GOVERNMENT OF WESTERN AUSTRALIA

Department of **Biodiversity, Conservation and Attractions**

Spatial and temporal analysis of ghost bat populations in the Pilbara using non-invasive sampling methodologies: Towards a robust genetic monitoring protocol. Phase 1: Refining our genetic methodology

Version: 2.0 Approved by:

Last Updated: 26 April 2024

Custodian: Rujiporn Thavornkanlapachai Review date:

Version number	Date approved DD/MM/YYYY	Approved by	Brief description
1.0	04/12/2023	Kym Ottewell	Draft report for comment
2.0	26/04/2024	Kym Ottewell	Final version

Biodiversity and Conservation Science

Spatial and temporal analysis of ghost bat populations in the Pilbara using non-invasive sampling methodologies: Towards a robust genetic monitoring protocol.

Phase 1: Refining our genetic methodology

Rujiporn Thavornkanlapachai, Kym Ottewell and Shelley McArthur

Final report

April 2024

GOVERNMENT OF WESTERN AUSTRALIA

Department of **Biodiversity, Conservation and Attractions**

Department of Biodiversity, Conservation and Attractions
Locked Bag 104
Bentley Delivery Centre WA 6983
Phone: (08) 9219 9000
Fax: (08) 9334 0498

www.dbca.wa.gov.au

April 2024

This report was prepared by Dr Rujiporn Thavornkanlapachai and Dr Kym Ottewell

Questions regarding the use of this material should be directed to:
Dr Kym Ottewell
Biodiversity and Conservation Science
Department of Biodiversity, Conservation and Attractions
Locked Bag 104
Bentley Delivery Centre WA 6983
Phone: (08) 9219 9086
Email: kym.ottewell@dbca.wa.gov.au

The recommended reference for this publication is:
Thavornkanlapachai, R., Ottewell, K., and McArthur, S. 2024. *Spatial and temporal analysis of ghost bat populations in the Pilbara using non-invasive sampling methodologies: Towards a robust genetic monitoring protocol. Phase 1: Refining our genetic methodology*. Department of Biodiversity, Conservation and Attractions. Kensington, Western Australia.

This document is available in alternative formats on request.

Contents

Appendices

Figures

Tables

Acknowledgements

We would like to thank Dr Kyle Armstrong from The University of Adelaide for providing DArTseq data on Pilbara ghost bat populations, which were developed into SNP markers. We wish to thank Trent Peters, Tamara MacDonald and Kimberley Rogl from Australian Genome Research Foundation (AGRF) for assistance with SNP panel development and SNP genotyping. Thanks also to Chris Knuckey, Brighton Downing and others from Biologic Environmental Survey for providing samples and assistance with sample metadata.

Summary

This report demonstrates the transition from microsatellite markers to SNP markers for ghost bat genetic monitoring, including alignment of genetic identifications from the two datasets and evaluation of the new technology's performance using the South Flank 2019 case study. We also report other developments and refinements to our analysis workflow including the addition of molecular sexing markers, changes in data processing, and SNP panel modifications.

In summary:

- 329 out of 444 unique ghost bat genotypes (74%) identified using microsatellite markers were successfully converted to SNP genotypes. A small number of South Flank putative resident bats (n = 6) were not able to be genotyped.
- 169 males, 146 females, 7 potential females and 7 undetermined sexes were identified using the TaqMan probe molecular sexing assay.
- Alignment of microsatellite and SNP genotypes detected 21 bat genotype IDs requiring reassignment where they had previously been identified as different individuals by microsatellite markers.
- Three SNP panels were developed, but one panel of ~50 SNPs was sufficient for individual identification from scats. Client reports from January 2020 onwards focus on analyses from this panel only.
- A 2-step data analysis approach to firstly detect bats from scats and, secondly, identify recaptures was resolved:
    - We found scats with high sample and locus amplification rates provided greater certainty in identifying individual bats, though fewer are detected;
    - We changed our approach from using a consensus genotype from multiple scats to using a single 'best' representative scat to identify recaptures from our database.
- To validate the transition between marker types, we found comparable numbers of identified bats and levels of roost activity generated from SNP genotypes compared to microsatellite markers in the South Flank 2019 case study (Ottewell et al. 2021).
- Following assessment, we improved SNP Panel 1 by removing failed markers, adding high-performing markers from Panels 2 and 3 to replace these, and adding one sexing marker (Zfx/y#1). Genotypes were highly reproducible between old and new panels, however, the sexing marker was less consistent.
- Molecular sexing using the TaqMan probe assay is recommended for ongoing analyses.

1 Background

Ghost bat biology

The ghost bat (*Macrodema gigas*) is listed as Vulnerable under IUCN criteria (Armstrong et al. 2021) and the EPBC Act (1999) and has a carnivorous diet consisting of small vertebrates including other bats (Churchill, 1998). Their distribution has contracted to sparser populations across northern Australia in response to increasing aridity and anthropogenic threats (Woinarski et al. 2014). Populations are impacted by habitat disturbance/loss, quarrying activities, habitat modification for livestock, cave tourism, competition with introduced foxes and cats and poisoning by cane toads with the current global population estimated to be <10,000 (Woinarski et al. 2014, Threatened Species Scientific Committee 2016). The Pilbara population is the most geographically and genetically isolated of the known ghost bat populations (Worthington Wilmer et al. 1994, Woinarski et al. 2014) and is estimated to consist of ~1200 bats (Armstrong and Anstee 2000). The ghost bat is an obligate troglodyte, and survival is critically dependent on finding natural roosts in caves, crevices, deep overhangs and artificial roosts such as abandoned mines (Hall et al. 1997). The dispersal pattern of this species is still largely unknown. Tracking studies of ghost bats indicate movements of 8 – 18 km during foraging bouts (Augusteyn et al. 2018, Bullen et al. 2023), although Toop's (1985) and Bullen et al. (2023) records of marked ghost bats indicate longer return flight of distances between 20 km and 50 km, with bats travelling as far as 90 km (Bullen et al. 2023) and 150 km (Toop 1985) over multiple nights, suggesting a high dispersal capacity for this species. With limited literature and knowledge gaps in habitat usage, critical roosts and landscape movements, additional long-term monitoring of ghost bats is needed to provide insight into their biology.

Challenges of monitoring ghost bats – Why scat genotyping?

Wildlife monitoring programs often use live capture, visual counts, or passive acoustics to estimate target species abundance but these methods can be unreliable due to heterogeneity in individual detection probability, low capture/recapture rates, trap shyness, or tag loss (Hoyle et al. 2001, Kunz and Fenton 2005, Augusteyn et al. 2018). Minimally invasive techniques such as visual counts or passive acoustics can suffer low precision due to the inability to distinguish individuals. 'Molecular tags', i.e. using individual genotypes identified from scats, can overcome many of these apparent issues as tags are permanent and capture/recapture doesn't rely on the target species encountering and entering traps. Thus, molecular tagging offers an alternative method with minimal interference to monitor ghost bat populations. Molecular tags using scat genotyping with microsatellite markers have been successfully applied in wildlife monitoring for many years (Piggott 2004, Berry et al. 2007, Sittenthaler et al. 2020). The benefit of microsatellite markers is that statistical power to identify individuals is achieved with few markers but many alleles within each marker. This technique is particularly useful for Capture-Mark-Recapture (CMR) methodologies which can provide information on bat populations, where recaptures over multiple CMR sessions/locations can be used to track movements of individuals and the identification

of new captures allows population size, longevity and recruitment of a population to be monitored.

What we have done in the past and why change

Microsatellite markers for ghost bats were first developed by Jane Hughes at Griffith University for monitoring the Queensland ghost bat population (Augusteyn et al. 2018) and, in 2015, were tested for application to the Pilbara population by Spencer and Tedeschi (2016) at Murdoch University. The genetic methodology employed by the Department of Biodiversity, Conservation and Attractions (DBCA) built on the Spencer and Tedeschi (2016) foundational work via isolation and refinement of additional microsatellite markers specific to the Pilbara population (Ottewell et al. 2020b). Between 2015 and 2018, DBCA generated individual genetic 'fingerprints' for ghost bat surveys via a panel of 11 microsatellite markers (Ottewell et al. 2020b).

Although microsatellite markers are useful for ghost bat scat genotyping, for low-quality and/or poor-quality DNA samples like scat DNA, laboratory artefacts such as stutter peaks, false alleles or allelic dropout can complicate genotype calling. While this can be ameliorated by applying additional laboratory protocols (e.g. a multi-tubes approach analysing multiple PCR replicates and calling a consensus genotype) and manually assessing genotypes to ensure correct and consistent calls, both are expensive and time-consuming (Taberlet et al. 1996, Frantz et al. 2003).

To improve the accuracy and efficiency of the genetic analyses, in 2019, Single Nucleotide Polymorphism (SNP) panels were developed and optimised to generate ghost bat genetic profiles (Thavornkanlapachai et al. 2024). SNPs are high-resolution molecular markers which produce binary data that are easy to score, analyse and reproduce. A large number of pre-selected markers can be surveyed simultaneously using automated genotype calling technology. SNP markers have been shown to genotype with higher precision, have lower genotyping error rates and require fewer repeats (Ekblom et al. 2021). In addition, DBCA has optimised a sex identification assay which is run alongside either the microsatellite or SNP panels (Ottewell et al. 2020b). Combined, these methods enable the identification of individuals and their sex from a single scat.

Development of ghost bat SNP panels

To develop SNP genotyping panels, we selected SNP markers from a previous genomic study of eight Pilbara populations generated through Diversity Arrays Technology (DArTseq) (K. Armstrong, unpublished data). From this data, 33,340 variable SNP loci were available that were subsequently screened for a single SNP per sequencing read, high read depth and high information content (heterozygosity, minor allele frequency) to identify high-quality loci for individual identification and discarding loci failing tests for Hardy-Weinberg equilibrium and linkage disequilibrium (Thavornkanlapachai et al. 2024). Initially, 611 informative SNPs were identified but, due to constraints on the multiplex primer design for MassArray genotyping (max n = 50 markers per panel), only 147 SNP loci were selected for multiplexing. These were arrayed across three MassArray SNP genotyping panels (Panel 1 = 50 SNPs, Panel 2 = 50, Panel 3 = 47). Probability of Identity (P_{ID}) analysis indicated using at least 20

loci is required to distinguish related individuals with a high level of certainty (PID < 0.0001; Thavornkanlapachai et al. 2024). For genetic monitoring purposes, we utilised the first panel as an Individual Identification (ID) panel to cluster genotypes and identify individuals. To increase the resolution of genetic diversity estimates and resolve genetic structure, use of all three panels is required. Further details are available in Thavornkanlapachai et al. (2024).

