Animal &
Plant Health
Agency

Asiantaeth
Iechyd Anifeiliaid
a Phlanhigion

Estimating abundance and collecting evidence for population change in badgers, following trapping and testing interventions on Welsh farms

Animal and Plant Health Agency Report to the Welsh Government

Project TBOG0235

(Year 1)

Animal and Plant Health Agency (APHA) report: Estimating abundance and collecting evidence for population change in badgers, following trapping and testing interventions on Welsh farms.

Report for project TBOG0235 (Year 1)

Executive Summary

- The purpose of this work was to estimate the proportion of the target population of badgers trapped on each farm and to attempt to detect any signs of social perturbation arising from these small scale interventions.
- The choice of methods available was limited by the need to confine all field activities to the target farm, therefore data collection focused on genotyping hair samples from hair traps and captured badgers.
- We successfully extracted usable genetic profiles from 34% of hair samples collected remotely from hair traps (104, 66 and 40 on farms A, B and C).
- We successfully extracted usable genetic profiles from 78% of hair samples from trapped badgers (15, 10 and 6 on farms A, B and C).
- The results suggest trapping efficiency was 38%, 53% and 55% on the three respective farms.
- At this stage it is not possible to draw any conclusions regarding social perturbation. The small scale of the interventions (numbers of farms and numbers of badgers) remains a significant challenge to answering the question of whether perturbation is likely to be a consistent consequence of local interventions. However, the likelihood that any significant trends will be detected should increase if sufficient numbers of hair trap samples can be collected from each farm, and as the number of farms in the study increases in the future.

1. Abstract

As part of the 2017 Welsh Government (WG) TB Eradication Plan, a programme of work at selected persistent breakdown farms began, which involved trapping, testing and removing test-positive badgers (*Meles meles*). At the same time, work was conducted to estimate the proportion of the badger population that was trapped and tested at each farm, and to attempt to investigate signs of social perturbation after intervention work.

We used an established 'trap sample matching' approach, which involves matching trapped individuals to a representative sample of the background population on the basis of their genetic identities. The background sample was established by genotyping hairs captured using barbed wire hair traps deployed at setts and on badger runs. The trapped population was established by genotyping hairs plucked from captured animals. The percentage of individuals within the background samples that are also cage trapped is an estimate of the percentage of the total badger population that has been trapped and tested.

One intervention was carried out on each of three farms in 2017. Each intervention included a 14-day phase of remote hair trapping, followed by cage trapping, followed by another 14-day phase of remote hair trapping. A total of 210 hair samples collected from hair traps were genotyped and used for trap sample matching (104, 66 and 40 from farms A, B and C respectively). This equates to a genotyping success rate of 34%. A total of 31 of the hair samples from cage trapped badgers (15, 10 and 6 hair samples from Farms A, B and C respectively) were successfully genotyped and used for trap sampling matching, equating to a genotyping success rate of 78%.

Trap sample matching was carried out based on an exact match for 16 out of 18 possible alleles. A total of 34, 15 and 11 unique genotypes were identified at Farms A, B and C respectively, of which, 13, 8 and 6 included cage-trapped animals. These estimates suggest a cage-trapping success of 38%, 53% and 55% on the three farms.

Strict criteria were used to classify genotypes as useable for analysis, resulting in the exclusion of 64% of samples. To improve on the current analysis two actions need to be completed. First, the output from replicate genotypes that are inconsistent needs to be re-assessed, and second, if replicates are still inconsistent then additional replicate PCRs need to be run.

Preliminary analysis shows that the number of unique genotypes identified on farms was much greater pre-intervention than post-intervention. In the post intervention phase, new unique profiles were identified on two of the three farms. Indicators of change in badger movement or relatedness can only be investigated by additional monitoring over a number of years and on a larger sample size of farms.

2. Introduction

In 2017, following public consultation, the Welsh Government (WG) published its Wales bovine tuberculosis (TB) Eradication Programme and its associated Wales TB Eradication Programme Delivery Plan (Welsh Government 2017). WG's Delivery Plan states that: "As part of the ongoing Action Plan process, where the Welsh Government views that badgers are contributing to the persistence of disease in chronic herd breakdowns, badgers will be trapped and tested on the breakdown farm and test positive badgers will be humanely killed. Persistent

herd breakdowns will be focused on initially". The Delivery Plan also states that "WG will continue to assess the most appropriate deployment of the badger BCG vaccine if and when it becomes available".

The Animal and Plant Health Agency (APHA) was tasked by WG to develop a programme of work to implement these proposals through trapping, testing and removing test-positive badgers on persistent breakdown farms. As part of the work, research was instigated that aimed to estimate the proportion of badgers trapped and tested at each farm and to gather data to investigate evidence for social perturbation in badger populations before and after intervention work.

A number of different methodologies were considered for this purpose. Changes in badger spatial organisation have been investigated using radio-tracking (e.g. Tuyttens et al. 2000) and bait marking (e.g. Cheeseman et al., 1993). However, these methods are not feasible in this study because access to land surrounding the study farms (necessary in order to track animal movements or to map bait returns), was not permitted. Given this constraint, the only methodology that could be conducted within farm boundaries involved collecting badger hair samples directly from trapped animals or remotely by using hair trapping devices deployed around setts and along runs. The hair was genotyped, and the data used to distinguish individual animals following Frantz *et al.* (2004). The same methodology has previously been used to generate estimates of the percentage of badgers removed during industry-led badger culling in 2013 in England (AHVLA 2014) and to estimate the percentage of badgers vaccinated in Wales in 2015 (APHA 2017). In the present study we aim to investigate the following two questions:

- What proportion of badgers using the farm are successfully trapped and tested?
  Any trapping based intervention will be subject to less than 100% trapping efficiency, leaving a proportion of animals untrapped. By genotyping hair from remote hair traps and from captured badgers it is possible to estimate the proportion of badgers using the farm land that have been trapped and tested during sequential trapping operations.

