UNIVERSIDADE NOVA DE LISBOA

MScCBBi

MASTER IN
COMPUTATIONAL BIOLOGY & BIOINFORMATICS

Diogo Alexandre Magalhães Rodrigues

BSc in Biology

UNDERSTANDING THE GENETIC CONSEQUENCES OF EXPANSIONS FROM MULTIPLE REFUGIA DURING GLACIAL CYCLES THROUGH SPATIAL SIMULATIONS

September, 2024

NOVA SCHOOL OF SCIENCE & TECHNOLOGY
NOVA MEDICAL SCHOOL
NOVA IMS Information Management School
DESDE 1902 INSTITUTO DE HIGIENE E MEDICINA TROPICAL UNIVERSIDADE NOVA DE LISBOA

DEPARTAMENT OF
LIFE SCIENCES

UNDERSTANDING THE GENETIC CONSEQUENCES OF EXPANSIONS FROM MULTIPLE REFUGIA DURING GLACIAL CYCLES THROUGH SPATIAL SIMULATIONS

Diogo Alexandre Magalhães Rodrigues

BSc in Biology

Adviser: Dr. Lounès Chikhi
Principal Investigator (PI) of the Population and Conservation Genetics group, Instituto Gulbenkian de Ciência

Co-advisers: Dr. Patrícia Brito
Senior researcher, UCIBIO, FCT-NOVA, NOVA University Lisbon

MASTER IN COMPUTATIONAL BIOLOGY & BIOINFORMATICS

NOVA University Lisbon
September, 2024

UNDERSTANDING THE GENETIC CONSEQUENCES OF EXPANSIONS FROM MULTIPLE REFUGIA DURING GLACIAL CYCLES THROUGH SPATIAL SIMULATIONS

Acknowledgments

Primarily, I would like to express my gratitude to my main supervisor, Dr. Lounès Chikhi for their guidance, and insightful feedback throughout the development of this thesis. Their expertise and feedback have been essential in bringing this work to fruition and helping me navigate through the challenges this project presented.

I would also like to thank Dr. Patrícia Brito, who, while not directly involved in the research, regularly checked in on my progress, deadlines, and was always willing to offer her support when needed.

A heartfelt thank you to both past and present colleagues at the PCG group in IGC, that in one way or another were a source of motivation via the stimulating meetings, collaborative spirit, and lunch times. A special thanks to a particular member, Ravi Vishwakarma. While he was not my official supervisor, he was responsible for providing weekly feedback and unwavering support as a friend. Without his guidance, presence, and camaraderie this thesis would simply not have been possible. If I ever have a MSc student under my responsibility like he did, I would be totally happy to be just like Ravi was for me.

Lastly, I am immensely grateful to my family and girlfriend for their love, encouragement, and patience. To my parents, thank you for believing in me and supporting me through every step of my academic journey, and to my girlfriend for always being there for me, endlessly supporting me.

Thank you all for helping me reach this milestone

Abstract

Human activities, along with natural climatic and geological events, have caused habitat expansions and contractions with consequences on the range of many species. Understanding the genetic consequences of these habitat changes is crucial for addressing biodiversity loss and informing conservation strategies. While many studies use spatial simulations to explore the genetic consequences of spatial processes, the role of life-history traits and multiple refugia is often overlooked. Vishwakarma et al., 2024 investigated how one ca. 14,000 years cycle of habitat expansion and contraction influence the dynamics of genetic diversity, in a refugium, for different contraction speeds and for species varying in life-history traits. However, they only simulated expansions starting from a single refuge. The current study expands on that work by simulating the same habitat changes across a second cycle and using more than one refugium. We also varied contraction speeds, generation times, and dispersal abilities, while analysing genetic diversity and differentiation patterns and trajectories over millennia. Our results reveal three distinct genetic differentiation dynamics in response to the same habitat changes. These dynamics can be predicted by the F_{st} values at the start of contraction, relative to equilibrium within refugia (F_{st}^{refuge}) or across the landscape ($F_{st}^{landscape}$). Additionally, our findings indicate that genetic differentiation between refugia at the beginning of the expansion have more impact on the genetic differentiation for both longer generation length species and for faster contraction periods. Additionally, geographic location of refugia significantly affects genetic differentiation, with the timing of refugial connections being crucial to their genetic dynamics.

Keywords: genetic differentiation, genetic diversity, spatial processes, multiple refugia, generation time

Resumo

As atividades humanas, juntamente com eventos climáticos e geológicos naturais, têm causado a expansão e contração dos habitats com consequências na distribuição das espécies. Para combater a perda de biodiversidade e criar estratégias eficazes de conservação, é crucial compreender as implicações que as mudanças de habitat têm nos genes das espécies que sofrem essas mudanças. Embora existam vários estudos que utilizam simulações espaciais para explorar as consequências genéticas de alterações de habitats, muitos negligenciam o papel dos *life-history traits* e a existência de diversos refúgios. Em Vishwakarma et al., 2024 foi estudado de que forma um ciclo com cerca de 14000 anos de expansão e contração de habitat, impacta as dinâmicas de diversidade genética, num refúgio, para diferentes velocidades de contração e para espécies com diferentes *life-history traits*. No entanto, esse estudo focou-se em iniciar o processo de expansão a partir de um único refúgio. Esta tese pretende expandir os resultados obtidos nesse outro trabalho simulando as mesmas alterações de habitat num segundo ciclo e usando mais do que um refúgio. Para além disto variámos as velocidades de contração do habitat, assim como o tempo de geração e capacidade de dispersão das espécies, enquanto analisámos padrões de diferenciação e diversidade genética ao longo de milénios. Os nossos resultados revelam que existem essencialmente três dinâmicas temporais distintas de diferenciação genética. Essas dinâmicas podem ser previstas com base nos valores de F_{st} medidos no início da contração, quando comparados aos valores de equilíbrio dentro dos refúgios (F_{st}^{refuge}) ou ocupando toda a paisagem ($F_{st}^{landscape}$). Para além disso, as nossas descobertas indicam que a diferenciação genética entre refúgios no início da expansão têm mais impacto na diferenciação genética tanto para espécies com maior tempo de geração como para períodos mais curtos de contração. Por fim, os resultados demonstram um impacto direto da localização geográfica dos refúgios nos padrões de diferenciação genética, já que o momento em que os refúgios entram em contacto é fulcral para a prever as suas dinâmicas genéticas.

Palavras-chave: diferenciação genética, diversidade genética, mudanças de habitat, múltiplos refúgios, tempo de geração

CONTENTS

List of Figures

List of Tables

GLOSSARY

Population Genetics	The study of genetic variation within and among populations and the evolutionary factors that explain this variation.
Biodiversity (biological diversity)	The variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems.
Genetic diversity	The biological variation that occurs within species. It makes it possible for species to adapt when the environment changes.
Spatially explicit models	Explicit representations of geographic space. Among spatial entities (patches, cells, points) there are defined spatial relations.
Refugia/microrefugia	Small area(s) with local favourable environmental features, in which small populations can survive outside their main distribution area, protected from the unfavourable regional environmental conditions.
mutation-drift-migration equilibrium	Maintenance of equilibrium level of genetic frequencies due to acting forces of mutation, drift and migration
Allometry	Relation between the size of an organism and aspects of its physiology, morphology, and life history

.

ACRONYMS

IUCN	International Union for Conservation of Nature
SINS	Simulating Individuals in Space (program)
MNA	Mean number of alleles
GL	Generation Length
IAS	Invasive alien species
SW	Southwest (refuge)
SE	Southwest (refuge)
N	North (refuge)
t_{cont}	Time/duration of contraction

SYMBOLS

K	Carrying capacity
F	Friction
H_e	Expected heterozygosity
H_e^{refuge}	Expected heterozygosity at mutation-drift-migration equilibrium, when only refugia are populated
$H_e^{landscape}$	Expected heterozygosity at mutation-drift-migration equilibrium, when the entire landscape is populated
F_{ST}	Fixation index
F_{ST}^{refuge}	Fixation index at mutation-drift-migration equilibrium, when only refugia are populated
$F_{ST}^{landscape}$	Fixation index at mutation-drift-migration equilibrium, when the entire landscape is populated
r	Growth rate
m	Migration rate
u	Mutation rate

NOTES

I would like to acknowledge the use of ChatGPT, which assisted in rephrasing some sections of my own text and provided guidance with coding in R, particularly for enhancing the aesthetic aspects of visualizations. In order to write section 2.1 I used previously written materials and methods regarding SINS in papers from the PCG team at IGC, namely: Rasteiro et al., 2012 , Sgarlata et al., 2022a,b and Vishwakarma, R. et al., 2024.

1

INTRODUCTION

Humanity's unprecedented economic and technological growth in recent decades has come at a steep price to our planet. Earth's ecosystems are now under extreme pressure as our demand for energy, food, fiber, water, and land arises. This relentless human impact is pushing our planet to a tipping point, potentially altering these ecosystems function beyond repair (Steffen et al., 2015). As a consequence, biodiversity is severely under threat.

Ironically, the very diversity of life forms that has driven our growth, is now at risk: plants, responsible for photosynthesis; bacteria, fungi, and invertebrates who enrich soils by breaking down organic matter; or pollinators, vital for securing food production, etc. These are just a few examples of the countless ecosystem services that biodiversity provides, and we are highly dependent on (*Biodiversity Loss: What Is Causing It and Why Is It a Concern? | Topics | European Parliament*, n.d.).

Among major causes for biodiversity loss are the exploitation of forests, seas, and other natural habitats for human activities. According to the United Nations report on world's forests, out of the 28,000 species currently at risk of extinction, overexploitation specifically poses a threat to over 85 percent of them (*"The State of the World's Forests 2020,"* 2020); secondly, the introduction of invasive alien species (*IAS*), responsible for nearly 40 percent of all animal extinctions since the 17th century, when there is a known cause (United Nations Environment Programme, n.d.; Secretariat of the Convention on Biological Diversity, n.d.); pollution is another significant driver of biodiversity loss, for example the misuse of non-selective insecticides that devastates insect and plant populations (United Nations Environment Programme, n.d.); and lastly, climate change (gas emissions have doubled since 1980 and are responsible for a 0.7 degrees Celsius increase in global temperatures) (Nature's Dangerous Decline 'Unprecedented' Species Extinction Rates 'Accelerating,' n.d.).

Forest exploitation is leading to a decline in species habitats, which, in turn, causes reductions in population sizes (Finn et al., 2023). In addition to these current threats, species habitats were also modified in the past by major climatic and geological events, such as glacial-interglacial cycles (Hewitt, 2004). These processes—both past and present—can result in habitat loss and fragmentation, forcing species to experience range expansions, contractions, or shifts.

To ensure the healthy functioning of ecosystems, it is urgent to conserve biodiversity at three levels: genetic diversity, species diversity, and ecosystem diversity, as recognized by the International Union for Conservation of Nature (IUCN)..

The IUCN Red List of Threatened Species is the world's most comprehensive tool for assessing the conservation status and extinction risk of species (*The IUCN Red List of Threatened Species*, n.d.). Several criteria are used to determine if a taxon belongs or not to the Red List threatened categories. These include population size reduction (Criterion A), geographic range (Criterion B), small population size and decline (Criterion C), very small or restricted population (Criterion D) and quantitative analysis such as population viability analysis (Criterion E). Although acknowledged, the current approach overlooks genetic diversity (Norderhaug et al., 2024; Hoban et al., 2023). To this date, genetic diversity is not a measure incorporated into the risk assessment of the IUCN Red List (Schmidt, C., et al., 2023). The incorporation of genetic diversity can help to inform about red-list status and vice-versa. Red list can help in identifying species at risk of rapid genetic erosion (Canteri et al., 2021; Garner et al., 2020), and gathering data on genetic diversity of certain species might improve red-list assessments (Brüniche-Olsen et al., 2021; Petit-Marty et al., 2021).

Genetic diversity is essential to determine a species' ability to adapt to different environments as variation provides the raw material for natural selection to act upon, leading to the emergence of adaptive traits (Hopkins, G. R., et al., 2012). Additionally, low genetic variability in populations can lead to an inbreeding phenomenon. This is a process characterized by the mating between relatives, leading to an increase in homozygosity and the expression of recessive alleles, ultimately resulting in inbreeding depression. Inbreeding can lead to reductions in phenotypic values and fitness due to the over-expression of deleterious homozygotes, impacting population dynamics under stressful conditions (Alvarez, G., et al., 2011; Brodie, E. D., 2007). Populations with abundant genetic variation can better align their adaptive plasticity with environmental selection pressures, enhancing their fitness and reducing extinction risks (Walter et al., 2024). Therefore, understanding the genetic consequences of changing habitats, both past and present, is crucial for addressing biodiversity loss and developing effective conservation strategies.

Several meta-analyses have examined the genetic effects of habitat loss and fragmentation focusing on various taxonomic groups: mammals (Lino et al., 2019), tetrapod (Rivera-Ortiz et al., 2015), broad range of plants and animals (Schlaepfer et al., 2018). These studies consistently found that habitat loss and fragmentation have negative consequences on genetic diversity. Notably, they observed that the mean number of alleles (a measure of genetic variation) shows a stronger response to habitat changes compared to expected heterozygosity. While these empirical studies provide valuable insights, analysing the impact of changing habitats on genetic diversity empirically, presents many challenges. It is undeniably difficult to conduct controlled experiments, collect data and manage multiple variables over various timescale in natural settings. To address these limitations, researchers often turn to simulations and analytical models. These approaches provide a flexible, controlled approach to explore how various factors influence genetic diversity. Analytical models are particularly useful as they help interpret genetic variations at both molecular and population-level, complementing empirical studies and possibly offering predictions for scenarios that are difficult to observe in nature.

Population genetics uses different models to study genetic variation and evolutionary processes, with the choice of models depending upon the scenario investigated. Simpler models may represent populations as single isolated units without geographic structure, while more complex models incorporate geographic aspects. These include continent-island model, n-island model and stepping stone model in order of increasing complexity and number of parameters. Moreover, theoretical models like the Wright-Fisher model and coalescent models (forward and backward in time, respectively) are used to track genetic evolution and genealogical relationships within populations (Nordborg, M., 2019 and Yuan, X., et al., 2012).

Spatially explicit models are particularly useful since they may provide a better representation of natural populations by incorporating geographical structure. However, investigating genetic diversity in context of complex spatial-temporal dynamics, such as population expansions, contractions, shifts or dispersal can be challenging analytically (DeAngelis, D. L., & Yurek, S., 2017). Thus, spatial simulations are useful to complement, validate and extend analytical results.

Several studies have utilized spatial simulations to gain genetic insights into genetic consequences of spatial processes. For example, Klopfstein, S., et al., 2006, investigated the phenomenon of mutations arising during range expansions, focusing on how these mutations can spread more efficiently compared to stable populations. In another study Arenas et al., 2012, used simulations to study the molecular consequences of range contractions and shifts in species due to climatic changes. These studies have significantly contributed to our understanding

of how spatial processes can affect genetic diversity. However, many of these studies usually don't incorporate life history traits (nonetheless see Rasteiro et al., 2012; Garnier and Lafontaine, 2022). It remains to be investigated how genetic diversity patterns vary across different species undergoing same habitat changes.

More recently, Vishwakarma, R. et al., 2024 examined genetic diversity patterns across different species undergoing the same habitat changes inspired by glacial-interglacial cycles. They started by focusing on varying generation length and studying its influence on genetic diversity, under alternating speeds of contraction. After that, simulations included more realistic combinations of life-history trait parameters, including dispersal distance and population density values based on reported allometric relationships for mammalian species. Their work featured one cycle of expansion followed by a stationary and a contraction period, with the original population starting in one refugium. This thesis builds on the work of Vishwakarma, R. et al. 2024 by investigating the dynamics of genetic diversity in multiple expansion and contraction scenarios with two refugia. More specifically, we simulated scenarios inspired by climatic oscillations of the Quaternary. During these periods, many species experienced repeated changes in their habitats.

Species move and track their habitats during periods of climatic change. Darwin described these effects in 1859 "On the Origin of Species": "As the climate becomes more and more temperate, the northern forms will tend to migrate northward, closely followed up by the inhabitants of the temperate regions, and these again by the inhabitants of the warmer regions.". In general, this would mean that during colder times, Artic species would colonize more southern regions while temperate species would go further south, having their habitat decreased. Warmer periods meant that species from colder regions got their habitat significantly reduced while temperate species expanded their habitat towards the north (Hofreiter, M., and Stewart, J., 2009). In regions of complex geography, such as Europe, with mountainous regions and peninsulas bordering the Mediterranean, these back-and-forth movements led to formation of isolated populations and multiple refugia/microrefugia. By definition these are "a small area with local favourable environmental features, in which small populations can survive outside their main distribution area, protected from the unfavourable regional environmental conditions" (Rull, Valentí, 2010).

Even though this Glacial-Refugium Theory is widely accepted, ancient DNA reveal a more complex picture: not all species outside of the southern refugia died during a glacial cycle, meaning that there are certain species able to inhabit refugia situated in more northern parts of the landscape (Hofreiter, M., and Stewart, J., 2009).

Understanding the role of multiple refugia is important since it allows different populations to persist in separate areas and reconnect during favourable periods. This process of cyclic isolation and connections are known to provide high excess of genetic diversity compared to what is expected at equilibrium (Alcala, N. and Vuilleumier, S.; 2014).

