Fishing for DNA? Designing baits for population genetics in target enrichment experiments: guidelines, considerations and the new tool supeRbaits

Belen Jimenez Mena¹, Hugo Flávio¹, Romina Henriques¹, Alice Manuzzi¹, Miguel Ramos², Dorte Meldrup¹, Janette Edson³, Snaebjorn Palsson⁴, Gubjörg Ásta Ólafsdóttir⁴, Jennifer Ovenden³, and Einar Nielsen¹

¹Technical University of Denmark ²University of Porto ³The University of Queensland ⁴University of Iceland

July 12, 2021

Abstract

Targeted sequencing is an increasingly popular Next Generation Sequencing (NGS) approach for studying populations, through focusing sequencing efforts on specific parts of the genome of a species of interest. Methodologies and tools for designing targeted baits are scarce but in high demand. Here, we present specific guidelines and considerations for designing capture sequencing experiments for population genetics for both neutral genomic regions and regions subject to selection. We describe the bait design process for three diverse fish species: Atlantic salmon, Atlantic cod and tiger shark, which was carried out in our research group, and provide an evaluation of the performance of our approach across both historical and modern samples. The workflow used for designing these three bait sets has been implemented in the R-package supeRbaits, which encompass our considerations and guidelines for bait design to benefit researchers and practitioners. The supeRbaits R package is user-friendly and versatile. It is written in C++ and implemented in R. supeRbaits and its manual are available from Github: https://github.com/BelenJM/supeRbaits

Hosted file

baitDesignPaper_SUBMITTED.docx available at https://authorea.com/users/425076/articles/529987-fishing-for-dna-designing-baits-for-population-genetics-in-target-enrichment-experiments-guidelines-considerations-and-the-new-tool-superbaits