- Will the interventions cause perturbation?
  There is compelling evidence that culling can give rise to social perturbation in badger populations (see Carter *et al.* 2007), but we have little information on the extent of any such response to very low levels of badger removal, or what the phenomenon looks like at the local scale. By genotyping hair samples from trapped or hair trapped animals we can attempt to quantify changes in dispersal or relatedness (see Pope *et al.* 2007). The application of this approach to try to identify social perturbation at the farm-level has not been attempted previously and is contingent on achieving sufficient sample sizes (numbers of hair samples) from individual farms, although the likelihood of detecting consistent responses (if they exist) would be expected to rise as the number of farms involved increased. The small scale of the intervention work (numbers of farms and/or numbers of badgers) remains a significant challenge to answering the question of whether perturbation is likely to be a consistent consequence of the local interventions.

3. Method

3.1 General approach

Trap sample matching involves matching trapped individuals to a representative sample of the background population on the basis of their genetic identities. The background sample is established by genotyping hairs captured using barbed wire hair traps deployed at setts and on runs. The trapped population is established by genotyping hairs plucked from every captured animal. The percentage of individuals within the background samples that are also cage trapped is an estimate of the percentage of the total badger population that has been trapped and tested. Potential indicators of social perturbation in the badger population post intervention are changes in the unique genotypes (i.e. unique individuals) that use the farm, and changes in the genetic relatedness of individuals within a farm.

3.2 Hair collection

Three farms were identified for initial interventions by WG in conjunction with APHA veterinary field staff. One intervention was carried out on each farm between August and November 2017. Each intervention took approximately 8 to 9 weeks to complete, from the initial badger activity survey to the end of post-intervention monitoring. Details of intervention methodology and results are reported elsewhere (APHA 2018).
In brief each intervention consisted of:

- Badger activity survey.
- Pre-treatment hair trapping to estimate abundance.
- Cage trapping and sampling trapped badgers.
- Post-treatment hair trapping to monitor for any evidence of social perturbation.

Each farm was surveyed for badger activity by experienced field staff. Following surveys for badger activity, hair traps were set for 14 days before cage traps were deployed. Hair traps were deployed on badger runs, at setts and elsewhere on the farm. Hair traps consist of strands of barbed wire suspended across sett entrances and/or nearby badger runs, using natural features where possible or wooden stakes if necessary. Each hair trap was labelled with a unique identifier, and was checked on each of the 14 days of deployment. All the hairs found on each hair trap on each visit were collected into a barcoded bag along with a sachet of desiccant. If hairs were caught on multiple barbs of the same trap, then separate small bags were used for each barb. All of these bags were enclosed in a larger single barcoded bag. The contents of each larger barcoded bag therefore represented a specific hair trap-day combination of samples. Once samples had been collected, hair traps were decontaminated by brief exposure to a naked flame.

Following hair trapping, cage traps were positioned at locations where there was most badger activity. Cage traps were pre-baited with peanuts for 8 to 10 days per farm prior to setting them to catch for 2 to 4 days. A pluck of approximately 10 guard hairs was taken from the rump of every trapped badger. The hair pluck was collected into a barcoded bag along with a sachet of desiccant.

Following cage trapping, hair traps were reinstalled at the same locations as before, and hair was collected every day for a further 14 days, before being dispatched to the laboratory for genotype analysis.

3.3. Genotyping

DNA was extracted from hairs using a suspension of chelex resin (Frantz *et al.* 2004) using the Qiagen DNeasy® Blood and Tissue kit. For hair samples plucked directly from a badger, ten hairs were selected from each barcoded sample bag. For samples collected from a hair trap, one hair was selected from each barcoded sample bag for analysis. One sample was selected because it is likely that the hair on a hair trap could contain contributions from more than one individual animal and if these individuals were presented in the hairs selected for analysis, then this could result in a mixed genetic profile. Selection of hair was based on the size of the hair follicle, which is the source of the DNA. For hair trap samples, hairs from different barbs of the same trap were collected into individual small barcoded sample bags. Thus hairs from the same trap on the same day were analysed if they were collected from different barbs of the trap.

Genetic profiles were obtained by amplifying ten microsatellites: *Mel-103, Mel-104, Mel-105, Mel-107, Mel-110, Mel-113, Mel-114, Mel-115, Mel-116, Mel-117*; (Carpenter *et al.* 2003). Microsatellite fragments were detected on an Applied Biosystems 3730xl Genetic Analyser and were analysed and sized using GeneMapper® Software (version 5).

3.4 Data quality control

Microsatellite genotyping is prone to error (e.g. Hoffman and Amos 2005), particularly when profiles are amplified and scored from poor quality DNA samples. Previous work, for example, estimates a 31% error rate for hair samples (Gagneux *et al.* (1997). See Jones and Conyers (2019) for further explanation. In an attempt to reduce error rates in the present study the following steps were taken.

- DNA profiles were replicated multiple times. Previous work (e.g. Taberlet and Luikart, 1999) recommends repeating each sample seven times, while other researchers (e.g. Huck *et al.,* 2008) repeated genotyping until the same alleles were observed at least twice in a heterozygous individual, or seven times in a homozygous individual. A more pragmatic approach was used in a similar study (AHVLA 2014), where all samples were genotyped at least twice and up to six times, and this was the method adopted for the present study.

- Automated allele calls performed in GeneMapper® were checked by two operators. This involved visual inspection of sample electropherograms (EPG), the graphical representation of amplified DNA used for determining genotype. Allele calls followed rules of the Standard Operating Procedure (Food and Environment Research Agency 2019). Final allele designation was determined by consensus between operators.

- Microsatellite *Mel-116* was omitted from analyses as previous work indicated it can be unreliable to assign a score to (Huck *et al.* 2008; AHVLA 2014), and following assessment of a subset of our data Jones and Conyers (2019) recommended that it be excluded from our analysis.