In this work, we used spatial simulations to investigate the dynamics of genetic diversity and differentiation in scenarios of multiple range expansion followed by contraction. Specifically, we examined *i)* the temporal dynamics of genetic diversity during contraction from two refugia, *ii)* the impact of refugium location on genetic diversity dynamics, and *iii)* the contrasting effects of contraction speed on genetic differentiation at the end of contraction. Our simulations aimed to clarify how different levels of genetic differentiation between refugia influence these dynamics.

2

MATERIALS AND METHODS

2.1 Simulation tool

Simulations were performed by employing SINS (Simulating Individuals in Space), a program built in-house. SINS allows simulating both population demographics and genetic data over non-overlapping discrete generations.. SINS utilizes a forward-time approach to simulate diploid individuals (both males and females) across a two-dimensional grid of demes, resembling a 2D stepping stone lattice (Kimura and Weiss, 1964). However, its design is primarily based on the principles of SPLATCHE (Currat et al., 2004). Both programs share numerous features, but they also differ in others, thereby offering complementary solutions (more details on the comparison between SINS and SPLATCHE can be found in Rasteiro et al., 2012 and Sgarlata et al., 2022a,b).

In a very simplistic manner, each deme has individual *K* (carrying capacity) and *F* (friction) values which are user-defined. The population in a deme is regulated by logistic growth formula, with a slight change accounting for a limitation of growth by the number of reproductive females (Smith and Slatkin,1973). The following equation represents how to calculate the expected number of individuals in generation *t+1*:

Equation 1

$$E(N_{t+1}) = 2N_{ft}\frac{1+r}{1+r\left(\frac{2N_{ft}}{K}\right)}$$

where N_{ft} represents the number of females at generation *t*, *r* is the intrinsic growth rate, and *K* corresponds to the carrying capacity of a deme. Note that, deme population size N_{t+1} is not computed deterministically as the number of individuals in generation *t+1* is drawn from a

Poisson distribution with mean $E(N_{t+1})$. SINS assumes random mating within demes, where one reproductive individual of each sex is randomly selected to produce offspring. Each offspring inherits one allele from each parent at each genetic locus. Mating continues in this manner each generation until the population reaches N_{t+1}.

Each deme can have up to four neighbouring demes, one in each cardinal direction, with which it can exchange migrants at a certain rate. Regarding migrations, they can only occur towards demes with habitats. The number of individuals that migrate from the focal deme at each generation is drawn from a Poisson distribution with mean $M = \frac{N_t m n_d}{4}$, where m is the migration rate and n_d is the number of available neighbouring demes. The probability that one migrant from the focal deme i chooses to move to a neighbouring deme j depends on the friction values of all neighbouring demes as

Equation 2

$$P_{ij} = \frac{1 - F_j}{\sum_{k \in \Gamma(i)} 1 - F_k}$$

where $\Gamma(i)$ denotes the neighbourhood of deme i in the lattice. The number of migrants moving to each neighbouring deme from a focal deme is then sampled from a multinomial distribution with probabilities given by the previous equation (Equation 2). While several types of molecular markers can be simulated, for the current study, we exclusively used independent microsatellite loci, mutating according to the Stepwise Mutation Model (Ohta and Kimura, 1973).

2.2 Methodological details

The simulations were performed in a landscape composed of 18 x 18 demes, plus two refugia of 2 x 2 demes (Figure 1A). The landscape is oriented north-south and the two refugia are located in the southwestern and southeastern corners (hereafter referred to as SW refugium and SE refugium, respectively). In all simulated scenarios, both refugia are populated at the start. One of the demes inside the refugia was arbitrarily chosen (as its choice does not impact the results) and expansion starts from that deme. We chose the top-right deme of the SW refugium and to make the expansion symmetrical, the starting point for the SE refugium is in top-left deme (both slightly darker orange demes in Figure 1A). Both populations begin with one allele per locus, meaning that the mean number of alleles (hereafter *MNA*) is 1, the

expected heterozygosity (from now *He*) is 0. This implies that F_{st} between both refugia is very low (< 0.1). The size of these ancestral populations is equal to $4K*2$, where K is carrying capacity according to the simulated scenario. When the whole landscape has been colonized, population size is $(18 \times 18 + 8) \times K = 332 \times K$ and when contraction ends, only the two refugia are habitable, therefore having a population of size $4K$ for each refugium. In total, this corresponds to an increase of more than (and a decrease of around) two orders of magnitude of the total population size. For instance, when $K = 50$, the total population size in the refugia corresponds to 400 individuals. The population increases to approximately 16600 individuals when whole landscape is colonized. We used a growth rate of $r = 0.8$ across all simulations. For each individual we simulated 50 independent microsatellite loci per individual, with a mutation rate of 10^{-3} per locus per generation. This is an estimated rate (on the higher end) reported in the literature of several mammalian species (Schlöttereretal. 1998).

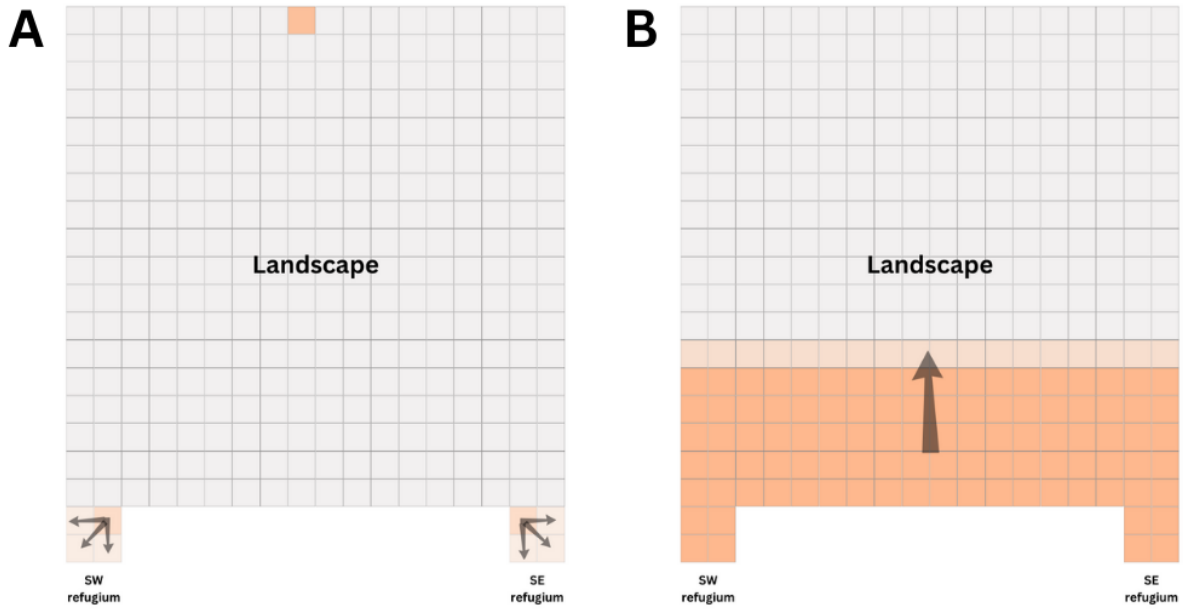

Figure 1 - **Simulated landscape:** A) The simulated 2D landscape is composed by a main area where there are 18 x 18 demes (grey) as well as two additional refuge areas (SW and SE refugium) of 2 x 2 demes (faded orange). Inside both refugia you can identify one deme with a slightly more saturated orange. These correspond to the starting point of the ancestral populations from which the spatial expansion starts. The faded arrows indicate the occupation of all demes inside the refugium. B) After certain number of years, the habitat has already been expanded, both refugia have been occupied as well as a reasonable amount of demes. The faded orange demes show the next expansion step (next demes to be colonized) and the arrow shows the direction of expansion.

The most natural way to measure time in population genetic studies is in terms of generations. However, since we wanted to simulate species with different generation length, we

choose to model habitat changes using “years” as time unit. Using "years" as time unit allows for comparison between different species experiencing same habitat changes.

Since our simulations began with two populated refugia, we varied F_{st} values between the SW and SE refugia to simulate different duration of isolation between them. These F_{st} values indicate the level of genetic differentiation between the refugia. For instance, prolonged isolation would result in higher F_{st} values, while shorter isolation would produce lower values of F_{st}. To simulate these different values of F_{st}, we modelled habitat expansion at different times. For low F_{st} values, habitat expansion began at time t = 0 generation, whereas for high F_{st} values, we allowed the two refugia to evolve in isolation for 50,000 generations before initiating habitat expansion.

After establishing different timepoints at which expansion starts, we modelled habitat expansion as follows. The expansion to whole landscape occurs sequentially from south to north direction in 18 steps. Each expansion step occurs every 20 years such that habitat expansion lasts for a total duration of 360 years. Each expansion step involves creating a new horizontal line of 18 demes as represented in Figure 1B (6th step, light orange). We assume that each deme represents an area of 100 x 100 km^2, which results in at least 1800 km separating the demes in the south from the northernmost. The habitat expansion speed is set at approximately 5 km/year, which is higher but comparable to the estimated rate of spread observed in European pine, hazel, and alder (estimated to be 1.5 to 2 km/year according to Hewitt, 1996). The expansion phase allows the habitat to increase, making it available to be colonized by the individuals. Following the expansion, the habitats remain stable for 2,500 years, during which the landscape was fully available for colonization without changes to configuration of the habitat. The rate of colonization is going to vary from one species to the other, depending on migration and growth rates. Of all parameter combinations tested, during this steady period, all simulations were able to colonize the entire landscape. Next, landscape contraction was implemented by updating the values of K to 0 and F to 1 for each horizontal line of 18 demes, turning them unhabitable, one row at a time. The contraction occurred from north to south direction. We modelled different intervals for contraction steps, resulting in three different contraction speeds (check Table 2 for a summary of speeds). The fastest contraction speed is going to reduce the habitat each 20 years, resulting in a 340 year-time period of contraction. The total duration of range contraction (henceforth t_{cont}) for the intermediate contraction is 4080 years (intervals of 240 years) and the slowest one lasts for 10880 years (which means the interval between steps is 640 years) (Figure 2). These durations of contractions correspond to contraction speeds of approximately 5, 0.4, and 0.15 km/year, respectively.

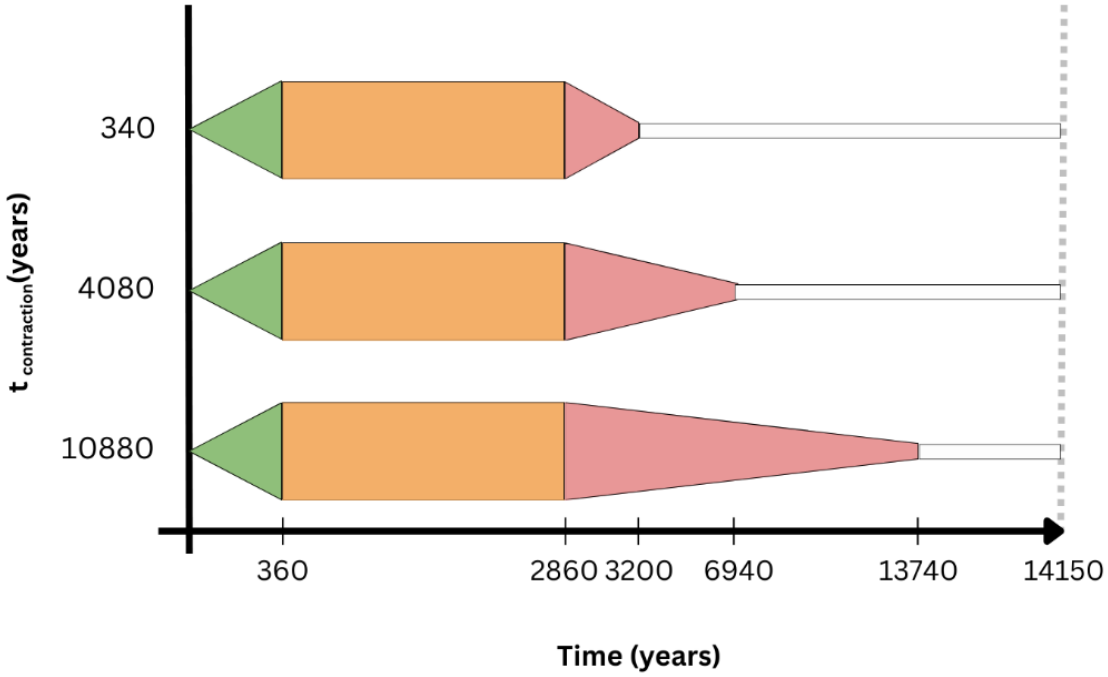

Figure 2 - **Schematic representation of habitat changes:** This diagram illustrates the temporal evolution of habitats, depicting an initial phase of expansion (shown in green), followed by a stationary phase (depicted in orange), and finally, a contraction phase (highlighted in red). Three distinct scenarios were examined (for each starting F_{st} condition) all featuring identical expansion and stationary periods but varying in contraction speed and duration (t_{cont}). Following contraction, refuge areas were identified where populations could persist. Sampling was conducted exclusively within these refuge areas (further details are provided in the text).

2.3 Analysing Genetic Diversity

We simulated 15 independent simulations for each scenario of contraction speed and two initial F_{st} values between refugia. We sampled a total of 80 individuals (10 per deme, accumulating to 40 from each refugia) across 66 separate time points. This number results from the fact that the timing of sampling (in years) must remain the same across the 3 different contraction speeds as well as the initial F_{ST} values. Because sampling this much amount of data notably compromises the speed of simulations, we needed to give special emphasis to sampling in certain regions rather than others. Therefore, we opted for sampling mainly in the beginning of expansion and the ending of the 3 types of contraction (if you check plots in Figure 3 you can observe that there are more data points near those referred regions). In order to compute the molecular diversity of the individuals throughout the simulations, we used R software (version 4.1.2). Three key statistics were then calculated: Mean number of alleles (*MNA*) was

computed using the adegenet R package (v2.1.10) (Jombart, 2008), while expected heterozygosity (H_e) and genetic differentiation among both refugia (F_{st}) was obtained via hierfstat R package (v0.5.11) (Goudet, 2005).

3

RESULTS

3.1 Dynamics of H_e from two different refugia (SW and SE)

Figure 1 illustrates the expansion phase, where habitats expand from the SW and SE refugia in a south-to-north direction. As previously described, there are three phases that compose a simulation: expansion, stationary and contraction period. Figure 2 represents these phases in different colours and shows three contraction durations (represented in the y-axis).

Figure 3 portrays dynamics of genetic diversity for two different species: one of them with two generations per year (GL=0.5 years, Panel A), and the other with a new generation every five years (GL=5 years, Panel B). Each panel has three differently coloured lines that correspond to values of expected heterozygosity associated to differing contraction durations: *i)* t_{cont} = 340 years (red), *ii)* t_{cont} = 4080 years (green), or *iii)* t_{cont} = 10880 years (blue). For both panels, simulations had the same parameters of carrying capacity (K = 50) and migration rate (m = 0.2). We observe that shorter GL's respond faster to the same habitat changes. At any time point during the contraction phase, we are able to witness that the faster contraction speed reaches lower genetic diversity values compared to slow contraction scenarios. During that period, for species with GL = 0.5 years and t_{cont} = 10880, we observe that H_e leads to minor increase before decreasing. In contrast, species with GL = 5 years keeps on increasing for long duration before decreasing. This difference in temporal dynamics seems to be related to the amount of genetic diversity accumulated in the SW and SE refugia before the start of contraction (time = 2860 years, black vertical line), with respect to the expected mutation-drift-migration equilibrium values (horizontal lines in Figure 3). Specifically, for the parameter combinations considered, these equilibria correspond to $H_e^{refuge} \simeq 0.36$ when the species is isolated in

both refugia. Additionally, $H_e^{landscape} \simeq 0.77$ represents the value of H_e measured in the SW and SE refugia populations when the entire landscape has been colonized for extended periods.

For the shorter generation length (GL = 0.5 years), genetic diversity is closer to equilibrium of landscape at the start of contraction ($H_e \sim 0.75$). In contrast, for species with GL = 5 years, $H_e \sim 0.5$ (above H_e^{refuge}, below $H_e^{landscape}$).

Figure 3 shows 2 out of the 3 possible scenarios (check Table 1) that allow to predict genetic diversity trajectory based on equilibria. These results are consistent with those reported by Vishwakarma, R. et al., 2024 (results from the original paper are present in Figure A.1 from supplementary materials), although there was slight difference in the simulated scenario. It is important to note that Vishwakarma et al., 2024 simulated expansion from single refugium whereas the results presented below were obtained when expansion originates from two refugia. However, we do not expect significant differences between the two scenarios, as refugia have the same population when habitat expands (low values for F_{st}).

