UNIVERSIDADE D
COIMBRA

Ramona-Elena Irimia

INVASIVE SPECIES – ECOLOGICAL AND GENOMIC APPROACHES TOWARDS UNDERSTANDING LOCAL ADAPTATION AND EARLY STAGES OF ALLOPATRIC SPECIATION

Tese de Doutoramento em Biociências, especialização em Ecologia, orientada pelo Doutor Daniel Montesinos Torres, Doutor Adrian Christopher Brennan e Doutor João Carlos Mano Castro Loureiro, apresentada ao Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

October 2019

Invasive species - ecological and genomic approaches towards understanding local adaptation and early stages of allopatric speciation

Espécies invasoras - abordagens ecológicas e genômicas para compreender a adaptação local e os estágios iniciais da especiação alopátrica

Ramona-Elena Irimia

Department of Life Sciences of the Faculty of Sciences and Technology
Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia

Dissertation submitted to the Department of Life Sciences of the Faculty of Sciences and Technology of the University of Coimbra in partial fulfillment of the requirements for the degree Doctor of Philosophy in Biosciences, specialization Ecology, supervised by Dr. Daniel Montesinos Torres (Department of Life Sciences, University of Coimbra and Australian Tropical Herbarium, James Cook University), Dr. Adrian Christopher Brennan (Department of Biosciences, Durham University) and Prof. João Carlos Mano Castro Loureiro (Department of Life Sciences, University of Coimbra)

Dissertação apresentada ao Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra para prestação de provas de Doutoramento em Biociências, especialização em Ecologia, orientada pelo Doutor Daniel Montesinos Torres (Departamento de Ciências da Vida, Universidade de Coimbra e Herbário Tropical Australiano, James Cook University), Doutor Adrian Christopher Brennan (Departamento de Biociências, Universidade de Durham) e Prof. João Carlos Mano Castro Loureiro (Departamento de Ciências da Vida, Universidade de Coimbra)

October 2019

UNIVERSIDADE D
COIMBRA

Invasive species - ecological and genomic approaches towards understanding local adaptation and early stages of allopatric speciation

Espécies invasoras - abordagens ecológicas e genômicas para compreender a adaptação local e os estágios iniciais da especiação alopátrica

Ramona-Elena Irimia

This work was carried out mainly at the Centre for Functional Ecology in the Department of Life Sciences of the University of Coimbra and partially in the Department of Biosciences at Durham University (UK), and was funded by POPH/ESF and POCH/ESF funds, by the Portuguese Foundation for Science and Technology through a Doctoral Grant (SFRH/BD/110987/2015) and the exploratory project PTDC/BIA-PLA/0763/2014 by FCT/MEC through national funds and co-funding by FEDER, within the PT2020 Partnership Agreement, and by COMPETE 2020 within the project UID/BIA/04004/2013

Este trabalho foi realizado principalmente no Centro de Ecologia Funcional do Departamento de Ciências da Vida da Universidade de Coimbra e parcial no Departamento de Biociências da Universidade de Durham e foi financiado por fundos POPH/FSE e POCH/FSE, pela Fundação para a Ciência e Tecnologia através de uma bolsa de Doutoramento (SFRH/BD/110987/2015) e do projeto exploratório PTDC/BIA-PLA/0763/2014 pela FCT/MEC através de fundos nacionais e cofinanciamento FEDER, no âmbito do Acordo de Parceria PT2020, e pelo COMPETE 2020, no âmbito do projeto UID/BIA/04004/2013

CENTRE FOR FUNCTIONAL ECOLOGY
SCIENCE FOR PEOPLE & THE PLANET

Cofinanciado por:

REPÚBLICA PORTUGUESA | CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR
FCT Fundação para a Ciência e a Tecnologia
POCH
PORTUGAL 2020
UNIÃO EUROPEIA Fundo Social Europeu

1 2 9 0
UNIVERSIDADE D COIMBRA

“We must make no mistake, we are seeing one of the great historical convulsions in the world’s fauna and flora”.

Charles Elton – The ecology of invasions by animals and plants

Acknowledgments

I am deeply thankful to many wonderful people I have met along this journey and without whom this work would not have been possible.

I would like to acknowledge my thesis advisor Dr. Daniel Montesinos for his patience, time, trust and continuous support in directing me throughout this venture. Muchas gracias Daniel for being a source of inspiration with your positive thoughts and contagious energy and for giving me the opportunity to work on such an exciting project! I am very grateful to my co-advisor Dr. Adrian Brennan and his lab at Durham University, UK, for his guidance and input in regards to the genomic part of this thesis. I also thank my co-advisor at UC, Dr. João Loureiro for assistance with the flow-cytometry and helpful guidance with many aspects of my thesis. A big thank you to Prof. Ian Sanders and his lab. Thank you to all my co-authors for stimulating scientific discussions. I thank the PhD thesis committee members for their valuable input and comments to improve this work.

I acknowledge the Portuguese Fundação para a Ciência e a Tecnologia (FCT), Ministério da Ciência, Tecnologia e Ensino Superior for the doctoral scholarship SFRH/BD/110987/2015 and for funding the exploratory project PTDC/BIA-PLA/0763/2014 that allowed to carry out this work.

I would like to thank the Centre for Functional Ecology (CFE) for logistics and facilities and Durham Genomics Sequencing and Analysis Facility for DNA sequencing service. Obrigada Andreia Jorge, Daniela Tavares, Soraia Branco, Miguel Barreira, João Filipe as well as other colleagues in the lab for your help in the greenhouse and with various other parts of my work.

I thank Prof. Rus Hoelzel at Durham for inspiring chats on ddRAD troubleshooting and qPCR. Thank you Alireza Foroozani, Menno de Jong and Daniel Moore for sharing your experience on DNA library preparation and Linux. Many thanks also to Anurag Chaturvedi for very helpful insights on NGS data analysis.

Lastly, a big thank you to my dear friend Bu, for always being there to support, encourage and help me stay sane and motivated during these years! Thanks to my parents and my brother who have provided me with continuous love, encouragement and support. To those lost but never forgotten: thanks grandma Ana, for your unconditional love. I miss you deeply!

Table of Contents

List of Abbreviations

2n – sporophytic

2x – diploid

2C – holoploid genome size

AIC – Akaike information criterion

A_r – allelic richness

ATP – adenosine triphosphate

BIOSIS – Biosciences Information Service

BLAST – basic local alignment search tool

bp – base pair

CI – confidence interval

CH_2Cl_2 – dichloromethane

CV – coefficient of variation

CTAB – cetyl trimethylammonium bromide

ddRADSeq – double digest restriction-site associated DNA

DAPC – discriminant analysis of principal components

DNA – deoxyribonucleic acid

d.f. – degrees of freedom

EICA – evolution of increased competitive ability

EDTA Na_2 x 2H2O – disodium ethylenediaminetetraacetate dehydrate

EST-SSR – expressed sequence tags simple sequence repeats

FI – fertility index

F_{is} – inbreeding coefficient

F_{ST} – fixation index (a measure of population differentiation due to genetic structure)

GBIF – global biodiversity information facility

GC-MS – gas chromatography mass spectrometry

GzLM – generalized linear model

HiSeq – high throughput sequencing

HSD – honestly significant difference

H_o – observed heterozygosity

H_e – expected heterozygosity

i.e. – id est, in other words

ISI – International Scientific Indexing

ISO – International Organization for Standardization

mRNA – messenger ribonucleic acid

MEDLINE – Medical Literature Analysis and Retrieval System Online

$MgCl_2$ x $6H_2O$ – magnesium chloride hexahydrate

MSD – mass spectrometric detector

mW – milliwatt

NaCl – sodium chloride

N_e – effective population size

NCBI – National Center for Biotechnology Information

NRCS – Natural Resources Conservation Service

ng – nanogram

nm – nanometer

NPK ratio – volume of nitrogen, phosphorus and potassium

NWH – novel weapons hypothesis

OECD – The Organisation for Economic Co-operation and Development

PCA – principal component analysis

PCR – polymerase chain reaction

P_{ST} – phenotypic divergence index

PstI – type II restriction endonuclease isolated from *Providencia stuartii*

PVP – polyvinylpyrrolidone

ppm – parts per million

pg – picograms

PI – propidium iodide

Q_{ST} – standardized measure of the genetic differentiation of a quantitative trait

QTL – quantitative trait locus

qPCR – quantitative polymerase chain reaction

rRNA – ribosomal ribonucleic acid

RNA – ribonucleic acid

RNase – ribonuclease

RMSE – root mean square error

SCI – self compatibility index

SDH – shifting defense hypothesis

SciELO – The Scientific Electronic Library Online

SD – standard deviation

SE – standard error

S:O – seed:ovule ratio

SNPs – single-nucleotide polymorphisms

Tris.HCl – tris(hydroxymethyl)aminomethane

Tris.HCl-EDTA – ethylenediaminetetraacetic acid-Tris(hydroxymethyl)aminomethane

tRNA – transfer ribonucleic acid

TAIR – Arabidopsis Information Resource

USDA – United States Department of Agriculture

US EPA – United States Environmental Protection Agency

μm – micrometer

WGS – World Geodetic System

WPB – woody plant buffer

WoS – Web of Science

w/v – weight by volume

v/v – volume per volume

$\chi^2_{(1)}$ – chi square distribution

List of Figures

List of Tables

List of Appendices

Abstract

Invasive species are a threat to biodiversity and economy, and extremely challenging to eradicate once established. Much research in invasion ecology had focused on understanding the factors behind invaders' success in new environments relative to the native range, formulate invasion theories, predict species abundance and occurrence, and provide management solutions. Nowadays, the new advancements in genomic tools and sequencing technologies make possible to disentangle some of the underlying evolutionary processes of trait divergence and rapid adaptation characterizing species introduction and invasion in new habitats.

The aim of this thesis was to explore the ecological and genomic basis of rapid adaptation and evolution in the invasive annual forb *Centaurea solstitialis* (yellow star-thistle), across its worldwide distribution, including sites in the native range (Turkey and Spain) and introduced range (Argentina, Chile, USA and Australia). The thesis focuses on four main directions: i) assessing the role of polyploidy and genome size on species success; ii) testing the allelopathic potential of leaf leachates iii) screening for reproductive isolation in allopatric populations, and iv) evaluating the adaptive phenotypic and genomic potential of introduced populations.

In the introduction, I developed a brief systematic review by compiling information on the studies published in *Centaurea solstitialis* and indexed in the Web of Science during the past 70 years, to assess the current state of knowledge in this species. I identified a number of 365 relevant papers mostly with an ecological focus and having USA as the geographical region of the study.

In the first chapter, I used flow cytometry to test the hypothesis that variation in genome size and changes in ploidy levels promote *C. solstitialis* invasion in the introduced range. I found no shifts in cytotype and similar genome size across native and non-native ranges, excluding the contribution of these factors to species invasiveness.

Chapter II explored biogeographical variation in leaf allelochemical production and its effects on phytometer species between the two ranges, using three different leaf extract concentrations (0.25%, 0.5% and 0.75%) and germination bioassays tests. I found that *C. solstitialis* leaf leachates can have alellopathic potential, exhibiting substantial variation in chemical composition and inhibitory effects across regions. Results suggest that different

selection pressures can act on the biochemical profiles in different regions.

Chapter III utilized a novel approach in this study system by experimentally producing F_1 hybrids of within-region and inter-region crosses to test for reproductive isolation. Results revealed a global mosaic of reproductive incompatibilities and fertilities with asymmetrical responses to inter-continental gene flow. Most negative and strong fitness interactions occurred in the Americas suggesting local adaptation and reinforcement against foreign pollen. In contrast, native Spain showed a preference for non-native pollen resulting in boosts in fertility. Results from this study show that reproductive isolation can emerge relatively fast in allopatry.

The final chapter of my thesis explored the role of natural selection in species evolution in the introduced range by measuring neutral genetic differentiation (F_{ST}) at thousands of genome wide single nucleotide polymorphisms (SNPs) markers and comparing it with phenotypic differentiation (P_{ST}), in a common garden experiment. I also screened for SNPs under selection and performed gene annotation. Based on these data, introduced populations in California, USA had significantly higher P_{ST} for seed mass than F_{ST}, compared to the two native regions as well as non-native Chile, suggesting that increased seed size had evolved post-introduction in California. Moreover, phenotypic divergence in flowering time and spine length exceeded neutral expectation in the comparison between California and Australia. My research also shed light on the genes likely to be involved in invasiveness, revealing genes associated to regulatory processes and response to environmental stressors. This suggests local adaptation in introduced populations, and reveals the importance of traits related to reproduction and possibly of epigenetic factors in shaping *C. solstitialis* evolution.

Taken together, the results of my PhD thesis have demonstrated the crucial aspect of incorporating biogeography in studying biological invasions, in order to capture variability in factors and processes important for species success in different regions, and pioneers the use of experimental inter-regional crosses to illustrate how gene flow and local adaptation interact across large geographical scales.

Keywords: invasive species, polyploidization, allelopathy, reproductive isolation, speciation, population genetics, divergent selection

Resumo

As espécies invasoras são uma ameaça para a biodiversidade e a economia, sendo extremamente complicado erradicá-las uma vez estabelecidas. A maioria da investigação em ecologia da invasão focou-se em perceber os fatores inerentes ao sucesso das espécies invasoras em novos habitats em comparação com a área nativa, em formular teorias sobre a invasão, em prever a abundância e ocorrência de espécies, e em providenciar soluções de gestão. Hoje em dia, os avanços em ferramentas de genómica e em tecnologias de sequenciação tornam possível desvendar alguns dos processos evolutivos inerentes à divergência de características assim como perceber a adaptação rápida que caracteriza a introdução de espécies e a invasão de novos habitats.

Os objetivos desta tese foram explorar a base ecológica e genómica da adaptação rápida e evolução na espécie invasora anual *Centaurea solstitialis* (cardo-estrelado amarelo), ao longo da sua distribuição à escala global, incluindo populações das áreas nativa (Turquia e Espanha) e invadidas (Argentina, Chile, EUA e Austrália). A tese foca-se em quatro direções principais: i) avaliação do papel da poliploidia e tamanho de genoma no sucesso da espécie; ii) teste do potencial alelopático de lixiviados foliares; iii) avaliação do isolamento reprodutivo em populações alopátricas; e iv) avaliação da potencial adaptativo fenotípico e genómico das populações introduzidas.

Na introdução, desenvolvi uma breve revisão sistemática através da compilação de informação sobre os estudos publicados em *Centaurea solstitialis* e indexados na "Web of Science" durante os últimos 70 anos, para avaliar o estado atual do conhecimento nesta espécie. Identifiquei 365 artigos relevantes com um foco ecológico e tendo os EUA como área geográfica de estudo.

No primeiro capítulo, utilizei a citometria de fluxo para testar a hipótese que a variação no tamanho do genoma e alterações nos níveis de ploidia promovem a invasão de *C. solstitialis* na área onde foi introduzida. Não foram encontradas alterações no citotipo, enquanto que o tamanho do genoma foi semelhante ao longo das áreas nativas e não nativas, excluindo a possibilidade destes fatores contribuírem para a invasão da espécie.

No capítulo II explorei a variação biogeográfica na produção aleloquímica das folhas e os

seus efeitos nas espécies fitométricas entre as duas áreas, usando três concentrações diferente de extrato foliar (0,25%, 0,5% e 0,75%) e testes de bioensaios de germinação. Descobri que lixiviados foliares de *C. solstitialis* podem ter potencial alelopático, exibindo variação substancial na composição química e efeitos inibitórios em todas as regiões. Os resultados sugerem que diferentes pressões seletivas podem atuar nos perfis bioquímicos nas diferentes regiões.

No capítulo III utilizei uma nova abordagem neste sistema de estudo, produzindo experimentalmente híbridos F1 de cruzamentos dentro e entre regiões para testar o isolamento reprodutivo. Os resultados revelaram um mosaico global de incompatibilidades e fertilidades reprodutivas com respostas assimétricas ao fluxo génico intercontinental. A maioria das interações de fitness negativas ocorreram nas Américas, sugerindo adaptação local e reforço contra pólen estrangeiro. Por outro lado, a área nativa de Espanha mostrou preferência pelo pólen não nativo, resultando num aumento da fertilidade. Os resultados deste estudo mostram que o isolamento reprodutivo pode emergir relativamente rápido em alopatria.

O capítulo final da minha tese explorou o papel da seleção natural na evolução das espécies na área invadida medindo a diferenciação genética neutra (F_{ST}) em milhares de marcadores de polimorfismos de nucleotídeo único (SNPs) ao longo de todo o genoma e sua comparação com a diferenciação fenotípica (P_{ST}), numa experiência de estufa. Também examinei SNPs sob seleção e realizei anotação de genes. Com base nestes dados, as populações introduzidas na Califórnia, nos EUA, apresentaram P_{ST} significativamente maior para o peso das sementes do que o F_{ST}, em comparação com as duas regiões nativas e o Chile não nativo, sugerindo que o aumento do tamanho das sementes evoluiu após a introdução na Califórnia. Além disso, a divergência fenotípica no tempo de floração e comprimento dos espinhos excedeu a expectativa neutra na comparação entre a Califórnia e a Austrália. A minha investigação também lançou novos dados sobre os genes que estão provavelmente envolvidos na invasão, revelando genes associados a processos regulatórios e com a resposta a estressores ambientais. Isto sugere adaptação local em populações introduzidas e revela a importância de características relacionadas com reprodução e, possivelmente, fatores epigenéticos relacionados com a evolução de *C. solstitialis*.

Em suma, os resultados da minha tese de doutoramento demonstraram que é crucial a

incorporação da biogeografia no estudo de invasões biológicas, a fim de capturar a variabilidade em fatores e processos importantes para o sucesso de espécies em diferentes regiões, e é pioneira no uso de cruzamentos inter-regionais experimentais para ilustrar como o fluxo génico e a adaptação local interagem ao longo de grandes escalas geográficas.

Palavras chave: espécies invasoras, poliploidização, alelopatia, isolamento reprodutivo, especiação, genética populacional, seleção divergente

General Introduction

Invasive species and evolution

Invasive species can cause significant damage to the environment by changing the structure and composition of communities and altering ecosystem processes and services (Vilá et al. 2011). They can also act as strong filters affecting resource availability which can lead to biotic homogenization (i.e. community simplification) with unknown consequences for future ecosystem stability and function (Stotz et al. 2019). Worldwide global estimates of the costs associated to the impact of invasive species are sporadic and incomplete, but see Bradshaw et al. 2016 who estimated a global cost of $70 billion per year, for invasive insects only. Human dimension of biological invasions has become an important component in the way invasions are perceived and managed especially nowadays when trade and globalisation have made alien species to cross biogeographic barriers that could not have been possible by natural colonization (Wilson et al. 2016). Understanding the processes and mechanisms of invasions for different species is crucial in taking management and policy actions.

Biological invasions stimulate a multitude of research questions related to evolutionary changes of the invader and of the recipient community, and should consider biogeography as a variable in order to understand patterns of species distribution and abundance and the processes behind it (Hierro et al. 2005). Introduced species are present in areas to which they might be poorly adapted and encounter diverse selection pressures to which they must respond quickly in order to establish and increase in abundance (Hänfling and Kollmann 2002). The success of the introduced species in recipient communities depends on their traits interaction with local filters, relative to natives and can translate into different range of outcomes such as failure to establish, naturalization or invasion (Pearson et al. 2018). Studying these factors and interactions can provide useful insights into the drivers of increased performance of invasive populations compared to their native conspecifics, the rate of adaptation and phenotypic change as well as genetic changes underlying it. Still, determinants of plant establishment and invasions remain a key question in ecology as they are likely dependent on the environment and can change over time (Kempel et al. 2013).

For the past 200 years, there has been an increase in species introductions globally, for all taxonomic groups, with no signs of saturation (Seebens et al. 2017). Asia and Europe have

been the major donors of alien species to all other continents (van Kleunen et al. 2015) and species with broad native habitat ranges were able to naturalize and invade new areas more easily (Kalusová et al. 2017). A global dataset of natural area plant invaders showed that most of them belong to the Poaceae, Fabaceae and Hydrocharitaceae families. The Asteraceae were not significantly over-represented, although a large number of Asteraceae were found among agricultural weeds and natural area invaders, probably as a function of the large number of species in this family (Daehler 1998). The genus *Centaurea* in the Asteraceae has been identified as having a propensity for invasiveness (Schmidt and Drake, 2011), and has attracted much research interest in invasion science with different species used as study systems to understand coevolution patterns in plant communities, climate niche shifts during invasion, belowground plant-plant allelopathic interactions (Kueffer et al. 2013) or test for incipient speciation in allopatry (Montesinos et al. 2012). Species in the *Centaurea* genus occur mainly in the Mediterranean and Irano-Turanian region and are commonly known as thistles, being associated with livestock farming and crop cultivation (Garcia-Jacas et al. 2006). Many of the *Centaurea* species were introduced throughout the world, and some became highly invasive, particularly in the North America (Skinner et al. 2000; Lejeune et al. 2002), South America (Busso et al. 2013; Montenegro et al. 1991) but also in Australia and South Africa where several species are listed as prohibited, noxious or as sleeper weeds (NEMBA 2016; Grice 2005).

This PhD thesis focuses on *Centaurea solstitialis*, an invasive aster forb and aims to gain a more comprehensive understanding of the ecological and genetic factors contributing to species invasive success. It combines data from a literature survey, a worldwide sampling of *C. solstitialis* natural populations in most representative native and introduced areas, pollination experiments under common garden conditions, phenotypic data and genomic tools. The thesis is structured in four chapters corresponding to four original papers (referred in the text by their Roman numerals I-IV), a general introduction and a general discussion section. Chapter I explores the putative role of genome size and changes in ploidy levels in the species success; Chapter II characterizes the leaf surface chemistry and its allelopathic potential; Chapter III tests for reproductive isolation between native and neo-allopatric populations of *C. solstitialis* and Chapter IV evaluates plant adaptive genetic and phenotypic potential.

Study system

Centaurea solstitialis (yellow star-thistle) is a diploid annual forb (Irimia et al. 2017) and an obligate out-crosser, with dimorphic seeds, pollinated predominantly by the European honey bee, *Apis mellifera* (Maddox et al. 1996). It forms dense monoculture infestations which affect the value of grasslands, displace native plant species by altering soil water availability and is also poisonous to horses (Pitcairn et al. 2006). The plant is a rich source of high quality pollen and nectar, and therefore is well valued by the bee industry (Pellet 1976), which has led some authors to hypothesize that honey bee could promote *C. solstitialis* invasion in the western USA (Barthell et al 2001). It is suggested that *C. solstitialis* originated in the Eastern Mediterranean and the Caucasus and is naturalized in the southern and central Europe (Maddox et al. 1985). The species was introduced to South America (central Argentina and Chile), California, South Africa and Australia around 1800, most likely as a contaminant of alfalfa seeds and due to its short presence there, these regions are regarded as non-native (Hierro et al. 2009). *Centaurea solstitialis* has become invasive in central Argentina and California, two regions with contrasting climates and plant functional communities. Central Argentina has a continental climate and its vegetation is dominated by perennial bunchgrasses whereas California has a more Mediterranean-like climate characterized by summer drought and winter rain and Eurasian annual grasslands (Hierro et al. 2006). *C. solstitialis* has also been catalogued as a serious weed of agricultural crops and native shrublands (mattoral) in central Chile (Gómez-González et al. 2009). In Australia, the species has a sparse and localized distribution (particularly in the New South Wales) and is not regarded as highly invasive (Hay et al. 2006).

Recent genomic data indicate that *C. solstitialis* was introduced to Chile and California from a single seed source of Spanish origin (Barker et al. 2017). In contrast, several introductions from different sources (such as Asia, Eastern Europe and Western Europe) were detected in the Pacific Northwest (i.e. Oregon, Idaho and Washington). This study also found clear population genetic structure in the native and introduced ranges but similar levels of genetic diversity in terms of allelic richness and heterozygosity between native and non-native ranges suggesting that *C. solstitialis* has not suffered from severe genetic bottlenecks post introduction in the

Americas (Barker et al. 2017). Hybridization with other *Centaurea* species has been documented for *C*. *solstitialis* in Turkey (Wagenitz 1955) and Oregon (Roché and Susanna, 2011), but Barker et al. 2017 did not find compelling evidence that hybridization has contributed to the species invasive success.

Many invasion hypothesis were tested in *C*. *solstitialis*, some with inconsistent results. Most research focused on traits associated with invasiveness or phenotypic shifts in invasive plants and showed that non-native genotypes frequently had superior fitness values compared to natives for germination, growth rates, total biomass, seed size, abundance and fecundity (Montesinos et al. 2018, Graebner et al. 2012; Hierro et al. 2006; 2013; Garcia et al. 2013). Other studies found evidence of increased competitive ability (Montesinos et al. 2019; Montesinos and Callaway 2017), escape from natural enemies (Hierro et al. 2006; Andonian et al. 2012; Waller et al. 2016), and exploitation of an empty niche, where the species is able to take advantage of deep water layers (Young et al. 2011; Dlugosch et al. 2015). Trade-offs between seed dormancy and dispersal have been observed in common garden experiments for Argentinean genotypes of *C*. *solstitialis* experiencing variation in winter precipitation and can be viewed as a ´´bet-hedging´´ strategy, contributing to species survival in heterogeneous environments (Hierro et al. 2009). Another study of the same species demonstrated no competition-colonization trade-offs in seed morph, which is likely to play a role in *C*. *solstitialis* abundance and distribution (Miguel et al. 2017). In addition to the dormancy dispersal trade-offs detected in Argentinean populations, Kaczowka et al. 2017, also found support for growth defense trade-offs in Californian genotypes of *C*. *solstitialis* which exhibited increased plant size and lower resistance to pathogen infection compared to native (Kaczowka et al. 2017).

On the other hand, Hierro et al. 2017, showed that factors limiting *C*. *solstitialis* performance can be encountered both abroad and at home, questioning the assumption that invasive species benefit from better conditions in the introduced areas. They also highlighted that disturbance and plant community resilience may play an important role in *C*. *solstitialis* abundance (Hierro et al. 2017). Moreover, *C. solstitialis* appears to have retained similar ecological requirements and climatic niche across different geographical ranges, although invading populations in California seem to occupy the warmest and seasonally dry habitats

available, which in the native area are only found marginally (Dlugosch et al. 2015).

It is clear that factors determining naturalization and invasion success differ between different *C. solstitialis* geographical areas. The species has been the focus of hundreds of studies, but lacks coherent systematic data to encompass global information across its native and non-native ranges in a biogeographical context. Therefore, as an addition and alternative to a classical introduction, we performed a systematic review on the available *C. solstitialis* scientific literature to categorize the studies published, identify major information gaps, link it to our own work on this study system and outline future research directions in the discussion section.

Centaurea solstitialis – a literature database

We performed a literature search in ISI Web of Science All Databases (i.e. Web of Science Core Collection, BIOSIS Citation Index, Current Contents Connect, Data Citation Index, Derwent Innovations Index, KCI-Korean Journal Database, MEDLINE, Russian Science Citation Index, SciELO Citation Index and Zoological Record) using the keywords "*Centaurea solstitialis*" OR "*yellow star-thistle*" for the period 1864 – 8 August 2019 and retrieved a total of 698 studies. We reviewed abstracts of each article and excluded studies whose species of interest was not *C. solstitialis*, the conference abstracts, and the papers published in languages other than English. The final database comprised 365 relevant studies. For each study, we recorded the following variables: i) the country of origin of the plant genotype used in the experiment; ii) the geographical location where the study was conducted; iii) duration of the experiment; iv) *Centaurea solstitialis* life stage used; v) type of research (whether theoretical or experimental); vi) type of habitat where the experiment was conducted; vii) main hypothesis; vii) whether it measured any fitness components and vii) whether it provided information about causal factors involved in the invasion success. Our main goal was to assess the current state of knowledge in *C. solstitialis* by identifying what has been published in the literature on this topic, and compile a list of ecological and evolutionary factors involved in its invasion success. We also build a database in excel comprinsing all this information, for more complex systematic research and future meta-analysis. We designed the systematic review following guidelines by Lowry et al. 2013. Until the date of this thesis submission to the University, we managed to retrieve complete information on 262

(out of 365) papers published during the past 23.5 years (1995 – 8 August 2019) (see Appendix 1).

Numbers of studies published

The first study identified in the ISI Web of Science (WoS) All Databases was published in 1948. Prior to the 1990 publication numbers were low (i.e. 43 papers) (Fig 1). Studies nearly doubled in the 90´ (79 papers, average of 7.9 papers per year) and continued to increase throughout the 2000s (129 papers, average of 12.9 papers per year) and in the following decade 2010 – 8 August 2019 (114 papers) (Fig 1).

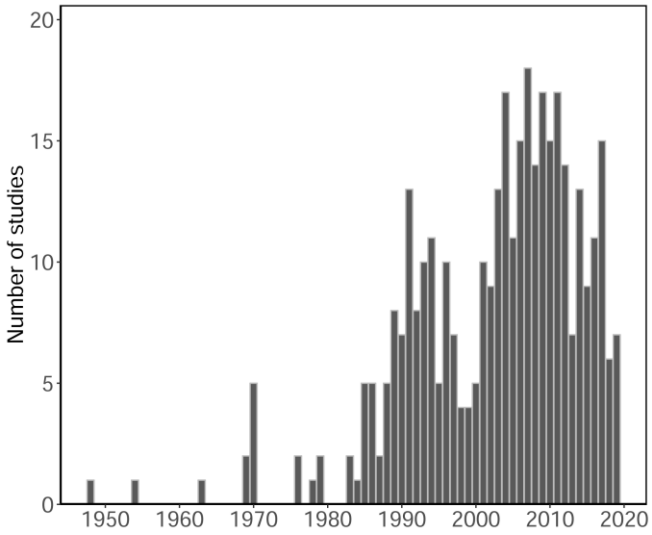


Figure 1. Number of studies (N = 365) returned from the ISI Web of Science (WoS) All Databases, all years (1864-2019), search topic "*Centaurea solstitialis*" OR "*yellow star-thistle*", after filtering for irrelevant articles. For 2019, only papers indexed in WoS prior to August 8 were included in the database. Height of the bars refers to the number of studies

Study focus

A large number of studies were concerned with methods for controlling invasive species such as:

- identification of biological control agents;
- interactions among biocontrol agents;

- chemical control;
- mechanical control;
- combined management practices;
- herbicide resistance and natural herbicides products;
- *Centaurea solstitialis* demographic models;
- weed detection methods and weed survey;
- weed management in general and economic impact.

Further categorization included studies trying to determine the causes of invasion such as:

- enemy release hypothesis;
- novel weapon hypothesis (NWH);
- evolution of increased competitive ability (EICA);
- evolution in general;
- invasive traits and phenotypic plasticity;
- plant soil feedback and plant-microbe interactions or plant-plant interactions;
- invasion facilitation;
- disturbance;
- hybridization;
- ploidy levels;
- others.

Another category addressed the phytochemical and medicinal applications of *C. solstitialis* i.e.

- detection and identification of phytochemical compounds;
- toxicological studies in horses;
- biological effects in ulcerogenic animal models;
- scavenging potential;
- anticancer activity;
- antibacterial effects.

The remaining papers included studies on the biology, phenology and genetics of *C. solstitialis*, community resistance to invasion, restoration ecology and impact on wildlife.

Study type

Collectively, experimental studies conducted in controlled or semi-controlled environments (greenhouse, growth chamber, lab or outdoor microcosms) were the most common type of research, followed by field manipulative and field observational studies. Reciprocal transplant experiments were highly underrepresented (only 1 study). A number of 43 studies out of 262 were conducted in more than one environment (16%). Very few studies involved theoretical modelling or reviews (0.06% respectively 0.02%) (Fig 2).

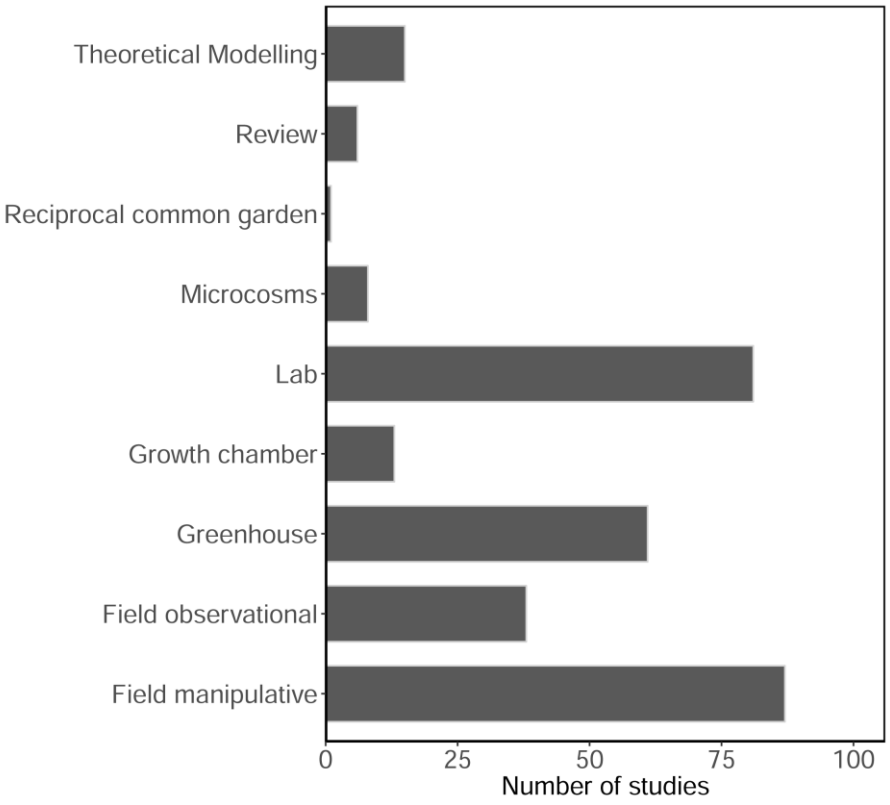


Figure 2. The type of research method used in the studies included in the database

Plant genotype origin and geographic location of study

Most studies were focused on a single genotype of *C. solstitialis* which was predominately of North American origin (mainly from California, some from Idaho and Washington and a few from Oregon and Nevada) followed by genotypes from Turkey, Argentina and France. By comparison, a number of 60 studies out of 262 included more than one *C. solstitialis* genotype

(22%), spanning species distribution in the native range (i.e. Armenia, Bulgaria, Croatia, Crete, Greece, Hungary, Italy, Republic of Georgia, Romania, Russia, Sardinia, Sicily, Spain, Uzbekistan, Ukraine and Turkey) as well as the introduced range (i.e. Argentina, Chile, USA and Australia). Globally, studies were clustered in North America followed by native Turkey, with very few studies elsewhere (Fig 3). Only 12 studies out of 262 (0.05%) were conducted in more than one geographical location (commonly 2 locations, rarely 4 or 5), including sites in the native species range in Eurasia and introduced range (represented by Argentina, Chile, USA and Canada) (Fig 3).

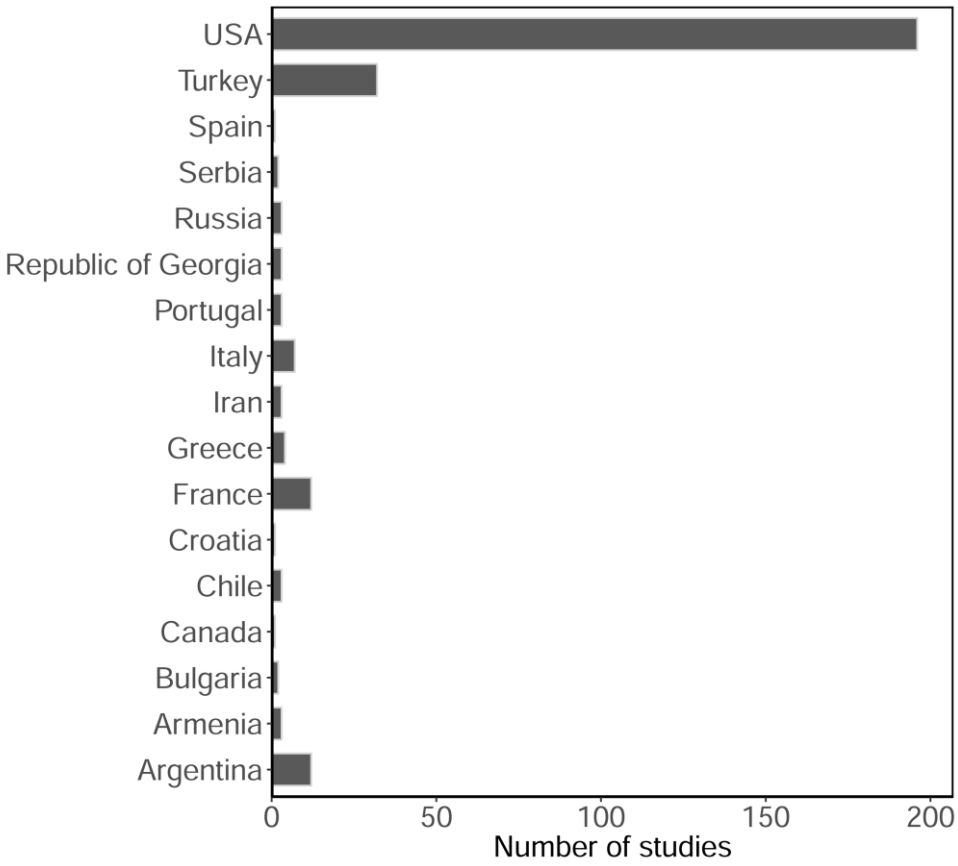


Figure 3. Geographical locations of the studies included in the database

Duration of the study and plant life stage tested

Information about the experiment duration could be retrieved for 174 (out of 262) studies only. In general, studies conducted in the field took longer to complete with most of them spanning across a period of 2 to 4 years, rarely 7 to 11 years in the case of studies that aimed to

assess the impact of restoration programmes or the success of different weed control methods. In contrast, studies conducted in controlled or semi-controlled environments (i.e lab, greenhouse, microcosms) lasted for a few months, a growing season and seldom for 1 or 2 years. The most studied *C. solstitialis* life stages were bloom, rosette and seeds, with a very limited numbers of studies investigating all the plant life stages.

Habitat

Field experiments were conducted in several types of *C. solstitialis* invaded habitats, predominately in grasslands (dominated by annual or exotic species, serpentine and non-serpentine grasslands, canyon grasslands), but also in man-made habitats (urban, roadsides, field margins, highly disturbed sites, abandoned sites and fallow lands, farmland, pasture, orchards, agricultural and botanical gardens) or in naturals sites such as prairies, savannah, rangelands, woodlands, foothills, dry interior valleys and watersheds. Additionally, some studies were conducted in coastal wet habitats, mountain / intermountain areas, island habitats or semi-desert and semi-arid open forests. A few studies were also conducted in stressful environments characterized by different light, shading and temperature regimes, drought or elevated CO_2.

Hypothesis

The largest number of studies were concerned with selecting natural enemies (mostly insects) in their native range and releasing them in the field to control *C. solstitialis* infestation in the western USA, and included host specificity testing, risk assessment and effectiveness (Fig. 4). Many studies also examined the effects of various management practices (herbicide application, mowing, fire, grazing) in relation to reducing the emergence, plant cover and *C. solstitialis* biomass. Another large category of studies tested diferent invasion hypothesis. The most well represented invasion hypothesis were Evolution of Increased Competitive Ability (EICA); Evolution in general; Enemy Release, Ecosystem Engineers and Disturbance. Other studies aimed to test the pharmacological and biological properties of different *C. solstitialis* extracts and individual compounds or gain a better understanding of the species reproductive biology and phenology. Lastly, some hypothesis were difficult to categorize and were placed under the

´´Others´´ section (Fig 4).

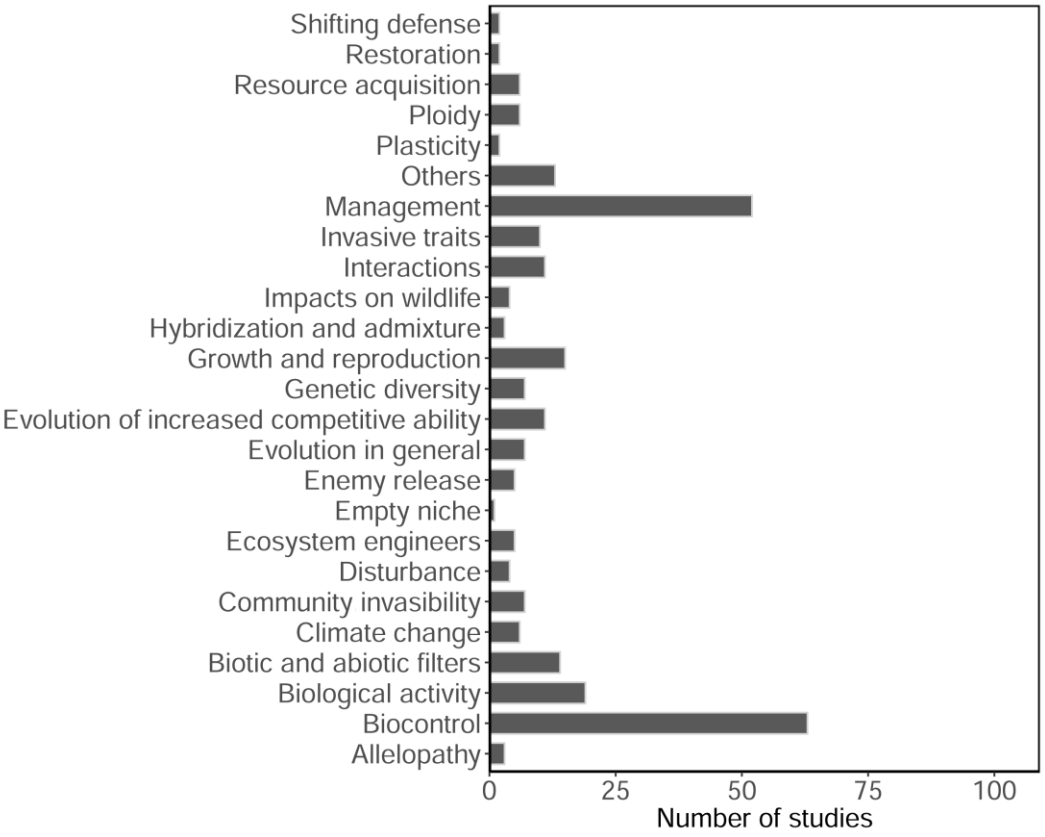


Figure 4. Hypothesis evaluated in the *C. solstitials* studies included in the database

Fitness components measured

A number of 193 studies out of 262 (73%) quantified plant fitness. Most studies incorporated multiple measurements of fitness and biomass was the most frequent estimate. Other trait categories were related to plant physiology (leaf chlorophyll and nitrogen content, water potential, leaf epinasty, chlorosis, photosynthesis, transpiration and stomatal conductance, net assimilation rates, necrosis, Ca tissue content and seeds nitrogen content), leaf area allocation (leaf area ratio, leaf area index), shoot allocation (leaf stem ratio, root shoot ratio), growth rate (radicle elongation, relative growth rates, root growth), size (plant height, height of the tallest leaf tip, spine length, style filament length, heights of bolting shoots, capitulum diameter, rosette diameter, leaf length and width, stem diameter and length,

involucre diameter, apical meristem height, bud diameter, seed size, root size), fecundity (capitula, ovule and seed number), pathogen and herbivory resistance, and very few studies measured offspring recruitment (seed and seedling).

Ecological and evolutionary explanations for *Centaurea solstitialis* invasive success

A number of 80 studies out of 262 (30%) were concerned with identifying causal factors behind *C. solstitialis* invasive success. Rapid adaptation, evolution and invasive traits in general were the most common explanatory variables for species establishment and survival in the introduced ranges (Fig 5). Two other factors (disturbance and empty niche) were also assumed to play an important role in the ecological success of the invader.

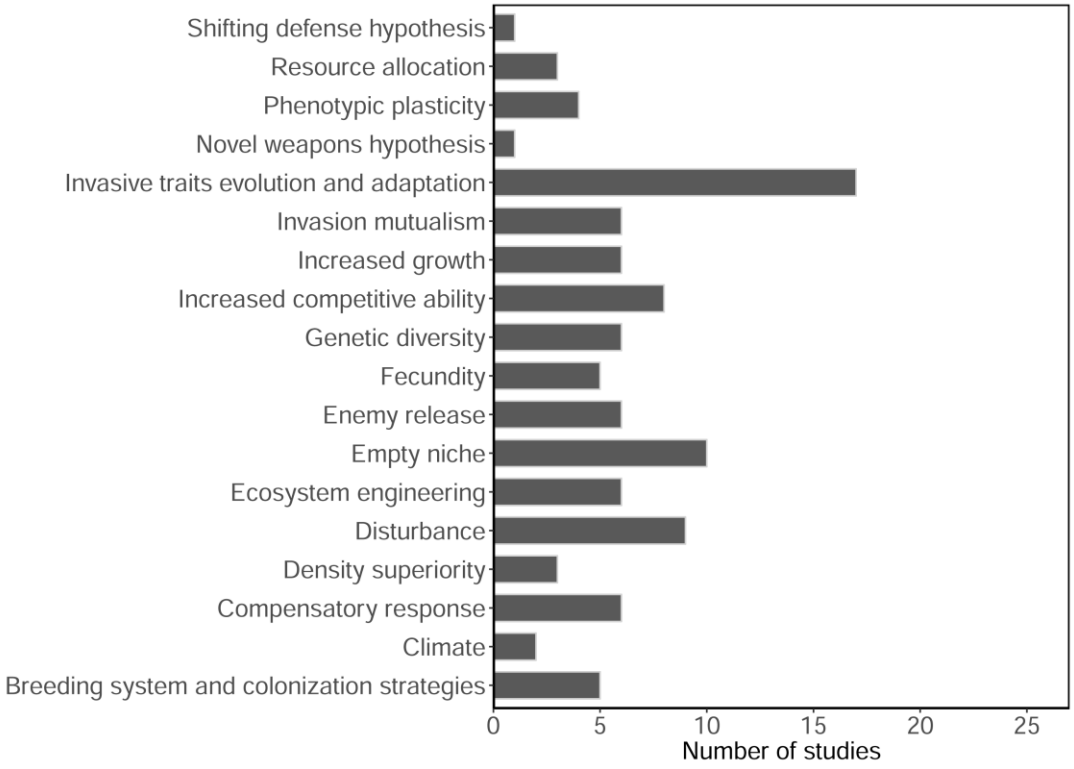


Figure 5. Major ecological and evolutionary explanations in *C. solstitialis* invasion

Other factors included evolution of increased competitive ability (EICA); compensatory responses to herbivory and mechanical control, mutualism and invasion facilitation by honey bee, release from enemies (i.e. soil microbes and pathogens), ecosystem engineering (i.e.

modifying soil aggregate stability, soil engineering to obtain sufficient nutrients for optimal growth and modifying soil microbiota in the detriment of competing species) and increased growth and genetic diversity (Fig 5).

Objectives and structure of the PhD thesis

The vast literature published in *C. solstitialis* based on the review we performed above suggests the emergence of this species as an optimum study system in invasion science. Owing to its successful establishment and colonization of multiple allopatric geographical locations, it offers an opportunity to assess the importance of cytological, biochemical and genomic factors in explaining its invasion success. Perhaps the most obvious pattern emerging from the review of the existent literature on *C. solstitialis* was the scarcity of biogeographical studies. In general, papers were focused mainly on a particular geographic area and a single plant genotype, e.g. western USA, a region where *C. solstitialis* is highly invasive, and showed poor representation in other introduced and native species areas. As advocated by Hierro et al. (2005), a biogeographic approach comparing native and introduced populations is key to understand plant invasions. Our sampling design for all the experiments presented in this thesis included several populations in the species' native area (Turkey and Spain) as well as important historical introductions in South America (Argentina and Chile), North America (California) and Australia, thus, ensuring a good representation at a broad biogeographical scale.

Some of the hypothesis tested in *C. solstitialis* showed inconclusive support and where not able to distinguish whther the importance of factors tested varied across native and non-native ranges. For example, there were conflicting reports of chromosome numbers and ploidy levels in the literature (Kuzmanov et al. 1990; Widmer et al. 2007; Inceer et. al. 2007) and only one estimate of genome size for one *C. solstitialis* population in the introduced range. **Chapter I** of my PhD thesis addresses this issue, by performing a large flow-cytometry screen on 52 *C. solstitialis* natural populations in the native and introduced range, to calculate genome size and unequivocally assess ploidy levels.

Allelopathy has been assumed to play a role in the invasive success of *C. solstitialis* in western USA but only two previous studies explored this possibility and focused exclusively on

root exudates with no evidence of any toxic compounds released (Carpenter 2007; Qin et al. 2007). However, all aerial parts of *C. solstitialis* contain potent sesquiterpene lactones, with well -known biological inhibitory activity (Alvarez 2008; Sotes *et al*. 2015), that could leak into the soil and act as phytotoxic agents on neighbor plants germination. In **Chapter II**, I tested for biogeographical differences in leaf surface chemistry and allelochemical production by high performance liquid chromatography coupled with mass spectrometry (HPLC-MS) and assessed the effects of leaf-extracts on a phytometer species by performing phytotoxicity tests and germination bioassays.

Populations of *C. solstitialis* occur allopatrically, and approximate introduction dates are known for most of the introduced regions, allowing to explore the emergence of divergent characters within known timeframes. A previous study detected an incipient degree of reproductive isolation between one native and one non-native *C. solstitialis* region, suggesting a role for reinforcement in maintaining local adaptation (Montesinos et al. 2012). In **Chapter III**, I tested the effects of admixture among geographically isolated ranges on reproductive outputs in *C. solstitialis* by performing experimental crosses and calculating region pair-wise fertility indices based on seed:ovule ratio.

The literature review also revealed a poor representation of evolutionary genetic studies (0.04%), although the past 7 years showed a positive trend for this type of research. Molecular tools such as allozymes, isozymes, microsatellites and recently SNPs (single nucleotide polymorphisms) markers were used to compare genetic diversity of native and introduced populations, elucidate population genetic structure and reconstruct invasion routes, assess effective population size in range expanding populations in the western USA and test for evolution and adaptation in native and non-native populations. Only one study explicitly tested for the role of selection in the invasion success of *C. solstitialis* in California and Argentina (Eriksen et al. 2012). Expanding on this information, **Chapter IV**, focuses on genetic signals of divergent selection between native and introduced *C. solstitialis* populations by performing a comparison between phenotypic trait differentiation (P_{ST}) and neutral genetic differentiation (F_{ST}) in a common garden experiment, using reduced representation DNA sequencing (ddRADSeq). It also investigates the genetic bases of adaptation by identifying and annotating outlier SNPs and

generates genomic data for *C*. *solstitialis* populations in Australia, which were never included in molecular studies before.

Chapter I - Extensive analysis of native and non-native *Centaurea solstitialis* L. populations across the world shows no traces of polyploidization

Chapter section published as an original research article in *PeerJ*:

Irimia RE, Montesinos D, Eren Ö, Lortie CJ, French K, Cavieres LA, Sotes GJ, Hierro JL (2017) Extensive analysis of native and non-native *Centaurea solstitialis* L. populations across the world shows no traces of polyploidization. PeerJ 5:e3531. 10.7717/peerj.3531.

ABSTRACT

Centaurea solstitialis L. (yellow starthistle, Asteraceae) is a Eurasian native plant introduced as an exotic into North and South America, and Australia, where it is regarded as a noxious invasive. Changes in ploidy level have been found to be responsible for numerous plant biological invasions, as they are involved in trait shifts critical to invasive success, like increased growth rate and biomass, longer life-span, or polycarpy. *C. solstitialis* had been reported to be diploid (2*n* = 2*x* = 16 chromosomes), however, actual data are scarce and sometimes contradictory. We determined for the first time the absolute nuclear DNA content by flow cytometry and estimated ploidy level in 52 natural populations of *C. solstitialis* across its native and non-native ranges, around the world. All the *C. solstitialis* populations screened were found to be homogeneously diploid (average 2C value of 1.72 pg, SD = ± 0.06 pg), with no significant variation in DNA content between invasive and non-invasive genotypes. We did not find any meaningful difference among the extensive number of native and non-native *C. solstitialis* populations sampled around the globe, indicating that the species invasive success is not due to changes in genome size or ploidy level.

Keywords: flow cytometry; genome size; hybridization; invasiveness; ploidy level; yellow starthistle

INTRODUCTION

Changes in ploidy level have been reported to be important for the invasive success of some plants species (Te Beest et al. 2011), by altering morphological, physiological and ecological parameters which can confer hybrid vigor, stress resistance, competitive advantages, or increased phenotypic plasticity, like in the case of the North American tetraploids of *Centaurea stoebe* L. (Hahn et al. 2012). Additionally, there are a series of associated "genome size constrained traits", related mostly to reproduction and dispersal, which dictate the ecological niche a species can access (Te Beest et al. 2011). In contrast, several studies support the hypothesis that a smaller genome can contribute to some species invasive potential by boosting early plant growth and enhancing competitive ability (Bennett et al. 1998; Grotkopp et al. 2004; Beaulieu et al. 2007; Lavergne et al. 2010; Suda et al. 2015). For instance, *Phalaris arundinacea* L. (reed canary grass, Poaceae) in the USA underwent a quick and significant reduction in genome size compared to the native European genotype, which was correlated with some advantageous phenotypic effects and enhanced aggressiveness (Lavergne et al. 2010). A list comparing the ploidy level of 128 worst invasive plant species worldwide, was recently made available by Te Beest et al. (2011), indicating that a quarter of them possess at least two different ploidy levels. An interesting example is *C. stoebe* (spotted knapweed) which occurs both as a diploid and tetraploid, with only the latter cytotype becoming invasive in the Western parts of the USA (Mráz et al. 2011). However, for many invasive species, ploidy levels and genome size are unknown or have not been thoroughly investigated.

