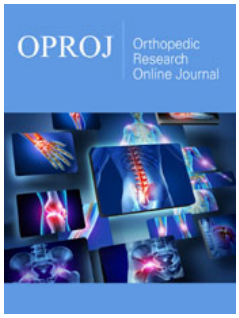
Crimson Publishers
Wings to the Research

Opinion

Using DNA for Estimating Animal Population Sizes

ISSN: 2576-8875



George A F Seber*

Emeritus Professor of Statistics, University of Auckland, New Zealand

***Corresponding author**: George A F Seber, Emeritus Professor of Statistics, University of Auckland, Science Building, 38 Princes Street, Auckland 1010, New Zealand

Submission: May 13, 2022

Published: May 23, 2022

Volume 9 - Issue 4

How to cite this article: George A F Seber. Using DNA for Estimating Animal Population Sizes. Ortho Res Online J. 9(4). OPROJ. 000717. 2022.
DOI: 10.31031/OPROJ.2022.09.000717

Opinion

With the rapid advances in DNA technology, it is not surprising that DNA methods are being used for the estimation of the sizes of animal populations and numbers of species, particularly with on-sight field methods. Also related to this is the estimation of biodiversity, a popular topic. Animal population sizes are a good litmus test as to how climate change and changing habitats through human intervention are affecting population sizes. We are losing a large number of species every year, as well as having to deal with invasive species, whether animals or plants. Freshwater ecosystems are among the most endangered habitats on Earth, with thousands of animal species known to be threatened or already extinct. We therefore need to keep an eye on endangered species and to check on population changes.

The method involves using DNA that is left behind on several occasions as a unique identifier for an animal of a given species. Each time the animal's DNA turns up, we have effectively "recaptured" it without having to capture and initially tag and recapture it. This can have an ongoing effect on the animal. The method is referred to statistically as a non-invasive Capture-Recapture (CR) method. If DNA samples turn up several times, we have a recapture sequence for each animal. If this process is carried out for a population using random selection methods, we can then build a statistical model to use the captures and recaptures to estimate the numbers not caught, and therefore estimate the population size. This simple idea has led to a huge number of extensions and very complex mathematical models to take care of a variety of factors such as sex, age, breeding status, death, birth, and migration. One can use information from dead animals as well as from live ones. Computer methods and packages are extensively used and are continuously being developed. CR methods have also been used for human populations as well such as the estimation of the number of diabetics, or for an elusive population the number of drug addicts.

To introduce a little mathematics, suppose we have N animals of which M are initially marked or tagged, (or DNA taken) and a random sample of size n yields m tagged (DNA recognized). Because the sample is random so that every individual has the same chance of being selected, the sample proportion m/n of tagged recaptures will be a suitable estimate of the population proportion M/N. Then N can be estimated by nM/m. This simple idea generally underlies the CR method, but with possible complex variations. For example, suppose there is mortality between the first and the recapture sample with P the proportion surviving, which is assumed to be the same for tagged and not tagged. Then the population proportion marked before the second sample will be MP/NP=M/N as before, so that our estimate of N remains the same. However, we still need to estimate P, which requires further data obtained from further recaptures involving such unknowns as the probability of capture each time. Im addition, not all animals may be equicatchable, in other words the probability of capture is heterogeneous

and varies from individual to individual. This demonstrates how the model can readily become more complex.

The type of molecular marker generally used for individual identification is microsatellite DNA, which has an intrinsically high mutation rate due to strand slippage during replication while undergoing meiosis. Nucleotide polymorphisms are also used. The idea is to genotype each sample at multiple molecular loci, and matching genotypes are assumed to be recaptures while non-matching are new captures. Genetic methods can be costly, not only for the time spent, but also the cost due to information loss when discarding samples that contain some degree of uncertainty in their identification.

Let us consider a couple of simple DNA examples. Bears like to scratch on a scratching post such as a tree trunk with barbed wire and leave behind fur, which can then be analysed for DNA. There is a variety of methods of collecting DNA such as using hair, saliva, feces, egg shells, sloughed skin, and feathers. Pellet droppings are frequently used for larger species such as, for example, wolves, badgers, elephants, and turtles, among others. This can sometimes be augmented with the presence of automatic cameras (called camera trapping). With fisheries, DNA fragments can be filtered from the water to determine species' presence. This DNA is usually referred to as eDNA or environmental DNA as it is extracted from an environmental sample without isolating the target organism. It has been characterized as a mixture of genomic DNA from many different organisms, which is often degraded into small fragments. Short DNA fragments allow the use of degraded DNA from environmental samples. It has been noted that because eDNA persists in water for days to weeks after organisms are removed from controlled experimental systems, it contains a catalog of species present in the recent past. Clearly there is an optimum number of samples that need to take to allow for species present but not detected, and for species misidentified.

There are some problems with DNA methods. Firstly, DNA can be easily degraded by environmental factors, primarily moisture and UV irradiation, so it is important that a proper protocol is followed, and samples from the field are correctly handled and stored as soon as soon as possible to prevent degradation. Secondly, using too few or insufficiently variable loci results in different individuals seeming to be the same individual, which results in the underestimation of population size; a phenomenon known as the "shadow effect". This is a source of heterogeneity where some animals not captured previously are believed to be recaptures due to their DNA profile being an indistinguishable shadow of previously captured animals. This problem can arise when only a subset of the genome is examined. Thirdly, genotyping errors can result in specimens from the same individual seeming to have different molecular tags, leading to an overestimation of population size. Contamination from other individuals can occur.

The problem of genotyping errors is more pronounced in studies using the non-invasive sources of DNA because of the low quantity of DNA these sources sometimes provide. The solution to the shadow effect is to use many highly variable loci to reduce the probability that individuals will share a molecular tag. However, the more loci used the higher the probability a genotyping error of some sort will be made. Also, there can be sequencing errors that can arise through "amplification failure" (the most severe problem with non-invasive sampling), when there is a copying failure of a piece of the genome, allelic dropout when one allele at a locus fails to "amplify" or is not present in the pipetted DNA sample, mutations during amplification, and individual heterogeneity such as variability in cell-shedding rates. Statistical methods for estimating genotypic error are available.

Genetic information has been utilized not just from "recapture of self" but also from the recapture of closely related kin. An individual is marked by its presence in the sample, and "recaptured" if the sample contains one or more close relatives. A topic still developing.

As genetic techniques improve, DNA will be used more and more, which will add further to the subject of estimating animal numbers [1], which is split into "open populations" where there are population changes such as mortality. migration etc. [2] and closed populations which do not have changes in population size during the sampling period such as plants, trees, and some animals. A book on this is over 80% written and has over 1000 references so far. The large number of references indicates how quickly the subject is expanding.

References

1. Seber GAF (1982) The estimation of animal abundance, (2nd edn), Reprinted in paperback by the Blackburn press, Caldwell, NJ, Griffin, London, UK.
2. Seber GAF, Schofield MR (2019) Capture-recapture: Parameter estimation for open animal populations. Springer Nature, Switzerland.