

PHD COURSE IN LIFE AND ENVIRONMENTAL SCIENCES

Report Form for PhD student annual evaluation (XXXVI and XXXVII cycles)

Name of PhD student: Francesco Giannelli

Title of PhD research: Adaptation genomics of Lessepsian invaders.

Name of PhD supervisor: Emiliano Trucchi, Emanuela Fanelli

Research lab name: Laboratorio di Genetica e Genomica

Cycle:

XXXVI

XXXVII

PhD Curriculum::

Marine biology and ecology

Biomolecular Sciences

Civil and environmental protection

DISVA instrumentation labs/infrastructure eventually involved in the project:

Actea Mobile Laboratory

Advanced Instrumentation lab

Aquarium

MassSpec lab

MaSBiC

Simulation/informatics lab

Other.

ABSTRACT (1000 characters, including spaces):

In recent years increasing human activities are threatening the biodiversity of the Mediterranean Sea, making it one of the most impacted seas in the world. The spread of invasive species, from the Red Sea into the Mediterranean, (Lessepsian alien species) is one important driver of biodiversity loss, as invasive species are negatively interfering with native species. In this project we want to study the impact of adaptive and deleterious genetic variation of successful Lessepsian species in the invasion of the Mediterranean Sea: test whether they share genomic characteristics that predispose a population to easily spread in a new environment and understand the consequences of the invasive processes on the genetic variability of these species. We expect to detect signatures of both neutral and selective processes in the invasive population and, in general, the dynamics and consequences of the invasive event.

Part 1. Scientific case of the PhD Research (2 to 3 pages, including figures)

- BACKGROUND

The Mediterranean Sea is considered as a highly heterogeneous environment with great variety in climatic and hydrological conditions, to such an extent that it is defined as a *miniature ocean* [1]. The high diversity in habitats and the isolation of this sea are reflected in high levels of biodiversity and in the presence of numerous endemic species [2]. In recent years numerous factors are threatening the biodiversity of this environment, making the Mediterranean one of the most impacted seas in the world [2]. The invasion by exotic species is one crucial point, as the list of species invading the Mediterranean environment is continuously growing. After the opening of the Suez canal, many species, called Lessepsian invaders, are entering the mediterranean, and some of them are successfully spreading all over the basin. The spread of these species is a serious threat for biodiversity [3]. This invasion seems also to be favoured by the impact of global warming on the Mediterranean [2].

Thanks to the increasing availability of sophisticated genomic approaches, the study of marine invaders has recently focused on understanding the genomic basis of successful invaders [3, 4, 5, 6]. The role of pre-adaptations and post-invasion selection in the invasive success of some species has been considered as a fundamental aspect to investigate, but the studies considering these factors, although increasing, are still limited [4]. Recent studies on Lessepsian invaders found evidence of directional selection acting on the Mediterranean population [5, 6].

However, genomic studies on Lessepsian species have not thoroughly investigated the role and impact of adaptive and deleterious variability in the colonisation of the Mediterranean Sea [4]. Nonetheless, the investigation of such variability has been considered as crucial to fully understand the factors underlying the success of invasive species in the new environment and to allow the design of predictive models [3].

SCIENTIFIC AIMS

Using Lessepsian fish species as a model, our aims are to categorise genetic variability based on its function: adaptive, deleterious and neutral, and to investigate the trajectories of these different types of genetic variability in the colonisation of new habitats, in order to answer the following questions:

- Are there ongoing selective processes acting in the invasive range?
- Are some of the genetic variants under selection in the new environment present in the source population as possible preadaptation?
- How did the invasion of the Mediterranean by Lessepsian species occur? And what are the dynamics of expansion in the invasive range?

- WORKPLAN AND RESEARCH ACTIVITIES

WP 1. Sampling activity

The main objective for the first year of the PhD project was to obtain samples of the Lessepsian invasive species chosen for the project.

