DNA extraction and amplification of regurgitate and faecal samples from Buller's Shearwater (*Puffinis bulleri*)

Report

By

Doyle, E. [1], & Adams, N. [1]

[1]Unitec Institute of Technology, Private Bag 92025, Victoria Street West, Auckland, New Zealand.

APPLIED MOLECULAR SOLUTIONS

Executive Summary

Research which aims to identify the prey items found in the diet of seabirds by analysing stomach contents, faeces, and other remains strongly biased towards detecting organisms which are slow to digest, while soft bodied prey may be fully digested and remain undetected. DNA-based methods of prey item identification have been used in a number of studies to address this issue. The intention of the current research was to test the feasibility of applying DNA-based prey item identification to analysing the diet of the Buller's shearwater (*Puffinis bulleri*). DNA was extracted from 17 samples of regurgitate and faecal material collected from Buller's shearwaters, received from the Northern New Zealand Seabird Trust. Following DNA extraction, PCR amplification is used to selectively amplify short sections of the 16S gene which can be used for identification. PCR products were run on an agarose gel, as a success/fail test for amplification. Only one of the DNA samples tested could be amplified. While this result suggests that DNA-based analysis may not be suitable for research on the diet of the Buller's shearwater, the outcome may be due to the small scale of the test.

Introduction

Analysis of faeces and stomach contents allows researchers to determine the diet of an organism. In some cases, identifying the contents of regurgitate or faecal samples may be possible though morphological methods, examining hard materials which survive digestion such as bones, otoltihs, and squid beaks (Barrett et al., 2007). However, these hard items may be difficult to identify to the species level, while softer tissues are often fully digested and so underrepresented (Barrett et al., 2007; Casper et al., 2007; Dunshea, 2009). DNA-based methods of sample identification have been used to aid in identifying the contents of these challenging samples in a number of studies (e.g. Deagle et al., 2007; Casper et al., 2007; Bowser et al., 2013).

The methods used for DNA extraction vary between studies, and are typically chosen with particular consideration for the type of material in the sample. In the case of partially or fully digested samples, in which only small amounts of highly degraded DNA may be extracted (Deagle et al., 2006) or the sample is small, the final volume of fluid in which the DNA is suspended in may be reduced in order to increase the concentration of the DNA extract (e.g. Bowser et al, 2007). Following extraction, a section of the DNA is amplified using the polymerase chain reaction (PCR). PCR primers are used to control which section of the genome is amplified in the reaction, such as the 16S gene, which is found in all species but sufficiently variable that it can be used for the identification of unknown specimens. This gene is found on the mitochondrial genome and multiple copies are present in each cell (Dunshea, 2009), increasing the chances of recovering it from a sample compared to a gene in the nuclear DNA of which there is only one copy per cell. This may be of particular importance when working with mixed and degraded samples. Short sections of the gene can be used for species level identification. PCR reactions which target short regions of the genome are preferred when working with potentially degraded DNA, as is expected with faecal material and regurgitate (Dunshea, 2009; Bowser et al., 2013). While it is possible to target particular species, and exclude the amplification of the predator's own DNA, using prey species specific PCR primers (Barrett et al., 2007), 'universal' primers which are capable for amplifying a range of taxa may be of particular use (Dunshea, 2009) as these allow for a wide variety of prey DNA sequences to be amplified and may allow for unexpected prey items to be found.

A recent study carried out by Unitec researchers (unpublished) was able to obtain amplifiable DNA from regurgitate and faecal samples collected from adult Australasian gannets (*Morus serrator*) in the Hauraki Gulf, and identify the contents of these samples using next generation sequencing. The intention here, was to test the efficacy of the methods used in the gannet study using samples collected from Buller's Shearwaters. The initial steps of DNA extraction and amplification were undertaken, with PCRs targeting a short section of the 16S gene which would be used if these samples were to be sequenced on the next generation sequencing Illumina platform.