The transition from DBCA microsatellite database to SNP database and refinement of our genetic methodology

DBCA maintains a database of ghost bat genotypes (unique individuals) from which 'recaptures' are identified. Ghost bat individuals between 2015 and 2018 were identified based on 11 microsatellites profiled from scats. Of 2,422 scats that passed the quality control (QC) filters, at least 444 bats were identified. With the development of the ghost bat SNP panels in 2019, we switched to genotyping scats on the MassArray automated SNP genotyping system. This report aims to ensure the transition from the ghost bat capture database based on microsatellite IDs to those based on SNPs. Firstly, we describe the development and refinement of molecular sexing assays using custom-made sexing markers (three Y-linked and one X-linked markers modified from Ottewell et al. 2020b). We then re-genotyped previously identified individuals from the microsatellite database with three SNP panels to obtain their SNP profiles and assign sex with the custom-made sexing markers. We investigated different approaches to matching novel genotypes from the South Flank 2019 project to the DBCA ghost bat database to identify the most accurate and efficient way to identify recaptures going forward. Third, we assessed the reproducibility of results and outcomes between microsatellite and SNP markers by using data from the previous South Flank 2019 report (Ottewell et al. 2021). Finally, we described the refinement of the first SNP panel as well as the data processing procedure to streamline analysis workflow.

2 Development and refinement of molecular sexing assays

2.1 Development of alternative molecular sexing approaches

Molecular sexing markers for ghost bats were initially identified by Ottewell et al. (2020). The original sexing assay outlined in this publication involved design of standard Polymerase Chain Reaction (PCR) markers targeting DNA fragments of four sex-linked genes (Y-linked genes: DDX3Y, SRY, Zfy, and X-linked gene: Zfx) and was designed for in-house sex identification using fluorescent-labelled primers and fragment analysis. This methodology, however, was time-consuming (approximately one week from sample to result) and required manual post-processing of genetic data. To improve the efficiency of sexing analyses, we investigated the viability of two additional methodologies: (1) development of TaqMan probes for sexing on a real-time quantitative PCR machine (qPCR) and (2) development of sexing assays for inclusion in the MassArray SNP panels to enable automated genotyping.

TaqMan probes are fluorescently-labelled oligonucleotides that are located within PCR product fragments (Figure 1) and designed for automated detection on a real-time PCR machine. The amount of fluorescence detected in the real-time PCR assay is directly proportional to the amount of DNA template present in the PCR. Thus, we can calibrate the amount of fluorescence detected in Y-linked markers against that of X-linked markers to determine sex. That is, Y-linked markers will only fluoresce in males and X-linked markers will only fluoresce at half the amount in males as in females. Real-time PCR generates results within two hours, increasing laboratory efficiency in both time and cost. We designed TaqMan probes with fluorescent dyes attached (DDX3Y-FAM, SRY-HEX, Zfy-TEX615, and Zfx-Cy5) to work with the standard sex-linked PCR markers from Ottewell et al. (2020b). However, because we routinely send samples to genotype on MassArray, we considered it more cost-efficient to develop sexing markers to include on MassArray panels.

*Figure 1 Ghost bat Zfx and Zfy sexing markers targeting a section of the potential sex-linked genes. Horizontal bars are DNA sequences that match the reference genome (*Macroderma gigas*) and small verticals of different colour inserts are SNP differences to the reference genome. The red and blue arrows are the PCR products of Zfx and Zfy genes respectively. The lines with two coloured circles at each end represent TaqMan probes.*

MassArray sexing markers were designed to target a single SNP position difference between males and females (Figure 2). Females and males are expected to have different allele combinations e.g. presence/absence for SRY and DDX3Y, GA in females and GG in males for DDX3X, and GG in females and GA or GT in males for Zfx/y. We conducted 2 trials on these markers, but they did not produce consistent results when compared to the TaqMan probe sexing array. Consequently, molecular sexing results presented in this report are obtained through typing using the TaqMan probe sexing assay.

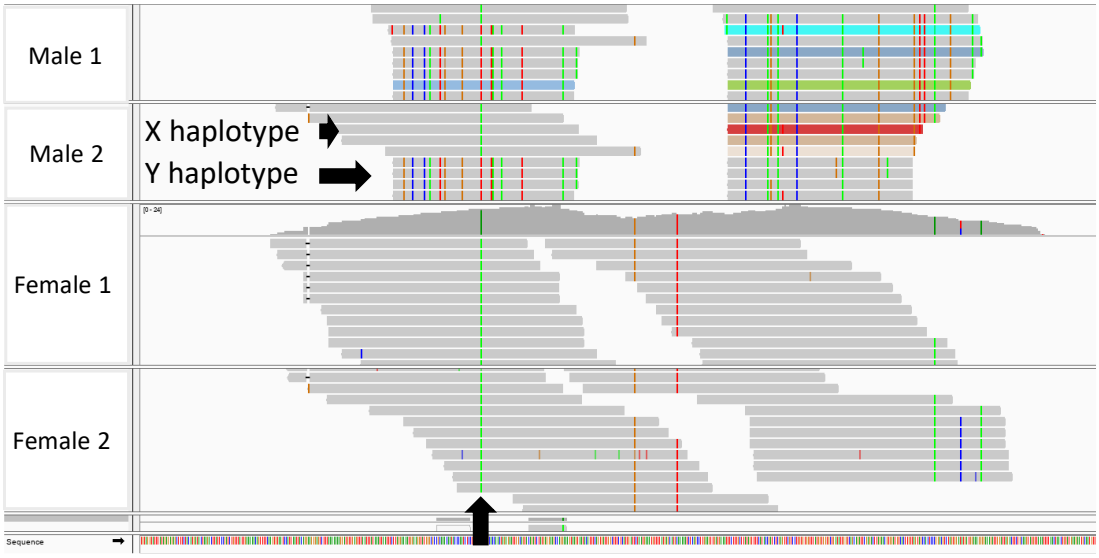


*Figure 2 Ghost bat DDX3Y MassArray marker targeting a single position SNP of a potential sex-linked gene. Horizontal bars are DNA sequences that match the reference genome (*Macroderma gigas*) and small vertical of different colour inserts are SNP differences to the reference genome. X haplotype on the left in males is similar to haplotypes observed in females while Y haplotype on the left and right fragments are unique to males. Potential SNPs are those that are unique at alternate alleles in each sex.*

2.2 TaqMan molecular sexing assay and sex-identification

TaqMan molecular sexing assay consists of custom-designed ghost bat sex-linked primers and probes for DDX3Y, SRY, Zfy, and Zfx which were arranged in a real-time PCR multiplex. Primers were amplified in 10 µl reactions using the PrimeTime™ Gene Expression Master Mix (Cat No: 1055772) following the manufacturer's instructions with an annealing temperature of 60°C, 40 amplification cycles and 4 µl of unconcentrated DNA. The reactions were run on the CFX96™ Real-Time System C1000 Touch Thermal Cycle (BIO-RAD, Singapore) and analysed in BioRad CFX Maestro software (BIO-RAD, Singapore). We repeated qPCR twice to ensure consistency in sex identification.

We found that scat quality can affect amplification consistency of sexing markers, particularly as the amount of DNA template from the Y chromosome is half of that from autosomes or X chromosomes in females. Thus, we developed several criteria to assign sex based on TaqMan assay results. First, amplification was considered successful if the qPCR Relative Fluorescence Units (RFU) ≥ 50. Samples were considered male if they met additional criteria as follows: a ratio of Y- to X-linked RFU > 0.1 and consistently assigned an individual to the same sex in multiple samples. 'Likely' sex is defined as a group of samples with small disagreement between markers and/or samples, and the sex selected made up the majority of the result. 'Undetermined' is defined as a group of samples with an amplification signal below 50 RFU from multiple markers or sex that cannot be confidently assigned.

3 SNP genotyping on previously identified individuals from microsatellite markers

3.1 DNA extraction and genotyping

To transition and calibrate microsatellite and SNP genotyping projects we selected one representative scat DNA sample from 444 individuals previously identified by the microsatellite markers for SNP genotyping. Since amplification rate is often associated with high DNA quality and/or quantity (Berry et al. 2007, Carpenter and Dziminski 2017, Sittenthaler et al. 2020), we selected the scat DNA sample with the highest microsatellite amplification rate for re-genotyping. We first sexed the samples with the TaqMan probe sexing assay as a form of screening for successful PCR amplification. If the selected sample failed to amplify, we replaced it with the next best sample until we had a working sample. However, due to variation in sample quality, a different DNA extraction kit being used and potential sample degradation during storage, we found many samples failed to work despite sample replacement. We identified an issue with the presence of ethylenediaminetetraacetic acid (EDTA) in the QIAamp® Fast DNA Stool Mini kit elution buffer (Qiagen, Germany, Cat No: 51604) used in genotyping projects between 2015 – 2018. EDTA interfered with the sexing qPCR reaction (as well as MassArray) resulting in failed or low amplification signal in many samples. In subsequent projects, we changed the DNA extraction kit to the Omega Biotek Mag-Bind Stool DNA 96 kit (Omega, USA, Cat No: M4016-01) and substituted half of the elution buffer with sterile deionised water. This significantly improved the average amplification success rate at SNP loci from 68.6% to 92.1% (Appendix 1).

Out of 444 individuals, 372 samples were successfully amplified and sent for MassArray genotyping using a commercial service (AGRF) with the three designed SNP array panels (147 loci). We repeated PCR and SNP genotyping on 43 samples to calculate the genotyping error rate. Across the 372 re-genotyped samples, the average amplification rate was 67.8 ± 0.6%. Genotyping error rate was low for allelic dropout (1.0 ± 0.1%) indicating that replicate genotypes were highly consistent with each other.

3.2 Re-genotyping success rate

Raw genotyping data from MassArray requires some manual quality control to remove low-performing samples and/or loci before being suitable for formal analysis. We applied a relaxed sample quality retention threshold of ≥ 30% amplification rate to both markers and samples to accommodate variation in sample quality and because we know that these genotypes are from unique individuals with a low genotyping error rate. With the 30% relaxed amplification threshold, we retained 329 out of 372 samples from 117 loci for further analysis.

From the molecular sexing assay previously undertaken on these samples, we identified 169 males, 146 females, 7 potential females and 7 undetermined sexes (Figure 3).

Samples excluded from the reference database either due to sample failure (grey squares, Figure 3) or not passing quality control thresholds (yellow squares, Figure 3) are listed in Appendix 2. Almost half of failed samples (55/115) were individuals detected in the South Flank area. The great majority of these individuals were detected at only one time point (49/55 individuals, Table A2.1), however, 6 individuals had previously been detected on multiple occasions and/or in multiple locations suggesting they may be resident bats in the area.

Collectively, the failure to obtain replacement SNP genotypes for all previously identified bats means that some individuals may be recorded in subsequent sampling events (2019 onwards) as new individuals rather than recaptures. Given that most ghost bats detected at South Flank are only recorded using the area once, estimated recapture and recruitment rates are likely to be minimally impacted by this change. This effect should disappear quickly over time as any resident bats are likely to be re-genotyped with SNP markers in following years and recapture rates can be estimated going forward with consistent SNP genotyping.