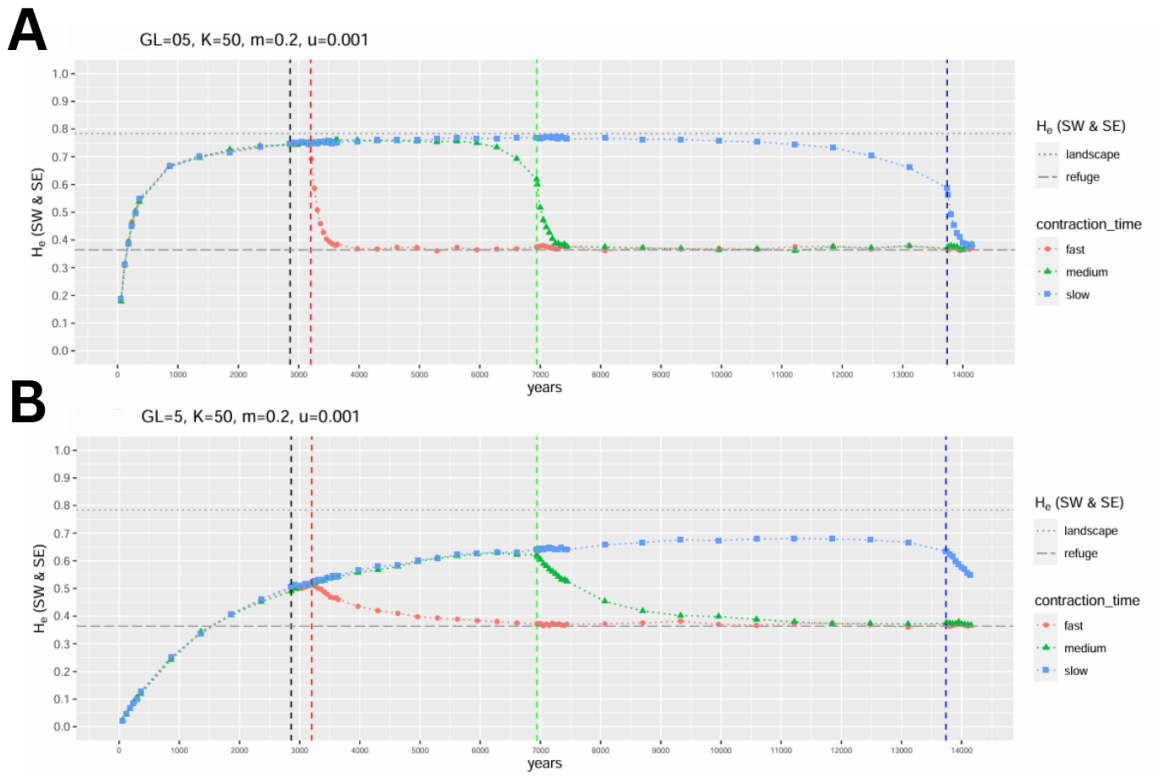

Figure 3 - **Dynamics of H_e comparing both SW and SE refugia for GL05 & GL5:** The same parameters for the simulations were considered but panels differ in the generation length from each species. Panel A presents the dynamics for a species with GL=05, whereas Panel B the species considered has GL=5. In both panels, data is plotted as circles, squares, and tringles that represent the means measured in all 15 simulations. Their shapes and colours vary depending on the speed of contraction (or duration of contraction). The vertical dashed black (x = 2860 years) specifies the starting point of all those contractions and the following vertical-coloured lines indicate the end of contraction associated with that colour (red, green and blue respectively correspond to t=3200, t=6940 and t=13740).

Horizontally dashed grey coloured lines represent the values of H_e at mutation-migration-drift equilibrium when the refugia are totally isolated (refuge) or when the entire landscape is occupied (landscape). Both equilibria are computed by maintaining stationary conditions throughout 100,000 generations.

Additionally, we compared the relationship between heterozygosity at the end of contraction and speed of contraction. Figure 4 illustrates the relationship between H_e at the end of contraction and contraction speeds for two generation lengths. The figure shows expected heterozygosity values at seven different timepoints, after contraction phase is finished (10, 60, 110, 160, 210, 260, 310 years). Just like the previous figure (Figure3), each colour is associated to a particular contraction speed. For shorter GL = 0.5 years (Figure 4A), faster contraction speeds (red bars) tend to preserve more genetic diversity than slower contraction ones just after the end of contraction (10 and 60 years post contraction). The opposite is true for longer GL's (Figure 4B), where slower contraction speeds (blue) preserve more genetic diversity than faster speeds (see discussion for explanation), suggesting that some species might be able to accumulate genetic diversity even when their habitat is being reduced. Additionally, GL = 5 years preserves both higher values for H_e after 310 years, but also keeps the variation of H_e between contraction speeds relatively constant when compared to GL = 0.5 . Once again, the produced results align with the findings in Vishwakarma, R. et al. (2024).

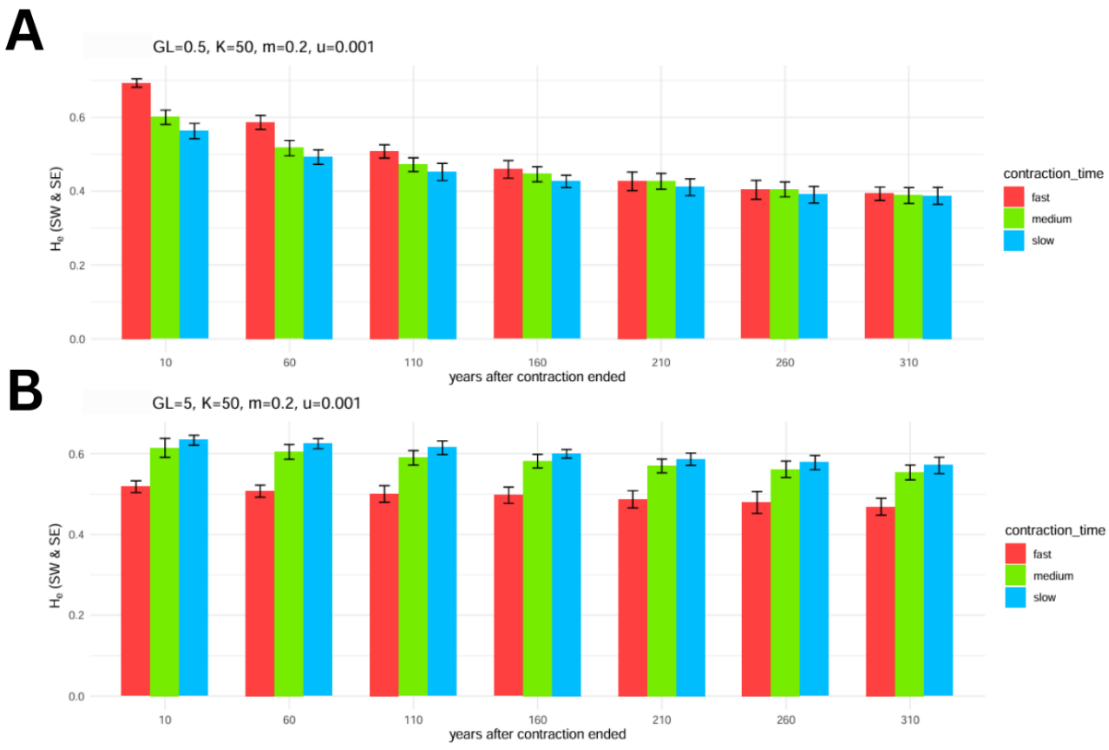

Figure 4 - **H_e of both SW and SE refugia after contraction has ended:** The same parameters for the simulations were considered but Panel A shows the GL = 0.5 years and Panel B shows GL = 5 years. The x-axis values represent the number of years that have passed since the end of contraction. Colours in both panels represent the duration (or speed) of the contraction phase for red, green, and blue scenarios (t_{cont} = 340, 4080, 10880 years, respectively). Standard deviations were measured across 15 simulations and are present just above the histogram bars.

The patterns described for H_e, are also consistent for MNA as a measure of genetic diversity. These plots support the results obtained previously and they are present in Figure A.2 from supplementary materials section. We summarized the results obtained in this section in Table 1, focusing on different heterozygosity scenarios in the population at the onset of contraction.

Table 1 — Scenarios of H_e at the beginning of contraction

Scenario of heterozygosity when contraction starts	Implications	Examples
$H_e \simeq H_e^{landscape}$	H_e is saturated and will decrease towards H_e^{refuge}. Faster contraction speeds retain more genetic diversity at the end of contraction.	Figure 3A & 4A
$H_e^{landscape} > H_e > H_e^{refuge}$	Initial difference between $H_e^{landscape}$ and H_e, allows genetic diversity to keep increasing after contraction starts in the longer contraction scenarios. Slower or intermediate contractions accumulate more H_e at the end of contraction than faster contractions.	Figure 3B & 4B
$H_e < H_e^{refuge}$	Increase of H_e until reaching H_e^{refuge}, regardless of population and habitat decrease	Figure 11A (fast contraction speed)

3.2 Dynamics of H_e from two different refugia, under different starting F_{ST} conditions

As discussed in the introduction, empirical evidence supports the existence of multiple refuge areas (Rowe, K. et al., 2004; Vega, R. et al., 2010; Bartáková, V. et al., 2013). Therefore, we modelled an expansion-contraction scenario beginning from two refugia with different levels of differentiation between them. We explored whether the level of differentiation between refugia affected the dynamics of expected heterozygosity. For "high F_{ST}" scenario we start expansion after isolation of 50,000 generations between the refugia. For K = 50, m =0.2 and u = 0.001, "low F_{ST}" scenarios initialise expansion with values of F_{ST} between refugia ~ 0.02, contrasting with "high F_{ST}" scenarios that start at ~ 0.63.

In Figure A.3 representing the F_{ST} dynamics of species with GL = 0.5 years, we observe that varying the initial F_{ST} conditions had very small influence on the dynamics of H_e. Due to the fact that there is a limit to which H_e can grow (mutation-migration-drift equilibrium at the landscape) and existence of many individuals per generation, H_e from the "low F_{ST}" scenario rapidly catches up to the values of "high F_{ST}" scenario. But Figure 5 illustrates that the same is not true for a longer generation length species (GL = 5 years).

Figure 5 is composed of three panels, where each panel represents one of the different contraction speeds that have been previously analysed (fast, intermediate and slow). Panels A, B and C correspond to the occurrence of contraction steps every 20, 240 and 640 years, respectively. In all panels, we can observe a clear difference between the starting points of the red lines ("high F_{ST}") and the blue lines ("low F_{ST}"). Although in each panel both curves present a similar trajectory, *i.e.* the shape of the curves is similar, the initial discrepancy in the starting points, leads to the observation of different values regarding H_e when we compare initial F_{ST} conditions. We are also able to observe that as the contraction steps get longer, the effect of initial genetic divergence between refugia on expected heterozygosity dynamics is diminished (something further explored in the results and discussion sections). These results suggest that for longer contraction periods, the fact that populations were or not isolated from one another previously, seem to have less impact in the genetic diversity of species.

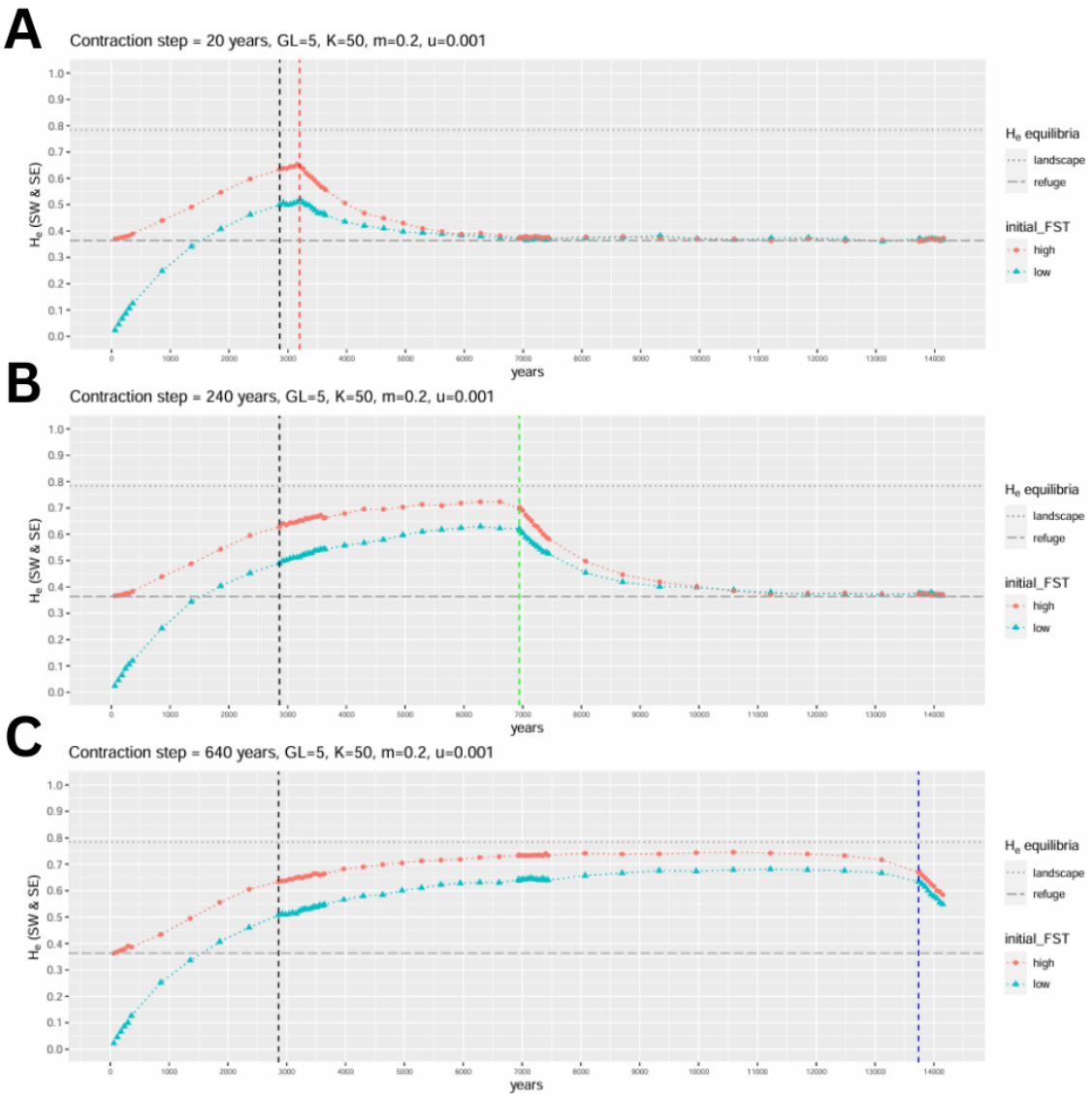

Figure 5 - **Dynamics of H_e for GL=5, in both "low F_{ST}" & "high F_{ST}" scenarios:** Populations are expanding from SW and SE refugia simultaneously for three contraction speeds and identical parameters (GL=5, K=50, m=0.2 and u=0.001). Each panel's total duration is going to match the duration of contractions we saw until now in previous figures: Panel A takes 340 years to contract, Panel B takes 4080 years, and Panel C takes 10880 years. For each panel, there are 2 scenarios corresponding to different initial Fst between refuge: in blue, "low F_{ST}" whereas in red, "high F_{ST}", and only after that period, expansion starts. In both panels, data is plotted as circles, squares, and tringles that represent the means measured in all 15 simulations. The vertical dashed black line (x = 2860 years) specifies the starting point of all those contractions and the following vertical-coloured lines indicate the end of contraction associated with that colour (red, green and blue respectively correspond to t=3200 years, t=6940 years and t=13740 years). Horizontally dashed grey coloured lines represent the values of H_e at mutation-migration-drift equilibrium when the refugia are totally isolated (refuge) or when the entire landscape is occupied (landscape). Both equilibria are computed by maintaining stationary conditions throughout 100,000 generations.

3.3 Characterization of F_{ST}

3.3.1 Dynamics of F_{ST} in a two-cycle expansion and effect of contraction speed on F_{ST} at the end of contraction

In this section we take a closer look at F_{ST} as a measure of genetic differentiation, alike what was earlier described for expected heterozygosity.

Figure 6 portrays two different scenarios, where Panel A represents the low F_{ST} starting point between refugia, while Panel B illustrates the high F_{ST} starting point for the species with GL=5 years, (check Methodological details). The two horizontal lines represent the expected F_{st} values at mutation-migration-drift equilibrium. The $F_{st}^{landscape}$ (represented by dotted line) corresponds to the equilibrium F_{st} when the whole landscape is occupied whereas F_{st}^{refuge} (represented by dashed line) corresponds to the equilibrium F_{st} values when both refugium are isolated for long durations. It is important to note that $F_{st}^{landscape}$ is lower than F_{st}^{refuge}, since F_{st} values are lower when refugium are connected. This is different from the equilibrium heterozygosity values where $H_e^{landscape}$ was greater than H_e^{refuge}. This simulation has the same parameters as previous simulations: carrying capacity (K = 50) and migration rate (m = 0.2).In Figure 6A, during the expansion phase, the level of F_{st} between refugia starts low, and during expansion phase, increases slightly towards the values at $F_{st}^{landscape}$. When contraction starts (black dotted line in Figure 6), population reduction at some demes will increase overall F_{st} because both refugia exchange gradually less migrants, although the sharp increase in this value will be more pronounced after the end of contraction (complete isolation of SW and SE refugia). If we look at Figure 6B, the opposing scenario is noticeable: we see F_{st} between SE and SW decreasing and it keeps that trend throughout the stationary phase. For a fast contraction speed, fixation index keeps going down, then suddenly increases when contraction ends and keeps on growing for the rest of the simulation trying to reach F_{st}^{refuge}. Both intermediate and slow contraction speeds in Figure 6B follow the same trend, but we can clearly see a more gradual increase of F_{st} before contraction terminates, as that shift is not as sudden: intermediate speed's F_{st} decreases, reaching $F_{st}^{landscape}$ values, and starts to increase back again before contraction ends. For the slower contraction its F_{st} values stay relatively stable for ~4000 years and the ongoing increase before the end of contraction is even more pronounced (blue line in Figure 6B).

