# Supplementary file S2: optimal sampling design using a power analysis on pilot data

Power analysis is a component of a hypothesis test that quantifies the probability that the test will reject a false null hypothesis (Ho), so that power (*pwr*) is the chance of not making a Type II error (*β*) (*pwr* = 1- *β*) (Cohen 2013). For this study, it is the probability that the tests used will detect a significant difference between the captive and wild status groups. A power analysis involves the relationship between four statistical parameters: power (*pwr*), significance criterion (*α*), sample size (*n*) and effect size (*ES*) (Nakagawa and Cuthill 2007). Therefore, the sample size n can be determined by setting the remaining parameters, which was the purpose of this analysis.

The power analysis was performed using “simr” (Green and MacLeod 2016). Linear models were created using values and variances from a pilot study measuring five turtles, which allowed for the creation of simulated datasets. Nitrogen was set as the dependent variable, with status as a fixed effect and individual and scute as random effects. Effect size was set to smallest individual means between individuals from different groups (i.e. smallest captive and largest wild values), and final parameters were set as: *α* = 0.05 and *pwr* = 0.8. The variance component of the effect size was calculated using a repeatability measure (using R package “rptr”), which has the advantage of being able to identify errors surrounding the variance component (Stoffel et al. 2017).

Sample sizes were manually set before each simulation, varying the number of individuals and scutes in each run. Simulated datasets were created a set number of times (in this study, 1000 times), and the final power result was the number of simulated datasets which showed a significant difference between wild and captive groups (*α* = 0.05)

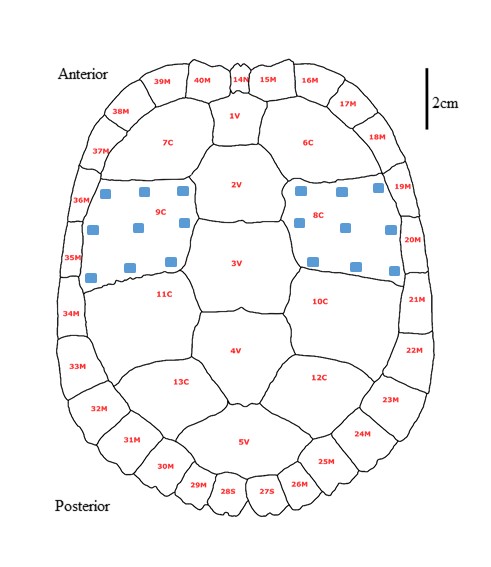
*Results*

Power simulations indicated that adding more scutes to simulations did not consistently improve power values; however, adding more individuals significantly improves power (Table S2). Variation between scutes was minimal when compared to variation between individuals. While all power values were over 80%, adding more individuals will account for potentially greater variation after more samples are added to the database.

**Table S2.1:** Total power (%) calculated using simulated datasets with increasing individuals and scutes calculated with simulations.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | Individuals per group | | | | |
| 2 | 3 | 4 | 5 | 6 |
| Scutes | 2 | 84.40 ± 2.40% | 90.50 ± 1.75% | 96.30 ± 1.08% | 98.90 ± 0.86% | 99.40 ± 0.70% |
| 3 | 82.10 ± 2.45% | 93.00 ± 1.50% | 96.50 ± 1.05% | 99.20 ± 0.77% | 99.70 ± 0.57% |
| 4 | 80.00 ± 2.44% | 92.30 ± 1.58% | 96.20 ± 1.38% | 98.60 ± 0.94% | 99.80 ± 0.52% |
| 5 | 100.0 ± 0.37% | 90.90 ± 1.71% | 96.30 ± 1.36% | 98.60 ± 0.94% | 99.90 ± 0.46% |

The determined sampling method included focusing sampling across all nine available individuals, rather than increased sampling on few individuals (Table S1). Furthermore, sampling two costal scutes (8C and 9C) and nine samples per scute (focused around the perimeter) was adequate to capture most potential variation within individuals (Figure S1).



**Figure S1:** Sampling design decided after power analysis on pilot data. All available individuals were sampled, concentrating on two scutes and sampling around the perimeter. Scutes 8C and 9C were selected as they were easy to sample.

# Literature Cited

Cohen J (1977) Chapter 1: the concepts of power analysis. In: Cohen J (Ed) Statistical power analysis for the behavioral sciences. Academic Press, 1-17. doi: 10.1016/B978-0-12-179060-8.50006-2

Green P, MacLeod CJ (2016) SIMR: an R package for power analysis of generalized linear mixed models by simulation. Methods in Ecology and Evolution 7: 493-498. doi:10.1111/2041-210X.12504

Nakagawa S, Cuthill IC (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists. Biological reviews 82: 591-605. doi:10.1111/j.1469-185X.2007.00027.x

Stoffel MA, Nakagawa S, Schielzeth H (2017) rptR: Repeatability estimation and variance decomposition by generalized linear mixed‐effects models. Methods in Ecology and Evolution 8: 1639-1644. doi:10.1111/2041-210X.12797