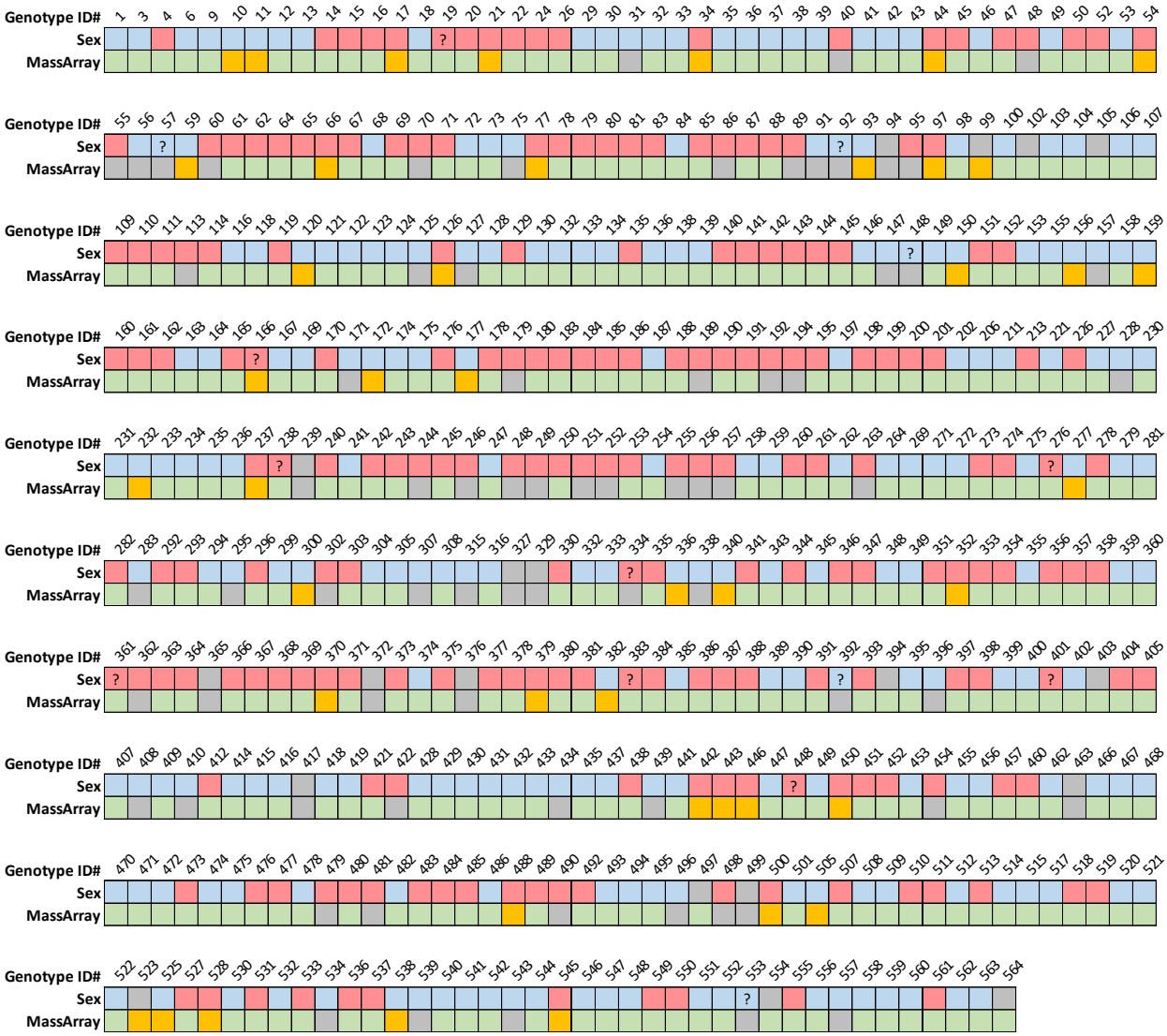


Figure 3 Detail of the unique bats identified through microsatellite analysis (Genotype ID#), their Sex and whether genotyped for MassArray SNP analysis. Molecular sexing was undertaken using the TaqMan probe sexing assay. Males and females are shown in blue and pink, respectively, while grey indicates failed amplification. '?' indicates potential sex where the molecular sexing result was unclear. Samples re-genotyped for SNP panels are shown in the bottom row. Grey, yellow and green colours represent samples with failed amplification, samples removed after quality control, and samples used in the analysis, respectively.

3.3 Conversion success from microsatellite to SNP markers

Given the potential increased resolution of SNP markers to identify unique individuals over microsatellite loci, we wished to evaluate the resultant reference SNP genotypes to confirm all individuals identified as unique through microsatellite analysis were also unique at SNP loci.

Firstly, we performed a preliminary comparative analysis of genotype clustering to identify individuals based on the full dataset of 117 loci compared to using a single panel of 50 loci (Panel 1). In genotyping studies (e.g. Sastre et al. 2023), as the number of genotyped loci increases the probability of genotyping errors also increases, making it more problematic to match and cluster sample genotypes. Thus, we compared different SNP clustering thresholds (allowing 2 to 20 SNP differences amongst samples within clusters) to identify individuals whereby scat samples are assigned to the same individual if they have less than the specified number of SNP differences. The comparison indicated that a threshold of 10 – 12 SNP differences for the 117-locus dataset produced genotype clustering equivalent to a threshold of 4 SNP differences on Panel 1. Subsequently, we decided to use a threshold of 10 SNP differences in clustering analyses when using the full dataset (117 loci) as it provided a comparable grouping of samples to Panel 1. Genotype clustering analysis was undertaken in our custom R package ScatMatch (see section 6), using the relaxed setting of 30% amplification threshold to both sample and locus amplification rates and assigning scat samples with 10 SNP differences or below to the same individual.

We found that 21 individuals considered unique genotypes in microsatellite analyses were clustered with other individuals in SNP markers (Table 1). For these samples, we manually compared microsatellite genotypes and checked sampling locations for each scat. We merged samples where differences in microsatellite genotypes could be the result of allelic dropout and if samples were collected in the same roosts at around the same time.

As a result of these analyses, we created a new database containing 308 SNP profiles (unique individuals) to act as our reference database for recapture analysis going forward.

Table 1 Previously identified genotype IDs by microsatellite markers that have been merged into the same individuals by SNP markers. Only the earliest IDs are kept (Maintained ID). Note that reports provided to BHP after 2021 will be impacted by these changes.

Maintained ID	Sex	Merged IDs	Affected reports
1	M	1,6	Ottewell et al. (2017)
29	M	29,269	Ottewell et al. (2017)
33	M	33,35	Ottewell et al. (2017)
79	F	79,85	Ottewell et al. (2017)
80	M	80,84	Ottewell et al. (2017)
81	F	81,88	Ottewell et al. (2017)
100	M	100,104	Ottewell et al. (2017)
107	M	107,116	Ottewell et al. (2017)
121	M	121,122	Ottewell et al. (2017)
135	F	135,136,142	Ottewell et al. (2017)
140	F	140,141	Ottewell et al. (2017)
161	F	161,191	Ottewell et al. (2017)
231	M	231,233	Ottewell et al. (2017)
274	F	274,452	Ottewell et al. (2020a)
275	M	275,279	Ottewell et al. (2017)
278	F	278,296	Ottewell et al. (2017)
348	M	348,349	Ottewell et al. (2017)
363	F	363,367	Ottewell et al. (2018)
366	F	366,371	Ottewell et al. (2018)
385	M	385,389	Ottewell et al. (2018)
398	F	398,403	Ottewell et al. (2019)

Ottewell, K., S. McArthur, S. V. Leeuwen, and M. Byrne. 2017. Population genetics of the Ghost Bat (*Macroderma gigas*) in the Pilbara bioregion. Final report prepared for Biologic Environmental Survey Pty Ltd., Department of Biodiversity, Conservation and Attractions, Kensington, Western Australia.

Ottewell, K., S. McArthur, S. V. Leeuwen, and M. Byrne. 2018. Cave use by the Ghost bat (*Macroderma gigas*) at the West Angelas mine site. Final report to Biologic Pty Ltd. Department of Biodiversity, Conservation and Attractions, Kensington, Western Australia.

Ottewell, K., S. McArthur, S. V. Leeuwen, and M. Byrne. 2019. Cave use by the Ghost Bat (*Macroderma gigas*) at the West Angelas mine site: Survey results October 2018. Department of Biodiversity, Conservation and Attractions, Kensington, Western Australia.

Ottewell, K., S. McArthur, S. V. Leeuwen, and M. Byrne. 2020a. Ghost bat (*Macroderma gigas*) genetic monitoring: South Flank 2017-2018. Final report to Biologic Environmental Survey. Department of Biodiversity, Conservation and Attractions, Kensington, Western Australia.

Ottewell, K., R. Thavornkanlapachai, and S. McArthur. 2021. Ghost bat (*Macroderma gigas*) genetic monitoring: South Flank 2019. Final report to Biologic Environmental Survey. Department of Biodiversity, Conservation and Attractions, Kensington, Western Australia.

4 Data processing from scats to genotypes

To evaluate the performance of the SNP markers in comparison to the microsatellite markers, we reanalysed the 2019 SNP data from the South Flank project (Ottewell et al. 2021). We selected this project because the area has been monitored over multiple years using microsatellites and provides a reliable base comparison to assess the transition to SNP genotyping. In this section, we investigated various aspects of the transition to SNP genotyping, including the influence of scat quality on bat detection, the methodologies employed to assign scats to individual bats, the exploration of different approaches to identify recaptures from the newly created SNP database, and lastly, a comparative analysis of reporting information generated from SNP data in contrast to that derived from microsatellite data.

4.1 Impact of scat quality on ghost bats detection

Scat quality has a significant impact on the detection of bats in the caves. The quality and quantity of DNA obtained from scats depend on environmental conditions and scat age. Exposure to sunlight, high temperature, and precipitation accelerate DNA degradation which subsequently decrease genotyping success and increase genotyping errors (Nsubuga et al. 2004, Piggott 2004, Berry et al. 2007, Carpenter and Dziminski 2017, Sittenthaler et al. 2020). Degradation of DNA typically occurs in the first 5 – 7 days post-deposition (Panasci et al. 2011, Skrbinšek 2020) but, in some species, it can occur at a slower rate, 2 – 3 weeks (Piggott 2004, Carpenter and Dziminski 2017). Here, we demonstrate the impact of scat quality on bat detection and how we modify our protocol to handle scats of differing quality.

For our analysis, we focused on the 2019 South Flank case study (Ottewell et al. 2021), which involved the collection of 712 samples primarily from caves situated within the core South Flank / Area C region (SF, AC, ACW, M1 caves), along with samples from the outlying OB35-02 and CATH-09 caves. Out of these, 697 samples were successfully DNA-extracted and subsequently genotyped with SNP Panel 1. To determine the genotyping error rate, a subset of 93 randomly selected samples (~13% of the total sample size) underwent genotyping twice. The average amplification rate of successfully genotyped samples including repeats was 93.1 ± 0.3% ranging from 54.5% to 100% (n = 782) indicating a high-quality batch of samples. The allelic dropout rate was low with an average of 1.0 ± 0.1% (n = 89) suggesting replicate samples' genotypes were nearly identical. All amplified samples and only one of two replicated samples were then grouped by our custom R package ScatMatch (Huntley 2021).