Results for GL=0.5 years were omitted for the sake of clarity, but they can be found in Figure A.4.

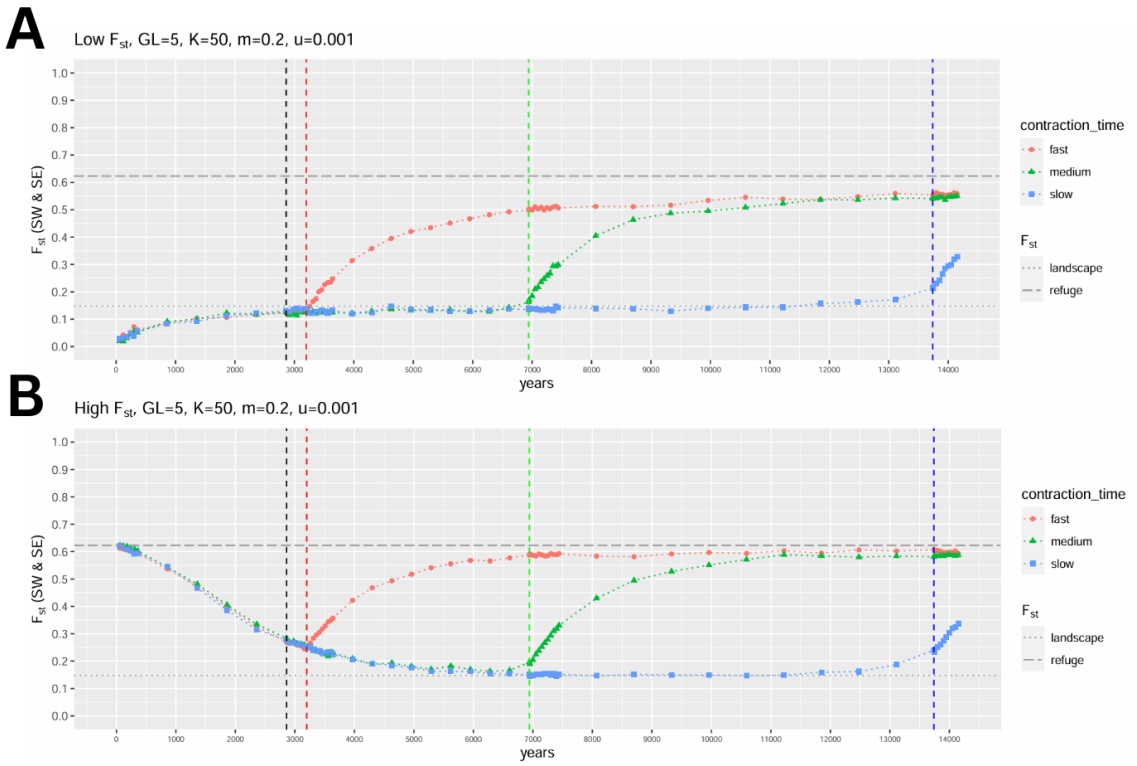

Figure 6 - **Dynamics of F_{st} between both SW and SE refugia:** The two panels differ in the starting levels of F_{st}. Panel A presents a situation where expansion of habitat started at t = 0 (low F_{ST}) whereas Panel B has both refugia living isolated from one another for 50k generations, and only after that period, expansion starts (high F_{ST}). In both panels, data is plotted as circles, squares, and tringles that represent the means measured in all 15 simulations. Their shapes and colours vary depending on the speed of contraction (or duration of contraction). The vertical dashed black (x = 2860 years) specifies the starting point of all those contractions and the following vertical-coloured lines indicate the end of contraction associated with that colour (red, green and blue respectively correspond to t=3200, t=6940 and t=13740). Horizontally dashed grey coloured lines represent the values of F_{st} at mutation-migration-drift equilibrium when the refugia are totally isolated (refuge) or when the entire landscape is occupied (landscape). Both equilibria are computed by maintaining stationary conditions throughout 100,000 generations.

Some observations extracted from H_e and *MNA* still hold for F_{st}, such as the idea that shorter GLs respond faster to habitat changes when compared to longer GLs. Regarding dynamics of F_{st} for contraction speed variation, it is noticeable that at any moment in the contraction phase, faster contraction steps result in higher levels of genetic differentiation between refugia, when compared to slow contraction scenarios. But perhaps most importantly, the trajectory of fixation index can be predicted if we have access to the values of F_{st} at the equilibria (this is explored in the discussion).

As mentioned in the materials and methods section as well as displayed in Figure 6, the simulations display particularly interesting F_{st} patterns at the end of various contractions (after red, green and blue dotted lines). Based on these results and the work of Arenas et al. 2012,

the F_{st} values at the end of contraction were further analysed and the results are presented in Figure 7. For all plots, seven different timepoints after the end of contraction were considered, as well as the three different contraction speeds associated with each scenario. For GL = 5 years, we can observe that the low F_{st} starting scenario shows that faster contraction time period (t_{cont} = 340 years, red bars) will preserve lower levels of F_{st} when compared to slower contraction (t_{cont} = 10880 years, blue bars), indicating that SW and SE refugia will be more genetically distinct, at the end of a slower contraction speed (results are in agreement with Arenas et al. 2012). If we maintain all parameters, but we look instead at a high F_{st} starting scenario, the outcome is different: we see that intermediate contraction speeds have lower F_{st} compared to fast and slow contraction speeds. By analysing Figure 6B this is explained by the continuous decrease of F_{st} for both intermediate and slow contractions, reaching towards $F_{st}^{landscape}$, and the rise of this metric before slow contraction ends. Similar to what happened for H_e, in genetic differentiation dynamics there is importance in the values of accumulated F_{st} in the SW and SE refugia before the start of contraction, in respect to values at equilibria (F_{st}^{refuge} and $F_{st}^{landscape}$).

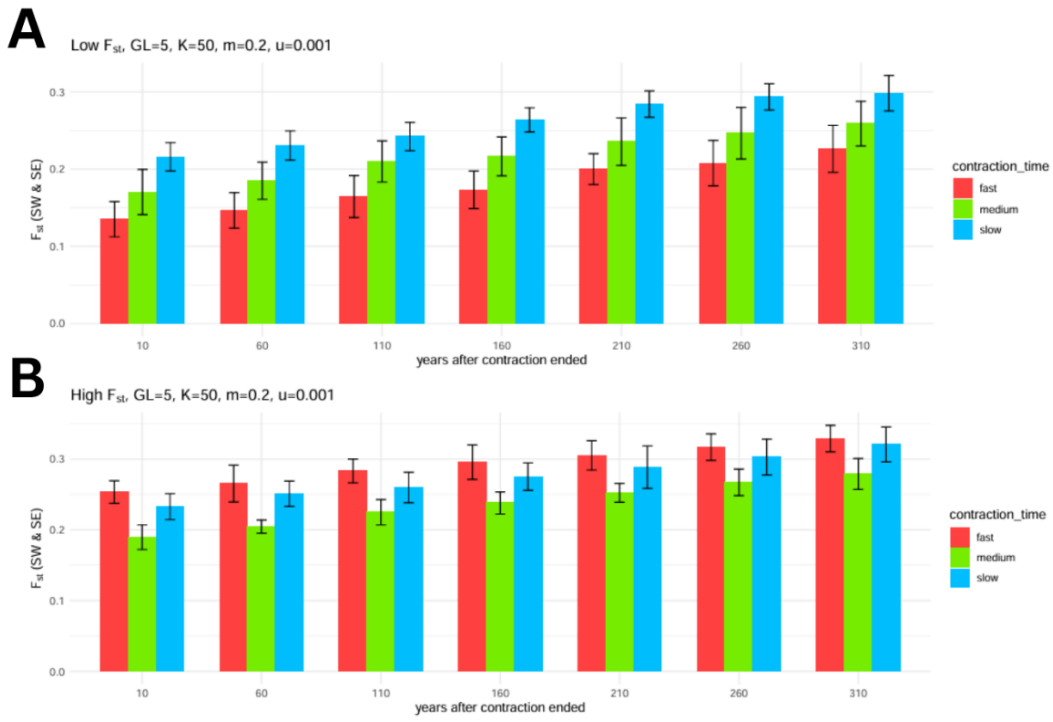

Figure 7 - **F_{st} between SW and SE refugia after the contraction has ended for GL5:** The same parameters for the simulations were considered as in Fig 6, with Panel A showing the low F_{st} scenario and Panel B showing the high F_{st} scenario. Colours in both panels represent the duration (or speed) of the contraction phase for red, green, and blue scenarios (t_{cont} = 340, 4080, 10880 years, respectively). Standard deviations were measured across 15 simulations and are present just above the histogram bars

In the previous section we witnessed that initial genetic differentiation between refugia had more impact in the H_e dynamics for a longer GL, but also that its effect diminished as we rose the length of the contraction steps (time of contraction is extended). The same phenomena are present in F_{st} dynamics, as no significant differences were found at GL=0.5 years (Figure A.4). On the contrary, Figure 8 displays the results for GL=5 years, showing that different initial F_{st} conditions have a clear impact when it comes to faster contraction speeds, but they become a lot less pronounced for slower contraction steps. By the end of each respective contraction, we can see that Figure 8A has a clear separation between the curves representing the two distinct initial F_{st} conditions, while for Figure 8B and 8C that is not so true.

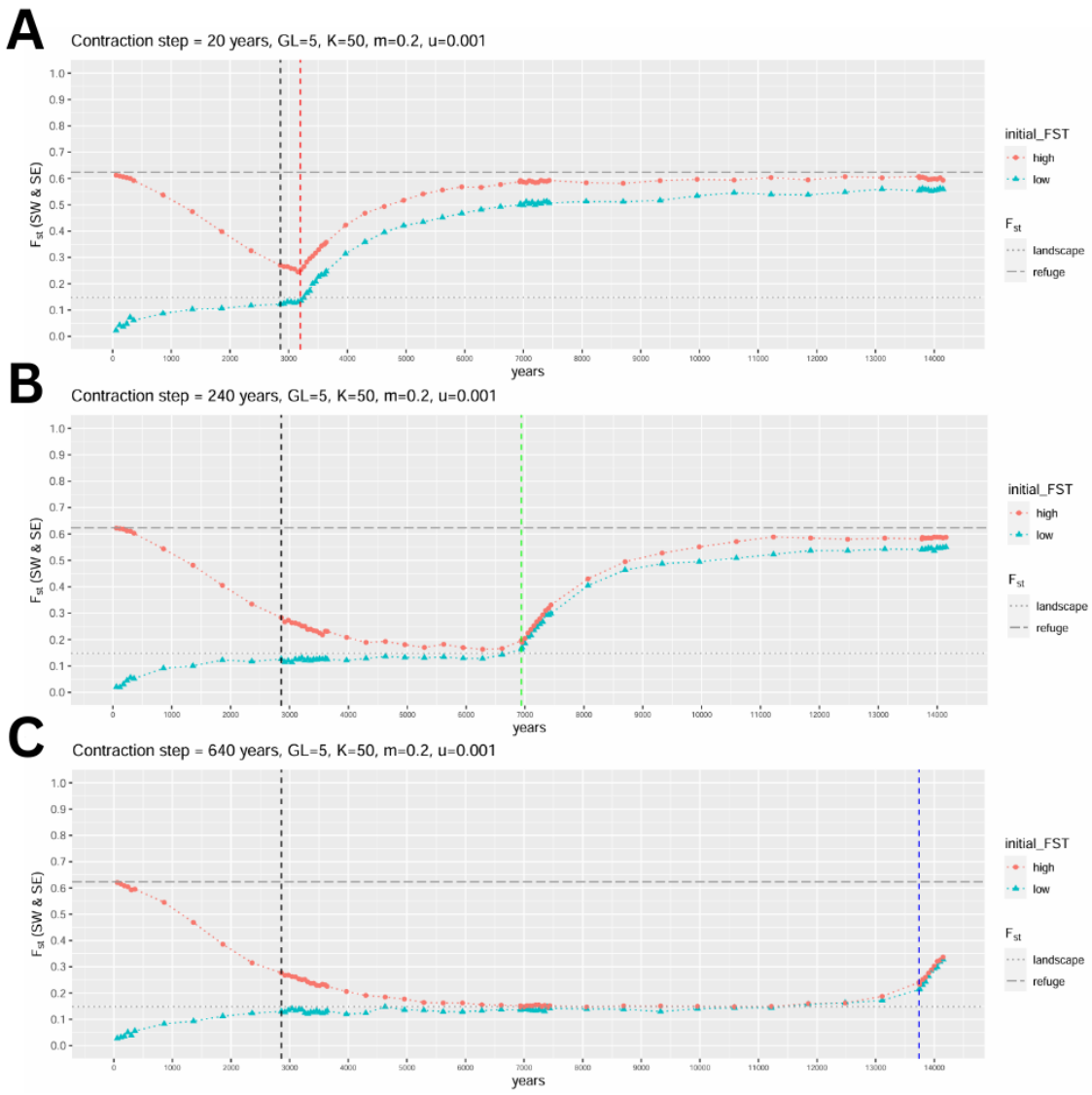

Figure 8 - **Dynamics of F_{st} for GL=5, in both "low F_{ST}" & "high F_{ST}" scenarios:** Each panel has identical parameters (GL=5, K=50, m=0.2 and u=0.001), only differing in the duration of each contraction step: Panel A takes 340 years to contract, Panel B takes 4080 years, and the last Panel takes 10880 years. For each panel there are 2 scenarios: in

blue, expansion of habitat started at t = 0 (low F_{st} scenario), whereas in red both refugia lived isolated from one another for 50k generations (high F_{st} scenario), and only after that period, expansion starts. In both panels, data is plotted as circles, squares, and tringles that represent the means measured in all 15 simulations. The vertical dashed black line (x = 2860 years) specifies the starting point of all those contractions and the following vertical-coloured lines indicate the end of contraction associated with that colour (red, green and blue respectively correspond to t=3200, t=6940 and t=13740). Horizontally dashed grey coloured lines represent the values of F_{st} at mutation-migration-drift equilibrium when the refugia are totally isolated (refuge) or when the entire landscape is occupied (landscape). Both equilibria are computed by maintaining stationary conditions throughout 100,000 generations.

3.3.2 Investigating the effect of connectivity between refugia

The previous subsection aimed to describe the patterns of F_{st} observed across various initial scenarios and contraction speeds, as described and tested in this work. In order to further comprehend this metric of genetic differentiation in the context of a two-cycle expansion and contraction, we decided to study one particularly interesting characteristic found: the increase of the fixation index metric as we get closer to the end of contraction. Looking at Figures 6 and 8 we see that, if F_{st} is relatively close to the equilibrium of the landscape, it will see an increase before the disconnection of both refugia (logically this increase is sped up as soon as contraction ends and refugia are physically isolated). Moreover, F_{st} stays constant for several years (for example in intermediate and slow contraction speeds), despite the habitat decreasing. This suggests that there is a set amount of demes that connect refugia, in order to have an impact in genetic differentiation between refugia.