Method

DNA extraction & quantification

DNA extraction methods were chosen to suit the varying compositions of the regurgitate samples. DNA was extracted from samples which consisted of identifiable bone and fish tissue using the Qiagen DNeasy® Blood and Tissue DNA extraction kit, as per the manufacturer's instructions. This approach was consistent with that used in the gannet study, in which pieces of prey tissue were dissected out of each regurgitate collection and sequenced as separate samples. The remaining regurgitate samples, which were not made up of large pieces of undigested material, were first gravity filtered through a sterile glass fibre filter membrane in order to separate the sample from the preserving ethanol. This approach was used to ensure any small pieces of organic material were not lost along with the ethanol, as may have occurred had the ethanol been tipped out of the sample. DNA was extracted from the material collected in the filters using the Qiagen PowerWater® DNA isolation kit, following the manufacturer's recommendations, with a final elution volume 50µL.

DNA from faecal samples were extracted using the Qiagen QIAamp® DNA Stool kit following the manufacturer's specifications, with a final elution volume of100µL. Unlike the gannet research, in which there was limited faecal material available to be used in each extraction, the Buller's shearwater faecal samples were fairly large. The larger sample size is beneficial, however samples which are collected fresh from the bird as opposed to collected from around nesting sites, may be less degraded.

Following DNA extraction, the total nucleotide concentrations of all the samples were estimated using a spectrophotometer. A spectrophotometrer gives an indication of the concentration of nucleotides in a sample by measuring the light which it absorbs. However, this is not a direct measurement of the DNA itself and does not give an indication of DNA degradation. There is no minimum nucleotide concentration needed for PCR as it is a highly sensitive reaction, which allows for samples with nucleotide concentrations too low to be detected by spectrophotometer to be successfully amplified.

DNA amplification

DNA amplification by PCR is required prior to sequencing. Following the protocols used in the gannet study, DNA samples were PCR amplified in four separate reactions using primers which target the 16S region, one designed to amplify the DNA of Chordata and the other Malacostraca. By using two primers, a wider range of prey species could be amplified.

In two reactions, a 155bp section of the 16S gene in Chordata was amplified using the primers Chord_16S_F_TagA (5'- ATG CGA GAA GAC CCT RTG GAG CT) and Chord_16S_R_Short (5'- CCT NGG TCG CCC CAA C) (Deagle et al., 2009). In one reaction, DNA samples were amplified in 20µL PCR reactions containing of 10µL GoTaq® PCR Master Mix (Promega), 0.8µL the forward primer (10µM), 0.8µl of the reverse primer

(10µM), 1.0 µL BSA, 0.2 µL MgCl (25mM), 3µL of DNA template, and 4.2µl of MQ water. A negative control without DNA, in which 3µL of MQ water was substituted for the template DNA, was included in each PCR. A positive control using 3µL of whitebait (*Galaxis sp.*) was also included. A second reaction was carried out in which the volume of DNA template was increased to 4µL and volume of MQ water was reduced to 3.2µL to keep the total volume of 20µL consistent across all reactions. Two further PCR reactions were carried out using the primers Mala_16S1F (5'-TGA CGA TAA GAC CCT) and Mala_16S2R (5'- CGC TGT TAT CCC TAA AGT AAC T) (Deagle et al., 2005), which target a 200bp section of the 16S gene in Malacostraca. These PCR reactions used the same method as those using Chordata primers, with arrow squid (*Nototodorus gouldi*) DNA used as the positive control.

With all four reactions, thermal cycling was conducted in a Surecycler 8800 (Agilent Technologies). An initial denaturation period of 15 minutes at 94°C was followed by 33 cycles of denaturation for 20 seconds at 94°C, annealing for 90 seconds at 48.7°C, and extension for 45 seconds at 72°C, followed by a final extension for 2 minutes at 72°C. PCR products were run on a 2% agarose E-Gel™ (Invitrogen™) and visualised in an UVIDOC HD6 Touch (UVITEC Cambridge), to determine the amplification success of each of the samples.

Testing for PCR inhibitors

Substances which inhibit PCR, such as salts, may be co-extracted with DNA. If these inhibitors are present in high concentrations they can prevent the sample from being amplified. To test for the presence of PCR inhibitors in a DNA sample which has failed to amplify, DNA from a sample which had previously been successfully amplified is added to the reaction mix, in addition to the DNA of interest. If there are high concentrations of inhibitors in the DNA sample being tested they will prevent PCR amplification of both the DNAs incorporated in the reaction.