In ScatMatch, we used a combination of amplification rate (i.e., amplification success), genotyping error rate, and clustering analysis to assign scats to groups (or putative individuals). A good quality batch would have a high average amplification rate and low genotyping error rate as we observed in the 2019 South Flank project. Samples with low amplification rates are likely to be older and contain more errors (Figure 4). Thus, we routinely assess and adjust amplification thresholds to both samples and markers (loci) in order to eliminate poorly amplified samples or loci. By using only high-quality scats, it increases certainty in individual assignment. This can be demonstrated

from a greater separation of the frequency distribution of SNP differences between pairs of scats collected from the same individual (caused by genotyping errors) and scats from different individuals (caused by biological variation) (Figure 5a). The overlap in frequency distributions reduces with a higher amplification threshold (Figure 5a, blue arrows). The group (individual) assignment identified by the clustering analysis also becomes clearer when only samples with high amplification rates are used (Figure 5b). Once a suitable threshold for sample and locus amplification rates is selected, we used the hierarchical cluster analysis to assign scats to individuals based on the number of SNP differences or allele mismatches (h). Scats with the fewest SNP differences are grouped together and h determines the number of SNP differences allowed between groups. Selecting h becomes increasingly challenging as data quality decreases (Figure 5b). The consequences of incorrectly assigning h could be over-merging scats of genetically similar individuals or over-splitting scats that belong to the same individual caused by genotyping errors. This can subsequently impact mark recapture analysis if the wrong number of individuals are identified. Although the temptation exists to utilize all available sample genotypes and permit a higher h value for grouping scats with genotyping errors to their most likely individuals, such an approach must be avoided as it would compromise the reliability of the results. To demonstrate this, we varied amplification thresholds for both samples and loci from 50% to 70% to see how the rate of missing data affects the number of groups (or putative number of bats) detected.

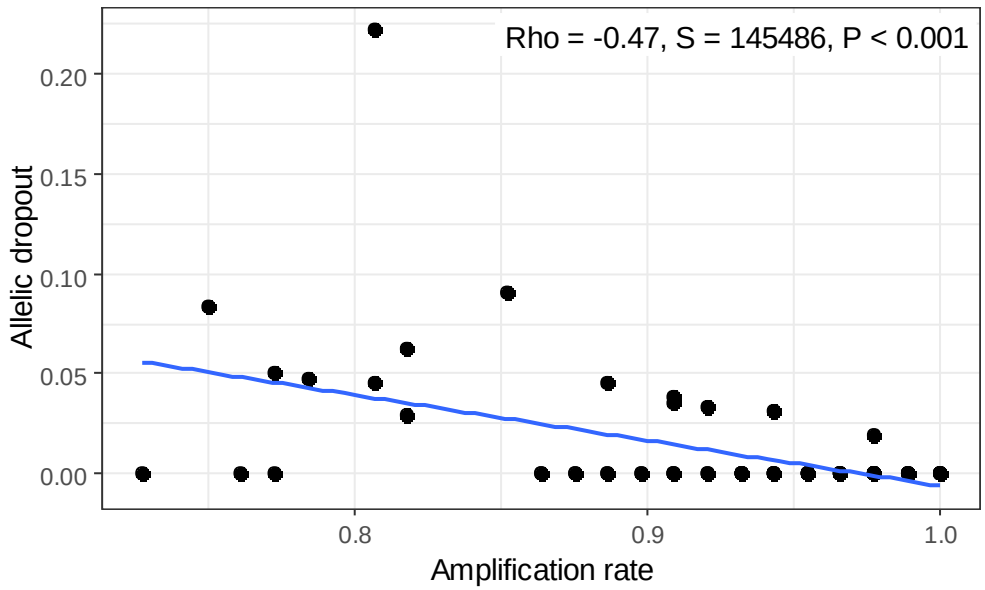


Figure 4 Comparison of the genotyping quality of South Flank 2019 project samples with increasing average amplification rate per sample. The average rate of allelic dropout was higher in samples with lower amplification rate.

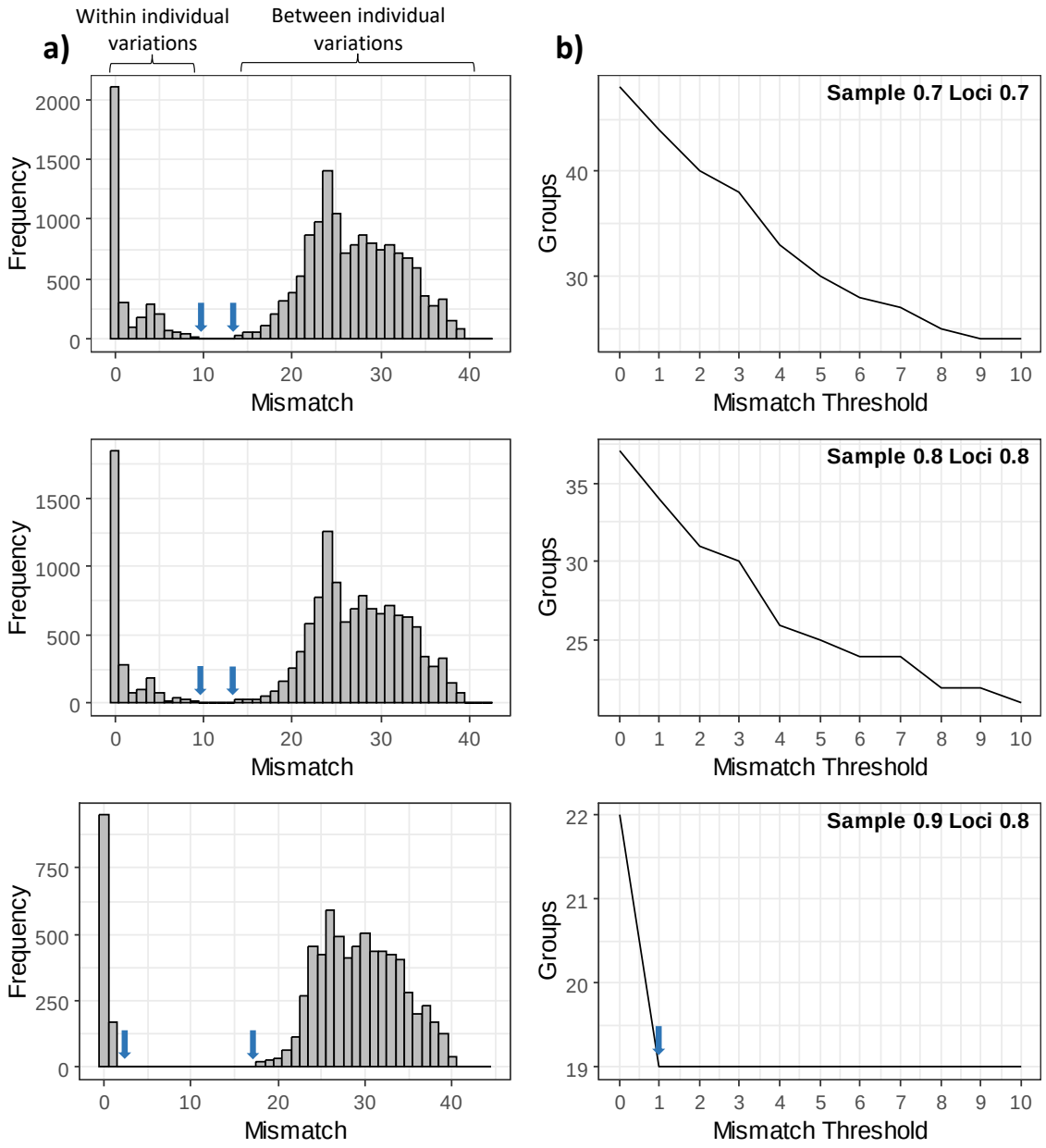


Figure 5 Scat and loci amplification rates (a) and threshold of SNP mismatch number to call scats from the same individual (b). The filtering used is indicated on the top right and the filter setting for the case study is shown in the last panels. The number of mismatches is calculated from pairwise comparison between scats with genotyping scores. Assuming the allelic frequency in a) follows a binomial distribution, allele mismatches forming the main peak on the right are likely to be the biological variations between individuals while allele mismatches on the left are likely to be variations between scat samples from the same individual. This figure is regenerated from Thavornkanlapachai et al. (2024) as it contains a larger variation of scat quality.

With lower amplification thresholds, we detected more bats. There are two possible reasons for this. First, by using a relaxed sample amplification threshold, we allow lower-quality scats to be included in the analysis. Lower quality scats are expected to be older (Carpenter and Dziminski 2017, Cowen et al. 2021) and may be more likely to be associated with transient individuals detected only once (fresh scats are likely to be biased towards resident bats). Second, older scats are found to have higher genotyping errors (Carpenter and Dziminski 2017, Cowen et al. 2021). This increases the number of SNP differences between scats from the same individuals and can be mistaken as separate individuals. To counter this problem, we generally increase h (number of SNP differences cutoff to call unique individuals) in low-quality batches of samples to accommodate for errors so scats can still be assigned correctly. For example, in Figure 5, the chosen h for the 90% sample amplification threshold is 1 while h for the 80% and 70% sample amplification threshold would be 4 and 7 respectively. As part of our analyses, we also compare individual assignments of scats at different amplification thresholds and h values to ensure consistency of assignment under different settings. However, to demonstrate the effect of the amplification threshold difference on the number of bats detected, we set h = 4 in all amplification thresholds. As shown in Figure 6 the assignment of scats among resident bats (~50 bats identified from 3 scats or more) were identical in all amplification thresholds. We found that a higher amplification threshold provides more reliable results, thus, a high amplification rate threshold (80%+) has been routinely utilised since 2019.

