THESIS

Presented for the degree of
DOCTOR IN PLANT BIOTECHNOLOGY

Supervised by:
Pontificia Universidad Católica de Chile, Facultad de Ciencias Biológicas,
Facultad de Agronomía e Ingeniería Forestal

Co-supervised by:
University of Bordeaux,
École Doctorale des Sciences de la vie et de la Santé

-

ANALYSIS OF THE METABOLIC FEATURES OF PLANT EXTREMOPHILE SPECIES FROM THE ATACAMA DESERT

-

Thomas DUSSARRAT

Publicly defended on 20 April 2022

Director	Mr. Rodrigo Antonio GUTIÉRREZ – Professor, *Pontificia Universidad Católica de Chile*
Co-directors	Mr. Dominique ROLIN – Professor, *University of Bordeaux* Mr. Pierre PÉTRIACQ – Associate Professor, *University of Bordeaux*

Thesis Committee

President	Mr. José O'BRIEN – Professor, *Pontificia Universidad Católica de Chile*
Rapporteurs	Ms. Anne Emmanuelle HAY DE BETTIGNIES – Associate Professor, *University Claude Bernard, Lyon 1* Mr. Cédric BERTRAND – Professor, *University of Perpignan*
Examinator	Mr. Michel HERNOULD – Professor, *University of Bordeaux*
Guests	Ms. Inés Marlene ROSALES – Professor, *Pontificia Universidad Católica de Chile* Mr. Rodrigo Antonio GUTIÉRREZ – Professor, *Pontificia Universidad Católica de Chile*

ACKNOWLEDGEMENTS

ACKNOWLEDGEMENTS OF THE FUNDING SOURCE

This work was supported by the VRI and Cotutelle doctoral fellowships. I warmly thank all the members of the Pontificia Universidad Católica de Chile and the University of Bordeaux who participated directly or indirectly in the establishment of this Cotutelle project. In addition, I am sincerely grateful to Rodrigo Gutiérrez's laboratory for its financial support during the last months of this project.

We thank the Fondo Nacional de Desarrollo Científico y Technológico (FONDECYT 1180759), the Fondo de Desarrollo de Áreas Prioritarias (FONDAP), DOI EVONET Proyect and ANID – Millennium Science Initiative Program – iBio ICN17_022 for financial support to Rodrigo A. Gutiérrez. We thank MetaboHUB (ANR-11-INBS-0010) and PHENOME (ANR-11-INBS-0012) for financial support to INRAE. Finally, we are grateful to the Comunidad de Talabre for accepting our fieldwork in the Atacama Desert over the past three years.

ACKNOWLEDGEMENTS

First, I would like to thank the members of the jury, José O'Brien, Inés Marlene Rosales, Cédric Bertrand, Anne Emmanuelle Hay De Bettignies and Michel Hernould for accepting to evaluate my work.

I would also like to thank the directors of the two teams who welcomed me in their laboratory and allowed me to evolve in a framework favouring my professional and personal development. Special thanks to Yves Gibon, head of the *Meta* team, with whom I spent several hours discussing science, music, culture, and personal and professional future. Our discussions at the doctoral meetings stimulated my scientific curiosity and brought new ideas or challenges to this wonderful project. Our informal exchanges have always been a source of profound reflections (which I greatly appreciate) on the place of science in society and on my place in the scientific world. Besides, these discussions have shown me that no matter how high the scientific level reached, it is possible to keep it simple and preserve that fabulous curiosity and desire to embark on new scientific puzzles. This curiosity is shared with Rodrigo Gutiérrez, head of the Plant Systems Biology lab, who agreed to supervise a thesis in an unknown metabolic terrain carried out by a student who presented himself at a conference with the simple question "would you be interested in analysing the metabolic profiles of plants in the Atacama desert?". Dear lab managers, I can only be thankful for your management qualities and your open-mindedness, which greatly facilitated my adaptation to your laboratory.

I think there are no words to describe how lucky I am to have had Rodrigo Gutiérrez, Dominique Rolin and Pierre Pétriacq as thesis directors. I am a very simple person, who likes science because it is an endless source of enigmas and challenges and because it allows me to interact and learn from others. This thesis allowed me to meet these three men who used their scientific and human knowledge to incredibly boost my scientific and personal learning. Thanks to them I felt free and supported all the time. I remember one of the sentences pronounced by one of them: "Better to have friends than a paper, no matter how good the paper is", a sentence I won't forget. Besides, trust has been the central point of this collaboration. This ability to guide a student without hindering his ideas is innate in you and has enabled the projects detailed in this manuscript to flourish. In my opinion, two things define a great researcher and teacher: scientific knowledge and the ability to transmit this knowledge. I thank you infinitely for excelling in these two aspects. Rodrigo, I admire your professionalism, your willingness to make science accessible to all and the support you give to your students. Dominique, I have always been fascinated by your ability to perceive people's qualities and to establish a precise plan to extract the best from each student or colleague. I consider you my scientific father and I am deeply grateful to you for being who you are. Pierre, you are for me a model of scientific success that I would like to reach one day. I thank you wholeheartedly for your kindness, your charisma, your legendary working capacity and your unfailing support during these three years. There will come a time when it becomes my turn to transmit my knowledge, and you can be sure that when that day comes, I will inspire myself with each of your virtues and advice.

I would also like to share some words on other brilliant scientists who have helped me throughout this project. Claudio Latorre and Francisca Díaz, what journeys we have made in the Atacama Desert! I am so grateful to you for sharing your incredible knowledge of ecology with me. Francisca, your kindness, your trust and your support helped me a lot with the interpretation of the results and allowed the birth of a project combining ecology and metabolomics! My favourite bioinformaticians: Tomás Moyano and Sylvain Prigent. You offered me your help in bioinformatics, which not only made it easier to handle the data but also allowed better exploration of the data produced. What a journey it's been since I first entered your office with a rather freaky question: "Would it be possible to predict plant elevation from metabolism?", right Sylvain? I don't know how to thank you either. Just know that you have contributed to my scientific fulfilment and surely to my career. Thank you superkind Cookie! I would also like to warmly thank Amélie Flandin who helped me from day one with the grinding, LCMS analysis and figure design! I will miss your *joie de vivre* and your stories!

I would also like to thank all the people with whom I interacted with during this project: Sophie Colombie, Jérôme Joubes, Ricardo Nilo, Jean Baptiste Ramon, Fabien Jourdan, Guillaume Tcherkez, Christiane Gallet, Anne Emmanuelle Hay. Some of these exchanges have led to additional projects that allowed me to broaden my skills while learning from your analytical technique, whereas others have

simply shared their ideas, comments and advice. Science is also about helping each other and I thank you for these enriching exchanges!

I am extremely grateful to the people who helped me directly or indirectly to establish the Cotutelle project. First of all, thanks to Michel Hernould and José O'Brien for giving me the opportunity to be involved in the joint Master programme. This opportunity marked the beginning of my scientific career and I hope that one day I will be able to help this collaboration and its future students in some way. I also would like to thank Antoine de Daruvar, without your kindness and passion for Chile, I am not sure that this project would have been possible. Finally, a special thanks to Wendy Wong Jimenez! What a wonderful person you are. I don't know what I would have done without you and your patience with the Chilean administration!

I am also grateful to all the people who helped me with this project. Cédric Cassan, I thank you so much for all the help you have given me! The chemical assays were not always easy but you remained as nice as ever... and we had a lot of fun anyway, right?! Melissa Aguilar and Jonathan Maldonado, a warm thanks to you as well! You helped me a lot with my adaptation in the lab and your positive attitude in the lab or the desert was truly nice. Stéphane Bernillon, I am sincerely grateful for all the knowledge you have transmitted to me. Learning how the LCMS works has been great fun with you! Catherine Deborde, I would also like to thank you for your help with the NMR analysis, we are certainly going to do something great with this! A special thanks to many other people who participated in this project: Mickael, Pierrot, Daniel, Annick, Carolina, Grace, Viviana, Jonathan, Soledad, Karem, Marie-Lou, Florence... And more generally, all the team of both laboratories. The thesis was a real pleasure for me, and it is largely thanks to you. And by the way... thanks for teaching me Spanish!

A HUGE thank you to my friends from the PhD students' office, Gabriel, Guillaume (it was a pleasure to work with you man!), Alice, Chloé, JJ and Théo. With you, I never had time to get bored, either in the laboratory or outside. I have no doubt that this is not the last time we will share a conversation or a drink, but I sincerely thank you for making this office a benevolent and funny place. For the younger ones, good luck for the future! I am sure with you the future of this office is assured!

Finally, I thank more than anything else my parents and my sister who have made me who I am and who have always given me the resources to achieve my studies and projects. Copain, Benjam, Joy’s, Vin’s, Claire, Syndelle, Pauline, you have always been there for me and you allowed me to take my mind off things whenever I needed it! I promise Neurons, in a few days we'll be back together for some crazy adventures!

MERCI POUR TOUT!

I hope you enjoy these few lines and this manuscript, which describes the results of our project.

ABSTRACT OF THE THESIS

Extreme lands lying at the edges of at least one abiotic gradient permit the survival of extremely few species. These so-called extremophile species (literally loving "*philos*" the extremes) harbour a unique reservoir of genetic and biochemical adaptations that has always attracted human curiosity. Previous studies have shown a high degree of species-specificity for plant adaptation to hostile biomes, thus explaining that successful transfers of protective mechanisms to crops remain scant. However, generic adaptive strategies may also exist. In this context, I propose to carry out a comprehensive approach from the ecosystem to the metabolites to investigate the biochemical adjustments of extremophile plant species from the Atacama Desert, the driest non-polar desert on earth. Plants were collected in their natural environment that spans an elevation gradient from 2500 to 4500m. Multiple metabolomic approaches were combined with machine learning to unveil a generic toolbox for plant resilience to harsh conditions. Subsequently, reaction and pathway enrichment analyses identified genetic legacies underlying convergent biochemical strategies selected through evolution. Finally, the role of positive interactions with the cactus *Maihueniopsis camachoi* in the adaptation of various plant species to harsh environments was explored. Results yielded a better mechanistic understanding of facilitation processes and the discovery of an intriguing set of metabolites able to predict the interaction status. Overall, while this study provided significant insights into our comprehension of adaptive mechanisms underlying plant resilience to extreme climates, our multi-species approach foreshadows promising studies and discoveries in agronomy and ecology.

RESUMEN DE LA TESIS

Los sitios extremos situados en los márgenes de al menos un gradiente abiótico permiten la supervivencia de muy pocas especies. Estas especies denominadas extremófilas (literalmente que aman los extremos) albergan una reserva única de adaptaciones genéticas y bioquímicas que siempre han atraído la curiosidad humana. Estudios anteriores han demostrado un alto grado de especificidad para la adaptación de las especies de plantas a biomas hostiles, lo que explica que las transferencias exitosas de esas adaptaciones a los cultivos sigan siendo escasas. Sin embargo, también pueden existir estrategias adaptativas genéricas o más generales. En este contexto, me propongo utilizar un enfoque integral desde el ecosistema hasta los metabolitos para investigar las adaptaciones bioquímicas de las especies vegetales extremófilas del desierto de Atacama, el desierto no polar más seco del planeta. Las plantas se recogieron en su entorno natural, que abarca un gradiente de altitud de 2500 a 4500m. Se combinaron múltiples enfoques metabolómicos con el aprendizaje automático o "machine learning" para develar una serie de herramientas genéricas para la resistencia de las plantas a las duras condiciones del Atacama. Posteriormente, los análisis de enriquecimiento de reacciones y vías metabólicas identificaron los legados genéticos subyacentes a las estrategias bioquímicas convergentes seleccionadas evolutivamente. Por último, se exploró el rol de las interacciones positivas con el cactus *Maihueniopsis camachoi* en la adaptación de varias especies a las extremas condiciones ambientales. Los resultdos permitieron comprender mejor los procesos de facilitación y descubrir un novedoso conjunto de metabolitos capaces de predecir el estado de la interacción. Finalmente, este estudio aporta información importante para comprender los mecanismos de adaptación que subyacen a la resistencia de las plantas a los climas extremos, y nuestro enfoque multiespecífico presagia estudios y descubrimientos prometedores en agronomía y ecología.

RÉSUMÉ DE LA THÈSE

Les terres extrêmes situées à la limite d'au moins un gradient abiotique permettent la survie de très peu d'espèces. Ces espèces dites extrêmophiles (littéralement, aimant "philos" les extrêmes) abritent un réservoir unique d'adaptations génétiques et biochimiques qui a toujours attiré la curiosité de l'homme. Des études antérieures ont montré un haut degré de spécificité à l'espèce pour l'adaptation des plantes aux écosystèmes hostiles, ce qui explique que les transferts réussis de mécanismes de résistance vers les cultures agronomiques restent rares. Cependant, des stratégies adaptatives génériques pourraient également exister. Dans ce contexte, je propose de mener une approche compréhensive, de l'écosystème aux métabolites, afin d'étudier les ajustements biochimiques des espèces végétales extrêmophiles du désert d'Atacama, le désert non polaire le plus sec de la planète. Les plantes ont été collectées dans leur environnement naturel qui s'étend sur un gradient d'altitude de 2500 à 4500m. De multiples approches métabolomiques ont été combinées avec le "machine learning" pour dévoiler une boîte à outils générique prédisant la résilience des plantes aux conditions environnementales difficiles. Par la suite, des analyses d'enrichissement des réactions et des voies métaboliques ont permis d'identifier des héritages génétiques gouvernant des stratégies biochimiques convergentes sélectionnées au cours de l'évolution. Enfin, le rôle des intéractions positives avec le cactus *Maihueniopsis camachoi* dans l'adaptation de diverses espèces végétales aux milieux inhospitaliers a été exploré. Les résultats ont permis une meilleure compréhension du processus de facilitation et la découverte d'un ensemble intrigant de métabolites capables de prédire le statut d'intéraction. Dans l'ensemble, cette étude a permis de mieux comprendre les mécanismes d'adaptation qui sous-tendent la résilience des plantes aux climats extrêmes. Par ailleurs, notre approche multi-espèces représente une nouvelle stratégie analytique qui ouvre la voie à des études et des découvertes prometteuses en agronomie et en écologie.

TABLE OF CONTENTS

ABBREVIATIONS

ANOVA: Analysis of variance

***A. thaliana*:** *Arabidopsis thaliana*

BABA: β-aminobutyric acid (BABA)

BEX: Extraction blank

BSA: Bovine serum albumin

C: Carbon

DMAPP: Dimethylallyl diphosphate

ESI: Electrospray ionisation

FAIR: Findable, accessible, interoperable, reusable

FAMEs: Fatty acid methyl esters

FBMN: Feature-based molecular network

GABA: γ-aminobutyric acid

GCFID: Gas chromatography-flame ionisation detector

GCMS: Gas chromatography-mass spectrometry

GLM: Generalised linear models

gLM: General linear models

GNPS: Global natural product social molecular networking

GOT: Glutamate-oxoglutarate aminotransferase

GS-GOGAT: Glutamine and glutamate synthase cycle

HCA: Hierarchical cluster analysis

HK: Hexokinase

HMDB: Human metabolome database

HRP: Horseradish peroxidase

INV: Invertase

IPP: Isopentenyl pyrophospate

KEGG: Kyoto Encyclopedia of Genes and Genomes

LCMS: Liquid chromatography-mass spectrometry

LDH: D-lactate dehydrogenase

LTQ-Orbitrap: Linear trap quadrupole Orbitrap

MDH: L-malate dehydrogenase

MEP: Methylerythritol phosphate

MSE: Mean square error

MSI: Metabolomics standards initiative confidence

MVA pathway: Mevalonate pathway

m/z: Mass on charge ratio

NAD(P): Nicotinamide adenine dinucleotide (Phosphate)

NMR: Nuclear magnetic resonance

N-related compounds: Nitrogen-related compounds

OD: Optical density

OPLS: Orthogonal partial least squares

OPLS-DA: Orthogonal partial least squares discriminant analyses

oxPPP: Oxidative pentose phosphate pathway

PA: Polyamines

PCA: Principal component analysis

PC regression: Principal component regression

PGI: Phosphoglucose isomerase

PLS: Partial least squares

PLSr: Partial least square regression

PLS-DA: Partial least squares discriminant analysis

PVPP: Polyvinylpolypyrrolidone

PVDF: Polyvinylidene fluoride

QACs: Quaternary ammonium compounds

QC: Quality control

Redox: Reduction-oxidation

ROS: Reactive oxygen species

Rpm: Rotations per minute

SLA: Specific leaf area

sPLS: Sparse partial least squares

SS(distances): Sum of the squared distances

T: Temperature

TCA cycle: Tricarboxylic acid cycle

TLT: Talabre-Lejía transect

XCMS: Various forms (X) of chromatography mass spectrometry

GENERAL CONTEXT OF THE PHD

The bedrock of this project is the establishment of a collaboration between the Pontificia Universidad Católica de Chile (PUC, Santiago, Chile) and the University of Bordeaux (UBx, Bordeaux Biologie AgroSciences Master, Bordeaux, France), which has led to the development of a joint program named plant biotechnology program. My experience as the first French student involved in this program allowed me to meet Rodrigo Gutiérrez (head of the Plant Systems Biology lab) during a congress in late 2018. The excellent expertise of his laboratory in transcriptomic and bioinformatic analyses enabled investigating of various fields ranging from the study of how *Arabidopsis (A.) thaliana* senses and responds to nutrients (and especially nitrogen) to the exploration of the transcriptome of pioneer and/or exotic plants. With the help of Prof. Claudio Latorre (Department of ecology, PUC), an ambitious and daring adventure was born almost ten years ago to decipher the plant processes required for adaptation in the Atacama Desert. Several years of fieldwork were necessary to discover and characterise the Talabre-Lejía transect (TLT), which offers unique plant biodiversity in the Atacama Desert along an altitudinal gradient from 2400 to 4500 m. First experiments unveiled genetic clusters relevant for adaptation and suggested an exciting role for specific biochemical pathways. Hence, this PhD project aims to explore the convergent and divergent metabolic strategies employed by Atacama plants to thrive under the extreme conditions of their natural environment. Multiple fields of expertise are required to embark on this multidisciplinary project, which implies diverse analyses and unpredictable results. Thus, this study required the development of a collaboration between the Plant Systems Biology laboratory (R. Gutiérrez, PUC) and the Meta team from Bordeaux Inrae (D. Rolin and P. Pétriacq, UBx) for their expertise and knowledge in plant metabolism and metabolomics. Finally, this project represents the first cotutelle contract between UBx and PUC universities in the field of plant sciences. Besides, this project has also required skills in ecology and evolutionary biology (knowledge shared between R. Gutiérrez and Claudio Latorre from PUC) to preserve the ecological context and better interpret all results. An introductory chapter will first provide the main keys (metabolomics, bioinformatics, plant life in extreme lands) to understand the different analyses. Then, this manuscript describes and discusses how a large spectrum of data covering metabolomics (chapter 1), transcriptomics (chapter 2) and ecology (chapter 3) was generated and integrated to unveil some of the main secrets enabling multiple plant species to thrive in the fascinating Atacama Desert. Notably, this collaboration has already resulted in a review (Chapter 1) (Dussarrat et al., 2021) and a scientific article (Chapter 3, minor revision, New Phytologist, Dussarrat et al., 2022). Additionally, two articles will be submitted in the coming days (Chapter 4) and weeks (Chapter 5).

General context of the PhD

CHAPTER 1

METABOLOMICS OF PLANT SPECIES ADAPTED TO EXTREME ENVIRONMENTS

I. METABOLOMICS IN PLANT SCIENCE

Plants are living organisms capable of converting environmental energy into chemical matter through a miscellaneous series of chemical reactions. The allocation of this chemical matter to developmental or defensive processes depends on the evolutionary trajectory of the organism (*i.e.* genome) and its environment. This dialogue between plant and environment is controlled by a precise orchestration of the different chemical elements in interaction (*e.g.* genes, proteins, small molecules). Biotic or abiotic variations result in a hierarchical response initiated by the plant genome, which ends in an adapted metabolic response after having integrated the regulations of the previous biological levels. In parallel, various regulations such as post-translational modifications can occur without encompassing genome shifts, which complexifies the analysis of the plant response.

These short affirmations result from centuries of research in plant science and have not always been directly accepted by scientists. Hence, this section first seeks to describe the fantastic evolution of both technologies and approaches during the 20th and 21st centuries, which enabled the evolution from a reductionist to an integrative approach. Then, the place of metabolomics in the study of the plant response to the environment is discussed and the complexity of the plant metabolome is addressed. Finally, we present why mathematics earned a central place in modern biological research and how this bioinformatics science allows the discovery of meaningful links between environment, plant biochemistry and plant phenotype.

I.1. From reductionist to integrative biology

Life on Earth, as we perceive it, is the interpretable picture of a constant chemical flow between environmental resources and living organisms. Plants are a complex chemical system that results from the interaction of chemical elements capable of aggregating to create molecules, cells, and finally observable physical matter like organs and tissues. Notably, the fascinating relation between plant phenotype (*i.e.* observable state of the organism or specific traits of this organism) and plant chemistry has always attracted human curiosity. Pioneer scientists started to explore the different organisation levels of the living scale from the entire organism to cells and molecules until the discovery of “something” that determines the physical properties of an organism (Johannsen, 2014). The words “gene” and “genotype” then replaced the “something” and were linked with the phenotype (Roll-Hansen, 2014). The foundation of biochemical genetics arose from the “one gene-one enzyme” theory developed by George W. Beadle and Edward L. Tatum, where a specific mutation led to the production of one enzyme that allowed adaptation to the new conditions (Beadle and Tatum, 1941). Besides, the newly discovered relationship between genes, DNA, mRNAs and proteins greatly enhanced the scientific interest in looking for specific gene-to-phenotype relationships (Brenner et al., 1961; Jacob

and Monod, 1961). During the 20th and 21st centuries, tremendous advances in genetics and sequencing technologies have made possible the analysis at the genome-scale (*i.e.* the analysis on the total amount of genes referred to as gene-"omics") (Holtorf et al., 2002). Model plants such as *Arabidopsis thaliana* were sequenced (The Arabidopsis Genome Initiative, 2000) and used to speed up the analyses and demonstrate a plethora of functional pairs (*i.e.* association between one gene and a phenotypic trait) (Boyes et al., 2001). Hence, the beginning of the 21st century is a prolific period where multiple descriptions of molecular mechanisms have greatly improved our understanding of the plant system. However, numerous studies have led to real breakthroughs, such as the discovery that a gene can generate multiple mRNA and proteins, or conversely, not encode for an enzyme (Stark, 1977; Nawa et al., 1984). Together with the fact that multiple proteins were found to have diverse functions (Jeffery, 1999), these discoveries first highlighted the weaknesses of the "one gene-one enzyme-one function" theory. In parallel, breeding programs have nicely used these significant discoveries to develop highly productive varieties, but subsequently faced two already known phenomena: inter-species variability and phenotypic plasticity (Aubin-Horth and Renn, 2009). Phenotypic trait variations can be observed based on the interaction between genotype and environment (a phenomenon called phenotypic plasticity), suggesting the impact of other molecular mechanisms (Bradshaw, 1965). At the extreme level, the same genotype can lead to dramatically different phenotypes, such as the caterpillar to butterfly transformation. Altogether, the rupture of the oversimplistic view of "one gene-one enzyme", the intra- and inter-species variability in gene function, and the phenotypic plasticity greatly contributed to the conceptual scientific movement towards the integrative approach (Aubin-Horth and Renn, 2009). Hence, integrative and reductionist approaches are complementary techniques that enabled significant scientific advances.

The use of new technologies subsequently allowed the breakdown of cellular chemical constituents between genotype and phenotype, from DNA to mRNA and proteins (Fig. I.1). Interestingly, the combination of these omics strategies allowed for comprehensive analyses of gene function, moving from the misconception of one gene-one phenotype to a pyramidal concept (Fig. I.1) (Holtorf et al., 2002). While the cumulative utilisation of these methods allowed to attribute gene function on a scale previously unimaginable, multiple contradictions were confirmed. For instance, gene expression levels were not always correlated to protein abundance (Vélez-Bermúdez and Schmidt, 2014). Besides, the integrative scale of this pyramid is associated with an increase in complexity. While genotype depends on 4 nucleobases, the theoric number of possibilities proteomics is defined by the combination of 21 proteinogenic amino acids in eukaryotes. However, although more tedious, integrated analysis can elucidate the chemical mechanisms underlying a given phenotypic variation (Fig. I.1) (Schmitt, 2003). For instance, proteome analysis allows the comprehension of what causes the phenotypic change by integrating not only genomic and transcriptomic levels, but also post-translational modification and protein-protein interactions (Altelaar et al., 2013).

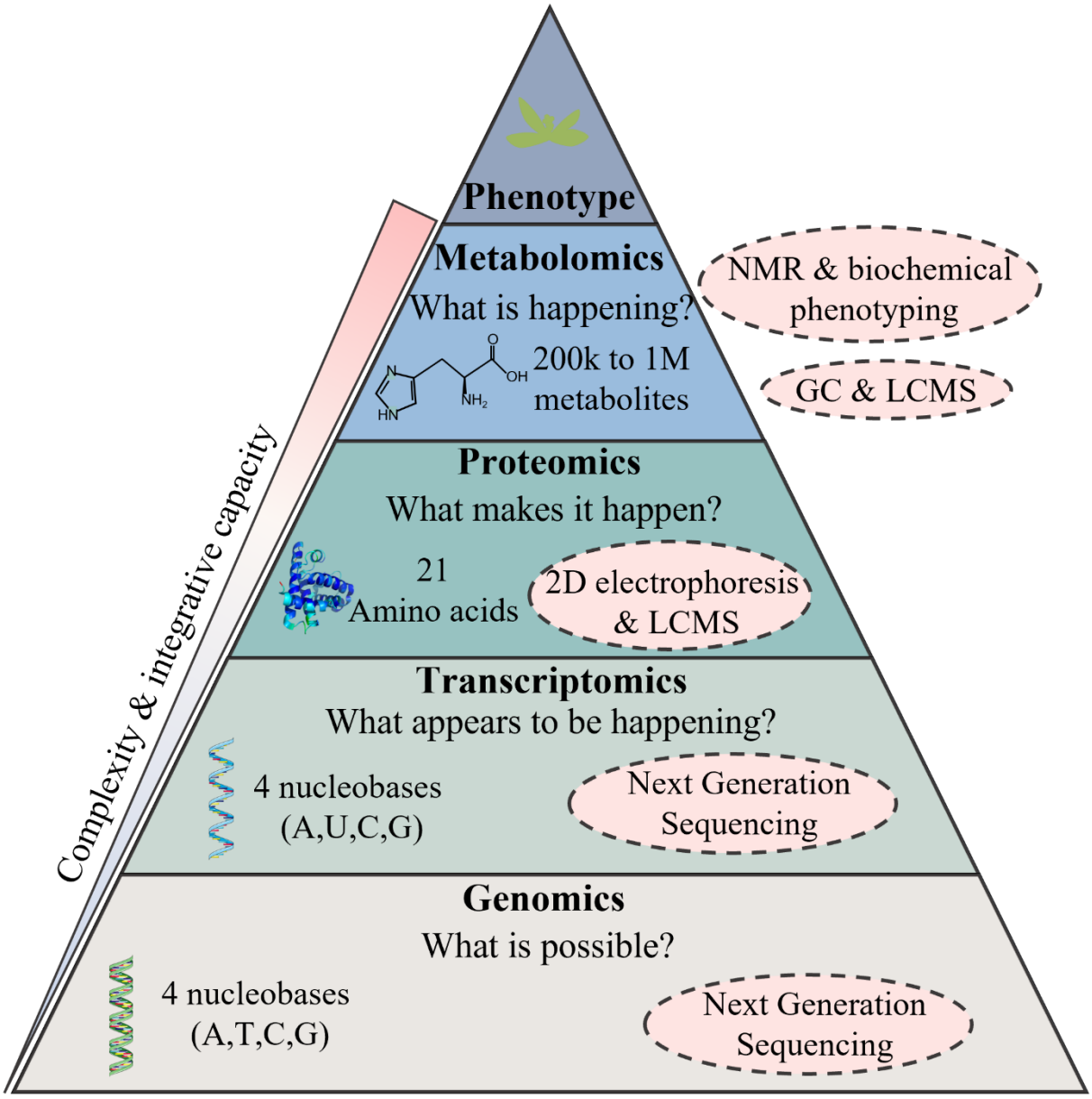

Fig. I.1 | The pyramidal concept of omics sciences. *GC/LCMS*: gaz or liquid chromatography coupled to mass spectrometry. The DNA, RNA and protein images were downloaded from Smart Server Medical Art.

Even more entertaining, technological advances allowed scientists to move one step further and access the direct analysis of the matter composition (Sumner et al., 2003). The analysis of the plant biochemical diversity was then considered as the most integrative technology by being the closest omics to the phenotype (Rochfort, 2005). Hence, these small molecules called metabolites create a biological link between genotype and phenotype, giving a glimpse of the myriad of opportunities offered by this science (Fig. I.1). Consequently, the large-scale study of the metabolites at any organisation level in a biological matrix, so-called metabolomics, has earned a central place in plant science since its introduction between 2 and 3 decades ago (Alseekh and Fernie, 2018).

I.2. The complexity of plant metabolism

The metabolome represents the entire set of endogenous and exogenous (xenobiotics) molecules harboured by an organism. Endogenous metabolites are typically divided into central, primary molecules (where central and primary metabolism are commonly grouped and referred to as primary metabolism), secondary compounds and regulators, which technically could be either primary or secondary metabolites (Erb and Kliebenstein, 2020). The first group involves highly conserved metabolites, which are shared between plant and animal species, and required for growth, survival and reproduction. Central pathways represent the backbone of plant metabolism and are responsible for the uptake and process of environmental resources. Besides, central compounds like hexoses from glycolysis or organic acids from the tricarboxylic acid cycle (TCA) fulfil the synthesis of other compounds. Primary metabolites are therefore compounds directly related compounds to central metabolism, including amino acids and lipids, for instance. Subsequently, primary pathways lead to the production of vital compounds that serve as building blocks for the synthesis of a miscellaneous series of secondary metabolites allowing plants to cope with their environment (Fig. I.2). Interestingly, whilst primary compounds have been extensively characterised and analysed during the last decades, knowledge of the chemical diversity and biological functions of the secondary metabolism remains partial. The following section describes the main classes of plant metabolism as well as the major analytical methods to study them.

Central metabolism

Major physiological functions like the uptake and process of environmental resources are managed by the glycolysis, TCA cycle, and oxidative pentose phosphate pathways (oxPPP) (Plaxton, 1996; Fernie et al., 2004; Araújo et al., 2012) considered as central metabolism. These pathways are composed of a succession of chemical reactions involving hexoses, hexose or pentose phosphates (glycolysis and oxPPP), organic acids (TCA cycle), and some amino acids (*i.e.* anaplerotic pathway). Energy (ATP), carbon (C) blocks for biosynthesis of various compounds (*e.g.* pyruvate, 2-oxoglutarate) and reducing power (NADH, NADPH) are generated via the hexose oxidations, the TCA cycle, and the oxPPP. Also, central metabolism governs plant development via its interactions with the photosynthetic and photorespiratory processes, for instance (Plaxton, 1996; Fernie et al., 2004; Araújo et al., 2012). Besides, the central place of these pathways in managing C allocation between growth and defence is pinpointed by the synergy observed between primary and secondary metabolisms (Fig. I.2). Notably, the transformation of pyruvate to acetyl-CoA (which is considered as a major building block of lipids), the relation between 2-oxoglutarate and amino acid biosynthesis, and the connection between oxPPP and the shikimate acid pathway (at the origin of several secondary compounds), are some examples of these reported metabolic crosstalks (Reid et al., 1977; Kruger and von Schaewen, 2003; Oliver et al., 2009; Araújo et al., 2012).

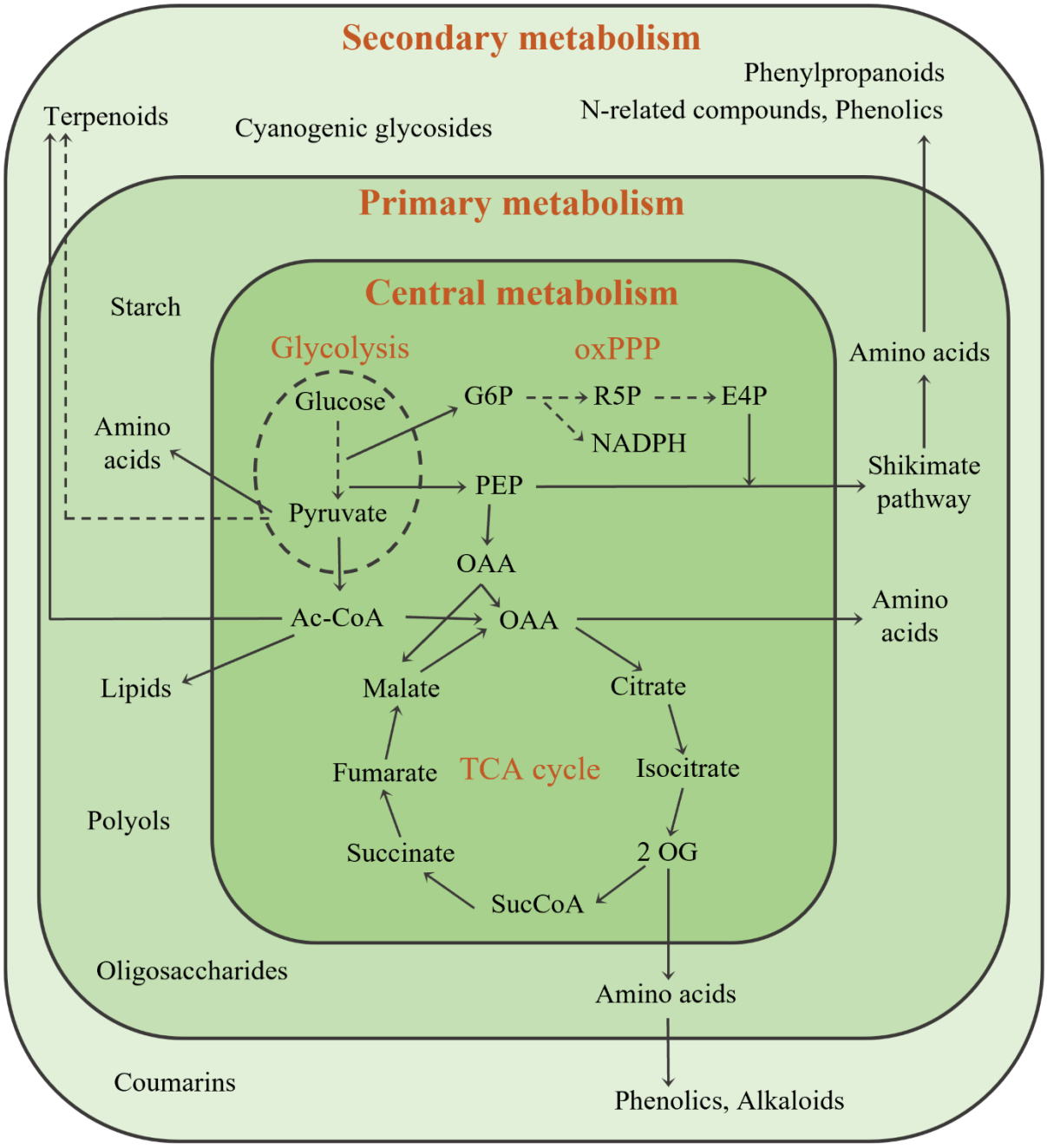

Fig. I.2 | Simplified scheme illustrating the complexity of plant metabolome and presenting some of the main crosstalks between central, primary and secondary compounds. *Ac-CoA*: acetyl-CoA, *E4P*: erythrose 4 phosphate, *G6P*: glucose 6 phosphate, *NADPH*: nicotinamide adenine dinucleotide phosphate, *OAA*: oxaloacetic acid, *oxPPP*: oxidative pentose phosphate pathway, *PEP*: phosphoenol pyruvate, *R5P*: ribulose 5 phosphate, *SucCoA*: succinyl-CoA, TCA cycle: tricarboxylic acid cycle, *2 OG*: 2-oxoglutarate.

Primary metabolism

Primary pathways are subjected to strong selective pressure owing to their vital properties for plant life, thus explaining their ubiquity (except for some exceptions) in the plant kingdom. Interestingly, the majority of those molecules are involved in plant growth and plant defence and include multiple polymers such as carbohydrates, lipids and monomers as amino acids. Besides, primary compounds act as a bridge between central and secondary pathways.

- *Compounds related to monosaccharides* -

Additional carbohydrates refer to primary metabolism such as oligosaccharides and polyols (Fig. I.2). Monosaccharides like glucose and fructose can assemble to create compounds known as oligo ("few" in Greek) -saccharides as sucrose (Kandler and Hopf, 1980). Starch is a polymer of sugars osmotically inactive acting as an accurate indicator of plant fitness, especially under abiotic constraints (Thalmann and Santelia, 2017). Additionally, several hundreds of oligosaccharides have been described (*e.g.* raffinose family oligosaccharides) and perform various functions such as energy storage and stress resistance (Kandler and Hopf, 1980). Polyols are reduced forms of sugars involving at least 3 C and bearing OH groups (*e.g.* inositol) that accomplish diverse roles in C storage and transport from source to sink tissues, or plant defence against abiotic stress (Noiraud et al., 2001).

- *Compounds related to fatty acyls* -

Lipids not only define the viability of biological membranes but are also key actors in plant development and defence (Harwood, 1996). Remarkably, the classification of molecules in the "lipid" class is still ambiguous since they are defined as any organic compounds that are insoluble in water and extractable by non-polar organic solvents (Ohlrogge and Browse, 1995). Lipids are divided into 8 classes: fatty acyls (synthesised from acetyl-CoA and malonyl-CoA), glycerolipids (including a glycerol group), glycerophospholipids (including a phosphate group), sphingolipids (including a nitrogen atom in the lipidic chain), saccharolipids (including a sugar attached to the fatty acyl group), polyketides, and sterols or prenol lipids (from dimethylallyl or isopentenyl pyrophosphate) (Fahy et al., 2011) (Fig. I.2).

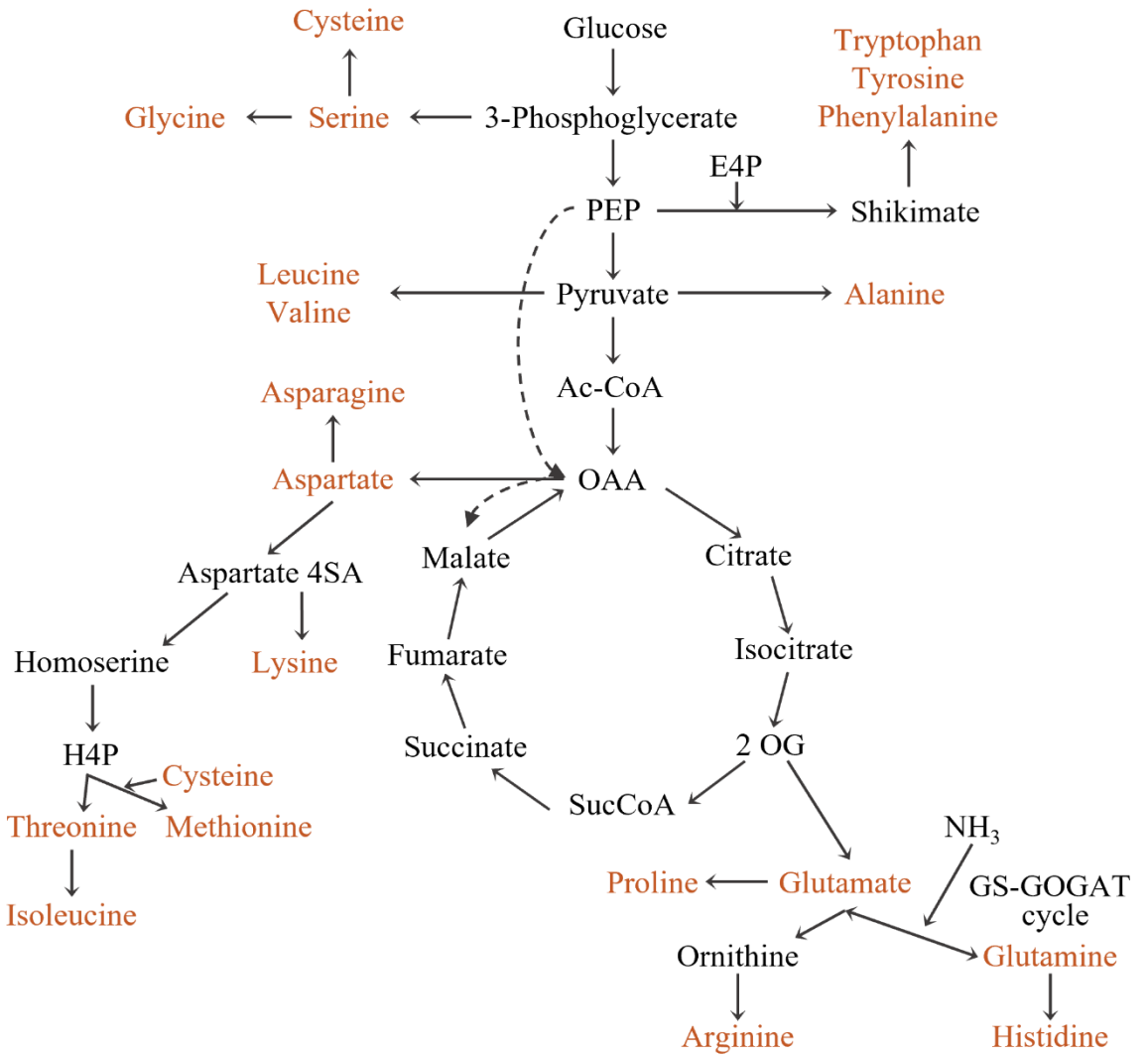

Fig. I.3 | Biosynthesis of amino acids. *Ac-CoA*: acetyl-CoA, *E4P*: erythrose 4 phosphate, *GS-GOGAT*: glutamine synthase-glutamine:2-oxoglutarate amidotransferase, *H4P*: homoserine 4-phosphate, *OAA*: oxaloacetic acid, *PEP*: phosphoenol pyruvate, *SucCoA*: succinyl-CoA, *2 OG*: 2-oxoglutarate. Figure inspired by Forde and Lea., 2007 and Yang et al., 2020.

- *Compounds related to proteins* -

Amino acids serve as constituents of proteins and are thus required for plant growth and survival (Hildebrandt et al., 2015). Among this class of compounds, 20 were defined as products of protein hydrolysis and qualified as "proteinogenic". Since their discovery started in the early 19th century (wonderfully described in (Vickery and Schmidt, 1931)), plenty of amino acids have been highlighted and categorised as non-proteinogenic and classified as secondary compounds (*e.g.* γ-aminobutyric acid (GABA)) (Jander et al., 2020) or novel phytohormone like β-aminobutyric acid (BABA) (Thevenet et al., 2017). Notably, ubiquitous proteinogenic compounds ensure additional functions like nitrogen assimilation and storage in glutamine, arginine, asparagine, glutamate and aspartate (Forde and Lea, 2007). Inorganic nitrogen is incorporated into glutamate and glutamine via the glutamine and glutamate synthase (known as glutamine oxoglutarate aminotransferase) cycle (GS-GOGAT), which thereafter provide aspartate and asparagine (which can also provide from TCA cycle), as well as arginine and proline (Forde and Lea, 2007) (Fig. I.3). Besides, aspartic acid is the precursor of lysine, threonine, methionine, isoleucine and asparagine (Azevedo et al., 2006; Yang et al., 2020) (Fig. I.3). Similarly, sulfate assimilation leads to the synthesis of cysteine, which subsequently generates methionine in coordination with aspartate biosynthesis (Hesse et al., 2004). Whilst 2-oxoglutarate is incorporated into GS-GOGAT cycle, other precursors come from central pathways like pyruvate for the production of branched amino acids (*e.g.* valine, leucine) and alanine (Binder, 2010; Xu et al., 2017). Also, oxPPP acts as a first step for the biosynthesis of i) aromatic amino acids (*i.e.* phenylalanine, tyrosine and tryptophan) via the shikimate pathway, and ii) histidine which is linked to the 5'-phosphoribosyl 1-pyrophosphate derived from ribose-5-phosphate (Rees et al., 2009; Tzin and Galili, 2010). Finally, different biosynthetic pathways were described for the production of glycine and serine with, among them, the 3 steps pathway metabolising the 3-phosphoglycerate into serine (Fig. I.3) (Ros et al., 2014).

Hence, central and primary pathways form the core of the plant metabolic system, ensuring the assimilation of essential chemical compounds (*e.g.* C, N, S) and their distribution according to physiological needs and environmental constraints. Even more fascinating, these compounds serve as basic pieces of a more complex puzzle known as secondary metabolism. Thus, fatty acids and amino acids are, for example, precursors of a plethora of secondary compounds (Fig. I.2) and the source of interactions with phytohormones that are considered regulators (*e.g.* phenolics, jasmonate and ethylene biosynthesis) (Oliver et al., 2009; Schaller and Stintzi, 2009; Häusler et al., 2014; Hasanuzzaman et al., 2018).

Secondary metabolism

The extraordinary diversity of plant secondary metabolism has been estimated between 200 000 and 1 000 000 compounds (Rai et al., 2017). Interestingly, the evolutionary theories of secondary metabolism may explain this functional diversity and redundancy. Secondary metabolites would have a dual origin from adaptation (*i.e.* positive selection) or non-selective processes (*e.g.* demographic history), leading joint and specific functions through evolution (Scossa and Fernie, 2020). Gene duplication (from polyploidisation, transposon activity) and gene fusion (*i.e.* fusion of independent cistrons which provide multifunctional proteins) are the main driving forces of secondary metabolism evolution. Duplicated genes (also called paralogous genes) then form gene families where the original gene maintains its function, while other copies acquire new roles based on, for instance, a shift in substrate preference or a change in expression pattern in time and space (Fondi et al., 2009). The successive combination of gene duplication and neo-functionalisation is the cornerstone of the patchwork hypothesis. The patchwork hypothesis is based on the fact that primitive enzymes were able to inefficiently react with various substrates, therefore leading to a myriad of potential reactions (Scossa and Fernie, 2020). Primitive enzymes were then recruited and specialised through gene duplication and neo-functionalisation to accomplish novel functions in emergent pathways (Caetano-Anollés et al., 2009; Scossa and Fernie, 2020). Overall, evolution led to a highly diverse metabolism composed of generic compounds (ubiquitous between plant families and at the basis of chemical pathways) and specific pathways that occurred in relation to environmental pressure and evolutionary trajectory. The immensity of this metabolic world led to complex classification systems based on the structure or the function of the molecules, for instance (Fig. I.4).

- *N-related compounds* -

N-related compounds cover an extremely high diversity of metabolites, among which alkaloids, polyamines and quaternary ammonium compounds play determinant roles in environmental stress mitigation and defence against biotic threats.

- Alkaloids represent a very large group of compounds divided into three major divisions based on their structures: i) *stricto sensu* alkaloids (heterocyclics), which contain an intracyclic nitrogen and derive from amino acids, ii) proto-alkaloids (non-heterocyclics), synthesised from amino acids (*e.g.* tryptophan and phenylalanine) and including a nitrogen atom outside of the ring, and iii) pseudo-alkaloids that are not derived from amino acids (Gutiérrez-Grijalva et al., 2020).

- Polyamines (PA) are low molecular products composed of two or more amino groups (Bachrach, 2010). While different biosynthesis pathways have been revealed, the amino acids ornithine and arginine (produced from glutamate) remain the major precursors. Interestingly, multiple roles have been attributed to the diamine “putrescine”, the triamine “spermidine” and the tetramine “spermines” in both

plant development and abiotic stress tolerance, but precise mechanisms of action are still poorly described ((Slocum, 2005; Bachrach, 2010), section II).

- Quaternary ammonium compounds (QACs) have been extensively characterised since highly resistant plants accumulate these compounds ((Storey et al., 1977), section II). These compounds are defined based on their constantly positively charged methylated nitrogen atom (Rhodes and Hanson, 1993) (Fig. I.4). Notably, the biosynthesis of these metabolites are diverse and sometimes poorly documented. Besides, betaines like proline and glycine betaine are nitrogenous osmolytes that accumulate in plants under stressful conditions but remain specific to several plant families (Trinchant et al., 2004).

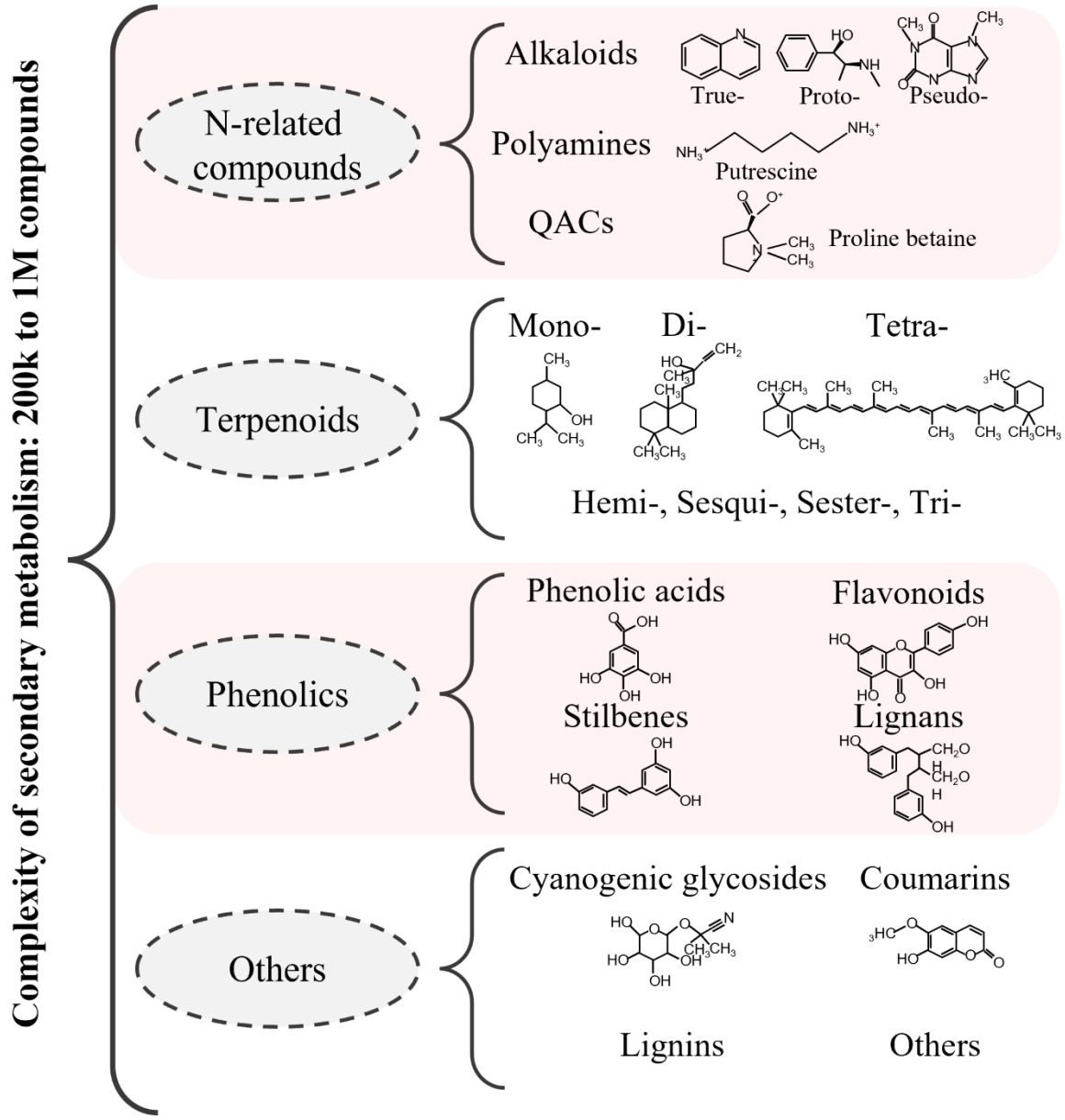

Fig. I.4 | Simplified scheme of the major classes of secondary metabolites.

- *Terpenoids* -

Terpenoids, or isoprenoids, have been massively studied for their great interest in the pharmaceutical and agronomic industries and represent one of the main classes of secondary metabolites in terms of biodiversity (Tetali, 2019) (Fig. I.4). Besides, terpenes are closely related to plenty of compounds required for plant development and defence (Fig. I.5). The precursors of all terpenoids are the five-carbon unit called isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) which are produced by two distinct pathways taking place in the plastid (the methylerythritol phosphate pathway named MEP) or cytoplasm (the mevalonate pathway named MVA) (Fig. I.5) (Nagegowda and Gupta, 2020). Thus, the subclasses are defined based on the number of isoprene(s) as follows: hemi- (1 isoprene), mono- (2), sesqui- (3), di- (4), sester- (5), tri-(6), and tetra- (8) -terpenoids. MVA starts with the conversion of acetyl-CoA to IPP and DMAPP which fulfil the C reserve for sesqui- and triterpenes production in the cytoplasm. In parallel, pyruvate and glyceraldehyde-3-phosphate undergo multiple reactions until biosynthesis of hemi-, mono-, sesqui-, di- and tetra-terpenes via the MEP in plastids (Tholl, 2015). Interestingly, the structural richness of isoprenoids is illustrated by the great diversity of chemical properties covered by these compounds (Nagegowda and Gupta, 2020). Diterpenes include compounds of primary interest such as phytol (vitamin precursor), whereas triterpenoids and tetraterpenes comprise phytosterols and carotenoids, respectively. Even more interesting, phytohormones such as gibberellic acids and abscisic acid as members of the terpene chemical family and are referred to as either secondary or regulator compounds (Tholl, 2015). Finally, other major compounds come down from the MVA pathway as brassinosteroids. Besides, the MEV pathway leads to the formation of strigolactones and cytokinins (Nagegowda and Gupta, 2020) (Fig. I.5).

- *Phenolics* -

Phenolics are another major group of secondary metabolites classified according to their chemical structure and biochemical origin (Lattanzio, 2013) (Fig. I.4). Pentose phosphate and glycolysis pathways provide erythrose 4-phosphate and phosphoenol pyruvate as building blocks for polyphenol biosynthesis (Naikoo et al., 2019) (Fig. I.6). These primary compounds are then used through the shikimate pathway to generate shikimic acid, which is then metabolised through the phenylpropanoid pathway to produce major aromatic amino acids like tyrosine, tryptophan and phenylalanine (Fig. I.6). These amino acids then undergo subsequent chemical reactions to synthesise the majority of phenolic compounds, which can however derive from the malonate acetate pathway in plants (Maeda and Dudareva, 2012). Like terpenes, phenolics encompass a wide range of compounds from simple aromatic rings to complex molecules. The five main classes, namely phenolic acids, stilbenes, flavonoids, lignans and others, have demonstrated pleiotropic roles in plant growth and resistances (López-Fernández et al., 2020). Besides, their role in mitigating oxidative stress has been widely described (Decros et al., 2019).

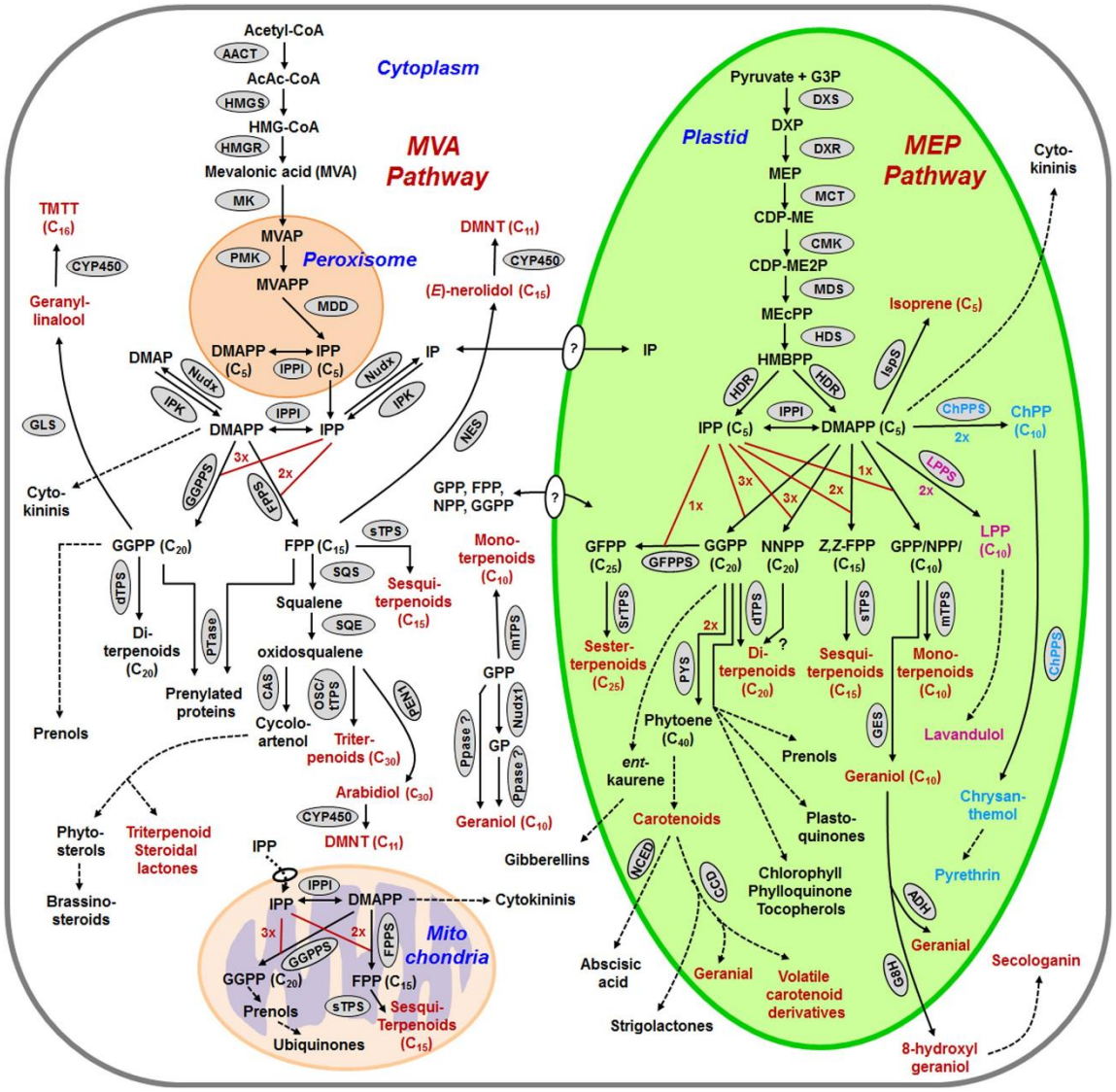

Fig. I.5 | Terpenoids and related compounds. Terpenoids biosynthesis, the figure was developed by D. Nagegowda and P. Gupta 2020**.**

- Phenolic acids are produced through the phenylpropanoid pathway as other phenylpropanoid compounds as flavonoids (Mandal et al., 2010). However, phenolic acids can also derive from the monolignol pathway or the breakdown of lignin compounds (Mandal et al., 2010). Notably, phenolic acids cover C6-C1 compounds like gallic acid, C6-C2 metabolites, and C6-C3 compounds like p-coumaric acid (also called 4-hydroxycinnamic acid), which acts as a precursor for the biosynthesis of other compounds (*e.g.* hydroxycoumarins) (Lattanzio, 2013).

- The phenylpropanoid pathway fulfils the 4-coumaroyl-CoA reserves used for flavonoid biosynthesis. This pathway is thus at the origin of a miscellaneous series of flavonoids, which are divided into multiple subclasses based on the chemical linkages between the classic C6-C3-C6 (Fig. I.6) (Lattanzio, 2013). Interestingly, some of these subclasses encompass a myriad of compounds of interest for their roles in plant defence as quercetin and kaempferol in flavonols, tannins in flavanols and cyanidin in anthocyanins (Treutter, 2006).

- Stilbenes are C6-C2-C6 phenolics massively studied for their benefits in human health (Chong et al., 2009). Their production through the phenylpropanoid pathway is based on the synthesis of the widely studied resveratrol from 4-coumaroyl-CoA (Fig. I.6).

- Lignans are another class of phenolics defined as phenylpropanoid dimers linked by the C8 carbon and produced via the phenylpropanoid pathway.

- Other classes of phenolics include chemicals derived from the phenylpropanoid pathway such as lignins (heteropolymer of three monolignols called p-coumaryl, coniferyl and sinapyl alcohols), cyanogenic glycosides (synthesised from amino acids and involved in plant defence), and coumarins (derived from phenolic acids), for instance (Vetter, 2000; Boerjan et al., 2003; Sarker and Nahar, 2017) (Fig. I.6).

The biological functions of secondary metabolites are as diverse as their chemical structures. Whilst these compounds were initially considered as by-products of primary pathways that do not participate in the development (Sachs, 1874), it is now clear that these compounds fulfil a plethora of roles (*e.g.* plant-soil or plant-insect interactions, reproduction, plant defence…) (Naikoo et al., 2019). However, common roles and properties emerged from multiple classes of chemicals like the antioxidant capacity of several terpenes and phenolics, therefore contributing to reduction-oxidation processes (Naikoo et al., 2019).

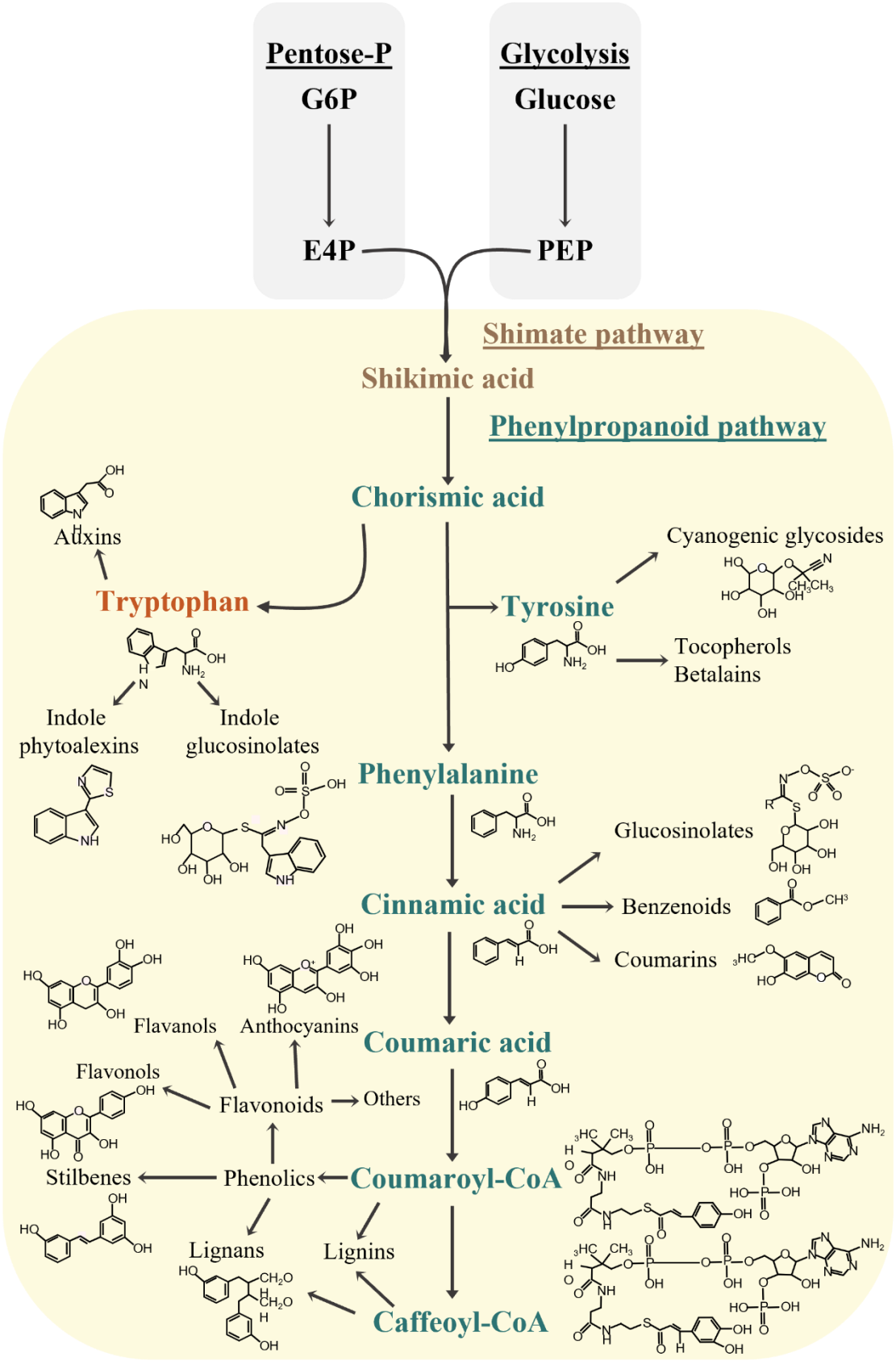

Fig. I.6 | Biosynthesis of polyphenols from shikimate pathway. *G6P*: glucose 6 phosphate, *E4P*: erythrose 4 phosphate, *PEP*: phosphoenol pyruvate.

Figure adapted from H. Maeda and N. Dudareva 2012.

Redox compounds

Plants are aerobic organisms and therefore require oxygen to survive. Paradoxically, the use of oxygen produces reactive oxygen species (reduced forms of oxygen (ROS) like the hydrogen peroxide, the superoxide anion and the hydroxyl radical), mostly through photosynthesis and mitochondrial respiration (Decros et al., 2019). Ascorbate (also called vitamin C in Humans) and glutathione are at the forefront of plant defence against oxidative damages. The main ascorbate biosynthesis pathway (called the D-mannose/L-galactose pathway) uses the D-glucose as a precursor and nicely illustrates the patchwork evolution theory since the first steps are common to cell wall precursors while the last steps are specific to ascorbate biosynthesis (Ishikawa et al., 2006). Glutathione is a thiol (sulfur-containing compound) produced via the combination of glutamate, cysteine and glycine (Noctor et al., 2012). Finally, the synthesis of pyridine nucleotides is derived from aspartate, which is subjected to multiple reactions in chloroplast and cytosol to finally generate nicotinamide adenine dinucleotide (NAD). NAD is then the precursor of the other forms of pyridine nucleotides called NAD phosphate (NADP) and their reduced forms (NADH and NADPH) (Gakière et al., 2018). Fascinatingly, evolution led to a vital cycle called Foyer-Asada-Halliwill cycle (*i.e.* ascorbate-glutathione pathway) where ascorbate and glutathione act as the two main ROS processors and maintain their redox state using the reduction power of pyridine nucleotides (*i.e.* NADPH and NADH). Additionally, non-enzymatic processing is performed via multiple secondary compounds like carotenoids and flavonoids (Decros et al., 2019). More precisely, secondary compounds including phenolics and terpenoids act as a significant antioxidant system processing excess of ROS and thus contribute to the redox balance (Decros et al., 2019). Overall, the rapid evolution of our understanding of redox processes is based on an elegant integrative approach where scientists have overcome the reductionist concepts “one gene-one enzyme-one function” or “primary and secondary metabolism”. For instance, this redox process first characterised for the generation of the toxic ROS responsible for cellular damages is now considered a major player in (i) physiological development and (ii) response to biotic or abiotic stress ((Das et al., 2015; Decros et al., 2019), see section II).

Altogether, the fascinating diversity of plant metabolism and the unimaginable amount of interactions between primary and secondary compounds perfectly demonstrates the need to evolve from a reductionist to an integrative approach (Fang et al., 2019). The redundance of their chemical structures and roles in plant defence or development, as well as their ubiquitous or specialised nature, represent a great challenge in metabolomics. The development of technical facilities now enables for precise *(i.e.* quantitative) or in-depth analysis of the chemical diversity, therefore providing a massive amount of data that have to be precautionarily handled to avoid misinterpretation. Hence, the development of adapted bioinformatic tools to manage plant chemical diversity is the *sine qua non* condition to extend the successes of the integrative approach.

Strategies for metabolite detection and quantification

The chemical composition (*i.e.* the nature and number of each chemical element) and the organisation of these elements in space define the chemical and physical properties of each molecule. All analytical techniques are using some of these properties (*e.g.* mass, absorbance, charge, size, volatility, spin) to detect and quantify the different metabolites. Besides, among the major analytical tools in metabolomics, biochemical phenotyping is the technique that requires the most biological details (description of chemical pathways, enzymes..) since it relies on the use of enzymes and cofactors for the quantification of different target metabolites (Gibon et al., 2004). This technique can then be combined with robotics to allow for a high-throughput dimension of the analysis (Gibon et al., 2004). Conversely, other techniques like nuclear magnetic resonance (NMR) or mass spectrometry (MS) solely rely on physical properties such as nuclear spin or the mass of the molecules (Zhang et al., 2012).

However, plant metabolomics deals with complex mixtures, including a wide range of molecules from small polar compounds to large hydrophobic metabolites. Notably, no analytical tool is currently capable of detecting such a broad spectrum, therefore making each metabolomic experiment an incomplete analysis. Interestingly, two major approaches are used to cope with this wonderful chemical diversity. First, each metabolic analysis must select the type of metabolites analysed via the choice of the extraction technique. For instance, while water extraction favours the extraction of polar compounds, fatty acyls are preferentially extracted using highly organic solvents like hexane (Domergue et al., 2010). Ethanolic or methanolic extractions remain the most widely used extraction methods in metabolomics studies due to their capacity to capture a relevant diversity of metabolites (mostly semi-polar) (Luna et al., 2020). Second, chromatography techniques are combined with mass spectrometry to decomplexify this biological matrix before MS analysis and therefore maximises the potential coverage. The nature of the column used in liquid or gas chromatography coupled to mass spectrometry (GC and LCMS) thus governs the type of detected molecules. For instance, C18 columns are capable of retaining metabolites from the mobile phase based on their hydrophobicity, leading to a sequenced arrival of semi-polar compounds (Zhang et al., 2012).

Finally, analytical techniques are defined by their sensitivity (Tian et al., 2016). For instance, whilst biochemical assays are limited to a few well-known metabolites, NMR technique enables the detection of major polar compounds (detection limit of around 150 µM) and LC or GCMS permit access to biochemical diversity (pM to fM) (Sumner et al., 2003). However, while the sensitivity of LCMS techniques explains its massive use in metabolomic analyses, this property rapidly becomes a handicap when it comes to annotating and quantifying metabolites. Consequently, bioinformatics tools for handling, processing and analysing the data vary depending on the analytical facilities employed and can range from comparing the intensity of a few known metabolites to integrating thousands of unknown chemical compounds within models and metabolic networks.

I.3. Systems biology

"Mathematics is biology's next microscope" (Cohen, 2004). The development of omics sciences, and especially metabolomics, has led to a sharp increase in the amount of data generated per experiment. The bidirectional evolution of metabolomics has fostered the emergence of targeted and untargeted approaches that can be combined or performed independently to investigate the molecular features and mechanisms underlying phenotypic trait variations (Gorrochategui et al., 2016). Besides, the arrival of diverse data sets has rapidly required the development of computational tools to meet the new analytical challenges. Analytical workflows are generally divided into a first explanatory phase, which seeks to describe the metabolic dataset and establish new hypotheses, and a correlation statistical analysis, which varies between analyses (Hendriks et al., 2011). Interestingly, one of the most critical challenges is to provide meaningful biological advances while using noisy, heterogeneous and collinear datasets. Modelling approaches have received particular attention in recent years due to their ability to extract the explanatory variables underlying the variation of phenotypic traits while preserving the biological context. Hence, mathematical (Belouah et al., 2019), stoichiometric (reaction-based) and kinetic (enzymatic-based) (Beauvoit et al., 2018), as well as predictive (correlation-based) (Nelder and Wedderburn, 1972; Luna et al., 2020) models sit at the top of multivariate analyses in the topic of plant metabolomics.

Bottom-up and top-down approaches towards the understanding of phenotypic variations

Depending on the analytical technique, metabolomics offers various options from the detailed analysis of precise chemical compounds and pathways to the production of a metabolic fingerprint. The first analysis can provide a quantitative analysis (via biochemical assays or the use of standards injected in NMR or LCMS) of a small set of known metabolic compounds from primary metabolism or some well-described secondary pathways (Beauvoit et al., 2018). This targeted approach can be used either in combination with untargeted analysis or on its own. When used alone, the targeted analyses are the backbone of "bottom-up" approach, which starts from studying key molecular compounds to reconstruct a chemical pathway and describe its role in modulating a phenotypic trait of interest (Fig. I.7). Thus, the objectives can be either to analyse the response of a metabolic pathway to environmental perturbations or to characterise a known metabolic pathway in a new organism or plant tissue. The second approach maximises the detection of primary and secondary compounds that may be known or unknown (Luna et al., 2020). While this untargeted approach opens the gates to the unknown, this study is strongly limited by the annotation process and its inability to quantify unknown compounds. In return, the untargeted analysis offers the unique opportunity of working in a "top-down" approach, where potential markers or predictors of phenotypic traits are unveiled from a global set of compounds via machine learning (Fig. I.7). Both approaches underpin distinct workflows and statistical analyses. Hence, the first challenge in systems biology is the definition of the biological question. For instance, specific questions

like how redox fluxes influence plant defence against high light intensity will (mainly) lead to targeted analyses of redox compounds with various related hypotheses. In contrast, questions like "what are the metabolic features of extreme plants" will drive the metabolic analysis towards a "without *a priori*" workflow, which sometimes does not necessarily require an initial hypothesis. Then, the design of the biological and technical workflow should include the biological context of the study and define the technical approach. Finally, once the experimental work is performed, a plethora of statistical tools can be used to extract relevant metabolites and provide new insights into the problem studied (Hendriks et al., 2011). Since the question of this PhD has been tackled using a "without *a priori*" approach, statistical analysis of the bottom-up strategy will only be described briefly.

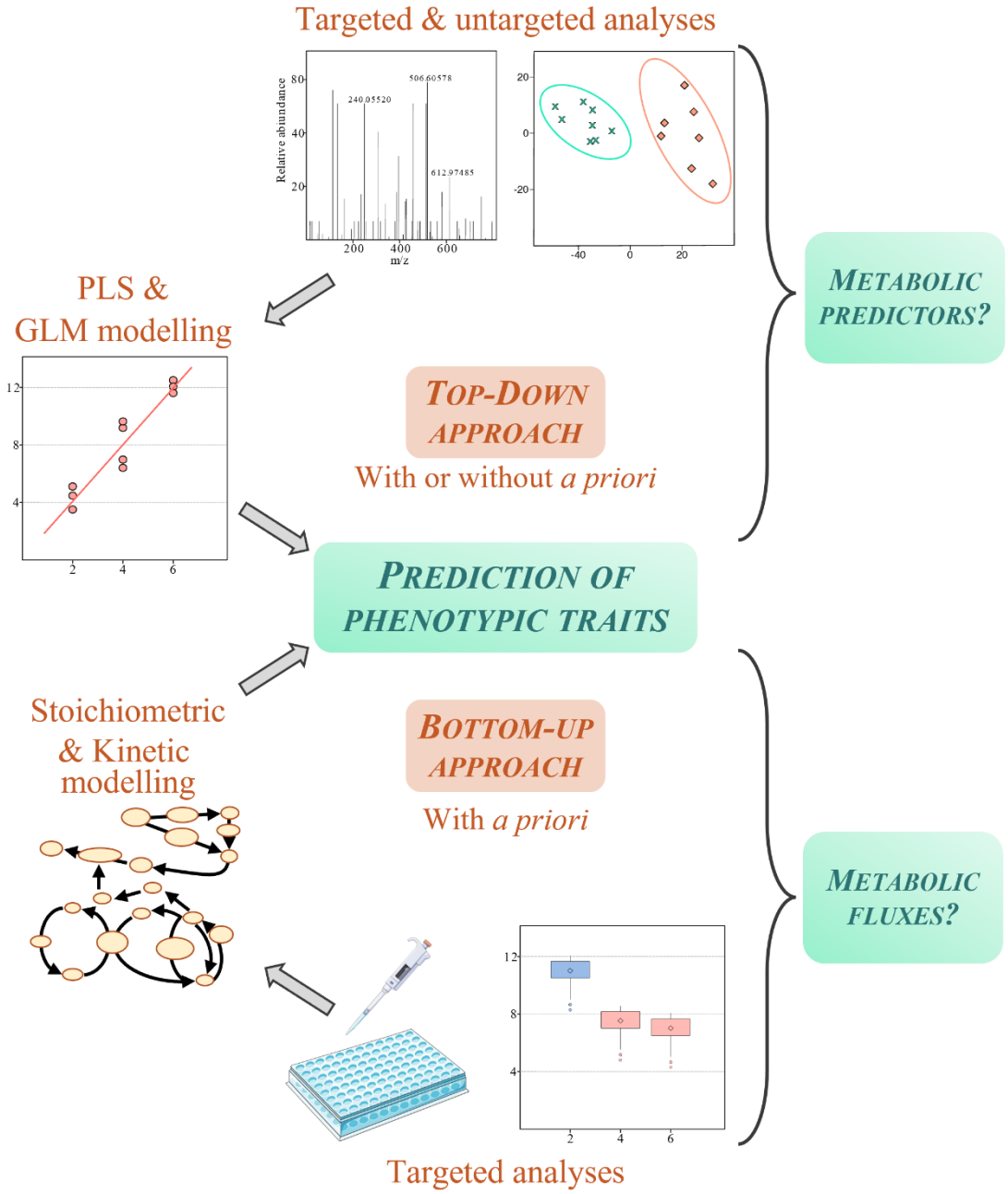

Fig. I.7 | Bottom-up and top down approaches in metabolomics. Pipette and microplate images were obtained on Smart Server Medical Art.

Univariate methods

Univariate statistics are generally used to test the significance of specific metabolic variables. Following data normalisation, univariate analyses are thus performed to screen the metabolic dataset and extract the informative compounds from the uninformative ones via t-test or analysis of variances (ANOVA) (Saccenti et al., 2014). Subsequently, correction tests for false discovery rates are generally performed since the probability of revealing false positives increases with the number of univariate tests performed (hundreds to thousands when focusing on untargeted analyses) (Benjamini and Hochberg, 1995). Hence, univariate analyses can be deployed to reduce the complexity of the metabolic dataset while preserving "informative" compounds related to the phenotypic trait under investigation. Notably, the use of univariate statistics to clean up the metabolic dataset before applying discriminant or regression analyses is quite valuable to limit the noise effect.

Exploratory statistics

Untargeted analyses provide complex multivariate datasets characterised by high chemical diversity, high noise and collinearity levels. Consequently, the first essential step into statistical analysis is data normalisation and scaling, which allows for a multivariate comparison between samples and variables and transform the dataset so statistical assumptions for univariate tests are fulfilled. Whilst multiple normalisations exist, median normalisation and Pareto scaling (*i.e.* each variable being mean-centred and divided by the square root of its standard deviation, which reduces noise influence compared to autoscaling) remains the most widely used method in untargeted metabolic datasets (Di Guida et al., 2016). However, other data scaling might be more suitable when analysing a small number of chemical variables like auto-scaling (*i.e.* each variable being mean-centred and divided by its standard deviation) (van den Berg et al., 2006; Di Guida et al., 2016).

Exploratory methods encompass two main unsupervised multivariate statistical methods: 1) principal component analysis (PCA), which highlights the global trends of omic impacts as well as quality control and potential outliers, and 2) hierarchical clustering, which explores the statistical relationships between samples and/or variables based on multivariate data (Fig. I.8). Briefly, PCA decomplexifies a highly dimensional dataset into a few uncorrelated (orthogonal) vectors called components that maximise the variance between samples (Hendriks et al., 2011). In short, PCA plots the samples into a p-dimensions space (p= number of variables), and defines the principal components by evaluating the sum of the squared distances (SS(distances)) from the projected samples to the origin. Thus, the first component called PC1 combines metabolic variables that provide the largest SS(distances) and therefore explains most of the variation between samples. Finally, other components can contribute to the description of the variation (PC2, PC3…) and create together with PC1 a new space with k-dimensions (k<p) (Fig. I.9). Clustering methods like hierarchical cluster analysis (HCA) provide

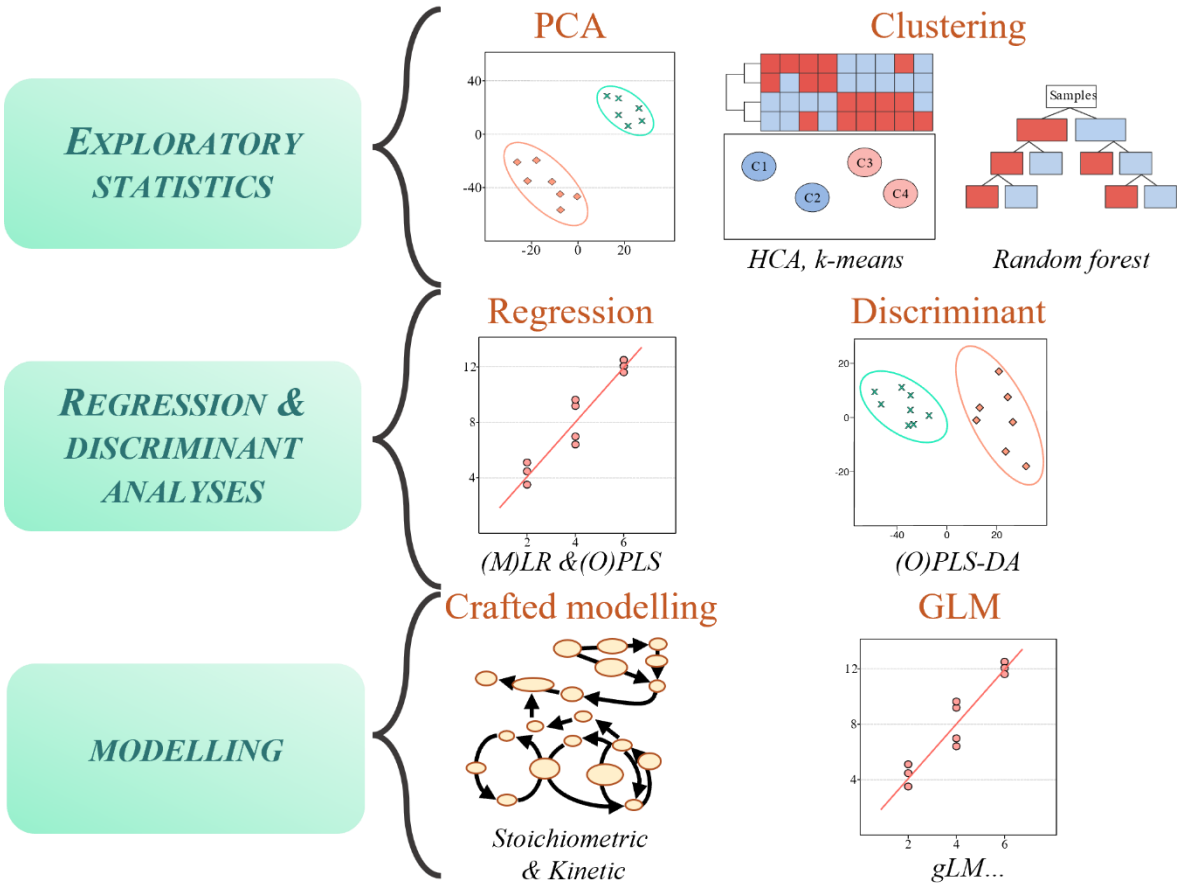

Fig. I.8 | Multivariate statistical analyses in untargeted metabolomics. *PCA*: principal component analysis, *HCA*: hierarchical clustering analysis, *(M)LR*: (multi) linear regression, *(O)PLS*: (orthogonal) partial least squares, *DA*: discriminant analysis, *GLM*: generalised linear modelling.

another unsupervised method to explore and describe the metabolic dataset by grouping metabolic variables and samples based on correlations (Hendriks et al., 2011) (Fig. I.8).

Complementary, supervised methods such as k-means clustering and random forest analysis. Shortly, k-means clustering is based on the definition of the "k" which represents the defined number of clusters in which samples will be classified (thus differing from HCA by trying to classify samples into k clusters). Then, classification relies on the objective of decreasing the total variation (*i.e.* the sum of variation between each sample and the mean value of the cluster) between the different groups of samples (Jain, 2010). Random forests are based on decision trees that classify samples based on simple successive questions (starting from the most discriminant to the less discriminant one) like "does concentration in glucose is higher than Xmg/gDW?". Then, random forest performs multiple decision tree analyses using different subsets of samples and metabolic variables to classify the separative capacity of the metabolic variables, which make the classification more efficient and accurate (Hastie et al., 2009).

Combining supervised and unsupervised exploratory methods is quite interesting first to classify samples among classes and then explore the metabolic traits that allow for this discrimination. In addition, other analyses question the predictive capacity of metabolic variables. Although multiple techniques exist to assign a "weight" to each metabolic variable and therefore perform variable selection (Hastie et al., 2009), we here focus on linear regression methods. To better explain the principles of regression analyses, the terms X and Y will refer to the explanatory matrix and the dependant or response variable, respectively.

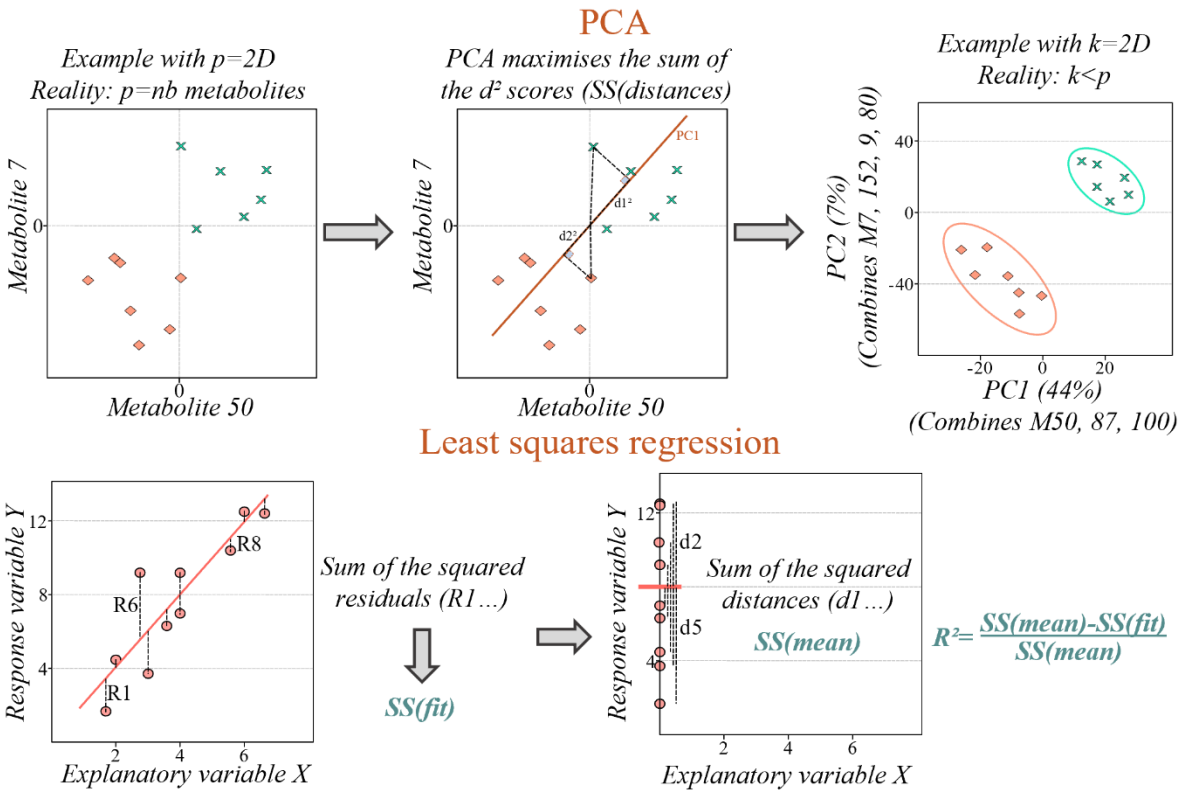

Fig. I.9 | Decomposition of Principal component analysis and least squares regression. *SS*: sum of squared. Inspired by Josh Starmer's works.

Regression and discriminant analyses

Linear regression uses the least square method (and are thus also called least square regression) to define and quantify the link between two variables. The objective of linear regression is to draw a line between two variables that will present the lowest sum of squared residuals (*i.e.* the "least squares"), where square residuals represent the distance between the data and the line (Fig. I.8 and I.9). In other words, linear regression will optimise the linear equation ($Y = aX + b$) to minimise the sum of squared residuals. In comparison, multiple regression (*i.e.* defining Y value from multiple X variables) follows the same logic under the equation $Y = slope_1X_1 + slope_2X_2 + slope_nX_n + intercept$ or in mathematical terms $Y = ß_1X_1 + ß_2X_2 + ß_nX_n + intercept$.

While these simple regression methods are very efficient in targeted analysis, their effectiveness is low when dealing with complex sets of noisy and collinear data (*i.e.* untargeted analyses). Thus, other methods were developed to deal with this chemical biodiversity.

Partial least square analysis (PLS) and orthogonal PLS (OPLS) enjoy large popularity in metabolomics since they greatly decomplexify the initial dataset before applying least square regression. PLS approach decomplexifies the initial set of explanatory variables by defining orthogonal components that explain the maximum of covariance between X and Y. These components (also called latent variables) thus explain both X and Y variances. In consequence, PLS regression models extract the components that (i) encompass the maximum variation in observable variables from X and that (ii) model the Y response (making PLS model unique when compared to other approaches as multilinear regression or principal component regression (PC regression)) (Haenlein and Kaplan, 2004). Similarly to PCA, coefficients and loading scores of the metabolic features underlying the latent variables can be extracted for further investigations. In the scenario of a single Y-variable, OPLS will provide a simpler view of the predictive model by splitting all latent variables (*i.e.* total component produced via PLS) into predictive (correlated to Y) and orthogonal (all components not related to Y) components while preserving the same predictive power. Besides, when the Y response is categorical (in opposition to numeric value), PLS and OPLS can be used as discriminant analyses (*i.e.* PLS-DA and OPLS-DA). Hence, while PLS and OPLS regression models will, *in fine*, aim to establish a quantitative relationship between the response factor (Y) and the metabolic data (X matrix), discriminant analyses seek to define a linear model to predict the class of the samples from their metabolic data (X matrix) (Rohart et al., 2017; Ruiz-Perez et al., 2020). Remarkably, whilst PLS, OPLS and their related discriminant analyses stand at the forefront of multivariate statistical analyses in metabolomics, the selection of the meaningful variables remain tedious since principal components can combine a wide range of compounds. Once again, mathematicians tackled this variable selection challenge via the development of Sparse PLS models (sPLS). sPLS makes a sparsity assumption based on the hypothesis that one can consider that only a few variables are managing the variation of the response variable Y. Hence, this method allows for a more accessible variable selection by selecting only a subset of correlated variables within the latent variables (Chun and Keles, 2010). The choice is based on lasso penalisation (a principle explained in the following paragraph) which computes a "0" coefficient to irrelevant variables, thus considered "non-essential" (Chun and Keles, 2010). Although lasso penalisation demonstrated outstanding robustness, this method is only partially efficient when the number of variables greatly exceeds the number of samples and/or (ii) when a high rate of correlations is observed in the dataset (Zou and Hastie, 2005). Thus, while parsimony is one of the main objectives in modelling approaches, variable selection remains quite challenging, especially for complex omics datasets (Engelhardt et al., 2016). Consequently, tremendous efforts were performed to develop new modelling methods able to palliate these model errors and improve the variable selection.

Even more entertaining, an entire conceptual movement called systems biology emerged to reinforce the link between modelling and wet lab experiments in both targeted and untargeted analyses (Engelhardt et al., 2016).

Modelling: the error-based learning approach

Systems biology is based on the fact that models (either predictive or reaction-based) simplify reality. For instance, kinetic models used in targeted approaches can only handle known metabolic pathways, while the dataset from untargeted strategies referenced only a part of the plant metabolome (depending on the experimental conditions) (Engelhardt et al., 2016). Besides, biological systems are still largely unknown, and much of metabolic biodiversity and interactions remain inaccessible, yielding incomplete and incorrect models. That said, modelling is a potent tool that could be defined as a simplification of reality that aims to predict or explain a specific outcome using explanatory variables. Thus, while models first unveil predictive capacity and/or the role of the essential elements, they also emphasize the place of non-essential and inaccessible variables in the outcome variation via determination coefficient R^2 in predictive models (explanatory variables rarely explained 100% of the response variable), or via external parameters in kinetic modelling, for instance. Hence, modelling is a statistical approach generating experimental hypotheses that result from (i) the contextualisation of explanatory variables and (ii) the unexplainable part of the outcome parameter variation (*i.e.* non-essential and/or inaccessible/unknown parameters). The systems biology concept represents this iteration between modelling and experimentation, where experimental hypotheses yield new experiments that produce new molecular data to continuously improve our comprehension of the plant system (Engelhardt et al., 2016). Here, we first present the latest advances in modelling methods with a specific focus on generalised linear models (GLM). Then, the development of innovative analytical tools that integrate the best predictors are described.

- "*Crafted" modelling approaches* -

Reductionist and integrative approaches greatly enriched our knowledge of plant metabolism. The discovery of major biochemical pathways arose from the integration of chemical compounds and their interactions into a scheme representing a flow of chemical reactions mainly catalysed by enzymes (Fernie et al., 2004). Similarly, several modelling tools are now used to gather and organise the vast amount of data obtained on primary and secondary metabolism. These statistical tools are mostly related to targeted approaches and permit the modelling of biochemical pathways and their activities as a constant chemical flow. The construction of these models is mostly "crafted" and in the form of a "step-by-step" process, where the addition of a reaction, an enzyme activity, or an unknown external factor is analysed according to its impact on the output variable (Bordbar et al., 2014; Beauvoit et al., 2018). Although other techniques are emerging, the two most artisanal modelling techniques are either

constraint-based (*i.e.* stoichiometric modelling) or enzyme activity-based (*i.e.* kinetic modelling) (Fig. I.8).