Centaurea L. is one of the most species rich genera in the Asteraceae (Bremer, 1994). Numerous *Centaurea* species have been introduced into new non-native regions, where many of them have become invasive. For instance, the US Federal Noxious Weeds list (USDA, NRCS, The PLANTS Database, 2017), includes no fewer than 13 taxa, but ploidy level for many of these is unknown or uncertain. In particular, *C. solstitialis* is a Eurasian native annual herb which was introduced into the Americas and Australia during the last two centuries (Barker et al. 2017) and became an impactful invader in the former case. In the invaded ranges, *C. solstitialis* forms dense stands that displace native plants species and reduce considerably livestock grazing capacity and forage value (Eagle et al. 2007).

It alters ecosystem functions by depleting soil water and nutrients through an extensive root system (DiTomaso, 2000), and can cause a neurological disorder in horses similar to human Parkinson (Chang et al. 2011). The species causes significant economic damage in the western USA and has been the subject of intensive research, and significant differentiation between native and non-native ranges have been reported for plant size (Eriksen et al. 2012; Graebner et al. 2012; García et al. 2013; Dlugosch et al. 2015), growth rates (Graebner et al. 2012), germination (Hierro et al. 2009), competitive ability (Montesinos and Callaway, 2017), and reproduction (Montesinos et al. 2012), among others. Such changes suggest diverging local adaptation occurring among native and non-native ranges, and hypothetical changes in genome size and ploidy level could be potentially responsible for at least some of the observed trait-shifts.

Until now, only three genome size estimates were available in the literature for *C. solstitialis*: two from the native range (Bulgaria: 1.74 pg/2C, one accession, in Bancheva and Greilhuber, 2006; and Croatia: 1.95 pg/2C, five accessions, in Carev et al. 2017) and another from an invasive population in western USA: 1.66 pg/2C, thirty accessions (Miskella, 2014). Based on these few studies, *C. solstitialis* had been reported to be diploid (Dlugosch et al. 2013; Rice et al. 2015) with $2n = 2x = 16$ chromosomes. However, records of $2n = 2x = 18$ chromosomes were published more than 30 years ago from the native range of Bulgaria (Jasiewicz and Mizianty, 1975; Kuzmanov et al. 1990) and recently from one accession from Sicily and the other one from Sardinia (Widmer et al. 2007). Furthermore, Inceer et al. 2007, reported tetraploids in seeds (single accession) sampled in northern Turkey, but none of those observations, made in only a handful of individuals, have been confirmed since then. Consequently, it was still unclear whether ploidy could have played a role in at least some of the *C. solstitialis* invaded ranges. To fill this knowledge gap for such an important species, we aimed to thoroughly sample and assess *C. solstitialis* ploidy level and genome size in a representative number of populations from around the world, including native Turkey, the ancestral origin of the species; native Spain, the main source of American populations; and all the known non-native regions represented by Argentina, Chile, USA and Australia.

MATERIALS AND METHODS

Seed collection

A total of 477 accessions from 52 natural populations (Appendix 2) of *C. solstitialis* were investigated in this study, for genome size and ploidy level assessment. Within the native area, we sampled ten populations from Turkey, near the Caucasus region, where high genetic diversity has been detected, and is regarded as the site of origin of the species (Wagenitz, 1955; Gerlach, 1997a; Uygur et al. 2004; Dlugosch et al. 2013; Eriksen et al. 2014), and ten populations from Spain, considered as the primary source of seeds to have colonized Chile and Argentina (Hijano and Basigalup, 1995; Eriksen et al. 2012; Eriksen et al. 2014; Dlugosch et al. 2013; Barker et al. 2017) in the nineteenth century (Gerlach, 1997b). For the non-native regions, we included ten populations from Argentina and California, eight from Australia and four from Chile. Seeds were extracted from mature flower heads collected in the wild from ten individuals per population between 2009 and 2014. Ten seeds from each individual were germinated in plant growing trays, under common greenhouse conditions, in early spring 2016 at the Botanical Garden of the University of Coimbra, Portugal.

Flow cytometry

Young and intact leaves of 4–6 weeks-old plants were sampled and screened by flow cytometry. Since analyses were based on leaves of small plants, which were destroyed by leaf sampling, no voucher specimens could be collected. Nuclei were isolated following the chopping method of Galbraith et al. (1983). Briefly, about 1 cm^2 of leaf tissue was co-chopped with a razor blade together with the same amount of reference standard (*Raphanus sativus* L. 'Saxa', 2C = 1.11 pg, Doležel et al. 1992) in 1 mL of woody plant buffer (WPB): 0.2 M Tris×HCl, 4 mM $MgCl_2 \times 6H_2O$, 2 mM EDTA $Na_2 \times 2H_2O$, 86 mM NaCl, 10 mM sodium metabisulfite, 1% polyvinylpyrrolidone (PVP-10) (w/v) and 1% Triton X-100 (v/v), with pH of the buffer adjusted to 7.5 (Loureiro et al. 2007). The resulting homogenate was filtered through a 50 µm nylon filter into a sample tube to remove large debris. Nuclei were stained with 50 mg/mL propidium iodide (PI; Fluka, Buchs, Switzerland), and 50 mg/ml of RNAse (Fluka, Buchs, Switzerland) was added to prevent the staining of double stranded RNA.

Samples were kept at room temperature and analyzed immediately on a Partec CyFlow Space flow cytometer (Partec GmbH, Görlitz, Germany) equipped with a 532 nm green solid-state laser, operating at 30 mW.

Data collection and analysis

Results were acquired using Partec FloMax software (v2.4d) (Partec GmbH, Münster, Germany) in the form of six graphics: fluorescence pulse integral in linear scale (FL); forward light scatter (FS) *vs*. side light scatter (SS), both in logarithmic (log) scale; FL *vs*. time; FL *vs*. fluorescence pulse height; FL *vs*. FS in log scale and FL *vs*. SS in log scale. Mean fluorescence values and coefficient of variation (CV value) of the fluorescence of both sample and standard were obtained for at least 1,300 nuclei in each G_1 peak, whenever possible. Samples with CV values above 5% were discarded, prepared and ran again. At least three individuals from every population were used to estimate genome size (Appendix 3), in different days, to account for the variation generated by the flow cytometer. The remaining individuals were analyzed in pool (three or four individuals) to determine ploidy level (Appendix 4), only. The absolute DNA content of a sample was calculated based on the following formula: 2C nuclear DNA content of the sample = (sample G_1 peak mean) / (standard G_1 peak mean) × 2C DNA content of standard. Descriptive statistics were calculated for genome size data (mean, standard deviation of the mean, standard error, coefficient of variation and minimum and maximum values) using Microsoft Excel 2016. Differences in average genome size values among regions were assessed by means of Linear Mixed-Effect Models with the formulation of Laird and Ware (1982), with a region as fixed factor and population within region as a random nested factor, in R-3.2.0 (R Development Core Team, 2010). Data was plotted in BoxPlotR (Spitzer et al. 2014).

RESULTS

Analysis of fresh leaf tissue sampled from seedlings germinated from wild seeds of individuals from 52 populations from Turkey, Spain, Argentina, Chile, USA and Australia (Appendix 2), showed no significant differences in genome size ($F_{5, 44} = 0.58$; $p = 0.716$) among regions (Fig. 6). All individuals (N = 477) were found to be diploid, presumably with $2n = 16$

chromosomes. Average genome size ranged from 1.70 pg/2C (SD = 0.06 pg) in Australia and Spain (SD = 0.06 pg) to 1.71 pg/2C (SD = 0.06 pg) in Chile, 1.72 pg/2C (SD = 0.06 pg) in Argentina and California (SD = 0.07 pg) and 1.73 pg/2C (SD = 0.07 pg) in Turkey (Table 1).

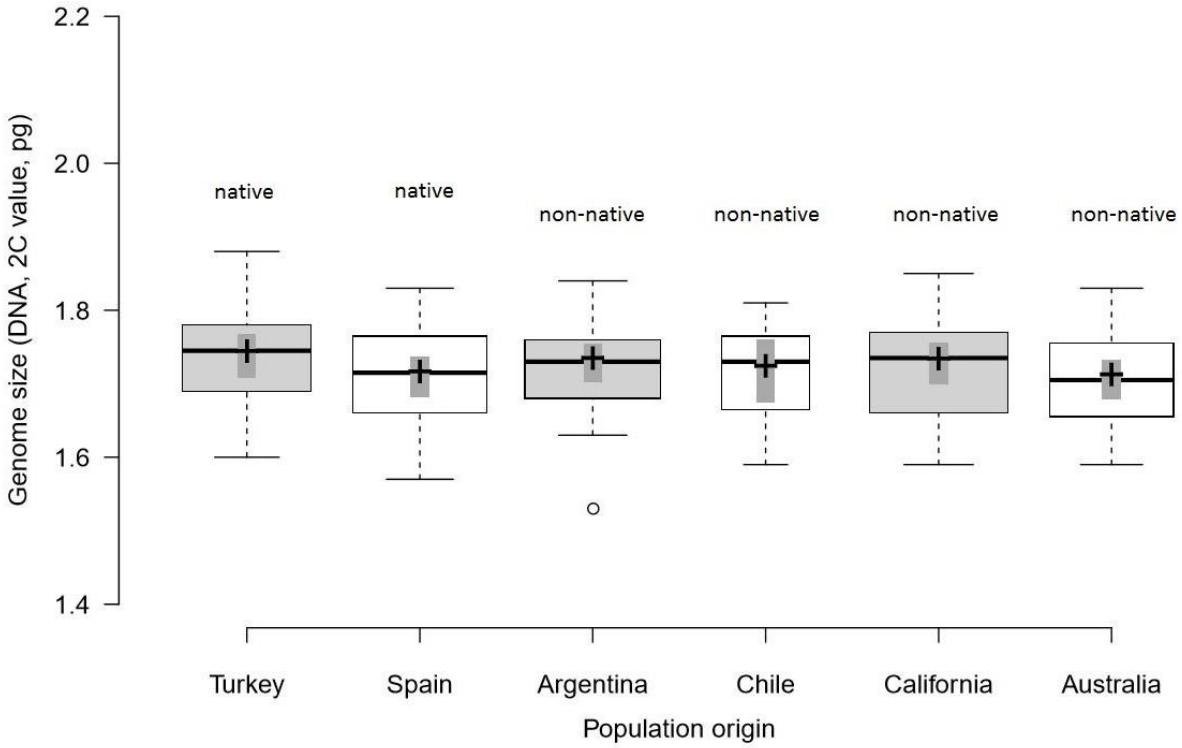


Figure 6. Comparison of genome size among native and non-native genotypes of *Centaurea solstitialis*. Black center lines represent the medians, crosses indicate sample means, box limits indicate the 25th and 75th percentiles, whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, bars show 95% confidence intervals of the means and outliers are represented by empty dots. Width of the boxes is proportional to the square root of sample size, n = 26, 28, 29, 12, 30, 24 sample points

Table 1. **Genome size estimations in *Centaurea solstitialis,* across the six regions**

Region	Genome size (2C, pg)					N
	Mean	SD	SE	Min	Max	
Argentina	1.727	0.067	0.012	1.53	1.84	29
Australia	1.705	0.061	0.012	1.59	1.83	24
California	1.727	0.074	0.013	1.59	1.85	30
Chile	1.717	0.065	0.018	1.59	1.81	12
Spain	1.709	0.069	0.013	1.57	1.83	28
Turkey	1.737	0.070	0.013	1.60	1.88	26
Total	**1.72**	**0.068**	**0.014**	**1.57**	**1.84**	**149**

Note: Values are given as mean, standard deviation and standard error of the mean. The minimum and maximum values and the number of analyzed individuals (N) for genome size estimations are also provided

Genome size variation among populations within regions (see Appendix 3) was also not significantly different, as indicated by very small standard deviations for the intercept and the residual obtained for the random effects ($SD_{intercept}$ = 0.024; $SD_{residual}$=0.063).

DISCUSSION

We found no traces of polyploidization events in the *C. solstitialis* populations investigated and geographic differences in genome size were negligible. A previous record of isolated tetraploids (one accession) in Northern Turkey (Inceer et al. 2007) is intriguing, since further genomic sampling in the area (e.g., less than 40 km from the initial site, Barker et al. 2017) did not validate the findings. Further investigation is also required to clarify the reported putative hybridization (Barker et al. 2017) with *Centaurea nicaeensis L.* (2*n* = 20 chromosomes, Guinochet and Foissac, 1962) since inter-specific hybridization does not seem to have played a significant role in the past invasion history of *C. solstitialis* (Barker et al. 2017). Formerly, a single natural

hybrid of *Centaurea* × *moncktonii* CE Britton and *C. solstitialis* was described from Oregon, USA (Roché and Susanna, 2010) and found to be a sterile triploid (Miskella, 2014). The genome size value we obtained for California (1.72 pg/2C, SD = 0.07 pg) was similar to the one previously reported for Southwestern Oregon (1.66 pg/2C, SD = 0.07 pg), by Miskella (2014) and, overall, genome sizes were similar among the six world regions.

CONCLUSIONS

In conclusion, our thorough sampling of the most representative native and non-native populations across the world's distribution of *C. solstitialis* indicates that its invasive success is not due to changes in genome size or ploidy level. We cannot discard that some individuals in some unsampled populations could present some degree of polyploidy, but their role in invasive success, to date, would have been of minor importance.

Chapter II – Biogeographic differences in the allelopathy of leaf surface extracts of an invasive weed

Chapter section published as an original research article in *Biological Invasions*:

Irimia RE, Lopes SMM, Sotes G, Cavieres LA, Eren Ö, Lortie CJ, French K, Hierro JL, Rosche C, Callaway RM, Pinho e Melo TMVD, Montesinos D (2019). Biogeographic differences in the allelopathy of leaf surface extracts of an invasive weed. https://doi.org/10.1007/s10530-019-02038-1.

ABSTRACT

Allelopathy, the release of chemicals by plants that inhibit the germination and growth of competing species, can be an important trait for invasive success. However, little is known about potential biogeographical differences in allelopathy due to divergent regional eco-evolutionary histories. To test this, we examined the allelochemical potential of the highly invasive species *Centaurea solstitialis* from six world regions including native (Spain, Turkey) and non-native ranges (Argentina, Chile, California and Australia). Seeds from several populations in each region were collected and grown under common garden conditions. Allelopathic potential and chemical composition of three leaf extract concentrations of *C. solstitialis* from each region: 0.25%, 0.5% 0.75% (w/v−1) were assessed on the phytometer *Lactuca sativa*. The main allelochemicals present in the leaf-surface extract were sesquiterpene lactones that varied in major constitutive compounds across regions. These leaf extracts had strong inhibitory effects on *L. sativa* seed germination and net growth. Summed across regions, the 0.25% concentration suppressed germination by 72% and radicle elongation by 66%, relative to the controls. At the 0.5% concentration, no seeds germinated when exposed to extracts from the non-native ranges of Argentina and Chile, whereas germination and radicle growth were reduced by 98% and 89%, respectively, in the remaining regions, relative to controls. Germination and seedling growth were completely inhibited at the 0.75% concentration extract for all regions. Some non-native regions were characterized by relatively lower concentrations of allelochemicals, suggesting that there is biogeographical variation in allelopathic expression. These findings imply that rapid selection on the biochemical signatures of an exotic invasive plant species can be highly region-specific across the world.

Keywords: biogeographical contrasts, biotic resistance, leaf-surface chemicals, phytotoxicity, sesquiterpene lactones, yellow star-thistle

INTRODUCTION

The release of compounds in leaf leachates, root exudates, volatiles, and from decaying plant material (Weir et al. 2004) can contribute to competitive interactions and plant defense (Müller-Schärer et al. 2004; Aschehoug et al. 2016). Allelopathic compounds can defend against pathogens (Meepagala et al. 2006; Zhang et al. 2013), inhibit germination (Chon and Nelson 2010), suppress other plant species (Ridenour and Callaway 2001; Lankau 2012), deter herbivores (Thelen et al. 2005), enhance nutrient acquisition (Tharayil et al. 2009), and ameliorate abiotic stressors such as high light intensity (Izhaki 2002). Additionally, some of these compounds, like emodine, appear to have multiple functions at a time, such as insect anti-feeding agent, seed dispersal facilitator and plant growth inhibitor (Hasan 1998; Inoue et al. 1992; Izhaki 2002).

Allelopathy has been linked to successful exotic invasion by non-native species (Callaway and Ridenour 2004). For example, the root extracts of the invasive *Centaurea diffusa* suppress *A. thaliana* seedlings (Quintana et al. 2009), whereas its shoot extracts inhibit the germination and growth of several *Lolium* species (Muir and Majak 1983). In some cases, phytotoxic compounds, such as polyacetylenes have been isolated from the roots of exotic invaders (*Acroptilon repens*, Quintana et al. 2008) and from soils occupied by exotics (Dayan et al. 2010). Collectively, this evidence suggests that plant compounds are powerful agents that can influence the outcome of interactions with other plant competitors and with consumers, and play an important role in at least some invasion processes such as interference or competitive interactions with other resident species.

Invasive species can rapidly develop different sets of adaptations to the different non-native regions they colonize (Maron et al. 2004; Callaway and Ridenour 2004; Graebner et al. 2012). If trait adaptation to local conditions is common, allelopathy should also experience selective forces both directly and indirectly via trade-offs and, in fact, the few studies that have compared allelopathy in native and non-native ranges have found significant differences (Yuan et al. 2012; Gruntman et al. 2015). However, previous studies tended to include one or two regional sites within the native or introduced ranges and found either increased allelopathy in the non-native range, or no differences (Gruntman et al. 2015). There is no reason to expect that all non-native regions will present similar levels of allelopathic effects, since diverging selective

forces and trade-offs could result in biogeographical changes in the relative importance of allelopathy in each range and region. Additionally, there may be variation in metabolic profiles and allelochemical production due to biogeographic history (Callaway and Aschehoug 2000; Schemske 2010; Rosche et al. 2018) or variation in soil nutrients, drought or thermal stress (Sampaio et al. 2016), potentially resulting in different trait shifts specific to ecologically unique regions. Therefore, it is crucial to conduct biogeographical experiments that compare allelochemical production and allelochemical effects for both native and introduced populations of invasive species.

Centaurea solstitialis (yellow starthistle, Asteraceae) is an ideal study system to investigate local adaptation rates and subsequent trait divergences. It is a Eurasian winter annual forb introduced throughout the world and highly invasive in the Americas. During its colonization history, the species had evolved many trait differences in its introduced areas, relative to the native range, such as increased seed size, germination timing and adult plant size (Graebner et al. 2012; Hierro et al. 2013; Barker et al. 2017), and increased competitive ability (Montesinos and Callaway 2017), among others. Other species in the genus were found to be allelopathic (Callaway and Aschehoug 2000; Ni et al. 2010; Chen et al. 2013; Aschehoug et al. 2014), but previous experiments did not find evidence of root mediated allelopathy in *C. solstitialis* (Carpenter 2007; Qin et al. 2007); however, some studies suggest that leaf leachates may be allelopathic (Zamora 1984; Gómez-González et al. 2009; Filipe et al. 2016). All aerial parts of *C. solstitialis* possess several sesquiterpene lactones (Stevens and Merrill 1985; Alvarez 2008; Sotes et al. 2015) which have well known pharmacological and biological inhibitory activity (Cheng et al. 1992; Özçelik et al. 2009), and that can act as an inducible chemical defense against herbivory (Beck et al. 2008; Smith and Beck 2013; Oster et al. 2015) or mediate competitive interactions with neighboring plants. In the Asteraceae, sequiterpenes are usually secreted and stored in glandular trichomes of the leaves or in epidermic folds or cavities due to their high toxicity (Göpfert et al. 2005) but they may also occur in secretory ducts inside the leaf mesophyll (Bartoli et al. 2011). Nevertheless, this aspect remains poorly studied in *C. solstitialis*. Stevens and Merrill (1985) tested the individual effects of several sesquiterpene lactones isolated from fresh leaves of *C. solstitialis* from California on lettuce seeds germination and seedling growth and found the

to possess growth regulatory activity, with low concentrations having stimulatory effects and higher concentrations displaying inhibitory effects. Hence, the chemicals present on the leaf surfaces of *C. solstitialis*, especially terpenoids, are potentially biologically relevant. They may present allelopathic activity, particularly as some of them have hydrophilic properties and can accumulate in the soil through rain or litter decomposition (Sotes et al. 2015). However, it is unknown whether *C. solstitialis* leaf surface chemicals are involved in allelopathic inhibition of other plants, or if there is any biogeographical variation in the composition and activity of leaf surface chemicals.

We addressed knowledge gaps relative to the biogeography of the allelochemical composition and allelopathic activity of invasive species by assessing (1) *C. solstitialis* leaf-surface chemistry for plants from six different native and non-native regions across the world, to determine whether there are biogeographical differences in allelochemical production, and (2) effects of leaf-extracts from these plants on seed germination and root growth of *Lactuca sativa*, a ''phytometer'' commonly used in allelopathic experiments, to test whether *C. solstitialis* leaf chemicals are effectively allelopathic.

MATERIALS AND METHODS

Plant material

Mature seeds of *Centaurea solstitialis* were collected in the field from several different maternal individuals from five natural populations in each of six different world regions, in the species' native range (Turkey and Spain) and non-native ranges (California, Argentina, Chile and Australia) (Fig. 7; Table 2). Turkey is regarded as the geographic center of the species origin (Uygur et al. 2004), and Spain the primary source of seeds that colonized South America (Eriksen et al. 2014; Barker et al. 2017), while the remaining four regions represent important introductions, where the species is considered invasive. Populations in each region were situated at least 30 km apart. Thus, our sampling was broad and considered the invasion pathways and historical relationships among different geographic regions. Seeds from ten individuals from each of the 30 populations selected were germinated in a glasshouse at the Botanical Garden of the University of Coimbra, Portugal, in 50 cell plug trays containing standard potting soil

(Substratos Profissionais, Leal and Soares S.A., Portugal).

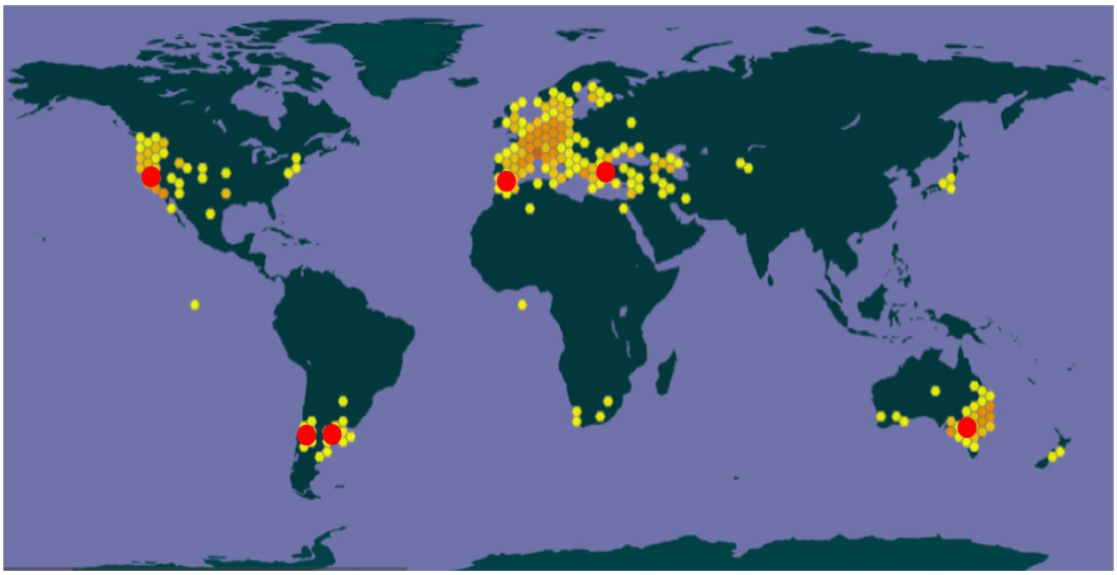


Figure 7. World distribution of *C. solstitialis* according to GBIF (yellow–orange dots), and sampled regions (red dots). Map adapted from GBIF (2017)

After 3 weeks, plants were transplanted into larger 2 L square plastic pots (1individual per pot), filled with standard potting soil and kept in the glasshouse on top of flow benches for the remainder of the experiment. Plants experienced natural sunlight and Mediterranean warm climate in Coimbra and were watered three times per week to guarantee no water shortage. Chemical fertilizer was supplied once during the experiment (Fertiberia Jardin, Spain, NPK: 8-4-6 plus microelements) by applying 25 ml of solution to each pot. We noticed no sign of nutrient deficiency, drought, or herbivory, that could have potentially altered the production of chemical compounds in the leaves. Commercial seeds of *Lactuca sativa* (lettuce; Bionda Degli Ortolani, Vilmorin, France) were used as model species for testing the effects of leaf extracts on seed germination and radicle growth. This species is recommended for the assessment of ecological effects of toxic compounds by several agencies and protocols, including the US Environmental Protection Agency (US EPA), the Organisation for Economic Co-operation and Development (OECD), and International Organisation for Standardization (ISO) (US EPA 1996; ISO 1995; OECD 2003).

Leaf-surface extracts

Three months after germination of *C. solstitialis*, the five largest fresh and healthy-looking rosette leaves (typically 120–150 mm long by 35–40 mm wide), were harvested from each of ten individuals per population, from 5 populations in each region. Leaves from all ten individuals within each population were pooled and their surface chemicals solvent-extracted by soaking leaves in 500 mL of CH_2Cl_2 (dichloromethane) at room temperature for 30 s, enough time to extract surface chemical compounds, but not enough to destroy tissues that would release internal leaf components (Sotes et al. 2015). Consequently, our leaf-surface extracts targeted those chemicals that are more likely to interact with other organisms. In contrast, crushed leaf extracts tend to include other constituent chemical compounds which could be less important for allelopathic interactions (Inderjit and Dakshini 1995).

Extracts from each of the 30 world populations were then filtered through 598 Whatman filter paper to remove debris. A final leaf-surface extract per region, n = 1 pooled sample for each of the six regions, was obtained by mixing together the extracts of the five populations within that region and concentrated using a vacuum rotary evaporator (IKA RV 8, Wilmington, USA). Plants and populations were pooled in order to obtain enough homogeneous extract from each region to conduct bioassays. Note that replication at the population level would have been desirable, because this would provide the opportunity to account for within-region variation (e.g., among and within populations). However, despite a design that accounted for within-region variation, the quantity of leaves obtained from individuals did not provide sufficient extract concentrations to accurately test for this, therefore we pooled extracts from populations within each region in order to develop the phytotoxicity bioassays.

Between 2 and 4 mg of each of the six pooled extracts were analyzed by gas chromatography coupled with mass spectrometry (GC–MS) on an Agilent Technologies 7820A System gas chromatograph linked to an Agilent Technologies 5975 Series MSD mass spectrometric detector as described by Sotes et al. (2015). The dried extract was stored in sterile glass vials at 4°C and used within one week in bioassays. Three different concentrations: 0.25%, 0.5% and 0.75% were obtained by re-suspending the dried extract in dichloromethane.

Table 2. GPS coordinates of the *C. solstitialis* populations sampled in this study (WGS84 datum)

Pop	Region	Range	Province	Locality	Latitude	Longitude
1	Turkey	native	Denizli	Pınarkent	37.802833	29.19525
2	Turkey	native	Burdur	Burdur	37.616083	30.146167
3	Turkey	native	Denizli	Serinhisar	37.531556	29.300861
4	Turkey	native	Izmir	Beydağ	38.085861	28.215472
5	Turkey	native	Izmir	Bozdağ	38.301361	28.049861
1	Spain	native	Cuenca	Moncalvillo de Huete	40.24159	-2.687453
2	Spain	native	Lleida	L´Espluga-Calba	41.50499	1.005857
3	Spain	native	Burgos	La Horra	41.728801	-3.834349
4	Spain	native	Valladolid	Castronuño	41.392459	-5.276957
5	Spain	native	Zaragoza	Sástago	41.408225	-0.289773
1	Argentina	non-native	La Pampa	Paraje El Tropezón, R14	-36.709	-64.831055
2	Argentina	non-native	La Pampa	El Durazno, R14 km 189	-36.700077	-65.391416
3	Argentina	non-native	La Pampa	Rucanelo, R11 y R10	-36.708944	-64.830833
4	Argentina	non-native	La Pampa	Quehué, R 35 y R18	-37.121611	-64.286611
5	Argentina	non-native	La Pampa	Unanue, R35 km 215	-37.559666	-64.2915
1	Chile	non-native	Talagante	Padre Hurtado	-33.570833	-70.855277
2	Chile	non-native	Talagante	El Monte	-33.689444	-71.055277
3	Chile	non-native	Santiago	Maipú	-33.524722	-70.751666
4	Chile	non-native	Santiago	Lo Barnechea	-33.37	-70.429722
1	California	non-native	Sacramento	Folsom	38.64215	-121.17596
2	California	non-native	Solano	Green Valley	38.20954	-122.14631
3	California	non-native	Napa	Napa	38.339041	-122.15467
4	California	non-native	Solano	Vacaville	38.41059	-121.934338
5	California	non-native	Napa	Napa	38.45353	-122.152875
1	Australia	non-native	NSW	Hume, 5km N of Holbrook	-36.677787	147.369684
2	Australia	non-native	NSW	Gundagai	-35.067197	148.108528
3	Australia	non-native	NSW	18 km N of Cudal	-33.22193	148.90875
4	Australia	non-native	NSW	Koorawatha	-34.016784	148.56897
5	Australia	non-native	NSW	Murringo	-34.319168	148.493212
6	Australia	non-native	NSW	Yass	-34.873953	148.908326

Lactuca sativa seed germination and radicle growth

Our experiment involved 18 treatments (6 regions x 3 different concentrations) and two controls (positive and negative controls). Each treatment had three replicates and each control had five replicates. Although extracts from each population were used only for three replicates each, the number of populations (N = 30), and the number of treatments resulted in a total number of experimental units of 64 (18 x 3 + 2 x 5). Thirty-six *L*. *sativa* seeds were placed on a Petri dish lined with one 90 mm Whatman filter paper (GE Healthcare) and moistened with 4 mL of plant extract, re-suspended in dichloromethane. We used distilled water as a negative control and dichloromethane as a positive control, with 4 mL of each applied to the filter paper, in order to control for potential effects of the extracting solvent in the germination tests. For all treatments, but the negative control, the organic solvent was evaporated in the fume hood followed by re-moistening of the filter paper with 3.5 mL of distilled water, following the standard protocol for germination bioassays (Vrchotová et al. 2011). Plates were sealed with parafilm to keep moisture in and stored at room temperature (25°C) in a completely randomized design. Germination was recorded daily for 14 days. A seed was categorized as germinated when a radicle of 1 mm was present. At the end of the experiment, we measured the length of the radicle of the germinated seeds to the nearest mm.

Data analysis

To quantify variation in chemical compounds among regions, we did a principal component analysis (standardized and centered PCA) on the chemical compounds that we found in our GC–MS analyses using the vegan 2.5-1 package (Oksanen et al. 2018) in R.3.5.1 (R Core Team 2018), employing the prcomp function to compute principal component scores. We then plotted our PCA with the ggbiplot 0.55 package (Vu 2011) and colored the regions in the PCA plot according to their assignment to the native and nonnative ranges. Further, we used the psych and stats packages in R (Revelle 2018) together with the panel.cor, cor.test and panel.hist functions to test for collinearity in the chemical dataset, generate pairwise scatter plots of all the variables and compute the Pearson correlation. Variables exhibiting significantly high correlation coefficients (P value ≤ 0.01) were dropped from the model and a PCA was performed on the

remaining variables to assess the levels of variance explained by the two axes and the phytochemical profiles of each region.

We calculated two germination indices: percent germination and days to germination using spreadsheets and commands developed by Ranal et al. (2009). Data were checked for heteroscedasticity and normality by Levene's and Shapiro–Wilk tests (Levene 1960; Shapiro and Wilk 1965). Data on seed germination (binomial) were analyzed with generalized linear models (GzLM), whereas data on untransformed radicle length (Gaussian) and days to germination (Poisson) were analyzed with linear models. To test for differences in the three response variables among regions, we ran models for each concentration with region as the fixed factor. If the models indicated significant differences among regions, we applied Tukey HSD post hoc tests with P-values ≤ 0.05 to infer which regions differed. To test for differences in the response variables between native versus non-native ranges, we generated generalized linear mixed-effects models (one for each concentration) with range as fixed factor and region as a random factor.

RESULTS

Chemical compounds in leaf-surface extracts

GC–MS phytochemical analysis of *C. solstitialis* leaf surface extracts identified three major classes of chemical compounds: alkanes, sesquiterpene lactones and pentacyclic triterpenoids (Table 3). The biologically important sesquiterpene lactones were the main class of substances. In total, we detected 24 compounds, including thirteen sesquiterpene lactones, there pentacyclic triterpenoids, and eight alkanes. The total unidentified fraction of the extracts ranged between 3.0–11% and the mass spectra of these extracts were similar to that of lactones. The sesquiterpene lactone fraction represented more than 64% of the final extract for all regions, except California where it was 56%. *Centaurea solstitialis* from California exhibited the largest diversity of sesquiterpene lactones of all regions (12 out of a total of 13 sesquiterpene lactones detected; Table 4). In contrast, 11 sesquiterpene lactones were identified in the Spanish extract, nine in Turkey, Argentina and Chile, and eight in Australia (Table 4).

Table 3. Major classes of chemical compounds (%), identified in the *C. solstitialis* final leaf-surface extract by GC-MS analysis

Compound (%)	**Turkey**	**Spain**	**Argentina**	**Chile**	**California**	**Australia**
Alkanes	18.34	18.13	14.54	10.18	28.66	14.69
Sesquiterpene lactones	70.95	64.61	76.71	78.26	55.85	78.32
Pentacyclic triterpenoids	3.51	5.39	4.54	4.96	5.64	2.4
Total identified	92.8	88.13	95.79	93.4	90.15	95.41
Total unidentified	6.7	10.89	4.16	5.74	6.75	3.31
Grand total	99.5	99.02	99.95	99.14	96.9	98.72

Repin was the major compound of the sesquiterpene lactone fraction in extracts from plants from Argentina (32%) and Spain (26%) followed by the *repin*'s isomer *subluteolide*, in extracts from Australia (31%) and Turkey (27%). Conversely, *epoxyrepdiolide* (19%), *janerin* (19%), *subluteolide* (15%) and *repin* (12%) were the major compounds in the Chilean extract, respectively *repin* (13%) and *epoxyrepdiolide* (12%) of the Californian extract. Likewise, *solstitialin A-13 acetate* concentrations displayed higher values in extracts of native plants from Turkey and Spain (8% and 10%) than in non-native plants from the remaining four regions (1–2%), whereas *solstitialin A-3 acetate* was found only in the Turkish and Spanish extracts, but in much smaller amounts (under 1%). *Cynaropicrin* derivatives were found exclusively in leaf extracts from non-native regions. Overall, *desacylcynaropicrin*, *aguerin B*, *cebellin F*, *cynaropicrin 3*, *4-diacetate* and *cynaropicrin 4-acetate* were present in very small amounts (under 1% of the final extract composition) across all regions (Table 4).

Table 4. Sesquiterpene lactones identified in the *C. solstitialis* final leaf-surface extract (%) by GC-MS analysis. The only compounds that appear to have clear biogeographical differences are shown in bold. Regions are abbreviated by two letters country code or three letters country code (Chile)

Compound (%)	TR	SP	AR	CHI	CA	AU
$C_{15}H_{18}O_4$ desacylcynaropicrin	0.32	0.37	0.79	0.94	0.82	0
$C_{17}H_{22}O_6$ solstitialin A-3 acetate	**0.94**	**0.32**	0	0	0	0
$C_{17}H_{22}O_6$ solstitialin A-13 acetate	**10.3**	**7.73**	2.02	1.12	0.85	1.62
$C_{19}H_{23}ClO_7$ acroptilin	2.33	3.09	3.28	5.75	3.77	2.46
$C_{19}H_{22}O_5$ aguerin B	0	0.14	0	0.7	0.47	0
$C_{19}H_{22}O_6$ epoxyrepdiolide	9.57	8.16	7.51	19.07	11.59	11.46
$C_{19}H_{22}O_7$ repin	12.22	26.04	31.58	12	12.88	9.31
$C_{19}H_{22}O_7$ subluteolide	26.5	8.28	9.01	14.57	6.4	30.55
$C_{19}H_{22}O_7$ janerin	8.94	9.25	14.11	18.87	8.38	19.85
$C_{20}H_{24}O_6$ cebellin F	0.12	0.14	0	0	0.22	0.1
$C_{21}H_{24}O_7$ cynaropicrin 3-acetate	0	0	**7.51**	**5.94**	**8.83**	**2.97**
$C_{21}H_{24}O_7$ cynaropicrin 4´-acetate	0	0	0	0	0.16	0
$C_{23}H_{26}O_8$ cynaropicrin 3,4´-diacetate	0	0.1	0.26	0	0.21	0

To our knowledge, this is the first evidence of *cebellin F* presence in *C. solstitialis*, although *cebellin C* had been previously reported (Bruno et al. 2013). The concentrations of pentacyclic triterpenoids (waxes) were similar among regions (Table 5). Californian plants had high concentrations of alkanes (29%), whereas Chilean plants had lower concentrations (10%) comprising just three alkanes (all linear) out of a total of eight identified in this study (Tables 6). Pentacosan ($C_{25}H_{52}$) was identified only in the Californian extract. *Dotriacontane* ($C_{32}H_{66}$) and tritriacontane ($C_{33}H_{68}$) are reported for the first time in C. solstitialis (Table 6).

Table 5. Pentacyclic triterpenoids (%) identified in the *C. solstitialis* final leaf-surface extract per region, by GC-MS analysis

Compound (%)	Turkey	Spain	Argentina	Chile	California	Australia
$C_{30}H_{50}O$ β-Amyrin	1.29	1.86	1.57	1.46	1.83	0.81
$C_{30}H_{50}O$ α-Amyrin	1.01	2.55	1.80	2.27	2.56	1.19
$C_{30}H_{50}O$ taraxasterol	1.21	0.98	1.17	1.23	1.25	0.40

Table 6. Alkanes identified in the *C. solstitialis* final leaf-surface extract per region (%) by GC-MS analysis

Compound (%)	Turkey	Spain	Argentina	Chile	California	Australia
$C_{25}H_{52}$ (Linear)	0.00	0.00	0.00	0.00	0.32	0.00
$C_{27}H_{56}$ (Linear)	2.10	2.53	1.34	1.63	3.44	1.42
$C_{29}H_{60}$ (Linear)	6.77	4.14	2.48	5.30	8.41	2.17
$C_{29}H_{60}$ (Branched)	0.00	4.26	3.92	0.00	6.12	3.34
$C_{31}H_{64}$ (Linear)	2.14	1.60	1.41	3.25	2.12	2.02
$C_{31}H_{64}$ (Branched)	4.74	4.17	3.91	0.00	5.88	4.57
$C_{32}H_{66}$ (Branched)	1.45	0.45	0.39	0.00	0.80	0.00
$C_{33}H_{68}$ (Branched)	0.60	0.98	1.09	0.00	1.67	1.17

The first and second axes of the PCA explained 66% of the inertia variance (Fig. 8). The first axis was mainly correlated with the compounds *janerin*, *epoxyrepdiolide*, and *cebellin F*, and therefore represents gradients in sesquiterpene lactones. The second axis was mainly correlated with *desacylcynaropicrin*. The PCA revealed a clustering of the two native regions in the lower right quarter, corresponding to high contents of *solstitialin A3-acetate*, *solstitialin A13-acetate*, and *hentriacontane* ($C_{31}H_{64}$). In contrast, the non-native regions were distributed widely across the other three quarters of the PCA plot. The collinearity test showed a significant positive relationship between *solstitialin A3-acetate* and *solstitialin A13-acetate* (r = 0.939), *acroptilin* and *aguerin B* (r = 0.922), *cynaropycrin 3-acetate* and *nanocosan* (branched $C_{29}H_{60}$) (r = 0.963) and *cynaropicrin 4-acetate* and *pentacosan* ($C_{25}H_{52}$) (r = 1) (Appendix 5). After dropping out *solstitialin A3-acetate*, *aguerin B*, *nanocosan* and *cynaropicrin 4-acetate* from the model, the PCA explained 66% variance in the two axes, similar to the initial PCA on all the variables, but revealed a slightly different region clustering based on chemical compounds (see Appendix 6).

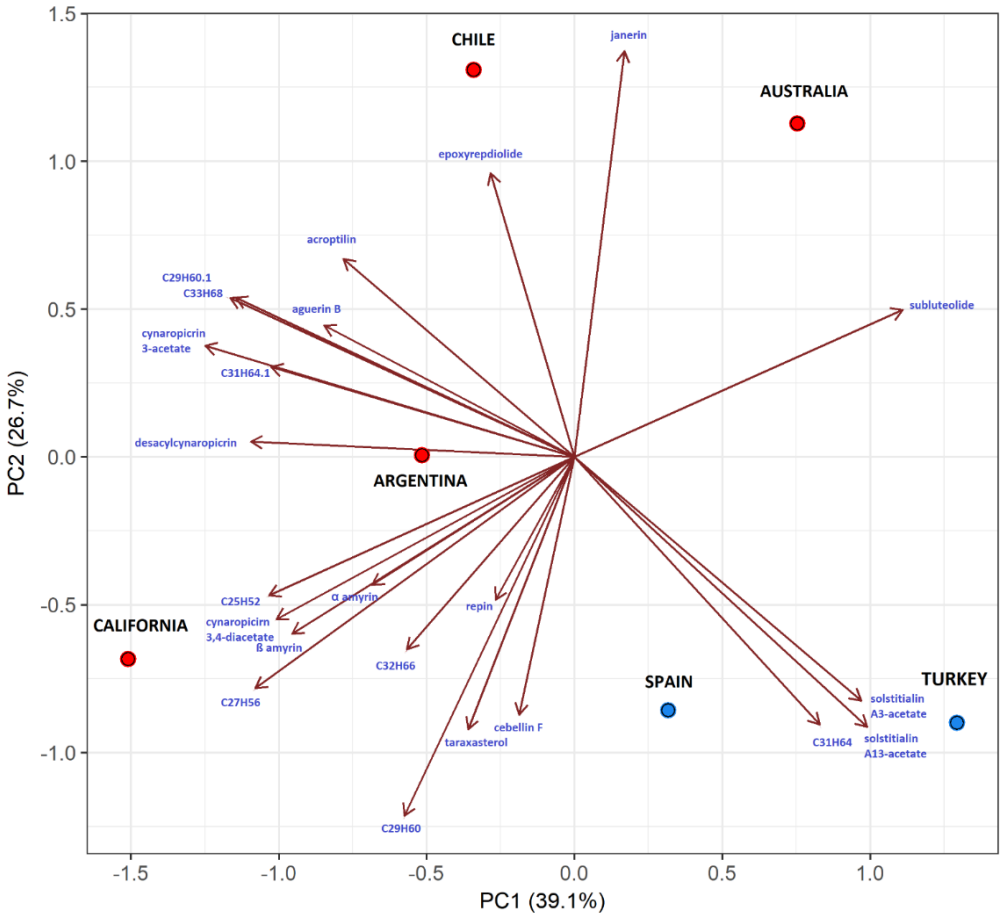


Figure 8. Plot of Principal Component Analysis scores (centered and standardized PCA) representing the multivariate space in chemical compounds found in six *C. solstitialis* world regions. Blue dots represent native regions, and red dots non-native region

Effects of leaf-extracts on *Lactuca sativa* germination

Lactuca sativa control seeds germinated at an overall rate of 99% ± 0.2% (mean ± SE), thus neither deionized water nor dichloromethane affected germination. Overall, the average germination rate for *L. sativa* seeds exposed to the lowest leaf-extract concentration (0.25%) was 28%, followed by an overall germination rate of 2.0% for the 0.5% extract, and 0% germination for seeds exposed to an extract concentration of 0.75% (Fig. 9). The lowest concentration of leaf extract (0.25%), drove differences in germination probability among regions ($\chi^2_{(1)} = 122.9$, P = 0.001; AIC null model = 102.3 vs. AIC full model = 215.2). However, there were no differences between the native and non-native ranges in germination probability ($\chi^2_{(1)} = 0.167$,

P = 0.682; AIC null model = 120.3 vs. AIC full model = 118.5). Post-hoc tests revealed that germination in Australia and California extracts was higher than in extracts from the other regions (60%, and 45%, respectively; see Appendix 7, Fig. 9). Leaf-extracts of plants from Turkey (21%), Argentina (19%), and Spain (17%) resulted in stronger germination inhibition; and plants from Chile resulte resulted in the highest germination inhibition (4.0% germination). Germination of seeds exposed to an intermediate concentration of leaf extract (0.5%) was different among regions overall ($\chi^2_{(1)}$ = 19.14, P = 0.001; AIC null model = 38.65 vs. AIC full model = 47.79) and there were differences between the native and non-native ranges too ($\chi^2_{(1)}$ = 4.569, P = 0.032; AIC null model = 40.71 vs. AIC full model = 43.28).

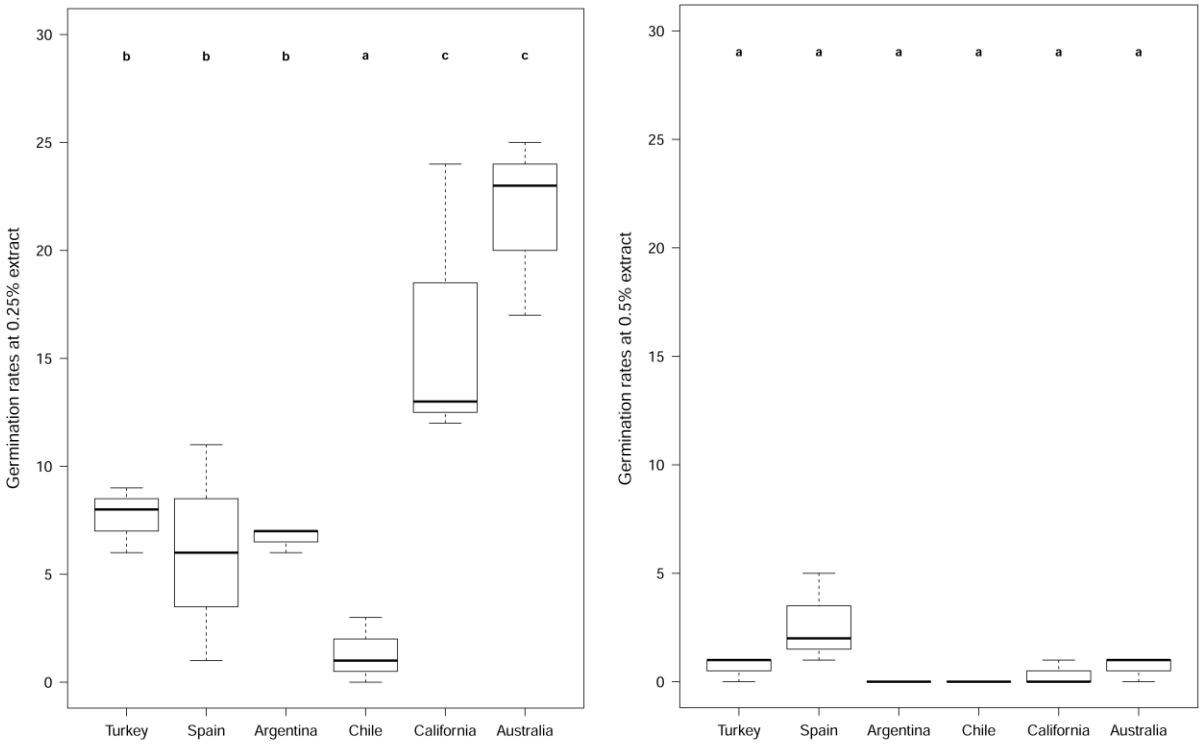


Figure 9. Effect of two different concentrations of *C. solstitialis* leaf extract on lettuce seed germination. Different letters indicate statistically significant differences among groups ($p < 0.05$). Both positive and negative control had a germination rate close to 100%

Leaf-extracts from Spain resulted in the highest germination (7.4%), which was still lower than any of the germination rates in the 0.25% concentration. Leaf-extracts from all other regions suppressed germination to < 2% with no seeds germinating when exposed to extracts from plants from Chile and Argentina. However, post hoc tests showed no significant difference among the groups tested (P-value > 0.05), (see Appendix 8, Fig. 9). At this concentration, non-native populations suppressed germination more than natives (0.7% vs. 4.62%), and this was generally determined by populations from Argentina and Chile which had zero germination. Exposure of seeds to the highest leaf-extract concentration (0.75%) completely suppressed seed germination for *C. solstitialis* from all population sources.

Lettuce seeds in both control groups took an average of 2.2 days to germinate after being sown, reaching a peak in the third day (Appendix 9). Seeds exposed to leaf-extract concentrations took longer times to germinate. The number of days to germination for seeds exposed to 0.25% leaf extract was different among regions (F = 3.091, df = 5, 171; P = 0.010, AIC null model = 192.3 vs. AIC full model = 197.7), with seeds treated with the Australian extract emerging first (4.7 days ± 0.21) followed by seeds exposed to Californian, Spanish, Chilean and Turkish extracts (all longer than 5 days) whereas seeds treated with Argentinean extract took the longest to germinate (6.0 days ± 0.47), (Appendix 9, Appendix 10, Fig. 10).

However, no differences were found between the native and non-native ranges ($\chi^2_{(1)}$ = 2.332, P = 0.126, AIC null model = 723.8 vs. AIC full model = 724.2). Seeds exposed to the 0.5% extract showed no differences in days to germination neither among regions (F = 0.397; df = 3, 9; P = 0.757, AIC null model = 13.84 vs. AIC full model = 9.460), nor between ranges ($\chi^2_{(1)}$ = 0.226, P = 0.634, AIC null model = 57.66 vs. AIC full model = 55.88) (Appendix 11, Appendix 12, Fig. 10).

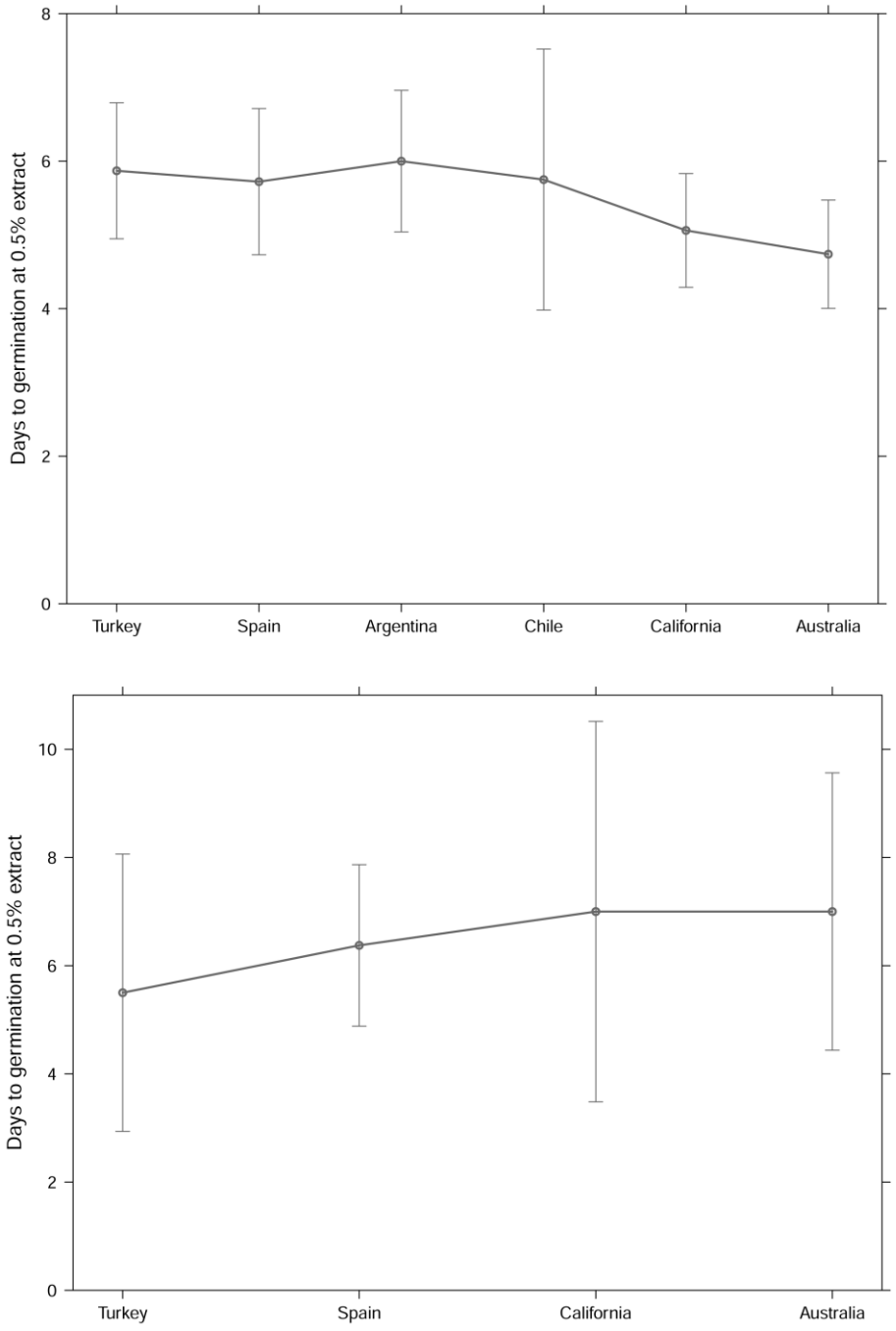


Figure 10. Effect of two different concentrations of *C. solstitialis* leaf extract on days to germination. Both positive and negative control took 2.2 days to germinate. No seeds from Chile and Argentina germinated at 0.5% extract

Effects of leaf-extracts on *Lactuca sativa* radicle length

Control seedlings achieved an average root length of 12.0 ± 0.06 mm (mean ± SE) for positive control and 9.3 ± 0.04 mm for negative control. Radicle lengths of germinants exposed to 0.25% leaf extracts were much shorter than in the controls, and significantly differed among regions overall (F region = 4.091, df = 5, 171; P = 0.001; AIC null model = - 421.9 vs. AIC full model = - 411.9). No significant differences were found between the native and non-native ranges ($\chi^2_{(1)}$ = 0.321, P = 0.570; AIC null model = 89.08 vs. AIC full model = 87.40). *Lactuca sativa* seeds exposed to 0.25% Australian, Californian and Spanish leaf extracts produced roots that were 60% shorter than those of the controls; and seeds exposed to extracts from Chile, Turkey and Argentina produced roots that were about 80% shorter than the controls (Appendix 13, Fig. 11). The average radicle length of the seeds in the 0.25% extract was 3.6 ± 0.02 mm.