The chosen sampling locations were: Red Sea, Cyprus and Greece (Figure 1); while the invasive species on which the sampling activities were focused were: *Siganus luridus*, *S. rivulatus*, *Pterois miles*, *Lagocephalus sceleratus*, *Fistularia commersonii* (Figure 1).

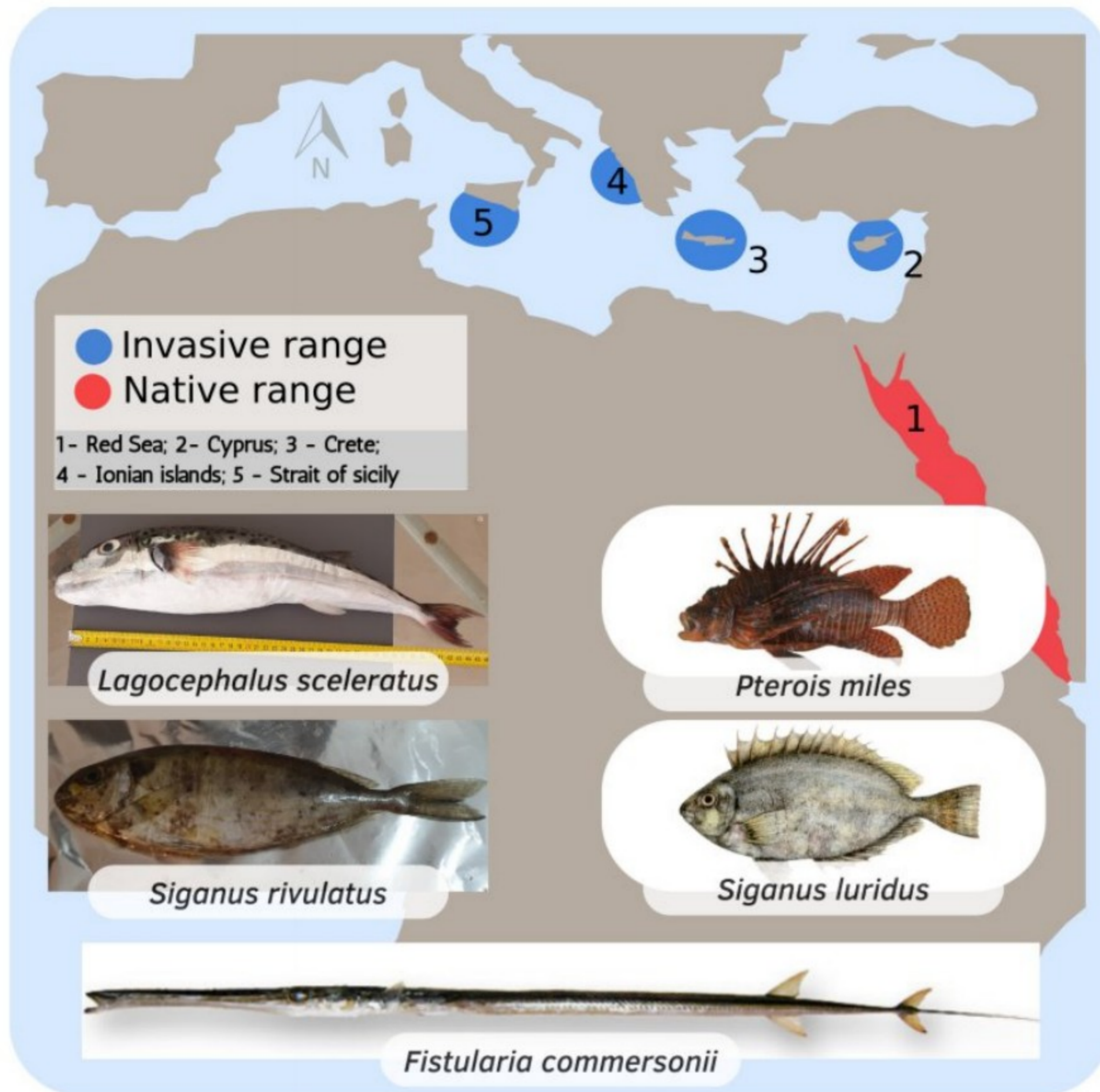


Figure 1. Sampling locations and chosen species for the sampling.