Stoichiometric models represent a scheme of successive reactions generally catalysed by enzymes founded on the assumption of a pseudo-steady state where the balance of the different chemical elements is maintained (Bordbar et al., 2014). Besides, this list of chemical reactions is associated with multiple constraints like input/output rates to provide a functional and dynamic scheme of metabolic fluxes (Bordbar et al., 2014). However, while stoichiometric models are often applied at the genome-scale, they mostly remain focused on one biochemical pathway or cellular compartment biology due to the complexity of this biological system (*e.g.* compartmentalisation) (Colombié et al., 2017). Kinetic models reconstruct dynamic metabolic networks using enzyme activities (Rohwer, 2012). This approach requires the description of the stoichiometry of the studied pathway, but also the kinetic properties as well as related subcellular compartmentalisation notions. Reactant concentration and kinetic properties (*e.g*. maximal activity) are then used to estimate and predict metabolic fluxes through the use of kinetic rate laws like the Michaelis-Menten equation (Rohwer, 2012). Hence, kinetic models provide a dynamic view of plant metabolic pathways, which offers the unique opportunity of (i) identifying potential targets for plant engineering and (ii) explaining and even predicting the variation of substrates and enzyme activities in response to environmental perturbation, for instance (Beauvoit et al., 2018). Besides, these dynamic models usually include external parameters to improve the fit between mathematical predictions and experimental data. These external parameters (*e.g.* compartmentalisation, transport) represent a golden resource by pinpointing the unknown parameters that need to be studied to further improve our understanding of the plant system. However, although artisanal modelling has provided critical successes in plant biology, the construction of these dynamic models is based on the use of well-known compounds and reactions (Rohwer, 2012; Bordbar et al., 2014).

- *Generalised linear models (GLM)* -

The amount of data generated has become exponential since the development of omics, leading to significant challenges like “how to handle variables whose only known characteristics are their mass-to-charge (*m/z*) ratio and their retention time?” or like “how to extract the explanatory variables from a highly noisy and collinear dataset?”. The advantages of GLM methods are multiples and cover (i) the possibility to handle datasets that do not follow a normal distribution, (ii) a sophisticated penalisation system to improve variable selection using various penalty systems, and (iii) the possibility to preserve the biological context (Nelder and Wedderburn, 1972; Zou and Hastie, 2005; Stroup, 2015). Generalised linear models are built from (i) a systematic component (*i.e.* the function $Y = ß_1X_1 + ß_2X_2 + ..$) which is equal to the one from multilinear models, (ii) a link function that links the linear model ($Y= ß_1X_1 \ldots$) to the response variable using a function (*e.g.* $Y=\log(X)$, or $Y=X^2$ where X is the systematic component), and (iii) a random component which is the distribution type of the data (*e.g.* normal, Poisson) (Fig. I.10). For instance, general linear models (gLM) are an example of a generalised linear model (GLM) where

the link function is Y=1X and therefore preserve the linearity of the model and is therefore exclusive to normalised datasets. Similarly, a binomial or Poisson link function will be if we analyse two groups responses (*e.g.* blue or yellow flowers) or if we analyse discrete values (*e.g.* number of times that a plant survives the stress occurrence) (Nelder and Wedderburn, 1972; Stroup, 2015).

GLM MODELS

$$Y = Link(\underbrace{\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_n X_n}_{X})$$

Name	Systematic component	Link function	Random component
General linear model (gLM)	$\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_n X_n$	$1X$	Normal
Poisson	$\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_n X_n$	e^X	Poisson

Fig. I.10 | Generalised linear models (GLM). The three components of a GLM model. Random component represents the distribution of the data. General linear models (gLM) are an example of generalised linear models (GLM) where the random component follows a normal distribution.

The second advantage of GLM models is their use of various penalty systems when defining the systematic component (*i.e.* line function as mentioned in the previous paragraph). In contrast, other multivariate analyses as sPLS only use one. Variable selection is based on the balance between obtaining a great fit (*i.e.* great R^2 score) with maximum parsimony while avoiding overfitting. General linear models (gLM) are characterised as regularised models based on the use of penalty methods (also called regularisation methods) divided into ridge, lasso and elastic net methods (Zou and Hastie, 2005; Bunea et al., 2011). As explained before, linear and multilinear regression aims to find a line that results in the minimum sum of squared residuals (Fig. I.8). However, minimising the least squared residuals greatly improve the risk of overfitting the data, which means that the model will be trained to fit the data greatly but will fail to fit additional data (Chicco, 2017). gLM models are always performed using a training set and a testing set to tackle this overfit risk through the modulation of two parameters called “bias” (*i.e.* the inability to capture the true relationship between two variables) and “variance” (*i.e.* the difference of the sum of square errors between training and testing set) (Fig. I.11). Thus, overfitting is characterised by a low bias but high variance. gLM models tend to limit the possibility of overfitting by using penalties, which will increase the bias and decrease the variance (Fig. I.11). Ridge and lasso penalty values are defined as $\lambda \times slope^2$ or $\lambda \times |slope|$, where λ is the penalty parameter that can go from 0 to $+\infty$. Besides, the λ is defined using cross-validations, which consist in varying 2 to 10 times the sample distribution between training and testing sets (Chicco, 2017). Multiple λ values from 0 to $+\infty$ are tested in each condition and the one allowing the lowest variance is chosen. Importantly, while the line equation (*i.e.* systematic component) of linear regression aims to limit the sum of square residuals,

the equation from ridge and lasso regression aims to minimise the sum of square residuals + the ridge or lasso penalty (Fig. I.11). Similarly, when using general multilinear model with ridge penalty, the equation of the line ($Y = \text{intercept} + \text{slope}_1X_1 + \text{slope}_2X_2 + \text{slope}_nX_n$) aim to limit the sum of the square residuals + λ x ($\text{slope}_1^2 + \text{slope}_2^2 + \text{slope}_n^2$). The major difference between lasso and ridge penalty is that by using a penalty of λ x slope^2 (also called L2-penalty), ridge method will shrink the slope of the variables but will not exclude variables, while lasso penalty (λ x | slope |, also called L1-penalty) allows a slope equal to 0 (Schmidt, 2005; Zou and Hastie, 2005) (Fig. I.11). Thus, lasso penalty allows variables exclusion while ridge regression only shrinks the predictive impact of variables. However, L1-penalty is not truly performant when dealing with more predictors than samples and high collinearity among variables. Besides, when dealing with thousands of variables, the choice between ridge and lasso regularisation method is tedious (Zou and Hastie, 2005). Consequently, a last regularisation method called elastic net method allow to combine both L1 and L2 penalty automatically under the penalty: (λ_1 x ($\text{slope}_1^2 + \text{slope}_n^2$)) + ($\lambda_2$ x (| slope_1 | + | slope_n |)). Thus, the regularised slope with elastic net method aims to minimise the sum of the squared residuals + elastic net penalty. In consequence, elastic-net regression allows a good compromise between shrinking and removing variables. Hence, elastic net method applied to gLM allows a great parsimony while preserving the correlated variables, which finally leads to a better variable selection (Zou and Hastie, 2005) (Fig. I.11).

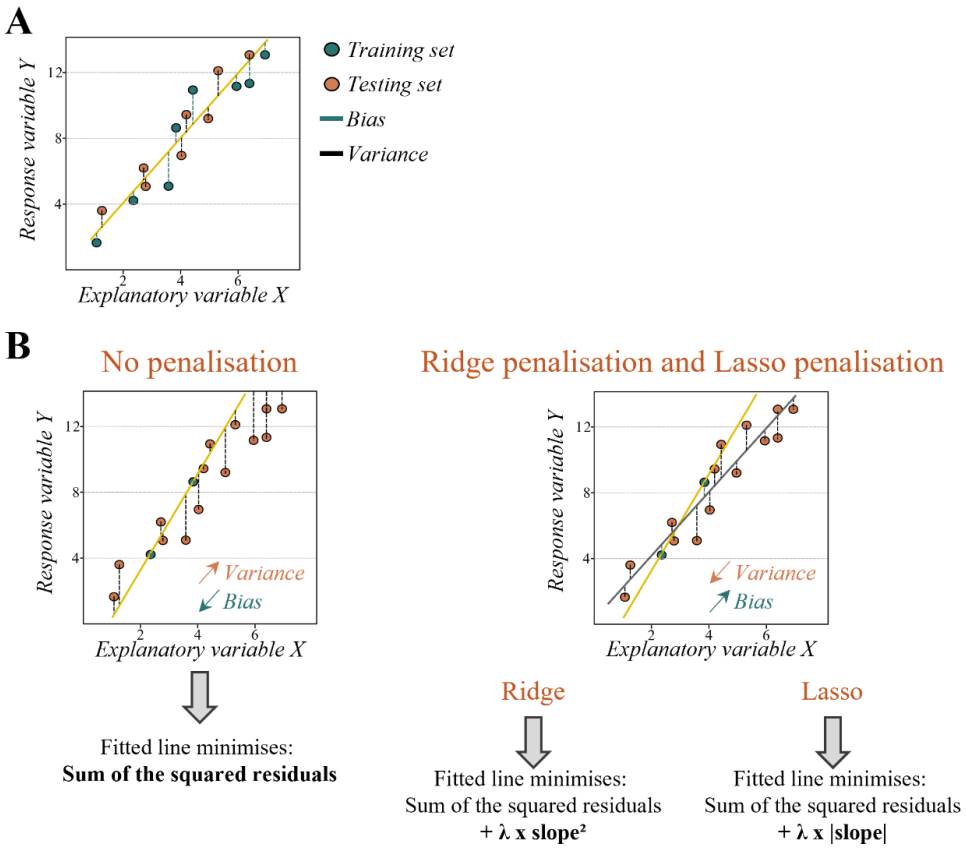

Fig. I.11 | Penalisation methods in general linear models (gLM). A. Division of the total sample set into a training set (80% dataset) and a testing set (20% dataset). **B.** Effect of penalisation method on variance and bias. Inspired by Josh Starmer's works.

Overall, we have seen that gLM models improve variable selection on normally distributed datasets using cross-validation and regularisation (penalty). However, an ultimate validation is required to perform "real" predictions (Fig. I.12). The predictive capacity of a model is defined when applying the final model equation (i.e. systematic component + link function of 1 in the case of a general linear model) built with both training and testing datasets on the validation set (Fig. I.12). Finally, permutation datasets are developed (where Y scores are randomly assigned to each sample) to test the likelihood of spurious predictions (Luna et al., 2020). Based on its capacity to limit overfit possibilities and improve variable selection, GLM approach is thus considered the best modelling method to provide meaningful discoveries in plant biology. For instance, GLM was used to unveil metabolic compounds that underpin resistance against biotic stress (Simmons and Gurr, 2004; Luna et al., 2020), correlate antioxidant capacity and phenolic profile (Loupit et al., 2020), predict plant adaptation to extreme ecosystems (Dussarrat et al., 2022, see Chapter 3).

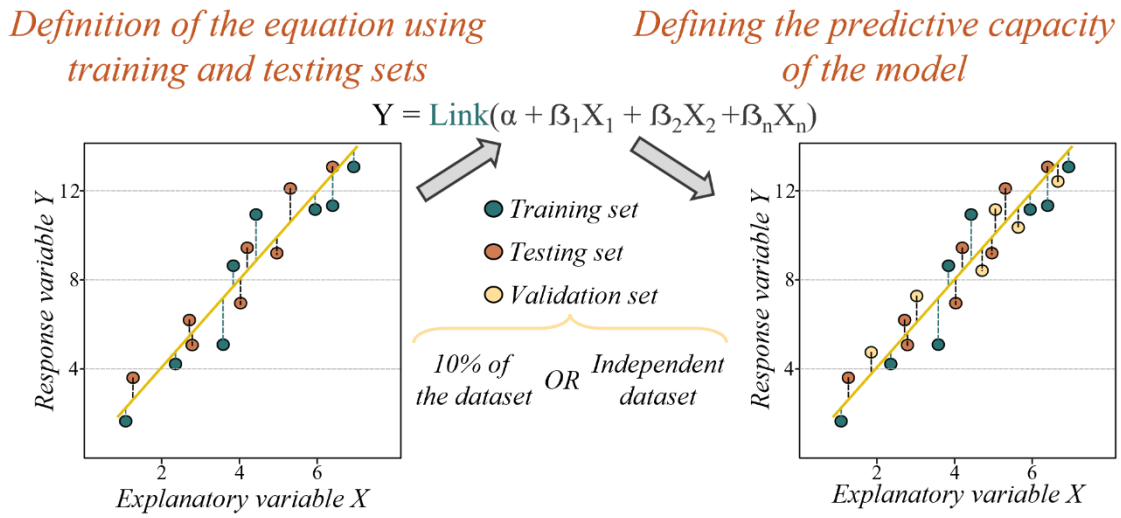

Fig. I.12 | Validation of the predictive capacity of the gLM models using a validation set.

Metabolic networks

Finally, complementary techniques are used to integrate the best predictors within metabolic networks to contextualise these markers and provide another level of biological interpretation. The integration of the best predictors within plant metabolism can be achieved via two principal ways described in Chapter 2. First, best predictors can be directly plotted into pre-existent metabolic networks via MetExplore (https://metexplore.toulouse.inrae.fr) (Cottret et al., 2010). Hence, MetExplore allows to observe and analyse the connections between the metabolite of interest and various metabolic pathways at the genome-scale. Alternatively, untargeted datasets produced via LC or GCMS/MS can be processed through GNPS platform (https://gnps.ucsd.edu) using a Feature-based molecular network method (FBMN). FBMN networks are based on the assumption that two compounds with a similar chemical structure will be interconnected (Nothias et al., 2020). Thus, structural-based networks facilitate the integration of the best predictors within metabolic clusters. Interestingly, whilst MetExplore require annotated compounds, FBMN networks is a valuable bioinformatics tool that (i) offers the possibility to easily observe if several compounds from the same biochemical pathway are involved in the outcome studied and (ii) facilitate the annotation of the best predictors (Nothias et al., 2020). Hence, the combination of untargeted approach using highly sensitive techniques with GLM statistical analysis and the integration of the best markers within metabolic networks offers a direct way from the unknown to the discovery of meaningful predictors. Besides, the integration of these predictors into the metabolome is the cornerstone for the comprehension of the metabolic mechanisms governing the plant response to the studied parameter. The top-down approach thus represents a pyramidal concept where a maximum of chemical diversity is first captured and then decomplexified into essential and non-essential elements to finally unveil integrated metabolic mechanisms underlying the plant response.

Overall, a fantastic scientific road has been travelled since discovering the “one gene-one enzyme-one function” concept. This initial concept led to the description of a myriad of relationships between genes, proteins and phenotypes. Despite a reductionist view, this concept is the root of the emergence of omics science and the integrative approach. The integration of genomics, transcriptomics, proteomics and metabolomics data revealed the fascinating complexity of the plant system and the diversity of its metabolism (Scossa and Fernie, 2020). The subsequent implementation of bioinformatics tools was later at the forefront of the development of systems biology (Engelhardt et al., 2016). Systems biology, which refers to the iteration between modelling and experiments, permitted this chemical diversity to be assembled and organised into metabolic networks (Rohwer, 2012; Bordbar et al., 2014). Furthermore, the combination of the latest technological (MS/MS and MS^n) and statistical (GLM models) advances have made possible the impossible by enabling the extraction of the essential elements from complex and poorly known datasets. Thus, the use of the tandem untargeted-modelling approach revealed key metabolites capable of predicting the variation of various phenotypic traits (Luna et al., 2020). Besides, the integration of these predictors into metabolic networks allows for a better perception

of the metabolic mechanisms involved (Nothias et al., 2020). Despite its advantages, this analytical tandem is still poorly used, possibly owing to particular bottlenecks such as the annotation process and the difficulty of characterising the structure of unknown molecules. However, there is no doubt that the future development of analytical and statistical techniques will overcome these challenges and finally make the untargeted-modelling tandem an indispensable approach that will improve our understanding of complex living systems. For instance, this approach is ideal for investigating unknown biological systems and thus allows the analysis of complex output variables such as adaptation to extreme environments.

II. SPATIO-TEMPORAL ADAPTATIONS IN EXTREME PLANTS

Wild species are a precious resource for crops improvement, as evolution has naturally fixed stress-tolerant traits that allow plant survival under changing conditions (Yolcu et al., 2020). This is particularly evident for species that thrive at the edges of plant life compatible gradient and are termed "lovers" (philos in Ancient Greek) of "extremes" (*i.e.* extremophiles). While abiotic stresses are becoming more frequent and intense, the analysis of the adaptive mechanisms of extreme plants holds great promise for improving our understanding of plant resistance to hostile conditions (Eshel et al., 2021; Hasegawa et al., 2021). Interestingly, studies highlighted temporal and spatial adaptations (*i.e.* life cycle adaptation and biological matter adaptation, respectively). Besides, metabolism has been brought to the forefront of the adaptation process through the discovery of a myriad of biochemical compounds that underpin structural and physiological adjustments (Kumari et al., 2020). However, the transfer of this knowledge to agronomic plants yielded few results (Turner, 2018). In addition, selective approaches that focused on a unique or limited number of species and environments led to the discovery of highly specific adaptive mechanisms (Scossa and Fernie, 2020).

The objective of the following review was to provide a meta-analysis of the adaptive mechanisms developed by extreme plants, from the morpho-anatomical to the metabolic scale (Dussarrat et al., 2021). Besides, this review presents the main successes of the research in extreme environments. We then explored the reasons and challenges explaining the gap between the discovery of adaptation markers and their transfer to crop species. Finally, this article exposed the potential strategies to meet these challenges, and therefore enhance our understanding of plant strategies for extreme climate resilience. We here present several parts of this review that have been slightly modified for better readability. However, the complete version is available in Annex I.A1.

II.1. Metabolic adaptations at the heart of plant adaptation to harsh lands

Metabolism is the cornerstone of plant responses to environmental changes. Hence, the metabolic adaptations improving survival chances in hostile ecosystems have received considerable attention over the past few decades. Using both targeted and untargeted metabolomics techniques, great advances have been achieved to describe these metabolic features by (i) comparing extreme species to related crops or model plants (Lugan et al., 2010; Yobi et al., 2013), (ii) analysing plant metabolic profiles in their natural environment (Tipirdamaz et al., 2006), and (iii) characterising plant metabolic responses through a gradient of abiotic stresses (Kumari et al., 2020). Untargeted and targeted analyses on extremophile plants transferred to controlled conditions allowed isolating one environmental variable to compare its effect on extremophile *versus* model species and highlight different metabolic response strategies. Complementarily, metabolic profiling of extremophile plant species described the different biochemical compounds accumulated within a given ecosystem, while the study of these organisms through a stress gradient (*e.g.* altitude, salinity) improves our understanding of how extreme plant metabolomes adapt to abiotic stress under extreme conditions. This review sought to summarise these works, covering 5 different ecosystems (*i.e.* deserts, mountains, frozen lands, saline lands and metal-contaminated sites), thereby pinpointing the convergences of the metabolic responses between species in one ecosystems and between ecosystems (Fig. I.13). This meta-analysis included 69 species and revealed a dynamic response of central, primary, secondary and redox metabolisms. However, only 31.6 % of the referenced metabolites referred to secondary metabolism (Fig. I.13A). This suggests that these specialised metabolic pathways are often overlooked, which urges for enhancing untargeted analyses with wide metabolome coverage of extremophile plant species.

Central metabolism

In extreme environments, the adjustment of central pathways under abiotic stress involved compounds that were up-or down-regulated depending on the plant species and ecosystem (Fig. I.13B). Total soluble sugars, including glucose, fructose and sucrose, were likely to be induced with the altitude gradient (Hashim et al., 2020). However, while high sucrose concentration was reported independently of the plant species and environment, glucose and fructose regulations were quite variable (Fig. I.13B). Interestingly, the same trends were observed within the TCA cycle where malate accumulated for all conditions and species, whilst citrate, fumarate and succinate levels fluctuated, an observation perhaps explained by the fact that major organic acid content depends on the plant species (Chia et al., 2000; Mikulic-Petkovsek et al., 2012; Igamberdiev and Eprintsev, 2016). These results support the central place of these pathways in managing carbon (C) resource allocation between primary and secondary pathways and in determining C allocation between plant growth and defence under extreme environmental conditions.

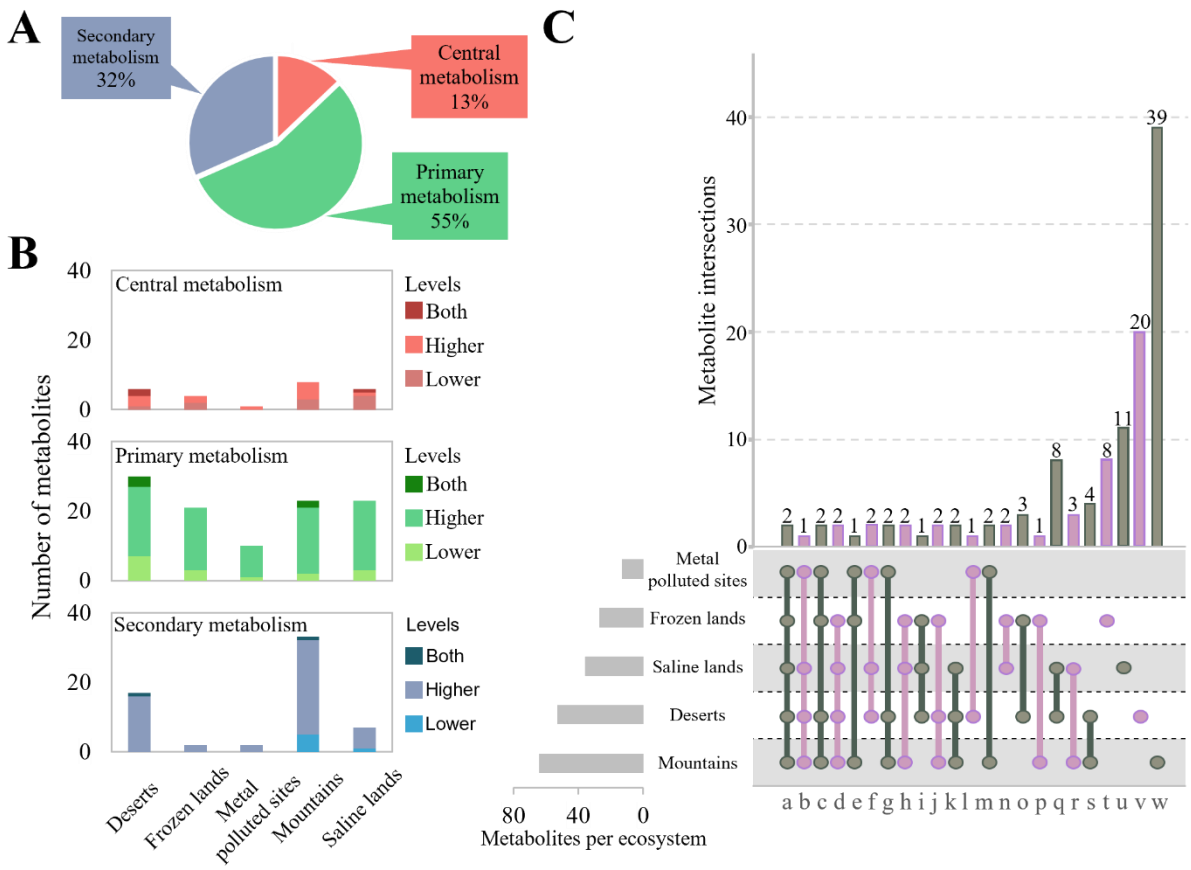

Fig. I.13 | A comprehensive meta-analysis of the metabolic features observed in plants thriving under extreme environments. A. Distribution of affected metabolic pathways. **B.** Details of the distribution in 5 extreme ecosystems and response of molecules to environmental pressures: Higher or Lower concentrations of compounds, or Both, depending on the plant species. **C.** UpSet plot of the metabolic overlaps between ecosystems. The bottom left side shows the number of metabolites described for each ecosystem. The right side shows the possibilities and levels of intersections of these molecules between the 5 extreme lands. Letters refer to compounds asoociated with the intersection (one letter per column). a: aspartate, inositol; b: proline; c: ascorbate, galactinol; d: glucose, sucrose; e: trehalose; f: glutamate, glutamine; g: glutathione, histidine; h: fructose, threonine; i: GABA, j: malate, raffinose; k: citrate, phenylalanine; l: mannitol; m: sorbitol, starch; n: pinitol, valine; o: stachyose, tocopherol, xylose; p: melibiose; q: alanine, asparagine, choline, choline-O-sulfate, fumarate, glycine, putrescine, serine; r: glycine betaine, lysine, rhamnose; s: arabinose, arginine, spermidine, zeaxanthin; t: arachidonic acid, catechin, digalactosylglycerol, linoleic acid, linolenic acid, oleic acid, ononitol, tryptophan; u: beta-alanine betaine, fucose, isoleucine, kaempferol, laricitrin, leucine, malonic acid, norvaline, quercetin, succinate, tyrosine; v: allantoin, apigenin, galactose, guanosine, hydroxyproline, isocitrate, lutein, luteolin, naringenin, neoxanthin, ribose, spermine, tri-O-galloylquinic acid, verbascose, violaxanthin, γ-glutamylisoleucine, γ-glutamylleucine, γ-glutamylmethionine, γ-glutamylphenylalanine, γ-glutamylthreonine; w: 2-oxoglutarate, abscisic acid, caffeoylquinic acids, cinnamic acid, coumaric acid, farnesene, ferulic acid, gentiobiose, glycerol, hypoxanthine, jacareubin, kokusaginine, maltose, melezitose, MHDglycerol, neohesperidin, PicrosideI, PicrosideII, PicrosideIII, PicrosideIV, quebrachitol, ranunculin, rebeccamycin, romucosine B, sagecoumarin, sagerinic acid, seychellene, tagatose, tannin, thapsigargin, thujone, total carotenoids, total flavonoids, total sugards, total phenolic content, total xantophyll pigments, tricetin, xanthosine, xylitol.
Adapted rom Dussarrat et al. 2021.

Primary metabolism

Compatible solutes, organic osmolytes responsible for osmotic balance and compatible with cellular metabolism (Galinski, 1993), have shown to accumulate in both drought- and freezing-tolerant plant species for their role in osmotic adjustment and cryoprotection (Chen and Murata, 2002; Bhandari and Nayyar, 2014). Under extreme low temperatures, adapted woody plants have shown high concentrations of oligosaccharides, which regulate viscosity in the cytoplasm and therefore prevent deleterious effects of freezing temperatures (Stushnoff et al., 1997; Strimbeck et al., 2015). Contents in disaccharides (*e.g.* trehalose), as well as RFOs like raffinose, were heightened under extreme temperatures, drought and heavy metal contamination (Fig. I.13C). Also, inositol, pinitol, mannitol, sorbitol and galactinol were over-represented within desert, cold-tolerant and hyperaccumulator plant species, and likely to be upregulated under each abiotic stress (Fig. I.13C) (Slama et al., 2015). This observation agrees with the possible roles of sugar alcohols as carbohydrate reserves or the thermoprotective function of sorbitol in higher plants (Moing, 2000).

Lipid profiling of resurrection plants has revealed an adjustment in lipid metabolism when submitted to drought stress (Quartacci, 2002), resulting in higher unsaturation levels (Tshabuse et al., 2018). Similarly, a positive correlation was observed between polyunsaturated fatty acid levels from seabuckthorn species and altitude in the Himalayan mountains (Sharma et al., 2020), supporting their role in maintaining membrane fluidity (Upchurch, 2008). These observations in both freezing- and drought-tolerant extremophile plants suggest a central role of lipid remodelling in adaptation to extreme environments, by countering the effects of direct drought and cellular dehydration initiated by extracellular ice formation (Moellering et al., 2010).

Besides sugars, plant resistance under extreme conditions is thought to be partially related to the induction of amino acids, which are important metabolic intermediates for the synthesis of environment-responsive, specialised metabolites (Chouhan et al., 2017). Twenty-six extremophile species accumulated proline under various stressful conditions. Notably, the central role of proline in plant tolerance to multiple stresses could result from i) its synthesis that limits the reducing power, thus leading to an imbalance of photosynthetic activity, and ii) the benefits of proline degradation that provides C to the TCA cycle and thus contributes to respiratory activity (Kaur and Asthir, 2015). Besides, the shikimic acid pathway activity resulted in the accumulation of tyrosine and phenylalanine in some of these plants. These aromatic amino acids act as precursors for the biosynthesis of flavonoids, known as secondary antioxidant compounds (Chouhan et al., 2017), suggesting a possible role of secondary pathways in adaptation.

Secondary metabolism

Plant secondary metabolism has been already recognised as a major actor of plant-environment interaction, involving many specialised metabolites that accumulate under stressful conditions (Akula and Ravishankar, 2011). For instance, linear accumulations of polyphenols were revealed within several altitudinal gradients for different species (Zidorn, 2010; Monschein et al., 2015; Cirak et al., 2017). In extreme conditions, the impact of abiotic stress on secondary metabolism was observed with 29 plant species from different extreme environments (Fig. I.13B). The potential role in the adaptation of several molecules including quaternary ammonium compounds, terpenes and phenolics were also mentioned for different plant families. Interestingly, these compounds were not specific to extremophile species and were further present in crop and model plants (Parida et al., 2018). This suggests that secondary metabolites found in various plants could play important roles in stress mitigation of harsh climates.

- *N-related compounds* -

Molecular mechanisms by which polyamines alleviate plant tolerance to abiotic stress are not fully understood but several works have reported interesting properties and results for such compounds. The application of these compounds could regulate the size of potassium channels and therefore the aperture of pores in the plasma membrane, which suggests a role in water loss control (Alcázar et al., 2010). Besides, polyamine accumulation promotes ROS degradation by raising antioxidant enzyme activities and possibly affects ion transport under salt stress (Saha et al., 2015).

Remarkably, polyamine profiles in extremophiles appear up- or down-regulated depending on the environmental stress (Fig. I.13B). This observation could be explained by the fact that plant resistance is more likely to be associated with a high ratio of (Spermidine + Spermine)/Putrescine, suggesting that this protective role mainly involves higher polyamines (Zapata et al., 2004; Chen et al., 2019). Another hypothesis is that polyamines could conjugate with other molecules, such as coumaric or caffeoyl acids, and lead to complex roles in plant defence (Alcázar et al., 2010; Burt et al., 2019). However, few details on the occurrence and function of such conjugated forms in extremophile plant species are available.

Quaternary ammonium compounds have been controversially considered as an adaptive response of halophytes or other extremophile plants that improves plant tolerance to drought, salt and low-temperature stress (Ashraf and Foolad, 2007). Their input in plant resistance mechanisms was illustrated by the high concentrations of glycine betaine and choline-O-sulfate found in Alpine, desert, and halophyte plant species (Fig. I.13C). These compounds could be involved in osmoregulation by modulating Na^+ and K^+ content, leading to a higher K^+/Na^+ ratio, which alleviates salt tolerance in higher plants (Hu et al., 2012). Additionally, glycine betaine has shown a relevant role in maintaining membrane integrity as well as enzyme activities (Annunziata et al., 2019).

Hence, while several extreme plants displayed important levels of N related compounds in hostile environments suggesting a possible role in adaptation, great variabilities were observed about polyamine accumulation and the understanding of how both polyamines and quaternary ammonium compounds are integrated in the response to abiotic stress of the different plant species. These observations raise the critical need of analysing metabolic features in a more holistic approach where tolerance mechanisms would be integrated into the different plant systems thriving in hostile ecosystems.

- *Terpenoids* -

The role of carotenoids and tocopherols in photoprotection through antioxidant activity has been already widely characterised, but new insights were highlighted when describing the mechanism by which ROS can cause the oxidative cleavage of carotenoids leading to hormonal compounds such as phytohormones (*e.g.* strigolactone or abscisic acid) (Havaux, 2014). Interestingly, the impact of abiotic stress on photosynthetic pigments and tocopherols is verified in extreme environments where both compound classes were accumulated in plants from desert and mountain ecosystems (Fig. I.13C). Besides, the increased levels of zeaxanthin and abscisic acid found in Alpine (Fig. I.13) and most resurrection plants (Rascio and Rocca, 2005) possibly illustrate the link between carotenoids and phytohormones. Altogether, these observations in extremophile plants emphasise the pivotal role for terpenes in stress signalling through hormonal responses, and stress mitigation via the processing of excess ROS in response to extreme temperatures and radiation levels, for instance.

- *Phenolics* -

The contributions of phenolics to extreme environmental stress responses were reflected with the high polyphenol concentrations in plants of Alpine and desert environments (Fig. I.13C). Furthermore, phenolics accumulated in a wide biodiversity of extremophile species and medicinal plants, some of which thrive under harsh conditions (Li et al., 2020b; Najjaa et al., 2020). These observations corroborate the fact that flavonoids hold an important antioxidant function improving photoprotection and reducing damage caused by UV radiations and frost (Agati and Tattini, 2010; Schulz et al., 2016). Recently, the UV-B protective function was extended to a global enhancement of ROS-processing activity independently of the solar wavelength proportions, based on the upregulation of flavonoids in response to ROS accumulation and an imbalance of redox homeostasis (Di Ferdinando et al., 2012). Additionally, these same authors argue in favour of this hypothesis by highlighting the interaction between cold and N stress that leads to the same flavonoid upregulation profiles as for cold and high light conditions. On the other hand, the accumulation of metabolites like cinnamic acid under extreme conditions may have a role in the production of lignin compounds known to be upregulated under abiotic stress to reinforce secondary cell walls (Le Gall et al., 2015).

Finally, it is noteworthy that the biosynthesis of several phenolics, proline, and polyols consume NADPH (Loescher and Everard, 2000; Szabados and Savouré, 2010; Caretto et al., 2015a), and thus participate in the control of cellular redox homeostasis by limiting the excess of reducing power. Hence, the increase of the $NADP^+$/NADPH ratio would possibly enhance the oxidative pentose phosphate pathway activity, which provides precursors for phenolic compound production (Caretto et al., 2015a). Altogether, these observations suggest that not only the proper function of each primary or secondary compound matters but also both their biosynthesis and degradation, which therefore emphasises the need for a more integrated approach to study the metabolic features of extreme plants.

Tuning redox metabolism upon extreme climates

Plant central metabolism produces ROS mostly via three sources that include chloroplastic photosynthesis, mitochondrial respiration and peroxisomal photorespiration (Schertl and Braun, 2014) (Fig. I.14A). In photosynthetic tissues, ROS mainly originate from the photosynthetic electron transport chain (Foyer, 2018), while other sources are important for different organs such as fruit tissues (Decros et al., 2019). Therefore, when redox homeostasis becomes unbalanced, a lack of ROS or ROS accumulation leads to reductive or oxidative stress, respectively, characterised by, for instance, DNA damages, oxidation of cysteine residues in proteins and lipid peroxidation inducing retrograde signalling or in some cases cell death (Fig. I.14B). Harmonious plant growth thus requires a finely tuned redox homeostasis to avoid oxidative or reductive stress. Due to their sessile lifestyle, plants have developed numerous antioxidative defence pathways and strategies to control their redox state (Fig. I.14A).

In harsh environments, plants growing at high altitudes undergo extreme temperature variations, water limitation, nutrient deficiency and high levels of irradiation inducing a reduced photosynthetic activity, which results in a surplus of reducing power (NADPH) and higher (photo)respiration (Fernández-Marín et al., 2020). All these abiotic stresses have been shown to exacerbate ROS production by central metabolism in model or agronomical species grown under laboratory conditions (Choudhury et al., 2017; Pandey et al., 2017). To prevent oxidative damages, Alpine plants adapt their photosynthetic defence machinery by increasing their free radical-scavenging capacity through the accumulation of photoprotective metabolites such as carotenoids, flavonoids and phenolics (Ma et al., 2015; Cui et al., 2019; Hashim et al., 2020). These specialised plant compounds are powerful antioxidants that process ROS, consume reducing power and can also avoid UV-induced damages (Bieza and Lois, 2001; Caretto et al., 2015b; Young and Lowe, 2018) (Fig. I.14A). A first field study of nine Alpine plants from different altitudes described a higher content in total leaf antioxidants, especially in ascorbate (Wildi and Lutz, 1996), suggesting an important role of redox homeostasis in plant acclimation. More recently, biochemical and proteomic analysis of Tibetan plants highlighted a positive correlation between the content of soluble antioxidants (ascorbate and phenolics), ROS-processing enzyme activity and altitude (Ma et al., 2015; Cui et al., 2019).

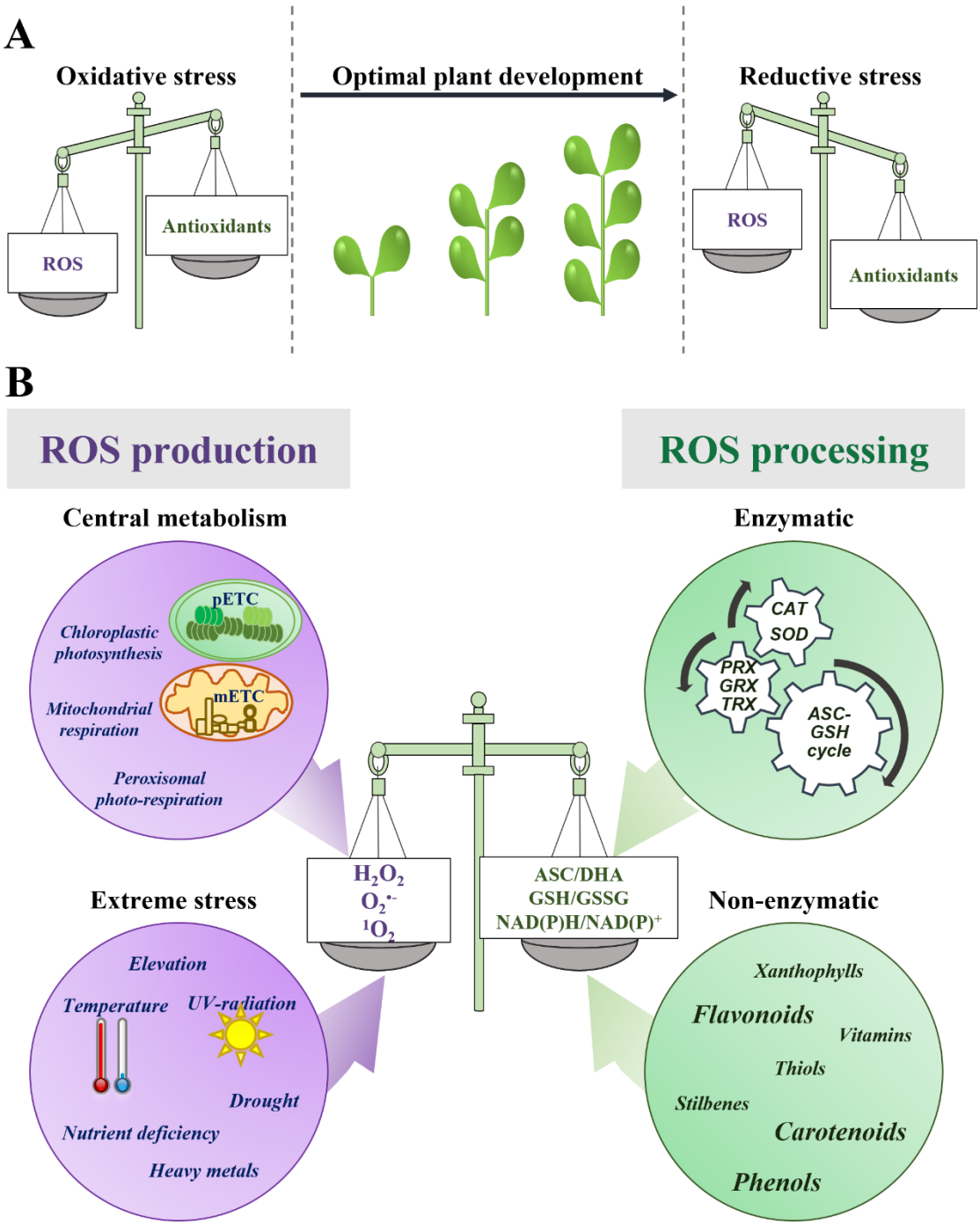

Fig. I.14 | Redox poise is pivotal to plant growth and acclimation. A. Harmonious plant growth requires a finely tuned redox homeostasis to avoid oxidative or reductive stress when the ROS/antioxidant balance is altered. **B.** Plants produce ROS and other redox signals during growth and in response to environmental stimuli. Redox homeostasis relies on the balance between ROS production (left side) and processing (right side). *ASC*: ascorbate; *CAT*: catalase; *DHA*: dehydroascorbate; *GRX*: glutaredoxins; *GSH*: glutathione; *GSSG*, disulfide glutathione; *mETC*: mitochondrial electron transport chain; *pETC*: photosynthetic electron transport chain; *PRX*: peroxiredoxins; *ROS*: reactive oxygen species; *SOD*: superoxide dismutase; *TRX*: thioredoxins. Adapted from Dussarrat et al. 2021.

Altogether, these studies of endemic species from the Tibetan plateau reported a strong correlation between the altitude gradient and the plant redox metabolism, which was characterised by higher enzymatic and non-enzymatic ROS processing capacities. Likewise, desert species also harboured higher contents in carotenoids and flavonoids, and increased enzymatic activities of the ascorbate-glutathione cycle, particularly glutathione reductase, suggesting a more important role of glutathione- and thiol-related signalling in the adaptation to desert lands (Streb et al., 1997; Talbi et al., 2015; Wang et al., 2016). Additionally, the high oxidation state of glutathione and ascorbate have been correlated with desiccation and rehydration tolerance of two resurrection plants (Kranner et al., 2002; Jiang et al., 2007). Finally, a comparative study reported that the native Antarctic species *Colobanthus quitensis* harboured a total antioxidant activity twenty times higher than its genetically-related species *Dianthus chinensis* combined with a two-fold increased respiration rate and activity of alternative oxidase after cold exposure (Clemente-Moreno et al., 2020). This illustrates the plasticity of redox pathways to maintain the redox poise depending on the environment and plant botanical taxa.

Hence, cellular redox homeostasis is a key factor that accompanies plant growth and responses to the environment, more remarkably within the activity of the ascorbate-glutathione cycle. In addition, antioxidant secondary metabolism (*e.g.* carotenoids, flavonoids and phenols) further appears as a relevant pathway to stimulate plant oxidative defence capacity and thus participates in the acclimation of plants to harsh environments. Currently, the paradigm of redox biology tends to display a bigger and clearer picture of the redox network occurring in plants, where multiple sources of ROS are possible and associated with many "ROS processing systems" (Noctor et al., 2018). Spatial, temporal, metabolic and antioxidant specificities are multiple factors that can influence redox signalling. While knowledge on redox biology in plants living in extreme environments is still fragmentary, the concepts that originate from model and agronomic species are useful to study the redox metabolism for plant acclimation.

The study of plants subjected to extreme environments undoubtedly demonstrates a reorchestration of plant metabolism for primary, secondary and redox pathways. However, the metabolic data available so far only provide a fragmentary knowledge and a more holistic, global overview of plant metabolome would benefit our understanding of plant acclimation to harsh climates. This could be addressed through untargeted metabolomics approaches to encompass a greater diversity of plant compounds, more specifically for specialised metabolites and redox compounds, which are likely to be central to mitigate stress. More importantly, it is now crucial to define the steps that should be further addressed to move from the description of metabolic features towards (i) the identification of metabolic divergences that could result from species-specific adaptation, and (ii) the convergences between plant species and between ecosystems to pinpoint generic mechanisms underpinning adaptive strategies.

II.2. Challenges and perspectives

The discovery of the main adaptive mechanisms underlying this metabolic reorchestration is the real challenge of research in harsh ecosystems. However, a global change in analytical strategy is required to achieve this objective (Fig. I.15). The first challenge will therefore be to use systems biology approaches by which the metabolic adaptations of extremophiles can be contextualised (Fig. I.15). Because models are limited to transform input variables into output variables, it will be particularly important to deal with the most interesting models, especially when it comes to defining performance. Hence, the concept of "performance = yields" must be complemented by the ability of a plant to survive and thrive in its ecosystem (Fernandez et al., 2016).

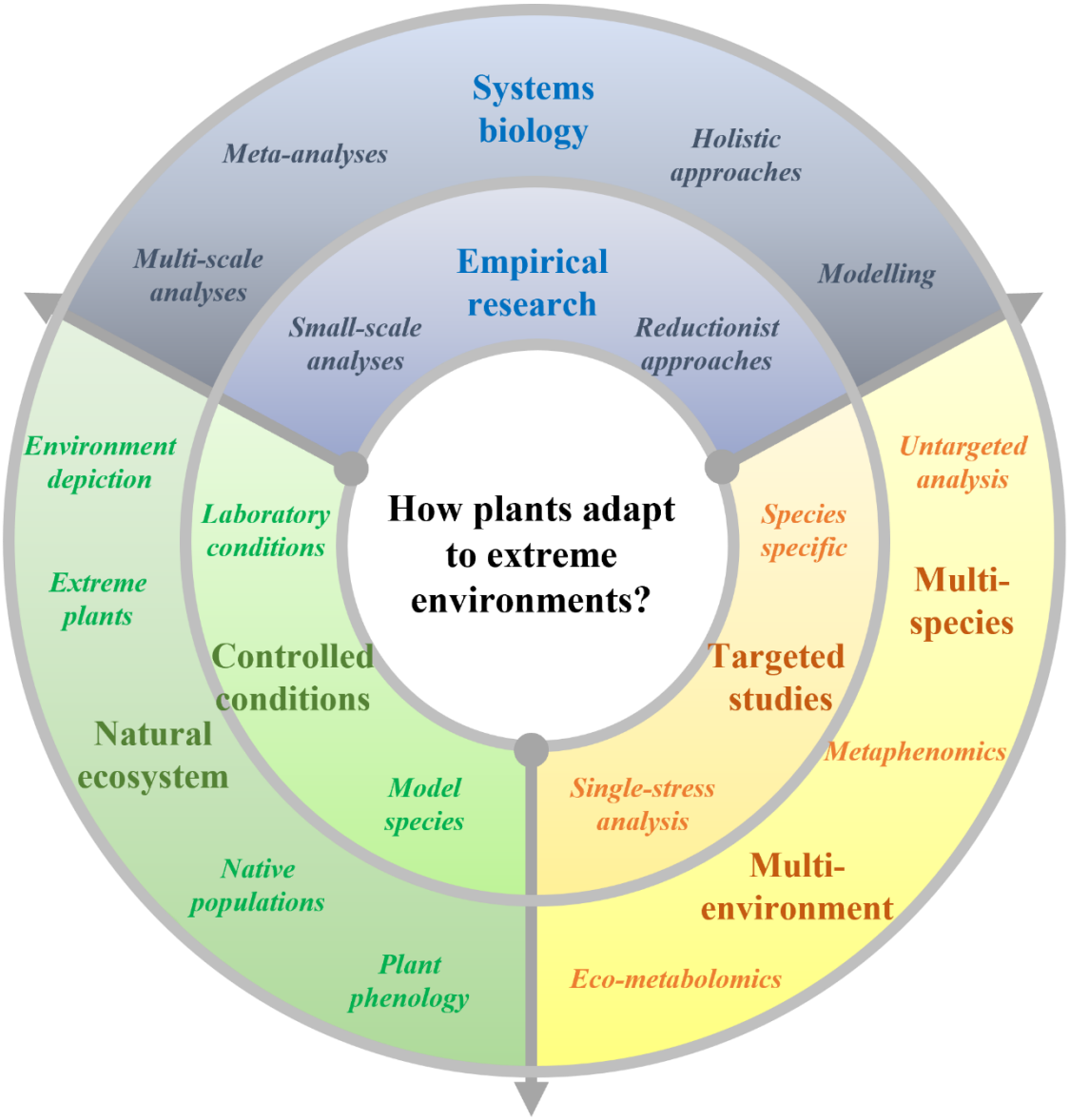

Figure I.15: Perspectives for a better understanding of plant adaptations to extreme habitats. Comprehensive exploration of the metabolome of multiple species in different environments combined with physiological and ecological data using systems biology approach can lead to new breakthroughs in the understanding of plant adaptation to extreme environments. Adapted from Dussarrat et al. 2021.

The second challenge will be to make the most relevant observations possible. Extreme biomes could be used as natural laboratories, in which the environment would be thoroughly monitored by characterising a maximum of environmental variables and by defining the analysis period (*e.g.* season, time of the day, plant organ and developmental stage). However, even if given factors could vary (*e.g.* adding fertiliser or watering), a major threat here lies in poor reproducibility of the growth conditions. Thus, in-deep knowledge of the plant environment is a prerequisite for performing ecological metabolomics (*i.e.* the analysis of the environmental impact on metabolic responses) (Díaz et al., 2016, 2019). In addition, the use of integrated and untargeted approaches appears as a requirement when studying plant adaptation at the ecological level. Exciting advances were provided when addressing these challenges. Plant responses to extreme environmental constraints were divided into temporal (*e.g.* adapted life-cycle) and structural (*e.g.* stomatal control or root phenotype adjustment) adaptations (Kumari et al., 2020). The metabolic strategies that govern the captivating capacity of plants to thrive under extreme conditions were shown to rely on a reorchestration of primary, secondary and redox metabolisms (Fig. I.13). However, the high biochemical diversity of metabolites that are abundant in extremophiles (Peters et al., 2018) including many that are still unknown (Gagneul et al., 2007; Sanchez et al., 2011), and the absence of success in engineering the accumulation of compatible solutes in crops (Turner, 2018), are still making the understanding of metabolic adaptations to extreme environments very difficult. Consequently, we now need to enlarge the coverage of the metabolome when studying extremophiles, in particular by using untargeted analytics and improving our annotation capacities (Allard et al., 2017) (Fig. I.15). Then, data should be integrated by using both unsupervised and supervised statistical approaches to highlight and then confirm hypotheses linking metabolism and adaptation (*e.g.* relations between metabolic traits and performance). These efforts will directly enrich the basis of the metabolic features from extremophile organisms that will thereafter be crucial to moving from their description to the comprehension of adaptive mechanisms harboured by both therophytes and perennial plants.

The third challenge will be to reconsider the concept of model species (Fig. I.15). Recently, the use of intra-species genetic diversity has emerged as a powerful tool to study metabolism and better understand how it participates in plant performance (Clancy et al., 2018). The so-called metaphenomics approach goes even further by researching mechanisms within panels of species (Poorter et al., 2010) (Fig. I.15). This approach has been performed to quantify combinations of phenotypic performance in different stressful environments (Wright et al., 2004; Poorter et al., 2010), and to analyse the biomass allocation in multiple species, in response to the environment (Poorter et al., 2012, 2015). Strikingly, the introduction of biological functions such as photosynthesis within these meta-analyses has provided promising results by correlating the plasticity to light intensity, plant density and plant environments (Poorter et al., 2019). Similarly, intriguing convergences were uncovered by this meta-analysis (Fig. I.13). However, only a handful of studies have studied the metabolic features of extreme plant species

using a multi-species untargeted approach (Defossez et al., 2021), and even fewer attempted to associate these features to plant adaptation.

III. GENERAL OBJECTIVES AND SPECIFIC GOALS OF THE PhD

The development of omics techniques and the integrative approach allows the analysis of the molecular mechanisms that govern the evolution of adaptive phenotypic traits. Wild species from hostile biomes lying at the edges of plant compatible gradients harbour the molecular responses to current abiotic constraints. For instance, the Atacama is the driest non-polar desert on Earth. Nevertheless, tens of plant species invaded the Atacama Desert (Díaz et al., 2019). The Talabre-Lejía transect (TLT) (lat

22°-24°S) hosts multiple plant lineages and spans an elevational cline from 2400 to 4500m.a.s.l in the Atacama Desert (Eshel et al., 2021). This transect covers three vegetations belts: the sparsely vegetated Prepuna (2400-3300m.a.s.l), the Puna (3300-4000m.a.s.l) and the high Andean Steppe (4000-4500m.a.s.l) (Eshel et al., 2021). This unique ecosystem is characterised by extremely low precipitation levels, which range from 20mm/year in the Prepuna to 160mm/year in the Steppe. Besides, critical nitrogen levels (average 9mg/kg), high solar irradiance (average 600W/m²/d) as well as daily negative temperatures and high salinity contribute to hindering plant life (Eshel et al., 2021). Interestingly, a previous study revealed 265 positively selected genes in Atacama species that cover adaptive mechanisms involved in the plant-soil interaction (*e.g.* nitrogen-fixing bacteria) and the protection against high light intensities, nitrogen starvation and osmotic stress (Eshel et al., 2021).

In parallel, biological biodiversity can palliate metabolic diversity through multi-species and multi-environment approaches, and therefore be used to discover generic markers (Dussarrat et al., 2021). For instance, results from the meta-analysis pinpointed certain evolutive convergences of biochemical pathways (*e.g.* flavonoids, amino acids…) (Dussarrat et al., 2021). Overall, these previous studies raise the following questions: (i) what is the role of metabolic processes in adaptation to extreme environments and (ii) what is their level of specialisation?

In this context, my PhD project aims to combine systems biology (and more specifically, untargeted-modelling approach) with ecology to investigate the convergent and divergent metabolic mechanisms relevant for extreme climate resilience. More precisely, this project investigates the role of metabolic processes in the adaptation of multiple Atacama plant species (Fig. I.16) to extreme habitats.

For this purpose:

- The existence and role of convergent metabolic strategies in plant resilience to extreme climates were assessed (Chapter 3). Unprecedented untargeted metabolomics analysis was performed to characterise the chemical diversity of 24 Atacama plant species. This precious dataset encompassed a quantitative analysis of major physiological markers (*e.g.* starch, organic acids, amino acids, major fatty acids) and a semi-quantitative fingerprint of both semi-polar compounds (LCMS) and fatty acyls (GCMS). Then, predictive metabolomics through GLM approach was deployed to unveil a generic toolbox for plant resilience to harsh conditions (Dussarrat et al., 2022).

- Subsequently, evolutionary trajectories of plant chemical compounds in multiple Atacama plant lineages were studied to gain insights into the genetic mechanisms that manage convergent metabolic strategies and ensure plant survival (Chapter 4). A reaction and pathway enrichment analysis used previously established transcriptomics data (Eshel et al., 2021) to compare gene family expansion and gene expression patterns between 32 Atacama plant species and their 32 closest non-adapted sequenced species. This strategy allowed the identification of genetic legacies underlying convergent biochemical strategies selected through evolution. This article will be submitted in the following weeks.

- While previous chapters focused on the discovery and characterisation of convergent strategies, additional traits and factors contribute to plant adaptation to harsh biomes. For instance, microhabitats commonly occur in extreme environments and provide favourable conditions for specific plant species (Flores and Jurado, 2003). Here, untargeted metabolic analysis, GLM and mathematical modelling were deployed to characterise the "puffer" effect of *Maihueniopsis camachoi* (Cactaceae), which enables *Atriplex imbricata* (Amaranthaceae) survival at high elevation levels (Chapter 5, submission expected for April 2022).

Besides, based on newly developed skills in metabolomics and predictive modelling, I was also given the opportunity to participate in various side projects. I used PLS regression modelling to explore the response strategy of boxwood to herbivory (col. Christiane Gallet and Anne-Emmanuelle Hay). I further analysed the metabolic features underlying the allelopathic properties of *Welwitschia mirabilis* in the Namibian desert (col. Jean Baptiste Ramond). Furthermore, I studied the impact of foliar biotic stress on the rhizosphere chemistry of *Arabidopsis thaliana* using PLS regression modelling and structural metabolic network analysis. This study revealed that foliar infection could be predicted with 80% accuracy based on the rhizochemicals. Finally, additional data were produced from the 24 Atacama plant species to deeply characterise their primary metabolism using NMR and investigate the impact of the environment on Atacama plant rhizochemicals. These data are under analysis.

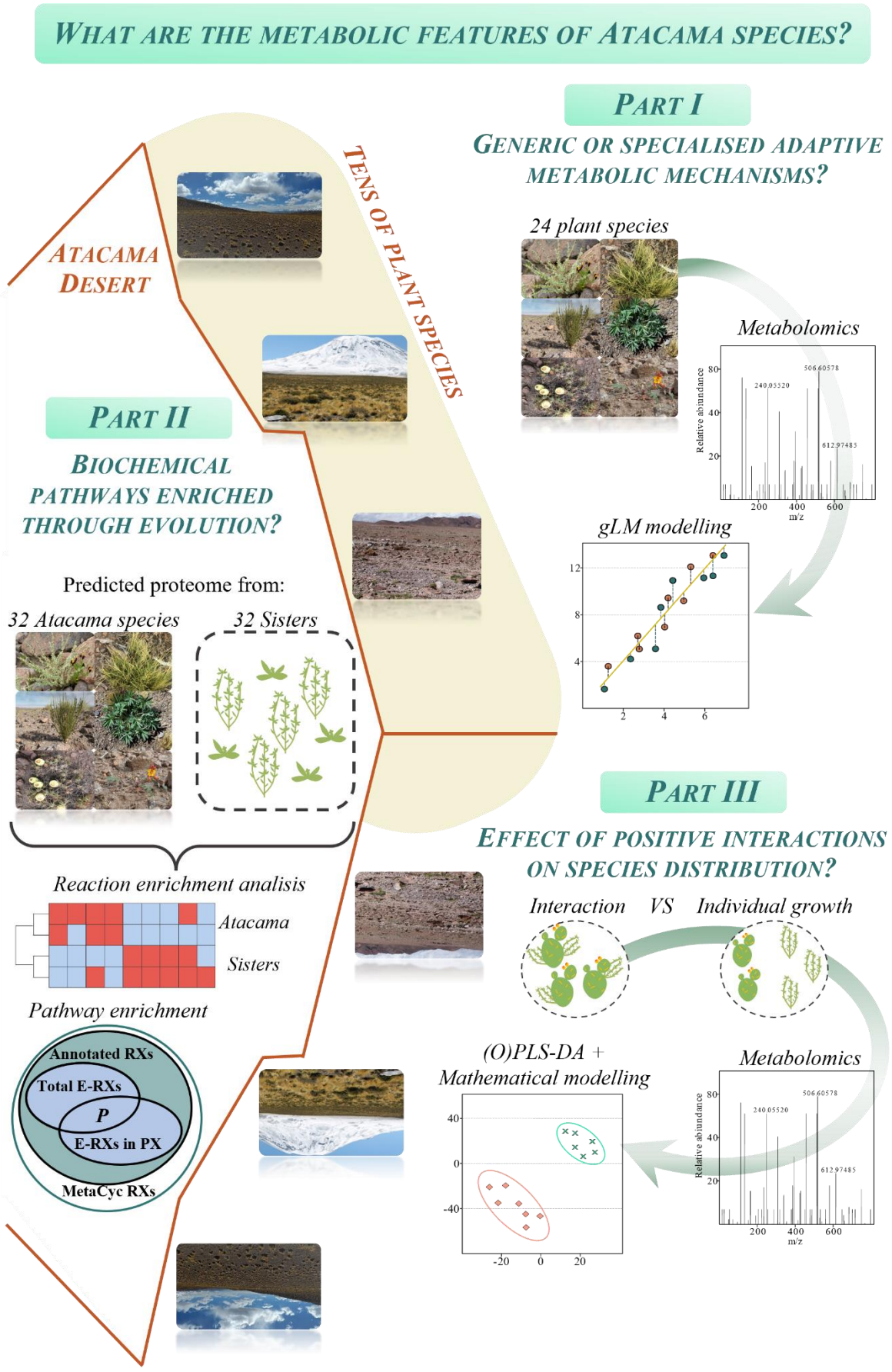

Fig. I.16 | Validation of the predictive capacity of the gLM models using a validation set.

CHAPTER 2

MATERIALS AND METHODS DEPLOYED TO UNLOCK THE SECRETS OF ATACAMA PLANTS

I. ENVIRONMENTAL DATA AND BIOLOGICAL MATERIALS

The investigation of the adaptive traits allowing Atacama plants to thrive under extreme conditions requires an in-depth knowledge of the environment and the phenotypic properties of the species present. To preserve the diversity of this fragile ecosystem, only plants with significant coverage and biomass, which was referenced for over 10 years, were collected in their natural environment.

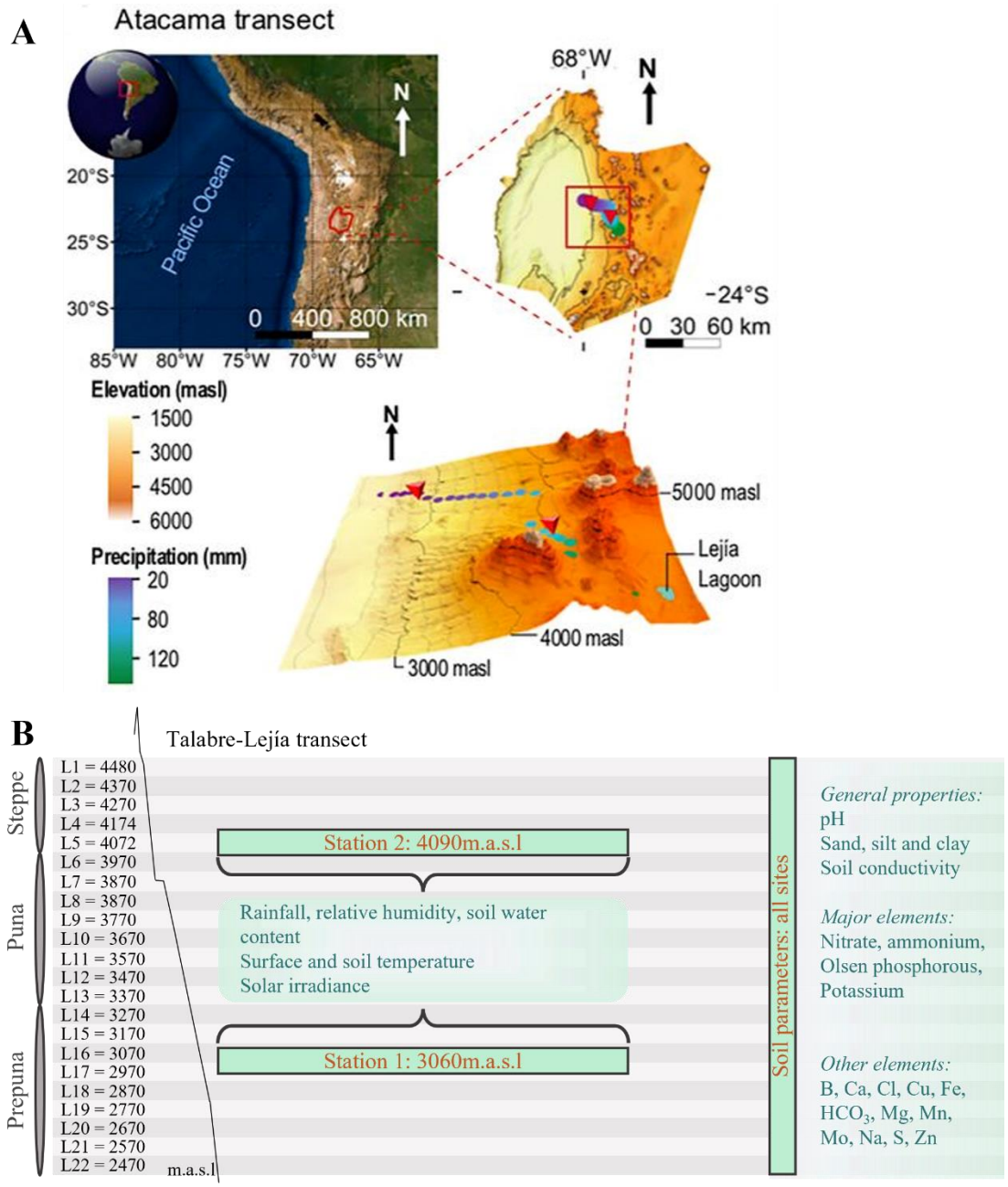

Fig. II.1 | Decomposition of the Talabre-Lejía transect and characterisation of the environment. A. Location of the elevational gradient from the Talabre-Lejía transect (TLT) (Adapted from Eshel et al., 2021). **B.** Depiction of the climatic and edaphic parameters from the TLT. *B*: boron, *Ca*: calcium, *Cl*: chlorine, *Cu*: copper, *Fe*: iron, HCO_3: bicarbonate salt, *Mg*: magnesium, *Mn*: manganese, *Mo*: molybdenum, *Na*: sodium, *S*: sulfur, *Zn*: zinc, *m.a.s.l*: meters above sea level.

I.1. Characterisation of the environment

The TLT transect area was divided into 22 sites (one level every 100m) to investigate the environmental impact on the metabolome of Atacama plant species (Fig. II.1). Two meteorological stations were installed at the Prepuna-Puna (3060m.a.s.l) and Puna-Steppe (4090m.a.s.l) junctions to record hourly environmental parameters between 2016 and 2022 including surface and soil temperature, relative humidity, rainfall, solar irradiance, soil water content (Fig. II.1). The linearity of major abiotic parameters like soil water content and temperature were tested and validated in 2019 by measuring these parameters at different elevations (Fig. II.2). Hence, theoretical values were assigned to the 22 levels by defining linear models in R (R Core Team, 2020). Soil properties included levels of nitrate, ammonium, Olsen phosphorous, potassium, copper, manganese, magnesium, zinc, iron, boron, calcium, sulfur, chlorine, sodium, molybdenum, magnesium, bicarbonate salt, sand, silt and clay as well as soil conductivity and pH were evaluated during three consecutive years in the 22 levels. Previous results presented great stability over years (Eshel et al., 2021). All values are presented in Annex II.A1.

I.2. Plant material

Plant samples from Eshel et al., 2021 collected in 2014 were complemented by three distinct environmental campaigns in April 2019, April 2021 and July 2021. Aerial parts of plants were sampled and directly snap-frozen in liquid nitrogen. Depending on the needs and availabilities, each species was collected in 1 to 6 elevation levels with a minimum of three biological replicates (Fig. II.3). Samples were then transported into dry ice until the laboratory and stored at -80°C until freeze-drying. Dried samples were ground into a fine powder and stored at -80°C until chemical extraction.

The main molecular and phenotypical properties of the 35 studied Atacama species including species and family name, life form, lifespan, endemism and carbon fixation system is provided in Annex II.A2. Additionally, a set of 11 agronomic and ornamental plant species were grown in multiple conditions in France and collected following the same protocol. These species included: *Beta vulgaris*, *Capsicum annuum*, *Helianthus annuus*, *Nicotiana tabacum*, *Phaseolus vulgaris*, *Pisum sativum*, *Portulaca oleracea*, *Solanum lycopersicum*, *Spinacia oleracea*, *Vicia faba* and *Zea mays*, which covers 5 of the 14 Atacama plant families: *Amaranthaceae*, *Asteraceae*, *Fabaceae*, *Poaceae* and *Solanaceae*.

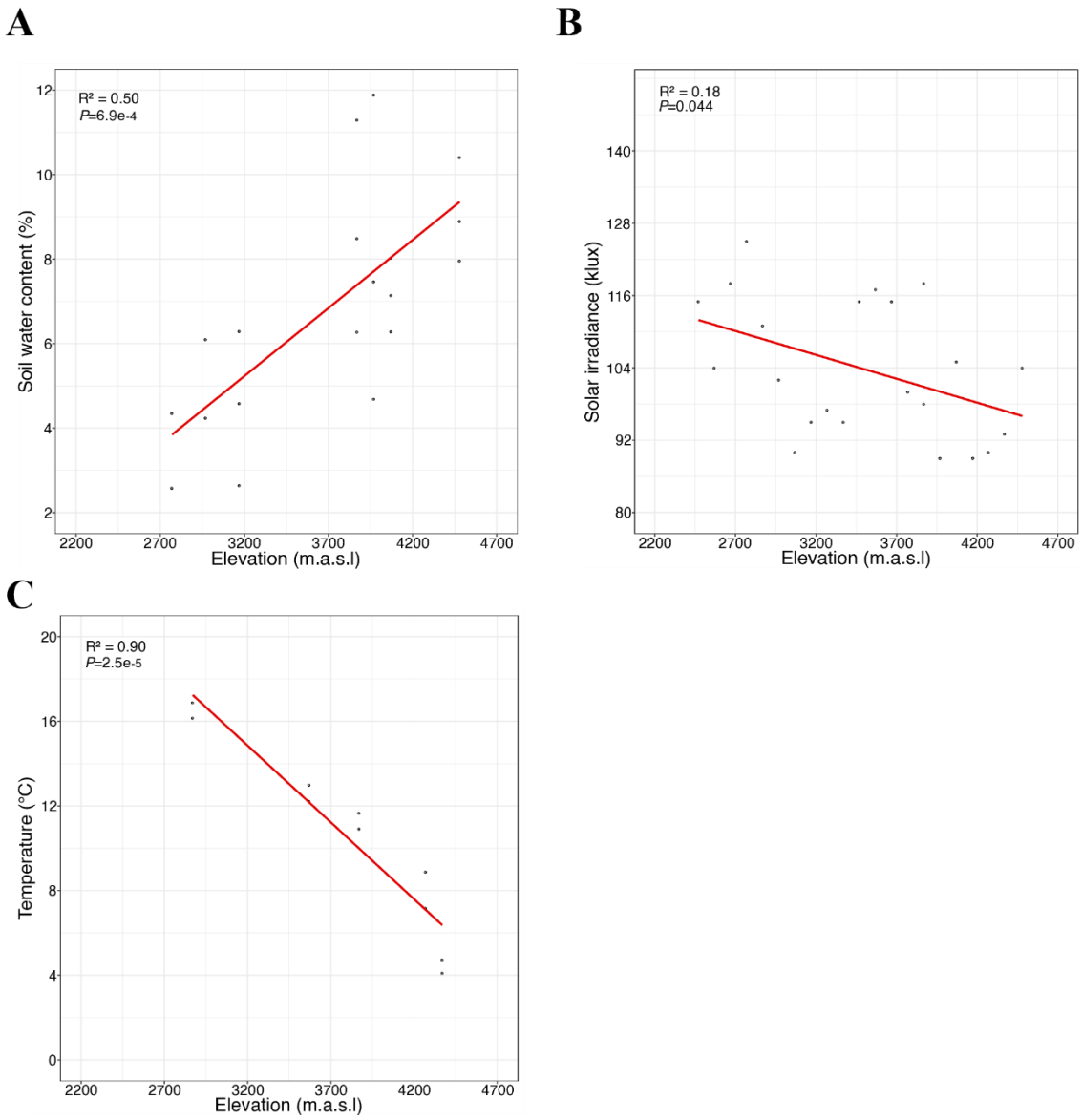

Fig. II.2 | Validation of the environmental prediction. Distribution of (**A**) soil water content, (**B**) solar irradiance and (**C**) temperature along the elevation gradient (Pearson correlation, $P<0.05$). Soil samples were collected near the plants of interest. The moisture content of the soil was determined by dividing the fresh weight by the dry weight after one week of freeze drying. Solar irradiance was measured using a lux meter near the plants of interest. Temperature was measured using two thermometers. Figures from Dussarrat et al., 2022.

II. FIELDWORK ANALYSES

Plant coverage across the TLT has been measured every year after the rainy season (March or April) since 2011 by Claudio Latorre, Francisca Díaz and Rodrigo Gutiérrez. To get insights into plant-plant interaction mechanisms, we performed additional coverage measurements in July 2014, 2015 and 2021. More precisely, 20 species were found interacting with the cactus *M. camachoi* in at least one elevation level over the years (Fig. II.3). Hence, these measurements aimed to characterise the facilitation intensity (*i.e.* the number of plants growing inside minus the number of plants growing outside the cushion). Twenty replicates were performed per elevation level. The diameter of these twenty independent cacti was reported and associated with the total diameter (cm) of the connected species (Table V.S1). For comparison, the coverage (cm) of the different species developing without cactus was measured for a surface equal to the diameter of the cactus for each replicate. The coverage of the Atacama species "with" and "without" *M. camachoi* was thus expressed in centimetres by meter squared of *M. camachoi* and defined in the different sites (Table V.S1). In parallel, three to four thermometers were installed at 3000, 3400 and 3800m.a.s.l to measure ground temperature for at least 24 hours in August 2016 and 2021. The resulting 2021 dataset combined coverage and temperature measurements performed in 2021 while the 2016 dataset was developed using 2016 temperature reports and coverage measured in 2014 and 2015. The 2016 dataset was used to validate observations and analyses conducted on the 2021 dataset (Chapter 5).

III. EXTRACTION

Extraction methods are based on the chemical properties of the target molecules. The use of polar or apolar solvents is therefore destined for the analysis of different molecules. A non-targeted approach requires access to the chemical biodiversity of the biological system studied. Hence, different chemical extraction methods were used to extract major physiological markers, explore semi-polar secondary metabolism and investigate fatty acyl profiles.

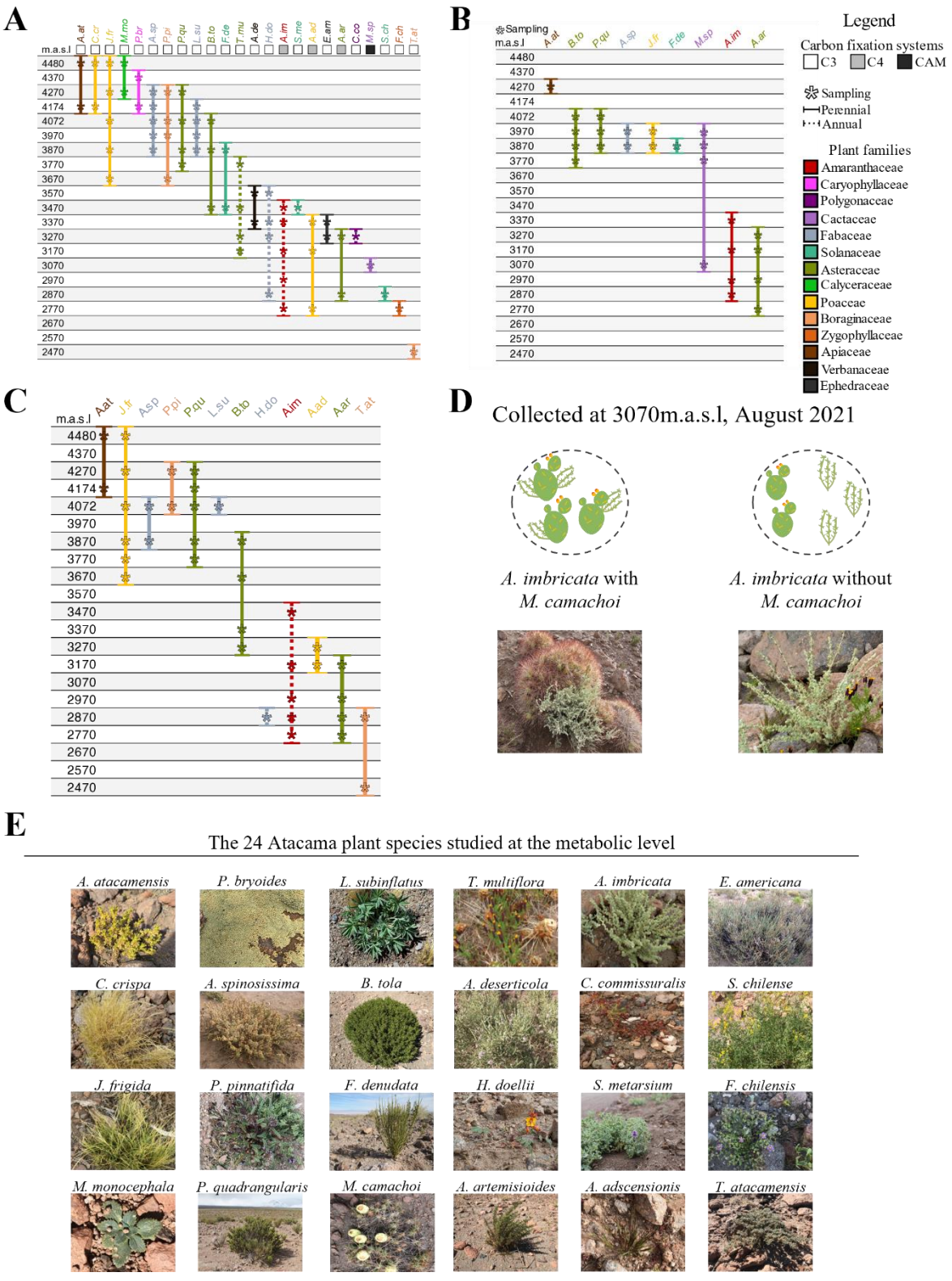

Fig. II.3 | Plant material for the different experiments. Results of sampling performed in April 2019 (**A**), April 2014 (**B**), April 2021 (**C**) and July 2021 (**D**). **E.** Pictures of the 24 plants species studied at the metabolic level.

III.1. Preliminary test and ethanolic extraction

Four distinct extraction methods were tested to analyse plant secondary metabolism via liquid chromatography-mass spectrometry (LCMS). Ethanolic, methanolic and water extractions were tested before LCMS analysis. Ethanolic extraction was used for its capacity to extract both major primary markers like organic acids and hexoses and semi-polar compounds (Luna et al., 2020). Hence, a solution of HEPES/KOH 10mM at pH 6 with 80% ethanol solution (EtOH: water, 80:20, v/v) was added to the 20mg of dry powder and heated at 80°C for 20 minutes as previously described (Luna et al., 2020). Samples were then centrifugated for 5 minutes at 14000 rotations per minute (rpm) and the supernatant (S1) was collected. The extraction process was repeated twice (once using EtOH 80% and once using EtOH 50%) and supernatants S2 and S3 were added to S1. Extracted samples, as well as pellets, were then stored at -20°C until robotised biochemical assays. However, preliminary assays highlighted high levels of polyphenols that disturbed the measurement of organic acids and soluble sugars. Thus, the same extraction method was applied on another microplate containing 20mg of samples plus 20mg of polyvinylpolypyrrolidone (PVPP) which adsorbs the polyphenols. Subsequent metabolic analyses were performed on the first or second set of ethanolic extracts (*i.e.* with or without PVPP) depending on the objective.

A distinct ethanolic extraction protocol was performed for LCMS analysis. First, the extraction solvent was 80% ethanol with 0.1% formic acid. Methyl vanillate (250 µg/mL) was used as a chemical standard to evaluate the quality of the LCMS injections. Besides, 300 µL of the extraction solution was added to each sample and the heating extraction step was replaced with ice bath sonication extraction for 15 minutes. The extraction process was repeated (EtOH 80%) and supernatants 1 and 2 were pooled. Finally, filtrations were performed using 96 wells 0.22µM Millipore plates (reference MSGVS2210). Filtered solutions were stored at -80°C until LCMS analysis (Luna et al., 2020).

Three other extraction methods were tested to evaluate their ability to cover the secondary metabolism of 24 plant species from the Atacama. Two methanolic extractions (methanol 70% or 95%) were realised as well as one hot water extraction method. In brief, a methanolic solution of 70% (Methanol: water 70:30, v/v) or 95% (95:5, v/v) was added to 20mg of dry powder and sonicated for 15 minutes in an ice bath. Extracted samples were then centrifugated at 14,000 rpm for 5 minutes and filtered using 0,2 µm polyvinylidene fluoride (PVDF) filters. Filtered supernatants were collected and stored at -80°C until LCMS analysis. Hot water extraction was performed as previously described (Cocuron and Alonso, 2014). In brief, 1mL of hot water (100°C) was added to 20mg of dry powder. Samples were then placed into a water bath at 100°C for 5 minutes, vortexed and replaced into the water bath for 5 additional minutes. After centrifugation (5 minutes at 14,000 rpm), samples were filtered using 0,2 µm PVDF filters and stored at -80°C until LCMS analysis.

Raw LCMS data from the extraction test were processed via XCMS (v 4.2, an acronym for various forms (X) of chromatography mass spectrometry) in R (v 3.6.1) using a noise threshold of 30,000 (as explained in sections V and VI). Results illustrated a similar coverage capacity (Fig. II.4) in positive and negative electrospray ionisation (ESI) mode. Besides, only 1% of the detected ions using methanolic or water extraction were not observed when using ethanolic extraction. Hence, ethanolic extraction was selected for further LCMS analysis due to its high coverage, safety and usability compared to other techniques.

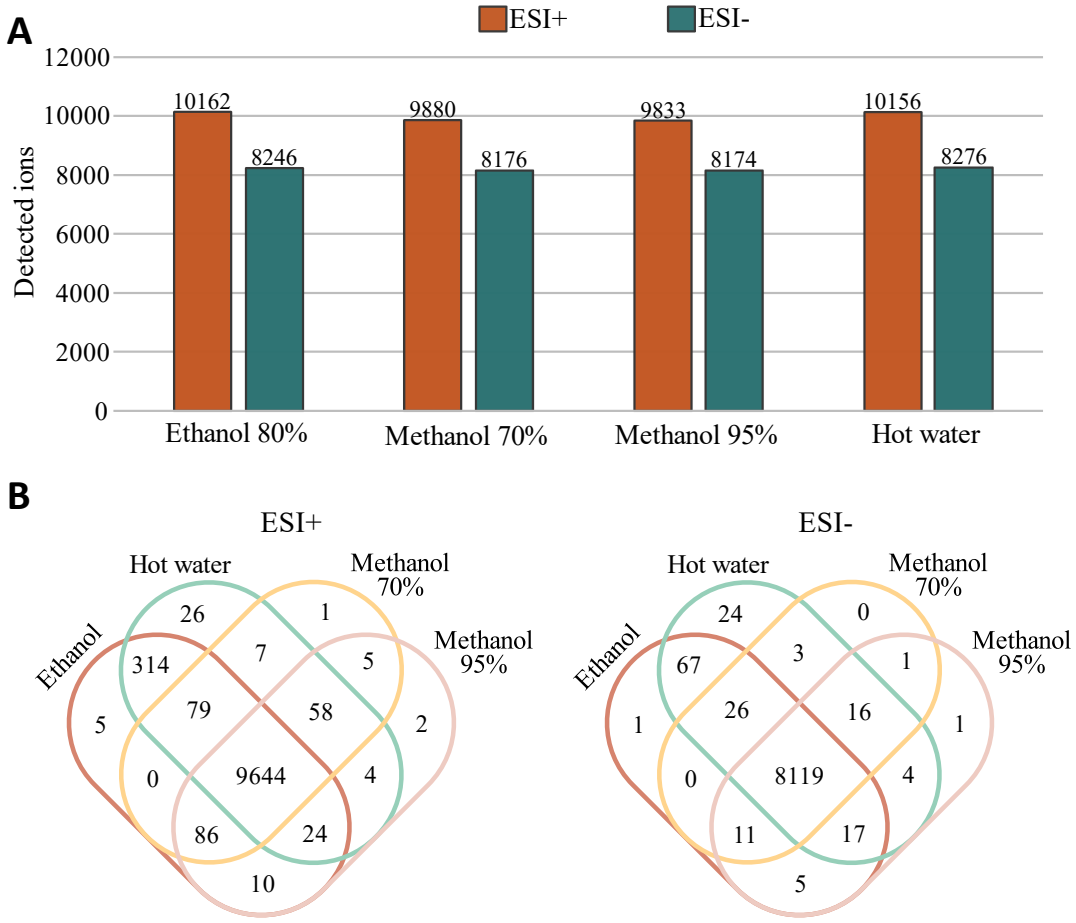

Fig. II.4 | Detected ions by LCMS analysis following ethanolic, methanolic or water extraction. **A.** Number of detected ions using a noise threshold of 30,000. **B.** Depiction of the detected ions using the different extraction methods.

III.2. Fatty acyl extraction

Hexane was used as an apolar solvent to extract fatty acyls from the 24 Atacama plant species as previously described (Domergue et al., 2010). Shortly, this protocol allows the extraction of fatty acyls from glycerolipids (including phospholipids, galactolipids and triglycerides) and sphingolipids which are hydrolysed and trans-esterified into fatty acid methyl esters (FAMEs) using a transesterification solution. This solution was applied on 10mg of dry powder and was composed of methanol and 5% of sulfuric acid (H_2SO_4). When analysing plant extracts, the C17:0 fatty acid is used as a chemical standard (20 µg/ml in the transesterification solution). The hydrolysis of the chemical bond (ester or amine bond for lipids and sphingolipids respectively) occurred when adding 1mL of the transesterification solution per sample. Samples are then kept at 80-85°C for 3 hours. After cooling, 1mL of NaCl 2.5% and 400µL of hexane 99% were added to the solution to extract fatty acyls. After agitation and centrifugation, the upper phase is collected in GC vials and stored at -20°C (Domergue et al., 2010).

IV. TARGETED BIOCHEMICAL PHENOTYPING

Robotised microplate assays were performed to quantitively evaluate levels of major physiological markers. The ethanolic extraction led to the separation of a soluble (supernatant) and an insoluble fraction (pellet). The two distinct fractions include different compounds that were measured using various robotised microplate assays as previously described (Roch et al., 2020).

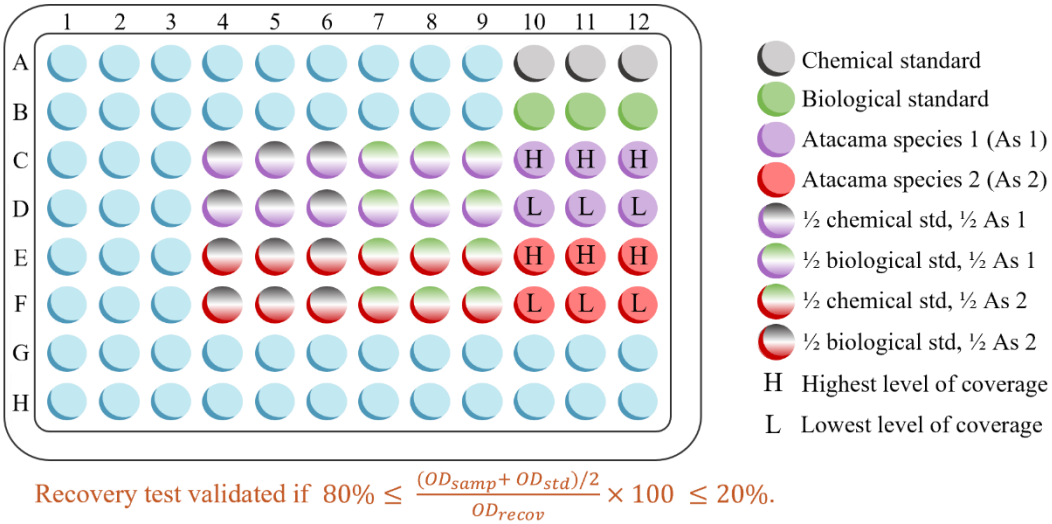

Fig. II.5 | Recovery test. Chemical standard varies depending on the chemical compounds studied (*e.g.* mix of glucose, fructose, sucrose when analysing sugar content). The variety M82 of *S. lycopersicum* was used as a biological standard. A threshold of 20% (*i.e.* 80%≤Δ recovery≤120%) was used to accept or reject the recovery test for each assay.

IV.1. Recovery test

Recovery tests were performed to optimise and control the quality of the measurements (Fig. II.5). Briefly, one chemical and one biological control were used to realise the recovery test. Since each of the 24 Atacama plant species was collected at distinct elevation levels, one sample from the highest and one sample from the lowest elevation level were subjected to the recovery test. For the different assays, the optical density (OD) was measured for XµL of chemical or biological standard (OD_{std}), XµL of sample extract (OD_{samp}) and for the mix of X/2µL of sample extract plus X/2µL of chemical or biological standard (OD_{recov}) (Fig. II.5). Then, the recovery test consists in comparing the mean OD between OD_{std} and OD_{samp} with OD_{recov} (Fig. II.5). Recovery test was validated if $80\% \leq \frac{(OD_{samp}+OD_{std})/2}{OD_{recov}} \times 100 \leq 120\%$. The complete list of dilutions and volumes used for the different species and assays is displayed in Annex II.A3. All OD measurements were performed on a 96-well plate reader spectrophotometer (SAFAS MP96). Results were then normalised by the sample weights used for the ethanolic extraction (20mg ± 1) before the corresponding assay to obtain a concentration (µmol) of amount (µg) of compound per unit of dry mass (*e.g.* µmol/g of dry weight).

IV.2. Soluble fraction

Following ethanolic extraction, the levels of chlorophyll, major soluble sugars (glucose, fructose, sucrose), organic acids (malate, fumarate, citrate), nitrates, total free amino acids and total polyphenols were measured using the soluble fraction (*i.e.* supernatant).

Measurement of chlorophyll levels

Chlorophyll content was determined directly after ethanolic extraction by adding 120µL of ethanol 98% to the corresponding volume of sample extract (Annex II.A3) and reading the OD at 645 and 665nm. Then, levels of chlorophyll a and b were estimated using the following formulas: chlorophyll a (µg/well) = 5.21 OD_{665} - 2.07 OD_{645} and chlorophyll b (µg/well) = 9.29 OD_{645} - 2.74 OD_{665} (Arnon, 1949).

Determination of major soluble sugar levels

As previously described (Stitt et al., 1989; Roch et al., 2020), levels of major soluble sugars like glucose, fructose and sucrose were accessed using successive addition of hexokinase (HK), phosphoglucose isomerase (PGI) and invertase (INV). In brief, 160 µL of a solution of 0.1M of HEPES/KOH pH 7 with 3mM of MgCl2, 3µM of ATP, 1.4µM of NADP and 3.4u/mL of glucose-6-phosphate dehydrogenase was added to the optimised volume of ethanolic extract (Annex II.A3). After a first measurement of the blank level at 340nm (at which NADPH is detected), 1µL of HK (900u/mL) was added to the mixture and the OD at 340nm was measured again. Similarly, 1µL of PGI (1000 u/mL) and 1µL of INV (30 000 u/mL) were successively added to evaluate the levels of fructose and sucrose respectively (Fig. II.6). Levels of metabolites are proportional to the produced NADPH. The concentration of NADPH at each step is evaluated using the following equation: NADPH (µmol) = $\frac{\Delta OD_{340nm}}{l \times \varepsilon}$ where l represents the optical path length in a well (equal to 2850 cm.L^{-1}) and where ϵ represents the extinction coefficient of NADPH (6.22 L.mol^{-1}.cm^{-1}).

Determination of malate and fumarate levels

The levels of malate and fumarate are measured through the use of malate dehydrogenase and fumarase which catalyse the transformation of malate to oxaloacetate and fumarate to malate, respectively (Fig. II.6). Since the metabolisation of malate into oxaloacetate uses NAD^+ as a cofactor, the amount of fumarate and malate are estimated through the production of NADH (detected at 340nm). Briefly, 90µL of the reactional solution that included 69µL of tricine/KOH 0.1M at pH 9, 10µL NAD^+ 30mM, 10µL of glutamate 20mM and 1µL of glutamate-oxoglutarate aminotransferase (GOT) at 200u/mL in 100mM tricine/KOH at pH 9 was added to the corresponding volume of ethanolic extract. Then, 2µL of MDH and 1µL of fumarase were successively added and the OD at 340nm was measured between each analytical step (blank, after addition of MDH and after addition of fumarase).

The difference in intensity between the different stages (*e.g.* OD after addition of MDH minus OD from blank) was used to determine the levels of malate and fumarate.

Citrate assay

Citrate levels were determined using the Citric Acid Assay Kit (K-CITR, Megazyme). The reaction catalysed by the citrate lyase metabolises citrate into oxaloacetate and acetate. Oxaloacetate is then transformed into malate by the L-malate dehydrogenase (MDH). Notably, some oxaloacetate can be converted into pyruvate if some oxaloacetate decarboxylase is present in the sample. Hence, the Citric Acid Assay Kit includes the D-lactate dehydrogenase (LDH) to transform the newly produced pyruvate into D-lactate. LDH and MDH employ NADH as a cofactor for the reaction and the quantity of citrate is therefore deduced from the loss of this compound (Fig. II.6).

Measurement of nitrate level

Nitrate levels were defined using the nitrate reductase as previously described (Mori, 2000). For a defined volume of ethanolic extract (Annex II.A3), 10µL of potassium phosphate buffer 1M at pH7.5, 1µL of NADPH 50mM, 1µL of nitrate reductase 5U/mL, and 83 µL of water were added. The resulting solution was mixed and conserved at room temperature for 30 minutes in dark conditions. Then, 15µL of phenazine methosulfate 0.25mM was added to the solution, which was then mixed and conserved at room temperature for 20 minutes in dark conditions. Subsequently, 60µL of sulphanilamide (1% w/v, diluted in phosphoric acid 3M) and 60µL of N(1-Naphtyl)ethylenediamine dihydrochloride 0.02% w/v were added to the solution. The OD was measured at 540nm after 10 minutes at room temperature in dark conditions. A blank was realised (without nitrate reductase) to access the nitrite amount in samples. Then, nitrate levels were obtained by subtracting the OD of the blank for each sample and using the calibration curve of KNO_3 from 0 to 1.6mM.

Determination of the total free amino acid levels

The total level of free amino acids was measured via the fluorescamine method previously described (Bantan-Polak et al., 2001). Shortly, 115 µL of borate buffer which included 15µL of sodium borate buffer 0.1M at pH8 with 100µL of water was added to the corresponding volume of ethanolic extract (Annex II.A3). Then, 90µL of fluorescamine 0.1% in acetonitrile was added to the mixture and the resulting solution was kept at room temperature for 5 minutes. The fluorescence was determined with an excitation at 405nm and emission at 485nm. Levels of total free amino acids were deduced using a calibration curve of glutamate from 0 to 1000µM.