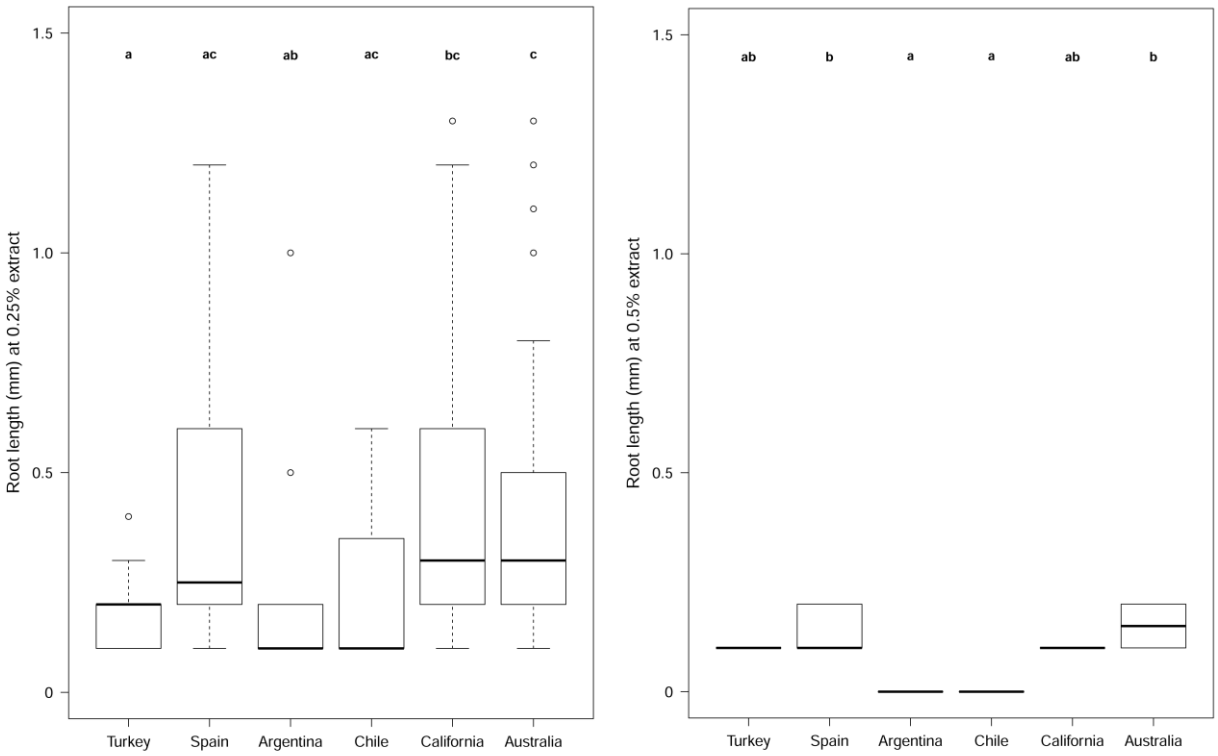


Figure 11. Impact of two different concentrations of *C. solstitialis* leaf extract on lettuce seedling radicle length (mm). Different letters indicate statistically significant groups (p<0.05)

Seeds exposed to 0.5% treatment also showed significant differences among regions overall (F region = 8.117, df = 5, 13; P = 0.001, AIC null model = - 115.0 vs. AIC full model = - 98.09), but there were no differences between the native and non-native ranges ($\chi 2(1) = 1.625$, P = 0.202; AIC null model = - 49.41 vs. AIC full model = - 49.79). No seeds exposed to extracts from Argentina or Chile germinated, but root length decreased more than 85% in response to extracts from Australia and Spain, and more than 90% in response to extracts from Turkey and California, compared to the controls (Appendix 14, Fig. 11). Average radicle length was 1.3 ± 0.01 mm for this treatment.

DISCUSSION

Our findings supported the potential allelopathic effects of *C. solstitialis* leaf leachates, and showed substantial variation in leaf-surface chemical composition and inhibitory effects among world regions. Allelopathic effects were strikingly weaker in the non-native regions of California and Australia than in any of the other four studied regions worldwide. Extracts from introduced populations in Chile and Argentina showed strong inhibitory effects, similar to those from native populations in Spain and Turkey. This suggests that *C. solstitialis* may have developed different leaf chemical profiles in different parts of the world where it has been introduced.

To the best of our knowledge, this is one of the first studies to report lower allelopathic effects in introduced regions of an invasive compared to native regions (but see Lankau et al. 2009). This suggests that chemical defenses might not be as important for invasive success in these regions as they are in other non-native regions and in the native range, potentially due to shifting selection pressures such as biogeographical differences in the relative importance of competition or herbivore release (Gruntman et al. 2017). However, confirming such a conclusion would require expanding the target species to include those that are dominant in each particular non-native region.

Geographically, leaf-surface extracts of *C. solstitialis* were highly variable in the composition of alkanes, sesquiterpene lactones, and pentacyclic triterpenoids; but with a clear predominance in all regions (≥ 64%) of the biologically important sesquiterpene lactones. Leaf-surface composition of sesquiterpene lactones as a group was similar between regions, but

sesquiterpene lactone sub-types differed among regions. Most of the sesquiterpene lactones identified in this study have been shown to possess a wide spectrum of biological activities and to be active at very low concentrations which makes them good candidates as allelopathic compounds. The mode of action of sesquiterpene lactones is not well understood, but we know that they can interfere with cell membrane functions and protein synthesis, and cause oxidative stress (Bachelier et al. 2006). For example, *acroptilin* and *repin* from *C. solstitialis* were found to promote lettuce root elongation at 10 ppm, but inhibit elongation at 80 ppm (Stevens and Merrill 1985). Both *repin* and *subluteolide*, the major sesquiterpene lactones identified in our study are highly reactive epoxides (Burrows and Tyrl 2013), with repin surpassing the toxicity of *subluteolide* > *janerin* > *cynaropicrin* > *acroptilin* > *solstitialin* in an *in vitro* cytotoxicity bioassay (Riopelle and Stevens 1993). *Repin* was also found to inhibit the growth of plants in vitro (Stevens et al. 1990) and interfere with insect larval growth and deter insect herbivory (Rosinski et al. 1988). Another class of compounds, *cynaropicrin* derivatives have well-documented antimicrobial and anti-insect activities (Bachelier et al. 2006; Bhattacharyya et al. 1996; Cis et al. 2006) and in this study were found exclusively in non-native regions. Similarly, *solstitialin*, a chemical specific to *C. solstitialis* (Heywood et al. 1977) could potentially represent a novel toxin in the non-native areas, particularly because of its high toxicity against eukaryotic and prokaryotic cells (Cheng et al. 1992). Although both *solstitialin A 3-acetate* and *solstitialin A 13-acetate* were present in the leaf extracts, only the latter is known to exhibit biological activity (Özçelik et al. 2009), but its concentrations where higher in natives plants relative to non-natives. Lastly, *janerin*, a compound displaying promising insect antifeedant activity (Cis et al. 2006), also exhibited levels similarly high across all regions. Our PCA analysis indicated that each region had somewhat unique concentrations of many of the major potential allelochemicals, such as *solstitialin A3-acetate* and *solstitialin A13-acetate* in Spain and Turkey, *janerin* and *subluteolide* in Australia, *repin* and *cebellin F* in Argentina and California or a unique ‘‘cocktail’’ of different chemicals such as *epoxyrepdiolide*, *acroptilin*, *aguerin B*, and *desacylcynaropicrin* in Chile. Note that our PCA analysis was based on 24 variables (representing all the chemical compounds identified in the extract), and explained 66% of the inertia variance in the first two axes (PC1 and PCA2). For this type of chemical compound analysis it is well known that the more variables, the

less explanation in the multivariate ordination. Other studies on plant phytochemical profiles reported similar or lower values of variance for constituents of essential oil of *Dracocephalum kotschyi* (PC1 + PC2 = 68%) (Jalaei et al. 2015) and *Citharexylum spinosum* (PC1 + PC2 = 60%) (El Ayeb-Zakhama et al. 2017), and for chemical compounds of *Alliaria petiolata* leaves (PC1 + PC2 = 58%) (Barto et al. 2010). Sotes et al. (2015) also profiled the leaf chemistry of *C. solstitialis* from four of the six world regions included in our study, but they did not test for allelopathic effects. That previous study found a similar sesquiterpene lactone fraction as described in this study, with the exception of California, where sesquiterpene lactones concentration as a group was higher. We found fewer different alkanes than Sotes et al. (2015), and for the total amount of alkanes we found 74% more in Spain and California, 44% more in Turkey, and 23% more in Chile. These differences could be due to phenological changes in the chemical content of leaves (see Alvarez 2008; Locken 1985; Geppert et al. 1983), since we sampled rosette leaves in pre-reproductive adult plants, whereas Sotes et al. (2015) sampled flowering plants.

We found substantial geographical variation in chemical profiles, but we could not unequivocally connect specific chemical compound concentrations and the effects of extracts on germination and radicle growth rates. This may be because many different chemical compounds have similar phytotoxic effects or that interactions between different chemicals create effects that are not explained by the individual chemical concentrations. For example, triterpenes can also disrupt cell membranes (Almeida et al. 2002). Moreover, some of the chemical compounds present in the extracts could not be identified (3–11%), although their absorption spectra indicated similarity to lactones, and these may have enhanced or diminished the biological potency of the different extracts. A recent study has found reduced diversity in leaf microbiome in invading *C. solstitialis* plants compared to natives (Lu-Irving et al. 2019). Our plants were subjected to uniform growing conditions and it is unlikely that some extracts were potentiated by the presence of co-extracted bacterial metabolites.

Our study targeted all chemical compounds on the leaf surfaces, but under natural conditions different compounds may have different rates of release and accumulation into the soil and may not act simultaneously. Additionally, we used a target species which is known to be sensitive to phytochemicals at low concentrations, but the response might have varied with other

bioassay species. For instance, a 0.5% w/v aqueous extracts of intact green, senesced leaves and combined ground leaves of *C. solstitialis* from Idaho did not affect the germination of *Bromus tectorum*, *Agropyron spicatum*, *Agropyron intermedium*, *Festuca idahoensis* and *C. solstitialis* itself under greenhouse conditions but negatively impacted radicle elongation for all species (Zamora 1984). An important point to consider in allelopathy is the concentration and persistence of the terpenoids in soil. This can be transient and influenced by the chemico-physical properties of the substrate and the microbial communities (Kokalis-Burelle and Rodríguez-Kábana 2006). For instance, Picman (1987) found that the compound isoalantolactone disappeared completely from soil after 3 months. Another study showed that parthenin can be rapidly degraded in the soil, in the course of a few days, depending on the dose and the experimental conditions (Belz et al. 2009). Similarly, very small quantities of cnicin were found in soil infested with C. maculosa (Locken and Kelsey 1987).

The allelochemical concentrations detected in our study may only partially reflect the real field situation where different biotic and abiotic factors come into play. Certain sesquiterpene lactones in *C. solstitialis* were found to vary seasonally in their concentration in plants from California, with an increase in these compounds as the plants matured (bolting and flowering stage), compared to the rosette stage (Alvarez 2008). However, they found no difference in chemical composition and concentration between different plant parts tested (e.g., leaves, stems and flowers) suggesting that they may all contribute to a potential phytotoxicity during plant life and after plant decomposition; and given the high biomass production in this species it may also result in residual allelopathy (Alvarez 2008). In California, *C. solstitialis* has an extended reproductive season, and flowers after the early season native annuals have set seeds (Roche et al. 1994). The plant starts to lose its leaves by early fall and seeds begin to germinate with the arrival of the fall rains (Sheley et al. 1999). This corresponds with the emergence of native and other exotic species, which makes forms of interference possible. In naturally occurring *C. solstitialis* infestations, averaging a density of 240 plants/m2, decaying shoot tissue represents 1.38% w/w of the soil sample (Zamora 1984). Soil mixed with ground *C. solstitialis* shoots and roots and supplemented with nutrients showed increasing inhibition of several plant species with increasing residue concentration in the soil (Zamora 1984). Likewise, soil from field plots with

natural vegetation, *C*. *solstitialis* alone and no vegetation showed that litter removal from soil significantly increased the root length of *B*. *tectorum* and *C*. *solstitialis* (Zamora 1984) but it was not clear whether this was due to nutrient deficiency or allelopathic compounds present in the soil.

CONCLUSIONS

Our bioassays showed that *C*. *solstitialis* leaf leachates can have allelopathic effects, but those effects present substantial biogeographical differences. These differences suggest that there can be lower allelopathic investments in some non-native regions where *C*. *solstitialis* is nonetheless highly invasive, compared to the native regions. We found differences in allelopathic effects among *C*. *solstitialis* regions, with plants from some of the non-native regions having weaker effects on lettuce germination, days to germination, and radicle elongation than plants from the native regions. The negative effects of leaf extracts from *C*. *solstitialis* from non-native California and Australia were far lower than for plants from the other regions, with extracts from Chile having stronger negative effects than plants from the other regions. Collectively, this suggests that rapid selection on the biochemical signatures of an exotic invasive plant species can be highly region-specific in different regions globally. Alternatively, founding effects due to introductions from different source populations in different world regions might have resulted in inherently different chemical signatures among regions. Our study is among the first to focus on biogeographic differences in the chemistry composition of allelochemicals and allelopathic effects, thereby providing fundamental insights into plant–plant chemical interactions in an invasion context. The difficulties intrinsic to international sampling and the high amounts of leaf tissue needed for chemical extraction are likely the reason why, to our knowledge, no other study has been able to assess biogeographic variations of allelopathy in more than 2–3 world regions previously. Further studies are now needed to lay emphasis on within region-variation, i.e., within and among populations of each distinct biogeographical region.

Chapter III - Experimental admixture between neo-allopatric regions of an invasive plant yields a global mosaic of reproductive incompatibility and heterosis

Chapter section in preparation for submission as an original research article to SCI journal:

Irimia RE, Hierro JL, Branco S, Sotes G, Cavieres LA, Eren Ö, Lortie CJ, French K, Callaway RM, Montesinos DM. Experimental admixture between neo-allopatric regions of an invasive plant yields a global mosaic of reproductive incompatibilities and heterosis.

ABSTRACT

Invasive species can rapidly adapt to new introduced regions. Classic evolutionary theory predicts that the accumulation of genetic differences over time in allopatric isolation can lead to reproductive incompatibilities resulting in decreases in reproductive success and, eventually, to speciation. However, the evidence for this theoretical prediction in the context of invasive species is anecdotal. We tested the hypothesis that allopatry influences reproductive output in geographically isolated and divergent genotypes of invasive species, when regional backcrossing occurs. We conducted a greenhouse experiment outcrossing *Centaurea solstitialis* individuals originating from 20 source populations in the native range and 30 source populations in four different non-native ranges and evaluated reproductive success as seed set. This invasive weed has shown some incipient degree of reproductive isolation between populations in California and Spain previously, and represents a good candidate to test for early signs of speciation. We found mixed fitness effects, ranging from increases to decreases to no net difference in fertility. Non-native populations in the Americas experienced a pattern of reduced fertility ranging between 18% to 42% when crossed with paternal genotypes from the native regions. Moreover, these populations also displayed up to 50% decrease in fertility when mixed with other non-native populations in the Americas suggesting population divergence there. By contrast, native maternal plants from Spain benefited the most from crosses with all the non-native regions, displaying increases in fertility from 44% to 70%, compared with crosses within region. Our results indicate an asymmetrical response to inter-regional gene flow, with non-native regions displaying mostly detrimental fitness effects indicative of reproductive incompatibilities while certain native regions displayed a considerable increase in fertility suggesting heterosis. The sensitivity and specificity of each region to admixture varied extensively, and it was not predicted by geographic distance. Our results provide a uniquely comprehensive view of the effects of geographical isolation on reproduction, and suggest that potential introductions of new genotypes into different non-native ranges could show unpredictable reproductive output.

Keywords: biological invasions, admixture, reproductive isolation, gene flow, allopatry, fertility, yellow starthistle

INTRODUCTION

Invasive species provide insights into evolutionary processes because they often show remarkable adaptive potential (Maron 2004, Callaway and Maron 2006; Lavergne and Molofsky 2007; HilleRisLambers et al. 2013; Rosche et al. 2019). Invasive plants prove that local adaptation could occur much faster than initially postulated (Oduor et al. 2016). The ability to adapt to new environments depends on sufficient genetic variability, which is why many successful invaders appear to experience multiple introductions before they successfully establish and expand (Ellstrand and Schierenbeck 2000; Marrs et al. 2008). In the initial colonization stage, low genetic variability is common, and the introduction of new genotypes can increase reproductive outputs and fitness by reducing post-colonization inbreeding depression, i.e. heterosis (Dlugosch and Parker 2008; van Kleunen et al. 2015). However, once a species becomes invasive, accumulation of habitat specific adaptations over time can lead to increasing degrees of reproductive isolation (Lowry et al 2008; Montesinos et al. 2012). This isolation can be reinforced by selecting against immigrant genotypes if local genotypes have higher fitness in local habitats (Hopkins 2013).

Experimental crosses within and between geographically isolated lineages or populations of the same species can provide insight into the effects of gene flow on fitness and the rates at which reproductive isolation develops (Waser and Prince, 1989). Some authors have regarded cross-population reproductive success as roughly equivalent to the degree of reproductive isolation between populations (Verrell and Arnold, 1989; Sapir and Mazzucco, 2015). But, only a few studies have addressed the mating success and potential consequences of admixture between native and introduced populations of invasive species (Wolfe et al. 2007; Montesinos et al. 2012; van Kleunen et al. 2015, Dlugosch et al. 2015, Dlugosch et al. 2018; Shi et al. 2018, Filipe and Montesinos 2016). Admixture is regarded as intraspecific hybridization between different source populations and can either enhance progeny performance across multiple generations or lead to reduced fitness benefits when highly divergent populations are mixed (Verhoeven et al. 2011). For example, crosses between populations from one native and two non-native regions of *Mimulus guttatus* led to an increase in biomass and seed production in the first generation progeny (van Kleunen et al. 2015). Another study found a 50% reduction in seed-sets of intercontinental crosses between populations from one native and one non-native range

of *Centaurea solstitialis,* compared to seed-sets of crosses within regions (Montesinos et al. 2012). Additionally, F1 progenies of *C. solstitialis* derived from crosses among geographically distinct populations in the native range, exhibited mainly positive fitness interactions in plant growth rates, consistent with heterosis (Barker et al. 2019).

Centaurea solstitialis is native to Eurasia, and was introduced to Australia and the Americas, where it is an aggressive invader (DiTomaso et al. 2006). Genomic data and historical records indicate that *C. solstitialis* populations from Western Europe were the primary source of introductions to Chile, and from Chile the species was introduced to California and Argentina. Several introduced populations in western North America appear to be the result of admixture with native populations, as a consequence of past multiple introduction events (Barker et al. 2017). In less than 200 years since its introduction, the species had evolved strong phenotypic divergence for numerous traits between native and introduced areas including divergence in seed size, germination, and plant size (Hierro et al. 2009; Graebner et al. 2012, Hierro et al. 2013), higher seed starch content (Widmer et al. 2007), and early flowering phenology (Dlugosch et al 2015). Moreover, there is evidence for incipient reproductive isolation between one native and one non-native range (Montesinos et al. 2012).

Allopatric evolution predicts that accumulation of adaptive differences among geographically isolated populations should lead to reproductive incompatibilities, resulting in decreases in reproductive success, a path to speciation (Via 2009). Invasive species provide an optimal opportunity to test this prediction, since approximate introduction times are known and their distributions are typically composed of geographically isolated ranges in different parts of the world. Here, we test the hypothesis that newly developed allopatry can result in relatively fast divergence in reproductive success for crosses between different ranges of invasive plant species. We predict that admixture between newly allopatric populations can result in differences in seed set due to either heterosis or incipient reproductive isolation. We collected seeds of the invasive weed *C. solstitialis* from numerous individuals from each of 50 populations from six native and non-native regions around the world, and germinated and grew them to maturity under common garden conditions in a greenhouse. Crosses within and between regions were used to assess the effect of admixture among geographically isolated ranges on

reproductive outputs.

MATERIALS AND METHODS

Study species

Centaurea solstitialis is a globally distributed annual herbaceous plant species found in disturbed habitats, meadows and fields. *Centaurea solstitialis* has composite flowers (capitulum) consisting of 20–80 single florets (Leong et al. 2014) and is predominantly self-incompatible (Maddox et al. 1996) relying on outcrossing to set seeds, although sporadic self-fertilization may sometimes occur too (Petanidou et al. 2012; Sun and Ritland, 1998). Irimia et al. (2017) found that *C. solstitialis* is diploid across its native and non-native ranges, ruling out polyploidy as a contributing factor for invasive success or reproductive incompatibilities. The flower heads produce two types of seeds: one that have a short plume (pappus) and the other lacking that plume (Maddox et al. 1996; Hierro et al. 2009). Populations of *C. solstitialis* occur allopatrically and are distinguished by a number of features including density of individuals in the field, seed size, germination timing, growth rate, adult plant size and competitive ability (Graebner et al. 2012, Hierro et al. 2013, Montesinos and Callaway 2017). In spite of such phenotypic differences, populations do not differ notably in genomic diversity and successful crosses between some geographically separated populations have been previously generated (Montesinos et al. 2012, Barker et al. 2019).

Pollination experiment

Mature seeds of *C. solstitialis* were collected in the field between 2009-2014, from 50 natural populations spanning localities in the native Turkey and Spain and in the non-native ranges of Argentina, Chile, California and Australia. Seeds originating from up to ten different maternal individuals per population (N = 339, see Appendix 15) were germinated in spring 2017 and grown to senescence in 2L pots filled with commercial soil (Substratos Profissionais, Leal and Soares S.A., Portugal), in a pollinator-excluded glasshouse at the Botanical Garden of the University of Coimbra, Portugal. Plants were watered daily and chemical fertilizer was supplied three times during the experiment, before the flowering onset (Fertiberia Jardin, Spain, NPK: 8-4-6 plus microelements). We sprayed the plants twice to stop fungal infection, each before the

onset of the anthesis, with Tebuconazol (Luna Experience, Bayer, containing 200g/L or 17.7% p/p fluopyram and 200 g/L or 17.7% p/p tebuconazol).

Controlled manual cross-pollinations were made by rubbing mature capitula (flower heads) to each other. Maturity was assessed based on the presence and abundance of pollen and on the receptiveness of stigmas, and preliminary tests across the flower anthesis period developed during previous years were used to determine the optimal phenological stage for manual cross-pollination. Only one pollen donor was used on each individual capitula. Treated capitula were grown until ripe and harvested for ovule and seed counting. Experimental crosses were conducted between the end of May to the end of September 2017.

Five different pollination treatments were applied to each individual plant: 1) spontaneous self-pollination; 2) manual self-pollination); 3) manual within-population pollination (crosses between two different individuals of the same population); 4) manual between-population within-region pollination (crosses between two individuals from different populations within the same region) and 5) manual between-region pollination (crosses between two individuals from two different randomly selected populations in two different world regions). All individual plants (339) used in this experiment acted as both pollen donors and pollen recipients. Since all treatments were replicated on each maternal plant, we eliminated possible confounding maternal effects and cross effects. In total, we performed 30 different between-region crosses combinations and six within-region crosses (Table 7), totaling 3922 yellow star-thistle capitula pollinated (excluding spontaneous and forced selfing) from 339 different plants.

Table 7. Matrix indicating the different types of manual crosses. The origin of the pollen donor is shown by columns, and the origin of the pollen receptor by lines. Grey indicates crosses within region

♀ \ ♂	**Turkey**	**Spain**	**Argentina**	**Chile**	**California**	**Australia**
Turkey	TRxTR	TRxSP	TRxAR	TRxCHI	TRxCA	TRxAU
Spain	SPxTR	SPxSP	SPxAR	SPxCHI	SPxCA	SPxAU
Argentina	ARxTR	ARxSP	ARxAR	ARxCHI	ARxCA	ARxAU
Chile	CHIxTR	CHIxSP	CHIxAR	CHIxCHI	CHIxCA	CHIxAU
California	CAxTR	CAxSP	CAxAR	CAxCHI	CAxCA	CAxAU
Australia	AUxTR	AUxSP	AUxAR	AUxCHI	AUxCA	AUxAU

Fruit harvesting and seed counting

Individual capitula were collected when ripe, approximately two weeks after pollination. Flower heads were tagged, stored in paper bags, and dissected in the laboratory to count the total number of ovules with pappus and ovules without pappus in a capitulum, and the number of viable and non-viable seeds present in each capitula. Seed viability was established visually, under a light microscope by gently pressing with tweezers on the seed to test for toughness and resistance to pressure. Preliminary tests involving seed dissection and germination of seeds collected in the wild allowed us to easily determine seed viability based on these attributes. We counted total ovule numbers of 1944 capitula from 339 plants, including: 225 spontaneous self-pollination, 313 manual self-pollination, 265 within population, 263 between populations within region, and 878 between region crosses.

Data analysis

All analysis were conducted in R v3.5.2. Data on ovule numbers and proportion of viable seeds *vs*. total number of ovules per capitula (seed:ovule ratio) were analyzed with generalized linear mixed models (GzLMMs) as implemented in the *glmer* function of the lme4 package in R, assuming Poisson distribution for ovule counts and binomial distribution for proportion of viable seeds *vs*. number of ovules, with region as fixed factor and population as a random factor. If the model indicated significant differences among regions, we applied Tukey HSD post- hoc tests with p-values < 0.05 to infer which regions differed, using the *multcomp* package. To test for

differences in ovule numbers between native *vs*. non-native ranges, we generated GzLMMs with range as fixed factor and region as a random factor. Data on manual selfing rates were grouped per individual plant and included the seeds produced after selfing and the seeds produced after outcrossing. We then calculated the self-compatibility index (SCI) (Lloyd and Schoen, 1992) for every region and each individual plant after the manual selfing treatment based on the formula below:

Formula 1: SCI = seed-set after selfing / seed-set after outcrossing

SCI ranges from 0 to 1, with 0.75 as the threshold for a plant to be considered self-compatible and 1 showing full self-compatibility (Lloyd and Schoen, 1992).

We also calculated an index of fertility adapted after Ramsey et al. (2003), by dividing the mean seed:ovule ratio (S:O) value of the between-region cross to the mean S:O value of the within-population within-region cross based on the formula below:

Formula 2: $FI = (S_{between}/O_{between}) / (S_{within}/O_{within})$

where $S_{between}$ and $O_{between}$ indicate the number of viable seeds and total ovules per capitulum for the between-region crosses, and S_{within} and O_{within} indicate the number of viable seeds and total ovules for the within-region crosses, respectively. Index values on or around 1 indicate no differences in fertility between within-and between-range crosses; values below 1 indicate lower fertility for crosses between regions relative to crosses within region, and values above 1 indicate higher fertility for between- than within-region crosses.

RESULTS

Ovule counts

We found differences in total ovule production among the six regions of the maternal genotype ($\chi^2(1) = 18.931$, P = 0.001; $AIC_{null\ model}$ =18969 *vs* $AIC_{full\ model}$ =18978) (Fig 12 A), but no distinction between native and non-native ranges ($\chi^2(1) = 0.0158$, P = 0.899, $AIC_{null\ model}$ =19957 *vs* $AIC_{full\ model}$ =19956) (Fig 12 B). Argentina produced significantly higher number of ovules incomparison to Turkey (post hoc test, p = 0.001), Chile (p = 0.001) and California (p = 0.01) (see Fig 12A).

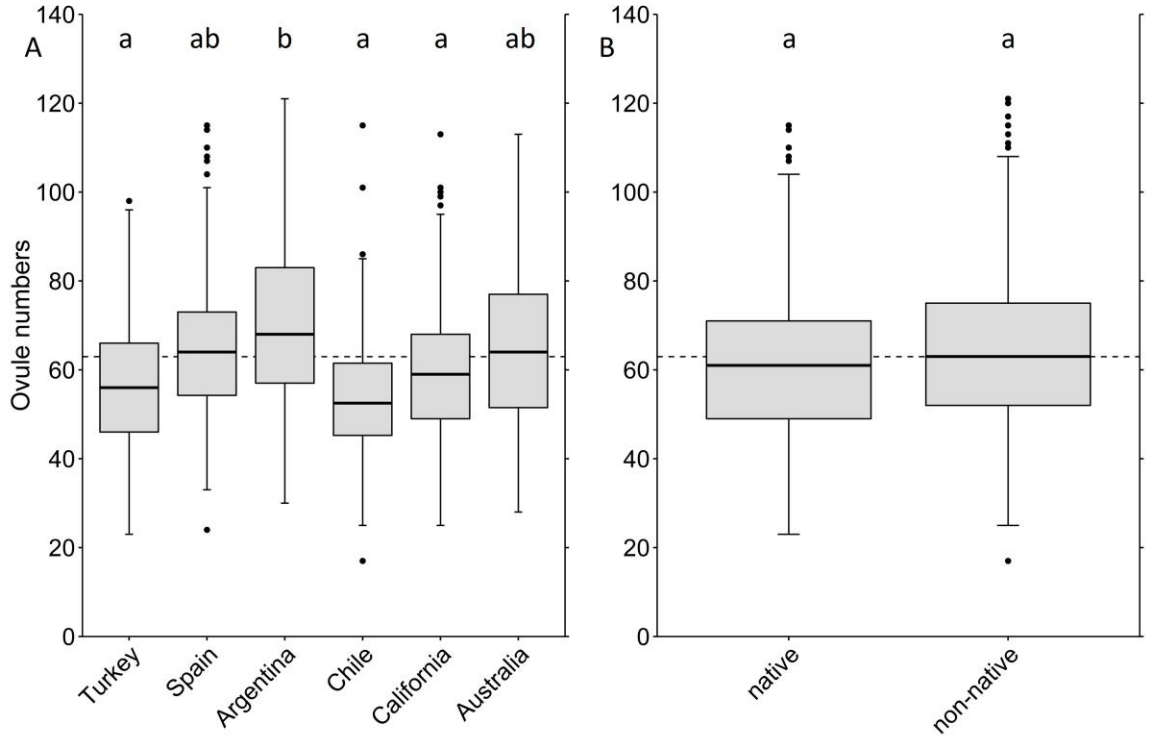


Figure 12. Ovule production per capitulum (Mean + SE) in *C. solstitialis*: A) comparison among the six regions; and B) comparison between the native *vs* non-natives ranges

Selfing rate

We counted total ovule numbers in 313 capitula from 169 individuals subjected to manual self-pollinations. Overall, the reproductive success for manual self-pollinations was 10%, with just 17 individuals (Turkey: 1, Spain: 4, Argentina: 6, California: 3, and Australia: 3) producing at least one viable seed (global mean of 1.94 + 0.34 SE viable seeds, with a maximum of 5 viable seeds per capitulum) after manual self-fertilization. The SCI index was consistently low across all six regions (under 0.01), indicating strong self-incompatibility. At the individual level, only four plants (2 from Spain and 2 from Australia) displayed full self-compatibility, with a SCI > 0.75. In the case of the negative control group (spontaneous self-fertilization), we analyzed 225 capitula from a total of 189 individuals and found only three capitula in three individuals from different geographic regions that set at least one viable seed (Turkey, California and Australia: 1 capitula with one viable seed each), with a reproductive success of 2%.

Fertility rates for crosses within region

We found that the interaction between treatment and range of the maternal genotype was significant and that it affected fertility rates ($\chi^2_{(1)}$ = 3.93, P = 0.047, AIC $_{\text{null model}}$ = 7291.7 vs AIC $_{\text{full model}}$ = 7293.6). In the non-native range, crosses within population displayed increased fertility compared to crosses between populations (P = 0.016). There were also differences in reproductive success among regions of the maternal genotype (Appendix 16). In Argentina and Chile, crosses between distant populations were disadvantageous relative to crosses within populations, resulting in 26% and 59% decrease in fertility, respectively. Conversely, crosses between distinct populations in Turkey and Australia showed 22% and 47% increase in fertility compared to crosses within population (Fig. 13).

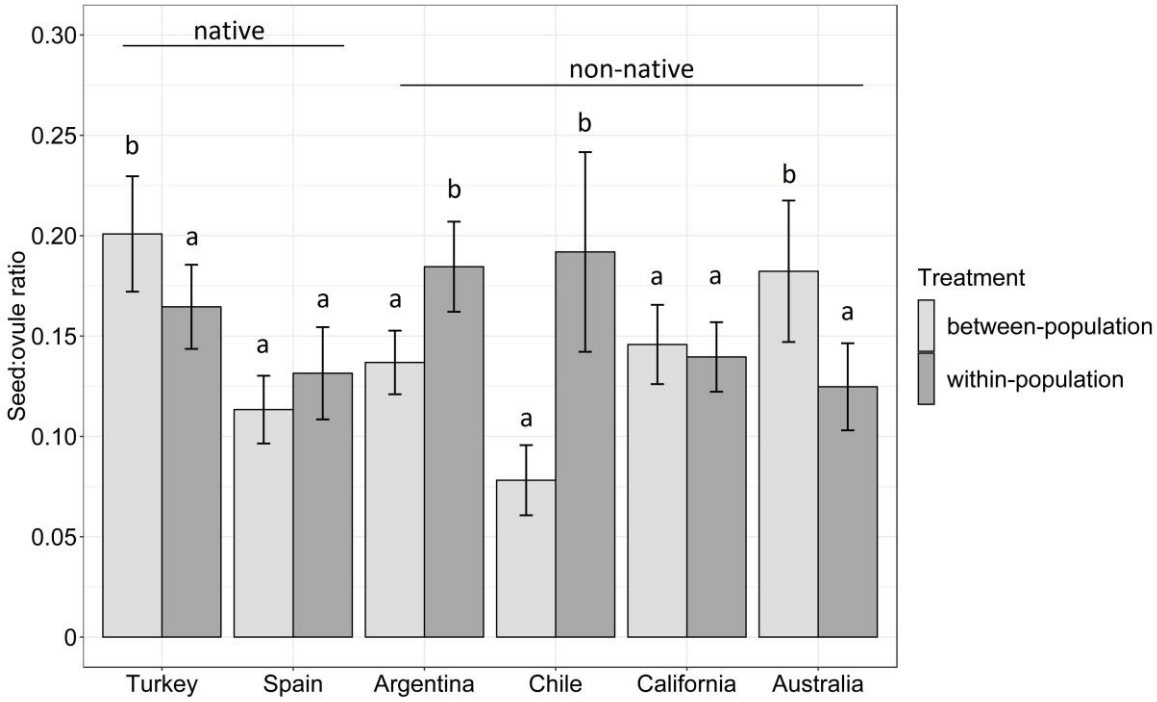


Figure 13. Seed to ovule ratio for treatments within regions (within population *vs* between populations). Seed to ovule ratio is given as the percentage of ovules that developed into viable seeds from the total number of ovules available in the capitula treated. Mean and standard error of the mean are also shown. Different letters indicate significant difference between the two treatments within each region at $p < 0.05$

Fertility rates for crosses between regions

Globally, we found differences in fertility rates for crosses between ranges ($\chi^2_{(1)}$ = 31.068, P < 0.001, AIC $_{\text{full model}}$ = 12666 *vs* AIC $_{\text{null model}}$ = 12691). Crosses involving native maternal plants and non-native paternal plants were more fertile than all other treatment combinations (i.e. native ♀ + native ♂; non-native ♀ + non-native ♂, and non-native ♀ + native ♂). Additionally, among the 30 region pair-wise cross combinations tested, 15 differed significantly in terms of fertility rates (Appendix 17). Spain demonstrated the highest number of positive fitness interactions with significant increase in fertility whenever pollen from any other non-native region was used (Table 8, Figs 14 and 15). In several non-native regions, fertility was significantly affected by pollen source, particularly in the Americas, but the direction of the interaction was idiosyncratic to each pair of regions (Table 8, Fig. 15). Australia was the only non-native region that exhibited positive fitness interactions when crossed to other non-native regions. The indices of fertility for crosses between regions ranged between 1.18 and 1.68 (increased fertility) and 0.51 and 0.85 (decreased fertility) (Table 8).

Table 8. Fertility indices indicating higher (above 1) or lower (below 1) fertility for between-region than for within-region crosses, for each pair of regions. Bold numbers indicate statistically significant differences (p < 0.05). Columns indicate the region of origin of the paternal genotypes, and lines indicate the origin of the maternal genotypes

♀ \ ♂	Turkey	Spain	Argentina	Chile	California	Australia
Turkey		0.98	1.26	**0.57**	1	0.96
Spain	1.16		**1.52**	**1.48**	**1.43**	**1.68**
Argentina	0.92	**0.58**		0.89	**0.66**	0.78
Chile	**0.82**	1	**0.85**		**0.69**	0.74
California	0.88	**0.77**	0.87	**0.51**		**1.18**
Australia	1.04	0.98	**1.43**	1.02	**1.29**	

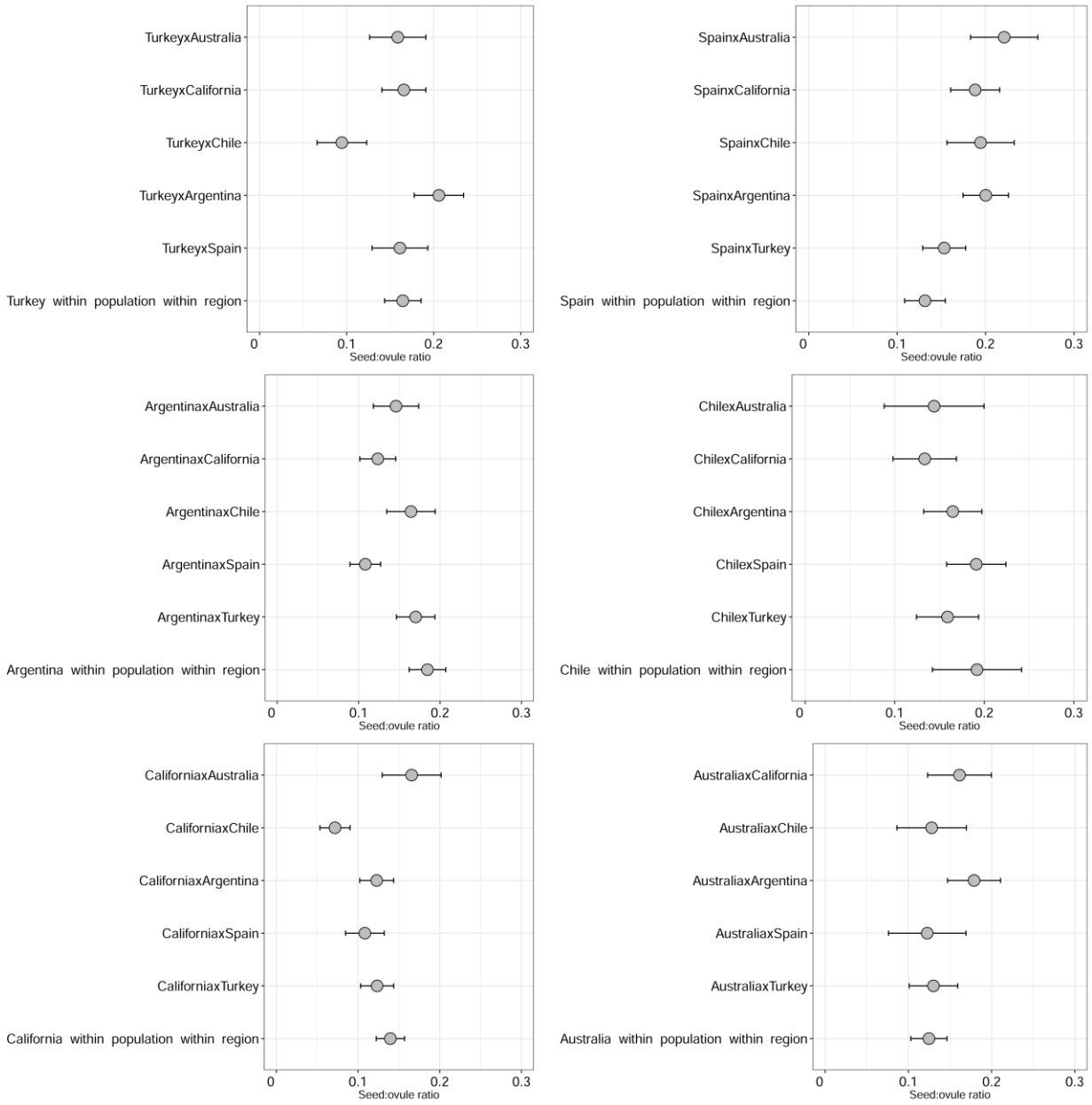


Figure 14. Seed to ovule ratio (S:O) for treatments between regions, and comparison to within populations within regions values (control), for each of the six maternal regions and their pair-wise combinations expressed as means and SEs of the means

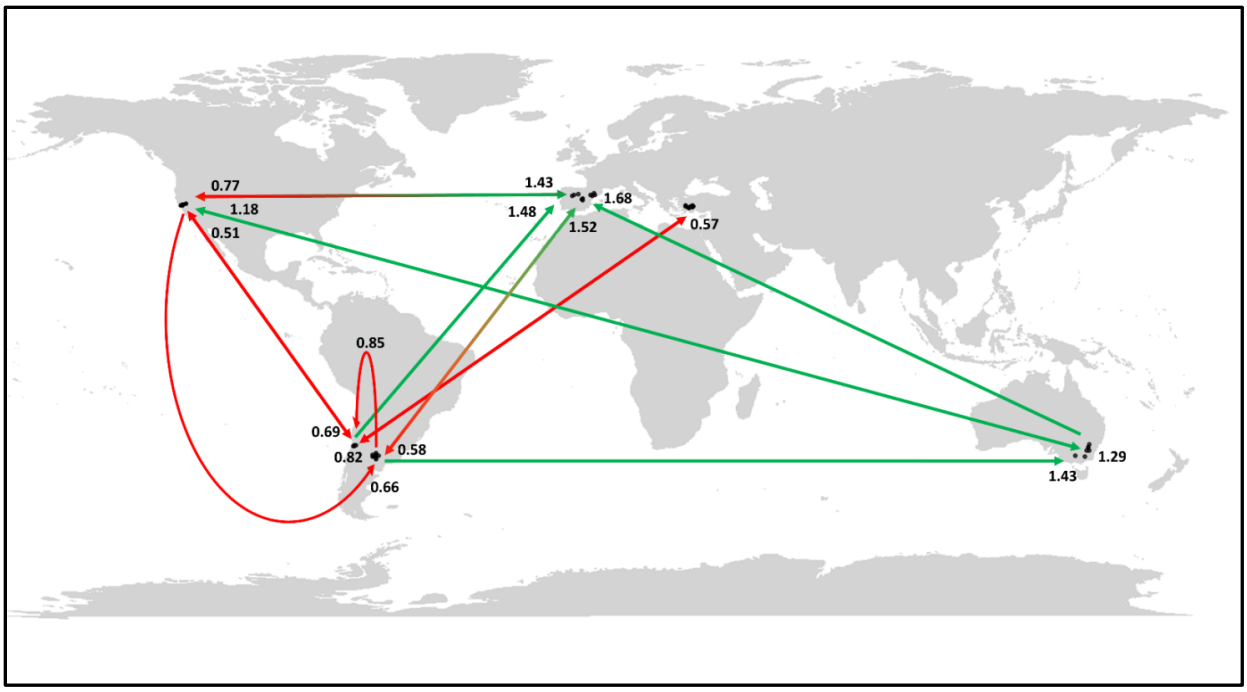


Figure 15. The global mosaic of fertility based on the fertility indices shown in Table 8. Fertility indices indicate higher (above 1) or lower (below 1) fertility for between-region than for within-region crosses, for each pair of regions. Red lines indicate a decrease in fertility and green lines show an increase in fertility. Lines with arrowheads at both ends indicate a bidirectional effect on fertility whereas gradient lines (half red half green) show that the cross generated negative effects in one direction and positive effects in the other direction. For clarity, only significantly different interactions are shown. Black dots indicate the seed sources of the *C. solstitialis* populations included in this study

DISCUSSION

Allopatry is an important process in ecology and evolution, and it is highly relevant to the study of invasive biology because invasive species have repeatedly shown to rapidly develop adaptations to newly colonized allopatric regions. Experimental crosses among *C. solstitialis* populations from six different native and non-native regions of the world demonstrated different reproductive success depending on the biogeographic origin of pollen donors and receivers. A previous study detected an incipient degree of reproductive isolation between Californian

maternal genotypes and Spanish paternal genotypes (Montesinos et al. 2012). The present study confirms that pattern with a much larger sample size and expands the methodology to encompass most known world regions where *C. solstitialis* occurs. In our experiment, viable seeds were obtained in both crossing directions and for all the outcross treatment combinations but exhibited a high degree of variation across regions, ranges and treatments, depending on the direction of the backcross. *Centaurea solstitialis* is an obligate outcrosser, as confirmed by our manual selfing treatments, indicating that the contribution of selfing to seed set is negligible and should not constitute a source of bias in interpreting the results of the outcross treatments conducted.

We found no differences in reproductive investment between native and non-native ranges in terms of total number of ovules produced. At the global level, crosses within and between the native and non-native ranges tended to yield similar fertility rates except for when crossing native maternal plants with non-native fathers which resulted in significant increase in fertility. However, this may have been driven by results from Spain alone, which exhibited large reproductive success when crossed with all other non-native regions. We found pronounced changes in fertility rates in 50% of our pair-wise controlled crosses between native and non-native *C. solstitialis* world regions, of which seven interactions were positive (heterosis: increased fertility) and eight were negative (decreased fertility for inter-regional seed-sets when compared to within region seed-sets). All but one negative fitness interactions (i.e. seven), occurred in the Americas, when maternal plants from these regions received native pollen or pollen from another non-native region in the Americas. This supports the idea that populations in the Americas are undergoing rapid evolutionary changes and starting to develop incipient reproductive barriers. In contrast, very strong heterotic effects (increases in seed-set) were observed for Spanish mother plants each time they were crossed with each non-native region.

According to evolutionary theory, reproductive isolation between populations of the same species are expected to arise when geographically isolated populations start to accumulate divergent local adaptations (Widmer et al. 2009). When these populations are brought back into contact, their potential to interbreed may be hampered by incipient reproductive barriers, and selection against maladapted individuals from a different population can favour the

reinforcement of that process. Admixture between these isolated populations may result in different outcomes depending on whether population differentiation is driven by genetic drift or by local adaptation. Under the first scenario, gene flow between populations is expected to be beneficial by masking deleterious alleles, increasing the standing genetic variation, and releasing the inbreeding load, thus, playing a significant role in the establishment phase of the invaders (Ellstrand and Schierenbeck, 2000). However, in a population that is locally adapted, admixture may be selected against because it could disrupt the co-adapted gene complexes leading to genetic incompatibilities (Verhoeven et al. 2011). Thus, selection against non-local genotypes may act as a barrier to gene flow and lead to a decline in the reproductive compatibility. Here, we observed mixed performance for the admixed crosses and also noticed that the identitiy of the maternal plant seemed to have played a major role in the cross outcome. There is compelling body of evidence that populations in the Americas have evolved adaptively which lead to the accumulation of different ecological fitness benefits compared to native genotypes (e.g. Eriksen et al. 2012, Graebner et al. 2012, García et al. 2013, Barker et al. 2016, Filipe et al. 2016, Irimia et al. 2019, Montesinos et al. 2019, Montesinos et al. 2018, Montesinos et al. 2017). Although we lack molecular data relative to the degree of divergence (e.g. F_{ST} index) for these specific populations we sampled in this study, another paper showed clear differentiation between populations in Asia (Turkey) and those in Western Europe (Spain), but instead found little divergence between Western Europe and populations in the two Americas (Barker et al. 2017). So, genetic divergence alone does not seem to offer a satisfactory explanation for all our results.

The strong negative fitness interactions detected in the non-native range of the Americas seem to suggest that intraspecific hybridization here may disrupt locally co-adapted allele combinations via negative epistasis (interaction of alleles at different loci (Barker et al. 2019). Historical records indicate that *C. solstitialis* was introduced to South America from Spain first into Chile around 1600 and from there to California at about 1850 (Gerlach 1997, Pitcairn et al. 2006) from a single genepool (Barker et al. 2016) and ultimately to Argentina approximately 1870 (Hijano and Basigalup 1995). Perhaps similarity in introduction times could contribute to the extent of the fitness effects observed here. So far, only cross-fertility barriers were reported in *C. solstitialis* (Montesinos et al. 2012) with no indication of negative pistil-pollen interactions or

structural modification in floral morphology that could influence seed output (Montesinos, personal communication). The fertility rates for crosses between regions for American populations were frequently asymmetric. Argentina, California and Chile experienced moderate fertility reduction when treated with native pollen from Spain (the first two regions) and Turkey (the latter). Some of these results are consistent with the findings of Montesinos et al. 2012, who detected a reduction in fertility for Californian mothers treated with Spanish pollen, although in our case the percentage decrease was half of the value reported there. The dissimilarity might be attributed to a methodological difference as here we analyzed the seed:ovule ratio, whereas in the original experiment only the number of viable seeds was measured. However, populations here demonstrated among the strongest fertility reductions compared to control when treated with other non-native populations in the America. Both Argentina and Chile showed reduced fertility when treated with Californian pollen and same was true for Californian mothers receiving Chilean pollen. According to Barker et al. 2017, California and Chile belong to different genetic clusters, so the genetic distance may be a potential explanation for the reciprocal incompatibilities observed between these two regions. Additionally, a weak but significant negative interaction was also detected between populations in South America when Chilean mothers received pollen from Argentina. *Centaurea solstitialis* populations in these two regions appear to be part of the same genetic cluster (Barker et al. 2017), but more sequencing effort of individuals from Argentina is required to tell whether there is any fine population sub-structure in South-America where the Andes may act as a natural barrier to gene flow. On the whole, these results seem to suggest that populations in the America were more prone to develop negative fitness interaction than any other region when backcrossed to native populations as well as with other non-native populations. In contrast, we found only one negative interaction in the native area – the cross between Turkish maternal genotypes and Chilean paternal genotypes. These two regions are characterized by low levels of genetic divergence between them, based on structure analysis of hundreds of SNPs (Barker et al. 2017) which suggests that other mechanisms should be responsible for the reciprocal decrease in fertility. For instance, the effects of admixture on the first generation of hybrids between different source populations of *C. solstitialis* from the native area concluded that both the

identity of the maternal parent and the genetic diversity present within the maternal population were important for the performance of the hybrid (Barker et al. 2019). They hypothesized that some of the negative fitness effects might have been due to cyto-nuclear incompatibilities, as a consequence of a rapidly evolving plastid genome (Barker et al. 2019).

The large increase in reproductive success when Spanish populations served as the maternal parent in crosses with all other non-native regions might be explained by the fact that populations in this region where the primary contributors of seeds that colonized the Americas, potentially serving as a "bridgehead" (Barker et al. 2017, Barker et al. 2019). The invasive bridgehead effect refers to a particular population that serves as the proximate origin of successful introductions into several other areas (Lombaert et al. 2010). Barker et al. (2019) also found very strong positive fitness interactions in F1 progenies derived from maternal genotypes from Western Europe plants, as was the case in this study. *Centaurea solstitialis* populations in Western Europe are the result of a recent admixture event between Asian and eastern European populations (Barker et al 2017), but experimental crosses in this region seem to provide additional fitness benefits. Among non-natives, California and Australia were the only two regions involved in positive fitness interactions, with a bidirectional seed increase between California and Australia and unidirectional seed increase between Australia (♀) and Argentina (♂). The population structure and colonization history of *C. solstitialis* in Australia, which is believed to have been introduced there around 1892 (Kloot 1983), has not been yet resolved with molecular data. Our results suggest that populations in the Americas might have been contributing sources to the colonization of Australia.