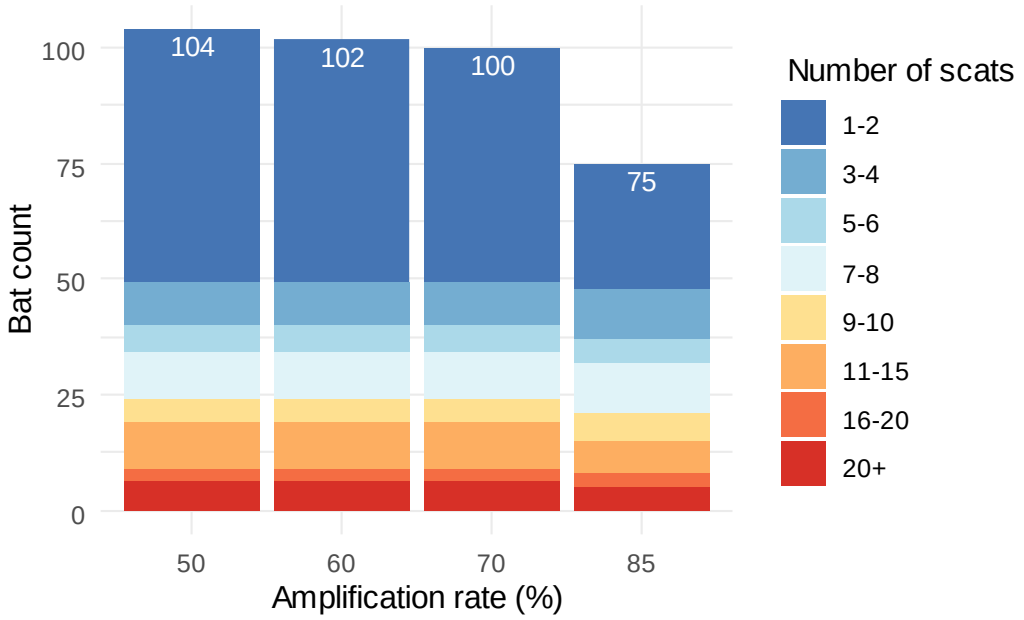


Figure 6 Effect of amplification rate and number of bats detected coloured by the number of scats assigned to each bat. We find that more individuals are detected from single scats when using lower (< 85%) amplification thresholds. Note that the sample amplification threshold rate of 85% was used in the Ottewell et al. (2021) report. The numbers on top of the bars are the total bat count.

4.2 Different approaches to identifying recaptures

Identifying recaptures between projects based on molecular tags can be challenging when genotyping errors contribute to differences observed between individuals. Retaining and matching only genotypes with complete information can exclude samples with partial matches that may yet be informative. Allowing a small amount of missing data in the analysis will increase sample size which is suggested to subsequently increase statistical power for mark-recapture analysis (Lukacs and Burnham 2005). In our routine reporting, DBCA typically retains the most complete genotype for each identified individual in the reference database. We considered that adding *all* raw genotypes (rather than retaining the most complete genotype) to the database could help impute the most likely alleles for missing loci resulting in a better detection of recaptures. Here, we investigated three different approaches using both identified unique genotypes from the refined SNP database and the raw genotypes to increase matching efficiency, and how these approaches affected clustering of individuals and identification of recaptures. We evaluate these approaches on detected bats identified from the South Flank 2019 project as follows:

1. Use the **database's unique genotypes** and identify recaptures from the **project's unique genotypes**
2. Use the **database's unique genotypes** and identify recaptures from the **project's raw scat genotypes**
3. Use the **database's raw genotypes** and identify recaptures from the **project's raw scat genotypes**

For Approach 1, the unique project genotypes were generated from the South Flank 2019 project using the same clustering threshold settings reported in Ottewell et al. (2021). Those are 85% sample, 80% loci amplification thresholds and SNP differences threshold (h) = 2 which is equivalent to h = 4 after manual assessment of potential allelic dropouts. Out of 595 usable sample genotypes, 75 individuals were identified. To identify which individuals were a recapture, these 75 genotypes were then matched with the SNP database which consists of one representative sample genotype for each genotype ID (from section 3.3). We used a 50% sample amplification threshold to accommodate for the lower overall amplification rate of samples in the SNP database (67.8%). This threshold results in a minimum of 22 loci available for matching which is above the recommended 20 loci needed to discriminate between related individuals by the P_{ID} analysis. We assigned samples with 4 SNP differences and below (h = 4) as recaptures and anything above as new individuals. For Approaches 2 and 3, we used raw unique scat genotypes from the South Flank 2019 project data (697 samples) and identified recaptured individuals from either the SNP database with only unique genotypes (Approach 2) or the raw scat genotypes collated for the SNP database (Approach 3). We used the same setting as Approach 1, that is, 50% amplification threshold for both sample and loci amplification rates and h = 4.

All three approaches generated identical groups of scat genotypes except that Approach 1 produced slightly fewer recaptures and fewer scats assigned to each bat (Appendix 3). Contradictory to the expectation, allowing scat samples with more

missing genotypes in the recapture analysis (Approach 2 and 3) did not result in more recaptures detected across times or caves. However, by having a lower project amplification threshold, we detected 2 additional recaptures (Genotype ID 128 – 2 scats 84.1% and 75% amplification rates and 276 – 1 scat 79.5% amplification rate, Appendix 3). Since using raw genotypes did not result in more recaptures detected, we considered it prudent to use only cleaned and unique genotypes in the recapture analysis.

4.3 Performance comparison between SNP and microsatellite markers

To investigate whether changing the type of genetic marker influences bat detection, we used results from Ottewell et al. (2021) and performed repeated-measures ANOVA on the number of unique genotypes detected per cave against various factors. First, we tested a full model with the number of bats per cave as a responding factor against the following fixed factors; cave ID, genotyping method, year of collection, number of scats collected, number of scats filtered, and frequency of cave visited (Table 2). We also ran a reduced model where non-significant variables were removed or assigned as random factors. We found that caves had the most influence on the number of bats detected inside the caves (Table 2, Figure 7). Roosts M-01, SF-05 and SF-08 consistently had more bats detected across the years (Figure 7). The method of genotyping had a smaller but significant impact on bat detection. The SNP method had a lower detection average of 4.18 bats/cave when compared to the microsatellite method 6.59 bats/cave (Welch two-sample t-test, $t = -2.117$, $df = 58.3$, $P = 0.039$). The differences came from the stringent amplification thresholds used on the SNP data (85% sample and 80% loci amplification rates) compared to the microsatellite data (2016 amplification rate ranged 36% – 100%, average 93.4 ± 0.4%; 2018 amplification rate ranged 63.6% – 100%, average 87.9 ± 0.5%). After relaxing the SNP amplification threshold to 60% for both loci and sample amplification rates, the mean differences in the numbers of bats detected per cave between different methods became smaller and not statistically different (SNP mean 5.64 bats/cave, microsatellite mean 6.59 bats/cave, Welch two-sample t-test, $t = -0.761$, $df = 57.8$, $P = 0.45$). In summary, the South Flank 2019 SNP data case study showed similar bat numbers and comparable levels of activities in multiple roosts to the previous surveys which employed microsatellite markers. Although converting to SNP markers may result in slightly fewer bats detected, we can be more assured that only high-quality scats were selected for the analysis and that genotypes generated contain fewer errors.

Table 2 Repeated-measures ANOVA to evaluate the effect of cave, genotyping methods (SNP versus microsatellite), year of collection, number of scats collected, filtered sample number, and frequency of cave visits on the number of bats detected.

	Df	Sum Sq	Mean Sq	F value	Pr (>F)	
Cave	21	915.9	43.6	2.6	0.006	**
Genotyping method	1	74.3	74.3	4.4	0.043	*
Year of collection	1	0.3	0.3	0.0	0.900	
No. of scat collected	1	23.5	23.5	1.4	0.244	
No. of filtered scat	1	3.3	3.3	0.2	0.659	
No. of visitation	1	11.7	11.7	0.7	0.410	
Residuals	34	569.3	16.7			

Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05

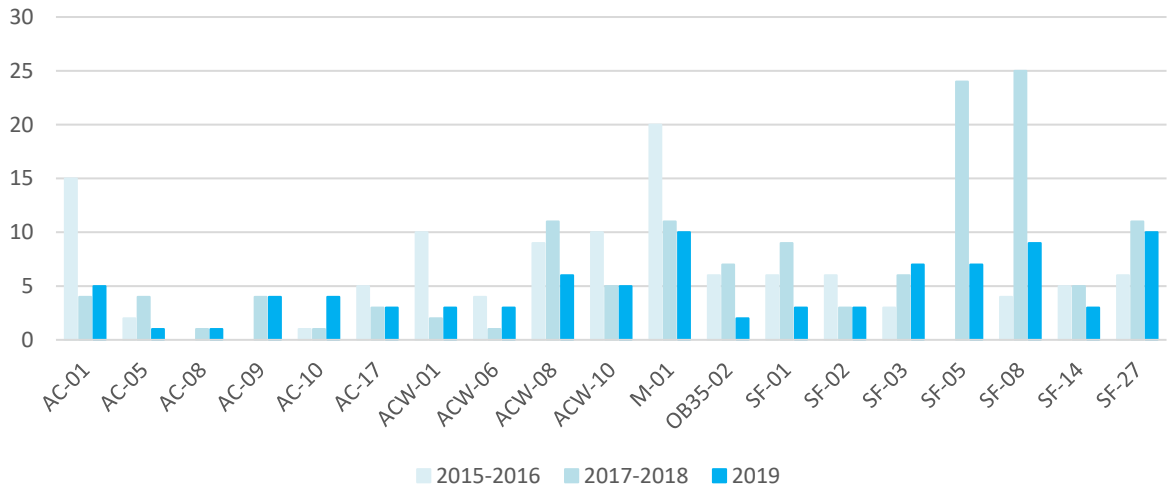


Figure 7 Comparison of the numbers of individuals detected in each cave over three sampling periods corresponding to data collected 2015 – 2019 (Ottewell et al. 2021). In the first two periods, scat samples were genotyped by 11 microsatellite markers and the last period was genotyped by 44 SNP markers.

Furthermore, the addition of sex information to genetic monitoring projects provides insight into ghost bat population dynamics, including assessment of putative maternity caves and estimates of sex-biased dispersal and habitat use (Prada et al. 2023). For example, with additional sexing information, we found that females commonly moved between SF caves, particularly SF-05, SF-08 and SF-27, which coincided with the high numbers of females detected in these caves (Figure 8).

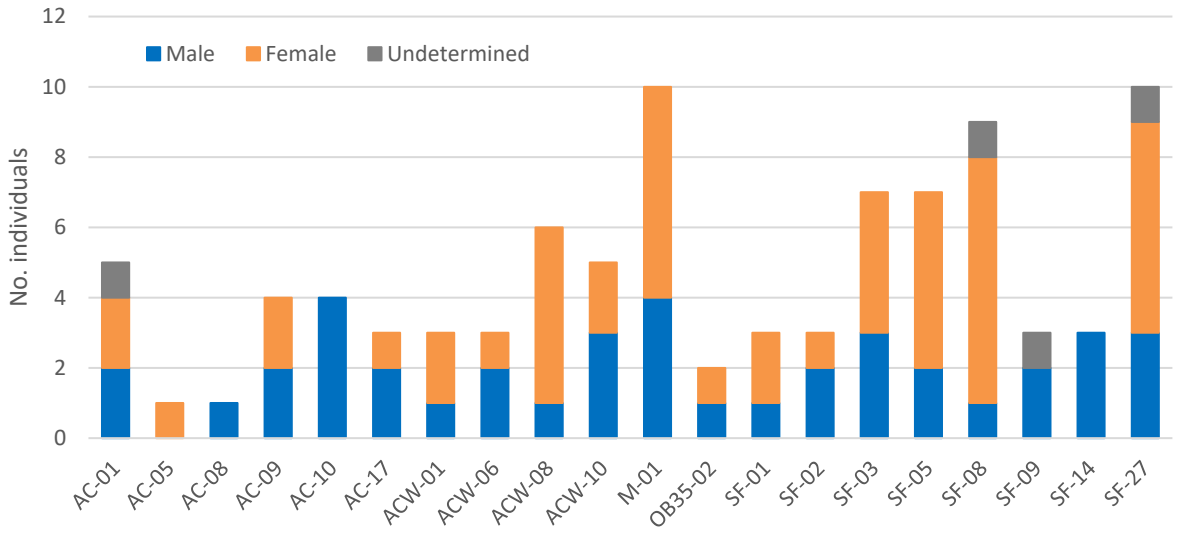


Figure 8 Numbers of male and female ghost bats detected in caves in the South Flank 2019 study area (Ottewell et al. 2021).