To test for the presence of PCR inhibitors, two reactions were carried out, one using the Chordata primers and one using the Malacostraca primers. DNA samples were amplified in 20µL PCR reactions containing 10µL GoTaq® PCR Master Mix (Promega), 0.8µL the forward primer (10µM), 0.8µl of the reverse primer (10µM), 1.0 µL BSA, 0.2 µL MgCl (25mM), 1.2µl of MQ water, and 3µL of whitebait DNA in each reaction in addition to 3µL of the DNA template being tested. A negative control without DNA, in which 6µL of MQ water was substituted for the DNA, was included in each PCR. A positive control using a total of 6µL of whitebait DNA was also included. With both reactions, thermal cycling was conducted in a Surecycler 8800 (Agilent Technologies). An initial denaturation period of 15 minutes at 94°C was followed by 33 cycles of denaturation for 20 seconds at 94°C, annealing for 90 seconds at 48.7°C, and extension for 45 seconds at 72°C, followed by a final extension for 2 minutes at 72°C. PCR products were run on a 2% agarose E-Gel™ (Invitrogen™) and visualised in an UVIDOC HD6 Touch (UVITEC Cambridge), to determine the amplification success of each of the samples.

Results

The mean nucleotide concentration of the DNA samples extracted from regurgitate was 8.6 ng/µL, while the mean for the samples extracted from faeces was slightly higher, at 10.2 ng/µL (Figure 1).

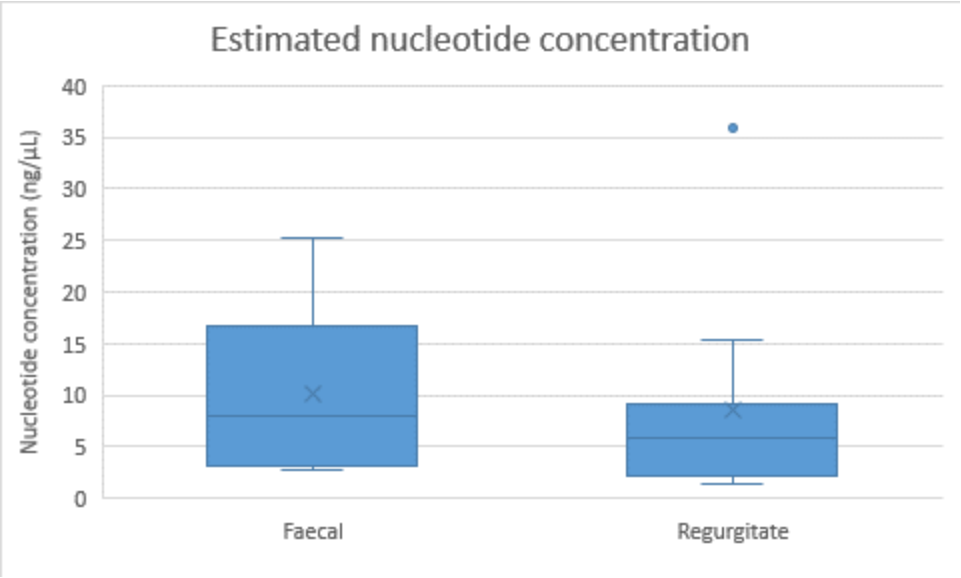

Figure 1: Spectrophotometer readings of DNA samples.

Both the PCR reactions using 3µL and 4µL of template DNA failed to amplify as expected for any of the regurgitate or faecal samples tested, with the exception of the regurgitated fish sample which was less digested than the other samples. The different primers used in the reactions, targeting different taxa, produced similarly negative results when visualised under UV light. However, each of the samples were successfully PCR amplified when additional DNA from whitebait was added into each reaction in order to test for the presence of inhibitors.

Discussion

The mean nucleotide concentrations of 8.6 ng/µL and 10.2 ng/µL for the regurgitate and faecal samples, respectively, suggest that the extraction methods used were sufficient to isolate DNA from each of the samples. The higher concentrations found in the faecal samples were likely due to the presence of contaminating non-target DNA from the birds' gut flora. Theoretically, the estimated nucleotide concentrations of both the faecal and regurgitate samples was sufficiently high for PCR amplification to work as PCR is an extremely sensitive reaction, capable of amplifying DNA from samples with nucleotide concentrations too low to be detected by spectrophotometry (Doyle, 2015).