CONCLUSIONS

Our experiment revealed a biogeographic mosaic of fertility among half of the pair-wise experimental crosses between allopatric regions of *C. solstitialis*, showing evidence of both heterosis and reproductive incompatibilities. Fitness interactions appeared to be asymmetrical and mostly unidirectional although some crosses generated bidirectional interactions too, with a fertility index exhibiting large variation. These results expand current evidence of rapid evolution in this invasive weed after relatively short time since introduction in non-native areas.

Additionally, our results also support the idea that, at this stage of invasion, potential admixture in some *C. solstitialis* non-native ranges could have detrimental effects on local populations, but that in some others (Australia) could result in increased fertility.

Chapter IV - Reproductive and defense traits facilitate rapid invasion in a weed species

Chapter section in preparation for submission as an original research article to SCI journal by Irimia et. al

ABSTRACT

Invasive species often possess a great capacity to evolve rapidly, and to adapt to local conditions post introduction, while exhibiting a large amount of spatial trait variation. Phenotypic differentiation in multiple traits can be due to several factors, such as varying selection, genetic drift, or plastic responses to the environment. We explored the factors driving geographic differentiation of several phenotypic traits in the highly invasive weed *Centaurea solstitialis* by measuring neutral genetic differentiation (F_{ST}) at 1975 genome-wide single nucleotide polymorphisms (SNPs) and comparing it with phenotypic differentiation (P_{ST}) in a common garden experiment in individuals across two native (Turkey and Spain) and four introduced regions (Argentina, Chile, USA, and Australia). Adaptation due to divergent selection is likely to have occurred where P_{ST} is significantly greater than F_{ST}. The genetic data was also screened for SNPs showing significantly elevated F_{ST} relative to the genomic background as these markers might be linked to genes under selection. We found clear genetic structure between native and non-native areas, with further evidence of divergence between Argentina and Chile in South America. Native plants tended to produce more flower heads than non-natives, whereas non-natives produced larger seeds of the plumed morph. This observation was also reflected in the global $P_{ST} – F_{ST}$ comparison, which demonstrated that seed mass is putatively an adaptive trait. In contrast, traits linked to growth (i.e. bolting day) demonstrated lower P_{ST} values compared to the mean F_{ST}, suggesting stabilizing selection. P_{ST} values for days to flower and spine length exceeded neutral F_{ST} in the pair-wise comparison between non-native California and Australia, indicating divergent selection between these regions. Gene ontology analysis of 19 highly divergent SNPs mapped to 17 genes revealed their functional involvement in gene modulation and regulation, suggesting a possible role of epigenetics in shaping adaptive differentiation between native and invasive genotypes. Our study indicated that rapid adaptive evolution has contributed to the success of *C. solstitialis* and provides new insights into traits that can contribute to fitness advantage in non-native populations.

Keywords: invasive plants, yellow-starthistle, single nucleotide polymorphisms, neutral genetic variation, biogeography, P_{ST}-F_{ST}

INTRODUCTION

Different populations of a single species occur in a range of different environments and can therefore experience different local selection pressures which in turn can lead to intraspecific variation and local adaptation (Blanquart et al. 2013). Many factors, such as gene flow, genetic drift, mutations, and standing genetic variation, influence the extent and rate of population differentiation and local adaptation (North et al. 2011; Lai et al. 2019). For introduced invasive species, detecting quantitative trait differentiation requires comparing the fitness performance of different genotypes in their novel environments relative to that of their native environments ("home vs away" comparison; Pigliucci, 2001; Blanquart et al. 2013).

A powerful approach to identify potentially adapted traits, relative to divergence caused by other neutral processes, is to compare the genetically-determined quantitative trait differentiation among populations (termed $Q_{ST,}$ Spitze 1993; Leinonen et al. 2013) with neutral genetic differentiation (F_{ST}) (Wright 1951) in common garden experiments. However, estimation of Q_{ST} requires prior knowledge about relatedness between sample individuals to account for within and among-population additive genetic variance, usually achieved by studying the progeny of controlled crosses, which for some species might not be feasible. An alternative to Q_{ST} is P_{ST}, when species are not suitable for breeding designs that allow the estimation of additive genetic variance. P_{ST} uses purely phenotypic data to compare the influence of genetic adaptation, phenotypic plasticity and genetic drift as causes of population differentiation (Leinonen et al. 2006; Leinonen et al. 2013). Under the neutral expectation, P_{ST} equals F_{ST}, with any significant deviation indicative of an additional influence of selection on P_{ST}. If $P_{ST} > F_{ST}$, divergent selection favors different phenotypes in different populations, while if $P_{ST} < F_{ST}$, stabilizing selection favors the same phenotypes in different populations. A disadvantage of P_{ST} compared to Q_{ST} is that it does not isolate additive genetic variance from environmental variance as effectively so that P_{ST} estimates tend to be larger than Q_{ST} estimates (Pujol et al. 2008). Nonetheless, P_{ST} can still prove informative as an exploratory tool for assessing selection, provided that highly conservative approaches are taken in interpreting the results of the P_{ST}-F_{ST} comparison (see Brommer 2011; O´Hara and Merilä 2005; Leinonen et al. 2013).

Invasive species represent interesting study systems to explore different evolutionary

processes in a contemporary time frame. When exotic species are introduced into new areas, they need to cope with novel biotic and abiotic pressures and adapt quickly in order to survive and establish (Bossdorf et al. 2005). Their adaptive potential depends on population genetic characteristics and phenotypic plasticity responses in climatically diverse areas and successful invasion can lead to the creation of a mosaic of geographically adapted populations (Matesanz et al. 2012). Reproductive isolation may also begin to evolve between allopatrically isolated populations as a byproduct of this genotypic and phenotypic differentiation (see Lowry et al 2008; Montesinos et al. 2012). Many hypothesis have been proposed to explain the factors contributing to biological invasions (Jeschke et al. 2012) but little research has been done to investigate the molecular mechanisms underlying rapid adaptation of invasive species. In the work that has been done in this area, studies have shown that genes associated with photosynthesis, energy metabolism and stress response to pathogens are likely to be up-regulated in invasive species and contribute to higher adaptive potential (Prentis et al. 2010; Guo et al. 2018).

Centaurea solstitialis is a widely distributed invader, native to Eurasia and highly invasive in the South and North Americas where it has been relatively recently introduced in the last 400 and 200 years, respectively (DiTomaso et al. 2006). Genetic analysis suggest that *C. solstitialis* populations in California and Chile originate from a single introduction event derived from Western Europe (Spain) and that high levels of genetic diversity in California could be due to increased connectivity and population size during range expansion as well as post-introduction evolution (Barker et al. 2017). This weed has spread across a wide range of elevations in California from coastal areas to the interior part of the state and is currently expanding its range into mid and high elevations of the Sierra Nevada Mountains (Pitcairn et al. 2006; Swope and Parker 2010). Additionally, its worldwide distribution spans contrasting climates such as Mediterranean climates with summer drought and winter rain in California and Chile and continental climates with predominately summer rain in Argentina (Hierro et al. 2009). These differences in climatic regime could impose divergent selection pressures and lead to the rapid evolution of variation in phenotypic and physiological characters. Indeed, evidence of rapid adaptation in the introduced areas has been documented in *C. solstitialis* (Hierro et al. 2009;

Graebner et al. 2012; García et al. 2013; Eriksen et al. 2012; Dlugosch et al. 2015; Barker et al. 2017; Montesinos and Callaway 2018)..A Q_{ST}-F_{ST} comparison of *C. solstitialis* populations from Republic of Georgia, Turkey, Argentina and California based on seven EST-SSR loci found evidence of selection for increased size related traits (leaf length and plant height) in non-native areas (Eriksen et al. 2012). Similarly, other studies found support for acceleration of growth, earlier flowering time and increased reproduction in invasive populations of *C. solstitialis* from California compared to native populations (Dlugosch et al. 2015), as well as increased allocation for spine growth (García et al. 2013) and even incipient reproductive incompatibilities (Montesinos et al. 2012). Meanwhile, a reciprocal home and away common garden experiment showed twofold increase in seed mass for invasive Argentinean populations compared to native Turkish populations (Hierro et al 2013). Plant size, density and final establishment were also greater for plants from Argentina grown in the common garden in Argentina, whereas no differences related to source population were observed in the common garden in Turkey (Hierro et al 2013). These results support that invasive Argentinian *C. solstitialis* exhibits local adaptation rather than phenotypic plasticity (see also Montesinos and Callaway 2018).

In this study, we search for genetic signals of divergent selection in native and non-native populations of *C. solstitialis* by measuring F_{ST} across 1975 single nucleotide polymorphisms (SNPs) and comparing it with phenotypic differentiation (P_{ST}) for multiple ecologically relevant traits (plant size, reproductive success and physical defense). We conducted a common glasshouse experiment with several population accessions from Turkey and Spain (native area) and from Argentina, Chile, California and Australia (non-native regions) to quantify trait variation among native and non-native genotypes. We used double digest restriction site associated sequencing (ddRADseq) to generate genome wide SNP marker genotypes to identify *C. solstitialis* population structure, assess genetic diversity across regions and to measure F_{ST} for comparison with P_{ST}. Notably, we generated molecular data for *C. solstitialis* in Australia, an invaded region never sampled before in genomic studies and we further improved population genetic resolution in South America. We also tested for signals of selection across SNP loci and investigated the genetic basis of adaptation by identifying outlier SNPs within genes based on the *C. solstitialis*

assembled transcriptome and conducting functional annotation of these genes according to *Arabidopsis thaliana* gene ontology annotation.

MATERIALS AND METHODS

Study species and sample collection

Centaurea solstitialis is a highly invasive diploid annual plant in the Asteraceae. The phylogeographic history of *C. solstitialis* is well understood and is characterized by a series of human mediated colonization and admixture events (Erikson et al. 2014; Barker et al. 2017). The geographic origin of the species is considered to be Turkey (Uygur et al. 2004, Garcia-Jacas et al. 2006) from where the species went through a slow stepwise range expansion into Western Europe, with populations from Spain serving as the main source of native seeds for introductions in the Americas (Barker et al. 2017). For this study a total of 46 *C. solstitialis* populations were sampled in the wild between 2009 and 2014. Ten populations were from Turkey, nine from Spain, ten from Argentina, four from Chile, seven from California and six from Australia (Appendix 18). Mature seeds were collected from up to ten different maternal individuals in each population. Populations spanned sites in the native area represented by Asia (eastern Turkey) and Western Europe (Spain), as well as introductions in South America (Argentina and Chile) and North America (California) and Australia.

Phenotypic traits

Seeds were germinated in a glasshouse at the Botanical Garden of the University of Coimbra in 50 cell plug trays containing commercial soil (Substratos Profissionais Leal and Soares S.A., Portugal) and seedlings were transplanted into 2L square plastic pots (1 plant per pot) filled with the same type of potting soil about three weeks after germination. Plants were kept in the glasshouse, where they experienced natural sunlight and Mediterranean climate in Coimbra and were grown to maturity and senescence (March to November 2017). We scored seven morphological traits on each individual that reached reproductive stage and survived through senescence (N = 330) including (1) days to bolting; (2) days from bolting to flower, (3) length of the largest spine; (4) adult plant height; (5) number of inflorescences; (6) ovule number and

(7) mass of the plumed seeds (fibrous outgrowths). Plant height and spine length were measured to the nearest mm using a flexible ruler. A plant was considered to have initiated bolting when a flowering stem of ~ 5 cm tall started to extend from the basal rosette. At the end of the reproduction period, we counted all the inflorescence and lateral floral buds on each individual. Ovule counts in a capitulum included the number of aborted and viable seeds. Seeds were stored in paper bags at room temperature for 6 months, before weighing on a Kern ALJ analytical balance (Balingen, Germany) to the nearest mg.

ddRADSeq library preparation

We selected 194 individuals sampled across the six regions for reduced representation sequencing (ddRADSeq). Genomic DNA was isolated from silica gel dried leaves sampled from 8 weeks old individual plants grown under common glasshouse conditions, using the CTAB protocol (Doyle and Doyle 1987). Leaf tissue was ground to a fine powder with the Tissue Lyser II (Qiagen) for 2 minutes at 25Hz. Following extraction, DNA quality and quantity was assessed on a Nanodrop 2.0 spectrophotometer (Thermo Fisher) and a Qubit 2.0 fluorometer (Invitrogen Life Technologies) using the QuantiFlour dsDNA sample kit (Promega, Madison, USA). A starting quantity of 500 ng of purified DNA from each individual was digested with the restriction enzymes, Pst1 (recognized sequence and cut site: *CTGCA^G*), and Mse1 (*T^TAA*), followed by the ligation of unique paired combinations of individual P1 and P2 barcoded adapters (kindly supplied by K. Dlugosch, Arizona University, see Appendix 19 for adapter and barcode sequences), following the protocol by Barker et al. 2017. The barcoded individuals were pooled in a single library that was purified using SpeedBead Magnetic Carboxylate modified particles (GE Healthcare UK) and size-selected for DNA fragments between 300 and 550 bp on a Pippin Prep automated size selection system (Sage Science, Beverly, MA, USA) using a 2% agarose gel cassette (dye free 100-600bp, DNA size range collection with external marker L; Sage Science). The size selected DNA library was amplified by 11 PCR cycles using Phusion High-Fidelity DNA polymerase (New England BioLabs, Massachusetts) on a Prime Thermal Cycler (Midwest Scientific, St. Louis) (program: initial denaturation: 2 min – 98°C, followed by 11 cycles of denaturation: 98°C – 10s, annealing: 65°C – 30s, extension: 30s – 72°C and a final extension of 10 min – 72°C), to increase

the concentration of properly ligated DNA fragments, and the resulting product was purified with SpeedBead modified particles and then eluted in 10 uM TrisHCl-EDTA solution. Library fragment size distribution was visualized on an Agilent 2200 TapeStation System by testing 1ul of purified DNA library using the D1000 high sensitivity ScreenTape and D1000 Sample buffer (Agilent Technologies). Quantitative PCR (qPCR) was done using the NEBNext Library quantification kit for Illumina (New England Biolabs, Massachusetts) to measure the concentration of the pooled library on a BioRad CFX connect real time system (BioRad, California). Libraries were sequenced on five separate lanes of an Illumina HiSeq 2500 (Illumina, Inc, San Diego, CA, USA) at the University of Durham Genomics Sequencing and Analysis Facility, UK, to generate 125 bp paired-end reads. In total, we obtained ~ 1.26 billion pair-end Illumina reads across 194 individuals. Accessions and run information are available in NCBI (data will be submitted by publication).

Data processing

Morphological traits and phenotypic divergence index

Data were checked for heteroscedasticity and normality by Levene's and Shapiro–Wilk tests (Levene 1960; Shapiro and Wilk 1965), and log transformed when these assumptions were not met. To test the differentiation of phenotypic traits between regions we generated generalized linear mixed-effects models in R 3.5.2 (R Development Core Team 2014) by employing the lme4 package (Bates et al. 2015), with region as fixed factor and population as a random factor. If the models indicated significant differences among regions, we applied Tukey HSD post hoc tests with P-values < 0.05 to infer which regions differed and graphically summarized the data using boxplots. We performed a principal component analysis (standardized and centered PCA) to visualize trait differences between native and non-native ranges of *C. solstitialis* by employing the prcomp function to compute principal component scores and plotted the PCA with the ggbiplot 0.55 package (Vu 2011). Missing data (less than 15% for ovule number and plumed seed mass and less than 5% for the remaining traits) were conservatively replaced with mean values for that particular trait. We used the psych and stats packages in R (Revelle 2018) to test for character correlations, generate pairwise scatter plots

and compute Pearson correlation coefficients. To compare the level of phenotypic differentiation with the level of genotypic differentiation at neutral markers we calculated P_{ST}, an index which assesses for differentiation among populations in phenotypic traits. P_{ST} for each trait and each paired region combination was calculated using a Bayesian approach following Leinonen et al. 2006 by fitting a linear model with population as a random effect and the trait of interest as a covariate using a Gibbs sampler implemented in the software WinBUGS 1.4.3 (Lunn et al. 2000) and specifying the corresponding mean value of F_{ST}. All phenotypic traits met normality assumptions except number of inflorescences, which was log transformed before fitting it into the model. Posterior distributions were obtained by running five independent chains (50 000 iterations) after a burn in of 1000 iterations. Bayesian credibility intervals were estimated for both P_{ST} and $P_{ST} - F_{ST}$ difference. If the credibility interval of the $P_{ST} - F_{ST}$ difference was higher or lower than zero, we regarded it as an indication that the trait tested is putatively under divergent or convergent selection, respectively. Alternatively, when the $P_{ST} - F_{ST}$ credibility interval overlaps zero, the observed degree of differentiation at quantitative traits could be the outcome of genetic drift solely.

De novo SNP discovery

Sequencing data was demultiplexed or assigned to each sample individual according to their unique paired barcode sequences and reads were quality filtered using Stacks 2.2 (Catchen et al. 2011; Catchen et al. 2013) to remove reads containing adapter sequence and low-quality base scores (a phred score below 10). The average number of reads per individual after demultiplexing and filtering was 6.5M (range 300K - 44M). All reads were end trimmed to 115 bp to ensure that they were of the same length before running it through the denovo_map pipeline. We used only the complete paired reads in the mate files 1.fq and 2.fq to perform the alignment in denovo_map.pl with the following set of parameters: a minimum coverage depth of five to create a stack (−*m* = 5), a maximum mismatch distance of two nucleotides between loci when processing a single individual (−*M* = 2), a maximum of two stacks at a single locus (−*X* = 2), and a mismatch distance of two nucleotides between loci (−*n* = 2), to account for the possibility of fixed differences at loci in individuals when creating the catalog of loci, according to Barker et.

et al. 2017. The *C. solstitialis* individuals were grouped into six geographic regions by supplying the population map into the denovo pipeline. We obtained 526.533 variant sites (unfiltered) across 194 individuals. Fifty individuals had sequenced very poorly with less than 500,000 reads each, so we decided to remove them and keep 144 individuals (24 per region x 6 regions) for all the subsequent analysis. VCFtools were used to filter the genotype data for the highest quality genotype calls by filtering out indels, including only bi-allelic sites (--min-alleles 2 --max-alleles 2), including only a minimum genotyping proportion per population of 0.9, and retaining only sites with a minor allele frequency greater than 0.05 in the whole dataset. These filtering steps retained high quality genotypes for 2138 variant sites (SNPs) across 144 individuals.

Genomic signatures of selection

We used three methods to conduct global outlier analysis across regions to identify variant sites that reliably showed signals of selection across multiple methods. First, we used BAYESCAN v.2.1 (Foll and Gaggiotti, 2008) to run a logistic regression model that decomposes F_{ST} into a locus- and population-specific component. The method allows for different migration rates and effective population sizes among subpopulations thus relaxing the conditions of the island model applied by other selection detection methods. We conducted 10 pilot runs each consisting of 5,000 iterations, with a burn-in of 50,000 iterations, a thinning interval of 10 and prior odds for neutral model set to 10, for a total of 3 replicates runs. Bayescan detected 11 putative outlier loci (q-value < 0.05). Secondly, we used OutFlank v0.2 package in R (Whitlock and Lotterhos 2015) to calculate a theoretical distribution of F_{ST} by trimming out loci in the upper and lower tails of the distribution and using the remaining loci to infer the distribution of F_{ST} for neutral markers by assigning q-values to each locus to detect outlier loci as candidate for local adaptation. OutFlank retrieved 36 putative outlier loci with a q-value < 0.05. Lastly, we used the PCAadapt package 4.0.3 in R (Luu et al. 2017) to compute the covariance matrix on the centered and scaled genotype matrix (linear combinations of allele counts). PCAadapt assumes that markers related to population structure are candidates for local adaptation. Test statistics and p-values were computed based on correlations between SNPs and the first 10 K principal components by implementing a robust Mahalanobis distance. We used the qvalue package

(Storey et al. 2019) in R to transform p-values into q-values and retrieved 138 SNPs with q-values < 0.05 as potential outlier candidates. We compiled a list of 163 putative outlier loci detected by at least one of these methods out of 2138 total variant sites.

Neutral genetic structure

Genetic diversity of each region was calculated at 1975 neutral loci as H_o (observed heterozygosity), H_e (expected heterozygosity) and Ar (allele richness) using the function divBasic in the R package diveRsity (Keenan et al. 2013). Global and pairwise population differentiation and its 95% CI was calculated using the standardized allelic variance F_{ST} with the diveRsity package (Keenan et al. 2013). Effective population size (N_e) was estimated using NeEstimator V2.1 (Do et al. 2014) using the linkage disequilibrium method assuming random mating and excluding sites with minor allele frequencies per population less than 0.05. Finally, we assessed neutral population structure at 1975 loci, in STRUCTURE 2.3.4 (Pritchard et al. 2000) by implementing a model of correlated allele frequencies (Falush et al. 2003) and admixture, and applying the default setting for all other parameters. Ten independent runs for all values of K (number of genetic clusters) between 1 and 8 were run using an MCMC length of 1000000 generations following a burn-in of 100000 generations. STRUCTURE HARVESTER (Earl and von Holdt 2012) program was used to carry out downstream processing of STRUCTURE results to calculate Evanno's Δ*k* value (Evanno et al. 2005) and prepare an input file for CLUMPAK (Kopelman et al. 2015) to generate bar graphs of population structure. Population structure at 1975 neutral loci was also visualized using a discriminant analysis of principal components (DAPC) (Jombart et al. 2010) as implemented in the R packages, ADEGENET 2.1.1 (Jombart 2008) and ADE4 (Dray and Dufour 2007). We performed an initial DAPC cross validation on the 144 PCs using the xvalDapc() function in adegenet with default parameters and 30 replicates to get an idea of where to focus more intense cross-validation runs. There was a peak around 40 PC (Root Mean Square Error = 0.09) so the search was narrowed by specifying around 40 PC, and doing 100 replicates each.

Gene ontology

The functional annotation of outlier SNPs based on gene ontology were performed in two

steps: (i) RAD-tag sequences harboring outlier SNPs were aligned to the assembled transcriptome of *C. solstitialis* (Dlugosch et al. 2013; Hodgins et al. 2015) using BLAST tool (Madden 2002) to determine whether such loci are localized within genes. (ii) Transcript sequences of the gene hits were functionally annotated with InterProScan (Jones et al. 2014) that performed BLASTX searches against the NCBI non-redundant protein database and *Arabidopsis thaliana* protein sequences on TAIR (https://www.arabidopsis.org/). The top protein hit for each transcript and associated information about gene ontology was retrieved, including information on biological process, molecular function, cellular component, growth and developmental stage and plant structure.

RESULTS

Trait differentiation

Globally, six traits out of seven (except days to bolting) showed significant differences among regions: (1) days from bolting to flower: ($AIC_{difference\ between\ models} = 7.5$); (2) number of inflorescences: ($AIC_{difference\ between\ models} = 6.5$); (3) length of the largest spine: ($AIC_{difference\ between\ models} = 7.05$); (4) ovules number : ($AIC_{difference\ between\ models} = 3.4$); (5) mass of plumed seeds: ($AIC_{difference\ between\ models} = 9.76$); and (6) final plant height: ($AIC_{difference\ between\ models} = 2.8$) (Fig 16). Detailed global statistics are provided in Appendix 20. Plants from Chile and California flowered 7.5 and 10.5 days earlier than those from Spain and Australia ($p < 0.05$), respectively, whereas plants from Australia took on average 8 days longer to initiate flowering compared to plants from Turkey and Argentina ($p < 0.001$) (Fig 16, Appendix 21). *Centaurea solstitialis* plants from the two native regions, Turkey and Spain, produced 37% more inflorescences per plant compared to Argentina, Chile and Australia *(*$p < 0.05$) (Fig 16, Appendix 21). By contrast, plants from Turkey produced 24% less ovules and 35% lighter seeds than plants in Argentina, California and Australia ($p < 0.05$). Similarly, plants in Spain and Chile also demonstrated a reduction of 31% and 43% in plumed seed mass compared to Argentina and California ($p < 0.05$). Australia presented spines that were 25% shorter compared to plants from all other regions except Spain ($p < 0.05$). Lastly, Chilean plants were 23% shorter than Turkish, Spanish and Argentinean plants ($p < 0.05$) (Fig 16, Appendix 21).

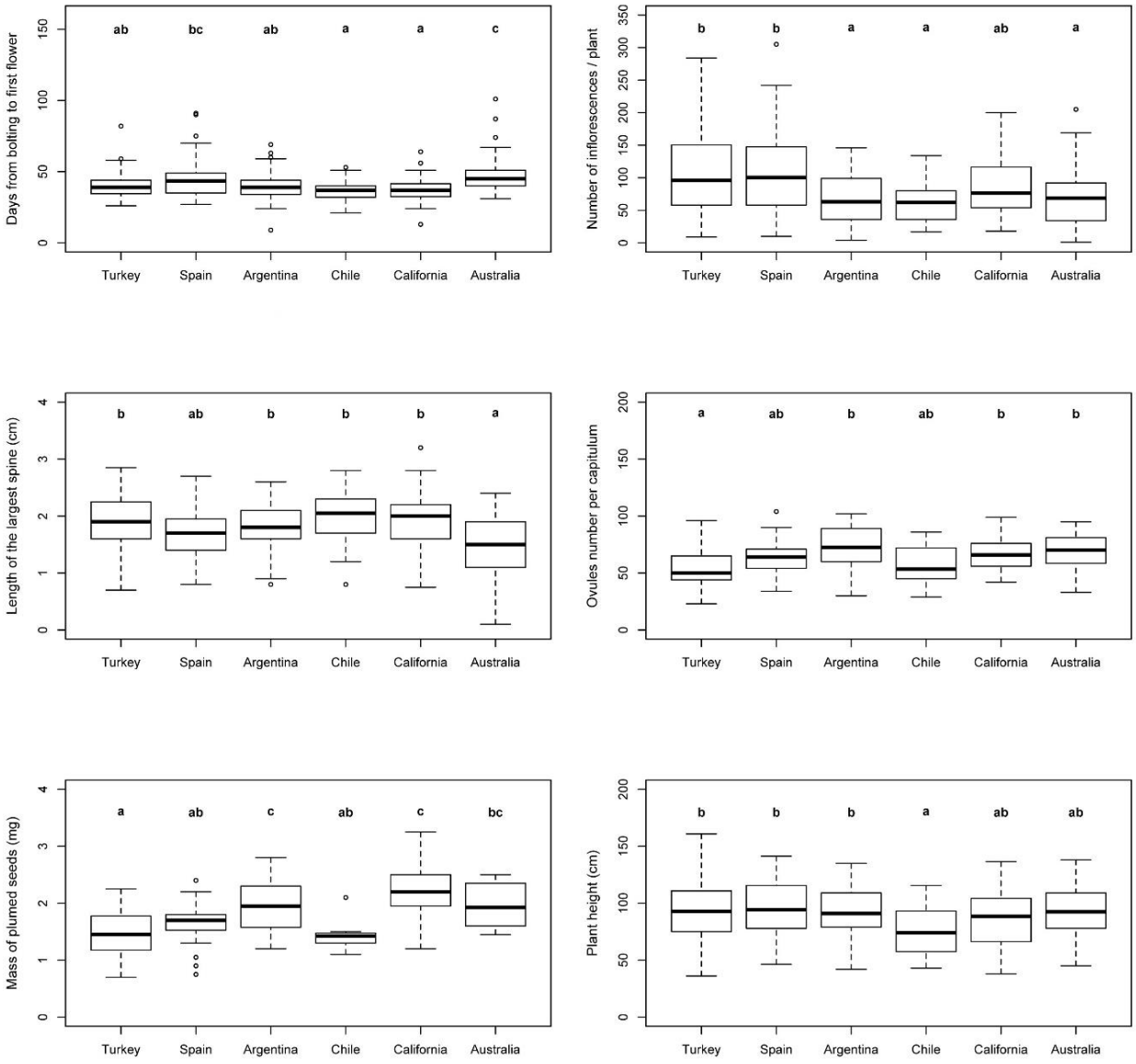


Figure 16. Boxplot showing summary data for phenotypic traits (mean ± SE). Turkey and Spain are the native regions and Argentina, Chile, California and Australia, the non-native regions

Principal component analysis of native vs. non-natives ranges

Principal component analysis revealed a tendency towards distinct clustering of *C. solstitialis* trait variation between the native and introduced ranges, particularly in terms of the second principle component (Fig. 17). Native plants were characterized by higher number of

of inflorescences whereas non-native plants distinguished themselves by higher numbers of ovules and heavier plumed seeds. PCA_1 and PCA_2 together explained 46.8% of the inertia variance in the two axes.

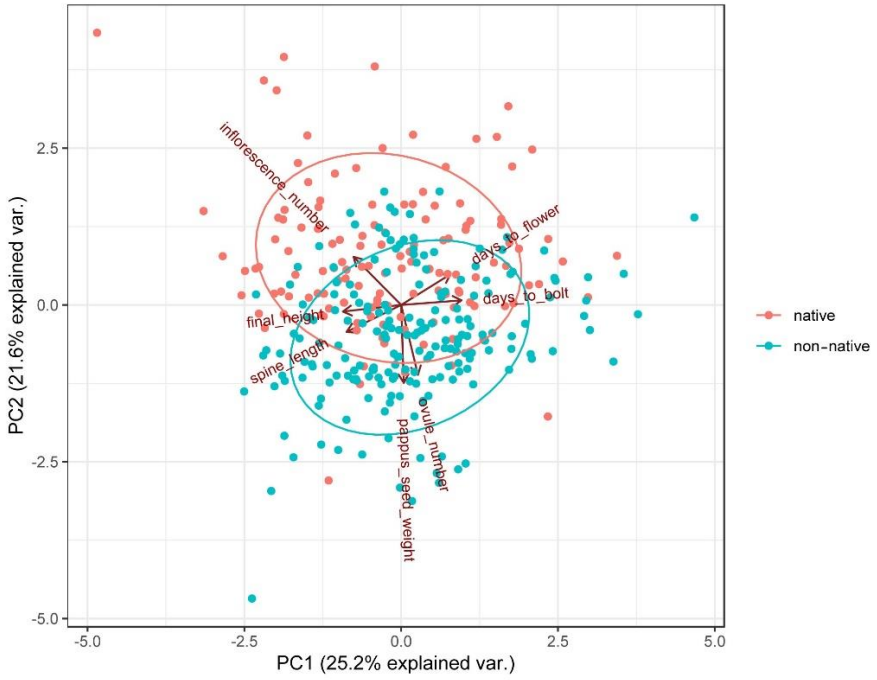


Figure 17. Principal component analysis of the phenotypic traits measured in *C. solstitialis*. A comparison between native and non-native regions

We found a moderate positive correlation between ovule number and mass of plumed seeds (r = 0.367, p = 0.001) and plant height and number of inflorescence (r = 0.33, p = 0.001). The remaining traits did not showed significant correlations (see Appendix 22).

PST divergence index

P_{ST} exceeded neutral F_{ST} for one trait associated with reproductive success, namely mass of the plumed seeds. Traits related to growth such as day to bolting displayed P_{ST} values lower than the mean F_{ST} suggesting convergent selection (a single phenotype favored across several

regions). Two traits: days to flower and spine length demonstrated slightly higher P_{ST} values than $F_{ST,}$ indicating that there is variation in these traits among regions, however their credibility intervals did not have enough statistical power to indicate strong evidence of divergent selection (Table 9, Figure 18).

Table 9. Global phenotypic and neutral genetic differentiation (P_{ST}-F_{ST} comparison) at seven morphological traits in *C*. *solstitialis*. Traits putatively under selection are highlighted in bold

Trait	Mean P_{ST} (95% CI)	Mean F_{ST}	P_{ST} - F_{ST} Bayesian credibility interval (95% CI)
Days to bolting	0.00 (0.00 – 0.03)	0.04 (0.03 – 0.05)	-0.04 (-0.05 – (-0.01)
Days to flower	0.05 (0.02 – 0.10)	0.04 (0.03 – 0.05)	0.01 (-0.03 – 0.06)
Final plant height	0.01 (0.00 – 0.05)	0.04 (0.03 – 0.05)	-0.03 (-0.05 – 0.01)
Number of inflorescence	0.03 (0.01 – 0.07)	0.04 (0.03 – 0.05)	-0.01 (-0.04 – 0.02)
Number of ovules	0.05 (0.01 – 0.11)	0.04 (0.03 – 0.05)	0.00 (-0.04 – 0.07)
Mass of plumed seeds	**0.19 (0.09 – 0.29)**	0.04 (0.03 – 0.05)	**0.14 (0.05 – 0.25)**
Spine length	0.06 (0.01 – 0.10)	0.04 (0.03 – 0.05)	0.01 (-0.02 – 0.05)

In the pair-wise P_{ST} - F_{ST} comparisons, no differences were observed between regions in terms of bolting days, plant height, inflorescence and ovule numbers. Differences in flowering time and spine length were evident between California and Australia (with plants from California demonstrating significantly earlier flowering time and larger spines) as also shown by the significantly higher P_{ST} values for these traits compared to F_{ST} *(*$P_{ST\ \text{days to flowering}}$ = 0.16 (CI: 0.041-0.29); $P_{ST\ \text{spine length}}$ = 0.13 (CI: 0.03 – 0.27). Furthermore, P_{ST} values for plummed seed mass exceeded F_{ST} values between California and the two native regions (Turkey and Spain) and California and Chile, with California presenting considerably larger seeds. These P_{ST} values for plumed seed mass were all above 0.30 (see Appendix 23 for Bayesian credibility intervals).

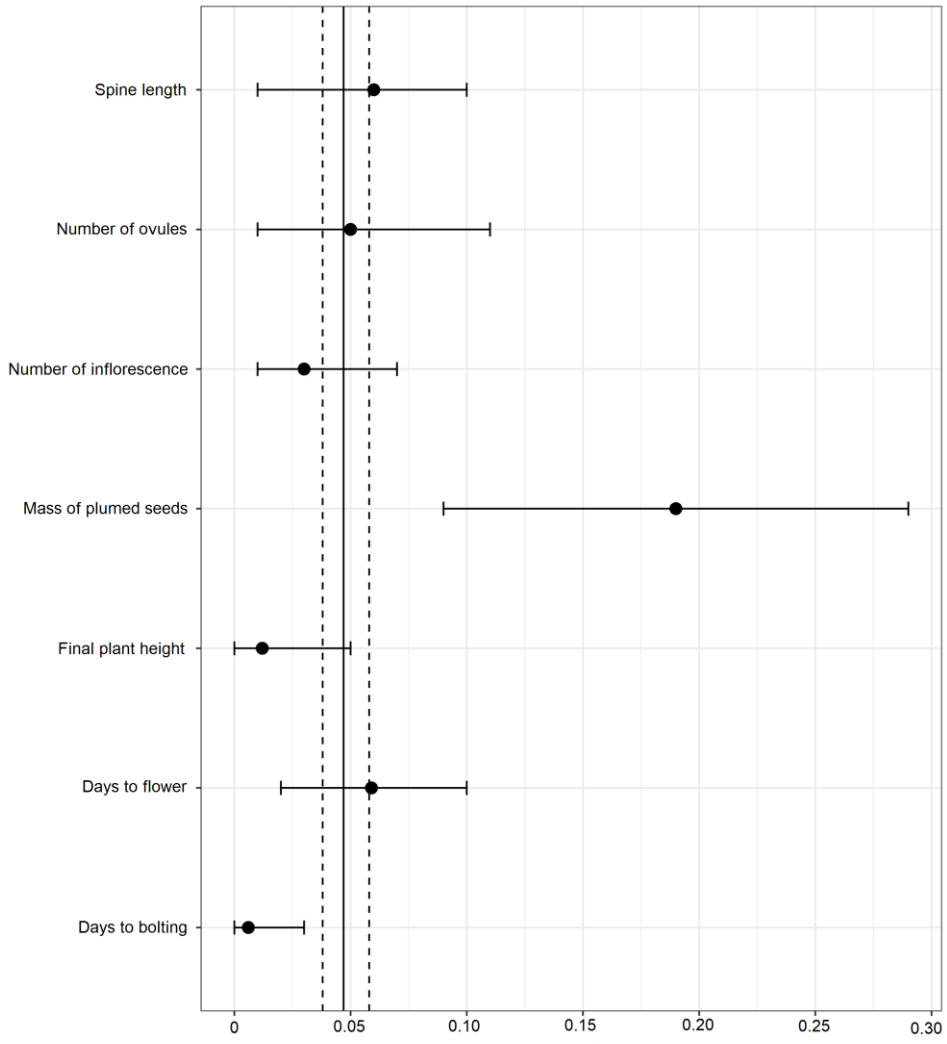


Figure 18. Global phenotypic trait divergence (P_{ST}) and neutral genetic differentiation (F_{ST}) based on 1975 neutral SNP loci. The vertical solid line represents mean global F_{ST} value and its lower and upper 95% CI (vertical dash lines). Horizontal bars indicate mean P_{ST} values (black dots) and its lower and upper CI

Population neutral genomic differentiation

In general, genomic diversity was similar across native and introduced ranges in terms of allelic richness, observed heterozygosity and expected heterozygosity. In the non-native range, populations from Argentina displayed the highest value for effective population size. By comparison, populations from Chile exhibited the lowest value across all regions while the two

native regions had similar values for this index. Moreover, Australia, California and Spain demonstrated higher inbreeding coefficients compared to the rest of the regions (Table 10).

Table 10. Summary statistics calculated based on 1975 neutral single nucleotide polymorphism loci of *C. solstitialis* in the native (Turkey and Spain) and non-native ranges (Argentina, Chile, California and Australia)

Region	Range	N	A_r	H_o	H_e	N_e	F_{IS}
Turkey	native	24	1.76	0.17	0.18	54.8	0.07
Spain	native	24	1.83	0.17	0.19	53.0	0.10
Argentina	non-native	24	1.86	0.18	0.19	177.5	0.05
Chile	non-native	24	1.8	0.18	0.19	16.8	0.03
California	non-native	24	1.85	0.17	0.19	33.7	0.11
Australia	non-native	24	1.81	0.15	0.18	34.8	0.13

Table legend: N, number of individuals analysed, A_r – allelic richness, H_o – observed heterozygosity for polymorphic loci, H_e – expected heterozygosity, N_e – effective population size, F_{IS} – inbreeding coefficient for a population

Pairwise region comparison of F_{ST} calculated from 1975 genome wide SNPs markers revealed low to moderate genetic differentiation between different *C. solstitialis* regions ranging from 0.02 to 0.09 (CI: 0.009 – 0.11). Pairwise comparisons between Turkey and all other regions demonstrated the highest values of F_{ST} (Table 11).

Table 11. Pairwise comparisons of F_{ST} values based on 1975 neutral SNP loci

Region	Turkey	Spain	Argentina	Chile	California	Australia
Turkey		0.087 (0.06-0.11)	0.078 (0.05-0.10)	0.089 (0.06-0.11)	0.064 (0.04-0.08)	0.087 (0.06-0.11)
Spain			0.029 (0.01-0.04)	0.042 (0.02-0.06)	0.036 (0.02-0.05)	0.037 (0.02-0.05)
Argentina				0.027 (0.01-0.04)	0.022 (0.01-0.03)	0.021 (0.00-0.03)
Chile					0.035 (0.02-0.05)	0.033 (0.02-0.05)
California						0.025 (0.01-0.04)
Australia						

Population structure

Discriminant analysis of principal components revealed population structuring in the native and non-native ranges and offered support for the presence of three genetic groups (Fig. 19). Individuals from native Turkey and non-native Chile were separated from all other regions, and each was forming an independent group, whereas individuals from the remaining geographic areas were less distinct and showed overlap. One point to consider is that DAPC offers a low-dimensional way of visualizing the data and may be limited in capturing fine population structuring.

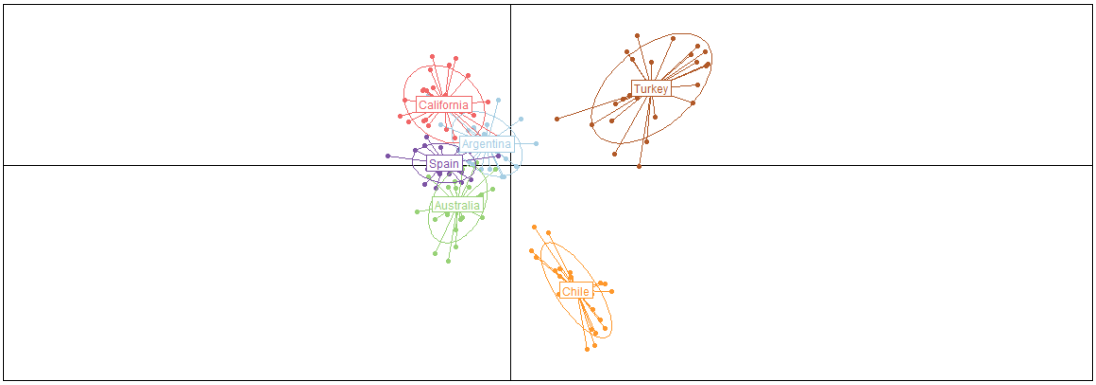


Figure 19. Discriminant analysis of principal components (DAPC) based on single nucleotide polymorphisms and using regions as prior clusters. Ovals are 95% inertia ellipses. Lines connect each individual to the regional mean value

STRUCTURE analysis identified K = two as the most probable number of genetic clusters (see Appendix 24 for ΔK values). In general, results of STRUCTURE were consistent with those of DAPC, except for Chile, which in this case appeared to be part of a single genetic cluster grouping populations from Spain and the introduced range (Fig. 20). In the native area, individuals from Turkey were clearly distinct from those in Spain, although some individuals in Turkey were not assigned with such high certainty. An additional genetic group defined Spain and individuals from the rest of the invasive range. In the introduced range, individuals from California showed slightly higher residual assignment to Turkey than Argentina and Australia.

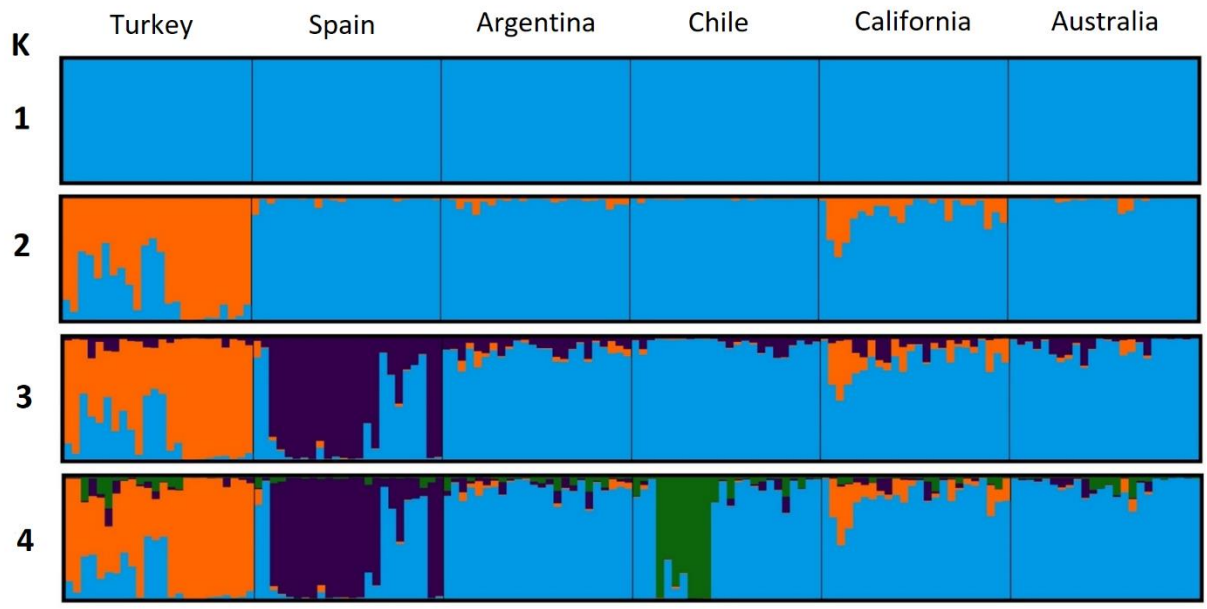


Figure 20. Individual assignments from STRUCTURE analysis based on 1975 neutral SNP loci of 144 individuals of *C. solstitialis*. STRUCTURE bar plots of assignment probabilities for the best estimate of K (2), for k = 1 to 4, where K is the number of genetic groups. Each vertical bar shows the proportional representation of the estimated group membership for a single individual

Outlier loci detection

Out of 2138 SNP markers, 163 were identified as putative outliers. Three outlier loci were identified by all three methods. Another three outliers were identified both by BayeScan and OutFLANK and one outlier both by BayeScan and PCAdapt. Lastly, twelve outliers were identified

both by OutFlank and PCAdapt. Consequently, 19 loci in total were identified by more than one method. The remaining 144 loci were identified by one method only as following: BayeScan: 4 loci, OutFLANK: 18 and PCAdapt: 122 (Fig. 21, Appendix 25).

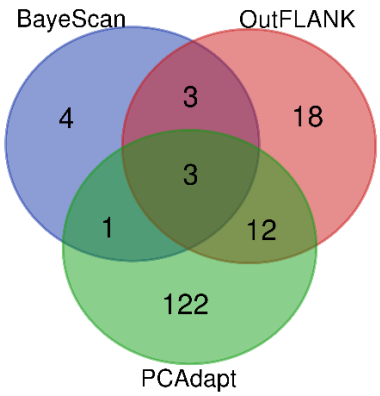


Figure 21. Venn diagram of putatively outlier loci identified by BayeScan, OutFLANK and PCAdapt

Gene ontology

For the gene ontology analysis, we only selected those outlier loci that were identified by more than one method (19 in total). All these loci where successfully annotated as coding for genes after subjecting their sequences to BLASTx. Outlier gene annotation were grouped into several categories of the Biological Process GO including: response to stimuli (i.e. response to heat, response to oxidative stress and response to light), regulation (rRNA and tRNA methylation, transcription regulation, chaperon cofactor protein refolding, rRNA processing), reproduction (mega-gametogenesis) and physiology and development (proteolysis, catabolism, oxidation reduction processes, heme biosynthesis, ubiquitin dependent protein catabolic process) (Fig. 22, Appendix 26). Genes were localized by cellular component to either cell or organelle as shown in Fig. 22. Out of these, 5 genes were assigned to the nucleus, 1 to Golgi complex, 6 to chloroplast, 3 to mitochondrion, 6 to cytoplasm, 4 to cytosol and 9 to membrane category. In most of the cases, genes showed multiple localizations. As for molecular function, the annotation of the three outlier loci detected by all three methods identified Transducin/WD40 repeat-like superfamily protein, SMAD/FHA domain-containing protein and putative S-adenosyl-L-

methionine-dependent methyltransferase genes (Appendix 26). The first two genes are involved in protein binding while the later possess protein methyl-transferase activity. Similarly, the rest of the outlier genes identified by two outlier methods were found to be involved in DNA binding, ATP binding, mRNA binding, chaperone binding, metal ion binding, ion sulfur cluster binding, protein binding, or to possess methyl-transferase activity, peptidase-, transferase-, phosphatase-, oxidoreductase- activities as well as protein dimerization activity (Fig. 22, Appendix 26).

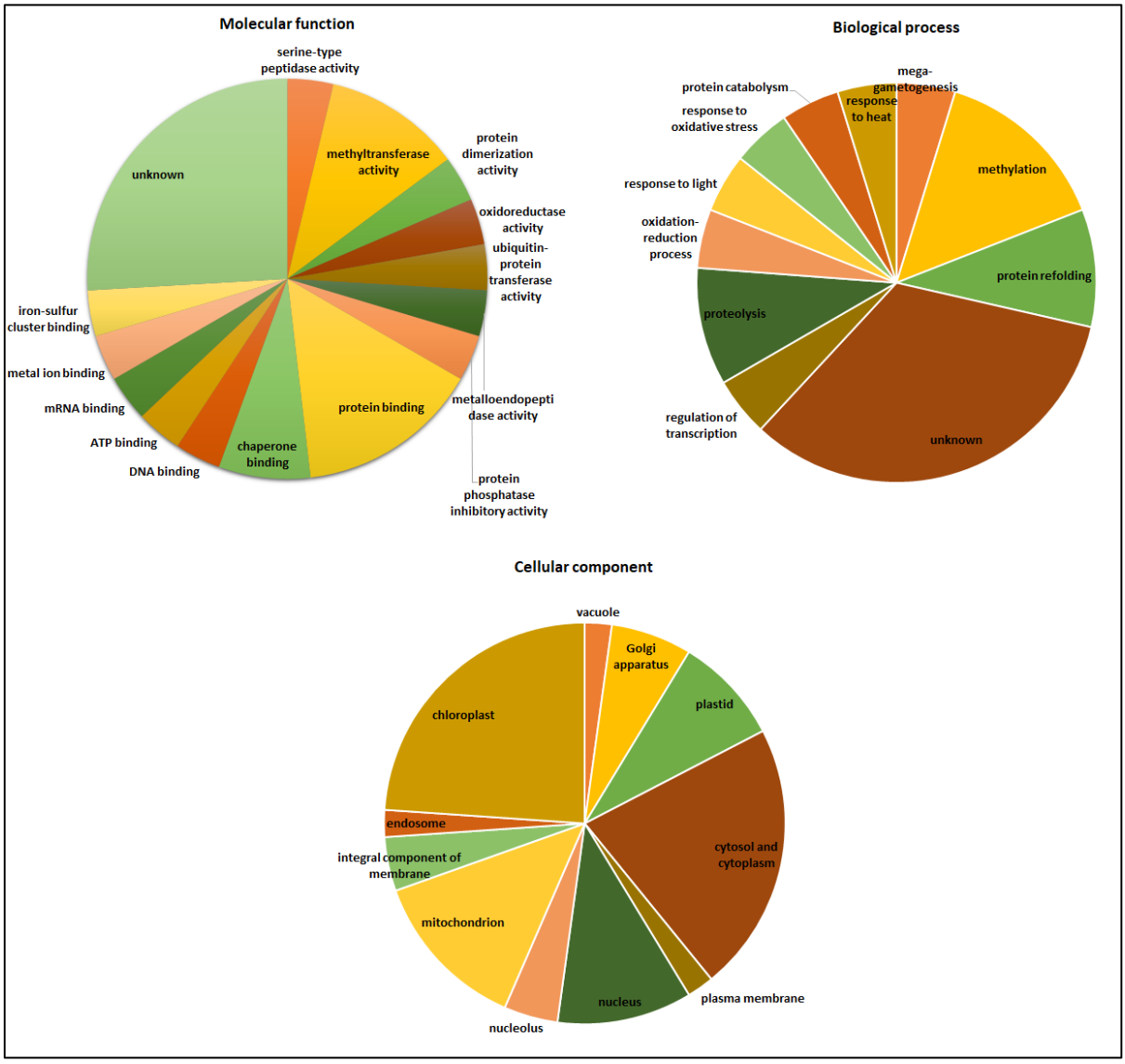


Figure 22. Outlier gene ontology annotation of 19 outlier genes identified in *Centaurea solstitialis*

DISCUSSION

In this study we used multiple plant traits linked to growth, reproduction and defense; and genome wide SNP markers to search for signatures of natural selection at phenotypic and genomic levels in the highly invasive *C. solstitialis*. We found little genetic differentiation between *C. solstitialis* regions and detected several loci putatively under selection associated with stress responses and regulation of physiological processes or other genes (i.e methylation and transcription regulation, rRNA processing and protein refolding). Additionally, we detected strong evidence of selection for increased seed mass in non-native regions compared to natives as well as shifts in flowering time and spine defense between two non-native regions. Taken together this suggests that rapid local adaptation via both genetic and epigenetic mechanisms are shaping the observed differentiation between invasive *C. solstitialis* populations in different regions.

We found evidence of significant differentiation in morphological traits between *C. solstitialis* regions for six out of seven traits measured, with individuals from the native areas tending to produce greater number of inflorescences whereas non-native individuals tended to produce more ovules and heavier plumed seeds. These results support previous findings of increased seed mass in *C. solstitialis* non-native range compared to native (Graebner et al. 2012; Hierro et al. 2013) but do not comply with the hypothesis that *C. solstitialis* had evolved towards larger plant size in the introduced area (Widmer et al. 2007; Eriksen et al 2012; Garcia et al. 2013; Barker et al. 2017). Genetic differentiation between regions was low to moderate, but with clear population structure between native and non-native regions, resulting in the discrimination of four genetic groups. Two genetic groups were present in the native area differentiating *C. solstitialis* individuals in Turkey from those in Spain. A second genetic group comprised individuals from Spain, Argentina, Chile, California and Australia. In general, our results were consistent with those reported before in the literature for this species (Eriksen et al. 2014; Barker et al. 2017). However, contrary to those two previous studies which detected a single genetic group in South America, our DAPC analysis (but not structure) found evidence of two distinct groups distinguishing Chile from Argentina. These differences may be partially accounted for by the use of different molecular markers across studies (EST-SSR in Eriksen et al. 2014) or by

previous very low sampling in Argentina (Barker et al. 2017).