5 Refinement of SNP genotyping array

5.1 SNP panel modifications

The SNP panels underwent modification to increase efficiency and reduce cost by removing consistently failing markers, replacing them with high-performing markers from the additional SNP panels and adding sexing markers. The previous design only had 44 out of 50 markers consistently working in Panel 1, 33 out of 50 markers working in Panel 2 and 29 out of 47 markers working in Panel 3. To reduce the genotyping cost, we removed failed or poorly amplified markers and added four potential sexing markers (two Zfx/y and two SRY). Note that these sexing markers are different from the routinely used TaqMan probe sexing array as they target a single SNP difference rather than targeting the whole DNA fragment between males and females (from section 2.1). After the redesign to include better-performing markers, AGRF could only accommodate one sexing marker (Zfx/y) and the reorganisation of five autosomal markers from Panels 2 and 3 into Panel 1.

We have not further pursued the redesign of Panels 2 and 3 because we found that the cost of genotyping Panels 2 and 3 outweighed the benefit to genetic diversity estimation of genotyping more markers. We compared genetic diversity estimates obtained from increasing numbers of markers (50 – 200) in Pilbara populations (Thavornkanlapachai et al. 2024). We found that a larger number of markers only reduces the variance in genetic diversity estimates while the mean value is consistent and is only useful in the population genetic structure analysis if the populations are genetically isolated which is not the case for the Pilbara populations (Thavornkanlapachai et al. 2024). For full-scale population genetic studies, high-quality and quantity DNA such as tissue DNA is preferred (e.g. Sovic et al. 2016, Umbrello et al. 2022). New genomic approaches for a larger number of targeted SNP markers, DArTag (e.g. Arbon et al. 2021), may offer an opportunity to undertake large-scale population genetic studies on the ghost bat from scat samples. However, it's worth noting that this technology is relatively recent in development, and there is still uncertainty surrounding its performance in comparison to MassArray. Additionally, it comes at a higher cost per sample, with DArTag costing $18 as opposed to $11 for MassArray. For the purpose of individual identification, our current targeted SNP genotyping with one MassArray SNP panel is sufficient to identify and monitor bats through via non-invasive sampling.

5.2 Upgraded SNP panel performance

The upgraded panel was tested for its consistency and efficiency by re-genotyping high-quality scats from 15 ghost bat individuals.

All samples were successfully genotyped with an average amplification rate of 97% (ranging from 88% – 100%) compared to 94% (ranging from 89% – 100%) from the previous panel version. The amplification rates of individual markers were similar across the two panel versions. From all samples with genotypes (611 pairwise comparisons – excluding genotype pairs that one or both samples failed), only two were mismatched as a result of allelic dropout in one panel or another. Unfortunately, the sexing marker (Zfx/y#1) did not show a consistent result with only half of the genotyped individuals (7/15) showing the same sexing result as determined with the Taqman probe sexing assay.

We recommend using the modified panel version in future projects but continue identifying sex with the TaqMan probe sexing assay to ensure reliability.

6 Refinement of data processing

6.1 R custom script to R package

To determine the number of unique individuals from microsatellite markers, we previously used the software COLONY (Jones and Wang 2010) to cluster identical scat genotypes. Data processing from filtering and assignment requires manual handling to ensure correct assignment which can be time-consuming. To improve the process, we developed R scripts to handle SNP data which filter samples by amplification rate, convert SNP to numeric data, assign scats to individuals, and provide visual aids to guide the user's decision of which threshold to assign scats to individuals and output data in the format that is ready for downstream analyses.

Last year, we released an R package which is an upgrade of our customed R scripts (Huntley 2021). In addition to the functions described above, we added data quality assessment and visualisation, finer-scale data filtering for both samples and loci, ability to process microsatellite data, additional assessment of scat assignment thresholds, and a map function to generate collecting locations and number of unique individuals assigned to each site. The R package reduces data processing time and removes manual handling so the results will be consistent between users given that they use the same setting. We also developed a website with tutorials to guide a new user through the package (Figure 9).

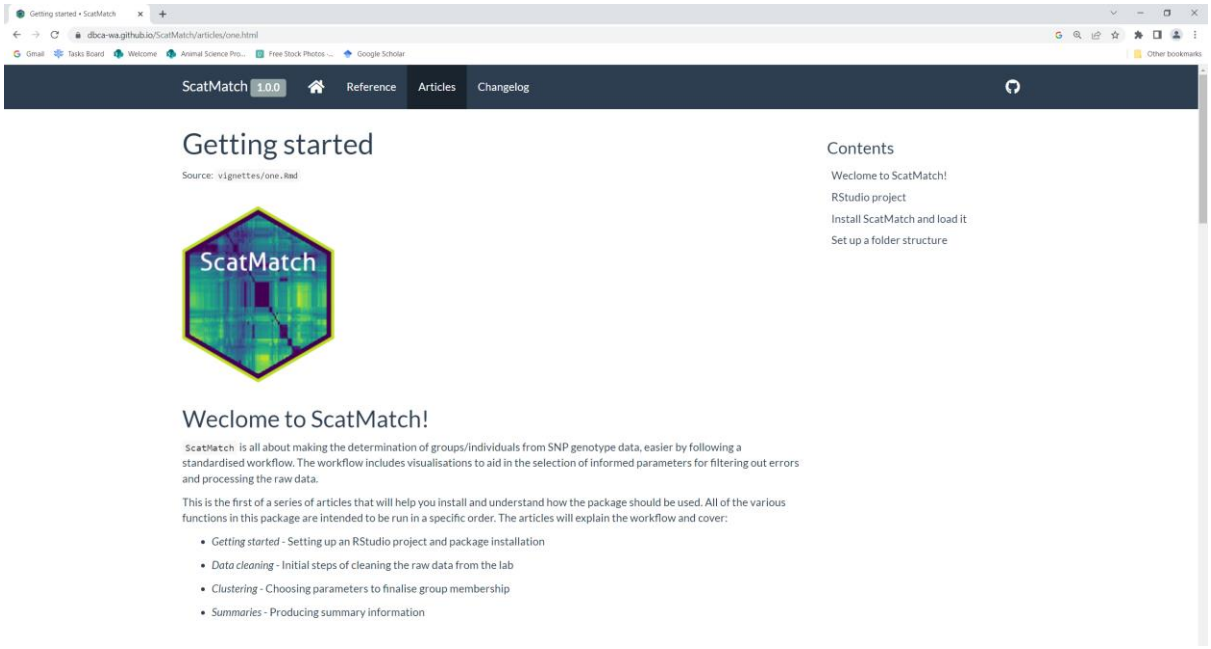


Figure 9 ScatMatch website. https://dbca-wa.github.io/ScatMatch/articles/one.html

7 Conclusion and future directions

The gap in ghost bat dispersal literature highlights the need for regular long-term monitoring (Cramer et al. 2022). Molecular tagging using scats offers an alternative monitoring method to live capture with minimal interference to individuals. Early molecular tagging studies of ghost bats were accomplished using microsatellite markers. While they have proven to be informative, there are limitations associated with low-quality and/or poor-quality scat DNA. To improve accuracy and efficiency, we developed custom Single Nucleotide Polymorphism (SNP) array panels using the MassArray automated genotyping technology. In addition, we streamlined our molecular sexing assay using Taqman probes. Combined we have a streamlined system to enable efficient identification of individuals and their sex from a single scat.

We were able to successfully transfer 74% of microsatellite genotypes to SNP genotypes, although the conversion rate was lower than expected due to an issue with EDTA inhibition of the multiplex MassArray reaction from the QIAamp® Fast DNA Stool Mini DNA extraction kit (Appendix 1; Schrader et al. 2012). The remaining 26% of genotypes, if recaptured, will be identified as new individuals in the SNP data. As the majority of unsampled genotypes were from single scats (most likely transitory animals), the effect on mark-recapture rates is likely to be minimal but should be considered.

The ability to identify recaptures and new individuals in the South Flank 2019 dataset (Ottewell et al. 2021) showed that the SNP array technology worked effectively and the identification of bats from scats using SNP markers was successful. A comparable number of bats was detected in each cave compared to the previous surveys that were based on microsatellite analysis. The patterns of activity of recaptured bats was also consistent. With additional sexing information that was not available at the time of the previous report, we were able to show levels of cave usage by different sexes.

There are advantages and limitations to both SNP and sexing arrays. The SNP array panels may have an advantage over microsatellite markers in that they consist of a larger number of pre-selected informative markers. Thus, a few failed markers are less likely to impact individual identification when compared to microsatellite markers which rely on high levels of allelic variation per marker, therefore, failure of microsatellite markers will reduce statistical power to assign individuals confidently. While it is tempting to use samples with low amplification rates to increase sample sizes (Lukacs and Burnham 2005), retaining DNA samples with a high fail rate has its limitations. In the SNP array study of von Thaden et al. (2017), poorer quality DNA had a lower amplification rate and higher genotyping errors. Genotyping errors increase the perceived variation between scats which could lead to over-estimating the number of individuals present and affect subsequent Mark Recapture analyses. Therefore, using samples with a high amplification rate is recommended, but determination of a suitable amplification threshold needs to be assessed on a project-by-project basis based on the observed genotyping error rate to retain enough samples to identify new individuals and recaptures. In the case of a project with poor scat quality, using a higher number of SNP mismatches will allow alleles caused by genotyping errors to

be assigned correctly, but care must be taken for the risk of merging scats from genetically similar or related individuals.

In addition, we found that sexing markers are more sensitive to poor quality/quantity DNA because our target genes are located on only one chromosome, providing only one copy for PCR amplification. Berry et al. (2007) also found that sexing markers amplification rate and accuracy declined with scat age. Allelic dropout and failed amplification are common among scats older than three weeks (Berry et al. 2007). Therefore, to enhance the confidence of sex assessment, we recommend employing a consensus sex determination method using multiple scat samples and limiting the analysis to those exhibiting strong amplification signals.

Since the development of SNP array panels for the ghost bat, we have undertaken multiple refinements to our sexing and SNP genotyping workflows to improve the performance and quality of data processing. First, we refine the sexing markers from basic PCR markers to TaqMan probes to increase efficiency and reduce processing costs. Second, we refined the SNP panel by removing failed markers and replacing them with working markers from other panels. As only one panel is necessary for individual identification, we focussed on upgrading only the first panel. Now we have a fully functional panel of 49 autosomal markers instead of 44 markers from the first panel design. Third, we have an R package to increase the efficiency of data handling and quality assessment/filter which allows flexibility and consistency in data handling.