The length of time between when a bird feeds and when regurgitate and faecal samples are collected may be a barrier to the success of this project, as it is expected that more highly digested food material would contain less amplifiable DNA. Here, a relatively short section of 100-200 bases was chosen for amplification as it was expected that the DNA in the samples would be fragmented, making it less likely for large sections of the genome to be present in the samples. Similarly, increases made to the volume of DNA template used in the PCR reactions were intended to improve the likelihood of successful amplification. The successful amplification seen in the samples which had whitebait DNA added to the PCR mix indicates that any inhibitors present in the DNA extracted from the regurgitate and faecal samples were at too low a concentration to prevent amplification. This indicates that the negative outcomes of the original PCR amplifications were due to the DNA being too highly degraded by digestion for them to be successfully amplified. DNA extracted from the regurgitate sample which consisted of a large piece of undigested fish was successfully amplified. This type of sample, were digestion has been minimal and large amounts of sample can be used in the DNA extraction, appears to be more suitable for genetic analysis. However, these results are based a small number of samples; seven faecal samples, including faecal matter scrapped from the surface of a dried eggshell, and ten regurgitate samples (including the fish and bone samples sourced from regurgitate). In contrast, the gannet study which this project was modelled after included regurgitate samples from 141 birds, 97 (86.8%) of which could be PCR amplified using the Mala primers, and 81 (57.4%) using the Chordata primers. These primers were also used to amplify DNA from 40 faecal samples, with only 5 (12.5%) successfully amplifying using the Mala primers, while 34 (85%) were amplified using the Chordata primers. This rate of amplification failure is consistent with that found in the literature, for example, in analysing the diet of Macaroni Penguins (*Eudyptes chrysolophus*), Deagle et al. (2007), found that less than half of the faecal samples tested could be PCR amplified.

While this study was not successful in amplifying DNA from the majority of the samples and suggests that DNA-based analyse may not be suited to the study of the Buller's Shearwater's diet, a high failure rate is common in the literature. It is possible that increasing the number of samples used may have had an improved the success rate.

References

Alonso, H., Granadeiro, J. P., Waap, S., Xavier, J., Symondson, W. O., Ramos, J. A., & Catry, P. (2014). An holistic ecological analysis of the diet of Cory's shearwaters using prey morphological characters and DNA barcoding. Molecular ecology, 23(15), 3719-3733.

Barret, R. T., Camphuysen, K., Anker-Nilssen, T., Chardine, J.W., Furness, R.W., Garthe, S., … Viet, R.R. (2007). Diet studies of seabirds: a review and recommendations. *ICES Journal if marine Science, 64*, 1675-1691.

Bowser, A. K., Diamond, A. W., & Addison, J. A. (2013). From puffins to plankton: a DNA-based analysis of a seabird food chain in the northern Gulf of Maine. PLoS One, 8(12), e83152.

Casper, R. M., Jarman, S. N., Gales, N. J., & Hindell, M. A. (2007). Combining DNA and morphological analyse of faecal samples improves insight into trophic interactions: a case study using a generalist predator. *Marine Biology, 152*, 815-825.

Deagle, B. E., Jarman, S. N., Pemberton, D., Gales, N. J. (2005). Genetic screening for prey in the gut contents from a giant squid (Architeuthis sp.). *Journal of Heredity 96*(4), 417-423.

Deagle, B. E., Eveson, J. P., & Jarman, S. N. (2006). Quantification of damage in DNA recovered from highly degraded samples- a case study on DNA in faeces. *Frontiers in Zoology*, 3(11).

Deagle, B. E., Kirkwood, R., & Jarman, S. N. (2009). Analysis of Australian fur seal diet by pyrosequencing prey DNA in feces. *Molecular Ecology 18*, 2022-2038.

Deagle, B. E., Gales, N. J., Evans, K., Jarman, S. N., Robinson, S., Trebilco, R., & Hindell, M. A. (2007). Studying seabird diet through genetic analysis of faeces: a case study on macaroni penguins (Eudyptes chrysolophus). PLoS One, 2(9), e831.

Doyle, E. J. (2015). *Assessing the use of molecular techniques to determine terrestrial invertebrate diversity in New Zealand mangroves and detection of blue ducks in New Zealand streams*. (Unpublished master's thesis). University of Waikato, Hamilton, New Zealand.

Dunshea, G. (2009). DNA-based diet analysis for any predator. *PLoS ONE, 4*(4).