P_{ST} estimates overlapped F_{ST} for most of the phenotypic traits investigated, except plummed seed mass, which demonstrated greater P_{ST} than neutral genetic differentiation, suggesting a role of divergent natural selection in shaping the morphological variation in this trait. In contrast, days to bolting exhibited significantly lower P_{ST} than F_{ST}, indicating convergent selection. Interestingly, a previous Q_{ST}-F_{ST} comparison of *C. solstitialis* genotypes from Republic of Georgia, Turkey, Argentina and California found evidence for selection of increased plant size in the non-native regions as well as large variation in days to bolting across all regions (Eriksen et al. 2012). Evidence from literature is mixed, with some studies reporting no differences in plant size between native and non-native ranges (Hierro et al. 2006, Andonian et al. 2011) while others showed increased growth and higher biomass in *C. solstitialis* introduced range (Graebner et al. 2012; García et al. 2013; Dlugosch et al. 2015; Montesinos and Callaway 2018). These contrasting patterns might be due to natural variability present within different populations of *C. solstitialis* or reflect variable trait expression (phenotypic plasticity) under different environmental conditions used by the different studies. Nonetheless, our results support previous observations that seed mass is putatively an adaptive trait in some of the species introduced ranges (Hierro et al. 2011; Hierro et al. 2013). The specific selective agent(s) behind increased seed mass in *C. solstitialis* introduced ranges remains unknown, but several explanations could account for this trait variation. Seed mass can be related to climatic conditions as it was found in the invasive *Echium plantagineum*, where populations sourced from hot, arid sites produced heavier seeds than populations from wetter sites, as a strategy to ensure reproductive success in arid environments (Konarzewski et al. 2012). Larger seeds usually produce larger seedlings, which is positively associated with survival, increased competitive ability and the capacity to withstand different hazards (i.e. herbivory, low nutrients, pathogen attack) (Coomes and Grubb, 2003; Hierro et al. 2013). Additionally, larger seeds might have been preferentially introduced during colonization of new ranges (Buckley et al. 2003). Levels of genetic diversity were similar between regions, lending support to the recent colonization and admixture scenario, as shown before by previous molecular studies on *C. solstitialis* (Eriksen et al. 2014; Barker et al. 2017). By comparison, other successful invaders such as *Impatiens glandulifera* and *Acacia saligna* were

found to possess limited genetic diversity in their introduced areas compared to the native ranges, despite several introductions (Hagenblad et al. 2015; Le Roux et al. 2011). Estimates of effective population size tended to be somewhat lower in Chile, consistent with a founder effect population bottleneck following introductions from Spain (Barker et al. 2017). In this study, individuals from Chile exhibited significantly lower fitness in terms of growth and reproduction compared to natives but also relative to other non- native populations in the Americas (i.e. Argentina and California), possibly as a consequence of the bottleneck and low effective population size. The rest of the N_e estimates were in the range previously reported for this species (Braasch et al. 2018), except for Argentina, which demonstrated higher effective population size, that could be regarded as a signal of admixture of different source populations associated with introductions. Historical records show that *C. solstitialis* introduction in the Americas had been linked to the alfalfa farming. An increase in demand for this crop during the first two decades of the twentieth century lead to extensive international trade and massive imports of cheap alfalfa seeds from various sources in Eurasia, which was contaminated with different weed seeds, including those of *C. solstitialis* (Gerlach 1997). The presence of increased levels of genetic diversity within these regions (Sun 1997; Dlugosch et al. 2013; Eriksen et al. 2014) as well as no trends in inbreeding coefficients moving from native to introduced regions argue that *C. solstitialis* invasion in this part of the Americas did not involve strong founder effect population bottlenecks (Barker et al. 2017, see also Zhu et al. 2019). Additionally, our observation of no increase in inbreeding in invasive regions confirms that no shifts in reproductive system from outcrossing towards selfing have occurred across the introduced range of *C. solstitialis*.

The pairwise F_{ST} comparisons showed relatively low genetic differentiation throughout the *C. solstitialis* introduced range, indicating recent colonization. STRUCTURE and DAPC results suggested the presence of some genetic substructure in South America, with Chilean and Argentinean populations belonging to two distinct genetic groups. Low N_e in Chile individuals, coupled with the distinctive phenotype and an incipient degree of reproductive incompatibilities detected between these two regions (Irimia et al. unpublished), as well as the presence of geographic barriers (the Andes mountains), seem to support that they are evolving

independently. Similarly, P_{ST}-F_{ST} analysis at the regional level indicated divergent selection for increased plumed seed mass in California compared to the two native regions (Turkey and Spain) as well as to non-native Chile. Additionally, P_{ST} for days to flower and spine length were significantly different between California and Australia, with fewer days spent in the vegetative state and larger spine length for Californian individuals relative to Australian ones. Earlier flowering phenology is an adaptive trait and could offer an advantage in the Mediterranean dry climate of California where summer rainfall is uncertain (see also Dlugosch et al. 2015). Adaptive phenotypic divergence in genes underlying photoperiodic flowering phenology was demonstrated in *Sisymbrium austriacum* during the establishment phase, in response to seasonal differences between native and invasive ranges (Vandepitte et al. 2014). Likewise, increased spine length in California may have evolved as a response against mammalian herbivores (García et al. 2013) or to deter illegitimate flower insects i.e. nectar robbing lepidopterans (Agrawal et al. 2000). Since our plants were grown under common greenhouse conditions, we can exclude major environmental effects on the traits we measured. Additionally, maternal effects are probably minor in *C. solstitialis* as in other common garden studies that tested for maternal effects in *C. solstitialis*, no differences in seed attributes and germination rates were found between the progeny of wild sampled plants or the progeny of cultivated plants (Widmer et al. 2007; Hierro et al. 2009). No estimation of trait heritability has been published to date in *C. solstitialis*. Nonetheless, despite all these challenges, our P_{ST} estimates were highly conservative and informative about the traits under divergent selection.

This study is the first to generate molecular markers data for non-native Australia. *Centaurea solstitialis* was first reported in Australia around 1856, although the circumstance of its introduction there remains unknown (Parsons and Cuthbertson 2001). Our genetic structuring analyses showed that Australia clustered with two other non-native regions: California and Argentina while DAPC also showed genetic overlap with Spain. Our results indicate that Australia has been colonized from elsewhere in the invasive range, but the lack of genetic differentiation between most of the invasive regions, limits resolving the source population further. Individuals from this region exhibited similarly larger plumed seeds compared to those from Argentina and California. Future Bayesian simulation modelling of genetic data might be able to shed further

light on the invasive spread of *C. solstitialis* to Australia.

Our study also made progress identifying the genetic targets of selection underlying some of the evolutionary changes that have been detected in invasive *C. solstitialis*. None of the putatively divergently selected gene functions has an obvious link with the measured traits, but instead point to more subtle traits related to regulation of gene expression to modulate developmental processes or stress tolerance responses through DNA binding, ATP binding, methyltransferase activity or regulation of transcription. For example, S-adenosyl-L-methionine dependent methyltransferases catalyzes protein methylation and the post-translational modification can serve diverse functions in plant growth and development (Yuan et al. 2015). Transducin/WD40 repeat proteins mediate diverse protein–protein interactions and act as key regulators of several plant developmental events (Gachomo et al. 2014). SMAD/FHA domain containing protein is known to be involved in DNA damage repair and signal transduction (Yu et al. 2008). HBP5 is a hem binding protein involved in defense against oxidative stress (Lee et al. 2012). This protein has been observed to enable weak chloroplast movement under blue light and is thought to be involved in chloroplast photo-relocation movement and positioning, protecting against photo-oxidative stress (Kodama et al. 2011). Similarly, FtsH proteases are involved in protein quality control during photosynthesis and avoidance of photo-inhibition (Kato and Sakamoto 2018).

Further studies of these candidate genes may aid in the understanding of their ecological roles in *C. solstitialis* invasion success. For example, stress induced gene expression (i.e. heat, drought) could lead to differences in transcriptional regulation and molecular physiology between native and introduced plant genotypes. Under drought conditions, invasive populations of *Centaurea diffusa* in USA showed higher levels of gene expression related to energy production, compared to native populations, suggesting greater ability to modulate fitness homeostasis (Turner et al. 2017). Genome scans in combination with quantitative trait loci mapping (QTL) can lead to the identification of genomic regions associated with ecologically important traits involved in successful invasions but currently there is a paucity of these studies in invasion science (Prentis et al. 2008).

CONCLUSIONS

In summary, our findings show that genetic drift is probably not the sole force causing the observed phenotypic variation between *C. solstitialis* ranges, indicating also a role of divergent selection. Additionally, this study also provides support for rapid evolution in the introduced range and highlights the presence of different evolutionary and ecological dynamics in each region.

General Discussion

Invasive species represent grand but unplanned experiments across large spatial and temporal scales that can provide unique data about different ecological and evolutionary processes and contribute to a better understanding of the natural world (Sax et al. 2007). The field of invasion ecology has been very prolific in formulating and developing new theories but much less inclined or able to reject those that have found little support (Davis 2011). There is a pressing need to elucidate the "nuts and bolts" of invasive organisms to reduce associated hazards and undertake management actions. As outlined in the general introduction, this thesis takes a biogeographical approach and aims to address some of the questions related to the ecology and genomic of rapid adaptation and evolution in an invasive thistle (Asteraceae), and pioneers on the use of experimental crosses of plants originating from different world regions. Below, I summarize the main findings and the insights gained from this work together with the limitations and future perspectives.

***Centaurea solstitialis* systematic review**

A review of the scientific literature over the past 70 years showed that studies published in *C. solstitialis* were relatively limited until 1990s and greatly expanded in the next two decades, consistent with a rising trajectory of the literature on biological invasions in general, and an increase in public awareness of the issues related to invasive species (Lowry et al. 2013). A large proportion of the published studies included experimental work performed either under controlled conditions or on natural populations in the field, but the duration of the study varied extensively across studies. In general, ecological factors were common in explaining species success, with realtively few studies categorized as having an evolutionary focus. The literature search revealed a longstanding interest on invasive plant control, which was also reflected in the high number of studies regularly published on this topic (44%). This is not surprising as *C. solstitialis* control in the western US remains elusive despite an integrated management approach and decades after the introduction of several capitula feeding insects and a foliar rust pathogen, probably because of compensatory mechanisms in seed production and rapid re-infestation of cleared areas (Swope and Parker 2010, Kyser et al. 2013). Interestingly, although *C. solstitialis* is listed as highly invasive in central Argentina too, we identified no studies in the

WoS to focus on control measures in that area except for a database on invasive plant species and a weed survey (Busso et al. 2013; Scursoni et al. 2017). Another 30% of the papers aimed to study the ecological dynamics of introduced populations in relation to enemies, mutualists and competitors, in trying to explain their demographic success (Mitchell et al. 2006). About 13% of the remaining articles focused on the phytochemical and pharmacological properties of *C. solstitialis*. Aerial parts of the plant are harvested for their perceived medicinal properties by indigenous people in the Middle East (i.e. Turkey, Iran), and used to treat different aliments (Yeşilada et al. 1999; Bahmani et al. 2014). At the same time, the plant is known to cause a fatal neurological disorder in horses (Moret et al. 2005).

The most obvious pattern emerging from the review of the existent literature on *C. solstitialis* was the scarcity of field parallel studies in the native and non-native species ranges (five studies only). The few studies of this kind aimed to assess the importance of several factors such as herbivory, disturbance, fire and interaction with pathogens on the species performance under field conditions (Hierro et al. 2006; 2017; Andonian et al. 2011) or to document differences in plant abundance and demography (Hierro et al. 2017; Andonian et. al 2011) and investigate pollinator guilds (Barthell et al 2009). These studies were limited to a few geographical regions and mostly involved comparisons of introduced populations in Argentina and California with the *C. solstitialis* ´´core region´´ (i.e. Caucasus and Turkey) but did not include sites in Spain which is regarded as the primary source of seeds that have colonized the Americas (Barker et al. 2017). In contrast, studies comparing the phenotypes or competitive ability of several *C. solstitialis* native and non-native populations grown together in the same environment (i.e. greenhouse, garden, growth chamber) in the introduced or native range were much more abundant (as also pointed out by Hierro et al. 2005), and found mixed support for increased performance of introduced genotypes (Hierro et al. 2006, Andonian et al. 2011), possibly reflecting genotype by environment interactions or phenotypic plasticity.

So, despite the vast amount of literature published in *C. solstitialis* to date, several questions remain unanswered particularly those related to the importance and variation in factors controlling species abundance across native and non-native areas.

Assessing the role of polyploidy and genome size on species success

In Chapter I, I provided strong evidence that *C. solstitialis* invasion in the Americas and Australia was not associated with shifts in ploidy levels and that the species occurs as a single (diploid) cytotype across its world distribution. In addition, I found no significant difference in genome size between ranges (Irimia et al. 2017), suggesting that selection has had little impact on genome size evolution in this species. There is a growing interest in the literature to use genome size as a predictor of plant invasiveness with several studies supporting the idea that invaders are likely to have a reduced genome size compared to non-invasive plants (Suda et al. 2015, for a literature review). However, intraspecific comparisons of genome size variation in invasive species and its relationship with traits favouring invasiveness are scarce (see Lavergne et al. 2010, Pyšek et al. 2018; Martinez et al. 2018), and my study provides new insights into this topic. Although genome size estimates were similar across *C. solstitialis* native and introduced populations, it does not rule out the role of other factors such as DNA structural and regulatory modifications (i.e. transposable elements, epigenetic changes), but currently the lack of a reference genome makes it challenging to investigate. For example, transposable elements were found to be highly enriched in the invasive *Capsella rubella* compared to its sister species *C. grandiflora* and were associated to natural variation in flowering time, trait correlated to fitness and adaptation (Niu et al. 2018).

Testing the allelopathic potential of leaf leachates

In contrast, in Chapter II, I showed that *C. solstitialis* leaf surface leachates possess allelopathic potential and vary extensively in metabolic profiles and inhibitory effects on phytometer species, across species range (Irimia et al. 2019). This suggests that different selection pressures may act on biochemical signatures in different world regions i.e. release from competitors or herbivores and some non-native populations can perform better than others. Additionally, I found that inhibitory effects were significantly weaker in two of the non-native regions compared to native, challenging the general assumption in invasion biology that introduced genotypes tend to outperform native genotypes. My results open up new avenues for understanding plant-plant interactions, and also point out the need to conduct

biogeographical field studies to confirm allelopathic effects on resident plant communities and assess the role of herbivores on leaf defence investment.

Screening for reproductive isolation in allopatric populations

Reproductive isolation is fundamental to speciation (Coyne and Orr 2004) and invasive species are good candidates to test for the presence of reproductive barriers because they have been introduced to new regions with different environmental conditions which sets the stage for evolutionary diversification among allopatric populations (Vellend et al. 2007). In Chapter III, I investigated the geographical patterns of reproductive success and fertility rates across native and non-native populations of *C. solstitialis* and observed an asymmetrical response to inter-regional gene flow characterized by variable fitness effects and a global mosaic of reproductive incompatibilities and fertilities. Two strong patterns emerged from these backcrosses: i) native *C. solstitialis* populations from Spain showed a preference for non-native pollen, which translated into up to 70% increase in fertility rates and ii) non-native populations in the Americas appeared to reject pollen from both native as well as other non-native regions exhibiting up to 50% reduction in reproductive output. Reproductive barriers in the Americas were moderate, leaving the opportunity for reinforcement to occur and character displacement. This study can be regarded as a test for early stages of ecological speciation and indicates that reproductive barriers can emerge relatively quickly in allopatry (i.e. less than 400 and 200 years since *C. solstitialis* introduction to different regions in the Americas). Future experimental studies to test the fate and fitness of F_1 and F_2 progenies obtained from these controlled backcrosses coupled with genomic studies could clarify the nature of the reproductive incompatibilities detected here.

Evaluating the adaptive phenotypic and genomic potential of introduced populations

In Chapter IV, I tested for signatures of natural selection at phenotypic and genomic levels in *C. solstitialis* native and introduced regions by using phenotypic data from a greenhouse experiment and genome wide SNP markers. I found that genetic diversity was similar between native and introduced populations, providing evidence of no major genetic bottlenecks in the non-native range, post-introduction. A number of 19 SNPs markers appeared to be top

candidates for divergent selection, all located within genes associated to regulation of gene expression of developmental processes, regulation of transcription and stress tolerance responses. There was also strong indication that seed mass is an adaptive trait. Compared to native populations, California displayed significantly larger seeds, whereas in the non-native area, Australia exhibited a significant shift in flowering time towards later flowering and decrease in spine length relative to California. These findings support the conclusion that selection and possibly epigenetic changes shaped the evolution of *C. solstitialis* in the introduced ranges. Further work is needed to establish the causative selective agent(s) behind these traits divergence and when it arose. A previous study showed that increased seed size in Argentinean populations of *C. solstitialis* could ensure better survival under non-native conditions (Hierro et al. 2013) whereas earlier flowering in California could help to avoid summer drought (Dlugosch et al. 2015). Genomic data supports that Chilean populations were the source of seeds introduced to California (Barker et al. 2017) and in our study, Chilean populations presented significantly smaller seeds compared to California, suggesting that larger seeds in this region likely arose post-introduction, possibly as a result of new biotic interactions (Dlugosch et al. 2015). Future studies could also attempt to search for differences in gene expression patterns between native and introduced populations of *C. solstitialis* under different conditions or stressors to identify candidate genes involved in invasiveness. Another research direction would consist in performing genome wide association studies to test if genetic regions are associated with specific phenotypic traits that confer a fitness advantage.

Conclusions

This thesis contributes to our understanding of ecological and evolutionary changes in one of the most noxious weeds in the Americas and Australia and represents an effort to integrate cytogenetic, chemical, phenotypic and genomic data from an unprecedented high numbers of populations and sites in both the native and introduced species range, to provide new insights into the global mosaic of local adaptation. This research also constitutes some of the first intraspecific investigations into the emergence of reproductive isolation in geographically isolated populations of an invasive species.

Literature cited

Alvarez, P., 2008. Sheep grazing to control yellow starthistle: role of sesquiterpene lactones. PhD thesis, University of California.

Almeida, A.B.A., Miotto, A.M., Nunes, D.S., Spadari Bratifisch R.C., Brito A.R.M., 2002. Mechanism of antiulcerogenic activity of semi-synthetic crotonin obtained from *Croton cajucara* Benth. *Revista Brasileira de Farmacognosia* 12:105-110. http://dx.doi.org/10.1590/S0102-695X2002000300050.

Andonian, K., Hierro, J.L., Khetsuriani, L., Becerra, P., Janoyan, G., Villarreal, D., Cavieres, L., Fox, L.R., Callaway, R.M., 2011. Range expanding populations of a globally introduced weed experience negative plant soil feedbacks. *Plos One* 6: e20117. https://doi.org/10.1371/journal.pone.0020117.

Agrawal, A.A., Rudgers, J.A., Botsford, L.W., Cutler, D., Gorin, J.B., Lundquist, C.J., Spitzer, B.W., Swann, A.L., 2000. Benefits and constraints on plant defense against herbivores: spines influence the legitimate and illegitimate flower visitors of yellow star thistle Centaurea solstitialis L (Asteraceae). *The Southwestern Naturalist* 45(1):1-5. https://doi.org/10.2307/3672545.

Aschehoug, E.T., Brooker, R., Atwater, D.Z., Maron, J.L., Callaway, R.M., 2016. The mechanisms and consequences of interspecific competition among plants. *Annual Review of Ecology, Evolution and Systematics* 47:263-281. https://doi.org/10.1146/annurev-ecolsys-121415-032123.

Aschehoug, E.T., Callaway, R.M., Newcombe, G., Tharayil, N., Chen, S., 2014. Fungal endophyte increases the allelopathic effects of an invasive forb. *Oecologia* 175:285-291. https://doi.org/10.1007/s00442-014-2891-0.

Baack, E., Melo, M.C., Rieseberg, L.H., Ortiz-Barrientos, D., 2015. The origins of reproductive isolation in plants. *New Phytologist* 207:968-984. https://doi.org10.1111/nph.13424.

Bachelier, A., Mayer, R.M, Klein, C.D., 2006. Sesquiterpene lactones are potent and irreversible inhibitors of the antibacterial target enzyme MurA. *Bioorganic and Medicinal Chemistry Letters* 16:5605-5609. https://doi.org/10.1016/j.bmcl.2006.08.021.

Bahmani, M., Zargaran, A., Rafieian-Kopaei, M., 2014. Identification of medicinal plants of Urmia for treatment of gastrointestinal disorders. *Revista Brasileira de Farmacognosia* 24:468

-480. https://doi.org/10.1016/j.bjp.2014.08.001.

Bancheva, S., Greilhuber, J., 2006. Genome size in Bulgarian *Centaurea* s.l. (Asteraceae). *Plant Systematics and Evolution* 257:95–117. https://doi.org/10.1007/s00606-005-0384-7.

Barker, B.S., Andonian, K., Swope, S.M., Luster, D.G., Dlugosch, K.M., 2017. Population genomic analysis reveal a history of range expansion and trait evolution across the native and invaded range of yellow starthistle (*Centaurea solstitialis*). *Molecular Ecology* 26(4):1131-1147. https://doi.org/10.1111/mec.13998

Barker, B.S., Cocio, J.E., Anderson, S.R., Braasch, J.E., Cang, F.A., Gillette, H.D., Dlugosch, K.M., 2019. Potential limits to the benefits of admixture during biological invasion. *Molecular Ecology* 28(1):100-113. https://doi.org/10.1111/mec.14958

Barthell, J.F., Randall, J.M., Thorp, R.W., Wenner, A.M., 2001. Promotion of seed set in yellow star-thistle by honey bees: evidence of an invasive mutualism. *Ecological Applications* 11(6):1870-1883.

Barthell, J.F., Clement, M.L., Song, D.S., Savitski, A.N., Hranitz, J.M., Petanidou, T., Thorp, R.W., Wenner, A.M., Griswold, T.L., Wells, H., 2009. Nectar secretion and bee guild characteristics of yellow star-thistle on Santa Cruz island and Lesvos: where have the honey bees gone? *Uludağ Bee Journal* 9:109-121

Bartoli, A., Galati, B.G., Tortosa, R.D., 2011. Anatomical studies of the secretory structures: Glandular trichomes and ducts, in Grindelia pulchella Dunal (Astereae, Asteraceae). *Flora* 206:1063-1068. https://doi.org/10.1016/j.flora.2011.07.012.

Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1):1-48. https://doi.org/10.18637/jss.v067.i01.

Beaulieu, J.M., Moles, A.T., Leitch, I.J., Bennett, M.D., Dickie, J.B., Knight, C.A., 2007. Correlated evolution of genome size and seed mass. *New Phytologist* 173:422–437. https://doi.org/10.1111/j.1469-8137.2006.01919.x.

Beck, J.J., Smith, L., Merrill, G.B., 2008. In situ volatile collection, analysis and comparison of three *Centaurea* species and their relationship to biocontrol with herbivorous insects *Journal of Agricultural and Food Chemistry* 56:2759-2764. https://doi.org/10.1021/jf073383u.

Belz, R.C., van der Laan, M., Reinhardt, C.F., Hurle, K., 2009. Soil Degradation of parthenin – does

it contradict the role of allelopathy in the invasive weed Parthenium hysterophorus L.? *Journal of Chemical Ecology* 35:1137-1150. https://doi.org/10.1007/s10886-009-9698-1.

Bennett, M.D., Leitch, I.J., Hanson, L. 1998. DNA amounts in two samples of angiosperm weeds. *Annals of Botany* 82 (Supplement A):121–134. https://doi.org/10.1006/anbo.1998.0785.

Benefield, C.B., DiTomaso, J.M., Kyser, G.B., Tschohl, A., 2001. Reproductive biology of yellow starthistle: maximizing late-season control. *Weed Science* 49:83-90. https://doi.org/10.1614/0043-1745(2001)049[0083:RBOYSM]2.0.CO;2.

Bhattacharyya, P.R., Barua, N.C., Ghosh, A.C., 1996. Cynaropicrin from *Trichlorepis glaberrima*: a potential insect feeding deterrent compound. *Industrial Crops and Products* 4:291-294. https://doi.org/10.1016/0926-6690(95)00044-5.

Blanquart, F., Kaltz, O., Nuismer, S.L., Gandon, S., 2013. A practical guide to measuring local adaptation. *Ecology Letters* 16(9):1195-1205. https://doi.org/10.1111/ele.12150.

Bossdorf, O., Auge, H., Lafuma, L., Rogers, W.E., Siemann, E., Prati, D., 2005. Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* 144:1-11. https://doi.org/10.1007/s00442-005-0070-z.

Braasch, J., Barker, B.S., Dlugosch, K.M., 2018. Expansion history and environmental suitability shape effective population size in a plant invasion. *Molecular Ecology* 28(10):2546-2558. https://doi.org/10.1111/mec.15104.

Bradshaw, C.J.A., Leroy, B., Bellard, C., Roiz, D., Albert, C., Fournier, A., Barbet-Massin, M., Salles, J.M., Simard, F., Courchamp, F., 2016. Massive yet grossly underestimated global costs of invasive insects. *Nature Communications* 7:12986. http://dx.doi.org/10.1038/ncomms12986.

Bremer, K., 1994. Asteraceae – cladistics and classification. Portland: Timber press.

Bruno, M., Bancheva, S., Rosselli, S., Maggio, A., 2013. Sesquiterpenoids in subtribe Centaureinae (Cass.) Dumort (tribe Cardueae, Asteraceae): Distribution, ^{13}C NMR spectral data and biological properties. *Phytochemistry* 95:19-93. https://doi.org/10.1016/j.phytochem.2013.07.002.

Brommer, J.E., 2011. Whither P_{ST}? The approximation of Q_{ST} by P_{ST} in evolutionary and conservation biology. *Journal of Evolutionary Biology* 24(6):1160-1168.

https://doi.org/10.1111/j.1420-9101.2011.02268.x.

Buckley, Y.M., Downey, P., Fowler, S.V., Hill, R., Memmot, J., Norambuena, H., Pitcairn, M., Shaw, R., Sheppard, A.W., Winks, C., Wittenberg, R., Rees, M., 2003. Are invasive bigger? A global study of seed size variation in two invasive shrubs. *Ecology* 84(6):1434-1440. https://doi.org/10.1890/0012-9658(2003)084[1434:AIBAGS]2.0.CO;2.

Busso, C.A., Bentivenga, D.J., Fernández, O.A., 2013. A review on invasive plants in rangelands of Argentina. *Interciencia* 38(2): 95-103.

Busch, J.W., 2011. Demography, pollination, and Baker´s law. *Evolution* 65(5):1511-1513. https://doi.org/10.1111/j.1558-5646.2011.01224.x.

Burrows, G.E., Tyrl, R.J., 2013. Toxic plants of North America. 2nd Edition. Wiley-Blackwell. Asteraceae Martinov pp 150–256.

Callaway, R.M., Aschehoug, E.T., 2000. Invasive plants versus their new old neighbors: a mechanism for exotic invasion. *Science* 290:521-523. https://doi.org/ 10.1126/science.290.5491.521.

Callaway, R.M., Ridenour, W.M., 2004. Novel weapons: invasive success and the evolution of increased competitive ability. *Frontiers in Ecology and the Environment* 2:436-443. https://doi.org/10.1890/1540-9295(2004)002[0436:NWISAT]2.0.CO;2.

Callaway, R.M., Maron, J.L., 2006. What have exotic plant invasions taught us over the past 20 years? *Trends in Ecology and Evolution* 21(7):369-374. https://doi.org/10.1016/j.tree.2006.04.008.

Carev, I., Ruščić, M., Skočibušić, M., Maravić, A., Silijak-Yakovlev, S., Politeo, O., 2017. Phytochemical and cytogenetic characterization of *Centaurea solstitialis* L. (Asteraceae) from Croatia. *Chemistry and Biodiversity* 14, e1600213. https://doi.org/10.1002/cbdv.201600213.

Carpenter, D.J., 2007. The role of the novel weapons and enemy release hypothesis on the invasiveness of knapweeds, *Centaurea* spp., in North America. MSc Thesis. Carleton University, Ottawa, Canada: 95 pp.

Catchen, J., Amores, P., Hohenlohe, P., Cresko, W., Postlethwait, J., 2011. *Stacks: building and genotyping loci de novo from short-read sequences*. *G3: Genes, Genomes, Genetics*,

1:171-182. https://doi.org/10.1534/g3.111.000240.

Catchen, J., Hohenlohe, P., Bassham, S., Amores, A., Cresko, W., 2013. *Stacks: an analysis tool set for population genomics*. Molecular Ecology. https://doi.org/10.1111/mec.12354.

Chang, H.T., Rumbeiha, W.K., Patterson, J.S., Puschner, B., 2011. Toxic equine Parkinsonism: an immuno-histochemical study of 10 horses with nigropallidal encephalomalacia. *Veterinary Pathology* 49(2):398-402. https://doi.org/10.1177/0300985811406885.

Chen, S., Xiao, S., Callaway, R.M., 2013. Light intensity alters the allelopathic effects of an exotic invader. *Plant Ecology and Diversity*, 5:521-526. https://doi.org/ 10.1080/17550874.2012.736548.

Cheng, C.H.K., Costal, B., Hamburger, M., Hostettmann, K., Naylor, R.J., Wang, Y., Jenner, P., 1992. Toxic effects of solstitialin A 13 acetate and cynaropicrin from *Centaurea solstitialis* L (Asteraceae) in cell cultures of fetal rat brain. *Neuropharmacology* 31:271-277. https://doi.org/10.1016/0028-3908(92)90177-Q.

Chon, S.U., Nelson, C.J., 2010. Allelopathy in Compositae plants. A review. *Agronomy for Sustainable Development* 30:349-358. https://doi.org/10.1051/agro/2009027.

Cis, J., Nowak, G., Kisiel, W., 2006. Antifeedant properties and chemotaxonomic implications of sesquiterpene lactones and syringing from *Rhaponticum pulchrum*. *Biochemical Systematics and Ecology* 34:862-867. https://doi.org/10.1016/j.bse.2006.05.019.

Coomes, D.A., Grubb, P.J., 2003. Colonization, tolerance, competition and seed size variation within functional groups. *Trends in ecology and evolution* 18(6):283-291. https://doi.org/10.1016/S0169-5347(03)00072-7.

Coyne, J.A., Orr, H.A., 2004. Speciation. Sinauer Associates, Sunderland (Massachusetts). pp 545.

Daehler, C.C., 1998. The taxonomic distribution of invasive angiosperm plants: ecological insights and comparison to agricultural weeds. *Biological Conservation* 84:167-180. https://doi.org/10.1016/S0006-3207(97)00096-7.

Dayan, F.E., Howell, J.L., Weidenhamer, J.D., 2010. Dynamic root exudation of sorgoleone and its in planta mechanism of action. *Journal of Experimental Botany*, 60:2107-2117. https://doi.org/10.1093/jxb/erp082.

Davis, M.A. 2011. Researching invasive species 50 years after Elton: a cautionary tale. In Fifty

years of invasion ecology: The legacy of Charles Elton. 1st edition. Ed. Richardson DM. Blackwell Publishing Ltd. New Jersey. pp 269-276.

Devoto, M., Medan, D., 2004. Preliminary evidence of landscape-level structure in a population of perennial herb, Cypellaherbertii (Iridaceae). *Ecología Aplicada* 3:188-191.

DiTomaso, J., Kyser, G.B., Pitcairn, M.J., 2006. Yellow starthistle management guide. In: Cal-IPC Publication 2006-03, p. 78. California Invasive Species Council, Berkeley, California.

DiTomaso, J.M., 2000. Invasive weeds in rangelands: species, impacts, and management. *Weed Science* 48(2):255-265. https://doi.org/10.1614/0043-1745(2000)048[0255:IWIRSI]2.0.CO;2.

Dlugosch, K.M., Parker, I.M., 2008. Founding Events in Species Invasions: Genetic Variation, Adaptive Evolution, and the Role of Multiple Introductions. *Molecular Ecology* 17 (1): 431–49. https://doi.org/10.1111/j.1365-294X.2007.03538.x.

Dlugosch, K.M., Lai, Z., Benin, A., Hierro, J.L., Rieseberg, L.H., 2013. Allele identification for transcriptome based population genomics in the invasive plant *Centaurea solstitialis. G3 Genes Genomes Genetics* 3:359–367. https://doi.org/10.1534/g3.112.003871.

Dlugosch, K.M., Cang, F.A., Barker, B.S., Andonian, K., Swope, S.M., Rieseberg, L.H., 2015. Evolution of invasiveness through increased resource use in a vacant niche. *Nature Plants* 1:15066. https://doi.org/10.1038/nplants.2015.66.

Dlugosch, K.M., Anderson, S.R., Braasch, J., Cang, A.F., Gillette, H.D., 2015. The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology* 24: 2095-2111. https://doi.org/10.1111/mec.13183.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., Ovenden, J.R. 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources* 14(1):209-214. https://doi.org/10.1111/1755-0998.12157.

Dobzhansky, T., 1940. Speciation as a stage in evolutionary divergence. *American Naturalist* 74:312–321.

Doležel, J., Sgorbati, S., Lucretti, S., 1992. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum* 85:625

631. https://doi.org/ 10.1111/j.1399-3054.1992.tb04764.x.

Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19:11-15.

Dray, S., Dufour, A.B., 2007. The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical software 22(4): 1-20. http://hdl.handle.net/10.18637/jss.v022.i04.

Eagle, A.J., Eiswerth, M.E., Johnson, W.S., Schoenig, S.E., van Kooten, C.G., 2007. Costs and losses imposed on California ranchers by yellow starthistle. *Rangeland Ecology and Management* 60:369-377.https://doi.org/10.2111/1551-5028(2007)60[369:CALIOC]2.0.CO;2.

Earl, D.A., von Holdt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genet Resources*. https://doi.org/10.1007/s12686-011-9548-7.

Ellstrand, N.C., Schierenbeck, K., 2000. Hybridization as a stimulus for the evolution of invasiveness in plants? *PNAS* 97:7043-7050. https://doi.org/10.1073/pnas.97.13.7043.

Eriksen, R.L., Desronvil, T., Hierro, J.L., Kesseli, R. 2012. Morphological differentiation in a common garden experiment among native and non-native specimens of the invasive weed yellow starthistle (*Centaurea solstitialis*). *Biological Invasions* 7:1459–1467 https://doi.org/10.1007/s10530-012-0172-6.

Eriksen, R.L., Hierro, J.L., Eren, O., Andonian, K., Török, K., Becerra, P.I., Montesinos, D., Khetsuriani, L., Diaconu, A., Kesseli, R., 2014. Dispersal pathways and genetic differentiation among worldwide populations of the invasive *Centaurea solstitialis* L. (Asteraceae). *PLoS One* 9:e114786 https://doi.org/10.1371/journal.pone.0114786.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x.

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164(4):1567-1587.

Filipe, J.C., Jorge, A., Özkan, E., Sotes, G., Hierro, J., Montesinos, D., 2016. Invasive and non-invasive congeneric Centaurea (Asteraceae) show contrasting patterns of herbivory by

snails. *Plant Ecology and Evolution*. 149:228-232. https://doi.org/10.5091/plecevo.2016.1231.

Filipe, J.C., Montesinos, D., 2016. Inter-regional hybrids of native and non-native *Centaurea sulphurea* inherit increased competitive ability from the non-natives. *Plant Ecology and Diversity*, 9(4), 381–386. https://doi.org/10.1080/17550874.2016.1261950.

Foll, M., Gaggiotti, O.E., 2008. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180, 977–993. https://doi.org/10.1534/genetics.108.092221.

Gachomo, E.W., Jimenez-Lopez, J.C., Baptiste, L.J., Kotchoni, S.O., 2014. GIGANTUS1 (GTS1), a member of Transducin/WD40 protein superfamily, controls seed germiantion, growth and biomass accumulation through ribosome-biogenesis protein interactions in *Arabidopsis thaliana*. *BMC Plant Biology* 14:37. https://doi.org/10.1186/1471-2229-14-37.

Garcia, Y., Callaway, R.M., Diaconu, A., Montesinos, D., 2013. Invasive and non-invasive congeners show similar trait shifts between their same native and non-native ranges. *PlosOne*:8(e82281). https://doi.org/10.1371/journal.pone.0082281.

Galbraith, D.W., Harkins, K.R., Maddox, J.M., Ayres, N.M., Sharma, D.P., Firoozabady, E., 1983. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* 220:1049 – 1051 https://doi.org/10.1126/science.220.4601.1049.

Garcia-Jacas, N., Uysal, T., Romashchenko, K., Suárez-Santiago, V.N., Ertuğrul, K., Susanna, A. 2006. *Centaurea* revisited: a molecular survey of the Jaceae group. *Annals of Botany* 98:741-753. doi:10.1093/aob/mcl157.

GBIF Secretariat 2017. *Centaurea solstitialis* L. in GBIF Backbone Taxonomy. Checklist dataset https://doi.org/10.15468/39omei accessed via GBIF.org on 2018-07-18.

Geppert, B., Drozdz, B., Kielczewski, M., Holub, M., 1983. Sesquiterpene lactones. XXIII. Isolation of sesquiterpene lactones from *Centaurea* L. species. *Acta Societatis Botanicorum Poloniae* 52:23-34. https://doi.org/10.5586/asbp.1983.003.

Gerlach, J.D., 1997a. How the west was lost: reconstructing the invasion dynamics of yellow starthistle and other plant invaders of western rangelands and natural areas. *California*

Exotic Pest Plant Council, Symposium Proceedings 3:67–72.

Gerlach, J.D., 1997b. The introduction, dynamics of geographic range expansion and ecosystem effects of yellow starthistle (*Centaurea solstitialis*). *Proceedings of Californian Weed Science Society* 49:236–241.

Gómez-González, S., Cavieres, L.A., Torres Morales, P., Torres-Díaz, C., 2009. Competitive effects of the alien invasive *Centaurea solstitialis* L. on two Chilean *Baccharis* species at different life-cycle stages. *Gayana Botanica* 66:71–83. https://doi.org/10.4067/S0717-66432009000100007.

Göpfert, J.C., Heil, N., Conrad, J., Spring, O., 2005. Cytological development and sesquiterpene lactone secretion in capitate glandular trichomes of sunflower. *Plant Biology* 7:148–155. 10.1055/s-2005-837575.

Graebner, R.C., Callaway, R.M., Montesinos, D., 2012. Invasive species grows faster, competes better, and shows greater evolution toward increased size and growth than exotic non-invasive congeners. *Plant Ecology* 213:545–553. https://doi.org/10.1007/s11258-012-0020-x.

Grice, T., 2005. Weeds of significance to the grazing industries of Australia. Project number COMP.045, Final Report. CRC for Australian Weed Management, University of Adelaide. 114 pp.

Grotkopp, E., Rejmánek, M., Sanderson, M.J., Rost, T.L., 2004. Evolution of genome size in Pines (*Pinus*) and its life-history correlates: supertree analyses. *Evolution* 58:1705–1729. https://doi.org/10.1554/03-545.

Gruntman, M., Zieger, S., Tielbörger, K., 2015. Invasive success and the evolution of enhanced weaponry. *Oikos* 125:59-65. https://doi.org/10.1111/oik.02109.

Gruntman, M., Segev, U., Glauser, G., Tielbörger, K., 2017. Evolution of plant defences along an invasion chronosequence: defence is lost due to enemy release – but not forever. *Journal of Ecology* 105(1):255-264. https://doi.org/10.1111/1365-2745.12660.

Guinochet, M., Foissac, J., 1962. Sur les Caryotypes de quelques espèces du genre *Centaurea* L. et leur signification taxonomique. *Bulletin de la Société Botanique de France* 109:373–389. https://doi.org/10.1080/00378941.1962.10838114.

Gürbüz, I., Özçelik, B., Karaoğlu, T., Yeşilada, E., 2006. Antiviral and antimicrobial activities of three sesquiterpene lactones from *Centaurea solstitialis* L. ssp *solstitialis*. *Planta Medica* 72:1004. https://doi.org/10.1016/j.micres.2007.05.006.

Guo, W., Liu, Y., Lun, N.G.W., Liao, P.C., Huang, B.H., Li, W., Li, C., Shi. X., Huang, Y., 2018. Comparative transcriptome analysis of the invasive weed *Mikania micrantha* with its native congeners provides insights into genetic basis underlying successful invasion BMC Genomics 19:392. https://doi.org/10.1186/s12864-018-4784-9.

Hagenblad, J., Hülskötter, J., Acharya, K.P., Brunet, J., Chabrerie, O., Cousins, S.A.O., Dar, P.A., Diekmann, M., De Frenne, P., Hermy, M., Jamoneau, A., Kolb, A., Lemke, I., Plue, J., Reshi, Z.A., Graae, B.J., 2015. Low genetic diversity despite multiple introductions of the invasive plant species *Impatiens glandulifera* in Europe. *BMC Genetics* 16(103):1-16. 10.1186/s12863-015-0242-8

Hahn, M.A., Buckley, Y.M., Müller-Schärer, H., 2012. Increased population growth rate in invasive polyploid *Centaurea stoebe* in a common garden. *Ecology Letters* 15(9):947–954 https://doi.org/10.1111/j.1461-0248.2012.01813.x.

Hänfling, B., Kollmann, J., 2002. An evolutionary perspective of biological invasions. *Trends in Ecology and Evolution* 17(12):545-546. https://doi.org/10.1016/S0169-5347(02)02644-7.

Hasan, H.A.H., 1998. Studies on toxigenic fungi in roasted food stuff (salted seed) and halotolerant activity of emodin-producing *Aspergillus wentii*. *Folia Microbiologica* 43:383–391. https://doi.org/10.1007/BF02818578.

Hathaway, L., Andersson, S., Prentice, H.C., 2009. Experimental crosses within European *Silene latifolia* (Caryophyllaceae): intraspecific differentiation, distance effects and sex ratio. *Botany* 87:231-240. https://doi.org/10.1139/B08-137.

Hay, G., Facelli, J.M., Panetta, D.F., 2006. Invasive potential and competitive ability of the Eurasian herb *Centaurea solstitialis* L. 15th Australian Weeds Conference: Papers and Proceedings: Managing Weeds in a Changing Climate. In: Adelaide: Weed Management Society of South Australia.

Heywood, V.H., Harborne, J.B., Turner, B.L., 1977. The Biology and Chemistry of the Compositae. Academic Press Inc. New York.

Hierro, J.L., Maron, J.L., Callaway, R.M., 2005. A biogeographical approach to plant invasions: the importance of studying exotics in their introduced and native range. *Journal of Ecology* 93:5-15. 10.1111/j.0022-0477.2004.00953.x.

Hierro, J.L., Villareal, D., Özkan, E., Graham, J.M., Callaway, R.M., 2006. Disturbance facilitates invasion: The effects are stronger abroad than at home. *American Naturalist* 168:144-156. http://dx.doi.org/10.1086/505767.

Hierro, J.L., Özkan, E., Khetsuriani, L., Diaconu, A., Török, K., Montesinos, D., Andonian, K., Kikodze, D., Janoian, L., Villarreal, D., Estanga-Mollica, M.E., Callaway, R.M. 2009. Germination responses of an invasive species in native and non-native ranges. *Oikos* 118:529-538. 10.1111/j.1600-0706.2009.17283.x.

Hierro, J.L., Lortie, C.L., Villarreal, Diego., Estanga-Mollica, M.E., Callaway, R.M., 2011. Resistance to *Centaurea solstitialis* invasion from annual and perennial grasses in California and Argentina. *Biological Invasions* 13:2249-2259. http://dx.doi.org/10.1007/s10530-011-0037-4.

Hierro, J.L., Özkan, E., Villareal, D., Chiuffo, M.C., 2013. Non-native conditions favor non-native populations of invasive plant: demographic consequences of seed size variation? *Oikos* 122:583-590. http://dx.doi.org/10.1111/j.1600-0706.2012.00022.x.

Hierro, J.L., Khetsuriani, L., Andonian, K., Eren, O., Villareal, D., Janoian, G., Janoian, G., Reinhart, K.O., Callaway, R.M., 2017. The importance of factors controlling species abundance and distribution varies in native and non-native ranges. *Ecography* 40:991-1002. http://dx.doi.org/10.1111/ecog.02633.

Hijano, E., Basigalup, D., 1995. El cultivo de la alfalfa en la República Argentina. In E. Hijano and A. Navarro, ed. La Alfalfa en la Argentina. Instituto Nacional de Tecnología Agropecuaria, Buenos Aires, Argentina, 13-18, in Spanish.

HilleRisLambers, J., Ettinger, A.K., Ford, K.R., Haak, D.C., Horwith, M., Miner, B.E., Rogers, H.S., Sheldon, K.S., Tewksbury, J.J., Waters, S.M., Yang, S., 2013. Accidental experiments: ecological and evolutionary insights and opportunities derived from global change. *Oikos*, 122(12), 1649–1661. https://doi.org/10.1111/j.1600-0706.2013.00698.x.

Hodgins, K.A., Bock, D.G., Hahn, M.A., Heredia, S.M., Turner, K.G., Rieseberg, L.H., 2015.

Comparative genomics in the Asteraceae reveals little evidence for parallel evolutionary change in invasive taxa. *Molecular Ecology* 24(9):2226-2240. https://doi.org/10.1111/mec.13026.

Hopkins, R., 2013. Reinforcement in plants. *New Phytologist* 197(4):1095-1103. https://doi.org/10.1111/nph.12119.

Howell, J.T., 1959. Distributional data on weedy thistles in western North America. *Leaflets of Western Botany* 9:17-32.

Inceer, H., Hayirlioglu-Ayaz, S., Ozcan, M., 2007. *Chromosome numbers of the twenty-two Turkish plant species. Caryologia* 60:349-357. https://doi.org/10.1080/00087114.2007.10797958.

Inderjit, Dakshini, K.M.M., 1995. On laboratory bioassays in allelopathy. *Botanical Review* 61:28-44. https://doi.org/10.1007/BF02897150.

Inoue, N., Nishimura, H., Li, H.H., Mizutani, J. 1992. Allelochemicals from *Polygonum sachalinese* Fr, Schm. (Polygonaceae). *Journal of Chemical Ecology* 18:1833–1840. https://doi.org/10.1007/BF02751107.

Irimia, R.E., Montesinos, D., Eren, O., Lortie, C.J., French, K., Cavieres, L.A., Sotes, G.J., Hierro, J.L., Jorge, A., Loureiro, J., 2017. Extensive analysis of native and non-native *Centaurea solstitialis* L. populations across the world shows no traces of polyploidization. *PeerJ* 5:e3531.

Irimia, R.E., Lopes, S.M.M., Sotes, G., Cavieres, L.A., Eren, Ö., Lortie, C.J., French, K., Hierro, J.L., Rosche, C., Callaway, R.M., Pinho e Melo, T.M.V.D., Montesinos, D., 2019. Biogeographic differences in the allelopathy of leaf surface extracts of an invasive weed. Biological Invasions. https://doi.org/10.1007/s10530-019-02038-1.

Irimia, R.E., Hierro, J.L., Branco, S., Sotes, G., Cavieres, L.A., Eren, Ö., Lortie, C.J., French, K., Callaway, R.M., Montesinos, D. (in review). Experimental admixture between neo-allopatric regions of an invasive plant yields a global mosaic of reproductive incompatibilities and heterosis.

ISO 1995. Soil quality—determination of the effects of pollutants on soil flora. Part 2: Effects of chemicals on the emergence of higher plants. International Organization for

Standardization, Genève, Switzerland.

Izhaki, I., 2002. Emodin: A secondary metabolite with multiple ecological functions in higher plants. *New Phytologist* 155:205-217. https://doi.org/10.1046/j.1469-8137.2002.00459.x.

Jasiewicz, A., Mizianty, M., 1975. Chromosome numbers of some Bulgarian plants. *Fragmenta Floristica et Geobotanica* 21:277–288.

Jeschke, J.M., Aparicio, L.G., Haider, S., Heger, T., Lortie, C.J., Pyšek, P., Strayer, D.L., 2012. Support for major hypothesis in invasion biology is uneven and declining. *NeoBiota* 14:1-20. https://doi.org/ 10.3897/neobiota.14.3435.

Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* 11:94. https://doi.org/10.1186/1471-2156-11-94.

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers *Bioinformatics* 24: 1403-1405. doi: 10.1093/bioinformatics/btn129.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.Y., Lopez, R., Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics.* doi:10.1093/bioinformatics/btu031.

Kaczowka, A.M., Lu-Irving, P., Baltrus, D.A., Dlugosch, K.M., 2017. Variation in the oxidative burst in response to wounding and bacterial infection among native and invading genotypes of yellow starthistle (*Centaurea solstitialis*). BioRxiv https://doi.org/10.1101/156083.

Kalusová, V., Chytrý, M., van Kleunen, M., Mucina, L., Dawson, W., Essl, F., Kreft, H., Pergl, J., Weigelt, P., Winter, M., Pyšek, P., 2017. Naturalization of European plants on other continents: The role of donor habitats. *PNAS* 114(52):13756-13761. https://doi.org/10.1073/pnas.1705487114.

Kato, Y., Sakamoto, W., 2018. FtsH protease in the thylakoid membrane: physiological functions and the regulation of protease activity. *Frontiers in Plant Science* 9:855. https://doi.org/10.3389/fpls.2018.00855.

Keenan, K., Mcginnity, P., Cross, T.F., Crozier, W.W., Prodöhl, P.A., 2013. DiveRsity: An R package

for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, 4(8), 782-788. https://doi.org/10.1111/2041-210X.12067.

Kempel, A., Crobock, T., Fischer, M., Rohr, R.P., van Kleunen, M., 2013. Determinants of plant establishment success in a multispecies introduction experiment with native and alien species. *PNAS* 110(31):12727-12732. https://doi.org/10.1073/pnas.1300481110.

Kloot, P.M., 1983. Early records of alien plants naturalized in South Australia. *Journal of Adelaide Botanical Garden* 6(2):93-131.

Kodama, Y., Suetsugu, N., Wada, M., 2011. Novel protein-protein interaction family proteins involved in chloroplast movement response. *Plant signaling and Behavior* 6(4):483-490. https://doi.org/10.4161/psb.6.4.14784.

Kokalis-Burelle, N., Rodríguez-Kábana, R., 2006. Allelochemicals as biopesticides for management of plant parasitic nematodes. In: Inderjit, Mukerji KG (Eds), Allelochemicals: Biological control of plant pathogens and diseases. Springer, Dordrech, The Netherlands, pp 15-29.

Konarzewski, T.K., Murray, B.R., Godfree, R.C., 2012. Rapid development of adaptive, climate driven clinal variation in seed mass in the invasive annual forb Echium plantagineum L. *Plos One* 7(12): e49000. https://doi.org/10.1371/journal.pone.0049000.

Kopelman, N.M., Mayzel, L., Jakobsson, M., Rosenberg, N.A., Mayrose, I., 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* 15:1179-1191.

Kuzmanov, B.A., Jurukova-Grancarova, P.D., Georgieva, S.B., 1990. Chromosome numbers of Bulgarian angiosperms. *Fitologiya* (Sofia) 38:92.

Kueffer, C., Pyšek, P., Richardson, D.M., 2013. Integrative invasion science: model systems, multi-site studies, focused meta-analysis and invasion syndromes. *New Phytologist* 200(3):615-633. https://doi.org/10.1111/nph.12415.

Kyser, G.B., Hazebrook, A., DiTomaso, J.M., 2013. Integration of prescribed burning, aminopyralid and reseeding for restoration of yellow starthistle (*Centaurea solstitialis*) infested rangeland. *Invasive Plant Science and Management* 6:480-491. http://dx.doi.org/10.1614/IPSM-D-12-00094.1.

Laird, N.N.M., Ware, J.J.H., 1982. Random-effects models for longitudinal data. *Biometrics* 38:963–974.

Lai, Y.T., Yeung, C.K.L., Omland, K.E., Pang, E.L., Hao, Y., Liao, B.Y., Cao, H.F., Zhang, B.W., et al 2019. Standing genetic variation as the predominant source for adaptation of a songbird. *PNAS* 116(6):2152-2157. https://doi.org/10.1073/pnas.1813597116.

Lankau, R.A., 2012. Coevolution between invasive and native plant driven by chemical competition and soil biota. *PNAS* 109:11240-11245. https://doi.org/10.1073/pnas.1201343109.

Lankau, R.A., Nuzzo, V., Spyreas, G., Davis, A.S., 2009. Evolutionary limits ameliorate the negative impact of an invasive plant. *PNAS* 106(36): 15362-15367. https://doi.org/10.1073/pnas.0905446106.

Lavergne, S., Muenke, N.J., Molofsky, J., 2010. Genome size reduction can trigger rapid phenotypic evolution in invasive plants. *Annals of Botany* 105:109–116. https://doi.org/ 10.1093/aob/mcp271.

Lavergne, S., Molofsky, J., 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. *PNAS* 104(10):3883-3888. https://doi.org/10.1073/pnas.0607324104.

Lee, H.J., Mochizuki, N., Masuda, T., Buckhout, T.J., 2012. Disrupting the bimolecular binding of the haem-binding protein 5(AtHBP5) to haem oxygenase 1 (HY1) leads to oxidative stress in *Arabidopsis* 63(16):5967-5978. https://doi.org/10.1093/jxb/ers242.

Leinonen, T., Cano, J.M., Mäkinen, H., Merilä, J., 2006. Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of three spine sticklebacks. Journal of Evolutionary Biology 19, 1803-1812. https://doi.org/10.1111/j.1420-9101.2006.01182.x.

