstrong, S. D., Davidson, A. J., Roberts, S. A., Gomez-Baena, G.,
Smadja, C. M. and Ganem, G. 2017. Molecular heterogeneity in major urinary proteins of
*Mus musculus* subspecies: potential candidates involved in speciation. – Scientific Reports
10.1038/srep44992.

Hurst, J. L., Thom, M. D., Nevison, C. M., Humphries, R. E. and Beynon, R. J. 2005. MHC odours are not required or sufficient for recognition of individual scent owners. – Proc. R. Soc. B 272: 715-724.


Sheehan, M. J. and Nachman, M. W. 2014. Morphological and population genomic evidence that human faces have evolved to signal individual identity. – Nature Communications doi:10.1038/ncomms5800.


Figure 1.

Global number of genotypes (red symbols) and mean genotypes per patch (blue symbols) across a range of levels of population fragmentation (x-axis) at the end of the model run. Large symbols: genotypes visible to females and evolving under sexual selection; small symbols: genotypes invisible to females and evolving only under neutral processes. Data are means from 100 replicate model runs with standard deviations. The number of patches in the global population is shown on the x-axis, with six rates of dispersal between patches on separate panels. Both globally and locally, genotypes visible to selection had significantly higher numbers of variants than genotypes invisible to selection at all levels of fragmentation and dispersal. Note that for the single-patch system, global and local genotype variability are necessarily identical. The top axis rug (red ticks) marks levels of fragmentation at which global signal diversity is significantly different from signal diversity in the single-patch system; maximum global signal diversity for each dispersal rate is marked with an arrow. Because local signal diversity at all levels of fragmentation was significantly different from signal diversity in the single-patch system, the x-axis rug (blue ticks) instead marks points at which local signal diversity is significantly different (p < 0.05) from local signal diversity at the immediately preceding level of fragmentation. Rugs were calculated using linear models with number of genotypes as the response variable and number of patches (restricted to the two levels of interest) as a factor.
Figure 1

Dispersal = 0
Dispersal = 0.0001
Dispersal = 0.001
Dispersal = 0.01
Dispersal = 0.1
Dispersal = 0.5

Number of variants +/- SD

Number of patches in system