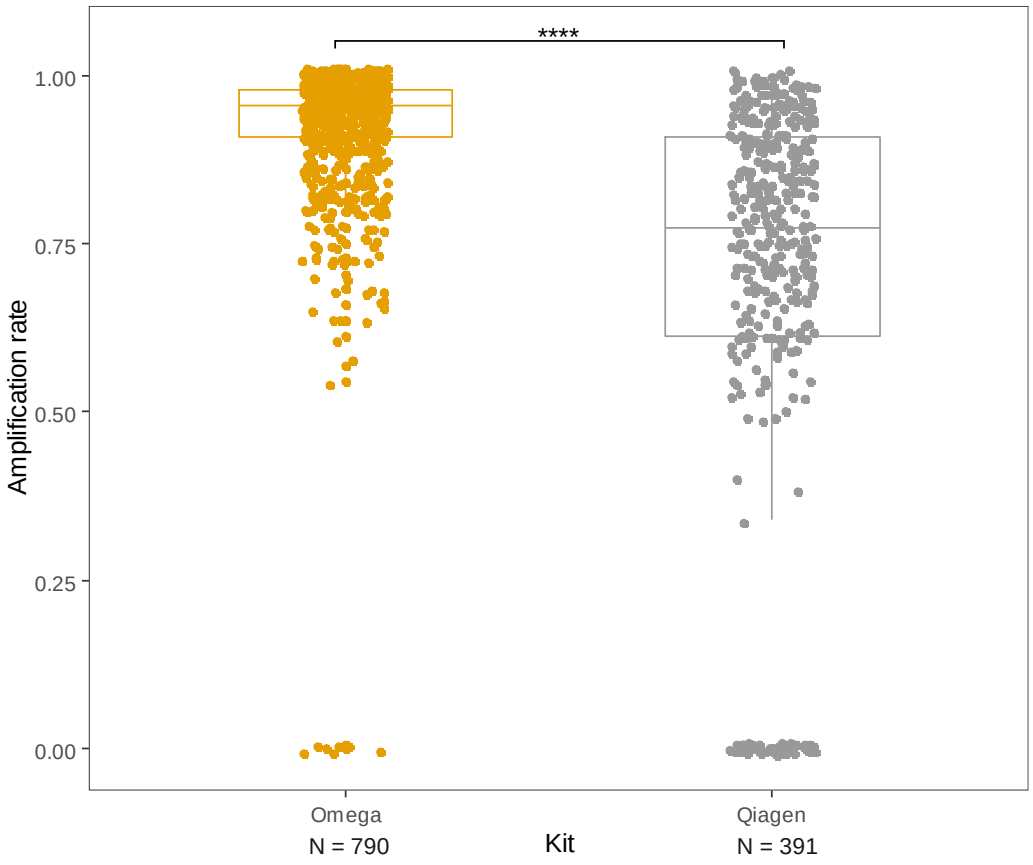
There are multiple technologies that could be developed to fill in research priorities for the ghost bat discussed in Cramer et al. (2022). First, the development of molecular ageing markers, if feasible from scats, holds the potential to significantly enhance Capture-Mark-Recapture analysis. This advancement could lead to improved recruitment and survival estimates (e.g. Jarman et al. 2015, Wright et al. 2018). Second, seasonal diet analysis will offer insight into food sources critical to the survival of ghost bats and insect or plant communities that need to be protected in addition to their roosts (Claramunt et al. 2018). Lastly, Cramer et al. (2022) pointed out the importance of a public online database which collates the trapping and monitoring of ghost bats to make this information accessible to regulatory agencies, proponents and others for impact assessment. Finding or developing a platform to host the DBCA-held ghost bat monitoring database could help encourage collaborations between stakeholders, private and public sectors and fill the gap in knowledge of ghost bat biology.

In conclusion, this report showed a successful transition from microsatellites to the SNP array for genotyping of ghost bat faecal samples. With the addition of sexing information, sex-specific habitat usage can be obtained for each project. The refinement of data processing led to improved efficiency, reproducibility and accuracy in identifying unique individuals. The transition provides a stable and robust approach to genetic monitoring of ghost bats, but with continuing developments in genomic technologies mean continuous improvement may lead to further refinements in the future, particularly if molecular ageing techniques can be developed for the species. This exciting prospect highlights the potential for even more precise and effective genetic monitoring methods on the horizon.

Appendices

Appendix 1 Comparison of amplification rate for different DNA extraction kits

*Boxplots of ghost bat faecal sample amplification rates of the identification panel (44 loci) extracted with the Omega Biotek Mag-Bind Stool DNA kit (South Flank 2019) and QIAamp® Fast DNA Stool Mini kit (Database). One dot represents one sample. Pairwise Wilcoxon test was performed as denoted above the boxplots as * P < 0.05, ** P < 0.01, *** P < 0.001, and **** P < 0.0001.*



Appendix 2 Details of microsatellite genotype IDs not converted to SNP IDs

Table A2.1 Details of microsatellite GenotypeIDs detected in South Flank roosts for which SNP genotypes could not be obtained, including information on the individual's roost location and years detected.

Genotype ID	Roost/s	Year/s	Region
10	ACW-01	2016	South Flank
11	ACW-01	2016	South Flank
17	ACW-06	2016	South Flank
21	ACW-08	2016	South Flank
31	ACW-08	2015	South Flank
34	ACW-10	2016, 2017	South Flank
40	ACW-10	2016	South Flank
86	AreaC-01	2016	South Flank
89	AreaC-01	2016	South Flank
91	AreaC-01	2016	South Flank
92	AreaC-01	2016	South Flank
93	AreaC-01	2016	South Flank
94	AreaC-01	2016	South Flank
95	AreaC-01	2016	South Flank
97	AreaC-03	2016	South Flank
99	AreaC-03	2016	South Flank
113	AreaC-08	2016	South Flank
120	AreaC-13	2016	South Flank
125	AreaC-17	2016	South Flank
126	AreaC-17	2016	South Flank
127	AreaC-17	2016	South Flank
277	M-01	2016	South Flank
283	M-01	2015, 2016	South Flank
295	M-01	2016	South Flank
300	Marillana-12/MARXX1	2015, 2016, 2018	South Flank
302	MARXX1	2015	South Flank
315	SF-01	2016	South Flank
327	SF-01	2015	South Flank
329	SF-01	2015	South Flank
334	SF-02	2016	South Flank
336	SF-02	2016	South Flank
338	SF-02	2015	South Flank
340	SF-03/SF-05/SF-14/SF-27	2016, 2018	South Flank
352	SF-15	2016	South Flank
408	SF-08/L3	2017, 2018	South Flank/West Angelas
410	SF-08	2017	South Flank

439	ACW-06	2018	South Flank
443	ACW-07	2017	South Flank
446	M-01	2018	South Flank
450	SF-14	2018	South Flank
454	SF-22	2017	South Flank
463	SF-05	2018	South Flank
488	M-01	2018	South Flank
490	M-01	2018	South Flank
496	SF-27	2018	South Flank
498	SF-05	2018	South Flank
500	SF-05/SF-08	2018	South Flank
505	SF-08/SF-27	2017	South Flank
523	SF-05/SF-08	2018	South Flank
525	SF-08	2017	South Flank
528	SF-27	2018	South Flank
534	SF-08	2017	South Flank
538	SF-05/SF-08	2017, 2018	South Flank
539	M-01	2018	South Flank
545	M-01	2018	South Flank

Table A2.2 Details of microsatellite GenotypeIDs detected outside of South Flank roosts for which SNP genotypes could not be obtained, including information on the individual's roost location for BHP-funded projects only and years detected.

Genotype ID	Roost/s	Year/s	Region
228	GU-02	2016	Eastern Hamersley
232	GU-03	2016	Eastern Hamersley
237	JIN-14	2016	Eastern Hamersley
239	K-01	2016	Eastern Hamersley
244	K-01	2016	Eastern Hamersley
246	K-01	2016	Eastern Hamersley
248	K-01	2016	Eastern Hamersley
249	K-01	2016	Eastern Hamersley
251	K-01	2016	Eastern Hamersley
252	K-01	2016	Eastern Hamersley
307	NT-03	2016	Eastern Hamersley
417	BHP0B35-02	2018	Eastern Hamersley
422	KOO04	2017	Eastern Hamersley
499	BHP0B35-02	2017	Eastern Hamersley
147	Call-01	2016	Chichester
148	Call-01	2016	Chichester
150	Call-01	2016	Chichester
156	Call-01	2016	Chichester
157	Call-01	2016	Chichester
159	Call-01	2016	Chichester
166	Call-01	2016	Chichester
171	Call-03	2017	Chichester
172	CathedralGorge-06	2016	Chichester
177	CattleGorge-02	2016	Chichester
179	CattleGorgeCulvert-01	2016	Chichester
189	CattleGorgeCulvert-06	2016	Chichester
192	CattleGorgeCulvert-07	2016, 2017	Chichester
194	CattleGorgeCulvert-08	2016	Chichester
255	KlondykeQueen	2017	Chichester
256	KlondykeQueen	2017	Chichester
257	KlondykeQueen	2017	Chichester
263	LallaRoohk	2017	Chichester
434	W007b	2018	Chichester
442	W007b	2018	Chichester
479	KOO04	2017	Chichester
481	KOO04	2017	Chichester
543	W007b	2018	Chichester

Appendix 3 Recaptured bats' scat counts identified from different approaches to match project genotypes to the SNP database

Each column represents different approaches of matching individuals from the South Flank 2019 project to the database. Approach 1: Matching between unique genotypes from the database and project. Approach 2: Matching between unique genotypes from the database and raw scat genotypes from the project. Approach 3: Matching between raw scat genotypes from both database and project.

Genotype ID	Scat count	
	Approach 1	Approach 2 & 3
4	2	3
24	9	11
26	2	2
80_84	20	22
81	2	2
83	12	12
118	11	13
128[a]	-	2
133_492	14	15
272	11	11
276[a]	-	1
293	1	1
308_618	8	8
343	7	7
346	8	8
358	12	14
412	1	1
416_665[b]	6	9
428	4	6
430	2	2
431	1	1
437	72	79
447	1	1
453	1	1
460	4	4
467	8	11
473	1	1
489	9	11
497	1	1
508	9	9
515	4	4
518	7	7

519	3	4
520	1	2
527	12	13
530	1	1
531_659[b]	13	16
415_532	10	10

[a] Not reported in Ottewell et al. (2021) due to the amplification rates of these scats below the threshold of 85%

[b] Upon reanalysing, we identified Genotype ID 665 and 659 as recaptures.

[d] Genotype ID 533 was reported as a recapture in Ottewell et al. (2021), but identified as a new individual in this report because Genotype ID 533 in the database has an amplification rate below 50% and was removed from the database.