Leinonen, T., Scott McCairns, R.J., O´Hara, R.B., Merilä, J., 2013. Q_{ST}-F_{ST} comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nature Reviews Genetics* 14(3):179-190. https://doi.org/10.1038/nrg3395.

Lejeune, K.D., Seastedt, T.R., 2002. *Centaurea* species: the forb that won the west. *Conservation Biology* 15(6):1568:1574. https://doi.org/10.1046/j.1523-1739.2001.00242.x.

Leong, M., Kremen, C., Roderick, G.K., 2014. Pollinator interactions with yellow starthistle (*Centaurea solstitialis*) across urban, agricultural and natural landscapes. *PLoSONE* 9(1): e86357. https://doi.org/10.1371/journal.pone.0086357.

Le Roux, J.J., Brown, G.K., Byrne, M., Ndlovu, J., Richardson, D.M., Thompson, G.D., et al. 2011. Phylogeographic consequences of different introduction histories of invasive Australian Acacia species and *Paraserianthes lophantha* (Fabaceae) in South Africa. *Diversity and Distributions* 17:861–71. https://doi.org/10.1111/j.1472-4642.2011.00784.x.

Levene, H., 1960. Robust testes for equality of variances. In: Olkin I (ed) Contributions to probability and statistics. Stanford University Press, Palo Alto, pp 278–292.

Li, Y., Stift, M., van Kleunene, M., 2018. Admixture increases performance of an invasive plant beyond first-generation heterosis. *Journal of Ecology* 106:1595-1606. https://doi.org/10.1111/1365-2745.12926.

Lloyd, D.G., Schoen, D.J., 1992. Self-fertilization in plants. 1. Functional dimensions. International *Journal of Plant Science* 153:358-369.

Locken, L.J., 1985. Cnicin concentrations in spotted knapweed (*Centaurea maculosa* Lam.) and associated soils. MSc Thesis. University of Montana.

Locken, L.J., Kelsey, R.G., 1987. Cnicin concentrations in *Centaurea maculosa*, spotted knapweed. *Biochemical Systematics and Ecology* 15(3):313-320. https://doi.org/10.1016/0305-1978(87)90005-6.

Lombaert, E., Guillemaud, T., Cornuet, J.M., Malausa, T., Facon, B., Estoup, A., 2010. Bridgehead effect in the worldwide invasion of the biocontrol Harlequin ladybird. *PlosOne* 5(3): e9743. https://doi.org/10.1371/journal.pone.0009743.

Loureiro, J., Rodriguez, E., Doležel, J., Santos, C., 2007. Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. *Annals of Botany* 4:875–488. https://doi.org/10.1093/aob/mcm152.

Lowry, D.B., Cotton, Rockwood, R., Willis, J.H., 2008. Ecological reproductive isolation of coast and inland races of *Mimulus guttatus*. *Evolution* 62:2196-2214. https://doi.org/10.1111/j.1558-5646.2008.00457.x.

Lowry, E., Rollinson, E.J., Laybourn, A.J., Scott, T.E., Aiello-Lammens, M.E., Gray, SM., Mickley, J.,

Gurevitch, J., 2013. Biological invasions: a field synopsis, systematic review and database of the literature. *Ecology and Evolution* 3(1):182-196. 10.1002/ece3.431.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS – A Bayesian modelling framework: Concepts, structure and extensibility. *Statistics and Computing* 10(4):325-337. https://doi.org/10.1023/A:1008929526011.

Lu-Irving, P., Harenčár, J.G., Sounart, H., Welles, S.R., Swope, S.M., Baltrus, D.A., Dlugosch, K.M., 2019. Native and invading yellow starthistle (*Centaurea solstitialis*) microbiomes differ in composition and diversity of bacteria. *mSphere* 4:e00088-19. https://doi.org/10.1128/mSphere.00088-19.

Luu, K., Bazin, E., Blum, M.G.B., 2017. Pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources* 17:67-77.

Maddox, D.M., Joley, D.B., Supkoff, D.M., Mayfield, A., 1996. Pollination biology of yellow starthistle (*Centaurea solstitialis*) in California. *Canadian Journal of Botany* 74(2):262-267. https://doi.org/10.1139/b96-031.

Maddox, D.M., Mayfield, A., Poritz, N.H., 1985. Distribution of yellow starthistle (*Centaurea solstitialis*) and Russian Knapweed (*Centaurea repens*). *Weed Science* 33:315-327.

Maddox, D.M., 1981. Introduction, phenology and density of yellow starthistle in coastal, intercoastal and central valley situations in California. Albany, CA:USDA ARS ARR-W-20. Pp 33.

Madden, T., 2002. The BLAST Sequence Analysis Tool. Oct 9 [Updated 2003 Aug 13]. In: McEntyre J, Ostell J, editors. The NCBI Handbook [Internet]. Bethesda (MD): National Center for Biotechnology Information (US). Available from: http://www.ncbi.nlm.nih.gov/books/NBK21097/.

Maron, J.L., Vilà, M., Bommarco, R., Elmendorf, S., Beardsley, P., 2004. Rapid evolution of an invasive plant. *Ecological Monographs*, 74(2), 261–280. https://doi.org/10.1890/03-4027.

Marrs, R.A., Sforza, R., Hufbauer, R.A., 2008. Evidence for multiple introductions of *Centaurea stoebe micranthos* (spotted knapweed, Asteraceae) to North America. *Molecular Ecology* 17(19):4197-4208. https://doi.org/10.1111/j.1365-294X.2008.03903.x.

Martinez, M.A., Baack, E.J., Hovich, S.M., Whitney, K.D., 2018. A reassessment of the genome size

invasiveness relationship in reed canarygrass (*Phalaris arundinaceae*). *Annals of Botany* 121(7):1309-1318. https://doi.org/10.1093/aob/mcy028.

Matesanz, S., Horgan-Kobelski, T., Sultan, S.E., 2012. Phenotypic plasticity and population differentiation in an ongoing species invasion. *PLoS One* 7(9):e44955. https://doi.org/10.1371/journal.pone.0044955.

Mayr, E., 1947. Ecological factors in speciation. Evolution 1:263:288.

Meepagala, K.M., Osbrink, W., Sturtz, G., Lax, A., 2006. Plant derived natural products exhibiting activity against Formosan subterranean termites (*Coptotermes formosanus*). *Pest Management Science* 62:565-570. https://doi.org/10.1002/ps.1214.

Miguel, M.F., Lortie, C.J., Callaway, R.M., Hierro, J.L., 2017. Competition does not come at the expense of colonization in seed morphs with increased size and dispersal. American Journal of Botany 104(9):1323-1333. https://doi.org/10.3732/ajb.1700266.

Miskella, J., 2014. Hybridization between yellow starthistle (*Centaurea solstitialis*) and meadow knapweed (*Centaurea* x *moncktonii*). Master´s Thesis. Oregon State University. Retrieved from ScholarsArchive@OSU.

Mitchell, C.E., Agrawal, A.A., Bever, J.D., Gilbert, G.S., Hufbauer, R.A., Klironomos, J.N., Maron, J.L., Morris, W.l., Parker, I.M., Power, A.G., Seabloom, E.W., Torchin, M.E., Vásquez, D.P., 2006. Biotic interactions and plant invasions. *Ecology Letters* 9(6):726-740. https://doi.org/10.1111/j.1461-0248.2006.00908.x.

Montesinos, D., Santiago, G., Callaway, R.M., 2012. Neo-Allopatry and rapid reproductive isolation. *American Naturalist* 180(4):529-533. https://doi.org/10.1086/667585.

Montesinos, D., Callaway, R.M., 2017. Inter-regional hybrids of native and invasive *Centaurea solstitialis* display intermediate competititve ability. *Ecography* 40(801:802). https://doi.org/10.1111/ecog.02653.

Montesinos, D., Callaway, R.M., 2018. Traits correlate with invasive success more than plasticity: A comparison of three *Centaurea* congeners. *Ecology and Evolution* 8:7378-7385. 10.1002/ece3.4080.

Montesinos, D., Graebner, R.C., Callaway, R.M., 2019. Evidence for evolution of increased competitive ability for invasive *Centaurea solstitialis*, but not for naturalized *C*.

calcitrapa. *Biological Invasions* 21:99-110. https://doi.org/10.1007/s10530-018-1807-z.

Montenegro, G., Teillier, S., Arce, P., Poblete, V., 1991. Introduction of plants into the Medierranean-type climate area of Chile. In: Groves RH, Di Castri F (eds). Biogeography of Mediterranean invasions. Cambridge University Press, Cambridge, pp 103–114.

Moret, S., Populin, T., Conte, L.S., Cosens, G., 2005. HPLC determination of free nitrogenous compounds of *Centaurea solstitialis* (Asteaceae), the cause of equine nigropallidal encephalomalacia.*Toxicon* 46:651-657. http://dx.doi.org/10.1016/j.toxicon.2005.07.006.

Morjan, C.L., Rieseberg, L.H., 2004. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology* 13(6):1341-1356. https://doi.org/10.1111/j.1365-294X.2004.02164.x.

Mráz, P., Garcia-Jacas, N., Gex-Fabry, E., Susanna, A., Barres, L., Müller-Schärer, H., 2011. Allopolyploid origin of highly invasive *Centaurea stoebe* s.l. (Asteraceae). *Molecular Phylogenetics and Evolution* 62:612–623. https://doi.org/10.1016/j.ympev.2011.11.006.

Muir, A., Majak, W., 1983. Allelopathic potential of diffuse knapweed (*Centaurea diffusa*) extracts. Canadian Journal of Plant Science, 63:989-996. https://doi.org/10.4141/cjps83-124.

Müller-Schärer, H., Schaffner, U., Steinger, T., 2004. Evolution in invasive plants: implications for biological control. *Trends in Ecology and Evolution* 19:417–422. https://doi.org/: 10.1016/j.tree.2004.05.010.

Ni, G., Schaffner, U., Peng, S., Callaway, R.M., 2010. *Acroptilon repens*, an Asian invader, has stronger competitive effects on species from America than species from its native range. *Biological Invasions* 12:3653-3663. https://doi.org/10.1007/s10530-010-9759-y.

Niu, X.M., Xu, Y.C., Li, Z.W., Bian, Y.T., Hou, X.H., Chen, J.F., Zou, Y.P., Jiang, J., Wu, Q., Ge, S., Balasubramanian, S., Guo, Y.L., 2018. Transposable elements drive rapid phenotypic variation in *Capsella rubella*. *PNAS* 116(14):6908-6913. https://doi.org/10.1073/pnas.1811498116.

North, A., Pennanen, J., Ovaskainen, O., Laine, A.L., 2011. Local adaptation in a changing world: the roles of gene-flow, mutation and sexual reproduction. *Evolution* 65:79-89. https://doi.org/10.1111/j.1558-5646.2010.01107.x.

OECD 2003. Organisation for Economic Cooperation and Development: OECD guideline 208. Terrestrial plant test: 208; Seedling emergence and seedling growth test.

Oduor, A.M.O., Leimu, R., van Kleunen, M., 2016. Invasive plant species are locally adapted just as frequently and at least as strongly as native plant species. *Journal of Ecology* 104:957-968. https://doi.org/10.1111/1365-2745.12578.

O´Hara, R.B., Merilä, J., 2005. Bias and precision in QST: Problems and some solutions. *Genetics* 171:1331-1339. https://doi.org/10.1534/genetics.105.044545.

Oksanen, J., Blanchet, G.F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H., Szoecs, E., Wagner, H., 2018. vegan: Community Ecology Package. R package version 2.5-1. https://CRAN.R-project.org/package=vegan.

Oster, M., Beck, J.J., Furrow, R.E., Yeung, K., Field, C.B., 2015. In-field yellow star thistle (*Centaurea solstitialis*) volatile composition under elevated temperature and CO_2 and implications for future control. *Chemoecology* 25:313-323. https://doi.org/10.1007/s00049-015-0200-y.

Özçelik, B., Gürbüz, I., Karaoğlu, T., Yeşilada, E., 2009. Antiviral and antimicrobial activities of three sesquiterpene lactones from *Centaurea solstitialis* L ssp. *solstitialis*. *Microbiological Research* 164:545-552. https://doi.org/10.1016/j.micres.2007.05.006.

Parsons, W.T., Cuthbertson, E.G., 2001. Noxious weed of Australia. 2nd Edition. Csiro Publishing, Collingwood Australia, 672 pp.

Pearson, D.E., Ortega, Y.K., Eren, Ö., Hierro, J.L., 2018. Community assembly theory as a framework for biological invasions. *Trends in Ecology and Evolution* 33(5). https://doi.org/10.1016/j.tree.2018.03.002.

Pellet, F.C., 1976. American honey plants. Dadant and Sons. Hamilton. Illinois. USA.

Petanidou, T., Godfree, R.C., Song, D.S., Kantsa, A., Dupont, Y.L., Waser, N.M., 2012. Self-compatibility and plant invasiveness: comparing species in native and invasive ranges. *Perspectives in Plant Ecology, Evolution and Systematics* 14:3-12. https://doi.org/10.1016/j.ppees.2011.08.003.

Pigliucci, M., 2001. Phenotypic plasticity: beyond nature and nurture. Baltimore, MD. The

Johns Hopkins University Press.

Picman, A.K., 1987. Persistence and disappearance of the sesquiterpene lactone, isoalantolactone, in soils. *Biochemical Systematics and Ecology* 15(3):361-363. https://doi.org/10.1016/0305-1978(87)90012-3.

Pitcairn, M., Schoenig, S., Yacoub, R., Gendron, J., 2006. Yellow starthistle continues its spread in California. California Agriculture, 60(2):83–90. https ://doi.org/10.3733/ca.v060n 02p83.

Prentis, P.J., Sigg, D.P., Raghu, S., Dhileepan, K., Pavasovic, A., Lowe, A.J., 2009. Understanding invasion history: genetic structure and diversity of two globally invasive plants and implications for their management. *Diversity and Distributions* 15: 822–830. https://doi.org/10.1111/j.1472-4642.2009.00592.x.

Prentis, P.J., Wilson, J.R.U., Dormontt, E.E., Richardson, D.M., Lowe, A.J., 2008. Adaptive evolution in invasive species. *Trends in Plant Science* 13(6):288-294. https://doi.org/10.1016/j.tplants.2008.03.004.

Prentis, P.J., Woolfit, M., Thomas-Hall, S.R., Ortiz-Barrientos, D., Pavasovic, A., Lowe, A.J,, Schenk, P.M., 2010. Massively parallel sequencing and analysis of expressed sequence tags in a successful invasive plant. *Annals of Botany* 106(6):1009-1017. https://doi.org/10.1093/aob/mcq201.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155(2):945-959.

Pujol, B., Wilson, A.J., Ross, I.C., Pannell, J.R., 2008. Are Q_{ST}-F_{ST} comparisons for natural populations meaningful? *Molecular Ecology* 17:4782-4785. https://doi.org/10.1111/j.1365-294X.2008.03958.x.

Pyšek, P., Skálová, H., Čuda, J., Guo, W.Y., Suda, J., Doležal, J., Kauzál, O., Lambertini, C., Lučanová, M, Mandáková, T., Moravcová, L., Pyšková, K., Brix, H., Meyerson, L.A., 2018. Small genome separates native and invasive populations in an ecologically cosmopolitan grass. *Ecology* 99(1):79:90. https://doi.org/10.1002/ecy.2068.

Qin, B., Lau, J.A., Kopshever, J., Callaway, R.M., McGray, H., Perry, L.G., Weir, T.L., Paschke, M.W., Hierro, J.L., Vivanco, J.M., Strauss, S., 2007. No evidence for root-mediated allelopathy in

Centaurea solstitialis, a species in a commonly allelopathic genus. *Biological Invasions* 9:897–907. https://doi.org/10. 1007/s10530-007-9089-x.

Quintana, N., El Kassis, E.G., Stermitz, F.R., Vivanco, J.M., 2009. Phytotoxic compounds from roots of *Centaurea diffusa* Lam. *Plant Signaling and Behavior* 4: 9-14. https://doi.org/ 10.4161/psb.4.1.7487

Quintana, N., Weir, T.L., Du, J., Broeckling, C.D., Rieder, J.P., Stermitz, F.R., Paschke, M.W., Vivanco, J.M., 2008. Phytotoxic polyacetylenes from roots of Russian knapweed *Acroptilon repens* (L.) DC). *Phytochemistry* 69:2572-2578. https://doi.org/10.1016/j.phytochem.2008.07.015.

R Core Team 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ranal, M.A., Garcia de Santana, D., Resende Ferreira, W., Mendes-Rodrigues, C., 2009. Calculating germination measurements and organizing spreadsheets. *Revista Brasileira de Botânica*, 32:849-855. https://doi.org/10.1590/S0100-84042009000400022.

Ramsey, J., Bradshaw, H.D., Schemske, D.W., 2003. Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M*. *cardinalis* (Phrymaceae). *Evolution*, 57: 1520-1534. DOI: https://doi.org/10.1111/j.0014-3820.2003.tb00360.x.

Revelle, W., 2018. psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.8.12.

Rice, A., Glick, L., Abadi, S., Einhorn, M., Kopelman, N.M., Salman-Minkov, A., Mayzel, J., Chay, O., Mayrose, I., 2015. The Chromosome Counts Database (CCDB) – a community resource of plant chromosome numbers. *New Phytologist* 206:19–26 https://doi.org/10.1111/nhp.13191.

Ridenour, W.M., Callaway, R.M., 2001. The relative importance of allelopathy in interference: the effects of an invasive weed on a native bunchgrass. *Oecologia*, 126:444-450. https://doi.org/10.1007/s004420000533.

Riopelle, R.J., Stevens, K.L., 1993. In vitro neurotoxicity bioassay: neurotoxicity of sesquiterpene lactones. In Bioactive Natural Products – Detection, Isolation and Structural

Determination. Colegate SM, Molyneaux RJ, eds. CRC Press. Boca Raton, FL, pp 457-463.

Roché, B.F., 1991. Achene dispersal in yellow starthistle (*Centaurea solstitialis* L). *Northwest Science* 66(2):62-65.

Roché, C.T., Thill DC., Shafii, B., 1997. Reproductive phenology in yellow starthistle (*Centaurea solstitialis*). *Weed Science* 45:763-770. https://doi.org/10.1017/S0043174500088949.

Roché, C.T., Susanna, A., 2011. New habitats, new menaces: *Centaurea* x *kleinii* (*C. moncktonii* x *C. solstitialis*), a new hybrid species between two alien weeds. *Collectanea Botanica*, 29, 17–23.

Roché, B.F., Roche, C.T., Chapman, R.C., 1994. Impacts of grassland habitat on yellow starthistle (*Centaurea solstitialis* L.) invasion. *Northwest Science*. 68(2): 86-96.

Rosche, C., Hensen, I., Lachmuth, S., 2018. Local pre-adaptation to disturbance and inbreeding environment interactions affect colonisation abilities of diploid and tetraploid *Centaurea stoebe. Plant Biology* 20:75-84. https://doi.org/10.1111/plb.12628.

Rosche, C., Hensen, I., Schaar, A., Zehra, U., Jasieniuk, M., Callaway, R.M., Khasa, D.P., Al-Gharaibeh, M.A., Lekberg, Y., Nagy, D.U., 2019. Climate outweights native vs. non-native range-effects for genetics and common garden performance of a cosmopolitan weed. *Ecological Monographs*. https://doi.org/10.1002/ecm.1386.

Rosinski, G., Bloszyk, E., Harmatha, J., Knapik, A., 1988. Changes produced by sesquiterpenes lactones in some physiological processes in *Tenebrio molitor* L. In F. Sehnal, A. Zabza, and D.L. Denlinger (Eds.), *Endocrinological Frontiers in Physiological Insect Ecology* (pp. 91-94). Wroclaw Technical University Press, Wroclaw.

Sampaio, B.L., Edrada-Ebel, R.A., Batista Da Costa, F., 2016. Effect of the environment on the secondary metabolic profile of *Tithonia diversifolia*: a model for environmental metabolomics of plants. *Scientific Reports* 6: 29265. https://doi.org/10.1038/srep29265.

Sapir, Y., Mazzucco, R., 2015. Post-zygotic reproductive isolation among populations of *Iris atropurpurea*: The effect of spatial distance among crosses and the role of inbreeding and outbreeding depression in determining niche width. Interim Report IR-12-064. International Institute for Applied Systems Analysis, Austria.

Sax, D.F., Stachowicz, J.J., Brown, J.H., Bruno, J.F., Dawson, M.N., Gaines, S.G., Grosberg, R.K., Hastings, A., Holt, R.D., Mayfield, M.M., O´Connor, M.I., Rice, W.R., 2007. Ecological and evolutionary insights from species invasions. *Trends in Ecology and Evolution* 22(9). https://doi.org/10.1016/j.tree.2007.06.009.

Schmidt, J.P., Drake, J.M., 2011. Why are some plant genera more invasive than others? *Plos One* 6(4): e18654. https://doi.org/10.1371/journal.pone.0018654.

Schemske, D.W., 2010. Adaptation and the origin of species. American Naturalist, 176:S4–S25. https://doi.org/10.1086/657060.

Schleuning, M., Templin, M., Huamán, V., Vadillo, G.P., Becker, T., Durka, W., Fischer, M., Matthies, D., 2011. Effects of Inbreeding, Outbreeding and supplemental pollen on the reproduction of a hummingbird-pollinated clonal Amazonian herb. *Biotropica* 43:183-191. https://doi.org/10.1111/j.1744-7429.2010.00663.x.

Scursoni, J.A., Gigón, R., Martín, A.N., Vigna, M., Leguizamón, E.S., Istilart, C., López, R., 2017. Changes in weed communities of spring wheat crops of Buenos Aires province of Argentina. *Weed Science* 61(1):51-62. https://doi.org/10.1614/WS-D-12-00141.1.

Seebens, H., Blackburn, T.M., Dyer, E.E., Genovesi, P., Hulme, P.E., Jeschke, J.M., Pagad, S., Pyšek, P., Winter, M., Arianoutsou, M., Bacher, S., 2017. No saturation in the accumulation of alien species worldwide. *Nature Communication* 8(14435). https://doi.org/10.1038/ncomms14435.

Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52:591-611. https://doi.org/10.2307/2333709.

Sheley, R.L, Larson, L.L., Jacobs, J.S., 1999. Yellow starthistle. In: Sheley, RL.; Petroff, JK., eds. Biology and management of noxious rangeland weeds. Corvallis, OR: Oregon State University Press: 408-416.

Shi, J., Macel, M., Tielbörger, K., Verhoeven, K.J.F. 2018. Effects of admixture in native and invasive populations of *Lythrum salicaria*. *Biological Invasions* 20(9):2381-2393. https://doi.org/10.1007/s10530-018-1707-2.

Skinner, K., Smith, L., Rice, P., 2000. Using noxious weed lists to prioritize targets for developing weed management strategies. *Weed Science* 48: 640-644.

https://doi.org/10.1614/0043-1745(2000)048[0640:UNWLTP]2.0.CO;2.

Smith, L., Beck, J.J., 2013. Effect of mechanical damage on emission of volatile organic compounds from plant leaves and implications for evaluation of host plant specificity of prospective biological control agents of weeds. *Biocontrol Science and Technology*, 23:880–907. https://doi.org/10.1080/09583157.2013.807908.

Sotes, G.J., Cavieres, L.A., Montesinos, D., Pereira, Coutinho, A.X., Pelaez, W.J., Lopes, S.M., Pinho e Melo, T.M.M.V.D., 2015. Inter-regional variation on leaf surface defenses in native and non-native *Centaurea solstitialis* plants. *Biochemical Systematics and Ecology*, 62:208–218. https://doi.org/10.1016/j.bse.2015.09.003.

Soltis, D.E., Albert, V.A., Leebens-Mack, J., Bell, C.D., Paterson, A.H., Zheng, C., Sankoff, D., dePamphilis, C.W., Wall, P.K., Soltis, P.S., 2009. Polyploidy and angiosperm diversification. *American Journal of Botany* 96(1):336–348. https://doi.org/10.3732/ajb.0800079.

Spitzer, M., Wildenhain, J., Rappsilber, J., Tyres, M., 2014. BoxPlotR: a web tool for generation of box plots. *Nature Methods* **11**:121–122 https://doi.org/10.1038/nmeth.2811.

Spitze, K., 1993. Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics* 135(2):367-374.

Stevens, K.L., Merrill, G.B., 1985. Sesquiterpene lactones and allelochemicals from *Centaurea* species. In A.C. Thompson (Eds.), ACS Symposium Series, *The Chemistry of Allelopathy.* American Chemical Society: Washington DC.

Stevens, K.L., Riopelle, R.J., Wong, R.Y., 1990. Repin, a sesquiterpene lactone from *Acroptilon repens* possessing exceptional biological activity. *Journal of Natural Products*, 53:218-221. https://doi.org/10.1021/np50067a038.

Storey, J.D., Bass, A.J., Dabney, A., Robinson, D., 2019. *qvalue: Q-value estimation for false discovery rate control*. R package version 2.16.0, http://github.com/jdstorey/qvalue.

Stotz, G.C., Gianoli, E., Chill, J.F., 2019. Biotic homogenization within and across eight widely distributed grasslands following invasion by *Bromus inermis*. *Ecology* 100(7):e02717. https://doi.org/10.1002/ecy.2717.

Sun, M., Ritland, K., 1998. Mating system of yellow starthistle (*Centaurea solstitialis*) a successful colonizer in North America. *Heredity* 80:225-232.

Sun, M., 1997. Population genetic structure of yellow starthistle (*Centaurea solstitialis*), a colonizing weed in the western United States. *Canadian Journal of Botany* 75:1470-1478. https://doi.org/10.1139/b97-861.

Suda, J., Meyerson, L.A., Leitch, I.J., Pyšek, P., 2015. The hidden side of plant invasions: the role of genome size. *New Phytologist* 205: 994–1007. https://doi.org/10.1111/nph.13107.

Swope, S.M., Parker, I.M., 2010. Widespread seed limitation affects plant density but not population trajectory in the invasive plant *Centaurea solstitialis*. *Oecologia* 164(1)-117:128. https://doi.org/10.1007/s00442-010-1641-1.

TAIR The Arabidopsis Information Resource. Available at https://www.arabidopsis.org/. Accessed on Aug 20, 2019.

te Beest, M., Le Roux, J.J., Richardson, D.M., Brysting, A.K., Suda, J., Kubešová, M., Pyšek, P., 2011. The more the better? The role of polyploidy in facilitating plant invasions. *Annals of Botany* 109:19–45. https://doi.org/10.1093/aob/mcr277.

Tharayil, N., Bhowmik, P., Alpert, P., Walker, E., Amarasiriwardena, D., Xing, B., 2009. Dual purpose secondary compounds: phytotoxin of *Centaurea diffusa* also facilitates nutrient uptake. *New Phytologist* 181:424-434. https://doi.org/10.1111/j.14698137.2008.02647.x.

The National Environmental Management: Biodiversity Act (10/2004) NEMBA. Draft Alien and Invasive Species Lists (2016). Department of Environmental Affairs, Government Gazette 82 (40166).

Thelen, G.C., Vivanco, J.M., Newingham, B., Good, W., Bais, H.P., Landres, P., Caesar, A., Callaway, R.M., 2005. Insect herbivory stimulates allelopathic exudation by an invasive plant and the suppression of natives. *Ecology Letters* 8:209-217. https://doi.org/10.1111/j.1461-0248.2004.00713.x.

Turner, K.G., Nurkowski, K.A., Rieseberg, L.H., 2016. Gene expression and drought response in an invasive thistle. *Biological invasions* 19(3):1-19. https://doi.org/10.1007/s10530-016-1308-x.

USDA, NRCS. The PLANTS Database. 2017. National Plant Data Team, Greensboro. http://plants.usda.gov. Accessed 10 January 2017.

U.S. EPA, 1996. Ecological effects test guidelines (OPPTS 850.4200): Seed germination/root elongation toxicity test.

Uygur, S., Smith, L., Uygur, F.N., Cristofaro, M., Balciunas, J., 2004. Population densities of yellow starthistle (*Centaurea solstitialis*) in Turkey. *Weed Science* 52: 746–753. https://doi.org/10.1614/WS-03-150R1.

van Kleunen, M., Dawson, W., Essl, F., Pergl, J., Winter, M., Weber, E., Kreft, H., Weigelt, P., Kartesz, J., Nishino, M., 2015. Global exchange and accumulation of non-native plants. *Nature* 525:100–103. https://doi.org/10.1038/nature14910.

van Kleunen, M., Röckle, M., Stift, M., 2015. Admixture between native and invasive populations may increase invasiveness of *Mimulus guttatus*. *Proceedings of the Royal Society B* 282:20151487. https://doi.org/10.1098/rspb.2015.1487.

Vandepitte, K., de Meyer, T., Helsen, K., van Acker, K., Roldán-Ruiz, I., Mergeay, J., Honnay, O., 2014. Rapid genetic adaptation precedes the spread of an exotic plant species. *Molecular Ecology* 23(9):2157-2164. https://doi.org/10.1111/mec.12683.

Vellend, M., Harmon, L.J., Lockwood, J.L., Mayfield, M.M., Hughes, A.R., Wares, J.P., Sax, D.F., 2007. Effects of exotic species on evolutionary diversification. *Trends in Ecology and Evolution* 22(9):481-488. https://doi.org 10.1016/j.tree.2007.02.017.

Verrell, P.A., Arnold, S.J., 1989.Behavioral observations of sexual isolation among allopatric populations of the mountain dusky salamander, *Desmognathus ochrophaeus*. *Evolution* 43(4):745-755.

Verhoeven, K.J., Macel, M., Wolfe, L.M., Biere, A., 2011. Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. *Proceedings of the Royal Society B-Biological Sciences* 278(1702):2-8. http://dx.doi.org/10.1098/rspb.2010.1272.

Via, S., 2009. Natural selection in action during speciation. *PNAS* 106(1):9939-9946. https://doi.org/10.1073/pnas.0901397106.

Vilá, M., Espinar, J.L., Hejda, M., Hulme, P.E., Jarošik, V., Pergl, J., Schaffner, U., Sun, Y., Pyšek, 2011. Ecological impacts of invasive alien plants: a meta-analysis of their effects on

species, communities and ecosystems *Ecology Letters* 14:702–708. https://doi.org/10.1111/j.1461-0248.2011.01628.x.

Vrchotová, N., Šerá, B., Krejčová, 2011. Allelopathic activity of extracts from *Impatiens* species. *Plant*, *Soil and Environment* 57:57-60. https://doi.org/10.17221/156/2010-PSE.

Vu, V.Q., 2011 ggbiplot: a ggplot2 based biplot. R package version 0.55. http://github.com/vqv/ggbiplot. Accessed 10 Jun 2019.

Wagenitz, C., 1955. Pollenmorphologie and Systematik in der gattung *Centaurea* L. s.l. *Flora* 142:213–275, in German.

Waller, L.P., Callaway, R.M., Klironomos, J.N., Ortega, Y.K., 2016. Reduced mycorrhizal responsiveness leads to increased competitive tolerance in an invasive exotic plant. *Journal of Ecology*. 104:1599:1607. https://doi.org/10.1111/1365-2745.12641.

Waser, N.M., Price, M.V., 1989. Optimal outcrossing in *Ipomopsis aggregata*: seed set and offspring fitness. *Evolution* 43(5):1097-1109.

Waser, N.M., Williams, C.F., 2001. Inbreeding and Outbreeding. In Evolutionary Ecology. Concepts and Case Studies (C.W. Fox, D.A. Roff and D.J. Fairbaine, eds), pp. 84-96. Oxford: Oxford University Press.

Weir, T.L., Park, S.W., Vivanco, J.M., 2004. Biochemical and physiological mechanisms mediated by allelochemicals. Current Opinion in Plant Biology 7:472-479. https://doi.org/: 10.1016/j.pbi.2004.05.007.

Widmer, T.L., Guermache, F., Dolgovskaia, M.Y., Reznik, S., 2007. Enhanced growth and seed properties in introduced vs. native populations of yellow starthistle (*Centaurea solstitialis*). *Weed Science* 55: 465-473. https://doi.org/10.1614/WS-06-211R.1.

Widmer, A., Lexer, C., Cozzolino, S., 2008. Evolution of reproductive isolation in plants. *Heredity* 102:31-38. https://doi.org/10.1038/hdy.2008.69.

Wilson, J.R.U., García-Díaz, P., Cassey, P., Richardson, D.M., Pyšek, P., Blackburn, T.M., 2016. Biological invasions and natural colonisations are different – the need for invasion science. *NeoBiota* 31:87-98. https://doi.org/10.3897/neobiota.31.9185.

Williamson, M., Fitter, A., 1996. The varying success of invaders. *Ecology* 77(6):1661-1666. dx.doi.org /10.2307/2265769.

Whitlock, M.C., Lotterhos, K.E., 2015. Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model through Trimming the Distribution of FST. *American Naturalist*. 186(S1):S24-S36.

Wolfe, L.M., Blair, A.C., Penna, B.M., 2007. Does intraspecific hybridization contribute to the evolution of invasiveness? an experimental test. *Biological Invasions* 9(5):515-521. https://doi.org/10.1007/s10530-006-9046-0.

Wright, S., 1951. The genetical structure of populations. *Annals of Human Genetics* 15:323–354. https://doi.org/10.1111/j.1469-1809.1949.tb02451.x.

Yeşilada, E., Gürbüz, I., Shibata, H., 1999. Screening of Turkish anti-ulcerogenic folk remedies for anti-*Helicobacter pylori* activity. *Journal of Ethnopharmacology* 66:289-293.

Young, S.L., Kyser, G.B., Barney, J.N., Claassen, V.P., Ditomaso, J.M., 2011. The role of light and soil moisture in plant community resistance to invasion by yellow starthistle (*Centaurea solstitialis*). *Restoration Ecology*: 19(599:606). 10.1111/j.1526-100X.2010.00686.x.

Yuan, Y., Wang, B., Zhang, S., Tang, J., Tu, C., Hu, S., Yong, J.W.H., Chen, X., 2012. Enhanced allelopathy and competitive ability of invasive plant *Solidago canadensis* in its introduced range. *Journal of Plant Ecology* 6(3):253-263. https://doi.org/10.1093/jpe/rts033.

Yuan, Y., Qi, L., Yu, J., Wang, X., Huang, L., 2015. Transcriptome wide analysis of SAMe superfamily to novelty phosphoethanolamine N-methyltransferase copy in *Lonicera japonica*. *International Journal of Molecular Science* 16(1):521-534. https://doi.org/10.3390/ijms16010521.

Yu, B., Zheng, B., Ji, L., Chevalier, D., Agarwal, M., Ramachadran, V., Li, W., Lagrange, T., Walker, J.C., Chen, X., 2008. The FHA domain protein DAWDLE in *Arabidopis* and SNIP1 in humans act in small RNA biogenesis. *PNAS* 105(29):10073-10078. https://doi.org/10.1073/pnas.0804218105.

Zamora, D.L., 1984. The allelopathic potential of yellow starthistle (*Centaurea solstitialis* L.). MSc thesis, University of Idaho, Idaho.

Zhang, H., Mallik, A., Zeng, R.S., 2013. Control of Panama disease of banana by rotating and

intercropping with Chinese chive (*Allium tuberosum* Rottller): role of plant volatiles. *Journal of Chemical Ecology* 39:243-252. https://doi.org/10.1007/s10886-013-0243-x.

Zhu, X., Gopurenko, D., Serrano, M., Spencer, M.A., Pieterse, P.J., Skoneczny, D., Lepschi, B.J., Reigosa, M.J., Gurr, G.M., Callaway, R.M., Weston, L.A., 2019. Genetic evidence for plural introduction pathways of the invasive weed Paterson´s curse (*Echium plantagineum* L.) to southern Australia. *PlosOne*. https://doi.org/10.1371/journal.pone.0222696

Appendices

Appendix 1. *Centaurea solstitialis* database comprising the studies included in the systematic review presented in the general introduction, based on WoS search for the period 8 August 2019 – 1995 (262 studies). The plant origin column indicate the number of *C. solstitialis* genotypes sampled from different countries

Author(s)	Publication Year	Plant origin	Country of study	Study type	Hypothesis	Fitness	Factors responsible for *C. solstitialis* success
Braasch et al	2019	6	USA	greenhouse	genetic diversity	no	land use, increased growth, earlier flowering
Barker et al	2019	8	USA	greenhouse	admixture benefits	yes	admixture
Alper and Gunes	2019	1	Turkey	lab	anti-cancer, anti-inflamatory activity	no	no
Woodley et al	2019	1	USA	field observational	biocontrol	yes	climate
Montesinos et al	2019	2	USA	greenhouse	competition resistance	yes	EICA
Lu Irving et al	2019	4	USA	lab	microbiome interactions	no	enemy release
Fitzpatrick et al	2018	3	USA	lab	microbial communities diversity	no	no
Tasar et al	2018	1	Turkey	lab	chromosome number	no	no
Hulvey and Teller	2018	1	USA	field manipulative	invasion resistance	yes	empty niche, density
Eastburn et al	2018	1	USA	field manipulative	invasion resistance	yes	resource interactions, competition for niche occupation
Montesinos and Callaway	2018	2	USA	greenhouse	phenotypic plasticity	yes	high growth rates, trait shifts
Woodley et al	2017	1	USA	field observational	biocontrol	no	no
Gutierrez et al	2017	4	USA	theoretical modelling	biocontrol	no	no
Swope et al	2017	1	USA	field observational	biocontrol	yes	plant density
Miguel et al	2017	1	Argentina	greenhouse and field observational	competition	yes	dispersal dormancy trade-offs, endosperm reserve
Spotswood et al	2017	1	USA	field manipulative	establishment limitation	yes	empty niche
Davy et al	2017	1	USA	field manipulative	resistance to invasion	yes	life cycle
Barker et al	2017	15	USA	greenhouse	genetic variation and invasion routes	yes	increase in plant size

Skowronek et al	2017	1	USA	field observational	weed detection by hyperspectral imagery	no	no
Montesinos and Callaway	2017	2	USA	greenhouse	EICA	yes	trait shifts, local adaptation
Young et al	2017	NA	USA	theoretical modelling	invasion triangle	no	no
Carev et al	2017	1	Croatia	lab	karyotype, genome size, antioxidant and antimicrobial activity	yes	no
Gonzalez et al	2017	1	Turkey	field observational	pollinators	yes	no
Irimia et al	2017	6	Portugal	lab	polyploidy, changes in genome size	yes	no
Disciglio et al	2017	1	Italy	lab	wild edible plants	no	no
Hierro et al	2017	5	Republic of Georgia, Armenia, Turkey, USA, Argentina	field manipulative	factors controlling species abundance	yes	disturbance
Erenler et al	2016	1	Turkey	lab	anticancer activity	no	no
Bruckart et al	2016	1	USA	field observational	biocontrol	yes	no
Baeza et al	2016	1	Chile	lab	chromosome number	no	no
Waller et al	2016	5	USA	greenhouse	enemy release	yes	enemy release
Perkins et al	2016	1	USA	greenhouse	plant soil feedback	yes	plant-soil feedback
Bradley	2016	1	USA	theoretical modelling	predicting species abundance	no	no
Xiao et al	2016	2	USA	theoretical modelling and greenhouse	role of different ecological factors in invasion	yes	disturbance, increased fertility, escape from competition
Filipe et al	2016	6	Portugal	greenhouse	shifting defense	yes	shifting defense
Ciler	2016	1	Turkey	lab	taxonomy	no	no
Uner et al	2016	1	Turkey	lab	thermoplastics	no	no
Bahraminejad et al	2015	1	Iran	lab	anti-fungal activity	no	no
Koc et al	2015	1	Turkey	lab	antioxidant activity	no	no
DiTomaso and Kyser	2015	1	USA	field manipulative	chemical control	yes	no

Lincoln and Beck	2015	1	USA	lab	chemical defense	yes	no
Oster et al	2015	1	USA	field manipulative	climate change	yes	no
Davy et al	2015	1	USA	field manipulative	grazing efficiency	yes	no
Hodgins et al	2015	2	USA	lab	parallel evolution	yes	multiple genetic solutions to adapt to environmental changes
Beck et al	2015	1	USA	lab and field experimental	plant insect chemical interaction	yes	no
Sotes et al	2015	4	Portugal	lab	shifting defense hypothesis	yes	traits shifts, novel chemicals
Dlugosch et al	2015	9	USA	greenhouse	vacant niche	yes	empty niche
Pearson et al	2014	1	Argentina	field manipulative	biotic resistance	yes	disturbance
Bahmani et al	2014	1	Iran	survey	identify medicinal plants	no	no
Oster et al	2014	1	USA	lab and greenhouse	chemical defense	yes	no
Eriksen et al	2014	10	USA	greenhouse	invasion routes, genetic diversity	no	multiple introductions, admixture
Eskelinen and Harrison	2014	1	USA	field manipulative	invasion resistance	yes	disturbance
Swope	2014	1	USA	field manipulative	biocontrol	yes	no
Leong et al	2014	1	USA	field observational	plant pollinator interactions	yes	no
Pearson et al	2014	1	Argentina	field manipulative	response to local community filters	yes	disrupted local filters, plant provenance
Hulvey and Aigner	2014	1	USA	field manipulative	restoration success	no	no
Spencer et al	2014	1	USA	lab and greenhouse	invasive traits	yes	seed weight
Gutierrez and Ponti	2014	1	USA	theoretical modelling	species distribution and abundance	no	no
Scursoni et al	2014	1	Argentina	field observational	weed communities changes	yes	no
Dlugosch et al	2013	7	USA	lab	allelic variation	no	heterozygosity, increased genetic diversity
Spencer et al	2013	1	USA	outdoor microcosms	weed management	yes	deep root system, compensatory response
Cristofaro et al	2013	2	Italy	field manipulative	biocontrol	yes	no

Hierro et al	2013	2	Argentina, Turkey	reciprocal common garden	trait shifts, disturbance	yes	increased seed size
Garcia et al	2013	2	USA	greenhouse	trait shifts	yes	shifts in key traits
Smith and Beck	2013	1	USA	lab	volatile organic compounds	yes	no
Kyser et al	2013	1	USA	field manipulative	weed management	no	no
Graebner et al	2012	2	USA	greenhouse	competition	yes	increased growth rates, EICA
Eriksen et al	2012	4	USA	greenhouse and lab	evolution	yes	local adaptation
Ozcan et al	2012	1	Turkey	lab	honey mineral content	no	no
Smith	2012	1	USA	lab	biocontrol	yes	no
Atanaska et al	2012	1	Bulgaria	greenhouse and field manipulative	biocontrol	yes	no
Lai et al	2012	2	USA	lab	introgression, hybridization	no	no
Hulvey and Zavaleta	2012	1	USA	field manipulative	invasion resistance	yes	empty niche, EICA
Swope and Satterthwaite	2012	1	USA	field manipulative	biocontrol	yes	no
Swope and Stein	2012	1	USA	field manipulative	biocontrol	yes	no
Swope and Parker	2012	1	USA	field observational	biocontrol	yes	no
Andonian et al	2012	5	USA	greenhouse	plant soil microbe interactions	yes	enemy release
Barthell et al	2012	1	Greece	field observational	pollinators guilds	yes	invasion mutualism
Montesinos et al	2012	2	USA	greenhouse	reproductive isolation	yes	local adaptation
Petanidou et al	2012	1	Greece	field manipulative	self-compatibility	yes	self-compatibility
Altay et al	2012	1	Turkey	field observational	urban vegetation	no	no
Matzek and Shannon	2012	1	USA	greenhouse and field manipulative	weed management	yes	no
Uygur	2011	2	Turkey, USA	lab and greenhouse	allelopathy	yes	no
Tekeli et al	2011	1	Turkey	lab	antibacterial effects	no	no
Bruckart et al	2011	1	USA	lab	biocontrol	yes	no
Fisher et al	2011	NA	USA	theoretical modelling	climatic factors	no	no

Hierro et al	2011	2	Argentina, USA	field manipulative and growth chamber	community invasibility	yes	disturbance, empty niche
Young et al	2011	1	USA	field manipulative	community invasibility	yes	empty niche
Spencer et al	2011	1	USA	outdoor microcosms	competition	yes	resource partitioning, empty niche
Bell et al	2011	1	USA	greenhouse	herbicide efficiency	no	no
Eskandari et al	2011	2	USA, France	manipulative and field observational	biocontrol	yes	no
Andonian and Hierro	2011	3	USA	greenhouse	interactions	yes	soil engineering
Chang et al	2011	1	USA	lab	neuropathology	no	no
Gucker and Bunting	2011	1	USA	field observational	plant community composition	no	facilitation
Andonian et al	2011	5	Argentina, Chile, USA, Armenia, Republic of Georgia	greenhouse and field observational	plant soil microbes interactions	yes	soil engineering
Dukes et al	2011	1	USA	field manipulative	response to environmental changes	yes	no
Saatkamp et al	2011	1	France	field manipulative	seed mortality	yes	no
Kyser et al	2011	1	USA	field manipulative	weed management	yes	no
Woods et al	2010	1	USA	field manipulative	biocontrol	yes	no
Birdsall and Markin	2010	1	USA	field observational and lab	biocontrol	yes	no
O´Brian et al	2010	1	USA	field manipulative	biocontrol	yes	no
Munshaw and Lortie	2010	1	USA	field manipulative	density series	yes	positive responses to intraspecific density
Rector et al	2010	1	France	field manipulative and greenhouse	biocontrol	no	no
Roche and Susanna	2010	1	USA	descriptive	hybridization	no	no
Brown and Rice	2010	1	USA	field manipulative	invasion resistance	yes	site characteristics
Kaya et al	2010	1	Turkey	lab	morphological study	no	no
Swope and Parker	2010	1	USA	field manipulative	biocontrol	yes	no

O´Brian et al	2010	1	USA	field manipulative	biocontrol	yes	no
Young et al	2010	1	USA	field manipulative	resource acquisition	yes	no
Goehring et al	2010	1	USA	field manipulative	restoration	yes	no
Lortie et al	2010	1	USA	field observational	seed limitation	yes	seed bank densities
Swope and Parker	2010	1	USA	field manipulative	seed recruitment	yes	compensatory response
EspanchinNiell et al	2010	1	USA	theoretical modelling	weed management	no	no
Tonkel and Piper	2009	1	USA	field observational	biocontrol	yes	no
Ozturk et al	2009	1	Turkey	lab	chromosome number	no	no
Garren and Strauss	2009	1	USA	field manipulative	compensatory responses	yes	compensatory response, plasticity
GomezGonzalez et al	2009	1	Chile	growth chamber and greenhouse	competition	yes	increased fertility, high competitive ability, trait plasticity
Lortie et al	2009	2	Canada	greenhouse	density effects	yes	disturbance, local adaptation and high seed density
Hierro et al	2009	10	USA	greenhouse and growth chamber	germination response	yes	adaptive changes in germination strategies
Barthell et al	2009	2	USA, Greece	field observational	invasion mutualism	yes	invasion mutualism
Akkol et al	2009	1	Turkey	lab	medicinal properties	no	no
Young et al	2009	1	USA	field manipulative	community invasibility	yes	no
Mciver et al	2009	1	USA	field observational	pollinator guilds	yes	invasion mutualism
Bradley et al	2009	1	USA	lab	predicting invasive plant distribution	no	no
Saatkamp et al	2009	1	France	field manipulative	seed bank persistence	yes	no
Woods et al	2009	1	USA	greenhouse and field manipulative	biocontrol	yes	no
Fisher et al	2009	1	USA	lab and field manipulative	biocontrol	yes	no
Aslan et al	2009	NA	USA	theoretical modelling	weed management	no	no
Grimsrud et al	2009	NA	USA	lab	weed management	no	no

Lau et al	2008	1	USA	lathe house and greenhouse	allelopathy	yes	no
Kolomiets et al	2008	1	Russia	lab	biocontrol	yes	no
Woods et al	2008	1	USA	field manipulative	biocontrol	yes	no
Senatore et al	2008	1	Italy	lab	biological activity	no	no
Janackovic et al	2008	1	Serbia	lab	cytotoxicity assay	no	no
Beck et al	2008	1	USA	greenhouse	chemical defense	yes	no
Gutierrez et al	2008	1	USA	lab	climate model projections	no	no
Frost et al	2008	1	USA	lab	forage quality	no	no
Wallace et al	2008	1	USA	field manipulative	grazing efficiency	yes	no
Azirak and Karaman	2008	1	Turkey	lab	inhibitory effects	yes	no
Gultekin et al	2008	1	Turkey	lab and field manipulative	biocontrol	yes	no
Tekeli et al	2008	1	Turkey	lab	scavenging potential	no	no
Gueltekin et al	2008	2	Turkey	field observational, greenhouse, lab	biocontrol	yes	no
Naab et al	2008	1	Argentina	lab	honey characteristics	no	no
Fisher et al	2008	1	USA	greenhouse and field manipulative	biocontrol	yes	no
Qin et al	2007	1	USA	greenhouse	allelopathy	yes	EICA
Gurbuz and Yesilada	2007	1	Turkey	lab	anti-ulcerogenic activity	no	no
Fisher et al	2007	1	USA	field manipulative	biocontrol	yes	no
Balciunas and Korotyaev	2007	5	Armenia, Republic of Georgia, Greece, Russia, Turkey	field observational	biocontrol	yes	no
Julia et al	2007	1	USA	lab	bio-economy	no	no
Inceer et al	2007	1	Turkey	lab	chromosome number	no	no
Pinay et al	2007	1	France	greenhouse	climate change and soil microbial processes	yes	no
Zavaleta and Hulvey	2007	1	USA	outdoor microcosms	community invasibility	yes	no

Fritz and Collinge	2007	1	USA	greenhouse	community invasibility	yes	no
Eagle et al	2007	1	USA	lab	economic impact	no	no
Berner et al	2007	2	USA	greenhouse	biocontrol	yes	no
Balciunas and Villegas	2007	1	USA	lab and field observational	biocontrol	yes	no
Smith	2007	1	USA	lab	biocontrol	no	no
Miao et al	2007	1	USA	field observational	weed detection by hyperspectral imagery	no	no
Ge et al	2007	1	USA	field manipulative	weed detection by hyperspectral imagery	yes	no
Muth and Pigliucci	2007	1	USA	greenhouse	phenotypic plasticity	yes	phenotypic plasticity
Widmer et al	2007	9	France, Russia	growth chamber and field manipulative	resource allocation	yes	seed starch allocation
Torres and Galetto	2007	1	Argentina	lab	style morphology	no	no
Morghan and Rice	2006	1	USA	field manipulative	abiotic filters	yes	empty niche
Widmer	2006	1	France	lab	biocontrol	yes	no
Widmer and Guermache	2006	2	France	lab	biocontrol	yes	no
Gurbuz et al	2006	1	Turkey	lab	biological activity	no	no
Callaway et al	2006	1	USA	greenhouse	compensatory responses	yes	compensatory responses, EICA
Hierro et al	2006	4	USA, Argentina, Turkey	field manipulative and greenhouse	disturbance	yes	disturbance, enemy release
Esmaeili et al	2006	1	Iran	lab	essential oil	no	no
Smith and Drew	2006	1	USA	lab and greenhouse	biocontrol	yes	no
Smith et al	2006	2	Turkey	field manipulative	biocontrol	yes	no
Miao et al	2006	1	USA	field observational and theoretical modelling	weed detection by hyperspectral imagery	no	no
Shaokui et al	2006	1	USA	field observational and lab	weed detection by hyperspectral imagery	no	no
Bruckart	2006	1	USA	greenhouse	biocontrol	yes	no

Batten et al	2006	1	USA	lab	rhizosphere microbial communities, ecosystem engineers	no	changes in soil microbial communities
Muth and Pigliucci	2006	2	USA	greenhouse	traits shifts	yes	traits interactions
Ditomaso et al	2006	1	USA	field manipulative	weed management	yes	no
Young et al	2005	3	USA	lab	achene germination profiles	yes	no
Connett and McCaffrey	2005	1	USA	lab and field observational	biocontrol	yes	no
Berner et al	2005	1	USA	greenhouse	biocontrol	yes	no
Uygur et al	2005	1	Turkey	field manipulative and field observational	biocontrol	yes	no
Enloe et al	2005	1	USA	field manipulative	combined weed management	yes	no
Gutierrez et al	2005	1	USA	theoretical modelling	biocontrol	no	compensatory response
Moret et al	2005	3	Italy	lab	neurotoxic compounds	no	no
Gelbard and Harrison	2005	1	USA	field manipulative	community invasibility	yes	disturbance
Morghan and Rice	2005	1	USA	field manipulative	resistance to invasion	yes	compensatory response
Batten et al	2005	1	USA	lab	ecosystem engineers	yes	soil engineering
Yesilada et al	2004	1	Turkey	lab	anti-ulcerogenic effects	no	no
Faggioli et al	2004	1	Italy	lab	biocontrol	yes	no
Widmer	2004	1	France	lab	biocontrol		no
Duncan et al	2004	1	USA	review	economic impact	no	no
Blank and Young	2004	1	USA	greenhouse	ecosystem engineers	yes	soil engineering
Gerlach	2004	1	USA	field manipulative	ecosystem soil moisture resources	yes	empty niche
Young	2004	1	USA	field manipulative	herbicide efficiency	yes	no
Uygur et al	2004	1	Turkey	field observational	field densities	yes	no
Carrithers et al	2004	1	USA	field manipulative	herbicide efficiency	yes	no
Lindbloom and Zager	2004	1	USA	field manipulative	impact on birds	yes	no
Widmer and Guermache	2004	1	France	lab	plant soil microbe interactions	yes	enemy release, thicker seed coats