- All replicate genotype profiles were compared, and a genotype was deemed useable if the following criteria were met:
    - Genotyping had been repeated until the same alleles were observed at least twice in a heterozygous individual, or in three PCR replicates in a homozygous individual (Jones and Conyers 2019).
    - A genotype profile was complete (i.e. 9 markers were successfully genotyped). In a similar badger hair analysis, profiles were rejected when one or more

markers had no allele scores. Two or more missing markers is likely to reduce the values of Pident (the probability of two badgers being identical at all genotyped loci) and Psib (the probability of two siblings being identical at all genotyped loci) below an acceptable threshold, thereby raising the probability of falsely identifying two samples as coming from the same individual.

- Finally, as recommended by Jones and Conyers (2019), data were tested for null alleles using the program CERVUS (Marshall *et al.* 1998). A null allele is any allele that cannot be detected by the assay used to genotype individuals at a particular locus (Appendix 1) and should be excluded from some types of genetic analyses. CERVUS was also used to measure heterozygosity of genotype data from each farm, as this can be used as a measure of the diversity of a locus: loci with low heterozygosity are less informative for identifying unique profiles.

3.5 Trap sample matching

Genetic profiles of trapped badgers were matched to the hair trapped population to estimate the percentage of the population that had been caught. Matching was carried out using the statistical package Allelematch (Galpern *et al.* 2012), executed in R3.3.2 (R development Core Team 2018). Samples were matched at nine loci; *Mel-116* having been dropped from the analysis. Profiles that differed from one another by two mutations (i.e. matching at 16 of a possible 18 alleles) were assigned to the same individual (Appendix 2).

Error rates were calculated from the data to account for false positive rates (matching samples from different animals) and false negative rates (two genetic samples failing to match despite coming from the same individual).

Error rates were used to adjust the raw number of matched genotypes in order to produce final estimates of the percentage of badgers trapped and tested (Appendix 3). Since our calculated error rates were based on a small sample size, we decided to use published error rates from a larger sample size to estimate confidence intervals. The error rates were obtained from APHA (2017). This study showed that genotyping was unsuccessful for 51 out of 1118 cage trapped badger hair samples, providing a false negative rate of 4.6%. The study also showed that the proportion of cage trapped badgers that matched another cage trapped badger was 47/1025, giving a false positive rate of 4.6%.

In addition, to account for the fact that natural badger movements may have resulted in hair trapped individuals being unavailable at the time of cage trapping, a movement rate was estimated for the population (Rogers *et al.* 1998). An estimated movement rate of between 0 and 10% was used as this equates to the proportion of animals detected as having moved between badger social groups during trapping events predominantly over distances of less than 2 km (Rogers *et al.* 1998). Since all of the genotypes from captured badgers in our study originated from traps that were within 2 km of a farm boundary, the potential for movement into and out of the population was a consideration for all animals. The movement rate was also used to adjust the raw number of matched genotypes in order to produce final estimates of the percentage of badgers trapped and tested (Appendix 3).

4. Results

4.1 Summary of samples collected and genotyped

A total of 100, 86 and 40 hair traps were placed at Farms A, B and C respectively. The number of samples collected on each farm in the pre and post trapping phases are shown in Table 1. Complete useable genotypes were returned for 34% (210/624) of the samples that originated from hair traps, and for 78% (31/40) of samples that originated for cage trapped animals. This rate of genotyping is comparable to a previous study which reported success rates for genotyping hair trap samples of 39.3% (APHA 2017), although is lower than the 70.6% achieved in another study (AHVLA 2014).

Table 1 Summary of hair samples collected and genotyped from three farms in 2017

	Farm A		**Farm B**		**Farm C**		**Total**
Number of hair traps	100		86		40		
	Pre	Post	Pre	Post	Pre	Post	
Number of hair samples genotyped	221	77	111	49	139	27	624
Number of useable profiles	82	22	45	21	35	5	210
Number of cage trapped badgers +recaptures	16+1		11+1		10+1		37+3
Number of cage trapped badger hair samples genotyped	17		12		11		40
Number of useable profiles	15		10		6		31

4.2 Presence of null alleles and allelic diversity

Analysis suggested null alleles in the genotypes of farm A at *Mel-103*, and in the genotypes of farm C at *Mel-115* (Appendix 1). It is better to exclude loci with high null allele frequencies (0.05 or more) from some types of genetic analyses, however when data is only being used to assign hairs to an individual, null alleles are arguably not a problem (Jones and Conyers 2019). It is for this reason that these alleles were not excluded. In addition, there was a low level of heterozygosity in some loci in the genotype data, indicating that some loci are less informative when attempting to identify unique profiles.

4.3 Trap sample matching

Trap sample matching was carried out at nine loci based on an exact match for at least 16 of a possible 18 alleles. Analysis identified 60 unique profiles from 241 hair samples. There were 34, 15 and 11 unique genotypes on Farms A, B and C respectively.

At farm A 13 of the 34 unique genotypes included a cage trapped badger, suggesting that an estimated 38% of the badgers that used the farm were trapped. At farm B, 8 of the 15 unique genotypes (53%) included a cage trapped badger, and at farm C, 6 of the 11 unique genotypes (55%) included captured badgers. These results provide uncorrected estimates of the percentage of badgers in the population that were cage trapped. These values are equivalent to badger density estimates of approximately 8, 10 and 13 badgers per km^2 on farms A, B and C respectively.

The false negative rate was estimated by calculating the number of instances where multiple unique genotypes matched the same cage trapped genotype (i.e. one cage trapped sample appearing in more than one unique genotype). In our analysis this false negative rate was zero. A false negative rate also occurs if genotyping fails to produce full genotypes. Of the original 40 cage-trapped samples genotyping was successful for 31 samples, giving a false negative rate of 22.5%. The false positive rate was estimated by assessing the percentage of cage trapped badger genotypes that matched the genotype of another cage trapped badger. There were only two pairs of cage trap genotypes, out of 31 genotypes matched at 16 alleles, giving a false positive rate of 6.5%.