References

Arbon, P. M., C. N. S. Silva, D. B. Jones, D. Jaccoud, M. Gervis, D. R. Jerry, and J. M. Strugnell. 2021. Development and validation of a SNP-based genotyping tool for pedigree establishment in Australian greenlip abalone *Haliotis laevigata* Donovan, 1808. Aquaculture Reports **20**:100746.

Armstrong, K. N., and S. D. Anstee. 2000. The ghost bat in the Pilbara: 100 years on. Australian Mammalogy **22**:93-101.

Armstrong, K. N., J. C. Z. Woinarski, N. M. Hanrahan, and A. A. Burbidge. 2021. *Macroderma gigas* (amended version of 2019 assessment). The IUCN Red List of Threatened Species 2021: e.T12590A209530568, https://dx.doi.org/10.2305/IUCN.UK.2021-3.RLTS.T12590A209530568.en.

Augusteyn, J., J. Hughes, G. Armstrong, K. Real, and C. Pacioni. 2018. Tracking and tracing central Queensland's *Macroderma* – determining the size of the Mount Etna ghost bat population and potential threats. Australian Mammalogy **40**:243-253.

Berry, O., S. D. Sarre, L. Farrington, and N. Aitken. 2007. Faecal DNA detection of invasive species: the case of feral foxes in Tasmania. Wildlife Research **34**:1-7.

Bullen, R., S. Reiffer, and J. Trainer. 2023. Satellite tracking Ghost Bats (*Macroderma gigas*) in the Pilbara, Western Australia. Records of the Western Australian Museum **1**:001-010.

Carpenter, F. M., and M. A. Dziminski. 2017. Breaking down scats: degradation of DNA from greater bilby (*Macrotis lagotis*) faecal pellets. Australian Mammalogy **39**:197-204.

Claramunt, A. M. A., N. E. White, M. Bunce, M. O'Connell, R. D. Bullen, and P. R. Mawson. 2018. Determination of the diet of the ghost bat (*Macroderma gigas*) in the Pilbara region of Western Australia from dried prey remains and DNA metabarcoding. Australian Journal of Zoology **66**:195-200.

Cowen, S., M. Smith, S. McArthur, K. Rayner, C. Jackson, G. Anderson, and K. Ottewell. 2021. Novel microsatellites and investigation of faecal DNA as a non-invasive population monitoring tool for the banded hare-wallaby (*Lagostrophus fasciatus*). Australian Journal of Zoology **69**:55-66.

Cramer, V. A., K. N. Armstrong, R. D. Bullen, S. L. Cross, L. Gibson, N. Hanrahan, C. G. Knuckey, K. Ottewell, S. Reiffer, L. Ruykys, R. E. Shaw, R. Thavornkanlapachai, S. A. Thompson, S. Wild, and S. van Leeuwen. 2022. Research priorities for the ghost bat (*Macroderma gigas*) in the Pilbara region of Western Australia. Australian Mammalogy **45**:1-12.

Ekblom, R., M. Aronsson, F. Elsner-Gearing, M. Johansson, T. Fountain, and J. Persson. 2021. Sample identification and pedigree reconstruction in Wolverine (*Gulo gulo*) using SNP genotyping of non-invasive samples. Conservation Genetics Resources **13**:261-274.

Frantz, A. C., L. C. Pope, P. J. Carpenter, T. J. Roper, G. J. Wilson, R. J. Delahay, and T. Burke. 2003. Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. Molecular Ecology **12**:1649-1661.

Hall, L., G. C. Richards, N. McKenzie, and N. Dunlop. 1997. The importance of abandoned mines as habitat for bats. Pages 326-334 *in* P. Hale and D. Lamb, editors. Conservation Outside Nature Reserves. Centre for Conservation Biology, University of Queensland, Brisband, Australia.

Hoyle, S. D., A. R. Pople, and G. J. Toop. 2001. Mark–recapture may reveal more about ecology than about population trends: Demography of a threatened ghost bat (*Macroderma gigas*) population. Austral Ecology **26**:80-92.

Huntley, B. 2021. ScatMatch: Functions to aid error filtering from SNP genotype data and determine group majorities. R package version 1.0.0.

Jarman, S. N., A. M. Polanowski, C. E. Faux, J. Robbins, R. De Paoli-Iseppi, M. Bravington, and B. E. Deagle. 2015. Molecular biomarkers for chronological age in animal ecology. Molecular Ecology **24**:4826-4847.

Jones, O. R., and J. Wang. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources **10**:551-555.

Kunz, T., and M. B. Fenton. 2005. Bat Ecology. University of Chicago Press, United Kingdom.

Lukacs, P. M., and K. P. Burnham. 2005. Review of capture–recapture methods applicable to noninvasive genetic sampling. Molecular Ecology **14**:3909-3919.

Nsubuga, A. M., M. M. Robbins, A. D. Roeder, P. A. Morin, C. Boesch, and L. Vigilant. 2004. Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. Molecular Ecology **13**:2089-2094.

Ottewell, K., S. McArthur, S. V. Leeuwen, and M. Byrne. 2017. Population genetics of the Ghost Bat (*Macroderma gigas*) in the Pilbara bioregion. Final report prepared for Biologic Environmental Survey Pty Ltd., Department of Biodiversity, Conservation and Attractions, Kensington, Western Australia.

Ottewell, K., S. McArthur, S. V. Leeuwen, and M. Byrne. 2018. Cave use by the Ghost bat (*Macroderma gigas*) at the West Angelas mine site. Final report to Biologic Pty Ltd. Department of Biodiversity, Conservation and Attractions, Kensington, Western Australia.

Ottewell, K., S. McArthur, S. V. Leeuwen, and M. Byrne. 2019. Cave use by the Ghost Bat (*Macroderma gigas*) at the West Angelas mine site: Survey results October 2018. Department of Biodiversity, Conservation and Attractions, Kensington, Western Australia.

Ottewell, K., S. McArthur, S. V. Leeuwen, and M. Byrne. 2020a. Ghost bat (*Macroderma gigas*) genetic monitoring: South Flank 2017-2018. Final report to Biologic Environmental Survey. Department of Biodiversity, Conservation and Attractions, Kensington, Western Australia.

Ottewell, K., R. Thavornkanlapachai, and S. McArthur. 2021. Ghost bat (*Macroderma gigas*) genetic monitoring: South Flank 2019. Final report to Biologic Environmental Survey. Department of Biodiversity, Conservation and Attractions, Kensington, Western Australia.

Ottewell, K., R. Thavornkanlapachai, S. McArthur, P. Spencer, J. Tedeschi, B. Durrant, C. Knuckey, K. Armstrong, and M. Byrne. 2020b. Development and optimisation of molecular assays for microsatellite genotyping and molecular sexing of non-invasive samples of the ghost bat, *Macroderma gigas*. Molecular Biology Reports **47**:5635-5641.

Panasci, M., W. B. Ballard, S. Breck, D. Rodriguez, L. D. Densmore, D. B. Wester, and R. J. Baker. 2011. Evaluation of fecal DNA preservation techniques and effects of sample age and diet on genotyping success. Journal of Wildlife Management **75**:1616-1624.

Piggott, M. P. 2004. Effect of sample age and season of collection on the reliability of microsatellite genotyping of faecal DNA. Wildlife Research **31**:485-493.

Prada, D., K. Ottewell, C. Lohr, and L. Gibson. 2023. Spatial and temporal analysis of ghost bat populations in the Pilbara using non-invasive sampling methodologies: Towards a robust genetic monitoring protocol. Phase 1: Capture-Recapture methods. Final report to BHP. Department of Biodiversity, Conservation and Attractions.

Sastre, N., A. Mercadé, and J. Casellas. 2023. SNP+ to predict dropout rates in SNP arrays. Conservation Genetics Resources **15**:113-116.

Schrader, C., A. Schielke, L. Ellerbroek, and R. Johne. 2012. PCR inhibitors – occurrence, properties and removal. Journal of Applied Microbiology **113**:1014-1026.

Sittenthaler, M., E. M. Schöll, C. Leeb, E. Haring, R. Parz-Gollner, and K. Hackländer. 2020. Factors influencing genotyping success and genotyping error rate of Eurasian otter (*Lutra lutra*) faeces collected in temperate Central Europe. European Journal of Wildlife Research **67**:2.

Skrbinšek, T. 2020. Effects of different environmental and sampling variables on the genotyping success in field-collected scat samples. Acta biologica slovenica **63**:89-98.

Sovic, M. G., B. C. Carstens, and H. L. Gibbs. 2016. Genetic diversity in migratory bats: Results from RADseq data for three tree bat species at an Ohio windfarm. PeerJ **4**:e1647.

Spencer, P. B. S., and J. Tedeschi. 2016. An initial investigation into the genetic diversity, structure and short-range spatial-use by ghost bat in the Hamersley subregion of the Pilbara. Murdoch University, Perth.

Taberlet, P., S. Griffin, B. Goossens, S. Questiau, V. Manceau, N. Escaravage, L. P. Waits, and J. Bouvet. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic acids research **24**:3189-3194.

Thavornkanlapachai, R., K. N. Armstrong, C. Knuckey, B. Huntley, N. Hanrahan, and K. Ottewell. 2024. Species-specific SNP arrays for non-invasive genetic monitoring of a vulnerable bat. Scientific Reports **14**:1847.

Threatened Species Scientific Committee. 2016. Conservation advice *Macroderma gigas* ghost bat. Department of the Environment and Energy, http://www.environment.gov.au/biodiversity/threatened/species/pubs/174-conservation-advice-05052016.pdf.

Toop, G. J. 1985. Habitat requirements, survival strategies and ecology of the ghost bat, *Macroderma gigas*, Dobson (Microchiroptera, Megadermatidae) in central coastal Queensland. Macroderma **1**:37-41.

Umbrello, L., R. Bullen, R. Shaw, S. McArthur, M. Byrne, S. van Leeuwen, and K. Ottewell. 2022. Extensive gene flow in a threatened bat (*Rhinonicteris aurantia*) in an arid landscape. Global Ecology and Conservation **37**:e02154.

von Thaden, A., B. Cocchiararo, A. Jarausch, H. Jüngling, A. Karamanlidis, A. Tiesmeyer, C. Nowak, and V. Muñoz-Fuentes. 2017. Assessing SNP genotyping of noninvasively collected wildlife samples using microfluidic arrays. Sci Rep **7**:10768-10768.

Woinarski, J. C. Z., A. A. Burbidge, and P. L. Harrison. 2014. The action plan for Australian mammals 2012. CSIRO Publishing, Collingwood.

Worthington Wilmer, J., C. Moritz, L. Hall, J. Toop, and J. D. Pettigrew. 1994. Extreme population structuring in the threatened ghost bat, *Macroderma gigas*: Evidence from mitochondrial DNA. Proceedings of the Royal Society of London. Series B: Biological Sciences **257**:193-198.

Wright, P. G. R., F. Mathews, H. Schofield, C. Morris, J. Burrage, A. Smith, E. L. Dempster, and P. B. Hamilton. 2018. Application of a novel molecular method to age free-living wild Bechstein's bats. Molecular Ecology Resources **18**:1374-1380.