Uygur	2004	1	Turkey	field observational	population density and natural enemies	yes	no
Shafii et al	2004	1	USA	theoretical modelling	models to estimate weed occurrence	no	no
Enloe et al	2004	1	USA	field manipulative	soil water use patterns	yes	empty niche
Christopherson and Morrison	2004	1	USA	field manipulative	impacts on nocturnal rodents	yes	no
Menke and Muir	2004	1	USA	field observational	impacts on endangered plants	yes	no
Vidal et al	2004	1	France	lab	biocontrol	yes	no
Lindbloom et al	2003	1	USA	field manipulative	impacts on birds	yes	no
DiTomaso et al	2003	1	USA	field manipulative	soil moisture depletion	yes	soil engineering
Joley et al	2003	1	USA	lab and field manipulative	germination patterns	yes	no
Shinn and Thill	2003	1	USA	field manipulative	chemical control	yes	no
Berner and Paxson	2003	6	USA	dew chamber and greenhouse	biocontrol	yes	no
Morghan et al	2003	1	USA	field manipulative	herbicide efficiency	yes	no
Gelbard and Harrison	2003	1	USA	field observational	factors controlling invasive species	yes	no
Toso and Skliar	2003	1	Argentina	lab	gastric cyto-protective action	no	no
Gerlach and Rice	2003	1	USA	field manipulative and outdoor microcosms	life history traits	yes	plasticity, EICA, disturbance
Sabba et al	2003	2	USA	greenhouse	herbicide resistance	yes	no
Shafii et al	2003	1	USA	theoretical modelling	prediction models to estimate occurrence	no	no
Gastine and Leadley	2003	1	France	field manipulative	ecosystem functioning	yes	no
Ziska	2003	1	USA	growth chamber	climate change	no	no
Bruckart and Eskandari	2002	1	USA	lab	biocontrol	no	no
Shinn and Thill	2002	1	USA	field manipulative	chemical control	yes	no
Dukes	2002	1	USA	outdoor microcosms	community invasibility	yes	no

Riba et al	2002	1	Spain	lab and greenhouse	factors controlling invasive species	yes	no
Kyser and DiTomaso	2002	1	USA	field manipulative	fire as control method	yes	no
ValenzuelaValenzuela et al	2002	2	USA	greenhouse	herbicide efficiency	yes	no
Dukes	2002	1	USA	outdoor microcosms	climate change	yes	no
de Lillo et al	2002	1	Turkey	field observational	biocontrol	yes	no
Eiswerth and van Kooten	2002	1	USA	theoretical modelling	weed management	no	no
Benefield et al	2001	1	USA	growth chamber, field observational and field manipulative	seedbank dynamics	yes	no
Roche et al	2001	1	USA	lab	biocontrol	yes	no
Roche and Thill	2001	NA	USA	review	invasive traits	no	invasive traits, increased fertility, disturbance
Balciunas and Villegas	2001	1	USA	lab	biocontrol	yes	no
Dukes	2001	1	USA	outdoor microcosms	community invasibility	yes	no
Sterling and Murray	2001	2	USA	greenhouse	competition	yes	no
Connett et al	2001	1	USA	field observational	biocontrol	yes	no
Miller et al	2001	1	USA	greenhouse and field manipulative	herbicide resistance	yes	no
ValenzuelaValenzuela et al	2001	2	USA	growth chamber and greenhouse	herbicide resistance	yes	no
Barthell et al	2001	1	USA	field observational and manipulative	invasion mutualism	yes	invasion mutualism
Agrawal et al	2000	1	USA	field manipulative	cost and benefits of plant defense	yes	no
Lass et al	2000	1	USA	modelling and field observational	weed detection by hyperspectral imagery	yes	no
Skinner and Smith	2000	NA	USA	review	noxious weed database	no	no
Jennings et al	2000	1	USA	greenhouse and lab	biocontrol	yes	no

DiTomaso	2000	1	USA	review	weed management	no	no
Gratton and Welter	1999	1	USA	greenhouse and field manipulative	enemy release	no	no
Smith et al	1999	NA	USA	review	management	no	no
DiTomaso et al	1999	1	USA	field manipulative	fire as control method	yes	no
Yesilada et al	1999	1	Turkey	lab	anti-ulcerogenic effects	no	no
Campobasso et al	1998	4	Italy	greenhouse	biocontrol	yes	no
Sun and Ritland	1998	3	USA	field manipulative and lab	mating system	yes	mating system, high level of genetic diversity
Sabba et al	1998	2	USA	greenhouse	herbicide efficiency	yes	no
Roche et al	1997	1	USA	growth chambers and field manipulative	reproductive phenology	yes	no
Sun	1997	3	USA	lab	genetic structure and colonizing routes	yes	high levels of genetic diversity
Joley et al	1997	1	USA	growth chambers	germination response	yes	no
Larson and Kiemnec	1997	2	USA	environmental chamber and field manipulative	germination response	yes	dimorphic seeds, exploiting soil moisture
Sheley and Larson	1997	1	USA	field manipulative	resource acquisition	yes	resource utilization in deep soils
Maxwell and Sheley	1997	NA	USA	theoretical modelling	population dynamics	no	no
Fuerst et al	1996	2	USA	greenhouse and growth chamber	herbicide efficiency	yes	no
Shishkoff and Bruckart	1996	1	USA	greenhouse and dew chamber	biocontrol and drought stress	yes	no
Turner et al	1996	2	USA	field observational	biocontrol	yes	no
Popay and Field	1996	NA	USA	review	grazing efficiency	no	no
Fornasari and Turner	1996	2	Italy, USA	lab	biocontrol	yes	no
Maddox et al	1996	1	USA	field manipulative, outdoor pots and greenhouse	pollination biology	yes	bees
Lass et al	1996	1	USA	field manipulative	weed detection by hyperspectral imagery	no	no

Sheley and Larson	1995	1	USA	field manipulative	plant-plant interactions	yes	no
Harrod and Taylor	1995	1	USA	field manipulative	reproduction and pollination strategy	yes	versatile breeding system
Roy et al	1995	1	USA	lab	neuropathology	no	no

Appendix 2. Geographic locations of *Centaurea solstitialis* populations investigated and their average genome size values. N is the number of individuals analyzed in each population for genome size and ploidy level. Geographic coordinates are given in WSG84 datum. Mean genome size per population in units of mass (picograms), standard deviation and standard error are also provided

ID	Region	Province/County	Location	Latitude	Longitude	Sample size (N)	Mean genome size (pg)	Standard Deviation	Standard Error
1	Turkey	Denizli	Pınarkent	37.80283	29.19525	10	1.73	0.03	0.01
2	Turkey	Afyonkarahisar	Dazkırı	37.95083	29.84033	9	1.72	0.07	0.05
3	Turkey	Isparta	Isparta	37.89844	30.43828	9	1.81	0.09	0.06
4	Turkey	Burdur	Burdur	37.61608	30.14617	8	1.72	0.02	0.02
5	Turkey	Denizli	Serinhisar	37.53156	29.30086	7	NA	NA	NA
6	Turkey	Afyonkarahisar	Dinar to Çay	38.15392	30.23417	9	1.64	0.05	0.04
7	Turkey	Izmir	Beydağ	38.08586	28.21547	9	1.65	0.04	0.02
8	Turkey	Izmir	Bozdağ	38.30136	28.04986	10	1.78	0.03	0.01
9	Turkey	Aydın	Akçaköy	37.95292	28.02997	4	1.80	0.05	0.02
10	Turkey	Aydın	Geyre	37.71231	28.69269	9	1.72	0.05	0.02
11	Spain	Cuenca	Moncalvillo de Huete	40.24159	-2.68745	10	1.66	0.08	0.04
12	Spain	Teruel	Noguera de Albarracín	40.46163	-1.61551	7	NA	NA	NA
13	Spain	Tarragona	Batea	41.06882	0.334305	10	1.70	0.09	0.05
14	Spain	Lleida	L´Espluga-Calba	41.50499	1.005857	9	1.70	0.04	0.02
15	Spain	Salamanca	Castellanos de Moriscos	41.02614	-5.60518	10	1.65	0.10	0.05
16	Spain	Burgos	La Horra	41.7288	-3.83435	10	1.66	0.04	0.02
17	Spain	Burgos	Burgos	42.38389	-3.67894	10	1.72	0.05	0.02

18	Spain	Valladolid	Castronuño	41.39246	-5.27696	10	1.76	0.04	0.02
19	Spain	Zaragoza	Sástago	41.40823	-0.28977	10	1.73	0.09	0.05
20	Spain	Salamanca	Doñinos de Salamanca	40.96853	-5.77332	10	1.76	0.03	0.02
21	Chile	Talagante	Comuna Padre Hurtado	-33.5708	-70.8553	9	1.68	0.04	0.03
22	Chile	Talagante	Comuna El Monte	-33.6894	-71.0553	10	1.73	0.05	0.02
23	Chile	Santiago	Comuna Maipú	-33.5247	-70.7517	8	1.71	0.09	0.04
24	Chile	Santiago	Comuna Lo Barnechea	-33.37	-70.4297	9	1.73	0.03	0.02
25	Argentina	La Pampa	Paraje El Tropezón, R14	-36.709	-64.8311	9	1.75	0.12	0.09
26	Argentina	La Pampa	El Durazno, R14 km 189	-36.7001	-65.3914	10	1.72	0.04	0.02
27	Argentina	La Pampa	Rucanelo, R11 y R10	-36.7089	-64.8308	10	1.72	0.03	0.02
28	Argentina	La Pampa	Victorica	-36.2501	-65.4543	10	1.72	0.08	0.04
29	Argentina	La Pampa	Winifreda	-36.224	-64.2813	10	1.72	0.05	0.03
30	Argentina	La Pampa	Santa Rosa	-36.6166	-64.2563	10	1.73	0.09	0.05
31	Argentina	La Pampa	Trenel, R35 km 426	-35.7246	-64.2693	10	1.78	0.03	0.01
32	Argentina	La Pampa	Quehué, R 35 y R18	-37.1216	-64.2866	10	1.67	0.12	0.07
33	Argentina	La Pampa	Unanue, R35 km 215	-37.5597	-64.2915	10	1.71	0.07	0.04
34	Argentina	La Pampa	Lonquimay, R1 km 203	-36.5655	-63.6647	10	1.75	0.01	0.005
35	California	Sacramento	Folsom	38.64215	-121.176	10	1.77	0.04	0.02
36	California	Sacramento	Folsom	38.68293	-121.181	1	1.64	NA	NA
37	California	Santa Clara	San José	37.24298	-122.871	8	1.64	0.05	0.03
38	California	Marin	Novato	38.15622	-122.693	4	1.66	NA	NA
39	California	Solano	Green Valley	38.20954	-122.146	10	1.76	0.07	0.04
40	California	Sonoma	Petaluma	38.22456	-122.534	10	1.71	0.04	0.02
41	California	Sonoma	Petaluma	38.23643	-122.564	8	1.68	0.04	0.02
42	California	Napa	Napa	38.33904	-122.155	10	1.69	0.07	0.03
43	California	Solano	Vacaville	38.41059	-121.934	10	1.82	0.04	0.02
44	California	Napa	Napa	38.45353	-122.153	10	1.76	0.04	0.02
45	Australia	NSW	Colbinabbin	-36.5955	144.7363	8	1.66	0.05	0.02
46	Australia	NSW	Hume, 5km N of Holbrook	-36.6778	147.3697	7	1.68	0.05	0.03
47	Australia	NSW	Gundagai	-35.0672	148.1085	9	1.72	0.02	0.01
48	Australia	NSW	Gundagai	-34.8447	148.295	10	1.8	0.02	0.01

49	Australia	NSW	18 km N of Cudal	-33.2219	148.9088	2	1.72	0.03	0.02
50	Australia	NSW	Koorawatha	-34.0168	148.569	10	1.67	0.02	0.01
51	Australia	NSW	Murringo	-34.3192	148.4932	10	1.68	0.09	0.05
52	Australia	NSW	Yass	-34.874	148.9083	10	1.71	0.05	0.03

Appendix 3. Genome size and coefficient of variation values for every individual analyzed. Genome size (G.s), mean fluorescence values (FL) and coefficient of variation (CV) for the individuals analysed for genome size in each population

ID	Origin	Individual	Replicate	FL Sample	FL Standard	DI	G.s. (pg)	CV Sample	CV Standard	Obs
1	Turkey	1	1	164.48	106.8	1.54	1.71	2.1	2.65	
1	Turkey	2	1	195.5	125.26	1.56	1.73	3.71	3.97	
1	Turkey	3	1	180.86	113.66	1.59	1.77	3.76	3.72	
2	Turkey	1	1	183.37	122.04	1.5	1.67	4.42	4.68	
2	Turkey	2	1	188.19	117.05	1.61	1.78	3.5	3.72	
3	Turkey	1	2	193.33	122.82	1.57	1.75	2.61	2.39	
3	Turkey	2	1	211.32	124.48	1.7	1.88	3.84	3.78	
4	Turkey	1	1	180.73	117.9	1.53	1.7	3.69	4.14	
4	Turkey	2	1	207.74	132.7	1.57	1.74	2,07	2,37	
5	Turkey	0	0	0	0	0	0	0	0	ploidy level only
6	Turkey	1	1	177.24	122.58	1.45	1.6	3.95	4	
6	Turkey	2	1	164.94	108.8	1.52	1.68	2.45	3.01	
7	Turkey	1	1	160.34	110.12	1.46	1.62	3.59	3.74	
7	Turkey	2	1	167.8	113.8	1.47	1.64	3.6	4.98	
7	Turkey	3	1	175.1	114.47	1.53	1.7	3.85	3.68	
8	Turkey	1	1	175.67	111.31	1.58	1.75	2.4	3.78	
8	Turkey	2	1	184.13	113.73	1.62	1.8	3.6	3.7	
8	Turkey	3	1	190.35	116.85	1.63	1.81	3.4	3.85	
9	Turkey	1	1	182.61	115.63	1.58	1.75	2.19	3.21	
9	Turkey	2	1	197	121.97	1.62	1.79	2.39	2.61	

9	Turkey	3	1	178.27	109.32	1.63	1.81	2.94	3.7	
9	Turkey	4	1	182.04	108.28	1.68	1.87	3.91	2.77	
10	Turkey	1	1	168.87	113.48	1.49	1.65	4.47	4.24	
10	Turkey	2	2	173.05	113.36	1.53	1.69	3.18	3.09	
10	Turkey	3	1	180.62	115.19	1.57	1.74	3.97	3.86	
10	Turkey	4	1	221.28	138.72	1.6	1.77	2.02	2.42	
10	Turkey	5	1	190.87	118.86	1.61	1.78	4.13	3.9	
11	Spain	1	1	159.83	113.13	1.41	1.57	3.65	4.7	
11	Spain	1	1	171.52	113.52	1.51	1.68	3.4	4.65	
11	Spain	3	1	199.47	127.64	1.56	1.73	4.81	5.32	
12	Spain	0	0	0	0	0	0	0	0	ploidy level only
13	Spain	1	1	165.5	113.43	1.46	1.62	3.58	4.03	
13	Spain	2	1	168.63	111	1.52	1.69	4.65	5.47	
13	Spain	3	1	176.33	108.69	1.62	1.8	2.91	3.49	
14	Spain	1	1	168.88	111.93	1.51	1.67	3.07	3.28	
14	Spain	2	1	183.54	120.23	1.53	1.69	4.73	7.11	
14	Spain	3	2	170.08	107.29	1.59	1.76	2.7	3.27	
15	Spain	1	1	154.41	109.41	1.41	1.57	3.45	4.86	
15	Spain	2	1	163.68	111.13	1.47	1.63	2.16	2.71	
15	Spain	3	1	183.72	115.41	1.59	1.77	2.48	3.25	
16	Spain	1	1	163.83	111	1.48	1.64	3.39	5.05	
16	Spain	2	1	168.12	113.53	1.48	1.64	3.36	3.83	
16	Spain	3	1	171.99	111.26	1.55	1.72	2.28	2.28	
17	Spain	1	1	175.47	115.32	1.52	1.69	3.94	3.93	
17	Spain	2	1	166.42	108.5	1.53	1.7	3.59	4.07	
17	Spain	3	1	173.79	112.88	1.54	1.71	3.67	3.46	
17	Spain	4	3	183.62	113.08	1.62	1.8	3.92	4.7	
18	Spain	1	1	181.05	116.56	1.55	1.72	2.73	3.49	
18	Spain	2	1	190.89	118.41	1.61	1.79	3.83	3.58	
18	Spain	3	1	180.57	111.79	1.62	1.79	3.68	4.24	
19	Spain	1	1	170.12	114.67	1.48	1.65	3.13	3.88	

19	Spain	2	1	169.75	109.65	1.55	1.72	3	3.4	
19	Spain	3	1	184.62	111.75	1.65	1.83	4.82	4.2	
20	Spain	1	1	176.51	113.36	1.56	1.73	3.1	4.95	
20	Spain	2	1	174.96	110.61	1.58	1.76	2.97	3.42	
20	Spain	3	1	196.95	121.58	1.62	1.8	1.95	2.63	
21	Chile	1	1	181.28	122.2	1.48	1.65	3.62	3.81	
21	Chile	2	1	172.97	111.91	1.55	1.72	3.73	4.56	
22	Chile	1	1	176.37	116.85	1.51	1.68	3.91	4.67	
22	Chile	2	1	195.5	124.36	1.57	1.74	2.16	2.54	
22	Chile	3	1	189.2	117.81	1.61	1.78	3.95	4.21	
23	Chile	1	1	172.14	120.21	1.43	1.59	3.16	4.19	
23	Chile	2	1	167.06	113.31	1.47	1.64	3.11	4.12	
23	Chile	3	1	165.79	104.34	1.59	1.76	2.9	3.95	
23	Chile	4	1	165.33	103.57	1.6	1.77	2.76	2.93	
23	Chile	5	1	199.34	122.44	1.63	1.81	2.33	3.18	
24	Chile	1	1	174.8	113.59	1.54	1.71	2.8	3.95	
24	Chile	2	1	193.22	122.1	1.58	1.76	2.36	2.69	
25	Argentina	1	1	193.2	129.4	1.49	1.66	4.97	6.1	
25	Argentina	2	1	202.94	122.45	1.66	1.84	2.96	4.05	
26	Argentina	1	1	168.64	111.31	1.52	1.68	3.26	3.63	
26	Argentina	2	1	183.15	118.17	1.55	1.72	3.86	4.41	
26	Argentina	3	1	190.45	120.26	1.58	1.76	3.73	4.07	
27	Argentina	1	1	194.03	127.88	1.52	1.68	4.23	5.76	
27	Argentina	2	1	177.85	114.07	1.56	1.73	2.88	3.7	
27	Argentina	3	1	181.46	115.29	1.57	1.75	2.29	2.82	
28	Argentina	1	1	172	115.04	1.5	1.66	4.12	5.1	
28	Argentina	2	1	172.01	112.31	1.53	1.7	3.17	3.95	
28	Argentina	3	1	204.3	124.54	1.64	1.82	2.26	2.29	
29	Argentina	1	1	174.98	115.4	1.52	1.68	3.17	3.55	
29	Argentina	2	1	187.13	122.21	1.53	1.7	2.66	2.78	
29	Argentina	3	1	193.76	120.42	1.61	1.79	2.6	3.17	

30	Argentina	1	1	161.84	109.49	1.48	1.64	3.36	4.54	
30	Argentina	2	1	170.96	110.09	1.55	1.72	4.24	5.1	
30	Argentina	3	1	185.27	112.62	1.65	1.83	4.24	4.31	
31	Argentina	1	1	167.28	105.94	1.58	1.75	2.48	2.96	
31	Argentina	2	1	198.94	124.18	1.6	1.78	4.2	4.87	
31	Argentina	3	1	188.17	115.34	1.63	1.81	4.49	4.22	
32	Argentina	1	1	157.16	113.9	1.38	1.53	3.1	4.99	
32	Argentina	2	1	208.6	133.71	1.56	1.73	1.92	2.72	
32	Argentina	3	1	182.25	114.88	1.59	1.76	2.34	3.01	
33	Argentina	1	1	154.56	105.43	1.47	1.63	4.27	5.63	
33	Argentina	2	1	189.67	121.55	1.56	1.73	2.29	3	
33	Argentina	3	1	189.64	118.22	1.6	1.78	3.07	3.37	
34	Argentina	1	1	184.08	117.33	1.57	1.74	2.62	3.43	
34	Argentina	2	1	180.99	115.04	1.57	1.75	4.53	6.1	
34	Argentina	3	1	182.57	115.42	1.58	1.76	3.57	3.87	
35	California	1	1	205.77	130.75	1.57	1.75	2.43	3.64	
35	California	2	1	227.92	144.46	1.58	1.75	2.12	3.13	
35	California	3	2	180.59	113.28	1.59	1.77	4.03	3.9	
35	California	4	1	185.24	111.93	1.65	1.84	4.87	5.4	
36	California	1	1	157.36	106.74	1.47	1.64	3.04	5.52	
37	California	1	1	154.83	107.94	1.43	1.59	3.03	4.16	
37	California	2	1	159.27	107.5	1.48	1.64	2.89	3.83	
37	California	3	1	163.03	106.67	1.53	1.7	3.33	4.49	
38	California	1	1	157.84	105.24	1.5	1.66	3.58	3.67	
39	California	1	1	176.97	116.42	1.52	1.69	4.2	4.68	
39	California	2	1	183.04	115.04	1.59	1.77	5.23	3.93	
39	California	3	1	187.19	112.98	1.66	1.84	5.77	4.45	
40	California	1	1	174.35	116.52	1.5	1.66	4.62	4.05	
40	California	2	1	186.04	119.19	1.56	1.73	4.77	4.66	
40	California	3	1	182.54	116.77	1.56	1.74	3.77	3.71	
41	California	1	1	170.33	116.86	1.46	1.62	3.42	3.91	

41	California	2	1	174.46	114.22	1.53	1.7	2.42	2.5	
41	California	3	1	171.83	112.47	1.53	1.7	4.49	5.45	
41	California	4	1	197.62	127.63	1.55	1.72	2.34	3.16	
42	California	1	1	170.78	118.18	1.45	1.6	2.53	3.08	
42	California	2	1	177.82	119.88	1.48	1.65	4.72	4.27	
42	California	3	1	186.78	119.35	1.56	1.74	3.82	3.38	
42	California	4	1	181.39	113.84	1.59	1.77	3.58	4.15	
43	California	1	1	165.09	103.18	1.6	1.78	4.47	4.64	
43	California	2	1	187.48	112.54	1.67	1.85	4.7	4.43	
43	California	3	1	188.69	113.18	1.67	1.85	5.31	5.04	
44	California	1	2	194.57	123.02	1.58	1.76	4.34	5.16	
44	California	2	2	191.43	120	1.6	1.77	4.29	4.9	
44	California	3	1	113.09	543.61	0.21	1.83	5.6	3.71	
44	California	4	1	116.96	234.5	0.5	1.71	6.4	4.59	
45	Australia	1	1	169.22	115.63	1.46	1.62	4.21	4.46	
45	Australia	2	1	180.64	122.22	1.48	1.64	3.78	5.81	
45	Australia	3	1	175.13	117.83	1.49	1.65	2.53	2.8	
45	Australia	4	1	197.17	124.93	1.58	1.75	1.85	2.45	
46	Australia	1	1	177.12	120.31	1.47	1.63	4.6	6.41	
46	Australia	2	1	162.03	106.78	1.52	1.68	3.72	4.24	
46	Australia	3	1	208.68	133.46	1.56	1.74	2.07	2.43	
47	Australia	1	1	179.09	116.44	1.54	1.71	3.52	5.23	
47	Australia	2	1	187.89	119.66	1.57	1.74	2.96	2.87	
48	Australia	1	2	187.56	117.21	1.6	1.78	3.97	3.55	
48	Australia	2	1	186.38	113.25	1.65	1.83	4.34	4.00	
48	Australia	3	1	188.41	116.69	1.61	1.79	4.01	3.93	
49	Australia	1	1	188.91	124.17	1.52	1.69	3.47	4.87	
49	Australia	2	1	169.52	108.77	1.56	1.73	3.2	3.84	
49	Australia	3	1	201.44	127.09	1.59	1.76	3.68	3.85	
50	Australia	1	1	168.93	113.34	1.49	1.65	4.51	4.53	
50	Australia	2	1	173.56	116.21	1.49	1.66	2.45	3.81	

50	Australia	3	1	187.05	121.85	1.54	1.7	4.4	4.27	
51	Australia	1	1	161.75	112.75	1.43	1.59	3.21	4.08	
51	Australia	2	1	162.4	107.14	1.52	1.68	2.83	4.57	
51	Australia	3	1	175.53	109.16	1.61	1.78	3.36	3.5	
52	Australia	1	1	180.54	121.08	1.49	1.66	3.72	5.62	
52	Australia	2	1	175.74	113.94	1.54	1.71	2.68	3.38	
52	Australia	3	1	190.89	119.61	1.6	1.77	2.54	3.25	

Appendix 4. Genome size (G.s), mean fluorescence values (FL) and coefficient of variation (CV) for the individuals analyzed for ploidy level

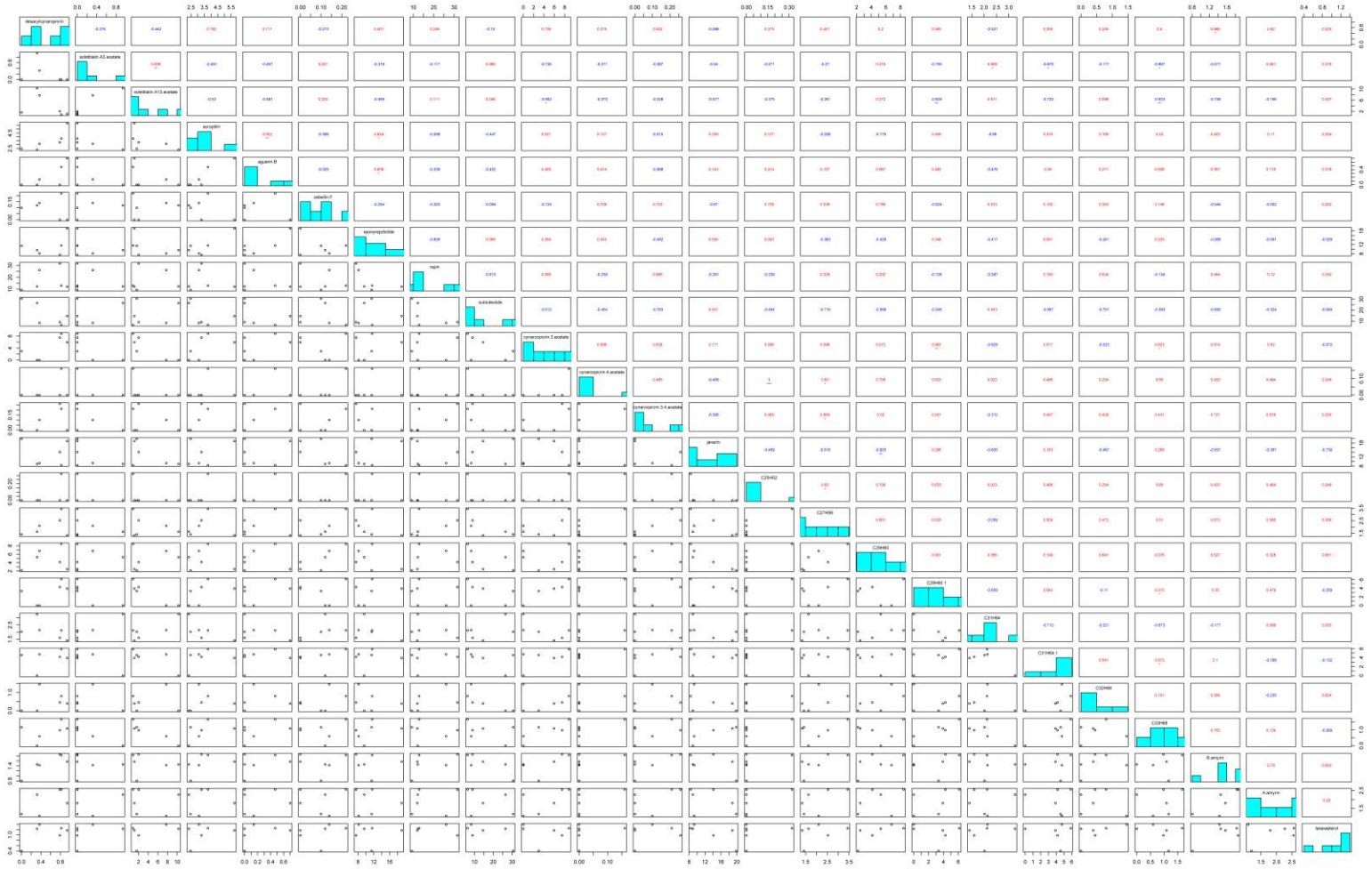
ID	Origin	Sample size (individuals in pool)	FL Sample	FL Standard	DI	G.s. (pg)	CV Sample	CV Standard	Ploidy	Obs
1	Turkey	4	197.52	123.78	1.60	1.77	4.33	4.26	diploid	
1	Turkey	3	195.66	121.89	1.61	1.78	3.78	3.87	diploid	
2	Turkey	3	178.29	110.77	1.61	1.79	3.83	4.14	diploid	
2	Turkey	4	182.37	117.38	1.55	1.72	4.12	3.76	diploid	
3	Turkey	3	158.13	105.18	1.50	1.67	3.35	3.82	diploid	
3	Turkey	4	154.65	104.93	1.47	1.64	3.70	4.24	diploid	
4	Turkey	3	153.99	105.07	1.47	1.63	3.79	4.95	diploid	
4	Turkey	3	154.27	104.75	1.47	1.63	4.32	6.28	diploid	
5	Turkey	4	151.54	100.00	1.52	1.68	3.39	5.99	diploid	
5	Turkey	3	149.32	105.33	1.42	1.57	6.14	4.90	diploid	
6	Turkey	3	175.35	116.87	1.50	1.67	2.88	3.15	diploid	
6	Turkey	4	165.08	108.61	1.52	1.69	4.34	4.48	diploid	
7	Turkey	3	179.64	124.64	1.44	1.60	4.99	5.81	diploid	
7	Turkey	4	169.26	113.27	1.49	1.66	4.24	6.17	diploid	

8	Turkey	3	178.42	109.64	1.63	1.81	3.44	3.57	diploid	
8	Turkey	4	189.55	120.64	1.57	1.74	4.17	4.93	diploid	
9	Turkey	0	0	0	0	0	0	0	NA	genome size only
10	Turkey	4	179.34	110.53	1.62	1.80	5.23	4.56	diploid	
11	Spain	3	184.17	116.37	1.58	1.76	4.77	5.95	diploid	
11	Spain	4	181.36	115.36	1.57	1.75	3.74	4.76	diploid	
12	Spain	3	156.51	102.90	1.52	1.69	3.84	4.87	diploid	
12	Spain	4	164.00	112.45	1.46	1.62	3.83	4.60	diploid	
13	Spain	4	171.00	111.26	1.54	1.71	3.10	3.80	diploid	
13	Spain	3	179.94	120.00	1.50	1.66	3.83	4.91	diploid	
14	Spain	4	172.76	114.64	1.51	1.67	3.59	4.78	diploid	
14	Spain	3	155.53	101.86	1.53	1.69	3.71	4.87	diploid	
15	Spain	3	163.47	109.37	1.49	1.66	4.83	5.03	diploid	
15	Spain	4	164.58	107.89	1.53	1.69	3.40	3.84	diploid	
16	Spain	4	162.94	116.06	1.40	1.56	4.40	5.77	diploid	
16	Spain	3	173.30	115.64	1.50	1.66	4.44	5.34	diploid	
17	Spain	4	161.29	106.50	1.51	1.68	4.27	4.38	diploid	
17	Spain	3	175.69	116.73	1.51	1.67	4.41	4.08	diploid	
18	Spain	4	168.41	114.35	1.47	1.63	4.89	4.98	diploid	
18	Spain	3	165.80	105.92	1.57	1.74	3.80	4.31	diploid	
19	Spain	4	165.83	105.89	1.57	1.74	3.75	3.24	diploid	
19	Spain	3	171.83	110.61	1.55	1.72	3.26	3.91	diploid	
20	Spain	4	193.93	124.53	1.56	1.73	4.10	3.05	diploid	
20	Spain	3	176.72	113.46	1.56	1.73	4.12	4.49	diploid	
21	Chile	4	181.31	113.51	1.60	1.77	5.93	7.40	diploid	
21	Chile	3	184.78	117.55	1.57	1.74	4.86	5.66	diploid	
22	Chile	3	189.21	119.28	1.59	1.76	3.92	4.35	diploid	
22	Chile	4	186.82	116.51	1.60	1.78	4.00	3.55	diploid	
23	Chile	3	156.20	100.94	1.55	1.72	3.47	5.02	diploid	
24	Chile	3	163.38	112.56	1.45	1.61	4.44	5.02	diploid	
24	Chile	4	158.57	105.66	1.50	1.67	2.69	3.18	diploid	

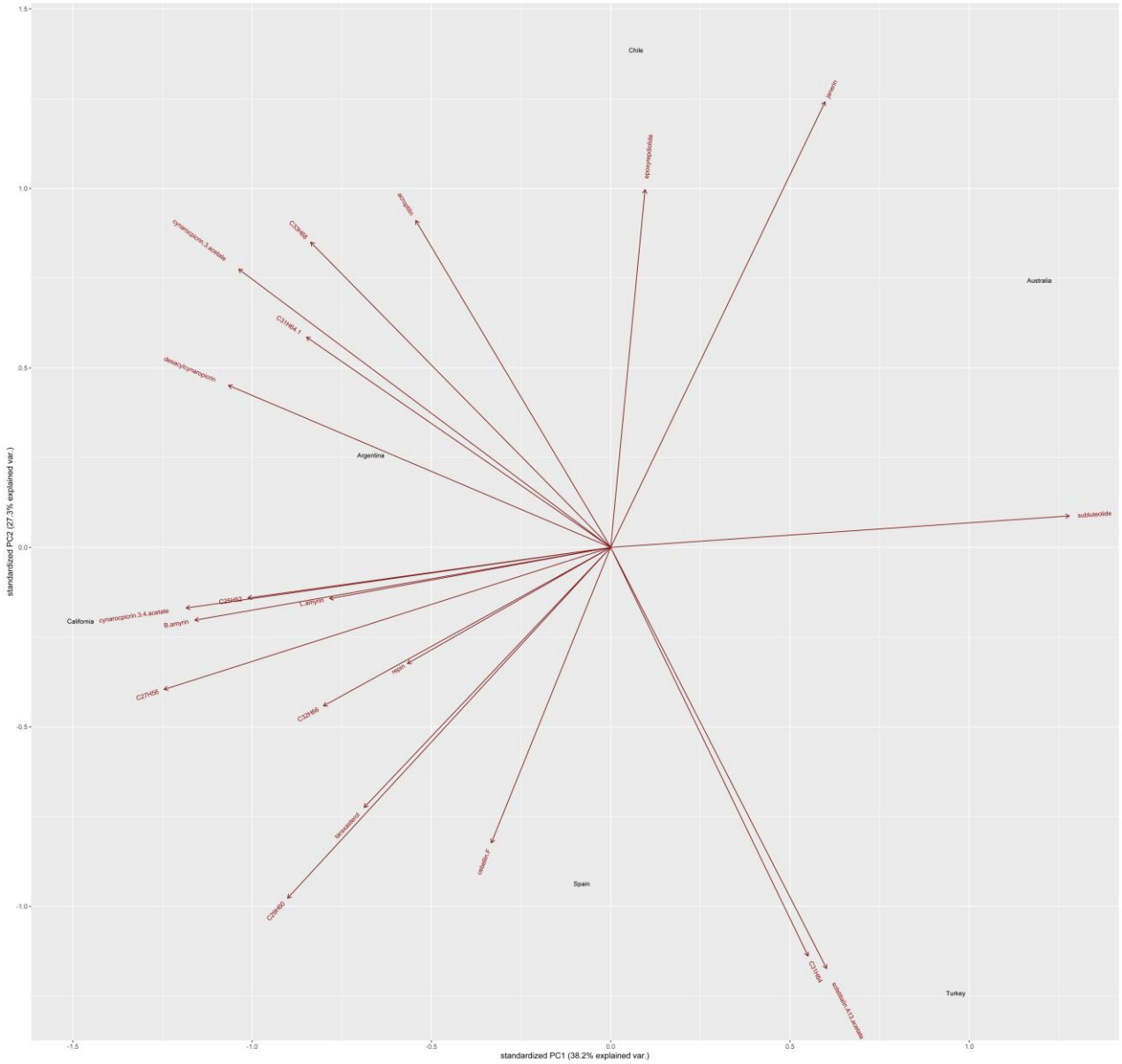
25	Argentina	4	156.06	106.37	1.47	1.63	4.24	6.31	diploid	
25	Argentina	3	178.97	117.50	1.52	1.69	5.10	6.59	diploid	
26	Argentina	4	177.84	112.01	1.59	1.76	3.67	4.25	diploid	
26	Argentina	3	184.09	114.94	1.60	1.78	4.02	3.79	diploid	
27	Argentina	4	172.38	115.19	1.50	1.66	3.37	4.31	diploid	
27	Argentina	3	144.15	100.99	1.43	1.58	4.76	7.41	diploid	
27	Argentina	3	184.99	118.64	1.56	1.73	3.83	3.83	diploid	
28	Argentina	4	158.73	103.59	1.53	1.70	3.01	3.76	diploid	
28	Argentina	3	154.78	106.40	1.45	1.61	3.31	4.48	diploid	
29	Argentina	3	184.61	116.39	1.59	1.76	2.68	2.59	diploid	
29	Argentina	4	174.96	107.14	1.63	1.81	3.60	3.73	diploid	
30	Argentina	4	179.54	110.82	1.62	1.80	4.17	4.51	diploid	
30	Argentina	3	166.37	112.31	1.48	1.64	3.76	5.25	diploid	
31	Argentina	3	188.87	117.39	1.61	1.79	4.32	4.04	diploid	
31	Argentina	4	178.91	112.40	1.59	1.77	4.42	4.56	diploid	
32	Argentina	4	167.53	118.13	1.42	1.57	4.30	5.32	diploid	
32	Argentina	3	174.04	121.08	1.44	1.60	5.21	5.82	diploid	
33	Argentina	4	167.60	109.58	1.53	1.70	3.22	4.05	diploid	
33	Argentina	3	163.57	109.37	1.50	1.66	3.23	3.78	diploid	
34	Argentina	3	175.61	112.76	1.56	1.73	3.93	4.61	diploid	
34	Argentina	4	182.96	116.90	1.57	1.74	3.54	4.55	diploid	
35	California	4	194.45	119.05	1.63	1.81	4.50	4.82	diploid	
35	California	3	192.10	118.19	1.63	1.80	4.87	4.89	diploid	
36	California	0	0	0	0	0	0	0	NA	genome size only
37	California	4	163.16	106.75	1.53	1.70	4.12	4.97	diploid	
38	California	4	152.62	101.37	1.51	1.67	5.02	5.64	diploid	
39	California	4	180.49	114.91	1.57	1.74	5.35	4.76	diploid	
39	California	3	176.05	116.69	1.51	1.67	4.96	4.31	diploid	
40	California	4	173.29	115.66	1.50	1.66	5.00	4.49	diploid	
40	California	3	178.23	121.40	1.47	1.63	4.61	4.10	diploid	
41	California	4	167.40	110.73	1.51	1.68	3.54	3.84	diploid	

42	California	3	171.98	118.02	1.46	1.62	5.66	4.72	diploid	
42	California	4	168.72	112.21	1.50	1.67	4.57	4.20	diploid	
43	California	4	189.14	115.80	1.63	1.81	5.73	5.67	diploid	
43	California	3	189.15	115.15	1.64	1.82	5.02	4.98	diploid	
44	California	3	186.42	118.30	1.58	1.75	4.67	4.87	diploid	
44	California	4	179.12	110.44	1.62	1.80	5.50	6.24	diploid	
45	Australia	4	150.45	99.94	1.51	1.67	4.54	5.03	diploid	
46	Australia	4	166.88	110.75	1.51	1.67	3.73	5.09	diploid	
47	Australia	3	168.87	112.46	1.50	1.67	3.43	4.91	diploid	
47	Australia	4	161.69	108.12	1.50	1.66	3.45	4.26	diploid	
48	Australia	0	0	0	0	0	0	0	NA	genome size only
49	Australia	3	177.92	110.89	1.60	1.78	4.44	5.72	diploid	
49	Australia	4	174.70	106.83	1.64	1.82	4.81	5.17	diploid	
50	Australia	4	175.66	107.08	1.64	1.82	4.77	5.28	diploid	
50	Australia	3	181.71	115.85	1.57	1.74	4.67	5.51	diploid	
51	Australia	3	174.61	110.43	1.58	1.76	3.62	4.39	diploid	
51	Australia	4	167.47	110.32	1.52	1.69	3.67	4.73	diploid	
52	Australia	4	158.45	108.56	1.46	1.62	3.66	7.13	diploid	
52	Australia	3	166.02	112.68	1.47	1.64	4.09	6.08	diploid	

Appendix 5. Pairwise scatter plots of all the chemical variables. The x axis in each scatter plot represents the column variable, the y axis the row variable. The value of the correlation and the significance level is indicated by asterisks (0.1 = ".", 0.05 = "*", 0.01 = "**", 0.001 = "***") on the top of the diagonal. Positive correlations are displayed in red and negative correlations in blue colour. A higher resolution image can be accessed from https://link.springer.com/article/10.1007%2Fs10530-019-02038-1.



Appendix 6. Plot of Principal Component Analysis scores (centered and standardized PCA) representing the multivariate space in chemical compounds found in six *C. solstitialis* world regions after performing the collinearity test and dropping the highly correlated variables from the model A higher resolution image can be accessed from https://link.springer.com/article/10.1007%2Fs10530-019-02038-1.



Appendix 7. Tukey post-hoc comparisons of the effects of a 0.25% *C. solstitialis* leaf extract on the germination rates of lettuce seeds. N = number of lettuce seeds germinated. SE = standard error of the mean

Group	N	Mean	SE	Turkey	Spain	Argentina	Chile	California
Turkey	23	7.66	0.88					
Spain	18	6	2.88	> 0.05				
Argentina	20	6.66	0.33	> 0.05	> 0.05			
Chile	4	1.33	0.88	**0.006**	**0.04**	**0.04**		
California	49	16.3	3.84	**0.002**	**< 0.001**	**< 0.001**	**< 0.001**	
Australia	65	21.6	2.4	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	> 0.05

Appendix 8. Tukey post-hoc comparisons of the effects of a 0.5% *C. solstitialis* leaf extract on the germination rates of lettuce seeds. N = number of lettuce seeds germinated. SE = standard error of the mean

Group	N	Mean	SE	Turkey	Spain	Argentina	Chile	California
Turkey	2	0.66	0.33					
Spain	8	2.66	1.2	> 0.05				
Argentina	0	NA	NA	> 0.05	> 0.05			
Chile	0	NA	NA	> 0.05	> 0.05	> 0.05		
California	1	0.33	0.33	> 0.05	> 0.05	> 0.05	> 0.05	
Australia	2	0.66	0.33	> 0.05	> 0.05	> 0.05	> 0.05	> 0.05

Appendix 9. Tukey post-hoc comparisons of the effects of a 0.25% *C. solstitialis* leaf extract on days to germination in lettuce

Group	N	Mean	SE	Turkey	Spain	Argentina	Chile	California
Turkey	23	5.86	0.32					
Spain	18	5.72	0.32	> 0.05				
Argentina	20	6	0.47	> 0.05	> 0.05			
Chile	4	5.75	0.75	> 0.05	> 0.05	> 0.05		
California	49	5.06	0.22	> 0.05	> 0.05	> 0.05	> 0.05	
Australia	65	4.73	0.21	> 0.05	> 0.05	**0.038**	> 0.05	> 0.05

Appendix 10. Percent germination and days to germination for lettuce seeds receiving 0.25% *C. solstitialis* leaf extract. PC = positive control, NC = negative control

Region	PC	NC	Turkey	Spain	Argentina	Chile	California	Australia
Percent germination	99	99	21	17	19	3.7	45	60
Days to germination	2.1	2.2	5.8	5.7	6	5.7	5	4.7

Appendix 11. Tukey post-hoc comparisons of the effects of a 0.5% *C. solstitialis* leaf extract on days to germination in lettuce

Group	N	Mean	SE	Turkey	Spain	Argentina	Chile	California
Turkey	2	5.5	1.5					
Spain	8	6.37	0.53	> 0.05				
Argentina	0	NA	NA	NA	> 0.05			
Chile	0	NA	NA	NA	> 0.05	> 0.05		
California	1	7	NA	> 0.05	> 0.05	> 0.05	> 0.05	
Australia	2	7	NA	> 0.05	> 0.05	> 0.05	> 0.05	> 0.05

Appendix 12. Percent germination and days to germination for lettuce seeds receiving 0.5% *C. solstitialis* leaf extract. PC = positive control, NC = negative control

Region	PC	NC	Turkey	Spain	Argentina	Chile	California	Australia
Percent germination	99	99	1.9	7.4	0	0	0.9	1.9
Days to germination	2.1	2.2	5.5	6.3	NA	NA	7	7

Appendix 13. Tukey post-hoc comparisons of the effects of a 0.25% *C. solstitialis* leaf extract on the growth of lettuce root (mm). N = number of lettuce seeds germinated. SE = standard error of the mean

Group	N	Mean	SE	Turkey	Spain	Argentina	Chile	California
Turkey	23	0.18	0.01					
Spain	18	0.41	0.08	> 0.05				
Argentina	18	0.19	0.05	> 0.05	> 0.05			
Chile	4	0.22	0.12	> 0.05	> 0.05	> 0.05		
California	49	0.41	0.04	**0.03**	> 0.05	> 0.05	> 0.05	
Australia	65	0.43	0.03	**0.007**	> 0.05	**0.027**	> 0.05	> 0.05

Appendix 14. Tukey post-hoc comparisons of the effects of a 0.5% *C. solstitialis* leaf extract on the growth of lettuce root (mm)

Group	N	Mean	SE	Turkey	Spain	Argentina	Chile	California
Turkey	2	0.1	0					
Spain	8	0.13	0.01	> 0.05				
Argentina	0	NA	NA	> 0.05	**0.003**			
Chile	0	NA	NA	> 0.05	**0.003**	> 0.05		
California	1	0.1	NA	> 0.05	> 0.05	> 0.05	> 0.05	
Australia	2	0.15	0.05	> 0.05	> 0.05	**0.018**	**0.018**	> 0.05

Appendix 15. Geographical coordinates of the populations included in this study. N = number of replicates per population, each representing a different maternal plant. Datum is WGS84.