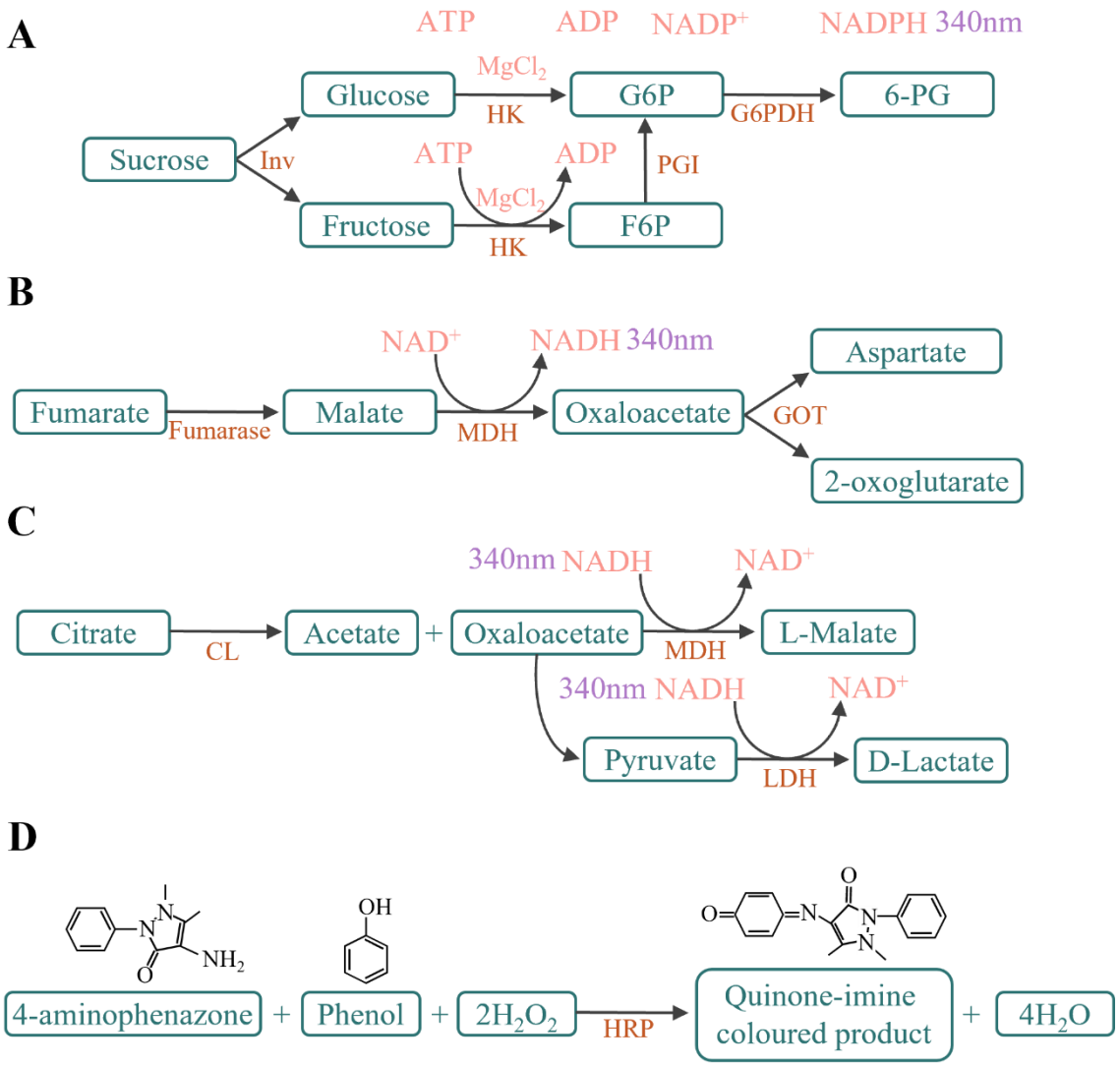

Fig. II.6 | Robotised microplate assays. A. Determination of glucose, fructose and sucrose levels. **B.** Determination of malate and fumarate levels. **C.** Citrate assay. **D.** Determination of total polyphenol levels. *CL*: citrate lyase, *F6P*: fructose-6-Phosphate, *G6PDH*: glucose-6-phosphate dehydrogenase, *G6P*: glucose-6-phosphate, *GOT*: glutamate-oxoglutarate aminotransferase, *HK*: hexokinase, *HRP*: horseradish peroxidase, *Inv*: invertase, *LDH*: lactate dehydrogenase, *MDH*: malate dehydrogenase; *PGI*: phosphoglucose isomerase, *6PG*: 6-phosphogluconolactone.

Determination of the total polyphenol levels

Polyphenol levels were estimated using the horseradish peroxidase (HRP) as previously described (Stevanato et al., 2004). Shortly, 71µL of water, 10µL of tricine/KOH 0.1M at pH 8, 3µL of H_2O_2 100mM and 6µL of 4-aminoantipyrine 200mM are added to the corresponding volume of ethanolic extract (without PVPP). The OD was measured at 405nm to define the baseline and the mixture was supplemented with 1µL of HRP to generate a coloured compound that is proportional to the phenol content (based on the saturation of H_2O_2 and 4-aminoantipyrine) (Fig. II.6). Polyphenol content is then determined using the calibration curve of pyrogallic acid from 0 to 8g/L.

IV.3. Insoluble fraction

Following ethanolic extraction, the levels of starch and the quantity of total soluble proteins were measured using the insoluble fraction (*i.e.* pellet). Besides, residual compounds contained in the pellet were weighed to estimate the cell wall content.

Measurement of the total amount of soluble proteins

The content of soluble proteins was measured via the Bradford reagent as previously described (Bradford, 1976). Soluble proteins by adding 400µL of NaOH 0.1M in the pellet. The solution was then heated at 95°C for 30 minutes and centrifugated at 2500 rpm for 5 minutes. Then, 180µL was added to the corresponding volume of ethanolic extract (Annex II.A3). After 5 minutes at room temperature, the OD was measured at 595nm and the soluble protein content was estimated using the calibration curve of bovine serum albumin (BSA).

Starch assay

Starch content was determined following a previously established process (Hendriks et al., 2003). Following the protein assay, 80µL of HCl 0.5M, acetate/NaOH 0.1M at pH 4.9 was added to the pellet (which has already been supplemented with 400µL of NaOH during the protein assay). After controlling the pH, 100µL of a degradation mix composed of amyloglucosidase (1.5mL) and α-amylase (15µL) were added. The resulting solution was mixed and incubated at 37°C for 16 hours. Then the starch content was assessed using the same protocol as the glucose assay.

Estimation of the cell wall content

Following the previously described assays, cell wall content was estimated after two washing steps. Samples (including NaOH and HCl solutions) were centrifugated at 2500rpm for 10 minutes. The supernatant was discarded and 250µL of NaOH 0.5M was added to the pellet. The resulting mixture was mixed and heated at 95°C for 20 minutes and centrifugated again. A second washing step was performed with 250µL of NaOH 0.5M. The washed pellets were then freeze-dried overnight and the cell wall content was represented by the delta in mass between the sample tube with and without the dried pellet.

V. METABOLOMICS ANALYSES

Robotized microplate assays were complemented with the investigation of fatty acyl and secondary compounds using various techniques. The high sensitivity of gas chromatography coupled to a flame ionisation detector (GCFID) or a mass spectrometric detector (GCMS) as well as LCMS gave access to the chemical diversity of the Atacama plant species.

V.1. Analysis of fatty acyls via GCFID and GCMS

Two analytical techniques were used to explore the lipid profile of multiple plant species from the Atacama Desert. First, a quantitative estimation of the FAME content was performed using GC (Hewlett-Packard 5890 series II) coupled to an FID and a capillary column HP-5MS of 30mX0.25mm as previously described (Domergue et al., 2010). A constant flux of helium was used as a carrier gas. A temperature gradient was used to optimise the separation of the FAMEs where 50°C was maintained during the first minute and followed by a gradual increase of 25°C per minute until 150°C. This temperature was maintained for 2 minutes and increased again with a rate of 10°C per minute to 320°C, which was held for 6 minutes. The temperature of the injector and detector was 250°C. The quantification of the FAMEs within Atacama samples was performed using the internal standard C17:0. Apolar extracts were subjected to untargeted fatty acyl profiling using a GC (Agilent 6850) equipped with a 30mX0.25mm HP-5MS column coupled to a mass spectrometer (Agilent 5975; 70 eV; m/z ratio from 50 to 750 Da) with a flux of 1.5mL/min of helium as carrier gas (Domergue et al., 2010). The same temperature gradient was used.

V.2. Analysis of semi-polar compounds via LCMS

Ethanolic extracts were subjected to untargeted analysis via an Ultimate 3000 ultra-high-pressure liquid chromatography (UHPLC) system coupled to a LTQ-Orbitrap Elite mass spectrometer as previously described (Luna et al., 2020). Samples were ionised via electrospray ionisation source (ESI) in positive or negative modes. A C18 column (C18-Gemini, 2.0x150mm, 3 µm, 110Å, Phenomenex, USA) was employed to separate semi-polar compounds before mass spectra acquisition. A gradient of 18 minutes with a flow of 0.350ml/min composed of 3% of solvent B (acetonitrile LCMS grade) and 97% of solvent A (water with 0.1% of formic acid) between 0 to 0.5 minutes, 10% of solvent B at T=1min, 50% at T=9min, 100% at T=13 minutes, and 3% from T=14.5min to the end. Besides, injection parameters were optimised as follows: draw and dispense speed of 600nl/s, draw and dispense delay of 30ms, waste and wash speed of 8 µl/s. Full scan analysis was realised at 240k resolution power at m/z = 200 Da. Extraction blanks were used to control the quality of the ethanolic extraction and quality control samples (QC, which is a mix of all injected samples) were injected every 10 samples to control the quality of the analysis. Similarly, one biological sample was injected every 30 samples to evaluate the repeatability of the analysis. Subsequently, MS/MS spectra were acquired in high energy collisional dissociation (HCD) mode at a normalised collision energy of 25% and 45% or 50% and 70% in positive and negative ionisation mode, respectively.

VI. PROCESSING OF THE METABOLIC DATA

The processing step enables the extraction, integration, organisation and correction of complex sets of informatics data produced via GCFID or GC/LCMS called raw data. This step is therefore essential and creates a data matrix that can be easily manipulated (*e.g.* csv table). Various processing strategies were deployed due to the high variety of analytical techniques used in this project.

VI.1. GCFID and GCMS spectra

Processing of GCFID and GCMS spectra was performed using the Agilent2 method (adapted from the Agilent method), which automatically performed the peak picking step. The intensity threshold was defined as 1% of the major peak. Results of this method were controlled and corrected manually. The annotation of the different peaks was carried out using an internal library.

The quantity of FAMEs was calculated using the formula: FAME (µg/g dry weight) = $\frac{(A_{samp} - A_{blk}).(M_{CS})}{A_{CS}} \times \frac{1}{M_{samp}}$ where A_{samp} and A_{blk} represent the area of the corresponding peak, M_{CS} and M_{samp} represent the quantity of chemical standard (µg) or dry sample powder (g) used for the analysis.

VI.2. LCMS spectra

Raw LCMS data can be processed through various software that have been developed recently. Here, we focus on the two processes that were used during the PhD project.

XCMS in R

Firstly, the raw LCMS data were converted into an mzML format compatible with the XCMS software. Then, XCMS software performs the following processing steps: peak picking (which allows the detection of the peaks), grouping (grouping the peaks according to their m/z ratio), RT correction (where the potential drift of retention time during the LCMS run is corrected), gap-filling (*i.e.* integrating baseline values for missing peaks) (Smith et al., 2006). Here, optimised parameters were used to process raw LCMS data as previously described (Luna et al., 2020). These parameters included: a centWave method, a m/z window of 0.01 (based on the resolution of the Orbitrap), a noise threshold that varied between 20,000 to 100,000 depending on the analysis and a signal on noise ratio threshold of 4. Finally, the XCMS software provides an Excel table summarising the intensity of the different ions for each sample (including biological samples, blanks and QC).

MS-DIAL

MS-DIAL is a user-friendly open-source software pipeline available allowing the processing of raw GC or LCMS data, which can be used without prior transformation (Tsugawa et al., 2015). Similar settings to those used on XCMS software were used to process the raw LCMS data and included: a MS1 tolerance of 0.01, a MS2 tolerance of 0.025 and a minimum peak height of 20,000.

Independently of the processing software, additional steps were performed to select the most stable and meaningful ions across the processed dataset (Fig. II.7). An example is provided with the analysis performed in Dussarrat et al., 2022. The total number of ions detected shows the number of ions that passed the processing step. The delta between the minimum and maximum RT and m/z should not exceed 60s and 0.01Da, respectively. Besides, ions should not be present in the extraction blanks (BEX). Finally, the coefficient of variation of an ion in the different injected QCs (one every 10 samples) must not exceed 30% to ensure satisfactory stability (Fig. II.7).

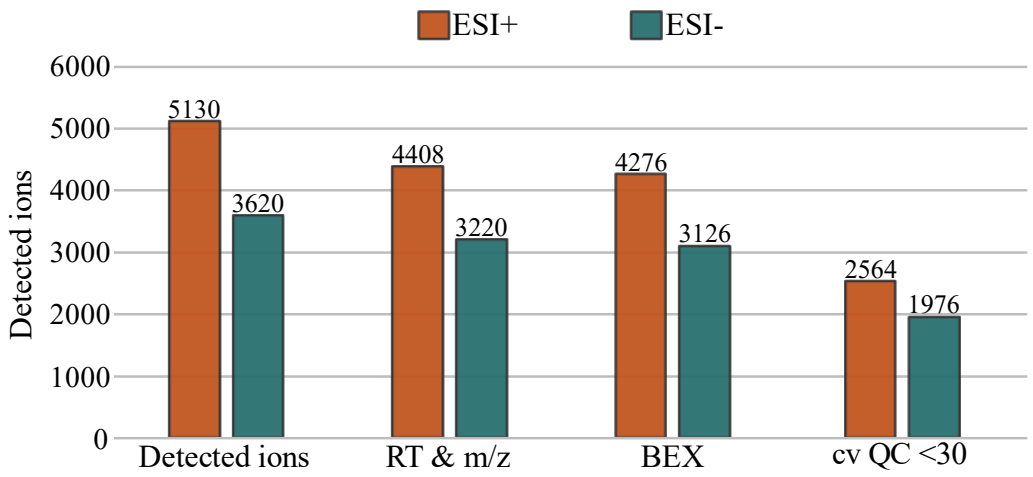

Fig. II.7 | Pre-processing of LCMS raw data. Detected ions represent the amount of ions detected using a noise threshold of 100,000. Total ions were then filtered by removing ions with a retention time (RT) delta (*i.e.* RTmax – RTmin) than 60s or delta m/z higher than 0.025 Da. Then, ions detected in extraction blanks (BEX) are removed, as well as ions with a coefficient of variation higher than 30% in quality control samples (QCs).

VII. STATISTICAL ANALYSES

Mathematics permits hundreds of variables (*e.g.* ions) to be comparable between hundreds of samples through, for instance, normalisation. In addition, statistics enable the extraction of the explanatory variables from a complex and abstract set of data. Here, a plethora of statistics was used to (i) meet the different constraints induced by an untargeted analysis on multiple unknown plant systems and to (ii) extract the explainable variables relevant to the biological question from tens of thousands of metabolic variables produced by the different experiments. These statistics can be divided into univariate, multivariate (including exploratory and regression analyses) and modelling approaches. All datasets were normalised by median normalisation, cube-root transformation and Pareto scaling using MetaboAnalyst (https://www.metaboanalyst.ca) as recommended for untargeted metabolomics studies (Xia et al., 2015; Di Guida et al., 2016).

VII.1. Univariate and multivariate analyses

Univariate statistics were deployed either to reduce the complexity of the processed dataset before applying multivariate and multivariate analyses or to confirm the relationship of the explanatory variables with the phenotypic trait studied. ANOVA or t-test were realised using the agricolae package (Mendiburu, 2021) on R or directly on MetaboAnalyst to select the metabolic compounds (*e.g.* ions)

significantly linked to the response factor studied (*i.e.* temperature parameter). This approach allowed either to screen the metabolic dataset and clean up variables that did not respond to the factor studied or to validate the potential of the discovered explanatory variables. Besides, various correlation tests were realised using Hmisc and corrplot packages to explore the relationships between metabolic and environmental variables or between multiple metabolic variables (Jr, 2021; Wei and Simko, 2021).

Exploratory statistics like PCA were performed to decomplexify a highly dimensional dataset and provide preliminary hypotheses via the factoextra and FactoMineR packages (Lê et al., 2008; Kassambara and Mundt, 2020). Other exploratory statistics such as clustering approaches and visualisation through heatmaps were performed via the pheatmap and cluster packages in R (Kolde, 2019). Pearson correlation and Ward algorithm were used to cluster the samples based on the consequent number of biological replicates.

PLS, OPLS and associated discriminant analyses (*i.e.* (O)PLS-DA) were performed to unveil the predictive components when the number of samples was not sufficient to perform gLM approach. Partial least square regression (PLSr) were realised using the pls package in R with the leave-one-out cross-validation approach as previously described (Meacham-Hensold et al., 2019; Mevik et al., 2020). To evaluate the predictive capacity of the PLS model, the total sample set was divided into a training set (80%) to develop the equation which was then applied to the testing set (20%). Models were performed 50 times to cover the different combinations offered by the stratified sampling method used to define the training and testing sets. The most predictive components were extracted from the predictive components based on their coefficient within the linear equation. Finally, the predictive capacity of the best metabolic markers was tested via additional PLSr models. Discriminant analyses ((O)PLS-DA) were realised directly on MetaboAnalyst. The likelihood of spurious predictions was systematically tested by developing permuted datasets where the values of the response factor are randomly redistributed between samples. In addition to this statistical validation, a biological validation was performed by exploring the predictive capacity of the selected markers on an independent dataset. To illustrate and facilitate the interpretation of the results, various plots (including scatter plots, box plots, lollipop plots, pie charts…) were designed using ggplot package in R (Wickham, 2016) and Inkscape software (https://inkscape.org).

VII.2. Generalised multilinear models (GLM)

GLMs were used to (i) explore the correlation between metabolic variables and environmental data and (ii) extract the best predictors of the studied response factor while preserving the biological context. More specifically, we here deployed general multilinear models (gLM) (*i.e.* a generalised linear model where the distribution is normal, see Chapter I) based on the normal distribution of the data using the glmnet package available in R (Friedman et al., 2010). Stratified sampling was used to define the training set (70%) and the testing set (20%) which were used to establish the model equation and the validation set (10%) to test the predictive capacity of the model. Variable selection was ensured via lasso, ridge and elastic net penalisation systems where a thousand penalty values were tested between 0 and 1 (the closer the penalty value is to 1, the smaller the number of variables used in the models). The final equation corresponded to the most parsimonious model ensuring the lowest mean square error (MSE) within one standard error of the minimal MSE. In total, 500 models were developed to cover the multiple sampling possibilities arising from stratified partitioning. Whilst internal cross-validations were performed to establish the equation and limit the risks of overfitting, 500 permuted datasets were created to test the likelihood of spurious predictions. Besides, statistical validation was complemented with biological validation. In short, the equation of the model (as well as variable selection) was established using the first biological dataset and then applied to an independent dataset to validate the predictive capacity of the best metabolic markers. Finally, figures were developed as explained in the previous paragraph (section VII.1).

VII.3. Mathematical modelling

Mathematical modelling was used to depict the effect of temperature on *A. imbricata* coverage. We first assessed temperature under the cushion and in open areas as mentioned previously (Chapter 2, section II) to determine that the thermal protective properties of M. camachoi occurred between 7 pm and 9 am (Chapter 5, Fig. V.3). While the temperature delta between cushions and open areas were quite stable over the years, we calculated the average temperature reported between June and July to avoid a potential “day effect” when introducing the temperature parameters in the mathematical model. Hence, temperatures recorded between 7 pm and 9 am every day between June and July were collected from the two weather stations available at 3060 and 4090 m.a.s.l. The linearity of the temperatures across elevation was then assumed to calculate a theoretical value for each elevation level as previously established (Dussarrat et al., 2022).

These temperatures (T) were used in the mathematical model performed in Chapter V to predict the cover (c) of A. imbricata from T data with the equation (1) adapted from Yan and Hunt, 1999.

$$c = C_{max}\left(\frac{F_{max}-T}{F_{max}-T_{opt}}\right)\left(\frac{T}{T_{opt}}\right)\text{^}\left(\frac{T_{opt}}{F_{max}-T_{opt}}\right) \qquad (1)$$

In this equation *c* the *Atriplex* cover to predict (cm/m² of *M. camachoi*), C_{max} the maximum coverage observed in July 2021 (cm/m² of *M. camachoi*), T_{opt} the optimal temperature for growth (°C), T the measured temperature (°C) and where F_{max} the maximum value of an arbitrary parameter that limited plant life at the lowest elevation levels. We assumed that temperature was not limiting at low elevation levels (Dussarrat et al., 2022), and therefore considered that water and nitrogen scarcity, as well as salinity, could be represented through this F_{max} factor. Temperatures recorded in 2021 at elevations suitable for maximum coverage with and without interaction with the cactus was used to define T_{opt} (as mentioned in Chapter 5). Model accuracy was then confirmed using an independent dataset collected in 2016. The predictive capacity of the model was assessed by varying the T parameter according to temperatures measured in 2016, while values of the other parameters established with the 2021 dataset (*i.e.* F_{max}, T_{opt}, C_{max}) were maintained.

Note: Results of the model presented in Chapter 5 were obtained using 2020 temperatures since measurements of 2021 only became available very recently. However, although the results are thus considered preliminary, the average temperatures between June and July did not show major changes between 2016 and 2020.

VIII. METABOLIC NETWORKS

The annotation of the best predictors and their integration within the plant metabolome is an essential step to provide a biological interpretation of the results. Depending on the biological question, various software and chemical databases were used to perform these essential steps.

VIII.1. Annotation process

The annotation of a chemical compound is performed at MS and MS/MS levels and should follow the metabolomics standards initiative confidence level (MSI) (Sumner et al., 2014). The annotation of the MS spectrum can lead to the attribution of the chemical formula of the molecule through two distinct analytical ways. First, raw spectra can be directly scrutinised to determine the presence or absence of chemical elements like nitrogen (N), carbon (C), oxygen (O), sulfur (S). For instance, the FreeStyle function from Xcalibur software proposed putative chemical formula annotations

based on accurate m/z ratio and the occurrence of isotopic patterns (^{13}C, ^{18}O, ^{15}N and ^{34}S). In parallel, the accurate m/z ratios from the best metabolic predictors can be screened through chemical databases such as METLIN (Xue et al., 2020) to directly speculate on the chemical formula. Finally, the putative chemical structure is hypothesised by exploring the metabolic candidates on chemical databases such as Chebi, METLIN, DNP (http://dnp.chemnetbase.com) and Knapsack (de Matos et al., 2010; Afendi et al., 2012; Xue et al., 2020). The MS/MS spectra were used to sharpen the annotation and allow great confidence into the chemical structure (and therefore the molecule) by screening the fragmentation pattern of the studied molecule on various libraries such as MzCloud (https://www.mzcloud.org), METLIN and Massbank (Horai et al., 2010).

Pathway analyses were performed to question whether biological predictors belonged to similar biochemical routes or conversely covered various pathways across primary and secondary metabolism. The KEGG identifiers were obtained via KEGG database and the related metabolism, biochemical pathway and sub-pathways were assigned using MetaboAnalyst and PlantReactome (Kanehisa et al., 2014; Naithani et al., 2019).

VIII.2. Metabolic networks

In this project, the best metabolic predictors were integrated into metabolic networks to (i) help interpret their relevance in the phenomenon studied and/or (ii) evaluate the response of the related compounds. First, top chemical markers were integrated into a pre-existing *A. thaliana* metabolic network using KEGG identifiers on MetExplore to appreciate their potential role in plant metabolism (Cottret et al., 2010). Alternatively, structural metabolic networks were developed using the GNPS platform. The feature-based molecular network (FBMN) method creates metabolic networks based on the assumption that structurally-related molecules belong to the same pathway (Nothias et al., 2020). In short, the correlation level between two molecules is defined by the cosine score which calculates the similarity of the fragmentation spectra (from MS/MS analysis) between 0 and 1 (where 1 refers to a perfect match). Optimised parameters were used and included an ion mass tolerance of 0.01Da and a cosine score threshold of 0.7 with a minimum of 4 corresponding fragment ions. Finally, the newly developed metabolic network was uploaded into Cytoscape (version 3.8.2) to improve visualisation (Shannon, 2003).

IX. REACTION AND PATHWAY ENRICHMENT ANALYSES

Wild species may have selected the most efficient metabolic strategies to cope with extreme abiotic conditions through millions of years of evolution within the Atacama Desert. Transcriptomics data from Eshel et al., 2021 were used to investigate the reaction and pathway enrichments within 32 Atacama plant species. A comparison between the transcriptome of the 32 Atacama species and the 32 closest non-adapted sequenced species was deployed to identify expanded and over-expressed gene families. The material and method presented below was summarised from Chapter 4, which will be submitted in the following weeks.

IX.1. Generation of annotated reactions and pathways

Reactions and pathways for each of the 64 species transcriptome were built using the annotation results generated by the e2p2v4 enzymes annotation tool (Schläpfer et al., 2017). Next, PTOOLS v24.5 was used to infer reactions and pathways using default parameter values (Karp et al., 2021).

IX.2. Reaction and pathway enrichment analyses

Reaction enrichment analysis was performed by assessing the expansion of gene families and the variation of the top-expressed genes within Atacama plant species compared to phylogenetically related species. First, the total number of genes per reaction, as well as the total number of top-expressed genes per reaction (*i.e.* genes underlying the top 10% intense transcripts), were extracted from the transcriptomics data previously developed (Eshel et al., 2021).

Thereafter, pairwise comparisons (*i.e.* Atacama-Related species couples) were performed to characterise a reaction as enriched when at least 3 times more genes per reaction were observed (total genes per reaction or top-expressed genes per reaction). Results were concatenated to define the percentage of occurrence (*i.e.* the number of species in which a given reaction was enriched). In addition, a Fisher's exact test was performed to explore the enrichment of entire biochemical pathways using the phyper function in R as previously described (Wieder et al., 2021). Finally, Benjamini-Hochberg *P* correction was employed to identify significantly enriched pathways using the p.adjust function available on R (Benjamini and Hochberg, 1995). The most relevant reactions and pathways were subjected to an annotation process using the MetaCyc, KEGG and HMDB databases.

X. MANAGEMENT OF SAMPLES, DATA AND METADATA

Untargeted analysis of unknown plants collected in different years produces a large amount of data. Besides, various collaborations were developed to meet the needs of the diverse angles of analysis. Furthermore, a large amount of data remain unexplored and should be available for all partners. Finally, subsequent projects are under development and will probably require samples collected in 2019 or 2014 to, for example, serve as an independent dataset for validation. Hence, we propose this section to clearly specify the location of (i) the remaining samples and their characteristics, (ii) raw metabolic data already processed and (iii) raw data under current investigation.

X.1. Samples

All samples collected in 2019 were subjected to untargeted LCMS analysis (20 mg of dry weight), untargeted GCMS (10 mg of dry weight), targeted robotised microplate assays (40 mg of dry weight in total). Besides, 30 mg of dry powder from several samples were used for NMR analysis. In addition, multiple tests were performed on various samples. All remaining samples are stored at -80°C in four distinct plastic bags which contain silica gel (two bags tagged “Atacama 2019-2014 TD” and two bags tagged “Atacama 2021”). Depending on the time of the next analysis, we perhaps recommend redoing the lyophilisation step to ensure the absence of water in the samples.

X.2. Data availability

Published metabolic data and sample metadata from samples collected in 2014 and 2019 have been made available for the scientific community following the Findable, Accessible, Interoperable, Reusable (FAIR) principles (Jacob et al., 2020). Transcriptomics data became publicly available since the publication of the paper from Eshel et al., 2021. Reaction and pathway enrichment data as well as metabolic data and sample metadata developed for the analysis of the interaction between *A. imbricata* and *M. camachoi* will become available following the FAIR method once published. All additional data that are still under investigation are available for both laboratories in private servers. At the Meta team from INRAE Villenave d’Ornon, data are stored in the NAS (folder Thomas Dussarrat). At the Plant Systems Biology Lab, Facultad de Ciencias Biológicas, Departamento de Genética Molecular y Microbiología, data are stored in the lab’s Google drive (folder Thomas Dussarrat).

CHAPTER 3

A CORE SET OF METABOLITES PREDICTS EXTREME CLIMATE RESILIENCE

I. TOWARDS THE DISCOVERY OF GENERIC MECHANISMS

This chapter has resulted in a scientific paper accepted in New Phytologist in February 2022. Here, we proposed a short introduction (Section I) to set the topic in the context of the PhD. The manuscript of this article was then implemented in a Word format in Section II. Finally, a short conclusion was provided to integrate the into the PhD context (Section III).

All supplemental tables are available at the following link:
https://drive.google.com/drive/folders/1Z3HLMY0Hb281HEu56MM82MHtY_YQ9tkE?usp=sharing

The adaptation of an organism is inherent to heritable genetic changes that allow survival in a given environment (Borowitzka, 2018). These random mutation events result in the adjustment of metabolic pathways which enhance the capacity of a plant to perform in its new environment (Scossa and Fernie, 2020). Whilst most studies focus on the analysis of model plants in controlled conditions, wild species have naturally adapted their metabolome to their environment (Castañeda-Álvarez et al., 2016). Harsh ecosystems such as deserts, high mountains and saline lands provide a unique reservoir of adapted species that possess the fascinating ability to cope with extreme conditions while most species can not survive (Tipirdamaz et al., 2006; Díaz et al., 2019; Kumari et al., 2020). Decades of research on extreme plants have nicely improved our understanding of the metabolic mechanisms underlying adaptation (Turner, 2018). Interestingly, altitudinal clines across extreme lands are of major interest since they provide the opportunity to decipher the metabolic mechanisms that underpin adaptation as well as their response to a gradient of environmental constraints (Walker et al., 2022). This was exemplified by a great correlation between phytochemical and elevation in Alpine mountains (Defossez et al., 2021). Besides, the first meta-analyses which compiled individual studies performed on a limited or unique plant species revealed general trends among stress-resistant organisms (Dussarrat et al., 2021; Walker et al., 2022). The response of phenotypic traits such as the specific leaf area (SLA) was linked to both elevation level and global metabolic trends (Walker et al., 2022). Primary compounds like amino acids and secondary compounds like terpenoids and flavonoids were increasingly synthesised as leaf area decreased with altitude (Walker et al., 2022). However, most studies sought to identify the genetic or metabolic differences between one agronomic or model plant and its related extreme species, yielding highly species-specific markers (Dussarrat et al., 2021). Our approach adopted the opposite analytical angle to question the role of metabolic convergences in plant resilience to harsh climates. We explored the biochemical diversity of 24 Atacama plant species which flourish across an elevational gradient (from 2400 to 4500 m.a.s.l) using robotised microplate assays, GCFID, GCMS and LCMS techniques. Subsequent machine learning unveiled a generic toolbox predicting plant environment, independently of the plant species and sampling year.

II. A METABOLIC TOOLBOX PREDICTING ENVIRONMENT

The following article was accepted in New Phytologist in February 2022. However, the PDF format is not yet produced. Here, we present the Word format that will be published online in the coming weeks.

Predictive metabolomics of multiple Atacama plant species unveils a core set of generic metabolites for extreme climate resilience

Thomas Dussarrat[1,2], Sylvain Prigent[2,3], Claudio Latorre[4,5], Stéphane Bernillon[2,3], Amélie Flandin[2,3], Francisca P. Díaz[1], Cédric Cassan[2,3], Pierre Van Delft[6], Daniel Jacob[2,3], Kranthi Varala[7,8], Jérôme Joubes[6], Yves Gibon[2,3], Dominique Rolin[2,3], Rodrigo A. Gutiérrez[1]*, Pierre Pétriacq[2,3]*

[1]: Departamento de Genetica Molecular y Microbiologia, Pontificia Universidad Catolica de Chile. FONDAP Center for Genome Regulation and ANID Millennium Science Integrative Program-Millenium Institute for Integrative Biology (iBio), Av Libertador Bernardo O'Higgins 340, Santiago, Chile; [2]: Univ. Bordeaux, INRAE, UMR1332 BFP, 33882 Villenave d'Ornon, France; [3]: Bordeaux Metabolome, MetaboHUB, PHENOME-EMPHASIS, 33140 Villenave d'Ornon, France; [4]: Departamento de Ecología, Pontificia Universidad Católica de Chile, Av Libertador Bernardo O'Higgins 340, Santiago, Chile; [5]: Institute of Ecology and Biodiversity (IEB), Las Palmeras, 3425, Ñuñoa, Santiago, Chile; [6]: Univ. Bordeaux, CNRS, Laboratoire de Biogenèse Membranaire, UMR 5200, F-33140 Villenave d'Ornon, France; [7]: Dept. of Horticulture and Landscape Architecture, Purdue University, IN 47907, USA; [8]: Center for plant biology, Purdue University, IN 47907, USA

* Author for correspondence:
Pierre Pétriacq and Rodrigo A. Gutiérrez.
Email: pierre.petriacq@inrae.fr and rgutierrez@bio.puc.cl

Summary

- Current crop yield of the best ideotypes is stagnating and threatened by climate change. In this scenario, understanding wild plant adaptations in extreme ecosystems offers an opportunity to learn about new mechanisms for resilience. Previous studies have shown species specificity for metabolites involved in plant adaptation to harsh environments.

- Here, we combined multi-species ecological metabolomics and machine learning-based generalised linear model predictions to link the metabolome to the plant environment in a set of 24 species and belonging to 14 families growing along an altitudinal gradient in the Atacama Desert.

- Thirty-nine common compounds predicted the plant environment with 79% accuracy, thus establishing the plant metabolome as an excellent integrative predictor of environmental fluctuations. These metabolites were independent of the species and validated both statistically and biologically using an independent dataset from a different sampling year. Thereafter, using multiblock predictive regressions, metabolites were linked to climatic and edaphic stressors like freezing temperature, water deficit and high solar irradiance.

- These findings indicate that plants from different evolutionary trajectories use a generic metabolic toolkit to face extreme environments. These core metabolites, also present in agronomic species, provide a unique metabolic goldmine for improving crop performances under abiotic pressure.

Keywords: adaptation, extreme environments, multiple species, plant metabolism, predictive metabolomics.

Introduction

Humans domesticated plants 10,000 years ago in the hostile environments of the Fertile Crescent (Riehl *et al.*, 2012; Dai *et al.*, 2012). Over the years, selected crops have been improved by a variety of methods. However, current yields of domesticated therophytes are stagnating and threatened by climate change despite significant efforts to develop abiotic stress tolerance for the best ideotypes (Long *et al.*, 2015). Wild plants naturally evolved mechanisms to meet abiotic constraints in natural habitats from which they cannot escape (Fatima *et al.*, 2020; Signori-Müller *et al.*, 2021). In this scenario, returning to wild plant species that live and thrive in some of the harshest environments on Earth offers an opportunity to find new strategies for crop improvement (Castañeda-Álvarez *et al.*, 2016). Recent studies pinpointed relevant metabolic clusters in adaptation to extreme environments in plants harvested in high mountains, deserts and salt lands (Dussarrat *et al.*, 2021). These adaptive mechanisms involved the accumulation of amino acids (Lugan *et al.*, 2010) as precursors of secondary metabolites, and carotenoids (Cui *et al.*, 2019) and polyphenols (Hashim *et al.*, 2020) as processors of reactive oxygen species (ROS). Besides, most studies were carried out on a unique or limited number of species (Dussarrat *et al.*, 2021), which, combined with high biochemical diversity, led to highly specific metabolic markers involved in adaptive mechanisms exclusive to species or environment (Peters *et al.*, 2018; Dussarrat *et al.*, 2021).

The metabolome is an excellent integrative system to predict plant environment since it carries imprints of omic inferences and environmental influences (Kosmacz *et al.*, 2020; Lewis & Kemp, 2021). Ecological metabolomics aims to study the environmental impact on metabolic responses, acclimation and adaptation processes in natural ecosystems. Applying untargeted ecological metabolomics on multiple species could unravel universal plant adaptive strategies to abiotic factors in their natural environment (Poorter *et al.*, 2012; Umair *et al.*, 2019; Sardans *et al.*, 2020; Wong *et al.*, 2020). However, this approach has primarily focused on phytochemical diversity analysis. Plant metabolomes were recently used to predict phenotypic traits such as yield and stress resistance (Luna et al., 2020; Szymański et al., 2020; Zhu et al., 2018) within specific species. By exploiting multiple species, previous studies reported strong relationships between growth rate and biomass composition (Roch *et al.*, 2020) and between phytochemical diversity and environmental conditions of plants growing in alpine regions (Defossez *et al.*, 2021). Interestingly, several phenotypic traits predicted from plant metabolism were further used to predict complex output like plant fitness but remain tedious to collect (Laughlin *et al.*, 2012; Laughlin & Messier, 2015). Thus, ecological metabolomics could be used to uncover readily measurable soft traits that can predict complex outputs such as plant fitness. Adaptation to extreme environments is thought to rely on specialised secondary metabolic pathways often considered to be species-specific (Moghe & Last, 2015; Scossa & Fernie, 2020). However, generic mechanisms may also exist. To test this hypothesis, large scale metabolomics in multiple wild species

are needed to unveil general metabolic interactions with environmental factors and propose adaptive roles for specific metabolites (Wong *et al.*, 2020).

The Atacama Desert is the driest non-polar desert on Earth. In addition to extreme aridity, the Atacama is characterised by high solar radiation, extreme daily temperature oscillations, high soil salinity and low nitrogen content (Eshel *et al.*, 2021). Although multiple abiotic factors are intense enough to severely limit plant life, this desert hosts tens of plant species (Jordan & Kirk-Lawlor, 2014; Díaz *et al.*, 2016, 2019), thus bestowing a unique opportunity to analyse adaptive metabolic plant responses to abiotic stress in an entire ecosystem. The present study aimed to characterise the metabolic profiles of 24 dominant plant species in 19 different sites along an altitudinal transect in the Atacama Desert. Biological and environmental diversity was used to question the extent adaptation to extreme environments relies on generic metabolic mechanisms.

To meet these ambitious objectives, multi-platform metabolomics covering primary compounds including carbohydrates, amino and organic acids, fatty acids and secondary metabolites revealed metabolic features that participate in environmental adaptation. Subsequent machine learning modelling of this comprehensive dataset via a generalised multilinear-based statistical approach established that the metabolome of these 24 extremophile plants was an excellent integrative predictor of plant environments. Moreover, our analysis uncovered a common set of metabolites associated with extreme climate resilience.

Materials and Methods

Plant materials and sampling. Aerial parts of 24 plant species belonging to 14 plant families (Table III.S1) were collected in their natural conditions from the Chilean Atacama Desert (Talabre-Leja transect (Díaz *et al.*, 2016), lat 22°-24°S). For each species, a minimum of three biological replicates composed of multiple plants was collected. Each species has been collected in 1 to 6 distinct elevation levels (Fig. III.1) depending on the biological availability, directly snap frozen into liquid nitrogen, brought back to the laboratory on dry ice and stored at -80°C until freeze-drying. Sampling was performed during two consecutive days (06-07/04/2019) between 9.30 am to 5.30 pm. The time variation between samplings in different environments did not impact starch content, suggesting stable central metabolism during the sampling period (Fig. III.S1). Additionally, crops and ornamental plant species including *Capsicum annuum*, *Phaseolus vulgaris*, *Spinacia oleracea*, *Vicia faba*, *Pisum sativum*, *Beta vulgaris*, *Portulaca oleracea*, *Helianthus annuus*, *Zea mays*, *Nicotiana tabacum* and *Solanum lycopersicum* were grown in multiple natural conditions in France. The aerial parts of those plants were harvested, snap-frozen into liquid nitrogen and stored at -80°C until freeze-drying. All freeze-dried material from extremophiles and common plants were kept at -80°C until further analysis.

Environmental data. Climatic conditions were characterised using two meteorological stations (at 3060 and 4090 meters above sea level (m.a.s.l)), which measured temperature, humidity and solar irradiance as well as precipitation or soil moisture levels every hour throughout the year 2018-2019 (Eshel *et al.*, 2021). Besides, soil chemical properties including pH and contents of nitrate, ammonium, Olsen phosphorous, zinc, potassium, manganese, copper, iron, boron, molybdenum, sulfur, calcium, manganese, sodium, chlorine, bicarbonate salt and silt were measured and described for over three years (Eshel *et al.*, 2021).

Metabolite extraction. Using 20 mg of lyophilised plant material (from crops, ornamental and Atacama plants), robotised extractions of metabolites were performed according to an ethanol fractionation protocol (Luna et al., 2020), which targets a wide range of semi-polar plant biochemicals including primary compounds (soluble sugars and starch, organic and amino acids, total proteins) and specialised metabolites (terpenes, phenolics, alkaloids). In parallel, 10 mg of lyophilised plant samples were used to extract fatty acyls from total lipids as described previously (Domergue *et al.*, 2010).

Metabolomics. Ethanol extracts were screened for multiple compounds (chlorophyll, glucose, fructose, sucrose, malate, free amino acids, nitrate, total proteins and starch) based on coupled enzyme assays (Luna et al., 2020). The same extracts were also subjected to untargeted metabolic profiling by UHPLC-LTQ-Orbitrap mass spectrometry (LCMS) using an Ultimate 3000 ultra-high-pressure liquid chromatography (UHPLC) system coupled to a LTQ-Orbitrap Elite mass spectrometer interfaced with an electrospray ionisation (ESI) source (ThermoScientific, Bremen, Germany) operating in both negative and positive ion modes as described previously (Luna et al., 2020). The separation was performed using a C18 column (C18-Gemini, 2.0x150mm, 3 µm, 110Å, Phenomenex, USA). Full scan high resolution MS spectra were acquired at 240k resolution power at *m/z* = 200 Da. Besides, LCMS/MS acquisitions were acquired in higher-energy collisional dissociation (HCD) mode at a normalised collision energy of 60% and 35% (ESI- and + respectively). Fatty acyls were analysed using gas chromatography coupled to a flame ionisation detector (GCFID) or a mass spectrometric detector (GCMS) as detailed previously (Domergue *et al.*, 2010). Biochemical phenotyping and LCMS experiments were performed for all plants (*i.e.* Atacama plants from 2014 and 2019, crops and ornamental plants), while GCFID and GCMS experiments were performed for 2019 Atacama plants and crops and ornamental plants exclusively.

Processing of metabolomic data. Raw LCMS data were processed via XCMS (v 4.2) in R (v 3.6.1) (Smith *et al.*, 2006) using in-house optimised parameters (Luna et al., 2020) yielding 8750 detected RT-*m/z* pairs for 5130 ESI- and for 3620 ESI+ modes. Subsequent data cleaning (blank check, $\Delta_{RT} < 60$ s, $\Delta_{m/z} < 0.025$ Da, coefficient of variation in quality controls < 30%) generated 4540 metabolic variables (2564 ESI- and 1976 ESI+) that were retained for chemometric analyses. Both untargeted and targeted metabolomics data were first normalised by median normalisation, cube-root transformation

and Pareto scaling using MetaboAnalyst v.3 (Xia *et al.*, 2015) before applying multivariate and univariate statistical analyses. The non-normalised dataset obtained after preprocessing is available in Supplemental Table III.2 and deposited online (see Section Data availability).

Generalised multilinear models (GLM). Generalised linear modelling was performed to appreciate the quantitative correlation between metabolism and elevation levels used as a proxy of the plant environment. All metabolic variables that could not be measured based on detection limitations were inputted as 0 in the data matrix. The linear models were generated using the glmnet package (Friedman *et al.*, 2010) in the R software (R Core Team, 2020) (version 3.6.1). Three model types were constructed (lasso, elastic net and ridge) by varying the penalty value of elastic net as a proxy to modulate the number of variables used by the models. Thousand values ranging from 0 to 1 were tested. Internal cross-validation was performed for the construction of the models to mitigate the overfitting. The best model was chosen based on Mean Square Error (MSE), and the most parsimonious model within one standard error of the minimal MSE was selected to perform predictions. The datasets were divided into three parts: 70% of the plants in a "training" set, 20% in a "test" set and 10% in the "validation" set to perform real predictions using the best model developed with both training and testing sets. Stratified sampling was used to perform a uniform sampling of the individuals based on the measured elevation levels. Due to this random partitioning, 500 different simulations were performed to sample the solution space of possible predictions. Besides, 500 sets of randomly assigned elevation levels were created to test the likelihood of spurious predictions. Student tests were performed to compare the 500 results from models performed with permutated and real elevation levels. The occurrences of the metabolic variables among the 500 simulations were analysed to extract the best predictors (Table III.1).

Finally, statistical validation was complemented with biological confirmation to validate the predictive capacity of the metabolic markers using an independent sample set harvested in 2014 (Eshel *et al.*, 2021) as for 2019. Using the equation calculated on the entire 2019 dataset, a validation model predicted the elevation for each plant from the 2014 dataset. The quality of the prediction was evaluated by both the coefficient of determination and the *P*-value observed when comparing real and predicted altitudes (Fig. III.2). The same permutation protocol was used to test the likelihood of spurious prediction of 2014 validation models.

Multivariate statistical analyses. The normalised dataset was processed through multivariate analysis like PCA via FactoMineR package (Lê *et al.*, 2008) from R software (version 3.6.1) and O2PLS via SIMCA 16.0.1 (Umetrics, Sweden). Besides, Tukey's tests were performed to compare the expression of the metabolic markers between species or between environments using agricolae package (Mendiburu, 2020) with a threshold of significativity established at $P<0.01$. Finally, box plots, scatter plots, correlation plots and heatmaps were realised using ggplot2, ggpubr, Hmisc and pheatmap

packages (Wickham, 2016; Kolde, 2019; Jr *et al.*, 2020; Kassambara, 2020) (Pearson correlation, Ward algorithm) from R software, respectively.

Annotation. The best metabolic predictors were annotated using two different methods. Firstly, MS spectra were used to analyse the isotopic patterns (^{13}C, ^{18}O, ^{15}N and ^{34}S) and speculate on the ion composition. Besides, all putative chemical formulas were calculated by the FreeStyle function from Xcalibur 4.2 software with the following minimal and maximal constraints on chemical elements: ^{14}N 0-60, ^{16}O: 0-6, ^{12}C: 0-100, ^{1}H: 0-200, ^{32}S: 0-60, ^{35}Cl: 0-60, ^{31}P: 0-60 and a mass tolerance at 10 ppm. Thus, the best candidates were chosen based on the MS spectra analysis and screened on chemical databases (Chebi (de Matos *et al.*, 2010), METLIN (Xue *et al.*, 2020), DNP (http://dnp.chemnetbase.com), Knapsack (Shinbo et al., 2006)). In parallel, accurate *m/z* values of the most discriminant monoisotopic ions were screened using METLIN database (Smith *et al.*, 2005) for putative annotation. The resulting outputs from both methods were compared to select the best putative annotation for each ion (Table III.1). Besides, MS/MS spectra of all samples were used to improve the annotation level by comparing experimental fragments with experimental MS/MS spectra available in multiple libraries like Massbank (Horai *et al.*, 2010), MzCloud (https://www.mzcloud.org) and METLIN (Xue *et al.*, 2020). The annotation level of each predictor was therefore attributed following the metabolomics standards initiative confidence level (MSI levels) (Sumner *et al.*, 2014).

Pathway analysis and metabolic networks. The 39 predictors were screened through chemical databases and integrated into a metabolic network to better interpret their role in the plant response to extreme conditions. The KEGG identifiers were identified using KEGG database (Kanehisa *et al.*, 2014) on the 39 molecules if available or on chemically related compounds (Table III.S3). Thereafter, a pathway enrichment analysis was realised using MetaboAnalyst (Xia *et al.*, 2015) and PlantReactome (Naithani *et al.*, 2019) databases, and combined with the best markers integration into a pre-existing *A. thaliana* metabolic network via MetExplore (Cottret *et al.*, 2010).

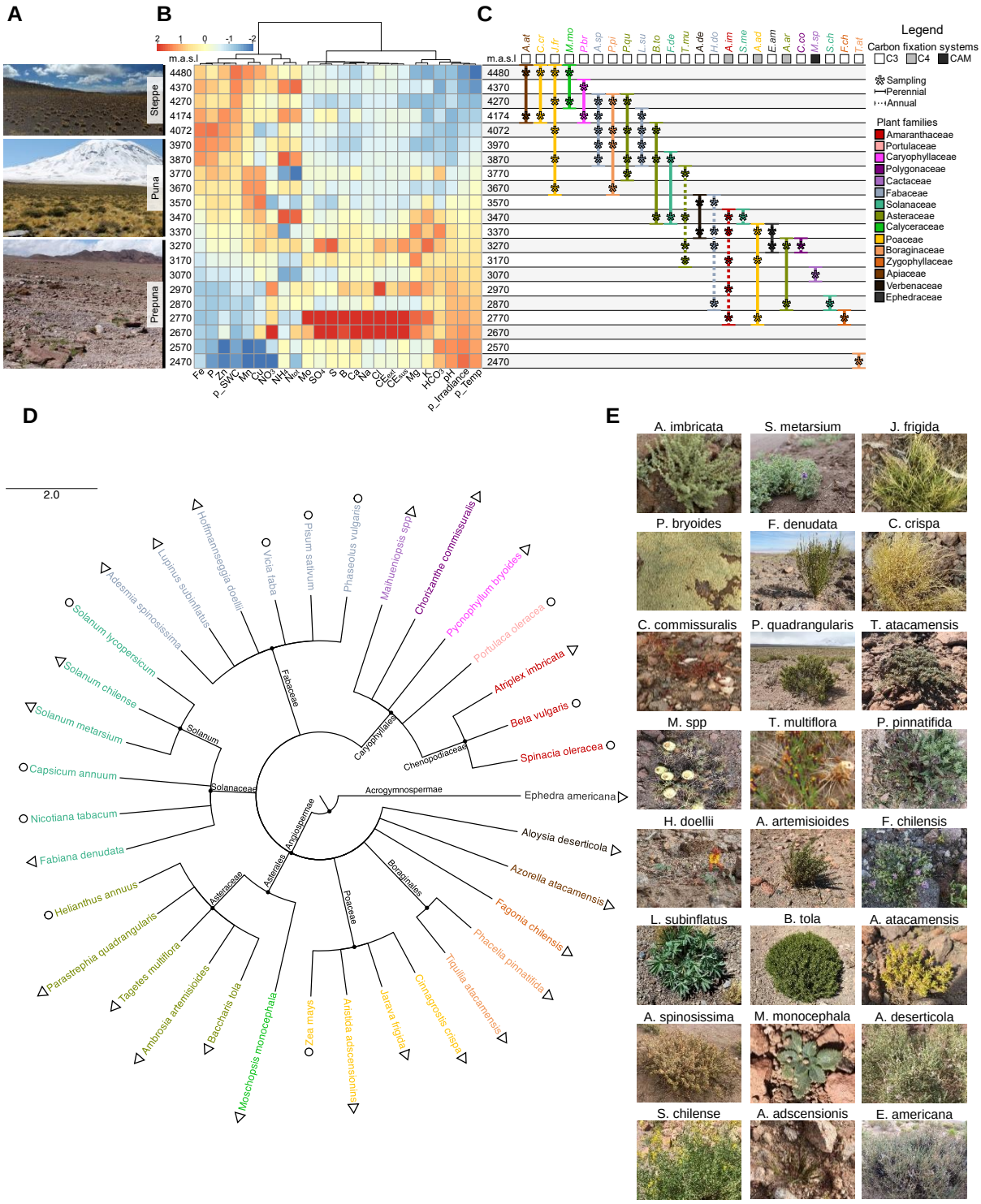

Fig. III.1 | Depiction of Atacama plant diversity despite extreme conditions. A. Picture of the three vegetation belts. **B.** Description of the environmental conditions observed along the elevation gradient (Pearson correlation, $P<0.05$). *SWC*: soil water content, *Ntot*: total nitrogen, *CE*: electrical conductivity, *Temp*: temperature, *p_* represents a partially predicted parameter. **C.** Description of the sampling site ranges and main characteristics (carbon fixation systems or lifespan) of the collected plant species. **D.** Analysis of the taxonomic relationships between Atacama species and between Atacama and agronomic or ornamental plant species. Triangles represent the Atacama plants while circles represent the agronomic and ornamental species. **E.** Pictures of the Atacama plant species collected. Adapted from Dussarrat et al., 2022, *New Phytologist*.

Results

Plant diversity in the extreme conditions of the Atacama Desert. The Atacama Desert represents one of the harshest environments for plant life (Jordan & Kirk-Lawlor, 2014; Díaz *et al.*, 2016), where plants must endure the major abiotic stresses currently threatening agriculture. The Talabre-Lejía Transect (TLT) spans an elevation gradient covering three different plant communities: the poorly vegetated Prepuna (2400-3300 m.a.s.l), the Puna shrubland (3300-4000 m.a.s.l) and the high Andean Steppe (4000-4500 m.a.s.l) (Fig. III.1A) (Díaz *et al.*, 2016; Eshel *et al.*, 2021). Water availability increases and temperature decreases with altitude, while high solar irradiance and very low nitrogen levels are critical constraints throughout the TLT transect (Eshel *et al.*, 2021). Rainfall ranges from 20 mm/year in the prepuna to 160 mm/year in the steppe, illustrating the extreme aridity as compared to other plant-sheltering deserts (Báez & Collins, 2008; Li *et al.*, 2015; Díaz *et al.*, 2016; Ziaco *et al.*, 2018). The daily average solar irradiance of 600 W/m²/d along this transect is three times higher than many deserts and high mountain ecosystems (Bo *et al.*, 2009; Zhang *et al.*, 2010; Arancibia-Bulnes *et al.*, 2014). Besides, low total nitrogen (average 9 mg/kg) throughout the transect, low phosphorus levels (6-20 mg/kg), and high salinity in Prepuna sites add to the harsh conditions plants must endure. Nonetheless, plant life in this ecosystem of Atacama can be traced back to 45,000 years ago (Latorre *et al.*, 2002; Díaz *et al.*, 2019) and likely thrived in such extreme conditions since probably 12 million years ago (Jordan & Kirk-Lawlor, 2014). Hence, this ecosystem represents a unique resource of adaptive mechanisms potentially relevant to engineer crop resilience. Interestingly, deep-sequencing of 32 dominant species representing the major clades highlighted common and specific strategies relevant for plant survival (Eshel *et al.*, 2021). In this context, we collected 21 of these 32 plant species based on their coverage in their natural ecosystem. We complemented this set with one Cactaceae, one Solanaceae and one Boraginaceae to finally represent relevant biodiversity covering annual and perennial plants, different carbon fixation systems (*i.e.* C3, C4 and CAM) and different lifespans like shrubs and herbs (Fig. III.1C and Table III.S1). Clear distinctions regarding the distribution of life-form and carbon fixation types have been highlighted where all annuals and C4 plants were observed under an elevation of 3870 m.a.s.l. Additionally, while some species were relatively specific to a single environment (*e.g. Moschopsis monocephala*), other species had a wide distribution along the transect area that we divided into 19 sites (each 100 m.a.s.l) (Fig. III.1B). We also selected 11 agronomic and ornamental species based on their plant family to analyse and compare using the same experimental procedures (as explained in the Methods section). A taxonomic analysis performed on the Atacama and agronomic plant species via NCBI taxonomy browser unveiled the relationships between the 14 Atacama plant families (Fig. III.1D). Interestingly, this sample set of 23 angiosperms and one gymnosperm included well known resilient plant families like *Cactaceae* and *Boraginaceae* (Ma *et al.*, 2010) together with species of economic interest such as *Poaceae*, *Asteraceae*, *Fabaceae* and

Solanaceae. Besides, the 11 agronomic and ornamental plant species covered 5 of the 14 Atacama plant families (including the most widespread ones like *Poaceae*, *Fabaceae*, *Asteraceae* and *Solanaceae*).

Predictive metabolomics reveals a core metabolic set in multiple resilient species. To get insight into the mechanisms by which these extremophile plants (i) adapt to the extreme conditions of the Atacama Desert and (ii) respond to environmental variations, we performed multi-platform metabolomics to screen both primary and secondary metabolisms from the aerial tissues of Atacama plants and agronomic species (Fig. III.2A). Quantitative evaluation of 10 major compounds by biochemical phenotyping (used as key physiological indicators) and 26 fatty acids by gas chromatography coupled to a flame ionisation detector (GCFID) highlighted a significant reduction in chlorophyll, nitrate and protein content in Atacama species when compared to 11 known crops and ornamental plants (Fig. III.S2). In addition, the unknown biochemical diversity of these extreme Atacama plants was analysed through untargeted metabolomics using GCMS and liquid chromatography-mass spectrometry (LCMS), which resulted in 335 acyl chains and 4540 semi-polar features after preprocessing (Fig. III.S2). Given that the phytochemical diversity fluctuated with environmental conditions along the elevation gradient (400-2000 m.a.s.l) (Defossez *et al.*, 2021), generalised linear modelling (GLM) was deployed to test whether the metabolome (4911 variables) could predict environmental conditions (Fig. III.2A). Elevation represents the integration of abiotic factors (Carpenter, 2005), among which climatic and edaphic factors have been previously described (Eshel *et al.*, 2021) (Fig. III.1B). Thus, the elevation level of the 19 sampling spots was used as a proxy of the 19 environmental conditions analysed. First, the possibility of calculating the elevation levels from five different plant species selected based on both their biomass and coverage along the elevation gradient was evaluated. For each species, 80% of the sample set (*i.e.* training sets) was used for the regression analysis. The equation was then used to calculate elevation for the 20% of the sample set remaining (*i.e.* testing set). Interestingly, the resulting average R^2 from 500 models (*i.e.* fits between calculated and measured elevation) ranged between 0.88 and 0.96 depending on the species (Fig. III.2B). These results indicate the plant metabolome integrates environmental variations. Environmental conditions elicited characteristic metabolic patterns where compounds correlate with elevation allowing us to infer the altitude from which the sample was collected.

Moreover, estimating the altitude (*i.e.* resulting as an environment proxy) from metabolic data alone for species from several plant families raised the question of whether generic mechanisms serve as a basis for adaptation to extreme environments such as the Atacama Desert. To address this question, we used GLM on the entire dataset divided into (i) a training set (70%), (ii) a testing set (20%) and a validation set (10%) since the total size of the dataset was sufficient (n = 224). We thus predicted the plant environment (*i.e.* elevation level) and highlighted the shared metabolic predictors (Fig. III.2A). A first modelling step determined the predictive capacity of the 4911 metabolic variables, represented by their percentage of use in the models. Consequently, a threshold of 40% (*i.e.* variables used in more than

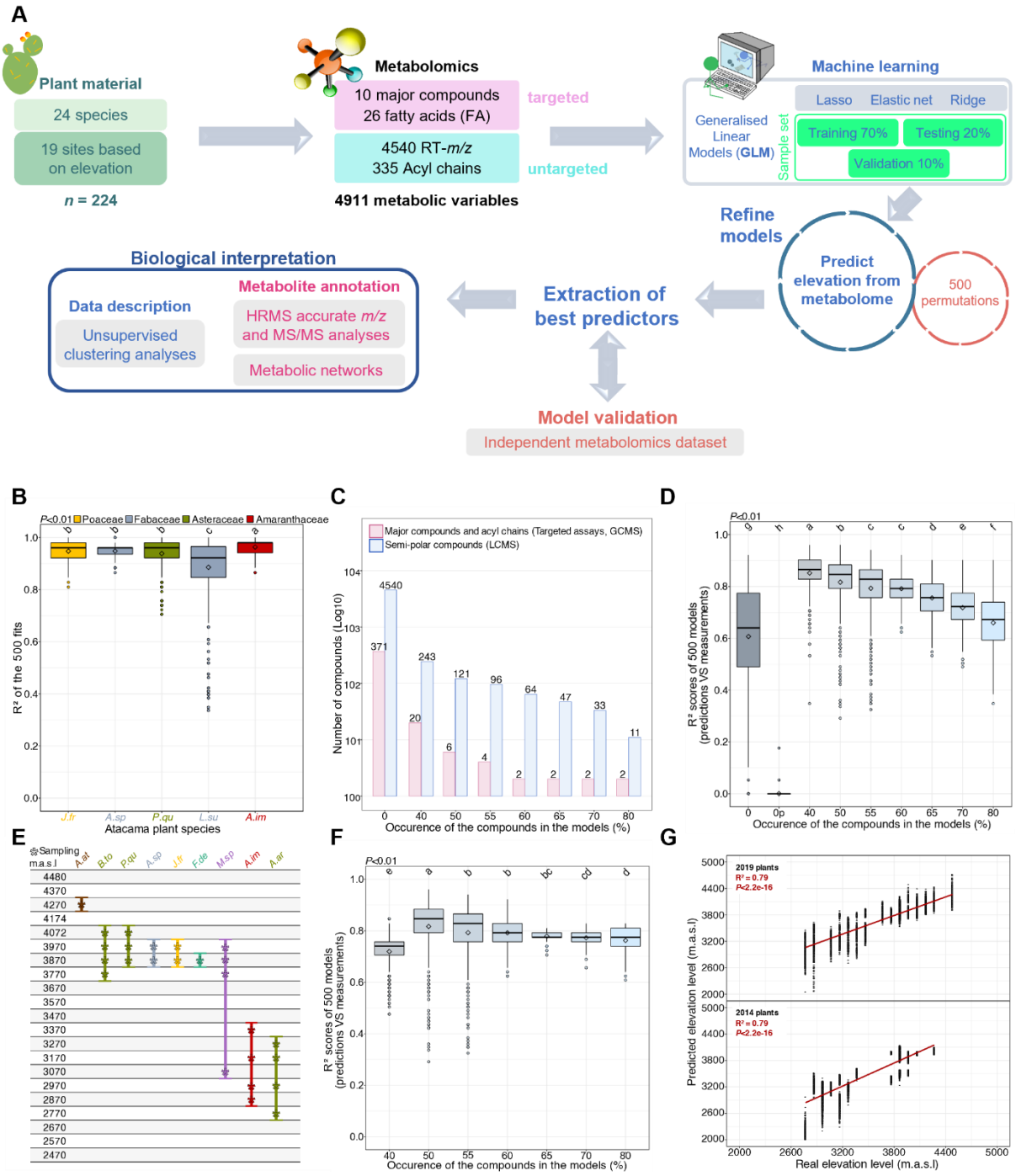

Fig. III.2 | Predictive metabolomics of Atacama plants. A. A simplified scheme of the predictive metabolomics approach used in this study. **B.** Species-specific level: R^2 scores of the fit between calculated and real elevation levels with letters indicating statistical significance (Tukey's test, $P<0.01$). Theoretical elevation levels were calculated from plant metabolome. **C.** Global level: threshold of the variable occurrence defined by 500 models performed on all variables for all species. The 13 variables used in 80% represent the most relevant compounds for predicting elevation. Numbers referred to the precise number of compounds for the corresponding occurrence. **D.** R^2 scores depending on the variable occurrence threshold (Tukey's test, $P<0.01$). **E.** Biological validation using an independent sample set from 2014. **F.** R^2 scores obtained by predicting the elevation level from 2014's plants using the multilinear equation calculated on 2019's plants depending on the variable occurrence threshold (Tukey's test, $P<0.01$), **G.** Predicted elevations from 2019 and 2014 plants using the best 66 markers (Pearson correlation). Compounds in c, d and f refer to metabolic variables *stricto sensu* prior to annotation.
Adapted from Dussarrat et al., 2022, *New Phytologist*.

40% of the 500 models) included 263 features while 80% involved the best 13 metabolic predictors (Fig. III.2C). Subsequently, each threshold was processed to (i) exclude the non-predictive features and (ii) tightly select the best ratio between the predictive capacity and the number of metabolic variables. The plant environment was considerably predictable at 66% and 79% using 13 or 66 markers, respectively (Fig. III.2D). Besides, lower thresholds (*e.g.* 40%) allowed better predictions but yielded less robust predictors (*i.e.* higher standard deviations). Importantly, 500 permutation sets involving randomly assigned elevation levels were developed to test the likelihood of spurious predictions, which led to a mean R^2 of 0% and thus statistically validated the GLM-based modelling approach. Hence, we demonstrated that common features could greatly predict plant environments (79%), independently of the species and family.

To further test the robustness of such predictions, we biologically confirmed the predictive capacity of the metabolic features using an independent dataset composed of 9 Atacama plant species harvested in 2014 and covering 12 environments (2770 to 4270 m.a.s.l) (Fig. III.2E). The linear equation developed using the 2019 samples was then applied to the 2014 dataset to estimate elevation levels, thereby resulting in similar predictive patterns between 2019 and 2014 (Fig. III.2F). Altogether, both mean R^2 prediction and standard deviation results pinpointed towards an ideal threshold of 60% (66 variables), which allowed a prediction at 79% for both years ($P < 2.2e^{-16}$) (Fig. III.2G). These results hence confirm that plants harbour a core set of metabolites to adapt to the environmental constraints.

A not-so-specialised set of secondary metabolites also detected in agronomic and ornamental plant species. Next, we annotated the best 66 predictors using both accurate *m/z* values and MS/MS analysis. This annotation process allowed excluding the fragments observed among the 66 features (Table III.S4), finally retaining 39 metabolic predictors without remarkable impact on the average R^2 (Fig. III.S2A). The MSI annotation level for each predictor is presented in Table III.1. Notably, the best predictor was starch, while 37 metabolites referred to semi-polar compounds (Table III.1). Only 6 markers were positively correlated to elevation, while the intensity of the remaining compounds decreased with the elevation (Fig. III.S3B). Remarkably, predictors in Atacama plant species were also found in several agronomic and ornamental plants (Table III.S5 and III.S6), demonstrating the ubiquitous nature of these metabolites. These 39 compounds were queried in biochemical databases (*e.g.* Kegg, PlantReactome) to perform a pathway analysis (Table III.S3) and placed into a pre-existing *A. thaliana* metabolic network available on MetExplore(Kanehisa *et al.*, 2014) (Fig. III.S4). More than half of the markers were involved in secondary metabolism (56%), while primary metabolism and regulators (*e.g.* jasmonates) covered 31% in total (Fig. III.3). The remaining 13% included 3 unknown compounds and 2 salt artefacts that combined sodium and magnesium to formic acid, suggesting salt hyperaccumulation processes (Fig. III.3 and III.S3A). Notably, starch, trehalose and amino acid-related pathways were involved in crosstalk with the biosynthesis of secondary

metabolites, while the central place of raffinose was highlighted in galactose metabolism involving other oligosaccharides known for their role in abiotic stress tolerance (Vinson *et al.*, 2020) (Fig. III.S4A).

Table III.1 | Annotation of the 39 best metabolic markers. *ND*: not determined.

VariableID	Occurrence in the model	Correlation	Observation	Detected *m/z*	Detected RT	P value FDR	Ion type	Predicted *m/z*	Δ m/z (ppm)	Putative formula	MSI level	Putative compound
Starch	98.6	neg	/	/	/	<1.64E-15	/	/	/	/	/	/
n_2561	90.0	neg	Mg(C2H2O4)n + HCOO-	386.93908	0.9	<1.64E-15	/	/	/	Mg(C2H2O4)n + HCOO-	/	Mg(C2H2O4)n + HCOO-
p_1777	87.4	neg	Na(HCO2)n + HCOO-	274.87218	1.1	<1.64E-15	/	/	/	Na(HCO2)n + HCOO-	/	Na(HCO2)n + HCOO-
n_0601	86.6	neg	/	233.10296	3.9	<1.64E-15	[M+H]-	234.11024	0	C10H18O6	MS3	3,6-Dihydroxy-2,7-dimethyloctanedioic acid
p_2029	85.7	neg	/	298.09560	3.6	1.03E-12	[M+H]-	297.08780	4.8	C11H15N5O3S	MS2	5'-Methylthioadenosine
n_0615	84.8	neg	/	236.05628	5.6	<1.64E-15	[M+H]-	237.06370	0.4	C11H11NO5	MS3	N-Benzoylaspartic acid
p_2329	83.7	pos	/	323.07473	7.8	8.91E-08	[M+H]+	322.06890	4.9	C15H14O8	MS3	Leucodelphinidin
p_0179	82.7	neg	Fragment of 251.16	116.05708	11.5	1.64E-15	[M+H]+	250.15640	4.2	C15H22O3	MS3	Xanthoxin
Qt027	82.5	pos	/	/	/	1.01E-10	/	/	/	/	/	ND
p_3344	82.0	neg	/	584.27182	9.1	<1.64E-15	[M+H]+	583.26820	5.9	C34H37N3O6	MS2	N1,NE,N10-Tricoumaroyl spermidine
n_3183	81.4	neg	/	436.22628	5.4	<1.64E-15	[M-H]-	437.02315	0.8	C25H31O4N3	MS2	*N1,N10-Dicoumaroylspermidine*
n_2074	79.3	neg	/	349.15000	7.3	<1.64E-15	[M-H]-	350.15727	0.4	C15H26O9	MS3	Azelaic acid glycoside
n_4005	78.2	pos	/	503.16151	1.4	1.10E-05	[M-H]-	504.16900	0.4	C18H32O16	MS2	Raffinose
p_0421	79.2	neg	/	144.10132	1.3	<1.64E-15	[M+H]+	143.09460	3.7	C7H13O2N	MS2	Proline betaine
n_2571	78.3	neg	/	387.16572	5.2	1.64E-15	[M-H]-	388.17330	1.3	C19H28O9	MS2	7-Epi-12-hydroxyjasmonic acid glucoside
n_4749	77.2	neg	M: 625.14020, ESI Relative: p_3426	627.14616	6.0	<1.64E-15	[M-H]-	626.14830	1.3	C27H30O17	MS3	3,3',4',5,7,8-Hexahydroxyflavone, 7-O-α-L-Rhamnopyranoside, 8-O-β-D-glucopyranoside
n_1843	77.1	neg	M: 331.04562	332.04902	8.9	<1.64E-15	[M-H]-	332.05320	0.9	C16H12O8	MS3	Quercetagetin methyl ether
p_0586	73.6	neg	/	161.09152	1.1	<1.64E-15	[M+H]+	161.09207	3.9	C3H7NO2	MS2	D-Alanyl-D-alanine
p_0184	73.0	neg	M: 116.06969	117.07322	1.3	6.79E-05	[M+H]+	115.06330	0.5	C5H9O2N	MS2	Proline
n_2208	71.1	pos	/	315.04863	11.1	0.000000151	[M+H]+	314.04270	3.8	C16H10O7	MS3	Wedelolactone
n_2574	70.2	neg	/	387.20213	5.2	1.96E-14	[M-H]-	388.20970	0.9	C19H32O8	MS3	9,13-Dihydroxy-4-megastigmen-3-one 9-glucoside
n_5122	69.4	neg	/	925.47889	10.0	<1.64E-15	[M-H]-	926.48750	1.1	C47H74O18	MS3	Araloside A
n_4791	69.0	neg	/	639.13506	8.1	1.09E-08	[M-H]-	640.14280	0	C31H28O15	MS3	Quercetin 3-(6''-feruylglucoside)
n_0657	68.3	pos	/	241.07186	5.0	8.10E-06	ND	ND	ND	ND	ND	ND
n_2605	67.6	neg	/	389.21773	8.4	<1.64E-15	[M-H]-	390.22540	0.4	C19H34O8	MS3	(3S,5R,6S,7E,9x)-7-Megastigmene-3,6,9-triol 9-glucoside
n_2530	66.9	neg	/	383.13785	5.6	<1.64E-15	[M-H]-	383.13562	2.1	C19H28O4S2	MS3	Unkwown
n_0378	65.0	neg	/	197.04552	3.9	<1.64E-15	[M-H]-	198.05280	0.3	C9H10O5	MS3	3-(3,4-dihydroxyphenyl)lactic acid
n_1078	64.4	pos	/	209.15278	5.0	1.20E-05	[M+H]+	208.14630	4.1	C13H20O2	MS3	4-Hydroxy beta ionone
p_2652	64.0	neg	/	365.10471	1.3	<1.64E-15	[M+Na]+	342.11622	1.6	C12H22O11	MS2	Trehalose
p_3452	63.6	neg	M1: 641.16797	642.17144	6.4	<1.64E-15	[M+H]+	640.16390	5.0	C28H32O17	MS3	Quercetagetin 7-methyl ether 3-neohesperidoside
n_1201	62.5	neg	/	288.06021	1.2	1.14E-11	[M-H]-	289.06740	1.2	C17H11N3S	MS4	2-[5-(Pyrimidin-4-yl)thiophen-2-yl]quinoline
n_0421	62.5	neg	/	204.08774	1.3	<1.64E-15	[M-H]-	205.09500	0.2	C8H15NO5	MS3	N-Acetyl-D-Lucosamine
n_4973	62.5	neg	M: 719.16089	720.16461	7.6	<1.64E-15	[M-H]-	720.16900	1.2	C36H32O16	MS2	Sagerinic acid
n_4127	62.2	neg	/	515.21332	7.7	2.34E-03	[M-H]-	516.21340	0.5	C24H36O12	MS3	C.cucurbitosideF
p_1908	61.4	neg	Fragment of 449.10553	287.05382	4.0	1.35E-12	[M+H]+	449.10501	6.2	C21H20O11	MS2	Luteolin-4'-O-glucoside
n_4969	61.3	neg	/	717.14574	7.7	<1.64E-15	[M-H]-	718.15340	0.2	C36H30O16	MS3	Salvianolic acid L
p_2660	60.8	neg	Fragment of 394.20529	366.17393	7.2	<1.64E-15	[M+H]+	393.19400	7.3	C24H27NO4	MS3	Tylophorone
p_0676	60.5	neg	/	160.09622	1.3	<1.64E-15	[M+H]+	159.08950	6.2	C7H13O3N	MS3	trans-4-hydroxy-L-proline betaine
p_1252	60.2	neg	/	227.12687	5.1	3.95E-13	[M+H]+	226.12050	4.6	C12H18O4	MS3	12-hydroxyjasmonic acid

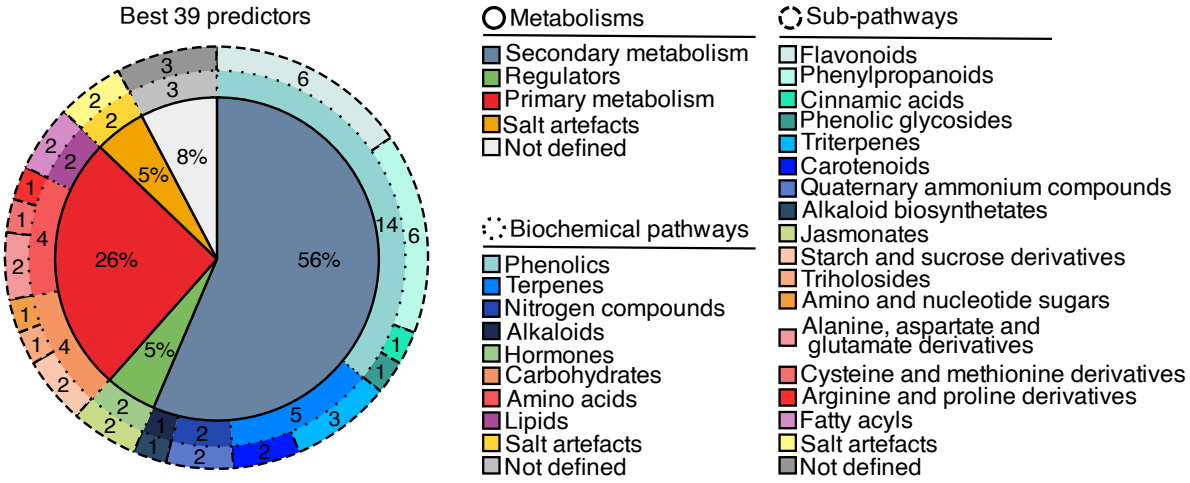

Fig. III.3 | Pathway analysis of the 39 markers. Metabolism, biochemical pathways and sub-pathways were elucidated by screening the KEGG identifiers through MetaboAnalyst, PlantReactome and MetExplore databases. Adapted from Dussarrat et al., 2022, *New Phytologist*.

Besides, phenolics represented the major enrichment observed in Atacama plant species with 14 of the 39 markers. While alkaloids and N-containing compounds (*e.g.* proline betaine, or polyamines combined with flavonoids) were included in the best markers, flavonoid, phenylpropanoid and terpenoid pathways were clearly overrepresented (Fig. III.3 and III.S4). Last but not least, despite their classification into primary or secondary metabolisms, a relevant part of these 39 markers also referred to redox homeostasis owing to their chemical nature or interactions with ascorbate or glutathione pathways (Fig. III.S4), suggesting its importance in the adaptation to hostile environments. Overall, predictive metabolomics reveals that plant metabolism greatly reflects environmental fluctuations in extreme ecosystems, also pinpointed by a core set of metabolites (involved in secondary, primary and redox pathways) capable of predicting at 79% the plant environment independently of the plant species. These findings thus confirm a central place of generic metabolic pathways underpinning plant adaptation to environmental constraints.

Plant metabolome is tailored to environmental constraints. Elevation integrates a wide range of abiotic factors, among which edaphic variables were measured in each of the 19 sampling spots. Besides, climatic variables like temperature, soil water content (SWC, representing the interaction between precipitation and soil properties), precipitations and solar irradiance were measured via two stations (at 3060 and 4090 m.a.s.l). Theoretical values of these factors along the elevation gradient for the 19 environments were predicted considering a linear distribution that was confirmed by field measurements (Fig. III.S5). Elevation greatly correlated with most environmental parameters in the Atacama Desert (Fig. III.S6). Thus, an analysis combining principal component (PCA) and two-way

orthogonal partial least square (O2PLS) was performed to (i) unravel the elevation factor and (ii) highlight the relationship between the 39 best predictors and environmental factors.

First, PCA was used to reveal the influence of the elevation on the climatic and edaphic conditions. The first two components of the PCA model explained 82.4% of the total variance of the dataset (Fig. III.4A) and showed clear discrimination of the plant communities (*i.e.* Prepuna, Puna and Steppe) along a multivariate vector that represented the elevation gradient. Also, a second plane defined by several minerals divided the different environments belonging to the Prepuna ecosystem, which did not occur with Steppe spots. Hence, the previously predicted elevation factor was here depicted by a multivariate vector represented mostly by edaphic variables (*e.g.* pH, P, K), temperature and solar irradiance.

Second, we predicted the covariation between environmental factors and the best 39 markers using an O2PLS analysis (Fig. III.4B). Remarkably, 95% of the variation observed in the environmental dataset was covered by the metabolic features (Table III.S7). Congruently with the correlation matrix and PCA (Fig. III.4), the best predictors were primarily distributed along the first component representing elevation and, to a lesser extent, linked to several edaphic factors like sulphur. The O2PLS biplot further highlighted a remarkable separation between metabolic compounds positively or negatively correlated with elevation along the first plane. In particular, 5 phytochemicals determined by Pearson clustering and including raffinose were plotted opposite of temperature. Hence, these results indicate that the discriminant capacity of elevation primarily resulted from temperature, solar irradiance, SWC and several edaphic factors. These PLS predictions were also confirmed by GLM, where the best predictions using the 39 best markers on independent environmental parameters were obtained for SWC, solar irradiance, pH, and temperature (Fig. III.4C).

Overall, we provide a valuable approach combining metabolomics, GLM and multivariate statistical analyses. Exploiting multiple species in a natural environment can successively unveil generic mechanisms of interest and disentangle complex systems into specific environmental parameters (*i.e.* climatic and edaphic).

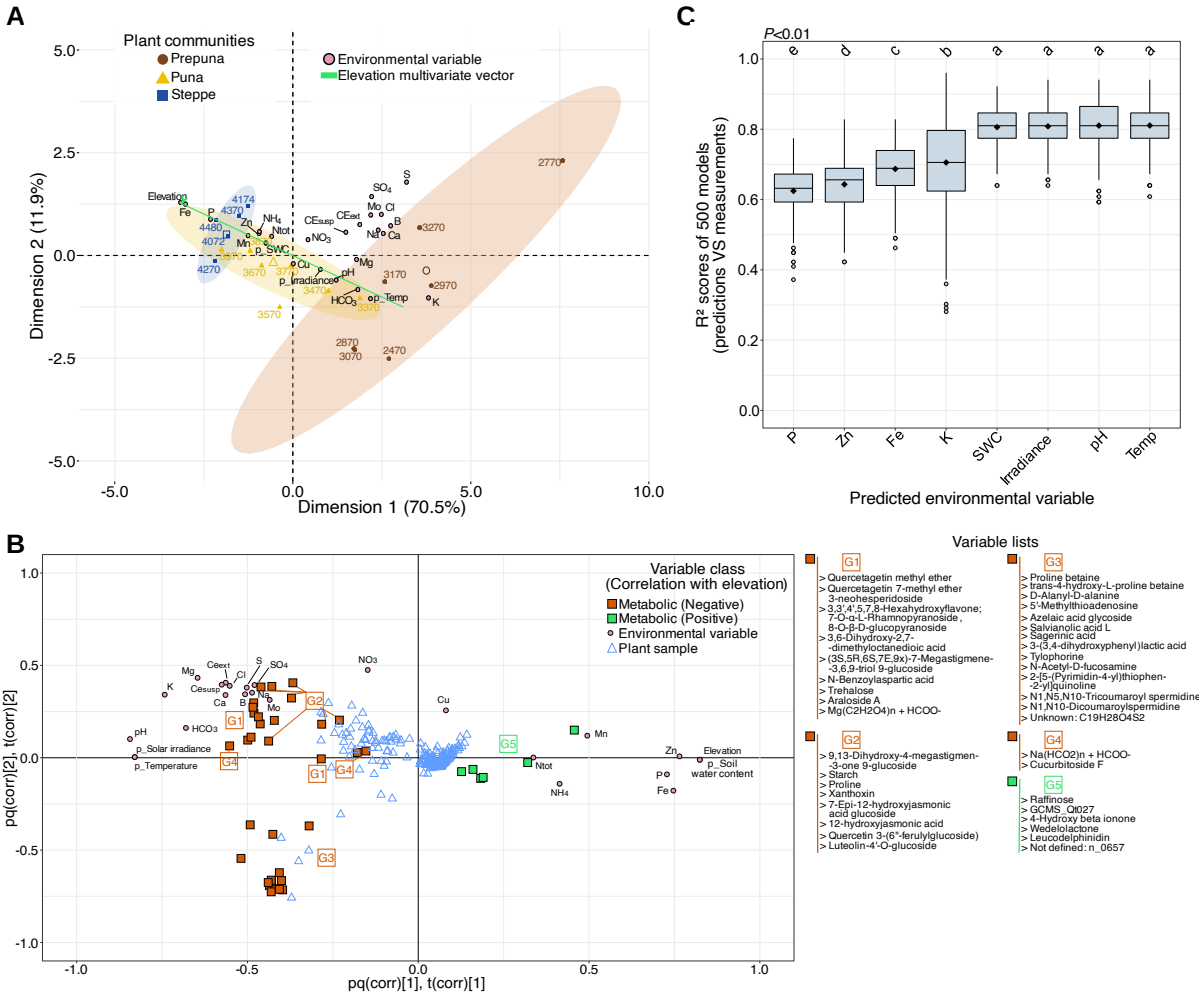

Fig. III.4 | Decomposition of the elevation factor and environment-metabolome covariation. **A.** Principal component analysis biplot. Discrimination of the sampling spots by the environmental data. *SWC* represents the soil water content while *p_* represents a partially predicted parameter. **B.** Two-way orthogonal partial least squares describing the covariation between environmental and metabolic data. Hierarchical clustering analysis was realised with Pearson correlation and Ward algorithm. **C.** Boxplot showing the average R^2 scores (500 models) performed on the best discriminant environmental variables using the 66 best metabolic markers. Letters indicate statistical significance (Tukey's test, $P<0.01$). *Soil W cont* represents the soil water content, *Temp* represents the temperature. The box in each box plot illustrates the lower, median and upper quartile values, and the vertical lines show the range of the R^2 variation in samples while squares and circles represent the mean R^2 and potential outliers, respectively.
Adapted from Dussarrat et al., 2022, *New Phytologist*.

Discussion

Predictive metabolomics demonstrates a generic metabolic toolbox for plant adaptation to extreme habitats. Ecological metabolomics, which allowed to study the interaction between plant metabolism and environment, has attracted scientific curiosity for over 50 years (Sardans *et al.*, 2020). While studies on single plant species led to limited results when transferred to crops, a meta-analysis of individual studies highlighted metabolic convergencies in extreme plant species (Dussarrat *et al.*, 2021), enticing plant researchers to move towards a more holistic approach. Strikingly, our approach combining ecological metabolomics with GLM-based modelling was able to predict the plant environment with an R^2 as high as 0.96 within given species (Fig. III.2B) and 0.79 between species (Fig. III.2). Such values are far above correlation coefficients usually obtained with phenotypic traits (Laughlin *et al.*, 2012; Poorter *et al.*, 2019), which are also more difficult to score (Laughlin & Messier, 2015), making metabolic markers ideal soft traits.

All Atacama plant species harboured low chlorophyll levels as compared to agronomic species, which could result from an adaptive response to high solar irradiance or the meagre availability of other resources like water or nitrogen (Hikosaka *et al.*, 2003), as further illustrated by the very low nitrate and protein contents (Fig. III.S2). Our results suggest that these 24 species, belonging to 14 families, also use a common metabolic toolbox, underpinned by the 39 metabolic markers revealed by our modelling approach, to cope with their environment (Fig. III.2G). Besides, this toolbox is certainly generic as the same metabolic traits were found in several agronomic and ornamental families, validating the ubiquitous nature of this core set. Main differences were observed for flavonoid and terpenoid related predictors, which greatly accumulated in Atacama plants (Fig. III.S7). In addition, levels of raffinose and 4 hydroxy beta-ionone (a compound related to carotenoid degradation) were higher in Steppe species than most temperate species, while Prepuna species accumulated proline derivative compounds. Several markers like quercetin glucoside and coumaroyl-spermidine relatives were only detectable using a lower threshold in agronomic families (Table III.S6). Conversely, several hormone and primary metabolism-related predictors (*e.g.* jasmonates, trehalose) did not present major changes between extreme and non-adapted species, except for chlorophylls and proteins which were lower in Atacama plants. Overall, these observations question the potential adaptive capacity of these agronomic and ornamental species, which naturally develop in mainly temperate regions. The possibility that the genome of agronomic plants would already permit the synthesis of metabolites relevant for plant survival in harsh lands is supported by the presence of *S. chilense* (closely related to the cultivated *S. lycopersicum)* in the 24 Atacama species studied. From an evolutionary point of view, this further suggests that it is easier to modify the regulation of existing metabolic pathways than to create new ones. Also, the high R^2 score prompts the question of how species-specific metabolic adaptations could provide a selective advantage. Hence, the relation between these metabolic markers and the genetic adaptations discovered in Atacama plants (Eshel *et al.*, 2021) deserves further investigation. In

particular, several species colonised a wide elevation gradient, which suggests high plasticity (Fig. III.1). Consequently, metabolic adjustments enabling the acclimation or adaptation to extreme conditions are not necessarily the result of a long evolutionary process.

Involvement of the best metabolic predictors in extreme environment adaptation. The influence of the elevation gradient on metabolic patterns is mainly reflected in a multivariate vector involving temperature, SWC, irradiance and pH (Fig. III.4A). However, extremophiles also face other environmental pressures like the mineral imbalance observed in the Atacama Desert that deserves closer examination (Fig. III.1B). Thus, the success of thriving in the Atacama Desert elevation gradient would result from the ability to (i) cope with daily freezing temperatures at the top of the transect and hyper-aridity and high salinity at the bottom or (ii) manage the balance between carbon input and access to other critical resources such as water.

Starch was the best predictor (Table III.1), while trehalose and raffinose were among the top predictors, confirming a suitable place for carbohydrates in the resilience mechanism (Fig. III.S4). In the Atacama Desert, solar irradiance is not a limiting factor for carbon entry, and even threatens plant survival (Eshel *et al.*, 2021). The shallow protein level observed in all Atacama plant species (Fig. III.S2) suggests that plant growth is very low (Elser *et al.*, 2008). Therefore, carbon that does not fuel plant growth and protein turnover could be transiently stored as starch or allocated to protective systems against oxidative stress induced by other environmental factors such as water availability, temperature and salinity. Starch, whose metabolism is known to play a major role against abiotic stress (Thalmann & Santelia, 2017), can be used as a carbon source for the synthesis of protective compounds when environmental conditions become harsher, while its accumulation could be linked to sodium scavenging in halophytes (Thalmann & Santelia, 2017), for instance. The strong negative correlation of starch with elevation would result from a trade-off with the production of osmolytes and other protective compounds required in the highest elevation levels, where daily freezing temperatures occur. Alternatively, the lower efficiency of transitory starch remobilisation under cold temperature could explain low starch contents at high levels. Conversely, raffinose negatively correlated with temperature, validating the central place of Raffinose Family Oligosaccharides in cold tolerance (Vinson *et al.*, 2020). Still, several predictors were fatty acyls within the primary metabolism and jasmonates, which supported the role of lipid remodelling in extreme environments (Cao *et al.*, 2016; Dussarrat *et al.*, 2021).

More than half of the 39 best markers were involved in secondary metabolism (Fig. III.3). Remarkably, phenolics (14/39 compounds) were increased at lower elevation levels, which would help plants cope with both the very low water availability and high salinity of these lands. This protective process has already been described in multiple extremophile plant species (Dussarrat *et al.*, 2021). Phenolic antioxidant properties, mainly for cinnamic acid and quercetin derivatives (extensively represented in the best predictors), enhance photoprotection and resilience to abiotic stresses (Agati &

Tattini, 2010). Regarding terpenoids, the presence of xanthoxin and 4-hydroxy beta-ionone (a carotenoid degradation product) within the 39 best markers supported the role of carotenoids *per se* as well as their degradation in extreme climate resilience (Table III.1). Despite their well-described antioxidant role, exciting studies have discussed the link between the catabolism of carotenoids and plant defence, as their cleavage leads to hormonal compounds (*e.g.* abscisic acid) or redox signalling (Havaux, 2014). Also, the accumulation of N-related compounds could be attributed to their role in osmoregulation (*e.g.* proline betaine). The contribution of phenolics that conjugate polyamines and other molecules (*e.g.* tricoumaroyl spermidine) is more complex, despite a growing body of evidence for their implication in stress mitigation (Pál *et al.*, 2018). Most importantly, our study linked plant survival under harsh conditions to redox metabolism since the majority of metabolic markers directly or indirectly involves redox homeostasis. This was exemplified by primary compounds of the glutathione and ascorbate pathways, metabolites for ROS processing including carotenoids, as well as potential links between the biosynthesis of several compounds and NAD metabolism (Fig. S4), all participating in oxidative stress signalling (Decros *et al.*, 2019; Dussarrat *et al.*, 2021). Among the best predictors is proline, whose accumulation in response to osmotic stress has been widely documented and recently attributed to redox homeostasis (Szabados & Savouré, 2010). Alternatively, accumulated levels of amino acids could serve as metabolic intermediates for the synthesis of more complex secondary metabolites with stress-responsive functions.

Overall, uncovering the metabolic characteristics of Atacama species highlighted (i) the linear encapsulation of environmental fluctuations by the plant metabolome (involving primary, secondary and redox pathways) and (ii) the use of generic metabolic mechanisms to adapt to extreme growth conditions. Such an approach (multi-species harvested in extreme environments) offers promising perspectives in both ecological chemistry and stress physiology worldwide. A fascinating perspective will be to research the genetic and molecular mechanisms that control the levels of these metabolic markers.

Acknowledgements

We acknowledge the GenOuest bioinformatics core facility (https://www.genouest.org) and the Genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul, https://doi.org/10.15454/1.5572369328961167E12) for providing computing infrastructures and resources. We also acknowledge the Comunidad de Talabre. We are grateful to Christiane Gallet, Guillaume Tcherkez and Fabien Jourdan for their analytical advice, and MetaboHUB (ANR-11-INBS-0010) and PHENOME (ANR-11-INBS-0012) for financial support to INRAE. We thank Fondo de Desarrollo de Áreas Prioritarias (FONDAP) Center for Genome Regulation (15200002), Fondo

Nacional de Desarrollo Científico y Tecnológico (FONDECYT 1180759), DOI EVONET Proyect and ANID – Millennium Science Initiative Program – iBio ICN17_022 for financial support to R.A.G.

Author contributions

T. Dussarrat, Y. Gibon, D. Rolin, R. Gutiérrez and P. Pétriacq designed and planned the project. T. Dussarrat, C. Latorre, F. Díaz and R. Gutiérrez performed the fieldwork. T. Dussarrat, S. Prigent, P. Pétriacq, S. Bernillon, A. Flandin, C. Cassan, P. Van Delft, K. Varala and J. Joubes conducted metabolic or bioinformatic experiments and analyses. D. Jacob uploaded all data online. T. Dussarrat, Y. Gibon, D. Rolin, R. Gutiérrez and P. Pétriacq integrated and analysed results of all the experiments. T. Dussarrat, R. Gutiérrez and P. Pétriacq wrote the paper with feedback of all the co-authors.

Data Availability

The metabolic dataset and all metadata were deposited online using Dataverse INRAE (https://dx.doi.org/10.15454/UUBXIF) and following the Findable, Accessible, Interoperable, Reusable (FAIR) principles (Jacob *et al.*, 2020).

All references are available at the end of the thesis manuscript.

Supplemental figures. Supplemental tables are available at the following link until publication:

https://drive.google.com/drive/folders/1Z3HLMY0Hb281HEu56MM82MHtY_YQ9tkE?usp=sharing

New Phytologist Supporting Information. Predictive metabolomics of multiple Atacama plant species unveils a core set of generic metabolites for extreme climate resilience.

Thomas Dussarrat, Sylvain Prigent, Claudio Latorre, Stéphane Bernillon, Amélie Flandin, Francisca P. Díaz, Cédric Cassan, Pierre Van Delft, Daniel Jacob, Kranthi Varala, Jérôme Joubes, Yves Gibon, Dominique Rolin, Rodrigo A. Gutiérrez, Pierre Pétriacq. 28 February 2022.

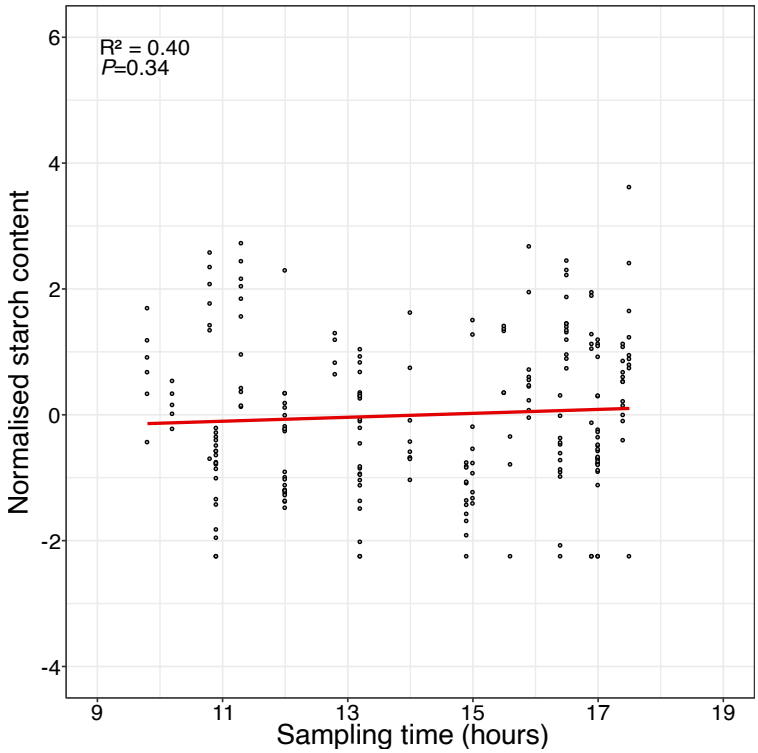

Fig. III.S1 | Correlation between starch content and sampling time. Absence of correlation between the time (in hours) of sampling and starch content in plant samples (Pearson correlation). Adapted from Dussarrat et al., 2022, *New Phytologist*.

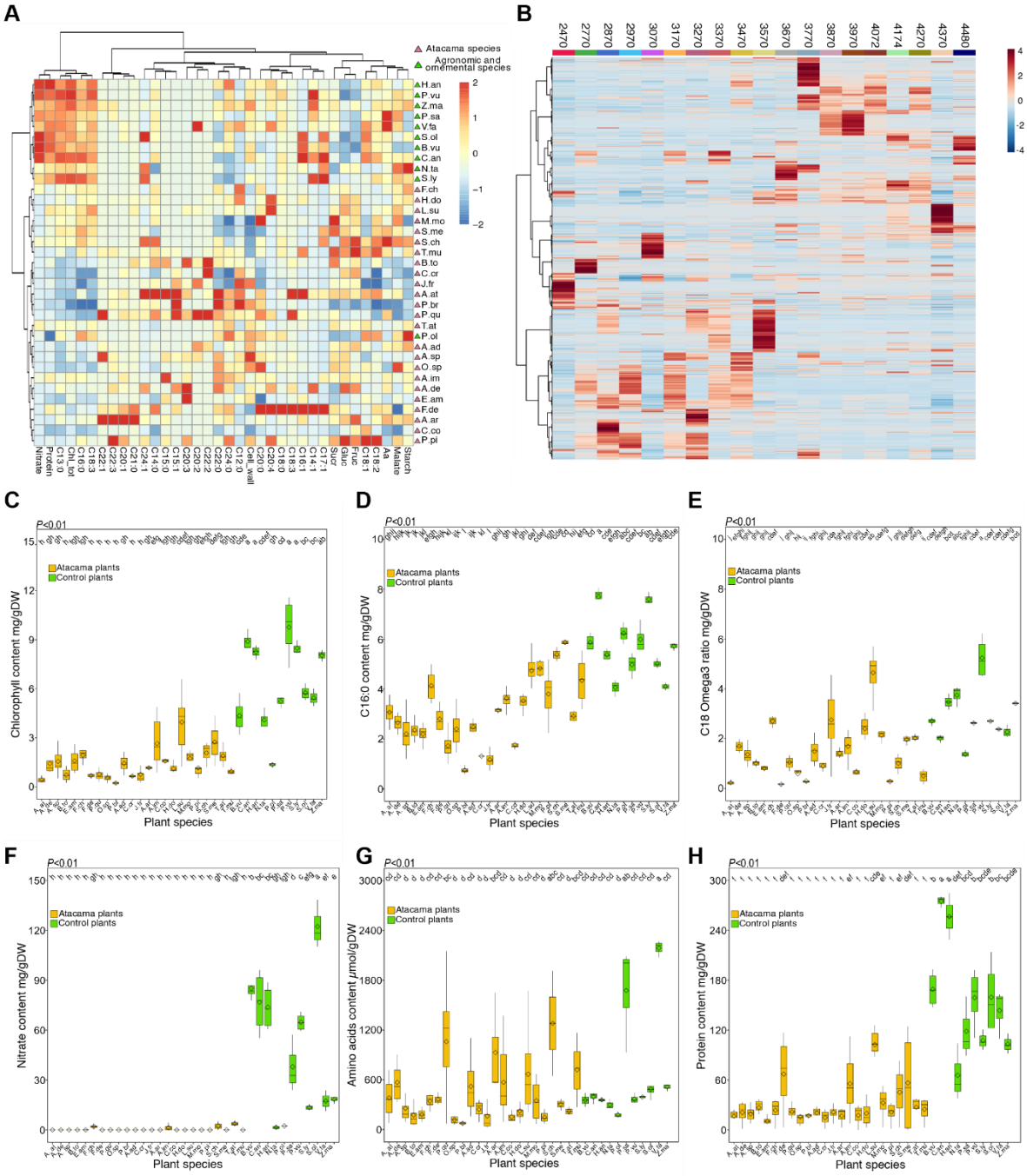

Fig. III.S2 | Changes in major compounds in Atacama plants. A. Two-dimensional clustering analysis (Pearson correlation, Ward algorithm) showing a heatmap of major compounds detected by targeted assays and GCFID of 24 Atacama plants *versus* 11 agronomic or ornamental species. **B.** Biochemical diversity observed in the 19 sites using targeted and untargeted analyses. Only significant variables were used (Tukey's test with $P<0.01$) **C-H.** Details of the major discriminant compounds. Tukey's test were performed with $P<0.01$. The box in each box plot illustrates the lower, median and upper quartile values, and the vertical lines show the range of the concentration variation in samples while squares and circles represent the mean concentration and potential outliers, respectively. Adapted from Dussarrat et al., 2022, *New Phytologist*.

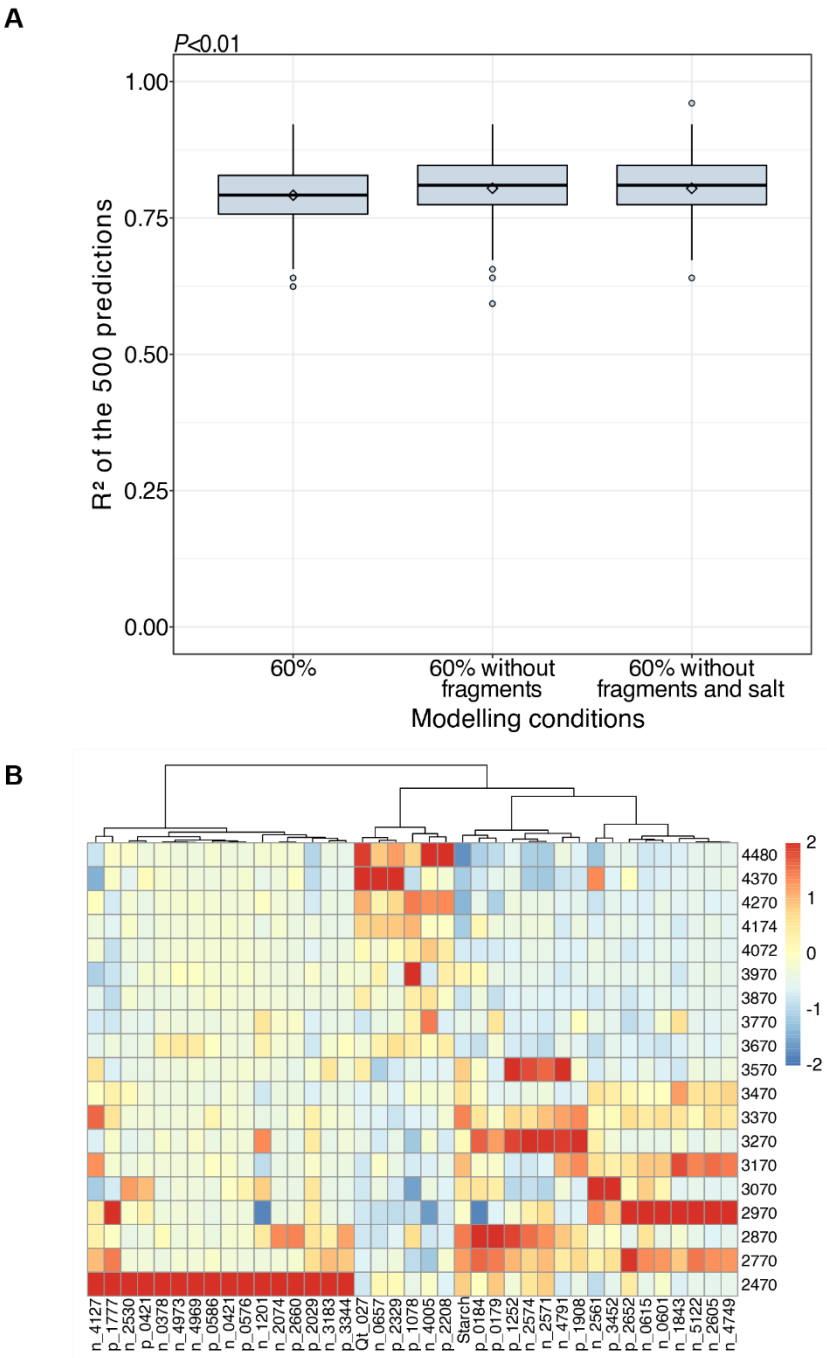

Fig. III.S3 | Best metabolomics predictors in Atacama plants. A. Boxplot showing R^2 of the 500 models performed on the best metabolic markers (threshold 60%) with or without fragments and Formic acid + salt compounds. The box in each box plot illustrates the lower, median and upper quartile values, and the vertical lines show the range of the R^2 variation in samples while squares and circles represent the mean R^2 and potential outliers, respectively. **B.** Clustering analysis (Pearson correlation, Ward algorithm) with a heatmap illustrating the relation between best metabolic predictors and elevation. Adapted from Dussarrat et al., 2022, *New Phytologist*.

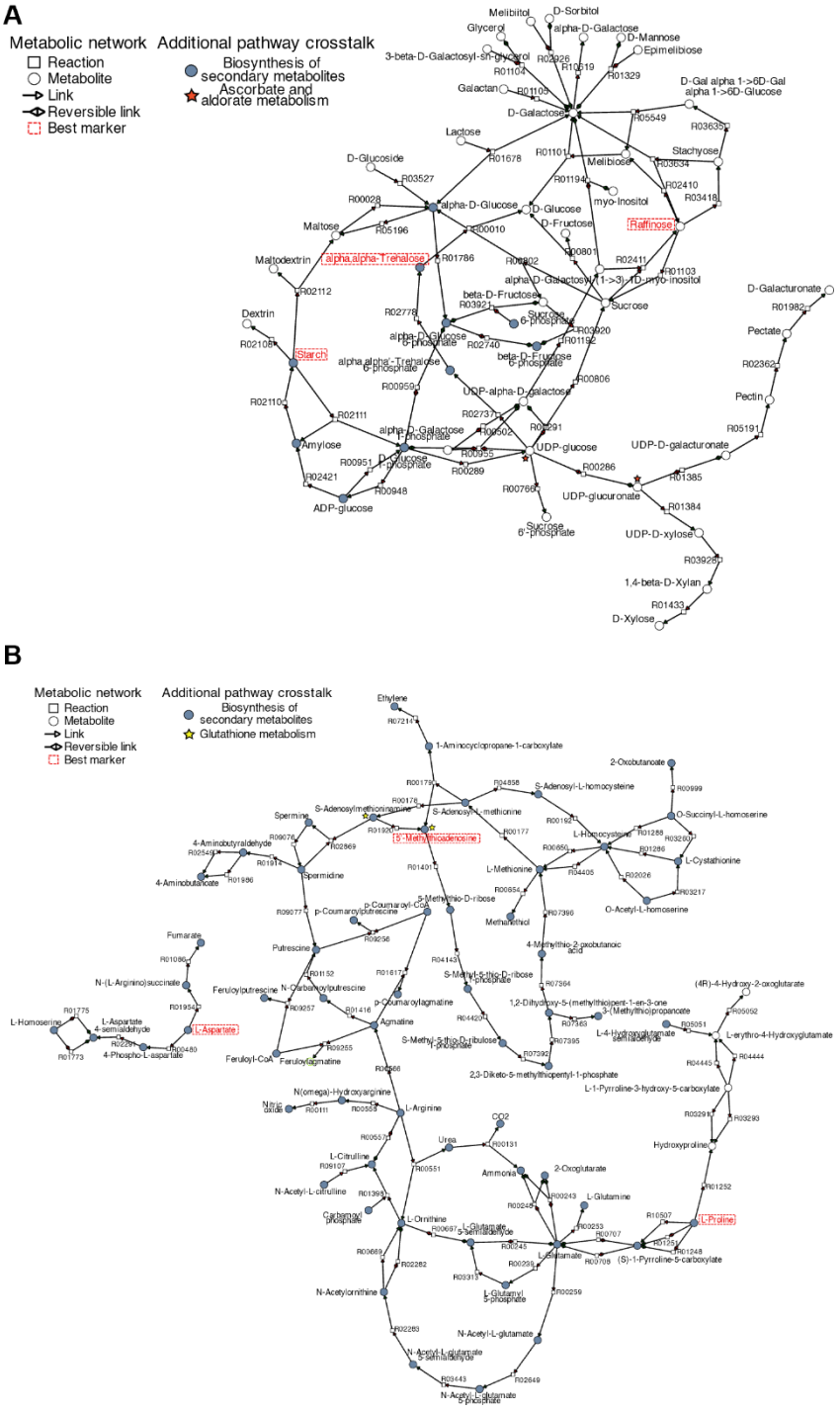

Fig. III.S4 | Metabolic networks. Best markers were mapped into a pre-existing *A. thaliana* metabolic network using MetExplore. **A.** Crosstalk between the sub-pathways "Starch and sucrose metabolism" and "Galactose metabolism". **B.** Crosstalk between the sub-pathways "Cysteine and methionine derivatives", "Arginine and proline derivatives" and "Alanine, aspartate and glutamate derivatives". **C.** "Flavonoid biosynthesis" sub-pathway. **D.** "Carotenoid biosynthesis" sub-pathway. **E.** "Phenylpropanoid biosynthesis" sub-pathway.
Adapted from Dussarrat et al., 2022, *New Phytologist*.

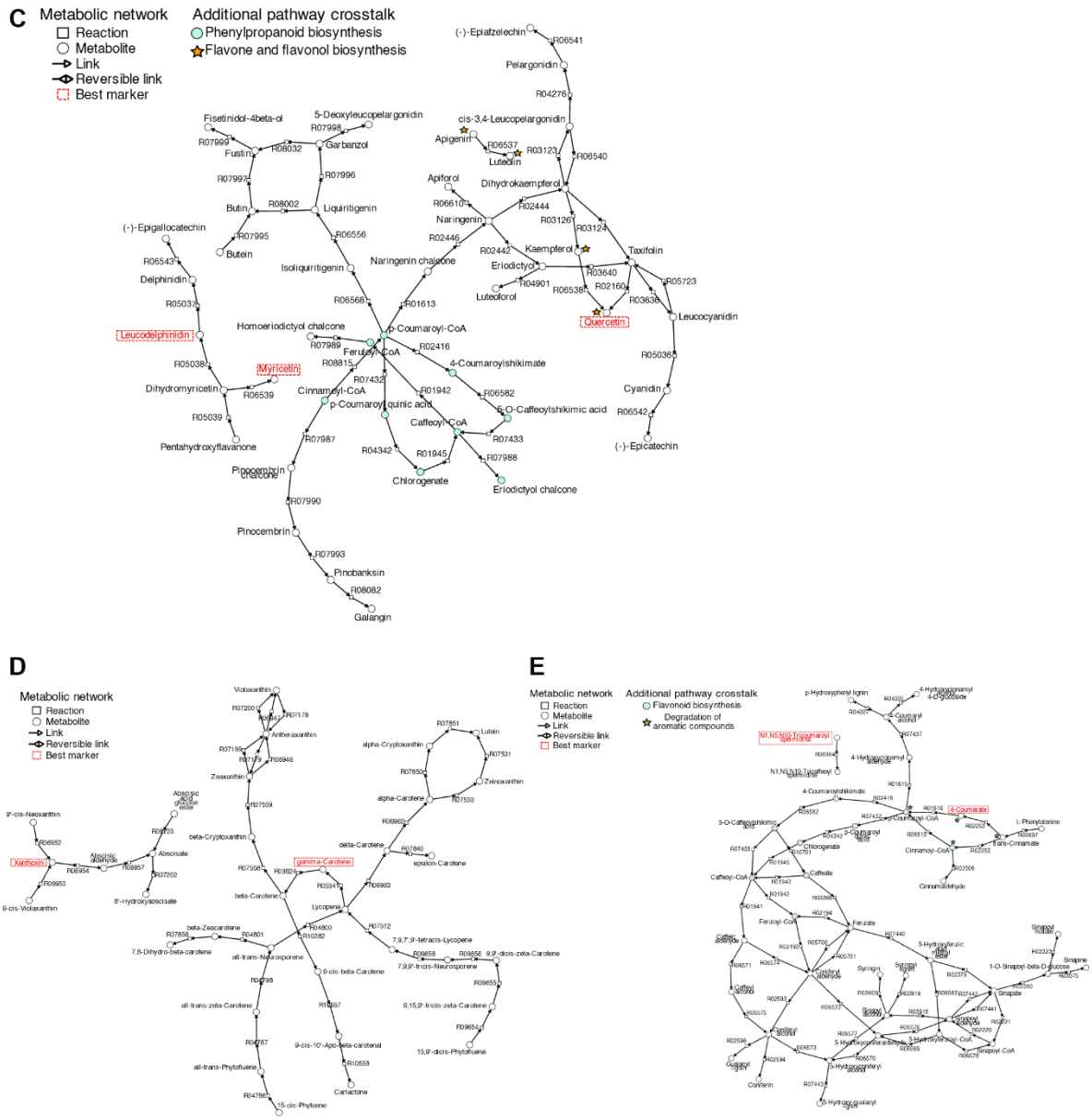

Fig. III.S4 | Metabolic networks. Best markers were mapped into a pre-existing *A. thaliana* metabolic network using MetExplore. **A.** Crosstalk between the sub-pathways "Starch and sucrose metabolism" and "Galactose metabolism". **B.** Crosstalk between the sub-pathways "Cysteine and methionine derivatives", "Arginine and proline derivatives" and "Alanine, aspartate and glutamate derivatives". **C.** "Flavonoid biosynthesis" sub-pathway. **D.** "Carotenoid biosynthesis" sub-pathway. **E.** "Phenylpropanoid biosynthesis" sub-pathway.
Adapted from Dussarrat et al., 2022, *New Phytologist*.

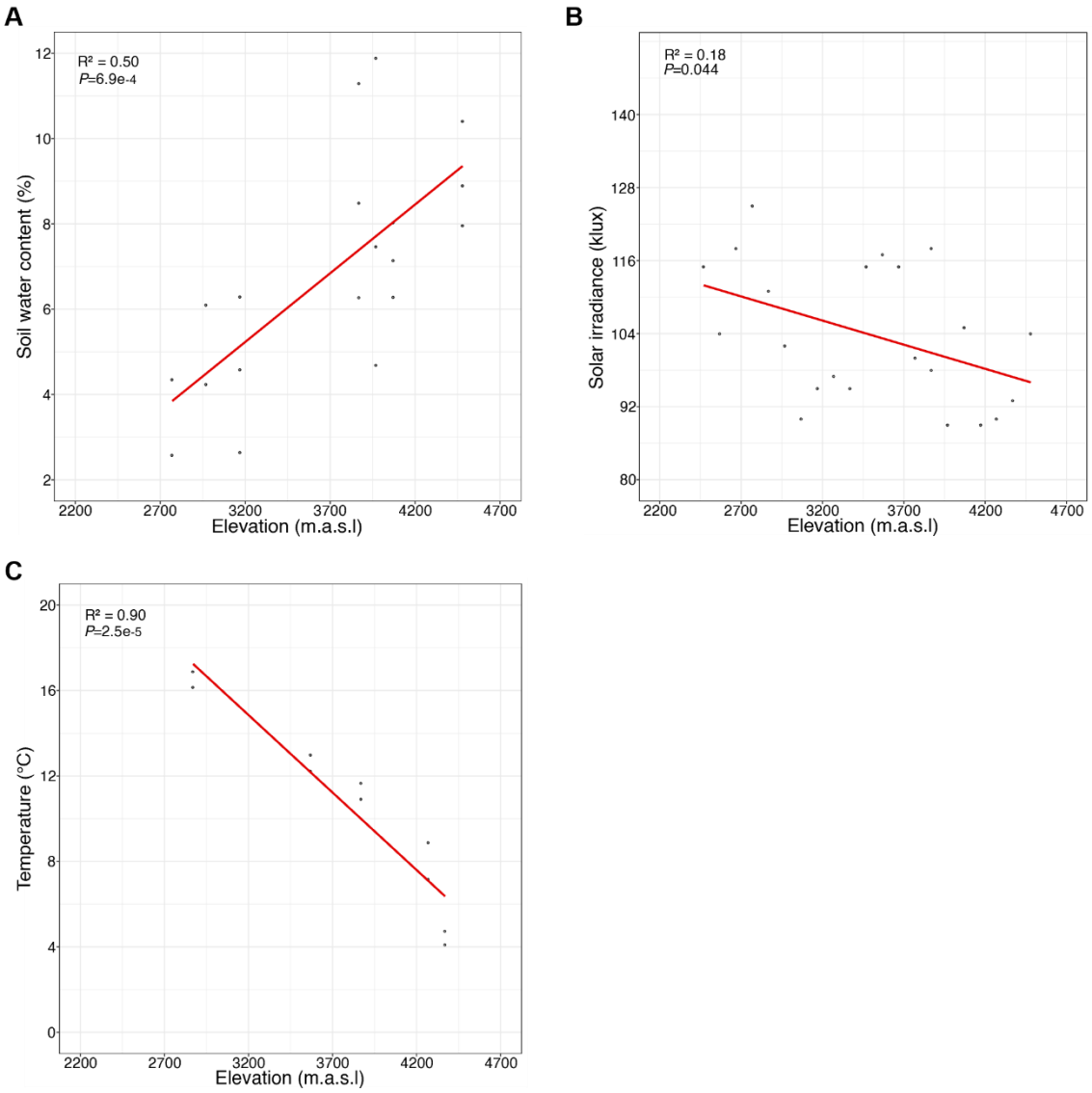

Fig. III.S5 | Validation of the environmental prediction. Distribution of (**A**) soil water content, (**B**) solar irradiance and (**C**) temperature along the elevation gradient (Pearson correlation, $P<0.05$). Soil samples were collected near the plants of interest. The moisture content of the soil was determined by dividing the fresh weight by the dry weight after one week of freeze drying. Solar irradiance was measured using a lux meter near the plants of interest. Temperature was measured using two thermometers. Adapted from Dussarrat et al., 2022, *New Phytologist*.

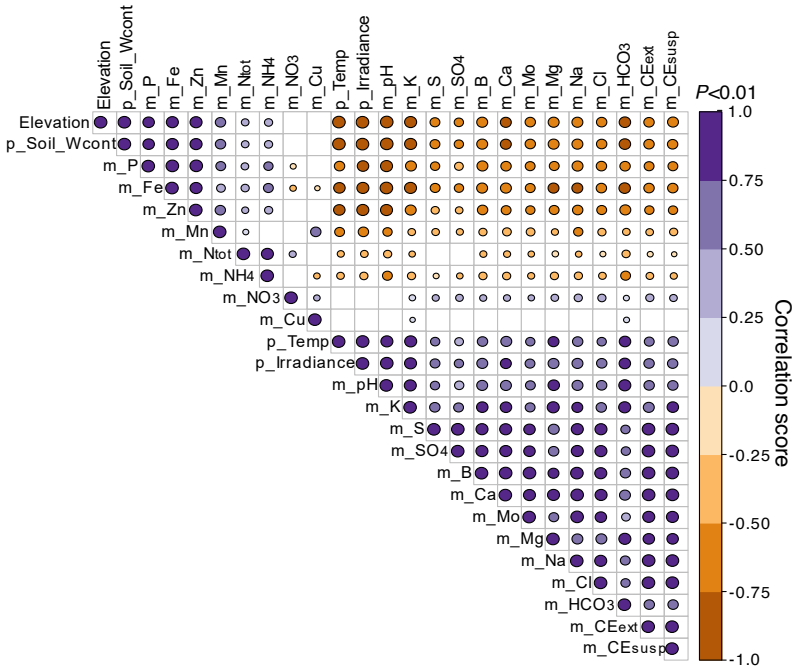

Fig. III.S6 | Decomposition of the elevation parameter. Correlation plot of the environmental data (*P*<0.01, Pearson correlation). Adapted from Dussarrat et al., 2022, *New Phytologist*.

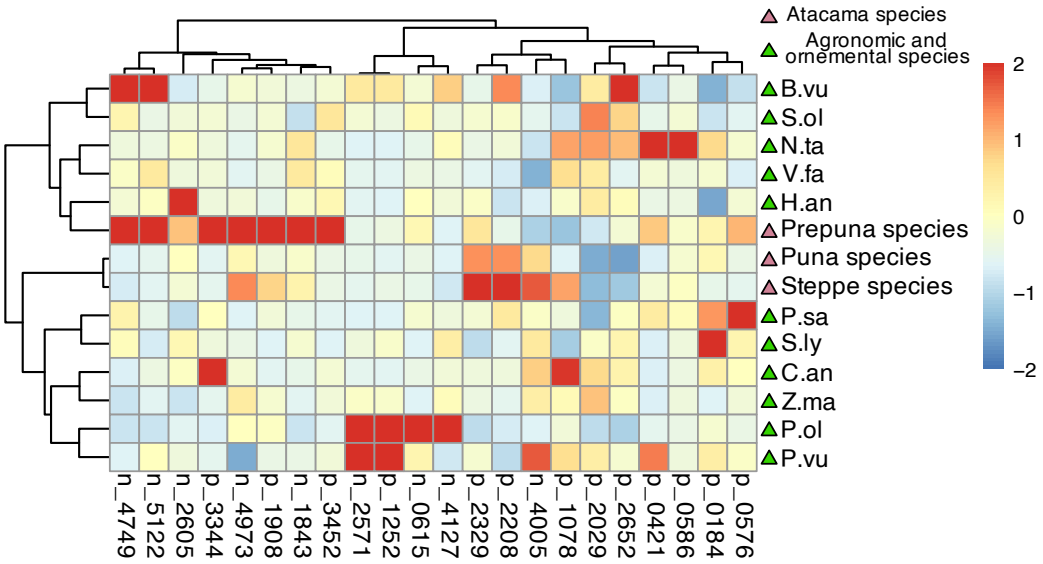

Fig. III.S7 | Best metabolic predictors in agronomic and ornamental plant species. Clustering analysis of the best metabolic predictors between Atacama and agronomic and ornamental plant species (Pearson correlation, Ward algorithm).
Adapted from Dussarrat et al., 2022, *New Phytologist*.

III. CONVERGENT STRATEGY TO FACE HARSH CONDITIONS

The combination of eco-metabolomics with machine learning highlighted the excellent integrative capacity of plant metabolome under extreme environmental conditions. This was first exemplified by the possibility of predicting the plant environment from metabolic traits with 90% accuracy in species thriving across a 500m elevation gradient. At the ecosystem level, our approach allowed the discovery of a generic toolbox composed of 39 metabolites predicting plant environment with 79% accuracy, independently of the plant species and sampling year. The high prediction rates in both global and species-specific analyses suggest that predictive metabolomics is a valuable technique to unveil adaptive metabolic markers. Besides, predictions greatly exceed previous results using phenotypic traits (Laughlin et al., 2012). Predictive metabolomics could therefore be deployed to uncover highly predictive soft traits for different purposes in agronomy or ecology (Fig. III.5). However, care is required since the biological interpretation of these markers depends on the biological level studied (*i.e*. from organism to ecosystem level). Here, predictive modelling at the species level was likely to highlight the best adaptive metabolic strategy for a given genome and environment. Meanwhile, a global scale modelling approach would pinpoint the generic metabolic mechanisms enabling Atacama plants to cope with the limiting abiotic constraints (*e.g*. temperature, solar irradiance, water availability).

Findings raise some troubling points when it comes to interpreting results at the evolution level. Predictions of the plant environment from 39 metabolites with 79% accuracy support the hypothesis that adaptation of Atacama plants is mediated by generic metabolic mechanisms (Fig. III.5). While previous reviews described similar chemical patterns among distinct plant species, our approach unveiled for the first time a generic metabolic toolbox with potential adaptive property for multiple plant species in the Atacama Desert. Besides, the underlying metabolites were congruent with previous studies. For instance, low chlorophyll levels were observed in all Atacama plants but also in plants from the Himalayan mountains to probably mitigate the impact of high solar irradiance (Cui et al., 2019). High levels of raffinose and proline were respectively detected in plants growing at high (negative temperature) or low (high salinity) elevations, as described in cold, saline or arid lands (Lugan et al., 2010; Strimbeck et al., 2015). Additional protective compounds like quercetin and jasmonates as well as the over-expression of protective reactions (*e.g*. 4-hydroxy beta-ionone from carotenoid degradation) were also related in other works (Dussarrat et al., 2021). By extension, the capacity of Atacama species to span the elevation gradient would depend on their capacity to regulate the levels of these different metabolic markers. Accordingly, some related compounds have been shown to evolve in a similar manner in other elevation clines (Kumari et al., 2020).

Last but not least, all predictors were detected in multiple agronomic and ornamental plant species, emphasising their potential for engineering resilient crop species (Fig. III.5). Altogether, the predictive capacity of these metabolites and their presence in agronomic plants suggest that adaptation

may lie in the regulation of pre-existing pathways rather than in the development of new ones. Hence, the analysis of genetic traits fixed through the evolutionary process is all the more crucial to confirm the existence of convergent evolutionary trajectories and the significance of these metabolic markers in adaptation.

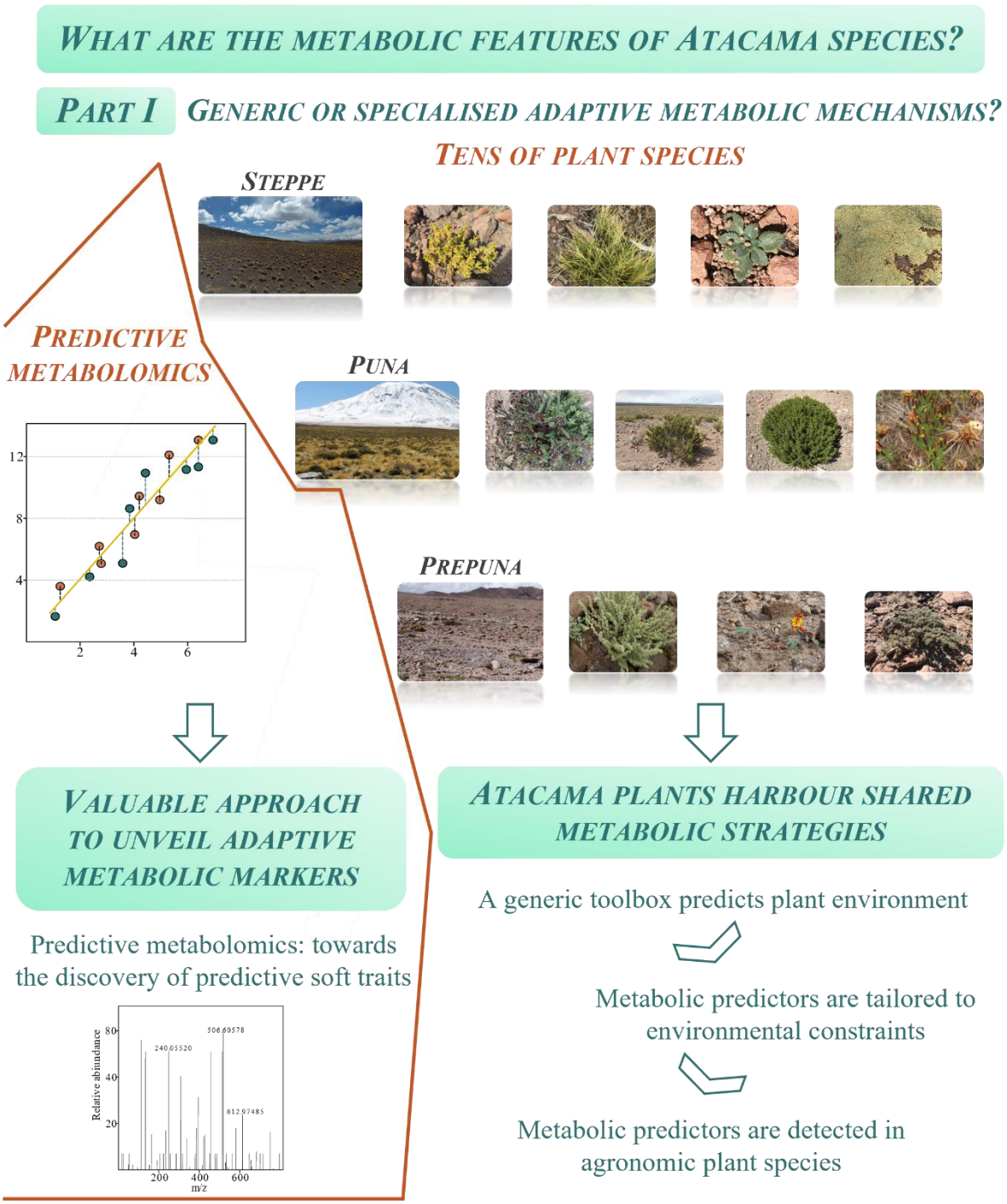

Fig. III.5 | The secrets of plant adaptation to the Atacama Desert.

Chapter 3: The use of a generic metabolic toolbox.

CHAPTER 4

CONVERGENT CHEMICAL EVOLUTION IN ATACAMA PLANT SPECIES

I. ENRICHMENT ANALYSES IN EXTREME PLANT SPECIES

This chapter has resulted in a scientific paper that will be submitted to the Journal of Experimental Botany in the following weeks. Here, we proposed a short introduction (Section I) to set the topic in the context of the PhD. The manuscript of this article was then implemented in a Word format in Section II. Finally, we provided a short conclusion that summarised the contribution of this study to the thesis project (Section III).

All supplemental tables are available at the following link:
https://drive.google.com/drive/folders/1Z3HLMY0Hb281HEu56MM82MHtY_YQ9tkE?usp=sharing

The genetic diversity of wild relatives of domesticated crops was used to enhance stress resilience in various crop varieties (Castañeda-Álvarez et al., 2016). This approach is even more meaningful when applied to plants growing in extreme biomes, through the analysis of expanded gene families, for instance (Bolger et al., 2014; Kang et al., 2020). As an example, an intriguing shift in gene expression patterns was observed in *Solanum pennellii*, a stress-tolerant species related to *Solanum lycopersicum*, and linked to a change in lipidic profile (Bolger et al., 2014). However, most works were performed on a limited number of species and used pairwise comparisons evaluating the genomic variations between one crop species and its closest sequenced wild species (Kashyap et al., 2020). Although these discoveries provided interesting insights, the resulting adaptive molecular markers were mainly species-specific (Dussarrat et al., 2021). Thus, adaptation to extreme lands is primarily considered as the result of a plant lineage that strongly depends on evolutionary processes (Chae et al., 2014; Scossa and Fernie, 2020). Conversely, recent studies suggested a fascinating role of generic molecular and metabolic mechanisms in the adaptation of multiple plant species in the Atacama Desert, the driest non-polar desert of the earth (Eshel et al., 2021; Dussarrat et al., 2022). For instance, 39 metabolic compounds, which encompassed primary compounds such as trehalose or proline and secondary compounds like quercetin, predicted the plant environment with 79% accuracy (Dussarrat et al., 2022). Hence, these recent findings offered strong support to the existence of a common metabolic strategy employed by Atacama plants to face harsh abiotic constraints, but the underlying genetic legacies remained unknown. Our approach sought to identify shared genetic traits developed through the evolutionary process by evaluating the reaction and pathway enrichments within 32 Atacama plant species compared to their phylogeneticd relative species. Besides, this analysis aimed to characterise whether these genetic legacies converged to the regulation of similar or divergent biochemical reactions, and therefore potentially validate the role of the previously described metabolic markers in adaptation. Computational analysis was deployed using PathwayTools and Fisher's exact test (Karp et al., 2021; Wieder et al., 2021) to explore gene expansion and gene expression patterns, which highlighted a set of reactions enriched in more than 50% of the plant species. These naturally selected reactions were tightly linked to environmental constraints and related to the predictive metabolic toolbox.

II. SHARED GENETIC LEGACIES GOVERNING PRECISE CHEMICAL PATHWAYS

The following article will be submitted to the Journal of Experimental Botany in the following weeks.

Phylogenetically diverse wild plant species use common biochemical strategies to face harsh abiotic conditions in the Atacama Desert.

Thomas Dussarrat[1,2], Ricardo Nilo-Poyanco[3], Tomás Moyano[1], Sylvain Prigent[2,4], Francisca P Díaz[1], Claudio Latorre[5,6], Viviana Araus[1], Guillaume Decros[2], Yves Gibon[2,4], Dominique Rolin[2,4], Pierre Pétriacq[2,4] and Rodrigo A Gutiérrez[1]*

[1]: Departamento de Genetica Molecular y Microbiologia, Pontificia Universidad Catolica de Chile. FONDAP Center for Genome Regulation and ANID Millennium Science Integrative Program-Millenium Institute for Integrative Biology (iBio), Av Libertador Bernardo O'Higgins 340, Santiago, Chile

[2]: Univ. Bordeaux, INRAE, UMR1332 BFP, 33882 Villenave d'Ornon, France

[3]: Universidad Mayor, Escuela de Biotechnologia, Facultad de Ciencias, Santiago, Chile

[4]: Bordeaux Metabolome, MetaboHUB, PHENOME-EMPHASIS, 33140 Villenave d'Ornon, France

[5]: Departamento de Ecología, Pontificia Universidad Católica de Chile, Av Libertador Bernardo O'Higgins 340, Santiago, Chile.

[6]: Institute of Ecology and Biodiversity (IEB), Las Palmeras, 3425, Ñuñoa, Santiago, Chile.

* Correspondence: rgutierrez@bio.puc.cl

Abstract.

Climate change is a serious threat to global agriculture. The best ideotype yields are under monitoring pressure due to increased aridity in many parts of the world. Developing new resilient varieties is urgent to face this threat. Understanding conserved molecular mechanisms in wild plants is key to develop new strategies for sustainable agriculture. Yet our knowledge of wild species is scant, particularly in extreme environments. We performed pathway and reaction enrichment analysis to understand the biochemical commonalities and differences of wild plant species in the Atacama Desert. To gain insights into the mechanisms that ensure plant survival in this extreme environment, we compared gene expansion and expression patterns between the annotated reactions from 32 Atacama plant species and 32 phylogenetically related plant species that do not live in Atacama. We found significant biochemical convergences in primary, secondary and redox metabolism characterised by reactions enriched in at least 50% of the species, independent of the plant lineage. Analysis of the annotation indicated potential advantages against drought, salinity, high solar irradiance and nitrogen starvation. These findings suggest adaptation in the Atacama Desert may result from shared genetic legacies governing the expression of key metabolic pathways to face harsh environmental conditions. Enriched reactions referred to ubiquitous compounds common to extreme and agronomic species. Hence, genes underlying these adaptive traits offer promising perspectives for improving abiotic stress resilience in crop species.

Keywords: Plants; metabolism; enrichment analysis; convergent mechanisms; extreme ecosystems; Atacama Desert; multi-species.

Introduction

Plants are sessile organisms which rely on the availability of local resources to live. Conversely, deficiencies or excesses can severely compromise plant growth, development and fitness (Fernandez et al., 2016; Prevéy et al., 2019). Substantial efforts in plant breeding programs have focused on developing increasingly productive and resistant ideotypes to environmental constraints such as drought (Voss-Fels et al., 2019). However, the accelerated changes in climate due to global warming is rapidly making these genetic improvements obsolete, resulting in a stagnation of crop yields worldwide (Bailey-Serres et al., 2019). Interestingly, a few wild plant species flourish under very harsh environmental conditions, representing a unique reservoir of adaptive mechanisms (Díaz et al., 2019). Random genetic mutations have tailored the plant genome to extreme ecosystems such as deserts, providing a selective advantage against high light intensities and nutrient or water deficiencies, for instance (Bolger et al., 2014). Understanding plant survival strategies from extreme environments could help unravel sustainable resistance mechanisms that would greatly benefit global food security. Although promising, studies on extreme wild plant species are scant and mostly performed under controlled laboratory conditions. Those settings lack ecological context and may hinder the identification of adaptive traits that are expressed only under natural conditions (Dussarrat et al., 2021).

Plants evolved temporal and spatial strategies to optimise the balance between development and defence in hostile ecosystems. For instance, multiple species avoid the harshest periods by modulating their life cycle and phenology (Prevéy et al., 2019). In contrast, the survival of plants that can not adapt in time depends on distinct genetic mechanisms underlying metabolic processes and adjustments as osmoregulation or tight management of the lipidic profile, respectively (Bolger et al., 2014; Turner, 2018). However, most of this knowledge is based on the study of one or a limited number of species. Understanding conserved mechanisms that act under natural conditions may increase the likelihood of success when transferred to other species, for example for crop improvement (Turner, 2018). Recent analyses on several hundred of thousands of plant species indicated that the appearance of several phenotypic traits strongly correlated with environmental variations (*e.g.* latitudinal gradient) worldwide (Joswig et al., 2021). Besides, the adjustment of phenotypic traits in response to abiotic stresses such as light intensity seemed to converge between species from different evolutionary trajectories (Poorter et al., 2019). Other classical examples of convergent evolution in plants, also related to the response to extreme conditions, are C4 and CAM photosynthesis (Edwards, 2019). Therefore, it is of interest to determine the extent to which convergent evolution occurred between molecular mechanisms. We have evidence that this can happen. A recent study from our group unveiled 265 positively selected genes (PSGs) from 32 species from the Atacama Desert, the driest non-polar desert on Earth (Eshel et al., 2021). These genes encompassed various molecular processes related to the protection against high solar irradiance, nitrogen starvation and osmotic stress, with a great part of these PSGs being shared among different plant lineages. Finally, top-expressed genes in the Atacama species were related to primary

and secondary chemical compounds (Eshel et al., 2021). Hence, the exciting possibility of a strong influence of shared adaptive genetic processes in adaptation and an intriguing over-expression of genes related to protective metabolite synthesis emerged (Eshel et al., 2021). These results thus pinpoint a potential existence of selected generic mechanisms managing key metabolic processes that govern plant life under major abiotic threats. However, the exact nature, role, and evolutionary trajectories of the enriched chemical reactions and pathways in Atacama plant species remain unclear.

Studies demonstrated that plant metabolism is an excellent predictor of environmental variation in harsh biomes (Fiehn, 2002; Kumari et al., 2020). For instance, the orchestration of primary and secondary metabolism led to the accumulation of amino acids (as precursors of secondary compounds), phenolics (*e.g.* flavonoids as quercetin), isoprenoids (*e.g.* carotenoids) and nitrogen-containing metabolites (*e.g.* quaternary ammonium compounds such as proline and glycine betaine) (Lugan et al., 2010; Dussarrat et al., 2021). Besides, while few studies performed an ecological metabolomic approach using multiple species, promising results demonstrated a significant correlation between phytochemical diversity and environmental variation, with secondary metabolism at the core of plant performance (Defossez et al., 2021). Interestingly, recent work from our group used a coprehensive analysis of 24 plant species thriving in the Atacama Desert to discover a metabolic toolbox composed of 39 metabolites predicting plant environment independently of the plant lineage. These generic predictors were also detected in agronomic plant species, raising great hope for their use in engineering crop resilience to harsh abiotic constraints (Dussarrat et al., 2022). However, the genetic traits governing the modulation of this generic toolbox remain unknown.

As detailed before, the Atacama Desert is a highly challenging environment where intensities of abiotic stresses reach extremes of the plant life-compatible gradients (Eshel et al., 2021). The onset of hyperaridity cycles 12 million years ago prompted the development of the Atacama Desert, which is currently characterised by extremely low precipitations (20 to 160 mm/year) and high solar irradiance (600 W/m²/d) compared to other deserts or high mountain ecosystems (Báez and Collins, 2008; Zhang et al., 2010; Jordan et al., 2014; Díaz et al., 2016; Ziaco et al., 2018). Besides, plants face extremely low levels of nutrients such as nitrogen and high levels of salinity (Eshel et al., 2021). Hence, the Atacama Desert offers opportunities to uncover molecular mechanisms determining plant performance under extreme conditions, from the genome to the metabolome. Development of a computer pipeline has made it possible to study the evolutionary trajectories of plant chemical compounds using genomic data (Chae et al., 2014; Schläpfer et al., 2017; Kang et al., 2020). This study aims to decipher the adaptive biochemical responses selected through the evolution of multiple plant lineages from the Atacama Desert. By extension, it is interesting to depict whether this evolution process led to the convergent fixation of various biochemical reactions and pathways or conversely resulted in distinct strategies. To meet these objectives, we first extracted biochemical reaction-related genes and annotated associated reactions and pathways from 32 Atacama species covering fourteen plant families. For comparison, we

performed a similar workflow in 32 plant species that were phylogenetically related to the Atacama species but lived in other milder environments (Eshel et al., 2021). We analysed gene family expansion and expression patterns and evaluated the enrichment of chemical reactions and pathways when comparing each pair of Atacama and related species using over-representation analysis (ORA) (Wieder et al., 2021). This computational strategy highlighted the convergent selective advantages of Atacama plant metabolomes. The most ubiquitous genetically enriched responses were related to protective mechanisms against major abiotic stresses observed in the Atacama Desert, such as drought, nitrogen deprivation and high light intensity. These findings provide fascinating new insights into adaptive mechanisms for plant survival in the Atacama Desert and new genetic targets for crop engineering for more resilient agriculture.

Materials and methods

Plant material. Reaction and pathway enrichment analyses were performed using previously described and available transcriptomics data from Atacama plant species (Talabre-Lejía transect, lat 22°-24°S) (Eshel et al., 2021). This natural environment spans an altitudinal cline from 2400 to 4500m and involves three vegetation belts defined based on the measured variations in water and nutrient availability, temperature and pH gradient (Carrasco-Puga et al., 2021). A set of 32 Atacama species were collected, directly snap-frozen and stored until transcriptomics analysis. These species covered 14 distinct plant families and flourished in the different vegetation belts. Additionally, transcriptomics data from the taxonomically-closest species available for each one of the 32 Atacama plants were extracted for phylogenomics analysis (Eshel et al., 2021). Here, those species are referred to as Sister species.

Generation of annotated reactions and pathways using PathwayTools. Reactions and pathways for each of the 64 species transcriptome were built using the annotation results generated by the e2p2v4 enzymes annotation tool (Schläpfer et al., 2017). Next, PTOOLS v24.5 was used to infer reactions and pathways using default parameter values (Karp et al., 2021).

Data treatment. To get insights into gene expansion patterns in Atacama plant species, we characterised the total number of genes per reaction for each annotated reaction. This process provided a first matrix which served as a basis for the reaction enrichment analysis through gene family expansion (Fig. IV.1, Table IV.S1). In parallel, the top 10% expressed genes from each of the 32 Atacama plant species, as well as the top expressed genes from 17 related Sister species (from which raw sequencing data were available), were extracted from the previous transcriptomics analysis (Eshel et al., 2021). Top expressed genes that were not associated with any reaction were removed. We then define the number of top-expressed genes per reaction. Hence, this number can range from 0 to "*n*" (no or "*n*" top-expressed

gene linked to the reaction *i*). This process produced the second matrix, which was used to decipher the enrichment of Atacama reactions based on gene expression patterns (Fig. IV.1, Table IV.S2).

Reaction enrichment analysis. Reaction enrichment analysis was performed on both gene family expansion and gene expression tables. We first focus our analysis on researching gene family expansion within Atacama plant species. Exclusive reactions from each ecosystem (*i.e.* either exclusive to the Atacama Desert or other lands) were extracted. Then, common reactions (detected in at least one species of each ecosystem) were scrutinised to determine their enrichment properties. Multiple comparison analyses were performed to compare the average number of genes per reaction in (i) all Atacama versus all Sister species, (ii) all species from specific vegetation belts (*i.e.* Prepuna or Steppe) versus related sisters. A high threshold was used to characterise reactions as enriched to avoid false positives. Thus, reactions were considered enriched if at least 3 times more genes per reaction were observed. Since this approach highlighted substantial variations in enriched reaction patterns between Prepuna and Steppe, pairwise comparisons between the 32 couples of Atacama-Sister species were performed to avoid the dilution effect. Results were then aggregated to calculate a percentage of occurrence per ecosystem (Atacama or Sister) or per vegetation belt (Prepuna or Steppe) (Table IV.S4).

Besides, since final transcript levels highly depend on genomics and transcriptomics interactions, a second analytical step focuses on gene expression levels. Reactions were considered enriched if at least 3 times more associated genes were observed in the top 10% of one Atacama species than its related sister. Similarly, average (*e.g.* average in all Atacama species versus average in all Sisters) and pairwise comparisons were performed for each available Atacama-Sister couple (Table IV.S5).

Pathway enrichment analysis. Pathway enrichment analysis was deployed to determine if entire pathways were enriched in Atacama plant species (Table IV.S6). The probability *P* of finding at least k reactions in pathway *i* was calculated using Fisher's exact test based on the hypergeometric distribution via phyper function available on R (version 4.0.4) as previously described (R Core Team, 2021; Wieder et al., 2021) (Fig. IV.1). Pathways were considered as potentially enriched when $P<0.05$. Then, Benjamini-Hochberg correction was applied via the p.adjust function available on R (version 4.0.4) to define significantly enriched pathways when corrected $P<0.05$ (Benjamini and Hochberg, 1995) (Tables IV.S7 and IV.S8). Finally, reactions and pathway enrichment analyses were performed using this same process to compare Steppe and Prepuna evolution (Tables IV.S9 and IV.S10).

Annotation of the enriched reactions and pathways. Reactions enriched in at least 50% of the plant species per vegetation belt (*i.e.* at least 50% of the species from Steppe, Prepuna or all Atacama plants) were annotated using the MetaCyc database (Caspi et al., 2020). Besides, metabolism, sub-pathways and biochemical pathways were determined using KEGG and HMDB databases (Kanehisa et al., 2014; Wishart et al., 2018).

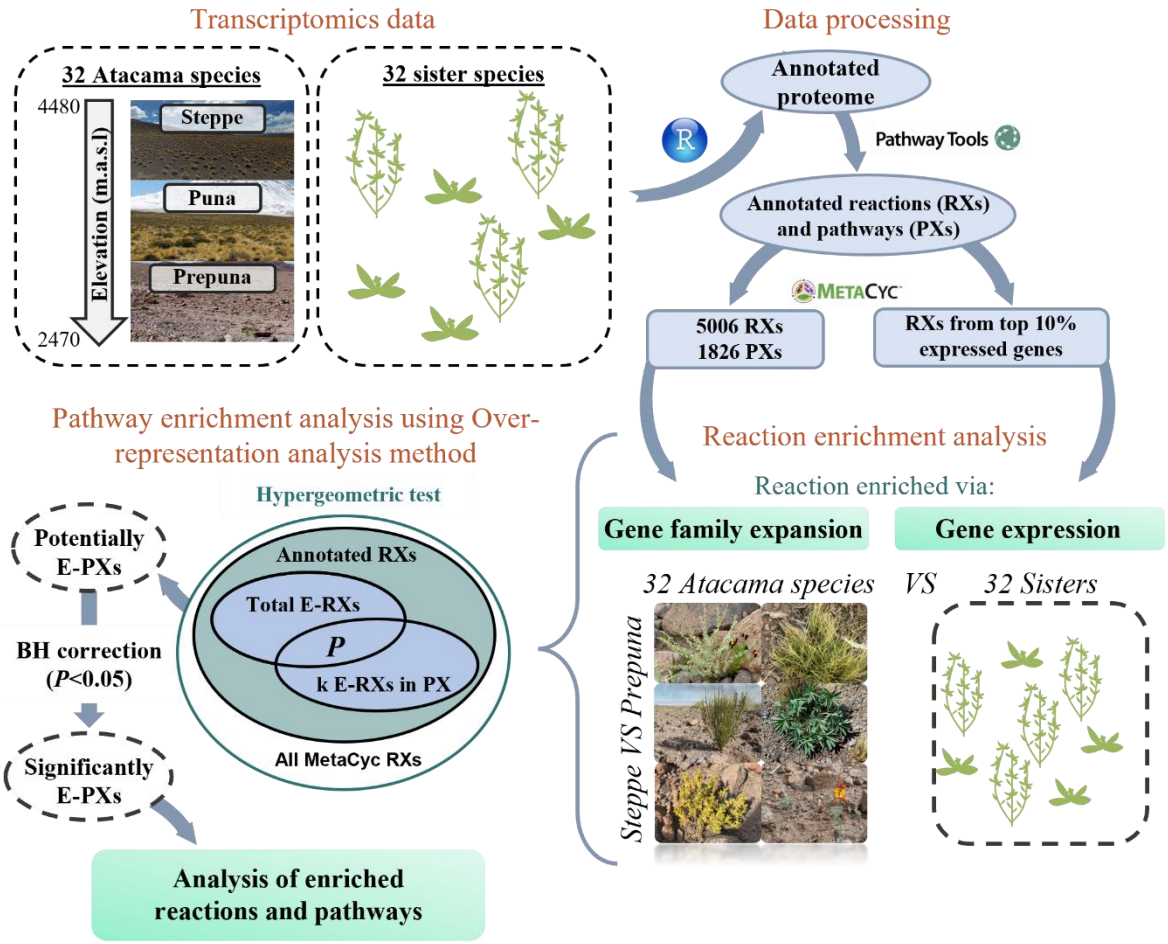

Fig. IV.1 | A simplified scheme of the reaction and pathway enrichment approach used in this study. *P* represents the probability of finding at least k reactions in pathway X (k E-RXs in PX) and was calculated using Fisher's exact test based on the hypergeometric distribution (Wieder et al., 2021). *E-RXs*: enriched reactions, *E-PXs*: enriched pathways.

Results

Atacama-exclusive reactions undergo species and environment specificity. Various plant lineages colonised and adapted their metabolism to perform in extreme ecosystems over millions of years of evolution (Guerrero et al., 2013). Studying the enrichment of chemical reactions and pathways on multiple species in harsh biomes could unravel universal or evolutionary convergent metabolic strategies relevant for adaptation. Remarkably, a phylogenomic analysis compared 32 Atacama species and 32 related species to unveil common and specific molecular mechanisms underlying plant survival (Eshel et al., 2021). These species encompass 14 plant families and flourished upon an elevation gradient from 2400 m to 4500 m in the Atacama Desert. This elevation gradient includes three vegetation belts or areas with distinct plant types: the Prepuna (low elevation, high salinity and low water availability), the Puna shrubland and the Steppe (high elevation, low temperatures) (Eshel et al., 2021 and Fig. IV.1). Here, we conducted reaction and pathway enrichment analyses that compared gene expansion and gene expression levels across the chemical reactions from 32 Atacama plant species and their related Sister species.

A total of five thousand and six annotated reactions (Fig. IV.1) were extracted from transcriptomics data available in the 64 species using PathwayTools (Eshel et al., 2021; Karp et al., 2021). The enrichment of chemical reactions may result from gene family expansion or a slightly different regulation that leads to a different expression of the plant genome (Kang et al., 2020; Scossa and Fernie, 2020). The expansion of enzyme-related genes through gene duplication, for instance, represents one of the main motors for eukaryote evolution (Lespinet et al., 2002). Through this phenomenon, gene copies will acquire a slightly different function as a shift in substrate preference (Fondi et al., 2009). Hence, reaction environment analysis was first performed by comparing gene family expansion and contraction across Atacama and Sister species (Fig. IV.1). To avoid false positives, reactions were considered enriched if the number of genes per reaction was at least 3 times higher in a given condition. In parallel, to access genome expression variations, we extracted and compared the top 10% expressed genes from Atacama and Sister species (Eshel et al., 2021 and Fig. IV.1). Then, we compiled the amount of top-expressed genes for each reaction and defined the enrichment property using the same threshold (*i.e.* at least three times more top-expressed genes per reaction). Interestingly, this comparative genomic evolution analysis highlighted an intriguing set of 463 exclusive reactions from Atacama species (Fig. IV.2). Although probably underestimated, these exclusive reactions were characterised by a high species-specificity level.

Environmental constraints have greatly impacted Atacama plant evolution, leading to different metabolic strategies. For instance, where exclusive reactions were equally divided between primary and secondary metabolism in Prepuna, 76% represented secondary pathways in Steppe species (Fig. IV.2). Fatty acyls, carbohydrates, amino acids and monoterpenes pathways have shown relevant activity to

cope with the extreme osmotic stressful conditions in Prepuna. Conversely, carotenoids and flavonoids were the main impacted pathways in Steppe, an ecosystem characterised by its low temperatures and high radiation levels (Fig. IV.2, Table IV.S3). Remarkably, multiple exclusive reactions were potentially related to regulating specific metabolites that have been characterised as relevant for adaptation to extreme Atacama conditions (*e.g.* quercetin, proline) (Dussarrat et al., 2022, Table IV.S3).

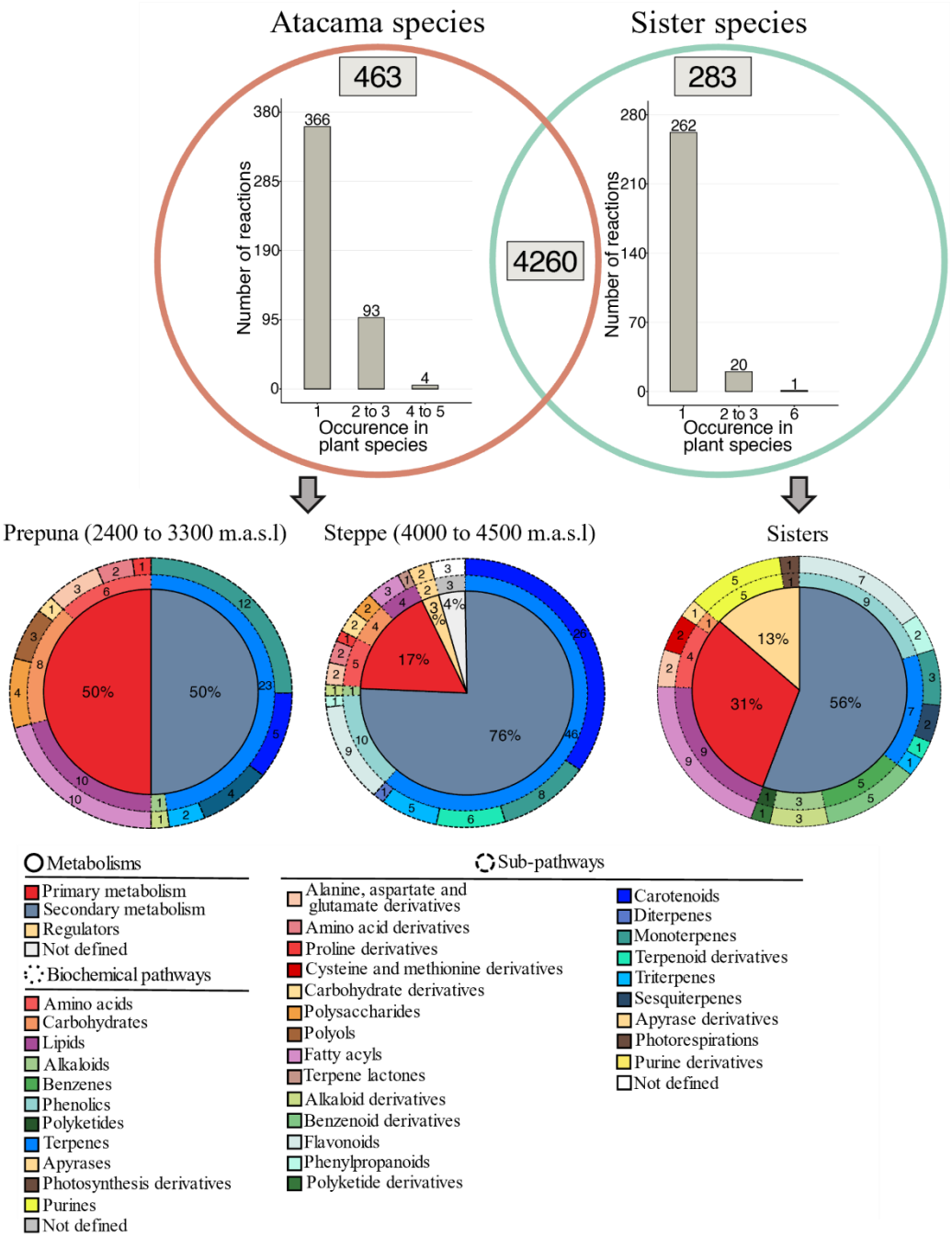

Fig. IV.2 | Analysis of reactions to Atacama and Sister species. The Venn diagram describes the ecosystem-specific reactions from the 32 Atacama plant species or their closely related Sister species. Since Puna (3300-4000 m.a.s.l) has shown intermediate results, this figure illustrates the exclusive reactions from Prepuna (2400-3300 m.a.s.l) and Steppe (4000-4500 m.a.s.l) species. Metabolism, biochemical pathways and sub-pathways were defined based on MetaCyc and KEGG data. *m.a.s.l*: meters above sea level.

Enrichment analyses unveil convergent chemical strategies shaped by environmental constraints. Next, a global comparison of the gene expansion and gene expression levels (*i.e.* genes per reaction and top-expressed genes per reaction respectively) between all Atacama plants *versus* all related species was performed to evaluate the enrichment among the 4260 shared reactions. Overall, although multiple convergences, reactions enriched via gene expansion and/or gene expression showed great environment specificity (Fig. IV.3A and IV.3B). Hence, pairwise comparisons were performed to avoid a potential dilution effect. Whilst gene expansion patterns were analysed on the 32 couples, 17 couples were used to study gene expression levels since raw sequencing data were only available for 17 species. In total, 2507 reactions presented three times more genes per reaction, and 1549 reactions involved three times more top-expressed genes per reaction in at least one Atacama species (Tables IV.S4 and IV.S5). To test the hypothesis of whether the Atacama plant evolution led to the convergent enrichment of biochemical reactions among the different plant lineages, we extracted the reactions enriched in at least 50% of the species in a given ecosystem (Fig. IV.3C and IV.S2). Hence, the gene expansion analysis results illustrate the reactions enriched in at least 16 species when considering all Atacama plants or at least seven species when considering individual vegetation belts (*i.e.* Prepuna or Steppe) (Fig. IV.3C).

Similarly, results from gene expression analysis encompassed reactions enriched in at least nine or four species since raw sequencing data were available for 17 Sister species. Excitingly, results pinpoint a global reorchestration of both primary and secondary metabolism. First, Atacama plants likely limit carbon entry by regulating chlorophyll levels, a protective process observed in other extreme plants (Cui et al., 2019) (Fig. IV.3C and IV.S2). Chlorophyll b reductase was overexpressed in more than half of the Atacama plant species, an enzyme whose protective role against high light intensity was described as essential (Sato et al., 2015). Besides, results highlighted a synthesis of a battery of protective primary and secondary compounds while a great activity of hormones related to plant development and growth was observed. Gibberellins (*e.g.* gibberellin oxidase), cytokinins (*e.g.* cytokinin oxidase but also cytokinin-activating enzymes), and jasmonates (*e.g.* acyl-coenzyme A oxidase 1) were among the impacted chemical compounds in at least half of the plant species, probably due to their role in plant development or in boosting plant defences against biotic and abiotic stress. Reactions related to carbohydrate pathways (*e.g.* starch and polyols), lipids (*e.g.* waxes synthesis) and amino acids (*e.g.* proline) were enriched in the majority of Atacama plant species (Fig. IV.3C and IV.S2). Importantly, most of the highlighted reactions by our computational analysis were previously characterised for their function in mitigating various abiotic stresses. For example, the glucan/water dikinase and the disproportionating enzyme was associated with starch degradation and freezing tolerance (Yano et al., 2005), while an intriguing role in coordinating plant growth and tolerance to abiotic stress was highlighted for inositol phosphatases (Lou et al., 2007; Jia et al., 2019), which were greatly enriched in Atacama species (Tables IV.S4 and IV.S5). Furthermore, fatty aldehyde decarbonylase was linked to cuticular waxes synthesis and protection against drought (Zhou et al., 2013), while proline

accumulations due to delta 1-pyrroline-5-carboxylase synthase overexpression in response to various osmotic perturbations were extensively detailed (Strizhov et al., 1997).

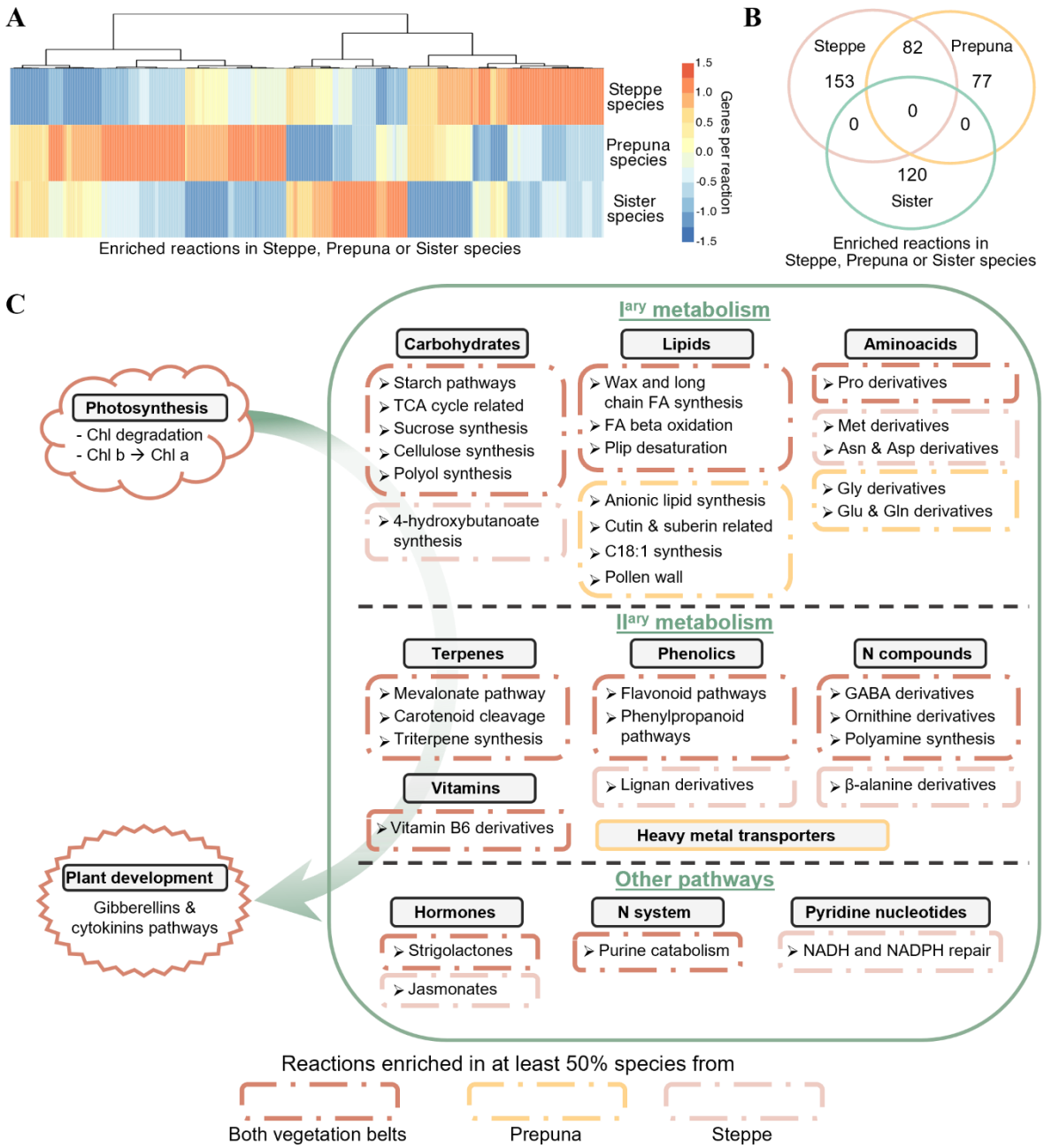

Fig. IV.3 | Identification and classification of enriched reactions from Atacama plant species into major chemical pathways. A. Global analysis of enriched reactions between vegetation belts based on gene family expansion analysis. **B.** Distribution of enriched reactions based on comparing top 10% expressed genes. Analyses were performed by comparing mean values of genes per reaction (A) or top 10% expressed genes per reaction (B) between Atacama VS Sisters or Prepuna or Steppe VS related sisters. C. Results of a pairwise comparison analysis and classification of the reactions enriched in at least 50% of the species. Each Atacama-Sister species pair was analysed individually to avoid a potential dilution effect. Results were then aggregated to calculate a percentage of occurrence per ecosystem (Atacama or Sister) or per vegetation belt (Prepuna or Steppe). *Asn*: asparagine, *Asp*: aspartate, *Chl*: chlorophyll, *FA*: fatty acids, *Gln*: glutamine, *Glu*: glutamate, *Gly*: glycine, *Met*: methionine, *Plip*: phospholipids, *Pro*: proline.

Secondary pathways were also highly represented and referred to carotenoid, triterpene, phenylpropanoid and flavonoid pathways (Fig. IV.3C and IV.S2). Enzymes involved in carotenoid synthesis and cleavage such as carotene hydroxylases and carotenoid dioxygenases, which were also linked to strigolactones synthesis, were observed among reactions enriched in at least 50% of the plant species. Concomitantly, flavonoid glucosyltransferase enzymes were extensively enriched in Atacama species, supporting the role of flavonoids in plant resilience to abiotic constraints (Di Ferdinando et al., 2012). Notably, the Atacama plant's evolution led to a natural enrichment of reactions related to uptake and process of nitrogen resources (*e.g.* reactions linked to glutamate, glutamine, GABA, polyamine) as well as nitrogen remobilisation through purine degradation. Strikingly, most of the underlying enzymes have shown interesting protective properties under various stressful conditions. As an example, the activity of glutamine, glutamate, GABA pathways, here represented by the presence of several enzymes such as glutaminases, γ-aminobutyrate aminotransferase and transaminases (Table IV.S5), was revealed in other extreme plants and undoubtedly linked to plant defence mechanisms (Kinnersley and Turano, 2000; Solomon and Oliver, 2002; Martinelli et al., 2007). Moreover, the enrichment of agmatine deiminase and spermidine synthase in Atacama plants was congruent with their pivotal role in plant defences (Kasukabe et al., 2004). Finally, evolution greatly impacted the cell wall content of Prepuna plants through lipidic profile modulation. Conversely, Steppe species emphasised the production of protective compounds as lignans (Fig. IV.3C and IV.S2, Tables IV.S4 and IV.S5). Besides, Steppe species favoured the tolerance to oxidative stress through 4-hydroxybutanoate synthesis and pyridine nucleotide repair via epimerases, which were recently considered as a relevant player in NAD(P)H metabolism (Breitkreuz et al., 2003; Gakière et al., 2018). Overall, reaction enrichment analysis unveiled relevant management of the resources (*i.e.* carbon and nitrogen) towards a more adapted balance between plant development and defence compared to related non-adapted species. Also, several reactions were specifically enriched in Steppe or Prepuna species to satisfy environmental demands. In contrast, Atacama plants have likely contracted or negatively regulated gene families involved in energy processes and hormone and terpene pathways (Fig. IV.S1, Table IV.S5).

Next, we performed Fisher's exact test based on the hypergeometric distribution to test whether entire biochemical pathways were enriched through evolution (Wieder et al., 2021 and Fig. IV.1). Interestingly, protective compounds were again over-represented and greatly influenced by environmental constraints. For instance, the production of protective compounds in Prepuna was mainly related to primary metabolism (*e.g.* polyunsaturated fatty acid, inositol synthesis). Conversely, the synthesis of phenolics, quaternary ammonium and cyanogenic glycoside compounds were prominent in Steppe. Additionally, multiple pathways were ubiquitously highlighted and referred to proline, ornithine, carotenoid and polyamine compounds (Fig. IV.4). Finally, to get insights into the variations in adaptive chemical strategies employed between vegetation belts, we conducted a reaction and pathway enrichment analysis comparing the 13 Steppe and 13 Prepuna species (Fig. IV.5, Tables IV.S9

and IV.S10). Results confirmed (i) the existence of similarities and divergences among the chemical strategies adopted by plants to respond to major environmental constraints and (ii) a significant homogeneity of enriched reactions and pathways across the different Prepuna or Steppe species. Overall, these findings identified multiple evolutionary convergences illustrated by the high proportion of reactions enriched in more than 50% of the plant species (Fig. IV.3, IV.4 and IV.5). Even more thrilling, these enriched reactions and pathways strongly coincide with the potential regulation of the metabolic toolbox employed by plants to face extreme Atacama conditions (Dussarrat et al., 2022, Table IV.S12). Hence, this genome-scale comparative analysis provided a unique goldmine of genetic targets for engineering resilient crops against major abiotic threats.

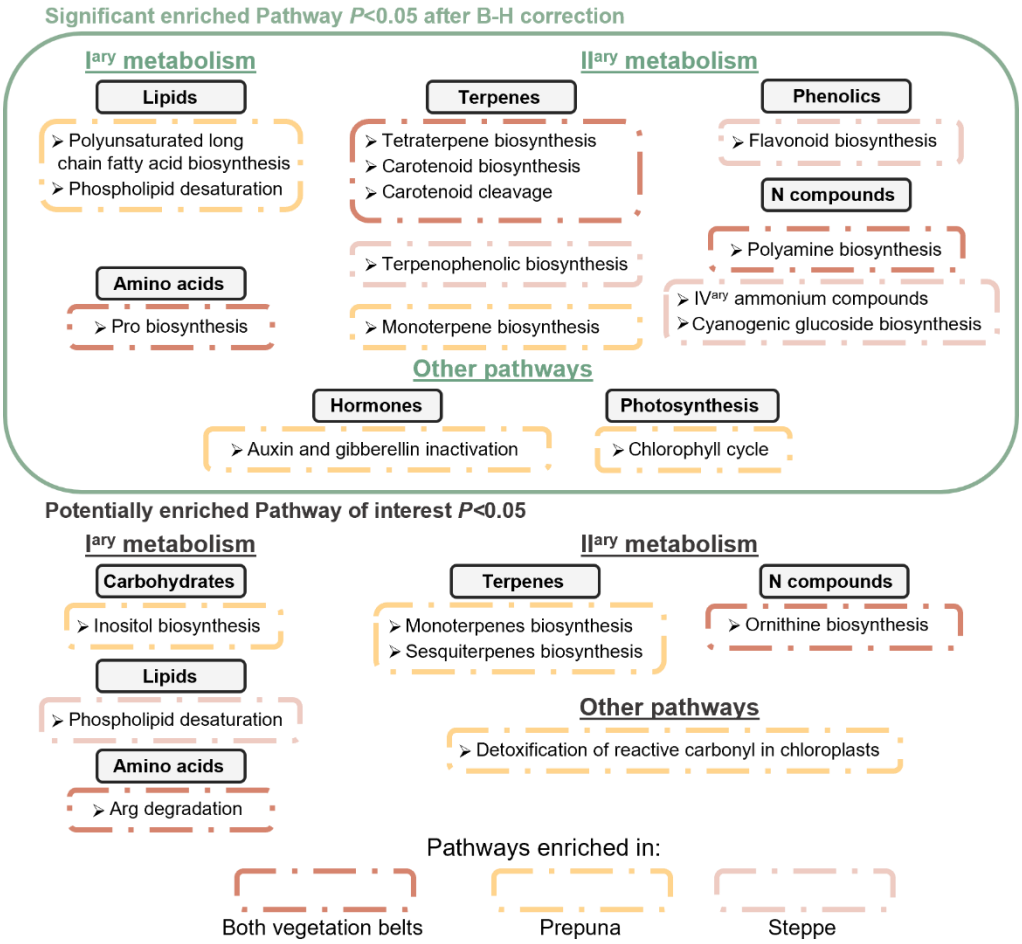

Fig. IV.4 | Characterisation of enriched pathways detected in Atacama plant species. All reactions (exclusive and common reactions) were used to perform pathway enrichment analysis. *B-H*: Benjamini-Hochberg, *Pro*: proline, *Arg*: arginine.

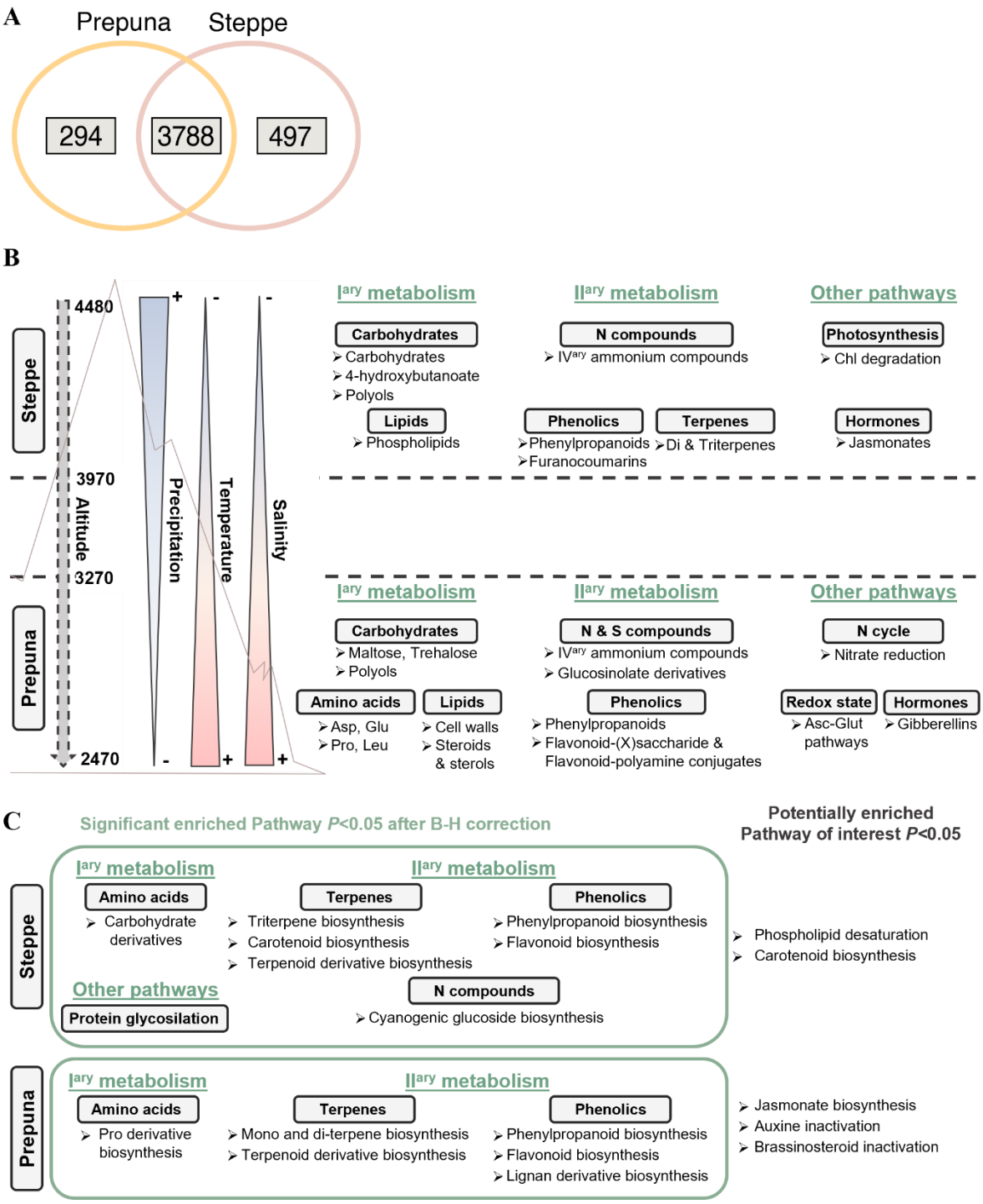

Fig. IV.5 | Reaction and pathway enrichment analysis comparing Steppe and Prepuna species. **A.** Venn diagram describing the distribution of the predicted annotated reactions in 32 Atacama plant species and their closely related sisters. Metabolism, biochemical pathways and sub-pathways of the ecosystem-specific reactions were defined based on MetaCyc and KEGG data. **B.** Depiction of the reactions enriched in at least 50% of the species from Steppe or Prepuna. **C.** Results of pathway enrichment analysis. *Asp*: aspartate, *Chl*: chlorophyll, *Glu*: glutamate, *Leu*: leucine, *Pro*: proline.

Discussion

Evolution led to generic metabolic strategies in extreme plants from diverse lineages. How plants adapt to their environment has been of great interest since the domestication of plants in the harsh environments of the Fertile Crescent 10,000 years ago (Riehl et al., 2012). Although great successes in plant breeding allowed feeding a growing population, the current acceleration of global warming strongly limits the efficiency of genetic improvement in the best ideotypes (Seneviratne et al., 2012; Bailey-Serres et al., 2019). We thus need innovative research strategies that break away from the current reductionist single-species approach to discover universal plant resilience mechanisms, which should be more easily transferable to crops. Presumably, the evolution process by natural selection should display some of the most efficient genetic traits for plant survival under harsh conditions. Remarkably, an eco-metabolomics approach unveiled the existence of a generic metabolic toolbox allowing plant resilience to harsh climates (Dussarrat et al., 2022). Importantly, our strategy interrogated the evolution of the genetic mechanisms underlying adaptive metabolic strategies of multiple Atacama plant species. This approach compared the gene expansion and expression levels at the genome-scale between 32 Atacama species and 32 related species, yielding the discovery of convergent and divergent biochemical evolutions relevant for plant survival.

Firstly, while the Atacama and Sister species allocate carbon and nitrogen resources differentially, it is noteworthy that these plants were not so fundamentally distinct. Very few exclusive reactions were observed in Atacama plants. Besides, the extreme majority (99%) of the 463 exclusive reactions were observed in only one, two or three species. Similarly, Atacama plants conserved most of the chemical reservoir from their ancestors since only 6% of the total annotated reactions were unique to Sisters (Fig. IV.2). Although probably underestimated since this approach can only access known reactions, these numbers prompt how common a random mutation event generates a new chemical pathway required for plant adaptation. While phylogenomics and metabolomics approaches permit the detection of these unknown traits, the majority of genes and metabolites highlighted as relevant for adaption of Atacama species referred to conserved processes (Eshel et al., 2021; Dussarrat et al., 2022). Considering the occurrence of vital genome innovations over the past 145 million years of angiosperm evolution, twelve million years (*i.e.* age of the Atacama Desert) represents a short time-scale for the development of new species-specific reactions allowing survival of 14 distinct plant lineages (Benton et al., 2021). Nevertheless, these exclusive reactions might have a role in improving plant fitness at the species level, which could be evaluated through an untargeted comparison of adapted and non-adapted species from the same genus (*e.g.* Atriplex).

Conversely, an exciting number of chemical convergences was shown as a result of the plant evolution process in the Atacama Desert among conserved reactions between Atacama and Sister species (Fig. IV.3 and IV.S2). Overall, these shared mechanisms demonstrated a high potential to modulate resource uptake and allocation between development and defence. More thrilling, the modulation of major metabolic processes was targeted by enriched reactions in at least 50% of the species studied, encompassing 14 different plant lineages. Besides, genes underlying these enriched chemical routes clearly point to the regulation of metabolites employed by Atacama plants to face extreme environmental conditions (Dussarrat et al., 2022). For instance, starch, proline, trehalose, jasmonic acids, 5'-methylthioadenosine, quercetin and quercetin glucoside, tricoumaroyl spermidine compounds as well as carotenoid cleavage and chlorophyll cycle (Table IV.S11). With carotenoid and chlorophyll-related pathways, quercetin was the most frequently observed compound in exclusive and enriched reactions. Interestingly, this flavonoid was linked to various roles as an antioxidant (response to high irradiance), a mediator of interaction with nitrogen-fixing bacteria (response to critical nitrogen levels), a protective compound against heavy metals, and its links with both hormones (*e.g.* abscisic acid and auxins) and redox buffers (*e.g.* glutathione) (Singh et al., 2021).

Hence, while adaptation was thought of as mainly species-dependent (Turner, 2018; Dussarrat et al., 2021), these findings strongly suggest that a relevant part of this is conversely the result of convergent evolution of regulatory processes. In other words, the development of adaptive traits providing a selective advantage is more likely to occur from the regulation of pre-existing compounds and pathways than from the emergence of new ones. Furthermore, most of these convergences represented conserved compounds among wild and crop species. Finally, some metabolic traits were highlighted using two independent analytical approaches and therefore hold great promise for potential crop engineering.

Strong environmental pressures define the evolutionary trajectories. Solar irradiance, water and nitrogen availability are three critical parameters for plant survival in the Atacama Desert (Eshel et al., 2021). Besides, other limiting parameters arise along the elevation gradient from 2400 to 4500m. Whilst Steppe undergoes negative daily temperatures, Prepuna species are subjected to extreme salinity (Díaz et al., 2019). Overall, the Atacama Desert represents a unique opportunity for studying plant response to current agronomical challenges (Pachauri et al., 2015). Our study unveiled several adaptive mechanisms that have been fixed through evolution and answer a great proportion of these challenging environmental constraints.

Carbon is not a limiting resource in desertic regions due to intense solar irradiance (Xu et al., 2016). In contrast, other edaphic and climatic parameters forced a global metabolic reorganisation. First, Atacama plants tend to limit the level of chlorophyll and increase the chlorophyll a/chlorophyll b ratio to regulate energy capture and mitigate oxidative stress, in agreement with other extreme plants (Cui et

al., 2019). Then, the allocation of carbon reserves between plant growth and defence was finely tuned through evolution with, for example, the modulation of gibberellin and cytokinin pathways (Fig. IV.3C). Besides, strong remobilisation of carbon reserves via starch degradation and the extensive synthesis of protective compounds suggested that Atacama plants evolved out of their physiological range (Tables IV.S4 and IV.S5). Hence, a series of enriched reactions in at least 50% of the plant species referred to as protective compounds from both primary and secondary metabolism could represent adaptive traits to face osmotic pressure as well as frost, high solar irradiance and nitrogen starvation (Fig IV.3C and IV.S2). Synthesis of oligosaccharides (*e.g.* sucrose) and other chemical constituents like polyols could be linked to the meagre water availability (Williamson et al., 2002; Pamuru et al., 2021). Lipids were also universally represented based on their role in plant defence against osmotic and cold stress. The synthesis of waxes and long fatty acyl chains can refer to how plants limit evapotranspiration and minimise the solar irradiance impact (Kolattukudy, 1970). Besides, the extensive activity of lipid metabolism illustrated by a shift in the degree of saturation could support the role of membrane fluidity in adaptation (Li et al., 2020a). Finally, these changes accompanied regulation of the TCA cycle and proteinogenic amino acid pathways, two major precursors of secondary compounds (Yang et al., 2020). Thus, a significant part of carbon reserves was used to produce secondary protective compounds. The presence of the mevalonate pathway and carotenoids synthesis and cleavage processes among the reactions enriched in more than 50% of the Atacama species is consistent with their role in stress mitigation and their links with hormonal and redox pathways (Havaux, 2014). Besides, nitrogen related compounds (*e.g.* GABA, polyamines and quaternary ammonium compounds) and phenolics are other metabolic features employed by plants to face a myriad of climate constraints (Dussarrat et al., 2021). Interestingly, a significant proportion of the enriched reactions is directly or indirectly linked to redox homeostasis through reactive oxygen species scavenging (*e.g.* polyphenols) or proline and carotenoid cleavage respectively, supporting its central place in adaptation.

Fascinatingly, the Atacama plant genome greatly integrated the specific requirement across the elevation gradient (Fig. IV.3C, IV.S2 and IV.4). This was exemplified by the over-representation of polyol and sterol synthesis in Prepuna, where drought and salinity are extremely severe (Rogowska and Szakiel, 2020; Eshel et al., 2021). Nitrogen starvation is of major importance in this vegetation belt, illustrated by the extensive genetic expansion and expression of glutamine, glutamate and nitrate reduction-related pathways. In addition, evolution led to an increased synthesis of some flavonoid-polyamine conjugates (*e.g.* tricoumaroyl spermidine) whose roles remain poorly described (Dussarrat et al., 2021). In contrast, Steppe species favoured the production of a plethora of secondary compounds (Fig. IV.4). Carotenoids and other isoprenoids could be used for their advantages to mitigate abiotic constraints (Havaux, 2014). Interestingly, plants from higher levels tend to produce a wide range of metabolites involved in biotic defence such as coumarins and cyanogenic glucosides (Gleadow and Woodrow, 2002; Stringlis et al., 2019).

Such a comparative approach allows investigating the plant response to major abiotic constraints while preserving the ecological and evolutionary context (Kang et al., 2020). Interrogating Atacama plant evolution highlighted that extreme plants uniqueness lay in the regulation of pre-existing pathways and unveiled a high degree of convergences between chemical strategies selected to face harsh climate conditions. Strikingly, these generic strategies included reactions and pathways relevant for plant resilience against osmotic (*e.g.* drought and salinity), frost and high irradiance stress as well as low soil suitability (*e.g.* salinity and nitrogen starvation). Hence, these findings pave the way for wider use of these generic metabolic mechanisms to provide sustainable solutions to improve global food security. A thrilling perspective will be to investigate whether these generic chemical reactions are also involved in adaptation to other extreme lands.

Acknowledgements.

We acknowledge the Comunidad de Talabre for providing us access to the Talabre-Lejía transect. We also thank Gil Eshel and Fabien Jourdan for their help and analytical advice. We thank Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT 1180759), DOI EVONET Proyect and ANID – Millennium Science Initiative Program – iBio ICN17_022 as well as Fondo de Desarrollo de Áreas Prioritarias (FONDAP) Center for Genome Regulation (15200002) for financial support to R. A. Gutiérrez. We thank Fondecyt Iniciación 11150107 for financial support to R. Nilo-Poyanco.

Author contributions.

RG and CL conceived the Atacama project. RG, CL, YG, DR, PP, TD and RNP participated in the conceptualisation of this enrichment analysis. TD and RNP produced the initial annotated reactions and pathways table. TD, RNP, VA, GD, TM, SP and RG analysed the data. TD, RNP, PP and RG wrote the manuscript with feedback from all co-authors.

All references are available at the end of the thesis manuscript.

Supplemental figures. Supplemental tables are available at the following link until publication: https://drive.google.com/drive/folders/1Z3HLMY0Hb281HEu56MM82MHtY_YQ9tkE?usp=sharing

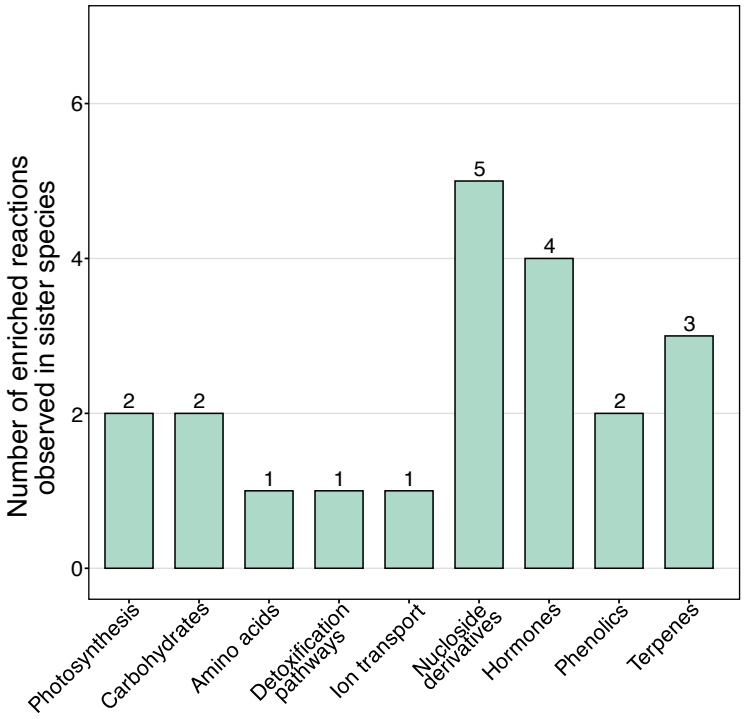

Fig. IV.S1 | Characterisation of enriched reactions from Sister species. Biochemical pathways were defined based on MetaCyc and KEGG data.

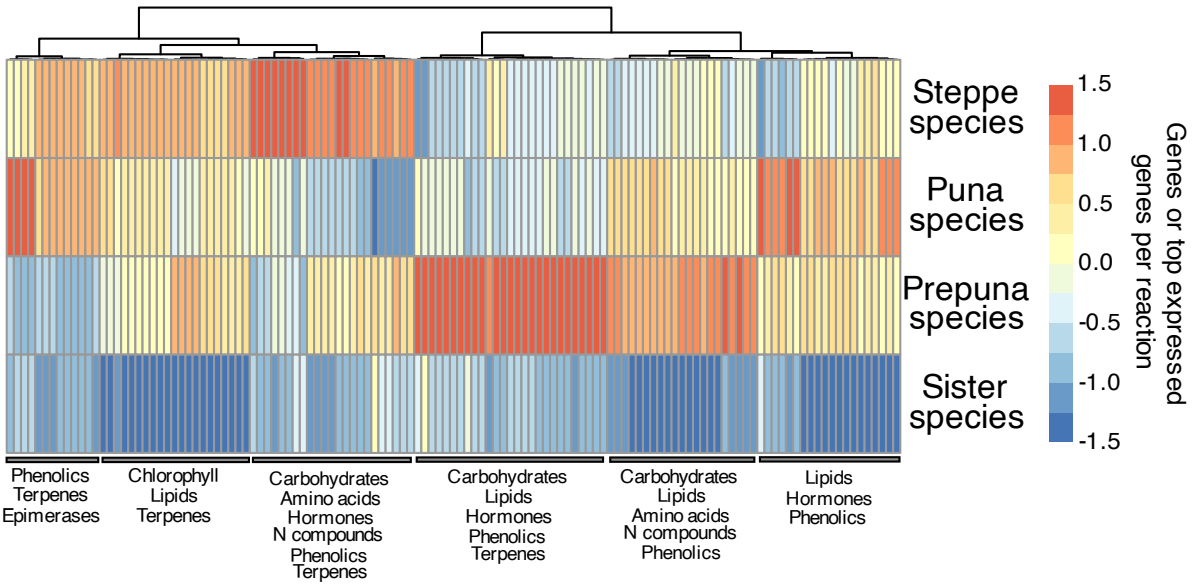

Fig. IV.S2 | Enriched reaction in Atacama species. Depiction of the main biochemical classes enriched in Atacama plant species. Only annotated reactions enriched in more than 50% of the species in a given ecosystem (*i.e.* Atacama, Steppe, Prepuna) are represented. *N compounds*: nitrogen-related compounds.

III. EXTREME CLIMATES CONDITION EVOLUTION TRAJECTORIES

The comparison of the gene expansion and expression patterns using PathwayTools (Kang et al., 2020) pinpointed an interesting set of reactions and pathways enriched in a broad range of Atacama plant species. In addition, the vast majority of reactions enriched in at least half of the plant species were detected in both extreme and non-adapted plant species, while Atacama exclusive reactions were mostly species-specific. Hence, based on the hypothesis that fixed genetic mutations provide a selective advantage, these findings confirm the existence and the significance of the convergent biochemical strategies in the adaptation of multiple Atacama plant species (Fig. IV.6). This was supported by a strong correlation between ubiquitously enriched reactions and major environmental constraints that challenge plant performance in Prepuna or Steppe. However, the potential role of the species-specific traits in adaptation can not be excluded and should be further studied. In this sense, a generic toolbox predicted the elevation level of 24 Atacama plant species with 79% accuracy, while models performed at the species level allowed a prediction with 90% accuracy (Dussarrat et al., 2022). These observations suggest that although generic strategies are likely to sit at the forefront of the adaptation process, the role of species-specific strategies should not be neglected. Notably, a plethora of causes could explain this 11% delta and the occurrence of divergent strategies (exclusive enriched reactions of species-specific markers). These markers could, for example, hide the existence of micro-environments within the Atacama Desert, a phenomenon that commonly occurs in harsh biomes (Flores and Jurado, 2003; Cavieres et al., 2006). For instance, epiphytes employ ingenious strategies to thrive in challenging environments such as orchids which developed specialized root tissues in tree canopies (Zotz and Winkler, 2013). Nurse plants favoured commensalism, where seeds benefit from a protected habitat providing a warm or moist environment (Flores and Jurado, 2003). Such conditions might occur in the Atacama Desert, and their analysis would broaden our knowledge of the adaptive strategies developed across the evolutionary history of this unique ecosystem.

Expanded and overexpressed gene families are congruent with the 39 metabolites capable of predicting the environment of multiple Atacama plant species (Dussarrat et al., 2022). The negative correlation of proline, quercetin glucoside and polyamine derivatives with the elevation level as well as the positive correlation of complex sugars like raffinose and carotenoids were validated by the genetic enrichment observed in Prepuna and Steppe species, respectively. Thus, findings from this computational approach first confirmed the existence of convergent biochemical evolution in different plant lineages, but also provided evidence for the significance of most metabolites included in the predictive toolbox at the evolutionary scale (Fig. IV.6). Interestingly, primary metabolism was much more represented as compared to the predictive metabolomics approach, pinpointing the need for complementary analyses towards polar, central compounds. Notably, amino acids such as glutamate and glutamine were observed among enriched reactions in Prepuna species. These traits could refer to an

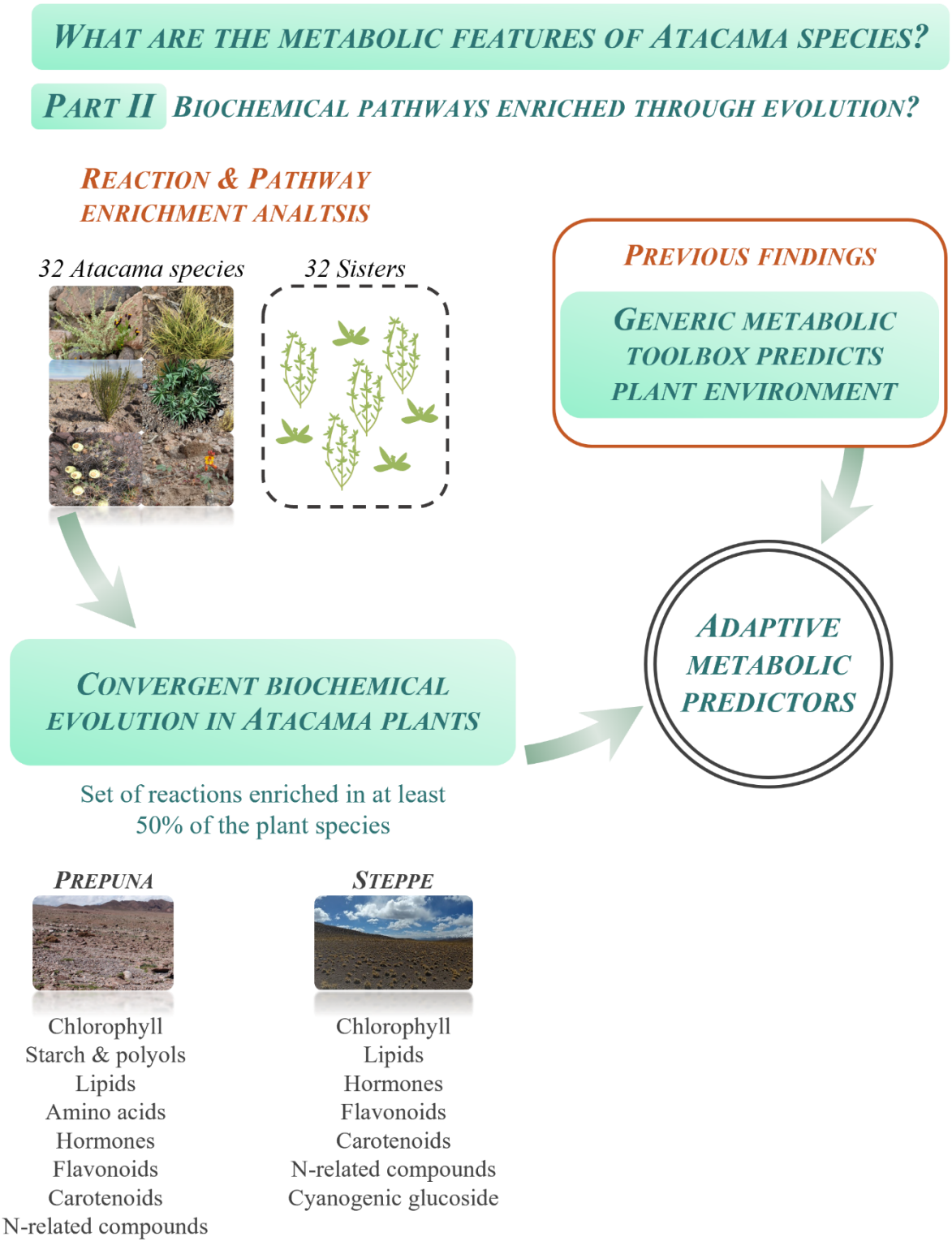

Fig. IV.6 | The secrets of plant adaptation to the Atacama Desert.

Chapter 4: A convergent biochemical evolution guided by environmental constraints.

adaptive response against nitrogen starvation, a critical limitation for plant life in the Atacama Desert (Díaz et al., 2016). While a high activity of amino acid metabolism was detected via robotised microplate assays (Dussarrat et al., 2022), complementary techniques such as NMR or hydrophilic chromatographic assays would permit a deeper investigation of these essential compounds (Kim et al., 2010). Finally, bottom-up approaches would help characterise the role of the metabolic markers validated at both metabolic and evolutionary levels. While only three species can be grown in the greenhouse (the Fabaceae *H. doellii*, the Solanaceae *S. chilense* and the Poaceae *A. adscensionis*), we could imagine an experiment to explore the response of these predictors under a temperature, nitrogen or drought gradient. Subsequently, the integrative modelling approach could be complemented by various molecular biology techniques to evaluate the impact of these compounds on the resilience of extreme plant species and relative crops and improve our comprehension of the underlying molecular mechanisms (Moore et al., 2009; Zhu et al., 2020).

CHAPTER 5

BIOTIC MICRO-ENVIRONMENTS SHAPE PLANT COMMUNITIES IN EXTREME LANDS

I. FROM INDIVIDUALISTIC TO ECOSYSTEM APPROACH

The publication of this chapter is expected for April 2022. Here, we proposed a short introduction (Section I) to set the topic in the context of the PhD. This analysis is being finalised, results are reported in a paper format and a short introduction is added to provide a general background of the facilitation process in anticipation of the valuation of this work (Section II). Finally, discussion and conclusion paragraphs were included in Section III.

The Atacama Desert is one of the harshest environments for plant life (Eshel et al., 2021). Such hostile conditions required a global reorganisation of carbon and nitrogen allocation to ensure a sufficient reserve for development and reproduction while delivering a precise biochemical defensive strategy (Dussarrat et al., 2021). A few species succeed in this fascinating challenge through adaptation of their metabolome, which encompassed remarkable metabolic strategies to cope with extreme drought, salinity, solar irradiation and daily sub-zero temperatures (Díaz et al., 2016). This biochemical adaptive response was surprisingly mainly mediated by convergent evolutions which enabled the regulation of pre-existing pathways. Thus, genetic legacies allowed subtle management of different metabolic reactions involving various compounds from primary and secondary metabolism. These metabolites predicted plant environment with 79% accuracy, independent of plant species and year (Dussarrat et al., 2022). Hence, these results strikingly supported the central place of convergent evolutions in adaptation to extreme lands. However, despite the so-called “multi-species” approach, these conclusions and analyses remained grounded in an individualistic approach. In other words, the comprehension of the metabolic strategies employed by plants to cope with abiotic stresses remained linked to the individual and not the population level. Interestingly, our perception of the plant system was affected by the development of the classification system and the analysis of their response to stress, which primarily led to the individualistic theory (Gleason, 1926; Huntley, 1991; Enquist and Leffler, 2001). This individualistic theory was initially supported by the fact that plant species variation and distribution across environmental gradients is continuous and rarely overlapping (Austin, 1985; Callaway et al., 2002). However, advances in ecological models proposed that the coverage of a given species was not simply linked to its capacity to regulate phenotypic and metabolic traits in response to environmental perturbation but rather conditioned by the ecosystem dynamics (Callaway, 1998; Bruno et al., 2003; Hu et al., 2021). For instance, microbial communities have an important role in plant adaptability, especially in arid lands (Soussi et al., 2016; Zhang et al., 2020). Besides, plant-plant interactions are a major driver of plant community structure (Lortie et al., 2004). While the potential role of communication between plants and soil microorganisms in adaptation to the Atacama Desert has been raised previously (Eshel et al., 2021), here we focus on the potential role of the facilitation process. We combined metabolomics and machine learning with ecology to investigate the role of plant-plant interactions in the adaptation of

Atacama plants. More precisely, great efforts were deployed to disentangle the thermophysical and metabolic mechanisms underlying the facilitation process from which multiple species benefited through the interaction with a cactus species called *Maihueniopsis camachoi*.

II. CACTUS “NURSE EFFECT”: FROM ECOSYSTEM TO METABOLISM

Introduction

Plant-plant interactions shape community structure by influencing plant diversity and abundance (Berlow, 1999; Lortie et al., 2004; Ploughe et al., 2019). Besides, this phenomenon is a crucial player in shaping ecosystem response to climate change (Brooker, 2006; Delgado-Sánchez et al., 2013; Sherwood and Fu, 2014; Åkesson et al., 2021). Previously neglected, positive relationships earned peculiar interest for their role in plant communities (Bruno et al., 2003; McIntire and Fajardo, 2014). Furthermore, nurse plant (*i.e.* a plant providing a positive effect to other species in close spatial association) properties have also shown promising results for the restoration of degraded environments (Padilla and Pugnaire, 2006; Zhao et al., 2007). While the effect of abiotic stress on facilitation processes is still debated, wildlands offer a unique opportunity to improve our understanding of facilitation processes, predict the impact of climate change on harsh biomes, and plan restoration practices (Brooker et al., 2007; Urza et al., 2019). Previous analyses highlighted an increased abundance of nurse species in arid and cold biomes compared to other environments (Flores and Jurado, 2003; Antonsson et al., 2009). The *Cactaceae* family includes many of those species, suggesting that their short stature and compact structure enable the trapping of ambient heat and moisture (Flores and Jurado, 2003; Yang et al., 2017). Nevertheless, analysing the integral picture of protective facets offered by cushions as thermal or hydric refuges is complex since these species generally experience only one of the above constraints (Körner, 2003). For instance, two meaningful studies performed on tens to hundred plant species spanning elevation gradients yielded opposite results, where positive associations decreased or increased with elevation (Callaway et al., 2002; Cavieres et al., 2006). Hence, an ecosystem-scale approach in an extreme environmental gradient combining temperature and hydric constraints is all the more crucial to provide a more inclusive understanding of facilitation processes on diverse plant lineages and lifespans.

Surprisingly, the potential benefits of nurse species on plant survival have been mainly reduced to the analysis of seedling establishment (Cavieres et al., 2006). Notably, two studies provided valuable information about the nurse effect on various ecophysiological and anatomical traits such as osmotic potential and chloroplast density, paving the way for further significant research (Delgado-Sánchez et al., 2013; van der Merwe et al., 2021). In contrast, the impact of these protected micro-environments on

plant metabolism remains unexplored. The metabolome has a tremendous ability to capture environmental variations and integrates past, present and future biochemical processes of plant life (Lewis and Kemp, 2021; Signori-Müller et al., 2021). Advances in analytical techniques and machine learning recently allowed the development of ecological metabolomics, which aimed to study the metabolic interactions between plants and their natural environment (Sardans et al., 2020). In parallel, one of the main goals in ecology is to extend individual traits to an ecosystem scale to provide a comprehensive view of plant communities (Lortie et al., 2004). We believe metabolomics could be combined with ecology to help address this objective by (i) providing readily measurable soft traits, which have already outperformed the predictive capacity of phenotypic traits (Dussarrat et al., 2022), and (ii) conferring a mechanistic understanding of the underlying molecular aspects (Walker et al., 2022). Yet, the use of metabolomics to investigate plant-plant connections seemed mostly limited to negative relationships, as exemplified by the wide description of allelopathic compounds (Weir et al., 2004). However, positive interactions were shown with the capacity of plants to activate the defences of their neighbours through the emission of volatile organic compounds (Baldwin et al., 2006). Therefore, it is questionable whether ecological metabolomics could be used to test the potential effect of facilitation processes on both host and protected species. Then, metabolomics and ecology should be combined with machine learning to unveil predictive metabolic markers explaining the plant response to such interaction.

The Atacama Desert is the driest non-polar desert on Earth and offers an elevation gradient from 2400 to 4500 m.a.s.l where plant life is challenged by drought, high solar radiation and daily negative temperatures, for instance (Eshel et al., 2021). Fascinatingly, tens of plant species flourish in this transect such as the cushion cacti *Maihueniopsis camachoi* (Díaz et al., 2016). Hence, this ecosystem provides a unique opportunity to analyse the impact of the facilitation process from the ecosystem to the metabolism. Here, the coverage of multiple plant species was evaluated in different years to highlight different patterns of association with *M. camachoi*. Fieldwork measurements were coupled with mathematical modelling to assess the potential benefits of the nurse effect that may vary between species. Targeted and untargeted metabolomics analyses were performed to identify the metabolic patterns underlying the interaction between *Atriplex imbricata* and *M. camachoi*. A generalised linear modelling approach was deployed to unveil metabolic signatures predicting interaction status independent of the sampling year.

Results

Facilitation patterns depend on plant species and elevation. The Talabre-Lejía Transect (TLT) is an elevation gradient that includes a linear range of extreme environmental constraints for plant life (Eshel et al., 2021). This transect covers three distinct vegetation belts as follows: the Prepuna (2400-

3300 m.a.s.l) characterised by very low rainfall (20 mm/year), the Puna (3300-4000 m.a.s.l), and the Steppe (4000-4500 m.a.s.l) defined by daily freezing temperatures. In addition, plants are subjected to very low soil suitability (*e.g.* high salinity in Prepuna and low nitrogen levels), and high solar irradiance (600 W/m²/d) (Eshel et al., 2021). Despite these hostile conditions, the Atacama Desert hosts a unique reservoir of plant biodiversity. Notably, several plant species flourish across a significant elevation gradient while others remain exclusive to specific levels (Díaz et al., 2019). In the present study, we focused on the 20 species that were able to (i) develop in winter and (ii) cover a sufficient elevation delta to study their interaction with *M. camachoi* (Fig. V.1A and Table V.S1).

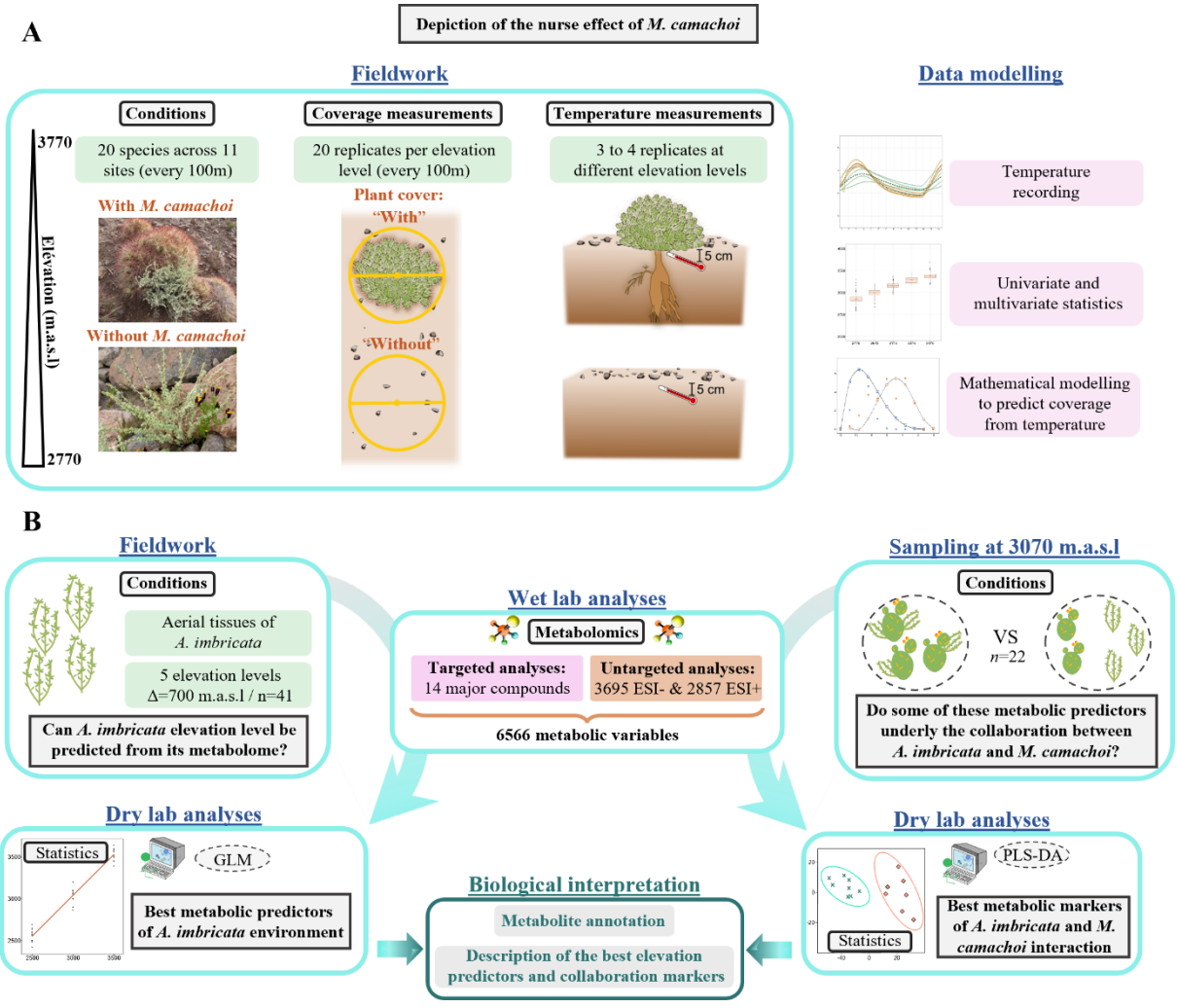

Figure V.1 | Global workflow. A. Fieldwork analysis. Depiction of the nurse effect provided by *M. camachoi*. Coverage measurements aimed to define the facilitation intensity. Plant cover “with” the cactus (cm/m² of cactus) was compared to the plant cover “without” cactus on an equivalent surface. **B.** Workflow of the metabolic analysis. *GLM*: Generalised multilinear models. *PLS-DA*: Partial least squares discriminant analysis. Both approach used the same metabolomic data (*i.e.* targeted and untargeted analyses).

The coverage of these species was evaluated at 11 sites across the TLT (*i.e.* one site every 100 meters from 2770 to 3770 m.a.s.l). This biodiversity encompassed C3, C4 and CAM carbon fixation systems, various lifespans (*i.e.* perennial and annuals) and life forms (*e.g.* shrubs, herbs) and plant families. Then, plant cover “with” cactus (cm/m² of cactus) was assessed for each species across the elevation gradients and compared to the related plant cover “without” cactus on an equivalent surface, which corresponded to the cactus surface (Fig. V.1A and Table V.S1). A total of 18 species (*i.e.* 90% of the studied species) were observed in interaction with *M. camachoi* in at least one elevation level. Three species called *Atriplex imbricata*, *Baccharis tola* and *Neuontobotrys tarapacana* established a significant positive relationship with *M. camachoi*, expanding their coverage at higher or lower elevations independent of year (Table V.S3). Additional species benefited from this relationship in one year and others presented clear tendencies (Table V.S3). These numbers were probably underestimated since most annual species from the Atacama Desert could not be observed in winter (Díaz et al., 2019). Overall, great inter-species variations were displayed with opposite trends between Prepuna (2400-3300 m.a.s.l) and Steppe (4000-4500 m.a.s.l) species (Fig. V.2). For instance, the C4 plant *Atriplex imbricata* showed a natural coverage between 2800 to 3700 m.a.s.l with a maximum cover between 2900 and 3200 m.a.s.l without positive interaction, while the C3 plant *Baccharis tola* developed between 3300 and 4000 m.a.s.l with a maximum cover between 3800 m.a.s.l and 4000 m.a.s.l. without positive interaction (Díaz et al., 2016). Thus, *Atriplex imbricata* favoured positive interaction at the upper limits of its life compatible gradient, while *Baccharis tola* evidenced the opposite pattern (Fig. V.2). Hence, the distribution of *A. imbricata* along the elevation cline seemed congruent with the hypothesis that temperature is a critical factor for C4 plant survival (Collins and Jones, 1986). In contrast, the nurse effect of the Cactaceae was likely to benefit *B. tola* at lower elevations where other abiotic factors are limiting (*e.g.* high salinity, water scarcity). Hence, these findings supported the stress-gradient hypothesis and illustrated the need of integrating phylogenic and phenotypic information into the analysis to provide a more integrative understanding of the facilitation phenomenon (Callaway et al., 2002).

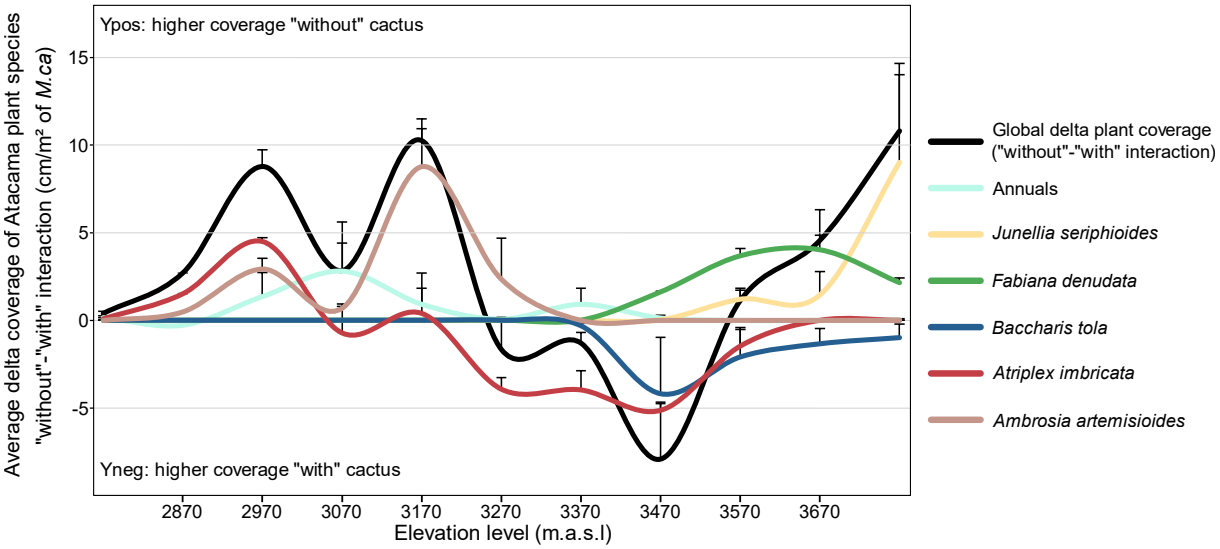

Figure V.2 | Depiction of the delta plant coverage with and without interaction with *M. camachoi* (*i.e.* coverage outside minus inside cactus). Average delta coverage of Atacama species significantly impacted by the interaction with *M. camachoi* (P<0.0001). Delta coverage were equal to the coverage “within” the cactus minus the coverage “with” the cactus. Thus, negative values corresponded to higher coverage with the cactus than without the cactus. Global delta plant coverage included the 20 analysed species. Other species interacted significantly with *M. camachoi* to a lesser extent (10 species at P<0.05) but were not represented in this graph (Table V.S4). Solid lines refer to discontinuous data.

The “nurse effect” of *M. camachoi* defines plant performances. To get insights into the potential benefits of the positive interaction of Prepuna species with *M. camachoi*, we (i) measured temperature at 5 cm of the soil surface over at least 24 hours at three elevation levels and (ii) deployed a mathematical modelling approach to study the coverage (Fig. V.1A). Interestingly, *M. camachoi* acted as a thermal buffer at various elevations, independent of the year (Fig. V.3). Notably, the protective effect of the facilitation process seemed to occur between approximately 19:00 and 9:00 based on the assumption that the potential benefit of this process resided in protection against cold temperature for Prepuna species. Besides, the puffer effect of *M. camachoi* was not linearly correlated and varied between 1.4 and 4.4°C (Table V.S2). This delta seemed related to cactus size (Fig. V.S1), in agreement with previous reports (Schöb et al., 2013).

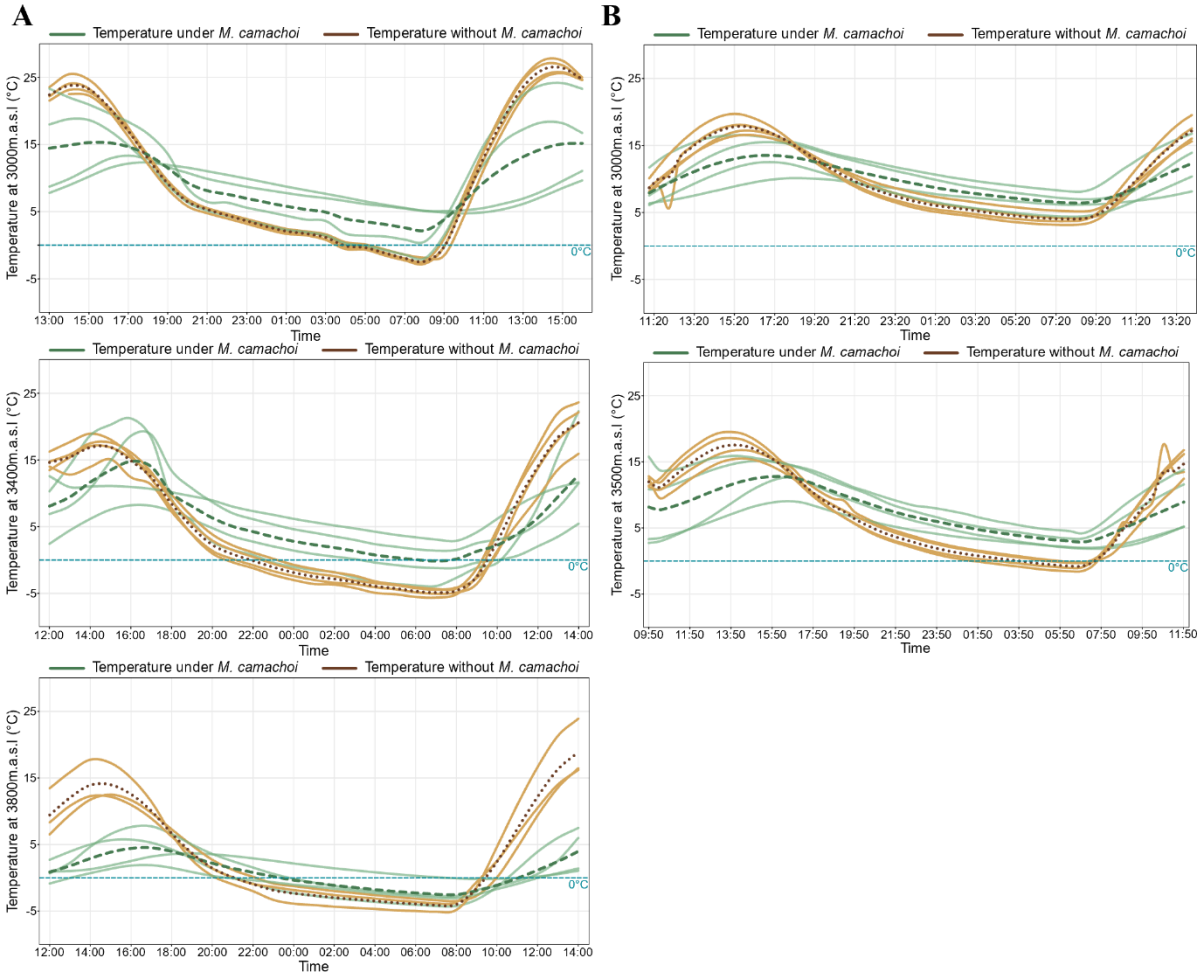

Figure V.3 | The nurse effect of *M. camachoi.* A. Characterisation of the soil temperature delta in presence or absence of *M. camachoi* in 2016. **B.** Characterisation of the soil temperature delta in presence or absence of *M. camachoi* in 2021. Temperatures were assessed at 3000, 3400, 3500, or 3800 m.a.s.l depending on the year. Straight lines represented the replicates (3 to 4 replicates per elevation site), while dotted lines illustrate the average of the replicates.

The interaction between *M. camachoi* and *A. imbricata* was then used to investigate the facilitation effect. Notably, the root system of the protected plant seemed to thrive in the decaying material of the cactus, suggesting an epiphyte relationship (Fig. V.S2). More importantly, the interaction status of *A. imbricata* showed a significant shift around 3170 m.a.s.l (Fig. V.4A). Mathematical modelling was then employed to get insights into the impact of temperature on the coverage of *A. imbricata* using a general equation established to simulate the temperature response of plants $c = C_{max}\left(\frac{F_{max}-T}{F_{max}-T_{opt}}\right)\left(\frac{T}{T_{opt}}\right)\text{^}(\frac{T_{opt}}{F_{max}-T_{opt}})$ (Yan and Hunt, 1999). Here, T_{opt} (optimal temperature for plant coverage) and C_{max} (maximum coverage of *A. imbricata*) were defined accordingly to fieldwork analyses (as detailed in Chapter 2 section VIII.3), T represented the temperature with or without cactus, while F*max* (theoretical maximum value for the arbitral limiting factor at low elevations) could here encapsulate water, salinity or nitrogen stress, for instance. First, the T_{opt} parameter was fixed as the temperature reported at 3370 m.a.s.l since the maximum coverage (C_{max}) of *A. imbricata* was observed between 3270 and 3470 m.a.s.l for 2021 and 2016 datasets (Fig. V.4A). Similarly, the F_{max} value was artificially defined as the temperature detected at 2770 m.a.s.l, thus forcing the model to limit plant life at the lower levels (Table V.S4). Importantly, relevant stability of the different parameters was noticed between the interaction status, which supported the viability of the model. For instance, the average C_{max} reached 5.25 and 3.97 cm/m² (average C_{max} measured in 2021 and 2016) and was related to an average T_{opt} of 8.08 and 8.17 °C under the “without” and “with” interaction conditions, respectively (Table V.S4). This equation was then used to calculate the theoretical coverage (c) score for each elevation level and therefore appreciate the impact of temperature on *A. imbricata* performances. Mathematical model predicted with plant coverage with 76% and 63% accuracy for the “without” and “with” conditions in 2021, respectively (Fig. V.4B and V.4D). Model accuracy was then confirmed using an independent dataset. The F_{max}, T_{opt} and C_{max} parameters established using 2021 data were preserved to characterise the temperature effect on plant coverage in 2016. Thus, temperatures recorded in 2016 were used to define the T parameter and predict Atriplex cover (c) in 2016. The predictive capacity of the model was then confirmed by R^2 scores of 72% and 54% using 2016 coverage and temperature data (Fig. V.4C and V.4D). Besides, predictions reached 72% and 78% accuracy if considering *Atriplex* coverage at 3070 (Fig. V.4C, green circle observed at 7 °C) as an outlier. Hence, these results showed that temperature alone was an excellent predictor of *A. imbricata* survival in the Atacama Desert. By extension, we here revealed that the puffer effect of *M. camachoi* facilitated and even enabled the extension of *A. imbricata* life compatible gradient to high elevations.

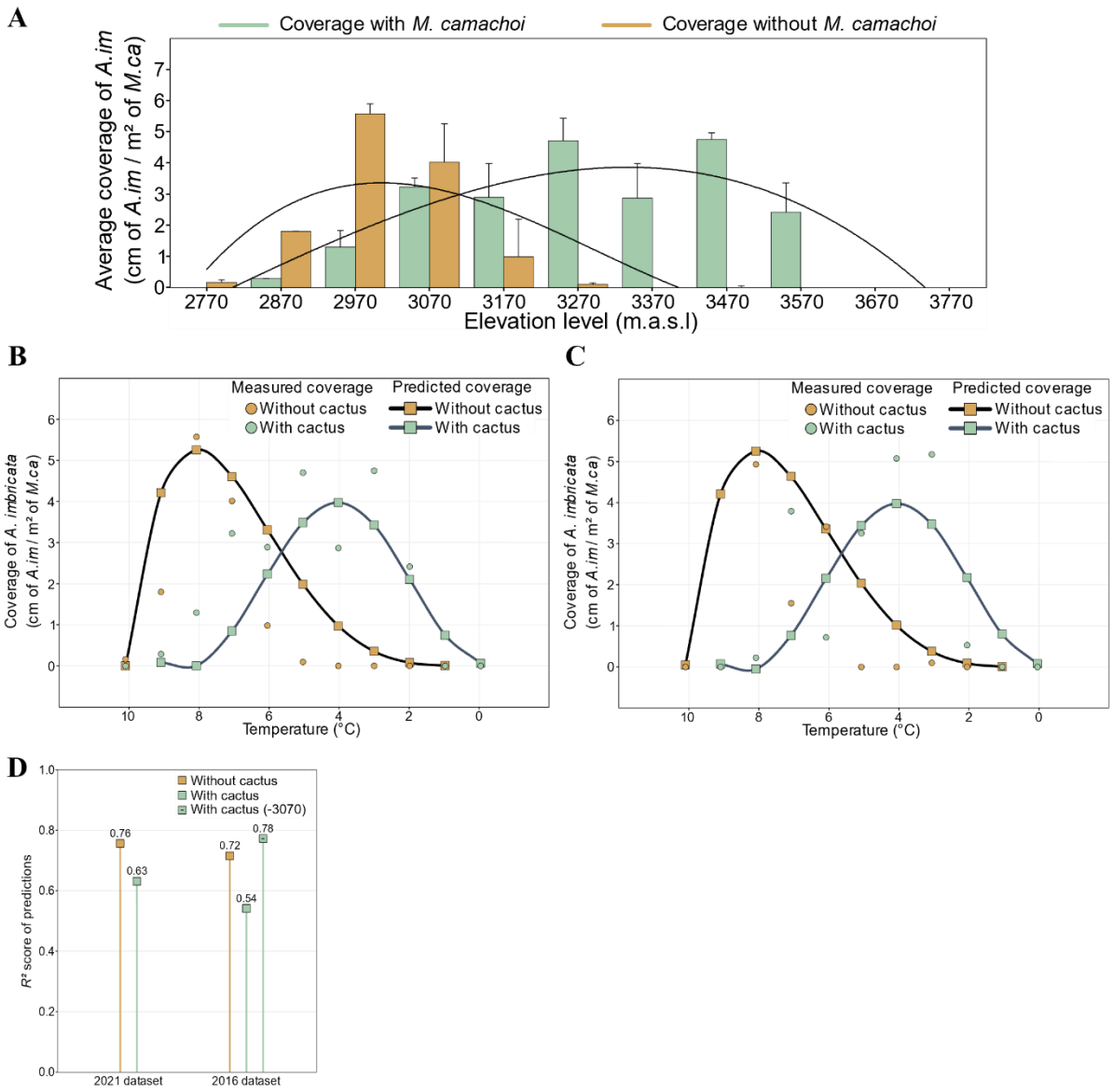

Figure V.4 | The nurse effect of *M. camachoi* extends *A. imbricata* survival to higher elevations. **A.** Average coverage of *A. imbricata* (*A.im*) with and without interaction with *M. camachoi* (*M.ca*) across the elevation gradient. Pearson standard deviations were calculated on the average coverage from 2021 and 2016 datasets. **B.** Prediction of *A. imbricata* cover using a mathematical model performed on the 2021 dataset. **C.** Biological validation of the nurse effect using an independent dataset. Values of T_{opt}, F_{max} and R_{max} defined on 2021 data were directly applied on 2016 data. **D.** R^2 scores of the predictions performed on 2021 and 2016. **Note**: Figure 4B and 4D were performed using 2020 temperatures and are thus under validation using 2021 real temperatures, which are currently under treatment (as detailed in Chapter 2 section VIII.3). However, average temperatures did not show major changes between the different years (from 2016 to 2020).

Plant metabolome predicts interaction status. The plant metabolome has the fascinating ability to integrate biotic and abiotic variations in complement to genome influences. Thus, an eco-metabolomics approach was conducted to explore the effect of the positive interaction between *A. imbricata* and *M. camachoi* (Fig. V.1B). We quantitatively evaluated 12 major compounds through robotised biochemical assays and assessed the biochemical diversity via liquid chromatography-mass spectrometry (LCMS) analysis to provide a fingerprint of the metabolome from *A. imbricata* collected in 2019 and 2021 (Table V.S5). Multivariate statistical analyses, orthogonal partial least squares discriminant analyses (OPLS-DA) and generalised linear modelling (GLM) methods were deployed for a thorough investigation of the metabolic features underlying the facilitation process (Fig. V.1B). First, the resulting metabolomics dataset (6556 variables after preprocessing) was processed through a GLM approach to test the extent to which *A. imbricata* could predict the environmental conditions as previously described (Dussarrat et al., 2022) (Fig. V.S3). Elevation level was used as a proxy of plant environment based on its integration of abiotic and biotic factors (Carpenter, 2005). The model equation was developed on 80% of the sample set (*i.e.* training and testing sets) composed of plants collected in April 2019 and 2021. The equation was then applied on the 20% left (*i.e.* validation set) to test the predictive capacity of the model. This first modelling step was used to select the 1% most predictive variables (*i.e.* 66 features) based on their occurrence in the model. Results showed that plant metabolome from *A. imbricata* predicted environmental conditions (*i.e.* elevation level) with an average accuracy of 87% (*i.e.* average R^2 from the 500 models) using 66 variables (Fig. V.S3). Finally, 500 permuted datasets were developed to test the likelihood of spurious predictions and yielded a mean R^2 of 0%, thus statistically validating the models. Notably, the intensities of the best 66 predictors suggested a bidirectional protective strategy marked by specific metabolic signatures at low (2770-2970 m.a.s.l) and high (3370-3470 m.a.s.l) elevations (Fig. V.S3). Hence, these findings (i) confirm the excellent integrative capacity of plant metabolome, and (ii) support the results of the previous ecophysiological analysis, which identified the 3170 m.a.s.l site as a critical elevation level requiring a shift in interaction status for survival (Fig. V.4).

To better understand the potential link between the protective strategy employed by *A. imbricata* at high elevations and the facilitation process, we combined metabolomics and machine learning to explore the impact of the positive interaction on *Atriplex* plants collected at 3070 m.a.s.l in July 2021 (Fig. V.1B). The biochemical biodiversity (6556 variables) was first processed using t-test analysis to identify 440 significant markers of the interaction with *M. camachoi* (Fig. V.5). Interestingly, several of the top 1% metabolic markers predicting elevation level were included in these 440 markers, and others showed relevant trends (Table V.S6). The discriminatory ability of these markers was assessed using unsupervised statistics (*i.e.* principal component analysis) and yielded a significant distinction of the interaction status, with a first component explaining almost 50% of the variance, for instance (Fig. V.S4). Besides, OPLS-DA allowed testing these 440 features towards the prediction of the interaction

status. Models displayed a tremendous predictive capacity with an R^2 score of 0.81 and a Q^2 score equal to 0.73 (Fig. V.5). Importantly, these models were statistically validated using permutations (Table V.S7).

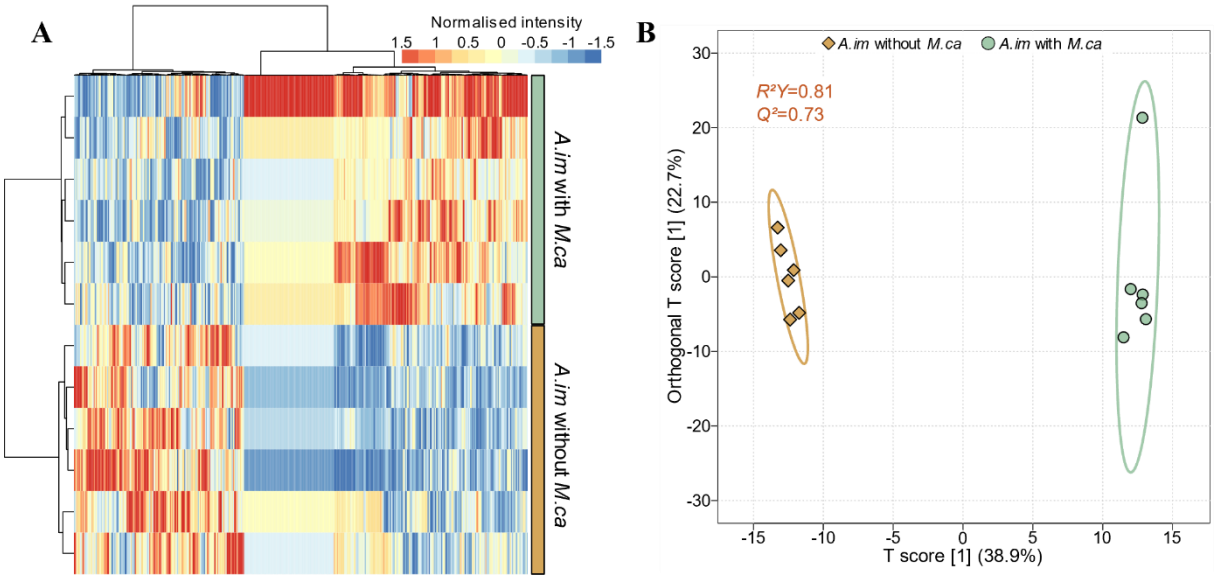

Figure V.5 | Metabolic pattern underlying *A. imbricata*-*M. camachoi* interaction. A. Depiction of the 440 significant metabolic features from *A. imbricata* underlying the interaction with *M. camachoi* (Pearson correlation, Ward algorithm). **B.** Prediction via OPLS-DA model of the *A. imbricata* interaction status using significant markers.

To improve the robustness of these markers and validate their significance over years, we tested their response on the entire set of *A. imbricata* plants (including plants collected in April 2019 and 2021). The interaction status of *A. imbricata* plants was defined according to their photographs (collected in the field during the sampling) when available (n=28), while plants without pictures were removed from the analysis. Notably, 98 features were expressed differentially between the interaction states and allowed a satisfactory classification of the samples, except for one sample that could be considered an outlier based on its metabolic profile (Fig. V.6 and Table V.S8). Ultimately, both independent datasets (*i.e.* July 2021 and April 2019 plus April 2021) were combined to test whether these 98 markers could efficiently predict the interaction status via GLM models. These markers provided a mean R^2 score of 83%, thus demonstrating a high predictive value independent of sampling year (Fig. V.6). Notably, 500 permutation tests were realised to validate the predictive capacity of these markers and yielded an average R^2 of 62.5%. This number was explained by the fact that the validation sets included 8 samples in total among which 5 were classified as “Without interaction” (*i.e.* 20% of the total dataset). Hence, models assigned the “Without interaction” status to all samples based on their incapacity to predict better, yielding an accuracy of 62.5% (5 good predictions on 8 samples). Our models were therefore statistically validated and confirmed via OPLS-DA analyses, which provided significant R^2 and Q^2

scores (Fig. V.6, Fig. S5 and Table S7). Finally, the previously described analytical workflow was applied on *M. camachoi* to speculate on the potential impact of the facilitation process on the host plant. Although the identified markers could not be validated biologically using an independent dataset, results suggested a significant impact of the interaction on 249 variables (Fig. V.S6). Besides, these predictors could efficiently predict the interaction state (Table V.S7). Overall, we here established for the first time that the puffer effect of *M. camachoi* induced a biochemical impact on protected plant species through the discovery of 440 potential markers. Even more entertaining, we highlighted a set of 98 variables capable of predicting the interaction status with 83% accuracy, independent of sampling year. Notably, these predictors hold great promise for improving our mechanistic understanding of the facilitation process.

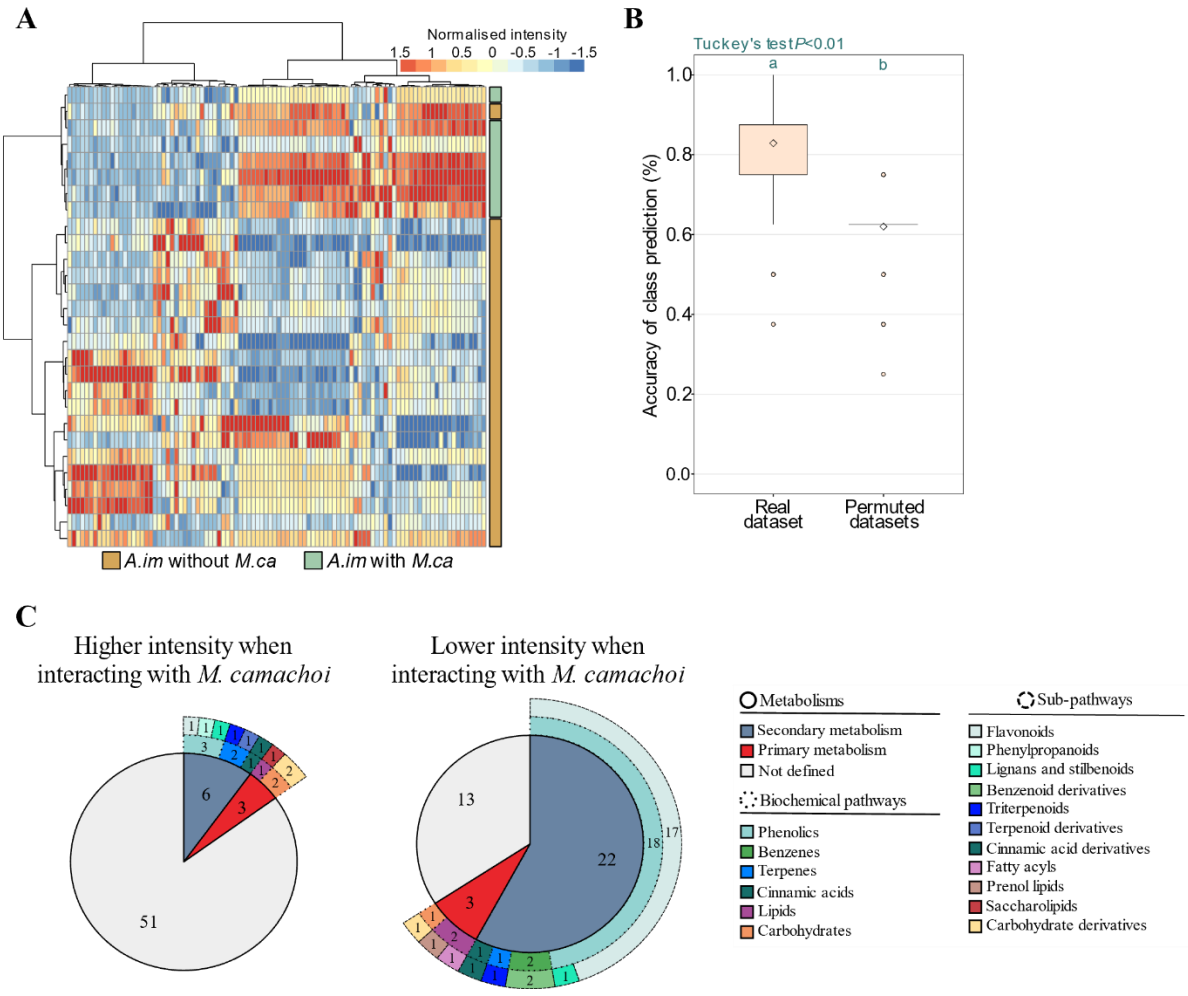

Figure V.6 | Predictive metabolomics unveiled metabolic markers predicting interaction with *M. camachoi, independent of sampling year*. A. Selection of the significant markers in independent datasets (*i.e. A. imbricata* collected in April 2019 and 2021) as well as in plants collected in August 2021 (Pearson correlation, Ward algorithm). ANOVA test was performed to highlight 98 significant markers (P<0.05). **B.** Average R^2 scores of 500 GLMs where the binomial variable “interaction state” was predicted from *A. imbricata* metabolome. Models were statistically validated using 500 permuted datasets (Tukey’s test, P<0.01). **C.** Annotation of the significant markers.

The facilitation process reshapes plant metabolic strategy. The best predictors of environmental conditions and biotic interactions were then annotated using accurate *m/z* values and MS/MS data (Fig. V.1 and Table V.S9). The annotation level was defined according to the metabolomics standards initiative confidence level (MSI levels) (Sumner et al., 2014). The KEGG identifier of the annotated molecules was then used to classify the markers in major biochemical pathways. First, more than half of identified 66 variables predicting elevation level belonged to secondary metabolism. *A. imbricata* accumulated a broad range of flavonoids and cinnamic acids at lower elevation levels, in agreement with previous results (Dussarrat et al., 2022). Besides, primary metabolites, including carbohydrates and organic acids, were negatively correlated with elevation level, and a significant shift in lipidic profiles was observed. Finally, results suggested a potential role of abscisic acid and jasmonates at high and low elevations, respectively (Table V.S9). Next, the 98 interaction markers from *A. imbricata* were classified as positive (*i.e.* higher intensity when interacting with *M. camachoi*) and negative (*i.e.* lower intensity when interacting with *M. camachoi*) predictors (Fig. V.6). Importantly, almost 70% of the annotated molecules negatively correlated with interaction were flavonoids, which agrees with previous eco-metabolomics studies (Defossez et al., 2021, Dussarrat et al., 2022). In addition, primary metabolites such as lipids were represented to a lesser extent among these negatively correlated predictors (Fig. V.6). Strikingly, a large proportion of these markers were related to the 39 metabolites predicting plant environment with 79% accuracy in multiple species thriving in the Atacama Desert (Dussarrat et al., 2022). For instance, secondary compounds such as quercetin, luteolin and coumaroyl derivatives were described as protective metabolites against osmotic stress in Prepuna and were negatively linked to the interaction with *M. camachoi* (Fig. V.6 and Table V.S9). In contrast, the extreme majority of the 59 positively correlated markers were unknown. Last but not least, the annotation of the best predictors from *M. camachoi* showed a similar profile (Fig. V.S6). Fascinatingly, while flavonoids represented the majority of the negatively correlated metabolites, a significant part of the positive markers remained unknown. However, an intriguing pattern arose from the analysis with the occurrence of terpenes, which accounted for more than 30% of the annotated positively correlated compounds (Fig. V.S6 and Table V.S9). Notably, additional markers from the 440 *Atriplex* compounds responding significantly to the interaction status were subsequently subjected to the annotation process. Flavonoids and terpenoids were the most represented biochemical class among those markers (Table V.S10). Altogether, these results suggested that the facilitation process not only favours seedling establishment but also reshapes the metabolic strategy employed by plant species to cope with their extreme environment.

Overall, this study further reinforces the potential benefits of an approach combining ecology and predictive metabolomics. The integration of ecophysiological, environmental and metabolic data has profoundly expanded our understanding of the facilitation process from the ecosystem to the metabolic scale. Besides, while phenotypic analyses highlighted the different advantages provided by

the positive interaction with *M. camachoi*, metabolomics deciphered its consequences on the biochemical level and unveiled a set of metabolites predicting the interaction status with 83% accuracy, independent of year.

III. THE PLACE OF FACILITATION PROCESS IN ATACAMA DESERT

The central place of biotic micro-environments in plant adaptation to extreme lands. The Atacama Desert, one of the oldest deserts on Earth, developed as a direct consequence of hyperaridity cycles 12 million years ago (Jordan et al., 2014). The TLT transect spans an elevation cline from 2700 to 4500m.a.s.l in the Atacama Desert and harbours fascinating plant biodiversity (Díaz et al., 2019). Interestingly, these species developed genetic and metabolic adaptation to face extreme stress gradients such as drought, temperature and salinity (Dussarrat et al., 2022; Eshel et al., 2021). However, although the evolution of genetic and metabolic strategies enabled the adaptation of multiple plant species to Atacama conditions, these organisms may not perform at their physiological optimum across the entire transect (Nicotra et al., 2010). The severity of the stress gradient progressively constraints the performance of the species until the critical point where the stress overcomes the phenotypic and metabolic plasticity of the species (Fig. V.2). At this point, the survival of plants depends on their interactions with biotic communities such as microorganisms or plants (Brooker et al., 2007).

Cushions seemed of major interest in structuring the Atacama plant community based on their significant influence on plant coverage across the different vegetation belts (Fig. V.2). The importance of niche construction by cushions has been widely described in alpine and arid regions (Callaway et al., 2002). The phenotypic traits of cacti confer protective thermal and hydric properties from which other species can benefit (Cavieres et al., 2006). These physicochemical properties have even earned the cactus a central place in environmental restoration programmes (Padilla and Pugnaire, 2006; Zhao et al., 2007). Here, a positive relationship with *M. camachoi* impacted significantly the coverage of five Atacama plant species (Table V.S3). Besides, the facilitation intensity (*i.e.* the difference between the number of individuals within and without the Cactaceae) rose dramatically between 3300 and 3550 m.a.s.l, while decreasing at lower and higher levels (Fig. V.2). Hence, these findings suggested that the requirement for positive interaction was not linked to elevation *per se* but the stress diversity and intensity. Notably, the high Andean Steppe was mainly colonised by C3 grasses and shrubs, whilst C4 species inhabited the Prepuna (Díaz et al., 2019). In agreement with previous studies, these observations strongly suggested a considerable influence of temperature and osmotic stress on the performances of Prepuna and Steppe species, respectively (Dussarrat et al., 2022). For instance, a consequence of the C4 carbon fixation system is a lower quantum yield (*i.e.* the initial slope of the photosynthetic light response curve) than C3 species at moderate to low temperature (Ehleringer et al., 1997). In this context, the

analysis of the puffer effect of *M. camachoi* was confirmed via fieldwork measurements and mathematical modelling (Fig. V.3 and V.4). In addition, the positive interaction with *M. camachoi* not only seemed to generate a facilitation process but rather represented a vital resource for extending the life compatible gradient of certain species (Fig. V.1). Results showed that the presence of *B. tola* and *A. imbricata* at the edge of their life gradient (both occurring around 3400-3500m.a.s.l) were dependent on this positive relationship (Fig. V.1 and V.4). Importantly, the highest diameters of *M. camachoi* reported between 3470 and 3570m.a.s.l coincided with the highest facilitation intensity level (Fig. V.S1 and V.1). These elevations corresponded to the Puna, where a major shift between Prepuna and Steppe species occured (Díaz et al., 2019), thus supporting the dynamic role of this nurse effect in determining the organisation and diversity of Atacama plant communities. Hence, these findings challenged the individualistic theory (*i.e.* the concept that the plant richness pattern of a given species is independent of other species) and suggested that the impact of facilitation on ecosystem dynamics remains underestimated (Gleason, 1926; Callaway et al., 2002).

Protective functions of the facilitation process are reflected by an adapted metabolic response. Plant-plant interactions are now recognised as an essential driver of terrestrial community organisation (Bruno et al., 2003). Although significant efforts allowed characterising the influence of facilitation on plant survival and seedling establishment, the impact at the metabolic level and the underlying biochemical mechanisms remain surprisingly unexplored (Cavieres et al., 2006; Brooker et al., 2007). Our ecological metabolomics approach combined with GLM learning provided, to our knowledge, the first analysis of these mechanistic aspects. Here, we showed that positive relationships with *M. camachoi* reshaped the metabolic strategy employed by plants to face extreme environmental constraints. Subsequently, this analysis unveiled a set of metabolites predicting the interaction status independent of year, thus supporting the potential of this approach to improve our comprehension of complex life systems.

The metabolome of *A. imbricata* was an excellent integrator of environmental variations (Fig. V.S3). Predictive metabolomics highlighted a bidirectional strategy employed by *A. imbricata* to face cold temperatures and osmotic stress (Fig. V.S3 and V.6). At high elevation levels (*i.e.* 3300 to 3600m.a.s.l), plant survival was linked to the positive interaction with *M. camachoi*. This relationship induced a significant response of 440 features in the protected plants. Strikingly, 98 markers predicted the interaction status at 83% accuracy, regardless of year. The annotation of these predictors pinpointed a global reorganisation of the secondary biochemical strategy (Fig. V.6). Results suggested that the micro-environment offered by *M. camachoi* provided favourable conditions under which the carbon allocation shifted from the production of protective compounds to other defensive or developmental mechanisms. For instance, flavonoids, cinnamic acids and benzoic acid derivatives were prominently represented among predictors negatively correlated with interaction status (Table V.S9). Besides, the synthesis of quercetin, luteolin and coumaroyl derivatives were lower in interacting plants, while

identified as a defensive response to osmotic stress in multiple Prepuna species from the Atacama Desert (Dussarrat et al., 2022). Remarkably, studies showed that the antioxidant properties of flavonoids mitigated a broad range of abiotic stress, such as low water availability and high salinity (Agati et al., 2012; Dussarrat et al., 2021). Intriguingly, an alternative hypothesis could reside in the potential link between flavonoids and plant-microorganisms interactions (Jeon et al., 2021; Yu et al., 2021). This hypothesis was supported by the potential epiphyte pattern of *A. imbricata* characterised by the root system developed in the decomposing cactus material (Fig. V.S2). Hence, living in this protected micro-environment could eventually lead to a lower interaction with soil micro-organisms. Importantly, the protective biochemical strategy of *A. imbricata* was, however, not restricted to the relationship with the Cactaceae. The prediction of the plant environment at 87% accuracy involved several compounds positively or negatively affected by elevation but uncorrelated to the interaction status (Fig. V.S3 and Table V.S9). For instance, accumulated levels of fatty acyls and flavonoids, as well as abscisic and ascorbic acids were exposed at high altitudes. These compounds could serve plant adaptation by mediating the defensive response against environmental constraints such as solar irradiance, which can not be alleviated by *M. camachoi* (Berli et al., 2009; Smirnoff, 2018). Similarly, contents in azelaic acid and jasmonate relatives increased at low sites, suggesting either higher herbivory or a higher osmotic pressure at these elevations (Arimura et al., 2000; Machado et al., 2013; Marti et al., 2013). Thus, these findings proposed that the adaptation of *A. imbricata* was underpinned by the evolution of (i) an individualistic metabolic strategy and (ii) a dependence on positive interaction with *M. camachoi* at the highest elevations. Interestingly, the benefits of the plant-plant interaction might not be restricted to the protected species but rather extended to the host as a similar metabolic response (*i.e.* a decrease in the synthesis of protective compounds) was observed in *M. camachoi* (Fig. V.S6) (Schöb et al., 2014).

Overall, our results characterised facilitation as a stress-gradient dependent process that profoundly influenced plant community structure by providing thermal and osmotic refuge for a wide range of plant species under harsh conditions (Fig. V.7). In the Atacama Desert, these micro-environments favoured the development of several plant species and inputted a greater resilience to major abiotic stresses such as temperature and drought. This enhanced resilience was illustrated by the decline in the production of several major defensive compounds. Thus, while these results reinforced the role of metabolites previously revealed as excellent predictors of the environment, they also highlighted the importance of interspecies dependence within this unique ecosystem (Fig. V.7). Besides, although it was difficult to speculate on the role that cacti played in the evolution of plant biodiversity tens of thousands of years ago, it is reasonable to say that this *Cactaceae* is currently assuming a pivotal position in the equilibrium of the plant community. Thus, understanding this facilitation process from ecosystem to metabolic levels is vital for anticipating the effect of increased global aridity on *M. camachoi* and, by extension, to predict the response of the entire Atacama system.

In addition, our eco-metabolomics approach unveiled (i) a set of metabolites predicting the interaction status with accuracy and (ii) a significant effect of facilitation on the metabolic strategy employed by *A. imbricata* to face extreme conditions. Hence, applying such an approach to multiple plant species would help unravel convergent metabolic mechanisms underlying plant-plant interactions, which seemed critical to develop more accurate and realistic ecological models.

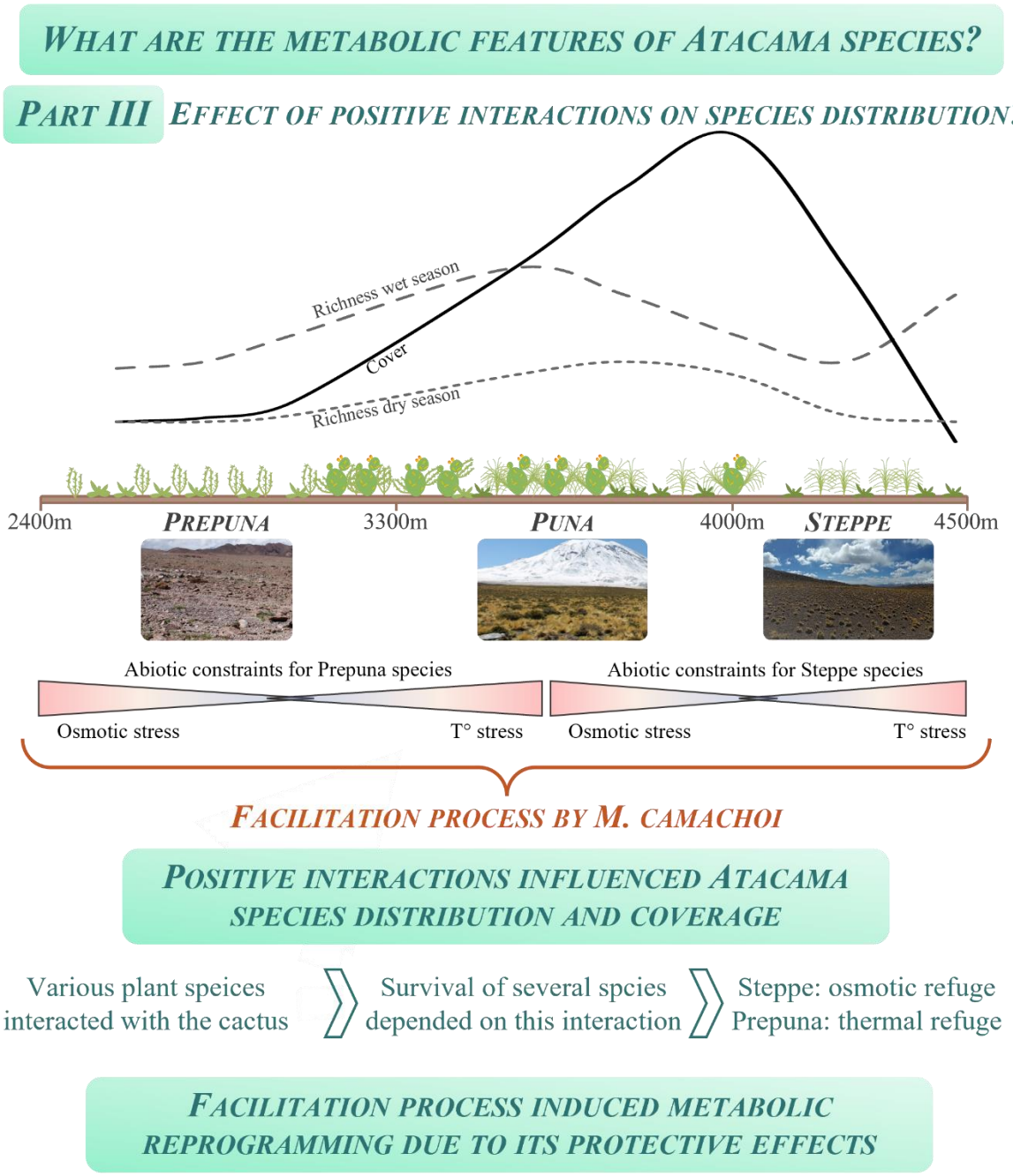

Fig. V.7 | The secrets of plant adaptation to the Atacama Desert.

Chapter 5: Positive interactions with *M. camachoi* influence Atacama species distribution. Schematic curves of plant cover and richness were adapted from Díaz et al., 2019.

Supplemental figures. Supplemental tables are available at the following link until publication:

https://drive.google.com/drive/folders/1Z3HLMY0Hb281HEu56MM82MHtY_YQ9tkE?usp=sharing

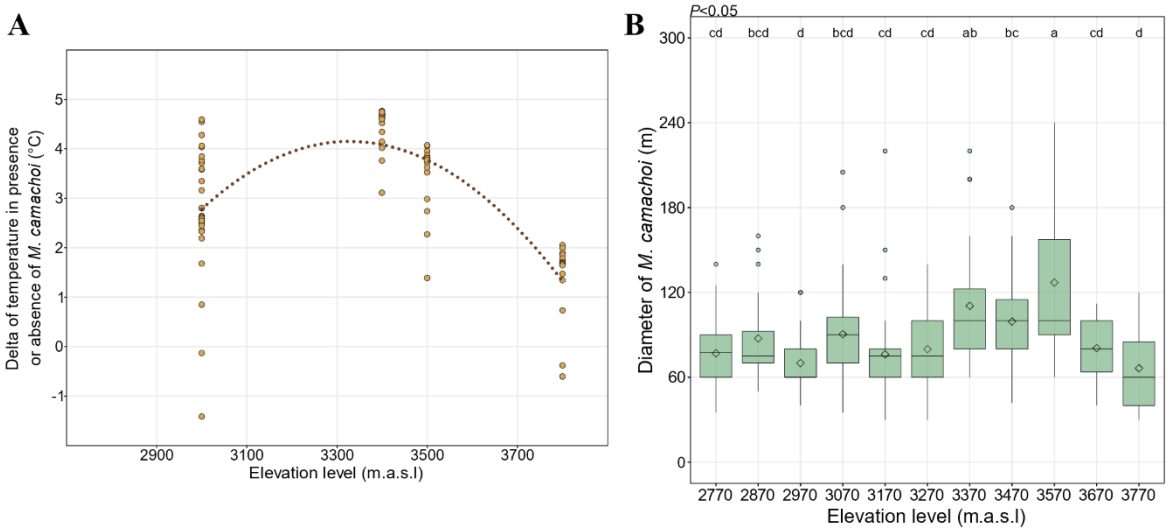

Figure V.S1 | Non-linearity of the delta of temperature potentially explained by the size of *M. camachoi.* A. Delta of temperature according to the elevation level. Dots represented the measured temperatures between 7 pm and 9 am where *M. camachoi* served as a thermal refuge. **B.** Depiction of the size of *M. camachoi* according to the elevation level.

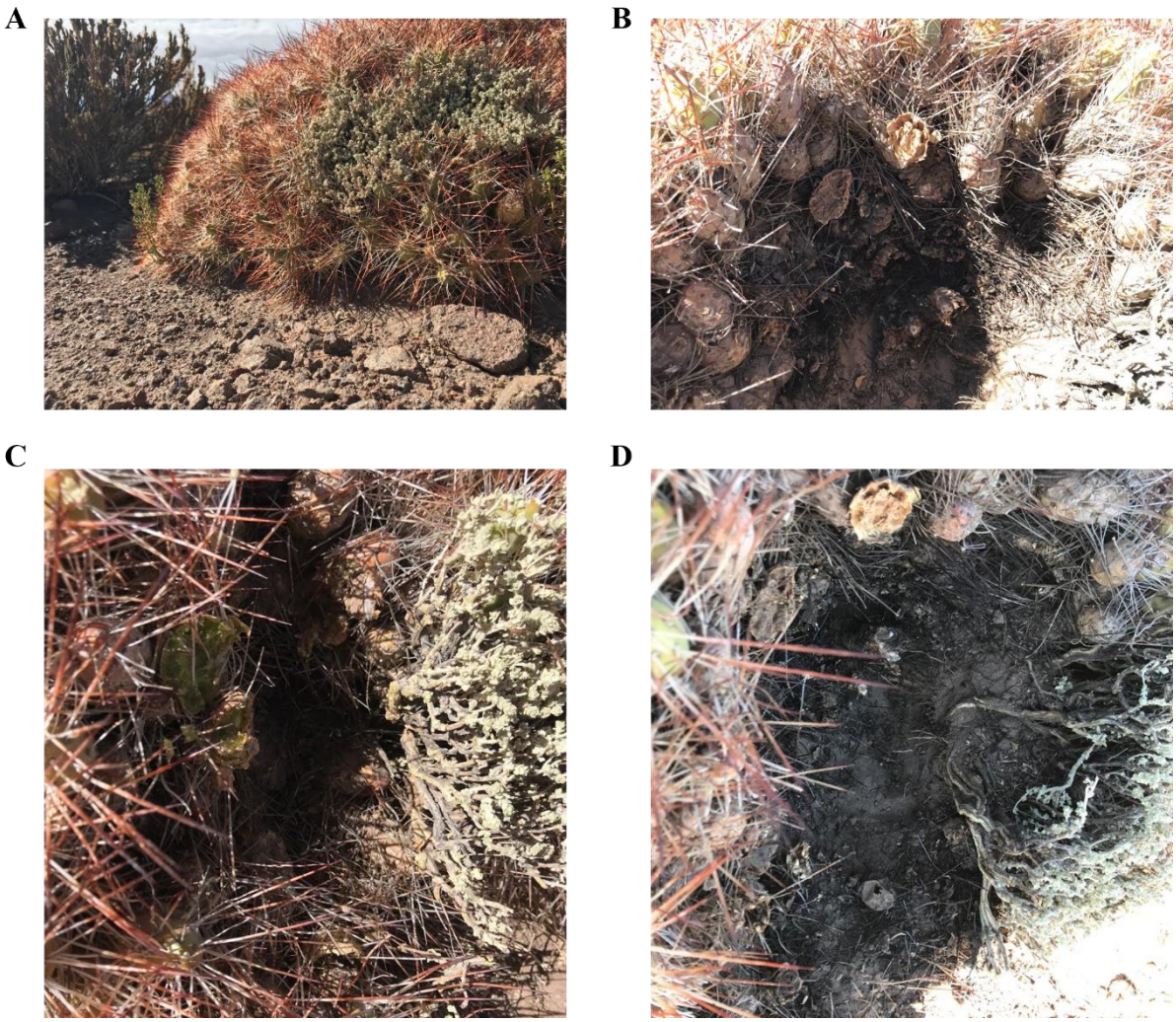

Figure V.S2 | Root system of *A. imbricata* within *M. camachoi*. Pictures illustrated the interaction of *A. imbricata* with *M. camachoi*. **A.** Positive interaction between the two species. **B.** The layered growth of *M. camachoi*. Living material seemed to blossom on the top of previously dead organic matter. **C** and **D.** The root system of *A. imbricata* seemed to thrive in the decaying material of the cactus.

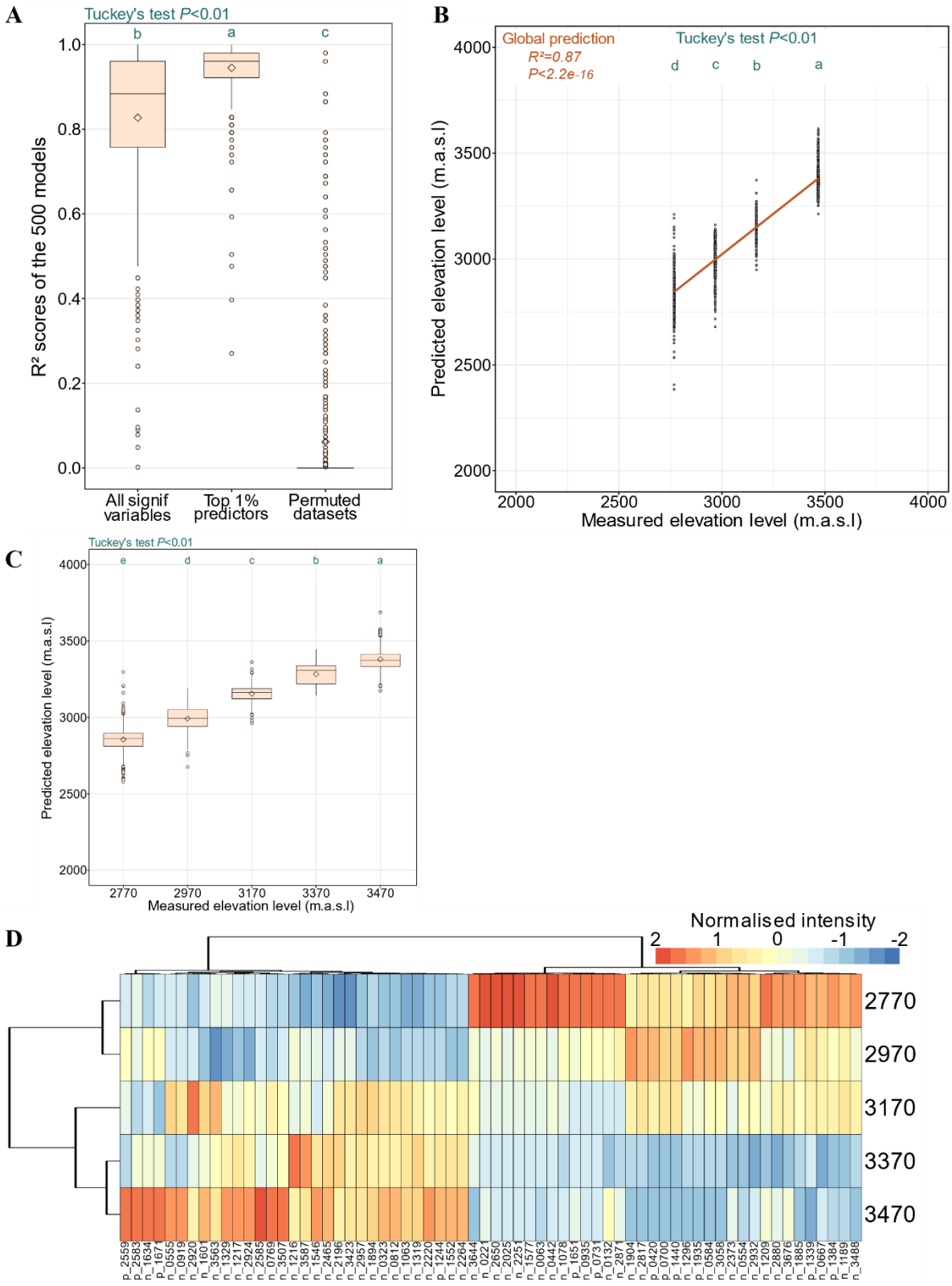

Figure V.S3 | Predictive metabolomics identified metabolic variables predicting elevation level of *A. imbricata*. **A.** R^2 scores of GLM models using all significant variables, top 1% predictors (*i.e.* 66 variables), or permuted datasets. 500 permuted datasets were created to test the likelihood of spurious predictions (Tuckey's test, $P<0.01$). **B.** Predicted elevation levels from the top 1% metabolic predictors (Pearson correlation). Comparison of the 500 predictions for each measured elevation level (Tuckey's test, $P<0.01$). **C.** Depiction of the fits (*i.e.* Predicted versus measured elevation levels) on the testing set using the top 1% predictors. **D.** Heatmap of the best 66 predictors (*i.e.* top 1%).

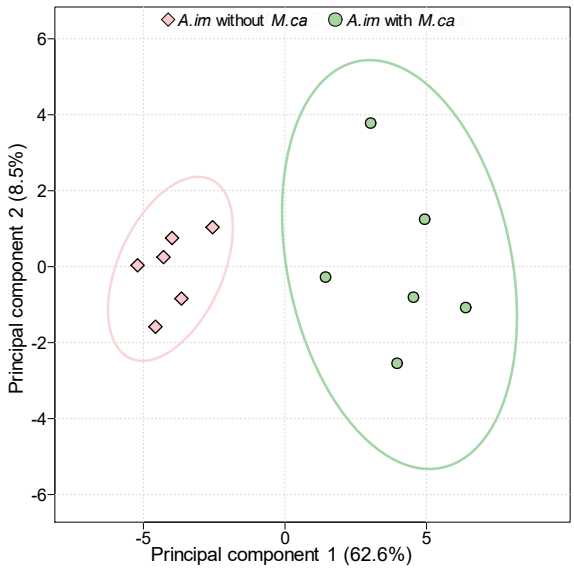

Figure V.S4 | Unsupervised method to analyse the discriminatory capacity of A. imbricata metabolism. Principal component analysis using the 440 significant markers**.**

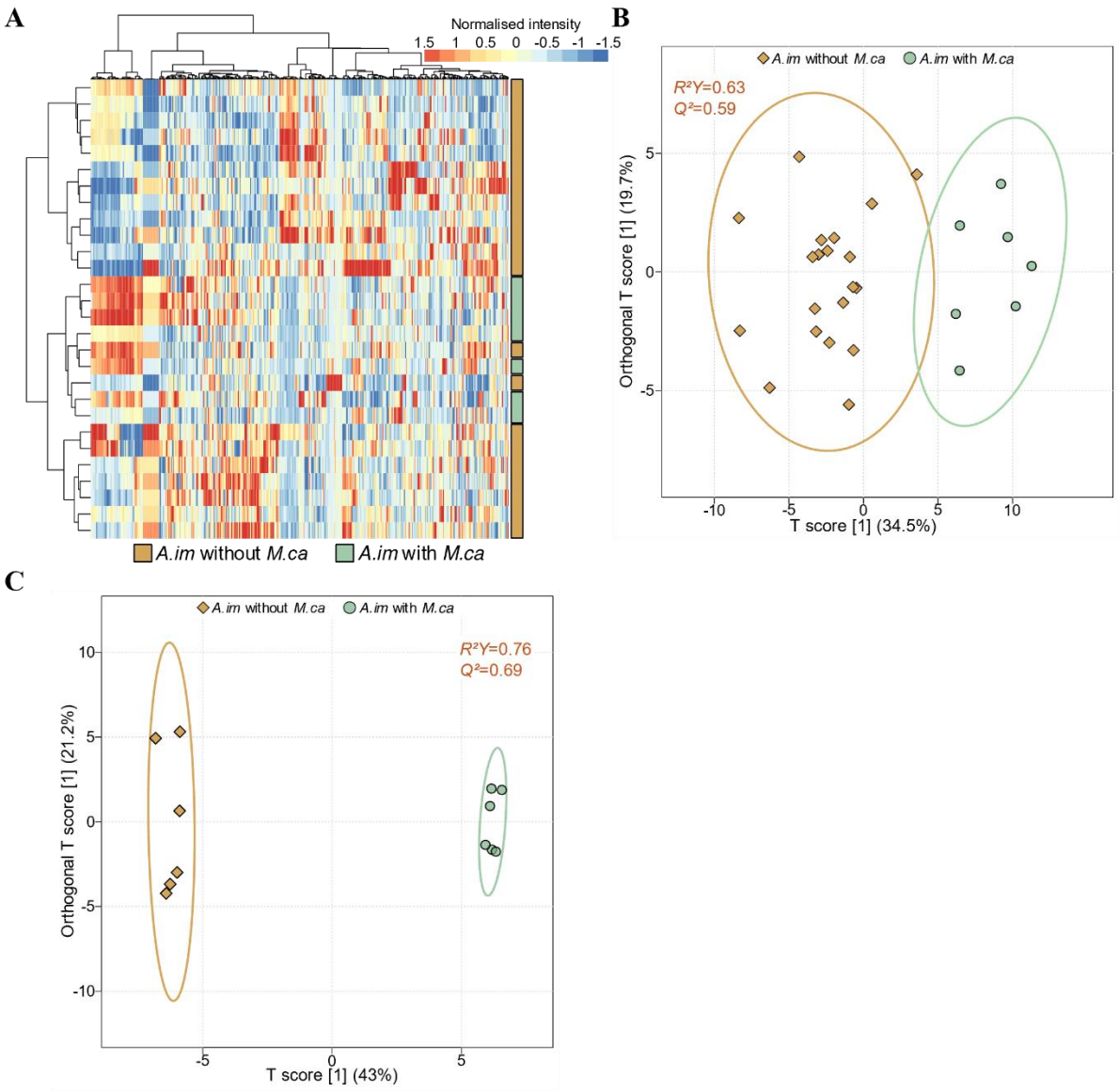

Figure V.S5 | Predicting the interaction status of *A. imbricata* from its metabolome over years. **A.** Clustering of the *A. imbricata* samples harvested in 2019 and 2021 in the Talabre-Lejía transect using the 440 significant compounds highlighted via the "*A. imbricata-M. camachoi* interaction" experiment (Pearson correlation, Ward algorithm). **B.** Prediction via OPLS-DA model of the *A. imbricata* interaction status using significant markers in all conditions. **C.** Prediction via OPLS-DA model on plants collected in August 2021 using the 98 significant markers.

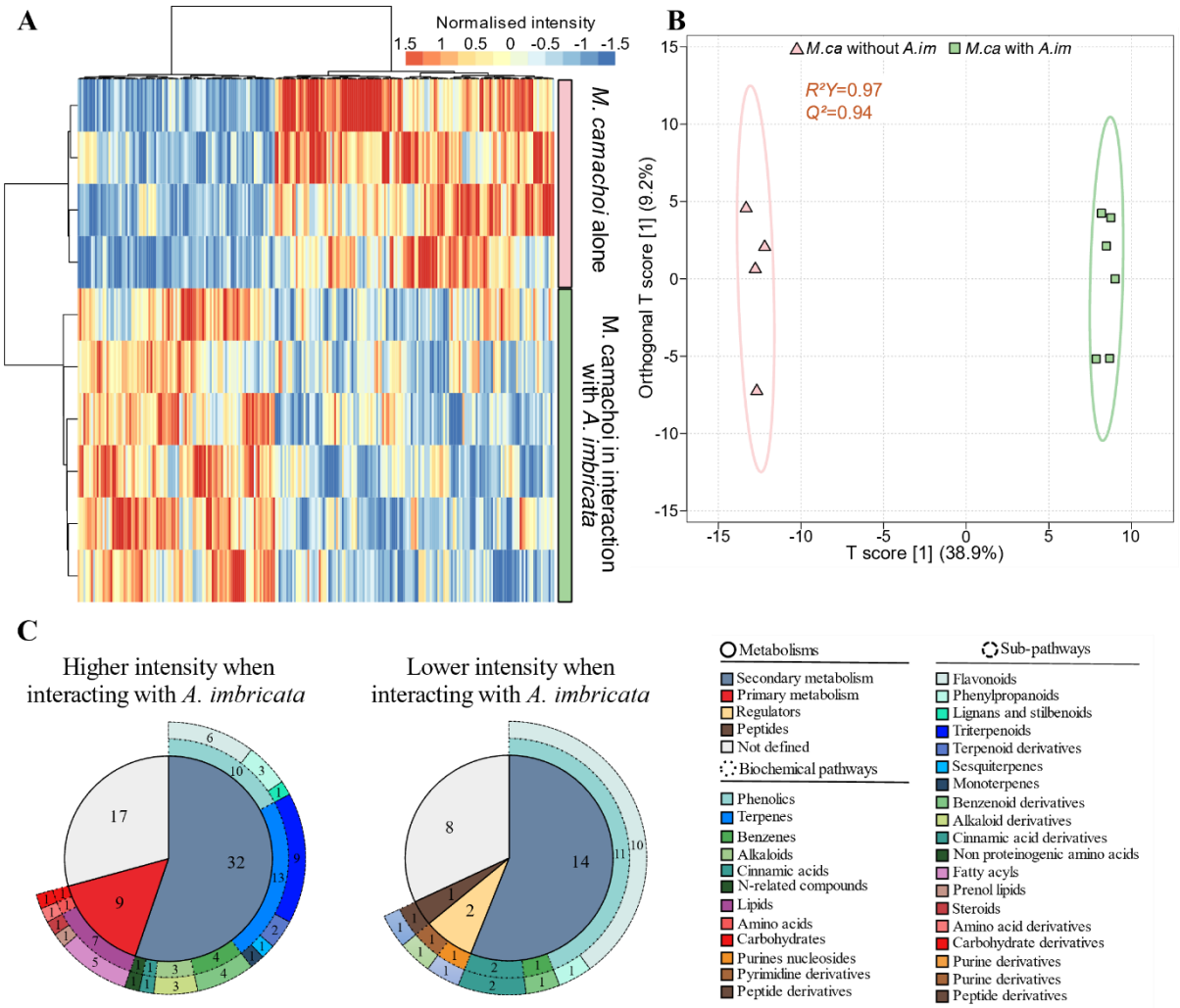

Figure V.S6 | *M. camachoi metabolome responds to the presence of A. imbricata*. A. Depiction of the 249 significant metabolic features from *M. camachoi* underlying the interaction with *A. imbricata* (Pearson correlation, Ward algorithm). **B.** Prediction of the *M. camachoi* interaction status using the 249 significant variables. **C.** Annotation of the significant markers

CHAPTER 6

GENERAL DISCUSSION. MAJOR METABOLIC FEATURES UNDER ATACAMA SPECIES ADAPTATION AND POTENTIAL FALLOUTS OF THE ANALYSIS

I. METABOLISM REORCHESTRATION: THE CONSEQUENCE AND CAUSE OF ATACAMA SPECIES ADAPTATION

In the Atacama Desert, intensities of abiotic stresses, which are a challenge for current agriculture, reach the extremes of the gradients of compatibility with plant life (Eshel et al., 2021). Adapted plant species embed genetic and metabolic innovations that provide sufficient resilience to perform under stressful conditions. In this framework, this PhD project aimed to explore the metabolic features of multiple Atacama species, which are of major interest for current agriculture. Although this objective could be addressed by distinct analytical strategies (Dussarrat et al., 2021), we selected a holistic, ecological approach that combined metabolomics, transcriptomics and machine learning techniques. This comprehensive approach preserved the ecological context of the study, but was also subjected to certain limitations. In this chapter, we first argued the main strengths and weaknesses of our approach. Subsequently, findings were integrated to discuss the metabolic strategies employed by Atacama plant species to cope with their environment. Finally, we conceived two potential perspectives that might provide complementary insights into our understanding of plant adaptation to extreme biomes.

I.1. An open question addressed by a holistic approach

A previous study suggested a significant place of metabolic processes in the adaptation of Atacama plant species (Eshel et al., 2021). In addition, a few studies described the biochemical diversity of Andean plants whilst others investigated specific genera, such as *Baccharis* and *Parastrephia* (two genera that were collected in the Atacama Desert) for pharmacological purposes (Padilla-González et al., 2017; Isla et al., 2021; Minteguiaga et al., 2021). Thus, the main limitations of this project were low availability of metabolic data on the selected Atacama species and the multi-stress environmental properties of this natural environment. Our works highlighted various methods to uncover adaptive markers of extreme plant species (Dussarrat et al., 2021). These approaches could be divided into two categories: the pairwise strategy, which aimed to compare pairs of extreme species with phylogenetically related species (mostly represented by crops or model plants), and a more global approach, which sought to explore general metabolic patterns of extreme plant species. In the context of Atacama species, the first approach would offer the possibility of analysing plant responses to specific stress gradients such as drought or salinity in natural or controlled conditions based on the opportunity to develop three species in greenhouses (*S. chilense*, *H. doellii* and *A. adscensionis*). Then, this strategy could be extended to some of the other 21 plant species collected. We preferred a more holistic approach, where various plant species were collected at different elevation levels and subjected to multi-platform metabolic analyses, for three reasons.

- First, previous results argued for the existence of convergent strategies to face harsh climates involved in protection against high light intensities, salt or nitrogen starvation, for instance (Eshel et al., 2021). Besides, strong relationships were observed between phytochemical diversity and environmental conditions, reinforcing the interest in using metabolomics to uncover protective biochemical traits against these environmental constraints (Defossez et al., 2021).

- Second, this holistic approach avoided potential acclimation processes that could occur under controlled conditions (Dussarrat et al., 2021).

- Third, this holistic strategy conserved biological diversity. Plant metabolome diversity in the plant kingdom was estimated between 200 000 and 1 million compounds, with up to 40 000 metabolites per species (Alseekh and Fernie, 2018). Although a myriad of compounds was classified as species-specific, a great proportion is shared between the majority of species in the plant kingdom (Hartmann, 1996; Alseekh and Fernie, 2018). For instance, primary metabolism and *basal* secondary metabolites are shared between plant species (Hartmann, 2007). Hence, it seems relevant to question to what extent these shared metabolites participate in adaptive processes, especially when considering these markers for a more generic breeding strategy. Our approach used biological diversity to highlight the most generic compounds as best predictors of the elevation levels based on (i) the need for shared compounds to predict the environment for all species and (ii) their high correlation with elevation.

Hence, a holistic eco-metabolomics approach appeared as a valuable strategy for reaching the project objectives and addressing the main analytical and environmental constraints. More excitingly, this approach is pioneering in several aspects. For instance, we provided an alternative strategy to study plant adaptation to harsh climates by showing that the application of eco-metabolomics to multiple species enabled the discovery of generic mechanisms predicting the distribution of plants at 79%. While these results will undoubtedly inspire future studies that will improve our understanding of adaptive mechanisms to extreme conditions, this approach is also likely to be increasingly used in agronomy and ecology (Chapter VI. Section II). However, such an approach was subjected to various limitations.

- First, a metabolic analysis integrating a broad range of species implied a large phenotypic and chemical variability that must be controlled. Significant efforts were deployed to characterise phenotypic specificities of each species (*e.g.* flowering, size, ratio between living and dead tissue). Similarly, such intra- and inter-species variability implied analytical compromises in chemical analyses. Analytical parameters (*e.g.* dilution, solvent, injection flow) were thus optimised via multiple time-consuming tests to (i) maximise the chemical coverage and (ii) ensure great comparability between species (Chapter 2).

- Besides, sufficient biodiversity was required to perform such an analysis and ensure a sufficient dynamic range of metabolic data. Hence, sampling was performed after the rainy season (*i.e.* March/April), where the maximum of species diversity and coverage was observed (Díaz et al., 2016,

2019). Consequently, our eco-metabolomics approach could not access the evolution of plant metabolome over a year. By extension, the protective role of the best metabolic predictors (*i.e.* 39 compounds predicting plant environment) across seasons could not be explored (Chapter 3). However, enrichment analyses confirmed their significance at the evolution level, thus rejecting the hypothesis of a transient role for these markers in resilience to extreme environmental constraints (Chapter 4).

- In addition, the efficiency of this strategy was based on the individualistic theory, according to which the distribution of the species is independent of other species (Callaway et al., 2002). We thus tested the influence of plant-plant interactions in the survival of Atacama plant species (Chapter 5).

- Finally, additional limitations were intrinsically linked to our analytical workflow. For instance, the freeze-drying process strongly restricted enzymatic, redox analyses and *in vivo* measurements such as photosynthesis activity for fresh materials.

Overall, while some of these limitations were overcome by complementary analyses, others represent blind spots of the project. Nevertheless, the benefits offered by this holistic approach led to significant discoveries that improved our understanding of the metabolic mechanisms underlying the adaptation of multiple plant species to extreme environments.

I.2. Metabolic strategies employed to face environmental constraints

This PhD project aimed to decipher the metabolic strategies employed by Atacama plant species to thrive in their extreme environment. For decades, the chemical properties enabling plant survival in harsh environments have attracted human curiosity (Turner, 2018). However, most studies focused on specific chemical classes known for their osmoprotective functions. In contrast, untargeted studies provided relevant information on the role of secondary compounds in adaptation, but have suffered from the limitations of the species-specific approach (Dussarrat et al., 2021). Hence, the first challenge of this project was to capture the biochemical diversity of these extreme Atacama species. Here, we deployed multi-platform metabolomics to provide a metabolic fingerprint of 24 Atacama species, which encompassed major primary compounds as well as fatty acyls and secondary metabolites. In addition, this metabolic matrix permitted to explore the impact of a multiple stress gradient (*e.g.* high solar irradiance, drought, salinity, freezing temperature) on the metabolism of these different species (19 sites between 2400 and 4500 m.a.s.l). Strikingly, our strategy enabled the discovery that (i) Atacama plant metabolome was an excellent integrator of environmental parameters, (ii) convergent biochemical strategies were conserved through evolution and ensured a predictable resilience against major abiotic constraints (Chapters 3 and 4), and (iii) plant community structure and interactions greatly influenced

plant distribution and coverage (Chapter 5). Importantly, these convergent metabolic strategies accurately reflected the major abiotic constraints (Fig. VI.1).

Regulation of energy entry. One of the few non-limiting resources in the Atacama Desert is light (Eshel et al., 2021). Instead, the high photon density was three times higher than other deserts and high mountains (Zhang et al., 2010; Arancibia-Bulnes et al., 2014). To face high solar irradiance, all Atacama species harboured low chlorophyll content (Fig. III.S1), in accordance with other analyses (Balaguer et al., 2002; Körner, 2003). This result suggested that plants mitigated solar irradiance damages by limiting energy entry. Unfortunately, photosynthetic and respiration rates could not be verified. It is generally admitted that the desynchronisation between photon entry and the photosynthetic electron transport chain disturbs redox homeostasis (Fernández-Marín et al., 2020). Studies revealed that high solar irradiance, as well as water or nutrient deficiency, reduced photosynthetic activity and increased ROS production (Decros et al., 2019; Dussarrat et al., 2021). The defect or excess of these environmental variables unbalanced the redox state, thereby resulting in a global reprogramming of plant metabolome.

Redefining priorities and performances based on carbon and nitrogen resources. While the objective of any plant in physiological conditions is to develop and reproduce, the primary objective of Atacama species is survival. In physiological or stressful conditions, carbon and nitrogen cycles (from uptake to transport and use) are extremely regulated at the molecular levels (Coruzzi and Zhou, 2001). This was even more evident in extreme plants that carefully managed carbon and nitrogen allocation between plant development and defence (Fig. IV.3). Nitrogen availability is one of the most vital limitations for plant life in the Atacama Desert (Díaz et al., 2016; Eshel et al., 2021). This was exemplified by critical nitrate levels compared to several agricultural and ornamental species (Fig. III.S1). Interestingly, Atacama species developed ingenious strategies to cope with nitrogen starvation. First, a previous transcriptomics analysis comparing 32 Atacama species and 32 phylogenetically related species revealed a strong interaction with growth-promoting bacteria, including nitrogen fixers (Eshel et al., 2021). Complementarily, more than 50% of these species naturally enriched reactions involved in the uptake and process of nitrogen resources (*e.g.* glutamate and glutamine related reactions), as well as nitrogen remobilisation (*e.g.* purine degradation) at the genomic level (Fig. IV.3). Although several aspects remained untackled, such as the Michaelis constant of nitrate transporters, these findings pinpointed the significance of adaptive processes to cope with inadequate nutrient conditions. Concomitantly, all Atacama plants harboured shallow protein levels (Fig. III.S1), and enrichment analyses highlighted a significant regulation of gibberellins and cytokinins pathways (Fig. IV.3). Hence, although phenotypic traits such as biomass and relative growth rate were not available, these observations indicated a strong restriction of plant growth processes. This idea was also supported by the position of starch as the best predictor of the plant environment (Fig. III.3), an adaptive role further confirmed by enrichment analyses (Fig. IV.3). Considering plant metabolome as a constant chemical

flux, carbon that did not fuel growth processes would be stored as osmotically inactive carbohydrates such as starch, especially under drought and saline conditions (Rontein et al., 2002).

In stark contrast, Atacama species invested remarkable efforts and resources in the synthesis of protective compounds. Importantly, the defence strategy of these species was precisely organised and tailored to (i) the common pressures of the different ecosystems in the transect and (ii) the vegetation belt-specific constraints. In addition to the nitrogen-related adaptations, protective strategies could be divided into three parts covering defence against high solar irradiance, low temperatures and osmotic constraints (including high salinity and drought).

Defence against high light intensity. Plants are sessile organisms that can not escape climatic parameters. The Atacama Desert does not offer shaded micro-environments that could protect annual and perennial plants from the constant solar irradiance pressure. High radiations induce dramatic effects on photosynthetic machinery (Wimalasekera, 2019). As a direct consequence, Atacama species developed adaptive mechanisms to face this abiotic constraint. Massive changes in lipidic profiles were likely to occur across Atacama species when compared to their 32 phylogenetically related species. For instance, extreme species tended to enhance the synthesis of long-chain fatty acids and waxes (Fig. IV.3), which demonstrated effective ultraviolet reflectance properties and additional protective functions against high light intensity (Holmes and Keiller, 2002; Santos et al., 2017). Besides, precursors of lipids (*e.g.* jasmonates, azelaic acid) figured among the best predictors of plant environment (Fig. III.3). Interestingly, the higher number of enriched reactions related to fatty acyl metabolism in Prepuna species coincided with the surprisingly highest solar irradiance at these levels, which was confirmed over the years (Eshel et al., 2021, Dussarrat et al., 2022). Complementarily, the adaptive role of carotenoids and flavonoids was pinpointed in gene expression enrichment analysis as well as in eco-metabolomics approaches carried out at the species-specific (*i.e. A. imbricata*) or global scales (Agati et al., 2011; Ramel et al., 2012). Alternatively, the excess of solar irradiance could be dissipated as heat through alternative oxidases and uncoupling proteins (Grant et al., 2009). Regrettably, mitochondrial photoprotective mechanisms (*e.g.* alternative oxidases that deviate the electron flow from mitochondrial complexes) have not been addressed.

Defence against freezing temperatures. Daily sub-zero temperatures challenge plant life in the high Andean Steppe, explaining the absence of C4 plants in these sites (Eshel et al., 2021, Dussarrat et al., 2022). Subfreezing air temperature can lead to extrinsic or intrinsic ice nucleation, cellular dehydration, membrane damages and plant death (Pearce, 2001; Neuner and Hacker, 2012). Primary metabolism emerged as a significant resource of protective compounds against freezing stress. Various sugars and polyols were significantly enriched in Steppe species (Fig. IV.3). Interestingly, previous studies showed their accumulation under cold acclimation to improve plant hardiness in a frost environment (Román-Figueroa et al., 2021). Notably, an intriguing shift in protection strategy occurred

in Steppe species, which favoured the production of complex sugars like raffinose (a compound observed among the 39 best predictors) over starch. The lower efficiency of starch remobilisation could explain this phenomenon under cold temperatures and by the relevant functions of raffinose family oligosaccharides to mitigate freezing effects (Peters and Keller, 2009). In addition, lipid remodelling (*e.g.* prevalence of polyunsaturated fatty acyls) observed in Atacama plants and other extreme species may contribute to cold stress tolerance by protecting membrane stability, for instance (Barrero-Sicilia et al., 2017). However, temperature effects were not restricted to high elevations. In contrast, the diversity and coverage of Prepuna species in Puna (*i.e.* between 3300 and 4000 m.a.s.l) were constrained by cold temperatures (Fig. V.2). Interestingly, the survival and development of these species like *H. doellii* and *A. imbricata* (*i.e.* annual or perrenial plant) were dependent on the positive interaction with *M. camachoi*, suggesting that their metabolome was not sufficiently adapted to cope with freezing temperatures. This facilitation process was responsible for a decrease in synthesis of protective compounds by acting as a thermal and hydric refuge (Fig. V.2 and V.6).

Defence against osmotic pressures. The Atacama Desert is the driest non-polar environment on earth (Eshel et al., 2021). While water limits plant life across the entire transect, other abiotic parameters negatively impact plant performances, such as salinity. Drought causes irreversible damage in plant cells and disturbs redox homeostasis, a scenario exacerbated by high salinity in Prepuna (Rizhsky et al., 2002; Parihar et al., 2015). Interestingly, the first observable consequence is the prevalence of C4 plants, as opposed to Steppe environments. The evolution of primary metabolism converged towards the synthesis of various sugars such as trehalose in Prepuna species (Fig. IV.5). The selective advantage of trehalose synthesis under osmotic pressure was confirmed by its predictive ability and its accumulation at the lower elevation levels (Fig. III.3), and by previous results (Bhattacharya and Kundu, 2020). Furthermore, a broad range of polyols accumulated in Atacama species, although some, like inositol, were more specific to Prepuna (Fig. IV.3 and IV.4). Concomitantly, the expression of genes related to lipid metabolism (*e.g.* pollen wall, phospholipids) was higher in the majority of low elevation species (Fig. IV.3 and IV.5), reinforcing their function in osmotic stress resilience (Barrero-Sicilia et al., 2017). Findings showed a significant place of various amino acids in adaptation of Steppe and Prepuna species. For instance, proline metabolism (*e.g.* proline, proline betaine) was systematically reported in the different analyses (Fig. III.3 and IV.5), consistent with its influence on redox homeostasis (Szabados and Savouré, 2010; Dussarrat et al., 2021). More generally, the accumulation of other amino acids could refer to their extensive relationship with secondary metabolism (Fig. I.6 and IV.3). Intriguing accumulation of phenolic-polyamine conjugates (*e.g.* tri-coumaroyl spermidine) was found in Prepuna species, congruent with resurrection species (Alcázar et al., 2011). Furthermore, phenolic pathways were extensively represented among the best markers of plant environment and generic enriched reactions in all Atacama plant species (Fig. III.3 and IV.3). Besides, their adaptive role was observed in other extreme species (Dussarrat et al., 2021). Importantly, their widely described antioxidant properties

contribute to photoprotection and mitigate the redox poise by scavenging ROS excess, thus improving plant resilience against osmotic stress (Agati and Tattini, 2010; Decros et al., 2019). Interestingly, quercetin relatives were the most widely represented compound in this study, a central adaptive place possibly explained by (i) its roles in antioxidant machinery, (ii) its interaction with soil chemistry and nitrogen fixers, and (iii) its link with growth and developmental processes (Singh et al., 2021). Finally, positive interactions between Steppe species and *M. camachoi* seemed to facilitate the development of these species in Puna, emphasising their protective role against osmotic stress (Fig. V.2). This function was also supported by previous analyses of facilitation processes in arid lands (Flores and Jurado, 2003).

In conclusion, while the diversity of the Atacama plant community encompassed multiple botanic families characterised by various evolutionary trajectories, the adaptation of these extreme species was underlined by convergent metabolic strategies fixed over evolution (Fig. VI.1). Here, we first provided an unprecedented metabolic resource uncovering the biochemical fingerprint of multiple extremophile plant species. Then, we confirmed the excellent integrative capacity of plant metabolome, which captured the most limiting abiotic parameters at both species and ecological levels. Next, our eco-metabolomics approach was combined with machine learning to unveil a generic toolbox allowing plant resilience to extreme climates. To our knowledge, this discovery represented the first evidence of the prevalence of convergent chemical strategies in adaptation to harsh lands. This finding was then validated at the evolutionary scale through the analysis of gene expansion and expression patterns. Importantly, while enabling survival of various plant lineages, these shared metabolic processes mainly involved secondary compounds, thus challenging the species-specificity of adaptive strategies previously expected. In addition, the annotation of these generic mechanisms has unlocked some of the secrets of the adaptation of Atacama species. Plant metabolome is undoubtedly tailored to environmental constraints. Atacama plants favoured the regulation of pre-existing ubiquitous pathways (*e.g.* quercetin and carotenoids chemicals) with various protective functions against high solar irradiance and drought, for instance. Furthermore, specific adaptations emerged to face freezing temperature (*e.g.* raffinose) or osmotic stress (*e.g.* proline, tri-coumaroyl spermidine) in Steppe and Prepuna species, respectively. Remarkably, a relevant part of adaptive compounds referred directly or indirectly (*i.e.* through interaction with redox pathways) to redox homeostasis, emphasising its central role in adaptation processes. Finally, this intricate metabolic reordering was complemented by positive biotic interactions, which enabled plant development beyond the limits of their life compatible gradients by providing protected micro-environments against temperature and osmotic pressures (Fig. VI.1). Overall, this study provided tremendous insights into our understanding of the adaptive responses of plants to extreme abiotic constraints that challenge current agriculture. Besides, our approach and findings offered promising opportunities and perspectives for further investigations to enhance our comprehension of plant-environment interactions.

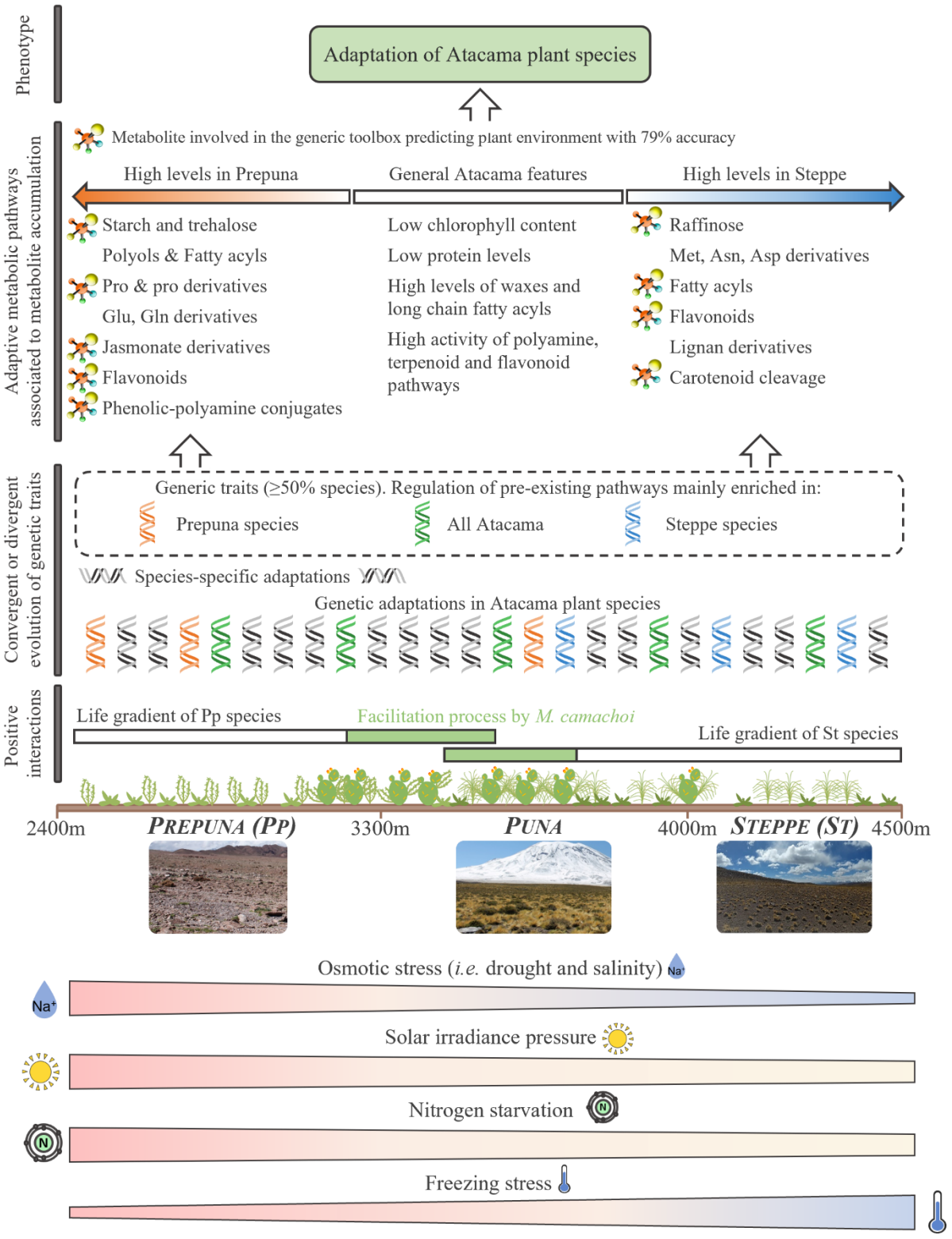

Figure VI.1 |

Evolution of adaptive traits in Atacama plant species: from the ecosystem to the metabolism.

Asn: Asparagine, *Asp*: Aspartate, *Gln*: Glutamine, *Glu*: Glutamate, *Met*: Methionine, *Pp*: Prepuna, *Pro*: Proline, *St*: Steppe.

II. GENERAL PERSPECTIVES: HOW TO GO FURTHER?

Atacama plants integrated molecular answers developed over tens of thousands of years of evolution. This PhD project unlocked several secrets hidden by this unique biodiversity. In addition, the multi-species eco-metabolomic approach offers promising perspectives for further research in plant sciences.

II.1. The mysteries of the Atacama: the quest for secrets goes on

The analyses performed throughout this project yielded a large number of perspectives already mentioned in the different chapters. For instance, complementary works should be conducted to (i) investigate the function of the chemical pathways and metabolites unveiled by both enrichment and metabolomics studies (Chapters 3 and 4) under environmental variations (*e.g.* testing their response to water or salt inputs), (ii) assess the photosynthetic and respiration rates, (ii) extend our comprehension of major primary compound responses to environmental variations in extreme lands (*e.g.* NMR analysis), and (iii) explore convergent metabolic mechanisms underlying plant-plant interaction processes (Chapter 5). Here, we detailed three complementary analyses that would help overcome the blind spots of our analysis.

Plant metabolic dynamics in the Atacama Desert. As mentioned above, the kinetic aspect of the metabolism of these plants is absent. Besides, although the adaptive role of the discovered metabolic traits was supported by two independent analyses, their behaviour through time remains unexplored. This notion of time could be addressed at two distinct scales. On short time scales (*i.e.* hours, days), studying the response of chemical fluxes to water or nitrogen inputs through enzymatic assays would provide significant insights into how Atacama plants process environmental resources. On a larger time scale (*i.e.* months), the trade-off between the synthesis of protective compounds and growth, and thus protein, could be examined. Intriguingly, all Atacama plants (including annuals) showed low protein contents compared to agronomic and ornamental plants (Dussarrat et al., 2022). Hence, implementing the behaviour of the best metabolic predictors to complementary data such as ecophysiological features (*e.g.* photosynthesis rates) and phenotypic traits (*e.g.* growth rate) over the year would provide precious information about how Atacama species manage carbon and nitrogen allocation to thrive under extreme conditions. For instance, comparing metabolic, proteomic and phenotypic profiles of plants collected in different seasons would enhance our comprehension of plant physiology under extreme conditions. Even more thrilling, such a study would help disentangle major roads from environment perturbation to metabolic integration and phenotypic consequences (Poorter et al., 2019; Walker et al., 2022). Lastly, a growing interest in the analysis of redox signalling resulted in a better understanding of its various roles in plant physiology and defence (Foyer and Noctor, 2016; Decros et al., 2019). Thus, understanding the

contribution of redox mechanisms in these trade-offs represents one of the most critical challenges for the future.

Plant-soil interactions under extreme abiotic conditions. Plant-soil feedback (*i.e.* reciprocal interaction between plant and soil) showed high potential in modulating plant community structure (Ke et al., 2015, 2021), and was strongly influenced by abiotic factors such as drought (Kaisermann et al., 2017). Remarkably, profound changes in microbial communities arose between the different vegetation belts (*i.e.* Prepuna, Puna, Steppe) in the Atacama Desert (Eshel et al., 2021). Besides, this same study highlighted the highest bacterial abundancy (≥2-fold) in plant root zones compared to bare soil patches, including growth-promoting and nitrogen-fixing microorganisms. These findings were not isolated since intriguing biodiversity of plant-microbe interactions was reported in other extreme ecosystems (Bang et al., 2018). Hence, it is highly probable that Atacama plant adaptation depends on plant-soil interactions and that these interactions may differ between species and vegetation belts. While the quest for convergent mechanisms underlying plant-microbe communications is still valuable between species from a given vegetation belt, a peculiar interest in the various strategies deployed by *Poaceae*, *Fabaceae* and *Asteraceae* (the three prominent plant families in the Atacama Desert) would be of major interest. Hence, a joint transcriptomics and metabolomics analysis applied to paired samples (*i.e.* aerial parts of plants with related plant root zones) from various plant species and elevation levels would represent an unprecedented analysis that aims to decipher the benefits of plant-soil interactions in extreme ecosystems and expand our understanding of the molecular mechanisms underlying these vital processes.

Behaviour of the best metabolic markers under various conditions. Although less extreme than the Atacama Desert, Californian drylands represent other types of harsh environments where plants face multiple stress gradients (Jacobsen et al., 2007). Interestingly, several plant genera collected in the Atacama Desert are observed in Californian drylands (*e.g. Atriplex*, *Ambrosia*, *Lupinus*). A tri-partite analysis would allow comparing the metabolomic profiles of an Atacama species, a Californian related species and a related agronomic species. This tri-partite system could be complemented with Atacama plant species developed under controlled conditions (*e.g. H. doellii*, *S. chilense*). Such a comparative study would address the following questions: “how do the best predictive markers of the plant environment respond to various natural and controlled environmental variations (*e.g.* in other ecosystems or under artificial drought gradient)?” and “do these metabolic markers provide a selective advantage in other challenging ecosystems?”.

II.2. Eco-metabolomics: from ecosystem to agronomic level

Climate warming alters plant functional traits and thus stability properties of ecological communities (Ma et al., 2017; Bjorkman et al., 2018). Biodiversity has the fascinating potential to mitigate the effects of increased frequency and intensity of abiotic events (Chapin III et al., 2000; Hisano et al., 2018). Moreover, species interactions are considered a major player in shaping climate change impact (Åkesson et al., 2021). However, ecological models are sorely lacking in a more mechanistic and molecular understanding of the ecological mechanisms that control the balance and response of plant communities.

In stark contrast, the current agricultural system is based on monospecific cultures of the best ideotypes, whose yields are directly dependent on the resilience of this species to climatic variations (Chai et al., 2021; Chaloner et al., 2021; Lesk et al., 2021). In consequence, the performances of these plants are stagnating and even decreasing over the years (Long et al., 2015), showing the difficulty agronomists have in finding sustainable solutions in the current breeding system. In this context, chemical and biological diversity might be helpful by delivering naturally selected strategies deployed by plants to cope with their environment. Altogether, current challenges in both ecology and biology illustrate the need to improve our ability to predict plant performances (*i.e.* resilience and yields) under a wide range of environmental conditions and perturbations.

We believe our predictive multi-species metabolomics approach represents an innovative and promising analytical strategy that could help meet these ambitious objectives through its ability to uncover highly generic metabolic predictors while preserving the ecological context. While providing clues for a better understanding of ecological models, such an approach could be used to unveil easily measurable soft traits predicting various phenotypic traits or biotic interaction patterns. In parallel, the multi-species approach would certainly serve the breeder's interests through its capacity to unveil generic strategies for various purposes. Finally, the already strong potential interest in the use of eco-metabolomics in agriculture and ecology is supported by high predictive performance, low costs and high throughput properties, thus offering a promising future to this integrative approach.

REFERENCES

Afendi, F. M., Okada, T., Yamazaki, M., Hirai-Morita, A., Nakamura, Y., Nakamura, K., et al. (2012). KNApSAcK family databases: integrated metabolite-plant species databases for multifaceted plant research. *Plant Cell Physiol.* 53, e1. doi:10.1093/pcp/pcr165.

Agati, G., Azzarello, E., Pollastri, S., and Tattini, M. (2012). Flavonoids as antioxidants in plants: Location and functional significance. *Plant Science* 196, 67–76. doi:10.1016/j.plantsci.2012.07.014.

Agati, G., Biricolti, S., Guidi, L., Ferrini, F., Fini, A., and Tattini, M. (2011). The biosynthesis of flavonoids is enhanced similarly by UV radiation and root zone salinity in L. vulgare leaves. *Journal of Plant Physiology* 168, 204–212. doi:10.1016/j.jplph.2010.07.016.

Agati, G., and Tattini, M. (2010). Multiple functional roles of flavonoids in photoprotection: Letters. *New Phytologist* 186, 786–793. doi:10.1111/j.1469-8137.2010.03269.x.

Åkesson, A., Curtsdotter, A., Eklöf, A., Ebenman, B., Norberg, J., and Barabás, G. (2021). The importance of species interactions in eco-evolutionary community dynamics under climate change. *Nat Commun* 12, 4759. doi:10.1038/s41467-021-24977-x.

Akula, R., and Ravishankar, G. A. (2011). Influence of abiotic stress signals on secondary metabolites in plants. *Plant Signaling & Behavior* 6, 1720–1731. doi:10.4161/psb.6.11.17613.

Alcázar, R., Altabella, T., Marco, F., Bortolotti, C., Reymond, M., Koncz, C., et al. (2010). Polyamines: molecules with regulatory functions in plant abiotic stress tolerance. *Planta* 231, 1237–1249. doi:10.1007/s00425-010-1130-0.

Alcázar, R., Bitrián, M., Bartels, D., Koncz, C., Altabella, T., and Tiburcio, A. F. (2011). Polyamine metabolic canalization in response to drought stress in Arabidopsis and the resurrection plant *Craterostigma plantagineum*. *Plant Signaling & Behavior* 6, 243–250. doi:10.4161/psb.6.2.14317.

Allard, P.-M., Genta-Jouve, G., and Wolfender, J.-L. (2017). Deep metabolome annotation in natural products research: towards a virtuous cycle in metabolite identification. *Current Opinion in Chemical Biology* 36, 40–49. doi:10.1016/j.cbpa.2016.12.022.

Alseekh, S., and Fernie, A. R. (2018). Metabolomics 20 years on: what have we learned and what hurdles remain? *The Plant Journal* 94, 933–942. doi:10.1111/tpj.13950.

Altelaar, A. F. M., Munoz, J., and Heck, A. J. R. (2013). Next-generation proteomics: towards an integrative view of proteome dynamics. *Nat Rev Genet* 14, 35–48. doi:10.1038/nrg3356.

Annunziata, M. G., Ciarmiello, L. F., Woodrow, P., Dell'Aversana, E., and Carillo, P. (2019). Spatial and Temporal Profile of Glycine Betaine Accumulation in Plants Under Abiotic Stresses. *Front. Plant Sci.* 10, 230. doi:10.3389/fpls.2019.00230.

Antonsson, H., Björk, R. G., and Molau, U. (2009). Nurse plant effect of the cushion plant *Silene acaulis* (L.) Jacq. in an alpine environment in the subarctic Scandes, Sweden. *Plant Ecology & Diversity* 2, 17–25. doi:10.1080/17550870902926504.

Arancibia-Bulnes, C. A., Peón-Anaya, R., Riveros-Rosas, D., Quiñones, J. J., Cabanillas, R. E., and Estrada, C. A. (2014). Beam Solar Irradiation Assessment for Sonora, Mexico. *Energy Procedia* 49, 2290–2296. doi:10.1016/j.egypro.2014.03.242.

Araújo, W. L., Nunes-Nesi, A., Nikoloski, Z., Sweetlove, L. J., and Fernie, A. R. (2012). Metabolic control and regulation of the tricarboxylic acid cycle in photosynthetic and heterotrophic plant tissues: TCA control and regulation in plant tissues. *Plant, Cell & Environment* 35, 1–21. doi:10.1111/j.1365-3040.2011.02332.x.

Arimura, G., Ozawa, R., Shimoda, T., Nishioka, T., Boland, W., and Takabayashi, J. (2000). Herbivory-induced volatiles elicit defence genes in lima bean leaves. *Nature* 406, 512–515. doi:10.1038/35020072.

Arnon, D. I. (1949). COPPER ENZYMES IN ISOLATED CHLOROPLASTS. POLYPHENOLOXIDASE IN *BETA VULGARIS*. *Plant Physiol.* 24, 1–15. doi:10.1104/pp.24.1.1.

Ashraf, M., and Foolad, M. R. (2007). Roles of glycine betaine and proline in improving plant abiotic stress resistance. *Environmental and Experimental Botany* 59, 206–216. doi:10.1016/j.envexpbot.2005.12.006.

Aubin-Horth, N., and Renn, S. C. P. (2009). Genomic reaction norms: using integrative biology to understand molecular mechanisms of phenotypic plasticity. *Molecular Ecology* 18, 3763–3780. doi:10.1111/j.1365-294X.2009.04313.x.

Austin, M. P. (1985). Continuum Concept, Ordination Methods, and Niche Theory. *Annu. Rev. Ecol. Syst.* 16, 39–61. doi:10.1146/annurev.es.16.110185.000351.

Azevedo, R. A., Lancien, M., and Lea, P. J. (2006). The aspartic acid metabolic pathway, an exciting and essential pathway in plants. *Amino Acids* 30, 143–162. doi:10.1007/s00726-005-0245-2.

Bachrach, U. (2010). The early history of polyamine research. *Plant Physiology and Biochemistry* 48, 490–495. doi:10.1016/j.plaphy.2010.02.003.

Báez, S., and Collins, S. L. (2008). Shrub Invasion Decreases Diversity and Alters Community Stability in Northern Chihuahuan Desert Plant Communities. *PLoS ONE* 3, e2332. doi:10.1371/journal.pone.0002332.

Bailey-Serres, J., Parker, J. E., Ainsworth, E. A., Oldroyd, G. E. D., and Schroeder, J. I. (2019). Genetic strategies for improving crop yields. *Nature* 575, 109–118. doi:10.1038/s41586-019-1679-0.

Balaguer, L., Pugnaire, F. I., Martínez-Ferri, E., Armas, C., Valladares, F., and Manrique, E. (2002). Ecophysiological significance of chlorophyll loss and reduced photochemical efficiency under extreme aridity in Stipa tenacissima L. *Plant and Soil* 240, 343–352. doi:10.1023/A:1015745118689.

Baldwin, I. T., Halitschke, R., Paschold, A., von Dahl, C. C., and Preston, C. A. (2006). Volatile Signaling in Plant-Plant Interactions: "Talking Trees" in the Genomics Era. *Science* 311, 812–815. doi:10.1126/science.1118446.

Bang, C., Dagan, T., Deines, P., Dubilier, N., Duschl, W. J., Fraune, S., et al. (2018). Metaorganisms in extreme environments: do microbes play a role in organismal adaptation? *Zoology* 127, 1–19. doi:10.1016/j.zool.2018.02.004.

Bantan-Polak, T., Kassai, M., and Grant, K. B. (2001). A Comparison of Fluorescamine and Naphthalene-2,3-dicarboxaldehyde Fluorogenic Reagents for Microplate-Based Detection of Amino Acids. *Analytical Biochemistry* 297, 128–136. doi:10.1006/abio.2001.5338.

Barrero-Sicilia, C., Silvestre, S., Haslam, R. P., and Michaelson, L. V. (2017). Lipid remodelling: Unravelling the response to cold stress in Arabidopsis and its extremophile relative Eutrema salsugineum. *Plant Science* 263, 194–200. doi:10.1016/j.plantsci.2017.07.017.

Beadle, G. W., and Tatum, E. L. (1941). Genetic Control of Biochemical Reactions in Neurospora. *Proceedings of the National Academy of Sciences* 27, 499–506. doi:10.1073/pnas.27.11.499.

Beauvoit, B., Belouah, I., Bertin, N., Cakpo, C. B., Colombié, S., Dai, Z., et al. (2018). Putting primary metabolism into perspective to obtain better fruits. *Annals of Botany* 122, 1–21. doi:10.1093/aob/mcy057.

Belouah, I., Nazaret, C., Pétriacq, P., Prigent, S., Bénard, C., Mengin, V., et al. (2019). Modeling Protein Destiny in Developing Fruit. *Plant Physiol.* 180, 1709–1724. doi:10.1104/pp.19.00086.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57, 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x.

Benton, M. J., Wilf, P., and Sauquet, H. (2021). The Angiosperm Terrestrial Revolution and the origins of modern biodiversity. *New Phytologist*, nph.17822. doi:10.1111/nph.17822.

Berli, F. J., Moreno, D., Piccoli, P., Hespanhol-Viana, L., Silva, M. F., Bressan-Smith, R., et al. (2009). Abscisic acid is involved in the response of grape (*Vitis vinifera* L.) cv. Malbec leaf tissues to ultraviolet-B radiation by enhancing ultraviolet-absorbing compounds, antioxidant enzymes and membrane sterols. *Plant, Cell & Environment*. doi:10.1111/j.1365-3040.2009.02044.x.

Berlow, E. L. (1999). Strong effects of weak interactions in ecological communities. *Nature* 398, 330–334. doi:10.1038/18672.

Bhandari, K., and Nayyar, H. (2014). "Low Temperature Stress in Plants: An Overview of Roles of Cryoprotectants in Defense," in *Physiological Mechanisms and Adaptation Strategies in Plants Under Changing Environment*, eds. P. Ahmad and M. R. Wani (New York, NY: Springer New York), 193–265. doi:10.1007/978-1-4614-8591-9_9.

Bhattacharya, S., and Kundu, A. (2020). "Sugars and Sugar Polyols in Overcoming Environmental Stresses," in *Protective Chemical Agents in the Amelioration of Plant Abiotic Stress*, eds. A. Roychoudhury and D. K. Tripathi (Wiley), 71–101. doi:10.1002/9781119552154.ch4.

Bieza, K., and Lois, R. (2001). An Arabidopsis Mutant Tolerant to Lethal Ultraviolet-B Levels Shows Constitutively Elevated Accumulation of Flavonoids and Other Phenolics. *Plant Physiol.* 126, 1105–1115. doi:10.1104/pp.126.3.1105.

Binder, S. (2010). Branched-Chain Amino Acid Metabolism in *Arabidopsis thaliana*. *The Arabidopsis Book* 8, e0137. doi:10.1199/tab.0137.

Bjorkman, A. D., Myers-Smith, I. H., Elmendorf, S. C., Normand, S., Rüger, N., Beck, P. S. A., et al. (2018). Plant functional trait change across a warming tundra biome. *Nature* 562, 57–62. doi:10.1038/s41586-018-0563-7.

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin Biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–546. doi:10.1146/annurev.arplant.54.031902.134938.

Bolger, A., Scossa, F., Bolger, M. E., Lanz, C., Maumus, F., Tohge, T., et al. (2014). The genome of the stress-tolerant wild tomato species Solanum pennellii. *Nat Genet* 46, 1034–1038. doi:10.1038/ng.3046.

Bordbar, A., Monk, J. M., King, Z. A., and Palsson, B. O. (2014). Constraint-based models predict metabolic and associated cellular functions. *Nat Rev Genet* 15, 107–120. doi:10.1038/nrg3643.

Borowitzka, M. A. (2018). The 'stress' concept in microalgal biology—homeostasis, acclimation and adaptation. *J Appl Phycol* 30, 2815–2825. doi:10.1007/s10811-018-1399-0.

Boyes, D. C., Zayed, A. M., Ascenzi, R., McCaskill, A. J., Hoffman, N. E., Davis, K. R., et al. (2001). Growth Stage–Based Phenotypic Analysis of Arabidopsis: A Model for High Throughput Functional Genomics in Plants. *Plant Cell* 13, 1499–1510. doi:10.1105/TPC.010011.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 72, 248–254. doi:10.1016/0003-2697(76)90527-3.

Bradshaw, A. D. (1965). "Evolutionary Significance of Phenotypic Plasticity in Plants," in *Advances in Genetics* (Elsevier), 115–155. doi:10.1016/S0065-2660(08)60048-6.

Breitkreuz, K. E., Allan, W. L., Van Cauwenberghe, O. R., Jakobs, C., Talibi, D., André, B., et al. (2003). A Novel γ-Hydroxybutyrate Dehydrogenase. *Journal of Biological Chemistry* 278, 41552–41556. doi:10.1074/jbc.M305717200.

Brenner, S., Jacob, F., and Meselson, M. (1961). An Unstable Intermediate Carrying Information from Genes to Ribosomes for Protein Synthesis. *Nature* 190, 576–581. doi:10.1038/190576a0.

Brooker, R. W. (2006). Plant–plant interactions and environmental change. *New Phytologist* 171, 271–284. doi:10.1111/j.1469-8137.2006.01752.x.

Brooker, R. W., Maestre, F. T., Callaway, R. M., Lortie, C. L., Cavieres, L. A., Kunstler, G., et al. (2007). Facilitation in plant communities: the past, the present, and the future. *J Ecology* 0, 070908024102002-??? doi:10.1111/j.1365-2745.2007.01295.x.

Bruno, J. F., Stachowicz, J. J., and Bertness, M. D. (2003). Inclusion of facilitation into ecological theory. *Trends in Ecology & Evolution* 18, 119–125. doi:10.1016/S0169-5347(02)00045-9.

Bunea, F., She, Y., Ombao, H., Gongvatana, A., Devlin, K., and Cohen, R. (2011). Penalized least squares regression methods and applications to neuroimaging. *NeuroImage* 55, 1519–1527. doi:10.1016/j.neuroimage.2010.12.028.

Burt, A. J., Arnason, J. T., and García-Lara, S. (2019). Natural variation of hydroxycinnamic acid amides in maize landraces. *Journal of Cereal Science* 88, 145–149. doi:10.1016/j.jcs.2019.06.002.

Caetano-Anollés, G., Yafremava, L. S., Gee, H., Caetano-Anollés, D., Kim, H. S., and Mittenthal, J. E. (2009). The origin and evolution of modern metabolism. *The International Journal of Biochemistry & Cell Biology* 41, 285–297. doi:10.1016/j.biocel.2008.08.022.

Callaway, R. M. (1998). Are Positive Interactions Species-Specific? *Oikos* 82, 202–207. doi:10.2307/3546931.

Callaway, R. M., Brooker, R. W., Choler, P., Kikvidze, Z., Lortie, C. J., Michalet, R., et al. (2002). Positive interactions among alpine plants increase with stress. *Nature* 417, 844–848. doi:10.1038/nature00812.

Caretto, S., Linsalata, V., Colella, G., Mita, G., and Lattanzio, V. (2015a). Carbon Fluxes between Primary Metabolism and Phenolic Pathway in Plant Tissues under Stress. *IJMS* 16, 26378–26394. doi:10.3390/ijms161125967.

Caretto, S., Linsalata, V., Colella, G., Mita, G., and Lattanzio, V. (2015b). Carbon Fluxes between Primary Metabolism and Phenolic Pathway in Plant Tissues under Stress. *IJMS* 16, 26378–26394. doi:10.3390/ijms161125967.

Carpenter, C. (2005). The environmental control of plant species density on a Himalayan elevation gradient: Plant species density on a Himalayan elevation gradient. *Journal of Biogeography* 32, 999–1018. doi:10.1111/j.1365-2699.2005.01249.x.

Carrasco-Puga, G., Díaz, F. P., Soto, D. C., Hernández-Castro, C., Contreras-López, O., Maldonado, A., et al. (2021). Revealing hidden plant diversity in arid environments. *Ecography* 44, 98–111. doi:10.1111/ecog.05100.

Caspi, R., Billington, R., Keseler, I. M., Kothari, A., Krummenacker, M., Midford, P. E., et al. (2020). The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Research* 48, D445–D453. doi:10.1093/nar/gkz862.

Castañeda-Álvarez, N. P., Khoury, C. K., Achicanoy, H. A., Bernau, V., Dempewolf, H., Eastwood, R. J., et al. (2016). Global conservation priorities for crop wild relatives. *Nature Plants* 2, 16022. doi:10.1038/nplants.2016.22.

Cavieres, L. A., Badano, E. I., Sierra-Almeida, A., Gómez-González, S., and Molina-Montenegro, M. A. (2006). Positive interactions between alpine plant species and the nurse cushion plant *Laretia acaulis* do not increase with elevation in the Andes of central Chile. *New Phytologist* 169, 59–69. doi:10.1111/j.1469-8137.2005.01573.x.

Chae, L., Kim, T., Nilo-Poyanco, R., and Rhee, S. Y. (2014). Genomic Signatures of Specialized Metabolism in Plants. *Science* 344, 510–513. doi:10.1126/science.1252076.

Chai, R., Mao, J., Chen, H., Wang, Y., Shi, X., Jin, M., et al. (2021). Human-caused long-term changes in global aridity. *npj Clim Atmos Sci* 4, 65. doi:10.1038/s41612-021-00223-5.

Chaloner, T. M., Gurr, S. J., and Bebber, D. P. (2021). Plant pathogen infection risk tracks global crop yields under climate change. *Nat. Clim. Chang.* 11, 710–715. doi:10.1038/s41558-021-01104-8.

Chapin III, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., et al. (2000). Consequences of changing biodiversity. *Nature* 405, 234–242. doi:10.1038/35012241.

Chen, D., Shao, Q., Yin, L., Younis, A., and Zheng, B. (2019). Polyamine Function in Plants: Metabolism, Regulation on Development, and Roles in Abiotic Stress Responses. *Front. Plant Sci.* 9, 1945. doi:10.3389/fpls.2018.01945.

Chen, T. H. H., and Murata, N. (2002). Enhancement of tolerance of abiotic stress by metabolic engineering of betaines and other compatible solutes. *Current Opinion in Plant Biology* 5, 250–257. doi:10.1016/S1369-5266(02)00255-8.

Chia, D. W., Yoder, T. J., Reiter, W.-D., and Gibson, S. I. (2000). Fumaric acid: an overlooked form of fixed carbon in Arabidopsis and other plant species. *Planta* 211, 743–751. doi:10.1007/s004250000345.

Chicco, D. (2017). Ten quick tips for machine learning in computational biology. *BioData Mining* 10, 35. doi:10.1186/s13040-017-0155-3.

C**hong, J., Poutaraud, A., and Hugueney, P.** (2009). Metabolism and roles of stilbenes in plants. *Plant Science* 177, 143–155. doi:10.1016/j.plantsci.2009.05.012.

Choudhury, F. K., Rivero, R. M., Blumwald, E., and Mittler, R. (2017). Reactive oxygen species, abiotic stress and stress combination. *Plant J* 90, 856–867. doi:10.1111/tpj.13299.

Chouhan, S., Sharma, K., Zha, J., Guleria, S., and Koffas, M. A. G. (2017). Recent Advances in the Recombinant Biosynthesis of Polyphenols. *Front. Microbiol.* 8, 2259. doi:10.3389/fmicb.2017.02259.

Chun, H., and Kele¸s, S. (2010). Sparse partial least squares regression for simultaneous dimension reduction and variable selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 72, 3–25. doi:10.1111/j.1467-9868.2009.00723.x.

Cirak, C., Radusiene, J., Jakstas, V., Ivanauskas, L., Seyis, F., and Yayla, F. (2017). Altitudinal changes in secondary metabolite contents of Hypericum androsaemum and Hypericum polyphyllum. *Biochemical Systematics and Ecology* 70, 108–115. doi:10.1016/j.bse.2016.11.006.

Clancy, M. V., Zytynska, S. E., Moritz, F., Witting, M., Schmitt-Kopplin, P., Weisser, W. W., et al. (2018). Metabotype variation in a field population of tansy plants influences aphid host selection: Plant chemical diversity in a plant-aphid system. *Plant Cell Environ* 41, 2791–2805. doi:10.1111/pce.13407.

Clemente-Moreno, M. J., Omranian, N., Sáez, P., Figueroa, C. M., Del-Saz, N., Elso, M., et al. (2020). Cytochrome respiration pathway and sulphur metabolism sustain stress tolerance to low temperature in the Antarctic species *Colobanthus quitensis*. *New Phytol* 225, 754–768. doi:10.1111/nph.16167.

Cocuron, J.-C., and Alonso, A. P. (2014). "Liquid Chromatography Tandem Mass Spectrometry for Measuring 13C-labeling in Intermediates of the Glycolysis and Pentose Phosphate Pathway," in *Plant Metabolic Flux Analysis* Methods in Molecular Biology., eds. M. Dieuaide-Noubhani and A. P. Alonso (Totowa, NJ: Humana Press), 131–142. doi:10.1007/978-1-62703-688-7_9.

Cohen, J. E. (2004). Mathematics Is Biology's Next Microscope, Only Better; Biology Is Mathematics' Next Physics, Only Better. *PLoS Biol* 2, e439. doi:10.1371/journal.pbio.0020439.

Collins, R. P., and Jones, M. B. (1986). The influence of climatic factors on the distribution of C4 species in Europe. *Vegetatio* 64, 121–129.

Colombié, S., Beauvoit, B., Nazaret, C., Bénard, C., Vercambre, G., Le Gall, S., et al. (2017). Respiration climacteric in tomato fruits elucidated by constraint-based modelling. *New Phytol* 213, 1726–1739. doi:10.1111/nph.14301.

Coruzzi, G. M., and Zhou, L. (2001). Carbon and nitrogen sensing and signaling in plants: emerging 'matrix effects.' *Current Opinion in Plant Biology* 4, 247–253. doi:10.1016/S1369-5266(00)00168-0.

Cottret, L., Wildridge, D., Vinson, F., Barrett, M. P., Charles, H., Sagot, M.-F., et al. (2010). MetExplore: a web server to link metabolomic experiments and genome-scale metabolic networks. *Nucleic Acids Res.* 38, W132-137. doi:10.1093/nar/gkq312.

Cui, G., Ji, G., Liu, S., Li, B., Lian, L., He, W., et al. (2019). Physiological adaptations of Elymus dahuricus to high altitude on the Qinghai–Tibetan Plateau. *Acta Physiol Plant* 41, 115. doi:10.1007/s11738-019-2904-z.

Das, P., Nutan, K. K., Singla-Pareek, S. L., and Pareek, A. (2015). Oxidative environment and redox homeostasis in plants: dissecting out significant contribution of major cellular organelles. *Front. Environ. Sci.* 2. doi:10.3389/fenvs.2014.00070.

de Matos, P., Alcántara, R., Dekker, A., Ennis, M., Hastings, J., Haug, K., et al. (2010). Chemical Entities of Biological Interest: an update. *Nucleic Acids Research* 38, D249–D254. doi:10.1093/nar/gkp886.

Decros, G., Baldet, P., Beauvoit, B., Stevens, R., Flandin, A., Colombié, S., et al. (2019). Get the Balance Right: ROS Homeostasis and Redox Signalling in Fruit. *Front. Plant Sci.* 10, 1091. doi:10.3389/fpls.2019.01091.

Defossez, E., Pitteloud, C., Descombes, P., Glauser, G., Allard, P.-M., Walker, T. W. N., et al. (2021). Spatial and evolutionary predictability of phytochemical diversity. *Proc Natl Acad Sci USA* 118, e2013344118. doi:10.1073/pnas.2013344118.

Delgado-Sánchez, P., Yáñez-Espinosa, L., Jiménez-Bremont, J. F., Chapa-Vargas, L., and Flores, J. (2013). Ecophysiological and Anatomical Mechanisms behind the Nurse Effect: Which Are More Important? A Multivariate Approach for Cactus Seedlings. *PLoS ONE* 8, e81513. doi:10.1371/journal.pone.0081513.

Di Ferdinando, M., Brunetti, C., Fini, A., and Tattini, M. (2012). "Flavonoids as Antioxidants in Plants Under Abiotic Stresses," in *Abiotic Stress Responses in Plants*, eds. P. Ahmad and M. N. V. Prasad (New York, NY: Springer New York), 159–179. doi:10.1007/978-1-4614-0634-1_9.

Di Guida, R., Engel, J., Allwood, J. W., Weber, R. J. M., Jones, M. R., Sommer, U., et al. (2016). Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. *Metabolomics* 12, 93. doi:10.1007/s11306-016-1030-9.

Díaz, F. P., Frugone, M., Gutiérrez, R. A., and Latorre, C. (2016). Nitrogen cycling in an extreme hyperarid environment inferred from $\delta^{15}N$ analyses of plants, soils and herbivore diet. *Scientific Reports* 6, 22226. doi:10.1038/srep22226.

Díaz, F. P., Latorre, C., Carrasco-Puga, G., Wood, J. R., Wilmshurst, J. M., Soto, D. C., et al. (2019). Multiscale climate change impacts on plant diversity in the Atacama Desert. *Global Change Biology* 25, 1733–1745. doi:10.1111/gcb.14583.

Domergue, F., Vishwanath, S. J., Joubès, J., Ono, J., Lee, J. A., Bourdon, M., et al. (2010). Three Arabidopsis Fatty Acyl-Coenzyme A Reductases, FAR1, FAR4, and FAR5, Generate Primary Fatty Alcohols Associated with Suberin Deposition. *Plant Physiol.* 153, 1539–1554. doi:10.1104/pp.110.158238.

Dussarrat, T., Decros, G., Díaz, F. P., Gibon, Y., Latorre, C., Rolin, D., et al. (2021). "Another Tale from the Harsh World: How Plants Adapt to Extreme Environments," in *Annual Plant Reviews online* (American Cancer Society), 551–603. doi:10.1002/9781119312994.apr0758.

Edwards, E. J. (2019). Evolutionary trajectories, accessibility and other metaphors: the case of C_4 and CAM photosynthesis. *New Phytol* 223, 1742–1755. doi:10.1111/nph.15851.

Ehleringer, J. R., Cerling, T. E., and Helliker, B. R. (1997). C 4 photosynthesis, atmospheric CO 2 , and climate. *Oecologia* 112, 285–299. doi:10.1007/s004420050311.

Engelhardt, B., Frőhlich, H., and Kschischo, M. (2016). Learning (from) the errors of a systems biology model. *Sci Rep* 6, 20772. doi:10.1038/srep20772.

Enquist, B. J., and Leffler, A. J. (2001). Long-term tree ring chronologies from sympatric tropical dry-forest trees: individualistic responses to climatic variation. *J. Trop. Ecol.* 17, 41–60. doi:10.1017/S0266467401001031.

Erb, M., and Kliebenstein, D. J. (2020). Plant Secondary Metabolites as Defenses, Regulators, and Primary Metabolites: The Blurred Functional Trichotomy. *Plant Physiol.* 184, 39–52. doi:10.1104/pp.20.00433.

Eshel, G., Araus, V., Undurraga, S., Soto, D. C., Moraga, C., Montecinos, A., et al. (2021). Plant ecological genomics at the limits of life in the Atacama Desert. *Proc Natl Acad Sci USA* 118, e2101177118. doi:10.1073/pnas.2101177118.

Fahy, E., Cotter, D., Sud, M., and Subramaniam, S. (2011). Lipid classification, structures and tools. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 1811, 637–647. doi:10.1016/j.bbalip.2011.06.009.

Fang, C., Fernie, A. R., and Luo, J. (2019). Exploring the Diversity of Plant Metabolism. *Trends in Plant Science* 24, 83–98. doi:10.1016/j.tplants.2018.09.006.

Fernandez, O., Urrutia, M., Bernillon, S., Giauffret, C., Tardieu, F., Le Gouis, J., et al. (2016). Fortune telling: metabolic markers of plant performance. *Metabolomics* 12, 158. doi:10.1007/s11306-016-1099-1.

Fernández-Marín, B., Gulías, J., Figueroa, C. M., Iñiguez, C., Clemente-Moreno, M. J., Nunes-Nesi, A., et al. (2020). How do vascular plants perform photosynthesis in extreme environments? An integrative ecophysiological and biochemical story. *Plant J* 101, 979–1000. doi:10.1111/tpj.14694.

Fernie, A. R., Carrari, F., and Sweetlove, L. J. (2004). Respiratory metabolism: glycolysis, the TCA cycle and mitochondrial electron transport. *Current Opinion in Plant Biology* 7, 254–261. doi:10.1016/j.pbi.2004.03.007.

Fiehn, O. (2002). "Metabolomics — the link between genotypes and phenotypes," in *Functional Genomics*, ed. C. Town (Dordrecht: Springer Netherlands), 155–171. doi:10.1007/978-94-010-0448-0_11.

Flores, J., and Jurado, E. (2003). Are nurse-protégé interactions more common among plants from arid environments? *Journal of Vegetation Science* 14, 911–916. doi:10.1111/j.1654-1103.2003.tb02225.x.

Fondi, M., Emiliani, G., and Fani, R. (2009). Origin and evolution of operons and metabolic pathways. *Research in Microbiology* 160, 502–512. doi:10.1016/j.resmic.2009.05.001.

Forde, B. G., and Lea, P. J. (2007). Glutamate in plants: metabolism, regulation, and signalling. *Journal of Experimental Botany* 58, 2339–2358. doi:10.1093/jxb/erm121.

Foyer, C. H. (2018). Reactive oxygen species, oxidative signaling and the regulation of photosynthesis. *Environmental and Experimental Botany* 154, 134–142. doi:10.1016/j.envexpbot.2018.05.003.

Foyer, C. H., and Noctor, G. (2016). Stress-triggered redox signalling: what's in pROSpect?: What's in pROSpect? *Plant, Cell & Environment* 39, 951–964. doi:10.1111/pce.12621.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* 33, 1–22.

Gagneul, D., Aïnouche, A., Duhazé, C., Lugan, R., Larher, F. R., and Bouchereau, A. (2007). A Reassessment of the Function of the So-Called Compatible Solutes in the Halophytic Plumbaginaceae Limonium latifolium. *Plant Physiol* 144, 1598–1611. doi:10.1104/pp.107.099820.

Gakière, B., Hao, J., Bont, L. de, Pétriacq, P., Nunes-Nesi, A., and Fernie, A. R. (2018). NAD+ Biosynthesis and Signaling in Plants. *Critical Reviews in Plant Sciences* 37, 259–307. doi:10.1080/07352689.2018.1505591.

Galinski, E. A. (1993). Compatible solutes of halophilic eubacteria: molecular principles, water-solute interaction, stress protection. *Experientia* 49, 487–496. doi:10.1007/BF01955150.

Gibon, Y., Blaesing, O. E., Hannemann, J., Carillo, P., Höhne, M., Hendriks, J. H. M., et al. (2004). A Robot-based platform to measure multiple enzyme activities in Arabidopsis using a set of cycling assays: comparison of changes of enzyme activities and transcript levels during diurnal cycles and in prolonged darkness. *Plant Cell* 16, 3304–3325. doi:10.1105/tpc.104.025973.

Gleadow, R. M., and Woodrow, I. E. (2002). Mini-Review: Constraints on Effectiveness of Cyanogenic Glycosides in Herbivore Defense. *Journal of Chemical Ecology* 28, 1301–1313. doi:10.1023/A:1016298100201.

Gleason, H. A. (1926). The Individualistic Concept of the Plant Association. *Bulletin of the Torrey Botanical Club* 53, 7. doi:10.2307/2479933.

Gorrochategui, E., Jaumot, J., Lacorte, S., and Tauler, R. (2016). Data analysis strategies for targeted and untargeted LC-MS metabolomic studies: Overview and workflow. *TrAC Trends in Analytical Chemistry* 82, 425–442. doi:10.1016/j.trac.2016.07.004.

Grant, N., Onda, Y., Kakizaki, Y., Ito, K., Watling, J., and Robinson, S. (2009). Two Cys or Not Two Cys? That Is the Question; Alternative Oxidase in the Thermogenic Plant Sacred Lotus. *Plant Physiology* 150, 987–995. doi:10.1104/pp.109.139394.

Guerrero, P. C., Rosas, M., Arroyo, M. T. K., and Wiens, J. J. (2013). Evolutionary lag times and recent origin of the biota of an ancient desert (Atacama-Sechura). *Proceedings of the National Academy of Sciences* 110, 11469–11474. doi:10.1073/pnas.1308721110.

Gutiérrez-Grijalva, E. P., López-Martínez, L. X., Contreras-Angulo, L. A., Elizalde-Romero, C. A., and Heredia, J. B. (2020). "Plant Alkaloids: Structures and Bioactive Properties," in *Plant-derived Bioactives*, ed. M. K. Swamy (Singapore: Springer Singapore), 85–117. doi:10.1007/978-981-15-2361-8_5.

Haenlein, M., and Kaplan, A. M. (2004). A Beginner's Guide to Partial Least Squares Analysis. *Understanding Statistics* 3, 283–297. doi:10.1207/s15328031us0304_4.

Hartmann, T. (1996). "Diversity and variability of plant secondary metabolism: a mechanistic view," in *Proceedings of the 9th International Symposium on Insect-Plant Relationships*, eds. E. Städler, M. Rowell-Rahier, and R. Bauer (Dordrecht: Springer Netherlands), 177–188. doi:10.1007/978-94-009-1720-0_42.

Hartmann, T. (2007). From waste products to ecochemicals: Fifty years research of plant secondary metabolism. *Phytochemistry* 68, 2831–2846. doi:10.1016/j.phytochem.2007.09.017.

Harwood, J. L. (1996). Recent advances in the biosynthesis of plant fatty acids. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* 1301, 7–56. doi:10.1016/0005-2760(95)00242-1.

Hasanuzzaman, M., Bhuyan, M. H. M. B., Mahmud, J. A., Nahar, K., Mohsin, S. M., Parvin, K., et al. (2018). Interaction of sulfur with phytohormones and signaling molecules in conferring abiotic stress tolerance to plants. *Plant Signaling & Behavior* 13, e1477905. doi:10.1080/15592324.2018.1477905.

Hasegawa, T., Sakurai, G., Fujimori, S., Takahashi, K., Hijioka, Y., and Masui, T. (2021). Extreme climate events increase risk of global food insecurity and adaptation needs. *Nat Food* 2, 587–595. doi:10.1038/s43016-021-00335-4.

Hashim, A. M., Alharbi, B. M., Abdulmajeed, A. M., Elkelish, A., Hozzein, W. N., and Hassan, H. M. (2020). Oxidative Stress Responses of Some Endemic Plants to High Altitudes by Intensifying Antioxidants and Secondary Metabolites Content. *Plants* 9, 869. doi:10.3390/plants9070869.

Hastie, T., Tibshirani, R., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*. 2nd ed. New York, NY: Springer.

Häusler, R. E., Ludewig, F., and Krueger, S. (2014). Amino acids – A life between metabolism and signaling. *Plant Science* 229, 225–237. doi:10.1016/j.plantsci.2014.09.011.

Havaux, M. (2014). Carotenoid oxidation products as stress signals in plants. *Plant J* 79, 597–606. doi:10.1111/tpj.12386.

Hendriks, J. H. M., Kolbe, A., Gibon, Y., Stitt, M., and Geigenberger, P. (2003). ADP-Glucose Pyrophosphorylase Is Activated by Posttranslational Redox-Modification in Response to Light and to Sugars in Leaves of Arabidopsis and Other Plant Species. *Plant Physiology* 133, 838–849. doi:10.1104/pp.103.024513.

Hendriks, M. M. W. B., Eeuwijk, F. A. van, Jellema, R. H., Westerhuis, J. A., Reijmers, T. H., Hoefsloot, H. C. J., et al. (2011). Data-processing strategies for metabolomics studies. *TrAC Trends in Analytical Chemistry* 30, 1685–1698. doi:10.1016/j.trac.2011.04.019.

Hesse, H., Nikiforova, V., Gakiere, B., and Hoefgen, R. (2004). Molecular analysis and control of cysteine biosynthesis: integration of nitrogen and sulphur metabolism. *Journal of Experimental Botany* 55, 1283–1292. doi:10.1093/jxb/erh136.

Hildebrandt, T. M., Nunes Nesi, A., Araújo, W. L., and Braun, H.-P. (2015). Amino Acid Catabolism in Plants. *Molecular Plant* 8, 1563–1579. doi:10.1016/j.molp.2015.09.005.

Hisano, M., Searle, E. B., and Chen, H. Y. H. (2018). Biodiversity as a solution to mitigate climate change impacts on the functioning of forest ecosystems: Biodiversity to mitigate climate change impacts. *Biol Rev* 93, 439–456. doi:10.1111/brv.12351.

Holmes, M. G., and Keiller, D. R. (2002). Effects of pubescence and waxes on the reflectance of leaves in the ultraviolet and photosynthetic wavebands: a comparison of a range of species: *Ultraviolet leaf reflectance*. *Plant, Cell & Environment* 25, 85–93. doi:10.1046/j.1365-3040.2002.00779.x.

Holtorf, H., Guitton, M.-C., and Reski, R. (2002). Plant functional genomics. *Naturwissenschaften* 89, 235–249. doi:10.1007/s00114-002-0321-3.

Horai, H., Arita, M., Kanaya, S., Nihei, Y., Ikeda, T., Suwa, K., et al. (2010). MassBank: a public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* 45, 703–714. doi:10.1002/jms.1777.

Hu, L., Hu, T., Zhang, X., Pang, H., and Fu, J. (2012). Exogenous Glycine Betaine Ameliorates the Adverse Effect of Salt Stress on Perennial Ryegrass. *J. Amer. Soc. Hort. Sci.* 137, 38–46. doi:10.21273/JASHS.137.1.38.

Hu, W., Ran, J., Dong, L., Du, Q., Ji, M., Yao, S., et al. (2021). Aridity-driven shift in biodiversity–soil multifunctionality relationships. *Nat Commun* 12, 5350. doi:10.1038/s41467-021-25641-0.

Huntley, B. (1991). How Plants Respond to Climate Change: Migration Rates, Individualism and the Consequences for Plant Communities. *Annals of Botany* 67, 15–22. doi:10.1093/oxfordjournals.aob.a088205.

Igamberdiev, A. U., and Eprintsev, A. T. (2016). Organic Acids: The Pools of Fixed Carbon Involved in Redox Regulation and Energy Balance in Higher Plants. *Front. Plant Sci.* 7. doi:10.3389/fpls.2016.01042.

Ishikawa, T., Dowdle, J., and Smirnoff, N. (2006). Progress in manipulating ascorbic acid biosynthesis and accumulation in plants. *Physiol Plant* 126, 343–355. doi:10.1111/j.1399-3054.2006.00640.x.

Isla, M. I., Zampini, I. C., and Cuello, A. S. (2021). "Parastrephia lucida (Meyen) Cabrera," in *Medicinal and Aromatic Plants of South America Vol. 2* Medicinal and Aromatic Plants of the World., eds. Á. Máthé and A. Bandoni (Cham: Springer International Publishing), 399–409. doi:10.1007/978-3-030-62818-5_31.

Jacob, D., David, R., Aubin, S., and Gibon, Y. (2020). Making experimental data tables in the life sciences more FAIR: a pragmatic approach. *GigaScience* 9, giaa144. doi:10.1093/gigascience/giaa144.

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* 3, 318–356. doi:10.1016/S0022-2836(61)80072-7.

Jacobsen, A. L., Pratt, R. B., Davis, S. D., and Ewers, F. W. (2007). Cavitation resistance and seasonal hydraulics differ among three arid Californian plant communities. *Plant Cell Environ* 30, 1599–1609. doi:10.1111/j.1365-3040.2007.01729.x.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters* 31, 651–666. doi:10.1016/j.patrec.2009.09.011.

Jander, G., Kolukisaoglu, U., Stahl, M., and Yoon, G. M. (2020). Editorial: Physiological Aspects of Non-proteinogenic Amino Acids in Plants. *Front. Plant Sci.* 11, 519464. doi:10.3389/fpls.2020.519464.

Jeffery, C. J. (1999). Moonlighting proteins. *Trends in Biochemical Sciences* 24, 8–11. doi:10.1016/S0968-0004(98)01335-8.

Jeon, J.-S., Carreno-Quintero, N., van Eekelen, H. D. L. M., De Vos, R. C. H., Raaijmakers, J. M., and Etalo, D. W. (2021). Impact of root-associated strains of three Paraburkholderia species on primary and secondary metabolism of Brassica oleracea. *Sci Rep* 11, 2781. doi:10.1038/s41598-021-82238-9.

Jia, Q., Kong, D., Li, Q., Sun, S., Song, J., Zhu, Y., et al. (2019). The Function of Inositol Phosphatases in Plant Tolerance to Abiotic Stress. *IJMS* 20, 3999. doi:10.3390/ijms20163999.

Jiang, G., Wang, Z., Shang, H., Yang, W., Hu, Z., Phillips, J., et al. (2007). Proteome analysis of leaves from the resurrection plant Boea hygrometrica in response to dehydration and rehydration. *Planta* 225, 1405–1420. doi:10.1007/s00425-006-0449-z.

Johannsen, W. (2014). The genotype conception of heredity. *Int. J. Epidemiol.* 43, 989–1000. doi:10.1093/ije/dyu063.

Jordan, T. E., Kirk-Lawlor, N. E., Blanco, N. P., Rech, J. A., and Cosentino, N. J. (2014). Landscape modification in response to repeated onset of hyperarid paleoclimate states since 14 Ma, Atacama Desert, Chile. *Geological Society of America Bulletin* 126, 1016–1046. doi:10.1130/B30978.1.

Joswig, J. S., Wirth, C., Schuman, M. C., Kattge, J., Reu, B., Wright, I. J., et al. (2021). Climatic and soil factors explain the two-dimensional spectrum of global plant trait variation. *Nat Ecol Evol*. doi:10.1038/s41559-021-01616-8.

Jr, F. E. H. (2021). *Hmisc: Harrell Miscellaneous*. Available at: https://CRAN.R-project.org/package=Hmisc.

Kaisermann, A., Vries, F. T., Griffiths, R. I., and Bardgett, R. D. (2017). Legacy effects of drought on plant–soil feedbacks and plant–plant interactions. *New Phytol* 215, 1413–1424. doi:10.1111/nph.14661.

Kandler, O., and Hopf, H. (1980). "Occurrence, Metabolism, and Function of Oligosaccharides," in *Carbohydrates: Structure and Function* (Elsevier), 221–270. doi:10.1016/B978-0-12-675403-2.50013-2.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucl. Acids Res.* 42, D199–D205. doi:10.1093/nar/gkt1076.

Kang, S.-H., Pandey, R. P., Lee, C.-M., Sim, J.-S., Jeong, J.-T., Choi, B.-S., et al. (2020). Genome-enabled discovery of anthraquinone biosynthesis in Senna tora. *Nat Commun* 11, 5875. doi:10.1038/s41467-020-19681-1.

Karp, P. D., Midford, P. E., Billington, R., Kothari, A., Krummenacker, M., Latendresse, M., et al. (2021). Pathway Tools version 23.0 update: software for pathway/genome informatics and systems biology. *Briefings in Bioinformatics* 22, 109–126. doi:10.1093/bib/bbz104.

Kashyap, S. P., Prasanna, H. C., Kumari, N., Mishra, P., and Singh, B. (2020). Understanding salt tolerance mechanism using transcriptome profiling and de novo assembly of wild tomato Solanum chilense. *Sci Rep* 10, 15835. doi:10.1038/s41598-020-72474-w.

Kassambara, A., and Mundt, F. (2020). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. Available at: https://CRAN.R-project.org/package=factoextra.

Kasukabe, Y., He, L., Nada, K., Misawa, S., Ihara, I., and Tachibana, S. (2004). Overexpression of Spermidine Synthase Enhances Tolerance to Multiple Environmental Stresses and Up-Regulates the Expression of Various Stress-Regulated Genes in Transgenic Arabidopsis thaliana. *Plant and Cell Physiology* 45, 712–722. doi:10.1093/pcp/pch083.

Kaur, G., and Asthir, B. (2015). Proline: a key player in plant abiotic stress tolerance. *Biologia plant.* 59, 609–619. doi:10.1007/s10535-015-0549-3.

Ke, P., Miki, T., and Ding, T. (2015). The soil microbial community predicts the importance of plant traits in plant–soil feedback. *New Phytol* 206, 329–341. doi:10.1111/nph.13215.

Ke, P., Zee, P. C., and Fukami, T. (2021). Dynamic plant–soil microbe interactions: the neglected effect of soil conditioning time. *New Phytol* 231, 1546–1558. doi:10.1111/nph.17420.

Kim, H. K., Choi, Y. H., and Verpoorte, R. (2010). NMR-based metabolomic analysis of plants. *Nat Protoc* 5, 536–549. doi:10.1038/nprot.2009.237.

Kinnersley, A. M., and Turano, F. J. (2000). Gamma Aminobutyric Acid (GABA) and Plant Responses to Stress. *Critical Reviews in Plant Sciences* 19, 479–509. doi:10.1080/07352680091139277.

Kolattukudy, P. E. (1970). Plant waxes. *Lipids* 5, 259–275. doi:10.1007/BF02532477.

Kolde, R. (2019). *pheatmap: Pretty Heatmaps*. Available at: https://CRAN.R-project.org/package=pheatmap.

Körner, C. (2003). *Alpine Plant Life*. Berlin, Heidelberg: Springer Berlin Heidelberg doi:10.1007/978-3-642-18970-8.

Kranner, I., Beckett, R. P., Wornik, S., Zorn, M., and Pfeifhofer, H. W. (2002). Revival of a resurrection plant correlates with its antioxidant status. *The Plant Journal* 31, 13–24. doi:10.1046/j.1365-313X.2002.01329.x.

Kruger, N. J., and von Schaewen, A. (2003). The oxidative pentose phosphate pathway: structure and organisation. *Current Opinion in Plant Biology* 6, 236–246. doi:10.1016/S1369-5266(03)00039-6.

Kumari, M., Joshi, R., and Kumar, R. (2020). Metabolic signatures provide novel insights to Picrorhiza kurroa adaptation along the altitude in Himalayan region. *Metabolomics* 16, 77. doi:10.1007/s11306-020-01698-8.

Lattanzio, V. (2013). "Phenolic Compounds: Introduction," in *Natural Products*, eds. K. G. Ramawat and J.-M. Mérillon (Berlin, Heidelberg: Springer Berlin Heidelberg), 1543–1580. doi:10.1007/978-3-642-22144-6_57.

Laughlin, D. C., Joshi, C., van Bodegom, P. M., Bastow, Z. A., and Fulé, P. Z. (2012). A predictive model of community assembly that incorporates intraspecific trait variation. *Ecol Lett* 15, 1291–1299. doi:10.1111/j.1461-0248.2012.01852.x.

Le Gall, H., Philippe, F., Domon, J.-M., Gillet, F., Pelloux, J., and Rayon, C. (2015). Cell Wall Metabolism in Response to Abiotic Stress. *Plants* 4, 112–166. doi:10.3390/plants4010112.

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR : An *R* Package for Multivariate Analysis. *J. Stat. Soft.* 25. doi:10.18637/jss.v025.i01.

Lesk, C., Coffel, E., Winter, J., Ray, D., Zscheischler, J., Seneviratne, S. I., et al. (2021). Stronger temperature–moisture couplings exacerbate the impact of climate warming on global crop yields. *Nat Food* 2, 683–691. doi:10.1038/s43016-021-00341-6.

Lespinet, O., Wolf, Y. I., Koonin, E. V., and Aravind, L. (2002). The Role of Lineage-Specific Gene Family Expansion in the Evolution of Eukaryotes. *Genome Res.* 12, 1048–1059. doi:10.1101/gr.174302.

Lewis, J. E., and Kemp, M. L. (2021). Integration of machine learning and genome-scale metabolic modeling identifies multi-omics biomarkers for radiation resistance. *Nat Commun* 12, 2700. doi:10.1038/s41467-021-22989-1.

Li, J., Liu, L.-N., Meng, Q., Fan, H., and Sui, N. (2020). The roles of chloroplast membrane lipids in abiotic stress responses. *Plant Signaling & Behavior* 15, 1807152. doi:10.1080/15592324.2020.1807152.

Li, Y., Kong, D., Fu, Y., Sussman, M. R., and Wu, H. (2020). The effect of developmental and environmental factors on secondary metabolites in medicinal plants. *Plant Physiology and Biochemistry* 148, 80–89. doi:10.1016/j.plaphy.2020.01.006.

Loescher, W. H., and Everard, J. D. (2000). "Regulation of Sugar Alcohol Biosynthesis," in *Photosynthesis* Advances in Photosynthesis and Respiration., eds. R. C. Leegood, T. D. Sharkey, and S. von Caemmerer (Dordrecht: Springer Netherlands), 275–299. doi:10.1007/0-306-48137-5_12.

Long, S. P., Marshall-Colon, A., and Zhu, X.-G. (2015). Meeting the Global Food Demand of the Future by Engineering Crop Photosynthesis and Yield Potential. *Cell* 161, 56–66. doi:10.1016/j.cell.2015.03.019.

López-Fernández, O., Domínguez, R., Pateiro, M., Munekata, P. E. S., Rocchetti, G., and Lorenzo, J. M. (2020). Determination of Polyphenols Using Liquid Chromatography–Tandem Mass Spectrometry Technique (LC–MS/MS): A Review. *Antioxidants* 9, 479. doi:10.3390/antiox9060479.

Lortie, C. J., Brooker, R. W., Choler, P., Kikvidze, Z., Michalet, R., Pugnaire, F. I., et al. (2004). Rethinking plant community theory. *Oikos* 107, 433–438. doi:10.1111/j.0030-1299.2004.13250.x.

Lou, Y., Gou, J.-Y., and Xue, H.-W. (2007). PIP5K9, an *Arabidopsis* Phosphatidylinositol Monophosphate Kinase, Interacts with a Cytosolic Invertase to Negatively Regulate Sugar-Mediated Root Growth. *The Plant Cell* 19, 163–181. doi:10.1105/tpc.106.045658.

Loupit, G., Prigent, S., Franc, C., De Revel, G., Richard, T., Cookson, S. J., et al. (2020). Polyphenol Profiles of Just Pruned Grapevine Canes from Wild *Vitis* Accessions and *Vitis* vinifera Cultivars. *Journal of Agricultural and Food Chemistry*, acs.jafc.9b08099. doi:10.1021/acs.jafc.9b08099.

Lugan, R., Niogret, M.-F., Leport, L., Guégan, J.-P., Larher, F. R., Savouré, A., et al. (2010). Metabolome and water homeostasis analysis of Thellungiella salsuginea suggests that dehydration tolerance is a key response to osmotic stress in this halophyte. *Plant J.* 64, 215–229. doi:10.1111/j.1365-313X.2010.04323.x.

Luna, E., Flandin, A., Cassan, C., Prigent, S., Chevanne, C., Kadiri, C. F., et al. (2020). Metabolomics to Exploit the Primed Immune System of Tomato Fruit. *Metabolites* 10, 96. doi:10.3390/metabo10030096.

Ma, L., Sun, X., Kong, X., Galvan, J. V., Li, X., Yang, S., et al. (2015). Physiological, biochemical and proteomics analysis reveals the adaptation strategies of the alpine plant Potentilla saundersiana at altitude gradient of the Northwestern Tibetan Plateau. *Journal of Proteomics* 112, 63–82. doi:10.1016/j.jprot.2014.08.009.

Ma, Z., Liu, H., Mi, Z., Zhang, Z., Wang, Y., Xu, W., et al. (2017). Climate warming reduces the temporal stability of plant community biomass production. *Nat Commun* 8, 15378. doi:10.1038/ncomms15378.

Machado, R. A. R., Ferrieri, A. P., Robert, C. A. M., Glauser, G., Kallenbach, M., Baldwin, I. T., et al. (2013). Leaf-herbivore attack reduces carbon reserves and regrowth from the roots via jasmonate and auxin signaling. *New Phytol* 200, 1234–1246. doi:10.1111/nph.12438.

Maeda, H., and Dudareva, N. (2012). The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants. *Annu. Rev. Plant Biol.* 63, 73–105. doi:10.1146/annurev-arplant-042811-105439.

Mandal, S. M., Chakraborty, D., and Dey, S. (2010). Phenolic acids act as signaling molecules in plant-microbe symbioses. *Plant Signaling & Behavior* 5, 359–368. doi:10.4161/psb.5.4.10871.

Marti, G., Erb, M., Boccard, J., Glauser, G., Doyen, G. R., Villard, N., et al. (2013). Metabolomics reveals herbivore-induced metabolites of resistance and susceptibility in maize leaves and roots: Plant-insect metabolomics. *Plant, Cell & Environment* 36, 621–639. doi:10.1111/pce.12002.

Martinelli, T., Whittaker, A., Bochicchio, A., Vazzana, C., Suzuki, A., and Masclaux-Daubresse, C. (2007). Amino acid pattern and glutamate metabolism during dehydration stress in the "resurrection" plant Sporobolus stapfianus: a comparison between desiccation-sensitive and desiccation-tolerant leaves. *Journal of Experimental Botany* 58, 3037–3046. doi:10.1093/jxb/erm161.

McIntire, E. J. B., and Fajardo, A. (2014). Facilitation as a ubiquitous driver of biodiversity. *New Phytol* 201, 403–416. doi:10.1111/nph.12478.

Meacham-Hensold, K., Montes, C. M., Wu, J., Guan, K., Fu, P., Ainsworth, E. A., et al. (2019). High-throughput field phenotyping using hyperspectral reflectance and partial least squares regression (PLSR) reveals genetic modifications to photosynthetic capacity. *Remote Sensing of Environment* 231, 111176. doi:10.1016/j.rse.2019.04.029.

Mendiburu, F. de (2021). *agricolae: Statistical Procedures for Agricultural Research*. Available at: https://CRAN.R-project.org/package=agricolae.

Mevik, B.-H., Wehrens, R., and Liland, K. H. (2020). *pls: Partial Least Squares and Principal Component Regression*. Available at: https://CRAN.R-project.org/package=pls.

Mikulic-Petkovsek, M., Schmitzer, V., Slatnar, A., Stampar, F., and Veberic, R. (2012). Composition of Sugars, Organic Acids, and Total Phenolics in 25 Wild or Cultivated Berry Species. *Journal of Food Science* 77, C1064–C1070. doi:10.1111/j.1750-3841.2012.02896.x.

Minteguiaga, M., Fariña, L., Cassel, E., Fiedler, S., Catalán, C. A. N., and Dellacassa, E. (2021). Chemical compositions of essential oil from the aerial parts of male and female plants of *Baccharis tridentata* Vahl. (Asteraceae). *Journal of Essential Oil Research* 33, 299–307. doi:10.1080/10412905.2020.1829720.

Moellering, E. R., Muthan, B., and Benning, C. (2010). Freezing Tolerance in Plants Requires Lipid Remodeling at the Outer Chloroplast Membrane. *Science* 330, 226–228. doi:10.1126/science.1191803.

Moing, A. (2000). "Sugar alcohols as carbohydrate reserves in some higher plants," in *Developments in Crop Science* (Elsevier), 337–358. doi:10.1016/S0378-519X(00)80017-3.

Monschein, M., Jaindl, K., Buzimkić, S., and Bucar, F. (2015). Content of phenolic compounds in wild populations of *Epilobium angustifolium* growing at different altitudes. *Pharmaceutical Biology* 53, 1576–1582. doi:10.3109/13880209.2014.993039.

Moore, J. P., Le, N. T., Brandt, W. F., Driouich, A., and Farrant, J. M. (2009). Towards a systems-based understanding of plant desiccation tolerance. *Trends in Plant Science* 14, 110–117. doi:10.1016/j.tplants.2008.11.007.

Mori, H. (2000). Direct Determination of Nitrate Using Nitrate Reductase in a Flow System. *JOURNAL OF HEALTH SCIENCE* 46, 385–388. doi:10.1248/jhs.46.385.

Nagegowda, D. A., and Gupta, P. (2020). Advances in biosynthesis, regulation, and metabolic engineering of plant specialized terpenoids. *Plant Science* 294, 110457. doi:10.1016/j.plantsci.2020.110457.

Naikoo, M. I., Dar, M. I., Raghib, F., Jaleel, H., Ahmad, B., Raina, A., et al. (2019). "Role and Regulation of Plants Phenolics in Abiotic Stress Tolerance," in *Plant Signaling Molecules* (Elsevier), 157–168. doi:10.1016/B978-0-12-816451-8.00009-5.

Naithani, S., Gupta, P., Preece, J., D'Eustachio, P., Elser, J. L., Garg, P., et al. (2019). Plant Reactome: a knowledgebase and resource for comparative pathway analysis. *Nucleic Acids Research*, gkz996. doi:10.1093/nar/gkz996.

Najjaa, H., Abdelkarim, B. A., Doria, E., Boubakri, A., Trabelsi, N., Falleh, H., et al. (2020). Phenolic composition of some Tunisian medicinal plants associated with anti-proliferative effect on human breast cancer MCF-7 cells. *The EuroBiotech Journal* 4, 104–112. doi:10.2478/ebtj-2020-0012.

Nawa, H., Kotani, H., and Nakanishi, S. (1984). Tissue-specific generation of two preprotachykinin mRNAs from one gene by alternative RNA splicing. *Nature* 312, 729–734. doi:10.1038/312729a0.

Nelder, J. A., and Wedderburn, R. W. M. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society. Series A (General)* 135, 370. doi:10.2307/2344614.

Neuner, G., and Hacker, J. (2012). "Ice Formation and Propagation in Alpine Plants," in *Plants in Alpine Regions: Cell Physiology of Adaption and Survival Strategies*, ed. C. Lütz (Vienna: Springer), 163–174. doi:10.1007/978-3-7091-0136-0_12.

Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E. J., Mathesius, U., et al. (2010). Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* 15, 684–692. doi:10.1016/j.tplants.2010.09.008.

Noctor, G., Mhamdi, A., Chaouch, S., Han, Y., Neukermans, J., Marquez-Garcia, B., et al. (2012). Glutathione in plants: an integrated overview: Glutathione status and functions. *Plant, Cell & Environment* 35, 454–484. doi:10.1111/j.1365-3040.2011.02400.x.

Noctor, G., Reichheld, J.-P., and Foyer, C. H. (2018). ROS-related redox regulation and signaling in plants. *Seminars in Cell & Developmental Biology* 80, 3–12. doi:10.1016/j.semcdb.2017.07.013.

Noiraud, N., Maurousset, L., and Lemoine, R. (2001). Transport of polyols in higher plants. *Plant Physiology and Biochemistry* 39, 717–728. doi:10.1016/S0981-9428(01)01292-X.

Nothias, L.-F., Petras, D., Schmid, R., Dührkop, K., Rainer, J., Sarvepalli, A., et al. (2020). Feature-based molecular networking in the GNPS analysis environment. *Nat Methods* 17, 905–908. doi:10.1038/s41592-020-0933-6.

Ohlrogge, J., and Browse, J. (1995). Lipid biosynthesis. *Plant Cell* 7, 957–970. doi:10.1105/tpc.7.7.957.

Oliver, D. J., Nikolau, B. J., and Wurtele, E. S. (2009). Acetyl-CoA—Life at the metabolic nexus. *Plant Science* 176, 597–601. doi:10.1016/j.plantsci.2009.02.005.

Pachauri, R. K., Mayer, L., (2015). *Climate change 2014: synthesis report*. Geneva, Switzerland: Intergovernmental Panel on Climate Change.

Padilla, F. M., and Pugnaire, F. I. (2006). The role of nurse plants in the restoration of degraded environments. *Frontiers in Ecology and the Environment* 4, 196–202. doi:10.1890/1540-9295(2006)004[0196:TRONPI]2.0.CO;2.

Padilla-González, G. F., Diazgranados, M., and Da Costa, F. B. (2017). Biogeography shaped the metabolome of the genus Espeletia: a phytochemical perspective on an Andean adaptive radiation. *Sci Rep* 7, 8835. doi:10.1038/s41598-017-09431-7.

Pamuru, R. R., Puli, C. O. R., Pandita, D., and Wani, S. H. (2021). "Sugar Alcohols and Osmotic Stress Adaptation in Plants," in *Compatible Solutes Engineering for Crop Plants Facing Climate Change*, eds. S. H. Wani, M. P. Gangola, and B. R. Ramadoss (Cham: Springer International Publishing), 189–203. doi:10.1007/978-3-030-80674-3_8.

Pandey, S., Fartyal, D., Agarwal, A., Shukla, T., James, D., Kaul, T., et al. (2017). Abiotic Stress Tolerance in Plants: Myriad Roles of Ascorbate Peroxidase. *Frontiers in Plant Science* 8. doi:10.3389/fpls.2017.00581.

Parida, A. K., Panda, A., and Rangani, J. (2018). "Metabolomics-Guided Elucidation of Abiotic Stress Tolerance Mechanisms in Plants," in *Plant Metabolites and Regulation Under Environmental Stress* (Elsevier), 89–131. doi:10.1016/B978-0-12-812689-9.00005-4.

Parihar, P., Singh, S., Singh, R., Singh, V. P., and Prasad, S. M. (2015). Effect of salinity stress on plants and its tolerance strategies: a review. *Environ Sci Pollut Res* 22, 4056–4075. doi:10.1007/s11356-014-3739-1.

Pearce, R. S. (2001). Plant Freezing and Damage. *Ann Bot* 87, 417–424. doi:10.1006/anbo.2000.1352.

Peters, K., Worrich, A., Weinhold, A., Alka, O., Balcke, G., Birkemeyer, C., et al. (2018). Current Challenges in Plant Eco-Metabolomics. *IJMS* 19, 1385. doi:10.3390/ijms19051385.

Peters, S., and Keller, F. (2009). Frost tolerance in excised leaves of the common bugle (*Ajuga reptans* L.) correlates positively with the concentrations of raffinose family oligosaccharides (RFOs). *Plant, Cell & Environment* 32, 1099–1107. doi:10.1111/j.1365-3040.2009.01991.x.

Plaxton, W. C. (1996). THE ORGANIZATION AND REGULATION OF PLANT GLYCOLYSIS. *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.* 47, 185–214. doi:10.1146/annurev.arplant.47.1.185.

Ploughe, L. W., Jacobs, E. M., Frank, G. S., Greenler, S. M., Smith, M. D., and Dukes, J. S. (2019). Community Response to Extreme Drought (CRED): a framework for drought-induced shifts in plant–plant interactions. *New Phytol* 222, 52–69. doi:10.1111/nph.15595.

Poorter, H., Jagodzinski, A. M., Ruiz-Peinado, R., Kuyah, S., Luo, Y., Oleksyn, J., et al. (2015). How does biomass distribution change with size and differ among species? An analysis for 1200 plant species from five continents. *New Phytol* 208, 736–749. doi:10.1111/nph.13571.

Poorter, H., Niinemets, Ü., Ntagkas, N., Siebenkäs, A., Mäenpää, M., Matsubara, S., et al. (2019). A meta-analysis of plant responses to light intensity for 70 traits ranging from molecules to whole plant performance. *New Phytol* 223, 1073–1105. doi:10.1111/nph.15754.

Poorter, H., Niinemets, Ü., Walter, A., Fiorani, F., and Schurr, U. (2010). A method to construct dose–response curves for a wide range of environmental factors and plant traits by means of a meta-analysis of phenotypic data. *Journal of Experimental Botany* 61, 2043–2055. doi:10.1093/jxb/erp358.

Poorter, H., Niklas, K. J., Reich, P. B., Oleksyn, J., Poot, P., and Mommer, L. (2012). Biomass allocation to leaves, stems and roots: meta-analyses of interspecific variation and environmental control: Tansley review. *New Phytologist* 193, 30–50. doi:10.1111/j.1469-8137.2011.03952.x.

Prevéy, J. S., Rixen, C., Rüger, N., Høye, T. T., Bjorkman, A. D., Myers-Smith, I. H., et al. (2019). Warming shortens flowering seasons of tundra plant communities. *Nat Ecol Evol* 3, 45–52. doi:10.1038/s41559-018-0745-6.

Quartacci, M. F. (2002). Plasma membrane lipids in the resurrection plant Ramonda serbica following dehydration and rehydration. *Journal of Experimental Botany* 53, 2159–2166. doi:10.1093/jxb/erf076.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing Available at: https://www.R-project.org/.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing Available at: https://www.R-project.org/.

Rai, A., Saito, K., and Yamazaki, M. (2017). Integrated omics analysis of specialized metabolism in medicinal plants. *Plant J* 90, 764–787. doi:10.1111/tpj.13485.

Ramel, F., Birtic, S., Ginies, C., Soubigou-Taconnat, L., Triantaphylides, C., and Havaux, M. (2012). Carotenoid oxidation products are stress signals that mediate gene responses to singlet oxygen in plants. *Proceedings of the National Academy of Sciences* 109, 5535–5540. doi:10.1073/pnas.1115982109.

Rascio, N., and Rocca, N. L. (2005). Resurrection Plants: The Puzzle of Surviving Extreme Vegetative Desiccation. *Critical Reviews in Plant Sciences* 24, 209–225. doi:10.1080/07352680591008583.

Rees, J. D., Ingle, R. A., and Smith, J. A. C. (2009). Relative contributions of nine genes in the pathway of histidine biosynthesis to the control of free histidine concentrations in *Arabidopsis thaliana*. *Plant Biotechnology Journal* 7, 499–511. doi:10.1111/j.1467-7652.2009.00419.x.

Reid, E. E., Thompson, P., Lyttle, C. R., and Dennis, D. T. (1977). Pyruvate Dehydrogenase Complex from Higher Plant Mitochondria and Proplastids. *Plant Physiol.* 59, 842–848. doi:10.1104/pp.59.5.842.

Rhodes, D., and Hanson, A. D. (1993). Quaternary Ammonium and Tertiary Sulfonium Compounds in Higher Plants. *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.* 44, 357–384. doi:10.1146/annurev.pp.44.060193.002041.

Riehl, S., Benz, M., Conard, N. J., Darabi, H., Deckers, K., Nashli, H. F., et al. (2012). Plant use in three Pre-Pottery Neolithic sites of the northern and eastern Fertile Crescent: a preliminary report. *Veget Hist Archaeobot* 21, 95–106. doi:10.1007/s00334-011-0318-y.

Rizhsky, L., Liang, H., and Mittler, R. (2002). The Combined Effect of Drought Stress and Heat Shock on Gene Expression in Tobacco. *Plant Physiology* 130, 1143–1151. doi:10.1104/pp.006858.

Roch, L., Prigent, S., Klose, H., Cakpo, C.-B., Beauvoit, B., Deborde, C., et al. (2020). Biomass composition explains fruit relative growth rate and discriminates climacteric from non-climacteric species. *Journal of Experimental Botany*, eraa302. doi:10.1093/jxb/eraa302.

Rochfort, S. (2005). Metabolomics Reviewed: A New "Omics" Platform Technology for Systems Biology and Implications for Natural Products Research. *J. Nat. Prod.* 68, 1813–1820. doi:10.1021/np050255w.

Rogowska, A., and Szakiel, A. (2020). The role of sterols in plant response to abiotic stress. *Phytochem Rev* 19, 1525–1538. doi:10.1007/s11101-020-09708-2.

Rohart, F., Gautier, B., Singh, A., and Lê Cao, K.-A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol* 13, e1005752. doi:10.1371/journal.pcbi.1005752.

Rohwer, J. M. (2012). Kinetic modelling of plant metabolic pathways. *Journal of Experimental Botany* 63, 2275–2292. doi:10.1093/jxb/ers080.

Roll-Hansen, N. (2014). The holist tradition in twentieth century genetics. Wilhelm Johannsen's genotype concept: Wilhelm Johannsen's genotype concept. *J Physiol* 592, 2431–2438. doi:10.1113/jphysiol.2014.272120.

Román-Figueroa, C., Bravo, L., Paneque, M., Navia, R., and Cea, M. (2021). Chemical products for crop protection against freezing stress: A review. *J Agro Crop Sci* 207, 391–403. doi:10.1111/jac.12489.

Rontein, D., Dieuaide-Noubhani, M., Dufourc, E. J., Raymond, P., and Rolin, D. (2002). The Metabolic Architecture of Plant Cells. *Journal of Biological Chemistry* 277, 43948–43960. doi:10.1074/jbc.M206366200.

Ros, R., Muñoz-Bertomeu, J., and Krueger, S. (2014). Serine in plants: biosynthesis, metabolism, and functions. *Trends in Plant Science* 19, 564–569. doi:10.1016/j.tplants.2014.06.003.

Ruiz-Perez, D., Guan, H., Madhivanan, P., Mathee, K., and Narasimhan, G. (2020). So you think you can PLS-DA? *BMC Bioinformatics* 21, 2. doi:10.1186/s12859-019-3310-7.

Saccenti, E., Hoefsloot, H. C. J., Smilde, A. K., Westerhuis, J. A., and Hendriks, M. M. W. B. (2014). Reflections on univariate and multivariate analysis of metabolomics data. *Metabolomics* 10, 361–374. doi:10.1007/s11306-013-0598-6.

Sachs, J. (1874). *Lehrbuch der botanik nach dem gegenwärtigen stand der wissenschaft. Bearb. von dr. Julius Sachs* ... Leipzig,: W. Engelmann, doi:10.5962/bhl.title.1332.

Saha, J., Brauer, E. K., Sengupta, A., Popescu, S. C., Gupta, K., and Gupta, B. (2015). Polyamines as redox homeostasis regulators during salt stress in plants. *Front. Environ. Sci.* 3. doi:10.3389/fenvs.2015.00021.

Sanchez, D. H., Pieckenstain, F. L., Escaray, F., Erban, A., Kraemer, U., Udvardi, M. K., et al. (2011). Comparative ionomics and metabolomics in extremophile and glycophytic Lotus species under salt stress challenge the metabolic pre-adaptation hypothesis. *Plant Cell Environ.* 34, 605–617. doi:10.1111/j.1365-3040.2010.02266.x.

Santos, M. A. Z., Colepicolo, P., Pupo, D., Fujii, M. T., de Pereira, C. M. P., and Mesko, M. F. (2017). Antarctic red macroalgae: a source of polyunsaturated fatty acids. *J Appl Phycol* 29, 759–767. doi:10.1007/s10811-016-1034-x.

Sardans, J., Gargallo-Garriga, A., Urban, O., Klem, K., Walker, T. W. N., Holub, P., et al. (2020). Ecometabolomics for a Better Understanding of Plant Responses and Acclimation to Abiotic Factors Linked to Global Change. *Metabolites* 10, 239. doi:10.3390/metabo10060239.

Sarker, S. D., and Nahar, L. (2017). "Progress in the Chemistry of Naturally Occurring Coumarins," in *Progress in the Chemistry of Organic Natural Products 106* Progress in the Chemistry of Organic Natural Products., eds. A. D. Kinghorn, H. Falk, S. Gibbons, and J. Kobayashi (Cham: Springer International Publishing), 241–304. doi:10.1007/978-3-319-59542-9_3.

Sato, R., Ito, H., and Tanaka, A. (2015). Chlorophyll b degradation by chlorophyll b reductase under high-light conditions. *Photosynth Res* 126, 249–259. doi:10.1007/s11120-015-0145-6.

Schaller, A., and Stintzi, A. (2009). Enzymes in jasmonate biosynthesis – Structure, function, regulation. *Phytochemistry* 70, 1532–1538. doi:10.1016/j.phytochem.2009.07.032.

Schertl, P., and Braun, H.-P. (2014). Respiratory electron transfer pathways in plant mitochondria. *Frontiers in Plant Science* 5. doi:10.3389/fpls.2014.00163.

Schläpfer, P., Zhang, P., Wang, C., Kim, T., Banf, M., Chae, L., et al. (2017). Genome-Wide Prediction of Metabolic Enzymes, Pathways, and Gene Clusters in Plants. *Plant Physiol.* 173, 2041–2059. doi:10.1104/pp.16.01942.

Schmidt, M. (2005). Least squares optimization with L1-norm regularization. *CS542B Project Report* 504, 195–221.

Schmitt, J. (2003). The Adaptive Evolution of Plasticity: Phytochrome-Mediated Shade Avoidance Responses. *Integrative and Comparative Biology* 43, 459–469. doi:10.1093/icb/43.3.459.

Schöb, C., Armas, C., Guler, M., Prieto, I., and Pugnaire, F. I. (2013). Variability in functional traits mediates plant interactions along stress gradients. *J Ecol* 101, 753–762. doi:10.1111/1365-2745.12062.

S**chöb, C., Michalet, R., Cavieres, L. A., Pugnaire, F. I., Brooker, R. W., Butterfield, B. J., et al.** (2014). A global analysis of bidirectional interactions in alpine plant communities shows facilitators experiencing strong reciprocal fitness costs. *New Phytol* 202, 95–105. doi:10.1111/nph.12641.

Schulz, E., Tohge, T., Zuther, E., Fernie, A. R., and Hincha, D. K. (2016). Flavonoids are determinants of freezing tolerance and cold acclimation in Arabidopsis thaliana. *Sci Rep* 6, 34027. doi:10.1038/srep34027.

Scossa, F., and Fernie, A. R. (2020). The evolution of metabolism: How to test evolutionary hypotheses at the genomic level. *Computational and Structural Biotechnology Journal* 18, 482–500. doi:10.1016/j.csbj.2020.02.009.

Seneviratne, S. I., Nicholls, N., Easterling, D., Goodess, C. M., Kanae, S., Kossin, J., et al. (2012). "Changes in Climate Extremes and their Impacts on the Natural Physical Environment," in *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, eds. C. B. Field, V. Barros, T. F. Stocker, and Q. Dahe (Cambridge: Cambridge University Press), 109–230. doi:10.1017/CBO9781139177245.006.

Shannon, P. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research* 13, 2498–2504. doi:10.1101/gr.1239303.

Sharma, B., Arora, S., Sahoo, D., and Deswal, R. (2020). Comparative fatty acid profiling of Indian seabuckthorn showed altitudinal gradient dependent species-specific variations. *Physiol Mol Biol Plants* 26, 41–49. doi:10.1007/s12298-019-00720-1.

Sherwood, S., and Fu, Q. (2014). A Drier Future? *Science* 343, 737–739. doi:10.1126/science.1247620.

Signori-Müller, C., Oliveira, R. S., Barros, F. de V., Tavares, J. V., Gilpin, M., Diniz, F. C., et al. (2021). Non-structural carbohydrates mediate seasonal water stress across Amazon forests. *Nat Commun* 12, 2310. doi:10.1038/s41467-021-22378-8.

Simmons, A. T., and Gurr, G. M. (2004). Trichome-based host plant resistance of Lycopersicon species and the biocontrol agent Mallada signata: are they compatible? *Entomologia Experimentalis et Applicata* 113, 95–101. doi:10.1111/j.0013-8703.2004.00210.x.

Singh, P., Arif, Y., Bajguz, A., and Hayat, S. (2021). The role of quercetin in plants. *Plant Physiology and Biochemistry* 166, 10–19. doi:10.1016/j.plaphy.2021.05.023.

Slama, I., Abdelly, C., Bouchereau, A., Flowers, T., and Savouré, A. (2015). Diversity, distribution and roles of osmoprotective compounds accumulated in halophytes under abiotic stress. *Annals of Botany* 115, 433–447. doi:10.1093/aob/mcu239.

Slocum, R. D. (2005). Genes, enzymes and regulation of arginine biosynthesis in plants. *Plant Physiology and Biochemistry* 43, 729–745. doi:10.1016/j.plaphy.2005.06.007.

Smirnoff, N. (2018). Ascorbic acid metabolism and functions: A comparison of plants and mammals. *Free Radical Biology and Medicine* 122, 116–129. doi:10.1016/j.freeradbiomed.2018.03.033.

Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006). XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Anal. Chem.* 78, 779–787. doi:10.1021/ac051437y.

Solomon, P. S., and Oliver, R. P. (2002). Evidence that γ-aminobutyric acid is a major nitrogen source during Cladosporium fulvum infection of tomato. *Planta* 214, 414–420. doi:10.1007/s004250100632.

Soussi, A., Ferjani, R., Marasco, R., Guesmi, A., Cherif, H., Rolli, E., et al. (2016). Plant-associated microbiomes in arid lands: diversity, ecology and biotechnological potential. *Plant Soil* 405, 357–370. doi:10.1007/s11104-015-2650-y.

Stark, G. R. (1977). Multifunctional proteins: one gene — more than one enzyme. *Trends in Biochemical Sciences* 2, 64–66. doi:10.1016/0968-0004(77)90133-5.

Stevanato, R., Fabris, S., and Momo, F. (2004). New Enzymatic Method for the Determination of Total Phenolic Content in Tea and Wine. *J. Agric. Food Chem.* 52, 6287–6293. doi:10.1021/jf049898s.

Stitt, M., Lilley, R. McC., Gerhardt, R., and Heldt, H. W. (1989). "[32] Metabolite levels in specific cells and subcellular compartments of plant leaves," in *Methods in Enzymology* (Elsevier), 518–552. doi:10.1016/0076-6879(89)74035-0.

Storey, R., Ahmad, N., and Wyn Jones, R. G. (1977). Taxonomic and ecological aspects of the distribution of glycinebetaine and related compounds in plants. *Oecologia* 27, 319–332. doi:10.1007/BF00345565.

Streb, P., Tel-Or, E., and Feierabend, J. (1997). Light stress effects and antioxidative protection in two desert plants. *Functional Ecology* 11, 416–424. doi:10.1046/j.1365-2435.1997.00105.x.

Strimbeck, G. R., Schaberg, P. G., Fossdal, C. G., Schröder, W. P., and Kjellsen, T. D. (2015). Extreme low temperature tolerance in woody plants. *Front. Plant Sci.* 6. doi:10.3389/fpls.2015.00884.

Stringlis, I. A., de Jonge, R., and Pieterse, C. M. J. (2019). The Age of Coumarins in Plant–Microbe Interactions. *Plant and Cell Physiology* 60, 1405–1419. doi:10.1093/pcp/pcz076.

Strizhov, N., Ábrahám, E., Ökrész, L., Blickling, S., Zilberstein, A., Schell, J., et al. (1997). Differential expression of two P5CS genes controlling proline accumulation during salt-stress requires ABA and is regulated by ABA1, ABI1 and AXR2 in Arabidopsis. *The Plant Journal* 12, 557–569. doi:10.1046/j.1365-313X.1997.00537.x.

Stroup, W. W. (2015). Rethinking the Analysis of Non-Normal Data in Plant and Soil Science. *Agron.j.* 107, 811–827. doi:10.2134/agronj2013.0342.

Stushnoff, C., Seufferheld, M. J., and Creegan, T. (1997). "Oligosaccharides as Endogenous Cryoprotectants in Woody Plants," in *Plant Cold Hardiness*, eds. P. H. Li and T. H. H. Chen (Boston, MA: Springer US), 301–309. doi:10.1007/978-1-4899-0277-1_27.

Sumner, L. W., Lei, Z., Nikolau, B. J., Saito, K., Roessner, U., and Trengove, R. (2014). Proposed quantitative and alphanumeric metabolite identification metrics. *Metabolomics* 10, 1047–1049. doi:10.1007/s11306-014-0739-6.

Sumner, L. W., Mendes, P., and Dixon, R. A. (2003). Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry* 62, 817–836. doi:10.1016/S0031-9422(02)00708-2.

Szabados, L., and Savouré, A. (2010). Proline: a multifunctional amino acid. *Trends Plant Sci.* 15, 89–97. doi:10.1016/j.tplants.2009.11.009.

Talbi, S., Romero-Puertas, M. C., Hernández, A., Terrón, L., Ferchichi, A., and Sandalio, L. M. (2015). Drought tolerance in a Saharian plant Oudneya africana: Role of antioxidant defences. *Environmental and Experimental Botany* 111, 114–126. doi:10.1016/j.envexpbot.2014.11.004.

Tetali, S. D. (2019). Terpenes and isoprenoids: a wealth of compounds for global use. *Planta* 249, 1–8. doi:10.1007/s00425-018-3056-x.

Thalmann, M., and Santelia, D. (2017). Starch as a determinant of plant fitness under abiotic stress. *New Phytol* 214, 943–951. doi:10.1111/nph.14491.

The Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* 408, 796–815. doi:10.1038/35048692.

Thevenet, D., Pastor, V., Baccelli, I., Balmer, A., Vallat, A., Neier, R., et al. (2017). The priming molecule β -aminobutyric acid is naturally present in plants and is induced by stress. *New Phytol* 213, 552–559. doi:10.1111/nph.14298.

Tholl, D. (2015). "Biosynthesis and Biological Functions of Terpenoids in Plants," in *Biotechnology of Isoprenoids* Advances in Biochemical Engineering/Biotechnology., eds. J. Schrader and J. Bohlmann (Cham: Springer International Publishing), 63–106. doi:10.1007/10_2014_295.

Tian, H., Lam, S. M., and Shui, G. (2016). Metabolomics, a Powerful Tool for Agricultural Research. *International Journal of Molecular Sciences* 17. doi:10.3390/ijms17111871.

Tipirdamaz, R., Gagneul, D., Duhazé, C., Aïnouche, A., Monnier, C., Özkum, D., et al. (2006). Clustering of halophytes from an inland salt marsh in Turkey according to their ability to accumulate sodium and nitrogenous osmolytes. *Environmental and Experimental Botany* 57, 139–153. doi:10.1016/j.envexpbot.2005.05.007.

Treutter, D. (2006). Significance of flavonoids in plant resistance: a review. *Environ Chem Lett* 4, 147–157. doi:10.1007/s10311-006-0068-8.

Trinchant, J.-C., Boscari, A., Spennato, G., Van de Sype, G., and Le Rudulier, D. (2004). Proline Betaine Accumulation and Metabolism in Alfalfa Plants under Sodium Chloride Stress. Exploring Its Compartmentalization in Nodules. *Plant Physiology* 135, 1583–1594. doi:10.1104/pp.103.037556.

Tshabuse, F., Farrant, J. M., Humbert, L., Moura, D., Rainteau, D., Espinasse, C., et al. (2018). Glycerolipid analysis during desiccation and recovery of the resurrection plant *Xerophyta humilis* (Bak) Dur and Schinz. *Plant Cell Environ* 41, 533–547. doi:10.1111/pce.13063.

Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., et al. (2015). MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat Methods* 12, 523–526. doi:10.1038/nmeth.3393.

Turner, N. C. (2018). Turgor maintenance by osmotic adjustment: 40 years of progress. *Journal of Experimental Botany* 69, 3223–3233. doi:10.1093/jxb/ery181.

Tzin, V., and Galili, G. (2010). New Insights into the Shikimate and Aromatic Amino Acids Biosynthesis Pathways in Plants. *Molecular Plant* 3, 956–972. doi:10.1093/mp/ssq048.

Upchurch, R. G. (2008). Fatty acid unsaturation, mobilization, and regulation in the response of plants to stress. *Biotechnol Lett* 30, 967–977. doi:10.1007/s10529-008-9639-z.

Urza, A. K., Weisberg, P. J., Chambers, J. C., and Sullivan, B. W. (2019). Shrub facilitation of tree establishment varies with ontogenetic stage across environmental gradients. *New Phytol* 223, 1795–1808. doi:10.1111/nph.15957.

v**an den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., and van der Werf, M. J.** (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* 7, 142. doi:10.1186/1471-2164-7-142.

van der Merwe, S., Greve, M., Olivier, B., and le Roux, P. C. (2021). Testing the role of functional trait expression in plant–plant facilitation. *Funct. Ecol.* 35, 255–265. doi:10.1111/1365-2435.13681.

Vélez-Bermúdez, I. C., and Schmidt, W. (2014). The conundrum of discordant protein and mRNA expression. Are plants special? *Front. Plant Sci.* 5. doi:10.3389/fpls.2014.00619.

Vetter, J. (2000). Plant cyanogenic glycosides. *Toxicon* 38, 11–36. doi:10.1016/S0041-0101(99)00128-2.

Vickery, H. Bradford., and Schmidt, C. L. A. (1931). The History of the Discovery of the Amino Acids. *Chem. Rev.* 9, 169–318. doi:10.1021/cr60033a001.

Voss-Fels, K. P., Stahl, A., Wittkop, B., Lichthardt, C., Nagler, S., Rose, T., et al. (2019). Breeding improves wheat productivity under contrasting agrochemical input levels. *Nat. Plants* 5, 706–714. doi:10.1038/s41477-019-0445-5.

Walker, T. W. N., Alexander, J. M., Allard, P., Baines, O., Baldy, V., Bardgett, R. D., et al. (2022). Functional Traits 2.0: The power of the metabolome for ecology. *Journal of Ecology* 110, 4–20. doi:10.1111/1365-2745.13826.

Wang, Y., An, Y., Yu, J., Zhou, Z., He, S., Ru, M., et al. (2016). Different responses of photosystem II and antioxidants to drought stress in two contrasting populations of Sour jujube from the Loess Plateau, China. *Ecol Res* 31, 761–775. doi:10.1007/s11284-016-1384-5.

Wei, T., and Simko, V. (2021). *R package "corrplot": Visualization of a Correlation Matrix*. Available at: https://github.com/taiyun/corrplot.

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

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York Available at: https://ggplot2.tidyverse.org.

Wieder, C., Frainay, C., Poupin, N., Rodríguez-Mier, P., Vinson, F., Cooke, J., et al. (2021). Pathway analysis in metabolomics: pitfalls and best practice for the use of over-representation analysis. Bioinformatics doi:10.1101/2021.05.24.445406.

Wildi, B., and Lutz, C. (1996). Antioxidant composition of selected high alpine plant species from different altitudes. *Plant Cell Environ* 19, 138–146. doi:10.1111/j.1365-3040.1996.tb00235.x.

Williamson, J. D., Jennings, D. B., Guo, W.-W., Pharr, D. M., and Ehrenshaft, M. (2002). Sugar Alcohols, Salt Stress, and Fungal Resistance: Polyols—Multifunctional Plant Protection? *jashs* 127, 467–473. doi:10.21273/JASHS.127.4.467.

Wimalasekera, R. (2019). "Effect of Light Intensity on Photosynthesis," in *Photosynthesis, Productivity and Environmental Stress*, eds. P. Ahmad, M. Abass Ahanger, M. Nasser Alyemeni, and P. Alam (Wiley), 65–73. doi:10.1002/9781119501800.ch4.

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., et al. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Research* 46, D608–D617. doi:10.1093/nar/gkx1089.

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., et al. (2004). The worldwide leaf economics spectrum. *Nature* 428, 821–827. doi:10.1038/nature02403.

Xia, J., Sinelnikov, I. V., Han, B., and Wishart, D. S. (2015). MetaboAnalyst 3.0--making metabolomics more meaningful. *Nucleic Acids Res* 43, W251-257. doi:10.1093/nar/gkv380.

Xu, S., Xu, Y., Gong, L., and Zhang, Q. (2016). Metabolomic prediction of yield in hybrid rice. *The Plant Journal* 88, 219–227. doi:10.1111/tpj.13242.

Xu, Z., Ma, J., Qu, C., Hu, Y., Hao, B., Sun, Y., et al. (2017). Identification and expression analyses of the alanine aminotransferase (AlaAT) gene family in poplar seedlings. *Sci Rep* 7, 45933. doi:10.1038/srep45933.

Xue, J., Guijas, C., Benton, H. P., Warth, B., and Siuzdak, G. (2020). METLIN MS2 molecular standards database: a broad chemical and biological resource. *Nat Methods* 17, 953–954. doi:10.1038/s41592-020-0942-5.

Yan, W., and Hunt (1999). An Equation for Modelling the Temperature Response of Plants using only the Cardinal Temperatures. *Annals of Botany* 84, 607–614. doi:10.1006/anbo.1999.0955.

Yang, T., Li, H., Tai, Y., Dong, C., Cheng, X., Xia, E., et al. (2020). Transcriptional regulation of amino acid metabolism in response to nitrogen deficiency and nitrogen forms in tea plant root (Camellia sinensis L.). *Sci Rep* 10, 6868. doi:10.1038/s41598-020-63835-6.

Yang, Y., Chen, J.-G., Schöb, C., and Sun, H. (2017). Size-Mediated Interaction between a Cushion Species and Other Non-cushion Species at High Elevations of the Hengduan Mountains, SW China. *Front. Plant Sci.* 08. doi:10.3389/fpls.2017.00465.

Yano, R., Nakamura, M., Yoneyama, T., and Nishida, I. (2005). Starch-Related α -Glucan/Water Dikinase Is Involved in the Cold-Induced Development of Freezing Tolerance in Arabidopsis. *Plant Physiology* 138, 837–846. doi:10.1104/pp.104.056374.

Yobi, A., Wone, B. W. M., Xu, W., Alexander, D. C., Guo, L., Ryals, J. A., et al. (2013). Metabolomic profiling in Selaginella lepidophylla at various hydration states provides new insights into the mechanistic basis of desiccation tolerance. *Mol Plant* 6, 369–385. doi:10.1093/mp/sss155.

Yolcu, S., Alavilli, H., and Lee, B. (2020). Natural Genetic Resources from Diverse Plants to Improve Abiotic Stress Tolerance in Plants. *IJMS* 21, 8567. doi:10.3390/ijms21228567.

Young, A., and Lowe, G. (2018). Carotenoids—Antioxidant Properties. *Antioxidants* 7, 28. doi:10.3390/antiox7020028.

Yu, P., He, X., Baer, M., Beirinckx, S., Tian, T., Moya, Y. A. T., et al. (2021). Plant flavones enrich rhizosphere Oxalobacteraceae to improve maize performance under nitrogen deprivation. *Nat. Plants* 7, 481–499. doi:10.1038/s41477-021-00897-y.

Zapata, P. J., Serrano, M., Pretel, M. T., Amorós, A., and Botella, M. Á. (2004). Polyamines and ethylene changes during germination of different plant species under salinity. *Plant Science* 167, 781–788. doi:10.1016/j.plantsci.2004.05.014.

Zhang, A., Sun, H., Wang, P., Han, Y., and Wang, X. (2012). Modern analytical techniques in metabolomics analysis. *Analyst* 137, 293–300. doi:10.1039/C1AN15605E.

Zhang, T., Wang, Z., Lv, X., Dang, H., and Zhuang, L. (2020). Variation of rhizosphere bacterial community diversity in the desert ephemeral plant Ferula sinkiangensis across environmental gradients. *Sci Rep* 10, 18442. doi:10.1038/s41598-020-75591-8.

Zhang, X., Gu, S., Zhao, X., Cui, X., Zhao, L., Xu, S., et al. (2010). Radiation partitioning and its relation to environmental factors above a meadow ecosystem on the Qinghai-Tibetan Plateau. *J. Geophys. Res.* 115, D10106. doi:10.1029/2009JD012373.

Zhao, H.-L., Zhou, R.-L., Su, Y.-Z., Zhang, H., Zhao, L.-Y., and Drake, S. (2007). Shrub facilitation of desert land restoration in the Horqin Sand Land of Inner Mongolia. *Ecological Engineering* 31, 1–8. doi:10.1016/j.ecoleng.2007.04.010.

Zhou, L., Ni, E., Yang, J., Zhou, H., Liang, H., Li, J., et al. (2013). Rice OsGL1-6 Is Involved in Leaf Cuticular Wax Accumulation and Drought Resistance. *PLoS ONE* 8, e65139. doi:10.1371/journal.pone.0065139.

Zhu, H., Li, C., and Gao, C. (2020). Applications of CRISPR–Cas in agriculture and plant biotechnology. *Nat Rev Mol Cell Biol* 21, 661–677. doi:10.1038/s41580-020-00288-9.

Ziaco, E., Truettner, C., Biondi, F., and Bullock, S. (2018). Moisture-driven xylogenesis in *Pinus ponderosa* from a Mojave Desert mountain reveals high phenological plasticity: Moisture-driven xylogenesis in *Pinus ponderosa*. *Plant Cell Environ* 41, 823–836. doi:10.1111/pce.13152.

Zidorn, C. (2010). Altitudinal variation of secondary metabolites in flowering heads of the Asteraceae: trends and causes. *Phytochem Rev* 9, 197–203. doi:10.1007/s11101-009-9143-7.

Zotz, G., and Winkler, U. (2013). Aerial roots of epiphytic orchids: the velamen radicum and its role in water and nutrient uptake. *Oecologia* 171, 733–741. doi:10.1007/s00442-012-2575-6.

Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. *J Royal Statistical Soc B* 67, 301–320. doi:10.1111/j.1467-9868.2005.00503.x.

ANNEX

ANNEX I.A1

Annual Plant Reviews (2021) **4**, 551–604
doi: 10.1002/9781119312994.apr0758

http://onlinelibrary.wiley.com

ANOTHER TALE FROM THE HARSH WORLD: HOW PLANTS ADAPT TO EXTREME ENVIRONMENTS

Thomas Dussarrat[1,2], Guillaume Decros[2], Francisca P. Díaz[1], Yves Gibon[2,3], Claudio Latorre[4,5], Dominique Rolin[2,3], Rodrigo A. Gutiérrez[1] and Pierre Pétriacq[2,3]

[1] *Departamento de Genética Molecular y Microbiología, Pontificia Universidad Católica de Chile, FONDAP Center for Genome Regulation and ANID–Millennium Science Initiative Program-Millennium Inst. for Integrative Biology (iBio), Santiago, Chile*
[2] *Univ. Bordeaux, INRAE, Villenave d'Ornon, France*
[3] *Bordeaux Metabolome, Villenave d'Ornon, France*
[4] *Departamento de Ecología, Pontificia Universidad Católica de Chile, Santiago, Chile*
[5] *Institute of Ecology and Biodiversity (IEB), Santiago, Chile*

Abstract: The environmental fluctuations of a constantly evolving world can mould a changing context, often unfavourable to sessile organisms that must adjust their resource allocation between both resistance or tolerance mechanisms and growth. Plants bear the fascinating ability to survive and thrive under extreme conditions, a capacity that has always attracted the curiosity of humans, who have discovered and improved species capable of meeting our physiological needs. In this context, plant research has produced a great wealth of knowledge on the responses of plants to a range of abiotic stresses, mostly considering model species and/or controlled conditions. However, there is still minimal comprehension of plant adaptations and acclimations to extreme environments, which cries out for future investigations. In this article, we examined the main advances in understanding the adapted traits fixed through evolution that allowed for plant resistance against abiotic stress in extreme natural ecosystems. Spatio-temporal adaptations from extremophile plant species are described from morpho-anatomical features to physiological function and metabolic pathways adjustments. Considering that metabolism is at the heart of plant adaptations,

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

a focus is given to the study of primary and secondary metabolic adjustments as well as redox metabolism under extreme conditions. This article further casts a critical glance at the main successes in studying extreme environments and examines some of the challenges and opportunities this research offers, especially considering the possible interaction with ecology and metaphenomics.

Keywords: extremophile, abiotic stress, adaptation, extreme environment, metaphenomics, metabolism, redox

1 Introduction

In a rapidly evolving world, environmental fluctuations create a changing context, often unfavourable to sessile organisms that must adjust their resource allocation between both resistance mechanisms and growth to thrive under stressful conditions. The stress concept was classically defined as any environmental factor capable of inducing a potentially injurious strain in living organisms (Selye, 1950; Levitt, 1980). Abiotic environmental stresses such as water deficit, temperature, excess light radiations, or soil depletion and contamination, negatively impact plant performances and likely become more prevalent and intense (Battisti and Naylor, 2009; Fedoroff et al., 2010). Indeed, the toxicological concept of '*Only the dose makes the poison*' (Paracelsus), could be applied when characterising soil suitability in which either water, mineral, and nutrient deficiency like aridity (Mishra and Singh, 2010), nitrogen (N), and phosphorous (P) starvation (Vance, 2001), or excess like salinity (Munns, 2002; Shahid et al., 2018), metal contamination (Nagajyoti et al., 2010; Tóth et al., 2016) and flooding (Loreti et al., 2016) are adverse for plants. Moreover, resistance to those stressors is further complicated when associated with low and high temperatures (Pearce, 2001; Neuner and Hacker, 2012; Qu et al., 2013) or intense UV radiations (Caldwell et al., 1989; Zandalinas et al., 2018), all of which can create conditions that challenge plant survival in such stressful constraints (Mittler, 2006).

Interestingly, some lands naturally harbour these stress combinations where both their intensity and frequency levels have led to the name of 'extreme environments'. Over the last few decades, extreme environments have been continuously discovered on Earth and permanently challenge our understanding of the limits at which life can exist. In terms of microbial life, these ecosystems include hot springs, glacial ice, and deep-sea vents (Merino et al., 2019). Plants do not colonise these environments, yet they can also grow and even thrive in areas where water is scarce, solar radiations are extremely high, temperatures can reach extremes of cold and/or heat, nutrients are almost completely absent, and salts accumulate (Rothschild and Mancinelli, 2001). Besides, extreme environments have been mostly characterised by specific environmental variables such as water resources, salinity, temperatures, and solar irradiance.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Water is of major limitation for plant growth and survival (Farooq et al., 2012). The aridity index, which refers to the net difference between precipitations and water losses through evapotranspiration, was used to characterise drylands (Levin et al., 2006; Girvetz and Zganjar, 2014). Hyperarid and arid zones present an aridity index below 0.03 and 0.03–0.20 with annual rainfall rarely exceeding 100 and 250 mm, respectively (Noy-Meir, 1973). Also, extremes usually involve scarce precipitations mainly occurring during the rainy season (Noy-Meir, 1973). Thus, this article refers to 'extreme aridity' lands where annual rainfalls are highly variable, with a large unpredictable component, and not exceeding 250 mm $year^{-1}$. Remarkably, several arid soils are also highly saline where halophytes have succeeded to survive (Allbed and Kumar, 2013). Halophytes have evolved from a wide range of plant families (Flowers et al., 2010) in several lands as deserts or sea coasts, and are defined as plants able to complete the life cycle in a salt concentration of at least 200 mM NaCl and constitute 1% of the world's flora (Flowers and Colmer, 2008).

Furthermore, drought and saline stress are usually combined with other parameters like the temperature in natural environments (Mittler, 2006). In most vascular plants, cold stress likely occurs once the subfreezing air temperature falls between −0.6 and −2.6 °C, where ice nucleation happens and causes extracellular freezing, resulting in cellular dehydration and plant death (Pearce, 2001; Neuner and Hacker, 2012; Hasanuzzaman et al., 2013). Adding the time parameter (Körner, 2016), extremely cold environments here referred to terrestrial areas with daily negative temperatures (e.g. polar circle, elevations above the climatic treeline, lands with long-term cold periods). Similarly, while the upper threshold temperature depends on plant species, an average maximum temperature of 35–40 °C was highlighted in crops, forage, rangeland, and wild species (Wahid et al., 2007; Hasanuzzaman et al., 2013; Dürr et al., 2015). Consequently, extreme heat stress would be considered as long-term 40 °C conditions, above which most plants are out of their optimal temperature ranges.

In addition, extremophile plants able to inhabit ecosystems like high mountains or deserts are facing low soil suitability (Larcher et al., 2010; Sun and Wang, 2016). This is particularly true for mineral nutrients such as N or phosphate for which soil deficiencies could be observed in natural conditions. An outstanding example is given with the Atacama desert in South America and South-western Australia lands characterised by extremely low N and phosphate soil availability, respectively (Lambers et al., 2011, 2013; Díaz et al., 2016). Nevertheless, these very harsh environmental conditions are rather rare as human activities have modelled most nutrient contents in the soil to suit the need for agricultural practices.

In stark contrast, light is not a limiting factor in desert or mountain ecosystems where the irradiance can even reach extreme levels (Piacentini et al., 2003). Further, the daily light integral (between 1 and 50 mol m^{-2} d^{-1}),

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

which is the photosynthetic photon flux density (400–700 nm) integrated over a day, has shown a positive relationship with multiple phenotypic traits (Poorter et al., 2019). However, expected ozone reduction could increase solar irradiance (Williamson et al., 2014) including (i) UV-B (280–315 nm) radiations that disturb plant performance directly by causing DNA and membrane damage or indirectly by increasing mutation occurrences (Zlatev et al., 2012) and (ii) UV-A (315–400 nm), which impact both photosynthetic level and plant growth (Verdaguer et al., 2017). A meta-analysis of the responses of woody and herbaceous plants from different latitudes to 40% elevated ambient UV-B radiations have shown deleterious effects on several phenotypic traits (Li et al., 2010).

Finally, while the previously mentioned abiotic variables naturally occur at extreme levels, human activity has led to the development of extreme lands where plants should harbour adaptive traits to cope with high heavy metal concentrations for instance. Soil contaminations by the overaccumulation of essential micronutrients [e.g. copper (Cu), cobalt (Co), manganese (Mn), molybdenum (Mo), and zinc (Zn)] or highly toxic elements [e.g. arsenic (As), cadmium (Cd), chrome (Cr), nickel (Ni), and lead (Pb)] are increasingly common and lead to diverse toxic effects on plants (Nagajyoti et al., 2010; Tóth et al., 2016). Metallophytes are plants capable of growing in metal-contaminated soils while the great majority of plants cannot (Antonovics et al., 1971). They represent 500 plant species among which most are obligate metallophytes (Pollard et al., 2014). Recently, a special interest has grown on a subgroup of metallophytes known as hyperaccumulators, which can survive while concentrating metals in shoots. Concentration thresholds in dry weight foliar tissue to define a plant as hyperaccumulator were nicely described for both essential micronutrients (e.g. Co and Cu > 300 $\mu g\, g^{-1}$; Mo > 1500 $\mu g\, g^{-1}$; Zn > 3000 $\mu g\, g^{-1}$; Mn > 10 000 $\mu g\, g^{-1}$) and highly toxic elements (e.g. Cd > 100 $\mu g\, g^{-1}$; Cr, Ni, Pb and As > 1000 $\mu g\, g^{-1}$) (Krämer, 2010; Boojar and Tavakkoli, 2011; Singh et al., 2013; Peng et al., 2020).

Altogether these observations suggest that an environmental variable could be defined as 'extreme' when its intensity lies at the edges of its organism-specific, life-compatible gradient, for a duration or periodicity high enough to allow extremely few species to survive and thrive in the ecosystem to which this variable belongs. Any environment lying at the edges (or 'extremes') of any abiotic gradient generates enough stress that would kill most organisms (Rothschild and Mancinelli, 2001). Consequently, adapted organisms able to thrive under such conditions are termed 'lovers' (*phílos* in Ancient Greek) of 'extremes' (i.e. extremophile). Accordingly, the adaptation of an organism to its environment would refer to any heritable change in genotype that improves survival and physiological activity at any level of organisation in response to an environmental variation, whereas acclimation would refer to plasticity, which involves both gene expression

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

and molecular mechanism variations in response to a change in the environment (Borowitzka, 2018). Nicotra et al. (2010) further stated that phenotypic plasticity was a feature that could be considered an adaptive trait if present over a long-term period. Besides, the adaptive traits of extremophile plant species positively influence the plant performance by increasing its capacity to survive, generate biomass, and even reproduce (Fernandez et al., 2016; Fridley, 2017). Indeed, the trade-off between survival and growth required under harsh environmental conditions are driven by specific adaptations on behalf of any organism to survive (Rothschild and Mancinelli, 2001).

Hitherto, several reviews have nicely summarised the progress into the analysis of the plant responses to abiotic stress (Bundy et al., 2008; Nakabayashi and Saito, 2015; Bowne et al., 2018) and biotic factors (Suzuki et al., 2014; Mhlongo et al., 2018; Tugizimana et al., 2018). Here, this review tackled the main advances in the understanding of the adapted traits fixed through evolution that allow for plant resistance against abiotic stress in extreme ecosystems. Spatio-temporal adaptations from extremophile plant species are described from morpho-anatomical features to physiological function and metabolic pathway adjustments, considering that plant resistance can be reached either by avoiding the intrusion or by tolerating its entry without impacting the plant (Levitt, 1980). Considering that metabolism is at the heart of plant adaptations, we particularly focus on the roles of both primary and secondary metabolic pathways, as well as the adjustments of redox metabolism under extreme conditions. Finally, we cast a critical glance at the main successes in studying extreme environments and examine some of the challenges and opportunities that this research offers, especially considering the possible interaction with ecology and phenomics.

2 Life Cycle and Morphological Adaptations, a Critical Step to Jump the Hurdles of Extreme Environments

Plant adaptation arises from changes in morphology and life cycle, and both are the product of development and evolution. Moreover, morphology is central to plant taxonomy and used to infer phylogeny. Nevertheless, the organisation and evolution of molecular diversity and phenotype are also structured and correlated with abiotic environmental heterogeneity and stress (Nevo, 2001).

Even though harsh environments present relatively low total species biodiversity compared to mild physical conditions (Brown, 1990), extreme or stressful environments tend to display high genetic diversity and phenotypic adaptations (Nevo, 2001).

This section seeks to summarise and contrast some examples of life cycle and morphology adaptations to survive in extreme environments and is by

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

no means an exhaustive list of all possible adaptations. Instead, it reflects some life cycle adaptations from a temporal avoidance perspective and some structural adaptations to tolerate or avoid stresses related to water resources, salinity, temperature, and radiation.

2.1 Temporal Avoidance Adaptations

Apart from molecular, physiological, and morphological adaptations, plant survival in harsh environments also relies on seasonal responses to variable environmental conditions. In this context, the temporal response to abiotic stress represents an escape or avoidance strategy with which plant life cycles synchronise with seasonal variations. On the other hand, seedling germination, establishment, and initial development are the most vulnerable and challenging aspects in a plant's life, imposing a vital population bottleneck. Finding the right time and space to germinate is crucial to avoid non-favourable growing seasons.

2.1.1 Short Life-forms

Plants (tissues or organs) may not need to be adapted to tolerate every season or the harshest environmental conditions. Geophytes (a perennial herb that propagates from an underground organ) or therophytes (short-life annual herbs), for example, are very dominant in deserts and steppes (Vidiella et al., 1999). These plants need effective mechanisms to sense environmental parameters, undergo rapid growth and development strategies to complete their life-cycle in a short period, and the ability of seeds or underground organs to survive during non-favourable seasons. Reproduction is also an essential trait under temporal regulation in extreme environments. Annual plants need to complete their life cycle in a short time period and, in some flood habitats, need to accelerate flowering during unpredictable short dry periods (Blom et al., 1990). In the winter-rainfall Succulent Karoo of South Africa, some perennials flower in autumn to avoid pollination competition with annuals during spring (Cowling et al., 1999).

2.1.2 Seed Adaptations

Seed dormancy is a common strategy to survive often unpredictable prolonged dry or cold periods. Another key strategy to increase seedling survival, however, is to spread germination over different time periods during transient favourable circumstances, but with overall unpredictable future conditions (Cohen, 1966; Gutterman, 2000). In some extreme cases, dispersal via seeds is not the most successful strategy, and some plants favour germination on the parent and disperse live seedlings or propagules (vivipary and pseudovivipary). Many plants use these germination strategies in salty shallow marine environments, such as mangroves or seagrasses, allowing some development (shoots, roots, floating systems) before being

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

adrift in salty and hypoxic water (Tomlinson and Cox, 2000; Alleman and Hester, 2011). This strategy is also used in arid environments. The epiphyte *Tillandsia recurvata* in the Southern Chihuahuan Desert of North America presents 'true-vivipary' as roots of seedlings may attach to hosts more readily than seeds without roots (Pérez-Noyola et al., 2020). The epiarenitic *Tillandsia landbeckii* survive exclusively off fog along the hyperarid coast of the Atacama Desert (Rundel et al., 1997) and produce asexual propagules that are released to the environment as small, fully developed, clones of the mother plants (pseudovivipary). This strategy allows them to quickly colonise favourable environments as fog water input fluctuates over time (Latorre et al., 2011).

2.2 Structural Tolerance and Avoidance Adaptations

Many extreme environments can be further defined by marked seasonality or by the absence or excess of one or more essential resources such as nutrients. Perennial plants that grow in extreme seasonal environments cannot escape temporally and therefore have evolved structural and morphological mechanisms to tolerate such stress. Rooting systems are adapted to use different strategies to cope with environmental stress. Leaf size and shape along with the density of stomata, epicuticular wax development, and trichomes are all influenced by environmental factors and are key components in determining the leaf boundary layer (Bickford, 2016), which has a significant impact on transpiration, leaf water status, and stomatal behaviour (Hill et al., 2015; Matthews and Lawson, 2018). Some of these adaptive mechanisms are constitutive (e.g. breathing roots, succulence, salt glands) and others are induced (e.g. supercooling, stomata control) responses that allow plants to maintain homeostasis despite environmental fluctuations.

2.2.1 Water Resources

In deserts, depending on the amount, seasonality, and predictability of rainfall, xerophytic plants develop diverse strategies to survive in arid environments. One root system, already mentioned with temporal scape strategies, is the underground perennating rootstocks typical of geophytes. In contrast, phreatophytic species use deep tap roots to reach the groundwater table and other desert species grow branched horizontal and shallow roots to collect surface water as primary sources. These root system architectural adaptations even modify community attributes. For instance, some desert perennials have extremely shallow (0.1–0.2 m) roots, making them vulnerable to rare episodes of lower-than-average rainfall, generating entire communities of relatively short-lived (5–10 years) shrubs (Von Willert et al., 1985; Cowling et al., 1999).

Leaf and stem aerial plant tissues have evolved diverse morphologies and structures to tolerate aridity, from dwarfism with a low leaf index to

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

waxy skin, hairs, and thorns to reduce water loss and to reflect heat. Some perennial plants even lose their leaves or entire shoots as conditions become drier. One of the better known structural adaptations is succulent leaves or stems, common in dehydration or salt stress environments. Succulents are plants with fleshy stems (e.g. Cactaceae, members of Euphorbiaceae) or leaves (e.g. species of the Aizoaceae, Crassulaceae, or Portulacaceae) that store more water than required for immediate metabolic needs (Cowling et al., 1999). Stem succulence illustrates the convergent evolution of functional adaptations in morphology. The external similarities contrast with a variable internal architecture, including the participation of different stem tissues in water storage (Eggli and Nyffeler, 2009).

Stomata control the uptake of CO_2 into the leaf along with water loss through transpiration and are critical for maintaining plant water balance and leaf temperature. Stomata size (guard cell length) and density determine their conductivity, but these structures adjust to the environment to balance the requirement for CO_2 entry against leaf dehydration. When water is limited, stomata modify their position and can be sunken and concentrated below the leaf surface or protected in leaves pressed into the stems (Sundberg, 1986; Dong and Zhang, 2000). Small stomata structures or densities will avoid water loss, but in many deserts, temperatures are high, so the increased leaf temperature as a result of stomatal closure can negatively impact plants (Matthews and Lawson, 2018). Xerophytic plants have evolved different strategies to deal with water stress, including increasing or decreasing their stomatal density depending on their particular environmental stresses, photosynthesis pathways (discussed below), and life-forms (e.g. relatively large, infrequent stomata on the surface of succulents) (Sundberg, 1986; Dong and Zhang, 2000).

In contrast to aridity, aquatic and flooded soil represent highly water-saturated soils that exclude oxygen, one of the fundamental requirements for plant life. Several anatomic adaptations facilitate gas transport in flooded soils, such as developing aerenchyma and gas-tight barriers in the epidermis and exodermis in roots decreasing radial oxygen losses (Aschi-Smiti et al., 2003). In mangroves, massive root systems appeared. Some plant species, such as *Avicennia* or *Sonneratia*, develop pneumatophores or breathing roots to obtain the scarce oxygen present in the mud (Scholander et al., 1955; Purnobasuki, 2011). Another structural root adaptation in mangroves is stilt roots that diverge from stems and branches and penetrate the soil away from the main stem to increase stability.

2.2.2 Salinity

Halophytes are adapted to live in salty environments, and they need to develop methods for salt exclusion or excretion and water conservation. The salinity tolerance of a species varies according to the plant's developmental status (Ball, 1988). Most halophyte root adaptations are physiological or

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

metabolic, but root structure also changes. Some mangrove root systems are isolated from the external solution by a hydrophobic barrier to apoplastic transport in the periderm and exodermis, restricting the access of salty water to a small distal proportion of the root periphery (Moon et al., 1986; Krishnamurthy et al., 2014).

Salt secretors or secretohalophytes are salt-tolerant plants that excrete excess salts through specialised glands on their leaf surfaces. This secretion is found in more than 50 species in 14 families and can be grouped into four structural classes sharing convergently evolved features that compartmentalise and excrete salt (Dassanayake and Larkin, 2017). Many grow in mangroves, but others are common to deserts including *Atriplex* species, where salt is excreted from salt glands into the central vacuole of the bladder cell (Fahn, 1988).

Arid soils are often highly saline. It is suggested that salinity-induced reduction in stomatal density represents a fundamental mechanism by which plants optimise water use efficiency under saline conditions. In the halophyte desertic plant *Chenopodium quinoa*, salt-grown plants showed a significant (approximately 30%) reduction in stomatal density observed in all leaves (Shabala et al., 2012).

2.2.3 Temperature

Temperature is one of the critical drivers of leaf size and shape worldwide, generating giant leaves that have fewer, smaller teeth in tropical plants, and tiny ones with more numerous teeth in deserts (Peppe et al., 2011). Aridity limits leaf size as the risk of overheating during daytime maximum temperatures increases as they grow larger. In contrast, freezing risks limits leaf size in wetter climates, especially freezing at night (Wright et al., 2017).

In high altitude and/or high latitude environments, leaves and stems tolerate or avoid freezing temperatures and potential ice formation within their tissues. Many plants survive to freeze through a process of cold acclimation. In freezing-tolerant tissues, the construction of extracellular ice by loss of cellular water lowers the cell's freezing point by increasing solute concentration as well as by producing osmolytes. In freezing-avoidance, tissues survive by supercooling, keeping cell solutions between the equilibrium freezing point and the homogeneous ice nucleation temperature of water (usually between −1 and −41 °C) (Wisniewski and Fuller, 1999). Freezing-tolerant species such as *Calluna vulgaris* have developed ice barriers tissue at the base of the pedicel, where the anatomical features of the pit membrane that are likely impermeable to ice and the presence of hydrophobic substances such as lignin, suberin, and cutin present in cell walls, represent a critical constriction for ice propagation into supercooled tissues (Kuprian et al., 2016).

In hot environments, plants can reduce their internal temperature using different structural adaptations, such as hairs and thorns to reduce overheating. Stomata control and an effective cuticular barrier are key plant

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

adaptations for thermal control. The plant cuticle consists of a matrix of polymeric cutin with cuticular waxes embedded or deposited in leaf surfaces and act as a transpiration barrier at elevated temperatures (Schuster et al., 2016). This plays a vital role in hot-desert plants, where it is imperative to reduce transpiration while lowering overall temperature. Plant cuticle layers are especially adapted in these habitats and show lower permeability than in non-desert habitats and more resistance to temperature increases (Schuster et al., 2016).

2.2.4 Radiation

The morphology of shoots and leaves is a key aspect for light capture and can be modified by ultraviolet (UV) radiation changes. When plants are exposed to excess UV-B radiation, overall biomass, and even root system morphology can exhibit sizable declines (Caldwell et al., 2007). In addition, UV-B has inhibitory effects on stem length, leaf size and leaf anatomy (Verdaguer et al., 2017). UV-A also has morphological effects, especially on leaf size and foliage.

The epicuticular wax and trichomes covering plant surfaces also play important roles in protecting plants against UV radiation. The dense trichomes often covering young leaves, in addition to other functions, transiently protect the underlying cells against UV-B radiation damage while other internal protective mechanisms develop (Karabourniotis et al., 1995).

3 Modulations of Plant Major Physiological Functions in Extreme Conditions

The evolution of plant phenotype and lifespan in extreme ecosystems requires the adaptation of some major physiological functions. In adequation to the life cycle and plant morphology features, these physiological functions allow for the modulation of both the uptake and transport of mineral elements (e.g. transport of salt from roots to glands in secretohalophytes) and the management of the allocation of C resources between resistance or growth mechanisms in response to abiotic perturbation, for instance. This section aims at summarising some of the physiological adaptations that extremophile plants harbour to face water scarcity, high solar irradiance, and low soil suitability.

3.1 Extremophile Plants Facing Water Scarcity

In extreme environments, the plant water status is indirectly dependent on precipitation levels and directly on their ability to draw water from the soil or fog deposition, and limit losses due to evaporative mechanisms

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

(Martin and von Willert, 2000; Xu et al., 2011; Schuster et al., 2016). Physical and chemical soil properties represent a major environmental variable that greatly participates in defining water resources (Hamblin, 1986) and therefore plant water status. The structure and texture of the soil govern the drying process during water shortage periods that affect plant growth in two ways (Zou et al., 2010). First, the dehydration state of the soil surface layers limits the microbial activity and the recycling of mineral nutrients, which therefore impact plant performance (Sardans and Peñuelas, 2005; Schimel, 2018). In contrast, long-term aridity leads to direct drought stress, known for its disastrous effects on plant survival and development (Schulze et al., 1980). Remarkably, assemblages of dry grasses, dwarf shrubs, and cushions or Alpine semi-desert plants, naturally thrive in some of the world's driest regions found in the Pamir desert between 3500 and 4500 m.a.s.l or in the Atacama desert (Pyankov et al., 1999; Díaz et al., 2016, 2019), where the possibility of surviving is not limited to 'extreme' plant families like Cactaceae or Boraginaceae, but extended in genera from crop families including Poaceae (e.g. *Jarava, Puccinellia*), Asteraceae (e.g. *Baccharis, Parastrephia*) and Fabaceae (e.g. *Lupinus, Oxytropis*). These ecosystems are characterised by low precipitations (lower than 250 mm $year^{-1}$), low winter snow cover, extreme surface desiccation of soils and low relative air humidity (lower than 40%) that force plants to cope with very little moisture (Edwards et al., 2007; Díaz et al., 2016). Great advances have been made on the understanding of the physical and biochemical mechanisms allowing those extremophiles like Alpine plants to extract water from the soil and limit their loss through evapotranspiration. In the course of evolution, plants have developed a set of hierarchical control procedures that allow both short and long-term control of water status (Körner, 2003). As an example, extreme plants tend to limit water deficits during the growing season by highly reducing the osmotic potential that could reach a level lower than −2 MPa (Seemann et al., 1986; Körner, 2003; Liu et al., 2003). The control of osmotic potential in extreme organisms, like for the Alpine plant *Potentilla saundersiana*, is underlined by adaptive mechanisms involving proteins, lipid peroxidation as well as metabolic adjustments (Ma et al., 2015). They revealed the elevation gradient impact on 118 proteins involved in antioxidant processes and on the adjustments of both primary (e.g. soluble sugars, proline) and secondary metabolites (e.g. anthocyanins). The stressful balance between the vital need of CO_2 entry and the danger of water loss has also led to precise control of stomatal density, length, and aperture (Li et al., 2014; Chaves et al., 2016). However, the regulation of these structures is not sufficient under extreme conditions for long-term control of plant water status that requires other mechanisms such as the adjustment of dry mater investments and phenology (e.g. restricting plant size, mass, and area to limit water losses). Indeed, the trend of increasing stomatal density may become reversed in elevation gradient (Körner et al., 1986, 1989) and

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

it is now clearly accepted that plants employ multiple strategies to enhance their adaptation to extreme environments. Thick cell wall structures, the first mechanical barrier to environmental stress, have been observed in leaf samples taken from plants grown at high elevations (Ma et al., 2015). Also, a positive correlation was described between abiotic stress intensity and positive interactions among Alpine plants (Callaway et al., 2002), where more than 90% of the plants shared the same root system at 5200 m of elevation (Ma et al., 2015).

3.2 Extremophile Plants Facing Energy Excess

In extreme conditions, plants are generally characterised by a short growing season where vegetative and reproductive growth compete for carbon (C) resources (Jordan and Nobel, 1979). In these challenging environments, phenotypic plasticity has been considered as an important adaptation response to harsh environments by improving fitness, suggesting a possible role in natural selection under a rapidly changing environment (Richards et al., 2006; Davidson et al., 2011; Godoy et al., 2011). The finding that invasive plant species were more plastic than native ones illustrates the high fragility of extreme environments where human activities could lead to a profound ecological shift by modifying both water and nutrient cycles that would enhance the ability of invaders to outcompete the existing vegetation (van Kleunen and Richardson, 2007; Davidson et al., 2011). Indeed, the invasion of non-native plants could restrict the access to environmental resources (e.g. light availability), which would ultimately limit plant growth of native species.

Plant biomass production is the net result of CO_2 uptake through photosynthesis and its loss through respiration and depends on environmental parameters like the photon flux density. Besides, CO_2 fixation in plants is governed by photosynthetic activity and leaf traits (Krall and Edwards, 1992; Reich et al., 1992). In the Alpine environment, specific leaf area (SLA, leaf area per leaf mass) and total leaf area per total plant mass tend to be lower at high elevation (Körner, 2003). Light is not a limiting factor of photosynthetic activity during the growing season in such environments (Berry and Bjorkman, 1980). Rather the opposite, these plants usually face an excess in photon flux densities, which can cause an imbalance in the photosynthetic machinery characterised by extreme rates of chloroplastic electron transport that finally induce the formation of reactive oxygen species (ROS, detailed in Section 4.3). Consequently, extreme species harbour relevant adaptations to adjust to this high photon flux by decreasing chlorophyll a and b levels, increasing the carotenoid/chlorophyll ratio, and developing effective photochemical quantum yield of PS II or photochemical quenching coefficient, for instance (Öncel et al., 2004; Cui et al., 2019). Also, two large databases on leaf traits of plant species from seven

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

different ecosystems have been analysed to determine whether mass-based metabolism rate was proportional to the surface area in plant leaves (Jin et al., 2008). They have shown that the ratio between mass-based photosynthetic capacity and specific leaf area is significantly higher (1.66) in the Alpine environment (low temperature and high light) than in the tropical forest (1.23, high temperature and high light). In parallel, changes in leaf structure and biochemistry have been observed with thermal acclimation of photosynthesis and respiration in cold-tolerant crops (Yamori et al., 2009), Antarctic vascular plants (Xiong et al., 2000), and tree species (Way and Oren, 2010). These observations pinpointed the impact of temperature variations on all aspects of the C cycle including photosynthesis and respiration processes (Zhang et al., 2015). This explains that temperature acclimation appears to be central to species distribution and growth. In plants, thermal acclimation can optimise C gain by (i) shifting the thermal optimum for net photosynthesis to improve photosynthetic capacity, (ii) modifying the respiration quotient (Q10) and/or the base respiration rate (Zhang et al., 2015).

Thus, these observations suggest that water availability as well as adequate temperatures are critical factors to ensure plant survival and growth while light does not seem to be a limiting environmental feature within the extreme habitats studied.

3.3 Extremophile Plants Facing Low Soil Suitability

In hostile environments, extremophiles not only have to cope with extreme temperatures and aridity but also with an imbalance between essential nutrients (e.g. N, P, and K) and deleterious elements to plant growth like salt and heavy metals (An et al., 2019). In extreme ecosystems, adapted plants are usually growing out of their physiological optimum, considering that both excess and deficiency in soil mineral nutrients are common situations (Lambers et al., 2011, 2013; Díaz et al., 2016). Consequently, plants growing with limited mineral nutrients allocate more C in root systems to improve the uptake capacity of the limiting soil resource (Li et al., 2019; Akram et al., 2020). Similarly, this adaptive strategy was confirmed on diverse perennial herbs that increasingly manage the biomass allocation towards belowground parts and especially storage organs as elevation increases in the Qinghai-Tibetan Plateau (Ma et al., 2010). Among mineral nutrients, N is one of the main nutrients for plant growth and considered a limiting factor for net primary production in all terrestrial ecosystems (Xia and Wan, 2008).

In most deserts, the N cycle is governed by the limitation of N that arises from the limited plant cover, from fluctuations in rainfall and microbial activities (León-Sobrino et al., 2019). The N cycle can be summarised as an exchange of N between the atmosphere and the biosphere (Gruber and

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Galloway, 2008). The atmospheric N_2 (78% of the atmosphere) is inaccessible to higher plants and only a small number of Archaea and Bacteria is able to fix N (Ramond et al., 2018). Thus, microbial activity became the key factor to provide N to desert plants (Hanna et al., 2013), but this microbial activity was controlled by water availability (Yuan et al., 2014). A comparative study performed in five arid Mars analogue environments of the genes related to the N cycle (nitrogenase and nitrite reductase, two functional markers for the identification of microorganisms that mediate N fixation and denitrification within any given community) and of the bacterial community composition (using 16S rRNA clone libraries), has reported that the soil content for almost all measured forms of C, N, and P was higher at the more humid site than at the drier one (López-Lozano et al., 2012). More recently, a study using the combinatory analysis of metagenomic and metatranscriptomic data indicated that the soil salinity shaped microbial communities, which play important roles in the N cycle providing C and N nutrients for higher plants in saline soils (Ren et al., 2018). Likewise, bacteria can break down and use P, which plays a fundamental role in the physiology and biochemistry of all living organisms, in different oxidation states and therefore contributes to ecosystem P cycling (Tapia-Torres et al., 2016). Finally, a meta-analysis showed a negative effect of an extended drought period on plant N and P contents, which suggested that water availability was the main factor governing plant growth rather than N and P, under long-term arid conditions (He and Dijkstra, 2014).

Furthermore, while soil nutrient availability limits plant development by deficiency, salinity restricts the performances by excess (Dodd and Donovan, 1999). As an example, salinity is a key determinant for microbial communities in deserts (Zhang et al., 2019). Besides, in low-nutrient ecosystems, nutrient resorption – the mechanism by which a plant withdraws nutrients from senescent leaves – is a key process to limit nutrient losses and therefore improve survival capacity (Lü et al., 2012). This mechanism was greatly impacted by abiotic stress like salinity in arid lands (Drenovsky et al., 2010). In addition, increasing salinity in desert ecosystems directly interferes with the uptake of cations such as K^+, Ca^{2+}, and Mg^{2+} and reduces N availability as well as P solubility (James et al., 2005). Consequently, extreme high salinity (up to 200 mM) required specific adaptation allowing to manage salt within the plant system. Halophyte plants have shown various remarkable sodium (Na) regulation strategies by enhancing the accumulation and compartmentation capacities, the ability to uptake and manage efficiently essential nutrients (e.g. P), and the ability to regulate transpiration under high salinity (Flowers et al., 2010). A first strategy to cope with saline soils is the one used by pseudohalophytes, which develop the capacity to limit the entry of saline ions within the transpiration stream. However, several halophytes developed adaptive strategies to respond to the massive entry of salt within leaf tissues. Indeed, while euhalophytes developed the possibility

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

of accumulating salt within foliar tissues and sequestering it in vacuoles (Flowers and Colmer, 2008), other halophytes are characterised by a secretion capacity (i.e. secretohalophytes) allowed by secreting glands (see Section 2.2.2) (Wang et al., 2017).

Remarkably, these adaptive traits were also observed in metallophytes (Verbruggen et al., 2009). As a consequence of human activity (agriculture, industry), the frequency of soil contamination by heavy metals steadily increases worldwide (Nagajyoti et al., 2010). Also, high concentrations of heavy metal were observed in ecosystems already naturally characterised as extreme (Zhang et al., 2013). The accumulation of essential micronutrients like Co and Zn, or highly toxic elements like Cd and Cr limits plant growth by impairing key physiological processes such as mineral nutrition, photosynthesis, and enzyme activities (Shanker et al., 2005; Yruela, 2009; Lin and Aarts, 2012), a destabilising context further complicated when coupled with other abiotic pressures in hostile ecosystems. The improvement of physiological mechanisms has allowed revegetation of lands characterised by high metal concentrations lying at the edge of the plant life-compatible gradient. Indeed, years of evolution have led to highly specialised mechanisms, which enable metallophyte tolerance by restricting the entry of metals into the shoot or by enhancing the ability to accumulate metals at high concentrations (Whiting et al., 2004). Excluders (i.e. metallophytes that maintain relatively low constant values of metals in the shoot) are able to thrive by limiting the level of metals translocated from roots to aerial parts until a critical soil value, above which the mechanism breaks down leading to plant death (Baker, 1981). Oppositely, hyperaccumulators (i.e. metallophytes capable of hyperaccumulating metals in aerial tissues) have been studied for their environmental role in detoxifying contaminated-lands, a process known as phytoremediation (Cunningham et al., 1995). Hyperaccumulators, which diverged from 34 different plant families, have shown enhanced capacities of metal uptake, long-distance transport, and detoxification (Peng et al., 2020). Vacuolar sequestration in the leaves is thought to be the main pathway of metal detoxification and involves an efficient transport from the roots to the leaf vacuole and the chelation of metals, which are bound to ligands such as organic acids or amino acids like histidine (Verbruggen et al., 2009). Interestingly, these adaptive processes require enhancing the expression of key genes that are shared with the non-hyperaccumulating relative species, suggesting that several plants could adapt to extreme environments by modifying basal functions rather than by developing entirely new biochemical pathways. Also, the evolution of both constitutive or induced barriers and physiological functions underlies specialised biochemical adaptations as well as supplemental adaptive metabolic features that together allowed a few plant species to cope with extreme environmental conditions.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

4 Metabolic Adaptations at the Heart of Plant Responses to Extreme Ecosystems

Metabolism is the cornerstone of plant responses to environmental changes. Hence, the metabolic adaptations improving survival chances in hostile ecosystems have received considerable attention over the past few decades (see Table S1 for references). Using both targeted and untargeted metabolomics techniques, great advances have been achieved to describe these metabolic features by (i) comparing extreme species to related crops or model plants (Lugan et al., 2010; Yobi et al., 2013), (ii) analysing plant metabolic profiles in their natural environment (Tipirdamaz et al., 2006), and (iii) characterising plant metabolic responses through a gradient of abiotic stresses (Kumari et al., 2020). Untargeted and targeted analyses on extremophile plants transferred to controlled conditions allowed isolating one environmental variable to compare its effect on extremophile versus model species and highlight different metabolic response strategies. Complementarily, metabolic profiling of extremophile plant species described the different biochemical compounds accumulated within a given ecosystem, while the study of these organisms through a stress gradient (e.g. elevation, salinity) improves our understanding of how extreme plant metabolomes adapt to abiotic stress under extreme conditions (see Table S1 for references). This review sought to summarise these works, covering five different ecosystems (i.e. deserts, mountains, frozen lands, saline lands, and metal-contaminated sites), thereby pinpointing the convergences of the metabolic responses between species in one ecosystems and between ecosystems (Figure 1). This meta-analysis included 69 species and revealed a dynamic response of central, primary, secondary, and redox metabolisms. However, only 31.6% of the referenced metabolites referred to secondary metabolism (Figure 1a). This suggests that these specialised metabolic pathways are often overlooked, which urges for enhancing untargeted analyses with wide metabolome coverage of extremophile plant species.

4.1 Primary Metabolism

4.1.1 Central Pathways

Glycolysis, tricarboxylic acid cycle, and oxidative pentose phosphate pathways are ubiquitous in the plant kingdom and essential for defining plant performance (plant growth and tolerance to biotic and abiotic stress) via its interactions with the photosynthetic and photorespiratory processes and amino-acid biosynthesis, for instance (Fernie et al., 2004). The glycolytic pathway provides carbohydrates carrying osmoprotectant functions that can also be used as a C source for polyol and oligosaccharide biosynthesis (Singh et al., 2015). The central place of the tricarboxylic acid cycle (TCA) cycle in

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

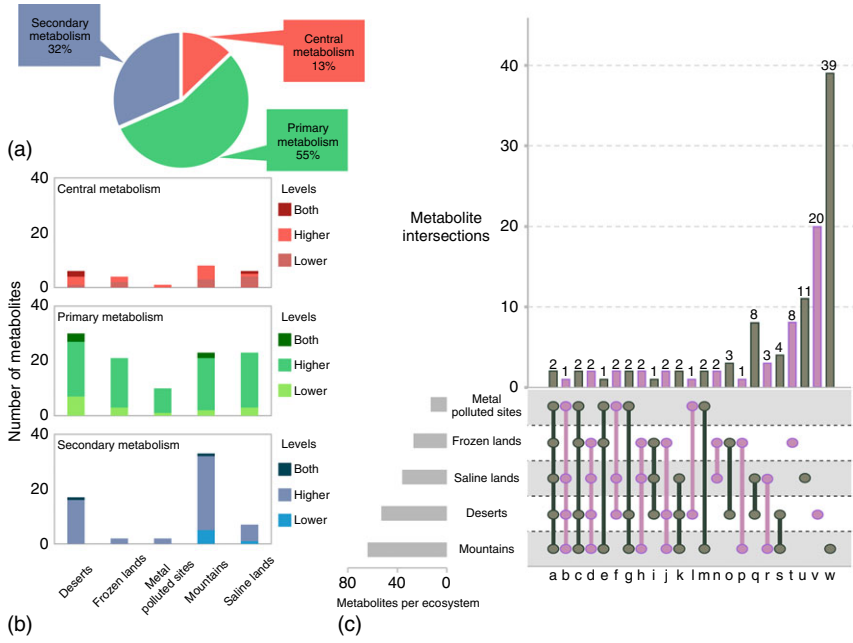

Figure 1 A comprehensive meta-analysis of the metabolic features observed in plants thriving under extreme environments. (a) Distribution of affected metabolic pathways. (b) Details of the distribution in five extreme ecosystems and the response of molecules to environmental pressures: Higher, higher concentration; Lower, lower concentration; Both, depending on the plant species. (c) UpSet plot of the metabolic overlaps between ecosystems. The bottom left side shows the number of metabolites described for each ecosystem. The right side shows the possibilities and levels of intersections of these molecules between the five extreme ecosystems. Letters refer to compounds associated with the intersection (one letter per column). a: aspartate, inositol; b: proline; c: ascorbate, galactinol; d: glucose, sucrose; e: trehalose; f: glutamate, glutamine; g: glutathione, histidine; h: fructose, threonine; i: GABA, j: malate, raffinose; k: citrate, phenylalanine; l: mannitol; m: sorbitol, starch; n: pinitol, valine; o: stachyose, tocopherol, xylose; p: melibiose; q: alanine, asparagine, choline, choline-*O*-sulfate, fumarate, glycine, putrescine, serine; r: glycine betaine, lysine, rhamnose; s: arabinose, arginine, spermidine, zeaxanthin; t: arachidonic acid, catechin, digalactosylglycerol, linoleic acid, linolenic acid, oleic acid, ononitol, tryptophan; u: β-alanine betaine, fucose, isoleucine, kaempferol, laricitrin, leucine, malonic acid, norvaline, quercetin, succinate, tyrosine; v: allantoin, apigenin, galactose, guanosine, hydroxyproline, isocitrate, lutein, luteolin, naringenin, neoxanthin, ribose, spermine, tri-*O*-galloylquinic acid, verbascose, violaxanthin, γ-glutamylisoleucine, γ-glutamylleucine, γ-glutamylmethionine, γ-glutamylphenylalanine, γ-glutamylthreonine; w: 2-oxoglutarate, abscisic acid, caffeoylquinic acids, cinnamic acid, coumaric acid, farnesene, ferulic acid, gentiobiose, glycerol, hypoxanthine, jacareubin, kokusaginine, maltose, melezitose, MHDglycerol, neohesperidin, Picroside I, Picroside II, Picroside III, Picroside IV, quebrachitol, ranunculin, rebeccamycin, romucosine B, sagecoumarin, sagerinic acid, seychellene, tagatose, tannin, thapsigargin, thujone, total carotenoids, total flavonoids, total sugards, total phenolic content, total xantophyll pigments, tricetin, xanthosine, xylitol.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

abiotic stress tolerance has been widely accepted and characterised by the role of organic acids as important players in osmoregulation, support of ionic gradients across membranes, acidification of extracellular spaces, and the maintenance of redox equilibrium (Igamberdiev and Eprintsev, 2016). TCA enzymes were also linked to photosynthetic activity and organic acids were embedded within a series of chemical reactions to produce or consume amino acids and secondary metabolites (Sweetlove et al., 2010).

In extreme environments, the adjustment of these central pathways under abiotic stress involved compounds that were up- or downregulated depending on the plant species and ecosystem (Figure 1b). Total soluble sugars, including glucose, fructose, and sucrose, were likely to be induced with the elevation gradient (Hashim et al., 2020). However, while high sucrose concentration was reported independently of the plant species and environment, glucose, and fructose regulations were quite variable (Figure 1b and Table S1). Interestingly, the same trends were observed within the TCA cycle where malate accumulated for all conditions and species, whilst citrate, fumarate, and succinate levels fluctuated, an observation perhaps explained by the fact that major organic acid content depends on the plant species (Chia et al., 2000; Mikulic-Petkovsek et al., 2012; Igamberdiev and Eprintsev, 2016). These results support the central place of these pathways in managing C resource allocation between primary and secondary pathways and in determining C allocation between plant growth and defence under extreme environmental conditions.

4.1.2 Primary Metabolites

Lipids are major constituents of all biological membranes that represent an interface between the cell and adverse environmental pressure (Mazliak, 1977; Anjum et al., 2015). The lipid remodelling observed in several plants like Arabidopsis is one critical mechanism to improve membrane stability and therefore limit damaging effects from freezing stress (Moellering et al., 2010).

Similarly, the shifts in lipid compositions were observed in extreme plants under abiotic stress like temperature (Zheng et al., 2011) or drought (Giarola et al., 2017). As an example, lipid profiling of resurrection plants has revealed an adjustment in lipid metabolism when submitted to drought stress (Quartacci, 2002), resulting in higher unsaturation levels (Tshabuse et al., 2018). Similarly, an increase in lipid unsaturation is a common response to cold stress (Barrero-Sicilia et al., 2017) and a positive correlation was observed between polyunsaturated fatty acid levels from seabuckthorn species and elevation in the Himalayan mountains (Sharma et al., 2020), supporting their role in maintaining membrane fluidity (Upchurch, 2008). Furthermore, the remobilisation of membrane lipids towards the synthesis of signalling lipids as phosphoinositides and phosphatidic acids was reported in the resurrection plant *Craterostigma plantagineum* (Gasulla et al., 2013) and accepted as

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

a major regulator of membrane properties, protein functions and phytohormone pathways (Hou et al., 2016). Altogether, these observations in both freezing- and drought-tolerant extremophile plants suggest a central role of lipid remodelling in adaptation to extreme environments, by countering the effects of direct drought and cellular dehydration initiated by extracellular ice formation (Moellering et al., 2010).

The role of primary metabolism in plant tolerance to osmotic stress was highlighted when describing the importance of osmotic adjustment in regulating plant cellular turgor and stomatal conductance (Blum, 2017). Compatible solutes, organic osmolytes responsible for osmotic balance and compatible with cellular metabolism (Galinski, 1993), have shown to accumulate in both drought- and freezing-tolerant plant species for their role in osmotic adjustment and cryoprotection (Chen and Murata, 2002; Bhandari and Nayyar, 2014). Moreover, the contributions in plant tolerance to abiotic stress of other compatible solutes, like the raffinose family oligosaccharides (RFO), have been extended to participate in stabilising proteins and membrane phospholipids (ElSayed et al., 2014).

Under extreme low temperatures, adapted woody plants have shown high concentrations of oligosaccharides, which regulate viscosity in the cytoplasm and therefore prevent deleterious effects of freezing temperatures (Stushnoff et al., 1997; Strimbeck et al., 2015). Contents in disaccharides (e.g. trehalose, melibiose), as well as RFOs like raffinose and stachyose, were heightened under extreme temperatures, drought, and heavy metal contamination (Figure 1c). Osmoprotective properties were also represented among polyol compounds like mannitol, allowing plants to cope with extreme salinity (Slama et al., 2015). Also, inositol, pinitol, mannitol, sorbitol, and galactinol were over-represented within desert, cold-tolerant and hyperaccumulator plant species, and likely to be upregulated under each abiotic stress (Figure 1c). This observation agrees with the possible roles of sugar alcohols as carbohydrate reserves or the thermoprotective function of sorbitol in higher plants (Moing, 2000).

Besides sugars, plant resistance under extreme conditions is thought to be partially related to the induction of amino acids, which are important metabolic intermediates for the synthesis of environment-responsive, specialised metabolites (Chouhan et al., 2017). In addition, a myriad of controversial functions was proposed for other amino acids, for example proline accumulation, including cytosolic pH buffer, protein structure stabiliser, osmotic adjustment, ROS scavenger, and metal chelator. However, recent efforts have allowed pinpointing the relationship of proline metabolism between either the pentose phosphate pathway or the mitochondrial electron transport (Kaur and Asthir, 2015). The central role of proline in plant tolerance to multiple stresses could result from (i) its synthesis that limits the reducing power, thus leading to an imbalance of photosynthetic activity, and (ii) the benefits of proline degradation that provides C to the

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

TCA cycle and thus contributes to respiratory activity (Kaur and Asthir, 2015).

Likewise, 26 extremophile species accumulated proline under various stressful conditions, and the shikimic acid pathway activity resulted in the accumulation of tyrosine and phenylalanine in some of these plants (Table S1). These aromatic amino acids act as precursors for the biosynthesis of flavonoids, known as secondary antioxidant compounds (Chouhan et al., 2017), suggesting a possible role of secondary pathways in adaptation. Furthermore, high concentrations of aspartate referenced in plant survivors from several extreme ecosystems (Figure 1c) could be linked to redox and C metabolisms by fuelling the synthesis of nicotinamide adenine dinucleotide (NAD), a redox cofactor also involved in stress responses (Gakière et al., 2018). However, pyridine nucleotide contents have been unfortunately overlooked in metabolic studies of extremophile plants. Altogether, these observations strongly suggest that the study of plant resistance against extreme abiotic stress should not be restricted to primary metabolism activity but broadened to a larger scale to include secondary and redox metabolisms.

4.2 Secondary Metabolism

Plant secondary metabolism has been already recognised as a major actor of plant–environment interaction, involving many specialised metabolites that accumulate under stressful conditions (Akula and Ravishankar, 2011). For instance, linear accumulations of polyphenols were revealed within several elevation gradients for different species (Zidorn, 2010; Monschein et al., 2015; Cirak et al., 2017). In extreme conditions, the impact of abiotic stress on secondary metabolism was observed with 29 plant species from different extreme environments (Table S1 and Figure 1b). The potential role in the adaptation of several molecules including quaternary ammonium compounds, terpenes, and phenolics was also mentioned for different plant families. Interestingly, these compounds were not specific to extremophile species and were further present in crop and model plants (Parida et al., 2018). This suggests that secondary metabolites found in various plants could play important roles in stress mitigation of harsh climates.

4.2.1 Nitrogen-related Compounds

Polyamines are N-containing secondary low molecular compounds involved in both plant development and stress resistance (Chen et al., 2019). The biosynthesis pathway of putrescine and higher polyamines as spermidine and spermine is ubiquitous and greatly affected by abiotic stress. Molecular mechanisms by which polyamines alleviate plant tolerance to abiotic stress are not fully understood but several works have reported interesting properties and results for such compounds. As an example, the application of these

compounds could regulate the size of potassium channels and therefore the aperture of pores in the plasma membrane, which suggests a role in water loss control (Alcázar et al., 2010). The integration of polyamines within both primary and secondary pathways was supported by the foliar application of putrescine that triggers the biosynthesis of amino acids and sugars (Chen et al., 2019). Finally, polyamine accumulation promotes ROS degradation by raising antioxidant enzyme activities and possibly affects ion transport under salt stress (Saha et al., 2015, see Section 4.3).

Remarkably, polyamine profiles in extremophiles appear up- or downregulated depending on the environmental stress (Figure 1b). This observation could be explained by the fact that plant resistance is more likely to be associated with a high ratio of (Spermidine + Spermine)/Putrescine, suggesting that this protective role mainly involves higher polyamines (Zapata et al., 2004; Chen et al., 2019). Another hypothesis is that polyamines could conjugate with other molecules, such as coumaric or caffeoyl acids, and lead to complex roles in plant defence (Alcázar et al., 2010; Burt et al., 2019). However, the proportions between free and conjugated polyamines vary among different plant species (Bagni and Tassoni, 2001), and few details on the occurrence and function of such conjugated forms in extremophile plant species are available.

Quaternary ammonium compounds have been controversially considered as an adaptive response of halophytes or other extremophile plants that improves plant tolerance to drought, salt, and low-temperature stress (Ashraf and Foolad, 2007). Their input in plant resistance mechanisms was illustrated by the high concentrations of glycine betaine and choline-*O*-sulfate found in Alpine, desert, and halophyte plant species (Figure 1c). Both compounds encountered in a wide range of extreme plants act as key compatible solutes. Glycine betaine is synthesised in the chloroplast from choline or glycine and could be involved in osmoregulation by modulating Na^+ and K^+ content, leading to a higher K^+/Na^+ ratio, which alleviates salt tolerance in higher plants (Hu et al., 2012). In addition, glycine betaine has shown a relevant role in maintaining membrane integrity as well as enzyme activities. Finally, accumulation of glycine betaine within different transgenic plant species could protect the photosynthetic machinery against salt stress damage (Giri, 2011). However, several plants like *Brassica napus* fail to accumulate glycine betaine (Gibon et al., 1997) and higher levels could negatively correlate with the production of other stress markers such as proline in some extremophiles (Tipirdamaz et al., 2006).

Hence, while several extreme plants displayed important levels of N-related compounds in hostile environments suggesting a possible role in adaptation, great variabilities were observed about polyamine accumulation and the understanding of how both polyamines and quaternary ammonium compounds are integrated in the response to abiotic stress of the different plant species. These observations raise the critical need (exposed in Section 5)

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

of analysing metabolic features in a more holistic approach where tolerance mechanisms would be integrated into the different plant systems thriving in hostile ecosystems.

4.2.2 Terpenoids

Terpenoids, or terpenes, represent one of the main class of secondary metabolites in terms of biodiversity and have been massively used as pharmaceuticals and industrial compounds (Tetali, 2019). The two compartmented pathways producing these compounds are the mevalonate (MVA) pathway in the cytoplasm and the methylerythritol phosphate (MEP) pathway in plastids, both leading to different subclasses of terpenoids (Cheng et al., 2007). The activity of these chemical pathways can be inversely affected in response to light conditions. While the expression of genes involved in the MVA pathway that produces sterols is downregulated, MEP pathway genes are upregulated under light conditions, enhancing carotenoid and tocopherol production (Tholl, 2015). The role of carotenoids and tocopherols in photoprotection through antioxidant activity has been already widely characterised (described in Section 4.3), but new insights were highlighted when describing the mechanism by which ROS can cause the oxidative cleavage of carotenoids leading to hormonal compounds such as phytohormones (e.g. strigolactone or abscisic acid) (Havaux, 2014). Finally, the roles of phytohormones in plant resistance against abiotic pressures like heavy metals contamination and osmotic stress are thoroughly examined in different reviews (Fahad et al., 2015; Singh et al., 2016; Sharma et al., 2019).

Interestingly, the impact of abiotic stress on photosynthetic pigments and tocopherols is verified in extreme environments where both compound classes were accumulated in plants from desert and mountain ecosystems (Figure 1c). Besides, the increased levels of zeaxanthin and abscisic acid found in Alpine (Figure 1) and most resurrection plants (Rascio and Rocca, 2005) possibly illustrate the link between carotenoids and phytohormones. Altogether, these observations in extremophile plants emphasise the pivotal role for terpenes in stress signalling through hormonal responses, and stress mitigation via the processing of excess ROS in response to extreme temperatures and radiation levels, for instance.

4.2.3 Polyphenols

Polyphenols are another major group of secondary metabolites presenting a wide biological diversity. Phenolics are synthesised from amino acids like phenylalanine via the shikimic acid pathway and classified according to their chemical structure. The five main classes, namely phenolic acids, stilbenes, flavonoids, lignans, and others, have demonstrated pleiotropic roles in both plant growth and resistance (López-Fernández et al., 2020). Their contributions to extreme environmental stress responses are reflected with the high polyphenol concentrations in plants of Alpine and desert

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

environments, which cover a great part of abiotic stresses, ranging from cold to high temperature, from nutrient deficiency to excess of salt, and both present high light intensities (Figure 1c). Furthermore, phenolics accumulated in a wide biodiversity of extremophile species (see Table S1 for references) and medicinal plants, some of which thrive under harsh conditions (Li et al., 2020; Najjaa et al., 2020).

Complementarily, an increased concentration of flavonoids mainly occurred in cold and high-light intensity or poor nutrient soil conditions (Table S1). This observation corroborates with the fact that flavonoids, and specially hydroxycinnamic acid derivatives (e.g. *p*-coumaric and ferulic acid) or quercetin derivatives, hold an important antioxidant function improving photoprotection and reducing damage caused by UV radiations and frost (Agati and Tattini, 2010; Schulz et al., 2016). Recently, the UV-B protective function was extended to a global enhancement of ROS-processing activity independently of the solar wavelength proportions, based on the upregulation of flavonoids in response to ROS accumulation and an imbalance of redox homeostasis (Di Ferdinando et al., 2012). In addition, these same authors argue in favour of this hypothesis by highlighting the interaction between cold and N stress that leads to the same flavonoid upregulation profiles as for cold and high light conditions. Finally, the comparison of flavonoid contents of several plants on different latitudes reported a qualitative change in flavonoids in response to temperature stress (Jaakola and Hohtola, 2010), which points the need of improving our understanding about the distinct roles of flavonoid classes. On the other hand, the accumulation of metabolites like cinnamic acid under extreme conditions may have a role in the production of lignin compounds known to be upregulated under abiotic stress to reinforce secondary cell walls (Le Gall et al., 2015).

Finally, it is noteworthy that the biosynthesis of several phenolics, proline, and polyols consume NADPH (Loescher and Everard, 2000; Szabados and Savouré, 2010; Caretto et al., 2015), and thus participate in the control of cellular redox homeostasis by limiting the excess of reducing power. Hence, the increase of the $NADP^+$/NADPH ratio would possibly enhance the oxidative pentose phosphate pathway activity, which provides precursors for phenolic compound production (Caretto et al., 2015). Altogether, these observations suggest that not only the proper function of each primary or secondary compound matters but also both their biosynthesis and degradation, which therefore emphasises the need of a more integrated approach to study the metabolic features of extreme plants.

4.3 Tuning Redox Metabolism Upon Extreme Climates

Due to the presence of ground-state oxygen (O_2) as a natural oxidant on earth, reduction-oxidation (redox) processes generate ROS, which encompass

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

highly reactive molecules that are partially reduced or excited forms of O_2 (e.g. $^1O_2$, H_2O_2, $O_2^{\bullet-}$ and $OH^{\bullet}$) (Mittler, 2017). Fascinatingly, photosynthetic organisms benefit from the redox potential of O_2 to produce energy (ATP) and reducing power (NAD(P)H) that fuel metabolic reactions. Hence, plant central metabolism produces ROS mostly via three sources that include chloroplastic photosynthesis, mitochondrial respiration, and peroxisomal photorespiration (Schertl and Braun, 2014) (Figure 2a). In photosynthetic tissues, ROS mainly originate from the photosynthetic electron transport chain (Foyer, 2018), while other sources are important for different organs such as fruit tissues (Decros et al., 2019). Study of redox metabolism in model plant species outlined a dual function for ROS both as toxic by-products of central metabolism and as powerful signals that modulate plant development and environmental responses (Kalia et al., 2017; Smirnoff and Arnaud, 2019). Therefore, when redox homeostasis becomes unbalanced, a lack of ROS or ROS accumulation leads to reductive or oxidative stress, respectively, characterised by, for instance, DNA damages, oxidation of cysteine residues in proteins and lipid peroxidation inducing retrograde signalling or in some cases cell death (Figure 2b). Harmonious plant growth thus requires a finely tuned redox homeostasis to avoid oxidative or reductive stress.

Due to their sessile lifestyle, plants have developed numerous antioxidative defence pathways and strategies to control their redox state (Figure 2a). Plants are natural biochemists that synthesise many different antioxidant molecules, although most plant families have developed their own range of specific antioxidant metabolites within their botanical taxa. Nevertheless, some major redox buffers are ubiquitous, like ferredoxins, carotenoids, vitamins (e.g. tocopherols), pyridine nucleotides, thioredoxins, glutathione, and ascorbate, which play fundamental roles in the development of plants and their responses to the environment (Geigenberger and Fernie, 2014; Gupta et al., 2016; Gakière et al., 2018). Besides antioxidant compounds, plants possess a common set of enzymes involved in ROS-processing (i.e. superoxide dismutase (SOD), catalase (CAT), peroxidases), in the control of the redox state (i.e. glutathione reductase (GR), dehydroascorbate reductase (DHAR), monodehydroascorbate reductase (MDHAR) and NAD(P)H-dehydrogenases), and signalling (e.g. glutaredoxins, thioredoxins) (Geigenberger et al., 2017; Martins et al., 2018). These redox features interplay within the ascorbate–glutathione cycle, also known as Foyer-Halliwell pathway, that involves a *ménage-à-trois* between ascorbate, glutathione and NAD(P), thereby participating actively in ROS processing and in controlling the cellular redox state (Foyer and Noctor, 2011; Decros et al., 2019) (Figure 2a).

Redox metabolism is particularly responsive to environmental changes, which has been extensively reviewed (Choudhury et al., 2017; Noctor et al., 2018; Decros et al., 2019). In harsh environments, plants growing at high elevation undergo extreme temperature variations, water limitation, nutrient

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

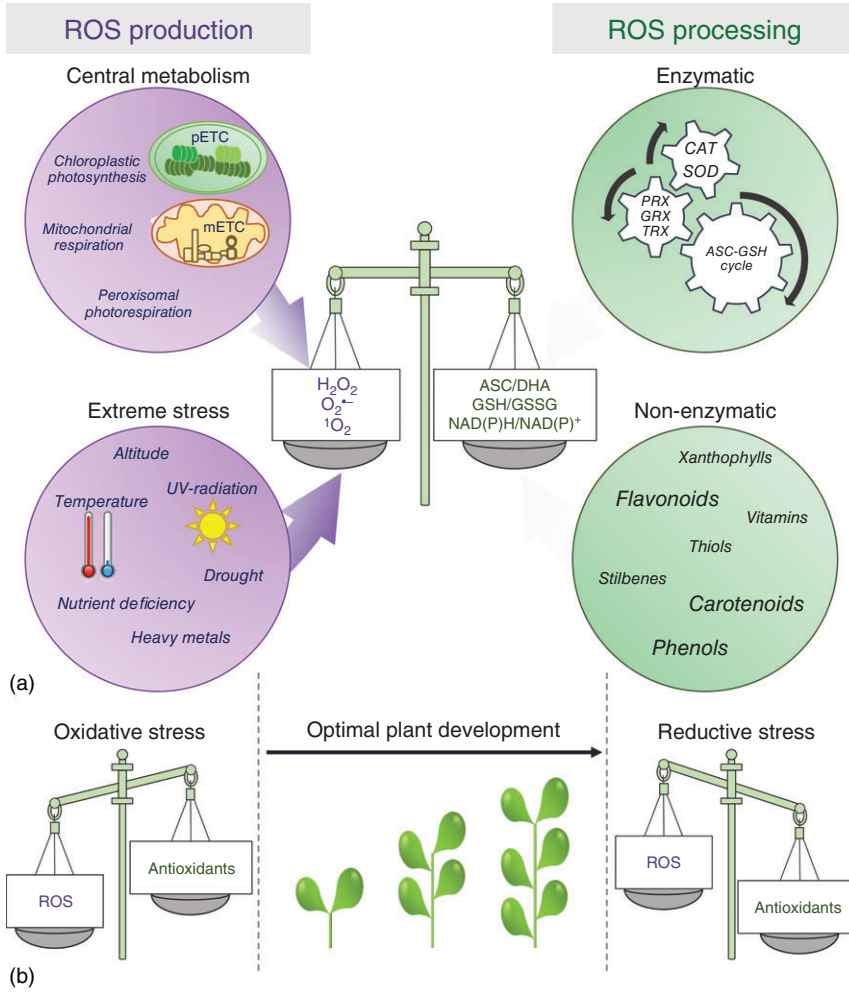

Figure 2 Redox poise is pivotal to plant growth and acclimation. (a) Plants produce ROS and other redox signals during growth and in response to environmental stimuli. Redox homeostasis relies on the balance between ROS production (left side) and processing (right side). This involves several enzymatic and non-enzymatic mechanisms, including antioxidant metabolites and major redox buffers (NAD/P(H), ASC and GSH). (b) Harmonious plant growth requires a finely tuned redox homeostasis to avoid oxidative or reductive stress when the ROS/antioxidant balance is altered. ASC, ascorbate; CAT, catalase; DHA, dehydroascorbate; GRX, glutaredoxins; GSH, glutathione; GSSG, disulfide glutathione; mETC, mitochondrial electron transport chain; pETC, photosynthetic electron transport chain; PRX, peroxiredoxins; ROS, reactive oxygen species; SOD, superoxide dismutase; TRX, thioredoxins.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

deficiency and high levels of irradiation inducing a reduced photosynthetic activity (see above), which results in a surplus of reducing power (NADPH) and higher (photo)respiration (Fernández Marín et al., 2020). All these abiotic stresses have been shown to exacerbate ROS production by central metabolism in model or agronomical species grown under laboratory conditions (Choudhury et al., 2017; Pandey et al., 2017). Generally, abiotic changes trigger an accumulation of antioxidative metabolites (e.g. tocopherols, carotenoids, ascorbate, and glutathione) associated with discrepancies in enzymatic activities of the ascorbate–glutathione cycle (Geigenberger et al., 2017; Noctor et al., 2018). A combination of water deficit, high irradiance, and temperature changes may induce a desynchronisation between photosynthetic electron transport chain (pETC) and light-harvesting reactions leading to electron sinks, which can result in photooxidation (Gollan et al., 2017) (Figure 2a). On the other hand, abiotic stress also influences the mitochondrial (photo)respiration level and thus mitochondrial ROS production, but to a lesser extent than chloroplast activity in photosynthetic tissues (Sevilla et al., 2015). Plant mitochondria possess alternative oxidases (AOX) and uncoupling proteins (UCP) that deviate the electron flow from mitochondrial complexes (Saha et al., 2016). Consequently, these proteins prevent excess mitochondrial ROS production by dissipating energy as heat, thus giving more adaptability to plant respiration and participation in mitochondrial signalling processes. Regrettably, a great majority of studies realised in extreme environments have reported redox responses that involve chloroplastic activity and photoprotective mechanisms, whereas few details on mitochondrial metabolism are available.

To adjust to the higher oxidative stress in extreme habitats, plants exploit sophisticated antioxidant systems to balance the redox poise. For instance, to prevent oxidative damages, Alpine plants adapt their photosynthetic defence machinery by increasing their free radical-scavenging capacity through the accumulation of photoprotective metabolites such as carotenoids, flavonoids, and phenolics (Ma et al., 2015; Cui et al., 2019; Hashim et al., 2020). These specialised plant compounds are powerful antioxidants that process ROS, consume reducing power and can also avoid UV-induced damages (Stapleton and Walbot, 1994; Bieza and Lois, 2001; Caretto et al., 2015; Young and Lowe, 2018) (Figure 2a). A first field study of nine Alpine plants from different elevations described a higher content in total leaf antioxidants, especially in ascorbate (Wildi and Lutz, 1996), suggesting an important role of redox homeostasis in plant acclimation. Besides, a comparative study of 18 steppes species (~1000 m) and 11 mountain species (~2000 m) identified an augmented antioxidant capacity in species growing at higher elevation, with specific respect to SOD activity and content in ascorbate and carotenoids (Öncel et al., 2004). More recently, biochemical and proteomic analysis of Tibetan plants highlighted a positive correlation between the content of soluble antioxidants (ascorbate and

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

phenolics), ROS-processing enzyme activity (SOD, CAT, and ascorbate peroxidase (APX)) elevation (Ma et al., 2015; Cui et al., 2019; Hashim et al., 2020). Altogether, these studies of endemic species from the Tibetan plateau reported a strong correlation between the elevation gradient and the plant redox metabolism, which was characterised by higher enzymatic and non-enzymatic ROS processing capacities.

Likewise, desert species also harboured higher contents in carotenoids and flavonoids, and increased enzymatic activities of the ascorbate–glutathione cycle, particularly glutathione reductase, suggesting a more important role of glutathione- and thiol-related signalling in the adaptation to desert lands (Streb et al., 1997; Talbi et al., 2015; Wang et al., 2016). In addition, the high oxidation state of glutathione and ascorbate have been correlated with desiccation and rehydration tolerance in *Myrothamnus flabellifolia* and *Boea hygrometrica*, two resurrection plants (Kranner et al., 2002; Jiang et al., 2007). Finally, even though most studies on extremophile plants demonstrated a higher antioxidant capacity related to the main redox metabolites and enzymes (ascorbate–glutathione cycle), a comparative study between the native Antarctic species *Colobanthus quitensis* and a genetically related species *Dianthus chinensis*, reported a normal antioxidant enzyme activity but a higher antioxidant capacity related to sulphur and secondary metabolisms (Clemente-Moreno et al., 2020). Conversely with the observation made in other extreme environments, CAT and APX activity did not increase in any of these species as well as glutathione content and oxidation state. Nonetheless, after cold exposure, *C. quitensis* harboured a total antioxidant activity 20 times higher than *D. chinensis* combined with a twofold increased respiration rate and activity of alternative oxidase. This illustrates the plasticity of redox pathways to maintain the redox poise depending on the environment and plant botanical taxa.

Hence, cellular redox homeostasis is a key factor that accompanies plant growth and responses to the environment, more remarkably within the activity of the ascorbate–glutathione cycle. In addition, antioxidant secondary metabolism (e.g. carotenoids, flavonoids, and phenols) further appears as a relevant pathway to stimulate plant oxidative defence capacity and thus participates in the acclimation of plants to harsh environments. Currently, the paradigm of redox biology tends to display a bigger and clearer picture of the redox network occurring in plants, where multiple sources of ROS are possible and associated with many 'ROS processing systems' (Noctor et al., 2018). Spatial, temporal, metabolic, and antioxidant specificities are multiple factors that can influence redox signalling. While knowledge on redox biology in plants living in extreme environments is still fragmentary, the concepts that originate from model and agronomic species are useful to study the redox metabolism for plant acclimation.

The study of plants subjected to extreme environments undoubtedly demonstrates a reorchestration of plant metabolism for primary, secondary,

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

and redox pathways. However, the metabolic data available so far only provide a fragmentary knowledge and a more holistic, global overview of plant metabolome would benefit our understanding of plant acclimation to harsh climates. This could be addressed through untargeted metabolomics approaches to encompass a greater diversity of plant compounds, more specifically for specialised metabolites and redox compounds, which are likely to be central to mitigate stress. More importantly, it is now crucial to define the steps that should be further addressed to move from the description of metabolic features towards (i) the identification of metabolic divergences that could result from species-specific adaptation, and (ii) the convergences between plant species and between ecosystems to pinpoint generic mechanisms underpinning adaptive strategies.

5 Challenges and Perspectives

The fascinating ability of plants to survive and thrive under extreme conditions has always attracted the curiosity of humans, who have discovered and improved plant species capable of meeting our physiological needs (Preece et al., 2017). Thus, the domestication of therophytes was central to the beginning of agriculture that took place over 10 000 years ago in the Fertile Crescent (Riehl et al., 2012), an area characterised by its extreme environments (Dai et al., 2012). These ancestors, that were adapted to harsh environmental conditions, led after many years of breeding and genetic improvements to the modern wheat and barley, currently two of the major sources of food for humans (Awika et al., 2011). We now rely on these and a few other domesticated species for our survival, which increases the pressure on our ability to reshape the best genetic ideotypes of these plants, whose yield is stagnating and even threatened by climate change (Long et al., 2015). It was, therefore, tempting to return to study extremophiles in order to identify resistance mechanisms that could then be transferred to cultivated plants. However, this strategy has not proven very successful so far. Thus, it has been more than 50 years since the accumulation of compatible solutes observed in many halophytic and drought-resistant species (Flowers, 1972; Jones and Gorham, 1983) raised great hopes, but the transfer of these metabolic properties into crops did not improve yield under abiotic stress so far (Turner, 2018). In other words, the reductionist approach of studying resistance mechanisms in isolation might not allow us to understand metabolic adaptation to extreme environments.

The first challenge will therefore be to use systems biology approaches by which the metabolic adaptations of extremophiles can be contextualised (Figure 3). By systems biology, we mean an iteration between experimentation and modelling (Engelhardt et al., 2016), the goal of which will be to

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

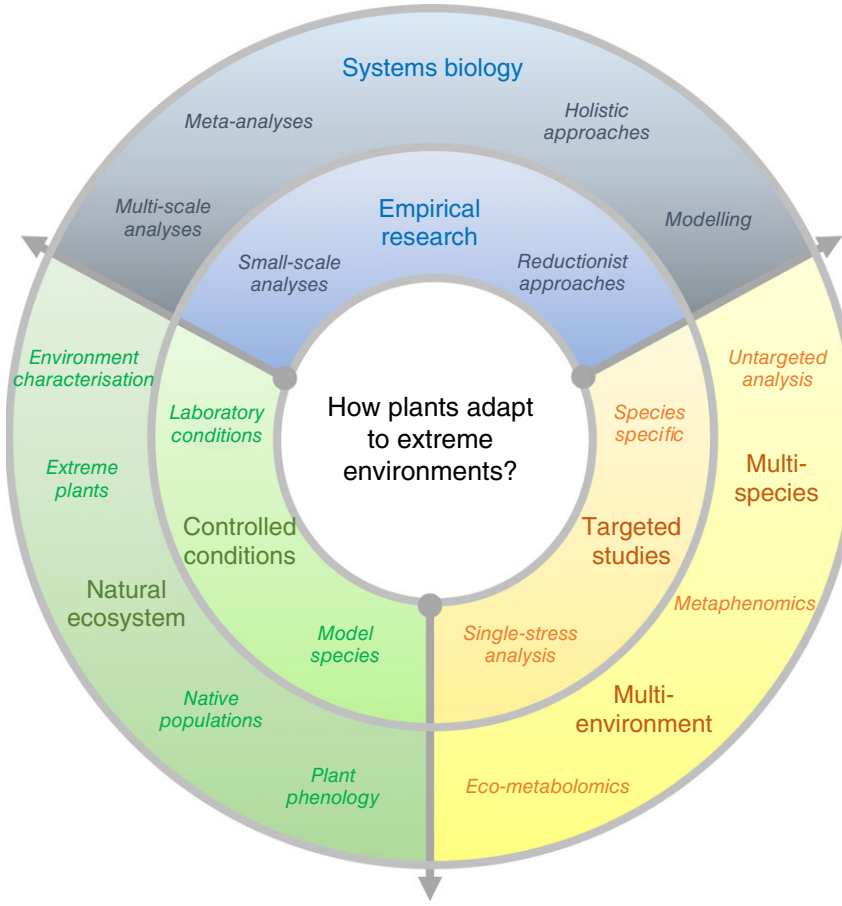

Figure 3 Perspectives for a better understanding of plant adaptations to extreme habitats. A comprehensive exploration of the metabolome of multiple species in different environments combined with physiological and ecological data using systems biology approaches can lead to new breakthroughs in the understanding of plant adaptation to extreme environments.

identify and then understand the mechanisms used by extremophiles to adapt. Because models are limited to transform input variables into output variables, it will be particularly important to deal with the most interesting models, especially when it comes to defining performance (Fernandez et al., 2016).

The second challenge will be to make the most relevant observations possible. It is probably useful to define performance first, then to choose the species to be studied accordingly. Thus, if one is interested in the ability

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

to maintain high growth under salt stress, it will be interesting to pick fast-growing halophytes. More generally, it will be important to clearly distinguish the ability to produce biomass, which is an important concept for agronomy, and fitness or reproductive performance, which is an ecological concept (Körner, 2018). Ecological and physiological optima can indeed be very different. Related to that, an important question will be to decide whether studies will be performed in the natural environment (Kumari et al., 2020) or under controlled growth conditions. Moving plants from their natural environment to controlled conditions provides the possibility to achieve reproducible experiments in which factors such as water supply or temperature can be tested (Figure 3). However, such an approach is likely to cause an acclimation process that could hide adaptive traits, unless the natural context is perfectly reproduced (e.g. soil composition, atmospheric pressure, environmental variable interactions). In contrast, ecosystems could be used as natural laboratories, in which the environment would be thoroughly monitored by characterising a maximum of environmental variables and by defining the analysis period (e.g. season, time of the day, plant organ, and developmental stage). However, even if given factors could vary (e.g. adding fertiliser or watering), a major threat here lies in poor reproducibility of the growth conditions. Thus, field experiments require an in-depth analysis of the environment, in which climate and edaphic variables and species biodiversity are monitored over different seasons or even years (Díaz et al., 2016, 2019). Plant phenology can also represent an issue when biomass and biodiversity are constantly fluctuating with seasons and weather conditions that regulate plant development and physiology (Nicotra et al., 2010) (Figure 3). Once these design steps have been tackled, the description of metabolic adaptations from both therophytes and perennial extreme plants has greatly improved (See Table S1 for references.) Interestingly, as knowledge of the environment increases, more and more metabolic convergence is found between plant species (Figure 2b), but also distinct strategies. Thus, the orchestration of primary pathways that leads to the accumulation of amino acids or their use as precursors of a range of secondary metabolites appears to be universally shared between extremophile plants (Lugan et al., 2010; Arbelet-Bonnin et al., 2020). Then, relatively high levels of the N-rich polyamines (e.g. putrescine, spermidine, and spermine), which can be conjugated to a range of secondary metabolites (Alcázar et al., 2010; Burt et al., 2019), and quaternary ammonium metabolites have been found in a wide range of extremophiles. Strikingly, these groups of metabolites are often negatively correlated with other stress markers such as proline (Tipirdamaz et al., 2006) and could even decrease performance in non-accumulating species (Gibon et al., 1997). Besides, increased amounts of isoprenoids like photosynthetic pigments and phenolics such as flavonoids were observed in harsh ecosystems (Yobi et al., 2012, 2013) and their levels correlated with elevation in alpine regions (Cui et al., 2019). Finally, a

central role has been suggested for redox metabolism, independently of plant species and environments (Eshel et al., 2017; Zhao et al., 2019; Hashim et al., 2020). However, the high biochemical diversity of metabolites that are abundant in extremophiles (Peters et al., 2018) including many that are still unknown (Gagneul et al., 2007; Sanchez et al., 2011), and the absence of success in engineering the accumulation of compatible solutes in crops (Turner, 2018), are still making the understanding of metabolic adaptations to extreme environments very difficult. Consequently, we now need to enlarge the coverage of the metabolome when studying extremophiles, in particular by using untargeted analytics and improve our annotation capacities (Allard et al., 2017) (Figure 3). Then, data should be integrated by using both unsupervised and supervised statistical approaches to highlight and then confirm hypotheses linking metabolism and adaptation (e.g. relations between metabolic traits and performance). These efforts will directly enrich the basis of the metabolic features from extremophile organisms that will thereafter be crucial to moving from their description to the comprehension of adaptive mechanisms harboured by both therophytes and perennial plants.

The third challenge will be to reconsider the concept of model species (Figure 3). Traditionally, physiological studies and later functional genomics have focused on a given species to study adaptation, betting that mechanisms found in one species could be relevant to others and transferred to crops (Cushman and Bohnert, 2000). More recently, the use of intra-species genetic diversity has emerged as a powerful tool to study metabolism and better understand how it participates in plant performance (Clancy et al., 2018; Tůmová et al., 2018). The so-called metaphenomics approach goes even further by researching mechanisms within panels of species (Poorter et al., 2010; Sardans et al., 2020) (Figure 3). Metaphenomics allows the integration of phenotypic traits (e.g. senescence, plant resistance, biomass production and allocation) and environmental data for a wide range of species via meta-analyses (Poorter et al., 2010). This approach has been performed to quantify combinations of phenotypic performance in different stressful environments (Wright et al., 2004; Poorter et al., 2010), and to analyse the biomass allocation in multiple species, in response to the environment (Poorter et al., 2012, 2015). Strikingly, the introduction of biological functions such as photosynthesis within these meta-analyses has provided promising results by correlating the plasticity to light intensity, plant density and plant environments (Poorter et al., 2019). Besides, the integration of data gathered from multiple species could be very useful for studying metabolism. For instance, a study carried out with ten species of fruit has shown a close link between fruit relative growth rate and the composition of the biomass (Roch et al., 2020). This approach could be particularly useful for studying the metabolic adaptations of extremophiles, which indeed seem to use convergent metabolic mechanisms to adapt. A more systematic approach

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

based on the comparison of the metabolome and other metabolic traits of panels of extremophile and non-extremophile species will allow us to address better the contributions of generic or specific mechanisms involved in metabolic adaptation. Fortunately, hostile ecosystems not only offer the opportunity to investigate physiological mechanisms under realistic conditions but also provide the possibility of coupling metabolomics to ecology (i.e. to connect metabolism with performance defined as fitness). Eco-metabolomics has been defined as 'the application of metabolomics techniques in ecological studies to characterise biochemical mechanisms underlying interactions of organisms with the environment and with other organisms across different spatial and temporal scales' (Peters et al., 2018) and appears as another great opportunity to move from model systems to native populations in the field (Nagler et al., 2018; figure 3). Evolution has led to a huge diversity of metabolites estimated at around one million compounds that have different or common functions and various levels of interaction (Kroymann, 2011; Afendi et al., 2012). Eco-metabolomic studies would allow employing this biodiversity as an advantage by investigating the adaptive metabolic features on a multi-species level, which would therefore emphasise the research on convergent adaptive mechanisms to cope with both biotic and abiotic stresses (Hennion et al., 2012; Sardans et al., 2020).

Finally, the discovery of convergent and divergent metabolic adaptations to extreme environments should be complemented by a better understanding of developmental processes that allow survival, growth and ultimately fitness. As an example, both flowering deserts of Atacama and Namaqualand after the rainy season offer a good representation of what plant evolution allowed in terms of structural, physiological, and metabolic responses to extreme conditions (Cowling et al., 1999; Vidiella et al., 1999). Indeed, while several annual plants have developed spectacular capacities of providing an entire life cycle in a short-time period under extreme conditions, suggesting both adequate sensors and efficient biological systems from uptake to development, perennial plants have to face a stressful decision between plant growth, energy storage, and plant defence. Watering and N-fertilisation experiments performed with desert extremophiles have shown the positive relationship between water and/or N supply and biomass production and further traits involved in plant performance (Gutierrez and Whitford, 1987; Brooks, 2003; Zhou et al., 2011). Moreover, the role played by metabolism in this gap between the ecological optimum and the physiological optimum remains very little explored. Also, an experiment performed with the apple of Sodom (*Calotropis procera*) in the Negev desert nicely provided new insights into the activity of both primary and secondary pathways following water input and pinpointed the challenge of linking plant performance, environmental perturbations, and metabolic responses (Ramadan et al., 2014). Finally, the multiplication of this type of experiment, in a wide range

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

of species, with an exhaustive collection of environmental and physiological data (in particular performance descriptors) as well as a deep exploration of the metabolome, should lead to the uncovering of many mechanisms and at the end a better understanding of adaptation (Figure 3). In this context, we can already speculate on the importance that the interoperability of metabolomics data and the annotation of metabolomes will have on the discovery process.

6 Concluding Remarks

Natural habitats are underpinned by numerous constraints, which often occur simultaneously. The interaction of these environmental pressures could result in adverse conditions that ultimately alter plant physiology. Nevertheless, plants, most particularly Angiosperms, have the captivating capacity to invade some of the most extreme and dynamic environments, including arid and cold biomes (Folk et al., 2020). Extremophile plants have been studied for decades and exciting progress has been made in the identification of the plant responses to some hostile habitats. Plant resistance to these challenging conditions was allowed by temporal avoidance strategies (e.g. short life-cycle), while other sessile organisms have developed a wide range of structural and physiological features complemented by adaptive metabolic mechanisms to adapt, therefore, through space. However, research was mostly restricted to selective approaches based, for instance, on the analysis of known adaptive mechanisms in a small number of species, thus addressing only facets of the total. Hence, further investigations from anatomical, physiological and molecular angles are necessary to enable comprehensive, systems-wide approaches (Figure 3). To a greater extent, metaphenomics allows an unprecedented connection between such analyses and will therefore provide novel insight into the exploration and understanding of the general strategies of plant adaptations.

Besides, the implementation of omics strategies based on high-throughput and holistic analytical techniques holds great promise for the study of multiple species, in various habitats, and through complementary, multi-scale aspects. This would help provide a better understanding of adaptive traits and address the question of acclimation in the context of adaptation. Indeed, the fact that extremophile plants rarely perform at their physiological optimum (Ramadan et al., 2014) raises the question of whether their capacity to thrive results from efficient metabolic responses to abiotic stress, replacing adaptation as the capacity to acclimate to changing environmental conditions? Furthermore, high-coverage metabolomics data, obtained from multiple platforms, provide a unique opportunity to gain fundamental insights into metabolic and even cellular regulations, which

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

occupy a critical place for plant adaptation. Likewise, other levels of omics, like transcriptomics and proteomics, will provide a significant boost to the study and the comprehension of the adaptive molecular processes of plants to extreme environments.

Acknowledgements

The authors would like to thank Dr Timothy Tranbarger (IRD, France) for his invitation to contribute a review article for *Annual Plant Reviews* online. The Pontifical Catholic University of Chile in Santiago (Chile) and the University of Bordeaux (France) are also acknowledged for providing financial support through doctoral grants to Thomas Dussarrat and Guillaume Decros, respectively.

List of Abbreviations

AOX	alternative oxidase
APX	ascorbate peroxidase
As	arsenic
C	carbon
CAT	catalase
Cd	cadmium
Co	cobalt
Cr	chrome
Cu	copper
DHAR	dehydroascorbate reductase
GR	glutathione reductase
H_2O_2	hydrogen peroxide
MDHAR	monodehydroascorbate reductase
Mo	molybdenum
Mn	Manganese
N	nitrogen
NAD	Nicotinamide adenine dinucleotide
NADP	NAD phosphate
NADP(H)	refers to both $NADP^+$ and NADPH
Ni	nickel
P	phosphorous
Pb	lead
ROS	reactive oxygen species
SOD	superoxide dismutase
Zn	zinc

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Related Articles

Introduction: Biochemistry, Role and Biotechnology of Secondary Metabolites (From APR Volume 2)
Modes of Action of Defensive Secondary Metabolites
Plant Pigments and Protection against UV-B Radiation
Abiotic Stress and Metabolomics
Abiotic Stress Responses are Governed by Reactive Oxygen Species and Age
Signalling Interactions in Flooding Tolerance
Regulating Root Aquaporin Function in Response to Changes in Salinity
Plant Desiccation Tolerance: A Survival Strategy with Exceptional Prospects for Climate-Smart Agriculture
Multi-omics Insights into the Evolution of Angiosperm Resurrection Plants
Leaf Trait Plasticity and Evolution in Different Plant Functional Types
To *Coldly* Go where No Grass has Gone Before: A Multidisciplinary Review of Cold Adaptation in Poaceae
Impact of Cell-wall Structure and Composition on Plant Freezing Tolerance

Supporting Information

Additional supporting information may be found online in the Supporting Information section in the HTML rendition of this article.

References

Afendi, F.M., Okada, T., Yamazaki, M. et al. (2012). KNApSAcK family databases: integrated metabolite-plant species databases for multifaceted plant research. *Plant & Cell Physiology* **53** (**2**): e1. doi: 10.1093/pcp/pcr165.

Agati, G. and Tattini, M. (2010). Multiple functional roles of flavonoids in photoprotection: Letters. *New Phytologist* **186** (**4**): 786–793. doi: 10.1111/j.1469-8137.2010.03269.x.

Akram, M.A., Wang, X., Hu, W. et al. (2020). Convergent variations in the leaf traits of desert plants. *Plants* **9** (**8**): 990. doi: 10.3390/plants9080990.

Akula, R. and Ravishankar, G.A. (2011). Influence of abiotic stress signals on secondary metabolites in plants. *Plant Signaling & Behavior* **6** (**11**): 1720–1731. doi: 10.4161/psb.6.11.17613.

Alcázar, R., Altabella, T., Marco, F. et al. (2010). Polyamines: molecules with regulatory functions in plant abiotic stress tolerance. *Planta* **231** (**6**): 1237–1249. doi: 10.1007/s00425-010-1130-0.

Allard, P.-M., Genta-Jouve, G., and Wolfender, J.-L. (2017). Deep metabolome annotation in natural products research: towards a virtuous cycle in metabolite identification. *Current Opinion in Chemical Biology* **36**: 40–49. doi: 10.1016/j.cbpa.2016.12.022.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Allbed, A. and Kumar, L. (2013). Soil salinity mapping and monitoring in arid and semi-arid regions using remote sensing technology: a review. *Advances in Remote Sensing*. doi: 10.4236/ars.2013.24040.

Alleman, L.K. and Hester, M.W. (2011). Reproductive ecology of Black Mangrove (*Avicennia germinans*) along the Louisiana coast: propagule production cycles, dispersal limitations, and establishment elevations. *Estuaries and Coasts* **34** (**5**): 1068. doi: 10.1007/s12237-011-9404-8.

An, H., Tang, Z., Keesstra, S. et al. (2019). Impact of desertification on soil and plant nutrient stoichiometry in a desert grassland. *Scientific Reports* **9** (**1**): 9422. doi: 10.1038/s41598-019-45927-0.

Anjum, N.A., Sofo, A., Scopa, A. et al. (2015). Lipids and proteins—major targets of oxidative modifications in abiotic stressed plants. *Environmental Science and Pollution Research* **22** (**6**): 4099–4121. doi: 10.1007/s11356-014-3917-1.

Antonovics, J., Bradshaw, A.D., and Turner, R.G. (1971). Heavy metal tolerance in plant. *Advances in Ecological Research*: 1–85. doi: 10.1016/S0065-2504(08)60202-0.

Arbelet-Bonnin, D., Blasselle, C., Rose Palm, E. et al. (2020). Metabolism regulation during salt exposure in the halophyte *Cakile maritima*. *Environmental and Experimental Botany* **177**: 104075. doi: 10.1016/j.envexpbot.2020.104075.

Aschi-Smiti, S., Chaibi, W., Brouquisse, R. et al. (2003). Assessment of enzyme induction and aerenchyma formation as mechanisms for flooding tolerance in Trifolium subterraneum "Park". *Annals of Botany* **91** (**2**): 195–204.

Ashraf, M. and Foolad, M.R. (2007). Roles of glycine betaine and proline in improving plant abiotic stress resistance. *Environmental and Experimental Botany* **59** (**2**): 206–216. doi: 10.1016/j.envexpbot.2005.12.006.

Awika, J.M., Piironen, V., and Bean, S. (2011). Major cereal grains production and use around the world. In: *Advances in Cereal Science: Implications to Food Processing and Health Promotion* (eds. J.M. Awika, V. Piironen and S. Bean), 1–13. American Chemical Society: Washington, DC. doi: 10.1021/bk-2011-1089.ch001.

Bagni, N. and Tassoni, A. (2001). Biosynthesis, oxidation and conjugation of aliphatic polyamines in higher plants. *Amino Acids* **20** (**3**): 301–317. doi: 10.1007/s007260170046.

Baker, A.J.M. (1981). Accumulators and excluders -strategies in the response of plants to heavy metals. *Journal of Plant Nutrition* **3** (**1–4**): 643–654. doi: 10.1080/01904168109362867.

Ball, M.C. (1988). Ecophysiology of mangroves. *Trees* **2** (**3**): 129–142. doi: 10.1007/BF00196018.

Barrero-Sicilia, C., Silvestre, S., Haslam, R.P. et al. (2017). Lipid remodelling: unravelling the response to cold stress in Arabidopsis and its extremophile relative *Eutrema salsugineum*. *Plant Science* **263**: 194–200. doi: 10.1016/j.plantsci.2017.07.017.

Battisti, D.S. and Naylor, R.L. (2009). Historical warnings of future food insecurity with unprecedented seasonal heat. *Science* **323** (**5911**): 240–244. doi: 10.1126/science.1164363.

Berry, J. and Bjorkman, O. (1980). Photosynthetic response and adaptation to temperature in higher plants. *Annual Review of Plant Physiology* **31** (**1**): 491–543. doi: 10.1146/annurev.pp.31.060180.002423.

Bhandari, K. and Nayyar, H. (2014). Low temperature stress in plants: an overview of roles of cryoprotectants in defense. In: *Physiological Mechanisms and Adaptation*

Strategies in Plants Under Changing Environment (eds. P. Ahmad and M.R. Wani), 193–265. New York: Springer. doi: 10.1007/978-1-4614-8591-9_9.

Bickford, C.P. (2016). Ecophysiology of leaf trichomes. *Functional Plant Biology* **43** (**9**): 807–814.

Bieza, K. and Lois, R. (2001). An arabidopsis mutant tolerant to lethal ultraviolet-B levels shows constitutively elevated accumulation of flavonoids and other phenolics. *Plant Physiology* **126** (**3**): 1105–1115. doi: 10.1104/pp. 126.3.1105.

Blom, C.W.P.M., Bögemann, G.M., Laan, P. et al. (1990). Adaptations to flooding in plants from river areas. *Aquatic Botany* **38** (**1**): 29–47. doi: 10.1016/0304-3770(90)90097-5.

Blum, A. (2017). Osmotic adjustment is a prime drought stress adaptive engine in support of plant production. *Plant, Cell & Environment* **40** (**1**): 4–10. doi: 10.1111/pce.12800.

Boojar, M.M.A. and Tavakkoli, Z. (2011). New molybdenum-hyperaccumulator among plant species growing on molybdenum mine – a biochemical study on tolerance mechanism against metal toxicity. *Journal of Plant Nutrition* **34** (**10**): 1532–1557. doi: 10.1080/01904167.2011.585209.

Borowitzka, M.A. (2018). The "stress" concept in microalgal biology—homeostasis, acclimation and adaptation. *Journal of Applied Phycology* **30** (**5**): 2815–2825. doi: 10.1007/s10811-018-1399-0.

Bowne, J., Bacic, A., Tester, M. et al. (2018). Abiotic stress and metabolomics. *Annual Plant Reviews Online*: 61–85. American Cancer Society. doi: 10.1002/9781119312994.

Brooks, M.L. (2003). Effects of increased soil nitrogen on the dominance of alien annual plants in the Mojave Desert: nitrogen effects on alien annual plants. *Journal of Applied Ecology* **40** (**2**): 344–353. doi: 10.1046/j.1365-2664.2003.00789.x.

Brown, J.H. (1990). Species diversity. In: *Analytical Biogeography* (eds. A.A. Myers and P.S. Giller), 57–89. The Netherlands: Springer. doi: 10.1007/978-94-009-0435-4_3.

Bundy, J.G., Davey, M.P., and Viant, M.R. (2008). Environmental metabolomics: a critical review and future perspectives. *Metabolomics* **5** (**1**): 3. doi: 10.1007/s11306-008-0152-0.

Burt, A.J., Arnason, J.T., and García-Lara, S. (2019). Natural variation of hydroxycinnamic acid amides in maize landraces. *Journal of Cereal Science* **88**: 145–149. doi: 10.1016/j.jcs.2019.06.002.

Caldwell, M.M., Teramura, A.H., and Tevini, M. (1989). The changing solar ultraviolet climate and the ecological consequences for higher plants. *Trends in Ecology & Evolution* **4** (**12**): 363–367. doi: 10.1016/0169-5347(89)90100-6.

Caldwell, M.M., Bornman, J., Ballaré, C. et al. (2007). Terrestrial ecosystems, increased solar ultraviolet radiation, and interactions with other climate change factors. *Photochemical & Photobiological Sciences* **6** (**3**): 252–266.

Callaway, R.M., Brooker, R.W., Choler, P. et al. (2002). Positive interactions among alpine plants increase with stress. *Nature* **417** (**6891**): 844–848. doi: 10.1038/nature00812.

Caretto, S., Linsalata, V., Colella, G. et al. (2015). Carbon fluxes between primary metabolism and phenolic pathway in plant tissues under stress. *International Journal of Molecular Sciences* **16** (**11**): 26378–26394. doi: 10.3390/ijms161125967.

Chaves, M.M., Costa, J.M., Zarrouk, O. et al. (2016). Controlling stomatal aperture in semi-arid regions—The dilemma of saving water or being cool? *Plant Science* **251**: 54–64. doi: 10.1016/j.plantsci.2016.06.015.

Chen, T.H.H. and Murata, N. (2002). Enhancement of tolerance of abiotic stress by metabolic engineering of betaines and other compatible solutes. *Current Opinion in Plant Biology* **5** (**3**): 250–257. doi: 10.1016/S1369-5266(02)00255-8.

Chen, D., Shao, Q., Yin, L. et al. (2019). Polyamine function in plants: metabolism, regulation on development, and roles in abiotic stress responses. *Frontiers in Plant Science* **9**: 1945. doi: 10.3389/fpls.2018.01945.

Cheng, A.-X., Lou, Y.-G., Mao, Y.-B. et al. (2007). Plant terpenoids: biosynthesis and ecological functions. *Journal of Integrative Plant Biology* **49** (**2**): 179–186. doi: 10.1111/j.1744-7909.2007.00395.x.

Chia, D.W., Yoder, T.J., Reiter, W.-D. et al. (2000). Fumaric acid: an overlooked form of fixed carbon in Arabidopsis and other plant species. *Planta* **211** (**5**): 743–751. doi: 10.1007/s004250000345.

Choudhury, F.K., Rivero, R.M., Blumwald, E. et al. (2017). Reactive oxygen species, abiotic stress and stress combination. *The Plant Journal* **90** (**5**): 856–867. doi: 10.1111/tpj.13299.

Chouhan, S., Sharma, K., Zha, J. et al. (2017). Recent advances in the recombinant biosynthesis of polyphenols. *Frontiers in Microbiology* **8**: 2259. doi: 10.3389/fmicb.2017.02259.

Cirak, C., Radusiene, J., Jakstas, V. et al. (2017). Altitudinal changes in secondary metabolite contents of *Hypericum androsaemum* and *Hypericum polyphyllum*. *Biochemical Systematics and Ecology* **70**: 108–115. doi: 10.1016/j.bse.2016.11.006.

Clancy, M.V., Zytynska, S.E., Moritz, F. et al. (2018). Metabotype variation in a field population of tansy plants influences aphid host selection: plant chemical diversity in a plant-aphid system. *Plant, Cell & Environment* **41** (**12**): 2791–2805. doi: 10.1111/pce.13407.

Clemente-Moreno, M.J., Omranian, N., Sáez, P. et al. (2020). Cytochrome respiration pathway and sulphur metabolism sustain stress tolerance to low temperature in the Antarctic species *Colobanthus quitensis*. *New Phytologist* **225** (**2**): 754–768. doi: 10.1111/nph.16167.

Cohen, D. (1966). Optimizing reproduction in a randomly varying environment. *Journal of Theoretical Biology* **12** (**1**): 119–129.

Cowling, R.M., Esler, K.J., and Rundel, P.W. (1999). Namaqualand, South Africa – an overview of a unique winter-rainfall desert ecosystem. *Plant Ecology* **142** (**1**): 3–21. doi: 10.1023/A:1009831308074.

Cui, G., Ji, G., Liu, S. et al. (2019). Physiological adaptations of *Elymus dahuricus* to high altitude on the Qinghai–Tibetan Plateau. *Acta Physiologiae Plantarum* **41** (**7**): 115. doi: 10.1007/s11738-019-2904-z.

Cunningham, S.D., Berti, W.R., and Huang, J.W. (1995). Phytoremediation of contaminated soils. *Trends in Biotechnology* **13** (**9**): 393–397. doi: 10.1016/S0167-7799(00)88987-8.

Cushman, J.C. and Bohnert, H.J. (2000). Genomic approaches to plant stress tolerance. *Current Opinion in Plant Biology* **3** (**2**): 117–124. doi: 10.1016/S1369-5266(99)00052-7.

Dai, F., Nevo, E., Wu, D. et al. (2012). Tibet is one of the centers of domestication of cultivated barley. *Proceedings of the National Academy of Sciences of the United States of America* **109** (**42**): 16969–16973. doi: 10.1073/pnas.1215265109.

Dassanayake, M. and Larkin, J.C. (2017). Making plants break a sweat: the structure, function, and evolution of plant salt glands. *Frontiers in Plant Science* **8**. doi: 10.3389/fpls.2017.00406.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Davidson, A.M., Jennions, M., and Nicotra, A.B. (2011). Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis: Invasive species have higher phenotypic plasticity. *Ecology Letters* **14** (**4**): 419–431. doi: 10.1111/j.1461-0248.2011.01596.x.

Decros, G., Baldet, P., Beauvoit, B. et al. (2019). Get the balance right: ROS homeostasis and redox signalling in fruit. *Frontiers in Plant Science* **10**. doi: 10.3389/fpls.2019.01091.

Di Ferdinando, M., Brunetti, C., Fini, A. et al. (2012). Flavonoids as antioxidants in plants under abiotic stresses. In: *Abiotic Stress Responses in Plants* (eds. P. Ahmad and M.N.V. Prasad), 59–179. New York: Springer. doi: 10.1007/978-1-4614-0634-1_9.

Díaz, F.P., Frugone, M., Gutiérrez, R.A. et al. (2016). Nitrogen cycling in an extreme hyperarid environment inferred from $\delta^{15}N$ analyses of plants, soils and herbivore diet. *Scientific Reports* **6**: 22226. doi: 10.1038/srep22226.

Díaz, F.P., Latorre, C., Carrasco-Puga, G. et al. (2019). Multiscale climate change impacts on plant diversity in the Atacama Desert. *Global Change Biology* **25** (**5**): 1733–1745. doi: 10.1111/gcb.14583.

Dodd, G.L. and Donovan, L.A. (1999). Water potential and ionic effects on germination and seedling growth of two cold desert shrubs. *American Journal of Botany* **86** (**8**): 1146–1153. doi: 10.2307/2656978.

Dong, X. and Zhang, X. (2000). Special stomatal distribution in *Sabina vulgaris* in relation to its survival in a desert environment. *Trees* **14** (**7**): 369–375.

Drenovsky, R.E., James, J.J., and Richards, J.H. (2010). Variation in nutrient resorption by desert shrubs. *Journal of Arid Environments* **74** (**11**): 1564–1568. doi: 10.1016/j.jaridenv.2010.05.030.

Dürr, C., Dickie, J.B., Yang, X.-Y. et al. (2015). Ranges of critical temperature and water potential values for the germination of species worldwide: contribution to a seed trait database. *Agricultural and Forest Meteorology* **200**: 222–232. doi: 10.1016/j.agrformet.2014.09.024.

Edwards, A.C., Scalenghe, R., and Freppaz, M. (2007). Changes in the seasonal snow cover of alpine regions and its effect on soil processes: a review. *Quaternary International* **162–163**: 172–181. doi: 10.1016/j.quaint.2006.10.027.

Eggli, U. and Nyffeler, R. (2009). Living under temporarily arid conditions - succulence as an adaptive strategy. *Bradleya*: 13–36. British Cactus and Succulent Society. doi: 10.25223/brad.n27.2009.a10.

ElSayed, A.I., Rafudeen, M.S., and Golldack, D. (2014). Physiological aspects of raffinose family oligosaccharides in plants: protection against abiotic stress. *Plant Biology* **16** (**1**): 1–8. doi: 10.1111/plb.12053.

Engelhardt, B., Frőhlich, H., and Kschischo, M. (2016). Learning (from) the errors of a systems biology model. *Scientific Reports* **6** (**1**): 20772. doi: 10.1038/srep20772.

Eshel, G., Shaked, R., Kazachkova, Y. et al. (2017). *Anastatica hierochuntica*, an Arabidopsis desert relative, is tolerant to multiple abiotic stresses and exhibits species-specific and common stress tolerance strategies with its halophytic relative, *Eutrema (Thellungiella) salsugineum*. *Frontiers in Plant Science* **7**. doi: 10.3389/fpls.2016.01992.

Fahad, S., Hussain, S., Matloob, A. et al. (2015). Phytohormones and plant responses to salinity stress: a review. *Plant Growth Regulation* **75** (**2**): 391–404. doi: 10.1007/s10725-014-0013-y.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Fahn, A. (1988). Secretory tissues in vascular plants. *New Phytologist* **108** (**3**): 229–257. doi: 10.1111/j.1469-8137.1988.tb04159.x.

Farooq, M., Hussain, M., Wahid, A. et al. (2012). Drought stress in plants: an overview. In: *Plant Responses to Drought Stress* (ed. R. Aroca), 1–33. Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-32653-0_1.

Fedoroff, N.V., Battisti, D.S., Beachy, R.N. et al. (2010). Radically rethinking agriculture for the 21st century. *Science (New York, N.Y.)* **327** (**5967**): 833–834. doi: 10.1126/science.1186834.

Fernández Marín, B., Gulías, J., Figueroa, C.M. et al. (2020). How do vascular plants perform photosynthesis in extreme environments? An integrative ecophysiological and biochemical story. *The Plant Journal* **101** (**4**): 979–1000. doi: 10.1111/tpj.14694.

Fernandez, O., Urrutia, M., Bernillon, S. et al. (2016). Fortune telling: metabolic markers of plant performance. *Metabolomics* **12** (**10**): 158. doi: 10.1007/s11306-016-1099-1.

Fernie, A.R., Carrari, F., and Sweetlove, L.J. (2004). Respiratory metabolism: glycolysis, the TCA cycle and mitochondrial electron transport. *Current Opinion in Plant Biology* **7** (**3**): 254–261. doi: 10.1016/j.pbi.2004.03.007.

Flowers, T.J. (1972). The effect of sodium chloride on enzyme activities from four halophyte species of chenopodiaceae. *Phytochemistry* **11** (**6**): 1881–1886. doi: 10.1016/S0031-9422(00)90147-X.

Flowers, T.J. and Colmer, T.D. (2008). Salinity tolerance in halophytes. *New Phytologist* **179** (**4**): 945–963. doi: 10.1111/j.1469-8137.2008.02531.x.

Flowers, T.J., Galal, H.K., and Bromham, L. (2010). Evolution of halophytes: multiple origins of salt tolerance in land plants. *Functional Plant Biology* **37** (**7**): 604. doi: 10.1071/FP09269.

Folk, R.A., Siniscalchi, C.M., and Soltis, D.E. (2020). Angiosperms at the edge: extremity, diversity, and phylogeny. *Plant, Cell and Environment*. doi: 10.1111/pce.13887.

Foyer, C.H. (2018). Reactive oxygen species, oxidative signaling and the regulation of photosynthesis. *Environmental and Experimental Botany* **154**: 134–142. doi: 10.1016/j.envexpbot.2018.05.003.

Foyer, C.H. and Noctor, G. (2011). Ascorbate and glutathione: the heart of the redox hub. *Plant Physiology* **155** (**1**): 2.

Fridley, J.D. (2017). Plant energetics and the synthesis of population and ecosystem ecology. *Journal of Ecology* **105** (**1**): 95–110. doi: 10.1111/1365-2745.12693.

Gagneul, D., Aïnouche, A., Duhazé, C. et al. (2007). A reassessment of the function of the so-called compatible solutes in the halophytic Plumbaginaceae *Limonium latifolium*. *Plant Physiology* **144** (**3**): 1598–1611. doi: 10.1104/pp. 107.099820.

Gakière, B., Fernie, A.R., and Pétriacq, P. (2018). More to NAD+ than meets the eye: a regulator of metabolic pools and gene expression in Arabidopsis. *Free Radical Biology & Medicine* **122**: 86–95. doi: 10.1016/j.freeradbiomed.2018.01.003.

Galinski, E.A. (1993). Compatible solutes of halophilic eubacteria: molecular principles, water-solute interaction, stress protection. *Experientia* **49** (**6–7**): 487–496. doi: 10.1007/BF01955150.

Gasulla, F., vom Dorp, K., Dombrink, I. et al. (2013). The role of lipid metabolism in the acquisition of desiccation tolerance in *Craterostigma plantagineum*: a comparative approach. *The Plant Journal* **75** (**5**): 726–741. doi: 10.1111/tpj.12241.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Geigenberger, P. and Fernie, A.R. (2014). Metabolic control of redox and redox control of metabolism in plants. *Antioxidants & Redox Signaling* **21** (**9**): 1389–1421. doi: 10.1089/ars.2014.6018.

Geigenberger, P., Thormählen, I., Daloso, D.M. et al. (2017). The unprecedented versatility of the plant thioredoxin system. *Trends in Plant Science* **22** (**3**): 249–262. doi: 10.1016/j.tplants.2016.12.008.

Giarola, V., Hou, Q., and Bartels, D. (2017). Angiosperm plant desiccation tolerance: hints from transcriptomics and genome sequencing. *Trends in Plant Science* **22** (**8**): 705–717. doi: 10.1016/j.tplants.2017.05.007.

Gibon, Y., Bessieres, M.A., and Larher, F. (1997). Is glycine betaine a non-compatible solute in higher plants that do not accumulate it? *Plant, Cell and Environment* **20** (**3**): 329–340. doi: 10.1046/j.1365-3040.1997.d01-82.x.

Giri, J. (2011). Glycinebetaine and abiotic stress tolerance in plants. *Plant Signaling & Behavior* **6** (**11**): 1746–1751. doi: 10.4161/psb.6.11.17801.

Girvetz, E.H. and Zganjar, C. (2014). Dissecting indices of aridity for assessing the impacts of global climate change. *Climatic Change* **126** (**3–4**): 469–483. doi: 10.1007/s10584-014-1218-9.

Godoy, O., Valladares, F., and Castro-Díez, P. (2011). Multispecies comparison reveals that invasive and native plants differ in their traits but not in their plasticity: phenotypic plasticity and invasiveness. *Functional Ecology* **25** (**6**): 1248–1259. doi: 10.1111/j.1365-2435.2011.01886.x.

Gollan, P.J., Lima-Melo, Y., Tiwari, A. et al. (2017). Interaction between photosynthetic electron transport and chloroplast sinks triggers protection and signalling important for plant productivity. *Philosophical Transactions of the Royal Society, B: Biological Sciences* **372** (**1730**): 20160390. doi: 10.1098/rstb.2016.0390.

Gruber, N. and Galloway, J.N. (2008). An Earth-system perspective of the global nitrogen cycle. *Nature* **451** (**7176**): 293–296. doi: 10.1038/nature06592.

Gupta, D.K., Palma, J.M., and Corpas, F.J. (2016). *Redox State as a Central Regulator of Plant-Cell Stress Responses*. Cham: Springer International Publishing. doi: 10.1007/978-3-319-44081-1.

Gutierrez, J.R. and Whitford, W.G. (1987). Chihuahuan desert annuals: importance of water and nitrogen. *Ecology* **68** (**6**): 2032–2045. doi: 10.2307/1939894.

Gutterman, Y. (2000). Environmental factors and survival strategies of annual plant species in the Negev Desert, Israel. *Plant Species Biology* **15** (**2**): 113–125. doi: 10.1046/j.1442-1984.2000.00032.x.

Hamblin, A.P. (1986). The influence of soil structure on water movement, crop root growth, and water uptake. *Advances in Agronomy*: 95–158. Elsevier. doi: 10.1016/S0065-2113(08)60674-4.

Hanna, A.L., Youssef, H.H., Amer, W.M. et al. (2013). Diversity of bacteria nesting the plant cover of north Sinai deserts, Egypt. *Journal of Advanced Research* **4** (**1**): 13–26. doi: 10.1016/j.jare.2011.11.003.

Hasanuzzaman, M., Nahar, K., and Fujit, M. (2013). Extreme temperature responses, oxidative stress and antioxidant defense. In: *Plants. Abiotic Stress – Plant Responses and Applications in Agriculture* (ed. K. Vahdati). doi: 10.5772/54833.

Hashim, A.M., Alharbi, B.M., Abdulmajeed, A.M. et al. (2020). Oxidative stress responses of some endemic plants to high altitudes by intensifying antioxidants and secondary metabolites content. *Plants* **9** (**7**): 869. doi: 10.3390/plants9070869.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Havaux, M. (2014). Carotenoid oxidation products as stress signals in plants. *The Plant Journal* **79** (**4**): 597–606. doi: 10.1111/tpj.12386.

He, M. and Dijkstra, F.A. (2014). Drought effect on plant nitrogen and phosphorus: a meta-analysis. *New Phytologist* **204** (**4**): 924–931. doi: 10.1111/nph.12952.

Hennion, F., Bouchereau, A., Gauthier, C. et al. (2012). Variation in amine composition in plant species: how it integrates macroevolutionary and environmental signals. *American Journal of Botany* **99** (**1**): 36–45. doi: 10.3732/ajb.1100211.

Hill, K.E., Guerin, G.R., Hill, R.S. et al. (2015). Temperature influences stomatal density and maximum potential water loss through stomata of *Dodonaea viscosa* subsp. angustissima along a latitude gradient in southern Australia. *Australian Journal of Botany* **62** (**8**): 657–665.

Hou, Q., Ufer, G., and Bartels, D. (2016). Lipid signalling in plant responses to abiotic stress: lipid signalling in plant responses to abiotic stress. *Plant, Cell & Environment* **39** (**5**): 1029–1048. doi: 10.1111/pce.12666.

Hu, L., Hu, T., Zhang, X. et al. (2012). Exogenous glycine betaine Ameliorates the adverse effect of salt stress on perennial ryegrass. *Journal of the American Society for Horticultural Science* **137** (**1**): 38–46. doi: 10.21273/JASHS.137.1.38.

Igamberdiev, A.U. and Eprintsev, A.T. (2016). Organic acids: the pools of fixed carbon involved in redox regulation and energy balance in higher plants. *Frontiers in Plant Science* **7**. doi: 10.3389/fpls.2016.01042.

Jaakola, L. and Hohtola, A. (2010). Effect of latitude on flavonoid biosynthesis in plants: effect of latitude on flavonoid biosynthesis. *Plant, Cell and Environment*. doi: 10.1111/j.1365-3040.2010.02154.x.

James, J.J., Tiller, R.L., and Richards, J.H. (2005). Multiple resources limit plant growth and function in a saline-alkaline desert community. *Journal of Ecology* **93** (**1**): 113–126. doi: 10.1111/j.0022-0477.2004.00948.x.

Jiang, G., Wang, Z., Shang, H. et al. (2007). Proteome analysis of leaves from the resurrection plant *Boea hygrometrica* in response to dehydration and rehydration. *Planta* **225** (**6**): 1405–1420. doi: 10.1007/s00425-006-0449-z.

Jin, D., Dai, Y., Sun, L. et al. (2008). Is mass-based metabolism rate proportional to surface area in plant leaves? A data re-analysis. *Journal of Integrative Plant Biology* **50** (**6**): 673–681. doi: 10.1111/j.1744-7909.2008.00674.x.

Jones, R.G.W. and Gorham, J. (1983). Aspects of salt and drought tolerance in higher plants. In: *Genetic Engineering of Plants* (eds. T. Kosuge, C.P. Meredith, A. Hollaender and C.M. Wilson), 355–370. Boston, MA: Springer. doi: 10.1007/978-1-4684-4544-2_24.

Jordan, P.W. and Nobel, P.S. (1979). Infrequent Establishment of Seedlings of Agave Deserti (Agavaceae) in the Northwestern Sonoran Desert. *American Journal of Botany* **66** (**9**): 1079–1084. doi: 10.1002/j.1537-2197.1979.tb06325.x.

Kalia, R., Sareen, S., Nagpal, A. et al. (2017). ROS-induced transcription factors during oxidative stress in plants: a tabulated review. In: *Reactive Oxygen Species and Antioxidant Systems in Plants: Role and Regulation under Abiotic Stress* (eds. M.I.R. Khan and N.A. Khan), 129–158. Singapore: Springer. doi: 10.1007/978-981-10-5254-5_6.

Karabourniotis, G., Kotsabassidis, D., and Manetas, Y. (1995). Trichome density and its protective potential against ultraviolet-B radiation damage during leaf development. *Canadian Journal of Botany* **73** (**3**): 376–383.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Kaur, G. and Asthir, B. (2015). Proline: a key player in plant abiotic stress tolerance. *Biologia Plantarum* **59** (**4**): 609–619. doi: 10.1007/s10535-015-0549-3.

van Kleunen, M. and Richardson, D.M. (2007). Invasion biology and conservation biology: time to join forces to explore the links between species traits and extinction risk and invasiveness. *Progress in Physical Geography: Earth and Environment* **31** (**4**): 447–450. doi: 10.1177/0309133307081295.

Körner, C. (2003). *Alpine Plant Life: Functional Plant Ecology of High Mountain Ecosystems*, 2e. Berlin, New York: Springer.

Körner, C. (2016). Plant adaptation to cold climates. *F1000Research* **5**: 2769. doi: 10.12688/f1000research.9107.1.

Körner, C. (2018). Concepts in empirical plant ecology. *Plant Ecology and Diversity* **11** (**4**): 405–428. doi: 10.1080/17550874.2018.1540021.

Körner, C., Bannister, P., and Mark, A.F. (1986). Altitudinal variation in stomatal conductance, nitrogen content and leaf anatomy in different plant life forms in New Zealand. *Oecologia* **69** (**4**): 577–588. doi: 10.1007/BF00410366.

Körner, C., Neumayer, M., Menendez-Riedl, S.P. et al. (1989). Functional morphology of mountain plants Dedicated to Prof. H. Meusel, on the occasion of his 80th birthday. *Flora* **182** (**5–6**): 353–383. doi: 10.1016/S0367-2530(17)30426-7.

Krall, J.P. and Edwards, G.E. (1992). Relationship between photosystem II activity and CO_2 fixation in leaves. *Physiologia Plantarum* **86** (**1**): 180–187. doi: 10.1111/j.1399-3054.1992.tb01328.x.

Krämer, U. (2010). Metal hyperaccumulation in plants. *Annual Review of Plant Biology* **61** (**1**): 517–534. doi: 10.1146/annurev-arplant-042809-112156.

Kranner, I., Beckett, R.P., Wornik, S. et al. (2002). Revival of a resurrection plant correlates with its antioxidant status. *The Plant Journal* **31** (**1**): 13–24. doi: 10.1046/j.1365-313X.2002.01329.x.

Krishnamurthy, P., Jyothi-Prakash, P.A., Qin, L. et al. (2014). Role of root hydrophobic barriers in salt exclusion of a mangrove plant *Avicennia officinalis*. *Plant, Cell & Environment* **37** (**7**): 1656–1671. doi: 10.1111/pce.12272.

Kroymann, J. (2011). Natural diversity and adaptation in plant secondary metabolism. *Current Opinion in Plant Biology* **14** (**3**): 246–251. doi: 10.1016/j.pbi.2011.03.021.

Kumari, M., Joshi, R., and Kumar, R. (2020). Metabolic signatures provide novel insights to *Picrorhiza kurroa* adaptation along the altitude in Himalayan region. *Metabolomics* **16** (**7**): 77. doi: 10.1007/s11306-020-01698-8.

Kuprian, E., Tuong, T.D., Pfaller, K. et al. (2016). Persistent supercooling of reproductive shoots is enabled by structural ice barriers being active despite an intact xylem connection. *PLoS One* **11** (**9**). doi: 10.1371/journal.pone.0163160.

Lambers, H., Brundrett, M.C., Raven, J.A. et al. (2011). Plant mineral nutrition in ancient landscapes: high plant species diversity on infertile soils is linked to functional diversity for nutritional strategies. *Plant and Soil* **348** (**1–2**): 7–27. doi: 10.1007/s11104-011-0977-6.

Lambers, H., Ahmedi, I., Berkowitz, O. et al. (2013). Phosphorus nutrition of phosphorus-sensitive Australian native plants: threats to plant communities in a global biodiversity hotspot. *Conservation Physiology* **1** (**1**): cot010. doi: 10.1093/conphys/cot010.

Larcher, W., Kainmüller, C., and Wagner, J. (2010). Survival types of high mountain plants under extreme temperatures. *Flora – Morphology, Distribution, Functional Ecology of Plants* **205** (**1**): 3–18. doi: 10.1016/j.flora.2008.12.005.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Latorre, C., González, A.L., Quade, J. et al. (2011). Establishment and formation of fog-dependent *Tillandsia landbeckii* dunes in the Atacama desert: evidence from radiocarbon and stable isotopes. *Journal of Geophysical Research: Biogeosciences* **116** (**G3**). doi: 10.1029/2010jg001521.

Le Gall, H., Philippe, F., Domon, J.-M. et al. (2015). Cell wall metabolism in response to abiotic stress. *Plants* **4** (**1**): 112–166. doi: 10.3390/plants4010112.

León-Sobrino, C., Ramond, J.-B., Maggs-Kölling, G. et al. (2019). Nutrient acquisition, rather than stress response over Diel cycles, drives microbial transcription in a hyper-arid Namib desert soil. *Frontiers in Microbiology* **10**: 1054. doi: 10.3389/fmicb.2019.01054.

Levin, N.E., Cerling, T.E., Passey, B.H. et al. (2006). A stable isotope aridity index for terrestrial environments. *Proceedings of the National Academy of Sciences* **103** (**30**): 11201–11205. doi: 10.1073/pnas.0604719103.

Levitt, J. (1980). Responses of plants to environmental stress. In: *Volume 1: Chilling, Freezing, and High Temperature Stresses*, 2e, 3–19. Academic Press. doi: 10.1016/B978-0-12-445501-6.50006-3.

Li, F.-R., Peng, S.-L., Chen, B.-M. et al. (2010). A meta-analysis of the responses of woody and herbaceous plants to elevated ultraviolet-B radiation. *Acta Oecologica* **36** (**1**): 1–9. doi: 10.1016/j.actao.2009.09.002.

Li, X., Yang, Y., Ma, L. et al. (2014). Comparative proteomics analyses of *Kobresia pygmaea* adaptation to environment along an elevational gradient on the Central Tibetan Plateau. *PLoS One* **9** (**6**): e98410. doi: 10.1371/journal.pone.0098410.

Li, C., Zheng, Z., Peng, Y. et al. (2019). Precipitation and nitrogen addition enhance biomass allocation to aboveground in an alpine steppe. *Ecology and Evolution* **9** (**21**): 12193–12201. doi: 10.1002/ece3.5706.

Li, Y., Kong, D., Fu, Y. et al. (2020). The effect of developmental and environmental factors on secondary metabolites in medicinal plants. *Plant Physiology and Biochemistry* **148**: 80–89. doi: 10.1016/j.plaphy.2020.01.006.

Lin, Y.-F. and Aarts, M.G.M. (2012). The molecular mechanism of zinc and cadmium stress response in plants. *Cellular and Molecular Life Sciences* **69** (**19**): 3187–3206. doi: 10.1007/s00018-012-1089-z.

Liu, M.Z., Jiang, G.M., Li, Y.G. et al. (2003). Leaf osmotic potentials of 104 plant species in relation to habitats and plant functional types in Hunshandak Sandland, Inner Mongolia, China. *Trees* **17** (**6**): 554–560. doi: 10.1007/s00468-003-0277-8.

Loescher, W.H. and Everard, J.D. (2000). Regulation of sugar alcohol biosynthesis. In: *Photosynthesis* (eds. R.C. Leegood, T.D. Sharkey and S. von Caemmerer), 275–299. Dordrecht, Netherlands: Springer. doi: 10.1007/0-306-48137-5_12.

Long, S.P., Marshall-Colon, A., and Zhu, X.-G. (2015). Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* **161** (**1**): 56–66. doi: 10.1016/j.cell.2015.03.019.

López-Fernández, O., Domínguez, R., Pateiro, M. et al. (2020). Determination of polyphenols using liquid chromatography–tandem mass spectrometry technique (LC–MS/MS): a review. *Antioxidants* **9** (**6**): 479. doi: 10.3390/antiox9060479.

López-Lozano, N.E., Eguiarte, L.E., Bonilla-Rosso, G. et al. (2012). Bacterial communities and the nitrogen cycle in the gypsum soils of Cuatro Ciénegas Basin, Coahuila: a Mars analogue. *Astrobiology* **12** (**7**): 699–709. doi: 10.1089/ast.2012.0840.

Loreti, E., van Veen, H., and Perata, P. (2016). Plant responses to flooding stress. *Current Opinion in Plant Biology* **33**: 64–71. doi: 10.1016/j.pbi.2016.06.005.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Lü, X.-T., Freschet, G.T., Flynn, D.F.B. et al. (2012). Plasticity in leaf and stem nutrient resorption proficiency potentially reinforces plant-soil feedbacks and microscale heterogeneity in a semi-arid grassland: fertility-nutrient resorption coupling. *Journal of Ecology* **100** (**1**): 144–150. doi: 10.1111/j.1365-2745.2011.01881.x.

Lugan, R., Niogret, M.-F., Leport, L. et al. (2010). Metabolome and water homeostasis analysis of *Thellungiella salsuginea* suggests that dehydration tolerance is a key response to osmotic stress in this halophyte. *The Plant Journal: For Cell and Molecular Biology* **64** (**2**): 215–229. doi: 10.1111/j.1365-313X.2010.04323.x.

Ma, W., Shi, P., Li, W. et al. (2010). Changes in individual plant traits and biomass allocation in alpine meadow with elevation variation on the Qinghai-Tibetan Plateau. *Science China Life Sciences* **53** (**9**): 1142–1151. doi: 10.1007/s11427-010-4054-9.

Ma, L., Sun, X., Kong, X. et al. (2015). Physiological, biochemical and proteomics analysis reveals the adaptation strategies of the alpine plant *Potentilla saundersiana* at altitude gradient of the Northwestern Tibetan Plateau. *Journal of Proteomics* **112**: 63–82. doi: 10.1016/j.jprot.2014.08.009.

Martin, C.E. and von Willert, D.J. (2000). Leaf epidermal hydathodes and the ecophysiological consequences of foliar water uptake in species of *Crassula* from the Namib desert in southern Africa. *Plant Biology* **2** (**2**): 229–242. doi: 10.1055/s-2000-9163.

Martins, L., Trujillo-Hernandez, J.A., and Reichheld, J.-P. (2018). Thiol based redox signaling in plant nucleus. *Frontiers in Plant Science* **9**. doi: 10.3389/fpls.2018.00705.

Matthews, J.S. and Lawson, T. (2018). Climate change and stomatal physiology. *Annual Plant Reviews Online*: 713–752.

Mazliak, P. (1977). Glyco- and phospholipids of biomembranes in higher plants. In: *Lipids and Lipid Polymers in Higher Plants* (eds. M. Tevini and H.K. Lichtenthaler), 48–74. Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-66632-2_3.

Merino, N., Aronson, H.S., Bojanova, D.P. et al. (2019). Living at the extremes: extremophiles and the limits of life in a planetary context. *Frontiers in Microbiology* **10**: 780. doi: 10.3389/fmicb.2019.00780.

Mhlongo, M.I., Piater, L.A., Madala, N.E. et al. (2018). The chemistry of plant–microbe interactions in the rhizosphere and the potential for metabolomics to reveal signaling related to defense priming and induced systemic resistance. *Frontiers in Plant Science* **9**. doi: 10.3389/fpls.2018.00112.

Mikulic-Petkovsek, M., Schmitzer, V., Slatnar, A. et al. (2012). Composition of sugars, organic acids, and total phenolics in 25 wild or cultivated Berry species. *Journal of Food Science* **77** (**10**): C1064–C1070. doi: 10.1111/j.1750-3841.2012.02896.x.

Mishra, A.K. and Singh, V.P. (2010). A review of drought concepts. *Journal of Hydrology* **391** (**1**): 202–216. doi: 10.1016/j.jhydrol.2010.07.012.

Mittler, R. (2006). Abiotic stress, the field environment and stress combination. *Trends in Plant Science* **11** (**1**): 15–19. doi: 10.1016/j.tplants.2005.11.002.

Mittler, R. (2017). ROS are good. *Trends in Plant Science* **22** (**1**): 11–19. doi: 10.1016/j.tplants.2016.08.002.

Moellering, E.R., Muthan, B., and Benning, C. (2010). Freezing tolerance in plants requires lipid remodeling at the outer chloroplast membrane. *Science* **330** (**6001**): 226–228. doi: 10.1126/science.1191803.

Moing, A. (2000). Sugar alcohols as carbohydrate reserves in some higher plants. *Developments in Crop Science*. Elsevier **26**: 337–358. doi: 10.1016/S0378-519X(00)80017.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Monschein, M., Jaindl, K., Buzimkić, S. et al. (2015). Content of phenolic compounds in wild populations of *Epilobium angustifolium* growing at different altitudes. *Pharmaceutical Biology* **53** (**11**): 1576–1582. doi: 10.3109/13880209.2014.993039.

Moon, G., Clough, B., Peterson, C. et al. (1986). Apoplastic and symplastic pathways in *Avicennia marina* (Forsk.) Vierh. roots revealed by fluorescent tracer dyes. *Functional Plant Biology* **13** (**5**): 637–648.

Munns, R. (2002). Comparative physiology of salt and water stress: comparative physiology of salt and water stress. *Plant, Cell & Environment* **25** (**2**): 239–250. doi: 10.1046/j.0016-8025.2001.00808.x.

Nagajyoti, P.C., Lee, K.D., and Sreekanth, T.V.M. (2010). Heavy metals, occurrence and toxicity for plants: a review. *Environmental Chemistry Letters* **8** (**3**): 199–216. doi: 10.1007/s10311-010-0297-8.

Nagler, M., Nägele, T., Gilli, C. et al. (2018). Eco-metabolomics and metabolic modeling: making the leap from model systems in the lab to native populations in the field. *Frontiers in Plant Science* **9**: 1556. doi: 10.3389/fpls.2018.01556.

Najjaa, H., Abdelkarim, B.A., Doria, E. et al. (2020). Phenolic composition of some Tunisian medicinal plants associated with anti-proliferative effect on human breast cancer MCF-7 cells. *The EuroBiotech Journal* **4** (**2**): 104–112. doi: 10.2478/ebtj-2020-0012.

Nakabayashi, R. and Saito, K. (2015). Integrated metabolomics for abiotic stress responses in plants. *Current Opinion in Plant Biology* **24**: 10–16. doi: 10.1016/j.pbi.2015.01.003.

Neuner, G. and Hacker, J. (2012). Ice formation and propagation in alpine plants. In: *Plants in Alpine Regions: Cell Physiology of Adaption and Survival Strategies* (ed. C. Lütz), 163–174. Springer: Vienna. doi: 10.1007/978-3-7091-0136-0_12.

Nevo, E. (2001). Evolution of genome-phenome diversity under environmental stress. *Proceedings of the National Academy of Sciences* **98** (**11**): 6233–6240. doi: 10.1073/pnas.101109298.

Nicotra, A.B., Atkin, O.K., Bonser, S.P. et al. (2010). Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* **15** (**12**): 684–692. doi: 10.1016/j.tplants.2010.09.008.

Noctor, G., Reichheld, J.-P., and Foyer, C.H. (2018). ROS-related redox regulation and signaling in plants. *Seminars in Cell & Developmental Biology* **80**: 3–12. doi: 10.1016/j.semcdb.2017.07.013.

Noy-Meir, I. (1973). Desert ecosystems: environment and producers. *Annual Review of Ecology and Systematics* **4** (**1**): 25–51. doi: 10.1146/annurev.es.04.110173.000325.

Öncel, I., Yurdakulol, E., Keleş, Y. et al. (2004). Role of antioxidant defense system and biochemical adaptation on stress tolerance of high mountain and steppe plants. *Acta Oecologica* **26** (**3**): 211–218. doi: 10.1016/j.actao.2004.04.004.

Pandey, S., Fartyal, D., Agarwal, A. et al. (2017). Abiotic stress tolerance in plants: Myriad roles of ascorbate peroxidase. *Frontiers in Plant Science* **8**. doi: 10.3389/fpls.2017.00581.

Parida, A.K., Panda, A., and Rangani, J. (2018). Metabolomics-guided elucidation of abiotic stress tolerance mechanisms in plants. In: *Plant Metabolites and Regulation Under Environmental Stress* (eds. P. Ahmad, M.A. Ahanger, V. Singh, et al.), 89–131. Elsevier. doi: 10.1016/B978-0-12-812689-9.00005-4.

Pearce, R.S. (2001). Plant freezing and damage. *Annals of Botany* **87** (**4**): 417–424. doi: 10.1006/anbo.2000.1352.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Peng, J.-S., Guan, Y.-H., Lin, X.-J. et al. (2020). Comparative understanding of metal hyperaccumulation in plants: a mini-review. *Environmental Geochemistry and Health*. doi: 10.1007/s10653-020-00533-2.

Peppe, D.J., Royer, D.L., Cariglino, B. et al. (2011). Sensitivity of leaf size and shape to climate: global patterns and paleoclimatic applications. *New Phytologist* **190** (**3**): 724–739. doi: 10.1111/j.1469-8137.2010.03615.x.

Pérez-Noyola, F.J., Flores, J., Yáñez-Espinosa, L. et al. (2020). Complete vivipary behavior detected in the epiphytic *Tillandsia recurvata* L.(Ball moss) in the Chihuahuan Desert in two continuous years. *Journal of Arid Environments* **174**: 103993.

Peters, K., Worrich, A., Weinhold, A. et al. (2018). Current challenges in plant eco-metabolomics. *International Journal of Molecular Sciences* **19** (**5**): 1385. doi: 10.3390/ijms19051385.

Piacentini, R.D., Cede, A., and Bárcena, H. (2003). Extreme solar total and UV irradiances due to cloud effect measured near the summer solstice at the high-altitude desertic plateau Puna of Atacama (Argentina). *Journal of Atmospheric and Solar-Terrestrial Physics* **65** (**6**): 727–731. doi: 10.1016/S1364-6826(03)00084-1.

Pollard, A.J., Reeves, R.D., and Baker, A.J.M. (2014). Facultative hyperaccumulation of heavy metals and metalloids. *Plant Science* **217–218**: 8–17. doi: 10.1016/j.plantsci.2013.11.011.

Poorter, H., Niinemets, Ü., Walter, A. et al. (2010). A method to construct dose–response curves for a wide range of environmental factors and plant traits by means of a meta-analysis of phenotypic data. *Journal of Experimental Botany* **61** (**8**): 2043–2055. doi: 10.1093/jxb/erp358.

Poorter, H., Niklas, K.J., Reich, P.B. et al. (2012). Biomass allocation to leaves, stems and roots: meta-analyses of interspecific variation and environmental control: Tansley review. *New Phytologist* **193** (**1**): 30–50. doi: 10.1111/j.1469-8137.2011.03952.x.

Poorter, H., Jagodzinski, A.M., Ruiz-Peinado, R. et al. (2015). How does biomass distribution change with size and differ among species? An analysis for 1200 plant species from five continents. *New Phytologist* **208** (**3**): 736–749. doi: 10.1111/nph.13571.

Poorter, H., Niinemets, Ü., Ntagkas, N. et al. (2019). A meta-analysis of plant responses to light intensity for 70 traits ranging from molecules to whole plant performance. *New Phytologist* **223** (**3**): 1073–1105. doi: 10.1111/nph.15754.

Preece, C., Livarda, A., Christin, P. et al. (2017). How did the domestication of Fertile Crescent grain crops increase their yields? *Functional Ecology* **31** (**2**): 387–397. doi: 10.1111/1365-2435.12760.

Purnobasuki, H. (2011). Structure of lenticels on the pneumatophores of *Avicennia marina*: as aerating device deliver oxygen in mangrove's root. *Jurnal BIOTA XVI* **2**: 309–315.

Pyankov, V.I., Kondratchuk, A.V., and Shipley, B. (1999). Leaf structure and specific leaf mass: the alpine desert plants of the Eastern Pamirs, Tadjikistan. *New Phytologist* **143** (**1**): 131–142. doi: 10.1046/j.1469-8137.1999.00435.x.

Qu, A.-L., Ding, Y.-F., Jiang, Q. et al. (2013). Molecular mechanisms of the plant heat stress response. *Biochemical and Biophysical Research Communications* **432** (**2**): 203–207. doi: 10.1016/j.bbrc.2013.01.104.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Quartacci, M.F. (2002). Plasma membrane lipids in the resurrection plant *Ramonda serbica* following dehydration and rehydration. *Journal of Experimental Botany* **53** (**378**): 2159–2166. doi: 10.1093/jxb/erf076.

Ramadan, A., Sabir, J.S.M., Alakilli, S.Y.M. et al. (2014). Metabolomic response of *Calotropis procera* growing in the desert to changes in water availability. *PLoS One* **9** (**2**). doi: 10.1371/journal.pone.0087895.

Ramond, J.-B., Woodborne, S., Hall, G. et al. (2018). Namib Desert primary productivity is driven by cryptic microbial community N-fixation. *Scientific Reports* **8** (**1**): 6921. doi: 10.1038/s41598-018-25078-4.

Rascio, N. and Rocca, N.L. (2005). Resurrection plants: the puzzle of surviving extreme vegetative desiccation. *Critical Reviews in Plant Sciences* **24** (**3**): 209–225. doi: 10.1080/07352680591008583.

Reich, P.B., Walters, M.B., and Ellsworth, D.S. (1992). Leaf life-span in relation to leaf, plant, and stand characteristics among diverse ecosystems. *Ecological Monographs* **62** (**3**): 365–392. doi: 10.2307/2937116.

Ren, M., Zhang, Z., Wang, X. et al. (2018). Diversity and contributions to nitrogen cycling and carbon fixation of soil salinity shaped microbial communities in Tarim basin. *Frontiers in Microbiology* **9**: 431. doi: 10.3389/fmicb.2018.00431.

Richards, C.L., Bossdorf, O., Muth, N.Z. et al. (2006). Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters* **9** (**8**): 981–993. doi: 10.1111/j.1461-0248.2006.00950.x.

Riehl, S., Benz, M., Conard, N.J. et al. (2012). Plant use in three Pre-Pottery Neolithic sites of the northern and eastern Fertile Crescent: a preliminary report. *Vegetation History and Archaeobotany* **21** (**2**): 95–106. doi: 10.1007/s00334-011-0318-y.

Roch, L., Prigent, S., Klose, H. et al. (2020). Biomass composition explains fruit relative growth rate and discriminates climacteric from non-climacteric species. *Journal of Experimental Botany*: eraa302. doi: 10.1093/jxb/eraa302.

Rothschild, L.J. and Mancinelli, R.L. (2001). Life in extreme environments. *Nature* **409**: 1092–1101. doi: 10.1038/35059215.

Rundel, P.W., Palma, B., Dillon, M. et al. (1997). *Tillandsia landbeckii* in the coastal Atacama desert of northern Chile. *Revista Chilena de Historia Natural* **70**: 341–349.

Saha, J., Brauer, E.K., Sengupta, A. et al. (2015). Polyamines as redox homeostasis regulators during salt stress in plants. *Frontiers in Environmental Science* **3**. doi: 10.3389/fenvs.2015.00021.

Saha, B., Borovskii, G., and Panda, S.K. (2016). Alternative oxidase and plant stress tolerance. *Plant Signaling & Behavior* **11** (**12**): e1256530. doi: 10.1080/15592324.2016.1256530.

Sanchez, D.H., Pieckenstain, F.L., Escaray, F. et al. (2011). Comparative ionomics and metabolomics in extremophile and glycophytic Lotus species under salt stress challenge the metabolic pre-adaptation hypothesis. *Plant, Cell & Environment* **34** (**4**): 605–617. doi: 10.1111/j.1365-3040.2010.02266.x.

Sardans, J. and Peñuelas, J. (2005). Drought decreases soil enzyme activity in a Mediterranean *Quercus ilex* L. forest. *Soil Biology and Biochemistry* **37** (**3**): 455–461. doi: 10.1016/j.soilbio.2004.08.004.

Sardans, J., Gargallo-Garriga, A., Urban, O. et al. (2020). Ecometabolomics for a better understanding of plant responses and acclimation to abiotic factors linked to global change. *Metabolites* **10** (**6**): 239. doi: 10.3390/metabo10060239.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Schertl, P. and Braun, H.-P. (2014). Respiratory electron transfer pathways in plant mitochondria. *Frontiers in Plant Science* **5**. doi: 10.3389/fpls.2014.00163.

Schimel, J.P. (2018). Life in dry soils: effects of drought on soil microbial communities and processes. *Annual Review of Ecology, Evolution, and Systematics* **49** (**1**): 409–432. doi: 10.1146/annurev-ecolsys-110617-062614.

Scholander, P.F., van Dam, L., and Scholander, S.I. (1955). Gas exchange in the roots of mangroves. *American Journal of Botany* **42** (**1**): 92–98. doi: 10.2307/2438597.

Schulz, E., Tohge, T., Zuther, E. et al. (2016). Flavonoids are determinants of freezing tolerance and cold acclimation in *Arabidopsis thaliana*. *Scientific Reports* **6** (**1**): 34027. doi: 10.1038/srep34027.

Schulze, E.-D., Hall, A.E., Lange, O.L. et al. (1980). Long-term effects of drought on wild and cultivated plants in the Negev desert: I. Maximal rates of net photosynthesis. *Oecologia* **45** (**1**): 11–18. doi: 10.1007/BF00346700.

Schuster, A.-C., Burghardt, M., Alfarhan, A. et al. (2016). Effectiveness of cuticular transpiration barriers in a desert plant at controlling water loss at high temperatures. *AoB Plants* **8**: plw027. doi: 10.1093/aobpla/plw027.

Seemann, J.R., Downton, W.J.S., and Berry, J.A. (1986). Temperature and leaf osmotic potential as factors in the acclimation of photosynthesis to high temperature in desert plants. *Plant Physiology* **80** (**4**): 926–930. doi: 10.1104/pp. 80.4.926.

Selye, H. (1950). Stress and the general adaptation syndrome. *British Medical Journal* **1** (**4667**): 1383–1392.

Sevilla, F., Jiménez, A., and Lázaro, J.J. (2015). What do the plant mitochondrial antioxidant and redox systems have to say under salinity, drought, and extreme temperature? In: *Reactive Oxygen Species and Oxidative Damage in Plants Under Stress* (eds. D.K. Gupta, J.M. Palma and F.J. Corpas), 23–55. Springer International Publishing. doi: 10.1007/978-3-319-20421-5_2.

Shabala, L., Mackay, A., Tian, Y. et al. (2012). Oxidative stress protection and stomatal patterning as components of salinity tolerance mechanism in quinoa (*Chenopodium quinoa*). *Physiologia Plantarum* **146** (**1**): 26–38. doi: 10.1111/j.1399-3054.2012.01599.x.

Shahid, S.A., Zaman, M., and Heng, L. (2018). Soil salinity: historical perspectives and a world overview of the problem. In: *Guideline for Salinity Assessment, Mitigation and Adaptation Using Nuclear and Related Techniques* (eds. M. Zaman, S.A. Shahid and L. Heng), 43–53. Springer International Publishing. doi: 10.1007/978-3-319-96190-3_2.

Shanker, A., Cervantes, C., Lozatavera, H. et al. (2005). Chromium toxicity in plants. *Environment International* **31** (**5**): 739–753. doi: 10.1016/j.envint.2005.02.003.

Sharma, A., Shahzad, B., Kumar, V. et al. (2019). Phytohormones regulate accumulation of osmolytes under abiotic stress. *Biomolecules* **9** (**7**): 285. doi: 10.3390/biom9070285.

Sharma, B., Arora, S., Sahoo, D. et al. (2020). Comparative fatty acid profiling of Indian seabuckthorn showed altitudinal gradient dependent species-specific variations. *Physiology and Molecular Biology of Plants* **26** (**1**): 41–49. doi: 10.1007/s12298-019-00720-1.

Singh, H.P., Mahajan, P., Kaur, S. et al. (2013). Chromium toxicity and tolerance in plants. *Environmental Chemistry Letters* **11** (**3**): 229–254. doi: 10.1007/s10311-013-0407-5.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Singh, M., Kumar, J., Singh, S. et al. (2015). Roles of osmoprotectants in improving salinity and drought tolerance in plants: a review. *Reviews in Environmental Science and BioTechnology* **14** (**3**): 407–426. doi: 10.1007/s11157-015-9372-8.

Singh, S., Parihar, P., Singh, R. et al. (2016). Heavy metal tolerance in plants: role of transcriptomics, proteomics, metabolomics, and ionomics. *Frontiers in Plant Science* **6**. doi: 10.3389/fpls.2015.01143.

Slama, I., Abdelly, C., Bouchereau, A. et al. (2015). Diversity, distribution and roles of osmoprotective compounds accumulated in halophytes under abiotic stress. *Annals of Botany* **115** (**3**): 433–447. doi: 10.1093/aob/mcu239.

Smirnoff, N. and Arnaud, D. (2019). Hydrogen peroxide metabolism and functions in plants. *New Phytologist* **221** (**3**): 1197–1214. doi: 10.1111/nph.15488.

Stapleton, A.E. and Walbot, V. (1994). Flavonoids can protect maize DNA from the induction of ultraviolet radiation damage. *Plant Physiology* **105** (**3**): 881–889. doi: 10.1104/pp. 105.3.881.

Streb, P., Tel-Or, E., and Feierabend, J. (1997). Light stress effects and antioxidative protection in two desert plants. *Functional Ecology* **11** (**4**): 416–424. doi: 10.1046/j .1365-2435.1997.00105.x.

Strimbeck, G.R., Schaberg, P.G., Fossdal, C.G. et al. (2015). Extreme low temperature tolerance in woody plants. *Frontiers in Plant Science* **6**. doi: 10.3389/fpls.2015.00884.

Stushnoff, C., Seufferheld, M.J., and Creegan, T. (1997). Oligosaccharides as endogenous cryoprotectants in woody plants. In: *Plant Cold Hardiness* (eds. P.H. Li and T.H.H. Chen), 301–309. Boston, MA: Springer. doi: 10.1007/978-1-4899-0277-1_27.

Sun, J. and Wang, H. (2016). Soil nitrogen and carbon determine the trade-off of the above- and below-ground biomass across alpine grasslands, Tibetan Plateau. *Ecological Indicators* **60**: 1070–1076. doi: 10.1016/j.ecolind.2015.08.038.

Sundberg, M.D. (1986). A comparison of stomatal distribution and length in succulent and non-succulent desert plants. *Phytomorphology* **36** (**1–2**): 53–66.

Suzuki, N., Rivero, R.M., Shulaev, V. et al. (2014). Abiotic and biotic stress combinations. *New Phytologist* **203** (**1**): 32–43. doi: 10.1111/nph.12797.

Sweetlove, L.J., Beard, K.F.M., Nunes-Nesi, A. et al. (2010). Not just a circle: flux modes in the plant TCA cycle. *Trends in Plant Science* **15** (**8**): 462–470. doi: 10.1016/j.tplants .2010.05.006.

Szabados, L. and Savouré, A. (2010). Proline: a multifunctional amino acid. *Trends in Plant Science* **15** (**2**): 89–97. doi: 10.1016/j.tplants.2009.11.009.

Talbi, S., Romero-Puertas, M.C., Hernández, A. et al. (2015). Drought tolerance in a Saharian plant *Oudneya africana*: role of antioxidant defences. *Environmental and Experimental Botany* **111**: 114–126. doi: 10.1016/j.envexpbot.2014.11.004.

Tapia-Torres, Y., Rodríguez-Torres, M.D., Elser, J.J. et al. (2016). How to live with phosphorus scarcity in soil and sediment: lessons from bacteria. *Applied and Environmental Microbiology* **82** (**15**): 4652–4662. doi: 10.1128/AEM.00160-16.

Tetali, S.D. (2019). Terpenes and isoprenoids: a wealth of compounds for global use. *Planta* **249** (**1**): 1–8. doi: 10.1007/s00425-018-3056-x.

Tholl, D. (2015). Biosynthesis and biological functions of terpenoids in plants. In: *Biotechnology of Isoprenoids* (eds. J. Schrader and J. Bohlmann), 63–106. Springer International Publishing. doi: 10.1007/10_2014_295.

Tipirdamaz, R., Gagneul, D., Duhazé, C. et al. (2006). Clustering of halophytes from an inland salt marsh in Turkey according to their ability to accumulate sodium and

nitrogenous osmolytes. *Environmental and Experimental Botany* **57** (**1**): 139–153. doi: 10.1016/j.envexpbot.2005.05.007.

Tomlinson, P.B. and Cox, P.A. (2000). Systematic and functional anatomy of seedlings in mangrove Rhizophoraceae: vivipary explained? *Botanical Journal of the Linnean Society*, Oxford Academic **134** (**1–2**): 215–231. doi: 10.1111/j.1095-8339.2000.tb02352.x.

Tóth, G., Hermann, T., Da Silva, M.R. et al. (2016). Heavy metals in agricultural soils of the European Union with implications for food safety. *Environment International* **88**: 299–309. doi: 10.1016/j.envint.2015.12.017.

Tshabuse, F., Farrant, J.M., Humbert, L. et al. (2018). Glycerolipid analysis during desiccation and recovery of the resurrection plant *Xerophyta humilis* (Bak) Dur and Schinz. *Plant, Cell & Environment* **41** (**3**): 533–547. doi: 10.1111/pce.13063.

Tugizimana, F., Mhlongo, M.I., Piater, L.A. et al. (2018). Metabolomics in plant priming research: the way forward? *International Journal of Molecular Sciences* **19** (**6**): 1759. doi: 10.3390/ijms19061759.

Tůmová, L., Tarkowská, D., Řehořová, K. et al. (2018). Drought-tolerant and drought-sensitive genotypes of maize (*Zea mays* L.) differ in contents of endogenous brassinosteroids and their drought-induced changes. *PLoS One* **13** (**5**): e0197870. doi: 10.1371/journal.pone.0197870.

Turner, N.C. (2018). Turgor maintenance by osmotic adjustment: 40 years of progress. *Journal of Experimental Botany* **69** (**13**): 3223–3233. doi: 10.1093/jxb/ery181.

Upchurch, R.G. (2008). Fatty acid unsaturation, mobilization, and regulation in the response of plants to stress. *Biotechnology Letters* **30** (**6**): 967–977. doi: 10.1007/s10529-008-9639-z.

Vance, C.P. (2001). Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world of declining renewable resources. *Plant Physiology* **127** (**2**): 390–397. doi: 10.1104/pp. 010331.

Verbruggen, N., Hermans, C., and Schat, H. (2009). Molecular mechanisms of metal hyperaccumulation in plants: Tansley review. *New Phytologist* **181** (**4**): 759–776. doi: 10.1111/j.1469-8137.2008.02748.x.

Verdaguer, D., Jansen, M.A.K., Llorens, L. et al. (2017). UV-A radiation effects on higher plants: exploring the known unknown. *Plant Science* **255**: 72–81. doi: 10.1016/j.plantsci.2016.11.014.

Vidiella, P.E., Armesto, J.J., and Gutiérrez, J.R. (1999). Vegetation changes and sequential flowering after rain in the southern Atacama desert. *Journal of Arid Environments* **43** (**4**): 449–458. doi: 10.1006/jare.1999.0565.

Von Willert, D., Brinckmann, E., Scheitler, B. et al. (1985). Availability of water controls Crassulacean acid metabolism in succulents of the Richtersveld (Namib desert, South Africa). *Planta* **164** (**1**): 44–55.

Wahid, A., Gelani, S., Ashraf, M. et al. (2007). Heat tolerance in plants: an overview. *Environmental and Experimental Botany* **61** (**3**): 199–223. doi: 10.1016/j.envexpbot.2007.05.011.

Wang, Y., An, Y., Yu, J. et al. (2016). Different responses of photosystem II and antioxidants to drought stress in two contrasting populations of Sour jujube from the Loess Plateau, China. *Ecological Research* **31** (**6**): 761–775. doi: 10.1007/s11284-016-1384-5.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Wang, L., Wang, L., He, W. et al. (2017). Nutrient resorption or accumulation of desert plants with contrasting sodium regulation strategies. *Scientific Reports* **7** (**1**): 17035. doi: 10.1038/s41598-017-17368-0.

Way, D.A. and Oren, R. (2010). Differential responses to changes in growth temperature between trees from different functional groups and biomes: a review and synthesis of data. *Tree Physiology* **30** (**6**): 669–688. doi: 10.1093/treephys/tpq015.

Whiting, S.N., Reeves, R.D., Richards, D. et al. (2004). Research priorities for conservation of metallophyte biodiversity and their potential for restoration and site remediation. *Restoration Ecology* **12** (**1**): 106–116. doi: 10.1111/j.1061-2971.2004.00367.x.

Wildi, B. and Lutz, C. (1996). Antioxidant composition of selected high alpine plant species from different altitudes. *Plant, Cell and Environment* **19** (**2**): 138–146. doi: 10.1111/j.1365-3040.1996.tb00235.x.

Williamson, C.E., Zepp, R.G., Lucas, R.M. et al. (2014). Solar ultraviolet radiation in a changing climate. *Nature Climate Change* **4** (**6**): 434–441. doi: 10.1038/nclimate2225.

Wisniewski, M. and Fuller, M. (1999). Ice nucleation and deep supercooling in plants: new insights using infrared thermography. In: *Cold-Adapted Organisms: Ecology, Physiology, Enzymology and Molecular Biology* (eds. R. Margesin and F. Schinner), 105–118. Berlin, Heidelberg: Springer. doi: 10.1007/978-3-662-06285-2_6.

Wright, I.J., Reich, P.B., Westoby, M. et al. (2004). The worldwide leaf economics spectrum. *Nature* **428** (**6985**): 821–827. doi: 10.1038/nature02403.

Wright, I.J., Dong, N., Maire, V. et al. (2017). Global climatic drivers of leaf size. *Science* **357** (**6354**): 917–921.

Xia, J. and Wan, S. (2008). Global response patterns of terrestrial plant species to nitrogen addition. *New Phytologist* **179** (**2**): 428–439. doi: 10.1111/j.1469-8137.2008.02488.x.

Xiong, F.S., Mueller, E.C., and Day, T.A. (2000). Photosynthetic and respiratory acclimation and growth response of Antarctic vascular plants to contrasting temperature regimes. *American Journal of Botany* **87** (**5**): 700–710. doi: 10.2307/2656856.

Xu, G.-Q., Li, Y., and Xu, H. (2011). Seasonal variation in plant hydraulic traits of two co-occurring desert shrubs, *Tamarix ramosissima* and *Haloxylon ammodendron*, with different rooting patterns. *Ecological Research* **26** (**6**): 1071–1080. doi: 10.1007/s11284-011-0858-8.

Yamori, W., Noguchi, K., Hikosaka, K. et al. (2009). Cold-tolerant crop species have greater temperature homeostasis of leaf respiration and photosynthesis than cold-sensitive species. *Plant and Cell Physiology* **50** (**2**): 203–215. doi: 10.1093/pcp/pcn189.

Yobi, A., Wone, B.W.M., Xu, W. et al. (2012). Comparative metabolic profiling between desiccation-sensitive and desiccation-tolerant species of *Selaginella* reveals insights into the resurrection trait: comparative metabolomics of desiccation tolerance. *The Plant Journal* **72** (**6**): 983–999. doi: 10.1111/tpj.12008.

Yobi, A., Wone, B.W.M., Xu, W. et al. (2013). Metabolomic profiling in *Selaginella lepidophylla* at various hydration states provides new insights into the mechanistic basis of desiccation tolerance. *Molecular Plant* **6** (**2**): 369–385. doi: 10.1093/mp/sss155.

Young, A. and Lowe, G. (2018). Carotenoids—antioxidant properties. *Antioxidants* **7** (**2**): 28. doi: 10.3390/antiox7020028.

Yruela, I. (2009). Copper in plants: acquisition, transport and interactions. *Functional Plant Biology* **36** (**5**): 409. doi: 10.1071/FP08288.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

Yuan, Y., Si, G., Wang, J. et al. (2014). Bacterial community in alpine grasslands along an altitudinal gradient on the Tibetan Plateau. *FEMS Microbiology Ecology* **87** (**1**): 121–132. doi: 10.1111/1574-6941.12197.

Zandalinas, S.I., Mittler, R., Balfagón, D. et al. (2018). Plant adaptations to the combination of drought and high temperatures. *Physiologia Plantarum* **162** (**1**): 2–12. doi: 10.1111/ppl.12540.

Zapata, P.J., Serrano, M., Pretel, M.T. et al. (2004). Polyamines and ethylene changes during germination of different plant species under salinity. *Plant Science* **167** (**4**): 781–788. doi: 10.1016/j.plantsci.2004.05.014.

Zhang, Z., Abuduwaili, J., and Jiang, F. (2013). Determination of occurrence characteristics of heavy metals in soil and water environments in Tianshan Mountains, Central Asia. *Analytical Letters* **46** (**13**): 2122–2131. doi: 10.1080/00032719.2013.784919.

Zhang, X.W., Wang, J.R., Ji, M.F. et al. (2015). Higher thermal acclimation potential of respiration but not photosynthesis in two Alpine Picea taxa in contrast to two lowland congeners. *PLoS One* **10** (**4**): e0123248. doi: 10.1371/journal.pone.0123248.

Zhang, K., Shi, Y., Cui, X. et al. (2019). Salinity is a key determinant for soil microbial communities in a desert ecosystem. *mSystems* **4** (**1**): e00225–e00218, /msystems/4/1/msys.00225-18.atom. doi: 10.1128/mSystems.00225-18.

Zhao, Y., Xu, F., Liu, J. et al. (2019). The adaptation strategies of *Herpetospermum pedunculosum* (Ser.) Baill at altitude gradient of the Tibetan plateau by physiological and metabolomic methods. *BMC Genomics* **20** (**1**): 451. doi: 10.1186/s12864-019-5778-y.

Zheng, G., Tian, B., Zhang, F. et al. (2011). Plant adaptation to frequent alterations between high and low temperatures: remodelling of membrane lipids and maintenance of unsaturation levels: lipid changes in high-low temperature alteration. *Plant, Cell & Environment* **34** (**9**): 1431–1442. doi: 10.1111/j.1365-3040.2011.02341.x.

Zhou, X., Zhang, Y., Ji, X. et al. (2011). Combined effects of nitrogen deposition and water stress on growth and physiological responses of two annual desert plants in northwestern China. *Environmental and Experimental Botany* **74**: 1–8. doi: 10.1016/j.envexpbot.2010.12.005.

Zidorn, C. (2010). Altitudinal variation of secondary metabolites in flowering heads of the Asteraceae: trends and causes. *Phytochemistry Reviews* **9** (**2**): 197–203. doi: 10.1007/s11101-009-9143-7.

Zlatev, Z.S., Lidon, J.C., and F. and Kaimakanova, M. (2012). Plant physiological responses to UV-B radiation. *Emirates Journal of Food and Agriculture* **24** (**6**). doi: 10.9755/ejfa.v24i6.14669.

Zou, T., Li, Y., Xu, H. et al. (2010). Responses to precipitation treatment for *Haloxylon ammodendron* growing on contrasting textured soils. *Ecological Research* **25** (**1**): 185–194. doi: 10.1007/s11284-009-0642-1.

Annual Plant Reviews Online, Volume 4. Edited by Jeremy Roberts.

ANNEX II.A1

Elevation (m.a.s.l)	CEext (mS/cm)	CEsusp (mS/cm)	Mn (mg/kg)	Cu (mg/kg)	Mo (mg/kg)	Mg (meq/100gr)	Cl (meq/L)	HCO3 (meq/L)	SO4 (meq/L)	Zn (mg/kg)	Na (meq/100gr)	B (mg/kg)
2470	0.650333333	0.135033333	0.397396392	1.448511433	2.011315058	0.801980652	1.25034517	2.628500271	3.229085591	0.1950541	0.302108155	3.673333333
2770	4.783333333	0.994333333	2.076129379	2.468382972	22.53898931	3.107199722	17.81692918	2.310896274	26.1768468	0.350999713	10.76401191	33.78666667
2870	0.4768	0.0788	1.634495296	2.621676601	1.82990847	0.702653889	0.762514051	1.644180804	1.839742607	0.300749448	1.061638545	2.523333333
2970	2.038	0.336333333	2.317031884	2.217274076	2.987919964	2.716741248	13.09146796	2.257750407	4.257776832	0.291051002	2.403141639	7.813333333
3070	0.538333333	0.099766667	2.818444007	3.112399853	1.400873527	1.254452258	1.111429611	2.309673871	2.035071554	0.342446334	0.794225683	2.983333333
3170	1.007	0.1664	4.140291122	3.095656875	2.943333333	2.714862012	3.502232361	1.792021071	5.084052686	0.404770343	1.059790528	3.743333333
3270	2.23	0.427	6.494634101	4.259703569	1.657085207	2.126035371	6.842191364	2.535028275	13.60602484	0.469483815	1.330679122	5.873333333
3370	0.961666667	0.164466667	5.987357286	3.991946934	1.395852965	2.640907851	1.675204655	3.455737083	4.072928879	0.484097257	0.787563914	4.056666667
3470	0.585333333	0.124566667	8.465660184	4.432047353	1.912534475	2.012301144	0.749622429	2.153292277	2.481037648	0.568404948	0.496685991	3.306666667
3670	0.421666667	0.073533333	11.63827508	4.887228661	1.176850944	0.524939868	0.71825696	0.504930669	3.127135866	0.5747853	0.269340641	2.04
3770	0.616333333	0.089	10.86233117	4.632271067	1.594117804	1.34303413	0.722192687	2.455916802	3.250429551	0.673536909	0.334090418	2.476666667
3870	0.452333333	0.084366667	6.489174693	2.342886184	1.33406522	0.790453488	0.825686702	0.56554187	2.60379154	0.629852427	0.270664329	1.55
3870	0.430333333	0.072533333	3.675167516	2.650500595	1.605650359	0.698181306	0.894923626	0.482583469	2.346283031	0.600516503	0.295324826	1.45
3970	0.339666667	0.061033333	5.62268913	2.742854664	0.982114469	0.384196145	0.689622429	0.441597335	1.939358464	0.770730491	0.2191604	1.05
4072	0.498566667	0.079733333	6.521607787	2.048113879	0.964870906	0.632230382	0.712353369	0.391597335	2.716114379	0.795305117	0.15455429	1.27
4174	0.577333333	0.091233333	6.127096395	3.444384753	1.09516019	0.831382145	0.936128414	0.577139205	4.671532772	0.818006338	0.262465036	1.29
4270	0.228333333	0.034133333	6.276561483	3.170670923	0.917149535	0.348377946	0.600987899	0.366902936	1.411802274	0.610398816	0.187205977	1.323333333
4370	0.4715	0.069866667	9.041746855	3.700310179	1.405769871	0.625924992	0.782353369	0.255916802	2.97205266	0.618663662	0.327608353	2.053333333
4480	0.497666667	0.072866667	12.01359178	4.694455457	1.212731149	0.32962435	1.224762944	0.334555736	2.739009236	0.696335449	0.192947024	1.403333333

Elevation (m.	Ca (meq/100gr)	Fe (mg/kg)	K (meq/100gr)	NH4 (mg/kg)	NO3 (mg/kg)	Ntot (mg/kg)	P (mg/kg)	pH (mg/kg)	S (mg/kg)	p_Radiation (W/m2)	p_Temp (°C)	p_Soil_water_content (mm3)
2470	4.552204327	1.846010029	226.1249186	5.291873964	1.116086235	6.407960199	2.333333333	8.65	16.7	317.09567	19.384651	0.8623748
2770	22.96237462	1.092581953	402.9629517	4.016583748	4.2039801	7.220563847	3.666666667	8.32	127.7	313.57067	16.634851	0.9452648
2870	2.854845171	1.413868635	228.6780904	5.119402985	3.898839138	9.018242123	3.666666667	8.153333333	8.266666667	312.39567	15.718251	0.9728948
2970	7.497274494	1.172366079	323.8258464	4.829187396	5.116086235	9.611940299	3.666666667	8.253333333	16.26666667	311.22067	14.801651	1.0005248
3070	3.365178886	1.885883032	290.6345978	3.189054726	3.072968491	6.262023217	5.666666667	7.936666667	6.09	310.0518467	13.88543369	1.0283
3170	5.556387794	2.341530636	258.3445028	4.059701493	3.174129353	7.233830846	7	7.63	27.38333333	308.87067	12.968451	1.0557848
3270	5.050877506	3.433810225	411.7142639	6.714759536	3.900497512	10.61525705	9.333333333	7.526666667	68.24	307.69567	12.051851	1.0834148
3370	5.27640905	3.888071674	305.0968424	3.041666667	4.466666667	7.841666667	11	7.626666667	12.02	306.52067	11.135251	1.1110448
3470	3.873471628	5.185174478	338.7590332	9.482587065	3.305140962	13.12106136	13.66666667	7.163333333	11.09333333	305.34567	10.218651	1.1386748
3670	1.959800455	12.35046099	129.9266917	6.379767828	3.348258706	10.06135987	14.33333333	6.056666667	6.993333333	302.99567	8.385451	1.1939348
3770	4.502046203	13.5444334	175.3853505	3.132669983	2.855721393	5.32172471	19.66666667	6.156666667	11.84	301.82067	7.468851	1.2215648
3870	2.036291106	20.61006025	129.58756	10.71310116	3.724709784	14.10447761	22.33333333	5.463333333	9.043333333	300.64567	6.552251	1.2491948
3870	1.863406365	21.60100975	129.58756	7.814262023	2.885572139	10.0331675	20	5.543333333	7.53	300.64567	6.552251	1.2491948
3970	1.208842917	23.40028833	94.48894755	6.744610282	3.739635158	9.817578773	23	5.39	5.2	299.47067	5.635651	1.2768248
4072	1.594578703	24.5208257	107.2563629	6.248756219	2.565505804	8.48092869	26.33333333	5.376666667	8.203333333	298.2803626	4.701226455	1.3052
4174	3.368670647	20.12390442	90.35840861	8.250414594	3.724709784	11.64179104	18.33333333	5.25	20.22333333	297.07367	3.765787	1.33319
4270	1.246868199	14.8021601	91.01326243	6.16252073	3.087893864	8.91708126	14.33333333	5.803333333	3.536666667	295.94567	2.885851	1.3597148
4370	2.262449682	18.24161733	142.9424337	8.484245439	4.058043118	12.87562189	18.33333333	5.68	10.85333333	294.77067	1.969251	1.3873448
4480	1.170737034	15.06143303	80.57911174	5.409618574	4.087893864	9.830845771	13.33333333	5.343333333	7.03	293.47817	0.960991	1.4177378

Environmental data.

ANNEX II.A2

Species	Family	Life form	Lifespan	Endemism	Carbon
Adesmia spinosissima Vogel	Fabaceae	shrub	perennial	endemic	C3
Aloysia deserticola (Phil.) Lu-Irving & O'Leary	Verbenaceae	shrub	perennial	native non-endemic	C3
Ambrosia artemisioides Meyen & Walp.	Asteraceae	herb	perennial	native non-endemic	C3
Aristida adscensionis L.	Poaceae	shrub	perennial	native non-endemic	C4
Atriplex imbricata (Moq.) D. Dietr.	Amaranthaceae	grass	annual	native non-endemic	C4
Azorella atacamensis G.M. Plunkett & A.N. Nicolas	Apiaceae	shrub	perennial	Native non-endemic	C3
Baccharis tola Phil.	Asteraceae	shrub	perennial	endemic	C3
Calamagrostis crispa (Rugolo & Villav.) Govaerts	Poaceae	grass	perennial	native non-endemic	C3
Chorizanthe commissuralis J. Rémy	Polygonaceae	herb	annual	native non-endemic	C3
Ephedra americana Humb. & Bonpl. ex Willd.	Ephedraceae	shrub	perennial	native non-endemic	C3
Fabiana denudata Miers	Solanaceae	shrub	perennial	endemic	C3
Fagonia chilensis Hook & Arn.	Zygophyllaceae	shrub	perennial	native non-endemic	C3
Hoffmannseggia doellii subsp. argentina	Fabaceae	herb	annual	native	C3
Jarava frigida (Phil.) F.Rojas	Poaceae	grass	perennial	native non-endemic	C3
Lupinus subinflatus C.P. Sm.	Fabaceae	herb	perennial	native non-endemic	C3
Maihueniopsis camachoi	Cactaceae	Cactus	perennial	NA	CAM
Moschopsis monocephala (Phil.) Reiche	Calyceraceae	herb	perennial	native non-endemic	C3
Parastrephia quadrangularis (Meyen) Cabr.	Asteraceae	shrub	perennial	native non-endemic	C3
Phacelia pinnatifida Griseb. Ex Wedd	Boraginaceae	herb	perennial	native non-endemic	C3
Pycnophyllum bryoides (Phil.) Rohrb.	Caryophyllaceae	cushion	perennial	native non-endemic	C3
Solanum chilense Dunal	Solanaceae	herb	perennial	endemic	C3
Solanum metarsium C.V.Morton,	Solanaceae	herb	annual	NA	C3
Tagetes multiflora Kunth	Asteraceae	herb	annual	native non-endemic	C3
Tiquilia atacamensis (Phil.) A.T.Richardson	Boraginaceae	herb	perennial	native	C3

Additional information on the 24 Atacama plant species studied at the metabolic level.

ANNEX II.A3

D: Dilution	*V*: Volume of extract (µL)							
Sample species	**Starch**	**Proteines**	**Polyphenol**	**Malate**	**Carbohydrates**	**Amino_acids**	**Chloro_V50**	**CitrateV20uL**
P.quadrangularis	No_dil_V10uL	No_dil_V5	D10_V2.5	No_dil_V2	No_dil_V2	D2_V2	No_dil_V50	No_dil_V2
C.commissuralis	No_dil_V10uL	No_dil_V2.5	D5_V5	No_dil_V10	No_dil_V4	D2_V2	No_dil_V50	No_dil_V4
H.doellii	No_dil_V10uL	No_dil_V2.5	D5_V5	No_dil_V2	No_dil_V10	D2_V2	No_dil_V25	No_dil_V4
A.erinacea	No_dil_V10uL	No_dil_V5	D5_V5	No_dil_V5	No_dil_V10	D2_V2	No_dil_V25	No_dil_V4
J.seriphioides	No_dil_V20uL	No_dil_V5	D5_V5	No_dil_V5	No_dil_V10	D2_V2	No_dil_V25	No_dil_V4
T.atacamensis	No_dil_V10uL	No_dil_V2.5	D5_V5	No_dil_V2	No_dil_V10	D2_V2	No_dil_V25	No_dil_V4
S.puchii	No_dil_V20uL	No_dil_V2.5	D5_V5	No_dil_V2	No_dil_V4	D2_V2	No_dil_V25	No_dil_V4
E.americana	No_dil_V10uL	No_dil_V5	D5_V5	No_dil_V2	No_dil_V2	D2_V2	No_dil_V50	No_dil_V4
A.deserticola	No_dil_V10uL	No_dil_V5	D5_V5	No_dil_V2	No_dil_V4	D2_V2	No_dil_V50	No_dil_V4
S.chilense	No_dil_V10uL	D5_V5	D5_V5	No_dil_V2	No_dil_V4	D10_V2	No_dil_V25	No_dil_V2
A.spinossisima	No_dil_V10uL	No_dil_V5	D5_V5	No_dil_V10	No_dil_V10	D2_V2	No_dil_V50	No_dil_V4
S.metarsium	No_dil_V10uL	D5_V5	D5_V5	No_dil_V2	No_dil_V4	D10_V2	No_dil_V25	No_dil_V2
B.tola	No_dil_V10uL	No_dil_V5	D5_V5	No_dil_V5	No_dil_V2	D2_V2	No_dil_V50	No_dil_V2
C.crispa	No_dil_V20uL	No_dil_V5	D5_V5	No_dil_V5	No_dil_V4	D2_V2	No_dil_V50	No_dil_V2
P.bryoides	No_dil_V10uL	No_dil_V5	D5_V5	No_dil_V10	No_dil_V10	D2_V2	No_dil_V50	No_dil_V4
A.atacamensis	No_dil_V20uL	No_dil_V5	D5_V5	No_dil_V2	No_dil_V2	D2_V2	No_dil_V25	No_dil_V4
A.imbricata	No_dil_V10uL	D5_V5	D5_V5	No_dil_V10	No_dil_V2	D2_V2	No_dil_V50	No_dil_V4
T.multiflora	D5_V20uL	No_dil_V2.5	D5_V5	No_dil_V5	No_dil_V2	D5_V2	No_dil_V25	No_dil_V4
Opuntia.sp	No_dil_V10uL	No_dil_V2.5	D5_V5	No_dil_V2	No_dil_V10	D2_V2	No_dil_V25	No_dil_V4
J.frigida	No_dil_V10uL	No_dil_V2.5	D5_V5	No_dil_V5	No_dil_V2	D2_V2	No_dil_V25	No_dil_V4
A.artemisioides	No_dil_V10uL	No_dil_V2.5	D5_V5	No_dil_V2	No_dil_V4	D5_V2	No_dil_V25	No_dil_V4
L.subinflatus	No_dil_V10uL	D5_V5	D5_V5	No_dil_V2	No_dil_V4	D5_V2	No_dil_V25	No_dil_V4
A.adscensionis	D5_V20uL	No_dil_V5	D5_V5	No_dil_V2	No_dil_V4	D5_V2	No_dil_V25	No_dil_V4
M.monocephala	No_dil_V10uL	No_dil_V2.5	D5_V5	No_dil_V2	No_dil_V4	D2_V2	No_dil_V25	No_dil_V4
F.denudata	No_dil_V10uL	D5_V5	D5_V5	No_dil_V2	No_dil_V2	D2_V2	No_dil_V25	No_dil_V2
P.pinnatifida	No_dil_V10uL	No_dil_V5	D5_V5	No_dil_V2	No_dil_V4	D2_V2	No_dil_V25	No_dil_V2
F.chilensis	No_dil_V10uL	No_dil_V5	D5_V5	No_dil_V5	No_dil_V10	D2_V2	No_dil_V25	No_dil_V2

Dilution and volume of ethanolic extract used for robotised biochemical assays.