The estimated proportion of badgers trapped and tested on each farm was corrected for previously published false positive and negative genotyping rates, and for badger movements. The resulting corrected estimates of the percentage of badgers trapped and tested on farms A, B and C was 32-45% (95% confidence interval, mean 39%), 40-67% (95% confidence interval, mean 54%), 45-67% (95% confidence interval, mean 55%) respectively.

4.4 Perturbation

Occurrence of unique genotypes during different phases of intervention are shown in table 2.

Table 2. Number of unique genotype profiles identified pre and post intervention on three Welsh farms in 2017.

Farm	Number of unique profiles recorded:		
	Pre intervention and trapped only	Pre and post intervention	Post intervention only
A	21	11	2
B	7	5	3
C	7	4	0

5. Discussion

5.1. Trapping efficiency

The aim of the present study was to estimate the proportion of badgers captured and tested on the three farms that underwent intervention in Wales in 2017. We also collected data that might be used to look for evidence of social perturbation in the badger population as a result of intervention work.

Identification of 34, 15 and 11 unique genotypes on farms A, B and C provide respective trapping efficiency estimates of 32-45%, 40-67% and 45-67% (95% confidence intervals). These values equate to the results of two previous studies that used the same method to estimate trapping efficiency as part of vaccine deployment (APHA 2017) and to estimate culling efficiency (AHVLA 2014). Both pieces of work occurred in a much large area and used a larger sample size of hairs. APHA (2017) estimated that 44-65% (95% confidence interval) of the resident badger population in the Intensive Action Area (IAA) was trapped and vaccinated in 2015. AHVLA (2014) estimated a trapping efficiency of 16.7-39% in Gloucestershire, and a trapping efficiency of between 20.9 and 46.8% in Somerset.

The estimates of badger density (8, 10 and 13 badgers per km^2 on farms A, B and C), are greater than those predicted by land class type (Judge *et al.* 2017), which peak at 4.6-7.4 badgers per km^2 (95% confidence interval) for the most favourable land class type. While our estimate was based on a small geographic scale (as trapping was only permissible within farm boundaries) the actual number of badgers caught in 2017 (16, 11 and 10 on farms A, B and C) confirms that respective densities were at least 4, 7 and 13 badgers per km^2 within the farm boundaries.

5.2 Trap sample matching assumptions

All methods used for estimating population size and changes in populations rely, to a certain extent, on assumptions. These must be properly understood if the results are to be correctly interpreted (Appendix 4). In particular, genotyping of hair samples from hair traps can be prone to errors and as a result we established strict criteria with which to classify genotypes as useable. We also used a Mismatch parameter of two to avoid mismatching between samples as far as possible. Genotypes that were excluded from the analyses could be considered useable if further replicates were genotyped and/or scored consistently. Hence, through some additional genotyping it may be possible to improve on the accuracy of the existing estimates of population size and proportion captured.

5.3 Social perturbation

One year of data is not sufficient to assess evidence of changes in unique genotypes, nor the genetic relatedness of individuals on a farm. This is a novel piece of work that has not been attempted previously at this scale. The only other work of a similar nature, that aimed to assess evidence for perturbation, was conducted during the Randomised Badger Culling Trial (RBCT). In that work, hair was sampled and genotyped from cage trapped animals in order to assess whether the intervention influenced the movement of animals (Pope *et al.* 2007). Analysis of the genetic signature of badgers revealed increased dispersal following culling. However, this work used better quality hair samples (taken directly from captured animals), was conducted on a far larger sample size, and over a substantially larger area than the current Welsh Government interventions. Genetic profiles of 3450 adult badgers were collected from eight RBCT proactive cull areas, and the profiles of badgers from the first cull (a comparatively undisturbed population) were compared with those taken during a second cull 5-22 months later. Whether the approach can consistently identify the presence or absence of perturbation in the context of the present study, may only become apparent if pre and post intervention hair trapping is conducted on multiple farms over time. This is however highly contingent on a sufficient sample size of hair being collected during each phase.

6. Appendix 1 - Allelic diversity and presence of null alleles

6.1 Calculation of allelic diversity and null alleles

Allelic diversity and null allele frequency were tested using the program CERVUS (Marshall *et al.* 1998), (Table A1).

Table A1 allelic diversity and null allele frequency in badger genotypes at three Welsh farms in 2017.

Allele	K	H obs	H exp	HW	F(Null)
Farm A; n = 119 genotype profiles					
Mel-103	5	0.412	0.579	*	0.1770
Mel-104	4	0.807	0.511	***	-0.2393
Mel-105	6	0.739	0.770	**	0.0222
Mel-107	4	0.630	0.609	NS	-0.0280
Mel-110	5	0.975	0.764	***	-0.1307
Mel-113	4	0.681	0.607	NS	-0.0439
Mel-114	4	0.252	0.230	ND	-0.0597
Mel-115	5	0.597	0.543	NS	-0.0324
Mel-117	3	0.655	0.562	*	-0.0986
Farm B; n = 76 genotype profiles					
Mel-103	3	0.811	0.548	***	-0.2191
Mel-104	5	0.851	0.686	***	-0.1363
Mel-105	3	0.730	0.639	NS	-0.0929
Mel-107	3	0.135	0.129	ND	-0.0256
Mel-110	3	0.973	0.620	***	-0.2583
Mel-113	4	0.811	0.684	**	-0.0936
Mel-114	3	0.189	0.176	ND	-0.0416
Mel-115	4	0.662	0.524	NS	-0.1299
Mel-117	6	0.486	0.431	ND	-0.0823
Farm C; n = 46 genotype profiles					
Mel-103	4	0.674	0.687	NS	-0.0091
Mel-104	5	1.000	0.750	ND	-0.1534
Mel-105	3	0.739	0.612	NS	-0.0886
Mel-107	3	0.130	0.124	ND	-0.0246
Mel-110	5	0.957	0.752	**	-0.1381
Mel-113	5	0.717	0.672	NS	-0.0753
Mel-114	2	0.022	0.022	ND	-0.0011
Mel-115	5	0.370	0.574	NS	0.2204
Mel-117	4	0.848	0.749	ND	-0.0738

K: Number of unique allele calls
H obs: Mean observed heterozygosity (i.e. proportion of alleles that are heterozygous)
H exp: Mean expected heterozygosity
HW: Hardy Weinberg equilibrium. ***/**/* indicate significant deviation from the HW equilibrium. ND indicates 'not done'. NS indicates 'not significant'.
F(null): Null allele frequency.