Population	N	Region	Province/County	Latitude	Longitude
1	10	Turkey	Denizli	37.802	29.195
2	6	Turkey	Afyonkarahisar	37.95	29.84
3	6	Turkey	Isparta	37.898	30.438
4	7	Turkey	Burdur	37.616	30.146
5	6	Turkey	Denizli	37.531	29.3
6	3	Turkey	Afyonkarahisar	38.153	30.234
7	10	Turkey	Izmir	38.085	28.215
8	7	Turkey	Izmir	38.301	28.049
9	5	Turkey	Aydın	37.952	28.03
10	5	Turkey	Aydın	37.712	28.692
11	8	Spain	Cuenca	40.241	-2.687
12	4	Spain	Tarragona	41.068	0.334
13	10	Spain	Lleida	41.505	1.005
14	6	Spain	Lleida	41.877	0.778
15	4	Spain	Cuenca	39.847	-2.501
16	8	Spain	Salamanca	41.026	-5.605
17	10	Spain	Burgos	41.728	-3.834
18	6	Spain	Valladolid	41.392	-5.277
19	4	Spain	Cuenca	40.014	-2.973
20	6	Spain	Zaragoza	41.408	-0.289
21	9	Argentina	La Pampa	-36.709	-64.831
22	6	Argentina	La Pampa	-36.7	-65.391
23	10	Argentina	La Pampa	-36.709	-64.831
24	9	Argentina	La Pampa	-36.25	-65.454
25	7	Argentina	La Pampa	-36.224	-64.281

26	9	Argentina	La Pampa	-36.617	-64.256
27	8	Argentina	La Pampa	-35.725	-64.269
28	8	Argentina	La Pampa	-37.122	-64.287
29	8	Argentina	La Pampa	-37.56	-64.292
30	5	Argentina	La Pampa	-36.566	-63.665
31	3	Chile	Talagante	-33.571	-70.855
32	8	Chile	Talagante	-33.689	-71.055
33	6	Chile	Santiago	-33.525	-70.752
34	8	Chile	Santiago	-33.37	-70.43
35	10	California	Sacramento	38.642	-121.18
36	2	California	Sacramento	38.683	-121.181
37	2	California	Marin	38.156	-122.69
38	8	California	Solano	38.209	-122.15
39	9	California	Sonoma	38.224	-122.53
40	2	California	Sonoma	38.236	-122.56
41	8	California	Napa	38.339	-122.15
42	6	California	Solano	38.41	-121.93
43	9	California	Napa	38.453	-122.15
44	5	Australia	NSW	-36.595	144.736
45	6	Australia	NSW	-36.678	147.37
46	7	Australia	NSW	-35.067	148.109
47	5	Australia	NSW	-33.222	148.909
48	7	Australia	NSW	-34.017	148.569
49	9	Australia	NSW	-34.319	148.493
50	9	Australia	NSW	-34.874	148.908

Appendix 16. Linear regression model output of fertility rates for crosses within region. $\chi 2$ – chi-squared distribution. AIC - Akaike information criterion**.** Significance levels: "*" $P \leq 0.05$ significant, "**" $P \leq 0.01$ highly significant, "***" $P \leq 0.001$ extremely significant. In bold are shown the crosses that differed significantly at $p < 0.05$

Region	$\chi^2(_1)$	AIC	P-value
Turkey	**10.338**	**AIC $_{\text{null model}}$ = 1590.3 *vs* AIC $_{\text{full model}}$ = 1598.6**	**0.01***
Spain	1.4531	AIC $_{\text{null model}}$ = 1012.0 *vs* AIC $_{\text{full model}}$ = 1011.5	0.228
Argentina	**25.28**	**AIC $_{\text{null model}}$ = 2260.3 *vs* AIC $_{\text{full model}}$ = 2283.6**	**< 0.001*** **
Chile	**51.69**	**AIC $_{\text{null model}}$ = 230.81 *vs* AIC $_{\text{full model}}$ = 280.50**	**< 0.001*** **
California	1.4681	AIC $_{\text{null model}}$ = 1092.6 *vs* AIC $_{\text{full model}}$ = 1092.1	0.2256
Australia	**16.323**	**AIC $_{\text{null model}}$ = 676.26*vs* AIC $_{\text{full model}}$ = 690.58**	**< 0.001*** **

Appendix 17. Linear regression model output of fertility rates for crosses between regions. In bold are shown the crosses that differed significantly at $p < 0.05$

Comparison	$\chi^2(1)$	AIC	P-value
Turkey vs Spain	0.216	AIC $_{\text{null model}}$ = 1127.5 *vs*AIC $_{\text{full model}}$ = 1125.8	0.641
Turkey vs Chile	**19.737**	**AIC $_{\text{null model}}$ = 1021.1*vs*AIC $_{\text{full model}}$ = 1038.9**	**< 0.001*****
Turkey vs California	0.1482	AIC $_{\text{null model}}$ = 1478.0 *vs* AIC $_{\text{full model}}$ = 1476.2	0.700
Turkey vs Australia	0.101	AIC $_{\text{null model}}$ = 1120.4*vs* AIC $_{\text{full model}}$ = 1118.5	0.749
Turkey vs Argentina	2.2597	AIC $_{\text{null model}}$ = 1273.5*vs* AIC $_{\text{full model}}$ = 1273.8	0.132
Spain vs Turkey	2.4185	AIC $_{\text{null model}}$ = 1143.5*vs*AIC $_{\text{full model}}$ = 1143.9	0.119
Spain vs Chile	**16.125**	**AIC $_{\text{null model}}$ = 868.25*vs*AIC $_{\text{full model}}$ = 882.38**	**< 0.001*****
Spain vs California	**48.812**	**AIC $_{\text{null model}}$ = 1687.1*vs* AIC $_{\text{full model}}$ = 1733.9**	**< 0.001*****
Spain vs Australia	**37.321**	**AIC $_{\text{null model}}$ = 1040.5*vs*AIC $_{\text{full model}}$ = 1075.8**	**< 0.001*****
Spain vs Argentina	**39.613**	**AIC $_{\text{null model}}$ = 1534.8*vs* AIC $_{\text{full model}}$ = 1572.4**	**< 0.001*****
Chile vs Turkey	**4.978**	**AIC $_{\text{null model}}$ = 362.36 *vs*AIC $_{\text{full model}}$ = 365.33**	**0.025***
Chile vs Spain	1.551	AIC $_{\text{null model}}$ = 349.52 *vs* AIC $_{\text{full model}}$ = 349.07	0.212
Chile vs California	**10.514**	**AIC $_{\text{null model}}$ = 419.81*vs*AIC $_{\text{full model}}$ = 428.32**	**0.001****
Chile vs Australia	0.4859	AIC $_{\text{null model}}$ = 321.45*vs*AIC $_{\text{full model}}$ = 319.93	0.485
Chile vs Argentina	**7.780**	**AIC $_{\text{null model}}$ = 401.76*vs* AIC $_{\text{full model}}$ = 407.54**	**0.005****
California vs Turkey	2.1573	AIC $_{\text{null model}}$ = 850.23*vs*AIC $_{\text{full model}}$ = 850.38	0.141
California vs Spain	**15.386**	**AIC $_{\text{null model}}$ = 787.01*vs*AIC $_{\text{full model}}$ = 800.40**	**< 0.001*****
California vs Chile	**20.202**	**AIC $_{\text{null model}}$ = 569.68*vs* AIC $_{\text{full model}}$ = 587.88**	**< 0.001*****
California vs Australia	**6.655**	**AIC $_{\text{null model}}$ = 760.73*vs*AIC $_{\text{full model}}$ = 765.39**	**0.009****
California vs Argentina	0.1133	AIC $_{\text{nullmodel}}$ =1078.7*vs*AIC $_{\text{full model}}$ =1076.8	0.7363

Australia vs Turkey	1.156	AIC $_{\text{null model}}$ = 575.67*vs*AIC $_{\text{full model}}$ = 574.83	0.282
Australia vs Spain	0.049	AIC $_{\text{null model}}$ = 418.07*vs*AIC $_{\text{full model}}$ = 416.11	0.824
Australia vs Chile	3.6012	AIC $_{\text{null model}}$ = 359.04*vs*AIC $_{\text{full model}}$ = 360.64	0.057
Australia vs California	**22.977**	**AIC $_{\text{null model}}$ = 734.26*vs* AIC $_{\text{full model}}$ = 755.24**	**< 0.001*****
Australia vs Argentina	**9.7344**	**AIC $_{\text{null model}}$ = 625.91*vs*AIC $_{\text{full model}}$ = 633.64**	**0.001****
Argentina vs Turkey	0.6758	AIC $_{\text{null model}}$ =2015.3*vs*AIC $_{\text{full model}}$ =2013.9	0.411
Argentina vs Spain	**25.1**	**AIC $_{\text{null model}}$ = 1377.3*vs*AIC $_{\text{full model}}$ = 1400.4**	**< 0.001*****
Argentina vs Chile	1.8289	AIC $_{\text{null model}}$ = 1470.8*vs*AIC $_{\text{full model}}$ = 1470.7	0.1763
Argentina vs California	**32.457**	**AIC $_{\text{nullmodel}}$ =1834.7*vs*AIC $_{\text{full model}}$ =1865.2**	**< 0.001*****
Argentina vs Australia	2.0988	AIC $_{\text{null model}}$ = 1567.6*vs* AIC $_{\text{full model}}$= 1567.7	0.147

Appendix 18. GPS coordinates of the *C. solstitialis* populations sampled in this study (Datum WGS84). Pop = population

Pop	Region	Province/County	Location	Latitude	Longitude
1	Turkey	Denizli	Pınarkent	37.80283	29.19525
2	Turkey	Afyonkarahisar	Dazkırı	37.95083	29.84033
3	Turkey	Isparta	Isparta	37.89844	30.43828
4	Turkey	Burdur	Burdur	37.61608	30.14617
5	Turkey	Denizli	Serinhisar	37.53156	29.30086
6	Turkey	Afyonkarahisar	Dinar to Çay	38.15392	30.23417
7	Turkey	Izmir	Beydağ	38.08586	28.21547
8	Turkey	Izmir	Bozdağ	38.30136	28.04986
9	Turkey	Aydın	Akçaköy	37.95292	28.02997
10	Turkey	Aydın	Geyre	37.71231	28.69269
11	Spain	Cuenca	Moncalvillo de Huete	40.24159	-2.687453
12	Spain	Tarragona	Batea	41.06882	0.334305
13	Spain	Lleida	L´Espluga-Calba	41.50499	1.005857
14	Spain	Lleida	Gerb	41.87776	0.778373
15	Spain	Salamanca	Castellanos de Moriscos	41.02614	-5.605175
16	Spain	Burgos	La Horra	41.7288	-3.834349
17	Spain	Valladolid	Castronuño	41.39246	-5.276957

18	Spain	Cuenca	Tarancón	40.01424	-2.9731
19	Spain	Zaragoza	Sástago	41.40823	-0.289773
20	Chile	Talagante	Padre Hurtado	-33.57083	-70.85528
21	Chile	Talagante	El Monte	-33.68944	-71.05528
22	Chile	Santiago	Maipú	-33.52472	-70.75167
23	Chile	Santiago	Lo Barnechea	-33.37	-70.42972
24	Argentina	La Pampa	Paraje El Tropezón, R14	-36.709	-64.83106
25	Argentina	La Pampa	El Durazno, R14, km 189	-36.70008	-65.39142
26	Argentina	La Pampa	Rucanelo, R11 y R10	-36.70894	-64.83083
27	Argentina	La Pampa	Victorica	-36.25011	-65.45425
28	Argentina	La Pampa	Winifreda	-36.22397	-64.28133
29	Argentina	La Pampa	Santa Rosa	-36.61658	-64.25625
30	Argentina	La Pampa	Trenel, R35 km 426	-35.72458	-64.26928
31	Argentina	La Pampa	Quehué, R 35 y R18	-37.12161	-64.28661
32	Argentina	La Pampa	Unanue, R35 km 215	-37.55967	-64.2915
33	Argentina	La Pampa	Lonquimay, R1 km 203	-36.56553	-63.66467
34	California	Sacramento	Folsom	38.64215	-121.176
35	California	Solano	Green Valley	38.20954	-122.1463
36	California	Sonoma	Petaluma	38.22456	-122.5341
37	California	Sonoma	Petaluma	38.23643	-122.5638
38	California	Napa	Napa	38.33904	-122.1547
39	California	Solano	Vacaville	38.41059	-121.9343
40	California	Napa	Napa	38.45353	-122.1529
41	Australia	NSW	Hume, 5km N of Holbrook	-36.67779	147.3697
42	Australia	NSW	Gundagai	-35.0672	148.1085
43	Australia	NSW	18 km N of Cudal	-33.22193	148.9088
44	Australia	NSW	Koorawatha	-34.01678	148.569
45	Australia	NSW	Murringo	-34.31917	148.4932
46	Australia	NSW	Yass	-34.87395	148.9083

Appendix 19. Adapters sequences used for ddRAD experiment. Asterisks represent modified nucleotides (phosphorothioate modification) to prevent nuclease degradation. Uppercase letters in DNA coding indicate adapter sequence while lower case letters indicate 6bp barcode sequence. N = complementary pairs of oligonucleotides to be annealed and used in the ligation step of DNA library preparation

ID	Sequence	N	Index
P1_PstI_top_i1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTgcatgaTGC*A	1	gcatga
P1_PstI_bottom_i1	/5Phos/tcatgcAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	1	tcatgc
P2_Mse_top_i1	/5Phos/TAgcatgaAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	1	gcatga
P2_Mse_bottom_i1	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTtcatg*c	1	tcatg
P1_PstI_top_i2	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTgatccgTGC*A	2	gatccg
P1_PstI_bottom_i2	/5Phos/cggatcAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	2	cggatc
P2_Mse_top_i2	/5Phos/TAgatccgAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	2	gatccg
P2_Mse_bottom_i2	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTcggat*c	2	cggat
P1_PstI_top_i3	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTtccagtTGC*A	3	tccagt
P1_PstI_bottom_i3	/5Phos/actggaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	3	actgga
P2_Mse_top_i3	/5Phos/TAtccagtAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	3	tccagt
P2_Mse_bottom_i3	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTactgg*a	3	actgg
P1_PstI_top_i4	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTgtctta TGC*A	4	gtctta
P1_PstI_bottom_i4	/5Phos/taagacAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	4	taagac
P2_Mse_top_i4	/5Phos/TAgtctta AGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	4	gtctta
P2_Mse_bottom_i4	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTtaaga*c	4	taaga
P1_PstI_top_i5	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTcggagtTGC*A	5	cggagt
P1_PstI_bottom_i5	/5Phos/actccgAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	5	actccg
P2_Mse_top_i5	/5Phos/TAcggagtAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	5	cggagt
P2_Mse_bottom_i5	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTactcc*g	5	actcc
P1_PstI_top_i6	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTcacgttTGC*A	6	cacgtt
P1_PstI_bottom_i6	/5Phos/aacgtgAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	6	aacgtg
P2_Mse_top_i6	/5Phos/TAcacgttAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	6	cacgtt
P2_Mse_bottom_i6	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTaacgt*g	6	aacgt

P1_PstI_top_i7	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTatacagTGC*A	7	atacag
P1_PstI_bottom_i7	/5Phos/ctgtatAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	7	ctgtat
P2_Mse_top_i7	/5Phos/TAatacagAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	7	atacag
P2_Mse_bottom_i7	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTctgta*t	7	ctgta
P1_PstI_top_i8	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTtgttacTGC*A	8	tgttac
P1_PstI_bottom_i8	/5Phos/gtaacaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	8	gtaaca
P2_Mse_top_i8	/5Phos/TAtgttacAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	8	tgttac
P2_Mse_bottom_i8	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTgtaac*a	8	gtaac
P1_PstI_top_i9	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTacgctcTGC*A	9	acgctc
P1_PstI_bottom_i9	/5Phos/gagcgtAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	9	gagcgt
P2_Mse_top_i9	/5Phos/TAacgctcAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	9	acgctc
P2_Mse_bottom_i9	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTgagcg*t	9	gagcg
P1_PstI_top_i10	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTttggcaTGC*A	10	ttggca
P1_PstI_bottom_i10	/5Phos/tgccaaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	10	tgccaa
P2_Mse_top_i10	/5Phos/TAttggcaAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	10	ttggca
P2_Mse_bottom_i10	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTtgcca*a	10	tgcca
P1_PstI_top_i11	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTagtaacTGC*A	11	agtaac
P1_PstI_bottom_i11	/5Phos/gttactAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	11	gttact
P2_Mse_top_i11	/5Phos/TAagtaacAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	11	agtaac
P2_Mse_bottom_i11	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTgttac*t	11	gttac
P1_PstI_top_i12	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTcaagcgTGC*A	12	caagcg
P1_PstI_bottom_i12	/5Phos/cgcttgAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	12	cgcttg
P2_Mse_top_i12	/5Phos/TAcaagcgAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	12	caagcg
P2_Mse_bottom_i12	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTcgctt*g	12	cgctt
P1_PstI_top_i13	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTatgctaTGC*A	13	atgcta
P1_PstI_bottom_i13	/5Phos/tagcatAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T	13	tagcat
P2_Mse_top_i13	/5Phos/TAatgctaAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA	13	atgcta
P2_Mse_bottom_i13	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTtagca*t	13	tagca

Appendix 20. Linear regression model output of global trait differentiation in C. solstitialis. χ2 – chi-squared distribution. AIC - Akaike information criterion**.** Significance levels: "*" P ≤ 0.05 significant, "**" P ≤ 0.01 highly significant, "***" P ≤ 0.001 extremely significant

Trait	χ2	P	AIC null model	AIC full model
Days to bolting	9.8493	0.079	2766.7	2766.5
Days to flower	18.448	0.003**	2514.1	2521.6
Number of inflorescence	16.55	0.005**	3645.2	3651.7
Spine length	17.044	0.004**	450.04	457.09
Number of ovules	13.437	0.019*	1443	1446.4
Mass of plumed seeds	19.764	0.001**	162.89	172.65
Final plant height	12.754	0.025*	3031.5	3034.3

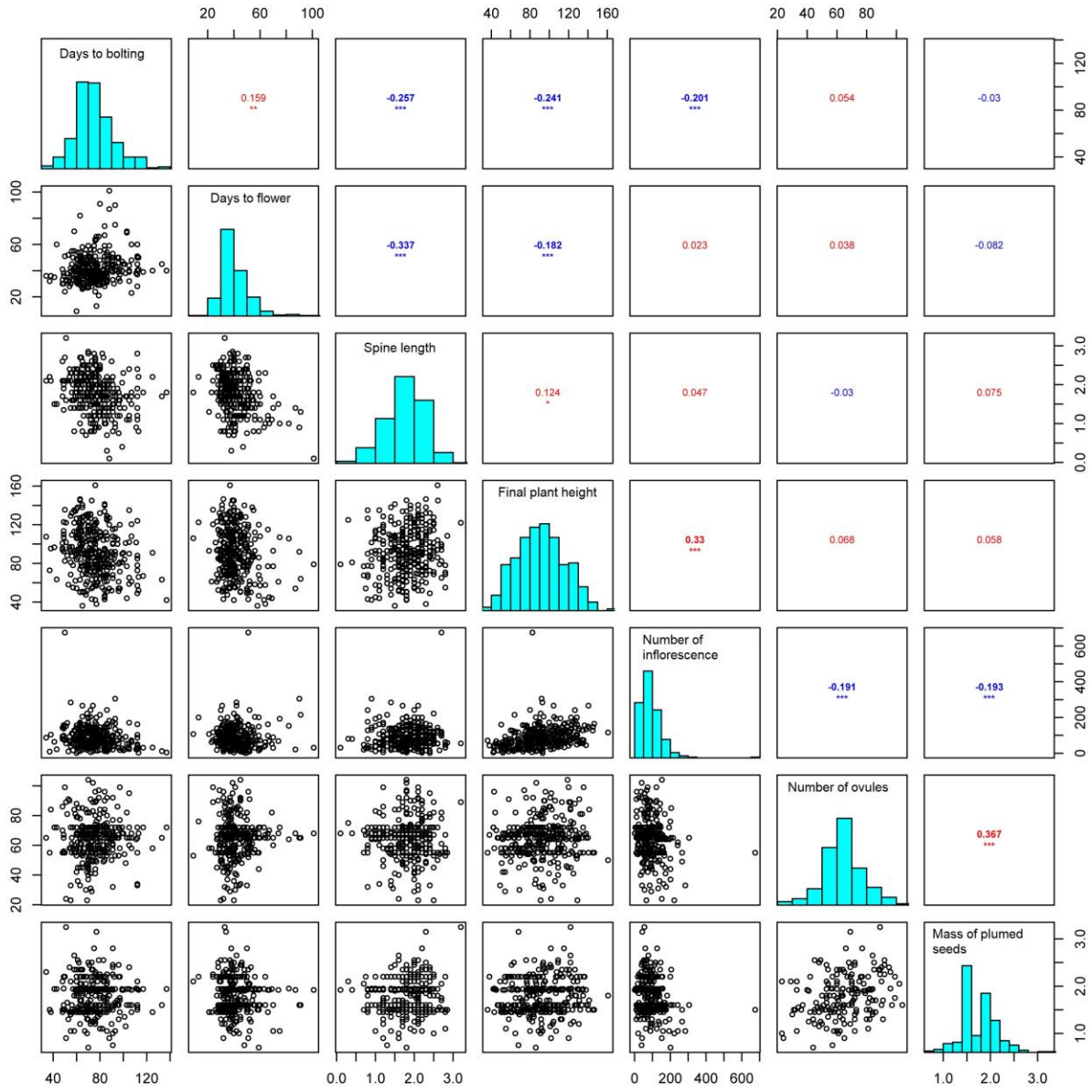
Appendix 21. Linear logistic regression model output for comparison of traits between C. solstitialis regions. Z value = measure of standard deviation, Pr = probability that the null hypothesis has been rejected

Trait	**Comparison**	**Estimate**	**Std. Error**	**Z value**	**Pr (>|z|)**
Days to bolting	Spain vs Turkey	6.567	2.733	2.402	0.152
	Argentina vs Turkey	2.687	2.704	0.994	0.918
	Chile vs Turkey	-1.071	3.606	-0.297	0.999
	California vs Turkey	-1.280	2.924	-0.438	0.997
	Australia vs Turkey	5.090	3.029	1.680	0.540
	Argentina vs Spain	-3.879	2.704	-1.435	0.701
	Chile vs Spain	-7.638	3.606	-2.118	0.273
	California vs Spain	-7.847	2.924	-2.684	0.076.
	Australia vs Spain	-1.476	3.029	-0.487	0.886
	Chile vs Argentina	-3.758	3.584	-1.049	0.899
	California vs Argentina	-3.968	2.896	-1.370	0.741
	Australia vs Argentina	2.403	3.003	0.800	0.966
	California vs Chile	-0.209	3.752	-0.056	1.000
	Australia vs Chile	6.161	3.836	1.606	0.589
	Australia vs California	6.371	3.202	1.989	0.343
Days to flower	Spain vs Turkey	4.289	1.834	2.339	0.175
	Argentina vs Turkey	-0.627	1.821	-0.345	0.999
	Chile vs Turkey	-3.091	2.429	-1.273	0.796
	California vs Turkey	-2.640	1.969	-1.341	0.758
	Australia vs Turkey	7.773	2.040	3.810	0.001**

	Argentina vs Spain	-4.917	1.814	-2.710	0.071.
	Chile vs Spain	-7.380	2.423	-3.045	0.027*
	California vs Spain	-6.929	1.963	-3.530	0.005**
	Australia vs Spain	3.484	2.034	1.713	0.518
	Chile vs Argentina	-2.463	2.414	-1.021	0.909
	California vs Argentina	-2.012	1.950	-1.032	0.905
	Australia vs Argentina	8.401	2.022	4.154	< 0.001 ***
	California vs Chile	0.450	2.527	0.178	0.999
	Australia vs Chile	10.864	2.583	4.206	< 0.001 ***
	Australia vs California	10.413	2.156	4.828	< 0.001 ***
Number of inflorescence	Spain vs Turkey	-9.879	10.179	-0.970	0.925
	Argentina vs Turkey	-46.775	10.107	-4.628	<0.001***
	Chile vs Turkey	-54.524	13.480	-4.045	<0.001***
	California vs Turkey	-30.265	10.929	-2.769	0.061.
	Australia vs Turkey	-42.725	11.323	-3.773	0.002**
	Argentina vs Spain	-36.896	10.069	-3.664	0.003**
	Chile vs Spain	-44.645	13.451	-3.319	0.011*
	California vs Spain	-20.387	10.894	-1.871	0.415
	Australia vs Spain	-32.846	11.289	-2.910	0.041*
	Chile vs Argentina	-7.749	13.396	-0.578	0.992
	California vs Argentina	16.509	10.826	1.525	0.643
	Australia vs Argentina	4.050	11.224	0.361	0.999
	California vs Chile	24.259	14.027	1.729	0.507
	Australia vs Chile	11.800	14.336	0.823	0.962
	Australia vs California	-12.969	11.969	-1.041	0.902
Number of ovules	Spain vs Turkey	9.875	3.918	2.520	0.113
	Argentina vs Turkey	17.256	3.762	4.587	<0.001***
	Chile vs Turkey	2.681	5.878	0.456	0.997
	California vs Turkey	11.874	4.091	2.903	0.041*
	Australia vs Turkey	13.502	4.655	2.901	0.041*
	Argentina vs Spain	7.381	3.847	1.918	0.381
	Chile vs Spain	-7.194	5.933	-1.212	0.825
	California vs Spain	1.999	4.169	0.479	0.996
	Australia vs Spain	3.627	4.724	0.768	0.971
	Chile vs Argentina	-14.575	5.831	-2.499	0.119
	California vs Argentina	-5.382	4.023	-1.338	0.756
	Australia vs Argentina	-3.754	4.595	-0.817	0.963
	California vs Chile	9.193	6.048	1.520	0.642
	Australia vs Chile	10.821	6.444	1.679	0.536
	Australia vs California	1.628	4.868	0.334	0.999
Mass of plumed seeds	Spain vs Turkey	0.113	0.106	1.070	0.888
	Argentina vs Turkey	0.464	0.099	4.677	<0.001***
	Chile vs Turkey	-0.032	0.160	-0.205	0.999

	California vs Turkey	0.738	0.107	6.890	<0.001***
	Australia vs Turkey	0.445	0.130	3.427	0.007**
	Argentina vs Spain	0.350	0.103	3.375	0.009**
	Chile vs Spain	-0.146	0.163	-0.895	0.945
	California vs Spain	0.624	0.111	5.603	<0.001***
	Australia vs Spain	0.332	0.133	2.486	0.122
	Chile vs Argentina	-0.497	0.159	-3.126	0.020*
	California vs Argentina	0.274	0.105	2.608	0.090.
	Australia vs Argentina	0.018	0.128	-0.145	0.999
	California vs Chile	0.771	0.164	4.699	<0.001***
	Australia vs Chile	0.478	0.179	2.660	0.079.
	Australia vs California	-0.292	0.134	-2.175	0.240
Spine length	Spain vs Turkey	-0.220	0.083	-2.648	0.084.
	Argentina vs Turkey	-0.073	0.081	-0.906	0.943
	Chile vs Turkey	0.041	0.109	0.382	0.998
	California vs Turkey	0.008	0.087	0.094	1.000
	Australia vs Turkey	-0.440	0.091	-4.812	<0.001***
	Argentina vs Spain	0.146	0.082	1.774	0.477
	Chile vs Spain	0.262	0.110	2.372	0.163
	California vs Spain	0.228	0.088	2.572	0.102
	Australia vs Spain	-0.219	0.092	-2.371	0.163
	Chile vs Argentina	0.115	0.109	1.059	0.895
	California vs Argentina	0.082	0.087	0.941	0.934
	Australia vs Argentina	-0.366	0.090	-4.029	<0.001***
	California vs Chile	-0.033	0.114	-0.295	0.999
	Australia vs Chile	-0.482	0.116	-4.123	<0.001***
	Australia vs California	-0.448	0.096	-4.640	<0.001***
Final plant height	Spain vs Turkey	1.485	4.130	0.360	0.999
	Argentina vs Turkey	-0.345	4.101	-0.084	1.000
	Chile vs Turkey	-16.933	5.470	-3.096	0.023*
	California vs Turkey	-6.803	4.434	-1.534	0.637
	Australia vs Turkey	-0.722	4.656	-0.155	1.000
	Argentina vs Spain	-1.831	4.085	-0.448	0.997
	Chile vs Spain	-18.418	5.458	-3.374	0.009**
	California vs Spain	-8.289	4.420	-1.875	0.412
	Australia vs Spain	-2.208	4.742	-0.476	0.996
	Chile vs Argentina	-16.587	5.436	-3.051	0.026*
	California vs Argentina	-6.458	4.393	-1.470	0.679
	Australia vs Argentina	-0.377	4.616	-0.082	1.000
	California vs Chile	10.129	5.692	1.780	0.474
	Australia vs Chile	16.210	5.866	2.763	0.062.
	Australia vs California	6.081	4.915	1.237	0.815

Appendix 22. Character correlation between traits. The x axis in each scatter plot represents the column variable, the y axis the row variable. The value of the correlation and the significance level is indicated by asterisks (0.1 = ".", 0.05 = "*", 0.01 = "**", 0.001 = ("***") on the top of the diagonal. Positive correlations are displayed in red and negative correlations in blue color



Appendix 23. P_{ST}-F_{ST} pair-wise region comparisons of phenotypic and neutral genetic differentiation at seven morphological traits in *C. solstitialis*. The significance of the differences for each comparison are indicated by the Bayesian credibility interval and highlighted in bold

Trait	comparison	mean P_{ST} (95% CI)	mean F_{ST} (95% CI)	P_{ST} - F_{ST} Bayesian credibility interval (95% CI)
Days to bolt	Turkey *vs* Spain	0.02 (0.00 – 0.10)	0.08 (0.06 – 0.11)	-0.06 (-0.08 – 0.01)
Days to bolt	Turkey *vs* Argentina	0.00 (0.00 – 0.03)	0.07 (0.05 – 0.10)	-0.07 (-0.07 – (-0.04)
Days to bolt	Turkey *vs* Chile	0.00 (0.00 – 0.02)	0.08 (0.06 – 0.11)	-0.08 (-0.08 – (-0.05)
Days to bolt	Turkey *vs* California	0.00 (0.00 – 0.02)	0.06 (0.04 – 0.08)	-0.06 (-0.06 – (-0.04)
Days to bolt	Turkey *vs* Australia	0.01 (0.00 – 0.08)	0.08 (0.06 – 0.11)	-0.07 (-0.08 – (-0.00)
Days to bolt	Spain *vs* Argentina	0.00 (0.00 – 0.04)	0.02 (0.01 – 0.04)	-0.02 (-0.02 – 0.01)
Days to bolt	Spain *vs* Chile	0.02 (0.00 – 0.13)	0.04 (0.02 – 0.06)	-0.01 (-0.04 – 0.09)
Days to bolt	Spain *vs* California	0.05 (0.00 – 0.14)	0.03 (0.02 – 0.05)	0.01 (-0.03 – 0.11)
Days to bolt	Spain *vs* Australia	0.00 (0.00 – 0.02)	0.03 (0.02 – 0.05)	-0.03 (-0.03 – (-0.01)
Days to bolt	Argentina *vs* Chile	0.06 (0.00 – 0.05)	0.02 (0.01 – 0.04)	-0.02 (-0.027 – 0.02)
Days to bolt	Argentina *vs* California	0.00 (0.00 – 0.05)	0.02 (0.01 – 0.03)	-0.01 (-0.02 – 0.03)
Days to bolt	Argentina *vs* Australia	0.00 (0.00 – 0.03)	0.02 (0.00 – 0.03)	-0.01 (-0.02 – 0.01)
Days to bolt	Chile *vs* California	0.00 (0.00 – 0.02)	0.03 (0.02 – 0.05)	-0.03 (-0.03 – (-0.00)
Days to bolt	Chile *vs* Australia	0.01 (0.00 – 0.11)	0.03 (0.02 –0.05)	-0.01 (-0.33 – 0.07)
Days to bolt	California *vs* Australia	0.03 (0.00 – 0.12)	0.02 (0.01– 0.04)	0.00 (-0.02 – 0.10)
Days to flower	Turkey *vs* Spain	0.01 (0.00 – 0.08)	0.08 (0.06 – 0.11)	-0.06 (-0.08 – (-0.00)
Days to flower	Turkey *vs* Argentina	0.00 (0.00 – 0.01)	0.07 (0.05 – 0.10)	-0.07 (-0.07 – (-0.06)
Days to flower	Turkey *vs* Chile	0.01 (0.00 – 0.08)	0.08 (0.06 – 0.11)	-0.07 (-0.08 – (-0.00)
Days to flower	Turkey *vs* California	0.00 (0.00 – 0.06)	0.06 (0.04 – 0.08)	-0.05 (-0.06 – (-0.00)
Days to flower	Turkey *vs* Australia	0.09 (0.00 – 0.20)	0.08 (0.06 – 0.11)	0.00 (-0.08 – 0.12)
Days to flower	Spain *vs* Argentina	0.02 (0.00 – 0.10)	0.02 (0.01 – 0.04)	-0.00 (-0.02 – 0.07)
Days to flower	Spain *vs* Chile	0.05 (0.00 – 0.18)	0.04 (0.02 – 0.06)	0.01 (-0.04 – 0.14)
Days to flower	Spain *vs* California	0.06 (0.00 – 0.17)	0.03 (0.02 – 0.05)	0.03 (-0.03 – 0.13)

Days to flower	Spain *vs* Australia	0.00 (0.00 – 0.05)	0.03 (0.02 – 0.05)	-0.03 (-0.03 – 0.01)
Days to flower	Argentina *vs* Chile	0.00 (0.00 – 0.06)	0.02 (0.01 – 0.04)	-0.01 (-0.02 – 0.03)
Days to flower	Argentina *vs* California	0.00 (0.00 – 0.04)	0.02 (0.01 – 0.03)	-0.01 (-0.02 – 0.02)
Days to flower	Argentina *vs* Australia	0.10 (0.01 – 0.22)	0.02 (0.00 – 0.03)	0.08 (-0.00 – 0.20)
Days to flower	Chile *vs* California	0.00 (0.00 – 0.02)	0.03 (0.02 – 0.05)	-0.03 (-0.03 – (-0.00)
Days to flower	Chile *vs* Australia	0.14 (0.00 – 0.30)	0.03 (0.02 – 0.05)	0.11 (-0.02 – 0.27)
Days to flower	**California *vs* Australia**	**0.16 (0.04 – 0.29)**	**0.02 (0.01 – 0.04)**	**0.13 (0.01 – 0.27)**
Final plant height	Turkey *vs* Spain	0.00 (0.00 – 0.02)	0.08 (0.06 – 0.11)	-0.08 (-0.08 – (-0.06)
Final plant height	Turkey *vs* Argentina	0.00 (0.00 – 0.02)	0.07 (0.05 – 0.10)	-0.07 (-0.07 – (-0.05)
Final plant height	Turkey *vs* Chile	0.08 (0.00 – 0.22)	0.08 (0.06 – 0.11)	0.00 (-0.08 – 0.13)
Final plant height	Turkey *vs* California	0.01 (0.00 – 0.06)	0.06 (0.04 – 0.08)	-0.05 (-0.06 – 0.00)
Final plant height	Turkey *vs* Australia	0.00 (0.00 – 0.02)	0.08 (0.06 – 0.11)	-0.08 (-0.08 – (-0.06)
Final plant height	Spain *vs* Argentina	0.00 (0.00 – 0.02)	0.02 (0.01 – 0.04)	-0.02 (-0.02 – 0.00)
Final plant height	Spain *vs* Chile	0.12 (0.00 – 0.26)	0.04 (0.02 – 0.06)	0.07 (-0.03 – 0.22)
Final plant height	Spain *vs* California	0.01 (0.00 – 0.08)	0.03 (0.02 – 0.05)	-0.01 (-0.03 – 0.05)
Final plant height	Spain *vs* Australia	0.00 (0.00 – 0.03)	0.03 (0.02 – 0.05)	-0.03 (-0.03 – (-0.00)
Final plant height	Argentina *vs* Chile	0.10 (0.00 – 0.24)	0.02 (0.01 – 0.04)	0.08 (-0.02 – 0.22)
Final plant height	Argentina *vs* California	0.01 (0.00 – 0.07)	0.02 (0.01 – 0.03)	-0.00 (-0.02 – 0.04)
Final plant height	Argentina *vs* Australia	0.00 (0.00 – 0.02)	0.02 (0.00 – 0.03)	-0.01 (-0.02 – 0.00)
Final plant height	Chile *vs* California	0.02 (0.00 – 0.13)	0.03 (0.02 – 0.05)	-0.00 (-0.03 – 0.10)
Final plant height	Chile *vs* Australia	0.13 (0.00 – 0.28)	0.03 (0.02 – 0.05)	0.09 (-0.02 – 0.25)
Final plant height	California *vs* Australia	0.00 (0.00 – 0.06)	0.02 (0.01 – 0.04)	-0.01 (-0.02 – 0.04)
Number of inflorescence	Turkey *vs* Spain	0.00 (0.00 – 0.02)	0.08 (0.06 – 0.11)	-0.08 (-0.08 – (-0.05)
Number of inflorescence	Turkey *vs* Argentina	0.07 (0.00 – 0.16)	0.07 (0.05 – 0.10)	-0.00 (-0.07 – 0.08)
Number of inflorescence	Turkey *vs* Chile	0.09 (0.00 – 0.23)	0.08 (0.06 – 0.11)	0.00 (-0.08– 0.14)
Number of inflorescence	Turkey *vs* California	0.01 (0.00 – 0.06)	0.06 (0.04 – 0.08)	-0.05 (-0.06 – (-0.00)
Number of inflorescence	Turkey *vs* Australia	0.05 (0.00 – 0.15)	0.08 (0.06 – 0.11)	-0.03 (-0.08 – 0.06)
Number of inflorescence	Spain *vs* Argentina	0.04 (0.00 – 0.13)	0.02 (0.01 – 0.04)	0.02 (-0.02 – 0.10)
Number of inflorescence	Spain *vs* Chile	0.06 (0.00 – 0.19)	0.04 (0.02 – 0.06)	0.02 (-0.04 – 0.15)
Number of inflorescence	Spain *vs* California	0.00 (0.00 – 0.03)	0.03 (0.02 – 0.05)	-0.02 (-0.03 – 0.00)

Number of inflorescence	Spain *vs* Australia	0.03 (0.00 – 0.12)	0.03 (0.02 – 0.05)	-0.00 (-0.03 – 0.08)
Number of inflorescence	Argentina *vs* Chile	0.00 (0.00 – 0.04)	0.02 (0.01 – 0.04)	-0.01 (-0.02 – 0.02)
Number of inflorescence	Argentina *vs* California	0.03 (0.00 – 0.12)	0.02 (0.01 – 0.03)	0.01 (-0.02 – 0.10)
Number of inflorescence	Argentina *vs* Australia	0.00 (0.00 – 0.02)	0.02 (0.00 – 0.03)	-0.01 (-0.02 – 0.00)
Number of inflorescence	Chile *vs* California	0.07 (0.00 – 0.21)	0.03 (0.02 – 0.05)	0.04 (-0.03 – 0.18)
Number of inflorescence	Chile *vs* Australia	0.00 (0.00 – 0.05)	0.03 (0.02 – 0.05)	-0.02 (-0.03 – 0.01)
Number of inflorescence	California *vs* Australia	0.02 (0.00 – 0.11)	0.02 (0.01 – 0.04)	0.00 (-0.02 – 0.08)
Number of ovules	Turkey *vs* Spain	0.04 (0.00 – 0.17)	0.08 (0.06 – 0.11)	-0.04 (-0.08 – 0.08)
Number of ovules	Turkey *vs* Argentina	0.16 (0.02 – 0.31)	0.07 (0.05 – 0.10)	0.08 (-0.05 – 0.23)
Number of ovules	Turkey *vs* Chile	0.00 (0.00 – 0.06)	0.08 (0.06 – 0.11)	-0.08 (-0.08 – (-0.02)
Number of ovules	Turkey *vs* California	0.07 (0.00 – 0.22)	0.06 (0.04 – 0.08)	0.01 (-0.06 – 0.16)
Number of ovules	Turkey *vs* Australia	0.07 (0.00 – 0.25)	0.08 (0.06 – 0.11)	-0.01 (-0.08 – 0.16)
Number of ovules	Spain *vs* Argentina	0.02 (0.00 – 0.13)	0.02 (0.01 – 0.04)	-0.00 (-0.02 – 0.10)
Number of ovules	Spain *vs* Chile	0.02 (0.00 – 0.16)	0.04 (0.02 – 0.06)	-0.02 (-0.04 – 0.12)
Number of ovules	Spain *vs* California	0.00 (0.00 – 0.04)	0.03 (0.02 – 0.05)	-0.03 (-0.03 – 0.00)
Number of ovules	Spain *vs* Australia	0.00 (0.00 – 0.06)	0.03 (0.02 – 0.05)	-0.02 (-0.03 – 0.02)
Number of ovules	Argentina *vs* Chile	0.09 (0.00 – 0.32)	0.02 (0.01 – 0.04)	0.06 (-0.02 – 0.29)
Number of ovules	Argentina *vs* California	0.01 (0.00 – 0.09)	0.02 (0.01 – 0.03)	-0.00 (-0.02 – 0.06)
Number of ovules	Argentina *vs* Australia	0.00 (0.00 – 0.06)	0.02 (0.00 – 0.03)	-0.01 (-0.02 – 0.04)
Number of ovules	Chile *vs* California	0.03 (0.00 – 0.23)	0.03 (0.02 – 0.05)	0.00 (-0.03 – 0.19)
Number of ovules	Chile *vs* Australia	0.03 (0.00 – 0.24)	0.03 (0.02 – 0.05)	0.00 (-0.03 – 0.21)
Number of ovules	California *vs* Australia	0.00 (0.00 – 0.04)	0.02 (0.01 – 0.04)	-0.01 (-0.02 – 0.02)
Spine length	Turkey *vs* Spain	0.04 (0.00 – 0.12)	0.08 (0.06 – 0.11)	-0.04 (-0.08 – 0.04)
Spine length	Turkey *vs* Argentina	0.00 (0.00 – 0.04)	0.07 (0.05 – 0.10)	-0.06 (-0.07 – (-0.03)
Spine length	Turkey *vs* Chile	0.00 (0.00 – 0.04)	0.08 (0.06 – 0.11)	-0.08 (-0.08 – (-0.04)
Spine length	Turkey *vs* California	0.00 (0.00 – 0.02)	0.06 (0.04 – 0.08)	-0.05 (-0.06 – (-0.03)
Spine length	Turkey *vs* Australia	0.13 (0.03 – 0.26)	0.08 (0.06 – 0.11)	0.05 (-0.05 – 0.17)
Spine length	Spain *vs* Argentina	0.02 (0.00 – 0.08)	0.02 (0.01 – 0.04)	-0.00 (-0.02 – 0.05)
Spine length	Spain *vs* Chile	0.06 (0.00 – 0.19)	0.04 (0.02 – 0.06)	0.02 (-0.04 – 0.14)
Spine length	Spain *vs* California	0.04 (0.00 – 0.14)	0.03 (0.02 – 0.05)	0.01 (-0.03 – 0.10)

Spine length	Spain *vs* Australia	0.03 (0.00 – 0.12)	0.03 (0.02 – 0.05)	0.00 (-0.03 – 0.09)
Spine length	Argentina *vs* Chile	0.01 (0.00 – 0.08)	0.02 (0.01 – 0.04)	-0.01 (-0.02 – 0.05)
Spine length	Argentina *vs* California	0.00 (0.00 – 0.05)	0.02 (0.01 – 0.03)	-0.01 (-0.02 – 0.02)
Spine length	Argentina *vs* Australia	0.11 (0.01 – 0.22)	0.02 (0.00 – 0.03)	0.08 (-0.00 – 0.20)
Spine length	Chile *vs* California	0.00 (0.00 – 0.04)	0.03 (0.02 – 0.05)	-0.02 (-0.03 – 0.01)
Spine length	Chile *vs* Australia	0.15 (0.01 – 0.31)	0.03 (0.02 – 0.05)	0.11 (-0.01 – 0.27)
Spine length	**California *vs* Australia**	**0.13 (0.03 – 0.27)**	**0.02 (0.01 – 0.04)**	**0.11 (0.00 – 0.24)**
Mass of plumed seeds	Turkey *vs* Spain	0.01 (0.00 – 0.10)	0.08 (0.06 – 0.11)	-0.06 (-0.08 – 0.01)
Mass of plumed seeds	Turkey *vs* Argentina	0.20 (0.05 – 0.37)	0.07 (0.05 – 0.10)	0.12 (-0.02 – 0.29)
Mass of plumed seeds	Turkey *vs* Chile	0.01 (0.00 – 0.11)	0.08 (0.06 – 0.11)	-0.06 (-0.08 – 0.02)
Mass of plumed seeds	**Turkey *vs* California**	**0.38 (0.19 – 0.54)**	**0.06 (0.04 – 0.08)**	**0.31 (0.12 – 0.48)**
Mass of plumed seeds	Turkey *vs* Australia	0.18 (0.00 – 0.40)	0.08 (0.06 – 0.11)	0.10 (-0.07 – 0.31)
Mass of plumed seeds	Spain *vs* Argentina	0.12 (0.00 – 0.29)	0.02 (0.01 – 0.04)	0.09 (-0.02 – 0.26)
Mass of plumed seeds	Spain *vs* Chile	0.03 (0.00 – 0.19)	0.04 (0.02 – 0.06)	-0.00 (-0.04 – 0.15)
Mass of plumed seeds	**Spain *vs* California**	**0.31 (0.12 – 0.50)**	**0.03 (0.02 – 0.05)**	**0.28 (0.08 – 0.46)**
Mass of plumed seeds	Spain *vs* Australia	0.12 (0.00 – 0.33)	0.03 (0.02 – 0.05)	0.08 (-0.03 – 0.29)
Mass of plumed seeds	Argentina *vs* Chile	0.20 (0.00 – 0.46)	0.02 (0.01 –0.04)	0.17 (-0.02 – 0.43)
Mass of plumed seeds	Argentina *vs* California	0.06 (0.00 – 0.20)	0.02 (0.01 – 0.03)	0.04 (-0.02 – 0.18)
Mass of plumed seeds	Argentina *vs* Australia	0.01 (0.00 – 0.07)	0.02 (0.00 – 0.03)	-0.00 (-0.02 – 0.05)
Mass of plumed seeds	**Chile *vs* California**	**0.37 (0.08 – 0.61)**	**0.03 (0.02 – 0.05)**	**0.33 (0.04 – 0.57)**
Mass of plumed seeds	Chile *vs* Australia	0.22 (0.00 – 0.52)	0.03 (0.02 – 0.05)	0.19 (-0.03 – 0.49)
Mass of plumed seeds	California *vs* Australia	0.06 (0.00 – 0.25)	0.02 (0.01 – 0.04)	0.04 (-0.02 – 0.22)

Appendix 24. Evanno method using delta K (rate of change in likelihood among models) to differentiate among number of populations inferred from structure analysis. K value demonstrating peak in delta K is bolded

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	ΔK
1	10	-181358.960000	2.341509	—	—	—
2	**10**	**-176524.470000**	**1.449176**	**4834.490000**	**2672.830000**	**1844.379137**
3	10	-174362.810000	8.395693	2161.660000	242.860000	28.926735
4	10	-172444.010000	10.642833	1918.800000	542.550000	50.977969
5	10	-171067.760000	11.445737	1376.250000	508.520000	44.428770
6	10	-170200.030000	127.892733	867.730000	145.810000	1.140096
7	10	-169478.110000	877.618196	721.920000	288.620000	0.328867
8	10	-168467.570000	33.725725	1010.540000	—	—

Appendix 25. Outlier loci identified by BayeScan, OutFLANK and PCAdapt. Locus naming system: digits before the low dash represent the tag number, digits after the low dash represent the position of SNP on the tag. Sequence information are available in NCBI (data will be submitted by publication)

Names	Total Loci	Locus name
BayeScan OutFLANK PCAdapt	3	27727_100, 1225_114, 29308_43
BayeScan OutFLANK	3	10004_58, 16569_113, 22138_114
BayeScan PCAdapt	1	4600_75
OutFLANK PCAdapt	12	12536_93, 22138_85, 4126_107, 801_101, 22566_60, 29948_96, 9430_93, 7770_46, 7770_73, 11870_114, 21765_16, 23976_118
BayeScan	4	737_45, 2498_84, 20745_61, 12279_26
OutFLANK	18	8691_20, 1911_54, 22973_109, 18176_10, 1911_94, 5441_75, 2719_83, 17802_114, 3854_32, 11958_91, 7306_115, 28262_8, 20213_10, 24865_33, 27727_14, 11958_56, 5441_108,1911_117

PCAdapt	122	24602_55, 46333_98, 22040_78, 770_10, 5906_37, 30939_83, 80_49, 3370_129, 25633_94, 26165_52, 925_39, 12147_112, 9744_90, 6768_68, 24577_80, 23302_66, 24833_108, 3276_66, 14843_20, 11156_91, 4600_96, 437_32, 17425_51, 25212_66, 22138_40, 3370_26, 1477_85, 3370_66, 11870_82, 3162_16, 516_117, 5527_45, 2128_46, 11240_47, 21609_40, 45761_59, 18473_10, 11870_88, 31478_28, 24339_44, 14280_107, 11495_74, 28398_45, 2669_72, 26271_111, 21567_8, 28110_85, 1917_72, 617949_57, 12289_23, 3790_34, 1065_80, 9708_54, 16866_37, 28320_58, 8464_76, 24142_70, 28110_69, 18788_42, 7421_8, 26276_25, 11821_121, 22176_78, 6595_91, 12536_81, 1335_96, 6370_22, 743_75, 17970_106, 26881_24, 39255_47, 16686_18, 6027_57, 192_96, 3855_114, 12059_86, 25618_7, 12536_42, 28589_110, 27027_83, 27147_49, 24009_74, 883207_47, 8464_112, 925_21, 29502_92, 12841_30, 13492_85, 879292_10, 38172_26, 25159_39, 22388_112, 26229_68, 5487_99, 1335_113, 408_95, 3642_31, 3973_37, 17970_96, 7374_6, 23469_96, 17863_76, 13086_10, 3642_13, 21077_98, 21077_22, 23302_30, 39255_12, 24731_61, 2004_35, 2004_90, 20907_98, 10084_49, 12535_63, 28631_28, 1024_74, 12466_64, 28592_74, 11471_87, 29902_92, 31330_89, 17147_7

Appendix 26. Gene ontology annotation of *Centaurea solstitialis* outlier genes based on The Arabidopsis Information Resource (https://www.arabidopsis.org/)

Locus ID	Method	SeqName	Description	Biological process	Cellular component	Molecular function	Growth and development stages	Plant structure
10004_58	OutFlank_Bayescan	WA_CGP_AR-13-24_1159	Serine carboxypeptidase S28 family protein	mega-gametogenesis, proteolysis	extracellular region, vacuole	dipeptidyl-peptidase activity, serine-type peptidase activity	L mature pollen stage, M germinated pollen stage, flowering stage	flower, pollen, pollen tube cell
16569_113	OutFlank_Bayescan	WA_CGP_AR-13-24_1085	Radical SAM superfamily protein	rRNA base methylation, tRNA methylation	cytoplasm	RNA methyltransferase activity, iron-sulfur cluster binding	LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage, LP.08 eight leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, M germinated pollen stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	carpel, cauline leaf, collective leaf structure, cotyledon, flower, flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, petal, petiole, plant embryo, plant sperm cell, pollen, root, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf
22138_85	OutFlank_Bayescan	WA_CGP_AR-13-24_6633	Chaperone DnaJ- domain superfamily protein	chaperone cofactor-dependent protein refolding	cytoplasm, cytosol	chaperone binding, unfolded protein binding	LP.02 two leaves visible stage, LP.06 six leaves visible stage, LP.10 ten leaves visible stage, M germinated pollen stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	carpel, cauline leaf, collective leaf structure, flower, flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf lamina base, petal, plant embryo, plant sperm cell, pollen, root, seed, sepal, shoot system, stamen, stem, vascular leaf
22138_114	OutFlank_Bayescan	WA_CGP_AR-13-24_6633	Chaperone DnaJ- domain	chaperone cofactor-dependent	cytoplasm, cytosol	chaperone binding,	LP.02 two leaves visible stage, LP.06 six leaves visible stage, LP.10 ten leaves visible stage,	carpel, cauline leaf, collective leaf structure, flower, flower pedicel,

			superfamily protein	protein refolding		unfolded protein binding	M germinated pollen stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	guard cell, hypocotyl, inflorescence meristem, leaf lamina base, petal, plant embryo, plant sperm cell, pollen, root, seed, sepal, shoot system, stamen, stem, vascular leaf
801_101	OutFlank_PCAdapt	AR_P13_24_c22586	Nucleolar/coiled-body phosphoprotein	biological process	chloroplast, nucleolus, nucleus	molecular function	LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage, LP.08 eight leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	carpel, cauline leaf, collective leaf structure, cotyledon, flower, flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, petal, petiole, plant embryo, root, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf
23976_118	OutFlank_PCAdapt	AR_P13_24_c23005	Neurofilament light protein	biological process	cytoplasm, cytosol, plastid	molecular function	LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage, LP.08 eight leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	carpel, cauline leaf, collective leaf structure, cotyledon, cultured plant cell, flower, flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, petal, petiole, plant embryo, pollen, root, rosette leaf, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf
29948_96	OutFlank_PCAdapt	AR_P13_24_s67867	Member of IQ67 (CaM binding)	biological process	mitochondrion, plasma membrane	molecular function	LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage,	carpel, cauline leaf, collective leaf structure, cotyledon, flower,

			domain containing family				LP.08 eight leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, petal, petiole, plant embryo, plant sperm cell, root, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf
12536_93	OutFlank_PCAdapt	WA_CGP_AR-13-24_9560	Transcription factor bHLH145	regulation of transcription, DNA-templated	chloroplast, nucleus	DNA binding, DNA-binding transcription factor activity, protein binding, protein dimerization activity	–	–
11870_114	OutFlank_PCAdapt	AR_P13_24_c11200	Encodes an FtsH protease that is localized to the chloroplast and the mitochondrion	PSII associated light-harvesting complex II catabolic process, proteolysis, response to heat	chloroplast, chloroplast envelope, chloroplast thylakoid membrane, integral component of membrane, mitochondrial inner membrane, mitochondrion, plastid	ATP binding, ATP-dependent peptidase activity, metal ion binding, metalloendopeptidase activity	LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage, LP.08 eight leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	carpel, cauline leaf, collective leaf structure, cotyledon, flower, flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, petal, petiole, plant callus, plant embryo, pollen, root, rosette leaf, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf
22566_60	OutFlank_PCAdapt	AR_P13_24_s50023	Encodes a homolog of COX15	heme a biosynthetic process, oxidation-reduction process	integral component of membrane, mitochondrial inner membrane, mitochondrion	oxidoreductase activity, acting on NAD(P)H, heme protein as acceptor, oxidoreductase activity, acting	LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage, LP.08 eight leaves visible stage, LP.10 ten leaves visible, LP.12 twelve leaves visible stage, flowering stage, mature plant embryo stage, petal	carpel, cauline leaf, collective leaf structure, cotyledon, flower, flower pedicel, guard cell, hypocotyl, leaf apex, leaf lamina base, petal, petiole, plant embryo, pollen, root,

						on the CH-CH group of donors	differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf
4126_107	OutFlank_PCAdapt	AR_P13_24_c21254	Weak chloroplast movement under blue light protein	chloroplast accumulation movement, chloroplast avoidance movement	cytosol	molecular function	L mature pollen stage, M germinated pollen stage, flowering stage, petal differentiation and expansion stage	carpel, collective leaf structure, flower, guard cell, petal, pollen, pollen tube cell
7770_46	OutFlank_PCAdapt	WA_CGP_AR-13-24_7344	Galactose oxidase kelch-repeat superfamily protein	biological process	chloroplast	molecular function	LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage, LP.08 eight leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	carpel, cauline leaf, collective leaf structure, cotyledon, flower, flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, petal, petiole, plant embryo, plant sperm cell, pollen, root, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf
7770_73	OutFlank_PCAdapt	WA_CGP_AR-13-24_7344	Galactose oxidase kelch-repeat superfamily protein	biological process	chloroplast	molecular function	LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage, LP.08 eight leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	carpel, cauline leaf, collective leaf structure, cotyledon, flower, flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, petal, petiole, plant embryo, plant sperm cell, pollen, root, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf
9430_93	OutFlank_PCAdapt	AR_P13_24_c39	Encodes a haem-binding	response to oxidative stress	chloroplast, chloroplast	heme binding, protein binding	LP.02 two leaves visible stage, LP.04 four visible stage, LP.06	carpel, cauline leaf, collective leaf structure,

			protein, HBP5. HBP5 binds haem and interacts with the haem oxygenase, HY1. Disrupting the binding of HBP5 to HY1 leads to oxidative stress		envelope, chloroplast thylakoid, chloroplast thylakoid membrane, cytoplasm, plastid		six leaves visible stage, LP.08 eight leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, flowering time, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	cotyledon, flower, flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, petal, petiole, plant embryo, pollen, rosette leaf, seed, shoot apex, shoot system, sepal, stamen, stem, vascular leaf
21765_16	OutFlank_PCAdapt	AR_P13_24_c9559	RNI-like superfamily protein	SCF-dependent proteasomal ubiquitin-dependent protein catabolic process, ubiquitin-dependent protein catabolic process	SCF ubiquitin ligase complex, cytoplasm	ubiquitin-protein transferase activity	LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage, LP.08 eight leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	carpel, cauline leaf, collective leaf structure, cotyledon, flower, flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, petal, petiole, plant embryo, pollen, root, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf
27727_100	OutFlank_PCAdapt_Bayescan	AR_P13_24_c6486	Transducin/WD40 repeat-like superfamily protein	biological process	Cul4-RING E3 ubiquitin ligase complex, nucleus	molecular function	LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage, LP.08 eight leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, M germinated pollen stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant	carpel, cauline leaf, collective leaf structure, cotyledon, flower, flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, petal, petiole, plant embryo, plant sperm cell, pollen, pollen tube cell, root, seed, sepal, shoot apex,

							embryo globular stage, vascular leaf senescent stage	shoot system, stamen, stem, vascular leaf
1225_114	OutFlank_PCAdapt_Bayescan	AR_P13_24_c11155	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	methylation	Golgi apparatus, Golgi trans cisterna, endosome, trans-Golgi network	methyl transferase activity	LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage, LP.08 eight leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	carpel, cauline leaf, collective leaf structure, cotyledon, flower, flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, petal, petiole, plant embryo, pollen, root, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf
29308_43	OutFlank_PCAdapt_Bayescan	AR_P13_24_c11469	SMAD/FHA domain-containing protein	biological process	nucleus	mRNA binding, protein phosphatase inhibitory activity	LP.04 four leaves visible stage, LP.12 twelve leaves visible stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	carpel, cauline leaf, collective leaf structure, cotyledon, flower, flower pedicel, guard cell, hypocotyl, inflorescence meristem, leaf lamina base, petal, plant embryo, root, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf
4600_75	PCAdapt_Bayescan	WA_CGP_AR-13-24_3855	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	metylation, rRNA processing	nucleolus, nucleus	S-adenosylmethionine dependent methyltransferase activity	LP.02 two leaves visible stage, LP.04 four leaves visible stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage	carpel, cauline leaf, collective leaf structure, flower, flower pedicel, guard cell, hypocotyl, leaf apex, petal, petiole, plant embryo, root, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf