University of Zagreb

Faculty of Science
Department of Biology

Sara Laura Šarančić

GENETIC BASIS OF ADAPTIVE DIVERGENCE IN AMETHYST MEADOW SQUILL (*Chouardia litardierei*, Hyacinthaceae)

DOCTORAL THESIS

Zagreb, 2025

University of Zagreb

Faculty of Science
Department of Biology

Sara Laura Šarančić

GENETIC BASIS OF ADAPTIVE DIVERGENCE IN AMETHYST MEADOW SQUILL (*Chouardia litardierei*, Hyacinthaceae)

DOCTORAL THESIS

Mentor:
Assoc. Prof. Ivan Radosavljević, PhD

Zagreb, 2025

Sveučilište u Zagrebu

Prirodoslovno-matematički fakultet
Biološki odsjek

Sara Laura Šarančić

GENSKA OSNOVA ADAPTIVNE DIVERGENCIJE LIVADNOG PROCJEPKA (*Chouardia litardierei*, Hyacinthaceae)

DOKTORSKI RAD

Mentor:
izv. prof. dr. sc. Ivan Radosavljević

Zagreb, 2025

This doctoral dissertation was made at the University of Zagreb, Faculty of Science, Department of Biology, under the supervision of Associate Professor Ivan Radosavljević, PhD, as part of the Postgraduate program of Biology at the Department of Biology, Faculty of Science, University of Zagreb. The doctoral thesis was prepared within the Croatian Science Foundation scientific research project "Amethyst meadow squill (*Chouardia litardierei*, Hyacinthaceae): a study system for ecological divergence" (HRZZ IP-2020-02-8099) led by Associate Professor Ivan Radosavljević, PhD.

Acknowledgements

I would like to begin by expressing my gratitude to my supervisor, Ivan Radosavljević, for his guidance, support, and trust throughout my research. His encouragement and invaluable advice have greatly shaped both this work and my personal development as a researcher.

I am also sincerely thankful to the members of my thesis committee for taking the time to review my work and for providing insightful comments and suggestions that contributed to its improvement.

Special thanks go to my colleague Damjan, with whom I had the opportunity to collaborate for a short period of time. I greatly enjoyed our work together and appreciate the experience we shared.

I am especially grateful to my colleague and friend Nikolina, from whom I have learned so much. Her support, patience, and willingness to share knowledge have been invaluable to me, both professionally and personally, and I am truly thankful for the opportunity to work alongside her and for the friendship we built during this journey.

I am deeply grateful to my dear friend Ana, who has supported me every step of the way, even from afar. Despite the distance, her presence and encouragement have been a constant source of strength and motivation.

To my dearest friend Lucija, I owe heartfelt thanks. She has always been there with unwavering support, understanding, and kindness, and her presence has been a constant source of encouragement and reassurance throughout this journey.

Finally, I wish to express my profound gratitude to my family. To my parents, Snježana and Jelisije, who made all of this possible through their endless love and unconditional support; to my brothers, Nenad and David, and their wives, Ana-Marija and Nikolina, who together created a complete support system around me; to my aunt Sandra, whose wholehearted support and kindness accompanied me at every step of this journey; and to every member of my family who stood by me in different ways throughout this path. Their encouragement, patience, and faith in me have been the foundation that carried me through to the end of this work.

Supervisor information

Ivan Radosavljević (born 22 December 1981 in Croatia) is an Associate Professor at the Department of Biology, Faculty of Science, University of Zagreb. He obtained his Master's degree in Biology and Chemistry Education (mag. educ. biol. et chem.) from the University of Zagreb in 2005, and later earned his PhD in Biology at the same institution in 2012. His academic career has been fully dedicated to the University of Zagreb, where he first worked as a PhD student and research assistant (2007–2012), followed by a postdoctoral researcher position (2012–2016), before being appointed Assistant Professor (2016–2022) and subsequently Associate Professor in 2022. He is actively engaged in teaching several courses at the Department of Biology, including Systematic Botany, Botany, Phylogeny and Systematics of Plants, Phylogeny and Molecular Systematics, and Fundamentals of Molecular Ecology, as well as field courses such as the Field Course in Biological, Geographical and Geological Environmental Protection and previously the Field Course in Botany and Vertebrates. Radosavljević has also been involved in a number of scientific projects. He recently led the Croatian Science Foundation (HRZZ) project Amethyst Meadow Squill (*Chouardia litardierei*, Hyacinthaceae): a study system for ecological divergence, and has contributed as a team member to several other initiatives, including the COST Action An integrated approach to conservation of threatened plants for the 21st Century (ConservePlants), the European Commission's Centre of Excellence for Biodiversity and Molecular Plant Breeding, and HRZZ projects on the genetic and epigenetic basis of diversity and bioactivity in Croatian plant species such as *Tanacetum cinerariifolium*, common bean landraces, and endemic *Salvia* species.

University of Zagreb
Faculty of Science
Department of Biology

Doctoral thesis

GENETIC BASIS OF ADAPTIVE DIVERGENCE IN AMETHYST MEADOW SQUILL (*CHOUARDIA LITARDIEREI*, HYACINTHACEAE)

SARA LAURA ŠARANČIĆ

Faculty of Science, Department of Biology

Local adaptation in *Chouardia litardierei* (Asparagaceae), a geophytic perennial endemic to the western Balkans, demonstrates how sharp ecological contrasts drive genomic and phenotypic divergence. The species inhabits three ecologically distinct environments: flooded karst poljes, arid dolomite slopes, and coastal salt marshes, providing a natural system to study how environmental heterogeneity influences adaptive differentiation. As an initial step towards understanding the genomic basis of local adaptation, a chromosome-level reference genome was assembled. Subsequently, ddRAD sequencing was employed to genotype individuals from selected populations, enabling genome-wide association studies (GWAS), genome–environment association analyses (GEA), and population genetic analyses. GWAS revealed high heritability for several phenological and reproductive traits, underscoring their strong genetic basis, while GEA identified precipitation during the coldest quarter as a key climatic driver with the strongest influence on genetic variation. Phenological patterns in the common garden experiment showed substantial overlap in flowering time and vegetative growth across habitat types, indicating no consistent differences among populations. Morphometric analyses additionally indicated reduced clonal investment in dolomite populations. Population genetic analyses revealed partial genetic structuring, with dolomite populations forming a distinct cluster, while seashore and meadow populations remained genetically undistinguished, likely due to shared ancestry or gene flow. Rather than constituting discrete ecotypes, populations seem to diverge through trait-specific, localised responses to environmental pressures. Overall, these findings illustrate how contrasts in ecological conditions influence fine-scale patterns of genomic and phenotypic divergence in *C. litardierei*, providing valuable genomic resources and a framework for future investigations of plant adaptation in heterogeneous South-European landscapes.

(159 pages, 4 figures, 164 references, original in English)

Keywords: *Chouardia litardierei*, local adaptation, phenological traits, reproductive strategies, GWAS, population genomics

Supervisor: Assoc. Prof. Ivan Radosavljević, PhD
Reviewers: Prof. Zlatko Liber, PhD
Prof. Zlatko Šatović, PhD
Assoc. Prof. Ivana Buj, PhD

Sveučilište u Zagrebu Doktorski rad

Prirodoslovno-matematički fakultet

Biološki odsjek

GENSKA OSNOVA ADAPTIVNE DIVERGENCIJE LIVADNOG PROCJEPKA (*CHOUARDIA LITARDIEREI*, HYACINTHACEAE)

SARA LAURA ŠARANČIĆ

Prirodoslovno-matematički fakultet, Biološki odsjek

Lokalna prilagodba livadnog procjepka (*Chouardia litardierei*, Asparagaceae), geofitne višegodišnje biljke endemične za zapadni Balkan, oblikovana je djelovanjem izraženih ekoloških kontrasta koji utječu na njegovu genetsku i fenotipsku divergenciju. Vrsta nastanjuje tri ekološki izrazito kontrastna staništa: poplavljena krška polja, sušne dolomitne padine i slane priobalne močvare, čime predstavlja pogodan model za istraživanje utjecaja okolišne heterogenosti na adaptivnu divergenciju. Kao prvi korak prema razumijevanju genetske osnove lokalne prilagodbe, sastavljen je referentni genom na razini kromosoma, a zatim su sve jedinke iz odabranih populacija genotipizirane pomoću ddRAD sekvenciranja. Ovi genomski podaci omogućili su provođenje cjelogenomske studije povezanosti (*engl. genome-wide association study*, GWAS), analize povezanosti genoma i okoliša (*engl. genome-envirionment association*, GEA) te populacijsko-genetičkih analiza. GWAS analiza pokazala je visoke vrijednosti nasljednosti za više fenoloških i reproduktivnih svojstava, što upućuje na njihovu snažnu genetsku osnovu. GEA analizom izdvojena je količina oborina tijekom najhladnijeg tromjesečja kao klimatska varijabla s najznačajnijim utjecajem na genetsku varijaciju. Usporedno s time, fenološki obrasci uočeni tijekom vrtnog pokusa pokazali su znatno preklapanje u vremenu cvatnje i vegetativnog rasta među različitim tipovima staništa, što ukazuje na izostanak konzistentnih razlika među populacijama. Morfometrijske analize dodatno su pokazale smanjeni udio klonalnog razmnožavanja u populacijama s dolomitnih staništa. Populacijsko-genetičke analize pokazale su da populacije s dolomitnih staništa tvore zaseban genetski klaster, što se može objasniti dugotrajnom izolacijom, ekološkim stresom i ograničenim protokom gena, dok su priobalne i livadne populacije genetski ne razlikuju, vjerojatno zbog zajedničkog porijekla ili kontinuiranog protoka gena. Umjesto jasno definiranih ekotipova, pokazuju specifične odgovore na lokalne selekcijske pritiske. Ovo istraživanje pruža temeljne uvide u procese biljne adaptacije u kompleksnim mediteranskim staništima te osigurava vrijedne genomske resurse za buduća evolucijska istraživanja.

(159 stranica, 4 slike, 164 literaturnih navoda, jezik izvornika: engleski)

Ključne riječi: *Chouardia litardierei*, lokalna prilagodba, fenološka svojstva, reproduktivne strategije, GWAS, populacijska genetika

Mentor: Izv. prof. dr. sc. Ivan Radosavljević
Ocjenjivači: Prof. dr. sc. Zlatko Liber
Prof. dr. sc. Zlatko Šatović
Izv. prof. dr. sc. Ivana Buj

Abbreviations and Acronyms

AHI	Average Height of Inflorescences
BC	Bulb Count
BOF	Beginning of Flowering
BOS	Beginning of Sprouting
BSLMM	Bayesian Sparse Linear Mixed Model
ddRAD-seq	Double Digest Restriction Site-Associated DNA Sequencing
EGGNog	Evolutionary Genealogy of Genes: Non-supervised Orthologous Groups
FPD	Flowering Period Duration
GEA	Genome-Environment Association
GEMMA	Genome-wide Efficient Mixed Model Association
Gbp	Gigabase pairs
GLMM	Generalized linear mixed model
GMMAT	Generalized Mixed Model Association Tests
GWAS	Genome-Wide Association Study
LMM	Linear Mixed Model
mGWAS	Multivariate genome-wide association study
NT	Near Threatened
PCA	Principal Component Analysis
PIP	Posterior Inclusion Probability
PVE	Proportion of Variance Explained
RDA	Linear Model Redundancy Analysis
sNMF	Sparse Non-negative Matrix Factorization
SNP	Single Nucleotide Polymorphism
TFC	Total Flower Count
TGS	Third Generation Sequencing
VPD	Vegetation Period Duration
WGS	Whole-Genome Sequencing

List of publications

Radosavljević, I., Križanović, K., **Šarančić, S. L.**, and Jakše, J. (2023). Towards the Investigation of the Adaptive Divergence in a Species of Exceptional Ecological Plasticity: Chromosome-Scale Genome Assembly of *Chouardia litardierei* (Hyacinthaceae). Int J Mol Sci 24, 10755. doi: 10.3390/ijms241310755

Šarančić, S. L., Pleić, N., Križanović, K., Surina, B., Mitić, D., and Radosavljević, I. (2025a). Uncovering the genomic basis of phenological traits in *Chouardia litardierei* (Asparagaceae) through a genome-wide association study (GWAS). Front Plant Sci 16, 1571608. doi: 10.3389/FPLS.2025.1571608

Šarančić, S. L., Pleić, N., Mitić, D., Križanović, K., Surina, B., and Radosavljević, I. (2025b). Genome-wide association study (GWAS) provides insights into the genomic basis of reproduction-related traits in *Chouardia litardierei* (Asparagaceae). BMC Plant Biology 2025 25:1 25, 1–25. doi: 10.1186/S12870-025-06617-4

Thesis summary

Local adaptation plays a crucial role in shaping genetic and phenotypic diversity within species, particularly in ecologically heterogeneous regions such as the Balkan Peninsula. Environmental elements, such as temperature, precipitation, and soil composition, impose divergent selective pressures that drive the evolution of locally advantageous traits. In sessile organisms like plants, these pressures often lead to differentiation in phenological or morphological traits, setting the stage for adaptive divergence and speciation. The Balkan Peninsula, one of Europe's most topographically and climatically complex regions, offers a unique setting to investigate how fine-scale environmental heterogeneity influences genomic variation and adaptive trait evolution. While numerous phylogeographic studies of local taxa were conducted, integrative research combining population genomics, ecological data, and phenotypic analyses remains scarce.

This thesis focuses on *Chouardia litardierei* (Asparagaceae), a geophytic perennial of the western Balkans. The species is notable for its pronounced ecological plasticity and occupies three ecologically distinct habitats: (1) seasonally flooded karst poljes, characterized by deep, fertile soils; (2) arid dolomite slopes, marked by shallow, rocky substrates, nutrient-poor soils, and extreme thermal fluctuations; and (3) low-lying coastal seashore meadows and salt marshes influenced by salinity, periodic tidal flooding, and low oxygen availability. These sharply contrasting environments impose divergent selective pressures on survival, reproduction, and phenology, offering a natural experimental framework for studying local adaptation. This combination of habitat heterogeneity and pronounced phenotypic plasticity makes *C. litardierei* an excellent model for studying local adaptation to contrasting environments. Its occurrence across ecologically contrasting yet logistically accessible areas makes the species exceptionally well-suited for field sampling. Moreover, as a small, bulbous perennial, it is readily cultivated, allowing for common garden experiments that minimize variations of environmental conditions and reveal heritable trait differences. Together, these features provide an ideal system for investigating how selective pressures shape morphological and genomic variation in natural plant populations.

Despite its ecological distinctiveness, *C. litardierei* has not been genomically characterized prior to this work. Earlier studies were limited to karyology or morphology analyses with some questionable conclusions (e.g., classification of dolomite populations as a distinct species). This thesis addresses the existing knowledge gap by investigating the genomic foundations of local adaptation, with a particular focus on phenology and reproduction-related morphological traits.

This research was guided by objectives spanning genome assembly, population structure of studied populations, and genotype–phenotype–environment associations, with a central goal of understanding the genomic basis of local adaptation in a previously uncharacterized species.

The research is presented across three publications. In **Publication I**, a high-quality, chromosome-level genome of *C. litardierei* was assembled with a total size of 3.7 Gbp, consisting of 13 pseudochromosomes, which is consistent with previous cytogenetic analyses. Comparative analyses confirmed its phylogenetic relationship to some other sequenced representatives of the Asparagales, underscoring the value of this resource for monocot evolutionary studies. Importantly, this reference genome served as a critical foundation for all subsequent genomic and population-level analyses in this thesis, enabling detailed investigations of genetic variation, trait architecture, and genome-environmental associations.

Subsequent genome-wide association study (GWAS) analyses required two complementary datasets: genotyping and phenological and morphological traits results obtained from the common garden experiment. The common garden experiment was established by transplanting a total of 214 individuals from selected populations.

Publication II employed a common garden experiment to investigate phenological divergence among nine selected populations, three originating from each of three distinct habitat types. Although dolomite populations flowered somewhat earlier and had shorter vegetative periods when compared to some other tested populations, the observed pattern was not exclusive to the dolomite populations, as non-dolomite populations overlapped in flowering time, suggesting incomplete phenological divergence. GWAS revealed numerous significant loci for key phenological traits such as the beginning of sprouting (BOS), beginning of flowering (BOF), and vegetative period duration (VPD). Narrow-sense heritability (h^2) was high for all traits, particularly VPD (86.95%), suggesting that much of the observed phenological variation among populations is genetically determined.

Building on this, **Publication III** integrated GEA analyses with GWAS to investigate links between genomic variation, environmental gradients, and morphological reproductive traits. Through RDA analysis, GEA identified winter precipitation (BIO19) as the climatic element with the strongest influence on genetic variation, pointing to moisture availability as a major driver of local adaptation. GWAS uncovered loci associated with tested reproductive traits, including bulb count (BC), total flower count (TFC), and average inflorescence height (AHI), all of which showed

high narrow-sense heritability (e.g., 71.95% for AHI). Genomic regions affecting nitrogen metabolism, hormone signaling, and flowering regulation were identified. Morphometric analysis also revealed variation in reproductive traits, such as bulb number, across populations, with dolomite populations typically producing fewer bulbs. This pattern suggested adaptive trade-offs favoring sexual over clonal reproduction in these drought-prone environments where flooding is absent. In contrast, populations from karst poljes and seashore meadows experience regular and prolonged flooding, which can hinder successful sexual reproduction and likely favors clonal reproduction, contributing to observed differences in reproductive strategies across habitats.

Population-genetic analyses revealed that dolomite populations of *C. litardierei* form a distinct genetic cluster, likely shaped by long-term isolation and absence of gene flow with other populations due to the patchiness of the specific habitat it inhabits. This pattern of partial genetic divergence, first hinted at in **Publication I** through whole-genome comparisons of single representatives, was confirmed in **Publication III** using broader population-level data. In contrast, groups of populations from seashore meadows and inland wet grasslands exhibited little to no genetic structuring and were indistinguishable from each other, likely due to either recent shared ancestry or contemporary gene flow among populations. Despite the long-standing informal designation of ecotypes based on habitat, the lack of consistent genetic and morphological differentiation among most of the studied groups suggests that divergence is incomplete. The only signals of genetic differentiation were observed in the dolomite group, suggesting that local adaptation in *C. litardierei* is not strictly tied to broad habitat types. Consequently, we adopted a conservative interpretation and refer to these population groups as specific habitat-associated groups, rather than clearly defined ecotypes.

Together, these studies provide a comprehensive view of the genetics underlying local adaptation in *C. litardierei*. This thesis presents the first GWAS and GEA conducted in this species and represents one of the earliest integrative genomic studies of a native South-European monocot. The identification of polygenic variation underlying both phenological and reproductive divergence highlights the complex interplay of natural selection, gene flow, and trait evolution. More broadly, this work contributes to our understanding of how plants adapt across heterogeneous landscapes and offers a valuable genomic and ecological framework for future research on adaptation and diversification in various South-European plant lineages.

Prošireni sažetak

Lokalna prilagodba ima ključnu ulogu u oblikovanju genetske i fenotipske raznolikosti unutar vrsta, osobito u ekološki heterogenim regijama poput Sredozemlja. Ekološki čimbenici poput temperature, količine oborina i sastava tla stvaraju selekcijske pritiske koji oblikuju svojstva prilagođena lokalnim uvjetima. U sesilnih organizama, poput biljaka, takvi pritisci često rezultiraju diferencijacijom fenoloških i morfoloških osobina, čime nerijetko započinje divergencija, a zatim i specijacija. Balkanski poluotok, kao jedna od topografski i klimatski najsloženijih regija Europe, pruža iznimnu priliku za istraživanje utjecaja heterogenog okoliša na genetsku varijabilnost i evoluciju adaptivnih svojstava. Bez obzira na porast broja filogeografskih istraživanja lokalnih vrsta, istraživanja koja povezuju populacijsku genetiku, ekološke podatke i fenotipske analize i dalje su rijetka.

Predmet istraživanja ove disertacije je livadni procjepak (*Chouardia litardierei*, Asparagaceae), višegodišnja geofitna vrsta s prirodnim arealom na zapadnom Balkanu. Vrsta se odlikuje izraženom ekološkom plastičnošću, a naseljava tri ekološki kontrastna staništa: (1) sezonski poplavljena krška polja s dubokim i plodnim tlima; (2) sušne dolomitne padine s plitkim, kamenitim tlima, niskim udjelom hranjivih tvari i značajnim toplinskim oscilacijama; te (3) povremeno plavljenim priobalnim livadama s niskom koncentracijom kisika i slanim močvarama. Raznolikost staništa koja *C. litardierei* nastanjuje stvara izražene selekcijske pritiske koji utječu na uspješnost opstanka, razmnožavanja i fenoloških obrazaca biljaka, pružajući pritom prirodan model za istraživanje lokalne prilagodbe. Zbog rasprostranjenosti u ekološki kontrastnim, ali terenski pristupačnim područjima, vrsta je prikladna za uzorkovanje. Dodatno, kao mala lukovičasta višegodišnja biljka, lako se uzgaja u kontroliranim uvjetima, što omogućuje provedbu vrtnih pokusa i razlučivanje genetskih od okolišnih utjecaja.

Unatoč svojoj ekološkoj specifičnosti, *C. litardierei* prije ovog istraživanja nije bila genomski okarakterizirana. Dosadašnja saznanja uglavnom su se temeljila na kariološkim analizama i morfološkim pretpostavkama, uključujući raniji prijedlog o taksonomskom izdvajanju dolomitnih populacija kao zasebne vrste, no takvi su zaključci ostali nepotkrijepljeni. Ova disertacija ima za cilj razjasniti nedostatno istražene aspekte genomske osnove lokalne prilagodbe, s posebnim naglaskom na fenologiju i morfološke osobine povezane s razmnožavanjem. Istraživanje je usmjereno na ostvarivanje niza znanstvenih ciljeva, uključujući sastavljanje referentnog genoma, filogenetsku analizu, analizu populacijske strukture te ispitivanje povezanosti

genetske varijabilnosti, fenotipskih obilježja i okolišnih čimbenika, s krajnjim ciljem dubljeg razumijevanja genomske osnove lokalne prilagodbe u dosad nedovoljno istraženoj biljnoj vrsti.

Rezultati istraživanja predstavljeni su kroz tri znanstvene publikacije. U okviru **prve publikacije** izrađen je visokokvalitetan referentni genom vrste, ukupne veličine 3,7 Gbp, koji se sastoji od 13 pseudokromosoma, što je u skladu s prethodnim citogenetskim analizama. Komparativne analize potvrdile su filogenetsku vezu vrste s drugim do sada sekvenciranim predstavnicima reda Asparagales, čime je dodatno naglašena znanstvena vrijednost ovog genoma za buduća evolucijska istraživanja. Referentni genom poslužio je kao temeljna infrastrukturna osnova za sve naknadne analize populacijske strukture, identifikaciju genetskih varijanti povezanih s kvantitativnim fenotipskim osobinama te ispitivanje povezanosti genetskih obilježja s okolišnim čimbenicima.

U **drugoj publikaciji** proveden je vrtni pokus radi procjene fenološke divergencije među devet odabranih populacija, pri čemu su po tri populacije pripadale svakom od triju različitih tipova staništa. Populacije s dolomitnih staništa imale su raniji početak cvatnje i kraće trajanje vegetacijskog razdoblja, što je u skladu s prilagodbom na sušne uvjete, međutim, ne-dolomitne populacije pokazivale su preklapanje u analiziranim svojstvima, što upućuje na nepotpunu fenološku divergenciju. Cjelogenomskom studijom povezanosti (*engl. genome-wide association study*, GWAS) identificirani su značajni lokusi povezanim s ključnim fenološkim svojstvima, uključujući početak nicanja (*engl. beginning of sprouting*, *BOS*), početak cvatnje (*engl. beginning of flowering*, *BOF*) i trajanje vegetacijskog perioda (*engl. vegetative period duration*, *VPD*). Procjene nasljednosti (*engl. narrow-sense heritability,* h^2) pokazale su visoke vrijednosti za sve analizirane osobine, pri čemu je VPD imao najvišu vrijednost h^2 (86,95 %), što ukazuje na snažnu genetsku osnovu opaženih fenoloških varijacija. Funkcionalna anotacija genomskih prozora koji okružuju značajne polimorfizme jednog nukleotida (*engl. single nucleotide polymorphism, SNP*) otkrila je regije koje kodiraju ključne proteinske obitelji uključene u regulaciju vremena cvatnje, vegetativnog rasta i odgovora na stres. Dobiveni nalazi sugeriraju da se fenološke razlike kod *C. litardierei* najbolje objašnjavaju kao populacijski specifične prilagodbe na lokalne uvjete.

U okviru **treće publikacije** predstavljeni su rezultati praćenja morfoloških reproduktivnih svojstava vrste te su provedene integrirane analize povezanosti genoma i okoliša (*engl. genome-environment association*, GEA) te cjelogenomska analiza povezanosti, s ciljem istraživanja odnosa između genetske varijabilnosti, okolišnih gradijenata i reproduktivnih svojstava. Pomoću RDA

analize, GEA analizom izdvojena je količina oborina tijekom najhladnijeg tromjesečja (BIO19) kao najznačajnija klimatska varijabla povezana s genetskom strukturom populacija, što ukazuje na potencijalno ključnu ulogu vlage u procesu lokalne prilagodbe. Paralelno, cjelogenomskom analizom povezanosti otkrivena je značajna povezanost između pojedinih genetskih lokusa i reproduktivnih svojstava poput broja lukovica (*engl. bulb count*, *BC*), ukupnog broja cvjetova (*engl. total flower count*, *TFC*) i prosječne visine cvatova (*engl. average height of inflorescence*, *AHI*), s visokom nasljednošću svih osobina (npr. 71,95 % za AHI). Funkcionalnom anotacijom otkrivene su genomske regije uključene u biološke procese poput metabolizma dušika, hormonske signalizacije i regulacije cvatnje, što dodatno podupire njihovu ulogu u prilagodbi na heterogene ekološke uvjete. Morfometrijska analiza također je pokazala varijabilnost reproduktivnih svojstava među populacijama, osobito u broju lukovica, pri čemu su dolomitne populacije u pravilu proizvodile manji broj lukovica. Ovakav obrazac upućuje na prisutnost adaptivnih kompromisa koji pogoduju spolnom razmnožavanju nauštrb klonalnog u sušnim staništima. Suprotno tome, populacije koje nastanjuju krška polja i priobalne livade izložene su redovitim i dugotrajnim poplavama, što može otežati uspješno spolno razmnožavanje te vjerojatno potiče oslanjanje na klonalne strategije, pridonoseći tako uočenim razlikama u reproduktivnim strategijama između različitih staništa.

Populacijsko-genetičke analize pokazale su da populacije s dolomitnih staništa tvore zaseban genetski klaster, što se može objasniti njihovom izoliranošću i odsutnošću protoka gena s ostalim populacijama, uvjetovanom fragmentiranošću specifičnih staništa koja vrsta nastanjuje. Ovaj obrazac djelomične genetske divergencije, koji je prvi put uočen u prvoj publikaciji usporedbom pojedinačnih genoma, potvrđen je širim analizama u trećoj publikaciji. Nasuprot tome, populacije s poplavljenih krških polja i priobalnih livada nisu pokazivale jasnu genetsku diferencijaciju, što vjerojatno odražava njihovo nedavno zajedničko podrijetlo ili kontinuirani protok gena. Unatoč dugogodišnjoj neformalnoj uporabi pojma ekotip za opisivanje različitih stanišnih skupina iste vrste, nedostatak konzistentne genetske i morfološke diferencijacije između većine proučavanih grupa, sugerira da divergencija nije potpuna. Jedini obrasci genetske izolacije i morfološke diferencijacije zabilježeni su kod populacija s dolomitnih staništa, što podupire interpretaciju da lokalna prilagodba nije povezana s određenim tipovima staništa kao što je to bilo očekivano. Stoga primjenjujemo konzervativniji pristup u interpretaciji te ove skupine opisujemo kao populacije povezane sa specifičnim staništima, a ne kao jasno definirane ekotipove. Za razliku

od širokih filogeografskih istraživanja koja obuhvaćaju velika geografska područja, ova je disertacija usmjerena na proučavanje divergencije unutar jedne vrste, istražujući različita staništa unutar relativno uskog geografskog prostora te pružajući uvid u mehanizme lokalne prilagodbe na prostornim i genetskim razinama.

Ova disertacija predstavlja prvi korak u razumijevanju genetske osnove lokalne prilagodbe kod *C. litardierei*. Predstavljena su prva GWAS i GEA istraživanja provedena na ovoj vrsti, kao i jedno od prvih genetsko-ekoloških istraživanja na nativnoj jednosupnici južne Europe. Identifikacija poligenske kontrole i visoke nasljednosti fenoloških i morfoloških reproduktivnih osobina ističe složenost interakcija između prirodne selekcije, protoka gena i evolucije adaptivnih svojstava. U širem kontekstu, ova disertacija doprinosi boljem razumijevanju načina na koje se biljne vrste prilagođavaju heterogenim okolišima te pruža vrijednu genetsku i ekološku osnovu za buduća istraživanja adaptacije i diferencijacije u različitim razvojnim linijama južnoeuropskih biljnih vrsta.

Table of Contents

1 Introduction

1.1 Genomic basis of adaptation

One of the central aims of evolutionary biology is to elucidate the genomic basis of adaptation, particularly how genetic variation enables organisms to cope with environmental challenges (Flood and Hancock, 2017; Campbell et al., 2018). Environmental heterogeneity imposes selective pressures that shape the genetic architecture of traits, enhancing survival and reproduction (Gregory, 2009; Hoban et al., 2016; Paschoa et al., 2023). These pressures drive changes in allele frequencies, promoting local adaptation by favoring variants that enhance fitness under specific environmental conditions (Hu et al., 2020; Walter et al., 2022; Lee et al., 2023). Over time, successive rounds of selection stabilize these adaptive responses, allowing populations to persist under heterogeneous environments (Savolainen et al., 2013; McKown et al., 2014). As populations inhabit different environments, such as coastal, mountainous, or inland regions, they often evolve distinct phenological or morphological traits that enhance their survival and reproduction. This divergence is frequently shaped by environmental gradients, including temperature, precipitation, and soil composition (Turesson, 1922; Todesco et al., 2020).

Over evolutionary timescales, local adaptation can set the stage for speciation, during which a group of individuals diverges into two or more distinct phylogenetic lineages (Clausen, 1951). In populations that are initially indistinguishable, whether genetically or morphologically, progressive genetic differentiation can gradually lead to new species (Rundle and Nosil, 2005; Tittes and Kane, 2014; Cortés et al., 2018). When gene flow across habitat boundaries is limited, local adaptation to contrasting environments can drive genetic divergence and form reproductive barriers—key steps in the process of ecological speciation (Rundle and Nosil, 2005; Lowry, 2012). At intermediate stages of divergence, ecological differentiation becomes more apparent, often culminating in the formation of ecotypes, genetically and morphologically distinct groups adapted to specific ecological niches rather than to specific geographic areas (Rundle and Nosil, 2005; Cortés et al., 2018). Although the role of ecotypes in speciation continues to be questioned (Lowry, 2012; Fernández-Meirama et al., 2022), many studies highlight their contribution in promoting genetic divergence along ecological gradients (Lowry et al., 2008a; Brandrud et al., 2017; Cortés et al., 2018; Bakhtiari et al., 2019). Because gene flow typically homogenizes genetic differences among populations, adaptive divergence is generally expected to arise between geographically isolated populations. As a result, most studies of local adaptation have focused on populations separated by

tens to hundreds of kilometers, leaving microgeographic divergence relatively understudied (Bamba et al., 2019). Detecting adaptive divergence at such fine spatial scales is further complicated by phenotypic plasticity, the ability of a single genotype to express different phenotypes under varying environmental conditions, which can mask underlying genetic differences (De Villemereuil et al., 2015). To mitigate this issue, common garden experiments are often employed to minimize the influence of environmental variation (McKay et al., 2001; Kawakami et al., 2011; Brachi et al., 2013; Toräng et al., 2015). By growing individuals from contrasting environments under uniform conditions, common garden experiments remove environmental confounding, thereby isolating genetic effects on complex, polygenic traits and providing a robust test for signals of local adaptation (De Villemereuil et al., 2015). Among the many genetically based traits shaped by local adaptation, reproductive strategies represent a key axis of divergence, particularly in heterogeneous environments where the costs and benefits of different types of sexual and asexual reproduction may vary.

1.2 Reproductive strategies as mechanisms of divergence

The development of different reproductive strategies, encompassing both sexual and asexual modes, plays a pivotal role in plant adaptation and divergence. In particular, clonal reproduction, or vegetative propagation, may arise under biotic or abiotic stresses that constrain sexual reproduction and provide numerous ecological advantages (Silvertown, 2008; Barrett, 2015). By enabling plants to forage for water and nutrients in patchy environments and to produce larger propagules capable of rapid establishment in new sites, clonality enhances survival and supports the successful establishment of populations after colonization (Klimeš et al., 1997; Stuefer et al., 2002; Orive, 2020). It is estimated that up to 80% of angiosperms utilize asexual reproduction, and its prevalence may vary among populations, especially in heterogeneous landscapes shaped by variations in climate, soil composition, hydrology, and biotic interactions (Price and Marshall, 1999). In such contexts, flexible reproductive systems allow plants to navigate diverse selective pressures and promote genetic differentiation across habitats.

However, the ecological flexibility afforded by clonality comes with important trade-offs. Investment in clonal growth may restrict the resources available for flowering and seed production (Van Drunen and Dorken, 2012). Additionally, high levels of clonality can elevate geitonogamous self-pollination, potentially accelerating reproductive isolation and speciation (Gong et al., 2010;

Vallejo-Marín et al., 2010). Over time, if sexual reproduction becomes extremely limited or absent, the accumulation of somatic mutations may further erode fertility, possibly resulting in the complete loss of sexual function (Barrett, 2015).

Beyond direct impacts on fertility, clonality profoundly alters spatial genetic structure by concentrating identical genotypes locally, which limits gene flow and promotes divergence (Vekemans and Hardy, 2004). These effects are further shaped by associated life-history traits such as bulb formation, which serve as critical adaptations. Bulbs act as critical storage organs, enabling plants to endure periods of dormancy and buffer against adverse conditions; this function is essential for maintaining reproductive capacity in habitats characterized by fluctuating conditions (Kleijn et al., 2005; Atif et al., 2020; Ma et al., 2020). The number of bulbs produced and the morphology of inflorescences can influence pollination success and seed output, directly affecting fitness and local adaptation (Ohashi and Yahara, 2009; Suetsugu et al., 2015).

Taken together, clonal and sexual reproduction represent complementary strategies that enable plants to cope with habitat variability, shaping both the genetic structure and adaptive potential of populations.

1.3 The evolutionary significance of phenological adaptation

Phenology, the seasonal timing of key life cycle events such as sprouting and flowering, is both a sensitive indicator of climate change and a fundamental mechanism through which plants adapt to environmental variability (CaraDonna et al., 2014; Schwartz, 2024). As sessile organisms, plants depend on phenological timing to synchronize growth and reproduction with favorable seasonal windows, thereby optimizing fitness (Mertens et al., 2021).

Among phenological traits, flowering time is particularly critical, as it integrates plant responses to both abiotic (e.g., temperature and moisture) and biotic (e.g., pollinator availability) cues, serving as a key expression of adaptive strategy (Pau et al., 2011; Wolkovich et al., 2014). It directly influences reproductive success by mediating trade-offs between fecundity and survival, shaping interactions with pollinators, regulating seed output, and aiding in stress avoidance (Anderson et al., 2012; Collins et al., 2025).

The diversity of factors influencing flowering time reflects the complex environmental landscape that plants must navigate. While temperature is often the primary driver, affecting development rates, metabolism, and water loss (Körner, 2006; Linder, 2020), other cues such as photoperiod (Adole et al., 2019; Wang et al., 2020), water availability (Zhou et al., 2024), salinity

(Li et al., 2007; Lee et al., 2023), and pollinator-mediated selection (Sandring and Ågren, 2009) also shape phenological timing (Cook et al., 2012; Schwartz, 2024). Importantly, moisture availability can be shaped not only by rainfall but by local soil characteristics and rooting depth (Cortés-Flores et al., 2017); in wetlands, hydrological regimes (e.g., groundwater levels or seasonal flooding) often outweigh precipitation as drivers of phenological timing (Mihevc et al., 2010).

From an evolutionary perspective, phenological traits such as flowering time can evolve rapidly under selection, particularly in seasonally variable or environmentally heterogeneous habitats (Gaudinier and Blackman, 2020). In such contexts, divergent flowering schedules may facilitate ecological speciation by enabling populations to exploit distinct pollinator communities, climatic niches, or microhabitat (Heslop-Harrison, 1964; Levin, 2006). However, not all species can adjust their phenology at a rate sufficient to match environmental change. When plasticity fails to buffer against altered pollinator availability, resource availability, or interspecific competition, reproductive success may decline, imposing selective pressures that favor genetic shifts in timing traits (Visser et al., 2003; Pau et al., 2011). While many phenological shifts begin as plastic responses, the presence of heritable variation is crucial for long-term evolutionary adaptation (Visser and Both, 2005).

Given the strong link between phenological timing and local adaptation (Rathcke and Lacey, 1985), investigating the genomic basis of these traits within ecologically dynamic contexts provides a powerful lens for understanding adaptive divergence (Bernatchez et al., 2023). Expanding this research beyond classical model systems such as *Arabidopsis thaliana* (Engelmann and Purugganan, 2006; Kinmonth-Schultz et al., 2021) to encompass a broader diversity of taxa is essential for capturing the complex genetic architectures that underlie phenological variation across contrasting environments (Molla, 2022; Song et al., 2023; Vicentini et al., 2023).

1.4 Study system: *Chouardia litardierei* (Asparagaceae)

Chouardia litardierei (Breist.) Speta, commonly known as the amethyst meadow squill, is a perennial bulbous plant species currently placed in the family Asparagaceae, following the APG III system (Bremer et al., 2009), but was formerly classified within Hyacinthaceae. *C. litardierei* is distributed across the Dinaric Alps in the western Balkan Peninsula, ranging from Slovenia in the northwest to Montenegro in the southeast (Gaži-Baskova, 1962; Petkovšek and Seliškar, 1978).

It is a rare example of a plant thriving across markedly different habitat types, with three major population groups identifiable based on habitat type, reflecting its notable ecological plasticity (Figure 1).

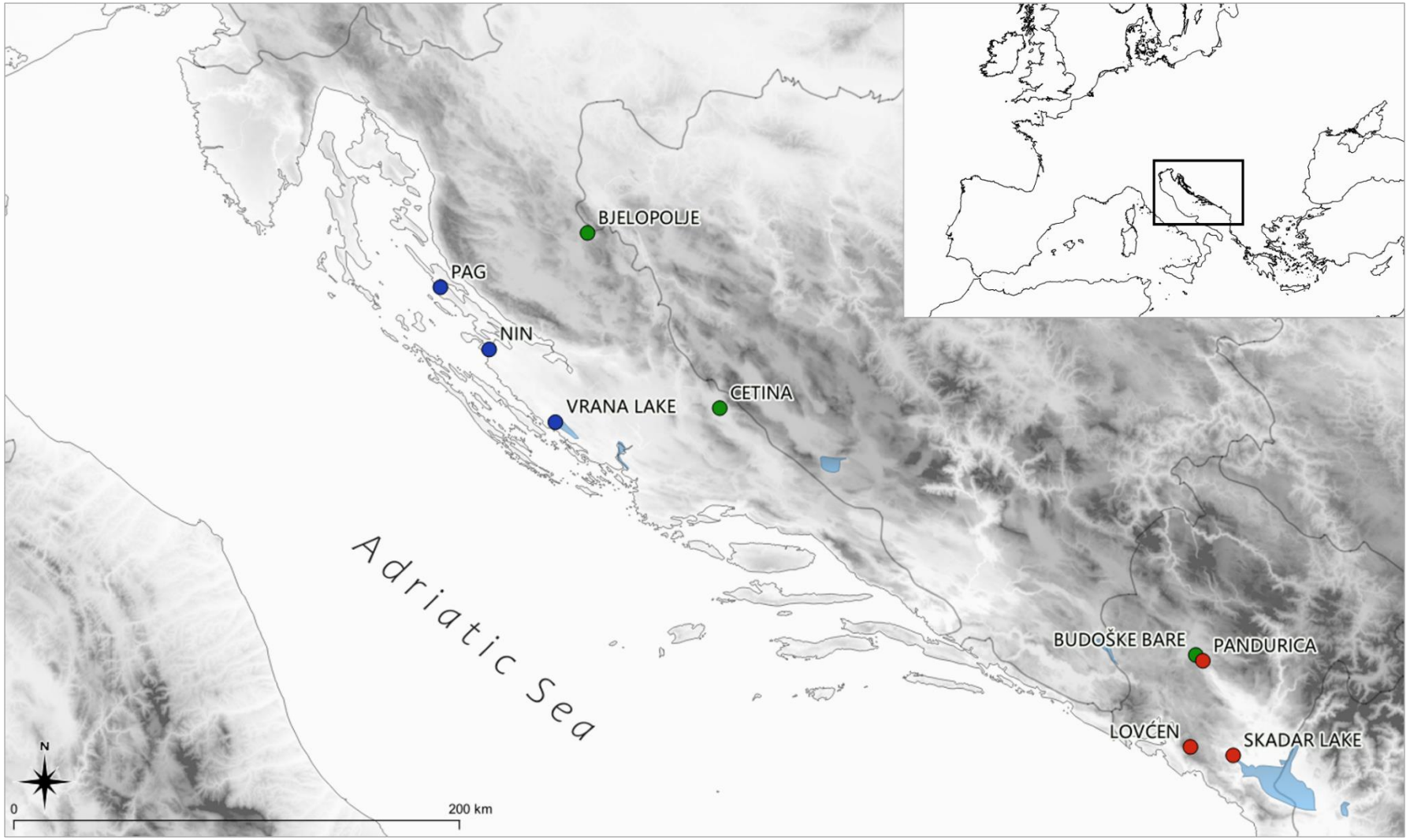


Figure 1. *Sampling locations of Chouardia litardierei populations across habitat types. Blue, green, and red circles indicate populations from seashore, meadow, and dolomite habitats, respectively.*

The largest group occupies karst *poljes*[1] (Figure 2), geomorphologically and hydrologically distinctive depressions within the Dinaric landscape (Prohic et al., 1998). These depressions are characterized by deep, fertile soils, temperature inversion, and seasonal flooding in early spring, driven by rapid snowmelt in surrounding mountains, together with their geomorphology and limited internal drainage (Mihevc, 2010; Bonacci, 2014; Marcin et al., 2021). In contrast with the surrounding karst terrain, typically rocky, shallow-soiled, and drought-prone, poljes serve as localized ecological refugia. The formation and ecological role of poljes is closely tied to the broader Dinaric karst system, the carbonate region of the Dinaric Mountains (Dinarides), marked by exceptional geological and climatic diversity (Mihevc, 2010). Spanning the western Balkan Peninsula and the entire Adriatic littoral zone, the Dinaric karst is recognized as one of the most ecologically diverse landscapes (Day and Chenoweth, 2013; Zupan Hajna, 2019). The dissolution of limestone and dolomite has produced sinkholes, caves, plateaus, and complex underground drainage, creating rugged topography and sharp ecological contrasts over small areas (Prohic et al., 1998). This topographic and hydrological complexity underpins the pronounced ecological heterogeneity characteristic of the region.

Figure 2. *Chouardia litardierei in its natural karst polje habitat (Budoške Bare, Montenegro). Photo taken by Ivan Radosavljević.*

[1] *Polje* is the Slavic word for "field", still commonly used in South Slavic languages without necessarily implying karst terrain (Ford and Williams, 2013). In geomorphology, however, the term refers internationally to large, flat-floored, closed depressions with karstic drainage, typically bordered by steep slopes (Gams, 1978; Jennings, 1985). These basins often provide the only arable land in karst regions.

In contrast, the second group of populations occupies exposed dolomite slopes and dry mountainous grasslands at altitudes of up to 2000 m (Figure 3), characterized by a thin layer of stony, nutrient-poor soils, prone to strong seasonal and daily temperature fluctuations and limited water availability, all of which impose intense ecological stress and demand physiological and phenological resilience (Mota et al., 2021).

Figure 3. *Chouardia litardierei in its natural dolomite habitat (Mt. Lovćen, Montenegro). Photo taken by Ivan Radosavljević.*

These habitats exemplify the so-called "dolomite phenomenon" or dolomitophily, a set of ecological and floristic traits associated with dolomitic substrates (Merlo et al., 2021; Mota et al., 2021). Dolomite-derived soils have poor water retention, high pH, low concentrations of Fe, P, and K, and are rich in Ca and Mg elements (Fekete et al., 1989; Pignatti and Pignatti, 2014; Markoski et al., 2016). The substrate's high thermal conductivity, intense solar radiation, and elevated summer temperatures create drought-prone microhabitats that limit the establishment of competitive vegetation such as forests (Thomas et al., 1973; Waples and Waples, 2004; Parker et al., 2024). The resulting open glades are further shaped by a short growing season, typically restricted to 3–4 months due to prolonged snow cover in higher elevations, which coincides with extreme daily microclimatic fluctuations: intense heating and low humidity during the day, followed by sharp nighttime cooling, with temperatures occasionally dropping below freezing

(Pignatti and Pignatti, 2014). Plants inhabiting these areas must cope with intense *edaphic stress*[2], and the vegetation is typically composed of drought-tolerant, stress-adapted perennials, some of which may exhibit relictual or endemic distributions (Merlo et al., 2021). As such, dolomite slopes challenge survival and act as ecological filters, promoting distinctive plant communities that contribute disproportionately to regional biodiversity.

A third, much smaller group of *C. litardierei* populations occurs in the salt marshes of northern Dalmatia, along the eastern Adriatic coast. In this region, low-lying coastal wetlands are rare due to steep, rocky shorelines, making these geographically limited marshes especially valuable for their specialized vegetation adapted to periodic tidal flooding and elevated soil salinity (Bui, 2013). These intertidal areas, shaped by dynamic sedimentation and erosion, are characterized by fine-scale topography, waterlogged soils with low oxygen availability, and species-poor ecosystems dominated by halophytes adapted to salinity fluctuations (Beeftink, 1977; Silvestri et al., 2005; Pedersen et al., 2017). Unlike the broader, more continuous salt marshes of the western and northern Adriatic, eastern Adriatic marshes are both spatially restricted and increasingly vulnerable to anthropogenic pressures such as agriculture and coastal tourism (Pandža et al., 2007). According to Croatia's national protection strategy, marshy coasts rank among the most critically endangered coastal habitats, making their preservation a top conservation priority (Martinić, 2000).

Reflecting the broader conservation concerns, *C. litardierei* has been classified as Near Threatened (NT) in Croatia (Nikolić and Topić, 2005). This status is attributed to habitat degradation driven by alterations in hydrological regimes, vegetation succession, and land-use changes. Socio-economic shifts in recent decades have led to the abandonment of hay meadows and pastures, accelerating succession and reducing habitat suitability. Additional pressures include drainage and reclamation of wetlands and infrastructure development (e.g., roads, highways), which collectively represent ongoing and future threats to the species' survival (Alegro, 2013). Despite these pronounced environmental differences, previous studies have documented no clear phenotypic differentiation among populations (Šilić, 1990).

However, as a geophyte, *C. litardierei* follows a marked seasonal cycle, remaining dormant from mid-summer through early spring. With the onset of spring, it produces leaves, soon followed

[2] Edaphic stress refers to adverse soil conditions that hinder plant growth, including limited availability of essential resources such as water, nutrients, and oxygen; the presence of toxic elements like excess salts, aluminum, heavy metals, or boron; and physical constraints on root function caused by factors such as mechanical impedance and extreme soil temperatures (Lynch, 2022).

by a racemose inflorescence bearing numerous radially symmetrical flowers. Flowering typically occurs between late April and early June, depending on population and habitat conditions, and is followed by fruit development and complete senescence by mid-July or August. During this phase, the plant reallocates resources to an underground bulb for overwintering. Although no specialized morphological adaptations for pollination have been documented, field observations suggest the species is open-pollinated. In addition to sexual reproduction, *C. litardierei* also reproduces clonally by forming multiple bulbs surrounding the central bulb, enabling persistence and expansion in suitable habitats. This dual reproductive strategy, combined with its ecological plasticity across sharply contrasting environments (Figure 1), makes it a robust model for investigating the genetic and phenotypic bases of local adaptation and adaptive divergence. Its small, bulbous perennial form allows easy transplantation and cultivation in common garden experiments, minimizing confounding effects of phenotypic plasticity. Finally, the species' broad yet accessible distribution across the Dinaric Alps in the western Balkan Peninsula (Ritter-Studnička, 1954; Gaži-Baskova, 1962; Šilić, 1990) facilitates comprehensive field sampling. Together, these attributes make *C. litardierei* a powerful system for exploring the evolutionary processes driving ecological specialization and divergence.

2 Theoretical Framework

2.1 Advances in plant genomics

In recent years, advances in plant genomics have fundamentally reshaped the approaches available for studying ecological and evolutionary processes. The release of the *A. thaliana* genome in 2000 marked a major milestone in plant biology (Kaul et al., 2000). Since then, rapid improvements in sequencing technologies have overcome the historical challenges posed by large genome sizes, their repetitive content, and variable ploidy levels (Marks et al., 2021; Schley et al., 2021). The emergence of third-generation sequencing (TGS) platforms such as PacBio HiFi and Oxford Nanopore has enabled fast, accurate, long-read sequencing without PCR amplification, reducing technical biases and assembly fragmentation, and making chromosome-scale assemblies feasible even for complex plant genomes (Zmienko et al., 2023; Scarano et al., 2024). Over the past decade, these advances have greatly expanded the number of *de novo* plant genome assemblies in public databases, opening the door to high-resolution genomic studies of wild and non-model taxa (Zmienko et al., 2023).

Although whole-genome sequencing (WGS) offers unparalleled resolution and the highest marker density, its cost and computational demands remain limiting factors for studies involving large numbers of individuals, especially in species with large and complex genomes (Ray and Satya, 2014; Huang et al., 2025). In such cases, reduced-representation methods such as double-digest restriction site–associated DNA sequencing (ddRAD-seq) provide a cost-effective alternative for genotyping of non-model species at a population scale (Harvey et al., 2016). By using a pair of restriction enzymes to cut DNA at consistent recognition sites, ddRAD-seq targets a reproducible subset of the genome, which is then barcoded, amplified, and sequenced (Peterson et al., 2012). However, the accuracy and consistency of this representation are critical, as biases in restriction site presence or amplification efficiency can affect genomic coverage (Magbanua et al., 2023). By sequencing a targeted subset of the genome, ddRAD-seq yields thousands of single-nucleotide polymorphisms (SNPs) that serve as genetic markers for assessing diversity, population structure, and enabling downstream analyses such as evolutionary inference and genotype–phenotype associations (Esposito et al., 2020).

Taken together, integrating ddRAD-seq with reference genome assemblies greatly improves resolution in detecting genetic variation and linking it to adaptive, ecological, and evolutionary processes.

Prior to recent genomic efforts, research on *Chouardia litardierei* focused mainly on cytogenetic characterization and preliminary taxonomic assessments. Only two individuals from contrasting habitats were analyzed karyologically (Siljak-Yakovlev, 2010), and there was an early proposal to classify the dolomite group as a distinct taxon (Šilić, 1990). However, these efforts offered a very limited view of the species' genetic and ecological differentiation, thus leaving the underlying mechanisms of local adaptation unexplored.

2.2 Genome-wide association study (GWAS)

Since Darwin, one of the central challenges in evolutionary biology has been to uncover the genetic basis of adaptation and speciation—a pursuit that has become increasingly tractable with the advent of genome-wide association study (GWAS) (Bamba et al., 2019). Approaches for detecting adaptive evolution generally fall into two categories: phenotype-first (top-down) and genotype-first (bottom-up) approaches (Ross-Ibarra et al., 2007). In the top-down approach, an adaptive phenotype is first identified, and then genetic association methods, such as GWAS, are used to pinpoint the underlying genetic factors, particularly when variation is linked to environmental differences. In contrast, bottom-up approaches examine genome-wide patterns of genetic variation to detect regions potentially under selection, without prior knowledge of specific adaptive traits. While this method is not restricted to predefined phenotypes, further analyses are required to interpret their functional significance (Flood and Hancock, 2017).

As a key top-down method, GWAS leverages naturally occurring genetic variation, typically in the form of SNPs, to detect loci statistically associated with observable traits (Groen and Whiteman, 2016). First developed in human genetics (Hirschhorn and Daly, 2005), GWAS is now widely applied, from model species like *Arabidopsis* to diverse non-model taxa (Korte and Farlow, 2013). The first step in such analyses involves defining clear, consistently measurable phenotypes and selecting a study population with sufficient genetic and phenotypic variation to enable robust associations (Alseekh et al., 2021). After genome-wide genotyping of selected individuals, the resulting sequences are aligned to a reference genome, followed by SNP calling to identify genetic variants. Downstream association testing then examines the relationship between SNPs and traits across many individuals, facilitating the discovery of genomic regions linked to adaptive variation (Korte and Farlow, 2013; Uffelmann et al., 2021). This framework is particularly valuable in ecological genomics, as linking phenotypic variation to underlying genotypes can

clarify the genetic architecture of traits influenced by natural selection, while allowing the reuse of the same genotyping data across multiple traits in the same populations (Berhe et al., 2021).

To translate genotypic and phenotypic variation into meaningful biological insights, GWAS relies on robust statistical models capable of capturing the complex relationships between genetic markers and traits (Visscher et al., 2017). Both frequentist and Bayesian approaches have been widely adopted, often used in tandem, and while each offers distinct methodological advantages, the question of their relative superiority remains a topic of ongoing debate (Bayarri and Berger, 2004; Huisman, 2023; Phylactou, 2023). Typically, frequentist methods apply a single-locus framework, while Bayesian approaches often employ multi-locus models (Berhe et al., 2021). However, these approaches are often viewed as complementary, with their combined application enhancing the robustness and interpretability of association findings.

Within the frequentist single-locus framework, association testing typically involves evaluating each genetic marker independently using models tailored to the distributional properties of the trait. For traits approximating a normal distribution, linear mixed models (LMMs), such as those implemented in the GEMMA software package (Zhou and Stephens, 2012), are commonly employed (Onifade et al., 2022; John et al., 2024). These models account for confounding due to population structure and relatedness through a kinship matrix and are well-suited to continuous traits (Kang et al., 2010). However, for traits that deviate from normality, particularly count-based phenotypes, generalized linear mixed models (GLMMs) offer a more appropriate alternative (Onifade et al., 2022). In such cases, tools like the GMMAT software package (Chen et al., 2019) allow for the specification of Poisson error structures, thereby accommodating the discrete nature of the data. Although the computational complexity of LMMs traditionally scales cubically with the number of individuals, the development of optimized software such as GEMMA has significantly improved computational efficiency, enabling their application in large-scale GWAS (Onifade et al., 2022). The complementary use of LMMs and Poisson GLMMs enables a more accurate modeling of diverse trait distributions within the frequentist paradigm, enhancing the reliability of single-locus association findings. Linear mixed models have become a cornerstone of GWAS due to their computational efficiency, straightforward interpretability, and broad software support. However, achieving sufficient statistical power with LMMs generally necessitates large sample sizes, which may pose practical limitations, particularly in studies of non-model organisms (Onifade et al., 2022). Despite advances in software that have improved scalability, large-scale

analyses still require access to high-performance computing infrastructure with substantial memory capacity (Runcie and Crawford, 2019; Schönherr et al., 2024). Methodologically, the univariate nature of LMMs restricts their capacity to capture the shared effects of loci on multiple traits, limiting their utility in the context of highly pleiotropic variants.

In contrast, Bayesian multi-locus approaches, such as the Bayesian Sparse Linear Mixed Model (BSLMM) proposed by Zhou et al. (2013), offer a probabilistic framework that allows for the simultaneous estimation of all marker effects, rather than testing each SNP independently. This enables the modeling of both sparse, large-effect variants and the cumulative influence of many small-effect loci, making Bayesian approaches particularly well-suited for traits with complex or polygenic architectures. Rather than relying on *p*-values, these models generate posterior inclusion probabilities (PIPs), which quantify the likelihood that a given SNP is truly associated with the trait. This facilitates the ranking of variants based on biological relevance and statistical confidence, offering a more informative and interpretable alternative to traditional significance thresholds (Pleić et al., 2022). Additionally, BSLMM addresses population structure and relatedness using a kinship matrix, and it accounts for linkage disequilibrium (LD) by estimating SNP effect sizes while controlling for other SNPs in the model. In addition, BSLMM makes it possible to estimate how much of the variation in a trait can be explained by genetic factors, a concept known as narrow-sense heritability. This includes both the overall contribution of all genotyped SNPs, known as the proportion of variance explained by all available genotypes (PVE) or narrow-sense heritability, as well as an additional measure provided by BSLMM, referred to as the proportion of genetic variance explained by variants with major effect (PGE) (Alamin et al., 2022). This helps distinguish between traits influenced by many small-effect variants and those driven by a few major ones. Despite their interpretability and robustness, Bayesian models remain computationally intensive, which can pose practical challenges in resource-limited settings (Zhao et al., 2019; Sun et al., 2021).

In parallel, multi-locus genome-wide association analyses (mGWAS) using multivariate linear mixed models (mvLMMs), such as those implemented in GEMMA, offer an extension that enables the simultaneous analysis of multiple correlated traits. This approach facilitates the identification of shared genetic variants that influence several traits at once, enhancing power and biological insight. While LMMs yield *p*-values to assess statistical significance, they do not provide information about the strength of associations (effect sizes) or the likelihood that a specific genetic

variant is truly causal and thus offer limited insight into biological relevance or credibility of effect (Yoon et al., 2021). This limits their utility in interpreting complex, polygenic traits.

Combining frequentist and Bayesian approaches within the same framework enables cross-validation of associations, mitigates method-specific limitations, and supports more accurate characterization of genetic architectures underlying adaptation. As a final step in many plant GWA studies, functional annotation of genomic regions surrounding loci under selection frequently reveals genomic regions associated with key protein families involved in essential biological pathways.

2.3 Genome-environment association (GEA) analyses

In a letter to Karl Freiesleben in June 1799, the renowned naturalist and explorer Alexander von Humboldt wrote: “I shall endeavor to find out how nature's forces act upon one another, and in what manner the geographic environment exerts its influence on animals and plants. In short, I must find out about the harmony in nature.” Humboldt’s pioneering work laid the groundwork for the modern study of organism–environment interactions and continues to inspire contemporary methods such as genome–environment association (GEA) analysis.

GEA has become a key approach for identifying genetic variation underlying adaptation to natural environments (Halpin-McCormick et al., 2025). It builds on the principle that populations distributed across contrasting climates and habitats experience spatially varying selection pressures, including variations in temperature, precipitation, altitude, and soil composition, which shape allele frequencies (Kawecki and Ebert, 2004; Hoban et al., 2016). Extending the GWAS framework introduced in the previous chapter, GEA analyses operate similarly, but instead of associating genetic variation with phenotypic traits, they model the relationship between genetic markers and environmental variables across populations (Cortés et al., 2022). This allows the identification of genomic regions potentially involved in local adaptation, even in the absence of phenotypic data, making GEA particularly valuable for studying non-model organisms (Hancock et al., 2011; Forester et al., 2018). Moreover, by capturing subtle allele frequency shifts across ecological gradients, GEA provides critical insight into the evolutionary processes shaping genetic diversity within species (Pritchard and Di Rienzo, 2010; Via, 2012).

GEA analysis is typically preceded by assessing population genetic structure through approaches such as principal component analysis (PCA), ancestry inference (e.g., sNMF), and

phylogenetic reconstructions (e.g., Nei's distance) (Nei, 1972; McVean, 2009; Frichot et al., 2014), respectively. These steps are essential to account for background genetic variation and to minimize confounding effects in subsequent analyses (Dauphin et al., 2023). Once population structure is characterized, multivariate statistical frameworks, most notably redundancy analysis (RDA), are employed to examine the relationship between genetic variation and environmental gradients (Ter Braak, 1987; Capblancq and Forester, 2021). RDA is particularly effective for detecting subtle shifts in allele frequencies across ecological conditions while simultaneously controlling for population structure. To maximize the accuracy and interpretability of this method, environmental variables, often sourced from global bioclimatic datasets such as WorldClim, must be carefully screened for multicollinearity, retaining only uncorrelated predictors in the model (Legendre and Legendre, 1998).

Similar to the concluding phase of GWAS, loci showing strong associations with environmental variables in GEA are identified as outliers, and their functional context can be explored through functional annotation against several available databases (e.g., EggNog, NCBI, SwissProt, etc.). This analytical framework allows for a robust investigation of how genetic variation is shaped by environmental heterogeneity across landscapes.

3 Thesis Outline

This PhD thesis investigated the genomic basis of local adaptation and ecological divergence in *Chouardia litardierei*, a non-model perennial with exceptional ecological plasticity across heterogeneous karst habitats. The primary aim was to identify the genetic determinants of variation in phenological and reproduction-related morphological traits under contrasting environmental conditions. To achieve this, the study integrated high-quality genomic sequencing, extensive sampling, common garden experiment phenotyping, and statistical modelling. Genotypic data were integrated with phenotypic measurements for genome-wide association studies and with bioclimatic variables for genome–environment association analyses. This integrative framework enabled genome-scale characterization and fine-scale genotype–phenotype–environment interaction mapping, offering novel insights into the evolutionary processes shaping this understudied lineage.

3.1 Research objectives and hypotheses

This study is guided by overarching objectives and hypotheses that shaped its design, methodology, and interpretation, forming a cohesive framework for investigating evolutionary dynamics in *C. litardierei*. The following objectives define the scope of this research:

1. **Assemble and annotate the genome**

Generate a high-quality, chromosome-level reference genome for *C. litardierei*, including annotation of gene content to provide a foundation for population and functional genomic analyses.

2. **Identify genotype-phenotype associations (GWAS)**

Conduct univariate GWAS using both single-locus and multi-locus approaches, as well as multivariate GWAS to identify SNPs associated with phenological and reproductive trait variations, followed by functional annotation of significant genomic regions to identify candidate genes involved in trait regulation.

3. **Detect genome–environment associations (GEA)**

Conduct genome–environment association analyses to identify genomic regions associated with tested environmental variables, and functionally annotate them to uncover candidate genes involved in local adaptation.

Based on these objectives, the research was guided by the following hypotheses:

1. *C. litardierei* comprises three ecologically distinct groups of populations, each genetically and morphologically differentiated according to the habitats they inhabit (karst poljes, dolomite slopes, and coastal marshes).
2. Phenological and reproductive traits exhibit high heritability and are associated with specific genetic variants, suggesting a strong genetic basis for adaptive trait variation.
3. As a consequence of highly contrasting environmental conditions they inhabit, the trade-off between sexual and asexual reproduction differs substantially among groups of populations.

3.2 Overview of research publications

In order to test the hypotheses, a complex analysis was carried out using different molecular-ecological tools and approaches, such as phenotyping and phenological characterization based on a common garden experiment, genotyping, and complex statistical data processing. The results are presented in three standalone yet thematically integrated publications, each addressing specific objectives and hypotheses.

3.2.1 Publication I overview

This publication presented a chromosome-level genome assembly of *C. litardierei*, providing a foundational genomic resource for investigating ecological divergence and local adaptation in this non-model species. Using PacBio long-read sequencing combined with Hi-C scaffolding, a highly contiguous 3.7 Gbp genome was assembled, anchored to 13 chromosome-scale scaffolds, consistent with previously reported karyotypes. Genome annotation revealed that over 80% of the assembly is composed of repetitive elements, with LTR retrotransposons particularly abundant. This supports the assumption that genome size in *C. litardierei* is largely driven by repetitive content and that its chromosomal architecture aligns with previous cytogenetic observations. Beyond genome characterization, genome-scale data confirmed that *C. litardierei* is phylogenetically distinct from *Asparagus* and other Asparagales species with available draft genome assemblies.

Although comprehensive genotype–phenotype and genotype–environment analyses are presented in later chapters, it must be emphasized that this genome assembly established the essential platform for all downstream genomic analyses, including ddRAD-seq data generation and

processing, population structure analysis, GWAS, GEA analysis, and functional annotation of candidate genomic regions.

3.2.2 Publication II overview

This publication investigated the genetic basis of variation in phenological traits of *C. litardierei*, with particular attention to their ecological importance and potential role in local adaptation. A common garden experiment was established with 214 individuals from nine populations, three from each of the three presumed habitat groups. This setup allowed for the isolation of genetically based phenotypic differences and the assessment of four key phenological traits: Beginning of Sprouting (BOS), Beginning of Flowering (BOF), Flowering Period Duration (FPD), and Vegetation Period Duration (VPD). To investigate the genetic underpinnings of these traits, genome-wide SNP data were generated using ddRAD-seq across multiple populations included in the common garden experiment, and a GWAS was conducted using both single-locus and multi-locus models.

Phenological traits data did not reveal grouping of the studied populations following their habitat types. This analysis revealed numerous significant genotype–phenotype associations, supporting the hypothesis that phenological traits in *C. litardierei* are heritable and influenced by specific genetic variants. Narrow-sense heritability was high across all traits, with VPD reaching 86.95%, highlighting its potential role in adaptation and fitness. To assess the functional relevance of associated loci, genomic regions surrounding significant SNPs were annotated. This revealed variants in genes encoding protein families central to phenological regulation. Several SNPs were identified in genomic regions encoding proteins with key roles in phenological regulation, including LHP1 (chromatin-mediated flowering control), pentatricopeptide repeat proteins, PPR (flowering onset), PIF1 (sprouting and development), and cytokinin-related genes (flowering responses to nutrients and drought).

3.2.3 Publication III overview

This publication investigates the genomic basis of local adaptation and reproductive trait variation in *C. litardierei*, integrating environmental and phenotypic data to identify candidate loci underlying adaptive divergence. Building on genomic resources developed in previous work, the study combines GEA analyses and GWAS to address multiple research objectives.

To explore how genomic variation reflects ecological divergence, GEA analysis based on RDA was conducted on the same set of studied populations. Precipitation of the coldest quarter (winter precipitation) emerged as the strongest predictor, with numerous SNPs being significantly associated with the climatic gradients. Recognized genomic regions of significant importance were functionally annotated, revealing candidate genes involved in abiotic stress responses, including pathways related to drought and cold tolerance. Key examples include kinase domains linked to salt tolerance and ion homeostasis, MYB transcription factors enhancing water stress tolerance, START domain proteins involved in drought signaling, Rubisco-related genes associated with heat sensitivity, and PPR genes related to developmental regulation.

Population-genetic analyses revealed partial genetic clustering, with the dolomite group from dry, drought-prone habitats forming a distinct genetic cluster. In contrast, populations from seashore and meadow habitats showed weak or no genetic structuring, likely reflecting recent shared ancestry or ongoing gene flow. This pattern only partially supported the hypotheses, indicating that while the dolomite group forms a distinct group, the lack of consistent divergence across all habitat types points to incomplete habitat type-based genetic differentiation in *C. litardierei.*

To investigate the genetic basis of reproductive trait development, phenotypic data from 214 individuals from a common garden experiment, relocated from nine populations across three habitat groups, were paired with ddRAD-seq genotyping. The study focused on three reproductive traits: two related to sexual reproduction, Average Height of Inflorescences (AHI) and Total Flower Count (TFC), and one asexual trait, Bulb Count (BC). Using both single-locus and multi-locus GWAS models, multiple significant associations were identified across the genome. All traits showed high narrow-sense heritability (>55%), with AHI reaching 71.95%, suggesting a key role in reproductive fitness and adaptive differentiation. Similar to phenological traits, recognized variations in reproduction-related traits were, for the most part, unrelated to the three assumed groups of populations, but were mostly population-specific. However, results showed that clonal reproduction (reflected in BC) was substantially more prevalent in flood-prone habitats, pointing to adaptive shifts in reproductive strategy under certain environmental constraints.

Functional annotation of candidate loci associated with reproduction-related morphological traits identified key genes involved in nitrogen metabolism, phytohormone signaling, and floral organ development, including cytochrome P450 enzymes (gibberellin biosynthesis, plant stature),

sugar transporters (bulb formation and starch accumulation), sterol biosynthesis genes (tissue morphogenesis and reproductive structures), as well as arginase genes, CCHC-ZFPs, aspartic proteases, Complex I genes, receptor-like kinases, and C2 domain proteins—highlighting pathways central to growth, reproduction, and adaptation in diverse habitats.

A schematic summary of the methodological workflow for all three publications is provided in Figure 4, offering a visual guide to the overall study design.

3.3 Scientific contribution

This thesis presents the first high-quality, chromosome-scale genome assembly of *C. litardierei*, providing a foundational resource for genomic research on this ecologically distinctive and previously understudied species. By integrating this genomic resource with extensive ecological, phenotypic, and population-genomic data, the study delivers a comprehensive insight into the genetic architecture underlying local adaptation and reproductive strategies in a non-model species.

This work pioneers the application of GWAS and GEA approaches in *C. litardierei*, identifying significant associations between genetic variants and key phenological and reproductive traits, as well as some of the tested climatic drivers. The research also includes the first large-scale common garden experiment in this species, enabling robust heritability estimates under standardized conditions—a critical step in disentangling genetic effects from environmental influences.

The results reveal a complex pattern of incipient divergence, with only dolomite populations showing clear signs of genetic differentiation. Interestingly, virtually all tested traits were not specifically linked to any of the population groups, but rather to a specific population. These findings not only shed light on local adaptation in *C. litardierei* but also establish a powerful framework for investigating evolutionary processes in other South-European taxa inhabiting fragmented and heterogeneous environments. By linking highly heritable traits to genomic regions involved in stress response and developmental regulation, the study establishes a reference point for future work on adaptation, speciation, and trait evolution in geophytes and other plant species from contrasting habitats.

Together, these findings establish *C. litardierei* as one of the few Balkan endemics thoroughly characterized across genomic, ecological, and phenotypic dimensions, using GWAS

and GEA analyses based on a common garden experiment. The genomic resources and analytical framework developed here will support future evolutionary and ecological research, not only within *Chouardia* or the Asparagaceae, but more broadly across other biodiversity hotspots where adaptive processes remain poorly understood.

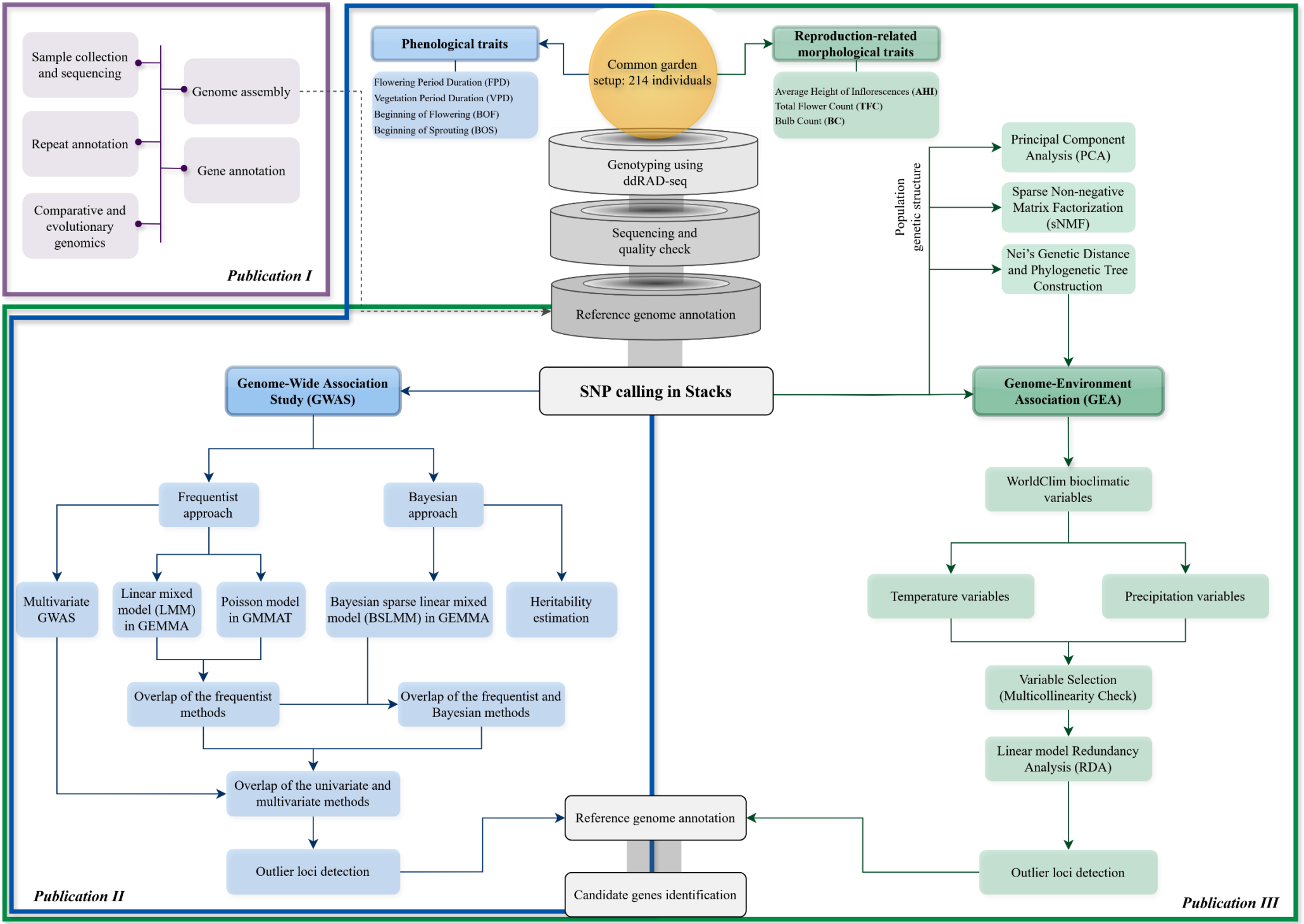


Figure 4. *Schematic overview of the research framework and workflow across the three publications. Purple outlines correspond to Publication I, blue to Publication II, and green to Publication III.*

4 Discussion

This thesis presents an integrative investigation of the genetic basis of local adaptation in *Chouardia litardierei*, a geophytic wild species native to the Balkan Peninsula. Through three original publications, it addresses key questions related to genome structure, population-genetic patterns, genotype-environment, and genotype-phenotype relationships. Collectively, these studies form a coherent framework for understanding early-stage divergence in species that occupy environmentally heterogeneous habitats.

4.1 Broader evolutionary context and implications

Despite limited prior research, *C. litardierei* presents a valuable system for investigating the early stages of ecological divergence in non-model species. Its distribution spans across highly heterogeneous environments, ranging from coastal marshes and wet meadows to high-elevation dolomite slopes (Gaži-Baskova, 1962; Šilić, 1990), making it a suitable model for examining how local selection pressures shape the genetics underlying reproductive strategies, phenology, and population structure (Lowry et al., 2008b; Cortés et al., 2018; Bakhtiari et al., 2019).

The results presented in this thesis point to early-stage ecological divergence in *C. litardierei*, particularly of the southern dolomite populations. Although the dolomite populations are genetically differentiated and uniform in terms of the geological substrate on which they grow, it should be noted that they also occur over a wide range of altitudes, thus experiencing different climatic conditions. This could easily have led to the development of variations that are specific to a particular population and not common across the entire group. These environmental differences render this genetically homogeneous dolomite group heterogeneous in terms of phenology and reproductive morphology. While divergence remains incomplete, the combination of environmental isolation and trait shifts suggests that these populations may be on an evolutionary trajectory toward ecological speciation. Such patterns are consistent with theoretical expectations under limited gene flow and strong local selection (Hoban et al., 2016; Rahbek et al., 2019).

The results underscore the value of integrating genomics, trait analysis, and environmental data to uncover cryptic evolutionary processes not yet reflected in morphological differentiation (Theissinger et al., 2023; Peng et al., 2025). In this context, *C. litardierei* contributes to a broader understanding of how adaptation unfolds in fragmented landscapes, where gene flow, environmental heterogeneity, and reproductive strategy interact in complex ways.

4.2 Building genomic infrastructure for ecological and evolutionary inference

As a wild species without economic value and no prior genomic resources, *C. litardierei* exemplifies the challenges and growing potential of applying modern genomic tools to ecologically important but understudied taxa. Its distribution across diverse habitats made it an ideal candidate for studying local adaptation, yet until recently, the absence of a reference genome limited deeper evolutionary and functional insights.

The development of a chromosome-scale reference genome (**Publication I**) addressed this limitation, providing the foundational resource needed to link genotypes to ecological and phenotypic variation. Using PacBio HiFi long-read sequencing combined with Hi-C scaffolding, we assembled a high-quality and complete genome of ~3.7 Gb, anchored to 13 pseudochromosomes, thus supporting previous cytogenetic data (Siljak-Yakovlev, 2010). Additionally, genome annotation revealed that over 80% of the assembly consists of repetitive elements, explaining the species' large genome size.

This reference genome enabled key downstream analyses, including assessments of population structure, trait heritability, GWAS, and GEA (**Publications II** and **III**). It has thus established the genomic framework necessary for linking adaptive traits to specific loci and for interpreting local adaptation at both the genetic and ecological levels.

Importantly, the successful assembly of a large, repetitive genome from a non-model species underscores the feasibility of generating high-quality genomic resources for other Balkan endemics, many of which remain poorly characterized despite their ecological significance (Quaresma et al., 2024). Beyond its immediate application in this thesis, the *C. litardierei* genome provides a springboard for future research on gene regulation, structural variation, and responses to environmental stress, offering a broader insight into plant adaptation in fragmented and climatically variable landscapes.

4.3 Ecotypic differentiation and genetic structure in *C. litardierei*

Publication I laid the foundation for further research by revealing preliminary evidence of divergence, suggesting an isolated lineage, and motivating broader analyses.

Publications II and **III** tested the genetic structuring of the studied populations to reveal that dolomite populations form a distinct lineage. In addition, as a prerequisite for the GWAS analysis that followed, the common garden experiment was set up, and the results obtained provided additional insight into the phenotypic structure of the studied populations. Population genetic analyses revealed that dolomite populations form a well-differentiated genetic cluster, whereas meadow and seashore populations were genetically indistinct from each other,

suggesting either a recent shared origin or substantial ongoing gene flow that prevents differentiation.

While these results clarify broad patterns of population structure, they do not support the idea that *C. litardierei* consists of three fully distinct ecotypes aligned with the major habitat types these groups inhabit. Although some divergence was evident, most notably for the dolomite group of populations, this variation was not consistent across all habitat groups. Instead, genetic structure and trait differentiation were more strongly associated with local environmental conditions and population-specific responses than with broad habitat categories. According to Lowry's (2012) definition, ecotypes require both genetic and phenotypic distinctiveness. Based on this criterion, *C. litardierei* does not qualify as a species with fully developed ecotypes. We therefore adopt a more conservative interpretation, referring to these as habitat-associated population groups, which acknowledges the presence of partial divergence without overstating its evolutionary significance. Morphologically and phenologically, dolomite and non-dolomite populations are largely indistinguishable, with the only notable difference being a substantially greater prevalence of clonal reproduction (reflected in BC trait) in flood-prone habitats—likely representing an adaptive shift in reproductive strategy under a set of specific environmental constraints.

In summary, population-genetic and phenotypic analyses revealed only partial divergence in *C. litardierei*, with clear genetic separation of dolomite populations but overall morphological similarity across habitats, apart from greater clonal reproduction in flood-prone sites. This mismatch between genetic structure and phenotypic expression can probably be explained by individual differences in the microclimatic and other environmental conditions to which individual populations are exposed, thus providing the evolutionary backdrop for interpreting the heritable basis and adaptive potential of phenological and reproductive traits in *C. litardierei*.

4.4 Trait variation and genetic architecture of adaptation

Patterns of phenological and reproductive trait variation in *C. litardierei* suggest that environmental pressures are shaping population-level divergence, as demonstrated through common garden experiment, heritability estimates, and genome-wide association studies (**Publications II** and **III**). By minimizing environmental noise in a controlled setting (De Villemereuil et al., 2015), we identified genetically based variation, though its alignment with ecological differences was not consistent across all traits or habitat types.

GWAS results revealed associations between both phenological and reproductive traits and numerous loci in genomic regions involved in flowering regulation, hormone signaling, stress response, and developmental pathways mechanisms broadly implicated in plant adaptation. Several SNPs were identified in genomic regions encoding protein families with key roles in phenological regulation. These included LHP1, a chromo domain protein controlling flowering time through chromatin-based regulation; pentatricopeptide repeat (PPR) proteins, which affect flowering onset via post-transcriptional processes; phytochrome-interacting factor 1 (PIF1), regulating sprouting and developmental transitions; and cytokinin-related genes mediating nutrient- and drought-linked flowering responses. Additional associations were detected with histidine phosphatases, which regulate hormone signaling and vegetative growth, and stress-related genes linked to drought and cold responses. Others were linked to reproduction-related morphological traits, including cytochrome P450 enzymes regulating gibberellin biosynthesis and plant stature, sugar transporters influencing bulb formation and starch accumulation, and sterol biosynthesis genes affecting tissue morphogenesis and reproductive structure development. Additional associations involved arginase genes central to nitrogen metabolism and growth, CCHC-ZFPs regulating development and stress adaptation, aspartic proteases driving rapid organ development, Complex I genes essential for growth at all life stages, receptor-like kinases mediating brassinosteroid signaling, and C2 domain proteins (e.g., QUIRKY, STRUBBELIG) crucial for intercellular communication in reproductive tissues.

These associations were reinforced by high narrow-sense heritability (h^2) estimates, which describe the degree of variation in a phenotypic trait that is due to genetic variation. For phenology, vegetative period duration (VPD) showed the highest h^2 (86.95%), suggesting strong genetic control and possible adaptive significance in environments with limited growing seasons. In contrast, flowering period duration (FPD) showed lower heritability (20.26%), implying greater plasticity or environmental sensitivity. Moreover, the BSLMM analysis revealed that 66.03% of the phenotypic variation in BOF (beginning of flowering) and 76.05% in BOS (beginning of sprouting) was explained by all genotypes. These findings demonstrate that both early-season sprouting and flowering onset are strongly shaped by genetic factors, underscoring the significant genetic control of key phenological transitions. This genetic regulation may play a critical role in enabling populations to synchronize their life cycle with the environmental constraints of their specific habitats.

These genomic findings partially align with patterns observed in the common garden experiment. On average, dolomite populations sprouted later, flowered earlier, and had shorter

vegetation periods than meadow and seashore populations, but notable exceptions were present. For example, the seashore population from Pag flowered at the same time as dolomite populations and significantly earlier than the nearby Vrana Lake population from the same habitat. Such variation indicated that phenological differences are not strictly determined by habitat group, suggesting additional influences from population-specific factors or microenvironmental variation. Together, these results suggest that phenological variation in *C. litardierei* occurs along a continuum of local adaptation rather than as discrete, habitat-specific shifts.

Reproductive traits exhibited complex patterns that varied more among populations than across broad habitat types, suggesting that localized selective pressures, such as pollinator communities or microhabitat variability, may exert stronger influence than overarching abiotic effects. This interpretation is supported by morphometric analyses, which revealed substantial population-level variation, and by heritability estimates showing that the average height of inflorescence (AHI) had the highest narrow-sense heritability (71.95%), followed by bulb count (BC, 69.87%) and total flower count (TFC, 55.89%). Although AHI showed high heritability and may be linked to pollination efficiency, its similar expression across habitat groups suggests that any potential selection acting on this trait is not strongly habitat-specific, making it difficult to draw firm conclusions about directional selection. In contrast, variation in TFC and BC appears more tightly linked to environmental conditions. For instance, dolomite populations tended to produce fewer bulbs, possibly reflecting a shift toward sexual reproduction in more hydrologically stable, drought-prone environments. Conversely, increased clonal reproduction in flood-prone habitats like karst poljes and coastal meadows may represent an adaptive response to unpredictable opportunities for sexual reproduction. This pattern is consistent with the hypothesis that clonal reproduction is more prevalent in flood-prone environments as a strategy to ensure persistence under unstable reproductive conditions.

Despite environmental contrasts, considerable overlap in trait values among habitat groups indicates that adaptive divergence in *C. litardierei* is not uniformly shaped by broad habitat categories. Instead, it reflects a mosaic of localized selection, genetic background, and the lasting effects of past evolutionary events. These findings support the hypothesis, which proposes that phenological and reproductive traits in *C. litardierei* exhibit high heritability and are associated with specific genetic variants. This nuanced pattern highlights how both trait types contribute to early-stage adaptation in complex and variable environments, reinforcing the importance of integrating ecological, genetic, and trait-based data when investigating evolutionary processes in non-model systems.

4.5 Genome–environment associations and local adaptation

To further understand the environmental factors driving genetic differentiation in *C. litardierei*, we explored GEA to identify loci potentially involved in local adaptation (**Publication III**). This approach complements trait-based and population genetic analyses by linking specific climatic variables to patterns of genetic variation (Faske et al., 2021; Dauphin et al., 2023). While GEA results can be affected by environmental collinearity and population structure, they can nonetheless reveal broad signals of climate-mediated selection.

Among the tested climatic predictors, precipitation during the coldest quarter (BIO19) was recognized as the most profound driver of the detected variation. BIO19 was linked to 131 of 256 SNP outliers, indicating that among the tested variables, winter moisture availability may be a key selective force shaping genomic divergence in *C. litardierei*. This pattern is consistent with the species' fragmented distribution across environments that differ strongly in seasonal water availability—from karst fields, where early-spring flooding results from rapid snowmelt in surrounding mountains combined with basin-like geomorphology and limited permeability of the substrate, to arid, drought-prone dolomite slopes (Horvatić, 1934; Bonacci, 2014). This aligns with the hypothesis, which proposes that genetic variation among populations is associated with environmental factors, indicating potential signatures of local adaptation.

To further explore the functional relevance of these associations, we annotated significant loci and identified over 80 genomic regions linked to abiotic stress responses and regulatory pathways, including salt and drought resistance, ion homeostasis, temperature resilience, photosynthesis, and developmental control. Key examples include C2 domains and protein kinase domains involved in salt tolerance and Na^+/K^+ homeostasis, MYB transcription factors enhancing salt and water stress tolerance, and START domain proteins linked to drought signaling. The presence of Rubisco-related genes, sensitive to heat stress, points to potential photosynthetic constraints in warmer environments, while DEAD-box and RRM1 genes indicate capacity for cold tolerance. Genes such as PPR and OBERON further highlight roles in developmental regulation. The alignment between these functional roles and contrasting ecological conditions suggests that local adaptation in *C. litardierei* is supported by a diverse physiological toolkit.

The diversity of detected genes indicates a polygenic basis of adaptation, with many small-effect loci collectively shaping population responses rather than a few large-effect genes (Hämälä et al., 2020). This polygenic pattern is common in species adapting to complex and

variable environments (Yeaman, 2015; Ehrlich et al., 2020). While these associations cannot confirm causal relationships, they provide a valuable starting point for future studies aiming to validate candidate gene functions and deepen our understanding of the molecular basis of local adaptation in *C. litardierei*.

4.6 Limitations and future objectives

While this thesis offers novel insights into the genomic and ecological landscape of *C. litardierei*, several limitations in study design, data quality, and analytical scope affect the certainty and generalizability of some conclusions. Addressing these gaps in future research will be essential for a more comprehensive understanding of the species' evolutionary dynamics.

Perhaps the major limitation lies in the genome assembly and the very limited accessibility of genome assemblies from other closely related species. Although the assembled genome was highly complete, only 44.5% of predicted genes were matched across all selected databases, indicating that more than half still lacked full functional annotation. This gap likely reflects the scarcity of genomic references for non-model monocots and the absence of closely related species for comparison. The large genome size and high repeat content (nearly 70%) also posed challenges for accurate annotation.

Sampling design and data limitations influenced both population- and trait-level analyses. Although the study spanned a wide ecological gradient, the number of populations and individuals per habitat type remained somewhat modest, possibly limiting the ability to detect rare variants, fine-scale genetic structure, and subtle patterns of divergence. The common garden experiment helped control for environmental effects, but focused primarily on phenological and reproductive traits, leaving other potentially adaptive morphological and physiological traits unexplored.

For the genotyping, the use of ddRAD-seq further constrained resolution. Because only a fraction of the genome was analyzed, this approach limited our ability to detect rare variants and selection signals that may lie outside the captured loci. These constraints had downstream effects, particularly for GWAS, where the combination of limited genomic coverage and small sample sizes reduced statistical power and increased the risk of confounding due to population structure. Despite efforts to control for multicollinearity, the potential influence of population structure and limited gene annotation constrained the interpretation of these results, especially regarding the functional significance of candidate loci recognized by GEA.

Several opportunities for future research were identified throughout this study. Based on the limitations outlined above, future work could further improve, extend, and validate the findings presented here. These directions include:

- Instead of the reduced-representation sequencing, implement whole-genome resequencing (WGS) of individuals across populations to detect rare variants. Individual-level WGS offers higher resolution of genetic diversity, gene flow, and demographic history.

- Expand sampling across the species' full ecological and geographic range. A broader sample set will better capture environmental gradients and potential ecotypic variation, improving inferences about adaptation.

- Analyze a broader set of traits beyond these tested ones, such as leaf morphology (size and number), root architecture, and bulb size. These traits may reveal additional axes of adaptation related to environmental stress.

- Include the analysis of the RNA through the implementation of various RNAseq approaches, to directly assess the gene expression in populations from different habitats.

- Establish reciprocal transplant experiments across habitats to test the fitness consequences of trait variation in contrasting environments. Such experiments can help distinguish phenotypic plasticity from genetic adaptation and provide direct evidence for local adaptation, which is crucial for confirming whether the observed habitat-associated groups represent true ecotypes.

- Incorporate soil variables into genome–environment association analyses to evaluate how contrasting edaphic conditions, from deep soils to exposed rocky substrates, shape genetic divergence.

5 Conclusion

The results of this thesis support most proposed hypotheses, revealing a complex interplay of genetic structure, trait variation, and early-stage ecological divergence in *Chouardia litardierei*. By integrating genomic, phenotypic, and environmental data, this work clarifies how selection and local adaptation interact in this ecologically diverse Balkan endemic.

The key findings of this PhD thesis can be summarized as follows:

1. **First chromosome-scale genome reveals evolutionary distinctiveness of *C. litardierei***

   A high-quality ~3.7 Gb genome was assembled, confirming the expected chromosome number and structure, and marking the first such resource for this species. Phylogenomic analyses placed *C. litardierei* as evolutionarily distant from *Asparagus* and other previously sequenced Asparagales, underscoring its phylogenetic uniqueness.

2. **Genetic structure separates dolomite populations as a distinct group**

   Population genomic analyses revealed two broad genetic clusters: one confined to dolomite habitats, and another spanning meadow and seashore sites. This pattern supports describing populations as habitat-associated groups rather than distinct ecotypes.

3. **Heritable trait variation is underpinned by specific genetic loci**

   GWAS revealed high heritability in phenological and reproductive traits and identified SNPs in genomic regions related to hormone signaling, stress response, and developmental timing. These findings suggested a strong genetic basis for variation in the tested adaptive traits and supported fine-scale natural selection across heterogeneous environments.

4. **Genetic differentiation is shaped by environmental factors, especially winter precipitation**

   GEA analyses identified winter precipitation as the environmental element with the strongest influence on genomic variations, indicating that water availability is a major selective force shaping local adaptation across the species' range.

5. **Clonal reproduction is more common in flood-prone habitats**

   Populations from flood-prone habitats, such as karst poljes and coastal meadows, showed a stronger tendency toward clonal reproduction. This supports a habitat-specific life-history trade-off, as clonality is often favoured in environments where, for some reason, sexual reproduction is threatened or unreliable.

6. Adaptive divergence is ongoing, partial, and shaped by local ecological conditions

Despite pronounced environmental contrasts, both genetic structure and trait differentiation remain incomplete, particularly among non-dolomite populations. This suggests that local adaptation in *C. litardierei* proceeds along a continuum shaped by different evolutionary forces, microhabitat variability, and historical demographic processes.

7. Adaptive traits appear to be governed by a polygenic basis

Environment- and trait-associated SNPs mapped to genes involved in drought tolerance, flowering regulation, and stress responses, indicating that adaptation in *C. litardierei* is shaped by shifts across multiple biological pathways rather than by single large-effect mutations.

6 Literature

Adole, T., Dash, J., Rodriguez-Galiano, V., and Atkinson, P. M. (2019). Photoperiod controls vegetation phenology across Africa. *Communications Biology 2019 2:1* 2, 1–13. doi: 10.1038/s42003-019-0636-7

Alamin, M., Humaira Sultana, M., Lou, X., Jin, W., and Xu, H. (2022). Dissecting Complex Traits Using Omics Data: A Review on the Linear Mixed Models and Their Application in GWAS. *Plants 2022, Vol. 11, Page 3277* 11, 3277. doi: 10.3390/PLANTS11233277

Alegro, A. (2013). Nacionalni programi za praćenje stanja očuvanosti vrsta u Hrvatskoj – Livadni procjepak (Chouardia litardierei (Breistr.) Speta).

Alseekh, S., Kostova, D., Bulut, M., and Fernie, A. R. (2021). Genome-wide association studies: assessing trait characteristics in model and crop plants. *Cellular and Molecular Life Sciences 2021 78:15* 78, 5743–5754. doi: 10.1007/S00018-021-03868-W

Anderson, J. T., Inouye, D. W., McKinney, A. M., Colautti, R. I., and Mitchell-Olds, T. (2012). Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. *Proceedings of the Royal Society B: Biological Sciences* 279, 3843–3852. doi: 10.1098/RSPB.2012.1051

Atif, M. J., Ahanger, M. A., Amin, B., Ghani, M. I., Ali, M., and Cheng, Z. (2020). Mechanism of Allium Crops Bulb Enlargement in Response to Photoperiod: A Review. *Int J Mol Sci* 21. doi: 10.3390/IJMS21041325

Bakhtiari, M., Formenti, L., Caggìa, V., Glauser, G., and Rasmann, S. (2019). Variable effects on growth and defense traits for plant ecotypic differentiation and phenotypic plasticity along elevation gradients. *Ecol Evol* 9, 3740–3755. doi: 10.1002/ECE3.4999

Bamba, M., Kawaguchi, Y. W., and Tsuchimatsu, T. (2019). Plant adaptation and speciation studied by population genomic approaches. *Dev Growth Differ* 61, 12–24. doi: 10.1111/DGD.12578

Barrett, S. C. H. (2015). Influences of clonality on plant sexual reproduction. *Proc Natl Acad Sci U S A* 112, 8859–8866. doi: 10.1073/PNAS.1501712112

Bayarri, M. J., and Berger, J. O. (2004). The Interplay of Bayesian and Frequentist Analysis. *https://doi.org/10.1214/088342304000000116* 19, 58–80. doi: 10.1214/088342304000000116

Beeftink, W. G. (1977). "Salt-marshes," in *Coastline*, ed. R. S. K. Barnes (New York: John Wiley and Sons), 93–121.

Berhe, M., Dossa, K., You, J., Mboup, P. A., Diallo, I. N., Diouf, D., et al. (2021). Genome-wide association study and its applications in the non-model crop Sesamum indicum. *BMC Plant Biology 2021 21:1* 21, 1–19. doi: 10.1186/S12870-021-03046-X

Bernatchez, L., Ferchaud, A. L., Berger, C. S., Venney, C. J., and Xuereb, A. (2023). Genomics for monitoring and understanding species responses to global climate change. *Nature Reviews Genetics 2023 25:3* 25, 165–183. doi: 10.1038/s41576-023-00657-y

Bonacci, O. (2014). "Ecohydrology of karst poljes and their vulnerability," in *Dinaric karst poljes - floods for life*, (Radolfzell am Bodensee: EuroNatur), 25–37.

Brachi, B., Villoutreix, R., Faure, N., Hautekèete, N., Piquot, Y., Pauwels, M., et al. (2013). Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in Arabidopsis thaliana. *Mol Ecol* 22, 4222–4240. doi: 10.1111/MEC.12396

Brandrud, M. K., Paun, O., Lorenzo, M. T., Nordal, I., and Brysting, A. K. (2017). RADseq provides evidence for parallel ecotypic divergence in the autotetraploid Cochlearia officinalis in Northern Norway. *Scientific Reports 2017 7:1* 7, 1–13. doi: 10.1038/s41598-017-05794-z

Bremer, B., Bremer, K., Chase, M. W., Fay, M. F., Reveal, J. L., Bailey, L. H., et al. (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* 161, 105–121. doi: 10.1111/J.1095-8339.2009.00996.X

Bui, E. N. (2013). Soil salinity: A neglected factor in plant ecology and biogeography. *J Arid Environ* 92, 14–25. doi: 10.1016/J.JARIDENV.2012.12.014

Campbell, C. R., Poelstra, J. W., and Yoder, A. D. (2018). What is Speciation Genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biological Journal of the Linnean Society* 124, 561–583. doi: 10.1093/BIOLINNEAN/BLY063

Capblancq, T., and Forester, B. R. (2021). Redundancy analysis: A Swiss Army Knife for landscape genomics. *Methods Ecol Evol* 12, 2298–2309. doi: 10.1111/2041-210X.13722

CaraDonna, P. J., Iler, A. M., and Inouye, D. W. (2014). Shifts in flowering phenology reshape a subalpine plant community. *Proc Natl Acad Sci U S A* 111, 4916–4921. doi: 10.1073/PNAS.1323073111

Chen, H., Huffman, J. E., Brody, J. A., Wang, C., Lee, S., Li, Z., et al. (2019). Efficient Variant Set Mixed Model Association Tests for Continuous and Binary Traits in Large-Scale Whole-Genome Sequencing Studies. *Am J Hum Genet* 104, 260–274. doi: 10.1016/j.ajhg.2018.12.012

Clausen, J. (1951). *Stages in the evolution of plant species.* Cornell University Press, Ithaca, New York.

Collins, C. G., Angert, A. L., Clark, K., Elmendorf, S. C., Elphinstone, C., and Henry, G. H. R. (2025). Flowering time responses to warming drive reproductive fitness in a changing Arctic. *Ann Bot* 135, 255–268. doi: 10.1093/AOB/MCAE007

Cook, B. I., Wolkovich, E. M., Davies, T. J., Ault, T. R., Betancourt, J. L., Allen, J. M., et al. (2012). Sensitivity of Spring Phenology to Warming Across Temporal and Spatial Climate Gradients in Two Independent Databases. *Ecosystems* 15, 1283–1294. doi: 10.1007/S10021-012-9584-5

Cortés, A. J., Garzón, L. N., Valencia, J. B., and Madriñán, S. (2018). On the causes of rapid diversification in the páramos: Isolation by ecology and genomic divergence in espeletia. *Front Plant Sci* 871, 408949. doi: 10.3389/FPLS.2018.01700

Cortés, A. J., López-Hernández, F., and Blair, M. W. (2022). Genome–Environment Associations, an Innovative Tool for Studying Heritable Evolutionary Adaptation in Orphan Crops and Wild Relatives. *Front Genet* 13, 910386. doi: 10.3389/FGENE.2022.910386

Cortés-Flores, J., Hernández-Esquivel, K., González-Rodríguez, A., and Ibarra-Manríquez, G. (2017). Flowering phenology, growth forms, and pollination syndromes in tropical dry forest species: Influence of phylogeny and abiotic factors. *Am J Bot* 104, 39–49. doi: 10.3732/AJB.1600305

Dauphin, B., Rellstab, C., Wüest, R. O., Karger, D. N., Holderegger, R., Gugerli, F., et al. (2023). Re-thinking the environment in landscape genomics. *Trends Ecol Evol* 38, 261–274. doi: 10.1016/J.TREE.2022.10.010

Day, M., and Chenoweth, S. (2013). Surface Roughness of Karst Landscapes. *Treatise on Geomorphology: Volume 1-14* 1–14, 157–163. doi: 10.1016/B978-0-12-374739-6.00108-1

De Villemereuil, P., Gaggiotti, O. E., Mouterde, M., and Till-Bottraud, I. (2015). Common garden experiments in the genomic era: new perspectives and opportunities. *Heredity 2016 116:3* 116, 249–254. doi: 10.1038/hdy.2015.93

Ehrlich, M. A., Wagner, D. N., Oleksiak, M. F., and Crawford, D. L. (2020). Polygenic Selection within a Single Generation Leads to Subtle Divergence among Ecological Niches. *Genome Biol Evol* 13, evaa257. doi: 10.1093/GBE/EVAA257

Engelmann, K., and Purugganan, M. (2006). The Molecular Evolutionary Ecology of Plant Development: Flowering Time in Arabidopsis thaliana. *Adv Bot Res* 44, 507–526. doi: 10.1016/S0065-2296(06)44013-1

Esposito, S., Cardi, T., Campanelli, G., Sestili, S., Díez, M. J., Soler, S., et al. (2020). ddRAD sequencing-based genotyping for population structure analysis in cultivated tomato provides new insights into the genomic diversity of Mediterranean 'da serbo' type long shelf-life germplasm. *Horticulture Research 2020 7:1* 7, 1–14. doi: 10.1038/s41438-020-00353-6

Faske, T. M., Agneray, A. C., Jahner, J. P., Sheta, L. M., Leger, E. A., and Parchman, T. L. (2021). Genomic and common garden approaches yield complementary results for quantifying environmental drivers of local adaptation in rubber rabbitbrush, a foundational Great Basin shrub. *Evol Appl* 14, 2881. doi: 10.1111/EVA.13323

Fekete, G., Tölgyesi, G., and Horánszky, A. (1989). Dolomite versus Limestone Habitats: a Study of Ionic Accumulation on a Broader Floristic Basis1)1)Dedicated to Prof. Dr. H. M Eusel on the ocassion of his 80th birthday. *Flora* 183, 337–348. doi: 10.1016/S0367-2530(17)31565-7

Fernández-Meirama, M., Rolán-Alvarez, E., and Carvajal-Rodríguez, A. (2022). A Simulation Study of the Ecological Speciation Conditions in the Galician Marine Snail Littorina saxatilis. *Front Genet* 13, 680792. doi: 10.3389/FGENE.2022.680792

Flood, P. J., and Hancock, A. M. (2017). The genomic basis of adaptation in plants. *Curr Opin Plant Biol* 36, 88–94. doi: 10.1016/J.PBI.2017.02.003

Ford, D., and Williams, P. (2013). Karst Hydrogeology and Geomorphology. *Karst Hydrogeology and Geomorphology*, 1–562. doi: 10.1002/9781118684986

Forester, B. R., Lasky, J. R., Wagner, H. H., and Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Mol Ecol* 27, 2215–2233. doi: 10.1111/MEC.14584

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., and François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics* 196, 973–983. doi: 10.1126/SCIENCE.1209244

Gams, I. (1978). The Polje: The Problem of Definition. With Special Reference to the Dinaric Karst. *Zeitschrift für Geomorphologie* 22, 170–181.

Gaudinier, A., and Blackman, B. K. (2020). Evolutionary processes from the perspective of flowering time diversity. *New Phytologist* 225, 1883–1898. doi: 10.1111/NPH.16205

Gaži-Baskova, V. (1962). Geografsko raširenje livadnog procjepka ili lučike (Scilla pratensis W. et K.). *Biološki Glas. 1962, 15, 49–54.* 15, 49–54.

Gong, W., Gu, L., and Zhang, D. (2010). Low genetic diversity and high genetic divergence caused by inbreeding and geographical isolation in the populations of endangered species Loropetalum subcordatum (Hamamelidaceae) endemic to China. *Conservation Genetics* 11, 2281–2288. doi: 10.1007/S10592-010-0113-9

Gregory, T. R. (2009). Understanding Natural Selection: Essential Concepts and Common Misconceptions. *Evolution: Education and Outreach* 2, 156–175. doi: 10.1007/S12052-009-0128-1

Groen, S. C., and Whiteman, N. K. (2016). Using Drosophila to study the evolution of herbivory and diet specialization. *Curr Opin Insect Sci* 14, 66–72. doi: 10.1016/j.cois.2016.01.004

Halpin-McCormick, A., Campbell, Q., Negrão, S., Morrell, P. L., Hübner, S., Neyhart, J. L., et al. (2025). Environmental genomic selection to leverage polygenic local adaptation in barley landraces. *Communications Biology 2025 8:1* 8, 1–11. doi: 10.1038/s42003-025-08045-4

Hämälä, T., Guiltinan, M. J., Marden, J. H., Maximova, S. N., Depamphilis, C. W., and Tiffin, P. (2020). Gene Expression Modularity Reveals Footprints of Polygenic Adaptation in Theobroma cacao. *Mol Biol Evol* 37, 110–123. doi: 10.1093/MOLBEV/MSZ206

Hancock, A. M., Brachi, B., Faure, N., Horton, M. W., Jarymowycz, L. B., Sperone, F. G., et al. (2011). Adaptation to climate across the Arabidopsis thaliana genome. *Science (1979)* 334, 83–86. doi: 10.1126/SCIENCE.1209244

Harvey, M. G., Smith, B. T., Glenn, T. C., Faircloth, B. C., and Brumfield, R. T. (2016). Sequence Capture versus Restriction Site Associated DNA Sequencing for Shallow Systematics. *Syst Biol* 65, 910–924. doi: 10.1093/SYSBIO/SYW036

Heslop-Harrison, J. (1964). Forty Years of Genecology. *Adv Ecol Res* 2, 159–247. doi: 10.1016/S0065-2504(08)60332-3

Hirschhorn, J. N., and Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics 2005 6:2* 6, 95–108. doi: 10.1038/nrg1521

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., et al. (2016). Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *Am Nat* 188, 379–397. doi: 10.1086/688018

Horvatić, S. (1934). Flora i vegetacija otoka Paga. *Prirodoslovna istraživanja Kraljevine Jugoslavije* 19, 116–372.

Hu, Z. M., Zhong, K. Le, Weinberger, F., Duan, D. L., Draisma, S. G. A., and Serrão, E. A. (2020). Linking Ecology to Genetics to Better Understand Adaptation and Evolution: A Review in Marine Macrophytes. *Front Mar Sci* 7, 545102. doi: 10.3389/FMARS.2020.545102

Huang, X., Duan, X., Ma, S., Sun, X., Geng, Q., Sun, X., et al. (2025). Optimizing genotype imputation strategies for low coverage whole genome and 2b-RAD sequencing in bivalve molluscs. *Aquaculture*, 742713. doi: 10.1016/J.AQUACULTURE.2025.742713

Huisman, L. (2023). Are P-values and Bayes factors valid measures of evidential strength? *Psychon Bull Rev* 30, 932–941. doi: 10.3758/S13423-022-02205-X

Jennings, J. N. (1985). Karst geomorphology. Second edition. *Karst geomorphology. Second edition.*

John, M., Korte, A., and Grimm, D. G. (2024). The benefits of permutation-based genome-wide association studies. *J Exp Bot* 75, 5377. doi: 10.1093/JXB/ERAE280

Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S. Y., Freimer, N. B., et al. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics 2010 42:4* 42, 348–354. doi: 10.1038/ng.548

Kaul, S., Koo, H. L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L. J., et al. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature 2000 408:6814* 408, 796–815. doi: 10.1038/35048692

Kawakami, T., Morgan, T. J., Nippert, J. B., Ocheltree, T. W., Keith, R., Dhakal, P., et al. (2011). Natural selection drives clinal life history patterns in the perennial sunflower species, Helianthus maximiliani. *Mol Ecol* 20, 2318–2328. doi: 10.1111/J.1365-294X.2011.05105.X

Kawecki, T. J., and Ebert, D. (2004). Conceptual issues in local adaptation. *Ecol Lett* 7, 1225–1241. doi: 10.1111/J.1461-0248.2004.00684.X

Kinmonth-Schultz, H., Lewandowska-Sabat, A., Imaizumi, T., Ward, J. K., Rognli, O. A., and Fjellheim, S. (2021). Flowering Times of Wild Arabidopsis Accessions From Across Norway Correlate With Expression Levels of FT, CO, and FLC Genes. *Front Plant Sci* 12, 747740. doi: 10.3389/FPLS.2021.747740

Kleijn, D., Treier, U. A., and Müller-Schärer, H. (2005). The importance of nitrogen and carbohydrate storage for plant growth of the alpine herb Veratrum album. *New Phytologist* 166, 565–575. doi: 10.1111/J.1469-8137.2005.01321.X

Klimeš, L., Klimešová, J., Hendriks, R. J. J., and Groenendael, J. M. (1997). Clonal plant architecture A comparative analysis of form and function.

Körner, C. (2006). Plant CO2 responses: an issue of definition, time and resource supply. *New Phytologist* 172, 393–411. doi: 10.1111/J.1469-8137.2006.01886.X

Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* 9, 1–9. doi: 10.1186/1746-4811-9-29/FIGURES/4

Lasky, J. R., Josephs, E. B., and Morris, G. P. (2023). Genotype-environment associations to reveal the molecular basis of environmental adaptation. *Plant Cell* 35, 125–138. doi: 10.1093/PLCELL/KOAC267

Lee, Z., Kim, S., Choi, S. J., Joung, E., Kwon, M., Park, H. J., et al. (2023). Regulation of Flowering Time by Environmental Factors in Plants. *Plants 2023, Vol. 12, Page 3680* 12, 3680. doi: 10.3390/PLANTS12213680

Legendre, P., and Legendre, L. (1998). *Numerical Ecology*., 2nd Edn. Amsterdam: Elsevier Science.

Levin, D. A. (2006). Flowering phenology in relation to adaptive radiation. *Syst Bot* 31, 239–246. doi: 10.1600/036364406777585928

Li, K., Wang, Y., Han, C., Zhang, W., Jia, H., and Li, X. (2007). GA signaling and CO/FT regulatory module mediate salt-induced late flowering in Arabidopsis thaliana. *Plant Growth Regul* 53, 195–206. doi: 10.1007/S10725-007-9218-7

Linder, H. P. (2020). The evolution of flowering phenology: an example from the wind-pollinated African Restionaceae. *Ann Bot* 126, 1141–1153. doi: 10.1093/AOB/MCAA129

Lowry, D. B. (2012). Ecotypes and the controversy over stages in the formation of new species. *Biological Journal of the Linnean Society* 106, 241–257. doi: 10.1111/J.1095-8312.2012.01867.X

Lowry, D. B., Modliszewski, J. L., Wright, K. M., Wu, C. A., and Willis, J. H. (2008a). Review. The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philos Trans R Soc Lond B Biol Sci* 363, 3009–3021. doi: 10.1098/RSTB.2008.0064

Lowry, D. B., Rockwood, R. C., and Willis, J. H. (2008b). Ecological reproductive isolation of coast and inland races of Mimulus guttatus. *Evolution* 62, 2196–2214. doi: 10.1111/J.1558-5646.2008.00457.X

Lynch, J. P. (2022). Edaphic stress interactions: Important yet poorly understood drivers of plant production in future climates. *Field Crops Res* 283, 108547. doi: 10.1016/J.FCR.2022.108547

Ma, R., Xu, S., Chen, Y., Guo, F., and Wu, R. (2020). Allometric relationships between leaf and bulb traits of Fritillaria przewalskii Maxim. grown at different altitudes. *PLoS One* 15. doi: 10.1371/JOURNAL.PONE.0239427

Magbanua, Z. V, Hsu, C.-Y., Pechanova, O., Ii, M. A., Grover, C. E., and Peterson, D. G. (2023). Innovations in double digest restriction-site associated DNA sequencing (ddRAD-Seq) method for more efficient SNP identification. *Anal Biochem* 662, 115001. doi: 10.1016/j.ab.2022.115001

Marcin, M., Raschmanová, N., Miklisová, D., and Kováč, Ľ. (2021). Microclimate and habitat heterogeneity as important drivers of soil Collembola in a karst collapse doline in the temperate zone. *Invertebrate Biology* 140, e12315. doi: 10.1111/IVB.12315

Markoski, M., Mitkova, T., Tanaskovik, V., Spalevic, V., and Zgorelec, Z. (2016). The Influence of the Parent Material on the Texture and Water Retention Curves in the Soil Formed upon Limestones and Dolomites. *The Journal "Agriculture and Forestry"* 62. doi: 10.17707/AgricultForest.62.4.20

Marks, R. A., Hotaling, S., Frandsen, P. B., and VanBuren, R. (2021). Representation and participation across 20 years of plant genome sequencing. *Nature Plants 2021 7:12* 7, 1571–1578. doi: 10.1038/s41477-021-01031-8

Martinić, I. (2000). *An Overview of the State of Biological and Landscape Diversity of Croatia with the Protection Strategy and Action Plans*. Zagreb: Ministry of Environmental Protection and Physical Planning, Nature Protection Division.

McKay, J. K., Bishop, J. G., Lin, J. Z., Richards, J. H., Sala, A., and Mitchell-Olds, T. (2001). Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rockcress. *Proc Biol Sci* 268, 1715–1721. doi: 10.1098/RSPB.2001.1715

McKown, A. D., Klápště, J., Guy, R. D., Geraldes, A., Porth, I., Hannemann, J., et al. (2014). Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of Populus trichocarpa. *New Phytologist* 203, 535–553. doi: 10.1111/NPH.12815

McVean, G. (2009). A Genealogical Interpretation of Principal Components Analysis. *PLoS Genet* 5, e1000686. doi: 10.1371/JOURNAL.PGEN.1000686

Merlo, E., Mendoza-Fernández, A. J., Salmerón-Sánchez, E., Martínez-Hernández, F., Ortiz-úbeda, A., and Mota, J. (2021). Elementome of endemic dolomitic flora: Pterocephalus spathulatus (lag.) coult. *Land (Basel)* 10, 1253. doi: 10.3390/LAND10111253/S1

Mertens, D., Boege, K., Kessler, A., Koricheva, J., Thaler, J. S., Whiteman, N. K., et al. (2021). Predictability of biotic stress structures plant defence evolution. *Trends Ecol Evol* 36, S0169-5347(20)30372–4. doi: 10.1016/J.TREE.2020.12.009

Mihevc, A. (2010). "Geomorphology," in *Introduction to the Dinaric Karst*, (Postojna, Slovenia: Karst Research Institute at ZRC SAZU), 30–43.

Mihevc, A., Prelovšek, M., and Zupan Hajna, N. (2010). *Introduction to the Dinaric Karst*. Postojna: Karst Research Institute at ZRC SAZU. doi: 10.3986/9789612541989

Molla, K. A. (2022). Flowering time and photoperiod sensitivity in rice: Key players and their interactions identified. *Plant Cell* 34, 3489–3490. doi: 10.1093/PLCELL/KOAC230

Mota, J., Merlo, E., Martínez-Hernández, F., Mendoza-Fernández, A. J., Pérez-García, F. J., and Salmerón-Sánchez, E. (2021). Plants on Rich-Magnesium Dolomite Barrens: A Global Phenomenon. *Biology 2021, Vol. 10, Page 38* 10, 38. doi: 10.3390/BIOLOGY10010038

Nei, M. (1972). Genetic Distance between Populations. *Am Nat* 106, 283–292. doi: 10.1086/282771

Nikolić, T., and Topić, J. eds. (2005). *Crvena knjiga vaskularne flore Hrvatske*. Zagreb: Ministarstvo kulture, Državni zavod za zaštitu, Republika Hrvatska.

Ohashi, K., and Yahara, T. (2009). Behavioural responses of pollinators to variation in floral display size and their influences on the evolution of floral traits. *Cognitive Ecology of Pollination*, 274–296. doi: 10.1017/CBO9780511542268.015

Onifade, M., Roy-Gagnon, M. H., Parent, M. É., and Burkett, K. M. (2022). Comparison of mixed model based approaches for correcting for population substructure with application to extreme phenotype sampling. *BMC Genomics* 23, 1–12. doi: 10.1186/S12864-022-08297-Y

Orive, M. E. (2020). "The Evolution of Sex," in *The Theory of Evolution*, ed. S. M. and M. D. P. Scheiner (University of Chicago Press), 273–295.

Pandža, M., Franjić, J., and Škvorc, Ž. (2007). The salt marsh vegetation on the East Adriatic coast. *Biologia (Bratisl)* 62, 24–31. doi: 10.2478/S11756-007-0003-X

Parker, T. H., Gerber, A., Campbell, E., Simonson, M., Shriver, R. K., and Persico, L. (2024). Solar Radiation Drives Potential Demographic Collapse in a Perennial Bunchgrass via Dramatically Reduced Seedling Establishment. *Rangel Ecol Manag* 92, 100–112. doi: 10.1016/J.RAMA.2023.10.006

Paschoa, R. P. da, Pinto, V. B., Pereira, J. P., Cavatte, P. C., Garbin, M. L., Godinho, T. de O., et al. (2023). Proteomic and Physiological Signatures of Altitude Adaptation in a Myrsine coriacea Population under Common Garden Conditions. *bioRxiv*, 2023.12.01.569649. doi: 10.1101/2023.12.01.569649

Pau, S., Wolkovich, E. M., Cook, B. I., Davies, T. J., Kraft, N. J. B., Bolmgren, K., et al. (2011). Predicting phenology by integrating ecology, evolution and climate science. *Glob Chang Biol* 17, 3633–3643. doi: 10.1111/J.1365-2486.2011.02515.X

Pedersen, O., Perata, P., and Voesenek, L. A. C. J. (2017). Flooding and low oxygen responses in plants. *Functional Plant Biology* 44, iii–vi. doi: 10.1071/FPV44N9_FO

Peng, J. C., He, Z., and Zhang, Z. Q. (2025). Standing genetic variation and introgression shape the cryptic radiation of Aquilegia in the mountains of Southwest China. *Communications Biology 2025 8:1* 8, 1–15. doi: 10.1038/s42003-025-08120-w

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., and Hoekstra, H. E. (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PLoS One* 7, e37135. doi: 10.1371/JOURNAL.PONE.0037135

Petkovšek, V., and Seliškar, A. (1978). Travniška morska čebulica, novost v slovenski flori. *Proteus* 40, 160–161.

Phylactou, P. (2023). P-Values or Bayes Factors? Whichever, but not Both. Available at: https://osf.io/dthns_v1 (Accessed May 26, 2025).

Pignatti, E., and Pignatti, S. (2014). Plant life of the dolomites: Vegetation structure and ecology. *Plant Life of the Dolomites: Vegetation Structure and Ecology*, 1–527. doi: 10.1007/978-3-642-31043-0/COVER

Pleić, N., Babić Leko, M., Gunjača, I., Boutin, T., Torlak, V., Matana, A., et al. (2022). Genome-Wide Association Analysis and Genomic Prediction of Thyroglobulin Plasma Levels. *Int J Mol Sci* 23, 2173. doi: 10.3390/IJMS23042173/S1

Price, E. A. C., and Marshall, C. (1999). Clonal plants and environmental heterogeneity: An introduction to the proceedings. *Plant Ecol* 141, 3–7. doi: 10.1023/A:1009838300691

Pritchard, J. K., and Di Rienzo, A. (2010). Adaptation - not by sweeps alone. *Nat Rev Genet* 11, 665–667. doi: 10.1038/NRG2880

Prohic, E., Peh, Z., and Miko, S. (1998). Geochemical characterization of a karst polje - An example from Sinjsko polje, Croatia. *Environmental Geology* 33, 263–273. doi: 10.1007/S002540050245

Quaresma, A., Ankenbrand, M. J., Garcia, C. A. Y., Rufino, J., Honrado, M., Amaral, J., et al. (2024). Semi-automated sequence curation for reliable reference datasets in ITS2 vascular plant DNA (meta-)barcoding. *Scientific Data 2024 11:1* 11, 1–11. doi: 10.1038/s41597-024-02962-5

Rahbek, C., Borregaard, M. K., Antonelli, A., Colwell, R. K., Holt, B. G., Nogues-Bravo, D., et al. (2019). Building mountain biodiversity: Geological and evolutionary processes. *Science (1979)* 365, 1114–1119. doi: 10.1126/SCIENCE.AAX0151

Rathcke, B., and Lacey, E. P. (1985). Phenological patterns of terrestrial plants. *Annual review of ecology and systematics. Vol. 16* 16, 179–214. doi: 10.1146/ANNUREV.ES.16.110185.001143

Ray, S., and Satya, P. (2014). Next generation sequencing technologies for next generation plant breeding. *Front Plant Sci* 5. doi: 10.3389/FPLS.2014.00367

Ritter-Studnička, H. (1954). Flora i vegetacija livada kraških polja Bosne i Hercegovine. *Godišnjak Biološkog instituta u Sarajevu* 7, 25–110.

Ross-Ibarra, J., Morrell, P. L., and Gaut, B. S. (2007). Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proc Natl Acad Sci U S A* 104, 8641–8648. doi: 10.1073/PNAS.0700643104

Runcie, D. E., and Crawford, L. (2019). Fast and flexible linear mixed models for genome-wide genetics. *PLoS Genet* 15, e1007978. doi: 10.1371/JOURNAL.PGEN.1007978

Rundle, H. D., and Nosil, P. (2005). Ecological speciation. *Ecol Lett* 8, 336–352. doi: 10.1111/J.1461-0248.2004.00715.X

Sandring, S., and Ågren, J. (2009). Pollinator-mediated selection on floral display and flowering time in the perennial herb Arabidopsis lyrata. *Evolution* 63, 1292–1300. doi: 10.1111/J.1558-5646.2009.00624.X

Savolainen, O., Lascoux, M., and Merilä, J. (2013). Ecological genomics of local adaptation. *Nat Rev Genet* 14, 807–820. doi: 10.1038/NRG3522

Scarano, C., Veneruso, I., De Simone, R. R., Di Bonito, G., Secondino, A., and D'Argenio, V. (2024). The Third-Generation Sequencing Challenge: Novel Insights for the Omic Sciences. *Biomolecules 2024, Vol. 14, Page 568* 14, 568. doi: 10.3390/BIOM14050568

Scheinfeldt, L. B., and Tishkoff, S. A. (2013). Recent human adaptation: genomic approaches, interpretation and insights. *Nature Reviews Genetics 2013 14:10* 14, 692–702. doi: 10.1038/nrg3604

Schley, R. J., Pellicer, J., Ge, X.-J., Barrett, C., Bellot, S., Guignard, M. S., et al. (2021). The Ecology of Palm Genomes: Repeat-associated genome size expansion is constrained by aridity. *bioRxiv*, 2021.11.04.467295. doi: 10.1101/2021.11.04.467295

Schönherr, S., Schachtl-Riess, J. F., Di Maio, S., Filosi, M., Mark, M., Lamina, C., et al. (2024). Performing highly parallelized and reproducible GWAS analysis on biobank-scale data. *NAR Genom Bioinform* 6, 15. doi: 10.1093/NARGAB/LQAE015

Schwartz, M. D. (2024). *Phenology: An Integrative Environmental Science*., 3rd Edn, ed. M. D. Schwartz. Milwaukee: Springer Nature Switzerland AG 2024. doi: 10.1007/978-94-007-0632-3

Šilić, Č. (1990). Morfologija, horologija, ekologija i fenologija dviju grupa populacija Scilla litardierei Breistr. (Syn.: S. pratensis Waldst. & Kit. non Bergeret). *Bilt. Društva Ekol. BiH* Ser. B, 107–116.

Siljak-Yakovlev, S. ; R. O. ; P. D. ; Š. Č. (2010). Cytogenetic characterization of two Chouardia species: Ch. lakusicii (Šilić) Speta and Ch. litardierei (Breistr.) Speta., in *Proceedings of the XIII Optima Meeting.*, ed. T. Ekim (Antalya: Flora Research Society.), 169. – 169.

Silvertown, J. (2008). The evolutionary maintenance of sexual reproduction: Evidence from the ecological distribution of asexual reproduction in clonal plants. *Int J Plant Sci* 169, 157–168. doi: 10.1086/523357/0

Silvestri, S., Defina, A., and Marani, M. (2005). Tidal regime, salinity and salt marsh plant zonation. *Estuar Coast Shelf Sci* 62, 119–130. doi: 10.1016/J.ECSS.2004.08.010

Song, X., Liu, L., Wang, C., Zhang, W., Li, Y., Zhu, J., et al. (2023). Real-time determination of flowering period for field wheat based on improved YOLOv5s model. *Front Plant Sci* 13, 1025663. doi: 10.3389/FPLS.2022.1025663

Stuefer, J. F., Erschbamer, B., Huber, H., and Suzuki, J.-I. (2002). The ecology and evolutionary biology of clonal plants: an introduction to the proceedings of Clone-2000. *Ecology and Evolutionary Biology of Clonal Plants*, 1–8. doi: 10.1007/978-94-017-1345-0_1

Suetsugu, K., Naito, R. S., Fukushima, S., Kawakita, A., and Kato, M. (2015). Pollination system and the effect of inflorescence size on fruit set in the deceptive orchid Cephalanthera falcata. *J Plant Res* 128, 585–594. doi: 10.1007/S10265-015-0716-9

Sun, S., Dong, B., and Zou, Q. (2021). Revisiting genome-wide association studies from statistical modelling to machine learning. *Brief Bioinform* 22, 1–10. doi: 10.1093/BIB/BBAA263

Ter Braak, C. J. F. (1987). *CANOCO-FORTRAN Program for Canonical Community Ordination by (Partial) Detrended (Canonical) Correspondence Analysis and Redundancy Analysis*. Wageningen: TNO Institute of Applied Computer Science, Statistical Department.

Theissinger, K., Fernandes, C., Formenti, G., Bista, I., Berg, P. R., Bleidorn, C., et al. (2023). How genomics can help biodiversity conservation. *Trends in Genetics* 39, 545–559. doi: 10.1016/J.TIG.2023.01.005

Thomas, J., Frost, R. R., and Harvey, R. D. (1973). Thermal conductivity of carbonate rocks. *Eng Geol* 7, 3–12. doi: 10.1016/0013-7952(73)90003-3

Tittes, S., and Kane, N. C. (2014). The genomics of adaptation, divergence and speciation: a congealing theory. *Mol Ecol* 23, 3938–3940. doi: 10.1111/MEC.12855

Todesco, M., Owens, G. L., Bercovich, N., Légaré, J. S., Soudi, S., Burge, D. O., et al. (2020). Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature 2020 584:7822* 584, 602–607. doi: 10.1038/s41586-020-2467-6

Toräng, P., Wunder, J., Obeso, J. R., Herzog, M., Coupland, G., and Ågren, J. (2015). Large-scale adaptive differentiation in the alpine perennial herb Arabis alpina. *New Phytol* 206, 459–470. doi: 10.1111/NPH.13176

Turesson, G. (1922). The genotypical response of the plant species to the habitat. *Hereditas* 3, 211–350. doi: 10.1111/J.1601-5223.1922.TB02734.X

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., et al. (2021). Genome-wide association studies. *Nature Reviews Methods Primers 2021 1:1* 1, 1–21. doi: 10.1038/s43586-021-00056-9

Vallejo-Marín, M., Dorken, M. E., and Barrett, S. C. H. (2010). The ecological and evolutionary consequences of clonality for plant mating. *Annu Rev Ecol Evol Syst* 41, 193–213. doi: 10.1146/ANNUREV.ECOLSYS.110308.120258

Van Drunen, W. E., and Dorken, M. E. (2012). Trade-offs between clonal and sexual reproduction in Sagittaria latifolia (Alismataceae) scale up to affect the fitness of entire clones. *New Phytologist* 196, 606–616. doi: 10.1111/J.1469-8137.2012.04260.X

Vekemans, X., and Hardy, O. J. (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol* 13, 921–935. doi: 10.1046/J.1365-294X.2004.02076.X

Via, S. (2012). Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367, 451. doi: 10.1098/RSTB.2011.0260

Vicentini, G., Biancucci, M., Mineri, L., Chirivì, D., Giaume, F., Miao, Y., et al. (2023). Environmental control of rice flowering time. *Plant Commun* 4, 100610. doi: 10.1016/J.XPLC.2023.100610

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., et al. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet* 101, 5–22. doi: 10.1016/J.AJHG.2017.06.005

Visser, M. E., Adriaensen, F., Van Balen, J. H., Blondel, J., Dhondt, A. A., Van Dongen, S., et al. (2003). Variable responses to large-scale climate change in European Parus populations. *Proc R Soc Lond B Biol Sci* 270, 367–372. doi: 10.1098/RSPB.2002.2244

Visser, M. E., and Both, C. (2005). Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings of the Royal Society B: Biological Sciences* 272, 2561–2569. doi: 10.1098/RSPB.2005.3356

Walter, G. M., Clark, J., Cristaudo, A., Terranova, D., Nevado, B., Catara, S., et al. (2022). Adaptive divergence generates distinct plastic responses in two closely related Senecio species. *Evolution* 76, 1229. doi: 10.1111/EVO.14478

Wang, H., Wang, H., Ge, Q., and Dai, J. (2020). The Interactive Effects of Chilling, Photoperiod, and Forcing Temperature on Flowering Phenology of Temperate Woody Plants. *Front Plant Sci* 11, 533650. doi: 10.3389/FPLS.2020.00443

Waples, D. W., and Waples, J. S. (2004). A review and evaluation of specific heat capacities of rocks, minerals, and subsurface fluids. Part 1: Minerals and nonporous rocks. *Natural Resources Research* 13, 97–122. doi: 10.1023/B:NARR.0000032647.41046.E7

Wolkovich, E. M., Cook, B. I., and Davies, T. J. (2014). Progress towards an interdisciplinary science of plant phenology: building predictions across space, time and species diversity. *New Phytologist* 201, 1156–1162. doi: 10.1111/NPH.12599

Yeaman, S. (2015). Local Adaptation by Alleles of Small Effect. *Am Nat* 186 Suppl 1, S74–S89. doi: 10.1086/682405

Yoon, S., Baik, B., Park, T., and Nam, D. (2021). Powerful p-value combination methods to detect incomplete association. *Scientific Reports 2021 11:1* 11, 1–11. doi: 10.1038/s41598-021-86465-y

Zhao, Y., Zhuw, H., Lu, Z., Knickmeyer, R. C., and Zou, F. (2019). Structured Genome-Wide Association Studies with Bayesian Hierarchical Variable Selection. *Genetics* 212, 397. doi: 10.1534/GENETICS.119.301906

Zhou, X., Carbonetto, P., and Stephens, M. (2013). Polygenic Modeling with Bayesian Sparse Linear Mixed Models. *PLoS Genet* 9, e1003264. doi: 10.1371/JOURNAL.PGEN.1003264

Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics 2012 44:7* 44, 821–824. doi: 10.1038/ng.2310

Zhou, Z., Feng, H., Ma, G., Ru, J., Wang, H., Feng, J., et al. (2024). Seasonal and vertical patterns of water availability and variability determine plant reproductive phenology. *Ann Bot* 135. doi: 10.1093/AOB/MCAE138

Zmienko, A., Wojciechowski, P., and Figlerowicz, M. (2023). Editorial: Resolving the complexity of plant genomes and transcriptomes with long reads, volume II. *Front Plant Sci* 14, 1326500. doi: 10.3389/FPLS.2023.1326500

Zupan Hajna, N. (2019). Dinaric karst—Geography and geology. *Encyclopedia of Caves, Third Edition*, 353–362. doi: 10.1016/B978-0-12-814124-3.00039-X

7 Appended publications

The publications that support the argumentation of this PhD thesis and present the detailed results are appended in the subsequent pages, in the following order:

Publication I

Radosavljević, I., Križanović, K., Šarančić, S. L., and Jakše, J. (2023). Towards the Investigation of the Adaptive Divergence in a Species of Exceptional Ecological Plasticity: Chromosome-Scale Genome Assembly of *Chouardia litardierei* (Hyacinthaceae). Int J Mol Sci 24, 10755. doi: 10.3390/ijms241310755

Publication II

Šarančić, S. L., Pleić, N., Križanović, K., Surina, B., Mitić, D., and Radosavljević, I. (2025a). Uncovering the genomic basis of phenological traits in *Chouardia litardierei* (Asparagaceae) through a genome-wide association study (GWAS). Front Plant Sci 16, 1571608. doi: 10.3389/FPLS.2025.1571608

Publication III

Šarančić, S. L., Pleić, N., Mitić, D., Križanović, K., Surina, B., and Radosavljević, I. (2025b). Genome-wide association study (GWAS) provides insights into the genomic basis of reproduction-related traits in *Chouardia litardierei* (Asparagaceae). BMC Plant Biology 2025 25:1 25, 1–25. doi: 10.1186/S12870-025-06617-4

7.1 Publication I

Radosavljević, I., Križanović, K., Šarančić, S. L., and Jakše, J. (2023). Towards the Investigation of the Adaptive Divergence in a Species of Exceptional Ecological Plasticity: Chromosome-Scale Genome Assembly of *Chouardia litardierei* (Hyacinthaceae). Int J Mol Sci 24, 10755. doi: 10.3390/ijms241310755

International Journal of
Molecular Sciences

MDPI

Article

Towards the Investigation of the Adaptive Divergence in a Species of Exceptional Ecological Plasticity: Chromosome-Scale Genome Assembly of *Chouardia litardierei* (Hyacinthaceae)

Ivan Radosavljević [1,*], Krešimir Križanović [2], Sara Laura Šarančić [1] and Jernej Jakše [3]

[1] Division of Botany, Department of Biology, Faculty of Science, University of Zagreb, Marulićev trg 9A, HR-10000 Zagreb, Croatia
[2] Department of Electronic Systems and Information Processing, Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, HR-10000 Zagreb, Croatia
[3] Department of Agronomy, Biotechnical Faculty, University of Ljubljana, Jamnikarjeva 101, SI-1000 Ljubljana, Slovenia
* Correspondence: ivanrad@biol.pmf.hr

Abstract: One of the central goals of evolutionary biology is to understand the genomic basis of adaptive divergence. Different aspects of evolutionary processes should be studied through genome-wide approaches, therefore maximizing the investigated genomic space. However, in-depth genome-scale analyses often are restricted to a model or economically important species and their closely related wild congeners with available reference genomes. Here, we present the high-quality chromosome-level genome assembly of *Chouardia litardierei*, a plant species with exceptional ecological plasticity. By combining PacBio and Hi-C sequencing technologies, we generated a 3.7 Gbp genome with a scaffold N50 size of 210 Mbp. Over 80% of the genome comprised repetitive elements, among which the LTR retrotransposons prevailed. Approximately 86% of the 27,257 predicted genes were functionally annotated using public databases. For the comparative analysis of different ecotypes' genomes, the whole-genome sequencing of two individuals, each from a distinct ecotype, was performed. The detected above-average SNP density within coding regions suggests increased adaptive divergence-related mutation rates, therefore confirming the assumed divergence processes within the group. The constructed genome presents an invaluable resource for future research activities oriented toward the investigation of the genetics underlying the adaptive divergence that is likely unfolding among the studied species' ecotypes.

Keywords: *Chouardia litardierei*; PacBio; Hi-C; chromosome-level genome; draft genome; local adaptation

Citation: Radosavljević, I.; Križanović, K.; Šarančić, S.L.; Jakše, J. Towards the Investigation of the Adaptive Divergence in a Species of Exceptional Ecological Plasticity: Chromosome-Scale Genome Assembly of *Chouardia litardierei* (Hyacinthaceae). *Int. J. Mol. Sci.* **2023**, *24*, 10755. https://doi.org/10.3390/ijms241310755

Academic Editors: Hon-Ming Lam, Sachiko Isobe and Man-Wah Li

Received: 24 May 2023
Revised: 25 June 2023
Accepted: 27 June 2023
Published: 28 June 2023

1. Introduction

Amethyst meadow squill (*Chouardia litardierei* (Breist.) Speta) (Figure 1A) is a bulbous perennial species of the Hyacinthaceae family. It grows naturally across the western and central parts of the Dinaric Alps in the Balkan Peninsula, occupying highly contrasting ecological niches [1,2] and therefore meadow, seashore, and mountainous ecotypes can be recognized (Figure 1B).

The meadow ecotype, distributed throughout the central and northern parts of the species distribution area, is found across karst fields at altitudes of up to 1000 m. These flat-floored and periodically flooded enclosed depressions are characterized by a unique microclimate and hydrological and geomorphological conditions compared to the surrounding areas [3]. The seashore ecotype occupies the lowlands of northern Dalmatia across the northwestern part of the species distribution range. These populations grow in salt marshes reaching the seashore, which experience Mediterranean climate conditions [4,5]. Finally, the mountainous ecotype is distributed throughout the southern parts of the species' distribution range and in comparison to the aforementioned two ecotypes, occupies a highly contrasting habitat. Its

populations inhabit arid, rocky slopes of high mountains with very little or virtually no soil in rock crevices at altitudes of up to 2000 m that are characterized by extreme seasonality of most climatic elements. Despite occupying contrasting environments, these groups of populations can hardly be distinguished from each other by any morphological trait. There was an attempt to describe the mountainous ecotype as a separate taxon based on morphological and phenological analyses [2], but the research was based on vague and unreliable approaches, therefore leaving room for justified doubts in the results. *C. litardierei* undoubtedly is a complex species characterized by very pronounced ecological plasticity. However, unlike in some other cases [6], it seems only the specific habitat, and not any morphological trait, can be used for reliable recognition of the ecotypes. We plan to use this species as a study system for a thorough investigation of the genetics underlying the ecological divergence and speciation process.

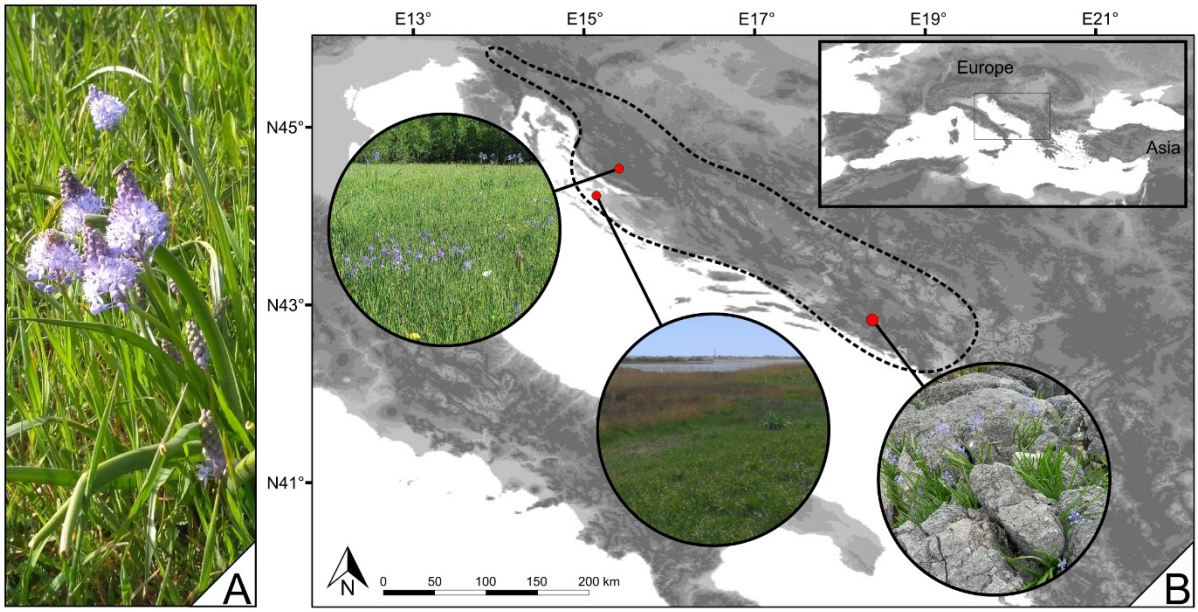


Figure 1. (**A**) *Chouardia litardierei* in full bloom, (**B**) the distribution area of *Chouardia litardierei* and contrasting habitat types it occupies. The distribution area of *Chouardia litardierei* is marked with a dotted line. In circles, from left to right, meadow, seashore, and mountainous ecotype habitats are shown.

To date, no significant research that investigated this species' ecological divergence or genetics has been performed. Besides the previously mentioned analyses by Šilić [2], the cytogenetic characterization of two individuals representing meadow and mountainous ecotypes was also performed [7]. Karyograms revealed that both ecotypes share the same number of chromosomes (2n = 26), with one long, two middle-sized, and ten small chromosome pairs. In addition, the 1C haploid genome size was estimated at 4.13 pg [8] or 4.039 Gbp according to the conversion by Doležel et al. [9].

During the process of speciation, a group of individuals diverges into two or more distinct phylogenetic lineages. In populations initially indistinguishable from each other, either genetically or morphologically, the accumulation of genetic differences can gradually lead to the emergence of a new species [10,11]. The type of speciation in which "barriers to gene flow evolve between populations as a result of ecologically based divergent selection" is referred to as ecological speciation [12]. As a consequence of organism adaptation to specific environmental conditions during ecological speciation, new morphologically and genetically divergent ecotypes found in a specific habitat rather than a specific geographic area, can emerge [13]. One of the central goals of evolutionary biology is to understand the genomic basis of adaptive evolution [14,15]. It is widely accepted that different aspects of evolutionary processes should be studied through genome-wide approaches, therefore maximizing the investigated genomic space. However, genome-scale analyses are often

restricted to a model or economically important species (and their closely related wild congeners) with available high-quality reference genomes [16–18]. In recent years, with the advancement of different NGS techniques and the inevitable increase in their affordability, more non-model species' genomes are being sequenced and assembled de novo [19–21].

Here, we present the high-quality chromosome-scale genome assembly for *C. litardierei*, which is also, to the best of our knowledge, the first reported genome assembly within the Hyacinthaceae family. By implementing PacBio HiFi sequencing and Hi-C scaffolding, a haploid 3.7 Gb genome organized in 13 pseudochromosomes was revealed. The obtained results represent the initial step in comprehensive research that will investigate the process of adaptive divergence and speciation that is likely unfolding among the ecotypes of the studied species. The availability of the species' genome assembly will enable the study of the ecotypes' genome architecture, genome–environment association (GEA), and genome-wide association studies (GWAS), which will elucidate the genomic mechanisms underlying the ongoing evolutionary processes in *C. litardierei*.

2. Results

2.1. Genome Sequencing and Assembly

After sequencing, high-quality PacBio CCS reads were obtained from subreads with a quality score of Q20 (1% error rate). More than 6.5 M PacBio HiFi reads were available with a total of 94.54 Gbp (23× genome coverage, genome size based on the *k*-mer analysis), producing an average read length of 14.5 Kbp. In addition, 861 M Hi-C read-pairs were obtained, resulting in 432 Gbp (105× genome coverage) in total. Based on *k*-mer analysis, the genome size of amethyst meadow squill was estimated at 4.085 Gbp. After processing the hifiasm assembly using Quast, the initial genome assembly of 3.67 Gbp with an average contig N50 of 12.9 Mbp was produced.

After processing the initial assembly and Hi-C data with 3D-DNA, the assembly results were moderately improved and the scaffold N50 measure topped 200 Mbp. The N50 measure obtained after the 3D-DNA pipeline should be considered reliable due to misjoins having been resolved by the pipeline. The rearrangement of scaffolds produced by the 3D-DNA pipeline with the Juicebox tool resulted in the recognition of 13 pseudochromosomes: one very long, two middle-sized, and ten small chromosomes (Figures 2 and 3).

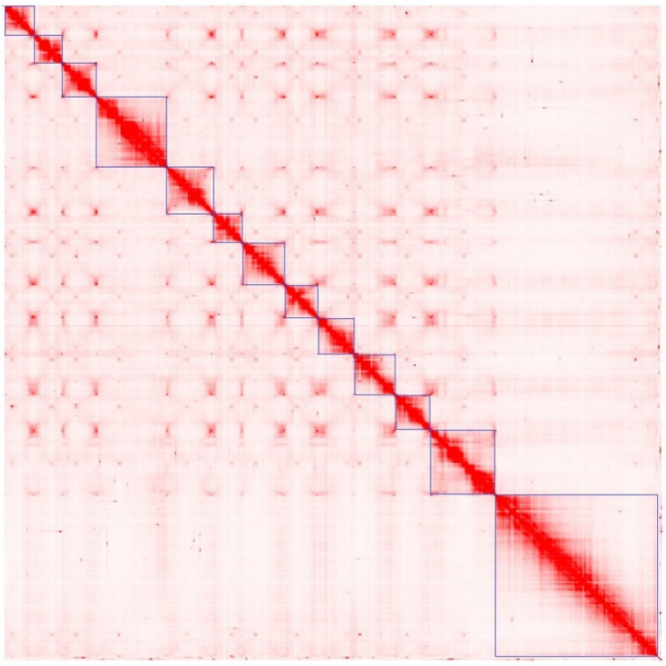


Figure 2. Heatmap showing Hi-C interactions for 13 pseudochromosomes of *Chouardia litardierei*.

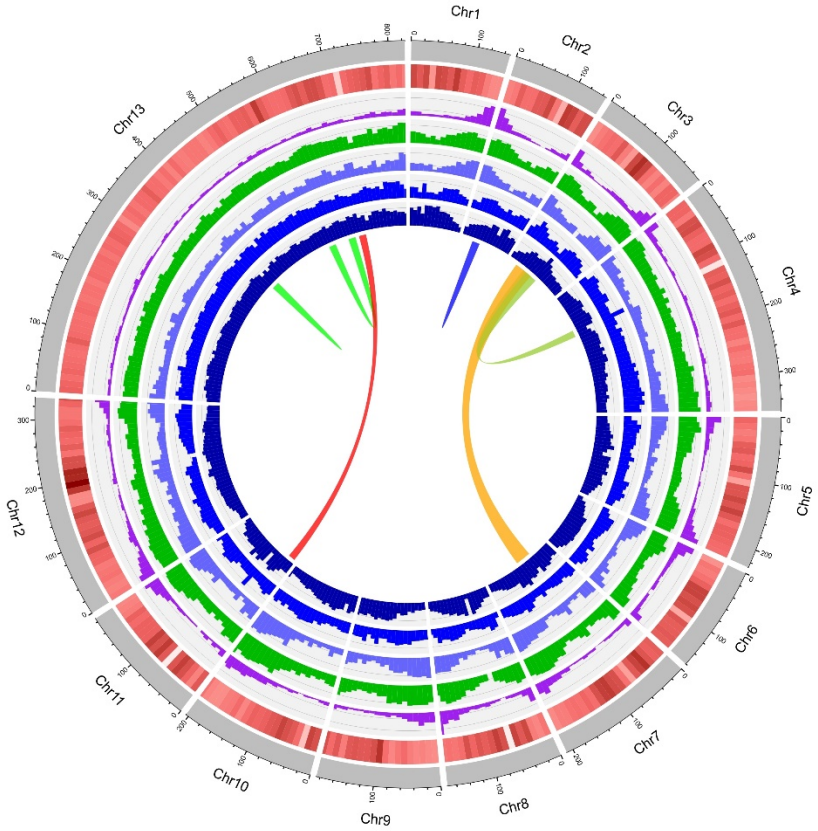


Figure 3. Genome features of 10 Mbp windows across the *Chouardia litardierei* genome. From outer to inner circles: chromosomes, GC content, gene density (purple), total repeats (green), DNA transposons density (light blue), *Copia* elements density (blue), *Gypsy* elements density (dark blue), and intra-genome syntenic blocks where the bandwidth is proportional to the syntenic block size.

The obtained assembly was polished using the HyPo tool, and the results are presented in Table 1. The N50 value reached more than 210 Mbp, and the largest scaffold was nearly 825 Mbp. The 13 largest scaffolds (representing pseudochromosomes) range from 146 Mbp to 825 Mbp, with a total size of 3.33 Gbp. This value represents 90% of the complete assembly and 81.6% of the predicted genome length. The rest of the assembly consists of numerous smaller sequences (2.3 Mbp and smaller) that did not successfully merge with the pseudochromosomes. Finally, the BUSCO completeness score of 97.4% confirmed the high quality of the obtained genome assembly. The summary statistics are presented in Table 1.

Table 1. Summary results for the final assembly of the *Chouardia litardierei* genome.

Sequence	
Assembly size (bp)	3,698,590,323
GC content (%)	42.90
Number of scaffolds	9916
Number of scaffolds (≥50 kbp)	1803
Longest scaffold (bp)	824,692,949
Scaffold N50 size (bp)	210,067,440
Number of contigs	3111
Number of contigs (≥50 kbp)	1611
Longest contig (bp)	54,979,118
Contig N50 size (bp)	12,914,002
Pseudochromosome	

Table 1. *Cont.*

Sequence	
Number	13
Size range (Mbp)	145.64–824.69
BUSCO score	
Complete BUSCOs (%)	97.4
Complete and single-copy BUSCOs (%)	89.9
Complete and duplicated BUSCOs (%)	7.5
Fragmented BUSCOs (%)	2.4
Missing BUSCOs (%)	0.2

2.2. Repetitive Elements Annotation

The annotation of repetitive elements revealed 2.99 Gbp of repetitive sequences representing 80.90% of the *C. litardierei* genome, with transposable elements (TEs) occupying 69.97% of the genome assembly. In addition, the analysis revealed that LTR retrotransposons were by far the most abundant repeat sequences (63.25% of the genome assembly), of which *Copia* and *Gypsy*, two superfamilies, account for 27.03% and 36.01% of the assembled sequences, respectively. Other detected repeat elements were unclassified elements (7.81%), DNA transposons (3.67%), long interspersed nuclear elements (LINEs; 2.99%), and others with lower abundances (Table 2).

Table 2. Classification of the repetitive elements in the *Chouardia litardierei* genome.

	Percent (%)	**Total Length (Mbp)**
Retrotransposons		
LINE	2.99	110.72
SINE	0.06	2.14
LTR	63.25	2339.37
DNA Transposons	3.67	135.60
Unclassified	7.81	288.98
Satellites	0.14	5.10
Simple repeats	1.42	52.63
Low complexity	0.31	11.53
Rolling circles	0.58	21.30
Small RNA	0.70	25.88
Total	80.90	2991.99

2.3. RNA Sequencing

The RNA sequencing yielded a total of 99.59 M raw reads. After trimming, 96.75 M reads with an average length of 135.6 bp were retained. The summary of the RNA sequencing results from different tissues is given in Table 3.

Table 3. RNA sequencing data from different *Chouardia litardierei* tissues.

	Root	**Leaf**	**Flower**	**Developing Fruit**
No. of raw reads	22,769,326	24,504,881	28,584,130	23,731,351
Total nucleotides [Mbp]	3013.5	3360.2	3918.0	3070.6
GC content [%]	47.90	49.18	49.65	51.26
Average length [bp]	123.0	137.1	137.1	129.4
Min-max length [bp]	8–383	8–381	8–381	8–384
No. of reads after trimming	22,079,991	23,884,368	27,738,059	23,053,022
Total nucleotides after trimming [Mbp]	2957.7	3309.0	3855.2	3018.6
Average read length after trimming [bp]	134.0	138.5	139.0	131.0

2.4. Gene Prediction and Annotation

By combining several approaches, we predicted 27,257 gene models, of which 23,297 were mapped to 13 pseudochromosomes, while the remaining 3960 were mapped to smaller scaffolds. Their average length, CDS length, and exon number were 3109.9 bp, 764.1 bp, and 4.2 bp, respectively (Table 4). Among the predicted genes, 23,398 were functionally annotated using the public databases Swiss-Prot, InterPro, NCBI NR, and EggNog (Figure 4A).

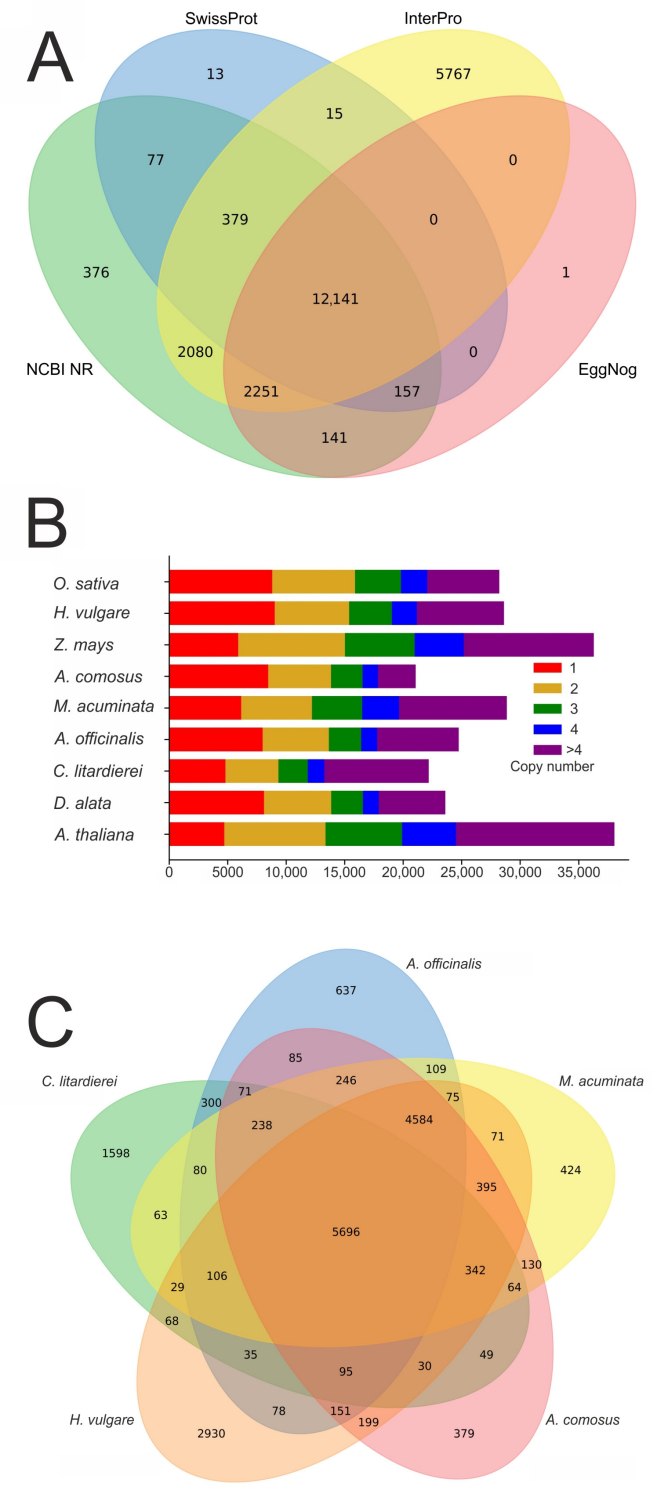


Figure 4. (**A**) Venn diagram showing the number of genes with functional annotation using multiple public databases, (**B**) number of gene copies among nine studied plant species, (**C**) Venn diagram of orthologous groups shared among selected species.

Table 4. Summary of the gene prediction and annotation results of *Chouardia litardierei*.

Gene Prediction	
Number of predicted genes	27,257
Number of predicted genes in 13 pseudochromosomes	23,297
Chr1	1237
Chr2	1152
Chr3	1477
Chr4	2137
Chr5	1757
Chr6	1309
Chr7	1429
Chr8	1513
Chr9	1435
Chr10	1589
Chr11	1344
Chr12	2373
Chr13	4545
Mean gene length (bp)	3109.9
Mean CDS length (bp)	764.1
Mean exon length (bp)	181.0
Mean intron length (bp)	728.0
Avg. exons per gene	4.2
Gene annotation	
NCBI NR annotated (%)	17,602
EggNog annotated (%)	14,691
InterPro annotated (%)	22,633
Swiss-Prot annotated (%)	12,782
Number of annotated genes	23,398
Proportion of annotated genes (%)	85.8%

2.5. Evolution Analysis

To elucidate the evolutionary history of *C. litardierei* within monocots, seven species across the group and one dicot (*A. thaliana* as an outgroup) were selected for the phylogenetic analysis. A total of 24,356 orthologous families of genes were identified: 377 single-copy families, 5189 shared by all studied species, 5486 shared only by monocots representatives, and 6621 shared by *C. litardierei* and *A. officinalis* (Figure 4C). For *C. litardierei* 1458 private gene families were recognized. Single-copy ortho-groups were used for the phylogenetic tree construction. Species formed groups that were in accordance with their already recognized phylogenetic relationships. *C. litardierei* paired with *A. officinalis* within the order Asparagales, while *Z. mays*, *H. vulgare*, and *O. sativa* grouped as representatives of the Poaceae family. As representatives of different families, *D. rotundata, M. acuminata*, and *A. comosus* were positioned separately, as was the case with *A. thaliana* as the sole representative of dicots that served as the outgroup. The divergence time between *C. litardierei* and *A. officinalis* was estimated at 49.9 Mya. The divergence times among the other analyzed species and gene family expansions and contractions are indicated in Figure 5.

2.6. Ecotypes Genomes Comparison

To perform a basic comparison of the different ecotypes' genomes, two additional samples, one representing the meadow, and another the mountainous ecotype, were sequenced. Illumina PE150 sequencing yielded 364 and 370 M reads for the meadow and mountainous ecotype individuals, respectively. However, the usability of such a short-read data set was limited and does not allow detailed comparative analyses of genomes characterized by very high proportions of repetitive elements. Nonetheless, we were able to calculate pairwise distances between the constructed genome assembly and the additional samples based on the total number of detected SNPs (Figure 6) and analyze

their distribution across the genomes (Figure 7). Additionally, the SNP abundances within genes and on the genome level were compared and expressed as the average distance between neighboring SNPs. The results showed that the mountainous ecotype was the most diverged one, while a substantially higher density of SNPs was detected within genes compared to the entire genome.

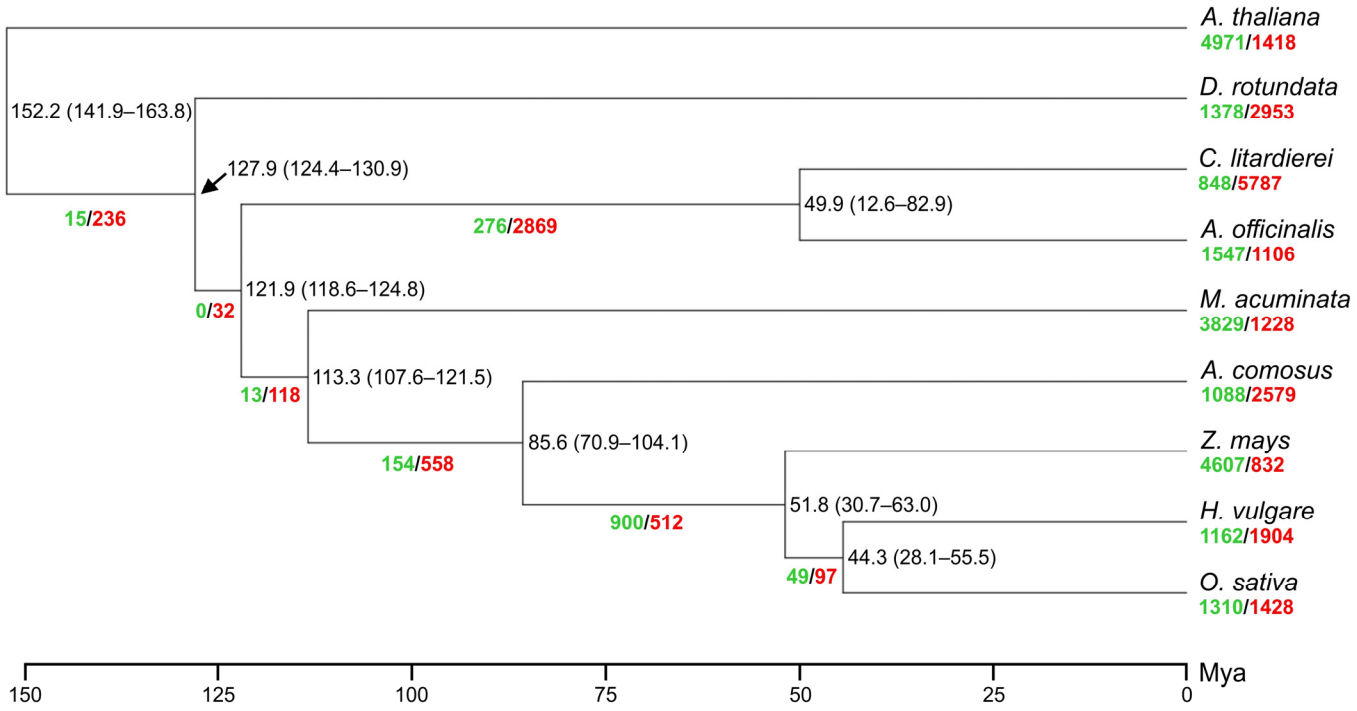


Figure 5. A phylogenetic tree showing topology, divergence times, and expansions/contractions of gene families for nine plant species including *Chouardia litardierei*. Numbers in green, red, and black represent expansions and contractions of gene families, and divergence times, respectively.

A	Assembly	Sample 1	Sample 2
Assembly	0	0.84	1.7
Sample 1	27.1	0	2.1
Sample 2	47.3	65.4	0

B	Assembly	Sample 1	Sample 2
Assembly	0	100.9	49.9
Sample 1	136.5	0	40.4
Sample 2	78.2	56.5	0

Figure 6. Genetic distances and distribution of SNPs among studied *Chouardia litardierei* genomes. Assembly—draft genome assembly of an individual belonging to the seashore ecotype; Sample 1—individual belonging to the meadow ecotype; Sample 2—individual belonging to the mountainous ecotype. (**A**) The total number of SNPs for the given sample pair is shown below the diagonal, and the number of SNPs detected in genes is shown above the diagonal (in millions). (**B**) Mean distance between neighboring SNPs throughout the genome for the given sample pair is shown below the diagonal, and the mean distance between neighboring SNPs within detected genes is shown above the diagonal (in base pairs).

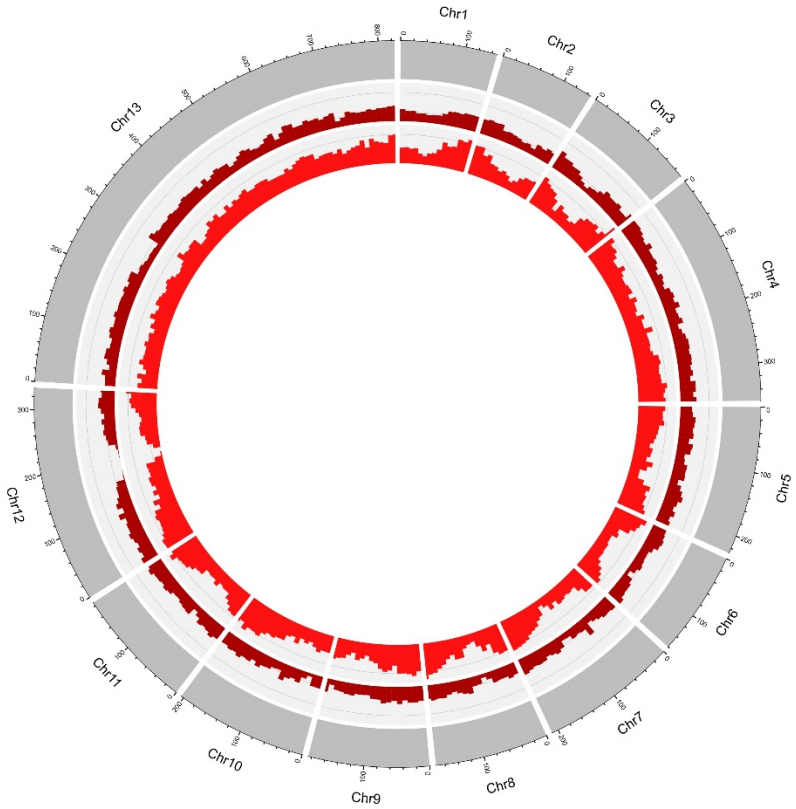


Figure 7. Distribution of meadow and mountainous ecotypes' SNPs in contrast to reference genome assembly. From outer to inner circles: chromosomes, meadow ecotype's SNPs, and mountainous ecotype's SNPs.

3. Discussion

Here, we present a draft genome assembly for *Chouardia litardierei*, a non-model monocot species from the Hyacinthaceae family. By combining long-read sequencing and the chromosome conformation capture method, we successfully assembled a high-quality 3.7 Gbp genome of *C. litardierei*, and the obtained result agrees with the previously reported genome size for the species [8]. By inspecting the Taxonomy Browser of the NCBI repository (https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/tree/?taxon=4447 (accessed on 17 April 2023)), it became obvious that, within monocots, most species with assembled genomes are either of substantial economic importance (maize, wheat, rice, pineapple, banana, asparagus, jams, onion, garlic, etc.) or their wild relatives. In a lower taxonomic rank, within the order Asparagales, assembled genomes of well-known groups of orchids (i.e., *Dendrobium*, *Vanilla*, and *Phalaenopsis*) and *Asparagus* prevail, once again showing bias towards species of economic importance. Of less closely related species to *C. litardierei* within Asparagales that have available genome assemblies, few can be mentioned. The genome assembly of *Asparagus setaceus* was 720 Mb in size and characterized by 1393 scaffolds and a 2.19 Mb N50 scaffold value [22]. The 1.19 Gb *Dendrobium nobile* genome assembly reached a 64.5 Mb N50 scaffold value [23], while the *Cymbidium goeringii* genome, of very similar size to the genome of *C. litardierei* (3.99 vs. 3.70 Gbp, respectively), had an N50 scaffold size of 178.2 Mb [24]. Since we reached the N50 scaffold value of more than 210 Mb, this indicates the high contiguity of the assembled genome. In addition, the BUSCO score of over 97% additionally supported this conclusion. Additionally, a revealed chromosome size distribution perfectly matches the only known karyotype for this species reported by Siljak-Yakovlev et al. [7].

The annotation of repetitive elements revealed that TEs occupy almost 70% of the genome, with LTR retrotransposons being the most abundant class. Such a result was not

surprising, as it is well known that genome size in plants greatly depends on these elements' abundance [25,26]. Our results are mostly consistent with those reported for other monocot species. For instance, the *Hordeum vulgare* ssp. *vulgare* genome (5.1 Gbp in size, Poaceae) consists of 72.8% TE elements [27], the genome of *Areca catechu* (2.6 Gbp, Arecaceae) of 80.4% [28], and that of *Allium fistulosum* (Amaryllidaceae 11.2 Gbp) of 69% [29]. At the same time, genomes of some other monocots, such as *Setaria italica* (423 Mbp, Poaceae) [30], *Trichopus zeylanicus* (713 Mbp, Dioscoreaceae) [31], and *Kobresia myosuroides* (400 Mbp, Cyperaceae) [32] reportedly harbor substantially fewer transposable elements, occupying 41%, 36%, and 44.9% of their genomes, respectively. As mentioned, since the abundance of TEs strongly influences the genome size, species characterized by smaller genomes usually have fewer TEs as well.

To reach high accuracy for the genome annotation, we implemented various approaches to annotate protein-coding genes. Out of the 27,257 predicted genes, most of them (85.8%) were matched with a functional annotation in at least one public database, while almost half of them (44.5%) were matched in all selected databases.

The genus *Prospero* represents a closely related group to *C. litardierei*. It formerly belonged to *Scilla*, and the same is true for the *Chouardia* studied here. *Prospero*, especially the *P. autumnale* s.l. group, is well known for its structural genome rearrangements and multiple ploidy levels and was used as the model group for research on the evolutionary implications of karyotype differentiation [33,34]. In addition, Siljak-Yakovlev et al. [7] hypothesized that the genome of *C. litardierei* could have originated through whole-genome duplication events. To verify whether the *C. litardierei* genome shares some characteristics with *P. autumnale* s.l., or has indeed originated through a whole-genome duplication event, we performed intra-genome syntenic gene block analysis. However, no clues supporting any of these assumptions were found, as it became clear that the *C. litardierei* genome did not undergo any such structural rearrangements since only a few gene blocks co-occurred on more than one position across the genome. In contrast to the limited distribution area of *C. litardierei*, *P. autumnale* s.l. stretches across the Mediterranean basin, so we can assume that the vast distances and subsequent geographical isolation eventually led to the establishment of groups of populations characterized by specific cytotypes.

The evolutionary analysis confirmed the positioning of *C. litardierei* and the entire Hyacinthaceae family within Asparagales. At the same time, it confirmed that the genus *Asparagus*, the closest relative to *C. litardierei* with the available draft genome, can hardly be treated as a close relative since the divergence time was estimated at around 50 Mya. This result further emphasizes the importance of our work, as *C. litardierei* is an obvious representative of, so far, a neglected phylogenetic group in terms of available genomic resources. Regarding other phylogenetic relationships and divergence times among the analyzed representatives of various monocot groups, our results were in high agreement with other similar studies [32,35,36].

Comparative analyses of the assembled genome and two individuals belonging to different ecotypes were of limited success. A shotgun-sequencing approach with a 150 bp read length greatly limited our abilities for in-depth analyses. Nonetheless, we were able to extract SNPs and analyze their distribution across the genomes. The results supported our initial assumption that a higher degree of relatedness is present between the seashore and meadow ecotypes, while the mountainous ecotype is more diverged and possibly represents a separate lineage. In addition, the analysis of SNP distribution within and outside protein-coding regions indicated an above-average density of variations within the coding regions. This result shows that some regions are evolving at a higher pace than others, possibly as a consequence of yet undetermined selective pressures. However, such a conclusion based on only three individuals is likely premature, as research that would include a substantially larger sample set is required for more reliable conclusions. The reasoning behind performing this analysis was to determine if there are any indications of ongoing divergence processes among the lineages, which in the end, we successfully identified.

4. Materials and Methods

4.1. Sample Collection, DNA Extraction, and Sequencing

Fresh leaf material from an individual belonging to the seashore ecotype of the studied species was collected and immediately placed in a silica gel for rapid desiccation. High-molecular-weight DNA extraction following the CTAB method [37], DNA quality control, PacBio HiFi, and Hi-C library preparation and sequencing were performed by Brigham Young University DNA Sequencing Center (Provo, UT, USA). In short, PacBio circular consensus sequencing (CCS) libraries were constructed and sequenced on the 8M SMRT cell of the PacBio Sequel II instrument (Pacific Biosciences of California, Menlo Park, CA, USA), while Hi-C libraries were constructed using a Dovetail® Omni-C® Kit and sequenced on an Illumina HiSeq platform (Illumina Inc., San Diego, CA, USA) to generate 2 × 250 paired-ends reads.

4.2. Genome Assembly

Before the assembly process, the genome size of *C. litardierei* was estimated using a *k*-mer counting method and the tool Jellyfish 2.3.0 [38]. PacBio HiFi reads were processed by Jellyfish to determine their *k*-mer distribution, and the *k*-mer size of 19 was selected. The genome size was estimated as the total number of counted *k*-mers divided by the highest frequency of *k*-mers that occurred. PacBio HiFi reads were assembled into contigs using hifiasm 0.16.1-r375 [39]. Racon 1.4.17 [40] was used in an attempt to improve read quality before the assembly process. The contigs obtained by hifiasm were polished using two rounds of consensus correction with Racon and PacBio HiFi reads.

The generated contigs were scaffolded into pseudochromosomes using Hi-C data. Hi-C reads were first processed following the Omni-C data analysis and quality control protocol, recording valid ligation events and removing PCR duplicates. After initial processing, the Hi-C reads were mapped to contigs using the Juicer tool [41], producing contact map information. To detect misjoins in contigs and to join contigs located on the same chromosomes, 3D-DNA v180922 [42] was used. For the manual rearrangement of obtained scaffolds into pseudochromosomes, we used the Juicebox tool [43]. The same software was also used to generate a FASTA file with sequences corresponding to 13 manually assembled chromosomes, with Ns filling the gaps between scaffolds within each chromosome. This final assembly was further polished with PacBio HiFi reads using the HyPo polisher [44]. HiFi reads were mapped to the final assembly using the minimap2 tool 2.23 [45] with the option "-x map-hifi".

The initial and the final assemblies' quality was assessed using Quast [46] and BUSCO 5.2.2 [47] to compare the assembly to the gene content of Viridiplantae_odb10 "https://busco-archive.ezlab.org/frame_plants.html (accessed on 7 December 2022)". For the genome assembly visualization, we used shinyCircos [48]. The GC content of the assembled genome was calculated using an in-house script. The density of total repeats, DNA transposons, *Copia* repeats, and *Gypsy* repeats was determined from the data obtained through the repetitive element annotation, as explained in the next subsection. Intra-genomic syntenic analysis was performed using SyMAP 5.4.0 [49] with the default parameters.

4.3. Repetitive Elements Annotations

First, the known repeat sequences of Viridiplantae were identified based on Dfam [50] hidden Markov Model (HMM) sequence profiles (release 3.6) using RepeatMasker 4.1.2-p1 [51] and the NCBI/RMBLAST search engine. Furthermore, the de novo repeat identification approach was implemented using RepeatModeler2 2.0.2 [52] with Tandem Repeats Finder 4.10 [53], RECON 1.0.8 [54], and RepeatScout 1.0.6 [55] which enabled LTR Structural analysis. RepeatClassifier (a module of RepeatModeler2) was implemented for further classification of de novo repeats into unknown and classified classes. All three groups of repeats were used in a combined masking step to construct the finally masked version of the genome.

The final BUSCO analysis against Viridiplantae_odb10 was performed on this version of the masked genome.

4.4. RNA Isolation and Sequencing

For support of the gene prediction, RNA-Seq data were generated. Total RNA was extracted from roots, leaves, flowers, and unripe fruit using a Monarch® Total RNA Miniprep Kit (New England BioLabs, Ipswich, MA, USA). The manufacturer's protocol, with an on-column DNAse digestion step, was followed. Eluted RNA was quantified utilizing spectrometry, and integrity was verified by Agilent Bioanalyzer 2100 electrophoresis using an RNA 6000 Nano Kit (Agilent Technologies, Santa Clara, CO, USA). RNA was stored at −80 °C until processing.

RNA sequencing was performed using the Ion Proton system. Total RNA was enriched for the poly-A mRNA fraction using a Dynabeads® mRNA DIRECT™ Micro Kit (Thermo Fisher Scientific, Waltham, MA, USA). The isolated mRNAs were used for RNA-Seq library preparation using the procedure for low-input RNA from the Ion Total RNA-Seq kit v2 (Thermo Fisher Scientific, Waltham, MA, USA). The RNA was fragmented using RNase III enzymatic digestion followed by ligation of Ion Adapters using four different barcodes to retain tissue specificity. The samples were reverse transcribed, purified, and cDNA amplified, and the obtained library was verified using the High Sensitivity DNA Kit (Agilent Technologies, Santa Clara, CO, USA). The libraries, in equimolar amounts, were pooled together and amplified by emulsion PCR using an Ion OneTouch™ 2 System and Ion PI Hi-Q OT2 200 Kit. Template-positive particles were enriched using Dynabeads® MyOne™ Streptavidin C1 beads (Thermo Fisher Scientific, Waltham, MA, USA) on an Ion OneTouch™ ES system. The obtained enriched particle samples were sequenced on PI™ Chip v3 using the Ion PI™ Hi-Q™ Sequencing 200 Kit (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer's protocol. The quality check of trimmed reads after processing was performed by the FastQC tool [56].

4.5. Gene Prediction and Annotation

To predict protein-coding sequences, we used several approaches implemented using different tools. First, gene models were developed with the MAKER genome annotation pipeline (MPI 3.01.04) [57] incorporating: (1) RNA-seq data, (2) protein-based evidence based on 139,388 Asparagales clade proteins downloaded from the NCBI RefSeq database "https://www.ncbi.nlm.nih.gov/refseq/ (accessed on 9 January 2023)", and (3) ab initio gene predictions obtained using SNAP 2006-07-28 [58] and Augustus 3.2.3 [59]. For SNAP software training, MAKER models with a max AED threshold of 0.25 and a minimum length of 50 amino acids were used, and for training Augustus, the BUSCO pipeline was employed following the method of Card et al. [60]. Three runs of MAKER were run iteratively to obtain most gene models with an AED score above 0.5.

Additional ab initio gene prediction was obtained using GeneMark-ES [61], followed by de novo and genome-guided transcriptome assembling using the Trinity 2.14.0 software [62] (default parameters). For the construction of the genome-guided transcriptome, the GMAP tool [63], and SAMtools 1.14 [64] were used to map the reads to the previously constructed genome assembly and to obtain a coordinate sorted bam file, respectively. The transcriptomes obtained by Trinity were used as inputs for the PASA alignment assembly pipeline 2.5.2 [65] (default parameters). The obtained transcriptome was further used to identify and extract likely coding regions using PASA's Transdecoder software. For homology-based gene prediction, the Asparagales protein set was used again. The proteins were mapped to the previously constructed genome using the miniprot tool [66].

Finally, the MAKER gene annotations together with the PASA transcriptome, PASA likely coding regions, protein alignments obtained by miniprot, and ab initio predictions obtained by GeneMark-ES, were analyzed using EVidenceModeler 2.0.0 [67], producing the final consensus gene set.

Recognized protein-coding genes were functionally annotated based on entries in the NCBI NR database [68], Swiss-Prot [69], InterPro [70], and EggNOG [71] databases, using BLASTP searches with an E-value cut-off of 1.0×10^{-5}. For the visualization of the obtained results, a Venn diagram was constructed.

4.6. Genome Evolution Analysis

Orthologous groups were identified using OrthoFinder 2.5.4 [72] and protein sequences from *Ananas comosus* (L.) Merr., *Arabidopsis thaliana* (L.) Heynh., *Asparagus officinalis* L., *Dioscorea rotundata* Poir., *Hordeum vulgare* L., *Musa acuminata* L., *Oryza sativa* L., and *Zea mays* L. Single-copy ortho-groups were collected and aligned using MUSCLE 3.8.1551 [73]. The alignments were concatenated into a super-alignment and filtered using Gblocks 0.91.1 [74]. The phylogenetic trees were constructed using RaxML-NG 0.9.0 [75].

Divergence time estimation was performed using the MCMCTree tool in the PAML 4.9j package [76]. Analyses were run using default settings (200,000 generations with a burn-in of 2000 iterations). The calibration points for the *O. sativa–H. vulgare* (42–62 Mya), *A. comosus–M. acuminata* (103–117 Mya), and *D. rotundata–A. thaliana* (142–164 Mya) were obtained from the TimeTree database [77] "http://www.timetree.org (accessed on 6 April 2023)". Finally, for the identification of gene families' expansions and contractions, CAFE5 [78] was implemented.

4.7. Intra-Species Comparison of the Genomes

In addition, to perform a basic comparative analysis of genomes from different ecotypes, two individuals, each from a distinct ecotype (meadow and mountainous ecotypes, Samples 1 and 2, respectively), were sampled. DNA was extracted from dried leaf material using the GenElute™ Plant Genomic DNA Miniprep Kit (Sigma–Aldrich, St. Louis, MO, USA) and sent to Novogene (UK) Company Limited for short-fragment libraries preparation and PE150 sequencing on an Illumina NovaSeq platform (Illumina Inc., San Diego, CA, USA). The paired-end reads were mapped to the constructed genome assembly using the BWA tool 0.7.17 [79], and the variants were called using the FreeBayes tool [80,81]. The obtained data were used to assess the pairwise genetic distances between analyzed individuals belonging to different ecotypes. In addition, the abundance of the SNPs within protein-coding regions was analyzed using an in-house script.

Author Contributions: I.R. conceived the idea for the research, collected all the samples, and wrote the manuscript. K.K. and J.J. conducted the statistical analyses. S.L.Š. assisted in figure preparation and writing the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Croatian Science Foundation. All the research activities related to the sampling and the preparation of analyzed samples and their sequencing were financed by the "Amethyst Meadow Squill (*Chouardia litardierei*, Hyacinthaceae): a study system for ecological divergence" (HRZZ-IP-2020-02-8099) project. Computational capacities needed for statistical analyses were provided through the project "Single-genome and metagenome de novo assembly" (HRZZ-IP-2018-01-5886).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the obtained data (PacBio, Hi-C, RNA-Seq, and WGS reads), as well as the final genome assembly and predicted gene models, are available in the NCBI database under the BioProject ID PRJNA974736.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gaži-Baskova, V. Geografsko raširenje livadnog procjepka ili lučike (*Scilla pratensis* W. et K.). *Biološki Glas.* **1962**, *15*, 49–54. (In Bosnian)
2. Šilić, Č. Morfologija, horologija, ekologija i fenologija dviju grupa populacija *Scilla litardierei* Breistr. (Syn.: *S. pratensis* Waldst. & Kit. non Bergeret). *Bilt. Društva Ekol. BiH Ser. B* **1990**, *5*, 107–116. (In Bosnian)
3. Mihevc, A. Geomorphology. In *Introduction to the Dinaric Karst*; Mihevc, A., Prelovšek, M., Zupan Hajna, N., Eds.; Karst Research Institute at ZRC SAZU: Postojna, Slovenia, 2010; pp. 30–43.
4. Horvatić, S. Vegetacijska karta otoka Paga s općim pregledom vegetacijskih jedinica Hrvatskog Primorja. *Prirodosl. Istraživanja* **1963**, *33*, 1–187. (In Croatian)
5. Dítě, D.; Dítě, Z.; Eliáš jun, P.; Šuvada, R. Rare plant species of salt marshes of the Croatian coast. *Hacquetia* **2018**, *17*, 221–234. [CrossRef]
6. Brandrud, M.K.; Paun, O.; Lorenzo, M.T.; Nordal, I.; Brysting, A.K. RADseq provides evidence for parallel ecotypic divergence in the autotetraploid Cochlearia officinalis in Northern Norway. *Sci. Rep.* **2017**, *7*, 5573. [CrossRef] [PubMed]
7. Siljak-Yakovlev, S.; Robin, O.; Papeš, D.; Šilić, Č. Cytogenetic characterization of two Chouardia species: *Ch. lakusicii* (Šilić) Speta and *Ch. litardierei* (Breistr.) Speta. In *XIII Optima Meeting Book of Abstract, Proceedings of the XIII Optima Meeting, Antalya, Turkey, 22–26 March 2010*; Ekim, T., Ed.; Flora Research Society: Grants Pass, OR, USA, 2010; p. 169.
8. Siljak-Yakovlev, S.; Pustahija, F.; Šolić, E.M.; Bogunić, F.; Muratović, E.; Bašić, N.; Catrice, O.; Brown, S.C. Towards a Genome Size and Chromosome Number Database of Balkan Flora: C-Values in 343 Taxa with Novel Values for 242. *Adv. Sci. Lett.* **2010**, *3*, 190–213. [CrossRef]
9. Doležel, J.; Čížková, J.; Šimková, H.; Bartoš, J. One Major Challenge of Sequencing Large Plant Genomes Is to Know How Big They Really Are. *Int. J. Mol. Sci.* **2018**, *19*, 3554. [CrossRef]
10. Futuyma, D.J.; Mayer, G.C. Non-allopatric speciation in animals. *Syst. Biol.* **1980**, *29*, 254–271. [CrossRef]
11. Coyne, J.A.; Orr, H.A. *Speciation*; Sinauer Associates: Sunderland, MA, USA, 2004.
12. Rundle, H.D.; Nosil, P. Ecological speciation. *Ecol. Lett.* **2005**, *8*, 336–352. [CrossRef]
13. Turesson, G. The genotypical response of the plant species to the habitat. *Hereditas* **1922**, *3*, 211–350. [CrossRef]
14. Flood, P.J.; Hancock, A.M. The genomic basis of adaptation in plants. *Curr. Opin. Plant Biol.* **2017**, *36*, 88–94. [CrossRef]
15. Campbell, C.R.; Poelstra, J.W.; Yoder, A.D. What is Speciation Genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biol. J. Linn. Soc.* **2018**, *124*, 561–583. [CrossRef]
16. Kardos, M.; Husby, A.; McFarlane, S.E.; Qvarnström, A.; Ellegren, H. Whole-genome resequencing of extreme phenotypes in collared flycatchers highlights the difficulty of detecting quantitative trait loci in natural populations. *Mol. Ecol. Resour.* **2016**, *16*, 727–741. [CrossRef]
17. Martinez Barrio, A.; Lamichhaney, S.; Fan, G.; Rafati, N.; Pettersson, M.; Zhang, H.; Dainat, J.; Ekman, D.; Höppner, M.; Jern, P.; et al. The genetic basis for ecological adaptation of the Atlantic herring revealed by genome sequencing. *Elife* **2016**, *5*, e12081. [CrossRef]
18. Hämälä, T.; Mattila, T.M.; Savolaine, O. Local adaptation and ecological differentiation under selection, migration, and drift in Arabidopsis lyrata. *Evolution* **2018**, *72*, 1373–1386. [CrossRef]
19. Yang, J.; Wariss, H.M.; Tao, L.; Zhang, R.; Yun, Q.; Hollingsworth, P.; Dao, Z.; Luo, G.; Guo, H.; Ma, Y.; et al. De novo genome assembly of the endangered *Acer yangbiense*, a plant species with extremely small populations endemic to Yunnan Province, China. *GigaScience* **2019**, *8*, giz085. [CrossRef]
20. Wang, M.; Zhang, L.; Tong, S.; Jiang, D.; Fu, Z. Chromosome-level genome assembly of a xerophytic plant, *Haloxylon ammodendron*. *DNA Res.* **2022**, *29*, dsac006. [CrossRef]
21. Feng, L.; Lin, H.; Kang, M.; Ren, Y.; Yu, X.; Xu, Z.; Wang, S.; Li, T.; Yang, W.; Hu, Q. A chromosome-level genome assembly of an alpine plant *Crucihimalaya lasiocarpa* provides insights into high-altitude adaptation. *DNA Res.* **2022**, *29*, dsac004. [CrossRef]
22. Li, S.F.; Wang, J.; Dong, R.; Zhu, H.W.; Lan, L.N.; Zhang, Y.L.; Li, N.; Deng, C.L.; Gao, W.J. Chromosome-level genome assembly, annotation and evolutionary analysis of the ornamental plant *Asparagus setaceus*. *Hortic. Res.* **2020**, *7*, 48. [CrossRef]
23. Xu, Q.; Niu, S.C.; Li, K.L.; Zheng, P.J.; Zhang, X.J.; Jia, Y.; Liu, Y.; Niu, Y.X.; Yu, L.H.; Chen, D.F.; et al. Chromosome-scale assembly of the *Dendrobium nobile* genome provides insights into the molecular mechanism of the biosynthesis of the medicinal active ingredient of *Dendrobium*. *Front. Genet.* **2022**, *13*, 844622. [CrossRef]
24. Chung, O.; Kim, J.; Bolser, D.; Kim, H.M.; Jun, J.H.; Choi, J.P.; Jang, H.D.; Cho, Y.S.; Bhak, J.; Kwak, M. A chromosome-scale genome assembly and annotation of the spring orchid (*Cymbidium goeringii*). *Mol. Ecol. Resour.* **2022**, *22*, 1168–1177. [CrossRef] [PubMed]
25. Grover, C.E.; Wendel, J.F. Recent insights into mechanisms of genome size change in plants. *J. Bot.* **2010**, *2010*, 382732. [CrossRef]
26. Kejnovsky, E.; Hawkins, J.S.; Feschotte, C. Plant transposable elements: Biology and evolution. In *Plant Genome Diversity Vol. 1: Plant Genomes, Their Residents and Their Evolutionary Dynamics*; Wendel, J.F., Greilhuber, J., Doležel, J., Leitch, I.J., Eds.; Springer: Vienna, Austria, 2012; pp. 17–33.
27. Sakkour, A.; Mascher, M.; Himmelbach, A.; Haberer, G.; Lux, T.; Spannagl, M.; Stein, N.; Kawamoto, S.; Sato, K. Chromosome-scale assembly of barley cv. 'Haruna Nijo' as a resource for barley genetics. *DNA Res.* **2022**, *29*, dsac001. [CrossRef] [PubMed]
28. Yang, Y.; Huang, L.; Xu, C.; Qi, L.; Wu, Z.; Li, J.; Chen, H.; Wu, Y.; Fu, T.; Zhu, H.; et al. Chromosome-scale genome assembly of areca palm (*Areca catechu*). *Mol. Ecol. Resour.* **2021**, *21*, 2504–2519. [CrossRef]

29. Liao, N.; Hu, Z.; Miao, J.; Hu, X.; Lyu, X.; Fang, H.; Zhou, Y.M.; Mahmoud, A.; Deng, G.; Meng, Y.Q.; et al. Chromosome-level genome assembly of bunching onion illuminates genome evolution and flavor formation in Allium crops. *Nat. Commun.* **2022**, *13*, 6690. [CrossRef]
30. Zhang, G.; Liu, X.; Quan, Z.; Cheng, S.; Xu, X.; Pan, S.; Xie, M.; Zeng, P.; Yue, Z.; Wang, W.; et al. Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. *Nat. Biotechnol.* **2012**, *30*, 549–554. [CrossRef]
31. Vadakkemukadiyil Chellappan, B.; Pr, S.; Vijayan, S.; Rajan, V.S.; Sasi, A.; Nair, A.S. High quality draft genome of Arogyapacha (*Trichopus zeylanicus*), an important medicinal plant endemic to western Ghats of India. *G3 Genes Genomes Genet.* **2019**, *9*, 2395–2404. [CrossRef]
32. Ning, Y.; Li, Y.; Dong, S.B.; Yang, H.G.; Li, C.Y.; Xiong, B.; Yang, J.; Hu, Y.K.; Mu, X.Y.; Xia, X.F. The chromosome-scale genome of *Kobresia myosuroides* sheds light on karyotype evolution and recent diversification of a dominant herb group on the Qinghai-Tibet Plateau. *DNA Res.* **2023**, *30*, dsac049. [CrossRef]
33. Jang, T.S.; Emadzade, K.; Parker, J.; Temsch, E.M.; Leitch, A.R.; Speta, F.; Weiss-Schneeweiss, H. Chromosomal diversification and karyotype evolution of diploids in the cytologically diverse genus Prospero (Hyacinthaceae). *BMC Evol. Biol.* **2013**, *13*, 136. [CrossRef]
34. Jang, T.S.; Parker, J.S.; Emadzade, K.; Temsch, E.M.; Leitch, A.R.; Weiss-Schneeweiss, H. Multiple origins and nested cycles of hybridization result in high tetraploid diversity in the monocot Prospero. *Front. Plant Sci.* **2018**, *9*, 433. [CrossRef]
35. D'hont, A.; Denoeud, F.; Aury, J.M.; Baurens, F.C.; Carreel, F.; Garsmeur, O.; Noel, B.; Bocs, S.; Droc, G.; Rouard, M.; et al. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* **2012**, *488*, 213–217. [CrossRef]
36. Sun, X.; Zhu, S.; Li, N.; Cheng, Y.; Zhao, J.; Qiao, X.; Lu, L.; Liu, S.; Wang, Y.; Liu, C.; et al. A chromosome-level genome assembly of garlic (*Allium sativum*) provides insights into genome evolution and allicin biosynthesis. *Mol. Plant* **2020**, *13*, 1328–1339. [CrossRef]
37. Doyle, J.J.; Doyle, J.L. Isolation of plant DNA from fresh tissue. *Focus* **1990**, *12*, 13–15.
38. Marçais, G.; Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of *k*-mers. *Bioinformatics* **2011**, *27*, 764–770. [CrossRef]
39. Cheng, H.; Concepcion, G.T.; Feng, X.; Zhang, H.; Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **2021**, *18*, 170–175. [CrossRef]
40. Vaser, R.; Sović, I.; Nagarajan, N.; Šikić, M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* **2017**, *27*, 737–746. [CrossRef]
41. Durand, N.C.; Shamim, M.S.; Machol, I.; Rao, S.S.; Huntley, M.H.; Lander, E.S.; Aiden, E.L. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **2016**, *3*, 95–98. [CrossRef]
42. Dudchenko, O.; Batra, S.S.; Omer, A.D.; Nyquist, S.K.; Hoeger, M.; Durand, N.C.; Shamim, M.S.; Machol, I.; Lander, E.S.; Aiden, A.P.; et al. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* **2017**, *356*, 92–95. [CrossRef]
43. Durand, N.C.; Robinson, J.T.; Shamim, M.S.; Machol, I.; Mesirov, J.P.; Lander, E.S.; Aiden, E.L. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst.* **2016**, *3*, 99–101. [CrossRef]
44. Kundu, R.; Casey, J.; Sung, W.K. HyPo: Super Fast & Accurate Polisher for Long Read Genome Assemblies. *bioRxiv* **2019**. [CrossRef]
45. Li, H. New strategies to improve minimap2 alignment accuracy. *Bioinformatics* **2021**, *37*, 4572–4574. [CrossRef] [PubMed]
46. Gurevich, A.; Saveliev, V.; Vyahhi, N.; Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **2013**, *29*, 1072–1075. [CrossRef] [PubMed]
47. Manni, M.; Berkeley, M.R.; Seppey, M.; Zdobnov, E.M. BUSCO: Assessing genomic data quality and beyond. *Curr. Protoc.* **2021**, *1*, e323. [CrossRef] [PubMed]
48. Yu, Y.; Ouyang, Y.; Yao, W. shinyCircos: An R/Shiny application for interactive creation of Circos plot. *Bioinformatics* **2018**, *34*, 1229–1231. [CrossRef] [PubMed]
49. Soderlund, C.; Bomhoff, M.; Nelson, W.M. SyMAP v3.4: A turnkey synteny system with application to plant genomes. *Nucleic Acids Res.* **2011**, *39*, e68. [CrossRef] [PubMed]
50. Wheeler, T.J.; Clements, J.; Eddy, S.R.; Hubley, R.; Jones, T.A.; Jurka, J.; Smit, A.F.A.; Finn, R.D. Dfam: A database of repetitive DNA based on profile hidden Markov models. *Nucleic Acids Res.* **2013**, *41*, 70–82. [CrossRef]
51. Tarailo-Graovac, M.; Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinform.* **2009**, *25*, 4–10. [CrossRef]
52. Flynn, J.M.; Hubley, R.; Goubert, C.; Rosen, J.; Clark, A.G.; Feschotte, C.; Smit, A.F. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 9451–9457. [CrossRef]
53. Benson, G. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580. [CrossRef]
54. Bao, Z.; Eddy, S.R. Automated de novo identification of repeat sequence families in sequenced genomes. *Genome Res.* **2002**, *12*, 1269–1276. [CrossRef]
55. Price, A.L.; Jones, N.C.; Pevzner, P.A. De novo identification of repeat families in large genomes. *Bioinformatics* **2005**, *21*, i351–i358. [CrossRef]
56. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010. Available online: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed on 27 June 2022).

57. Campbell, M.S.; Holt, C.; Moore, B.; Yandell, M. Genome annotation and curation using MAKER and MAKER-P. *Curr. Protoc. Bioinform.* **2014**, *48*, 4–11. [CrossRef]
58. Korf, I. Gene finding in novel genomes. *BMC Bioinform.* **2004**, *5*, 59. [CrossRef]
59. Keller, O.; Kollmar, M.; Stanke, M.; Waack, S. A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics* **2011**, *27*, 757–763. [CrossRef]
60. Card, D.C.; Adams, R.H.; Schield, D.R.; Perry, B.W.; Corbin, A.B.; Pasquesi, G.I.M.; Row, K.; Van Kleeck, M.J.; Daza, J.M.; Booth, W.; et al. Genomic basis of convergent island phenotypes in boa constrictors. *Genome Biol. Evol.* **2019**, *11*, 3123–3143. [CrossRef]
61. Lomsadze, A.; Ter-Hovhannisyan, V.; Chernoff, Y.O.; Borodovsky, M. Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res.* **2005**, *33*, 6494–6506. [CrossRef]
62. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **2013**, *8*, 1494–1512. [CrossRef]
63. Wu, T.D.; Reeder, J.; Lawrence, M.; Becker, G.; Brauer, M.J. GMAP and GSNAP for Genomic Sequence Alignment: Enhancements to Speed, Accuracy, and Functionality. *Methods Mol. Biol.* **2016**, *1418*, 283–334.
64. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and samtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef]
65. Haas, B.J.; Delcher, A.L.; Mount, S.M.; Wortman, J.R.; Smith, R.K., Jr.; Hannick, L.I.; Maiti, R.; Ronning, C.M.; Rusch, D.B.; Town, C.D.; et al. Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* **2003**, *31*, 5654–5666. [CrossRef]
66. Li, H. Protein-to-genome alignment with miniport. *Bioinformatics* **2023**, *39*, btad014. [CrossRef] [PubMed]
67. Haas, B.J.; Salzberg, S.L.; Zhu, W.; Pertea, M.; Allen, J.E.; Orvis, J.; White, O.; Buell, C.R.; Wortman, J.R. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* **2008**, *9*, R7. [CrossRef] [PubMed]
68. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [CrossRef] [PubMed]
69. Boeckmann, B.; Bairoch, A.; Apweiler, R.; Blatter, M.C.; Estreicher, A.; Gasteiger, E.; Martin, M.J.; Michoud, K.; O'Donovan, C.; Phan, I.; et al. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **2003**, *31*, 365–370. [CrossRef] [PubMed]
70. Finn, R.D.; Attwood, T.K.; Babbitt, P.C.; Bateman, A.; Bork, P.; Bridge, A.J.; Chang, H.Y.; Dosztányi, Z.; El-Gebali, S.; Fraser, M.; et al. InterPro in 2017—Beyond protein family and domain annotations. *Nucleic Acids Res.* **2017**, *45*, D190–D199. [CrossRef]
71. Huerta-Cepas, J.; Szklarczyk, D.; Heller, D.; Hernández-Plaza, A.; Forslund, S.K.; Cook, H.; Mende, D.R.; Letunic, I.; Rattei, T.; Jensen, L.J.; et al. eggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **2019**, *47*, D309–D314. [CrossRef]
72. Emms, D.M.; Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **2019**, *20*, 238. [CrossRef]
73. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [CrossRef]
74. Talavera, G.; Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **2007**, *56*, 564–577. [CrossRef]
75. Kozlov, A.M.; Darriba, D.; Flouri, T.; Morel, B.; Stamatakis, A. RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **2019**, *35*, 4453–4455. [CrossRef]
76. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [CrossRef]
77. Hedges, S.B.; Dudley, J.; Kumar, S. TimeTree: A public knowledge-base of divergence times among organisms. *Bioinformatics* **2006**, *22*, 2971–2972. [CrossRef]
78. Mendes, F.K.; Vanderpool, D.; Fulton, B.; Hahn, M.W. CAFE 5 models variation in evolutionary rates among gene families. *Bioinformatics* **2020**, *36*, 5516–5518. [CrossRef]
79. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef]
80. Garrison, E.; Marth, G. Haplotype-based Variant Detection from Short-Read Sequencing. *arXiv* **2012**, arXiv:1207.3907.
81. Garrison, E. FreeBayes Source Repository. 2012. Available online: https://github.com/ekg/freebayes (accessed on 15 March 2023).

7.2 Publication II

Šarančić, S. L., Pleić, N., Križanović, K., Surina, B., Mitić, D., and Radosavljević, I. (2025a). Uncovering the genomic basis of phenological traits in *Chouardia litardierei* (Asparagaceae) through a genome-wide association study (GWAS). Front Plant Sci 16, 1571608. doi: 10.3389/FPLS.2025.1571608

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 17 April 2025
DOI 10.3389/fpls.2025.1571608

OPEN ACCESS

EDITED BY
Rita Lourenço Costa,
National Institute for Agricultural and Veterinary Research (INIAV), Portugal

REVIEWED BY
Marianella Quezada,
Universidad de la República, Uruguay
Mubashir Abbas,
Chinese Academy of Agricultural Sciences (CAAS), China

*CORRESPONDENCE
Ivan Radosavljević
ivan.radosavljevic@biol.pmf.hr

RECEIVED 05 February 2025
ACCEPTED 01 April 2025
PUBLISHED 17 April 2025

CITATION
Šarančić SL, Pleić N, Križanović K, Surina B, Mitić D and Radosavljević I (2025) Uncovering the genomic basis of phenological traits in *Chouardia litardierei* (Asparagaceae) through a genome-wide association study (GWAS). *Front. Plant Sci.* 16:1571608. doi: 10.3389/fpls.2025.1571608

Uncovering the genomic basis of phenological traits in *Chouardia litardierei* (Asparagaceae) through a genome-wide association study (GWAS)

Sara Laura Šarančić[1], Nikolina Pleić[2], Krešimir Križanović[3], Boštjan Surina[4,5], Damjan Mitić[1] and Ivan Radosavljević[1]*

[1]Department of Biology, Faculty of Science, University of Zagreb, Zagreb, Croatia, [2]Department of Biology and Human Genetics, School of Medicine, University of Split, Split, Croatia, [3]Department of Electronic Systems and Information Processing, Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia, [4]Natural History Museum Rijeka, Rijeka, Croatia, [5]Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia

Chouardia litardierei (Asparagaceae) is a non-model, perennial species characterized by exceptional ecological plasticity. In this research, we studied the genetic architecture underlying several phenological traits in selected ecologically diverged populations of this species. We conducted a genome-wide association study (GWAS) to identify genomic regions linked to the following populations-specific phenological traits: Beginning of Sprouting (BOS), Beginning of Flowering (BOF), Flowering Period Duration (FPD), and Vegetation Period Duration (VPD). Combining phenological data from a common garden experiment with an SNP dataset obtained through the ddRAD-seq approach, we identified numerous loci associated with these traits using single- and multi-locus GWAS models. Narrow-sense heritability estimates were high for all traits, with the VPD trait showing the highest estimate (86.95%), emphasizing its importance for local adaptation. Functional annotation of associated genomic regions revealed key protein families involved in flowering time regulation, vegetative growth timing, and stress adaptation. These findings provide insights into the molecular mechanisms of local adaptation in *C. litardierei*'s populations from different habitats, emphasizing the role of genetic factors in phenological trait variation and ecological divergence across populations.

KEYWORDS

GWAS, local adaptation, adaptive traits, *Chouardia litardierei*, phenology

Introduction

Understanding the genetic basis of phenotypic variation is essential for evolutionary biology, as it elucidates mechanisms underlying speciation, biogeographical distributions, and fitness in natural populations (Savolainen et al., 2013; Mckown et al., 2014). Natural selection acts on allele frequencies within populations, shaping their variation and promoting adaptive traits that enhance survivability and reproductive success (Hu et al., 2020; Walter et al., 2022; Lee et al., 2023). As populations undergo local adaptation, ecological speciation may lead to the emergence of new ecotypes (Turesson, 1922; Todesco et al., 2020) — genetically distinct populations of the same species well-adapted to specific ecological niches (Rundle and Nosil, 2005; Cortés et al., 2018). Although the role of ecotypes in the speciation process remains debated (Lowry, 2012; Fernández-Meirama et al., 2022), several studies highlight their importance in driving genetic divergence along ecological gradients (Lowry et al., 2008; Brandrud et al., 2017; Cortés et al., 2018; Bakhtiari et al., 2019). Rapidly evolving lineages in heterogeneous environments offer valuable insights into the genetic mechanisms driving adaptation and speciation (Feder et al., 2011; Cortés et al., 2018).

Phenology is one of the key features of plants as sessile organisms. It determines the timing of life cycle phases and the duration of growth and reproduction (Schwartz, 2003). Although other factors like photoperiod (Adole et al., 2019; Wang et al., 2020), water availability (Zhou et al., 2024), or selection by pollinators (Sandring and Ågren, 2009) may play an important role as well, temperature is considered to be the environmental element with the most substantial impact on various phenological traits (Schwartz, 2003; Cook et al., 2012). Matching the growth and especially reproduction periods with the optimal environmental conditions is of exceptional evolutionary importance and is strongly influenced by natural selection (Duputié et al., 2015). Among phenological traits, flowering time is particularly sensitive to environmental factors, marking a critical transition from vegetative growth to reproduction (Hill and Li, 2016; Gaudinier and Blackman, 2020). In seasonally variable habitats, where the timing and duration of the vegetational season differ across landscapes, plants must initiate sprouting and flowering within a constrained annual timeframe (Anderson et al., 2012). Therefore, the regulation of flowering time emerges as a frequent target of evolutionary processes (Gaudinier and Blackman, 2020). Ecologically divergent taxa in numerous lineages often have different flowering times (e.g., Heslop-Harrison, 1964; Grant, 1981; Levin, 2000) suggesting that some niche shifts were predicated upon temporal change (Levin, 2006). Consequently, alterations in flowering schedules may allow populations to better exploit different groups of pollinators (e.g., Waser, 1983; Goldblatt and Manning, 1996; Johnson et al., 1998), while movements into new pollinator niches are accompanied by changes in floral attributes (Levin, 2006). Natural selection generally favours bigger individuals at maturity; however, the timing of flowering presents a trade-off between maximizing fecundity and ensuring reproductive completion before adverse conditions, such as drought or winter, occur (Anderson et al., 2012). Species facing water limitations often adjust their flowering phenology to align with peak moisture availability, taking advantage of optimal conditions (Settele et al., 2016). For example, Schmalenbach et al. (2014) found that late-flowering *Arabidopsis* plants coped better with drought by compensating for early growth losses with later recovery, while early-flowering plants, which may flower sooner to exploit available moisture before drought, exhibited lower fitness under the same conditions. High salinity also impairs plant growth in *Arabidopsis*, acting as a suppressive factor that delays flowering time (Li et al., 2007; Lee et al., 2023). Coupled with variation in mating opportunity, temporal variation in sexual phases of individual flowers may have a significant impact on reproductive success in dichogamous plants (Sargent and Roitberg, 2000). Since phenological traits display extensive variations in plants and are often related to local adaptation (Rathcke and Lacey, 1985), the analysis of their genetic background presents a great opportunity to study the mechanisms of the adaptive divergence process.

Investigating the genomic underpinnings of specific traits within the framework of environmental dynamics is essential for uncovering the mechanisms driving local adaptation and elucidating the complex relationship between phenological traits and adaptive responses (Bernatchez et al., 2023). Although much of our understanding of flowering regulation and vegetation duration derives from studies on model organisms such as *A. thaliana* (Engelmann and Purugganan, 2006; Kinmonth-Schultz et al., 2021), significant advancements have also been made in agriculturally important species (e.g., Molla, 2022; Vicentini et al., 2023; Flohr et al., 2017; Song et al., 2023). However, broadening this research beyond model organisms could increase our understanding of the diverse genetic mechanisms governing phenological variation in populations of wild, non-model species facing different ecological pressures in their habitats.

Here, we investigated the genetic basis of selected phenological traits in the amethyst meadow squill, *Chouardia litardierei* (Breist.) Speta; a small, bulbous, perennial species belonging to the Asparagaceae family [following the APG III system (Bremer et al., 2009)]. Being a typical geophyte, *C. litardierei* plants undergo a dormancy period, which usually stretches from mid-summer to late autumn or early spring, depending on the individual season's properties. During the spring, soon after the development of young leaves, inflorescence emerges. From late April to early June, depending on the population's location, the flowering phenophase will unfold, shortly followed by fruiting, which marks the beginning of dying back to an underground perennating organ, i.e., a bulb. *C. litardierei* produces a large racemose inflorescence, typically consisting of several dozen radially symmetrical flowers, without any apparent morphological adaptations for specific pollination mechanisms. While this has not been formally studied, it is expected to be an open-pollinated species (pers. obs.). In addition to sexual reproduction, it propagates clonally through the formation of bulbs surrounding the central bulb. *C. litardierei* populations are found across the Dinaric Alps in the western parts of the Balkan Peninsula (Ritter-Studnička, 1954; Gaži-Baskova, 1962). Throughout this region, populations inhabit

highly contrasting habitats, thus indicating a very pronounced ecological plasticity of the species (Figure 1).

In terms of habitat types, the most substantial contrast can be observed between southernmost populations, which are found on patches of exposed dolomite bedrocks or dry mountainous grasslands with very thin and sparse soil cover, on one side, and populations occupying lush meadows of karst poljes, enclosed depressions with deep and fertile soils, abundant in water, on another. These groups of populations cope with very different types of challenges. For the first group of the populations, the most substantial adaptation pressure is expected to come from limited resource availability accompanied by pronounced seasonality in water availability and temperature, which are usual for such a habitat (Mota et al., 2021). At the same time, the second group faces seasonal flooding that can last up to seven months each year (Mihevc et al., 2010; Bonacci, 2014). In addition to these two prevailing groups of populations, based on a habitat type, a third and the smallest group can be recognized, the one inhabiting deep-soiled marshes along the coastline in western parts of the species

FIGURE 1
Habitat types of the studied *Chouardia litardierei* populations, shown from top to bottom: **(A)** Karst poljes meadows (locality of Budoške Bare population), **(B)** Dry mountainous grasslands with exposed bedrock (Lovćen), and **(C)** Saline coastal marshes (Vrana Lake).

distribution range. These populations situated in proximity to the seashore are experiencing different climates [i.e., Cfa and Csb climate types according to Köppen classification (Köppen, 1918)] than other inland meadow-habitat populations, which, for the most part, are found in habitats characterized by Cfb type of climate. In addition, these seashore populations are exposed to periodical sea flooding, which causes an increase in soil salinity, one of the major factors in plant ecology (Bui, 2013). Nonetheless, it is essential to note that although this issue was already addressed by Šilić (1990), no clear differentiation, either phenotypic or genetic, among these groups of populations has yet been recognized.

To learn as much as possible about the genetic background of phenological traits of selected *C. litardierei* populations from across its distribution range and from different habitats, results from a common garden experiment and genotyping were processed through a set of comprehensive single- and multi-locus genome-wide association (GWA) models. Functional annotation of recognized candidate loci was further performed, thus enabling us to deepen our understanding of the complex genetics behind the phenological aspect of adaptive divergence and to analyze the extent to which differentiation of the studied populations has advanced.

Methods

Plant material, common garden experiment, and phenotyping

To establish the common garden experiment, 214 individuals were relocated from nine chosen populations of *C. litardierei*. Three populations were selected to represent each of the three presumed groups of populations from different habitat types, as illustrated in Figure 1.

During the sampling expeditions, 22 to 25 individuals were selected from each population, ensuring a minimum distance of 10 meters between them, following the 1:20 rule (Wagner, 1995). The geographic coordinates of the sampling locations are listed in Supplementary File 1. Leaf material from each individual was collected for DNA extraction and desiccated using silica gel. Each sampled individual (represented by a single bulb) was transplanted into a separate two-litre plastic container filled with a mixture of soil, sand, and perlite. The containers were placed in raised beds outdoors, creating a common garden setup that exposed the plants to a temperate continental climate (Cfb climate type) (Köppen, 1918; Zaninović et al., 2008). No additional interventions, such as supplemental watering or pesticide use, were applied, allowing the plants to grow under natural, undisturbed environmental conditions.

After two vegetative seasons of acclimatization, four phenological traits were selected for further research (Table 1): (i) Flowering Period Duration (FPD), calculated as the number of days from the appearance of the first flower to the last; (ii) Vegetation Period Duration (VPD), measured as the time from sprouting to the opening of the first capsule with ripened seeds, also in days; (iii) Beginning of Flowering (BOF), recorded using the earliest plant flowering dates a reference; and (iv) Beginning of Sprouting (BOS), noted by referencing the sprouting date of the first individual. All traits examined were measured on an individual genotype level and were considered polygenic.

To assess differences in phenological traits across individual populations and three groups of populations originating from different habitat types, Kruskal-Wallis tests were implemented in the PAST software (Hammer et al., 2001), were performed. We further performed pairwise comparisons using Mann-Whitney *post-hoc* tests with Bonferroni correction to identify significant trait variations. Since none of the variables followed a normal distribution, Spearman's correlation analysis was conducted to examine the relationships between FPD, VPD, BOF, and BOS variables using the "stats" package in R (R Core Team, 2016).

Sequencing, genomic data processing, and population genetic structure

DNA isolation was carried out using the GenElute™ Plant Genomic DNA Miniprep Kit (Sigma–Aldrich®). DNA concentrations were measured with the Qubit™ Fluorometer (Thermo Fisher Scientific, Wilmington, DE, USA), and samples were subsequently diluted to a concentration of 20 ng/µL.

For genotyping the studied *C. litardierei* populations, a ddRAD-seq approach was utilized (Peterson et al., 2012). DNA was initially digested with two restriction enzymes, AseI and NsiI (NEB #R0526L and #R0127L, respectively). The resulting fragments were then ligated with barcoded i5 and i7 adapters, allowing all samples to be multiplexed. Final amplification was carried out after nick repair using DNA polymerase I (NEB #M0209L). The resulting DNA libraries were double-sequenced (150 bp paired-end) on the Illumina HiSeq X platform.

The initial sequencing data underwent preprocessing for quality trimming and adapter removal using Trim Galore (Martin, 2011).

TABLE 1 Descriptive statistics of the *Chouardia litardierei* phenological traits examined in the study.

Overall			
Trait (days)	Description	Median (Q1 – Q3)	Min – max
FPD	Duration from the date of the first to the last flower for each genotype	17 (15 – 18)	9 - 25
VPD	Duration from genotype sprouting to the opening of the first capsule	97 (88 – 107)	55 - 162
BOF	Beginning of flowering considering the flowering date of the first genotype as a reference	11 (10 – 13)	1 - 33
BOS	Beginning of sprouting considering the sprouting date of the first genotype as a reference	56 (52 – 63)	1 - 88

All traits were measured in days. BOF, Beginning of Flowering; BOS, Beginning of Sprouting; FPD, Flowering Period Duration; max, maximum value; min, minimum value; VPD, Vegetation Period Duration; Q1, first quartile; Q3, third quartile.

Post-trimming, BAM files were generated by aligning the reads to the *C. litardierei* reference genome (Radosavljević et al., 2023) using the Burrows-Wheeler Aligner (Li and Durbin, 2009). SNP identification was performed with the Stacks software package v1.48 (Catchen et al., 2013). The ref_map.pl wrapper module was utilized, following Paris et al. (2017) recommendations, the *pstacks* module was executed to extract loci previously aligned to the reference genome, with a minimum coverage depth of three reads to ensure a reliable representation of loci across samples and reduce low-confidence genotype calls. The *cstacks* module then constructed a comprehensive catalogue of loci across populations, allowing a maximum of four mismatches among sample loci to minimize alignment errors. Subsequently, the *populations* module calculated population-level summary statistics. To ensure high data quality, loci were retained only if present in all nine populations and at least 70% of individuals within each population, with a maximum observed heterozygosity of 0.70. Additional filtering criteria included retaining only one SNP per locus and excluding loci with minor allele frequencies (MAF) below 1%. This stringent filtering approach focused on common and well-represented genetic variants, reducing the risk of inaccuracies due to sequencing or sampling errors. The resulting dataset, comprising high-quality genetic markers, was exported in .vcf format for downstream analysis.

To assess the neutral population genetic structure of the studied populations, we used a model-based clustering method implemented in ParallelStructure (Pritchard et al., 2000; Besnier and Glover, 2013). To overcome the issue of this analysis's high computational demands and lengthy processing time for such a large number of SNPs, we constructed a subset of 5,000 randomly selected SNPs. The analysis comprised ten runs for each of the ten clusters (K). Each run consisted of a burn-in period of 50,000 steps, followed by 500,000 Monte Carlo Markov Chain (MCMC) replicates. We used the StructureSelector online software (Li and Liu, 2018) to obtain the most likely number of clusters (K) following Evanno's method (Evanno et al., 2005) as well as to retrieve the final data through the clustering and averaging of the runs following the Clumpak algorithm (Kopelman et al., 2015). The obtained results were processed using CorelDRAW X7 v.17.1.0.572 software (Corel Corp., Ottawa, Canada) for improved visualization.

Genome-wide association analyses

Figure 2 illustrates a schematic representation of the methodological approach used in this study. All traits were treated as polygenic and GWAS analyses were carried out assuming an additive genetic model. Variants with a minor allele frequency (MAF) below 1% were excluded using the BCFtools software (Danecek et al., 2021). Two distinct statistical approaches were employed for each association analysis: the frequentist single-locus approach and the Bayesian multi-locus approach. In the frequentist single-locus approach, two distinct models were applied. A standard linear mixed model (LMM) was fitted using GEMMA 0.98.5 (Zhou and Stephens, 2012) for all four traits, keeping in mind that this approach assumes a normal trait distribution. Additionally, all traits were analyzed using GMMAT 1.4.2 (Chen et al., 2019), applying a Poisson generalized linear mixed model (GLMM), to account for their count-based distributions. The Poisson GLMM in GMMAT was selected because it effectively accounts for the non-normal distribution of count data, providing a complementary approach to the LMM analysis performed in GEMMA.

In the Bayesian multi-locus approach, a Bayesian sparse linear mixed model (BSLMM) (Zhou et al., 2013) was simultaneously fitted for all traits under analysis. Significant SNPs for each trait were identified by first intersecting the sets of significant SNPs obtained from GLMM and LMM, and then further intersecting the resulting set with those identified by BSLMM, ensuring consistency across both the frequentist and Bayesian approaches (Figure 2). Additionally, a multivariate linear mixed model (mvLMM) was performed in GEMMA to simultaneously analyze significantly correlated traits (FPD and VPD, as well as BOF and VPD) to identify shared association signals between them.

The results were visualized using Manhattan plots generated with the R package "qqman" (Turner, 2018) and "CM plot" (Yin et al., 2021). An *ad hoc* threshold of 1×10^{-3} was used for the frequentist GWAS analyses (GLMM, LMM, and mvLMM).

Generalized linear mixed model using a poisson distribution

The generalized linear mixed model (GLMM) with a Poisson distribution was applied using GMMAT, and the model is expressed as follows (Equations 1–3):

$$log(\mu_i) = \mathbf{W_i}\alpha + \mathbf{x_i}\beta + u_i \quad (1)$$

$$u \sim MVN_n(0,\ \lambda\mathbf{K}) \quad (2)$$

$$y_i \sim Poisson(\mu_i) \quad (3)$$

In this model, y_i represents the observed count for the i-th individual, while μ_i denotes the mean count, modeled as the exponential of the linear predictor. $\mathbf{W_i}$ is the i-th row of an $n \times c$ matrix of covariates (fixed effects), α is the corresponding vector of coefficients for these covariates, $\mathbf{x_i}$ represents the genotype of the i-th individual, and β denotes the effect size of the genetic marker. The random effects u are assumed to follow a multivariate normal distribution MVN_n (0,λK), where $\mathbf{K}$ is the relatedness matrix of size $n \times n$, and λ represents the ratio of variance components. The observed data y_i is assumed to follow a Poisson distribution with μ_i. This model incorporates individual-level random effects and a genetic relationship matrix $\mathbf{K}$ to account for population structure and relatedness. When assuming a normal distribution and an identity link function for continuous traits, GMMAT conducts association tests using linear mixed models (LMMs).

10.3389/fpls.2025.1571608

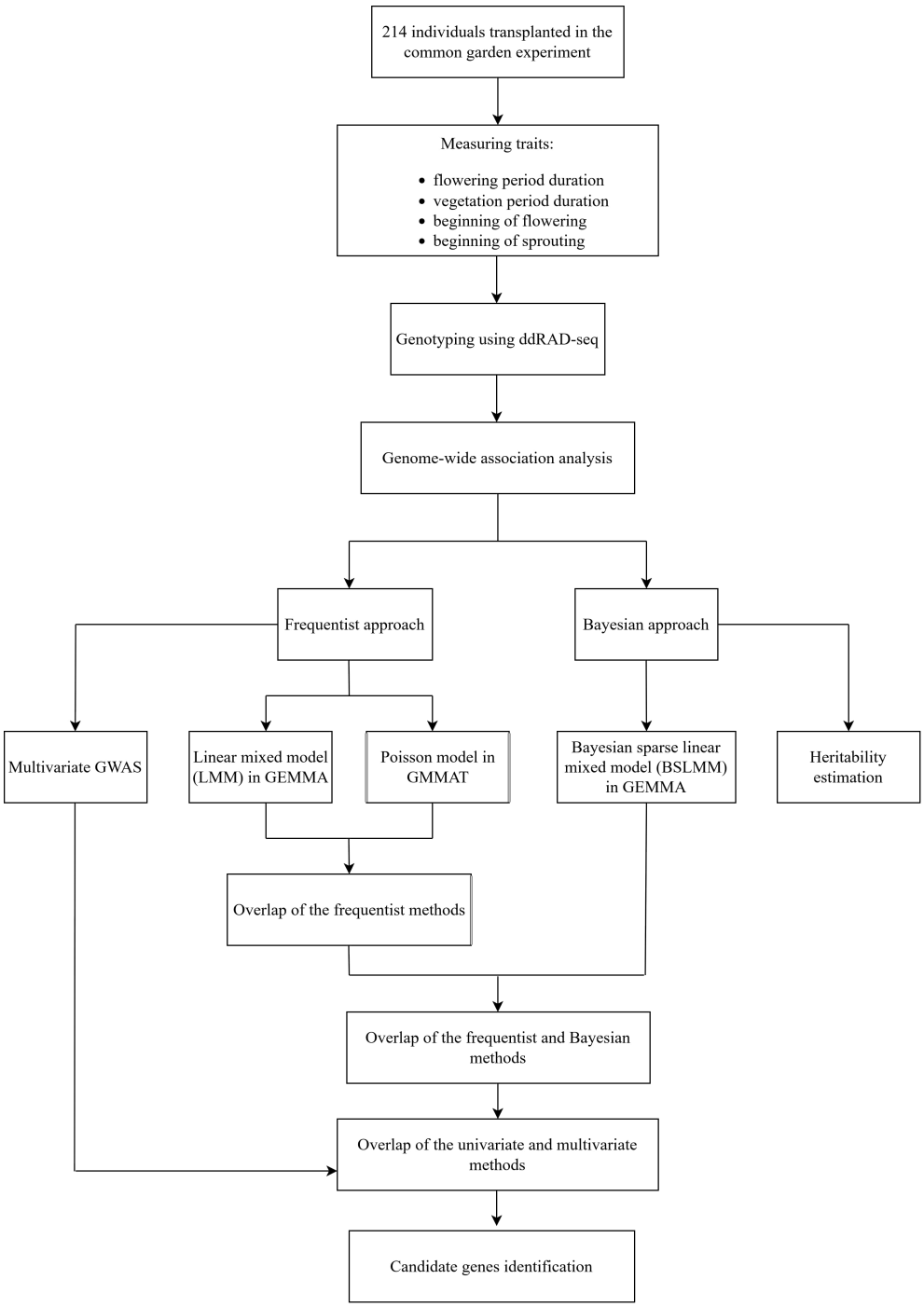


FIGURE 2
A schematic outline of the methodological approach employed to study the genetic basis of phenological traits in *Chouardia litardierei*.

Linear mixed model

The standard LMM was applied using GEMMA 0.98.5. in the following form:

$$\mathbf{y} = \mathbf{W}\alpha + \mathbf{x}\beta + \mathbf{u} + \varepsilon \quad (4)$$

$$\mathbf{u} \sim MVN_n(0,\ \lambda\tau^{-1}\mathbf{K}) \quad (5)$$

$$\varepsilon \sim MVN_n(0,\ \tau^{-1}\mathbf{I_n}) \quad (6)$$

Here, $\mathbf{y}$ represents a vector of trait values for 214 individuals, and $\mathbf{W}$ is an $n \times c$ matrix of covariates (fixed effects), which, in this case, consists of a column of 1s. Let α represent a c-vector of the intercept, $\mathbf{x}$ be an n-vector of marker genotypes, and β denote the effect size of the marker. Additionally, $\mathbf{u}$ is an n-vector of random

10.3389/fpls.2025.1571608

effects, ε is an n-vector of errors, τ^{-1} denotes the variance of the residual errors, and λ is the ratio between the two variance components. $\mathbf{K}$ is the known $n \times n$ relatedness matrix, and $\mathbf{I}_n$ is an $n \times n$ identity matrix. MVN_n refers to the n-dimensional multivariate normal distribution. The effect sizes indicate the change in trait values associated with each additional effect allele in the genotypes of individuals.

Bayesian framework

The LMM (Equations 4–6) implemented in GEMMA evaluates the alternative hypothesis H_1: $\beta \neq 0$ against the null hypothesis H_0: $\beta = 0$ for each SNP individually. Extensions of the LMM that account for the effects of variants across multiple loci simultaneously could improve the power to identify causal variants. Bayesian LMMs can model all markers simultaneously by assigning different prior distributions to the marker effects and sampling from their posterior distribution. These Bayesian models, designed for estimating SNP effect sizes, start with a basic linear model that links genotypes $\mathbf{X}$ to phenotypes $\mathbf{y}$:

$$\mathbf{y} = \mathbf{1_n}\,\mu + \mathbf{X}\beta + \varepsilon \tag{7}$$

$$\varepsilon \sim MVN_n(0,\ \tau - 1\mathbf{I_n}) \tag{8}$$

we let $\mathbf{y}$ be a vector of phenotypes observed on n individuals, and $\mathbf{X}$ be an $n \times p$ matrix of genotypes for these same n individuals at p genetic markers. The vector β represents the effects of genetic markers, $\mathbf{1_n}$ is an n-vector of 1s, μ is a scalar representing the mean phenotype, and ε is an n-vector of error terms with variance τ^{-1}. Our aim was to estimate the parameter β, which corresponds to the effects of the genetic markers. However, because the number of genetic markers in our study (p = 23,315) far exceeds the number of individuals (n = 214), certain modeling assumptions regarding SNP effect sizes β had to be made. These assumptions range from the infinitesimal (or polygenic) model, which posits that all SNPs have non-zero effects, to the sparse model, which assumes that only a small subset of SNPs affect the phenotype. The success of the model relies on the true genetic architecture of the trait being studied, although this is typically unknown. The most widely used polygenic model assumes that all SNPs impact the phenotype (i.e., have non-zero effects) with normally distributed effect sizes:

$$\beta \sim N(0, \sigma_\beta^2) \tag{9}$$

When Equations 7–8 are combined with the normality assumption (Equation 9) for effect sizes b, they result in the previously described LMM, as it incorporates a random effect term that represents the combined genetic effects.

Bayesian sparse linear mixed model

A more general assumption, which includes both polygenic and sparse modeling scenarios, suggests that effect sizes come from a mixture of two normal distributions.

$$\beta_i \sim \pi \mathrm{N}(0, \frac{\sigma_a^2 + \sigma_b^2}{p\tau}) + (1-\pi)\mathrm{N}(0, \frac{\sigma_b^2}{p\tau}) \tag{10}$$

In this model, π represents the proportion of SNPs with large effects, while $\sigma^2{}_\beta$ and $\sigma^2{}_\alpha$ correspond to the variances of small and large effects, respectively. The resulting BSLMM model combines polygenic and sparse effects in the prior distribution of effect sizes, allowing it to adapt to various genetic architectures of the traits being studied. BSLMM addresses population structure and relatedness by incorporating a genomic kinship matrix as a random effect term, and it accounts for linkage disequilibrium (LD) by estimating SNP effect sizes β while controlling for other SNPs in the model. The model uses a Markov chain Monte Carlo algorithm to sample from the posterior distribution and estimate SNP effect sizes. Unlike LMM, which provides p-values, BSLMM outputs a posterior inclusion probability (PIP) for each SNP, reflecting the likelihood that a marker is associated with the trait based on the data. This PIP is calculated as the proportion of chain iterations in which the SNP exhibits a large effect. SNPs with high PIPs are considered the most likely functional variants influencing the analyzed traits. We applied BSLMM to the same dataset (214 individuals and 23,315 variants) used in our primary frequentist association analysis to compare single-SNP and multi-SNP approaches and reduce false positives. The BSLMM chain was run with 1,000,000 sampling steps and 100,000 burn-in iterations. We used the estimated PIPs from BSLMM for additional fine-mapping of genomic regions identified in the frequentist analysis.

SNP heritability estimation

The proportion of variance in phenotypes accounted for by all available genotypes (PVE), also referred to as narrow-sense heritability (h^2), along with the proportion of genetic variance explained by variants with large effects (PGE), was estimated for the traits shown in Table 1. This estimation was based on the assumption that SNP effect sizes follow a mixture of two normal distributions (Equation 10), as implemented in GEMMA BSLMM.

Multivariate genome-wide association analyses

To identify common variants associated with the trait pairs showing the strongest statistically significant correlations, multivariate genome-wide association analyses were performed using a multivariate linear mixed model (mvLMM) in GEMMA. Specifically, multivariate GWAS was conducted for the VPD and BOS traits, as well as for the VPD and BOF traits, which exhibited the strongest statisticaly significant correlations. This approach enabled the simultaneous analysis of genetic effects on both trait pairs of traits by treating them as dependent variables. The mvLMM method accounts for population structure and relatedness among individuals, ensuring accurate identification of genetic variants contributing to the observed phenotypic variation in these traits.

Candidate genes prediction

After identifying phenotypic evidence for local adaptation in distinct *C. litardierei* populations and conducting GWAS analysis, efforts focused on pinpointing associated candidate genes. Using the reference genome, sequences were extracted spanning a total of 50 kilobases – including 25 kilobases upstream and downstream of each significant SNP identified through both statistical models, using SAMtools (Danecek et al., 2021). Functional annotations for these sequences were then obtained through the eggNOG-mapper v2 database, applying an *e*-value threshold of $< 1 \times 10^{-2}$ (Huerta-Cepas et al., 2019).

Results

Phenotyping

Figure 3 illustrates the phenological variations observed among *C. litardierei* populations in the common garden experiment.

Out of the 214 individuals sampled across nine populations, 204 flowered successfully. Consequently, all traits related to flowering [FPD, VPD (since its ending is related to the start of the fruiting phenophase), and BOF] were measured and subsequent analyses were performed on the set of 204 individuals, while the remaining 10 were discarded. At the same time, the BOS trait was analyzed across all 214 individuals. The FPD and the VPD ranged from 9 to 25 days, with a median of 17 days (Q1 – Q3: 15 – 18), and 55 to 162 days, with a median of 97 days (Q1 – Q3: 88 – 107), respectively. The BOF and BOS traits ranged from 1 to 33 days, with a median of 11 days (Q1 – Q3: 10 – 13), and 1 to 88 days, with a median of 56 days (Q1 – Q3: 52 – 63), respectively. All the obtained data are summarised in Table 1. Supplementary File 2 contains the results of Kruskal-Wallis and Mann-Whitney *post-hoc* tests for the studied phenological traits, showing significant differences at the population level and between the assumed population groups. The distribution of these phenological traits is visually represented using box plots in Figure 4.

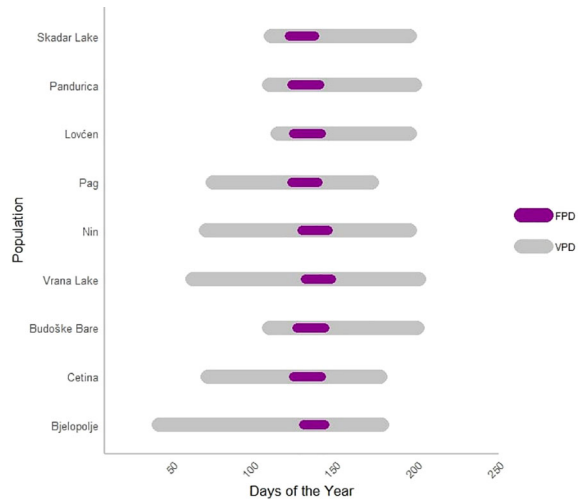


FIGURE 3
The horizontal bar plot illustrates the durations of Vegetation period Duration (VPD) and Flowering Period (FPD) across populations of the *Chouardia litardierei* during one vegetational season. The x-axis represents the days of the year, while the y-axis lists the populations being compared. The dark magenta bars indicate the FPD, which represents the duration from the date of the first to the last flower for each genotype. In contrast, the grey bars represent the VPD, denoting the duration from the genotype sprouting to the opening of the first capsule. Additionally, the figure provides a visual reference for the Beginning of Flowering (BOF) and the Beginning of Sprouting (BOS), where BOF and BOS are calculated relative to the individual that flowered or sprouted first, respectively.

A correlation analysis revealed several significant associations among the studied traits (Table 2). A weak positive correlation was observed between FPD and VPD, while a strong positive correlation was found between VPD and BOF.

Sequencing, genomic data processing, and population genetic structure

The sequencing process generated a total of 1,284,680,304 reads. After filtering the raw sequences and mapping them to the reference genome, 1,278,409,966 reads were retained. SNP identification and filtration were performed using the Stacks software, resulting in the detection of 24,660 SNPs. Following the application of the BCFtools MAF filter with a 1% threshold, 23,315 SNPs were kept for subsequent analysis.

The cluster analysis based on the Bayesian model implemented in the ParallelStructure software revealed that the most likely number of genetic clusters was two (Supplementary File 3). One cluster corresponded to the group of populations from the dolomite bedrock habitat, while the remaining populations formed the other cluster (Supplementary File 4). Such structuring reflects the environmental preferences of the studied populations only to some extent, as populations from seashore and meadow habitats remained grouped without any differentiation among them.

Genome-wide association analyses

The analysis of the FPD trait using LMM identified 48 significant SNPs, while GLMM detected 8. An overlap of these results revealed 8 SNPs that were significant across both methods. Further validation using BSLMM confirmed 3 of these SNPs as significant, with one located on each of chromosomes 10, 7, and 11. For the VPD trait, LMM identified 26 significant SNPs, while GLMM detected 54. Fourteen SNPs were found to overlap between the two methods. Subsequent analysis with BSLMM confirmed 2 of these SNPs as significant, located on chromosomes 4 and 12. In the case of the BOF trait, LMM and GLMM identified 17 and 29 SNPs, respectively, with 8 overlapping SNPs. BSLMM analysis confirmed 1 significant SNP located on chromosome 2. For the BOS trait, LMM identified 34 significant SNPs, while GLMM detected 162. Seven SNPs overlapped between the two methods, and BSLMM analysis confirmed 1 significant SNP

10.3389/fpls.2025.1571608

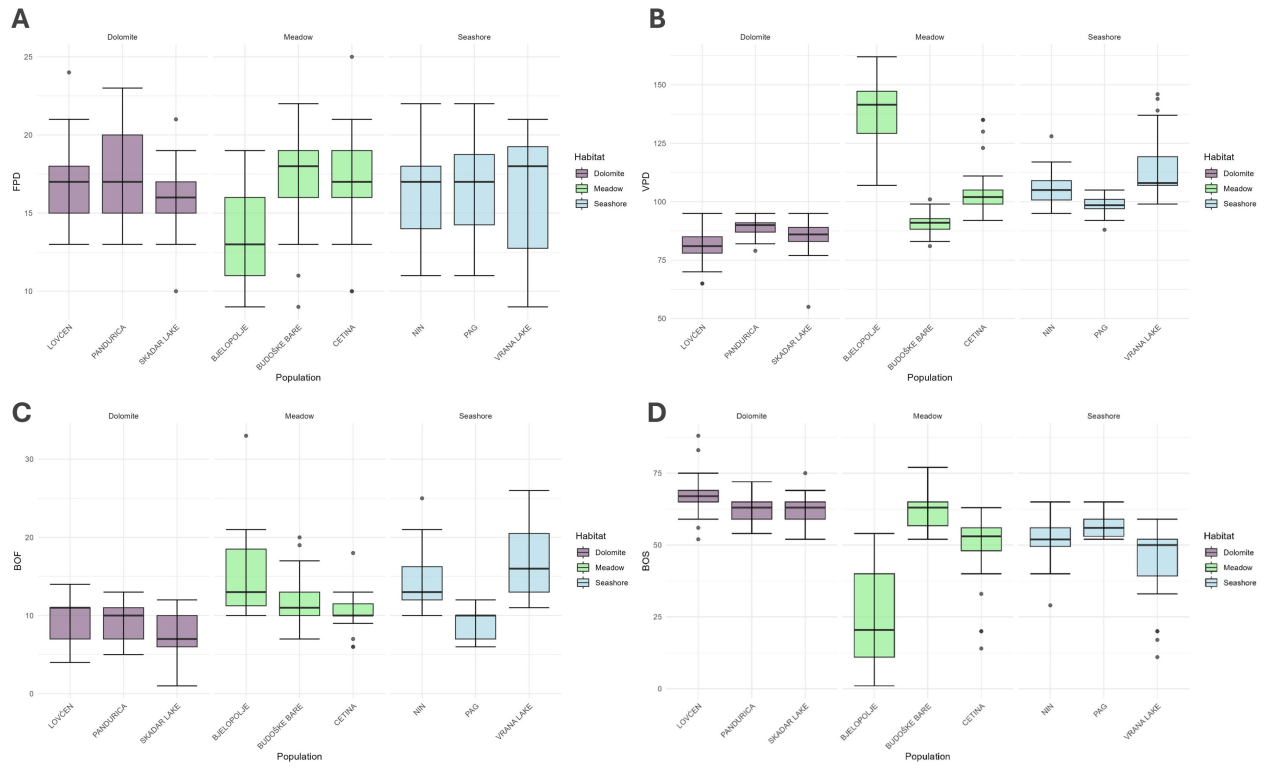


FIGURE 4
Box plots illustrate the obtained phenological results from a common garden experiment, depicting four phenological traits: **(A)** Flowering Period Duration (FPD) (top left), **(B)** Vegetation Period Duration (VPD) (top right), **(C)** Beginning of Flowering (BOF) (bottom left), and **(D)** Beginning of Sprouting (BOS) (bottom right) per genotype. Each box represents the interquartile range (IQR), with the horizontal line inside the box indicating the median. Whiskers extend to data points within 1.5 times the IQR, while dots represent outliers.

on chromosome 12. All SNPs passing the genome-wide significance threshold (1×10^{-3}) in both LMM and GLMM single-SNP LMM analysis are listed in Table 3. The results from the single-SNP association analysis conducted in GMMAT and GEMMA are presented together in Manhattan plots in Figure 5.

In the Bayesian association analysis, two SNPs were identified as having a major sparse effect on the FPD trait. These SNPs were estimated to have a sparse effect in at least 10% of the BSLMM chain iterations (posterior inclusion probability, PIP ≥ 0.099). Additionally, both SNPs showed a sparse effect in over 16% of the iterations (PIP ≥ 0.165), further highlighting their significance. In contrast, for the VPD trait, 75 SNPs displayed a major sparse effect in ≥10% of BSLMM chain iterations (PIP ≥ 0.095). In addition, the top four SNPs displayed a major sparse effect in more than 44% of iterations (PIP ≥ 0.447). Concerning the BOF trait, three SNPs were identified with a major sparse effect in ≥10% of iterations (PIP ≥ 0.098) and the top SNP had a major sparse effect in over 17% of iterations (PIP ≥ 0.172). Similarly, for the BOS trait, 26 SNPs exhibited a sparse effect in ≥10% of BSLMM chain iterations (PIP ≥ 0.095), with the top two SNPs showing a strong effect in over 82% of iterations (PIP > 0.829). The data outlined above is reported in Supplementary File 5.

A total of 7 SNPs passed the genome-wide significance threshold (1×10^{-3}) in the single-SNP LMM analyses and the posterior inclusion probability threshold (PIP ≥ 10%) in the Bayesian multi-SNP BSLMM analysis and are listed in Table 4. Manhattan plots from the BSLMM analysis are provided in Supplementary File 6.

TABLE 2 Spearman's correlation coefficients and *p*-values for the four *C. litardierei* phenological traits: FPD, VPD, BOF, and BOS.

Trait 1	Trait 2	Spearman's ρ	*p*-value
FPD	VPD	0.025	0.725
FPD	BOF	-0.241	0.0005
FPD	BOS	-0.069	0.324
VPD	BOF	0.430	$1.33e^{-10}$
VPD	BOS	-0.948	$< 2.2e^{-16}$
BOF	BOS	-0.241	0.0005

BOF, Beginning of Flowering; BOS, Beginning of Sprouting; FPD, Flowering Period Duration; VPD, Vegetation Period Duration.

SNP heritability estimation

The BSLMM analysis, performed using 23,315 SNPs, provided estimates of narrow-sense heritability (PVE) for the phenological traits studied, along with the proportion of genetic effect (PGE) and the count of variants with a major effect (n.gamma), as detailed in Table 5. The PVE estimate for the FPD revealed that 20.26% of the

10.3389/fpls.2025.1571608

TABLE 3 SNPs passing the genome-wide significance threshold (1×10^{-3}) in both GMMAT and GEMMA single-SNP LMM analyses for *Chouardia litardierei* traits: FPD, VPD, BOF, and BOS.

Trait	SNP	Chr	Position	Effect Allele	Referent Allele	MAF	Single-SNP LMM Analysis β (p-value) in GMMAT	Single-SNP LMM Analysis β (p-value) in GEMMA
FPD	275195_16	13	197688818	C	T	0.14	0.12 (2.41×10^{-4})	-0.52 (5.68×10^{-6})
FPD	131957_13	10	97222552	T	G	0.02	0.30 (2.79×10^{-4})	-1.13 (8.22×10^{-6})
FPD	750129_37	7	113120650	G	A	0.06	0.18 (3.24×10^{-4})	-0.76 (9.69×10^{-6})
FPD	688820_29	5	89511430	T	C	0.03	0.33 (3.68×10^{-4})	-1.39 (1.02×10^{-5})
FPD	134834_42	11	108997955	G	C	0.02	0.38 (3.69×10^{-4})	-1.49 (1.89×10^{-5})
FPD	445498_105	1	133595095	T	G	0.07	0.16 (5.24×10^{-4})	-0.66 (2.29×10^{-5})
FPD	53032_22	9	14725086	A	G	0.06	0.16 (6.05×10^{-4})	-0.68 (2.98×10^{-5})
FPD	380447_37	13	615321041	T	A	0.06	0.18 (8.73×10^{-4})	-0.77 (5.91×10^{-5})
VPD	565532_39	4	14626431	C	A	0.13	0.10 (5.23×10^{-6})	-0.46 (1.01×10^{-5})
VPD	65720_38	9	26233589	A	G	0.03	-0.14 (1.52×10^{-6})	0.60 (1.28×10^{-5})
VPD	305761_25	13	320423026	T	G	0.13	-0.08 (6.14×10^{-7})	0.33 (1.58×10^{-5})
VPD	167223_27	11	64125165	T	G	0.09	0.11 (1.62×10^{-4})	-0.46 (7.95×10^{-5})
VPD	221833_73	12	284678317	C	G	0.35	0.05 (6.52×10^{-4})	-0.22 (1.02×10^{-4})
VPD	210123_39	12	239066297	T	G	0.03	-0.13 (8.71×10^{-7})	0.51 (1.37×10^{-4})
VPD	618657_20	4	345766799	A	G	0.01	-0.23 (2.46×10^{-5})	0.93 (1.75×10^{-4})
VPD	334377_114	13	437172692	C	A	0.01	-0.14 (5.03×10^{-4})	0.75 (1.75×10^{-4})
VPD	76416_37	9	64503815	A	G	0.06	-0.08 (1.09×10^{-4})	0.35 (2.34×10^{-4})
VPD	57078_21	9	163441584	A	C	0.28	-0.06 (3.76×10^{-5})	0.23 (2.43×10^{-4})
VPD	774777_66	7	206933711	C	T	0.03	-0.14 (1.07×10^{-4})	0.56 (2.83×10^{-4})
VPD	635043_17	4	93373391	T	C	0.08	-0.09 (8.57×10^{-5})	0.37 (3.38×10^{-4})
VPD	790473_18	7	8052404	A	T	0.02	-0.20 (2.39×10^{-5})	0.81 (4.63×10^{-4})
VPD	272420_33	13	186297710	T	G	0.08	-0.10 (8.13×10^{-7})	0.33 (9.05×10^{-4})
BOF	445520_34	1	133744238	A	G	0.23	-0.19 (2.16×10^{-6})	0.41 (1.74×10^{-4})
BOF	504422_54	2	95535920	T	G	0.34	-0.18 (9.52×10^{-6})	0.49 (9.17×10^{-5})
BOF	623094_18	4	39016369	A	G	0.43	-0.14 (3.29×10^{-5})	0.30 (4.02×10^{-4})
BOF	477240_15	2	115724781	A	G	0.25	-0.15 (6.87×10^{-5})	0.34 (9.49×10^{-4})
BOF	768498_16	7	186900792	G	C	0.02	-0.33 (1.86×10^{-4})	0.89 (3.07×10^{-4})
BOF	252813_22	13	104630774	C	G	0.03	0.60 (4.04×10^{-4})	-1.26 (3.79×10^{-5})
BOF	455458_35	1	37036194	G	T	0.36	-0.28 (4.08×10^{-4})	0.59 (7.76×10^{-4})
BOF	186978_19	12	148693882	A	C	0.03	-0.30 (6.69×10^{-4})	0.81 (1.99×10^{-4})
BOS	210123_39	12	239066297	T	G	0.03	0.95 (5.88×10^{-22})	-0.64 (5.95×10^{-5})
BOS	65720_38	9	26233589	A	G	0.03	0.71 (2.04×10^{-14})	-0.64 (5.95×10^{-5})
BOS	57078_21	9	163441584	A	C	0.28	0.20 (1.19×10^{-10})	-0.61 (4.60×10^{-4})
BOS	774777_66	7	206933711	C	T	0.03	0.36 (6.64×10^{-6})	-0.27 (3.39×10^{-4})
BOS	221833_73	12	284678317	C	G	0.35	-0.12 (1.85×10^{-5})	-0.66 (5.88×10^{-4})
BOS	38821_38	8	94422071	G	A	0.27	0.12 (4.12×10^{-5})	0.27 (1.34×10^{-4})
BOS	333922_26	13	435879200	T	G	0.43	0.12 (4.58×10^{-5})	-0.24 (6.68×10^{-4})

Statistical analyses were performed with GEMMA and GMMAT LMM. p-values $< 1 \times 10^{-3}$ are considered genome-wide significant. BOS, Beginning of Sprouting; BOF, Beginning of Flowering; Chr, Chromosome; FPD, Flowering Period Duration; LMM, Linear Mixed Model; MAF, Minor Allele Frequency; SNP, Single Nucleotide Polymorphism; VPD, Vegetation Period Duration.

10.3389/fpls.2025.1571608

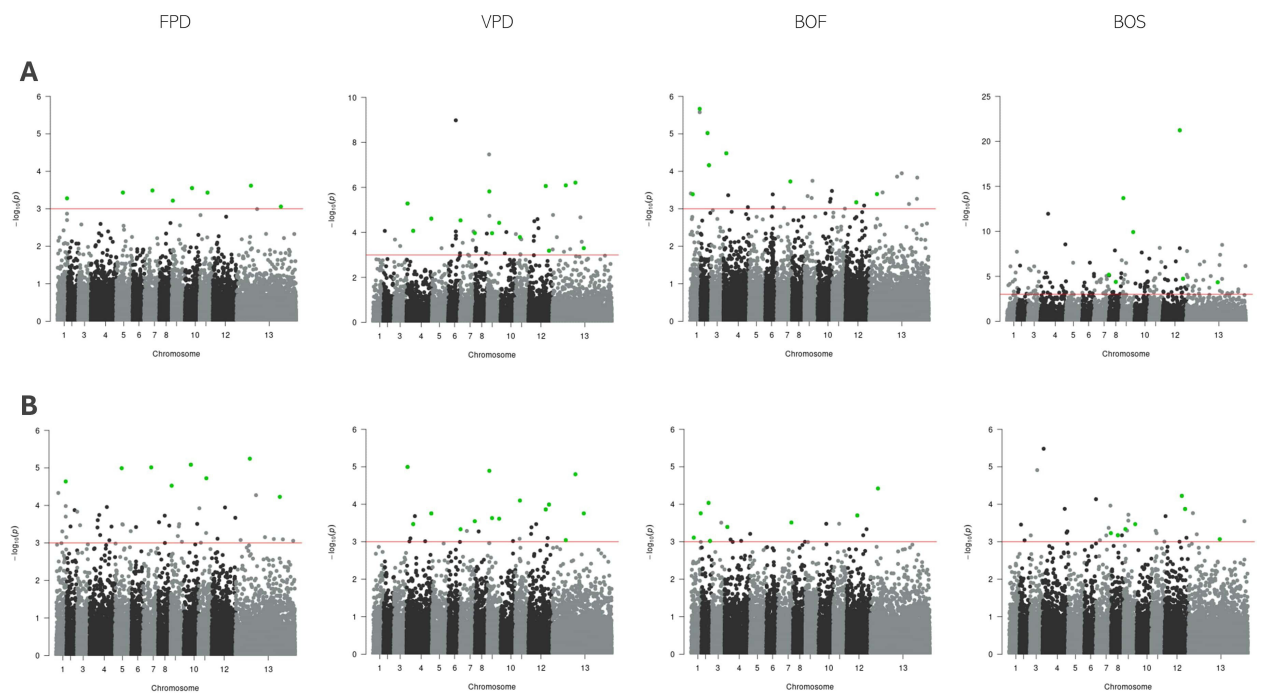


FIGURE 5
Manhattan plots of single-SNP association mapping of FPD, VPD, BOF, and BOS traits. Single-SNP analysis was conducted using **(A)** GMMAT (top row) and **(B)** GEMMA (bottom row) for each trait, where the x-axis represents the chromosomal positions of SNPs and the y-axis shows the –log10 (*p*-values) from the LMM analysis. The red horizontal line denotes the genome-wide significance threshold ($p = 1 \times 10^{-3}$). Each point on the Manhattan plot corresponds to a SNP, with stronger associations appearing higher due to lower *p*-values. Green dots indicate SNPs identified in both analyses.

phenotypic variation was explained by all available genotypes, with 47.22% attributed to 60 SNPs exhibiting significant phenotypic effects. Similarly, the PVE estimate for the VPD indicated that 86.95% of the phenotypic variation was explained by all genotypes, with 65.72% attributed to 111 SNPs exhibiting notable phenotypic effects. Moreover, the BSLMM analysis revealed that 66.03% of the phenotypic variation in BOF was explained by all genotypes, with 25.86% of this variation accounted for by 47 SNPs with significant effects. The PVE estimate for the BOS revealed that 76.05% of the phenotypic variation was explained by all available genotypes, with 63.19% attributed to 52 SNPs exhibiting significant phenotypic effects. Supplementary File 7 contains the means, medians, and 95% equal tail posterior probability intervals (95% ETPPIs) of the hyperparameters derived from the BSLMM.

TABLE 4 SNPs passing the genome-wide significance threshold (1×10^{-3}) in the single-SNP LMM analyses and the posterior inclusion probability treshold (PIP ≥ 10%) in the Bayesian multi-SNP BSLMM analysis.

Trait	SNP	Chr	Position	Effect Allele	Referent Allele	MAF	Single-SNP LMM Analysis β (*p*-value) in GMMAT	Single-SNP LMM Analysis β (*p*-value) in GEMMA	Multi-SNP BSLMM Analysis β (PIP)
FPD	131957_13	10	97222552	T	G	0.02	0.30 (2.79×10^{-4})	-1.13 (8.22×10^{-6})	-0.70 (0.17)
FPD	750129_37	7	113120650	G	A	0.06	0.18 (3.24×10^{-4})	-0.76 (9.69×10^{-6})	-0.48 (0.17)
FPD	134834_42	11	108997955	G	C	0.02	0.38 (3.69×10^{-4})	-1.49 (1.89×10^{-5})	-0.70 (0.06)
VPD	565532_39	4	14626431	C	A	0.13	0.10 (5.23×10^{-6})	-0.46 (1.01×10^{-5})	-0.33 (0.91)
VPD	210123_39	12	239066297	T	G	0.03	-0.13 (8.71×10^{-7})	0.51 (1.37×10^{-4})	0.33 (0.75)
BOF	504422_54	2	95535920	T	G	0.34	-0.18 (9.52×10^{-6})	0.49 (9.17×10^{-5})	0.32 (0.17)
BOS	210123_39	12	239066297	T	G	0.03	0.95 (5.88×10^{-22})	-0.64 (5.95×10^{-5})	-0.48 (0.83)

Statistical analyses were performed with GEMMA and GMMAT LMM and BSLMM. *p*-values< 1×10^{-3} are considered genome-wide significant. BOS, Beginning of Sprouting; BOF, Beginning of Flowering; BSLMM, Bayesian Sparse Linear Mixed Model; Chr, Chromosome; FPD, Flowering Period Duration; LMM, Linear Mixed Model; MAF, Minor Allele Frequency; PIP; Posterior Inclusion Probability; SNP, Single Nucleotide Polymorphism; VPD, Vegetation Period Duration. The table presents the single-SNP LMM *p*-values along with their corresponding posterior inclusion probabilities from the BSLMM analysis for *Chouardia litardierei* traits FPD, VPD, BOF, and BOS.

10.3389/fpls.2025.1571608

TABLE 5 Genetic architectures of *Chouardia litardierei* phenological traits identified using a BSLMM.

Trait	PVE/%	PGE/%	n.gamma
FPD	20.26	47.22	60
VPD	86.95	65.72	111
BOF	66.03	25.86	47
BOS	76.05	63.19	52

BOF, Beginning of Flowering; BOS, Beginning of Sprouting; FPD, Flowering Period Duration; n.gamma, number of variants with major effect; PGE, Proportion of Variance Explained by major effect variants; PVE, Proportion of Variance Explained by genetic data; VPD, Vegetation Period Duration.

Multivariate GWAS analysis

In the multivariate GWAS analysis, 113 SNPs surpassed the genome-wide significance threshold ($p = 1 \times 10^{-3}$) for the model with BOS and VPD traits as dependent variables (Supplementary File 8). This indicates shared genetic factors influencing these phenological traits across multivariate and univariate analyses. Five SNPs were significant in both LMM and GLMM univariate analyses for the BOS trait, and these same five were also significant for the VPD trait, along with an additional eight SNPs that were significant only for VPD, bringing the total to 13 (Table 6). In the multivariate GWAS analysis for the model with VPD and BOF traits as dependent variables, 36 SNPs exceeded the same threshold (Supplementary File 9). Among these, 10 SNPs were significant in LMM and GLMM univariate analyses for the VPD trait, while 4 showed significance for the BOF trait (Table 6). The multivariate GWAS findings for BOS and VPD, and BOF and VPD are plotted in Manhattan plots in Figure 6. The frequencies of effect alleles across populations for the significant SNPs (shown in Tables 4, 6) are depicted in a plot provided in Supplementary File 10.

GWAS candidate genes identification

The eggNOG tool provided detailed data clarifying the connection between individual SNPs/sequences and specific protein families (PFAM). To identify candidate genes potentially influencing phenological traits, we conducted eggNOG analysis on 7 SNPs that passed the genome-wide significance threshold (1×10^{-3}) in both the single-SNP LMM and multi-SNP BSLMM analyses of *C. litardierei* traits, including FPD, VPD, BOF, and BOS. This analysis identified 59 queries corresponding to sequences matched to the eggNOG database for functional annotation (Supplementary File 11). Using eggNOG, we further analyzed 13 SNPs that met the same significance threshold in the multivariate GWAS analysis of BOS and VPD, uncovering 114 additional queries (Supplementary File 12). Similarly, 14 SNPs passed the same significance threshold in the multivariate GWAS analysis of VPD and BOF, resulting in 173 additional queries (Supplementary File 13). The eggNOG analysis connected sequences to protein families, which we further explored through manual inspection and a literature review to identify specific genes and PFAM domains related to the traits being studied. Some domains were common to both the univariate and multivariate GWAS results, resulting in overlaps. The most significant findings, along with their biological functions and relevant references, are summarized in Table 7.

Discussion

This study aimed to advance our understanding of the genetic foundations of phenological adaptive traits in *C. litardierei*'s populations occupying contrasting habitats and shaped by distinct ecological pressures. To minimize the effects of phenotypic plasticity and identify heritable local adaptation traits as accurately as possible, individuals from divergent environments were grown under uniform conditions (Liu and El-Kassaby, 2019; Schwinning et al., 2022). This approach allowed the separation of genetic influences from environmental effects, revealing the heritable components driving local adaptation, where populations evolve toward optimal phenotypic and genetic configurations in response to local selective pressures (Montejo-Kovacevich et al., 2021).

Basic statistical analyses on the common garden experiment data were first performed to characterize the variations within the tested phenological traits and their potential importance for the local adaptation of studied populations in their natural habitats. Except for the duration of flowering (FPD), substantial variations in tested traits among the studied populations were revealed, highlighting their importance for adaptation to contrasting environmental pressures. However, although significant differences in phenological traits among studied groups and individual populations were present, there were many exceptions in the general pattern. For instance, although the dolomite-habitat population group began with flowering (BOF) before the remaining two groups, the Pag population from the seashore habitat was an exception, as it overlaped with all the dolomite-habitat populations. At the same time, the Pag population came into flowering significantly earlier than the Vrana Lake population, which is found in the same habitat and is even geographically closely positioned to the Pag population. Similarly, VPD was significantly shorter in the dolomite-habitat group of populations in contrast to other groups; however, the Budoške Bare population from karst poljes' meadow habitat joined this group due to having VPD also significantly shorter than any of the remaining populations from this and the seashore habitat. Such a result supports the earlier assumption that although groups of *C. litardierei* population thrive in highly contrasting habitats, their differentiation into well-differentiated ecotypes remains poorly supported. This was also partially confirmed by the obtained population genetic results (Supplementary Files 3, 4). Here, only the group of populations from the dolomite habitat was substantially differentiated and formed a well-defined genetic cluster, while all the remaining populations remained clustered together, without signs of differentiation between the seashore and meadow-habitat groups. Since the ecotypes are defined as groups of populations whose differentiation is supported both genetically and phenotypically (Lowry, 2012), the studied groups do not meet these criteria. Nonetheless, some trends can be observed in the obtained results that point to certain conclusions. The dolomite-habitat population

10.3389/fpls.2025.1571608

TABLE 6 SNPs passing genome-wide significance threshold (1×10^{-3}) in the multivariate GWAS mvLMM analysis of *Chouardia litardierei* phenological traits BOS and VPD, and BOF and VPD.

Trait	SNP	Chr	Position	Effect Allele	Ref. Allele	MAF	Beta1 (VPD)	Beta2 (BOS)	mvLMM in GEMMA (*p*-value)
BOS +VPD	305761_25	13	320423026	T	G	0.13	0.35	-0.37	7.85×10^{-6}
BOS +VPD	65720_38	9	26233589	A	G	0.03	0.61	-0.59	1.25×10^{-5}
BOS +VPD	565532_39	4	14626431	C	A	0.13	-0.45	0.55	1.73×10^{-5}
BOS +VPD	221833_73	12	284678317	C	G	0.35	-0.23	0.30	3.20×10^{-5}
BOS +VPD	334377_114	13	437172692	C	A	0.01	0.75	-0.62	9.93×10^{-5}
BOS +VPD	167223_27	11	64125165	T	G	0.09	-0.46	0.49	2.14×10^{-4}
BOS +VPD	790473_18	7	8052404	A	T	0.02	0.85	-1.13	3.34×10^{-4}
BOS +VPD	210123_39	12	239066297	T	G	0.03	0.51	-0.58	4.28×10^{-4}
BOS +VPD	774777_66	7	206933711	C	T	0.03	0.57	-0.72	5.10×10^{-4}
BOS +VPD	618657_20	4	345766799	A	G	0.01	0.93	-1.05	6.71×10^{-4}
BOS +VPD	57078_21	9	163441584	A	C	0.28	0.22	-0.21	6.98×10^{-4}
BOS +VPD	76416_37	9	64503815	A	G	0.06	0.35	-0.38	7.19×10^{-4}
BOS +VPD	635043_17	4	93373391	T	C	0.08	0.37	-0.44	8.20×10^{-4}
							Beta1 (VPD)	Beta2 (BOF)	
BOF +VPD	65720_38	9	26233589	A	G	0.03	0.59	0.25	8.71×10^{-6}
BOF +VPD	305761_25	13	320423026	T	G	0.13	0.34	0.01	2.04×10^{-5}
BOF +VPD	565532_39	4	14626431	C	A	0.13	-0.45	0.15	3.26×10^{-5}
BOF +VPD	334377_114	13	437172692	C	A	0.10	0.73	0.49	7.39×10^{-5}
BOF +VPD	774777_66	7	206933711	C	T	0.03	0.61	-0.27	2.41×10^{-4}
BOF +VPD	167223_27	11	64125165	T	G	0.09	-0.46	0.16	2.74×10^{-4}
BOF +VPD	221833_73	12	284678317	C	G	0.35	-0.22	0.11	3.14×10^{-4}
BOF +VPD	618657_20	4	345766799	A	G	0.01	0.95	-0.28	4.52×10^{-4}
BOF +VPD	76416_37	9	64503815	A	G	0.06	0.35	0.06	6.51×10^{-4}
BOF +VPD	210123_39	12	239066297	T	G	0.03	0.50	-0.01	6.84×10^{-4}
BOF +VPD	504422_54	2	95535920	T	G	0.34	0.23	0.49	2.16×10^{-5}
BOF +VPD	252813_22	13	104630774	C	G	0.03	0.10	-1.27	1.03×10^{-4}
BOF +VPD	186978_19	12	148693882	A	C	0.03	-0.22	0.82	4.17×10^{-4}
BOF +VPD	445520_34	1	133744238	A	G	0.23	0.06	0.39	5.72×10^{-4}

Statistical analyses were performed with GEMMA mvLMM. *p*-values$< 1 \times 10^{-3}$ are considered genome-wide significant. BOF, Beginning of Flowering; BOS, Beginning of Sprouting; Chr, Chromosome; FPD, MAF, Minor Allele Frequency; mvLMM, multivariate Linear Mixed Model; SNP, Single Nucleotide Polymorphism; VPD, Vegetation Period Duration. Listed SNPs were found to be significant in both GEMMA and GMMAT univariate analyses.

group sprouted later (BOS) and flowered earlier (BOF) but had a shorter vegetation period (VPD) than the remaining two groups. Such a shift in phenophases is likely to be significantly influenced by habitat properties. To understand how this specific habitat may affect this phenomenon, two aspects must be considered: the dolomite substrate properties and the influence of the local climate dynamics on the vegetation season. These southernmost populations of *C. litardierei* are usually found on bare dolomite bedrock or less frequently in dry, exposed mountainous grassland habitats developed on very shallow rendzina soils (Figure 1). Due to reduced water and nutrient capacity, accompanied by high levels of thermal conductivity and thermal capacity of the dolomite substrate (Thomas et al., 1973; Waples and Waples, 2004; Mota et al., 2021), these drought-prone habitats are known to induce heat stress in adjacent organisms and thus present a hostile environment for plant species (Mota et al., 2021). In addition, due to very sparse vegetation cover in such habitats, the substrate temperature can be expected to reach far greater values when compared to habitats covered with

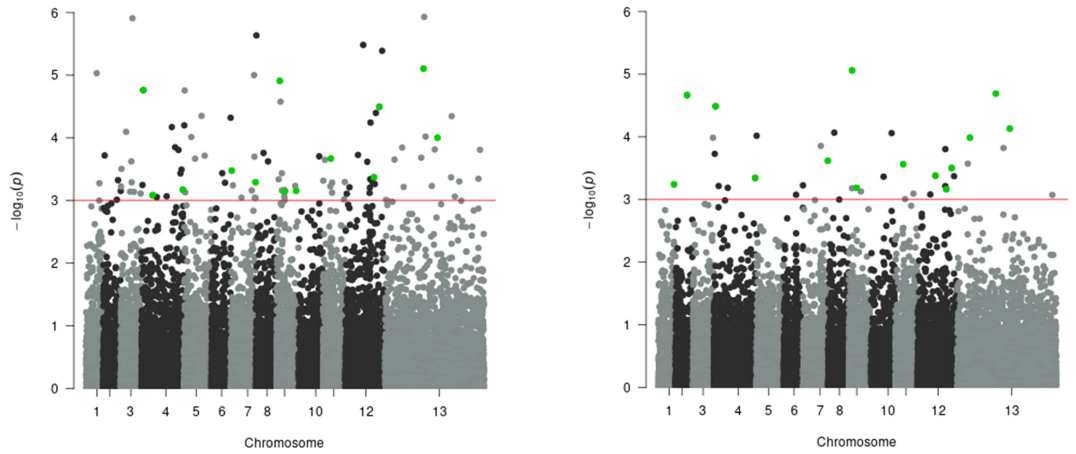


FIGURE 6
Manhattan plot of multivariate genome-wide association study (multi-GWAS) of **(A)** BOS and VPD (left) and **(B)** BOF and VPD traits (right). The red horizontal line indicates the genome-wide significance threshold ($p = 1 \times 10^{-3}$). Each dot on the Manhattan plot signifies a SNP. The strongest associations have the smallest *p*-values, so their negative logarithms will be the greatest, appearing higher on the plot. Green dots indicate SNPs identified as significant in the multivariate GWAS analysis as well as in both GEMMA and GMMAT univariate analyses for each of the two plots.

canopies or meadows (Oliver et al., 1987), thus further worsening already inhospitable conditions. Regarding the influence of regional climatic patterns on local vegetation, two peaks of ecosystem productivity have been observed in Mediterranean climate conditions across southern Europe – the larger one during spring and the less pronounced one during autumn. Such a modality has developed because of ecological constraints imposed by low winter temperatures on one side and summer droughts on another (Spano et al., 2013; Camarero et al., 2021), thus leaving relatively short time frames in spring and autumn suitable for development and reproduction. Consequently, it seems plausible that populations experiencing such climatic patterns, in combination with drought- and heat-stress-prone habitats, have developed short development- and reproduction-related phenophases. At the same time, the remaining *C. litardierei* populations inhabiting deep, moisture-retaining soils protected by dense vegetation layer which additionally reduces the increase of substrate temperature (Oliver et al., 1987), experience a less limited time frame for closing the sexual reproduction cycle. This is reflected in significant shifts in related phenophases toward later sprouting and the beginning of flowering, as well as a more extended vegetation period.

By emphasizing their heritable nature, the high PVE values observed in our study were suggested to indicate the great evolutionary importance of detected candidate loci in shaping the phenological adaptation of populations to local climatic conditions. The highest PVE value (86.95%) was exhibited by the trait VPD, suggesting that the length of the growing season in this species is predominantly determined by genetic factors. The high genetic variance observed in VPD could be reflective of adaptive mechanisms that allow *C. litardierei* to optimize its growth and reproductive success in response to environmental cues, such as climate and soil conditions, with strong natural selection acting on traits critical for survival in fluctuating environments. While a PVE for flowering time exceeding 95% has been reported for *Arabidopsis* from Cape Verde and Morocco (Neto and Hancock, 2023), highlighting the predominant genetic influence, the PVE for *C. litardierei* flowering period duration (FPD) was found to be 20.26%, indicating a more significant role of environmental or non-genetic factors. PVE values of 66.03% and 76.05% were exhibited by the BOF and BOS traits, respectively, indicating that genetic elements were exerting a greater influence than local environmental factors in shaping these traits. This was reinforced by the PGE values, with the highest PGE (65.72%) being observed in VPD, driven by a few major variants. In contrast, lower PGE values (25.86% and 63.19%) were found for BOF and BOS, respectively, reflecting the influence of numerous small-effect variants and a greater environmental impact. Overall, these heritability estimates and genetic findings provided evidence of the significant role played by genetic factors in shaping phenological traits in *C. litardierei*, emphasizing the complex interaction between genetics and environment and offering a strong foundation for future genetic, evolutionary, and adaptation studies.

In this study, multiple loci linked to phenological traits in *C. litardierei* were identified through univariate and multivariate GWAS approaches. The relatively low overlap of significant SNPs detected across the different GWAS models likely reflects inherent differences in their statistical assumptions and approaches to modelling genetic effects. While both frequentist methods (GLMM and LMM) applied a consistent significance threshold of $< 1 \times 10^{-3}$, the BSLMM relies on posterior inclusion probabilities, which are generally more conservative and not directly comparable to *p*-values. Importantly, each model is optimized for different data characteristics: LMM assumes normally distributed traits, whereas GLMM, using a Poisson distribution, is more appropriate for count-based traits with non-normal distributions. Applying trait-appropriate models increases the reliability and power of association detection, even if it results in a lower number of shared SNPs. Functional annotation of the genomic windows surrounding significant SNP loci revealed

10.3389/fpls.2025.1571608

TABLE 7 List of candidate genes for regions of strong association with FPD, VPD, BOF and BOS identified by the eggNOG-mapper v2 database.

Query	Method	*e*-value	Chr	EGGNog PFAM	Candidate Genes	Species	Relevant biological functions	References
H 9:113095650-113145650_6	GWAS	$4.55e^{-212}$	7	Chromo domain	LHP1	*Arabidopsis thaliana*	Chromatin regulation and flowering time control	Gaudin et al. (2001), Adrian et al. (2010)
H 9:206908711-206958711_54	mGWAS1	$6.8e^{-111}$	13					
H 2:95510920-95560920_36	mGWAS2	$3.82e^{-132}$	2					
H 9:113095650-113145650_4	GWAS	$3.1e^{-163}$	7	Histidine phosphatase protein family	Hd3a, ZCN8	*Oryza sativa*, *Zea mays*	Hormone signalling, development, stress response, and flowering	Cho et al. (2022)
H 9:113095650-113145650_44	GWAS	$2.31e^{-42}$	7	Aspartic Proteases (APs)	PvAP1	*Phaseolus vulgaris*	Drought stress adaptation and osmotic resistance	Contour-Ansel et al. (2010)
H 4:14601431-14651431_21	GWAS	$2.19e^{-80}$	4	CCHC-type zinc finger proteins (CCHC-ZFPs)	AtCSP4	*Arabidopsis thaliana*	Growth, development, and stress responses	Yang and Karlson (2011)
H 15:284653317-284703317_45	mGWAS1	$1.77e^{-306}$	12					
H 15:284653317-284703317_50	mGWAS2	$1.77e^{-306}$	12					
H 4:14601431-14651431_5	GWAS	$1.14e^{-23}$	4	Pentatricopeptide repeat (PPR) proteins	AT1G15480	*Arabidopsis thaliana*	Flowering time regulation	Emami and Kempken (2019)
H 4:14601431-14651431_4	mGWAS2	$9.61e^{-26}$	4					
H 14:108972955-109022955_19	GWAS	$2.14e^{-100}$	11	Phytochrome-interacting Factor 1 (PIF1)	PIF1	*Arabidopsis thaliana*	Sprouting control, growth, stress adaptation, and photosynthesis regulation	Yadav (2024), Soy et al. (2014), Li et al. (2024), Chen et al. (2013)
H 14:108972955-109022955_22	mGWAS1	$1.00e^{-308}$	11					
H 16:320398026-320448026_57	mGWAS1	$3.11e^{-22}$	13	MATE domain	OsMATE2, OsMATE4, OsMATE42, OsMATE46	*Oryza sativa*	Early salt stress response and drought stress resistance	Du et al. (2021)
H 16:320398026-320448026_58	mGWAS2	$3.11e^{-22}$	13					
H 15:284653317-284703317_6	mGWAS1	$1.08e^{-12}$	12	Protein kinase domain	SOS2/CIPK24	*Arabidopsis thaliana*	Salt stress responses and hormonal signaling	Chen et al. (2023)
H 4:345741799-345791799_9	mGWAS2	$2.09e^{-24}$	4					
H 1:133719238-133769238_38	mGWAS2	$1.33e^{-36}$	1	MLO protein family	OsMLO1-4, OsMLO9, OsMLO11	*Oryza sativa*	Heat and/or cold stress response	Nguyen et al. (2016)
H 15:148668882-148718882_58	mGWAS2	$5.59e^{-190}$	12	C2 domain	QUIRKY, STRUBBELIG	*Arabidopsis thaliana*	Promotes intercellular communication and tissue morphogenesis	Vaddepalli et al. (2014)

BOF, Beginning of Flowering; Chr, chromosome; BOS, Beginning of Sprouting; flowering period duration; GWAS, genome-wide association study; H, HiC scaffold; mGWAS1, multivariate Genome-Wide Association Study of BOS and VPD; mGWAS2, multivariate Genome-Wide Association Study of BOF and VPD; PFAM, protein family; VPD, Vegetation Period Duration. (*e*-value< 1×10^{-2}) in *Chouardia litardierei* based on the 7 recognized SNPs passing genome-wide significance threshold (1×10^{-3}) in the single-SNP LMM and multi-SNP BSLMM analysis as well as 13 SNPs passing the same threshold in the multivariate GWAS mvLMM analysis of BOS and VPD (mGWAS1), and 14 SNPs in BOF and VPD (mGWAS2). The names of the identified candidate genes associated with the SNPs, PFAMs, their relevant biological functions, and corresponding references are provided.

regions encoding key protein families involved in essential biological pathways related to phenological events. Among others, SNP loci were identified in regions encoding the chromo domain, which is crucial to plant chromatin-based gene regulation. In *Arabidopsis*, mutations in LHP1 (LIKE HETEROCHROMATIN PROTEIN 1), a gene encoding a chromo domain, have been shown to cause early flowering and reduced plant size (Gaudin et al., 2001). Overexpression of CONSTANS (CO), which activates FLOWERING LOCUS T (FT) in long-day conditions, has been found to alter chromatin at the FT locus by reducing LHP1 binding and increasing histone acetylation, suggesting LHP1 represses flowering through chromatin regulation (Adrian et al., 2010). SNP loci were also identified in regions encoding histidine phosphatase proteins, which are known to regulate plant development and stress responses, particularly through hormone signaling pathways like cytokinins that influence flowering and vegetative growth (Werner et al., 2001; Hai et al., 2020). For instance, it has been demonstrated that exogenous cytokinin

application extends the vegetative phase in rice and maize by inhibiting the expression of florigen genes, such as Hd3a and ZCN8, thus delaying flowering time (Cho et al., 2022). Additionally, cytokinins have been found to interact with environmental signals like nutrient sensing (Argueso et al., 2009; Prasad, 2022), potentially aiding plant adaptation to nutrient-poor and drought-prone habitats, like those inhabited by the southern group of *C. litardierei* populations. Similarly, cytokinin-deficient mutants have been observed to exhibit delayed flowering on nutrient-poor substrates, underscoring cytokinin's role in adaptation to nutrient-limited environments (Miyawaki et al., 2006). SNP loci within the genomic regions encoding aspartic proteases (APs) and CCHC-type zinc finger proteins (CCHC-ZFPs) were recognized as well. In drought-susceptible common bean cultivars, the PvAP1 gene exhibited significant upregulation under mild water stress, supporting the role of APs in drought responses (Contour-Ansel et al., 2010). CCHC-ZFPs are considered essential for growth and development, as demonstrated in *Arabidopsis*, where AtCSP4 has been identified as a key factor (Yang and Karlson, 2011). Additionally, SNP loci within the genomic region encoding pentatricopeptide repeat (PPR) proteins were identified. It has been reported that mutations in the *Arabidopsis* gene AT1G15480, encoding a P-class PPR protein, result in early flowering (Emami and Kempken, 2019). Furthermore, mutations were detected in genetic regions responsible for encoding phytochrome-interacting factor 1 (PIF1). In *Arabidopsis*, PIF1 has been found to play a major role in sprouting inhibition (Oh et al., 2004; Yadav, 2024), plant growth and development regulation (Soy et al., 2014), stress adaptation (Li et al., 2024), and regulation of photosynthesis initiation (Chen et al., 2013). In addition, SNP loci were identified within regions encoding the MATE domain, the protein kinase domain, and loci associated with the MLO protein family. Several MATE domain genes in *O. sativa* (OsMATE2, OsMATE4, OsMATE42, and OsMATE46) have been shown to regulate plant responses to abiotic stresses, such as salt and drought, through differential expression patterns (Du et al., 2021), while the protein kinase SOS2/CIPK24 has been recognized as a central regulator of salt stress response and hormonal signaling in *Arabidopsis* (Chen et al., 2023). Finally, the MLO protein family is considered crucial for temperature stress adaptation, as exemplified by several OsMLO proteins in *O. sativa* (Nguyen et al., 2016).

Here, we investigated the genetic background of phenological traits in *C. litardierei*, revealing significant associations between them and specific genetic variations across the genome. Our findings indicate that certain genomic regions may be instrumental in the adaptive responses of populations to contrasting environmental conditions. The genetic architecture of these phenological traits is complex, with multiple candidate loci contributing to phenotypic diversity across habitats. Using the ddRAD-seq approach and comprehensive GWAS analyses, we identified key candidate genes and multiple loci associated with phenological traits. However, the limited genome scan resolution of ddRAD-seq, particularly in large genomes like *C. litardierei* (3.7 Gb), leaves much genomic information unexplored. The relatively small sample size is a limitation of our study, particularly given that GWAS typically include larger cohorts to detect robust and reproducible associations. Nevertheless, our analysis revealed several biologically plausible signals, which, while requiring validation, provide a valuable foundation for future studies. These findings should be interpreted with caution, but they offer meaningful insights that can be further explored and confirmed in larger, independent populations. Functional annotation of the associated genomic regions revealed key protein families involved in vital biological pathways related to flowering time, vegetative growth, and stress adaptation. These protein families are crucial regulators of plant development, environmental responses, and abiotic stress adaptation. High narrow-sense heritability estimates indicated that genetic factors accounted for a significant portion of the phenotypic variance, with PVE ranging from 20.26% for flowering period duration (FPD) to 86.95% for vegetation period duration (VPD). This study underscores the complexity of the genetic architecture driving phenotypic diversity in plants, highlighting the critical role of genomic approaches in examining adaptive traits in non-model species exposed to diverse ecological pressures. Despite challenges in studying a wild, non-model species, this research advances our understanding of the genomic basis of adaptive divergence and ecological differentiation in *C. litardierei*. Expanding this research through a comprehensive Genome-Environment Association (GEA) study, incorporating more populations across the species' distribution range, could provide deeper insights into the genomic drivers of local adaptation and phenological divergence.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA1164649.

Author contributions

SLŠ: Conceptualization, Formal Analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. NP: Conceptualization, Formal Analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. KK: Methodology, Writing – original draft, Writing – review & editing. BS: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. DM: Visualization, Writing – original draft, Writing – review & editing. IR: Conceptualization, Funding acquisition, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was funded by the Croatian Science Foundation. All the research activities were

financed by the "Amethyst Meadow Squill (Chouardia litardierei, Hyacinthaceae): a study system for ecological divergence" (HRZZ-IP-2020-02-8099) project.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2025.1571608/full#supplementary-material

References

Adole, T., Dash, J., Rodriguez-Galiano, V., and Atkinson, P. M. (2019). Photoperiod controls vegetation phenology across Africa. *Commun. Biol.* 2, 1–13. doi: 10.1038/s42003-019-0636-7

Adrian, J., Farrona, S., Reimer, J. J., Albani, M. C., Coupland, G., and Turck, F. (2010). cis-regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in Arabidopsis. *Plant Cell* 22, 1425–1440. doi: 10.1105/TPC.110.074682

Anderson, J. T., Inouye, D. W., McKinney, A. M., Colautti, R. I., and Mitchell-Olds, T. (2012). Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. *Proc. R. Soc. B: Biol. Sci.* 279, 3843–3852. doi: 10.1098/RSPB.2012.1051

Argueso, C. T., Ferreira, F. J., and Kieber, J. J. (2009). Environmental perception avenues: the interaction of cytokinin and environmental response pathways. *Plant Cell Environ.* 32, 1147–1160. doi: 10.1111/J.1365-3040.2009.01940.X

Bakhtiari, M., Formenti, L., Caggìa, V., Glauser, G., and Rasmann, S. (2019). Variable effects on growth and defense traits for plant ecotypic differentiation and phenotypic plasticity along elevation gradients. *Ecol. Evol.* 9, 3740–3755. doi: 10.1002/ECE3.4999

Bernatchez, L., Ferchaud, A. L., Berger, C. S., Venney, C. J., and Xuereb, A. (2023). Genomics for monitoring and understanding species responses to global climate change. *Nat. Rev. Genet.* 25, 165–183. doi: 10.1038/s41576-023-00657-y

Besnier, F., and Glover, K. A. (2013). ParallelStructure: A R package to distribute parallel runs of the population genetics program STRUCTURE on multi-core computers. *PloS One* 8, e70651. doi: 10.1371/JOURNAL.PONE.0070651

Bonacci, O. (2014). "Ecohydrology of karst poljes and their vulnerability," in *Dinaric karst poljes - floods for life* (EuroNatur, Radolfzell am Bodensee), 25–37.

Brandrud, M. K., Paun, O., Lorenzo, M. T., Nordal, I., and Brysting, A. K. (2017). RADseq provides evidence for parallel ecotypic divergence in the autotetraploid Cochlearia officinalis in Northern Norway. *Sci. Rep.* 7, 1–13. doi: 10.1038/s41598-017-05794-z

Bremer, B., Bremer, K., Chase, M. W., Fay, M. F., Reveal, J. L., Bailey, L. H., et al. (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* 161, 105–121. doi: 10.1111/J.1095-8339.2009.00996.X

Bui, E. N. (2013). Soil salinity: A neglected factor in plant ecology and biogeography. *J. Arid. Environ.* 92, 14–25. doi: 10.1016/J.JARIDENV.2012.12.014

Camarero, J. J., Valeriano, C., Gazol, A., Colangelo, M., and Sánchez-Salguero, R. (2021). Climate differently impacts the growth of coexisting trees and shrubs under semi-arid Mediterranean conditions. *Forests* 12, 381. doi: 10.3390/F12030381

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., and Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22, 3124–3140. doi: 10.1111/MEC.12354

Chen, C., He, G., Li, J., Perez-Hormaeche, J., Becker, T., Luo, M., et al. (2023). A salt stress-activated GSO1-SOS2-SOS1 module protects the Arabidopsis root stem cell niche by enhancing sodium ion extrusion. *EMBO J.* 42, e113004. doi: 10.15252/EMBJ.2022113004

Chen, H., Huffman, J. E., Brody, J. A., Wang, C., Lee, S., Li, Z., et al. (2019). Efficient variant set mixed model association tests for continuous and binary traits in large-scale whole-genome sequencing studies. *Am. J. Hum. Genet.* 104, 260–274. doi: 10.1016/j.ajhg.2018.12.012

Chen, D., Xu, G., Tang, W., Jing, Y., Ji, Q., Fei, Z., et al. (2013). Antagonistic basic helix-loop-helix/bZIP transcription factors form transcriptional modules that integrate light and reactive oxygen species signaling in Arabidopsis. *Plant Cell* 25, 1657–1673. doi: 10.1105/TPC.112.104869

Cho, L. H., Yoon, J., Tun, W., Baek, G., Peng, X., Hong, W. J., et al. (2022). Cytokinin increases vegetative growth period by suppressing florigen expression in rice and maize. *Plant J.* 110, 1619–1635. doi: 10.1111/TPJ.15760

Contour-Ansel, D., Torres-Franklin, M. L., Zuily-Fodil, Y., and de Carvalho, M. H. C. (2010). An aspartic acid protease from common bean is expressed 'on call' during water stress and early recovery. *J. Plant Physiol.* 167, 1606–1612. doi: 10.1016/J.JPLPH.2010.06.018

Cook, B. I., Wolkovich, E. M., Davies, T. J., Ault, T. R., Betancourt, J. L., Allen, J. M., et al. (2012). Sensitivity of spring phenology to warming across temporal and spatial climate gradients in two independent databases. *Ecosystems* 15, 1283–1294. doi: 10.1007/S10021-012-9584-5

Cortés, A. J., Garzón, L. N., Valencia, J. B., and Madriñán, S. (2018). On the causes of rapid diversification in the páramos: Isolation by ecology and genomic divergence in espeletia. *Front. Plant Sci.* 871. doi: 10.3389/FPLS.2018.01700/BIBTEX

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10, 1–4. doi: 10.1093/GIGASCIENCE/GIAB008

Du, Z., Su, Q., Wu, Z., Huang, Z., Bao, J., Li, J., et al. (2021). Genome-wide characterization of MATE gene family and expression profiles in response to abiotic stresses in rice (Oryza sativa). *BMC Ecol. Evol.* 21, 1–14. doi: 10.1186/S12862-021-01873-Y/FIGURES/7

Duputié, A., Rutschmann, A., Ronce, O., and Chuine, I. (2015). Phenological plasticity will not help all species adapt to climate change. *Glob. Chang. Biol.* 21, 3062–3073. doi: 10.1111/GCB.12914

Emami, H., and Kempken, F. (2019). PRECOCIOUS1 (POCO1), a mitochondrial pentatricopeptide repeat protein affects flowering time in Arabidopsis thaliana. *Plant J.* 100, 265–278. doi: 10.1111/TPJ.14441

Engelmann, K., and Purugganan, M. (2006). The molecular evolutionary ecology of plant development: flowering time in Arabidopsis thaliana. *Adv. Bot. Res.* 44, 507–526. doi: 10.1016/S0065-2296(06)44013-1

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/J.1365-294X.2005.02553.X

Feder, J. L., Gejji, R., Powell, T. H. Q., and Nosil, P. (2011). Adaptive chromosomal divergence driven by mixed geographic mode of evolution. *Evolution* 65, 2157–2170. doi: 10.1111/J.1558-5646.2011.01321.X

Fernández-Meirama, M., Rolán-Alvarez, E., and Carvajal-Rodríguez, A. (2022). A simulation study of the ecological speciation conditions in the galician marine snail Littorina saxatilis. *Front. Genet.* 13. doi: 10.3389/FGENE.2022.680792/BIBTEX

Flohr, B. M., Hunt, J. R., Kirkegaard, J. A., and Evans, J. R. (2017). Water and temperature stress define the optimal flowering period for wheat in south-eastern Australia. *Field Crops Res.* 209, 108–119. doi: 10.1016/J.FCR.2017.04.012

Gaudin, V., Libault, M., Pouteau, S., Juul, T., Zhao, G., Lefebvre, D., et al. (2001). Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in Arabidopsis. *Development* 128, 4847–4858. doi: 10.1242/DEV.128.23.4847

Gaudinier, A., and Blackman, B. K. (2020). Evolutionary processes from the perspective of flowering time diversity. *New Phytol.* 225, 1883–1898. doi: 10.1111/NPH.16205

Gaži-Baskova, V. (1962). Geografsko raširenje livadnog procjepka ili lučike (Scilla pratensis W. et K.) *Biološki Glas* 15, 49–54.

Goldblatt, P., and Manning, J. C. (1996). Phylogeny and speciation in Lapeirousia subgenus Lapeirousia (Iridaceae: Ixioideae). *Ann. Missouri Bot. Garden* 83, 346–361. doi: 10.2307/2399865

Grant, V. (1981). *Plant Speciation. Second edition*. New York, NY: Columbia University Press 432.

Hai, N. N., Chuong, N. N., Tu, N. H. C., Kisiala, A., Hoang, X. L. T., and Thao, N. P. (2020). Role and regulation of cytokinins in plant response to drought stress. *Plants* 9, 422. doi: 10.3390/PLANTS9040422

Hammer, D. A. T., Ryan, P. D., Hammer, Ø., and Harper, D. A. T. (2001). *Past: Paleontological Statistics Software Package for Education and Data Analysis*. Available online at: http://palaeo-electronica.org/2001_1/past/issue1_01.html (Accessed February 3, 2025).

Heslop-Harrison, J. (1964). Forty years of genecology. *Adv. Ecol. Res.* 2, 159–247. doi: 10.1016/S0065-2504(08)60332-3

Hill, C. B., and Li, C. (2016). Genetic architecture of flowering phenology in cereals and opportunities for crop improvement. *Front. Plant Sci.* 7. doi: 10.3389/FPLS.2016.01906/BIBTEX

Hu, Z. M., Zhong, K., Weinberger, F., Duan, D. L., Draisma, S. G. A., and Serrão, E. A. (2020). Linking ecology to genetics to better understand adaptation and evolution: A review in marine macrophytes. *Front. Mar. Sci.* 7. doi: 10.3389/FMARS.2020.545102/BIBTEX

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314. doi: 10.1093/NAR/GKY1085

Johnson, S. D., Linder, H. P., and Steiner, K. E. (1998). Phylogeny and radiation of pollination systems in Disa (Orchidaceae). *Am. J. Bot.* 85, 402–411. doi: 10.2307/2446333

Kinmonth-Schultz, H., Lewandowska-Sabat, A., Imaizumi, T., Ward, J. K., Rognli, O. A., and Fjellheim, S. (2021). Flowering times of wild Arabidopsis accessions from across Norway correlate with expression levels of FT, CO, and FLC genes. *Front. Plant Sci.* 12. doi: 10.3389/FPLS.2021.747740/BIBTEX

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., and Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* 15, 1179–1191. doi: 10.1111/1755-0998.12387

Köppen, W. (1918). Klassifikation der Klimate nach Temperatur, Niederschlag und Jahreslauf. *Petermanns Geogr. Mitt* 64, 193–203.

Lee, Z., Kim, S., Choi, S. J., Joung, E., Kwon, M., Park, H. J., et al. (2023). Regulation of flowering time by environmental factors in plants. *Plants* 12, 3680. doi: 10.3390/PLANTS12213680

Levin, D. A. (2000). *The Origin, Expansion, and Demise of Plant Species*. New York, NY: Oxford University Press. doi: 10.1093/OSO/9780195127287.001.0001

Levin, D. A. (2006). Flowering phenology in relation to adaptive radiation. *Syst. Bot.* 31, 239–246. doi: 10.1600/036364406777585928

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/BIOINFORMATICS/BTP324

Li, Y. L., and Liu, J. X. (2018). StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol. Ecol. Resour.* 18, 176–177. doi: 10.1111/1755-0998.12719

Li, Z. Y., Ma, N., Zhang, F. J., Li, L. Z., Li, H. J., Wang, X. F., et al. (2024). Functions of phytochrome interacting factors (PIFs) in adapting plants to biotic and abiotic stresses. *Int. J. Mol. Sci.* 25, 2198. doi: 10.3390/IJMS25042198

Li, K., Wang, Y., Han, C., Zhang, W., Jia, H., and Li, X. (2007). GA signaling and CO/FT regulatory module mediate salt-induced late flowering in Arabidopsis thaliana. *Plant Growth Regul.* 53, 195–206. doi: 10.1007/S10725-007-9218-7/FIGURES/6

Liu, Y., and El-Kassaby, Y. A. (2019). Phenotypic plasticity of natural Populus trichocarpa populations in response to temporally environmental change in a common garden. *BMC Evol. Biol.* 19, 1–15. doi: 10.1186/S12862-019-1553-6/FIGURES/4

Lowry, D. B. (2012). Ecotypes and the controversy over stages in the formation of new species. *Biol. J. Linn. Soc.* 106, 241–257. doi: 10.1111/J.1095-8312.2012.01867.X

Lowry, D. B., Modliszewski, J. L., Wright, K. M., Wu, C. A., and Willis, J. H. (2008). Review. The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 363, 3009–3021. doi: 10.1098/RSTB.2008.0064

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/EJ.17.1.200

Mckown, A. D., Klápště, J., Guy, R. D., Geraldes, A., Porth, I., Hannemann, J., et al. (2014). Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of Populus trichocarpa. *New Phytol.* 203, 535–553. doi: 10.1111/NPH.12815

Mihevc, A., Prelovšek, M., and Zupan Hajna, N. (2010). *Introduction to the Dinaric Karst*. Postojna: Karst Research Institute at ZRC SAZU. doi: 10.3986/9789612541989

Miyawaki, K., Tarkowski, P., Matsumoto-Kitano, M., Kato, T., Sato, S., Tarkowska, D., et al. (2006). Roles of Arabidopsis ATP/ADP isopentenyltransferases and tRNA isopentenyltransferases in cytokinin biosynthesis. *Proc. Natl. Acad. Sci. U S A* 103, 16598–16603. doi: 10.1073/PNAS.0603522103

Molla, K. A. (2022). Flowering time and photoperiod sensitivity in rice: Key players and their interactions identified. *Plant Cell* 34, 3489–3490. doi: 10.1093/PLCELL/KOAC230

Montejo-Kovacevich, G., Salazar, P. A., Smith, S. H., Gavilanes, K., Bacquet, C. N., Chan, Y. F., et al. (2021). Genomics of altitude-associated wing shape in two tropical butterflies. *Mol. Ecol.* 30, 6387–6402. doi: 10.1111/MEC.16067

Mota, J., Merlo, E., Martínez-Hernández, F., Mendoza-Fernández, A. J., Pérez-García, F. J., and Salmerón-Sánchez, E. (2021). Plants on rich-magnesium dolomite barrens: A global phenomenon. *Biology* 10, 38. doi: 10.3390/BIOLOGY10010038

Neto, C., and Hancock, A. (2023). Genetic architecture of flowering time differs between populations with contrasting demographic and selective histories. *Mol. Biol. Evol.* 40. doi: 10.1093/MOLBEV/MSAD185

Nguyen, V. N. T., Vo, K. T. X., Park, H., Jeon, J. S., and Jung, K. H. (2016). A systematic view of the MLO family in rice suggests their novel roles in morphological development, diurnal responses, the light-signaling pathway, and various stress responses. *Front. Plant Sci.* 7. doi: 10.3389/FPLS.2016.01413/BIBTEX

Oh, E., Kim, J., Park, E., Kim, J., Kang, C., and Choi, G. (2004). PIL5, a phytochrome-interacting basic helix-loop-helix protein, is a key negative regulator of seed germination in Arabidopsis thaliana. *Plant Cell* 16, 3045–3058. doi: 10.1105/TPC.104.025163

Oliver, S. A., Oliver, H. R., Wallace, J. S., and Roberts, A. M. (1987). Soil heat flux and temperature variation with vegetation, soil type and climate. *Agric. For Meteorol.* 39, 257–269. doi: 10.1016/0168-1923(87)90042-6

Paris, J. R., Stevens, J. R., and Catchen, J. M. (2017). Lost in parameter space: a road map for stacks. *Methods Ecol. Evol.* 8, 1360–1373. doi: 10.1111/2041-210X.12775

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., and Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PloS One* 7, e37135. doi: 10.1371/JOURNAL.PONE.0037135

Prasad, R. (2022). Cytokinin and its key role to enrich the plant nutrients and growth under adverse conditions-an update. *Front. Genet.* 13. doi: 10.3389/FGENE.2022.883924/BIBTEX

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. doi: 10.1093/GENETICS/155.2.945

Radosavljević, I., Križanović, K., Šarančić, S. L., and Jakše, J. (2023). Towards the Investigation of the Adaptive Divergence in a Species of Exceptional Ecological Plasticity: Chromosome-Scale Genome Assembly of Chouardia litardierei (Hyacinthaceae). *Int. J. Mol. Sci.* 24, 10755. doi: 10.3390/ijms241310755

Rathcke, B., and Lacey, E. P. (1985). Phenological patterns of terrestrial plants. *Annu. Rev. Ecol. Syst.* 16, 179–214. doi: 10.1146/ANNUREV.ES.16.110185.001143/CITE/REFWORKS

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ritter-Studnička, H. (1954). Flora i vegetacija livada kraških polja Bosne i Hercegovine. *Godišnjak Biološkog instituta u Sarajevu* 7, 25–110.

Rundle, H. D., and Nosil, P. (2005). Ecological speciation. *Ecol. Lett.* 8, 336–352. doi: 10.1111/J.1461-0248.2004.00715.X

Sandring, S., and Ågren, J. (2009). Pollinator-mediated selection on floral display and flowering time in the perennial herb Arabidopsis lyrata. *Evolution* 63, 1292–1300. doi: 10.1111/J.1558-5646.2009.00624.X

Sargent, R. D., and Roitberg, B. D. (2000). Seasonal decline in male-phase duration in a protandrous plant: a response to increased mating opportunities? *Funct. Ecol.* 14, 484–489. doi: 10.1046/J.1365-2435.2000.00453.X

Savolainen, O., Lascoux, M., and Merilä, J. (2013). Ecological genomics of local adaptation. *Nat. Rev. Genet.* 14, 807–820. doi: 10.1038/NRG3522

Schmalenbach, I., Zhang, L., Reymond, M., and Jiménez-Gómez, J. M. (2014). The relationship between flowering time and growth responses to drought in the Arabidopsis Landsberg erecta x Antwerp-1 population. *Front. Plant Sci.* 5. doi: 10.3389/FPLS.2014.00609/ABSTRACT

Schwartz, M. D. (2024). *Phenology: An Integrative Environmental Science*. 3rd Edn, ed. M. D. Schwartz. USA. Milwaukee: Springer Nature Switzerland AG 2024. Available online at: https://link.springer.com/book/10.1007/978-3-031-75027-4 (Accessed January 10, 2025).

Schwinning, S., Lortie, C. J., Esque, T. C., and DeFalco, L. A. (2022). What common-garden experiments tell us about climate responses in plants. *J. Ecol.* 110, 986–996. doi: 10.1111/1365-2745.13887

Settele, J., Bishop, J., and Potts, S. G. (2016). Climate change impacts on pollination. *Nat. Plants* 2. doi: 10.1038/NPLANTS.2016.92

Šilić, Č. (1990). Morfologija, horologija, ekologija i fenologija dviju grupa populacija Scilla litardierei Breistr. (Syn.: S. pratensis Waldst. & Kit. non Bergeret). *Bilt. Društva Ekol. BiH Ser. B*. 5, 107–116.

Song, X., Liu, L., Wang, C., Zhang, W., Li, Y., Zhu, J., et al. (2023). Real-time determination of flowering period for field wheat based on improved YOLOv5s model. *Front. Plant Sci.* 13. doi: 10.3389/FPLS.2022.1025663/BIBTEX

Soy, J., Leivar, P., and Monte, E. (2014). PIF1 promotes phytochrome-regulated growth under photoperiodic conditions in Arabidopsis together with PIF3, PIF4, and PIF5. *J. Exp. Bot.* 65, 2925–2936. doi: 10.1093/JXB/ERT465

Spano, D., Snyder, R. L., and Cesaraccio, C. (2013). "Mediterranean Phenology" in *Phenology: An Integrative Environmental Science* 3rd Edn, ed. M. D. Schwartz. USA. Milwaukee: Springer Nature Switzerland AG 2024. Available online at: https://link.springer.com/book/10.1007/978-3-031-75027-4. (Accessed January 10, 2025).

Thomas, J., Frost, R. R., and Harvey, R. D. (1973). Thermal conductivity of carbonate rocks. *Eng. Geol.* 7, 3–12. doi: 10.1016/0013-7952(73)90003-3

Todesco, M., Owens, G. L., Bercovich, N., Légaré, J. S., Soudi, S., Burge, D. O., et al. (2020). Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature* 584, 602–607. doi: 10.1038/s41586-020-2467-6

Turesson, G. (1922). The genotypical response of the plant species to the habitat. *Hereditas* 3, 211–350. doi: 10.1111/J.1601-5223.1922.TB02734.X

Turner, S. D. (2018). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Source Softw.* 3, 731. doi: 10.21105/JOSS.00731

Vicentini, G., Biancucci, M., Mineri, L., Chirivì, D., Giaume, F., Miao, Y., et al. (2023). Environmental control of rice flowering time. *Plant Commun.* 4, 100610. doi: 10.1016/J.XPLC.2023.100610

Wagner, D. H. (1995). A Rule of Thumb for Botanists: The 1 in 20 Rule *Oregon Flora-On-Line Newsletter*. United States: Oregon State University.

Walter, G. M., Clark, J., Cristaudo, A., Terranova, D., Nevado, B., Catara, S., et al. (2022). Adaptive divergence generates distinct plastic responses in two closely related Senecio species. *Evolution* 76, 1229. doi: 10.1111/EVO.14478

Wang, H., Wang, H., Ge, Q., and Dai, J. (2020). The interactive effects of chilling, photoperiod, and forcing temperature on flowering phenology of temperate woody plants. *Front. Plant Sci.* 11. doi: 10.3389/FPLS.2020.00443/BIBTEX

Waples, D. W., and Waples, J. S. (2004). A review and evaluation of specific heat capacities of rocks, minerals, and subsurface fluids. Part 1: Minerals and nonporous rocks. *Natural Resour. Res.* 13, 97–122. doi: 10.1023/B:NARR.0000032647.41046.E7/METRICS

Waser, N. M. (1983). "The Adaptive Nature of Floral Traits: Ideas and Eviaence" in *Pollination Biology*, ed. L. Real (Academic Press), 241–285. doi: 10.1016/B978-0-12-583980-8.X5001-7

Werner, T., Motyka, V., Strnad, M., and Schmülling, T. (2001). Regulation of plant growth by cytokinin. *Proc. Natl. Acad. Sci.* 98, 10487–10492. doi: 10.1073/PNAS.171304098

Yadav, A. (2024). Branching paths: Unveiling functional divergence of PIF1 and PIF4 via promoter and protein variations. *Plant Cell* 36, 2749–2750. doi: 10.1093/PLCELL/KOAE152

Yang, Y., and Karlson, D. T. (2011). Overexpression of AtCSP4 affects late stages of embryo development in Arabidopsis. *J. Exp. Bot.* 62, 2079–2091. doi: 10.1093/JXB/ERQ400

Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., et al. (2021). rMVP: A memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics Proteomics Bioinf.* 19, 619–628. doi: 10.1016/J.GPB.2020.10.007

Zaninović, K., Gajić-Čapka, M., Tadić, M. P., Vučetić, M., Milković, J., Bajić, A., et al. (2008). *Climate Atlas of Croatia: 1961–1990: 1971–2000/Klimatski Atlas Hrvatske, Climate Atlas of Croatia: 1961–1990: 1971–2000* (Zagreb: Državni hidrometeorološki zavod).

Zhou, X., Carbonetto, P., and Stephens, M. (2013). Polygenic modeling with Bayesian sparse linear mixed models. *PloS Genet.* 9, e1003264. doi: 10.1371/JOURNAL.PGEN.1003264

Zhou, Z., Feng, H., Ma, G., Ru, J., Wang, H., Feng, J., et al. (2024). Seasonal and vertical patterns of water availability and variability determine plant reproductive phenology. *Ann. Bot.* 135. doi: 10.1093/AOB/MCAE138

Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* 44, 821–824. doi: 10.1038/ng.2310

Glossary

APs	Aspartic Proteases
BAM	Binary Alignment Map
BOF	Beginning of flowering
BOS	Beginning of sprouting
BSLMM	Bayesian Sparse Linear Mixed Model
CCHC-ZFPs	CCHC-type zinc finger proteins
Chr	Chromosome
ddRAD-seq	Double Digest Restriction Site-Associated DNA Sequencing
EGGNog	Evolutionary Genealogy of Genes: Non-supervised Orthologous Groups
FPD	Flowering period duration
GEA	Genome-Environment Association
GEMMA	Genome-wide Efficient Mixed Model Association
GMMAT	Generalized Mixed Model Association Tests
GWAS	Genome-Wide Association Study
LMM	Linear Mixed Model
MAF	Minor Allele Frequency
mGWAS	Multivariate genome-wide association study
PGE	Proportion of Genetic Effect
PPR proteins	Pentatricopeptide repeat proteins
PFAM	Protein Family
PIF1	Phytochrome-interacting Factor 1
PIP	Posterior Inclusion Probability
PVE	Proportion of Variance Explained
SNP	Single Nucleotide Polymorphism
VPD	Vegetation period duration

Table 1. The locations of sampled *Chouardia litardierei* populations and their corresponding habitat types.

Population/Location	Country	Latitude (N)	Longitude (E)	Habitat Type
Bjelopolje	Croatia	44.693754°	15.773682°	Meadow – karst poljes
Cetina (Paško polje)	Croatia	43.940922°	16.436367°	Meadow – karst poljes
Budoške Bare	Montenegro	42.743747°	18.926361°	Meadow – karst poljes
Pag (Kolansko blato)	Croatia	44.514886°	14.919922°	Seashore - grassland
Nin	Croatia	44.249564°	15.172015°	Seashore - grassland
Vrana Lake	Croatia	43.937292°	15.514689°	Seashore - grassland
Lovćen	Montenegro	42.377169°	18.843117°	Dolomite - bedrock
Skadar Lake	Montenegro	42.326486°	19.069464°	Dolomite - bedrock
Pandurica	Montenegro	42.721628°	18.962442°	Dolomite - bedrock

Supplementary File 2

Table 1. Mann-Whitney post-hoc test for four phenological traits of *Chouardia litardierei*.

trait	continental-litorarl	continental-dolomitic	litoral-dolomitic
BOF	0.5425	1.48E-06	7.32E-09
BOS	1	4.13E-12	1.04E-18
FPD	NA	NA	NA
VPD	0.918	1.34E-15	3.09E-23

Table 2. Kruskal-Wallis test for equal medians for four phenological traits of *Chouardia litardierei*.

trait	H (chi2)	Hc (tie corrected)	*p* (same)
BOF	41.69	42.4	6.22E-10
BOS	86.83	87.4	1.048E-16
FPD	0.1727	0.1746	0.9164
VPD	113.8	113.9	1.845E-25

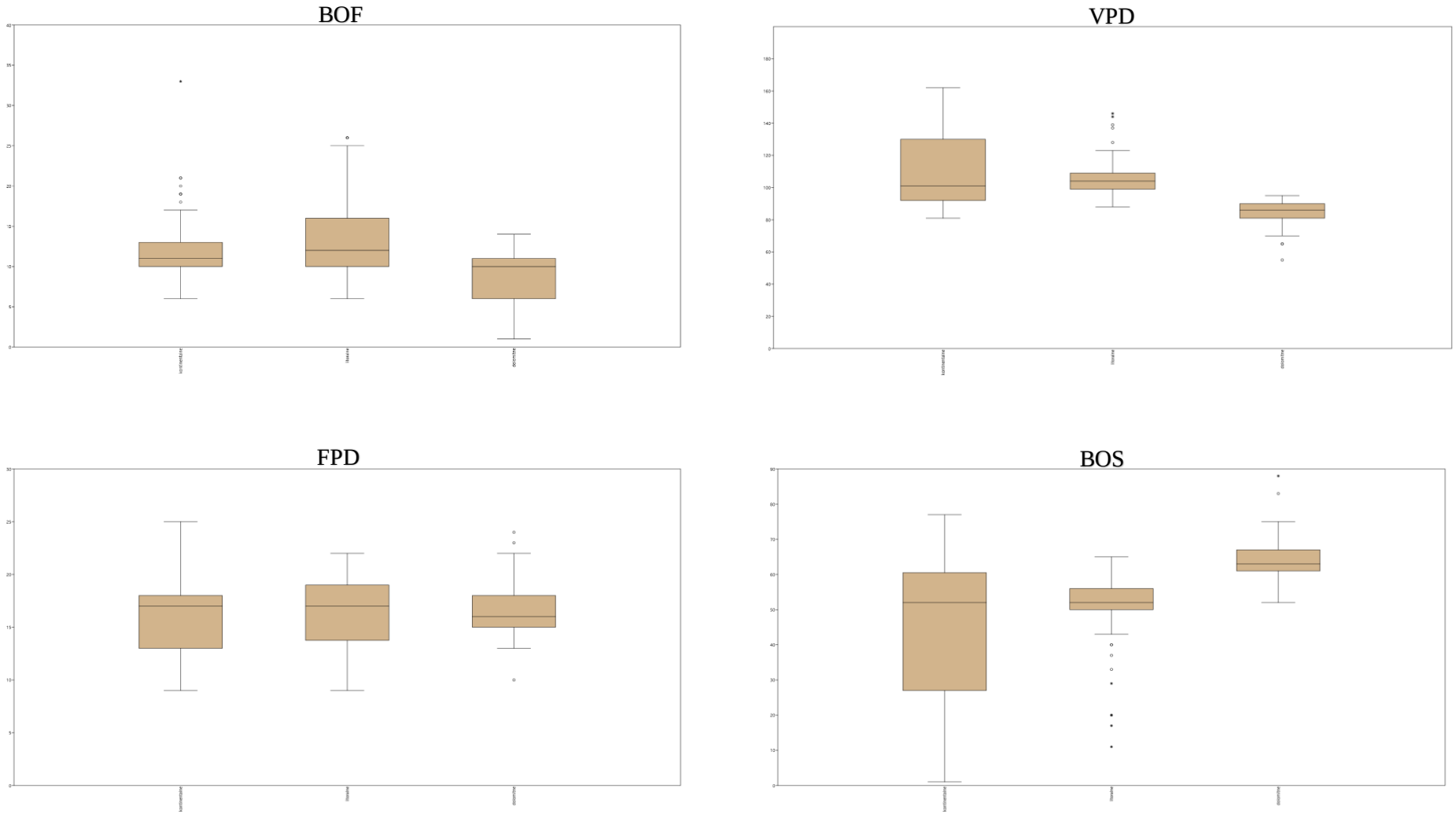


Figure 1. Box plots showing the distribution of four phenological traits (BOF, BOS, FPD, and VPD) in *Chouardia litardierei* populations across different habitat types. Each box plot represents the median, interquartile range (IQR), and variability of trait values across different populations. Whiskers indicate data within 1.5 times the IQR, while dots represent outliers.

Table 3. Kruskal-Wallis test for equal medians for four phenological traits in nine *Chouardia litardierei* populations.

trait	**H (chi2)**	**Hc (tie corrected)**	***p* (same)**
BOF	104	105.8	2.762E-19
BOS	141.5	142.5	7.243E-27
FPD	21.44	21.68	0.005543
VPD	168.2	168.3	2.961E-32

Table 4. Mann-Whitney post-hoc test for VPD trait in nine *Chouardia litardierei* populations.

	BJELOPOLJE	CETINA	BUDOŠKE BARE	VRANA LAKE	NIN	PAG	SKADAR LAKE	PANDURICA	LOVĆEN
BJELOPOLJE		5.72E-05	2.79E-06	0.003915	2.23E-05	2.74E-06	1.15E-06	3.84E-06	1.16E-06
CETINA	5.72E-05		1.18E-05	0.03173	1	0.4125	2.31E-07	1.73E-06	1.82E-07
BUDOŠKE BARE	2.79E-06	1.18E-05		2.88E-07	2.07E-06	0.00124	0.1842	1	0.0003949
VRANA LAKE	0.003915	0.03173	2.88E-07		0.5151	3.51E-06	7.26E-08	3.62E-07	7.28E-08
NIN	2.23E-05	1	2.07E-06	0.5151		0.00987	7.83E-08	3.93E-07	7.85E-08
PAG	2.74E-06	0.4125	0.00124	3.51E-06	0.00987		1.16E-06	1.91E-05	5.55E-07
SKADAR LAKE	1.15E-06	2.31E-07	0.1842	7.26E-08	7.83E-08	1.16E-06		1	0.2871
PANDURICA	3.84E-06	1.73E-06	1	3.62E-07	3.93E-07	1.91E-05	1		0.002412
LOVĆEN	1.16E-06	1.82E-07	0.0003949	7.28E-08	7.85E-08	5.55E-07	0.2871	0.002412	

Table 5. Mann-Whitney post-hoc test for FPD trait in nine *Chouardia litardierei* populations.

	BJELOPOLJE	CETINA	BUDOŠKE BARE	VRANA LAKE	NIN	PAG	SKADAR LAKE	PANDURICA	LOVĆEN
BJELOPOLJE		0.052	0.02663	0.2822	0.2316	0.2235	1	0.05792	0.03659
CETINA	0.052		1	1	1	1	0.6153	1	1
BUDOŠKE BARE	0.02663	1		1	1	1	0.3502	1	1
VRANA LAKE	0.2822	1	1		1	1	1	1	1
NIN	0.2316	1	1	1		1	1	1	1
PAG	0.2235	1	1	1	1		1	1	1
SKADAR LAKE	1	0.6153	0.3502	1	1	1		1	1
PANDURICA	0.05792	1	1	1	1	1	1		1
LOVĆEN	0.03659	1	1	1	1	1	1	1	

Table 6. Mann-Whitney post-hoc test for BOS trait in nine *Chouardia litardierei* populations.

	BJELOPOLJE	CETINA	BUDOŠKE BARE	VRANA LAKE	NIN	PAG	SKADAR LAKE	PANDURICA	LOVĆEN
BJELOPOLJE		0.0014	1.37E-06	0.03928	0.00014	2.82E-06	2.21E-07	1.704E-07	2.2E-07
CETINA	0.0014		0.000529	1	1	0.4914	8.21E-06	0.0000121	1.1E-06
BUDOŠKE BARE	1.4E-06	0.00053		5E-06	0.0012	0.04781	1	1	0.2271
VRANA LAKE	0.03928	1	5E-06		1	0.000133	2.89E-07	2.661E-07	1.9E-07
NIN	0.00014	1	0.001198	1		1	2.71E-05	6.456E-05	2.2E-06
PAG	2.8E-06	0.4914	0.04781	0.00013	1		0.000483	0.0009476	1.4E-05
SKADAR LAKE	2.2E-07	8.2E-06	1	2.9E-07	2.7E-05	0.000483		1	0.2374
PANDURICA	1.7E-07	1.2E-05	1	2.7E-07	6.5E-05	0.000948	1		0.04234
LOVĆEN	2.2E-07	1.1E-06	0.2271	1.9E-07	2.2E-06	1.41E-05	0.2374	0.04234	

Table 7. Mann-Whitney post-hoc test for BOF trait in nine *Chouardia litardierei* populations.

	BJELOPOLJE	CETINA	BUDOŠKE BARE	VRANA LAKE	NIN	PAG	SKADAR LAKE	PANDURICA	LOVĆEN
BJELOPOLJE		0.004181	0.7619	1	1	2.35E-05	1E-05	0.00142	0.002
CETINA	0.004181		1	3.701E-05	0.001553	1	0.01413	1	1
BUDOŠKE BARE	0.7619	1		0.01927	0.7324	0.009554	0.00026	0.6021	1
VRANA LAKE	1	3.701E-05	0.01927		1	4.78E-07	2.1E-07	1E-05	9E-06
NIN	1	0.001553	0.7324	1		3.73E-06	9.1E-07	0.00027	0.0003
PAG	2.35E-05	1	0.009554	4.775E-07	0.000003726		0.9649	1	1
SKADAR LAKE	1E-05	0.01413	0.000261	2.089E-07	9.085E-07	0.9649		0.7455	0.0772
PANDURICA	0.001416	1	0.6021	1.006E-05	0.0002744	1	0.7455		1
LOVĆEN	0.002024	1	1	8.607E-06	0.0003147	1	0.07717	1	

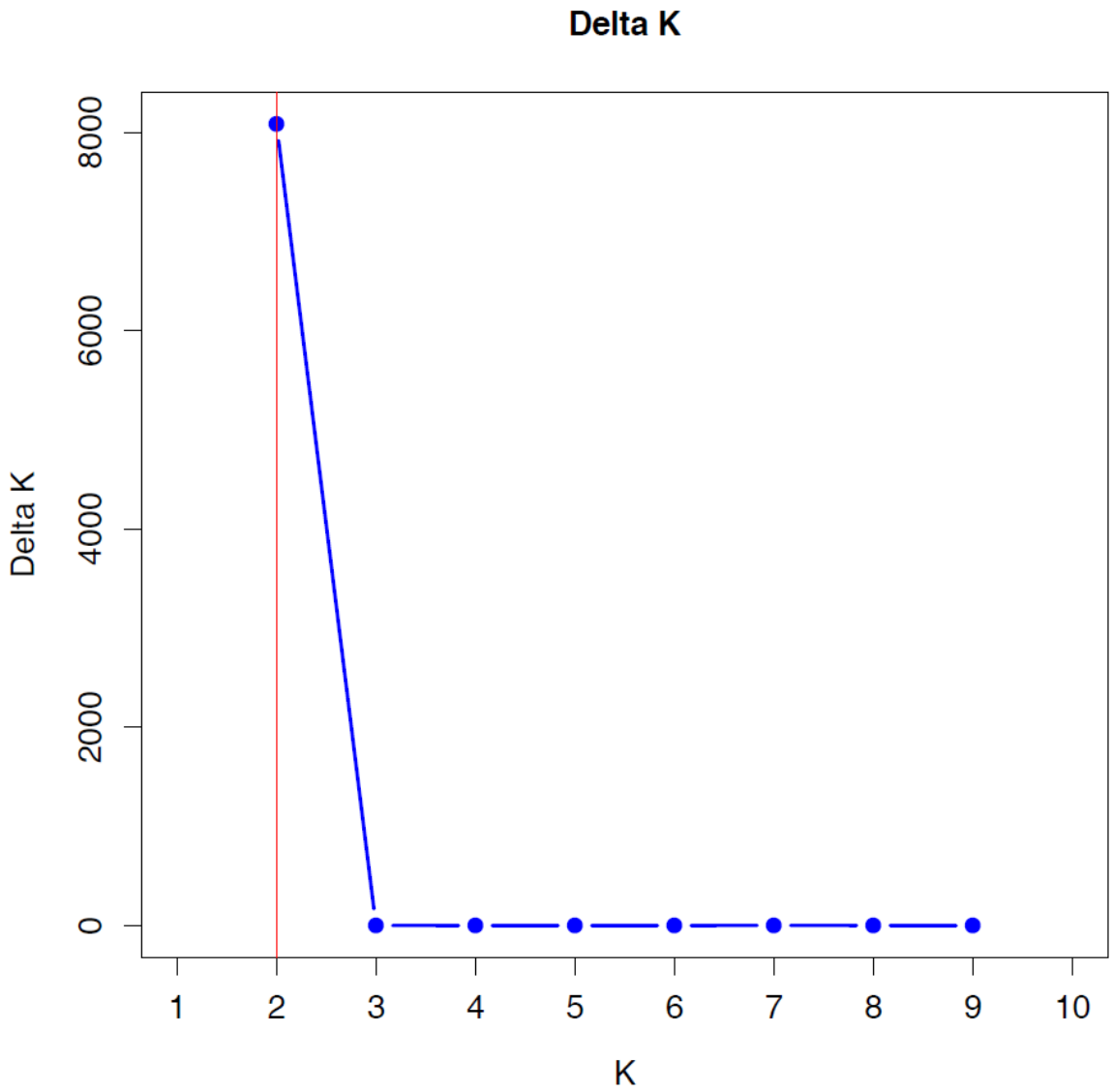


Figure 1. Delta K values as obtained by the STRUCTURE software.

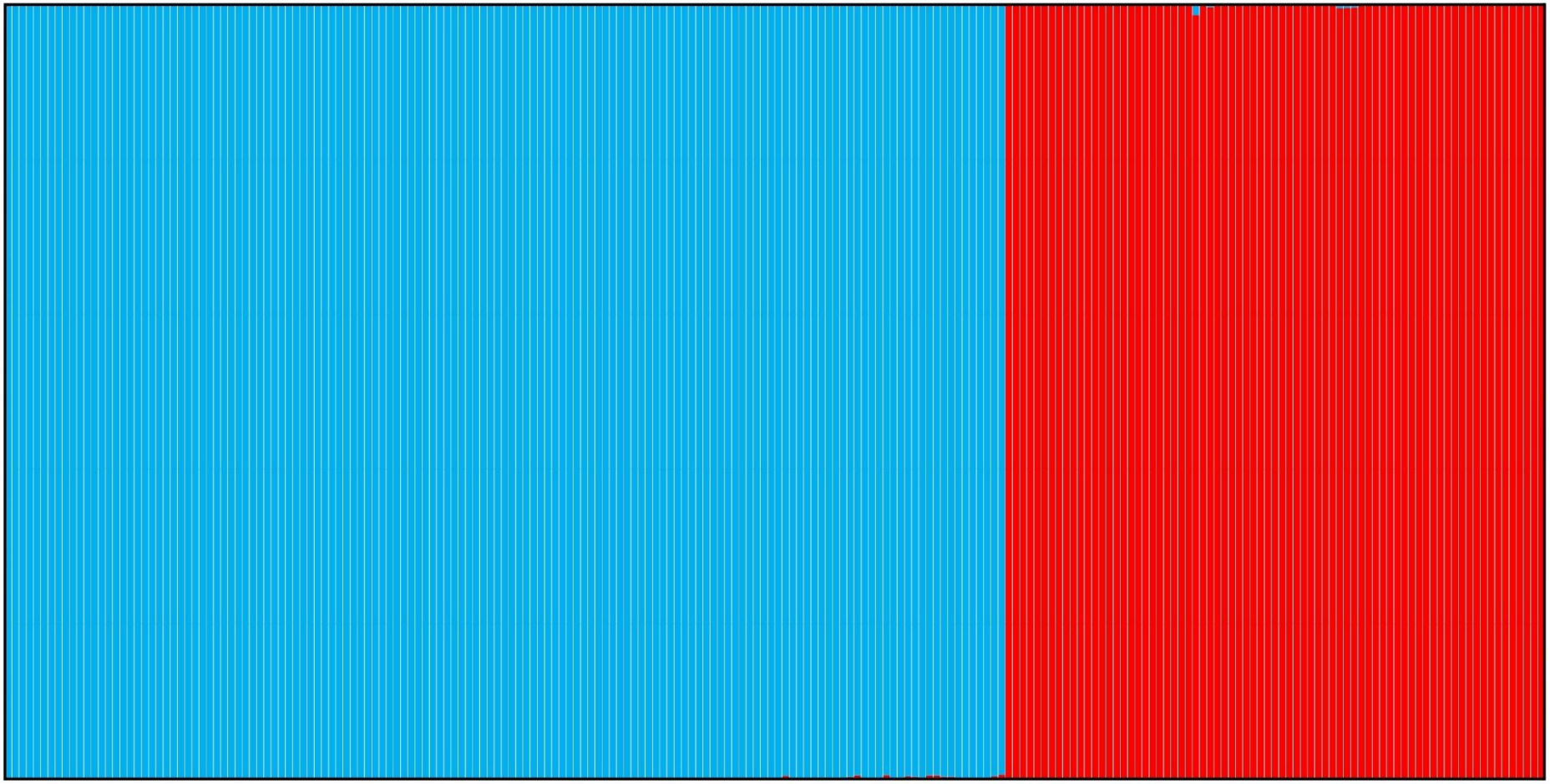


Figure 1. Population-genetic structure of studied *Chouardia litardierei* populations as revealed by the STRUCTURE software. Each stacked column represents a single individual. Individuals belonging to the meadow and the seashore groups of populations are marked with blue, and individuals from the dolomite habitat populations are marked with red.

Table 1. SNPs identified as having a major sparse effect (PIP > 0.1) on **FPD**, **VPD**, **BPF** and **BOS** traits in the multi-SNP Bayesian sparse linear mixed model (BSLMM) analysis.

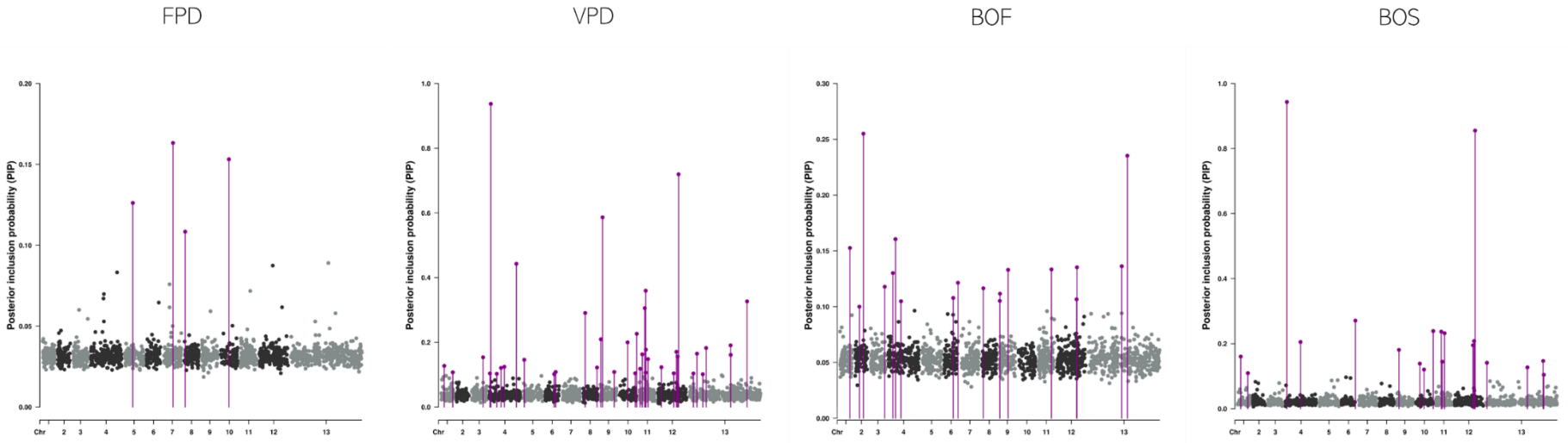
Trait	SNP	Chr	Position	Multi-SNP BSLMM Analysis β (PIP)
FPD	131957_13	10	97222552	-0.699 (0.167)
	750129_37	7	113120650	-0.476 (0.165)
VPD	565532_39	4	14626431	-0.327 (0.912)
	210123_39	12	239066297	0.328 (0.748)
	64746_61	9	23100056	0.222 (0.579)
	608473_26	4	312243109	-0.228 (0.447)
	169192_99	11	71192707	-0.212 (0.393)
	165641_27	11	59678762	-0.224 (0.308)
	22755_41	8	27370914	-0.229 (0.286)
	392671_80	13	668483795	0.188 (0.277)
	345151_44	13	477215139	0.159 (0.236)
	68440_39	9	3481745	-0.295 (0.232)
	114136_23	10	203465059	-0.244 (0.225)
	131957_13	10	97222552	0.181 (0.224)
	158923_36	11	31022405	-0.174 (0.224)
	651864_26	5	14713180	-0.161 (0.192)
	16906_26	8	164769480	-0.213 (0.188)
	274011_25	13	193652215	-0.183 (0.184)
	203695_51	12	214471601	-0.161 (0.180)
	562222_42	4	133794937	-0.135 (0.175)
	206941_26	12	227677762	-0.206 (0.175)
	431951_18	13	86851653	0.173 (0.167)
	518753_26	3	131097306	-0.206 (0.164)
	404599_18	13	718065214	0.176 (0.158)
	264719_36	13	154558733	0.211 (0.152)
	345740_18	13	478800231	-0.186 (0.152)
	169723_49	11	73174727	0.177 (0.148)
	455977_31	1	38506814	-0.187 (0.146)
	779448_31	7	37409277	-0.127 (0.141)
	453245_41	1	28792186	0.184 (0.141)
	175156_81	11	97958069	-0.176 (0.139)
	732487_30	6	50042305	-0.172 (0.137)
	236070_45	12	39684927	-0.245 (0.136)
	794077_90	7	99621809	0.159 (0.135)
	713226_25	6	119601819	0.135 (0.131)
	167443_28	11	6497513	0.177 (0.129)
	628242_25	4	62348852	0.144 (0.122)
	708419_55	6	102061527	-0.147 (0.122)
	108954_31	10	185953028	-0.140 (0.119)
	56223_19	9	159032132	0.169 (0.119)
	168723_26	11	69545761	0.165 (0.117)
	321566_29	13	384493611	0.190 (0.115)
	571637_57	4	172299280	-0.166 (0.115)
	63780_22	9	20268889	0.150 (0.115)
	144881_40	11	149677564	0.149 (0.115)
	196140_32	12	183140101	-0.194 (0.114)
	41769_13	9	100304229	0.123 (0.114)
	243670_21	12	66893770	-0.094 (0.114)
	446954_131	1	137876700	0.195 (0.111)
	625480_18	4	5030882	-0.155 (0.111)
	293282_20	13	27101492	-0.144 (0.109)
	760751_41	7	159527948	-0.123 (0.109)

	338697_27	13	45325469	-0.148 (0.108)
	597865_16	4	273124373	-0.140 (0.107)
	633475_14	4	84524636	0.130 (0.107)
	790359_35	7	80015952	-0.128 (0.107)
	584545_89	4	22256369	0.115 (0.106)
	76972_41	9	66872924	0.164 (0.106)
	278704_14	13	212393830	0.127 (0.104)
	723031_44	6	159155591	-0.119 (0.103)
	207532_23	12	229820496	0.106 (0.102)
	173326_23	11	91514076	-0.124 (0.102)
	577335_20	4	194281453	-0.129 (0.101)
	513067_18	3	110714483	0.098 (0.100)
	296028_31	13	282176585	0.169 (0.100)
	346507_22	13	481318344	0.092 (0.100)
	321869_23	13	385455641	0.131 (0.099)
	318805_22	13	374121017	-0.189 (0.099)
	541757_17	3	63012003	0.113 (0.099)
	373292_20	13	588262669	0.133 (0.099)
	541107_21	3	59731100	-0.211 (0.098)
	369203_29	13	572629490	0.096 (0.098)
	642210_29	5	103184092	-0.132 (0.097)
	486980_18	2	22661955	0.116 (0.096)
	162069_19	11	45040054	0.135 (0.096)
	634062_21	4	8802327	0.098 (0.096)
	171198_47	11	79969607	-0.140 (0.095)
BOF	504422_54	2	95535920	0.321 (0.172)
	337862_27	13	450832737	0.281 (0.156)
	633306_18	4	83549086	-0.378 (0.098)
BOS	565532_39	4	14626431	0.437 (0.959)
	210123_39	12	239066297	-0.482 (0.829)
	723031_44	6	159155591	0.307 (0.356)
	175156_81	11	97958069	0.342 (0.277)
	114136_23	10	203465059	0.376 (0.241)
	64746_61	9	23100056	0.229 (0.222)
	203695_51	12	214471601	0.279 (0.221)
	165641_27	11	59678762	0.310 (0.219)
	206941_26	12	227677762	0.391 (0.208)
	571637_57	4	172299280	0.353 (0.202)
	169192_99	11	71192707	0.245 (0.177)
	131957_13	10	97222552	-0.286 (0.172)
	121534_19	10	48645623	0.326 (0.169)
	455977_31	1	38506814	0.358 (0.156)
	252718_24	13	10389007	0.208 (0.143)
	441735_22	1	121138696	-0.292 (0.138)
	345740_18	13	478800231	0.284 (0.133)
	391526_58	13	661633656	0.300 (0.126)
	708419_55	6	102061527	0.239 (0.123)
	651208_30	5	143900923	-0.177 (0.123)
	732487_30	6	50042305	0.358 (0.121)
	392671_80	13	668483795	-0.236 (0.117)
	657278_42	5	171253039	0.193 (0.103)
	518753_26	3	131097306	0.264 (0.100)
	437888_20	1	105737657	0.228 (0.096)
	496686_32	2	61419316	-0.311 (0.095)
	565532_39	4	14626431	0.437 (0.959)
	210123_39	12	239066297	-0.482 (0.829)
	723031_44	6	159155591	0.307 (0.356)
	175156_81	11	97958069	0.342 (0.277)

	114136_23	10	203465059	0.376 (0.241)
	64746_61	9	23100056	0.229 (0.222)
	203695_51	12	214471601	0.279 (0.221)
	165641_27	11	59678762	0.310 (0.219)
	206941_26	12	227677762	0.391 (0.208)
	571637_57	4	172299280	0.353 (0.202)
	169192_99	11	71192707	0.245 (0.177)
	131957_13	10	97222552	-0.286 (0.172)
	121534_19	10	48645623	0.326 (0.169)
	455977_31	1	38506814	0.358 (0.156)
	252718_24	13	10389007	0.208 (0.143)
	441735_22	1	121138696	-0.292 (0.138)
	345740_18	13	478800231	0.284 (0.133)
	391526_58	13	661633656	0.300 (0.126)
	708419_55	6	102061527	0.239 (0.123)
	651208_30	5	143900923	-0.177 (0.123)
	732487_30	6	50042305	0.358 (0.121)
	392671_80	13	668483795	-0.236 (0.117)
	657278_42	5	171253039	0.193 (0.103)
	518753_26	3	131097306	0.264 (0.100)
	437888_20	1	105737657	0.228 (0.096)
	496686_32	2	61419316	-0.311 (0.095)

BSLMM was fitted on 23,315 SNPs.; BSLMM, Bayesian sparse linear mixed model; BOF, Beginning of Flowering; BOS, Beginning of Sprouting; Chr, Chromosome; FPD, Flowering Period Duration; PIP, Posterior Inclusion Probability; SNP, Single Nucleotide Polymorphism; VPD, Vegetation Period Duration.

Figure 1. Manhattan plots of the BSLMM analysis for the FPD, VPD, BOF, and BOS traits of the *Chouardia litardierei*. The x-axis represents the chromosomal position of SNPs, and the y-axis represents their posterior inclusion probabilities (PIPs).



BSLMM; Bayesian Sparse Linear Mixed Model, BOF, Beginning of Flowering; BOS, Beginning of Sprouting; FPD, Flowering Period Duration; VPD, Vegetation Period Duration.

Table 1. Means, medians, and 95% equal tail posterior probability intervals (95% ETPPIs) of hyperparameters estimated from the Bayesian sparse linear mixed model (BSLMM) in phenological trait **FPD**, **VPD**, **BOF** and **BOS**.

Trait	Hyperparameter	Mean	Median	2.5%	97.5%
FPD	h	0.3244	0.3128	0.0316	0.6830
	PVE	0.2026	0.1798	0.0171	0.5215
	rho	0.5220	0.5315	0.0275	0.9799
	PGE	0.4722	0.4776	-	0.9671
	pi	3.23×10^{-2}	1.06×10^{-2}	6.40×10^{-4}	1.42×10^{-1}
	n.gamma	60.67	20.00	-	267.00
VPD	h	0.8224	0.8305	0.6667	0.9312
	PVE	0.8695	0.8717	0.7803	0.9451
	rho	0.6511	0.6902	0.1185	0.9872
	PGE	0.6572	0.7720	0.0553	0.9947
	pi	5.86×10^{-2}	4.62×10^{-2}	2.18×10^{-3}	1.55×10^{-1}
	n.gamma	111.28	88.00	4.00	290.00
BOF	h	0.7046	0.7164	0.4483	0.8934
	PVE	0.6603	0.6622	0.4659	0.8455
	rho	0.3267	0.2735	0.0118	0.8882
	PGE	0.2586	0.1556	-	0.9121
	pi	2.52×10^{-2}	1.07×10^{-2}	6.14×10^{-4}	1.11×10^{-1}
	n.gamma	47.29	20.00	0.00	207.00
BOS	h	0.7319	0.7414	0.5371	0.8738
	PVE	0.7605	0.7616	0.6332	0.8829
	rho	0.6346	0.6475	0.2029	0.9796
	PGE	0.6319	0.6870	0.1598	0.9888
	pi	3.05×10^{-2}	2.25×10^{-2}	2.87×10^{-3}	1.09×10^{-1}
	n.gamma	52.58	39.00	5.00	189.00

BSLMM was fitted on 23,315 SNPs. BOF, Beginning of Flowering; BOS, Beginning of Sprouting; FPD, Flowering Period Duration; h, approximation to the proportion of phenotypic variance explained by variants; n.gamma, number of variants with major effect; PGE, Proportion of Genetic variance explained by variants with major effect; pi, proportion of variants with non-zero effects; PVE, proportion of phenotypic variance explained by variants; rho, approximation to the proportion of genetic variance explained by variants with major effect; VPD, Vegetation Period Duration.

Table 1. SNPs passing the genome-wide significance threshold ($p < 1 \times 10^{-3}$) in the multivariate linear mixed model (mvLMM) analysis for VPD and BOS traits of *Chouardia litardierei* in GEMMA multivariate GWAS.

SNP	Chr	Position	Effect allele	Reference allele	Beta1 (VPD)	Beta2 (BOS)	mvLMM Analysis in GEMMA (*p*-value)
475214_20	2	10692011	T	C	0.01	-0.16	6.68×10^{-10}
654470_57	5	159590069	G	A	-0.81	0.23	9.18×10^{-10}
689821_13	5	93776308	A	C	-0.77	0.22	1.66×10^{-8}
83332_22	9	9096582	G	A	0.18	-0.31	4.37×10^{-8}
558534_14	4	116336173	T	A	-0.92	0.48	6.36×10^{-8}
145174_73	11	15061603	C	G	-0.54	-0.01	1.02×10^{-7}
522007_32	3	141694985	C	T	-0.32	-0.07	2.72×10^{-7}
179782_120	12	117670370	T	C	-0.49	0.02	3.20×10^{-7}
248115_31	12	8479581	A	T	-0.51	-0.06	3.46×10^{-7}
28256_68	8	50959986	A	G	-0.38	0.67	4.19×10^{-7}
187632_54	12	151393801	T	G	-0.40	-0.01	5.14×10^{-7}
268970_17	13	171155860	A	G	-0.49	-0.03	7.85×10^{-7}
307326_19	13	326551046	A	G	-0.66	0.06	1.17×10^{-6}
550061_33	3	97695604	A	T	-0.46	0.88	1.24×10^{-6}
26676_33	8	4513527	C	G	-0.23	-0.11	2.32×10^{-6}
187238_23	12	149770349	T	C	-0.34	-0.08	3.29×10^{-6}
227369_31	12	307596886	C	A	-0.16	-0.03	4.09×10^{-6}
305761_25	13	320423026	T	G	0.35	-0.37	7.85×10^{-6}
467066_106	1	91683912	C	T	-0.55	0.07	9.29×10^{-6}
770850_26	7	195082174	C	T	-0.61	0.26	9.98×10^{-6}
65720_38	9	26233589	A	G	0.61	-0.59	1.24×10^{-5}
565532_39	4	14626431	C	A	-0.45	0.55	1.73×10^{-5}
674245_25	5	2834749	C	G	-0.30	0.03	1.75×10^{-5}
67781_20	9	32778369	T	A	-0.16	-0.19	2.65×10^{-5}
221833_73	12	284678317	C	G	-0.23	0.30	3.20×10^{-5}
214650_53	12	255323079	A	G	-0.74	0.46	4.01×10^{-5}
651220_77	5	143922570	T	G	-0.32	-0.10	4.46×10^{-5}
365394_49	13	555889659	C	T	-0.28	0.14	4.51×10^{-5}
723031_44	6	159155591	G	A	-0.27	0.52	4.78×10^{-5}
202818_127	12	210719984	C	T	-0.57	0.30	5.72×10^{-5}
622051_16	4	357265730	A	G	-0.31	0.81	6.32×10^{-5}
592384_37	4	253302099	T	C	-0.16	-0.43	6.73×10^{-5}
537684_52	3	44057957	T	G	-0.16	-0.06	8.02×10^{-5}
310101_18	13	337885640	T	C	-0.33	-0.05	9.55×10^{-5}
681559_42	5	58303261	G	T	-0.39	0.19	9.72×10^{-5}
334377_114	13	437172692	C	A	0.75	-0.62	9.93×10^{-5}
600093_25	4	281363631	A	G	-0.63	0.19	1.42×10^{-4}
261183_91	13	142406888	T	G	0.22	-0.48	1.43×10^{-4}
328618_13	13	414580639	T	A	-0.16	-0.18	1.53×10^{-4}
422540_18	13	792932133	A	G	-0.87	0.44	1.55×10^{-4}
607415_46	4	308484921	A	G	0.22	-0.36	1.56×10^{-4}
32158_19	8	64100022	T	A	-0.05	-0.21	1.74×10^{-4}
177171_18	12	105568874	A	G	-0.73	0.60	1.87×10^{-4}
475216_15	2	10692219	T	G	-0.01	-0.18	1.91×10^{-4}
657020_14	5	169994000	G	A	-0.27	0.02	1.93×10^{-4}
104384_40	10	170660529	G	A	-0.25	0.04	1.97×10^{-4}
771754_39	7	197927864	G	C	0.03	-0.20	2.00×10^{-4}
301480_19	13	302007580	A	G	-0.35	0.06	2.07×10^{-4}

167223_27	11	64125165	T	G	-0.46	0.49	2.14×10^{-4}
688909_13	5	89858040	A	C	-0.04	-0.20	2.15×10^{-4}
858479_18	13	78001418	T	C	0.36	-0.20	2.22×10^{-4}
143076_28	11	14367538	T	C	-0.27	-0.07	2.24×10^{-4}
548241_67	3	88011092	T	C	-0.25	-0.09	2.37×10^{-4}
583_81	8	102136191	C	A	-0.24	0.58	2.38×10^{-4}
196140_32	12	183140101	A	T	-0.74	0.37	2.41×10^{-4}
540381_39	3	5572300	A	T	0.14	-0.24	3.13×10^{-4}
614876_51	4	334032848	C	T	-0.01	-0.13	3.20×10^{-4}
790473_18	7	8052404	A	T	0.85	-1.13	3.34×10^{-4}
740200_45	6	89087646	C	A	-0.50	0.04	3.66×10^{-4}
71327_30	12	43693525	T	G	-0.41	0.02	3.67×10^{-4}
612469_18	13	325349603	A	C	0.14	0.16	3.70×10^{-4}
364317_22	9	551367550	A	G	-0.70	0.29	4.26×10^{-4}
210123_39	4	239066297	T	G	0.51	-0.58	4.28×10^{-4}
419317_22	13	780074412	T	C	-0.39	-0.09	4.49×10^{-4}
202068_64	12	207629434	T	C	-0.31	-0.02	4.57×10^{-4}
60130_20	9	17391855	C	T	-0.32	0.73	4.58×10^{-4}
478635_49	8	122304660	A	C	-0.30	0.04	4.73×10^{-4}
170834_35	13	774970	G	A	-0.35	0.11	4.82×10^{-4}
381580_38	5	619787450	T	G	-0.27	0.06	4.97×10^{-4}
154168_43	12	180827479	C	A	-0.27	0.57	5.07×10^{-4}
774777_66	3	206933711	C	T	0.57	-0.72	5.10×10^{-4}
548930_13	4	90980761	T	G	-0.53	0.23	5.11×10^{-4}
170124_137	13	74451579	A	G	-0.67	0.27	5.17×10^{-4}
711345_25	9	112404079	C	A	0.04	0.15	5.20×10^{-4}
207337_22	12	229211403	A	C	0.59	-0.14	5.22×10^{-4}
439975_31	12	114209920	A	G	0.01	-0.19	5.30×10^{-4}
212201_44	4	246320665	T	C	-0.04	-0.10	5.46×10^{-4}
631869_38	9	7683248	T	G	0.70	-0.46	5.65×10^{-4}
767107_28	9	18271185	A	G	-0.30	0.02	5.72×10^{-4}
55253_106	7	155581032	G	A	-0.34	0.08	5.81×10^{-4}
322826_18	12	390026768	A	T	-0.42	0.18	5.86×10^{-4}
165592_49	4	59562398	A	G	0.04	0.23	5.94×10^{-4}
529999_34	7	1714447	A	G	-0.35	-0.02	5.97×10^{-4}
264780_28	13	154888633	T	A	0.12	-0.62	6.07×10^{-4}
234677_87	7	34919398	G	C	-0.19	0.42	6.15×10^{-4}
201405_32	12	204415035	A	T	-0.27	-0.07	6.26×10^{-4}
618657_20	7	345766799	A	G	0.93	-1.05	6.71×10^{-4}
76756_18	12	65979867	C	A	-0.63	0.19	6.81×10^{-4}
71567_24	13	44564317	T	C	0.22	-0.02	6.89×10^{-4}
765142_28	9	175841496	A	C	-0.83	1.17	6.90×10^{-4}
57078_21	5	163441584	A	C	0.22	-0.21	6.98×10^{-4}
484023_15	3	144655732	G	A	-0.50	0.11	7.03×10^{-4}
78404_13	5	72519187	T	A	-0.57	0.29	7.06×10^{-4}
76416_37	7	64503815	A	G	0.35	-0.38	7.19×10^{-4}
545555_90	3	78147943	T	G	-0.14	-0.17	7.21×10^{-4}
512981_36	12	110347767	G	T	0.06	0.16	7.26×10^{-4}
684715_22	7	7273425	T	C	-0.34	0.49	7.38×10^{-4}
170122_38	12	74451257	A	G	-0.66	0.26	7.50×10^{-4}
200956_103	9	202665910	C	T	0.20	0.22	7.55×10^{-4}
528481_19	12	165845777	T	C	0.74	-0.55	7.72×10^{-4}
180221_31	5	11898341	C	A	-0.29	0.48	7.78×10^{-4}
635043_17	4	93373391	T	C	0.37	-0.44	8.20×10^{-4}
581314_17	13	210039723	A	G	-0.12	0.42	8.55×10^{-4}
371940_27	4	583317033	T	G	-0.55	0.16	8.66×10^{-4}
76460_82	9	64634814	G	C	-0.28	0.15	8.71×10^{-4}
560706_36	7	127372766	A	C	-0.30	-0.06	8.82×10^{-4}
792151_44	9	90140260	A	G	0.11	-0.63	9.00×10^{-4}
772653_20	12	200562469	C	T	-0.32	0.53	9.03×10^{-4}
775046_20	2	207737647	A	G	0.19	0.22	9.09×10^{-4}

151787_41	5	173804269	C	G	-0.11	-0.05	9.58×10^{-4}
476624_22	5	112342852	A	G	0.24	-0.06	9.68×10^{-4}
415244_13	9	7620614	A	G	0.13	-0.26	9.75×10^{-4}
77325_18	4	68217631	A	C	-0.49	0.80	9.94×10^{-4}

mvLMM in GEMMA was fitted on 23,315 SNPs. BOS, Beginning Sprouting; Chr, Chromosome; VPD, Vegetation Period Duration; mvLMM, multivariate Linear Mixed Model; SNP, Single Nucleotide Polymorphism.

Table 1. SNPs passing the genome-wide significance threshold ($p < 1 \times 10^{-3}$) in the multivariate linear mixed model (mvLMM) analysis for VPD and BOF traits of *Chouardia litardierei* in GEMMA multivariate GWAS.

SNP	Chr	Position	Effect allele	Reference allele	Beta1 (VPD)	Beta2 (BOF)	mvLMM Analysis in GEMMA (*p*-value)
65720_38	9	26233589	A	G	0.59	0.25	8.71×10^{-6}
305761_25	13	320423026	T	G	0.34	0.01	2.04×10^{-5}
504422_54	2	95535920	T	G	0.23	0.49	2.16×10^{-5}
565532_39	4	14626431	C	A	-0.45	0.15	3.26×10^{-5}
334377_114	13	437172692	C	A	0.73	0.49	7.39×10^{-5}
28256_68	8	50959986	A	G	-0.39	0.44	8.58×10^{-5}
104625_19	10	171714671	C	A	0.55	0.71	8.73×10^{-5}
622077_29	4	357364641	C	T	0.32	-0.62	9.61×10^{-5}
252813_22	13	104630774	C	G	0.10	-1.27	1.03×10^{-4}
528481_19	3	165845777	T	C	0.73	0.80	1.03×10^{-4}
757985_15	7	148582555	C	T	0.47	-0.77	1.40×10^{-4}
321869_23	13	385455641	T	C	0.36	0.79	1.51×10^{-4}
207869_16	12	231068248	T	A	0.23	0.60	1.57×10^{-4}
631869_38	4	7683248	T	G	0.70	0.80	1.87×10^{-4}
774777_66	7	206933711	C	T	0.61	-0.27	2.41×10^{-4}
431930_95	13	86813045	A	T	0.21	0.37	2.68×10^{-4}
167223_27	11	64125165	T	G	-0.46	0.16	2.74×10^{-4}
221833_73	12	284678317	C	G	-0.22	0.11	3.14×10^{-4}
186978_19	12	148693882	A	C	-0.22	0.82	4.17×10^{-4}
226324_53	12	303921633	A	G	0.55	-0.66	4.24×10^{-4}
86910_49	10	104918258	T	C	0.15	-0.50	4.33×10^{-4}
618657_20	4	345766799	A	G	0.95	-0.28	4.52×10^{-4}
445520_34	1	133744238	A	G	0.06	0.39	5.72×10^{-4}
723279_19	6	159836366	C	T	0.62	-0.54	5.95×10^{-4}
623516_30	4	40486247	T	G	0.31	-0.33	6.08×10^{-4}
207870_38	12	231068565	C	T	0.21	0.53	6.11×10^{-4}
558534_14	4	116336173	T	A	-0.92	0.30	6.51×10^{-4}
76416_37	9	64503815	A	G	0.35	0.06	6.51×10^{-4}
65785_19	9	26407191	G	A	-0.18	-0.13	6.61×10^{-4}
210123_39	12	239066297	T	G	0.50	-0.01	6.84×10^{-4}
41769_13	9	100304229	A	G	0.28	0.53	7.44×10^{-4}
144881_40	11	149677564	T	C	0.42	-0.81	8.05×10^{-4}
177171_18	12	105568874	A	G	-0.73	0.13	8.30×10^{-4}
708427_23	6	102085089	A	C	-0.24	-0.43	8.38×10^{-4}
422713_13	13	793523369	G	A	-0.32	-0.53	8.45×10^{-4}
171277_71	11	80221934	T	C	0.12	-0.63	9.85×10^{-4}

mvLMM in GEMMA was fitted on 23,315 SNPs. BOF, Beginning of Flowering; Chr, Chromosome; FPD, Flowering Period Duration; mvLMM, multivariate Linear Mixed Model; SNP, Single Nucleotide Polymorphism.

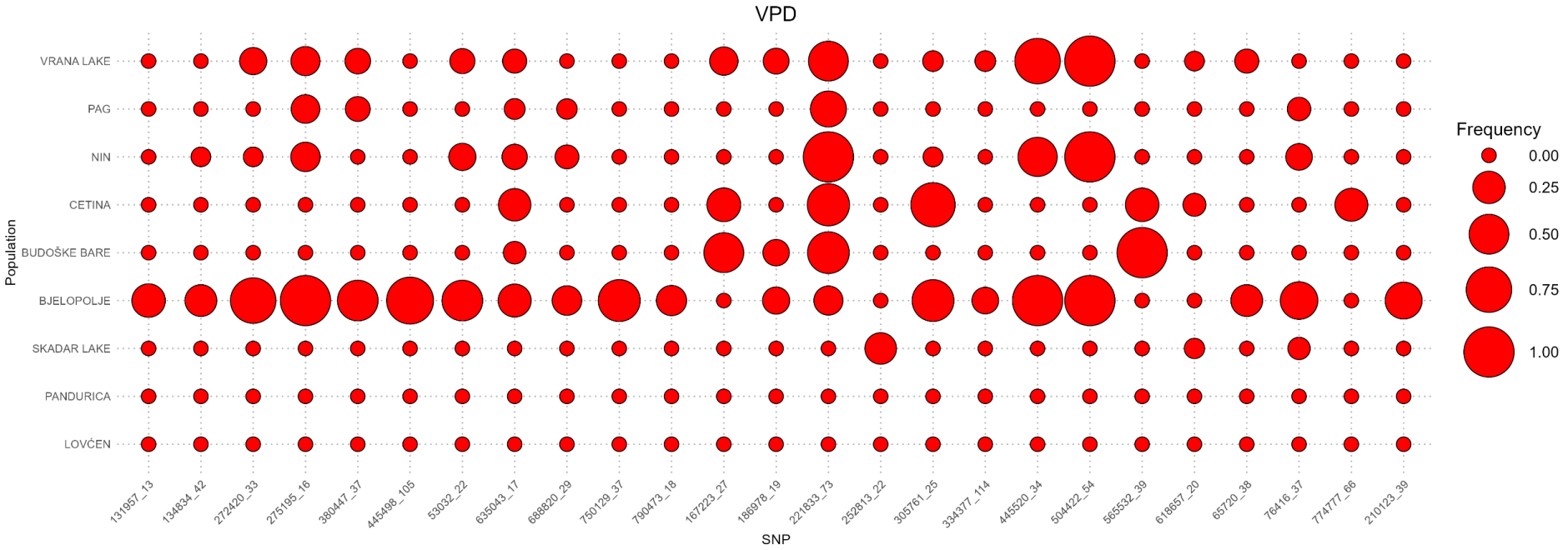


Figure 1. Frequency of effect alleles across populations for significant SNPs identified in the single-SNP LMM analysis (GEMMA and GMMAT), as well as the multi-SNP BSLMM analysis, all of which surpassed the genome-wide significance threshold (1×10^{-3}) for the Vegetation Period Duration (VPD) trait. The analysis also includes SNPs meeting the same threshold in the multivariate GWAS. The corresponding SNPs are detailed in Table 3 and Table 5 of the manuscript.

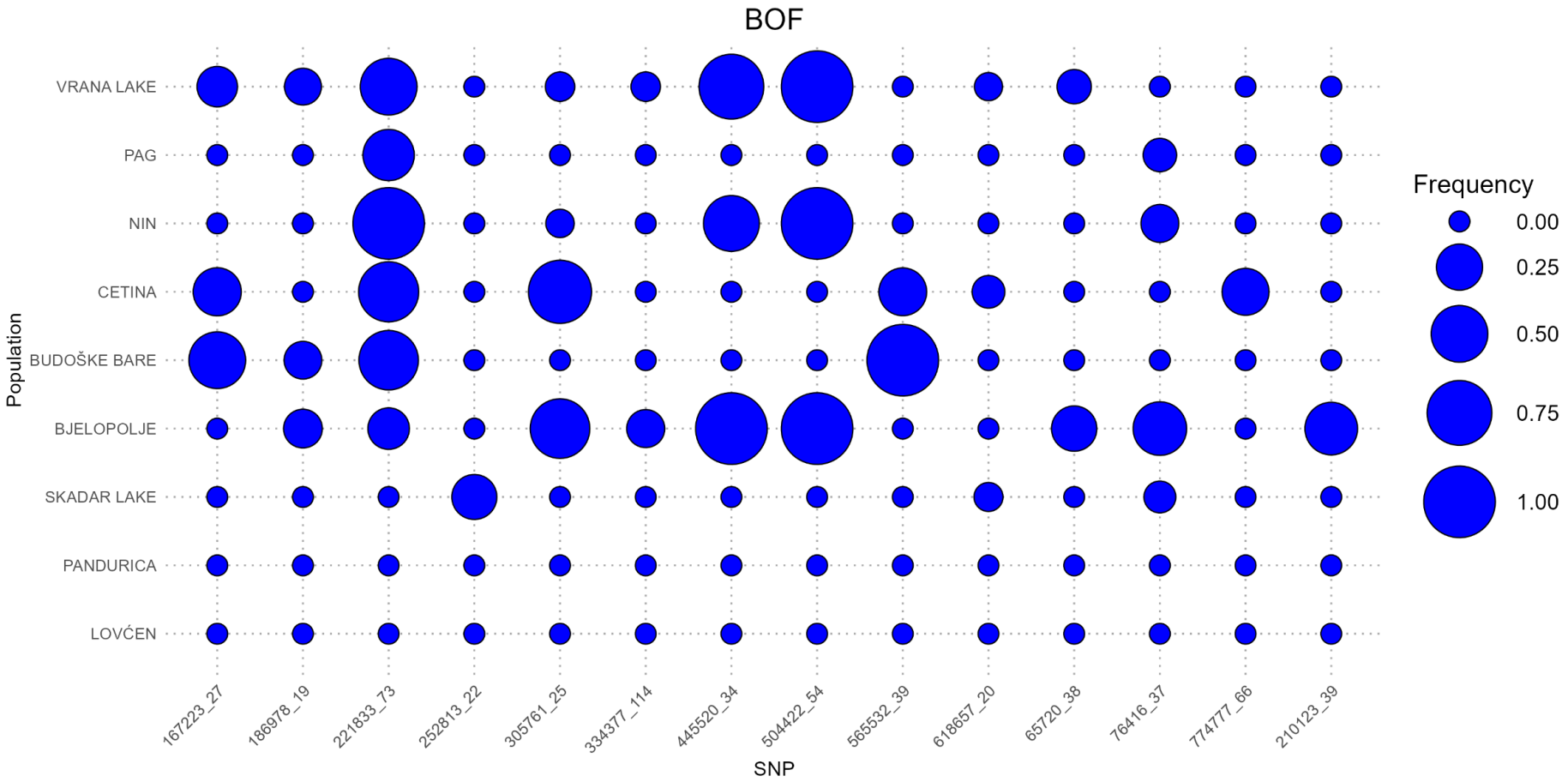


Figure 2. Frequency of effect alleles across populations for significant SNPs identified in the single-SNP LMM analysis (GEMMA and GMMAT), as well as the multi-SNP BSLMM analysis, all of which surpassed the genome-wide significance threshold (1×10^{-3}) for the Beginning of Flowering (BOF) trait. The analysis also includes SNPs meeting the same threshold in the multivariate GWAS. The corresponding SNPs are detailed in Table 3 and Table 5 of the manuscript.

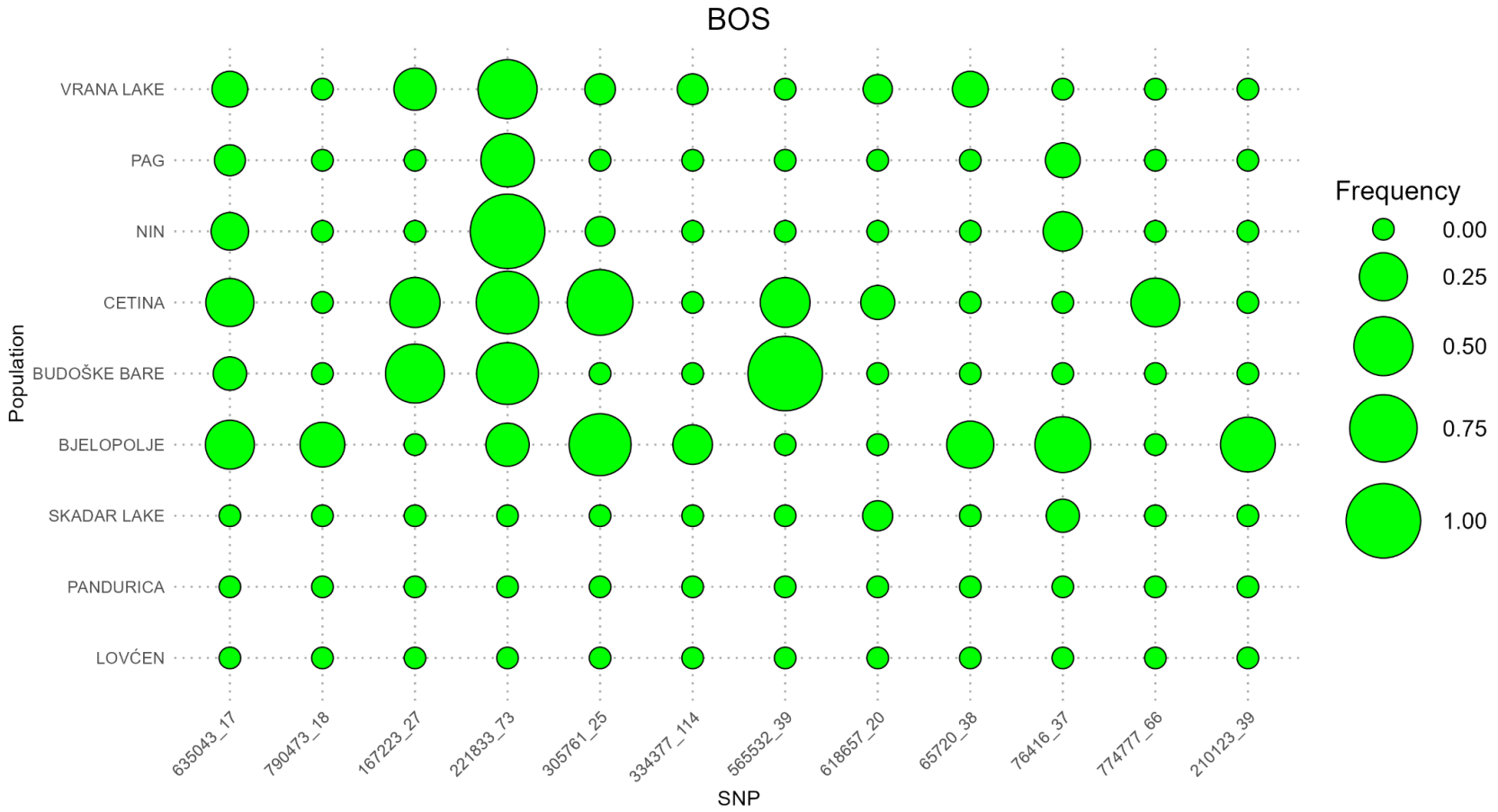


Figure 3. Frequency of effect alleles across populations for significant SNPs identified in the single-SNP LMM analysis (GEMMA and GMMAT), as well as the multi-SNP BSLMM analysis, all of which surpassed the genome-wide significance threshold (1×10^{-3}) for the Beginning of Sprouting (BOS) trait. The analysis also includes SNPs meeting the same threshold in the multivariate GWAS. The corresponding SNPs are detailed in Table 3 and Table 5 of the manuscript.

Figure 4. Frequency of effect alleles across populations for significant SNPs identified in the single-SNP LMM analysis (GEMMA and GMMAT), as well as the multi-SNP BSLMM analysis, all of which surpassed the genome-wide significance threshold (1×10^{-3}) for the Flowering Period Duration (FPD) trait. The analysis also includes SNPs meeting the same threshold in the multivariate GWAS. The corresponding SNPs are detailed in Table 3 of the manuscript.

Due to size constraints, the following supplementary files are not reproduced in this thesis but can be accessed at the journal's website (Frontiers in Plant Science, DOI: 10.3389/fpls.2025.1571608):

Supplementary File 11: EggNOG output file for 7 SNP loci that exceeded the genome-wide significance threshold (1×10^{-3}) in the multivariate GWAS analysis of the *Chouardia litardierei* traits: FPD, VPD, BOF, and BOS.

Supplementary File 12: EggNOG output file for 13 SNP loci that exceeded the genome-wide significance threshold (1×10^{-3}) in the multivariate GWAS analysis of the *Chouardia litardierei* traits: BOS and VPD.

Supplementary File 13: EggNOG output file for 14 SNP loci that exceeded the genome-wide significance threshold (1×10^{-3}) in the multivariate GWAS analysis of the *Chouardia litardierei* traits: BOF and VPD.

7.3 Publication III

Šarančić, S. L., Pleić, N., Mitić, D., Križanović, K., Surina, B., and Radosavljević, I. (2025b). Genome-wide association study (GWAS) provides insights into the genomic basis of reproduction-related traits in *Chouardia litardierei* (Asparagaceae). BMC Plant Biology 2025 25:1 25, 1–25. doi: 10.1186/S12870-025-06617-4

BMC Plant Biology

RESEARCH

Open Access

Genome-wide association study (GWAS) provides insights into the genomic basis of reproduction-related traits in *Chouardia litardierei* (Asparagaceae)

Sara Laura Šarančić[1], Nikolina Pleić[2], Damjan Mitić[1], Krešimir Križanović[3], Boštjan Surina[4,5] and Ivan Radosavljević[1*]

Abstract

Background *Chouardia litardierei*, commonly known as amethyst meadow squill, is a plant species characterized by profound ecological plasti vcity. As a wild, non-model species, it represents a suitable system for gaining insights into the genomic background of the local adaptation process. By implementing a genome-environment and genome-wide association studies, we sought to investigate the genomic regions related to the local adaptation and the development of several reproduction-related traits in *C. litardierei*: for sexual reproduction, Average Height of Inflorescences (AHI) and Total Flower Count (TFC) per genotype, and for asexual reproduction, Bulb Count (BC) per genotype.

Results A genome-environment association (GEA) study of selected *C. litardierei* populations revealed the precipitation of the coldest quarter as the bioclimatic variable with the most substantial influence on detected variability, with numerous candidate genes detected and functionally characterized. To evaluate the genetic basis of selected reproduction-related traits we combined phenotypic data of 214 individuals raised as a part of a common garden experiment with ddRADseq genotyping results. After implementing various single- and multi-locus GWAS models for all traits, multiple candidate loci affecting their development were recognized. In addition, high, narrow-sense heritability estimates indicated that genetic factors accounted for over 55% of the phenotypic variance in each trait. Notably, the highest heritability estimate was observed for the Average Height of Inflorescences (71.95%), suggesting its crucial role in reproductive success. Functional annotation of the associated genomic regions identified key protein families involved in reproduction-related biological pathways, including nitrogen metabolism, phytohormone regulation, and floral organs development.

Conclusion By implementing GEA and GWAS, we revealed a list of candidate loci significantly associated with adaptation to specific environmental variables and morphological traits related to sexual and asexual reproduction in *C. litardierei*. These findings provide a foundation for a deeper understanding of the molecular mechanisms driving the local adaptation processes occurring among *C. litardierei* populations from different habitat types. At the same time, the high heritability estimates of morphological traits further underscore the significance of genetic factors in the local adaptation process.

Keywords GWAS, GEA, Local adaptation, Adaptive traits, Reproduction, *Chouardia litardierei*

*Correspondence:
Ivan Radosavljević
ivan.radosavljevic@biol.pmf.hr
Full list of author information is available at the end of the article

Background

The intricate interplay between contrasting environmental conditions shapes the genetic architecture of various traits favored by natural selection [1, 2]. Natural selection acts on allele frequencies, driving populations toward local adaptation [3] through the development of distinct phenotypic variations within populations [4]. To cope with the highly contrasting environmental conditions, populations must adapt rapidly to survive and ultimately ensure reproductive success [5, 6]. Such changes not only enhance the species' adaptation to specific ecological niches but also play a pivotal role in the broader context of speciation, contributing to the emergence of unique populations with reproductive isolation potential [2].

The development of reproductive barriers presents a crucial point in the lineages' divergence and speciation. However, besides sexual reproduction, which is usually the focus of evolutionary biologists, as much as 80% of angiosperms reproduce asexually through vegetative propagation [7], also referred to as clonal growth. The clonal type of reproduction is considered to emerge in situations where different types of biotic or abiotic stress threaten the success of sexual reproduction [8, 9]. The balance between sexual and asexual reproduction may vary significantly among populations of the same species, strongly influencing the evolution of life history traits [10]. In addition, a trade-off between these reproduction types occurs since both require substantial resources, and different allocation patterns may develop [11, 12]. The impact of clonality on sexual reproduction is severe, as it can strongly influence its spatial patterns by causing the non-random distribution of genotypes and the development of a spatial genetic structure in affected populations [13]. In addition, since clonality positively influences levels of geitonogamy and, consequently, inbreeding, it can also directly influence the lineages' divergence process [14, 15]. Facing challenging environmental conditions, many plant species rely on bulbs as vital storage organs, which enable them to endure dormant periods, mitigate the effects of adverse environmental conditions [16], and maintain reproductive capacity across heterogeneous habitats [17], as reflected in the number of bulbs produced. Bulb formation is further regulated by internal signaling pathways that respond to the surrounding ecological conditions [18]. The number of flowers per inflorescence has been shown to affect pollination success and subsequent seed production [19]. Suetsugu et al. [20] demonstrated that inflorescence size influences pollinator behavior in the deceptive orchid *Cephalanthera falcata*, serving as both a visual attractant and a mechanism for enhancing pollen accumulation and deposition. Similarly, subtle variations in inflorescence height can influence pollinator accessibility and optimize pollen dispersal [21].

Among other evolutionary phenomena (e.g., genetic drift, complex genetic architecture, or demographic history of studied species), phenotypic plasticity presents one of the severe challenges when studying the genetic background of complex polygenic traits. One of the more efficient tools to overcome this challenge is the common garden experiment. By growing individuals originating from populations experiencing contrasting environmental conditions in a common environment, the idea is to control and restrain the expression of phenotypic plasticity, thus obtaining more reliable results [22]. Since it enables overcoming the hampering effect of different environmental conditions to the characterization of complex phenotypes' genetic basis, the common garden experiments were often used in various local adaptation studies (e.g., [23–26]. Extreme caution is also needed when performing genome-wide association studies in non-model species due to the confounding effect of phenotypic plasticity. However, common garden experiments can greatly help address this problem and are consequently being implemented in such studies [27–29].

Chouardia litardierei (Breist.) Speta (Asparagaceae) is a bulbous perrenial. It develops a sizeable racemose inflorescence, usually comprising several dozen radially symmetrical flowers with no specific pollination-related morphological adaptations. Although this has never been studied, it is presumingly an open-pollinated species. In addition to sexual reproduction, it reproduces clonally by producing numerous bulbs surrounding the central one. This species distribution area stretches across the Dinaric Alps karst environment in the western Balkans, from Slovenia in the north-west to Montenegro in the south-east [30, 31], a region known for its exceptional environmental heterogeneity and consequently, diverse spectrum of available ecological niches [32, 33]. Three groups of populations can be distinguished based on their habitat types. The largest group predominantly occupies karst poljes, flat-bottomed basins characterized by karstic drainage systems. These fields, typically enclosed by rugged dolomite and limestone mountains and characterized by deep and nutrient-reach soils, experience periodic floodings typically lasting for several months each year [34], thus presenting a hydrologically and geomorphologically unique environment [35, 36]. The fewest populations are found in the coastal salt marshes of northern Dalmatia, a habitat subjected to tidal flooding and dominated by salt-tolerant vegetation [37]. Finally, the southernmost group of populations inhabits highly contrasting habitat types: drought-prone dolomite slopes characterized by minimal amounts of soil typically present only in rock crevices. This hostile

environment is known for its reduced water and nutrient capacity, and pronounced seasonality in temperature and water availability [38]. While a previous attempt was made to characterize the dolomite group of populations as a distinct taxon [39], the reliability of the results was compromised due to indistinct approaches employed in the research, raising justified doubts about the validity of the results [40]. Despite the apparent ecological differences between these habitats, identifying consistent morphological distinctions among the assumed ecotypes remains challenging [39], highlighting the need for a deeper investigation into the genetic foundations underlying these specific morphological traits. This species presents a valuable study system for investigating local adaptation and speciation for several reasons. First, it exhibits marked ecological plasticity, with three groups of populations (Fig. 1) adapted to contrasting environmental conditions [40]. Second, being a small bulbous perennial makes it suitable for cultivation under controlled conditions, thus reducing phenotypic plasticity oscillations as a confounding factor in trait analysis [41]. Third, *C. litardierei* populations are, for the most part, distributed across easily accessible locations in the Dinaric Alps of the Balkan Peninsula [30, 39, 42], enabling comprehensive sampling.

Understanding the genomic basis of specific traits in the context of environmental dynamics is essential for uncovering the mechanisms underlying local adaptation and response to contrasting ecological pressures [43, 44]. As a foundational step in investigating the genomic basis of local adaptation in *C. litardierei*, we have already introduced a high-quality, chromosome-scale assembly of the *C. litardierei* genome [40]. Beyond a prior attempt to categorize the dolomite group of populations as a distinct taxon [39], limited research has been conducted on the ecological divergence or genetics of this species, aside from the cytogenetic characterization of two individuals presumed to belong to the meadow and dolomite groups of populations [45], and the chromosome-scale genome assembly mentioned above [40]. To advance our understanding of the genetic architecture underlying local adaptation, we have implemented both a genome-environment association (GEA) study based on available bioclimatic variables and a genome-wide association study (GWAS), which integrated morphometric data from a common garden experiment with ddRADseq genotyping. These analyses aimed to elucidate the genetic basis of local adaptation and the reproduction-related traits in selected populations of the wild, non-model monocot species *C. litardierei.*

Fig. 1 **A** *Chouardia litardierei* individual from the common garden experiment. Contrasting types of habitats of the studied *C. litardierei* populations; **B** Seashore grassland developed on deep soils, prone to occasional tidal floodings and salinization, experiencing the Mediterranean climate; **C** Inland karst poljes' meadows on deep and rich soils, exposed to seasonal floodings that can last up to several months, and **D** Drought- and heat-stress prone dolomite bedrocks habitat with very little available soil, characterized by highly unhospitable environmental elements

Methods

Plant material, common garden experiment, and phenotyping

To set up the common garden experiment, 214 individuals were transplanted from nine selected populations of *C. litardierei*, with three populations representing each of the three presumed habitat types (Fig. 1).

During the sampling expeditions, 22—25 individuals per population, situated no closer than 10 m from each other, were selected following the 1:20 rule [46]. The geographic coordinates of the sampling locations are provided in the Additional File 1. Simultaneously, leaf material from each individual required for DNA extraction was gathered and desiccated using silica gel. Sampled individuals (represented as a single bulb) were transplanted into a separate two-litre plastic container filled with soil, sand, and perlite. The containers were placed in raised beds outdoors as part of a common garden setup, allowing the plants to grow under temperate continental climate conditions (Cfb climate type according to Köppen classification) [47, 48]. No additional interventions, such as supplemental watering or pesticide application, were practiced, thus allowing plants to grow under undisturbed environmental conditions.

All voucher specimens were deposited in a publicly accessible herbarium at the Natural History Museum Rijeka (Index Herbariorum: NHMR), under the accession numbers NHMR 3306 (Budoške bare population), NHMR 3189 (Lovćen population), NHMR 2097 (Skadar lake population), NHMR 3151 (Pandurica population), NHMR 3247 (Cetina population), NHMR 3125 (Bjelopolje population), NHMR 3255 (Nin population), NHMR 3304 (Pag population), and NHMR 3305 (Vrana lake population). Voucher specimens were collected and identified by Ivan Radosavljević and Boštjan Surina.

To study the genetic background of selected reproduction-related traits in *C. litardierei*, we conducted a common garden experiment using nine populations from different habitat types across its distribution range. To test the ecological relatedness among studied populations in terms of prevailing climatic conditions they experience in their natural habitat and the relative positioning of the common garden experiment site, we ran the PC analysis based on 19 WorldClim bioclimatic variables for the 1950–2000 period in the 30 s resolution [49] using the basic prcomp R function. For the visualization of PCA results, R package "ggfortify" [50] was used. Measurements were carried out after two vegetational seasons of acclimatization to minimize carry-over effects from the original environment. Three distinct reproduction-related morphological traits (Table 1) were measured: (i) Total Flower Count (TFC), and (ii) the Average Height of Inflorescences (AHI) as pollination-related traits of great importance for sexual reproduction, and (iii) the Bulb Count per genotype (BC) as an indicator of asexual reproduction rate. TFC was determined as the number of flowers across all inflorescences per genotype. At the same time, AHI was measured using a graduated ruler with a precision of 0.1 cm, with the heights of individuals' inflorescences averaged. All of the studied traits were considered polygenic.

Pearson correlation analysis was conducted to examine the relationships between AHI and TFC variables with a

Table 1 Descriptive statistics of the *Chouardia litardierei* reproduction-related morphological traits examined in the study

Overall			
Trait	**Description**		**Mean ± SD**
TFC (number)	Count of flowers across all inflorescences per genotype		90.59 ± 45.59
AHI (cm)	Average height of inflorescences per genotype		18.00 ± 4.80
BC (number)	Count of bulbs per genotype		2.78 ± 3.06
By location			
Location	**BC (Mean ± SD)**	**AHI (Mean ± SD)**	**TFC (Mean ± SD)**
Bjelopolje	1.91 ± 1.54	14.68 ± 2.48	39.50 ± 21.33
Budoške Bare	2.14 ± 1.39	15.00 ± 3.20	82.00 ± 42.14
Cetina	4.67 ± 2.71	18.49 ± 3.09	81.17 ± 33.33
Lovćen	0.00 ± 0.00	15.81 ± 2.96	102.72 ± 44.22
Nin	4.83 ± 2.75	20.40 ± 3.54	97.00 ± 32.50
Pag	4.52 ± 2.91	15.78 ± 3.04	64.45 ± 38.68
Pandurica	0.36 ± 0.99	24.70 ± 4.43	115.90 ± 52.41
Skadar	0.56 ± 0.92	14.31 ± 2.94	109.24 ± 51.56
Vrana Lake	6.29 ± 3.57	22.63 ± 3.56	109.12 ± 36.30

AHI Average Height of Inflorescences, *BC* Bulb Count, *SD* Standard Deviation, *TFC* Total Flower Count

normal distribution using the "stats" package in R [51]. For BC, which does not follow a normal distribution, Spearman's correlation was performed using the same package.

Sequencing and genomic data processing

DNA isolation was performed using the GenElute™ Plant Genomic DNA Miniprep Kit (Sigma–Aldrich®). Concentrations were assessed using the Qubit™ Fluorometer (Thermo Fisher Scientific, Wilmington, DE, USA), and samples were diluted to a concentration of 20 ng/µL.

To perform genotyping of the studied *C. litardierei* populations, a ddRADseq approach was employed [52]. In short, DNA was first digested using restriction enzymes AseI and NsiI (NEB # R0526L and # R0127L, respectively). The resulting fragments were ligated with barcoded i5 and i7 adapters, after which all the samples were multiplexed. Final amplification was performed after nick repair using DNA polymerase I (NEB # M0209L). Obtained DNA libraries were double-sequenced (150 bp PE) on the Illumina HiSeq X platform.

The initial sequencing data was preprocessed with quality trimming and adapter removal using Trim Galore [53]. After trimming, BAM files were created by aligning the reads to the *C. litardierei* reference genome [40] through the Burrow-Wheelers Aligner [54]. SNP identification was done using the Stacks software package v1.48 [55]. The ref_map.pl wrapper module was employed, and in line with the suggestions of Paris et al. [56], the *pstacks* module was executed to extract loci previously aligned to the reference genome, with a minimum depth of coverage set at three. This ensures a reliable representation of loci across samples, reducing the risk of low-confidence genotype calls. Subsequently, the *cstacks* module generated a comprehensive catalogue of loci across populations, permitting a maximum of four mismatches among sample loci during its construction, further minimizing potential alignment errors. Finally, the *populations* module computed population-level summary statistics, requiring loci to be present in all nine populations and at least 70% of individuals within each population, with a maximum observed heterozygosity of 0.70. Further constraints were applied to retain only one SNP per locus and discard loci with minor allele frequencies (MAF) below 1%, ensuring the inclusion of high-quality, well-represented genetic markers. By focusing on common and stable genetic variants, this approach minimized the risk of inaccuracies arising from sequencing or sampling errors. The final dataset was generated in.vcf format for downstream analysis.

Population-genetic and genome-environment association analysis

Several methods were implemented to assess the genetic structure of the studied populations and deepen our understanding of phylogenetic relationships among them. First, the VCF file was converted into a genlight object using the gl.read.vcf function from the "dartR" v2.9.7. package [57]. We performed PCA using the prcomp R function, and the R package "plotly" v4.10.4 (https://plotly.com/r/, accessed on 9 Dec 2024) was used to construct the PCA plot. We used the hierarchical clustering method implemented in the R package FactoMineR v2.11 [58] to assess the optimal number of clusters of PCA data. The results were visualized with a dendrogram using the factoextra v1.0.7 package [59].

For the assessment of the genetic structuring of studied populations, we transformed genlight into a geno object using the gl2geno function in the "dartR" package. The sparse non-negative matrix factorization (sNMF) method was implemented using the "LEA" R package [60], with 100,000 iterations, 50% burnin, and 20 repetitions for K-values from 1 to 10. From the results, a cross-entropy values graph was constructed using a basic R function plot to select the optimum K-value. Furthermore, we constructed the phylogenetic tree based on Nei's genetic distance matrix to better appreciate phylogenetic relationships among studied populations. VcfR2genind function from the "vcfR" v1.15.0. package [61] was used to create a genind object, which was further transformed into a genpop object by using the genind2genpop function from "adegenet" package [62]. To generate Nei's genetic distance matrix, dist.genpop function from "adegenet" was used. We used the obtained matrix to create a bootstrapped phylogenetic tree (1,000 replicates) using the aboot function from the "poppr" v2.9.6 package [63]. "Ape" package [64] was used to convert the obtained tree to the "Newick" format that was used for the final visualization of the phylogenetic tree in the MEGA7 software [65].

To gain a more profound knowledge of the adaptation of the populations studied to local environmental conditions, we performed the RDA (linear model redundancy analysis) [66, 67]. Compared to other approaches often used for similar purposes of detecting the genetic signatures of local adaptation like generalized linear models (GLM) or latent factor mixed models (LFMM), the RDA was recognized as a superior method, as it is characterized by high true positive and low false positive rates [68, 69]. We started the procedure by downloading 19 available WorldClim bioclimatic variables for the 1950–2000 period in the 30 s resolution [49]. For the location of each sampled population, we used Qgis v3.16.0 (https://qgis.org/) to extract the data. We treated temperature and precipitation

variables (BIO1-11 and BIO 12–19, respectively) as separate datasets to test multicollinearity among variables. We used the vifstep function from the "usdm" package [70] to calculate variance inflation factors (VIF) for each variable, and only variables with a VIF < 10 were retained for further analysis. RDA implemented in the R package "vegan" (https://cran.r-project.org/web/packages/vegan/index.html Accessed 3 Dec 2024.) [71] was used for the characterization of the retained variables' influence on the genetic variations among studied populations. The optimal model was determined using the ordiR2 step function for a forward selection procedure with 10,000 permutations. To test the significance of the RDA model, we ran the anova.cca function with 10,000 permutations. We used the SNPs' loadings (i.e. coordinates) in the ordination space obtained through the RDA analysis to identify which loci are under selection. We considered loci as outliers if their loadings were more than 3 standard deviations away from the mean loading on either of the first two RDA axes (two-tailed p-value = 0.0027), following the recommendations of Forester et al. [69]. To assess the potential biological function of the genes positioned near the outliers, we extracted surrounding 50 kb DNA windows (25 kb upstream and downstream of the loci) using a custom Python script and the available draft genome assembly [40]. Finally, the obtained.fasta files were compared against the eggNOG database [72, 73]. Recognized candidate genes were manually inspected, and ones with functions seemingly of biological importance for the local adaptation to tested environmental variables were retained.

Genome-Wide Association Analyses (GWAS)

The schematic representation of the methodological approach we employed in GWAS analysis is presented in Fig. 2.

All traits were considered polygenic, and GWAS analyses were conducted assuming an additive genetic model. Imputed variants with MAF < 0.01 were excluded using the BCFtools program [74]. For each association analysis, two different statistical approaches were considered: the frequentist single-locus approach and the Bayesian multi-locus approach. Within the frequentist single-locus approach, different models were employed based on the distribution of the traits. For the trait AHI, which has an approximately normal distribution, a standard linear mixed model (LMM) was fitted using GEMMA 0.98.5 [75]. For the count-based traits BC and TFC, LMMs in GEMMA were also applied, recognizing that this approach assumes a normal trait distribution. Additionally, all three traits were analyzed using GMMAT 1.4.2 [76] with a GMMAT LMM fitted for AHI and a Poisson generalized linear mixed model (GLMM) applied for BC and TFC to account for their count-based distributions. The Poisson GLMM in GMMAT was specifically chosen for BC and TFC because it accurately models the non-normal distribution of count data, complementing the LMM analysis conducted in GEMMA.

In the Bayesian multi-locus approach, a Bayesian sparse linear mixed model (BSLMM) [77] was fitted in parallel for all analyzed traits. By intersecting the resulting sets of significant SNPs from the frequentists and Bayesian approaches, significant SNPs for each trait were consistently identified. In addition, a multivariate linear mixed model (mvLMM) was fitted to simultaneously analyze significantly correlated traits (AHI and TFC, as well as AHI and BC) to detect shared association signals between these traits.

To visualize the results, Manhattan plots were generated using the R package "qqman" [78] and "CM plot" [79].

Generalized Linear Mixed Model (GLMM) for count data using a Poisson distribution

The generalized linear mixed model (GLMM) with a Poisson distribution was fitted using GMMAT. The model can be expressed in the following form (Eqs. (1-3)):

$$\log(\mu_i) = \mathbf{W_i}\alpha + \mathbf{x_i}\beta + u_i \tag{1}$$

$$u \sim MVN_n(0,\ \lambda\mathbf{K}) \tag{2}$$

$$y_i \sim \text{Poisson}(\mu_i) \tag{3}$$

Here, y_i denotes the observed count for the i-th individual, and μ_i represents the mean count, which is modeled as the exponential of the linear predictor. $\mathbf{W_i}$ is the i-th row of an $n \times c$ matrix of covariates (fixed effects), α is the corresponding vector of coefficients for these covariates, $\mathbf{x_i}$ represents the genotype of the i-th individual, and β is the effect size of the genetic marker. The random effects u are assumed to follow a multivariate normal distribution MVN_n (0,λK), where K is the $n \times n$ relatedness matrix, and λ represents the variance component ratio. The observed data y_i is assumed to follow a Poisson distribution with μ_i. This model allows for integrating individual-level random effects and a genetic relationship matrix $\mathbf{K}$ to account for population structure and relatedness while analyzing count-based traits. If a normal distribution and an identity link function are assumed for continuous traits, GMMAT performs association tests based on linear mixed models (LMMs).

Linear Mixed Model (LMM)

The standard LMM was fitted using GEMMA 0.98.5. in the following form:

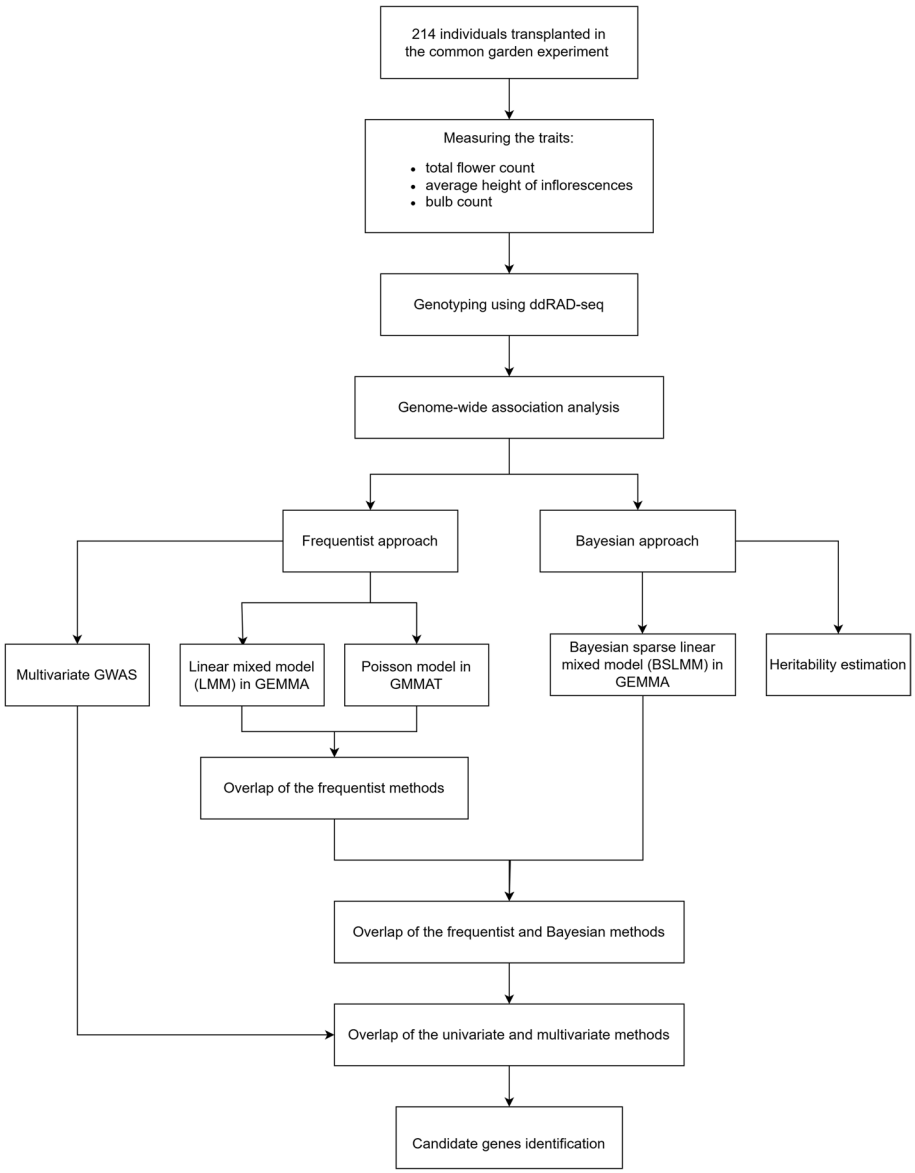


Fig. 2 Outline of the methodological approach used to investigate the genetic basis of reproduction-related traits in *Chouardia litardierei*

$$\mathbf{y} = \mathbf{W}_{\alpha} + \mathbf{x}\beta + \mathbf{u} + \varepsilon \quad (4)$$

$$\mathbf{u} \sim MVN_n\left(0,\ \lambda\tau^{-1}\mathbf{K}\right) \quad (5)$$

$$\varepsilon \sim MVN_n\left(0,\ \tau^{-1}\mathbf{I_n}\right) \quad (6)$$

where we let **y** be a vector of trait values for 214 individuals and **W** be an $n \times c$ matrix of covariates (fixed effects), which in our case is a column of 1 s. Let α represent a c-vector of the intercept, **x** be an n-vector of marker genotypes, and β denotes the effect size of the marker. Additionally, **u** is an n-vector of random effects, ϵ is an n-vector of errors, τ^{-1} represents the variance of the residual errors, and λ is the ratio between the two variance components. **K** is a known $n \times n$ relatedness matrix, and $\mathbf{I}_n$ is an $n \times n$ identity matrix. MVN_n denotes the n-dimensional multivariate normal distribution. Effect sizes represent the change

in trait levels for each additional effect allele in the genotypes of individuals.

Bayesian framework

The LMM (Eqs. (4-6)) implemented in GEMMA tests the alternative hypothesis H_1: $\beta \neq 0$ against the null hypothesis H_0: $\beta = 0$ for each SNP individually. Extensions of the LMM that simultaneously consider the effects of variants across multiple loci could enhance the power to detect causal variants. Bayesian LMMs can model all markers together by assuming different prior distributions on the marker effects and sampling from their posterior distribution. Bayesian models developed for estimating SNP effect sizes begin with a simple linear model that relates genotypes X to phenotypes y:

$$\mathbf{y} = \mathbf{1_n}\mu + \mathbf{x}\beta + \varepsilon \tag{7}$$

$$\varepsilon \sim MVN_n(0, \tau - 1\mathbf{I_n}) \tag{8}$$

we let $\mathbf{y}$ be a vector of phenotypes measured on n individuals and $\mathbf{X}$ be an $n \times p$ matrix of genotypes measured on these same n individuals at p genetic markers. The vector β represents the effects of genetic markers, $\mathbf{1_n}$ is an n-vector of 1 s, μ is a scalar representing the mean phenotype, and ϵ is an n-vector of error terms with variance τ^{-1}. Our goal was to estimate the parameter β, representing the effects of the genetic markers. However, since the number of genetic markers p in our study (23,315) greatly exceeds the number of individuals n (214), we needed to make certain modelling assumptions for SNP effect sizes β. These assumptions range from the infinitesimal (or polygenic) model, which assumes that all SNPs have non-zero effects, to the sparse model, which assumes that only a small proportion of SNPs affect the phenotype. The model's performance depends on the true underlying genetic architecture of the trait being studied. However, this genetic architecture is generally unknown. The most commonly used polygenic modeling approach assumes that all SNPs influence the phenotype (i.e., have non-zero effects) with normally distributed effect sizes:

$$\beta \sim N\left(0, \sigma_\beta^2\right) \tag{9}$$

When Eqs. (7 and 8) are combined with the normality assumption (Eq. (9)) for effect sizes β, results in the previously mentioned LMM due to the inclusion of a random effect term representing the combined genetic effects.

Bayesian Sparse Linear Mixed Model (BSLMM)

A broader assumption, encompassing both polygenic and sparse modeling scenarios, posits that effect sizes originate from a combination of two normal distributions.

$$\beta_i \sim \pi \mathrm{N}\left(0, \frac{\sigma_\alpha^2 + \sigma_\mathrm{b}^2}{\mathrm{p}\tau}\right) + (1 - \pi)\mathrm{N}\left(0, \frac{\sigma_\mathrm{b}^2}{\mathrm{p}\tau}\right) \tag{10}$$

In this model, π represents the proportion of SNPs with large effects, while $\sigma_\beta^2/\mathrm{p}\tau$ and $\sigma_\alpha^2/\mathrm{p}\tau$ represent the small and large effects variances, respectively. The resulting model, BSLMM, incorporates a combination of polygenic and sparse effects for the prior distribution of effect sizes, enabling adaptation to various genetic architectures of the studied traits. BSLMM accounts for relatedness and population stratification by including a genomic kinship matrix as a random effect term, and it handles linkage disequilibrium (LD) by estimating SNP effect sizes β while controlling for other SNPs in the model. The model uses a Markov chain Monte Carlo algorithm to sample from the posterior distribution and obtain SNP effect sizes. Unlike LMM, which provides p-values, BSLMM outputs a posterior inclusion probability (PIP) for each SNP, indicating the probability that a marker is associated with the trait, given the data, calculated as the proportion of chain iterations in which the SNP has a large effect. SNPs with high PIPs are the most likely candidates for functional variants affecting the analysed traits. We applied BSLMM to the same dataset (214 individuals and 23,315 variants) used in our primary frequentist association analysis to compare single-SNP and multi-SNP approaches and to reduce false positives. The BSLMM chain was run with 1,000,000 sampling steps and 100,000 burn-in iterations. We used the estimated PIPs from BSLMM for additional fine-mapping of genomic regions identified in the frequentist analysis.

SNP heritability estimation

The proportion of variance in phenotypes explained by all available genotypes (PVE), also known as narrow-sense heritability (h^2), as well as the proportion of genetic variance explained by variants with major effect (PGE), was estimated for traits listed in Table 1. This estimation was conducted assuming that the SNP effect sizes follow a mixture of two normal distributions (Eq. 10), as implemented in GEMMA BSLMM.

Multivariate genome-wide association analyses

To identify common variants associated with the AHI and the TFC traits, multivariate genome-wide association analyses were conducted using a multivariate linear mixed model (mvLMM) in GEMMA. Similarly,

multivariate GWAS was conducted using the same model for AHI and BC traits. This approach allowed for the simultaneous examination of genetic influences on both pairs of traits by considering them as dependent variables. The mvLMM method accounts for population structure and relatedness among individuals, ensuring robust identification of genetic variants contributing to the observed phenotypic variation in these traits.

Candidate genes prediction

Following identifying phenotypic evidence for local adaptation to diverse conditions in distinct *C. litardierei* populations and subsequent GWAS analysis, further efforts were directed toward identifying associated candidate genes. Utilizing the reference genome, sequences were generated encompassing a total of 50 kilobases, including 25 kilobases upstream and downstream of each significant SNP identified through both statistical models, using SAMtools [74]. Finally, functional annotations were obtained using the eggNOG-mapper v2 database (*e*-value $< 1 \times 10^{-2}$).

Results

Phenotyping

PCA based on the bioclimatic conditions studied populations are experiencing in their natural habitats showed exceptional diversity among sites. In addition, the common garden experiment site was equally environmentally differentiated from the sampling sites of studied populations, making it suitable for the purpose. The PCA results are provided in Additional File 2.

Phenotypic variations among *C. litardierei* populations in the common garden experiment are illustrated in Fig. 3.

For TFC per genotype, out of 214 individuals across nine populations, 204 flowered. Overall, the mean count of flowers across all inflorescences per genotype was 90.59 ± 45.59. The AHI per genotype was assessed across a cohort of 204 flowering individuals. Overall, the mean height of inflorescences per genotype was 18.00 ± 4.80 cm. The number of bulbs developed (BC), considered a very important indicator of asexual reproduction, had a mean count of 2.78 ± 3.06 bulbs per genotype. All the data mentioned above are summarized in Table 1.

A positive Pearson's correlation coefficient was observed between the TFC and the AHI traits ($r = 0.445$, p-value < 0.001, 95% CI [0.327, 0.549]). Similarly, a positive correlation was observed between the BC and the AHI traits (Spearman's $\rho = 0.172$, p-value $= 0.014$). At the same time, a weak negative correlation was observed between the BC and the TFC traits. However, this correlation was not statistically significant (Spearman's $\rho = -0.102$, p-value $= 0.146$).

Sequencing and genomic data processing

A total of 1,284,680,304 reads were obtained from the sequencing. After filtering the raw sequences and annotating against the reference genome, 1,278,409,966 reads were retained. SNP calling and filtration were performed using Stacks software, resulting in the identification of 24,660 SNP loci, which were subsequently processed. After applying the BCFtools MAF filter with a threshold of 0.01, 23,315 SNPs remained for further analysis.

Population-genetic and genome-environment association analysis

Results obtained using different approaches were highly congruent, thus supporting their high reliability. PC analysis revealed that populations from meadow and seashore habitats were genetically indistinguishable, forming a compact genetic cluster. At the same time, populations from rocky habitats were strongly differentiated, both from the populations from other habitats and each other (Fig. 4).

As in the sNMF analysis the selection of the final K number is somewhat arbitrary, we present results for both K = 2 and K = 3 as the two most reliable numbers of ancestral populations (Additional File 3). In both cases, populations from the meadow and seashore habitats were grouped together, forming a separate cluster without substantial admixture levels among populations, as shown in Fig. 5. For K = 2, dry-habitat populations form an individual cluster. However, this cluster was further structured at the K = 3 level, with the Pandurica population being differentiated from the remaining two populations.

We assessed Nei's inter-populations genetic distances to investigate the phylogenetic relationships among the studied populations, and we constructed the unrooted tree to visualize the results (Additional File 4). Once again, populations from the dry, rocky habitats have shown very strong differentiation from others characterized by substantially weaker differentiation levels. All nodes on the phylogenetic tree were statistically well supported.

After the variance inflation factors analysis, four temperature-related (BIO2—mean diurnal range, BIO4—temperature seasonality, BIO8—mean temperature of wettest quarter, and BIO9—mean temperature of driest quarter) and two precipitation-related variables (BIO17—precipitation of driest quarter and BIO19—precipitation of coldest quarter) were retained for further analysis. The RDA model was globally significant ($p <$ 0,001) and explained as much as 52.26% of the total variance (adjusted R2 = 0.509). The first RDA axis explained the majority (40.49%) of this variation, while the second explained a substantially smaller portion of just 4.94%. Consequently, most tested bioclimatic

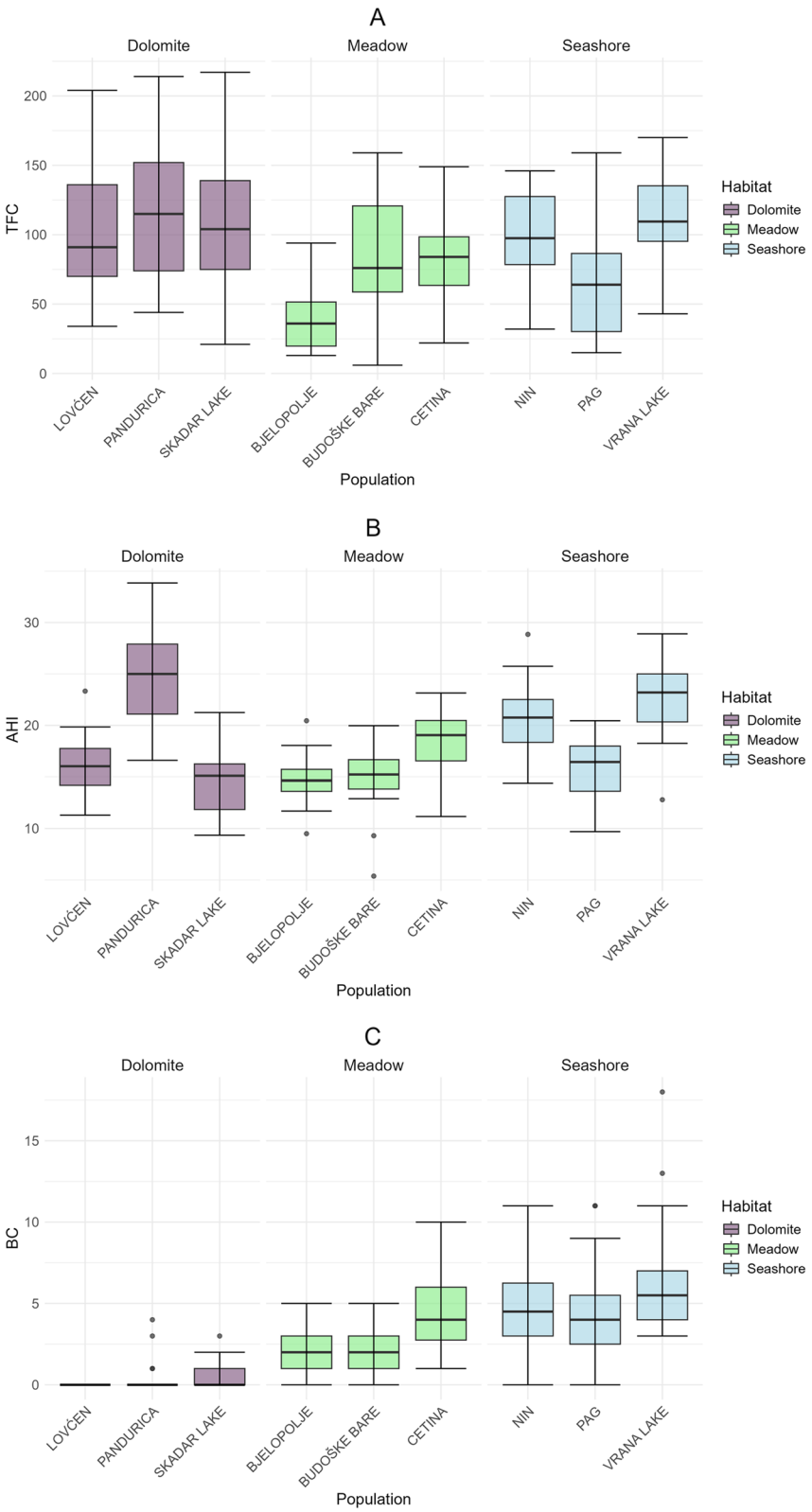


Fig. 3 Box plots illustrate the obtained morphometric results from a common garden experiment, depicting three reproduction-related morphological traits: **A** Total Flower Count (TFC), **B** Average Height of Inflorescences (AHI), and **C** Bulb Count (BC) per genotype. Each box represents the interquartile range (IQR), with the horizontal line inside the box indicating the median. Whiskers extend to data points within 1.5 times the IQR, while dots represent outliers

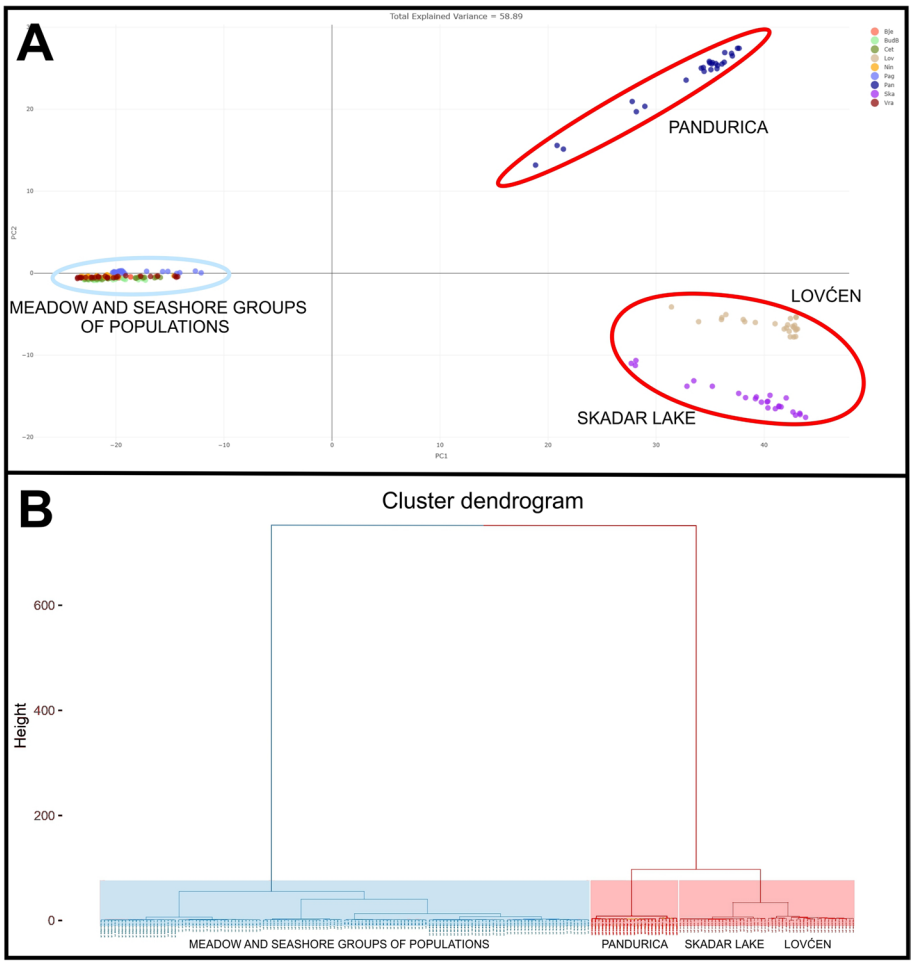


Fig. 4 **A** PCA visualization of studied *Chouardia litardierei* populations. Each dot represents a single sample. Ellipses indicate clusters identified by hierarchical clustering analysis. Blue ellipse encircles the meadow and the seashore populations, and red ones the dolomite populations. **B** Hierarchical cluster dendrogram of obtained PCA data of studied *C. litardierei* populations. In blue are individuals from the meadow and seashore populations, and in red are individuals belonging to the dolomite group of populations

variables were significantly associated with the RDA axis 1, with BIO19 being recognized as the variable with the most profound influence. Following such a result, most loci (131 out of 256) being recognized as outliers were linked to the BIO19 variable (precipitation of coldest quarter). In contrast, substantially fewer were linked to the remaining variables: two to BIO2, 65 to BIO4, six to BIO8, and 52 to BIO9, while none was linked to BIO17. The results of this analysis are visually represented in Fig. 6.

To assess the potential biological functions of genes surrounding 83 recognized loci, we compared the obtained 50 kb DNA windows to the eggNOG database. Querying of annotations revealed 324 genes linked to a specific metabolic function. After manually inspecting individual genes, we retained 82 with recognized biological functions seemingly associated with adaptation to tested bioclimatic variables (Additional File 5).

Genome-wide association analyses

GEMMA detected 26 significant SNPs for the AHI trait, while GMMAT identified 34. Overlapping these results revealed 26 common genome-wide significant SNPs. Subsequent analysis with BSLMM confirmed four of these SNPs as significant, one on chromosome 3 and three on chromosome 13. Similarly, for the TFC trait, GEMMA and GMMAT identified 18 and 43 SNPs, respectively, with nine overlapping SNPs. BSLMM analysis confirmed only one significant SNP on chromosome 1. In the case of the BC trait, GEMMA and GMMAT identified 86 and 96 SNPs, respectively, with 85 overlapping SNPs. BSLMM analysis confirmed seven significant SNPs, with three located on chromosome 13 and one SNP on each of chromosomes 1, 6, 9, and 12. All SNPs passing the genome-wide significance threshold (1×10^{-3}) in the single-SNP LMM analysis are reported in Additional File 6.

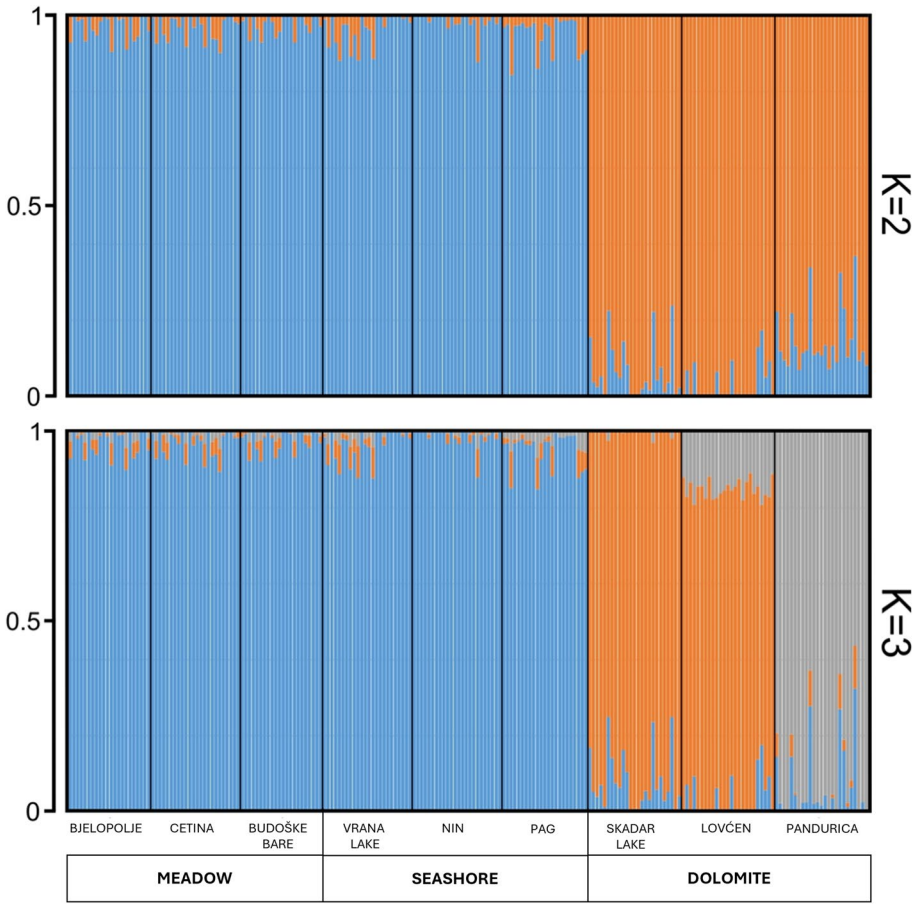


Fig. 5 Genetic structure of the studied *Chouardia litardierei* populations as determined by sNMF analysis at K = 2 and K = 3. Each stacked column represents an individual's ancestry coefficient, with populations separated by lines. Population names are labeled along the x-axis

In the Bayesian association analysis, 21 SNPs were identified as having a major sparse effect on the AHI trait, and these variants were estimated to have a sparse effect in ≥ 10% of BSLMM chain iterations (i.e., posterior inclusion probability, PIP ≥ 0.099). Moreover, the top five SNPs were identified as having a sparse effect on AHI in more than 20% of chain iterations (PIP > 0.21). Similarly, for the TFC trait, nine SNPs displayed a major sparse effect in ≥ 10% of BSLMM chain iterations (PIP ≥ 0.095). In addition, the top three SNPs displayed a major sparse effect in more than 12% of iterations (PIP ≥ 0.12). Concerning the BC trait, 14 SNPs were identified with a major sparse effect in ≥ 10% of iterations (PIP ≥ 0.097), and the top six SNPs had a major sparse effect in over 55% of iterations (PIP ≥ 0.55). The data outlined above is reported in Additional File 7.

Results from the single-SNP association analysis in GMMAT and GEMMA, alongside the multi-SNP association analysis (BSLMM) for all of the studied traits, are plotted in parallel in Manhattan plots in Fig. 7. Twelve SNPs reached genome-wide significance ($p = 1 \times 10^{-3}$) in the LMM analysis, mirroring their major sparse effects identified in the BSLMM analysis (Table 2).

Heritability estimation

The BSLMM analysis, conducted with 23,315 SNPs, yielded estimates of narrow-sense heritability (PVE) for the examined reproduction-related morphological traits, along with the PGE and the number of variants with major effect (n.gamma), as summarised in Table 3. The PVE estimate for the TFC revealed that 55.89% of the phenotypic variation in TFC was explained by all available genotypes, with 28.78% attributed to 78 SNPs exhibiting significant phenotypic effects. Similarly, the PVE estimate for the AHI indicated that 71.95% of the phenotypic variation in AHI was explained by all genotypes, with 37.47% attributed to 47 SNPs exhibiting notable phenotypic effects. Moreover, the BSLMM analysis revealed that 69.87% of the phenotypic variation in BC was explained by all genotypes, with 89.15% of this variation accounted for by 18 SNPs with significant effects. Additional File 8 contains the means, medians, and 95%

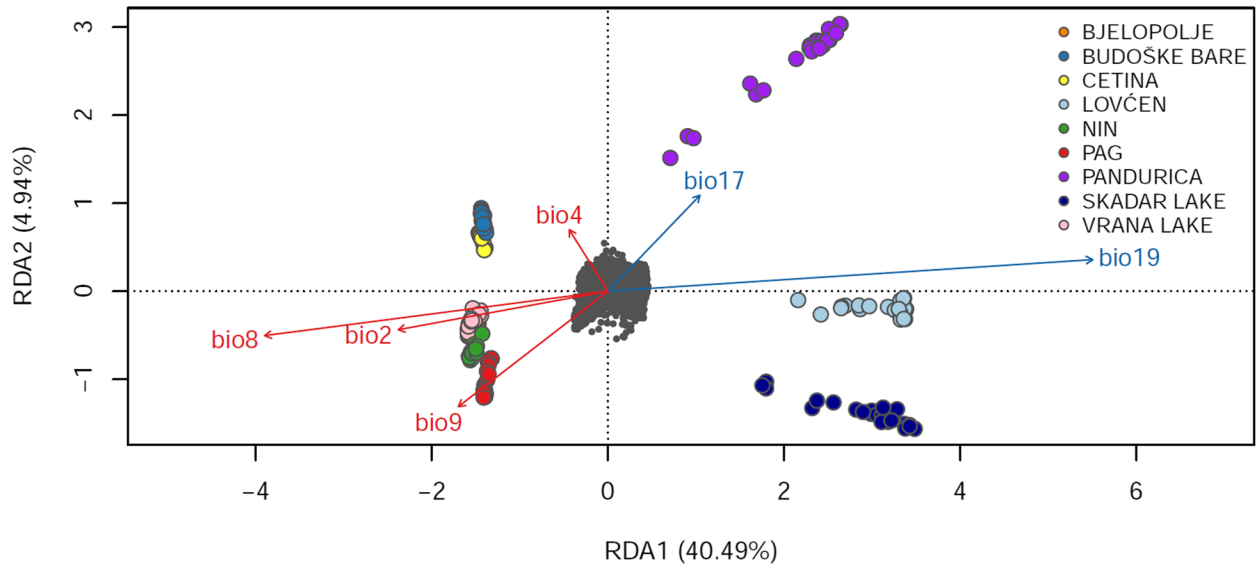


Fig. 6 A triplot based on six bioclimatic variables included in the optimal RDA model illustrating the relative contribution of bioclimatic variables in shaping the genetic structure of nine *C. litardierei* populations. Colored dots represent samples, while empty dots around the centre represent SNPs. Temperature-related variables (BIO2, BIO4, BIO8, and BIO9) are shown as red vectors, and precipitation-related variables (BIO17 and BIO19) as blue vectors

equal tail posterior probability intervals (95% ETPPIs) of the hyperparameters derived from the BSLMM.

Multivariate GWAS analysis

In the multivariate GWAS analysis, 42 SNPs surpassed the genome-wide significance threshold ($p = 1 \times 10^{-3}$) for AHI and TFC traits (Additional File 9). Among these, 10 SNPs were significant in GEMMA and GMMAT univariate analyses for the AHI trait, while only three SNPs showed significance for the TFC trait (Table 4). In the multivariate GWAS analysis for AHI and BC traits, 64 SNPs exceeded the same threshold (Additional File 10). Among these, two SNPs were significant in GEMMA and GMMAT univariate analyses for the BC trait, while none showed significance for the AHI trait (Table 4). This indicates shared genetic factors influencing these reproductive traits across multivariate and univariate analyses. The multivariate GWAS findings for the AHI and TFC, as well as AHI and BC traits, are plotted in Manhattan plots in Fig. 8. The frequencies of effect alleles across populations for the significant SNPs (shown in Tables 2 and 4) are depicted in a plot provided in Additional File 11.

GWAS candidate genes identification

The eggNOG tool provided comprehensive data elucidating the relationship between individual SNPs/sequences and distinct protein families (PFAM). To identify candidate genes potentially influencing reproduction-related morphological traits, we conducted eggNOG analysis on 12 SNPs that exceeded the genome-wide significance threshold (1×10^{-3}) in both the single-SNP LMM and multi-SNP BSLMM analyses of *C. litardierei* traits: TFC, AHI, and BC. This analysis identified 130 queries corresponding to sequences matched to the eggNOG database for functional annotation (Additional File 12). We utilized eggNOG to analyze 13 SNP loci that met the significance threshold in the multivariate GWAS analysis for AHI and TFC. This analysis identified 134 queries (Additional File 13) corresponding to sequences associated with functional roles in reproduction-related morphological traits. Similarly, eggNOG was employed to analyze 2 SNP loci meeting the same threshold in the multivariate GWAS analysis for AHI and BC, uncovering 18 additional queries (Additional File 14). The eggNOG analysis linked identified sequences to protein families, which we then further examined through manual inspection and a literature review to identify specific genes and PFAM domains associated with the traits under study. Some domains were shared between the univariate and multivariate GWAS results, leading to overlaps across the sets. The most relevant findings, along with their relevant biological functions and references, are summarized in Table 5.

Discussion

In our research, we took several approaches to gain insight into the genetic background of local adaptation of *C. litardierei* populations inhabiting contrasting

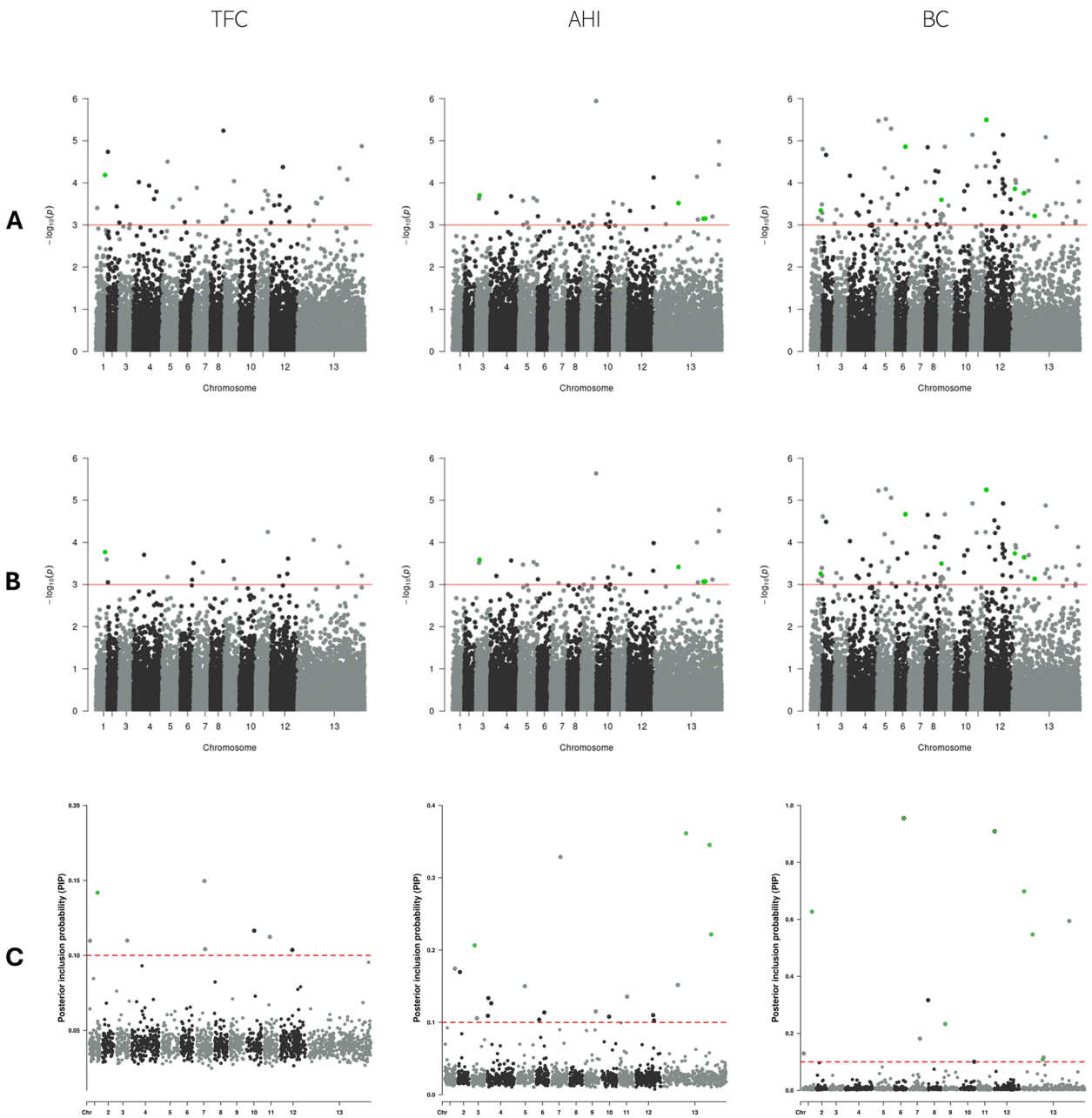


Fig. 7 Manhattan plots of single-SNP and multi-SNP association mapping of TFC, AHI and BC traits. **A** Manhattan plots of single-SNP analysis in GMMAT and **B** in GEMMA for each trait. The x-axis represents the chromosomal position of SNPs, and the y-axis represents their $-\log_{10}$ (*p*-values) obtained by the LMM analysis. The red horizontal line indicates the genome-wide significance threshold ($p = 1 \times 10^{-3}$). Each dot on the Manhattan plot signifies a SNP. Because the strongest associations have the smallest *p*-values, their negative logarithms will be the greatest, appearing higher on the plot. **C** Manhattan plots of multi-SNP BSLMM analysis for each trait. The x-axis represents the chromosomal position of SNPs, and the y-axis represents their posterior inclusion probabilities (PIPs) obtained by the BSLMM analysis. Green dots signify SNPs that are recognized in all three models

types of habitats. First, we coupled population-genetic analysis with the GEA study to characterize the genetic structure of studied populations, identify the bioclimatic variables predominantly influencing detected variability, and finally gather knowledge regarding the molecular mechanisms underlying populations' ability to cope with contrasting ecological conditions. Then, based on the common garden experiment, a comprehensive GWAS analysis was performed to elucidate the genetic background of heritable reproduction-related traits. This

Table 2 SNPs passing genome-wide significance threshold (1×10^{-3}) in the single-SNP LMM analysis and their corresponding PIPs from the multi-SNP BSLMM analysis of *Chouardia litardierei* traits TFC, AHI and BC

Trait	SNP	Chr	Position	Effect Allele	Referent Allele	MAF	Single-SNP GLMM Analysis β (p-value) in GMMAT	Single-SNP LMM Analysis β (p-value) in GEMMA	Multi-SNP BSLMM Analysis β (PIP)
TFC	439681_33	1	113,066,723	G	A	0.35	−0.89 (6.50×10^{-5})	−1.29 (1.69×10^{-4})	−0.53 (0.14)
AHI	536624_21	3	40,433,867	G	T	0.02	−1.02 (1.98×10^{-4})	−1.02 (2.57×10^{-4})	−0.41 (0.21)
AHI	299462_80	13	294,757,456	T	C	0.04	−0.67 (3.02×10^{-4})	−0.67 (3.81×10^{-4})	−0.43 (0.36)
AHI	383241_14	13	626,409,144	A	G	0.03	−0.73 (7.01×10^{-4})	−0.73 (8.43×10^{-4})	−0.42 (0.22)
AHI	377817_17	13	604,813,128	A	G	0.13	−0.62 (7.09×10^{-4})	−0.62 (8.51×10^{-4})	−0.43 (0.35)
BC	241986_29	12	6,101,088	G	A	0.01	1.28 (3.19×10^{-6})	1.28 (5.63×10^{-6})	0.94 (0.91)
BC	713226_25	6	119,601,819	T	A	0.06	0.50 (1.38×10^{-5})	0.50 (2.15×10^{-5})	0.45 (0.96)
BC	291775_18	13	26,477,605	T	C	0.02	0.72 (1.39×10^{-4})	0.72 (1.82×10^{-4})	0.57 (0.70)
BC	260150_22	13	138,653,995	G	A	0.35	−0.93 (1.75×10^{-4})	−0.93 (2.26×10^{-4})	−0.65 (0.55)
BC	64746_61	9	23,100,056	T	A	0.50	0.33 (2.51×10^{-4})	0.33 (3.17×10^{-4})	0.26 (0.23)
BC	441718_55	1	121,120,631	G	A	0.04	0.48 (4.50×10^{-4})	0.48 (5.49×10^{-4})	0.43 (0.63)
BC	293695_19	13	272,369,031	A	G	0.03	0.66 (6.08×10^{-4})	0.66 (7.30×10^{-4})	0.42 (0.11)

Statistical analyses were performed with GEMMA and GMMAT LMM, GLMM and BSLMM. p-values $< 1 \times 10^{-3}$ are considered genome-wide significant

AHI Average Height of Inflorescences, *BC* Bulb Count, *BSLMM* Bayesian Sparse Linear Mixed Model, *Chr* Chromosome, *GLMM* Generalized linear Mixed Model, *LMM* Linear Mixed Model, *MAF* Minor Allele Frequency, *PIP* Posterior Inclusion Probability, *SNP* Single Nucleotide Polymorphism, *TFC* Total Flower Count

Table 3 Genetic architectures of *Chouardia litardierei* reproduction-related morphological traits obtained using the BSLMM

Trait	PVE/%	PGE/%	n.gamma
TFC	55.98	28.78	78
AHI	71.95	37.47	47
BC	69.87	89.15	18

AHI Average Height of Inflorescences, *BC* Bulb Count, *n.gamma* number of variants with major effect, *PGE* Proportion of Variance Explained by major effect variants, *PVE* Proportion of Variance Explained by genetic data, *TFC* Total Flower Count

way, we provided comprehensive coverage of molecular mechanisms involved in the ecology-driven differentiation process observed among *C. litardierei* populations.

Population genetic structure

Population-genetic analyses only partially confirmed the assumed genetic structuring of the studied populations, where we anticipated that ecological differentiation would be coupled with the genetic one. While the populations from the dry, drought-prone habitats formed a separate, well-differentiated group, the same was not the case with the remaining two groups of populations, the one from the inland meadow habitats and the other from the seashore habitats. These populations were genetically indistinguishable on the level of presumed ecotypes and the individual population level as well (Figs. 4 and 5). Such results suggest they either recently originated from a common ancestral population or are experiencing profound contemporary inter-population gene flow, which acts against any substantial differentiation [80, 81]. In contrast, dolomite-habitat populations were characterized by high inter-population differentiation levels, which can likely be explained by strong fragmentation and patchiness of their habitat and subsequent lack of gene flow among them. However, although these robust results undoubtedly point to the general genetic structure and phylogenetic relationships among studied populations from different habitats, substantially more populations from across the entire species' distribution range should be included for more reliable and comprehensive results.

Genome-environment association analysis

We performed RDA to understand better the genetic mechanisms enabling the local adaptation of *C. litardierei* populations to specific bioclimatic conditions across the species distribution range. Due to the ubiquitous nature of environmental correlations, interpreting the obtained RDA result can easily lead to misleading conclusions [82]. Therefore, we observe the obtained results only as general patterns of local adaptation-related mechanisms and focus more on the genetic aspect of the obtained results rather than on details regarding specific bioclimatic variables. Of the tested variables, BIO19 (precipitation of the coldest quarter) was recognized as the most profound driver of the detected variation. Consequently, it is unsurprising that most outliers were linked to this variable. The functional annotation of genomic regions surrounding outliers identified numerous candidate genes potentially involved in local adaptations, including PFAM domains linked to stress responses and key

Table 4 SNPs passing genome-wide significance threshold (1×10^{-3}) in the multivariate GWAS mvLMM analysis of *Chouardia litardierei* reproduction-related morphological traits AHI and TFC, and AHI and BC. Listed SNPs were significant in both GEMMA and GMMAT univariate analyses

Trait	SNP	Chr	Position	Effect Allele	Ref. Allele	MAF	Beta1 (AHI)	Beta2 (TFC)	mvLMM in GEMMA (*p*-value)
AHI + TFC	62254_22	9	181,728,136	C	A	0.04	−1.02	−0.65	5.17×10^{-6}
AHI + TFC	423028_13	13	795,057,155	A	C	0.03	−0.80	−0.31	3.90×10^{-5}
AHI + TFC	230454_16	12	319,053,038	A	T	0.14	−0.53	−0.51	1.53×10^{-4}
AHI + TFC	423027_44	13	795,056,910	T	A	0.03	−0.76	−0.35	1.56×10^{-4}
AHI + TFC	730317_18	6	4,133,503	C	A	0.36	−0.72	0.04	1.76×10^{-4}
AHI + TFC	356033_30	13	518,039,464	A	C	0.35	−0.25	−0.74	1.98×10^{-4}
AHI + TFC	357122_13	13	523,783,140	C	G	0.27	0.38	0.31	2.77×10^{-4}
AHI + TFC	45968_38	9	118,116,906	C	G	0.06	−0.02	0.56	3.49×10^{-4}
AHI + TFC	275195_16	13	197,688,818	C	T	0.14	−0.22	−0.49	3.77×10^{-4}
AHI + TFC	593460_76	4	257,849,493	A	C	0.04	−0.64	−0.10	4.60×10^{-4}
AHI + TFC	669910_120	5	218,775,782	A	G	0.14	−0.41	−0.11	5.88×10^{-4}
AHI + TFC	679100_46	5	47,723,772	G	A	0.03	−0.82	−0.47	9.56×10^{-4}
AHI + TFC	299462_80	13	294,757,456	T	C	0.04	−0.68	−0.43	9.92×10^{-4}
							Beta1 (AHI)	**Beta2 (BC)**	
AHI + BC	178892_42	12	113,762,962	A	G	0.29	0.40	0.22	3.92×10^{-4}
AHI + BC	22031_53	8	25,314,160	A	G	0.29	0.14	1.07	9.88×10^{-4}

Statistical analyses were performed with GEMMA mvLMM. *p*-values $< 1 \times 10^{-3}$ are considered genome-wide significant

AHI Average Height of Inflorescences, *BC* Bulb Count, *Chr* Chromosome, *MAF* Minor Allele Frequency, *mvLMM* multivariate Linear Mixed Model, *SNP* Single Nucleotide Polymorphism, *TFC* Total Flower Count

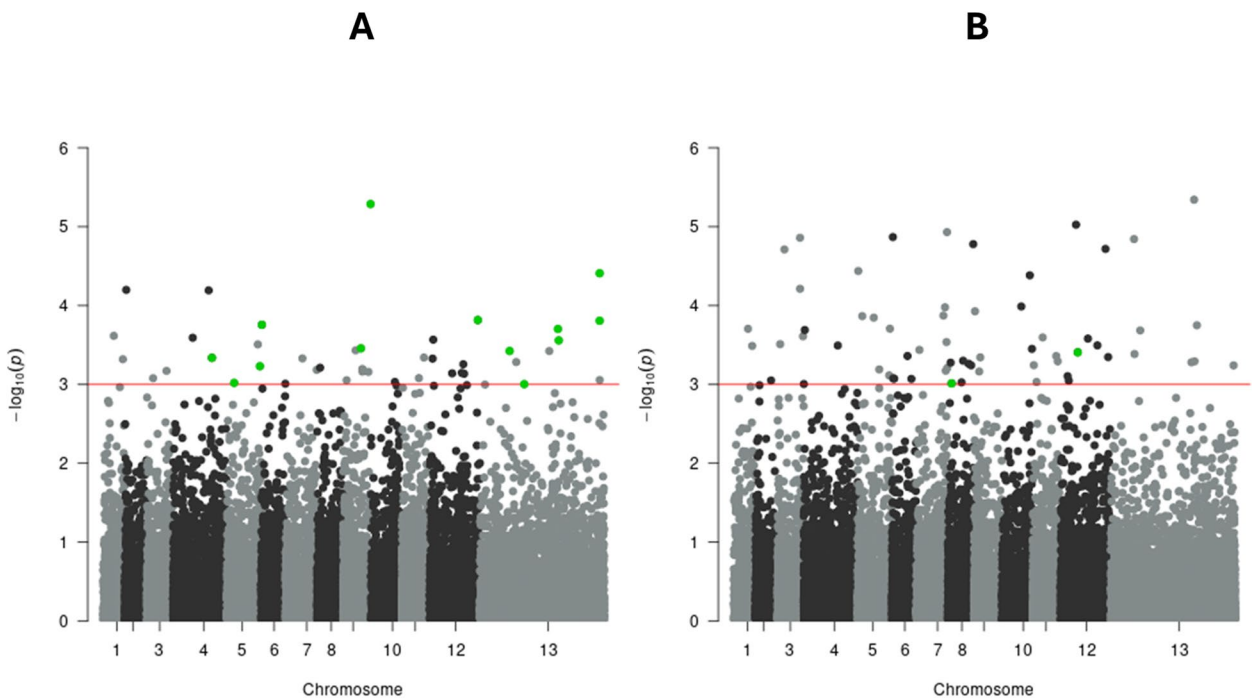


Fig. 8 Manhattan plot of multivariate genome-wide association study of (**A**) AHI and TFC traits and (**B**) AHI and BC traits. The red horizontal line indicates the genome-wide significance threshold ($p = 1 \times 10^{-3}$). Each dot on the Manhattan plot signifies a SNP. The strongest associations have the smallest *p*-values, so their negative logarithms will be the greatest, appearing higher on the plot. Green dots represent SNPs identified as significant in the multivariate GWAS analysis and in both GEMMA and GMMAT univariate analyses for each of the two plots

Table 5 List of genes and PFAM domains for regions of strong association with AHI, TFC, and BC identified by the eggNOG-mapper v2 database (*e*-value $< 1 \times 10^{-2}$) in *Chouardia litardierei*, based on the 12 recognized SNPs passing the genome-wide significance threshold ($p = 1 \times 10^{-3}$) in the single-SNP LMM and multi-SNP BSLMM analysis, 13 SNPs passing the same threshold in the multivariate GWAS mvLMM analysis of AHI and TFC (mGWAS1), and two SNPs passing the threshold in the multivariate GWAS mvLMM analysis of AHI and BC (mGWAS2). Names of identified candidate genes associated with the SNPs and PFAMs, along with their relevant biological functions and references, are provided

Query	Method	*e*-value	Chr	EGGNog PFAM	Candidate Genes	Species	Relevant biological functions	References
H 16:294,732,456-294782456_3	GWAS	$1.34e^{-40}$	13	Arginase family (ARG)	OsARG, ARGH1, ARGH2	*Oryza sativa*, *Arabidopsis thaliana*	Integral to nitrogen metabolism, amino acid metabolism, and photosynthesis	[102]
H 16:294,732,456-294782456_5	mGWAS1	$1.85e^{-49}$						[104]
H 16:604,788,128-604838128_3	GWAS	$1.88e^{-10}$	13	Cytochrome P450 (CYP450)	CYP701 A8, MdCYP716B1, CYP88	*Oryza sativa*, *Malus domestica*, *Hordeum vulgare* and *Zea mays*	Regulates biosynthesis and catabolism of phytohormones and metabolites, impacts plant stature, height, and bulb growth	[112–114]
H 16:294,732,456-294782456_13	GWAS	$4.07e^{-21}$	13	Complex 1	NDUFV1	*Arabidopsis thaliana*	Essential for growth and development at all stages	[119]
H 16:294,732,456-294782456_14	mGWAS1	$4.07e^{-21}$						
H 16:26,452,605-26502605_43	GWAS	$1.01e^{-212}$	13	CCHC-ZFP genes	TaCCHC-ZFP	*Triticum aestivum*	Regulates phytohormones and metabolites	[117]
H 16:795,032,155-795082155_54	mGWAS1	$4.28e^{-273}$						
H 10:25,289,160-25339160_2	mGWAS2	$2.45e^{-99}$	8					
H 3:40,408,867-40458867_73	GWAS	$2.06e^{-62}$	3	Aspartic proteases (APs)	PhAP	*Phyllostachys edulis*	Supports rapid growth and organ development	[118]
H 5:47,698,772-47748772_17	mGWAS1	$7.25e^{-47}$	5					
H 12:23,075,056-23125056_50	GWAS	$2.11e^{-50}$	9	Protein tyrosine kinase (PTK)	OsPTK2, OsPTK8, OsPTK13, OsPTK14, OsPTK18	*Oryza sativa*	Involved in abiotic stress tolerance, including cold, heat, and submergence	[123]
H 15:319,028,038-319078038_55	mGWAS1	$2.72e^{-29}$	12					
H 8:119,576,819-119626819_36	GWAS	$2.16e^{-143}$	6	C2 domain	QUIRKY, STRUBBELIG	*Arabidopsis thaliana*	Promotes intercellular communication and tissue morphogenesis	[124]
H 16:518,014,464-518064464_15	mGWAS1	$1.48e^{-85}$	13					
H 16:294,732,456-294782456_44	GWAS	$1.48e^{-137}$	13	Receptor-like protein kinases (RLK)	BRI1	*Arabidopsis thaliana*	Regulate numerous aspects of plant growth and development	[120–122]
H 10:25,289,160-25339160_15	mGWAS2	$1.43e^{-57}$	8	Sterol synthase	FACKEL (FK)	*Arabidopsis thaliana*	Regulates membrane integrity, cell division, and tissue patterning during plant development	[125]
H 10:25,289,160-25339160_17	mGWAS2	$9.51e^{-21}$	8	Sugar transporters	LohSTP8, LohSTP12, LflERD6.3	*Lilium spp.*	Essential for bulb formation	[126]

AHI Average Height of Inflorescences, *BC* Bulb Count, *Chr* Chromosome, *GWAS* Genome-Wide Association Study, *H* HiC scaffold, *mGWAS1* multivariate Genome-Wide Association Study of AHI and TFC, *mGWAS2* multivariate Genome-Wide Association Study of AHI and BC, *PFAM* Protein Family, *TFC* Total Flower Count

physiological processes. For example, we recognized the C2 domain has been shown to boost salt stress tolerance in soybean [83], the protein kinase domain that regulates Na^+/K^+ homeostasis under salt stress in *Arabidopsis* [84], and the MYB transcription factor that enhances stress tolerance under salt and water stress in sugarcane [85]. We also identified TPR16, which has been found to regulate stress responses and enhance drought tolerance in *Arabidopsis* [86] and START domain proteins that aid in drought stress signaling in chickpeas [87]. Additionally, we identified Rubisco, an enzyme whose activity is affected by heat stress, thereby limiting photosynthesis [88], while the recognized DEAD-box gene family and the RRM1 gene family in wheat and *Brassica rapa* enhance cold stress tolerance [89, 90]. PPR, OBERON, and ZF-C2H2 proteins were also identified as critical regulators of gene expression in growth and development, influencing RNA editing [91], interacting with WRKY factors [92], and modulating transcriptional networks [93].

Heritability and evolutionary significance of reproductive traits

Studying the genetic basis of adaptive traits presents significant challenges, mainly due to the complex interactions among polygenic backgrounds and diverse environmental factors [22]. Traits observed in natural settings may exhibit variability influenced by environmental conditions, rendering them less reliable for identifying local adaptation [44]. To mitigate the effects of phenotypic plasticity among the groups of studied populations of *C. litardierei*, we employed a common-garden experiment in which individuals from contrasting environments were cultivated under uniform conditions. This approach enabled us to differentiate genetic influences from environmental effects on trait expression [94], thereby elucidating aspects of the genetic background underlying the local adaptation process in the studied species.

Our morphometry analysis revealed substantial variations in reproduction-related traits among the studied populations but not the groups of populations from different habitats. Such results suggest these populations experience specific selection pressures in their surroundings, unrelated to the habitat types they originated from (i.e., seashore grasslands, karst poljes' meadows, and dolomite bedrocks). These findings are consistent with Exposito-Alonso et al. [95], who investigated *Arabidopsis* populations and found that the fitness heritability traits varied significantly between experimental sites due to contrasting natural selection pressures. Such variations across diverse environments contribute to developing the populations' adaptive potential and evolutionary trajectories, further reflected in our study's high PVE values. The substantial genetic contribution these values indicate underscores the heritable nature of reproductive traits in *C. litardierei*, emphasizing the solid genetic foundation critical for understanding evolutionary processes and local adaptation mechanisms. In contrast to moderate heritabilities previously reported for complex traits in some other taxa (e.g., *Pinus albicaulis* [96] and *Populus* [97]), our GWAS study identified a notably higher heritability for the average height of inflorescences (AHI). With PVE and PGE values of 71.95% and 37.47%, respectively, the high heritability of AHI suggests that inflorescence height, a trait important for pollination efficiency and reproductive success [98], holds significant evolutionary importance. Furthermore, the high PGE value for total flower count (TFC) emphasizes the critical role of major effect variants in shaping reproductive traits. This finding is consistent with studies in other species, such as *Silene dioica* and *Silene latifolia* [99], where high PGE values underscore the significant contribution of major effect loci to phenotypic variation in cumulative flowering. Although discussing our findings in the context of their biological meaning and importance for evolution and the local adaptation process in studied species is speculative, some assumptions can still be made. The trade-off between sexual and asexual reproduction is of major evolutionary importance and develops in response to various biotic and abiotic elements in different environments. It is considered that on the evolutionary scale, species orientation towards clonal reproduction will occur as a response to various pressures endangering the success of sexual reproduction [100]. In the case of *C. litardierei*, among many others, the habitats occupied by the studied population groups differ in a way that has significant ecological importance. Although coping with many challenges, populations growing on dolomite slopes will never experience floods of any intensity. On the other hand, populations from karst poljes or seashore meadows are flooded regularly for prolonged periods [34, 37], which puts their sexual reproduction at risk and makes it irregular. Consequently, the genetics underlying a bias toward clonal reproduction observed in populations from karst poljes and areas near the sea can be linked with the evolutionary shift favoring clonal reproduction. At the same time, the results obtained for the inflorescence height and the total number of flowers were even more population-specific, as their values overlapped substantially among all three studied groups of populations (Fig. 3). Such results suggested that genetic mechanisms underlying these pollination success-related variations have developed independently of perceived ecological pressures in different habitats. They are likely unrelated to abiotic variables we considered important when classifying these habitats (e.g., water and nutrients availability or drought and temperature stress). Instead, they are

possibly linked to biotic elements outside this research's scope, such as local pollinator assemblage or predominant vegetation type [101–103].

Genetic loci and functional pathways associated with reproductive traits in *C. litardierei*

Using univariate and multivariate GWAS approaches, we identified several loci associated with reproduction-related traits in *C. litardierei*. Functional annotation of the genomic regions surrounding these SNP loci revealed their association with genomic regions responsible for coding key protein families involved in crucial biological pathways related to reproduction. We observed the pivotal role of nitrogen metabolism mediated by arginase (ARG), which in *O. sativa* is encoded by OsARG, influencing plant height, growth, and development through its impact on amino acid metabolism and photosynthesis [104, 105]. The silencing of arginase genes ARGAH1 and ARGAH2 in *Arabidopsis* increased nitric oxide (NO) synthase activity, reducing nitrogen levels [106]. Given that nitrogen is essential for almost all plant metabolic processes, including the formation of macromolecules necessary for growth [107, 108], its deficiency can significantly impede plant productivity by inhibiting photosynthesis [109, 110], growth potential [111], CO_2 uptake, and carbohydrate synthesis [112]. We also identified SNP loci in regions encoding enzymes from the cytochrome P450 family involved in the biosynthesis and catabolism of phytohormones and metabolites [113]. For example, in *O. sativa*, CYP701 A8, a member of the cytochrome P450 family, regulates gibberellin (GA) phytohormone biosynthesis [114], while in *Malus domestica*, MdCYP716B1 influences plant height by modulating GA levels [115]. Similarly, mutations in CYP88 disrupt gibberellin biosynthesis, resulting in altered plant stature in barley and maize [116]. Our findings suggest that GAs could stimulate bulb growth by enhancing cell division and regulating processes such as sugar accumulation, which is essential for dormancy release and growth initiation in bulbs [117]. We identified significant SNP loci within genomic regions associated with CCHC-ZFP genes, which are known to play critical roles in plant growth, development, and responses to biotic and abiotic stresses [118]. Sun et al. [119] have further demonstrated that TaCCHC-ZFP genes in *Triticum aestivum* regulate plant growth and stress adaptation. Furthermore, we identified significant SNP loci within the genomic region associated with aspartic proteases (APs), crucial for rapid growth and organ development, as demonstrated in *Phyllostachys edulis* (Moso bamboo) and its associated PhAPs [120]. SNP loci within the genomic region encoding Complex I were also discovered; deficiencies in Complex I, specifically due to the absence of the NDUFV1 gene, are known to slow down growth and development at all life stages, as observed in *A. thaliana* mutants [121]. Additionally, we detected mutations in genetic regions encoding receptor-like kinases (RLKs), which are crucial for perceiving brassinosteroids (BR) and regulating essential growth processes, as exemplified by the BRI1 receptor in *A. thaliana* [122–124]. Regions encoding protein tyrosine kinases (PTKs) were also recorded in the functional annotation of *C. litardierei* sequences. In *O. sativa*, OsPTK2, OsPTK8, OsPTK13, OsPTK14, and OsPTK18 were identified as stress-responsive PTKs involved in abiotic stress tolerance, including cold, heat, and submergence [125]. In the context of flower development, we identified variations within the regions encoding the C2 domain, including the proteins QUIRKY and STRUBBELIG, which are crucial for intercellular communication and tissue morphogenesis in *A. thaliana*, processes vital for reproductive structure development [126]. We identified significant SNP loci within genomic regions associated with sterol synthase, and in *A. thaliana*, mutations in the FACKEL (FK) gene, which encodes a sterol C-14 reductase involved in sterol biosynthesis, disrupt cell division and tissue patterning, leading to stunted growth and abnormal development of key structures like cotyledons, hypocotyl, and meristem [127]. Also, we detected mutations in genetic regions encoding sugar transporters. Huang et al. [128] found that the expression of key sugar transporter genes, such as LohSTP8, LohSTP12, and LflERD6.3, was upregulated during critical stages of bulb formation, including bulblet initiation, suggesting these genes play vital roles in sucrose metabolism and starch accumulation during bulb development in lilies (*Lilium spp.*).

Study limitations and considerations for future research

Despite giving valuable insight into molecular mechanisms underlying local adaptation in studied *C. litardierei* populations, this research also has limitations worth mentioning, and perhaps the most important one is the selection of the ddRADseq approach for the DNA library preparation. As one of the most popular reduced representation sequencing approaches, known for its high robustness, flexibility, and cost-efficiency, ddRADseq has often been used in similar research [129–131]. However, ddRADseq, like other members of a RADseq family, has a significant limitation regarding the genome scan resolution [132]. When implementing any of the RADseqs for the DNA library preparation, a substantial portion of the genomic information remains unexplored, particularly as the size of the studied genome increases. Consequently, given the relatively large genome size of *C. litardierei* (3.7 Gb) [40], the complexity of our results is also influenced by this factor.

Conclusion

This study provides valuable insights into the genetic basis of local adaptation and reproduction-related traits in selected *Chouardia litardierei* populations across contrasting environments. Population-genetic analyses revealed partial genetic structuring, with populations from dry, drought-prone habitats forming a distinct, well-differentiated group. In contrast, those from inland meadows and seashore habitats showed no clear genetic structuring, likely due to recent common ancestry or contemporary gene flow. Precipitation in the coldest quarter was recognized as a key driver of adaptive genetic variation. Furthermore, the GEA study identified numerous genes linked mostly to various abiotic stress responses and key physiological processes, improving our understanding of molecular mechanisms enabling local adaptation of natural populations coping with contrasting environmental conditions. By implementing comprehensive GWAS approaches, we identified numerous loci significant for reproduction-related traits' development in studied populations. Functional annotation of the associated genomic regions revealed key protein families involved in vital biological pathways related to reproduction, including nitrogen metabolism, phytohormone regulation, and floral organ development. High narrow-sense heritability estimates indicated that genetic factors accounted for over 55% of the phenotypic variance in each trait. Among these, the average height of inflorescences (AHI) showed the highest heritability of 71.95%, underscoring its significant role in reproductive success. These findings enhance our understanding of the genetic mechanisms driving local adaptation in *C. litardierei* and establish a foundation for future plant adaptation and speciation studies. This research emphasizes the complexity of the genetic architecture driving phenotypic diversity in plants. It highlights the importance of genomic approaches in investigating adaptive traits in non-model species facing various ecological pressures.

Abbreviations

AHI	Average Height of Inflorescences
AP	Aspartic Protease
APG III	Angiosperm Phylogeny Group III
ARG	Arginase Family
BAM	Binary Alignment Map
BC	Bulb Count
BSLMM	Bayesian Sparse Linear Mixed Model
Chr	Chromosome
CYP450	Cytochrome P450
ddRADseq	Double Digest Restriction Site-Associated DNA Sequencing
EGGNog	Evolutionary Genealogy of Genes: Non-supervised Orthologous Groups
GEA	Genome-Environment Association
GEMMA	Genome-wide Efficient Mixed Model Association
GLMM	Generalized linear mixed model
GMMAT	Generalized Mixed Model Association Tests
GWAS	Genome-Wide Association Study
IRQ	Interquartile range
LFMM	Latent Factor Mixed Models
LMM	Linear Mixed Model
MAF	Minor Allele Frequency
mGWAS	Multivariate genome-wide association study
PCA	Principal Component Analysis
PGE	Proportion of Genetic Effect
PFAM	Protein Family
PIP	Posterior Inclusion Probability
PTK	Protein Tyrosine Kinase
PVE	Proportion of Variance Explained
RDA	Linear Model Redundancy Analysis
RKL	Receptor-Like protein Kinase
sNMF	Sparse Non-negative Matrix Factorization
SNP	Single Nucleotide Polymorphism
TFC	Total Flower Count

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06617-4.

Additional file 1. Locations of sampled *Chouardia litardierei* populations with their associated habitat types. Locations of sampled *C. litardierei* populations and their associated habitat types. This file provides the geographic coordinates (latitude and longitude) for each population site, along with the habitat type for each location. Populations were sampled from three distinct habitat types: meadow-karst poljes, seashore-grassland, and dolomite-bedrock. The table includes populations from both Croatia and Montenegro.

Additional file 2. PCA result based on 19 bioclimatic variables for the localities of the studied populations and the common garden experiment site. The principal component analysis (PCA) was conducted based on 19 bioclimatic variables to assess the environmental conditions of the localities of the studied populations and the common garden experiment site. These variables represent a comprehensive suite of climatic factors, including temperature and precipitation metrics, which are critical for understanding ecological and environmental diversity. The temperature-related variables analyzed include Annual Mean Temperature (BIO1), Mean Diurnal Range (BIO2), Isothermality (BIO3), Temperature Seasonality (BIO4), Maximum Temperature of the Warmest Month (BIO5), Minimum Temperature of the Coldest Month (BIO6), Temperature Annual Range (BIO7), and the mean temperatures for specific seasonal periods: Wettest Quarter (BIO8), Driest Quarter (BIO9), Warmest Quarter (BIO10), and Coldest Quarter (BIO11). The precipitation-related variables encompass Annual Precipitation (BIO12), Precipitation of the Wettest Month (BIO13), Precipitation of the Driest Month (BIO14), Precipitation Seasonality (BIO15), and the precipitation levels during the Wettest Quarter (BIO16), Driest Quarter (BIO17), Warmest Quarter (BIO18), and Coldest Quarter (BIO19).

Additional file 3. Cross-entropy vs. number of ancestral populations in sNMF analysis on *Chouardia litardirei*. Description of Data: This figure shows the relationship between the number of ancestral populations (K) and the cross-entropy values from an sNMF analysis. Cross-entropy decreases as the number of ancestral populations increases, eventually stabilizing, indicating an optimal K where the model best explains the genetic structure.

Additional file 4. Phylogenetic relationships among *Chouardia litardierei* populations based on Nei's genetic distances. The unrooted phylogenetic tree illustrates the relationships among *Chouardia litardierei* populations based on Nei's genetic distances, highlighting genetic divergence across habitats. The tree was constructed by calculating Nei's genetic distances using the "adegenet" package in R, followed by bootstrapping (1,000 replicates) with the "poppr" package. The final tree was visualized in MEGA7 after conversion to Newick format using the "ape" package.

Additional file 5. EggNOG output file for the 83 most significant SNP loci associated with four distinct bioclimatic variables, identified as being most relevant to the traits under investigation. The EggNOG output file provides functional annotations for the 83 most significant SNP loci associated with four key bioclimatic variables, identified as crucial to the traits under

investigation. These variables include bio4 (Temperature Seasonality), bio8 (Mean Temperature of Wettest Quarter), bio9 (Mean Temperature of Driest Quarter), and bio19 (Precipitation of Coldest Quarter). Using the reference genome, sequences were generated for each significant SNP, covering a 50-kilobase region, with 25 kilobases upstream and downstream of each SNP.

Additional file 6. SNPs passing the genome-wide significance threshold in the single-SNP linear mixed model (LMM) analysis in both GEMMA and GMMAT analyses for each reproduction-related morphological trait of *Chouardia litardierei*: TFC, AHI and BC. This file details the results of the single-SNP linear mixed model (LMM) analyses performed using both GEMMA and GMMAT for *C. litardierei*. It specifically presents the single nucleotide polymorphisms (SNPs) that passed the genome-wide significance threshold (p-values $< 1 \times 10^{-3}$) for each reproduction-related morphological trait, including Average Height of Inflorescences (AHI), Total Flower Count (TFC), and Bulb Count (BC) per genotype. The LMM was fitted on a dataset comprising 23,315 SNPs and 214 individuals.

Additional file 7. SNPs identified as having a major sparse effect (PIP > 0.1) on the AHI, TFC and BC traits of *Chouardia litardierei* in the multi-SNP Bayesian sparse linear mixed model (BSLMM) analysis. This file presents the results of the Bayesian sparse linear mixed model (BSLMM) analysis conducted on 23,315 single nucleotide polymorphisms (SNPs) across 214 individuals of *C. litardierei*. It highlights SNPs identified as having a major sparse effect (posterior inclusion probability, PIP > 0.1) on three key reproductive traits: Average Height of Inflorescences (AHI), Total Flower Count (TFC), and Bulb Count (BC) per genotype.

Additional file 8. Means, medians, and 95% equal tail posterior probability intervals (95% ETPPIs) of hyperparameters estimated from the Bayesian sparse linear mixed model (BSLMM) in reproduction-related morphological traits AHI, TFC and BC of *Chouardia litardierei*. This file presents the means, medians, and 95% equal tail posterior probability intervals (95% ETPPIs) of hyperparameters estimated from the Bayesian sparse linear mixed model (BSLMM) analysis focused on reproduction-related morphological traits in *C. litardierei*, specifically Average Height of Inflorescences (AHI), Total Flower Count (TFC), and Bulb Count (BC). The BSLMM was fitted using a dataset of 23,315 single nucleotide polymorphisms (SNPs) across 214 individuals. The file includes detailed estimates of hyperparameters such as h (the proportion of phenotypic variance explained by variants), n.gamma (the number of variants with major effect), pi (the proportion of variants with non-zero effects), PGE (the proportion of genetic variance explained by variants with major effect), PVE (the proportion of phenotypic variance explained by variants), and rho (the proportion of genetic variance explained by variants with major effect).

Additional file 9. SNPs passing the genome-wide significance threshold ($p < 1 \times 10^{-3}$) in the multivariate linear mixed model (mvLMM) analysis for AHI and TFC traits of *Chouardia litardierei* in GEMMA multivariate GWAS. This file lists the single nucleotide polymorphisms (SNPs) that passed the genome-wide significance threshold ($p < 1 \times 10^{-3}$) in the multivariate linear mixed model (mvLMM) analysis for the Average Height of Inflorescences (AHI) and Total Flower Count (TFC) traits of *C. litardierei*, using multivariate genome-wide association studies (GWAS) conducted with GEMMA. The mvLMM in GEMMA was fitted on a dataset comprising 23,315 SNPs from 214 individuals.

Additional file 10. SNPs passing the genome-wide significance threshold ($p < 1 \times 10^{-3}$) in the multivariate linear mixed model (mvLMM) analysis for AHI and BC traits of *Chouardia litardierei* in GEMMA multivariate GWAS. This file lists the single nucleotide polymorphisms (SNPs) that passed the genome-wide significance threshold ($p < 1 \times 10^{-3}$) in the multivariate linear mixed model (mvLMM) analysis for the Average Height of Inflorescences (AHI) and Bulb Count (BC) traits of *C. litardierei*, using multivariate genome-wide association studies (GWAS) conducted with GEMMA. The mvLMM in GEMMA was fitted on a dataset comprising 23,315 SNPs from 214 individuals.

Additional file 11. Frequency of effect alleles across populations for significant SNPs identified in the single-SNP LMM analysis (GEMMA and GMMAT), as well as the multi-SNP BSLMM analysis, all of which surpassed the genome-wide significance threshold (1×10^{-3}). The analysis also includes SNPs meeting the same threshold in the multivariate GWAS. The corresponding SNPs are detailed in Table 2 and Table 4 of the manuscript. Overlapping points of different colors represent SNPs associated with different traits, with each color corresponding to a specific trait. The overlap occurs because some SNPs are shared across traits, leading to their placement one in front of the other. AHI, Average Height of Inflorescences; BC, Bulb Count; TFC, Total Flower Count. The data represents the frequency of effect alleles for significant SNPs identified through different GWAS approaches, including single-SNP LMM analyses (GEMMA and GMMAT), multi-SNP BSLMM analysis, and multivariate GWAS. All SNPs included surpassed the genome-wide significance threshold of 1×10^{-3}. The dataset also captures overlaps of SNPs associated with different traits—AHI (Average Height of Inflorescences), BC (Bulb Count), and TFC (Total Flower Count)—with overlapping points indicating shared SNPs across traits. Frequencies are stratified by population, allowing for comparative analysis of allele distributions.

Additional file 12. EggNOG output file for 12 SNPs that exceeded the genome-wide significance threshold (1×10^{-3}) in both the single-SNP LMM and multi-SNP BSLMM analyses of the *Chouardia litardierei* traits: TFC, AHI, and BC. This file lists the results of eggNOG-mapper v2 analysis for regions associated with 12 SNPs that exceeded the genome-wide significance threshold (1×10^{-3}) in both the single-SNP LMM and multi-SNP BSLMM analyses of TFC, AHI, and BC traits in *C. litardierei*.

Additional file 13. EggNOG output file for 13 SNP loci that exceeded the genome-wide significance threshold (1×10^{-3}) in the multivariate GWAS analysis of the *Chouardia litardierei* traits: AHI and TFC. This file lists the results of eggNOG-mapper v2 analysis for regions associated with 13 SNPs that exceeded the genome-wide significance threshold (1×10^{-3}) in the multivariate GWAS BSLMM analysis of the *C. litardierei* traits: AHI and TFC.

Additional file 14: Table 1. EggNOG output file for 2 SNP loci that exceeded the genome-wide significance threshold (1×10^{-3}) in the multivariate GWAS analysis of the *Chouardia litardierei* traits: AHI and BC. This file lists the results of eggNOG-mapper v2 analysis for regions associated with 13 SNPs that exceeded the genome-wide significance threshold (1×10^{-3}) in the multivariate GWAS BSLMM analysis of the *C. litardierei* traits: AHI and BC.

Acknowledgements
Not applicable.

Clinical trial number
Not applicable.

Authors' contributions
Conceptualization, I.R., B.S., S.L.Š., and N.P.; methodology, S.L.Š, N.P., D.M., and I.R., and K.K.; funding acquisition, I.R.; supervision, I.R.; visualization, S.L.Š., N.P., and D.M.; writing: S.L.Š., N.P., and I.R.; review and editing, S.L.Š, N.P., I.R. and B.S. All authors read and approved the final manuscript.

Funding
This research was funded by the Croatian Science Foundation. All the research activities were financed by the "Amethyst Meadow Squill (*Chouardia litardierei*, Hyacinthaceae): a study system for ecological divergence" (HRZZ-IP-2020–02-8099) project.

Data availability
The FASTQ files used for this analysis are available in the NCBI database under BioProject accession number PRJNA1164649.

Declarations

Ethics approval and consent to participate
Experimental research and field studies complied with national and international guidelines and legislation. Permission to collect samples for the molecular analysis and the voucher specimens of *Chouardia litardierei* in

Croatia was granted by the Ministry of Economy and Sustainable Development (permission UP/I-612–07/21–48/27; March 8, 2021), and in Montenegro by the Environmental Protection Agency (U2-UPI-361/1; March 15, 2021).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Biology, Faculty of Science, University of Zagreb, Marulićev Trg 9 A, Zagreb 10000, Croatia. [2]Department of Biology and Human Genetics, School of Medicine, University of Split, Šoltanska 2, Split 21000, Croatia. [3]Department of Electronic Systems and Information Processing, Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, Zagreb 10000, Croatia. [4]Natural History Museum Rijeka, Lorenzov Prolaz 1, Rijeka 51000, Croatia. [5]Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Glagoljaška 8, Koper 6000, Slovenia.

Received: 11 October 2024 Accepted: 24 April 2025
Published online: 02 May 2025

References

1. Paschoa RP da, Pinto VB, Pereira JP, Cavatte PC, Garbin ML, Godinho T de O, et al. Proteomic and Physiological Signatures of Altitude Adaptation in a Myrsine coriacea Population under Common Garden Conditions. BioRxiv 2023:2023.12.01.569649. https://doi.org/10.1101/2023.12.01.569649.
2. Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, et al. Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. Am Nat. 2016;188:379–97. https://doi.org/10.1086/688018.
3. Hu ZM, Le ZK, Weinberger F, Duan DL, Draisma SGA, Serrão EA. Linking Ecology to Genetics to Better Understand Adaptation and Evolution: A Review in Marine Macrophytes. Front Mar Sci. 2020;7: 545102. https://doi.org/10.3389/FMARS.2020.545102/BIBTEX.
4. Walter GM, Clark J, Cristaudo A, Terranova D, Nevado B, Catara S, et al. Adaptive divergence generates distinct plastic responses in two closely related Senecio species. Evolution. 2022;76:1229. https://doi.org/10.1111/EVO.14478.
5. Rajpal VR, Rathore P, Mehta S, Wadhwa N, Yadav P, Berry E, et al. Epigenetic variation: A major player in facilitating plant fitness under changing environmental conditions. Front Cell Dev Biol. 2022;10:1020958. https://doi.org/10.3389/FCELL.2022.1020958/BIBTEX.
6. Crispo E. Modifying effects of phenotypic plasticity on interactions among natural selection, adaptation and gene flow. J Evol Biol. 2008;21:1460–9. https://doi.org/10.1111/J.1420-9101.2008.01592.X.
7. Klimeš L, Klimešová J, Hendriks R, Van Groenendael J. Clonal plant architecture: a comparative analysis of form and function. Ecol Evol Clonal Plants. 1997;1:29.
8. Silvertown J. The evolutionary maintenance of sexual reproduction: Evidence from the ecological distribution of asexual reproduction in clonal plants. Int J Plant Sci. 2008;169:157–68. https://doi.org/10.1086/523357/0.
9. Barrett SCH. Influences of clonality on plant sexual reproduction. Proc Natl Acad Sci U S A. 2015;112:8859–66. https://doi.org/10.1073/PNAS.1501712112.
10. Stuefer JF, Erschbamer B, Huber H, Suzuki J-I. The ecology and evolutionary biology of clonal plants: an introduction to the proceedings of Clone-2000. Ecology and Evolutionary Biology of Clonal Plants 2002:1–8. https://doi.org/10.1007/978-94-017-1345-0_1.
11. Prati D, Schmid B. Genetic differentiation of life-history traits within populations of the clonal plant Ranunculus reptans. Oikos. 2000;90:442–56. https://doi.org/10.1034/J.1600-0706.2000.900303.X.
12. Eriksson O. Clonal life histories and the evolution of seed recruitment. In: de Kroon H, van Groenendael J, editors. The Ecology and Evolution of Clonal Plants. Leiden: Backhuys Publishers; 1997. p. 211–26.
13. Vekemans X, Hardy OJ. New insights from fine-scale spatial genetic structure analyses in plant populations. Mol Ecol. 2004;13:921–35. https://doi.org/10.1046/J.1365-294X.2004.02076.X.
14. Gong W, Gu L, Zhang D. Low genetic diversity and high genetic divergence caused by inbreeding and geographical isolation in the populations of endangered species Loropetalum subcordatum (Hamamelidaceae) endemic to China. Conserv Genet. 2010;11:2281–8. https://doi.org/10.1007/S10592-010-0113-9/METRICS.
15. Vallejo-Marín M, Dorken ME, Barrett SCH. The ecological and evolutionary consequences of clonality for plant mating. Annu Rev Ecol Evol Syst. 2010;41:193–213. https://doi.org/10.1146/ANNUREV.ECOLSYS.110308.120258/CITE/REFWORKS.
16. Kleijn D, Treier UA, Müller-Schärer H. The importance of nitrogen and carbohydrate storage for plant growth of the alpine herb Veratrum album. New Phytol. 2005;166:565–75. https://doi.org/10.1111/J.1469-8137.2005.01321.X.
17. Ma R, Xu S, Chen Y, Guo F, Wu R. Allometric relationships between leaf and bulb traits of Fritillaria przewalskii Maxim. grown at different altitudes. PLoS One 2020;15. https://doi.org/10.1371/JOURNAL.PONE.0239427.
18. Atif MJ, Ahanger MA, Amin B, Ghani MI, Ali M, Cheng Z. Mechanism of Allium Crops Bulb Enlargement in Response to Photoperiod: A Review. Int J Mol Sci 2020;21. https://doi.org/10.3390/IJMS21041325.
19. Ohashi K, Yahara T. Behavioural responses of pollinators to variation in floral display size and their influences on the evolution of floral traits. Cognitive Ecology of Pollination 2009:274–96. https://doi.org/10.1017/CBO9780511542268.015.
20. Suetsugu K, Naito RS, Fukushima S, Kawakita A, Kato M. Pollination system and the effect of inflorescence size on fruit set in the deceptive orchid Cephalanthera falcata. J Plant Res. 2015;128:585–94. https://doi.org/10.1007/S10265-015-0716-9.
21. Trunschke J, Sletvold N, Ågren J. The independent and combined effects of floral traits distinguishing two pollination ecotypes of a moth-pollinated orchid. Ecol Evol. 2019;9:1191. https://doi.org/10.1002/ECE3.4808.
22. De Villemereuil P, Gaggiotti OE, Mouterde M, Till-Bottraud I. Common garden experiments in the genomic era: new perspectives and opportunities. Heredity 2016 116:3 2015;116:249–54. https://doi.org/10.1038/hdy.2015.93.
23. McKay JK, Bishop JG, Lin JZ, Richards JH, Sala A, Mitchell-Olds T. Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rockcress. Proc Biol Sci. 2001;268:1715–21. https://doi.org/10.1098/RSPB.2001.1715.
24. Kawakami T, Morgan TJ, Nippert JB, Ocheltree TW, Keith R, Dhakal P, et al. Natural selection drives clinal life history patterns in the perennial sunflower species. Helianthus maximiliani Mol Ecol. 2011;20:2318–28. https://doi.org/10.1111/J.1365-294X.2011.05105.X.
25. Brachi B, Villoutreix R, Faure N, Hautekèete N, Piquot Y, Pauwels M, et al. Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in Arabidopsis thaliana. Mol Ecol. 2013;22:4222–40. https://doi.org/10.1111/MEC.12396.
26. Toräng P, Wunder J, Obeso JR, Herzog M, Coupland G, Ågren J. Large-scale adaptive differentiation in the alpine perennial herb Arabis alpina. New Phytol. 2015;206:459–70. https://doi.org/10.1111/NPH.13176.
27. Slavov GT, Nipper R, Robson P, Farrar K, Allison GG, Bosch M, et al. Genome-wide association studies and prediction of 17 traits related to phenology, biomass and cell wall composition in the energy grass Miscanthus sinensis. New Phytol. 2014;201:1227–39. https://doi.org/10.1111/NPH.12621.
28. Liu Y, Erbilgin N, Cappa EP, Chen C, Ratcliffe B, Wei X, et al. Decoupling of height growth and drought or pest resistance tradeoffs is revealed through multiple common-garden experiments of lodgepole pine. Evolution (N Y). 2023;77:893–906. https://doi.org/10.1093/EVOLUT/QPAD004.
29. Compton S, Stackpole C, Dixit A, Sekhwal MK, Kolb T, De la Torre AR. Differences in heat tolerance, water use efficiency and growth among Douglas-fir families and varieties evidenced by GWAS and common garden studies. AoB Plants. 2023;15:1–15. https://doi.org/10.1093/AOBPLA/PLAD008.
30. Gaži-Baskova V. Geografsko raširenje livadnog procjepka ili lučike (Scilla pratensis W. et K.). Biološki Glas. 1962;15:49–54.

31. Petkovšek V, Seliškar A. Travniška morska čebulica, novost v slovenski flori. Proteus. 1978;40:160–1.
32. Mihevc A, Prelovšek M, Zupan Hajna N. Introduction to the Dinaric Karst. Postojna: Karst Research Institute at ZRC SAZU; 2010. https://doi.org/10.3986/9789612541989.
33. Deák B, Kovács B, Rádai Z, Apostolova I, Kelemen A, Kiss R, et al. Linking environmental heterogeneity and plant diversity: The ecological role of small natural features in homogeneous landscapes. Sci Total Environ. 2021;763: 144199. https://doi.org/10.1016/J.SCITOTENV.2020.144199.
34. Bonacci O. Ecohydrology of karst poljes and their vulnerability. Dinaric karst poljes - floods for life, Radolfzell Am Bodensee: EuroNatur; 2014, p. 25–37.
35. Mihevc A. Geomorphology. In: Zupan Hajna N, Mihevc A, Prelovšek M, editors. Introduction to the Dinaric Karst. Postojna: Inštitut za raziskovanje krasa ZRC SAZU; 2010. p. 30–43.
36. Prohić E, Peh Z, Miko S. Geochemical characterization of a karst polje: an example from Sinjsko Polje, Croatia. Environ Geology. 1998;33:263–73. https://doi.org/10.1007/s002540050245.
37. Horvatić S. Flora i vegetacija otoka Paga. Prirodoslovna Istraživanja Kraljevine Jugoslavije. 1934;19:116–372.
38. Mota J, Merlo E, Martínez-Hernández F, Mendoza-Fernández AJ, Pérez-García FJ, Salmerón-Sánchez E. Plants on Rich-Magnesium Dolomite Barrens: A Global Phenomenon. Biology. 2021;10:38. https://doi.org/10.3390/BIOLOGY10010038.
39. Šilić Č. Morphology, chorology, ecology and phenology of two population groups of Scilla litardierei Breistr. (syn.: S. pratensis Waldst. & Kit. non Bergeret). Bilten Društva Ekologa Bosne i Hercegovine, Serija B. 1990;5:107–16.
40. Radosavljević I, Križanović K, Šarančić SL, Jakše J. Towards the investigation of the adaptive divergence in a species of exceptional ecological plasticity: chromosome-scale genome assembly of Chouardia litardierei (Hyacinthaceae). Int J Mol Sci. 2023;24:10755. https://doi.org/10.3390/ijms241310755.
41. Arab MM, Brown PJ, Abdollahi-Arpanahi R, Sohrabi SS, Askari H, Aliniaeifard S, et al. Genome-wide association analysis and pathway enrichment provide insights into the genetic basis of photosynthetic responses to drought stress in Persian walnut. Hortic Res 2022;9. https://doi.org/10.1093/HR/UHAC124.
42. Ritter-Studnička H. Flora i vegetacija livada kraških polja Bosne i Hercegovine. Godišnjak Biološkog Instituta u Sarajevu. 1954;7:25–110.
43. Bernatchez L, Ferchaud AL, Berger CS, Venney CJ, Xuereb A. Genomics for monitoring and understanding species responses to global climate change. Nature Reviews Genetics 2023 25:3 2023;25:165–83. https://doi.org/10.1038/s41576-023-00657-y.
44. Montejo-Kovacevich G, Salazar PA, Smith SH, Gavilanes K, Bacquet CN, Chan YF, et al. Genomics of altitude-associated wing shape in two tropical butterflies. Mol Ecol. 2021;30:6387–402. https://doi.org/10.1111/MEC.16067.
45. Siljak-Yakovlev S. Cytogenetic characterization of two Chouardia species: Ch. lakusicii (Šilić) Speta and Ch. litardierei (Breistr.) Speta. In: Ekim T, editor. Proceedings of the XIII Optima Meeting. Antalya: Flora Research Society. 2010, p. 169.
46. Wagner DH. A Rule of Thumb for Botanists: The 1 in 20 Rule. Oregon Flora-On-Line Newsletter: Oregon State University; 1995. p. 1.
47. Köppen W. Klassifikation der Klimate nach Temperatur. Niederschlag und Jahreslauf Petermanns Geogr Mitt. 1918;64:193–203.
48. Zaninović K; G-ČM; PTM; VM; MJ; BA; V V. Climate Atlas of Croatia: 1961–1990: 1971–2000 / Klimatski Atlas Hrvatske, Climate Atlas of Croatia: 1961–1990: 1971–2000. Zagreb: Državni hidrometeorološki zavod; 2008.
49. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 2005;25:1965–78. https://doi.org/10.1002/JOC.1276.
50. Tang Y, Horikoshi M, Li W. Ggfortify: Unified interface to visualize statistical results of popular r packages. R Journal 2016;8:478–89. https://doi.org/10.32614/RJ-2016-060.
51. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2016.
52. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. PLoS ONE. 2012;7: e37135. https://doi.org/10.1371/JOURNAL.PONE.0037135.
53. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 2011;17:10–2. https://doi.org/10.14806/EJ.17.1.200.
54. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–60. https://doi.org/10.1093/BIOINFORMATICS/BTP324.
55. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set for population genomics. Mol Ecol. 2013;22:3124–40. https://doi.org/10.1111/MEC.12354.
56. Paris JR, Stevens JR, Catchen JM. Lost in parameter space: a road map for stacks. Methods Ecol Evol. 2017;8:1360–73. https://doi.org/10.1111/2041-210X.12775.
57. Mijangos JL, Gruber B, Berry O, Pacioni C, Georges A. dartR v2: An accessible genetic analysis platform for conservation, ecology and agriculture. Methods Ecol Evol. 2022;13:2150–8. https://doi.org/10.1111/2041-210X.13918.
58. Lê S, Josse J, Husson F. FactoMineR: An R Package for Multivariate Analysis. J Stat Softw 2008;25:1–18. https://doi.org/10.18637/JSS.V025.I01.
59. Kassambara A, Mundt F. factoextra: extract and visualize the results of multivariate data analyses. R package version 1.0.7. 2017.
60. Frichot E, François O. LEA: An R package for landscape and ecological association studies. Methods Ecol Evol. 2015;6:925–9. https://doi.org/10.1111/2041-210X.12382.
61. Knaus BJ, Grünwald NJ. vcfr: a package to manipulate and visualize variant call format data in R. Mol Ecol Resour. 2017;17:44–53. https://doi.org/10.1111/1755-0998.12549.
62. Jombart T, Ahmed I. adegenet 1.3–1: new tools for the analysis of genome-wide SNP data. Bioinformatics 2011;27:3070–1. https://doi.org/10.1093/BIOINFORMATICS/BTR521.
63. Kamvar ZN, Tabima JF, Gr̈unwald NJ. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2014;2:1–14. https://doi.org/10.7717/PEERJ.281.
64. Paradis E, Claude J, Strimmer K. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics. 2004;20:289–90. https://doi.org/10.1093/BIOINFORMATICS/BTG412.
65. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 2016;33:1870–4. https://doi.org/10.1093/MOLBEV/MSW054.
66. Legendre P, Legendre L. Numerical ecology. 3rd English edition Elsevier: Amsterdam. 2012; 1006 pp.
67. van den Wollenberg AL. Redundancy analysis an alternative for canonical correlation analysis. Psychometrika. 1977;42:207–19. https://doi.org/10.1007/BF02294050/METRICS.
68. Forester BR, Jones MR, Joost S, Landguth EL, Lasky JR. Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. Mol Ecol. 2016;25:104–20. https://doi.org/10.1111/MEC.13476.
69. Forester BR, Lasky JR, Wagner HH, Urban DL. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. Mol Ecol. 2018;27:2215–33. https://doi.org/10.1111/MEC.14584.
70. Naimi B, Hamm NAS, Groen TA, Skidmore AK, Toxopeus AG. Where is positional uncertainty a problem for species distribution modelling? Ecography. 2014;37:191–203. https://doi.org/10.1111/J.1600-0587.2013.00205.X.
71. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB et al.. vegan: Community Ecology Package. 2012.
72. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res 2019;47:D309–14. https://doi.org/10.1093/NAR/GKY1085.
73. Cantalapiedra CP, Herṇandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. Mol Biol Evol 2021;38:5825–9. https://doi.org/10.1093/MOLBEV/MSAB293.
74. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10:1–4. https://doi.org/10.1093/GIGASCIENCE/GIAB008.

75. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nature Genetics 2012 44:7 2012;44:821–4. https://doi.org/10.1038/ng.2310.
76. Chen H, Huffman JE, Brody JA, Wang C, Lee S, Li Z, et al. Efficient Variant Set Mixed Model Association Tests for Continuous and Binary Traits in Large-Scale Whole-Genome Sequencing Studies. Am J Hum Genet. 2019;104:260–74. https://doi.org/10.1016/j.ajhg.2018.12.012.
77. Zhou X, Carbonetto P, Stephens M. Polygenic Modeling with Bayesian Sparse Linear Mixed Models. PLoS Genet. 2013;9: e1003264. https://doi.org/10.1371/JOURNAL.PGEN.1003264.
78. Turner SD. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. J Open Source Softw 2018;3:731. https://doi.org/10.21105/JOSS.00731.
79. Yin L, Zhang H, Tang Z, Xu J, Yin D, Zhang Z, et al. rMVP: A Memory-efficient, Visualization-enhanced, and Parallel-accelerated Tool for Genome-wide Association Study. Genomics Proteomics Bioinformatics. 2021;19:619–28. https://doi.org/10.1016/J.GPB.2020.10.007.
80. Ellstrand NC. Is gene flow the most important evolutionary force in plants? Am J Bot. 2014;101:737–53. https://doi.org/10.3732/AJB.1400024.
81. Slatkin M. Gene flow and the geographic structure of natural populations. Science. 1987;236:787–92. https://doi.org/10.1126/SCIENCE.3576198.
82. Capblancq T, Forester BR. Redundancy analysis: A Swiss Army Knife for landscape genomics. Methods Ecol Evol. 2021;12:2298–309. https://doi.org/10.1111/2041-210X.13722.
83. Sun Y, Zhao JY, Li YT, Zhang PG, Wang SP, Guo J, et al. Genome-Wide Analysis of the C2 Domain Family in Soybean and Identification of a Putative Abiotic Stress Response Gene GmC2-148. Front Plant Sci. 2021;12: 620544. https://doi.org/10.3389/FPLS.2021.620544/BIBTEX.
84. Yang Y, Guo Y. Elucidating the molecular mechanisms mediating plant salt-stress responses. New Phytol. 2018;217:523–39. https://doi.org/10.1111/NPH.14920.
85. Prabu G, Prasad DT. Functional characterization of sugarcane MYB transcription factor gene promoter (PScMYBAS1) in response to abiotic stresses and hormones. Plant Cell Rep. 2012;31:661–9. https://doi.org/10.1007/S00299-011-1183-Y/METRICS.
86. Dittmer S, Kleine T, Schwenkert S. The TPR- and J-domain-containing proteins DJC31 and DJC62 are involved in abiotic stress responses in Arabidopsis thaliana. J Cell Sci 2021;134. https://doi.org/10.1242/JCS.259032.
87. Satheesh V, Chidambaranathan P, Jagannadham PT, Kumar V, Jain PK, Chinnusamy V, et al. Transmembrane START domain proteins: in silico identification, characterization and expression analysis under stress conditions in chickpea (Cicer arietinum L.). Plant Signal Behav 2016;11. https://doi.org/10.4161/15592324.2014.992698.
88. Salvucci ME, Crafts-Brandner SJ. Inhibition of photosynthesis by heat stress: the activation state of Rubisco as a limiting factor in photosynthesis. Physiol Plant. 2004;120:179–86. https://doi.org/10.1111/J.0031-9317.2004.0173.X.
89. Ru JN, Hou ZH, Zheng L, Zhao Q, Wang FZ, Chen J, et al. Genome-Wide Analysis of DEAD-box RNA Helicase Family in Wheat (Triticum aestivum) and Functional Identification of TaDEAD-box57 in Abiotic Stress Responses. Front Plant Sci. 2021;12: 797276. https://doi.org/10.3389/FPLS.2021.797276/BIBTEX.
90. Ma L, Tao X, Wang W, Jiao J, Pu Y, Yang G, et al. Genome-wide identification of RNA recognition motif (RRM1) in Brassica rapa and functional analysis of RNA-binding protein (BrRBP) under low-temperature stress. BMC Plant Biol. 2023;23:1–17. https://doi.org/10.1186/S12870-023-04639-4/FIGURES/8.
91. Qin T, Zhao P, Sun J, Zhao Y, Zhang Y, Yang Q, et al. Research Progress of PPR Proteins in RNA Editing, Stress Response. Plant Growth and Development Front Genet. 2021;12: 765580. https://doi.org/10.3389/FGENE.2021.765580/BIBTEX.
92. Du P, Wang Q, Yuan D, Chen S, Su Y, Li L, et al. WRKY transcription factors and OBERON histone-binding proteins form complexes to balance plant growth and stress tolerance. EMBO J. 2023;42. https://doi.org/10.15252/EMBJ.2023113639/SUPPL_FILE/EMBJ2023113639-SUP-0006-SDATAEV.ZIP.
93. Martin RC, Kronmiller BA, Dombrowski JE. Transcriptome Analysis of Lolium temulentum Exposed to a Combination of Drought and Heat Stress. Plants. 2021;10:2247. https://doi.org/10.3390/PLANTS10112247.
94. Ghalambor CK, McKay JK, Carroll SP, Reznick DN. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. Funct Ecol. 2007;21:394–407. https://doi.org/10.1111/J.1365-2435.2007.01283.X.
95. Exposito-Alonso M, Exposito-Alonso M, Gómez Rodríguez R, Barragán C, Capovilla G, Chae E, et al. Natural selection on the Arabidopsis thaliana genome in present and future climates. Nature. 2019;573:126–9. https://doi.org/10.1038/S41586-019-1520-9.
96. Lind BM, Friedline CJ, Wegrzyn JL, Maloney PE, Vogler DR, Neale DB, et al. Water availability drives signatures of local adaptation in white-bark pine (Pinus albicaulis Engelm.) across fine spatial scales of the Lake Tahoe Basin, USA. Mol Ecol 2017;26:3168–85. https://doi.org/10.1111/MEC.14106.
97. Bresadola L, Caseys C, Castiglione S, Buerkle CA, Wegmann D, Lexer C. Admixture mapping in interspecific Populus hybrids identifies classes of genomic architectures for phytochemical, morphological and growth traits. New Phytol. 2019;223:2076–89. https://doi.org/10.1111/NPH.15930.
98. Gumbert A, Kunze J. Inflorescence Height Affects Visitation Behavior of Bees–A Case Study of an Aquatic Plant Community in Bolivia1. Biotropica. 1999;31:466–77. https://doi.org/10.1111/J.1744-7429.1999.TB00389.X.
99. Gramlich S, Liu X, Favre A, Buerkle CA, Karrenberg S. A polygenic architecture with habitat-dependent effects underlies ecological differentiation in Silene. New Phytol. 2022;235:1641–52. https://doi.org/10.1111/NPH.18260.
100. Barrett SCH. Influences of clonality on plant sexual reproduction. Proc Natl Acad Sci U S A. 2015;112:8859–66. https://doi.org/10.1073/PNAS.1501712112/ASSET/2AFC16C4-07D5-4BE0-AC49-C94470796DC1/ASSETS/GRAPHIC/PNAS.1501712112FIG04.JPEG.
101. Herrera CM. Flower traits, habitat, and phylogeny as predictors of pollinator service: a plant community perspective. Ecol Monogr. 2020;90: e01402. https://doi.org/10.1002/ECM.1402.
102. Potts SG, Vulliamy B, Dafni A, Ne'eman G, Willmer P. Linking bees and flowers: How do floral communities structure pollinator communities? Ecology. 2003;84(10):2628–42.
103. Herrera CM. Microclimate and individual variation in pollinators: Flowering plants are more than their flowers. Ecology. 1995;76:1516–24. https://doi.org/10.2307/1938153.
104. Ma X, Cheng Z, Qin R, Qiu Y, Heng Y, Yang H, et al. OsARG encodes an arginase that plays critical roles in panicle development and grain production in rice. Plant J. 2013;73:190–200. https://doi.org/10.1111/J.1365-313X.2012.05122.X.
105. Siddappa S, Marathe GK. What we know about plant arginases? Plant Physiol Biochem. 2020;156:600–10. https://doi.org/10.1016/J.PLAPHY.2020.10.002.
106. Flores T, Todd CD, Tovar-Mendez A, Dhanoa PK, Correa-Aragunde N, Hoyos ME, et al. Arginase-Negative Mutants of Arabidopsis Exhibit Increased Nitric Oxide Signaling in Root Development. Plant Physiol. 2008;147:1936–46. https://doi.org/10.1104/PP.108.121459.
107. Chen Q, Wang Y, Zhang Z, Liu X, Li C, Ma F. Arginine Increases Tolerance to Nitrogen Deficiency in Malus hupehensis via Alterations in Photosynthetic Capacity and Amino Acids Metabolism. Front Plant Sci. 2022;12: 772086. https://doi.org/10.3389/FPLS.2021.772086/BIBTEX.
108. Wen B, Xiao W, Mu Q, Li D, Chen X, Wu H, et al. How does nitrate regulate plant senescence? Plant Physiol Biochem. 2020;157:60–9. https://doi.org/10.1016/J.PLAPHY.2020.08.041.
109. Bassi D, Menossi M, Mattiello L. Nitrogen supply influences photosynthesis establishment along the sugarcane leaf. Scientific Reports 2018 8:1 2018;8:1–13. https://doi.org/10.1038/s41598-018-20653-1.
110. Mu X, Chen Y. The physiological response of photosynthesis to nitrogen deficiency. Plant Physiol Biochem. 2021;158:76–82. https://doi.org/10.1016/J.PLAPHY.2020.11.019.
111. Winter TR, Rostás M. Nitrogen deficiency affects bottom-up cascade without disrupting indirect plant defense. J Chem Ecol. 2010;36:642–51. https://doi.org/10.1007/S10886-010-9797-Z/FIGURES/3.
112. Ciompi S, Gentili E, Guidi L, Soldatini GF. The effect of nitrogen deficiency on leaf gas exchange and chlorophyll fluorescence parameters in sunflower. Plant Sci. 1996;118:177–84. https://doi.org/10.1016/0168-9452(96)04442-1.

113. Xu J, Wang XY, Guo WZ. The cytochrome P450 superfamily: Key players in plant development and defense. J Integr Agric. 2015;14:1673–86. https://doi.org/10.1016/S2095-3119(14)60980-1.
114. Wang Q, Hillwig ML, Wu Y, Peters RJ. CYP701A8: A Rice ent-Kaurene Oxidase Paralog Diverted to More Specialized Diterpenoid Metabolism. Plant Physiol. 2012;158:1418–25. https://doi.org/10.1104/PP.111.187518.
115. Shi J, Zhang F, Wang Y, Zhang S, Wang F, Ma Y. The cytochrome P450 gene, MdCYP716B1, is involved in regulating plant growth and anthracnose resistance in apple. Plant Sci. 2023;335: 111832. https://doi.org/10.1016/J.PLANTSCI.2023.111832.
116. Werck-Reichhart D, Bak S, Paquette S. Cytochromes p450. Arabidopsis Book. 2002;1: e0028. https://doi.org/10.1199/TAB.0028.
117. Markovic M, Trifunovic Momcilov M, Uzelac B, Cingel A, Milosevic S, Jevremovic S, et al. Breaking the Dormancy of Snake's Head Fritillary (Fritillaria meleagris L.) In Vitro Bulbs—Part 1: Effect of GA3, GA Inhibitors and Temperature on Fresh Weight. Sprouting and Sugar Content Plants. 2020;9:1449. https://doi.org/10.3390/PLANTS9111449.
118. Li W-THMWJWY-P. Zinc finger protein (ZFP) in plants—a review. Plant Omics. 2013;6:474–80.
119. Sun A, Li Y, He Y, Zou X, Chen F, Ji RZ, et al. Comprehensive Genome-Wide Identification, Characterization, and Expression Analysis of CCHC-Type Zinc Finger Gene Family in Wheat (Triticum aestivum L.). Front Plant Sci 2022;13:892105. https://doi.org/10.3389/FPLS.2022.892105/BIBTEX.
120. Wang X, Yan X, Li S, Jing Y, Gu L, Zou S, et al. Genome-wide identification, evolution and expression analysis of the aspartic protease gene family during rapid growth of moso bamboo (Phyllostachys edulis) shoots. BMC Genomics. 2021;22:1–20. https://doi.org/10.1186/S12864-020-07290-7/FIGURES/9.
121. Kühn K, Obata T, Feher K, Bock R, Fernie AR, Meyer EH. Complete Mitochondrial Complex I Deficiency Induces an Up-Regulation of Respiratory Fluxes That Is Abolished by Traces of Functional Complex I. Plant Physiol. 2015;168:1537–49. https://doi.org/10.1104/PP.15.00589.
122. Clouse SD, Langford M, Mcmorris TC. A Brassinosteroid-Insensitive Mutant in Arabidopsis thaliana Exhibits Multiple Defects in Growth and Development'. Plant Physiol. 1996;11:671–8.
123. Li J, Chory J. A Putative Leucine-Rich Repeat Receptor Kinase Involved in Brassinosteroid Signal Transduction. Cell. 1997;90:929–38. https://doi.org/10.1016/S0092-8674(00)80357-8.
124. She J, Han Z, Kim TW, Wang J, Cheng W, Chang J, et al. Structural insight into brassinosteroid perception by BRI1. Nature. 2011;474:472–6. https://doi.org/10.1038/nature10178.
125. Allimuthu E, Dalal M, Kumar KG, Sellathdurai D, Kumar RR, Sathee L, et al. Characterization of Atypical Protein Tyrosine Kinase (PTK) Genes and Their Role in Abiotic Stress Response in Rice. Plants. 2020;9:664. https://doi.org/10.3390/PLANTS9050664.
126. Vaddepalli P, Herrmann A, Fulton L, Oelschner M, Hillmer S, Stratil TF, et al. The C2-domain protein QUIRKY and the receptor-like kinase STRUBBELIG localize to plasmodesmata and mediate tissue morphogenesis in Arabidopsis thaliana 2014. https://doi.org/10.1242/dev.113878.
127. Schaller H. The role of sterols in plant growth and development. Prog Lipid Res. 2003;42:163–75. https://doi.org/10.1016/S0163-7827(02)00047-4.
128. Huang Z, Gao C, Xu Y, Liu J, Kang J, Ren Z, et al. Identification and Expression Analysis of Putative Sugar Transporter Gene Family during Bulb Formation in Lilies. Int J Mol Sci. 2024;25. https://doi.org/10.3390/IJMS25063483.
129. Okada Y, Monden Y, Nokihara K, Shirasawa K, Isobe S, Tahara M. Genome-Wide Association Studies (GWAS) for Yield and Weevil Resistance in Sweet potato (Ipomoea batatas (L.) Lam). Plant Cell Rep. 2019;38:1383–92. https://doi.org/10.1007/S00299-019-02445-7.
130. Jaiswal V, Gupta S, Gahlaut V, Muthamilarasan M, Bandyopadhyay T, Ramchiary N, et al. Genome-Wide Association Study of Major Agronomic Traits in Foxtail Millet (Setaria italica L) Using ddRAD Sequencing. Sci Rep. 2019;9:1–11. https://doi.org/10.1038/s41598-019-41602-6.
131. Yang Y, Wu L, Wu X, Li B, Huang W, Weng Z, et al. Identification of Candidate Growth-Related SNPs and Genes Using GWAS in Brown-Marbled Grouper (Epinephelus fuscoguttatus). Mar Biotechnol (NY). 2020;22:153–66. https://doi.org/10.1007/S10126-019-09940-8.
132. Lowry DB, Hoban S, Kelley J, Lotterhos KE, Reed LK, Antolin MCF, et al. Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. Mol Ecol Resour. 2017;17:142–52. https://doi.org/10.1111/1755-0998.12635.

Publisher's Note

Table 1. Locations of sampled *Chouardia litardierei* populations with their associated habitat types.

Population/Location	Country	Latitude (N)	Longitude (E)	Habitat Type
Bjelopolje	Croatia	44.693754°	15.773682°	Meadow – karst poljes
Cetina (Paško polje)	Croatia	43.940922°	16.436367°	Meadow – karst poljes
Budoške Bare	Montenegro	42.743747°	18.926361°	Meadow – karst poljes
Pag (Kolansko blato)	Croatia	44.514886°	14.919922°	Seashore - grassland
Nin	Croatia	44.249564°	15.172015°	Seashore - grassland
Vrana Lake	Croatia	43.937292°	15.514689°	Seashore - grassland
Lovćen	Montenegro	42.377169°	18.843117°	Dolomite - bedrock
Skadar Lake	Montenegro	42.326486°	19.069464°	Dolomite - bedrock
Pandurica	Montenegro	42.721628°	18.962442°	Dolomite - bedrock

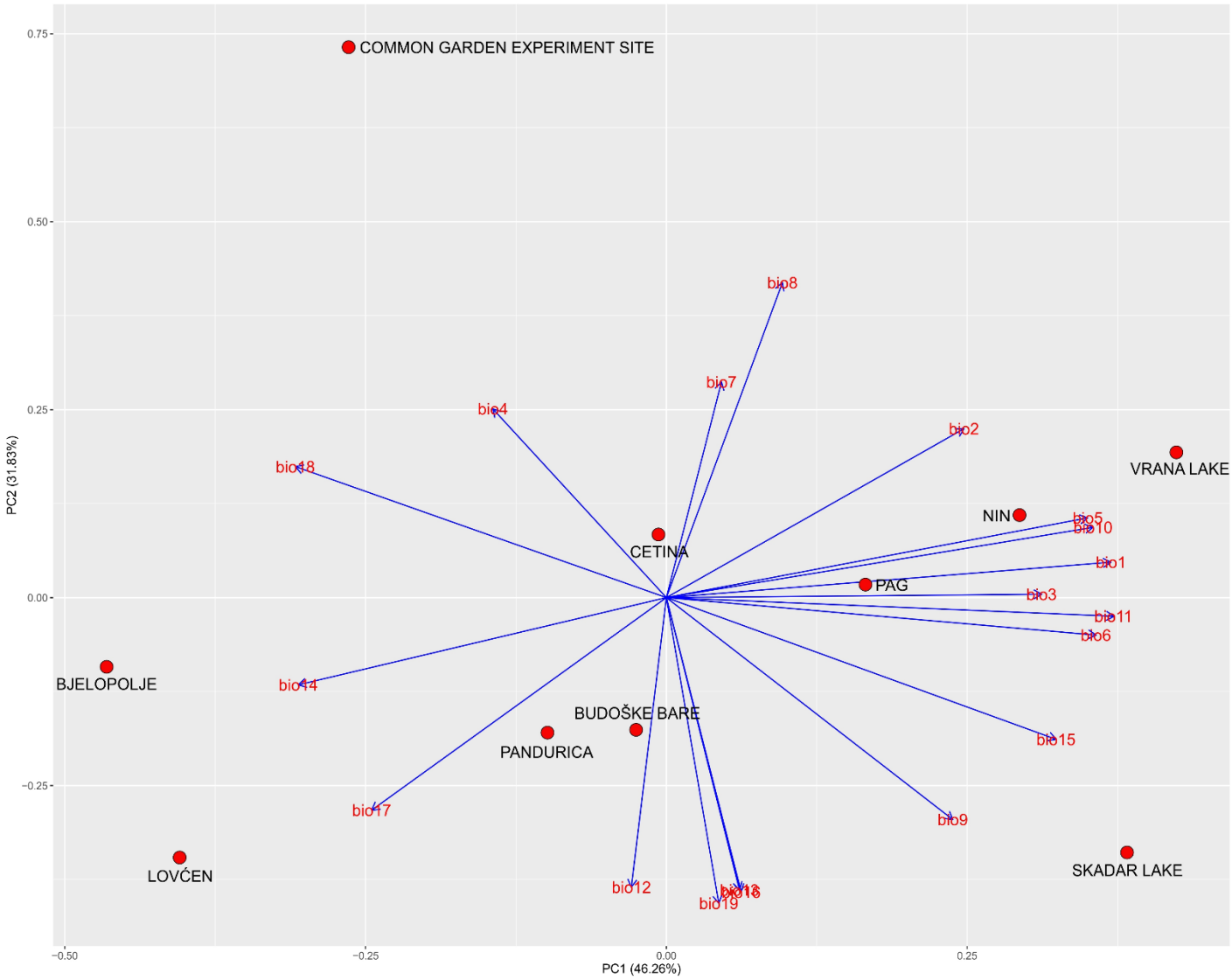


Figure 1. PCA result based on 19 bioclimatic variables for the localities of the studied populations and the common garden experiment site. BIO1 - Annual Mean Temperature; BIO2 - Mean Diurnal Range; BIO3 - Isothermality; BIO4 - Temperature Seasonality; BIO5 - Max Temperature of Warmest Month; BIO6 - Min Temperature of Coldest Month; BIO7 - Temperature Annual Range; BIO8 - Mean Temperature of Wettest Quarter; BIO9 - Mean Temperature of Driest Quarter; BIO10 - Mean Temperature of Warmest Quarter; BIO11 - Mean Temperature of Coldest Quarter; BIO12 - Annual Precipitation; BIO13 - Precipitation of Wettest Month; BIO14 - Precipitation of Driest Month; BIO15 - Precipitation Seasonality; BIO16 - Precipitation of Wettest Quarter; BIO17 - Precipitation of Driest Quarter; BIO18 - Precipitation of Warmest Quarter; BIO19 - Precipitation of Coldest Quarter

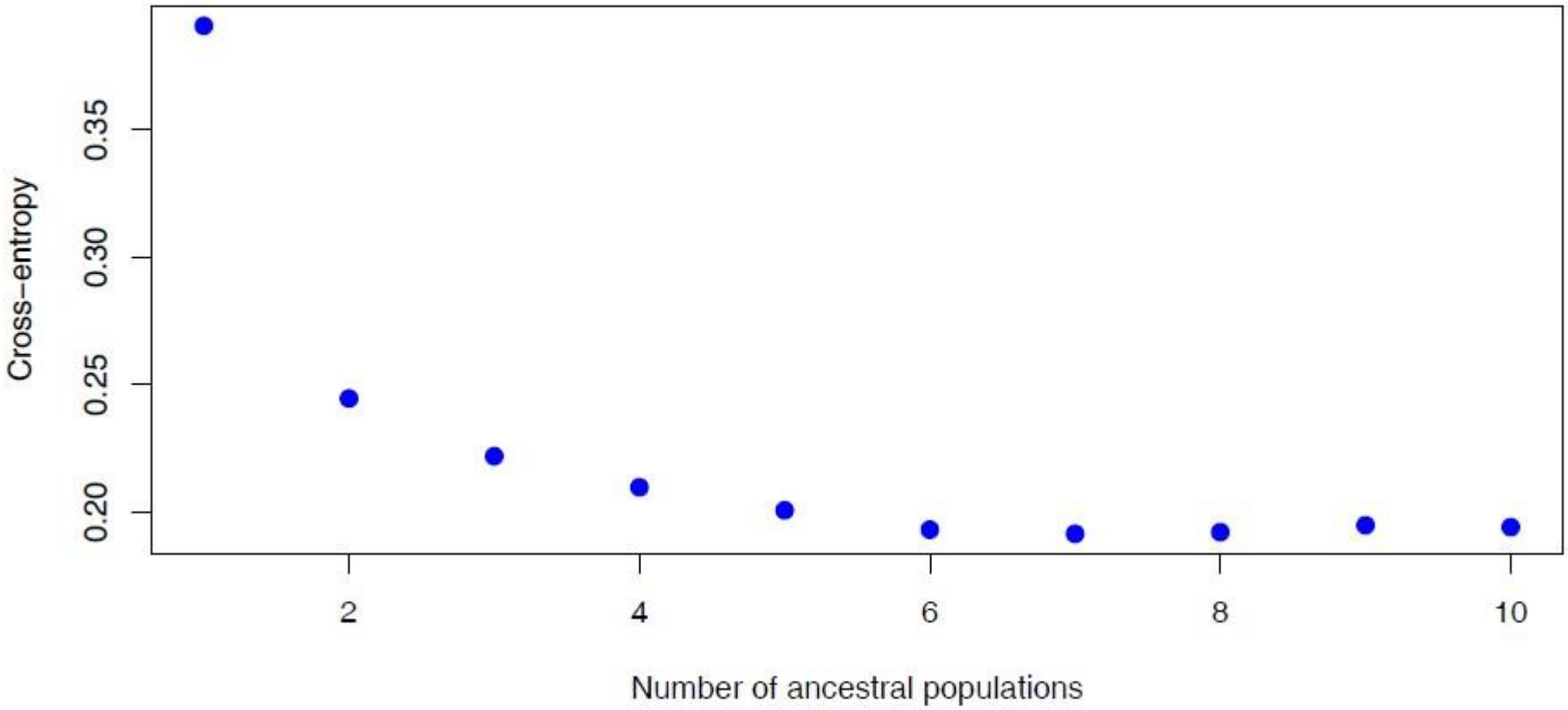


Figure 1. Cross-entropy vs. number of ancestral populations in sNMF analysis on *Chouardia litardierei*. Description of Data: This figure shows the relationship between the number of ancestral populations (K) and the cross-entropy values from an sNMF analysis. Cross-entropy decreases as the number of ancestral populations increases, eventually stabilizing, indicating an optimal K where the model best explains the genetic structure.

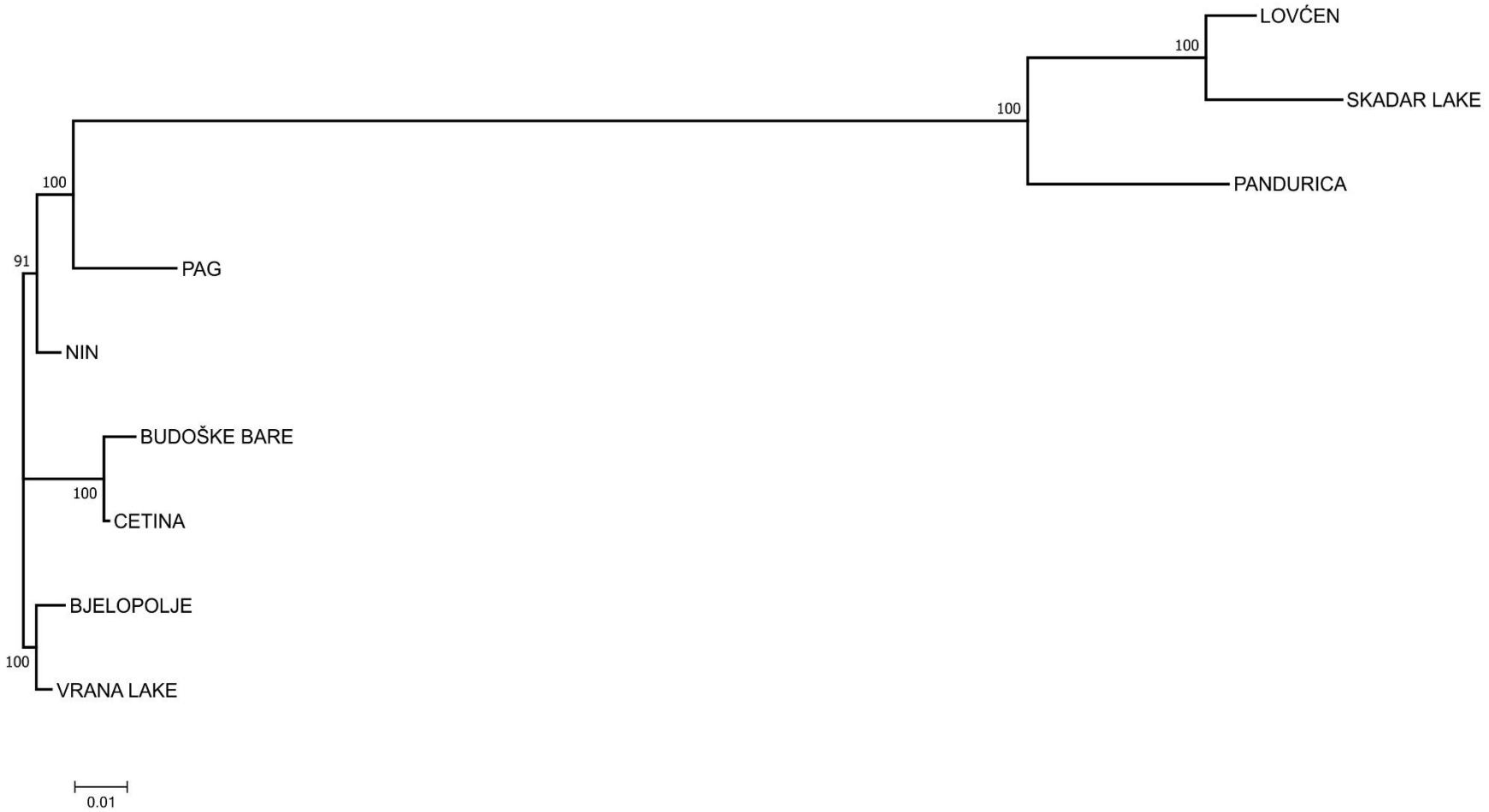


Figure 1. Phylogenetic relationships among *Chouardia litardierei* populations based on Nei's genetic distances. The unrooted phylogenetic tree illustrates the relationships among *C. litardierei* populations based on Nei's genetic distances, highlighting genetic divergence across habitats. The tree was constructed by calculating Nei's genetic distances using the "adegenet" package in R, followed by bootstrapping (1,000 replicates) with the "poppr" package. The final tree was visualized in MEGA7 after conversion to Newick format using the "ape" package.

Table 1. SNPs passing the genome-wide significance threshold in the single-SNP linear mixed model (LMM) analysis in both GEMMA and GMMAT analyses for each reproduction-related morphological trait of *Chouardia litardierei*: TFC, AHI, and BC.

Trait	SNP	Chr	Position	Reference allele	Effect allele	Single-SNP LMM Analysis β (p-value) in GMMAT	Single-SNP LMM Analysis β (p-value) in GEMMA
TFC	15052_36	8	158551682	C	G	5.76×10^{-6} (0.6915)	2.77×10^{-4} (−0.9771)
	475181_57	2	1067391	A	C	1.83×10^{-5} (0.4512)	8.85×10^{-4} (−0.5857)
	682731_21	5	63843877	G	A	3.12×10^{-5} (0.5882)	6.59×10^{-4} (−0.8690)
	356033_30	13	518039464	A	C	4.45×10^{-5} (0.4788)	1.25×10^{-4} (−0.7181)
	439681_33	1	113066723	G	A	6.50×10^{-5} (0.8912)	1.69×10^{-4} (−1.2873)
	380447_37	13	615321041	T	A	8.28×10^{-5} (0.4479)	3.06×10^{-4} (−0.7120)
	45968_38	9	118116906	C	G	9.07×10^{-5} (-0.3201)	7.32×10^{-4} (0.5426)
AHI	62254_22	9	181728136	C	A	1.14×10^{-6} (1.0287)	2.29×10^{-6} (−1.0287)
	423028_13	13	795057155	A	C	1.04×10^{-5} (0.7959)	1.69×10^{-5} (−0.7959)
	423027_44	13	795056910	T	A	3.69×10^{-5} (0.7564)	5.39×10^{-5} (−0.7564)
	357122_13	13	523783140	C	G	7.10×10^{-5} (-0.3847)	9.88×10^{-5} (0.3847)
	230454_16	12	319053038	A	T	7.47×10^{-5} (0.5503)	1.03×10^{-4} (−0.5503)
	536624_21	3	40433867	G	T	1.98×10^{-4} (1.0158)	2.57×10^{-4} (−1.0158)
	593460_76	4	257849493	A	C	2.08×10^{-4} (0.6376)	2.69×10^{-4} (−0.6376)
	659179_34	5	178097173	A	G	2.29×10^{-4} (-0.6336)	2.94×10^{-4} (0.6336)
	534889_25	3	32961390	T	C	2.37×10^{-4} (-0.4168)	3.03×10^{-4} (0.4168)
	679100_46	5	47723772	G	A	2.64×10^{-4} (0.8166)	3.36×10^{-4} (−0.8166)
	669910_120	5	218775782	A	G	2.64×10^{-4} (0.4081)	3.36×10^{-4} (−0.4081)
	167406_45	11	6489057	G	A	2.89×10^{-4} (0.6809)	3.66×10^{-4} (−0.6809)
	299462_80	13	294757456	T	C	3.02×10^{-4} (0.6737)	3.81×10^{-4} (−0.6737)
	137109_19	11	119021216	T	C	3.21×10^{-4} (-0.6327)	4.03×10^{-4} (0.6327)
	228909_34	12	313636308	A	T	3.77×10^{-4} (-0.3962)	4.70×10^{-4} (0.3962)
	218952_18	12	27381126	G	T	4.60×10^{-4} (0.5300)	5.66×10^{-4} (−0.5300)
	631549_20	4	75595465	C	T	5.10×10^{-4} (0.5409)	6.24×10^{-4} (−0.5409)
	96271_20	10	142377502	T	G	5.60×10^{-4} (0.9148)	6.82×10^{-4} (−0.9148)
	53067_33	9	147851683	A	G	5.85×10^{-4} (0.3589)	7.11×10^{-4} (−0.3589)
	730317_18	6	4133503	C	A	6.20×10^{-4} (0.7198)	7.50×10^{-4} (−0.7198)
	403881_18	13	716013518	G	A	6.29×10^{-4} (0.5012)	7.61×10^{-4} (−0.5012)
	383241_14	13	626409144	A	G	7.01×10^{-4} (0.7275)	8.43×10^{-4} (−0.7275)
	377817_17	13	604813128	A	G	7.09×10^{-4} (0.6227)	8.51×10^{-4} (−0.6227)
	360081_66	13	534790629	T	A	7.37×10^{-4} (0.5730)	8.84×10^{-4} (−0.5730)
	794075_23	7	99621462	A	G	7.69×10^{-4} (1.2110)	9.20×10^{-4} (−1.2110)
	104294_72	10	170303215	C	G	8.39×10^{-4} (-0.5422)	9.99×10^{-4} (0.5422)
BC	253435_30	13	107335971	G	A	6.87×10^{-8} (−0.4151)	1.83×10^{-7} (0.4151)
	642566_22	5	104574563	G	A	3.05×10^{-6} (1.0706)	1.98×10^{-7} (1.0706)

	241986_29	12	6101088	G	A	3.19×10^{-6} (−1.2847)	1.62×10^{-7} (1.2847)
	649437_15	5	13468654	A	G	3.36×10^{-6} (−0.8847)	2.31×10^{-7} (0.8847)
	656817_53	5	169059242	T	A	5.16×10^{-6} (−0.6242)	2.63×10^{-7} (0.6242)
	154558_21	11	18248593	T	C	7.21×10^{-6} (−0.5456)	1.92×10^{-7} (0.5456)
	203263_28	12	212399118	T	C	7.24×10^{-6} (−1.1158)	1.07×10^{-6} (1.1158)
	326909_111	13	408253490	T	C	8.22×10^{-6} (0.4218)	3.92×10^{-7} (0.4218)
	713226_25	6	119601819	T	A	1.38×10^{-5} (−0.5002)	4.66×10^{-7} (0.5002)
	76595_24	9	65171073	C	A	1.39×10^{-5} (−0.5538)	1.17×10^{-7} (0.5538)
	22031_53	8	25314160	A	G	1.42×10^{-5} (0.3055)	2.05×10^{-7} (0.3055)
	449864_18	1	146668342	G	T	1.57×10^{-5} (−0.6531)	3.10×10^{-7} (0.6531)
	177168_43	12	105567044	T	G	1.98×10^{-5} (−1.2990)	2.65×10^{-6} (1.2990)
	491467_27	2	39704043	T	G	2.17×10^{-5} (−0.3640)	4.76×10^{-7} (0.3640)
	362848_15	13	546154157	T	C	2.92×10^{-5} (0.5494)	6.82×10^{-7} (0.5494)
	188007_17	12	152850324	A	T	3.03×10^{-5} (−1.0653)	1.25×10^{-6} (1.0653)
	154104_24	11	180657987	A	G	3.97×10^{-5} (−0.4466)	2.81×10^{-7} (0.4466)
	168618_59	11	68980362	A	C	4.11×10^{-5} (0.4063)	2.31×10^{-7} (0.4063)
	178892_42	12	113762962	A	G	4.19×10^{-5} (−0.3068)	1.78×10^{-7} (0.3068)
	688782_16	5	89357080	A	G	4.50×10^{-5} (0.2832)	2.35×10^{-7} (0.2833)
	5022_15	8	122022233	A	C	5.14×10^{-5} (−0.3252)	2.51×10^{-7} (0.3252)
	14652_22	8	157168515	A	G	5.39×10^{-5} (−0.5751)	3.20×10^{-7} (0.5751)
	571070_26	4	17043373	T	C	6.74×10^{-5} (−0.3067)	2.14×10^{-7} (0.3067)
	661577_27	5	185906413	C	G	7.33×10^{-5} (0.8340)	1.29×10^{-6} (0.8341)
	824846_131	12	205849898	G	C	8.13×10^{-5} (−0.4234)	5.42×10^{-7} (0.4234)
	309278_17	13	33471278	A	C	8.52×10^{-5} (−0.5773)	7.51×10^{-7} (0.5773)
	775464_44	7	208947973	A	G	9.47×10^{-5} (0.3681)	4.56×10^{-7} (0.3681)
	236426_14	12	40747253	G	T	9.54×10^{-5} (−0.6723)	8.01×10^{-7} (0.6723)
	426598_87	13	810283382	A	T	9.55×10^{-5} (0.2838)	2.15×10^{-7} (0.2838)
	2730_51	8	110403271	A	C	9.68×10^{-5} (−0.2893)	2.24×10^{-7} (0.2893)
	323413_32	13	39372258	A	G	9.92×10^{-5} (−1.1485)	1.21×10^{-6} (1.1485)
	202604_47	12	210026545	G	A	1.02×10^{-4} (−0.4234)	5.42×10^{-7} (0.4234)
	101366_35	10	160674132	G	A	1.14×10^{-4} (−0.7825)	1.13×10^{-6} (0.7825)
	208646_19	12	234004964	C	T	1.18×10^{-4} (−0.3534)	3.71×10^{-7} (0.3534)
	717130_26	6	136955413	A	G	1.37×10^{-4} (−0.3974)	3.83×10^{-7} (0.3974)
	291775_18	13	26477605	T	C	1.39×10^{-4} (−0.7174)	9.13×10^{-7} (0.7174)
	201641_32	12	205775063	A	C	1.40×10^{-4} (−0.7385)	8.95×10^{-7} (0.7385)
	275195_16	13	197688818	C	T	1.53×10^{-4} (0.3184)	2.93×10^{-7} (0.3184)
	90763_61	10	121305153	G	T	1.57×10^{-4} (−0.2903)	2.45×10^{-7} (0.2903)
	261805_25	13	144317310	T	C	1.73×10^{-4} (−0.4124)	4.90×10^{-7} (0.4124)
	260150_22	13	138653995	G	A	1.75×10^{-4} (0.9271)	1.01×10^{-6} (0.9271)
	206213_23	12	225086121	A	T	1.76×10^{-4} (−0.6725)	7.81×10^{-7} (0.6725)
	726003_42	6	24894175	G	T	1.88×10^{-4} (0.2737)	1.91×10^{-7} (0.2737)
	572458_29	4	175624353	T	C	1.96×10^{-4} (−0.6042)	6.72×10^{-7} (0.6042)
	179046_40	12	114292861	C	T	2.01×10^{-4} (0.5540)	5.90×10^{-7} (0.5540)
	680794_16	5	55125312	A	C	2.27×10^{-4} (−0.5969)	8.73×10^{-7} (0.5969)
	226322_15	12	303921342	T	C	2.46×10^{-4} (−0.9300)	9.18×10^{-7} (0.9300)

	64746_61	9	23100056	T	A	2.51×10^{-4} (−0.3266)	3.06×10^{-7} (0.3266)
	430075_45	13	822183802	T	C	2.72×10^{-4} (−0.2848)	2.46×10^{-7} (0.2848)
	598489_119	4	275215624	C	G	2.84×10^{-4} (0.2950)	3.41×10^{-7} (0.2950)
	360276_31	13	535270309	T	A	3.01×10^{-4} (0.5274)	5.81×10^{-7} (0.5274)
	380447_37	13	615321041	T	A	3.16×10^{-4} (0.4729)	6.48×10^{-7} (0.4729)
	651099_28	5	143602840	A	T	3.17×10^{-4} (−0.4787)	7.36×10^{-7} (0.4787)
	445961_30	1	135087176	C	T	3.24×10^{-4} (−0.2652)	3.67×10^{-7} (0.2652)
	44175_36	9	110634998	T	C	3.44×10^{-4} (0.2755)	2.75×10^{-7} (0.2755)
	331309_70	13	425001652	G	C	3.53×10^{-4} (−0.9734)	1.29×10^{-6} (0.9734)
	252444_137	13	102435018	A	G	3.92×10^{-4} (−0.2657)	3.72×10^{-7} (0.2657)
	648614_50	5	130563585	T	C	3.97×10^{-4} (−0.3560)	4.01×10^{-7} (0.3560)
	89984_62	10	118298581	A	C	4.21×10^{-4} (−0.3812)	4.66×10^{-7} (0.3812)
	531929_17	3	21899113	T	C	4.32×10^{-4} (0.2056)	2.91×10^{-7} (0.2056)
	325231_32	13	401796448	T	A	4.45×10^{-4} (−0.2695)	3.11×10^{-7} (0.2695)
	441718_55	1	121120631	G	A	4.50×10^{-4} (−0.4841)	4.88×10^{-7} (0.4841)
	3438_34	8	113956553	A	G	4.66×10^{-4} (−0.4134)	5.14×10^{-7} (0.4134)
	591772_23	4	250745030	C	T	4.80×10^{-4} (−0.9506)	1.12×10^{-6} (0.9506)
	201876_77	12	206762918	T	C	4.92×10^{-4} (−0.6514)	8.24×10^{-7} (0.6514)
	554483_14	4	101633733	G	T	5.13×10^{-4} (0.5548)	5.94×10^{-7} (0.5548)
	447402_22	1	139232915	C	T	5.15×10^{-4} (−0.5579)	6.07×10^{-7} (0.5579)
	228909_34	12	313636308	A	T	5.37×10^{-4} (−0.3041)	3.42×10^{-7} (0.3041)
	769726_28	7	191343469	A	T	5.41×10^{-4} (0.3239)	2.34×10^{-7} (0.3239)
	66901_34	9	30152315	T	G	5.43×10^{-4} (−0.4763)	5.81×10^{-7} (0.4763)
	272420_33	13	186297710	T	G	5.44×10^{-4} (0.3702)	4.22×10^{-7} (0.3702)
	80554_31	9	7962018	A	G	5.69×10^{-4} (0.6022)	6.55×10^{-7} (0.6022)
	339452_85	13	455653812	A	C	5.71×10^{-4} (−0.2736)	2.35×10^{-7} (0.2736)
	560733_51	4	127515444	T	C	5.82×10^{-4} (0.5123)	5.17×10^{-7} (0.5123)
	546891_140	3	82875693	A	C	5.88×10^{-4} (−0.5194)	5.26×10^{-7} (0.5194)
	71435_14	9	44021754	G	T	5.98×10^{-4} (−0.2466)	3.77×10^{-7} (0.2466)
	293695_19	13	272369031	A	G	6.08×10^{-4} (−0.6642)	8.83×10^{-7} (0.6642)
	392720_57	13	668772912	G	T	6.32×10^{-4} (0.3679)	4.62×10^{-7} (0.3679)
	466588_35	1	89484395	A	C	6.72×10^{-4} (0.4678)	6.56×10^{-7} (0.4678)
	48910_53	9	12893230	T	G	7.13×10^{-4} (0.4002)	4.50×10^{-7} (0.4002)
	763026_20	7	1681982	T	G	7.41×10^{-4} (0.2608)	3.48×10^{-7} (0.2608)
	445498_105	1	133595095	T	G	7.64×10^{-4} (0.3789)	4.90×10^{-7} (0.3789)
	672852_31	5	23571109	A	C	7.82×10^{-4} (−0.7870)	9.53×10^{-7} (0.7870)
	659179_34	5	178097173	A	G	7.97×10^{-4} (−0.4594)	6.72×10^{-7} (0.4594)
	793910_83	7	98929040	C	T	8.09×10^{-4} (0.6528)	8.19×10^{-7} (0.6528)

Statistical analyses were performed with GEMMA and GMMAT. LMM was fitted on 23,315 SNPs. *p*-values < 1×10^{-3} are genome-wide significant. AHI, Average Height of Inflorescences per genotype; BC, Bulbs Count per genotype; Chr, Chromosome; LMM, Linear Mixed Model; SNP, Single Nucleotide Polymorphism; TFC, Total Flower Count per genotype.

Table 1. SNPs identified as having a major sparse effect (PIP > 0.1) on the AHI, TFC, and BC traits of *Chouardia litardierei* in the multi-SNP Bayesian sparse linear mixed model (BSLMM) analysis.

Trait	SNP	Chr	Position	Multi-SNP BSLMM Analysis β (PIP)
TFC	750129_37	7	113120650	-0.324 (0.150)
	750129_37	1	113066723	-0.527 (0.142)
	750129_37	10	105055239	0.295 (0.116)
	750129_37	5	74279195	-0.211 (0.112)
	750129_37	3	144638725	-0.226 (0.110)
	750129_37	1	13072555	-0.225 (0.110)
	750129_37	7	123155593	0.175 (0.104)
	750129_37	6	156375740	0.176 (0.104)
	750129_37	13	793523075	-0.306 (0.095)
AHI	299462_80	13	294757456	-0.429 (0.361)
	377817_17	13	604813128	-0.429 (0.345)
	752051_37	7	123155593	0.327 (0.329)
	383241_14	13	626409144	-0.422 (0.222)
	536624_21	9	40433867	-0.512 (0.206)
	447236_67	1	138736826	-0.307 (0.174)
	487582_35	8	25023093	0.249 (0.170)
	272531_23	13	186930464	-0.455 (0.152)
	643010_44	11	106406258	0.401 (0.150)
	175155_40	5	97957904	-0.387 (0.136)
	565532_39	10	14626431	-0.287 (0.134)
	626164_80	10	52896152	0.287 (0.126)
	51325_24	3	138708252	-0.357 (0.115)
	708355_18	12	101851366	0.398 (0.114)
	206941_26	6	227677762	-0.417 (0.110)
	634062_21	10	8802327	-0.252 (0.109)
	131957_13	4	97222552	-0.341 (0.108)
	543242_19	9	69802074	0.362 (0.106)
	728823_24	12	35185140	0.334 (0.104)
	210123_39	6	239066297	-0.321 (0.103)
	155402_49	5	18495521	-0.217 (0.099)
BC	713226_25	12	119601819	0.448 (0.955)
	241986_29	6	6101088	0.942 (0.909)
	291775_18	13	26477605	0.571 (0.698)
	441718_55	1	121120631	0.425 (0.627)
	381378_15	13	619145251	-0.510 (0.594)
	260150_22	13	138653995	-0.653 (0.547)
	26230_86	8	4273979	0.299 (0.316)
	64746_61	3	23100056	0.256 (0.233)
	755807_21	7	139325995	-0.447 (0.182)
	444730_73	1	13072555	-0.351 (0.130)
	296029_34	13	282177106	-0.469 (0.115)
	293695_19	13	272369031	0.418 (0.108)
	108823_24	4	185496013	-0.504 (0.100)
	490848_21	8	37042132	0.294 (0.097)

BSLMM was fitted on 23,315 SNPs. AHI, Average Height of Inflorescences per genotype; BC, Bulb Count per genotype; BSLMM, Bayesian Sparse Linear Mixed Model; Chr, Chromosome; PIP, Posterior Inclusion Probability; SNP, Single Nucleotide Polymorphism; TFC, Total Flower Count per genotype.

Table 1. Means, medians, and 95% equal tail posterior probability intervals (95% ETPPIs) of hyperparameters estimated from the Bayesian sparse linear mixed model (BSLMM) in reproduction-related morphological traits AHI, TFC, and BC of *Chouardia litardierei*.

Trait	Hyperparameter	Mean	Median	2.5%	97.5%
AHI	h	0.7850	0.7944	0.5869	0.9278
	PVE	0.7195	0.7206	0.5226	0.9037
	rho	0.4134	0.3846	0.0227	0.9378
	PGE	0.3747	0.3489	0.0012	0.9317
	pi	2.51×10^{-2}	1.75×10^{-2}	8.53×10^{-4}	8.40×10^{-2}
	n.gamma	47.40	33	1	159
TFC	h	0.6909	0.7119	0.3569	0.9104
	PVE	0.5598	0.5561	0.2571	0.8861
	rho	0.3299	0.2767	0.0115	0.8950
	PGE	0.2878	0.2415	0.0000	0.8627
	pi	4.13×10^{-2}	2.15×10^{-2}	6.31×10^{-4}	1.46×10^{-1}
	n.gamma	77.57	40	0	274
BC	h	0.6409	0.6537	0.3717	0.8404
	PVE	0.6987	0.7019	0.5546	0.8277
	rho	0.8856	0.9281	0.5359	0.9976
	PGE	0.8915	0.9596	0.2568	0.9986
	pi	9.96×10^{-3}	8.23×10^{-2}	1.96×10^{-3}	2.65×10^{-2}
	n.gamma	18.09	15	5	44

BSLMM was fitted on 23,315 SNPs. AHI, Average Height of Inflorescences per genotype; BC, Bulb Count per genotype; h, approximation to the proportion of phenotypic variance explained by variants (PVE); n.gamma, number of variants with major effect; PGE, Proportion of Genetic Variance explained by variants with major effect; pi, proportion of variants with non-zero effects; PVE, Proportion of Phenotypic Variance explained by variants; rho, approximation to the proportion of genetic variance explained by variants with major effect; TFC, Total Flower Count per genotype.

Table 1. SNPs passing the genome-wide significance threshold ($p < 1 \times 10^{-3}$) in the multivariate linear mixed model (mvLMM) analysis for AHI and TFC traits of *Chouardia litardierei* in GEMMA multivariate GWAS.

SNP	Chr	Position	Reference allele	Effect allele	Beta1 (AHI)	Beta2 (TFC)	mvLMM Analysis in GEMMA (*p*-value)
62254_22	9	181728136	A	C	-1.0233	-0.6464	5.17×10^{-6}
423028_13	13	795057155	C	A	-0.8016	-0.3108	3.90×10^{-5}
500624_24	2	7834046	A	G	-0.3194	0.7410	6.34×10^{-5}
588341_19	10	236169480	T	C	-0.2150	0.4102	6.42×10^{-5}
230454_16	12	319053038	T	A	-0.5337	-0.5133	1.53×10^{-4}
423027_44	13	795056910	A	T	-0.7603	-0.3483	1.56×10^{-4}
730317_18	6	4133503	A	C	-0.7169	0.0446	1.76×10^{-4}
356033_30	13	518039464	C	A	-0.2531	-0.7405	1.98×10^{-4}
463223_64	1	73180436	C	T	0.3163	-0.2853	2.42×10^{-4}
561279_13	10	130303684	G	T	0.4456	1.4653	2.56×10^{-4}
208200_70	12	23218154	T	G	-0.0628	0.5553	2.71×10^{-4}
357122_13	13	523783140	G	C	0.3769	0.3095	2.77×10^{-4}
666088_105	11	204738976	G	A	-0.2455	0.3554	3.11×10^{-4}
45968_38	9	118116906	G	C	-0.0231	0.5586	3.49×10^{-4}
81506_26	9	82807440	C	A	-0.3177	0.5778	3.69×10^{-4}
275195_16	13	197688818	T	C	-0.2214	-0.4944	3.77×10^{-4}
341030_70	13	460998239	G	T	-0.1489	1.0188	3.77×10^{-4}
144029_28	5	146787421	T	C	-0.2007	-0.4054	4.56×10^{-4}
593460_76	10	257849493	C	A	-0.6385	-0.1012	4.60×10^{-4}
750129_37	7	113120650	A	G	-0.5378	-0.6841	4.68×10^{-4}
197966_30	12	18977810	A	G	-0.2716	0.1541	4.71×10^{-4}
445498_105	1	133595095	G	T	-0.4377	-0.6521	4.79×10^{-4}
285797_40	13	241129049	G	T	0.0988	-0.4041	5.20×10^{-4}
205633_17	12	222557156	G	T	0.0332	0.6208	5.55×10^{-4}
669910_120	11	218775782	G	A	-0.4145	-0.1086	5.88×10^{-4}
20739_90	8	20177276	T	C	-0.2182	0.3899	6.14×10^{-4}
48553_26	9	127546435	C	T	-0.6902	0.0702	6.37×10^{-4}
774988_31	7	207552004	A	G	-0.2538	0.1140	6.54×10^{-4}
518385_24	9	129492878	T	C	-0.5431	0.1163	6.73×10^{-4}
48552_21	9	127546136	G	A	-0.6851	0.0752	6.88×10^{-4}
58193_25	12	167644306	G	T	-0.2927	0.4614	6.94×10^{-4}
203254_20	12	212353432	G	A	1.0385	1.9815	7.13×10^{-4}
186978_19	12	148693882	C	A	-0.6745	-0.8306	7.23×10^{-4}
206941_26	5	227677762	A	T	-0.9359	-0.9880	7.32×10^{-4}
135671_24	9	112993797	C	T	0.8214	0.9700	8.33×10^{-4}
536624_21	13	40433867	T	G	-0.9858	-0.8117	8.35×10^{-4}
423033_32	3	795167985	C	T	0.2732	0.3570	8.80×10^{-4}
64618_14	11	22823004	G	T	-0.1080	0.2438	8.86×10^{-4}
100510_16	6	157555197	G	T	-0.6082	0.0212	9.29×10^{-4}
679100_46	7	47723772	A	G	-0.8230	-0.4745	9.56×10^{-4}
723260_21	9	159782442	T	A	-0.4675	0.1541	9.79×10^{-4}
299462_80	13	294757456	C	T	-0.6833	-0.4259	9.92×10^{-4}

mvLMM in GEMMA was fitted on 23,315 SNPs. AHI, Average Height of Inflorescences per genotype; Chr, Chromosome; mvLMM, multivariate Linear Mixed Model; SNP, Single Nucleotide Polymorphism; TFC, Total Flower Count per genotype.

Table 1. SNPs passing the genome-wide significance threshold ($p < 1 \times 10^{-3}$) in the multivariate linear mixed model (mvLMM) analysis for AHI and BC traits of *Chouardia litardierei* in GEMMA multivariate GWAS.

SNP	Chr	Position	Reference allele	Effect allele	Beta1 (AHI)	Beta2 (BC)	mvLMM Analysis in GEMMA (*p*-value)
736175_27	6	67903776	C	G	0.5979	1.415	1.03×10^{-7}
364607_20	13	552521276	T	G	-15.809	0.2513	4.56×10^{-6}
176612_27	12	102085689	T	G	0.2323	0.5946	9.45×10^{-6}
773973_55	7	204445848	A	G	-12.131	0.3277	1.17×10^{-5}
733886_34	6	5721001	G	T	0.2418	0.7287	1.35×10^{-5}
524058_56	3	148761312	T	C	-0.4348	-0.3413	1.38×10^{-5}
264616_23	13	154205990	T	C	0.5185	0.8123	1.44×10^{-5}
18006_76	8	168175276	G	A	0.5963	0.7048	1.67×10^{-5}
224704_28	12	297710995	T	G	0.5639	0.7713	1.92×10^{-5}
538105_25	3	45117962	C	A	0.8985	13.156	1.95×10^{-5}
677181_33	5	3870016	G	A	-0.1287	-0.739	3.65×10^{-5}
108764_25	10	185320554	T	G	-0.3746	0.443	4.14×10^{-5}
524057_35	3	148761135	C	A	0.4288	0.3377	6.14×10^{-5}
92521_57	10	128449310	A	C	0.1445	0.4099	1.03×10^{-4}
769653_97	7	191131828	A	G	0.6619	0.7687	1.05×10^{-4}
77194_22	9	6766688	T	G	-0.4583	-0.1688	1.18×10^{-4}
766497_33	7	180578950	A	G	-0.02	0.3902	1.34×10^{-4}
674817_56	5	30528207	C	G	0.4579	0.7237	1.36×10^{-4}
643010_44	5	106406258	A	C	0.4711	-0.4387	1.43×10^{-4}
368751_23	13	570896987	T	A	0.8173	12.086	1.78×10^{-4}
668508_28	5	213959178	A	C	-0.5816	0.1643	1.97×10^{-4}
468248_25	1	96988857	A	T	-0.7511	0.6573	1.97×10^{-4}
630601_22	4	7135974	T	C	-0.2828	-0.4859	2.04×10^{-4}
274513_23	13	195179209	C	A	0.6461	-0.2094	2.07×10^{-4}
529123_19	3	168860131	T	A	0.8623	-0.0234	2.45×10^{-4}
167780_52	11	66195928	G	C	0.3489	0.8378	2.54×10^{-4}
195619_17	12	181011473	A	T	-0.3514	0.9959	2.63×10^{-4}
772778_13	7	200837811	C	A	-0.6746	0.1614	2.90×10^{-4}
529321_19	3	16953884	T	C	-0.8858	-0.7266	3.09×10^{-4}
211566_37	12	243839810	A	T	0.4811	0.0241	3.20×10^{-4}
585482_58	4	226159268	A	T	0.1996	0.7592	3.21×10^{-4}
442856_30	1	124869628	A	C	-0.3202	-0.4025	3.25×10^{-4}
112990_33	10	199377871	C	A	0.3393	-0.217	3.54×10^{-4}
775883_28	7	21276111	A	G	0.9731	0.6984	3.65×10^{-4}
178892_42	12	113762962	A	G	0.3957	0.2169	3.92×10^{-4}
265427_20	13	157460474	A	G	0.4658	-0.4725	4.12×10^{-4}
708607_31	6	102551803	G	A	0.6296	-0.3405	4.36×10^{-4}
146609_18	11	156264743	A	G	0.3416	0.0919	4.37×10^{-4}
229378_26	12	315229318	C	T	0.2088	0.3259	4.50×10^{-4}
70626_65	9	41503693	T	C	-0.3768	-10.127	4.54×10^{-4}
322_26	8	101414222	C	T	0.8316	-0.4206	5.01×10^{-4}
149663_69	11	166136268	T	C	0.8973	-0.1638	5.09×10^{-4}
365305_47	13	555442390	A	G	0.6108	11.891	5.10×10^{-4}
360080_19	13	534790477	C	T	0.6631	13.626	5.24×10^{-4}
20248_53	8	18385940	A	G	0.3844	0.5787	5.30×10^{-4}
9058_15	8	137969173	A	G	0.7381	0.8322	5.47×10^{-4}
170461_26	11	7610629	A	G	-0.3711	-0.1396	5.70×10^{-4}
13864_52	8	154615308	C	A	0.5852	0.2518	5.73×10^{-4}

427675_39	13	814235512	A	T	0.2918	0.6833	5.73×10^{-4}
774577_45	7	206386106	T	A	0.0031	0.636	6.18×10^{-4}
650985_26	5	143049003	C	T	0.4261	12.124	6.46×10^{-4}
771180_71	7	196132350	C	A	-0.5368	-0.3073	6.71×10^{-4}
68322_22	9	34484978	G	A	-0.4129	-0.225	6.85×10^{-4}
667179_22	5	208973401	T	G	-0.6511	0.0569	7.68×10^{-4}
238115_20	12	46592935	A	T	0.2443	10.283	7.86×10^{-4}
738313_35	6	8070864	G	T	0.2117	-0.3262	8.41×10^{-4}
718968_54	6	14451537	A	G	0.677	0.8042	8.45×10^{-4}
715340_33	6	128544256	T	A	-0.5495	0.35	8.51×10^{-4}
473972_18	2	102121380	A	G	0.1559	0.6482	8.90×10^{-4}
240119_70	12	54071168	T	A	-10.754	-0.0847	8.97×10^{-4}
157538_21	11	25491070	A	C	-0.1713	0.9383	9.30×10^{-4}
38147_41	8	90743188	G	T	0.1764	0.4746	9.42×10^{-4}
22031_53	8	25314160	A	G	-0.0928	0.2756	9.69×10^{-4}
634131_58	4	885019	A	G	0.1445	1.074	9.88×10^{-4}

mvLMM in GEMMA was fitted on 23,315 SNPs. AHI, Average Height of Inflorescences per genotype; Chr, Chromosome; BC, Bulb count; mvLMM, multivariate Linear Mixed Model; SNP, Single Nucleotide Polymorphism.

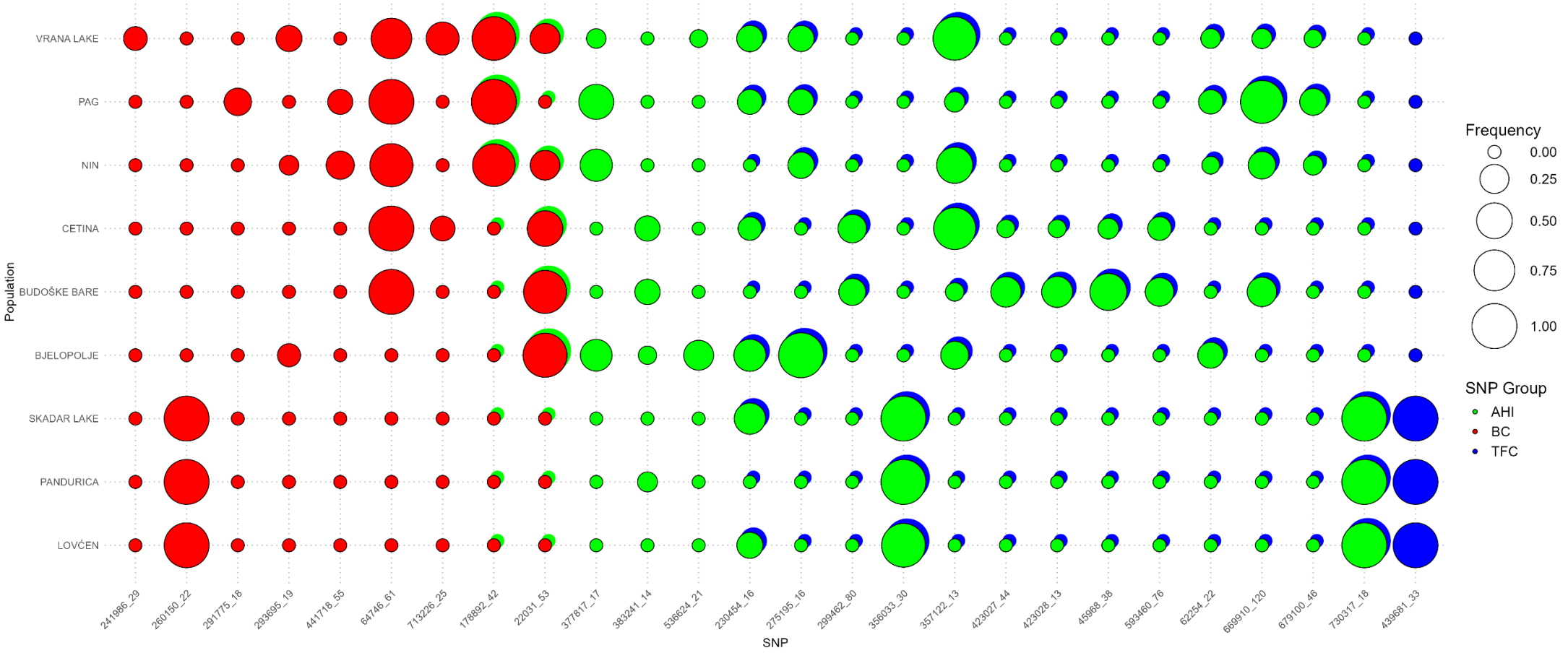


Figure 1. Frequency of effect alleles across populations for significant SNPs identified in the single-SNP LMM analysis (GEMMA and GMMAT), as well as the multi-SNP BSLMM analysis, all of which surpassed the genome-wide significance threshold (1×10^{-3}). The analysis also includes SNPs meeting the same threshold in the multivariate GWAS. The corresponding SNPs are detailed in Table 2 and Table 4 of the manuscript. Overlapping points of different colors represent SNPs associated with different traits, with each color corresponding to a specific trait. The overlap occurs because some SNPs are shared across traits, leading to their placement one in front of the other. AHI, Average Height of Inflorescences; BC, Bulb Count; TFC, Total Flower Count.

Due to size constraints, the following additional files are not reproduced in this thesis but can be accessed at the journal's website (BMC Plant Biology, DOI: 10.1186/s12870-025-06617-4):

Additional File 5: EggNOG output file for the 83 most significant SNP loci associated with four distinct bioclimatic variables, identified as being most relevant to the traits under investigation.

Additional File 12: EggNOG output file for 12 SNPs that exceeded the genome-wide significance threshold (1×10^{-3}) in both the single-SNP LMM and multi-SNP BSLMM analyses of the *Chouardia litardierei* traits: TFC, AHI, and BC.

Additional File 13: EggNOG output file for 13 SNP loci that exceeded the genome-wide significance threshold (1×10^{-3}) in the multivariate GWAS analysis of the *Chouardia litardierei* traits: AHI and TFC.

Additional File 14: EggNOG output file for 2 SNP loci that exceeded the genome-wide significance threshold (1×10^{-3}) in the multivariate GWAS analysis of the *Chouardia litardierei* traits: AHI and BC.

Curriculum vitae

Sara Laura Šarančić was born in Zagreb, Croatia, on April 24th, 1997. She completed both her elementary and high school education in Zagreb and went on to pursue higher education at the University of Zagreb, Faculty of Science, Department of Biology. During her studies, she received multiple recognitions, including the University of Zagreb Rector's Award, a STEM Scholarship, and the University of Zagreb Scholarship. She earned her master's degree in biology and chemistry education in 2021 from the same institution, completing her thesis in the Laboratory for Genetic Diversity, Phylogeny, and Molecular Plant Systematics under the mentorship of Prof. Zlatko Liber. Since 2021, Sara has been employed as a research assistant and enrolled in a PhD program in Biology at the Faculty of Science, University of Zagreb. She is currently working under the supervision of the Assoc. Prof. Ivan Radosavljević on a Croatian Science Foundation-funded project titled "Amethyst Meadow Squill (*Chouardia litardierei*, Hyacinthaceae): A Study System for Ecological Divergence" (HRZZ-IP-2020-02-8099). She has actively participated in seven scientific workshops and presented her research at international conferences, delivering two oral talks and contributing five posters. As a member of the Croatian Botanical Society, Sara is committed to furthering the understanding of plant evolutionary biology through innovative, data-driven research.

Scientific activities and publications

SCIENTIFIC ID in Croatian Science Bibliography: 37820

Original scientific papers:

Radosavljević, I., Križanović, K., **Šarančić, S. L.**, and Jakše, J. (2023). Towards the Investigation of the Adaptive Divergence in a Species of Exceptional Ecological Plasticity: Chromosome-Scale Genome Assembly of *Chouardia litardierei* (Hyacinthaceae). *Int J Mol Sci* 24, 10755. doi: 10.3390/ijms241310755

Šarančić, S. L., Pleić, N., Križanović, K., Surina, B., Mitić, D., and Radosavljević, I. (2025a). Uncovering the genomic basis of phenological traits in *Chouardia litardierei* (Asparagaceae) through a genome-wide association study (GWAS). *Front Plant Sci* 16, 1571608. doi: 10.3389/FPLS.2025.1571608

Šarančić, S. L., Pleić, N., Mitić, D., Križanović, K., Surina, B., and Radosavljević, I. (2025b). Genome-wide association study (GWAS) provides insights into the genomic basis of reproduction-related traits in *Chouardia litardierei* (Asparagaceae). *BMC Plant Biology 2025 25:1* 25, 1–25. doi: 10.1186/S12870-025-06617-4

Conference abstracts:

Šarančić, Sara Laura; Pleić, Nikolina; Križanović, Krešimir; Radosavljević, Ivan Unraveling the Genomic Blueprint: Insights into Phenological and Reproduction-Related Morphological Traits in *Chouardia litardierei* (Hyacinthaceae) via GWAS Analysis// 8th Faculty of Science PhD Student Symposium: Book of Abstracts / Posarić, Laura; Gmižić, Daria; Ostojić, Tea et al. (ur.).Zagreb: Faculty of Science, University of Zagreb, Zagreb, Croatia, 2024. str. 24-24

Šarančić, Sara Laura; Mitić, Damjan; Križanović, Krešimir; Radosavljević, Ivan Genome-environment association study uncovers loci associated with local adaptation in *Chouardia litardierei* (Hyacinthaceae) // XX International Botanical Congress IBC 2024, Spain: Book of Abstracts. Posters.Madrid: Fase 20 Ediciones, 2024. str. 334-334

Šarančić, Sara Laura; Pleić, Nikolina; Križanović, Krešimir; Radosavljević, Ivan. The genetic architecture of some phenological and reproduction-related morphological traits in *Chouardia litardierei* (Hyacinthaceae) as revealed by GWAS // XX International Botanical Congress IBC 2024, Spain: Book of Abstracts. Posters.Madrid: Fase 20 Ediciones, 2024. str. 333-333

Šarančić, Sara Laura; Križanović, Krešimir; Jakše, Jernej; Radosavljević, Ivan A chromosome-level genome assembly of *Chouardia litardierei* (Hyacinthaceae), a plant characterized by extreme ecological plasticity // Proceedings of Genetika 2022. Ljubljana: Genetic Society of Slovenia, Ljubljana, 2022. str. 101-101

Šarančić, Sara Laura; Surina, Boštjan; Dragičević, Snežana; Glasnović, Peter; Radosavljević, Ivan. Morphological and phenological characterization of selected *Chouardia litardierei* and *C. lakusicii* populations // Book of Abstracts. Zagreb, 2022. str. 57-57

Šarančić, Sara Laura; Križanović, Krešimir; Jakše, Jernej; Radosavljević, Ivan *Chouardia litardierei* (Hyacinthaceae) genome sequence // Book of Abstracts. Zagreb, 2022. str. 56-56

Scientific projects:

(2021 – 2025) Research Assistant on HRZZ-IP-2020–02-8099 project: Amethyst meadow squill (*Chouardia litardierei*, Hyacinthaceae): a study system for ecological divergence

Workshops:

(2024) Introduction to the Galaxy Platform, University Computing Centre SRCE, University of Zagreb

(2023) UNIX for Bioinformatics in the analysis of genetic data, Faculty of Forestry and Wood Technology, University of Zagreb

(2023) Introduction to Statistics and R, BICRO Biocenter-Exaltum, Zagreb

(2022) RADseq for phylogenOMICS, Georg-August-Universität Göttingen, online

(2022) Introduction to genome-wide association studies, Physalia Courses, online

(2022) Adaptation Genomics, Physalia Courses, online

(2022) Introduction to R syntax and its application in basic statistical and graphical data analysis, University Computing Centre SRCE, University of Zagreb