In order to evaluate the impact on F_{st} from varying the amount of lines of demes that connect both refugia, we simulated an intermediate speed, with expansion starting at t=0, GL=5, K=50 and m=0.2. Four scenarios were simulated, each differing in the number of steps of horizontal demes left after contraction ends. The nr_deme_line = 0 corresponds to complete isolation between the refugia left after contraction ends (as in previous section), whereas the other three still have the SW and SE refugia connected by a certain amount of lines of demes ("nr_deme_lines" in Figure 9). Note that each line corresponds to 18 demes. Figure 9 shows that one line of horizontal demes between SW and SE refuge allows for migrants to exchange genetic information, and consequently generate a significant reduction in the observed values of F_{st} between SW and SE (F_{st} ~ 0.3). For each subsequent addition of step of horizontal demes, F_{st} between SW and SE refugia decreases. For nr_deme_line = 3, we observe that F_{st} between SE and SW is close to $F_{st}^{landscape}$. Any further subsequent addition of horizontal line of demes would have very little reduction on F_{st} between refugia. But there is still a relatively significant

decrease in F_{st} values from one to two and even to three lines of demes connecting the refugia,. These results suggest that, for every single scenario, there is a specific set number of deme lines establishing contact with SW and SE, that is enough to keep genetic information flowing and for F_{st} values to be near equilibria

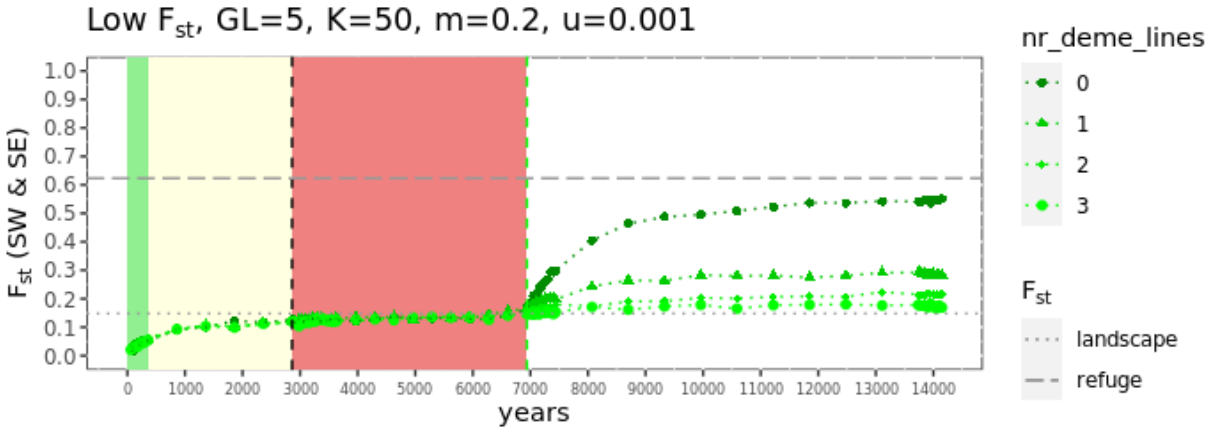

Figure 9 - **F_{st} between SW and SE refugia, varying the amount of demes connecting them both:** The plot corresponds to the "intermediate" contraction speed (t_{cont} = 4080 years), with expansion starting at t=0, GL=5, K=50 and m=0.2. Each horizontal line corresponds to "nr_deme_lines" which refers to the number of horizontal line of demes at the end of contraction

3.4 Effect of contraction speed on genetic diversity and differentiation at the end of longer contraction periods

So far, we examined three contraction durations: "fast" (t_{cont} = 340 years), "intermediate" (t_{cont} = 4080 years) and "slow" (t_{cont} = 10880 years). Simulations in this section intend to evaluate the veracity of previously obtained results, for longer contraction periods. We tested three slower speeds of contraction, where each contraction step took 1000, 2000 and 3000 years resulting in total contraction duration of 10880, 17000 and 34000 years respectively. We sampled data throughout the whole experiment, but focusing specially on the end of the 3 new contraction speeds (Table 2).

Our previous results pointed towards a bigger influence of starting F_{st} conditions for longer GLs and for shorter/faster contractions. On a short generation length species, like GL = 0.5 years, both plots for H_e and F_{st} (located on the left side of Figure 10) shows us that initial F_{st} conditions have very little impact in the dynamics of these same metrics, as the blue and red bars are nearly superimposed on each other. Furthermore, by inspecting both metrics on the right side of the panels (GL = 5 years) in Figure 10 we are able to notice that initial

conditions of genetic differentiation become more and more irrelevant as we increase the duration of habitat contraction (for contraction steps longer than 1000 years, lines representing the dynamics of H_e and F_{st} are starting to superimpose). This may indicate that if the habitat changes occur at a slower pace, the initial levels of genetic differentiation become less and less impactful.

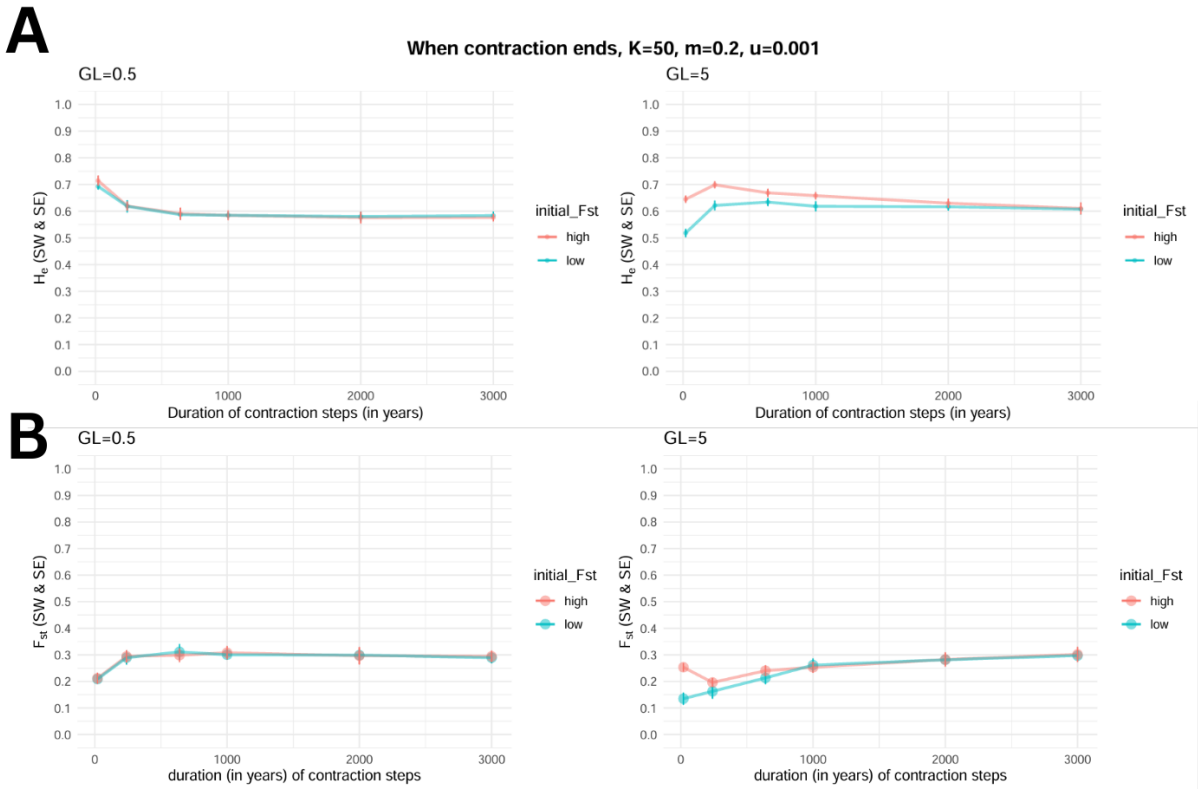

Figure 10 - **H_e and F_{st} between SW and SE refugia at the end of contraction as function of contraction duration:** The same parameters as previous simulations were considered. Plots on the left refer to GL=0.5 years while on the right they refer to GL=5 years. Previous time of contractions (t_{cont}) were 340, 4080, 10880 years, but slower contractions result in t_{cont} of 17000, 34000 and 51000 years. As a result, there are six timepoints in the x-axis that refer to the duration of a single contraction step: we have the 3 previous speeds at 20, 240, and 640 years, as well as the newly added 1000, 2000, and 3000 years. Panel A and Panel B show different metrics (H_e and F_{st}). Colours in both panels refer to initial different F_{st} conditions. Standard deviations were measured across 15 simulations and are presented as vertical lines on top of the plotted data points.

Table 2 — Summary of contraction speeds

Scenario	Time to eliminate one line of demes (contraction step)	Duration of contraction (t_{cont})	Where they were tested
"fast" contraction	20 years	340 years	Most simulations in this work
"intermediate" contraction	240 years	4080 years	
"slow" contraction	640 years	10880 years	
slower/longer contraction speeds	1000 years	17000 years	Longer Contractions
	2000 years	34000 years	
	3000 years	51000 years	

3.5 Confirming results for different carrying capacities and migration rates

So far, simulations have been varying in length of each generation, speed of contractions and initial genetic differentiation. Other than these parameters, simulations have maintained the same values for carrying capacity and for migration. Variance in life-history traits can end up showing totally different results from what we have witnessed so far. Thus, this section intends to explore dynamics of H_e and F_{st} and what happens at the end of contraction, with different K and m values.

We selected parameter values in accordance with the allometric relationship across mammalian species as described in Vishwakarma, R. et al., 2024. Species' body mass can be used to predict population density within a given area and their migration rates. Consequently, certain combinations of values align with these biological patterns, while others do not conform to such ecological principles (for more thorough explanation, check Vishwakarma, R. et al., 2024). Specifically, we choose two combinations of (K, m) values: i) (150, 0.059); ii) (400, 0.014). These parameter values correspond to the body mass of 76.34 and 20.65 kgs, respectively.

For expected heterozygosity, low vs high F_{st} scenarios did change the dynamics of H_e without necessarily changing the after-contraction plots (results in Figure 11). This is true due to the distance of measured H_e at the start of contraction, in relation to the equilibrium of landscape. As habitat is contracting, expected heterozygosity is still far away from reaching $H_e^{landscape}$ and for that reason genetic diversity keeps increasing until habitat finishes

contracting. Despite the change in dynamics, we still find ourselves inside the same second H_e scenario in Table 1 and as a result after contraction plots will keep the same shape.

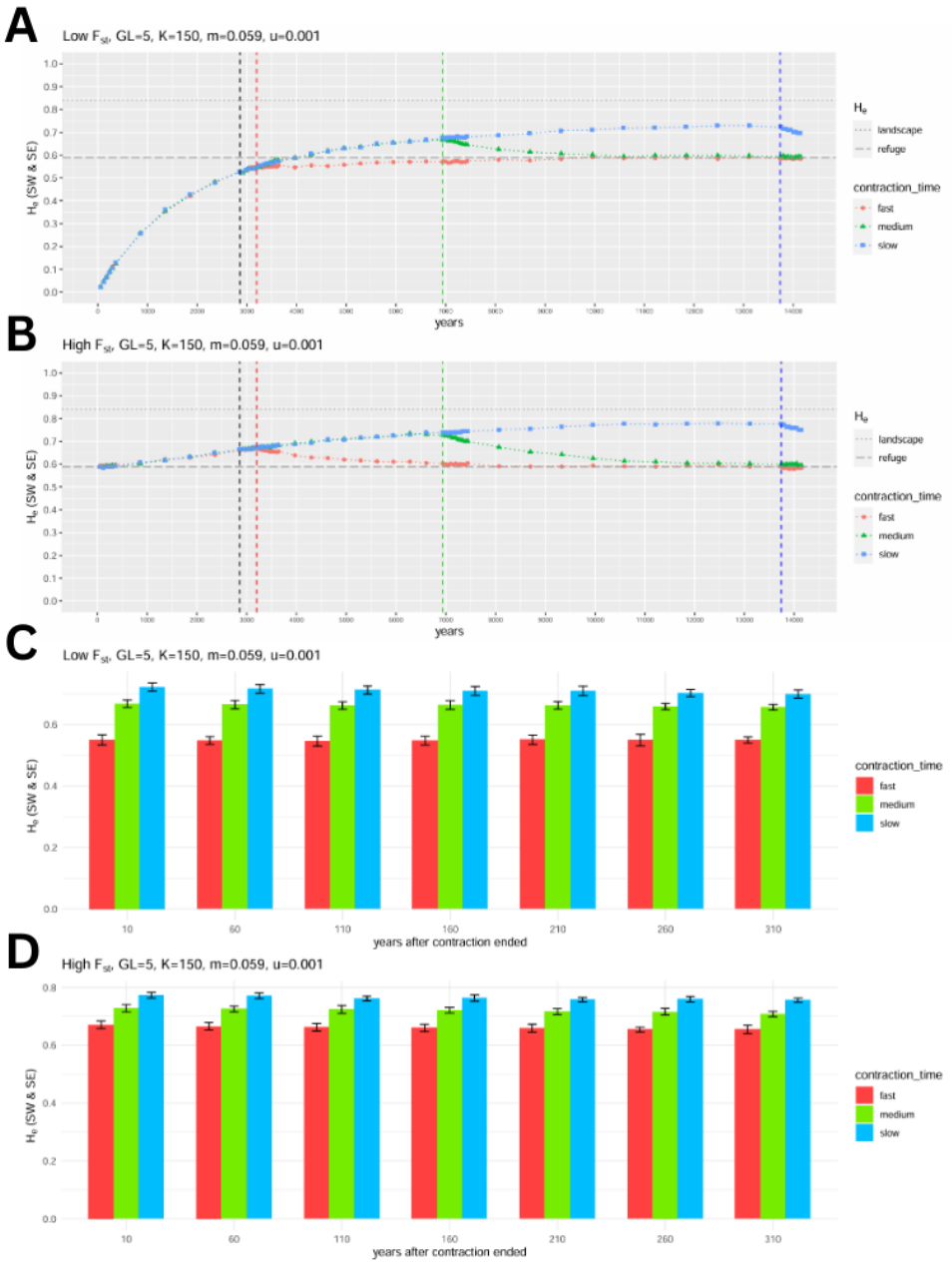

Figure 11 - **H_e metrics of K=150 & m=0.059 for both SW and SE refugia:** Simulations this time refer to K=150 & m=0.059, with Panel A and B touching on the dynamics throughout the expansion-contraction scenarios we have seen so far. Colour codes refer to the same 3 initial contraction speeds. Panel C and D are representations of expected heterozygosity x years after contraction has ended. Take note that these last 2 panels have very similar representations despite the respective dynamics looking very distinct.

When it comes to F_{st}, the picture is different as low and high F_{st} scenarios not only changed the dynamics appearance, but also the pattern of F_{st} after the end of contraction plots.

By examining the distance from F_{st} measured in the start of contraction relative to F_{st}^{refuge} and $F_{st}^{landscape}$, we observe that panels A and B are very distinct. For instance, in Panel A, since F_{st} between refuge < $F_{st}^{landscape}$, genetic differentiation keeps on increasing during contraction. Whereas in Panel B, $F_{st} \sim F_{st}^{refuge}$, F_{st} decreases and increases during contraction. Because genetic differentiation between refugia in low F_{st} scenario starts near landscape equilibrium (value measured at vertical black dotted line in Panel A), it will only go up from there, whereas genetic differentiation in high F_{st} scenario (value measured at vertical black dotted line in Panel B) starts off near refugia and comes down towards $F_{st}^{landscape}$. For the sake of simplicity, only results for F_{st} of K=150 & m=0.059 are shown in this section (Figure 12) but the observations are true for K=400 & m=0.014 as well (supplementary in Figure A.5).

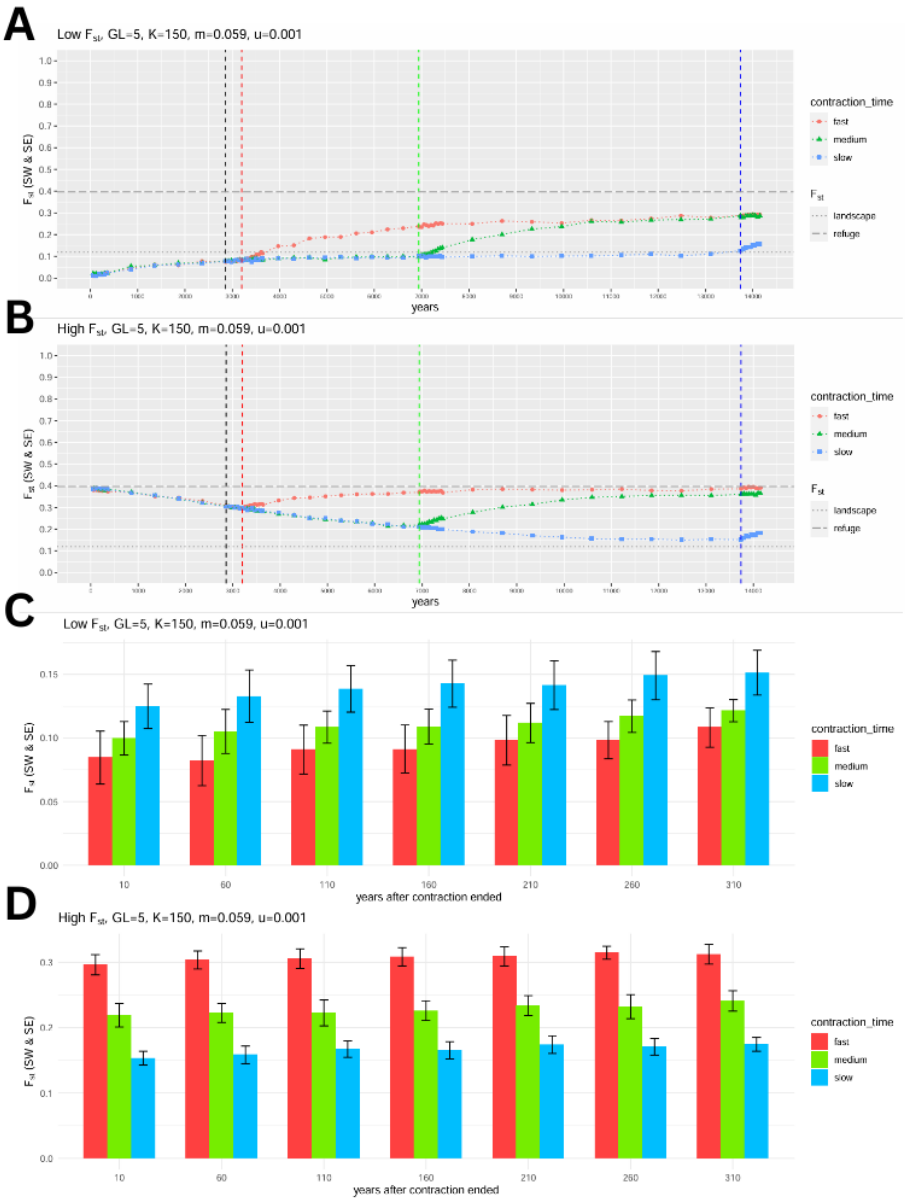

Figure 12 - **F_{st} metrics of K=150 & m=0.059 between SW and SE refugia:** Simulations refer to *K*=150 & *m*=0.059, with Panel A and B touching on the dynamics throughout the expansion-contraction scenarios, with different initial genetic differentiation values (F_{st}). Colour codes refer to the same 3 initial contraction speeds. Panel C and D are representations of expected F_{st} x years after contraction has ended. Contrasting with the results in the previous figure, these last 2 panels represent very different scenarios as Panel C has higher values for F_{st} in slower contraction speeds, when contraction ends, whereas in Panel D faster contraction speeds accumulate higher F_{st} values.

3.6 Northern refuge

We also wanted to investigate the influence of the geographical location of the refugium in our results. So far, results have been presented comparing Southwest and Southeast scenarios, but what happens when we have a Northern located refuge instead of Southeast? Are the results different or do they go in line with what was previously presented here?

One of the main differences is that in the SW and SE scenarios, both refugia came in contact immediately after expansion started. However, with a northern refuge, there is a delay between the start of the expansion phase and the connection between SW/SE and Northern refugia. Figure 13 represents the placement of the northern refuge within the landscape.

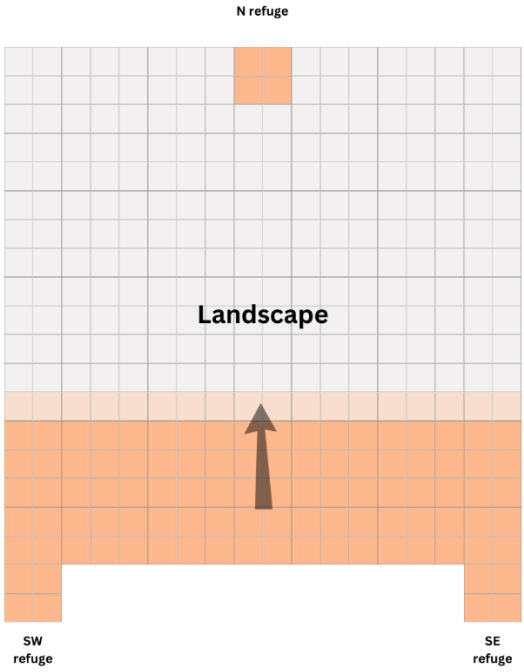

Figure 13 - **Simulated landscape with a Northern refuge:** The simulated 2D landscape is composed by a main area where there are 18 x 18 demes (grey), as well as two additional refuge areas (SW and SE refugia) of 2 x 2 demes. In addition, we also have a Northern located refuge inside of main area (four orange demes at the top. After certain number of years, the habitat has already been expanded, both refugia have been occupied as well as a reasonable amount of demes. The faded orange demes show the next expansion step (next demes to be colonized) and the arrow shows the direction of expansion. Note that the expansion starts in the SW and not in the North.

The H_e dynamics of SW (Figure 14) refuge resemble those shown in Figure 3A. In the N refuge, we observe that diversity increases during expansion and stationary phase, as expected. Notably, there is a sharp increase in diversity (around ~ 360 years) during expansion phase, due to contact between N refuge with population expanding from SW and SE refuge. Once

contraction starts, N refuge gets isolated, leading to a decrease in genetic diversity as it reduces to its mutation-drift-migration equilibrium. With faster contraction speeds, this isolation occurs earlier compared to slower contractions.

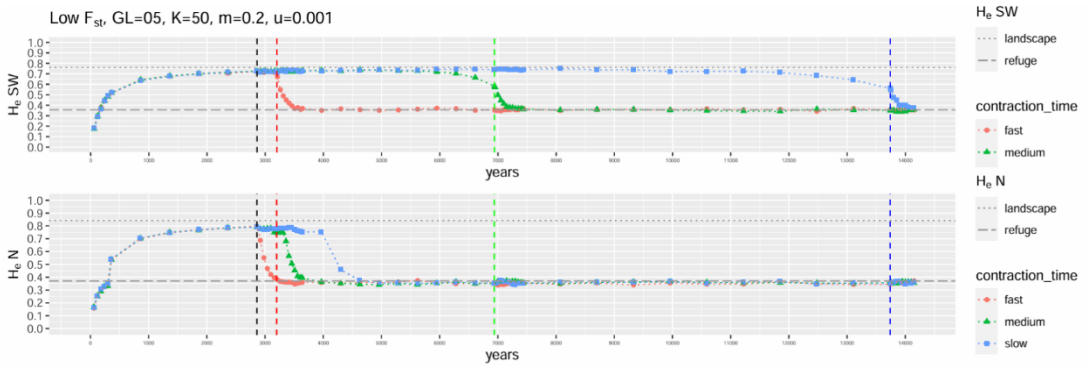

Figure 14 - **Separated dynamics of H_e comparing SW and N refugia for GL05:** The plot in the top half presents the metrics of H_e only in the Southwestern refugium, which are very similar to other dynamics present in this work. On the bottom half we have H_e for the Northern refugium, showing very abrupt decreases in relatively few years, even in slow contraction speeds (blue). This decrease comes from the fact that the 4 demes constituting the refugium get isolated from the rest of the landscape very early on.

Figure 15 represents the dynamics of F_{st} between SW and N refuge. During expansion and stationary phase, we observe the F_{st} increase in the first ~300 years. Since we start with a scenario where N and SW refuge were genetically similar, F_{st} between them is going to increase since habitat are disconnected. The moment habitats get connected, what follows is a sudden decrease in F_{st} between refugia. The evaluated metric stabilizes at $F_{st}^{landscape}$ in the stationary phase. After that, faster contractions are going to separate refugia at a faster rate and consequently rise the F_{st} values between N and SW also at a faster pace. As expected, the slower type of contraction (represented by a blue line) is going to take longer to perform the same separation, and genetic differentiation will not increase as rapidly as its faster counterparts. Nevertheless, there is a particularly interesting aspect in these results: for intermediate and slower contraction speeds, F_{st} values stays constant for a long period during contraction, before increasing again towards the end of contraction. For instance, for slow contraction curve in Figure 15 (blue curve), F_{st} stays constant from 5000 - 11500 years before increasing again. This plateau in F_{st} during contraction can be understood by H_e dynamics for SW and N refugia in Figure 14. We observe that during the 5000 - 11500 years period, H_e dynamics corresponding to slower contraction remain stable for both SW and N refuge. Considering the equation for fixation index as $F_{st} = 1 - \frac{H_s}{H_t}$, we witness that both parameters on the right (H_s and H_t) are also going to keep constant, leading to a plateau in the F_{st}.

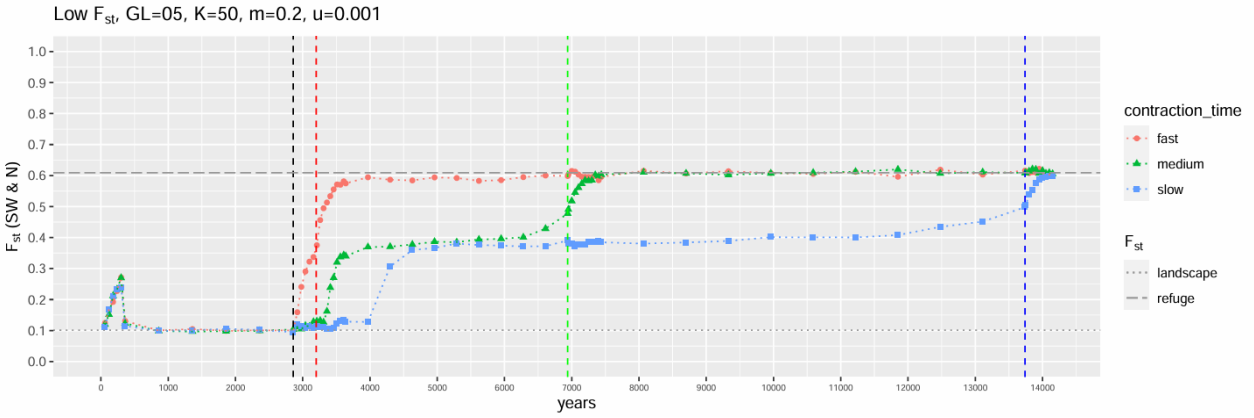

Figure 15 - **Dynamics of F_{st} comparing SW and N refugia for GL05:** Figure presents the F_{st} dynamics for a species with GL=05, across 3 different contraction speeds. Data is plotted as circles, squares, and tringles that represent the means measured in all 15 simulations. Their shapes and colours vary depending on the speed of contraction (or duration of contraction). The vertical dashed black (x = 2860 years) specifies the starting point of all those contractions and the following vertical-coloured lines indicate the end of contraction associated with that colour (red, green and blue respectively correspond to t=3200, t=6940 and t=13740). Horizontally dashed grey coloured lines represent the values of F_{st} at mutation-migration-drift equilibrium when the refugia are totally isolated (refuge) or when the entire landscape is occupied (landscape). Both equilibria are computed by maintaining stationary conditions throughout 100,000 generations.

The following figure (Figure 16) depicts this idea that both H_e (light and dark blue) stay in constant values from a certain point in time, leading to a plateau of F_{st}, even when habitat is contracting. This observation also correlates with the idea presented in 3.3.2 that there is a certain amount of demes that are enough to sustain certain amounts of diversity, not affecting the oscillation of genetic diversity metrics. The disconnection from N and SW is enough for F_{st} to slightly increase but it plateaus as long as H_e in refugia also becomes stable. There is a breaking point from where lines of demes contribute a lot for the maintenance of H_e and as soon as they start to be eliminated, F_{st} will rise (12000 years).

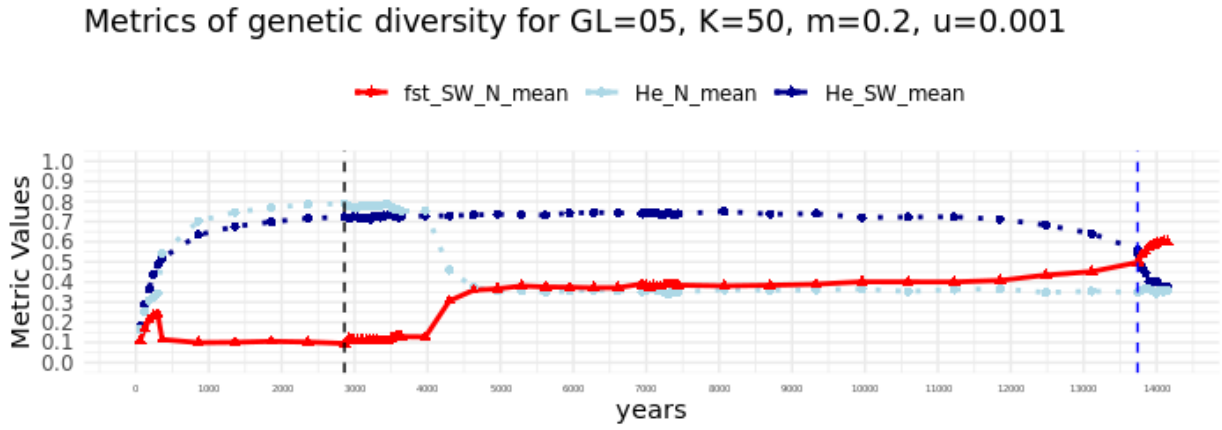

Figure 16 - **Dynamics of H_e and F_{st} altogether, for SW and N refugia (GL05 and slow contraction):** Light and dark blue lines represent genetic diversity (H_e) measured at the N and SW refugia, respectively. The red line corresponds to the genetic differentiation between both refugia throughout expansion and contraction of habitat. This case corresponds to a "slow" contraction (t_{cont} = 10880 years).

4

Discussion

Natural populations undergo changes in population size, geographical distribution and connectivity across various spatial and temporal scales. Understanding how these processes affect the genetic variation over time and across species is crucial for assessing the impacts of past environmental changes on genetic diversity, as well as predicting how future changes could affect it. This work aimed to investigate dynamics of genetic diversity (H_e, *MNA*) and differentiation (F_{st} between refugia) over multiple expansion-contraction cycles such as glacial-interglacial cycles, using spatial simulations across species with different traits. The results help clarify how initial F_{st} between refugia impacts genetic diversity and differentiation dynamics for different contraction speeds. Additionally, this work highlights the effect of contraction speed on diversity and differentiation at the end of contraction.

In section 3.1 of Results, we first tested how expansion from two refuge areas affect the dynamics of genetic diversity. Despite the complexity of dynamics, we found that the trajectory of genetic diversity could be predicted, by comparing the genetic diversity of the populations at the beginning of the simulated contractions to $H_e^{landscape}$ and H_e^{refuge}. These conclusions match with those from Vishwakarma et al., 2024. We then simulated scenarios where F_{st} between the refugia was higher (F_{st} ~ 0.63) and compared dynamics of genetic diversity with low initial F_{st} scenarios. We found similar temporal dynamics between the two F_{st} scenarios; however, the magnitude of genetic diversity was lower in scenarios with low initial F_{st} compared to high F_{st}.

Our major finding is related to the dynamics of F_{st} between refuge areas for different contraction speeds and initial F_{st} scenarios. We found three distinct temporal dynamics of F_{st} during contraction phase: *i)* increase *ii)* remains stable and then increase *iii)* could decrease initially and then increase. Similar to genetic diversity, dynamics during contraction can be

predicted by comparing the F_{st} between the refugia to the expected F_{st} between refugia at mutation-migration-drift equilibrium. We also identified contrasting effects of contraction speed on F_{st} between refuge at the end of contraction. The first pattern shown in Figure 7A, low initial F_{st} between refuge shows faster contraction scenarios lead to less genetic differentiation between the two refugia when compared to slower contractions. This result aligns with Arenas et al., 2012 and Vishwakarma et al., 2024. On the other hand, for high F_{st} between refugia (Figure 7B), fast and slow contraction speeds show higher values of F_{st} when compared to intermediate speeds. Sections 3.4 and 3.5 from Results serve to confirm the findings obtained thus far for longer contraction speeds and different parameters of carrying capacity and migration rate. In the final section we investigated the impact of having a differently located refuge (northern refuge individuals were sampled instead of southeast) on genetic diversity and genetic differentiation, across different species. Once again, even in the presence of an intricate scenario like this, we show that dynamics of genetic differentiation between refugia can still be predicted based on values at the mutation-drift-migration equilibrium. Additionally, we found that for intermediate and slower contraction speeds (4080 and 10880 years) genetic differentiation between refugia can remain constant for long periods of time due to the stabilisation of genetic diversity (H_e) within each refugium, despite the changes in habitat.

Overall, our simulations allow to predict dynamics of genetic diversity (H_e, MNA) and differentiation (F_{st} between refuge) over multiple expansion-contraction cycles, regardless of the complexity of the scenario presented, as long as we have access to those same metrics at equilibria. This could be highly significant for the conservation of endangered species residing in isolated refugia.

In next subsections we discuss and clarify some consequences of our work.

4.1.1 Temporal dynamics of F_{st} between refugia in contracting populations

Our results allow for the categorization of distinct "scenarios", depending on the distance from F_{st} between refugia at the beginning of contraction, with respect to equilibria of genetic differentiation measured as F_{st}^{refuge} and $F_{st}^{landscape}$. This classification allows to predict the behaviour of genetic differentiation between refugia, similarly to what was previously done for H_e (Table 1). Essentially, dynamics of F_{st} can be described according to Table 3:

Table 3 — Scenarios of F_{st} at the beginning of contraction

Scenario of F_{st} when contraction starts	Implications	Examples
$F_{st} \simeq F_{st}^{landscape}$	Population is at mutation-drift-migration equilibrium before contraction starts. Elimination of some demes will increase overall F_{st} although the big leap in this value will be more pronounced during and after the last line of demes that connects refugia is eliminated. Faster contraction speeds will have lower F_{st} values at the end of contraction compared to slower contraction speeds.	Figure 6A & 7A
$F_{st}^{landscape} < F_{st} < F_{st}^{refuge}$	F_{st} keeps decreasing for long periods of time despite habitat reduction, before increasing again towards the end of contraction. Regarding the effect of contraction speed on F_{st} at the end of contraction, intermediate speed shows lower F_{st} values at the end of contraction compared to fast or slow contraction speeds.	Figure 6B & 7B
$F_{st} \simeq F_{st}^{refuge}$	Decrease of F_{st} towards values of $F_{st}^{landscape}$, regardless of habitat and population reduction. The increase comes from the moment that refugia are separated. After contraction plots will have faster contractions with the highest F_{st} values.	Figure 12B & 12D

In addition to being able to predict dynamics of F_{st}, the creation of two different starting conditions (low or high F_{st}) enable us to understand to what extent those conditions influence different speeds of contraction as well as species with different generation lengths. Results in Figures 8 and 10 show how the effect of initial F_{st} becomes less pronounced for *i)* species with shorter generation lengths and *ii)* longer contraction periods. If we consider two

different mammalian species such as mouse (shorter GL) and a wolf (longer GL) for example, under the same scenarios of habitat changes, our results suggest that they would be very differently affected by ancient events. For temperate species, glacial-interglacial periods had long periods of isolation which could result in refugia with high F_{st} values. As habitat expands, we expect that genetic diversity in mice refugia to be less affected by initial F_{st} scenario compared to genetic diversity in wolves refugia. Since wolves take longer to reproduce, differences in initial F_{st} scenarios will have greater impact on their genetic diversity and differentiation dynamics. Moreover, longer contraction durations allow populations with longer GL's to have similar dynamics of diversity and differentiation. For this reason, these types of ancient events loose relevancy as contractions periods increase: populations are under the same conditions for longer and become more similar in dynamics of F_{st}.

These observations of differences across species with different GL go in line with the idea the findings from classical paper by Taberlet, et al., 1998, that showed that phylogeographic congruence was low across different species in Europe . Altogether, the findings emphasize that the genetic responses of species to habitat contractions and expansions are not uniform and are influenced by a combination of ecological, geographical, and historical factors. Our simulations also highlight uneven responses to expansion cycles, only by varying a small amount of parameters like generation length. This leads to the idea that this type of movements is very complex but studies like ours can help unravel dynamics of genetic diversity and differentiation in refugia populations.

4.1.2 Temporal dynamics of F_{st} with northern refuge

The sampling of individuals belonging to a northern refuge aimed to explore the temporal dynamics of genetic diversity and differentiation in samples from other regions within the landscape. Our observations for a more northern sited refuge, agree with what was found for SW and SE: F_{st} measured at the beginning of contraction will allow to estimate the trajectory of dynamics of genetic differentiation. However, we observe that F_{st} between SW and N refugia could remain steady during contraction for long durations. We also found that this is due to the fact that H_e values remain steady within each refugia during that period. These results show that connectivity between refugia is fundamental and that geographic location of refugia is an equally important factor to consider when analysing consequences of changing habitats in metrics of genetics for different species.

4.1.3 Limitations of the study

Our results highlighted the dynamics and patterns of F_{st} between different refuge areas (SW,SE and N) for different scenarios and parameters. Although we simulated many different parameters and scenarios, we still made many simplifying assumptions. For instance, the creation of two exactly identical populations (one in each refugia), with identical allele frequencies for each locus, in practical terms means that there was a strong bottleneck that created two exactly identical populations. Unfortunately, this is something quite unrealistic to happen to populations in nature. Nevertheless, the strategy of keeping both refugia isolated and letting individuals reproduce within them allowed to control the level of genetic differentiation populations have before the second cycle started. This allowed for a better characterization of F_{st} dynamics as well as describing what happens to these values towards the end of contraction. Despite taking into consideration allometric relationships in species, there were a considerable amount of variables kept constant, simplifying our initial assumptions. We assumed homogeneous carrying capacities, migration, mutation, and growth rates. In addition, we assumed random mating within demes for all species across the landscape and treated co-distributed species as ecologically independent, without accounting for species interactions. The connectivity between the refugium and the population across the landscape was assumed to remain constant throughout the simulation. The model assumes that the speed of range shifts is imposed by environmental changes that occur in constant intervals of time, which may not always be the case in real-world scenarios. We only simulated species with non-overlapping generations, nonetheless theoretical studies suggest that overlapping generations increase the time needed to reach equilibrium (Lloyd et al., 2013). Moreover, long distance dispersal events were not considered. These are known to affect patterns of diversity after a range expansion (Nichols and Hewitt, 1994; Ibrahim et al., 1996; Ray and Excoffier, 2010) and could almost certainly impact contractions as well.

These limitations suggest that while the study provides valuable insights, further research is needed to explore these dynamics in more complex and realistic contexts.

5

BIBLIOGRAPHY

Alvarez, G., Quinteiro, C., & C., F. (2011). Inbreeding and Genetic Disorder. *Advances in the Study of Genetic Disorders*. https://doi.org/10.5772/18373

Arenas, M., Ray, N., Currat, M., & Excoffier, L. (2012). Consequences of range contractions and range shifts on molecular diversity. *Molecular Biology and Evolution*, *29*(1), 207–218. https://doi.org/10.1093/molbev/msr187

Bartáková, V., Reichard, M., Janko, K., Polačik, M., Blažek, R., Reichwald, K., Cellerino, A., & Bryja, J. (2013). Strong population genetic structuring in an annual fish, Nothobranchius furzeri, suggests multiple savannah refugia in southern Mozambique. *BMC Evolutionary Biology*, *13*(1). https://doi.org/10.1186/1471-2148-13-196

Biodiversity loss: what is causing it and why is it a concern? | Topics | European Parliament. (n.d.). Topics | European Parliament. https://www.europarl.europa.eu/topics/en/article/20200109STO69929/biodiversity-loss-what-is-causing-it-and-why-is-it-a-concern

Brodie, E. D. (2007). Population size is not genetic quality. *Animal Conservation*, *10*(3), 288–290. https://doi.org/10.1111/j.1469-1795.2007.00134.x

Brüniche-Olsen, A., Kellner, K. F., Belant, J. L., & Dewoody, J. A. (2021). Life-history traits and habitat availability shape genomic diversity in birds: Implications for conservation. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1961). https://doi.org/10.1098/rspb.2021.1441

Canteri, E., Fordham, D. A., Li, S., Hosner, P. A., Rahbek, C., & Nogués-Bravo, D. (2021). IUCN Red List protects avian genetic diversity. *Ecography*, *44*(12), 1808–1811. https://doi.org/10.1111/ecog.05895

CBD Secretariat. (2011). Living in harmony with nature Invasive Alien Species. *Convention on Biological Diversity*, *10*(3), 28–46.

Currat, M., Ray, N., & Excoffier, L. (2004). SPLATCHE: A program to simulate genetic diversity taking into account environmental heterogeneity. *Molecular Ecology Notes*, *4*(1), 139–142. https://doi.org/10.1046/j.1471-8286.2003.00582.x

Darwin, C. & Kebler, L. (1859) On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life. London: J.

De Kort, H., Prunier, J. G., Ducatez, S., Honnay, O., Baguette, M., Stevens, V. M., & Blanchet, S. (2021). Life history, climate and biogeography interactively affect worldwide genetic diversity of plant and animal populations. Nature Communications, 12(1). https://doi.org/10.1038/s41467-021-20958-2

DeAngelis, D. L., & Yurek, S. (2017). Spatially Explicit Modeling in Ecology: A Review. *Ecosystems*, *20*(2), 284–300. https://doi.org/10.1007/s10021-016-0066-z

Finn, C., Grattarola, F., & Pincheira-Donoso, D. (2023). More losers than winners: investigating Anthropocene defaunation through the diversity of population trends. *Biological Reviews*, *98*(5), 1732–1748. https://doi.org/10.1111/brv.12974

Garner, B. A., Hoban, S., & Luikart, G. (2020). IUCN Red List and the value of integrating genetics. *Conservation Genetics*, *21*(5), 795–801. https://doi.org/10.1007/s10592-020-01301-6

Garnier, J., & Lafontaine, P. (2022). Life history traits and dispersal shape neutral genetic diversity in metapopulations. In *Journal of Mathematical Biology* (Vol. 84, Issue 6). Springer Berlin Heidelberg. https://doi.org/10.1007/s00285-022-01749-9

Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes, 5(1), 184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x

Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, *58*(3), 247–276. https://doi.org/10.1006/bijl.1996.0035

Hoban, S., Bertorelle, G., & Gaggiotti, O. E. (2012). Computer simulations: Tools for population and evolutionary genetics. *Nature Reviews Genetics*, *13*(2), 110–122. https://doi.org/10.1038/nrg3130

Hoban, S., Bruford, M. W., da Silva, J. M., Funk, W. C., Frankham, R., Gill, M. J., Grueber, C. E., Heuertz, M., Hunter, M. E., Kershaw, F., Lacy, R. C., Lees, C., Lopes-Fernandes, M., MacDonald, A. J., Mastretta-Yanes, A., McGowan, P. J. K., Meek, M. H., Mergeay, J., Millette, K. L., … Laikre, L. (2023). Genetic diversity goals and targets have improved, but remain insufficient for clear implementation of the post-2020 global biodiversity framework. Conservation Genetics, 24(2), 181–191. https://doi.org/10.1007/s10592-022-01492-0

Hofreiter, M., & Stewart, J. (2009). Ecological Change, Range Fluctuations and Population Dynamics during the Pleistocene. *Current Biology*, *19*(14), R584–R594. https://doi.org/10.1016/j.cub.2009.06.030

Hopkins, G. R., Gall, B. G., French, S. S., & Brodie, E. D. (2012). Interfamily variation in amphibian early life-history traits: Raw material for natural selection? *Ecology and Evolution*, *2*(7), 1637–1643. https://doi.org/10.1002/ece3.287

Ibrahim, K. M., Nichols, R. A., & Hewitt, G. M. (1996). Spatial patterns of genetic variation generated by different forms of dispe...: Lund University Libraries. Heredity, 77(November 1995), 282–291. http://eds.b.ebscohost.com.ludwig.lub.lu.se/eds/detail?sid=f178cf98-8cc5-4aeb-b6be-63d9c5c9e64b%40sessionmgr115&vid=1&hid=102&bdata=JnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3D%3D#db=a9h&AN=8854253

Ibrahim, M. L. A., Zango, O., Inoussa, M. M., Moussa, S., & Bakasso, Y. (2020). Multivariate Statistical Methods Used in Population Genetics. *Asian Journal of Probability and Statistics*, *7*(1), 59–70. https://doi.org/10.9734/ajpas/2020/v7i130173

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics, 24(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Kimura, M., & Weiss, G. H. (1964). the Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation With Distance. Genetics, 49(4), 561–576. https://doi.org/10.1093/genetics/49.4.561

Kimura, M., & Weiss, G. H. (1964). the Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation With Distance. Genetics, 49(4), 561–576. https://doi.org/10.1093/genetics/49.4.561

Klopfstein, S., Currat, M., & Excoffier, L. (2006). The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution*, *23*(3), 482–490. https://doi.org/10.1093/molbev/msj057

Kuipers, K. J. J., Hilbers, J. P., Garcia-Ulloa, J., Graae, B. J., May, R., Verones, F., Huijbregts, M. A. J., & Schipper, A. M. (2021). Habitat fragmentation amplifies threats from habitat loss to mammal diversity across the world's terrestrial ecoregions. *One Earth*, *4*(10), 1505–1513. https://doi.org/10.1016/j.oneear.2021.09.005

Lloyd, M. W., Campbell, L., & Neel, M. C. (2013). The Power to Detect Recent Fragmentation Events Using Genetic Differentiation Methods. PLoS ONE, 8(5). https://doi.org/10.1371/journal.pone.0063981Alcala, N., & Vuilleumier, S. (2014). Turnover and accumulation of genetic diversity across large time-scale cycles of isolation and connection of populations. Proceedings of the Royal Society B: Biological Sciences, 281(1794). https://doi.org/10.1098/rspb.2014.1369

Lloyd, M. W., Campbell, L., & Neel, M. C. (2013). The Power to Detect Recent Fragmentation Events Using Genetic Differentiation Methods. PLoS ONE, 8(5). https://doi.org/10.1371/journal.pone.0063981

Maié, T., Rasteiro, R., Curbelo, D., Chenin, N., Viaud, H., Bouttier, P.-A., Monier, D., Sousa, V., Thébaud, C., & Chikhi, L. (2017). SINS: forward simulation of individuals through time and space.

Nature's dangerous decline 'Unprecedented' species extinction rates 'Accelerating.' (n.d.). UN Environment. https://www.unep.org/news-and-stories/press-release/natures-dangerous-decline-unprecedented-species-extinction-rates

Nordborg, M. (2019). Coalescent theory. *Handbook of Statistical Genomics*, *1*, 145–175. https://doi.org/10.1002/9781119487845.ch5

Norderhaug, K. M., Knutsen, H., Filbee-Dexter, K., Sodeland, M., Jorde, P. E., Wernberg, T., Oomen, R., & Moland, E. (2024). The International Union for Conservation of Nature Red List does not account for intraspecific diversity. ICES Journal of Marine Science, 81(5), 815–822. https://doi.org/10.1093/icesjms/fsae039

Petit-Marty, N., Vázquez-Luis, M., & Hendriks, I. E. (2021). Use of the nucleotide diversity in COI mitochondrial gene as an early diagnostic of conservation status of animal species. *Conservation Letters*, *14*(1), 1–7. https://doi.org/10.1111/conl.12756

Pfeifer, M., Lefebvre, V., Peres, C. A., Banks-Leite, C., Wearn, O. R., Marsh, C. J., Butchart, S. H. M., Arroyo-Rodríguez, V., Barlow, J., Cerezo, A., Cisneros, L., D'Cruze, N., Faria, D., Hadley, A., Harris, S. M., Klingbeil, B. T., Kormann, U., Lens, L., Medina-Rangel, G. F., … Ewers, R. M. (2017). Creation of forest edges has a global impact on forest vertebrates. Nature, 551(7679), 187–191. https://doi.org/10.1038/nature24457

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rasteiro, R., Bouttier, P. A., Sousa, V. C., & Chikhi, L. (2012). Investigating sex-biased migration during the neolithic transition in Europe, using an explicit spatial simulation framework. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1737), 2409–2416. https://doi.org/10.1098/rspb.2011.2323

Ray, N., & Excoffier, L. (2010). A first step towards inferring levels of long-distance dispersal during past expansions. Molecular Ecology Resources, 10(5), 902–914. https://doi.org/10.1111/j.1755-0998.2010.02881.x

Rivera-Ortíz, F. A., Aguilar, R., Arizmendi, M. D. C., Quesada, M., & Oyama, K. (2015). Habitat fragmentation and genetic variability of tetrapod populations. Animal Conservation, 18(3), 249–258. https://doi.org/10.1111/acv.12165

Rowe, K. C., Heske, E. J., Brown, P. W., & Paige, K. N. (2004). Surviving the ice: Northern refugia and postglacial colonization. *Proceedings of the National Academy of Sciences of*

the United States of America, *101*(28), 10355–10359. https://doi.org/10.1073/pnas.0401338101

Schlaepfer, D. R., Braschler, B., Rusterholz, H. P., & Baur, B. (2018). Genetic effects of anthropogenic habitat fragmentation on remnant animal and plant populations: a meta-analysis. Ecosphere, 9(10). https://doi.org/10.1002/ecs2.2488

Schmidt, C., Hoban, S., Hunter, M., Paz-Vinas, I., & Garroway, C. J. (2023). Genetic diversity and IUCN Red List status. *Conservation Biology*, *37*(4), 1–10. https://doi.org/10.1111/cobi.14064

Sebbenn, A. M., Carvalho, A. C. M., Freitas, M. L. M., Moraes, S. M. B., Gaino, A. P. S. C., Da Silva, J. M., Jolivet, C., & Moraes, M. L. T. (2011). Low levels of realized seed and pollen gene flow and strong spatial genetic structure in a small, isolated and fragmented population of the tropical tree Copaifera langsdorffii Desf. *Heredity*, *106*(1), 134–145. https://doi.org/10.1038/hdy.2010.33

Sgarlata, G. M., Maié, T., De Zoeten, T., Rasteiro, R., & Chikhi, L. (2022a). The effect of habitat loss and fragmentation on isolation-by-distance and time. *BioRxiv*. https://doi.org/10.1101/2022.10.26.513874

Sgarlata, G. M., Maié, T., de Zoeten, T., Rasteiro, R., & Chikhi, L. (2022b). On the genetic consequences of habitat contraction: edge effects and habitat loss. *BioRxiv*, 1–65.

Smith, J. M., & Slatkin, M. (1972). The Stability of Predator-Prey Systems Author (s): J . Maynard Smith and M . Slatkin Published by : Ecological Society of America Stable URL : http://www.jstor.org/stable/1934346 . THE STABILITY OF PREDATOR-PREY SYSTEMS '. Ecology, 54(2), 384–391.

Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., & Cosson, J. F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, *7*(4), 453–464. https://doi.org/10.1046/j.1365-294x.1998.00289.x

The IUCN red list of threatened species. (n.d.). IUCN Red List of Threatened Species. https://www.iucnredlist.org/about/background-history

The State of the World's Forests 2020. (2020). In FAO and UNEP eBooks. https://doi.org/10.4060/ca8642en

United Nations Environment Programme. (n.d.). Five drivers of the nature crisis. UNEP. https://www.unep.org/news-and-stories/story/five-drivers-nature-crisis

United Nations Environment Programme. (n.d.-a). A roadmap towards a circular plastics economy. UNEP. https://www.unep.org/news-and-stories/speech/roadmap-towards-circular-plastics-economy

Vega, R., Fløjgaard, C., Lira-Noriega, A., Nakazawa, Y., Svenning, J. C., & Searle, J. B. (2010). Northern glacial refugia for the pygmy shrew sorex minutus in europe revealed

by phylogeographic analyses and species distribution modelling. *Ecography*, *33*(2), 260–271. https://doi.org/10.1111/j.1600-0587.2010.06287.x

Vishwakarma, R., Sgarlata, G. M., Soriano-Panos, D., Rasteiro, R., Maie, T., Paixao, T., Tournebize, R., & Chikhi, L. (2024). Life history traits influence the dynamics of genetic diversity in a refugium population undergoing expansion and contraction. BioRxiv, 2024.03.28.586951. https://doi.org/10.1101/2024.03.28.586951

Walter, G. M., Monro, K., Terranova, D., La Spina, E., Majorana, M., Pepe, G., Clark, J., Cozzolino, S., Cristaudo, A., Hiscock, S. J., & Bridle, J. (2024). Environmental effects on genetic variance are likely to constrain adaptation in novel environments. Evolution Letters, 8(3), 374–386. https://doi.org/10.1093/evlett/qrad065

Yuan, X., Miller, D. J., Zhang, J., Herrington, D., & Wang, Y. (2012). An overview of population genetic data simulation. *Journal of Computational Biology*, *19*(1), 42–54. https://doi.org/10.1089/cmb.2010.0188

A

SUPPLEMENTARY MATERIALS

A.1 Figures complementary to the Results section

In this section of the Supplementary Materials, I am showcasing figures that were mentioned in the Results section, merely to complement the work but not of extreme importance to comprehend it.

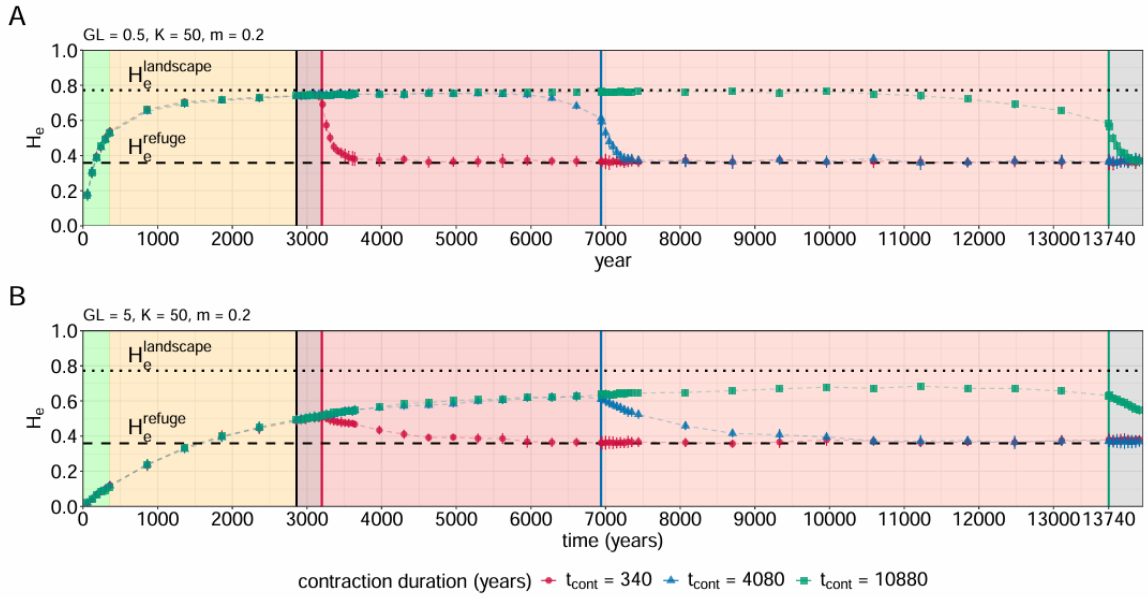

Figure A.1 - **Dynamics of genetic diversity in the Southwest (SW) refugium population:** Temporal evolution of H_e in the SW refugium population at discrete times. Two generation lengths are considered: GL=0.5years (Panel A) and GL=5years (Panel B). In both panels, each discrete dot (square, triangle or circle) represents the means while the error bars represent the standard deviations measured across 15 simulations. The discrete timepoints are connected by dashed lines and their symbols and colours represent the duration of the contraction phase for three scenarios, corresponding to *tcont* = 340 years (red circles), *tcont* = 4080 years (blue triangles) and *tcont* = 10880 years (green squares) respectively. The dotted (dashed) horizontal line represents the value of H_e at mutation-migration-drift equilibrium when the refugium (entire landscape) is occupied $H_e^{refugium}$ e ($H_e^{landscape}$). Both equilibria are computed

by keeping stationary conditions across 100,000 generations. The four vertical lines, from left to right, represent the following events respectively: i) start of the contraction phase (black, time=2860 years for all scenarios) and end of the contraction phase for (ii) *tcont* = 340 years(red, time=3200 years) iii) *tcont*=4080 years (blue, time=6940 years) and iv) *tcont*=10880 years (green, time=13740 years). The other parameter values for both plots are (K, m) = (50,0.2). Panel B shows that when contraction is slow (here 10880 years), H_e can keep increasing for more than 7000 years after contraction started and keep high and stable levels of genetic diversity that are much higher than H_e^{refuge} for more than 10000 years.

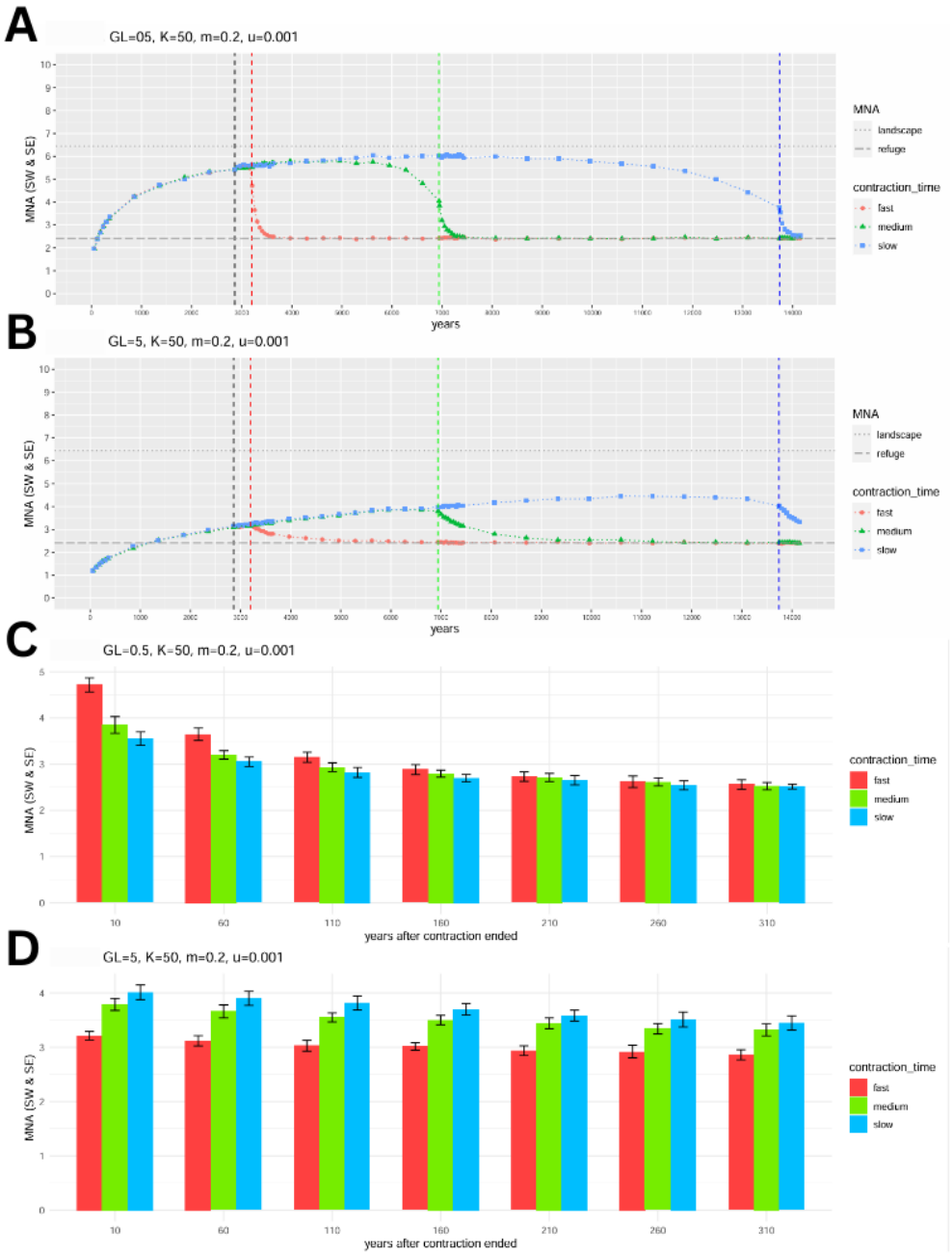

Figure A.2 - **Dynamics of Mean Number of Alleles (MNA) in the Southwest and Southeast (SW and SE) refugium population:** Same scenarios as depicted in Figures 3 and 4 but now measuring MNA instead of H_e

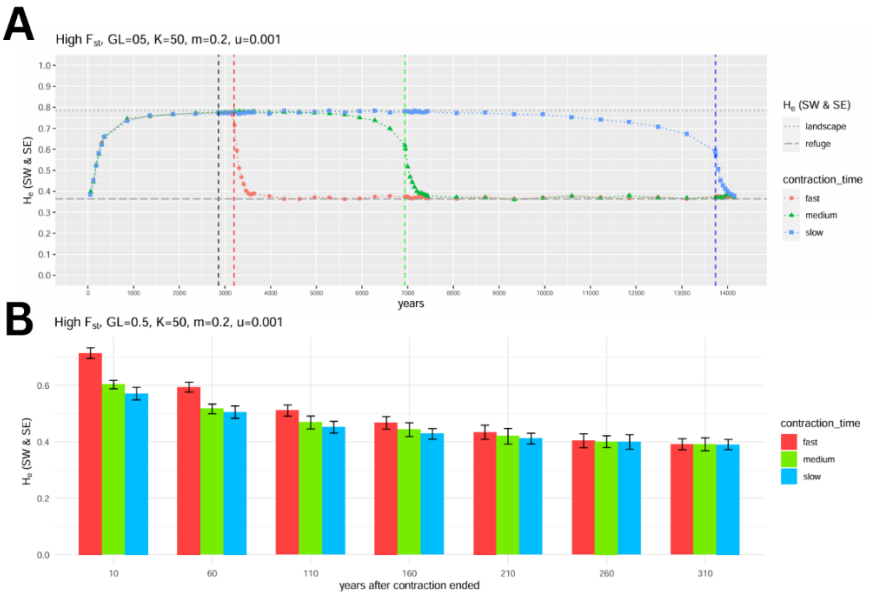

Figure A.3 - **Dynamics of H_e (MNA) in the Southwest and Southeast (SW and SE) refugia populations, for GL=05, in a "High F_{st}" scenario:** Same scenarios as depicted in Figure 3 and 4 but now populations have been reproducing inside each of their respective refugia, hence why "High F_{st}". The differences between starting conditions of genetic diversity do not impact the dynamics for this particular GL=05 (compare these results to GL=05 in Figures 3 and 4)

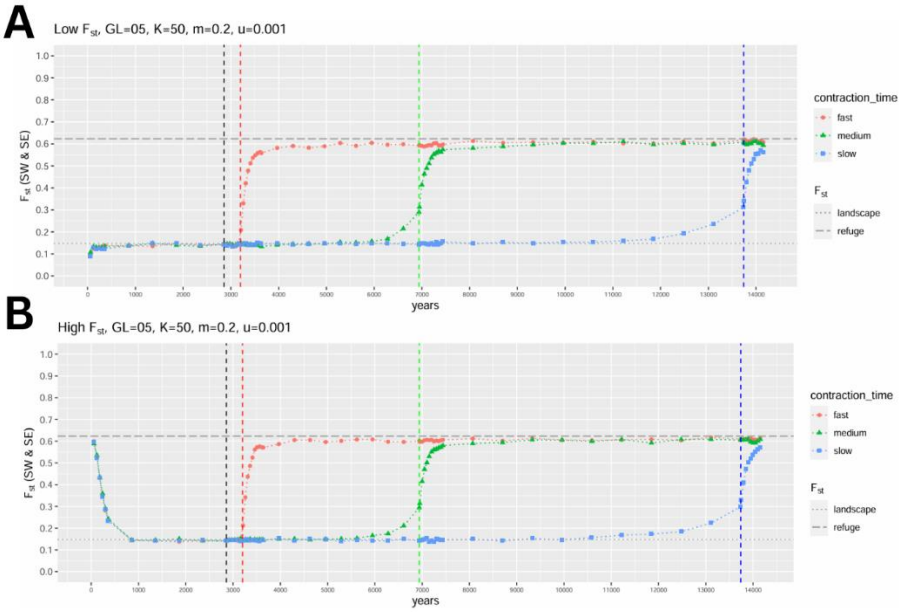

Figure A.4 - **Dynamics of F_{st} in the Southwest and Southeast (SW and SE) refugia populations, for GL=05, in both "High F_{st}" and "Low F_{st}" scenarios:** Same scenarios as depicted in Figure 6 but referring to GL=05, where we observe no significant impact from starting the expansion phase with different amounts of accumulated genetic diversity.

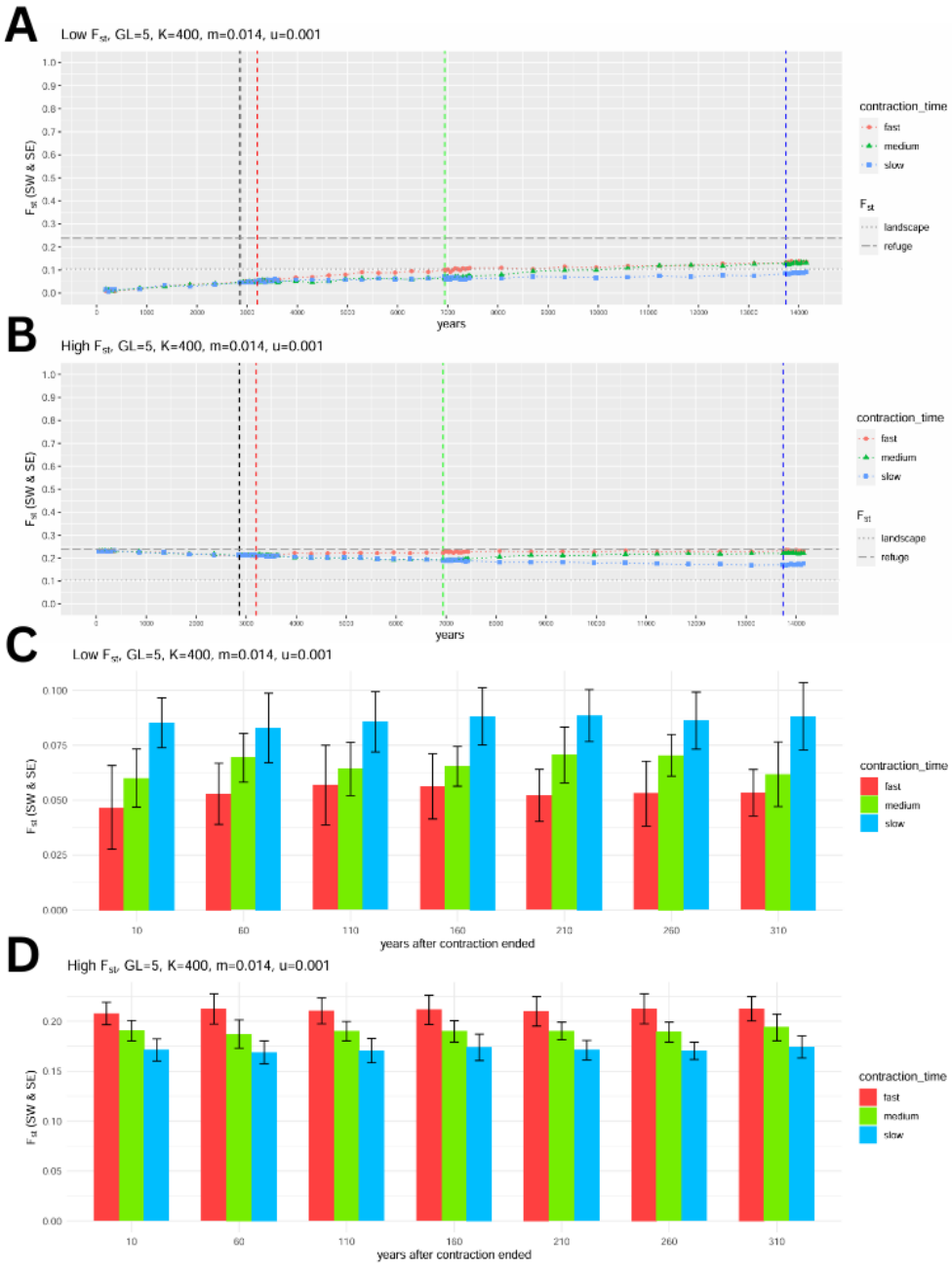

Figure A.5 - **F_{st} metrics of K=400 & m=0.014 for both SW and SE refugia:** Simulations refer to K=400 & m=0.014, with Panel A and B touching on the dynamics throughout the expansion-contraction scenarios, with different initial genetic differentiation values (F_{st}).

UNIVERSIDADE NOVA
DE LISBOA