6.2 Explanation of diversity measurements and null alleles

Observed and expected heterozygosity
Heterozygosity is a measure of the diversity of a locus: loci with low heterozygosity are less informative for identifying unique profiles. Loci with expected heterozygosity of 0.5 or less are in general not very useful for large-scale parentage analysis. Genotype data from each of the three farms indicated a low level of heterozygosity at some alleles, indicating that they are less informative when identifying unique profiles.

Hardy Weinberg (HW)
HW equilibrium describes the expected frequencies of genotypes in a population under random mating. In CERVUS, deviations from HW equilibrium are assessed using a chi-squared test, which compares observed genotype frequencies with expected genotype frequencies that are calculated from allele frequencies assuming HW equilibrium. Minor deviations from HW at one or two loci may occur because of natural selection. Such minor deviations are unlikely to bias likelihoods across all loci, and the loci can be used for parentage analysis as normal.

More commonly a deviation is an indicator of problems in genotyping that locus e.g. a segregating null allele, a failure to consistently distinguish alleles, biases towards typing particular genotypes, a locus that is sex-linked, cytoplasmically inherited or shows dominant inheritance. Ideally the locus should be excluded from analysis. Deviations at many or all loci are an indicator of population substructure. The population might be divided into a series of closely related or inbred family groups. In this case the data can still be used for parentage analysis, however it is necessary to interpret the confidence of parentage assignments with caution. If HW is not performed (ND=not done) it means that there were too few individuals to allow the test to proceed.

Data from Farms A, B and C, indicates that there is some deviation from HW equilibrium at some loci, although the cause of such deviation cannot be deduced here. Since this measure is of importance to parentage analysis, rather than assignment of individuals, no data was excluded.

Null Alleles
A null allele is any allele that cannot be detected by the assay used to genotype individuals at a particular locus. Null alleles are a common cause of apparent deviations from HW equilibrium (Pemberton et al 1995), particularly where only a single locus shows a deviation. With microsatellites, a null allele most often occurs because of mutations in one or both primer binding sites, sufficient to prevent effective amplification of the microsatellite allele.

Providing that at least 10 alleles are typed, CERVUS can estimate the frequency of any null allele segregating at each locus. In the absence of a null allele, the estimated frequency will be close to zero, and may be slightly negative (negative values imply an excess of observed

heterozygote genotypes). A locus with a large positive estimate of null allele frequency indicates an excess of homozygotes but does not necessarily imply that a null allele is present.

In the absence of known parent-offspring relationships, it is difficult to identify a null allele with certainty and it is important to consider other possible explanations. While it is advised to exclude loci with high null allele frequencies (0.05 or more) from parentage analysis; those with lower null allele frequencies should not cause any problems.

Analysis suggested null alleles in the genotypes of farm A at *Mel-103*, and in the genotypes of farm C and *Mel-115* (null allele frequency >0.05), however, since known parent-offspring relationships are absent from this data, it is not possible to assign null alleles with confidence. Data was not excluded, since null alleles are arguably not a problem when identifying individuals from hairs (Jones and Conyers 2019).

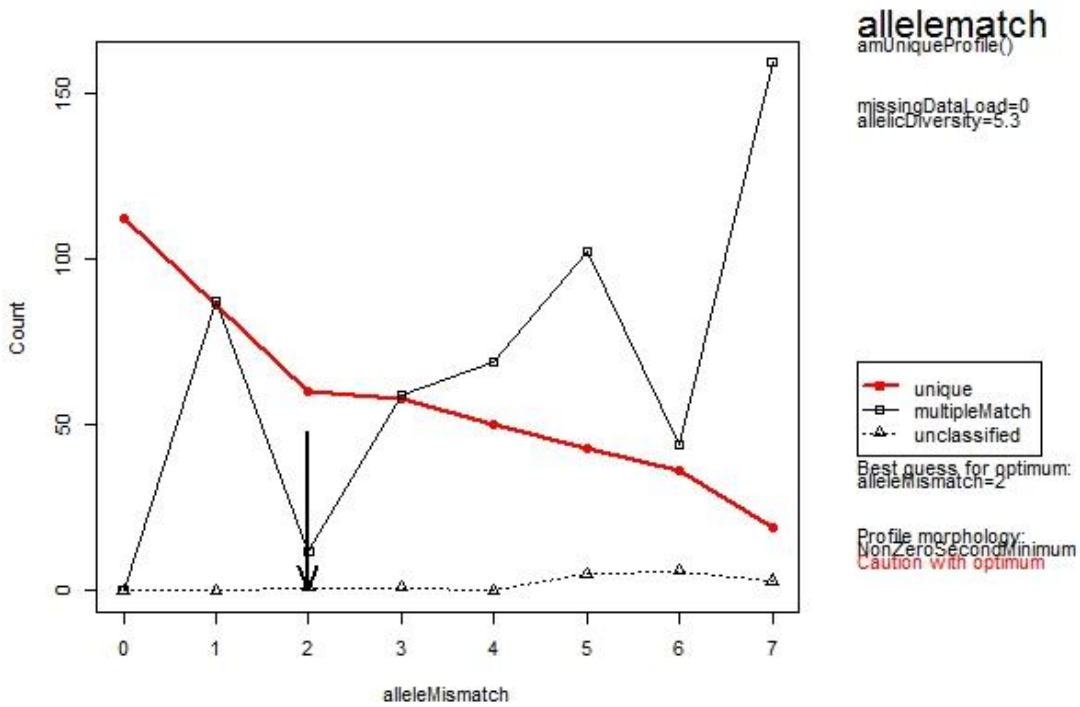
7. Appendix 2 – The Allele Mismatch parameter

The Allelematch program can be used to estimate the most appropriate number of alleles that are required to match in order to classify samples as the same unique genotype. This value is referred to as the allele mismatch parameter (Figure A1).

The allele mismatch parameter influences the number of unique profiles and the amount of multiple matching in the dataset. When the number of multiple match profiles approaches zero, this indicates that unique genotypes are sorting unambiguously into groups with minimal overlap. An allele mismatch parameter set too low inflates the number of unique genotypes because profiles with genotyping errors are declared unique. As a result a larger proportion of profiles will match multiple unique genotypes. Setting the parameter too high has the opposite effect, resulting in too many profiles identified as the same unique genotype when in fact they are different; the signature of this incorrect identification of unique genotypes is also a higher number of profiles that match multiple unique genotypes. Allelematch therefore declares the parameter value that achieves the least ambiguous sorting of profiles as optimum for the dataset.

For our dataset the optimum allele mismatch parameter is two. This means that two profiles will be declared the same if they match at 16 out of a possible 18 alleles.

Figure A1 Allele mismatch parameter (x axis) for genotype data from all farms combined



8. Appendix 3 – Details of trap sample matching

8.1 Principle of trap sample matching

Methodology for trap sample matching followed the same approach as two previous pieces of work (APHA 2017, AHVLA 2014). Hair samples are collected by hair trapping, hairs are genotyped providing a database of unique individual profiles, representing the 'target population' i.e. the badgers that use that farm. The target population is then compared with the DNA profiles from trapped animals to identify which of the target population had been trapped. Thus the percentage of badgers that were trapped and tested in the population can be calculated.

8.2 Identifying unique individuals and genetic matches.

Each individual, with the exception of identical twins, has a unique genetic code. Genetic profiling techniques aim to characterise enough of this code such that individual animals can be accurately identified. Hair samples were sequenced at 10 genetic markers, with one of these (*Mel-116*) excluded from the analysis due to the high presence of null alleles. Every marker yields two alleles (9 markers *2). Only hair samples with complete genetic profiles (where DNA amplified at all nine markers) were included in the analysis.

Genetic matches between hair samples were identified using the statistical package Allelematch (Galpern *et al.* 2012) using program R. Allelematch starts with pairwise comparisons between samples, it then estimates a similarity score between each pair of profiles, and then uses clustering to find groupings of similar profiles that likely belong to a single individual. Allelematch is particularly well suited to studies such as this one where genetic samples are obtained by remote sampling, which can result in variable sample quality and potential genotyping errors (Galpern *et al.* 2012). Genotyping errors ('stutter' or allelic dropout) may mean that genetic profiles differ slightly despite being obtained from the same individual. In this study samples were identified as being from the same individual if they matched at least 16 out of 18 alleles. The decision to match individuals at 16 alleles was in order to minimise the likelihood that genotyping errors would result in failure to match, or the creation of new false genotypes. Matching samples at one or a small number of alleles is commonly used in similar wildlife genetics studies (Hettinga *et al.* 2012).

Allelematch calculates Psib (the probability of the observed match occurring between siblings based on observed allele frequencies) for each genetic match and unique genotype. A threshold Psib value can also be used to remove matches which may have occurred by chance (Galpern *et al.* 2012) with a cut-off of <0.05 commonly used (Hettinga *et al.* 2012). All matched genotypes in the current study had Psib values of <0.05 (mean=0.001, min=0.0003, max=0.0024).

8.3 False negative (failure to match)

A false negative results can occur if two genetic samples fail to match despite coming from the same individual. For example if hair from an individual was hair-trapped and profiled, but when that badger was subsequently cage trapped, its genetic profile did not produce a match with its hair trap profile. Failure to match can lead to an overestimation of the population size, and therefore an underestimation of the proportion of badgers trapped. Failures to match could occur due to genotyping errors at more than 1 allele, although previous work has suggested that false negative rates may be low, since analysis of 749 badger hair samples did not find any genotyping errors (Frantz et al 2004).

A method with which to estimate false negative rates involves comparing the rate that distinct hair trapped genotypes match the same badger but not each other (AHVLA 2014). In such instances it can be assumed that the two hair genotypes are from the same individual but have failed to match. Thus, the effect of 'missed matches' on the percentage of badgers trapped can be estimated from the binomial proportion of the number of unique hair trap genotypes that match cage trap genotypes (AHVLA 2014). For example if 10 cage trapped individual animals matched 12 unique hair trap genotypes, then this would result in a false negative rate of 1-10/12 (9%). In our data, there were no instances were multiple unique genotypes matched the same cage trapped genotype.

False negatives also occur if full useable genetic profiles were not obtained from all cage trapped badgers. If this happened, then cage trapped individuals could not be matched to hair trapped individuals resulting in an effective false negative rate. In this analysis, nine of the cage trapped badger hair genotypes were not deemed useable profiles, thus the false negative rate here was 9/40 (22.5%). A previous study, using a larger sample size recorded a false negative rate of 51/1118 = 4.6% (APHA 2017). Since this value was based on a large sample size, we used it as an estimate of false negative rate in our analysis.

8.4 False positive rate

False positives occur where two samples are incorrectly matched to one another. This can occur because of genotyping errors, or if loosely related individuals are genetically identical at >= 16 alleles. False positives would inflate the number of matches between hair trapped genotypes and cage trapped genotypes, and thus result in overestimates of the percentage of badgers trapped. Thus the estimate of percentage of animals trapped were corrected for an estimated mismatch rate.

The false positive rate was estimated by examining the percentage of cage trapped badgers that matched the genotype of another cage trapped badger. In this study two pairs of animals had matching genotypes, out of 31 individuals. However, work involving a much larger sample size of hairs, showed a false positive rate of 47/1025=4.6%. We used this estimate in our analysis.

8.5 Movement rate

An assumption of trap sample matching is that hair trapped individuals are available to be cage trapped. However, the hair trapping and cage trapping was not carried out simultaneously. As a consequence, it is possible that hair trapped individuals are not available to be caught, either because they have moved out of the area before they could be cage trapped, or because they moved into the area after cage trapping. It is not possible to determine exact movement rates at each farm during this project. Long term monitoring of badger populations suggests that movement occurs between badger social groups in about 10% of trapping events (Rogers *et al.* 1998), with individuals moving on average 0.4-1km with most movements less than 2km. The rate of movement in the current study was estimated as 0-10% on each farm. This is because all traps were within 2km of a farm boundary and therefore movement into and out of the population was a possibility for all of these animals.

8.6 Estimating the percentage of badgers trapped

The percentage of badgers trapped was estimated as the percentage of hair trapped genotypes matching cage-trapped genotypes with correction for false positive match rate, false negative match rate and movement rate. This was implemented using a second order Monte Carlo estimate following (APHA 2017). In brief it involves the following steps:

1) Estimate a distribution for:
    a) false negative probability from missed matches (FN_1),
    b) false negative probability from failures to genotype (FN_2),
    c) false positive rate from matches amongst trapped animals (FP),
    d) probability of movement (FE).
2) Select independent random quantiles from each distribution.
3) Calculate the percentage of tested badgers as the number of captured animals (x)/ the number hair trapped individuals (n), using formula i.e. (x/n)*100.
4) Adjust the proportion of tested badgers, x, by the effective false positive rate eFP (random binomial draw with probability FP, with size x).
5) Adjust the number of hair trapped individuals n, by the effective false negative rates eFN_1 (random binomial draw with probability FN1, size n), eFN_2 (random binomial draw with probability FN_2, size n) and also by the effective population movement eFE (random binomial draw with probability FE and size n).
5) Calculate the adjusted percentage trapped x-eFP/(n-eFE-eFN_1eFN_2).
6) Repeat above steps 1000 times.
7) On each repeat, the population level estimate is a random quartile from the binomial proportion (x-eFP)/(n-eFE-eFN_1-eFN_2). The 95% confidence intervals are the 2.5^{th} and 97.5^{th} percentiles of this distribution.

9. Appendix 4 – Assumptions of data collection and analysis

Methodology and data analysis relied on assumptions that must be correctly understood (Appendix 4). The following assumptions were made:

- We used an allele mismatch parameter of two to match samples.
  Microsatellite genotyping is prone to errors, particularly when genotyping samples from animal hair. These errors cannot be avoided by good laboratory practice, however they can be, to a certain extent, accounted for during analysis. One likely genotyping error is the mismatching of samples from the same individual. A large number of such multiple matches would indicate some ambiguity in identifying unique genotypes. To avoid multiple matches we used a mismatch parameter of two, i.e. we allowed matches between genotypes that differed by two alleles. As a result there were 12 hair samples that presented as multiple matches, which were used to calculate a rate of false negative matches. A more stringent mismatch parameter of 1 was used in similar work (APHA 2017), however this would have resulted in much greater mismatching and an overestimate of the number of unique genotype profiles in our data.

  The low level of heterozygosity that occurred at some loci also likely influences the ability to identify unique genotypes. The reason for the low level of heterozygosity at some loci is unknown. One possibility is that it is a consequence of the limited size of the area where hairs were collected (i.e. the farm level), and perhaps indicates that some of the animals are related. It might also be due to difficulties in genotyping certain alleles, resulting in them consistently being excluded from analyses due to our stringent criteria for useable profiles. Heterozygosity in this dataset of genotypes may increase if excluded samples are re-tested until useable profiles are obtained.

- Strict conditions were used to deem a genotype profile useable. Specifically, we excluded any samples that were missing data or where three replicates did not consistently match. This resulted in exclusion of a large proportion of samples. Excluded samples could meet the useable criteria when laboratory staff have taken the following steps:
  - Assess the EPG of unuseable samples, to double-check allele calls at all loci, and to confirm readings when only two replicates for a sample could be completed. (Three replicates for some samples could not be completed when there was not sufficient DNA).
  - Identify samples with inconsistent replicates that require further genotyping. Samples that are homozygous are most difficult to score, and are the most likely to require multiple (more than 3) repeats to enable accurate scoring.

- We assumed that hair trapped badgers are representative of the target population. In an attempt to satisfy this assumption, hair traps were distributed throughout the farm area, and the number deployed was as high as practically possible. In previous work, hair trapping was conducted for 28 days (collecting a hair sample once every 2 days). This was done because it is likely that most, if not all, individuals would have been active at the main sett during a 28 day period (Scheppers *et al.* 2007) and hence available for trapping. This study, however, was subject to some practical limitations related to only having access to land within the farm boundary. As a result, although we hair trapped at setts wherever they were present within the farm boundary, a high proportion of traps had to be deployed remotely. Nevertheless, variation in the duration and intensity of

activity at the main sett will influence the probability of an animal being hair trapped and therefore, our result could be biased towards animals that are more active. It is also possible that our sample is biased towards adult badgers, as cubs may have passed beneath hair traps without contacting the barbed wire. This is a generic problem, but in this case, since sampling took place from August onwards, we avoided the time when very small cubs are present.

- We assumed that 14 days of hair collection would allow enough time to collect sufficient hair samples representative of the badger population. It was necessary to hair trap for this shorter time frame, rather than the 28 days used in comparable studies (AHVLA 2014, APHA 2017), in order to reduce costs and because resources were not available to conduct hair trapping for longer. We did however collect hair samples every day, rather than every alternate day. Before work began, we assessed the impact of reducing the hair collection period, and concluded that 14 days of collection should achieve a reasonable sample size. Specifically, analysis of hair trap data from Judge *et al.* (2017), where hair was collected every 2 days for 28 days, indicated that if the hair sample collection period was reduced to collection every 2 days for 14 days, then we would encounter at least 75% of the total sampled individuals at 50% of all setts (Smith pers. comm. 2017).

- The estimate of the movement rate (churn) of 0-10% of animals was derived from published rates of badger movement (APHA 2017). Incorporation of movement rate adds uncertainty to the estimate. If movement rates were higher than 10% then this would result in estimated percentage of badgers trapped being lower.

10. References

AHVLA. 2014. Animal Health and Veterinary Laboratories Agency (2014) report: Monitoring the efficacy of badger population reduction by controlled shooting during the first six weeks of the pilots. Report to Defra.

APHA. 2018. Animal and Plant Health Agency (APHA) report on the delivery of badger trap and test operations on chronic TB breakdown farms in Wales in 2017 (TBOG0235) (Year 1). APHA report to the Welsh Government.

APHA. 2017. Estimating the population size and the percentage of vaccinated badgers in the Intensive Action Area. APHA report to the Welsh Government.

Carpenter, P. J., D. A. Dawson, C. Greig, A. Parham, C. L. Cheeseman and T. Burke. 2003. Isolation of 39 polymorphic microsatellite loci and the development of a fluorescently labelled marker set for the Eurasian badger (*Meles meles*). Molecular Ecology Notes 3: 610-615.

Carter, S. P., R. J. Delahay, G. C. Smith, D. W. Macdonald, P. Riordan, T. R. Etherington, E. R. Pimley, N. J. Walker and C. L. Cheeseman. 2007. Culling-induced social perturbation in Eurasian badgers *Meles meles* and the management of TB in cattle: an analysis of a critical problem in applied ecology. Proceedings of the Royal Society of London Series B: Biological Sciences 274: 1471-2954.

Cheeseman, C. L., Mallinson, P. J., Ryan, J. and Wilesmith, J. W. 1993. Recolonisation by badgers in Gloucestershire. In Hayden, T. J. (ed.). *The Badger*, 78-93. Royal Irish Academy, Dublin.

Food and Environment Research Agency (FERA). 2019. Analysis of sample profiles generated by PCR amplification with DT1, and DT2 following electrophoresis on the 3130xls Genetic analyser. Standard Operating Procedure reference PLH/556 version 6 amended for use by APHA.

Frantz, A. C., M. Schaul, L. C. Pope, F. Fack, L. Schley, C. P. Muller, and T. J. Roper. 2004. Estimating population size by genotyping remotely plucked hair: the Eurasian badger. Journal of Applied Ecology 41: 985-995.

Gagneux, P., C. Boesch and D. S. Woodruff. 1997. Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. Molecular Ecology 6: 861-868.

Galpern, P., M. Manseau, P. Hettinga, K. Smith and P Wilson. 2012. Allelematch: an R package for identifying unique multilocus genotypes where genotyping error and missing data may be present. Molecular Ecology Resources 12: 771-778.

Hettinga, P. N., A. N. Amason, M. Manseau, D. Cross, K. Whaley and P. J. Wilson. 2012. Estimating size and trend of the North Interlake woodland Caribou population using Faecal DNA and capture-recapture models. The Journal of Wildlife Management 76: 1153-1164.

Hoffman, J. I. and W. Amos. 2005. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Molecular Ecology 14: 599-612.

Huck, M., A. C.Frantz, D. A. Dawson, T. Burke and T. J. Roper. 2008. Low genetic variability, female biased dispersal and high movement rates in an urban population of Eurasian badgers *Meles meles*. Journal of Animal Ecology 77: 905-915.

Jones, E. P. and Conyers, C. 2019. Troubleshooting of badger hair microsatellite data and analysis for APHA project TBOG0235. Report to APHA.

Judge, J., G. J. Wilson, R. Macarthur, R. A. McDonald, and R. J. Delahay. 2017. Abundance of badgers (*Meles meles*) in England and Wales. Scientific Reports 7.

Marshall, T. C., J. Slate, L. E. B. Kruuk and J. M. Pemberton. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7:639-655.

Pemberton, J. M., J. Slate, D. R. Bancroft and J. A. Barrett. 1995. Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. Molecular Ecology 4: 249-252.

Pope, L. C., R. K. Butlin, G. J. Wilson, R. Woodroffe, K. Erven, C. M. Conyers, T. Franklin, R. J. Delahay, C. L. Cheeseman and T. Burke. 2007. Genetic evidence that culling increases badger movement: implications for the spread of bovine tuberculosis. Molecular Ecology 16: 4919-4929.

Rogers, L. C., R. J. Delahay, C. L. Cheeseman, S. Langton, G. C. Smith and R. S. Clifton-Hadley. 1998. Movement of badgers (*Meles meles*) in a high density population: individual population and disease effects. Proceedings of the Royal Society of London B: 265: 1269-1276.

Scheppers, T. L. J., A. C. Frantz, M. Schaul, E. Engel, P. Breyne, L. Schley, and T. J. Roper. 2007. Estimating social group size of Eurasian badgers *Meles meles* by genotyping remotely plucked single hairs. Wildlife Biology 13: 195-207.

Taberlet, P., L. P. Waits and G. Luikart. 1999. Non-invasive genetic sampling: look before you leap. Trends in Ecology and Evolution 14: 323-327.

Tuyttens, F. A. M., R. J. Delahay, D. W. Macdonald, C. L. Cheeseman, B. Long and C. A. Donnelly. 2000. Spatial perturbation caused by a badger (*Meles meles*) culling operations: implications for the function of territoriality and the control of bovine tuberculosis (*Mycobacterium bovis*). Journal of Animal Ecology 69: 815-828.

Welsh Government. 2017. Wales TB eradication programme delivery plan. http://gov.wales/topics/environmentcountryside/ahw/disease/bovinetuberculosis/bovinetberadication/?lang=en

Wilson, G. J., R. J. Delahay, A. N. S. de Leeuw, P. D. Spyvee and D. Handoll 2003. Quantification of badger (*Meles meles*) sett activity as a method of predicting badger numbers. Journal of Zoology 259: 49-56.