The sampling design was structured to obtain 10 individuals from which to extract DNA (from muscle) and 5 individuals from which to extract RNA samples (from muscle, gonad, liver, brain, and kidney) for each species, from each sampling locality.

Regarding the Red Sea, we sought reliable contacts to obtain the necessary samples; while for the Mediterranean Sea the sampling was carried out by myself, Emanuela Fanelli (supervisor), Ernesto Azzurro and Emiliano Trucchi (supervisor). Samplings were carried out through the cooperation of local fishermen or by spearfishing.

We were able to obtain almost all of the samples we set out to get, from Cyprus, Crete and Kefalonia. Some species were more difficult to find, especially at the borders of the invasive range. At the end of the sampling, as was already planned, we decided to narrow the focus of our project to 3 species: *Siganus luridus*, *S. rivulatus*, and *Pterois miles*, as they were found at each sampling site in sufficient quantity for the planned analyses.

WP 2. Research of available data and analysis methodologies.

In addition to collecting samples, the objective of the first year was to search for genomic and transcriptomic data available in various online databases and identify the most suitable and up-to-date methodologies to perform our analyses.

The search for genomic data was done through the use of various online tools, such as VGP (Vertebrate Genome Project), NCBI (National Center for Biotechnology Information), ENA (European Nucleotide Archive) databases; but also through Google Scholar search. The search has revealed the limited availability of genomic and transcriptomic data regarding these species; however, on the other hand, for *P. miles* [7] and *L. sceleratus* [8] reference genomes are available online (*P. miles* in preparation). The rest of the necessary data will be produced by us during the project.

Regarding the methodologies for analysing data, some software has been identified. These software are advanced and widely used by the scientific community.

- The Cactus software [9] has been identified as an ideal tool for detecting mutations under selection by constructing multi-alignments among numerous species.
- GONE software [10] will likely be used to reconstruct recent demographic histories of these species
- XP-EHH software [11] will be used to go and detect selective signals along the genome of sampled individuals.
- EdgeR pipeline [12] will be used to detect highly expressed genes and to highlight different patterns of expression between genes.
- SLiM [13] simulative software will be used to validate expectations and test hypotheses.

Data preparation for the analyses will be produced through the use of standard-use software such as VCFtools [14], GATK [15] or Freebayes [16].

- REFERENCES

1. Bethoux, J. P., Gentili, B., Morin, P., Nicolas, E., Pierre, C., & Ruiz-Pino, D. (1999). The Mediterranean Sea: a miniature ocean for climatic and environmental studies and a key for the climatic functioning of the North Atlantic. *Progress in Oceanography*, 44(1-3), 131-146.
2. Templado, J. (2014). Future trends of Mediterranean biodiversity. In *The Mediterranean Sea* (pp. 479-498). Springer, Dordrecht.
3. Sherman, C. D. H., Lotterhos, K. E., Richardson, M. F., Tepolt, C. K., Rollins, L. A., Palumbi, S. R., & Miller, A. D. (2016). What are we missing about marine invasions? Filling in the gaps with evolutionary genomics. *Marine Biology*, 163(10), 1-24.
4. Chiesa, S., Azzurro, E., & Bernardi, G. (2019). The genetics and genomics of marine fish invasions: a global review. *Reviews in Fish Biology and Fisheries*, 29(4), 837-859.
5. Azzurro, E., Nourigat, M., Cohn, F., Souissi, J. B., & Bernardi, G. (2021). Right Out of The Gate: The Genomics of Lessepsian Invaders in The Vicinity of The Suez Canal. 10.21203/rs.3.rs-632020/v1.
6. Bernardi, G., Azzurro, E., Golani, D., & Miller, M. R. (2016). Genomic signatures of rapid adaptive evolution in the bluespotted cornetfish, a Mediterranean Lessepsian invader. *Molecular ecology*, 25(14), 3384-3396.
7. Christos V. Kitsoulis, Vasileios Papadogiannis, Jon B. Kristoffersen, Elisavet Kaitetzidou, Aspasia Sterioti, Costas S. Tsigenopoulos, & Tereza Manousaki. (2022). A high-quality reference genome assembly for the devil firefish, *Pterois miles*. Zenodo. <https://doi.org/10.5281/zenodo.6380502>
8. Danis, T., Papadogiannis, V., Tsakogiannis, A., Kristoffersen, J. B., Golani, D., Tsaparis, D., ... & Manousaki, T. (2021). Genome analysis of *Lagocephalus sceleratus*: unraveling the genomic landscape of a successful invader. *Frontiers in genetics*, 12.
9. Armstrong, J., Hickey, G., Diekhans, M., Fiddes, I. T., Novak, A. M., Deran, A., ... & Paten, B. (2020). Progressive Cactus is a multiple-genome aligner for the thousand-genome era. *Nature*, 587(7833), 246-251.
10. Santiago, E., Novo, I., Pardiñas, A. F., Saura, M., Wang, J., & Caballero, A. (2020). Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. *Molecular Biology and Evolution*, 37(12), 3642-3653
11. Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., ... & Lander, E. S. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature*, 449(7164), 913-918.
12. Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *bioinformatics*, 26(1), 139-140.
13. Haller, B. C., & Messer, P. W. (2019). SLiM 3: forward genetic simulations beyond the Wright-Fisher model. *Molecular biology and evolution*, 36(3), 632-637.
14. Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156-2158.

15. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, 20(9), 1297-1303.
16. Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. arXiv preprint arXiv:1207.3907.

Part 2. PhD student information on the overall year activity (courses/seminars/schools, mobility periods, participation to conferences)

List of attended courses/seminars/schools

1. 12/04/22 - UNIVPM - A Shot of Science: "Marine Benthic microalgae as indicators for environmental quality"
2. Online - SLiM Workshop - SLiM, an evolutionary simulation framework
3. Online - Linux Shell Scripting: A Project-Based Approach to Learning
4. Online - Python A-Z™: Python per Data Science
5. Online - SLiM Workshop - SLiM, an evolutionary simulation framework

List of periods spent abroad

1. 5-18 June 2022, Cesky Krumlov, Czech Republic: participation at the: "2022 WORKSHOP ON POPULATION AND SPECIATION GENOMICS, CESKY KRUMLOV"
2. 24 June - 8 July 2022, Larnaca - Cyprus: sampling activity in order to get samples of the invasive species that we are going to study
3. 1-11 October 2022, Cephalonia - Greece: sampling activity in order to get samples of the invasive species that we are going to study

List of conferences/workshops attended and of contributions eventually presented

1. 5-18 June 2022, Cesky Krumlov, Czech Republic: participation at the: "2022 WORKSHOP ON POPULATION AND SPECIATION GENOMICS, CESKY KRUMLOV"
2. 4-7 September 2022, Ancona, Italy: participation and presentation a poster at the: "9TH SIBE CONGRESS"
3. 21 September 2022, Online: participation at the conference: "The role of genomics for biodiversity conservation in a changing climate"

Part 3. PhD student information on publications

List of publications on international journals

In preparation - Study of the the relationship between mito-nuclear discordance and sex biased dispersal in natural populations, using a forward simulation method (SLiM).

List of publications on conference proceedings

none

List of other publications (books, book chapters, patents)

none

17/11/22

Student signature

Thomas Groll

Supervisor signature

Emmanuel

Groll
